Supplier Score Model Based on Pearson Correlation Coefficient and Entropy Weighting Method

Xiaolin Liu [1, *], Yuyang Qi [2]

[1] University of Jinan, School of Electrical Engineering, Jinan, 250000

[2] University of Jinan, Business School University of Jinan, Jinan, 250000

* Corresponding Author Email: bsquirrel@zohomail.cn

Both authors contributed as much to the article

Abstract. For the problem of ordering and transportation of raw materials in production enterprises, a comprehensive score of supplier supply characteristics is established based on the entropy value method, and a mathematical model reflecting the importance of guaranteeing the production of enterprises is established. Firstly, based on the raw material suppliers' supply data, we get the better 60 suppliers according to the total supply; secondly, secondly, based on the supply and ordering data, we excavate the indexes to measure the supply characteristics of each supplier, and carry out the Pearson correlation analysis, and get the four indexes which are weakly correlated with the type of materials supplied by the suppliers, order fulfillment, stability of supply, and enterprise purchasing. Finally, the entropy value method is used to determine the weight of each indicator, to construct the indicator system reflecting the importance of safeguarding the production of the enterprise, to obtain the comprehensive score of the supply characteristics of the 60 suppliers after screening, and to determine the top 50 suppliers such as S361, S218, etc. as the most important suppliers according to the descending order of the score.

Keywords: Pearson correlation analysis; entropy method; multi-objective planning models.

1. Introduction

The procurement and transportation costs of raw materials directly impact a company's production efficiency. To ensure normal production and operation, a company needs to quantitatively analyze the supply characteristics of its suppliers and establish a scientifically sound mathematical model that reflects the importance of safeguarding its own production. This will help in devising the best raw material procurement and transportation plans tailored to the specific circumstances of the suppliers.

In this context, the question provides data for the past five years, including the order quantity and supply quantity from 402 raw material suppliers, as well as the transportation loss rates from 8 transportation companies. The goal is to analyze the relevant data, taking into account practical considerations such as the company's production capacity and production demands, and address the following issue:

Considering factors such as the type of materials supplied and order fulfillment rates, quantitatively assess the supply characteristics of the 402 raw material suppliers. Simultaneously, establish a mathematical model that reflects the importance of safeguarding the company's production, using this model as the basis to identify the top 50 most critical suppliers.

2. Identification of Key Suppliers

2.1. Analysis of Approach

The first question requires data mining and processing of basic data, based on the order quantity and supply quantity data from raw material suppliers. This involves a comprehensive evaluation of supplier's operations, credibility, strength, and other aspects to derive four key indicators: material types supplied, order fulfillment rate, supplier supply stability, and enterprise procurement ratio. The reasonableness of these indicators is then validated through Pearson correlation analysis.

Subsequently, the entropy method is used to determine indicator weights, leading to the construction of an indicator system that reflects the importance of safeguarding the company's production. This process results in a composite score for the supply characteristics of 60 selected suppliers, which are then ranked in descending order to identify the top 50 most important suppliers.

2.2. Data Preprocessing

Due to long-term operational and transportation cost considerations, the company will not establish long-term cooperation with suppliers with low supply volumes, thus based on the formula1.

$$G = \sum_{i=1}^{240} g_i \quad (1)$$

Where: G is the total supply quantity from a single supplier; g_i is the supply quantity from a single supplier for the i-th week. The calculation results in the ascending order of the total supply quantity from 402 raw material suppliers over the past five years, as shown in Table 1.

Table 1: The total supply quantity from 402 raw material suppliers over the past five years

Supplier ID	Total Supply Quantity(m^3)
$S014$	28
$S015$	28
$S026$	28
$S056$	28
…	…
$S108$	240950
$S140$	302047
$S361$	328080
$S229$	354887

Through the analysis, it was found that there was a significant gap of 1095 m^3 between the total supply quantities of the 343rd and 344th suppliers when sorted by total supply quantity. Considering long-term operational and transportation cost factors, and without being influenced by other factors, the company tends to prefer suppliers with a larger total supply quantity. Therefore, based on the total supply quantity data, 60 optimal suppliers were selected with total supply quantities falling in the range of 1525 to 354,887 m^3, as shown in Table 2.

