Available online at www.sciencedirect.com

ScienceDirect

Journal of Environmental Sciences 2011, 23(1) 70–78

JOURNAL OF ENVIRONMENTAL SCIENCES
ISSN 1001-0742
CN 11-2629/X
www.jesc.ac.cn

Predicting As, Cd and Pb uptake by rice and vegetables using field data from China

Hongzhen Zhang[1,2], Yongming Luo[1,*], Jing Song[1], Haibo Zhang[1], Jiaqi Xia[3], Qiguo Zhao[1]

1. Key Laboratory of Soil Environment and Pollution Remediation, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China. E-mail: hongzhenzhang@126.com
2. Chinese Academy for Environmental Planning of Ministry of Environmental Protection, Beijing 100012, China
3. Nanjing Institute of Environmental Science, Ministry of Environmental Protection, Nanjing 210042, China

Received 10 May 2010; revised 22 June 2010; accepted 01 July 2010

Abstract

Plant uptake factor (PUF), single-variable regression of natural log-transformed concentrations in rice grain/vegetables versus natural log-transformed concentrations in soil and multiple-variable regression with soil concentrations and pH, was derived, validated and compared based on the paired crop and soil data collected from studies regarding As, Cd and Pb contaminated croplands in China. Results showed that the median value of PUF did not present deterministic prediction. But after natural logarithm transformation, the PUF followed Gaussian distribution which could be useful in risk assessment. The single-variable regression models were significant for As, Cd and Pb uptake both by rice and vegetables; however, the standard errors of all the regressions were comparatively large. Soil pH as a variable was generally significant but it only contributed positively to model fit for Cd uptake. After model comparison and selection, the upper 95% prediction limits of the multiple regression model for Cd uptake by rice was recommended to calculate screening value of Cd for paddy soil based on the limit for Cd concentration in rice grain.

Key words: trace elements; regression model; rice; vegetable; plant uptake factor; croplands

DOI: 10.1016/S1001-0742(10)60375-0

Citation: Zhang H Z, Luo Y M, Song J, Zhang H B, Xia J Q, Zhao Q G, 2011. Predicting As, Cd and Pb uptake by rice and vegetables using field data from China. *Journal of Environmental Sciences*, 23(1): 70–78.

Introduction

In recent years, trace elements transfer in the soil-crop system attracted more attention since it is the major exposure pathway of human to contaminated croplands. Understanding and prediction of heavy metals and metalloids accumulation in crops grown on contaminated croplands were motivated to secure crop quality and food safety (McLaughlin et al., 1999; Adams et al., 2004). Previous studies have shown that soil properties, agricultural management practices and seasonal conditions were crucial factors which determined availability of trace elements to crops (Kabata-Pendias and Pendias, 2001). Additionally, significant differences in heavy metal uptake were observed among cultivars (Wenzel et al., 1996). In view of the complex nature of trace element transfer mechanisms and financial or time constraints for site-specific measurements, empirical models were commonly developed to describe trace elements transfer in different soil-crop systems and over large geographical regions (Alsop et al., 1996; Efroymson et al., 2001; Chen et al., 2009). Inclusion of speciation and bioavailability of trace elements in soils can largely reduce model uncertainty, but its application was limited to specific contaminated sites due to data deficiency for large areas (Zarcinas et al., 2004; Simmons et al., 2008; Kuo et al., 2004). Thus, total soil trace element concentration and other routinely measured soil properties were commonly used in model development for generic risk assessment and/or derivation of soil screening benchmark.

Plant uptake factor (PUF), which was defined as the ratio of the element concentration in plant tissue to that in soils, was commonly used to characterize trace element transfer from soil to plant (Bingham, 1979; Cui et al., 2004; Sample et al., 1999; Suter et al., 2000). In risk assessment and soil benchmark derivation, PUF has been used to estimate the uptake of potentially toxic trace elements by plants from soils and applied to normalize the differences caused by soils properties thus standardize the assessments of human exposure to trace element from croplands (Chen et al., 2009). It is imperative that the PUF values can be characterized using field data (Wang et al., 2006).

The evidence has indicated that the uptake of trace

* Corresponding author. E-mail: ymluo@issas.ac.cn

elements is strongly affected by soil properties and it is better to model the uptake based on the soil characteristics which were most likely to control bioavailability and commonly reported in geochemical surveys and experimental datasets (soil trace element concentration, pH and organic matter) (Hough et al., 2003). Various regression models linking trace elements concentrations in soil to that in crops have been already developed, and soil pH, organic matter content, calcium content or other soil properties were included to improve the prediction ability of the model.

Previous studies showed significant regressions for the uptake of trace elements by plant leaves (Efroymson et al., 2001), vegetables (Hough et al., 2004) and earthworms (Sample et al., 1999; Nahmani et al., 2007) using log-transformed concentrations. Efroymson et al. (2001) recommended the multiple regression models with soil As, Cd, Cu, Pb, Hg, Ni, Se and Zn concentrations and soil pH to predict the uptake of these elements by plant leaves (R^2: 0.27–0.85). McBride (2002) suggested that combination of soil pH and total soil Cd was reasonably predictive of Cd concentration in above-ground plant tissue, and the variance of coefficients depended on the particular soil, climate and crop types.

There were also attempts in which soil basic properties were combined to explain trace elements uptake by rice, wheat and barley grain directly from soil total concentrations. The concentrations of Cd, Cu, Ni, Pb and Zn in maize and wheat were predicted using soil total concentrations and pH in a sewage disposal site with 41%–86% and 9%–92% of the variance were accounted for respectively (Hough et al., 2003).

The PUF methods and regression models have been employed widely to investigate the relationship between soil properties and uptake of trace elements under field conditions (McBride et al., 2002). Median PUF and regression models, the 90th percentile of PUF and the upper 95% prediction limits of regression models using total trace element concentrations in soil have been employed to describe trace element uptake by earthworms and plants leaves, and this method was justified for risk assessment and derivation of soil screening levels (Sample et al., 1999). However, the use of these models depends heavily on the ranges of data they were derived from and the reliability has to be identified by model validation before these models can be used (Sauve et al., 2000; Bonten et al., 2008).

