

ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA V PRAZE
FAKULTA ŽIVOTNÍHO PROSTŘEDÍ
KATEDRA PROSTOROVÝCH VĚD

Hluková mapa odvozená ze satelitních dat

DIPLOMOVÁ PRÁCE

Vedoucí práce: **Ing. David MORAVEC, Ph.D.**

Konzultant: **Ing. Vojtěch HAVLÍČEK, Ph.D.**

Diplomant: **Bc. Lenka MIKOVÁ**

ZADÁNÍ DIPLOMOVÉ PRÁCE

Bc. Lenka Miková

Prostorové vědy v životním prostředí

Název práce

Hluková mapa odvozená ze satelitních dat

Název anglicky

Noise map derived from remote sensing data

Cíle práce

V současné době je tvorba hlukových map velmi časově i finančně náročná, opírá se o mnoho vstupů z terénního sběru a zároveň není vytvářena na celkové ploše ČR. Tato práce si klade za cíl predikovat celoplošné hlukové mapy pomocí snadno dostupných dat dálkového a následně porovnat celkovou přesnost predikce oproti Strategickým hlukovým mapám z roku 2022.

Metodika

Nejprve bude provedena rešerše existujících řešení a dostupných hlukových map, které poslouží jako základní zdroj dat. Z těchto map budou vytvořena trénovací a validační data pro následnou analýzu. Následně bude provedena regresní analýza mezi hlukem a vtipovanými prediktory odvozenými ze satelitních map, která umožní identifikovat vztahy mezi hlukem a dostupnými geoinformačními parametry. Na základě této analýzy bude vytvořen model, který bude schopen odhadnout hladinu hluku pro následné vytvoření plošné hlukové mapy. Na závěr bude spolehlivost mapy ověřena pomocí validačních dat, a bude odhadnuta spolehlivost takového řešení.

Doporučený rozsah práce

40 stran

Klíčová slova

Hluk, Predikce, Satelitní data, Strojové učení, Model

Doporučené zdroje informací

- H. KOI , S-I. FUJITA , I. AOI , H. TAMURA & K. HIRAMATSU (1993) Regional noise prediction and management system using remote sensing, International Journal of Remote Sensing, 14:13, 2427-2443, DOI: 10.1080/01431169308904284
- Leonardo, Eicher, M., Mommert, D., Borth. (2022). Traffic Noise Estimation from Satellite Imagery with Deep Learning. doi: 10.1109/IGARSS46834.2022.9883463
- Liu Y, Goudreau S, Oiamo T, Rainham D, Hatzopoulou M, Chen H, Davies H, Tremblay M, Johnson J, Bockstael A, Leroux T, Smargiassi A. Comparison of land use regression and random forests models on estimating noise levels in five Canadian cities. Environ Pollut. 2020 Jan;256:113367. doi: 10.1016/j.envpol.2019.113367. Epub 2019 Oct 10. PMID: 31662255.
- MOHAMMED, Murtala Uba, et al. Towards Urban Sustainability: Developing Noise Prediction Model in an Informal Setting. Applied Sciences, 2022, 12.18: 9071.
- RAESS, Michelle, et al. Land use regression modelling of community noise in São Paulo, Brazil. Environmental research, 2021, 199: 111231.

Předběžný termín obhajoby

2024/25 LS – FŽP

Vedoucí práce

Ing. David Moravec, Ph.D.

Garantující pracoviště

Katedra prostorových věd

Elektronicky schváleno dne 11. 12. 2024**prof. Ing. Petra Šimová, Ph.D.**

Vedoucí katedry

Elektronicky schváleno dne 12. 12. 2024**prof. RNDr. Michael Komárek, Ph.D.**

Děkan

V Praze dne 21. 01. 2025

Čestné prohlášení

Prohlašuji, že jsem diplomovou práci na téma: **Hluková mapa odvozená ze satelitních dat** vypracovala samostatně a citovala jsem všechny informační zdroje, které jsem v práci použila a které jsem rovněž uvedla na konci práce v seznamu použitých informačních zdrojů.

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

Jsem si vědoma, že odevzdáním diplomové práce souhlasím s jejím zveřejněním podle zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů, ve znění pozdějších předpisů, a to i bez ohledu na výsledek její obhajoby.

Svým podpisem rovněž prohlašuji, že elektronická verze práce je totožná s verzí tištěnou a že s údaji uvedenými v práci bylo nakládáno v souvislosti s GDPR.

V Praze dne 26. 3. 2025

.....
Lenka MIKOVÁ

Poděkování

Předně bych chtěla poděkovat svému vedoucí práce Davidu Moravcovi za vedení, cenné rady a připomínky a také svému konzultantovi Vojtěchu Havlíčkovi. Dále pak svému manželovi a dětem za velkou podporu a mimořádnou trpělivost.

V Praze dne 26. 3. 2025

.....
Lenka MIKOVÁ

Abstrakt

Tato diplomová práce se zabývá možností vytvoření modelu, který by s využitím satelitních dat ve formě různých spektrálních indexů, termálních snímků, údajů o koncentraci atmosférických plynů a informací o intenzitě světelného znečištění dokázal odhadnout hladinu hluku na území České republiky. Pro tento účel byly aplikovány pokročilé modelovací metody, včetně Random Forest, Gradient Boosting a neuronových sítí. Dataset zahrnoval 1 400 bodů, které pokrývaly přibližně 7 % rozlohy České republiky. Výsledky ukazují, že přesnost modelu trénovaného na strategické hlukové mapě je nedostačující pro praktické využití. Nicméně tento přístup vykazuje potenciál, zejména pokud by byla k dispozici vhodnější hluková data zahrnující všechny úrovně hlukového zatížení, včetně nízkých hodnot pod 50 dB nebo reálně naměřené hodnoty hluku.

Klíčová slova

Hluk, predikce, satelitní data, strojové učení, model

Abstract

This thesis explores the possibility of developing a model that, using satellite data in the form of various spectral indices, thermal imagery, atmospheric gas concentration data, and information on light pollution intensity, can estimate noise levels across the Czech Republic. For this purpose, advanced modeling methods such as Random Forest, Gradient Boosting, and neural networks were applied. The dataset included 1,400 points covering approximately 7 % of the Czech Republic's area. The results indicate that the accuracy of the model trained on the Strategic Noise Map is insufficient for practical application. However, this approach shows potential, particularly if more suitable noise data were available, including all noise levels—especially those below 50 dB—or actual measured noise values.

Key words

Noise, prediction, satellite data, machine learning, model

Seznam zkratek

CNOSSOS-EU	Výpočtová metodika <i>Common Noise Assessment Methods in Europe</i>
Colab	Google Colaboratory
MDČR	Ministerstvo dopravy České republiky
MZČR	Ministerstvo zdravotnictví České republiky
SHM	Strategické hlukové mapy
SZÚ	Státní zdravotní ústav
WHO	World Health Organization (Světová zdravotnická organizace)
ZÚ Ostrava	Zdravotní ústav se sídlem v Ostravě

Obsah

1	Úvod.....	1
2	Cíle práce	3
	LITERÁRNÍ REŠERŠE	4
3	Hluk.....	4
3.1	Vlivy hluku	4
3.2	Zdroje hluku.....	5
3.3	Posouzení míry hluku	6
3.3.1	Hlukové indikátory a mezní hodnoty.....	6
3.4	Šíření hluku	7
4	Dostupnost a charakteristika hlukových dat.....	9
4.1	Přímé měření hluku.....	9
4.2	Výpočet hluku	9
4.2.1	Akustické studie.....	9
4.2.2	Hlukové mapy	9
4.3	Způsoby výpočty hluku.....	10
5	Hlukové mapy	12
5.1	Obecné informace.....	12
5.2	SHM a kola SHM.....	13
5.3	Vstupní data pro SHM	14
5.3.1	Ověření výpočtu.....	14
6	Satelitní data	15
6.1	Landsat 8 (NASA)	15
6.1.1	NDVI (<i>Normalized Difference Vegetation Index</i>)	17
6.1.2	NDBI (<i>Normalized Difference Built-up Index</i>).....	17
6.1.3	MNDWI (<i>Modified Normalized Difference Water Index</i>).....	18
6.1.4	UI (<i>Urban Index</i>)	18
6.1.5	ISA (<i>Impervious Surface Area</i>) <i>Index</i>	18
6.2	MODIS (NASA).....	19
6.3	Sentinel-2 (ESA).....	19
6.4	Sentinel-5P (ESA)	20
6.4.1	CO (<i>oxid uhelnatý</i>)	21
6.4.2	HCHO (<i>formaldehyd</i>)	21
6.4.3	NO ₂ (<i>oxid dusičitý</i>)	22

6.4.4	<i>O₃</i> (<i>ozon</i>).....	22
6.4.5	<i>SO₂</i> (<i>oxid siřičitý</i>)	22
6.4.6	<i>CH₄</i> (<i>metan</i>).....	22
6.5	VIIRS (EOG)	22
7	Metody zpracování dat.....	24
7.1	Statistická analýza	25
7.1.1	Lineární regrese a další regresní modely.....	25
7.2	Pokročilé modelovací metody	25
7.2.1	Random Forest.....	26
7.2.2	Gradient Boosting.....	26
7.2.3	Neuronové sítě	27
	METODIKA	29
8	Vstupní data	29
8.1	Hluková data	29
8.1.1	Aglomerace	31
8.1.2	Silnice	31
8.1.3	Železnice.....	32
8.1.4	Letiště.....	33
8.2	Data z dálkového průzkumu Země.....	33
8.2.1	Landsat 8	33
8.2.2	MODIS.....	36
8.2.3	Sentinel-2.....	37
8.2.4	Sentinel-5P	37
8.2.5	VIIRS	40
8.3	Zpracování dat.....	41
	VÝSTUPY	44
9	Analýza a tvorba modelu	44
9.1	Příprava dat.....	44
9.1.1	Roční data	44
9.1.2	Měsíční data.....	54
9.2	Tvorba modelu	55
9.2.1	Random Forest	55
9.2.2	Gradient Boosting.....	56
9.2.3	Neuronové sítě.....	58
10	Souhrn výsledků	60
11	Diskuse.....	61

12 Závěr.....	63
Seznam použité literatury	65
Seznam obrázků	76
Seznam tabulek.....	78
Seznam rovnic.....	78
Seznam příloh.....	78

1 Úvod

Hluk je považován hned po znečištění ovzduší za nejvýznamnější environmentální stresor dnešní doby, a to nejen v kontextu vlivu na člověka, ale také vlivu na životní prostředí. S ohledem na tuto skutečnost je nezbytné získávat co nejaktuálnější informace o oblastech, které jsou hlukem výrazně zasaženy. Aktuálně jsou v České republice lokality s vysokou úrovní hluku, kde je třeba aplikovat protihluková opatření, vybírána na základě strategické hlukové mapy. Tato mapa je však aktualizována v pravidelných pětiletých cyklech a pokrývá pouze 7,2 % území státu. Celý proces od zahájení přes vytvoření hlukové mapy až po implementaci akčních plánů na snížení hlukové zátěže je však také velmi časově náročný.

Tato práce volně navazuje na mou bakalářskou práci s názvem „Porovnání hlukových map a jejich kartografická vizualizace ve 2D a 3D“, která se zabývala porovnáním dvou hlukových map vytvořených ve stejném období. Ačkoliv rozdíly mezi těmito výpočtovými mapami byly většinou nevýznamné, v některých oblastech se odchylky přesahovaly i 10 dB. Možnou příčinou těchto rozdílů může být neaktuálnost vstupních dat nebo způsob jejich výpočtu.

Existuje několik způsobů získávání informací o hluku, avšak mnoho z těchto metod vyžaduje vstupní data z terénního měření, údaje o dopravě nebo další podklady, jejichž získání je časově i finančně náročné. Tyto faktory značně omezují možnost aplikace na rozsáhlejší území. Ve světě se proto k odhadu hluku na větším území využívají regresní modely založené na datech o využití krajiny (tzv. Land Use), případně metoda Random Forest (Raess et al., 2021; Liu et al., 2019; Staab et al., 2021; Mohammed et al., 2022). V Itálii byla například použita metoda využívající menší počet stanic ke zpřesnění rovnic pro konkrétní lokality (Zambon et al., 2018), zatímco v Číně byla testována možnost získávání hlukových dat prostřednictvím speciální mobilní aplikace zapojující veřejnost v rámci Citizen Science¹ (Shim et al., 2016). Tyto přístupy přinášejí zajímavé výsledky, avšak jsou buď omezeny na malá území, nebo vyžadují vstupní data, která nejsou vždy veřejně dostupná a aktualizovaná.

¹ Občanská věda – aktivní zapojení občanů do vědy (přispívání k rozvoji a poznání)

Nabízí se tedy možnost využití dat, která lze jednoduše získat pro jakékoliv území a která zároveň umožňují zkrácení časového intervalu pořízení vstupních dat. Tuto podmínu splňují satelitní data. Možnost predikce hluku na základě satelitních snímků již testoval Eicher et al. (2022), který využil neuronovou síť U-Net k modelování hlukové zátěže na základě snímků Sentinel-2.

Satelitní data jsou dnes široce dostupná a jejich kombinací lze získat nepřeberné množství informací o stavu naší planety v různých prostorových měřítkách. Moderní technologie navíc umožňují zpracovávat velký objem snímků bez nutnosti využití superpočítáčů. Rozvoj umělé inteligence (*Artificial Intelligence*, AI), zahrnující strojové učení (*Machine Learning*, ML) a hluboké učení (*Deep Learning*, DL), výrazně rozšířil možnosti analýzy těchto dat. Díky tomu může model sám testovat a optimalizovat své výsledky na základě zpracovaných dat, dokud nesplní požadované kritéria.

Tato práce využívá širokou škálu satelitních dat. Snímky z družic Landsat 8 a Sentinel-2 jsou použity k výpočtu spektrálních indexů, které umožňují detektovat vegetaci a zastavěné oblasti, jež ovlivňují šíření hluku. Na základě předpokladu, že zdroje hluku zároveň vyzařují teplo, byla zařazena data o teplotě zemského povrchu ze satelitu MODIS. Protože hlavními zdroji hluku jsou doprava a průmysl, byla zvolena také data ze Sentinel-5P, která mapují koncentrace atmosférických plynů související s těmito aktivitami. V neposlední řadě byla využita data o světelném znečištění, které koreluje s dopravní a průmyslovou činností, protože nejvyšší míra světelného znečištění je zaznamenávána v hustě zastavěných a průmyslových oblastech.

Doposud nebyla provedena žádná studie, která by zkombinovala data z družic Landsat 8, MODIS, Sentinel-2, Sentinel-5P a VIIRS s využitím pokročilých modelovacích metod k predikci hlukové zátěže na základě strategických hlukových map (SHM).

2 Cíle práce

Hlavním cílem této práce je vytvoření hlukové mapy na základě predikovaných dat získaných prostřednictvím dálkového průzkumu Země. Mezi dílčí cíle patří analýza současných metod nepřímého získávání údajů o intenzitě hluku z dostupných dat.

Dále bude v jazyce Python vytvořen skript pro poloautomatizované získávání vstupních dat, která přímo či nepřímo souvisejí s hlukovým zatížením. Tato data budou čerpána z následujících zdrojů:

- Landsat 8 – spektrální indexy (NDVI, NDBI, MNDWI, UI, ISA),
- MODIS – teplota povrchu,
- Sentinel-2 – spektrální indexy shodné s Landsat 8,
- Sentinel-5P – koncentrace atmosférických plynů (CO, HCHO, NO₂, SO₂, CH₄, O₃),
- VIIRS – úroveň světelného znečištění.

Získaná data budou dále zpracována v jazyce R, v němž bude vytvořen predikční model pro odhad hlukového zatížení vybraného území. K modelování budou využity pokročilé metody strojového učení (Machine Learning) a neuronových sítí. Model bude následně testován na datech strategického hlukového mapování z roku 2022. Po vyhodnocení jeho přesnosti bude vytvořena celoplošná hluková mapa České republiky a provedena její validace.

Očekávaným přínosem této práce je vyhodnocení vlivu vybraných spektrálních indexů, koncentrací atmosférických plynů, teploty povrchu a úrovně světelného znečištění na hlukové zatížení, přičemž hlavním referenčním zdrojem pro validaci výsledků bude dataset SHM 2022.

LITERÁRNÍ REŠERŠE

3 Hluk

Hluk je možné definovat různými způsoby, přičemž je vždy důležitý kontext, ve kterém je slovo **hluk** použito. Z pohledu akustiky jej Everest a Pohlmann (2009) považují za nosič informací, případně za nežádoucí šum, přičemž uznávají, že je obtížné rozlišit, kdy se jedná o nežádoucí šum a kdy o legitimní nosič informací. Státní zdravotní ústav (2023) nahlíží na hluk v kontextu zdraví jako na „*zvuky příliš časté, příliš silné nebo vyskytující se v nevhodnou dobu a situaci*“.

Fyzikové či akustici hluk považují za zvukovou vlnu. Mechlová et al. (1999) uvádí, že se akustika zabývá vlastnostmi zvukového vlnění, jeho vznikem, šířením a působením na tělesa a rovněž jeho účinky na sluchové orgány. Míra působení na sluchový orgán lze vysvětlit hlasitostí. Jedná se o subjektivní veličinu, která je úzce spojena s posluchačem. Za objektivní veličinu lze považovat intenzitu hluku. Tu lze rozdělit na práh slyšení a práh bolesti, které odlišují, kdy je intenzita nejmenší a kdy naopak intenzita nepříjemná až bolestivá.

Všechny definice a chápání pojmu hluku jsou shodné v tom, že se jedná o projev, jenž je vnímán negativně. Podle Paschalidou et al. (2019) a EEA (2023) je míra zatížení hlukem subjektivní a závisí na intenzitě hluku, jeho zdroji a délce expozice.

Hlavním právním předpisem v České republice je **zákon č. 258/2000 Sb., o ochraně veřejného zdraví a o změnách některých souvisejících zákonů** a prováděcí předpisy v podobě **vyhlášky č. 315/2018 Sb., o strategickém hlukovém mapování**. V neposlední řadě rovněž **nařízení vlády č. 272/2011 Sb., o ochraně zdraví před nepříznivými účinky hluku a vibrací**.

3.1 Vlivy hluku

Hlukové znečištění je celosvětovým problémem, který má vliv nejen na zdraví člověka, ale rovněž na životní prostředí. EEA (2020b) považuje za největšího původce zdravotních obtíží u lidí znečištění ovzduší, avšak hluk je v pořadí druhým nebezpečným faktorem, který negativně působí na zdraví populace.

Nepříznivé účinky hluku na lidské zdraví jsou podle MZČR (2015b) takové, které vedou ke zhoršení nebo poškození funkcí organismu, ke snížení odolnosti organismu vůči stresu či mohou vést ke zvýšení vnímavosti k jiným nepříznivým vlivům prostředí. Rovněž účinek hluku na spánek, a to jak ve fázi usínání, tak i na

délku či kvalitu spánku, může mít podle MZČR (2015c) za následek rozmrzelost, únavu, špatnou náladu, snížení výkonu či bolest hlavy. Světová zdravotnická organizace (WHO) se vlivy hluku na zdraví člověka (dospělých i dětí) zabývá dlouhodobě a uvádí, že účinky hluku mohou být jak krátkodobého, tak dlouhodobého charakteru, přičemž okolo 40 % populace EU je vystaveno hluku ze silniční dopravy přesahujícího 55 dB a 20 % populace EU je vystaveno hluku přesahujícího 65 dB během dne (WHO, 2024). EEA (2020b) uvádí, že dlouhodobá expozice hlukem může mít vliv na zdraví, a to ve formě rušení spánku, vlivu na kardiovaskulární či metabolický systém či může způsobovat poruchy kognitivních funkcí u dětí a další druhy poruch, například poruchu čtení. Vliv hluku na zdraví je předmětem mnoha studií. Cai et al. (2020) zkoumal souvislost s obezitou populace, Guha a Gokhale (2023) sledovali vliv na kardiovaskulární systém populace a Hong et al. (2022) se zabýval zdravotními dopady stavebního hluku na populaci. Všechny tyto studie prokazují negativní dopad hluku na lidské zdraví.

Co se týče hluku, životní prostředí je ohroženo nejvíce hlukem z dopravy. Dopravní hluk způsobuje živočichům obdobně jako lidem řadu fyzických i psychických problémů, které mohou vést k problémům při párení a rozmnožování, jak uvádí den Boer a Schrote (2007), a dále může ovlivnit či narušit komunikaci mezi živočichy (Lima et al., 2022). Antropogenní hluk ovlivňuje savce, hmyz, ryby a další druhy živočichů, ať už pochází z pozemní dopravy (Martin et al., 2022), nebo z lodní dopravy (Faria et al., 2022; Amorim et al., 2022; Woods et al., 2022).

Nepříznivé vlivy hluku na zdraví člověka a životní prostředí jsou dostatečně prokázány a je jim věnována patřičná pozornost ze strany jednotlivých států a organizací, a to jak evropských, tak světových. Mezi nástroje, které se využívají v boji proti hluku, patří například strategické hlukové mapy, na které navazuje tvorba akčních plánů a realizace protihlukových opatření. Tyto nástroje jsou blíže rozebrány v kapitole 4.

3.2 Zdroje hluku

Jak již bylo popsáno v předešlé podkapitole, při posuzování hluku jej dělíme na hluk ve vnitřním a venkovním prostředí. Celá tato práce se zabývá pouze **hlukem ve venkovním prostředí**. Podle směrnice 2002/49/ES patří mezi velké **zdroje hluku** silniční a železniční vozidla, infrastruktura, letadla, zařízení určená k použití ve venkovním prostředí, průmyslová, mobilní a strojní zařízení. Intenzita hluku se vyjadřuje hlukovým indikátorem, což je dle SZÚ (2023a) relativní logaritmická veličina označující se L , což je hladina akustického tlaku s jednotkou

dB. Beran (2010) uvádí, že zdroj hluku lze dále rozlišovat na to, zda se jedná o bodový zdroj, liniový (přímkový) zdroj nebo plošný akustický zdroj.

3.3 Posouzení míry hluku

Pro posouzení míry zasažení hlukem se využívají hlukové indikátory pro které jsou stanoveny mezní hodnoty. Kromě těchto parametrů je z pohledu vnímání hluku podstatná i tónová složka hluku.

3.3.1 Hlukové indikátory a mezní hodnoty

Podstatou zavedení hlukových indikátorů je objektivní posouzení míry zasažení hlukem s tím, aby bylo možné porovnání jak v čase, tak místě, a to za použití jednočíselné hodnoty. Pro posouzení míry hluku ve venkovním prostředí byly definovány **hlukové indikátory (ukazatele)**², které vychází ze směrnice 2002/49/EC a byly implementovány do vyhlášky č. 315/2018 Sb.³:

- L_{den} – hlukový indikátor pro celkové obtěžování hlukem den-večer-noc;
- L_d – hlukový indikátor pro obtěžování hlukem během dne, tj. v rozmezí 6:00–18:00⁴;
- L_e – hlukový indikátor pro obtěžování hlukem během večera, tj. v rozmezí 18:00–22:00⁵;
- L_n – hlukový indikátor pro noční období pro rušení spánku, tj. v rozmezí 22:00–6:00⁶;
- L_{dn} – hlukový indikátor pro den-noc – vážený energetický průměr s noční penalizací 10 dB z hodnot L_{Aeq} pro den (6:00–22:00) a pro noc (22:00–6:00)⁷.

Tyto indikátory představují dlouhodobou průměrnou hodnotu za období kalendářního roku. Jejich **mezní hodnoty** si určují jednotlivé členské státy, přičemž se mohou lišit pro různé zdroje hluku a měnit v čase (MZČR, 2015d). Aktuální mezní hodnoty jsou stanoveny ustanovením § 2 odst. 5 vyhlášky č. 315/2018 Sb. a jsou uvedeny v tabulce 1. Mezní hodnoty se stanovují pouze pro zdroje hluku, nikoliv pro aglomerace jako celek.

² V celé práci byly použity zkratky indikátorů vycházející z angličtiny, ve výše uvedeném přehledu indikátorů byly označeny tučně, tj. L_{den} , L_d , L_e , L_n , L_{dn} .

³ Uvedené ukazatele podle směrnice 2002/49/ES.

⁴ Většina evropských zemí používá interval pro den 7:00–19:00.

⁵ Většina evropských zemí používá interval pro večer 19:00–23:00.

⁶ Většina evropských zemí používá interval pro noc 23:00–7:00.

⁷ Jedná se o starší hlukový indikátor, který byl postupně nahrazen indikátorem L_{den} .

Tabulka 1: Mezní hodnoty hlukových ukazatelů (zdroj: Miková podle vyhlášky č. 315/2018 Sb.).

Zdroj hluku / Indikátor	L_{den} [dB]	L_n [dB]
Silniční doprava	70	60
Železniční doprava	70	65
Letecká doprava	60	50
Integrovaná zařízení	50	40

Nařízení vlády č. 272/2011 Sb. týkající se vnitřního prostředí pracuje s následujícími ukazateli:

- hladinou akustického tlaku označeného $L_{Aeq,T}$ – ekvivalentní hladina akustického tlaku zjištěná pomocí filtru A⁸;
- $L_{Aeq,16h}$ a $L_{Aeq,8h}$ – ukazatele pro hluk z dopravy na pozemních komunikacích, využívají se také na železniční dráhy a hluk z leteckého provozu pro denní (6:00–22:00) a noční hodiny (22:00–6:00).

Hlukový indikátor L_{den} je dle směrnice 2002/49/EC (2002) a European Commission (Kephelopoulos et al., 2012) definován rovnicí 1.

Rovnice 1 - Vzorec pro výpočet hlukového indikátoru L_{den} podle směrnice 2002/49/EC

$$L_{den} = 10 \cdot \log_{10} \left(\frac{12}{24} \cdot 10^{\left(\frac{L_d}{10}\right)} + \frac{4}{24} \cdot 10^{\left(\frac{L_e+5}{10}\right)} + \frac{8}{24} \cdot 10^{\left(\frac{L_n+10}{10}\right)} \right)^9.$$

Ve vzorci se pracuje s hodnotami jednotlivých ukazatelů pro den, večer a noc. U ukazatele pro večer se připočítává penalizace 5 dB a pro noc se připočítává 10 dB. Důvodem připočítávání penalizace je zohlednění závažnějších účinků hluku ve večerních a nočních hodinách (Vandasová a Fialová, 2016).

3.4 Šíření hluku

Beran (2010) uvádí, že na šíření hluku má vliv mnoho atmosférických faktorů a dalších vlivů, jako jsou například vítr, teplota, vlhkost a déšť, odraz a ohyb zvukových vln či průchod vln překážkami. Rychlosť a směr větru ovlivňují směr šíření hluku obdobně jako teplota, kdy při rostoucí teplotě rovněž roste rychlosť zvuku. Naproti tomu vlhkost prostředí způsobuje zeslabování zvuku. V obydlených oblastech je pak velice důležitý faktor odrazu, kdy část vlnění je pohlcena a část

⁸ Filtr A se používá pro přizpůsobení objektivně změřené hladiny subjektivně vnímané hlasitosti.

⁹ Koefficient, který se přičítá k indikátoru L_e a L_n je tzv. večerní (+ 5 dB) a noční penalizace (+ 10 dB).

odražena. Míra pohlcení a odražení závisí na materiálu, přičemž se materiály, jejichž pohltivost je menší, a tedy odrazivost větší, nazývají akusticky tvrdé povrchy. U vlhkosti je zase velice důležitá frekvence vlnění, kdy u vyšších frekvencí je útlum zvuku větší. Ohyb zvukových vln závisí na frekvenci - při vyšších frekvencích se zvuk lépe odstíní překážkou. Průchod vln překážkou souvisí se šířením zvukové vlny průchodem štěrbinou či jiným otvorem skrz překážku. Podstatným faktorem je rovněž rychlosť šíření hluku, která závisí na typu látky. Například rychlosť šíření hluku u vybraných pevných látek a plynů je uvedena v tabulce 2.

Tabulka 2: Rychlosť šíření hluku ve vybraných pevných látkách a plynach (zdroj: Beran, 2010)

		Rychlosť šíření [m/s]
PEVNÉ LÁTKY	Beton	1 700
	Bukové dřevo	3 400
	Dubové dřevo	3 400
	Sklo	5 200 ¹⁰
	CO ₂	258
PLÝNY	CO	337
	N ₂	336
	Vodní pára	405

¹⁰ Uvedená rychlosť se uvádí pro teplotu 20 °C.

4 Dostupnost a charakteristika hlukových dat

Údaje o hluku lze získat buď **přímým měřením**, anebo **výpočtem** za použití matematických modelů.

