



VUMOP

VÝZKUMNÝ ÚSTAV MONITORINGU
A OCHRANY PŮDY, v.v.i.

Jan Skála

Představení interaktivního znalostního portálu SoilPass

II. část – praktická ukázka práce s nástroji portálu



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

HODNOCENÍ ZNEČIŠTĚNÍ PŮDY

VSTUP DO APLIKACE

Vítejte v internetové aplikaci HODNOCENÍ ZNEČIŠTĚNÍ PŮDY (SoilPass = „Soil Pollution Assessment“) pro hodnocení hygienického stavu zemědělské půdy na Vámi zvolené lokalitě. Cílem je poskytnout věrohodné informace o obsahu rizikových látek v zemědělské půdě jako podklad pro přiměřenou ochranu ZPF v návaznosti na existující legislativní kritéria znečištění, která určují oprávněné legislativní požadavky pro regulaci vstupu rizikových látek na straně hygienické jakosti jak aplikovaných materiálů, tak cílových půd pro aplikaci. Naplnění požadavků lze objektivně stanovit pouze analyticky, proto tato aplikace poskytuje metodickou a informační podporu pro potencionální identifikaci vhodných zemědělských pozemků plnících s dostatečnou pravděpodobností legislativní požadavky. Výstupy analýz jsou výsledkem propojení existujících analytických výsledků s progresivními predikčními nástroji a jsou pouze informativní – tj. slouží pouze pro indikaci pravděpodobného stavu a optimalizaci výběru vhodných ploch.

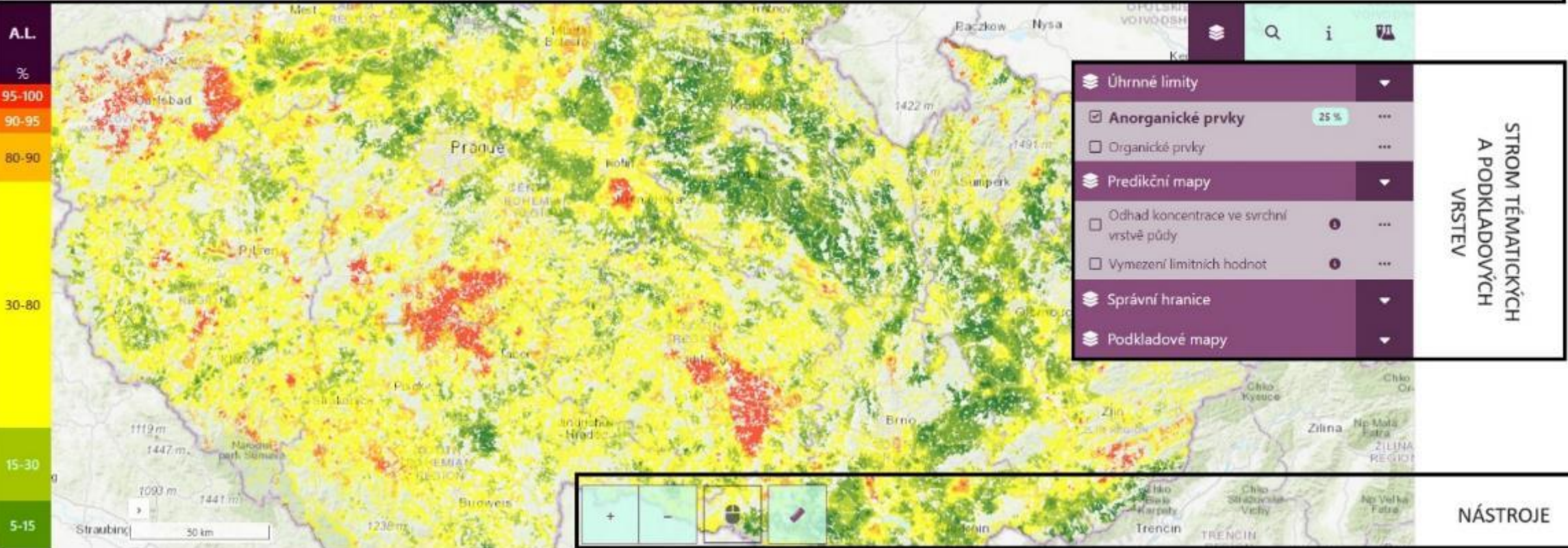
<https://soilpass.vumop.cz>

Interaktivní mapová aplikace

- mapová aplikace je založena na grafickém rozhraní, které vychází z uživatelsky ověřené podoby v rámci geoportálu VÚMOP
- základními ovládacími prvky interaktivního prostředí aplikace jsou:
 - strom tematických a podkladových vrstev,
 - interaktivní lišta,
 - panel nástrojů,
 - informační panel.



