Journal of Geo-Energy and Environment
http://dx.doi.org/10.62762/JGEE.2026.416866

RESEARCH ARTICLE

A High-Accuracy Cost Prediction Model for Shale Gas Drilling in Southern Sichuan Using PCA and BP Neural Network

Geli Ma[1], Rong Huang[1], Yichen Dong[1], Chao Lv[2,*] and Wenjing Lin[2]

[1] Sichuan Changing Natural Gas Development Ltd., Chengdu 610051, China
[2] Chengdu Kingray Information Technology Co. Ltd., Chengdu 610041, China

Abstract

Shale gas, as a typical low-quality marginal hydrocarbon resource, faces persistently high drilling costs, which have become one of the main bottlenecks restricting its large-scale development. The Southern Sichuan region of China holds enormous shale gas reserves and is a strategically important area for achieving cost-effective large-scale development. However, as production capacity construction intensifies and the volume of investment and cost data is increasing, traditional data processing methods can no longer meet the timeliness and accuracy requirements for handling massive data. Accurate prediction of oil and gas drilling costs will help in making scientific decisions and evaluations. In this study, based on the costs and engineering parameters of settled wells in the Southern Sichuan Block N shale gas field, we established a Back-Propagation (BP) neural network model incorporating principal component analysis (PCA) to achieve accurate prediction of single-well drilling costs. Results show that: (1) PCA can effectively extract useful information from the shale gas drilling cost influence factors. Specifically, the number of fracturing stages, drilling duration, well depth, total proppant volume, horizontal section length, etc., are identified as key parameters affecting single-well drilling cost. (2) Using Matlab programming and a graphical user interface (GUI), we developed an integrated shale gas single-well cost prediction software system that combines data import, model training, cost prediction, and results export. The BP neural network model's predictions achieved an average relative error of only -0.73%, demonstrating convenience, practicality, and high accuracy. This system can provide a basis for investment decision-making in the Southern Sichuan shale gas block and has value for commercial application.

Keywords: shale gas, principal component analysis, drilling cost prediction, BP neural network, software system.

1 Introduction

Shale gas is a typical unconventional natural gas resource characterized by low quality and low

Submitted: 07 January 2026
Accepted: 24 January 2026
Published: 04 February 2026

Vol. 2, **No.** 1, 2026.
10.62762/JGEE.2026.416866

***Corresponding author:**
Chao Lv
lc@kingray.cn

Citation
Ma, G., Huang, R., Dong, Y., Lv, C., & Lin, W. (2026). A High-Accuracy Cost Prediction Model for Shale Gas Drilling in Southern Sichuan Using PCA and BP Neural Network. *Journal of Geo-Energy and Environment*, 2(1), 46–55.

permeability [1], and its economically effective development has long faced the dual challenges of complex geological conditions and high engineering costs [2] Drilling costs account for a significant proportion of the total investment in shale gas development and are influenced by a web of complex factors, making cost variability high. Effective prediction and control of drilling costs have become a critical bottleneck for achieving scalable and cost-effective shale gas development. The Southern Sichuan region in China is rich in shale gas reserves with tremendous exploration and development potential, making it a key strategic area for ensuring national energy security and advancing the transition to clean energy [3, 4]. As production capacity construction tasks in this region increase, so do drilling activity and the volume of investment and settlement analysis data. Traditional data processing methods (such as analogy methods, factor analysis, incremental calculation, and ratio methods) cannot satisfy the requirements for timeliness, logic, and accuracy when dealing with such large datasets. Therefore, exploring an intelligent method capable of precise and efficient single-well drilling cost prediction is of great practical significance for reducing the overall development cost of the Southern Sichuan shale gas field and improving project economic efficiency.

Many experts at home and abroad have conducted extensive research and practical work on oil and gas drilling cost prediction, proposing numerous forecasting methods [5]. For example, Zhang et al. [6] applied grey theory to predict drilling costs; Fattahi et al. [7] applied support vector machine regression to drilling cost prediction; Ewees et al. [8] used a modified GM(1,1) model to achieve high-precision cost prediction for an oilfield; Yang et al. [9] utilized systematic clustering to categorize drilling cost components and built a total cost regression model via stepwise regression, and Xinhua et al. [10] applied learning curve methods to shale gas cost prediction. However, shale gas drilling costs are subject to a multitude of factors and more complex geological and engineering conditions, resulting in cost data that exhibit greater randomness, variability, and uncertainty [11–13]. The above statistical models have considerable limitations when applied to shale gas drilling cost prediction, and simple physical models require assumptions of normally distributed errors, which greatly affect result stability. In recent years, the rapid development of artificial intelligence—exemplified by artificial neural networks—has demonstrated clear advantages in handling fuzzy, random, and non-linear data, and is especially suitable for complex, poorly-understood systems [14–16]. Such approaches have been widely used in the petroleum and natural gas engineering field, providing a new solution for shale gas drilling cost prediction.

