



Multivariate linear regression models for predicting metal content and sources in leafy vegetables and human health risk assessment in metal mining areas of Southern Jharkhand, India

Soma Giri¹ · Mukesh Kumar Mahato¹ · Abhay Kumar Singh¹

Received: 9 November 2020 / Accepted: 11 January 2021 / Published online: 28 January 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH, DE part of Springer Nature 2021, corrected publication 2021

Abstract

The present study was intended to investigate the metal concentrations in the leafy vegetables, irrigation water, soil, and atmospheric dust deposition in the iron and copper mining areas of Southern Jharkhand, India. The study aimed to develop a multivariate linear regression (MVLRL) model to predict the concentration of metals in leafy vegetables from the metals in associated environmental factors and assessment of the risk to the local population through the consumption of leafy vegetables and other allied pathways. The developed species-specific MVLRL models were well fitted to predict the concentration of metals in the leafy vegetables. The coefficient of determination values (R^2) was greater than 0.8 for all the species-specific models. Risk assessment was carried out considering multiple pathways of ingestion, inhalation, and dermal contact of vegetables, soil, water, and free-fall dust. Consumption of leafy vegetables was the major route of metal exposure to the local population in both the metal mining areas. The average hazard index (HI) value considering all the metals and pathways was calculated to be 5.13 and 12.1, respectively for iron and copper mining areas suggesting considerable risk to the local residents. Fe, As, and Cu were the major contributors to non-carcinogenic risk in the Iron mining areas while in the case of copper mining areas, the main contributors were Co, As, and Cu.

Keywords Metals · Leafy vegetables · Health risk assessment · Multivariate linear regression model

Introduction

A significant quantity of tailings and wastewater is produced during the mining exploitation and ore smelting processes which results in severe metal pollution in the surface water, groundwater, soils, and food crops (Amoakwah et al. 2020; Rodríguez-Estival et al. 2019). The metals in the mining environment pose a health risk to the population in the vicinity through different pathways like ingestion, dermal contact, or inhalation of dust and soil; ingestion and dermal contact of

water; and consumption of food crops grown in contaminated soils or irrigated with polluted water (Doabi et al. 2018; Park and Choi 2013; Wu et al. 2015).

Consumption of metal-contaminated crops, particularly vegetables, from a mining dominated area form a major exposure pathway for the local consumers. Metal concentrations in plants grown in mine contaminated soils are considerably elevated compared to the reference soils (Cao et al. 2016). Nevertheless, in addition to soil, irrigation water and atmospheric deposition are impending sources of heavy metals in the vegetables (Bi et al. 2018; El-Radaideh and Al-Taani 2018; Li et al. 2012). However, in a similar environment, concentrations of metals in the vegetables are also dependent on plant categories and species.

Leafy vegetables accrue elevated metal concentrations compared to rootstalk and fruiting vegetables. Foliar deposition of metals through atmospheric dust emitted from anthropogenic sources form the governing pathway of metal uptake for the leafy vegetables (Luo et al. 2011; Noh and Jeong 2019; Xu et al. 2015). Also, the leafy vegetables have higher transportation rates of metals compared to other vegetable types

The original online version of this article was revised: The entries in the 3rd column of Table are overlapping.

Responsible Editor: Lotfi Aleya

✉ Soma Giri
soma0307@gmail.com

¹ Natural Resources and Environmental Management Group, CSIR-Central Institute of Mining and Fuel Research, Barwa Road, Dhanbad 826015, India

(Hu et al. 2017). Crop-specific prediction models have been used by many researchers to estimate the uptake and accumulation of metals by the crops and thus to predict the dietary hazards of metals (Eid et al. 2018, 2020; Kumar et al. 2019; Legind and Trapp 2010). The empirical models can also be used for practical applications like soil remediation and health risk assessment (Kumar et al. 2020; Wang et al. 2015).

Singhbhum craton and shear zones are the highly mineralized areas of India harboring one of the biggest sources of high-grade iron ore and copper, respectively. These areas are under immense load of metal mining activities and subsequently their processing industries which together pave the way for metal contamination of the environs and consequently the food chains. There are a number of research works on metals in soil-vegetable systems but not many deals with the multifaceted system comprising of soil, dust, water, and leafy vegetables in a mining environment. In this complex system, human health risk needs to be quantified considering all the applicable pathways. Since there were multiple pathways of metals exposure, accounting for all of them is mandatory for the overall risk assessment for the populace in the vicinity of mining areas. In addition, a stepwise multivariate linear regression (MVLRL) model to evaluate the concentration of metals in the leafy vegetables from metals in soil, irrigation water, and atmospheric dust was also imperative. The species-specific MVLRL models need to be versatile such that they can be applicable in diverse mining areas. The MVLRL model can establish the sources of the metals in the leafy vegetables also. In view of these facts, the present study was carried out to examine the metal concentrations in the locally growing leafy vegetables in iron and copper mining areas along with soil, irrigating water, and atmospheric dust.

The objectives of the study envisaged to (1) investigate the metal contamination in the intricate system of soils, irrigating water, atmospheric dust, and leafy vegetables in two metal mining environment; (2) determine the contribution of the different matrices, i.e., soil, water, and dust on the concentration of metals in the leafy vegetables using stepwise multivariate linear regression (MVLRL); (3) to predict the concentration of metals in the leafy vegetables using MVLRL; and (4) quantify the health risk to the local populace considering the multiple pathways of metal exposure, i.e., ingestion, dermal contact, or inhalation of dust and soil; ingestion and dermal contact of water; and consumption of leafy vegetables.

Materials and method

Study area

The study was taken up in the iron mining areas of Singhbhum craton and copper mining areas of Singhbhum Shear Zone known for their high-grade ores of iron and copper. Both the

study areas are situated in the Southern Jharkhand of Eastern India (Fig. 1) and form the India's most mineralized and industrialized zones.

The Singhbhum craton consists of Precambrian rocks that are circumscribed by the arcuate Singhbhum Shear Zone in the north and Sukinda thrust in the south. The craton is a part of volcano-sedimentary Archean Iron ore group (IOG) that harbors one of the largest iron and manganese deposits of India (Upadhyay et al. 2010). A number of phases of Mesoarchean granitic entities, i.e., Singhbhum granite and Bonai granite are found within the craton. The Singhbhum granite is the most widespread unit and includes synformal keels of banded iron formations (BIF)-bearing greenstone belts (Saha 1994). The minerals that are found in the area consist of hematite, magnetite, goethite, chromites, romanechite, cryptomelane, pyrolusite, gibbsite, kaolinite, and quartz.

The Singhbhum shear zone forms an arcuate vastly deformed linear zone in the Singhbhum crustal province and is known to be one of the most potential sulfide-bearing stretch of India. The copper mineralization along Singhbhum copper belt is located along the Dhanjori group of rocks at south of shear zone and Singhbhum group of rocks at north of shear zone. The main rock types of the region are soda granite, Dhanjori lavas, and epidiorite-schist. The sulfide minerals of the zone consists predominantly of chalcopyrite, followed by pyrite and pyrrotite, accompanied by magnetite and minor pentlandite, millerite, violarite, sphalerite, and molybdenite (Dunn and Dey 1942; Mukherjee 1968).