4.1 Přímé měření hluku

Přímá metoda spočívá v umístění měřicích zařízení přímo v terénu. MZČR vydalo *Metodický návod pro měření a hodnocení hluku v mimopracovním prostředí*, který stanovuje způsoby měření a podmínky, za nichž se provádí měření a následné zpracování dat pro jednotlivé zdroje hluku. S metodou přímého měření je možné se nejčastěji setkat při předkládání akustické studie v souvislosti s přípravou stavební dokumentace nebo při aktualizaci modelů pro výpočtové metody. Její výhodou je rychlosť pořízení, avšak její nevýhodou je omezená doba, kdy se může měření provádět, tj. pouze v určitých měsících v roce, aby naměřené hodnoty reprezentovaly průměrnou standardní situaci (MZČR, 2017).

4.2 Výpočet hluku

Na rozdíl od přímého měření lze hlukové zatížení stanovit také výpočtem za použití matematických modelů. Existuje více výpočtových metod, přičemž každá z nich vychází z odlišného matematického modelu v závislosti na zdroji hluku. Tento přístup se využívá například při strategickém hlukovém mapování, kde se hlukové zatížení stanovuje výpočtem. Metoda spočívá ve výpočtu hlukového zatížení v pravidelně rozmištěných výpočtových bodech na vybraném území (EEA, 2020a). Výsledkem je tedy opět shluk bodů s hodnotami hlukové intenzity v intervalech po 5 dB s označením LOW pro nižší hodnotu a HIGH pro vyšší hodnotu v rozpětí.

4.2.1 Akustické studie

Nejčastěji se vytvářejí **akustické studie** zaměřené na měření hluku v konkrétní oblasti. Akustické studie mohou být modelovány po vložení vstupních dat do predikčního softwaru, kde se hluk vypočítává na základě informací o dopravě v daném místě a dalších parametrech. Další možností je akustická studie založená na naměřených hodnotách získaných přímo v terénu. Akustické studie jsou komplexní, avšak obvykle se týkají pouze malého území.

4.2.2 Hlukové mapy

Hlukové mapy jsou obdobně jako některé akustické studie vypočítány v predikčním softwaru na základě vstupních dat týkajících se jednotlivých zdrojů

hluku. Hluková mapa má však větší prostorový rozsah, a to nejčastěji ve velikosti města, kraje či vybraných větších celků.

4.3 Způsoby výpočty hluku

Hlavním současným způsobem výpočtu je využití výpočtových metodik¹¹. Jedinou zavázanou metodikou v rámci EU je aktuálně CNOSSOS-EU¹², jež využívá velké množství vstupních dat (3D model terénu, umístění průmyslových zdrojů, informace o dopravních cestách, modely vedení komunikací a kolejí, dopravní informace konkrétně intenzita dopravy, informace o budovách a překážkách v šíření hluku, informace o způsobu využití krajiny, údaje o vzletech a přistání, informace o územním členění, údaje o dopravních modelech měst, údaje o počtu osob žijících v budovách a další).

V minulosti se různé studie pokoušely odvodit hlukové mapy z jiných vstupních dat. Tyto studie se zaměřují na různé zdroje hluku a liší se ve vstupních dat i metod jejich zpracování. Jedním z těchto přístupů je například použití **Land-Use regresního modelu** (LUR) či **Random Forest modelu** (RF). Raess et al. (2021) ve své studii uvádí, že jejich regresní model využívající plochy (LUR) k odhadu expozice hluku ve společenství v západní části Sāo Paula v Brazílii vykazoval dobrou shodu. Hlavními prediktory hluku byly index vegetace, oblasti slumů a vzdálenost od silnic, vzdělávacích zařízení a obytných oblastí. Tuto metodu autoři doporučují pro odhad hlukové zátěže v zemích s nízkými příjmy. Liu et al. (2019) ve své studii porovnávala RF model a LUR model v pěti kanadských městech. Regresní model rovněž využil Staab et al. (2021) na části území Německa (Koblenz a okolní regiony) a Mohammed et al. (2022) v Nigérii (Kano).

Zambon et al. (2018) využíval data přímo z terénu v kombinaci s neakustickými parametry z modelu dopravy. Výstupem byla akustická mapa Milána v reálném čase, vytvořená za pomoci malého počtu stanic pro monitorování hluku. Data z **monitorovacích senzorů** v reálném čase se použila k sestavení prediktivních rovnic pro lokální ekvivalentní hladinu hluku na daném místě. Tyto rovnice berou v úvahu příspěvky monitorovacích stanic a mapy statického hluku spojené se seskupenými silnicemi. Neméně zajímavým přístupem je využití **Citizen Science** a chytrých telefonů, jak ukazuje studie Shim et al. (2016) ze Soulu, kde za použití speciální aplikace nainstalované u dobrovolníků byla získávána data o hluku po

¹¹ <https://mzd.gov.cz/metodiky-vypoctu-shm/>

¹² Povinnost používat metodiku mají členské státy od 31. 12. 2018.

dobu 7 dní v 10minutových intervalech. Oba tyto přístupy však vyžadují přímé měření.

V neposlední řadě je nutné zmínit článek Koi et al. z roku 1993, který popisuje využití dat ze satelitu **Landsat-3**¹³ k odhadu hlukového znečištění. Při práci s daty byla použita metoda GMDH (Group Method of Data Handling), přičemž vyvinutý predikční systém odhaduje faktory, které ovlivňují šíření hluku na základě podílu půdy zastavěné budovami. Kromě satelitních dat byla dále použita data o intenzitě dopravy, vydefinované průmyslové oblasti a simulované hodnoty NO₂. Satelitní data zde byla použita pro stanovení Land-Use. Posledním zajímavým přístupem je využití satelitních dat Eicherem et al. (2022) s využitím **DL**, a to použitím segmentační modelu **U-Net** trénovaného na snímcích ze **Sentinel-2**. Eicher et al. (2022) na konferenci IGARSS 2022 prezentoval aplikaci modelu na data ze Sentinel-2 s prostorovým rozlišením 10 m pro predikci hluku ze silniční dopravy za použití dat o odhadu hluku silničního provozu ze Švýcarska.

¹³ Prostorové rozlišení vstupních data bylo 80 m.

5 Hlukové mapy

Volně dostupných dat o hluku, která by pocházela od jednoho poskytovatele, byla sbírána jednotnou metodologií a pokrývala co největší území ČR, je velmi omezené množství. Hlavním zdrojem takových dat jsou v ČR strategické hlukové mapy, jejich pořízení spadá do gesce Ministerstva zdravotnictví ČR. Jedná se o kategorická data obsahující interval hlukové hladiny v rozmezí 5 dB, přičemž první kategorie obsahuje interval 50–55 dB a poslední kategorie zahrnuje hodnoty hluku větší než 75 dB. Přestože pokrývá jen 7,2 % území ČR, jedná se o zdroj, jenž pokrývá největší část území.

5.1 Obecné informace

Směrnice 2002/49/ES stanovuje od roku 2007 povinnost všech členských států Evropské unie vytvářet **strategické hlukové mapy** (SHM), a to každých 5 let. Tyto pravidelné cykly se označují jako kola SHM (MZČR, 2015e). Směrnice stanovila, že mapy mají být vytvořeny pro všechny:

- hlavní silnice, po kterých projede za rok více než 6 000 000¹⁴ vozidel za rok,
- hlavní železniční tratě, po kterých projede více než 60 000¹⁵ vlaků za rok,
- hlavní letiště určená členským státem, které má více než 50 000 vzletů a přistání za rok,
- vybrané aglomerace s více než 250 000¹⁶ obyvateli.

Hlukové mapy pro vybrané aglomerace představují akustické součty všech jednotlivých zdrojů hluku (Ondřej Volf, 25. 7. 2022, in litt.). Ke každému zdroji hluku a kolu SHM je zpracována závěrečná zpráva obsahující informace o použitých datech a celém postupu.

Tyto mapy slouží nejen laické veřejnosti, ale rovněž dalším státním orgánům jako podklad v boji proti dopadu hluku na obyvatelstvo. Ministerstvo zdravotnictví České republiky je gestorem pro pořízení strategických hlukových map. To pověřilo svou přímo řízenou organizaci Zdravotní ústav se sídlem v Ostravě jejich přípravou a zpracováním. **Hlukové mapy** jsou následně využívány jako

¹⁴ Od 2. kola SHM byla hodnota snížena na 3 000 000 vozidel za rok.

¹⁵ Od 2. kola SHM byla hodnota snížena na 30 000 vlaků za rok.

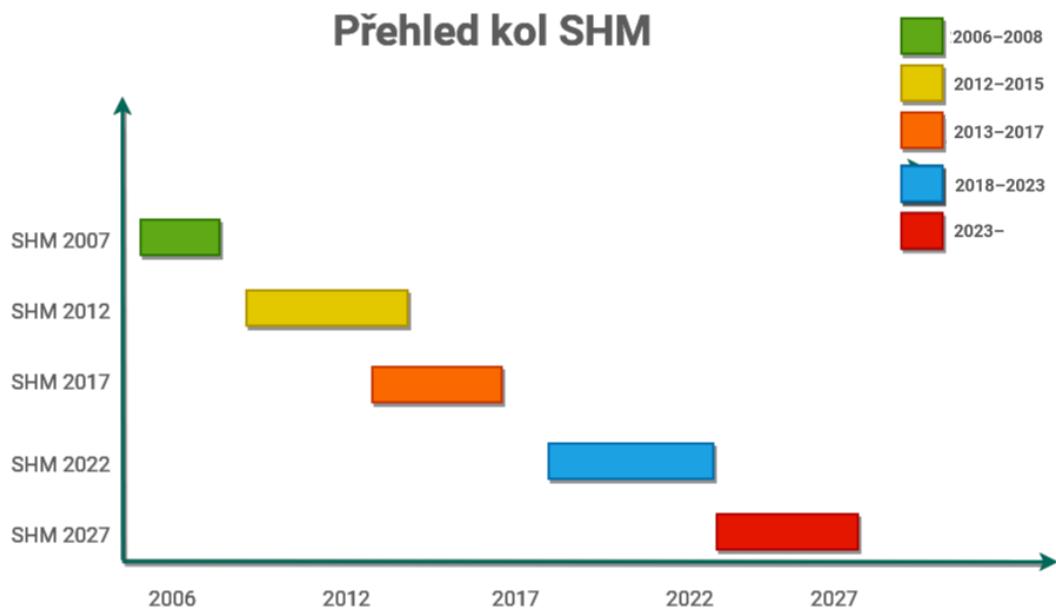
¹⁶ Od 2. kola SHM byla hodnota snížena na 100 000 obyvatel.

podklad pro tvorbu akčních plánů. Jedná se o strategický dokument popisující vybrané rizikové lokality, přehled opatření a termíny pro jejich realizaci včetně předpokládaného výsledku implementace příslušných opatření (MDČR, 2024).

Na základě směrnice Komise (EU) 2015/996 ze dne 19. května 2015 o stanovení společných metod hodnocení hluku podle směrnice Evropského parlamentu a Rady 2002/49/ES a její novely 2021/1226 jsou všechny členské státy povinny využívat pro výpočet strategických hlukových map od 31. 12. 2018 metodiku CNOSSOS_EU (Common Noise Assessment Methods in Europe). EEA (2020a) tento krok podnikla, aby bylo možné porovnat údaje mezi členskými státy a rovněž posuzování hlukových map bylo lépe harmonizováno.

5.2 SHM a kola SHM

Jak již bylo uvedeno výše, strategická hluková mapa je výpočtová mapa, která má strategický význam pro minimalizaci hluku ve venkovním prostředí, v němž je hlukem zasaženo velké množství lidí. Do doby odevzdání této práce byla ukončena celkem **4 kola strategického hlukového mapování** (viz obrázek 1). Veškeré informace o jednotlivých kolech jsou stále veřejně dostupné (MZČR, 2015a).



Obrázek 1: Přehled kol strategického hlukového mapování (zdroj: vlastní zpracování podle MZČR, 2015e).

5.3 Vstupní data pro SHM

Strategické hlukové mapy jsou získány z modelu, do kterého vstupuje velké množství dat od různých subjektů. SHM 2022 byly vypracovány v akustickém programu LimA Software Suite (ZÚ Ostrava, 2022), což je software podporující předpovědi hluku pro průmysl, silnice, železnice, letiště a větrné turbíny (SoftNoise, 2024). Všechna tato data se pravidelně aktualizují, avšak každá z datových sad v různých časových intervalech. V modelu nebyly zohledněny lokální meteorologické podmínky, neboť lokální meteorologické koeficienty nejsou dostupné (ZÚ Ostrava, @2022).

Získávání vstupních dat pro výpočtový model je velice náročný a zdlouhavý proces, jak je patrné z diagramu uvedeného v příloze 1, přičemž v datech se každoročně odhalí mnoho chyb, které je třeba dále zpřesnit, doplnit či opravit na základě ortofotomap, Google Street View a dalších zdrojů dat. V minulosti se jednalo například o doplnění vrstvy protihlukových stěn, úpravy liniového vedení komunikací, zpřesnění výškového profilu v okolí silnic a železnic, umístění mostů, segmentace tratí podle technických parametrů či revize vrstvy lůžkových zdravotnických a školských zařízení (ZÚ Ostrava, 2022).

Postupy práce s daty jsou podrobněji popsány v závěrečné zprávě ZÚ Ostrava (2022).

Vstupní data by se dala rozdělit do 2 oblastí, a to data využitelná přímo v modelu (např. sčítání dopravy, digitální model reliéfu apod.) a dále data doplňující (např. data o školských zařízeních, CORINE Land Cover¹⁷).

5.3.1 Ověření výpočtu

Jelikož jsou strategické hlukové mapy výstupem, jenž zohledňuje akustickou situaci v dlouhodobějším měřítku, je ověření dat v terénu problematické. Model byl však ve větším rozsahu ověřen pomocí dat z přímého měření v rámci SHM 2007, kde došlo k porovnání již vypočtených výsledků s novými výpočty (ZÚ Ostrava, 2022).

¹⁷ Komplexní, podrobný a harmonizovaný datový soubor o krajinném pokryvu a využití území na evropském kontinentu.

6 Satelitní data

Počet poskytovatelů satelitních dat se pozvolna navýšuje. Kromě veřejných poskytovatelů NASA a ESA přibývají rovněž komerční poskytovatelé jako Planet s PlanetScope a SkySat, Maxar Technologies s WorldView a GeoEye či Capella Space s Iceye. Kromě multispektrálních senzorů (Sentinel-2, Landsat) jsou dále hyperspektrální senzory (PRISMA, EnMAP, HICO), radarové senzory (SAR) (Sentinel-1, TerraSAR-X, Capella Space) a v neposlední řadě také LiDARové senzory. Senzory se liší nejen podle pokrývajících spektrálních pásem, ale i časovým a prostorovým rozlišením. Z veřejných poskytovatelů má například Sentinel-2 prostorové rozlišení 10 m a časové rozlišení 5 dní a z komerčních například PlanetScope má prostorové rozlišení 3 m a časové rozlišení 1 den, či dokonce WorldView-3 disponuje prostorovým rozlišením 30 cm s možností aktualizace dat na základě poptávky. V současnosti není již nutné snímky přímo stahovat, ale lze využít některou z platform pro zpracování dat, například Google Earth Engine (GEE) či Amazon Web Service.

Satelitní data mají kromě nespočtu výhod i určitá omezení, jako jsou oblačnost (u optických senzorů), přístupnost hyperspektrálních dat, náročnost pro jejich zpracování a výše nákladů pro získání dat (zejména komerční data s vysokým rozlišením) a také časová dostupnost.

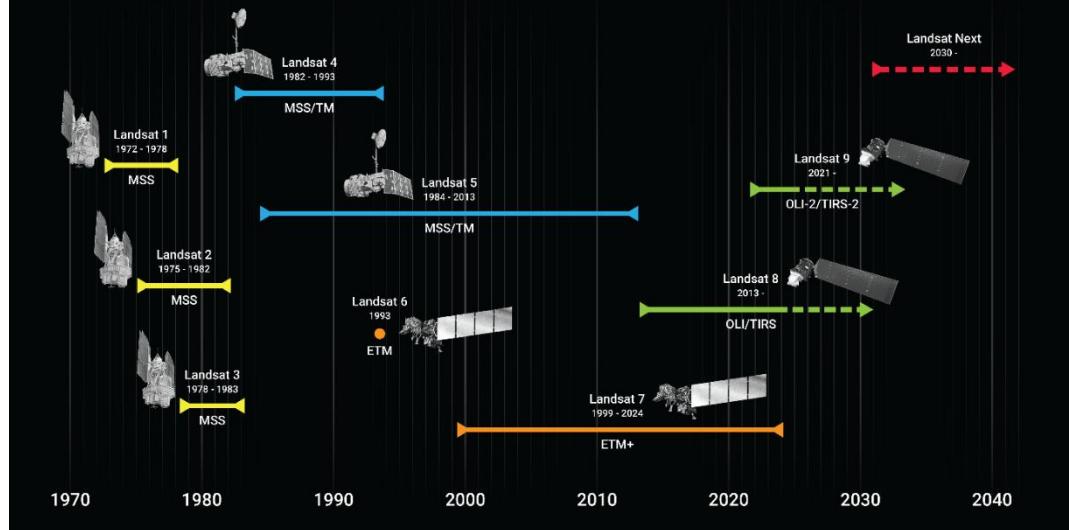
Vstupní data byla vybrána s cílem pokrýt co největší území ČR pomocí dat dostupných prostřednictvím GEE, možnost automatizované práce s daty bez nutnosti stahování celých snímků. Data pro analýzu byla dále zvolena podle schopnosti zaznamenat faktory ovlivňující šíření hluku (řešeno přes vybrané spektrální indexy a vybrané plyny), dále podle schopnosti zaznamenat aktivitu zdrojů hluku a jeho projevů (vybrané plyny, teplota či světelné znečištění).

6.1 Landsat 8 (NASA)

Landsat je program NASA/USGS¹⁸, což je americká agentura zodpovědná za civilní kosmický program. Aktuálně jsou aktivní mise Landsat 8 a Landsat 9 (viz obrázek 2). V roce 2030/2031 bude vypuštěna Landsat Next, která bude mít nejen lepší prostorové rozlišení, ale i větší spektrální pokrytí (NASA, 2024b).

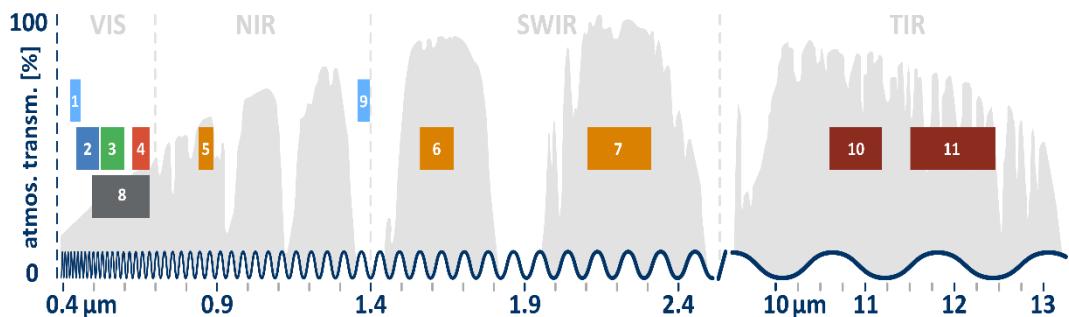
¹⁸ National Aeronautics and Space Administration/United States Geological Survey

BUILDING ON THE LANDSAT LEGACY



Obrázek 2: Mise Landsat (zdroj: <https://landsat.gsfc.nasa.gov/satellites/timeline/>)

Landsat 8 byl vypuštěn 11. 2. 2013 a poskytuje optická data (OLI¹⁹, TIRS²⁰) o globálním pokrytí pevniny. Satelit nese 2 senzory (i) **OLI** pokrývající 8 spektrálních pásem s prostorovým rozlišením **30 m** a panchromatické pásmo s prostorovým rozlišením **15 m**; (ii) **TIRS** pokrývající 2 spektrální pásmata TIR-1 (10 600–11 190 nm) a TIR-2 (11 500–12 510 nm) s prostorovým rozlišením **100 m**. K dispozici jsou úrovně Level 1, který zahrnuje geometrické a radiometrické kalibrace, a Level 2 zahrnující atmosférické korekce (NASA, 2024a). Grafické znázornění rozložení bandů je uvedeno na obrázku 3.



Obrázek 3: Landsat 8 rozložení bandů (zdroj: <https://blogs.fu-berlin.de/reseda/landsat-8/>)

Data z optického senzoru se používají například pro výpočet spektrálních indexů, jako jsou vegetační indexy, indexy zástavby a mnoho dalších. Informace

¹⁹ OLI – Operational Land Imager

²⁰ TIRS – Thermal Infrared Sensor

o indexech a způsobu jejich výpočtu lze nalézt například v Index DataBase²¹. V následujících podkapitolách jsou popsány vybrané indexy.

6.1.1 NDVI (*Normalized Difference Vegetation Index*)

NDVI je spektrální index používaný k detekci a kvantifikaci vegetace na povrchu Země. Tento index měří zelenost a pokrytí zemského povrchu vegetací zachycené na satelitním snímku. Využívá rozdílu v odrazivosti vegetace ve viditelném červeném pásmu a blízkém infračerveném pásmu (viz rovnice 2). Hodnota NDVI se pohybuje v rozmezí od -1 do 1. Hodnoty menší než 0 odpovídají vodě, hodnoty mezi 0–0,2 holé půdě či dalším nezeleným povrchem a hodnoty 0,2 do 1 značí hustou nebo zdravější vegetaci²². Podle EOS Data Analytics (2024a) záporné hodnoty vznikají hlavně v důsledku oblačnosti, vody a sněhu, hodnoty blízké nule jsou typické pro oblasti pokryté kameny a holou půdou. Zdravá vegetace silně odráží blízké infračervené záření (Near InfraRed – NIR)²³ a má nízkou odrazivost v červeném pásmu.

Rovnice 2 - Výpočet NDVI (vegetační index)

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

6.1.2 NDBI (*Normalized Difference Built-up Index*)

NDBI je spektrální index, který se používá k identifikaci zastavěných oblastí, včetně městských, průmyslových a komerčních ploch. Využívá rozdílu mezi hodnotami odrazivosti v blízkém infračerveném a krátkovlnném infračerveném pásmu (viz rovnice 3), kde zastavěné oblasti mívají vyšší odrazivost v krátkovlnném infračerveném pásmu (Short Wave InfraRed – SWIR²⁴) a nižší v pásmu NIR. Tento index pomáhá při plánování měst, sledování rozrůstání měst a analýze dopadu urbanizace na životní prostředí. NDBI se pohybuje, stejně jako NDVI v rozmezí od -1 do 1. Kladné hodnoty odpovídají zastavěným (urbanizovaným) plochám a záporné hodnoty odpovídají přírodním (přirozeným) povrchem (např. vegetace či voda) (Zha et al., 2005).

²¹ <https://www.indexdatabase.de/>

²² 0,2–0,3 – představují křoviny a louky; 0,5–0,8 mírné a tropické lesy

²³ NIR – Near-Infrared

²⁴ SWIR – Short-wave Infrared

Rovnice 3 - NDBI (index zástavby)

$$NDBI = \frac{SWIR - NIR}{SWIR + NIR}$$

6.1.3 MNDWI (*Modified Normalized Difference Water Index*)

MNDWI může být také užitečný pro identifikaci zastavěných oblastí, protože zlepšuje viditelnost rozdílů mezi vodními a zastavěnými plochami. Voda má vysokou odrazivost v zeleném pásmu a nízkou ve krátkovlnném infračerveném pásmu (viz rovnice 4), což umožňuje efektivně odlišit vodní plochy od jiných typů povrch. Rovněž nabývá hodnot od -1 do 1 , přičemž kladné hodnoty odpovídají vodním plochám a záporné hodnoty zase povrchům bez vody (vegetace, půda, zastavěné plochy) (Xiaobiao et al., 2018).

Rovnice 4 - MNDWI (modifikovaný vodní index)

$$MNDWI = \frac{Green - SWIR}{Green + SWIR}$$

6.1.4 UI (*Urban Index*)

Urban Index je dalším z nástrojů, který může být použit k detekci urbanizovaných ploch. Existuje několik variant tohoto indexu, které využívají různé kombinace spektrálních pásem (viz rovnice níže). Verze s využitím NIR, RED a GREEN (viz rovnice 5) je navržena k detekci a mapování urbanizovaných oblastí z multispektrálních dat (Kaur a Pandey, 2022). Stejně jako u NDBI kladné hodnoty odpovídají urbanizovaným plochám, naopak záporné odpovídají přírodním povrchům. UI může snížit falešné detekce vegetace nebo vody v městských oblastech.

Rovnice 5 - UI (index urbanizačních ploch)

$$UI = \frac{NIR - (Red + Green)}{NIR + (Red + Green)}$$

6.1.5 ISA (*Impervious Surface Area*) Index

ISA index se používá k odhadu nepropustných povrchů, které jsou typické pro průmyslové a komerční oblasti (viz rovnice 6). Těmito povrchy mohou být silnice, střechy a další umělé povrchy (Tian et al., 2018). ISA index se více zaměřuje na fyzikální vlastnosti povrchu. Je citlivý na povrchy, jako je beton, asfalt

čí střechy. Kladné hodnoty naznačují, že se jedná o urbanizované oblasti, a záporné hodnoty naznačují, že se jedná o vodu či vegetaci.

Rovnice 6 - *ISA (index nepropustnosti povrchů)*

$$\mathbf{ISA = NDBI - NDVI}$$

6.2 MODIS (NASA)

MODIS (*Moderate Resolution Imaging Spectroradiometer*) je klíčový snímač na palubě satelitů Terra (původně EOS²⁵ AM-1) a Aqua (původně EOS PM-1) vypuštěný agenturou NASA. Terra MODIS a Aqua MODIS sledují celý zemský povrch každé 1 až 2 dny a získávají data ve 36 spektrálních pásmech nebo skupinách vlnových délek. Tato data zlepšují chápání globální dynamiky a procesů probíhajících na pevnině, v oceánech a v nižší atmosféře (NASA, 2024c).

MODIS poskytuje data v prostorovém rozlišení 250 m pro viditelné spektrum **VIS** (620–670 nm, 841–876 nm), 500 m pro **VIS** (459–479 nm, 545–565 nm) a **NIR** (1 230–1 250 nm, 1 628–1 652 nm, 2 105–2 165 nm) a 1 000 m pro zbytek pásem, včetně **SLT**²⁶ (band 31–32). Data MODIS jsou udávána v kelvinech. MOD11 jsou produkty ze satelitu Terra a MYD11 jsou ze satelitu Aqua. Kategorie A je s rozlišením **1 km**, B je v rozlišení **6 km** a C v rozlišení **5,6 km**. Verze 1 představuje denní data, verze 2 jsou 8denní data a 3 jsou měsíční data (NASA, 2024d).

6.3 Sentinel-2 (ESA)

Satellyt Sentinel patří pod ESA²⁷, což je mezinárodní organizace pro využití vesmíru²⁸. Sentinel-2 je určen pro monitoring krajinného pokryvu a jeho změn. Mise Sentinel-2 se skládá ze 3 satelitů, a to 2A (vypuštěn 23. 6. 2015), 2B (vypuštěn 7. 3. 2017) a 2C (vypuštěn 5. 9. 2024). Poskytuje multispektrální snímky **MSI**²⁹ s časovým rozlišením 5 dní a pokryvající 13 spektrálních pásem (443–2 190 nm) s vysokým prostorovým rozlišením **10 m** pro 4 kanály (VIS, NIR), **20 m** pro 6 kanálů (RED/SWIR) a **60 m** pro 3 kanály (atmosférická korekční

²⁵ EOS – Earth Observing System

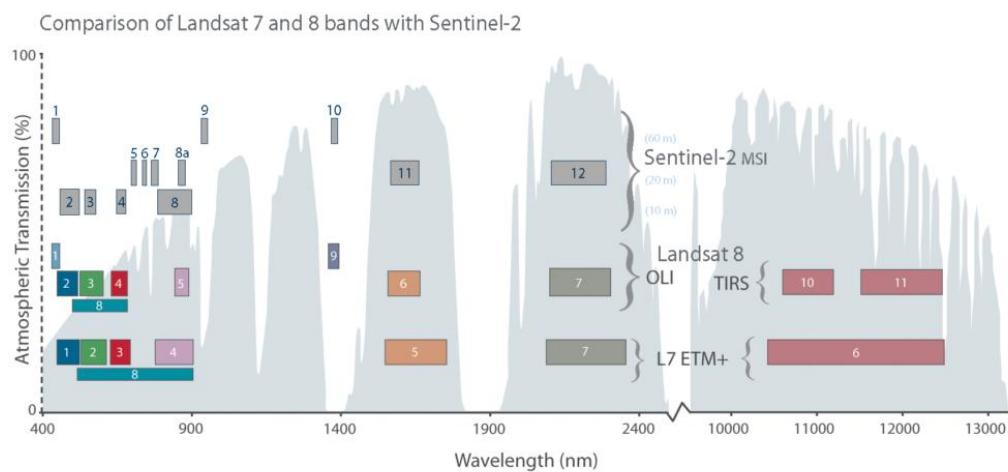
²⁶ SLT – Surface Land Temperature

²⁷ The European Space Agency

²⁸ Jejím členem je od roku 2008 i Česká republika.

