

Article



Predicting Bioaccumulation of Potentially Toxic Element in Soil–Rice Systems Using Multi-Source Data and Machine Learning Methods: A Case Study of an Industrial City in Southeast China

Modian Xie^{1,2}, Hongyi Li², Youwei Zhu^{3,*}, Jie Xue^{1,4}, Qihao You⁵, Bin Jin⁶ and Zhou Shi^{1,4}

- Key Laboratory of Environment Remediation and Ecological Health, Ministry of Education, College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310058, China; dd574884351@sina.com (M.X.); xj2019@zju.edu.cn (J.X.); shizhou@zju.edu.cn (Z.S.)
 Denattment of Land Resource Management Schement Schement and Ukban Management
 - Department of Land Resource Management, School of Tourism and Urban Management,
- Jiangxi University of Finance and Economics, Nanchang 330013, China; lihongyi@jxufe.edu.cn
- ³ Protection and Monitoring Station of Agricultural Environment, Bureau of Agriculture of Department of Rural and Agriculture of Zhejiang Province, Hangzhou 310020, China
- ⁴ Institute of Agricultural Remote Sensing and Information Technology Application, College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310058, China
- ⁵ Eco-Environmental Science & Research Institute of Zhejiang Province, Hangzhou 310012, China; 3130100204@zju.edu.cn
- ⁶ Ningbo Agricultural Food Safety Management Station, Ningbo 315000, China; bjin502500196@163.com
- * Correspondence: 13018941333@163.com; Tel./Fax: +86-130-1894-1333

Abstract: Potentially toxic element (PTE) pollution in farmland soils and crops is a serious cause of concern in China. To analyze the bioaccumulation characteristics of chromium (Cr), zinc (Zn), copper (Cu), and nickel (Ni) in soil-rice systems, 911 pairs of top soil (0-0.2 m) and rice samples were collected from an industrial city in Southeast China. Multiple linear regression (MLR), support vector machines (SVM), random forest (RF), and Cubist were employed to construct models to predict the bioaccumulation coefficient (BAC) of PTEs in soil-rice systems and determine the potential dominators for PTE transfer from soil to rice grains. Cr, Cu, Zn, and Ni contents in soil of the survey region were higher than corresponding background contents in China. The mean Ni content of rice grains exceeded the national permissible limit, whereas the concentrations of Cr, Cu, and Zn were lower than their thresholds. The BAC of PTEs kept the sequence of Zn (0.219) > Cu (0.093) > Ni (0.032) > Cr (0.018). Of the four algorithms employed to estimate the bioaccumulation of Cr, Cu, Zn, and Ni in soil-rice systems, RF exhibited the best performance, with coefficient of determination (R^2) ranging from 0.58 to 0.79 and root mean square error (RMSE) ranging from 0.03 to 0.04 mg kg⁻¹. Total PTE concentration in soil, cation exchange capacity (CEC), and annual average precipitation were identified as top 3 dominators influencing PTE transfer from soil to rice grains. This study confirmed the feasibility and advantages of machine learning methods especially RF for estimating PTE accumulation in soil-rice systems, when compared with traditional statistical methods, such as MLR. Our study provides new tools for analyzing the transfer of PTEs from soil to rice, and can help decision-makers in developing more efficient policies for regulating PTE pollution in soil and crops, and reducing the corresponding health risks.

Keywords: potentially toxic elements; soil-rice systems; bioaccumulation coefficient; machine learning methods; potential dominators

1. Introduction

Soil contamination arouse from potentially toxic element (PTE) such as chromium (Cr), lead (Pb), cadmium (Cd), mercury (Hg), arsenic (As), copper (Cu), zinc (Zn), nickel (Ni) in agricultural land has raised serious concerns worldwide [1–7], particularly in



Citation: Xie, M.; Li, H.; Zhu, Y.; Xue, J.; You, Q.; Jin, B.; Shi, Z. Predicting Bioaccumulation of Potentially Toxic Element in Soil–Rice Systems Using Multi-Source Data and Machine Learning Methods: A Case Study of an Industrial City in Southeast China. *Land* **2021**, *10*, 558. https://doi.org/ 10.3390/land10060558

Academic Editors: Evangelia Golia, Vassilios Diakoloukas and Nikolaos Tsiropoulos

Received: 14 April 2021 Accepted: 21 May 2021 Published: 26 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). China, which has experienced rapid industrialization and urbanization in the past four decades [8–13]. Apart from natural weathering from parent soil materials, anthropogenic activities (including industrial waste production, sewage irrigation, agricultural inputs, mining, and smelting) are a major source of PTE accumulation in farmland soils [14–16]. Several previous studies have reported PTE pollution in farmland soils from different areas of China, particularly in regions like the Yangtze River Delta [17–19], Pearl River Delta [20], Beijing-Tianjin-Hebei [21,22], and Northeast China [23,24].

PTE uptake by the soil–crop system is the dominant pathway of human exposure to PTEs [25–29]. Rice is one of the most important staple food crops worldwide. China is the world's largest producer of rice, and rice is the staple food for majority of the Chinese population, particularly in Southern China. Therefore, PTEs pollution in soil and their subsequent accumulation in these crops, and related health risks of human exposure have been extensively studied around the world [30–35].

The transfer and bioaccumulation of PTEs from soil to rice is a complicated process involving PTE release, dissolution, and bioaccumulation in different plant parts [36,37]. It is affected by various factors, including the total concentration of PTEs in soil, as well as soil physical and chemical properties [38–40]. Numerous studies have analyzed the shift of PTEs in soil-rice system [41–44]. Brus et al (2009) built a multiple linear regression (MLR) model to estimate the Cd concentration in rice grains using the total Cd concentration in soil, soil pH, soil clay, and soil organic matter (SOM) as predictors ($R^2 = 0.66$) [45]. Zhao et al. (2009) developed a linear regression (LR) model to predict the transfer of Cr, Zn, Cu, Pb, and Ni from soil to rice in Zhejiang Province of China. The total PTE content and soil pH were used as covariates, and the model R² ranged from 0.13 to 0.37 [46]. Chen et al. (2016) further used LR to predict the transfer of Cu, Pb, Zn, Cd, and Mercury (Hg) from soil to rice grains, using total PTE content, soil pH, and SOM as covariates, with a model R^2 of 0.24 to 0.63 [36]. Deng et al. (2020) predicted the transfer of Cr, Pb, Cd, Hg, and As from soil to rice grains using MLR, with the total soil Cd concentration and soil pH as covariates, and an \mathbb{R}^2 ranging from 0.109 to 0.456 [40]. Mu et al. (2020) also predicted the Cd concentration in rice grain using LR, with the total Cd concentration, soil pH, SOM as covariates, and obtained the R² ranging from 0.42 to 0.66 [47]. Liu et al. (2021) analyzed the transfer characteristic of Pb in different soil-crops system in Southwest of China and found that beans, peas and peanuts have a stronger ability of bioaccumulation of Pb [48]. Ma et al. (2021) found that TFe2O3, MnO have essential effects on the migration of Cd in soil-rice ecosystem in Southwest of China [49].

However, some limitations existed in current studies. Firstly, most of current researches concentrated on analyzing linear relationship between PTEs and others factors. For example, most studies conducted till date developed prediction models employing traditional statistical methods, such as LR and MLR, and it can only analyze the linear relationship between studied PTEs and others variables. Secondly, few quantitative covariates were included in the existing studies, which limits the value of the results since the migration of PTEs in the soil-crops system is affected by various factors. Thirdly, most of current studies can only analyze the effects of quantitative variables on transfer of PTEs in soil-crops system and not able to take qualitative variable into analysis. Finally, previous studies did not quantify the effects of different components on the transfer of PTEs in soilrice systems. Machine learning algorithms like support vector machine (SVM), random forest (RF), and Cubist, can predict non-linear relationships between different variables. In addition, RF and Cubist can manage quantitative and categorical variables simultaneously. Another advantage of machine learning methods is that they can provide information about the importance of variables in the model, which can aid the regulation of PTE pollution in soil-crop systems and reduce the health risks of PTE exposure [36,40,45-47].

Therefore, in the present study, we used three machine learning algorithms to build models to predict the bioaccumulation coefficient (BAC) of Cr, Cu, Zn, and Ni from soil to rice grains. A total of 20 covariates, including total soil PTE content, pH, CEC, and soil texture, were used as predictors to improve model performance. An MLR model

was constructed using the correlated quantitative covariates as predictors for comparison. The objectives of current research were to: (1) analyze the concentration and transfer characteristics of PTEs in soil–rice system; (2) build models to predict the transfer of PTEs from soil to rice grains using MLR, SVM, RF, and Cubist methods; (3) identify potential dominators of PTE bioaccumulation in soil–rice systems. The results are expected to aid the control of PTE contamination in soil and rice, as well as reduce the health risk of human consumption of such rice.