Based on this, the present paper focuses on Block N in the Southern Sichuan shale gas region. Using a large sample of completed wells' data, we built a single-well drilling cost prediction model based on engineering parameters via a BP neural network. We also integrated the BP neural network with a user interface, turning complex code commands into simple operations. The result is a shale gas single-well cost prediction software system that integrates data import, automated training, cost prediction, and result export, which achieves accurate predictions while meeting research needs. This provides a basis for investment decision-making in the Southern Sichuan shale gas block and possesses commercial promotion value.

2 Geological Setting

The Southern Sichuan region is a core area for shale gas exploration and development in China, blessed with rich resource endowment and enormous development potential (Figure 1) [17–19]. In particular, Block N is the main battleground for shale gas development in Southern Sichuan, with a favorable shale gas area of 840 km^2 and resources of about 6.922×10^{11} m^3. This provides a solid resource foundation for large-scale, cost-effective shale gas development. However, the region's geological structure is complex and engineering challenges are significant, leading to persistently high single-well drilling costs. Therefore, in the process of converting resource potential into high-efficiency production capacity, precise prediction and control of drilling costs is a key step to achieve economically effective development of shale gas in this area.

3 Analysis of Factors Influencing Single-Well Drilling Cost

3.1 Factors Influencing Drilling Cost

In recent years, oil companies have placed greater emphasis on cost control, striving for cost advantages in fierce market competition to ensure long-term development. Shale gas drilling costs are influenced by a multitude of complex factors including geological

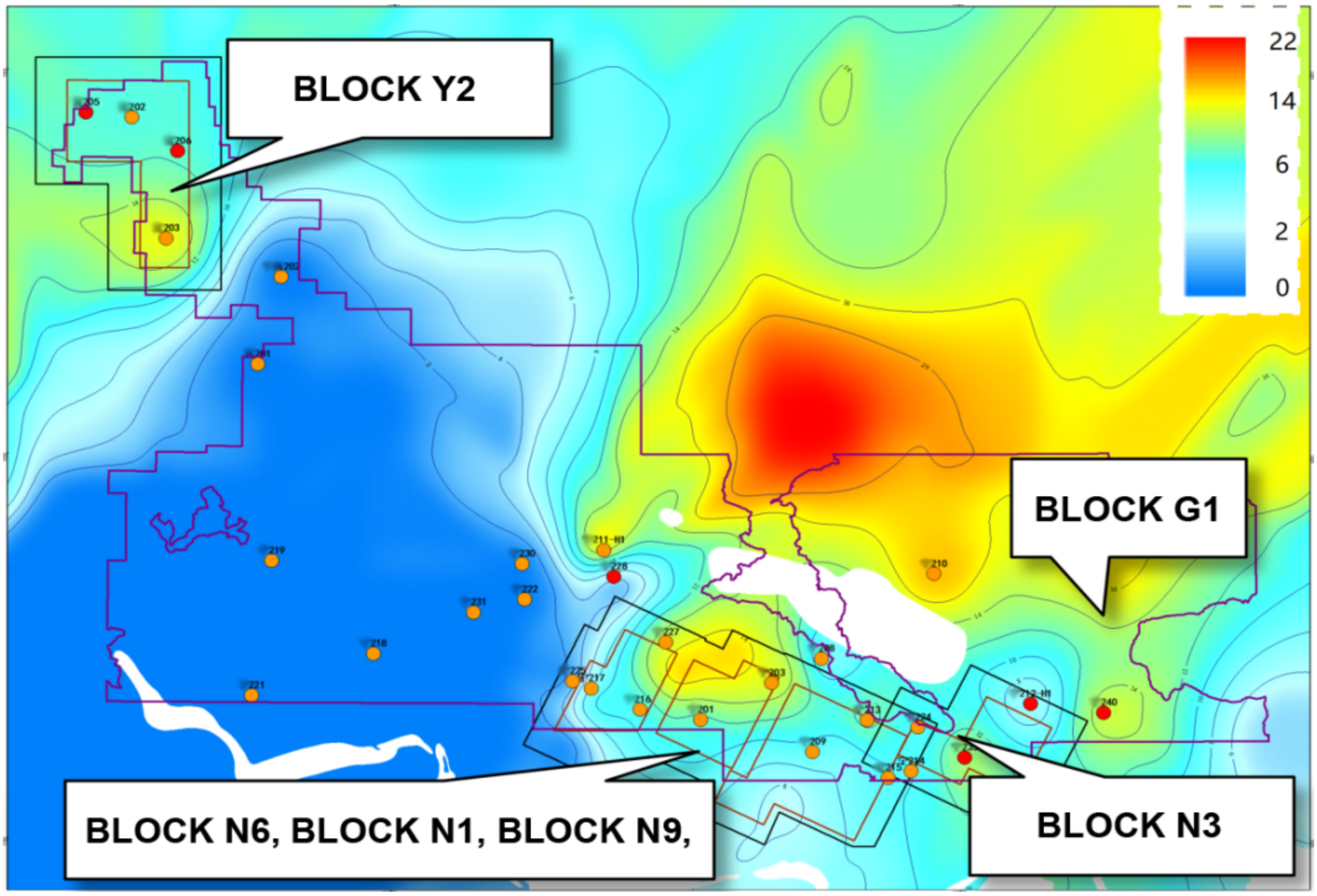

Figure 1. Distribution of main shale gas development areas in Southern Sichuan.

