Proceedings of the Sixth International Conference on
Railway Technology: Research, Development and Maintenance
Edited by: J. Pombo
Civil-Comp Conferences, Volume 7, Paper 23.3
Civil-Comp Press, Edinburgh, United Kingdom, 2024
ISSN: 2753-3239, doi: 10.4203/ccc.7.23.3

Railway Material Forecasting With Particle Swarm Optimization-Based Neural Network

Z. Huang[1], X. Cai[1], R. Jin[1], J. Sun[2], F. Wang[2], B. Liu[3] and M. Dai[3]

[1] School of Science and Engineering, The Chinese University of Hong Kong Shenzhen, China
[2] China Academy of Railway Sciences, China Academy of Railway Sciences Beijing, China
[3] Guangdong Sugo Technology Corporation Limited, Guangdong Sugo Technology Corporation Limited, China

Abstract

Railway material management represents a significant challenge in the daily operations of railway departments. Accurate and efficient prediction of railway material demand holds substantial practical and academic significance. This paper, based on a large amount of real railway material data provided by one large railway company in China, constructs a prediction model using a particle swarm optimization-based neural network to forecast railway material demand. The study finds that the particle swarm optimization-based neural network model not only possesses a high prediction accuracy but also exhibits strong generalization capabilities, making it particularly suitable for railway material data characterized by high discreteness and randomness.

Keywords: railway material, prediction model, particle swarm optimization, particle swarm optimization-based neural network, validation, forecasting.

1 Introduction

With the rapid development of China's railway transportation industry, efficient and accurate prediction of railway material demand has become increasingly important for ensuring the stable operation and long-term development of the railway system. Railway materials, including rails, indicators, spring plates, radiator assemblies, battery signs, lamp head connectors, etc., are indispensable components of daily railway department operations. This paper aims to explore the significance of railway material prediction in railway department operations and propose a design scheme for a prediction model based on the particle swarm optimization (PSO) algorithm, in hopes of providing more efficient and accurate material management and allocation plans for railway departments.

Firstly, the timely supply of railway materials is key to ensuring the safety of railway operations. A railway system lacking necessary materials may face serious safety risks. Secondly, the accuracy of material prediction directly affects the efficiency of railway operations. Moreover, accurate material prediction also helps optimize inventory management, reducing stockpiling and shortages, thus lowering storage costs and enhancing the efficiency of capital utilization.

Currently, there are many difficulties in the operation and management of railway materials in China. The first is uncertainty and deviation. Due to the long cycle in demand reporting and tender procurement, the demand is highly uncertain. The second is the issue of inventory capital occupation. In the absence of scientific demand forecasting, material management personnel often store an excessive amount of rarely used spare parts, leading to a large amount of capital being tied up in inventory, causing unnecessary financial waste. Lastly, the management demand is complex and cumbersome. Due to the lack of scientific data support in demand collection, review, and other links, the entire process is enormous and very complex, and the accuracy of demand determined based on experience is also difficult to satisfy.

In summary, railway material demand prediction plays a vital role in ensuring railway safety, improving operational efficiency, reducing costs, and supporting sustainable development. In view of this, this paper proposes a railway material demand prediction model based on a particle swarm optimization-based back propagation neural network (PSO-BPNN), which is especially suitable for the characteristics of railway material data with its high solution efficiency, strong generalization, and difficulty in falling into local optima. This paper aims to provide more effective decision-making support for railway departments by improving the accuracy of predictions.

The subsequent sections of this paper are as follows. Section 2 is a literature review. Section 3 explains the model methods. Section 4 displays data, model construction, parameter tuning, and comparison of results with different methods. Section 5 is a summary.

2 Literature Review

Railway material forecasting is essential for ensuring the efficient and safe operation of railway systems. Past scholarly research on railway material forecasting can be mainly categorized into the following areas.

