

PAPER • OPEN ACCESS

Development and application of water environment assessment model for lithium mining area based on python

To cite this article: Feng Yu and Yang Yu 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **227** 062002

View the [article online](#) for updates and enhancements.

Development and application of water environment assessment model for lithium mining area based on python

Feng Yu^{1,2} and Yang Yu^{1,3}

¹MNR Key Laboratory of Metallogeny and Mineral Assessment, Institute of Mineral Resources, Chinese Academy of Geological Sciences, Beijing 100037, China;

²China University of Geosciences, Beijing 100083, China;

³Email: yuyang_cags@sina.com

Abstract. Sixty-four surface water samples were selected to study the water quality around the lithium mining area in western Sichuan Province. Principal component analysis, in which 6 principal component indexes (Mn, Zn, As, Cl⁻, SO₄²⁻, NO₃⁻) with cumulative contribution rate >85% were selected as the assessment index. Based on fuzzy mathematics, the membership function is used to determine the attribution of all indexes, and the pollution contribution rate is used to determine the index weight. Using the national standard (*Environmental quality standards for surface water, GB3838-2002* and *Quality standard for ground water, GB/T 14848-93*) as a benchmark, a fuzzy comprehensive assessment model based on Python language was established to assess the water environment quality of the mining area. The result shows that the water environment classification of sixty-four sampling points in the mining area are all Grade I. This result indicates that the water environment quality of the mining area is very good, and the mineralization and mining engineering activities do not pollute the water environment around the mining area.

1. Introduction

As a rare metal, lithium is widely used in batteries, ceramics, glass, lubricants, refrigerants, nuclear industry, and optoelectronics. Lithium has important strategic value, especially in recent years in the field of new energy, and was known as "the energy metal of the 21st century". There are currently more than 150 kinds of lithium-containing minerals, including spodumene, lithium mica, lithium phosphite, lithium feldspar, iron-lithium mica, etc. The Li₂O content of spodumene, with the content from 5.8% to 8.1% [1], is the highest among the common lithium-containing minerals. Therefore, the spodumene resource enrichment area is one of the most important sources of lithium ore and has a very high development and utilization value. However, whether the development and utilization of spodumene will bring environmental pollution, especially water environment problems, has not been studied yet. The Jiajika Lithium Mine located in western Sichuan Province is the most important spodumene lithium mine base in China, which was selected as an example to establish a water environment assessment model for the spodumene resource enrichment area.

There is no recognized and unified system for the assessment of water environment [2]. As early as the 1960s, Jacobs proposed a water quality index for water quality assessment [3]. In the following decades, scholars at home and abroad have proposed dozens of different comprehensive indexes in the assessment of water environment. Comprehensive indexes are still heavily used because of the simple calculation and easy operation. Since the 1980s, the emergence of some uncertainties, nonlinear



theoretical and the popularization of computer technology have provided new theoretical basis and technical support for water environmental assessment [4].

2. Assessment model establishment

Fuzzy comprehensive assessment, a method for comprehensive assessment, was used to assess the water environment of the mining area in this paper. Based on fuzzy mathematics [5], fuzzy comprehensive assessment applied the principle of fuzzy relationship, and quantified some indexes with unclear boundaries. In order to realize the comprehensive assessment and ranking of water quality grades through the establishment of the index set, assessment set, membership function and weight set. The fuzzy comprehensive assessment could objectively reflect the actual situation of the water environment in the mining area, and also solve a series of objective fuzzy problems such as pollution degree and water quality grades. It has become a common method for comprehensive assessment of water environment, and already been considered as an effective means of studying the water environment.

2.1. Determination of assessment indexes

The comprehensive assessment indexes of the water environment in the Jiajika mining area include various ions and elements. The assessment of the water environment is a multi-factor comprehensive analysis problem. For this multivariate, multi-index, large sample assessment process, the dimensionality reduction of indexes is the primary problem to be solved. However, in many current assessment models, there are no uniform selection criteria for assessment indexes. Different assessment models use different indexes, such as the empirical index, which causes human factors to interfere with the assessment results, resulting different assessment results in the same place. It is impossible to provide effective decision-making information for decision makers, such as water resources conservation, sustainable development, utilization of land resources, etc. Principal component analysis is a computational method based on statistical principle, and there is no need to human intervention. Its calculation results are objective and have been widely used in the field of assessment [6]. Therefore, for the comprehensive assessment of the water environment around the mining area, principal component analysis can be used to determine the indexes.