conditions, engineering techniques, operational practices, and market conditions. Accurately identifying and quantifying the key engineering parameters that affect the drilling cost of a single shale gas well is a prerequisite for building a high-precision predictive model and achieving scientific cost management. Therefore, prior to constructing our prediction model, we conducted a systematic study and selection of potential cost-influencing parameters.

At present, extensive empirical and theoretical research on drilling cost drivers has been carried out by domestic and foreign scholars, providing an important basis for constructing the parameter set in this study.

For example, Krishna et al. [20] established a power-law equation relating comprehensive drilling unit cost (or drilling engineering unit cost) C to drilling rig monthly rate I. Hossain et al. [21] argued that well type, drilling duration, footage, and formation structure all affect drilling costs. Xinhua et al. [10] found through studies in the Sichuan shale gas blocks that the key factors influencing drilling cost are drilling duration and horizontal section length, while the factors influencing completion cost include fracturing stage length, fracturing fluid volume, and proppant volume. Yang et al. [9] indicated that drilling cost is affected by both macro and micro factors—the former includes commodity prices, inflation rate, international oil prices, and policy, while the latter includes well structure, geological factors, drilling technology, well depth, and drilling duration.

Foreign scholars have also developed models based on historical drilling data. For example, Augustine et al. [22] analyzed U.S. oil and gas well costs to create a depth-dependent drilling cost index and compared them to geothermal wells. Mistré et al. [11] established an exponential function model relating well depth and total cost. Kaiser [23] and Elkatatny [24] suggested that aside from well depth, geological factors have a strong influence on cost, as the geological environment determines drilling rate, the number of casing strings, and the frequency of drill string failures. Lukawski et al. [25] using data from U.S. oil and gas wells (1976–2009), developed a Cost Evaluation Index (CEI) with 9 sub-indicators to assess drilling costs.

Drawing on extensive research and analysis of financial reports, we identified 10 engineering parameters as the major potential factors influencing shale gas drilling costs: well depth, horizontal section length, number of fracturing stages, drilling duration, final drilled true vertical depth, curve length (build section length), fracturing scale, fracturing stage length, average pump pressure, and total proppant volume. To construct a cost prediction model suitable for the Southern Sichuan Block N, we used actual data from completed wells in this block as the basis. These ten parameters were taken as the initial inputs for the model, and we next performed principal component analysis to extract the key parameters that reflect the cost structure and engineering characteristics of this region.

3.2 Principal Component Analysis

A standard BP neural network algorithm typically requires many training iterations and the initially selected predictive parameters may be interrelated. If all are directly used as input variables for the prediction model, it not only increases the model's complexity but may also lead to overfitting and decreased prediction stability. To address this issue, we applied principal component analysis (PCA) for dimensionality reduction and key information extraction from the input factors, which can effectively mitigate the problems mentioned. PCA is a multivariate statistical method that transforms a set of correlated variables into a few comprehensive indices (principal components) [26, 27]. This yields a smaller number of new variables than the original factors, yet those principal components can explain most of the variance in single-well cost, thereby reducing the complexity of the model while retaining the majority of the information [28, 29].

Using SPSS 25.0 statistical software, we performed PCA on the initial factors. The resulting principal components' contribution rates and eigenvalues are shown in Table 1. The eigenvalues of the first three principal components are all greater than 1, and their cumulative variance contribution reaches 70.62%, indicating that these three principal components effectively represent the majority of information from the original parameters. Therefore, we selected the first three principal components as the factors for single-well cost evaluation. By analyzing the loading matrix of the principal components, we calculated the mean vector coefficient (weight) of each original parameter, as shown in Table 2. The magnitude of the absolute value of each coefficient directly reflects that parameter's explanatory power on the overall cost variation. The analysis results indicate that the number of fracturing stages, drilling duration, and well depth have the most significant impact on single-well cost, followed by total proppant volume and horizontal section length, among others. At the same time, we obtained a linear regression model for drilling cost prediction (equation 1):

$$\begin{aligned} Y = {} & 0.838X_1 + 0.830X_2 + 0.789X_3 + 0.711X_4 \\ & + 0.684X_5 + 0.604X_6 + 0.538X_7 + 0.429X_8 \\ & + 0.377X_9 - 0.123X_{10} \end{aligned} \tag{1}$$

where $X_1 \ldots X_{10}$ correspond to the ten original influencing factors in the order listed above. This provides the basis for selecting the engineering parameters as inputs to the BP neural network model, avoiding the difficulty of choosing input neurons, and thereby enabling accurate prediction of single-well drilling costs.

