Přírodovědecká fakulta

Univerzita Palackého v Olomouci

Katedra geoinformatiky

Studijní program: **P1314 Geografie**
Studijní obor: **1302V011 Geoinformatika a kartografie**

VYUŽITÍ DÁLKOVÉHO PRŮZKUMU ZEMĚ PŘI STANOVENÍ ZÁSOB UHLÍKU V KRAJINĚ

Disertační práce

Mgr. Pavel Vyvlečka

Školitel: prof. RNDr. Vilém Pechanec, Ph.D.

Olomouc 2026

ANOTACE

Disertační práce se zabývá využitím dat dálkového průzkumu Země pro prostorové modelování nadzemní biomasy a uhlíku v nadzemní biomase lesních porostů na území České republiky. Cílem práce bylo navrhnout, sestavit a otestovat metodický postup založený na volně dostupných družicových datech, referenčních pozemních měřeních a metodách strojového učení. V rámci práce byla zpracována data Sentinel-1 a Sentinel-2, vegetační indexy, texturní metriky a další prostorové prediktory, které byly využity pro trénování a validaci modelů nadzemní biomasy. Součástí metodického postupu bylo také testování různých sad prediktorů, porovnání vybraných algoritmů strojového učení, vyhodnocení výsledků vůči referenčním datům a srovnání s dostupnými globálními datovými sadami a údaji národní inventarizace lesů. Výsledkem práce jsou prostorové vrstvy nadzemní biomasy a uhlíku v nadzemní biomase, které jsou dostupné ve vědeckém repozitáři Zenodo. Práce přináší opakovatelný postup pro kvantifikaci nadzemní biomasy a uhlíku v lesních porostech a ukazuje možnosti využití družicových dat pro podporu monitoringu ekosystémových funkcí a uhlíkových zásob v krajině.

KLÍČOVÁ SLOVA

modelování nadzemní biomasy; uhlík v biomase; dálkový průzkum Země; strojové učení

ANNOTATION

The dissertation is focused on the use of remote sensing data for spatial modelling of above-ground biomass and biomass carbon in forest stands within the territory of the Czech Republic. The aim of the dissertation was to design, develop and test a methodological workflow based on freely available satellite data, reference field measurements and machine learning methods. Sentinel-1 and Sentinel-2 data, vegetation indices, texture metrics and other spatial predictors were processed and subsequently used for training and validating above-ground biomass models. The methodological workflow also included testing different sets of predictors, comparing selected machine learning algorithms, evaluating the results against reference data and comparing them with available global datasets and national forest inventory data. The outputs of the dissertation are spatial layers of above-ground biomass and biomass carbon, which are available in the Zenodo scientific repository. The dissertation provides a repeatable workflow for quantifying above-ground biomass and biomass carbon in forest stands and demonstrates the potential of satellite data for supporting the monitoring of ecosystem functions and carbon stocks in the landscape.

KEYWORDS

Above-ground biomass modelling; biomass carbon; remote sensing; machine learning

Prohlašuji, že

- disertační práci včetně příloh, jsem vypracoval samostatně s využitím uvedené literatury a dalších řádně deklarovaných zdrojů. Při jejím zpracování byl jako podpůrný nástroj využit generativní jazykový model ChatGPT společnosti OpenAI, a to v rozsahu a pro účely podrobně popsané v kapitole 3.4. Odborný obsah, výzkumná rozhodnutí, ověření výstupů, interpretace výsledků a konečná podoba práce jsou výsledkem mé vlastní činnosti.
- jsem si vědom, že na moji disertační práci se plně vztahuje zákon č.121/2000 Sb. - autorský zákon, zejména § 35 – využití díla v rámci občanských a náboženských obřadů, v rámci školních představení a využití díla školního a § 60 – školní dílo,
- beru na vědomí, že Univerzita Palackého v Olomouci (dále UP Olomouc) má právo nevýdělečně, ke své vnitřní potřebě, disertační práci užívat (§ 35 odst. 3),
- souhlasím, že údaje o mé disertační práci budou zveřejněny ve Studijním informačním systému UP,
- v případě zájmu UP Olomouc uzavřu licenční smlouvu s oprávněním užít výsledky a výstupy mé disertační práce v rozsahu § 12 odst. 4 autorského zákona,
- použít výsledky a výstupy mé disertační práce nebo poskytnout licenci k jejímu využití mohu jen se souhlasem UP Olomouc, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly UP Olomouc na vytvoření díla vynaloženy (až do jejich skutečné výše).

V Olomouci dne Mgr. Pavel VYVLEČKA

Děkuji školiteli práce prof. RNDr. Vilému Pechancovi, Ph.D. za odborné vedení, cenné rady a připomínky při vypracování disertační práce. Mé poděkování patří také za možnost spolupráce na výzkumných projektech a odborných publikacích, které pro mě byly cennou zkušeností a přispěly k mému odbornému růstu.

Dále děkuji doc. Ing. Stanislavu Hejdukovi, Ph.D. za poskytnutí pozemních dat luční biomasy. Za poskytnutí pozemních dat o nadzemní lesní biomase ze sítě DendroNetwork a růstových korekcí bych chtěl rovněž poděkovat Ústavu výzkumu globální změny Akademie věd České republiky – CzechGlobe.

OBSAH

ÚVOD 11
1. CÍLE PRÁCE 13
2. SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY 15
2.1. Ekosystémové funkce a služby 15
2.2. Klasifikace zemského povrchu a vymezení lesa 18
2.2.1. Datové zdroje krajinného pokryvu a lesních charakteristik 23
2.3. Dálkový průzkum ve výzkumu funkcí ekosystému 26
2.4. Nadzemní biomasa a uhlík v lesních porostech 29
2.4.1. Analýza publikační aktivity v oblasti nadzemní biomasy a uhlíku 29
2.4.2. Význam nadzemní biomasy a uhlíku v lesních ekosystémech 30
2.4.3. Přístupy k odhadu nadzemní biomasy z dat dálkového průzkumu Země 32
3. METODY A DATA 38
3.1. Zájmové území 38
3.2. Vstupní data a jejich zpracování 39
3.2.1. Sentinel-2 39
3.2.2. Sentinel-1 40
3.2.3. Vegetační indexy 41
3.2.4. Texturní metriky GLCM 42
3.2.5. Morfometrické charakteristiky 43
3.2.6. Výška lesního porostu 44
3.2.7. Maska lesního porostu 44
3.2.8. Referenční data nadzemní biomasy 45
3.2.9. Harmonizace referenčních dat pro rok 2023 48
3.2.10. Extrakce hodnot prediktorů pro referenční plochy 48
3.3. Modelování nadzemní biomasy 49
3.3.1. Trénování modelu a prostorově vnořená křížová validace 50
3.4. Využití nástrojů generativní umělé inteligence 54
4. VÝSLEDKY PRÁCE 55
4.1. Srovnání výkonnosti jednotlivých sad prediktorů 55
4.1.1. Statistické srovnání sad prediktorů 55

4.1.2. Prostorové hodnocení sad prediktorů ... 57
4.2. Srovnání algoritmů pro sadu prediktorů Sentinel-1 + Sentinel-2 ... 68
4.2.1. Porovnání na lesních plochách ... 69
4.2.2. Porovnání na rozšířeném referenčním souboru ... 72
4.2.3. Vztah mezi měřenou a predikovanou biomasou ... 75
4.2.4. Význam vstupních proměnných ... 76
4.2.5. Výběr finálního modelu Random Forest ... 78
4.3. Srovnání finálního modelu s externími daty ... 78
4.3.1. Srovnání s ESA CCI Biomass ... 78
4.3.2. Srovnání s daty národní inventarizace lesů (NIL) ... 80
4.4. Shrnutí dosažených výsledků ve vztahu k cílům práce ... 82
5. DISKUZE ... 84
6. ZÁVĚR ... 91
7. DOSTUPNOST DAT A ZDROJOVÝCH KÓDŮ ... 93
8. PŘÍLOHY ... 94
LITERATURA ... 99

SEZNAM ZKRATEK

ABA	Area-based Approach
AGB	Above-ground Biomass
AGBD	Above-ground Biomass Density
AI	Artificial Intelligence
ALOS	Advanced Land Observing
ANN	Artificial Neural Network
ARVI	Atmospherically Resistant Vegetation Index
ASM	Angular Second Moment
ATGI	Annual Tree Growth Index
AVHRR	Advanced Very-High-Resolution Radiometer
BiLSTM	Bidirectional Long Short-Term Memory
CCI	Climate Change Initiative
CI	Clumping Index
CICES	Common International Classification of Ecosystem Services
CMS	Carbon Monitoring System
CNN	Convolutional Neural Network
DEM	Digital Elevation Model
DMR	Digitální model reliéfu
DPZ	Dálkový průzkum Země
EBV	Essential Biodiversity Variables
ECV	Essential Climate Variables
EFB	Exclusive Feature Bundling
ENFIN	European National Forest Inventory Network
EO	Earth Observation
ESA	European Space Agency
EVI	Enhanced Vegetation Index
FAO	Food and Agriculture Organization
FAPAR	Fraction of Absorbed Photosynthetically Active Radiation
GAI	Green Area Index
GBDT	Gradient Boosting Decision Tree
GCOS	Global Climate Observing System
GEDI	Global Ecosystem Dynamics Investigation
GEMI	Global Environmental Monitoring Index
GEOBON	The Group on Earth Observations Biodiversity Observation Network

GIS	Geografický informační systém
GLC-SHARE	Global Land Cover-SHARE
GLCM	Grey-Level Co-Occurrence Matrix
GNDVI	Green Normalized Difference Vegetation Index
GOSS	Gradient-based One-Side Sampling
GPP	Gross Primary Production
GRD	Ground Range Detected
GRI	Global Reference Image
HRL	High Resolution Layer
IPVI	Infrared Percentage Vegetation Index
ISS	International Space Station
IW	Interferometric Wide Swath
kNN	K-Nearest Neighbours
LAI	Leaf Area Index
LGBM	LightGBM
MAE	Mean Absolute Error
MD	Mean Deviation
MEA	Millennium Ecosystem Assessment
MLPNN	Multi-Layer Perceptron Artificial Neural Network
MODIS	Moderate Resolution Imaging Spectroradiometer
MSAVI	Modified Soil Adjusted Vegetation Index
MSI	Multi-Spectral Instrument
NASA	National Aeronautics and Space Administration
NDHD	Normalised difference hotspot and darkspot index
NDI45	Normalized Difference Index 45
NDREI	Normalized Difference Red Edge Index
NDVI	Normalized Difference Vegetation Index
NFI	National Forest Inventory
NIL	Národní inventarizace lesů České republiky
NIR	Near Infrared
NOAA	National Oceanic and Atmospheric Administration
NPP	Net Primary Production
OBIA	Objektově založená klasifikace
OLI	Operational Land Imager
OOF	Pooled Out-of-fold
PALSAR	Phased Array Type L-band Synthetic Aperture Radar

PgC	Petagram uhlíku
PLO	Přírodní lesní oblasti
ReNDVI	Red-Edge NDVI
RF	Random Forest
RMSD	Root Mean Square Deviation
RMSE	Root Mean Square Error
RVI	Radarový vegetační index
SAR	Synthetic Aperture Radar
SD	Standard Deviation
SLT	Soubor lesních typů
SRTM	Shuttle Radar Topography Mission
SSM	Surface Soil Moisture
SVM	Support Vector Machine
SWIR	Short-Wavelength Infrared
TEEB	The Economics of Ecosystems and Biodiversity
THEOS	Thailand Earth Observation System
TIFF	Tagged Image File Format
TIR	Thermal Infrared
TM	Thematic Mapper
UAV	Unmanned Aerial Vehicle
USGS	United States Geological Survey
VH	Vertical-horizontal polarizace
VIIRS	Visible Infrared Imaging Radiometer Suite
VV	Vertical-vertical polarizace
WRS-2	Worldwide Reference System-2
XGB	Extreme Gradient Boosting
ZABAGED	Základní báze geografických dat
ÚHÚL	Ústav pro hospodářskou úpravu lesa

ÚVOD

Zemské ekosystémy hrají zásadní roli při podpoře života a poskytování základních funkcí organismům nebo ostatním ekosystémům. Mezi základní ekosystémové funkce patří zásoba pitné vody, potravy, materiálu, regulace klimatu a eroze nebo úkryt pro živočichy (De Groot a kol., 2002). Při špatném či nepřesném fungování ekosystémů dochází ke změně ekosystémových funkcí, které může vést ke globálním problémům. V současnosti dochází ke změnám v ekosystémech a tím se narušují i jeho funkce. Zvyšující se lidská populace klade čím dál vyšší nároky na dostatek pitné vody a potravy, zejména masa. Podle FAO (2023) se produkce masa za posledních 30 let téměř ztrojnásobila. Celkově roste i životní úroveň, nejen ve vyspělých zemích, ale i v rozvojových a méně rozvinutých zemích. Nastává tak větší spotřeba materiálů a elektrické energie, v důsledku toho se zvyšuje produkce skleníkových plynů, která vede ke globálnímu oteplování. Globální oteplování způsobuje tání ledovců, zvyšování hladiny oceánů a vede k úbytku pitné vody. Lze tak pozorovat dominový efekt mezi ekosystémy, kdy počáteční změna může ovlivnit celou řadu ekosystémů. V současné době existuje spousta globálních plánů v boji proti změnám klimatu, nicméně je nepravděpodobné, že by byl trend nárůstu teploty zastaven. Reálnějším řešením je zpomalení oteplování a nechat ekosystémy a organismy adaptovat se na nové podmínky. Proto je nyní velmi důležitým aspektem zkoumání funkcí ekosystémů a pochopení jejich dynamiky v závislosti na změně klimatu.

Aby bylo možné předcházet globálním klimatickým změnám, je důležitý monitoring ekosystémových funkcí. Funkce ekosystémů zahrnují širokou škálu procesů jako je sekvestrace uhlíku, koloběh vody nebo živin a mají zásadní význam pro pochopení zdraví ekosystémů, dopadů lidské činnosti nebo biologickou rozmanitost. Jedním z důležitých indikátorů ekosystémových funkcí je biomasa, která bude v disertační práci zkoumána. Tradiční metody sběru dat v terénu jsou časově velmi náročné, nákladné a prostorově omezené. V mnoha případech se navíc jedná o kontaktní metody, které mohou porušit nebo i zničit měřenou biomasu. Jako alternativní řešení se v tomto případě nabízí dálkový průzkum Země. Dálkový průzkum dokáže zachytit velké množství informací o povrchu na základě odrazivosti v různorodých elektromagnetických spektrech. Jedná se tak o perspektivní, nedestruktivní a bezkontaktní metodu, která dokáže pokrýt velké území bez ohledu na vlastnické právo či péči o krajinu. Přesto mají družicová data nadále svá omezení. Optický dálkový průzkum je výrazně ovlivněn atmosférickými podmínkami, zejména oblačností a aerosoly, které často znemožňují přímé pozorování povrchu. Radarové (SAR) snímání je naopak do značné míry nezávislé na oblačnosti a denní době, avšak jeho interpretaci komplikují zrnitost (*speckle*) a geometrická zkreslení. Zpětný rozptyl je navíc silně závislý na drsnosti povrchu a struktuře/hustotě vegetace, což může zvyšovat nejistotu klasifikace a odhadu biofyzikálních parametrů.

Řešením by v daném případě mohlo být použití více různých družicových dat, pozemní měření a pokročilé modelovací přístupy, které jsou schopny odhadnout a mapovat různé funkce ekosystémů. Kombinace dvou odlišných metod dálkového průzkumu dokáže částečně eliminovat slabiny a může pomoci ke komplexnějšímu porozumění zemského povrchu. Významným

parametrem pro úspěšnou kvantifikaci funkcí je i pozemní měření. Z družicových dat nelze kvantifikovat ekosystémové funkce přímo, ale pouze nepřímo za použití spektrálních pásem nebo z nich vytvořených vegetačních indexů. Pozemní měření tak slouží jako trénovací a validační množina pro vstupní družicová data. Posledním klíčovým aspektem je modelovací algoritmus. V dnešní době jsou používány algoritmy strojového učení, *deep learning* nebo neuronové sítě, které jsou efektivní v modelování či predikci téměř jakékoliv proměnné.

1. CÍLE PRÁCE

Cílem práce je sestavení a praktické otestování metodického postupu, který povede ke kvantifikaci zásob nadzemní biomasy a uhlíku v nadzemní biomase na území České republiky. Výsledků bude docíleno pomocí použití dat dálkového průzkumu Země a dalších doplňkových prediktorů. Základní hypotézou je předpoklad, že je možné nadzemní biomasu a uhlík modelovat pomocí volně dostupných dat. Výsledky práce by měly rovněž odpovědět na otázku do jaké míry může být přesnost odhadu zlepšena zahrnutím dalších vrstev.

Součástí výzkumné práce je vytvoření automatizovaného a opakovatelného postupu zahrnujícího přípravu a stažení vstupních dat a poté trénování a validaci modelů. Výsledky modelování budou následně porovnány s dostupnými referenčními údaji nebo již existujícími globálními datovými sadami. Hlavní cíl disertace je členěn do čtyř dílčích cílů.

DÍLČÍ CÍL 1 (DC1): Vymezení teoretického a datového rámce pro odhad nadzemní biomasy z dat dálkového průzkumu Země

Prvním dílčím cílem je vytvořit znalostní a metodický podklad pro odhad nadzemní biomasy lesních porostů pomocí dat dálkového průzkumu Země. Tento cíl zahrnuje vymezení základních pojmů souvisejících s nadzemní biomasou, uhlíkem v biomase a ekosystémovými funkcemi vegetace. Součástí je rovněž rešerše dostupných družicových dat, prostorových prediktorů a metod používaných pro kvantifikaci nadzemní biomasy v lesních ekosystémech. Na základě dostupné literatury a parametrů jednotlivých datových zdrojů jsou posouzena jejich omezení a možnosti využití pro prostorové modelování nadzemní biomasy v podmínkách České republiky. Důraz je kladen především na volně dostupná data vysokého rozlišení.

DÍLČÍ CÍL (DC2): Sestavení automatizovaného postupu pro přípravu a zpracování družicových dat

Druhým dílčím cílem je navrhnout a realizovat automatizovaný postup pro přípravu družicových dat využitelných při modelování nadzemní biomasy. Tento postup zahrnuje zejména zpracování dat v cloudové platformě Google Earth Engine, tvorbu bezoblačných obrazových kompozitů, výpočet vegetačních indexů a sjednocení vstupních vrstev do podoby vhodné pro další analýzu v prostředí GIS.

Součástí cíle je také příprava dalších prostorových prediktorů, jejich převzorkování na jednotné prostorové rozlišení, zarovnání na společnou rastrovou mřížku a prostorové ořezání pomocí lesní masky. Výsledkem tohoto kroku je konzistentní sada prediktorových vrstev, která umožňuje systematické testování různých kombinací vstupních dat a jejich vlivu na přesnost odhadu nadzemní biomasy.

DÍLČÍ CÍL (DC3): Modelování nadzemní biomasy lesních porostů a hodnocení prediktorových sad

Třetím dílčím cílem je otestovat vhodnost vybraných metod strojového učení pro odhad nadzemní biomasy lesních porostů na základě terénních referenčních dat a připravených prostorových prediktorů. Modelování je zaměřeno na posouzení, zda lze nadzemní biomasu odhadovat pouze

s využitím družicových dat a zda zahrnutí dalších prediktorových vrstev vede ke zlepšení přesnosti a prostorové stability výsledků.

V rámci tohoto cíle jsou porovnávány různé skupiny prediktorů. Důraz je kladen nejen na celkovou přesnost modelů, ale také na interpretaci významnosti jednotlivých prediktorů a na hodnocení jejich přínosu pro výslednou predikci nadzemní biomasy. Součástí modelování je rovněž kalibrace a validace modelů pomocí referenčních terénních dat.

DÍLČÍ CÍL (DC4): Tvorba výsledných vrstev nadzemní biomasy a uhlíku a jejich porovnání s existujícími datovými sadami

Čtvrtým dílčím cílem je vytvořit výslednou prostorovou vrstvu nadzemní biomasy lesních porostů pro území České republiky a z ní odvodit zásobu uhlíku v nadzemní biomase. Výsledné vrstvy jsou následně vyhodnoceny z hlediska prostorového rozložení hodnot, průměrných hektarových zásob a celkového odhadu nadzemní biomasy a uhlíku v rámci použité lesní masky.

Součástí tohoto cíle je porovnání výsledků s již existujícími datovými sadami nadzemní biomasy a s dostupnými inventarizačními referencemi. Toto porovnání slouží k posouzení míry shody navrženého modelového přístupu s nezávislými zdroji informací a k identifikaci hlavních rozdílů vyplývajících z použité metodiky, prostorového rozlišení, vstupních dat, referenčních měření a vymezení lesní plochy. Výsledkem je zhodnocení využitelnosti navrženého metodického postupu pro kvantifikaci nadzemní biomasy a uhlíku v lesních porostech České republiky.

2. SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY

Dálkový průzkum Země představuje jeden z hlavních nástrojů současného prostorového monitoringu krajiny, vegetace a vybraných ekosystémových charakteristik (Pettorelli a kol., 2014). Monitoring pokryvu Země lze obecně chápat jako systematické měření cílové oblasti v čase, které je prováděno opakovaně za účelem zachycení změn kvality nebo kvantity sledovaného objektu a pochopení procesů, které za těmito změnami stojí (Mücher, 2009). V kontextu této práce je pozornost zaměřena zejména na lesní ekosystémy, jejich prostorové vymezení, nadzemní biomasu a uhlík uložený v nadzemní biomase. Vzhledem k tomu, že tato práce pracuje s biofyzikálními charakteristikami lesních ekosystémů odvozenými z dat dálkového průzkumu Země, je nejprve vhodné vymezit širší rámec ekosystémových funkcí a služeb, do kterého sledované indikátory spadají. Na tento základ navazuje problematika klasifikace zemského povrchu, vymezení lesa a přehled možností dálkového průzkumu Země při sledování ekosystémových charakteristik. Následně je pozornost zaměřena na nadzemní biomasu a uhlík v lesních porostech a na přístupy využívané pro jejich prostorové modelování.

2.1. Ekosystémové funkce a služby

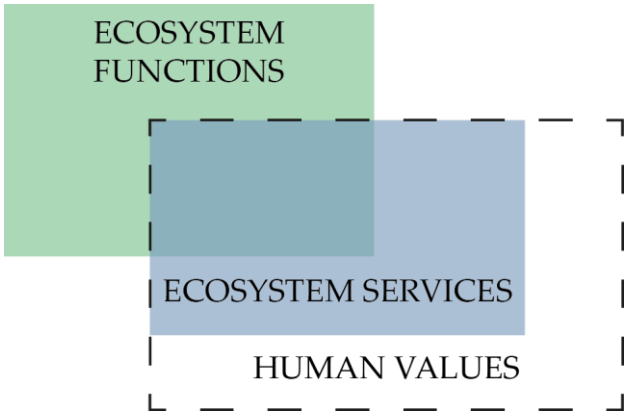
Přírodní ekosystémy jsou velmi komplexní systémy abiotických a biotických elementů. Ekosystémové funkce jsou základními interakcemi mezi komponenty ekosystémů nebo napříč ekosystémy a popisují procesy (přenos energie a materiálů) ekosystému (Lyons a kol., 2005). V dalším případě jsou funkce definovány ve třech bodech (Eugene, 1962):

- Míra proudění biologické energie v ekosystému;
- Míra koloběhu materiálů nebo živin (biogeochemické cykly);
- Biologická nebo ekologická regulace, zahrnující regulaci organismů prostředím tak regulaci prostředí organismy.

Funkce ekosystémů jsou vlastnosti spojené s fungováním ekosystémů, které vznikají důsledkem jednoho nebo více ekosystémových procesů (Lovett a kol., 2006). Jiná definice uvádí, že ekosystémové funkce jsou přímé či nepřímé benefity ekosystémových procesů pro řadu živočichů včetně člověka (Pettorelli a kol., 2018). Základem všech funkcí ekosystémů je energie, voda a uhlík. Tyto elementy jsou úzce propojeny. Nemá-li rostlina dostatek vody, mění se její tok energie, rostlina mění výpar a zvyšuje teplo, čímž se zpomalují základní procesy, jako je například vázání uhlíku (Forzieri a kol., 2020). Tyto prvky musí zůstat v rovnováze, protože se navzájem ovlivňují.

Ekosystémové služby jsou přínosem pro lidský blahobyt, který člověk přímo nebo nepřímo získává z funkcí ekosystému (Daily, 1997). Služby lze zjednodušeně definovat jako benefity, které člověk získává z ekosystémů (MEA, 2005). Jiným způsobem, jak lze definovat služby zní: specifické efekty funkcí ekosystémů, které přímo udržují a zlepšují lidský život (Brown a kol., 2007). Rozdíl mezi službami a funkcemi je definován tvrzením, že funkce mohou mít jak přírodní, tak potenciální antropocentrické hodnoty, zatímco služby jsou definovány pouze z hlediska jejich přínosů pro člověka (Petter a kol., 2012). Přestože jsou stanoveny hranice mezi funkcemi

a službami, klasifikace se budou vždy překrývat, jelikož člověk je součástí organismů na Zemi. Překryv vychází ze základních definic, jelikož služby jsou založeny na funkcích anebo se přímo z funkce stane služba, pokud ji člověk vyhodnotí jako prospěšnou pro jeho blahobyt. Typickým příkladem je kategorie „potraviny a materiály“, kterou lze považovat jak za funkci, tak za službu. Jako funkce poskytuje potravu a materiál pro všechny živé a neživé organismy, zatímco jako služba by byla přímým přínosem pro člověka, např. produkce palivového dříví, papíru, ovoce a lesních plodů nebo stavebních materiálů. Výsledky těchto služeb mohou být zpeněženy. The Economics of Ecosystems and Biodiversity (TEEB) považuje ekosystémové služby z ekonomického hlediska za dividendy, které společnost získává z přírodního kapitálu. Udržování zásob přírodního kapitálu tak umožňuje trvalé poskytování toků ekosystémových služeb, které zajišťují blahobyt lidí (TEEB, 2010). Služby vytvářejí přínosy a každá z nich má pro daného člověka jinou hodnotu (Hicks a kol., 2015). Jak již bylo uvedeno, ekosystémové služby a funkce však nejsou striktně odděleny a v některých případech se překrývají (Obrázek 1).

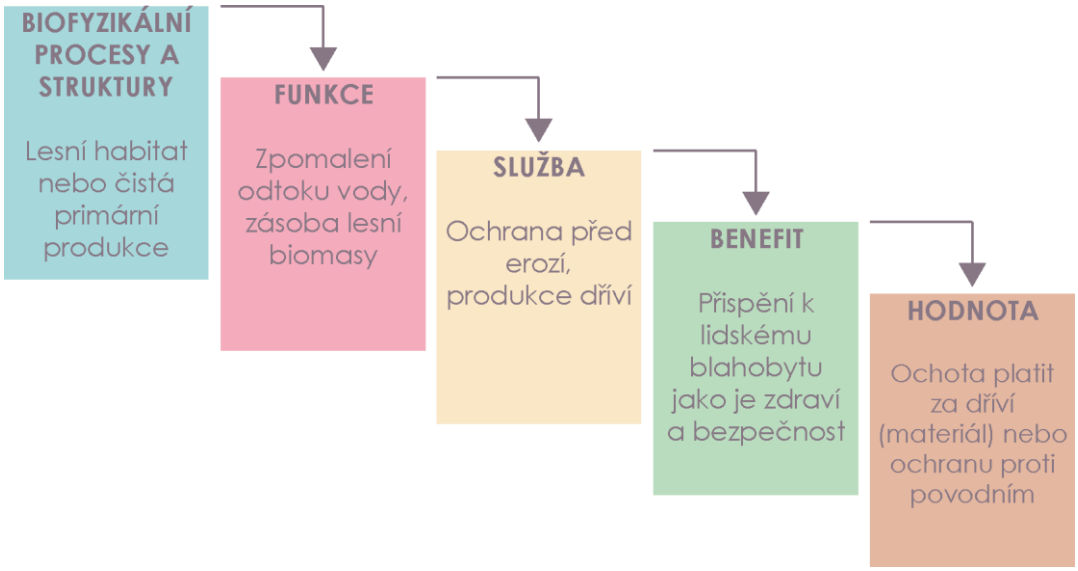


Obrázek 1 Vztah mezi ekosystémovými službami a funkcemi (upraveno podle Pettorelli a kol., 2018)

Povědomí o významu ekosystémových služeb a funkcí se začíná dostávat do strategických plánů států nebo mezinárodních organizací. Některé klíčové ukazatele ekosystémových služeb nebo funkcí byly zařazeny na seznam základních klimatických proměnných ECV (Essential Climate Variables). Seznam obsahuje 54 proměnných, které jsou rozděleny do 11 skupin. Ve skupině biosféra je osm proměnných: nadzemní biomasa, *land cover*, albedo, požáry, teplota povrchu, leaf area index (LAI), podíl absorbovaného fotosynteticky aktivního záření (FAPAR) a půdní uhlík (GCOS, 2022). Obecně se jedná o skupinu fyzikálních, chemických a biologických veličin, které se rozhodujícím způsobem podílejí na charakteristikách klimatu Země a jeho budoucím vývoji (Pettorelli a kol., 2018). Struktura ekosystému je také považována za jednu ze tříd EBV (Essential Biodiversity Variables). Tuto třídu tvoří následující tři proměnné: podíl živého pokryvu Země, rozložení ekosystému a vertikální profil ekosystému (GEOBON, 2022). Proměnné jsou sledovány na úrovni jednotlivých ekosystémů, ale nikoliv na úrovni jejich vzájemné interakce.

Problematikou klasifikace ekosystémových funkcí a služeb se zabývalo již několik autorů (Costanza a kol., 1997; De Groot a kol., 2002; Turkelboom a kol., 2013; Meyer a kol., 2015). De Groot identifikoval celkem 23 ekosystémových funkcí rozdělených do 4 základních kategorií: regulační, stanovištní, produkční a informační (De Groot a kol., 2002). Problematika klasifikace ekosystémových funkcí a služeb se dostala do širšího podvědomí díky Millennium Ecosystem

Assessment (MEA, 2005). Podobná klasifikace byla vytvořena iniciativou TEEB (TEEB, 2010), ta se nicméně zabývala především ekosystémovými službami z ekonomického hlediska. Dalším významným počinem v oblasti klasifikace ekosystémových služeb bylo vytvoření databáze CICES (Common International Classification of Ecosystem Services). V roce 2013 byla vydána verze databáze 4.3, která byla v roce 2018 nahrazena verzí 5.1. Databáze je rozdělena do tří hlavních kategorií: zásobovací, regulační a kulturní. Zmíněné kategorie se dále dělí podle biotických a abiotických ekosystémů. Přesné vymezení klasifikace funkcí a služeb ekosystémů je však stále vágní. Vágnost spočívá v tom, že některé ekosystémové služby jsou klasifikovány jako třídy ekosystémových funkcí a naopak. Mnoha autorům se nepodařilo přesně vymezit hranice mezi klasifikací funkcí a služeb (Daily, 1997; Meyer a kol., 2015; Oliver a kol., 2015). Závislost ekosystémových funkcí a lidským blahobytem je znázorněna na ekosystémové kaskádě (Haines-Young a Potschin, 2010). Biofyzikální procesy a funkce ekosystémů jsou v kaskádě považovány za podpůrné a zprostředkující služby, kdežto ekosystémové služby jsou považovány za finální služby. Finální služby jsou to, co lze z ekosystému sklidit nebo získat. Vzhledem k již překrývajícím se pojmům bych procesy neoznačoval jako podpůrné a zprostředkující služby. Ekosystémové funkce stojí ve druhém stupni kaskády (Obrázek 2) a určuje funkci biofyzikálních procesů a struktur. Výsledkem celé kaskády je hodnota, kterou každá služba představuje pro blahobyt člověka.



Obrázek 2 Ekosystémová kaskáda (upraveno podle Haines-Young a Potschin, 2010)

Nadzemní biomasa a uhlík v nadzemní biomase patří mezi biofyzikální charakteristiky, které jsou úzce spojeny s fungováním lesních ekosystémů. Biomasa vyjadřuje organickou hmotu lesních porostů a souvisí s produkční funkcí ekosystému, strukturou lesního porostu, dostupností stanovišť pro organismy i schopností vegetace ovlivňovat vodní, energetický a živinový cyklus. Uhlík uložený v biomase je zároveň významnou složkou uhlíkového cyklu a přímo souvisí s regulačními funkcemi ekosystémů, zejména se sekvestrací uhlíku a zmírňováním dopadů klimatické změny. Z tohoto důvodu jsou nadzemní biomasa a uhlík v nadzemní biomase vhodnými indikátory pro hodnocení vybraných ekosystémových funkcí lesních porostů. Jejich kvantifikace

umožňuje prostorově hodnotit nejen zásobu organické hmoty a uhlíku, ale také funkční význam lesních ekosystémů v krajině. Právě tato vazba mezi biomasou, uhlíkem a ekosystémovými funkcemi je důvodem, proč je v této práci vymezen obecný vztah mezi ekosystémovými funkcemi, službami a biofyzikálními indikátory.

2.2. Klasifikace zemského povrchu a vymezení lesa

Prostorové modelování nadzemní biomasy vyžaduje jednoznačné vymezení území, na kterém má být predikce aplikována. V případě lesní biomasy je proto klíčovým aspektem nejen vlastní modelování hodnot biomasy, ale také definice a prostorové vymezení lesa. Výsledná lesní maska ovlivňuje prostorový rozsah predikce a následně i celkový odhad národní zásoby biomasy a uhlíku. I při obdobné průměrné hektarové hodnotě biomasy mohou rozdílné masky lesa vést k odlišným celkovým zásobám.

Klasifikace zemského povrchu z dat dálkového průzkumu Země je dlouhodobě využívána pro mapování krajinného pokryvu a využití území. *Land cover*, v české terminologii označovaný jako půdní kryt (v širším smyslu zahrnuje i vegetační kryt), vyjadřuje biofyzikální pokryv zemského povrchu, přičemž půdní kryt lze chápat jako pedosféru vzniklou přeměnou svrchní části litosféry působením organismů za účasti slunečního záření, vzduchu a vody (Terminologická komise ČÚZK, 2021). Zahrnuje například lesní porosty, travní porosty, vodní plochy, holou půdu, skály, ledovce nebo zastavěné plochy. Tyto třídy jsou z hlediska dálkového průzkumu Země zpravidla rozlišovány na základě spektrálního, prostorového nebo texturního chování povrchu. *Land use* naproti tomu vyjadřuje způsob, jakým je daný povrch využíván člověkem, v českém překladu využití půdy, a je definován jako souhrn úprav, činností a vstupů, které člověk uskutečňuje v určitém typu půdního krytu. Může se jednat například o zemědělské, lesnické, sídelní, rekreační nebo průmyslové využití území. Pojmy *land cover* a *land use* jsou však v praxi často zaměňovány, a to i v rámci studií založených na datech dálkového průzkumu Země. Tato nejednoznačnost komplikuje interpretaci výsledků i integraci dat o území, protože oba koncepty jsou zásadně odlišné (Comber, 2008). *Land cover* je určován přímým pozorováním zemského povrchu, zatímco *land use* představuje socioekonomickou interpretaci aktivit, které se na tomto povrchu odehrávají (Fisher a kol., 2005). V praxi jsou však tyto dva přístupy často kombinovány, což se promítá do většiny existujících databází a klasifikačních systémů, takže čisté klasifikace *land cover* nebo *land use* jsou spíše výjimečné, přestože jsou často deklarovaným cílem (Di Gregorio a Jansen, 2005).

Vliv prostorového rozlišení na klasifikaci zemského povrchu byl předmětem řady odborných studií. Ve studii Suwanprasit a Srichai (2012) byla porovnávána data Landsat-5 TM s prostorovým rozlišením 30 m a THEOS s rozlišením 15 m při klasifikaci sedmi tříd v jižním Thajsku. Data byla klasifikována metodou Support Vector Machine. Celkové přesnosti klasifikace bylo dosaženo ve výši 89 % u Landsat-5 TM a 90,65 % u THEOS. Na základě výsledků bylo konstatováno, že vyšší prostorové rozlišení může významně zlepšit výsledky klasifikace, zároveň však byla zdůrazněna důležitost spektrální informace, zejména přítomnosti NIR pásem. Bylo rovněž uvedeno, že vyšší prostorové rozlišení snižuje problém smíšených pixelů a umožňuje detailnější zachycení prostorové heterogenity krajinného pokryvu a jednotlivých typů zemského povrchu. Další

publikací, ve které byl zkoumán vliv prostorového rozlišení na mapování využití země v jihovýchodní Brazílii, byla studie Fisher a kol. (2018). Jako data s vysokým prostorovým rozlišením byly využity snímky QuickBird a WorldView-2, ze kterých byla vybrána čtyři spektrální pásma (RED, GREEN, BLUE a NIR). Vzhledem k tomu, že prostorové rozlišení družice QuickBird činilo 0,6 m a družice WorldView-2 0,5 m, byla data převzorkována na jednotné rozlišení 1 m, mimo jiné z důvodu návaznosti na digitální model reliéfu využívaný v dalších analýzách. Jako data s nižším prostorovým rozlišením byly použity snímky Landsat-7 s rozlišením 30 m. Stejně jako v předchozí studii bylo klasifikováno sedm tříd: vodní plocha, holá půda, pastviny, rýžová pole, nepropustný povrch, plantáže a lesní porost. Celková přesnost klasifikace dosáhla 82,3 % u dat s vysokým rozlišením, zatímco u dat s nižším rozlišením bylo dosaženo přesnosti 75,1 %. Na základě výsledků bylo doporučeno využívat data s nižším rozlišením především v oblastech s homogennějším rozložením tříd, s pozvolnými změnami nadmořské výšky a s rozsáhlými krajinnými prvky, jejichž velikost převyšuje prostorové rozlišení použitých dat.

Některé studie se věnují spojitosti prostorového rozlišení a typu družicových snímků v závislosti na rozlišení habitatů (Nagendra a kol., 2013; Corbane a kol., 2015). Tyto práce slouží jako přehled již publikovaných studií, které se v daném rozlišení mapování biotopů věnovaly. Studie Nagendra a kol. (2013) uvádí, že nízké rozlišení (MODIS, AVHRR) není pro mapování biotopů vhodné. Střední rozlišení (Landsat), dokáže zachytit hrubý rozsah a spektrální chování biotopů. Jako příklad je uvedena studie od Brown De Colstoun a kol. (2003), která mapuje biotopy v USA a celková přesnost klasifikace s použitím dat Landsat-7 je 82 %. Pokud jsou ovšem použity snímky s vyšším rozlišením, lze od sebe rozlišit typy habitatů, které mají podobné spektrální chování (Lucas a kol., 2011). Nejlepších výsledků v multispektrálních senzorech dosahovaly senzory s velmi vysokým rozlišením (IKONOS, QuickBird, WorldView-2). V daném rozlišení je možné identifikovat a rozlišit homogenní habitaty, ekotony či plochy (Comber a kol., 2010). Studie zmiňuje, že habitaty lze rozlišit i za použití hyperspektrálních dat, či dat radarových nebo lidarových.

Ve studii Corbane a kol. (2015) je zvolen odlišný přístup k dělení habitatů. Přehled je rozdělen na jednotlivé rozlišovací úrovně. Při úrovni 1 je možné rozlišit různé typy habitatů mezi sebou, jako například louku, křoví a les. Úroveň 2 představuje rozlišení v dané třídě mezi jednotlivými druhy, například jehličnatý a listnatý les. Pro úroveň 1 jsou doporučené k použití všechny typy prostorového rozlišení, od nízkého po velmi vysoké. Pokud ale přejdeme k úrovni 2 a rozdělení mezi jednotlivými druhy, použití dat s nízkým rozlišením se nedoporučuje. To platí u všech čtyř zkoumaných habitatů, kterými jsou lesní porosty, louky, vřesoviště a mokřady. U středního rozlišení lze zachytit alespoň hrubý rozsah jednotlivých typů. Data s vysokým a velmi vysokým rozlišením jsou však pro danou úlohu zdaleka nejlepší. Studie vyvozuje závěr, že pokud se dbá pouze na identifikaci rozsahu jednotlivých typů habitatů a jejich změn, je potřeba, aby multispektrální data obsahovala NIR pásmo a jejich prostorové rozlišení bylo lepší než 15 m. Jako vhodná metoda identifikace je pro nejlepší výsledek určeno modelování v GIS (geografický informační systém) softwaru v kombinaci s řízenou klasifikací.

Velký potenciál pro mapování *land use/land cover* je připisován družici Sentinel-2. Díky prostorovému rozlišení, počtu spektrálních pásem a časovému rozlišení je považována za vhodný zdroj dat pro mapování zemského povrchu. Ve studii Forkuor a kol. (2018) byla porovnána data Landsat-8 a Sentinel-2 pro mapování *land use/land cover* v Burkině Faso. Oba produkty jsou volně dostupné a poskytují kvalitní informace o zemském povrchu. Výhodou Landsat-8 jsou termální pásma, která Sentinel-2 neobsahuje, zatímco u Sentinel-2 je k dispozici větší počet pásem v oblasti *red-edge* a blízkého infračerveného záření. Ve studii byla vybrána pásma ve viditelné, NIR a SWIR části spektra a následně byla převzorkována na jednotné prostorové rozlišení 10 m. Důležitou součástí předzpracování bylo odstranění rozdílů v geolokaci obou datových sad. Landsat-8 a Sentinel-2 využívají odlišné referenční systémy a kolokaci pixelů. Landsat-8 je založen na systému Worldwide Reference System-2 (WRS-2), který využívají také starší mise Landsat, zatímco Sentinel-2 používá Global Reference Image (GRI) vytvořený z ortorektifikovaných snímků Sentinel-2 (Irons a kol., 2012; Storey a kol., 2016). Při vzájemném porovnávání snímků tak může docházet k nesouososti pixelů a geolokačním rozdílům mezi oběma produkty. Ve studii proto byla data Landsat-8 dodatečně georeferencována pomocí kontrolních bodů využívaných pro Sentinel-2 a byla aplikována metoda *image-to-image registration*. Tento krok je považován za důležitý zejména při kombinaci více družicových zdrojů, protože nepřesná kolokace pixelů může negativně ovlivnit následnou klasifikaci i modelování.

Po dokončení předzpracování byla do klasifikace zahrnuta jak spektrální pásma, tak vegetační indexy. Testovány byly algoritmy Random Forest, Support Vector Machine a Stochastic Gradient Boosting. Nejvyšší celkové přesnosti bylo dosaženo algoritmem Stochastic Gradient Boosting, následovaným algoritmem Random Forest (Obrázek 3). Při porovnání jednotlivých datových sad byly nejlepší výsledky dosaženy s využitím pásem Sentinel-2, což bylo autory přisuzováno zejména přítomnosti *red-edge* pásem. Studie tak potvrzuje význam těchto pásem pro rozlišení tříd krajinného pokryvu a současně ukazuje, že při kombinaci dat z různých družicových misí je nutné věnovat pozornost nejen spektrálním vlastnostem dat, ale také jejich prostorové registraci a kolokaci pixelů.

	Overall accuracy			Kappa coefficient		
Experiment/Classification	RF	SVM	SGB	RF	SVM	SGB
Landsat-8	90.5	88.9	89.2	0.886	0.867	0.871
Sentinel-2 subset	93.1	92.9	93.6	0.918	0.915	0.923
Sentinel-2	92.5	93.4	94.3	0.910	0.920	0.932
Red-edge only	92.4	91.1	91.7	0.909	0.893	0.900
Landsat plus red edge	92.3	92.7	93.4	0.908	0.913	0.921

Obrázek 3 Srovnání produktů a metod klasifikace (Forkuor a kol., 2018)

Vedle pixelově orientovaných přístupů byly při klasifikaci zemského povrchu využívány také objektově orientované metody. V práci Lin a kol. (2015) byly porovnávány pixelově založený algoritmus Maximum Likelihood Classifier a objektově založený algoritmus Support Vector Machine při klasifikaci snímků WorldView-2. Zatímco pixelová klasifikace byla založena výhradně

na spektrálním chování jednotlivých pixelů, u objektově orientované klasifikace jsou kromě spektrálních charakteristik zohledňovány také prostorové, texturní a kontextové vlastnosti objektů (Weih a Riggan, 2010). Výhodou objektově orientovaného přístupu je možnost využití doplňujících informací o tvaru, velikosti a vzájemných vztazích objektů, které nejsou při klasifikaci jednotlivých pixelů k dispozici.

Srovnání pixelově, objektově a hybridně orientovaného přístupu bylo provedeno také Wangem a kol. (2004) na základě dat družice IKONOS v oblasti Panamy. Nejvyšší celkové přesnosti bylo dosaženo při použití hybridního přístupu, který kombinoval výhody obou metod. Výsledky studie naznačily, že propojení spektrálních informací s objektovými charakteristikami může vést ke zlepšení klasifikačních výsledků, zejména v heterogenním prostředí. Autoři dosáhli velmi vysoké přesnosti při klasifikaci tropického deštného lesa, zatímco výrazně nižší přesnost byla zaznamenána u mangrovových porostů, které představovaly jednu z nejproblematičtějších tříd z hlediska správné identifikace. Algoritmy Random Forest, Nearest Neighbour a Support Vector Machine byly následně testovány na datech Sentinel-2 (Thanh Noi a Kappas, 2018) a současně byl hodnocen vliv počtu trénovacích ploch na přesnost klasifikace. Nejvyšší celkové přesnosti bylo dosaženo pomocí algoritmu SVM (Support Vector Machine), následovaného algoritmy RF (Random Forest) a kNN (K-Nearest Neighbors). Současně bylo potvrzeno, že rozsah a reprezentativnost trénovacích dat významně ovlivňují výslednou kvalitu klasifikace.

Z uvedených studií vyplývá, že výsledná přesnost klasifikace je ovlivňována řadou faktorů, mezi které patří prostorové rozlišení dat, spektrální informace, kvalita a rozsah trénovacích dat, heterogenita zájmového území i zvolený klasifikační algoritmus. Významnou roli hraje rovněž způsob zpracování dat, tedy zda je využíván pixelově orientovaný, objektově orientovaný nebo hybridní přístup. Tyto poznatky jsou důležité i pro modelování nadzemní biomasy, protože volba lesní masky a její tematická přesnost přímo ovlivňují prostorový rozsah výsledné predikce a následnou interpretaci odhadů biomasy a uhlíku. V následující tabulce jsou srovnány celkové přesnosti klasifikací pro různé typy dat a zájmové oblasti (Tabulka 1).

Tabulka 1 Přehled použitých dat, metod a jejich přesnost při klasifikaci povrchu

Použitá data	Použité prostorové rozlišení	Metoda klasifikace	Celková přesnost klasifikace	Zájmová oblast	Zdroj
Landsat 5 TM	30 m	Support Vector Machine	89 %	Thajsko	Suwanprasit a Srichai, 2012
THEOS	15 m	Support Vector Machine	90,65 %		
QuickBird, Worldview 2	1 m	Řízená klasifikace	82,30 %	Brazílie	Fisher a kol., 2018
Landsat 7	30 m	Řízená klasifikace	75,10 %		
Sentinel-2	10 m	Random Forest	93,10 %	Burkina Faso	Forkuor a kol., 2018
		Support Vector Machine	93,40 %		
		Stochastic Gradient Boosting	94,30 %		
Landsat 8	10 m	Random Forest	90,50 %		

		Support Vector Machine	88,90 %		
		Stochastic Gradient Boosting	89,20 %		
Worldview-2	0,5 m	Maximum Likelihood Classifier	82,44 %	Tajwan	Lin a kol., 2015
		Support Vector Machine	89,36 %		
Worldview-2	2 m	Maximum Likelihood Classifier	78,99 %		
		Support Vector Machine	75,17 %		
SPOT 5	1 m	Řízená klasifikace	66,90 %	Arkansas, USA	Weih a Riggan, 2010
	10 m		64,40 %		
	1 m	Objektově založená klasifikace	82 %		
	10 m		66,10 %		
IKONOS	1 m	Maximum Likelihood Classifier	88,90 %	Panama	Wang a kol., 2004
		Nearest-Neighbour	80,40 %		
		Hybridní přístup	91,40 %		
Sentinel-2	10 m	Random Forest	90,16 %	Turecko	Colkesen a Kavzoglu, 2017
Landsat 8	15 m		86,18 %		
Sentinel-2	10 m	Maximum Likelihood Classifier	87,10 %	Chorvatsko	Gašparović a Jogun, 2018
		Neuronová síť	83,60 %		
		Objektově založená klasifikace (OBIA)	87,60 %		
Sentinel-2	20 m	Random Forest	90,71 % až 94,7 %	severní Vietnam	Thanh Noi a Kappas, 2018
		Nearest-Neighbour	89,80 % až 94,59 %		
		Support Vector Machine	92,35 % až 95,32 %		

V kontextu této disertační práce není klasifikace zemského povrchu cílem sama o sobě. Její význam spočívá především v podpoře výběru vhodné lesní masky pro následné modelování nadzemní biomasy. Volba lesní masky je zásadní nejen pro prostorové rozložení výsledné predikce, ale také pro výpočet celkové zásoby nadzemní biomasy a uhlíku. Mezinárodně používaná definice FAO vymezuje les jako území o rozloze větší než 0,5 ha se stromy vyššími než 5 m a korunovým zápojem přesahujícím 10 %, případně se stromy schopnými těchto parametrů dosáhnout *in situ*. Do této definice nejsou zahrnována území využívaná převážně pro zemědělské nebo urbanizované účely (FAO, 2012). Toto harmonizované vymezení, používané také evropskou sítí národních lesních inventarizací ENFIN (European National Forest Inventory Network), zvyšuje srovnatelnost lesnických statistik napříč Evropou.

