

Lesnická a dřevařská fakulta
Mendelova univerzita v Brně

BILANCE UHLÍKU V LESNÍCH POROSTECH VE VZTAHU K LESNICKÉMU MANAGEMENTU

Nmap – Specializovaná mapa s odborným obsahem

Výsledek projektu NAZV QK21010198 (Adaptace lesního hospodářství pro
udržitelné využívání přírodních zdrojů)

Brno 2026



BILANCE UHLÍKU V LESNÍCH POROSTECH VE VZTAHU K LESNICKÉMU MANAGEMENTU

Kolektiv autorů:

Ing. Aleš Kučera, Ph.D.*

Ing. Martin Valtera, Ph.D.**

Ing. Zdeněk Patočka, Ph.D.***

Ing. Václav Zouhar, Ph.D.****

Ing. Jan Pecháček, Ph.D.****

doc. Dr. Ing. Jan Kadavý***

* Ústav geologie a pedologie, Lesnická a dřevařská fakulta, Mendelova univerzita v Brně. Zemědělská 3, 613 00 Brno

** Ústav lesnické botaniky, dendrologie a geobiocenologie, Lesnická a dřevařská fakulta, Mendelova univerzita v Brně. Zemědělská 3, 613 00 Brno

*** Ústav hospodářské úpravy lesa a aplikované geoinformatiky, Lesnická a dřevařská fakulta, Mendelova univerzita v Brně. Zemědělská 3, 613 00 Brno

**** Národní lesnický institut, pobočka Brno, Vrázova 1, 616 00 Brno-Žabovřesky

Podkladová data:

© Národní lesnický institut

Formát map: A3

© Mendelova univerzita v Brně, 2026

1. vydání



Obsah

1	NÁZEV MAPY	6
2	POPIS NOVOSTI MAPY	7
3	INFORMACE K ROZSAHU VYUŽITÍ MAPY	8
4	INFORMACE O PŘÍNOSECH MAPY PRO UŽIVATELE.....	9
5	SEZNAM ODBORNÝCH PODKLADŮ.....	10
5.1	SEZNAM PUBLIKACÍ PŘEDCHÁZEJÍCÍCH VÝSLEDKU	10
5.2	POUŽITÁ LITERATURA	10
6	DEDIKACE.....	13
7	JMÉNA OPONENTŮ A NÁZVY JEJICH ORGANIZACÍ.....	14
8	VLASTNÍ POPIS VÝSLEDKU.....	15
8.1	MATERIÁL A METODY.....	16
8.1.1	<i>Vstupní data</i>	<i>16</i>
8.1.2	<i>Zpracování dat.....</i>	<i>16</i>
8.1.3	<i>Ověření významu prediktorů zásob uhlíku ve svrchních půdních vrstvách: Regresní analýza</i>	<i>19</i>
8.1.4	<i>Prostorová analýza: distribuce agregovaných porostních typů</i>	<i>20</i>
8.1.5	<i>Scénář změněné dřevinné skladby ve prospěch listnáčů u vybraných CHS a agregovaných porostních typů s nižším zastoupením jehličnanů</i>	<i>23</i>
8.1.6	<i>Tvorba prostorových modelů a map zásob TOC.....</i>	<i>26</i>
8.2	VÝSLEDKY A DISKUSE	29
8.2.1	<i>Výsledky smíšených regresních modelů zásob uhlíku v závislosti na agregovaném porostním typu 29</i>	
8.2.2	<i>Výsledné prostorové modely zásoby a bilance uhlíku</i>	<i>30</i>
8.2.3	<i>Výsledné mapy zásoby a bilance uhlíku.....</i>	<i>32</i>
9	ZÁVĚR	35
10	PŘÍLOHY.....	36

Seznam obrázků

OBR. 1. (A) AKTUÁLNÍ DISTRIBUCE AGREGOVANÝCH POROSTNÍCH TYPŮ (APT); (B) KARTOGRAM ZASTOUPENÍ APT V JEDNOTLIVÝCH PLO. ČERNÉ LINIE OZNAČUJÍ HRANICE PLO, POTAŽMO ČR.....	21
OBR. 2. PROCENTUÁLNÍ ZASTOUPENÍ AGREGOVANÝCH POROSTNÍCH TYPŮ (APT) V RÁMCI CÍLOVÝCH HOSPODÁŘSKÝCH SOUBORŮ (CHS) – PRŮMĚRNÉ HODNOTY ZA ČESKOU REPUBLIKU. (A) HODNOTY AKTUÁLNÍHO STAVU (NIL3); (B) HODNOTY V PŘÍPADĚ „LISTNATÉHO“ SCÉNÁŘE JAKO PODKLAD PRO NAVAZUJÍCÍ TVORBU MODELU „LISTNATÉHO SCÉNÁŘE“	25
OBR. 3. (P1): QQ DIAGRAM RESIDUÍ SMÍŠENÉHO REGRESNÍHO MODELU ZÁSOB UHLÍKU V ORGANICKÉ VRSTVĚ O S AGREGOVANÝM POROSTNÍM TYPEM JAKO FIXNÍ PROMĚNNOU A PLO A CHS JAKO NÁHODNÝMI PROMĚNNÝMI. NA HORIZONTÁLNÍ OSE JSOU TEORETICKÉ KVANTILY, NA VERTIKÁLNÍ OSE JSOU RESIDUA MODELU (N = 1.120); MODRÝ PÁS ZNAČÍ 95% INTERVAL SPOLEHLIVOSTI NORMÁLNÍHO ROZDĚLENÍ (SHAPIRO-WILK TEST $P = 0,489$).....	58
OBR. 4. (P2): QQ DIAGRAM RESIDUÍ SMÍŠENÉHO MODELU ZÁSOB UHLÍKU V ORGANOMINERÁLNÍ VRSTVĚ A S AGREGOVANÝM POROSTNÍM TYPEM JAKO FIXNÍ PROMĚNNOU A PLO A CHS JAKO NÁHODNÝMI PROMĚNNÝMI. NA HORIZONTÁLNÍ OSE JSOU TEORETICKÉ KVANTILY, NA VERTIKÁLNÍ OSE JSOU RESIDUA MODELU (N = 893); MODRÝ PÁS ZNAČÍ 95% INTERVAL SPOLEHLIVOSTI NORMÁLNÍHO ROZDĚLENÍ (SHAPIRO-WILK TEST $P = 0,217$).	59

Seznam tabulek

TAB. 1. KLASIFIKACE AGREGOVANÝCH POROSTNÍCH TYPŮ NA ZÁKLADĚ RELATIVNÍHO ZASTOUPENÍ JEHLIČNANŮ VE STROMOVÉM PATŘE.	16
TAB. 2. PODKLAD PRO VÝPOČET ZASTOUPENÍ CHS V RÁMCI PLO PŘI KALKULACI „LISTNATÉHO“ SCÉNÁŘE. HODNOTY UDÁVAJÍ PROCENTUÁLNÍ ZŮSTATKY AGREGOVANÝCH POROSTNÍCH TYPŮ (APT) V PŘÍSLUŠNÝCH CHS PŘI UPLATNĚNÍ „LISTNATÉHO“ SCÉNÁŘE.	24
TAB. 3. PROCENTUÁLNÍ ZMĚNY V ZASTOUPENÍ AGREGOVANÝCH POROSTNÍCH TYPŮ (APT) V RÁMCI CÍLOVÝCH HOSPODÁŘSKÝCH SOUBORŮ (CHS) V „LISTNATÉM SCÉNÁŘI“. CELOREPUBLIKOVÝ PRŮMĚR.	25
TAB. 4. STŘEDNÍ HODNOTY ZÁSOB UHLÍKU [T HA ⁻¹] V ORGANICKÝCH VRSTVÁCH (O) DLE AGREGOVANÉHO POROSTNÍHO TYPU (APT) ($P < 0,001$; $F = 9,76$; $R^2M = 0,030$) S NÁHODNÝMI INTERCEPTY PRO PLO A CHS VE SMÍŠENÉM REGRESNÍM MODELU ($R^2C = 0,328$; $AIC = 5120$).	29
TAB. 5. STŘEDNÍ HODNOTY ZÁSOB UHLÍKU [T HA ⁻¹] VE SVRCHNÍ ORGANOMINERÁLNÍ VRSTVĚ (A) DLE AGREGOVANÉHO POROSTNÍHO TYPU (APT) ($P = 0,002$; $F = 4,195$; $R^2M = 0,019$) S NÁHODNÝMI INTERCEPTY PRO PLO A CHS VE SMÍŠENÉM REGRESNÍM MODELU ($R^2C = 0,253$; $AIC = 519$).	29
TAB. 6. VÝSLEDKY PREDIKCE ZÁSOB UHLÍKU POMOCÍ MODELŮ RANDOM FOREST. R^2 – KOEFICIENT DETERMINANCE NA TRÉNOVACÍCH DATECH; RMSE (ROOT MEAN SQUARE ERROR) – ODMOCNĚNÁ STŘEDNÍ KVADRATICKÁ CHYBA [T/HA]; MAE (MEAN ABSOLUTE ERROR) – CELKOVÁ STŘEDNÍ CHYBA PREDIKCE [T/HA]. 30	
TAB. 7. (P1): STŘEDNÍ HODNOTY ZÁSOB UHLÍKU [T HA ⁻¹] V ORGANICKÝCH VRSTVÁCH (O) DLE AGREGOVANÉHO POROSTNÍHO TYPU (APT) ($P < 0,001$; $F = 17,5$; $R^2M = 0,050$) S NÁHODNÝM INTERCEPTEM PRO PLO VE SMÍŠENÉM REGRESNÍM MODELU ($R^2C = 0,313$; $AIC = 5140$).	58
TAB. 8. (P2): STŘEDNÍ HODNOTY ZÁSOB UHLÍKU [T HA ⁻¹] VE SVRCHNÍ MINERÁLNÍ VRSTVĚ (A) DLE AGREGOVANÉHO POROSTNÍHO TYPU (APT) ($P < 0,001$; $F = 4,77$; $R^2M = 0,021$) S NÁHODNÝM INTERCEPTEM PRO CHS VE SMÍŠENÉM REGRESNÍM MODELU ($R^2C = 0,211$; $AIC = 526$).	59

1 Název mapy

Bilance uhlíku v lesních porostech ve vztahu k lesnickému managementu.

2 Popis novosti mapy

Výsledek (specializovaná mapa s odborným obsahem) popisuje pro zájmové území České republiky, které bylo předmětem projektového řešení, zásobu uhlíku ve vztahu k lesnickému managementu. Vzhledem k výrazné diverzitě přístupů, zaváděných s nástupem novodobých hospodářských opatření, vzhledem k disponibilitě dat o lesních půdách a vzhledem ke komplexitě tématu vyžadující nezbytnou generalizaci, je problematika prezentována prostřednictvím zásoby uhlíku diversifikované hloubkově a porostním typem (PT), diferencovaně pro cílové hospodářské soubory (CHS) v rámci přírodních lesních oblastí (PLO). Porostní typ je sám o sobě vyjádřen poměrovým zastoupením jehličnatých a listnatých druhů dřevin ve stromovém patře. Toto pojetí poskytuje v daném měřítku optimální rozlišení, na kterém lze vyjádřit efektivitu lesnického managementu vzhledem k plnění celospolečenské a zároveň klimatické funkce lesa – k uhlíkovému sinku.

Dosažený cíl sleduje náplň plánovaného výsledku projektu. Ten je zaměřen na tzv. *topsoil*, svrchní půdní vrstvy, které utvářejí humusovou formu jako nedílnou ekosystémovou součást lesní půdy. Tato součást půdního tělesa zahrnuje horizonty s největší dynamikou a zároveň přímo odráží stav porostu – vývojové stádium, dřevinnou skladbu, management apod. Při tvorbě bilančních map bylo zohledněno hledisko změny druhové skladby lesů tak, jak již byla predikována v řadě předcházejících klimatických scénářů, přičemž však časová složka sama o sobě není součástí představovaných modelů. Z důvodů své komplexity a tím také obtížné predikovatelnosti jsou přírodní podmínky vnímány jako statická proměnná, uvnitř které dochází ke změně druhové skladby, resp. porostního typu směrem k vyššímu zastoupení listnatých dřevin, a to diverzifikovaně pro konkrétní CHS. A protože porostní typ byl shledán jako významnou determinantou zásoby organického uhlíku v rámci *topsoil*, bilance je potom vyjádřena změnou v plošném zastoupení PT v rámci CHS, individuálně pro každou PLO. Vlastní porostní typ je pro účely tvorby mapy vyjádřen jako „agregovaný porostní typ“ (APT), který v pětibodové škále sdružuje strukturované porostní typy (SPT) podle procentuálního zastoupení jehličnatých a listnatých dřevin. Dynamika zásoby uhlíku je tak prezentovaná ve vazbě na tzv. „listnatý scénář“ vyjádřený poměrově vyšším zastoupením APT s vyšším podílem listnáčů, na úkor jehličnatých APT.

3 Informace k rozsahu využití mapy

Cílovou skupinou jsou uživatelé, pro něž je výstup potenciálně významným příspěvkem do aktuálně probíhající diskuse nad otázkou celkové konceptualizace lesnického managementu. Nastavení odpovídajících priorit v prostředí měnícího se klimatu je tak otázkou komplexního naplňování celospolečenských požadavků na hospodářský les ve smyslu funkcí lesa a navazujících ekosystémových služeb. Z vědeckého hlediska jsou to odborné instituce, které se tématem uhlíkového sinku a klimatických scénářů dlouhodobě zabývají – vysoké školy, výzkumné ústavy, akademie věd. V této skupině je zcela významnou dílčí skupinou studentská obec a odborná i laická veřejnost, kdy osvěta a diseminace výsledků zásadně ovlivňuje veřejné mínění a náhled na les jako celospolečenskou součást.

Z aplikačního hlediska nachází mapa uplatnění u vlastníků a správců lesních majetků – soukromých vlastníků, LČR, VLS, obecních lesů, národních parků. Významnou skupinu odběratelů tvoří národní lesnický institut (NLI) a nevládní organizace a certifikační systémy (FSC, PEFC) včetně soukromého sektoru např. v oblasti uhlíkových offsetů.

Z administrativního hlediska mapa může sloužit jako podpůrný nástroj při nastavování národních strategií v oblasti klimatu, lesnictví, ochrany půdy a při přípravě dotačních titulů.

Mezinárodní přesah, přinejmenším v rámci střeoevropského regionu i v širším evropském kontextu zasahuje lesy mírného klimatického pásma v zemích sdílejících podobné výzvy z hlediska klimatických změn. V kontextu s dlouhodobým využíváním lesa jako součásti kulturní krajiny je zásadní také historický kontext vývoje hospodaření a půdních charakteristik. Téma půdního uhlíku je aktuálně intenzivně diskutováno v rámci evropských i globálních politik (EU Green Deal, IPCC) a podobná tematická data mohou přispět k harmonizaci postupů i navazující mezinárodní vědecké spolupráci.

4 Informace o přínosech mapy pro uživatele

Hlavním přínosem mapových výstupů je informace o možnostech ekologické stabilizace uhlíkových cyklů v rámci lesního ekosystému a tím o možnostech zvýšení ekologické stability lesních porostů jako celku. Vycházíme z předpokladu nezbytnosti uplatnění série adaptačních opatření napříč sektorem lesního hospodářství, které má důvod hledat novodobé přístupy ke správě lesů. Zcela zásadním aspektem lesnického managementu, který podmiňuje navazující činnost v lesích, je dřevinná skladba. Ta je také základním a v tomto díle ve výsledku jediným faktorem managementového přístupu, který lze cíleně usměrňovat, a u kterého se také projevuje jednoznačný význam ve smyslu role lesa jako úložiště organického uhlíku.

Vedlejší přínosy mohou být také ekonomické. Ty ovšem přímo závisí na možnostech ekonomického zhodnocení dřevní hmoty v přímé vazbě na dřevoprodukční funkce lesa, nepřímo potom jednak na vyšší stabilizaci ekosystémových služeb a jednak na případném zohlednění příslušných organizačních složek státu celospolečenského přínosu lesa v boji s klimatickými výkyvy vázanými na zvýšené koncentrace uhlíku v ovzduší.

5 Seznam odborných podkladů

Odborné podklady sestávaly z rozsáhlých datových sad Národní inventarizace lesů z období 2016–2020 (NIL), které vstupovaly přímo do kalkulací a statistických analýz (regresní analýzy, prostorové analýzy modelu Random Forest). Souvisejícími podklady byly také další datové sady o fyzikálních a chemických vlastnostech půdy, které byly využity pro přípravu souvisejících publikací. Bez těch by došlo k výraznému omezení využití dat NIL, která jsou metodicky limitovaná při výpočtu zásoby půdního uhlíku v minerálních půdních horizontech. Tyto výpočty totiž zahrnují také data o objemové hmotnosti redukované půdy, která je veličinou s omezenými možnostmi kolekce při zvýšené skeletnatosti půdy nebo při omezené mocnosti půdních horizontů. V těchto případech byly využity pedotransferové funkce publikované v pracích Zouhar et al. (2024) a Kučera et al. (2025), prostřednictvím kterých bylo možné využít kompletní datové sady NIL.

Dalším souvisejícím materiálem využitým při zpracování tohoto díla je certifikovaná metodika autorů Kadavý a kol. (2025), která zpracovává adaptační opatření v lesních porostech ve vazbě na globální klimatickou změnu. Uvedené principy a přístupy ve změněném prostředí navedly autory této mapy k nastavení tzv. „listnatého“ scénáře, který v reálných obrysech navrhuje možnosti úpravy dřevinné skladby a na té demonstruje přímý dopad do půdního prostředí ve smyslu ukládání organického uhlíku v rámci tzv. *topsoil*, tedy povrchových vrstev organických a organominerálních horizontů. Pracovní hypotézou potom při zpracování dat a tvorbě map bylo, že změna dřevinné skladby s sebou nese odezvu v podobě vertikální redistribuce organického uhlíku, a to směrem do větších hloubek půdy při zvýšeném podílu listnatých dřevin.

5.1 Seznam publikací předcházejících výsledku

Kučera A., Vavříček D., Drápela K., Zouhar V., Friedl M., Vranová V. 2025: An improved relationship between bulk density and organic carbon content for forest soils using generalised linear mixed-effect model. *Carbon Balance and Management* 20(1): 36. doi: 10.1186/s13021-025-00298-0.

Zouhar V., Kučera A., Drápela K. 2024: A GLMER-based pedotransfer function expressing the relationship between total organic carbon and bulk density in forest soils. *Journal of Forest Science* 70(12): 619–633. (dedikace na projekt)

Kadavý J., Čermák P., Pokorný R., Kučera A., Kupec P., Adamec Z., Březina D., Kneifl M., Knott R., Zemánek T., Hrib M., Cempírek M., Flora M. (2025). Adaptace lesních porostů na globální klimatickou změnu. Certifikovaná metodika, Mendelova univerzita v Brně.

5.2 Použitá literatura

Bečvářová, P., Horváth, M., Šarapatka, B., Zouhar, V., 2018. Dynamics of soil organic carbon (SOC) content in stands of Norway spruce (*Picea abies*) in central Europe. *iForest*, 11. 734-742.

Hanáková Bečvářová, P., Horvath, M., Šarapatka, B., Zouhar, V., 2022. The effect of stand characteristics on soil organic carbon content in spruce and deciduous stands. *Forest Systems*, 31.

Breiman, L. 2001 Random Forests. *Machine Learning* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>

Cienciala E., Exnerová Z., Macků J., Henžlík V. 2006. Forest topsoil organic carbon content in Southwest Bohemia region. *Journal of Forest Science* 52(9):387–398.

Culek M. (ed.) a kol. (1996): Biogeografické členění České republiky. Enigma, Praha.

Djahangard M., Zhang H., Yousefpour R. 2024. Assessing the ecological complexity and uncertainty of predicting forest ecosystem services under climate change. *Ecological Complexity* 60: 101106.

- Dorogush, A. V., Ershov, V., & Gulin, A. (2018). CatBoost: gradient boosting with categorical features support. *arXiv: Learning*. <https://arxiv.org/abs/1810.11363>
- García, A. 2025. Greedy algorithms: a review and open problems. *J Inequal Appl* 2025, 11. <https://doi.org/10.1186/s13660-025-03254-1>
- Horváth, M., Hanáková Bečvářová, P., Šarapatka, B., Vencálek, O., Zouhar, V., 2021. Potential relationships of selected abiotic variables, chemical elements and stand characteristics with soil organic carbon in spruce and beech stands. *iForest*, 14. 320-328.
- Horváth, M., Hanáková Bečvářová P., Šarapatka, B., Zouhar, V., 2022. The impact of forest naturalness and tree species composition on soil organic carbon content in areas with unnatural occurrence of Norway spruce in the Czech Republic. *Soil & Water Research*, 17. 139–148.
- Kuhn, M., & Wickham, H. 2020. Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. R package version 1.1.0. <https://CRAN.R-project.org/package=tidymodels>
- Lenth R (2025). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.11.1. <https://doi.org/10.32614/CRAN.package.emmeans>
- Máslo J., Adolt R., Kohn I., Kučera M. 2024. Mortalita kmenů v ČR. Výsledky třetího cyklu národní inventarizace lesů 2016–2020. *Lesnická práce* 2 (36–42). Dostupné na https://nli.gov.cz/wp-content/uploads/2024_02_01__6_mortalita_kmenu_nil3.pdf (cit. 30. 10. 2025).
- Maxwell C.R.M., Scheller M., Long J.W., Manley P. 2022. Forest management under uncertainty: the influence of management versus climate change and wildfire in the Lake Tahoe Basin, USA. *Ecology and Society* 27(2): 15.
- Němeček J. a kol. (2011): Taxonomický klasifikační systém půd ČR. Praha, ČZU.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., & Dubourg, V. 2011. Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, 12, 2825–2830.
- Pebesma, E. (2018). *Simple Features for R: Standardized support for spatial vector data*. *The R Journal*, 10(1), 439–446. <https://doi.org/10.32614/RJ-2018-009>
- Samec P. 2020. Geografická závislost dřevinné skladby lesů a půdních vlastností během environmentální změny. Mendelova univerzita, Brno.
- Schimel D.S., Braswell B.H., Holland E-A., Mc-Keown R., Ojima D.S., Painter T.H., Parton W.J., Townsend A.R. 1994. Climatic, edaphic, and biotic controls over storage and turnover of carbon in soils. *Global Biogeochemical Cycles*, 8: 279–293.
- Schulp C.J.E., Nabuurs G.-J., Verburg P.H., de Waal R.W. 2008. Effect of tree species on carbon stocks in forest floor and mineral soil and implications for soil carbon inventories. *Forest Ecology and Management* 256: 482–490
- SPT – NLI.GOV – Strukturované porostní typy – legenda ke kódům strukturovaných porostních typů. Dostupné na https://nli.gov.cz/wp-content/uploads/Druhova_skladba.pdf (cit. 30. 10. 2025)
- Šrámek V., Fadrhonsová V., Neudertová Hellebrandová K., Cienciala E., Borůvka L. 2024. Zásoby uhlíku v lesních půdách a lesní hospodářství – review. *Zprávy lesnického výzkumu* 69(1): 22–36.
- Temlyakov, V. 2011. Greedy Approximation. Cambridge: Cambridge University Press.
- URL [1] Geovědní mapa 1:50 000 (dostupné na <https://mapy.geology.cz/geocr50/>). Citováno 15. 5. 2022.
- URL [2] Půdní mapa 1:50 000 (dostupné na <https://mapy.geology.cz/pudy/>). Citováno 15. 5. 2022.
- Vašát R., Borůvka L., Novotný R., Šrámek V., Němeček K., Pavlů L., Fadrhonsová V., Neudertová Hellebrandová K., Tejnecký V. 2025. Předpokládané změny sekvestrace uhlíku v důsledku očekávané změny druhové skladby lesů. Mapy predikce vývoje zásob organického uhlíku v celém profilu lesních půd pro roky 2030 a 2050. Specializovaná mapa s odborným obsahem, VÚLHM, v.v.i., Jíloviště-Strnady.
- Vokoun J. a kol. 2002. Příručka pro průzkum lesních půd. Taxonomický klasifikační systém půd ČR (Jan Němeček a kol.) v lesnické praxi. Ústav pro hospodářskou úpravu lesů, Brandýs nad Labem.
- Vyhláška č. 298/2018 Sb. o zpracování oblastních plánů rozvoje lesů a o vymezení hospodářských souborů.

Wu Yi, Miao C., Fan X., Gou J., Zhang Q., Zheng H. 2022. Quantifying the uncertainty sources of future climate projections and narrowing uncertainties with bias corrections. *Earth's Future* 10(11), e2022EF002963.

ZOUHAR, Václav. Organický uhlík v lesních půdách v porostním a typologickém kontextu jako strategická půdní složka 21. století]. Online. Disertační práce. Brno: Mendelova univerzita v Brně, Lesnická a dřevařská fakulta. 2024. Dostupné z: <https://theses.cz/id/gcz4jr/>.

Žárník M. 2008. Návrh úpravy systému edafických kategorií. In: Samec P. (ed.), *Metody zpracování dat v lesnickém monitoringu*. Folia Forestalia Bohemica (proceedings) 2: 124-131.

Wright, M. N., & Ziegler, A. (2017). *ranger: A fast implementation of random forests for high dimensional data in C++ and R*. *Journal of Statistical Software*, 77(1), 1–17. <https://doi.org/10.18637/jss.v077.i01>

6 Dedikace

Výsledek byl dosažen v rámci řešení projektu NAZV-Země QK21010198 Adaptace lesního hospodářství pro udržitelné využívání přírodních zdrojů.

7 Jména oponentů a názvy jejich organizací

Oponent z oboru:

Ing. Miloš Hoferka

milos.hoferka@seznam.cz

Znalec v oboru Ekonomika (36/1967 Sb.), odvětví Ceny a odhady, specializace dřevo, lesní porosty, lesní pozemky.

Znalec v oboru Lesní hospodářství (36/1967 Sb.), odvětví Dříví – těžba, specializace hospodářská úprava lesů.

Licence odborný lesní hospodář č. 4/96, č.j. 127/96 ŽP-2/Di-221 ze dne 21. 2. 1996.

Dříve učitel Nauky o lesním prostředí, Hospodářské úpravy lesů a Oceňování lesů na ČLA Trutnov.