2. Materials and Methods

2.1. Study Area

The survey region was situated in Eastern China (120°55′–122°16′ E, 28°51′–30°33′ N) (Figure 1), occupying an area of around 9,800 square kilometers and with a population of 8.5 million. The survey city is one of the most developed cities in China, with the 12th highest gross domestic product (GDP) in 2019. The total industrial output value of Ningbo in 2019 was 578.3 billion Yuan. Three pillar industries in Ningbo are chemical industry, textile and clothing, and machinery industry. The survey region was situated in the lower reach of the Yangtze River in China, and has been an important area for rice cultivation, because of the favorable climate and terrain conditions. Rice plantations occupy more than 100,000 ha in this region and the rice yield was around 660,000 tons in 2018. Details of the study area have been provided Hu et al. [30].



Figure 1. Map of the study area and sampling sites.

2.2. Sampling and Chemical Analysis

Overall, 911 pairs of surface soil (0–0.2 m) and rice samples were collected in 2013. The rice samples were collected during the harvest season. The surface soil samples were collected after the harvest season at the same locations as rice samples. Each soil sample was a composite of 5 sub-samples collected from five locations within five meters using a stainless steel shovel. The sampling positions were documented using a GPS (Figure 1). Here, we presented the number of surface soil and rice sample collected from the study area in grid with size of 2 kilometers (Figure 1). The pH was determined by a pHS-3C digital pH meter (Shanghai REX Sensor Technology Co., Ltd., China). Soil samples were digested in acid (HCl–HNO₃–HClO₄), whereas the corresponding rice grain samples were digested

using the dry ashing method, and the concentration of Cr, Cu, Zn, and Ni were measured by inductively coupled plasma optical emission spectrometry (ICP-OES 6300, Thermo Fisher Scientific, Waltham, MA, USA). Quality control and standard quality assurance were conducted using soil standard reference materials (GBW07403 and GBW07404 obtained from the National Standard Detection Research Center, Beijing, China), duplicate samples, and procedure blanks during the measurements. Reagent blanks were also included in the analyses for quality control, and element recoveries were in the range of 90% to 110% [39].

2.3. Model Construction Algorithms

2.3.1. Multiple Linear Regression (MLR)

MLR [50] can predict the outcome of a response variable using multiple explanatory variables. MLR is used to construct model to describe linear relationships between the independent variables and the dependent variable. MLR has been described in detail by Myers (1990) [50].

2.3.2. Support Vector Machines (SVM)

SVMs aim to build a hyperplane or a set of hyperplanes in a high- or infinitedimensional space, which can then be used for classification and regression [51]. For regression analyses, SVMs can take the non-linearities of a natural system into account, including kernel functions, which function as constructing blocks for SVMs [52]. SVMs have previously been described by Cortes and Vapnik (1995) [51] and Bordoni et al. (2018) [53].

2.3.3. Random Forest (RF)

The RF is a non-linear, non-parametric algorithm developed by Breiman (2001) [54], which comprises a sequence of numerous individual predictor tree models trained from bootstrap samples of the data. The splits of each tree are decided based on a subset of predictor variables were chosen randomly from all available predictors [37]. The prediction results of all trees are then averaged to obtain the final prediction [54]. RF can also calculate the relative importance of a variable from the prediction error of out-of-bag (OOB) predictions [55]. RF perturbs each variable and estimates its importance as the change in the OOB error [54].

2.3.4. Cubist

Cubist is a rule-based data mining and prediction-oriented regression model [56,57]. It firstly creates a tree structure and then folds each path through the tree into a rule. After that, it fitted a regression model for each rule based on the subset of data described by the rules, and the prediction model is selected based on these rules. Prediction can be improved by generating several rule-based models, referred to as "committees", through boosting [58]. Subsequent models are built based on the corrections of the predictions made by previous models to minimize the prediction error.

2.4. Data Collection

We used 20 auxiliary variables, which have been proved to have effects on the transfer of PTEs from soil to the rice by previous studies to build a model to predict PTE bioaccumulation in soil–rice systems (Table 1) [36,40,45–49]. The auxiliary variables included soil properties (such as soil organic matter, pH, soil bulk density, content of PTE in soil, cation exchange capacity, soil sand content, soil clay content, soil silt content, soil coarse fraction), climatic factors (such as annual temperature, annual precipitation), terrain attributes (such as elevation), agricultural management practices (such as amount of phosphate fertilizer applied annually, amount of organic fertilizer applied annually, amount of nitrogen fertilizer applied annually, amount of potash fertilizer applied annually), geology (soil group, parent material), and land use (land use types), and population density are detailed in Table 1.

Auxiliary Variable	Abbreviation	Resolution	Type ^a	Source
Content of PTE in soil ^b	SC	_	Q	This study
Soil organic matter	SOM	_	С	This study
pH	pН	_	Q	This study
Soil group	SG	-	С	National Soil Survey Office ^c
Population density	PD	1 km	Q	REDC ^d
Land use types	LU	1 km	С	REDC ^d
Annual temperature	Tem	1 km	Q	REDC ^d
Annual precipitation	Preci	1 km	Q	REDC ^d
Elevation	DEM	_	Q	This study
Amount of phosphate fertiliser applied annually	PHF	_	Q	This study
Amount of organic fertiliser applied annually	OF	_	Q	This study
Amount of nitrogen fertiliser applied annually	NTF	-	Q	This study
Amount of potash fertiliser applied annually	POF	-	Q	This study
Soil bulk density	BD	250 m	Q	ISRIC SoilGrids ^e
Parent material	PM	-	С	National Soil Survey Office ^f
Cation exchange capacity	CEC	250 m	Q	ISRIC SoilGrids ^e
Soil sand content	Sand	250 m	Q	ISRIC SoilGrids ^e
Soil clay content	Clay	250 m	Q	ISRIC SoilGrids ^e
Soil silt content	Silt	250 m	Q	ISRIC SoilGrids ^e
Soil coarse fraction	Coarse	250 m	Q	ISRIC SoilGrids ^e

Table 1. Auxiliary variables used to predict PTE bioaccumulation in soil-rice systems.

^a Q: quantitative; C: categorical. ^b It means the content of individual target PTE in soil. ^c [59]. ^d REDC: Resource and Environmental Data Cloud Platform (http://www.resdc.cn/Default.aspx). ^e [60]. ^f [61].

2.5. Data Analysis

In this study the BAC for different PTEs was calculated to represent the migration of PTEs from soil to rice. The BACs were determined by:

$$BAC = \frac{C_{rice}}{C_{soil}} \tag{1}$$

where *C*_{rice} and *C*_{soil} represent the total PTE content in rice grain and soil, respectively.

Exploratory data analysis was performed using R (R Core Team, 2015). The MLR, RF, SVM, and Cubist algorithms in the *caret* package [62] for R were used to construct models for predicting BACs of PTEs in soil–rice systems. The flowchart of this study was presented in Figure 2.



Figure 2. Flowchart of the modeling of potentially toxic elements in soil-rice system.

3. Results

3.1. PTE Content in Soil and Rice Grains

The summary characteristics of PTEs in soil and rice grains are described in Table 2. Zn exhibited the highest mean content (115.50 mg kg⁻¹) in soil, followed by Cr (71.87 mg kg⁻¹), Cu (35.76 mg kg⁻¹), and Ni (29.94 mg kg⁻¹). The coefficient of variation (CV, %) indicates the variation degree of the PTE content [39] (Hu et al., 2019). The CVs of the soil PTE concentrations were in the order of: Ni (50.17%) > Cu (36.89%) > Cr (36.81%) > Zn (29.76%). All the soil samples have Cr and Cu content lower than risk screening value regulated by Chinese government (GB 15618–2018). Additionally, 1.32% of soil samples with Zn content higher than risk screening value (GB 15618–2018).

Table 2. Summary statistics for the PTEs in soil and rice grains (N = 911).

Element		Min (mg/kg)	Median (mg/kg)	Mean (mg/kg)	Max (mg/kg)	SD ^a	CV (%) ^b	Percentage above the National Standard
Cr	Soil	9.16	74.20	71.87	246.00	26.46	36.81	0
	Rice grain	0.01	0.52	0.79	13.00	0.92	117.08	20.75%
Cu	Soil	8.92	34.10	35.76	116.00	13.19	36.89	0
	Rice grain	0.15	3.00	2.98	6.90	0.80	26.92	0
Zn	Soil	34.30	110.00	115.50	714.00	34.39	29.76	1.32%
	Rice grain	1.30	24.00	23.89	52.00	4.31	18.04	0.11%
Ni	Soil	3.81	30.40	29.94	293.00	50.17	50.17	0.88%
	Rice grain	0.01	0.50	0.64	5.40	0.49	75.79	65.75%

Note: ^a denotes standard deviation; ^b denotes coefficient of variation.

The mean PTE contents in rice grains kept the sequence: Zn (23.89 mg kg⁻¹) > Cu (2.98 mg kg⁻¹) > Cr (0.79 mg kg⁻¹) > Ni (0.64 mg kg⁻¹). As outlined by the Chinese Government, the national standard values for Cr [63], Cu [64], Zn [65], and Ni [66] in rice grains are 1.00, 10.00, 50.00, and 0.40 mg kg⁻¹, respectively. The results revealed that the mean Ni concentration in rice grains exceeded the national threshold, whereas Cr, Cu, and Zn contents were below their national limits. The CVs of PTEs in the rice grain were in the order of: Cr (117.08%) > Ni (75.79%) > Cu (26.92%) > Zn (18.04) (Table 2). The percentage of rice samples with a concentration of Cr, Cu, Zn, and Ni higher than corresponding national standard value is 20.75%, 0, 0.11%, and 65.75%, respectively.