Obdobné vymezení používá také Národní inventarizace lesů České republiky, která rozlišuje několik kategorií pozemků. Kategorie les, která je rozlišena podle definice FAO zahrnuje také dočasně odlesněné plochy, například holiny, kalamitní holiny, požářiště nebo plochy po přípravě

půdy pro obnovu, u nichž existuje předpoklad budoucího dosažení požadovaného zápoje stromů. Naopak nejsou zahrnovány porosty s šířkou menší než 20 m ani porosty a pozemky s převážně zemědělským nebo městským využitím. NIL dále odlišuje nelesní porosty dřevin, ostatní pozemky s porostem stromů a ostatní pozemky (Kučera a Adolt, 2019). Výsledná rozloha lesa podle NIL je odvozena z inventarizačního šetření, tedy ze sítě inventarizačních ploch, na nichž je v terénu hodnocen skutečný stav území v době šetření. Získané informace jsou následně statisticky přepočítány na celé území České republiky, a proto jsou výsledky NIL uváděny jako odhady s příslušnou nejistotou.

Odlišný pohled poskytují údaje Českého úřadu zeměměřického a katastrálního, které vycházejí z katastru nemovitostí a evidence druhů pozemků. Z hlediska lesní plochy je klíčovou kategorií druh pozemku „lesní pozemek“. Takto vymezená plocha vyjadřuje především právně-evidenční stav pozemků v katastru, nikoliv inventarizační vyhodnocení skutečného stromového pokryvu nebo zápoje korun v době šetření. Katastrální evidence proto může zahrnovat pozemky vedené jako lesní, které se svým aktuálním stavem mohou odlišovat od inventarizačního pojetí lesa, a naopak některé plochy splňující inventarizační definici lesa nemusí být evidovány jako lesní pozemek. Rozdíly mezi údaji NIL a ČÚZK tak vyplývají především z rozdílného účelu obou zdrojů.

Další skupinu představují *land cover* databáze, které vymezují les především na základě aktuálně pozorovatelných vlastností zemského povrchu a použitého klasifikačního schématu. Do výsledné lesní třídy se promítá prostorové rozlišení dat, minimální mapovací jednotka, období snímkování, tematické rozlišení tříd, použitý klasifikační algoritmus a kvalita trénovacích nebo referenčních dat. Les v *land cover* databázi tak zpravidla reprezentuje obrazově nebo spektrálně identifikovaný stromový pokryv, který nemusí být totožný ani s inventarizační kategorií lesa podle NIL, ani s katastrální kategorií lesního pozemku podle ČÚZK.

Jiný typ informace poskytují lesnicko-typologické podklady, například vrstva souborů lesních typů. Lesnicko-typologický systém není primárně určen k obrazové detekci aktuálního korunového pokryvu, ale k diferenciaci růstových a stanovištních podmínek lesů a k popisu lesního prostředí (Plíva, 1987). Lesní typ představuje základní jednotku diferenciace růstových podmínek a soubor lesních typů sdružuje ekologicky příbuzné lesní typy s obdobnými hospodářsky a stanovištně významnými vlastnostmi. Soubory lesních typů jsou vymezovány kombinací půdních kategorií a lesních vegetačních stupňů (Viewegh, 2003). Vzhledem k návaznosti na lesnické pojetí území a pozemky určené k plnění funkcí lesa představují tyto podklady odlišný, ale pro lesnické analýzy relevantní způsob vymezení lesního prostředí. Na rozdíl od *land cover* produktů nejsou založeny pouze na aktuálně pozorovatelném stromovém pokryvu, ale reflektují stanovištní a typologické charakteristiky lesa. Při jejich využití je však nutné zohlednit, že výsledná plocha nemusí být totožná s inventarizačním odhadem rozlohy lesa, katastrální evidencí lesních pozemků ani obrazově odvozenými produkty lesního pokryvu.

2.2.1. Datové zdroje krajinného pokryvu a lesních charakteristik

Díky příchodu nových technologiích a zlepšení družicového dálkového průzkumu Země, je nyní možné, díky široké škále senzorů, kvantifikovat ekosystémové funkce. Jedná se o kvantifikaci

přímou (definovanou fyzikální jednotkou) anebo nepřímo (bezrozměrnou jednotkou, pomocí indexu).

Někteří operátoři nebo přidružené organizace satelitních misí vytvářejí již hotové datové sady indikátorů ekosystémových funkcí. Typickým příkladem je datová sada pozemního pokryvu nebo lesního pokryvu, díky kterým lze odhalit rozsah daného habitatu. Některé datové sady ovšem trpí nízkým prostorovým rozlišením. Dalším problémem může být neharmonizace klasifikačních schémat. Regionální a globální klasifikační schémata jsou odlišné kvůli požadavkům koncových uživatelů, a hlavně kvůli rozlišovací schopnosti použitých dat. Globální datové sady jsou určeny pro použití ve všech částech světa, kdežto regionální pouze v daném státě nebo vybrané klimatické oblasti. Zde je výhoda družicových dat s vysokým a velmi vysokým rozlišením, které dokážou odhalit variabilitu pokryvu Země i ve větších měřítcích. Přesnost, rozlišení a pokrytí se zvýšilo díky použití dálkového průzkumu Země oproti mapování prováděnému pozemními metodami. Nevýhodou těchto datových sad je pevně dané aktualizační období. Pokud je potřeba aktuální *land cover* území, je potřeba využít aktuální datové snímky a provést identifikaci. Globální nebo kontinentální datové sady o *land cover* poskytuje spousta organizací. ESA (European Space Agency) produkuje globální databázi pokryvu země ESA WorldCover s prostorovým rozlišením 10 m, která je dostupná pro referenční roky 2020 a 2021 a je odvozena ze společného využití dat Sentinel-1 a Sentinel-2. Její přesnost je v Evropě 77,9 +/- 1 % a celkově 76,7 +/- 0,5 % (Zanaga a kol., 2022). Příklady volně dostupných *land cover* databází jsou v tabulce 2.

Tabulka 2 Přehled *land cover* databází

Databáze	Operátor	Prostorové rozlišení	Počet tříd	Poslední verze	Pokrytí	Zdroj
GLC2000	European Environment Agency	250 m	22	2000	Svět	EC, 2003
GLC-SHARE	Food and Agriculture Organization	1000 m	11	2012	Svět	Latham a kol., 2014
Corine Land Cover	Copernicus	100 m	44	2018	Evropa	EU, 2022
National Land Cover Database	United States Geological Service	30 m	16	2024	USA	Dewitz, 2021; USGS, 2024
Globeland30	National Geomatics Center of China	30 m	10	2020	Svět	Chen a Chen, 2018
ESRI LULC	ESRI	10 m	10	2024	Svět	Karra a kol., 2021; ESRI, 2026
ESA World Cover	European Space Agency	10 m	11	2021	Svět	ESA, 2021

Kromě databází o celkovém *land cover* jsou dostupné i dílčí habitaty. K lesním biotopům jsou dostupné databáze o výšce lesa z roku 2019 v rozlišení 30 metrů globálně (Potapov a kol., 2021), změna pokryvnosti lesů v průběhu let globálně (Hansen a kol., 2013) nebo pro Evropu hustota zalesnění, dominantní typ lesa a lesní typ v rozlišení 10 metrů (Copernicus, 2022). Dalším

příkladem může být například odhad nadzemní biomasy nebo primární produkce. První globální databází nadzemní biomasy je GlobBiomass z roku 2010. Vytvořena byla z radarových, optických i lidarových dat. Většina validačních bodů se nachází v Evropě (42003 bodů) a rozlišení databáze je 100 metrů v oblasti rovníku. V Evropě je střední kvadratická chyba 31,3 $t \cdot ha^{-1}$, což je relativně 40,3 % (Santoro, 2018). Další globální databází nadzemní biomasy je ESA Biomass Climate Change Initiative. Aktuálně je dostupná verze 6.0, která poskytuje globální mapy nadzemní biomasy lesa (AGB) pro roky 2007, 2010 a každoročně 2015–2022 (Santoro a Cartus, 2025). Jedná se o datovou sadu s prostorovým rozlišením 100 m. Primárními vstupy pro tvorbu produktu jsou radarová data (zejména Sentinel-1), přičemž kvalita a konzistence produktů je dokumentována ve validační a interkomparační zprávě. Pro mapu 2022 je ve validační zprávě při porovnání s referenčními terénními plochami (Tier-1) na původním rozlišení uvedena průměrná odchylka MD = −14 $t \cdot ha^{-1}$ a RMSD = 168 $t \cdot ha^{-1}$ (n = 49 927). Při prostorové agregaci na 0,1° se RMSD snižuje a závisí na úrovni biomasy; například pro interval 0–50 $t \cdot ha^{-1}$ je RMSD 30 $t \cdot ha^{-1}$, zatímco pro 100–150 $t \cdot ha^{-1}$ je RMSD 74 $t \cdot ha^{-1}$. Stejně tak je dostupná i databáze primární produkce MODIS Gross Primary Production (GPP)/ Net Primary Production (NPP). Databáze má rozlišení 1 km (Running a Zhao, 2015). NASA ovšem plánuje ukončit sběr dat ze senzoru MODIS na družici Terra v roce 2027, přičemž alternativou k produktu MOD17 bude produkt VNP17 ze senzoru VIIRS (Visible Infrared Imaging Radiometer Suite). Přehled vybraných globálních a evropských produktů zaměřených na lesní charakteristiky, výšku porostu, nadzemní biomasu a uhlíkovou hustotu je uveden v Tabulce 3.

Tabulka 3 Dostupné produkty lesních charakteristik

Produkt (datová sada)	Sledovaná charakteristika	Prostorové a časové rozlišení	Vstupní data
Hansen Global Forest Change	Stromový pokryv, ztráta a přírůstek pokryvu	30 metrů; Od roku 2000 každoroční informace o ztrátě pokryvu, přírůstek pouze za období 2000–2012	Landsat 5, 7, 8, 9
Copernicus HRL Forest Layers	Hustota stromového pokryvu, dominantní listový typ a typ lesa	10 metrů; pravidelná aktualizace	Družicová data programu Copernicus
High Resolution 1m Global Canopy Height Maps (Tolan a kol., 2024)	Výška stromového porostu	1 metr; globální produkt z dat roku 2018 až 2020	Maxar, lidar GEDI
High-resolution Canopy Height Model of the Earth (Lang a kol., 2023)	Výška stromového porostu	10 metrů; globální produkt pro rok 2020	Sentinel-2, lidar GEDI
ESA CCI Biomass v6	Lesní nadzemní biomasa	100 metrů; dostupné roky 2007, 2010 a 2015 až 2022	Radarová data, zejména Sentinel-1 a další
GEDI L4B	Hustota nadzemní biomasy, nejedná se o souvislou mapu.	1000 metrů; Období 2019 až 2021	Lidar GEDI (ISS)
Global Aboveground and Belowground Biomass Carbon Density Maps for the Year 2010 (NASA CMS)	Hustota uhlíku v nadzemní a podzemní biomase	300 metrů, rok 2010	Mapy biomasy (Globbiomass), MODIS tree cover a ESA CCI Land Cover 2010, NDVI

Uvedené produkty poskytují cenný přehled o vybraných charakteristikách lesních ekosystémů, jejich přímé využití však může být omezeno rozdílným prostorovým rozlišením, referenčním obdobím a použitou definicí lesa. Pro detailní hodnocení na národní nebo regionální úrovni proto může být vhodné vytvořit vlastní model sledované biofyzikální charakteristiky. Ten lze přizpůsobit konkrétnímu území, dostupným pozemním datům a požadovanému prostorovému rozlišení. Přesnost vlastního modelu závisí především na množství, kvalitě a prostorové reprezentativnosti trénovacích a validačních dat, na variabilitě modelované veličiny a na výběru vhodných prediktorů. Mezi často využívané prediktory patří optická a radarová družicová data, vegetační indexy, texturní charakteristiky, informace o struktuře porostu a doplňkové prostorové vrstvy popisující topografické nebo stanovištní podmínky.

2.3. Dálkový průzkum ve výzkumu funkcí ekosystému

Optické senzory produkují širokou škálu dat. Hlavní výhodou dálkového průzkumu Země v procesu analýzy ekosystémových funkcí je sběr aktuálních, opakovatelných a nedestruktivních dat. Zpracovaná data umožňují provádět diskrétní i kontinuální monitoring. Data jsou vždy prostorově lokalizována a umožňují zachytit prostorovou variabilitu, protože v jednom okamžiku pokryjí velké území. Hlavním úskalím této metody jsou atmosférické vlivy. Na snímky je nutné použít atmosférické korekce. Multispektrální senzory se používají především pro družicový dálkový průzkum. Snímají nejen ve viditelné části, ale také v blízkém nebo střední infračervené části elektromagnetického spektra. V některých případech snímají v tepelném infračerveném nebo panchromatickém pásmu. Panchromatické data se vyznačují vysokým nebo velmi vysokým prostorovým rozlišením a pouze jedním snímaným pásmem. Pásmo je tvořeno celkovou světelnou energií ve viditelném spektru. Multispektrální data mají ve srovnání s panchromatickými daty vyšší spektrální rozlišení tvořené více spektrálními pásy, nicméně zpravidla jejich prostorové rozlišení je nižší. Panchromatické pásmo se tak využívá pro zvýšení prostorového rozlišení u multispektrálních snímků pomocí metody pan-sharpening. Další možností je použití hyperspektrálních snímků. Hyperspektrální snímky mají nejvyšší spektrální rozlišení (vysoký počet a velmi úzká spektrální pásma) a dokáží nejlépe pokrýt variabilitu odrazivosti. Podle Liang a kol. (2012) jsou data rozdělena podle prostorového rozlišení do čtyř kategorií:

- nízké rozlišení (> 1000 m):
    - Sentinel-5P, Meteosat MSG, DSCOVER, NOAA-AVHRR;
- střední rozlišení (100-1000 m):
    - Sentinel-3, NOAA-VIIRS, TERRA-MODIS;
- vysoké rozlišení (5-100 m):
    - Sentinel-2, Landsat-8, Landsat-9, SPOT 5;
- velmi vysoké rozlišení (<5 m):
    - Pléiades Neo-3, Pléiades-1, WorldView-4, QuickBird, Ikonos.

Jak bylo uvedeno výše, optické senzory produkují velké množství dat. Pokud se zaměříme na volně dostupná optická data, seznam se výrazně zmenší. V současné době je nejlepší množství využít data z družic Landsat nebo Sentinel. Kromě aktuálních snímků lze stáhnout i archivy snímků. V následující tabulce (Tabulka 4) jsou uvedeny parametry volně dostupných dat. Družice Landsat 7 není uvedena, jelikož jsou dostupné novější verze.

Tabulka 4 Volně dostupná optická data

Satelit	**Landsat 8**	**Landsat 9**	**Sentinel-2**	**TERRA**	**NOAA-21 (JPSS-2)**
Operátor	NASA, USGS	NASA, USGS	ESA	NASA	NOAA
Datum vyslání	11.02.2013	27.09.2021	S2A= 23. 6. 2015; S2B = 7. 3. 2017	18.12.1999	10.11.2022
Oběžná dráha	Polární ve výšce 705 km	Polární ve výšce 705 km	Polární ve výšce 786 km	Polární ve výšce 705 km	Polární ve výšce 833 km
Inklinace orbity	98,2 °	98,2 °	98,62 °	98.2098 °	98.75 °
Časové rozlišení	16 dní (8 dní v kombinaci s Landsat 9)	16 days (8 dní v kombinaci s Landsat 8)	10 dní pro S2A, S2B (5 dní společně)	1–2 dny	16 dní
Senzory	OLI; TIRS	OLI-2; TIRS-2	MSI	ASTER	VIIRS
Prostorové rozlišení	15 m a 30 m (OLI); 100 m (TIRS)	15 m a 30 m (OLI-2); 100 m (TIRS-2)	10 m, 20 m a 60 m	15 m, 30 m a 90 m	375 m a 750 m
Radiometrické rozlišení	12-bit (4096 potenciálních hodnot)	14-bit (16384 potenciálních hodnot)	12-bit (4096 potenciálních hodnot)	8-bit (256) pro VNIR, SWIR; 12-bit (4096) pro TIR	12-bit (4096 potenciálních hodnot)
Spektrální pásma	11 pásem	11 pásem	13 pásem	14 pásem	22 pásem
Zdroj	NASA, 2022	NASA, 2022	Copernicus Sentinel online, 2022	NASA MODIS, 2022	NASA TERRA, 2022

Významnou studií v daném směru představuje studie Pettorelli a kol. (2018). Podle studie byly ekosystémové funkce zkoumány na velkých plochách jen velmi zřídka. Biodiverzita byla sledována především podle strukturálních a kompozičních charakteristik, nikoliv však podle funkčních vlastností. Pokusy o sledování ekosystémových funkcí byly prováděny na malých plochách. Většina hodnocení ekosystémů pak nezohledňuje funkce kvůli nedostatku vhodných prostorových dat. Ekosystémové funkce mohou v některých případech reagovat rychleji na změny prostředí než strukturální nebo kompoziční atributy a mohly by patřit k nejcitlivějším indikátorům změn při globálním monitorování ekosystémů. Ve studii je pojednáváno i o tom, že vznikají diskuse na téma koordinovaného monitorování ekosystémů (Oliver a kol., 2015), nicméně praktická realizace stále chybí. Ve studii jsou shrnuty informace o rozdělení ekosystémových funkcí, příslušné indikátory, které mají možnost ekosystémovou funkci identifikovat a následně senzory, které jsou schopné daný indikátor zachytit. Přehled senzorů je rozdělen do dvou skupin, a to na již vzniklé studie s použitím daného senzoru a taky budoucí potenciální využití senzoru. Například družice Sentinel-2 je vhodná k určení NDVI (*Normalised*

Difference Vegetation Index), LAI (*Leaf Area Index*), FAPAR (*Fraction of Absorbed Photosynthetically Active Radiation*) nebo *land cover*, které jsou dále spojeny s ekosystémovými funkcemi jako primární produkce či sekvestrace uhlíku. Družice Landsat mají potenciál kvantifikovat ekosystémové funkce spojené s teplotou povrchu, Sentinel-1 pak nadzemní biomasu a vlhkost půdy (Copernicus, 2022). Podle Wagner a kol. (2010) má na určení vlhkosti půdy velký vliv drsnost a přítomná vegetace, které negativně ovlivňují hodnoty odrazivosti. Při určení vlhkosti holé půdy a kukuřice byla zjištěna odchylka 7 %, přičemž u ostatních typů půdního krytu byly zjištěny daleko větší odchylky (Esch, 2018). K dobrému určení vlhkosti půdy dochází v malých měřítkách, kde je redukován radarový šum. Copernicus Global Land Service vytváří produkt SSM (Surface Soil Moisture), který je v rozlišení 1 km.

Další studie s podobnou problematikou, kde se autoři zaměřují především na dálkový průzkum v oblasti zemědělství (Weiss a kol., 2020). Studie se zaměřuje na proměnné, které lze získat z dálkového průzkumu Země. Podle Nock a kol. (2016) jsou definovány proměnné jako vlastnosti, které se liší od jedince (rostlina), skupiny (plodina) až po společenství (oblast). Dané vlastnosti tak můžou být typologické (typ plodiny), fyzikální (teplota povrchu rostliny nebo vlhkost půdy), chemické (obsah dusíku v listech), biologické (fenologie rostlin), strukturální (sestavení listů) nebo geometrické (hustota rostliny). Autor zdůrazňuje, že žádná z těchto vlastností není přímo měřitelná pomocí dálkového průzkumu Země. Příbuznost mezi radiancí a vlastnostmi je nutno nějakým způsobem namodelovat. Jsou tak zkoumány příbuznosti mezi jednotlivými metodami dálkového průzkumu a vybranými agronomickými vlastnostmi. Například z dat LIDAR (metoda) lze zjistit výšku plodiny (vlastnost), z optických dat je možno zjistit Green Area Index nebo výnos plodiny, z radarových dat pak vlhkost půdy. Green Area Index a výška plodiny se dá považovat jako primární proměnnou (jsou přímo spojeny s procesem radiance), kdežto výnos plodin je sekundární proměnnou (je nepřímo dostupná z dat DPZ). Obecně se stále využívá ke sledování ekosystémových služeb primárních proměnných, jako NDVI, LAI, *land cover*, teplota povrchu nebo výška vegetace (Cord a kol., 2017; De Araujo Barbosa a kol., 2015). Pokud je zkoumána sekundární proměnná, ovlivňuje vztah mezi radiancí a danou proměnnou více faktorů. Další proměnnou, která má zásadní význam pro fotosyntézu, hydrologické procesy a přenos záření z korun stromů, ale není uvedena v ECV, je Clumping index (CI) (Fang, 2021). Ci souvisí s LAI a je definován jako poměr efektivního indexu listové plochy (LAIe) k indexu listové plochy (LAI) (Chen, 2018).

Hodnoty CI se pohybují mezi 0,3 (shluknuté listy) a 1 (náhodně rozmístěné listy). Pro různé typy vegetace se CI řídí pořadím tráva, plodiny, keře a lesy (Lesy jsou více shluknuté než tráva) (Wei a kol., 2019). Index lze získat přímou, alometrickou, nepřímou optickou nebo proxy metodou. Studie Fanga (2021) komplexně popisuje metody a využití CI. Index lze získat z dálkového průzkumu Země pomocí normalised difference hotspot and darkspot index (NDHD). Studie vyvozuje primární proměnné a která metoda sběru dat z DPZ je nejvhodnější. Jako nejvhodnější zdroj lze označit optický DPZ, který lze použít pro širší škálu proměnných, nicméně záleží na dané potřebě použití (Tabulka 5).

Tabulka 5 Nejvhodnější metody pro sběr primárních proměnných

Primární proměnné	Nejvhodnější data	Příklad použití
Plant Density	multispektrální/hyperspektrální data	Bai a kol., 2011
GAI/LAI	multispektrální/hyperspektrální data	Mananze a kol., 2018
Green Cover Fraction	multispektrální/hyperspektrální data	Filipponi a kol., 2018
Leaf Biochemical Content	multispektrální/hyperspektrální data	Zhang a kol., 2008
Orientace listů	Fotogrammetrie/LIDAR	Zheng a Moskal, 2012
Výška	Fotogrammetrie/LIDAR	Streutker a Glenn, 2006
FAPAR	multispektrální/hyperspektrální data	Tao a kol., 2016
Albedo	multispektrální/hyperspektrální data	Bright a Astrup, 2019
Teplota (půdy/vegetace)	multispektrální/hyperspektrální data	Kumar a Shekhar, 2015
Vlhkost půdy	multispektrální/hyperspektrální/radarová data	Gao a kol., 2017

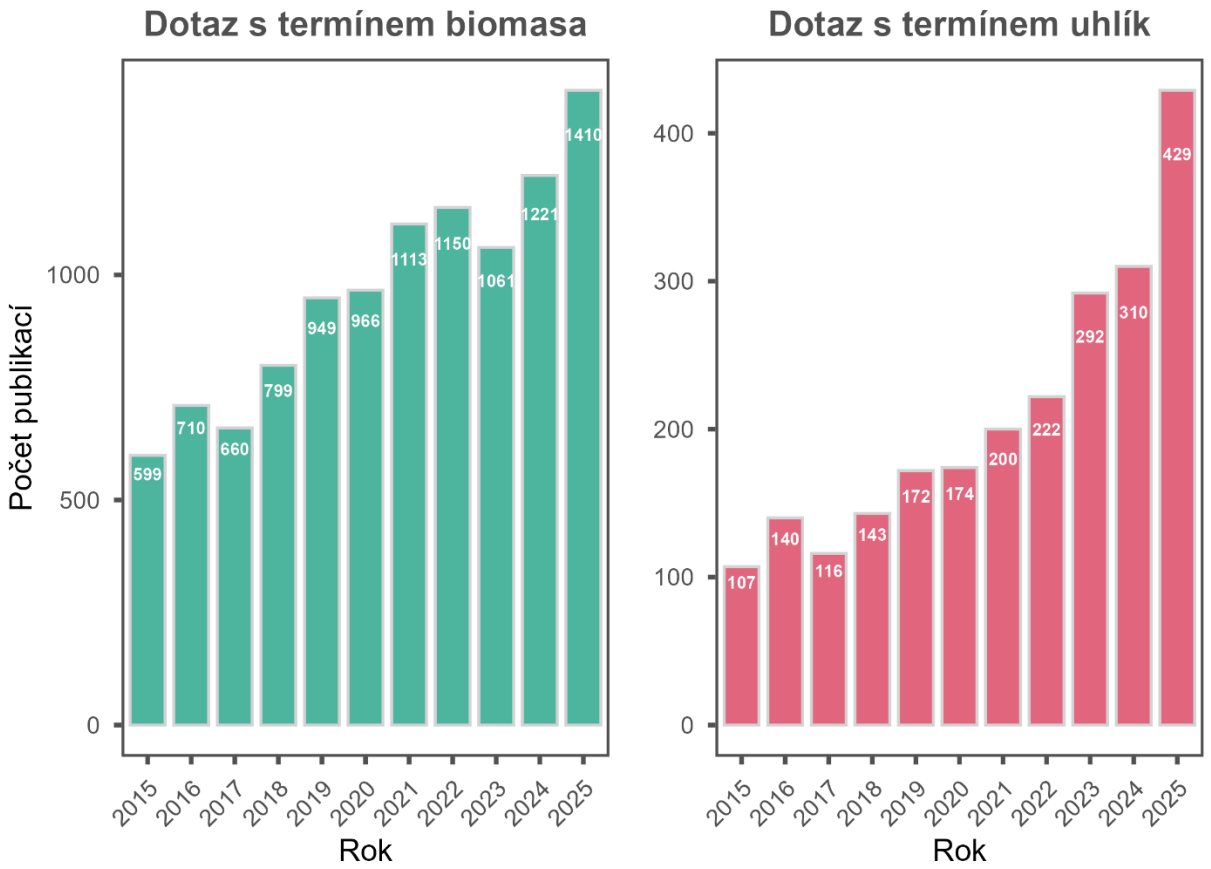
2.4. Nadzemní biomasa a uhlík v lesních porostech

Lesy představují komplexní ekosystémy tvořené vzájemně interagující vegetací, houbami a dalšími organismy, jejichž struktura a fungování se mění v závislosti na klimatu, druhové skladbě a stanovištních podmínkách. Jedním ze základních deskriptorů lesní struktury je nadzemní biomasa (above-ground biomass, AGB), která integruje informaci o produktivitě porostu, vývoji porostu a stavu ekosystému. Přesná kvantifikace lesní AGB je proto zásadní pro hodnocení ekosystémů, monitoring lesů i kvantifikaci suchozemských zásob uhlíku. Les je v definicích FAO, Evropské komise, Eurostatu a Evropské agentury pro životní prostředí vymezen minimálními prahovými hodnotami plochy, výšky stromů a zápoje korun (FAO, 2012). Toto harmonizované vymezení, používané rovněž evropskou sítí národních lesních inventarizací ENFIN, zvyšuje srovnatelnost lesnických statistik napříč Evropou, současně však klade nároky na odhad biomasy z dat dálkového průzkumu Země v prostorově a strukturně heterogenních krajinách; to je zvláště relevantní pro temperátní lesy střední Evropy. Hlavním cílem ENFIN je přispívat k naplňování nové Evropské lesní strategie do roku 2030 (European Commission, 2021).

2.4.1. Analýza publikační aktivity v oblasti nadzemní biomasy a uhlíku

Pro zmapování vývoje výzkumu zaměřeného na odhad biomasy a zásob uhlíku z dat DPZ byla provedena rešerše ve Web of Science Core Collection. Vyhledávání bylo provedeno dne 20. 2. 2026. Vyhledávání bylo omezeno na období 2015–2025 a bylo sestaveno na základě tří tematických bloků: sledovaný parametr, typ použitých dat a typ analytické úlohy. Byly použity dvě varianty dotazu, z nichž první byla zaměřena na biomasu a druhá na uhlík. V první variantě zahrnoval blok sledovaného parametru termíny *biomass*, *aboveground biomass* a *AGB*, zatímco ve druhé variantě byly použity termíny *carbon stock* a *carbon storage*. Blok datových zdrojů tvořily

termíny *remote sensing*, *earth observation*, *LiDAR*, *SAR*, *optical* a *multispectral*. Blok analytické úlohy zahrnoval výrazy *estimat**, *mapp**, *predict** a *quantif**. Pro variantu dotazu zaměřenou na biomasu bylo v období 2015–2025 ve Web of Science Core Collection nalezeno 10 638 záznamů. Počet publikací v čase převážně rostl, přičemž výraznější nárůst lze pozorovat zejména od roku 2021. V roce 2025 byl počet záznamů více než dvojnásobný oproti roku 2015. Pro variantu dotazu zaměřenou na uhlík bylo ve stejném období nalezeno 2 305 záznamů. I zde je patrný postupný růst počtu publikací, nicméně celkový počet nalezených záznamů byl výrazně nižší než v případě varianty zaměřené na biomasu. Ve srovnání s rokem 2015 byl počet záznamů v roce 2025 přibližně čtyřnásobný (Obrázek 4). Uvedené počty představují výsledky databázového vyhledávání na základě zvolených klíčových termínů a slouží především k orientačnímu zachycení vývoje publikační aktivity v dané oblasti.



Obrázek 4 Vývoj počtu záznamů ve Web of Science Core Collection v letech 2015–2025 pro variantu dotazu zaměřenou na biomasu a variantu dotazu zaměřenou na uhlík.

2.4.2. Význam nadzemní biomasy a uhlíku v lesních ekosystémech

Nadzemní biomasa (AGB) představuje klíčový ukazatel struktury a produktivity lesních porostů a současně jeden z hlavních vstupů pro odhad zásob uhlíku v biomase. Lesy fungují jako významné zásobníky uhlíku: běžně se uvádí, že živá biomasa ukládá přibližně 42 % lesního uhlíku, půda přibližně 44 %, zatímco mrtvé dřevo a opad přispívají zhruba 8 %, resp. 5 % (Pan a kol., 2011). Projekce dále naznačují, že relativní příspěvek temperátních a boreálních lesů k celkové

sekvestraci uhlíku může v budoucnu růst (Tagesson a kol., 2020). AGB lze stanovit destruktivním vzorkováním nebo nepřímo pomocí terénních měření a alometrických rovnic (Liepiņš a kol., 2018; Kučera a Adolt, 2019; Dutcă a kol., 2020). Pro plošné aplikace se uplatňuje tzv. plošně založený přístup (area-based approach, ABA), kde jsou AGB odhady odvozovány z prediktorů získaných z dálkového průzkumu Země a kalibrovány na terénní referenční data. Za jednu z nejpřesnějších metod je považováno letecké laserové skenování, které poskytuje bodová mračna využitelná pro odvození výškových a strukturních prediktorů (Brovkina a kol., 2022). V národním či kontinentálním měřítku však představují praktickou alternativu volně dostupná optická a radarová družicová data a také jejich kombinace s družicovými lidarovými misemi.

Pro srovnání s národně optimalizovanými produkty jsou relevantní i globální databáze. Mezi nejpoužívanější patří ESA CCI Global Forest Above-Ground Biomass (Santoro a Cartus, 2025) a globální databáze zásob uhlíku a toků (Ma a kol., 2023), která je odvozena především z lidarových pozorování GEDI (Global Ecosystem Dynamics Investigation) a ICESat-2. Zatímco ESA poskytuje AGB, produkty NASA CMS typicky poskytují odhady zásob uhlíku; pro převod AGB na uhlík se často používá obecný koeficient 0,50 (Baccini a kol., 2017), přičemž Národní inventarizace lesů v ČR rozlišuje koeficienty 0,51 pro jehličnaté a 0,49 pro listnaté dřeviny (Ústav pro hospodářskou úpravu lesů, 2024). Recentní studie dále naznačují, že zásoba uhlíku se neliší pouze mezi jehličnatými a listnatými porosty, ale může být zvýšena i ve smíšených porostech, kde komplementární růstové strategie podporují vyšší AGB a akumulaci uhlíku ve srovnání s porosty čistými (Xu a kol., 2024). Některé práce uvádějí, že oddělená kalibrace modelů pro jehličnaté a listnaté porosty může zvýšit predikční přesnost, s typickým zlepšením R^2 přibližně o 0,05–0,10 (Wegler a kol., 2025). V této disertaci je však záměrně uplatněn jednotný modelový rámec napříč typy lesa s cílem ověřit, do jaké míry lze pomocí volně dostupných dat Sentinel-1 a Sentinel-2 vytvořit robustní národní mapu AGB a následně i zásob uhlíku v biomase při omezeném rozsahu terénních referenčních dat.

Disertační práce je primárně zaměřena na stanovení nadzemní biomasy (AGB) a odvození zásob uhlíku v biomase, které představují klíčové strukturální charakteristiky lesních ekosystémů. AGB je současně úzce provázána s řadou ekosystémových funkcí a služeb, a proto je v této práci zasazena do širšího funkčního kontextu. V tomto rámci jsou stručně zmiňovány zejména vazby AGB na indikátory, jako je zásoba biomasy a uhlíku, produkce biomasy, sekvestrace uhlíku, rozloha a kvalita habitatů. Tyto indikátory spadají především do regulačních a zásobovacích (produkčních) funkcí ekosystémů, přičemž AGB na ně může přímo navazovat nebo je významně ovlivňovat.

Regulační funkce zmírňují výkyvy přirozeného i antropogenního původu a množství biomasy se v tomto kontextu pojí např. s regulací odtoku (O'Riordan a kol., 2021), eroze (Almagro a kol., 2019; Kort a kol., 1998), disturbancí (Rabin a kol., 2017), koncentrace oxidu uhličitého prostřednictvím sekvestrace uhlíku (Malhi a kol., 2015) a také s regulací teploty (Schwaab a kol., 2021). Zásobovací funkce pak zahrnují poskytování potravy, materiálu či úkrytu organismům. Skupiny funkcí spolu úzce souvisejí, protože jejich společným základem je zelená složka vegetace a s ní spojené ekofyziologické procesy. Tato složka je zároveň dobře zachytitelná optickými senzory. Na základě

spektrálního chování vegetace lze při dostatečném prostorovém rozlišení rozlišovat typy stanovišť a hodnotit jejich kvalitu či zdravotní stav, identifikovat vegetaci se zhoršenou kondicí a vymezovat prostorové hranice habitatů. Produkce biomasy představuje tok biomasy a vyjadřuje se v $kg \cdot m^{-2} \cdot rok^{-1}$. Zásoba a sekvestrace uhlíku jsou v rámci hodnocení ekosystémových služeb úzce související, avšak odlišné veličiny. Například přístup InVEST pracuje se zásobou uhlíku ve čtyřech základních rezervoárech, kterými jsou nadzemní biomasa, podzemní biomasa, půda a mrtvá organická hmota, přičemž sekvestrace představuje změnu této zásoby v čase (Tallis a kol., 2011). Zásoba biomasy a sekvestrace uhlíku jsou přitom úzce propojené: uhlík je během fotosyntézy vázán do rostlinné biomasy, kde se akumuluje. V podmínkách České republiky se pro převod nadzemní biomasy na uhlík používají koeficienty Národní inventarizace lesů, přičemž AGB je nejčastěji vyjadřována v $t \cdot ha^{-1}$ ($Mg \cdot ha^{-1}$). Pro kvantifikaci nadzemní biomasy na konkrétním stanovišti je nezbytné mít vymezené hranice daného habitatu. Výběr těchto souvislostí a jejich vazba na habitatové typy v této práci navazuje na přístup založený na expertních hodnotách, které byly přiřazeny jednotlivým habitatům a dále agregovány pomocí vážené metody (Pechanec a kol., 2022).

2.4.3. Přístupy k odhadu nadzemní biomasy z dat dálkového průzkumu Země

Optický dálkový průzkum snímá odražené záření a díky různým spektrálním pásmům poskytuje hodnoty záření v širším spektru. Každé pásmo má specifickou odrazivost pro určitý typ povrchu. Tato pásma dále vstupují do modelování ekosystémových funkcí jako prediktory. Některé primární proměnné lze určit pomocí vegetačních indexů (primární produkce, zásoba biomasy nebo c-faktor). Přesnějších výsledků se však dosáhne pomocí vlastního modelování vybrané proměnné, pokud jsou k dispozici přesná trénovací a validační data. Modelování se provádí pomocí modelovacích algoritmů jako jsou RF (Breiman, 2001), SVM (Noble, 2006), ANN (artificial neural network) (Jain a kol., 1996) nebo XGBoost (extreme gradient boosting) (Chen a kol., 2015). Vedle těchto přístupů se v odhadu nadzemní biomasy uplatňují také hluboké neuronové sítě a hybridní architektury, zejména v případech, kdy jsou kombinována multispektrální, radarová a lidarová data (Tian a kol., 2024; Jiang a kol., 2025).

Jak již bylo zmíněno, ekosystémové funkce nebo primární proměnné nejsou přímo měřitelné z dálkového průzkumu Země. V mnoha případech jsou funkce zjistitelné na základě vztahu k vegetačním indexům. Mnoho studií využívá vegetační indexy, zejména NDVI, k odhadu nadzemní biomasy (Van Der Meer a kol., 2000; Liu a kol., 2015; Gizachew a kol., 2016; Anwar a kol., 2018; Macedo a kol., 2018; Muhd-Ekhzarizal a kol., 2018; Berger a kol., 2020). Výsledky nejsou vyvozeny přímo z daných vegetačních indexů, ale jsou pomocí nich modelovány. Vegetační indexy slouží jako vstupní data (prediktory) pro modelování spolu se spektrálními pásy. Do modelování může vstupovat více proměnných, jako digitální model reliéfu (Gao a kol., 2018; Chen a kol., 2018), podíl vegetace (Chang a Shoshany, 2016; Lovynska a kol., 2020), půdní pokryv (Running a Zhao, 2015) nebo Grey-Level Co-Occurrence Matrix (GLCM) (Chen a kol., 2018; Li a kol., 2021). Tyto prediktory mohou přispět ke zlepšení výsledné přesnosti modelu. Nicméně, záleží na funkcích, které jsou pomocí prediktorů modelovány. Pro některé funkce jsou nezbytné další

prediktory, jako jsou bioklimatické proměnné, údaje o půdě nebo rychlost větru. V některých případech jsou postradatelné a nepřinášejí zlepšení. Odhad primární produkce pomocí pouze satelitních dat byl přesnější, než odhad kombinovaný s meteorologickými daty (Rossini a kol., 2012; Ma a kol., 2014). Některé studie uvádějí, že modelování primární produkce bez použití CI (clumping index) podhodnocuje výsledky, zejména v tropických oblastech. Byly tak porovnávány globální modely GPP bez použití a s použitím indexu a bylo zjištěno zvýšení primární produkce o 5,53 PgC·rok^{-1} při použití CI. Nicméně, model je účinnější v tropických oblastech (kde je vysoký LAI), mezi 20° s. š. a 20° j. š. došlo k nárůstu o 4,18 PgC ·rok^{-1} (Braghiere a kol., 2019). Jediná studie, která demonstrovala opačný účinek CI používala jiný model i přístup. S použitím CI tak došlo ke snížení GPP o 12,1 PgC·rok-1 (Chen a kol., 2012). Na rozdíl od předchozí studie nepoužívá více vrstev koruny stromů s různým podílem osluněných a zastíněných listů, ale pouze jednu vrstvu rozdělenou na osluněné a zastíněné listy. Další studie, která zkoumala vliv CI na přenos záření do porostu a míru fotosyntézy, poukázala na to, že použití globálních datasetů CI nadhodnocuje FAPAR hodnoty osluněných listů a podhodnocuje zastíněné listy. Při porovnání globálního GPP s použitím CI bylo opět zjištěno zvýšení o 1 PgC·rok-1 (Li a kol., 2023).

Přesnost odhadu nadzemní biomasy (AGB) se mezi studiemi výrazně liší, protože se mění vstupní data, počet a kvalita pozemních měření, predikční algoritmy, charakter zájmové oblasti i měřítko zpracování. Variabilita v rozsahu terénních dat je v literatuře značná: některé studie pracovaly pouze s 8 validačními měřeními (R^2 = 0,58; Clerici a kol., 2016), zatímco jiné využily 24 měření (R^2 = 0,75; Macedo a kol., 2018), 45 měření (R^2 = 0,84; Nuthammachot a kol., 2020), 70 měření (R^2 = 0,57; Yu a kol., 2022), 80 měření (R^2 = 0,87; Moradi a kol., 2022) nebo až 664 měření (R^2 = 0,41; Gao a kol., 2018). Přestože počet pozemních měření (trénovacích a validačních) představuje důležitý předpoklad spolehlivého odhadu, výsledná přesnost je často významněji určena volbou predikčních proměnných (vstupních dat), zvoleným modelovacím algoritmem, charakterem a heterogenitou zájmové oblasti, měřítkem zpracování a typem vegetačního/*land cover* prostředí. Z hlediska algoritmů se v literatuře opakovaně ukazuje, že metody strojového učení (zejména Random Forest a XGBoost) dosahují vyšší predikční výkonnosti než klasické lineární či vícenásobné lineární regrese (Anwar a kol., 2018; Wang a kol., 2019; Guerini Filho a kol., 2020; Yu a kol., 2022), zatímco logaritmické a polynomické regrese vykazují zpravidla pouze mírné zlepšení (Anwar a kol., 2018; Wang a kol., 2019). V přímém srovnání dosáhly v jedné ze studií algoritmy XGBoost a Random Forest podobných výsledků, a to R^2 = 0,87 a RMSE 112,43 g·m^{-2} pro XGBoost, R^2 = 0,84 a RMSE = 126,57 g·m^{-2} pro Random Forest (Li a kol., 2021). Uplatnění nacházejí také neuronové sítě, zejména při práci s větším počtem vstupů a rozsáhlými datovými soubory (Gao a kol., 2018; Moradi a kol., 2022). Tento trend potvrzují i další práce. Tian a kol. (2024) ukázali, že model CNN-LSTM při využití vícezdrojových dat dosáhl hodnoty R^2 = 0,74 a RMSE = 26,43 Mg·ha^{-1}, zatímco při využití pouze optických dat dosáhl R^2 = 0,73 a RMSE = 26,92 Mg·ha^{-1}; samotná SAR data vedla k nižší přesnosti (R^2 = 0,59; RMSE = 30,04 Mg·ha^{-1}). Autoři zároveň konstatují, že kombinace SAR a optických dat obecně zlepšuje přesnost odhadu oproti použití jednotlivých datových zdrojů samostatně. Obdobně Rajput (2025) při kombinaci *in situ* a družicových dat porovnával exponenciální regresi, lineární regresi, SVM a Random Forest; nejlepší výsledky vykázal Random Forest (RMSE = 1,61; MAE = 0,84; relRMSE = 0,10; R^2 = 0,51),

zatímco SVM dosáhl pouze R^2 = 0,25 a lineární regrese R^2 = −0,16. Dalším příkladem jsou hybridní hluboké modely. Jiang a kol. (2025) navrhli model SCE-BiLSTM, který integruje prostorovou a kanálovou pozornost a komponentu extreme learning machine; při využití 11 prediktorů dálkového průzkumu a dat GEDI L4A dosáhl model v hlavní oblasti R^2 = 0,91, snížil MAE o 3,59 Mg·ha^{-1} a RMSE o 6,46 Mg·ha^{-1} oproti srovnávacímu CNN-BiLSTM modelu a při validaci v oblasti Žluté řeky dosáhl R^2 = 0,83.

Významným zdrojem variability výsledků je rovněž kombinace datových zdrojů. Řada studií využívá fúzi radarových a multispektrálních dat, typicky Sentinel-1 a Sentinel-2, s cílem spojit spektrální informaci o stavu vegetace se strukturální citlivostí radarového signálu (Hoscilo a kol., 2018; Laurin a kol., 2018; Wang a kol., 2019; Nuthammachot a kol., 2020; Bucha a kol., 2021). Současně však bývá uváděno, že samotná radarová data obvykle poskytují výrazně nižší přesnost než optická data, např. při srovnání Sentinel-1 vs. Sentinel-2 bylo dosaženo hodnot R^2 0,34 a 0,82 (Nuthammachot a kol., 2020), a proto radar často vystupuje spíše jako doplňkový zdroj informace (Abrishamkar a Ahmadi, 2017). Současně se ukazuje, že další zlepšení přesnosti přináší zejména doplnění optických a radarových dat o výškovou informaci z lidarových systémů a o topografické proměnné. Guerra-Hernández a kol. (2022) pro středomořské lesy ukázali, že kombinace ICESat-2, Sentinel-1, Sentinel-2, ALOS-2/PALSAR-2 a topografických dat umožňuje přesnější extrapolaci biomasy do plošných map. Podobně Schwartz a kol. (2023) ukazují, že kombinace Sentinel-1, Sentinel-2 a GEDI může sloužit nejen k mapování výšky porostu, ale i k následnému odvození objemu a AGB. Při validaci vůči nezávislým datům francouzské národní inventarizace dosáhla mapa výšky porostu chyby MAE 2,94 m a R^2 = 0,69.

Klíčovou roli hraje také validační strategie. V prostorově autokorelovaných datech může standardní náhodná křížová validace nadhodnocovat generalizační schopnost modelu; naopak prostorová (bloková) validace zpravidla poskytuje realističtější odhad přenositelnosti výsledků (Le Rest a kol., 2014; Roberts a kol., 2017). Ploton a kol. (2020) například ukázali výrazný rozdíl mezi klasickou 10násobnou křížovou validací (R^2 = 0,53) a prostorovou 44násobnou validací (R^2 = 0,14). Neméně důležité je prostorové rozlišení vstupních dat: vyšší rozlišení umožňuje lépe zachytit variabilitu na úrovni jedinců či menších částí porostu, zatímco nižší rozlišení zvyšuje podíl smíšených pixelů a může zhoršovat přesnost zejména v heterogenním prostředí. Přesto lze při vhodné volbě metody a homogenním charakteru území dosahovat dobrých výsledků i se středním rozlišením (např. R^2 = 0,71 u TERRA-MODIS) (Wang a kol., 2018), zatímco u komplexnějších území a více typů habitatů se přesnost může významně měnit s měřítkem a datovým rozlišením (Rodríguez-Veiga a kol., 2019; Chang a Shoshany, 2016; Anwar a kol., 2018; Li a kol., 2021; Moradi a kol., 2022). V současných studiích navíc validační data často netvoří pouze terénní plochy, ale také data národních inventarizací, segmenty odvozené z ALS nebo footprinty družicových lidarových misí. Schwartz a kol. (2023) uvádějí, že pro převod výšky na AGBD (Above-ground Biomass Density) odvodili mocninné alometrické rovnice z dat francouzské národní inventarizace, přičemž francouzská NFI (National Forest Inventory) každoročně měří přibližně 6000 nových ploch. Při nezávislé validaci vůči NFI dosáhla jejich mapa výšky porostu MAE = 2,94 m a R^2 = 0,69; odhad AGBD odvozený z výšky pak při porovnání s NFI AGBD vykázal MAE = 61,7 Mg·ha^{-1} a R^2 = 0,4. Na úrovni lokálních porostních studií pak může vhodná kombinace

UAV RGB snímků, lidarových dat a terénních měření zlepšit odhad biomasy oproti klasickým regresním přístupům. Maesano a kol. (2022) v porostu borovice v jižní Itálii uvádějí při použití Random Forest zlepšení z R^2 = 0,81 na 0,86 a současně snížení RMSE z 45,5 na 31,7 Mg·ha^{-1} a MAE z 34,2 na 22,1 Mg·ha^{-1} ve srovnání s krokovou regresí.

Následující tabulka (Tabulka 6) shrnuje vybrané studie zaměřené na odhad zásoby biomasy a souvisejících ukazatelů z dat dálkového průzkumu Země, včetně použitých družicových dat, typických vstupních proměnných a modelovacích přístupů.