The studies mentioned above only included relatively small contaminated croplands or a few contamination sources and may not represent a wider geographical region compared with data collection scale in this study. In China, more than 20 million ha^2 of cropland has been contaminated by toxic trace elements coming from wastewater irrigation, sewage sludge, fertilizers, vehicle and other urban emissions, mine wastes and smelter depositions (SEPA, 2003). Intake of trace elements from dietary sources may represent a significant exposure pathway for human populations. However, the current Soil Environment Standard (GB15618-1995) of China came into force in 1995 and the tolerable values of trace element contents in soil were based on limited information. Prediction of trace element concentrations in edible part of crops based on trace element concentration in soil and other environmental factors are urgently required for human risk assessment and derivation of soil environmental benchmark croplands.

In this study, our aim is to compile paired soil and crop data from relative studies in China to derive plant uptake factors and regression models for generic human health risk assessment as well as derivation of national soil benchmark for croplands. We assumed that there was a significant relationship between natural log-transformed rice grain/vegetables trace element concentrations and natural log-transformed soil element concentrations with or without soil pH. The transfer of As, Cd and Pb in soil-crop system was predicted based on regression models developed from literature data. The best model was chosen to derive Cd benchmark for cropland of China based on crop Cd concentration limit.

1 Methods

1.1 Data collection

Data used in this study were collected from three sources. Firstly, data were collected from literature in the last three decades, in which trace element contamination of China's croplands were reported (1301 observations). Secondly, unpublished data sampled from a rice-growing valley covering an area of 10.9 km^2 in Fuyang, Zhejiang Province, Southeast China (504 observations). This valley was contaminated with a mixture of heavy metals due to Zn fertilizer production in the 1950s and widespread Cu smelting activities since the 1980s. Thirdly, we also obtained data from the Soil Environmental Capacity of Trace Element Research Program ((75)-60-02-03) in which the transfer of As, Cd, Cu and Pb to crops from soils of seven main types in China were investigated in the 1980s (272 observations). Based on the collected data, a database was compiled consisting of As, Cd and Pb concentrations in paired soils and crops, soil pH, and crop species. The soil trace element concentration ($Q_{\rm soil}$) and corresponding crop trace element concentration ($Q_{\rm crop}$) measurements were collected as essential requirements for inclusion in the database. There was obviously much variability among the various experimental protocols complied from more than 150 different studies. It is not practical to thoroughly review all of the individual methodologies and analytical techniques, but the data need to be screened and selected carefully and systematically before they can be used for model development.

Studies in which trace elements were freshly added to soils in the form of inorganic salts were excluded from the database because early studies have clearly shown that the bioavailability of trace elements differed significantly between field investigation data and salt-added experiment data (Efroymson et al., 2001). In this study, field studies included data derived from cropland survey or pot experiments in glasshouse where the crops were grown in real contaminated soils. Data from studies in which the growth media were non-soil media such as specific clays, oxides or sand, hazardous waste, sediment were excluded. Observations were included in the database if the total concentrations in soil and crop were measured by strong acid digestion accompanied by heat; concentrations in soils measured by extraction with diethylenetriamane pentaacetic acid-octreotide, acetic acid, or some other mild extraction methods were excluded.

Field investigation data were excluded if any other contaminant sources existed besides soil during crop growth season. For studies in which soil samples were taken at multiple depths of contaminated sites, the concentrations at the topsoil were recorded. Data from studies in which the plant material was washed using ultrapure water before the digestion were included. Data from studies in which measured concentrations were lower than detection thresholds were excluded.

For studies which reported trace element concentrations in rice grain, only if samples were collected at or shortly before harvest and the concentrations were expressed on dry weight basis were included. Different cultivars of rice were not distinguished because in most studies rice cultivars were not reported. For trace element concentrations of vegetables, many studies reported element concentrations on fresh weight basis using wet digestion methods. These concentrations were converted to concentrations on dry weight basis using mean water contents of each vegetable species reported by U.S. Environment Protection Agency (US EPA, 1997). Vegetables were divided into two subgroups: leaf/stem vegetables (e.g., spinach, Chinese cabbage, celery) and non-leaf/stem vegetables (e.g., carrot, radishe, pepper, towelgourd, cucumber). Only data from studies which reported leaf/stem vegetable concentrations were included because the plant uptake of trace elements as well as their distribution among different tissues varies dramatically between species (Wenzel et al., 1996).

1.2 Model development and validation

The transfer of trace elements from soil to plant tissues can be depicted by plant uptake factor, PUF, given as:

$$\mathrm{PUF} = \frac{Q_{\mathrm{crop}}}{Q_{\mathrm{soil}}} \tag{1}$$

where, Q_{crop} is the trace element concentration in rice grain and leaf/stem vegetables, and Q_{soil} is the corresponding trace element concentration in soil. The obtained PUF were transformed into natural logarithm form and then fitted to Gaussian distribution following the method described by Chen et al. (2009). The cumulative probability distribution, median value and other statistical characteristics of PUF were compared with results of other published PUF.

Single-variable regressions using ln-transformed Q_{crop} and Q_{soil} were developed (Eq. (2)). Multiple-variable regressions were also developed by the combination of soil pH (Eq. (3)).

$$\ln Q_{\mathrm{crop}} = \beta + \beta_2 \times \ln Q_{\mathrm{soil}} \tag{2}$$

$$\ln Q_{\mathrm{crop}} = \beta + \beta_1 \times \mathrm{pH} + \beta_2 \times \ln Q_{\mathrm{soil}} \tag{3}$$

where, β, β_1 and β_2 are empirical coefficients, and pH refers to the soil pH of the corresponding soil sample.

The collected data were divided into two subgroups randomly: the initial dataset (for original model development) and dataset for model validation (used to test the model accuracy and predictive utility). To evaluate the suitability of those models for general applications, estimation of trace element concentrations of rice grain and leaf/stem vegetables were predicted using the median PUF, single-variable regression and multiple regression models. Conservative estimates were also made by applying the 90th percentile PUF and the upper 95% prediction limits of both single-variable regression and multiple regression models.