AN. LÁTKY	arsen (As)	beryllium (Be)	kadmium (Cd)	kobalt (Co)	chrom (Cr)	měď (Cu)	rtuť (Hg)	nikl (Ni)	olovo (Pb)	vanad (V)	zinek (Zn)	INTERAKTIVNÍ LIŠTA
Org. látky	dichlordifenyl trichlorethan (DDT)		hexachlor cyklohexan (HCH)		polycyklické aromatické uhlovodíky (PAU)	polychlorované bifenyly (PCB)		polychlorované dibenzodioxiny (PCDD/F)		hexachlorbenzen (HCB)		



Interaktivita mapové aplikace

- lokalizace zájmového bodu kurzorem
- rychlé „skenování“ prediktivního povrchu
- místopisné vyhledávání
- po lokalizaci bodu kurzorem – rychlá kategorizace zájmových prvků na interaktivní liště pomocí ikon kategorizují prvky/látky dle potřeby pozornosti
- bližší informace o odhadu koncentrací a pravděpodobnosti překročení limitních hodnot lze získat v informačním panelu
- možnost změny zobrazeného povrchu volbou na interaktivní liště mezi jednotlivými parametry, zatímco zvolený bod zůstává zafixován a na informačním panelu dostává uživatel detailní informace o vždy zvoleném prvku/látce
- rozšířená práce s informačním panelem

- An. látky arsen (As) beryllium (Be) kadmium (Cd) kobalt (Co) **CHROM (CR)** měď (Cu) rtuť (Hg) nikl (Ni) olovo (Pb) vanad (V) zinek (Zn)
- Org. látky dichlordifenyl trichlorethan (DDT) hexachlor cyklohexan (HCH) polycyklické aromatické uhlovodíky (PAU) polychlorované bifenylly (PCB) polychlorované dibenzodioxiny (PCDD/F) hexachlorbenzen (HCB)



Odhad koncentrace v půdě

Chrom

180 mg/kg

S pravděpodobností 99,23 % je překročen preventivní limit.
více o preventivních limitech na webu MZe ČR