Tabulka 6 Přehled metod ke kvantifikaci nebo identifikaci ekosystémové funkce

Družice	Zdroj	Podrobnosti	Výpočet (pokud je dostupný)
GeoEye-1, Pleiades-1A	Clerici a kol., 2016	Odhad kmenové biomasy v Kolumbii (tropické klima). Nejlepší model pro odhad biomasy s použitím indexu RVI.	$Log\ AGB\ [t{\cdot}ha^{-1}] = -3{,}208 \times RVI + 2{,}185$
Landsat 5,7; ALOS PALSAR; LIDAR	Rodríguez-Veiga a kol., 2019	Hodnocení metod používaných při odhadu lesní biomasy v Polsku, Švédsku, Mexiku, Africe a Indonésii.	
TERRA – MODIS	Geng a kol., 2021	Odhad kukuřičné biomasy v Číně (mírné klima). Nejlepším algoritmem pro odhad biomasy byl XGBoost (R^2 = 0,78), následoval RF (R^2 = 0,77). SVM a ANN byly méně přesné.	
Sentinel-1, Sentinel-2	Hoscilo a kol., 2018	Modelování nadzemní lesní biomasy pomocí RF v Polsku (mírné klima). Model podhodnocuje větší hodnoty biomasy (větší než 250 t·ha^{-2}) a nadhodnocuje nižší hodnoty (menší než 100 t·ha^{-2}).	
Landsat, SPOT, RapidEye	Kross a kol., 2015	Odhad biomasy z východní části Ontaria v Kanadě (mírné klima). Biomasa kukuřice nejlépe korelovala s indexem SRe a biomasa sóji s MTVI2.	$AGB_{maize}[g{\cdot}m^{-1}] = 8{,}247 \times SRe - 360{,}98$ $AGB_{soybean}[g{\cdot}m^{-1}] = 17{,}26 \times MTVI2 - 82{,}339$
Sentinel-2	Guerini Filho a kol., 2020	Odhad luční biomasy v Brazílii (subtropický pás). Nejlepší výsledky s indexy NDREI a EVI.	$AGB[g{\cdot}m^{-1}] = 6{,}14 + (0{,}86 \times EVI) + (3{,}94 \times NDREI)$
TERRA – MODIS	Wang a kol., 2018	Odhad luční biomasy v Číně (semiaridní kontinentální klima). Nejlepší výsledky s použitím indexu MNDVI.	$AGB[g{\cdot}m^{-1}] = 51{,}747 \times \ln(MNDVI) + 155{,}87$
Sentinel-2	Li a kol., 2021	Odhad luční biomasy v Číně (subtropické klima). Modely XGBoost prokazovaly mírně lepší výsledky modelování než RF. Vstupy modelu: pásma Sentinel-2, vegetační indexy, GLCM (sousednost 7×7).	
Sentinel-2, Landsat	Puliti a kol., 2021	Srovnání mezi Landsat a Sentinel-2 při použití snímků z dvou různých období (2015, 2019) a pouze jednoho období (2019). Odhad lesní biomasy pro boreální	$AGB[t{\cdot}ha^{-1}] = -185{,}93 - (485{,}72 \times B4/B7) + (301 \times B4/B11)$

		les (88 % smrku ztepilého) v Norsku (mírné klima). Nejlepší výsledky byly dosaženy s použitím dat Sentinel-2.	
Sentinel-2	Lovynska a kol., 2020	Modelování nadzemní biomasy jehličnatého lesa na Ukrajině (mírné klima). Nejlepší modely s použitím vegetačních indexů NDVI a TVI nebo v druhém případě NDVI a FCOVER.	$AGB[t \cdot ha^{-1}] = 716{,}261 \times NDVI \times TVI^{-1{,}644}$
QuickBird	Macedo a kol., 2018	Odhad lesní biomasy pro dub okrouhlolistý v Portugalsku (středomořské klima). Nejlepší modely s použitím NDVI nebo SR.	$AGB[t \cdot ha^{-1}] = -37{,}026 + 223{,}808 \times NDVI$ $AGB[t \cdot ha^{-1}] = -47{,}108 + 42{,}582 \times SR$
Sentinel-2, Landsat 8	Anwar a kol., 2018	Modelování lesní biomasy v Pákistánu (vlhké subtropické klima). Nejlepší model s použitím NDVI.	$AGB[t \cdot ha^{-1}] = (2099{,}9 \times NDVI^2 - 1635{,}6 \times NDVI) + 331{,}04$
Sentinel-2, Sentinel-1	Nuthammachot a kol., 2020	Modelování lesní biomasy v Indonésii (tropické klima). Nejlepší model s použitím polarizace VH (Sentinel-1), NDI45 a pásma B6 (Sentinel-2).	$AGB[t \cdot ha^{-1}] = -418{,}95 + (542{,}23 \times NDI45) + (248{,}48 \times B6) - (6{,}56 \times VH)$
Sentinel-2, Landsat 8, Gaofen-2	Yu a kol., 2022	Modelování lesní biomasy v Číně (mírné klima). Nejlepší přesnost byla modelována z dat Gaofen-2 algoritmem RF (59,43 %). Sentinel-2 měl mírně vyšší přesnost než Landsat 8. Algoritmus RF dosahoval lepších výsledků, než vícenásobné lineární regrese bez ohledu na použitá družicová data.	
Landsat, ALOS PALSAR	Gao a kol., 2018	Modelování lesní biomasy v Číně (subtropické klima). Landsat dosahoval přesnějších odhadů biomasy než ALOS PALSAR. Nejlepším algoritmem byl ANN, zatímco RF a kNN byly méně přesné.	
Sentinel-2	Moradi a kol., 2022	Modelování lesní biomasy v západním Íránu (subtropické klima). Nejlepším algoritmem pro předpověď biomasy byla umělá neuronová síť MLPNN (Multi-Layer Perceptron Artificial Neural Network) s R^2 0,87. Vegetační indexy NDVI a IPVI měly nejvyšší korelaci s biomasou.	
Sentinel-2, Landsat 8	Berger a kol., 2020	Modelování luční biomasy v Krugerově národním parku v Jihoafrické republice (semiaridní subtropické klima). Nejlepším prediktorem pro modelování biomasy byl vegetační index NDVI s R^2 0,74, který měl lepší výsledky, než samotná pásma Sentinel-2 nebo Landsat 8.	
Sentinel-1, Sentinel-2	Chang a Shoshany, 2016	Odhad biomasy pro středomořské křoviny (subtropické klima). Sentinel-2 dosahoval vyšší přesnosti (R^2 0,72), než Sentinel-1 (méně než 0,6). Fúzní model zvýšil přesnost na 0,86. Sentinel-2 vykazuje vyšší přesnost	$AGB[\text{kg} \cdot \text{m}^{-1}] = (0{,}148 + 1{,}735 \times \text{NDVI}) \times \text{sqrt}\,(1 + (2{,}5\sigma\text{HV} - 1{,}2) \times \text{FVC}/0{,}5)$

		odhadu biomasy než Sentinel-1, nicméně Sentinel-1 má potenciál indikovat strukturu vegetace.	
SPOT-5	Muhd-Ekhzarizal a kol., 2018	Odhad biomasy mangrovových lesů v Malajsii (tropické klima). Nejlepší model byl vytvořen s použitím indexů NDVI a GEMI-NDVI.	$AGB[t \cdot ha^{-1}] = 793{,}676 \times \text{NDVI} / 574{,}770 \times \text{GEMI-NDVI-}574{,}219$
Sentinel-2	Chen a kol., 2018	Modelování lesní biomasy pomocí pásem Sentinel-2 a vegetačních indexů (subtropické monzunové klima).	
Sentinel-1, Sentinel-2, GEDI	Schwartz a kol., 2023	Francie. Produkt FORMS pro mapování výšky porostu, objemu a AGB v rozlišení 10–30 m. Využita vícezdrojová fúze dat a hluboké učení; validace vůči ALS a inventarizačním datům.	
ICESat-2, Sentinel-1, Sentinel-2, ALOS-2/PALSAR-2, topografie	Guerra-Hernández a kol., 2022	Středomořské lesy. Integrace lidarových, radarových, optických a topografických dat pro plošné mapování AGB. Studie ukazuje přínos vícezdrojového přístupu při extrapolaci biomasy.	
UAV RGB, LiDAR	Maesano a kol., 2022	Borový porost v jižní Itálii. Odhad AGB z UAV RGB snímků a LiDAR dat; Random Forest zlepšil výsledky oproti krokové regresi.	
Sentinel-1, Sentinel-2, ALOS-2, GaoFen-3, GaoFen-6, Landsat-8	Tian a kol., 2024	Čína. Odhad lesní AGB z vícezdrojových radarových a optických dat s využitím RF, CNN a CNN-LSTM. Kombinace dat zvýšila přesnost oproti jednosenzorovým vstupům.	
GEDI + DPZ prediktory	Jiang a kol., 2025	Regionální odhad lesní biomasy pomocí hybridního modelu SCE-BiLSTM. Studie dokládá využití hlubokého učení a attention mechanismů při modelování AGB.	
In situ + DPZ data	Rajput, 2025	Indie. Srovnání lineární regrese, exponenciální regrese, SVM a Random Forest při modelování lesní biomasy; nejlepší výsledky vykázal Random Forest.	

3. METODY A DATA

3.1. Zájmové území

Řešeným územím je celá oblast České republiky, která má rozlohu 78 871 km^2. Z celkové rozlohy území leží 52 817 km2 (66,97 %) v nadmořských výškách do 500 metrů, 25 222 km2 (31,98 %) v nadmořských výškách od 500 do 1 000 metrů a 827 km2 (1,05 %) nad 1 000 metrů (Český statistický úřad, 2023). Z geomorfologického hlediska je Česká republika rozdělena na dva systémy. Západní a střední části vyplňuje Česká vysočina, zatímco menší území ve východní části tvoří Západní Karpaty. Prostorová diferenciace nadmořské výšky, reliéfu a s ní souvisejících stanovištních podmínek se výrazně promítá do rozmístění lesních ekosystémů a jejich druhové skladby. Představuje proto jeden ze základních faktorů ovlivňujících regionální variabilitu nadzemní biomasy a zásoby uhlíku v lesích České republiky.

Lesnatost České republiky nyní činí 29 232 km^2 (37,1 %) (Kučera a Adolt, 2019). Jde o nárůst přibližně o 0,3 % celkové rozlohy ve srovnání s předchozími měřeními Národní inventarizace lesů. Podle třetího cyklu národní inventarizace zůstala rozloha lesů nezměněna. Výsledky třetího cyklu Národní inventarizace lesů (NIL3 2016–2020) vycházejí ze zjištění skutečného stavu na 39 407 inventarizačních plochách, což poskytuje robustní referenční rámec pro charakteristiku lesů na celostátní úrovni (Máslo a kol., 2023). Dominantní podíl na celkové ploše lesa má porostní půda (97,5 ± 0,3 %), přičemž bezlesí tvoří pouze 0,4 ± 0,1 %. Z hlediska nadmořské výšky se 25,6 ± 0,7 % lesa nachází do 400 metrů, 57,8 ± 0,8 % v nadmořské výšce od 400 do 700 metrů a 16,6 ± 0,6 % se nachází v polohách nad 700 metrů nad mořem. Pokud jsou porovnány přírodní lesní oblasti (PLO), nejnižší lesnatost je v Hornomoravském úvalu (PLO 34) a to 7,5 ± 1,7 %, zatímco nejvyšší lesnatost je v Novohradských horách (PLO 14) 82,4 %. Vysoká lesnatost Novohradských hor je způsobena zejména malým územím PLO, kde plocha lesa zabírá 11 200 hektarů. Vysoká lesnatost, přesahující 78 %, byla zaznamenána taky v Krkonoších (PLO 22), Hrubém Jeseníku (PLO 27) a Moravskoslezských Beskydech (PLO 40). Podle krajů ČR byla největší lesnatost zjištěna v Karlovarském kraji (50,2 %) následována Libereckým krajem. Tyto rozdíly dokládají prostorovou heterogenitu českých lesů.

V ČR jehličnaté lesy (57,6 %) převažují nad listnatými lesy (42,4 %). Z hlediska druhové skladby je nejvíce zastoupeným jehličnanem smrk ztepilý, následovaný borovicí lesní, zatímco nejvíce zastoupenou listnatou dřevinou je buk lesní, následovaný duby. Národní inventarizace lesů ukazuje, že lesy České republiky představují významný rezervoár nadzemní biomasy a uhlíku. Celková zásoba uhlíku vázaného v nadzemní biomase byla pro období NIL3 odhadnuta na 314,3 ± 5,5 milionů tun C, což po přepočtu odpovídá 110,3 ± 1,3 t C·ha^{-1} porostní půdy. Odhad celkové hmotnosti sušiny nadzemní biomasy dosahuje přibližně 625,1 mil. t, při zahrnutí stojících souší nad 10 cm by činil 625,6 mil. t. (Máslo a kol., 2024). Z hlediska prostorové variability byla nejvyšší hektarová zásoba uhlíku v nadzemní biomase zjištěna ve Zlínském kraji, což potvrzuje, že rozložení těchto veličin není na území České republiky homogenní.

3.2. Vstupní data a jejich zpracování

Primární datovou základnu této studie tvořila volně dostupná data programu Copernicus, konkrétně mise Sentinel-2 a Sentinel-1. Sentinel-2 poskytoval optickou multispektrální informaci o vegetačním pokryvu, zatímco Sentinel-1 poskytoval radarovou informaci citlivou na strukturu porostu a jeho fyzikální vlastnosti. Oba typy dat se z hlediska modelování nadzemní biomasy vhodně doplňují, protože optická data lépe vystihují spektrální vlastnosti vegetace, zatímco radarová data umožňují částečně zachytit strukturní rozdíly v porostu. Kromě družicových dat Sentinel byla jako doplňkový prediktor testována také existující vrstva výšky lesního porostu. Pro kalibraci a validaci modelů byla použita terénní referenční data nadzemní biomasy pocházející ze sítě DendroNetwork a z měření nízkobiomasových travních porostů. Přehled základních charakteristik hlavních družicových vstupů je uveden v tabulce 7.

Tabulka 7 Základní parametry použitých senzorů Sentinel-2 MSI a Sentinel-1 C-SAR

	Sentinel-2 MSI	**Sentinel-1 C-SAR**
Typ senzoru	Pasivní multispektrální optický senzor	Aktivní jednofrekvenční radar
Spektrální oblast/frekvence	13 spektrálních pásem (viditelné – blízké infračervené spektrum)	5,405 GHz (vlnová délka cca 5,55 cm)
Prostorové rozlišení	10, 20, 60 m podle pásma	Závislé na režimu snímání (5 – 40 m)
Šířka záběru	290 km	250 km (IW)
Časové rozlišení	10 dní pro jednu družici (5 dní pro dvoudružicovou konfiguraci)	12 dní pro jednu družici (6 dní pro dvoudružicovou konfiguraci)
Využití	Albedo, vegetační a vlhkostní indexy, informace o stavu vegetace	Citlivost na strukturu a vlhkost vegetace, nezávislost na oblačnosti
Omezení	Citlivost na oblačnost, aerosol a světelné podmínky	Přítomnost šumu (*speckle*), geometrické zkreslení

3.2.1. Sentinel-2

Mise Sentinel-2 představuje optickou složku programu Copernicus a je určena k systematickému monitorování zemského povrchu ve vysokém prostorovém a časovém rozlišení. V současné době je tvořena konfigurací dvou družic S2B a S2C, které obíhají ve stejné polární dráze se vzájemným fázovým posunem 180°. Družice S2A, jejíž úlohu měla S2C postupně převzít, byla na základě požadavků uživatelské komunity dočasně ponechána v provozu. V období od ledna do března 2025 byla manévrována do nové doplňkové orbitální pozice, vzdálené o 36° od S2B, odkud dále rozšiřuje pozorovací kapacitu mise. Tato tříčlenná konfigurace zvyšuje frekvenci snímání a zlepšuje dostupnost dat (Copernicus, 2025b).

Družice Sentinel-2 jsou vybaveny multispektrálním senzorem MSI (Multi-Spectral Instrument), který zaznamenává data ve 13 spektrálních pásmech s prostorovým rozlišením 10, 20 a 60 m. Pro analýzy byla využita data Sentinel-2 úrovně zpracování L2A, která představují povrchovou

odrazivost. Zpracování bylo provedeno v cloudové platformě Google Earth Engine (Gorelick a kol., 2017). Satelitní scény byly nejprve prostorově ořezány podle hranic České republiky, aby byla snížena výpočetní zátěž a velikost exportovaných dat. Při zpracování byly vyloučeny snímky s celkovou oblačností přesahující 40 %, aby se omezil vliv výrazně kontaminovaných scén. Maskování oblačnosti bylo provedeno pomocí algoritmu *s2cloudless*, přičemž byla použita prahová hodnota pravděpodobnosti výskytu oblačnosti 40 %. Tato hodnota představuje kompromis mezi spolehlivým odstraněním oblačnosti a zachováním dostatečného množství vstupních dat pro tvorbu mozaiky. Stíny oblaků byly identifikovány na základě prahové hodnoty tmavých pixelů v blízké infračervené oblasti 0,15. Masky oblačnosti i stínů byly následně rozšířeny o *buffer* 50 m, aby bylo zajištěno odstranění kontaminovaných pixelů a současně omezen vliv okrajových efektů.

Zvolené prahové hodnoty vycházejí ze zavedených postupů zpracování v prostředí Google Earth Engine a z předchozích velkoplošných studií využívajících data Sentinel-2, v nichž je z důvodu zachování prostorového pokrytí preferováno konzervativní maskování oblačnosti před příliš restriktivní filtrací. S cílem minimalizovat fenologickou variabilitu a zajistit srovnatelnost mezi jednotlivými lesními typy byly do analýzy zahrnuty pouze snímky pořízené v období vrcholné části vegetační sezóny. Časové období bylo zvoleno tak, aby byla zachycena doba, kdy je koruna stromů nejvíce vyvinutá, a zároveň aby výsledky nebyly ovlivněny jarním rašením ani pozdně sezónním stárnutím listů. Původně byl konec časového okna nastaven na 10.září, nicméně z hlediska chybějících scén bylo období prodlouženo do 20. září. Do výsledného sezónního mediánového kompozitu vstoupilo celkem 415 scén Sentinel-2. Exportovaný datový soubor Sentinel-2 zahrnoval následující parametry:

- Spektrální pásma: B11, B8, B7, B6, B5, B4, B3, B2;
- Prostorové rozlišení: 20 metrů;
- Časové okno: 1. června až 20. září 2023;
- Souřadnicový systém: EPSG:32633.

3.2.2. Sentinel-1

Mise Sentinel-1 představuje radarovou složku programu Copernicus a je určena k systematickému snímání zemského povrchu bez ohledu na denní dobu a oblačnost. Je založena na konstelaci družic nesoucích radar se syntetickou aperturou v pásmu C (C-SAR), které obíhají po téměř polární heliosynchronní dráze. Konfigurace dvou družic pracuje stejně jako Sentinel-2 ve stejné orbitální rovině se vzájemným fázovým posunem 180°. Původní družice Sentinel-1B ukončila misi v roce 2022 v důsledku technické poruchy, zatímco družice Sentinel-1C byla vypuštěna 5. prosince 2024 a postupně obnovila plnohodnotnou dvou družicovou konfiguraci mise společně se Sentinel-1A (Copernicus, 2025a). Jedna družice má opakovací cyklus 12 dní, zatímco konstelace dvou družic zajišťuje přesný šestidenní opakovací cyklus.

Sentinel-1 poskytuje radarová data v několika zobrazovacích režimech s prostorovým rozlišením až kolem 5 m a se záběrem až 400 km, přičemž podporuje i duální polarizaci. Radarový signál je vysílán v jedné polarizaci, horizontální (H) nebo vertikální (V), a přijímán může být v jedné nebo

ve dvou polarizačních rovinách podle konkrétního akvizičního režimu. Významnou předností mise je schopnost získávat konzistentní časové řady dat pro monitoring pevniny, mořského prostředí, ledu, sněhu, půdní vlhkosti i vegetace. Pro odvození prediktorů radarového zpětného rozptylu byla použita data Sentinel-1 GRD (Ground Range Detected) úrovně Level-1. Zpracování bylo provedeno v prostředí Google Earth Engine, kde byla využita kolekce `COPERNICUS/S1_GRD`. Na tuto kolekci je aplikován standardní předzpracovatelský postup zahrnující aplikaci orbitálních souborů, odstranění okrajového a termálního šumu, radiometrickou kalibraci, terénní korekci a převod hodnot zpětného rozptylu σ^0 do jednotek dB. Dodatečné filtrování šumu typu *speckle* nebylo provedeno, protože jeho vliv byl redukován prostřednictvím časového průměrování při kompozici většího počtu scén. Pro obě polarizace, VV a VH, byly samostatně vytvořeny kompozity průměrného zpětného rozptylu pro vzestupné a sestupné dráhy. Následně byly tyto orbitálně specifické kompozity zprůměrovány s cílem získat jednu reprezentativní vrstvu zpětného rozptylu pro každou polarizaci a současně omezit směrové efekty související s geometrií snímání. Pro vytvoření výsledné radarové mozaiky bylo celkem využito 207 scén Sentinel-1. Exportovaný datový soubor Sentinel-1 zahrnoval následující parametry:

- Polarizace: VV, VH;
- Orbitální uzel: vzestupný, sestupný;
- Prostorové rozlišení: 20 metrů;
- Časové okno: 1. června až 10. září 2023;
- Souřadnicový systém: EPSG:32633.

3.2.3. Vegetační indexy

Za účelem zvýraznění spektrálních charakteristik byly využity vegetační indexy zdůrazňující strukturu vegetace nebo obsah biomasy. Vegetační indexy, jako jsou NDVI (Normalized Difference Vegetation Index) a ReNDVI (Red-Edge NDVI), se běžně používají pro mapování nebo odhad nadzemní biomasy lesů (Vaglio Laurin a kol., 2016; Koerner a kol., 2022; Wai a kol., 2022). V této studii byl však index NDVI nahrazen indexem ARVI (Atmospherically Resistant Vegetation Index), který je čtyřikrát méně citlivý na atmosférické vlivy než NDVI. Pro ARVI je definován koeficient γ, přičemž při $\gamma = 0$ je ARVI totožný s NDVI. Hodnota $\gamma = 1$ je pro většinu aplikací považována za nejvhodnější (Kaufman a Tanré, 1992). Dalším použitým indexem byl ReNDVI, který je známý svou vyšší citlivostí na změny vegetace v oblastech charakterizovaných hustou vegetací a vysokými hodnotami biomasy. Index MSAVI (Modified Soil Adjusted Vegetation Index) omezuje vliv půdního pozadí a je účinný v oblastech s řídkou vegetací. Indexy GNDVI (Green Normalized Difference Vegetation Index) a NDI45 (Normalized Difference Index 45) jsou citlivější na obsah chlorofylu než NDVI. Celkem tak bylo použito pět vegetačních indexů (Tabulka 8).

Tabulka 8 Použité vegetační indexy

Vegetační index	Vzorec	Zdroj	
ARVI	$ARVI = \frac{B8 - B4 - \gamma(B4 - B2)}{B8 + B4 - \gamma(B4 - B2)}$	Kaufman and Tanré, 1992	(1)
GNDVI	$GNDVI = \frac{B8 - B3}{B8 + B3}$	Gitelson and Merzlyak, 1994	(3)
ReNDVI	$ReNDVI = \frac{B8 - B5}{B8 + B5}$	Gitelson and Merzlyak, 1994	(2)
MSAVI	$MSAVI = \frac{2 \times B8 + 1 - \sqrt{(2 \times B8 + 1)^2 - 8 \times (B8 - B4)}}{2}$	Qi a kol., 1994	(4)
NDI45	$NDI45 = \frac{B5 - B4}{B5 + B4}$	Delegido a kol., 2011	(5)

3.2.4. Texturní metriky GLCM

Texturní vlastnosti optického snímku Sentinel-2 byly vypočteny pomocí nástroje *Texture Analysis – GLCM* v softwaru ESA SNAP. Získání texturních charakteristik z obrazu umožňuje vytvořit doplňující obrazové vrstvy popsané různými statistickými vztahy mezi pixely. V rámci tohoto nástroje je nejprve nutné definovat velikost pohyblivého okna, které vymezuje oblast pro výpočet texturní analýzy. Následně se nastavují parametry, jako je posun (displacement), vyjadřující vzdálenost od referenčního pixelu pro porovnání se sousedními pixely, a úhel (angle), určující směr tohoto porovnání. V této práci byly použity parametry: pohyblivé okno o velikosti 13 × 13 pixelů, posun nastavený na hodnotu 2 a úhel specifikovaný pro všechny směry. Program SNAP využívá pro výpočet texturních charakteristik matici GLCM (Grey-Level Co-Occurrence Matrix), navrženou Haralickem a kol. (1973). Mezi vypočítávané texturní charakteristiky patří: Contrast (CON), Dissimilarity (DIS), Homogeneity (HOM), Angular Second Moment (ASM), Maximum Probability (MAX), Entropy (ENT), GLCM Mean, GLCM Variance a GLCM Correlation. Při využití osmi spektrálních pásem a devíti texturních charakteristik by bylo do modelování zahrnuto celkem 72 texturních rastrů. Vzhledem k silným korelacím mezi některými charakteristikami, ať již kladným či záporným, se však doporučuje vybrat pouze ty relevantní (Hall-Beyer, 2017). Počet charakteristik byl proto redukován na tři. První z nich představuje GLCM Mean, který je klíčový pro charakterizaci texturních vlastností vnitřních částí objektů. Druhou vybranou charakteristikou byl Contrast, použitý k odlišení hran, a třetí Entropy, která tyto informace doplňuje (Tabulka 9).

Tabulka 9 Použité texturní metriky GLCM

Texturní vlastnost	Vzorec	
Contrast (CON)	$CON = \sum_{i,j=0}^{N-1} iP_{ij}(i-j)^2$	(6)
Entropy (ENT)	$ENT = \sum_{i,j=0}^{N-1} iP_{ij}(-ln\ P_{ij})$	(7)
GLCM Mean (MEA)	$MEA = \sum_{i,j}^{N-1} iP_{ij}$	(8)

3.2.5. Morfometrické charakteristiky

Jako zdroj morfometrických charakteristik byl použit ZABAGED ® – Výškopis – DMR 4G. Jedná se o digitální model reliéfu České republiky 4. generace v souřadnicovém systému S-JTSK. Model je spravován Zeměměřičským ústavem a je pravidelně aktualizován v návaznosti na DMR 5G. Představuje zobrazení zemského povrchu ve formě rastrového formátu TIFF nebo pravidelné síti bodů. V disertační práci bylo využito rastrového formátu TIFF s rozlišením 5 × 5 metrů na pixel. Výškový souřadnicový systém produktu je Balt po vyrovnání (Bvp), přičemž střední chyba výšky je uváděna jako 0,3 m v odkrytém terénu a 1 m v zalesněném terénu. Model byl vytvořen v letech 2009 až 2013 a v roce 2016 byl nově vygenerován z dat DMR 5G. Nyní je aktualizován v návaznosti na aktualizaci DMR 5G. V disertační práci byla použita verze dostupná v roce 2023.

Z DMR 4G vznikly tři topografické prediktory: nadmořská výška, sklon svahu a orientace svahu. Topografické proměnné byly zařazeny do testování proto, že reliéf může ovlivňovat stanovištní podmínky. Jejich využití při modelování nadzemní biomasy je běžné i v předchozích studiích. V práci Zhou a kol. (2025) bylo uvedeno, že topografické faktory ovlivňují světelné, vlhkostní a půdní podmínky, a jsou proto významné pro predikci lesní nadzemní biomasy. Pro nadmořskou výšku byl v této studii použit digitální model reliéfu SRTM (Shuttle Radar Topography Mission), ze kterého byly dále odvozeny sklon a orientace svahu. V evropském prostředí byla zásoba nadzemní biomasy modelována například v práci Puliti a kol. (2020), kde byla využita data národní inventarizace, ArcticDEM a Sentinel-2. V této studii bylo ukázáno, že proměnné odvozené z ArcticDEM byly pro model důležité a v daném případě měly dokonce vyšší význam než proměnné Sentinel-2.

Před vstupem do modelování byly topografické vrstvy prostorově harmonizovány s ostatními prediktory. Byly převedeny do jednotného souřadnicového systému, zarovnány na společný rastr, převzorkovány na prostorové rozlišení 20 m a oříznuty podle použité lesní masky. Topografické proměnné nebyly součástí základní sady Sentinel-1 + Sentinel-2, ale byly testovány v rozšířené sadě S1+S2+DEM, aby bylo možné ověřit, zda doplnění informace o reliéfu zlepší predikci nadzemní biomasy oproti samotným radarovým a optickým prediktorům.

3.2.6. Výška lesního porostu

Kromě družicových snímků byla využita již existující datová sada výšky lesního porostu, konkrétně globální data výšky lesního porostu z roku 2020 s prostorovým rozlišením 1 m (Tolan a kol., 2024). Na rozdíl od datové sady Lang a kol. (2023), která vycházela z dat Sentinel-2 a GEDI, přičemž GEDI bylo využito pro validaci, zvolili Tolan a kol. (2024) odlišný přístup založený na přímém modelování vztahu mezi daty GEDI a snímky Maxar, což umožnilo mapování výšky stromů ve velmi vysokém prostorovém rozlišení. GEDI využívá technologii lidaru se senzory umístěnými na Mezinárodní vesmírné stanici (ISS). Jedná se o nejpřesnější lidarový systém nasazený na oběžné dráze pro monitoring vegetace. Přístroj GEDI je tvořen třemi lasery, které vytvářejí celkem osm paralelních pozemních stop. Jednotlivé laserové stopy mají průměr přibližně 25 m a podél dráhy jsou rozmístěny přibližně po 60 metrech.

Vrstva výšky lesního porostu byla do analýzy zařazena jako potenciální doplňkový prediktor nadzemní biomasy. Výška porostu může být s biomasou úzce spojena, protože zachycuje vertikální strukturu vegetace, kterou optická a radarová data Sentinel nemusí vždy dostatečně vystihnout. Současně však tato vrstva reprezentuje stav roku 2020, zatímco hlavní modelování AGB bylo vztaženo k vegetační sezóně roku 2023. Z tohoto důvodu byla výška porostu testována pouze v rámci sady prediktorů ALL a nebyla považována za základní vstup finálního modelu.

3.2.7. Maska lesního porostu

Vlastní prostorová predikce nadzemní biomasy byla provedena pouze v rámci masky lesních porostů. Lesní maska byla odvozena z vrstvy souborů lesních typů České republiky, která vychází z lesnicko-typologického klasifikačního systému používaného pro diferenciaci růstových a stanovištních podmínek lesů. V tomto systému představuje lesní typ základní jednotku diferenciace růstových podmínek, zatímco soubor lesních typů sdružuje ekologicky příbuzné lesní typy s obdobnými hospodářsky a stanovištně významnými vlastnostmi. Soubory lesních typů jsou v typologické síti vymezovány kombinací půdních kategorií a lesních vegetačních stupňů (Viewegh, 2003).

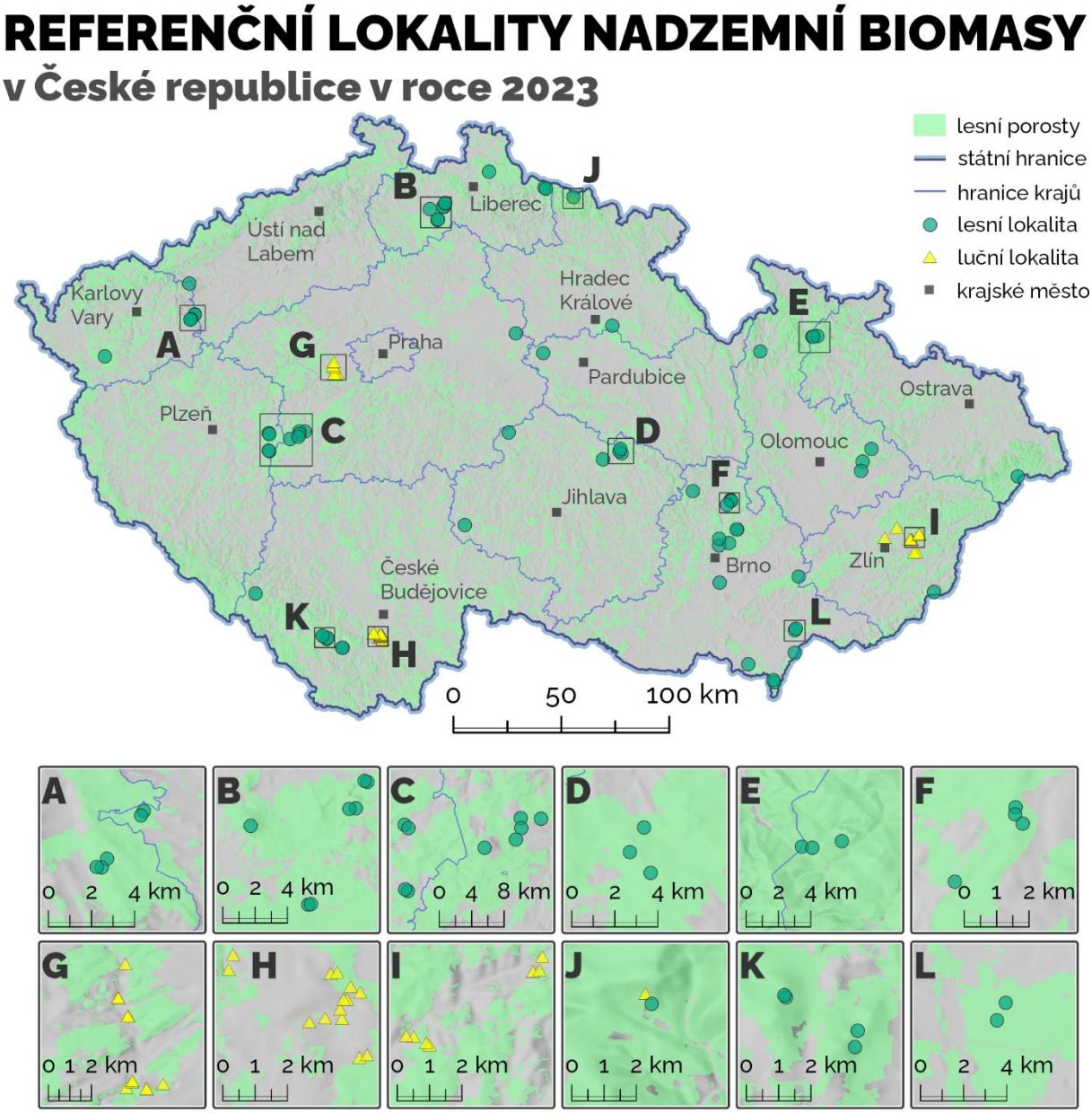
Výsledná modelovaná lesní vrstva měla rozlohu přibližně 2 717 tis. ha. Tato hodnota je nižší než odhad rozlohy lesa uváděný Národní inventarizací lesů. Podle výsledků NIL 3 dosahuje odhad plochy lesa na území České republiky 2 923,2 ± 37,6 tis. ha (Máslo a kol., 2023). Rozdíl mezi použitou lesní maskou a inventarizačním odhadem je nutné zohlednit zejména při interpretaci celkové zásoby nadzemní biomasy. Rozloha masky totiž přímo ovlivňuje celkový národní odhad AGB, i když průměrná hektarová hodnota může být velmi podobná výsledkům inventarizace.

Vedle vlastní lesní masky bylo v modelování ponecháno také několik maloplošných nelesních lokalit, na kterých bylo provedeno pozemní měření luční biomasy. Tyto plochy budou podrobněji popsány v samostatné metodické kapitole. V kontextu celkové rozlohy lesní masky šlo o plošně zcela zanedbatelnou plochu, která neměla za cíl rozšiřovat mapování AGB na nelesní území. Tyto lokality byly v datové sadě ponechány především z důvodu následného vizuálního a statistického srovnání výsledků práce a zároveň umožnily lépe zachytit velmi nízké hodnoty biomasy. Zahrnutí těchto nelesních ploch tedy nemění základní vymezení výsledné predikce jako predikce zaměřené

na lesní území. Jejich účelem bylo doplnit rozsah trénovacích a validačních hodnot v oblasti nízké biomasy a snížit riziko systematického nadhodnocování AGB v místech s řídkou vegetací, lesními holinami nebo porosty v raných fázích obnovy, které se mohou vyskytovat i uvnitř lesní masky.

3.2.8. Referenční data nadzemní biomasy

Referenční data pro modelování nadzemní biomasy byla tvořena terénními měřeními pocházejícími ze dvou zdrojů. Základní část souboru představovala lesní měření ze sítě DendroNetwork, která byla doplněna o měření biomasy travních porostů reprezentující velmi nízké hodnoty nadzemní biomasy. Cílem tohoto doplnění bylo přesnější zachycení spodní části rozsahu hodnot AGB. Hlavní část referenčního souboru tvořila data sítě DendroNetwork. Tato síť, založená v roce 2016, je určena k monitoringu lesních ekosystémů a ke sběru informací o dendrometrických charakteristikách porostů. Pro účely této práce bylo využito 71 lesních ploch měřených v období 2021–2023. Inventarizační jednotka každé plochy měla rozměr 30 × 30 m. Na jednotlivých plochách byly shromažďovány údaje o tloušťce stromů, odebraných vývrtech, struktuře porostu, půdní vlhkosti, vlhkosti vzduchu a teplotě vzduchu. Hodnoty nadzemní biomasy byly pro jednotlivé inventarizační plochy poskytnuty organizací CzechGlobe již v agregované podobě. Jejich odvození vycházelo z dendrometrických charakteristik jednotlivých stromů a použití alometrických rovnic pro stanovení nadzemní biomasy, která byla následně agregována na úroveň inventarizační plochy 30 × 30 m. Tyto plochy tvořily hlavní lesní referenční soubor pro kalibraci a validaci modelů nadzemní biomasy, přičemž zaznamenané souřadnice plochy byly reprezentovány středovým bodem inventarizační plochy.

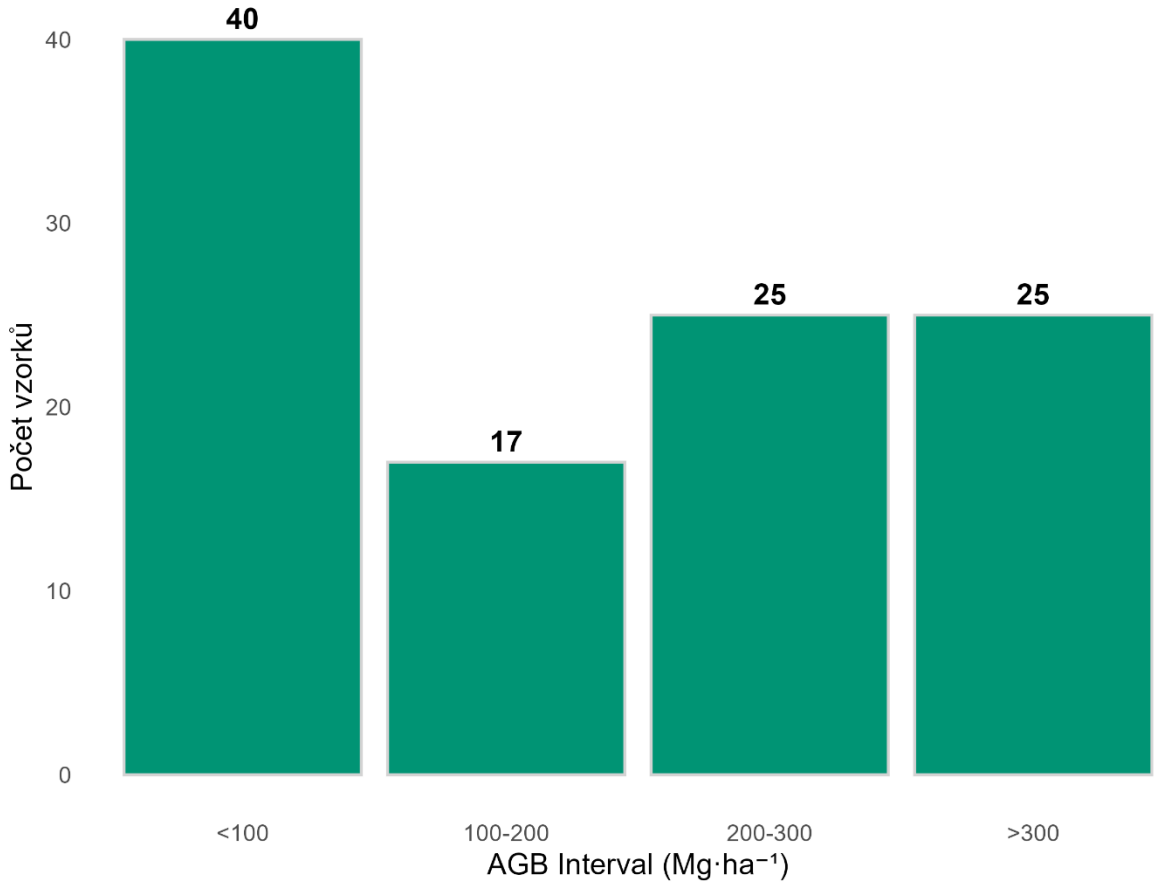


Obrázek 5 Lokality lesní a luční nadzemní biomasy použité jako referenční zdroj dat

Druhou složku referenčního souboru tvořila měření biomasy travních porostů, která reprezentovala nízké hodnoty nadzemní biomasy. Tato data byla získána jednak z vlastních terénních měření realizovaných v rámci disertační práce, jednak z dlouhodobého monitoringu travních porostů poskytnutého doc. Stanislavem Hejdukem z Mendelovy univerzity v Brně. Pro jednotlivé lokality byl stanoven víceletý průměr z měření provedených v období 2018–2023. Celkem bylo do této části souboru zahrnuto 36 lokalit. Měření biomasy travních porostů byla prováděna na typologicky homogenních lokalitách s menším rozsahem vzorkování. Biomasa byla odebírána buď ze dvou sousedních kvadrátů o ploše 1 m^2, nebo z jednoho kvadrátu o ploše 4 m^2, v závislosti na charakteru lokality. Plochy byly voleny tak, aby byla zajištěna vizuální a strukturní homogenita vegetace v okolním prostoru. Zařazení těchto nelesních ploch mělo za cíl lépe reprezentovat spodní část rozsahu hodnot nadzemní biomasy. Tyto situace mohou být relevantní

i uvnitř lesní masky, například v případě řídké vegetace, lesních holin nebo ploch v raných fázích obnovy. Svými spektrálními charakteristikami se takové plochy mohou více podobat travním porostům než zapojeným lesním korunám. Travní lokality tedy byly použity ke snížení rizika systematického nadhodnocování biomasy v oblastech s nízkou vegetační pokryvností.

Celkový referenční soubor zahrnoval 107 lokalit, z toho 71 lesních ploch a 36 nelesních ploch (Obrázek 5). Hodnoty nadzemní biomasy se v celém souboru pohybovaly v rozmezí 1,04 až 553,73 Mg·ha^{-1}, přičemž medián činil 173,62 Mg·ha^{-1}. Rozdělení hodnot bylo bimodální, což souviselo se zastoupením lokalit s velmi nízkou biomasou a lokalit s dřevinnou vegetací. Ve skupině lokalit s hodnotami AGB < 5 Mg·ha^{-1} bylo zahrnuto 36 pozorování, zatímco skupina s hodnotami AGB ≥ 5 Mg·ha^{-1} zahrnovala 71 pozorování. V rámci této druhé skupiny se hodnoty AGB pohybovaly v rozmezí 6,05 až 553,73 Mg·ha^{-1}, s mediánem 267,40 Mg·ha^{-1}. Nejpočetnější skupinu tvořily vzorky s hodnotou AGB do 100 Mg·ha^{-1} (n = 40), přičemž tato kategorie zahrnovala 36 nelesních a 4 lesní pozorování. Interval 100–200 Mg·ha^{-1} byl zastoupen 17 vzorky, zatímco intervaly 200–300 Mg·ha^{-1} a ≥ 300 Mg·ha^{-1} obsahovaly shodně 25 vzorků. Referenční soubor tedy pokrýval celé sledované rozpětí biomasy, avšak se zřetelně nerovnoměrným zastoupením jednotlivých intervalů (Obrázek 6).



Obrázek 6 Rozdělení referenčních hodnot nadzemní biomasy

3.2.9. Harmonizace referenčních dat pro rok 2023

Terénní data byla georeferencována a systematicky ověřena z hlediska polohové přesnosti. Vzhledem k tomu, že lesní porostní data byla sbírána nesynchronně v letech 2021 až 2023, bylo nutné všechna měření harmonizovat tak, aby odpovídala podmínkám roku 2023. K tomuto účelu byly na data z let 2021 a 2022 aplikovány růstové korekce pomocí Annual Tree Growth Index (ATGI). Pro jednoletý posun byla korekce provedena podle rovnice (9):

$$AGB_{2023} = AGB_{2022} \times (1 + \frac{Growth\ Index}{100}) \quad (9)$$

U měření pocházejících z roku 2021 byl postup aplikován analogicky pro harmonizaci na rok 2023. Annual Tree Growth Index byl poskytnut organizací CzechGlobe v rastrové podobě s prostorovým rozlišením 500 m. Raster obsahoval hodnoty v rozsahu 0–2, které vyjadřovaly procentuální roční přírůstek biomasy. Díky rastrové podobě bylo možné každé referenční ploše přiřadit prostorově explicitní a lokalitně specifickou hodnotu růstové korekce. Roční hodnoty ATGI zachycují prostorovou a meziroční variabilitu přírůstu biomasy, která může být ovlivněna klimatickými podmínkami i disturbancemi. Pro omezení možných zkreslení způsobených anomálními roky byl vypočten dlouhodobý průměr ATGI na základě dat z období 2010–2023. Cílem použití dlouhodobého průměru bylo získat stabilnější odhad typického ročního přírůstu pro jednotlivé lokality a omezit vliv meziročních výkyvů a anomálních hodnot v jednotlivých letech. Pro každou plochu byl vypočten víceletý průměr přírůstu, přičemž nejnižší zaznamenaný roční přírůstek činil 0,57 % a nejvyšší dosáhl 1,43 %. ATGI byl použit jako harmonizační faktor pro standardizaci měření na podmínky roku 2023. Použití dlouhodobého průměru přitom nepředstavuje rekonstrukci konkrétního meziročního vývoje jednotlivých porostů a nemůže explicitně zachytit případnou lokální disturbanci mezi rokem terénního měření a rokem 2023. Nejistota spojená s tímto růstovým indexem nebyla samostatně kvantifikována a je proto považována za jednu ze složek celkové nejistoty referenčních dat a následného modelování. Uvedené hodnoty jsou v souladu se širšími studiemi růstových trendů biomasy v Evropě. Podle první iterace integrovaného hodnocení biomasy činil průměrný roční přírůstek biomasy od roku 2000 přibližně 1,3 % (Camia a kol., 2018). Aktualizované hodnocení uvádí mírně širší rozmezí 1–2 % ročně (Avitabile a kol., 2023). Obě studie zároveň upozorňují na nedávné zpomalení akumulace biomasy v posledních pěti letech, které je přičítáno faktorům, jako je stárnutí lesů, změna klimatu a disturbance.

3.2.10. Extrakce hodnot prediktorů pro referenční plochy

Extrakce hodnot prediktorů pro jednotlivé referenční plochy byla provedena v prostředí ArcGIS Pro. Vstupem byly již připravené rastrové vrstvy prediktorů sjednocené do prostorového rozlišení 20 m a zarovnané na společnou rastrovou mřížku. Hodnoty prediktorů byly následně přiřazeny k referenčním lokalitám podle jejich prostorové polohy, přičemž způsob extrakce se lišil podle typu referenční plochy a její prostorové reprezentativnosti.

Lesní referenční plochy byly vymezeny jako čtvercové vzorkovací jednotky o velikosti 30 × 30 m, zatímco predikční proměnné odvozené z družicových dat Sentinel-1 a Sentinel-2 byly zpracovány v prostorovém rozlišení 20 m. Pro sladění tohoto prostorového nesouladu a zajištění prostorově více konzistentní reprezentace lesních měření byl použit přístup založený na lokálním průměrování hodnot prediktorů. Pro každou lesní plochu bylo definováno okolí odpovídající čtyřem pixelům o rozlišení 20 m, zahrnující pixel obsahující střed plochy a jeho nejbližší sousední pixely. Průměrná hodnota těchto čtyř pixelů byla následně použita jako hodnota predikční proměnné pro danou lesní referenční plochu. Tento postup byl upřednostněn před bilineární interpolací, která váží hodnoty centrálního a okolních pixelů podle vzdálenosti. Lokální průměrování lépe odpovídalo prostorové podpoře lesních vzorkovacích ploch a zároveň snižovalo citlivost na drobné subpixelové nesoulady mezi polohou terénního měření a rastrovými prediktory. Použití čtyř okolních pixelů proto nemělo představovat přesnou geometrickou shodu s hranicí plochy 30 × 30 m, ale lokální aproximaci prostorové podpory referenčního měření. Bylo přitom předpokládáno, že hodnota AGB stanovená pro inventarizační plochu je reprezentativní také pro její bezprostřední okolí. Lokální průměrování zároveň omezilo citlivost výsledku na drobné rozdíly v registraci rastrových dat a poloze středu inventarizační plochy.

Naproti tomu u měření biomasy travních porostů nebyla použita žádná fokální statistika. Vzhledem k malé prostorové ploše těchto měření a k homogenitě vybraných lokalit byly hodnoty prediktorů extrahovány přímo z pixelu, ve kterém se měření nacházelo. Odlišné strategie extrakce prediktorů pro lesní a travní referenční data tak odrážely rozdílné prostorové charakteristiky obou typů lokalit a zajišťovaly, že vztahy mezi prediktory a nadzemní biomasou byly odvozeny metodicky konzistentním způsobem. Zahrnutí nelesních vzorků s nízkou biomasou do trénovacího souboru zároveň umožnilo lépe reprezentovat přechod mezi vegetací s nízkou, střední a vysokou biomasou. Tento přístup je v souladu s doporučeními pro modelování AGB ve větším prostorovém měřítku, podle nichž by referenční soubory měly zahrnovat různé typy vegetace, včetně oblastí s nízkou nebo nulovou biomasou, aby se snížilo riziko systematického zkreslení a zlepšila schopnost modelu zobecňovat v heterogenní krajině (Avitabile a kol., 2016).