The predictive utility of those models were estimated using Wilcoxon signed-rank tests. Differences were considered significant if $p \leqslant 0.05$. Proportional deviation (PD) of the predicted values from measured values were used to estimate the relative accuracy and quality of different models (Eq. (4)):

$$\mathrm{PD} = \frac{Q_{\mathrm{ai}} - Q_{\mathrm{pi}}}{Q_{\mathrm{ai}}} \tag{4}$$

where, Q_{ai} is the measured Q_{crop} at soil concentration i and Q_{pi} is the predicted Q_{crop} at soil concentration i. Relative quality of the estimation methods was evaluated by the criteria from Sample et al. (1999) and Efroymson et al. (2001), with the following criteria donating increasing predicting quality: (1) median PD closest to 0 (indicates estimates center around measured values), (2) PD with narrowest range (indicates relative accuracy of method), (3) percentage overestimation closest to 50% (indicates that estimates center around measured values), and (4) difference between estimated and measured values not significant as determined by Wilcoxon signed-rank tests. With the four criteria, the models were evaluated by weight-of-evidence. The first three criteria were weighted more than the forth criterion since the Wilcoxion test can be influenced by sample size. Relative quality of conservative estimation methods was evaluated by the smallest, negative median PD value (indicates method overestimates while minimizing the degree of overestimation), and PD with the narrowest range (to minimize the degree of overestimation).

The initial regressions were also compared with similar models developed by validation datasets using F-test procedure for regression lines comparison. Differences were considered significant if $p \leqslant 0.05$. Following model validation, the initial datasets and the validation datasets were pooled to recalculate the final PUF and regression models. These final results were compared and the best model was recommended for risk assessment and/or derivation of soil environmental benchmark. All the statistical analysis was performed using SAS software package (SAS Institute, 1988).

2 Results and discussion

After data selection, the final database has 769 observations, including 331 observations for soil-rice system and 436 observations for soil-vegetable system. The number of observations used for development of plant uptake factors and regression models ranged from 77 for Pb in soil-rice system to 161 for Pb in soil-vegetable system. The number of studies incorporated in the models ranged from 8 for As in soil-rice system to 24 for Cd in soil-vegetable system. These studies covered 18 of the 32 provinces in China and included most of the soil types. Study locations were mainly in the middle-east China which is the main crop production sources.

2.1 PUF characterizing

The concentrations of As, Cd and Pb in rice grain and vegetables did not seem to correlate with the corresponding soil concentrations (Fig. 1). The As concentrations in soil ranged from 2.45 to 169 mg/kg while the majority of the corresponding As concentrations in rice grain were below 0.6 mg/kg, but the As concentrations in vegetables were much higher (ranged from 0.003 mg/kg to more than 20 mg/kg), resulting in great differences in As PUF ranges between rice grain and vegetables. For Cd, the soil concentrations ranged from 0.02 mg/kg to around 12 mg/kg, while the corresponding Cd concentrations in rice grain were lower than 3 mg/kg. The PUF discrepancy between Cd accumulation in rice grain and vegetables was not as dramatic as As since the majority of the corresponding Cd concentrations in vegetables were also lower than 2 mg/kg which showed a relatively higher Cd accumulation in rice grain than As. For Pb, the soil concentrations ranged from 25.6 to 840 mg/kg, while the corresponding concentrations in rice grain were mainly lower than 5 mg/kg and the corresponding concentrations in vegetables were mainly below 35 mg/kg, resulting in a PUF range of 0.003–0.17 which was comparable with the PUF range for As but much lower than the PUF range for Cd (Table 1).

The predictive utility of median PUF values calculated using initial data were validated based on validation datasets. Significant differences between measured and predicted Q_{crop} were observed for 5 out of the 6 median PUF (Table 2). Therefore, the PUF is probabilistic rather than deterministic over the range of soil As, Cd and Pb

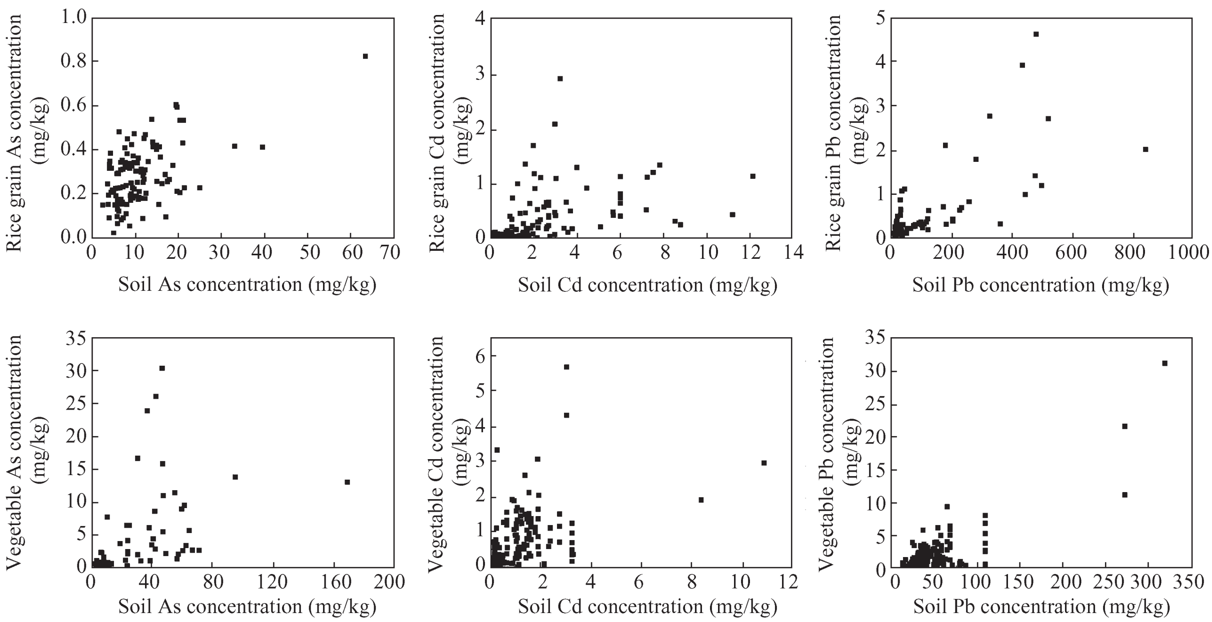

Fig. 1 Relationship between As, Cd and Pb concentrations in soil and crops.

