Taylor & Francis
Taylor & Francis Group

Human and Ecological Risk Assessment: An International Journal

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/bher20

Modeling metal uptake by selected vegetables from urban soils in Europe: uncovering key soil factors using partial least squares regression (PLS-R)

Maria Lundgren, Rupert Lloyd Hough, Luke Beesley, Mads Troldborg, Lukáš Trakal, Eduardo Moreno-Jiménez, Deliang Dai & Anna Augustsson

To cite this article: Maria Lundgren, Rupert Lloyd Hough, Luke Beesley, Mads Troldborg, Lukáš Trakal, Eduardo Moreno-Jiménez, Deliang Dai & Anna Augustsson (17 Feb 2025): Modeling metal uptake by selected vegetables from urban soils in Europe: uncovering key soil factors using partial least squares regression (PLS-R), Human and Ecological Risk Assessment: An International Journal, DOI: 10.1080/10807039.2025.2464109

To link to this article: https://doi.org/10.1080/10807039.2025.2464109

HUMAN AND ECOLOGICAL RISK ASSESSMENT: AN INTERNATIONAL JOURNAL
https://doi.org/10.1080/10807039.2025.2464109

Taylor & Francis
Taylor & Francis Group

Modeling metal uptake by selected vegetables from urban soils in Europe: uncovering key soil factors using partial least squares regression (PLS-R)

Maria Lundgren[a], Rupert Lloyd Hough[a,b], Luke Beesley[c,d], Mads Troldborg[e], Lukáš Trakal[c], Eduardo Moreno-Jiménez[f,g], Deliang Dai[h], and Anna Augustsson[a]

[a]Department of Biology and Environmental Science, Linnaeus University, Kalmar, Sweden; [b]The James Hutton Institute, Aberdeen, Scotland, UK; [c]Department of Environmental Geosciences, Czech University of Life Sciences Prague, Suchdol, Prague 6, Czech Republic; [d]School of Science, Engineering & Environment, University of Salford, Salford, UK; [e]The James Hutton Institute, Dundee, Scotland, UK; [f]Department of Agricultural and Food Chemistry, Universidad Autónoma de Madrid, Madrid, Spain; [g]Institute for Advanced Research in Chemical Sciences, Universidad Autónoma de Madrid, Madrid, Spain; [h]Department of Economics and Statistics, School of Business and Economics, Linnaeus University, Växjö, Sweden

ABSTRACT

Partial Least Squares Regression (PLS-R) was introduced as a method for modeling the uptake of six potentially toxic elements (PTEs)- Ba, Cd, Cu, Ni, Pb, and Zn- by lettuce, chard, and carrot. Data were obtained from a pot experiment where these crops were cultivated in urban soils of various characteristics. The models consider soil concentrations of PTE, Al, Ca, Fe, K, Mg, Mn, Na, P, S and pH, SOM, CEC, and soil texture as predictors. Initially, eighteen metal- and crop-specific models with all predictors were developed, using selectivity ratios (SRi) to identify influential variables for predicting PTE soil-to-crop transfer. Reduced models were then created using only predictors with high SRi. Key variables for predicting PTE soil-to-crop transfer included soil PTE concentration, pH, Fe and Mn soil concentrations, and soil texture. Out of eighteen models, sixteen were suitable for predicting correlations and assessing PTE accumulation in crops, while eight were accurate for quantitative predictions. This study shows that PLS-R is a robust method for modeling soil-to-crop transfer of metal contaminants, even with multicollinear predictors. PLS-R also helps identify key variables, providing insights into the mechanisms of PTE accumulation in crops, which is crucial for effective risk assessments.

ARTICLE HISTORY
Received 19 March 2024
Revised manuscript
Accepted 26 January 2025

KEYWORDS
Urban gardening; metal contaminant soil-to-crop transfer; risk assessments; partial least squares regression (PLS-R); SDG 11: sustainable cities and communities

Introduction

Urban soils are shaped by a combination of natural biogeochemical processes and human-driven impacts, including intensive land modifications, emissions from traffic

CONTACT Maria Lundgren maria.lundgren@lnu.se Department of Biology and Environmental Science, Linnaeus University, Universitetsplatsen 1, 391 82 Kalmar, Sweden.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/10807039.2025.2464109.

and industries, and waste disposal, making the urban soils fundamentally different from their rural or natural counterparts (Pavao-Zuckerman 2008; Liu et al. 2014; Riddle et al. 2022; Ribeiro et al. 2023). They often consist of a mix of original soil and excavated materials from other sites, often containing waste, resulting in considerable geochemical variability over short distances (Beesley et al. 2020). Consequently, urban soils often display elevated concentrations of many metals, or potentially toxic elements (PTEs), e.g., barium (Ba), cadmium (Cd), copper (Cu), nickel (Ni), lead (Pb) and zinc (Zn) (Orsini et al. 2013; Datko-Williams et al. 2014; Mitchell et al. 2014; Antisari et al. 2015; Rouillon et al. 2017; Beesley et al. 2020).

Nonetheless, the expanse of urban areas is significant, with an increasing share of the world's growing population settling in cities. As a fundamental component in shaping tomorrow's sustainable cities, urban farming is often highlighted (e.g. Bendt et al. 2013; Langemeyer et al. 2018, 2021). Even at the citizen level, there is a noticeable rise in involvement in urban gardening, in settings such as community gardens or allotments (Bieri et al. 2024). Factors contributing to this trend include easy access to affordable fresh produce, health benefits from the moderate physical activity associated with gardening, and the promotion of sustainable community development and social cohesion (Audate et al. 2019; Stubberfield et al. 2022). However, also soils used for the cultivation of crops in urban areas typically contain elevated concentrations of PTEs (Clark et al. 2006; Szolnoki et al. 2013; Wu et al. 2015). Therefore, initiatives to expand vegetable production in these environments should consider the extent to which contaminants in the soil may be absorbed by the crops, and which factors that influence PTE uptake. This is critical because, for non-occupationally exposed individuals, the soil-crop-human pathway is the primary route of metal contaminant exposure (EFSA 2009, 2010; Tchounwou et al. 2012). To assess health risks, reliable models for the prediction of PTE uptake in edible crops are vital. However, the transfer of PTEs from soils to crops is complex and not easily modeled. It varies between different crop tissues and cultivars (Alexander et al. 2006; Dalcorso et al. 2013; Lima et al. 2015), and different metals differ in their availability for crop uptake (Ge et al. 2000, 2002; Bielińska and Mocek-Pióciniak 2010; Alloway 2014). The availability for uptake depends heavily on the metals' solubility, which is determined by complex, interdependent sorption processes controlling the balance between soluble and adsorbed fractions. Main sorbents include organic matter, clay minerals and (hydr)oxides of iron (Fe), manganese (Mn) and aluminum (Al) (Gray and Mclaren 2006; Romic and Zovko 2011; Suda and Makino 2016; Wang et al. 2019). Other factors that have been shown to affect the speciation and retention of metals in soils are the soil pH, cation exchange capacity (CEC) and redox potential (Kabata-Pendias 2004; Alloway 2014; Antoniadis et al. 2017). The abundance of other cations, mainly the major elements calcium (Ca), magnesium (Mg), potassium (K) and sodium (Na), is influential as they compete for sorption sites in the soil matrix, and consequently, they play a critical role in determining the total cation exchange capacity (CEC) of soils (Kabata-Pendias 2004; Alloway 2014). These elements are also major nutrients, along with nitrogen (N), phosphorus (P) and sulfur (S), and as such decisive for the nutritional status of the soil, which, in turn, may influence the geochemical cycling of many PTEs (Kirkby 2012; Zakari et al. 2021).

Despite these complexities, various mathematical models have been developed by authoritative bodies and researchers to predict the transfer of metals from soils to food crops. Empirical models, typically developed based on fitting some sort of least squares regression to large data sets with paired soil:crop samples and incorporating various soil parameters have been found to be a practicable approach to predicting crop uptake for several divalent metals (Hough et al. 2004; Legind and Trapp 2010; Stubberfield 2017). The majority of previous studies aimed at developing models to simulate the uptake of PTEs by various crops, applying MLR, have considered the soil's total metal content and pH as predictor variables (e.g. Efroymson et al. 2001; McBride 2002). To enhance the models, some have also added measurements of additional parameters, e.g., the soil's CEC or its content of organic matter, clays, Fe and/or Mn (hydro)oxides, or exchangeable Ca and/or Mg (e.g. Hough et al. 2004; Römkens et al. 2009; Bacigalupo and Hale 2011; Liang et al. 2013; dos Santos-Araujo et al. 2017; Stubberfield 2017; Xiao et al. 2018). Some metals, however, have consistently proven more difficult to model (e.g., Pb and Cu) than others (e.g., Cd and Zn) (Römkens et al. 2009; Antoniadis et al. 2017; Lundgren et al. 2023).

The most commonly applied approach for modeling the influence of soil geochemistry on metal uptake in crops is multiple linear regression (MLR) (McBride 2002; Hough et al. 2004; Novotná et al. 2015; dos Santos-Araujo et al. 2017; Stubberfield 2017; Cavanagh et al. 2019). However, an assumption in linear regression modeling is the independence of all predictor variables (James et al. 2013). This aspect is crucial for ensuring the model's reliability, and violations can result in biased estimates and misleading conclusions. Multicollinearity, which occurs when independent variables are correlated, essentially violates the requirements of a multiple linear regression (MLR) model, rendering it invalid under these circumstances (James et al. 2013; Kuhn and Johnson 2013). Since the geochemical sorption/desorption processes that control PTE phytoavailability involve multiple interdependent soil parameters, partial least squares regression (PLS-R) is an appropriate analytical approach to incorporate them effectively. PLS-R combines and generalizes features from principal component analysis (PCA) and MLR and is suitable for analyzing datasets with mutually correlated predictors, or where the number of predictors is close to or exceeds the number of observations (Wold et al. 2001; Carrascal et al. 2009; Hastie et al. 2009; Kuhn and Johnson 2013). It has previously been applied successfully to model chemical reactions and complex ecological data, such as the relationships between environmental gradients and species distribution, making it a philosophically suitable approach for modeling processes that influence soil-to-plant metal transfer (Wold et al. 2001; Carrascal et al. 2009) The PLS-R method can also aid in identifying predictors highly relevant to the system under investigation – specifically, those variables that are crucial for explaining variations in the modeled response variable. The variable selection process enables the development of simpler, more effective models and may enhance our understanding of complex systems (Andersen and Bro 2010; Mehmood et al. 2012; Farrés et al. 2015; Mehmood et al. 2020#C/a>).

This study introduces partial least squares regression (PLS-R) to predict the transfer of six divalent metals (Ba, Cd, Cu, Ni, Pb, Zn) from soil to three common garden crops: lettuce (*Lactuca sativa* L. var. *Crispa*), chard (*Beta vulgaris* L. gr. Leaf Beet), and carrot

(*Daucus carota* L.) – while identifying key soil variables. We used data from a greenhouse experiment where the test crops were grown in urban soils with diverse geochemical characteristics (with several multicollinear variables) and varying levels of contamination.

