

Analysis of Mathematical Correlation Model between Engineering Parameters and Costs

Yingqian Tian

School of Computer Science and Technology, Huaibei Normal University, Huaibei, 235000, China

Abstract: This article studies the mathematical correlation model between engineering parameters and costs, and explores its core role in engineering project management. By analyzing the functional relationship between engineering parameters such as material usage, construction period, equipment specifications, and direct costs, indirect costs, and risk costs, a scientific basis is provided for project budget preparation, cost control, and decision optimization. Research has shown that the popularity of BIM technology and popularization of the Internet provide new possibilities for improving model accuracy.

Keywords: Engineering parameter, costs, BIM technology, Function relationship

1. Introduction

The study of mathematical correlation models between engineering parameters and costs is one of the core topics in engineering project management. Studying the functional relationship between quantitative analysis of engineering parameters (such as material usage, construction period, equipment specifications, etc.) and costs (including direct costs, indirect costs, and risk costs) can provide scientific basis for project budget preparation, cost control, and decision optimization. In fields such as architecture, manufacturing, and energy, the application of this correlation model can significantly improve resource allocation efficiency [1], for example, by adjusting concrete strength parameters to reduce material loss, or optimizing construction cycle parameters to reduce labor costs. Its research value is not only reflected in the economic benefits, but also in the transformation of management mode from experience driven to data-driven through dynamic mapping of parameters and costs.

Currently, with the popularization of BIM technology and the Internet of Things, real-time data collection provides new possibilities for improving model accuracy [2]. The research on the mathematical correlation model between engineering parameters and costs mainly revolves around three core aspects: firstly, the classification of parameter types is the foundation, including material parameters (such as steel strength, concrete grade), process parameters (such as welding temperature, pouring speed), and environmental parameters (such as geological conditions, climate factors), which directly affect material consumption, labor efficiency, and equipment loss. Secondly, in the analysis of cost composition, direct costs (such as raw material procurement and machinery leasing) are explicitly related to parameters, such as the linear relationship between concrete usage and structural volume; Indirect costs (such as management fees and regulatory fees) indirectly affect the total cost through the coupling effect between parameters. Finally, the correlation characteristics manifest as nonlinearity and dynamism: on the one hand, parameter changes may lead to exponential cost growth (such as an increase in wind load parameters for super high-rise buildings causing a sharp rise in structural costs); On the other hand, extending the construction period will amplify the total expenditure through the cumulative effect of time cost. This complexity requires the model to take into account the interaction between parameters and the interference of external variables (such as market price fluctuations) [3].

In the mathematical modeling of the relationship between engineering parameters and costs, multiple regression analysis, grey prediction model, and parameter optimization algorithm are three typical methods. Multiple regression analysis [4] establishes linear or nonlinear equations between cost and multiple engineering parameters (such as $C = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$), which is suitable for scenarios with strong independence between parameters, such as the direct correlation between material usage and labor costs. The grey prediction model [5] targets small sample data and generates weakened randomness through accumulation, making it particularly suitable for dynamic cost prediction of time series parameters such as construction period. Parameter optimization algorithm [6] uses genetic algorithm, particle swarm optimization and other methods to solve the extreme value of the cost function, such as balancing procurement cost and energy consumption parameters in equipment selection, to achieve the minimization of the entire life cycle cost. The application of these methods needs to be combined with data characteristics: regression analysis relies on the quality of historical data, grey models require high regularity of data, and optimization algorithms require clear constraints (such as project duration and safety standards).

In practical engineering, mixed models are often used to improve accuracy, such as combining the short-term cost trend of grey prediction with the sensitivity of regression analysis parameters to form a hierarchical prediction framework. For example, the study performed by Arian and Carine is based on greenhouse gas (GHG) emission data from Norwegian apartment buildings and proposes a three-level benchmark value system (target value, reference value, limit value) aimed at providing emission standards that meet national requirements and higher emission reduction targets for new multi-family residential buildings. The benchmark value is presented in two functional units: unit building area (GFA) and usable area (UFA). For example, the target value is 270 kg CO₂ eq./m² GFA (or 340 kg CO₂ eq./m² UFA), the reference value is 290 kg CO₂ eq./m² GFA (or 360 kg CO₂ eq./m² UFA), and the limit value is 350 kg CO₂ eq./m² GFA (or 440 kg CO₂ eq./m² UFA). The study also extended the full lifecycle (WLC) emission baseline, covering building elements 21-28 with a limit of 370 kg CO₂ eq./m² GFA (460 kg CO₂ eq./m² UFA), and if technical/electrical installation is included, the limit is raised to 450 kg CO₂ eq./m² GFA (560 kg CO₂ eq./m² UFA). These benchmark values can provide early emission reduction decision-making basis for designers and policy makers, and support the Norwegian construction industry in gradually achieving its 2030 and 2050 climate goals [7].

