metals

Article

Accurate detection of Cd^{2+} and Pb^{2+} concentrations in soils by stripping voltammetry peak areas under the mutual interference of multiple heavy metals

Wenshuai Ye[1]†, Ning Liu[2]†, Guo Zhao[3] and Gang Liu[1,2*]

†These authors share first authorship.
1 Key Lab of Smart Agriculture Systems, Ministry of Education, China Agricultural University, Beijing 100083, China
2 Key Laboratory of Agricultural Information Acquisition Technology, Ministry of Agriculture and Rural Affairs of China, China Agricultural University, Beijing 100083, China
3 College of Artificial Intelligence, Nanjing Agricultural University, Nanjing 210031, China
* Correspondence: pac@cau.edu.cn

2. Materials and Methods

2.2. Preparation of the Bi/GCE

First, the GCE was polished using alumina powder(0.05mm) on polishing paper made of deerskin. Then the GCE was sequentially cleaned by ultrasonic in 0.1 M HNO_3, 50% ethanol, and Millipore-Q water for 1 min, respectively, and dried using N_2 gas. Afterward, 300 μg/L of $Bi^{3+}$was added into the 25 mL of buffer solution (0.2 M, pH 5.0). Finally, the Bi^{3+}, together with other HMIs, were deposited into the surface of GCE at a potential of -1.3 V during the electrodeposition process, obtaining the in-situ bismuth film modified GCE (Bi/GCE).

2.3. SWASV measurement

The optimal experimental conditions for detection of Cd^{2+} and Pb^{2+} concentration using Bi/GCE by SWASV had been explored in our previous study[1], the optimal concentration of Bi^{3+} was 300 μg/L, the optimal pH value was 5.0, the optimal deposition potential was −1.3 V, and the optimal deposition time was 180 s. The sensor was characterized by CV analysis, and the specific CV plots was shown in Figure S1. The CV plots demonstrated that the electrode has considerable reversibility and can be used to detect the electrochemical signals of Zn^{2+}, Cd^{2+}, Pb^{2+}, Bi^{3+}, and Cu^{2+}. In addition, the repeatability, stability and electroanalytical performance of Bi/GCE had been validated, as presented in Fig S2, which demonstrated that the obtained SWASV dataset was reliable to build machine learning detection models.

Before performing SWASV measurements, dissolved oxygen was removed from the test solution by blowing N_2 gas for 2 mins. The main steps of SWASV were as follows. The HMIs were electro-deposited into the GCE surface for 180 s at the potential of −1.3 V, accompanied by stirring at a speed of 300 r/min. After an equilibration period for 10 s, the deposited HMIs were stripped off from the GCE surface by applying the square-wave excitation potential in the range of −1.4 to 0.2 V to obtain stripping currents. The square-wave frequency, potential amplitude, and potential increment of the excitation potential were 25 Hz, 25 mV, and 5 mV, respectively. A total of 320 data points were collected in the stripping potential range of −1.4 to 0.2 V for one SWASV data curve due to the potential increment of 5 mV.

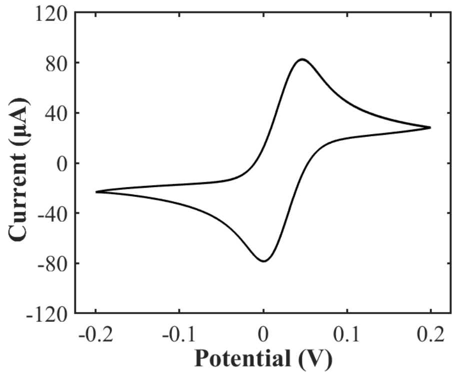

Figure S1. CV curve of the bare GCE in the 5 mM $[Fe(CN)_6]^{3-/4-}$ solution with 0.1 M KCl.