Materials and methods

Soil and crop data preparation

Soil sampling

Soil samples for the cultivation experiment were collected from active urban allotment plots in inner city areas of 6 European cities. The selection of sampling sites was affected by the timing of the project, which overlapped with the Covid pandemic and restricted access to many areas. Consequently, we were referred to allotment areas in the cities we aimed to include, based on where access was granted. This resulted in 22 different sites: 7 from Copenhagen in Denmark; 5 from Widnes in the UK; 4 from Malmö in Sweden; 3 from Madrid in Spain; 2 from Pribram in the Czech Republic, and 1 from Berlin in Germany. After removing litter and/or vegetation, topsoil samples were taken from the upper 20 cm using a stainless-steel hand trowel. Soil samples were taken from each allotment area in multiple replicates and then combined into a single composite sample for each of the 22 allotments. Upon arrival to the laboratory, bulk samples were air dried, sieved to $<$ 2 mm, and thoroughly mixed. Thereafter, three 1-liter pots were prepared for each soil-crop combination, and a sub-portion of each soil sample was set aside for chemical analyses.

Crop cultivation

The cultivation experiment was conducted indoors under controlled conditions, enabling more accurate regulation and investigation of soil factors and their influence. Cultivation of lettuce (*Lactuca sativa* L. var. *Crispa*), chard (*Beta vulgaris* L. gr. Leaf Beet), and carrot (*Daucus carota* L.) was carried out in a growth room at Linnaeus University in Kalmar, Sweden, from late December 2020 until early May 2021. Chard and lettuce were cultivated in all 22 soils and harvested after 8 wk, but only 21 soils produced sufficient biomass for analysis. For carrots, adequate biomass for further analysis was obtained from 17 soils. For carrots, which require more time to reach an edible size, harvesting was done after 16 wk. Details about the soil experimental set-up, soil preparation and vegetable cultivation are described by Qvarforth et al. (2022). After harvest, the crop samples were rinsed in deionized water prior to drying at 60 °C for 18 h. The dry matter content of the crop samples was determined during sample preparation.

Chemical analyses

Approximately 0.50 g of each soil and crop sample was placed in 50 mL poly-propylene tubes and sent to the commercial laboratory ALS Scandinavia AB in Luleå, Sweden, for digestion and determination of trace elements (Al, Ba, Ca, Cd, Cu, Fe, K, Mg, Mn, Na, Ni, P, Pb, S and Zn), using double-focusing sector field ICP-MS (ICP-SFMS). For the soil samples, pseudo-total concentrations were determined after Aqua Regia (AR) digestion, where the samples were mixed with 7.5 mL HCl and 2.5 mL HNO_3, both concentrated and of Suprapur grade. The samples were heated (130 °C) on a heating block for 2 h, followed by dilution to 50 mL using Milli-Q water. The crop samples were digested using 10.0 mL concentrated HNO_3, providing quantitative oxidation of the biologic material while achieving limited dissolution of resistant minerogenic fractions that may be associated with the plant materials (Rodushkin et al. 1999). The crop + acid mixtures were placed in a fume hood at 20 °C overnight, and thereafter on a heating block (120 °C) for 1 h. The extracts were diluted with Milli-Q water to 20 mL before the ICP analyses. Information on instrument settings and the procedure for Quality Assurance (QA) and Quality Control (QC) is provided by Qvarforth et al. (2022), and further details on the ICP-SFMS analysis can be found in e.g., Engström et al. (2004).

The analyses of grain size (clay and silt fractions) and effective cation exchange capacity (CEC) were also conducted at ALS Scandinavia AB, using approximately 100 g of material from each of the dried and sieved soils. The determination of effective CEC was done using a barium chloride solution according to CSN ISO 11260, while the analyses of particle size distribution (PSD) were done by combining wet sieve analysis (for fractions > 0.063 mm) and a laser diffraction method (for the finer fractions), according to method CZ_SOP_D06_07_120.

Soil pH and soil organic matter (SOM) were analyzed at Linnaeus University, Kalmar, Sweden. SOM was determined by loss on ignition (LOI) at 580 °C (Schulte and Hopkins 1996; Salehi et al. 2011), and pH in a suspension of soil and deionized water (Thomas 1996) using a portable multi probe (HACH HQ40d) designed for low ionic strength samples.

Statistical analysis

This study uses Partial Least Squares Regression (PLS-R) to develop predictive models for the soil-to-plant transfer of six potentially toxic elements commonly elevated in urban areas (Ba, Cd, Cu, Ni, Pb, and Zn), using soil concentrations of these PTEs along with Al, Ca, Fe, K, Mg, Mn, Na, P, and S, as well as pH, CEC, and clay and silt content as predictors. PLS-R addresses multicollinearity among predictors by reducing dataset dimensionality, creating a few uncorrelated latent variables (PLS components) from combinations of the correlated variables (Wold et al. 2001; Boulesteix and Strimmer 2007; Hastie et al. 2009). These components maximize the predictor variance relevant to the response, strengthening model accuracy by optimizing covariance between predictors and response(s) (Hastie et al. 2009; Abdi 2010). A short description of the algorithm behind PLS-R can be found in the supplement and more comprehensive descriptions of PLS-R in general can be found in Wold et al. (2001) and Hastie et al. (2009).

To assess the importance of each predictor in the PLS-R analyses, we used selectivity ratios (SR_i), which represents target-projected loadings within the validated PLS-R model (Rajalahti et al. 2009; Kvalheim 2010). SR_i is defined as the ratio of the explained predictive variance ($V_{expl,i}$) and the residual, or unexplained, variance ($V_{res,I}$) for each predictor i on the target projected component (Kvalheim 2010, 2020). SR_i can range from 0 to ∞, with $SR_i \approx 1$ indicating that a predictor's contribution to explaining the response variable is comparable to its residual variance. Predictors with low SR_i (close to 0) contribute little to explaining variance and may rather introduce noise into the model (Rajalahti et al. 2009). The decision on an SR_i-threshold to distinguish between informative or uninformative variables is generally data-dependent and context-specific (Kvalheim 2010). In this study, we calculated means of SR_i (Eq. (1)).

$$SR_{mean} = \frac{\sum SR_i}{n},\ i = 1, 2, 3, \ldots \tag{1}$$

The PLS-R analyses were conducted in R (version 4.1.3) using the "mdatools" package (Kucheryavskiy 2020; R Core Team 2022). An overview of the analysis workflow is shown in Figure 1. Prior to analysis, the dataset was divided into 18 sub-datasets, one for each combination of PTE metal and crop (Ba, Cd, Cu, Ni, Pb, Zn in chard, lettuce, and carrot). Univariate PLS-R was then applied to predict the concentration of each metal in the crops, considering the 15 soil predictors (concentrations of the target PTE, Al, Ca, Fe, K, Mg, Mn, Na, P, and S, as well as pH, SOM, CEC, and clay and silt content). All response variables and predictors, except soil pH and CEC, were log10-transformed to normalize distributions and ensure a linear relationship with the response (Kuhn and Johnson 2013). Predictors were also auto-scaled (mean-centered and scaled to unit variance) to provide equal weight in the analysis. Pearson correlation matrices were assessed to confirm multicollinearity among soil variables (Supplement Figure S1).

Step 1 in the PLS-R analyses

In the first step of the PLS-R analysis, all 15 predictor variables were included to create 18 full models (one for each PTE and crop combination). Cross-validation was used to select the optimal number of PLS components, and leave-one-out cross-validation (LOOCV) was applied to determine the best component count (Wakeling and Morris 1993; Westad and Marini 2015; Nengsih et al. 2019). The optimal number of PLS components in the PLS-R models was selected using the first local minimum of $RMSE_{cv}$ (Eq. (2)).

$$RMSE_{cv} = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{n}} \tag{2}$$

In Eq. (2), n represents the number of observations, y_i is the measured value of the response (target metal concentration in crop) and $\hat{y}_i$ is the model's predicted value from the cross-validation procedure. Inspection of the plots with RMSE for calibration and cross-validation results against the number of PLS components were also used to complement the automated selection described above (Kucheryavskiy 2020). In this step, SR_i for each predictor, and the SR_{mean} was calculated separately for each model (Eq. (1)).

Step 2 in the PLS-R analyses

In the second step, reduced PLS-R models were created for each PTE and crop combination, including only predictors with SR_i values above the SR_{mean} from step 1 (Mehmood et al. 2020; Westad and Marini 2022). The optimal number of PLS

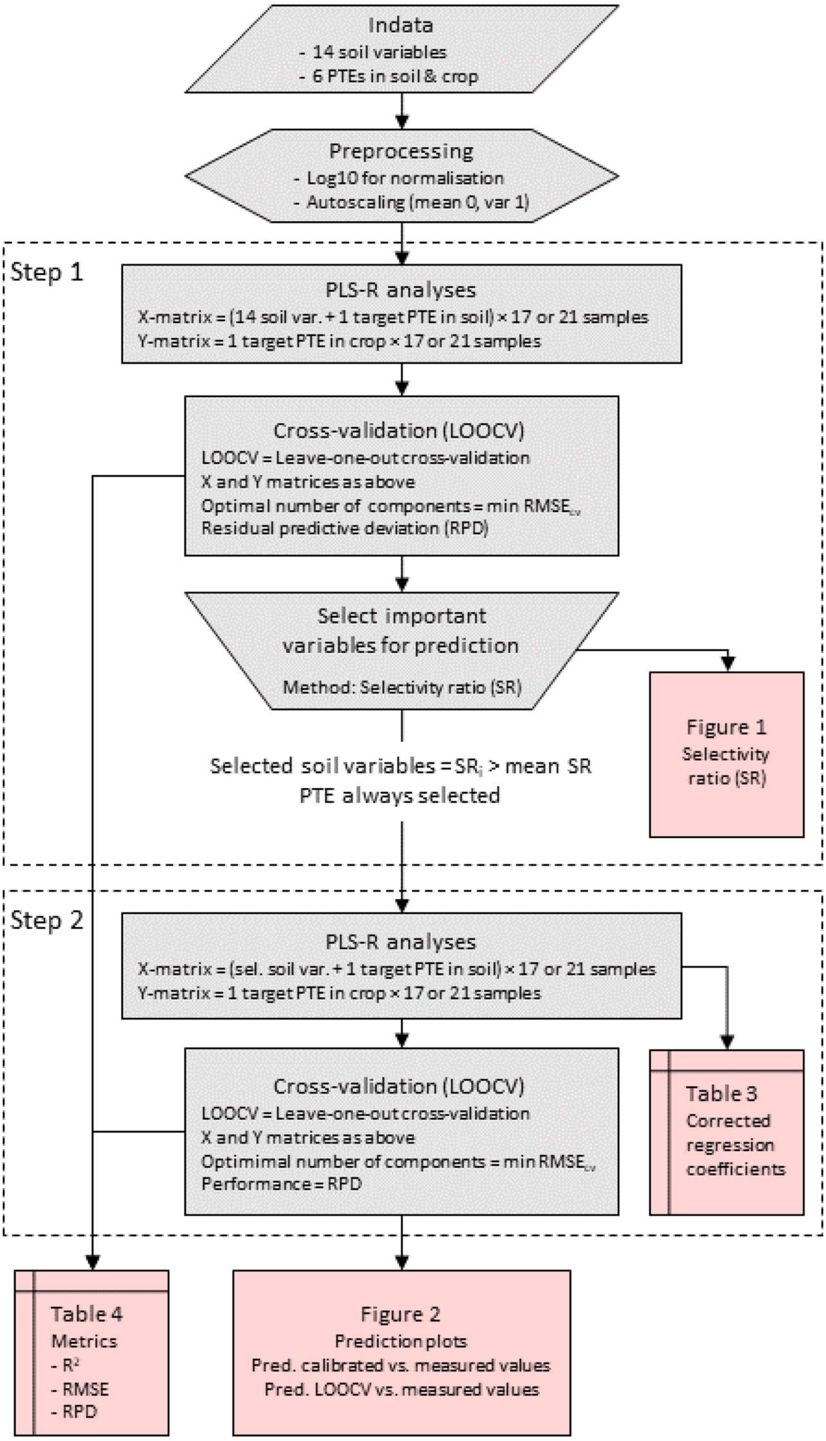

Figure 1. Flow chart of the PLS-R analysis.

