Journal of University of Science and Technology Beijing
1996, Vol 3, No 1 : 21 ~ 25

CORRELATION AMONG CORROSIVE FACTORS OF SOIL

LI CHANGRONG [1] QU ZUYU [1] WANG GUANGRONG [1]
CHEN JIANMIN [2] DING YI [2] LI CHUNGUANG [2]

1) Department of Surface Science and Corrosion Engineering, USTB, Beijing 100083
2) East China Administrative Bureau of Oil Transferring, Xuzhou 221008

(Received 1995-12-15)

ABSTRACT In this paper, through the data treatment of correlation coefficients and the clustering technique of pattern recognition, both the matrix of correlation coefficients and the pedigree of factor clusters show that some factors affecting soil corrosivity are interrelated closely, and some independent relatively. The data analytical method used in this paper has certain assistance to the selection of major soil factors during the prediction and evaluation of soil corrosivity.

KEY WORDS clustering analysis, corrosive factors, soil corrosion

The behavior of materials in soil corrosion depends on many soil factors. The correlation between the soil corrosivity and any single physical or chemical factor of soil is usually poor. Various factors must be considered. The German GW9 specification on the prediction of soil corrosivity considers twelve factors, including soil variety, soil condition, electric resistivity, water content, pH and so on [1]. The exploration survey of soil environment for the Chinese national soil corrosion network involves more than twenty factors, including physical, chemical and electrochemical properties [2].

In this paper, the data from an exploration report on an soil corrosion environment are analyzed by means of the multi-variable analytical method of mathematical statistics. The correlation coefficients between various soil factors are obtained through data processing. In addition, the clustering technique of pattern recognition is introduced to produce the pedigree of factor clusters. The purpose of this analysis is to go further into the correlation among the corrosive factors themselves in order to help the selection of major soil factors during the prediction and evaluation of soil corrosivity.

1 CORRELATION COEFFICIENTS BETWEEN SOIL CORROSIVE FACTORS

1.1 SELECTION OF CORROSIVE FACTORS

According to Reference [2] and with the full consideration of all the soil corrosive factors, the following 22 items are taken into account during the exploration survey of soil environment.

(1) Electrochemical properties : resistivity, conductivity, on-site pH, indoor pH, redox potential, steel/soil potential, pipe/soil potential.

(2) Physical properties: water capacity volume density, void volume, air capacity.

(3) Chemical properties: total salts, CO_3^{2-}, HCO_3^-, Cl^-, SO_4^{2-}, NO_3^-, Na^+, Mg^{2+}, Ca^{2+}, organic substances, cation exchange.

1.2 CORRELATION BETWEEN SOIL FACTORS

The correlation between factors is represented by either "similarity" or "dissimilarity". Only for the similarity coefficient, there are many different ways of definition. Correlation coefficient in common use is defined by the following expression to measure the similarity:

$$r_{i,j}=\sum_{k=1}^{N}(x_{ik}-\bar{x}_i)(x_{jk}-\bar{x}_j)/[\sum_{k=1}^{N}(x_{ik}-\bar{x}_i)^2\sum_{k=1}^{N}(x_{jk}-\bar{x}_j)^2]^{1/2} \quad (1)$$

where N—numbers of specimen; x_{ik}, x_{jk}—the ith, and jth property of the kth specimen; $\bar{x}_i=\sum_{k=1}x_{ik}/N$—the average value of the ith property; $\bar{x}_j=\sum_{k=1}x_{jk}/N$—the average value of the jth property.

The correlation coefficient following equation (1) can be either positive or negative. The absolute value indicates the degree of correlation; the sign the direction of concomitant change. According to the expression of the correlation coefficient, r_{ij}, the data from the exploration survey are processed, and the matrix of correlation coefficients, table 1, is obtained.