3.3. Modelování nadzemní biomasy

Metody strojového učení jsou vhodné pro modelování nadzemní biomasy (AGB) z dat dálkového průzkumu Země díky schopnosti vytvářet predikce na základě predikčních rastrů, jako jsou například spektrální pásma Sentinel-2. Strojové učení představuje podmnožinu umělé inteligence (AI) zaměřenou na tvorbu algoritmů a statistických modelů. Tyto algoritmy jsou navrženy tak, aby rozpoznávaly vzory a vztahy v datech a následně tyto poznatky využívaly k predikci výsledků nebo k rozhodování pro dříve nepozorovaná data. V disertační práci byly použity metody Random Forest (RF) a Extreme Gradient Boosting (XGB), které ve srovnání s alternativními algoritmy poskytly nejpřesnější výsledky modelování (Vyvlečka a Pechanec, 2023). Pro účely disertace byla využita regrese k predikci spojitých proměnných, přičemž stejné algoritmy lze použít i pro klasifikaci snímků (kategorické proměnné). Algoritmus RF (Breiman, 2001) trénuje mnoho rozhodovacích stromů, přičemž každý z nich je vytvořen na náhodně vybrané podmnožině trénovacích dat a na náhodné podmnožině prediktorů. Během trénování je každý strom pěstován

do maximální hloubky bez prořezávání. Při predikci algoritmus agreguje výstupy všech jednotlivých stromů. Jednou z klíčových výhod regresního RF je schopnost efektivně pracovat s vícerozměrnými daty i rozsáhlými datovými soubory, přičemž je zároveň méně náchylný k přeučení. Metoda XGB (Chen a Guestrin, 2016) pracuje iterativně: vytváří soubor slabých predikčních modelů, typicky rozhodovacích stromů, a postupně zvyšuje jejich výkon minimalizací zvolené ztrátové funkce. XGB je navrženo tak, aby zvyšovalo výkon i výpočetní efektivitu, a je proto vhodné pro rozsáhlé datové soubory. Třetím použitým algoritmem je LightGBM, který vychází z klasického gradient boosting decision tree (GBDT). Na rozdíl od standardního GBDT zahrnuje LightGBM dvě optimalizační techniky: Gradient-based One-Side Sampling (GOSS) a Exclusive Feature Bundling (EFB). Tato vylepšení umožňují algoritmu efektivně zpracovávat velké datové soubory a vysoký počet proměnných při kratších výpočetních časech ve srovnání s XGB (Ke a kol., 2017). Část metodického postupu a hlavních výsledků prezentovaných v této disertační práci byla publikována také v samostatné vědecké studii (Vyvlečka a Pechanec, 2026).

3.3.1. Trénování modelu a prostorově vnořená křížová validace

Modelování AGB bylo provedeno v prostředí ArcGIS Pro s využitím integrovaného rozhraní Python. Toto prostředí sloužilo jako hlavní platforma pro všechny fáze modelovacího postupu, od předzpracování dat a explorační analýzy až po trénování, validaci a predikci. Python prostředí bylo rozšířeno o všechny nezbytné knihovny pro aplikace strojového učení, přičemž klíčovou roli hrála knihovna *scikit-learn*. Ta poskytuje komplexní sadu nástrojů pro regresi, klasifikaci, shlukování a transformaci dat, což ji činí vhodnou pro vývoj a vyhodnocování predikčních modelů. ArcGIS Pro byl využíván především k zajištění konzistentní práce s rastry (projekce, zarovnání, maskování) a k přímé integraci výstupů modelů do GIS *workflow* používaného pro mapování v národním měřítku a následné zpracování.

Vzhledem k tomu, že některé referenční plochy biomasy byly prostorově blízko sebe a pravděpodobně vykazovaly prostorovou autokorelaci, byla konvenční náhodná k-fold křížová validace považována za nevhodnou, protože může vést k příliš optimistickým odhadům predikčního výkonu. Proto byl použit prostorově explicitní vnořený rámec křížové validace, který odděluje výběr modelu od hodnocení jeho výkonu a poskytuje robustnější a méně zkreslené odhady pro prostorově strukturovaná ekologická data (Krstajić a kol., 2014). Za účelem omezení prostorové závislosti mezi trénovacími a validačními vzorky byly všechny referenční plochy přiřazeny do prostorových skupin odvozených z pravidelné mřížky s velikostí buňky 1 × 1 km. Všechny referenční plochy nacházející se ve stejné buňce mřížky sdílely shodný identifikátor skupiny a byly proto při rozdělení pomocí GroupKFold vždy zařazeny do stejného foldu. Tím bylo zabráněno rozdělení prostorově blízkých vzorků ležících uvnitř stejného 1 km bloku mezi trénovací a validační část. Použitý postup však nepředstavuje pevný 1 km buffer mezi trénovacími a validačními lokalitami, protože vzorky v sousedních buňkách mohou ležet ve vzájemné vzdálenosti menší než 1 km. Velikost bloku 1 km představovala kompromis mezi omezením lokální prostorové závislosti a zachováním dostatečného počtu prostorových skupin pro realizaci pětinásobné vnější i vnitřní křížové validace při omezeném počtu referenčních lokalit.

Trénování a validace modelu byly provedeny pomocí vnořené prostorové křížové validace založené na přístupu *GroupKFold*. Vnořený návrh sestával ze dvou hierarchických smyček: vnější smyčky pro vyhodnocení modelu a vnitřní smyčky pro ladění hyperparametrů. Ve vnější smyčce byl datový soubor rozdělen do pěti prostorových foldů. V každé iteraci byl jeden fold ponechán stranou jako nezávislý testovací soubor, zatímco zbývající čtyři foldy byly použity pro kalibraci modelu. Tento vnější testovací fold nebyl použit v žádné fázi optimalizace hyperparametrů, a představuje tak plně nezávislé prostorové vyhodnocení výkonu modelu. V rámci každé vnější trénovací množiny byla pro ladění hyperparametrů použita vnitřní křížová validace *GroupKFold* s pěti foldy. Hyperparametry byly optimalizovány výhradně pomocí foldů vnitřní křížové validace, čímž bylo zajištěno striktní oddělení mezi výběrem modelu a jeho hodnocením.

Výkon modelu byl hodnocen pomocí tří komplementárních metrik: koeficientu determinace (R^2), střední kvadratické chyby (RMSE) a průměrné absolutní chyby (MAE). R^2 vyjadřuje podíl variability pozorované AGB vysvětlené modelem, zatímco RMSE klade větší důraz na větší chyby predikce a poskytuje interpretovatelnou míru průměrné odchylky v jednotkách $Mg \cdot ha^{-1}$. MAE byla zahrnuta jako doplňková metrika, protože nabízí robustnější odhad typické predikční chyby a je méně citlivá na extrémní hodnoty. Metriky výkonu byly počítány samostatně pro každý vnější fold. Výsledná přesnost modelu je primárně uváděna jako průměr a směrodatná odchylka R^2, RMSE a MAE napříč všemi vnějšími foldy, což odráží nejen celkovou predikční přesnost, ale také stabilitu modelu mezi prostorově oddělenými validačními bloky. Tento způsob hodnocení byl použit zejména pro srovnání jednotlivých sad prediktorů a algoritmů, protože každý vnější fold má stejnou váhu a je tak možné posoudit prostorovou robustnost modelu.

Doplňkově byly vypočteny také *pooled out-of-fold (OOF)* metriky. V tomto případě byly spojeny predikce ze všech vnějších testovacích foldů do jednoho souboru a metriky byly vypočteny nad všemi referenčními pozorováními současně. Každé pozorování bylo i v tomto případě predikováno modelem, který nebyl trénován na daném pozorování. OOF metriky proto nepředstavují nezávislý druh validace, ale jiný způsob agregace výsledků vnějších foldů. Byly použity především pro reziduální diagnostiku, *scatterploty* měřených a predikovaných hodnot, výpočet celkového *biasu* (predikovaná – měřená hodnota) a interpretaci chyb na úrovni celého referenčního souboru. Menší rozdíly mezi průměrnými *fold-wise* metrikami a OOF metrikami jsou očekávatelné, protože průměr přes foldy přisuzuje stejnou váhu každému validačnímu bloku, zatímco OOF metriky jsou ovlivněny počtem pozorování v jednotlivých foldech. Zároveň některé metriky, zejména R^2, nejsou aditivní napříč foldy.

Pro každý algoritmus byla optimalizace hyperparametrů provedena pomocí strategie `grid-search` ve vnitřní smyčce křížové validace. Model RF byl laděn s ohledem na počet stromů (`n_estimators`), maximální hloubku stromu (`max_depth`), minimální počet vzorků potřebný k rozdělení vnitřního uzlu (`min_samples_split`) a počet prediktorů zvažovaných při každém rozdělení (`max_features`). U modelu XGB se vyhledávání hyperparametrů zaměřilo na *learning rate*, hloubku stromu, poměr *subsamplingu*, *column subsampling* a regularizační parametry (`gamma, reg_alpha, reg_lambda`). Tyto parametry hrají klíčovou roli při řízení složitosti modelu a omezování přeučení, zejména při přítomnosti prostorové autokorelace. Model

LightGBM byl laděn napříč parametry řídícími strukturu stromu a strategii vzorkování, včetně `num_leaves`, `max_depth`, `min_data_in_leaf`, `feature_fraction` a `bagging_fraction`. LightGBM používá strategii růstu stromu *leaf-wise*, která umožňuje rychlé snižování ztráty, ale vyžaduje pečlivou regularizaci, aby se předešlo přeučení. Hyperparametrický prostor byl proto omezen tak, aby upřednostňoval stabilní a zobecnitelné konfigurace modelu.

Po dokončení vnořené křížové validace byla finální sada hyperparametrů vybrána na základě nejčastěji se vyskytující optimální konfigurace napříč foldy vnitřní křížové validace (Tabulka 10). Tento přístup zajišťuje, že finální parametry modelu nejsou ovlivněny žádným jednotlivým vnějším testovacím foldem. Důležité je, že výsledná predikční mapa neodpovídá žádnému jednotlivému foldu křížové validace. Naopak představuje model natrénovaný na všech dostupných datech, přičemž očekávaný predikční výkon a prostorová přenositelnost jsou kvantifikovány výsledky vnořené křížové validace.

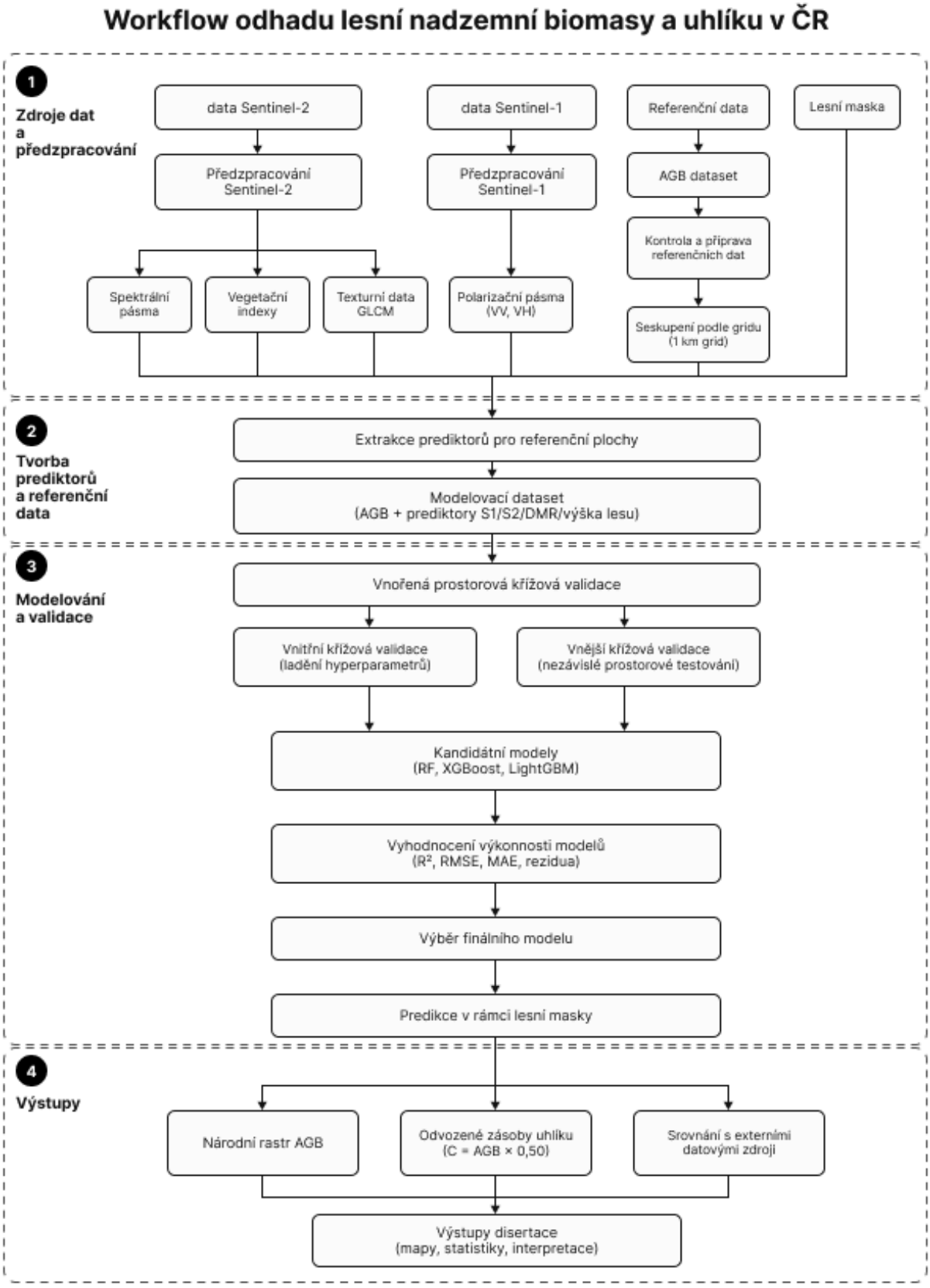
Tabulka 10 Finální hyperparametry modelů

Model	Finální hyperparametry
Random Forest	`max_depth=13; max_features=9; min_samples_split=7; n_estimators=20`
XGBoost	`colsample_bytree=0.8; gamma=2; learning_rate=0.05; max_depth=8; min_child_weight=4; n_estimators=100; reg_alpha=1; reg_lambda=4; subsample=0.8`
LightGBM	`boosting_type=gbdt; bagging_fraction=0.8; bagging_freq=1; feature_fraction=0.65; lambda_l1=5; lambda_l2=10; learning_rate=0.03; max_bin=255; max_depth=6; min_data_in_leaf=10; min_gain_to_split=0.01; n_estimators=300; num_leaves=40`

Původní sada kandidátních prediktorů byla před finálním srovnáním algoritmů iterativně redukována jako součást vývoje modelu. Nejprve byl celý modelovací a validační postup proveden s kompletní sadou dostupných prediktorů. Následně byly odstraněny proměnné s velmi nízkou významností a celý postup trénování, optimalizace hyperparametrů a prostorové validace byl zopakován. Redukovaná varianta byla zachována pouze tehdy, pokud vedla ke zvýšení R^2 a současně ke snížení RMSE a MAE. Tímto postupem byla získána méně komplexní sada prediktorů, která byla následně fixována a konzistentně použita pro finální srovnání algoritmů Random Forest, XGBoost a LightGBM. Finální sada prediktorů použitá pro modelování je uvedena v tabulce 11.

Tabulka 11 Finální prediktory

Finální prediktory	
Sentinel-1	VV, VH
Sentinel-2	B2, B3, B4, B5, B6, B8, B11
GLCM	B4 Entropy, B8 Contrast. B11 GLCM Mean
Vegetační indexy	ARVI, MSAVI, NDI45, ReNDVI



Obrázek 7 Schéma postupu disertační práce

3.4. Využití nástrojů generativní umělé inteligence

Při zpracování disertační práce byl jako podpůrný nástroj využíván generativní jazykový model ChatGPT společnosti OpenAI. Nástroj byl použit zejména při návrhu, optimalizaci a kontrole dílčích částí skriptů vytvořených v jazycích Python a R. Veškeré navržené skripty a jejich úpravy byly autorem následně kontrolovány, testovány a podle potřeby přepracovány. Výpočty, analýzy a modelování byly provedeny autorem v příslušných programových prostředích. Další využití nástroje ChatGPT bylo pro jazykové a stylistické úpravy textu. Nástroj byl využíván také ke kontrole bibliografických údajů. Kromě toho byl nástroj použit i pro sjednocení citační normy. Generativní umělá inteligence nebyla použita k vytváření, pozměňování nebo nahrazování výzkumných dat ani k rozhodování o použitých metodách a interpretaci výsledků. Autor nese plnou odpovědnost za správnost výpočtů, použitých metod, citací, interpretací a výsledné podoby disertační práce.

4. VÝSLEDKY PRÁCE

V rámci disertační práce byly testovány tři algoritmy strojového učení, a to Random Forest, XGBoost a LightGBM, vždy v kombinaci s různými sadami vstupních prediktorů. Cílem bylo posoudit, jak se změna prediktorů promítá do přesnosti odhadu nadzemní biomasy. Mezi vstupní prediktory byla zařazena spektrální pásma Sentinel-2 (B2, B3, B4, B5, B6, B7, B8 a B11), doplněná o vybrané vegetační indexy (ARVI, GNDVI, MSAVI, NDI45 a ReNDVI) a texturní metriky GLCM (Contrast, Entropy a GLCM Mean). Další skupinu prediktorů tvořila radarová data Sentinel-1, konkrétně polarizace VV a VH. Kromě aktuálních satelitních dat byly do modelování zahrnuty také topografické proměnné odvozené z digitálního modelu reliéfu 4. generace, konkrétně nadmořská výška, sklon a orientace svahu. Samostatnou kategorii prediktorů představoval rastr výšky lesního porostu, který byl dostupný pouze pro rok 2020, a vykazoval tedy určitou časovou nekonzistenci vzhledem k ostatním vstupním datům. Přesto byl do analýzy zařazen, protože jedním z dílčích cílů práce bylo ověřit, zda informace o výšce porostu povede ke zvýšení přesnosti predikce nadzemní biomasy, nebo zda naopak její přínos zůstane omezený vzhledem k časové nehomogenitě dat. Modelování bylo realizováno v prostředí ArcGIS Pro s využitím Pythonu, přičemž hlavní nástroje pro přípravu modelů, ladění hyperparametrů a vyhodnocení jejich výkonnosti byly založeny na knihovně *scikit-learn*.

4.1. Srovnání výkonnosti jednotlivých sad prediktorů

4.1.1. Statistické srovnání sad prediktorů

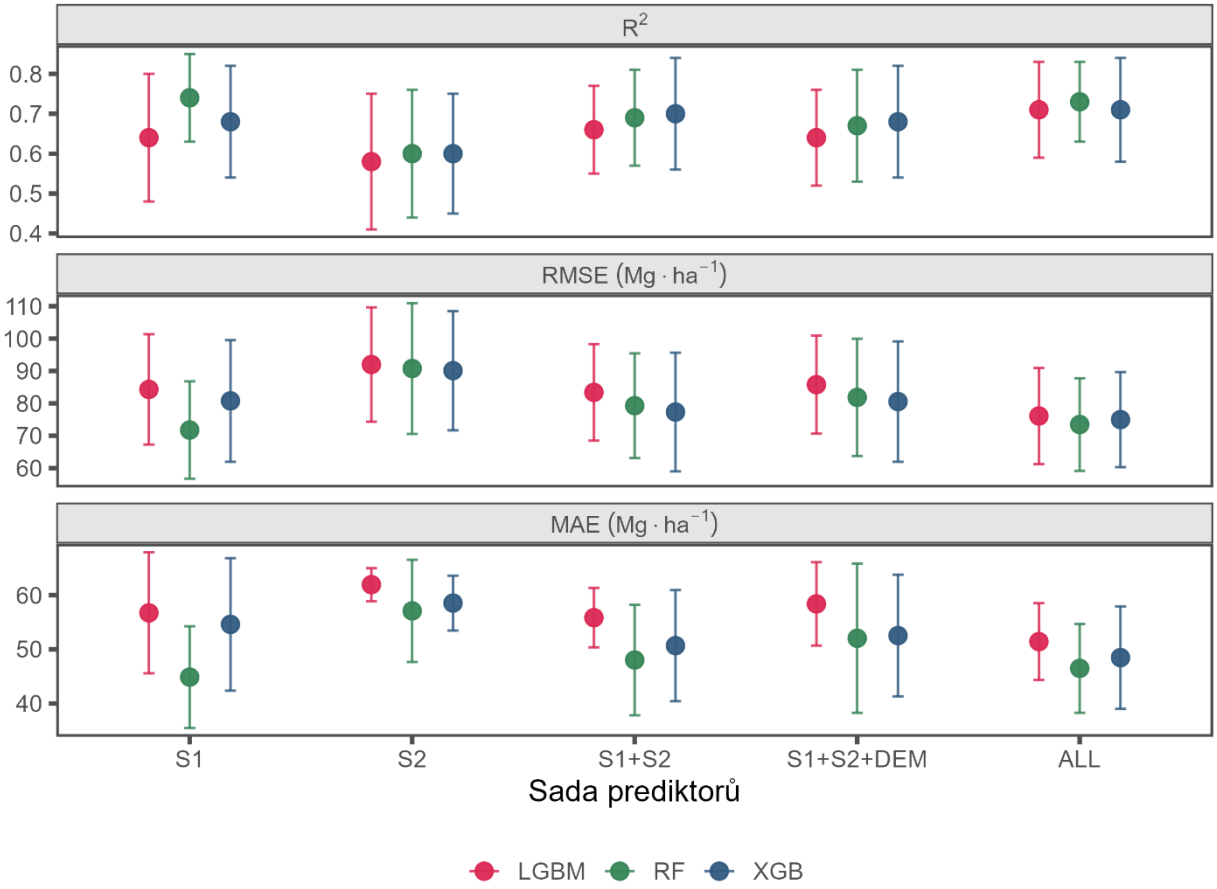
Výkonnost jednotlivých sad prediktorů byla hodnocena pomocí metrik R^2, RMSE a MAE, přičemž tabulka 12 uvádí průměrné hodnoty a směrodatné odchylky dosažené v rámci validace. Cílem této analýzy bylo posoudit, jak se mění přesnost odhadu nadzemní biomasy v závislosti na typu a kombinaci použitých vstupních proměnných.

Tabulka 12 Srovnání všech testovaných modelů

Model	Sada prediktorů	R^2 průměr ± SD	RMSE průměr ± SD	MAE průměr ± SD
RF	S1+S2	0,69 ± 0,12	79,28 ± 16,14	48,03 ± 10,20
RF	S1+S2+DEM	0,67 ± 0,14	81,84 ± 18,13	52,04 ± 13,77
RF	S1	**0,74 ± 0,11**	**71,75 ± 15,03**	**44,87 ± 9,39**
RF	S2	0,60 ± 0,16	90,76 ± 20,18	57,09 ± 9,42
RF	ALL	0,73 ± 0,10	73,46 ± 14,31	46,49 ± 8,21
XGB	S1+S2	0,70 ± 0,14	77,32 ± 18,34	50,68 ± 10,24
XGB	S1+S2+DEM	0,68 ± 0,14	80,56 ± 18,59	52,53 ± 11,24
XGB	S1	0,68 ± 0,14	80,76 ± 18,79	54,60 ± 12,23
XGB	S2	0,60 ± 0,15	90,10 ± 18,42	58,53 ± 5,07
XGB	ALL	**0,71 ± 0,13**	**74,98 ± 14,69**	**48,47 ± 9,43**
LGBM	S1+S2	0,66 ± 0,11	83,39 ± 14,91	55,84 ± 5,49
LGBM	S1+S2+DEM	0,64 ± 0,12	85,77 ± 15,11	58,38 ± 7,69

LGBM	S1	0,64 ± 0,16	84,30 ± 17,04	56,73 ± 11,15
LGBM	S2	0,58 ± 0,17	91,99 ± 17,67	61,93 ± 3,06
LGBM	ALL	**0,71 ± 0,12**	**76,10 ± 14,85**	**51,42 ± 7,09**

Grafické znázornění hodnot uvedených v tabulce 12 je uvedeno na obrázku 8. Obrázek umožňuje přehlednější porovnání výkonnosti jednotlivých sad prediktorů napříč testovanými algoritmy a současně ukazuje rozsah variability vyjádřený směrodatnou odchylkou.



Obrázek 8 Srovnání přesnosti modelů podle použitých sad prediktorů. Body znázorňují průměrné hodnoty metrik R^2, RMSE a MAE dosažené v rámci validace, chybové úsečky odpovídají směrodatné odchylce

Z výsledků je patrné, že mezi testovanými sadami prediktorů existovaly poměrně výrazné rozdíly. Nejhorších výsledků ve všech třech algoritmech dosahovala sada **S2** založená pouze na optických datech Sentinel-2. U modelu Random Forest tato sada dosáhla hodnoty R^2 = 0,60 ± 0,16; RMSE = 90,76 ± 20,18 Mg·ha^{-1} a MAE = 57,09 ± 9,42 Mg·ha^{-1}. Obdobný trend byl patrný také u XGBoost a LightGBM, kde se hodnoty R^2 taktéž pohybovaly pouze mezi 0,58 a 0,60 a RMSE se blížila hodnotě 90 Mg·ha^{-1}. Samotná optická data tedy v daném uspořádání neposkytovala dostatečně silný základ pro přesný odhad biomasy. Naopak velmi dobrých výsledků dosahovala sada **S1** založená pouze na radarových datech Sentinel-1, a to zejména u algoritmu Random Forest. Tato kombinace vykázala vůbec nejlepší numerické výsledky ze všech testovaných variant (R^2 = 0,74 ± 0,11; RMSE = 71,75 ± 15,03 Mg·ha^{-1}; MAE = 44,87 ± 9,39 Mg·ha^{-1}). U modelů XGBoost a LightGBM

však sada Sentinel-1 již tak výrazně nedominovala a její výsledky byly spíše srovnatelné s dalšími kombinacemi prediktorů. Tento výsledek naznačuje, že radarová data nesla v rámci testovaného souboru velmi silnou informaci o struktuře porostu, avšak jejich přínos nebyl napříč algoritmy konzistentní.

Kombinace radarových a optických dat v sadě **S1 + S2** vedla k vyrovnaným a stabilním výsledkům u všech tří algoritmů. U Random Forest dosáhla tato sada hodnot R^2 = 0,69 ± 0,12; RMSE = 79,28 ± 16,14 Mg·ha^{-1} a MAE = 48,03 ± 10,20 Mg·ha^{-1}, u XGBoost R^2 = 0,70 ± 0,14; RMSE = 77,32 ± 18,34 Mg·ha^{-1} a MAE = 50,68 ± 10,24 Mg·ha^{-1}. Také u LightGBM představovala tato sada jednu z nejlépe hodnocených variant. Lze tedy konstatovat, že kombinace Sentinel-1 a Sentinel-2 poskytovala napříč algoritmy robustní kompromis mezi přesností a stabilitou modelu. Za pozornost stojí i to, že rozšíření této kombinace o morfometrické parametry (**S1 + S2 + DEM**) nevedlo v žádném z testovaných algoritmů ke zlepšení modelu. Ve všech případech byly dosažené hodnoty mírně horší než u samotné kombinace S1 + S2. Například u Random Forest došlo k poklesu R^2 z 0,69 na 0,67 a ke zvýšení RMSE z 79,28 na 81,84 Mg·ha^{-1}. Podobný trend byl patrný i u XGBoost a LightGBM. Přidání topografických proměnných tedy v rámci testovaného hyperparametrického prostoru nepřineslo dodatečnou informaci, která by vedla ke zvýšení přesnosti odhadu. Velmi dobrých výsledků dosáhla také kompletní sada prediktorů (**ALL**), zejména u algoritmů XGBoost a LightGBM. U XGBoost představovala tato sada nejlepší variantu (R^2 = 0,71 ± 0,13, RMSE = 74,98 ± 14,69 Mg·ha^{-1}, MAE = 48,47 ± 9,43 Mg·ha^{-1}), obdobně u LightGBM (R^2 = 0,71 ± 0,12, RMSE = 76,10 ± 14,85 Mg·ha^{-1}, MAE = 51,42 ± 7,09 Mg·ha^{-1}). U Random Forest byla však sada ALL mírně horší než samotná sada Sentinel-1. Tyto výsledky naznačují, že rozšíření počtu prediktorů nemusí v testovaném nastavení modelů automaticky vést ke zvýšení přesnosti odhadu a že efekt přidání dalších proměnných je závislý jak na typu algoritmu, tak na charakteru daného datového souboru.

Celkově lze na základě statistického hodnocení konstatovat, že mezi nejsilnější kandidáty patřily sady S1, S1 + S2 a ALL, zatímco sada S2 vykazovala konzistentně nejslabší výsledky a sada S1 + S2 + DEM v rámci testovaného nastavení modelů nepřinesla oproti kombinaci S1 + S2 zlepšení. Současně je však patrné, že rozdíly mezi nejlepšími variantami nebyly ve všech případech zásadní a částečně se překrývaly i jejich směrodatné odchylky. Finální výběr sady prediktorů proto nebyl založen výhradně na globálních přesnostních metrikách, ale byl dále doplněn o prostorové hodnocení výsledných mapových výstupů.

4.1.2. Prostorové hodnocení sad prediktorů

Moranovo I

Pro každou hodnocenou variantu modelu, tj. pro každou kombinaci algoritmu a sady prediktorů, bylo nejprve vypočteno reziduum jako rozdíl mezi měřenou a predikovanou hodnotou nadzemní biomasy v referenčních lokalitách. Tato rezidua byla následně použita jako vstup pro analýzu globální prostorové autokorelace pomocí ukazatele Global Moran's I. Cílem této analýzy bylo ověřit, zda jsou chyby modelu v prostoru rozmístěny náhodně, nebo zda vytvářejí systematické prostorové shluky, které by naznačovaly, že model nezachytil část prostorové struktury variability nadzemní biomasy.

Výpočet byl proveden v prostředí ArcGIS Pro pomocí nástroje *Spatial Autocorrelation (Global Moran's I)*. Vzhledem k nerovnoměrnému a lokálně shlukovitému rozmístění referenčních lokalit byla jako hlavní způsob konceptualizace prostorových vztahů zvolena metoda k nejbližších sousedů (K nearest neighbors). Jako hlavní nastavení bylo použito k = 8, které zajišťuje, že každý bod vstupuje do analýzy se stejným počtem sousedů, a tím omezuje vliv rozdílné hustoty vzorkování mezi jednotlivými částmi území. Pro ověření robustnosti výsledků byla analýza následně zopakována také pro alternativní nastavení k = 4 a doplňkově rovněž při konceptualizaci prostorových vztahů typu inverse distance. Ve všech případech byly získány velmi podobné hodnoty Moranova I, z-score i p-hodnot, což naznačuje, že výsledná interpretace není citlivá na konkrétní volbu tohoto parametru ani na zvolený způsob definice prostorových vah. Výsledky ukázaly, že u žádné z testovaných variant nebyla zjištěna statisticky významná globální prostorová autokorelace reziduí. Hodnoty Moranova I se pohybovaly v blízkosti nuly a ve všech případech byly doprovázeny nevýznamnými p-hodnotami ($p > 0,05$). Na základě tohoto testu tedy nelze tvrdit, že by chyby některé z hodnocených variant vytvářely na globální úrovni systematický prostorový vzor. Současně nebyly zjištěny ani podstatné rozdíly mezi jednotlivými sadami prediktorů z hlediska globální prostorové struktury reziduí.

Současně je však třeba zdůraznit, že Global Moran's I hodnotí pouze celkovou prostorovou strukturu reziduí v referenčních bodech a nepostihuje plně prostorovou věrohodnost výsledné predikční mapy. Z tohoto důvodu bylo statistické hodnocení doplněno také vizuálním porovnáním mapových výstupů jednotlivých sad prediktorů ve vybraných územích.

Tabulka 13 Hodnoty globálního Moranova I pro rezidua jednotlivých kombinací algoritmů a sad prediktorů s použitím K nearest neighbour

Model	Sada prediktorů	Moranovo I	Z-score	p-hodnota
RF	S1	-0,024629	-0,364273	0,715654
RF	S2	-0,027615	-0,444677	0,656553
RF	S1+S2	-0,045817	-0,889476	0,373747
RF	S1+S2+DEM	-0,045361	-0,875123	0,381507
RF	ALL	-0,051805	-1,028234	0,30384
XGB	S1	-0,037991	-0,688457	0,491165
XGB	S2	-0,010064	-0,015424	0,987694
XGB	S1+S2	-0,048197	-0,945166	0,344574
XGB	S1+S2+DEM	-0,04967	-0,978414	0,32787
XGB	ALL	-0,021533	-0,291519	0,770655
LGBM	S1	-0,01497	-0,13361	0,893711
LGBM	S2	0,009646	0,464852	0,642037
LGBM	S1+S2	-0,028579	-0,465616	0,64149
LGBM	S1+S2+DEM	-0,040189	-0,744306	0,456691
LGBM	ALL	-0,017994	-0,205963	0,836819

Vizuální hodnocení datových výřezů

Přestože statistické hodnocení i analýza globální prostorové autokorelace reziduí poskytly důležitou informaci o celkové výkonnosti testovaných variant, samy o sobě nepostačují

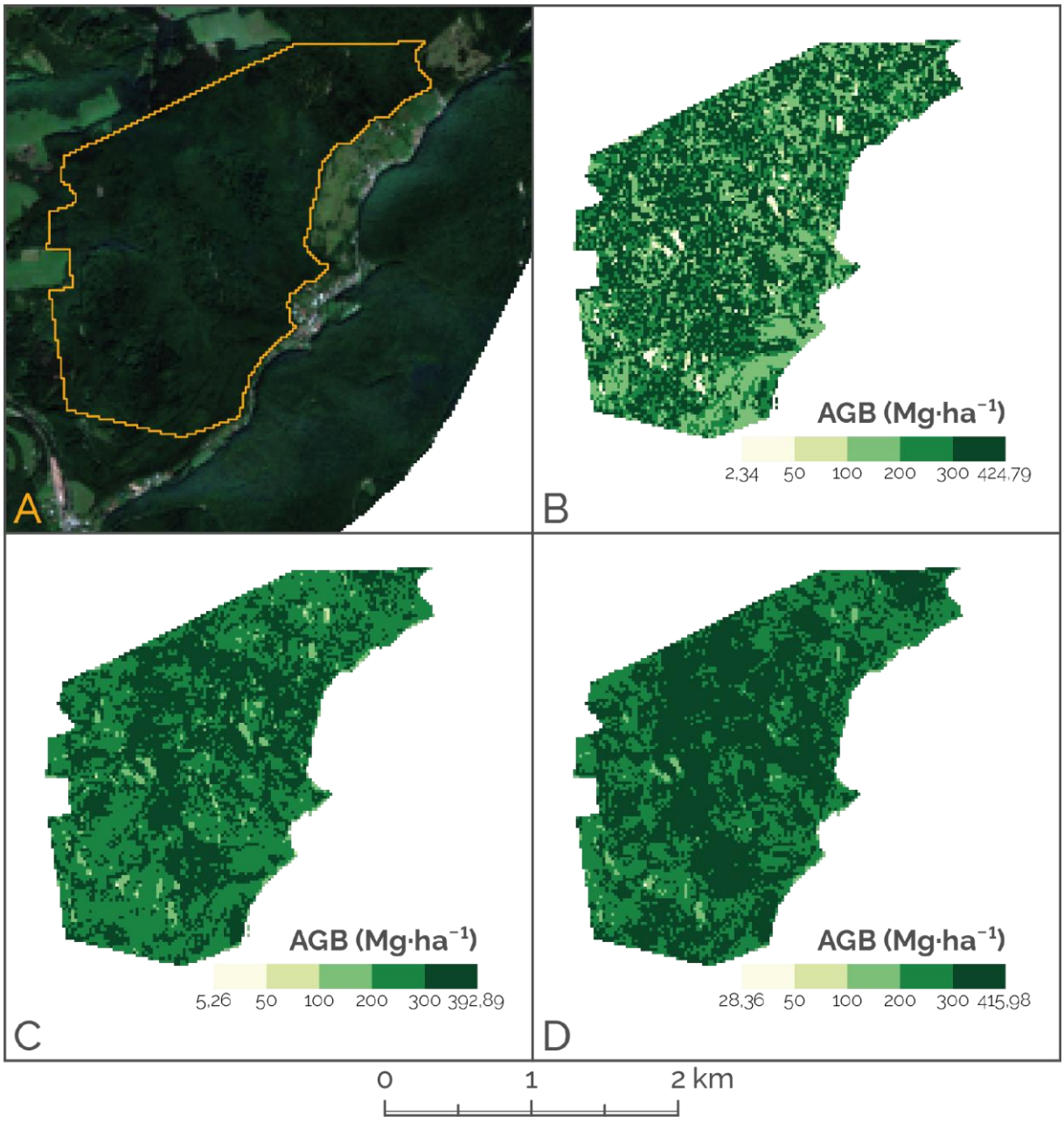
k posouzení kvality výsledného rastru nadzemní biomasy. Zatímco globální metriky charakterizují přesnost predikce na referenčních lokalitách a Global Moran's I hodnotí pouze souhrnnou prostorovou strukturu reziduí, pro účely mapování je důležité posoudit také prostorovou věrohodnost samotných predikčních výstupů. Tato část proto sloužila nejen k doplnění statistického hodnocení, ale především k rozhodnutí, která sada prediktorů představuje nejvhodnější základ pro navazující detailní srovnání algoritmů Random Forest, XGBoost a LightGBM. V této fázi tedy nebylo cílem definitivně určit nejlepší finální model, ale vybrat takovou kombinaci vstupních proměnných, která vedle příznivých hodnot R^2, RMSE a MAE poskytuje také prostorově interpretovatelnou a věrohodnou predikci. Hlavními kandidáty pro další testování byly na základě statistického hodnocení identifikovány zejména sady S1, S1 + S2 a ALL. Prostorové hodnocení mělo následně ověřit, zda numericky úspěšnější varianta zároveň vede k realistickému prostorovému rozložení biomasy, nebo zda některá z alternativních sad poskytuje vhodnější kompromis mezi statistickou přesností a kvalitou predikce.

Za tímto účelem bylo provedeno vizuální porovnání datových výřezů ve vybraných reprezentativních územích:

- Homogenní oblast s kompaktním starším lesním porostem,
- heterogenní oblast s mozaikou starších porostů, holin a regenerujících či mladších porostů,
- okrajovou část lesa v kontaktu s nelesní vegetací.

Jako referenční podklad bylo využito satelitních dat Sentinel-2 z období odpovídající prediktorům S2 (1. červen až 20. září 2023). Pro vizuální porovnání prostorové věrohodnosti sad prediktorů byly ve všech třech výřezech použity výstupy modelu Random Forest, aby bylo možné porovnat vliv sad S1, S1+S2 a ALL při shodném modelovacím přístupu. Všechny mapové výstupy byly současně klasifikovány do stejných intervalů AGB (0–50, 50–100, 100–200, 200–300 a >300 $Mg \cdot ha^{-1}$), aby byla zajištěna přímá vizuální porovnatelnost mezi jednotlivými variantami.

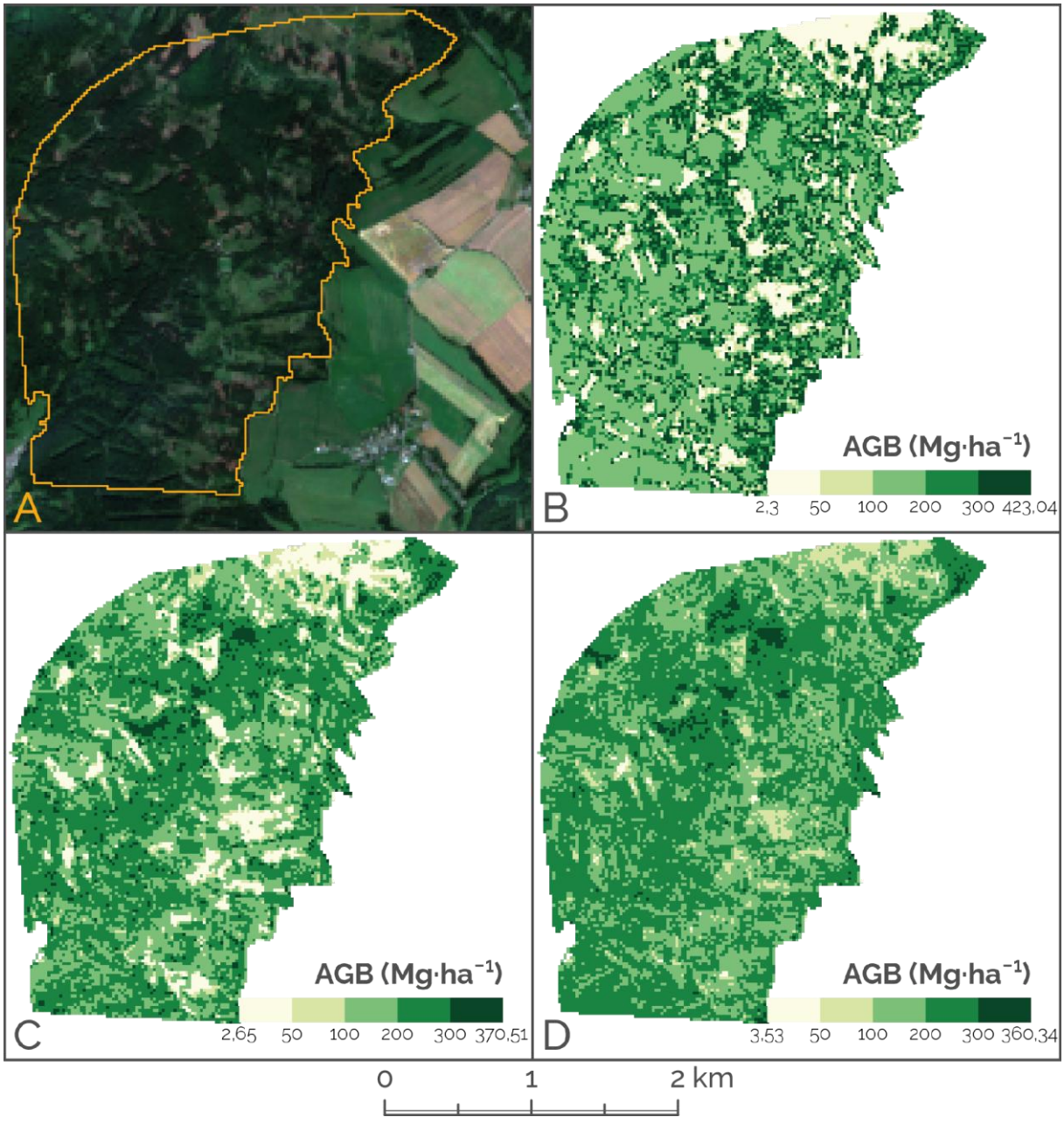
První výřez se nachází v Bílých Karpatech v okolí Vlárského průsmyku a celková plocha porovnávaného území činí přibližně 410 ha. Z typologického hlediska jde o relativně homogenní lesní celek s výraznou převahou SLT 4B a 3B, které dohromady zaujímají přibližně 93,6 % plochy výřezu. Po zahrnutí typu 4D činí podíl hlavních zastoupených typů přibližně 96,9 %. Výřez byl proto využit jako reprezentant převážně homogenního staršího porostu, vhodný pro hodnocení vnitřní variability a prostorové plynulosti modelových výstupů. V prvním výřezu (Obrázek 9), se rozdíly mezi sadami prediktorů projevily především ve vnitřní variabilitě ploch. Varianta S1 vykazovala zvýšenou pixelovou fragmentaci a méně přesvědčivě zachycovala souvislé porosty s vysokou biomasou. Oproti tomu kombinace S1+S2 vytvářela kompaktnější a interpretačně věrohodnější obraz homogenního porostu. Varianta ALL působila ještě vyhlazenějším dojmem.



Obrázek 9 Prostorové porovnání predikované AGB ve výřezu 1. A: RGB kompozice Sentinel-2 s vyznačeným územím. B: model Random Forest se sadou prediktorů S1. C: model Random Forest se sadou prediktorů S1+S2. D: model Random Forest se sadou prediktorů ALL.

Druhý výřez se nachází v Hanušovické vrchovině v okolí Zábřehu a jeho celková plocha činí přibližně 642,4 ha. Na rozdíl od prvního výřezu jde o území stanovištně i prostorově výrazně pestřejší. Největší zastoupení zde mají SLT 3S, 4S a 3H, doplněné širší mozaikou dalších typů s menším plošným podílem. Výřez tak reprezentuje heterogenní lesní prostředí s kombinací starších porostů, otevřených ploch a regenerujících částí lesa a byl použit k hodnocení schopnosti modelů zachytit mozaiku porostních stadií a disturbovaných ploch. Ve výřezu 2 (Obrázek 10) se projevily nejvýraznější rozdíly mezi sadami. Varianta S1 i zde vykazovala vyšší lokální fragmentaci a holiny či obnovující se porosty byly v některých částech rozlišeny méně zřetelně. Kombinace S1+S2 naopak poskytovala nejvyváženější výstup, který lépe zachycoval jak souvislé

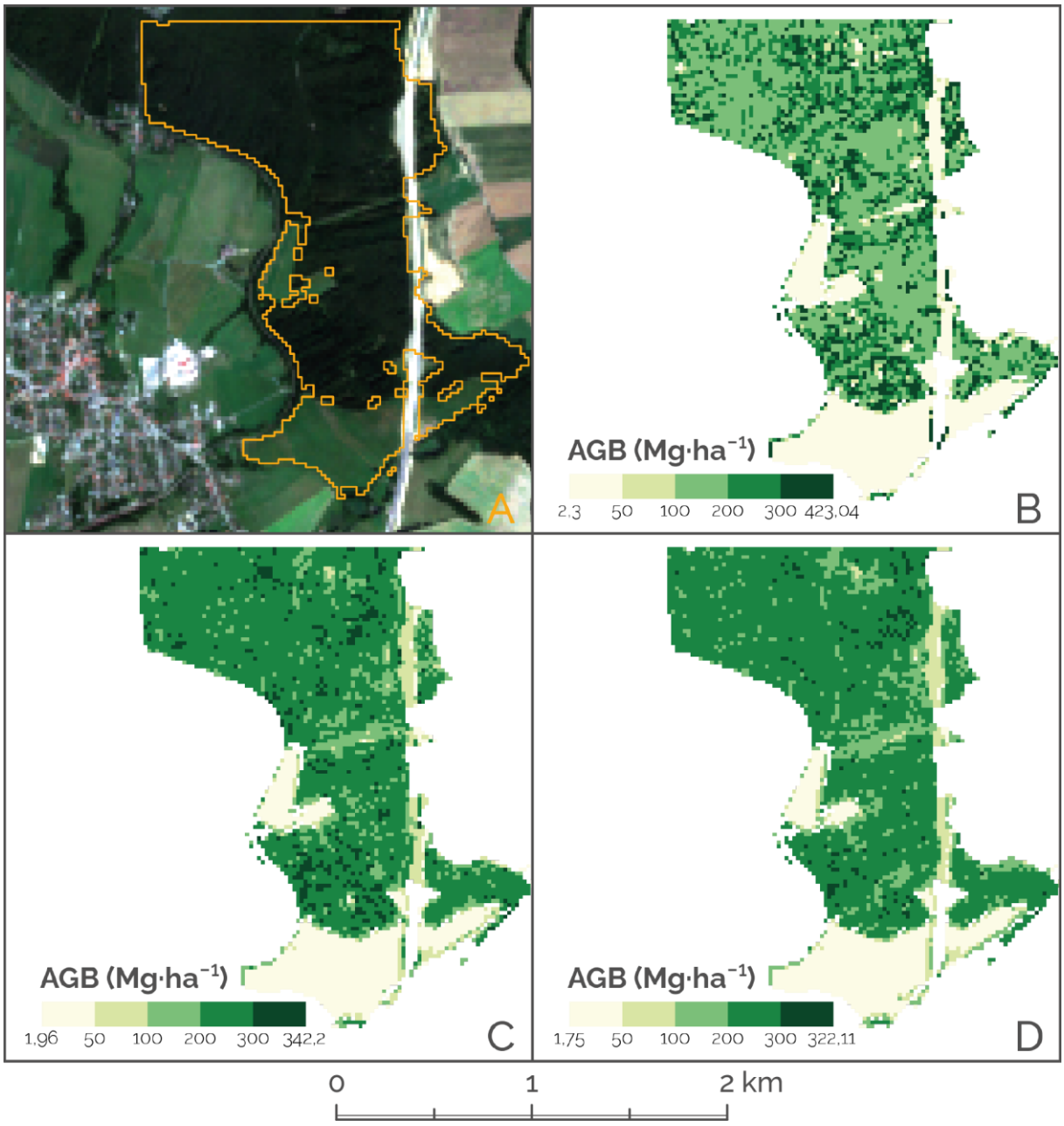
starší porosty, tak plochy s nízkou biomasou odpovídající disturbovaným nebo obnovovaným částem lesa. Varianta ALL působila prostorově kompaktněji, avšak současně výrazněji vyhlazovala jemnější vnitřní strukturu území a méně zřetelně rozlišovala plochy s nízkou biomasou.



Obrázek 10 Prostorové porovnání predikované AGB ve výřezu 2. A: RGB kompozice Sentinel-2 s vyznačeným územím. B: model Random Forest se sadou prediktorů S1. C: model Random Forest se sadou prediktorů S1+S2. D: model Random Forest se sadou prediktorů ALL.

Třetí výřez se nachází v okolí Kamenného Újezdu poblíž Českých Budějovic a jeho rozloha činí 233,76 ha. Tento výřez byl zvolen především pro posouzení chování modelů v přechodové zóně mezi lesním a nelesním prostředím. Hlavním cílem zde bylo ověřit, do jaké míry jednotlivé varianty zachycují hranici mezi lesem a loukou, tedy zda správně oddělují plochy s nízkou biomasou od zapojených lesních porostů a zda výsledný přechod působí prostorově věrohodně.

Ve výřezu 3, zaměřeném na přechod mezi lesním a nelesním prostředím (Obrázek 11), byly rozdíly mezi sadami prediktorů méně výrazné než u předchozích dvou výřezů. Varianta S1 místy naznačovala ostřejší oddělení ploch s nízkou biomasou na hraně lesa, tento efekt však nebyl zcela konzistentní a byl doprovázen zvýšenou pixelovou fragmentací uvnitř lesních částí. Kombinace S1+S2 poskytovala prostorově podobné zachycení přechodu mezi lesní a nelesní biomasou. Místy byl přechod o něco méně konzistentní než u první varianty. To stejné platí i u poslední varianty (ALL).