components was again based on the first local RMSE minimum. For each crop-metal pair, predictive feasibility was validated using LOOCV, and predictions were plotted against measured crop metal concentrations with coefficients of determination (R^2) for calibration (R^2_{cal}) and cross-validation (R^2_{cv}). Additionally, model performance was assessed using the residual predictive deviation (RPD), a ratio of sample standard deviation to prediction RMSE, Eq. (3) (Williams and Sobering 1993; Chang et al. 2001).

$$RPD = \frac{SD_{sample}}{RMSE_{cv}} \tag{3}$$

According to Chang et al. (2001), models with RPD values <1.0 have poor predictive ability and should not be used; values between 1.0 and 1.4 indicate poor models that can potentially distinguish high from low values; RPDs between 1.4 and 2.0 suggest moderate predictive ability; and RPDs >2.0 indicate models suitable for quantitative predictions.

Finally, corrected regression coefficients were extracted for each reduced model in step 2 (Kucheryavskiy 2020).

Results

Soil and crop chemical composition

All PTEs included in this study (Ba, Cd, Cu, Ni, Pb and Zn) were present at higher concentrations in the 22 analyzed inner-city allotment soils than in typical agricultural soils (Table 1), while the major elements (Al, Ca, Fe, K, Mg, Mn, Na, P, and S) fell within representative ranges (Table 1; Szolnoki et al. 2013; Alloway 2014). Compared to the baseline values for Swedish and European agricultural soils in Table 1, also the median values of soil CEC, pH and SOM in the analyzed allotment soils were slightly elevated. It is worth reiterating that the measured metal concentrations represent total (or pseudo-total) levels.

For PTEs in the cultivated lettuce and carrot samples, mean concentrations were higher than those reported for commercial produce by Jorheim et al. (2016), except for Ni in lettuce (Table 2). No specific reference data was available for chard. For Cd and Pb, which have permissible concentrations set by Commission Regulation (EU) 2023/91, several samples also exceeded the EU Commission thresholds for leafy vegetables and carrots (root and tuber vegetables) (Table 2).

Using PLS-R to identify influential soil variables for predicting the soil-to-plant transfer of PTEs

In the full models of step 1, the optimal number of PLS components (ncomp) utilized to define the models varied from only one component (in five models), to two (in six models) and three (in five models) (Figure 2). Five components were used in two models; for Cd in chard and Ni in carrot. The SR_{mean} values, which were used to identify the most important variables for prediction of the response, varied between 0.4 and 2.0. These values were, in general, lowest for chard (0.4–0.6) and highest for carrot (0.7–2.0) (Figure 2).

Table 1. Descriptive statistics for the analyzed target PTEs and basic soil geochemistry for all 22 study soils.

Variable	Unit	Soil samples ($n = 22$)					Sweden		Europe Median
		Mean	Min	Median	(PC25–PC75)	Max	Median	(PC25–PC75)	
pH (H_20)		7.2	5.0	7.3	(7.1–7.5)	8.0	4.9	(4.6–5.4)[a]	6.3[b]
SOM	%	8.4	4.5	8.6	(5.6–11)	14	4.5	(3.4–6.0)[c]	2.5[d]
CEC	$cmol\cdot kg^{-1}$	17	4.1	17	(14–21)	27	12	(7.7–18)[e]	13[b]
Clay	%	0.60	0.01	0.50	(0.24–0.80)	1.8	n.f.		n.f.
Silt	%	26	1.9	20	(16–32)	65	n.f.		n.f.
Al	$g\cdot kg^{-1}{}_{dw}$	10	5.6	9.4	(8.0–13)	21	9.2	(6.5–13)[a]	11[a]
Ca	$g\cdot kg^{-1}{}_{dw}$	17	2.1	16	(7.1–.26)	39	2.5	(1.6–130)[a]	3.0[a]
Fe	$g\cdot kg^{-1}{}_{dw}$	19	9.2	19	(11–23)	43	16	(9.2–21)[a]	17[a]
K	$g\cdot kg^{-1}{}_{dw}$	2.4	0.39	2.2	(1.8–3.0)	5.8	790	(370–1.4)[a]	1.2[a]
Mg	$g\cdot kg^{-1}{}_{dw}$	3.1	0.61	2.5	(2.3–3.2)	7.2	2.3	(1.3–3.9)[a]	2.9[a]
Mn	$g\cdot kg^{-1}{}_{dw}$	0.38	0.094	0.34	(0.24–0.42)	1.3	0.34	(0.21–0.53)[a]	0.44[a]
Na	$g\cdot kg^{-1}{}_{dw}$	0.37	0.22	0.31	(0.27–0.42)	0.76	0.059	(0.038–0.096)[a]	0.048[a]
P	$g\cdot kg^{-1}{}_{dw}$	1.5	0.54	1.3	(0.96–1.9)	3.3	0.76	(0.59–0.92)[a]	0.65[a]
S	$g\cdot kg^{-1}{}_{dw}$	0.60	0.24	0.57	(0.40–0.71)	1.2	0.29	(0.22–0.37)[a]	0.21[a]
Ba	$mg\cdot kg^{-1}{}_{dw}$	300	51	170	(84–370)	930	52	(36–57)[a]	62[a]
Cd	$mg\cdot kg^{-1}{}_{dw}$	1.2	0.092	0.71	(0.31–1.4)	7.4	0.15	(0.10–0.21)[a]	0.18[a]
Cu	$mg\cdot kg^{-1}{}_{dw}$	150	16	46	(24–230)	690	10	(7.0–16)[a]	15[a]
Ni	$mg\cdot kg^{-1}{}_{dw}$	23	4.9	15	(10–34)	70	9	(4–14)[a]	15[a]
Pb	$mg\cdot kg^{-1}{}_{dw}$	320	21	84	(43–310)	1 900	12	(9.0–16)[a]	16[a]
Zn	$mg\cdot kg^{-1}{}_{dw}$	440	61	190	(110–550)	2 000	47	(30–64)[a]	45[a]

European and Swedish medians and percentiles (PCs) for normal agricultural soils are provided at the right-hand end of the table for comparison. n.f. = no comparable results found.

[a] Andersson et al. (2014).

[b] Ballabio et al. (2019).

[c] Eriksson (2021).

[d] de Brogniez et al. (2015), SOM converted from organic carbon (OC) assuming $SOM = .\frac{OC}{0.58}$

[e] Eriksson et al. (2010).

Table 2. Descriptive statistics for the analyzed PTE concentrations ($mg\cdot kg^{-1}{}_{dw}$) in crops from the greenhouse experiment.

Crop	Metal	Crop concentration ($mg\cdot kg^{-1}{}_{dw}$)					Commercial produce Mean ($mg\cdot kg^{-1}{}_{dw}$)	Threshold[b] ($mg\cdot kg^{-1}{}_{dw}$)
		Mean	Min	Median	(PC25–PC75)	Max		
Chard	Ba	84	15	66	(38–120)	300		
$n = 21$	Cd	**6.1**	0.051	0.92	(0.29–1.9)	**95**		**2.5**
	Cu	18	7.0	18	(13–26)	37		
	Ni	0.59	0.067	0.32	(0.12–0.62)	3.3		
	Pb	**5.5**	0.39	0.92	(0.78–1.6)	**88**		**3.8**
	Zn	390	25	130	(59–620)	1600		
Lettuce	Ba	25	3.2	15	(6.8–24)	140	9.8[b]	
$n = 21$	Cd	2.9	0.16	0.55	(0.29–1.2)	**34**	0.16	**3.3**
	Cu	12	6.1	10	(9.2–18)	24	7. 7	
	Ni	0.49	0.092	0.35	(0.19–0.61)	1.8	0.81	
	Pb	1.7	0.29	0.64	(0.39–1.3)	**14**	0.12	**5.0**
	Zn	120	30	63	(46–190)	310	36	
Carrot	Ba	23	6.3	19	(14–31)	49	22[b]	
$n = 17$	Cd	0.27	0.016	0.15	(0.069–0.26)	**1.6**	0.19	**0.91**
	Cu	4.6	2.9	3.8	(3.2–5.5)	11	3.8	
	Ni	1.0	0.051	0.22	(0.092–0.68)	12	0.36	
	Pb	0.64	0.061	0.21	(0.11–**1.0**)	**2.8**	0.095	**0.91**
	Zn	40	12	22	(17–31)	280	15	

Bold figures indicate values exceeding the thresholds for Cd and Pb in commercial produce of commission Regulation (EU) 2023/915. Descriptive statistics for commercial produce (carrot and lettuce) were published by the Swedish food agency (Jorheim et al. 2016). PC = percentile.

[a] Means for commercial produce and the EU commission thresholds are converted from fresh weight, using the dry weight data from this study, i.e., 8% for chard, 6% for lettuce and 11% for carrot.

[b] Data derived from McBride et al. (2014) and refers to the median.