Table 2: 60 Optimal Suppliers Selected Based on Total Supply Quantity

Supplier ID	$S129$	$S023$	$S314$	$S150$	$S123$	$S266$
	$S005$	$S007$	$S154$	$S291$	$S078$	$S208$
	$S189$	$S292$	$S273$	$S114$	$S074$	$S003$
	$S218$	$S210$	$S229$	$S244$	$S086$	$S294$
	$S080$	$S346$	$S055$	$S367$	$S364$	$S338$
	$S040$	$S031$	$S365$	$S284$	$S126$	$S374$
	$S037$	$S247$	$S395$	$S307$	$S201$	$S143$
	$S352$	$S348$	$S194$	$S306$	$S268$	$S356$
	$S330$	$S308$	$S131$	$S139$	$S329$	$S275$
	$S282$	$S340$	$S151$	$S108$	$S140$	$S361$

2.3. Model Formulation and Solution

2.3.1 Quantitative Indicator Selection and Numerical Processing

The core of influencing the company's raw material procurement and transportation plan is how to select suppliers. The quantification of supplier supply characteristics is the first issue that needs to be addressed. In actual transactions, for enterprises, the type of materials supplied by suppliers and supply stability not only directly reflect the stability of the supplier's operations, but also indirectly

depict the supplier's strength. Additionally, the credibility of the supplier and their influence among similar suppliers are also important factors that cannot be ignored.

Based on the fundamental fact that companies will prioritize establishing cooperation with suppliers who are strong in strength, high in credibility, and influential among their peers, combined with the order quantity and supply quantity data of 402 raw material suppliers provided in Attachment 1 over the past five years, the main consideration is the four key indicators related to the supply characteristics of the supplier: the type of materials supplied, order fulfillment rate, supplier order stability, and enterprise procurement ratio.

1. The type of materials supplied.

As one of the primary factors for the selection of suppliers by the company, the type of materials supplied by the supplier should be given priority. The materials supplied by the supplier are classified into three levels: A, B, and C. The cost expenses for producing each cubic meter of product are used as the standard for numerical processing, based on Formula 2

$$M = P \times S \tag{2}$$

Where: M represents the production cost per cubic meter of the product; P is the unit price of raw materials; S is the quantity of various raw materials required to produce one cubic meter of the product. The cost of producing one cubic meter of the product using only A, B, or C materials can be calculated as shown in Formula 3.

$$\begin{cases} M_a = 0.72 \\ M_b = 0.726 \\ M_c = 0.72 \end{cases} \tag{3}$$

Given a certain supply volume from the supplier, the company will prioritize the purchase of Class C materials with a lower unit price. The category of materials supplied by the supplier is denoted as x_1 and is quantified as shown in Formula 4.

$$x_1 = \begin{cases} 1,\ A \\ 2,\ B \\ 3,\ C \end{cases} \tag{4}$$

2. Order Fulfillment

Based on the number of orders received from suppliers and the actual number of orders fulfilled, not all suppliers are able to fulfill their commitments by transacting with the enterprise within the specified time as agreed. Supplier order fulfillment is an important factor reflecting the credibility of suppliers, and the higher the order fulfillment rate, the higher the trust the enterprise has in the supplier, and the more willing the enterprise is to establish a stable cooperative relationship with them. Supplier order fulfillment is represented by x2, as per Formula 5.

$$x_2 = \frac{z}{d} \tag{5}$$

Whereas: z represents the total number of orders actually completed by the supplier; d denotes the total number of orders received by the supplier, to calculate the indicator value of order fulfillment for each supplier.