²⁹ MSI – Multispectral Imager

pásma) (ESA, 2024c; ESA_b; CollGS, 2024). Na obrázku 4 je dále porovnání rozložení pásem u Landsat 7 a Landsat 8 se Sentinel-2.



Obrázek 4: Porovnání bandů u Landsat 7 a 8 se Sentinel-2 (zdroj:
<https://landsat.gsfc.nasa.gov/article/sentinel-2a-launches-our-compliments-our-complements/>)

V této práci jsou využita data ze Sentinel-2, ze kterých byly vypočteny stejné spektrální indexy jako u Landsat 8.

6.4 Sentinel-5P (ESA)

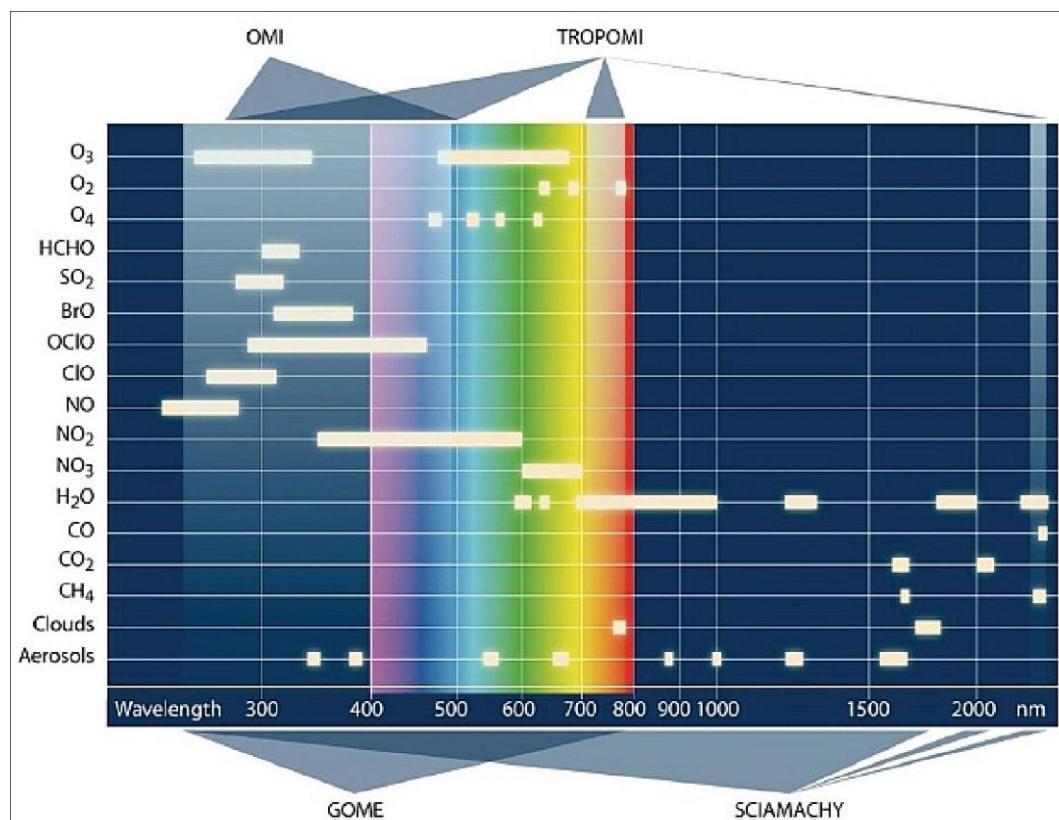
Sentinel-5P (Precursor) se zaměřuje na kvalitu ovzduší a interakci složení a podnebí s hlavními datovými produkty O_3 , NO_2 , SO_2 , $HCHO$, $CHOCHO$ a aerosoly. Kromě toho Sentinel-5P (rovněž SENTINEL-5) také obsahuje parametry kvality pro CO , CH_4 (díky SWIR pásmu) a stratosférický O_3 s denním globálním pokrytím pro klima, kvalitu ovzduší a ozónové či povrchové UV aplikace (CollGS, 2024a).

Mise Sentinel-5P se senzorem TROPOMI se skládá ze spektrometru s vysokým rozlišením v ultrafialovém až krátkovlnném infračerveném pásmu se 7 různými spektrálními pásmi: ultrafialové **UV-1³⁰** (270–300 nm), **UV-2** (300–370 nm), **VIS** (370–495 nm), **NIR-1** (675–710 nm), **NIR-2** (745–775 nm) a **SWIR** (2 305–2 385 nm). Na oběžnou dráhu byl vypuštěn 13. 10. 2017, jeho prostorové rozlišení je **7 km × 3,5 km** a časové rozlišení je 1 den. K dispozici je Level 1B, což jsou kalibrovaná radiometrická data, a Level 2, jehož data jsou odvozená z úrovni 1B

³⁰ UV – UltraViolet

a obsahují již zpracované koncentrace specifických atmosférických parametrů (ESA_a, b. r.a; CollGS, 2024).

Obrázek 5 znázorňuje, v jakých vlnových délkách jsou jednotlivé plyny detekovatelné.



Obrázek 5: Detektovatelnost plynů ze Sentinel-5P (TROPOMI) (zdroj: <https://www.eoportal.org/satellite-missions/copernicus-sentinel-5p#tropomi-tropospheric-monitoring-instrument>)

6.4.1 CO (oxid uhelnatý)

Oxid uhelnatý je jedovatý plyn vznikající nedokonalým spalováním paliv. Má krátkou životnost, ale významně ovlivňuje kvalitu ovzduší (SZÚ, 2023–2024). V přírodě se sice vyskytuje, ale jen ve velmi malém množství v atmosféře, případně se vyskytuje v souvislosti se sopečnou činností. Ve velké míře však pochází z nedokonalého spalování fosilních paliv.

6.4.2 HCHO (formaldehyd)

Formaldehyd je prekurzor vzniku troposférického ozonu, který je nejen skleníkovým plynem, ale také významným znečišťujícím činitelem ovzduší s negativními dopady na zdraví. Formaldehyd pochází zejména z lesních požárů,

vzniká během některých rozkladních a průmyslových procesů (spalovny odpadů, teplárny, tepelné elektrárny aj.) (SZÚ, 2023b).

6.4.3 NO_2 (*oxid dusičitý*)

Oxid dusičitý se tvoří při spalování paliv, především v dopravě. Způsobuje kyselé deště a přispívá k tvorbě troposférického ozonu (ESA, 2023b).

6.4.4 O_3 (*ozon*)

Ozon může mít pozitivní i negativní účinky, a to v závislosti na tom, kde se nachází. Vysoko ve stratosféře je důležitý, protože chrání život na Zemi před škodlivým ultrafialovým zářením. Avšak v troposféře je to látka znečišťující ovzduší jakožto hlavní složka městského smogu a je škodlivý pro zdraví lidí, rostlin a živočichů (ESA, 2018).

6.4.5 SO_2 (*oxid siřičitý*)

Zdrojem oxidu siřičitého je zejména antropogenní činnost, např. tepelné elektrárny, doprava, těžký průmysl. Škodí dýchacím cestám a způsobuje acidifikaci půd a vodních zdrojů. V přírodě se vyskytuje přirozeně, a to v souvislosti se sopečnou činností (ZÚ Ostrava, 2020).

6.4.6 CH_4 (*metan*)

Metan je skleníkový plyn, který má výrazný vliv na globální oteplování. Pochází zejména z ropných a plynárenských aktivit, zemědělství a rozkladu organické hmoty (ESA, 2023a).

6.5 VIIRS (EOG)

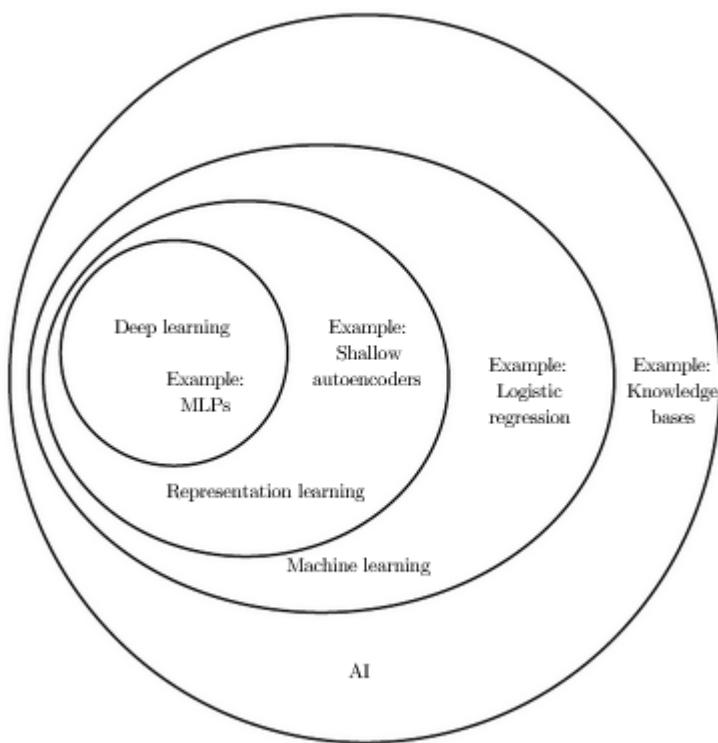
Earth Observation Group (EOG) při Payne Institute for Public Policy na Colorado School of Mines je průkopníkem v oblasti nočního dálkového průzkumu Země. Skupina se specializuje na globální pozorování nočního osvětlení a zdrojů spalování (Payne Institute for Public Policy, 2024). Družice vybavené senzorem VIIRS (*Visible Infrared Imaging Radiometer Suite*) s pásmem DNB (*Day/Night Band*) poskytují detailní data o nočním osvětlení. VIIRS je součástí satelitů Suomi NPP (*Suomi National Polar-orbiting Partnership*) a NOAA-20, které jsou součástí programu JPSS (*Joint Polar Satellite System*). Pásma DNB je citlivé na slabé světelné zdroje, což umožňuje detekci i velmi slabého osvětlení (NASA_a, b, r.a). Prostorové rozlišení VIIRS se liší podle spektrálních pásem: pásmo I mají rozlišení

375 m a pásma M 750 m. Data jsou dostupná v denních, měsíčních a ročních kompozitech, které vznikají průměrováním jednotlivých pozorování (NASA_b, b. r.a). EOG zpracovává tato data a vytváří globální kompozity nočního osvětlení, které jsou využívány pro výzkum a analýzy v oblastech, jako je monitorování urbanizace, ekonomická aktivita či environmentální změny (Payne Institute for Public Policy, 2024).

7 Metody zpracování dat

Existuje celá řada metod, které pomáhají s odhalováním souvislostí mezi daty. Mohou to být jak jednoduché statistické metody jako lineární regrese, tak pokročilejší metody za použití strojového učení jako Random Forest (RF), Gradient Boosting (GB) či neuronových sítí (NN) jako U-Net. Každá z metod je vhodná pro jiná data a má vždy své výhody i nevýhody. Výběr vhodné metody je velice důležitý a je nutné ji zhodnotit s ohledem na vstupní data.

V posledních letech se stále více využívají metody založené na umělé inteligenci (Artificial Intelligence – AI), strojové učení (*Machine Learning* – ML) nebo hluboké učení (*Deep Learning* – DL). AI se zaměřuje na napodobení kognitivních funkcí lidského myšlení, zejména prostřednictvím rozhodovacích procesů založených na datech a učení. Klíčovou součástí tohoto přístupu je strojové učení (ML), přičemž pro náročnější úlohy se využívá hluboké učení (DL) založené na umělých neuronových sítích. Na obrázku 6 se nachází Vennův diagram popisující propojení DL, ML a AI (Goodfellow et al., 2016).



Obrázek 6: Vennův diagram ukazující propojení hlubokého učení (DL), strojového učení (ML) a umělé inteligence (AI) (zdroj: <https://www.deeplearningbook.org/contents/intro.html>)

Pro predikci hluku byly využity různé přístupy, včetně metody LUR³¹, RF a neuronové sítě U-Net.

7.1 Statistická analýza

Pro tvorbu modelu ze satelitních dat je nezbytné porozumět vztahům mezi vstupními hodnotami (spektrálními indexy, koncentrací plynů apod.), které mohou ovlivnit míru hluku na vybraném území. Jelikož se jedná o kombinaci dat numerických (satelitní data) a kategorických (hluková data), bylo třeba zvolit vhodné nástroje pro práci a vyhodnocení těchto dat.

7.1.1 Lineární regrese a další regresní modely

Lineární regrese je jedna z nejpopulárnějších statistických metod (resp. algoritmus) užívaných k predikci číselných hodnot. Modeluje závislost proměnnou jako lineární kombinaci nezávislých proměnných. Výsledky této metody se snadno interpretují, její výpočet je rychlý a výpočetně nenáročný a lze ji jednoduše rozšířit o více nezávislých proměnných. Jak však název napovídá, je omezená pouze na lineární vztahy mezi proměnnými (James et al., 2021). Raess et al. (2021) použil smíšenou lineární regresi k vytvoření modelu LUR pro předpověď hladin hluku v São Paulu. Model měl dobré přizpůsobení s hodnotami $R^2 = 0,61$ pro L_{den} a $R^2 = 0,68$ pro L_n ve výcvikové sadě. Při křížové validaci (Leave-One-Site-Out Cross Validation – LOSOCV) byly hodnoty R^2 0,56 pro L_{den} a 0,63 pro L_n , což naznačuje stabilitu modelu. Chyby modelu RMSE pro L_{den} byly 3,0 dB na tréninkovém modelu a 3,8 dB u LOSOCV. Pro L_n byla hodnota RMSE 3,4 dB na tréninkovém modelu a 4,1 dB u LOSOCV.

Rovněž ve studii Liu et al. (2019) byly porovnávány modely LUR a Random Forests (RF) při odhadování hladin hluku v pěti kanadských městech. Model LUR v této studii dosáhl hodnot $R^2 = 0,47$ a RMSE = 4,99 dB. V Německu (Staab a kol., 2021) byla LUR použita pro predikci dopravního hluku (L_{den}) s vysokou přesností $R^2 = 0,702$ a střední absolutní chybou (MAE) = 4,24 dB.

7.2 Pokročilé modelovací metody

Mezi pokročilé metody modelování patří bezesporu strojové učení, jež je podmnožinou umělé inteligence (AI). Tyto algoritmy jsou trénovány k nalezení vzorů a korelací ve velkých množinách dat a k provádění nejlepších rozhodnutí

³¹ LUR – Land Use Regression

a předpovědí založených na této analýze. Hluboké učení a umělé neuronové sítě jsou komponenty strojového učení. Algoritmy strojového učení umožňují umělé inteligenci nejen zpracovávat data, ale používat je k učení a chytřejší práci, aniž by potřebovala přímé programování (Goodfellow et al., 2016). SAP (b. r.a) uvádí, že umělá inteligence je nadřazenou položkou všech podmnožin strojového učení, které jsou pod ní.

7.2.1 Random Forest

Random Forest (RF) je technika strojového učení, jež patří mezi tzv. *Ensemble Learning* algoritmy, tj. takové, které kombinují více algoritmů či modelů z důvodu dosažení vyšší přesnosti. Tato metoda slouží pro klasifikaci a regresi, přičemž kombinuje více rozhodovacích stromů do jednoho modelu. RF trénuje rozhodovací stromy nezávisle a paralelně. To by mělo zajistit lepší přesnost modelu a rovněž větší odolnost proti přetrénování modelu (Breiman, 2001). Jedná se o algoritmus, jenž je dostupný v podobě knihoven (např. *randomForest*, *ranger*, *caret*, *partykit*, *randomForestSRC*, *h2o*, *mlr3*), které umožňují ladění algoritmu prostřednictvím hyperparametrů. Jejich správné nastavení je stežejní pro model a ovlivňuje rychlosť tréninku, celkovou přesnost i schopnost modelu generalizovat nová data. Mezi hyperparametry patří *ntree* (počet stromů), *max_depth* (maximální hloubka stromu), *mtry* (počet náhodně vybraných proměnných), *min_samples_leaf* (minimální počet pozorování, jež musí být v koncových uzlech stromu) a *sample_size* (procento trénovacích dat použitých k trénování každého stromu) (Cha et al., 2021; Baladram, 2024).

Random Forest metoda byla využita například v Kanadě (Liu et al., 2019), kde model RF vykazoval lepší výkonnost než u LUR ($R^2 = 0,47$ a RMSE = 4,99 dB), a to hodnotou $R^2 = 0,58$ a RMSE = 4,44 dB. Tato zjištění naznačují, že model RF poskytuje přesnější odhadы hladin hluku než tradiční LUR model.

7.2.2 Gradient Boosting

Gradient Boosting (GB) je technika strojového učení, jež také patří mezi *Ensemble Learning* algoritmy, která vyhledává i nelineární vztahy mezi proměnnými. Tento algoritmus se umí vypořádat s chybějícími či odlehlymi hodnotami a umí pracovat i s kategorickými daty. Jedná se o algoritmus, jenž je dostupný v podobě knihoven, které umožňují ladění modelu prostřednictvím hyperparametrů. Stejně jako u RF je jejich nastavení pro model velice důležité. Mezi hyperparametry patří *interaction.depth* (hloubka stromu), *n.minobsinnode* (minimální počet pozorování

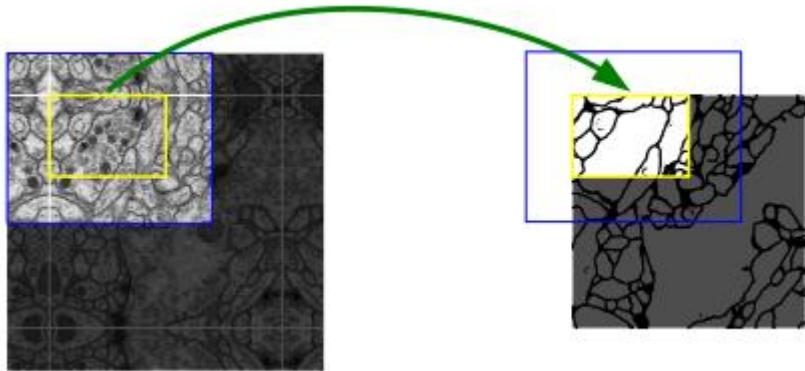
v každém listovém uzlu), *n.trees* (počet stromů v modelu), *shrinkage* (rychlosť učení a jak moc se váhy modelu aktualizují) a další (Cha et al., 2021; Baladram, 2024).

Podstatou algoritmu je tvorba několika slabých modelů (často rozhodovacích stromů) a jejich následná kombinace pro získání lepšího výkonu modelu. GB trénuje modely sekvenčně (na rozdíl od RF, který trénuje stromy paralelně) a nový model obsahuje opravy chyb předešlých modelů. Každým krokem se minimalizuje chybová funkce. Algoritmus využívá gradient této funkce, jenž zlepší jeho predikci. Výsledný model je součtem výstupů jednotlivých slabých modelů (Cha et al., 2021; Baladram, 2024).

7.2.3 Neuronové sítě

Neuronové sítě jsou pokročilé algoritmy strojového učení, jež jsou inspirovány biologickými neuronovými sítěmi v lidském mozku. Ty se skládají z vrstev neuronů, které jsou propojené pomocí vážených propojení synaptickými vazbami. Algoritmus pracuje se vstupní vrstvou (vstupy z dat), skrytou vrstvou (mezi vstupní a výstupní vrstvou, kde probíhají výpočty) a výstupní vrstvou (výstup modelu). Obdobně jako u RF či GB je algoritmus dostupný v podobě knihoven a je zde velice důležité nastavení hyperparametrů. Mezi hyperparametry patří *hidden_layer* (počet skrytých vrstev), *units/nodes* (počet neuronů v každé vrstvě), *activation* (funkce, jež přidává nelinearitu do neuronů), *learning_rate* (velikost kroku optimalizačního algoritmu), *epochs* (počet průchodů trénovací sadou) a další.

Jedním z představitelů modelu založených na neuronových sítích je U-Net, který byl představen v roce 2015 jako konvoluční neuronová síť určený pro segmentaci obrazu využitelného v medicíně, například pro segmentaci buněk. Cílem modelu U-Net je lokalizace a segmentace objektů na úrovni pixelu, tj. rozdelení snímku na různé části, kde každý pixel patří do jedné třídy (viz obrázek 7) (Ronneberger et al., 2015).



Obrázek 7: Překryvání dlaždic pro segmentaci velkých obrazů. Žlutá oblast je predikce segmentace vyžadující obrazový vstup v modré oblasti. (Zdroj: Ronneberger et al., 2015)

Tento model se skládá ze dvou částí, a to Encoder a Decoder. Encoder zpracovává vstupní snímek získáním důležitých vzorů ve snímku (konvoluce) a postupně snižuje rozlišení (pooling). Decoder pak rekonstruuje původní velikost snímku a zachycuje a kombinuje rysy v obrazu (Ronneberger et al., 2015).

METODIKA

8 Vstupní data

Pro tuto práci byly využity dvě skupiny dat, a to hluková data a satelitní data. Hluková data pochází od Ministerstva zdravotnictví ČR, a to strategické hlukové mapy ze 4. kola³² hlukového mapování pro celou Českou republiku. Jako vstupní satelitní data byla použita data Landsat, MODIS, Sentinel a dále také VIIRS. Přehled prostorového rozlišení vstupních dat je uveden v tabulce 3.

Tabulka 3: Prostorové rozlišení vstupních dat

Zdroj dat	Typ dat	Prostorové rozlišení	Jednotky
HLUKOVÁ DATA			
MZČR	Hluková data	10 m	dB
SATELITNÍ DATA			
Landsat 8	OLI-SR/T1_L2	30 m	
	TIRS/T1_L2	100 m	K
MODIS	Terra-MOD11	1 000 m	K
Sentinel-2	MSI-SR_H		
	VIS, NIR, RED	10 m	
	SWIR	20 m	
Sentinel-5P	TROPOMI	7 km x 3,5 km	
	CO	1 113,2 m	mol/m ²
	HCHO	1 113,2 m	mol/m ²
	NO ₂	1 113,2 m	mol/m ²
	SO ₂	1 113,2 m	mol/m ²
	CH ₄	1 113,2 m	mol fraction ³³
	O ₃	1 113,2 m	mol/m ²
EOG	VIIRS-VNL	500 m	nW/cm ² /sr ³⁴

8.1 Hluková data

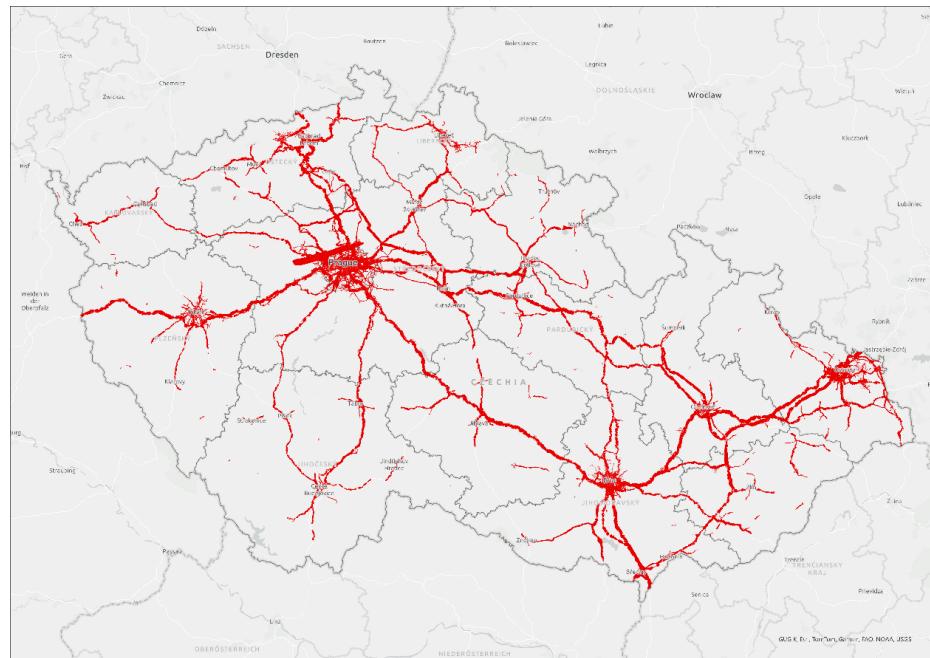
Hluková data pochází od MZČR a jsou ze SHM 2022. Stažený soubor dat byl rozdělen do několika shapefilů, a to aglomerace celek a dále pak podle zdroje

³² Staženo s geoportálu MZČR.

³³ Molární zlomek – bezrozměrná veličina, která udává poměr látkového množství jedné složky k celkovému látkovému množství všech složek směsi.

³⁴ Míra plošné radiance vyzařované ze Země.

hluku (letiště³⁵, průmysl, silnice, železnice). Pro jednotlivé zdroje hluku byla k dispozici kategorická data hlukového indikátoru L_{den} . Vstupní data měla podobu polygonů odpovídajících jednotlivým hlukovým pásmům. Tato vrstva byla klasifikována do 5dB intervalů (DB_Low, DB_High) na základě původní sítě výpočtových bodů, která měla rozestup 10 m a výpočtovou výšku 4 m nad terénem. Nejnižší kategorie hluku byla stanovena na 50 dB³⁶, zatímco nejvyšší kategorie zahrnovala hodnoty nad 75 dB. Zdravotní ústav se sídlem v Ostravě (2022) uvádí, že do hlukových map pro jednotlivé aglomerace jsou zahrnuta data ze silniční, tramvajové, železniční a letecké dopravy a průmyslových zdrojů. Hluková data pokrývají přibližně jen 7,2 % území ČR (viz obrázek 8), přičemž největší podíl hlukových dat připadá na hlavní silnice, a to 4,2 % (aglomerace a hlavní železniční tratě zaujmají podobný podíl, přibližně 1,4 % z celkové plochy ČR).



Obrázek 8: Pokrytí SHM 2022 na území České republiky (zdroj: vlastní zpracování v ArcGIS Pro)

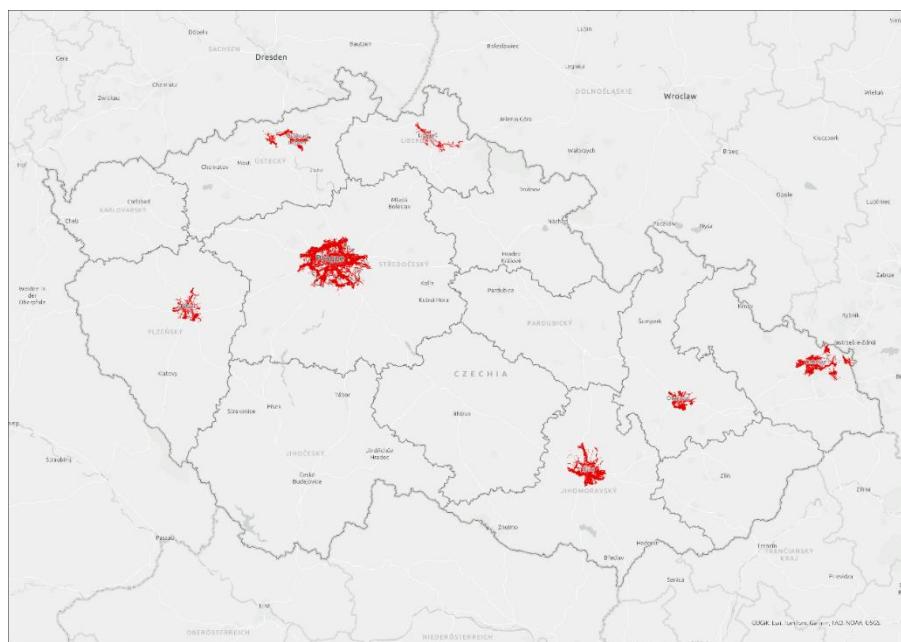
Všechna hluková data byla získána v souřadnicovém systému S-JTSK Krovak EastNorth (EPSG: 5514).