Table 1 Summary statistics of plant uptake factor (PUF)

	n	Mean	Standard deviation	Minimum	Median	90th percentile	Maximum	Mean of lnPUF	SD of lnPUF
PUF for rice									
As	120	0.030	0.017	0.0042	0.026	0.051	0.090	–3.62	0.58
Cd	134	0.26	0.29	0.014	0.15	0.74	1.47	–1.90	1.08
Pb	77	0.0086	0.0073	0.00094	0.0052	0.022	0.031	–5.08	0.89
PUF for vegetables									
As	117	0.093	0.012	0.0016	0.043	0.23	0.68	–3.09	1.24
Cd	158	1.03	1.30	0.045	0.70	2.14	12.3	–0.43	0.96
Pb	161	0.041	0.033	0.00071	0.037	0.078	0.17	–3.57	1.00

n: observations.

Table 2 Comparison of quality of general estimation methods as determined by the proportional deviation (PD) of the estimated values from measured values

	Median uptake factor				Single-variable regression model			Regression model with pH		
	n	Median UF	Median PD (range)	Over estimated (%)	n	Median PD (range)	Over estimated (%)	n	Median PD (range)	Over estimated (%)
Soil-rice grain system										
As	81	0.026	0.22[c] (–1.83~0.77)	29.6	81	0.13[a] (–2.05~0.69)	38.3	6	0.31[a] (–0.048~0.41)	16.7
Cd	46	0.14	–0.10[c] (–7.01~0.83)	56.5	46	0.21[a] (–5.21~0.89)	47.8	46	0.19[a] (–2.46~0.85)	43.5
Pb	24	0.0092	–1.49[d] (–4.28~–0.08)	100	24	–0.65[d] (–2.19~0.54)	95.8	5	–0.62[a] (–1.29~0.50)	60
Soil-vegetable system										
As	37	0.057	–0.72[d] (–18.7~0.92)	73.0	37	–0.069[b] (–10.56~0.93)	56.8	7	0.541[b] (0.31–0.91)	0
Cd	32	0.61	0.37[a] (–12.6~0.84)	68.8	32	–0.072[a] (–19.24~0.71)	46.9	16	–0.31[a] (–4.51~0.36)	72.7
Pb	23	0.041	–3.56[d] (–56.8~0.53)	95.7	23	–2.59[d] (–41.97~0.58)	82.6	45	0.50[d] (–3.41~0.88)	97.8

[a] Wilcoxon signed-rank test ($p < 0.05$); [b] Wilcoxon signed-rank test ($0.01 < p \leqslant 0.05$); [c] Wilcoxon signed-rank test ($0.001 < p \leqslant 0.01$); [d] Wilcoxon signed-rank test ($p \leqslant 0.001$).
n: datasets used for model validation.

concentrations the data defined. Conservative predictive utility was also validated using the 90th percentile PUF, with percent overestimates ranging from 81.3% to 100% (Table 3).

The probabilistic distribution of PUF was judged using the statistical method described by Chen et al. (2009) and the results showed that both the PUF collected in China and obtained in California croplands followed Gaussian distribution (Fig. 2 and Table 4). The fitness to Gaussian distribution was better for As and Cd than Pb, and better for the PUF of vegetables than PUF of rice grain (Fig. 3). The median PUF for all the three trace elements were less than one. The distribution of PUF for As, Cd and Pb spanned at least three orders of magnitude. The PUF ranges of Cd were a little more than three orders of magnitude, and the PUF ranges of As and Pb were four orders of magnitude. The resulting PUF was summarized and compared with PUF published by other publications in Table 5.

The PUF obtained in this study were quite compatible with the PUF published in other articles (Table 5). The means of PUF obtained in this study were similar to the PUF employed by CDFA (2002) and reported by Chen et al. (2009). But the ranges of PUF obtained in this study were about one order bigger than the ranges of PUF employed in California. This is reasonable because of a larger range of element concentrations in soils and the diversity of soil types and contamination sources in the database collected in this study.

The PUF obtained in this study were also consistent with PUF that have been used for risk assessment and soil benchmark derivation in other countries. For example, PUF employed in Belgium, the Netherlands, Sweden and Norway for soil clean-up standard for residential use were 0.015–0.03 for arsenic, 0.15–0.7 for cadmium and 0.001–0.03 for lead on fresh weight (Provoost et al., 2006). Those values were within the ranges of PUF presented in this study (Table 1).

Table 3 Comparison of quality of conservative estimation methods as determined by the proportional deviation (PD) of the estimated values from measured values

	90th percentile uptake factor				Upper 95% prediction limit for single-variable regression model			Upper 95% prediction limit for regression model with pH		
Chemical	n	Median UF	Median PD (range)	Over estimated (%)	n	Median PD (range)	Over estimated (%)	n	Median PD (range)	Over estimated (%)
Soil-rice grain system										
As	81	0.050	–0.88[d] (–5.79~0.45)	87.6	81	–2.31[d] (–10.63~–0.21)	100	6	–0.32b (–1.03~–0.13)	100
Cd	46	0.59	–4.71[d] (–40.48~0.11)	93.5	46	–4.52[d] (–42.32~0.21)	95.7	46	–3.48d (–17.85~0.18)	95.7
Pb	24	0.022	–4.99[d] (–11.73~–1.60)	100	24	–0.43[d] (–13.48~–1.16)	100	5	–10.23a (–13.63~–2.48)	100
Soil-vegetable system										
As	37	0.26	–6.82[d] (–88~0.62)	97.3	37	–11.39[d] (–133~0.21)	97.3	7	–8.93b (–13.20~–4.86)	100
Cd	32	2.00	–1.13[d] (–43.96~0.48)	81.3	32	–4.85[d] (–113~0.54)	100	16	–4.94c (–24.20~–1.58)	100
Pb	23	0.084	–10.19[d] (–140.73~–0.16)	100	23	–19.88[d] (–247~–1.41)	100	45	–2.42d (–28.88~0.16)	100

[a] Wilcoxon signed-rank test ($p < 0.05$); [b] Wilcoxon signed-rank test ($0.01 < p \leqslant 0.05$); [c] Wilcoxon signed-rank test ($0.001 < p \leqslant 0.01$); [d] Wilcoxon signed-rank test ($p \leqslant 0.001$).
n: datasets used for model validation.