Table 1. Eigenvalues and contribution rates of principal components for single-well cost evaluation.

Principal Component	Contribution Rate /%	Cumulative Contribution /%	Eigenvalue
F_1	39.833	39.833	3.983
F_2	19.178	59.011	1.918
F_3	11.605	70.616	1.161
F_4	9.541	80.157	0.954
F_5	6.526	86.683	0.653
F_6	5.658	92.341	0.566
F_7	4.523	96.864	0.452
F_8	1.881	98.745	0.188
F_9	1.253	99.998	0.125
F_{10}	8.808E-15	100.000	8.808E-16

Table 2. Principal component vector coefficients (weights) of original cost-influencing parameters.

Cost Influence Parameter	Average Vector Coefficient
Number of fracturing stages (X_1)	0.838
Drilling duration (X_2)	0.830
Well depth (X_3)	0.789
Total proppant volume (X_4)	0.711
Horizontal section length (X_5)	0.684
Final drilled vertical depth (X_6)	0.604
Curve section length (X_7)	0.538
Average pump pressure (X_8)	0.429
Fracturing stage length (X_9)	0.377
Fracturing fluid volume (X_{10}) – "frac scale"	–0.123

Note: Positive/negative signs of coefficients indicate the direction of influence on cost. Magnitude reflects the strength of influence.

Table 3. Single-well engineering parameters and costs for Southern Sichuan Block N (partial data).

Well	Depth/m	Horizontal/m	Frac Stages /count	Drilling Days/d	Proppant/t	Cost/$\times10^4$ RMB
N_1	4070	1400	18	53	1939	5110
N_2	4035	1350	14	40	1385	5003
N_3	4500	1500	18	49	2020	5182
N_4	4600	1500	23	42	2684	5636
N_5	4570	1800	23	34	2665	5939
N_6	4522	1510	18	60	1673	4706
N_7	4980	1500	22	100	2001	5615
N_8	4360	1500	22	90	2194	5425
N_9	4230	1500	22	80	2484	5491
N_{10}	4800	1500	22	73	2367	5495

4 Construction of Single-Well Drilling Cost Prediction Model Based on BP Neural Network

A BP neural network is a multilayer feed-forward network trained by error back-propagation. The core idea of the BP algorithm is to use gradient search techniques to minimize the mean squared error between the network's actual output and the desired output [30, 31]. Typically, one or more hidden layers of neurons are added between the input and output layers; signals propagate forward through each hidden layer, and errors propagate backward through weight connections during training [32]. Therefore, the key to establishing a BP neural network model for single-well drilling cost prediction lies in determining its topology, i.e. the number of input layers, output layers, and hidden layers (and neurons in each).

4.1 Data and Preprocessing

From the PCA results, we identified the relative importance of engineering factors on single-well cost: the number of fracturing stages, drilling duration, and well depth have the most significant impacts, followed by total proppant volume and horizontal section length. Accordingly, we selected these five parameters ($X_1 \sim X_5$ in Table 2) as the input layer neurons of the network.

We took a sample of 250 completed wells in the Southern Sichuan Block N (with their costs and engineering parameters) as the dataset for model development (see Table 3 for partial data). Because the various engineering factors in the prediction system have different units, scales, and trends, it was necessary to normalize the sample data for consistency. We applied an effect coefficient method to standardize and unify the trends of the data, i.e. perform min-max normalization. Using Equation (2), each input factor was rescaled to a [0,1] range, and then Equation (3) was used in the output layer to convert the predicted value back to the original scale.

$$\bar{x}_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{2}$$

$$x_i = \bar{x}_i(x_{max} - x_{min}) + x_{min} \tag{3}$$

where x_i is the normalized value, x_i is the original input data, and x_{min} and x_{max} are the minimum and maximum values of that factor in the sample, respectively.

4.2 Model Structure

According to Kolmogorov's theorem, a three-layer BP neural network (with one hidden layer) can approximate any continuous function with sufficient training [33, 34]. Therefore, we adopted a three-layer BP network for single-well drilling cost prediction. The constructed network topology is 5–10–1, meaning it has 5 input layer neurons (corresponding to the five selected parameters: number of frac stages, drilling duration, well depth, total proppant, horizontal length), 1 output neuron (the target output, i.e. the predicted single-well cost), and 10 hidden layer neurons. The number of hidden neurons (10) was determined using an empirical formula and by analyzing the minimum mean squared error of the network. Specifically, we employed the empirical Equation (4) and trial-and-error to choose an appropriate hidden layer size:

$$m = \sqrt{n + l} + a \tag{4}$$

where m is the number of hidden layer nodes, n is the number of input layer nodes, l is the number of output

Table 4. Comparison of prediction results for 10 sample wells (partial validation data).