³⁵ Vrstva Letiště nebyla použita, neboť obsahuje pouze Letiště Václava Havla Praha, které nezasahuje do vybrané oblasti.

³⁶ Nejnižší hodnota (DB_Low) pro ukazatel L_{den} .

8.1.1 Aglomerace

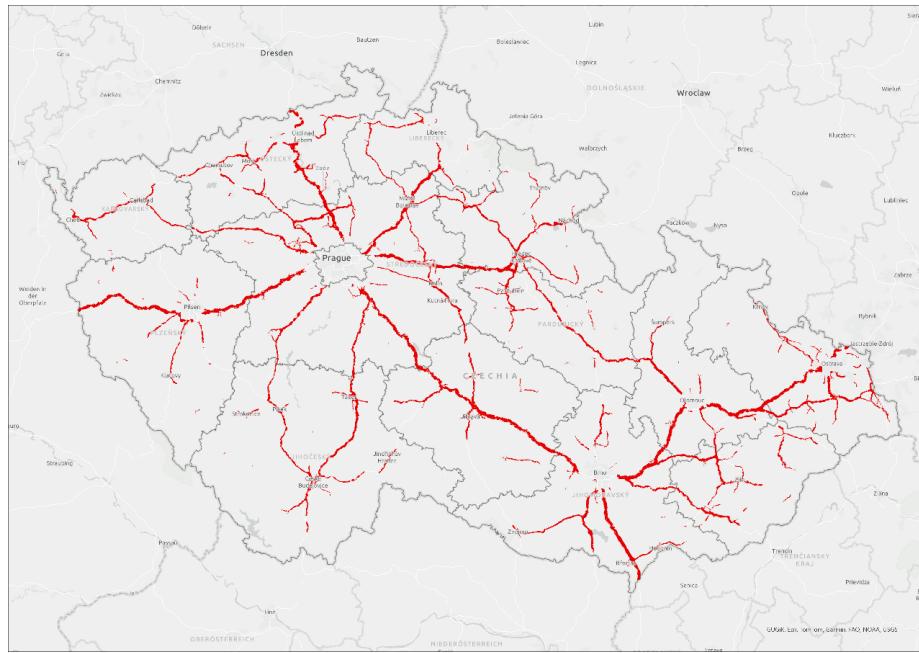
Data za vybrané aglomerace pokrývají 1 121,72 km² území České republiky, což odpovídá necelým 1,4 % celkové rozlohy (viz obrázek 9). Do aglomerací spadají města Ústí nad Labem a Teplice, Liberec a Jablonec nad Nisou, Praha, Plzeň, Brno, Olomouc, Ostrava s Bohumínem, Karvinou a Havířovem.



Obrázek 9: Pokrytí SHM 2022 – Aglomerace na území České republiky (zdroj: vlastní zpracování v ArcGIS Pro).

8.1.2 Silnice

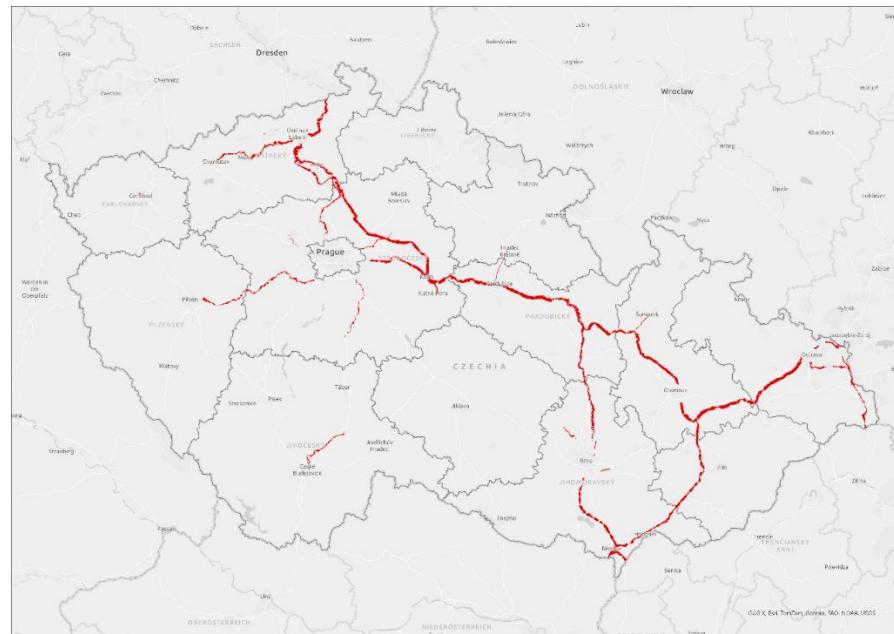
Data hlavních silničních tahů pokrývají 3 334,81 km² území České republiky, což odpovídá necelým 4,2 % (viz obrázek 10).



Obrázek 10: Pokrytí SHM 2022 – Silnice na území České republiky (zdroj: vlastní zpracování v ArcGIS Pro)

8.1.3 Železnice

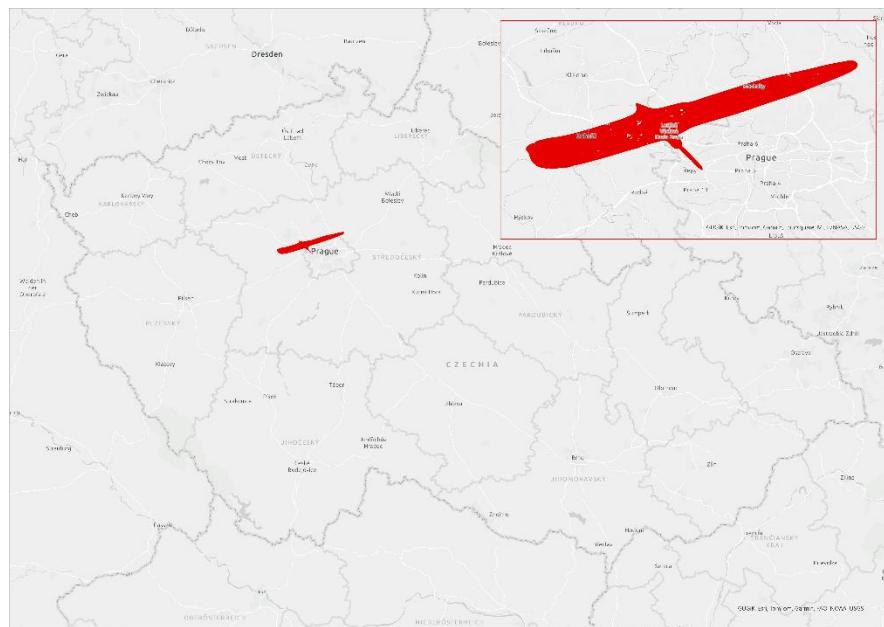
Data hlavních železničních tratí pokrývají $1\ 145,41\text{ km}^2$ území České republiky, což odpovídá necelým 1,5 % (viz obrázek 11).



Obrázek 11: Pokrytí SHM 2022 – Železnice na území České republiky (zdroj: vlastní zpracování v ArcGIS Pro)

8.1.4 Letiště

Data týkající se hlavního letiště pokrývají 125,29 km² území České republiky, což odpovídá necelým 0,2 % (viz obrázek 12).



Obrázek 12: Pokrytí SHM 2022 – Letiště na území České republiky (zdroj: vlastní zpracování v ArcGIS Pro)

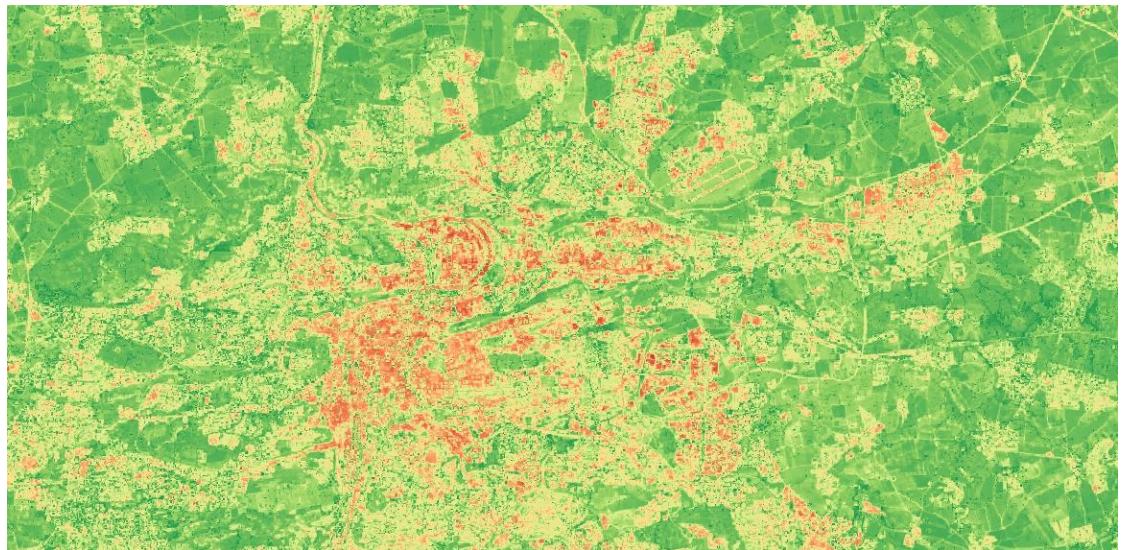
8.2 Data z dálkového průzkumu Země

Výběr vstupních dat vychází z faktorů ovlivňujících šíření hluku. Materiály vyskytující se v prostředí jsou reprezentovány NDVI, NDBI, MNDWI, UI a ISA indexy. Další faktory jsou reprezentovány teplotou povrchu, jež může být projevem některého z hlavních zdrojů hluku. Neméně důležitými faktory jsou koncentrace vybraných atmosférických plynů CO, HCHO, NO₂, O₃, SO₂ a CH₄ či SLT, jejichž výskyt souvisí s dopravou i průmyslem obdobně jako světelné znečištění.

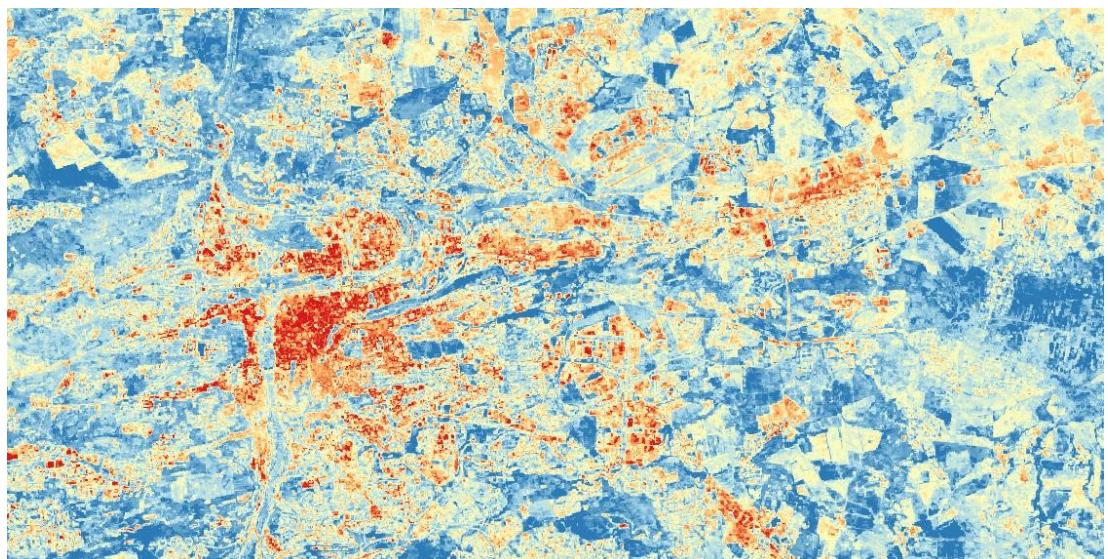
8.2.1 Landsat 8

Data byla použita ve formě vypočtených hodnot spektrálních indexů pro rozlišení a kategorizaci zástavby území, vegetace a průmyslu, a to NDVI, NDBI, MNDWI, UI a ISA indexy pro vybrané časové období (1. 1. 2022 – 31. 12. 2022). Data byla získána v souřadnicovém systému WGS84/UTM, zóna 33N (EPSG: 32633)³⁷. Na obrázcích 13–18 jsou ukázky zobrazení jednotlivých indexů vytvořené v prostředí Google Colab za použití Google Earth Engine (viz příloha 2).

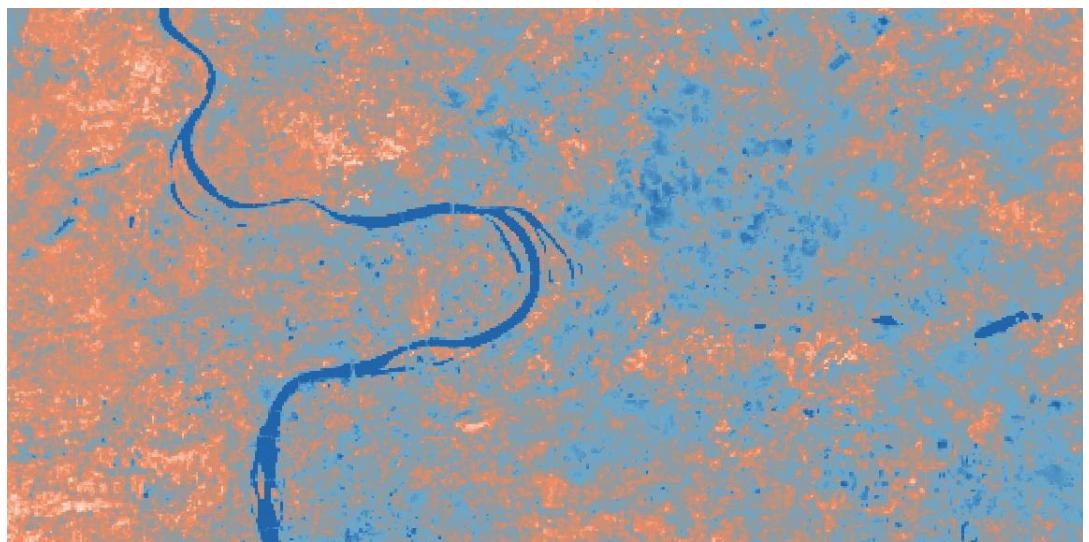
³⁷ <https://www.usgs.gov/media/files/landsat-8-data-users-handbook>



Obrázek 13: Výstup z GEE pro vypočtený index NDVI Max, oblast Praha (zdroj: vlastní zpracování v Google Colab)



Obrázek 14: Výstup z GEE pro vypočtený index NDBI, oblast Praha (zdroj: vlastní zpracování v Google Colab).



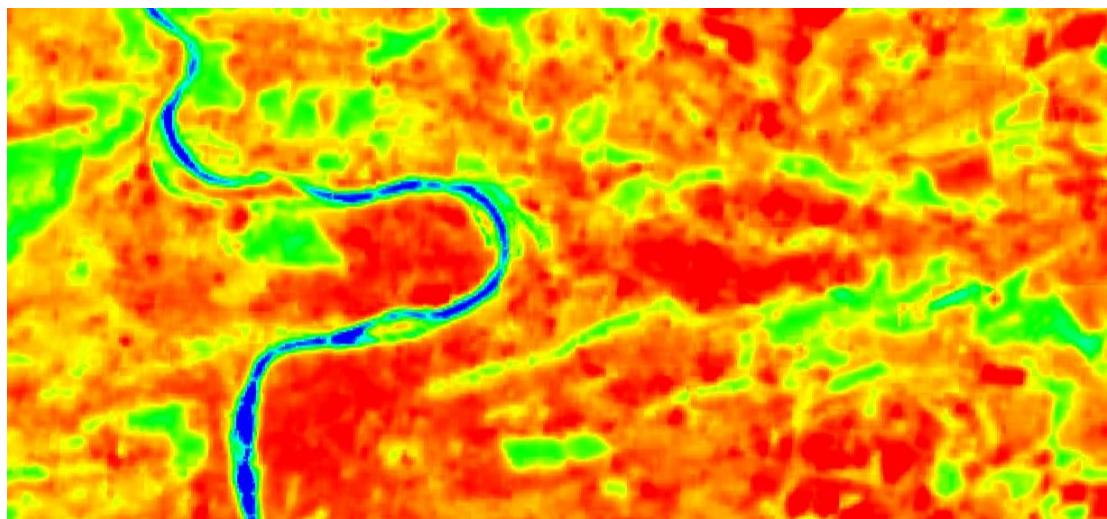
Obrázek 15: Výstup z GEE pro vypočtený index MNDWI, oblast Praha (zdroj: vlastní zpracování v Colab).



Obrázek 16: Výstup z GEE pro vypočtený index UI, oblast Praha (zdroj: vlastní zpracování v Colab).



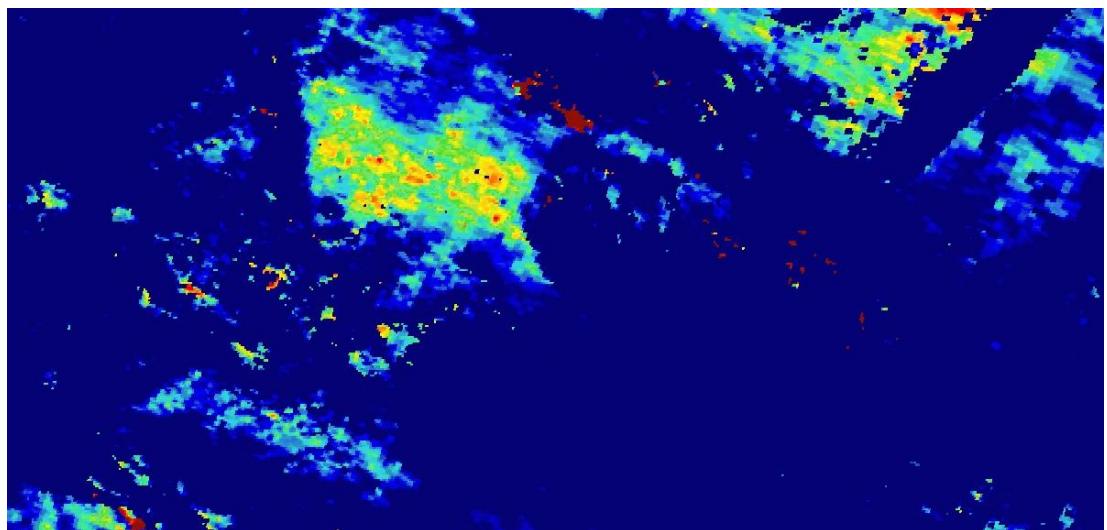
Obrázek 17: Výstup z GEE pro vypočtený index ISA, oblast Praha (zdroj: vlastní zpracování v Colab).



Obrázek 18: Výstup z GEE pro data TIR v kelvinech, oblast Praha (zdroj: vlastní zpracování v Colab).

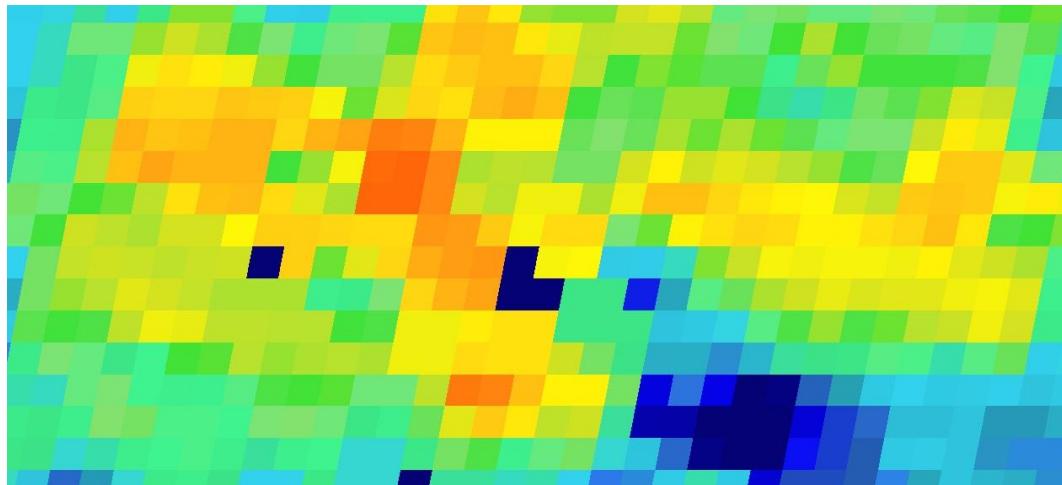
8.2.2 MODIS

Denní data MODIS A1 z Terra (viz obrázek 19 a 20) byla stažena v kelvinech jako maximum z denních hodnot v celočíselné sinusové projekci (SR-ORG:6974)³⁸ a byla získána v prostředí Google Colab za použití Google Earth Engine (viz příloha 2).



Obrázek 19: Výstup z GEE pro data z MODIS, oblast ČR (zdroj: vlastní zpracování v Colab).

³⁸ <https://icess.eri.ucsb.edu/modis/LstUsrGuide/usrguide.html>



Obrázek 20: Výstup z GEE pro data z MODIS, oblast Praha (zdroj: vlastní zpracování v Colab).

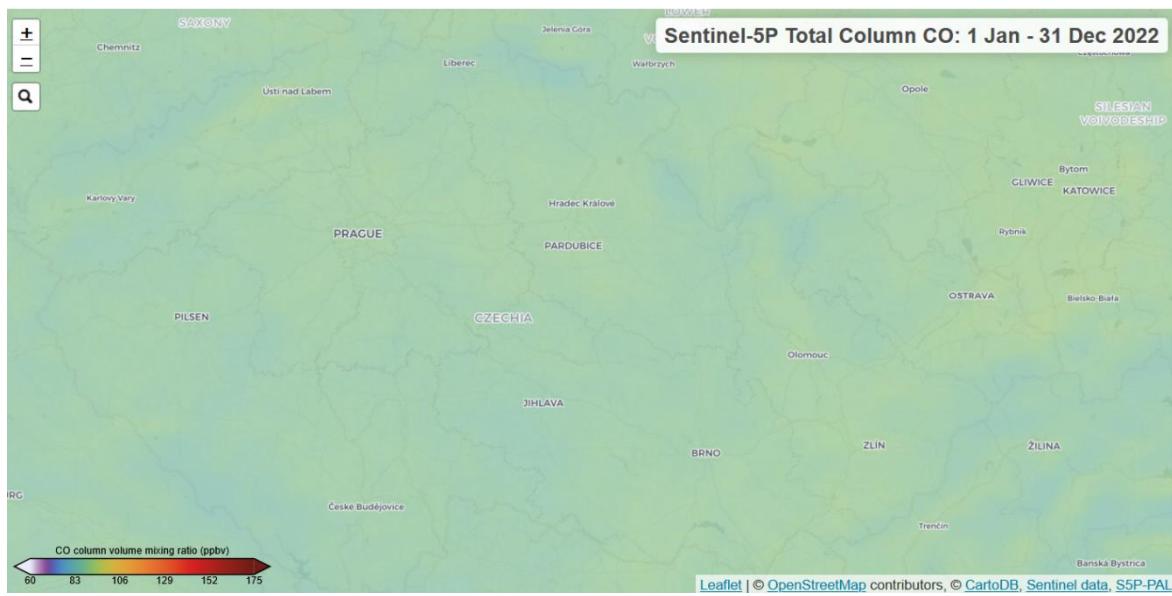
8.2.3 Sentinel-2

Data ze Sentinel-2 byla stažena ve formě stejných spektrálních indexů jako u Landsat 8, a to v souřadnicovém systému WGS84/UTM, zóna 33N (EPSG: 32633) a byla získána prostřednictvím skriptu za použití Google Earth Engine (viz příloha 2).

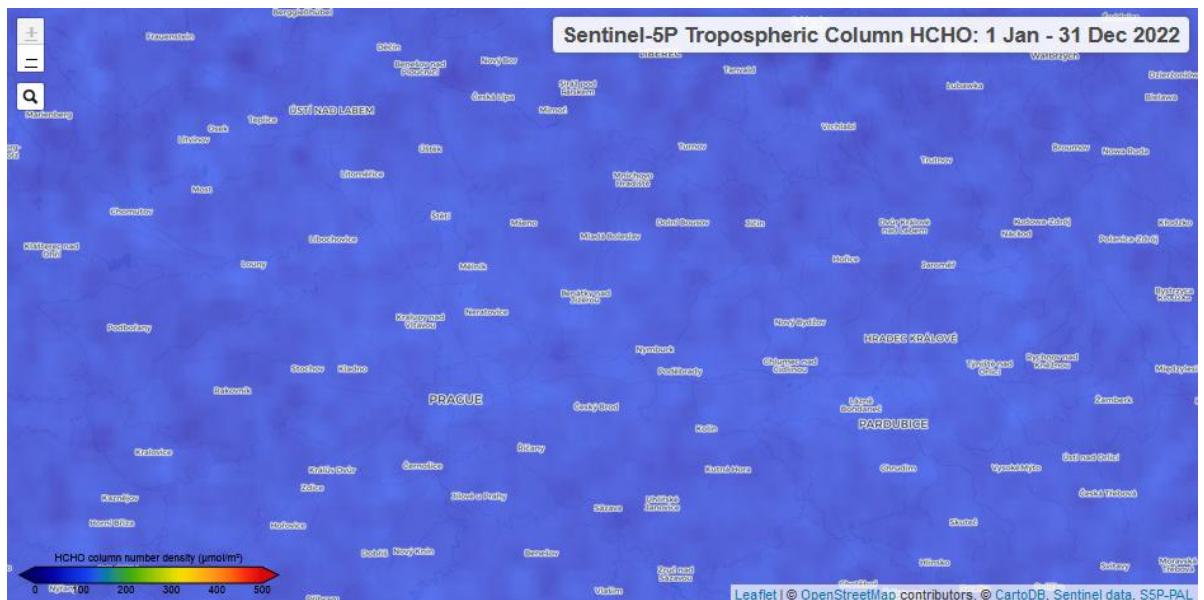
8.2.4 Sentinel-5P

Byla použita data pro různé atmosférické plyny, a to CO, HCHO, NO₂, O₃, SO₂ a CH₄ za vybrané časové období. Výsledná data jsou průměry za vybrané období. Data byla spočítána pro každý plyn zvlášť a poté byly snímky zkombinovány do jednoho vícepásmového snímku. Na obrázcích 21–26 jsou ukázky zobrazení jednotlivých plynů.

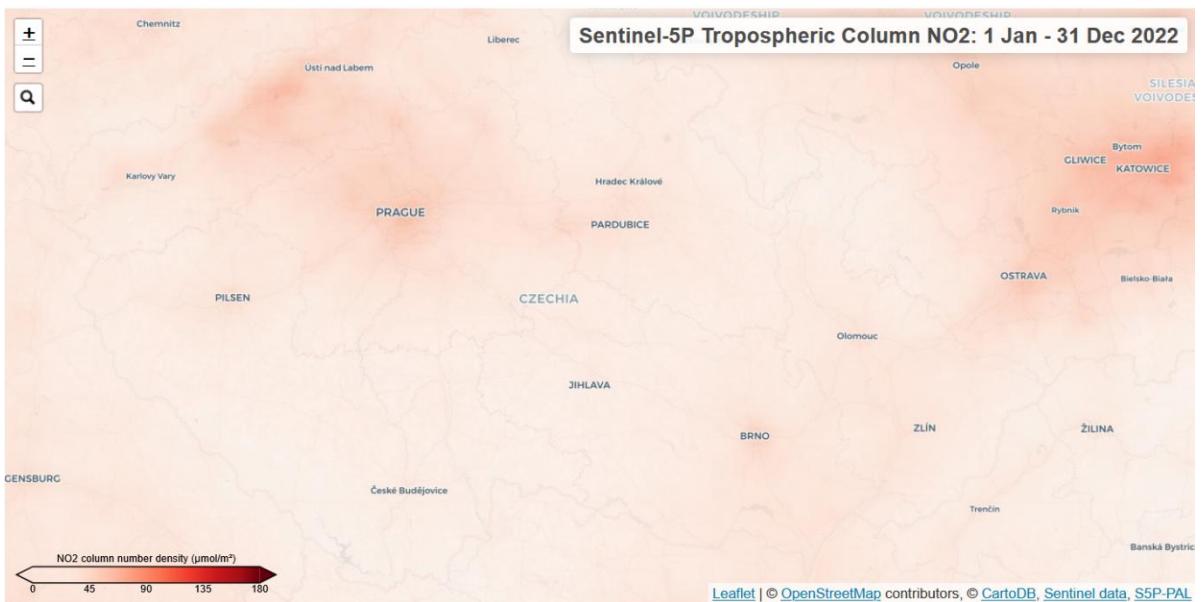
Data ze Sentinel-5P byla získána v souřadnicovém systému WGS84 (EPSG: 4326) a byla získána prostřednictvím skriptu za použití Google Earth Engine (viz příloha 2).



