CHEMICAL ENGINEERING TRANSACTIONS

VOL. 67, 2018

A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.aidic.it/cet

Guest Editors: Valerio Cozzani, Bruno Fabiano, Davide Manca

ISBN 978-88-95608-64-8; **ISSN** 2283-9216

DOI: 10.3303/CET1867109

Land Metal Pollution and Improvement Based on SVM Vector Machine

Cunan Su[a], Lanhui Su[b], Bo Wu[b], Jing Wang[b]

[a]School of Hydraulic Engineering, Changsha University of Science &Technology, Changsha 410114, China
[b]Hunan Guodi Water Conservancy and Hydropower Design Consulting Company, Changsha 410114, China
cunnansu1283@21cn.com

To sample, analyze and measure the heavy metal content in the soil, a comprehensive evaluation and analysis is conducted on the heavy metal content in the soil of the region. Then, the evaluation results is analyzed and compared with the help of scientific methods to provide scientific basis for the treatment of heavy metal pollution in the soil. Methods: The Hefei City is taken as an example and the soil sample from the surrounding urban areas of Huaihua, Hubei Province are selected to determine the heavy metal content, such as Ni, As, Cu, in the soil. Then, the svm vector machine is used for the comprehensive evaluation of the heavy metal pollution in the soil, and at the same time, the contrastive analysis is conducted on the results combined with the Nemerow integrated pollution index method and the Hakanson method. Results: Through the application of svm vector machine, it is finally found that the variation coefficient of Hg in the soil is 0.648, which is the highest among all heavy metals while the variation coefficient of Cr is the lowest, namely 54.7 mk·kg-1, indicating that the variation coefficient of heavy metal content in the soil in this region is large. Conclusion: Compared with the Nemerow integrated pollution index method and the Hakanson method, the results obtained by the svm vector machine evaluation method are more accurate. Especially considering the multiple influencing factors, the evaluation results are more authentic and more reliable suing the latter method.

1. Introduction

In recent years, with the development of social economy and the acceleration of urbanization, the problems faced in the process of modern urban development are also increasing, such as population growth pressure, traffic congestion pressure and economic development pressure. Especially the environmental problems that have arisen during the economic development and expansion process of cities have become increasingly prominent and the worsening environment has exerted great impact on the production and life of urban residents. Although domestic scholars have made many contributions to the research on environmental protection, these researches are basically aimed at water pollution and atmospheric pollution. The research concerning the heavy metal pollution in urban land needs to be improved. At present, the research on this aspect mainly includes the evaluation method, influencing factor and existence form of soil pollution mainly related to this research in China are involved. The method used has certain flaws and is easily subject to subjective factors.

In order to better improve the research quality of metal pollution in the soil, this paper proceeds from the svm vector machine and is based on strict mathematics. At the same time, with the assistance of many small sample data, this paper takes the soil around the urban area of Huaihua City as an example and conducts statistical analysis of the heavy metal content in the soil in this area through sampling, analysis and determination. Moreover, combined with the Nemerow integrated pollution index method and the Hakanson method, the pollution evaluation results of the svm method are comprehensively compared to find out the differences, advantages and disadvantages of different evaluation methods and to provide a more scientific basis for the improvement of heavy metal pollution.

Please cite this article as: Su C., Su L., Wu B., Wang J., 2018, Land metal pollution and improvement based on svm vector machine, Chemical Engineering Transactions, 67, 649-654 DOI: 10.3303/CET1867109

2. Literature review

Contaminated soil refers to a soil that is exposed to corrosive media and changes its original physicochemical properties. The corrosive medium here mainly includes waste gas, waste liquid, waste residue generated in industrial production process, and organic pesticides used in agricultural production (Barzegar et al., 2018). Generally speaking, it can be divided into two categories: heavy metal contaminated soil and organic pollutant soil. Mining, industrial waste accumulation, and wastewater discharge often cause excessive levels of heavy metals in the foundation soil. Organic pollution sources mainly include gas stations, airport aviation engine oils and some chemical factories. From the appearance, the color and state of the contaminated soil are often different from the uncontaminated soil. It is mostly brown, black and has an odor. The state is mostly soft plastic or flow plastic (Fan et al., 2017). The soil structure is mostly honeycomb and loose, with poor stability and low bearing capacity. Complex physical and chemical reactions often occur inside the soil, which will cause structural damage to the soil and cause a series of engineering construction problems such as uneven settlement of buildings. From the field of environmental protection and engineering construction, the classification of contaminated soil is the basis for its improvement and restoration. Classification criteria based on risk assessment of contaminated soil sites are also widely used. European and American countries have proposed different classification criteria for contaminated sites. However, the international environmental geotechnical engineering community has not yet formed a relatively uniform and perfect classification standard for contaminated soil engineering.