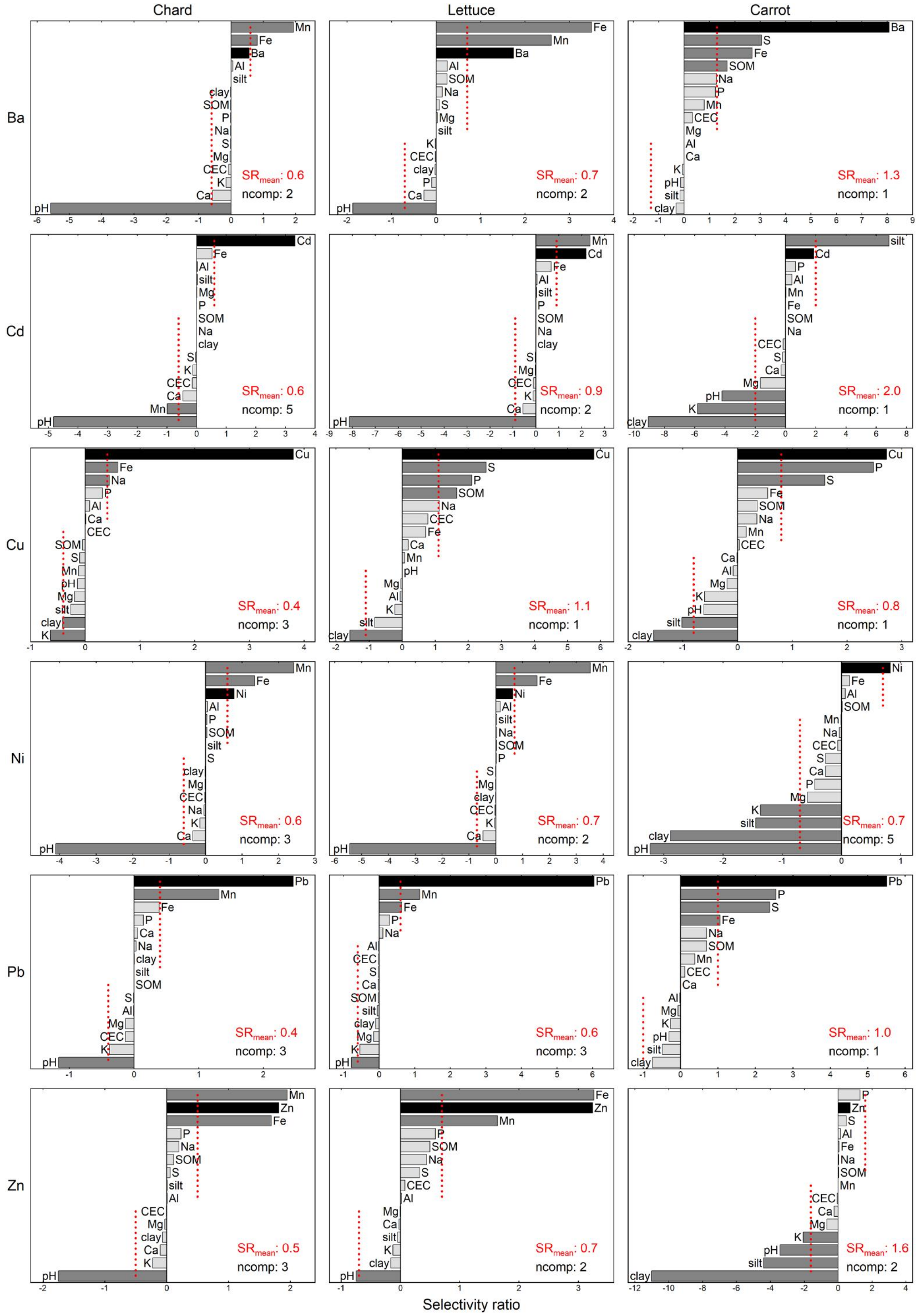

Figure 2. Selectivity ratio (SR) for soil variables in the crop and PTE specific PLS-R models developed in step 1. Black baris the target PTE in the soil (always selected to be included in step 2), dark grey bars are soil variables with $SR_i > SR_{mean}$, i.e. those selected to be included in step 2, and light grey bars are soil variables excluded in step 2. Negative bars show variables with negative signs and positive bars show variables with positive signs for the PLS-R coefficient. Ncomp refers to the number of PLS components used in the analyses.

The SR_is for the individual predictors in all models were between 0 and 11, but values above 3 were rare. When examining individual predictor's SR_is relative to the models' SR_{mean}, with the inclusion criterion for a predictor variable in step 2 being that its SR_i in step 1 exceeded the SR_{mean}, it was observed that higher SR_is were particularly associated with the soil concentration of the focal PTE, as visualized through the black bars in Figure 2. The PTE concentration showed $SR_i > SR_{mean}$ in 14 of 18 models (Figure 2). The SR_i was close to the mean values also for the last 4 models; Ba in chard, Ni in lettuce, and Cd and Zn in carrot. In addition, soil pH had SR_is above the mean in 13 of the 18 PLS-R models, and the concentrations of Fe and Mn were found significant enough for inclusion in step 2 in 10 cases (Figure 2). Four predictors had SR_is below SR_{mean} in all step 1 models, and were thus not included at all in step 2. These predictors were CEC, Mg, Al, and Ca.

Reduced models compared to full models

The reduced PLS-R models in step 2 are developed from data on the target metal in the soil together with the variables with SR_i above SR_{mean}. The optimal number of PLS components in the reduced models (ncomp in Table 3) was between 1 and 3 for all models except that for Ba in carrot, which was based on 4 optimal components. Table 3 also shows the PLS-R corrected regression coefficients from the second step, which can be applied directly to raw data (without auto-scaling).

A comparison between the predictive performance of the full and reduced models, based on R^2_{cv}, $RMSE_{cv}$ and RPD (Table 4), indicates that the reduced models, in general, have a better potential for predictive purposes. This is implied by higher R^2_{cv} values and higher RPDs in 10 out of 18 of the reduced models and lower $RMSE_{cv}$ values in 11 out of 18. Improved predictive performances was clearly shown in all Pb and Zn models.

Performance of reduced models

In general, our models' R^2_{cal} values ranged from 0.64 to 0.94 (Figure 3), indicating a high model fit. Our PLS-R models for predicting Cd uptake in lettuce and chard showed especially high R^2_{cal} values, 0.81 and 0.94, respectively. The predicted concentration of Cd, Ni and Zn in crops generally align more closely with the 1:1 reference line and have higher R^2 values (both for calibration and cross-validation) compared to the models for Cu, Pb and Ba (Figure 3). In addition, the models for Pb and Ba showed the highest variation in performance depending on crop type. The models developed for Cd in chard, Zn in chard and Ni in carrot were the most accurate, with predictions from the cross-validation (LOOCV) closely aligned to the 1:1 reference line. Accordingly, these models also returned a high R^2_{cv} and high scores for RPD (Table 4). The opposite was shown for all three models predicting Cu uptake, as well as for the models predicting Pb and Ba in chard. The R^2 for the cross-validation (R^2_{cv}) for the PLS-R models in this study, which, together with RPD (Table 4), was used as a validation of the models' potential predictive performances, averaged 0.68 (range 0.36–0.90).

Table 3. Corrected regression coefficients from the PLS-R analyses in step 2.

Target metal (T_{met}) in crop	Crop type	ncomp	Intercept	Regression coefficients for selected soil variables (X)														
				T_{met} [a]	pH	Fe	Mn	Clay	K	Silt	S	P	SOM	Na	Mg	Al	Ca	CEC
Ba	chard	1	−0.034	0.192	−0.167	0.420	0.327											
	lettuce	1	−2.23	0.316	−0.184	0.666	0.456											
	carrot	4	4.34	0.706		−0.087					−1.30		0.637					
Cd	chard	3	8.52	0.857	−0.967		−0.575											
	lettuce	1	0.901	0.423	−0.414		0.807											
	carrot	3	2.86	0.685	−0.431			−0.03	−0.078	0.414								
Cu	chard	1	−0.145	0.099		0.236		−0.092	−0.183					0.236				
	lettuce	1	−0.324	0.069				−0.073			0.128	0.137	0.141	0.186				
	carrot	3	0.294	0.202				−0.080		−0.088	−0.016	−0.080						
Ni	chard	3	2.650	0.450	−0.659	0.34	−0.156											
	lettuce	2	−1.20	0.152	−0.294	0.412	0.371											
	carrot	2	0.596	0.841	−0.287			−0.283	−0.297	0.189								
Pb	chard	1	−0.284	0.284	−0.271		0.686											
	lettuce	3	0.791	0.721	0.004	−0.883	0.517											
	carrot	3	−3.62	0.703		0.74					−1.00	0.412						
Zn	chard	3	6.22	0.770	−0.777	0.286	−0.553											
	lettuce	3	2.06	0.531	−0.236	0.089	−0.024											
	carrot	1	2.31	0.091	−0.114			−0.156	−0.251	−0.221								
№ of models the variable was included in				18	13	10	10	6	4	4	3	3	2	2	0	0	0	0
Percentage				100%	72%	56%	56%	33%	22%	22%	17%	17%	11%	11%	0%	0%	0%	0%

Ncomp is the optimal number of components selected to be used in the PLS-R models. The two bottom rows summarize the number and percentage of models each variable was included in. Before the PLS-R analyses, all predictors, except pH and CEC, and the responses were $\log_{10}$ transformed.

[a] T_{met} was always included in step 2.

Table 4. Comparative model metrics for the PLS-R models from step 1 and 2.

Model		Step 1 models all X-variables			Step 2 models SR-selected X-variables			R^2_{cv} Step 2 ≥ R^2_{cv} Step 1	$RMSE_{cv}$ Step 2 ≤ $RMSE_{cv}$ Step1	RPD Step 2 ≥ RPD Step 1
Metal	Crop	R^2_{cv}	$RMSE_{cv}$	RPD	R^2_{cv}	$RMSE_{cv}$	RPD			
Ba	Chard	0.65	0.180	1.73	0.56	0.213	1.55	FALSE	FALSE	FALSE
	Lettuce	0.50	0.321	1.45	0.62	0.273	1.66	TRUE	TRUE	TRUE
	Carrot	−0.01	0.226	1.03	0.76	0.112	2.10	TRUE	TRUE	TRUE
Cd	Chard	0.92	0.201	3.57	0.90	0.232	3.28	FALSE	FALSE	FALSE
	Lettuce	0.77	0.280	2.14	0.75	0.290	2.07	FALSE	FALSE	FALSE
	Carrot	0.46	0.395	1.40	0.63	0.358	1.71	TRUE	TRUE	TRUE
Cu	Chard	0.65	0.125	1.73	0.52	0.124	1.48	FALSE	TRUE	FALSE
	Lettuce	0.46	0.132	1.40	0.43	0.135	1.36	FALSE	FALSE	FALSE
	Carrot	0.47	0.106	1.42	0.61	0.100	1.66	TRUE	TRUE	TRUE
Ni	Chard	0.79	0.209	2.27	0.75	0.228	2.06	FALSE	FALSE	FALSE
	Lettuce	0.76	0.161	2.1	0.74	0.169	2.01	FALSE	FALSE	FALSE
	Carrot	0.87	0.280	2.96	0.87	0.244	2.84	FALSE	TRUE	FALSE
Pb	Chard	0.19	0.449	1.14	0.35	0.430	1.27	TRUE	TRUE	TRUE
	Lettuce	0.34	0.346	1.26	0.64	0.259	1.71	TRUE	TRUE	TRUE
	Carrot	0.52	0.325	1.49	0.69	0.211	1.86	TRUE	TRUE	TRUE
Zn	Chard	0.80	0.253	2.28	0.86	0.209	2.77	TRUE	TRUE	TRUE
	Lettuce	0.66	0.199	1.75	0.68	0.195	1.81	TRUE	TRUE	TRUE
	Carrot	0.59	0.232	1.63	0.74	0.214	2.02	TRUE	TRUE	TRUE

Discussion

Soil and crop chemical composition

Although the selection of the 22 test soils in this project was largely influenced by accessibility, all are from inner-city allotments within major metropolitan regions. Consequently, they are all impacted to varying degrees by a mix of traffic emissions and historical industrial activities. These study soils may not represent the full range and diversity of urban soils used for crop cultivation across Europe; however, the results reveal a wide variety of soil characteristics and elevated concentrations of potentially toxic elements (PTEs). This variability in soil composition provides a solid foundation for studying PTE uptake in crops specifically grown in urban soils designated for cultivation.