Collection of samples and processing for metal analysis

Locally grown leafy vegetables (31 samples) that are commonly consumed by the local population were collected comprising of 4 varieties (Red amaranthus, Spinach, Drumstick leaves, and Mustard leaves) from six villages each from both the study areas. Supplementary Table 1 displays the local, English, and scientific names and the collection sites for each. For each leafy vegetable, multiple picks were collected and then composited to one sample. The leaves were washed many times with the tap water to remove the dust adhered on to the leaf surfaces. Final wash was done with the ultrapure water, and then, the samples were air-dried followed by drying in an oven at 80 °C to a constant weight. 0.5 g of the grinded and homogenized leaves were subjected to acid digestion in a microwave digestion system (model: Anton Paar Multiwave 3000) using concentrated HNO₃ and 30% H₂O₂.

Composite soil samples were prepared for each site. About 1 kg soil sample was collected from the plough layer (0–15 cm) from various points from where the vegetables were collected. These soil samples were mixed together, air-dried at room temperature (25–30 °C), and sieved through 2 mm

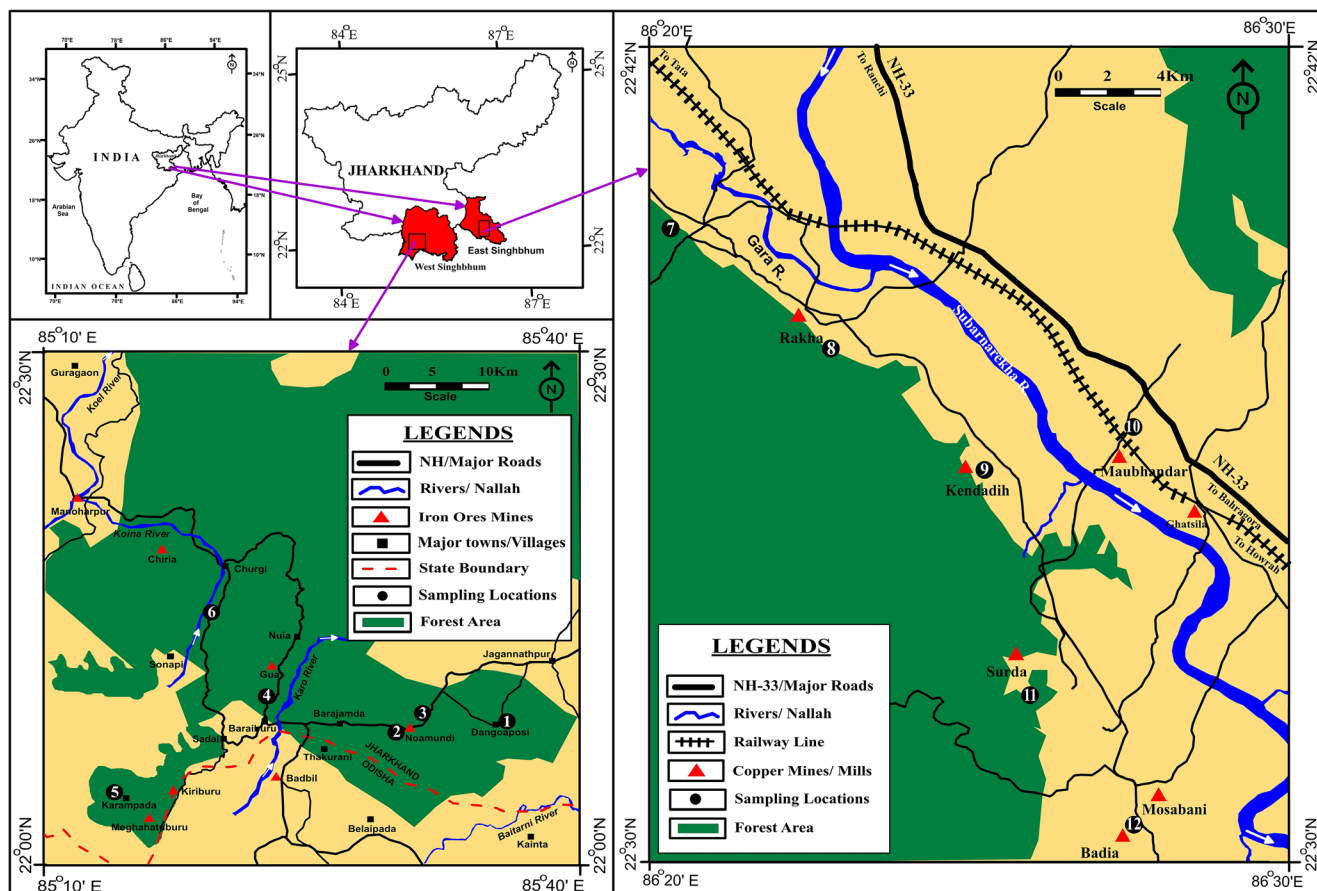


Fig. 1 Map of the study area with sampling locations (modified from Giri et al. 2020)

screen to separate out the pebbles. The fraction that passed through the 2 mm screen was further quarter coned to reduce the size to 100 g which was oven-dried at 110 °C for 24 h and crushed and the whole quantity was passed through a 200 mesh screen. 0.25 g of the sieved soil samples was digested in a microwave digestion system using a digestion solution of HNO₃ (4.5 ml), HF (2 ml), HCl (1 ml), and H₂O₂ (0.5 ml) using the EPA 3052 method (USEPA 1996).

Atmospheric dust samples were also collected from the selected sites using the gravity technique (BIS 2006). Dustfall sampling devices were assembled with high-density plastic buckets mounted on an iron tripod stand 1.2 m above the sampling floor to avoid the collection of dust picked up by wind eddies. It was exposed to the atmosphere for a month with a bird-shield. After the collection, the samples were oven-dried at 110 °C for 24 h and crushed, and the whole quantity was passed through a 200 mesh screen. The digestion procedure was the same as the soil samples.

From each site, the water used for irrigating the vegetable crops was collected, filtered, and preserved at pH < 2 by adding suprapure HNO₃ for heavy metal analyses (Radojevic and Bashkin 1999).

Heavy metal analyses

As, Cd, Co, Cr, Cu, Fe, Mn, Ni, Pb, and Zn were determined in the processed vegetable, soil, dust, and water samples using inductively coupled plasma mass spectrometer (ICP-MS, Model: Perkin Elmer Elan DRC-e). The stability of the calibration curve was checked and confirmed after every 15th sample using calibration blank and a standard not used in the calibration.