3. Supply Stability

Supplier supply stability reflects the operational stability of the supplier and is also an important metric for quantifying supplier delivery characteristics within a company. Based on the supply volume data from raw material suppliers, the supplier's supply stability is determined by considering the difference between the average supply volume (m^3/week) and the maximum supply volume

(m^3/week) of the raw material supplier. Supplier supply stability is represented by x_3, as defined by Equation 6.

$$x_3 = \max(g_i) - \frac{G}{f} \tag{6}$$

Where: G is the total supply volume from a single supplier; f is the total number of weeks where the supplier's supply volume is not zero; g_i represents the supply volume of the individual supplier for the i-th week. This calculation yields the indicator value for the supply stability of each supplier.

4. Enterprise Procurement Ratio

Procurement is a core component of enterprise management and represents a crucial resource for profit generation. It holds a pivotal position in the normal operations of an enterprise. The ratio of the number of times an enterprise purchases raw materials from a specific supplier to the total number of procurement occasions not only reflects the supplier's influence among similar suppliers but also demonstrates the frequency of interaction between the enterprise and that supplier. The enterprise's procurement ratio is denoted as x_4 and can be calculated using formula 7.

$$x_4 = \frac{c_m}{c_z} \tag{7}$$

Wherein, c_m represents the number of times an enterprise purchases raw materials from a specific supplier, and c_z represents the total number of procurement occasions. These values are used to calculate the enterprise's procurement ratio for various suppliers.

The results of numerical processing for the four indicators are shown in Table 3.

Table 3 Numerical Processing Results of the Four Indicators

Supplier ID	The type of materials supplied	Order Fulfillment	Supply Stability	Enterprise Procurement Ratio
	x_1	x_2	x_3	x_4
$S129$	3	0.945812808	135.0572917	0.003465862
$S023$	3	0.83125	115.7368421	0.003706768
$S314$	3	1	32.89361702	0.003795403
$S150$	2	1	559.475	0.004104489
$S123$	2	1	87.12916667	0.014656618
…	…	…	…	…

2.3.2 Indicator Validity Test

To further validate the rationality of quantitative indicator selection, an analysis was conducted on the four correlated indicators provided, namely material type, order completion rate, supply stability, and enterprise procurement ratio, in order to assess the degree of correlation between two of these indicators. This was accomplished using Formula 8.

$$\rho = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \div \sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{8}$$

The heatmap in Figure 1 displays the Pearson correlation coefficients calculated between the indicators pairwise.

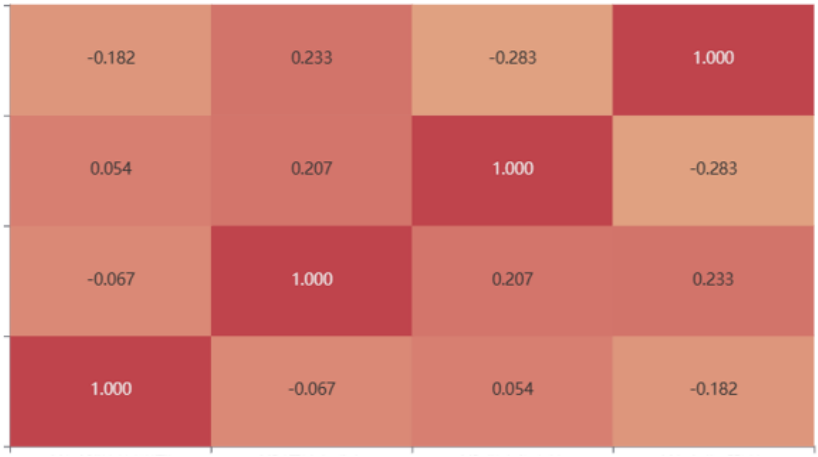

Figure 1 Correlation Coefficient Heatmap

The correlation coefficient heatmap indicates that the maximum value of |ρ| is below 0.3, suggesting a weak correlation between pairwise indicators. This implies that the four indicators, namely material type, order fulfillment rate, supply stability, and enterprise procurement ratio, have a quantifiable ability to characterize raw material supplier supply characteristics, and this finding is reasonably justified.