Oponent za pracovníka z odborného orgánu státní správy:

Ing. Norbert Buchta

Ministerstvo zemědělství České republiky

Odbor státní správy, hospodářské úpravy lesů a ochrany lesů

Oddělení ochrany lesů

8 Vlastní popis výsledku

Obsah organického uhlíku v půdě představuje jeden ze základních indikátorů vitality a „zdraví“ terestrických ekosystémů (Cienciala et al. 2006). V cyklu uhlíku půda přitom hraje zcela zásadní roli, a to jak v globálním, tak i v lokálním měřítku (Schimel et al. 1994). Zásoba uhlíku v půdě přitom podléhá řadě faktorů, přičemž nejvyšší dynamikou se vyznačují svrchní vrstvy půdy, na které je také zaměřen předkládaný výsledek. Tyto faktory lze rozdělit do dvou skupin: na (1) vnější podmínky, zejména proměnlivost klimatu a na (2) vnitřní podmínky – jednak edafické ve smyslu hydrického režimu, mineralogie půdy, topografie aj., a jednak porostní ve smyslu druhové skladby (Vašát a kol. 2025) nebo souvisejícího lesnického managementu, jako je např. chemická meliorace vápněním, která je obecně problematickou formou úpravy půdního prostředí v lesních porostech s komplexními důsledky na obsah uhlíku v různých hloubkách půdy (Šrámek et al. 2025).

Ve vazbě na druhovou skladbu lesů dosavadní studie předpokládají celkové zvýšení sekvestrace uhlíku při zvyšujícím se podílu listnatých dřevin. Minoritní pozornost je však věnována vertikální distribuci dynamiky organického půdního uhlíku (TOC) při detailnějším pojetí tzv. *topsoil* – svrchních partií půdního tělesa. Ty se však vyznačují různou mírou proměn organické hmoty a různou hloubkovou alokací právě ve vazbě na druhovou skladbu lesního porostu (Schulp et al. 2008). V citované studii přitom mnohé případy zaznamenaly situaci vyšší míry uhlíkového sinku v jehličnatých porostech, s různou kvantifikací právě při porovnání organických a organominerálních horizontů.

V lesních půdách ČR je odhadovaná zásoba uhlíku v rámci celého profilu na 472,2 mil. t C (Vašát a kol. 2025). Dosavadní modely sestavované za účelem predikce vývoje zásob uhlíku v půdě ukazují spíše na jeho pokles, a to navzdory předpokladu o zvyšujícím se podílu listnáčů na druhové skladbě lesních porostů. Důvodem uváděným v citované studii je vysoká významnost faktoru klimatu a v souvislosti intenzivněji probíhajících dekompozičních procesů v porovnání s druhovým složením lesů, jehož význam je v porovnání minoritní.

Na druhou stranu, dnes rozšířené klimatické scénáře hůře reflektují lokálnost a specifickou sezónnost podmínek prostředí, která je řídicím faktorem ve smyslu vodovzdušného a teplotního režimu. V tomto ohledu je komplexnost stanovištních podmínek určující z hlediska adaptační kapacity daného ekosystému na topické úrovni. Samotné klima se tak jeví jako méně robustní z hlediska přesnosti predikce, ale i jako zatížené nespecifickou mírou nejistoty (Djahangard et al. 2024; Wu et al. 2022). Alternativní a komplementární přístup (Maxwell et al. 2022) potom může zahrnovat celistvé jednotky definující ekotop (klimatop + edatop) (lesní typy a nadstavbové jednotky jako lesní vegetační stupně a edafické kategorie, soubory lesních typů, hospodářské soubory), vyjadřující výrazně vyšší vnitřní nosnou kapacitu prostředí v porovnání s rutinními klimatickými ukazateli. Již v současnosti je známá vnitřní variabilita růstových podmínek v intervalu dané typologické jednotky (Samec 2020; Žárník 2008), která tak musí reflektovat potřebu redefinice cílové druhové skladby i prostorové distribuce v kontextu měnících se podmínek prostředí.

Z výše uvedeného vyplývá, že predikce zásob organického uhlíku v půdě by měla být postavena nejen na klimatických parametrech, ale i na komplexním pojetí stanovištních podmínek. Tento přístup může přispět k širšímu pochopení problematiky uhlíkového sinku a bilance uhlíku v terestrických systémech i k přesnější identifikaci rizik a možností adaptačních strategií uplatňovaných v lesních ekosystémech střední Evropy.

Předkládaný výsledek řeší danou problematiku na úrovni tzv. trvalých stanovištních podmínek, definovaných pro Českou republiku lesnicko-typologickým systémem (Žárník 2008), který na nejvyšší úrovni popisuje cílové hospodářské soubory jako hlavní jednotku pro hospodaření v lesích (Vyhl. č. 298/2018 Sb.), a který zároveň lépe vystihuje lokálnost stanovištních podmínek v porovnání s globálními klimatickými modely.

8.1 Materiál a Metody

8.1.1 Vstupní data

Pro tvorbu mapy byla z Národního lesnického institutu (NLI) vyžádána data z třetího cyklu Národní inventarizace lesů (NIL3) o zásobě svrchní organické vrstvy [kg m^{-2}], koncentraci uhlíku (total organic carbon – TOC; %), objemové hmotnosti půdy a mocnosti organominerálního horizontu a také lesnicko-typologické klasifikaci stanovišť a druhovém složení porostů na inventarizačních plochách. Z hlediska druhové skladby byly inventarizační plochy klasifikovány do jedné z kategorií **agregovaného porostního typu (APT; Tab. 1)** (SPT – NLI.GOV):

Tab. 1. Klasifikace agregovaných porostních typů na základě relativního zastoupení jehličnanů ve stromovém patře.

Agregovaný porostní typ	Zastoupení jehličnanů [%]
Listnatý	≤ 10
Smišený listnatý	$> 10; \leq 30$
Smišený	$> 10; < 70$
Smišený jehličnatý	$\geq 70; < 90$
Jehličnatý	≥ 90

Souřadnice odběrů byly znáhodněny formou zneprěsněné generované X a Y souřadnice inventarizačního bodu (souřadnicový systém WGS84 – EPSG:4326), která odpovídá poloze středu 1 x 1 km buňky INSPIRE (souřadnicový systém ETRS89-LAEA – EPSG: 3035), v níž se nachází daný inventarizační bod.

Dále byla vyžádána plošná data pro celou ČR, a to jak data SLHP, OPRL, tak i produkty dálkového průzkumu Země, které jsou umístěny na mapovém serveru NLI. Konkrétně byla využita vektorová data lokalizující cílové hospodářské soubory (F_A_CIHO.shp), přírodní lesní oblasti (F_A_Lesni_oblast.shp), lesní typy (F_A_Lesni_typ.shp), lesní vegetační stupně (F_A_Lesni_vegetacni_stupen.shp), hranice PUPFL (F_A_LHC_PUPFL.shp), smíšenosti v porostech (F_A_SMISENOST_POR.shp), strukturované porostní typy (F_A_SPT.shp) a klasifikaci využití půdy v rámci projektu CORINE Land Cover 2018 (CLC18_CZ.shp). Dále byla využita rastrová data zobrazující jehličnaté a listnaté porosty (lesy_jehlicnate_listnate.tif), rastrová data zobrazující hlavní lesní dřeviny z roku 2022 (lesni_dreviny_2022_5514.tif), rastrová data normalizovaného digitálního modelu povrchu (nebo také modelu výšek korun) (nDSM.tif) a digitálního modelu povrchu 5. generace (DMR_5G.tif).

8.1.2 Zpracování dat

Zásoby uhlíku [t C ha^{-1}] byly sumarizovány s ohledem na metodický přístup NIL pro jednotlivé **půdní vrstvy**:

- i. Svrchní organické vrstvy (O): $ZTOC_o$

- ii. Svrchní organominerální vrstva (A): $ZTOC_A$
- iii. Celková zásoba svrchních vrstev půdy (O+A): $ZTOC_{Top}$

Taxonomická klasifikace jednotlivých horizontů vychází z vnitřních metodických pokynů Národního lesnického institutu (NLI) (Vokoun a kol. 2002), které vylišují následující půdní horizonty:

- vrstva opadu zahrnuje kategorie horizontů Ln (nové opadanky), Lv (změněné opadanky) a L (opadanky bez vylišení);
- organická vrstva F zahrnuje kategorie horizontů, Fz (zoogenní), Fm (mykogenní), Fa (amfigenní), F (fermentační bez vylišení);
- organická vrstva H zahrnuje kategorie horizontů Hh (humusový), Hz (zoogenní), Ha (amfigenní), Hr (reziduální) a H (horizont měli bez vylišení)
- svrchní organominerální půda zahrnuje kategorie horizontů Ai (iniciální), Ah (humózní lesní), Ahe (koloidy ochuzený humózní lesní), Ahk (humózní lesní s karbonáty), Ahg (hydrogenní humózní lesní bez bročků), Ahn (hydrogenní humózní lesní s bročky, Ac (černický), Ad (humózní drnový), Am (melanický), Ame (melanický ochuzený), Amk (melanický s karbonáty), Amg (hydrogenní melanický bez bročků), Amn (hydrogenní melanický s bročky), Au (umbrický), Aug (hydrogenní umbrický bez bročků), Aun (hydrogenní umbrický s bročky), Az (antropický, Azp (vytvořený hlubokou kultivací), Azx (navrstvený humózní u koluvizemí), Ap (orniční), At (zrašelinělý).

V **organických horizontech** (L, F, H) byla zásoba uhlíku spočítána jako součet zásob jednotlivých organických horizontů v závislosti na způsobu odběru:

- pro L horizont $ZTOC_L$, viz vzorec (1);
- pro F horizont $ZTOC_F$, viz vzorec (1) v případě absence horizontu H a vzorec (2) a (3) v případě sekvence F–H;
- pro H horizont $ZTOC_H$;
- pro L+F+H horizonty $ZTOC_O$

pomocí vzorce

$$ZTOC_O = \frac{ZH \times A}{m} \times TOC_w \quad [t \text{ ha}^{-1}] \quad (1),$$

kde:

ZH je sušina vzorku odebraného na inventarizační ploše z kovového rámu 25 × 25 cm [kg m⁻²];

A je přepočít na plochu 1 ha (10 000)

m je přepočít na t (1 000);

TOC_w je koncentrace uhlíku v daném horizontu [g g⁻¹]

- pro horizont L $TOC_{wL} = 0,4405$;
- pro horizont F $TOC_{wF} = 0,3585$;
- pro horizont H TOC_{wH} analyticky stanovená z půdního vzorku v případě mocnosti H > 2 cm a v případě absence analytické hodnoty $TOC_{wH} = 0,2377$.

V případě sekvence F–H (je-li přítomný horizont H kdy jsou tyto horizonty vzorkovány jako směsný FH, avšak zásoby TOC jsou z nich vyjádřeny odděleně podle jejich mocností), byla hodnota zásoby TOC stanovena podle mocnosti jednotlivých horizontů (viz vzorce (2) a (3)).

$ZTOC_F$ (v horizontu F) byla v takovém případě stanovena jako

$$ZTOC_F = \frac{ZH \times A}{m \times (h_F + 2 \times h_H)} \times h_F \times TOC_{WF} \quad [\text{t ha}^{-1}] \quad (2)$$

kde:

h_F je mocnost horizontu F [cm];

h_H je mocnost horizontu H [cm]

$ZTOC_H$ (v horizontu H) byla stanovena jako

$$ZTOC_H = \frac{ZH \times A}{m \times (h_F + 2 \times h_H)} \times (2 \times h_H) \times TOC_{WH} \quad [\text{t ha}^{-1}] \quad (3)$$

Ve svrchních **organominerálních horizontech** (A) byla zásoba uhlíku spočítána na základě analyticky stanovené hodnoty koncentrace uhlíku v daném půdním vzorku a objemové hmotnosti redukované, pro danou konkrétní mocnost horizontu a při daném objemovém podílu skeletu pomocí vzorce

$$ZTOC_A = \frac{TOC_w \times BD \times h \times A \times Sk}{m} \quad [\text{t ha}^{-1}] \quad (4),$$

kde

TOC_w je koncentrace uhlíku zjištěná analyticky [g g^{-1}];

BD je (bulk density) objemová hmotnost redukovaná [g cm^{-3}];

h je mocnost horizontu [cm];

A je přepočítání na plochu 1 ha (10^8);

Sk je korekce na skelet: objemový podíl jemnozeme vzhledem k celkovému objemu půdy, kdy $Sk = 1 - \text{objemový podíl skeletu}$ [g g^{-1}];

m je přepočítání na t (10^6).

Neporušené půdní vzorky nebyly odebírány, pokud se vzorkovaný horizont vyznačoval: (1) skeletností $> 25\%$, při skeletnosti ve formě hrubého písku (2–4 mm) $> 50\%$; (2) mocností < 4 cm. V případě absence údaje o objemové hmotnosti redukované byla hodnota objemové hmotnosti redukované vypočtena pomocí pedotransferové funkce (Zouhar et al. 2024):

$$BD_{PTF} = \frac{1}{0,668562 + TOC_w \times 0,088349} \quad [\text{g cm}^{-3}] \quad (5).$$

Pro každou inventarizační plochu byla navíc vypočtena celková zásoba uhlíku pro svrchní vrstvy půdy zahrnující organické i organominerální horizonty:

$$ZTOC_{Top} = ZTOC_o + ZTOC_A \quad [t \text{ ha}^{-1}] \quad (6).$$

8.1.3 Ověření významu prediktorů zásob uhlíku ve svrchních půdních vrstvách: Regresní analýza

Tvorbě prostorového modelu zásoby uhlíku v půdě předcházela statistická analýza na datech NIL, kdy byl testován význam jednotlivých prediktorů zásob půdního organického uhlíku, konkrétně agregovaného porostního typu a jednotek diferencujících stanovištní podmínky i management lesů. Závislost mezi zásobami uhlíku v jednotlivých vrstvách půdy a typem porostu byla testována pomocí lineárních smíšených regresních modelů s náhodným absolutním členem v programu R (verze 4.5.0; R Core Team, 2025). U smíšených modelů byl jako fixní proměnná použit **agregovaný porostní typ** a jako náhodné proměnné **přírodní lesní oblast (PLO)**, **lesní vegetační stupeň (LVS)**, **edafická kategorie (EK)** a **cílový hospodářský soubor (CHS)** dle následující obecné rovnice:

$$ZTOC_{ij} = \mu + \alpha_j + b_i + \varepsilon_{ij} \quad [t \text{ ha}^{-1}] \quad (7),$$

kde

$ZTOC$ je zásoba uhlíku v dané půdní vrstvě inventarizační plochy,

μ je celkový průměr (intercept),

α_j je fixní efekt skupiny j (v našem případě daného agregovaného porostního typu),

b_i je náhodný intercept jednotky i , modelované jako náhodný efekt (např. PLO),

ε_{ij} je reziduální chyba.

Tento přístup umožňuje modelovat variabilitu $ZTOC$ v rámci jednotek (fixní efekt) odděleně od variability $ZTOC$ mezi jednotkami (náhodný efekt). V případě více náhodných efektů byl model rozšířen následovně:

$$ZTOC_{ijk} = \mu + \alpha_j + b_i + c_k + \varepsilon_{ijk} \quad [t \text{ ha}^{-1}] \quad (8),$$

kde

c_k je intercept pro jednotku k dalšího náhodného efektu (např. CHS).

Pro splnění obecných předpokladů lineárního modelu byla data transformována a normalita reziduí byla ověřena pomocí Shapiro-Wilkova testu normality a QQ diagramu. Konkurenční modely byly porovnány pomocí Akaikeho informačního kritéria (AIC) a testu poměru pravděpodobnosti (Likelihood Ratio; LR test). Odhad středních hodnot pro jednotlivé agregované porostní typy (fixní proměnnou) byl proveden pomocí balíčku "emmeans" (Lenth 2025) s následným mnohonásobným porovnáním a Šidákovou korekcí pro hodnoty P . Nejvyšší akceptovaná hladina významnosti všech testů byla stanovena na $P = 0,05$ (tj. 5 %). R^2_m a R^2_c udávají podíl rozptylu dat vysvětleného pomocí fixní proměnné (agregovaný porostní typ) a celého modelu (fixní i náhodné vlivy). SE je v textu užíváno pro označení směrodatné chyby a "df" pro stupně volnosti.

8.1.4 *Prostorová analýza: distribuce agregovaných porostních typů*

Výpočet hodnot prediktorů při zneprůhlednění souřadnicích dat NIL3:

Z důvodu znáhodnění souřadnic odběrů NIL v gridu 1x1 km bylo potřeba provést nový výpočet hodnot prediktorů z dat poskytnutých NLI. Nejdříve byl vytvořen čtvercový buffer o velikosti 1x1 km kolem bodů odběrů pomocí Python skriptu a knihovny arcpy (skript viz Příloha 1.1).

Pro každý vzniklý čtverec byla vypočítána jedna hodnota prediktoru převládajícího na daném čtverci. K tomu byl v případě vektorových dat (vrstvy obsahující PLO, CHS, SPT) využit nástroj softwaru ArcGIS Pro Spatial Join s parametrem Largest Overlap, který do atributové tabulky přidal hodnotu polygonu s největším překryvem. Pro rastrová data (normalizovaný digitální model povrchu - nDSM) byl využit nástroj Zonal Statistics as Table, který pro daný čtverec vypočítal průměr, maximum a směrodatnou odchylku hodnot dané rastrové vrstvy. Rastrové vrstvy byly ořezané maskou lesa ještě před výpočtem zonální statistiky. Pro nasazení modelu na data ČR byl vytvořen grid o rozměru 1x1 km pomocí nástroje Create Fishnet a prediktory pro tento grid byly získány obdobným způsobem.

Vytvoření redukce počtu prediktorů a agregace porostních typů:

Byl vytvořen skript pro model Random Forest, v jehož rámci byla pomocí metody postupné redukce vstupních proměnných (**Recursive Feature Elimination, RFE**) určena minimálně optimální sada prediktorů. Výsledky ukázaly, že jako signifikantní prediktory je vhodné použít vylišení území (**PLO**), specifikaci přírodních resp. stanovištních podmínek (**CHS**), lesnický management vyjádřený strukturovaným porostním typem (**SPT**) a dále využít hodnoty aritmetického průměru a směrodatné odchylky výšek porostu získaných z normalizovaného digitálního modelu povrchu (**nDSM**). Výška porostů byla použita jako náhradní veličina místo věku porostů, který nebyl dostupný pro celou ČR. Aby bylo možné zavést scénáře změny druhové skladby lesních porostů, byly SPT, s četností cca 68 tis./ČR sloučeny do tzv. agregovaných porostních typů (**APT**) (pětibodová škála) s identickým relativním zastoupením jehličnanů ve stromovém patře jako v případě regresních modelů (Tab. 1). Vzhledem k tomu, že v první verzích map byly patrné ostré přechody hodnot TOC mezi jednotlivými PLO, bylo provedeno sloučení prediktorů PLO a CHS do nového prediktoru PLO_CHS.

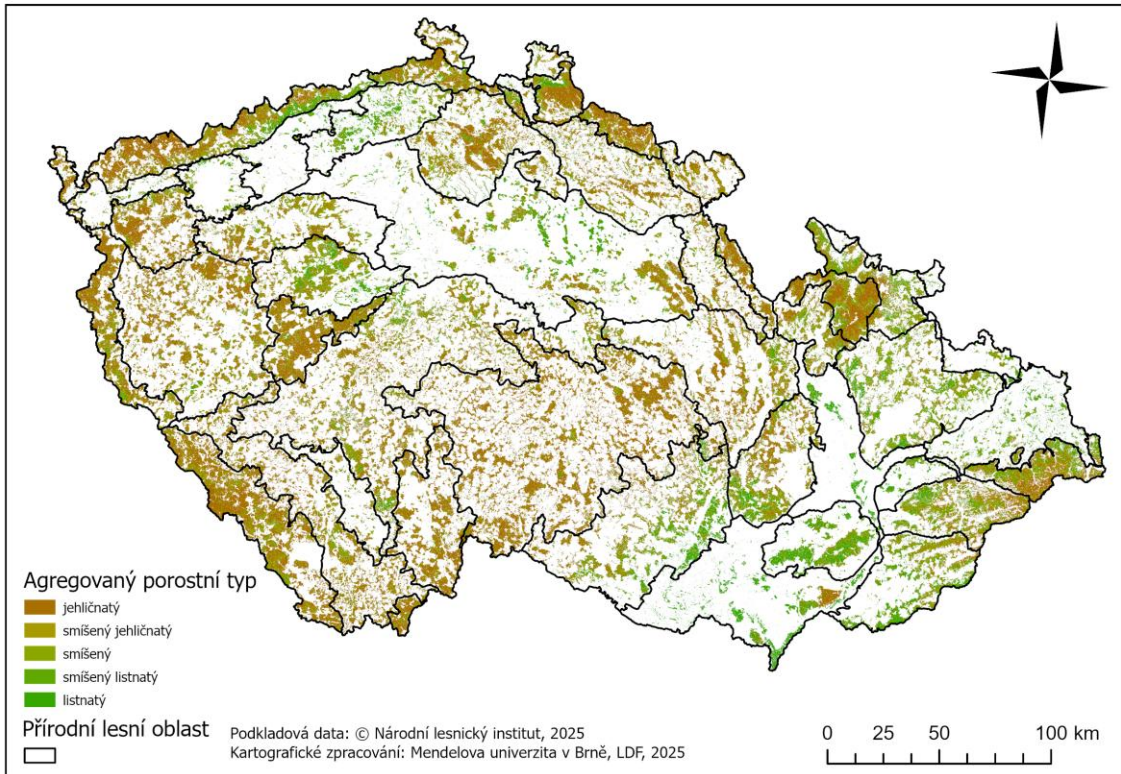
Pomocí Python skriptu byly SPT rozparsovány a zjištěny jednotlivé druhy dřevin:

- Kódy **1, 1e, 2, 2e, 3, 3e, 4, 4k, 4x** = jehličnany
- Kódy **5, 5e, 6, 7, 7e, 8, 9, 9e, 9x** = listnáče

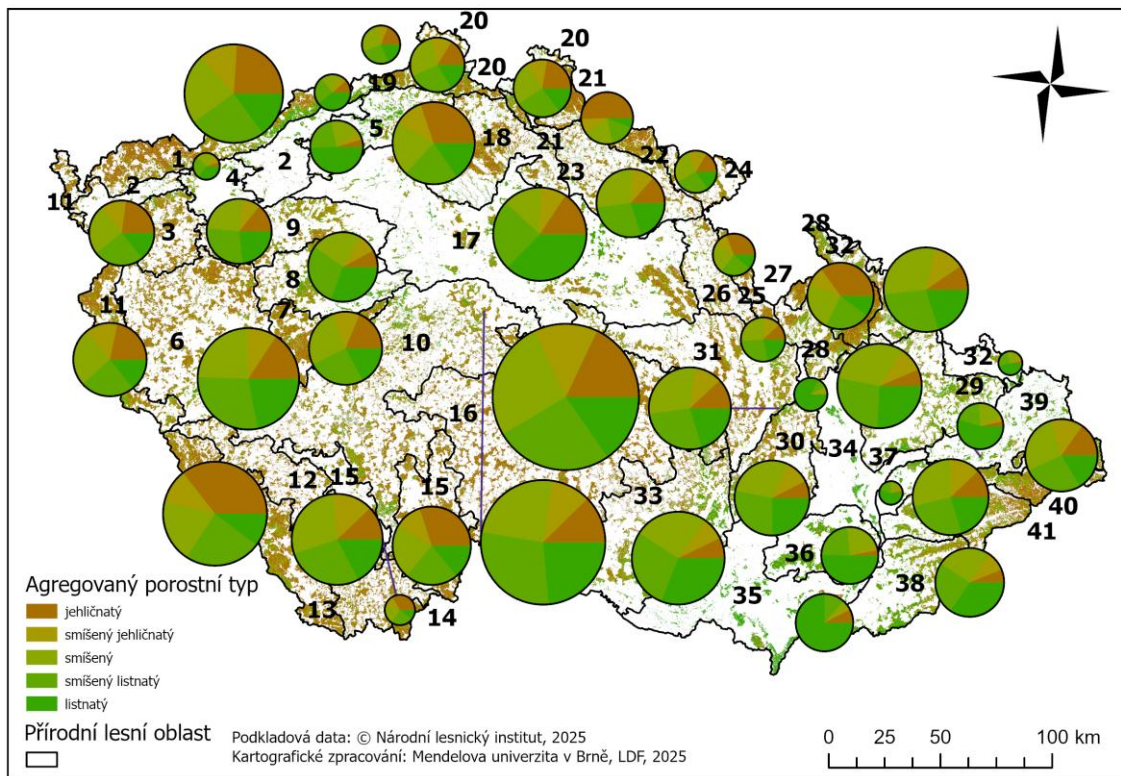
Dále byly rozklíčovány jednotlivé charaktery smíšení:

- Charakter smíšení (**C, D, M, Z, P, V**) (**C** = čistý ≥ 90 %, **D** = dominantní 70–90 %, **M** = majoritní 50–70 %, **Z** = základní 30–50 %, **P** = příměs 10–30 %, **V** = vtroušené <10 %)

Aktuální distribuce APT vyjádřených na základě výše uvedeného postupu je uvedena na Obr. 1. Výsledný Python skript viz Příloha 1.2.



a



b

Obr. 1. (a) Aktuální distribuce agregovaných porostních typů (APT); (b) Kartogram zastoupení APT v jednotlivých PLO. Černé linie označují hranice PLO, potažmo ČR.

Odhad obsahu uhlíku v půdě: Random Forest predikce TOC – rychlá implementace

Pro odhad obsahu uhlíku v lesních půdách jsme použili algoritmus strojového učení **Random Forest** v prostředí R (balíčky `tidymodels`, `ranger`, `sf`, `future`, `purrr`, `readr`, `jsonlite`) (Kuhn et Wickham, 2020; Wright et Ziegler, 2017; Pebesma, 2018). Cílem bylo natrénovat tři modely – zvlášť pro organickou vrstvu (horizont O), organominerální horizont (horizont A) a pro jejich **součet (O+A)** – a aplikovat je na predikční grid.

Vstupní data

- Bodové vrstvy v souřadnicovém systému S-JTSK / Křovák East North (EPSG: 5514):
 - `body_O.shp` (měření v horizontu O),
 - `body_A.shp` (měření v horizontu A),
 - `body_A_O.shp` (součet O+A).
- Prediktory:
 - spojité: **MEAN**, **STD** (průměr a směrodatná odchylka z nDSM),
 - kategoriální: **CHS**, **PLO**, **APT**.
- Predikční grid: `grid_TOC.shp`.

Datová příprava

- Všechny vrstvy se načítají funkcí `read_sf_5514()`, která doplní chybějící CRS (nastaví EPSG:5514) a jinak provede transformaci do EPSG:5514.
- Funkce `unify_predictors()` vybírá prediktory pouze podle „master listu“ (MEAN, STD, MAX, APT, PLO, CHS) a řeší kompatibilitu vstupů (včetně ošetření list-sloupců přes `squash_list_cols()`). Funkce `cast_types()` převádí numerické prediktory na numeric (včetně podpory desetinné čárky) a kategoriální na factor.
- Pro každý dataset se provádí diagnostika typů prediktorů funkcí `inspect_pred_types()` (deklarovaný typ vs. detekovaný typ, počty NA, unikáty/úrovně). Report může být uložen do CSV (`type_report_*.csv`).