Contaminants can change the original physical and chemical properties of the soil, thus changing the engineering environment characteristics of the foundation soil and its buried facilities. At present, the focus of research by scholars at home and abroad is the impact of acid and alkali and organic matter. Research on heavy metal contaminated soils is still relatively rare. Heavy metals generally refer to metals having a specific gravity (density) greater than 5, and there are about 45 species. Typical heavy metals that are harmful to the human body include lead, mercury, chromium, arsenic, and cadmium. These heavy metals cannot be decomposed in water and are extremely toxic. It combines with other toxins in the water to produce more toxic organic or inorganic substances, such as cadmium and lead in the combined state, and elemental mercury (Li et al., 2017). It is estimated that the amount of Cd, Cu, Pb, Ni, and Zn produced by the Earth has reached 500,000 tons, 307 million tons, 241 million tons, 17 million tons, and 250 million tons, respectively. The total amount of these metals diffused through the atmosphere is about 320,000 tons, 2.2 million tons, 20 million tons, 1 million tons, and 14 million tons, respectively. At present, the world emits about 3.4 million tons of Cu per year, about 15,000 tons of Hg, about 5 million tons of Pb, about 1 million tons of Ni, and about 15 million tons of Mn. According to the National Sewage Irrigation District Survey of the Ministry of Agriculture, in the 1.4 million hectares of sewage irrigation areas, the area of land contaminated by heavy metals accounts for 64.8% of the area of sewage irrigation areas. Among them, 46.7% were mildly polluted, 9.7% were moderately polluted, and 8.4% were heavily polluted (Liu et al., 2016). These heavy metal pollutants accumulate deposits on the surface and migrate with surface water and groundwater systems. Global land faces a serious threat of heavy metal pollution. Heavy metal ions will affect the environment with the movement of underground water, harm the growth and development of animals and plants, affect the quality of crops, and further harm the survival and safety of human beings through the amplification of biologic chain. For example, from 1953 to 1956, Minamata City, Kumamoto Prefecture, Japan, caused water rickets due to the discharge of mercury-containing wastewater from petrochemical plants. In 2008, 73 villagers near a chemical plant in Huaihua City, Hunan Province suffered from varying degrees of arsenic poisoning. In the second half of 2009, more than 200 children in Dongchuan District of Kunming City, Yunnan Province were diagnosed with excessive lead. At the end of 2010, some children in Gaohe Town, Huaining County, Anqing City, Anhui Province were found to have excessive lead levels (Meng et al., 2017).

Heavy metal ion contaminated foundations may also cause damage to buildings due to changes in soil properties. For example, in the 1990s, electrolytic Zn occurred in Liuzhou City, and the red clay gene in the workshop was contaminated by $ZnSO_4$ and the strength was significantly reduced, which led to the shutdown of the workshop. In the 1960s, the Nanjing Investigation Company of the Ministry of Chemical Industry discovered that the foundation soil was contaminated during the renovation of some old factories, causing damage to buildings. In addition, urban underground structures and deep buried equipment are increasing. Toxic heavy metal ions can reduce the stability and durability of such engineering structures and underground facilities to affect project life (Rodriguez-Galiano et al., 2015). To prevent soil pollution, protect the ecological environment, ensure agricultural production, and maintain human health, the corresponding soil environmental quality safety standards have been established. China's soil environmental quality standards are divided into three levels. The primary standard is the limit value of the soil quality that protects the natural ecology of the region and maintains the natural background. The secondary standard is the soil limit value for ensuring agricultural production and maintaining human health. The third-level standard is the soil critical value for ensuring agricultural and forestry production and normal plant growth. According to this, the soil environmental quality standard value is divided into three grades. Foreign soil environmental quality standards are more detailed, which emphasizes the use of contaminated sites (Singh et al., 2014). The limits for heavy metal safety standards specified in the laws and regulations issued by governments are only a reference value. The migration characteristics and toxicity of heavy metals are affected by soil properties (such as soil pH, mineral composition, clay content, organic matter, and salts). Therefore, the above effects should be considered when evaluating the actual safety performance of the soil. For potential contaminated sites, on-site field surveys and sampling were carried out and the composition of the soil was analyzed. The soil is classified and evaluated. The distribution, migration characteristics and related geological parameters of pollutants are clarified. A targeted contaminated site restoration program was proposed.