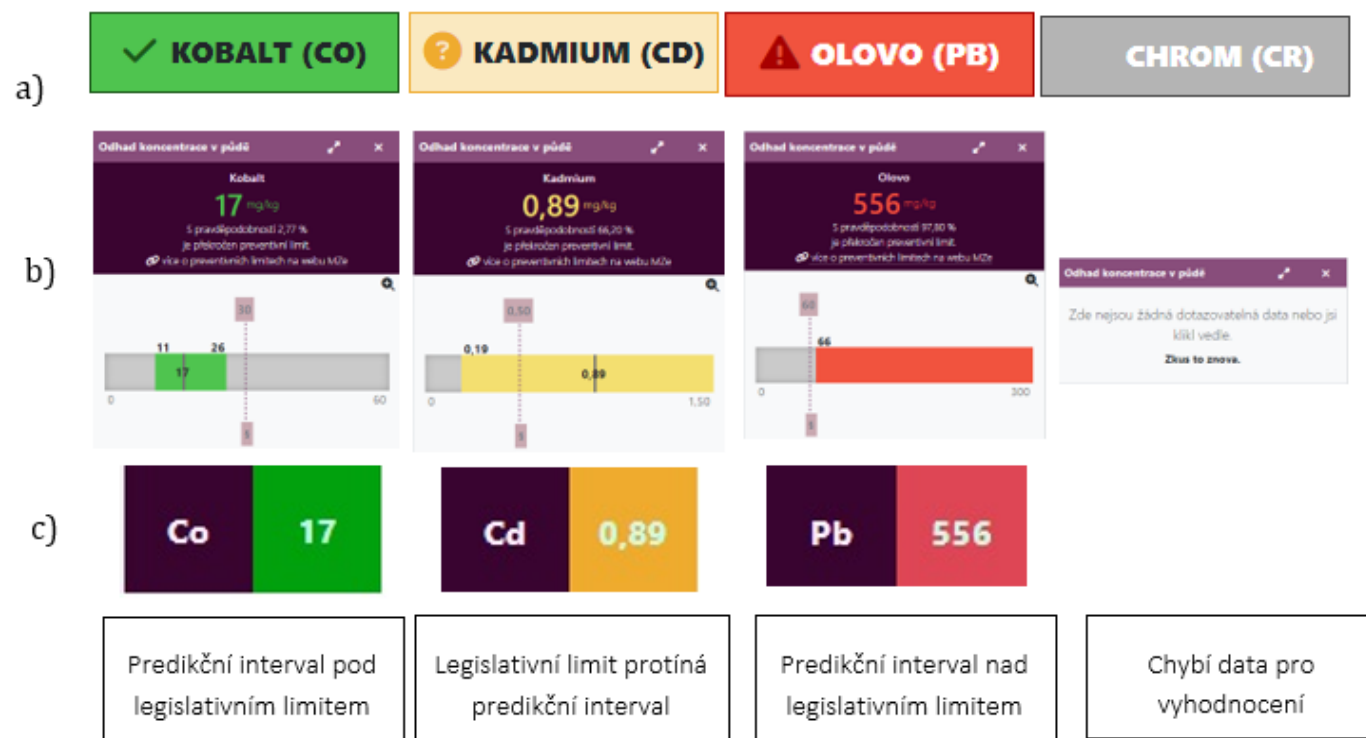
Pravděpodobnost Referenční hodnoty

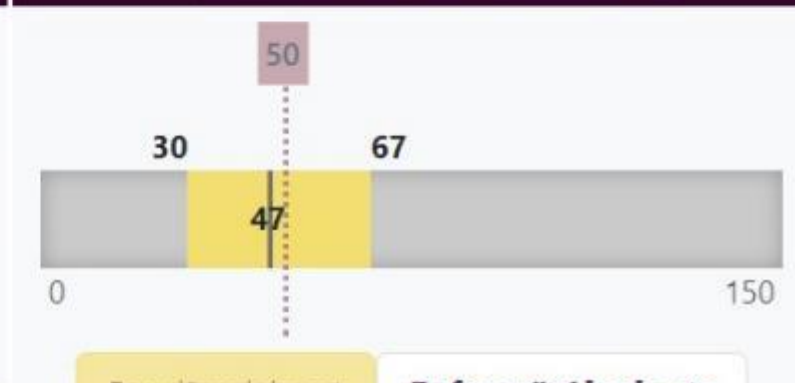
Spodní hranice predikčního intervalu je **97** mg/kg. Střední hodnota predikce, která udává nejpravděpodobnější modelový odhad koncentrace je **180** mg/kg. Horní hranice

INFORMAČNÍ PANEL



Sémantický význam barev a ikon, který je zachován ve všech částech aplikace - interaktivní lišta - a), mapový panel - b), dynamická legenda - c)





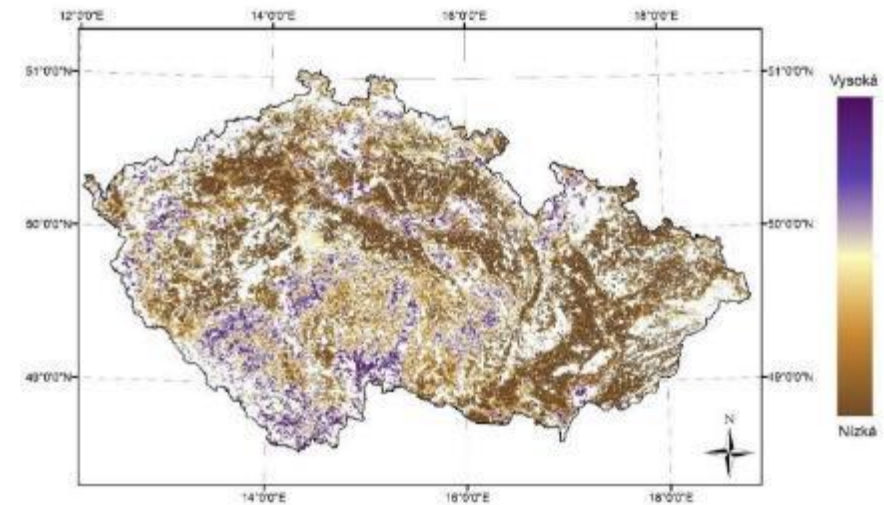
Spodní hranice predikčního intervalu je **30 mg/kg**. Střední hodnota predikce, která udává nejpravděpodobnější modelový odhad koncentrace je **47 mg/kg**. Horní hranice predikčního intervalu je **67 mg/kg**.

Legislativní limit v místě je **50 mg/kg**.

Predikční interval překrývá legislativní limit.
[Více zde.](#)

Pravděpodobnost	Referenční hodnoty
Preventivní limit pro lehké půdy	45 mg/kg
Preventivní limit pro běžné půdy	50 mg/kg
Svrchní mez statistické variability koncentrací v ČR	38 mg/kg
Lokální svrchní mez variability koncentrací	84 mg/kg
Pravděpodobnost příslušnosti do třídy lehkých půd	32 %

Pravděpodobnost, že půda má obsah částic o velikosti < 0,01 mm nižší než 20 % - modelováno na základě dat KPP (základní sondy)



Tab. 1 Diferenciace preventivních limitních hodnot zinku podle půdní textury dle Vyhlášky č. 153/2016 Sb.

	As	Be	Cd	Co	Cr	Cu	Hg	Ni	Pb	V	Zn
Běžné půdy	20	2.0	0.5	30	90	60	0,3	50	60	130	120
Lehké půdy	15	1.5	0.4	20	55	45	0,3	45	55	120	105