These cases validate the applicability of parametric models in complex engineering: they can capture the explicit correlation between material usage and cost, as well as reveal the implicit impact of parameters such as project duration on indirect costs. The study of mathematical correlation models between engineering parameters and costs still faces multiple challenges. Firstly, data quality and integrity are the core bottlenecks, for example, key parameter records may be missing in historical engineering data (such as construction parameters under special geological conditions), or there may be human errors in filling in, resulting in insufficient representativeness of model training samples. Secondly, the complexity of the interaction between parameters has not been fully resolved, such as the coupling effect between material parameters (such as steel strength) and process parameters (such as welding temperature), which may cause nonlinear cost changes. Existing models are mostly based on independent variable assumptions, making it difficult to accurately capture such correlations. In addition, the adaptability to dynamic environments is insufficient. When market prices fluctuate or policies are adjusted (such as upgrading environmental standards), static models are difficult to update parameter weights in real time, leading to prediction bias. Future research can focus on three breakthroughs: firstly, integrating the Internet of Things and BIM technology to achieve real-time parameter collection, constructing dynamic databases to improve data timeliness; The second is to introduce machine learning algorithms (such as neural networks) to handle the nonlinear relationships between high-dimensional parameters, enhancing the model's self-learning ability; The third is to develop a hybrid prediction framework that combines the small sample advantage of grey models with the parameter sensitivity of regression analysis, while embedding risk factor modules (such as supply chain interruption probability) to enhance robustness. These directions are expected to promote the evolution of the model from "post analysis" to "prediction in advance regulation in the event" closed-loop management.

2. The Commonly-used Models

2.1 Grey prediction model

The grey prediction model (GM model) [8] is a small sample prediction method based on grey system theory, suitable for predicting dynamic systems with incomplete information and limited data. Its core is to process raw data through cumulative generation (cumulative sequence), reduce randomness, and explore potential patterns, ultimately establishing a differential equation model for trend prediction.

The most basic first-order univariate model, in the form of

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b. \quad (1)$$

By accumulating and generating sequences to solve parameters a (development coefficient) and b (grey application amount), it is suitable for short-term forecasting.

2.2 Regularized dynamic parameter estimation

One can determine the optimal time distribution $p(t)$ of the parameters for each experimental dataset, so that the predictions of model (1) best match the corresponding experimental measurements. This is expressed as the following dynamic optimization problem, solved independently for each dataset [9]:

$$\begin{aligned}
& \min_{x, y, p, z, x_0} \sum_{j \in \{1, \dots, N_{meas}\}} (z_j - \%z_j)^T \cdot W_j^z \cdot (z_j - \%z_j) + R(p(t)), \\
& s.t. \frac{dx}{dt}(t) = f(x(t), y(t), u(t), p(t)), \forall t \in \tau \\
& \quad 0 = g(x(t), y(t), u(t), p(t)), \forall t \in \tau \\
& \quad x(t_1^{meas}) = x_0 \\
& \quad z_j = h(x(t_j), y(t_j)), \forall t_j \in \tau^{meas} \\
& \quad 0 \geq c(x(t), y(t), u(t), p(t)), \forall t \in \%
\end{aligned} \tag{2}$$

Here, $\tau = [t_1^{meas}, t_{N_{meas}}^{meas}]$.

2.3 Static correlation model

The improved Cobb Douglas production function is an extension or optimization of the classical Cobb Douglas production function, aimed at addressing the limitations of the original model in specific application scenarios. The following are its core improvement directions and typical applications.

The original form is:

$$Y = A \cdot L^\alpha \cdot K^\beta \tag{3}$$

Where Y denotes output volume. A denotes technical level (total factor productivity), L is the Labor input and K is the Capital investment. α and β are the output elasticity coefficient of labor and capital, usually satisfying the characteristic of return to scale.

Improved form usually replaces constant A with a time variable $A(t)$, for example

$$Y = A(t) \cdot L^\alpha \cdot K^\beta \tag{4}$$

With $A(t) = A_0 \cdot e^{\gamma t}$. γ is the technological progress rate. The purpose is to reflect the dynamic impact of technological progress on output and better fit long-term economic growth analysis.

2.4 Dynamic prediction model

Constructing LSTM Neural Network Time Series Prediction Framework has two key steps.

The first one is Model construction logic

- **Input layer:** Engineering parameters (material strength, construction period, environmental factors, etc.)
- **Processing layer:** mathematical transformations and relationship mining
- **Output layer:** Cost forecast values and optimization suggestions

The second one is key technical methods which include Multiple linear regression model, Machine learning models and Dynamic Cost Model.

(1) The formula of the Multiple linear regression model is

$$\text{Cost} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon \tag{5}$$

Where X_i is the engineering parameters, β_i is the regression coefficient.