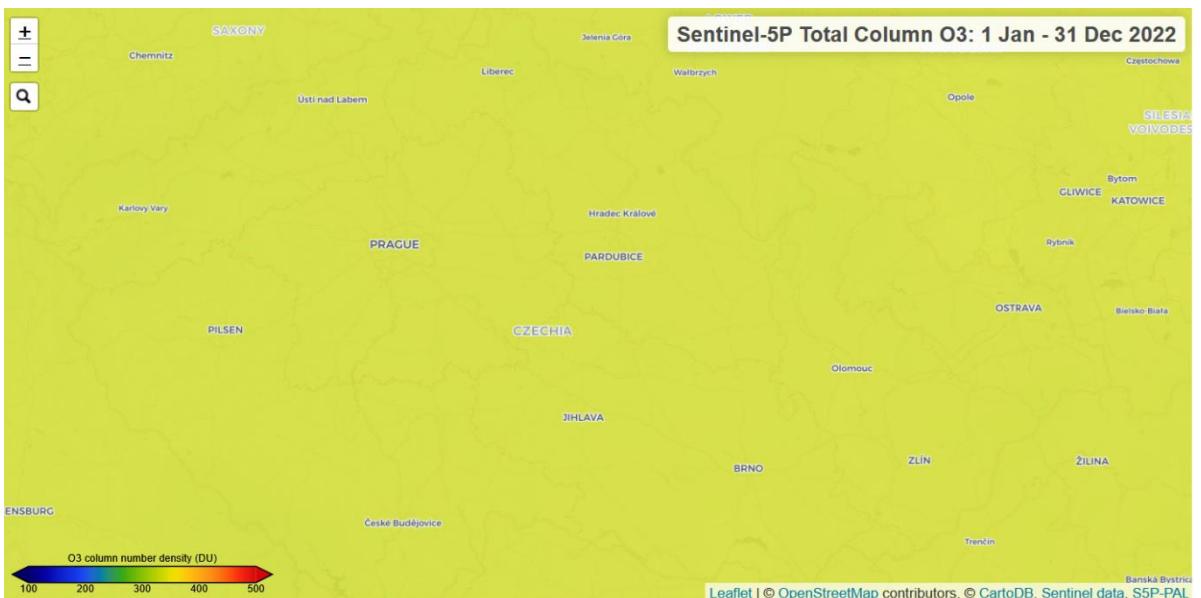
Obrázek 21: Zobrazení plynu CO na území ČR za rok 2022 (Zdroj: <https://maps.s5p-pal.com/co/year/>).



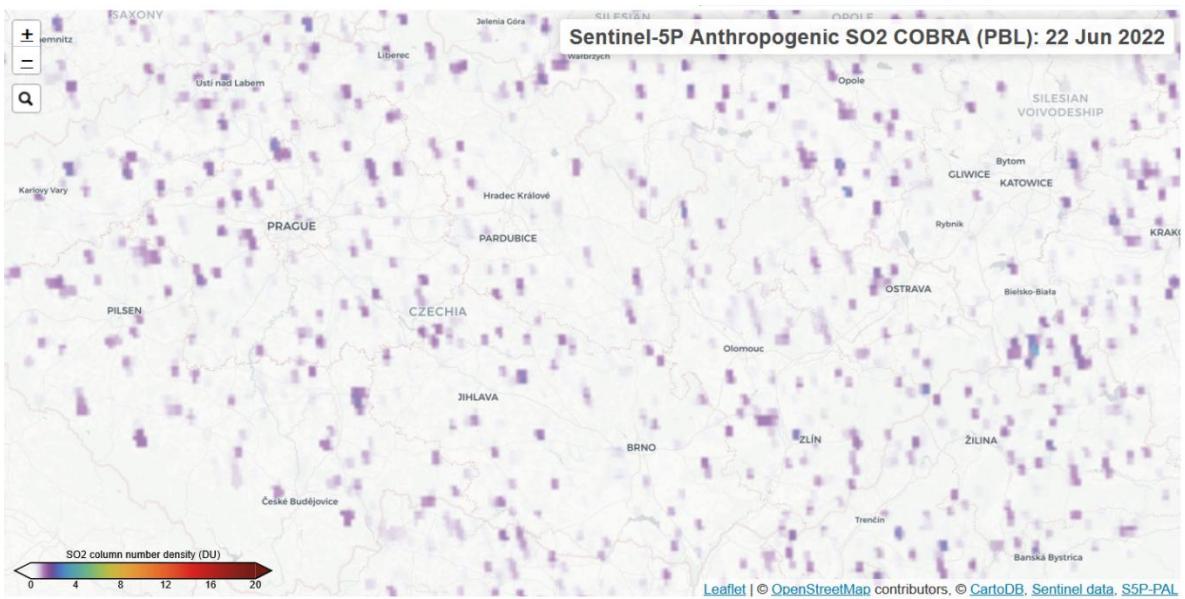
Obrázek 22: Zobrazení plynu HCHO na území ČR za rok 2022 (zdroj: <https://maps.s5p-pal.com/hcho/year/>).



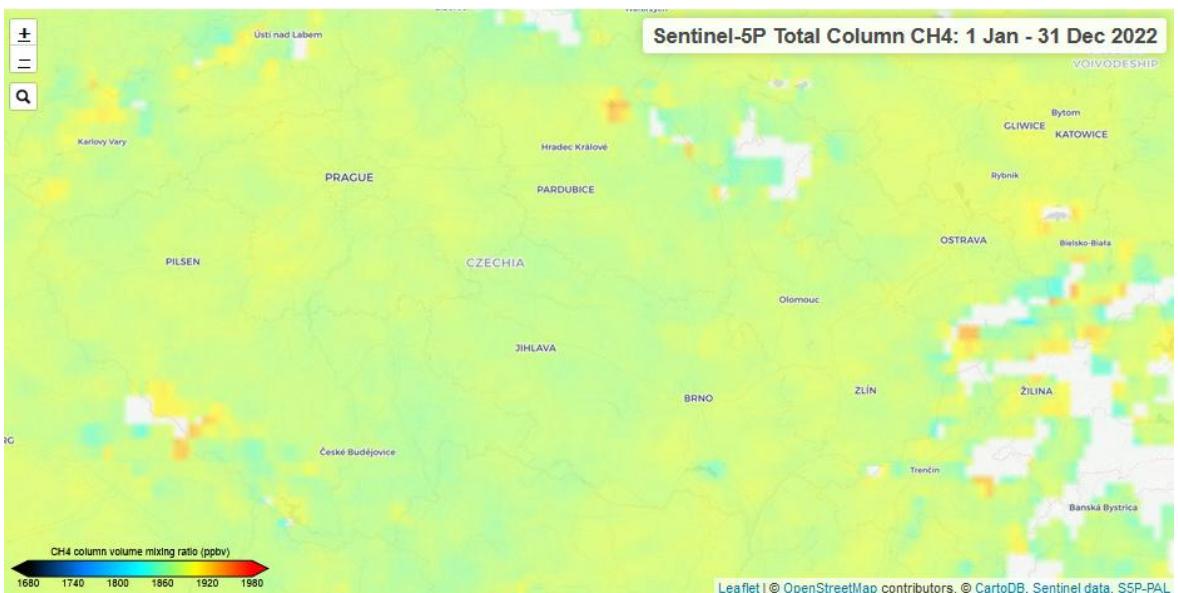
Obrázek 23: Zobrazení plynu NO₂ na území ČR za rok 2022 (zdroj: <https://maps.s5p-pal.com/no2/year/>).



Obrázek 24: Zobrazení plynu O₃ na území ČR za rok 2022 (zdroj: <https://maps.s5p-pal.com/o3/year/>).



Obrázek 25: Zobrazení plynu SO₂ na území ČR z 22. 6. 2022 (zdroj: <https://maps.s5p-pal.com/so2-cobra-pbl/day/>).



Obrázek 26: Zobrazení plynu CH₄ na území ČR za celý rok 2022 (zdroj: <https://maps.s5p-pal.com/ch4/>).

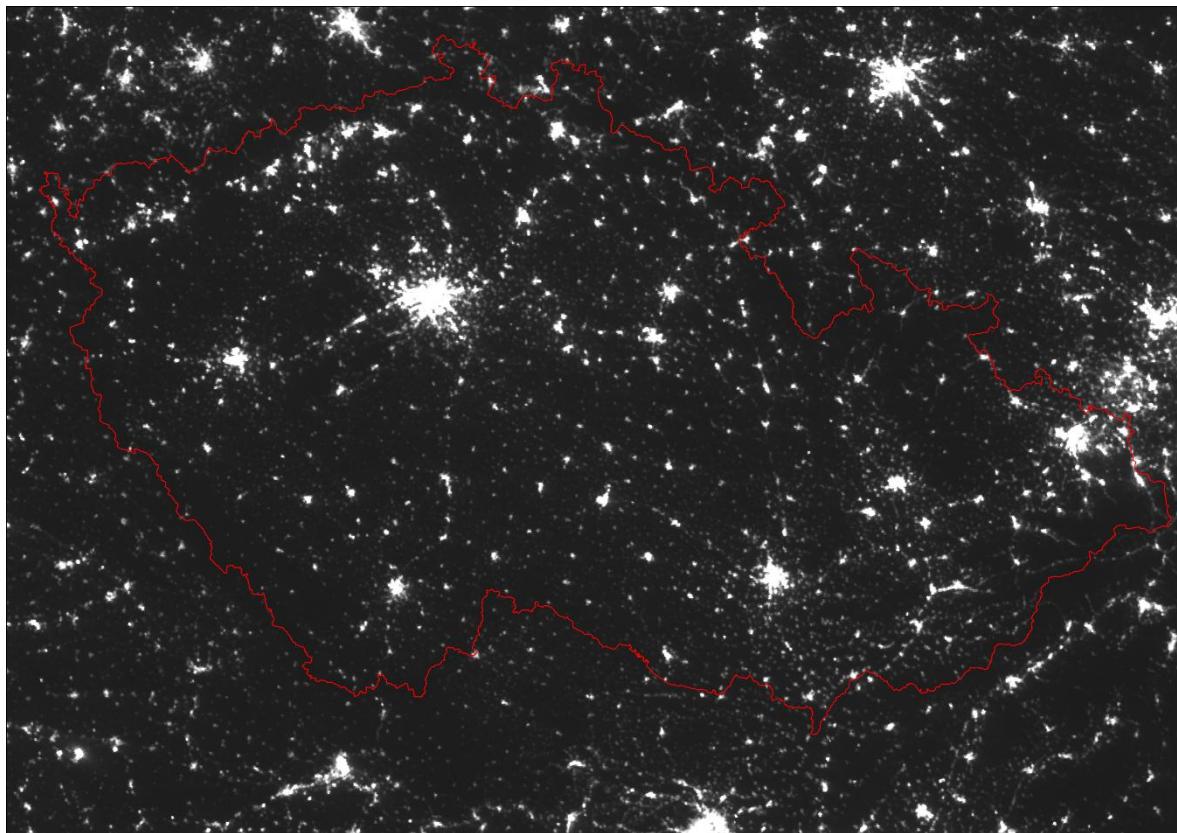
8.2.5 VIIRS

Data o světelném znečištění byla použita v celoroční mozaice od EOGdata, VNL verze v2.2, konkrétně **vcmslcfg**³⁹ (viz obrázek 27). Jedná se o konzistentně zpracovanou časovou řadu ročních globálních nočních světel VIIRS (*Visible Infrared Imaging Radiometer Suite*) za rok 2022. Kompozit byl vytvořen z měsíčních bezoblačných průměrů, které byly následně spojeny do hrubého

³⁹ Data korigovaná rozptylem světla, mají větší pokrytí dat směrem k pólům, ale mají sníženou kvalitu.

ročního kompozitu. U dat došlo k filtrování a odstranění slunečních a měsíčních paprsků a pixelů pokrytých mraky. Pro měsíční data byla použita měsíční mozaika od EOGdata, SVDNB, verze v10, rovněž vcmslcfg, pro měsíce únor, květen, srpen a listopad (EOG, 2021).

Veškerá data byla získána v souřadnicovém systému WGS84 (EPSG: 4326) a byla stažena prostřednictvím skriptu za použití Google Earth Engine (viz příloha 2).

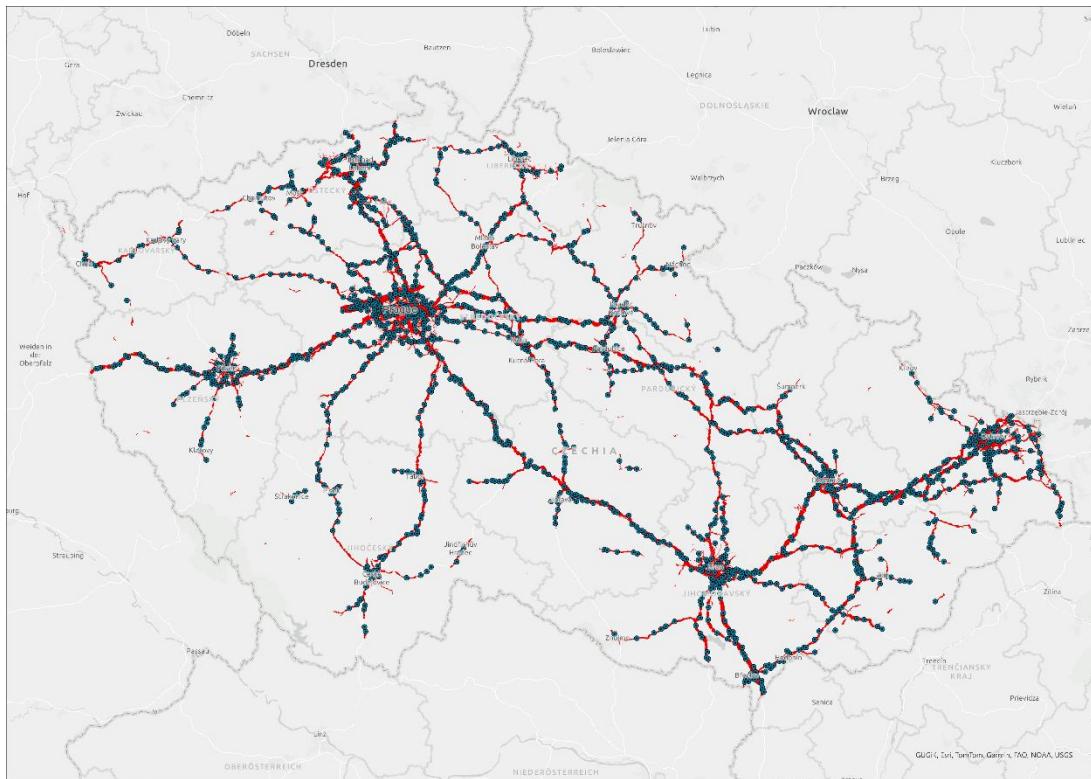


Obrázek 27: Zobrazení světelného znečištění z VIIRS (celoroční mozaika) na území České republiky (zdroj: vlastní zpracování v ArcGIS Pro).

8.3 Zpracování dat

Z důvodu různého prostorového rozlišení vstupních dat byla zvolena extrakce hodnot pro předem definované body. S ohledem na to, že kromě hlukových dat lze všechna data získat na území celé České republiky, byla zvolena možnost vytvoření náhodných bodů na území pokrytého hlukovými mapami. Bylo náhodně vytvořeno 1 400 bodů (*SHM_Body_1000m_SJ*), jak je znázorněno na obrázku 28. Vzdálenost mezi body byla stanovena v závislosti na různorodosti prostorového rozlišení vstupních dat (10–1 000 m) na 1 000 m. Počet bodů vychází z původního předpokladu získat data ze 14 krajů po 100 bodech. Jelikož byl pro práci s daty využit nástroj Google Earth Engine, nebylo nutné převádět data na společný

souřadnicový systém, jelikož tuto operaci již provedl nástroj od Google. Vektorová data byla převedena z EPSG: 5514 do EPSG: 4326.



Obrázek 28: Pokrytí území České republiky náhodnými body (zdroj: vlastní zpracování v ArcGIS Pro).

Na začátku analýzy bylo stanoveno období pro stažení dat od 1. 1. 2022 do 31. 12. 2022. Původně bylo zvažováno omezení pouze na období mimo vegetační útlumu, avšak pro větší variabilitu dat bylo nakonec zahrnuto období celého roku.

S ohledem na skutečnost, že hluková data jsou kategorická a roční průměry (případně maximální hodnoty) by mohly vést ke zkreslení dat, byla kromě celoročních dat získána rovněž měsíční data reprezentující jednotlivá roční období. Pro zajištění sezónní variability byla data shromážděna za následující období: 1. 2. 2022 – 28. 2. 2022, 1. 5. 2022 – 31. 5. 2022, 1. 8. 2022 – 31. 8. 2022 a 1. 11. 2022 – 30. 11. 2022.

Indexy byly spočítány pro každý satelitní snímek (Landsat 8 a Sentinel-2) a následně došlo ke sloučení do jedné mozaiky. Z vypočtených indexů byl následně vypočten medián za sledované období a v případě NDVI také maximální hodnota. Z důvodu výpočetní náročnosti byla Česká republika rozdělena na jednotlivé dlaždice.

Z Landsat 8 byla rovněž získána data o teplotě z bandu TIR (ST_B10), který však má hodnoty v rozmezí 1–65 535 s hodnotou scale 0,00341802 a offset 149. Původní hodnoty ve formátu Digital Number byly převedeny na hodnoty v kelvinech pomocí rovnice 7.

Rovnice 7 - Převod Digital Number na hodnoty v kelvinech

$$T_{\text{Kelvin}} = DN \times \text{scale} + \text{offset}$$

Jak již bylo uvedeno, pro získání satelitních dat byl použit skript v jazyce Python (viz příloha 2), a to prostřednictvím webového nástroje [Colab](#). Pro stažení a práci s daty byly nezbytné balíčky⁴⁰ os, geemap, earthengine-api (ee), openeo, geopandas, pandas, rasterio, ijson, gdal, numpy, shutil, shapely a time. Získání dat a jejich zpracování vyžadovalo připojení na externí služby, a to Google Drive (nahrány všechny vstupní vrstvy oblasti), Google Earth Engine (stažení dat z Landsat 8, MODIS, Sentinel-2, Sentinel-5P, VIIRS) a také přístup k externím geodatovým službám, například ke stažení vektorových vrstev a satelitních snímků.

Jelikož oblast zájmu, Česká republika, byla příliš velká, bylo nutné oblast rozdělit na menší části (dlaždice o velikosti 32 × 32 pixelů), a to z důvodu snazšího zpracování dat a menší výpočetní náročnosti.⁴¹ Satelitní data byla načtena pro vybrané období a oblast a pro každý zdroj dat byla vytvořena mozaika pro sjednocení nejlepších hodnot pixelů. Výsledky byly následně vzorkovány pro předem vygenerované body (SHM_Body1400). Výstup byl exportován do CSV souborů pro další zpracování v jazyce R.

⁴⁰ Geemap, eartengine-api a openeo – načítání a zpracování satelitních dat; geopandas – práce s geoprostorovými daty, pandas – manipulace s tabulkovými daty; time – práce s daty a časem; sentinelhub – stahování dat ze SentinelHub.

⁴¹ Bylo nutné zvýšit kapacitu GPU počítače, na kterém probíhaly výpočty, a tedy došlo na nákup výpočetních jednotek v rámci Google Colaboratory.

VÝSTUPY

9 Analýza a tvorba modelu

Pro analýzu a tvorbu modelu byl použit programovací jazyk R, konkrétně knihovny *dplyr*, *pkgconfig*, *data.table*, *ggplot2*, *randomForest*, *caret*, *smotefamily*, *patchwork*, *purrr*, *corrplot*, *car*, *ranger*, *gbm* a *neuralnet*. Analýza byla prováděna v prostředí RStudio. Kompletní skript je uveden v příloze 3.

9.1 Příprava dat

Všechna vstupní data (SHM2022, spektrální indexy z Landsat 8 a Sentinel-2⁴², termální data z MODIS, koncentrace plynů ze Sentinel-5P⁴³ a data o světelném znečištění) byla načtena z CSV souborů. Kapitola 7 podrobně popisuje zdroje vstupních dat a jejich předzpracování. Všechna data byla sloučena do jedné tabulky (*inputdata_all*) podle identifikačního sloupce ID (odpovídající bodům z vrstvy *SHM_Body_1000_SJ*). Pro každou proměnnou byly vytvořeny histogramy a boxploty, aby bylo možné vizualizovat rozložení hodnot a detektovat odlehlé hodnoty. U vybraných proměnných byla provedena transformace dat.

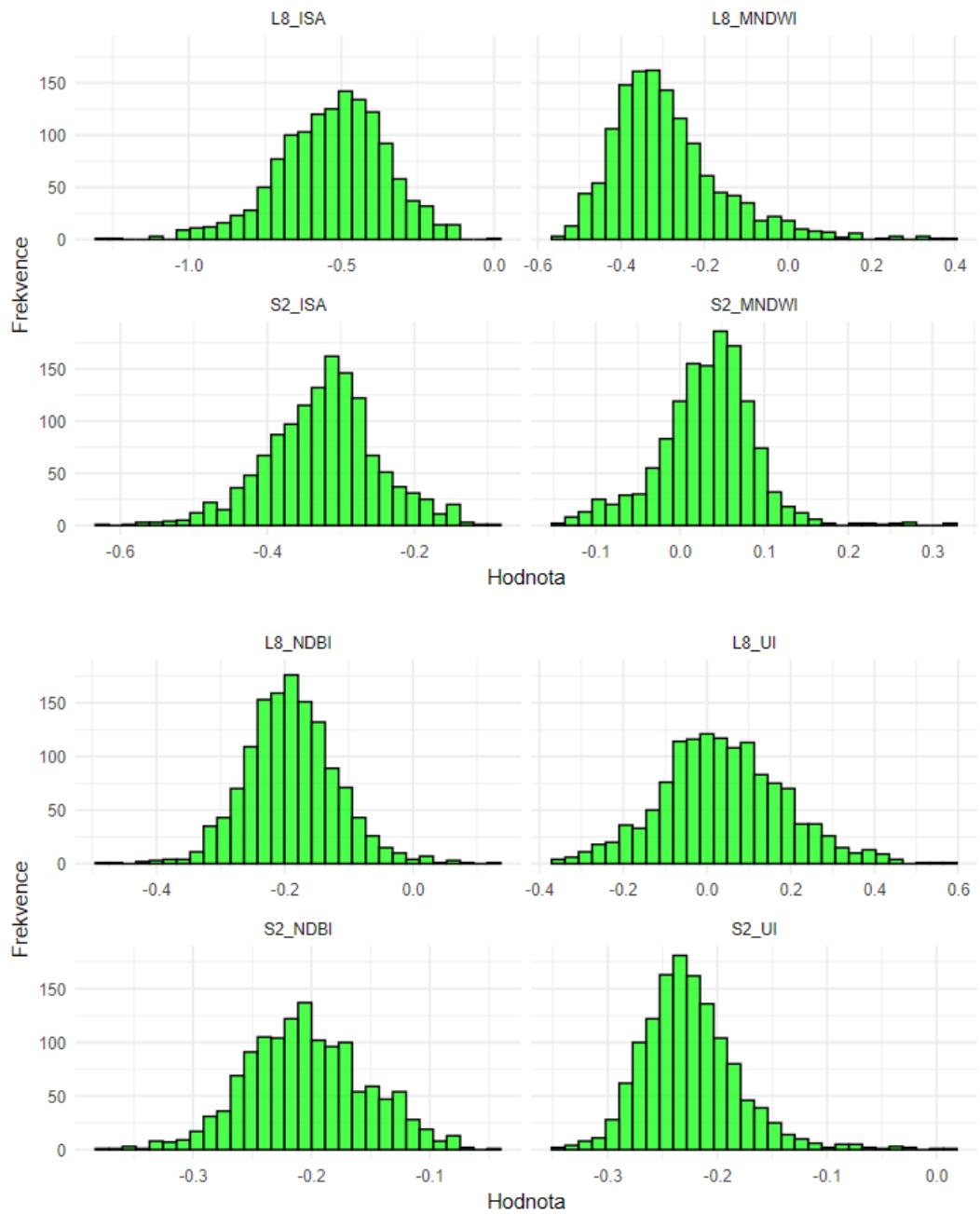
9.1.1 Roční data

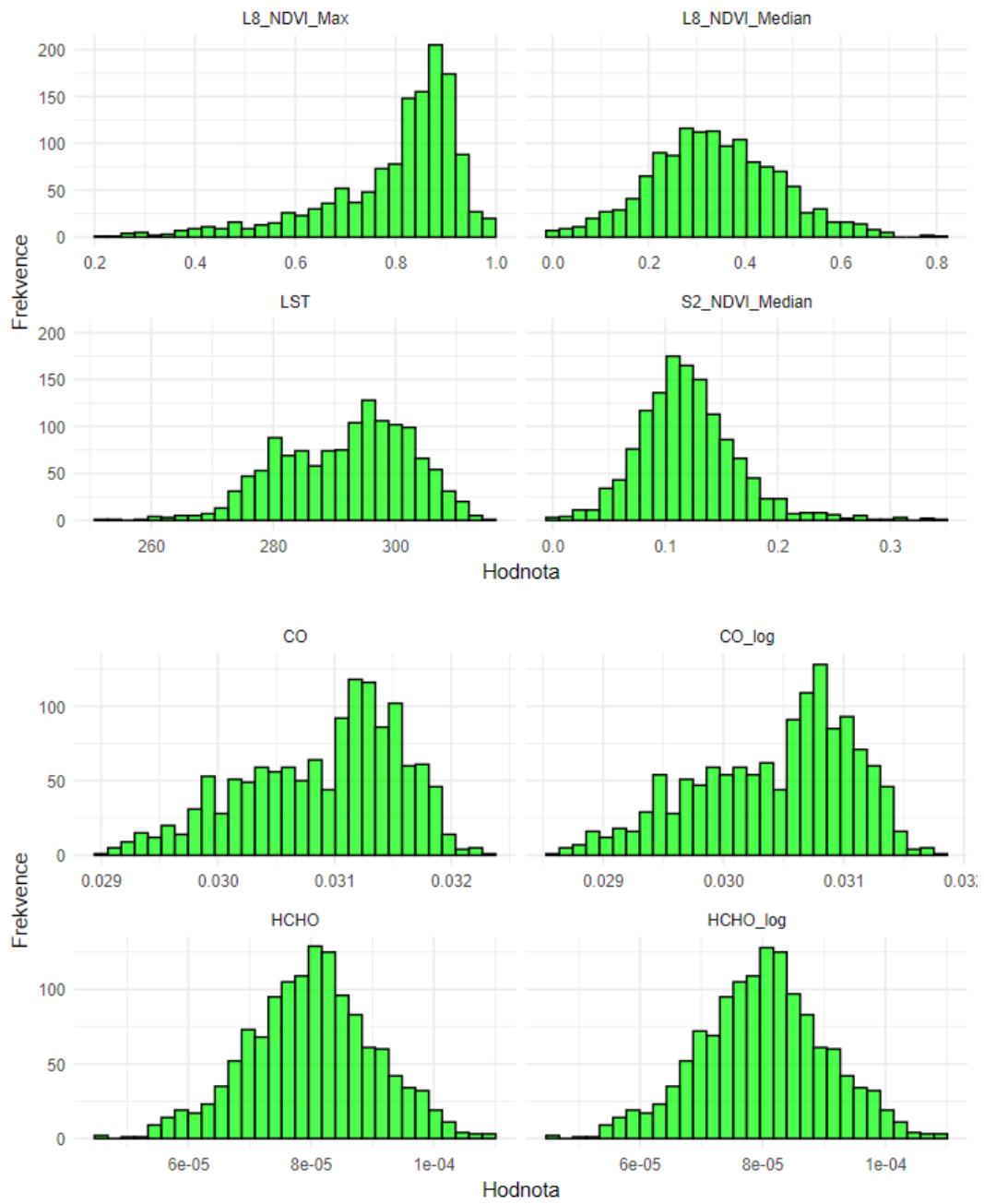
Primárně byla vytvořena vstupní datová sada obsahující mediány dat za celý kalendářní rok 2022. **Histogramy** (viz obrázek 29) ukazují rozložení jednotlivých proměnných. U některých proměnných (VIIRS, CO, CH₄, HCHO a NO₂) bylo pozorováno šikmě rozložení, což by mohlo negativně ovlivnit výkon lineárního modelu⁴⁴. Pro možnost použití lineárních metod byly tyto proměnné transformovány logaritmickou funkcí. Histogramy po transformaci ukázaly, že u VIIRS, CH₄, HCHO a NO₂ byla transformace účinná, zatímco u CO byla její efektivita nižší. U HCHO sice došlo ke zlepšení, avšak proměnná stále vykazovala určitou asymetrii.