Modelování

Recept (recipe) zahrnuje imputaci chybějících hodnot, sjednocení kategorií a odstranění prediktorů s nulovou variabilitou.

Pro Random Forest se používá balíček `ranger`, nastavení: max. hloubka 30, vzorek 70 %, ignorování neznámých kategorií.

Tuning: probíhá přes **10-fold CV** a **latin hypercube** výběr **30** kombinací. Laděné parametry:

- **trees:** 400–1200,
- **mtry:** 1 až ~50 % počtu prediktorů po `bake`,
- **min_n:** 2–20.

V tuningu je významnost prediktorů vypnutá (**importance="none"**) pro zrychlení výpočtu.

Finální model se trénuje s nejlepšími hyperparametry a až zde se počítá permutation importance proměnných. Interpretace modelu je založena primárně na **CV permutační významnosti** (`cv_permutation_importance()`): v testovací části každého fold-u se prediktor náhodně permutuje a měří se zhoršení **RMSE** (delta RMSE). Tento přístup je férovější vůči numerickým i kategoriálním proměnným a je reportován jako průměr a SD delta RMSE přes foldy a opakování (`repeats`). Doplnkově je z finálního modelu exportována i **impurity importance** z `rangeru` (**importance="impurity"**).

Paralelizace běží pouze uvnitř tuningu přes balíček future.

Výstupy

Pro tréninková data se ukládají resub metriky (R^2 , RMSE, MAE, nRMSE_mean, nRMSE_range).

Na predikční grid se doplní tři nové atributy:

A_pred (horizont A), O_pred (horizont O), SUM_pred (součet).

Výsledky se ukládají jako shapefile v EPSG:5514 s UTF-8 kódováním (.cpg soubor).

Do CSV se ukládají i sumarizační metriky a tabulky importance proměnných (raw i agregované podle prediktorů). Finální workflow modely se ukládají jako .rds a metadata jako .json do složky models_rf_full

Shrnutí

Tento skript představuje **rychlejší verzi Random Forest predikce TOC**, která:

- zajišťuje jednotnou typovou kontrolu prediktorů,
- minimalizuje dobu ladění parametrů (zúžený grid, méně foldů),
- a současně poskytuje robustní výstupy (predikční mapu, metriky, proměnnou důležitost).

Výhodou je plná reprodukovatelnost v prostředí R a možnost snadno upravit počet stromů, rozsah ladění nebo přidat nové prediktory podle dostupných dat.

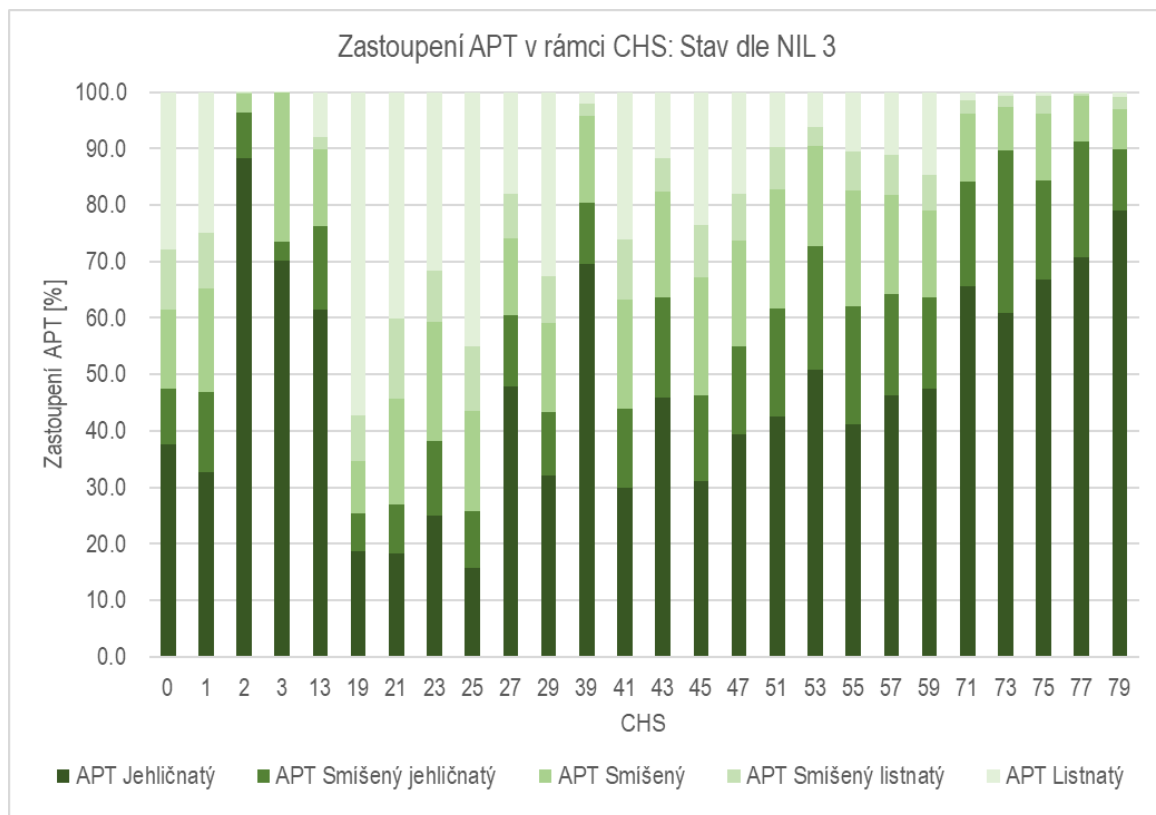
Výsledný R skript pro trénování a nasazení Random Forest modelu viz Příloha 1.3.

8.1.5 *Scénář změněné dřevinné skladby ve prospěch listnáčů u vybraných CHS a agregovaných porostních typů s nižším zastoupením jehličnanů*

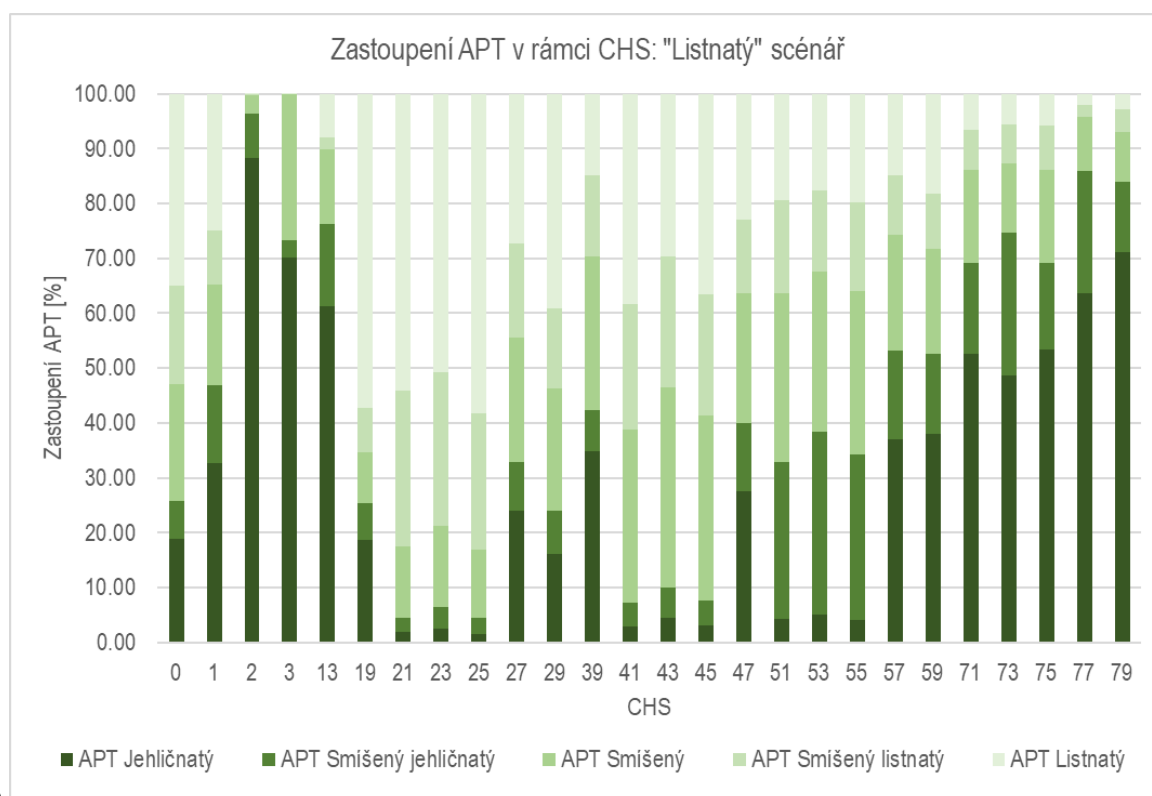
Byla vytvořena tabulka zobrazující procentuální zastoupení jednotlivých agregovaných porostních typů (APT) v rámci CHS – průměry pro celorepublikovou úroveň (Obr. 2a). Tato procentuální zastoupení byla změněna podle zadaného scénáře, kdy v současnosti klimaticky resp. stanovištně rizikové APT vykazovaly mezi lety 2016 až 2022 zhoršenou vitalitu a předčasný přechod do fáze rozpadu (Máslo a kol. 2024). Na základě expertních odhadů tak byl navržený scénář předpokládaného procentuálního uchování porostů s dominancí jehličnanů diferencovaně pro jednotlivé CHS (Tab. 2) s tím, že úbytek zastoupení vyjádřený zbytkovým procentem byl ve prospěch listnatých porostních typů proporcionalně rozpočítán mezi nedotčené APT. Pro celorepublikovou úroveň zastoupení agregovaných porostních typů v jednotlivých CHS ukazuje Obr. 2b. Tyto změny nebyly vztaženy na CHS 01; 02 a 03 se specifickým lesnickým managementem, na CHS 13 (přirozená borová stanoviště a stanoviště borových doubrav) a 19 (přirozená lužní stanoviště nižších poloh) a na reziduální plochy PUPFL s minoritní reprezentací označené v Tab. 2 a na Obr. 2 jako "(0)". Vlastní procentuální změny v zastoupení APT v rámci CHS uvádí Tab. 3.

Tab. 2. Podklad pro výpočet zastoupení CHS v rámci PLO při kalkulaci „listnatého“ scénáře. Hodnoty udávají procentuální zůstatky agregovaných porostních typů (APT) v příslušných CHS při uplatnění „listnatého“ scénáře.

CHS	APT Jehličnatý	APT Smíšený jehličnatý	APT Smíšený	APT Smíšený listnatý	APT Listnatý
(0)	50	70	-	-	-
01	100	100	-	-	-
02	100	100	-	-	-
03	100	100	-	-	-
13	100	100	-	-	-
19	100	100	-	-	-
21	10	30	70	-	-
23	10	30	70	-	-
25	10	30	70	-	-
27	50	70	-	-	-
29	50	70	-	-	-
39	50	70	-	-	-
41	10	30	-	-	-
43	10	30	-	-	-
45	10	30	-	-	-
47	70	80	-	-	-
51	10	-	-	-	-
53	10	-	-	-	-
55	10	-	-	-	-
57	80	90	-	-	-
59	80	90	-	-	-
71	80	90	-	-	-
73	80	90	-	-	-
75	80	90	-	-	-
77	90	-	-	-	-
79	90	-	-	-	-



a



b

Obr. 2. Procentuální zastoupení agregovaných porostních typů (APT) v rámci cílových hospodářských souborů (CHS) – průměrné hodnoty za Českou republiku. (a) hodnoty aktuálního stavu (NIL3); (b) hodnoty v případě „listnatého“ scénáře jako podklad pro navazující tvorbu modelu „listnatého scénáře“.

Tab. 3. Procentuální změny v zastoupení agregovaných porostních typů (APT) v rámci cílových hospodářských souborů (CHS) v „listnatém scénáři“. Celorepublikový průměr.

CHS	APT Jehličnatý	APT Smíšený jehličnatý	APT Smíšený	APT Smíšený listnatý	APT Listnatý
(0)	0	0	0	0	0
01	0	0	0	0	0
02	0	0	0	0	0
03	0	0	0	0	0
13	0	0	0	0	0
19	0	0	0	0	0
21	-16,5	-28,0	-5,6	25,1	25,1
23	-22,4	-22,1	-6,3	25,4	25,4
25	-14,2	-31,5	-5,3	25,5	25,5
27	-23,9	-5,4	9,8	9,8	9,8
29	-16,0	-9,8	8,6	8,6	8,6
39	-34,8	-0,6	11,8	11,8	11,8
41	-27,0	-18,2	15,1	15,1	15,1
43	-41,2	-8,3	16,5	16,5	16,5
45	-27,9	-16,5	14,8	14,8	14,8
47	-11,8	-3,6	5,1	5,1	5,1
51	-38,2	9,6	9,6	9,6	9,6
53	-45,8	11,4	11,4	11,4	11,4
55	-37,0	9,3	9,3	9,3	9,3
57	-9,2	-1,1	3,5	3,5	3,5
59	-9,5	-1,5	3,6	3,6	3,6
71	-13,1	-0,2	4,4	4,4	4,4
73	-12,2	-0,1	4,1	4,1	4,1
75	-13,4	-0,1	4,5	4,5	4,5
77	-7,1	1,8	1,8	1,8	1,8
79	-7,9	2,0	2,0	2,0	2,0

Hodnoty z Tab. 3 byly použity pro výpočet nové plochy zastoupení APT v rámci CHS, zvlášť pro každou PLO: Byl vytvořen XLSX soubor s převodem aktuálního procentuálního [%] i absolutního [ha] zastoupení APT na zastoupení v „listnatém“ scénáři pro danou kombinaci PLO a CHS. Následně byl navržen skript v jazyku **Python** s využitím knihoven `arcpy`, `pandas` a základních funkcí jazyka Python. Účelem bylo přepočítat dřevinnou skladbu porostních **polygonů** podle nově definovaných výměr.

Základní vstup tvoří:

- shapefile strukturovaných porostních typů obohacený o atributy APT, CHS a PLO,
- list Excelu, kde jsou pro kombinace (PLO, CHS) uvedeny výměry jednotlivých kategorií APT (např. *jehličnatý, smíšený, listnatý*),

Výstupem je pak shapefile s aktualizovanou dřevinnou skladbou (změněný vstupní shapefile strukturovaných porostních typů).

8.1.6 Tvorba prostorových modelů a map zásob TOC

Postup zpracování

1. Načtení dat a kontrola vstupů

Skript ověřuje existenci vstupních souborů a klíčových polí (APT, CHS, PLO), případně varianty s apendixem _1, které vznikají při operaci s shapefiley nebo třídy prvků geodatabáze.

Následně vytvoří kopii původního shapefile, aby nedocházelo ke ztrátě originálních dat, a přidá nové pole APT_OLD, ve kterém zachová původní hodnoty dřevinné skladby.

2. Zpracování tabulky scénáře

Tabulka xlsx s listem nových výměr se načte pomocí knihovny `pandas`.

Výměry jednotlivých kategorií APT se převádějí na **podíly v rámci každé kombinace (PLO, CHS)**, aby bylo možné scénář aplikovat nezávisle na jednotkách výměry (ha, m²).

Součty ploch se přepočítají na jednotkové hodnoty (0–1) a uloženy jako tzv. *cílové podíly* pro každou dvojici klíčů.

3. Výpočet reálných výměr v mapě

Pomocí kurzoru `arcpy.da.SearchCursor` se pro každý polygon zjistí aktuální hodnota APT a plocha z geometrie (`SHAPE@AREA`).

Všechny prvky se seskupí podle kombinace (LES_OB, CHS), aby bylo možné porovnat jejich aktuální stav s cílovým rozložením podle scénáře.

4. Ekologická posloupnost dřevinných typů

Dřevinné typy jsou uspořádány do fixní posloupnosti odpovídající ekologickému gradientu:

jehličnatý → smíšený jehličnatý → smíšený → smíšený listnatý → listnatý

Každému typu je přiřazen index (0–4), který umožňuje modelovat přechod směrem k listnatějším typům.

5. Greedy algoritmus pro přerozdělení

Vlastní přerozdělení ploch probíhá pomocí greedy (chamtivého) algoritmu (Temlyakov, 2011, García 2025), který v každém kroku vybírá lokálně „nejlepší“ možné řešení – polygon se přiděluje té kategorii dřevinné skladby, která má největší zbývající nenaplněnou výměru podle scénáře.

Princip výpočtu

Pro každou dvojici (PLO_1, CHS) se spočítá součet ploch všech polygonů v gridu. Cílová výměra pro jednotlivé typy APT se určí jako

$$A_{\text{target},i} = S_i \times A_{\text{total}} \quad (9),$$

kde:

- * $A_{\text{target},i}$ je cílová plocha daného typu APT,
- * s_i je cílový podíl ze scénáře,
- * A_{total} je celková plocha kombinace (PLO, CHS).

Každý polygon se zpracovává v pořadí od největší plochy k nejmenší, čímž se nejdříve přiřazují prvky s největším vlivem na výsledek.

Polygon může přejít pouze směrem k listnatému typu, tedy do třídy s vyšším indexem v posloupnosti ($i_{\text{new}} \geq i_{\text{old}}$). Není však omezen krokem o ± 1 – může přeskočit více úrovní, pokud to vyžaduje naplnění cílových výměr. Současně je pro změny zamčen APT pro CHS 01, 02, 03, 13, 19. Dále byly testovány varianty algoritmu s různými podmínkami:

1. Stejně omezení směru, a navíc ještě omezení kroku o maximálně jeden APT (např. z jehličnatého může přejít maximálně do smíšeného jehličnatého).
2. Bez jakéhokoli omezení, důležité je pouze naplnit požadované nové výměry APT. V každém kroku algoritmus vyhodnotí kandidátní třídy (tj. všechny typy „více listnaté“ než aktuální), z nich vybere tu, která má nejvyšší zbývající hodnotu ($A_{\text{target},i} - A_{\text{current},i}$), přiřadí polygon do této kategorie a odpovídající cílová výměra se sníží o jeho plochu.

Proces pokračuje, dokud nejsou přiřazeny všechny polygony dané skupiny, čímž se minimalizuje rozdíl mezi skutečnými a cílovými podíly dřevinných typů.

6. Diagnostika a kontrola výsledků

Po přerozdělení se nově přiřazené hodnoty uloží do pole APT, zatímco původní zůstávají v APT_OLD.

Skript následně vypočítá pro každou kombinaci (PLO, CHS) průměrnou absolutní odchylku mezi dosaženým a cílovým podílem jednotlivých APT kategorií. Tato hodnota (typicky 0–0,1) slouží jako indikátor kvality dosažení scénáře.

V této variantě skriptu, která byla finálně použita pro generování nových APT, bylo změněno APT u 1 269 636 z 3 278 975 prvků a byla dosažena průměrná absolutní odchylka podílů (0–1): 0,0306. V případě alternativní varianty skriptu s omezením směru a omezením kroku o jeden APT bylo změněno APT u 1 076 674 z 3 278 975 prvků a průměrná absolutní odchylka podílů byla 0,1044. U varianty bez jakéhokoli omezení, pouze s cílem dosažení nových výměr bylo změněno 2 492 927 z 3 278 975 prvků a dosažena průměrná absolutní odchylka podílů 0.0070.

Python skript pro výše pospaný postup uvádí Příloha 1.4.

Vzniklý shapefile s novými agregovanými porostními typy byl připojen do původního gridu 1x1 km. K tomu byl znovu využit nástroj softwaru ArcGIS Pro Spatial Join s parametrem Largest Overlap, který do atributové tabulky přidal hodnotu polygonu s největším překryvem. Prostorová predikce obsahu uhlíku v půdě byla provedena pomocí již natrénovaných modelů Random Forest, uložených ve formátu RDS. Modely byly aplikovány na nový predikční grid prostřednictvím samostatného skriptu (Příloha č. 1.5), který zajišťuje zachování shodné struktury a datových typů prediktorů jako při tréninku. Predikce byla provedena odděleně pro horizont A, horizont O a jejich součet. Výsledná polygonová vrstva byla následně převedena do rastrů s rozlišením 5 m. Pole použitá pro výpočet hodnot rastrů byly hodnoty sekvestrovaného uhlíku v příslušném horizontu (O, A, O+A). Vzhledem k velikosti souborů při použití uvedeného postupu (např. pro odlišné scénáře změny v zastoupení jehličnatých dřevin) důrazně doporučujeme převést všechny potřebné vrstvy a tabulky do geodatabáze kvůli lepší indexaci. Výsledné rastry byly ořezány pomocí hranice PUPFL, respektive polygonové třídy prvků se strukturovanými porostními typy. Následně byly predikované hodnoty zásoby uhlíku v půdních horizontech O, A a jejich součtu (O+A) agregovány z predikčního gridu o rozlišení 1 × 1 km na úroveň porostních polygonů. K tomuto účelu byl použit nástroj Tabulate Intersection, který umožňuje kvantifikovat plošný překryv mezi porostními polygonovými jednotkami a jednotlivými buňkami predikčního gridu. Výsledná tabulka obsahovala identifikátor porostu, identifikátor gridové buňky a plochu jejich vzájemného překryvu. Predikované hodnoty zásoby uhlíku v jednotlivých horizontech byly následně k této tabulce připojeny pomocí nástroje Join Field a pro každý porost byl vypočten plošně vážený průměr zásoby uhlíku. Váhovým koeficientem byla plocha překryvu mezi porostem a příslušnou gridovou buňkou. Agregace proběhla pomocí součtu plošně vážených příspěvků a jejich následným přepočtem na jednotkovou plochu porostu. Výsledkem byly průměrné hektarové zásoby uhlíku v horizontech O, A a O+A přiřazené ke každému porostnímu polygonu. Tyto hodnoty byly následně využity pro výpočet průměrných zásob sekvestrovaného uhlíku v rámci kombinací porostních charakteristik **PLO, CHS a APT**.

Aplikace scénáře pomocí greedy algoritmu představuje prostorovou nejistotu, jelikož záleží na tom, jakým způsobem se nové APT rozmístí v rámci dané kombinace PLO a CHS. Technicky správnější by bylo vytvořit scénář pomocí přirozené dřevinné skladby v rámci CHS a PLO.

8.2 Výsledky a Diskuse

8.2.1 Výsledky smíšených regresních modelů zásob uhlíku v závislosti na agregovaném porostním typu

7. Organické půdní vrstvy

Výsledný model dokládá významné rozdíly v zásobách uhlíku mezi APT, kde nejnižší zásoby byly pozorovány v lese listnatém ($14,4 \pm 1,8$ t ha⁻¹) a nejvyšší v lesích s vyšším zastoupením jehličnanů, od smíšeného po jehličnatý ($19,5$ až $22,8 \pm 2,2$, t ha⁻¹; Tab. 4). Výsledný model o dvou náhodných proměnných (PLO a CHS) poskytoval významně lepší predikční schopnost (LR test $P < 0,001$) v porovnání s modelem o jedné náhodné proměnné (Příloha 2 – Tab. P1), při čemž spolehlivost predikce modelu byla ověřena na reziduích (Příloha 2 – Obrázek P1).

Tab. 4. Střední hodnoty zásob uhlíku [t ha⁻¹] v **organických vrstvách** (O) dle agregovaného porostního typu (APT) ($P < 0,001$; $F = 9,76$; $R^2_m = 0,030$) s náhodnými interceptry pro PLO a CHS ve smíšeném regresním modelu ($R^2_c = 0,328$; $AIC = 5120$).

APT	Odhad	SE	df	Interval spolehlivosti	
				2,5%	97,5%
Listnatý	14,4 ^a	(1,8)	76,7	9,7	19,0
Smíšený listnatý	20,0 ^{ab}	(2,9)	210,3	12,4	27,5
Smíšený	20,3 ^b	(2,3)	91,1	14,1	26,4
Smíšený jehličnatý	19,5 ^b	(2,2)	74,7	13,8	25,2
Jehličnatý	22,8 ^b	(2,2)	52,2	17,0	28,7

^{a,b} Rozdílná písmena značí statistické rozdíly mezi APT (Šídák t-test; $P < 0,05$); společná písmena značí neprůkaznost nebo absenci rozdílu.

B) Organominerální vrstva

V organominerální vrstvě byly pozorovány nejnižší zásoby uhlíku v lese jehličnatém ($20,9 \pm 2,3$ t ha⁻¹) a smíšeném jehličnatém ($22,0 \pm 2,6$ t ha⁻¹) a naopak nejvyšší v lese listnatém ($28,4 \pm 3,4$ t ha⁻¹; Tab. 5). Výsledný model o dvou náhodných proměnných (PLO a CHS) poskytoval významně lepší predikční schopnost (LR test $P < 0,003$) v porovnání s modelem o jedné náhodné proměnné (Příloha 2 – Tabulka P2), při čemž uspokojivá spolehlivost predikce modelu byla ověřena na reziduích (Příloha 2 – Obrázek P2).

Tab. 5. Střední hodnoty zásob uhlíku [t ha⁻¹] ve **svrchní organominerální vrstvě** (A) dle agregovaného porostního typu (APT) ($P = 0,002$; $F = 4,195$; $R^2_m = 0,019$) s náhodnými interceptry pro PLO a CHS ve smíšeném regresním modelu ($R^2_c = 0,253$; $AIC = 519$).

APT	Odhad	(SE)	df	Interval spolehlivosti	
				2,5%	97,5%
Listnatý	29,0 ^b	(3,2)	41	20,4	37,5
Smíšený listnatý	25,1 ^{ab}	(3,7)	123	15,6	34,7
Smíšený	24,8 ^{ab}	(2,9)	53	17,2	32,5
Smíšený jehličnatý	22,6 ^{ab}	(2,5)	43	15,9	29,4
Jehličnatý	21,4 ^a	(2,1)	28	154,5	27,2

^{a,b} Rozdílná písmena značí statistické rozdíly mezi APT (Šídák t-test; $P < 0,05$); společná písmena značí neprůkaznost nebo absenci rozdílu.

C) Suma zásob organické a organominerální vrstvy

Pro celkovou zásobu uhlíku svrchních vrstev (O+A) nebyl nalezen spolehlivý regresní model (SW test $P < 0,001$), což naznačuje potřebu použít jiného přístupu anebo modelovat zásoby uhlíku ve svrchních organických a organominerálních vrstvách zvlášť.

8.2.2 Výsledné prostorové modely zásoby a bilance uhlíku

Tab. 6 uvádí výsledky predikce pomocí modelu Random Forest s uvedením konkrétních parametrů odděleně pro jednotlivé vrstvy půdy. Nejvýznamnější prediktory jsou přitom vyjádřeny jako „permutation importance“ – tedy: relativnímu zhoršení predikční přesnosti modelu po náhodném promíchání hodnot daného prediktoru, při zachování ostatních proměnných beze změny.