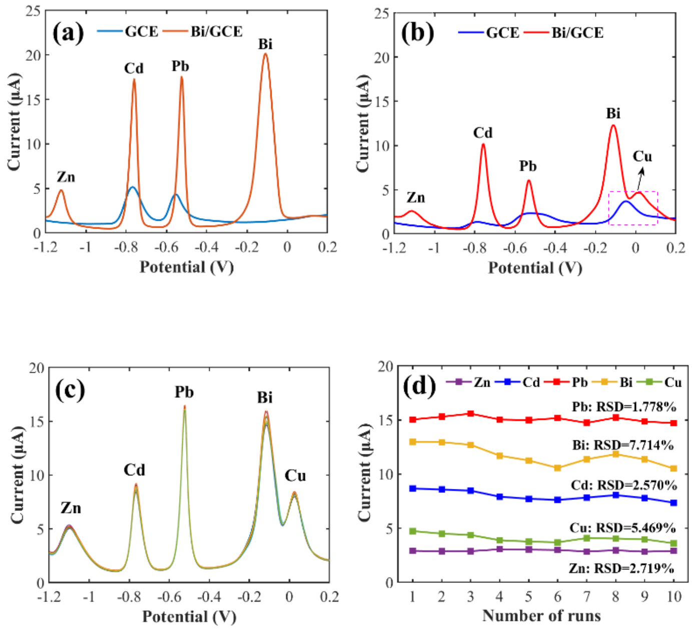

Figure S2. SWASV responses of (**a**) 20 μg/L Zn^{2+}, Cd^{2+}, and Pb^{2+} on GCE and Bi/GCE, (**b**) 20 μg/L Zn^{2+}, Cd^{2+}, Pb^{2+}, and Cu^{2+} on GCE and Bi/GCE, and (**c-d**) 40 μg/L Zn^{2+}, Cd^{2+}, Pb^{2+}, and Cu^{2+} on Bi/GCE for ten repetitive measurements.

2.4. Establishment of Experimental Datasets

Table S1. the excogitation of 64 sets of orthogonal experiments by the L64(8^4) orthogonal table.

Experiment number	**Concentrations of heavy metal ions (μg/L)**			
	Cd^{2+}	**Pb^{2+}**	**Cu^{2+}**	**Zn^{2+}**
1	2	5	0	0
2	2	25	25	50
3	2	50	50	100
4	2	75	75	150
5	2	100	100	200
6	2	125	125	250
7	2	150	150	300
8	2	175	175	350
9	5	50	0	50

10	5	75	25	0
11	5	5	50	150
12	5	25	75	100
13	5	150	100	250
14	5	175	125	200
15	5	100	150	350
16	5	125	175	300
17	10	100	0	100
18	10	125	25	150
19	10	150	50	0
20	10	175	75	50
21	10	5	100	300
22	10	25	125	350
23	10	50	150	200
24	10	75	175	250
25	15	150	0	150
26	15	175	25	100
27	15	100	50	50
28	15	125	75	0
29	15	50	100	350
30	15	75	125	300
31	15	5	150	250
32	15	25	175	200
33	20	125	0	200
34	20	100	25	250
35	20	175	50	300
36	20	150	75	350
37	20	25	100	0
38	20	5	125	50
39	20	75	150	100
40	20	50	175	150
41	25	175	0	250
42	25	150	25	200
43	25	125	50	350
44	25	100	75	300
45	25	75	100	50
46	25	50	125	0
47	25	25	150	150
48	25	5	175	100
49	30	25	0	300
50	30	5	25	350
51	30	75	50	200
52	30	50	75	250
53	30	125	100	100
54	30	100	125	150
55	30	175	150	0
56	30	150	175	50
57	35	75	0	350
58	35	50	25	300
59	35	25	50	250

60	35	5	75	200
61	35	175	100	150
62	35	75	125	100
63	35	125	150	50
64	35	100	175	0

2.6. Detection models

For multivariate regression, it was representation samples that chosen to build model was important. Based on the Euclidean distance between two samples, the K-S algorithm, a kind of unsupervised selection methods, was often used in the area of chemometrics for presenting good results, and for being easy to be applied and understood[2,3]. The process of K-S algorithm was as follows. first, the samples with the largest Euclidean distance in the sample were calculated and determined as the calibration set, and then repeatedly from the remaining samples selected the samples that had the shortest and largest distance with calibration set. Until the number of samples met the requirements of calibration set. Calculation formula (1) as follow.