Obrázek 11 Prostorové porovnání predikované AGB ve výřezu 3. A: RGB kompozice Sentinel-2 s vyznačeným územím. B: model Random Forest se sadou prediktorů S1. C: model Random Forest se sadou prediktorů S1+S2. D: model Random Forest se sadou prediktorů ALL.

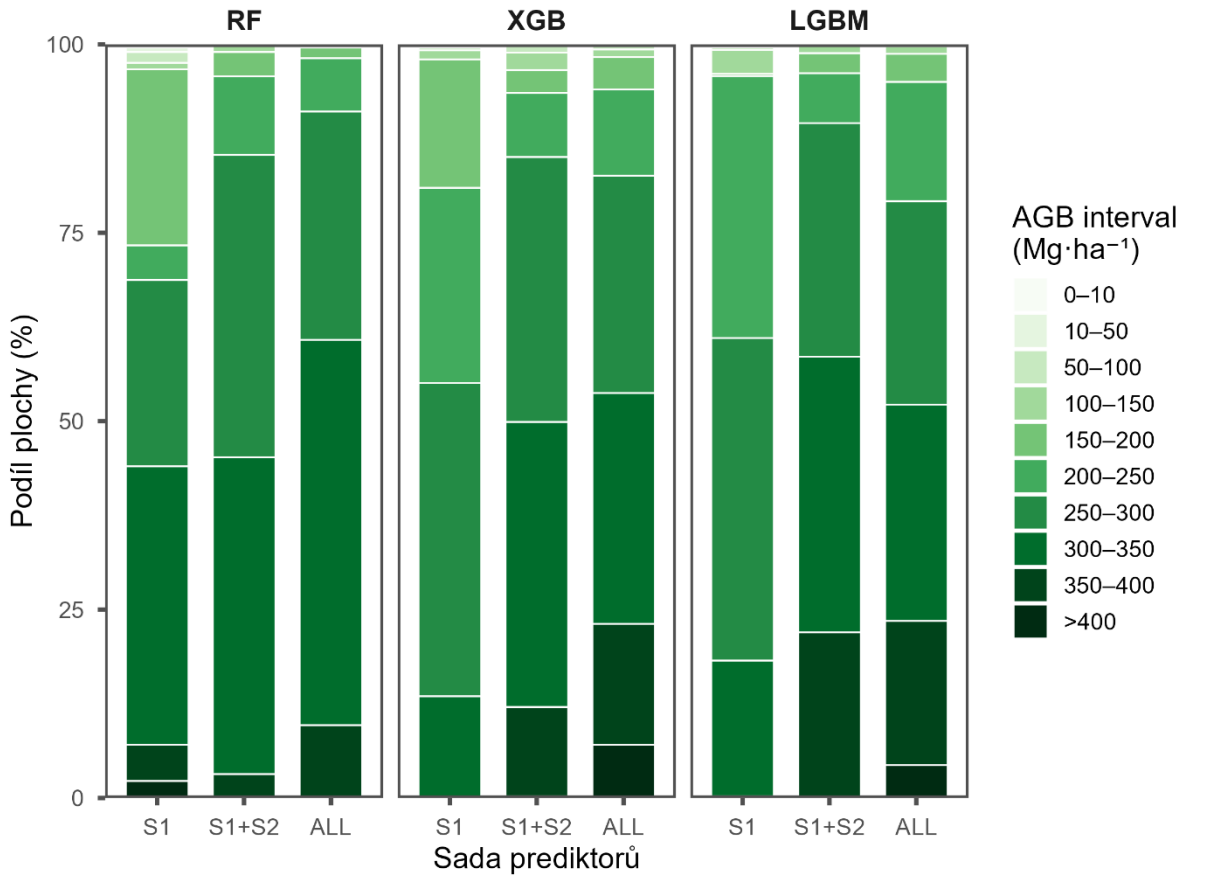
Vizuální porovnání datových výřezů ukázalo, že samotná sada S1 sice v některých případech vedla k výraznějšímu kontrastu mezi plochami s nízkou a vysokou biomasou, avšak současně

vykazovala vyšší míru pixelové fragmentace a méně stabilní prostorový vzor, zejména v homogenních lesních porostech. Sada ALL naopak poskytovala prostorově klidnější a kompaktnější výstupy, avšak v heterogenních územích i na přechodech mezi různými typy vegetačního krytu často potlačovala jemnější vnitřní diferenciaci. Jako nejvyváženější se z hlediska prostorové věrohodnosti jevila kombinace S1+S2, která lépe zachycovala souvislé homogenní porosty, současně však umožňovala rozlišit holiny, obnovující se plochy i přechodové zóny mezi lesem a nelesním prostředím.

Kvantitativní hodnocení datových výřezů

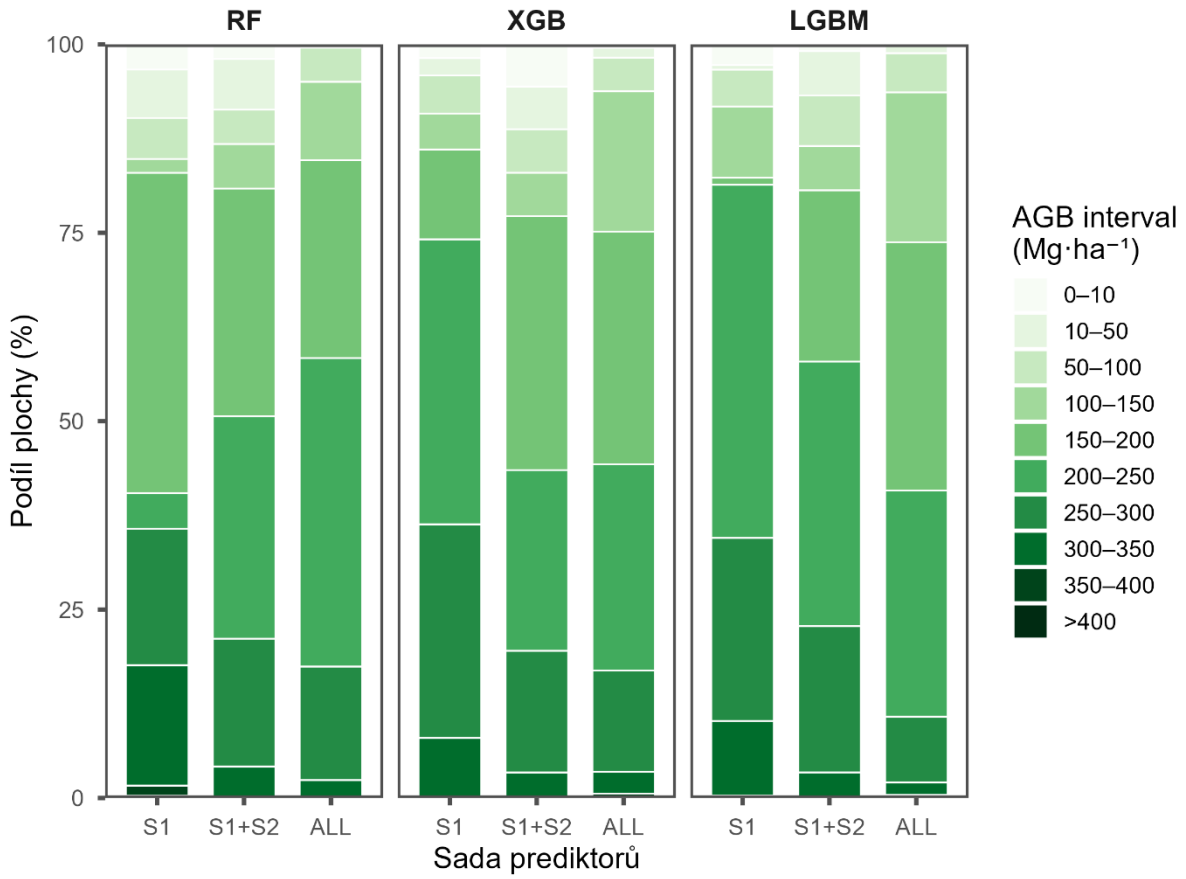
Pro každý výřez a každou hodnocenou variantu byly dále vyčísleny základní souhrnné statistiky predikované AGB (minimum, maximum, průměr, medián a směrodatná odchylka) a plošné zastoupení hodnot biomasy v předem definovaných intervalech. Tyto charakteristiky sloužily k doplnění vizuálního hodnocení a k porovnání toho, jak jednotlivé varianty reprezentují strukturu biomasy v odlišných typech prostředí. Na základě prostorového hodnocení tří reprezentativních výřezů se ukázalo, že mezi sadami prediktorů S1, S1+S2 a ALL existují systematické rozdíly, které globální metriky samy o sobě plně nevystihují. Podrobnější analýza výřezů potvrdila rozdíly ve vnitřní variabilitě výstupů i v rozdělení plochy do jednotlivých intervalů AGB.

V prvním výřezu, který reprezentoval převážně homogenní starší lesní porost, vykazovala sada S1 ve všech třech algoritmech nižší průměrné hodnoty biomasy než varianty S1+S2 a ALL. U modelu RF činila průměrná hodnota pro S1 266,48 $Mg \cdot ha^{-1}$, zatímco u S1+S2 to bylo 289,36 $Mg \cdot ha^{-1}$ a u ALL 304,45 $Mg \cdot ha^{-1}$. Obdobný trend byl patrný i u XGB (250,53 vs. 293,17 a 303,66 $Mg \cdot ha^{-1}$) a u LightGBM (260,50 vs. 306,90 a 301,26 $Mg \cdot ha^{-1}$). Stejně tak podíl plochy ve vyšších třídách biomasy nad 250 $Mg \cdot ha^{-1}$ byl u varianty S1 systematicky nižší než u S1+S2 a ALL. Například u XGB zaujímaly tyto třídy při použití S1 přibližně 225,84 ha, zatímco u S1+S2 348,84 ha a u ALL 338,52 ha. V homogenním starším porostu tedy sada S1 vedla spíše k posunu predikce do nižších tříd biomasy a méně přesvědčivě reprezentovala souvislé porosty s vysokou biomasou. Kombinace S1+S2 a ALL naopak lépe odpovídaly charakteru kompaktního staršího lesa.



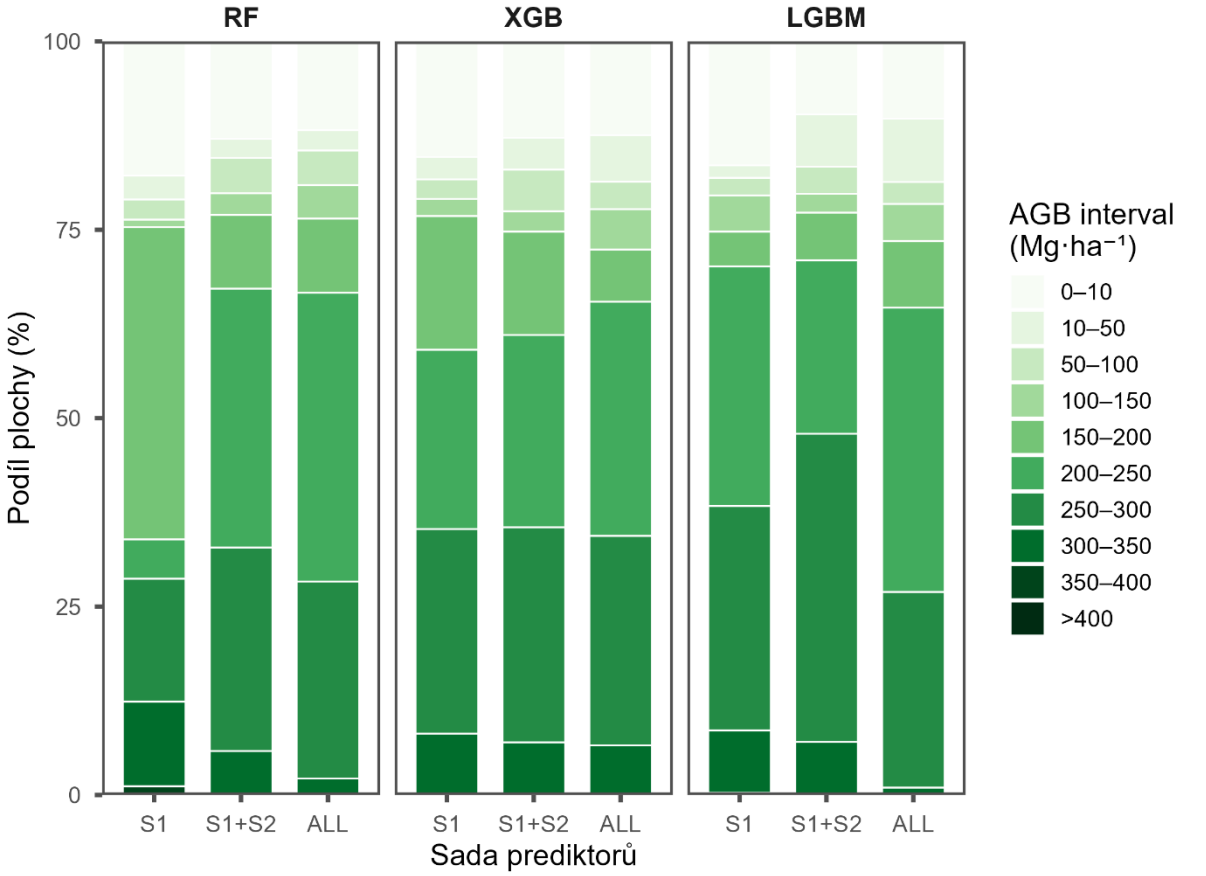
Obrázek 12 Rozdělení ploch do intervalů AGB pro sady prediktorů v homogenním starším lesním porostu

Ve druhém výřezu, který představoval heterogenní lesní území s mozaikou starších porostů, holin a regenerujících ploch, se rozdíly mezi sadami prediktorů projevily odlišně. Varianta ALL zde ve všech třech algoritmech vykazovala nejnižší směrodatnou odchylku a nejvíce koncentrovala plochu do středních tříd biomasy, což naznačuje spíše vyhlazený výstup. Například u RF činila směrodatná odchylka u ALL 53,22 Mg·ha^{-1}, zatímco u S1+S2 74,70 Mg·ha^{-1} a u S1 dokonce 91,61 Mg·ha^{-1}. Podobně u XGB byla směrodatná odchylka pro ALL 61,89 Mg·ha^{-1} oproti 80,03 Mg·ha^{-1} u S1+S2 a 72,21 Mg·ha^{-1} u S1. Sada ALL tak působí v heterogenním prostředí kompaktněji, ale současně méně zvýrazňuje rozdíly mezi plochami s nízkou a vysokou biomasou. Naproti tomu kombinace S1+S2 v heterogenním výřezu zachycovala širší a ekologicky uvěřitelné spektrum biomasy, zejména lépe reprezentovala plochy s nízkou biomasou odpovídající holinám a mladším porostům. U XGB například třídy do 100 Mg·ha^{-1} zaujímaly při použití S1+S2 109,56 ha, zatímco u S1 pouze 59,04 ha a u ALL 40,00 ha. Podobně u LightGBM dosahoval součet tříd do 100 Mg·ha^{-1} při použití S1+S2 86,72 ha, zatímco u S1 53,04 ha a u ALL 40,96 ha. Varianta S1 se v tomto výřezu často posouvala do vyšších tříd biomasy a v některých případech působila méně stabilně, zejména u RF, kde se projevila velmi vysokou vnitřní variabilitou. Z hlediska heterogenního lesního prostředí se tedy jako nejvyváženější jevila sada S1+S2, zatímco ALL měl tendenci skutečnou heterogenitu spíše vyhlazovat.



Obrázek 13 Rozdělení ploch do intervalů AGB pro sady prediktorů v heterogenním lesním prostředí

Ve třetím výřezu, zaměřeném na přechod mezi lesem a loukou, byly rozdíly mezi sadami prediktorů menší než ve výřezech 1 a 2. Všechny varianty zde vykazovaly vysokou vnitřní variabilitu, což je vzhledem k charakteru území očekávatelné. Sada S1 v některých případech vedla k o něco většímu zastoupení nízkých tříd biomasy, zejména u Random Forest, kde třídy do 100 Mg·ha^{-1} zaujímaly 55,32 ha oproti 47,12 ha u S1+S2 a 44,60 ha u ALL. Současně však u stejného modelu klesla průměrná hodnota biomasy na 173,39 Mg·ha^{-1} oproti 196,70 Mg·ha^{-1} u S1+S2 a 192,84 Mg·ha^{-1} u ALL. Tento výsledek může naznačovat ostřejší oddělení nízko biomasových ploch na hraně lesa, avšak tento efekt se nepotvrdil konzistentně i u XGB a LightGBM, kde byly rozdíly mezi sadami menší. Na základě třetího výřezu proto nelze obecně tvrdit, že by samotná sada S1 byla jednoznačně lepší pro zachycení hran mezi lesním a nelesním prostředím.



Obrázek 14 Rozdělení ploch do intervalů AGB pro sady prediktorů při přechodu lesní a nelesní plochy

Celkově lze konstatovat, že sada S1 sice v některých případech dosahovala velmi dobrých globálních statistických výsledků, avšak z hlediska prostorového chování byla méně stabilní a méně konzistentní napříč různými typy prostředí. V homogenním starším lese měla tendenci posouvat predikci do nižších tříd biomasy, v heterogenním prostředí vykazovala u některých algoritmů nadměrnou variabilitu nebo posun do vyšších tříd a na přechodu lesa a louky její případná výhoda nebyla prokázána jednoznačně. Sada ALL naopak poskytovala prostorově kompaktnější a vyhlazenější výstupy, ale v heterogenních územích částečně potlačovala zastoupení nízko biomasových ploch a tím i jemnější vnitřní diferenciaci porostních stadií. Jako nejvyváženější se proto jeví kombinace S1+S2, která v homogenním lese lépe reprezentovala vysokou biomasu, v heterogenním výřezu zachovávala širší a věrohodnější spektrum biomasy a v přechodové zóně poskytovala stabilní a interpretovatelný výstup bez výrazného nárůstu šumu. Vizuální porovnání představovalo doplňkovou kvalitativní diagnostiku prostorového chování výsledné vrstvy, nikoliv nezávislou validaci jeho přesnosti. Z tohoto důvodu lze sadu S1+S2 považovat za nejvhodnější kandidát pro další detailní porovnání algoritmů RF, XGB a LightGBM.

Současně však výsledky naznačují, že samotná sada S1 by neměla být z dalšího testování zcela vyloučena. Její dobré číselné výsledky ukazují, že radarová data nesou významnou informaci o biomase, avšak jejich prostorový výstup může být citlivější na lokální variabilitu a šum. Proto je

možné uvažovat i o navazující optimalizaci této sady, například úpravou radarových vstupů pomocí vyhlazení, pokud by se tím podařilo zlepšit prostorovou konzistenci výsledné mapy.

Úprava sady S1 pomocí *edge-preserving* filtrů

Jelikož výsledná AGB z datové sady S1 dosahovala nejlepších statistických výsledků, ale v prostorové věrohodnosti nebyla tak přesvědčivá, byly na ní otestovány dva filtry zachovávající hrany (*edge-preserving*): *bilateral filter* a *guided filter*. *Low-pass* filtry, které shlazují obraz rovnoměrně ani nebyly použity, právě kvůli ostrým přechodům mezi lesem a holinami. *Bilateral filter* kombinuje prostorovou vzdálenost a podobnost hodnot pixelů, takže potlačuje lokální šum a současně zachovává hrany mezi oblastmi s výrazně odlišnými hodnotami biomasy (Tomasi a Manduchi, 1998). *Guided filter* vychází z lokálně lineárního modelu a rovněž rozlišuje mezi homogenními a vysoce proměnlivými oblastmi, přičemž hrany zachovává lépe než bilinear filtr (He a kol., 2012). Parametry byly přizpůsobeny dynamickému rozsahu dat AGB. U bilateral filtru bylo použito okno 9 × 9 pixelů (d = 9) se `sigmaSpace = 5` a `sigmaColor = 50`, což při velikosti pixelu 20 m odpovídá prostorovému dosažení přibližně 100 m a potlačení lokálních rozdílů v řádu desítek Mg·ha^{-1}. U guided filtru bylo použito okno 11 × 11 pixelů (`radius = 5`) a `eps = 2500`, tedy regularizační práh odpovídající přibližně kontrastu 50 Mg·ha^{-1}. Oba filtry tak byly nastaveny k potlačení *salt and pepper* efektu u sady S1 v rámci homogenních oblastí bez ztráty klíčových hran mezi holinami, mladými porosty a zapojeným lesem. Účinek filtrů byl následně hodnocen na základě vizuálního porovnání ve třech reprezentativních výřezech a dále pomocí souhrnných statistik rastru a rozdělení plochy do intervalů AGB. Pozornost byla věnována zejména změně vnitřní variability vrstvy, zachování homogenních porostů, rozlišitelnosti holin a mladých porostů a charakteru přechodů mezi lesním a nelesním prostředím.

Po aplikaci filtrů bylo ve všech třech analyzovaných výřezech patrné výrazné potlačení *salt and pepper* efektu, který byl charakteristický pro původní S1 predikci. Současně se však mezi oběma přístupy objevily zřetelné rozdíly v intenzitě vyhlazení a zachování lokální prostorové variability biomasy. Bilateral filtr představoval konzervativnější přístup, který zachovával širší dynamický rozsah hodnot AGB i jemnější prostorovou heterogenitu uvnitř porostů. Ve druhém výřezu došlo po aplikaci obou filtrů ke snížení extrémních hodnot. Maximální hodnota původní vrstvy činila 423,04 Mg·ha^{-1}, po bilateral filtraci 387,26 Mg·ha^{-1} a po guided filtraci 390,38 Mg·ha^{-1} (Tabulka 14). Výraznější rozdíl mezi filtry byl patrný ve směrodatné odchylce, která u bilateral filtru dosáhla 83,64 Mg·ha^{-1}, zatímco u guided filtru pouze 72,38 Mg·ha^{-1}. Bilateral filtr tak zachoval větší část původní lokální variability. Současně lépe zachovával jemnější hranice mezi jednotlivými porostními segmenty a lokální rozdíly uvnitř heterogenních lesních struktur. Guided filtr naopak vykazoval výrazně silnější homogenizační efekt. Ve všech výřezech docházelo k výraznějšímu omezení lokální variability a k redukci izolovaných extrémních pixelů. Současně systematicky zvyšoval minimální hodnoty biomasy, čímž potlačoval jednotlivé nízkohodnotové pixely a vytvářel kompaktnější výstup. Tento efekt byl zvláště patrný v homogenních starších porostech, kde výsledná vrstva působila prostorově velmi souvisle a vizuálně věrohodně. Snížení směrodatné odchylky však nelze interpretovat jako jednoznačné zlepšení kvality predikce, protože příliš intenzivní vyhlazení může vést ke ztrátě reálné prostorové variability biomasy.

Nižší rozptyl hodnot, proto v tomto případě představuje spíše indikátor intenzity filtrace než přímý ukazatel vyšší přesnosti modelu. Guided filtr tak sice produkoval vizuálně homogennější výstup, avšak místy již docházelo k částečné ztrátě jemné vnitřní struktury porostu a k většímu zobecnění heterogenních oblastí.

Ve třetím výřezu, zaměřeném na přechody mezi lesním a nelesním prostředím, oba *edge-preserving* filtry současně zachovaly hlavní hranice mezi zapojeným lesem, holinami a otevřenými plochami. Na rozdíl od běžných *low-pass* filtrů tedy nedocházelo k výraznému rozmazání okrajů porostů ani ke ztrátě hlavních prostorových struktur. Výsledky tak potvrzují, že *edge-preserving* filtrace představuje vhodný nástroj pro zvýšení vizuální interpretovatelnosti radarově založených AGB vrstev. Celkově filtrace nepředstavovala jednoznačné zlepšení modelového výstupu, ale spíše kompromis mezi redukcí radarového šumu a zachováním prostorové informace. Guided filtr poskytoval vizuálně hladší výstup, avšak za cenu větší homogenizace vrstvy. Bilateral filtr naproti tomu zachovával větší část původní variability a z interpretačního hlediska se jevil jako vhodnější varianta pro opatrné vyhlazení radarově založené predikce AGB. Je však pravděpodobné, že při odlišném nastavení parametrů, zejména nižší hodnotě regularizačního parametru eps nebo menším okně filtrace, by guided filtr mohl lépe zachovat jemnější prostorovou heterogenitu a představovat konkurenceschopnou alternativu bilateral filtru. Přestože filtrované varianty působily prostorově věrohodněji než původní S1 raster, jednalo se stále pouze o následnou úpravu výsledné vrstvy, nikoliv o zlepšení samotné predikční schopnosti modelu.

Tabulka 14 Srovnání původních verzí modelu S1 s upravenými verzemi s bilateral nebo guided filtrem

Verze modelu		Min	Max	Průměr	Medián	Směrodatná odchylka
Výřez 1	Původní S1 model	2,34	424,79	266,48	299,45	71,07
	Aplikace bilateral filtru	12,74	396,66	267,08	296,01	59,54
	Aplikace guided filtru	44,72	384,98	265,23	285,61	49,58
Výřez 2	Původní S1 model	2,29	423,04	201,23	182,27	91,61
	Aplikace bilateral filtru	4,42	387,26	200,63	185,18	83,64
	Aplikace guided filtru	11,91	390,38	201,04	191,63	72,38
Výřez 3	Původní S1 model	2,29	423,04	173,39	177,15	104,62
	Aplikace bilateral filtru	2,45	377,11	172,02	180,60	98,24
	Aplikace guided filtru	4,98	379,03	174,88	187,02	86,31

4.2. Srovnání algoritmů pro sadu prediktorů Sentinel-1 + Sentinel-2

Po předchozím vyhodnocení jednotlivých sad prediktorů byla jako nejvhodnější kompromis mezi statistickou přesností a prostorovou věrohodností vybrána kombinace Sentinel-1 a Sentinel-2. V dalším kroku proto bylo provedeno detailní srovnání tří algoritmů strojového učení, Random Forest, XGBoost a LightGBM, a to již při fixované sadě prediktorů S1+S2. Cílem této části bylo určit, který z algoritmů poskytuje nejvhodnější rovnováhu mezi přesností a robustností pro predikci lesní nadzemní biomasy na úrovni celé České republiky. Výkonnost modelů byla hodnocena pomocí vnořené prostorové křížové validace, která poskytuje realističtější odhad generalizační

schopnosti než náhodná křížová validace. Tento přístup je praktičtější v úlohách dálkového průzkumu Země, kde se prostorově blízké vzorky často vyznačují podobnou strukturou porostu, stanovištními podmínkami a spektrálním chováním. Náhodné dělení dat by mohlo vést k nadhodnocení přesnosti modelu.

Při hodnocení na základě průměrných hodnot z vnějších validačních bloků dosáhl nejvyšší průměrné predikční přesnosti model XGBoost, který vykázal hodnoty R^2 = 0,70 ± 0,14, RMSE = 77,3 ± 18,3 Mg·ha^{-1} a MAE = 50,7 ± 10,2 Mg·ha^{-1}. Model Random Forest dosáhl velmi podobných výsledků, konkrétně R^2 = 0,69 ± 0,12, RMSE = 79,3 ± 16,1 Mg·ha^{-1} a MAE = 48,0 ± 10,2 Mg·ha^{-1}, a vyznačoval se srovnatelnou stabilitou mezi jednotlivými validačními bloky. Naproti tomu LightGBM vykázal ze srovnávaných přístupů nejslabší výsledky, a to R^2 = 0,66 ± 0,11, RMSE = 83,4 ± 14,9 Mg·ha^{-1} a MAE = 55,8 ± 5,5 Mg·ha^{-1}.

Tabulka 15 Porovnání výkonnosti algoritmů pro sadu prediktorů Sentinel-1 + Sentinel-2 při prostorově vnořené křížové validaci.

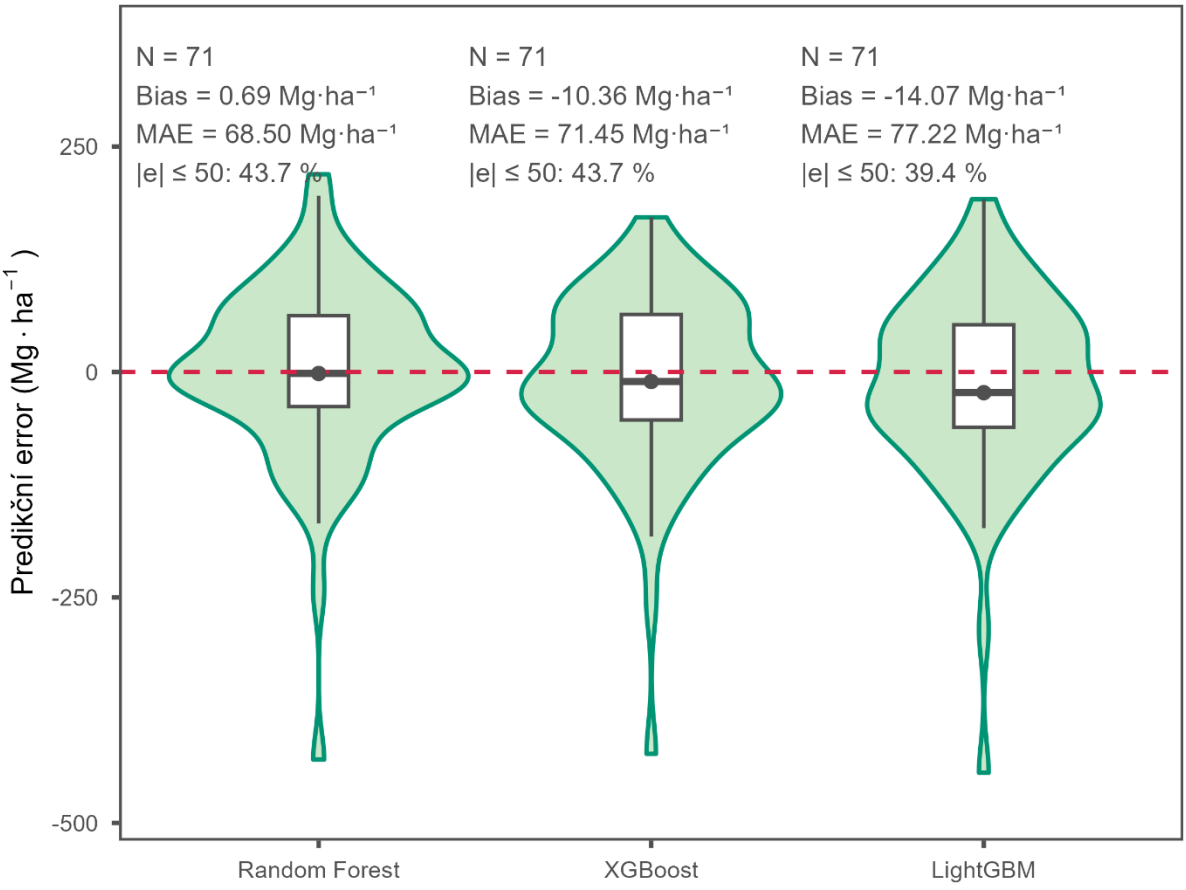
Model S1+S2	R^2 (průměr ± SD)	RMSE (průměr ± SD)	MAE (průměr ± SD)
Random Forest	0,69 ± 0,12	79,3 ± 16,1	48,0 ± 10,2
XGBoost	0,70 ± 0,14	77,3 ± 18,3	50,7 ± 10,2
LightGBM	0,66 ± 0,11	83,4 ± 14,9	55,8 ± 5,5

Přestože XGBoost dosáhl nejvyšší průměrné shody a nejnižší RMSE, rozdíly mezi XGBoost a Random Forest byly malé. Random Forest naopak vykazoval nižší MAE, což ukazuje na menší typickou absolutní chybu na úrovni jednotlivých referenčních ploch. LightGBM byl ve srovnání s oběma ostatními algoritmy méně přesný a dosahoval nejvyšší MAE. Výsledky tedy ukazují, že z hlediska celkové přesnosti byl nejsilnější XGBoost, zatímco Random Forest představoval velmi blízkou a stabilnější alternativu. Tento rozdíl byl dále posouzen pomocí analýzy chyb na lesním a rozšířeném referenčním souboru.

4.2.1. Porovnání na lesních plochách

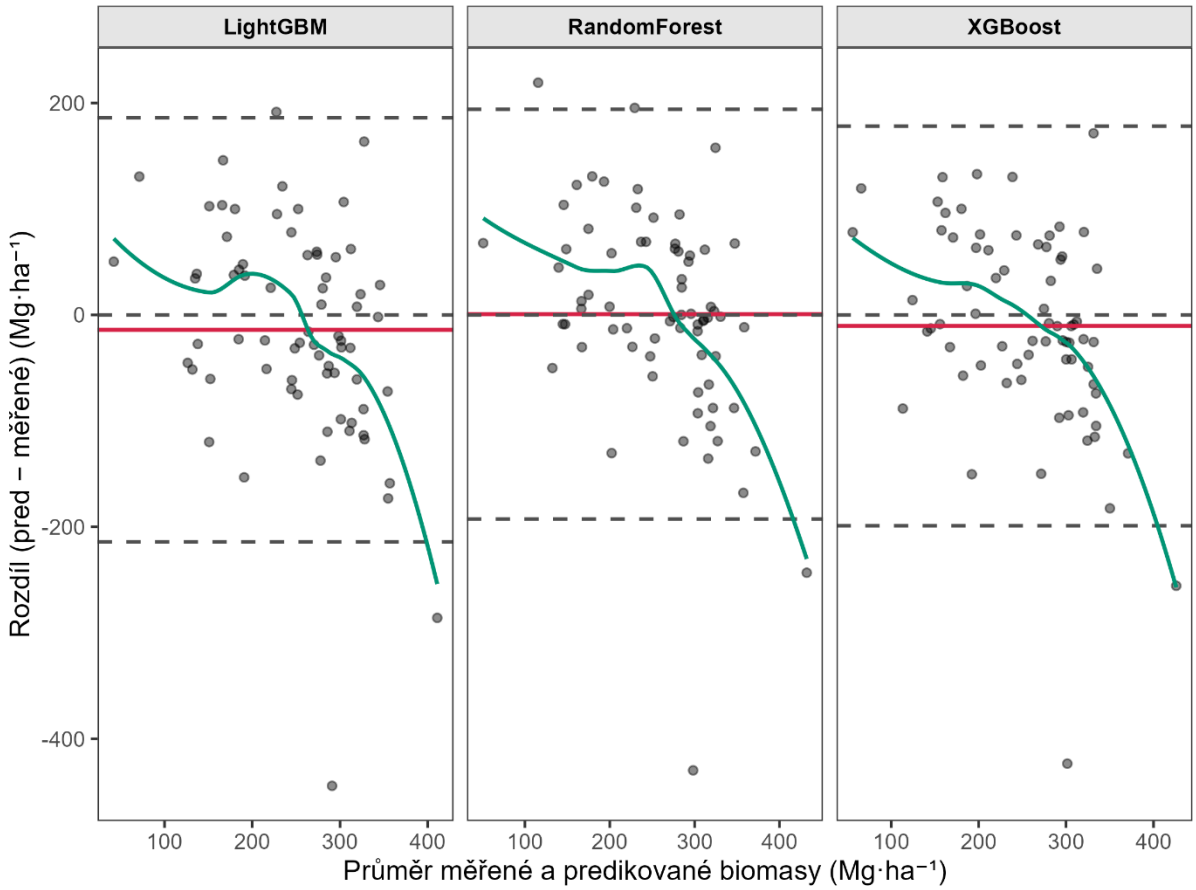
Pro doplnění agregovaných validačních metrik bylo porovnání algoritmů dále provedeno na samostatné podmnožině lesních ploch (N = 71). Tato část hodnocení je důležitá, protože přímo odpovídá hlavnímu účelu modelu, jímž je mapování nadzemní biomasy lesních porostů. Na rozdíl od rozšířeného referenčního souboru zde nebyly zahrnuty nelesní plochy s velmi nízkou biomasou, a bylo tak možné lépe posoudit chování modelů v rámci samotného gradientu lesní biomasy. *Violin* grafy predikčních chyb ukazují, že rozdělení chyb se mezi modely lišilo nejen velikostí typické chyby, ale také polohou nejvyšší hustoty hodnot. U Random Forest byla největší šířka rozdělení soustředěna v okolí nulového predikční chyby, zatímco u XGBoost a LightGBM se nejširší část rozdělení posouvala do záporných hodnot (Obrázek 15). Nejnižší *bias* tak měl Random Forest (0,69 Mg·ha^{-1}), zatímco XGBoost a LightGBM vykazovaly záporný *bias* −10,36 Mg·ha^{-1} a −14,07 Mg·ha^{-1}, tedy průměrnou tendenci k podhodnocování hodnot lesní biomasy. Random Forest zároveň dosáhl nejnižší MAE (68,50 Mg·ha^{-1}), následovaný XGBoostem (71,45 Mg·ha^{-1}) a LightGBM (77,22 Mg·ha^{-1}). Podíl predikcí s absolutní chybou do 50 Mg·ha^{-1} byl u RF

a XGBoost shodný (43,7 %), zatímco u LightGBM byl nižší (39,4 %). Z pohledu typické chyby a celkového *biasu* tak Random Forest vykazoval na lesních plochách nejvyváženější chování.



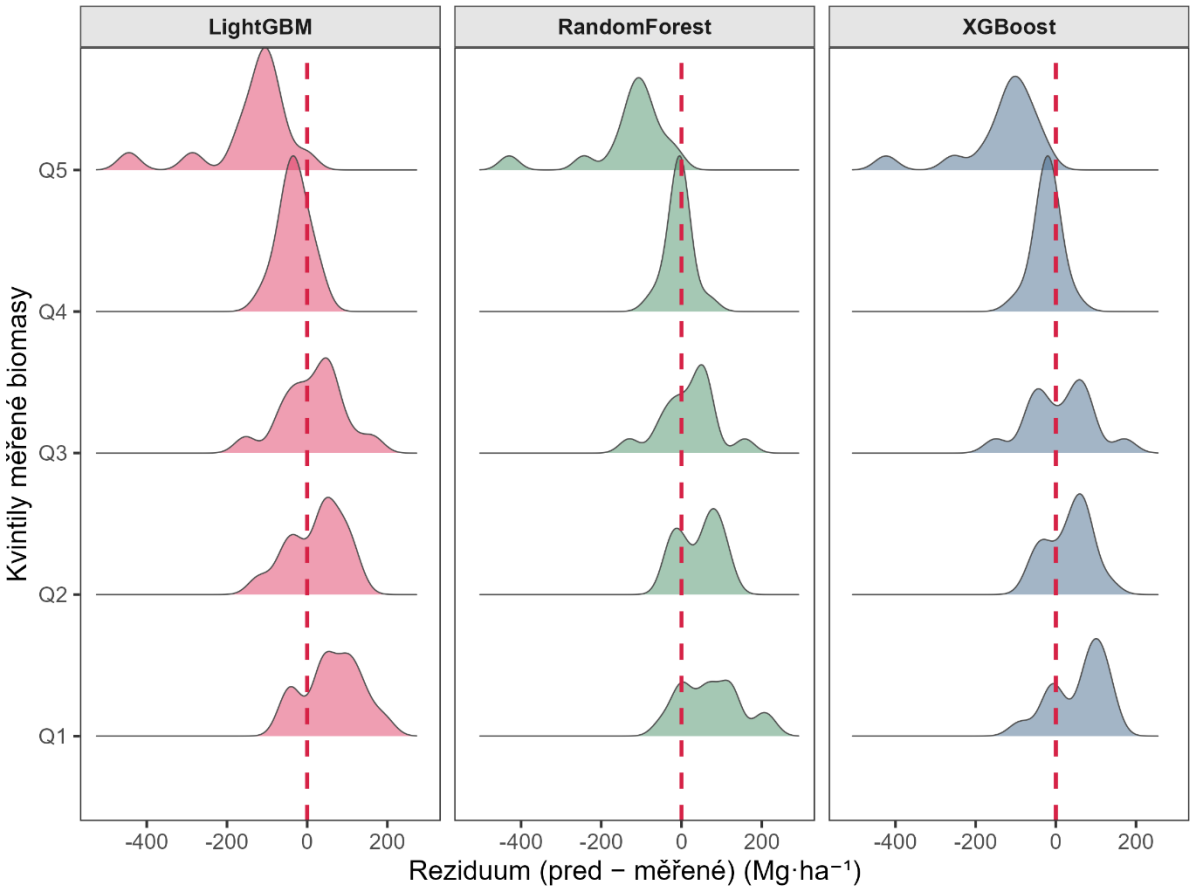
Obrázek 15 Srovnání predikčních chyb pomocí *violin plotu*

Bland–Altman grafy ukazují, že chyba modelů nebyla konstantní napříč celým rozsahem biomasy. Plná červená čára vyjadřuje průměrný rozdíl mezi predikovanou a měřenou hodnotou, tedy celkový *bias* modelu. U Random Forest byla tato hodnota nejblíže nulové linii, což potvrzuje nejmenší systematickou odchylku na lesních plochách. U všech tří algoritmů byl současně patrný podobný trend: při nižších hodnotách biomasy se častěji objevovalo nadhodnocování, zatímco při vyšších hodnotách biomasy se modely posouvaly do záporných predikčních chyb, tedy k podhodnocování. Nejvyrovnanější chování bylo patrné přibližně ve střední části rozsahu, zhruba kolem 250–300 Mg·ha^{-1}, kde se chyba nejvíce přibližovala nulové hodnotě. Nad touto úrovní již predikce všech modelů začínaly výrazněji zaostávat za měřenými hodnotami, což naznačuje omezenou citlivost použitých radarových a optických Sentinel prediktorů při vysokých hodnotách biomasy.



Obrázek 16 Bland–Altman graf pro porovnání tří modelů na lesních plochách

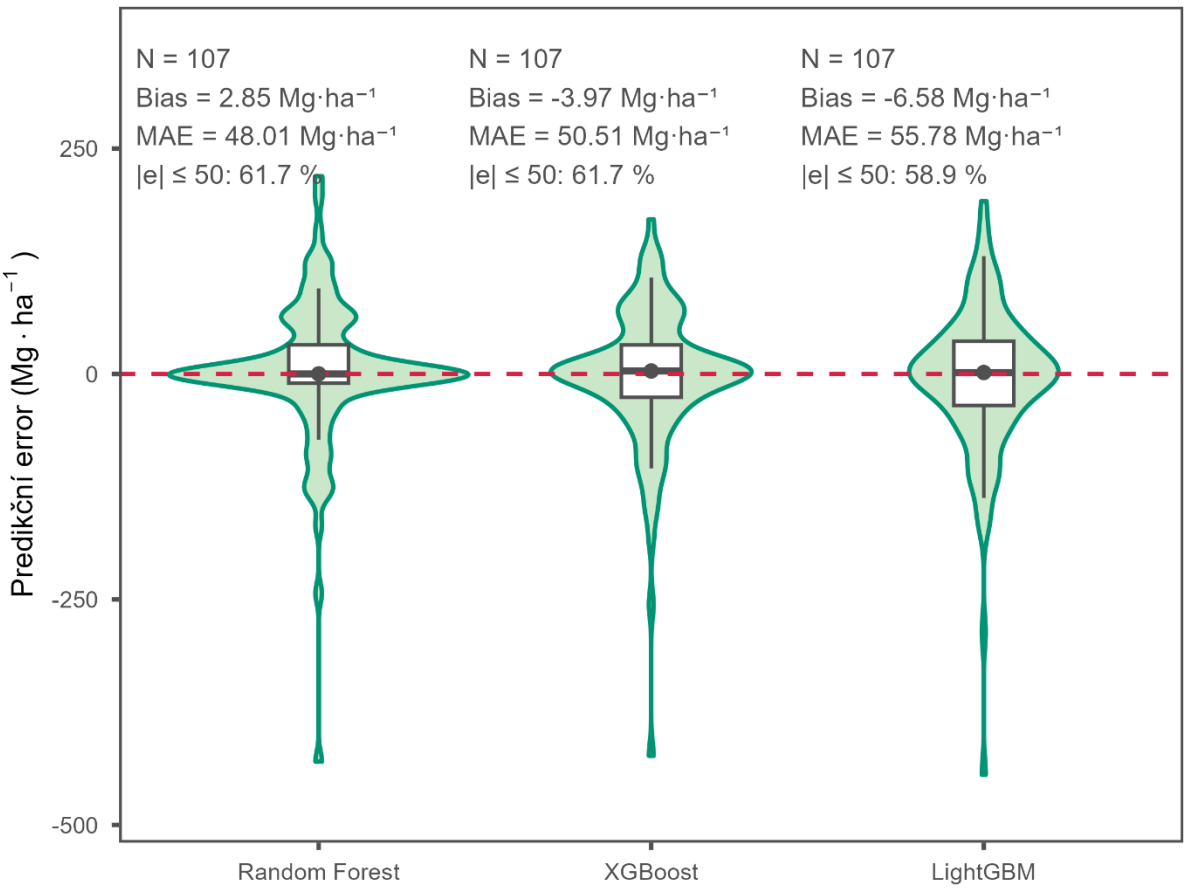
Kvintilové rozdělení reziduí umožňuje přesněji rozlišit, v jakém rozsahu hodnot lesní biomasy jednotlivé modely chybovaly nejvíce. V nejnižších kvintilech (Q1–Q2) byly reziduální distribuce u všech modelů posunuty převážně do kladných hodnot, což znamená nadhodnocování ploch s nižší měřenou biomasou. Tento posun byl mezi modely poměrně podobný, nicméně u Random Forest se rozdělení od středních kvintilů jevilo mírně blíže nulové hodnotě než u XGBoost a LightGBM. Ve středním rozsahu hodnot biomasy (zejména Q3–Q4) vykazoval Random Forest nejstabilnější chování, protože rezidua byla soustředěna nejblíže nulové linii a rozdělení nebylo výrazně posunuto ani do kladných, ani do záporných hodnot. XGBoost a LightGBM si v těchto kvintilech zachovávaly mírnější systematický posun, což ukazuje na méně vyrovnané chování při středních hodnotách biomasy. Největší rozdíly se projevily v nejvyšším kvintilu (Q5), kde se distribuce reziduí u všech modelů výrazně posunula do záporných hodnot. To potvrzuje systematické podhodnocování porostů s nejvyšší biomasou a zároveň ukazuje, že hlavní omezení modelů neleží ve středních hodnotách, ale především při odhadu vysokých hodnot AGB. Celkově tedy kvintilové rozdělení podporuje interpretaci, že Random Forest byl nejvyrovnanější zejména ve středních hodnotách lesní biomasy, zatímco všechny modely měly omezenou schopnost zachytit nejvyšší hodnoty biomasy.



Obrázek 17 Kvintily a reziduum pro měřené modely

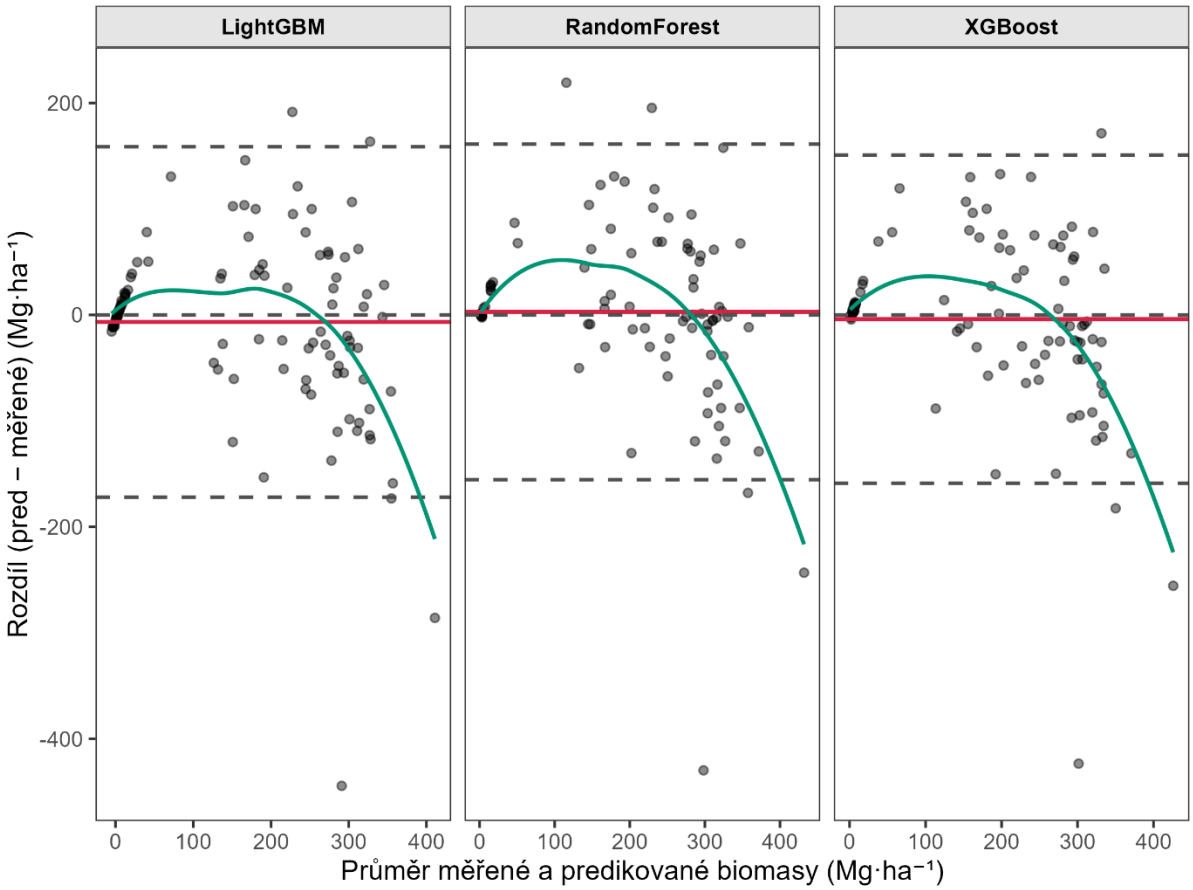
4.2.2. Porovnání na rozšířeném referenčním souboru

Druhá část porovnání algoritmů byla provedena na rozšířeném referenčním souboru (N = 107), který kromě lesních ploch zahrnoval také plochy s velmi nízkou biomasou, zejména travní a otevřenější nelesní stanoviště. Smyslem tohoto rozšířeného souboru bylo lépe ukotvit spodní část rozsahu hodnot AGB a omezit tendenci modelu nadhodnocovat biomasu v oblastech s řídkou vegetací, světlinami a čerstvě disturbovanými plochami. *Violin* grafy ukazují, že po zahrnutí nízkobiomasových ploch se celkové rozdělení chyb ve srovnání s čistě lesním souborem zúžilo a zvýšil se podíl predikcí s absolutní chybou do 50 Mg·ha^{-1}. U Random Forest i XGBoost tento podíl dosáhl 61,7 %, zatímco u LightGBM 58,9 %. Random Forest měl nejnižší MAE (48,01 Mg·ha^{-1}) a nízký *bias* 2,85 Mg·ha^{-1}. XGBoost dosáhl mírně vyšší celkové shody, ale měl vyšší MAE (50,51 Mg·ha^{-1}) a záporný *bias* −3,97 Mg·ha^{-1}. LightGBM zůstal nejslabší variantou s MAE 55,78 Mg·ha^{-1} a *biasem* −6,58 Mg·ha^{-1}. Rozdíly mezi RF a XGBoost však byly malé a oba modely vykazovaly velmi podobné celkové chování. Hlavní rozdíl spočíval v tom, že RF byl z hlediska *biasu* a typické absolutní chyby mírně vyrovnanější.