Sample	Depth/m	Horiz./m	Stages	Days	Proppant/t	Actual Cost /$\times10^4$ RMB	LR Pred /$\times10^4$ RMB	BP Pred /$\times10^4$ RMB	LR Error /%	BP Error /%
1	4880	2000	27	98	2899	6594	7515	6696	13.98	1.55
2	5060	1500	21	118	2430	6094	8341	5756	36.88	-5.54
3	4780	1500	21	106	2335	5456	7536	5224	38.13	-4.25
4	4800	1500	22	127	2119	5605	7275	5496	29.79	-1.95
5	5350	1800	24	119	2544	5845	7897	5790	35.10	-0.94
6	5380	1900	27	122	2728	6241	8704	6295	39.48	0.87
7	5200	1800	24	96	2490	5756	7589	5680	31.84	-1.32
8	4460	1500	20	77	1583	4681	5841	4621	24.77	-1.29
9	4206	1506	11	57	657	3559	4459	3817	25.31	7.25
10	4115	1500	21	74	1684	4781	5197	4699	8.69	-1.72
Average value									28.39	-0.73

layer nodes, and a is an integer between 1 and 10. Using this formula (with $n = 5, l = 1$) and verifying via mean squared error analysis, we set $m = 10m$ as the optimal hidden layer size in our model.

We randomly selected 200 sets of data out of the 250 well samples as the modeling dataset, of which 70% were used as training samples and 30% as validation (testing) samples. This yielded the BP neural network model for single-well drilling cost prediction. The remaining 50 wells' data were reserved for final model validation. Figure 2 shows the structure of the BP neural network model established in this study.

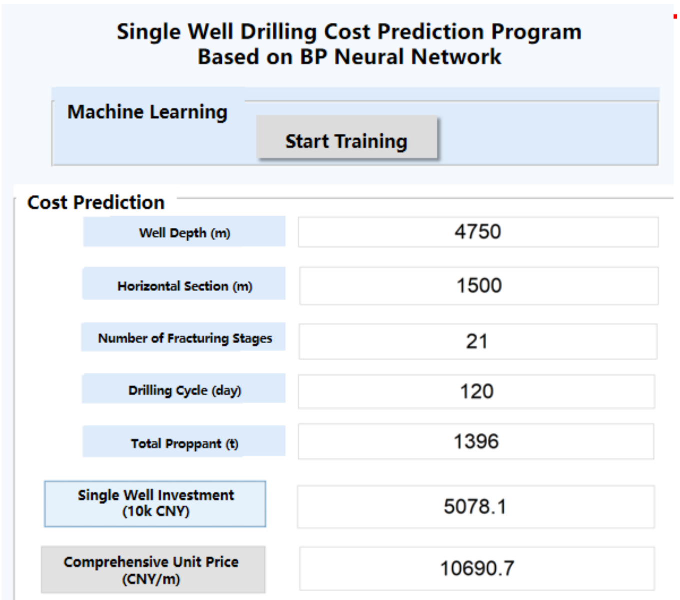

Figure 2. Structure of the BP neural network model developed in this paper.

4.3 Implementation of the BP Neural Network Model

We implemented the BP neural network model in MATLAB and conducted training. Using MATLAB's Neural Network Toolbox, we created the network with the newff function and trained it with a momentum BP algorithm via the traingdm() function, which achieves fast convergence and high training accuracy. Before training the model, the network weights and biases were initialized [35, 36]. The training utilized the traincgf (conjugate gradient with Fletcher-Reeves updates) function for the hidden layer (or a similar trainsig sigmoid training function as mentioned) and the purelin linear transfer function for the output layer. Based on engineering experience, the main network parameters were set as follows: the learning rate was 0.1, the target minimum mean squared error was 1×10^{-4}, and the maximum number of training epochs was 1000.

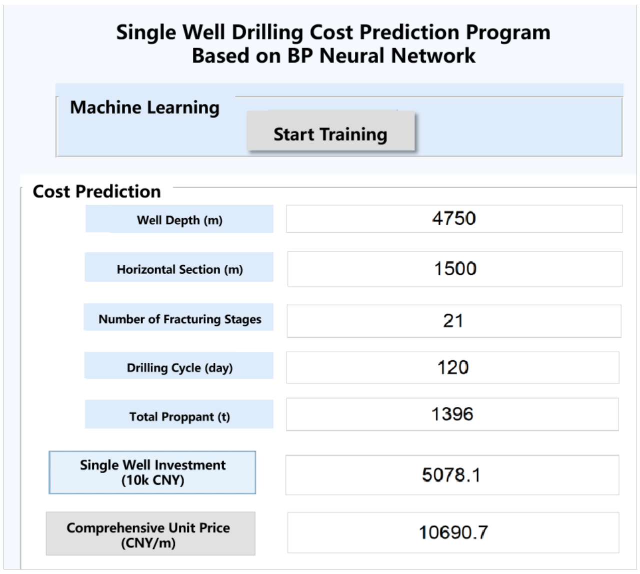

Figure 3. GUI interface of the BP neural network-based cost prediction program.