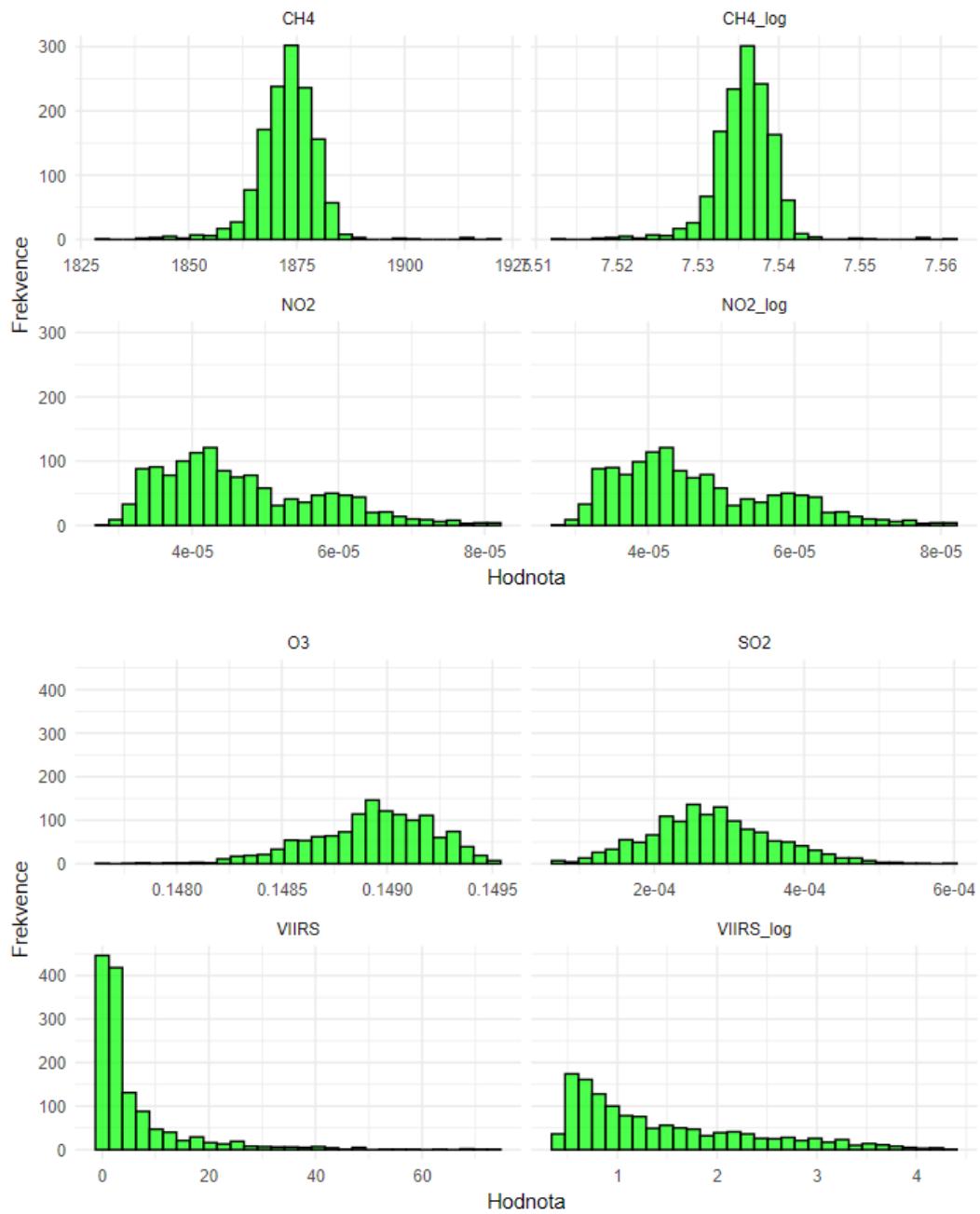
⁴² NDVI, NDBI, MNDWI, UI, ISA

⁴³ CO, HCHO, NO₂, O₃, SO₂, CH₄

⁴⁴ Lineární modely předpokládají normální rozdělení, přičemž nelineární modely jsou flexibilnější a lépe se dokáží vypořádat s asymetrií dat.

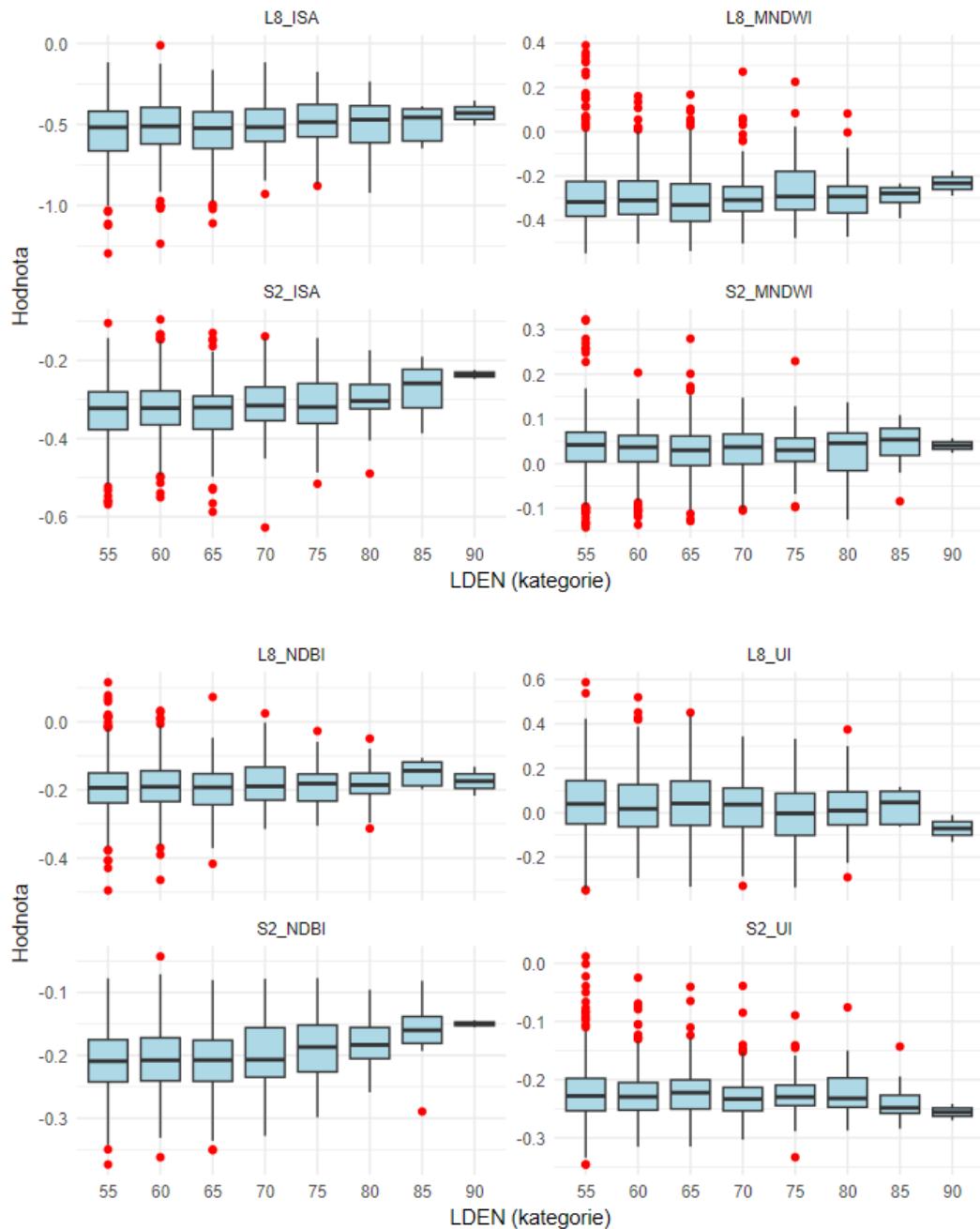


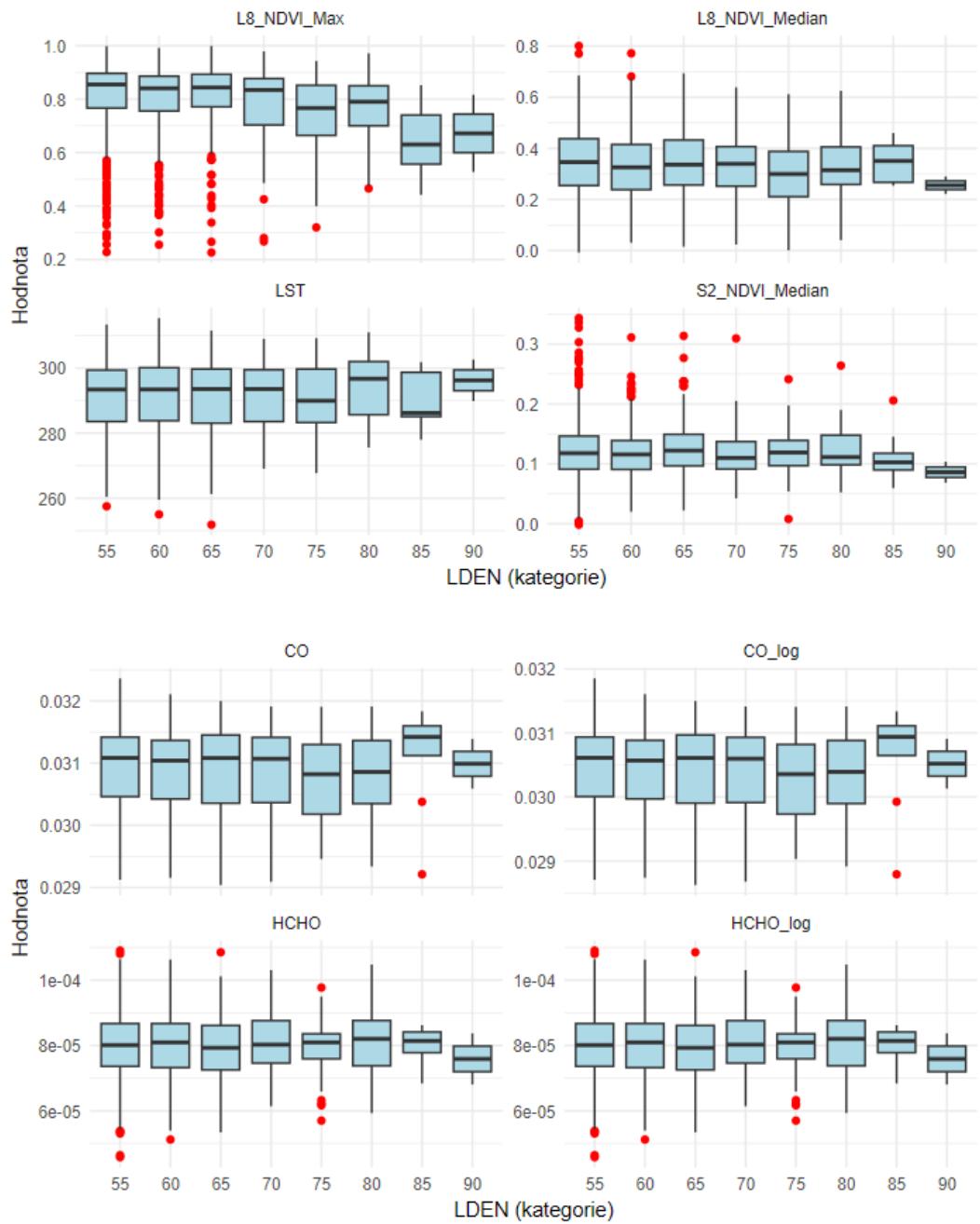


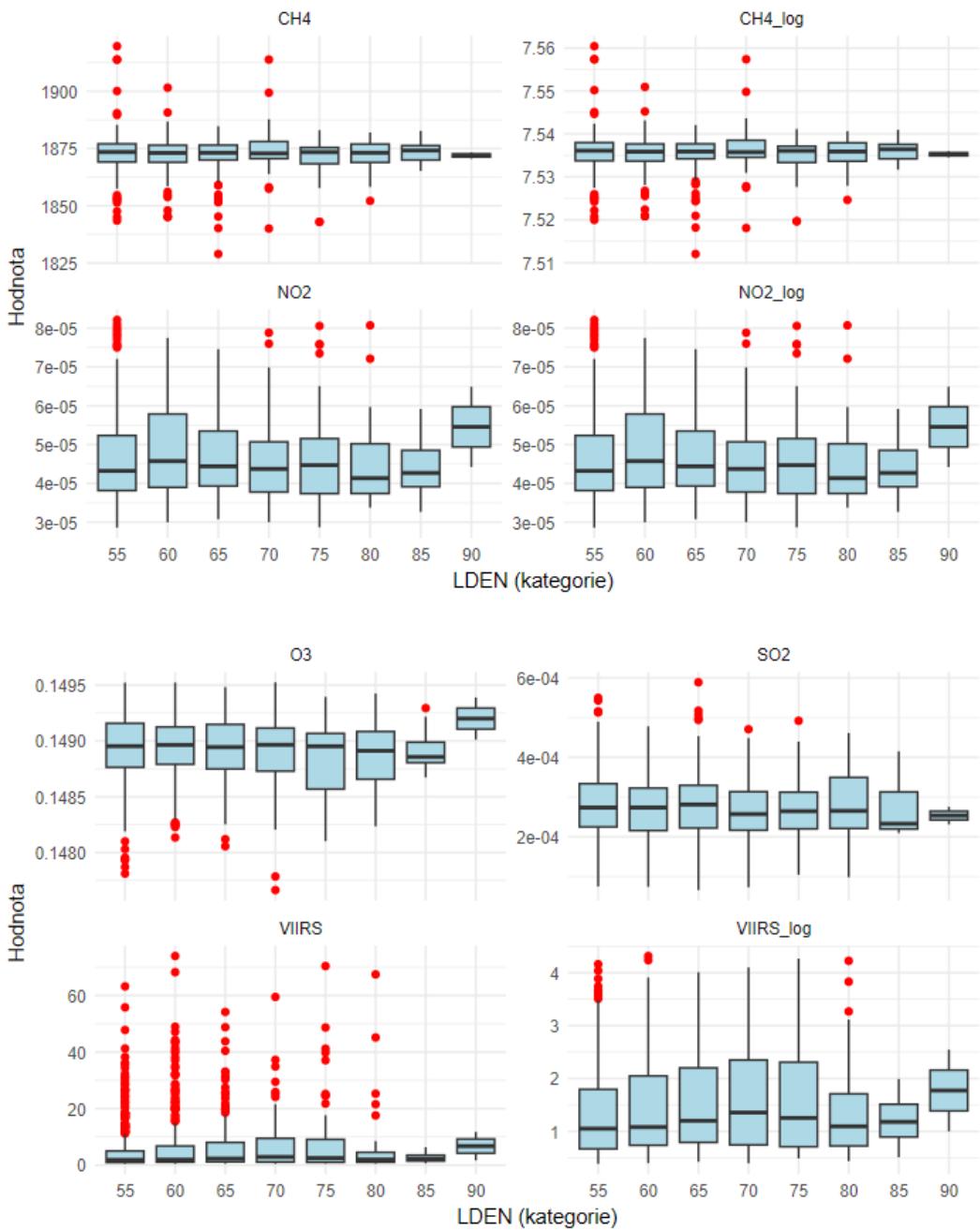


Obrázek 29: Histogramy všech proměnných včetně transformovaných proměnných (zdroj: vlastní zpracování v R Studio).

Boxploty proměnných ve vztahu k hlukovému indikátoru L_{den} (obrázek 30) ukazují, že některé proměnné vykazují signifikantní rozdíly mezi kategoriemi hluku.





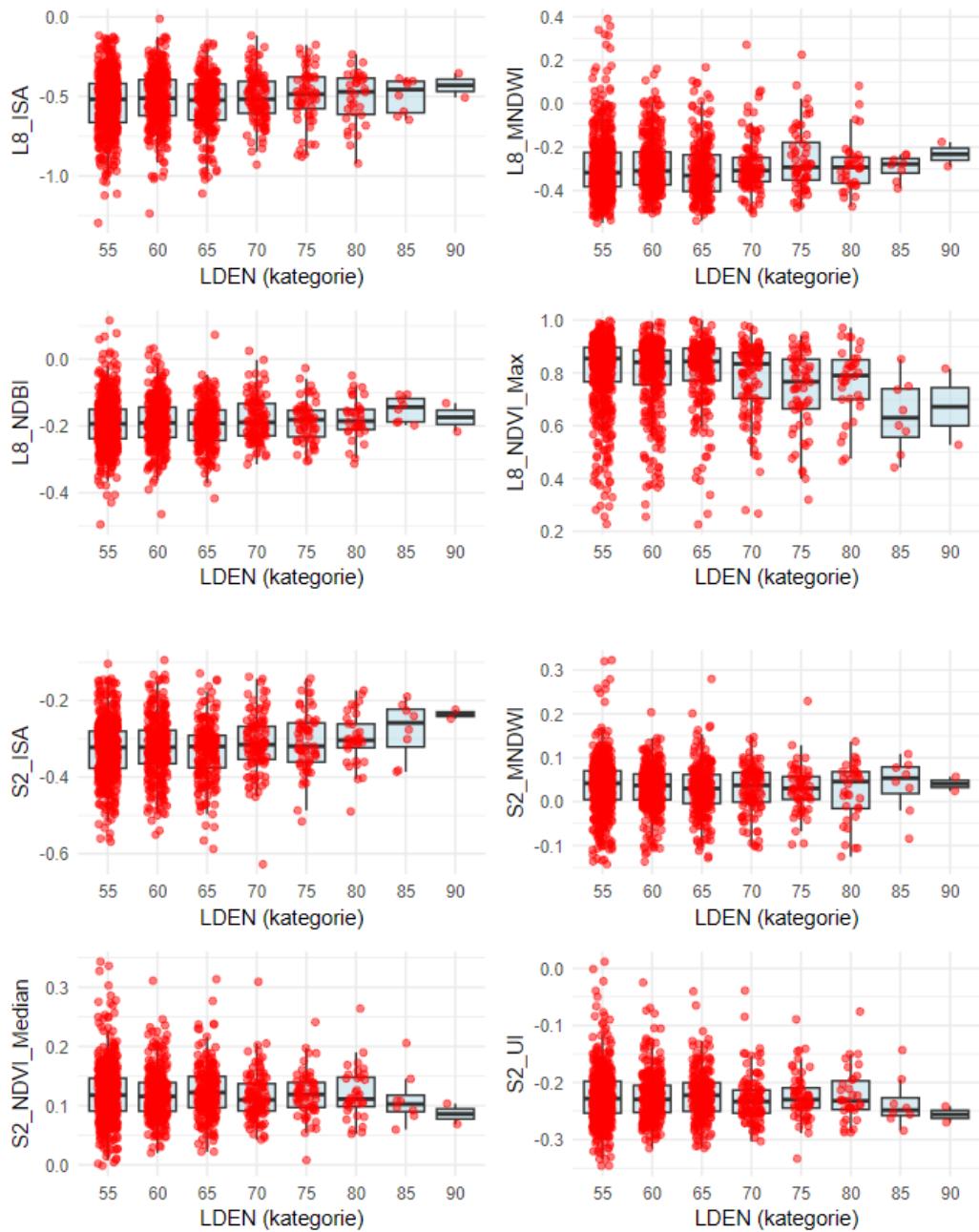


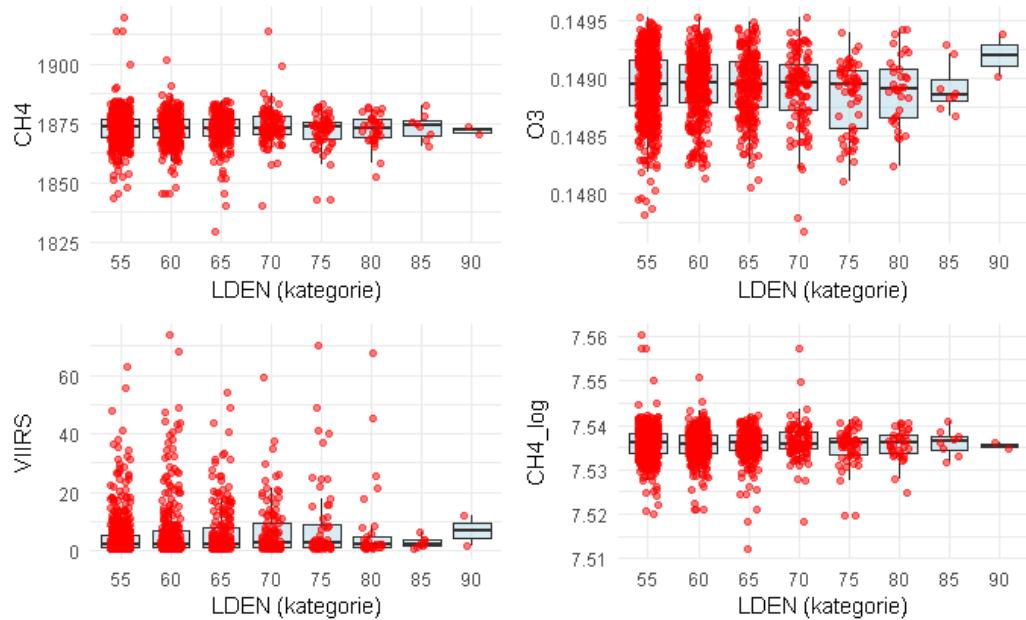
Obrázek 30: Boxploty proměnných (zdroj: vlastní zpracování v R Studio).

Například u proměnných *L8_NDVI_Max*, *S2_ISA*, *VIIRS_log* a *S2_NDVI_Median* jsou patrné rozdíly mezi jednotlivými úrovněmi hluku. Vyšší hodnoty hluku jsou spojeny s nižšími hodnotami NDVI (nižší vegetační pokryv) a vyššími hodnotami ISA (vyšší urbanizace). To naznačuje, že tyto proměnné mohou být vhodné pro prediktivní modelování. Menší, přesto pozorovatelné rozdíly v hodnotách napříč kategoriemi hlukové zátěže byly zaznamenány u proměnných *CO_log*, *CH4_log*, *HCHO_log* a *NO2_log*. Tyto ukazatele se tedy pro model jeví jako relevantní.

Pro zajištění kvality dat a odstranění případných anomálií byl dataset před analýzou podroben několika úpravám. Z datasetu byly odstraněny sloupce obsahující pouze NA hodnoty, čímž se zvýšila kvalita datasetu a eliminovaly se proměnné bez informační hodnoty. U sloupců s nulovou variabilitou (tedy těch, které obsahují konstantní hodnoty napříč všemi pozorováními) byl přidán minimální šum ($1e^{-6}$), aby se zabránilo problémům se singularitou při statistických výpočtech. Odlehlé hodnoty byly identifikovány na základě mezikvartilového rozpětí (Interquartile Range – IQR) a následně byly odstraněny, čímž byl vytvořen dataset očištěný od extrémních pozorování.

Pro ověření, zda odlehlé hodnoty jsou skutečně statisticky významné anomálie, byl proveden Grubbsův test na celý původní dataset. Výsledkem bylo rozdělení odlehlých hodnot na statisticky významné anomálie a nevýznamné odlehlé hodnoty. K identifikaci příčin těchto odlehlých hodnot byly následně vytvořeny scatter ploty s jitterem. Z těchto vizualizací (obrázek 32) je patrné, že L_{den} silně koreluje s vegetací (NDVI) a nepropustností povrchů (ISA), zatímco vodní indexy (MNDWI) a urbanizační index (UI) nejsou významně ovlivněny.





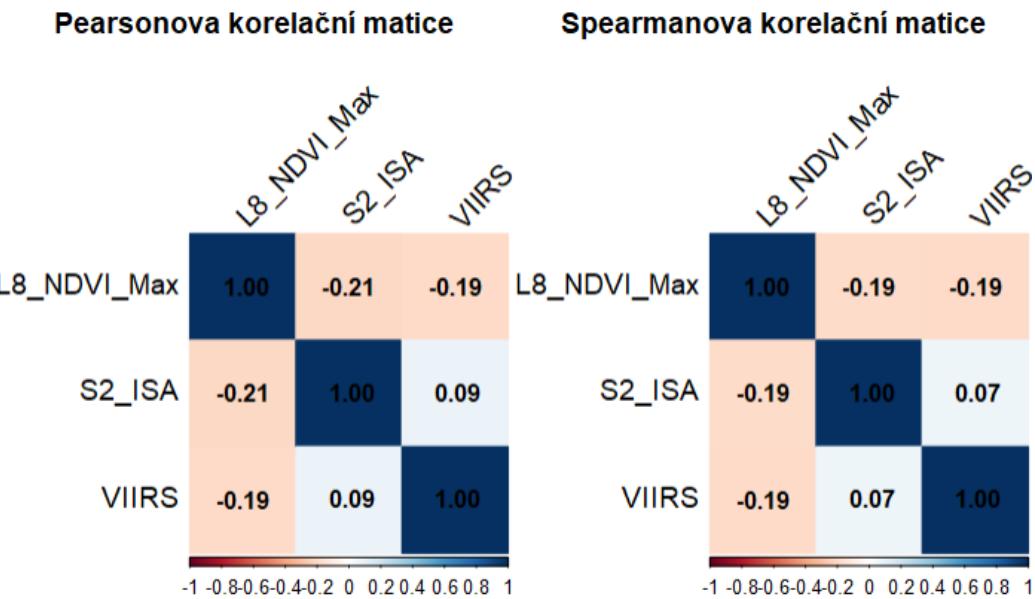
Obrázek 31: Jitter plot s Boxplot pro proměnné s odlehlými hodnotami (zdroj: vlastní zpracování v R Studio)

S ohledem na výsledky bylo pro další analýzu vhodné odstranit proměnnou O₃ z důvodu malého rozptylu hodnot v jednotlivých kategoriích hluku a CH₄ kvůli extrémním hodnotám. Byla provedena regresní analýza, přičemž pro porovnání rozdílů mezi kategoriemi hluku byla zvolena ANOVA. Jako vhodné proměnné pro analýzu se ukázaly L8_NDVI_Max (silný trend poklesu vegetace s L_{den}), S2_NDVI_Median (potvrzení NDVI z jiného satelitu), L8_ISA (vyšší podíl umělých povrchů v oblastech s vyšší hladinou hluku), S2_ISA (potvrzení trendu ISA z jiného zdroje dat), CH₄_log (transformovaná verze CH₄, redukující extrémy) a VIIRS (rostoucí trend nočního osvětlení s L_{den}).

ANOVA otestovala, zda existují statisticky významné rozdíly mezi kategoriemi hluku pro vybrané proměnné. Pro proměnné L8_NDVI_Max, S2_ISA a VIIRS byly zjištěny statisticky významné rozdíly mezi kategoriemi hluku. To potvrzuje, že v místech, kde je vyšší hluk, je méně vegetace, více nepropustných povrchů (zastavěné oblasti) a vyšší úroveň světelného znečištění. Naopak proměnné S2_NDVI_Median, L8_ISA a CH₄_log se neukázaly jako statisticky nevýznamné. Důvodem může být nevhodnost proměnných jako prediktoru nebo nedostatečné/nesprávné rozlišení satelitních dat.

Pro vybrané proměnné byly vytvořeny Pearsonova (zachycuje lineární vztahy) a Spearmanova (odolnější vůči odlehlým hodnotám) korelační matici (obrázek 33), aby bylo možné zhodnotit, zda odlehlé hodnoty ovlivňují výsledek. Z korelačních matic je patrné, že odlehlé hodnoty nemají zásadní vliv na vztahy

mezi proměnnými. Dále je patrné, že žádná proměnná nemá silnou korelaci s proměnnou L_{den} . To značí, že hladina hluku není s žádnou z těchto proměnných silně lineárně korelovaná, a tedy žádná z nich s ní nesdílí výrazný lineární vztah. Neznamená to však, že by nemohly být relevantní v predikčním modelu při kombinaci s dalšími proměnnými.



Obrázek 32: Korelační matice vybraných proměnných (zdroj: vlastní zpracování v R Studio).