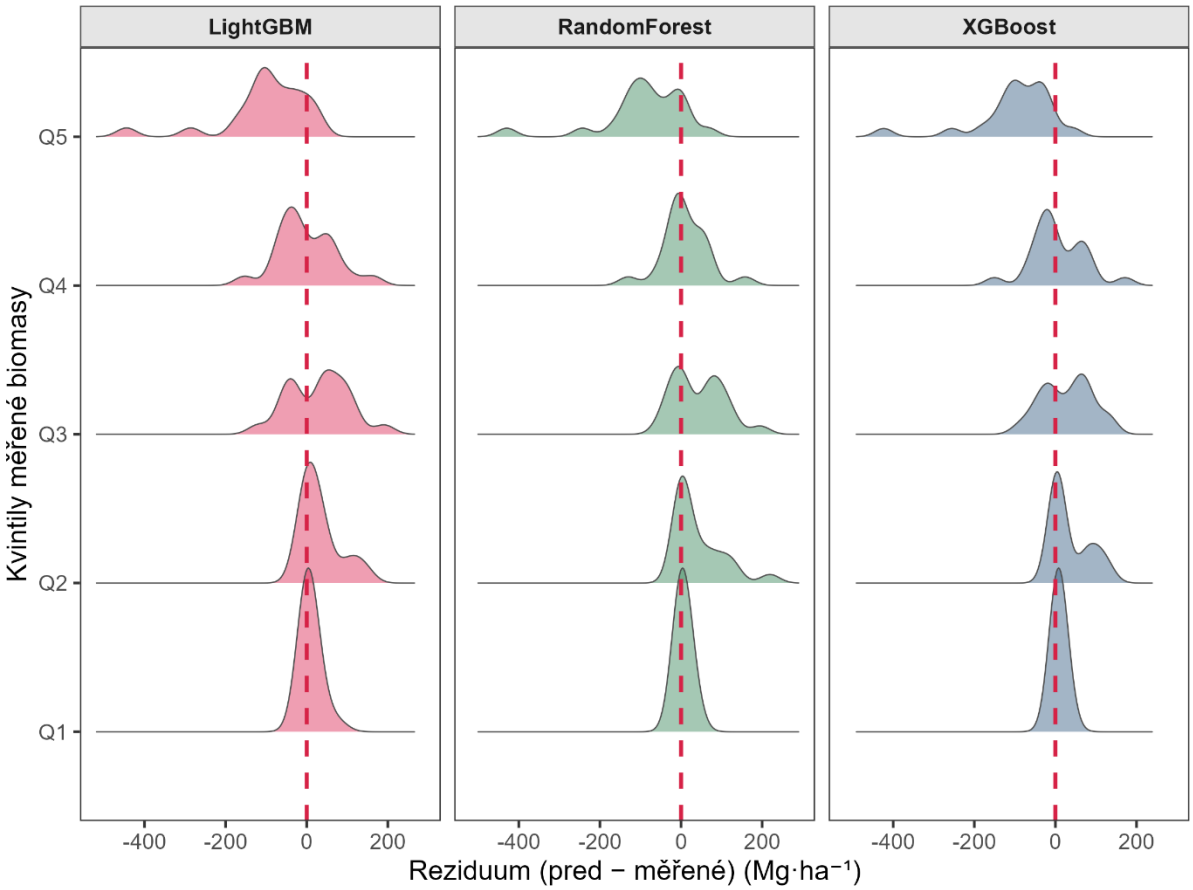
Obrázek 18 Srovnání predikčních chyb jednotlivých modelů na rozšířeném referenčním souboru pomocí *violin* grafu

Bland–Altman grafy pro rozšířený soubor ukazují, že zařazení nízkobiomasových ploch zlepšilo ukotvení modelů v dolní části rozsahu hodnot biomasy. Plná červená čára opět vyjadřuje celkový *bias* modelu. U Random Forest byla tato hodnota nejblíže nulové linii, zatímco XGBoost a LightGBM vykazovaly mírný posun do záporných hodnot. V oblasti velmi nízké biomasy byly chyby u všech modelů soustředěny blízko nuly, což naznačuje, že rozšířený referenční soubor pomohl stabilizovat predikce nízkých hodnot AGB. Přesto zůstal zachován celkový trend patrný již u lesního souboru: s rostoucí biomasou se modely postupně posouvaly k záporným chybám, tedy k podhodnocování. Nejvyrovnanější chování bylo opět patrné přibližně ve středním rozsahu hodnot biomasy, zatímco při vysokých hodnotách AGB se chyba zvětšovala.



Obrázek 19 Bland–Altman graf pro porovnání tří modelů na rozšířeném referenčním souboru

Kvintilové rozdělení reziduí dále ukazuje, že v nejnižším kvintilu byly chyby u všech modelů koncentrovány velmi blízko nulové hodnotě. To potvrzuje přínos nízkobiomasových referenčních ploch, které pomohly omezit nadhodnocování spodní části rozsahu AGB. Ve středních kvintilech byly rozdíly mezi modely poměrně malé, přičemž Random Forest se opět jevil jako mírně vyrovnanější, protože jeho rezidua byla soustředěna blíže nulové linii. Ve vyšších kvintilech se rozptyl reziduí zvětšoval a v nejvyšším kvintilu docházelo u všech modelů k posunu do záporných hodnot. Tento výsledek potvrzuje, že přidání nízkobiomasových ploch zlepšilo chování modelů při nízké biomase, ale neodstranilo hlavní omezení při odhadu nejvyšších hodnot AGB. Celkově tedy rozšířený referenční soubor vedl k lepšímu ukotvení spodní části rozsahu biomasy, zatímco relativní rozdíly mezi modely zůstaly podobné jako u čistě lesních ploch.



Obrázek 20 Rozdělení reziduí podle kvintilů měřené biomasy na rozšířeném referenčním souboru

4.2.3. Vztah mezi měřenou a predikovanou biomasou

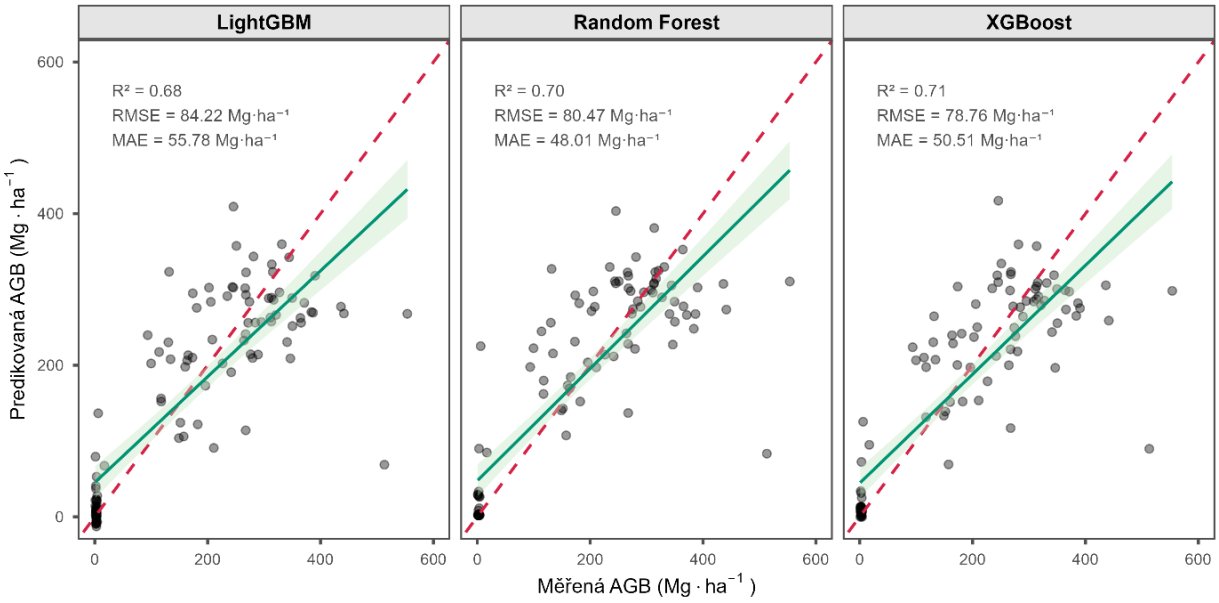
Pro doplnění validačních metrik bylo provedeno také vizuální porovnání vztahu mezi měřenou a predikovanou biomasou pro všechny tři testované algoritmy. *Scatterploty* vycházejí z *pooled out-of-fold* predikcí, tedy z predikcí spojených ze všech vnějších testovacích foldů. Každá referenční plocha tak byla porovnána s predikcí modelu, který nebyl trénován na daném pozorování. Tento způsob zobrazení umožňuje posoudit nejen celkový soulad mezi měřenými a modelovanými hodnotami, ale také případné systematické odchylky v různých částech rozsahu hodnot AGB.

Tabulka 16 *Pooled out-of-fold* metriky modelů pro sadu prediktorů Sentinel-1 + Sentinel-2

Model S1+S2	R^2	RMSE (Mg·ha^{-1})	MAE (Mg·ha^{-1})
Random Forest	0,70	80,47	48,01
XGBoost	0,71	78,76	50,51
LightGBM	0,68	84,22	55,78

U všech tří modelů byl patrný zřetelný pozitivní vztah mezi měřenou a predikovanou biomasou, přičemž celkové pořadí modelů odpovídalo dříve uvedeným OOF metrikám. *Scatterploty* proto slouží především jako vizuální doplnění předchozího hodnocení a ukazují, že RF a XGBoost měly velmi podobný vztah mezi měřenými a predikovanými hodnotami, zatímco LightGBM vykazoval

širší rozptyl chyb. Ve všech případech se nejvyrovnanější predikce soustředily přibližně ve středním rozsahu hodnot biomasy, zejména kolem 250–300 Mg·ha^{-1}. Při velmi nízké biomase zůstávaly predikce většinou blízko nulovým hodnotám, i když se objevovaly jednotlivé případy nadhodnocení. Naopak při hodnotách nad přibližně 350 Mg·ha^{-1} docházelo k častějšímu podhodnocování, kdy část bodů ležela pod linií 1:1. Světle zelený 95% interval spolehlivosti kolem regresní linie vyjadřuje nejistotu odhadu průměrného trendu mezi měřenou a predikovanou biomasou; nejde tedy o interval očekávané chyby jednotlivých bodů, ale o míru nejistoty samotné regresní linie. Jeho rozšíření při vyšších hodnotách AGB ukazuje, že vztah mezi měřenou a predikovanou biomasou je v této části rozsahu méně jistý, což souvisí s menším počtem referenčních ploch s velmi vysokou biomasou a větším rozptylem predikcí. Malý počet výraznějších odlehlých hodnot pravděpodobně souvisí s lokálním nesouladem mezi referenčními plochami a prediktory dálkového průzkumu Země. Celkově *scatterploty* potvrzují závěry předchozí reziduální analýzy: RF a XGBoost poskytovaly velmi podobnou celkovou shodu, zatímco LightGBM byl méně přesný, a hlavní omezení všech modelů se projevovalo především při odhadu nejvyšších hodnot AGB.



Obrázek 21 Vztah mezi měřenou a predikovanou biomasou pro modely Random Forest, XGBoost a LightGBM

4.2.4. Význam vstupních proměnných

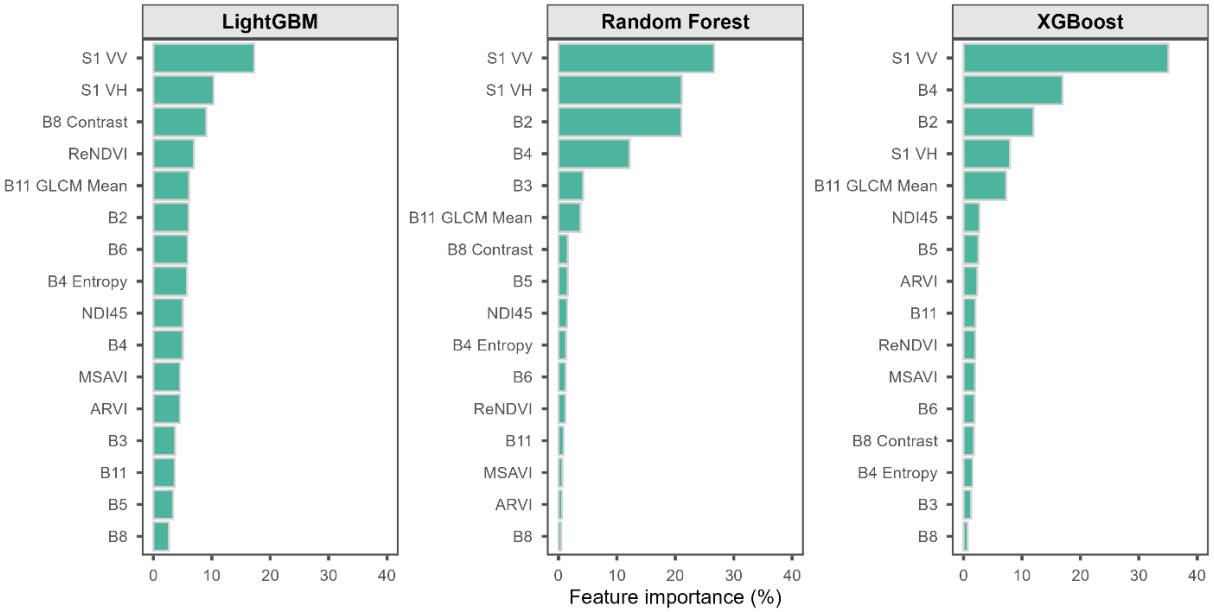
Kromě přesnostních metrik a reziduální diagnostiky byl porovnán také význam vstupních proměnných u jednotlivých algoritmů. Toto srovnání neposkytuje přímou informaci o přesnosti modelu, ale umožňuje posoudit, do jaké míry jsou jednotlivé modely závislé na konkrétních prediktorech a zda využívají radarovou i optickou složku datové sady S1+S2 vyváženě. Z hlediska provozní použitelnosti je důležité, aby model nebyl nadměrně závislý pouze na jediném vstupním kanálu, protože taková závislost může snižovat robustnost při změně podmínek snímání nebo při lokálních anomáliích ve vstupních datech. Vzhledem k tomu, že LightGBM poskytuje významnost proměnných v jiné absolutní škále než RF a XGBoost, byly hodnoty *feature importance* pro účely

společného grafického porovnání normalizovány v rámci každého modelu na procentuální podíl z celkové důležitosti. V případě LightGBM představovala původní absolutní hodnota počet použití daného prediktoru při větvení rozhodovacích stromů (split importance). Nejde tedy přímo o procentuální podíl, ale o počet splitů, ve kterých byla daná proměnná využita. Pro společný graf byly tyto hodnoty převedeny na procenta vzhledem k celkovému součtu všech splitů. Graf proto neporovnává absolutní hodnoty importance mezi algoritmy, ale relativní rozložení významu prediktorů uvnitř každého modelu.

Random Forest vykazoval relativně vyvážené využití radarových i optických prediktorů. Nejvyšší význam měly proměnné S1 VV (26,6 %) a S1 VH (21,1 %), avšak velmi podobně významné bylo také optické pásmo B2 (21,0 %) a dále B4 (12,2 %). Tyto čtyři proměnné dohromady tvořily přibližně 81 % celkové důležitosti modelu. Ostatní prediktory měly výrazně nižší podíl, nejvýše B3 (4,2 %) a B11 GLCM Mean (3,8 %), zatímco většina vegetačních indexů a dalších pásem se pohybovala přibližně pod 1,5 %.

U XGBoost byla významnost proměnných více koncentrována do menšího počtu vstupů. Nejvýznamnějším prediktorem byl S1 VV (35,1 %), následovaný pásmy B4 (17,0 %) a B2 (12,0 %). S1 VH dosáhl 7,9 % a B11 GLCM Mean 7,3 %. Prvních pět prediktorů tak tvořilo přibližně 78 % celkové důležitosti modelu. Ostatní proměnné, včetně vegetačních indexů a zbývajících spektrálních pásem, měly převážně doplňkový význam, obvykle v rozmezí přibližně 1–3 %. XGBoost tedy dosáhl velmi dobré přesnosti, ale jeho rozhodování bylo více soustředěno na dominantní radarový prediktor S1 VV.

LightGBM měl po normalizaci rozloženější profil významnosti než RF a XGBoost. Nejvyšší podíl měl opět S1 VV (17,3 %; absolutně 246 splitů), následovaný S1 VH (10,3 %; 147 splitů), B8 Contrast (9,1 %; 129 splitů), ReNDVI (7,0 %; 99 splitů), B11 GLCM Mean (6,1 %; 87 splitů) a B2 (6,0 %; 86 splitů). Ve srovnání s RF a XGBoost tak LightGBM více využíval také texturní proměnné a vegetační indexy. Nejnižší relativní význam měly zejména B8 (2,7 %; 38 splitů), B5 (3,4 %; 48 splitů), B11 (3,6 %; 52 splitů) a B3 (3,7 %; 53 splitů). Přestože tento profil působí vyváženěji, v kombinaci s nižší přesností a vyšší MAE nepředstavoval LightGBM vhodnější alternativu pro finální mapování. Vyvážený profil významnosti u RF proto nepředstavoval hlavní, ale podpůrný argument pro jeho výběr jako finálního modelu.



Obrázek 22 Význam vstupních proměnných pro modely Random Forest, XGBoost a LightGBM při použití sady prediktorů Sentinel-1 + Sentinel-2

4.2.5. Výběr finálního modelu Random Forest

Na základě kombinace validačních metrik, analýzy predikčních chyb, vztahu mezi měřenou a predikovanou biomasou, významnosti vstupních proměnných a prostorové použitelnosti výsledného rastru byl jako finální model pro stanovení nadzemní biomasy vybrán Random Forest. Přestože XGBoost dosáhl mírně vyšší celkové shody a nepatrně nižší RMSE, Random Forest vykazoval nižší MAE, tedy menší typickou absolutní chybu na úrovni ploch. Zároveň měl velmi nízký *bias* na lesních plochách a jeho výkon mezi prostorovými validačními bloky byl srovnatelně stabilní. Důležitým podpůrným argumentem byl také vyváženější profil významnosti proměnných. RF kombinoval informaci z radarových i optických prediktorů, zatímco XGBoost byl výrazněji závislý na jednom dominantním radarovém prediktoru. Dalším praktickým důvodem bylo, že XGBoost v celoplošném rastrovém výstupu produkoval záporné hodnoty AGB, které musely být následně reklasifikovány na nulu. Random Forest tak představoval nejvhodnější kompromis mezi přesností, stabilitou, interpretovatelností a fyzikální věrohodností výsledné vrstvy. Následující srovnání s externími daty je proto provedeno již pouze pro finální model Random Forest.

4.3. Srovnání finálního modelu s externími daty

4.3.1. Srovnání s ESA CCI Biomass

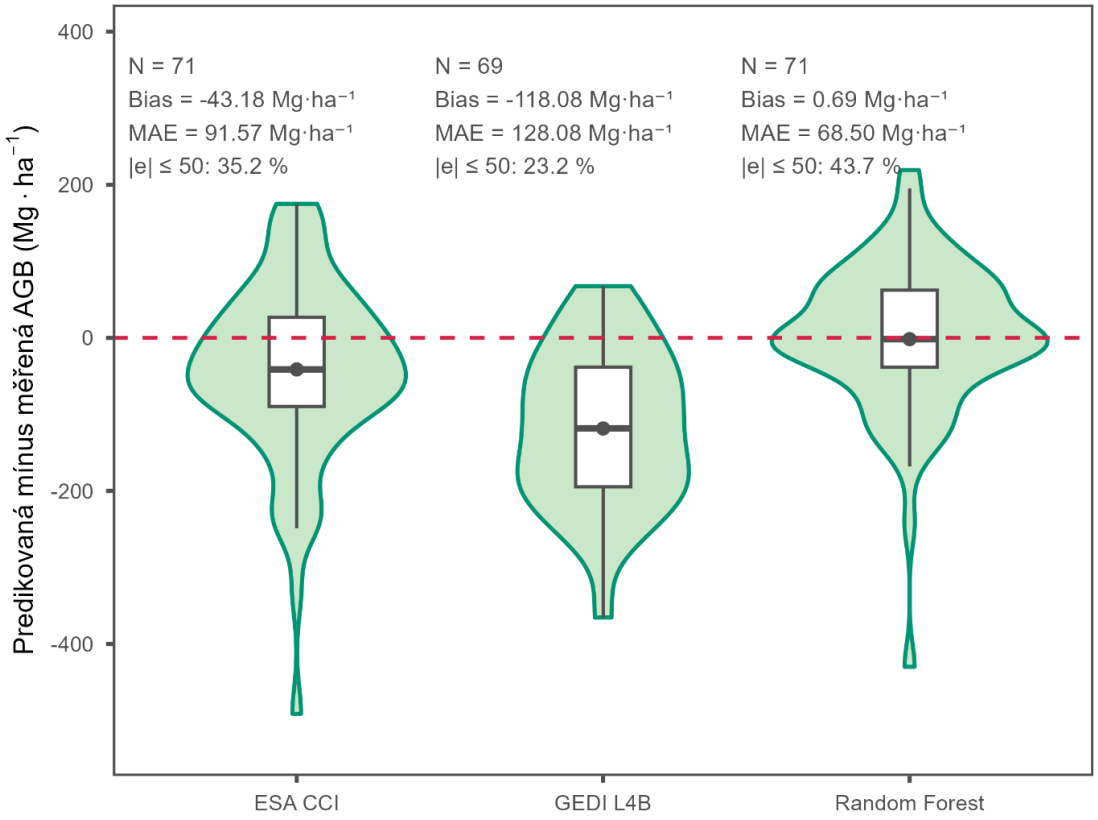
Výsledná vrstva Random Forest byla porovnána s globálními produkty ESA CCI Biomass v6 a GEDI L4B. RF vrstva dosáhla průměrné hodnoty AGB 213,63 $Mg \cdot ha^{-1}$, zatímco ESA CCI Biomass v6 dosáhla 178,72 $Mg \cdot ha^{-1}$ a GEDI L4B 125,47 $Mg \cdot ha^{-1}$. RF vrstva měla zároveň vyšší medián i horní kvartil, což ukazuje, že lépe reprezentovala porosty se střední až vyšší biomasou. Globální

produkty naopak vykazovaly nižší centrální hodnoty a zejména GEDI L4B měl výrazně užší rozptyl hodnot.

Tabulka 17 Statistiky odhadů AGB z RF modelu a globálních produktů biomasy

	1. kvartil	Medián	3. kvartil	Průměr	Směrodatná odchylka
ESA CCI v6	106	182	255	178,72	93,38
GEDI L4B	81,08	124,57	164,60	125,47	55,98
Random Forest	175,12	226,29	274,94	213,63	76,71

Rozdíly mezi produkty lze vysvětlit především odlišným prostorovým rozlišením, časovým obdobím a účelem jednotlivých datových sad. RF vrstva reprezentuje stav vegetační sezóny 2023 v rozlišení 20 m, zatímco ESA CCI Biomass v6 reprezentuje rok 2022 v rozlišení 100 m a GEDI L4B poskytuje gridované odhady v rozlišení 1 km založené na pozorováních z období 2019–2021. Globální produkty tedy nemohou zachytit jemnější prostorovou strukturu českých lesů v takovém detailu jako model přizpůsobený podmínkám České republiky. Porovnání na lesních referenčních plochách ukázalo, že RF byl ze srovnávaných produktů nejbližší terénním hodnotám. RF měl *bias* 0,69 Mg·ha^{-1} a MAE 68,50 Mg·ha^{-1}. ESA CCI vykazovala výraznější podhodnocení s *biasem* −43,18 Mg·ha^{-1} a MAE 91,57 Mg·ha^{-1}. GEDI L4B podhodnocoval nejsilněji, s *biasem* −118,08 Mg·ha^{-1} a MAE 128,08 Mg·ha^{-1}. Tento výsledek potvrzuje, že globální produkty jsou užitečné pro širší kontext, ale pro národní mapování biomasy v českých lesích jsou příliš hrubé a mají tendenci podhodnocovat vyšší hodnoty AGB.



Obrázek 23 Porovnání rozdílů mezi RF modelem a globálními produkty ESA CCI Biomass v6 a GEDI L4B na lesních referenčních plochách

4.3.2. Srovnání s daty národní inventarizace lesů (NIL)

Dalším referenčním rámcem bylo srovnání s Národní inventarizací lesů České republiky (NIL), která poskytuje národní odhady nadzemní biomasy a zásob uhlíku. Podle výsledků třetího inventarizačního cyklu činila průměrná AGB českých lesů 219,4 ± 2,5 Mg·ha^{-1} a celková zásoba AGB 625,1 ± 10,9 milionu Mg. Finální RF model odhadl průměrnou hodnotu AGB 213,63 Mg·ha^{-1}, což je hodnota velmi blízká národní inventarizační referenci. Přestože je odhad RF mírně nižší, rozdíl v průměrné hektarové hodnotě je relativně malý a výsledek ukazuje, že model přizpůsobený podmínkám České republiky poskytuje odhad ve srovnatelném rozsahu jako NIL. Současně je nutné zohlednit časový nesoulad, protože třetí cyklus NIL reprezentuje období 2016–2020, zatímco výsledná RF vrstva odpovídá vegetační sezóně roku 2023.

Výraznější rozdíl se projevil u celkové zásoby AGB. RF model odhadl celkovou zásobu AGB na 580,6 milionu Mg, zatímco NIL uvádí 625,1 ± 10,9 milionu Mg. Tento rozdíl nelze interpretovat pouze jako chybu modelu, protože je významně ovlivněn také rozdílným vymezením lesní plochy. Česká NIL reportovala celkovou lesní plochu 2 892,9 ± 53 tis. ha v roce 2015 a 2 923 ± 37,6 tis. ha v roce 2020. Naproti tomu v této práci byla modelována lesní vrstva o rozsahu přibližně 2 717 tis. ha. Již samotný rozdíl mezi lesní plochou použitou v NIL a rozsahem modelované vrstvy tedy významně ovlivňuje celkový odhad zásoby biomasy. Jinými slovy, i při velmi podobné průměrné hektarové hodnotě může být celková zásoba AGB nižší, pokud je výpočet aplikován na menší lesní plochu. Rozdíly v lesní ploše nejsou specifické pouze pro tuto práci, ale souvisejí obecně s odlišnými definicemi lesa, metodikou mapování a účelem jednotlivých datových zdrojů. Pro srovnání, FAO uvádí pro Českou republiku průměrnou AGB 163,23 Mg·ha^{-1} a lesní plochu 2 677,09 tis. ha pro rok 2015. Tato nižší hodnota pravděpodobně odráží rozdíly v reportovacím rámci, definici lesa a harmonizované mezinárodní metodice, která je určena především pro mezistátní srovnatelnost. Podobně nedávná studie z České republiky, která použila metodu RF pro klasifikaci lesa, dosáhla celkové přesnosti 98,3 % a reportovala lesní plochu 2 681,47 tis. ha (Kovárník a Janová, 2025). Tyto hodnoty jsou bližší modelované lesní vrstvě použité v této práci než ploše uváděné NIL, což ukazuje na systematický rozdíl mezi odhady lesní plochy získanými metodami dálkového průzkumu Země a pozemními inventarizačními metodami.

Při převodu biomasy na uhlík pomocí konverzního faktoru 0,50 odpovídal RF odhad přibližně 106,82 Mg·ha^{-1} C v AGB, zatímco NIL uvádí 110,3 ± 1,3 Mg·ha^{-1}. Také zde je hektarová hodnota RF velmi blízká národní inventarizační hodnotě. Z hlediska celkové zásoby uhlíku však RF odhadl přibližně 290,3 milionu Mg C, zatímco NIL uvádí 314,3 ± 5,5 milionu Mg C. Podobně jako u celkové zásoby AGB je tento rozdíl ovlivněn nejen částečným podhodnocováním vysokých hodnot biomasy, ale také menší rozlohou modelované lesní vrstvy.

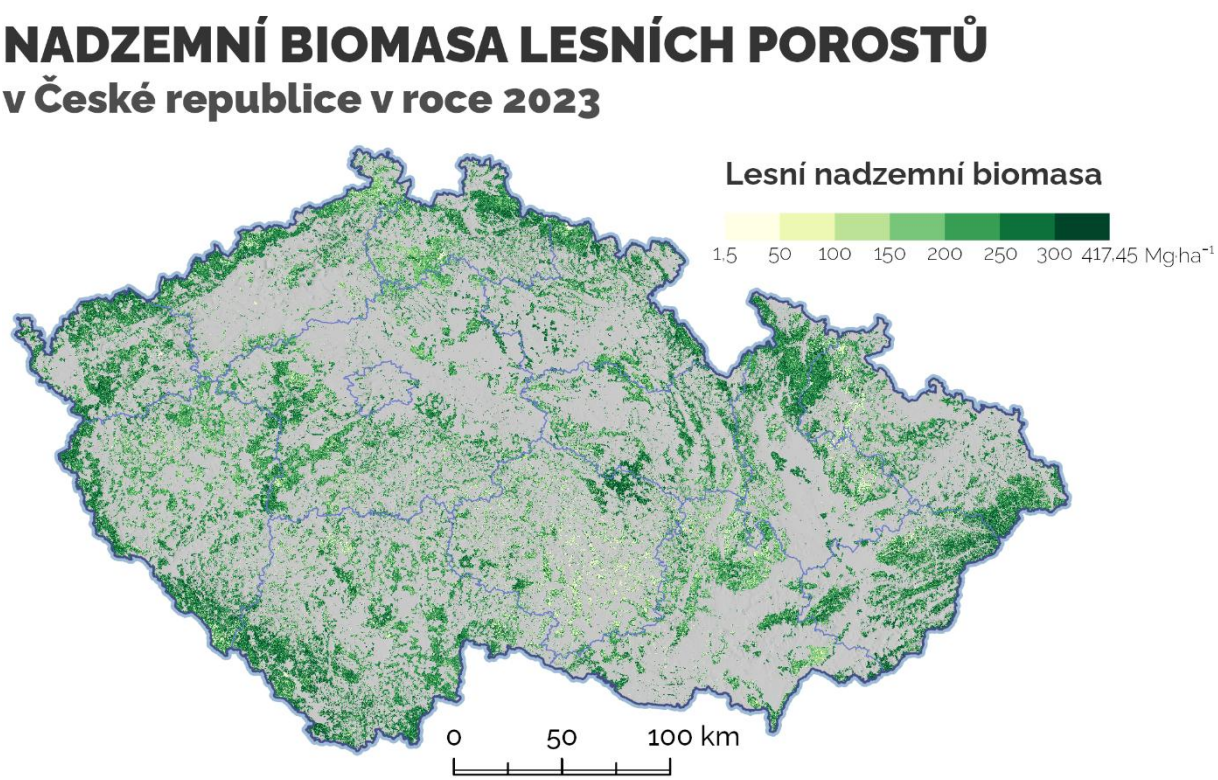
Při interpretaci uvedených hodnot je nutné zohlednit také rozdílný způsob hodnocení nejistoty obou přístupů. Intervalové odhady uváděné NIL vycházejí z výběrového rámce národní inventarizace a vyjadřují statistickou nejistotu odhadu příslušných populačních charakteristik lesa založeného na rozsáhlém souboru inventarizačních ploch. Přesnost těchto odhadů je ovlivněna variabilitou sledované veličiny, velikostí a uspořádáním výběrového souboru a použitým výběrovým designem (Cochran, 1977; McRoberts, 2010). V případě EO (Earth

Observation) modelu nebyl v této práci odvozován přímo srovnatelný interval spolehlivosti národního průměru AGB nebo uhlíku v AGB. Predikční schopnost a velikost predikčních chyb modelu byly namísto toho hodnoceny prostřednictvím prostorově vnořené křížové validace a out-of-fold predikcí pomocí metrik R^2, RMSE, MAE a biasu. Intervaly uváděné NIL a validační charakteristiky RF modelu tedy představují metodicky odlišné způsoby vyjádření nejistoty a nelze je přímo číselně porovnávat.

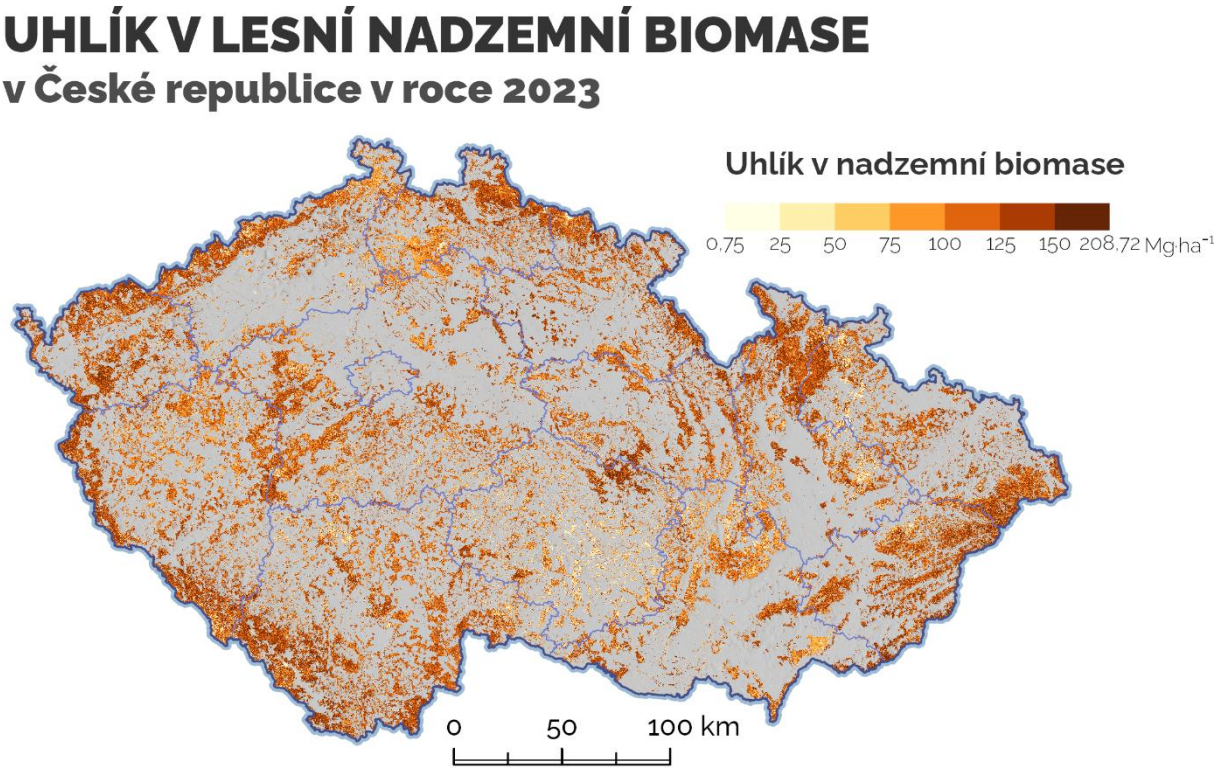
Cílem vytvořené vrstvy proto není přesně reprodukovat inventarizační statistiku NIL, ale vytvořit prostorově spojitý národní produkt AGB, který je s národní referencí konzistentní a zároveň umožňuje prostorovou analýzu rozdílů uvnitř území České republiky. Z tohoto hlediska lze RF vrstvu považovat za vhodný doplněk k NIL: inventarizace poskytuje statisticky robustní národní odhady, zatímco model dálkového průzkumu Země umožňuje detailní prostorové zobrazení a interpretaci regionálních rozdílů v zásobě nadzemní biomasy.

Tabulka 18 Srovnání průměrné AGB a uhlíku v AGB mezi modely a Národní inventarizací lesů

Metoda	Průměr C v AGB (Mg·ha^{-1})	Průměrná AGB (Mg·ha^{-1})
Random Forest	106,82	213,63
XGBoost	103,64	207,28
LightGBM	110,38	220,76
NIL (3. cyklus)	110,3 ± 1,3	219,4 ± 2,5
FAO (2020)	76,72	163,23



Obrázek 24 Nadzemní biomasa lesních porostů



Obrázek 25 Uhlík v nadzemní biomase lesních porostů

4.4. Shrnutí dosažených výsledků ve vztahu k cílům práce

Výsledková část práce poskytla jednoznačnou odpověď na základní předpoklad formulovaný v cílech práce. **Nadzemní biomasu a uhlík v nadzemní biomase lesních porostů lze na území České republiky modelovat pomocí volně dostupných dat dálkového průzkumu Země,** referenčních pozemních měření a metod strojového učení. Přesnost výsledků však byla výrazně ovlivněna volbou vstupních prediktorů, typem modelovacího algoritmu a způsobem validačního hodnocení.

Prvním hlavním výsledkem práce je vytvoření opakovatelného postupu pro přípravu a zpracování vstupních dat. V rámci práce byly datové zdroje Sentinel-1 a Sentinel-2 a vegetační indexy zpracovány pomocí skriptů v prostředí Google Earth Engine. Texturní metriky, ostatní prostorové prediktory a celkové sjednocení prediktorů do podoby pro modelování proběhlo v desktop aplikacích. Následné modelování a systematické testování jednotlivých prediktorových sad proběhlo pomocí Pythonu v prostředí ArcGIS Pro. Vytvořený postup tak umožnil opakovatelné zpracování vstupních dat, trénování modelů a porovnání vlivu jednotlivých kombinací prediktorů na přesnost odhadu nadzemní biomasy.

Druhým hlavním výsledkem je zjištění, že samotné **rozšíření počtu prediktorů nevedlo** automaticky **ke zvýšení přesnosti** odhadu nadzemní biomasy. Nejslabších výsledků dosahovala sada založená pouze na optických datech Sentinel-2, zatímco radarová data Sentinel-1 vykazovala silnou informační hodnotu zejména u algoritmu Random Forest. **Kombinace Sentinel-1 a Sentinel-2 poskytovala stabilní výsledky** napříč algoritmy, zatímco přidání morfometrických

charakteristik z DMR 4G nepřineslo zlepšení. Velmi dobrých výsledků dosahovala také kombinace všech sad prediktorů, její využití však bylo omezeno časovým nesouladem vrstvy výšky lesního porostu, a proto nebyla tato kombinace zahrnuta do dalšího finálního hodnocení.

Třetím hlavním výsledkem je výběr finálního modelovacího přístupu a vytvoření prostorových vrstev nadzemní biomasy a uhlíku v nadzemní biomase lesních porostů České republiky pro rok 2023. Jako **finální model byl vybrán Random Forest** využívající **kombinaci prediktorů Sentinel-1 a Sentinel-2.** Finální výběr nebyl založen pouze na hodnotách přesnostních metrik, ale také na prostorovém hodnocení výsledných rastrových výstupů, stabilitě modelu, významnosti vstupních proměnných a fyzikální věrohodnosti výsledné vrstvy. Přestože XGBoost dosáhl v některých metrikách mírně lepších hodnot, Random Forest vykazoval nižší typickou absolutní chybu, velmi nízký *bias* na lesních plochách, vyváženější využití radarových a optických prediktorů a neprodukoval záporné hodnoty nadzemní biomasy v celoplošném rastrovém výstupu. Na základě finálního modelu byla vytvořena prostorová vrstva nadzemní biomasy, která zachycuje rozložení biomasy v lesních porostech v rámci použité lesní masky a slouží jako podklad pro odvození zásoby uhlíku v nadzemní biomase. Finální model Random Forest odhadl **průměrnou nadzemní biomasu lesních porostů** na **213,63 $\mathbf{Mg \cdot ha^{-1}}$** a celkovou **zásobu nadzemní biomasy na 580,6 milionu Mg**. Z těchto hodnot byla odvozena **průměrná zásoba uhlíku v nadzemní biomase přibližně 106,8 $\mathbf{Mg \cdot ha^{-1}}$** a celková **zásoba uhlíku přibližně 290,3 milionu Mg C.**

Čtvrtým hlavním výsledkem je srovnání vytvořené vrstvy s nezávislými národními a globálními datovými sadami. Finální vrstva Random Forest dosáhla vyšší průměrné hodnoty nadzemní biomasy než globální produkty ESA CCI Biomass v6 a GEDI L4B. Ve srovnání s Národní inventarizací lesů České republiky byla **průměrná hektarová hodnota nadzemní biomasy velmi blízká inventarizační referenci**. **Výraznější rozdíl se projevil** především **u celkové zásoby nadzemní biomasy a uhlíku**, což souviselo zejména s **odlišným vymezením lesní plochy, rozdílným prostorovým rozlišením, použitou metodikou a rozsahem referenčních dat**.

5. DISKUZE

Cílem této disertační práce bylo ověřit, zda lze nadzemní biomasu lesních porostů a z ní odvozený uhlík v nadzemní biomase modelovat v národním měřítku pomocí volně dostupných dat dálkového průzkumu Země. Výsledky naznačují, že tento cíl se podařilo naplnit. Kombinace družicových dat Sentinel-1 a Sentinel-2 umožnila vytvořit model nadzemní biomasy pro území České republiky s uspokojivou predikční schopností i při použití prostorově založené validace. Práce tak potvrzuje, že volně dostupná data DPZ mohou být využita nejen pro lokální nebo regionální odhady, ale také pro tvorbu národního produktu AGB v podmínkách mírného pásma středoevropských lesů. Toto zjištění je však nutné interpretovat s ohledem na několik zásadních omezení. Výsledný model nelze chápat jako bezchybný popis skutečné zásoby biomasy v každém pixelu, ale jako prostorově souvislý odhad založený na dostupných prediktorech, omezeném souboru referenčních dat a zvolené validační strategii. Hlavním přínosem práce proto není pouze samotný dataset AGB, ale také ověření chování modelu při prostorově odděleném hodnocení a identifikace hlavních limitů při převodu pozemních měření do národního odhadu.

Naplnění výzkumných cílů a význam výsledků

Práce přímo odpovídá na stanovenou výzkumnou otázku, zda je možné modelovat nadzemní biomasu lesních porostů a z ní odvozený uhlík v nadzemní biomase pomocí volně dostupných dat dálkového průzkumu Země. Výsledky ukazují, že tento přístup je v podmínkách České republiky realizovatelný a umožňuje vytvořit prostorově souvislý odhad AGB v národním měřítku. Z hlediska praktického využití výsledná vrstva AGB zachycuje hlavní prostorové rozložení biomasy lesních porostů v České republice a umožňuje odvodit také prostorové rozložení uhlíku v nadzemní biomase. Je však nutné zdůraznit, že uhlík nebyl modelován jako nezávislá veličina, ale byl odvozen z AGB pomocí konstantního přepočetního faktoru. Nejistota vrstvy uhlíku je proto přímo závislá na nejistotě modelované AGB a navíc zahrnuje zjednodušující předpoklad konstantního podílu uhlíku v biomase. Výsledky tedy potvrzují využitelnost použitých dat a metod pro tvorbu národního odhadu AGB. Současně je však nutné je interpretovat s ohledem na omezení vstupních dat a metodiky. Výsledný produkt proto představuje konzistentní prostorový odhad biomasy v národním měřítku, nikoliv náhradu detailního inventarizačního šetření na úrovni jednotlivých porostů.

Referenční data jako hlavní zdroj nejistoty

Jednou z největších limitací této práce byl rozsah a struktura dostupných referenčních dat. Pro modelování nadzemní biomasy v měřítku celé České republiky by bylo ideální mít rozsáhlý, rovnoměrně prostorově rozmístěný a stratifikovaný soubor pozemních ploch, který by pokrýval celé spektrum hodnot AGB, různé typy lesních porostů, věkové třídy, stanovištní podmínky i různé stupně disturbance. Takový soubor se však nepodařilo získat. Přestože bylo jednáno s více institucemi působícími v České republice, například s Národním lesnickým institutem, Výzkumným ústavem pro krajinu a Univerzitou Karlovou, dostupný vzorek pozemních dat zůstal omezený.

Toto omezení se projevilo zejména nerovnoměrným zastoupením jednotlivých tříd biomasy. V referenčním souboru byly relativně dobře zastoupeny střední hodnoty AGB, zatímco velmi nízké hodnoty, například holiny, raně regenerující porosty nebo řídké mladé porosty, a naopak extrémně vysoké hodnoty biomasy byly zastoupeny omezeně. Právě tato nevyváženost může významně ovlivnit chování modelu na okrajích rozsahu hodnot. Modely strojového učení jsou obecně silně závislé na rozsahu hodnot, který je obsažen v trénovacích datech. Pokud určitá část rozsahu biomasy není dostatečně zastoupena, model ji nedokáže spolehlivě naučit a má tendenci predikovat hodnoty blíže k dominantnímu rozsahu trénovacích dat. Z tohoto důvodu byla pro ukotvení spodní části gradientu biomasy použita doplňková pozemní data luční biomasy. Jejich zahrnutí představuje metodické rozhodnutí, které má opodstatnění, ale zároveň i své limity. Luční biomasa samozřejmě nepředstavuje lesní porost ve strukturálním smyslu, ale její spektrální projev může být v některých případech podobný nízké vegetaci, holinám nebo raně regenerujícím porostům, zejména při použití optických dat. V rámci disertační práce bylo testováno i modelování bez těchto doplňkových nízkobiomasových vzorků. Výsledky tohoto testu ukázaly, že jejich vynechání vedlo k výraznému nadhodnocování ploch s nízkou biomasou, zejména holin, regenerujících porostů a nízkých porostů. Zařazení těchto vzorků proto pomohlo stabilizovat spodní část predikčního rozsahu a snížit systematické nadhodnocení nízkých hodnot AGB.

Současně je však nutné přiznat, že tím do modelu vstupuje určitý kompromis. Luční vzorky byly využity jako nízkobiomasový referenční bod, nikoliv jako přímá náhrada lesních ploch. Jejich funkce byla především metodická: zabránit tomu, aby model interpretoval všechny nízké nebo narušené lesní plochy jako porosty se střední zásobou biomasy. Toto rozhodnutí je proto obhajitelné, ale mělo by být chápáno jako důsledek omezené dostupnosti vhodnějších nízkobiomasových lesních referencí.

Dalším zdrojem nejistoty je samotná nejistota pozemního měření. U lučních vzorků měřených v rámci doktorského studia byla odchylka odhadována přibližně na 5–10 % v závislosti na velikosti a charakteru plochy. Tato nejistota je vzhledem k celkovému rozsahu AGB relativně nízká. Větší význam má pravděpodobně nejistota lesních referenčních dat. U první verze dat DendroNetwork byla relativní směrodatná odchylka AGB, vyjádřená jako koeficient variability, v průměru 17,4 %, s mediánem 12,8 %. Po vyloučení dvou velmi nízkých hodnot AGB, u nichž relativní variabilita přesahovala 50 %, činila průměrná relativní odchylka přibližně 14,0 %. Většina hodnot vykazovala relativní SD nižší než 20 %, což naznačuje převážně nízkou až střední variabilitu odhadů AGB. Je však nutné zdůraznit, že uvedené hodnoty se vztahovaly k první verzi dat, která byla následně aktualizována. U aktualizované verze již směrodatná odchylka nebyla k dispozici, a proto nebylo možné nejistotu pozemních dat přímo zahrnout do modelování ani provést formální propagaci nejistoty do výsledné vrstvy. Lze tedy předpokládat, že určitá část nejistoty referenčních dat byla přenesena do modelu, ale její velikost nelze v této práci spolehlivě kvantifikovat. Toto představuje důležité omezení výsledků.

Nejistota vstupních referenčních hodnot byla dále ovlivněna aplikací růstových korekcí ATGI. Ty byly poskytnuty v rastrovém formátu a samy pravděpodobně obsahují další složku nejistoty. Navíc se jednalo o data s prostorovým rozlišením 500 m, což je ve vztahu k velikosti pozemních

ploch a k výslednému 20m mapování poměrně hrubé rozlišení. Z tohoto důvodu nebyla pro každou plochu použita hodnota růstové korekce pouze z konkrétního roku, ale průměr za období 2010–2023. Cílem tohoto rozhodnutí bylo získat stabilnější odhad průměrného přírůstku biomasy na dané lokalitě a omezit vliv meziročních výkyvů v rastrovém produktu. Vzhledem k tomu, že se jednalo o sledované plochy sítě DendroNetwork, u nichž nebyl předpoklad těžby nebo náhlého odstranění porostu, bylo toto řešení zvoleno jako vhodnější než použití jednotlivých ročních hodnot, které mohly lokálně vykazovat i záporné přírůstky. I toto rozhodnutí však představuje další zdroj nejistoty, který nebylo možné samostatně kvantifikovat.