Quality assurance

Spinach leaves standard reference material (NIST SRM 1570a) was used for validation of the analytical procedure. The analytical results for water were validated using standard reference material of natural water (NIST-1640a) while for soil and dust, estuarine sediment (NIST-1646a) was used. The % recovery of the metals for NIST-1570a, NIST-1640a, and NIST-1646a varied from 95.6 to 105.9, 97.1 to 104.0, and 90.9 to 108.8, respectively (Supp. Table 2).

Non-carcinogenic risk assessment

There are multiple pathways of metal exposure to the local population. The pathways which are considered in the study are as follows:

- i. Ingestion pathway for the leafy vegetables
- ii. Ingestion and dermal pathways for water
- iii. Ingestion, dermal, and inhalation pathways for soil
- iv. Ingestion, dermal, and inhalation pathways for dust

Non-carcinogenic risk posed by each metal present in different matrices (vegetable, water, soil, and dust) was typified by a hazard quotient (HQ) which is the ratio of average daily dose (ADD) to the reference doses for ingestion, dermal, or inhalation pathways (RfD_{ing} or RfD_{derm} or RfC_i) (Eq. 1)

$$HQ = ADD/RfD_{ing} \text{ or } RfD_{derm} \text{ or } RfC_i \tag{1}$$

Expressions for calculating HQ for different pathways have been listed in Table 1. Reference doses for the ingestion pathway (RfD_{ing} in mg/kg/day), dermal pathway (RfD_{derm} in mg/kg/day), and inhalation pathway (RfC in mg/m³) are obtained from USEPA (2011). These represent the safe doses with reference to a particular metal that a human body can bear without any probable adverse health effects. The input parameters for calculating HQ are provided in Supp. Table 3. All the input values originated from USEPA (USEPA 2002, 2004, 2011, 2017) except for IR_{veg}, IR_{water}, and BW which

were taken in Indian context from NSSO 2014; Dang et al. (1994); and NFI 2010, respectively.

ADD and RfD of different metals are different in different pathways; thus, corresponding HQ(s) should also be different. Thus, total hazard index (HI) considering multiple metals and multiple pathways of exposure comprising “n” number of pathways and “m” number of metals may be expressed as Eq. 2,

$$HI = \sum_{j=1}^n \sum_{k=1}^m HQ_j^k \tag{2}$$

Thus, if we consider the water matrix, comprising of four heavy metals, namely, Fe, Mn, Zn, and Cu, the comprehensive hazard index arising from ingestion and dermal absorption pathways might be expressed as,

$$HI = [HQ_{ing}^{Fe} + HQ_{ing}^{Mn} + HQ_{ing}^{Zn} + HQ_{ing}^{Cu}] + [HQ_{derm}^{Fe} + HQ_{derm}^{Mn} + HQ_{derm}^{Zn} + HQ_{derm}^{Cu}]$$

Likewise, the HI for all the considered matrices and pathways may be calculated.

HQ_j^k > 1 implies that the health risk associated with a particular metal in a particular matrix through a particular pathway is more than the accepted level.

Table 1 Calculation of hazard quotient (HQ) (USEPA 2004, 2011)

Sl. No	Hazard quotient	Formula	Classifications of exposure parameters
1	HQ _{ing(veg)} (through the ingestion of leafy vegetable)	$HQ_{ing} = C_{veg} \times \frac{IR_{veg} \times EF \times ED}{BW \times AT \times R_f D_{ing}}$	HQ _{ing} = hazard quotient through ingestion pathway
2	HQ _{ing(water)} (through the Ingestion of water)	$HQ_{ing} = C_{water} \times \frac{IR_{water} \times EF \times ED}{BW \times AT \times R_f D_{ing}} \times 10^{-3}$	HQ _{inh} = hazard quotient through inhalation pathway
3	HQ _{dermal(water)} (through the dermal contact of water)	$HQ_{dermal} = C_{water} \times \frac{SA \times K_p \times ET \times EF \times ED}{BW \times AT \times R_f D_{dermal}} \times 10^{-6}$	HQ _{dermal} = hazard quotient through dermal contact
4	HQ _{ing(soil/dust)} (through the ingestion of soil/dust)	$HQ_{ing} = C_{soil/dust} \times \frac{IR_{soil/dust} \times EF \times ED}{BW \times AT \times R_f D_{ing}} \times 10^{-6}$	C _{veg} = concentration of the metal in vegetable (mg/kg)
5	HQ _{dermal(soil/dust)} (through the dermal contact of soil/dust)	$HQ_{dermal} = C_{soil/dust} \times \frac{SA \times AF \times ABS \times EF \times ED \times EV}{BW \times AT \times R_f D_{dermal}} \times 10^{-6}$	C _{water} = concentration of the metal in water (µg/L)
6	HQ _{inh(soil/dust)} (through the inhalation of soil/dust)	$HQ_{inh} = C_{soil/dust} \times \frac{EF \times ED}{PEF \times AT \times R_f C_i}$	C _{soil/dust} = concentration of the metal in soil/dust (mg/kg)

IR_{veg} = ingestion rate of vegetables (kg/day)
 IR_{water} = ingestion rate of water (L/day)
 IR_{soil/dust} = ingestion rate of soil/dust (mg/day)
 EF = exposure frequency (days/year)
 ED = exposure duration (years)
 ET = exposure time (h/day)
 EV = events (events/day)
 BW = body weight (Kg)
 AT = averaging time (days)
 K_p = dermal permeability coefficient in water (cm/hr)
 PEF = particle emission factor (m³/kg)
 SA = surface area of exposed skin (cm²)
 AF = skin adherence factor for dust/soil (mg/cm² event)
 ABS = dermal absorption factor
 R_fD_{ing} = reference dose for ingestion (mg/kg/day)
 R_fC_i = inhalation reference concentration (mg/m³)
 R_fD_{dermal} = reference dose for dermal contact (mg/kg/day)

Similarly, $HI > 1$ also implies that total risk exceeds the acceptable limit even though each of its constituent components (HQ_j^k) may be < 1 .

Proposed metal transport model

Stepwise linear multiple regression analyses were used to recognize the influential environmental parameters, i.e., cultivated soils, irrigation waters, and atmospheric dust to ascertain models for predicting metal concentrations in the leafy vegetables (Xu et al. 2015; Yang et al. 2014). Stepwise MVLRL checks the best combination of the independent variables to predict the dependent variable and discards the insignificant variables. F test was used in regression analysis to compare the disparity between the actual data and the modeled data derived from the fitted equation. The Durbin Watson values are also studied to observe the autocorrelation in the residuals.

The simplest possible model for describing transport of metal from soil, water, and dust to the leafy vegetables was envisaged. Since the sources of metals in vegetables are primarily soil along with irrigation water and atmospheric dust, their inter relationship may be functionally expressed as,

$$V_i = f(X_i, Y_i, Z_i) \quad (3)$$

where V_i , X_i , Y_i , and Z_i are the i^{th} heavy metal concentration in the vegetable, water, soil, and dust, respectively.

