



ŽIVOTNÍ PROSTŘEDÍ – PROSTŘEDÍ PRO ŽIVOT – 2024

Představení interaktivního znalostního portálu SoilPass pro informace o hygienickém stavu zemědělských půd v ČR

Mgr. Jan Skála, Ph.D.

Výzkumný ústav monitoringu a ochrany půdy, v.v.i.

Číslo a název projektu: SS03010364 Systém na podporu rozhodování při hodnocení kvality půdy z hlediska obsahu rizikových látek v zemědělských půdách České republiky

Prioritní téma programu: ochrana vody, půdy, horninového prostředí a dalších přírodních zdrojů

Klíčová slova: půda, digitální mapování půd, znečištění, strojové učení, znalostní portál

Prezentace a praktická ukázka v příspěvku představí stěžejní výstup projektu SS03010364 v podobě interaktivního znalostního portálu „SoilPass = „Soil Pollution Assessment“ (<https://soilpass.vumop.cz/>), který je výsledkem prostorového prediktivního modelování. Naprogramované uživatelské rozhraní integruje postup, jak ze zdrojových dat dostupných v ČR získat požadované a srozumitelné informace o předpokládaném hygienickém stavu půdy, tedy simuluje expertní průzkum a využívá existující datové zdroje v ČR k poskytnutí srozumitelného výstupu, aniž by aplikace poskytovala individuální data jednotlivých gestorů. Zároveň se interaktivní portál inovativním způsobem vypořádává s poskytnutím kompletních informací o výstupech z prediktivního prostorového modelu včetně práce s nejistotou vlastních výstupů, což není běžná praxe při zveřejnění mapových výstupů. Pro každý prediktivní povrch obsahu rizikových látek/prvků je k dispozici stažitelný metainformační záznam, kde jsou zpracovány ucelené informace o vstupních datech, postupu tvorby a vyhodnocení prediktivního modelu, což umožňuje reprodukování postupu tvorby map. Při tvorbě map byl používán algoritmus kvantilové formy náhodných lesů ze skupiny metod strojového učení, jehož výhodou je schopnost odhadnout kvantily modelované proměnné pro všechna místa predikce, a tím získat i distribuční parametry pravděpodobné hodnoty cílové proměnné. Proto použitý model umožňuje lokální hodnocení spolehlivosti modelu, pomocí šířky predikčního intervalu, která doplňuje celkovou míru přesnosti modelu kvantifikovanou z rozdílů predikovaných a skutečně naměřených hodnot v souboru dat odejmutých z primárního vzorku před vlastním trénováním modelu. Principem algoritmů strojového učení je automatické extrakce („naučení se“) cenné informace z trénovacích dat, a naučené vzory aplikovat na prostorové prediktivní modelování mimo trénovací data. V případě (geo)chemického prediktivního mapování jsou tyto vzory odvozovány na základě vztahu mezi půdními obsahy prvků/látek a dalšími popisnými proměnnými životního prostředí postihující půdně-geologické, klimatické, geomorfologické podmínky či parametrizují zdroje a vstupy prvků a cizorodých látek do prostředí. Výše uvedeným algoritmem byly natrénovány prediktivní modely nejen pro odhad koncentrací prvků/látek v půdě, ale také pravděpodobnostně indikátorové modely překročení relevantních limitních hodnot pro obsahy v půdě. Vytvořené interaktivní rozhraní pro zpřístupnění digitálních map podstatným způsobem přispívá ke zvýšení současné úrovně znalostí o hygienickém stavu půd, neboť srovnatelný informační nástroj prozatím nebyl v České republice vytvořen.