Tab. 6. Výsledky predikce zásob uhlíku pomocí modelů Random Forest. R^2 – koeficient determinance na trénovacích datech; RMSE (Root Mean Square Error) – odmocněná střední kvadratická chyba [t/ha]; MAE (Mean Absolute Error) – celková střední chyba predikce [t/ha].

Model / Horizont	R^2	RMSE	nRMSE	MAE	Nejvýznamnější prediktory (pořadí)
horizont O	0.296	16.507	0,104	12.054	1. PLO_CHS (0.315), 2. APT (0.228), 3. MEAN (0.189), 4. MAX (0.070), 5. STD (-0.008)
horizont A	0,304	28.290	0,048	16.221	1. APT (0.266), 2. STD (0.138), 3. MAX (0.126), 4. PLO_CHS (0.045), 5. MEAN (-0.008)
součet O+A	0.276	29.496	0,050	18.920	1. MAX (0.244), 2. PLO_CHS (0.214), 3. STD (0.099), 4. MEAN (0.090), 5. APT (0.056)

Výsledky ukazují, že kvalita predikce obsahu uhlíku pomocí algoritmu Random Forest se mezi jednotlivými půdními horizonty liší. V horizontu O dosáhl model střední úrovně predikční schopnosti ($R^2 = 0,296$), přičemž hodnoty normalizované střední kvadratické chyby (nRMSE = 0,104) odpovídají trendově dosaženým hodnotám koeficientu determinance (Tab. 6). Relativně vyšší hodnota nRMSE, normalizované podle rozsahu hodnot, naznačuje vyšší variabilitu predikční chyby v tomto horizontu, což může souviset s prostorově heterogenním charakterem opadové vrstvy. Pro horizont A vykazuje model srovnatelný podíl vysvětlené variability ($R^2 = 0,304$), avšak absolutní chyba predikce je vyšší (RMSE = 28,29 t C ha⁻¹). Nižší hodnota MAE ve srovnání s RMSE naznačuje přítomnost několika odlehlých hodnot s výrazně vyšší chybou predikce. Při normalizaci RMSE podle rozsahu hodnot dosahuje tento horizont nižší relativní chyby (nRMSE = 0,048) než horizont O, což ukazuje na stabilnější predikční výkon v rámci celého rozpětí hodnot. Model predikující součet obsahu uhlíku v horizontech O a A dosáhl mírně nižší hodnoty koeficientu determinance ($R^2 = 0,276$), přičemž absolutní chyba predikce byla nejvyšší (RMSE = 29,50 t C ha⁻¹). Normalizace RMSE podle rozsahu hodnot (nRMSE = 0,050) však ukazuje, že relativní chybovost tohoto modelu je srovnatelná s horizontem A. Pokud je RMSE normalizována podle průměrné hodnoty predikované proměnné, vykazuje model pro součet horizontů O+A relativně nižší chybu (nRMSE_mean = 0,650) ve srovnání s horizontem A (nRMSE_mean = 0,708) a horizontem O (nRMSE_mean = 0,938). Přestože model pro součet horizontů O+A vykazuje vyvážené hodnoty relativní chybovosti, nebyl tento model použit pro mapovou prezentaci obsahu uhlíku. Důvodem je skutečnost, že samostatný model pro součet O+A nedosahoval stejné konzistence a interpretační stability jako prostý součet predikcí samostatných modelů pro horizont O a horizont A. Z tohoto důvodu byl pro tvorbu map obsahu uhlíku použit rastr vzniklý součtem

predikovaných hodnot modelu pro opadovou vrstvu a modelu pro organominerální horizont, obdobně jako v případě regresního přístupu. K tomu byl využit nástroj Plus v softwaru ArcGIS Pro. Lze rovněž vytvořit model bez prediktoru APT, který by ale nebylo možné použít pro scénář vyššího zastoupení listnáčů, ani pro bilanci uhlíku.

Analýza významnosti prediktorů, vyjádřené pomocí permutation importance, ukazuje rozdílné mechanismy ovlivňující predikci obsahu uhlíku v jednotlivých půdních horizontech. V horizontu O se jako nejvýznamnější prediktory uplatnily proměnná PLO_CHS a APT, následované spojitými proměnnými popisujícími strukturu porostu (MEAN, MAX a STD). To naznačuje, že u opadové vrstvy mají výrazný vliv stanovištní a hospodářské charakteristiky zachycené v PLO_CHS, které doplňuje informace o typu porostu (APT), zatímco strukturální metriky hrají podpůrnou roli. V horizontu A byla nejvyšší významnost zaznamenána u proměnné APT, následované ukazateli variability a extrémů výškové struktury porostu (STD a MAX). Proměnná PLO_CHS zde vykazuje nižší významnost, což může naznačovat, že obsah uhlíku v organominerálním horizontu je více vázán na charakter porostu a jeho strukturní heterogenitu než na stanovištní kategorizaci. Naopak proměnná MEAN se v tomto modelu uplatnila pouze marginálně. U modelu predikujícího součet obsahu uhlíku v horizontech O a A dominovala zejména spojitá proměnná MAX, přičemž bezprostředně za ní následovala proměnná PLO_CHS. Dalšími prediktory byly STD a MEAN, zatímco APT měla v tomto modelu relativně nižší významnost. Tento výsledek naznačuje, že při agregaci obou horizontů se současně uplatňuje jak vliv celkové výškové struktury porostu (zejména horní výšky porostu), tak stanovištní a hospodářské charakteristiky zahrnuté v prediktoru PLO_CHS.

Záporné hodnoty permutation importance (např. u STD v modelu pro horizont O nebo u MEAN v modelu pro horizont A) neindikují chybu modelu ani nevhodnost daných proměnných. Mohou být důsledkem kolinearity mezi prediktory, přítomnosti šumu nebo slabého, avšak stabilního vlivu proměnné, který se projeví pouze v kombinaci s ostatními prediktory. Z tohoto důvodu nejsou prediktory s nízkou či zápornou hodnotou delta_RMSE automaticky vyřazovány, protože jejich odstranění nemusí vést ke zlepšení generalizační schopnosti modelu a současně by mohlo snížit konzistenci a interpretovatelnost modelů mezi jednotlivými horizonty.

Jedním z hlavních limitů použité metodiky je znáhodnění souřadnic jednotlivých lokalit odběru v gridu 1×1 km, v jehož rámci bylo nutné asociovat prediktory k jednotlivým lokalitám prostřednictvím jejich průměrování v ploše čtverce 1×1 km, případně v případě vektorových dat přiřazením hodnoty prediktoru s největším plošným překryvem. Tento aspekt se projevil zejména při vyhodnocení bilance zásob uhlíku pro vybrané cílové hospodářské soubory (např. CHS 1, 2, 3, 13 a 19), u nichž se dle zadaného scénáře nemění dřevinná skladba, a tedy ani očekávaná zásoba uhlíku. Prostorově sousedící CHS však mohou v rámci scénáře dřevinnou skladbu měnit, přičemž tato změněná hodnota v rámci gridu 1×1 km může částečně zasahovat i do CHS, které se nemění. Výsledkem jsou minoritní změny v celkové i průměrné bilanci zásoby uhlíku jednotlivých CHS v rámci příslušné přírodní lesní oblasti.

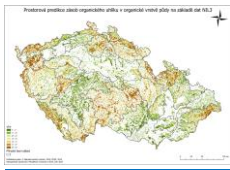
Zvýšení predikční přesnosti modelování zásob uhlíku by bylo možné dosáhnout zavedením dalších prediktorů nebo využitím prostorově explicitních variant algoritmu Random Forest (např. spatialRF). Takový přístup by však mohl vést k potlačení vlivu dřevinné skladby ve prospěch prostorové autokorelace, čímž by se model stal méně vhodným pro scénářové hodnocení bilance uhlíku. Alternativní technologické přístupy založené na prostředí Python a knihovně scikit-learn by rovněž umožnily využití širší škály metod

strojového učení, jako jsou Support Vector Machine, Gaussian Process Regression, Gradient Boosting Machine nebo metody typu XGBoost, LightGBM a CatBoost (Pedregosa et al., 2011). Zejména algoritmus CatBoost se jeví jako perspektivní díky své schopnosti efektivně pracovat s kombinací spojitéch a kategoričkových prediktorů a stabilnímu chování při jejich vzájemné interakci (Dorogush et al., 2018).

8.2.3 Výsledné mapy zásoby a bilance uhlíku

Výsledné mapy zásob a bilance uhlíku jsou uvedeny v přílohách na následujících odkazech:

[Aktuální zásoby organického uhlíku v organické vrstvě půdy \(O\)](#)



[Aktuální zásoby organického uhlíku v organominerální vrstvě půdy \(A\)](#)

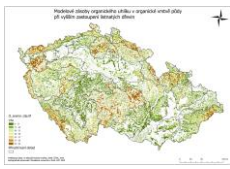


[Aktuální zásoby organického uhlíku v organické a organominerální vrstvě půdy \(O+A\)](#)



[Příloha 4 Zásoby organického uhlíku v povrchových vrstvách lesních půd pro „listnatý scénář“](#)

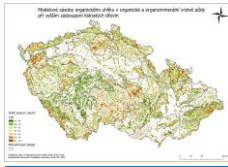
[Modelové zásoby organického uhlíku v organické vrstvě půdy \(O\) při vyšším zastoupení listnatých dřevin](#)



[Modelové zásoby organického uhlíku v organominerální vrstvě půdy \(A\) při vyšším zastoupení listnatých dřevin](#)



Modelové zásoby organického uhlíku v organické a organominerální vrstvě půdy (O+A) při vyšším zastoupení listnatých dřevin



Příloha 5 Bilance zásob organického uhlíku v povrchových vrstvách lesních půd pro „listnatý“ scénář

Bilance zásob organického uhlíku v organické vrstvě půdy (O) při vyšším zastoupení listnatých dřevin („listnatý scénář“ minus aktuální stav)



Bilance zásob organického uhlíku v organominerální vrstvě půdy (A) při vyšším zastoupení listnatých dřevin („listnatý scénář“ minus aktuální stav)



Bilance zásob organického uhlíku v organické a organominerální vrstvě půdy (O+A) při vyšším zastoupení listnatých dřevin („listnatý scénář“ minus aktuální stav)



Podobně jako z lineárních regresních modelů, i z map vyplývá význam dřevinné skladby pro zásoby uhlíku v jednotlivých vrstvách půdy.

V aktuálním stavu (ke stavu NIL3) zásoby uhlíku v lesní půdě ČR dosahují cca 52,7 mil. t TOC v organické vrstvě a 85,8 mil. t TOC v organominerálním horizontu. Z odhadované celkové zásoby 472,2 mil. t TOC v celém půdním profilu (Vašát a kol. 2025) tak zásoba ve sledovaných povrchových horizontech (O + A) činí téměř 30 %. Změna dřevinné skladby ve prospěch listnatých porostních typů potom vede ke snížení zásoby TOC v organické vrstvě

cca o 3,168 mil. t TOC (o 6,02 % vzhledem k aktuálnímu stavu) a ke zvýšení zásoby TOC v organominerálním horizontu cca o 4,479 mil. t TOC (o 4,48 % vzhledem k aktuálnímu stavu). Celková bilance svrchních půdních vrstev v tomto modelu byla kladná, a to o +1,31 mil. t TOC ve prospěch organominerálních půdních vrstev (navýšení o 0,95 % vzhledem k aktuálnímu stavu). Sumární bilanci za organické a organominerální horizonty je však nutné vnímat pouze jako orientační vzhledem ke komplexitě ekosystémové odezvy na změny v dřevinné skladbě v daných podmínkách ekotopu definované edatopem a ekotopem.

Záporná bilance v organické vrstvě byla nad očekávání prakticky ve všech PLO. Celkové zásoby a také bilance pro jednotlivé PLO jsou v logice zvolené metodiky a přírodních podmínek ČR definovány zejména rozlohou PLO, s mírnými nuancemi. Největší hodnoty výsledné bilance jsou tak v PLO 16 (Českomoravská vrchovina), 10 (Středočeská pahorkatina) a 6 (Západočeská pahorkatina) a v dalších PLO jednak s větší rozlohou, ale také s georeliéfovou dominancí středních a vyšších poloh, v nichž dochází k největšímu projevu dřevinné skladby směrem k úbytku zásoby TOC v organických horizontech (Příloha 6). V organominerálním horizontu byla bilance výhradně kladná pro všechny PLO, až na výjimky se shodným pořadím jako v případě organického horizontu. Průměrné hektarové zásoby přitom pro všechny PLO nabývají velmi podobných hodnot v rozsahu 20,17–21,68 t/ha pro opadový horizont a 33,05–36,52 t/ha pro organominerální horizont.

Z hlediska CHS byla bilance v organické vrstvě prakticky výhradně záporná (Příloha 7), s nejvýraznější odezvou v CHS ve středních polohách. V organominerálním horizontu byla bilance zásoby TOC ve všech CHS kladná, s nejvyššími hodnotami v CHS 41 až 47 a 53 až 57. Tyto hodnoty nelze z hlediska metodického přístupu přisoudit výhradně samotnému „listnatému“ scénáři, ale také procentuálnímu zastoupení konkrétních CHS v rámci ČR, nicméně zde se výrazněji odrážejí jednak přírodní podmínky a jednak nepřírozená dřevinná skladba, ve středních polohách výrazně směřovaná k jehličnanům resp. ke smrku ztepilému.

Navzdory poměrně nízkému vlivu APT na velikost zásob půdního uhlíku, v porovnání s náhodnými prediktory (stanovištní podmínky – PLO, CHS a další nadstavbové typologické jednotky), simulované změny se změnou zastoupení APT „uvnitř“ nejvíce ohrožených CHS v rámci PLO jsou z pohledu regionálního velmi markantní. Zároveň je těmito výsledky podtržen a demonstrován meliorační význam listnatých dřevin, které výrazně přispívají k přesunu zásob půdního uhlíku z nadložní organické do svrchní organominerální vrstvy půdy a tím k jejich potenciální stabilizaci a vyšší odolnosti vůči narůstajícímu riziku negativních dopadů disturbancí na lesní ekosystém v souvislosti s klimatickou změnou, jako jsou např. požáry, nebo také odlesnění.

9 Závěr

Mapa se specializovaným obsahem „*Bilance uhlíku v lesních porostech ve vztahu k lesnickému managementu*“ vznikla jako výstup projektu NAZV *Adaptace lesního hospodářství pro udržitelné využívání přírodních zdrojů*. Mapa resp. soubor map zahrnuje informace o aktuální zásobě TOC a o modelovaných změnách těchto zásob ve vazbě na lesnický management, resp. na změny v dřevinné skladbě. Výsledky jsou zobrazeny pro dvě svrchní vrstvy lesní půdy: organickou (O) a organominerální (A) a také souhrnně pro O + A. Pro vyjádření zásoby a bilance TOC byla široká škála potenciálních prediktorů zúžena do následujících tří, při přijatelném, resp. minimálním poklesu predikční síly modelů:

(1) agregovaný porostní typ, což je pracovní jednotka vytvořená za účelem agregace početných strukturovaných porostních typů na pětibodovou škálu;

(2) cílové hospodářské soubory, vyjadřující stálé resp. blízké stanovištní podmínky a tím také produkční sílu stanoviště;

(3) a přírodní lesní oblasti územně vymezuující přírodě blízké podmínky.

Celkově výsledky potvrzují, že kategorie prostředí (CHS, PLO) jsou zásadními prediktory prostorového rozložení TOC, zatímco spojitě spektrální ukazatele (MEAN, STD) přispívají spíše k upřesnění variability v rámci obvodů. Slabší predikční schopnost v horizontu A lze vysvětlit menší konzistencí mezi dálkovými prediktory a skutečným obsahem uhlíku v minerální vrstvě půdy.

Z výsledků vyplývá, že změna druhové skladby ve smyslu podpory zastoupení listnatých dřevin vede ke snížení zásoby organického uhlíku v organické půdní vrstvě, avšak také k výraznému zvýšení zásoby v organominerálním horizontu. Celková bilance v rámci dané části půdního tělesa je tak kladná.

10 Přílohy

PŘÍLOHA 1.	SKRIPTY POUŽITÉ PŘI PROSTOROVÉM MODELOVÁNÍ ZÁSOB UHLÍKU	37
PŘÍLOHA 1.1	PYTHON SKRIPT S KNIHOVNOU ARCPY POUŽITÝ PRO VÝPOČET HODNOT PREDIKTORŮ Z DAT SE ZNÁHODNĚNÝMI SOUŘADNICEMI NLI3.	37
PŘÍLOHA 1.2	PYTHON SKRIPT PRO ROZPARSOVÁNÍ STRUKTUROVANÝCH POROSTNÍCH TYPŮ (SPT) NA AGREGOVANÉ POROSTNÍ TYPY (APT).....	38
PŘÍLOHA 1.3	R SKRIPT PRO TRÉNOVÁNÍ A NASAZENÍ RANDOM FOREST MODELU – PREDIKCE ZÁSOPY TOC V ODDĚLENĚ PRO O HORIZONT, A HORIZONT A O+A.	40
PŘÍLOHA 1.4	PYTHON SKRIPT PRO NOVÉ PROPORČNÍ ROZDĚLENÍ AGREGOVANÝCH POROSTNÍCH TYPŮ (APT) DO PŘÍSLUŠNÝCH CHS V RÁMCI PLO PŘI UPLATNĚNÍ „LISTNATÉHO“ SCÉNÁŘE.	49
PŘÍLOHA 1.5	SKRIPT PRO NASAZENÍ ULOŽENÉHO MODELU RANDOM FOREST NA SCÉNÁŘOVÝ GRID	53
PŘÍLOHA 2.	DOPLŇKOVÉ TABELÁRNÍ A GRAFICKÉ VÝSTUPY Z REGRESNÍ ANALÝZY	58
PŘÍLOHA 3.	MAPOVÉ VÝSTUPY ZÁSOPY UHLÍKU VE SVRCHNÍCH VRSTVÁCH PŮDY: ZÁSOPY ORGANICKÉHO UHLÍKU V POVRCHOVÝCH VRSTVÁCH LESNÍCH PŮD.....	60
PŘÍLOHA 3.1	AKTUÁLNÍ ZÁSOPY ORGANICKÉHO UHLÍKU V ORGANICKÉ VRSTVĚ PŮDY (O)	60
PŘÍLOHA 3.2	AKTUÁLNÍ ZÁSOPY ORGANICKÉHO UHLÍKU V ORGANOMINERÁLNÍ VRSTVĚ PŮDY (A)	61
PŘÍLOHA 3.3	AKTUÁLNÍ ZÁSOPY ORGANICKÉHO UHLÍKU V ORGANICKÉ A ORGANOMINERÁLNÍ VRSTVĚ PŮDY (O+A)	62
PŘÍLOHA 4.	ZÁSOPY ORGANICKÉHO UHLÍKU V POVRCHOVÝCH VRSTVÁCH LESNÍCH PŮD PRO „LISTNATÝ SCÉNÁŘ“	63
PŘÍLOHA 4.1	MODELOVÉ ZÁSOPY ORGANICKÉHO UHLÍKU V ORGANICKÉ VRSTVĚ PŮDY (O) PŘI VYŠŠÍM ZASTOUPENÍ LISTNATÝCH DŘEVIN	63
PŘÍLOHA 4.2	MODELOVÉ ZÁSOPY ORGANICKÉHO UHLÍKU V ORGANOMINERÁLNÍ VRSTVĚ PŮDY (A) PŘI VYŠŠÍM ZASTOUPENÍ LISTNATÝCH DŘEVIN	64
PŘÍLOHA 4.3	MODELOVÉ ZÁSOPY ORGANICKÉHO UHLÍKU V ORGANICKÉ A ORGANOMINERÁLNÍ VRSTVĚ PŮDY (O+A) PŘI VYŠŠÍM ZASTOUPENÍ LISTNATÝCH DŘEVIN	65
PŘÍLOHA 5.	BILANCE ZÁSOB ORGANICKÉHO UHLÍKU V POVRCHOVÝCH VRSTVÁCH LESNÍCH PŮD PRO „LISTNATÝ“ SCÉNÁŘ	66
PŘÍLOHA 5.1	BILANCE ZÁSOB ORGANICKÉHO UHLÍKU V ORGANICKÉ VRSTVĚ PŮDY (O) PŘI VYŠŠÍM ZASTOUPENÍ LISTNATÝCH DŘEVIN („LISTNATÝ SCÉNÁŘ“ MÍNUS AKTUÁLNÍ STAV).....	66
PŘÍLOHA 5.2	BILANCE ZÁSOB ORGANICKÉHO UHLÍKU V ORGANOMINERÁLNÍ VRSTVĚ PŮDY (A) PŘI VYŠŠÍM ZASTOUPENÍ LISTNATÝCH DŘEVIN („LISTNATÝ SCÉNÁŘ“ MÍNUS AKTUÁLNÍ STAV).....	67
PŘÍLOHA 5.3	BILANCE ZÁSOB ORGANICKÉHO UHLÍKU V ORGANICKÉ A ORGANOMINERÁLNÍ VRSTVĚ PŮDY (O+A) PŘI VYŠŠÍM ZASTOUPENÍ LISTNATÝCH DŘEVIN („LISTNATÝ SCÉNÁŘ“ MÍNUS AKTUÁLNÍ STAV).....	68
PŘÍLOHA 6.	BILANCE ZÁSOB ORGANICKÉHO UHLÍKU (TOC) NA ÚROVNI ČR PRO „LISTNATÝ“ SCÉNÁŘ DLE PŘÍRODNÍCH LESNÍCH OBLASTÍ (PLO): O – ORGANICKÁ VRSTVA, A – ORGANOMINERÁLNÍ VRSTVA PŮDY.....	69
PŘÍLOHA 7.	BILANCE ZÁSOB ORGANICKÉHO UHLÍKU (TOC) NA ÚROVNI ČR PRO „LISTNATÝ“ SCÉNÁŘ) DLE CÍLOVÝCH HOSPODÁŘSKÝCH SOUBORŮ (CHS): O – ORGANICKÁ VRSTVA, A – ORGANOMINERÁLNÍ VRSTVA PŮDY. HODNOTY ZOBRAZUJÍ BILANCE PRO CELOREPUBLIKOVÉ VÝMĚRY CHS.	71

Příloha 1. Skripty použité při prostorovém modelování zásob uhlíku

Příloha 1.1 Python skript s knihovnou arcpy použitý pro výpočet hodnot prediktorů z dat se znáhodněnými souřadnicemi NLI3.

```
# =====
# GENEROVÁNÍ ČTVERCOVÝCH POLYGONŮ KOLEM BODŮ (např. 1x1 km)
# -----
# Účel:
#   Z bodové vrstvy (např. reprezentující vzorkovací body)
#   vytvoří polygonovou vrstvu čtverců o definované velikosti.
#   Každý bod je středem svého čtverce.
#
# Typické využití:
#   - tvorba gridu pro prostorové agregace (např. 1 km2)
#   - definice okolí měření nebo odběrných míst
#   - vizualizace rozsahu vlivu bodů
#
# Vstup:
#   - body.shp ... bodová vrstva s definovaným souřadnicovým systémem
#
# Výstup:
#   - body_ctverec_1km.shp ... polygonová vrstva čtverců (1x1 km)
#
# Poznámky:
#   - Vzdálenosti jsou v mapových jednotkách daného souřadnicového systému
#     (např. S-JTSK → metry).
#   - Proměnná ,d' určuje polovinu délky strany čtverce.
#     Pro 1 km čtverec tedy d = 500 (m).
# =====

import arcpy

# ---- Cesty a parametry ----
in_fc = "body.shp" # vstupní bodová vrstva
out_fc = "body_ctverec_1km.shp" # výstupní polygonová vrstva
d = 500 # polovina délky strany čtverce (m)

# ---- Kontrola a vytvoření výstupní vrstvy ----
sr = arcpy.Describe(in_fc).spatialReference # převzít SRS z bodů
arcpy.management.CreateFeatureclass(
    out_path = arcpy.env.workspace or ".", # výstupní workspace
    out_name = out_fc,
    geometry_type = "POLYGON",
    spatial_reference = sr
)

# ---- Hlavní část: pro každý bod vytvoř polygon čtverce ----
with arcpy.da.SearchCursor(in_fc, ["SHAPE@XY"]) as cursor, \
     arcpy.da.InsertCursor(out_fc, ["SHAPE@"]) as icursor:

    for (x, y) in cursor:
        # Souřadnice rohů čtverce (v pořadí proti směru hodinových ručiček)
        array = arcpy.Array([
            arcpy.Point(x - d, y - d), # jihozápadní roh
            arcpy.Point(x - d, y + d), # severozápadní roh
            arcpy.Point(x + d, y + d), # severovýchodní roh
            arcpy.Point(x + d, y - d), # jihovýchodní roh
            arcpy.Point(x - d, y - d) # návrat pro uzavření polygonu
        ])

        # Vytvoření polygonu a vložení do výstupní vrstvy
        polygon = arcpy.Polygon(array, sr)
        icursor.insertRow([polygon])

print("Hotovo: vytvořeny čtverce 1x1 km kolem bodů.")
```

Příloha 1.2 Python skript pro rozparsování strukturovaných porostních typů (SPT) na agregované porostní typy (APT).

```
# =====
# Automatické určení agregovaného porostního typu (APT)
# z kódu porostního typu (SPT) ve vrstvě porostních skupin
# -----
# Účel:
# Z textového pole SPT (např. „C2 D5 M9e“) odvodí:
# - procentuální podíl jehličnanů (PCT_J)
# - odpovídající kategorie APT
#
# Princip:
# 1) SPT je tvořen kombinacemi znaků:
# [C,D,M,Z,P,V] = podíl hospodářského významu (váha)
# + kód porostního typu (např. 2, 5e, 9x)
# 2) Každé písmeno má váhu dle významu dřeviny:
# C=95, D=80, M=60, Z=40, P=20, V=5
# 3) Porostní typy jsou rozděleny na:
# - jehličnaté (1-4)
# - listnaté (5-9)
# 4) Z textu se spočítá vážený podíl jehličnanů v %.
# 5) Podíl se převede na pětistupňovou kategorii APT:
# ≥90 % ..... jehličnatý
# 70-89 % ..... smíšený jehličnatý
# 30-69 % ..... smíšený
# 10-30 % ..... smíšený listnatý
# <10 % ..... listnatý
#
# Vstup:
# - feature class (např. z GDB) s polem 'SPT'
# Výstup:
# - doplněné pole 'PCT_J' (DOUBLE) = % jehličnanů
# - doplněné pole 'APT' (TEXT) = slovní typ
# =====

import arcpy, re

# ---- Cesty a názvy polí -----
in_fc = r"D:\NLI\Nmap_toc_nli\Nmap_toc_nli.gdb\SPT_CIHO_PLO"
field_spt = "SPT" # vstupní textový kód (např. "C2 D5 M9e")
field_apt = "APT" # výstupní kategorie (např. "smíšený listnatý")
field_pct = "PCT_J" # výstupní % jehličnanů

# ---- Definice porostních typů -----
# Jehličnaté a listnaté porostní typy dle číselníku SPT
conifer_pts = {"1", "1e", "2", "2e", "3", "3e", "4", "4k", "4x"}
broadleaf_pts = {"5", "5e", "6", "7", "7e", "8", "9", "9e", "9x"}
valid_pts = conifer_pts.union(broadleaf_pts)

# Váhy podle významu dřeviny (vyšší = větší zastoupení)
weight_map = {"C":95, "D":80, "M":60, "Z":40, "P":20, "V":5}

# Regulární výraz pro rozpoznání páru (písmeno + kód PT)
# např. "C2" nebo "M9e"
token_regex = re.compile(r'([CDMZPV])\s*([0-9]+[ekx]?)', re.I)

# =====
# Funkce: výpočet podílu jehličnanů (%)
# =====
def conifer_share_from_spt(spt):
    """
    Spočítá vážený podíl jehličnatých porostních typů (v %)
    z textového kódu SPT, např. 'C2 D5 M9e'.
    """
    if not spt:
        return 0.0
    tokens = token_regex.findall(str(spt))
    tot = conf = 0
    for ch, pt in tokens:
        w = weight_map.get(ch.upper(), 0)
        if pt.lower() in valid_pts:
            tot += w
        if pt.lower() in conifer_pts:
```

```

        conf += w
    return 100 * conf / tot if tot > 0 else 0.0

# =====
# Funkce: kategorizace APT dle % jehličnanů
# =====
def categorize(pct):
    """
    Vrací slovní kategorii APT podle podílu jehličnanů.
    """
    if pct >= 90:         return "jehličnatý"
    if 70 <= pct < 90:   return "smíšený jehličnatý"
    if 30 < pct < 70:    return "smíšený"
    if 10 < pct <= 30:   return "smíšený listnatý"
    return "listnatý"

# =====
# Hlavní část: kontrola a doplnění polí + výpočet
# =====

# 1) Přidej pole, pokud neexistují
if field_apt not in [f.name for f in arcpy.ListFields(in_fc)]:
    arcpy.management.AddField(in_fc, field_apt, "TEXT", field_length=30)
if field_pct not in [f.name for f in arcpy.ListFields(in_fc)]:
    arcpy.management.AddField(in_fc, field_pct, "DOUBLE")

# 2) Projdi všechny řádky a doplň výpočty
with arcpy.da.UpdateCursor(in_fc, [field_spt, field_pct, field_apt]) as cur:
    for spt, pct, apt in cur:
        p = conifer_share_from_spt(spt) # podíl jehličnanů
        a = categorize(p)                # slovní kategorie
        cur.updateRow([spt, p, a])

print("Hotovo: pole APT a PCT_J byla doplněna.")