Tato slabá korelace mezi hodnotou hluku (L_{den}) a ostatními proměnnými naznačuje, že by bylo vhodnější zvolit nějaký nelineární model, neboť vztah mezi proměnnými může být složitější. S ohledem na tuto skutečnost je vhodné použít vhodnou pokročilejší modelovací techniku.

9.1.2 Měsíční data

Pro porovnání vhodnosti vstupních dat pro modelování byly vytvořeny čtyři sady měsíčních dat. Hluková data, která slouží jako cílová proměnná, nejsou spojena s konkrétním časovým obdobím – je pouze známo, že byla spočítána za rok 2022. Z tohoto důvodu bylo zvoleno kratší časové období pro vstupní proměnné, aby bylo možné testovat jejich vztah k průměrnému ročnímu hluku. Každá sada byla sestavena jako měsíční mediány pro vybrané měsíce, a to únor, květen, srpen a listopad. Tato volba se ukázala jako nedostačující, neboť po odstranění NA hodnot z původních 1 400 hodnot zůstalo pro únorová data 641 hodnot, pro květnová data 855 hodnot, a pro srpnová data dokonce pouhých 60 hodnot. U volby měsíčních dat by tak došlo k výraznému snížení vzorku vstupních dat.

v důsledku nutnosti odstranění NA hodnot. S ohledem na výše uvedené byly pro další analýzu použity pouze roční agregované údaje.

9.2 Tvorba modelu

Vstupní datová sada byla rozdělena na trénovací, validační a testovací data, a to v poměru 70 : 15 : 15. Na těchto datech pak byly použity níže uvedené metody. Aby bylo zajištěno rovnoměrné rozložení ve všech kategoriích, byla na testovací data použit balíček *smotefamily*, jenž využívá generování syntetických dat k vyvážení tříd. Pro některé hlukové kategorie bylo k dispozici malé množství hodnot (70 dB – 47 hodnot, 75 dB – 19, 80 dB – 12, 85 dB – 2, 90 dB – 1).

9.2.1 Random Forest

Pro tvorbu modelu byla použita metoda Random Forest (RF) (viz příloha 3). Model vykazuje velmi vysokou chybovost 55,78 % (viz obrázek 34), přičemž pro kategorie, kde je malý počet hodnot, činí 100 %. To znamená, že tyto kategorie není schopen model správně klasifikovat. Sloučení kategorií hluku 70–90 dB by mohlo problém vyřešit.

```
Call:  
randomForest(formula = LDEN ~ ., data = balanced_data, mtry = 2,      ntree = 300, nodesize =  
10, importance = TRUE, na.action = na.omit)  
Type of random forest: classification  
Number of trees: 300  
No. of variables tried at each split: 2  
  
OOB estimate of error rate: 54.94%  
Confusion matrix:  
   90 55 60 65 70 75 80 85 class.error  
90  0  1  0  0  0  0  0  0  1.0000000  
55  0 230 34 13  2  0  0  0  0.1756272  
60  0 102 36  3  3  0  0  0  0.7500000  
65  0  74 14  3  2  0  0  0  0.9677419  
70  0  31 12  4  0  0  0  0  1.0000000  
75  0  15  2  0  2  0  0  0  1.0000000  
80  0   9  2  0  1  0  0  0  1.0000000  
85  0   1  1  0  0  0  0  0  1.0000000
```

Obrázek 33: Výsledky trénování RF modelu (zdroj: vlastní zpracování v R Studio).

Jak je patrné z obrázku 34, nemělo sloučení kategorií na výsledek žádný vliv, výsledek byl stejný jako v předchozím případě.

```

Call:
randomForest(formula = LDEN ~ ., data = balanced_data, mtry = 2,      ntree = 300, nodesize =
15, importance = TRUE, na.action = na.omit)
      Type of random forest: classification
                  Number of trees: 300
No. of variables tried at each split: 2

OOB estimate of  error rate: 54.94%
Confusion matrix:
 70+ 55 60 65 class.error
70+  8 58 15  0  0.9012346
55   11 230 33  5  0.1756272
60   4 107 30  3  0.7916667
65   4  72 16  1  0.9892473

```

Obrázek 34: Výsledky trénování modelu po sloučení kategorií hluku (zdroj: vlastní zpracování v R Studio).

Po navýšení počtu náhodně vybraných proměnných (*mtry*) na 3 došlo ke slabému snížení chybovosti na 54,27 %. V případě zvýšení počtu stromů (*ntree*) na 500 se chybovost lehce snížila na 53,77 %. Snížení počtu stromů na 400 se chybovost mírně zvýšila na 53,94 % a v případě navýšení na 600 stromů došlo opět ke zvýšení na 54,1 %. Z výsledků je patrné, že ani po úpravě různých hyperparametrů nedochází k podstatnému snížení chybovosti modelu Random Forest, zůstává velmi vysoká.

Tabulka 4: Statistika modelu při použití Random Forest (zdroj: vlastní zpracování v R Studio).

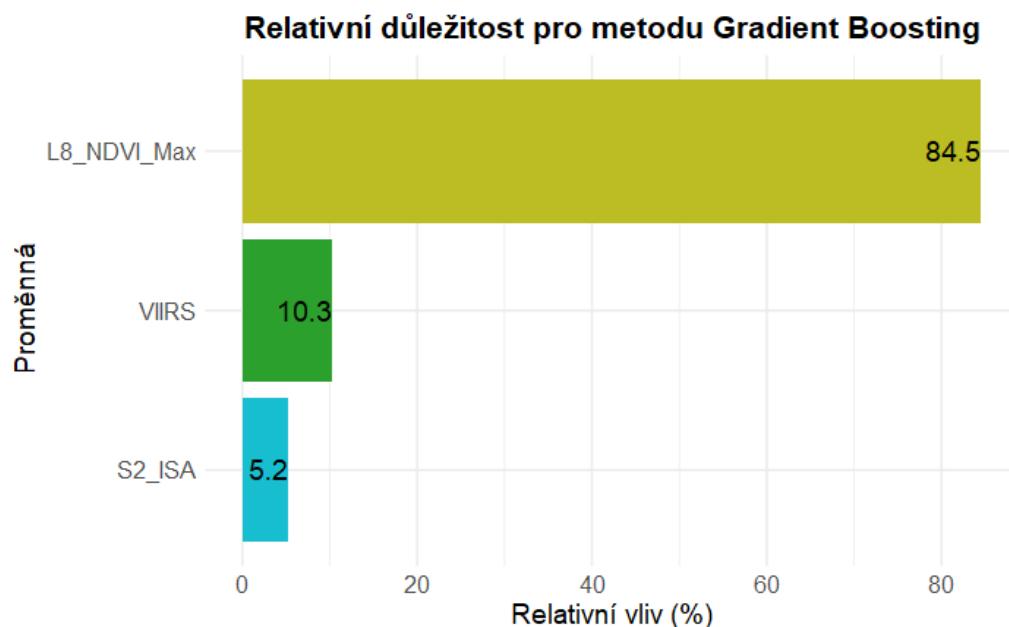
Random Forest			
	Train Data	Test Data	Valid Data
RMSE	1,46	1,44	1,61
MAE	1,17	1,13	1,12
R²	0,00002	0,00038	0,00199

Při výpočtu statistik modelu (tabulka 4) na základě trénovacích, testovacích a validačních dat je patrné, že model má velice nízkou prediktivní sílu (R^2 se blíží nule). Přesto však dosahuje poměrně nízké střední kvadratická chyba okolo 1,5 dB. To naznačuje, že predikce se sice příliš neliší od průměrů, ale model nevystihuje variabilitu dat.

9.2.2 Gradient Boosting

Další použitou metodou pro sestavení modelu byl Gradient Boosting (GB) (viz příloha 4). U modelu byla spuštěna funkce pro optimalizaci vstupních parametrů modelu, která navrhla následující nastavení: počet stromů v modelu 100 (*n.trees*), hloubka stromů 1 (*interaction.depth*), rychlosť učení modelu 0,01 (*shrinkage*) a minimální počet vzorků na uzlu 20 (*n.minobsinnode*). Nastavení počtu stromů na 100 značí, že trénink modelu bude rychlejší, avšak s možností omezení výkonu. Hloubka interakce 1 představuje tu nejjednodušší formu, kdy

model využívá pouze jeden prediktor a jednu proměnnou. Hodnota 0,01 u rychlosti učení poukazuje na pomalé učení, avšak snižuje pravděpodobnost přeucení. Minimální počet vzorků v uzlu 20 omezuje hloubku větvení stromů a tím zlepšuje generalizaci modelu. Test významu proměnných identifikoval (obrázek 35) jako nejdůležitější prediktor proměnnou *L8_NDVI_Max* s podílem 84,5 % důležitosti. Ostatní prediktory vykazují výrazně nižší míru významnosti.



Obrázek 35: Relativní důležitost proměnných pro model Gradient Boosting (zdroj: vlastní zpracování v R Studio).

Při výpočtu statistik modelu (tabulka 5) byly výsledky stejné jak pro optimalizované, tak pro ručně upravované parametry. Změna hyperparametrů neměla na výsledky vliv.

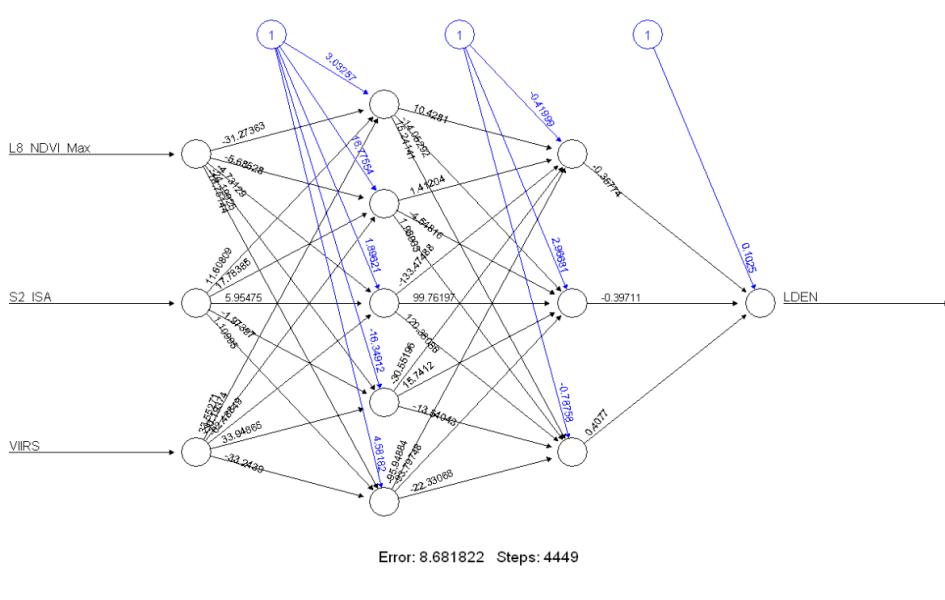
Tabulka 5: Statistika modelu při použití Gradient Boosting (zdroj: vlastní zpracování v R Studio).

Gradient Boosting			
	Train Data	Test Data	Valid Data
RMSE	1,27	1,35	1,49
MAE	0,99	1,07	1,07
R²	0,04712	0,00515	0,00194

Výsledky pro trénovací, testovací i validační data jsou vzájemně srovnatelné. Model však vykazuje velmi nízkou predikční schopnost, což dokládají hodnoty koeficientu determinace R². Přestože je střední kvadratická chyba poměrně malá – přibližně 1,4 dB. Tento rozpor naznačuje, že model dobře odhaduje průměr, ale nedokáže zachytit variabilitu hodnot.

9.2.3 Neuronové sítě

Posledním zvoleným modelem byla neuronová síť. Data pro model byla nejdříve normalizována. U modelu byla spuštěna funkce pro vyhledání nejlepší kombinace skryté vrstvy. Ve výběru pro testování byly struktury $(5, 3)$ ⁴⁵, $(10, 5)$, $(20, 10, 5)$ ⁴⁶, $(10, 5, 3)$, $(20, 10, 5)$ a $(30, 15, 5)$. Nejvhodnější se ukázala struktura $(5, 3)$. Výstup modelu je znázorněn na obrázku 36.



Obrázek 36: Výsledek modelu neuronové sítě pro strukturu $(5, 3)$ (zdroj: vlastní zpracování v R Studio).

Výsledky pro trénovací data byly lepší než pro testovací a validační data (tabulka 6).

Tabulka 6: Statistika modelu s využitím neuronové sítě (zdroj: vlastní zpracování v R Studio).

Neuronové sítě			
	Train Data	Test Data	Valid Data
RMSE	5,97	7,01	7,78
MAE	4,66	5,53	5,49
R²	0,15559	0,00210	0,01137

Nízké hodnoty R^2 naznačují, že model má omezenou predikční schopnost, což může být způsobeno buď přeúčením (overfittingem), nebo vysokou variabilitou v datech. U trénovacích dat model vysvětlil přibližně 16 % variability, zatímco u validačních dat pouze 1,1 % a u testovacích dokonce jen 0,2 %. Přestože R^2 je

⁴⁵ Struktura $(5, 3)$ – 2 skryté vrstvy s 5 a 3 neurony (celkem 8 neuronů).

⁴⁶ Struktura $(20, 10, 5)$ – 3 skryté vrstvy s 20, 10 a 5 neurony (celkem 35 neuronů).

velmi nízké, střední absolutní chyba predikce zůstává u všech datasetů v přijatelném rozmezí 6–7,8 dB.

10 Souhrn výsledků

Na základě provedené analýzy a testování různých modelovacích přístupů (Random Forest, Gradient Boosting a neuronové sítě) je zřejmé, že žádná z metod nebyla schopna uspokojivě vysvětlit variabilitu hlukových dat (L_{den}) na základě vybraných prediktorů.

Random Forest i Gradient Boosting dosáhly velmi nízkých hodnot koeficientu determinace R^2 (tabulka 7), což naznačuje, že prediktory nemají dostatečnou vypovídací schopnost pro predikci hluku v městském prostředí. I když chyby RMSE a MAE byly relativně nízké (pohybují se kolem 1,3–1,6 dB), skutečnost, že modely nevysvětlují téměř žádnou variabilitu v datech, znamená, že predikce jsou nepřesné a nejsou spolehlivé.

Nejlepšího R^2 na trénovacích datech dosáhl model neuronových sítí (15,56 %), avšak na testovacích a validačních datech klesla hodnota pod 2 %, což naznačuje přeучení modelu. RMSE a MAE byly u neuronových sítí výrazně vyšší (6–8 dB), což ukazuje na horší predikční schopnost modelu ve srovnání s Random Forest a Gradient Boosting.

Tabulka 7: Tabulka souhrnu výsledků použitých metod RF, GB a NN (zdroj: vlastní zpracování v MS Excel).

	RMSE Train	RMSE Test	RMSE Valid	R ² Train	R ² Test	R ² Valid
Random Forest	1,46	1,44	1,61	0,00 %	0,04 %	0,20 %
Gradient Boosting	1,27	1,35	1,49	4,72 %	0,52 %	0,19 %
Neuronové sítě	5,97	7,01	7,78	15,56 %	0,21 %	1,14 %

Na základě těchto výsledků lze konstatovat, že žádný z testovaných modelů neposkytl uspokojivou predikční schopnost pro vysvětlení hlukové zátěže na základě dostupných satelitních a environmentálních dat.

11 Diskuse

V současnosti neexistuje žádný senzor na oběžné dráze Země, který by přímo měřil hluk. Hluk je z vesmíru aktuálně neměřitelný, nicméně satelitní data poskytují široké spektrum environmentálních informací, které mohou s hlukovou zátěží korelovat a potenciálně umožnit její odhad.

Cílem této práce bylo **ověřit možnost predikce hluku** na základě různých environmentálních proměnných získaných ze satelitních senzorů. Výstupem sice není funkční model s dostatečnou predikční schopností, nicméně výsledky naznačují **možné důvody neúspěchu**, především ve **vhodnosti výběru vstupních dat a jejich charakteru**.

Prostorové a metodické omezení vstupních dat

Jedním z hlavních limitujících faktorů byla různorodost prostorového rozlišení vstupních dat. Například termální data byla získána ze senzoru MODIS s prostorovým rozlišením 1 000 m, avšak vhodnější alternativou by byl Sentinel-3 SLSTR s prostorovým rozlišením 500 m. Stahování těchto dat však není možné dávkově (např. použitím Google Earth Engine, Sentinel Hub). Rovněž ostatní vstupní data mají různá prostorová rozlišení v rozmezí od 10 m do 1 000 m, což mohlo způsobit potíže při jejich vzájemném sladění a ovlivnit přesnost modelu.

Dalším problémem byla vstupní hluková data. Prvním zásadním nedostatkem je, že data neobsahují kategorie hluku pod 50 dB, což znamená, že tišší oblasti nejsou ve vstupním souboru vůbec zastoupeny. To mohlo ovlivnit schopnost modelu správně odhadnout vztah mezi prediktory a hladinou hluku. Druhým problémem je samotná metoda výpočtu hlukového ukazatele L_{den} , který zahrnuje penalizaci za hluk ve večerních hodinách (+5 dB) a v noci (+10 dB). To **může výrazně zkreslovat hodnoty hlukových dat**, přestože by se penalizace měly v ideálním případě rovnoměrně rozložit v celém datasetu.

Metodické srovnání s literaturou

Výběr metod Random Forest, Gradient Boosting a neuronové sítě zohledňoval metody využité v úspěšných studiích, v nichž je využili k predikci hluku. V minulosti byly v různých výzkumech kombinovány environmentální a urbanizační proměnné, například:

- Raess et al. (2021) využil vegetační indexy pro predikci hluku.

- Koi et al. (1993) testoval predikční modely zahrnující zastavěnost území a simulované hodnoty NO₂.
- Eicher et al. (2022) využil snímky Sentinel-2 v kombinaci s U-Net neuronovou sítí.
- Lia et al. (2019) zjistila, že Random Forest dosahuje lepších výsledků než Land Use Regression (LUR) při modelování hluku, neboť dokáže zachytit složitější nelineární vztahy mezi proměnnými.
- Staab et al. (2021) využil LUR model pro mapování hluku v Německu, přičemž dosáhl uspokojivých výsledků.
- Mohammed et al. (2022) testoval regresní model v Nigérii (Kano) a potvrdil, že hustota obyvatelstva je klíčovým faktorem při vysvětlování změn hladiny hluku.

Možnosti dalšího výzkumu a doporučení

S ohledem na dosažené výsledky by se budoucí práce měla zaměřit na rozšíření a úpravu vstupních dat:

1. Použití měřených hlukových dat namísto modelovaných hodnot.
2. Zahrnutí dalších prediktorů, zejména dopravních údajů, socioekonomických faktorů (hustota obyvatelstva) a meteorologických parametrů.
3. Vylepšení prostorového rozlišení vstupních dat, například použitím dat s jednotným rozlišením 100 m.
4. Testování pokročilejších modelů, jako jsou hluboké neuronové sítě nebo hybridní metody kombinující strojové učení a tradiční statistické přístupy.

Výsledky této práce jasně ukazují, že současný přístup není vhodný pro predikci hluku s využitím dat ze SHM, což lze shrnout slovy Járy Cimrmana:

„Někdo musel slepou uličku lidského poznání ohledat a ohlásit světu: Tudy ne, přátelé.“

12 Závěr

Tato práce se zaměřila na využití různých zdrojů satelitních dat a pokročilých metod modelování pro predikci hlukové zátěže na území České republiky. Hlavním cílem bylo ověřit, zda je možné efektivně využít dostupná data z družic Landsat 8, MODIS, Sentinel-2, Sentinel-5P a VIIRS – konkrétně spektrální indexy, teplotu povrchu, koncentrace atmosférických plynů a světelné znečištění – pro odhad hladin hluku z dat ze SHM 2022.

Výsledky analýzy ukázaly, že použité metody (Random Forest, Gradient Boosting a neuronové sítě) nedosáhly uspokojivé přesnosti predikce hlukové zátěže. Hlavní příčinou tohoto výsledku je pravděpodobně kombinace několika faktorů, a to zejména nízké kvality a nehomogenity vstupních dat, rozdílů v prostorovém rozlišení datových zdrojů, absence kategorií hluku pod 50 dB a možného zkreslení způsobeného penalizací ukazatele L_{den} .

Přestože se nepodařilo vytvořit funkční predikční model, práce identifikovala klíčové problémy spojené s využitím satelitních dat pro predikci hluku a zároveň navrhla možná zlepšení pro budoucí výzkum. Mezi doporučení patří:

- Zvýšení kvality vstupních dat, například využitím termálních snímků Sentinel-3 SLSTR s jemnějším rozlišením.
- Rozšíření datasetu o přesnější a aktuálnější hluková data, která zahrnují i nižší hladiny hluku.
- Využití pokročilejších modelovacích metod, například hlubokého učení pracujícího přímo s rastrovými daty.

Význam této práce spočívá v identifikaci limitů a překážek, které je nutné překonat při snaze o predikci hlukové zátěže pomocí satelitních dat. Ačkoli nebylo dosaženo výsledku v podobě spolehlivého predikčního modelu, zjištění této práce přináší hodnotné poznatky pro budoucí výzkum v oblasti modelování hluku. Práce tak otevírá prostor pro další inovace a naznačuje směr, kterým by se mohly ubírat snahy o efektivnější mapování hlukové zátěže s využitím moderních technologií.

Tato práce ukazuje, že cesta využití kombinace satelitních dat a modelovacích metod pro predikci hluku vyžaduje zásadní úpravy a vylepšení, než bude možné ji efektivně implementovat.

Součástí výstupu této práce je také interaktivní Story Mapa, která slouží jako nástroj pro prezentaci výsledků široké veřejnosti i odborné komunitě. Story Mapa obsahuje základní informace o práci, postupu a výsledků s odkazem na kompletní znění práce včetně zdrojových kódů. Jejím cílem je nejen přiblížit problematiku prostorové analýzy hlukové zátěže, ale zároveň umožnit opakování postupu v jiném území a s jinými daty.

DP Story Maps



Seznam použité literatury

Legislativa

- [1] Commission recommendation **2003/613/ES**, concerning the guidelines on the revised interim computation methods for industrial noise, aircraft noise, road traffic noise and railway noise, and related emission data.
- [2] Nařízení vlády č. **272/2011 Sb.**, o ochraně zdraví před nepříznivými účinky hluku a vibrací, ve znění Nařízení vlády č. 217/2016 Sb. a Nařízení vlády č. 241/2018 Sb.
- [3] Směrnice Evropského parlamentu a Rady **2002/49/ES** o hodnocení a řízení hluku ve venkovním prostředí.
- [4] Směrnice Komise (EU) **2015/996**, o stanovení společných metod hodnocení hluku podle směrnice Evropského parlamentu a Rady 2002/49/ES.
- [5] Smlouva o založení Evropského společenství (konsolidované znění 2002) **2002/C 325/01**.
- [6] Vyhláška č. **315/2018 Sb.**, o strategickém hlukovém mapování, v platném znění.
- [7] Zákon č. **258/2000 Sb.**, o ochraně veřejného zdraví a o změně některých souvisejících zákonů.

Balíčky (Python a R)

- [8] Spatenka, J. ***openeo***: Python Client for the openEO API. R package version 0.14.0, 2023. Dostupné z: <<https://pypi.org/project/openeo/>>
- [9] Gillies, S. ***Rasterio***: Geospatial Raster Data Processing Library. R package version 1.3.6, 2023. Dostupné z: <<https://pypi.org/project/rasterio/>>
- [10] Wu, Q. ***Geemap***: Interactive Geospatial Analysis and Visualization with Google Earth Engine. R package version 0.28.0, 2023. Dostupné z: <<https://geemap.org/>>
- [11] Google Developers. ***Earth Engine Python API***. R package version 0.1.341, 2023. Dostupné z: <<https://pypi.org/project/earthengine-api/>>
- [12] Cortesi, I. ***ijson***: Iterative JSON Parser for Python. R package version 3.2.3, 2022. Dostupné z: <<https://pypi.org/project/ijson/>>
- [13] Hunter, J. D. ***Matplotlib***: A 2D Graphics Environment. Computing in Science & Engineering, 9(3), 2007, pp. 90–95. Dostupné z: <<https://pypi.org/project/matplotlib/>>
- [14] Warmerdam, F. ***GDAL***: Geospatial Data Abstraction Library. Open Source Geospatial Foundation, 2023. Dostupné z: <<https://pypi.org/project/GDAL/>>
- [15] Jordahl, K. ***GeoPandas***: Python Tools for Geographic Data Analysis. Journal of Open Source Software, 4(43), 2019, p. 1686. Dostupné z: <<https://pypi.org/project/geopandas/>>
- [16] McKinney, W. ***pandas***: A Foundational Python Library for Data Analysis. Python Software Foundation, 2023. Dostupné z: <<https://pypi.org/project/pandas/>>
- [17] Python Software Foundation. ***datetime***: Date/Time Data Type and Manipulation. Dostupné z: <<https://docs.python.org/3/library/datetime.html>>
- [18] Harris, C. R., et al. Array programming with ***NumPy***. Nature, 585(7825), 2020, pp. 357–362. Dostupné z: <<https://pypi.org/project/numpy/>>
- [19] Python Software Foundation. ***shutil***: High-Level File Operations. Dostupné z: <<https://docs.python.org/3/library/shutil.html>>
- [20] Reitz, K. ***Requests***: HTTP for Humans. Python Software Foundation, 2023. Dostupné z: <<https://pypi.org/project/requests/>>

- [21] Gillies, S. ***Shapely***: Geometric Operations in the Cartesian Plane. R package version 2.0.2, 2023. Dostupné z: <<https://pypi.org/project/shapely/>>
- [22] Wickham, Hadley, François, Romain, Henry, Lionel, and Müller, Kirill. ***dplyr***: A Grammar of Data Manipulation. R package version 1.0.10, 2022. Dostupné z: <<https://CRAN.R-project.org/package=dplyr>>
- [23] Lang, Gabor Csardi. ***pkgconfig***: Private Configuration for R Packages. R package version 2.0.3, 2019. Dostupné z: <<https://CRAN.R-project.org/package=pkgconfig>>
- [24] Dowle, Matt and Srinivasan, Arun. ***data.table***: Extension of data.frame. R package version 1.14.2, 2021. Dostupné z: <<https://CRAN.R-project.org/package=data.table>>
- [25] Wickham, Hadley. ***ggplot2***: Elegant Graphics for Data Analysis. New York: Springer-Verlag, 2016. ISBN 978-3-319-24277-4. Dostupné z: <<https://ggplot2.tidyverse.org/>>
- [26] Liaw, Andy and Wiener, Matthew. ***randomForest***: Breiman and Cutler's Random Forests for Classification and Regression. R package version 4.6-14, 2021. Dostupné z: <<https://CRAN.R-project.org/package=randomForest>>
- [27] Kuhn, Max. ***caret***: Classification and Regression Training. R package version 6.0-90, 2021. Dostupné z: <<https://CRAN.R-project.org/package=caret>>
- [28] Zhao, Zhang. ***smotefamily***: A Collection of Oversampling Techniques for Class Imbalance Problem Based on SMOTE. R package version 1.3.1, 2020. Dostupné z: <<https://CRAN.R-project.org/package=smotefamily>>
- [29] Pedersen, Thomas Lin. ***patchwork***: The Composer of ggplots. R package version 1.1.1, 2020. Dostupné z: <<https://CRAN.R-project.org/package=patchwork>>
- [30] Henry, Lionel and Wickham, Hadley. ***purrr***: Functional Programming Tools. R package version 1.0.0, 2022. Dostupné z: <<https://CRAN.R-project.org/package=purrr>>
- [31] Wei, Taiyun and Simko, Viliam. ***corrplot***: Visualization of a Correlation Matrix. R package version 0.92, 2021. Dostupné z: <<https://CRAN.R-project.org/package=corrplot>>
- [32] Fox, John and Weisberg, Sanford. ***car***: An R Companion to Applied Regression. 3rd edition. Thousand Oaks, CA: Sage, 2019. ISBN 978-1-5443-3646-9. Dostupné z: <<https://CRAN.R-project.org/package=car>>

- [33] Wright, Marvin and Ziegler, Andreas. **ranger**: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. Journal of Statistical Software, 77(1), 2017, pp. 1–17. DOI: 10.18637/jss.v077.i01. Dostupné z: <<https://CRAN.R-project.org/package=ranger>>
- [34] Greenwell, Brandon M., Boehmke, Bradley C., Cunningham, Jeffrey, and Developers, GBM. **gbm**: Generalized Boosted Regression Models. R package version 2.1.8, 2020. Dostupné z: <<https://CRAN.R-project.org/package=gbm>>
- [35] Günther, Frauke and Fritsch, Stefan. **neuralnet**: Training of Neural Networks. R package version 1.44.2, 2016. Dostupné z: <<https://CRAN.R-project.org/package=neuralnet>>