In recent years, domestic scholars have carried out a series of exploration and research on the harmfulness of heavy metal pollution components to the soil, and achieved certain results. According to the research on adsorption equilibrium test, although red clay has certain purification ability to Zn^{2+}, when Zn^{2+} content exceeds its adsorption capacity, excess Zn^{2+} will crystallize and precipitate in soil cracks and gaps. Due to the expansion and wedging effect, the soil particle gap increases, the soil structure is destroyed, and the strength of the foundation soil is reduced, which jeopardizes the building (Tan et al., 2018). The change law of the physico-mechanical properties of cohesive soil in the late stage of adsorption of inorganic ions was studied. The permeability of cohesive soil increases with the increase of ion content in the soil, while the strength decreases with the increase of ion content.

In summary, in the exploration of heavy metal contaminated soil reaction mechanism and repair technology, there is still a big gap between domestic scholars and foreign scholars. It is still in its infancy and lacks systematic theoretical understanding. In terms of the treatment of contaminated sites, the classification and identification of the contaminated sites and the evaluation criteria for the survey are unclear, and the treatment and repair techniques are immature. Developed countries in Europe and America have carried out a lot of research work on the repair technology and basic properties of heavy metal contaminated soil, and achieved good results. Based on the above research status, the content of heavy metals in soil is sampled, analyzed and measured. A comprehensive evaluation and analysis is conducted on the heavy metal content in the soil of the region. Then, the evaluation results are analyzed and compared with the help of scientific methods to provide scientific basis for the treatment of heavy metal pollution in the soil.

3. Research methods

3.1 Sampling selection of research objects

In this paper, Hecheng District, Zhongfang County and Hongjiang City of Huaihua City are taken as the main study areas and the total area of these three administrative regions is about 867 km^2. Among them, the sampling site in Hecheng District is in the irrigated area and irrigation type of the farmland is diversion irrigation. The sampling site in Zhongfang County is located in areas where heavy chemical enterprises are widely distributed and the land-use type is mainly forest land, grassland and cultivated land; the sampling site in Hongjiang City is farmland and the farmland is mainly irrigated by sewage in the ditch.

3.2 Sample collection

In accordance with the sampling design, combined with the current situation of land use, traffic conditions and pollution sources during the sampling, the principle of “randomness”, “equivalence” and “integration of multiple points” shall be followed Samples shall be taken along certain route. The sampling should be conducted using the plum-shaped spot and the special parts such as roadsides, field ridges, trenches and heaps of fertilizers should be avoided. After positioning with GPS, soil samples are collected using a quaternary method after collecting 10 points of 0-20 cm topsoil, which is the sample in this sampling point. Altogether a total of 80 samples are collected. After the registration and numbering, the soil samples are packaged in a PVC plastic bag. After natural withering in the laboratory, the impurities are extracted and the sample is ground. After the screening of 200 mesh, the sample is blended, waiting to be tested.

3.3 Measurement of research samples

The heavy metal elements in the soil include Cu, Zn, Pb, Cr, Ni, Cd, As and Hg, among which the measurement of heavy metal elements of Cu, Zn, Pb, Cr and Ni uses the ZSX Primu II X-ray fluorescence spectrometer which takes the gas flow gas proportional counter (F- PC) as the detector; the measurement of Cd element content uses the more common graphite furnace atomic absorption spectrophotometry; the primary color measurement of As, Hg uses the XDY-2A dual-channel atomic fluorescence spectrometer with good sensitivity and reproducibility produced by Beijing Geometer and the detection method is reducing gasification - atomic fluorescence spectrometry. The local land and resources testing center has measured the content of 8 heavy metal elements in 80 samples and the detection limit and precision of the test method are obtained through the experiment, as is shown in Table 1.