2.3.3 Establishment of a Model Reflecting the Importance of Ensuring Enterprise Production Based on the Entropy Weighting Method

Step 1: Standardization

$$X^* = \frac{X - \min(X)}{max(X) - \min(X)} \tag{9}$$

$$X^* = \frac{max(X) - X}{max(X) - min(X)} \tag{10}$$

Step 2: Calculating P-Values

$$P_{ij} = X_{ij}^* \div \sum_{i=1}^{n} X_{ij}^* \tag{11}$$

Step 3: Calculating Entropy (e-Entropy)

$$e_j = -k \sum_{i=1}^{n} (P_{ij} \times lnP_{ij)} \tag{12}$$

Step 4: Calculating d

$$d_j = 1 - e_j \tag{13}$$

Step 5: Calculating w to Obtain the Weights of Each Variable

$$w_j = d_j \div \sum_{i=1}^{n} d_j \tag{14}$$

Step 6: Calculating the Score

$$Score = \sum_{j=1}^{n} \left(w_j \times X_{ij}^* \right) \tag{15}$$

The weights of various indicators calculated using the entropy method and the composite scores of the supply characteristics of the top 60 better-performing companies are shown in Figure 2 and Table 4.

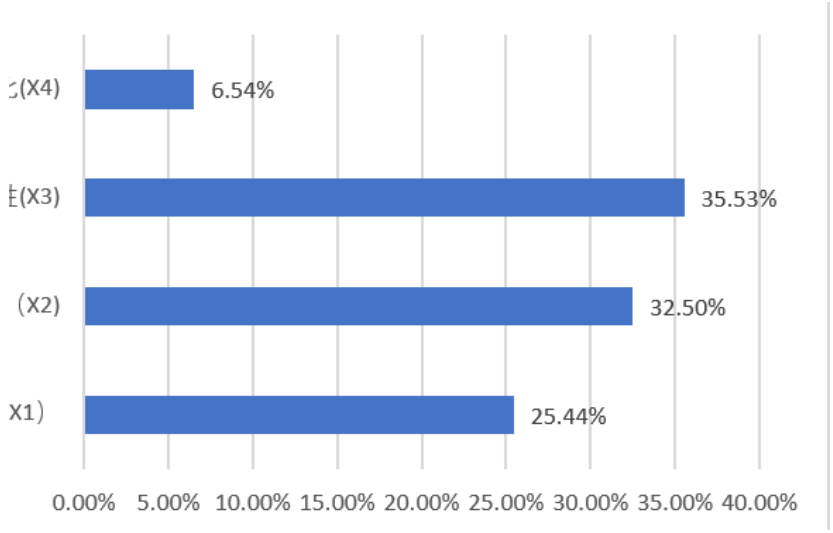

Figure 2 The weights of each indicator

Table 4 Composite scores for supply characteristics of the 60 better firms

Supplier ID	Composite scores
*S*361	0.981045379
*S*268	0.956582036
*S*306	0.95393836
…	…
*S*140	0.538586064
*S*348	0.449902366
*S*201	0.224329456

According to the comprehensive score of 60 better companies, the suppliers with high scores are prioritized, so the top 50 are taken as the 50 most important suppliers, and the results are shown in Table 5.

Table 5 50 most important suppliers

*S*361	*S*218	*S*329	*S*114	*S*340
*S*268	*S*314	*S*282	*S*291	*S*131
*S*306	*S*284	*S*352	*S*151	*S*031
*S*194	*S*086	*S*210	*S*189	*S*346
*S*356	*S*003	*S*266	*S*078	*S*040
*S*247	*S*129	*S*123	*S*005	*S*364
*S*365	*S*074	*S*007	*S*273	*S*367
*S*294	*S*037	*S*023	*S*126	*S*055
*S*080	*S*229	*S*150	*S*307	*S*338
*S*244	*S*275	*S*143	*S*374	*S*395

3. Summary

First of all, based on the original data raw material supplier supply data, according to the total supply of high and low to get a better 60 suppliers; Secondly, based on the supply and order quantity data mining to measure the supply characteristics of each supplier's indicators, and Pearson correlation analysis, to get the degree of correlation between the suppliers of the type of material supplied by the weaker degree of completion of the order of the four indicators; Finally, the use of entropy method to determine the weight of each indicator, to build a reflective guarantee of the importance of the enterprise's production of the indicator system, to get the screened 60 suppliers to

supply the characteristics of the composite score, according to the score of the descending order of the determination of the first 50 for the most important suppliers.