3.2. BAC of PTEs from Soil to Rice

The BACs of the different PTEs in rice grains exhibited remarkable variations (Figure 3). The BACs of Cr, Cu, Zn, Ni ranged from 0.001 to 0.890, 0.002 to 0.485, 0.013 to 0.729, and 0.003 to 0.783, respectively. The average BACs of the different PTEs followed the sequence: Zn (0.219) > Cu (0.093) > Ni (0.032) > Cr (0.018) (Figure 3). The CVs of the BACs of Cr, Cu, Zn, and Ni in rice grains were 270.34%, 45.90%, 29.87%, and 168.68%, respectively, demonstrating that the BACs of Cr and Ni varied greatly in the study area.

3.3. Modeling the Transfer of PTEs from Soil to Rice

We employed four different methods (Cubist, RF, SVMs, and MLR) to build models for predicting PTE bioaccumulation from soil to rice grains with input from multiple ancillary sources. The sample dataset was firstly randomly divided as training dataset and validation dataset with ratio of 4:1. Firstly the models were trained using a training dataset and 5-fold cross-validation was employed to evaluate the model performance on training dataset. Then the constructed models were validated independently using a validation dataset. Performance metrics of the constructed models are listed in Table 3. The R^2 of the Cubist models for the different PTEs varied between 0.05 and 0.72, whereas the RMSE varied between 0.04 and 3.22 mg kg⁻¹. The R^2 of RF models varied between 0.58 and 0.79, whereas the RMSE varied between 0.03 and 0.04 mg kg⁻¹(Figure 4). For the SVM models, the R² spanned 0.05 and 0.69, and the RMSE ranged from 0.05 to 2.95 mg kg⁻¹. For the MLR models, the R² varied between 0.46 and 0.67, and the RMSE ranged from 0.05 to 2.95 mg kg⁻¹. The RF algorithms markedly outperformed the other three methods, with the highest R² and Lin's concordance correlation coefficient (CCC) and the lowest RMSE and bias for all the tested PTEs (Table 3).



Figure 3. Summary statistics of the BACs of PTEs from soil to rice grains (N = 911).

	Table 3. (Comparison o	of model performance	validated using a	a validation dataset ($N = 182$).	
--	------------	--------------	----------------------	-------------------	-------------------------------------	--

Model	Index	Cr	Cu	Zn	Ni
Cubist	R ²	0.72	0.05	0.09	0.20
	CCC ^a	0.83	0.21	0.29	0.35
	RMSE (mg kg ^{-1}) ^b	0.04	0.69	3.22	0.36
	Bias (mg kg ^{-1})	-5.31E-03	6.33E-02	-2.23E-02	-7.07E - 02
RF	\mathbb{R}^2	0.79	0.58	0.66	0.74
	CCC ^a	0.86	0.69	0.77	0.85
	RMSE (mg kg ^{-1}) ^b	0.03	0.04	0.04	0.04
	Bias (mg kg ^{-1})	-4.41E-03	1.91E - 03	-3.58E - 03	7.93E-07
SVM	R ²	0.69	0.05	0.13	0.21
	CCC ^a	0.62	0.15	0.24	0.35
	RMSE (mg kg $^{-1}$) ^b	0.05	0.66	2.95	0.36
	Bias (mg kg ^{-1})	-1.10E - 02	5.45E - 02	-1.81E - 01	-8.26E - 02
MLR	\mathbb{R}^2	0.67	0.49	0.51	0.46
	CCC ^a	0.72	0.66	0.67	0.63
	RMSE (mg kg $^{-1}$) ^b	0.04	0.05	0.06	0.06
	Bias (mg kg ^{-1})	-5.67E - 03	-1.11E - 15	-3.11E - 15	-3.55E - 15

Note: ^a indicates Lin's concordance correlation coefficient; ^b indicates root mean square error.



Figure 4. Comparison between the measured and RF-predicted BACs of Cr (**a**), Cu (**b**), Zn (**c**) and Ni (**d**) in validation dataset (N = 182).

3.4. Variable Importance for Modeling PTE Bioaccumulation in Soil-Rice Systems

We used 20 auxiliary variables (Table 1), including terrain attributes, climatic factors, soil properties, geological factors, and geomorphic factors, to estimate PTE bioaccumulation from soil to rice grains. As described in Section 3.3, the RF model yielded the best estimation of PTE bioaccumulation from soil to rice grains. Therefore, the potential dominators for PTE bioaccumulation were based on the relative variable importance determined by RF.

The relative importance of variables for estimation of PTE bioaccumulation in soil–rice systems is presented in Figure 5. The total concentration of Cr in soil, CEC, and the amount of organic fertilizer applied annually were identified as the three most important variables for predicting the BAC of Cr. Soil Cu content, annual average precipitation, and the amount of potash applied annually were the three principle variables for modeling the BAC of Cu. Soil Zn content and annual average precipitation and temperature were the three primary variables for modeling the BAC of Zn. Soil Ni content, CEC, and the amount of nitrogen fertilizer applied annually were the three most critical variables for modeling the BAC of Ni.



Figure 5. Relative importance of variables for modeling BAC of Cr (**a**), Cu (**b**), Zn (**c**) and Ni (**d**) (Abbreviations: SC, PTE content in soil; SOM, soil organic materials; CEC, Cation exchange capacity; Preci, annual average precipitation; Tem, annual average temperature; PHF, amount of phosphate fertilizer applied annually; OF, amount of organic fertilizer applied annually; NTF, amount of nitrogen fertilizer applied annually; POF, amount of potash fertilizer applied annually; BD, soil bulk density; DEM, elevation; PD, population density; SG, soil group; LU, land use; PM, parent material).

4. Discussion

4.1. PTE Content in Soil-Rice Systems

In this study, the mean contents of all PTEs in the region under survey were clearly higher than their national background concentrations in China, indicating PTE accumulation resulting from anthropogenic activities (Table 4). In comparison with other surveys conducted in regions like Liaoning, Jilin, Jiangsu, Fujian, Xinjiang, Shanghai, Guangdong and Hunan Province in China [42,67–70] (listed in Table 4), the soil Cr content was relatively high, whereas soil Cu, Zn, and Ni were present in moderate levels. The mean PTE contents in rice grains from the survey region were notably lower than the national permissible values, except Ni. The concentration of Cr in rice grains observed in this study was the highest among those reported in the studies listed in Table 4. The Zn content of rice grains was also higher than that noted in other surveys, whereas the Cu content of rice grains was lower. The maximum values of Cr, Zn, and Ni contents exceeded the national permissible limits, indicating that some of the rice crops are already affected by PTE pollution (Table 2). Especially, 20.75% of rice samples have Cr content higher than national threshold, moreover 65.75% of rice samples have Ni content higher than national threshold. This indicates that it is urgent task to take measures to control the bioaccumulation of PTEs in rice.

Location	Cr (mg kg ⁻¹) Soil/Rice/BAC	Cu (mg kg ⁻¹) Soil/Rice/BAC	Zn (mg kg ⁻¹) Soil/Rice/BAC	Ni (mg kg ⁻¹) Soil/Rice/BAC	Source
Zhejiang	71.37/0.79/0.018	35.76/2.98/0.093	115.50/23.89/0.219	29.94/0.64/0.032	This study
Zhuhai, Guangdong	-	49.34/3.98/0.081	120.2/21.51/0.179	-	[71]
Qingyuan, Guangdong	-	96.9/5.23/0.054	104/25.1/0.241	8.07/0.83/0.103	[72]
Shengyang, Liaoning	-	-	109.5/18.4/0.168	-	[23]
Wenzhou, Zhejiang	74.8/0.61/0.008	52.6/3.51/0.067	144.0/26.8/0.186	35.0/0.41/0.012	[73]
Hanzhong, Hubei	-	32.9/0.40/0.012	217/22.5/0.104	-	[74]
Jiangsu, Zhejiang, Shanghai	64.3/0.19/0.003	30.47/11.77/0.386	102.21/22.79/0.223	-	[69]
Huzhou, Zhejiang	-	31.06/2.49/0.080	106.82/14.28/0.134	32.14/0.12/0.004	[72]
Shaoxing, Zhejiang	-	28.64/2.98/0.104	98.74/22.41/0.227	27.03/0.35/0.013	[72]
Wenzhou, Zhejiang	-	41.13/3.09/0.075	98.74/20.69/0.210	27.03/0.22/0.008	[72]
Guizhou	-	-	135/11.56/0.086	40.5/1.57/0.039	[75]
Shantou, Guangdong	60.2/0.21/0.003	78.4/3.01/0.038	111.9/17.32/0.155	37.8/1.37/0.036	[68]
Changsha, Hunan	53.6/0.44/0.008	23.9/3.69/0.154	82.7/17.7/0.214	23.3/0.34/0.015	[43]
Jiangsu, Zhejiang, Shanghai	-	38.7/5.02/0.130	105/22.09/0.210	-	[36]
Shaoguan, Guangdong	29.1/0.34/0.012	67.2/3.63/0.054	129/29.1/0.226	15.1/0.83/0.055	[67]
China	54.6 ^a /1.0 ^b	23.5 ^c /10 ^d	82.1 ^e /50 ^f	28 ^g /0.4 ^h	[63–66,76]

Table 4. Comparison between mean PTE contents in soil-rice systems obtained in the presented research and those reported in previous studies from China.