Ostatní zdroje

- [36] ALI A., NAYYAR Z.A., 2021: A Modified Built-up Index (MBI) for automatic urban area extraction from Landsat 8 Imagery. Infrared Physics & Technology, Volume 116, DOI: 10.1016/j.infrared.2021.103769.
- [37] AMORIM M. C. P., VIEIRA M., MEIRELES G., NOVAIS S. C., LEMOS M. F. L., MODESTO T., ALVES D., ZUAZU A., LOPES A. F., MATOS A. B., FONSECA P. J., 2022: Boat noise impacts Lusitanian toadfish breeding males and reproductive outcome. Science of The Total Environment, Volume 830, DOI: 10.1016/j.scitotenv.2022.154735.
- [38] BALADRAM S., MEDIUM, 2024: Random Forest, Explained: A Visual Guide with Code Examples (online)[cit.2025.01.16], dostupné z <<https://medium.com/towards-data-science/random-forest-explained-a-visual-guide-with-code-examples-9f736a6e1b3c>>
- [39] BHATTI S.S., TRIPATHI N.K., 2014: Built-up area extraction using Landsat 8 OLI imagery. GIScience & Remote Sensing, 51(4), S 445-467, DOI: 10.1080/15481603.2014.939539.
- [40] BENNOCI R., ROMAN H.E., ZAMBON G., 2021: Optimized Sensors Network and Dynamical Maps for Monitoring Traffic Noise in a Large Urban Zone. Applied Sciences, S. 11, 8363, DOI: 10.3390/app11188363
- [41] BERAN V., 2010: Chvění a hluk. Západočeská univerzita v Plzni, Vydavatelství. Plzeň. 204 s.
- [42] BREIMAN L., 2001: Random Forests. *Machine Learning*, Volume 45, S5–32, DOI 10.1023/A:1010933404324
- [43] CAI Y., ZIJLEMA W. L., SØRGJERD E. P., DOIRON D., de HOOGH K., HODGSON S., WOLFFENBUTTEL B., GULLIVER J., HANSELL A. L.,

- NIEUWENGUISEN M., RAHIMI K., KVALØY K.**, 2020: Impact of road traffic noise on obesity measures: Observational study of three European cohorts. Environmental Research, Volume 191, DOI: 10.1016/j.envres.2020.110013.
- [44] **CollGS**, 2024a: Sentinel-5P. Koordinační rady ministra dopravy pro kosmické aktivity. (online) [cit.2024.01.12], dostupné z: <<https://collgs.czechspaceportal.cz/dalsi-mise-copernicus/sentinel-5/>>
- [45] **CollGS**, 2024b: Sentinel-3. Koordinační rady ministra dopravy pro kosmické aktivity. (online) [cit.2024.01.12], dostupné z: <<https://collgs.czechspaceportal.cz/flotila-druzic-sentinel/sentinel-3/>>
- [46] **den BOER L. C., SCHROten A.**, 2007: Traffic noise reduction in Europe. T&E Brussels: CE Delft (online) [cit.2024.04.03], dostupné z <https://www.transportenvironment.org/wp-content/uploads/2021/05/2008-02_traffic_noise_ce_delft_report.pdf>.
- [47] **Encyklopédie fyziky**, 2024: Akustika a její dělení (online) [cit.2024.03.08], dostupné z <http://fyzika.jreichl.com/main.article/view/214-akustika-a-jeji-deleni>
- [48] **EOG**, 2021: Products. VNL. Earth Observation Group. (online) [cit.2024.06.17], dostupné z <<https://eogdata.mines.edu/products/vnl/>>.
- [49] **EEA**, 2023: Noise. European Environment Agency. (online) [cit.2024.03.08], dostupné z <<https://www.eea.europa.eu/en/topics/in-depth/noise>>.
- [50] **EEA**, 2020a: Environmental noise in Europe -2020. European Environment Agency. (online) [cit.2024.03.08], dostupné z <<https://www.eea.europa.eu/publications/environmental-noise-in-europe>>.
- [51] **EEA**, 2020b: Hlukové znečištění představuje závažný problém, a to jak pro lidské zdraví, tak pro životní prostředí. European Environment Agency. (online) [cit.2024.04.03], dostupné z <https://www.eea.europa.eu/ds_resolveuid/a79db279ea56476d8de032b1a63c70f3>.
- [52] **EOS Data Analytics**, 2024a: NDVI: Normalized Difference Vegetation Index. (online)[cit.2024-11-13], dostupné z <<https://eos.com/make-an-analysis/ndvi/>>
- [53] **ESA**, ESA_a: Sentinel-5P, Facts and figures. (online) [cit.2024.04.25], dostupné z: <https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-5P/Facts_and_figures>

- [54] **ESA**, ESA_b: Sentinel-2, Facts and figures. (online) [cit.2024.06.17], dostupné z:
[<https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2/Facts_and_figures>](https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2/Facts_and_figures)
- [55] **ESA**, 2018: Copernicus Sentinel-5P ozone boosts daily forecasts. (online) [cit.2024.10.22], dostupné z: <
[<https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-5P/Copernicus_Sentinel-5P_ozone_boosts_daily_forecasts>](https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-5P/Copernicus_Sentinel-5P_ozone_boosts_daily_forecasts)
- [56] **ESA**, 2023a: Trio of Sentinel satellites map methane super-emitters. (online) [cit.2024.10.22], dostupné z: <
[<https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Trio_of_Sentinel_satellites_map_methane_super-emitters>](https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Trio_of_Sentinel_satellites_map_methane_super-emitters)
- [57] **ESA**, 2023b: Sentinel-5P data map nitrogen dioxide in Finland. (online) [cit.2024.10.22], dostupné z:
[<https://sentinels.copernicus.eu/fi/web/success-stories/-/sentinel-5p-data-map-nitrogen-dioxide-in-finland>](https://sentinels.copernicus.eu/fi/web/success-stories/-/sentinel-5p-data-map-nitrogen-dioxide-in-finland)
- [58] **ESA**, 2024a: S2 Products. (online) [cit.2024.11.14], dostupné z:
[<https://sentiwiki.copernicus.eu/web/s2-products>](https://sentiwiki.copernicus.eu/web/s2-products)
- [59] **EUR-Lex**, 2023: Future Noise Policy, European Commission Green Paper (online) [cit.2023.01.24], dostupné z:
[<https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwju6Y6e09_8AhVdQ_EDHbHrAPwQFnoECBqQAQ&url=https%3A%2F%2Feur-lex.europa.eu%2FLexUriServ%2FLexUriServ.do%3Furi%3DCOM%3A1996%3A0540%3AFIN%3AEN%3APDF&usq=AOvVaw1aKw-rA5ocbU1L1fNDYcQH>.](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwju6Y6e09_8AhVdQ_EDHbHrAPwQFnoECBqQAQ&url=https%3A%2F%2Feur-lex.europa.eu%2FLexUriServ%2FLexUriServ.do%3Furi%3DCOM%3A1996%3A0540%3AFIN%3AEN%3APDF&usq=AOvVaw1aKw-rA5ocbU1L1fNDYcQH)
- [60] **EICHER L., MOMMERT M., BORTH D.**, 2022: Traffic Noise Estimation from Satellite Imagery with Deep Learning. 5937-5940. DOI: 10.1109/IGARSS46834.2022.9883463
- [61] **EVEREST F.A., POHLMANN K.C.**, 2009: Master Handbook od Acoustics. The McGraw-Hill Companies, Inc.. California, 510 s.
- [62] **FARIA A., FONSECA P. J., VIEIRA M., ALVES L. M. F., LEMOS M. F. L., NOVAIS S. C., MATOS A. B., VIEIRA D., AMORIM M. C. P.**, 2022: Boat noise impacts early life stages in the Lusitanian toadfish: A field experiment. Science of The Total Environment, Volume 811. DOI: 10.1016/j.scitotenv.2021.151367.

- [63] **GOODFELLOW I., BENGIO Y., COURVILLE A.**, 2016: Deep Learning. MIT Press. (online), dostupné z: <<https://www.deeplearningbook.org/>>
- [64] **GUHA A. K., GOKHALE S.**, 2023: Urban workers' cardiovascular health due to exposure to traffic-originated PM2.5 and noise pollution in different microenvironments. *Science of The Total Environment*, Volume 859, Part 1, DOI: 10.1016/j.scitotenv.2022.160268.
- [65] **HONG J., KANG H., HONG T., PARK, H. S., LEE D.**, 2022: Construction noise rating based on legal and health impacts. *Automation in Construction*, Volume 134, DOI: 10.1016/j.autcon.2021.104053.
- [66] **CHA G.W., MOON H.J., KIM Y.C.**, 2021: Comparison of Random Forest and Gradient Boosting Machine Models for Predicting Demolition Waste Based on Small Datasets and Categorical Variables. *Int J Environ Res Public Health*. Volume 18(16). DOI: 10.3390/ijerph18168530.
- [67] **JAMES G., WITTEN D., HASTIE T., TIBSHIRANI R.**, 2021: An Introduction to Statistical Learning with Applications in R. 2nd Edition. Springer (online)[cit.2024-11-25], dostupné z: <<https://www.statlearning.com/>>
- [68] **KAUR, R., PANDEY, P. A.**, 2022: A review on spectral indices for built-up area extraction using remote sensing technology. *Arab J Geosci* 15, 391. DOI: 10.1007/s12517-022-09688-x
- [69] **KEPHALOPOULOS S., PAVIOTTI M., ANFOSSO-LÉDÉE F.**, 2012: Common Noise Assessment Methods in Europe (CNOSSOS-EU) (online) [cit.2024.04.03], dostupné z <<https://publications.jrc.ec.europa.eu/repository/handle/JRC72550>>.
- [70] **KOI H., FUJITA S-I., AOI I., TAMURA H. & HIRAMATSU K.**, 1993: Regional noise prediction and management system using remote sensing. *International Journal of Remote Sensing*, Volume 14, Number 13, s.2427-2443, DOI: 10.1080/01431169308904284
- [71] **KŘIVÁNEK V., HABLOVIČOVÁ B.**, 2002-2024: Terénní ověření snížení hlučnosti různými typy protihlukových stěn v ČR. (online) [cit.2024.04.03], dostupné z <<https://silnice-zeleznice.cz/zajimavosti/terenni-overeni-snizeni-hlucnosti-ruznymi-typy-protihlukovych-sten-v-cr-674>>.
- [72] **LIMA N. D. A. P., REBOUÇAS L. F. T., LISBOA C. S.**, 2022: Influence of urban noise in call traits of the Atlantic Forest treefrog *Bokermannohyla hylax*. *Zoologischer Anzeiger*, Volume 300. S 41–46. DOI: 10.1016/j.jcz.2022.07.002.
- [73] **LIU Y., GOUDREAU S., OIAMO T., RAINHAM D., HATZOPOULOU M., CHEN H., DAVIES H., TREMBLAY M., JOHNSON J., BOCKSTAEL A.**,

- LEROUX T., SMARGIASSI A.**, 2020: Comparison of land use regression and random forests models on estimating noise levels in five Canadian cities. Environmental Pollution, Volume 256. DOI: 10.1016/j.envpol.2019.113367.
- [74] **LU L., GUO H., WANG C., PESARESI M.**, 2014: Monitoring bidecadal development of urban agglomeration with remote sensing images in the Jing-Jin-Tang area, China. Journal of Applied Remote Sensing, DOI: 10.1117/1.JRS.8.084592.
- [75] **MARTIN M., GRIDLEY T., ELWEN S. H., CHARRIER I.**, 2020: Assessment of the impact of anthropogenic airborne noise on the behaviour of Cape fur seals during the breeding season in Namibia. Journal of Experimental Marine Biology and Ecology, Volume 550. DOI: 10.1016/j.jembe.2022.151721.
- [76] **MDČR**, 2024: Akční plány ke snížení hluku z dopravy (online) [cit.2024.04.03], dostupné z <<https://www.mdcr.cz/Dokumenty/Strategie/Hluk/Akcní-plany-ke-snizení-hluku-z-dopravy>>.
- [77] **MEDIUM**, 2018: Understanding Gradient Boosting Machines (online)[cit.2025.01.16], dostupné z <<https://towardsdatascience.com/understanding-gradient-boosting-machines-9be756fe76ab>>
- [78] **MECHLOVÁ E. a kol.**, 1999: Výkladový slovník fyziky pro základní vysokoškolský kurz fyziky. Prometheus. Praha, 588 s.
- [79] **MOHAMMED M.U., BADAMASI M.M., USMAN F., ZANGO Z.U., DENNIS J.O., ALJAMEEL A.I., MOHAMMED Ali M.K., ALDAGHRI O.A., IBNAOUF K.H., HUSSEIN T.M.**, 2022: Towards Urban Sustainability: Developing Noise Prediction Model in an Informal Setting. Applied Sciences, Volume 12, Issue 18, DOI: 10.3390/app12189071.
- [80] **MZČR**, 2007: Zpráva o zpracování Strategické hlukové mapy ČR (online) [cit.2024.04.03], dostupné z <https://www.mzcr.cz/wp-content/uploads/wepub/11088/24694/Zprava_o_zpracovani_Strategicke_hlukove_mapy_CR.pdf>.
- [81] **MZČR**, 2015a: Hlukové mapy (online) [cit.2024.03.08], dostupné z <<https://www.mzcr.cz/category/ochrana-verejneho-zdravi/hlukove-mapy/popis-shm/>>.
- [82] **MZČR**, 2015b: Nepříznivé účinky hluku na člověka (online) [cit.2024.04.03], dostupné z <<https://www.mzcr.cz/nepriznive-ucinky-hluku-na-cloveka/>>.

- [83] **MZČR**, 2015c: Nepříznivé ovlivnění spánku (online) [cit.2024.04.03], dostupné z <<https://www.mzcr.cz/nepriznive-ovlivneni-spanku/>>.
- [84] **MZČR**, 2015d: Hlukové indikátory (online) [cit.2024.04.03], dostupné z <<https://www.mzcr.cz/hlukove-indikatory/>>
- [85] **MZČR**, 2015e: Cykly (kola) SHM (online) [cit.2024.03.08], dostupné z <<https://www.mzcr.cz/cykly-kola-shm/>>
- [86] **MZČR**, 2015f: Základní popis 1. kola SHM (online) [cit.2024.04.03], dostupné z <<https://www.mzcr.cz/zakladni-popis-1-kola-shm/>>
- [87] **MZČR**, 2015g: Základní popis 2. kola SHM (online) [cit.2024.04.03], dostupné z <<https://www.mzcr.cz/zakladni-popis-2-kola-shm/>>
- [88] **MZČR**, 2017: Metodický návod pro měření a hodnocení hluku v mimopracovním prostředí (online) [cit.2022.06.08], dostupné z <<https://www.mzcr.cz/vestnik/vestnik-c-11-2017/>>.
- [89] **MZČR**, 2022: Geoportál MZČR: Hlukové mapy 2022 (online) [cit.2024.02.16], dostupné z <<https://geoportal.mzcr.cz/shm/?locale=cs>>.
- [90] **NASA**, NASA_a: Visible Infrared Imaging Radiometer Suite (VIIRS) (online)[cit.2025.02.10], dostupné z <https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/viirs/?utm_source=chatgpt.com>
- [91] **NASA**, NASA_b: VIIRS Overview(online)[cit.2025.02.10], dostupné z: <https://lpdaac.usgs.gov/data/get-started-data/collection-overview/missions/s-npp-nasa-viirs-overview/?utm_source=chatgpt.com>
- [92] **NASA**, 2024a: Landsat 8 (online)[cit.2024.06.17], dostupné z <<https://landsat.gsfc.nasa.gov/satellites/landsat-8/>>
- [93] **NASA**, 2024b: Landsat Next (online)[cit.2024.11.13], dostupné z <<https://landsat.gsfc.nasa.gov/satellites/landsat-next/>>
- [94] **NASA**, 2024c: MODIS (online)[cit.2024.11.13], dostupné z <<https://modis.gsfc.nasa.gov/about/>>
- [95] **NASA**, 2024d: MODIS Land Surface Temperature and Emissivity (MOD11) (online)[cit.2024.11.14], dostupné z <<https://modis.gsfc.nasa.gov/data/dataproduct/mod11.php>>
- [96] **PASCHALIDOU A. K., KASSOMENOS P., CHONIANAKI F.**, 2018: Strategic Noise Map and Action Plans for the reduction of population exposure in a Mediterranean port city. *Science of the Total Environment*, S. 144–153. DOI: 10.1016/J.SCITOTENV.2018.11.048.
- [97] **The Payne Institute for Public Policy**, 2024: Earth Observation Group (online)[cit.2024.10.22], dostupné z <<https://payneinstitute.mines.edu/eog/>>

- [98] **RAESS, M., BRENTANI, A., De ANTAS De CAMPOS, BL, FLÜCKIGER, B., De HOOGH, K., FINK, G., a RÖÖSLI, M.**, 2021: Land use regression modelling of community noise in São Paulo, Brazil. Environmental Research, Volume 199, DOI: 10.1016/j.envres.2021.111231
- [99] **RONNEBERGER O., FISCHER P., BROX T.**, 2015: U-Net: Convolutional Networks for Biomedical Image Segmentation. ArXiv, abs/1505.04597., DOI: 10.48550/arXiv.1505.04597.
- [100] **SAP**: Co je strojové učení? (online)[cit.2025.01.16], dostupné z <<https://www.sap.com/cz/products/artificial-intelligence/what-is-machine-learning.html>>
- [101] **SHIM E., KIM D., WOO H., CHO Y.**, 2016: Designing a Sustainable Noise Mapping System Based on Citizen Scientists Smartphone Sensor Data. PLoS ONE 11(9), DOI: 10.1371/journal.pone.0161835
- [102] **SoftNoise**, 2024: Predictor-LiaA (online)[cit.2024.03.22], dostupné z <<https://softnoise.com/products/predictor-lima/>>
- [103] **STAAB J., SCHADY A., WEIGAND M. et al.**, 2021: Predicting traffic noise using land-use regression—a scalable approach. J Expo Sci Environ Epidemiol 32, S. 232–243, DOI: 10.1038/s41370-021-00355-z
- [104] **SUDOP PRAHA**, 2019: Optimalizace trati Černošice (včetně) – odb. Berounka (mimo). Akustická studie, měření hluku a vibrace. (nepublikováno).
- [105] **SZÚ**, 2023a: Zdroje hluku a jak se měří (online) [cit.2024.03.08], dostupné z <<https://szu.cz/temata-zdravi-a-bezpecnosti/zivotni-prostredi/hluk/zdroje-hluku-a-jak-se-meri/>>.
- [106] **SZÚ**, 2023b: Formaldehyd. (online) [cit.2024.10.22], dostupné z <<https://szu.cz/wp-content/uploads/2023/03/Formaldehyd.pdf>>
- [107] **SZÚ**, 2023-2024: Oxid uhelnatý (online) [cit.2024.10.22], dostupné z <<https://szu.cz/odborna-centra-a-pracoviste/centrum-zdravi-a-zivotniho-prostredi/oddeleni-hygieny-ovzdusi-2/vnitri/oxid-uhelnaty/>>.
- [108] **TIAN Y., CHEN H., SONG Q., ZHENG K.**, 2018: A Novel Index for Impervious Surface Area Mapping: Development and Validation. Remote Sensing, Volume 10, Issue 10, 1521. DOI: 10.3390/rs10101521
- [109] **VANDASOVÁ Z., FIALOVÁ A.**, 2016: Vztahy mezi hlukovými ukazateli Ldn a Ldn. (online) [cit.2022.08.01⁴⁷], dostupné

⁴⁷ Článek nebyl k 3.4.2024 již na nových stránkách Státního zdravotního ústavu nalezen

z <<http://www.szu.cz/tema/zivotni-prostredi/vztahy-mezi-hlukovymi-ukazateli-ldvn-a-ldn>>.

- [110] **WHO**, 2024: Noise (online) [cit.2024.04.03], dostupné z <<http://www.who.int/europe/news-room/fact-sheets/item/noise>>.
- [111] **WOODS M. B., BROWN N. A. W., NIKOLICH K., HALLIDAY W. D., BALSHINE S., JUANES F.**, 2022: Context-dependent effects of anthropogenic noise on nest defence in a singing toadfish. *Animal Behaviour*, Volume 191. S 105–115. DOI: 10.1016/j.anbehav.2022.06.018.
- [112] **XIAOBIAO W., SHUNPING X., XUELIANG Z., CHENG Ch., HAO G., JINKANG D., ZHENG D.**, 2018: A robust Multi-Band Water Index (MBWI) for automated extraction of surface water from Landsat 8 OLI imagery, *International Journal of Applied Earth Observation and Geoinformation*, Volume 68, S 73-91, DOI: 10.1016/j.jag.2018.01.018.
- [113] **ZAMBON G., ROMAN H., SMIRAGLIA M., BENOCCI R.** 2018: Monitoring and Prediction of Traffic Noise in Large Urban Areas. *Applied Sciences*. Vol. 8. S 251. DOI: 10.3390/app8020251.
- [114] **Zdravotní ústav se sídlem v Ostravě**, 2022: Závěrečná zpráva. Strategické hlukové mapy aglomerací. IV. kolo. 2022. (online) [cit.2024.03.08], dostupné z <<https://geoportal.mzcr.cz/portal/apps/storymaps/stories/e51bc98c0bc841d0b17e244c67659be7>>.
- [115] **Zdravotní ústav se sídlem v Ostravě**, 2020: Oxid siřičitý (SO₂). (online) [cit.2024.10.22], dostupné z <<https://www.zuova.cz/home/Clanek/oxid-siricity-so2>>.
- [116] **ZHA Y., GAO J., NI S.** 2005: Use of normalized difference built-up index in automatically mapping urban areas from TM imagery, *International Journal of Remote Sensing*, Vol.24, No.3, 2005, pp. 583-594

Seznam obrázků

- Obrázek 1 - Přehled kol Strategického hlukového mapování (zdroj: vlastní zpracování podle MZČR, 2015e).
- Obrázek 2 - Mise Landsat (zdroj: <https://landsat.gsfc.nasa.gov/satellites/timeline/>)
- Obrázek 3 – Landsat 8 rozložení bandů (zdroj: <https://blogs.uberlin.de/reseda/landsat-8/>)
- Obrázek 4 - Porovnání bandů u Landsat 7 a 8 se Sentinel 2 (zdroj: <https://landsat.gsfc.nasa.gov/article/sentinel-2a-launches-our-compliments-our-complements/>)
- Obrázek 5 - Detektovatelnost plynů ze Sentinel 5P (TROPOMI) (zdroj: <https://www.eoportal.org/satellite-missions/copernicus-sentinel-5p#tropomi-tropospheric-monitoring-instrument>)
- Obrázek 6 – Vennův diagram ukazující propojení hlubokého učení (DL), strojového učení (ML) a umělé inteligence (AI) (zdroj: <https://www.deeplearningbook.org/contents/intro.html>)
- Obrázek 7 - Překrývání dlaždic pro segmentaci velkých obrazů. Žlutá oblast je predikce segmentace, vyžadující obrazový vstup v modré oblasti. (Zdroj: Ronneberger et al., 2015)
- Obrázek 8 – Pokrytí SHM 2022 na území České republiky (zdroj: vlastní zpracování v ArcGIS Pro)
- Obrázek 9 – Pokrytí SHM 2022 Aglomerace na území České republiky (zdroj: vlastní zpracování v ArcGIS Pro).
- Obrázek 10– Pokrytí SHM 2022 Silnice na území České republiky (zdroj: vlastní zpracování v ArcGIS Pro)
- Obrázek 11 – Pokrytí SHM 2022 Železnice na území České republiky (zdroj: vlastní zpracování v ArcGIS Pro)
- Obrázek 12 – Pokrytí SHM 2022 Letiště na území České republiky (zdroj: vlastní zpracování v ArcGIS Pro)
- Obrázek 13 - Výstup z GEE pro vypočtený index NDVI Max, oblast Praha (zdroj: vlastní zpracování v Google Colab)
- Obrázek 14 - Výstup z GEE pro vypočtený index NDBI, oblast Praha (zdroj: vlastní zpracování v Google Colab)
- Obrázek 15 - Výstup z GEE pro vypočtený index MNDWI, oblast Praha (zdroj: vlastní zpracování v Colab)
- Obrázek 16 - Výstup z GEE pro vypočtený index UI, oblast Praha (zdroj: vlastní zpracování v Colab)
- Obrázek 17 - Výstup z GEE pro vypočtený index ISA, oblast Praha (zdroj: vlastní zpracování v Colab)
- Obrázek 18 - Výstup z GEE pro data TIR v Kelvinech, oblast Praha (zdroj: vlastní zpracování v Colab)
- Obrázek 19 – Výstup z GEE pro data z MODIS, oblast ČR (zdroj: vlastní zpracování v Colab)

Obrázek 20 – Výstup z GEE pro data z MODIS, oblast Praha (zdroj: vlastní zpracování v Colab)

Obrázek 21 - Zobrazení plynu CO na území ČR za rok 2022 (Zdroj: <https://maps.s5p-pal.com/co/year/>).

Obrázek 22 - Zobrazení plynu HCHO na území ČR za rok 2022 (zdroj: <https://maps.s5p-pal.com/hcho/year/>).

Obrázek 23 - Zobrazení plynu NO₂ na území ČR za rok 2022 (zdroj: <https://maps.s5p-pal.com/no2/year/>).

Obrázek 24 - Zobrazení plynu O₃ na území ČR za rok 2022 (zdroj: <https://maps.s5p-pal.com/o3/year/>).

Obrázek 25 - Zobrazení plynu SO₂ na území ČR za vybraný den roku 2022 (zdroj: <https://maps.s5p-pal.com/so2-cobra-pbl/day/>).

Obrázek 26 - Zobrazení plynu CH₄ na území ČR za celý rok 2022 (zdroj: <https://maps.s5p-pal.com/ch4/>).

Obrázek 27 - Zobrazení světelného znečištění z VIIRS (celoroční mozaika) na území České republiky (zdroj: vlastní zpracování v ArcGIS Pro)

Obrázek 28 - Pokrytí území České republiky náhodnými body (zdroj: vlastní zpracování v ArcGIS Pro)

Obrázek 29 - Histogramy všech proměnných včetně transformovaných proměnných (zdroj: vlastní zpracování v R Studio)

Obrázek 30 - Boxploty proměnných (zdroj: vlastní zpracování v R Studio)

Obrázek 31 – Jitter plot s Boxplot pro proměnné s odlehlými hodnotami (zdroj: vlastní zpracování v R Studio)

Obrázek 32 - Korelační matice vybraných proměnných (zdroj: vlastní zpracování v R Studio)

Obrázek 33 - Výsledky trénování RF modelu (zdroj: vlastní zpracování v R Studio)

Obrázek 34 - Výsledky trénování modelu po sloučení kategorií hluku (zdroj: vlastní zpracování v R Studio)

Obrázek 35 - Relativní důležitost proměnných pro model Gradient Boosting (zdroj: vlastní zpracování v RStudio)

Obrázek 36 - Výsledek modelu neuronové sítě pro strukturu (5, 3) (zdroj: vlastní zpracování v R Studio)

Seznam tabulek

Tabulka 1 - Mezní hodnoty hlukových ukazatelů (zdroj: Miková podle vyhlášky č. 315/2018 Sb.).

Tabulka 2 - Rychlosť šírenia hľuku v rôznych pevných látkach a plynech (zdroj: Beran, 2010)

Tabulka 3 - Prostorové rozlišení vstupných dat

Tabulka 4 - Statistika modelu pri použití Random Forest (zdroj: vlastní zpracování v RStudio)

Tabulka 5 - Statistika modelu pri použití Gradient Boosting (zdroj: vlastní zpracování v RStudio)

Tabulka 6 - Statistika modelu s využitím neuronovej sítě (zdroj: vlastní zpracování v RStudio)

Tabulka 7 - Tabuľka souhrnu výsledkov použitých metod RF, GB a NN (zdroj: vlastní zpracovanie v MS Excel)

Seznam rovnic

Rovnica 1 - Vzorec pro výpočet hlukového indikátoru L_{den} podle smernice 2002/49/EC

Rovnica 2 - Výpočet NDVI (vegetační index)

Rovnica 3 - NDBI (index zástavby)

Rovnica 4 - MNDWI (modifikovaný vodní index)

Rovnica 5 - UI (index urbanizačních ploch)

Rovnica 6 - ISA (index nepropustnosti povrchů)

Rovnica 7 - Převod Digital Number na hodnoty v kelvinech

Seznam príloh

Pričočka 1: Python skript pre staženie satelitných dát.

Pričočka 2: R skript práce s dátami a tvorba modelu