```

Příloha 1.3 R skript pro trénování a nasazení Random Forest modelu – predikce zásoby TOC v odděleně pro O horizont, A horizont a O+A.

```
# =====
# Prostorová predikce zásoby půdního organického uhlíku (TOC)
# Random Forest (tidymodels + ranger) – pouze varianta FULL
#
# Účel:
# Skript natrénuje tři samostatné regresní modely Random Forest pro:
# (1) A-horizont (TOC_A_new)
# (2) O-horizont (TOC_O_new)
# (3) Součet A+O (SUM_TOC_ne)
# a následně aplikuje natrénované modely na polygonový grid pro tvorbu mapových
# výstupů (atributy A_pred, O_pred, SUM_pred).
#
# Vstupy:
# - body_A.shp ... bodová tréninková data pro A-horizont
# - body_O.shp ... bodová tréninková data pro O-horizont
# - body_A_O.shp ... bodová tréninková data pro součet A+O
# - grid_TOC.shp ... cílová polygonová vrstva (grid) s prediktory
#
# Prediktory (FULL):
# Numerické: MEAN, STD, MAX
# Kategoriální: APT, PLO_CHS
#
# Souřadnicový systém:
# Všechny vrstvy jsou sjednoceny do EPSG:5514 (S-JTSK / Křovák East North).
#
# Modelování:
# - Hyperparametry (trees, mtry, min_n) jsou laděny pomocí 10-fold CV
# (latinský hyperkrychlový výběr kombinací).
# - Finální model je následně fitován na všech dostupných tréninkových datech
# (po ošetření typů a chybějících hodnot).
#
# Hodnocení přesnosti:
# - Reportované metriky: R2, RMSE, MAE, nRMSE_mean, nRMSE_range.
# - Metriky jsou počítány na trénovacích datech (resubstituce) jako rychlá
# kontrola kvality. Výběr hyperparametrů probíhá přes CV (optimalizace RMSE).
#
# Interpretace modelu (významnost proměnných):
# (1) Permutační významnost v křížové validaci (CV permutation importance)
# - prediktory se permutují v testovací části foldů
# - významnost = zhoršení RMSE po permutaci (delta RMSE)
# - férovější vůči numerickým i kategoriálním proměnným
# - pouze report; žádné proměnné se v této verzi nevyřazují
# (2) Impurity importance z finálního modelu
# - doplňkový report z enginu ranger (může být zkreslený / biased)
#
# Výstupy:
# - Nový shapefile s predikcemi na gridu (nepřepisuje původní):
# A_pred, O_pred, SUM_pred
# - Souhrn metrik do CSV (A, O, A+O)
# - CSV s významností proměnných:
# * cv-perm (robustnější) a impurity (doplňkové)
# - Uložení modelů (workflow .rds) + metadat (JSON) do models_rf_full/
#
# Reprodukovatelnost a ochrana proti přepsání:
# Skript cíleně zapisuje do NOVÝCH souborů (nový název vrstvy), aby nedošlo
# k přepsání původních vstupních dat nebo předchozích výstupů.
# =====

suppressPackageStartupMessages({
  library(sf); library(dplyr); library(stringr)
  library(tidymodels); library(ranger); library(readr)
  library(future); library(tidyr); library(purrr)
  library(jsonlite)
})

# ---- CESTY K DATŮM A VÝSTUPŮM -----
# base_dir: pracovní adresář s daty a výstupy (shapefile, CSV, modely)
base_dir <- "D:/NLI/final"

# Tréninková data (bodové vrstvy) pro jednotlivé cílové proměnné
path_A <- file.path(base_dir, "body_A.shp")
```

```

path_O      <- file.path(base_dir, "body_0.shp")
path_A0     <- file.path(base_dir, "body_A_0.shp")

# Cílový grid (polygonová vrstva) s prediktory pro mapovou predikci
path_grid  <- file.path(base_dir, "grid_TOC.shp")

# Výstupní adresář (zde shodný s base_dir)
out_dir    <- base_dir

# ---- NÁZVY NOVÝCH VÝSTUPŮ (OCHRANA PŘED PŘEPSÁNÍM) ----
# out_layer/out_shp: nový název výstupní vrstvy s predikcemi
out_layer  <- "grid_TOC_pred_FULLL2"
out_shp    <- file.path(out_dir, paste0(out_layer, ".shp"))

# Souhrn metrik pro všechny tři modely (A, O, A+O)
metrics_csv <- file.path(base_dir, "rf_metrics_summary_FULLL2.csv")

# Složka pro uložení finálních modelů a metadat
models_dir <- file.path(base_dir, "models_rf_full")
dir.create(models_dir, showWarnings = FALSE, recursive = TRUE)

# ---- DEFINICE PREDIKTORŮ (FULL) ----
# Master seznam prediktorů používaných ve všech modelech
num_preds <- c("MEAN", "STD", "MAX")
cat_preds <- c("APT", "PLO_CHS")
all_preds <- c(num_preds, cat_preds)

# ---- NASTAVENÍ PARALELIZACE ----
# Hlavní běh skriptu je sekvenční (kvůli stabilitě), paralelizace se používá
# cíleně pouze ve výpočetně náročných částech (CV/importance).
future::plan(sequential)

# =====
# UTILITNÍ FUNKCE: ČTENÍ, SJEDNOCENÍ CRS, TYPY, KONTROLY
# =====

# Načtení vektorové vrstvy a sjednocení CRS do EPSG:5514.
read_sf_5514 <- function(path){
  s <- st_read(path, quiet = TRUE)
  if (is.na(st_crs(s))) s <- st_set_crs(s, 5514)
  if (!identical(st_crs(s)$epsg, 5514L)) s <- st_transform(s, 5514)
  s
}

# Převod list-sloupců (vznikají např. při čtení některých SHP) na skalární hodnoty,
# aby bylo možné bezpečně provádět převody typů a modelování.
squash_list_cols <- function(df){
  for(nm in names(df)){
    if (is.list(df[[nm]])) {
      df[[nm]] <- vapply(df[[nm]], function(z){
        if (length(z)==0 || is.null(z)) NA_character_ else as.character(z[[1]])
      }, FUN.VALUE = character(1))
    }
  }
  df
}

# Výběr prediktorů z atributové tabulky podle seznamu `keep`
# a převod případných list-sloupců na skalární hodnoty (kvůli kompatibilitě).
unify_predictors <- function(sfobj, keep = all_preds){
  df <- sf::st_drop_geometry(sfobj) |> as.data.frame(stringsAsFactors = FALSE)
  keep <- intersect(names(df), keep)
  df <- df[, keep, drop = FALSE]
  squash_list_cols(df)
}

# Převod datových typů:
# - numerické prediktory -> numeric (s podporou desetinné čárky)
# - kategoriální prediktory -> factor
cast_types <- function(df, num_cols = num_preds, cat_cols = cat_preds){
  df <- squash_list_cols(df)
  for (nm in intersect(num_cols, names(df))) {
    df[[nm]] <- suppressWarnings(as.numeric(gsub(",", ".", as.character(df[[nm]]))))
  }
}

```

```

}
for (nm in intersect(cat_cols, names(df))) {
  df[[nm]] <- as.factor(as.character(df[[nm]]))
}
df
}

# Diagnostika typů prediktorů: deklarovaný vs. skutečný typ + počty NA/unikátů/úrovní.
# mismatch = TRUE označuje sloupec, kde deklarace (num/cat) neodpovídá datovému typu.
inspect_pred_types <- function(df, num_cols = num_preds, cat_cols = cat_preds) {
  nm <- names(df)
  declared <- ifelse(nm %in% num_cols, "numeric",
                    ifelse(nm %in% cat_cols, "categorical", "other"))
  tibble::tibble(
    column      = nm,
    declared    = declared,
    class       = vapply(df, function(x) paste(class(x), collapse="/"), character(1)),
    detected    = ifelse(vapply(df, is.factor, logical(1)), "categorical",
                        ifelse(vapply(df, is.numeric, logical(1)), "numeric", "other")),
    n_levels    = ifelse(vapply(df, is.factor, logical(1)), vapply(df, nlevels, integer(1)),
NA_integer_),
    n_unique    = vapply(df, function(x) dplyr::n_distinct(x, na.rm = TRUE), integer(1)),
    n_na        = vapply(df, function(x) sum(is.na(x)), integer(1)),
    example_lvls= ifelse(vapply(df, is.factor, logical(1)),
                        vapply(df, function(x) paste(utils::head(levels(x), 6), collapse=","),
character(1)),
                        NA_character_),
    mismatch    = ifelse(declared == "numeric" & !vapply(df, is.numeric, logical(1)), TRUE,
                        ifelse(declared == "categorical" & !vapply(df, is.factor, logical(1)), TRUE,
FALSE))
  ) |>
  dplyr::arrange(dplyr::desc(mismatch), declared, column)
}

# Kontrola kvality vstupních dat pro daný dataset:
# - sjednocení a převod prediktorů
# - výpis diagnostiky typů do konzole
# - volitelně uložení reportu do CSV (audit dat)
check_dataset <- function(sfobj, label, keep_preds = all_preds, out_dir = NULL){
  X_raw <- unify_predictors(sfobj, keep = keep_preds)
  X      <- cast_types(X_raw)
  rpt    <- inspect_pred_types(X)
  cat("\n--- KONTROLA TYPŮ ---", label, "---\n"); print(rpt, n = nrow(rpt))
  if (any(rpt$mismatch)) warning("Nesoulad typů v ", label, " (viz sloupec 'mismatch').")
  if (!is.null(out_dir)) readr::write_csv(rpt, file.path(out_dir, paste0("type_report_", label,
".csv")))
  invisible(X)
}

# Výběr cílového sloupce (target) s možností fallbacků:
# - pokud target_col neexistuje, použije se první dostupný z fallback seznamu
# - pokud není k dispozici ani fallback, skript skončí s chybou a vypíše názvy sloupců
pick_target_col <- function(sfobj, target_col, fallbacks = NULL) {
  cols <- names(sf::st_drop_geometry(sfobj))
  if (target_col %in% cols) return(target_col)
  if (!is.null(fallbacks)) {
    hit <- fallbacks[fallbacks %in% cols]
    if (length(hit) > 0) {
      message("Target '", target_col, "' nenalezen. Používám fallback: '", hit[1], "'")
      return(hit[1])
    }
  }
  stop("Target sloupec '", target_col, "' nebyl nalezen. Dostupné sloupce např.: ",
      paste(head(cols, 50), collapse = ", "), if (length(cols) > 50) ", ..." else "")
}

# Definice receptu (pre-processing) pro RF:
# - imputace numerik mediánem
# - ošetření nových/neznámých kategorií (novel/unknown)
# - sloučení málo častých kategorií (threshold 1 %)
# - odstranění prediktorů s nulovou variabilitou
rf_recipe_for <- function(train_df, target_col){
  recipe(stats::as.formula(paste(target_col, "~ .")), data = train_df) |>
  step_impute_median(all_numeric_predictors()) |>

```

```

    step_novel(all_nominal_predictors()) |>
    step_unknown(all_nominal_predictors()) |>
    step_other(all_nominal_predictors(), threshold = 0.01) |>
    step_zv(all_predictors())
  }

# Výpočet reportovaných metrik kvality predikce:
# - R2, RMSE, MAE
# - nRMSE_mean = RMSE / mean(obs)
# - nRMSE_range = RMSE / (max(obs)-min(obs))
# Pozn.: nRMSE je bezrozměrné, umožňuje srovnání mezi horizonty s odlišným rozsahem.
calc_metrics <- function(obs, pred){
  obs <- as.numeric(obs); pred <- as.numeric(pred)
  ok <- is.finite(obs) & is.finite(pred)
  obs <- obs[ok]; pred <- pred[ok]
  if (length(obs) < 2) {
    return(tibble(R2=NA_real_, RMSE=NA_real_, MAE=NA_real_, NRMSE_mean=NA_real_, NRMSE_range=NA_real_))
  }
  rmse <- sqrt(mean((obs - pred)^2))
  mae <- mean(abs(obs - pred))
  r2 <- 1 - sum((obs - pred)^2) / sum((obs - mean(obs))^2)
  mu <- mean(obs)
  rng <- max(obs) - min(obs)
  nrmse_mean <- if (is.finite(mu) && abs(mu) > 0) rmse / mu else NA_real_
  nrmse_range <- if (is.finite(rng) && abs(rng) > 0) rmse / rng else NA_real_
  tibble(R2=r2, RMSE=rmse, MAE=mae, NRMSE_mean=nrmse_mean, NRMSE_range=nrmse_range)
}

# =====
# CV PERMUTAČNÍ VÝZNAMNOST (ROBUSTNĚJŠÍ) – POUZE REPORT
# =====
# Cíl:
# Odhadnout významnost prediktorů férovým způsobem pro numerické i kategoriální
# proměnné: prediktor se v testovací části foldů náhodně permutuje a měří se,
# o kolik se zhorší RMSE (delta RMSE).
#
# Poznámky:
# - Používá se stejná struktura pre-processing receptu jako při trénování.
# - Výpočet probíhá nad v-fold CV (default 10) a pro stabilitu se opakuje
#   několikrát (repeats).
# - Slouží výhradně k interpretaci; žádná redukce prediktorů se zde neprovádí.
cv_permutation_importance <- function(train_df, target_col, final_params,
                                     v = 10, repeats = 3, seed = 123,
                                     num_threads = max(1, future::availableCores() - 1)){
  set.seed(seed)
  folds <- vfold_cv(train_df, v = v)

  # Specifikace RF s fixními parametry (z tuning best) a bez interní importance,
  # protože významnost se zde odvozuje externě permutací.
  spec <- rand_forest(
    trees = final_params$trees,
    mtry = final_params$mtry,
    min_n = final_params$min_n
  ) |>
  set_mode("regression") |>
  set_engine("ranger",
             importance = "none",
             num.threads = num_threads,
             respect.unordered.factors = "ignore")

  rec <- rf_recipe_for(train_df, target_col)

  # Výpočet pro jeden fold:
  # - fit na trénovací části
  # - baseline RMSE na testu
  # - pro každý prediktor: permutace + RMSE => delta RMSE
  one_fold <- function(split){
    tr <- analysis(split)
    te <- assessment(split)

    wf <- workflow() |>
      add_recipe(rec, blueprint = hardhat::default_recipe_blueprint(allow_novel_levels = TRUE)) |>
      add_model(spec)
  }
}

```

```

fit_wf <- fit(wf, data = tr)

pred0 <- predict(fit_wf, new_data = te)$pred
base_rmse <- sqrt(mean((te[[target_col]] - pred0)^2))

preds <- setdiff(names(te), target_col)

map_dfr(seq_len(repeats), function(r){
  map_dfr(preds, function(pn){
    te_perm <- te
    te_perm[[pn]] <- sample(te_perm[[pn]])
    predp <- predict(fit_wf, new_data = te_perm)$pred
    rmse_p <- sqrt(mean((te[[target_col]] - predp)^2))
    tibble(var = pn, delta_rmse = rmse_p - base_rmse, repeat_id = r)
  })
})

# Paralelizace přes foldy (multisession), aby byl výpočet importance rychlejší.
oplan <- future::plan(multisession, workers = max(1, future::availableCores() - 1))
on.exit(future::plan(oplan), add = TRUE)

imp_long <- purrr::map_dfr(folds$splits, one_fold)

# Souhrn: průměr a variabilita delta RMSE; navíc orientační t-statistika/p-hodnota
# (pouze informativní, nikoliv formální test významnosti pro publikaci).
imp_sum <- imp_long |>
  group_by(var) |>
  summarise(
    mean_delta_rmse = mean(delta_rmse, na.rm = TRUE),
    sd_delta_rmse = sd(delta_rmse, na.rm = TRUE),
    n = sum(is.finite(delta_rmse)),
    t_stat = ifelse(n > 1 && sd_delta_rmse > 0, mean_delta_rmse / (sd_delta_rmse / sqrt(n)),
NA_real_),
    p_value = ifelse(n > 1 && sd_delta_rmse > 0, stats::pt(-t_stat, df = n-1), NA_real_),
    .groups = "drop"
  ) |>
  arrange(desc(mean_delta_rmse))

list(long = imp_long, summary = imp_sum)
}

# =====
# IMPURITY IMPORTANCE Z FINÁLNÍHO MODELU – POUZE DOPLŇKOVÝ REPORT
# =====
# Cíl:
# Získat interní "variable importance" z rangeru (impurity-based).
# Upozornění:
# Tento typ importance může být zkreslený (např. ve prospěch proměnných s vyšší
# variabilitou či více kategoriemi). Proto slouží pouze jako doplňkové vodítko.
extract_impurity_importance <- function(wf_fit){
  eng <- workflows::extract_fit_engine(wf_fit)
  vi <- eng$variable.importance
  tibble(
    var = names(vi),
    impurity = as.numeric(vi)
  ) |>
  arrange(desc(impurity))
}

# =====
# TUNING + FINÁLNÍ FIT (FULL) + ULOŽENÍ IMPORTANCE (REPORT)
# =====
# Postup:
# 1) Kontrola a převod prediktorů, výběr cílové proměnné (target)
# 2) Tuning hyperparametrů přes 10-fold CV (RMSE)
# 3) Fit finálního modelu na plných tréninkových datech s best parametry
# 4) Výpočet metrik na tréninku (resubstituce) + export impurity importance
tune_and_fit_full <- function(train_sf, target_col, id_label,
  keep_preds = all_preds,
  v_tune = 10, grid_size = 30, seed = 123){

# Kontrola datových typů prediktorů + případný audit do CSV

```

```

X <- check_dataset(train_sf, label = id_label, keep_preds = keep_preds, out_dir = out_dir)

# Ověření existence cílové proměnné a případný fallback (kvůli kompatibilitě názvů)
target_real <- pick_target_col(
  train_sf, target_col,
  fallbacks = c("TOC_A", "TOC_A_new", "TOC_O", "TOC_O_new", "SUM_TOC", "SUM_TOC_ne")
)

# Převod cílové proměnné na numeric (podpora desetinné čárky)
y_raw <- sf::st_drop_geometry(train_sf)[[target_real]]
y_num <- suppressWarnings(as.numeric(gsub(",", ".", as.character(y_raw))))

if (length(y_num) != nrow(X)) stop("Target délka nesedí pro ", id_label, " (", length(y_num), " vs ",
nrow(X), ").")

# Sestavení trénovací tabulky (prediktory + target), odstranění neplatných hodnot
train_df <- X
train_df[[target_real]] <- y_num
train_df <- train_df |> filter(is.finite(.data[[target_real]]))
if (nrow(train_df) < 10) stop("Málo tréninkových řádků v ", id_label, ": ", nrow(train_df))

# Definice receptu (pre-processing)
rec <- rf_recipe_for(train_df, target_real)

# Zjištění počtu prediktorů po "bake" (kvůli rozumnému rozsahu mtry)
prep_tmp <- prep(rec, training = train_df, verbose = FALSE)
x_tmp <- bake(prepare_tmp, new_data = train_df)
p <- ncol(x_tmp) - 1L

# RF specifikace pro tuning (importance vypnuta; běží po 1 vlákně kvůli paralelizaci přes resampl.)
spec_tune <- rand_forest(trees = tune(), mtry = tune(), min_n = tune()) |>
  set_mode("regression") |>
  set_engine("ranger",
    importance = "none",
    num.threads = 1,
    sample.fraction = 0.70,
    max.depth = 30,
    respect.unordered.factors = "ignore")

# Workflow = recept + model
wf_tune <- workflow() |>
  add_recipe(rec, blueprint = hardhat::default_recipe_blueprint(allow_novel_levels = TRUE)) |>
  add_model(spec_tune)

# Resampling schéma pro tuning
set.seed(seed)
folds <- vfold_cv(train_df, v = v_tune)

# Definice rozsahů laděných parametrů
param_set <- parameters(spec_tune) |>
  update(
    mtry = mtry(c(1L, max(2L, floor(0.5 * p)))),
    trees = trees(c(400L, 1200L)),
    min_n = min_n(c(2L, 20L))
  )

# Latinský hyperkrychlový výběr kombinací (efektivnější než plná mřížka)
set.seed(seed)
grid <- dials::grid_latin_hypercube(param_set, size = grid_size)

# Paralelizace pro tuning přes resamplování
oplan <- future::plan(multisession, workers = max(1, future::availableCores() - 1))
on.exit(future::plan(oplan), add = TRUE)

# Řízení tuningu: predikce se neukládají (šetří paměť), paralelizace přes resamplování
ctrl <- control_grid(save_pred = FALSE, parallel_over = "resamples", verbose = TRUE)

message("=== TUNING (FULL): ", id_label, " | preds: ", paste(keep_preds, collapse=", "), " ===")
set.seed(seed)
tuned <- tune_grid(wf_tune, resamples = folds, grid = grid,
  metrics = metric_set(rmse), control = ctrl)

# Výběr nejlepší kombinace dle RMSE
best_params <- select_best(tuned, metric = "rmse")

```

```

print(best_params)

# Finální model: importance = "impurity" pouze kvůli doplňkovému reportu
spec_final <- rand_forest(
  trees = best_params$trees,
  mtry = best_params$mtry,
  min_n = best_params$min_n
) |>
  set_mode("regression") |>
  set_engine("ranger",
    importance = "impurity",
    num.threads = max(1, future::availableCores() - 1),
    respect.unordered.factors = "ignore")

# Finální workflow a fit na celé trénovací sadě
final_wf <- workflow() |>
  add_recipe(rec, blueprint = hardhat::default_recipe_blueprint(allow_novel_levels = TRUE)) |>
  add_model(spec_final)

set.seed(seed)
final_fit <- fit(final_wf, data = train_df)

# Metriky na trénovacích datech (resubstituce) – kontrolní report
pred_fit <- predict(final_fit, new_data = train_df)$pred
mfit <- calc_metrics(train_df[[target_real]], pred_fit)

cat(sprintf("[%s] TRAIN (resub): R2=%3f | RMSE=%3f | MAE=%3f | NRMSE_mean=%3f |
NRMSE_range=%3f\n",
  id_label, mfit$R2, mfit$RMSE, mfit$MAE, mfit$NRMSE_mean, mfit$NRMSE_range))

# Doplňková impurity importance
imp_impurity <- extract_impurity_importance(final_fit)

list(
  fit = final_fit,
  train_df = train_df,
  target = target_real,
  predictors = keep_preds,
  tuned = tuned,
  best = best_params,
  metrics_fit = mfit,
  metrics_cv = collect_metrics(tuned) |> arrange(.metric, mean),
  impurity_importance = imp_impurity
)
}

# Uložení modelu a metadat pro audit a reprodukovatelnost:
# - model: uložen jako RDS (workflow)
# - metadata: JSON (target, prediktory, best parametry, počet řádků, metriky)
save_model_bundle <- function(obj, tag, id_label){
  rds_path <- file.path(models_dir, paste0("rf_", tag, "_full_workflow.rds"))
  meta_path <- file.path(models_dir, paste0("rf_", tag, "_full_meta.json"))

  saveRDS(obj$fit, rds_path)

  meta <- list(
    id_label = id_label,
    variant = "FULL",
    target_col = obj$target,
    predictors = obj$predictors,
    best_params = obj$best,
    train_rows = nrow(obj$train_df),
    metrics_fit = as.list(obj$metrics_fit)
  )
  writeLines(jsonlite::toJSON(meta, auto_unbox = TRUE, pretty = TRUE), meta_path)

  list(rds = rds_path, meta = meta_path)
}

# Predikce na grid:
# - sjednocení a převod prediktorů v gridu (stejná pravidla jako v tréninku)
# - aplikace finálního workflow (včetně receptu)
predict_on_grid <- function(final_fit, grid_sf, keep_preds){
  grid_X <- unify_predictors(grid_sf, keep = keep_preds) |> cast_types()