Note: ^a national soil background value of Cr content in China; ^b national regulation value of Cr content in rice grain in China; ^c national soil background value of Cu content in China; ^d national regulation value of Cu content in rice grain in China; ^e national soil background value of Zn content in China; ^f national regulation value of Zn content in China; ^g national soil background value of Ni content in China; ^h national regulation value of Ni content in rice grain in China; ^g national regulation value of Ni content in rice grain in China; ^h national regulation value of Ni content in rice grain in China; ^h national regulation value of Ni content in rice grain in China.

Taking the results of other studies into consideration, the BACs of Cr, Cu, Zn, and Ni ranged from 0.003 to 0.012, 0.038 to 0.386, 0.086 to 0.241, and 0.004 to 0.103, respectively (Table 4). Our results found that migration capacity of Zn is the strongest while Cr is the weakest in soil–rice system which is consistent with the study reported by Li et al. (2021). The BAC of Cr calculated in this research was higher than that reported by previous study listed in Table 4, whereas the BACs of Cu, Zn and Ni were mid-range. The BAC of Zn was evidently higher than that of the other PTEs, which was consistent with previous study reported by Wang et al. (2016) [67], Chen et al. (2018) [26], Du et al. (2018) [42], and He et al. (2019) [39]. Cakmark (2008) [77] noted that application of Zn fertilizers enhances the amount of available Zn in the soil solution, leading to bioaccumulation in grains during the productive growth stage. Nan et al. (2002) [78], Mohammad and Moheman (2010) [79]; demonstrated the interactions occurring during the accumulation of Cd and Zn and the synergistic effects of the bioaccumulation of these two PTEs in soil–crop systems. These factors may contribute to the higher BAC of Zn observed in soil-rice systems.

4.2. Model Performance Employing the Different Methods and Elements

Clear differences were observed in the model performances. Generally, RF had the best performance, followed by MLR, SVM and Cubist. MLR aims to model the linear relationship between the explanatory variables and response variable. MLR has previously been used to predict soil organic carbon (SOC) [80], soil pH [81], soil bulk density [82], and soil texture [50]. However, MLR is not suitable if non-linear relationships are present in the data [83]. This may explain the poor performance of MLR when compared with that of RF and Cubist. Furthermore, the relationship between the covariates and BAC of PTEs in soil–rice systems is highly complex and difficult to model using linear relations alone, which further lower the performance of MLR. SVMs have also previously been employed to predict soil SOC stock [84], phosphorus [85], and clay [85]. They were reported to be more suitable for classification than quantitative prediction [86]. Cubist has proven to be a feasible method for predicting soil properties, such as SOC [87], salinity [88], silt [89], and concentration of soil Cr, Pb, Cu, Zn, Ni, and As [90]. The RF algorithm is based on decision trees and is currently the most widely used machine learning method for predicting soil properties. It has several advantages, such as ease of application, insensitivity to data

size, and the ability to model non-linear relationships in the data set [37,60,91], which may contribute to the better performance of RF.

The model constructed using RF yielded the best estimation of Cr bioaccumulation in soil–rice systems. The R^2 of this model was 0.79 (Table 3 and Figure 4), and it significantly outperformed the linear models constructed by Zhao et al. (2009) [46] ($R^2 = 0.13$), Zeng et al. (2011) [92] ($R^2 = 0.22$), and Deng et al. (2020) [40] ($R^2 = 0.46$) (Table 5). Our models for predicting Cu, Zn, and Ni were also more accurate than those reported previously in similar case studies (Table 5). The results confirmed the ability of RF to efficiently predict PTE bioaccumulation in soil–rice systems with multi-source covariates.

Table 5. Comparison of model performance for the estimation of PTE bioaccumulation in soil-rice systems obtained in the presented study and in previous studies.

Element	Method	R ²	Study Area	Covariates	Source
Cr	LR	0.456	Shaoxing, China	pH, SC	[46]
Cu, Zn, Ni	LR	0.52, 0.52, 0.55	Zhejiang China	pH, SOM, EC, sand, silt, clay	[27]
Cu, Zn	MLR	0.24, 0.63	YRD, China	SC, pH, SOM	[36]
Cr, Cu, Zn, Ni	MLR	0.13, 0.15, 0.37, 0.20	Zhejiang China	SC, pH	[46]
Cr, Cu, Zn	SR	0.22, 0.06, 0.37	Zhejiang China	SC, pH	[92]
Cr, Cu, Zn, Ni	RF	0.79, 0.58, 0.66, 0.74	-	Table 1	This study

Notes: LR, Linear Regression model; SR: Stepwise regression model; MLR, Multiple Linear Regression; YRD, Yangtze River Delta; SC, soil PTE content; EC, electrical conductivity.

4.3. Potential Dominators for PTE Bioaccumulation in Soil-Rice Systems

The total PTE content of soil was found to be the principal variable for estimating the BAC of PTEs in soil–rice systems, as evident from Figure 4. In addition, CEC and annual average precipitation contributed significantly to BAC estimation. Cr, Cu, Zn, and Ni are indispensable elements for rice growth. The PTEs can be absorbed by rice plants from the soil. Several studies have demonstrated the strong relationship between PTEs in the soil and in crops, including rice [39,41,47,49,93,94], which may explain the importance of SC in modeling PTE bioaccumulation in rice.

The close relationship between CEC and absorption of PTEs by crops from the soil has been reported by several researchers [37,95–97]. Vega et al. (2010) [98] revealed that CEC plays important roles in the sorption and retention processes of PTEs in soil–crop systems. Gupta et al. (2008) [99] reported that high CEC could significantly slow down PTE uptake by crops from the soil. Shen et al. (1998) [100] also found that high CEC allows the minerals in soil to absorb PTE ions from soil solutions through ion-exchange processes. Gu et al. (2011) [101] noted that an increase in soil CEC could promote PTE precipitation and complexation in the soil solution.

The annual average precipitation may affect PTE bioaccumulation in soil–rice systems by promoting crop growth, subsequently enhancing PTE absorption by the crop. Other covariates, such as SOM, appeared to be less important than expected, and these warrant further investigation.

4.4. Policy Recommendations

Our results indicate that some measures could be taken to curb the bioaccumulation of PTEs in rice. Firstly, great effort still necessary to reduce accumulation of PTEs in soil which could then reduce the bioaccumulation of PTEs from soil to rice. Secondly, more reasonable agricultural land management measures for example reasonable application of fertilization, reasonable crop rotation were expected to curb transfer of PTEs from soil to rice. After that, liming of acidic soils is recommended especially in the areas serous polluted by PTEs such as Cd, Pb since some PTEs such as Cd and Pb are more easily absorbed by crops in acidic soil [102]. Finally, breeding crops, such as rice cultivars with low accumulation, is another efficient way to reduce the human health risk caused by PTEs pollution in food or vegetables.

4.5. Limitations and Perspectives

Although the model developed using RF could efficiently predict PTE bioaccumulation in soil–crop systems, some limitations remain to be addressed. The present study did not take the rice variety into consideration while building the model. Some studies have reported significant differences in the BAC of PTEs among different rice varieties [47,103,104]. Furthermore, some of our ancillaries, such as CEC, soil bulk density, annual average temperature, annual average precipitation, and soil texture, were sourced from maps available online. The resolution of these data is 1000 m, which is still coarse to provide accurate information on the spatial variation of these variables (Table 1). This may negatively affect model performance. In addition, the agricultural practices, such as application of lime, duration of flooding, can affect PTE bioaccumulation in soil–rice systems [105,106].

Several measures could be employed to improve the estimation accuracy. Firstly, the differences between the rice varieties must be taken into consideration in future works. Secondly, maps with high resolution data of CEC, soil bulk density, annual average temperature, annual average precipitation, and soil texture, recorded during additional surveys should be used as covariates. Greater information related to soil management practices should also be included in the model to improve prediction accuracy. Finally, in our current study, the model was built using the total PTE content as one of the core covariates. However, the available form of PTEs in soil was reported to be more closely related to the PTEs absorbed by crops from the soil [39,107–110]. Therefore, future studies should employ the available form of PTEs in soil for constructing the model, instead of the total PTE content.

5. Conclusions

Results obtained in current study revealed that the mean contents of Cr, Cu, Zn, and Ni in farmland soils in the region under survey were higher than their background contents in China. The average Ni content in rice grains clearly exceeded the national permissible limit, whereas the Cr, Cu, and Zn contents were lower than their thresholds. The mean value of BAC of Zn (0.219) was the highest, followed by that of Cu (0.093), Ni (0.032), and Cr (0.018).