First, the evolution of methods in railway material forecasting. As a key aspect of railway system management, the methods for railway material forecasting have undergone significant evolution over the past decades. Initially relying on empirical judgment and basic statistical techniques, recent years have seen an adoption of advanced forecasting models. For instance, [1] discussed the planning of resource maintenance in railway bridge construction, including the forecasting of material needs and delivery schedule formulation. Moreover, [2] demonstrated the application of deep learning methods in predicting the moisture of high-speed railway subgrade materials, marking a step forward in railway material forecasting towards advanced data analysis and machine learning technologies.

Second, material research and environmental sustainability. Railway material forecasting is not limited to quantifying demand and supply but also involves the selection and use of materials. [3] highlighted the application of modern high-strength alloy steels in railway transportation, emphasizing the importance of material innovation in reducing vehicle weight and enhancing efficiency. Similarly, [4] analyzed the recyclability and recoverability of different types of railway vehicles, underscoring the importance of material recycling in railway systems for profound environmental sustainability impacts.

Third, safety and risk management. Another critical aspect of railway material forecasting is safety and risk management. [5] provided a risk analysis of transporting explosives by rail, revealing potential safety risks in high population density areas. [6] focused on predicting fire risks of flammable items in railway transport, advocating for enhanced fire prevention measures. These studies show that safety and risk management are indispensable components of railway material forecasting.

Fourth, practical applications and future trends. Case studies of practical applications of railway material forecasting provide a bridge between theoretical research and practice. The railway wheel wear forecasting method based on universal Kriging estimation proposed by Cremona et al. demonstrated the effectiveness of forecasting techniques in real-world applications [7]. [8] explored the application of remote sensing materials and geographic information technology in railway design, indicating the potential of modern technologies in railway planning and management. These studies not only showcase current practices but also offer insights into future trends in railway material forecasting.

Lastly, beyond the railway sector, many industries face challenges in material and product forecasting. [9] proposed a material demand forecasting model based on the convolutional neural network (CNN) algorithm, offering significant references for businesses to improve efficiency and promote development. Additionally, [10] introduced grey relational analysis into grey forecasting models for predicting magnesium

material demand. Furthermore, [11] presented a PSO-BPNN model for aircraft material demand forecasting, achieving commendable results.

In summary, the field of railway material forecasting has transitioned from traditional methods to modern technologies, encompassing material research, environmental sustainability, safety, and risk management. With the continual development of technology, forecasting methods are becoming more precise and efficient, providing strong support for the optimized management of railway systems.

3 Method

3.1 Particle Swarm Optimization (PSO)

The PSO algorithm is a group intelligence-based optimization technique proposed by Eberhart and Kennedy in 1995 [12]. It is inspired by the social behavior of bird flocks or fish schools. In PSO, each potential solution to an optimization problem is considered a 'particle' in the search space. Each particle has its own position and velocity, which are adjusted based on the particle's own experience and the experiences of other particles in the group.

The particle swarm algorithm assumes the existence of N particles (population size), each with K-dimensional attributes (number of features). For each particle i, its position in the K-dimensional space is represented as a vector $X_i = (x_1, x_2, \cdots, x_K)$, and its flight velocity as a vector $V_i = (v_1, v_2, \cdots, v_K)$. Each particle has a fitness value determined by the objective function and is aware of the best position it has discovered so far (pbest) and its current position X_i. This can be seen as the particle's own flying experience. In addition, each particle knows the best position discovered so far by all particles in the group (gbest), which can be considered as the experience of its companions. The particle decides its next move based on its own experience and the best experience among its companions.

PSO is initialized with N random particles (random solutions). Then, it iterates to find the optimal solution. In each iteration, particle i updates itself by tracking two 'extremes' (pbest, gbest). After finding these two optimal values, the particle updates its velocity and position using the following formulas:

$$v_i = w \cdot v_i + c_1 \cdot rand() \cdot (pbest_i - x_i) + c_2 \cdot rand() \cdot (gbest_i - x_i) \tag{1}$$

$$x_{i+1} = x_i + v_i \tag{2}$$

where v_i is the particle's velocity, x_i is the particle's position, $rand()$ is a random number between (0,1), c_1 is the self-cognitive factor, c_2 is the social-cognitive factor, w is the inertia weight, and c_1, c_2, w are all greater than 0. Conventionally, c_1 and c_2 are generally set between (0,2), and w is typically set between (0.5,0.9).