Table 1 Matrix of correlation coefficients

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
(1)	1.00																					
(2)	0.00	1.00																				
(3)	−0.07	−0.29	1.00																			
(4)	−0.14	−0.43	0.40	1.00																		
(5)	−0.04	−0.05	−0.59	−0.05	1.00																	
(6)	0.20	0.00	−0.43	−0.20	0.24	1.00																
(7)	−0.04	0.05	−0.20	−0.20	0.24	0.17	1.00															
(8)	0.11	−0.16	0.01	−0.05	−0.11	0.00	−0.40	1.00														
(9)	−0.10	0.16	0.00	0.04	0.11	0.01	0.40	−0.20	1.00													
(10)	0.54	0.04	0.20	0.25	−0.21	−0.20	−0.05	−0.00	0.04	1.00												
(11)	0.07	0.20	−0.20	−0.44	−0.01	0.05	0.01	−0.17	0.15	0.05	1.00											
(12)	−0.00	−0.27	0.50	0.79	−0.00	−0.17	−0.10	−0.21	0.21	0.23	−0.27	1.00										
(13)	0.10	−0.20	−0.20	0.13	0.20	0.05	0.22	0.10	−0.16	−0.27	−0.20	0.05	1.00									
(14)	−0.04	0.00	−0.27	−0.44	−0.15	0.05	0.07	−0.17	0.17	0.00	0.00	−0.20	−0.20	1.00								
(15)	−0.07	0.77	−0.20	−0.20	0.07	0.01	0.05	−0.11	0.10	−0.04	0.70	−0.20	−0.13	0.50	1.00							
(16)	0.20	0.22	0.05	−0.20	−0.04	−0.05	−0.22	0.01	−0.05	0.24	0.25	−0.10	−0.22	0.07	−0.24	1.00						
(17)	0.05	0.04	−0.17	−0.70	−0.05	−0.04	−0.00	0.07	−0.00	0.04	0.04	−0.02	−0.20	0.22	0.02	0.40	1.00					
(18)	−0.21	0.71	−0.10	−0.05	−0.17	−0.21	0.05	0.05	−0.05	−0.15	0.07	−0.20	−0.05	0.04	0.40	0.25	0.70	1.00				
(19)	0.05	0.04	−0.20	−0.13	0.02	0.11	0.04	−0.25	0.27	0.07	0.00	−0.00	−0.00	0.02	0.71	−0.05	0.15	0.20	1.00			
(20)	−0.15	−0.24	0.25	0.07	−0.20	0.05	0.00	0.24	−0.24	−0.21	−0.27	0.17	−0.04	−0.15	−0.41	0.07	−0.20	−0.15	−0.22	1.00		
(21)	0.44	−0.04	0.20	0.15	−0.15	−0.25	−0.75	0.20	−0.25	0.74	−0.02	−0.04	−0.21	−0.10	−0.13	0.40	0.14	−0.05	−0.12	−0.23	1.00	
(22)	0.20	−0.13	−0.20	−0.10	0.10	0.27	−0.00	−0.13	0.12	0.17	−0.13	−0.13	0.15	−0.07	−0.10	−0.14	−0.10	−0.42	0.01	0.00	−0.05	1.00

(1)—Resistivity; (2)—Conductivity; (3)—On-site pH; (4)—Indoor pH; (5)—Organic Substances; (6)—Cation-exchange; (7)—Water Capacity; (8)—Volume Density; (9)—Void Volume; (10)—Air Capacity; (11)—Total Salts; (12)—Redox Potential; (13)—Steel/Soil Potential; (14)—Pipe/Soil Potential; (15)—CO_3^{2-}; (16)—HCO_3^-; (17)—Cl^-; (18)—SO_4^{2-}; (19)—NO_3^-; (20)—Ca^{2+}; (21)—Mg^{2+}; (22)—Na^+

1.3 PRELIMINARY ANALYSIS

Although the factors affecting the soil corrosivity are many, the analysis of the correlation shows that there exists certain kind of interrelation among the physical, chemical, and electro-chemical properties of soil. Some correlations between

electro－chemical and chemical properties, chemical and chemical properties, as well as physical and physical properties are very intimate. At the same time, there are also a few physical properties which are scarcely related to other chemical or electro－chemical properties. For example, the correlation coefficient between conductivity (electro－chemical) and total salt (chemical) is high up to 99%; as for volume density (physical) and void volume (physical), though the correlation coefficient between themselves is －99.5%, all the other coefficients related to these two properties are relatively small. Besides, due to the way of sampling, transporting and storage, along with different test methods, the differences are caused between the on－site and the indoor measured properties.

2 SYSTEMATIC CLUSTERING OF SOIL CORROSIVE FACTORS

2.1 DEFINITION OF DISTANCE BETWEEN CLUSTERS AND STEPS OF CLUSTERING

The basis for the systematic clustering is the distance between clusters. There are several ways of defining the cluster distance [3] The way of the shortest distance is used in this paper, i.e., the distance between the pair of points which are from two different clusters separately and with the shortest distance among all the possible pairs is regarded as the distance between these two clusters. The steps of clustering are as follows:

The first step: take each specimen as one cluster. The distance between two clusters is right the distance between the two specimens. Combine the nearest points as a new cluster.