Odhad koncentrace v půdě

Nikl

47 mg/kg

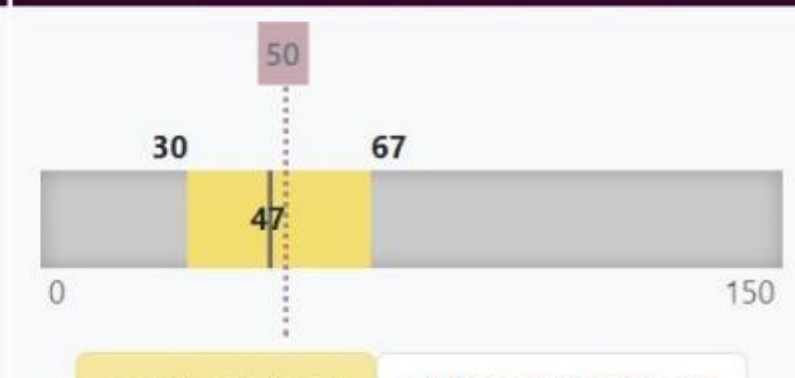
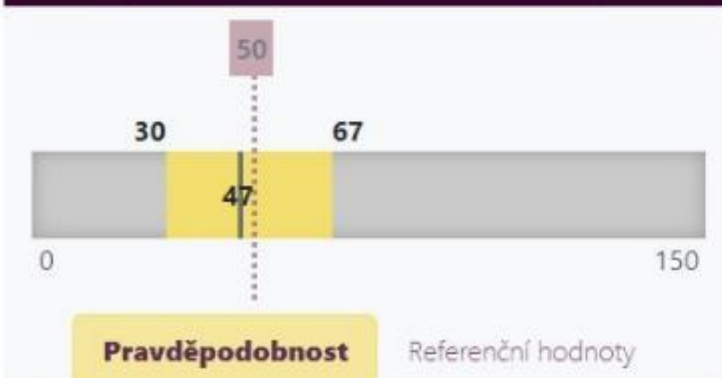
S pravděpodobností 76,08 % je překročen preventivní limit. více o preventivních limitech na webu MZe ČR

Odhad koncentrace v půdě

Nikl

47 mg/kg

S pravděpodobností 76,08 % je překročen preventivní limit. více o preventivních limitech na webu MZe ČR



Pravděpodobnost Referenční hodnoty

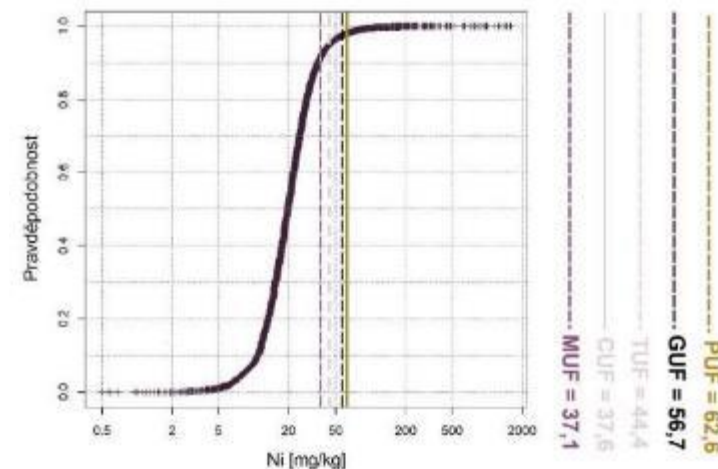
Spodní hranice predikčního intervalu je **30 mg/kg**. Střední hodnota predikce, která udává nejpravděpodobnější modelový odhad koncentrace je **47 mg/kg**. Horní hranice predikčního intervalu je **67 mg/kg**.

Legislativní limit v místě je **50 mg/kg**.

Predikční interval překrývá legislativní limit.
[Více zde.](#)

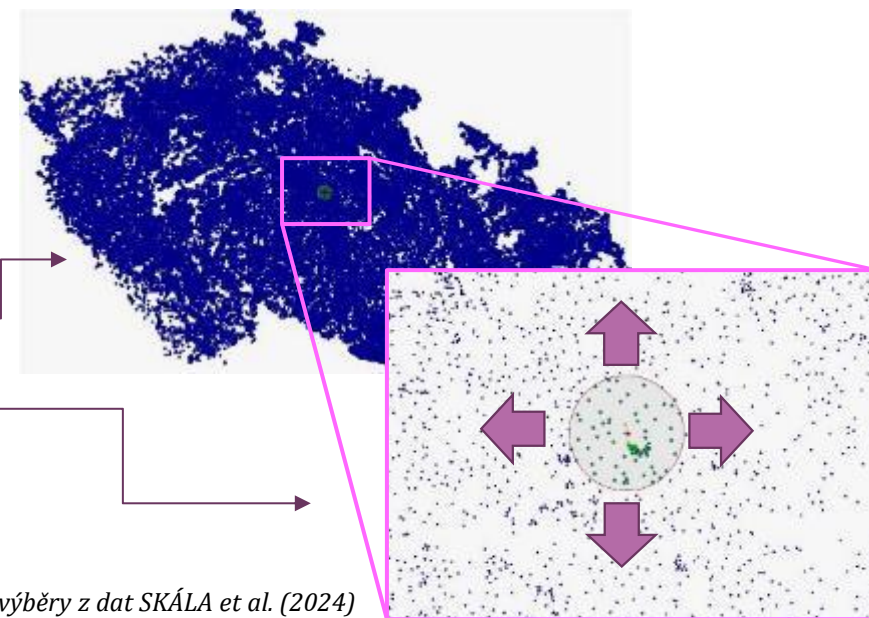
Preventivní limit pro lehké půdy	45 mg/kg
Preventivní limit pro běžné půdy	50 mg/kg
Svrchní mez statistické variability koncentrací v ČR	38 mg/kg
Lokální svrchní mez variability koncentrací	84 mg/kg
Pravděpodobnost příslušnosti do třídy lehkých půd	32 %