The flowchart for the standard PSO algorithm is as follows.

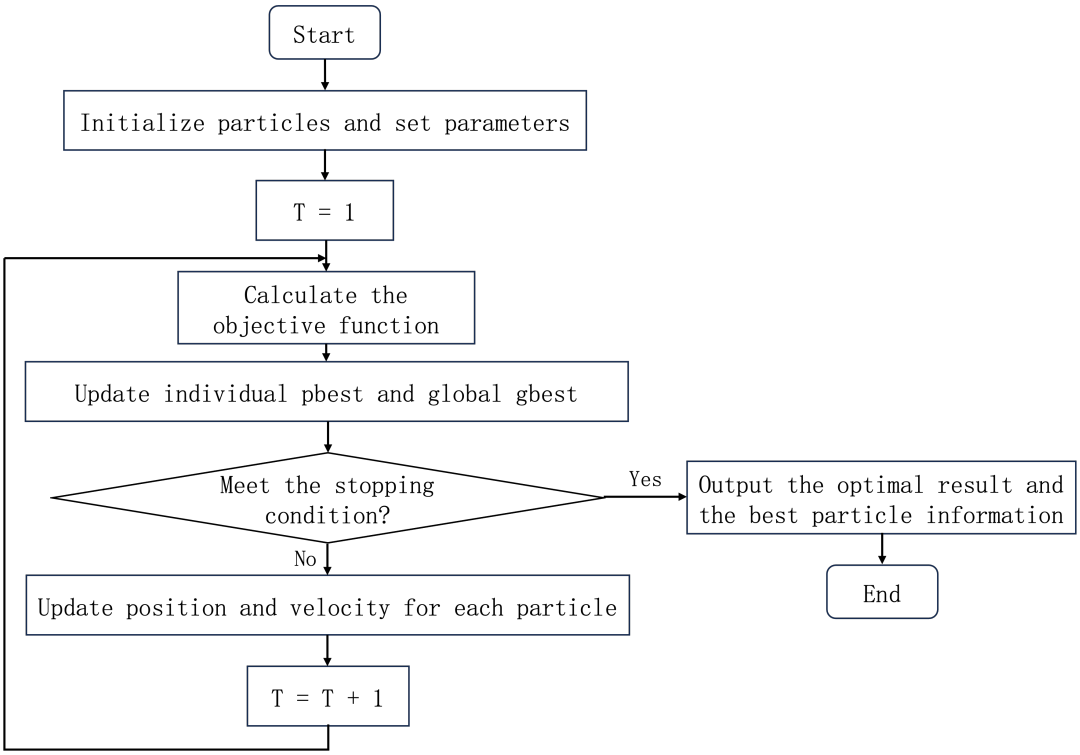

Figure 1: PSO algorithm flow chart.

3.2 Back Propagation Neural Network (BPNN)

Neural Networks (NN) are computational models that mimic the workings of the human brain, used to solve a variety of complex pattern recognition and prediction problems. They consist of a large number of nodes (or neurons), typically arranged in layers. Each neuron is connected to multiple other neurons through weights, simulating the synapses in biological neural systems. The Backpropagation Neural Network (BPNN) is a type of multi-layer feedforward neural network, trained using the backpropagation algorithm, primarily used in supervised learning scenarios. It includes an input layer, one or more hidden layers, and an output layer.

The working principle of BPNN includes two parts. First is the forward propagation part. Input signals propagate from the input layer through the hidden layers to the output layer within the neural network. Each neuron receives weighted inputs from neurons of the previous layer, then processes them through an activation function to produce an output. This is expressed by the formula:

$$a^{(l+1)} = f(W^{(l)} \cdot a^{(l)} + b^{(l)}) \tag{3}$$

where $a^{(l)}$ is the activation value of the l-th layer, $W^{(l)}$ and $b^{(l)}$ are the weights and biases of the l-th layer, respectively, and f is the activation function.