Before using principal component analysis, the original indexes must be dimensionless. Non-dimensionalization is also called data standardization. It eliminates the influence of the dimension of each index through mathematical changes, making the indexes comparable.

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_j} (i = 1, 2, \dots, m; j = 1, 2, \dots, n) \tag{1}$$

Where, x_{ij} indicates the value of the index j of the sampling point i , \bar{x}_j and σ_j represent the mean and standard deviation of the index j , respectively. The data normalized by the standard deviation has a mean of 0 and a standard deviation of 1.

Using principal component analysis to analyze the indexes, first calculate the covariance matrix Σ :

$$\Sigma = \begin{bmatrix} s_1^2 & \text{cov}(1,2) & \dots & \text{cov}(1,n) \\ \text{cov}(2,1) & s_2^2 & \dots & \text{cov}(2,n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(n,1) & \text{cov}(n,2) & \dots & s_n^2 \end{bmatrix} \tag{2}$$

Where, $\text{cov}(x, x) = s_x^2$

$$\text{cov}(x, y) = \text{cov}(y, x) = \frac{1}{m-1} \begin{bmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \dots & \gamma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \dots & \gamma_{nn} \end{bmatrix} \tag{3}$$

Where, $\gamma_{xy} = \frac{\text{cov}(x, y)}{s_x s_y}$, Then calculate the eigenvalues and eigenvectors of the correlation matrix

R. According to the characteristic equation:

$$|R - \lambda I| = 0 \tag{4}$$

Calculated eigenvalue λ_i , Corresponding eigenvector U.

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nm} \end{bmatrix} = (U_1, U_2, \dots, U_n) \tag{5}$$

The principal component is calculated based on the previous calculation results. The linear combination of the first r principal components is:

$$\begin{aligned} F_1 &= U_1 X^T = u_{11}x_1 + u_{12}x_2 + \cdots + u_{1n}x_n \\ F_2 &= U_2 X^T = u_{21}x_1 + u_{22}x_2 + \cdots + u_{2n}x_n \\ &\vdots \\ F_r &= U_r X^T = u_{r1}x_1 + u_{r2}x_2 + \cdots + u_{rn}x_n \end{aligned} \tag{6}$$

Where, F_1, F_2, \dots, F_r are irrelevant, the variance contribution rate and the cumulative variance contribution rate of the feature values are calculated.

$$e_i = \lambda_i / \sum_{i=1}^n \lambda_i (i = 1, 2, \dots, n) \tag{7}$$

According to formula (7), the contribution rate of the principal component can be calculated, and the cumulative variance contribution rate of the first r principal components:

$$E_r = \sum_{k=1}^r \lambda_k / \sum_{i=1}^n \lambda_i \tag{8}$$

Judging based on the cumulative variance contribution rate E_r : When $E_r > 85\%$, the assessment indexes are determined by the cumulative variance contribution rate of the first r principal components. It is indicated that the first r principal components basically contain the information of all the data, and the higher scores among the r principal components are selected as the assessment indexes.

2.2. Establishment of fuzzy assessment model

There are four models for calculating the weighted membership degree by fuzzy comprehensive assessment [7]: $M(\wedge, \vee)$, $M(\circ, \vee)$, $M(\wedge, \oplus)$ and $M(\circ, +)$. Among them, only the $M(\circ, +)$ model is considered reasonable for water environment assessment. The target membership degree is taken as the "weighted sum" of the index membership degree and the weight of the indexes [8]. The fuzzy comprehensive assessment method can be expressed by the following mathematical model:

$$B = A \circ R \tag{9}$$

Where A is the input, an $l \times n$ order matrix obtained by normalizing the weight of the parameter assessment index; R is a "fuzzy converter", which is an $n \times m$ order fuzzy matrix composed of each single index matrix; B is the output, which is the assessment result obtained, and is an $l \times m$ order matrix.