The model developed using RF significantly outperformed the MLR, SVM, and Cubist models, and could efficiently predict the bioaccumulation of Cr, Cu, Zn, and Ni in soil–rice systems, with an R^2 varying between 0.58 and 0.79 and RMSE varying between 0.03 and 0.04 mg kg⁻¹. Total PTE content in soil, CEC, and annual average precipitation were the principal components for the estimation of BACs of PTEs in soil–rice systems. This study confirmed the feasibility of RF for predicting PTE bioaccumulation in soil–rice systems and identifying potential dominators for the transportation of PTEs in soil–rice systems. The findings of this study are expected to enhance our knowledge of the PTEs accumulation in soil and rice grains, contributing to food safety and reducing the human health risk of consuming PTE-polluted rice.

Author Contributions: Conceptualization, Z.S., H.L. and Y.Z.; methodology, M.X. and H.L.; software, M.X. and J.X.; validation, M.X., Q.Y. and J.X.; formal analysis, M.X. and H.L.; investigation, Y.Z. and B.J.; resources, Z.S. and Y.Z.; data curation, M.X.; writing—original draft preparation, M.X. and H.L.; writing—review and editing, M.X. and H.L.; visualization, M.X.; supervision, Z.S. and H.L.; project administration, Y.Z. and Z.S.; funding acquisition, Y.Z. and Z.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Key Research and Development Program of Zhejiang Province (2015C02011), the National Natural Science Foundation of China (41771244, 42071068), and Key Program of the Natural Science Foundation of Zhejiang Province (LZ21D010002).

Acknowledgments: This work was supported by Key Research and Development Program of Zhejiang Province (2015C02011), the National Natural Science Foundation of China (41771244, 42071068), and Key Program of the Natural Science Foundation of Zhejiang Province (LZ21D010002).

Conflicts of Interest: The authors have declared that no competing interests exist.

References

- 1. Hu, B.F.; Chen, S.C.; Hu, J.; Xia, F.; Xu, J.F.; Li, Y.; Shi, Z. Application of portable XRF and VNIR sensors for rapid assessment of soil heavy metal pollution. *PLoS ONE* **2017**, *12*, e0172438. [CrossRef]
- Lequy, E.; Saby, N.P.A.; Ilyin, I.; Bourin, A.; Sauvage, S.; Leblond, S. Spatial analysis of trace elements in a moss bio-monitoring data over France by accounting for source, protocol and environmental parameters. *Sci. Total. Environ.* 2017, 590–591, 602–610. [CrossRef]
- 3. Marchant, B.P.; Saby, N.P.A.; Arrouays, D. A survey of topsoil arsenic and mercury concentrations across France. *Chemosphere* **2017**, *181*, 635–644. [CrossRef]
- Ono, K.; Yasutaka, T.; Hayashi, T.I.; Kamo, M.; Iwasaki, Y.; Nakamori, T.; Fujii, Y.; Kamitani, T. Model construction for estimating potential vulnerability of Japanese soils to cadmium pollution based on intact soil properties. *PLoS ONE* 2019, 14, e0218377. [CrossRef] [PubMed]
- Xia, F.; Hu, B.F.; Shao, S.; Xu, D.Y.; Zhou, Y.; Zhou, Y.; Huang, M.X.; Li, Y.; Chen, S.C.; Shi, Z. Improvement of Spatial Modeling of Cr, Pb, Cd, As and Ni in Soil Based on Portable X-ray Fluorescence (PXRF) and Geostatistics: A Case Study in East China. *Int. J. Environ. Res. Public Health* 2019, 16, 2694. [CrossRef]
- 6. Fu, T.T.; Zhao, R.Y.; Hu, B.F.; Jia, X.L.; Wang, Z.G.; Zhou, L.Q.; Huang, M.X.; Li, Y.; Shi, Z. Novel framework for modelling the cadmium balance and accumulation in farmland soil in Zhejiang Province, East China: Sensitivity analysis, parameter optimisation, and forecast for 2050. *J. Clean. Prod.* **2021**, *279*, 123674. [CrossRef]
- 7. Infante, E.F.; Dulfo, C.P.; Dicen, G.P.; Hseu, Z.Y.; Navarrete, I.A. Bioaccumulation and human health risk assessment of chromium and nickel in paddy rice grown in serpentine soils. *Environ. Sci. Pollut. Res.* **2021**, *28*, 17146–17157. [CrossRef] [PubMed]
- Hu, B.F.; Zhou, Y.; Jiang, Y.F.; Ji, W.J.; Fu, Z.Y.; Shao, S.; Li, S.; Huang, M.X.; Zhou, L.Q.; Shi, Z. Spatio-temporal variation and source changes of potentially toxic elements in soil on a typical plain of the Yangtze River Delta, China (2002–2012). *J. Environ. Manag.* 2020, 271, 110943. [CrossRef] [PubMed]
- 9. Shao, S.; Hu, B.F.; Fu, Z.Y.; Wang, J.Y.; Lou, G.; Zhou, Y.; Jin, B.; Li, Y.; Shi, Z. Source Identification and Apportionment of Trace Elements in Soils in the Yangtze River Delta, China. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1240. [CrossRef]
- Jia, X.L.; Hu, B.F.; Marchant, B.P.; Zhou, L.Q.; Shi, Z.; Zhu, Y.W. A methodological framework for identifying potential sources of soil heavy metal pollution based on machine learning: A case study in the Yangtze Delta, China. *Environ. Pollut.* 2019, 250, 601–609. [CrossRef] [PubMed]
- 11. Xia, F.; Hu, B.F.; Zhu, Y.W.; Ji, W.J.; Chen, S.C.; Xu, D.Y.; Shi, Z. Improved Mapping of Potentially Toxic Elements in Soil via Integration of Multiple Data Sources and Various Geostatistical Methods. *Remote Sens* **2020**, *12*, 3775. [CrossRef]
- 12. Yang, S.; Zhao, J.; Chang, S.X.; Collins, C.; Xu, J.; Liu, X. Status assessment and probabilistic health risk modeling of metals accumulation in agriculture soils across China: A synthesis. *Environ. Int.* **2019**, *128*, 165–174. [CrossRef]
- 13. Tang, M.; Lu, G.; Fan, B.; Xiang, W.; Bao, Z. Bioaccumulation and risk assessment of heavy metals in soil-crop systems in Liujiang karst area, Southwestern China. *Environ. Sci. Pollut. Res.* **2021**, *28*, 9657–9669. [CrossRef]
- 14. Fei, X.F.; Lou, Z.; Xiao, R.; Ren, Z.; Lv, X.N. Contamination assessment and source apportionment of heavy metals in agricultural soil through the synthesis of PMF and GeogDetector models. *Sci. Total Environ.* **2020**, 747, 141293. [CrossRef]
- 15. Guo, G.; Zhang, D. Source apportionment and source-specific health risk assessment of heavy metals in size-fractionated road dust from a typical mining and smelting area, Gejiu, China. *Environ. Sci. Pollut. Res.* **2021**, *28*, 9313–9326. [CrossRef]
- 16. Liu, J.; Li, Y.; Zhang, M.; Zhang, N.M.; Han, D.J. Health risk assessment and benchmark of lead pollution in agricultural soils in East Yunnan, China. *Trans. Chin. Soc. Agric. Eng.* **2021**, *37*, 241–250, (In Chinese with English abstract).
- Hu, B.F.; Shao, S.; Ni, H.; Fu, Z.Y.; Hu, L.S.; Zhou, Y.; Min, X.X.; She, S.F.; Chen, S.C.; Huang, M.X.; et al. Current status, spatial features, health risks, and potential driving factors of soil heavy metal pollution in China at province level. *Environ. Pollut.* 2020, 266, 114961. [CrossRef] [PubMed]
- 18. Jia, X.L.; Fu, T.T.; Hu, B.F.; Zhou, L.Q.; Shi, Z.; Zhu, Y.W. Identification of the potential risk areas for soil heavy metal pollution based on the source-sink theory. *J. Hazard Mater.* **2020**, *393*, 122424. [CrossRef] [PubMed]
- 19. Yang, S.H.; Qu, Y.J.; Ma, J.; Liu, L.L.; Wu, H.W.; Liu, Q.Y.; Gong, Y.W.; Chen, Y.X.; Wu, Y.H. Comparison of the concentrations, sources, and distributions of heavy metal (loid) s in agricultural soils of two provinces in the Yangtze River Delta, China. *Environ. Pollut.* **2020**, *264*, 114688. [CrossRef]
- 20. Zhang, J.R.; Li, H.Z.; Zhou, Y.Z.; Dou, L.; Cai, L.M.; Mo, L.P.; You, J. Bioavailability and soil-to-crop transfer of heavy metals in farmland soils: A case study in the Pearl River Delta, South China. *Environ. Pollut.* **2018**, 235, 710–719. [CrossRef]
- 21. Men, C.; Liu, R.; Xu, F.; Wang, Q.; Guo, L.; Shen, Z. Pollution characteristics, risk assessment, and source apportionment of heavy metals in road dust in Beijing, China. *Sci. Total Environ.* **2018**, *612*, 138–147. [CrossRef]
- 22. Yang, L.; Yang, M.; Wang, L.; Peng, F.; Li, Y.; Bai, H. Heavy metal contamination and ecological risk of farmland soils adjoining steel plants in Tangshan, Hebei, China. *Environ. Sci. Pollut. Res.* **2018**, 25, 1231–1242. [CrossRef]
- 23. Lian, M.H.; Wang, J.; Sun, L.N.; Xu, Z.; Tang, J.X.; Yan, J.; Zeng, X.F. Profiles and potential health risks of heavy metals in soil and crops from the watershed of Xi River in Northeast China. *Ecotoxicol. Environ. Saf.* **2019**, *169*, 442–448. [CrossRef]
- 24. Xiao, Q.; Zong, Y.; Malik, Z.; Lu, S. Source identification and risk assessment of heavy metals in road dust of steel industrial city (Anshan), Liaoning, Northeast China. *Hum. Ecol. Risk Assess. Int. J.* **2020**, *26*, 1359–1378. [CrossRef]
- 25. Liu, X.M.; Song, Q.J.; Tang, Y.; Li, W.L.; Xu, J.M.; Wu, J.J.; Wang, F.; Brooks, P.C. Human health risk assessment of heavy metals in soil–vegetable system: A multi-medium analysis. *Sci. Total Environ.* **2013**, *463–464*, 530–540. [CrossRef]