References

[1] Cheng Qiyun, Sun Caixin, Zhang Xiaoxing, et al. Short-Term load forecasting model and method for power system based on complementation of neural network and fuzzy logic. Transactions of China Electrotechnical Society, 2004, 19(10): 53-58.

[2] Fangfang. Research on power load forecasting based on Improved BP neural network. Harbin Institute of Technology, 2011.

[3] F. An, B. Zhao, B. Cui, et al., "DC Cascaded Energy Storage System Based on DC Collector with Gradient Descent Method," IEEE Trans. Industrial Electronics, vol. 71, no. 2, pp. 1594-1605, Feb. 2024.

[4] Amjady N. Short-term hourly load forecasting using time series modeling with peak load estimation capability. IEEE Transactions on Power Systems, 2001, 16(4): 798-805.

[5] Ma Kunlong. Short term distributed load forecasting method based on big data. Changsha: Hunan University, 2014.

[6] An, F., Zhao, B., Cui, B., Ma, Y., Zhang, X., Tang, X., & Dong, L. (2022). Asymmetric Topology Design and Quasi-Zero-Loss Switching Composite Modulation for IGCT-Based High-Capacity DC Transformer. IEEE Transactions on Power Electronics, 38(4), 4745-4759.

[7] SHI Biao, LI Yu Xia, YU Xhua, YAN Wang. Short-term load forecasting based on modified particle swarm optimizer and fuzzy neural network model. Systems Engineering-Theory and Practice, 2010, 30(1): 158-160.

[8] F. An, B. Zhao, B. Cui, Y. Chen, L. Qu, Z. Yu, R. Zeng, "Selective Virtual Synthetic Vector Embedding for Full-Range Current Harmonic Suppression of the DC Collector," IEEE Trans. Power Electronics., vol. 38, no. 2, pp. 2577-2588, Feb. 2023.

[9] Fangfang. Research on power load forecasting based on Improved BP neural network. Harbin Institute of Technology, 2011.

[10] Amjady N. Short-term hourly load forecasting using time series modeling with peak load estimation capability. IEEE Transactions on Power Systems, 2001, 16(4): 798-805.

[11] An, F., Song, W., Yang, K., Yang, S., & Ma, L. (2018). A simple power estimation with triple phase-shift control for the output parallel DAB DC–DC converters in power electronic traction transformer for railway locomotive application. IEEE Transactions on Transportation Electrification, 5(1), 299-310.

[12] Ma Kunlong. Short term distributed load forecasting method based on big data. Changsha: Hunan University, 2014.

[13] An, F., Zhao, B., Cui, B., & Bai, R. (2021). Multi-functional DC collector for future ALL-DC offshore wind power system: Concept, scheme, and implement. IEEE Transactions on Industrial Electronics, 69(8), 8134-8145.

[14] SHI Biao, LI Yu Xia, YU Xhua, YAN Wang. Short-term load forecasting based on modified particle swarm optimizer and fuzzy neural network model. Systems Engineering-Theory and Practice, 2010, 30(1): 158-160.

[15] Fangfang. Research on power load forecasting based on Improved BP neural network. Harbin Institute of Technology, 2011.

[16] Amjady N. Short-term hourly load forecasting using time series modeling with peak load estimation capability. IEEE Transactions on Power Systems, 2001, 16(4): 798-805.

[17] Ma Kunlong. Short term distributed load forecasting method based on big data. Changsha: Hunan University, 2014.

[18] SHI Biao, LI Yu Xia, YU Xhua, YAN Wang. Short-term load forecasting based on modified particle swarm optimizer and fuzzy neural network model. Systems Engineering-Theory and Practice, 2010, 30(1): 158-160.