The second part is the backward propagation. Neural networks are trained through the backpropagation algorithm, which involves adjusting weights and biases to mini-

mize the difference between the predicted output and actual output. The loss function L is used to compute the error, a common one being mean squared error (MSE):

$$L = \frac{1}{n}\sum(y_{pred} - y_{true})^2 \tag{4}$$

where n is the number of samples, y_{pred} is the network's predicted value, and y_{true} is the actual value.

3.3 Particle Swarm Optimization Back Propagation Neural Networks (PSO-BPNN)

The training process of BPNN depends on the selection of initial weights. However, as the network weights are randomly initialized, this can lead to randomness in training and a propensity to fall into local optima, resulting in low reliability and poor generalization ability of the trained network. A PSO-BPNN utilizes the particle swarm algorithm to optimize the initial weights of the network. The particle swarm algorithm is simple to operate and highly efficient in optimization, characterized by its global optimization and high computational precision. Compared to other metaheuristic algorithms (such as genetic glgorithms, GA), the PSO algorithm is simpler, requiring less time per iteration, which is particularly important for larger datasets. Therefore, the particle swarm algorithm can be used to optimize BPNN, enhancing the network's global search capability and generalization ability, thus making the optimized network more accurate in prediction.

The basic steps for training a neural network using PSO are as follows:

1. Set the PSO hyperparameters, including c_1, c_2, w, the number of particles N, the number of iterations, etc., as well as the neural network hyperparameters, including the number of network layers, the number of neurons in each layer, activation functions.

2. Calculate the total number of weight parameters K of the neural network, use K as the dimension of the particle swarm space, and randomly initialize N K-dimensional particles.

3. Calculate the fitness value of each particle (using MSE as the metric), and determine pbest and gbest.

4. Update the particles' velocity and position using formulas and recalculate the new fitness value.

5. Repeat steps 3 and 4 until the iteration stop condition is met. Retain the optimal particle feature parameters as the optimal neural network weights.

6. Use the optimal neural network weights for evaluation and prediction.

The flowchart for PSO-BPNN is as follows.

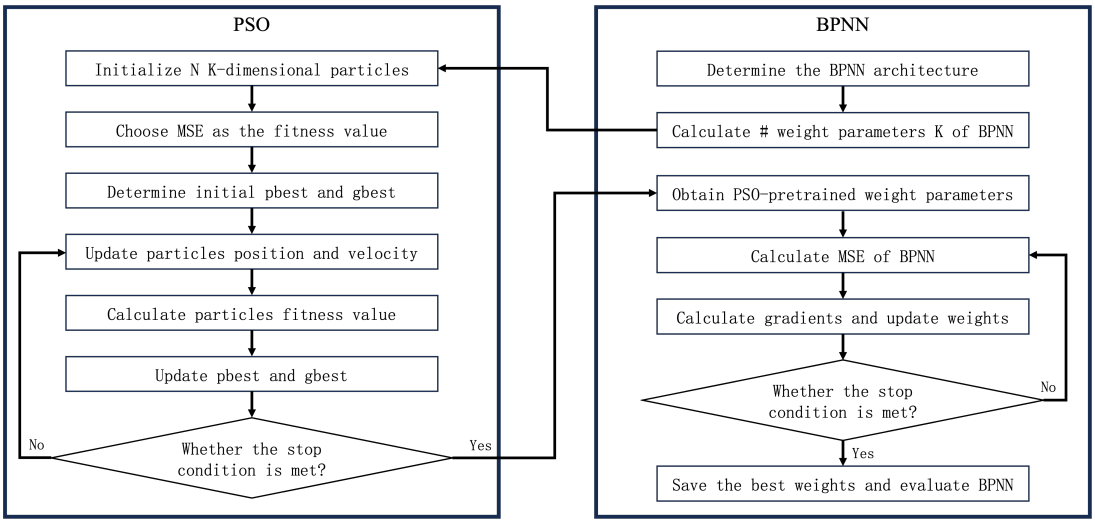

Figure 2: PSO-BPNN model flow chart.