- (2) Machine learning models include the random forest and Support Vector Machine. The random forest can handle nonlinear relationships and feature importance analysis. The support vector machine has high precision prediction on small sample data.
- (3) Dynamic Cost Model can introduce time series analysis (ARIMA) to address the impact of schedule fluctuations and establish parameter cost partial differential equations to describe dynamic changes.

3. Conclusions

This article systematically studies the mathematical correlation model between engineering parameters and costs, and draws the following core conclusions and future research directions:

3.1 Model innovation and algorithm development

Dynamic prediction capability: Establish a parameter cost model based on power law relationships, develop a cluster of dynamic cost prediction algorithms that integrate time series analysis (such as ARIMA) and

machine learning (such as LSTM neural networks) to cope with dynamic disturbances such as market fluctuations and schedule changes.

Advantages of hybrid modeling: The effectiveness of hybrid models (such as combining grey prediction with regression analysis) in improving prediction accuracy has been verified, especially when dealing with small sample data and nonlinear relationships.

3.2 Standardization and tool application

Parameter Management Standards: Propose a draft of engineering parameter classification and management standards, clarify the quantitative specifications for material parameters (such as steel strength), process parameters (such as welding temperature), and environmental parameters (such as geological conditions), and provide a unified reference for the industry.

Development of analysis tools: Construct parameter sensitivity analysis matrix and cost warning threshold system to assist project managers in identifying key cost drivers and formulating optimization strategies.

3.3 Technology integration and data-driven approach

Integration of Internet of Things and BIM: Through real-time data collection technology (such as IoT sensors) and BIM platform, dynamic updates of engineering parameters and adaptive adjustment of models are achieved, solving the problem of static model lag.

Deepening the application of machine learning: Introducing algorithms such as neural networks and random forests to handle high-dimensional parameter interactions, enhancing the model's analytical ability for nonlinear relationships (such as material process coupling effects).

3.4 Challenges and Future Directions

Data quality improvement: It is necessary to address the issues of missing historical data and manual input errors. It is recommended to establish a full lifecycle engineering database.

Dynamic environmental adaptability: Develop a hybrid prediction framework that embeds risk factors (such as supply chain interruption probability) to enhance the model's real-time response capability to policy adjustments and market price fluctuations.

Transformation of closed-loop management: Promote the evolution of models from "post analysis" to "prediction regulation" closed-loop, ultimately achieving the intelligence and automation of engineering cost management.

This study provides theoretical support and methodological tools for cost optimization of engineering projects, and in the future, interdisciplinary cooperation can further expand the application boundaries of the model in fields such as green building and intelligent manufacturing.

Acknowledgements

This work was partially supported by the Horizontal Scientific Research Funds in Huaibei Normal University (No. 22100415).

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1]. Peng B, Zhang N. Research on the seismic behavior of subway station crossed by overpass. *Journal of Railway Engineering Society*, 34(1),118-123, (2017).
- [2]. Juan M, Luigi B, Mattia P, Juan E. Nieto-Julián, Optimizing best-fit algorithms for complex cross-vault geometries in HBIM generation using point cloud data, *Automation in Construction*, 176, 106274, (2025).
- [3]. Dariia B, Maria S, Andrii B, Olexandr I, Ihor M. Numerical methods of calculation of stress–strain state of combined elements of sheeting, *Journal of Mechanics*, 41, 82–102, (2025).
- [4]. Muhammad A, Hafiz M. Applications of machine learning and traditional prediction techniques for estimating resilient modulus values of unbound granular materials incorporating reclaimed asphalt pavement (RAP) as a primary component, *Construction and Building Materials*, 473, 140900, (2025).
- [5]. ZengB, TuoY. Multivariate grey prediction model with fractional time-lag parameter and its application, *Applied Mathematical Modelling*, 141, 115940, (2025).

- [6]. Marcel N, Claudiu-Ionel N, Dan S, Dorin Ş. Sensorless PMSM control system based on multi-agent reinforcement learning and fractional order LADRC for rapid control prototyping, Alexandria Engineering Journal, 130, 292-310, (2025).
- [7]. Arian L, Carine L. Setting benchmark values for apartment blocks in Norway as the basis for GHG emission reduction along the building process, Building and Environment, 280,113127, (2025).
- [8]. GouX, Mi C, YangY, ZengB. A nonlinear mixed-frequency grey prediction model with two-stage lag parameter optimization and its application. Applied Mathematical Modelling, 150, 116360, 2026.
- [9]. Adrian C, Thomas B, Sarah F, Daniel L, Maarten N, Franzisko W, Merle W, Constantinos C. Dynamic hybrid modeling: Incremental identification and model predictive control. Computers & Chemical Engineering, Volume 204, 109413, (2026).