Table 1: Soil heavy metals analysis limits of detection and precision

Element	The detection limit	Precision (%)
Cd	0.029	6.64
As	0.42	3.09
Cu	1	5.54
Ni	2	3.73
Pb	2	5.01
Cr	5	3.25
Hg	0.003	1.80
Zn	2	1.62

4. Evaluation method of heavy metal pollution in the soil

4.1 Traditional evaluation methods

In this paper, the Nemerow untegrated pollution index and Hakanson index are used to evaluate the level of eight kinds of heavy metal pollution. The expression of Nemerow untegrated pollution index method is:

$$P_i = \sqrt{\frac{(C_i IS_i)_{\max}^2 + (C_i IS_i)_{av}^2}{2}} \tag{1}$$

Among them, Pi is the integrated pollution index, and (Ci/Si)max is the maximum value of the pollution index of each heavy metal element; (Ci/Si)av is the arithmetic mean value of the pollution index of each heavy metal element; Ci is the actual measured value of each heavy metal element; and Si is the reference value of each heavy metal element. In this paper, the local soil element background value is used as the reference value, as is shown in Table 2:

Table 2: Soil heavy metal pollution evaluation criteria

Rating level	Pb	Cu	Ni	As	Cd	Cr	Hg	Zn
2A	≤40.26	≤27.33	≤30.40	≤8.39	≤0.196	≤72.8	≤0.147	≤79.4
2B	40.26	27.33	30.40	8.39	0.196	72.8	0.147	79.4
2C	≥63.11	≥37.47	≥40.30	≥13.37	≥0.296	≥90.9	≥0.269	≥100.2
Background values	14.7	22.9	29.9	9.1	0.102	55.3	0.023	63.5
Standard value	80	100	100	25	0.8	250	1.5	300

The evaluation criteria of each element in the above table are used as the basis and then different evaluation criteria are brought into formula (1), obtaining the evaluation grade based on the Nemero integrated pollution index. The specifics are shown in Table 3:

Table 3: The specifics

Nemerow Pollution Index	Rating level
≤8.61	2A
8.61-12.47	2B
≥12.47	2C

The calculation process of the Hakanson index method is simple, considering the damage degree of each pollution factor to give weight, which can truly reflect the actual pollution situation. The specific implementation method is as follows:

$$RI = \sum E_i = \sum T_i (C_s^i / C_n^i) \tag{2}$$

The Csi is the actual measured value of the heavy metal i and Cni is the reference value of the heavy metal i. In this paper, the background value of the local soil element is selected. Ti refers to the risk coefficient of the heavy metal i and the toxicity coefficient of different elements is 5, 5, 2, 10, 30, 2, 40 and 1 respectively. Ei is the potential ecological risk factor for each heavy metal element and the grading standard is shown in Table 4:

Table 4: Potential ecological risk index and rating

Potential ecological risk index	Rating level
≤150	2A
150-350	2B
≥250	2C

4.2 Support vector machine

The SVM method converts the indivisible variables in the low-dimensional space into a high-dimensional space through the operation on the kernel function, thereby obtaining an optimal classifying hyperplane and then conducting the classification. When the two-dimensional space is linearly separable, it not only requires to classify the samples without errors, but also achieves the straight line with the largest interval of the classification between the two. Only when achieving these two points can the empirical risk and the confidence risk reach the minimum, so that the classification is accurate. This is called the optimal classification line in the SVM, as is shown in Figure 1:

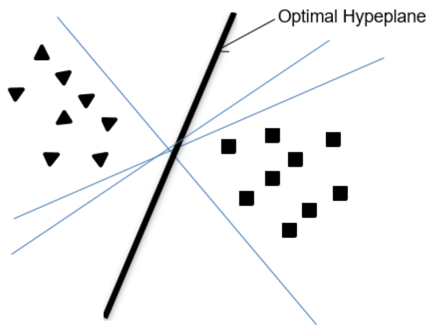

Figure 1: Optimal classification hyperplane

For the specific implementation process of this paper, firstly, the training and test samples are obtained according to evaluation criteria interpolation in Matlab. After that, the training samples are trained in Matlab and the cross-validation method should be used in the training process to select the penalty parameter c and the kernel function g. The test sample is taken to test the accuracy of the trained SVM model. When the accuracy of the model reaches the maximum, the measured data is classified using the trained model.

5. Analysis of results

5.1 Analysis statistical results

Table 5: Statistical analysis of measured values of soil elements

Element	Number	MIN	MAX	Average	Deviation	Coefficient of variation
Pb	80	17.4	86.0	26.29	10.75	0.409
Cu	80	17.2	47.6	28.87	5.31	0.184
Ni	80	20.5	50.2	29.76	5.13	0.172
As	80	4.5	17.8	10.96	2.83	0.258
Cd	80	0.1	0.4	0.21	0.05	0.248
Cr	80	54.7	109.0	73.69	9.28	0.126
Hg	80	0.02	0.39	0.12	0.08	0.648
Zn	80	58.6	121.0	86.08	11.76	0.137

Table 5 shows the measured results of the heavy metal elements obtained from the statistical analysis.
It can be seen from the above table that the variation coefficient of heavy metal elements in the soil is Hg, Pb, As, Cd, Cu, Ni, Zn, Cr from high to low. The Hg content is between 0.02 and 0.39; the variation coefficient is 0.648; the maximum value is 19.5 times the minimum value. The Pb content in the soil is between 17.4 and 86; the variation coefficient is 0.409. The maximum value of Cr element with the smallest variation coefficient is 109 and the minimum value is 54.7. Thus, it can be seen that the variation of each element in the soil in the area is significant.