- 26. Chen, H.P.; Yang, X.; Wang, P.; Wang, Z.X.; Li, M.; Zhang, F.J. Dietary cadmium intake from rice and vegetables and potential health risk: A case study in Xiangtan, southern China. *Sci. Total Environ.* **2018**, *639*, 271–277. [CrossRef] [PubMed]
- Yu, Y.J.; Zhu, X.H.; Li, L.Z.; Lin, B.G.; Xiang, M.D.; Zhang, X.H.; Chen, X.C.; Yu, Z.L.; Wang, Z.D.; Wan, Y. Health implication of heavy metals exposure via multiple pathways for residents living near a former e-waste recycling area in China: A comparative study. *Ecotoxicol. Environ. Saf.* 2019, 169, 178–184. [CrossRef]
- 28. Hu, B.F.; Shao, S.; Fu, T.T.; Fu, Z.Y.; Zhou, Y.; Li, Y.; Qi, L.; Chen, S.C.; Shi, Z. Composite assessment of human health risk from potentially toxic elements through multiple exposure routes: A case study in farmland in an important industrial city in East China. *J. Geochem. Explor.* **2020**, *210*, 106443. [CrossRef]
- 29. Liu, X.; Yu, T.; Yang, Z.F.; Hou, Q.Y.; Yang, Q.; Li, C.; Ji, W.B.; Li, B.; Duan, Y.R.; Zhang, Q.Z.; et al. Transfer mechanism and bioaccumulation risk of potentially toxic elements in soil–rice systems comparing different soil parent materials. *Ecotoxicol. Environ. Saf.* **2021**, *216*, 112214. [CrossRef]
- 30. Hu, B.F.; Jia, X.L.; Hu, J.; Xu, D.Y.; Xia, F.; Li, Y. Assessment of heavy metal pollution and health risks in the soil-plant-human system in the Yangtze river delta, China. *Int. J. Environ. Res. Public Health* **2017**, *14*, 1042. [CrossRef] [PubMed]
- 31. Lu, A.; Li, B.; Li, J.; Chen, W. Heavy metals in paddy soil-rice systems of industrial and township areas from subtropical China: Levels, transfer and health risks. *J. Geochem. Explor.* **2018**, *194*, 210–217. [CrossRef]
- 32. Aslam, M.W.; Ali, W.; Meng, B.; Abrar, M.M.; Lu, B.; Qin, C.Y.; Zhao, L.; Feng, X.B. Mercury contamination status of rice cropping system in Pakistan and associated health risks. *Environ. Pollut.* 2020, 263, 114625. [CrossRef]
- Bocardi, J.M.B.; Pletsch, A.L.; Quinaia, S.P. Quality reference values for heavy metals in soils developed from basic rocks under tropical conditions. J. Geochem. Explor. 2020, 217, 106591. [CrossRef]
- 34. Zhao, L.; Meng, B.; Feng, X.B. Mercury methylation in rice paddy and accumulation in rice plant: A review. *Ecotoxicol. Environ. Saf.* **2020**, *195*, 110462. [CrossRef] [PubMed]
- 35. Shao, S.; Hu, B.F.; Tao, Y.; You, Q.H.; Huang, M.X.; Zhou, L.Q.; Chen, Q.X.; Shi, Z. Comprehensive source identification and apportionment analysis of five heavy metals in soils in Wenzhou City, China. *Environ. Geochem. Health* **2021**, 1–24. [CrossRef]
- Chen, H.Y.; Yuan, X.Y.; Li, T.Y.; Sun, H.; Ji, J.F.; Wang, C. Characteristics of heavy metal transfer and their influencing factors in different soil–crop systems of the industrialization region, China. *Ecotoxicol. Environ. Saf.* 2016, 126, 193–201. [CrossRef] [PubMed]
- Hu, B.F.; Xue, J.; Zhou, Y.; Shao, S.; Fu, Z.Y.; Li, Y.; Chen, S.C.; Qi, L.; Shi, Z. Modelling bioaccumulation of heavy metals in soil-crop ecosystems and identifying its controlling factors using machine learning. *Environ. Pollut.* 2020, 262, 114308. [CrossRef] [PubMed]
- Römkens, P.F.A.M.; Guo, H.Y.; Chu, C.L.; Liu, T.S.; Chiang, C.F.; Koopmans, G.F. Prediction of cadmium uptake by brown rice and derivation of soil–plant transfer models to improve soil protection guidelines. *Environ. Pollut.* 2009, 157, 2435–2444. [CrossRef] [PubMed]
- 39. Hu, B.F.; Shao, S.; Fu, Z.Y.; Li, Y.; Ni, H.; Chen, S.C.; Zhou, Y.; Jin, B.; Shi, Z. Identifying heavy metal pollution hot spots in soil-rice systems: A case study in South of Yangtze River Delta, China. *Sci. Total Environ.* **2019**, *658*, 614–625. [CrossRef] [PubMed]
- 40. Deng, M.H.; Zhu, Y.; Shao, K.; Zhang, Q.; Ye, G.H.; Shen, J. Metals source apportionment in farmland soil and the prediction of metal transfer in the soil-rice-human chain. *J. Environ. Manag.* **2020**, *260*, 110092. [CrossRef]
- 41. Deng, S.; Yu, J.; Wang, Y.T.; Xie, S.Q.; Ran, Z.X.; Wei, W. Distribution, transfer, and time-dependent variation of Cd in soil-rice system: A case study in the Chengdu plain, Southwest China. *Soil Till. Res.* **2019**, *195*, 104367. [CrossRef]
- 42. Du, F.; Yang, Z.G.; Liu, P.; Lin, W. Accumulation, translocation, and assessment of heavy metals in the soil-rice systems near a mine-impacted region. *Environ. Sci. Pollut. Res.* **2018**, *25*, 32221–32230. [CrossRef] [PubMed]
- 43. Wan, Y.; Huang, Q.; Wang, Q.; Ma, Y.; Su, D.; Qiao, Y.; Li, H. Ecological risk of copper and zinc and their different bioavailability change in soil-rice system as affected by biowaste application. *Ecotoxicol. Environ. Saf.* **2020**, *192*, 110301. [CrossRef] [PubMed]
- Zhang, X.F.; Liu, T.X.; Li, F.B.; Li, X.M.; Du, Y.H.; Yu, H.Y.; Wang, X.Q.; Liu, C.P.; Feng, M.; Liao, B. Multiple effects of nitrate amendment on the transport, transformation and bioavailability of antimony in a paddy soil-rice plant system. *J. Environ. Sci.* 2021, 100, 90–98. [CrossRef]
- 45. Brus, D.J.; Li, Z.B.; Song, J.; Koopmans, G.F.; Temminghoff, E.J.M.; Yin, X.B.; Yao, C.X.; Zhang, H.B.; Luo, Y.M.; Japenga, J. Predictions of spatially averaged cadmium contents in rice grains in the Fuyang Valley, PR China. *J. Environ. Qual.* **2009**, *38*, 1126–1136. [CrossRef]
- 46. Zhao, K.L.; Zhang, W.W.; Zhou, L.; Liu, X.M.; Xu, J.M.; Huang, P.M. Modeling transfer of heavy metals in soil–rice system and their risk assessment in paddy fields. *Environ. Earth Sci.* 2009, *59*, 519–527. [CrossRef]
- 47. Mu, T.; Wu, T.; Zhou, T.; Li, Z.; Ouyang, Y.; Jiang, J.; Zhou, D.; Hou, J.Y.; Wang, Z.Y.; Luo, Y.M.; et al. Geographical variation in arsenic, cadmium, and lead of soils and rice in the major rice producing regions of China. *Sci. Total Environ.* **2020**, 677, 373–381. [CrossRef]
- 48. Liu, X.; Gu, S.; Yang, S.; Deng, J.; Xu, J. Heavy metals in soil-vegetable system around E-waste site and the health risk assessment. *Sci. Total Environ.* **2021**, 779, 146438. [CrossRef] [PubMed]
- 49. Ma, H.H.; Peng, M.; Guo, F.F.; Liu, F.; Tang, S.Q.; Yang, Z.; Zhang, F.G.; Zhou, Y.L.; Yang, K.; Li, K.; et al. Factors Affecting the Translocation and Accumulation of Cadmium in a Soil-Crop System in a Typical Karst Area of Guangxi Province, China. *Environ. Sci.* **2021**, *42*, 1514–1522, (In Chinese with English abstract).
- 50. Myers, R.H. Classical and Modern Regression with Applications; Duxbury Press: Belmont, CA, USA, 1990.