Národní výpočet – všechna data dostupná v ČR



$$CUF = Q_{50} + 1.5 (Q_{75} - Q_{25})$$

GRYGAR et al. (2023)



Lokální výpočet – geograficky vážené výběry z dat SKÁLA et al. (2024)



Moving from geochemical to contamination maps using incomplete chemical information from long-term high-density monitoring of Czech agricultural soils

Tomáš Matys Grygar^{1,2} · Jitka Elznicová² · Stěpánka Tůmová^{1,2} · Tomáš Kylich² · Jan Skála³ · Karel Hron⁴ · Miguel Ángel Álvarez-Vázquez⁵

Received: 6 May 2022 / Accepted: 5 December 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

The aim of this work was to show how to construct maps of anthropogenic contamination of agricultural soils in Czech Republic by risk elements. The geochemical datasets for this work originated from state monitoring utilising acid extractions according to past and present national legislative requirements. The goal was to distinguish contamination from natural variability respecting knowledge in environmental geochemistry and available information sources. Conventional approaches to geochemical maps, such as plotting sampling points where element concentrations exceed Tukey or Carling upper fences (boxplot approach), can be used to visualise only extreme contamination, such as historical ore mining and processing. The challenge starts when weak and/or diffuse contamination is of interest and should be distinguished from natural variability. Geogenic anomalies in Czech Republic are represented by mafic volcanic rocks (Cd, Cu, Zn), metamorphic rocks (As, Zn), felsic intrusive volcanic rocks (Pb, Zn), and variegated rocks showing volcanic components (Cd, Pb, Zn). Lithological anomalies are typical for floodplain sediments of lowland rivers. Each cumulation of concentrations above the Tukey or Carling upper fences within the whole-Czech dataset, i.e. potential contamination hotspot, should be examined in detail to judge possible natural controls. Pleistocene and Holocene sediments, in particular aeolian and fluvial deposits with their specific grain size, represent an important controlling factor in such detailed maps. Element concentration ratios in rational subcompositions, e.g. including Co, Cu, Pb, and Zn, were found useful to separate geogenic and lithogenic anomalies. In this subcomposition, Co is promising reference element for datasets obtained by conventional acid extractions as a surrogate for missing analyses of lithogenic elements. There is no automated way of distinguishing anthropogenic contamination from natural variability for weak contamination, expert opinion is indispensable to distinguish natural and anthropogenic factors. The larger (more heterogeneous) the mapped areas, the more complicated interpretation of their geochemical maps and less reliable identification of anthropogenic contamination. Zooming in and examination of empirical cumulative distribution of element concentrations for the zoomed areas is the most powerful tool in converting geochemical to contamination maps, assuming the zoomed areas are covered by relatively homogeneous soils, with small number of soil-forming bedrock and not much geomorphic heterogeneity.

Keywords Heavy metals · Agricultural soils · Contamination · Pollution · Data exploration

✉ Tomáš Matys Grygar
grygar@iic.cas.cz

¹ Institute of Inorganic Chemistry ASCR, 250 68 Řež, Czech Republic

² Faculty of Environment, J. E. Purkyně University in Ústí nad Labem, Ústí nad Labem, Czech Republic

³ Research Institute for Soil and Water Conservation, Závovřeska 250, 156 27 Prague, Czech Republic

⁴ Department of Mathematical Analysis and Applications of Mathematics, Faculty of Science, Palacký University, 17. Listopadu 12, 771 46 Olomouc, Czech Republic

⁵ GEAAT Group, Department of History, Art and Geography, Area of Physical Geography, University of Vigo, 32004 Ourense, Spain

Published online: 17 December 2022



Novel definition of local baseline values for potentially toxic elements in Czech farmland using adaptive spatial weighting

Jan Skála^{a,*}, Tomáš Matys Grygar^{b,c}, Alla Achasova^a

^a Research Institute for Soil and Water Conservation, CZ 156 00, Prague, Czech Republic

^b Institute of Inorganic Chemistry ASCR, CZ 250 68, Řež, Czech Republic

^c Faculty of Environment, J. E. Purkyně University in Ústí nad Labem, Czech Republic

ARTICLE INFO

Editorial handling by: Q Gong

Keywords

Potentially toxic elements
Soil geochemistry
Geochemical background
Geogenic anomalies
Diffuse soil contamination