The comprehensive assessment model of water environment consists of four parts: index set U, weight set A, assessment set V and fuzzy matrix R. The index set U takes the index obtained based on the principal component analysis. An index set consisting of n elements $U = \{U_1, U_2, \dots, U_n\}$, where index U_n is the sampling point of the water environment assessment sample.

$$\begin{aligned} U_1 &= \{u_{11}, u_{12}, \dots, u_{1m}\} \\ U_2 &= \{u_{21}, u_{22}, \dots, u_{2m}\} \\ &\vdots \\ U_n &= \{u_{n1}, u_{n2}, \dots, u_{nm}\} \end{aligned} \tag{10}$$

Where u_{nm} represents the value of the index m of the sampling point n. The assessment set V can be established according to the selected assessment criteria. For the weight set A, the pollution contribution rate method is generally used to objectively calculate the weight coefficients of each index. The mathematical formula is:

$$A_i = \frac{\bar{u}_i}{s_i} \left(\sum_{i=1}^n \frac{\bar{u}_i}{s_i} \right)^{-1} \tag{11}$$

Where u_{nm} represents the average of all measured values of the index i; \bar{s}_i represents the mean of the different grades of the index i in the assessment criteria; n is the number of assessment indexes for each sample; and A_i is the weight of the index i.

The membership degree is the degree to which the index belongs to different standards. The characteristic is that the assessment result is not affirmative or negative but is represented by a fuzzy set. Extend the value range from the original value of only 0 and 1 to [0,1] [9]. The membership degree is determined by obscuring the assessment index by using the membership function. Commonly used membership functions include many different kinds of distributions, such as semi-rectangular distribution, semi-normal distribution, asemi-trapezoidal distribution, semi-concave (convex) distribution, lower-half-west distribution, etc. In this study, the membership function is established by using a semi-trapezoidal distribution. The formulas are as follows:

Assume that the assessment criteria are divided into five grades. If the sample belongs to the first grade (j=1), the membership function is:

$$r_{ij} = \begin{cases} 1 & 0 \leq u_i \leq s_{ij} \\ \frac{s_{ij+1} - u_i}{s_{ij+1} - s_{ij}} & s_{ij} \leq u_i \leq s_{ij+1} \\ 0 & s_{ij+1} \leq u_i \end{cases} \tag{12}$$

If it belongs to Grade II~IV(j=2~4), the membership function is:

$$r_{ij} = \begin{cases} 1 & u_i = s_{ij} \\ \frac{u_i - s_{ij-1}}{s_{ij} - s_{ij-1}} & s_{ij-1} \leq u_i < s_{ij} \\ \frac{s_{ij+1} - u_i}{s_{ij+1} - s_{ij}} & s_{ij} < u_i \leq s_{ij+1} \\ 0 & x_i > s_{ij+1}, x_i < s_{ij-1} \end{cases} \tag{13}$$

If it belongs to Grade V(j=5), the membership function is:

$$r_{ij} = \begin{cases} 0 & 0 < u_i < s_{ij-1} \\ \frac{u_i - s_{ij-1}}{s_{ij} - s_{ij-1}} & s_{ij-1} < u_i < s_{ij} \\ 1 & u_i > s_{ij} \end{cases} \tag{14}$$

In the above formulas, x_i is the measured value of the index i, and s_{ij-1} , s_{ij} , s_{ij+1} are the standard values of the grade j-1, j, j+1 of the index i, respectively. Finally, according to the model A°R synthesis operation, and then according to the fuzzy mathematics principle and the maximum membership degree principle to determine the assessment of the water environment.

3. Assessment model application

This study takes the Jiajika spodumene mining area as an example. The spodumene in the Jiajika mining area is mainly distributed in the granitic pegmatite. The predecessors have studied the geological characteristics, ore body characteristics, metallogenic age and metallogenic regularity of

the mining area [10-15]. However, the environment of the mining area has not been clearly assessed, especially the assessment models of the water environment are now in the blank stage of research.

Table 1. Total variance interpretation.

	Variance contribution rate%	Cumulative contribution rate %
Factor 1	79.539	79.539
Factor 2	13.481	93.020

Using the calculated 6 assessment indexes, the index set U can be established according to formula (10).

Table 2. Factor loadings (loading >0.2 marked in bold).