Table 4 Parameters for the probability distribution of PUF of As, Cd and Pb for rice grain and vegetables

Rice	x_0*	b**	R^2	Vegetable	x_0*	b**	R^2
As	–3.61(–3.66)	0.54(0.58)	0.84	As	–3.09(–3.09)	1.10(1.24)	0.92
Cd	–1.91(–1.91)	1.21(1.05)	0.86	Cd	–0.41(–0.43)	0.98(0.96)	0.93
Pb	–5.18(–5.08)	1.08(0.82)	0.58	Pb	–3.26(–3.57)	0.93(1.00)	0.76

* Number in parentheses refer to the natural logarithmic mean of the data set; ** number in parentheses refer to the standard deviation of the data set (after natural logarithmic transformation).

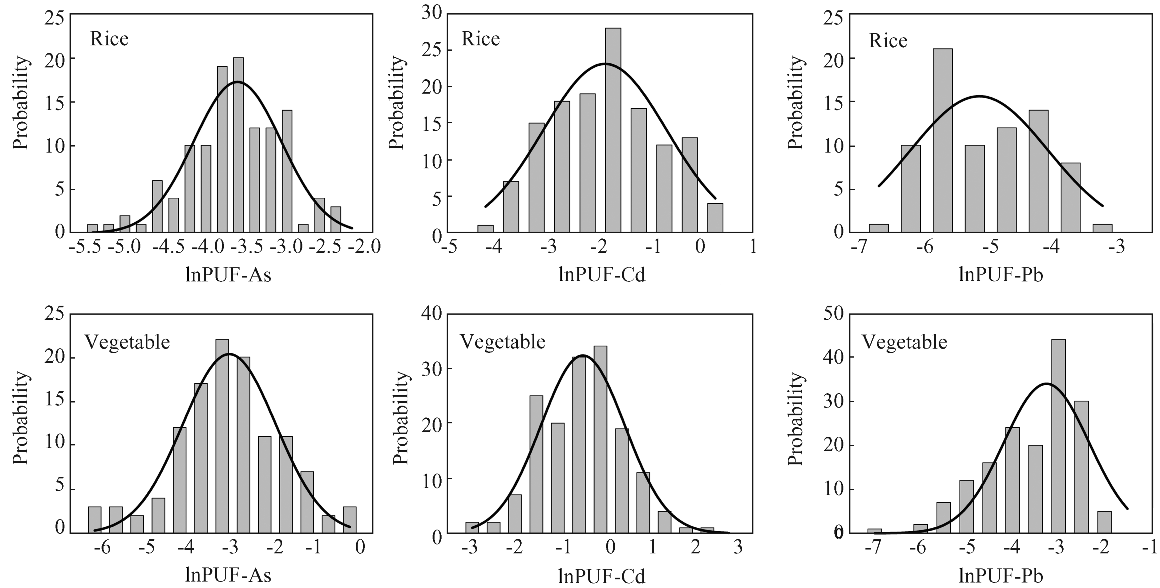

Fig. 2 Probability distributions of PUF of As, Cd and Pb for rice grain and vegetables. The data was first transformed into natural logarithm form. Then the data was fitted to Gaussian distribution (with three parameters) that: $y = a \times \exp(-0.5 \times (x - x_0)/b)^2)$, where y is the probability of PUF at value of x and a is a constant.

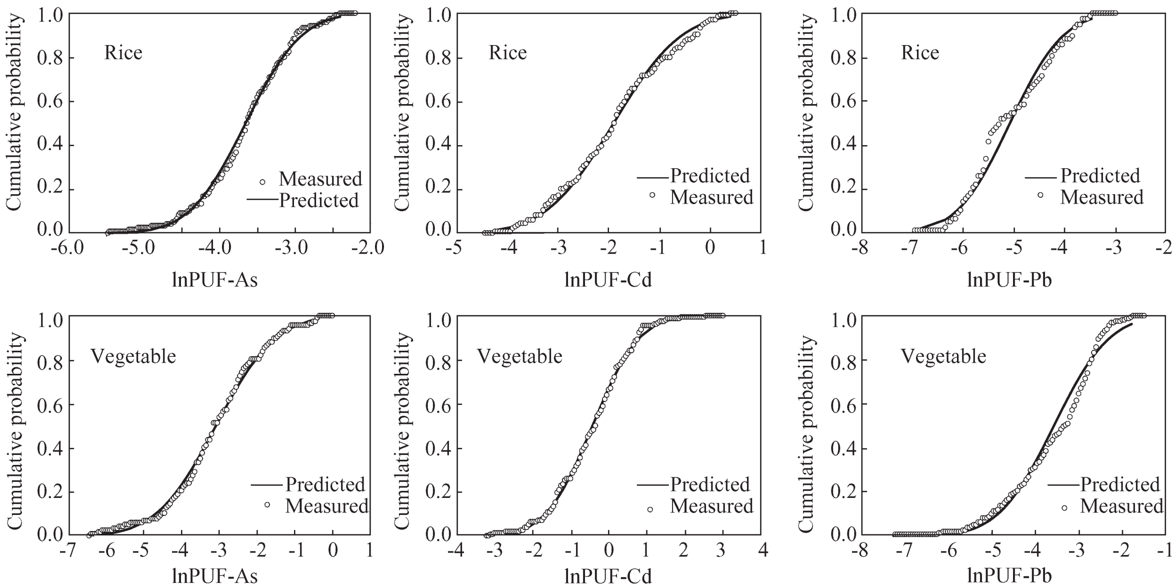

Fig. 3 Measured and predicted cumulative probability distribution of As, Cd and Pb for rice grain and vegetables. The cumulative probability distribution of PUF was predicted based on the mean(x_0) and standard deviation (b) of the population that: $F(x) = 0.5 + 0.5\mathrm{erf}((x - x_0)/\sqrt{2}b)$, where $F(x)$ donates the cumulative probability that PUF $\leqslant x$; erf: error function.