ABSTRACT

Exploratory data analysis is commonly used in geoscientific research to identify various data populations within datasets. Frequently, the distribution-wise metrics for the statistical centre and spread are combined to define normal (background) ranges of topsoil contents of potentially toxic elements. When decreasing the geographical scale, the survey areas turn too heterogeneous for the statistical definition of a single background range. The traditional solution is the domain approach wherein various data populations (and their statistical parameters) can be attributed to contextual (geological, ecological) information. Nevertheless, summarising the entire data set from large areas as a single statistical entity would provide too much data reduction which would decrease sensitivity of detecting localised anthropogenic contamination and work wrong in areas of geogenic anomalies spatially larger than contamination. In this paper, we tested a novel numerical solution for deriving local distribution-wise baseline values via spatially limited sliding window combined with geographical weighting. Considering environmental variables (soil and topographical properties) at every analysed soil sample point, we extended the geographical kernel weighting approach which considers only spatial dimension (given by geographical coordinates). The advanced version combines two similarity modes to assign highest weights to the nearest points expected to share similar environmental contexts within the user-defined moving kernel. The method was implemented for data-mining in the Czech high-density monitoring data for agricultural soils which had to be firstly regressed to achieve analytical harmony between two distinct extraction methods employed in that monitoring, in particular cold diluted nitric acid and hot aqua regia. After the reliable harmonisation, local baseline values were delivered as the localised outer limits of variation using the proposed double-weighted kernel approach. We compared the estimated localised background ranges for 10 potentially toxic elements (As, Be, Cd, Co, Cr, Cu, Ni, Pb, V, and Zn) with those based on the conventional substrate-wise domain approach and single nationwide legislation thresholds. The comparison was also efficient in identifying an inappropriate aggregation of some geological units. Finally, the kernel approach delivered regional outer limits of variability sensitive to subtle regional variations of topsoil geochemistry.

1. Introduction

The chemical composition of the soil generally and topsoil especially results from a long-term interaction of natural soil formation and different types of anthropogenic influences. As a result, geochemistry launched research efforts to define typical (–baseline) or natural (–background) concentration ranges of elements in soils (Reimann and Garrett, 2005). The formulation of methodical tools to statistically distinguish various population densities is an ongoing process subjected to continuing refinement due to emerging research findings (Reimann

and de Caritat, 2017). The thresholds are used to distinguish between naturally occurring concentrations of the element from those resulting from diffuse anthropogenic sources, or even to differentiate between diffuse and point anthropogenic sources (McIlwaine et al., 2014). As geochemical research shows (Reimann et al., 2013; Grygar et al., 2024), even the robustly derived, single background range for large areas provides too much data reduction about variation over the survey area because it is usually too heterogeneous for the statistical definition of single background range. Hence, the site-specific local/regional reference values should be preferred to national/global counterparts (Boente

* Corresponding author.
E-mail address: skala.jan@vumop.cz (J. Skála).

<https://doi.org/10.1016/j.apgeochem.2024.106082>

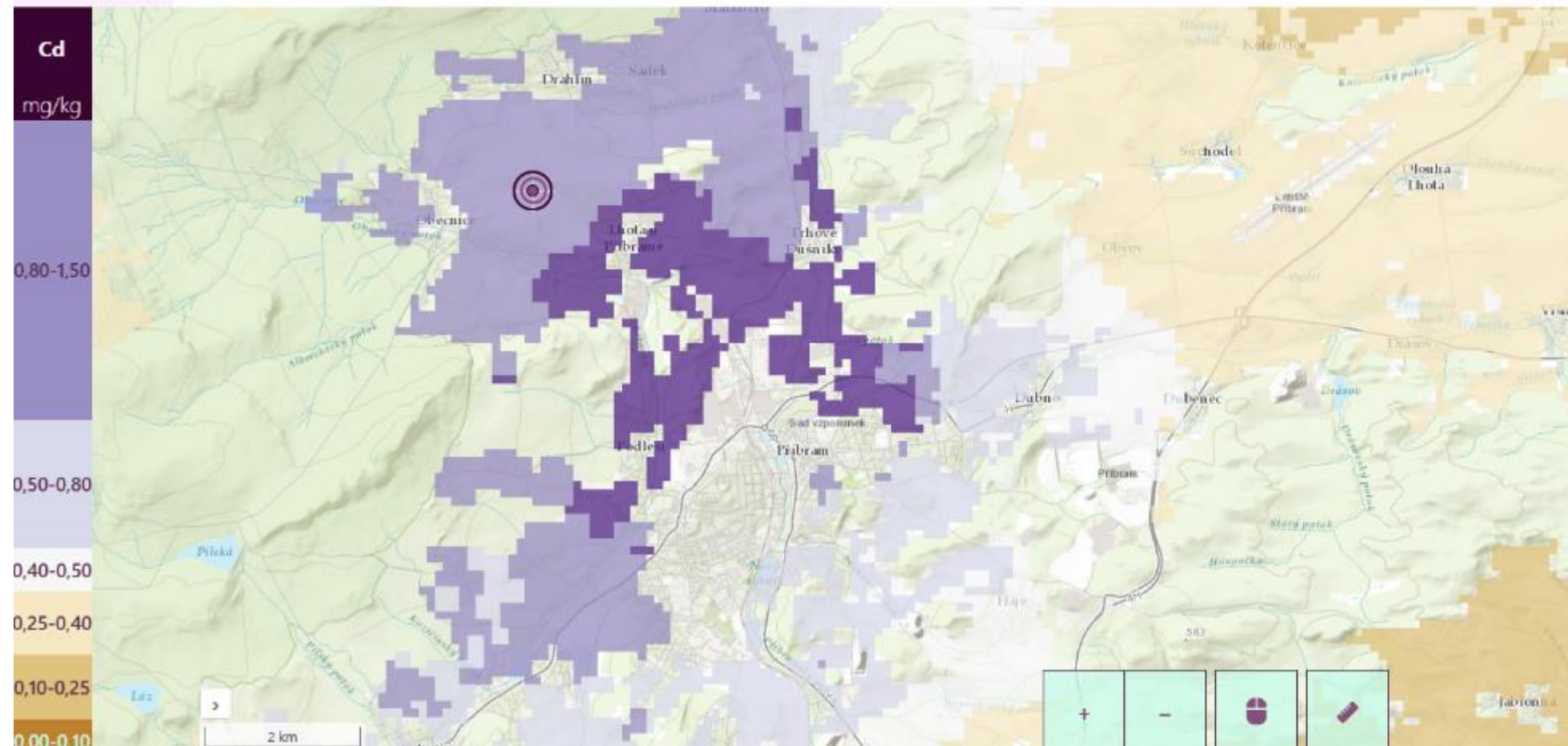
Received 29 February 2024; Received in revised form 14 June 2024; Accepted 19 June 2024