Celkově lze tedy říci, že hlavní nejistota výsledného modelu neleží pouze v samotném modelovacím algoritmu, ale už v referenčních datech. Ta určují rozsah hodnot, který se model může naučit, prostorovou reprezentativnost tréninku i kvalitu následné validace. Z tohoto hlediska je omezený počet a nerovnoměrná distribuce referenčních ploch pravděpodobně nejdůležitější limitací celé studie.

Výběr modelovacích algoritmů a ladění hyperparametrů

Pro modelování byly zvoleny neparametrické algoritmy Random Forest, XGBoost a LightGBM. Jejich použití je vhodné vzhledem k charakteru vstupních dat, u nichž nelze předpokládat normální rozdělení, lineární vztahy ani jednoduchou vazbu mezi spektrálními, radarovými a topografickými prediktory a nadzemní biomasou. Tyto algoritmy jsou schopné zachytit nelineární vztahy, interakce mezi prediktory a rozdílný význam proměnných v různých částech datového prostoru (Gao a kol., 2018; Wai a kol., 2022; Zhou a kol., 2025). Z tohoto důvodu jsou často využívány v úlohách environmentálního a dálkově-průzkumného modelování.

Modely byly laděny pomocí hyperparametrů, přičemž pro každý algoritmus byl definován rozsah možných hodnot a následně byla vybrána kombinace vedoucí k nejlepším validačním výsledkům. V práci byl použit *grid search*, tedy systematické vyhledávání nejlepší kombinace hyperparametrů z předem definované množiny hodnot. Výhodou tohoto přístupu je jeho transparentnost a reprodukovatelnost. Je však nutné zdůraznit, že *grid search* nenachází obecně nejlepší možnou kombinaci hyperparametrů, ale pouze nejlepší kombinaci v rámci předem zadaného diskrétního seznamu hodnot. Pokud by byla hledací množina rozšířena nebo jemněji odstupňována, mohly by být nalezeny i jiné kombinace. *Grid search* je zároveň výpočetně náročný, zejména při kombinaci více algoritmů, většího počtu prediktorů a vnořené křížové validace. V průběhu práce byl testován také *random search*, který je výpočetně rychlejší, ale výběr kombinací je náhodný. Literatura zároveň upozorňuje, že *random search* může být v mnoha případech efektivnější než *grid search*, protože neplýtvá výpočty na kombinace méně významných parametrů a lépe pokrývá prostor hodnot při omezeném počtu pokusů (Bergstra a Bengio, 2012). Do budoucna by bylo vhodné otestovat také bayesovskou optimalizaci hyperparametrů. Ta postupně modeluje vztah mezi volbou hyperparametrů a validačním výkonem modelu a na základě předchozích výsledků vybírá další vhodné kombinace k otestování. Tento přístup může být výpočetně efektivnější než *grid search*, zejména pokud je trénink modelu časově náročný nebo pokud je prostor hyperparametrů rozsáhlý (Snoek a kol., 2012). V této práci však byl *grid search* zvolen kvůli jeho přehlednosti, jednoznačnosti a reprodukovatelnosti v rámci definovaného rozsahu hodnot.

Validace modelu a prostorová závislost dat

Kromě samotného tréninku modelu je klíčovou součástí modelování také validace. V počátečních fázích práce bylo testováno běžné náhodné rozdělení dat na trénovací a validační množinu v poměru 70:30. Tento přístup však vedl k velmi optimistickým výsledkům, které pravděpodobně neodpovídaly reálné schopnosti modelu predikovat AGB mimo známé lokality. Následně byla testována také náhodná sedminásobná křížová validace. Ta poskytovala realističtější výsledky než jednoduché rozdělení 70:30, ale stále neřešila problém prostorové blízkosti některých pozemních měření.

V případě prostorových dat je časté, že vzorky ležící blízko sebe sdílejí podobné stanovištní podmínky, druhové složení, věkovou strukturu, historii hospodaření nebo spektrální vlastnosti. Pokud se takové prostorově blízké vzorky náhodně rozdělí mezi trénovací a validační část, model může být hodnocen na datech, která jsou velmi podobná datům použitým při tréninku. Výsledkem je umělé zvýšení predikční výkonnosti. Tento problém je v ekologickém a environmentálním modelování dobře známý a bloková nebo prostorově strukturovaná validace je obecně doporučována tam, kde je cílem hodnotit přenositelnost modelu do nových lokalit (Roberts a kol., 2017). Z tohoto důvodu byla v práci použita prostorově vnořená křížová validace se seskupením podle 1 km mřížky. Tento postup zajišťuje, že vzorky náležející do stejného 1 km prostorového bloku nejsou rozděleny mezi trénovací a validační část. Výsledky jsou proto konzervativnější, ale pravděpodobně realističtější než u náhodné validace. To je zvláště důležité pro národní odhad AGB, kde model není určen pouze k predikci v okolí známých ploch, ale k aplikaci na celé území České republiky. Použití vnořené křížové validace navíc umožnilo oddělit proces ladění hyperparametrů od finálního hodnocení modelu. Vnitřní část validace sloužila k výběru optimálních hyperparametrů, zatímco vnější část poskytla nezávislejší odhad predikční schopnosti modelu. Tento postup snižuje riziko nadhodnocení výsledků způsobené tím, že by stejná data byla použita jak k ladění modelu, tak k jeho hodnocení. Je nutné uvést, že ani prostorová validace neřeší všechny problémy přenositelnosti. Pokud se v mapovaném území vyskytují kombinace prediktorů, které nebyly dostatečně zastoupeny v trénovacích datech, predikce v těchto oblastech zůstávají nejisté. V tomto směru by budoucí práce mohla využít koncept oblasti aplikovatelnosti modelu, který umožňuje identifikovat oblasti, kde je predikce založena na podobném datovém prostoru jako trénovací data, a oblasti, kde model pravděpodobně extrapoluje (Meyer a Pebesma, 2021).

Výběr prediktorů a jejich vliv na výsledky

Další důležitou částí modelování byl výběr prediktorů. V průběhu práce byly opakovaně testovány různé kombinace vstupních proměnných, validačních schémat a modelovacích algoritmů. Na základě výsledků modelů a významnosti prediktorů byly postupně identifikovány proměnné, které dlouhodobě přinášely pouze malou informační hodnotu. Ty byly systematicky odstraňovány, což vedlo ke zjednodušení modelu a ve většině případů i ke zvýšení jeho výkonnosti.

Původní soubor optických prediktorů Sentinel-2 zahrnoval spektrální pásma, vegetační indexy a texturové charakteristiky GLCM. Z tohoto širšího souboru byly odstraněny některé prediktory,

které vykazovaly nízkou důležitost napříč různými modelovacími konfiguracemi. Příkladem je GNDVI, který ve většině testů přinášel pouze omezené množství dodatečné informace. Podobně byly redukovány i některé texturové vrstvy odvozené z pásem B4, B8 a B11. Přestože tato pásma patří v mapování vegetace mezi důležitá, ne všechny z nich přispívaly v texturové podobě ke zlepšení modelu. Tento výsledek ukazuje, že větší počet prediktorů automaticky neznamená lepší model. Některé proměnné mohou být redundantní, silně korelované s jinými prediktory nebo mohou do modelu přinášet šum. Vzhledem k omezenému počtu referenčních ploch je navíc důležité udržet model dostatečně jednoduchý a robustní. Významnou roli v modelu sehrála kombinace radarových dat Sentinel-1 a optických dat Sentinel-2. Radarové prediktory VV a VH poskytují informaci o struktuře porostu, zatímco optická data zachycují spektrální vlastnosti povrchu, stav korunového zápoje a případně podíl odkrytého půdního povrchu. Kombinace obou typů dat je proto vhodná, protože zachycuje různé aspekty lesního porostu.

Zároveň se ukázalo, že některé vegetační indexy měly po zahrnutí původních spektrálních pásem a radarových proměnných pouze omezený dodatečný přínos. To neznamená, že vegetační indexy nejsou pro mapování biomasy užitečné obecně, ale v této konkrétní situaci mohly částečně opakovat informaci již obsaženou ve spektrálních pásmech. Podobně i topografické a strukturální prediktory je nutné hodnotit nejen podle jejich obecné relevance, ale podle konkrétního přínosu v daném modelovacím rámci.

Výběr finálního modelu

Z testovaných algoritmů dosáhly nejlepší výsledky Random Forest a XGBoost, zatímco LightGBM byl v této konfiguraci slabší. Přestože XGBoost dosáhl v některých metrikách mírně lepšího výkonu, pro finální mapování byl zvolen Random Forest. Toto rozhodnutí nebylo založeno pouze na jedné hodnotě R^2 nebo RMSE, ale na celkovém chování modelu, interpretovatelnosti výsledků a stabilitě výstupu. Random Forest poskytoval vyrovnaný výkon, nižší typické absolutní chyby a zároveň produkoval fyzikálně smysluplné hodnoty bez nutnosti dodatečných úprav záporných predikcí. Oproti tomu XGBoost v některých případech generoval záporné hodnoty AGB, které musely být dodatečně ořízнuty na nulu. Ačkoli je takový post-processing technicky možný, pro datový produkt určený k dalšímu využití je vhodnější model, který přirozeně poskytuje konzistentní a fyzikálně smysluplné hodnoty. Důležitým argumentem pro výběr Random Forest byla také jeho vyváženější závislost na radarových a optických prediktorech. Model, který je extrémně silně závislý na jedné proměnné, může být citlivější na specifické podmínky akvizice, vlhkostní stav, geometrii snímání nebo lokální spektrální anomálie. Vyváženější využití více zdrojů informace může naopak přispět k robustnějšímu chování modelu při aplikaci na heterogenní území.

Struktura chyb a limity modelování extrémních hodnot AGB

Analýza reziduí ukázala, že model má tendenci nadhodnocovat nízké hodnoty AGB a naopak podhodnocovat velmi vysoké hodnoty biomasy. Tento vzorec je v dálkovém průzkumu Země očekávatelný. U nízkých hodnot může být problémem omezené zastoupení takových ploch v referenčním souboru a spektrální podobnost různých typů nízké vegetace. U vysokých hodnot AGB se naopak projevuje saturace optických i radarových dat, zejména pokud je porost hustý

a další nárůst biomasy již nevede k jednoznačné změně spektrální nebo radarové odezvy. Tento problém je důležitý zejména u starších, hustších nebo strukturálně složitých porostů, kde se může skutečná biomasa výrazně lišit, ale dostupné satelitní prediktory tuto variabilitu nedokážou plně rozlišit. Model pak predikuje hodnoty blíže střednímu rozsahu dat a extrémní hodnoty jsou částečně potlačeny. Výsledná vrstva tak pravděpodobně nejlépe zachycuje střední rozsah AGB, zatímco nejnižší a nejvyšší hodnoty jsou zatíženy větší nejistotou. Tento závěr podporuje také skutečnost, že referenční data byla nerovnoměrně rozmístěna v biomase i prostoru. Pokud by bylo k dispozici více ploch s velmi nízkou a velmi vysokou AGB, bylo by možné lépe naučit model chování na okrajích hodnotového rozsahu. Rozšíření referenční databáze je proto jedním z hlavních předpokladů pro další zlepšení modelu.

Možnosti budoucího rozvoje

Budoucí výzkum by měl směřovat především ke zlepšení referenčních dat. Největší přínos by mělo rozšíření počtu pozemních ploch a jejich lepší stratifikace podle hodnot AGB, lesních typů, věkových tříd, druhového složení, stanovištních podmínek a disturbance. Zvláštní pozornost by měla být věnována plochám s nízkou biomasou, jako jsou holiny, mladé regenerující porosty a řídké porosty, a také plochám s velmi vysokou biomasou. Právě tyto části gradientu představují pro modelování největší problém. Metodicky by bylo vhodné dále rozvíjet také práci s nejistotou. V ideálním případě by budoucí modelování nemělo pracovat pouze s jednou hodnotou AGB na referenční ploše, ale také s odhadem její nejistoty. To by umožnilo lépe odlišit chybu modelu od nejistoty vstupních dat a případně vytvořit i prostorovou vrstvu nejistoty výsledného mapového produktu. V současné práci to nebylo možné kvůli omezené dostupnosti informací o nejistotě aktualizovaných referenčních dat.

Dalším perspektivním směrem je využití strukturálních prediktorů, zejména výšky porostu. V této práci byla výšková data testována, nicméně kvůli časovému nesouladu s dalšími prediktory nebyly zahrnuty do finálního výběru. Nicméně po jejich přidání k ostatním prediktorům (model ALL) dosahoval model lepších výsledků. Do budoucna by bylo vhodné vytvořit vlastní výškový model založený například na datech GEDI, případně kombinovat GEDI s dalšími zdroji tak, aby vznikl časově konzistentní strukturální prediktor. Takový přístup by mohl pomoci zejména při modelování vyšších hodnot AGB, kde optická a radarová data narážejí na problém saturace.

Zajímavou možnost představují také nové produkty typu *satellite embeddings*. Například dataset Google Satellite Embedding v prostředí Google Earth Engine poskytuje globální roční 10m vrstvy, v nichž je každý pixel reprezentován 64rozměrným vektorem shrnujícím roční povrchové podmínky z více zdrojů EO dat. Pro budoucí modelování AGB by bylo možné otestovat, zda tyto reprezentace dokážou zachytit komplexní prostorovou, spektrální a časovou informaci lépe než ručně vytvořené kombinace jednotlivých pásem, indexů a textur. Vzhledem k tomu, že jde o relativně nový produkt dostupný od roku 2025, nebyl součástí původní metodiky této práce.

Další možností je rozdělení modelování podle lesních typů nebo dominantních dřevinných skupin. Jednotný model pro všechny lesní porosty má výhodu jednoduchosti a umožňuje vytvořit konzistentní národní dataset. Na druhou stranu však může částečně potlačovat rozdíly mezi jehličnatými, listnatými a smíšenými porosty, které se mohou lišit strukturou, spektrální odezvou

i vztahem mezi výškou, zápojem a biomasou. Budoucí práce by proto mohla porovnat jednotný model s oddělenými modely podle typů porostů.

Shrnutí a zhodnocení práce

Tato práce prokázala, že nadzemní biomasu lesních porostů a uhlík v nadzemní biomase lze v národním měřítku modelovat pomocí volně dostupných dat dálkového průzkumu Země. Výsledky jsou významné tím, že byly hodnoceny pomocí prostorově vnořené křížové validace, která poskytuje realističtější odhad predikční výkonnosti než běžné náhodné validační postupy. Hlavní limitací práce však zůstává omezený a nerovnoměrně stratifikovaný soubor referenčních dat. Nedostatečné zastoupení nízkých a vysokých hodnot AGB ovlivňuje schopnost modelu správně predikovat krajní části biomasy, což se projevuje nadhodnocováním nízkých a podhodnocováním vysokých hodnot. Další nejistota vstupuje do modelování prostřednictvím pozemních měření, růstových korekcí a nemožnosti formálně propagovat nejistotu referenčních dat do výsledného mapového produktu.

Navzdory těmto omezením výsledky ukazují, že použitý přístup je funkční a využitelný pro tvorbu národního datasetu AGB a uhlíku. Práce tak představuje důležitý krok směrem k systematickému využívání volně dostupných družicových dat pro hodnocení lesní biomasy a uhlíku v České republice. Současně jasně ukazuje, že další zlepšení přesnosti nebude záviset pouze na sofistikovanějších algoritmech, ale především na dostupnosti kvalitnějších, početnějších a lépe stratifikovaných referenčních dat.

6. ZÁVĚR

Cílem disertační práce bylo sestavit a prakticky otestovat metodický postup pro kvantifikaci nadzemní biomasy a uhlíku v nadzemní biomase lesních porostů na území České republiky s využitím volně dostupných dat dálkového průzkumu Země, referenčních pozemních měření a metod strojového učení. Na základě dosažených výsledků lze konstatovat, že **hlavní cíl práce i všechny čtyři dílčí cíle byly splněny**. Byl vymezen teoretický a datový rámec řešené problematiky, sestaven opakovatelný postup přípravy vstupních dat, otestovány různé sady prediktorů a tři algoritmy strojového učení a vytvořeny výsledné vrstvy nadzemní biomasy a z ní odvozeného uhlíku, které byly porovnány s nezávislými globálními produkty a údaji Národní inventarizace lesů. **Základní hypotéza práce byla potvrzena: nadzemní biomasu lesních porostů lze v národním měřítku modelovat pomocí volně dostupných dat a z výsledné vrstvy lze následně odvodit prostorové rozložení uhlíku v nadzemní biomase.** Uhlík přitom nebyl modelován jako samostatná veličina, ale byl vypočten z odhadu nadzemní biomasy pomocí přepočetního faktoru 0,50.

Významným výsledkem práce je propojení jednotlivých kroků do uceleného a opakovatelného pracovního postupu. Data Sentinel-1 a Sentinel-2 a vegetační indexy byla připravena pomocí skriptů v prostředí Google Earth Engine, další prediktory byly harmonizovány na společné prostorové rozlišení a rastrovou mřížku a modelování bylo provedeno v prostředí Pythonu. Důležitou součástí postupu byla prostorově vnořená křížová validace se seskupením referenčních ploch podle mřížky s velikostí buňky 1 km. Oddělení vnitřní smyčky pro ladění hyperparametrů od vnější smyčky pro nezávislé hodnocení omezilo riziko nadhodnocení predikčního výkonu. Dosažené hodnoty přesnosti jsou proto konzervativnější než při náhodném dělení dat, ale lépe vystihují schopnost modelu přenášet odhad do prostorově oddělených a dříve nepozorovaných lokalit.

Výzkumná otázka týkající se přínosu dalších vstupních vrstev byla zodpovězena odlišně. **Rozšíření počtu vstupních prediktorů nevedlo automaticky ke zvýšení přesnosti** ani ke kvalitnější finální vrstvě. Samotná optická data Sentinel-2 poskytovala nejslabší výsledky, zatímco radarová data Sentinel-1 nesla v použitém referenčním souboru silnou informaci o struktuře porostu a u algoritmu Random Forest dosáhla nejlepších numerických metrik. Výsledný raster založený pouze na Sentinel-1 však vykazoval zvýšenou pixelovou fragmentaci a nižší prostorovou věrohodnost. **Kombinace Sentinel-1 a Sentinel-2 naopak představovala nejvyváženější kompromis mezi přesností, stabilitou napříč algoritmy** a schopností zachytit souvislé porosty, holiny, regenerující plochy i přechody mezi lesním a nelesním prostředím. **Přidání morfometrických charakteristik nepřineslo zlepšení** a kompletní **sada zahrnující také výšku porostu sice u některých konfigurací přesnost zvýšila**, její použití však omezoval časový nesoulad s ostatními prediktory. Přínos doplňkových vrstev je tedy podmíněn jejich skutečnou informační hodnotou, časovou konzistencí a prostorovým chováním výsledné predikce, nikoli pouze jejich počtem.

Při detailním srovnání algoritmů pro zvolenou kombinaci Sentinel-1 a Sentinel-2 dosáhly nejlepších a vzájemně velmi podobných výsledků Random Forest a XGBoost, zatímco LightGBM byl v použité konfiguraci méně přesný. **Jako finální byl vybrán model Random Forest**. V prostorově vnořené křížové validaci dosáhl v průměru R^2 = 0,69 ± 0,12, RMSE = 79,28 ± 16,14 Mg·ha^{-1} a MAE = 48,03 ± 10,20 Mg·ha^{-1}; *pooled out-of-fold* hodnocení poskytlo R^2 = 0,70, RMSE = 80,47 Mg·ha^{-1} a MAE = 48,01 Mg·ha^{-1}. XGBoost vykázal mírně vyšší R^2 a nižší RMSE, Random Forest však dosáhl nižší typické absolutní chyby, velmi nízkého systematického vychýlení na lesních plochách, vyváženějšího využití radarových a optických prediktorů a v celoplošném výstupu neprodukoval záporné hodnoty. **Výběr finálního modelu tak nebyl založen pouze na jediné statistické metrice, ale také na stabilitě, interpretovatelnosti, prostorové kvalitě a fyzikální věrohodnosti výsledku.**

Finální model byl aplikován na lesní masku o rozloze přibližně 2 717 tisíc hektarů a poskytl prostorově souvislou vrstvu nadzemní biomasy pro vegetační sezónu roku 2023 v rozlišení 20 m. Průměrná AGB byla odhadnuta na 213,63 Mg·ha^{-1} a celková zásoba na 580,6 milionu Mg. Tomu odpovídala průměrná zásoba uhlíku v nadzemní biomase 106,82 Mg·ha^{-1} a celková zásoba přibližně 290,3 milionu Mg C. Průměrná hektarová hodnota AGB byla velmi blízká výsledku Národní inventarizace lesů 219,4 ± 2,5 Mg·ha^{-1}. Nižší celkový odhad oproti inventarizaci souvisel především s menší rozlohou použité lesní masky a částečně také s podhodnocováním nejvyšších hodnot biomasy. Vytvořená vrstva současně dosáhla vyšší průměrné AGB a lepší shody s pozemními referencemi než globální produkty ESA CCI Biomass v6 a GEDI L4B. Výsledky tak potvrzují přínos modelu kalibrovaného pro podmínky České republiky, zejména pokud je cílem zachytit jemnější prostorovou variabilitu lesních porostů.

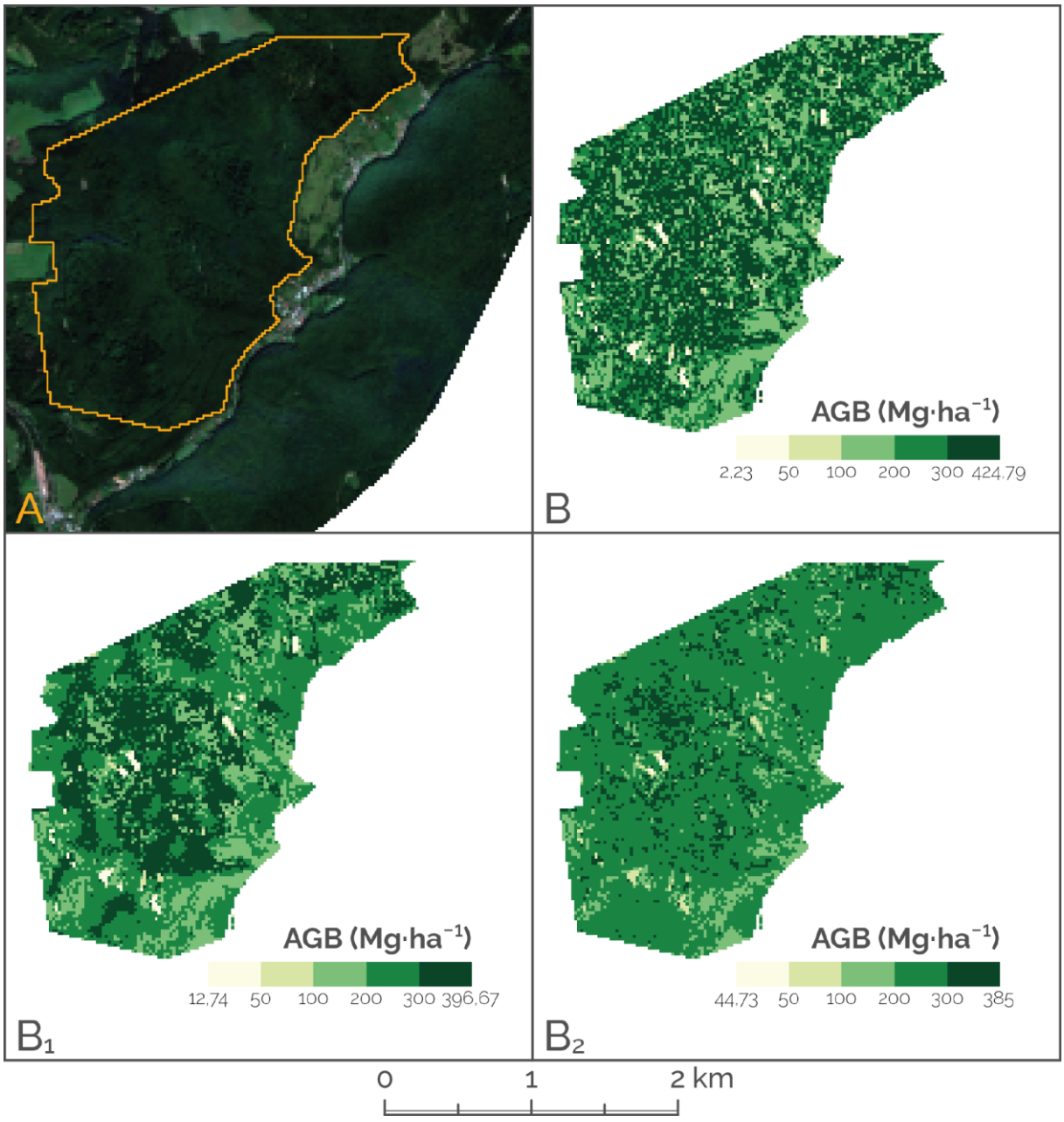
Přínos disertační práce je metodický, poznatkový i datový. Metodicky přináší opakovatelný postup od přípravy radarových a optických dat po prostorovou validaci modelů. Z poznatkového hlediska ukazuje, že nejvyšší numerická přesnost nemusí vést k nejvhodnějšímu mapovému produktu a že je nutné hodnotit i jeho prostorovou strukturu, fyzikální smysluplnost a časovou konzistenci vstupů. Datovým výstupem jsou národní vrstvy AGB a uhlíku v nadzemní biomase, zpřístupněné ve vědeckém repozitáři. Nenahrazují Národní inventarizaci lesů, ale představují prostorově podrobný doplněk pro analýzu regionálních rozdílů, monitoring uhlíkových zásob a další výzkum. Další zpřesnění vyžaduje početnější a lépe stratifikovaná referenční data a časově konzistentní strukturální prediktory. Práce však potvrzuje využitelnost volně dostupných družicových dat pro národní mapování AGB a z ní odvozeného uhlíku.

7. DOSTUPNOST DAT A ZDROJOVÝCH KÓDŮ

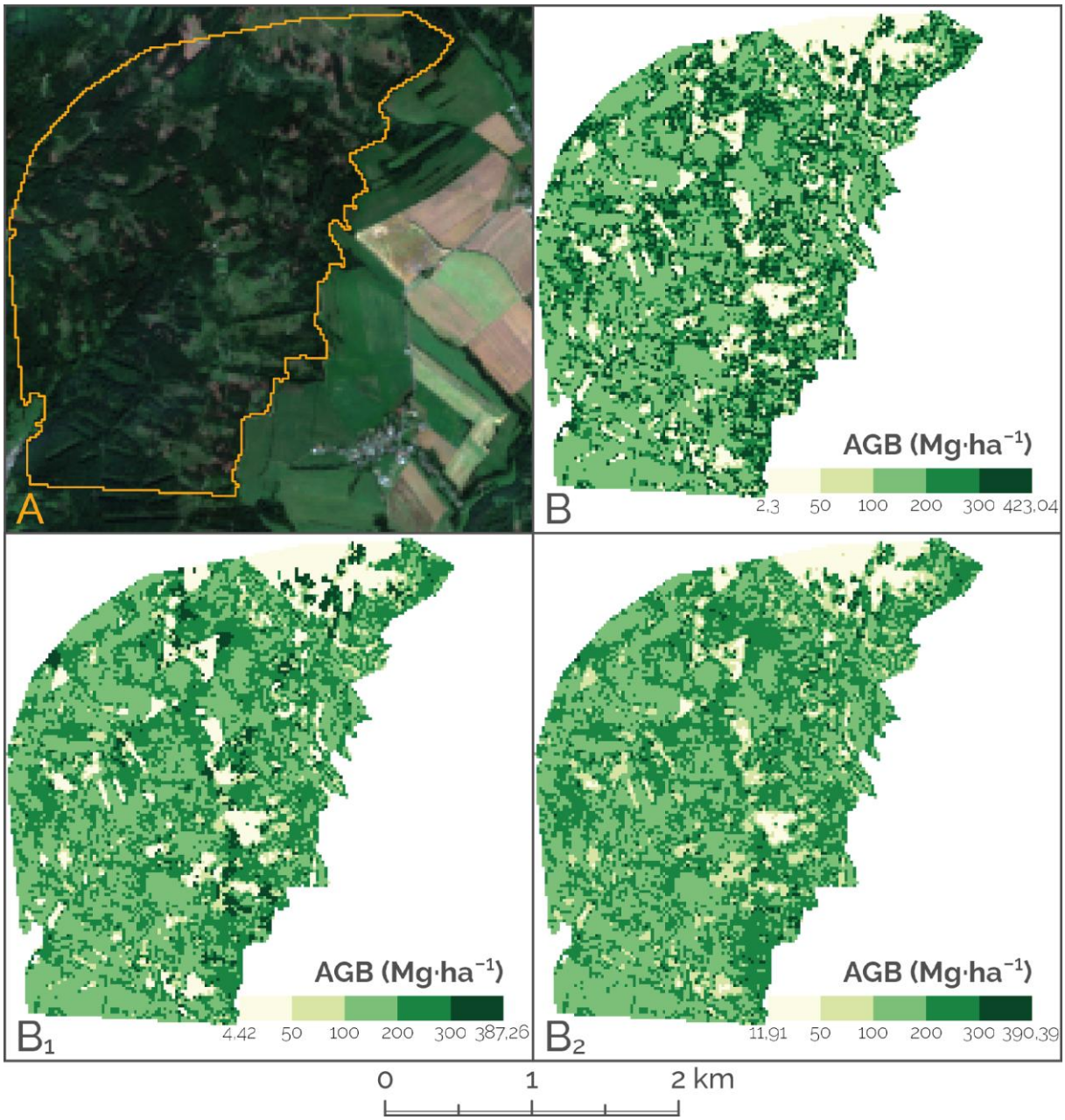
Výsledný rastr nadzemní biomasy lesních porostů České republiky pro rok 2023 je volně dostupný v repozitáři Zenodo: https://doi.org/10.5281/zenodo.19627063. Stejně tak je dostupný i rastr uhlíku v nadzemní biomase lesních porostů České republiky pro rok 2023: https://doi.org/10.5281/zenodo.21260614. Oba rastry jsou dostupné ve formátu TIFF.

Zdrojový kód pro zpracování dat Sentinel-2, vytvoření sezónní mozaiky, vypočtení vegetačních indexů a exportu do Google Drive je dostupný v prostředí Google Earth Engine Code Editor: https://code.earthengine.google.com/1a90825e9a5e125eeb8da6ed8834b2c5. Další kód slouží ke zpracování dat Sentinel-1 a výšky lesního porostu s následným exportem do Google Drive: https://code.earthengine.google.com/2efd2ca901d28c95299d2fb6c1cd0144. Zdrojové kódy použité pro trénování modelů, optimalizaci hyperparametrů, prostorově vnořenou křížovou validaci a tvorbu rastrových predikcí jsou dostupné v repozitáři GitHub: https://github.com/Pajis/DP_vyvlecka.

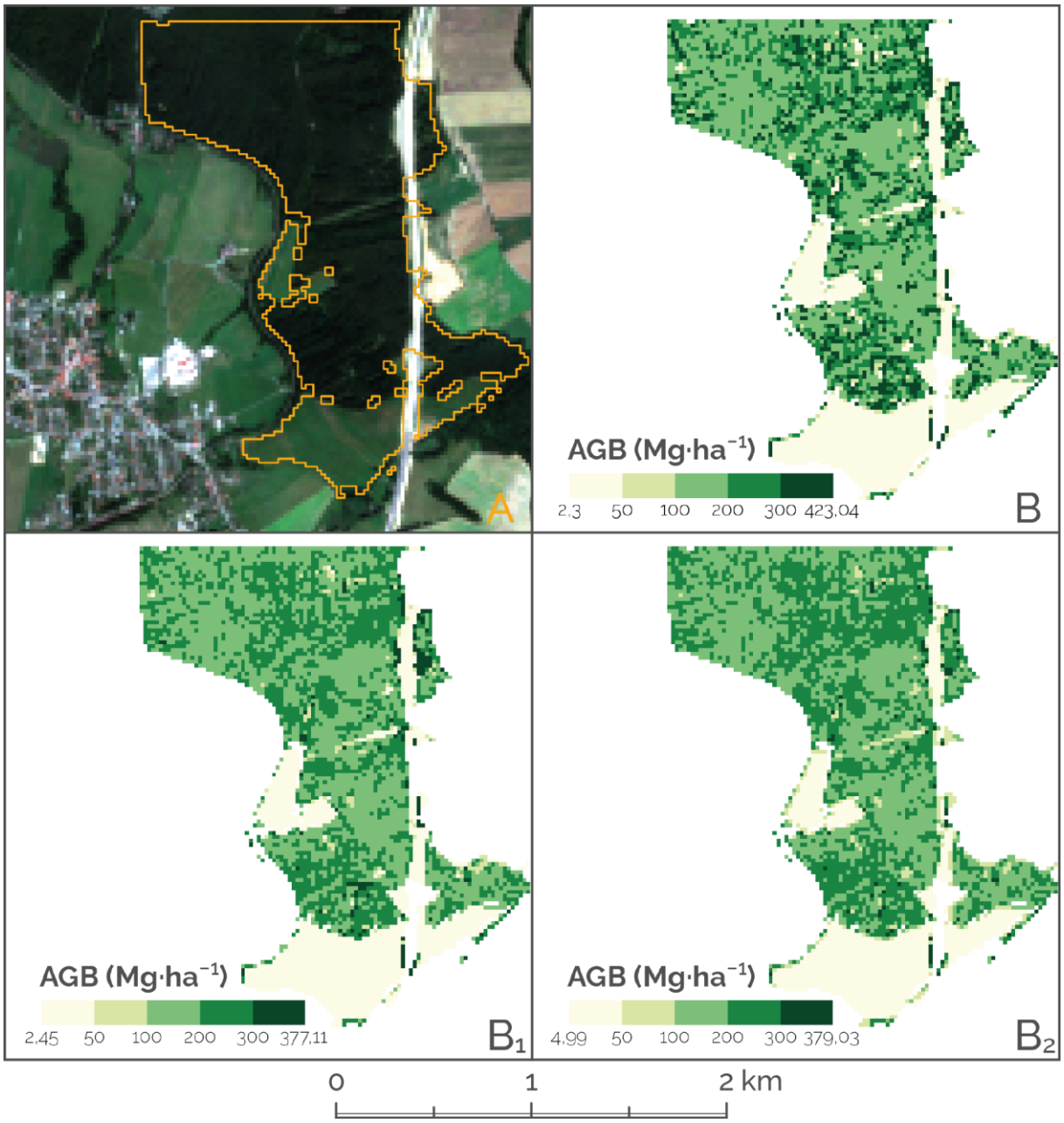
8. PŘÍLOHY



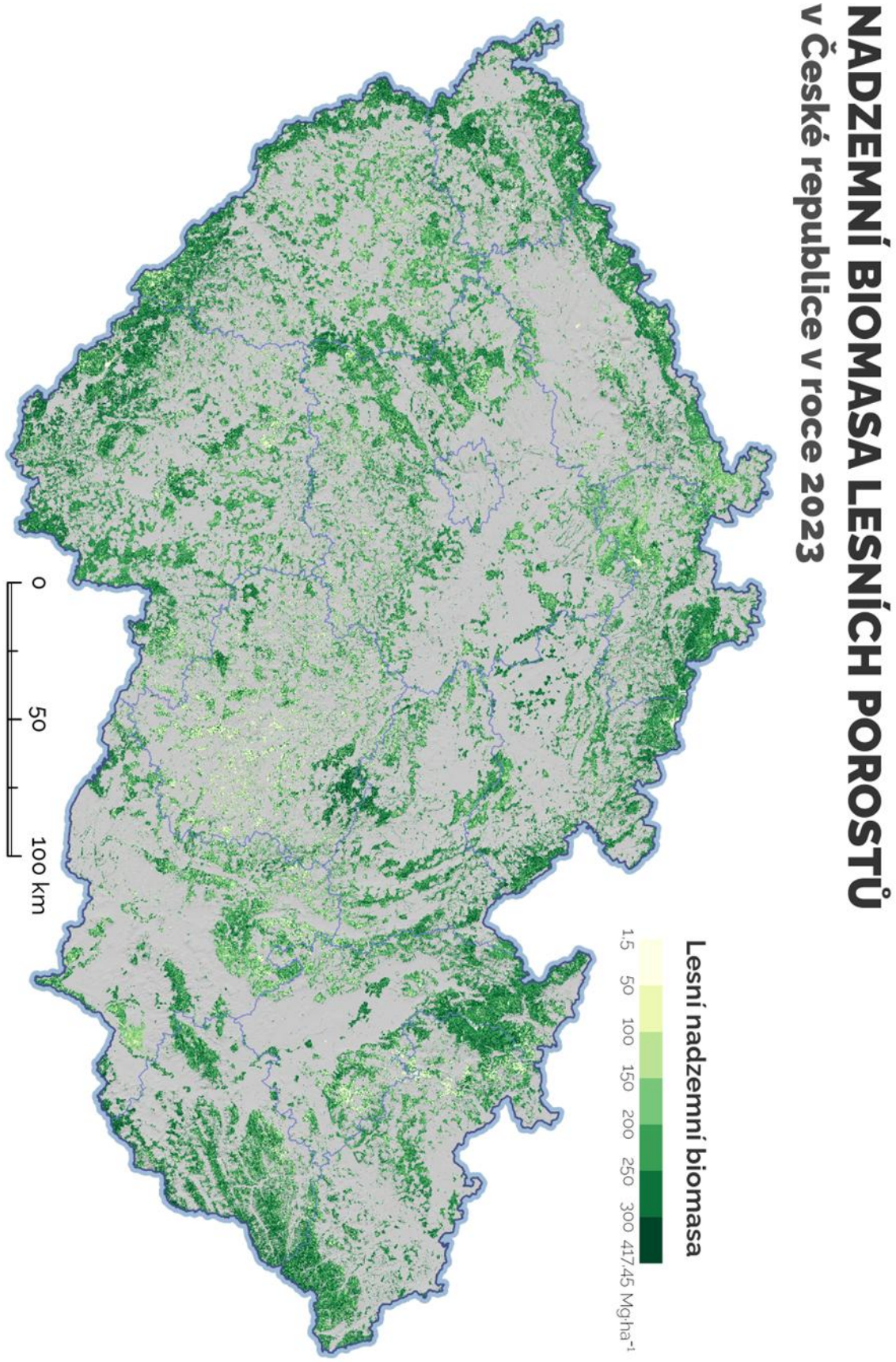
Obrázek 26 Prostorové porovnání predikované AGB ve výřezu 1. A: RGB kompozice Sentinel-2 s vyznačeným územím. B: model Random Forest se sadou prediktorů S1. C: upravený model Random Forest S1 pomocí bilateral filtru. D: upravený model Random Forest S1 pomocí guided filtru.



Obrázek 27 Prostorové porovnání predikované AGB ve výřezu 2. A: RGB kompozice Sentinel-2 s vyznačeným územím. B: model Random Forest se sadou prediktorů S1. C: upravený model Random Forest S1 pomocí bilateral filtru. D: upravený model Random Forest S1.



Obrázek 28 Prostorové porovnání predikované AGB ve výřezu 3. A: RGB kompozice Sentinel-2 s vyznačeným územím. B: model Random Forest se sadou prediktorů S1. C: upravený model Random Forest S1 pomocí bilateral filtru. D: upravený model Random Forest S1.



Obrázek 29 Lesní nadzemní biomasa v České republice v roce 2023

Obrázek 30 Uhlík v nadzemní biomase lesních porostů v České republice v roce 2023

LITERATURA

Abrishamkar, M.; Ahmadi, A. Evapotranspiration Estimation Using Remote Sensing Technology Based on SEBAL Algorithm. Iran. J. Sci. Technol. - Trans. Civ. Eng. 2017, 41, 65–76, doi:10.1007/S40996-016-0036-X/FIGURES/7.

Almagro, A.; Thomé, T.C.; Colman, C.B.; Pereira, R.B.; Marcato Junior, J.; Rodrigues, D.B.B.; Oliveira, P.T.S. Improving Cover and Management Factor (C-Factor) Estimation Using Remote Sensing Approaches for Tropical Regions. *Int. Soil Water Conserv. Res.* 2019, *7*, 325–334, doi:10.1016/J.ISWCR.2019.08.005.

Altman, N. S. 1992. "An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression." American Statistician 46 (3): 175–85. https://doi.org/10.1080/00031305.1992.10475879.

Anwar, A.; Saleem, U.; Shaiza, B.; Naveed, A.; Asad, A.; Awais, K.; A, A.; S, U.; S, B.; N, A.; et al. Quantifying Forest Carbon Stocks by Integrating Satellite Images and Forest Inventory Data. 2018, 2, 93–117.

Avitabile, V., Baldoni, E., Baruth, B., Bausano, G., Boysen-Urban, K., Caldeira, C., Camia, A., Cazzaniga, N., Ceccherini, G., De la Vega, L., Doerner, H., Giuntoli, J., Gras, M., Guillen, G.J., Gurria, P., Hassegawa, M., Jasinevičius, G., Jonsson, R., Konrad, C., Kupschus, S., La Notte, A., M'barek, R., Mannini, A., Migliavacca, M., Mubareka, S., Patani, S., Pilli, R., Rebours, C., Ronchetti, G., Ronzon, T., Rougieux, P., Sala, S., Sanchez Lopez, J., Sanye Mengual, E., Sinkko, T., Sturm, V., Van Leeuwen, M., Vasilakopoulos, P., Verkerk, P.J., Virtanen, J., Winker, H. and Zulian, G. 2023. Biomass production, supply, uses and flows in the European Union. Luxembourg: Publications Office of the European Union. https://doi.org/10.2760/811744

Avitabile, V., Herold, M., Heuvelink, G.B.M., Lewis, S.L., Phillips, O.L., Asner, G.P., Armston, J., Ashton, P.S., Banin, L., Bayol, N., Berry, N.J., Boeckx, P., de Jong, B.H.J., Devries, B., Girardin, C.A.J., Kearsley, E., Lindsell, J.A., Lopez-Gonzalez, G., Lucas, R., Malhi, Y., Morel, A., Mitchard, E.T.A., Nagy, L., Qie, L., Quinones, M.J., Ryan, C.M., Slik, F.J.W., Sunderland, T., Vaglio Laurin, G., Cazzolla Gatti, R., Valentini, R., Verbeeck, H., Wijaya, A., Willcock, S. 2016. An Integrated Pan-Tropical Biomass Map Using Multiple Reference Datasets. Global Change Biology, 22(4), 1406–1420. https://doi.org/10.1111/GCB.13139

Baccini, A., Walker, W., Carvalho, L., Farina, M., Sulla-Menashe, D. and Houghton, R.A. 2017. Tropical forests are a net carbon source based on aboveground measurements of gain and loss. Science, 358(6360), 230–234. https://doi.org/10.1126/science.aam5962

Bai, J.; Li, J.; Li, S. Monitoring the Plant Density of Cotton with Remotely Sensed Data. IFIP Adv. Inf. Commun. Technol. 2011, 345 AICT, 90–101, doi:10.1007/978-3-642-18336-2_11/COVER.

Berger, C.; Lux, H.; Urban, M.; Schmullius, C.; Baade, J.; Thiel, C.; Wigley-Coetsee, C.; Smit, I. Annual Grass Biomass Mapping with Landsat-8 and Sentinel-2 Data over Kruger National Park, South Africa. *Int. Geosci. Remote Sens. Symp.* 2020, 4323–4326, doi:10.1109/IGARSS39084.2020.9324413.

Bergstra, J.; Bengio, Y. Random Search for Hyper-Parameter Optimization. Journal of Machine Learning Research. 2012, 13(10), 281–305.

Braghiere, R.K.; Quaife, T.; Black, E.; He, L.; Chen, J.M. Underestimation of Global Photosynthesis in Earth System Models Due to Representation of Vegetation Structure. *Global Biogeochem. Cycles* 2019, *33*, 1358–1369, doi:10.1029/2018GB006135.

Breiman, L. Random Forests. *Mach. Learn.* 2001, *45*, 5–32, doi:10.1023/A:1010933404324.

Bright, R.M.; Astrup, R. Combining MODIS and National Land Resource Products to Model Land Cover-Dependent Surface Albedo for Norway. Remote Sens. 2019, Vol. 11, Page 871 2019, 11, 871, doi:10.3390/RS11070871.

Brovkina, O., Navrátilová, B., Novotný, J., Albert, J., Slezák, L. and Cienciala, E. 2022. Influences of vegetation, model, and data parametres on forest aboveground biomass assessment using an area-based approach. Ecological Informatics, 70, 101754. https://doi.org/10.1016/j.ecoinf.2022.101754

Brown De Colstoun, E. C., Story, M. H., Thompson, C., Commisso, K., Smith, T. G., & Irons, J. R. (2003). National Park vegetation mapping using multitemporal Landsat 7 data and a decision tree classifier. Remote Sensing of Environment, 85(3), 316–327. https://doi.org/10.1016/S0034-4257(03)00010-5

Brown, T.C.; Bergstrom, J.C.; Loomis, J.B. Defining, Valuing, and Providing Ecosystem Goods and Services. *Nat. Resour. J.* 2007, *47*, 329–376.

Bucha, T.; Papčo, J.; Sačkov, I.; Pajtík, J.; Sedliak, M.; Barka, I.; Feranec, J. Woody Above-Ground Biomass Estimation on Abandoned Agriculture Land Using Sentinel-1 and Sentinel-2 Data. Remote Sens. 2021, Vol. 13, Page 2488 2021, 13, 2488, doi:10.3390/RS13132488.

Camia, A., Robert, N., Jonsson, R., Pilli, R., García-Condado, S., López-Lozano, R., van der Velde, M., Ronzon, T., Gurría, P., M'barek, R., Tsiropoulos, I., Cristiano, R., Parisi, C., Gómez-Barbero, M., Pérez Domínguez, I. and Fellmann, T. 2018. *Biomass production, supply, uses and flows in the European Union: first results from an integrated assessment*. Luxembourg: Publications Office of the European Union. https://doi.org/10.2760/539520

Clerici, N.; Rubiano, K.; Abd-Elrahman, A.; Hoestettler, J.M.P.; Escobedo, F.J. Estimating Aboveground Biomass and Carbon Stocks in Periurban Andean Secondary Forests Using Very High Resolution Imagery. For. 2016, Vol. 7, Page 138 2016, 7, 138, doi:10.3390/F7070138.

Cochran, W.G. 1977. Sampling techniques. 3rd ed. New York: John Wiley & Sons.

Colkesen, Ismail, and Taskin Kavzoglu. 2017. "Ensemble-Based Canonical Correlation Forest (CCF) for Land Use and Land Cover Classification Using Sentinel-2 and Landsat OLI Imagery." Remote Sensing Letters 8 (11): 1082–91. https://doi.org/10.1080/2150704X.2017.1354262.

Comber, A. J. 2008. "The Separation of Land Cover from Land Use Using Data Primitives." Journal of Land Use Science 3 (4): 215–29. https://doi.org/10.1080/17474230802465173.

Comber, Alexis, Katie Medcalf, Richard Lucas, Peter Bunting, Alan Brown, Daniel Clewley, Johanna Breyer, and Steve Keyworth. 2010. "Managing Uncertainty When Aggregating from Pixels to Objects: Habitats, Context-Sensitive Mapping and Possibility Theory." International Journal of Remote Sensing 31 (4): 1061–68. https://doi.org/10.1080/01431160903246691.

Copernicus Global Land Service - Surface Soil Moisture Available online: https://land.copernicus.eu/global/products/ssm.

Copernicus High Resolution Layer. 2022, online: https://land.copernicus.eu/pan-european/high-resolution-layers.

Copernicus Sentinel Online Available online: https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi.

COPERNICUS. 2025a. From launch to images: the roadmap to Sentinel-1C operations [online]. 6. 2. 2025 [cit. 2026-02-27]. Dostupné z: https://sentinels.copernicus.eu/-/from-launch-to-images-the-roadmap-to-sentinel-1c-operations

COPERNICUS. 2025b. OBSERVER: Sentinel-2A extending operations to meet user needs [online]. 5. 2. 2025 [cit. 2026-02-27]. Dostupné z: https://www.copernicus.eu/en/news/news/observer-sentinel-2a-extending-operations-meet-user-needs

Corbane, Christina, Stefan Lang, Kyle Pipkins, Samuel Alleaume, Michel Deshayes, Virginia Elena García Millán, Thomas Strasser, Jeroen Vanden Borre, Spanhove Toon, and Förster Michael. 2015. "Remote Sensing for Mapping Natural Habitats and Their Conservation Status - New Opportunities and Challenges." International Journal of Applied Earth Observation and Geoinformation 37 (May): 7–16. https://doi.org/10.1016/j.jag.2014.11.005.

Cord, A.F.; Brauman, K.A.; Chaplin-Kramer, R.; Huth, A.; Ziv, G.; Seppelt, R. Priorities to Advance Monitoring of Ecosystem Services Using Earth Observation. Trends Ecol. Evol. 2017, 32, 416–428, doi:10.1016/J.TREE.2017.03.003.

Cortes, Corinna, and Vladimir Vapnik. 1995. "Support-Vector Networks." Machine Learning 20 (3): 273–97. https://doi.org/10.1007/bf00994018.