4 Experiments

4.1 Data

We collected daily operational material outflow data from vehicle and locomotive depots under the Chinese railway group, including major provinces in China, like Guangdong and Hunan. The data span from January 2019 to December 2023. After data processing and cleaning, we obtained 21,032 records, which are used for model analysis. In the analyzed data, the 'Quantity' field is the dependent variable (y), while the 'Material Number' and 'Order Date' fields are the independent variables. The analyzed data is shown in the following table.

Material Number	Order Date	Quantity
910000010043	2019-01-09	6.0
910000010043	2019-01-16	4.0
910000010043	2019-01-18	4.0
910000010043	2019-01-23	2.0
910000010043	2019-01-29	4.0
...	...	...
945300200047	2023-04-14	4.0
945300200047	2023-04-23	4.0
945300200047	2023-04-28	3.0
945300200047	2023-05-06	4.0
945300200047	2023-05-24	1.0

Table 1: Railway operational material outflow data.

After one-hot encoding of categorical variables and considering previous periods of

'Quantity' as features (lag order = 15), we ultimately yield a combination of numerical variables and binary variables, totaling over a hundred dimensions as features X (in this case, the dimensionality of X is 163, i.e., $dom(X) = 163$).

Finally, we take the most recent two months of data (from November 1, 2023, to December 30, 2023) as the prediction set. From the remaining data, 80% is used as the training set and 20% as the validation set. Consequently, the sample sizes for the training set, validation set, and prediction set are 15,093, 3,774, and 665, respectively.

4.2 Model Construction

The experiment findings indicate that a neural network comprising three hidden layers is already capable of adequately performing the task of material forecasting. As a result, a five-layer fully connected neural network is constructed, with the input layer dimension equal to the feature X dimension. The first hidden layer contains 100 neurons, the second hidden layer has 50 neurons, and the third hidden layer comprises 25 neurons. Except for the output layer, which employs the LeakyReLu activation function to ensure non-negative network output values, all other layers utilize ReLu as the activation function.

We use MSE as the Loss function and calculate the total number of neural network weight parameters K as follows:

$$K = (n_{input} + 1) * 100 + (100 + 1) * 50 + (50 + 1) * 25 + (25 + 1) * n_{output} \quad (5)$$

where n_{input} and n_{output} mean the number of neurons in the input and output layers of network. In this case, we have $n_{input} = dom(X) = 163$, $n_{output} = 1$ and $K = 22751$.

4.3 Parameters Tuning

Based on experience, we set the particle swarm parameter w to follow a uniform distribution between 0.5 and 0.9, and $N = 100$. The parameters c_1 and c_2 are found through grid search with a step of 0.2 within the interval (0.2,1.2) to find the optimal parameter values. After several comparative experiments, we set the number of PSO iterations to 300 (training 300 times is sufficient to reach convergence). Additionally, the weights pretrained by PSO are directly used as the final weights of the neural network (continuing to train with gradient descent on the basis of PSO pretraining not only does not improve model performance but may even worsen the results). We also calculate the mean absolute percentage error (MAPE) for both the validation set and the test set to measure the performance of different parameter combinations as shown in the following table.

We found that the model performs optimally when the (c_1, c_2) parameter combination is set to either (0.6,0.8) or (1,0.2), i.e., validation MAPE is the lowest. Since the model yields satisfactory results when parameters are around (1,0.2)—given that the validation MAPE values for (0.4,0.2), (0.6,0.2), and (0.8,0.2) are relatively low—we choose (1,0.2) as the final setting for the PSO hyperparameters.