2.2 Regression models, development and validation

Based on validation datasets, significant differences between measured and predicted Q_{crop} were observed in 3 out of the 6 single-variable regression models (Table 2). This difference was found in 2 out of the 6 multiple regression models with pH. The conservative prediction utility was also validated. The 95% upper prediction limits for both single-variable and multiple regression models can significantly overestimate measured Q_{crop} and give quite good conservative estimates (Table 3).

A comparison of regression models from initial and validation data showed that 3 out of 6 single-variable regression models and 2 out of 6 multiple regression models were statistically significantly different (data not shown). Both single-variable regressions and multiple regressions produced significant model fits for As, Cd and Pb in both soil-rice and soil-vegetable systems ($p < 0.05$). The slopes of all regression models were positive. Final regression models that incorporated both initial and validation data are presented in Tables 6 and 7.

For Cd and Pb, the final single-variable regression models of Q_{crop} against Q_{soil} in soil-rice system were found to be significant ($p < 0.0001$) and explained 54%, 49% of the variance in the concentrations of Cd and Pb in rice grain, respectively. For As, the final single-variable regression model only explained 17% of the variance. The Standard errors of single-variable regressions for soil-rice

Table 5 Comparison of the means and ranges of PUF for different crops used in different publications

	Grain	Vegetable	Reference
As	0.020 (0.0044–0.060)	0.024 (0.0015–0.360)	CDFA, 2002
As	0.022 (0.0050–0.055)	0.013 (0.0012–0.069)	Chen et al., 2009
As	0.030 (0.0042–0.090)	0.093 (0.0016–0.680)	This study[a]
Cd	1.77 (0.110–8.71)	0.68 (0.004–13.3)	CDFA, 2002
Cd	0.22 (0.025–1.29)	0.22 (0.110–8.71)	Chen et al., 2009
Cd	0.26 (0.014–1.47)	1.03 (0.045–12.3	This study[a]
Pb	0.052 (0.0010–0.21)	0.014 (0.0001–0.39)	CDFA, 2002
Pb	0.084 (0.0450–0.13)	0.052 (0.0010–0.21)	Chen et al., 2009
Pb	0.009 (0.0009–0.03)	0.041 (0.0007–0.17)	This study[a]

[a] The UPF of rice grain represents Grain; the UPF of leaf/stem vegetables represent Vegetable.

Table 6 Parameter values for single-variable regression model

Chemical	n	$\beta \pm$ SE	$\beta_2 \pm$ SE	R^2_{adj}	SE
Soil-rice system					
As	120	-2.36 ± 0.20^c	0.43 ± 0.08^c	0.17	0.50
Cd	134	-1.91 ± 0.09^c	0.75 ± 0.06^c	0.54	1.01
Pb	77	-3.38 ± 0.28^c	0.58 ± 0.07^c	0.49	0.67
Soil-vegetable system					
As-leaf	117	-4.33 ± 0.33^c	1.48 ± 0.12^c	0.56	1.17
Cd-leaf	158	-0.58 ± 0.07^c	0.61 ± 0.07^c	0.35	0.87
Pb-leaf	161	-4.46 ± 0.53^c	1.27 ± 0.14^c	0.33	0.99

[c] $p \leqslant 0.001$. β and β_2 are the constants (± SE) from equation $\ln Q_{crop} = \beta + \beta_2 Q_{soil}$. n: observations.

system were quite large (0.50 for As, 1.01 for Cd and 0.67 for Pb). The results of single-variable regression models for vegetables were similar to model results for rice.

Sample size decreased due to short of soil pH values when this variable is included in the multiple regression model (Table 7). In this case, the multiple regression models can not be compared with single-variable regression models directly. As shown in Table 7, the multiple regressions resulted in significant model fits with R^2 values ranged from 0.12 to 0.71. The significance test results showed that soil pH was a significant factor in 5 out of 6 multiple regression models (Table 7). The combination of Q_{soil} and soil pH resulted in improved R^2 (from 0.54, 0.35 to 0.79, 0.70 for rice and vegetable respectively) and lower standard errors (from 1.01, 0.87 to 0.79, 0.70 respectively) for Cd uptake in both rice grain and vegetables (Table 6, 7). But for As and Pb, the combination of Q_{soil} and soil pH did not lead to significant improvement of the regression results (Tables 6 and 7).

In this study, soil pH significantly contributed to the regression model fits for Cd transfer in soil-crop systems. However, the standard errors (SE) of the multiple regression model (0.70–0.79) were obviously larger than SE reported in other publications where the cultivar was identified (Adams et al., 2004; Yu et al., 2006) or the Cd sources in croplands were restricted only to sewage sludge or fertilizer application (Hough et al., 2003). Similar results were found by Wenzel et al. (1996), who indicated Cd concentration in wheat grain was significantly influenced by soil basic characteristics and by cultivars. Significant difference in Cd accumulation in wheat and barley grain among cultivars was also found in a sewage disposal farm in UK (Adams et al., 2004).

2.3 Deriving soil benchmark for Cd in croplands

In this study, data from various sources were pooled and different statistical methods were compared to find the optimal approach to characterize and predict the plant uptake of As, Cd and Pb from contaminated croplands in China. Using the model selection criteria mentioned above, the relatively best estimates for As concentrations in both rice grain and vegetables were provided by the single-variable regression models although only 17% of the variance was accounted for, and the best estimates of Cd and Pb concentrations in both rice grain and vegetables were provided by the multiple regression models (Table 2). Among conservative prediction models, the 95% upper prediction limit for multiple regression models produced the best conservative estimates of As, Cd and Pb uptake by rice grain and vegetables except for Pb uptake by rice grain where the 95% upper prediction limit for single-variable regression models gave the best conservative estimates (Table 3). Even though the standard errors of the multiple regression models were large, the combination of soil pH improved the prediction of Cd uptake significantly (Fig. 4).