Available online 21 June 2024

0883-2927/© 2024 Published by Elsevier Ltd.



- An. prvky arsen (As) beryllium (Be) **KADMIUM (Cd)** kobalt (Co) chrom (Cr) měď (Cu) rtuť (Hg) nikl (Ni) olovo (Pb) vanad (V) zinek (Zn)
- Org. látky dichlordifenyl trichlorethan (DDT) hexachlor cyklohexan (HCH) polycyklické aromatické uhlovodíky (PAU) polychlorované bifenyls (PCB) polychlorované dibenzodioxiny (PCDD/F) hexachlorbenzen (HCB)



Pravděpodobnost Referenční hodnoty

Spodní hranice predikčního intervalu je **0,70 mg/kg**. Střední hodnota predikce, která udává nejpravděpodobnější modelový odhad koncentrace je **1,19 mg/kg**. Horní hranice predikčního intervalu je **1,83 mg/kg**.

Legislativní limit v místě je **0,50 mg/kg**.

Predikční interval překračuje legislativní limit.

[Více v popisu predikčního intervalu](#)
[Metainformační list pro: cd](#)

Horní mez modelového predikčního intervalu **1,83 mg/kg** je vyšší než 1,50 mg/kg a nachází se mimo rozsah osy.

Cd
kadmium

48

Metainformační list prediktivní geochemická mapa Kadmia

Výstup projektu TAČR č. SS03010364

System na podporu rozhodování při hodnocení kvality
půdy z hlediska obsahu rizikových látek v
zemědělských půdách České republiky



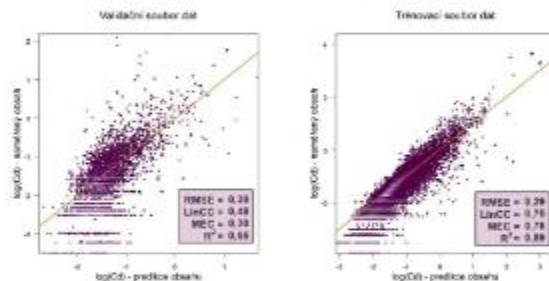
Výzkumný ústav meliorací a ochrany půdy, v.v.i.

T A
Č R



skutečné měřených, tj. z primárních dat, které byly měřeny v legislativně závazné extrakci pomocí lučavky královské. Graf modelem predikovaných a skutečně naměřených hodnot v logaritmické škále je zobrazen na Obr. 6. Globální míry přesnosti validace ukazují, že modely a zvolené environmentální faktory nejsou schopny popsat celou míru variability geochemických podmínek obsahu kadmia na území celé ČR, která se vyznačuje velkou různorodostí přírodních podmínek. Koeficienty spolehlivosti modelu vyjádřené jako korelace mezi naměřenými a modelovanými daty $R^2 = 0,55$ a jako MEC = 0,3 ukazují, že model je ve validaci po data splitu slabším modelem. Nicméně pokud se podíváme na výsledky modelu v trénovacím souboru, tak křížová validace ukazuje mnohem příznivější výsledky (MEC = 0,78).