```

```

    as.numeric(suppressWarnings(predict(final_fit, new_data = grid_X))$.pred)
  }

# =====
# BĚH PRO JEDEN CÍL (FULL): TUNING+FIT → IMPORTANCE (REPORT) → ULOŽENÍ
# =====
# Funkce provede celý proces pro jednu cílovou proměnnou:
# - tuning + finální fit
# - výpočet CV-perm importance (robustnější) a uložení do CSV
# - extrakce impurity importance (doplňkové) a uložení do CSV
# - uložení modelu a metadat (RDS + JSON)
run_one_full <- function(train_sf, target_col, tag, id_label,
  keep_preds_full = all_preds,
  v_tune = 10, grid_size = 30,
  imp_v = 10, imp_repeats = 3, seed = 123){

  # Tuning a finální fit pro FULL model
  full <- tune_and_fit_full(
    train_sf, target_col, paste0(id_label, " (FULL)"),
    keep_preds = keep_preds_full,
    v_tune = v_tune, grid_size = grid_size, seed = seed
  )

  # CV permutační importance (férovější) – export do CSV
  imp_cv <- cv_permutation_importance(
    train_df = full$train_df, target_col = full$target, final_params = full$best,
    v = imp_v, repeats = imp_repeats, seed = seed
  )
  imp_cv_csv <- file.path(out_dir, paste0("rf_importance_cvperm_", tag, "_FULL.csv"))
  write_csv(imp_cv$summary, imp_cv_csv)

  # Impurity importance (doplňkové) – export do CSV
  imp_imp_csv <- file.path(out_dir, paste0("rf_importance_impurity_", tag, "_FULL.csv"))
  write_csv(full$impurity_importance, imp_imp_csv)

  # Uložení modelu + metadat
  saved <- save_model_bundle(full, tag, id_label)

  list(
    full = full,
    imp_cv = imp_cv$summary,
    files = list(
      imp_cv_csv = imp_cv_csv,
      imp_impurity_csv = imp_imp_csv,
      model_rds = saved$rds,
      meta_json = saved$meta
    )
  )
}

# =====
# HLAVNÍ BĚH: NAČTENÍ DAT, TRÉNINK 3 MODELŮ, PREDIKCE NA GRID
# =====

# Načtení tréninkových vrstev a gridu + sjednocení CRS (EPSG:5514)
train_A <- read_sf_5514(path_A)
train_0 <- read_sf_5514(path_0)
train_A0 <- read_sf_5514(path_A0)
grid_sf <- read_sf_5514(path_grid)

# Trénink modelu pro A-horizont (TOC_A_new)
res_A <- run_one_full(
  train_A, target_col = "TOC_A_new", tag = "A", id_label = "A-horizont",
  v_tune = 10, grid_size = 30, imp_v = 10, imp_repeats = 3
)

# Trénink modelu pro 0-horizont (TOC_0_new)
res_0 <- run_one_full(
  train_0, target_col = "TOC_0_new", tag = "0", id_label = "0-horizont",
  v_tune = 10, grid_size = 30, imp_v = 10, imp_repeats = 3
)

# Trénink modelu pro součet A+0 (SUM_TOC_ne)
res_A0 <- run_one_full(

```

```

train_AO, target_col = "SUM_TOC_ne", tag = "AO", id_label = "O+A součet",
v_tune = 10, grid_size = 30, imp_v = 10, imp_repeats = 3
)

# =====
# PREDIKCE NA GRID (FULL) – VÝSTUPNÍ ATRIBUTY PRO MAPY
# =====
# Pozn.: predikce probíhá přes workflow, takže se automaticky aplikuje stejný
# recept (imputace, ošetření kategorií atd.) jako při tréninku.
grid_sf$A_pred <- predict_on_grid(res_A$full$fit, grid_sf, res_A$full$predictors)
grid_sf$O_pred <- predict_on_grid(res_O$full$fit, grid_sf, res_O$full$predictors)
grid_sf$SUM_pred <- predict_on_grid(res_AO$full$fit, grid_sf, res_AO$full$predictors)

# =====
# EXPORT METRIK (FULL) – SOUHRNNÁ TABULKA PRO REPORT/ČLÁNEK
# =====
metrics_tbl <- bind_rows(
  res_A$full$metrics_fit |> mutate(model="TOC_A_new", variant="FULL"),
  res_O$full$metrics_fit |> mutate(model="TOC_O_new", variant="FULL"),
  res_AO$full$metrics_fit |> mutate(model="SUM_TOC_ne", variant="FULL")
) |>
  select(model, variant, everything())

write_csv(metrics_tbl, metrics_csv)

# =====
# ZÁPIS VÝSLEDNÉHO SHAPEFILE (NOVÝ NÁZEV, BEZ PŘEPŠÁNÍ PŮVODNÍCH DAT)
# =====
# - výstup se zapisuje pod novým názvem vrstvy (out_layer)
# - před zápisem se provede kontrola geometrické validity a případná oprava
# - předchozí soubory se stejným názvem se odstraní, aby nevznikl mix starých/nových částí
dir.create(out_dir, showWarnings = FALSE, recursive = TRUE)
grid_sf <- st_transform(grid_sf, 5514)
if (any(!st_is_valid(grid_sf))) grid_sf <- st_make_valid(grid_sf)

# Smazání případných zbytků starších výstupů (shp je sada souborů s různými příponami)
exts <- c(".shp", ".shx", ".dbf", ".prj", ".cpg", ".qmd")
invisible(file.remove(file.path(out_dir, paste0(out_layer, exts))))

# Zápis shapefile s UTF-8 kódováním (kvůli diakritice)
st_write(
  grid_sf,
  dsn = out_dir,
  layer = out_layer,
  driver = "ESRI Shapefile",
  layer_options = "ENCODING=UTF-8",
  delete_layer = TRUE,
  quiet = TRUE
)
writelines("UTF-8", file.path(out_dir, paste0(out_layer, ".cpg")))

# Návrat k sekvenčnímu plánu (hygiena paralelizace)
future::plan(sequential)

# =====
# ZÁVĚREČNÉ SHRNUTÍ BĚHU (KONZOLE)
# =====
cat(
  "\n=== HOTOVO (FULL ONLY) ===\n",
  "Výstupní SHP: ", out_shp,
  "\nMetriky: ", metrics_csv,
  "\nModely: ", models_dir,
  "\nImportance CSV:\n rf_importance_cvperm_[A|O|AO]_FULL.csv (permutační; robustnější)\n",
  rf_importance_impurity_[A|O|AO]_FULL.csv (impurity; doplňkové, může být biased)\n",
  "\nPole ve SHP:\n A_pred, O_pred, SUM_pred\n",
  sep=""
)

```

Příloha 1.4 Python skript pro nové proporční rozdělení agregovaných porostních typů (APT) do příslušných CHS v rámci PLO při uplatnění „listnatého“ scénáře.

```
# =====
# ArcGIS Pro / Python Window
# Proporční přerozdělení APT (výměry -> podíly) se směrem k listnatému
# VARIANTA 3 + LOCK CHS: u CHS {1,2,3,13,19} se APT NEMĚNÍ
# -----
# Účel:
# Podle scénáře (Excel, list "nova_procenta") přepočítat dřevinnou skladbu (APT)
# v polygonové vrstvě tak, aby v každé kombinaci (LES_OBL[_1], CHS) odpovídaly
# dosažené podíly cílovým podílům ze scénáře. Přerozdělení je jednosměrné -
# polygon může přejít jen „k listnatějšímu“ typu v pevné posloupnosti.
#
# Vstup:
# - Shapefile s poli: APT (kategorie), CHS (kód), LES_OBL_1 (nebo LES_OBL),
# geometrie polygonů (plocha se bere z SHAPE@AREA).
# - Excel (scenar.xlsx), list "nova_procenta" obsahuje VÝMĚRY APT
# pro kombinace (LES_OBL[_1], CHS). Jednotky (ha/m²) nejsou podstatné,
# skript vždy převádí výměry na podíly v rámci dané dvojice klíčů.
#
# Výstup:
# - Kopie vstupního SHP s názvem ..._var3_LOCK.shp.
# - Pole APT_OLD = snapshot původní APT, pole APT = nové (po přerozdělení).
# - Log s diagnostikou: počet zamčených skupin/prvků, počet změněných prvků,
# průměrná absolutní odchylka podílů (jen pro nezamčené CHS).
#
# Klíčové kroky:
# 1) Načtení a kontrola vstupů, vytvoření kopie dat, záloha APT do APT_OLD.
# 2) Načtení scénáře, převod výměr APT na podíly v rámci (LES,CHS).
# 3) Seskupení polygonů podle (LES,CHS). V každé skupině:
# - Cílová výměra APT = podíl_A * (součet ploch polygonů ve skupině).
# - Greedy alokace polygonů od největších k menším: každý polygon se
# přiřadí té „povolené“ kategorii s nejvyšší zbývajícím nenaplněnou
# cílovou plochou. „Povolené“ = kategorie s indexem >= aktuální kategorie.
# 4) Zápis výsledků do shapefileu a diagnostické metriky.
#
# Omezení / poznámky:
# - U CHS ∈ {1,2,3,13,19} se APT nikdy nemění (na úrovni skupiny i řádku).
# - Názvy APT ve scénáři musí přesně odpovídat hodnotám APT v datech
# (bez automatického přejmenování). Nepřítomné sloupce se ignorují.
# - Greedy přístup minimalizuje odchylku k cíli heuristicky; není to globálně
# optimální algoritmus. Pro účely scénářů je však rychlý, stabilní a interpretovatelný.
# =====

import arcpy, os, re, unicodedata
import pandas as pd
from collections import defaultdict

# ---- Cesty k souborům (přizpůsob dle potřeby) -----
base_dir = r"D:/NLI/final"
in_shp = os.path.join(base_dir, "SPT_CIHO_PLO.shp")
excel_xls = os.path.join(base_dir, "scenar.xlsx")
sheet_new = "nova_procenta"
out_shp = os.path.join(base_dir, "SPT_CIHO_PLO_APT_scenar_var3_LOCK.shp")

# ---- „Zámek“ CHS: kombinace, u kterých APT zůstává beze změny -----
# Pozn.: CHS se v kódu normalizuje na text (např. 1 -> "1"), proto zde řetězce.
LOCK_CHS = {"1", "2", "3", "13", "19"}

# ---- Pomocné funkce -----
def norm_txt(s):
    """Lehká normalizace textu (lowercase, bez diakritiky, bez speciálů).
    Používá se pro spolehlivé párování klíčů (LES, CHS) napříč zdroji."""
    if s is None: return ""
    s = str(s).strip().lower()
    s = unicodedata.normalize('NFKD', s)
    s = "".join(ch for ch in s if not unicodedata.combining(ch))
    s = re.sub(r"^[^a-z0-9 ]+", "", s)
    s = re.sub(r"\s+", " ", s).strip()
    return s
```

```

def pick_col(cols, candidates):
    """Vybere první existující název sloupce z dodaného seznamu kandidátů
    (case-insensitive). Umožňuje podporovat LES_OBL_1 i LES_OBL."""
    low = {c.lower(): c for c in cols}
    for cand in candidates:
        c = low.get(cand.lower())
        if c: return c
    return None

def msg(*a):
    """Zapíše zprávu do konzole a zároveň (pokud je to možné) do geoprocessing logu."""
    text = " ".join(str(x) for x in a)
    print(text)
    try:
        arcpy.AddMessage(text)
    except Exception:
        pass

# ---- 1) Kontroly vstupů a vytvoření kopie -----
if not arcpy.Exists(in_shp):
    raise FileNotFoundError(f"Vstupní shapefile nenalezen: {in_shp}")
if not os.path.exists(excel_xls):
    raise FileNotFoundError(f"Excel nenalezen: {excel_xls}")

fields = {f.name for f in arcpy.ListFields(in_shp)}
if "APT" not in fields: raise ValueError("V shapefilu chybí pole 'APT'.")
if "CHS" not in fields: raise ValueError("V shapefilu chybí pole 'CHS'.")

# Preferuj LES_OBL_1, jinak použij LES_OBL
LES_F = "LES_OBL_1" if "LES_OBL_1" in fields else ("LES_OBL" if "LES_OBL" in fields else None)
if not LES_F:
    raise ValueError("V shapefilu chybí 'LES_OBL_1' i 'LES_OBL'.")

# Bezpečná práce s kopií (originál se NEMĚNÍ)
if arcpy.Exists(out_shp):
    arcpy.management.Delete(out_shp)
arcpy.management.CopyFeatures(in_shp, out_shp)
msg("Vytvořena kopie:", out_shp)

# Snapshot původní APT (pro audit a možnost návratu)
out_fields = {f.name for f in arcpy.ListFields(out_shp)}
if "APT_OLD" not in out_fields:
    arcpy.management.AddField(out_shp, "APT_OLD", "TEXT", field_length=50)
arcpy.management.CalculateField(out_shp, "APT_OLD", "!APT!", "PYTHON3")
msg("Pole APT_OLD naplněno původní hodnotou APT.")

# ---- 2) Kategorie APT dostupné v gridu -----
# Budeme z nich skládat cíle; názvy ve scénáři musí těmto hodnotám odpovídat.
apt_set = set()
with arcpy.da.SearchCursor(out_shp, ["APT"]) as cur:
    for (apt,) in cur:
        if apt is not None and str(apt).strip():
            apt_set.add(str(apt))
if not apt_set:
    raise ValueError("V poli APT nejsou žádné neprázdné hodnoty.")
msg("APT kategorie (z gridu):", ", ".join(sorted(apt_set)))

# ---- 3) Načtení scénáře a převod výměr -> podíly -----
# V listu "nova_procenta" jsou výměry APT pro (LES,CHS); konkrétní jednotka je jedno.
df = pd.read_excel(excel_xls, sheet_name=sheet_new)

# Detekce názvů sloupců pro klíče
les_col = pick_col(df.columns, ["LES_OBL_1", "LES_OBL"])
chs_col = pick_col(df.columns, ["CHS"])
if not les_col or not chs_col:
    raise ValueError("V Excelu nejsou sloupce 'LES_OBL_1/LES_OBL' a/nebo 'CHS'.")

# APT sloupce: vezmeme jen ty, které přesně odpovídají hodnotám APT v gridu
apt_cols = [c for c in df.columns if c not in (les_col, chs_col)]
exact_apt_cols = [c for c in apt_cols if str(c) in apt_set]
ignored_cols = [c for c in apt_cols if str(c) not in apt_set]
if not exact_apt_cols:
    raise ValueError(

```

```

        f"V Excelu nebyly nalezeny žádné APT sloupce shodné s gridem.\nGrid: {sorted(apt_set)}\nExcel:
{apt_cols}"
    )
    if ignored_cols:
        msg("Ignorovány APT sloupce (nejsou v gridu):", ", ".join(map(str, ignored_cols)))

# Číselná interpretace výměr (NaN -> 0)
for c in exact_apt_cols:
    df[c] = pd.to_numeric(df[c], errors="coerce").fillna(0.0)

# Normalizované klíče pro robustní párování
df["_LES_N"] = df[les_col].map(norm_txt)
df["_CHS_N"] = df[chs_col].map(norm_txt)

# Agregace (kdyby v Excelu byly duplicity) a převod na podíly v rámci (LES,CHS)
agg = df.groupby(["_LES_N", "_CHS_N"], as_index=False)[exact_apt_cols].sum()
shares = {}
for _, r in agg.iterrows():
    vals = {apt: float(r[apt]) for apt in exact_apt_cols}
    total = sum(vals.values())
    if total > 0:
        # Podíl APT_i = vymera_i / suma_vymer
        shares[(r["_LES_N"], r["_CHS_N"])] = {apt: vals[apt] / total for apt in exact_apt_cols}
msg("Načteno (LES,CHS) kombinací ve scénáři:", len(shares))

# ---- 4) Příprava gridových dat (plochy, klíče) -----
oid_f = arcpy.Describe(out_shp).OIDFieldName
rows = []
with arcpy.da.SearchCursor(out_shp, [oid_f, LES_F, "CHS", "APT", "SHAPE@AREA"]) as cur:
    for oid, les, chs, apt, area in cur:
        rows.append({
            "oid": oid,
            "les_n": norm_txt(les),
            "chs_n": norm_txt(chs), # Pozor: CHS se porovnává jako text (např. "1")
            "apt": str(apt),
            "area": float(area or 0.0)
        })

# Ekologická posloupnost APT (zleva jehličnatější → vpravo listnatější)
apt_order = ["jehličnatý", "smíšený jehličnatý", "smíšený", "smíšený listnatý", "listnatý"]
apt_to_idx = {a: i for i, a in enumerate(apt_order)}

# ---- 5) Greedy přerozdělení (jen „k listnatému“) + zámeček CHS -----
# Seskupení polygonů podle klíčů; zamčené CHS se přeskočí už zde.
by_group = defaultdict(list)
for r in rows:
    by_group[(r["les_n"], r["chs_n"])].append(r)

changes = {}
locked_groups = 0

for key, feats in by_group.items():
    les_n, chs_n = key

    # ZÁMEK: celé skupiny s CHS v LOCK_CHS se NEMĚNÍ
    if chs_n in LOCK_CHS:
        locked_groups += 1
        continue

    # Pokud pro skupinu není ve scénáři cíl, nic neměníme
    if key not in shares:
        continue

    # Cílové podíly a z nich odvozené cílové VÝMĚRY (v m²) v rámci skupiny
    target_share = shares[key] # {apt: share}
    total_area = sum(f["area"] for f in feats)
    target_area = {apt: target_share.get(apt, 0) * total_area for apt in apt_set}
    remain = dict(target_area) # zbývající cílová výměra

    # Stabilita: největší polygony alokujeme dříve (větší vliv na bilanci)
    feats_sorted = sorted(feats, key=lambda f: -f["area"])

    for f in feats_sorted:
        cur_apt = f["apt"]

```

```

cur_idx = apt_to_idx.get(cur_apt, None)

# Dovolené cíle: jen „k listnatějším“ (index >= aktuální); neznámé -> všechny
candidates = apt_order[cur_idx:] if cur_idx is not None else apt_order[:]

# Greedy krok: vyber APT s nejvyšším zbývajícím cílem (remain)
best = max(candidates, key=lambda a: remain.get(a, 0.0))
if remain.get(best, 0.0) > 0:
    changes[f["oid"]] = best
    remain[best] -= f["area"] # odečti plochu právě alokovaného polygonu

# ---- 6) Zápis změn (s druhou pojistkou zámku na úrovni řádku) -----
n_total = n_changed = n_locked_skipped = 0
with arcpy.da.UpdateCursor(out_shp, [oid_f, LES_F, "CHS", "APT"]) as cur:
    for oid, les, chs, apt in cur:
        n_total += 1
        chs_n = norm_txt(chs)

        # Druhá pojistka: pokud je prvek v zamčeném CHS, prostě přeskočit
        if chs_n in LOCK_CHS:
            n_locked_skipped += 1
            continue

        # Zapiš změnu, pokud byla pro OID navržena a skutečně mění kategorii
        new_val = changes.get(oid)
        if new_val and new_val != apt:
            cur.updateRow((oid, les, chs, new_val))
            n_changed += 1

msg(f"Zamčených CHS skupin (bez změn): {locked_groups} (CHS in {sorted(LOCK_CHS)})")
msg(f"Zamčených prvků přeskočeno: {n_locked_skipped}")
msg(f"Změněno APT u {n_changed} z {n_total} prvků.")

# ---- 7) Orientační kontrola kvality (jen pro NEZAMKNUTÉ CHS) -----
# Po akci znovu sečteme dosažené podíly a spočítáme průměrnou absolutní odchylku
# vůči cíli (0-1; čím blíže 0, tím lépe).
ach = defaultdict(lambda: defaultdict(float))
with arcpy.da.SearchCursor(out_shp, [LES_F, "CHS", "APT", "SHAPE@AREA"]) as cur:
    for les, chs, apt, area in cur:
        chs_n = norm_txt(chs)
        if chs_n in LOCK_CHS:
            continue
        ach[(norm_txt(les), chs_n)][str(apt)] += float(area or 0.0)

sum_abs = count = 0
for (les_n, chs_n), tgt in shares.items():
    if chs_n in LOCK_CHS: # vynech zamčené
        continue
    if (les_n, chs_n) not in ach:
        continue
    tot = sum(ach[(les_n, chs_n)].values()) or 1.0
    got = {apt: ach[(les_n, chs_n)].get(apt, 0.0) / tot for apt in apt_set}
    for a in apt_set:
        sum_abs += abs(got.get(a, 0) - tgt.get(a, 0))
        count += 1

if count:
    msg(f"Průměrná absolutní odchylka podílů (0-1) pro NEZAMKNUTÉ CHS: {sum_abs/count:.4f}")
else:
    msg("Pro NEZAMKNUTÉ CHS nebyl spočten žádný pár (LES,CHS) pro porovnání.")

msg("HOTOVO. Výstup:", out_shp)

```

Příloha 1.5 Skript pro nasazení uloženého modelu Random Forest na scénářový grid

```
# =====  
# APPLY SAVED RF FULL MODELS (tidymodels workflows) ON NEW GRID  
# =====  
#  
# Účel skriptu:  
# Aplikace již natrénovaných plných Random Forest modelů  
# (uložených jako tidymodels workflows) na nový predikční grid.  
#  
# Kontext:  
# Modely byly natrénovány zvlášť pro:  
# - horizont A (minerální vrstva),  
# - horizont O (organická vrstva),  
# - součet A + O.  
#  
# Tento skript neprovádí trénink modelů, pouze:  
# - načte uložené modely,  
# - připraví prediktory ve stejném formátu jako při tréninku,  
# - provede prostorovou predikci,  
# - zapíše výsledky do nové shapefile vrstvy.  
#  
# Výstupy:  
# - polygonová vrstva s novými atributy:  
#   A_pred ... predikce TOC v horizontu A  
#   O_pred ... predikce TOC v horizontu O  
#   SUM_pred ... predikce TOC (A + O)  
#  
# =====
```

```
suppressPackageStartupMessages({  
  # Práce s vektorovými daty a atributy  
  library(sf)  
  library(dplyr)  
  
  # Tidymodels - nutné pro predikci z uložených workflow objektů  
  library(tidymodels)  
  
  # Pomocné balíčky  
  library(readr)  
})
```

```
# =====  
# CESTY K DATŮM  
# =====  
#  
# models_dir:  
# Adresář s uloženými plnými (FULL) RF modely ve formátu .rds  
#  
# grid_in:  
# Vstupní polygonová vrstva (predikční grid),  
# obsahuje pouze prediktory - bez měřených hodnot TOC.  
#  
  
models_dir <- "D:/NLI/final/models_rf_full"  
grid_in <- "D:/NLI/final/grid_TOC_scenar_leden.shp"  
  
# Výstupní shapefile  
out_dir <- "D:/NLI/final"  
out_layer <- "grid_TOC_scenar_leden_pred_FULL"  
out_shp <- file.path(out_dir, paste0(out_layer, ".shp"))
```

```

# =====
# DEFINICE PREDIKTORŮ
# =====
#
# Prediktory MUSÍ přesně odpovídat těm,
# které byly použity při tréninku modelů:
#
# - numerické proměnné (odvozené z rastrů):
#   MEAN, STD, MAX
#
# - kategoriální proměnné (lesnické charakteristiky):
#   PLO_CHS ... lesní vegetační stupeň
#   APT      ... typ porostu
#
num_preds <- c("MEAN", "STD", "MAX")
cat_preds <- c("PLO_CHS", "APT")
all_preds <- c(num_preds, cat_preds)

# =====
# CESTY K MODELŮM
# =====
#
# Každý model je uložen jako tidymodels workflow,
# který obsahuje:
# - předzpracování dat (recept),
# - vlastní Random Forest model.
#
mA_path <- file.path(models_dir, "rf_A_full_workflow.rds")
mO_path <- file.path(models_dir, "rf_O_full_workflow.rds")
mAQ_path <- file.path(models_dir, "rf_AO_full_workflow.rds")

# Kontrola existence vstupních souborů
stopifnot(file.exists(mA_path), file.exists(mO_path), file.exists(mAQ_path))
stopifnot(file.exists(grid_in))

# =====
# POMOCNÉ FUNKCE
# =====

# -----
# Načtení vrstvy a kontrola souřadnicového systému
# -----
#
# Zajišťuje:
# - korektní přiřazení S-JTSK / Křovák (EPSG:5514),
# - případnou reprojekci vstupních dat.
#
read_sf_5514 <- function(path){
  s <- st_read(path, quiet = TRUE)
  if (is.na(st_crs(s))) s <- st_set_crs(s, 5514)
  if (!identical(st_crs(s)$epsg, 5514L)) s <- st_transform(s, 5514)
  s
}

# -----
# Ošetření listových sloupců
# -----
#
# Některé shapefile atributy mohou být načteny jako listy
# (např. po geoprocessingu v ArcGIS).