The multiple regression model of Cd uptake by rice grain was run iteratively to estimate 95% upper prediction limit values of Q_{soil} which gave rise to Q_{crop} values approaching maximum levels of Cd in human food (Fig. 5). Tolerance limit of Cd in rice grain is 0.2 mg/kg on fresh weight basis defined by Chinese Standard (GB2762-2005). Figure 5 presented a comparison of Cd values in soils derived from the 95% upper prediction limits of multiple regression model with current Chinese Environmental Quality Standard for Cd in agricultural soils (GB 15618-1995). GB 15618-1995 was quite conservative when cropland was used for rice production (pH > 6.5).

Table 7 Parameter values for multiple regression model

	n	$\beta \pm$ SE	$\beta_1 \pm$ SE	$\beta_2 \pm$ SE	R^2_{adj}	SE
Soil-rice system						
As	76	-1.68 ± 0.32^c	-0.03 ± 0.04^{ns}	0.25 ± 0.07^b	0.12	0.32
Cd	91	3.12 ± 0.78^c	-0.78 ± 0.11^c	1.00 ± 0.09^c	0.71	0.79
Pb	40	-9.55 ± 1.73^c	0.69 ± 0.19^b	0.82 ± 0.14^c	0.47	0.76
Soil-vegetable system						
As-leaf	27	-11.7 ± 2.96^b	0.79 ± 0.31^a	2.19 ± 0.66^b	0.31	1.26
Cd-leaf	59	-1.91 ± 0.79^a	-0.34 ± 0.11^b	0.86 ± 0.09^c	0.61	0.70
Pb-leaf	95	-11.4 ± 0.88^c	0.93 ± 0.10^c	1.213 ± 0.12^c	0.65	0.87

[a] $0.01 < p \leqslant 0.05$; [b] $0.001 < p \leqslant 0.01$; [c] $p \leqslant 0.001$; ns $p > 0.05$. β, β_1 and β_2 are the constants (±SE) from equation $\ln Q_{crop} = \beta + \beta_1 \text{pH} + \beta_2 \ln Q_{soil}$. n: observations.

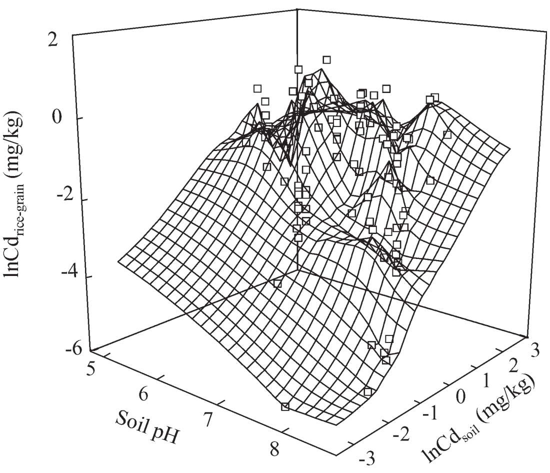

Fig. 4 Cd concentrations in rice grain as a function of soil Cd concentrations and soil pH. The graphic surfaces were obtained using correlation gridding method.

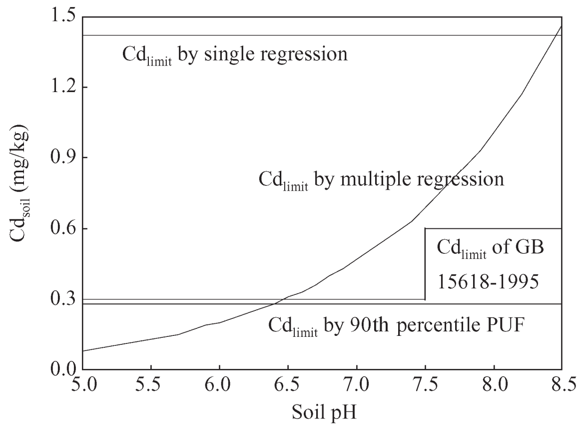

Fig. 5 Comparison of Cd values of National Standard (GB 15618-1995) with maximum permissible Cd values calculated by 95% upper prediction limits of multiple regression models with soil pH, limited Cd values calculated by the 90th percentile PUF and limited Cd values calculated by the 95% upper prediction limits of single-variable regression models.

The model results were over conservative when pH value was low (< 6.5) since the predicted values were even lower than the background value of Cd concentration in China's agricultural soils (0.187 mg/kg, Liu, 1996). The Cd values derived by the 90th percentile of PUF and the 95% upper prediction limit of single-variable regression as shown in Fig. 5 were over conservative and less conservative respectively compared with the Cd values derive from the multiple regression model.

3 Conclusions

Paired data of As, Cd, and Pb concentrations in soil and rice grain and leaf/stem vegetables were collected from studies which reported China's contaminated croplands to predict the transfer of trace elements from soil to edible parts of crops. After data selection, plant uptake factor, single-variable and multiple regression models were developed and compared. The best model was then used to establish the soil benchmark for Cd in paddy soils.

(1) The median PUF failed to give reasonable prediction of trace element uptake by crops. But quite conservative prediction results were presented by the 90th percentile of PUF which were commonly used for generic risk assessment and soil benchmark derivation. The PUF were probabilistic and fitted Gaussian distribution over the soil concentration range of this study.

(2) Both the single-variable regression and multiple regression models only explained less than 71% of the variance in As, Cd and Pb accumulations in crops. The standard errors of regression models were relatively large probably because different cultivars of rice were ignored and leaf/stem vegetables were polled together for the purpose of this study.

(3) Combination of soil pH improved the regression results for the prediction of Cd uptake. The 95% upper prediction limit of multiple regression model for Cd uptake by rice grain was recommended as the optimal approach to derive soil Cd benchmark for paddy soils.

(4) There are several pros and cons with varying the variables that ultimately influence the simulated results obtained. It is suggested to include all the possible variants in simulation in future and try to obtain the optimal results with more sufficient data.