5.2 Comparison of evaluation results SVM vector machine with traditional methods

The evaluation results of the support vector machine are compared with two traditional integrated evaluation methods and the results of the Hankson index are compared with the SVM. There are 28 samples with differences in the results. Then, the actual measurement values of heavy metal elements are analyzed and the evaluation results in five soil samples are more accurate with the Hankson index while the evaluation results in the remaining 23 sample are more accurate with the SVM.

6. Conclusion

6.1 Research conclusions

It can be seen from the comparison of the evaluation results that the sample number of 2A, 2B, and 2C in the evaluation results of Nemerow integrated pollution index method, Hakanson's index method and SVM are: 41, 47 and 45; 37, 29 and 33; 2, 4 and 2 respectively. The evaluation results are significantly different, but it can be discovered form the comparison of the evaluation results of each sample point that compared with SVM, the Nemerow integrated pollution index method and Hakanson's index method have 22 and 28 samples with different evaluation results and that the same rate of evaluation results reaches 70% and 65%. The generalization ability of SVM method is better, which requires fewer number of parameters to be set. The classification surface is simple and the fitting accuracy is high, which can analyze inconsistent results.

6.2 Outlook

It can be found through the analysis of three evaluation methods that compared with traditional methods, the support vector machine reduces the impact of human subjective judgment on the results and that the evaluation results are more in line with the actual situation. The support vector machine has a strict mathematical basis, good generalization ability, simple implementation process and thus has a wide application prospect in the evaluation of heavy metal pollution in the soil.

References

Barzegar R., Moghaddam A.A., Deo R., Fijani E., Tziritis E., 2018, Mapping groundwater contamination risk of multiple aquifers using multi-model ensemble of machine learning algorithms, Science of the Total Environment, 621, 697-712, DOI: 10.1016/j.scitotenv.2017.11.185

Fan J., Wu J., Kong W., Zhang Y., Li M., Zhang Y., Meng W., 2017, Predicting Bio-indicators of Aquatic Ecosystems Using the Support Vector Machine Model in the Taizi River, China, Sustainability, 9(6), 892, DOI: 10.3390/su9060892

Li H., Wang J., Wang Q.G., Tian C., Qian X., Leng X.Z., 2017, Magnetic properties as a proxy for predicting fine-particle-bound heavy metals in a support vector machine approach, Environmental science & technology, 51(12), 6927-6935, DOI: 10.1021/acs.est.7b00729

Liu Y., Wang H., Zhang H., Liber K., 2016, A comprehensive support vector machine-based classification model for soil quality assessment, Soil and Tillage Research, 155, 19-26, DOI: 10.1016/j.still.2015.07.006

Meng D., Zhao N., Wang Y., Ma M., Fang L., Gu Y., Liu J., 2017, On-line/on-site analysis of heavy metals in water and soils by laser induced breakdown spectroscopy, Spectrochimica Acta Part B: Atomic Spectroscopy, 137, 39-45, DOI: 10.1016/j.sab.2017.09.011

Rodriguez-Galiano V., Sanchez-Castillo M., Chica-Olmo M., Chica-Rivas M., 2015, Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines, Ore Geology Reviews, 71, 804-818, DOI: 10.1016/j.oregeorev.2015.01.001

Singh S.K., Srivastava P.K., Gupta M., Thakur J.K., Mukherjee S., 2014, Appraisal of land use/land cover of mangrove forest ecosystem using support vector machine, Environmental earth sciences, 71(5), 2245-2255, DOI: 10.1007/s12665-013-2628-0

Tan K., Wang H., Zhang Q., Jia X., 2018, An improved estimation model for soil heavy metal (loid) concentration retrieval in mining areas using reflectance spectroscopy, Journal of Soils and Sediments, 18(5), 2008-2022, DOI: 10.1007/s11368-018-1930-6