Material selection and optimization scheme of Seabed Data Center

Kechao Xu1,*, Xiaotong Gu2, Shangyi Jiao3, Boyang Ren2, Zhengyu Tian2, Yutong Fan2, Yan An4, Yuhao Lin2, Hanxiao Wang2

1Shandong University of Science and Technology, College of Energy and Mining Engineering, Qingdao, Shandong, 266590
2Shandong University of Science and Technology, School of Mathematics and Systems Science, Qingdao, Shandong, 266590
3Shandong University of Science and Technology, College of Geodesy and Geomatics, Qingdao, Shandong, 266590
4Shandong University of Science and Technology School of Materials Science and Engineering, Qingdao, Shandong, 266590

*Corresponding author: kechaoxu@sdust.edu.cn

Abstract. With the advent of the era of big data, the number of global data centers increases rapidly. But at the same time, it also brings some conflicts of resources. Because electronics use a lot of energy to dissipate heat. In this paper, in order to select the most suitable materials for Seabed Data Center, we established a material screening model based on weight and entropy weight method, and finally obtained the optimal design scheme.

Keywords: entropy weight method; multiple time series segmentation method.

1. Introduction
With the development of the Internet, various industries are increasingly dependent on data, and the construction of big data centers is becoming more and more important. Heat treatment is an important link in the operation of big data centers. Most of the data centers in China are built in inland and coastal areas, which occupy a lot of land and cost a lot of money. Heat dissipation also consumes a lot of power and cooling water resources. Therefore, the construction of undersea data centers has become the development trend of China's big data centers in the future. Sea water is a strong corrosive medium. In order to improve the pressure resistance and corrosion resistance of the container shell, the material analysis model is established and the material selection is carried out through the information of different materials in the attachment, so as to achieve the optimal material selection scheme and the optimal design of the seabed depth.

2. Model building and solution

2.1. Data preprocessing
First of all, we imported the data in the attachment into the Excel table for data preprocessing. Through
cluster analysis, we divided the materials into metal and non-metal categories according to the different types of materials. And they removed outliers. According to the requirements of the topic, we divide the factors that affect the selection of materials into two categories: corrosion resistance and compression resistance. After the analysis, the elastic modulus, flexural strength and tensile strength are finally divided into the criteria to judge the compressive property of materials, and the criteria to judge the corrosion resistance of seawater potential and the corroded state. Due to the parameters of the concrete data is given in appendix compressive sex more, so we give priority to with pressure resistance standard judgment, comprehensive judgment result of corrosion resistance, make a comprehensive analysis, so as to choose the most appropriate materials, and ensure the quality of cost processing scheme is given at the same time, to reduce the cost, on the basis of optimization for container radiator bottom depth.

2.2. A material screening model

2.2.1. The data is processed. Due to the large differences in the data given, we use SPSS software for data processing to eliminate the influence of outliers and make the solution result more accurate.

2.2.2. Model principle. TOPSIS method, namely the solution distance method between good and bad, is a kind of ordering method close to the ideal. In the operation process, it can make full use of the information of the original formula to accurately evaluate each scheme and give the ideal score. In view of the fact that the original data in the appendix are relatively complicated and the topic requires the use of data to screen out the most appropriate production materials, we decided to establish a screening model based on TOPSIS method. The principle is as follows:

The first step is to convert the matrix forward, that is, to convert all index types into extremely large indexes. The formula is as follows:

$$\text{max} = x$$  

The second step standardizes the forward matrix, that is, in order to eliminate the influence of different dimensional indicators, the formula is as follows:

$$z_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{n}x_{ij}^2}}$$  

The third step Calculate scores and normalize them:

Define a maximum:

$$Z^{\max}$$

Define a minimum:

$$Z^{\min}$$

Define the distance between each evaluation object in the I and the maximum value:

$$D_i^{\max} = \sqrt{\sum_{j=1}^{n}(Z_{ij}^{\max} - z_{ij}^{\max})^2}$$  

Define the distance between each evaluation object in the I and the minimum value:
\[ D_i^{\text{min}} = \sqrt{\sum_{j=1}^{n} (Z_j^{\text{min}} - z_j)^2} \]  

(4)

From the above formula, the unnormalized score of each evaluation object in the \( i \)th can be obtained:

\[ S_i = \frac{D_i^{\text{min}}}{D_i^{\text{max}}} \]  

(5)

Normalize the score:

\[ S_i' = \frac{S_i}{\sum_{i=1}^{n} S_i} \]  

(6)

2.2.3. Model establishment and solution. First, we do pressure-resistant screening for material information. The elastic modulus, flexural strength and tensile strength were selected as the judging basis, and the superior solution distance method was used to screen the materials. The specific steps are as follows:

The pre-processed data was imported into MATLAB, which was firstly processed by forward processing and then standardized processing to calculate the score. It can be obtained by solving in MATLAB:

The comparison can be 410 stainless steel score the highest, that is, the best pressure resistance. The detailed data are shown in the figure:

![Figure 1 410 Stainless steel](image)

Secondly, we did corrosion resistance analysis on the material, the same principle as above, and the material with the highest score was butyl rubber. The detailed data are shown in the figure:

![Figure 2 Butyl rubber](image)
2.3. Material Selection Scheme

According to the solution results of the model, the material with the best pressure resistance is 410 stainless steel, and the material with the best corrosion resistance is butyl rubber. Therefore, we choose 410 stainless steel as the main shell material, and butyl rubber is added on the surface of stainless steel, so as to ensure the pressure resistance and increase the corrosion resistance of the material. According to the detailed information of the material, both of them have good cooling effect, and the price of 410 stainless steel is 14,000 yuan/ton, the price of butyl rubber is 17,500 yuan/ton, compared with the common materials, the price is lower. Therefore, on the basis of reducing the cost and increasing the service life, the model optimizes the heat dissipation effect of the container, and the solution results are reasonable.

2.4. Model optimization scheme based on weight method and entropy weight method

In the calculation of distance by TOPSIS, the weight of each index is the same by default, while in practical problems, the importance of different indexes may be different. Therefore, in practical problems, we consider to give weight to the index, that is, to introduce the weight into TOPSIS method to calculate the distance formula.

At the same time, in order to reduce the subjectivity given by weight, we introduce the entropy weight method to calculate the weight, in order to increase the objectivity given by weight and improve the accuracy of the calculation results. The judgment criteria are as follows: the greater the information entropy of each index is, the greater its weight will be; On the contrary, the smaller the information entropy of an index is, the smaller the variation degree of the index value is, and the smaller its weight is.

The specific steps are as follows:

The first step is to standardize the data of each index:

\[ X_{ij} = \frac{x_{ij}-\min(x_i)}{\max(x_i)-\min(x_i)} \]  

(7)

The second step is to calculate the information entropy of each index:

\[ \delta_j = -\frac{1}{\ln n} \sum_{i=1}^{n} p_{ij} \ln p_{ij} \]  

(8)

\[ p_{ij} = \frac{x_{ij}}{\sum_{i=1}^{n} x_{ij}} \]  

(9)

The third step is to determine the weight of each index:

\[ W_i = \frac{1-\delta_i}{\kappa-\sum_{i=1}^{\kappa} \delta_i} (i = 1, 2, \cdots, \kappa) \]  

(10)

Finally, the distance between each evaluation object in the ith and the maximum value is redefined:

\[ D_{i}^{\max} = \sqrt{\sum_{j=1}^{n} W_j (Z_{j}^{\max} - z_j)^2} \]  

(11)

The distance between each evaluation object in the I and the minimum value is redefined:

\[ D_{i}^{\min} = \sqrt{\sum_{j=1}^{n} W_j (Z_{j}^{\min} - z_j)^2} \]  

(12)
The unnormalized score of the $i$th evaluation object can be obtained by recalculating the above formula:

$$S_i = \frac{D_i^{\min}}{D_i^{\max} - D_i^{\min}}$$  \hspace{1cm} (13)

Normalize the score:

$$S_i' = \frac{S_i}{\sum_{i=1}^{n} S_i}$$  \hspace{1cm} (14)

2.5. Subsea depth optimization design

First, MATLAB is used to solve the curve of water velocity changing with temperature when the container is fully loaded with the server, as can be seen from the figure, the faster the flow rate of seawater is, the faster the heat dissipation rate of the container will be. Due to the influence of monsoon and temperature difference, the deeper the water level is, the lower the flow velocity is generally. Therefore, it is necessary to select an appropriate seabed depth [1], which takes into account both the temperature difference between the container and the seawater and the seawater velocity.

The second step is to use MATLAB and reference data to calculate the variation of seawater velocity along the depth:

![Figure 3. Seawater velocity varies with depth](image1)

![Figure 4. The seawater velocity varies quarterly](image2)
So the best place to build seabed data is somewhere between the upper and deeper parts of the ocean. Construction on the seabed has more advantages. First laying does not require braces, so it can be built faster and with less investment. Secondly, the submarine is less affected by the destruction of natural factors and human production activities, so it has better stability and reliability than the land cable. In the case of suitable materials, anti-interference ability and confidentiality are enhanced.

2.6. Analysis of tide and seasonal factors
Tides can be divided into semi-diurnal, diurnal and mixed tides according to the period. Under the action of tide, the container shell will be in an alternately wet and dry environment for a long time. The "wet" process provides conditions for corrosive electrolytes, while the "dry" process provides more oxygen and intensifies the corrosion intensity. The tide will affect the velocity of sea water to some extent, and then affect the heat dissipation effect of the radiator.

When the tide is large, the flow rate of the tide is faster and the heat is transported farther. Neap tides, on the other hand, move more slowly and carry heat closer. Seasonal changes mainly affect sea temperature, which in turn affects convective heat transfer. And in different seasons, the water level will also change, so as to reduce or increase the pressure of the container, affecting the shell [2].