A multivariate linear regression (MVLRL) model was chosen as the function $f(X_i, Y_i, Z_i)$ which may be expressed as

$$V_i = a_0 + a_1.X_i + a_2.Y_i + a_3.Z_i \quad (4)$$

where a_0 is the constant and a_1 , a_2 , and a_3 are the coefficients for water, soil, and dust, respectively.

Model validation

Multiple regression modeling is a statistical tool used extensively for the prediction analysis and examines the effect of more than one independent variable over a response (DeForest et al. 2018). However, it is imperative to make certain that the prediction model works appropriately with real data and predicted data. For the performance evaluation, three assessment tools were adopted:

- i. A regression plot (scatter plot) of actual values against the modeled values allows a visual assessment of the model performance. More the scatters tend to concentrate close to the equiline (denoted as 1:1 line), better is the agreement between the actual and predicted values, and thus, better is the model.
- ii. For the validation process, model efficiency (ME) is widely used to minimize error in the selection and feasibility of a model (Eid et al. 2018). It is also known as

Nash and Sutcliffe coefficient of efficiency. The ME of regression models was calculated using Eq. 5:

$$ME = 1 - \frac{\sum_{i=1}^N (O_i - P_i)^2}{\sum_{i=1}^N (O_i - \bar{O})^2} \quad (5)$$

where O_i and P_i are the observed values and predicted values, respectively, while \bar{O} is the mean of the observed values. The sample size is “ N .”

- iii. To quantify the prediction error, the root mean square error (RMSE) is used which is calculated as Eq. 6

$$\sqrt{\frac{\sum_{i=1}^N (O_i - P_i)^2}{N}} \quad (6)$$

where the variables O_i , P_i , and N are same as in Equation 5.

The criteria for the goodness of fit evaluation for ME and RMSE are given in Supplementary Table 4.

Results and discussion

Heavy metals in leafy vegetables, irrigation water, soil, and atmospheric dust samples

The concentrations of the 10 metals (As, Cd, Co, Cr, Cu, Fe, Mn, Ni, Pb, and Zn) in the different environmental matrices for both the mining areas are depicted in Tables 2 and 3. The samples were collected from six locations each from the iron mining areas of West Singhbhum (Table 2) and copper mining areas of East Singhbhum (Table 3). It is a common observation that the concentrations of the metals in the environmental matrices do not follow a normal distribution which can be attributed to the extreme values that are often encountered. Since the central tendency is better represented by the geometric mean for a skewed data, it has been used in the study to signify the average concentrations of the metals. Compared to the Indian standards for food (Awasthi 2000), none of the metal concentrations in the leafy vegetables from both the areas exceeded the standards except for Cu in some samples of Red amaranthus collected from copper mining areas.

Water, soil, and free-fall atmospheric dust samples were collected from the same locations at the same time as that of the collection of the leafy vegetables. The metal concentrations of the soil and dust samples were compared to the metals in the average shale values (Turekian and Wedepohl 1961) to reveal that almost all the metals surpassed the values which

Table 2 Concentration of metals in leafy vegetables, water, soil, and dust in the iron mining areas of West Singhbhum (unit: water—µg/L; soil, dust, food—mg/kg)

Matrix		As	Cd	Co	Cr	Cu	Fe	Mn	Ni	Pb	Zn
Drumstick leaf	Minimum	0.009	0.0005	0.052	0.386	1.14	140.6	12.0	0.102	0.007	4.47
	Maximum	0.284	0.096	0.161	0.772	2.92	424.5	88.8	0.695	0.417	9.68
	Geomean	0.049	0.001	0.076	0.590	1.74	217.5	33.4	0.325	0.057	6.23
Red Amaranthus leaf	Minimum	0.006	0.032	0.031	0.106	0.469	91.9	12.3	0.019	0.005	6.04
	Maximum	0.053	0.898	0.304	0.481	2.553	250.1	67.8	1.137	0.138	12.04
	Geomean	0.017	0.091	0.083	0.225	1.039	162.3	24.0	0.097	0.045	7.86
Mustard leaf	Minimum	0.003	0.0002	0.017	0.123	0.455	55.8	2.48	0.057	0.003	2.712
	Maximum	0.019	0.144	0.040	0.357	2.336	202.0	22.32	0.200	0.144	9.708
	Geomean	0.010	0.007	0.028	0.199	0.810	123.0	8.29	0.122	0.027	4.042
Spinach	Minimum	0.004	0.043	0.022	0.229	0.486	44.2	3.30	0.213	0.005	3.06
	Maximum	0.010	0.278	0.060	0.721	1.465	73.6	11.49	1.025	0.047	7.04
	Geomean	0.007	0.100	0.039	0.356	0.993	59.5	6.23	0.424	0.020	5.26
Water	Minimum	0.030	0.011	0.276	0.895	0.357	58.6	5.1	1.90	0.025	3.6
	Maximum	0.887	0.054	1.443	1.173	3.69	1426.2	114.0	5.23	0.164	684.4
	Geomean	0.202	0.026	0.582	0.982	0.618	436.6	13.6	3.37	0.067	33.9
Soil	Minimum	49.0	0.21	12.3	48.7	66.2	41135.9	510.8	40.1	6.7	130.9
	Maximum	66.9	3.15	51.4	295.8	147.5	105542.6	2547.4	101.3	13.9	803.4
	Geomean	61.2	0.77	24.0	124.8	84.1	75145.3	1156.0	68.8	8.7	233.6
Dust	Minimum	39.1	0.19	2.90	28.7	85.5	36245.6	298.3	18.3	63.8	118.1
	Maximum	60.6	0.54	4.05	278.6	240.6	126913.7	1069.3	24.8	253.2	991.5
	Geomean	47.7	0.36	3.50	93.3	153.5	69138.5	542.3	20.5	115.9	401.3
Indian Standards Food ^a		1.1	1.5			30				2.5	50
Average shale ^b		13	0.3	19	90	45	47200	850	68	20	95
Irrigating water standards ^c		100			100	200		200	200		2000

^a Awasthi 2000

^b Turekian and Wedepohl 1961

^c FAO 1985

can be attributed to the metal bearing deposits of both the study areas in conjunction with the mining and industrial activities. The irrigation water that was used for the vegetable crops in the region was mostly groundwater, which was also used for drinking purposes. However, in one location of copper mining area (Kendadih), the mine water from the copper mine was used for the irrigation. Mine water is the water emanating from the mines and is often loaded with metals in the case of metal mines since it is being dewatered from the working face of the mines. The metal concentrations did not exceed the irrigation water standards (FAO 1985) for any of the locations but for the mine water from Kendadih which exceeded the standards for Cu, Mn, and Ni. There is a marked difference in the metal concentrations between the two study areas as for most of the matrices; higher Cu, Ni, and Co concentrations were observed in the copper mining areas while Fe and Mn were higher in the iron mining areas. These groups of metals are associated with the particular geological formations of both areas along with the respective metal mining activities.