Costanza, R.; D'Arge, R.; De Groot, R.; Farber, S.; Grasso, M.; Hannon, B.; Limburg, K.; Naeem, S.; O'Neill, R. V.; Paruelo, J.; et al. The Value of the World's Ecosystem Services and Natural Capital. Nature 1997, 387, 253–260, doi:10.1038/387253a0.

Český statistický úřad. 2023. Statistická ročenka České republiky 2023. Praha: Český statistický úřad.

ČESKÝ ÚŘAD ZEMĚMĚŘICKÝ A KATASTRÁLNÍ. ZABAGED® – Výškopis – DMR 4G. Digitální model reliéfu České republiky 4. generace v S-JTSK, Bpv [datová sada]. Praha: Zeměměřický úřad, Geoportál ČÚZK. Použitá verze dat: 2023. Dostupné z: Geoportál ČÚZK. [cit. 26. 5. 2026].

Daily, G.C. Introduction: What Are Ecosystem Services? Isl. Press 1997, 1–10.

De Araujo Barbosa, C.C.; Atkinson, P.M.; Dearing, J.A. Remote Sensing of Ecosystem Services: A Systematic Review. Ecol. Indic. 2015, 52, 430–443, doi:10.1016/J.ECOLIND.2015.01.007.

De Groot, R.S.; Wilson, M.A.; Boumans, R.M.J. A Typology for the Classification, Description and Valuation of Ecosystem Functions, Goods and Services. Ecol. Econ. 2002, 41, 393–408, doi:10.1016/S0921-8009(02)00089-7.

Delegido, J.; Verrelst, J.; Alonso, L.; Moreno, J. 2011. Evaluation of Sentinel-2 Red-Edge Bands for Empirical Estimation of Green LAI and Chlorophyll Content. *Sensors*, 11(7), 7063–7081, https://doi.org/10.3390/s110707063

Dewitz, J. (2021). National Land Cover Database (NLCD) 2019 Products [NLCD 2019]. U.S. Geological Survey. https://doi.org/10.5066/P9KZCM54

Di Gregorio, A. a Jansen, L. J. M. (2005): *Land Cover Classification System: Classification concepts and user manual. Software version 2*. Environment and Natural Resources Series No. 8. Food and Agriculture Organization of the United Nations, Rome. ISBN 978-92-5-105327-0.

Dutcă, I., Zianis, D., Petriţan, I.C., Bragă, C.I., Ştefan, G., Yuste, J.C. and Petriţan, A.M. 2020. Allometric biomass models for European beech and silver fir: testing approaches to minimize the demand for site-specific biomass observations. Forests, 11(11), 1136. https://doi.org/10.3390/f11111136

ESA WorldCover project / Contains modified Copernicus Sentinel data (2021) processed by ESA WorldCover consortium

Esch, S. Determination of Soil Moisture and Vegetation Parameters from Spaceborne C-Band SAR on Agricultural Areas, University of Cologne: Cologne, 2018.

ESRI. *Esri Land Cover* [online]. ArcGIS Living Atlas of the World. [cit. 2026-02-27]. Dostupné z: https://livingatlas.arcgis.com/landcover/

Eugene, P.O. RELATIONSHIPS BETWEEN STRUCTURE AND FUNCTION IN THE ECOSYSTEM. 日本生態学会誌 1962, *12*, 108–118, doi:10.18960/seitai.12.3_108.

European Commission (EC). Global Land Cover 2000 database, Joint Research Centre, 2003, https://forobs.jrc.ec.europa.eu/products/glc2000/glc2000.php".

European Commission. 2021. New EU forest strategy for 2030. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. COM(2021) 572 final. Brussels: European Commission.

European Union (EU), Copernicus Land Monitoring Service 2022, European Environment Agency (EEA), Copenhagen, Dennmark, 2022.

Fang, H. Canopy Clumping Index (CI): A Review of Methods, Characteristics, and Applications. *Agric. For. Meteorol.* 2021, *303*, 108374, doi:10.1016/J.AGRFORMET.2021.108374.

FAO. 2012. Forest resources assessment working paper: terms and definitions. Rome: Food and Agriculture Organization of the United Nations.

FAO. 2014. Global Land Cover (GLC-SHARE) Beta-Release 1.0 Database, Land and Water Division, John Latham, Renato Cumani, Ilaria Rosati and Mario Bloise.

FAO. 2020. Global forest resources assessment 2020: country report, Czech Republic. Rome: Food and Agriculture Organization of the United Nations. Available at: https://fra-data.fao.org/assessments/fra/2020/CZE/home/overview

FAO. 2023. FAOSTAT – Crops and livestock products. License: CC BY-NC-SA 3.0 IGO. Extracted from: https://www.fao.org/faostat/en/#data/QCL. Data of Access: [17-07-2023].

Filipponi, F.; Valentini, E.; Xuan, A.N.; Guerra, C.A.; Wolf, F.; Andrzejak, M.; Taramelli, A. Global MODIS Fraction of Green Vegetation Cover for Monitoring Abrupt and Gradual Vegetation Changes. Remote Sens. 2018, Vol. 10, Page 653 2018, 10, 653, doi:10.3390/RS10040653.

Fisher, Jonathan R.B., Eileen A. Acosta, P. James Dennedy-Frank, Timm Kroeger, and Timothy M. Boucher. 2018. "Impact of Satellite Imagery Spatial Resolution on Land Use Classification Accuracy and Modeled Water Quality." Remote Sensing in Ecology and Conservation 4 (2): 137–49. https://doi.org/10.1002/rse2.61.

Fisher, Peter, Alexis Comber, and Richard Wadsworth. 2005. "Land Use and Land Cover: Contradiction or Complement." Re-Presenting GIS. https://www.academia.edu/14493546/Land_use_and_Land_cover_Contradiction_or_Complement.

Forkuor, Gerald, Kangbeni Dimobe, Idriss Serme, and Jerome Ebagnerin Tondoh. 2018. "Landsat-8 vs. Sentinel-2: Examining the Added Value of Sentinel-2's Red-Edge Bands to Land-Use and Land-Cover

Mapping in Burkina Faso." GIScience and Remote Sensing 55 (3): 331–54. https://doi.org/10.1080/15481603.2017.1370169.

Forzieri, G.; Miralles, D.G.; Ciais, P.; Alkama, R.; Ryu, Y.; Duveiller, G.; Zhang, K.; Robertson, E.; Kautz, M.; Martens, B.; et al. Increased Control of Vegetation on Global Terrestrial Energy Fluxes. *Nat. Clim. Chang. 2020 104* 2020, *10*, 356–362, doi:10.1038/s41558-020-0717-0.

Friedman, Jerome H. 2002. "Stochastic Gradient Boosting." Computational Statistics and Data Analysis 38 (4): 367–78. https://doi.org/10.1016/S0167-9473(01)00065-2.

Gao, Q.; Zribi, M.; Escorihuela, M.J.; Baghdadi, N. Synergetic Use of Sentinel-1 and Sentinel-2 Data for Soil Moisture Mapping at 100 m Resolution. Sensors 2017, Vol. 17, Page 1966 2017, 17, 1966, doi:10.3390/S17091966.

Gao, Y.; Lu, D.; Li, G.; Wang, G.; Chen, Q.; Liu, L.; Li, D. Comparative Analysis of Modeling Algorithms for Forest Aboveground Biomass Estimation in a Subtropical Region. *Remote Sens. 2018, Vol. 10, Page 627* 2018, *10*, 627, doi:10.3390/RS10040627.

Gašparović, Mateo, and Tomislav Jogun. 2018. "The Effect of Fusing Sentinel-2 Bands on Land-Cover Classification." International Journal of Remote Sensing 39 (3): 822–41. https://doi.org/10.1080/01431161.2017.1392640.

GCOS. Essential Climate Variable. Available online: https://gcos.wmo.int/en/essential-climate-variables [20-10-2022].

Geng, L.; Che, T.; Ma, M.; Tan, J.; Wang, H. Corn Biomass Estimation by Integrating Remote Sensing and Long-Term Observation Data Based on Machine Learning Techniques. *Remote Sens.* 2021, *13*.

GEOBON Essential Biodiversity Variables Available online: https://geobon.org/ebvs/what-are-ebvs/ [20-10-2022].

Gitelson, A.A.; Merzlyak, M.N. 1994. Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *Journal of Photochemistry and Photobiology B: Biology*, 22(3), 247–252. https://doi.org/10.1016/1011-1344(93)06963-4

Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sensing of Environment 1996, 58(3), 289–298. https://doi.org/10.1016/S0034-4257(96)00072-7

Gitelson, A.A.; Peng, Y.; Masek, J.G.; Rundquist, D.C.; Verma, S.; Suyker, A.; Baker, J.M.; Hatfield, J.L.; Meyers, T. Remote Estimation of Crop Gross Primary Production with Landsat Data. Remote Sens. Environ. 2012, 121, 404–414, doi:10.1016/J.RSE.2012.02.017.

Gizachew, B.; Solberg, S.; Naesset, E.; Gobakken, T.; Bollandsås, O.M.; Breidenbach, J.; Zahabu, E.; Mauya, E.W. Mapping and Estimating the Total Living Biomass and Carbon in Low-Biomass Woodlands Using Landsat 8 CDR Data. Carbon Balanc. Manag. 2016, 11, doi:10.1186/s13021-016-0055-8.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R. 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Guerini Filho, M.; Kuplich, T.M.; Quadros, F.L.F.D. Estimating Natural Grassland Biomass by Vegetation Indices Using Sentinel 2 Remote Sensing Data. https://doi.org/10.1080/01431161.2019.1697004 2020, 41, 2861–2876, doi:10.1080/01431161.2019.1697004.

Guerra-Hernández, J., Narine, L. L., Pascual, A., Gonzalez-Ferreiro, E., Botequim, B., Malambo, L., … Godinho, S. (2022). Aboveground biomass mapping by integrating ICESat-2, SENTINEL-1, SENTINEL-2, ALOS2/PALSAR2, and topographic information in Mediterranean forests. *GIScience & Remote Sensing*, *59*(1), 1509–1533. https://doi.org/10.1080/15481603.2022.2115599

Haines-Young, R.; Potschin, M. The Links between Biodiversity, Ecosystem Services and Human Well-Being. In Ecosystem Ecology; Cambridge University Press, 2010; pp. 110–139.

Hall-Beyer, M. 2017. Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. International Journal of Remote Sensing, 38(5), 1312–1338. https://doi.org/10.1080/01431161.2016.1278314

Hansen, M.C.; Potapov, P. V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman, S. V.; Goetz, S.J.; Loveland, T.R.; et al. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science (80-.).* 2013, *342*, 850–853, doi:10.1126/SCIENCE.1244693/SUPPL_FILE/HANSEN.SM.PDF.

Hantson, S.; Pueyo, S.; Chuvieco, E. Global Fire Size Distribution Is Driven by Human Impact and Climate. Glob. Ecol. Biogeogr. 2015, 24, 77–86, doi:10.1111/GEB.12246/SUPPINFO.

Haralick, R.M., Dinstein, I. and Shanmugam, K. 1973. Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics, SMC-3(6), 610–621. https://doi.org/10.1109/TSMC.1973.4309314

Harwood, T.D.; Donohue, R.J.; Williams, K.J.; Ferrier, S.; McVicar, T.R.; Newell, G.; White, M. Habitat Condition Assessment System: A New Way to Assess the Condition of Natural Habitats for Terrestrial Biodiversity across Whole Regions Using Remote Sensing Data. Methods Ecol. Evol. 2016, 7, 1050–1059, doi:10.1111/2041-210X.12579/ABSTRACT.

He, K., Sun, J., & Tang, X. (2012). Guided image filtering. *IEEE transactions on pattern analysis and machine intelligence*, *35*(6), 1397-1409.

Hicks, C.C.; Cinner, J.E.; Stoeckl, N.; McClanahan, T.R. Linking Ecosystem Services and Human-Values Theory. *Conserv. Biol.* 2015, *29*, 1471–1480.

Hoscilo, A.; Lewandowska, A.; Ziólkowski, D.; Sterenczak, K.; Lisanczuk, M.; Schmullius, C.; Pathe, C. Forest Aboveground Biomass Estimation Using a Combination of Sentinel-1 and Sentinel-2 Data. Int. Geosci. Remote Sens. Symp. 2018, 2018-July, 9026–9029, doi:10.1109/IGARSS.2018.8517965.

Chang, J.; Shoshany, M. Mediterranean Shrublands Biomass Estimation Using Sentinel-1 and Sentinel-2. *Int. Geosci. Remote Sens. Symp.* 2016, *2016-November*, 5300–5303, doi:10.1109/IGARSS.2016.7730380.

Chen, J.; Chen, J. GlobeLand30: Operational Global Land Cover Mapping and Big-Data Analysis. *Sci. China Earth Sci.* 2018, *61*, 1533–1534, doi:10.1007/S11430-018-9255-3/METRICS.

Chen, J.M. Remote Sensing of Leaf Area Index and Clumping Index. *Compr. Remote Sens.* 2018, *1–9*, 53–77, doi:10.1016/B978-0-12-409548-9.10540-8.

Chen, J.M.; Mo, G.; Pisek, J.; Liu, J.; Deng, F.; Ishizawa, M.; Chan, D.; Chen, C.:; Mo, G.; Pisek, J.; et al. Effects of Foliage Clumping on the Estimation of Global Terrestrial Gross Primary Productivity. *Global Biogeochem. Cycles* 2012, *26*, doi:10.1029/2010GB003996.

Chen, T. and Guestrin, C. 2016. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), 13–17 August 2016, San Francisco, CA, 785–794. https://doi.org/10.1145/2939672.2939785

Chen, T.; He, T.; Benesty, M.; Khotilovich, V.; Tang, Y.; Cho, H.; Chen, K. Xgboost: Extreme Gradient Boosting. *R Packag. version 0.4-2* 2015, *1*, 1–4.

Chen, Y.; Li, L.; Lu, D.; Li, D. Exploring Bamboo Forest Aboveground Biomass Estimation Using Sentinel-2 Data. *Remote Sens. 2019, Vol. 11, Page 7* 2018, *11*, 7, doi:10.3390/RS11010007.

Irons, James R., John L. Dwyer, and Julia A. Barsi. 2012. "The next Landsat Satellite: The Landsat Data Continuity Mission." Remote Sensing of Environment 122 (July): 11–21. https://doi.org/10.1016/j.rse.2011.08.026.

Jain, A.K.; Mao, J.; Mohiuddin, K.M. Artificial Neural Networks: A Tutorial. *Computer (Long. Beach. Calif).* 1996, *29*, 31–44, doi:10.1109/2.485891.

Jiang, B., Zhao, Z., Wang, H. *et al.* SCE-BiLSTM: a hybrid deep learning model for regional forest biomass estimation with spatial-channel attention and extreme learning. *SIViP* **19**, 805 (2025). https://doi.org/10.1007/s11760-025-04414-y

Jiao, Z.; Dong, Y.; Schaaf, C.B.; Chen, J.M.; Román, M.; Wang, Z.; Zhang, H.; Ding, A.; Erb, A.; Hill, M.J.; et al. An Algorithm for the Retrieval of the Clumping Index (CI) from the MODIS BRDF Product Using an Adjusted Version of the Kernel-Driven BRDF Model. *Remote Sens. Environ.* 2018, *209*, 594–611, doi:10.1016/J.RSE.2018.02.041.

Karra, Kontgis, et al. "Global land use/land cover with Sentinel-2 and deep learning." IGARSS 2021-2021 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2021.

Kaufman, Y.J. and Tanré, D. 1992. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. IEEE Transactions on Geoscience and Remote Sensing, 30(2), 261–270. https://doi.org/10.1109/36.134076

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q. and Liu, T.Y. 2017. LightGBM: a highly efficient gradient boosting decision tree. In: Advances in Neural Information Processing Systems 30 (NIPS 2017). Long Beach, CA: Curran Associates, 3146–3154.

Koerner, L.M., Chadwick, M.A. and Tebbs, E.J. 2022. Mapping invasive strawberry guava (Psidium cattleianum) in tropical forests of Mauritius with Sentinel-2 and machine learning. International Journal of Remote Sensing, 43(3), 841–872. https://doi.org/10.1080/01431161.2021.2020364

Kort, J.; Collins, M.; Ditsch, D. A Review of Soil Erosion Potential Associated with Biomass Crops. *Biomass and Bioenergy* 1998, *14*, 351–359, doi:10.1016/S0961-9534(97)10071-X.

Kovárník, R. and Janová, J. 2025. Validation of Sentinel-2 based machine learning models for Czech National Forest Inventory. Ecological Informatics, 87, 103133. https://doi.org/10.1016/j.ecoinf.2025.103133

Kross, A.; McNairn, H.; Lapen, D.; Sunohara, M.; Champagne, C. Assessment of RapidEye Vegetation Indices for Estimation of Leaf Area Index and Biomass in Corn and Soybean Crops. Int. J. Appl. Earth Obs. Geoinf. 2015, 34, 235–248, doi:10.1016/J.JAG.2014.08.002.

Krstajić, D., Buturović, L.J., Leahy, D.E. and Thomas, S. 2014. Cross-validation pitfalls when selecting and assessing regression and classification models. Journal of Cheminformatics, 6(1), 10. https://doi.org/10.1186/1758-2946-6-10

Kučera, M., & Adolt, R. (2019). *Národní inventarizace lesů v České republice: výsledky druhého cyklu 2011–2015.* Ústav pro hospodářskou úpravu lesů Brandýs nad Labem.

Kumar, D.; Shekhar, S. Statistical Analysis of Land Surface Temperature–Vegetation Indexes Relationship through Thermal Remote Sensing. Ecotoxicol. Environ. Saf. 2015, 121, 39–44, doi:10.1016/J.ECOENV.2015.07.004.

Kundu, S.; Aggarwal, S.P.; Kingma, N.; Mondal, A.; Khare, D. Flood Monitoring Using Microwave Remote Sensing in a Part of Nuna River Basin, Odisha, India. Nat. Hazards 2015, 76, 123–138, doi:10.1007/S11069-014-1478-8/TABLES/3.

Lang, N., Jetz, W., Schindler, K. and Wegner, J.D. 2023. A high-resolution canopy height model of the Earth. Nature Ecology & Evolution, 7(11), 1778–1789. https://doi.org/10.1038/s41559-023-02206-6

Laurin, G.V.; Balling, J.; Corona, P.; Mattioli, W.; Papale, D.; Puletti, N.; Rizzo, M.; Truckenbrodt, J.; Gaia, M.U.; Laurin, V.; et al. Above-Ground Biomass Prediction by Sentinel-1 Multitemporal Data in Central Italy with Integration of ALOS2 and Sentinel-2 Data. https://doi.org/10.1117/1.JRS.12.016008 2018, 12, 016008, doi:10.1117/1.JRS.12.016008.

Le Rest, K.; Pinaud, D.; Monestiez, P.; Chadoeuf, J.; Bretagnolle, V. Spatial Leave-One-out Cross-Validation for Variable Selection in the Presence of Spatial Autocorrelation. *Glob. Ecol. Biogeogr.* 2014, *23*, 811–820, doi:10.1111/GEB.12161/SUPPINFO.

Li, C.; Zhou, L.; Xu, W. Estimating Aboveground Biomass Using Sentinel-2 MSI Data and Ensemble Algorithms for Grassland in the Shengjin Lake Wetland, China. Remote Sens. 2021, Vol. 13, Page 1595 2021, 13, 1595, doi:10.3390/RS13081595.

Li, F.; Hao, D.; Zhu, Q.; Yuan, K.; Braghiere, R.K.; He, L.; Luo, X.; Wei, S.; Riley, W.J.; Zeng, Y.; et al. Vegetation Clumping Modulates Global Photosynthesis through Adjusting Canopy Light Environment. *Glob. Chang. Biol.* 2023, *29*, 731–746, doi:10.1111/GCB.16503.

Liang, S.; Li, X.; Wang, J. Advanced Remote Sensing; Academic Press, 2012; ISBN 9780123859549.

Liepiņš, J., Lazdiņš, A. and Liepiņš, K. 2018. Equations for estimating above- and belowground biomass of Norway spruce, Scots pine, birch spp. and European aspen in Latvia. Scandinavian Journal of Forest Research, 33(1), 58–70. https://doi.org/10.1080/02827581.2017.1337923

Lin, Chinsu, Chao Cheng Wu, Khongor Tsogt, Yen Chieh Ouyang, and Chein I. Chang. 2015. "Effects of Atmospheric Correction and Pansharpening on LULC Classification Accuracy Using WorldView-2 Imagery." Information Processing in Agriculture 2 (1): 25–36. https://doi.org/10.1016/j.inpa.2015.01.003.

Liu, S.; Su, X.; Dong, S.; Cheng, F.; Zhao, H.; Wu, X.; Zhang, X.; Li, J. Modeling Aboveground Biomass of an Alpine Desert Grassland with SPOT-VGT NDVI. 2015, 52, 680–699, doi:10.1080/15481603.2015.1080143.

Loboda, T. V.; Giglio, L.; Boschetti, L.; Justice, C.O. Regional Fire Monitoring and Characterization Using Global NASA MODIS Fire Products in Dry Lands of Central Asia. Front. Earth Sci. 2012 62 2012, 6, 196–205, doi:10.1007/S11707-012-0313-3.

Lovett, G.M.; Jones, C.G.; Turner, M.G.; Weathers, K.C. Ecosystem Function in Heterogeneous Landscapes. In Ecosystem Function in Heterogeneous Landscapes; Lovett, G.M., Turner, M.G., Jones, C.G., Weathers, K.C., Eds.; Springer: New York, NY, USA, 2006; pp. 1–4.

Lovynska, V.; Buchavyi, Y.; Lakyda, P.; Sytnyk, S.; Gritzan, Y.; Sendziuk, R. Assessment of Pine Aboveground Biomass within Northern Steppe of Ukraine Using Sentinel-2 Data. J. For. Sci. 2020, 66 (2020), 339–348, doi:10.17221/28/2020-JFS.

Lucas, Richard, Katie Medcalf, Alan Brown, Peter Bunting, Johanna Breyer, Dan Clewley, Steve Keyworth, and Philippa Blackmore. 2011. "Updating the Phase 1 Habitat Map of Wales, UK, Using Satellite Sensor Data." ISPRS Journal of Photogrammetry and Remote Sensing 66 (1): 81–102. https://doi.org/10.1016/j.isprsjprs.2010.09.004.

Lukeš P., Strejček R., Křístek, Š., Mlčoušek, M. Forest Health Assessment in Czech Republic Using Sentinel-2 Satellite Data. Certified Methodology. 2018, 65.

Lyons, K.G.; Brigham, C.A.; Traut, B.H.; Schwartz, M.W. Rare Species and Ecosystem Functioning. *Conserv. Biol.* 2005, *19*, 1019–1024.

Ma, L., Hurtt, G.C., Tang, H., Lamb, R., Lister, A.J., Chini, L.P., Dubayah, R.O., Armston, J., Campbell, E., Duncanson, L., Healey, S.P., O'Neil-Dunne, J., Ott, L., Poulter, B. and Shen, Q. 2023. Global forest aboveground carbon stocks and fluxes from GEDI and ICESat-2, 2018–2021. Oak Ridge, TN: ORNL DAAC. https://doi.org/10.3334/ORNLDAAC/2180

Ma, X.; Huete, A.; Yu, Q.; Restrepo-Coupe, N.; Beringer, J.; Hutley, L.B.; Kanniah, K.D.; Cleverly, J.; Eamus, D. Parameterization of an Ecosystem Light-Use-Efficiency Model for Predicting Savanna GPP Using MODIS EVI. Remote Sens. Environ. 2014, 154, 253–271, doi:10.1016/J.RSE.2014.08.025.

Macedo, F.L.; Sousa, A.M.O.; Gonçalves, A.C.; Marques da Silva, J.R.; Mesquita, P.A.; Rodrigues, R.A.F. Above-Ground Biomass Estimation for Quercus Rotundifolia Using Vegetation Indices Derived from High Spatial Resolution Satellite Images. https://doi.org/10.1080/22797254.2018.1521250 2018, 51, 932–944, doi:10.1080/22797254.2018.1521250.

Maesano M, Santopuoli G, Moresi FV, Matteucci G, Lasserre B, Scarascia Mugnozza G (2022). Above ground biomass estimation from UAV high resolution RGB images and LiDAR data in a pine forest in Southern Italy. iForest 15: 451-457. - doi: 10.3832/ifor3781-015

Malhi, Y.; Doughty, C.E.; Goldsmith, G.R.; Metcalfe, D.B.; Girardin, C.A.J.; Marthews, T.R.; del Aguila-Pasquel, J.; Aragão, L.E.O.C.; Araujo-Murakami, A.; Brando, P.; et al. The Linkages between Photosynthesis, Productivity, Growth and Biomass in Lowland Amazonian Forests. *Glob. Chang. Biol.* 2015, *21*, 2283–2295, doi:10.1111/GCB.12859.

Mananze, S.; Pôças, I.; Cunha, M. Retrieval of Maize Leaf Area Index Using Hyperspectral and Multispectral Data. Remote Sens. 2018, Vol. 10, Page 1942 2018, 10, 1942, doi:10.3390/RS10121942.

MÁSLO, Jan, Radim ADOLT, Ivo KOHN a Miloš KUČERA. Nadzemní biomasa a zásoba uhlíku v ČR: výsledky třetího cyklu Národní inventarizace lesů 2016–2020. Lesnická práce: časopis pro lesnicko-dřevařskou vědu a praxi. 2024, roč. 103, č. 3, s. 42.

MÁSLO, Jan, Radim ADOLT, Ivo KOHN a Miloš KUČERA. Plocha lesa v ČR: výsledky třetího cyklu Národní inventarizace lesů 2016–2020. Lesnická práce: časopis pro lesnicko-dřevařskou vědu a praxi. 2023, roč. 102, č. 6, s. 42.

McRoberts, R.E. 2010. Probability- and model-based approaches to inference for proportion forest using satellite imagery as ancillary data. Remote Sensing of Environment, 114(5), 1017–1025. https://doi.org/10.1016/j.rse.2009.12.013

MEA. Millennium Ecosystem Assessment. Ecosystems and Human Well-Being; Island Press: Washington, DC, USA, 2005.

Meyer, H., & Pebesma, E. (2021). Predicting into unknown space? Estimating the area of applicability of spatial prediction models. *Methods in Ecology and Evolution, 12*(9), 1620–1633. doi:10.1111/2041-210X.13650.

Meyer, S.T.; Koch, C.; Weisser, W.W. Towards a Standardized Rapid Ecosystem Function Assessment (REFA). Trends Ecol. Evol. 2015, 30, 390–397, doi:10.1016/J.TREE.2015.04.006.

Moradi, F.; Mohammad Moein Sadeghi, S.; Beygi Heidarlou, H.; Deljouei, A.; Boshkar, E.; Alexandru Borz, S.; M M, S.S.; B, H.H. Above-Ground Biomass Estimation in a Mediterranean Sparse Coppice Oak Forest Using Sentinel-2 Data. *Ann. For. Res* 2022, *65*, 165–182, doi:10.15287/afr.2022.2390.

Muhd-Ekhzarizal, M.E.; Mohd-Hasmadi, I.; Hamdan, O.; Mohamad-Roslan, M.K.; Noor-Shaila, S. Estimation of Aboveground Biomass in Mangrove Forests Using Vegetation Indices from SPOT-5 Image. J. Trop. For. Sci. 2018, 30, 224–233, doi:10.26525/JTFS2018.30.2.224233.

Mücher, C. A., Hennekens, S. M., Bunce, R. G. H., Schaminée, J. H. J., & Schaepman, M. E. (2009). Modelling the spatial distribution of Natura 2000 habitats across Europe. Landscape and Urban Planning, 92(2), 148–159. https://doi.org/10.1016/J.LANDURBPLAN.2009.04.003

Nagendra, H.; Lucas, R.; Pradinho Honrado, J.; Jongman, R.H.; Tarantino, C.; Adamo, M.; Mairota, P. Remote Sensing for Conservation Monitoring: Assessing Protected Areas, Habitat Extent, Habitat Condition, Species Diversity, and Threats. Ecol. Indic. 2013, 33, 45–59, doi:10.1016/j.ecolind.2012.09.014.

NASA Landsat Science Available online: https://landsat.gsfc.nasa.gov/ [08-08-2022].

NASA MODIS – Moderate Resolution Imaging Spectroradiometer Available online: https://modis.gsfc.nasa.gov/ [08-08-2022].

NASA TERRA – The EOS Flagship Available online: https://terra.nasa.gov/ [08-08-2022].

Neumann, C. Habitat Sampler—A Sampling Algorithm for Habitat Type Delineation in Remote Sensing Imagery. Divers. Distrib. 2020, 26, 1752–1766, doi:10.1111/DDI.13165.

Noble, W.S. What Is a Support Vector Machine? *Nat. Biotechnol.* 2006, *24*, 1565–1567, doi:10.1038/nbt1206-1565.

Nock, C.A.; Vogt, R.J.; Beisner, B.E. Functional Traits. eLS 2016, 1–8, doi:10.1002/9780470015902.A0026282.

Nuthammachot, N.; Askar, A.; Stratoulias, D.; Wicaksono, P. Combined Use of Sentinel-1 and Sentinel-2 Data for Improving above-Ground Biomass Estimation. https://doi.org/10.1080/10106049.2020.1726507 2020, 37, 366–376, doi:10.1080/10106049.2020.1726507.

O'Riordan, R.; Davies, J.; Stevens, C.; Quinton, J.N.; Boyko, C. The Ecosystem Services of Urban Soils: A Review. *Geoderma* 2021, *395*, 115076, doi:10.1016/J.GEODERMA.2021.115076.

Oliver, T.H.; Heard, M.S.; Isaac, N.J.B.; Roy, D.B.; Procter, D.; Eigenbrod, F.; Freckleton, R.; Hector, A.; Orme, C.D.L.; Petchey, O.L.; et al. Biodiversity and Resilience of Ecosystem Functions. Trends Ecol. Evol. 2015, 30, 673–684, doi:10.1016/J.TREE.2015.08.009.

Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., Canadell, J.G., Ciais, P., Jackson, R.B., Pacala, S.W., McGuire, A.D., Piao, S., Rautiainen, A., Sitch, S. and Hayes, D. 2011. A large and persistent carbon sink in the world's forests. Science, 333(6045), 988–993. https://doi.org/10.1126/science.1201609

Pechanec, V.; Štěrbová, L.; Purkyt, J.; Prokopová, M.; Včeláková, R.; Cudlín, O.; Vyvlečka, P.; Cienciala, E.; Cudlín, P. Selected Aspects of Carbon Stock Assessment in Aboveground Biomass. *L. 2022, Vol. 11, Page 66* 2022, *11*, 66, doi:10.3390/LAND11010066.

Petter, M.; Mooney, S.; Maynard, S.M.; Davidson, A.; Cox, M.; Horosak, I. A Methodology to Map Ecosystem Functions to Support Ecosystem Services Assessments. *Ecol. Soc.* 2012, *18*, doi:10.5751/ES-05260-180131.

Pettorelli, N., Laurance, W.F., O'Brien, T.G., Wegmann, M., Nagendra, H. and Turner, W. (2014), Satellite remote sensing for applied ecologists: opportunities and challenges. J Appl Ecol, 51: 839-848. https://doi.org/10.1111/1365-2664.12261

Pettorelli, N.; Schulte to Bühne, H.; Tulloch, A.; Dubois, G.; Macinnis-Ng, C.; Queirós, A.M.; Keith, D.A.; Wegmann, M.; Schrodt, F.; Stellmes, M.; et al. Satellite Remote Sensing of Ecosystem Functions: Opportunities, Challenges and Way Forward. Remote Sens. Ecol. Conserv. 2018, 4, 71–93, doi:10.1002/rse2.59.

PLÍVA, Karel. *Typologický klasifikační systém ÚHÚL*. Brandýs nad Labem: Ústav pro hospodářskou úpravu lesů, 1987.

Ploton, P.; Mortier, F.; Réjou-Méchain, M.; Barbier, N.; Picard, N.; Rossi, V.; Dormann, C.; Cornu, G.; Viennois, G.; Bayol, N.; et al. Spatial Validation Reveals Poor Predictive Performance of Large-Scale Ecological Mapping Models. *Nat. Commun. 2020 111* 2020, *11*, 1–11, doi:10.1038/s41467-020-18321-y.

Potapov, P.; Li, X.; Hernandez-Serna, A.; Tyukavina, A.; Hansen, M.C.; Kommareddy, A.; Pickens, A.; Turubanova, S.; Tang, H.; Silva, C.E.; et al. Mapping Global Forest Canopy Height through Integration of GEDI and Landsat Data. *Remote Sens. Environ.* 2021, *253*, 112165, doi:10.1016/J.RSE.2020.112165.

Puliti, S., Hauglin, M., Breidenbach, J., Montesano, P., Neigh, C.S.R., Rahlf, J., Solberg, S., Klingenberg, T.F. and Astrup, R. 2020. Modelling above-ground biomass stock over Norway using national forest inventory data with ArcticDEM and Sentinel-2 data. Remote Sensing of Environment, 236, 111501. https://doi.org/10.1016/j.rse.2019.111501

Puliti, S.; Breidenbach, J.; Schumacher, J.; Hauglin, M.; Klingenberg, T.F.; Astrup, R. Above-Ground Biomass Change Estimation Using National Forest Inventory Data with Sentinel-2 and Landsat. Remote Sens. Environ. 2021, 265, 112644, doi:10.1016/J.RSE.2021.112644.

Qi, J.; Chehbouni, A.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. 1994. A modified soil adjusted vegetation index. *Remote Sensing of Environment*, 48(2), 119–126,DOI:10.1016/0034-4257(94)90134-1

Rabin, S.S.; Melton, J.R.; Lasslop, G.; Bachelet, D.; Forrest, M.; Hantson, S.; Kaplan, J.O.; Li, F.; Mangeon, S.; Ward, D.S.; et al. The Fire Modeling Intercomparison Project (FireMIP), Phase 1: Experimental and Analytical Protocols with Detailed Model Descriptions. *Geosci. Model Dev.* 2017, *10*, 1175–1197, doi:10.5194/GMD-10-1175-2017.

Rajput, P. Machine learning approach for forest biomass modelling with in-situ and remote sensing data in Narmadapuram central India. *Model. Earth Syst. Environ.* **11**, 350 (2025). https://doi.org/10.1007/s40808-025-02527-4

Richards, F. S. G. 1961. "A Method of Maximum-Likelihood Estimation." Journal of the Royal Statistical Society: Series B (Methodological) 23 (2): 469–75. https://doi.org/10.1111/j.2517-6161.1961.tb00430.x.

Roberts, D.R.; Bahn, V.; Ciuti, S.; Boyce, M.S.; Elith, J.; Guillera-Arroita, G.; Hauenstein, S.; Lahoz-Monfort, J.J.; Schröder, B.; Thuiller, W.; et al. Cross-Validation Strategies for Data with Temporal, Spatial, Hierarchical, or Phylogenetic Structure. *Ecography (Cop.).* 2017, *40*, 913–929, doi:10.1111/ECOG.02881.

Rodríguez-Veiga, P.; Quegan, S.; Carreiras, J.; Persson, H.J.; Fransson, J.E.S.; Hoscilo, A.; Ziółkowski, D.; Stereńczak, K.; Lohberger, S.; Stängel, M.; et al. Forest Biomass Retrieval Approaches from Earth Observation in Different Biomes. Int. J. Appl. Earth Obs. Geoinf. 2019, 77, 53–68, doi:10.1016/J.JAG.2018.12.008.

Rossini, M.; Cogliati, S.; Meroni, M.; Migliavacca, M.; Galvagno, M.; Busetto, L.; Cremonese, E.; Julitta, T.; Siniscalco, C.; Morra Di Cella, U.; et al. Remote Sensing-Based Estimation of Gross Primary Production in a Subalpine Grassland. Biogeosciences 2012, 9, 2565–2584, doi:10.5194/BG-9-2565-2012.

Running, S.W.; Zhao, M. User's Guide Daily GPP and Annual NPP (MOD17A2/A3) Products NASA Earth Observing System MODIS Land Algorithm. 2015.

Santoro, M. GlobBiomass - Global Datasets of Forest Biomass 2018.

Santoro, M.; Cartus, O. (2025): ESA Biomass Climate Change Initiative (Biomass_cci): Global datasets of forest above-ground biomass for the years 2007, 2010, 2015, 2016, 2017, 2018, 2019, 2020, 2021 and 2022, v6.0. NERC EDS Centre for Environmental Data Analysis, *17 April 2025*. doi:10.5285/95913ffb6467447ca72c4e9d8cf30501.

Santoro, M.; Cartus, O. ESA Biomass Climate Change Initiative (Biomass_cci): Global Datasets of Forest above-Ground Biomass for the Years 2010, 2017 and 2018, V3. *NERC EDS Cent. Environ. Data Anal.* 2021, doi:10.5285/5f331c418e9f4935b8eb1b836f8a91b8.

Santos, M.J.; Rosalino, L.M.; Santos-Reis, M.; Ustin, S.L. Testing Remotely-Sensed Predictors of Meso-Carnivore Habitat Use in Mediterranean Ecosystems. Landsc. Ecol. 2016, 31, 1763–1780, doi:10.1007/S10980-016-0360-3.

Senay, G.B.; Velpuri, N.M.; Alemu, H.; Md Pervez, S.; Asante, K.O.; Kariuki, G.; Taa, A.; Angerer, J. Establishing an Operational Waterhole Monitoring System Using Satellite Data and Hydrologic Modelling: Application in the Pastoral Regions of East Africa. Pastoralism 2013, 3, 1–16, doi:10.1186/2041-7136-3-20/FIGURES/9.

Schwaab, J.; Meier, R.; Mussetti, G.; Seneviratne, S.; Bürgi, C.; Davin, E.L. The Role of Urban Trees in Reducing Land Surface Temperatures in European Cities. *Nat. Commun. 2021 121* 2021, *12*, 1–11, doi:10.1038/s41467-021-26768-w.

Schwartz, M., Ciais, P., De Truchis, A., Chave, J., Ottlé, C., Vega, C., Wigneron, J.-P., Nicolas, M., Jouaber, S., Liu, S., Brandt, M., and Fayad, I.: FORMS: Forest Multiple Source height, wood volume, and biomass maps in France at 10 to 30 m resolution based on Sentinel-1, Sentinel-2, and Global Ecosystem Dynamics Investigation (GEDI) data with a deep learning approach, Earth Syst. Sci. Data, 15, 4927–4945, https://doi.org/10.5194/essd-15-4927-2023, 2023.

Souza, M.C.; Biudes, M.S.; Danelichen, V.H.D.M.; Machado, N.G.; De Musis, C.R.; Vourlitis, G.L.; Nogueira, J. de S. Estimation of Gross Primary Production of the Amazon-Cerrado Transitional Forest by Remote Sensing Techniques. *Rev. Bras. Meteorol.* 2014, *29*, 01–12, doi:10.1590/S0102-77862014000100001.

Snoek, J.; Larochelle, H.; Adams, R. P. Practical Bayesian Optimization of Machine Learning Algorithms. In: Advances in Neural Information Processing Systems 25. 2012, s. 2951–2959.

Storey, James, David P. Roy, Jeffrey Masek, Ferran Gascon, John Dwyer, and Michael Choate. 2016. “A Note on the Temporary Misregistration of Landsat-8 Operational Land Imager (OLI) and Sentinel-2 Multi Spectral Instrument (MSI) Imagery.” Remote Sensing of Environment 186 (December): 121–22. https://doi.org/10.1016/j.rse.2016.08.025.

Streutker, D.R.; Glenn, N.F. LiDAR Measurement of Sagebrush Steppe Vegetation Heights. Remote Sens. Environ. 2006, 102, 135–145, doi:10.1016/J.RSE.2006.02.011.

Suwanprasit, Chanida, and Naiyana Srichai. 2012. “Impacts of Spatial Resolution on Land Cover Classification.” Proceedings of the Asia-Pacific Advanced Network 33 (0): 39. https://doi.org/10.7125/apan.33.4.

Tagesson, T., Schurgers, G., Horion, S., Ciais, P., Tian, F., Brandt, M., Ahlström, A., Wigneron, J.P., Ardö, J., Olin, S., Fan, L., Wu, Z. and Fensholt, R. 2020. Recent divergence in the contributions of tropical and boreal forests to the terrestrial carbon sink. Nature Ecology & Evolution, 4(2), 202–209. https://doi.org/10.1038/s41559-019-1090-0

Tallis, H.; Ricketts, T.H.; Daily, G.C.; Polasky, S. *Natural Capital: Theory and Practice of Mapping Ecosystem Services*; Oxford University Press, 2011; ISBN 0199588996.

Tao, X.; Liang, S.; He, T.; Jin, H. Estimation of Fraction of Absorbed Photosynthetically Active Radiation from Multiple Satellite Data: Model Development and Validation. Remote Sens. Environ. 2016, 184, 539–557, doi:10.1016/J.RSE.2016.07.036.

TEEB. The Economics of Ecosystems and Biodiversity. Mainstreaming the Economics of Nature: A Synthesis of the Approach, Conclusions and Recommendations of TEEB; TEEB: Bonn, Germany, 2010.

Terminologická komise Českého úřadu zeměměřického a katastrálního. 2021. Terminologický slovník zeměměřictví a katastru nemovitostí. Dostupné z: https://www.slovnikcuzk.eu/hledej.php [cit. 2021-11-05].

Thanh Noi, Phan, and Martin Kappas. 2018. “Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery.” Sensors (Basel, Switzerland) 18 (1): 18. https://doi.org/10.3390/s18010018.

Tian, Xin, Jiejie Li, Fanyi Zhang, Haibo Zhang, and Mi Jiang. 2024. "Forest Aboveground Biomass Estimation Using Multisource Remote Sensing Data and Deep Learning Algorithms: A Case Study over Hangzhou Area in China" *Remote Sensing* 16, no. 6: 1074. https://doi.org/10.3390/rs16061074

Tolan, J., Yang, H.I., Nosarzewski, B., Couairon, G., Vo, H.V., Brandt, J., Spore, J., Majumdar, S., Haziza, D., Vamaraju, J., Moutakanni, T., Bojanowski, P., Johns, T., White, B., Tiecke, T. and Couprie, C. 2024. Very high resolution canopy height maps from RGB imagery using self-supervised vision transformer and convolutional decoder trained on aerial lidar. Remote Sensing of Environment, 300, 113888. https://doi.org/10.1016/j.rse.2023.113888

Tomasi, C., & Manduchi, R. (1998, January). Bilateral filtering for gray and color images. In *Sixth international conference on computer vision (IEEE Cat. No. 98CH36271)* (pp. 839-846). IEEE.

Turkelboom, F.; Raquez, P.; Dufrêne, M.; Raes, L.; Simoens, I.; Jacobs, S.; Stevens, M.; De Vreese, R.; Panis, J.A.E.; Hermy, M.; et al. CICES Going Local: Ecosystem Services Classification Adapted for a Highly Populated Country. Ecosyst. Serv. Glob. Issues, Local Pract. 2013, 223–247, doi:10.1016/B978-0-12-419964-4.00018-4.

U.S. GEOLOGICAL SURVEY (USGS). *Annual National Land Cover Database* [online]. Earth Resources Observation and Science (EROS) Center, 26 September 2024 [cit. 2026-02-27]. Dostupné z: https://www.usgs.gov/centers/eros/science/annual-national-land-cover-database

Ústav pro hospodářskou úpravu lesů. 2024. *Národní inventarizace lesů – třetí cyklus 2016–2020*. Brandýs nad Labem: Ústav pro hospodářskou úpravu lesů Brandýs nad Labem.

Vaglio Laurin, G., Puletti, N., Hawthorne, W., Liesenberg, V., Corona, P., Papale, D., Chen, Q. and Valentini, R. 2016. Discrimination of tropical forest types, dominant species, and mapping of functional guilds by hyperspectral and simulated multispectral Sentinel-2 data. Remote Sensing of Environment, 176, 163–176. https://doi.org/10.1016/j.rse.2016.01.017

Van Der Meer, F.; Bakker, W.; Scholte, K.; Skidmore, A.; De Jong, S.; Clevers, J.; Epema, G. VEGETATION INDICES, ABOVE GROUND BIOMASS ESTIMATES AND THE RED EDGE FROM MERIS. 2000.

VIEWEGH, J. Klasifikace lesních rostlinných společenstev: se zaměřením na Typologický systém ÚHÚL. Praha: Česká zemědělská univerzita, 2003.

Vyvlečka, P.; Pechanec, V. 2023. Optical remote sensing in provisioning of ecosystem-functions analysis: review. Sensors, 23(10), 4937. https://doi.org/10.3390/s23104937

Vyvlečka, P.; Pechanec, V. 2026. National-scale mapping of temperate forest above-ground biomass in the Czech Republic using Sentinel-1 and Sentinel-2 data. European Journal of Remote Sensing, 59(1), 2701998. https://doi.org/10.1080/22797254.2026.2701998

Wagner, W.; Sabel, D.; Doubkova, M.; Bartsch, A.; Pathe, C. THE POTENTIAL OF SENTINEL-1 FOR MONITORING SOIL MOISTURE WITH A HIGH SPATIAL RESOLUTION AT GLOBAL SCALE. *ESA Spec. Publ.* 2010.

Wai, P., Su, H. and Li, M. 2022. Estimating aboveground biomass of two different forest types in Myanmar from Sentinel-2 data with machine learning and geostatistical algorithms. Remote Sensing, 14(9), 2146. https://doi.org/10.3390/rs14092146

Wang, G.; Liu, S.; Liu, T.; Fu, Z.; Yu, J.; Xue, B. Modelling Above-Ground Biomass Based on Vegetation Indexes: A Modified Approach for Biomass Estimation in Semi-Arid Grasslands. https://doi.org/10.1080/01431161.2018.1553319 2018, 40, 3835–3854, doi:10.1080/01431161.2018.1553319.

Wang, J.; Xiao, X.; Bajgain, R.; Starks, P.; Steiner, J.; Doughty, R.B.; Chang, Q. Estimating Leaf Area Index and Aboveground Biomass of Grazing Pastures Using Sentinel-1, Sentinel-2 and Landsat Images. ISPRS J. Photogramm. Remote Sens. 2019, 154, 189–201, doi:10.1016/J.ISPRSJPRS.2019.06.007.

Wang, L., W. P. Sousa, and P. Gong. 2004. "Integration of Object-Based and Pixel-Based Classification for Mapping Mangroves with IKONOS Imagery." International Journal of Remote Sensing 25 (24): 5655–68. https://doi.org/10.1080/014311602331291215.

Wegler, M., Kacic, P., Thonfeld, F., Holzwarth, S., Jaggy, N., Gessner, U. and Kuenzer, C. 2025. Tree species from space: a new product for Germany based on Sentinel-1 and -2 time series. International Journal of Remote Sensing. https://doi.org/10.1080/01431161.2025.2530236

Wei, S.; Fang, H.; Schaaf, C.B.; He, L.; Chen, J.M. Global 500 m Clumping Index Product Derived from MODIS BRDF Data (2001–2017). *Remote Sens. Environ.* 2019, *232*, 111296, doi:10.1016/J.RSE.2019.111296.

Weih, Robert C, and Norman D Riggan. 2010. “Object-Based Classification vs. Pixel-Based Classification: Comparative Importance of Multi-Resolution Imagery.” Remote Sensing and Spatial Information Sciences.

Weiss, M.; Jacob, F.; Duveiller, G. Remote Sensing for Agricultural Applications: A Meta-Review. Remote Sens. Environ. 2020, 236, doi:10.1016/J.RSE.2019.111402.

Xu, W., Zhang, B., Xu, Q., Gao, D., Zuo, H., Ren, R., Diao, K. and Chen, Z. 2024. Enhanced carbon storage in mixed coniferous and broadleaf forest compared to pure forest in the north subtropical–warm temperate transition zone of China. Forests, 15(9), 1520. https://doi.org/10.3390/f15091520

Yan, J.; Ma, Y.; Zhang, D.; Li, Z.; Zhang, W.; Wu, Z.; Wang, H.; Wen, L.; Yan, J.; Ma, Y.; et al. High-Resolution Monitoring and Assessment of Evapotranspiration and Gross Primary Production Using Remote Sensing in a Typical Arid Region. L. 2021, Vol. 10, Page 396 2021, 10, 396, doi:10.3390/LAND10040396.

Yu, Y.; Pan, Y.; Yang, X.; Fan, W. Spatial Scale Effect and Correction of Forest Aboveground Biomass Estimation Using Remote Sensing. *Remote Sens. 2022, Vol. 14, Page 2828* 2022, *14*, 2828, doi:10.3390/RS14122828.

Zanaga, D.; Van De Kerchove, R.; Daems, D.; De Keersmaecker, W.; Brockmann, C.; Kirches, G.; Wevers, J.; Cartus, O.; Santoro, M.; Fritz, S.; et al. ESA WorldCover 10 m 2021 V200. 2022, doi:10.5281/ZENODO.7254221.

Zhang, Y.; Chen, J.M.; Miller, J.R.; Noland, T.L. Leaf Chlorophyll Content Retrieval from Airborne Hyperspectral Remote Sensing Imagery. Remote Sens. Environ. 2008, 112, 3234–3247, doi:10.1016/J.RSE.2008.04.005.

Zheng, G.; Moskal, L.M. Leaf Orientation Retrieval from Terrestrial Laser Scanning (TLS) Data. IEEE Trans. Geosci. Remote Sens. 2012, 50, 3970–3979, doi:10.1109/TGRS.2012.2188533.

Zhou, J., Zan, M., Zhai, L., Yang, S., Xue, C., Li, R. and Wang, X. 2025. Remote sensing estimation of aboveground biomass of different forest types in Xinjiang based on machine learning. Scientific Reports, 15(1), 6187. https://doi.org/10.1038/s41598-025-90906-3