c_1	c_2	Validation MAPE	Testing MAPE
0.2	1.2	0.76	0.87
0.4	1.2	0.92	1.00
0.6	1.2	0.97	1.04
0.8	1.2	0.96	1.07
1.0	1.2	0.67	0.89
1.2	1.2	1.00	1.10
0.2	1.0	0.89	1.04
0.4	1.0	0.67	0.84
0.6	1.0	0.81	0.97
0.8	1.0	0.67	0.85
1.0	1.0	0.66	0.86
1.2	1.0	0.68	0.86
0.2	0.8	0.87	1.02
0.4	0.8	0.85	0.95
0.6	0.8	**0.64**	**0.84**
0.8	0.8	0.75	0.95
1.0	0.8	0.90	1.06
1.2	0.8	0.67	0.88
0.2	0.6	0.98	1.06
0.4	0.6	0.66	0.85
0.6	0.6	0.78	0.91
0.8	0.6	0.67	0.86
1.0	0.6	0.68	0.83
1.2	0.6	0.74	0.89
0.2	0.4	0.75	0.95
0.4	0.4	0.74	0.85
0.6	0.4	0.86	0.98
0.8	0.4	0.65	0.85
1.0	0.4	0.67	0.79
1.2	0.4	0.68	0.85
0.2	0.2	0.87	1.03
0.4	0.2	0.67	0.82
0.6	0.2	0.68	0.82
0.8	0.2	0.69	0.86
1.0	0.2	**0.64**	**0.84**
1.2	0.2	0.73	0.91

Table 2: Performance of PSO-BPNN with various c_1 and c_2 parameters.

4.4 Comparisons between Different Models

To ensure the reliability of our model results, we conducted both vertical (with SGD-trained, PSO+SGD-trained BPNN) and horizontal (with models from different do-

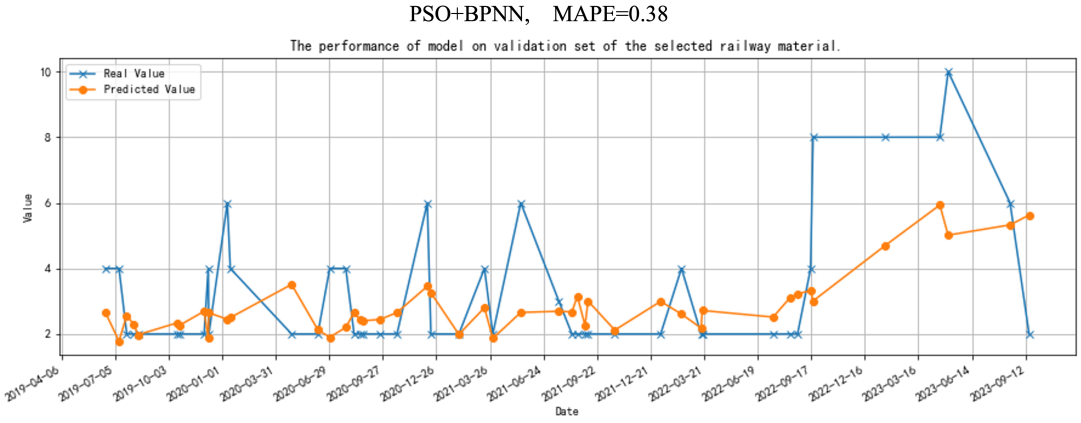

Figure 3: The fitting performance of the PSO+BPNN model on the validation set for a randomly selected material item.

mains) comparisons of the performance of various forecasting methods. Specifically, we compared traditional machine learning methods (SVM, Random Forest, XGBoost), conventional deep learning methods (RNN, LSTM), traditional time series methods (Panel Data Regression), and the same neural network architecture but trained using traditional methods (SGD, PSO+SGD) in terms of MAPE on both the validation and testing sets, as shown in the table below.

Model	Validation MAPE	Testing MAPE
PSO+BPNN (Ours)	**0.64**	**0.84**
SGD+BPNN	1.06	1.25
PSO+SGD+BPNN	0.94	1.22
SVM	0.68	0.78
Random Forest	0.66	0.89
XGBoost	0.67	0.90
Panel Data Regression	0.67	0.84
RNN	0.88	1.26
LSTM	1.10	1.19

Table 3: Comparison of forecasting methods on validation and testing sets.

From the table above, it is evident that our PSO+BPNN model has the highest prediction accuracy. The performance of pretraining with PSO followed by further training with SGD is significantly inferior to training exclusively with PSO. This demonstrates that PSO has stronger generalization capabilities, while SGD may lead to poorer model predictions due to overfitting issues.