Human health risk assessment

The health risk was assessed for 9 different pathways for the local population. The pathways comprised of ingestion of vegetables, water, soil, and dust; dermal contact with water, soil, and dust; and inhalation pathway for soil and dust. The hazard quotients for the metals for considering all the pathways of exposure did not exceed unity for the iron ore mining area except for Co and Cr at one location and Fe at two locations through ingestion of vegetables (Fig. 2a). In the case of the copper mining areas, the HQ for the metals As, Co, Cu, and Mn surpassed the reference value of one for some of the pathways at some of the locations. All of these incidences were related to ingestion of leafy vegetables and water which implied considerable risk to the local population due to the consumption of metal-contaminated vegetables and water (Fig. 2b).

The combined risk due to the multiple metals was evaluated by summation of the HQs to obtain the hazard index (HI) (Table 4). The HI for the different pathways for the different

Table 3 Concentration of metals in leafy vegetables, water, soil, and dust in the copper mining areas of East Singhbhum (unit: water— $\mu\text{g/L}$; soil, dust, food— mg/kg)

Matrix		As	Cd	Co	Cr	Cu	Fe	Mn	Ni	Pb	Zn
Drumstick leaf	Minimum	0.095	0.003	0.044	0.274	3.01	62.1	3.26	0.428	0.049	4.07
	Maximum	0.799	0.059	0.214	0.881	24.13	184.3	10.69	1.670	0.821	18.62
	Geomean	0.264	0.028	0.116	0.428	8.39	103.4	5.64	0.792	0.191	7.47
Red Amaranthus leaf	Minimum	0.012	0.003	0.103	0.384	2.6	58.2	5.57	0.357	0.011	5.05
	Maximum	0.724	0.033	0.514	0.547	35.2	238.7	15.05	2.144	0.095	18.10
	Geomean	0.050	0.011	0.214	0.449	13.5	118.5	11.53	1.051	0.030	9.30
Mustard leaf	Minimum	0.013	0.029	0.035	0.357	4.83	54.3	11.1	0.392	0.079	9.54
	Maximum	0.028	0.051	0.123	0.625	22.6	145.4	14.9	2.361	0.097	13.72
	Geomean	0.019	0.038	0.066	0.472	10.5	88.8	12.9	0.962	0.088	11.44
Water	Minimum	0.247	0.009	0.554	0.678	1.167	411.7	0.714	2.851	0.014	4.96
	Maximum	1.300	0.112	244.5	4.319	1874.1	1547.6	977.1	916.0	0.158	232.0
	Geomean	0.677	0.037	2.882	1.131	23.07	794.6	34.29	14.51	0.071	19.18
Soil	Minimum	14.2	0.09	13.9	62.5	72.6	20415.0	119.2	34.1	21.5	65.2
	Maximum	61.0	0.73	54.9	396.5	2965.3	58584.6	882.3	285.6	67.7	368.2
	Geomean	28.7	0.29	37.2	118.9	422.1	37719.7	464.6	118.3	41.4	184.8
Dust	Minimum	37.1	0.49	5.9	97.6	219.6	14454.6	244.6	10.2	32.0	215.5
	Maximum	50.9	2.38	23.9	138.3	7858.7	75562.5	372.8	68.0	201.7	430.5
	Geomean	41.1	0.93	14.3	127.7	2966.6	28717.1	310.8	35.6	51.6	356.5
Indian Standards Food ^a		1.1	1.5			30				2.5	50
Average shale ^b		13	0.3	19	90	45	47200	850	68	20	95
Irrigating water standards ^c		100			100	200		200	200		2000

^a Awasthi 2000^b Turekian and Wedepohl 1961^c FAO 1985

locations of the two study areas depicted that the most important route of exposure to the metals was the ingestion of the leafy vegetables, irrespective of locations and area. Intake of the leafy vegetables contributed 71.9% and 80.7% of the total risk for the iron and copper mining areas, respectively. In the iron mining areas, the next significant pathway was ingestion of soil (8.6% of total risk) followed by ingestion of dust and drinking of water. In the area, the groundwater is used for irrigating the vegetable fields which is also used for drinking as well. In the case of copper mining areas, the risk apportionment is a little different, where the second most relevant pathway for exposure of metals was the drinking water pathway contributing 10.6% of the total risk. This has been calculated excluding the risk from the water pathway of the Kendadih location where mine water was used for irrigation and which is not used for drinking. The other important pathways were the ingestion of dust and soil. Ingestion of soil and dust is found to be an important pathway of metal exposure for both the study areas. Ingestion exposure can occur via the intentional or inadvertent non-dietary ingestion of soil and dust on surfaces or objects that are contacted via hand-to-mouth or object-to-mouth activity (USEPA 2017).

Considering the HIs, it can be inferred that for both the mining areas, the ingestion pathway is the primary exposure route irrespective of the matrix. Dermal and inhalation routes form the secondary exposure mode. Considering both the mining areas, the overall risk was higher for the locations of the copper mining area compared to the iron mining areas. The overall risk as represented by HI was more than double for the copper mining area (12.1) than the iron mining area (5.13) which may be attributed to the higher concentrations of Cu and Co in environs of the copper mining areas. Cu and Co are known to have a greater risk on human health compared to Fe and Mn which are found to be at large in iron mining areas.

Regression models

Regression models were formed to estimate the metal concentrations in the leafy vegetables based on the metal concentrations in different environmental matrices. Stepwise multivariate linear regression analysis was performed on the basis of the three factors, i.e., metal concentrations in the soil, irrigation water, and atmospheric dust. The results and prediction precisions of the models are depicted in Table 5. The ANOVA

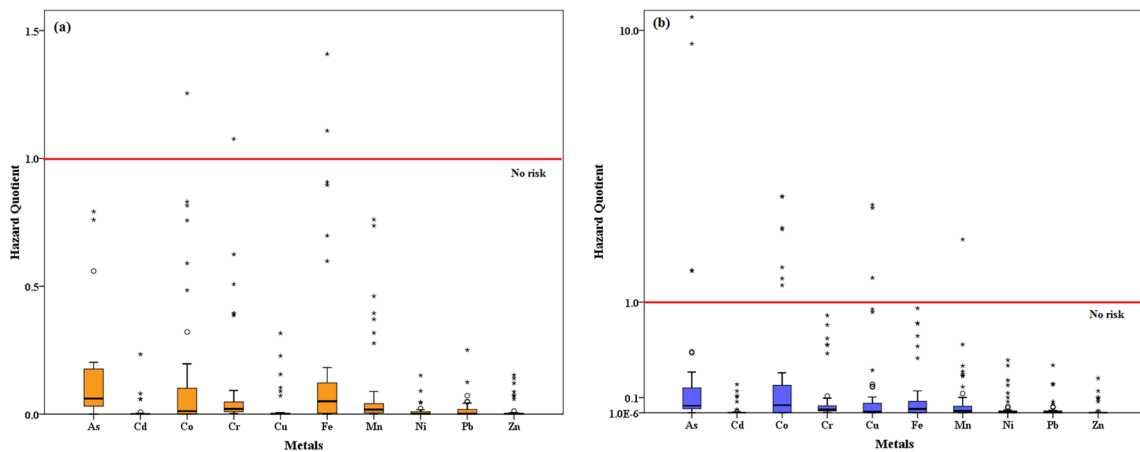


Fig. 2 Boxplots of hazard quotient (HQ) of metals considering the 9 pathways in **a** iron mining areas of West Singhbhum, **b** copper mining areas of East Singhbhum