	Factor 1	Factor 2	Communalities
Li	0.120	-0.178	0.046
Sc	0.126	-0.110	0.028
Mn	0.999	-0.037	0.999
Ni	-0.007	-0.380	0.144
Zn	0.256	0.933	0.936
As	0.377	0.751	0.706
Se	-0.020	0.046	0.003
Rb	0.003	-0.048	0.002
Sr	0.062	-0.072	0.009
Ba	0.020	0.114	0.013
K ⁺	-0.154	-0.324	0.129
Na ⁺	0.027	-0.288	0.084
Ca ²⁺	0.014	-0.093	0.009
Mg ²⁺	0.116	-0.144	0.034
HCO ₃ ⁻	-0.138	-0.088	0.027
Cl ⁻	-0.104	0.540	0.302
SO ₄ ²⁻	-0.059	0.295	0.091
NO ₃ ³⁻	0.008	0.205	0.042
F ⁻	-0.111	0.085	0.020

3.1. Sample collection and testing

In this research work, a total of sixty-four samples of the water environment around the mining area were collected. The water samples are all preserved in ultra-cleaned polyethylene jars. There are two types of water samples collected in the field. One type is sealed directly with PARAFILM sealing film after collection, used for the determination of major elements and ions. The other type was filtered through a polyethersulfone (PES) membrane having a pore size of 0.45 μm , and 1:1 pure nitric acid was added, sealed with a PARAFILM sealing membrane for the determination of trace element content. The purpose of adding nitric acid is to slow the hydrolysis and redox of the compound, and to reduce evaporation and components loss.

Inductively coupled plasma mass spectrometry (PE300D) was used for determination the concentration of trace elements and major anions (F⁻, Cl⁻, HCO₃³⁻, SO₄²⁻, NO₃³⁻, NO₂²⁻) in water samples. The concentration of main cations (Ca²⁺, K⁺, Mg²⁺, Na⁺) and next elements: Al, P, and B were determined by a plasma spectrometer (PE8300). As and Se concentration was determined by atomic fluorescence spectrometry.

3.2. Selection of index and construction of index set U and assessment set V

The Statistical Package for the Social Sciences (SPSS) was used to analyze the data of sixty-four sampling points in the mining area. After calculation, two principal components are extracted, and the cumulative contribution rate of the two principal components is 93.02%>85%, results are shown in Table 1 and Table 2.

On the basis of calculating the water environment assessment index, the assessment set V of six indexes is established with reference to the national standard (*Environmental quality standards for surface water, GB3838-2002* and *Quality standard for ground water, GB/T 14848-93*). Results are shown in Table 3.

Table 3. Assessment set of 6 indexes grading assessment ($\mu\text{g/L}$).

	Mn	Zn	As	Cl ⁻	SO ₄ ²⁻	NO ₃ ⁻
Grade I	50	50	50	50000	50000	2000
Grade II	50	1000	50	150000	150000	5000
Grade III	100	1000	50	250000	250000	20000
Grade IV	1000	2000	100	350000	350000	30000
Grade V	2000	2000	100	450000	450000	50000

3.3. Membership calculation and grading assessment

```

File Edit Format Run Options Window Help
# -*- coding: utf-8 -*-
"""
Created on Fri Aug 10 10:29:53 2018
@author: yufeng
"""

import numpy as np
import csv

with open('data.csv', 'r', encoding = 'UTF-8') as file_1:
    reader = csv.reader(file_1)
    t = np.array(list(reader))
    file_1.close()

po = np.array(t[1:, 0], dtype = str)
u = np.array(t[1:, 1:], dtype = float)

Mn = np.array(u[:, 0])
Zn = np.array(u[:, 1])
As = np.array(u[:, 2])
Cl = np.array(u[:, 3])
SO4 = np.array(u[:, 4])
NO3 = np.array(u[:, 5])

with open('standard.csv', 'r', encoding = 'UTF-8') as file_2:
    reader = csv.reader(file_2)
    t = np.array(list(reader))
    file_2.close()

r_Min = np.zeros((len(Mn), 5))
r_Zn = np.zeros((len(Zn), 5))
r_As = np.zeros((len(As), 5))
r_Cl = np.zeros((len(Cl), 5))
r_SO4 = np.zeros((len(SO4), 5))
r_NO3 = np.zeros((len(NO3), 5))

j = 1
for i in range(len(Mn)):
    if Mn[i] < float(t[j, 1]):
        r_Min[i, j - 1] = 1
    
```

Figure 1. The Python calculation code for fuzzy comprehensive assessment.

According to the formula (12)~(14), the membership degree of each index of each sampling point is calculated by Python, and then the six indexes of each sampling point are weighted and summed to calculate the membership degree of the sampling point. Figure 1 shows the Python calculation code for fuzzy comprehensive assessment.