Obr. 6 Graf modelem predikovaných a skutečně naměřených hodnot v logaritmické škále pro koncentrace kadmia predikovaných finálním modelem uQRF po rozdělení dat – validací soubor dat (70 % ze souboru přímých analytických koncentrací extrakcí lučavkou královskou), trénovací soubor dat (instanční data bez validací souboru)



To značí, že větší chyba predikce se obvykle vyskytuje v místech, která nejsou dobře pokryta trénovacími daty („přetřénování modelu na vstupní data“) nebo se nacházejí v podmínkách, které nejsou vhodně zastoupeny v trénovacím souboru. Dalším zdrojem chyby je využití různých dat – z různých období, s mírnými diferenciemi v hloubkách odběrů a rozdíly v tzv. „support size“ – tj. rozsah plochy, které vzorky odebrané jako směs, reprezentují. Podíl na nižší věrohodnosti modelu může mít také využití regresně harmonizovaných dat, nicméně ani test modelu pouze s měřenými daty nevedl k příznivějším validačním metrikám. Naopak příznivější metricky byly dosaženy pro indikátorové pravděpodobnostní modely. Pravděpodobnosti překročení limitních hodnot jsou tedy zatíženy menší nejistota než predikce vlastních koncentrací. Proto v případě výstupů modelování kadmia doporučujeme sledovat i výstupy pravděpodobnostních modelů překročení limitních hodnot. Celkově doporučujeme sledovat jak globální metricky obou povrchů (koncentrační a pravděpodobnostní mapy), tak zejména horní a dolní interval predikce (Q5 a Q95), které vymezují šířku predikčního intervalu, která

23

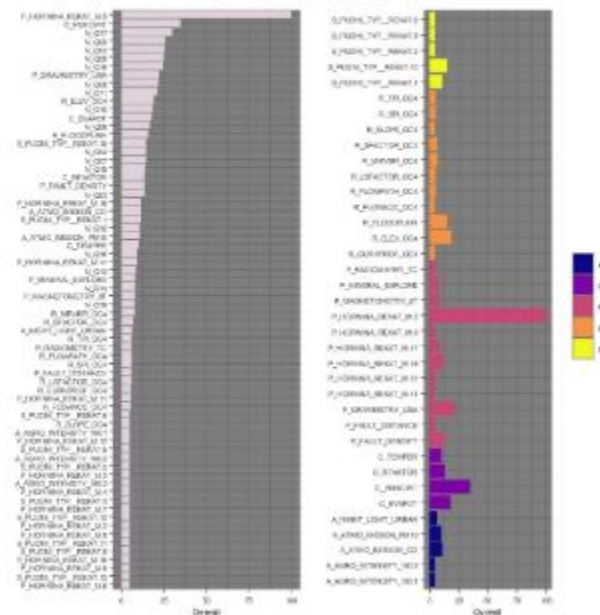


T A
Č R



ukazuje mimo jiné lokální kvalitu predikce, neboť širší predikční interval znamená, že jednotlivé modely v rámci QRF algoritmu měly větší rozptyl výsledných predikcí, a tedy i větší pravděpodobnost možné chyby predikce.

Obr. 7. Pořadí prediktorů mezi 60 % nejnižšími environmentálními prediktory úje Vápu (Drevál) včetně prostanové autokorelační struktury a výkonu jednotlivých environmentálních prediktorů po vyřazení prediktorů prostanové autokorelační struktury z tzv. „buffer distance“ – řazené dle výkonu prediktorů jako v Tabulce 3 - 5 – pádky parametry společné prořezávací, R – odvozené parametry digitálního modelu terénu, P – prediktory geologické/národnostní variace, C – faktory klimatické, A – antropogenní činnosti, N – faktory geografické pozice



24



T A
Č R

Děkuji za pozornost



Jan Skála



+420 257 027 241



skala.jan @vumop.cz



www.vumop.cz



Literatura

- *Bartoňová A. 2004. Aktuální otázky znečištění ovzduší. Praha: Univerzita Karlova, 216 s.*
- *Brunsdon Ch, Fotheringham S, Charlton M. 1998. Geographically Weighted Regression. Journal of the Royal Statistical Society: Series D (The Statistician), 47 (3).*
- *Elvidge CD, Zhizhin M, Ghosh T, Hsu FC, Taneja J. 2021. Annual time series of global VIIRS nighttime lights derived from monthly averages: 2012 to 2019. Remote Sensing 13(5).*
- *Grygar MT, Elznicová J, Tůmová Š, Kylich T, Skála J, Hron K, Álvarez-Vázquez MA. 2023. Moving from geochemical to contamination maps using incomplete chemical information from long-term high-density monitoring of Czech agricultural soils. Environ Earth Sci 82.*
- *Meinshausen N. 2006. Quantile regression forests. J. Mach. Learn. Res. 7.*
- *Reimann C, Birke M, Demetriades A, Filzmoser P, O'Connor P. 2014. Chemistry of Europe's agricultural soils – Part B: General background information and further analysis of the GEMAS data set. Geologisches Jahrbuch (Reihe B 103), Schweizerbarth, Hannover, 352 pp.*
- *Salminen R et al. 2005. FOREGS Geochemical Atlas of Europe, Part 1: Background Information, Methodology and Maps. Geological Survey of Finland, 526 p.*
- *Skála J, Grygar MT, Achasova A. 2024. Novel definition of local baseline values for potentially toxic elements in Czech farmland using adaptive spatial weighting. Applied Geochemistry 170.*
- *Zbírál J, Honsa I, Malý S, Čížmár D. 2004. Analýza půd III. Ústřední kontrolní a zkušební ústav zemědělský, Brno.*