- 51. Cortes, C.; Vapnik, V. Support-vector networks. Mach. Learn. 1995, 20, 273-297. [CrossRef]
- 52. Cristianini, N.; Shaw-Taylor, J. An Introduction to Support Vector Machines and Other Kernel Based Learning Methods; Cambridge University Press: Cambridge, UK, 2000.
- Bordoni, M.; Bittelli, M.; Valentino, R.; Chersich, S.; Persichilloet, M.G.; Meisina, C. Soil water content estimated by support vector machine for the assessment of shallow landslides triggering: The role of antecedent meteorological conditions. *Environ. Model. Assess.* 2018, 23, 333–352. [CrossRef]
- 54. Breiman, L. Random forests. Mach. Learn. 2001, 45, 5–32. [CrossRef]
- 55. Pouladi, N.; Møller, A.B.; Tabatabai, S.; Greve, M.H. Mapping soil organic matter contents at field level with Cubist, Random Forest and kriging. *Geoderma* **2019**, *342*, 85–92. [CrossRef]
- 56. Quinlan, J.R. C4.5: Programs for Machine Learning; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 1993.
- 57. Peng, J.; Biswas, A.; Jiang, Q.S.; Zhao, R.Y.; Hu, B.F.; Shi, Z. Estimating soil salinity from remote sensing and terrain data in southern Xinjiang Province, China. *Geoderma* **2019**, *337*, 1309–1319. [CrossRef]
- 58. Kuhn, M.; Wing, J.; Weston, S.; Williams, A.; Keefer, C.; Engelhardt, A.; Cooper, T.; Mayer, Z.; Kenkel, B.; R Core Team; et al. Package 'caret'. *R J.* **2020**, 223.
- 59. National Soil Survey Office. Chinese Soil Genus Records; China Agriculture Press: Beijing, China, 1995. (In Chinese)
- Hengl, T.; de Jesus, J.M.; Heuvelink, G.B.; Gonzalez, M.R.; Kilibarda, M.; Blagotić, A.; Shangguan, W.; Wright, M.N.; Geng, X.Y.; Bauer-Marschallinger, B.; et al. SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE* 2017, 12, e0169748. [CrossRef]
- 61. National Soil Survey Office. Chinese Soils; China Agriculture Press: Beijing, China, 1998; p. 1e1252. (In Chinese)
- 62. Kuhn, M. Caret: Classification and Regression Training. In *Astrophysics Source Code Library*; 2015. Available online: https://ascl.net/1505.003 (accessed on 20 April 2021).
- 63. Ministry of Health of the People's Republic of China (MHPRC). *National Hygienic Standard for Food in China GB2762-2017;* China Standard Press: Beijing, China, 2017. Available online: http://www.jlfsstd.net/db/files/Std1_6735986214439267.pdf (accessed on 14 April 2021).
- Ministry of Health of the People's Republic of China (MHPRC). Maximum Levels of Contaminants in Foods GB15199-1994; China Standard Press: Beijing, China, 1994. Available online: http://www.bzxzk.net/gjbz/24122011/88880.html (accessed on 14 April 2021).
- Ministry of Health of the People's Republic of China (MHPRC). Maximum Levels of Contaminants in Foods GB13106-1991; China Standard Press: Beijing, China, 1991. Available online: http://www.bzxzk.net/gjbz/24122011/88784.html (accessed on 14 April 2021).
- 66. Zhu, Z.Q.; Zhu, Y.W.; Shi, Z. *Research of Agricultural Soil Environment and Agricultural Product Safety*; China Agriculture Press: Beijing, China, 2009.
- Wang, X.; Zeng, X.; Chuanping, L.; Li, F.; Xu, X.; Lv, Y. Heavy metal contaminations in soil-rice system: Source identification in relation to a sulfur-rich coal burning power plant in Northern Guangdong Province, China. *Environ. Monit. Assess.* 2016, 188, 460. [CrossRef]
- 68. Yin, Y.M.; Zhao, W.T.; Huang, T.; Cheng, S.G.; Zhao, Z.L.; Yu, C.C. Distribution characteristics and health risk assessment of heavy metals in a soil-rice system in an e-waste dismantling area. *Huan Jing Ke Xue* **2018**, *39*, 916–926. (In Chinese)
- 69. Mao, C.; Song, Y.; Chen, L.; Ji, J.; Yuan, X.; Yang, Z.F.; Ayoko, A.G.; Frost, R.L.; Theiss, F. Human health risks of heavy metals in paddy rice based on transfer characteristics of heavy metals from soil to rice. *Catena* **2019**, *175*, 339–348. [CrossRef]
- 70. Hu, B.F.; Shao, S.; Ni, H.; Fu, Z.Y.; Huang, M.X.; Chen, Q.X.; Shi, Z. Assessment of potentially toxic element pollution in soils and related health risks in 271 cities across China. *Environ. Pollut.* **2021**, 270, 116196. [CrossRef]
- Liu, S.; Zhao, H.F.; Wu, K.N.; Zhang, Z.; Hou, Y.X.; Chen, T.Y.; Jin, Q. Evaluation of heavy metal distribution characteristics of agricultural soil–rice system in a high geological background area according to the influence index of comprehensive quality (IICQ). *Environ. Sci. Pollut. Res.* 2020, 27, 20920–20933. [CrossRef]
- Yu, Z.Y.; Dong, J.Q.; Fu, W.J.; Ye, Z.Q.; Li, W.Y.; Zhao, K.L. The transfer characteristics of potentially toxic trace elements in different soil-rice systems and their quantitative models in southeastern China. *Int. J. Environ. Res. Public Health* 2009, *16*, 2503. [CrossRef]
- He, M.J.; Shen, H.R.; Li, Z.T.; Wang, L.; Wang, F.; Zhao, K.L.; Liu, X.M.; Wendroth, O.; Xu, J.M. Ten-year regional monitoring of soil-rice grain contamination by heavy metals with implications for target remediation and food safety. *Environ. Pollut.* 2019, 244, 431–439. [CrossRef]
- 74. Xiao, R.; Guo, D.; Ali, A.; Mi, S.S.; Liu, T.; Ren, C.Y.; Li, R.H.; Zhang, Z.Q. Accumulation, ecological-health risks assessment, and source apportionment of heavy metals in paddy soils: A case study in Hanzhong, Shaanxi, China. *Environ. Pollut.* **2019**, *248*, 349–357. [CrossRef] [PubMed]
- 75. Kong, X.Y.; Liu, T.; Yu, Z.H.; Chen, Z.; Lei, D.; Wang, Z.W.; Zhang, H.; Li, Q.H.; Zhang, S.S. Heavy metal bioaccumulation in rice from a high geological background area in Guizhou Province, China. Int. J. Environ. Res. Public Health 2018, 15, 2281. [CrossRef]
- 76. CNEMC (China National Environmental Monitoring Center). *The Background Concentrations of Soil Elements of China;* China Environmental Science Press: Beijing, China, 1990. (In Chinese)
- 77. Cakmak, I. Enrichment of cereal grains with Zinc: Agronomic or generic biofortification. Plant Soil 2008, 302, 1–17. [CrossRef]
- 78. Nan, Z.; Li, J.; Zhang, J.; Cheng, G. Cadmium and zinc interactions and their transfer in soil-crop system under actual field conditions. *Sci. Total Environ.* **2002**, *285*, 187–195. [CrossRef]