To more vividly display the results of different models, we randomly select an item of materials and compare the fitting effects of various methods on the validation set for this item. The results are shown in the following two figures.

Figure 4: The fitting performance of other methods on the validation set for the same material item.

From Figure 3 and Figure 4, it is evident that our PSO+BPNN model not only has the highest fitting accuracy (MAPE=0.38) but also shows fluctuation trends most consistent with the real trends. Among other models, both XGBoost and random forest models show fluctuation trends that are relatively consistent with the actual trends, but their MAPE values are higher than our model. The MAPE of the SVM model (MAPE=0.39) is closest to that of our model, yet its fluctuation trend significantly differs from the real trend.

5 Conclusion

This paper investigates the demand forecasting problem in the railway material domain, employing the PSO+BPNN method to study the material outflow data of various regions from one large railway company in China. The study finds that, compared to traditional gradient-based optimization methods, PSO exhibits stronger generalization capabilities due to its metaheuristic search characteristics, is less prone to falling into local optima, and has a higher optimization efficiency. Through horizontal and vertical comparisons of different forecasting methods, we find that PSO+BPNN achieves the highest prediction accuracy, and the overall time series it predicts aligns more closely with the actual time series. This proves the practical value of PSO+BPNN in the field of railway material forecasting.

Acknowledgment

The author would like to extend sincere gratitude to the China Academy of Railway Sciences Corporation Limited for their support of the research project (No. 2022YJ352).

References

[1] M. Klykov, N. Grigoryev, and A. Zemskov, "Planning of resource maintenance at building of railway bridge," vol. 106, p. 05002, 2017.

[2] L. Chen, J. Chen, C. Wang, Y. Dai, R. Guo, and Q. Huang, "Modeling of moisture content of subgrade materials in high-speed railway using a deep learning method," *Advances in Materials Science and Engineering*, 2021.

[3] W. Stawecki, M. Bajerlein, and P. Daszkiewicz, "Modern high-resistance alloy steel used in railway transport," *Combustion Engines*, 2017.

[4] M. Delogu, F. D. Pero, L. Berzi, M. Pierini, and D. Bonaffini, "End-of-life in the railway sector: Analysis of recyclability and recoverability for different vehicle case studies.," *Waste management*, vol. 60, pp. 439–450, 2017.

[5] J.-H. Lee, D. Song, and S. kyung Lee, "The risk analysis for the rail transport of explosives," vol. 15, pp. 33–39, 2011.

[6] L. Koroleva, A. G. Khaidarov, G. K. Ivakhnuk, T. Antoshina, and A. A. Boyakhchyan, "Comparative analysis of methods for determining the flammability of municipal solid waste as railway transport cargo," *IOP Conference Series: Earth and Environmental Science*, vol. 938, 2021.

[7] M. A. Cremona, B. Liu, Y. Hu, S. Bruni, and R. Lewis, "Predicting railway wheel wear under uncertainty of wear coefficient, using universal kriging," *Reliab. Eng. Syst. Saf.*, vol. 154, pp. 49–59, 2016.

[8] O. Andreeva, N. Konon, and M. Ratinski, "On the use of remote sensing in the design of railways," *Geodesy and Cartography*, 2019.

[9] Z. Tang and Y. Ge, "Cnn model optimization and intelligent balance model for material demand forecast," *International Journal of System Assurance Engineering and Management*, vol. 13, no. Suppl 3, pp. 978–986, 2022.

[10] Y.-C. Hu, P. Jiang, Y.-J. Chiu, and Y.-W. Ken, "Incorporating grey relational analysis into grey prediction models to forecast the demand for magnesium materials," *Cybernetics and Systems*, vol. 52, no. 6, pp. 522–532, 2021.

[11] X. Chen, T. Wang, and W. Liang, "General aircraft material demand forecast based on pso-bp neural network," *International Journal of Control and Automation*, vol. 9, no. 5, pp. 407–418, 2016.

[12] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95-international conference on neural networks*, vol. 4, pp. 1942–1948, IEEE, 1995.