Even though total concentrations exceed those of the potentially phytoavailable fraction – particularly for metals with low solubility (such as Pb or Cu) compared to those with higher solubility (like Zn or Cd) – they offer the most reliable measure of the soil's "metal content" for subsequent modeling. Consequently, many soil-to-plant transfer models define the soil metal content based on concentrations after strong extractions, utilizing e.g. Aqua Regia or concentrated $HNO_3 + H_2O_2$, and then use geochemical information to estimate the plant available fraction from the near total measure (McBride 2002; Hough et al. 2004; Novotná et al. 2015; dos Santos-Araujo et al. 2017; Stubberfield 2017). This has been shown to be a more reliable approach to modeling uptake than using more direct measures of bioavailability (Hough et al. 2005). Weaker extractions, which aim to target phytoavailable concentrations, often yield more variable results that are strongly influenced by the composition of the solid material being tested, thereby reducing their reliability.

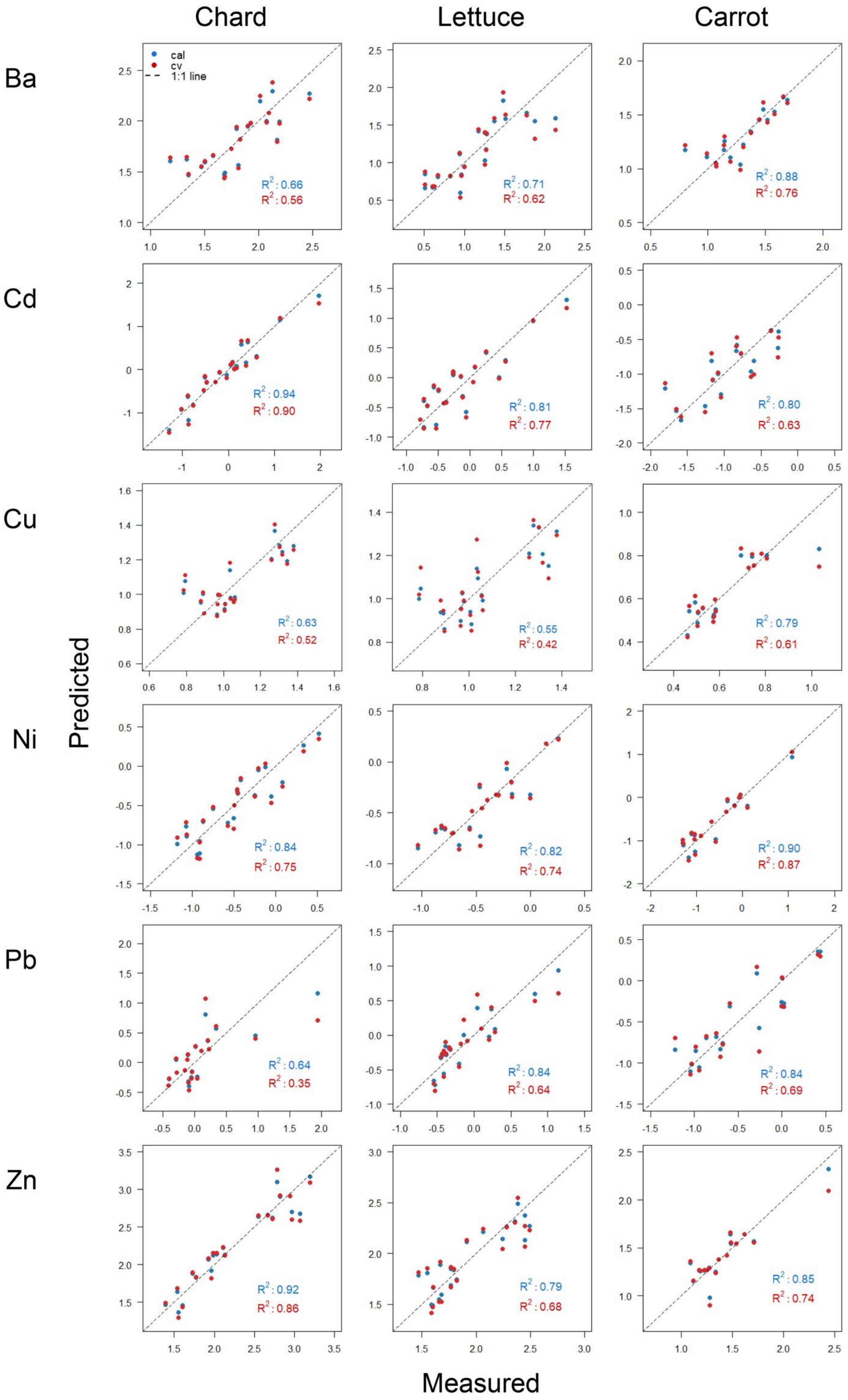

Figure 3. Predicted vs. measured $\log_{10}$ metal concentration ($mg{\cdot}kg^{-1}{}_{dw}$) in crop for the PLS-R models developed on selected soil variables in step 2. Model predictions and R^2 for calibration (cal) are shown in blue and for cross-validation (cv) in red. The 1:1 line (black dotted line) indicates total agreement between predicted and measured metal concentration in crop.

Using PLS-R to identify influential soil variables for predicting the soil-to-plant transfer of PTEs

Predictors that were defined as influential in the Step 1 PLS-R components for different PTE/crop combinations, and subsequently included in Step 2, are indicated in the matrix of Table 3. Since the concentration of the studied PTE in the soil was an important predictor in the majority of the models, it implies that the use of total (or pseudo-total) metal concentrations to define the soils' "metal content" is applicable.

The significance of soil pH

The soil pH had a clear effect on the uptake of all metals except Cu (Table 3), despite a relatively narrow pH span of the study soils. Additionally, the pH effect was greater for the uptake in lettuce and chard than in carrots. The pH effect on metal solubility in soils and – consequently – their subsequent absorption by crops is well known (McBride 2002; McBride et al. 2014; Adamczyk-Szabela et al. 2015) and is explained by the strong influence of pH on the availability of negative surface charges in the soil matrix, where the high concentration of protons at low pH values effectively prevents electrostatic sorption of other cations (McBride 1989). Consequently, the solubility and availability for crop uptake of cations increases as pH decreases.

The significance of Fe and Mn (hydr)oxides

In this study, we use pseudo-total Fe and Mn concentrations as indicators for the abundance of their (hydr)oxide forms. Since the Aqua Regia digestion used before ICP analysis also dissolves other fractions, such as carbonates and sulfides, it's important to note that the (hydr)oxide abundance may be overestimated. Nevertheless, this approximation has been employed by others before; for instance, both Bešter et al. (2013) and Yi et al. (2020) successfully used pseudo-total Mn and Fe concentrations in soil to model Cd uptake into vegetables.

With a pattern resembling that seen for pH, the concentrations of Fe and Mn in our studied allotment soils also had a significant impact on the uptake of all PTEs except Cu, and a lesser effect on uptake in carrots than in leafy vegetables (Table 3). The generally high relevance of Fe and Mn in our study aligns with previous findings, which show that, in the presence of Mn/Fe (hydr)oxides, a significant portion of adsorbed trace metals typically binds to these oxides, resulting in reduced solubility (Gray and Mclaren 2006; Suda and Makino 2016; Antoniadis et al. 2017). However, different trace elements bind to the surfaces of Mn/Fe (hydr)oxides with different strengths; for example, Pb ions are in most cases more strongly bound than e.g. Cu, Ni and Cd (Kabata-Pendias 2010). The binding strength of contaminant metals to Mn/Fe usually also depends on the pH, and typically decreases with pH (Suda and Makino 2016). This may explain the similar role seen for pH and the concentrations of Mn/Fe in our models, which can be described as part of the underlying processes controlling the phytoavailability of PTE, which PLS-R can capture in the analyses (Wold et al. 2001).

The significance of clay and silt

In our PLS-R models, clay content, and to some extend silt, was particularly important for predicting Cu uptake, and in several of the carrot models (Table 3). Our models of PTE uptake in carrot indicate that clay and silt, often together with pH were key parameters influencing the processes controlling phytoavailability of the target PTE (Figure 2). However, overall, clay content has relatively low importance for predicting metal uptake, which contrasts with previous studies that identify clay as a key factor in metal adsorption in soils and uptake by crops (Römkens et al. 2004; dos Santos-Araujo et al. 2017). The high surface area and abundance of negatively charged sites on clay minerals are crucial for the soil's ability to buffer and retain metal cations (McBride 1989; Kabata-Pendias 2004; Alloway 2014). For this reason, clay content has been used as a predictor in models of metal transfer from soil to crops in several studies (Lijzen et al. 2001; Otte et al. 2001; dos Santos-Araujo et al. 2017).

Other variables

In general, SOM, together with pH, is the most widely used variable in regression models for prediction of the transfer of PTEs from soil to different crops (Hough et al. 2004; Novotná et al. 2015; dos Santos-Araujo et al. 2017; Lundgren et al. 2023). However, in contradiction to the well-documented association between soil organic matter (SOM) and the phytoavailability of trace elements in soils (Sauvé et al. 2000; Römkens et al. 2004; Violante et al. 2010; Antoniadis et al. 2017; dos Santos-Araujo et al. 2017), SOM was of low importance for predicting the uptake of PTEs in our PLS-R analyses (Figure 2). One possible explanation for why SOM did not emerge as one of the most influential parameters for predicting PTE uptake in this study could be its multicollinearity with S and P. The strong correlation between S, P, and SOM is evident in the correlation matrices provided in the Supplement Figure S1. Similarly, in the loading plots, S and P often appear clustered together with SOM (Supplement Figure S2), indicating that they exhibit similar relationships with PTE concentrations in crops. It is also possible that the determination of SOM (%LOI) was less accurate than the measurements of S and P concentrations, resulting in a clearer association between the latter two and the response. Salehi et al. (2011) noted that %LOI may be more suitable for rough estimations of organic matter than for measurements requiring high accuracy.