```

```

# Funkce zajistí převod na jednoduché atomické hodnoty.
#

squash_list_cols <- function(df){
  for(nm in names(df)){
    if (is.list(df[[nm]])) {
      df[[nm]] <- vapply(df[[nm]], function(z){
        if (length(z)==0 || is.null(z)) NA_character_
        else as.character(z[[1]])
      }, FUN.VALUE = character(1))
    }
  }
  df
}

# -----
# Výběr a sjednocení prediktorů
# -----
#
# - odstraní geometrii,
# - ponechá pouze prediktory použité při tréninku,
# - sjednotí názvy polí (pokud se liší mezi verzemi dat).
#

unify_predictors <- function(sfobj, keep = all_preds){
  df <- sf::st_drop_geometry(sfobj) |>
  as.data.frame(stringsAsFactors = FALSE)

  # Historická kompatibilita názvů (pokud by se vyskytly)
  if (!("PLO" %in% names(df)) && ("PLO_1" %in% names(df)))
    df$PLO <- df$PLO_1

  keep <- intersect(names(df), keep)
  df <- df[, keep, drop = FALSE]

  squash_list_cols(df)
}

# -----
# Přetypování proměnných
# -----
#
# - numerické proměnné → numeric
# - kategoriální proměnné → factor
#
# Nutné pro bezchybné použití uložených workflow modelů.
#

cast_types <- function(df, num_cols = num_preds, cat_cols = cat_preds){
  df <- squash_list_cols(df)

  for (nm in intersect(num_cols, names(df))) {
    df[[nm]] <- suppressWarnings(
      as.numeric(gsub(",", ".", as.character(df[[nm]])))
    )
  }

  for (nm in intersect(cat_cols, names(df))) {
    df[[nm]] <- as.factor(as.character(df[[nm]]))
  }
}

```

```

    df
  }

# -----
# Vlastní predikce na gridu
# -----
#
# - připraví prediktory,
# - aplikuje uložený workflow,
# - vrátí numerický vektor predikcí.
#

predict_on_grid <- function(wf_fit, grid_sf, keep_preds){
  grid_X <- unify_predictors(grid_sf, keep = keep_preds) |>
    cast_types()

  as.numeric(
    suppressWarnings(
      predict(wf_fit, new_data = grid_X)$`.pred`
    )
  )
}

# =====
# NAČTENÍ MODELŮ A PREDIKCE
# =====

mod_A <- readRDS(mA_path)
mod_O <- readRDS(mO_path)
mod_AO <- readRDS(mAO_path)

grid_sf <- read_sf_5514(grid_in)

# Výpočet predikcí
grid_sf$A_pred <- predict_on_grid(mod_A, grid_sf, all_preds)
grid_sf$O_pred <- predict_on_grid(mod_O, grid_sf, all_preds)
grid_sf$SUM_pred <- predict_on_grid(mod_AO, grid_sf, all_preds)

# =====
# ZÁPIS VÝSTUPU
# =====

# Vytvoření výstupního adresáře
dir.create(out_dir, showWarnings = FALSE, recursive = TRUE)

# Kontrola geometrické validity
grid_sf <- st_transform(grid_sf, 5514)
if (any(!st_is_valid(grid_sf)))
  grid_sf <- st_make_valid(grid_sf)

# Odstranění případných starých souborů shapefile
exts <- c(".shp", ".shx", ".dbf", ".prj", ".cpg", ".qmd")
invisible(file.remove(file.path(out_dir, paste0(out_layer, exts))))

# Zápis nové vrstvy
st_write(
  grid_sf,
  dsn = out_dir,
  layer = out_layer,
  driver = "ESRI Shapefile",
  layer_options = "ENCODING=UTF-8",

```

```
delete_layer = TRUE,
quiet = TRUE
)

# Explicitní nastavení kódování
writeLines("UTF-8", file.path(out_dir, paste0(out_layer, ".cpg")))

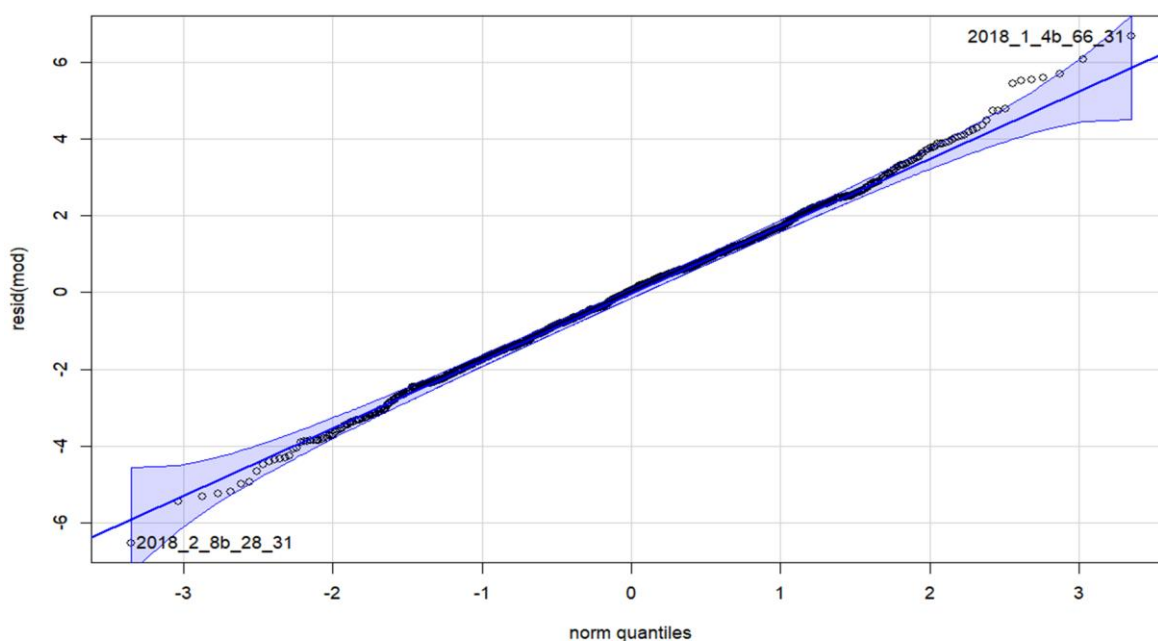
# Informativní výpis
cat(
  "\n=== HOTOVO ===\n",
  "Input grid:  ", grid_in,
  "\nOutput SHP:  ", out_shp,
  "\nPred fields: A_pred, O_pred, SUM_pred\n",
  sep = ""
)
```

Příloha 2. Doplnkové tabelární a grafické výstupy z regresní analýzy

Tab. 7. (P1): Střední hodnoty zásob uhlíku [t ha^{-1}] v organických vrstvách (O) dle agregovaného porostního typu (APT) ($P < 0,001$; $F = 17,5$; $R^2_m = 0,050$) s náhodným interceptem pro PLO ve smíšeném regresním modelu ($R^2_c = 0,313$; $AIC = 5140$).

APT	Odhad	(SE)	df	Interval spolehlivosti	
				2,5%	97,5%
Listnatý	11,9 ^a	(1,3)	78	8,5	15,4
Smíšený listnatý	17,5 ^{abc}	(2,4)	281	11,2	23,8
Smíšený	18,5 ^{bc}	(1,9)	13	13,5	23,6
Smíšený jehličnatý	17,8 ^b	(1,7)	13	13,2	22,4
Jehličnatý	22,2 ^c	(1,8)	47	17,4	27,0

^{a,b,c} Rozdílná písmena značí statistické rozdíly mezi APT (Šidák t-test $P < 0,05$); společná písmena značí neprůkaznost nebo absenci rozdílu.

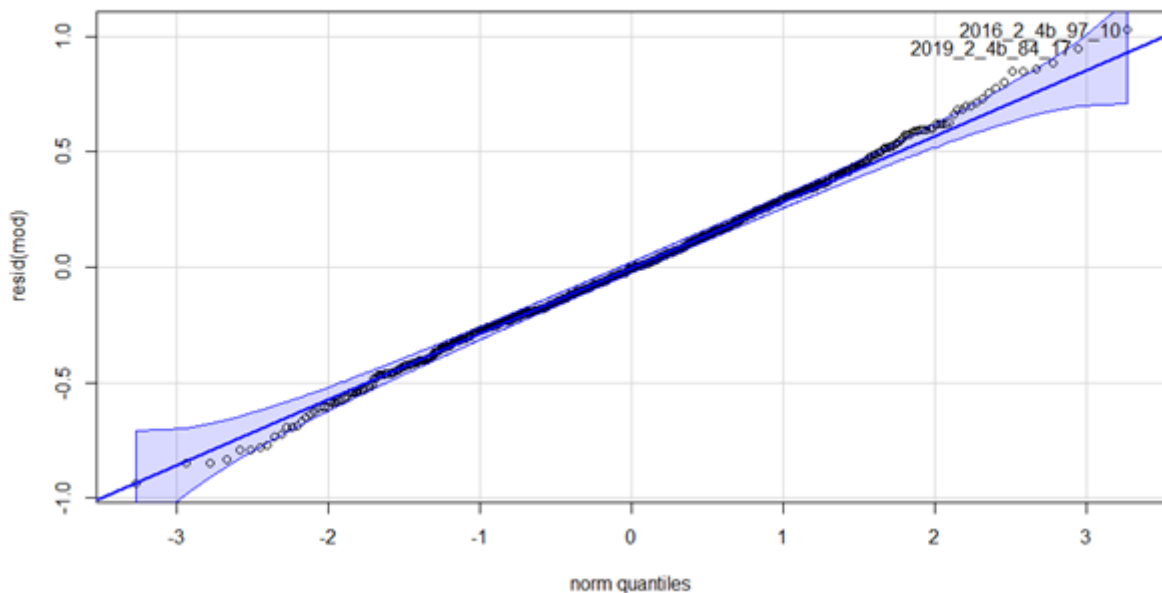


Obr. 3. (P1): QQ diagram residuí smíšeného regresního modelu zásob uhlíku v organické vrstvě O s agregovaným porostním typem jako fixní proměnnou a PLO a CHS jako náhodnými proměnnými. Na horizontální ose jsou teoretické kvantily, na vertikální ose jsou residua modelu ($n = 1.120$); modrý pás značí 95% interval spolehlivosti normálního rozdělení (Shapiro-Wilk test $P = 0,489$).

Tab. 8. (P2): Střední hodnoty zásob uhlíku [t ha⁻¹] ve **svrchní minerální vrstvě (A)** dle agregovaného porostního typu (APT) ($P < 0,001$; $F = 4,77$; $R^2_m = 0,021$) s náhodným interceptem pro CHS ve smíšeném regresním modelu ($R^2_c = 0,211$; $AIC = 526$).

APT	Odhad		(SE)	df	Interval spolehlivosti	
					2,5%	97,5%
Listnatý	29,5	b	(3,1)	37	21,1	37,8
Smíšený listnatý	24,7	ab	(3,5)	122	15,5	33,9
Smíšený	24,8	ab	(2,8)	50	17,4	32,2
Smíšený jehličnatý	22,9	ab	(2,4)	39	16,4	29,5
Jehličnatý	21,3	a	(2,0)	23	15,8	26,9

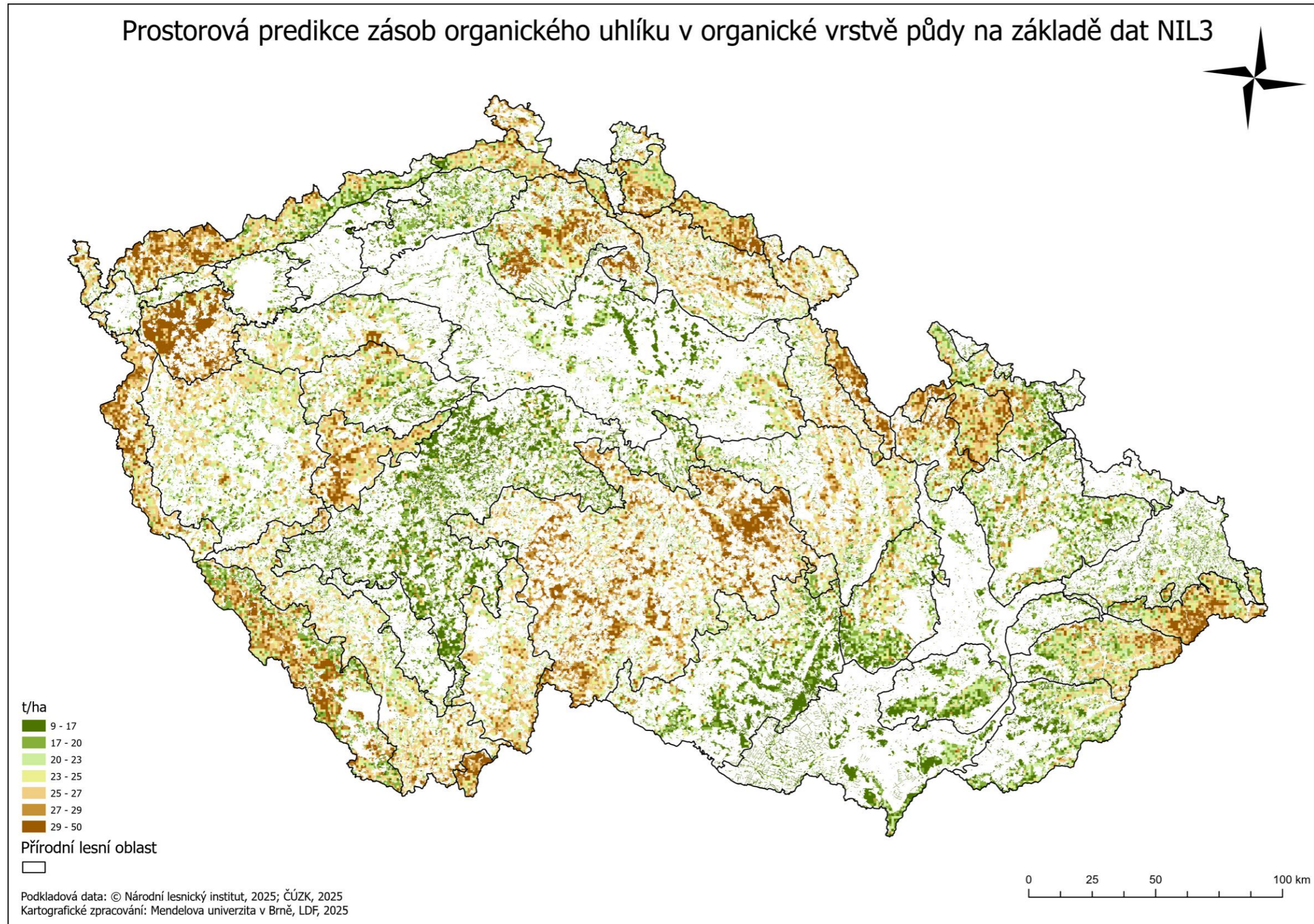
^{a,b} Rozdílná písmena značí statistické rozdíly mezi APT (Šidák t-test $P < 0,05$); společná písmena značí neprůkaznost nebo absenci rozdílu.

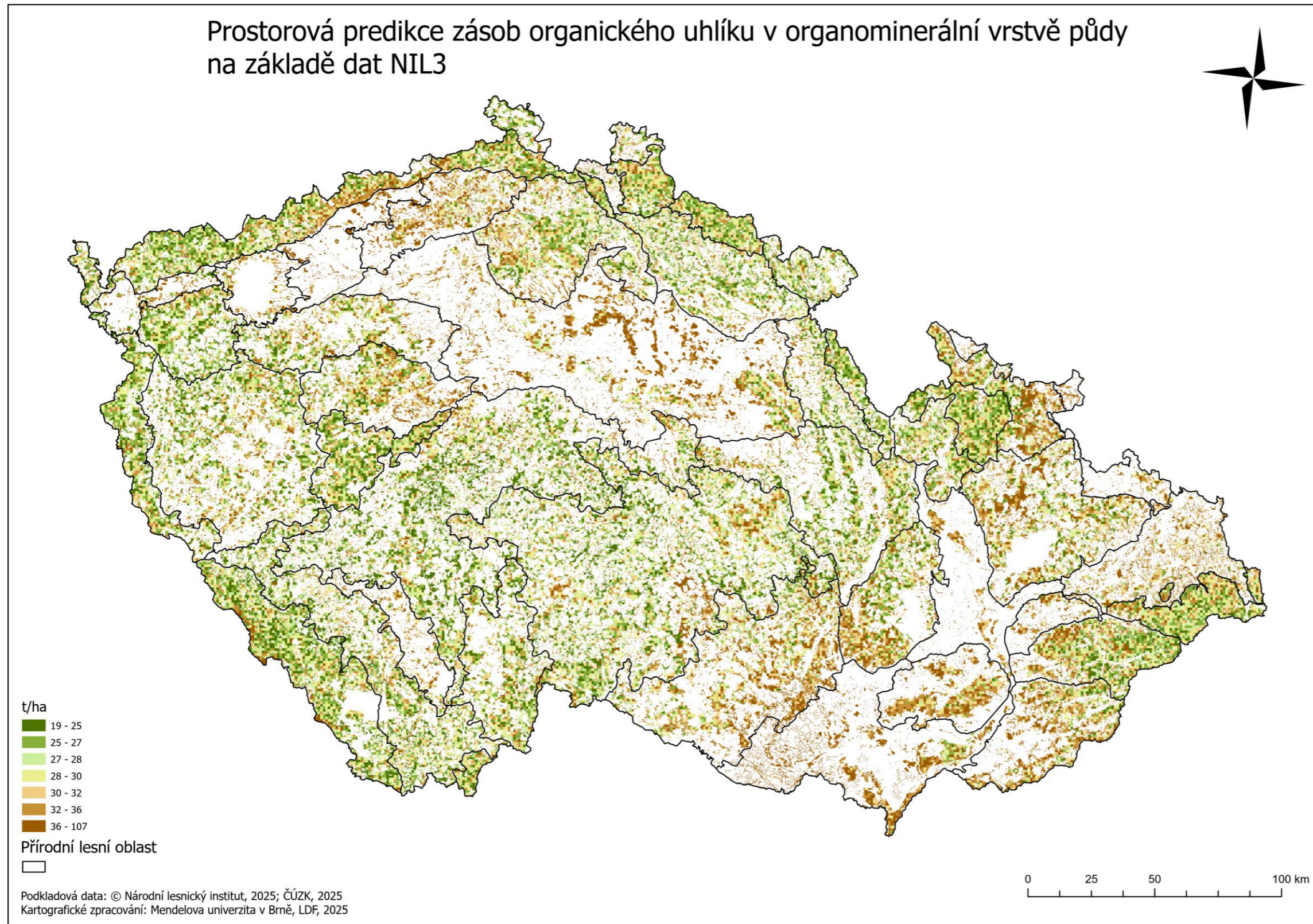


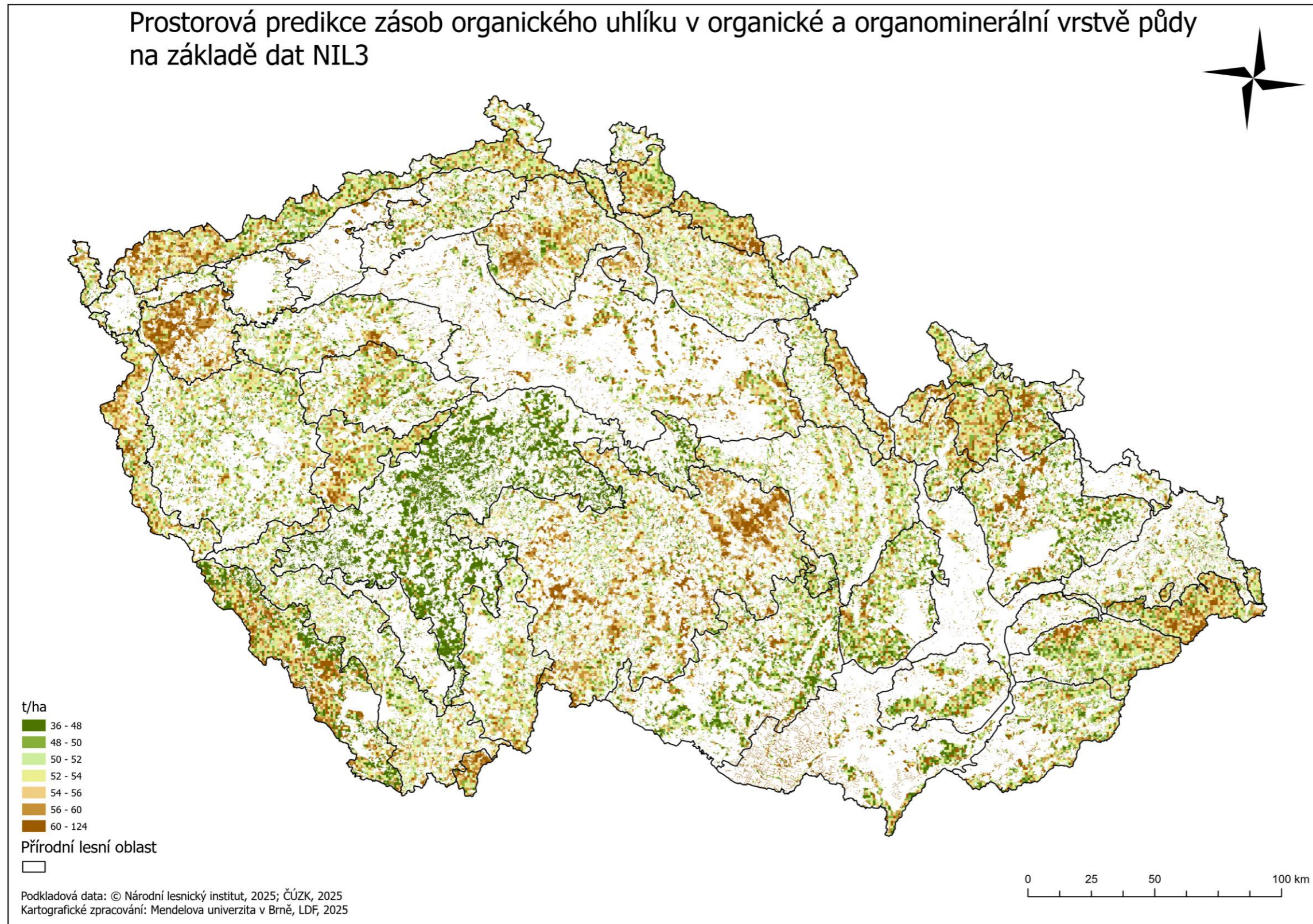
Obr. 4. (P2): QQ diagram residuí smíšeného modelu zásob uhlíku v organominerální vrstvě A s agregovaným porostním typem jako fixní proměnnou a PLO a CHS jako náhodnými proměnnými. Na horizontální ose jsou teoretické kvantily, na vertikální ose jsou residua modelu ($n = 893$); modrý pás značí 95% interval spolehlivosti normálního rozdělení (Shapiro-Wilk test $P = 0,217$).

Příloha 3. Mapové výstupy zásoby uhlíku ve svrchních vrstvách půdy: Zásoby organického uhlíku v povrchových vrstvách lesních půd

Příloha 3.1 Aktuální zásoby organického uhlíku v organické vrstvě půdy (O)

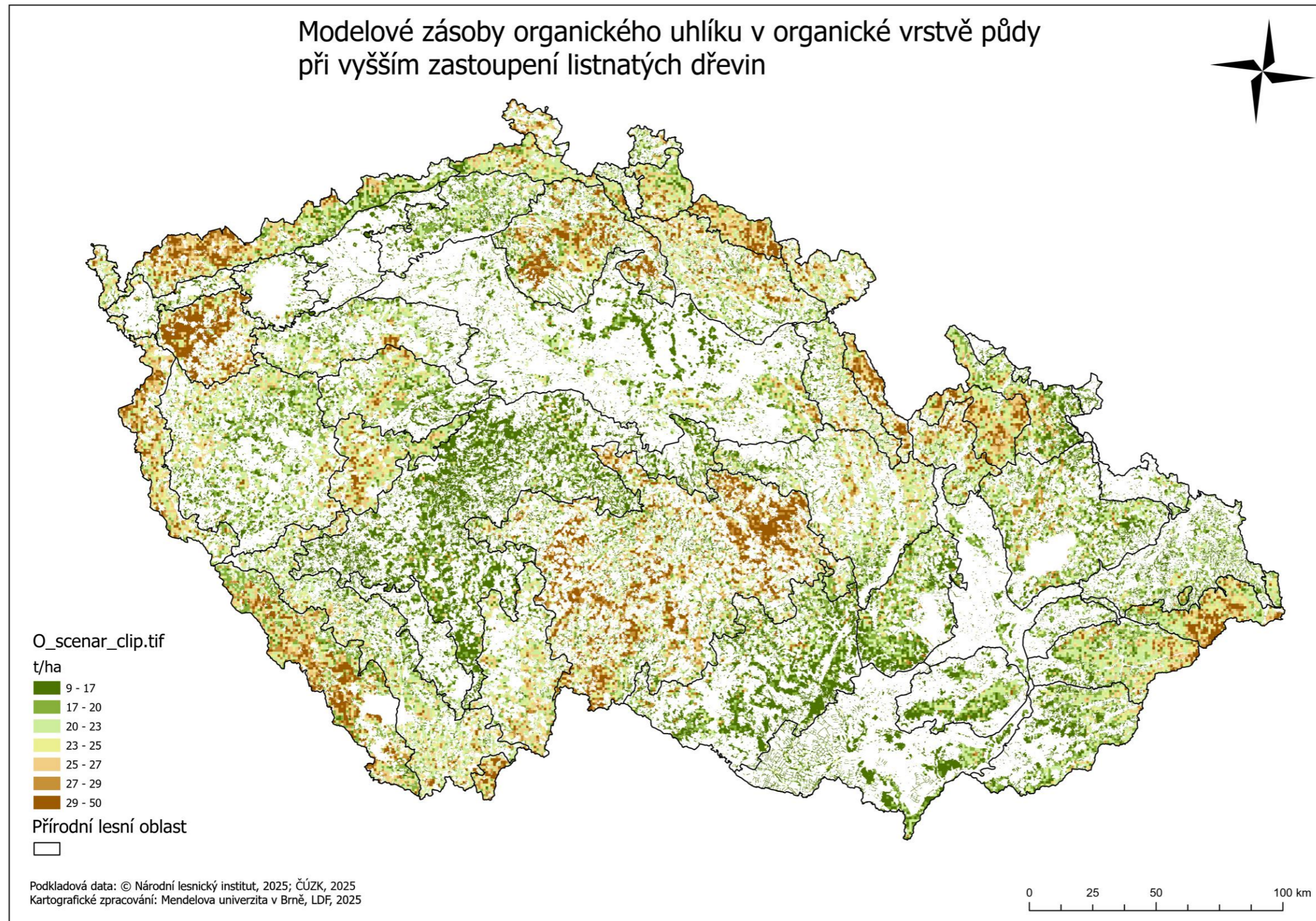




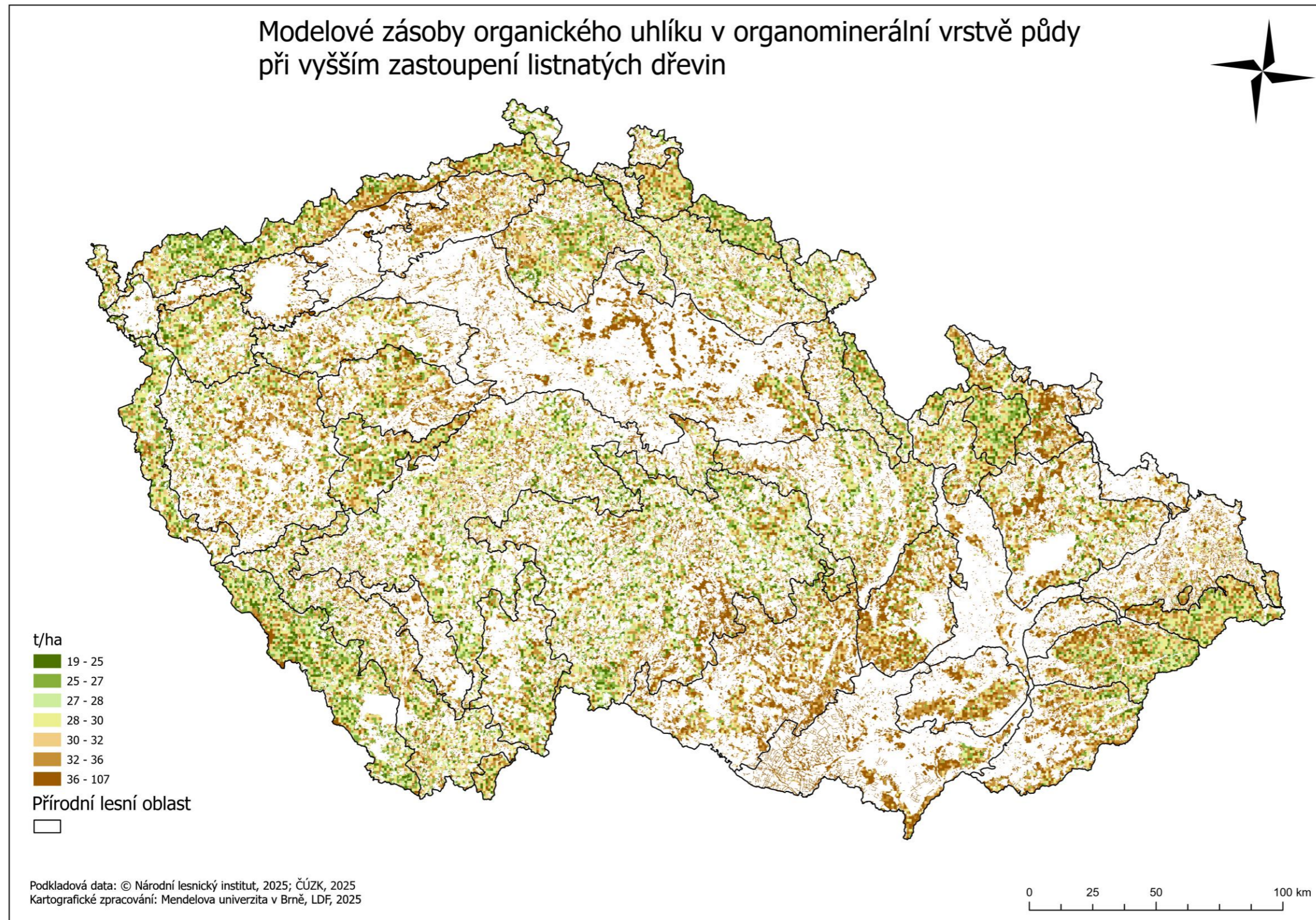


Příloha 4. Zásoby organického uhlíku v povrchových vrstvách lesních půd pro „listnatý scénář“