![Figure 5 Water level seasonal variation curve](image)

As can be seen from seasonal changes in the course of the journey, the temperature of sea water is higher in summer, and the fluctuation of sea water changes greatly. In early spring and late winter, the sea is cooler. Seawater temperature changes relatively steadily. Meanwhile, the change of seasonal factors belongs to the change of time series. So we have to find a way to do multidimensional analysis.

2.7. Optimization scheme of analysis model based on multiple time series segmentation
Firstly, based on the temperature changes caused by tides and seasons, we established the fitting temperature data equation, so as to observe the temperature changes more intuitively.

\[
y = n_n x^n + n_{n-1} x^{n-1} + \cdots + n_2 x^2 + nx + a
\]  

(15)

For multidimensional changes in time and space. We adopt a piecewise method of multivariate time series based on factor model and dynamic programming [3].

1. Carry out variable clustering. In order to deal with the parameters of higher latitude, we calculate the Person correlation coefficient between any two variables to get the normalized correlation coefficient, investigate the correlation between variables, and then get the clustering result, so as to achieve dimensional reduction processing and reduce the processing difficulty.

**Person** calculation formula of the correlation coefficient is as follows:

\[
corr(e_i(t), e_j(t)) = \frac{\sum_{t=1}^{T} [e_i(t), \bar{x}_i] \cdot [e_j(t), \bar{x}_j]}{\sqrt{\sum_{t=1}^{T} [e_i(t), \bar{x}_i]^2 \sum_{t=1}^{T} [e_j(t), \bar{x}_j]^2}}
\]  

(16)
2. Extract common factor sequences. C clusters obtained from the original data after clustering all contain one or more variable sequences. For the clusters containing univariate sequences, the dimensionality reduction process is no longer carried out, and the cluster is directly added to the multivariate time series with low dimensions. For the cluster with multiple variable sequences, dynamic factor extraction is selected.

The multivariable sequence \( Z^{(m_1)}(t_1) \) in the \( m_1 \) cluster can be modeled by the dynamic factor:

\[
Z^{(m_1)}(t_1) = A^{(m_1)}f^{(m_1)}(t_1) + v^{(m_1)}(t_1) \tag{17}
\]

The hysteresis operator polynomial is:

\[
\psi(m_1)(A)f^{(m_1)}(t_1) = \omega^{(m_1)}(t_1) \tag{18}
\]

3. Segmentation of multivariate time series. For each multivariable sequence data set, a common factor sequence is extracted, and univariate sequences that are not clustered with other variables are combined to form a low-dimensional multivariate time series of a new dimension. The low dimensional multivariate time series of the new dimension are segmented.

The BIC criterion was used to determine the optimal value of the autoregressive order and the number of segments of the autoregressive parameter:

\[
BIC(p, N) = \log\left( \sum_{i=1}^{L(N)} \log(J_{p-1}) + \frac{\log(J-p)}{j-p} \times J \times c \times (c \times p + 1) \right) \tag{19}
\]

2.8. Analysis of model results

Since seasonal variations in land and ocean temperatures coincide well, we fitted the outdoor temperature.

![Temperature fitting diagram](image)

Figure 6 Temperature fitting diagram

In the tidal area, it is recommended to choose neap tide area with short period to relatively reduce the influence of corrosion on the shell. In terms of season, it is suggested to intensify the inspection and maintenance of the data center in summer. To ensure the security of the subsea data centers.

3. Model evaluation and generalization

3.1. Advantages of model

The TOPSIS method optimized based on weight method and entropy weight method reduces the subjectivity of evaluation criteria and can be used for real value processing, which has a certain promotion and use value.
3.2. Model shortcomings
During data preprocessing, due to the limitations of the model itself and the complex and missing data given, the selection of materials is relatively one-sided.

The seasonal and tidal factors are treated separately, and the correlation between variables is less considered, which may limit the analysis of the problem in some cases.

3.3. Generalization of the model
Because the current research of submarine data center is still in the development stage, and the heat dissipation problem is the premise to solve the technical problems of development. Therefore, the decision analysis model established by us has certain popularization and application value for the treatment of heat dissipation problem, and can provide reference and reference for the development of data center industry.

Reference
[1] Zha Miao, Wang Yingjian, Wang Zhenyu et al. Research on Buried Depth of Undersea Optical Cable. Optical Fiber and Cable and Its Application Technology, 2015, 000(002):33-36,42.
[2] Zhang Hong, Jiang Zhiting, Liu Enxu. Seasonal variation characteristics of water temperature in the Yellow Sea and East China Sea [J]. Journal of China Water Transport, 2016, 16(011): 167-170.
[3] Wang Ling, Xu Peipei, Peng Kaixiang. Multivariate time series segmentation method based on factor model and dynamic programming [J]. Control and Decision, 35(1):35.