Among the predictors that were not included in any of the reduced models (Table 3), CEC is occasionally found to be useful for predicting the uptake of various metals by plants (Xiao et al. 2018). The CEC should, however, be viewed as a somewhat imprecise measure as it reflects the overall abundance of negatively charged surfaces within the soil matrix, and as such, it is influenced by a multitude of other soil factors and processes that intricately interact, e.g., the content of SOM and clay minerals, pH (Kabata-Pendias 2004; Violante et al. 2010), redox potential and base cation saturation (McBride 1989; Alloway 2014). In our study CEC and SOM were strongly correlated, $r = 0.86$ and 0.85 in the soil compositions for chard + lettuce and carrot models, respectively (Supplement S1), and CEC often occurs close to SOM in the loading plots (Supplement Figure S2).

Reduced models compared to full models

A relatively large fraction of the reduced models did not perform better than the full models. This implies that the SR_{mean} threshold may have excluded variables that contained important information for predicting the response. This was evident for Cd in chard and lettuce, as well as for all Ni models, where the reduced models, despite having RPDs > 2, showed lower predictive performance compared to the full models.

Filter methods, such as using SR_{mean} values to select soil variables, provide a simple and computationally efficient way to rank predictors. However, the reliance on a fixed threshold can be a limitation, particularly for small datasets (Mehmood et al. 2020; Westad and Marini 2022). However, in most cases, our results show that the SR filter effectively removed irrelevant or noisy variables, which improved the predictive performance. For nearly all reduced models, RPDs ≥ 1.4 still indicated their suitability for assessing trends and correlations. The exceptions were the models for Pb in chard and Cu in lettuce.

Performance of reduced models

The good performance of our PLS-R models for predicting Cd uptake in chard and lettuce align well with previous studies, e.g., the ones by McBride (2002) and Brown et al. (1998). The former reported R^2 values between 0.79 and 0.87, while the latter presented an R^2 of 0.94. For the other PTEs, our PLS-R models generally displayed higher R^2 values than previously presented models describing the control of soil geochemistry on PTE uptake by the same crop types (Hough et al. 2004; Bacigalupo and Hale 2011; dos Santos-Araujo et al. 2017; Stubberfield 2017; Cavanagh et al. 2019). Our findings align well with other studies, indicating that metals with higher solubility (e.g. Cd and Zn) are generally easier to model than those with lower solubility (e.g. Cu and Pb) (Legind and Trapp 2010; McBride et al. 2014).

The results from the LOOCV indicate that our PLS-R models will have a generally good performance also on unseen (i.e. future) samples (Figure 3, Table 4). However, caution should be taken to the small sample size, the use of a dataset with a rather limited range for some of the soil variables, and the lack of validation on independent data. Studies that have tested soil-to-plant transfer models on independent data have generally shown low predictive model performance, even when the model parameters have been within the model calibration range (Legind and Trapp 2010; Novotná et al. 2015; Stubberfield 2017; Lundgren et al. 2023). Unreliable predictions have to some extent been attributed to many models being too “site specific,” i.e. based on soil types with low variability, or local geochemistry not fully covered by the parameterization data (McLaughlin et al. 2011). Moreover, low variability of longitude, latitude, altitude, climate, temperature, humidity, precipitation, contamination sources/origin and history, and other factors might influence soil properties and their influence on the soil-to-crop PTE transfer (dos Santos-Araujo et al. 2017). Testing the models on independent data is therefore of great value in the process of model evaluation, though the use of a single test dataset has limited ability to characterize the uncertainty of the results. Given it is advantageous to use as much available data as possible to develop statistical models of this type, the availability of independent datasets is limited, and resampling methods, such as the leave-one-out cross-validation (LOOCV) applied in our study, can produce reasonable predictions of the model performance on future samples (Kuhn and Johnson 2013).

Future perspectives

A key finding of the study is that the PLS-R approach resulted in a high agreement between measured and modeled uptake of most PTEs by lettuce, chard and carrot. PLS-R can also be used as a variable selection tool, to identify predictors with high importance when explaining the variation in the modeled response variable. We used SR_{mean} as a filter to select the soil variables of highest importance for the development of reduced models, but since the full models in some cases performed better than the reduced ones, other alternative variable selection methods should be evaluated (Farrés et al. 2015; Mehmood et al. 2020; Westad and Marini 2022). To comprehensively assess the precision and accuracy of the PLS-R approach in predicting soil metal uptake by crops, forthcoming studies should also prioritize subjecting PLS-R models to independent data. Additionally, enhancing the generalizability of PLS-R models in the soil-to-crop transfer context necessitates utilizing larger datasets with a wider range of soil and crop types. It would also be valuable to use this method to evaluate different types of soil treatments, like adding organic matter, sewage sludge and biochar.

Other perspectives to explore are the uptake of other not so well documented PTEs, the use of gentler metal extraction methods and also environmental variables that could potentially affect the uptake of PTEs in crops, e.g., contamination sources, aerial deposition, climate, etc.

Conclusions

This study explores the applicability of Partial Least Squares Regression (PLS-R) in developing predictive models for crop uptake of metal contaminants, using soil geochemical variables as predictors. (Mehmood et al. 2020). Key conclusions drawn include:

- PLS-R demonstrates an ability to model and identify key variables for predicting the soil-to-crop transfer of potentially toxic elements in the presence of multicollinearity among predictor through the creation of latent variables (PLS components).
- Key variables influencing the processes controlling PTE uptake were soil pH, along with the concentration of the target PTE in the soil, often in combination with the Fe and Mn concentration in soil.
- The majority of our reduced models in step 2 of the workflow, which were based on the predictor variables that were assessed most influential in step 1, were found suitable for assessing trends and correlations, and some exhibit quantitative predictive capabilities.

In risk assessments of soil-to-crop PTE transfer, a wide range of soil parameters are often analyzed on a rather small number of soil samples. PLS-R has the ability to and is particularly suited for analysing datasets where the number of predictors is close to or exceeds the number of observations and when complex systems need to handle multicollinear parameters, which we took advantage of in this study. Furthermore, PLS-R can contribute to deeper insights about the key controlling soil variables that influence PTE accumulation in crops cultivated in contaminated soils, e.g. in urban environments, which can be used in the developments of risk assessment models, leading to higher accuracy and precision and, thus, a base for better decisions for stakeholders.

Acknowledgments

We express our gratitude to the Faculty of Health and Life Sciences at Linnaeus University for funding this study. For assistance during the soil sampling, we thank the volunteers in Copenhagen, Malmö, and Widnes. Special thanks to Thomas Nehls for coordinating Berlin soil collection, and to Agneta Wadmark of Malmö Municipality for further soil collection coordination. We appreciate the support of Eva Pohl, Karina Adbo, Terese Uddh-Söderberg, Anders Månsson, Henrik Hallberg, and Leon Norén at Linnaeus University during the greenhouse experiment. Our thanks go also to the analysts at ALS Scandinavia, particularly Ilia Rodushkin, Cora Paulukat, and Emma Engström, for their professionalism in the chemical analyses. Lastly, we acknowledge Sergey Kucheryavski at Aalborg University for his early-stage guidance on statistical analyses using the mdatools bookdown tutorial.

Author contributions

M.L. – Conceptualization and methodology, Sample collection, Data analysis, Result interpretation, Visualization, Writing – original draft preparation, Writing – review and editing. R.L.H – Conceptualization and methodology, Supervision, Content review, Writing – review and editing, Language editing. L.B. – Sample collection, Writing – review and editing. M.T. – Conceptualization and methodology, Content review, Writing – review and editing. L.T. – Sample collection, Writing – review and editing. E.M.J. – Sample collection, Writing – review and editing. D.D. – Data analysis, Content review, Writing – review and editing. A.A. – Funding acquisition, Conceptualization and methodology, Sample collection, Supervision, Content review, Writing – review and editing.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was funded by the Faculty of Health and Life Sciences at Linnaeus University.

Data availability statement

The raw data used in this study are, together with descriptions, is available at the data repository the Swedish National Data Service: https://doi.org/10.5878/mpyy-9j29

References

Abdi H. 2010. Partial least squares regression and projection on latent structure regression (PLS Regression). WIREs computational statistics [Internet]. 2(1):97–106. doi: 10.1002/wics.051.

Adamczyk-Szabela D, Markiewicz J, Wolf WM. 2015. Heavy metal uptake by herbs IV. Influence of soil pH on the content of heavy metals in Valeriana officinalis (L.). Water Air Soil Pollut. 226(4):106. doi: 10.1007/s11270-015-2360-3.

Alexander PD, Alloway BJ, Dourado AM. 2006. Genotypic variations in the accumulation of Cd, Cu, Pb and Zn exhibited by six commonly grown vegetables. Environ Pollut. 144(3):736–745. doi: 10.1016/j.envpol.2006.03.001.

Alloway BJ. 2014. Heavy metals in soils: trace metals and metalloids in soils and their bioavailability. New York, NY: Springer Verlag.

Andersen CM, Bro R. 2010. Variable selection in regression-a tutorial. J Chemom. 24(11–12): 728–737. doi: 10.1002/cem.1360.

Andersson M, Carlsson M, Ladenberger A, Morris G, Sadeghi M, Uhlbäck J. 2014. Geokemisk atlas över Sverige [Geochemical atlas of Sweden]. Uppsala: Elanders Sverige AB. http://resource.sgu.se/dokument/mineralnaring/Geokemisk atlas/text/geokemisk-atlas-over-sverige.pdf.

Antisari LV, Orsini F, Marchetti L, Vianello G, Gianquinto G. 2015. Heavy metal accumulation in vegetables grown in urban gardens. Agron Sustain Dev. 35(3):1139–1147. doi: 10.1007/s13593-015-0308-z

Antoniadis V, Levizou E, Shaheen SM, Ok YS, Sebastian A, Baum C, Prasad MNV, Wenzel WW, Rinklebe J. 2017. Trace elements in the soil-plant interface: phytoavailability, translocation, and phytoremediation—a review. Earth Sci Rev. 171:621–645. doi: 10.1016/j.earscirev.2017.06.005.

Audate PP, Fernandez MA, Cloutier G, Lebel A. 2019. Scoping review of the impacts of urban agriculture on the determinants of health. BMC Public Health. 19(1):672. doi: 10.1186/s12889-019-6885-z.

Bacigalupo C, Hale B. 2011. Soil-plant transfer factors for garden produce from contaminated soils: site specific versus generic estimates for As and Pb. Human Ecol Risk Assess. 17(2):394–413. doi: 10.1080/10807039.2011.552396.

Ballabio C, Lugato E, Fernández-Ugalde O, Orgiazzi A, Jones A, Borrelli P, Montanarella L, Panagos P. 2019. Mapping LUCAS topsoil chemical properties at European scale using Gaussian process regression. Geoderma. 355:113912. doi: 10.1016/j.geoderma.2019.113912.