results revealed that the correlation between the measured and predicted metal concentrations by model equations was statistically significant with high *F* values and low *P* values (level of significance $P < 0.001$). The Durbin Watson values are also between 1.5 and 2.5 suggesting the absence of autocorrelation in the residuals (Rafiee et al. 2010). The high (> 0.7) coefficient of determination values (R^2) which is an indicative of the goodness of fit of the developed models is supportive for the prediction of effective metal uptake by the leafy vegetables from the different environmental components. The R^2 of 0.725 was obtained for the model when metals for all the samples of leafy vegetables irrespective of the species were considered cumulatively. However, the R^2 values further improved when the different species of the leafy vegetables were considered individually thus advocating for species-specific regression models for metal uptake. The R^2 for all species-

specific models ranged between 0.821 and 0.980, the highest being for the Spinach.

The constant term or the intercept (a_0) is insignificant for all the models and can be omitted implying that the concentration of metals in the leafy vegetables is totally dependent on the concentrations of metals in soil, water, and dust. The possibility of other sources of metals in the leafy vegetables is negligible. In view of the coefficients of the models, it can be suggested that each and every one of the predictor variables (metal concentrations of water, soil, and dust) did not have a significant contribution to the metal uptake of all the leafy vegetables. The concentration of metals in the mustard leaves and red amaranthus was dependant on all three matrices. However, the highest contribution is made from the metals in the irrigation water. Same is the case when the model is developed irrespective of the species. However, in the case of

Table 4 Location wise hazard index for different pathways for iron mining areas of West Singhbhum and copper mining areas of East Singhbhum

S. no.	Locations	Leafy vegetable Ingestion	Water		Soil		Dust			Total HI	
			Ingestion	Dermal contact	Ingestion	Dermal contact	Inhalation	Ingestion	Dermal contact		Inhalation
West Singhbhum (iron mining areas)											
1	Dangoaposi	3.587	0.307	0.007	0.491	0.069	0.027	0.250	0.036	0.009	5.44
2	Noamundi	2.709	0.165	0.008	0.439	0.071	0.019	0.517	0.067	0.020	4.80
3	Kotgarh	4.760	0.545	0.011	0.393	0.071	0.041	0.415	0.058	0.017	6.98
4	Bichaikiri	2.921	0.501	0.010	0.570	0.095	0.046	0.550	0.078	0.011	5.72
5	Karampada	4.496	0.189	0.008	0.465	0.084	0.015	0.363	0.053	0.010	6.39
6	Chota Nagra	3.667	0.893	0.029	0.282	0.053	0.015	0.266	0.039	0.010	5.72
	Mean	3.69	0.43	0.01	0.44	0.07	0.03	0.40	0.06	0.01	5.13
East Singhbhum (copper mining areas)											
7	Bhatin	5.361	0.895	0.043	0.136	0.021	0.005	0.283	0.041	0.009	7.12
8	Bodamdera	6.838	0.576	0.008	0.390	0.063	0.024	0.452	0.058	0.011	9.11
9	Kendadih	16.31	63.48*	0.546	0.463	0.070	0.020	0.460	0.051	0.009	18.7
10	Kutludih	7.420	3.743	0.143	0.226	0.027	0.012	0.435	0.053	0.011	12.6
11	Sohada	6.754	0.677	0.011	0.442	0.062	0.022	0.466	0.050	0.009	9.25
12	Badia	15.656	0.506	0.023	0.464	0.060	0.018	0.530	0.054	0.011	18.2
	Mean	9.72	1.28	0.13	0.36	0.05	0.02	0.44	0.05	0.01	12.1

*Not included in HI calculations

Table 5 Results and prediction precisions of the species-specific multivariate linear regression (MVLRL) models for predicting metals in leafy vegetables

S. no.	Species	Model summary							Coefficients			
		<i>n</i>	<i>F</i>	<i>p</i> value	<i>R</i> ²	ME	RMSE	Durbin Watson	<i>a</i> ₀ (constant)	<i>a</i> ₁ (water)	<i>a</i> ₂ (soil)	<i>a</i> ₃ (dust)
1	All samples	310	268.8	2.02E–85	0.7250	0.720	26.56	1.869	n.s.	0.02113	0.00128	0.00057
2	Mustard leaves	70	212.3	7.79E–34	0.9061	0.903	12.43	1.849	n.s.	0.0284	0.00058	0.00094
3	Red amaranthus	90	324.4	9.17E–47	0.9188	0.915	14.57	2.120	n.s.	0.02199	0.00079	0.00162
4	Spinach	30	669.0	9.97E–24	0.9802	0.959	3.711	2.221	n.s.	0.058	0.00032	n.s.
5	Drumstick leaves	120	541.9	6.23E–46	0.8212	0.810	26.12	1.942	n.s.	n.s.	0.00269	n.s.

n.s. not significant

spinach, the metal uptake was contributed from water and soil and the role of atmospheric dust was insignificant. Foliar deposition of metals through atmospheric dust is more pronounced in rough leaf surfaces compared to smooth ones (Schreiber and Schönherr 1992). The spinach leaves are smooth compared to red amaranthus and mustard leaves, and thus, the effect of atmospheric dust was insignificant in spinach leaves. For the drumstick leaves, the metal content in the soil was the only causative factor of the metal concentration in the leaves. The three leafy vegetables other than drumstick leaves are short life-cycled plants; however, the drumstick is perennial trees with smooth leaves which are used as leafy vegetable by the local populace. Being a perennial tree, it is not needed to be irrigated as in the case of other three species. This might be the explanation for fact that the stepwise regression model for drumstick leaves only retained the metal content in the soil as a contributing factor eliminating the water and dust factors. The same models relevant for both the mining areas prove its versatility of being useful in different environmental conditions with variable concentrations of metals in different matrices like soil, water, and atmospheric dust. However, the models need to be species-specific. Thus, the models for the prediction of metals in the 4 considered leafy vegetables are depicted as Equations 7–10.

$$\text{Metal (mustard leaves)} = 0.0284[\text{water}] + 0.00058[\text{soil}] + 0.00094[\text{dust}] \quad (7)$$

$$\text{Metal (R.amaranthus)} = 0.02199[\text{water}] + 0.00079[\text{soil}] + 0.00162[\text{dust}] \quad (8)$$

$$\text{Metal (spinach)} = 0.058[\text{water}] + 0.00032[\text{soil}] \quad (9)$$

$$\text{Metal (drumstick leaves)} = 0.00269[\text{soil}] \quad (10)$$

where [water] is concentration of metals in irrigating water; [soil] is concentration of metals in soil; and [dust] is concentration of metals in atmospheric dust;