$$\boldsymbol{d_x(p,q)} = \sqrt{\sum_{j=1}^{N}\left[x_p(j) - x_q(j)\right]^2} \quad \mathbf{p}, \mathbf{q} \in [\mathbf{1}, \mathbf{N}] \tag{S1}$$

The SPXY algorithm, a kind of supervised selection, required all samples to be labeled, sequentially, the information provided by which could efficiently select representative samples[4]. The SPXY algorithm was developed on the basis of K-S algorithm, and the difference from K-S algorithm was that SPXY algorithm took both x (five peak heights or peak areas) and y (the real concentrations of target ions) variables into account when calculating the distance between samples. Calculation formula (2,3) as follow[5]. Therefore, SPXY were used to efficient select representative samples to accurately detective Cd^{2+} and Pb^{2+} concentration under Zn^{2+} and Cu^{2+} interference.

$$\boldsymbol{d_y(p,q)} = \sqrt{(y_p - y_q)^2} \quad \mathbf{p}, \mathbf{q} \in [\mathbf{1}, \mathbf{N}] \tag{S2}$$

$$\boldsymbol{d_{xy}(p,q)} = \frac{d_x(p,q)}{\max\limits_{\mathbf{p},\mathbf{q}\in[\mathbf{1}.\mathbf{N}]} d_x(p,q)} + \frac{d_y(p,q)}{\max\limits_{\mathbf{p},\mathbf{q}\in[\mathbf{1}.\mathbf{N}]} d_y(p,q)} \quad \mathbf{p}, \mathbf{q} \in [\mathbf{1}, \mathbf{N}] \tag{S3}$$

2.8. Preparation of soil extract samples

The soil extract solution was prepared as follows. 1 g of dried soil was mixed with 40 mL of acetate buffer (0.2 M, pH 5.0) in a glass bottle. The soil suspension was shaken in an end-over-end shaker at room temperature for 24 h. The mixture was centrifuged at a rotation speed of 3000 r/min for 20 min to achieve solid-liquid separation. The super natant was subjected to ultraviolet photolysis for 10 min in a photolysis tank, obtaining the soil extract solution. The 10 mL of soil extract was transferred into a beaker to conduct SWASV measurement after blowing N2 gas for 2 min.

Table S2. Detailed information on real soil samples.

Soil sample	**Total contents of heavy metals (mg/kg)**				**Granularity**	**Soil pH**	**Sampling location**
	Cd	**Pb**	**Cu**	**Zn**			
Sample 1	**0.160**	**20.100**	23.300	64.200	<0.099 mm	8.450	Luochuan city, Shaanxi province
Sample 2	0.272	47.400	36.500	90.400	<0.074 mm	6.620	Suzhuo city, Jiangsu province

3. Results and discussion

3.1. Analysis of SWASV Signals and Input Variables

3.1.2. Comparison of Peak Heights and Peak Areas of Multiple HMIs Acquired by Different Methods

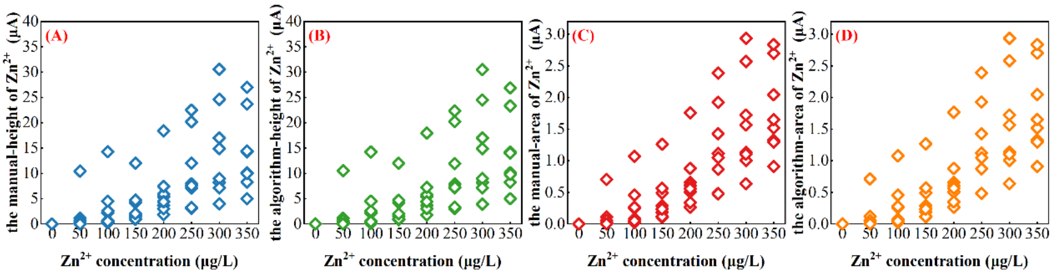

Figure S3. I_{manu} (A), I_{algo} (B), A_{manu} (C) and A_{algo} (D) of Zn^{2+} acquired from the SWASV curves of 64 experiments. (Blue scatters represent the I_{manu}, green scatters represent the I_{algo}, red scatters represent the A_{manu}, orange scatters represent the A_{algo}).