Beesley L, Moreno-Jiménez E, Jenn P, Lepp NW. 2020. Carbon and metal(loid)s in parkland and road verge surface soils in the city of Liverpool, UK. Agronomy. 10(3):335. doi: 10.3390/agronomy10030335.

Bendt P, Barthel S, Colding J. 2013. Civic greening and environmental learning in public-access community gardens in Berlin. Landsc Urban Plan. 109(1):18–30. doi: 10.1016/j.landurbplan.2012.10.003.

Bešter PK, Lobnik F, Eržen I, Kastelec D, Zupan M. 2013. Prediction of cadmium concentration in selected home-produced vegetables. Ecotoxicol Environ Saf. 96:182–190. doi: 10.1016/j.ecoenv.2013.06.011.

Bielińska EJ, Mocek-Płóciniak A. 2010. Impact of ecochemical soil conditions on selected heavy metals content in garden allotment vegetables. Pol J Environ Stud. 19(5):896–900.

Bieri D, Joshi N, Wende W, Kleinschroth F. 2024. Increasing demand for urban community gardening before, during and after the COVID-19 pandemic. Urban For Urban Green. 92:128206. doi: 10.1016/j.ufug.2024.128206.

Boulesteix AL, Strimmer K. 2007. Partial least squares: a versatile tool for the analysis of high-dimensional genomic data. Brief Bioinform. 8(1):32–44. doi: 10.1093/bib/bbl016.

de Brogniez D, Ballabio C, Stevens A, Jones RJA, Montanarella L, van Wesemael B. 2015. A map of the topsoil organic carbon content of Europe generated by a generalized additive model. Eu J Soil Sci. 66(1):121–134. doi: 10.1111/ejss.12193.

Brown SL, Chaney RL, Angle JS, Ryan JA. 1998. The phytoavailability of cadmium to lettuce in long-term biosolids-amended soils. J Env Qual. 27(5):1071–1078. doi: 10.2134/jeq1998.00472425002700050012x.

Carrascal LM, Galván I, Gordo O. 2009. Partial least squares regression as an alternative to current regression methods used in ecology. Oikos. 118(5):681–690. doi: 10.1111/j.1600-0706.2008.16881.x.

Cavanagh JAE, Yi Z, Gray CW, Munir K, Lehto N, Robinson BH. 2019. Cadmium uptake by onions, lettuce and spinach in New Zealand: implications for management to meet regulatory limits. Sci Total Environ. 668:780–789. doi: 10.1016/j.scitotenv.2019.03.010.

Chang C-W, Laird DA, Mausbach MJ, Hurburgh CR. 2001. Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties. Soil Sci Soc Am J. 65(2):480–490. doi: 10.2136/sssaj2001.652480x.

Clark HF, Brabander DJ, Erdil RM. 2006. Sources, sinks, and exposure pathways of lead in urban garden soil. J Environ Qual. 35(6):2066–2074. doi: 10.2134/jeq2005.0464.

Dalcorso G, Manara A, Furini A. 2013. An overview of heavy metal challenge in plants: from roots to shoots. Metallomics. 5(9):1117–1132. doi: 10.1039/c3mt00038a.

Datko-Williams L, Wilkie A, Richmond-Bryant J. 2014. Analysis of U.S. soil lead (Pb) studies from 1970 to 2012. Sci Total Environ. 468–469:854–863. doi: 10.1016/j.scitotenv.2013.08.089.

Efroymson RA, Sample BE, Suter GW. 2001. Uptake of inorganic chemicals from soil by plant leaves: regressions of field data. Environ Toxicol Chem. 20(11):2561–2571. doi: 10.1002/etc.5620201123.

EFSA. 2009. Scientific opinion of on cadmium in food, EFSA panel on contaminants in the food chain (CONTAM). EFSA J. 7(3):980. doi: 10.2903/j.efsa.2009.980.

EFSA. 2010. Scientific opinion on lead in food, EFSA panel on contaminants in the food chain (CONTAM). EFSA J. 8(4):1–151. doi: 10.2903/j.efsa.2010.1570.

Engström E, Stenberg A, Senioukh S, Edelbro R, Baxter DC, Rodushkin I. 2004. Multi-elemental characterization of soft biological tissues by inductively coupled plasma–sector field mass spectrometry. Anal Chim Acta. 521(2):123–135. doi: 10.1016/j.aca.2004.06.030.

Eriksson J. 2021. Tillståndet i svensk åkermark och gröda. Uppsala. Sveriges lantbruksuniversitet. (Ekohydrologi, 168). http://pub.epsilon.slu.se.

Eriksson J, Mattsson L, Söderström M. 2010. Tillståndet i svensk åkermark och gröda. Naturvårdsverkets Rapport 6349. Stockholm.

Farrés M, Platikanov S, Tsakovski S, Tauler R. 2015. Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation. J Chemom. 29(10):528–536. doi: 10.1002/cem.2736.

Ge Y, Murray P, Hendershot WH. 2000. Trace metal speciation and bioavailability in urban soils. Environ Pollut. 107(1):137–144. doi: 10.1016/S0269-7491(99)00119-0.

Ge Y, Murray P, Sauvé S, Hendershot W. 2002. Low metal bioavailability in a contaminated urban site. Environ Toxicol Chem. 21(5):954–961. doi: 10.1002/etc.5620210509.

Gray CW, Mclaren RG. 2006. Soil factors affecting heavy metal solubility in some New Zealand soils. Water Air Soil Pollut. 175(1–4):3–14. doi: 10.1007/s11270-005-9045-2.

Hastie T, Tibshirani R, Friedman J. 2009. The elements of statistical learning: data mining, inference, and prediction. 2nd ed. New York, NY: Springer. doi: 10.1007/978-0-387-84858-7

Hough RL, Breward N, Young SD, Crout NMJ, Tye AM, Moir AM, Thornton I. 2004. Assessing potential risk of heavy metal exposure from consumption of home-produced vegetables by urban populations. Environ Health Perspect. 112(2):215–221. doi: 10.1289/ehp.5589.

Hough RL, Tye AM, Crout NMJ, McGrath SP, Zhang H, Young SD. 2005. Evaluating a “Free Ion Activity Model” applied to metal uptake by Lolium perenne L. grown in contaminated soils. Plant Soil. 270(1):1–12. doi: 10.1007/s11104-004-1658-5.

James G, Witten D, Hastie T, Tibshirani R. 2013. Linear regression. In: An introduction to statistical learning Springer texts in statistics. Vol. 103. New York: Springer; p. 59–126. doi: 10.1007/978-1-4614-7138-7_3.

Jorheim L, Åstrand C, Sundström B, Engman J, Kollander B. 2016. Frukt, bär, grönsaker och svamp. Metaller i livsmedel - fyra decenniers analyser. Swedish Food Agency Report 2016:10 [Internet]. Stockholm; [accessed 2023 Nov 22]. https://www.livsmedelsverket.se/globalassets/publikationsdatabas/rapporter/2016/metaller-i-livsmedel–-fyra-decenniers-analyser-av-frukt-bar-gronsaker-och-svamp_rapport_10_2016.pdf.

Kabata-Pendias A. 2004. Soil–plant transfer of trace elements—an environmental issue. Geoderma. 122(2–4):143–149. doi: 10.1016/j.geoderma.2004.01.004.

Kabata-Pendias A. 2010. Trace elements in soil and plants. 4th ed. Boca Raton: CRC Press.

Kirkby E. 2012. Chapter 1 – Introduction, definition and classification of nutrients. In: Marschner P, editor. Marschner's mineral nutrition of higher plants. 3rd ed. San Diego: Academic Press; p. 3–5. doi: 10.1016/B978-0-12-384905-2.00001-7.

Kucheryavskiy S. 2020. mdatools – R package for chemometrics. Chemom Intell Lab Syst. 198: 103937. doi: 10.1016/j.chemolab.2020.103937.

Kuhn M, Johnson K. 2013. Applied predictive modeling with applications in R. y. New York: Springer. doi: 10.1007/978-1-4614-6849-3.

Kvalheim OM. 2010. Interpretation of partial least squares regression models by means of target projection and selectivity ratio plots. J Chemom. 24(7–8):496–504. doi: 10.1002/cem.1289.

Kvalheim OM. 2020. Variable importance: comparison of selectivity ratio and significance multivariate correlation for interpretation of latent-variable regression models. J Chemom. 34(4): e3211. doi: 10.1002/cem.3211.

Langemeyer J, Camps-Calvet M, Calvet-Mir L, Barthel S, Gómez-Baggethun E. 2018. Stewardship of urban ecosystem services: understanding the value(s) of urban gardens in Barcelona. Landsc Urban Plan. 170:79–89. doi: 10.1016/j.landurbplan.2017.09.013.

Langemeyer J, Madrid-Lopez C, Mendoza Beltran A, Villalba Mendez G. 2021. Urban agriculture—a necessary pathway towards urban resilience and global sustainability? Landsc Urban Plan. 210:104055. doi: 10.1016/j.landurbplan.2021.104055.

Legind CN, Trapp S. 2010. Comparison of prediction methods for the uptake of As, Cd and Pb in carrot and lettuce. SAR QSAR Environ Res. 21(5–6):513–525. doi: 10.1080/1062936X.2010.502296.

Liang Z, Ding Q, Wei D, Li J, Chen S, Ma Y. 2013. Major controlling factors and predictions for cadmium transfer from the soil into spinach plants. Ecotoxicol Environ Saf. 93:180–185. doi: 10.1016/j.ecoenv.2013.04.003.

Lijzen J, Baars A, Otte P, Rikken M, Swartjes F, Verbruggen E, van Wezel A. 2001. Technical evaluation of the intervention values for soil/sediment and groundwater. RIVM Report. 711701(February); p. 147.

Lima FS, Nascimento CWA, Sousa CS. 2015. Lead and nutrient allocation in vegetables grown in soil from a battery site. Sem Ci Agr. 36(4):2483–2491. doi: 10.5433/1679-0359.2015v36n4p2483.

Liu Z, He C, Zhou Y, Wu J. 2014. How much of the world's land has been urbanized, really? A hierarchical framework for avoiding confusion. Landsc Ecol. 29(5):763–771. doi: 10.1007/s10980-014-0034-y.

Lundgren M, Troldborg M, Stubberfield J, Augustsson A, Hough RL. 2023. Predictive modeling of plant uptake of Pb and Cd: implications of aerial deposition and the origin of parameterisation data. Environ Chall. 12:100734. doi: 10.1016/j.envc.2023.100734.

McBride MB. 1989. Reactions controlling heavy metal solubility in soils. Advances in Soil Science. 10:1–56. doi: 10.1007/978-1-4613-8847-0_1.

McBride MB. 2002. Cadmium uptake by crops estimated from soil total Cd and pH. Soil Sci. 167(1):62–67. https://journals.lww.com/soilsci/Fulltext/2002/01000/CADMIUM_UPTAKE_BY_CROPS_ESTIMATED_FROM_SOIL_TOTAL.6.aspx.