Model validation

The validation of the models was performed using equiline scatter plot of actual and modeled values, model efficiency (ME) values, and root mean square error (RMSE) values. The scatter plot (Fig. 3) depicts that most of the points are near the equiline with some exceptions which may be considered outliers. However, the outliers have been retained in the models since the extreme values (outliers) are often encountered in the environmental studies. Nevertheless, the outliers in our study did not affect the models drastically. The plot provides a high agreement in the positive correlation between the observed and predicted values for all five models. The ME values ranged from 0.720 to 0.959 which depicted good to very good goodness of fit for the models (Cheng et al. 2017). The ratio of standard deviation of observed values to the RMSE ranged from 1.89 (considering all species) to 5.03

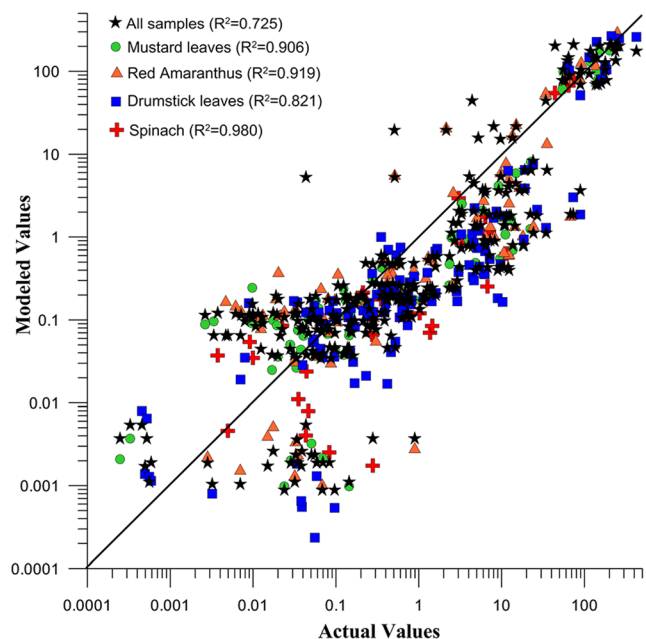


Fig. 3 Equiline scatter plot of actual and modeled values of regression models

(Spinach) which suggest that the models are acceptable to very good based on the quantification of prediction error.

Conclusions

From the assessment of metal concentrations in the leafy vegetables and other associated environmental components collected from the metal mining areas, we could understand the impact of the mining and industrial activities on the food chain and subsequently the potential risk to the human health due to the contaminated environs inclusive of food items. Health risks of metals were assessed for multiple pathways to the local populace including ingestion of leafy vegetables. The hazard quotient of the metals exceeded the reference value of unity for some metals, however, more frequently due to ingestion pathways for vegetables and water. The combined risk due to multiple metals as assessed by hazard index suggested that the ingestion of leafy vegetables was the major route of exposure of metals contributing 71.9% and 80.7% of total risk for the iron and copper mining areas, respectively. The average HI value considering all the metals and pathways was calculated to be 12.1 and 5.13, respectively for copper and iron mining areas suggesting considerable risk to the local residents.

The present study also aimed to develop species-specific multivariate regression models for predicting the concentrations of metals in leafy vegetables from their concentration in the soil, irrigating water, and atmospheric dust. Correlations between the measured and predicted metal values, high model efficiency values, and low root mean square values reflected the goodness of fit of the model. Species-specific models for the leafy vegetables were found to be better in the prediction of metals uptake from the associated environs. The models will be useful for the prediction of metal concentrations in the leafy vegetables, and thus, the possible human risks can be identified. The models would also help to understand the effect of environmental factors on the concentration of metals in the leafy vegetables. Also, the results suggest screening of irrigating water to ensure the safety of food in particular leafy vegetables with respect to metals.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11356-021-12494-9>.

Acknowledgments The authors are thankful to the Department of Science and Technology and Science and Engineering Research Board for providing the necessary funding for the study under the DST-Young Scientist Scheme (Grant No. YSS/2015/001211) and National Post Doctoral Fellowship (Grant No. PDF/2017/000953/EAS), respectively. Also, the authors are grateful to the Director and the Research group of Natural Resources and Environment Management (NREM) of CSIR-Central Institute of Mining and Fuel Research, Dhanbad, for providing the needed laboratory amenities and other logistic support to carry out the study.

Authors' contributions SG and MKM were the principal investigators of the both the projects and were involved in sampling, sample preparation, analysis, data compilation, and statistical analysis. The data handling and writing of the present manuscript were done by SG. AKS were the mentor for both the projects. He was also involved in correction of the manuscript and providing valuable inputs for further enhancement of the manuscript.

Funding The study has been funded by Department of Science and Technology and Science and Engineering Research Board under the DST-Young Scientist Scheme (Grant No. YSS/2015/001211) and National Post Doctoral Fellowship (Grant No. PDF/2017/000953/EAS), respectively.

Data availability The data is provided as tables and supplementary tables. For any other information, corresponding author may be contacted.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval and consent to participate Not applicable since the study does not involve any human subjects.

Consent for publication All of the authors have read and approved the paper.

References

- Amoakwah E, Ahsan S, Rahman MA, Asamoah E, Essumang DK, Ali M, Islam KR (2020) Assessment of heavy metal pollution of soil-water-vegetative ecosystems associated with artisanal gold mining. *Soil Sed Contamin: Inter J* 29(7):788–803
- Awasthi SK (2000) Prevention of Food Adulteration Act No. 37 of, 1954 Central and State rules as amended for 1999. Ashoka Law House, New Delhi
- Bi C, Zhou Y, Chen Z, Jia J, Bao X (2018) Heavy metals and lead isotopes in soils, road dust and leafy vegetables and health risks via vegetable consumption in the industrial areas of Shanghai, China. *Sci Total Environ* 619:1349–1357
- BIS (2006) Indian standard methods for measurement of air pollution, Bureau of Indian Standards, Part 1 Dust Fall, IS:5182.
- Cao C, Chen X, Ma Z, Jia H, Wang J (2016) Greenhouse cultivation mitigates metal-ingestion-associated health risks from vegetables in waste water irrigated agroecosystems. *Sci Total Environ* 560: 204–211
- Cheng K, Lien Y, Wu Y, Su Y (2017) On the criteria of model performance evaluation for real-time flood forecasting. *Stoch Env Res Risk A* 31(5):1123–1146
- Dang HS, Jaiswal DD, Parameswaran M, Krishnamony S (1994) Physical anatomical, physiological and metabolic data for reference Asian man—a proposal, BARC (Bhabha Atomic Research Centre), Report No. BARC/1994/FE/043.
- DeForest DK, Brix KV, Tear LM, Adams WJ (2018) Multiple linear regression models for predicting chronic aluminum toxicity to freshwater aquatic organisms and developing water quality guidelines. *Environ Toxicol Chem* 37(1):80–90
- Doabi SA, Karami M, Afyuni M, Yeganeh M (2018) Pollution and health risk assessment of heavy metals in agricultural soil, atmospheric dust and major food crops in Kermanshah province, Iran. *Ecotoxicol Environ Saf* 163:153–164