To facilitate easier use of the model, we designed a GUI (graphical user interface). MATLAB's GUI functionality allows one to call program functions and design user interfaces in a simple and convenient manner. Developing a visual interface as a multi-functional human–computer interaction platform is the most direct, effective, and rapid means to elevate the program to an application level. The GUI design consists of two parts: interface layout design and program function implementation [37]. We created interface elements such as buttons and text boxes, and for each control we wrote callback functions in the code and set appropriate properties. For the single-well drilling cost prediction functionality in this project, the platform interface was designed as shown in Figure 3. As shown in Figure 3, the user can input the required engineering parameters and click the "Start Training" button to initiate machine learning; once training is complete, the results - including the single-well prediction outcome and the corresponding investment cost - are displayed as numerical values in the output fields on the interface.

5 Model Verification

We used the trained BP neural network model to predict the drilling costs of the remaining 50 wells (the validation dataset). For comparison, we also applied a conventional linear regression model (the one derived in section 2.2) to predict the costs of these wells. The prediction results of both models, along with their errors, are shown in Table 4 and Figure 4. The linear regression model's relative errors ranged from 8.69% to 41.33%, with an average error of 28.39%, whereas the BP neural network model's relative errors ranged from –8.80% to 11.99%, with an average of only –0.73%. The prediction accuracy of the BP neural network is therefore far higher than that of the linear regression model, and its predicted costs are in close agreement with the actual single-well costs. These results indicate that the single-well drilling cost prediction model established in this paper is highly reliable for the shale gas block in question and can be applied for wider use.

6 Conclusions

(1) Based on a sample of data from over 200 completed wells in the Southern Sichuan Block N shale gas field, we determined through principal component analysis the relative significance of engineering factors affecting shale gas single-well drilling cost. Among these factors, the number of fracturing stages, drilling duration, and well depth have the most pronounced impact on single-well cost, followed by total proppant volume, horizontal section length, etc. PCA effectively extracted useful information from the shale gas drilling cost indicators, significantly reducing the dimensionality of inputs and simplifying the prediction model's input vector.

(2) Using the sample of 200+ wells' costs and parameters, we established a BP neural network