- 79. Mohammad, A.; Moheman, A. The effects of cadmium and zinc interactions on the accumulation and tissue distribution of cadmium and zinc in tomato (*Lycopersicon esculentum* Mill.). *Arch. Agron. Soil Sci.* **2010**, *56*, 551–561. [CrossRef]
- 80. Bonfatti, B.R.; Hartemink, A.E.; Giasson, E.; Tornquist, C.G.; Adhikari, K. Digital mapping of soil carbon in a viticultural region of Southern Brazil. *Geoderma* **2016**, *261*, 204–221. [CrossRef]
- 81. Mosleh, Z.; Salehi, M.H.; Jafari, A.; Borujeni, I.E.; Mehnatkesh, A. The effectiveness of digital soil mapping to predict soil properties over low-relief areas. *Environ. Monit. Assess.* **2016**, *188*, 195. [CrossRef] [PubMed]
- 82. Mahmoudabadi, E.; Karimi, A.; Haghnia, G.H.; Sepehr, A. Digital soil mapping using remote sensing indices, terrain attributes, and vegetation features in the rangelands of northeastern Iran. *Environ. Monit. Assess.* 2017, 189, 500. [CrossRef]
- 83. da Silva Chagas, C.; de Carvalho Junior, W.; Bhering, S.B.; Calderano Filho, B. Spatial prediction of soil surface texture in a semiarid region using random forest and multiple linear regressions. *Catena* **2016**, *139*, 232–240. [CrossRef]
- 84. Ottoy, S.; De Vos, B.; Sindayihebura, A.; Hermy, M.; Van Orshoven, J. Assessing soil organic carbon stocks under current and potential forest cover using digital soil mapping and spatial generalization. *Ecol. Indic.* **2017**, *77*, 139–150. [CrossRef]
- Campbell, P.M.D.M.; Fernandes Filho, E.I.; Francelino, M.R.; Demattê, J.A.M.; Pereira, M.G.; Guimarães, C.C.B.; Pinto, L.A.D.S.R. Digital Soil Mapping of Soil Properties in the "Mar de Morros" Environment Using Spectral Data. *Rev. Bras. Ciência Solo* 2018, 42, 42. [CrossRef]
- Barman, U.; Choudhury, R.D. Soil texture classification using multi class support vector machine. *Information Inf. Process. Agric.* 2020, 7, 318–332. [CrossRef]
- 87. Peng, Y.; Xiong, X.; Adhikari, K.; Knadel, M.; Grunwald, S.; Greve, M.H. Modeling soil organic carbon at regional scale by combining multi-spectral images with laboratory spectra. *PLoS ONE* **2015**, *10*, e0142295. [CrossRef] [PubMed]
- 88. Peng, H.; Chen, Y.L.; Weng, L.P.; Ma, J.; Ma, Y.L.; Li, Y.T.; Islam, M.S. Comparisons of heavy metal input inventory in agricultural soils in North and South China: A review. *Sci. Total Environ.* **2019**, *660*, 776–786. [CrossRef]
- 89. Rossel, R.V.; Chen, C.; Grundy, M.J.; Searle, R.; Clifford, D.; Campbell, P.H. The Australian three-dimensional soil grid: Australia's contribution to the GlobalSoilMap project. *Soil Res.* 2015, *53*, 845–864. [CrossRef]
- 90. Peng, Y.; Kheir, R.B.; Adhikari, K.; Malinowski, R.; Greve, M.B.; Knadel, M.; Greve, M.H. Digital mapping of toxic metals in Qatari soils using remote sensing and ancillary data. *Remote Sens.* **2016**, *8*, 1003. [CrossRef]
- 91. Khaledian, Y.; Miller, B.A. Selecting appropriate machine learning methods for digital soil mapping. *Appl. Math. Model.* **2020**, *81*, 401–418. [CrossRef]
- 92. Zeng, F.; Ali, S.; Zhang, H.; Ouyang, Y.; Qiu, B.; Wu, F.; Zhang, G. The influence of pH and organic matter content in paddy soil on heavy metal availability and their uptake by rice plants. *Environ. Pollut.* **2011**, *159*, 84–91. [CrossRef]
- Osman, H.E.M.; Abdel-Hamed, E.M.W.; Al-Juhani, W.S.M.; Al-Maroai, Y.A.O.; El-Morsy, M.H.E.M. Bioaccumulation and human health risk assessment of heavy metals in food crops irrigated with freshwater and treated wastewater: A case study in Southern Cairo, Egypt. *Environ. Sci. Pollut. R* 2021, 1–13, Online ahead of print. [CrossRef]
- 94. Wang, Y.; Xu, W.; Li, J.; Song, Y.; Hua, M.; Li, W.; He, X. Assessing the fractionation and bioavailability of heavy metals in soil–rice system and the associated health risk. *Environ. Geochem. Health* **2021**, 1–18, Online ahead of print. [CrossRef]
- Cao, X.R.; Wang, X.Z.; Tong, W.B.; Gurajala, H.K.; Liu, M.; Hamid, Y.; Feng, Y.; He, Z.L.; Yang, X.E. Distribution, availability and translocation of heavy metals in soil-oilseed rape (*Brassica napus* L.) system related to soil properties. *Environ. Pollut.* 2019, 252, 733–741. [CrossRef] [PubMed]
- 96. Zhou, Y.J.; Jia, Z.Y.; Wang, J.X.; Chen, L.; Zou, M.M.; Li, Y.; Zhou, S.L. Heavy metal distribution, relationship and prediction in a wheat-rice rotation system. *Geoderma* **2019**, *354*, 113886. [CrossRef]
- 97. Ata-Ul-Karim, S.T.; Cang, L.; Wang, Y.; Wang, Y.; Zhou, D. Interactions between nitrogen application and soil properties and their impacts on the transfer of cadmium from soil to wheat (*Triticum aestivum* L.) grain. *Geoderma* **2020**, 357, 113923. [CrossRef]
- Vega, F.A.; Andrade, M.L.; Covelo, E.F. Influence of soil properties on the sorption and retention of cadmium, copper and lead, separately and together, by 20 soil horizons: Comparison of linear regression and tree regression analyses. *J. Hazard Mater.* 2010, 174, 522–533. [CrossRef]
- 99. Gupta, S.; Nayek, S.; Saha, R.N.; Satpati, S. Assessment of heavy metal accumulation in macrophyte, agricultural soil, and crop plants adjacent to discharge zone of sponge iron factory. *Environ. Earth Sci.* **2008**, *55*, 731–739. [CrossRef]
- 100. Shen, X.Y.; Chen, S.G.; Wang, Y.; Wang, Y.M.; Cai, W.X. Study on different clays as adsorbents in heavy metals-containing waste water treatment. *Environ. Pollut. Control* **1998**, *20*, 15–18.
- 101. Gu, H.H.; Qiu, H.; Tian, T.; Zhan, S.S.; Deng, T.H.B.; Chaney, R.L.; Wang, S.Z.; Tang, Y.T.; Morel, J.L.; Qiu, R.L. Mitigation effects of silicon rich amendments on heavy metal accumulation in rice (*Oryza sativa* L.) planted on multi-metal contaminated acid soil. *Chemosphere* 2011, *83*, 1234–1240. [CrossRef]
- 102. Zhao, F.J.; Ma, Y.; Zhu, Y.G.; Tang, Z.; McGrath, S.P. Soil contamination in China: Current status and mitigation strategies. *Environ. Sci. Technol.* **2015**, *49*, 750–759. [CrossRef]
- 103. Zhou, H.; Zeng, M.; Zhou, X.; Liao, B.H.; Peng, P.Q.; Hu, M.; Zhu, W.; Wu, Y.J.; Zou, Z.J. Heavy metal translocation and accumulation in iron plaques and plant tissues for 32 hybrid rice (*Oryza sativa* L.) cultivars. *Plant Soil* 2015, 386, 317–329. [CrossRef]
- 104. Zhang, M.; Shan, S.D.; Chen, Y.G.; Wang, F.; Yang, D.Y.; Ren, J.K.; Lu, H.Y.; Ping, L.F.; Chai, Y.J. Biochar reduces cadmium accumulation in rice grains in a tungsten mining area-field experiment: Effects of biochar type and dosage, rice variety, and pollution level. *Environ. Geochem. Health* 2019, 41, 43–52. [CrossRef]

- 105. Guo, F.Y.; Ding, C.F.; Zhou, Z.G.; Huang, G.X.; Wang, X.X. Effects of combined amendments on crop yield and cadmium uptake in two cadmium contaminated soils under rice-wheat rotation. *Ecotoxicol. Environ. Saf.* **2018**, *148*, 303–310. [CrossRef]
- Huang, Y.; Sheng, H.; Zhou, P.; Zhang, Y.Z. Remediation of Cd-contaminated acidic paddy fields with four-year consecutive liming. *Ecotoxicol. Environ. Saf.* 2020, 188, 109903. [CrossRef]
- 107. Wuana, R.A.; Okieimen, F.E. Heavy metals in contaminated soils: A review of sources, chemistry, risks and best available strategies for remediation. *ISRN Ecol.* 2011, 2011, 1–20. [CrossRef]
- 108. Liu, P.; Liu, Z.H.; Hu, Y.M.; Shi, Z.; Pan, Y.C.; Wang, L. A hybrid back propagation neural network and particle swarm optimization for estimating soil heavy metal contents using hyper-spectral data. *Sustainability* **2019**, *11*, 419. [CrossRef]
- Wen, Y.B.; Li, W.; Yang, Z.F.; Zhuo, X.X.; Guan, D.X.; Song, Y.Y.; Guo, C.; Ji, J.F. Evaluation of various approaches to predict cadmium bioavailability to rice grown in soils with high geochemical background in the karst region, Southwestern China. *Environ. Pollut.* 2020, 258, 113645. [CrossRef] [PubMed]
- 110. Liu, B.; Mo, C.H.; Zhang, Y. Using cadmium bioavailability to simultaneously predict its accumulation in crop grains and the bioaccessibility in soils. *Sci. Total Environ.* **2019**, *665*, 246–252. [CrossRef]