The second step: combine the nearest clusters as a new cluster. The distance between two clusters is in accordance with the above definition.

Then, repeat the second step over and over again. In this way, the varying degrees of clustering can be obtained.

2.2 CLUSTERING OF SOIL CORROSIVE FACTORS

The systematic clustering of the soil corrosive factors is based on two basic hypotheses. First, each factor affecting soil corrosivity is taken as one generalized coordinate point of a specimen. Secondly, the reciprocal of the correlation coefficient between two factors, $1/|r_{ij}|$ is taken as the generalized distance between these two factors. In fact, this is another way of considering the correlation between the soil corrosive factors. Figure 1 gives the pedigree of clustering formed by the change of correlation coefficient with the soil corrosive factors. The absolute value of the correlation coefficient is used in the figure. The result of the clustering shows that the 22 soil corrosive factors are clustered into seven distinct clusters under the correlation coefficient value above 0.5 (see Tab. 2).

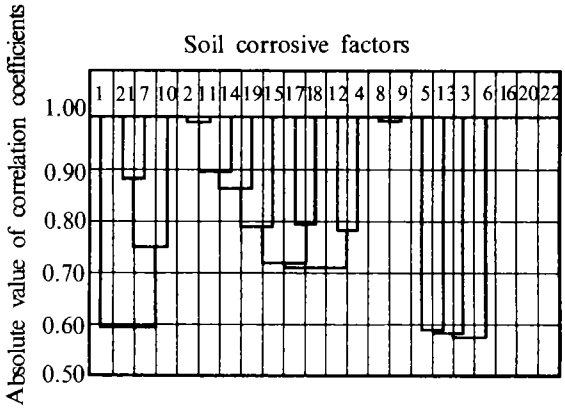

Fig. 1 Pedigree of factor clusters

Table 2 The resultant clusters

Clusters	Soil corrosive factors
1	resistivity, steel/soil potential, water capacity, air capacity
2	conductivity, total salt, Cl^-, Na^+, SO_4^{2-}, Ca^{2+}, Mg^{2+}, CO_3^{2-} indoor pH
3	volume density, void volume
4	organic substances, HCO_3^-, on−site pH, cation Exchange
5	NO_3^-
6	redox potential
7	pipe/soil potential

3 COMMENTS AND CONSIDERATIONS

In this paper, the correlation data processing and the clustering analysis of the soil corrosive factors show the following:

(1) Some of the factors affecting soil corrosivity are interrelated closely. They are clustered together in the generalized distance space. In the same category, the relationship among all the factors of this category is in accordance with the common analysis, such as the relationship between volume density and void volume, total salt and the contribution of its most ions to the conductivity, and so on. Some of the factors are relatively independent.

(2) To some extent, the value of correlation coefficient represents the precision of the expression of one factor by the other. The correlation data processing exposes some problems occurred during the exploration survey of soil corrosive factors. It is necessary to analyze and improve various possible steps during which the differences may be caused, so that the correlation between the properties measured indoors and on the spots can be represented accurately and reliably.

(3) Certain kind of correlation among some factors is achieved by the application of the correlation data processing and the systematic clustering to the analysis of soil corrosive factors. In order to carry on the further research, chosing the representatives from the categories of factors is the dimensional reduction to the many corrosive factors. This is helpful to the selection of major soil factors and the prediction and evaluation of soil corrosive factors.

REFERENCES

1 DIN 50929−85, DEUTSCHE NORM. Berlin: Beuth Verlag GmbH, 1985

2 Chinese National Experimental Network for Soil Corrosion. Experimental Method for Soil Corrosion of Materials. Beijing: Academic Press, 1990

3 Wang Biquan, Chen Zhuyin. Pattern Recognition—Theory, Method and Application. Beijing : Seismic Press, 1989.

土壤腐蚀性因素的相关性

李长荣[1)] 屈祖玉[1)] 王光雍[1)] 陈建民[2)] 丁 乙[2)] 李春光[2)]

1) 北京科技大学表面科学与腐蚀工程系，北京 100083 2) 华东输油管理局 徐州 221008

摘要 本文通过相关性数据处理和模式识别的系统聚类方法，对影响土壤腐蚀性的各种因素之间的相互关系进行了分析讨论. 相关系数矩阵和聚类谱系图两者都表明有些因素之间的相关性很大，有些却相对独立. 本文采用的分析处理方法有助于从众多因素中选取主要特征，进而对土壤腐蚀性作出预测和评价.

关键词 聚类分析，腐蚀性因素，土壤腐蚀