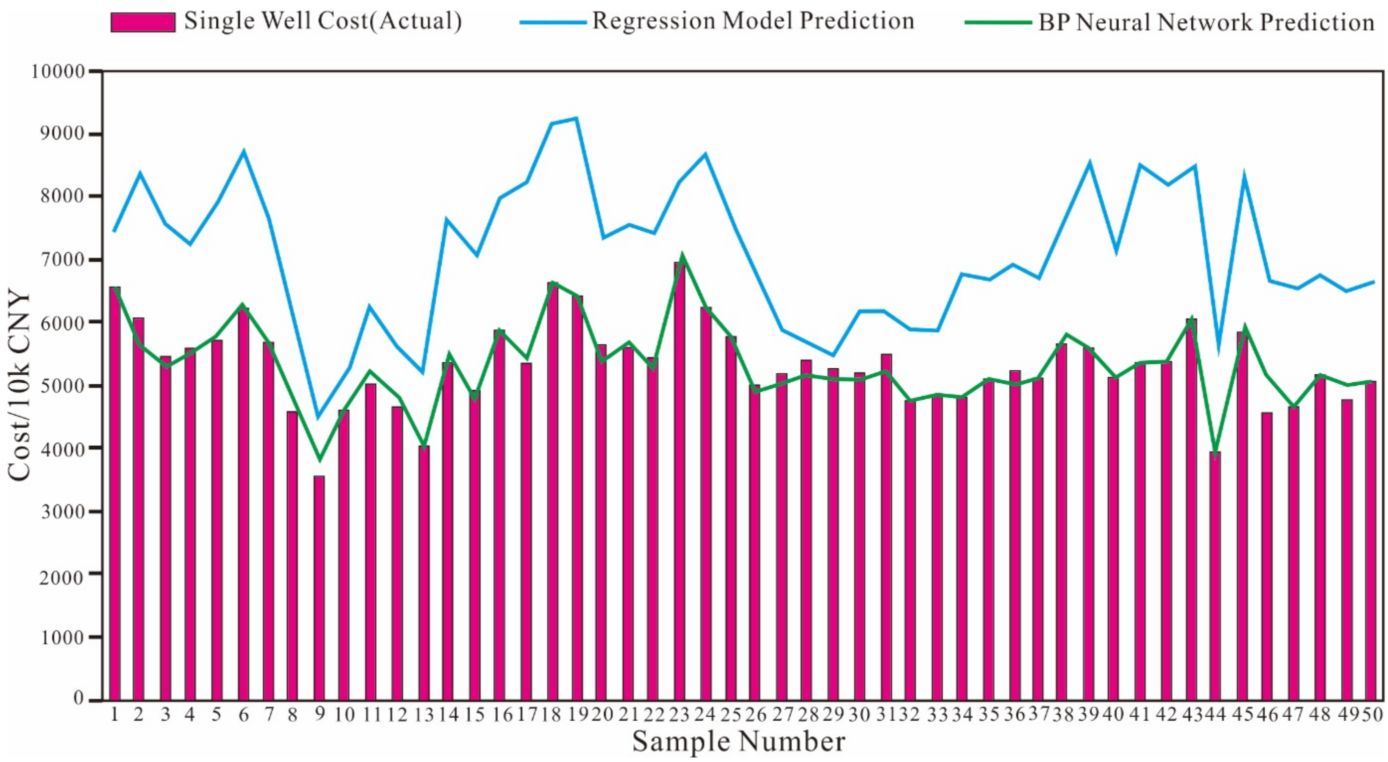

Figure 4. Comparison of actual vs. predicted single-well cost results for the linear regression model and BP neural network model.

model for single-well cost prediction. The model was programmed in MATLAB, and in combination with the software's GUI tools, a visual interface was designed. We developed an integrated shale gas single-well cost prediction software system that incorporates data import, model training, cost prediction, and result export all in one platform, providing a user-friendly human–computer interaction.

(3) The shale gas single-well cost prediction model based on the BP neural network was validated with real-world data, showing an average relative error of only –0.73%. This accuracy is far superior to that of a linear regression model (which had an average error of 28.39%), and the BP model's predictions closely matched the actual costs. The significantly improved prediction precision offers strong support for decision-making and evaluation by investors and corporate managers.

(4) The BP neural network-based cost prediction system for single wells is convenient, practical, and accurate, and thus has value for commercial promotion. With continued advancements in artificial intelligence algorithms, this system can be readily expanded with new functions. In combination with fuzzy mathematics, the algorithms can be continuously improved to achieve better integration with MATLAB and even higher prediction accuracy.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

Geli Ma, Rong Huang, and Yichen Dong are affiliated with the Sichuan Changing Natural Gas Development Ltd., Chengdu 610051, China; Chao Lv and Wenjing Lin are affiliated with the Chengdu Kingray Information Technology Co. Ltd., Chengdu 610041, China. The authors declare that the listed affiliations did not influence the study design, data collection, analysis, interpretation, or the decision to publish, and that no other competing interests exist.

AI Use Statement

The authors declare that no generative AI was used in the preparation of this manuscript.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Li, J., Li, H., Jiang, W., Cai, M., He, J., Wang, Q., & Li, D. (2024). Shale pore characteristics and their impact on the gas-bearing properties of the Longmaxi Formation in the Luzhou area. *Scientific Reports, 14*(1), 16896. [CrossRef]

[2] Ge, H., Wang, X., & Zhang, Y. (2013). A Technical Approach to Reduce Shale Gas Development Cost. *Petroleum Drilling Techniques, 41*(6), 1–5. [CrossRef]

[3] Zhang, L., He, X., Li, X., Li, K., He, J., Zhang, Z., ... & Liu, W. (2022). Shale gas exploration and development in the Sichuan Basin: Progress, challenge and countermeasures. *Natural Gas Industry B, 9*(2), 176-186. [CrossRef]

[4] Zecheng, W. A. N. G., Yizuo, S. H. I., Long, W. E. N., Wuren, X. I. E., Rong, L. I., Hui, J. I. N., & Zengmin, Y. A. N. (2022). Exploring the potential of oil and gas resources in Sichuan Basin with Super Basin Thinking. *Petroleum Exploration and Development, 49*(5), 977-990. [CrossRef]

[5] Chukwuma-Eke, E. C., Ogunsola, O. Y., & Isibor, N. J. (2022). A conceptual approach to cost forecasting and financial planning in complex oil and gas projects. *International Journal of Multidisciplinary Research and Growth Evaluation, 3*(1), 819-833. [CrossRef]

[6] Zhang, Z. X., Kang, X. F., & Zhang, F. F. (2013). Application of grey theory for oil and gas reservoir evaluation program optimization. *Advanced Materials Research, 616*, 1008-1012. [CrossRef]

[7] Fattahi, H., & Babanouri, N. (2017). Applying optimized support vector regression models for prediction of tunnel boring machine performance. *Geotechnical and Geological Engineering, 35*(5), 2205-2217. [CrossRef]

[8] Ewees, A. A., Al-qaness, M. A., Thanh, H. V., AlRassas, A. M., & Elaziz, M. A. (2025). Optimized neural networks for efficient modeling of crude oil production. *Knowledge and Information Systems*, 1-22. [CrossRef]

[9] Yang, T., Liang, Y., Wang, Z., & Ji, Q. (2024). Dynamic Prediction of Shale Gas Drilling Costs Based on Machine Learning. *Applied Sciences (2076-3417), 14*(23). [CrossRef]

[10] Xinhua, M. A., & Jun, X. I. E. (2018). The progress and prospects of shale gas exploration and development in southern Sichuan Basin, SW China. *Petroleum Exploration and Development, 45*(1), 172-182. [CrossRef]