McBride MB, Shayler HA, Spliethoff HM, Mitchell RG, Marquez-Bravo LG, Ferenz GS, Russell-Anelli JM, Casey L, Bachman S. 2014. Concentrations of lead, cadmium and barium in urban garden-grown vegetables: the impact of soil variables. Environ Pollut. 194:254–261. doi: 10.1016/j.envpol.2014.07.036.

McLaughlin MJ, Smolders E, Degryse F, Rietra R. 2011. Uptake of metals from soil into vegetables – chapter 8. In: Swartjes FA, editor. Dealing with contaminated sites from theory towards practical application. Dordrecht: Springer; p. 325–367.

Mehmood T, Liland KH, Snipen L, Sæbø S. 2012. A review of variable selection methods in partial least squares regression. Chemom Intell Lab Syst. 118:62–69. doi: 10.1016/j.chemolab.2012.07.010.

Mehmood T, Sæbø S, Liland KH. 2020. Comparison of variable selection methods in partial least squares regression. J Chemom. 34(6):e3226. doi: 10.1002/cem.3226.

Mitchell RG, Spliethoff HM, Ribaudo LN, Lopp DM, Shayler HA, Marquez-Bravo LG, Lambert VT, Ferenz GS, Russell-Anelli JM, Stone EB, et al. 2014. Lead (Pb) and other metals in New York City community garden soils: factors influencing contaminant distributions. Environ Pollut. 187:162–169. doi: 10.1016/j.envpol.2014.01.007.

Nengsih TA, Bertrand F, Maumy-Bertrand M, Meyer N. 2019. Determining the number of components in PLS regression on incomplete data set. Stat Appl Genet Mol Biol. 18(6):20180059. doi: 10.1515/sagmb-2018-0059.

Novotná M, Mikeš O, Komprdová K. 2015. Development and comparison of regression models for the uptake of metals into various field crops. Environ Pollut. 207:357–364. doi: 10.1016/j.envpol.2015.09.043.

Orsini F, Kahane R, Nono-Womdim R, Gianquinto G. 2013. Urban agriculture in the developing world: A review. Agron Sustain Dev. 33(4):695–720. doi: 10.1007/s13593-013-0143-z

Otte P, Lijzen J, Otte J, Swartjes F, Versluijs C. 2001. Evaluation and revision of the CSOIL parameter set. RIVM Report 711701021 [Internet]. (March); p. 1–125. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:RIVM+report+711701021+Evaluation+and+revision+of+the+CSOIL+parameter+set#0.

Pavao-Zuckerman MA. 2008. The nature of urban soils and their role in ecological restoration in cities. Restor Ecol. 16(4):642–649. doi: 10.1111/j.1526-100X.2008.00486.x.

Qvarforth A, Lundgren M, Rodushkin I, Engström E, Paulukat C, Hough RL, Moreno-Jiménez E, Beesley L, Trakal L, Augustsson A. 2022. Future food contaminants: an assessment of the plant uptake of technology-critical elements versus traditional metal contaminants. Environ Int. 169:107504. doi: 10.1016/J.ENVINT.2022.107504.

R Core Team. 2022. R: a Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

Rajalahti T, Arneberg R, Kroksveen AC, Berle M, Myhr KM, Kvalheim OM. 2009. Discriminating variable test and selectivity ratio plot: quantitative tools for interpretation and variable (biomarker) selection in complex spectral or chromatographic profiles. Anal Chem. 81(7):2581–2590. doi: 10.1021/ac802514y.

Ribeiro A, Madureira L, Carvalho R. 2023. Evidence on how urban gardens help citizens and cities to enhance sustainable development. Review and bibliometric analysis. Landsc Urban Plan. 236:104766. doi: 10.1016/j.landurbplan.2023.104766.

Riddle RL, Siebecker MG, Weindorf DC, Shaw RK, Scharenbroch BC. 2022. Soils in urban and built environments: pedogenic processes, characteristics, mapping, and classification. In: Sparks DL, editor. Advances in agronomy. Vol. 173. Academic Press Inc.; p. 227–255. doi: 10.1016/bs.agron.2022.02.004.

Rodushkin I, Ruth T, Huhtasaari Å. 1999. Comparison of two digestion methods for elemental determinations in plant material by ICP techniques. Anal Chim Acta. 378(1–3):191–200. doi: 10.1016/S0003-2670(98)00635-7.

Romic M, Zovko M. 2011. Soil contamination by trace metals: geochemical behaviour as an element of risk assessment. In: Dar IA, Dar MA, editors. Earth and environmental sciences [Internet]. Rijeka: IntechOpen. doi: 10.5772/25448.

Römkens PFAM, Groenenberg J, Bonten LTC, Vries W, Bril J. 2004. Derivation of partition relationships to calculate Cd, Cu, Ni, Pb, Zn solubility and activity in soil solutions. Wageningen: Alterra. Alterra-report 305:1–72.

Römkens PFAM, Guo HY, Chu CL, Liu TS, Chiang CF, Koopmans GF. 2009. Prediction of Cadmium uptake by brown rice and derivation of soil-plant transfer models to improve soil protection guidelines. Environ Pollut. 157(8–9):2435–2444. doi: 10.1016/j.envpol.2009.03.009.

Rouillon M, Harvey PJ, Kristensen LJ, George SG, Taylor MP. 2017. VegeSafe: a community science program measuring soil-metal contamination, evaluating risk and providing advice for safe gardening. Environ Pollut. 222:557–566. doi: 10.1016/j.envpol.2016.11.024.

Salehi MH, Beni OH, Harchegani HB, Borujeni IE, Motaghian HR. 2011. Refining soil organic matter determination by loss-on-ignition. Pedosphere. 21(4):473–482. doi: 10.1016/S1002-0160(11)60149-5.

dos Santos-Araujo SN, Swartjes FA, Versluijs KW, Moreno FN, Alleoni LRF. 2017. Soil–plant transfer models for metals to improve soil screening value guidelines valid for São Paulo, Brazil. Environ Monit Assess. 189(12):615. doi: 10.1007/s10661-017-6298-3.

Sauvé S, Hendershot W, Allen HE. 2000. Solid-solution partitioning of metals in contaminated soils: dependence on pH, total metal burden, and organic matter. Environ Sci Technol. 34(7): 1125–1131. doi: 10.1021/es9907764.

Schulte EE, Hopkins BG. 1996. Estimation of organic matter by weight loss-on-ignition. In: Magdoff FR, Tabataba MA, Hanlon EA, editors. Soil organic matter: analysis and interpretation, SSSA Special Publication Number 46. Madison: SSSA; p. 21–31.

Stubberfield J. 2017. The health benefits and risks of growing-your-own produce in an urban environment [Ph D thesis] [Internet]. Nottingham, UK: University of Nottingham. [accessed 2024 Mar 16]. https://eprints.nottingham.ac.uk/49345/.

Stubberfield J, Troldborg M, Ander L, Crout N, Young SD, Hough RL. 2022. Exercise, urban food production, preparation and consumption: implications, benefits and risks to grow-your-own (GYO) gardeners. Agronomy. 12(1):181. doi: 10.3390/agronomy12010181.

Suda A, Makino T. 2016. Functional effects of manganese and iron oxides on the dynamics of trace elements in soils with a special focus on arsenic and cadmium: A review. Geoderma. 270: 68–75. doi: 10.1016/j.geoderma.2015.12.017.

Szolnoki Z, Farsang A, Puskás I. 2013. Cumulative impacts of human activities on urban garden soils: origin and accumulation of metals. Environ Pollut. 177:106–115. doi: 10.1016/j.envpol.2013.02.007.

Tchounwou PB, Yedjou CG, Patlolla AK, Sutton DJ. 2012. Heavy metal toxicity and the environment. Exp Suppl. 101:133–164. doi: 10.1007/978-3-7643-8340-4_6.

Thomas GW. 1996. Soil pH and soil acidity. In: Sparks DL, Page AL, Helmke PA, Loeppert RH, Soltanpour PN, Tabatabai MA, Johnston CT, Sumner ME, editors. Methods of soil analysis. John Wiley & Sons, Ltd; p. 475–490. doi: 10.2136/sssabookser5.3.c16.

Violante A, Cozzolino V, Perelomov L, Caporale AG, Pigna M. 2010. Mobility and bioavailability of heavy metals and metalloids in soil environments. J Soil Sci Plant Nutr. 10(3):268–292. doi: 10.4067/S0718-95162010000100005.

Wakeling IN, Morris JJ. 1993. A test of significance for partial least squares regression. J Chemom. 7(4):291–304. doi: 10.1002/cem.1180070407.

Wang J, Wang PM, Gu Y, Kopittke PM, Zhao FJ, Wang P. 2019. Iron-manganese (oxyhydro)oxides, rather than oxidation of sulfides, determine mobilization of Cd during soil drainage in paddy soil systems. Environ Sci Technol. 53(5):2500–2508. doi: 10.1021/acs.est.8b06863.

Westad F, Marini F. 2015. Validation of chemometric models – A tutorial. Anal Chim Acta. 893: 14–24. doi: 10.1016/j.aca.2015.06.056.

Westad F, Marini F. 2022. Variable selection and redundancy in multivariate regression models. Front Anal Sci. 2:897605. doi: 10.3389/frans.2022.897605.

Williams PC, Sobering DC. 1993. Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. J Near Infrared Spectrosc. 1(1): 25–32. doi: 10.1255/jnirs.3.

Wold S, Sjöström M, Eriksson L. 2001. PLS-regression: a basic tool of chemometrics. Chemom Intell Lab Syst. 58(2):109–130. doi: 10.1016/S0169-7439(01)00155-1.

Wu S, Peng S, Zhang X, Wu D, Luo W, Zhang T, Zhou S, Yang G, Wan H, Wu L. 2015. Levels and health risk assessments of heavy metals in urban soils in Dongguan, China. J Geochem Explor. 148:71–78. doi: 10.1016/j.gexplo.2014.08.009.

Xiao W, Ye X, Zhang Q, Chen D, Hu J, Gao N. 2018. Evaluation of cadmium transfer from soil to leafy vegetables: influencing factors, transfer models, and indication of soil threshold contents. Ecotoxicol Environ Saf. 164(March):355–362. doi: 10.1016/j.ecoenv.2018.08.041.

Yi Z, Lehto NJ, Robinson BH, Cavanagh JAE. 2020. Environmental and edaphic factors affecting soil cadmium uptake by spinach, potatoes, onion and wheat. Sci Total Environ. 713:136694. doi: 10.1016/j.scitotenv.2020.136694.

Zakari S, Jiang X, Zhu X, Liu W, Allakonon MGB, Singh AK, Chen C, Zou X, Akponikpè PBI, Dossa GGO, et al. 2021. Influence of sulfur amendments on heavy metals phytoextraction from agricultural contaminated soils: a meta-analysis. Environ Pollut. 288:117820. doi: 10.1016/j.envpol.2021.117820.