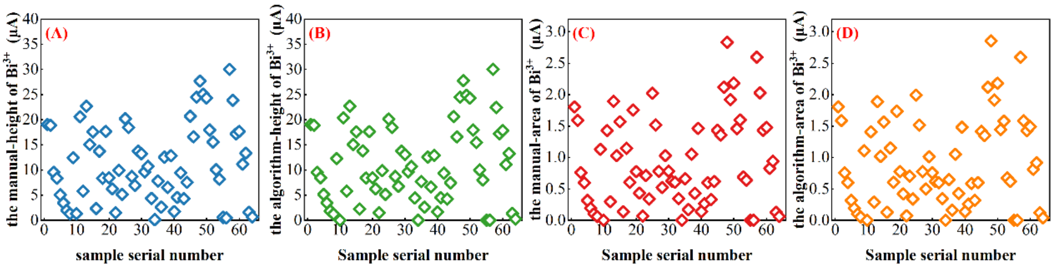

Figure S4. I_{manu} (A), I_{algo} (B), A_{manu} (C) and A_{algo} (D) of Bi^{3+} acquired from the SWASV curves of 64 experiments. (Blue scatters represent the I_{manu}, green scatters represent the I_{algo}, red scatters represent the A_{manu}, orange scatters represent the A_{algo}).

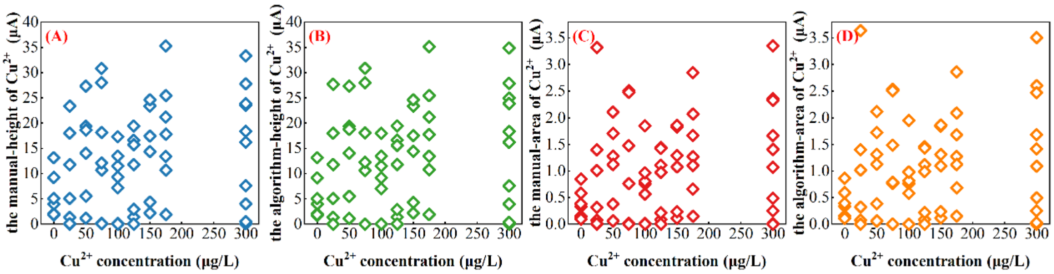

Figure S5. I_{manu} (A), I_{algo} (B), A_{manu} (C) and A_{algo} (D) of Cu^{2+} acquired from the SWASV curves of 64 experiments. (Blue scatters represent the I_{manu}, green scatters represent the I_{algo}, red scatters represent the A_{manu}, orange scatters represent the A_{algo}).

Table S3. Detailed R_θ^2 and $\cos\theta$ of the peak heights and peak areas

.HMI	peak heights		peak areas	
	R_θ^2	$\cos\theta$	R_θ^2	$\cos\theta$
Zn^{2+}	0.9999	1.0000	1.0000	1.0000
Cd^{2+}	1.0000	1.0000	1.0000	1.0000
Pb^{2+}	0.9998	0.9999	0.9999	1.0000
Bi^{3+}	0.9995	0.9998	0.9970	0.9989
Cu^{2+}	0.9984	0.9993	0.9989	0.9995