Acknowledgments

This work was supported by the Key Project of the National Natural Science Foundation of China (No. 40432005), the Program of Knowledge Innovative Engineering of the Chinese Academy of Sciences (No. CXTD-Z2005-4), and the Chinese Ministry of Science and Technology (No. 2006DFA91940).

References

Adamsa M L, Zhao F J, McGratha S P, Nicholsonb F A, Chambersb B J, 2004. Predicting cadmium concentrations in wheat and barley grain using soil properties. *Journal of Environmental Quality*, 33: 532–541.

Alsop W R, Hawkins E T, Stelljes M E, Collins W, 1996. Comparison of modeled and measured tissue concentrations for ecological receptors. *Human and Ecological Risk Assessment*, 2(3): 539–557.

Bingham F T, 1979. Bioavailability of Cd to food crops in relation to heavy metal content of sludge-amended soil. *Environmental Health Perspectives*, 28: 39–43.

Bonten L T C, Jan E, Groenenberg J E, Weng L P, van Riemsdijk W H, 2008. Use of speciation and complexation models to estimate heavy metal sorption in soils. *Geoderma*, 146: 303–310.

CDFA (California Department of Food and Agriculture), 1998. Development of Risk based Concentrations for Arsenic, Cadmium, and Lead in Inorganic Fertilizers. The Report of the Heavy Metal Task Force, Agricultural Commodities and Regulatory Services, Department of Food and Agriculture, California.

Chen W P, Li L Q, Chang A C, Wu L S, Chaney R L, Smith R et al., 2009. Characterizing the solid-solution partitioning coefficient and plant uptake factor of As, Cd, and Pb in California croplands. *Agriculture, Ecosystems and Environment*, 129: 212–220.

Cui Y J, Zhu Y G, Zhai R H, Chen D Y, Huang Y Z, Qiu Y et al., 2004. Transfer of metals from soil to vegetables in an area near a smelter in Nanning, China. *Environment International*, 30: 785–791.

Efroymson R A, Smaple B E, Suter G W, 2001. Uptake of inorganic chemicals from soil by plant leaves: regressions of field data. *Environmental Toxicology and Chemistry*, 20(11): 2561–2571.

Hough R L, Breward N, Young S D, Crout N M J, Tye A M, Moir A M et al., 2004. Assessing potential risk of heavy metal exposure from consumption of home-produced vegetables by urban populations. *Environmental Health Perspectives*, 112(2): 215–221.

Hough R L, Young S D, Crout N M J, 2003. Modelling of Cd, Cu, Ni, Pb and Zn uptake, by winter wheat and forage maize, from a sewage disposal farm. *Soil Use and Management*, 19(1): 19–27.

Kabata-Pendias A, Pendias H, 2001. Trace Elements in Soils and Plants (2nd ed.). CRC Press, Boca Raton, Florida.

Kuo S, Huang B, Bembenek R, 2004. The availability to lettuce of zinc and cadmium in a zinc fertilizer. *Soil Science*, 169(5): 363–373.

Liu Z, 1996. Soil Trace Elements in China. Jiangsu Science and Technology Press, Nanjing.

McBride M, 2002. Cadmium uptake by crops estimated from soil total Cd and pH. *Soil Science*, 167(1): 62–67.

McLaughlin M J, Parker D R, Clarke J M, 1999. Metals and micronutrients–food safety issues. *Field Crops Research*, 60: 143–163.

Nahmani J, Hodson M E, Black S, 2007. A review of studies performed to assess metal uptake by earthworms. *Environmental Pollution*, 145(2): 402–424.

Provoost J, Cornelis C, Swartjes F, 2006. Comparison of soil clean-up standards for trace elements between countries: Why do they differ? *Journal of Soils and Sediments*, 6(3): 173–181.

Sample B E, Suter G W, Beauchamp J J, Efroymson R A, 1999. Literature-derived bioaccumulation models for earthworms: Development and validation. *Environmental Toxicology and Chemistry*, 18(9): 2110–2120.

SAS Institute, 1988. SAS/STAT User's Guide, Release 6.03 ed. Cary, NC, USA.

Sauve S, Hendershot W, Allen H E, 2000. Solid-solution partitioning of metals in contaminated soils: dependence on pH, total metal burden, and organic matter. *Environmental Science and Technology*, 34(7): 1125–1131.

Simmons R W, Noble A D, Pongsakul P, Sukreeyapongse O, Chinabut N, 2008. Analysis of field-moist Cd contaminated paddy soils during rice grain fill allows reliable prediction of grain Cd levels. *Plant and Soil*, 302: 125–137.

SEPA (State Environmental Protection Agency), 2003. Investigation report of eco-environmental situation in China's Mid-East Regions. *Environmental Protection*, 8: 3–8.

Suter G W, Efroymson R A, Sample B E, Jones D S, 2000. Ecological Risk Assessment for Contaminated Sites. CRC Press LLC, Florida.

US EPA, 1997. Exposure Factors Handbook (Final Report) U.S. EPA/600/P-95/002F a-c, Washington DC.

Wang G, Su M Y, Chen Y H, Lin F F, Luo D, Gao S F, 2006. Transfer characteristics of cadmium and lead from soil to the edible parts of six vegetable species in southeastern China. *Environmental Pollution*, 144: 127–135.

Wenzel W W, Blum W E H, Brandstetter A, Jockwe F, Kochl A, Oberforster M et al., 1996. Effects of soil properties and cultivar on cadmium accumulation in wheat grain. *Zeitschrift für Pflanzenernaehr Bodenkd.*, 159: 609–614.

Yu H, Wang J L, Fang W, Yuan J G, Yang Z Y, 2006. Cadmium accumulation in different rice cultivars and screening for pollution-safe cultivars of rice. *Science of the Total Environment*, 370: 302–309.

Zarcinas B A, Ishak C F, McLaughlin M J, Cozens G, 2004. Heavy metals in soils and crops in southeast Asia. 1. Peninsular Malaysia. *Environmental Geochemistry and Health*, 26: 343–357.