4. Conclusions

It is a very effective to evaluate the water environment of the spodumene resource enrichment area by fuzzy comprehensive assessment. The assessment indexes could be obtained by principal component analysis. The establishment of the assessment set could refer to the corresponding national standards. Use the fuzzy membership function and the pollution contribution rate to determine the classification and weight of the assessment index. This paper established an effectively assessment model for the water environment of the spodumene resource enrichment area.

For the Jiajika mining area, six assessment indexes (Mn, Zn, As, Cl⁻, SO₄²⁻, NO₃⁻) were selected for the assessment of the water environment by principal component calculation. The assessment set was established according to the national standard (*Environmental quality standards for surface water, GB3838-2002* and *Quality standard for ground water, GB/T 14848-93*), and the water environment grading assessment result of sixty-four sampling points was finally obtained by using Python scientific calculation. The water environment classification of all the sixty-four sampling points in the mining area is Grade I, indicating that the water environment of the mining area is very good, and the mining activities have not affected the water environment in the area.

Acknowledgement

The authors thank two anonymous reviewers and editor for constructive comments that greatly improved this manuscript. This work was supported jointly by the National Natural Science Foundation of China (No. 41202254) and the Central Public-interest Scientific Institution Basal Research Fund (No. K1209).

References

- [1] Editorial Board of the Mineral Resources Industry Requirements Manual 2010 Mineral Resources Industry Requirements Manual *Geological Publishing House*
- [2] Jin J L, Wei Y M, Ding J 2001 Projection Pursuit Model for Comprehensive Evaluation of Water Quality *Journal of Environmental Science* **21(4)** 431-4
- [3] Jacobs, Harold L, Gabrielson, Ira N, Horton, Robert K, Lyon, Walter A, Hubbard, Earle C, McCallum, Gordon E 1965 Water Quality Criteria-Stream vs. Effluent Standards *Journal (Water Pollution Control Federation)* 292-315
- [4] Li R Z 2005 Research progress and trend analysis of theoretical model of water quality assessment *Journal of Hefei University of Technology (Natural Science Edition)* **28(4)** 369-73
- [5] Feng Y D, Lou S B 1983 Fuzzy mathematics method and application *Seismological press*
- [6] Feng L H 2003 Principal component analysis of environmental quality *Journal of Mathematics in Practice and Theory* **33(8)** 32-5
- [7] Lu Y S 1999 Environmental Assessment - 2nd Edition *Tongji University Press*
- [8] Liu X, Liu K D, Li C J, Pang Y J, Li W G 2009 Index Weight and Membership Conversion Algorithm in Water Quality Evaluation *Journal of Lanzhou University of Technology* **35(1)** 63-6
- [9] Wang J F, Lu Z D 2000 Method for determining membership function in fuzzy control *Henan Sciences* **18(4)** 348-51
- [10] Fu X F, Yuan L P, Wang D H, Hou L W, Pan M, Hao X F, Liang B, Tang Y 2015 Mineralization characteristics and prospecting model of newly discovered X03 rare metal vein in Jiajika orefield, Sichuan *Mineral Deposits* **34(6)** 1172-86
- [11] Liu L J, Fu X F, Wang D H, Hao X F, Yuan L P, Pan M 2015 Geological characteristics and metallogeny of Jiajika-style rare metal deposits *Mineral Deposits* **34(6)** 1187-98
- [12] Wang D H, Wang C H, Sun Y, Li J K, Liu S B, Rao Y K 2017 Research progress and related issues of lithium-niobium deposits in China *Geological Survey of China* **4(5)** 1-8
- [13] Hao X F, Fu X F, Liang B, Yuan L P, Pan M, Tang Y 2015 Formation ages of granite and X03 pegmatite vein in Jiajika western Sichuan, and their geological significance *Mineral Deposits* **34(6)** 1199-208
- [14] Liu L J, Fu X F, Wang D H, Hao X F, Yuan L P, Pan M 2015 Geological characteristics and metallogeny of Jiajika-style rare metal deposits *Mineral Deposits*
- [15] Wang D H, Li J K, Fu X F 2005 40Ar/39Ar dating for the Jiajika pegmatite-type rare metal deposit in western Sichuan and its significance *Geochimica*