3.2. Detection of Cd^{2+} and Pb^{2+} Concentrations by Peak Heights

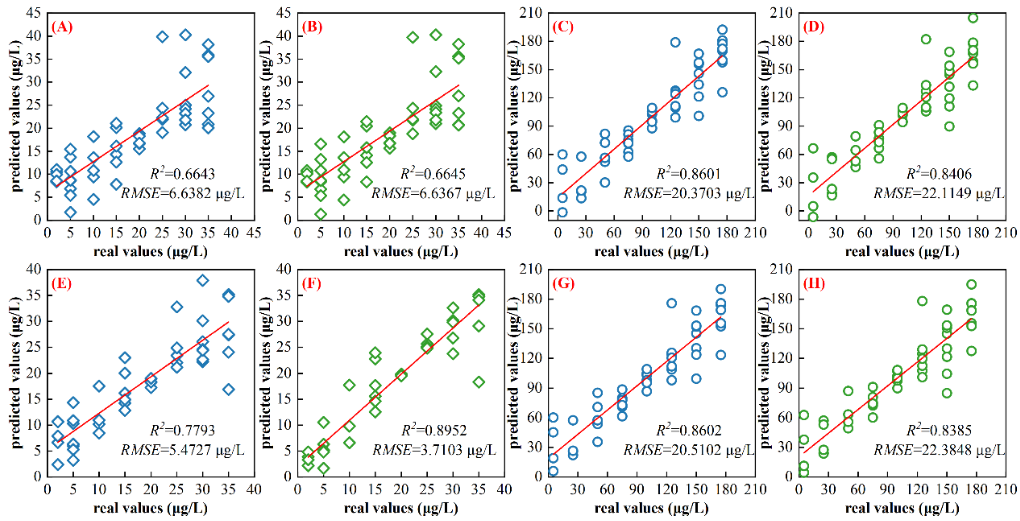

Figure S6. PLSR models (A, B, C and D) and SVR models (E, F, G and H) results for Cd^{2+} and Pb^{2+} concentrations detection in the calibration dataset. (Diamond scatters represent Cd^{2+} concentrations, circular scatters represent Pb^{2+} concentrations, blue scatters represent the model input as I_{manu}, green scatters represent the model input as I_{algo}.).

Table S4. The R^2 values and RMSE values of the models built by peak heights.

machine learning model	HMIs	input	calibration set		validation set		parameters		
			RMSE (μg/L)	R^2	RMSE (μg/L)	R^2	c	g	mse
PLSR	Cd^{2+}	I_{manu}	6.6382	0.6643	5.4463	0.7375	-	-	-
		I_{algo}	6.6367	0.6645	5.4838	0.7323	-	-	-
	Pb^{2+}	I_{manu}	20.3703	0.8601	19.2392	0.8645	-	-	-
		I_{algo}	22.1149	0.8406	21.9047	0.7766	-	-	-
SVR	Cd^{2+}	I_{manu}	5.4727	0.7793	5.3916	0.7650	9.0970	0.0380	0.1857
		I_{algo}	3.7103	0.8952	3.9922	0.8424	9.3231	0.1977	0.1260
	Pb^{2+}	I_{manu}	20.5102	0.8602	20.0015	0.8791	343.0406	0.0010	0.0605
		I_{algo}	22.3848	0.8385	21.3431	0.8229	239.3599	0.0010	0.0762

3.3. Detection of Cd^{2+} and Pb^{2+} Concentrations by Peak Areas

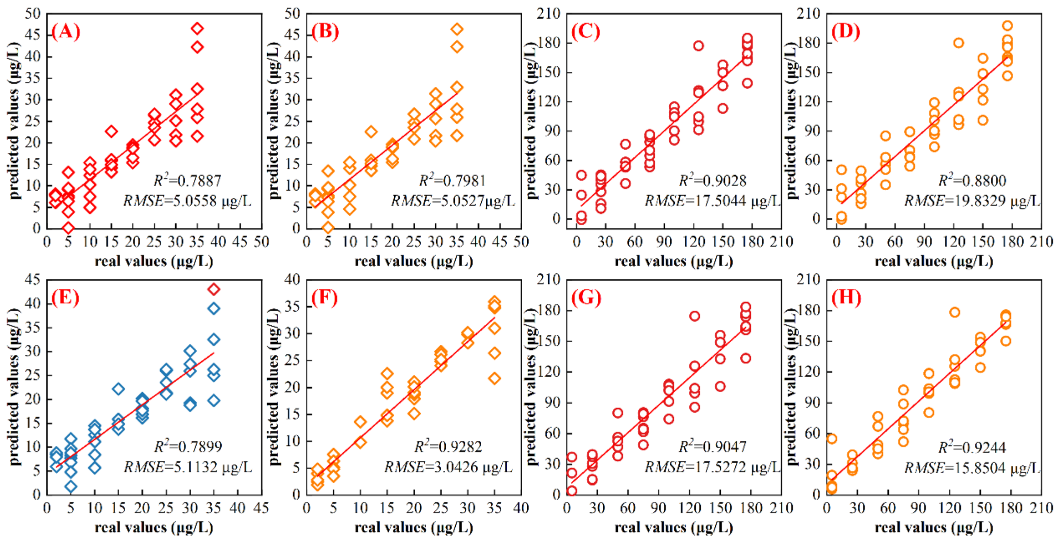

Figure S7. PLSR models (A, B, C and D) and SVR models (E, F, G and H) results for Cd^{2+} and Pb^{2+} concentrations detection in the calibration dataset. (Diamond scatters represent Cd^{2+} concentrations, circular scatters represent Pb^{2+} concentrations, red scatters represent the model input as A_{manu}, orange scatters represent A_{algo} as the model input.).