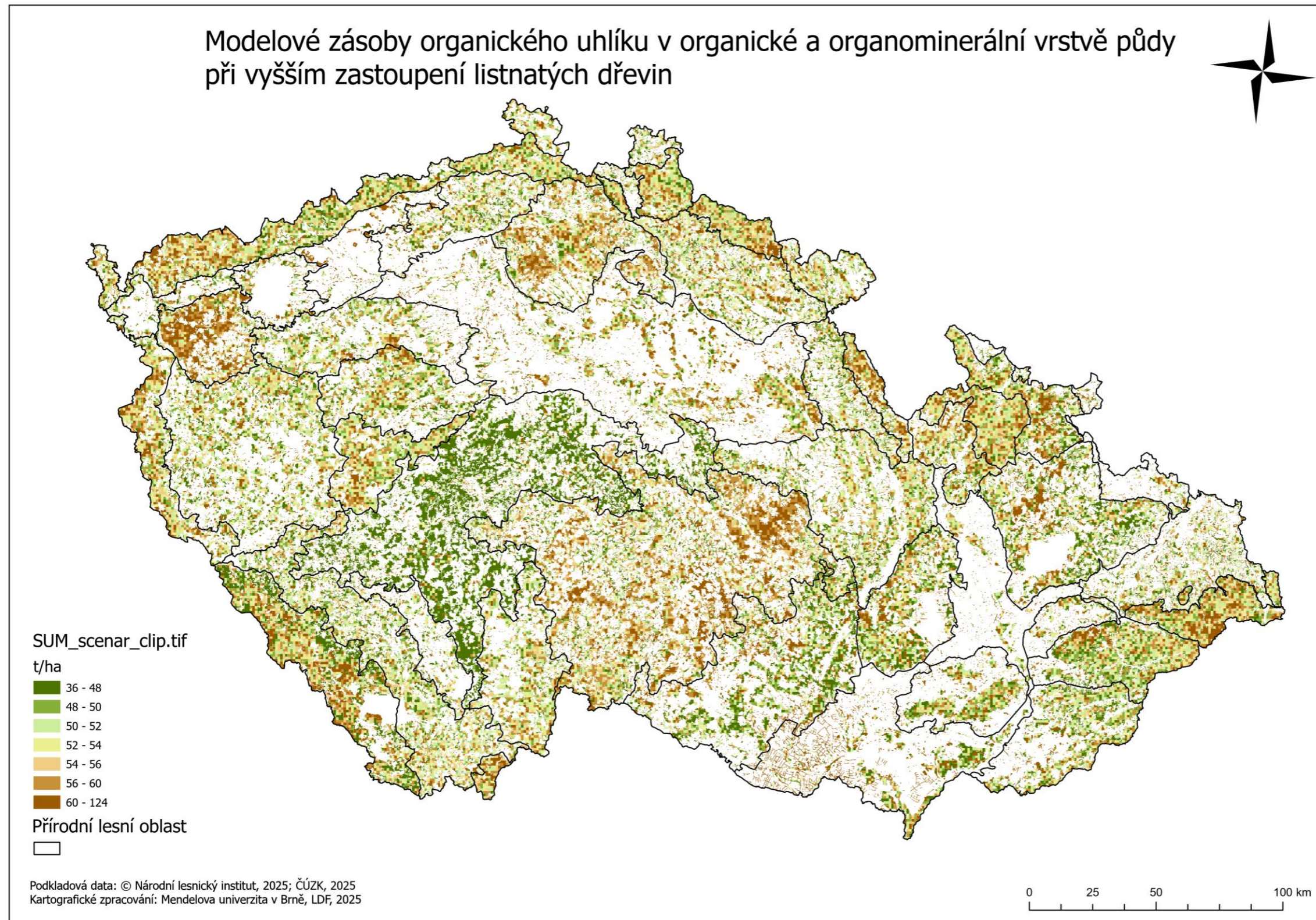
Příloha 4.1 Modelové zásoby organického uhlíku v organické vrstvě půdy (O) při vyšším zastoupení listnatých dřevin



Příloha 4.2 Modelové zásoby organického uhlíku v organominerální vrstvě půdy (A) při vyšším zastoupení listnatých dřevin

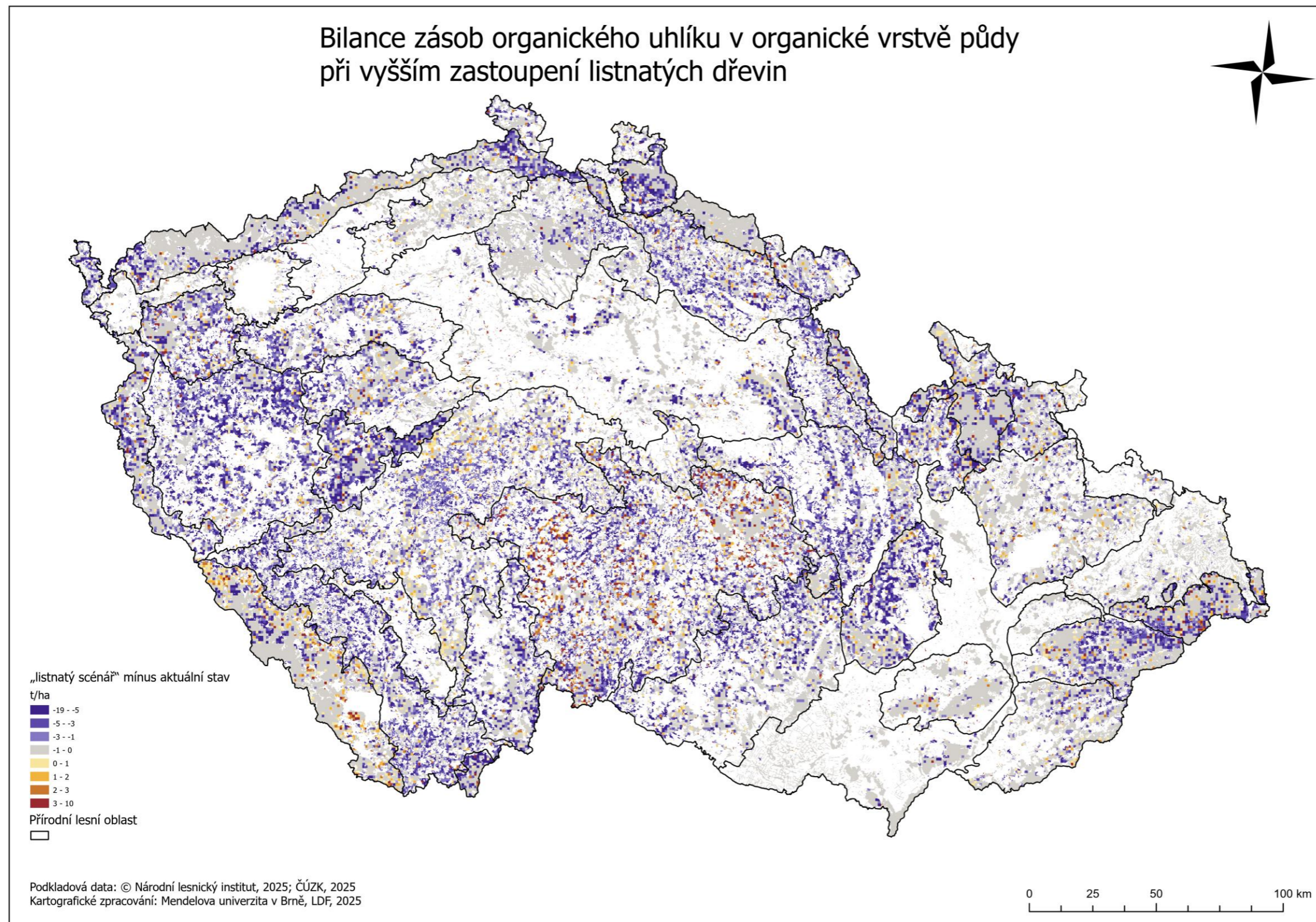


Příloha 4.3 Modelové zásoby organického uhlíku v organické a organominerální vrstvě půdy (O+A) při vyšším zastoupení listnatých dřevin

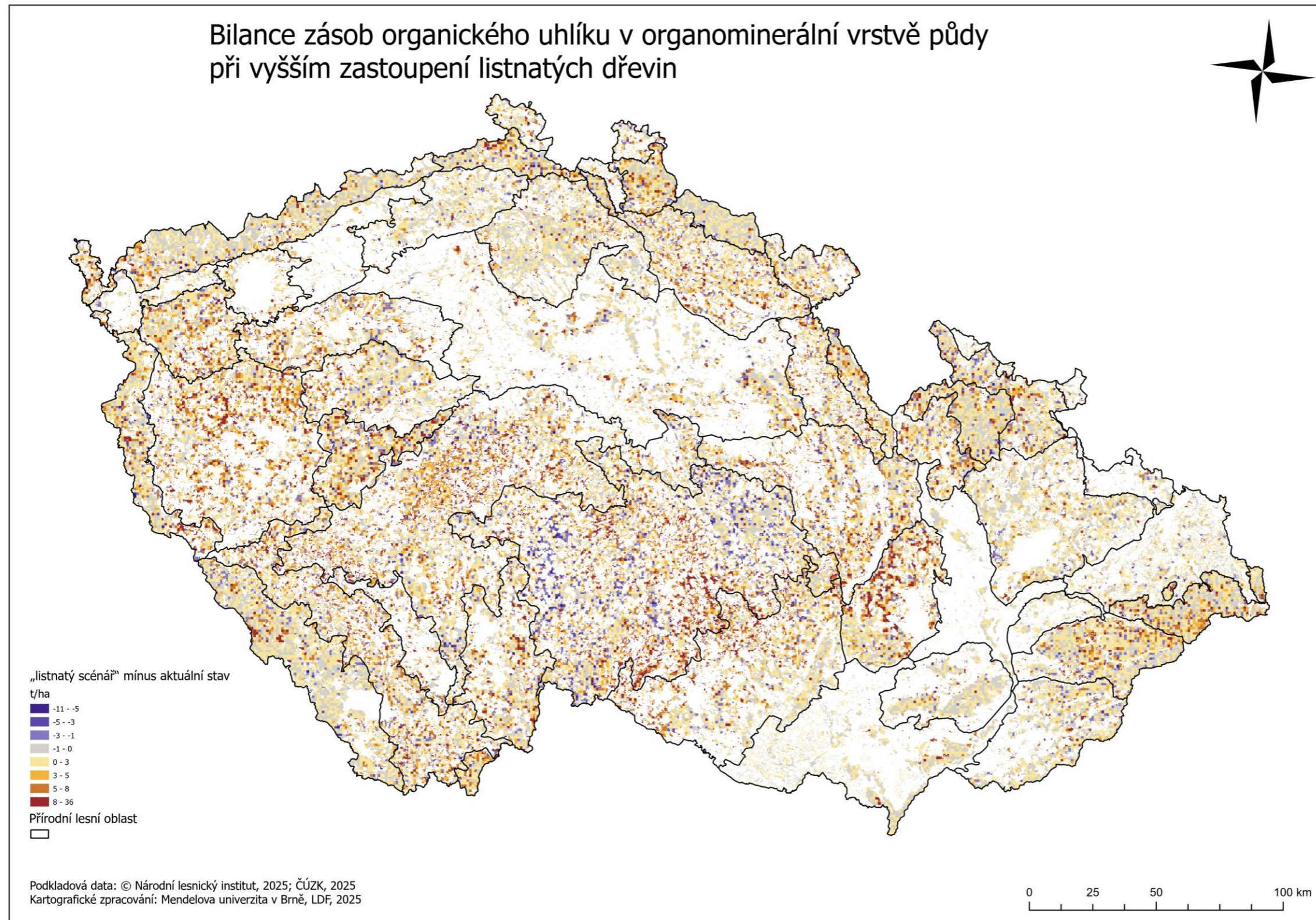


Příloha 5. Bilance zásob organického uhlíku v povrchových vrstvách lesních půd pro „listnatý“ scénář

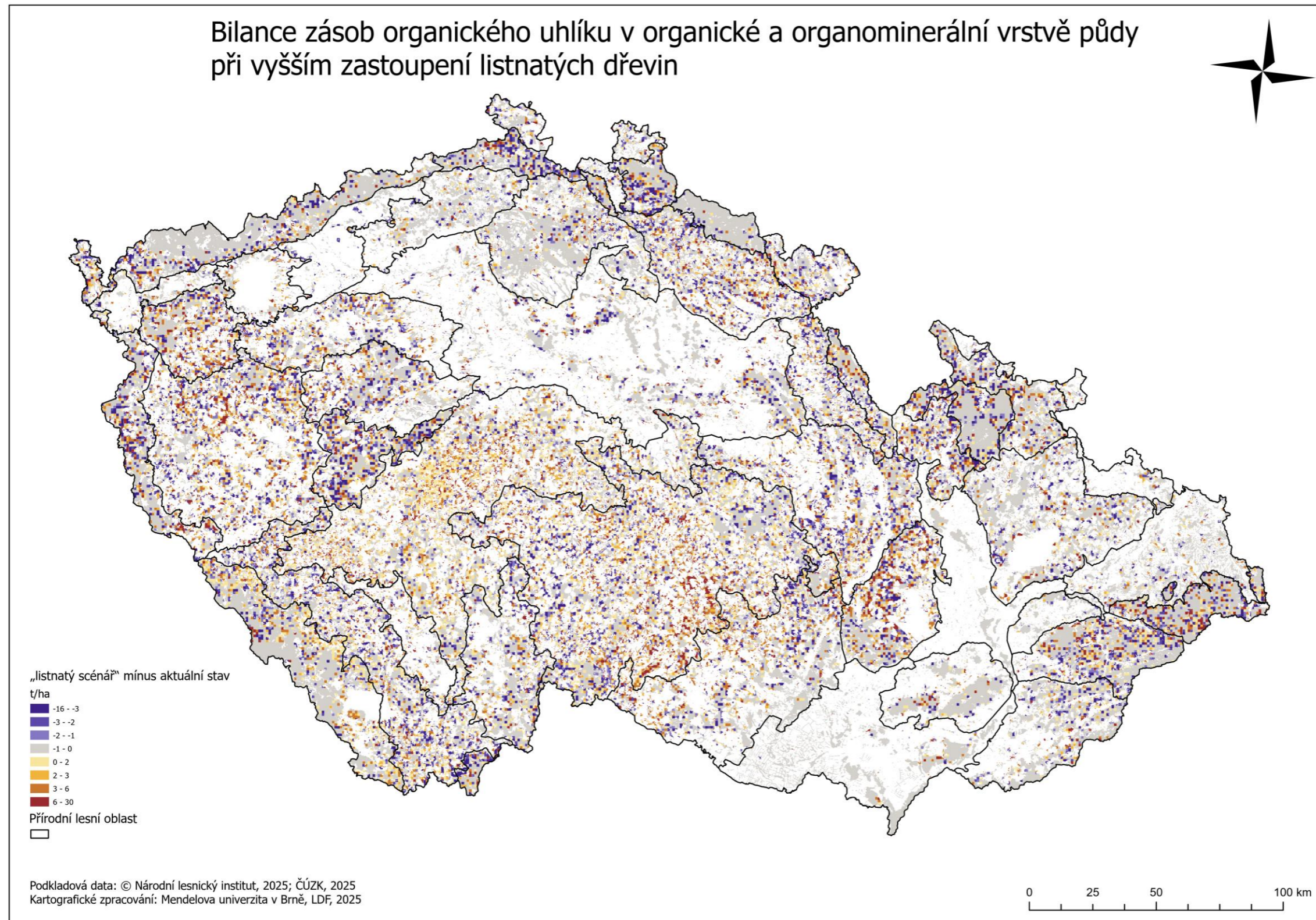
Příloha 5.1 Bilance zásob organického uhlíku v organické vrstvě půdy (O) při vyšším zastoupení listnatých dřevin („listnatý scénář“ minus aktuální stav)



Příloha 5.2 Bilance zásob organického uhlíku v organominerální vrstvě půdy (A) při vyšším zastoupení listnatých dřevin („listnatý scénář“ minus aktuální stav)



Příloha 5.3 Bilance zásob organického uhlíku v organické a organominerální vrstvě půdy (O+A) při vyšším zastoupení listnatých dřevin („listnatý scénář“ minus aktuální stav)



Příloha 6. Bilance zásob organického uhlíku (TOC) na úrovni ČR pro „listnatý“ scénář dle přírodních lesních oblastí (PLO): O – organická vrstva, A – organominerální vrstva půdy.

část (a) Bilance pro celkové rozlohy PLO.

PLO	Název	O	A [t TOC]	O+A
1	Krušné hory	-137 901	188 117	50 216
2	Podkrušnohorské pánve	-19 657	28 918	9 261
3	Karlovarská vrchovina	-57 663	79 251	21 588
4	Doupovské hory	-12 853	15 278	2 425
5	České středohoří	-29 552	42 171	12 618
6	Západočeská pahorkatina	-158 488	229 746	71 257
7	Brdská vrchovina	-85 571	129 476	43 904
8	Křivoklátsko a Český kras	-75 820	101 904	26 083
9	Rakovnicko-kladenská pahorkatina	-53 904	78 834	24 930
10	Středočeská pahorkatina	-264 709	369 506	104 797
11	Český les	-93 716	135 928	42 213
12	Předhoří Šumavy a Novohradských hor	-122 593	180 030	57 437
13	Šumava	-149 318	195 836	46 518
14	Novohradské hory	-8 541	17 830	9 289
15	Jihočeské pánve	-81 793	109 187	27 393
16	Českomoravská vrchovina	-347 875	501 229	153 354
17	Polabí	-123 574	178 316	54 743
18	Severočeská pískovcová plošina a Český ráj	-84 565	120 005	35 440
19	Lužická pískovcová vrchovina	-40 914	58 180	17 267
20	Lužická pahorkatina	-19 388	25 554	6 166
21	Jizerské hory a Ještěd	-38 063	43 854	5 791
22	Krkonoše	-34 711	52 579	17 867
23	Podkrkonoší	-63 108	76 463	13 355
24	Sudetské meziohří	-24 290	25 976	1 685
25	Orlické hory	-26 537	44 242	17 705
26	Předhoří Orlických hor	-29 972	45 760	15 787
27	Hrubý Jeseník	-57 840	81 942	24 102
28	Předhoří Hrubého Jeseníku	-114 557	164 471	49 914
29	Nízký Jeseník	-115 660	171 027	55 367
30	Drahanská vrchovina	-97 474	131 869	34 395
31	Českomoravské meziohří	-107 472	151 773	44 301
32	Slezská nížina	-8 073	11 848	3 774
33	Předhoří Českomoravské vrchoviny	-109 606	160 092	50 486
34	Hornomoravský úval	-16 415	23 258	6 844
35	Jihomoravské úvaly	-40 834	56 769	15 935
36	Středomoravské Karpaty	-37 546	55 433	17 887
37	Kelečská pahorkatina	-10 220	12 632	2 412
38	Bílé Karpaty a Vizovické vrchy	-61 269	86 436	25 166
39	Podbeskydská pahorkatina	-39 293	55 551	16 259
40	Moravskoslezské Beskydy	-83 029	108 871	25 842
41	Hostýnskovsetínské vrchy a Javorníky	-83 729	132 806	49 077
CELKEM		-3 168 096	4 478 948	1 310 852

část (b) Bilance hektarových zásob – průměrné hodnoty za PLO.

PLO	Název	O	A	O+A
			[t TOC / ha]	
1	Krušné hory	-1,213	1,655	0,442
2	Podkrušnohorské pánve	-1,279	1,882	0,603
3	Karlovarská vrchovina	-1,161	1,596	0,435
4	Doupovské hory	-1,517	1,803	0,286
5	České středohoří	-0,901	1,285	0,385
6	Západočeská pahorkatina	-1,317	1,908	0,592
7	Brdská vrchovina	-1,368	2,070	0,702
8	Křivoklátsko a Český kras	-1,339	1,800	0,461
9	Rakovnicko-kladenská pahorkatina	-1,091	1,595	0,504
10	Středočeská pahorkatina	-1,461	2,040	0,579
11	Český les	-1,486	2,156	0,669
12	Předhoří Šumavy a Novohradských hor	-1,267	1,861	0,594
13	Šumava	-1,181	1,549	0,368
14	Novohradské hory	-0,793	1,655	0,862
15	Jihočeské pánve	-1,141	1,523	0,382
16	Českomoravská vrchovina	-1,391	2,004	0,613
17	Polabí	-1,214	1,751	0,538
18	Severočeská pískovcová plošina a Český ráj	-1,071	1,520	0,449
19	Lužická pískovcová vrchovina	-1,174	1,669	0,495
20	Lužická pahorkatina	-1,128	1,487	0,359
21	Jizerské hory a Ještěd	-0,979	1,128	0,149
22	Krkonoše	-1,092	1,654	0,562
23	Podkrkonoší	-1,155	1,399	0,244
24	Sudetské mezíhoří	-1,163	1,244	0,081
25	Orlické hory	-1,270	2,118	0,848
26	Předhoří Orlických hor	-1,302	1,987	0,686
27	Hrubý Jeseník	-1,118	1,584	0,466
28	Předhoří Hrubého Jeseníku	-1,371	1,969	0,597
29	Nízký Jeseník	-1,385	2,048	0,663
30	Drahanská vrchovina	-1,482	2,005	0,523
31	Českomoravské mezíhoří	-1,364	1,926	0,562
32	Slezská nížina	-1,193	1,751	0,558
33	Předhoří Českomoravské vrchoviny	-1,089	1,590	0,501
34	Hornomoravský úval	-1,275	1,807	0,532
35	Jihomoravské úvaly	-1,057	1,470	0,413
36	Středomoravské Karpaty	-0,988	1,458	0,471
37	Kelečská pahorkatina	-1,565	1,934	0,369
38	Bílé Karpaty a Vizovické vrchy	-1,108	1,563	0,455
39	Podbeskydská pahorkatina	-1,576	2,228	0,652
40	Moravskoslezské Beskydy	-1,357	1,779	0,422
41	Hostýnskovsetínské vrchy a Javorníky	-1,241	1,969	0,728
CELKEM		-1,263	1,786	0,523

Příloha 7. Bilance zásob organického uhlíku (TOC) na úrovni ČR pro „listnatý“ scénář) dle cílových hospodářských souborů (CHS): O – organická vrstva, A – organominerální vrstva půdy. Hodnoty zobrazují bilance pro celorepublikové výměry CHS.

část (a) Bilance pro celkové rozlohy CHS.

CHS	O	A [t TOC]	O+A
0	-35 219	38 766	3 547
01	-153 963	222 382	68 419
02	-36 810	45 190	8 380
03	-7 941	17 566	9 625
13	-89 694	128 266	38 572
19	-39 663	58 021	18 358
21	-47 967	65 093	17 126
23	-155 371	209 500	54 129
25	-119 943	166 272	46 329
27	-43 831	62 230	18 399
29	-118 679	173 147	54 469
39	-15 323	21 672	6 348
41	-200 803	296 131	95 329
43	-331 835	467 523	135 688
45	-405 747	584 658	178 911
47	-321 472	457 053	135 581
51	-134 396	192 077	57 681
53	-266 425	364 082	97 657
55	-143 850	211 192	67 342
57	-167 307	244 544	77 237
59	-199 897	275 375	75 478
71	-12 255	13 515	1 260
73	-39 055	51 056	12 000
75	-12 053	18 764	6 710
77	-16 246	24 477	8 230
79	-52 350	70 396	18 045
CELKEM	-3 168 096	4 478 948	1 310 852

část (b) Bilance hektarových zásob – průměrné hodnoty za CHS.

CHS	O	A [t TOC / ha]	O+A
0	-1,219	1,341	0,123
01	-1,219	1,761	0,542
02	-1,416	1,738	0,322
03	-0,474	1,049	0,575
13	-1,103	1,577	0,474
19	-1,246	1,823	0,577
21	-1,312	1,781	0,469
23	-1,241	1,674	0,432
25	-1,215	1,684	0,469
27	-1,269	1,802	0,533
29	-1,254	1,829	0,575
39	-1,155	1,633	0,478
41	-1,289	1,902	0,612
43	-1,320	1,859	0,540
45	-1,322	1,906	0,583
47	-1,281	1,822	0,540
51	-1,292	1,846	0,554
53	-1,230	1,680	0,451
55	-1,252	1,838	0,586
57	-1,257	1,838	0,580
59	-1,313	1,809	0,496
71	-1,188	1,310	0,122
73	-1,303	1,704	0,400
75	-0,971	1,512	0,541
77	-1,142	1,721	0,579
79	-1,280	1,722	0,441
CELKEM	-1,263	1,786	0,523