- Dunn JA, Dey AK (1942) The geology and petrology of Eastern Singhbhum and surrounding areas. *Mere Geol Surv India* 69
- Eid EM, Alrumman SA, Farahat EA, El-Bebany A (2018) Prediction models for evaluating the uptake of heavy metals by cucumbers (*Cucumis sativus* L.) grown in agricultural soils amended with sewage sludge. *Environ Monit Assess* 190:501
- Eid EM, Shaltout KH, Alamri SA, Sewelam NA, Galal TM, Brima EI (2020) Prediction models for evaluating heavy metal uptake by *Pisum sativum* L. in soil amended with sewage sludge. *J Environ Sci Health, Part A* 55(2):151–160
- El-Radaideh NM, Al-Taani AAK (2018) Geo-environmental study of heavy metals of the agricultural highway soils, NW Jordan. *Arab J Geosci* 11(24):787
- FAO (1985) Water quality for agriculture. Food and agriculture organization, irrigation and drainage paper no. 29, Rev. 1. United Nations, Rome.
- Giri S, Mahato MK, Bhattacharjee S, Singh AK (2020) Development of a new noncarcinogenic heavy metal pollution index for quality ranking of vegetable, rice, and milk. *Ecol Indic* 113:106214
- Hu W, Huang B, Tian K, Holm PE, Zhang Y (2017) Heavy metals in intensive greenhouse vegetable production systems along Yellow Sea of China: levels, transfer and health risk. *Chemosphere* 167: 82–90
- Kumar V, Thakur RK, Kumar P (2019) Assessment of heavy metals uptake by cauliflower (*Brassica oleracea* var. *botrytis*) grown in integrated industrial effluent irrigated soils: a prediction modeling study. *Sci Hortic* 257:108682
- Kumar V, Thakur RK, Kumar P (2020) Predicting heavy metals uptake by spinach (*Spinacia oleracea*) grown in integrated industrial wastewater irrigated soils of Haridwar, India. *Environ Monit Assess* 192(11):709
- Legind CN, Trapp S (2010) Comparison of prediction methods for the uptake of As, Cd and Pb in carrot and lettuce. *SAR QSAR Environ Res* 21:513–525
- Li FL, Liu CQ, Yang YG, Bi XY, Liu TZ, Zhao ZQ (2012) Natural and anthropogenic lead in soils and vegetables around Guiyang city, southwest China: a Pb isotopic approach. *Sci Total Environ* 431: 339–347
- Luo C, Liu C, Wang Y LX, Li F, Zhang G, Li X (2011) Heavy metal contamination in soils and vegetables near an e-waste processing site, south China. *J Hazard Mater* 186:481–490
- Mukherjee B (1968) Genetic significance of trace elements in certain rocks of Singhbhum, India. *Min Mag* 36(281):661–670
- NFI (2010) Bulletin of Nutrition Foundation of India (Nutrition Foundation of India). 31(1).
- Noh K, Jeong BR (2019) Particulate matter in the cultivation area may contaminate leafy vegetables with heavy metals above safe levels in Korea. *Environ Sci Pollut Res* 26(25):25762–25774
- NSSO (2014) National Sample Survey Office, Household consumption of various goods and services in India 2011–12, NSS 68th Round, Ministry of Statistics and Programme Implementation, Government of India.
- Park JH, Choi KK (2013) Risk assessment of soil, water and crops in abandoned Geumryeong mine in South Korea. *J Geochem Explor* 128:117–123
- Radojevic M, Bashkin VN (1999) Practical environmental analysis. Royal Society of Chemistry, London
- Rafiee S, Avval SHM, Mohammadi A (2010) Modeling and sensitivity analysis of energy inputs for apple production in Iran. *Energy* 35(8): 3301–3306
- Rodríguez-Estival J, Morales-Machuca C, Pareja-Carrera J, Ortiz-Santaliestra ME, Mateo R (2019) Food safety risk assessment of metal pollution in crayfish from two historical mining areas: accounting for bioavailability and cooking extractability. *Ecotoxicol Environ Saf* 185:109682
- Saha AK (1994) Crustal evolution of Singhbhum-North Orissa, eastern India: geological society of India. *Memoir* 27:341
- Schreiber L, Schönherr J (1992) Uptake of organic chemicals in conifer needles: surface adsorption and permeability of cuticles. *Environ Sci Technol* 26:153–159
- Turekian KK, Wedepohl KH (1961) Distribution of the elements in some major units of the earth's crust. *Amer Geol Soc Bull* 72:175–182
- Upadhyay RK, Venkatesh AS, Roy S (2010) Iron ore deposits of eastern India: geological aspects and their mineralogical Characteristics. *Resour Geol* 60:203–211
- USEPA (1996) EPA Method 3052: Microwave assisted acid digestion of siliceous and organically based matrices. Office of Solid Waste and Emergency Response. United States Environmental Protection Agency, Washington
- USEPA (2002) Supplemental guidance for developing soil screening levels for superfund sites.: Office of soil waste and emergency response; [OSWER 9355.4–24]. United States Environmental Protection Agency, Washington
- USEPA (2004) Risk assessment guidance for superfund (RAGS), vol. I: human health evaluation manual (Part E, supplemental guidance for dermal risk assessment). United States Environmental Protection Agency, Washington
- USEPA (2011) Exposure factors handbook edition (Final). <http://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deidD236252>, United States Environmental Protection Agency, Washington. Accessed 15 March 2020
- USEPA (2017) Update for chapter 5 of the exposure factors handbook, soil and dust ingestion, EPA/600/R-17/384F, United States Environmental Protection Agency, Washington
- Wang C, Wu S, Zhou S, Wang H, Li B, Chen H (2015) Polycyclic aromatic hydrocarbons in soils from urban to rural areas in Nanjing: concentration, source, spatial distribution, and potential human health risk. *Sci Total Environ* 527–528:375–383
- Wu S, Peng S, Zhang X, Wu D, Luo W, Zhang T, Zhou S, Yang G, Wan H, Wu L (2015) Levels and health risk assessments of heavy metals in urban soils in Dongguan, China. *J Geochem Explor* 148:71–78
- Xu L, Lu A, Wang J, Ma Z, Pan L, Feng X, Luan Y (2015) Accumulation status, sources and phytoavailability of metals in greenhouse vegetable production systems in Beijing, China. *Ecotox Environ Saf* 122: 214–220
- Yang L, Huang B, Hu W, Chen Y, Mao M, Yao L (2014) The impact of greenhouse vegetable farming duration and soil types on phytoavailability of heavy metals and their health risk in eastern China. *Chemosphere* 103:121–130

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.