Table S5. The R^2 values and RMSE values of the models.

machine learning model	HMIs	input	calibration set RMSE (μg/L)	calibration set R^2	validation set RMSE (μg/L)	validation set R^2	parameters c	parameters g	parameters mse
PLSR	Cd^{2+}	I_{manu}	5.0558	0.7887	3.8996	0.8828	-	-	-
		I_{algo}	5.0527	0.7981	3.7414	0.891	-	-	-
	Pb^{2+}	I_{manu}	17.5044	0.9028	15.1333	0.9272	-	-	-
		I_{algo}	19.8329	0.8800	12.9264	0.9668	-	-	-
SVR	Cd^{2+}	I_{manu}	5.1132	0.7899	3.9824	0.8785	400.0000	0.0010	0.1312
		I_{algo}	3.0426	0.9282	2.9906	0.9204	18.5183	0.1254	0.0795
	Pb^{2+}	I_{manu}	17.5272	0.9047	13.3444	0.9359	1000.0000	0.0010	0.0503
		I_{algo}	15.8504	0.9244	13.1574	0.9756	1.1370	0.1736	0.0655

3.5. Application of A_{algo}-SVR for Detecting Cd^{2+} and Pb^{2+} Concentrations in Soil Extract Solution

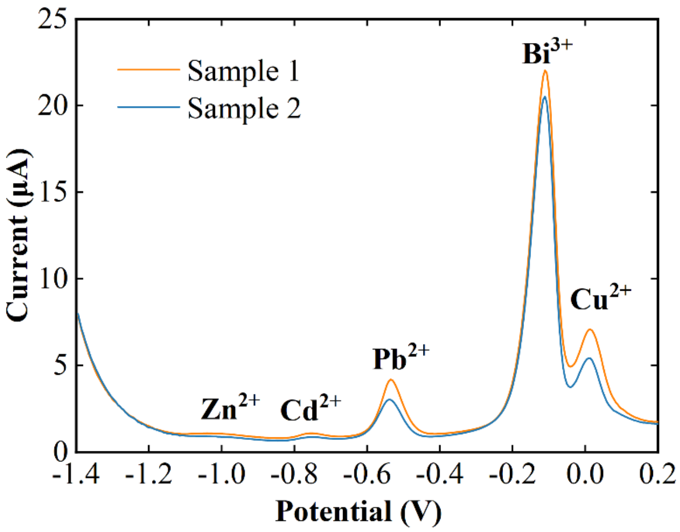

Figure S8. SWASV response curves of soil extract samples.

References:

1. Liu, N., Ye, W., Liu, G., Zhao, G., 2022a. Improving the accuracy of stripping voltammetry detection of cd2+ and pb2+ in the presence of cu2+ and zn2+ by machine learning: understanding and inhibiting the interactive interference among multiple heavy metals. Anal. Chim. Acta 1213, 339956.
2. Li, X., Fan, P., Li, Z., Chen, G., Qiu, H., Hou, G., 2021. Soil classification based on deep learning algorithm and visible near-infrared spectroscopy. J. Spectrosc. 2021, 1-11.
3. Ferreira, R.D.A., Teixeira, G., Peternelli, L.A., 2022. Kennard-stone method outperforms the random sampling in the selection of calibration samples in snps and nir data. Ciência Rural 52(5).
4. He, Z., Song, S., Shen, K., Zhang, X., 2022. Performance enhancement - based active learning sample selection method. J. Chemometr. 36(3).
5. Wang, S.F., Han, P., Cui, G.L., Wang, D., Liu, S.S., Zhao, Y., 2019. The nir detection research of soluble solid content in watermelon based on spxy algorithm. Spectrosc. Spect. Anal. 39(3), 738-742.