[11] Mistré, M., Crénes, M., & Hafner, M. (2018). Shale gas production costs: historical developments and outlook. *Energy Strategy Reviews, 20*, 20-25. [CrossRef]

[12] Bai, K., Dong, A., Zhan, C., Zhang, W., & Tu, B. (2023). Drilling Parameter Optimization of Shale Gas Wells Based on Saw-Tooth Genetic Algorithm to Reduce Drilling Costs. *Geofluids, 2023*(1), 5647442. [CrossRef]

[13] Zhang, G., Chen, R., Hu, G., Huang, W., Zhang, X., & Liu, H. (2019). Low-cost drilling technology for horizontal wells with atmospheric shale gas in the outer margin of Sichuan Basin. In *IOP Conference Series: Earth and Environmental Science, 295*, 042098. IOP Publishing. [CrossRef]

[14] Abdulraheem, A. (2022). Generation of synthetic sonic slowness logs from real-time drilling sensors using artificial neural network. *Journal of Energy Resources Technology-Transactions of the ASME, 144*(1), 013201. [CrossRef]

[15] Szoplik, J., & Muchel, P. (2022). Using an artificial neural network model for natural gas compositions forecasting. *Energy, 263*(Part D), 126001. [CrossRef]

[16] Szoplik, J. (2015). Forecasting of natural gas consumption with artificial neural networks. *Energy, 85*, 208–220. [CrossRef]

[17] Ma, X. (2019). Enrichment laws and scale effective development of shale gas in the southern Sichuan Basin. *Natural Gas Industry B, 6*(3), 240-249. [CrossRef]

[18] Xu, J., Li, J., Yang, W., Chen, G., Liu, Y., Verdecchia, A., ... & Tang, J. (2025). Shallow Lingering and Deep Transient Seismicity Related to Hydraulic Fracturing in the Changning Shale Gas Field, Sichuan Basin, China. *Journal of Geophysical Research: Solid Earth, 130*(4), e2024JB030279. [CrossRef]

[19] Liu, J., Xue, F., Dai, J., Yang, J., Wang, L., Shi, X., ... & Liu, C. (2024). Waveform features and automatic discrimination of deep and shallow microearthquakes in the Changning shale gas field, Southern Sichuan Basin, China. *Journal of Applied Geophysics, 241*, 105850. [CrossRef]

[20] Krishna, S., Ridha, S., Vasant, P., Ilyas, S. U., & Ofei, T. N. (2020). Simplified predictive model for downhole pressure surges during tripping operations using power law drilling fluids. *Journal of Energy Resources Technology, 142*(12), 123001. [CrossRef]

[21] Hossain, M. E. (2015). Drilling costs estimation for hydrocarbon wells. *Journal of Sustainable Energy Engineering, 3*(1), 3-32. [CrossRef]

[22] Augustine, C., Tester, J. W., Anderson, B., Petty, S., & Livesay, B. (2006, January). A comparison of geothermal with oil and gas well drilling costs. In *Proceedings of the 31st Workshop on Geothermal Reservoir Engineering* (pp. 5-19). New York, New York: Curran Associates Inc.

[23] Kaiser, M. J. (2007). A survey of drilling cost and complexity estimation models. *International Journal of Petroleum Science and Technology, 1*(1), 1–22.

[24] Elkatatny, S. (2021). Real-time prediction of rate of penetration while drilling complex lithologies using artificial intelligence techniques. *Ain Shams Engineering Journal, 12*(1), 917-926. [CrossRef]

[25] Lukawski, M. Z., Anderson, B. J., Augustine, C., Capuano, L. E., Jr., Beckers, K. F., Livesay, B., & Tester, J. W. (2014). Cost analysis of oil, gas, and geothermal well drilling. *Journal of Petroleum Science and Engineering, 118*, 1–14. [CrossRef]

[26] Heidary, M. (2015). The use of kernel principal component analysis and discrete wavelet transform to determine the gas and oil interface. *Journal of Geophysics and Engineering, 12*(3), 386–399. [CrossRef]

[27] Bhosale, S., Manigiri, R., Choudhury, R. P., & Bhakthavatsalam, V. (2020). High resolution mass spectrometry and principal component analysis for an exhaustive understanding of acidic species composition in vacuum gas oil samples. *Energy & Fuels, 34*(3), 2800–2806. [CrossRef]

[28] Han, Y., Liu, J., Liu, F., & Geng, Z. (2022). An intelligent moving window sparse principal component analysis-based case based reasoning for fault diagnosis: Case of the drilling process. *ISA Transactions, 128*(Part A), 242–254. [CrossRef]

[29] Komadja, G. C., Westman, E., Rana, A., & Vitalis, A. (2025). Predicting rock mass strength from drilling data using synergistic unsupervised and supervised machine learning approaches. *Earth Science Informatics, 18*(3), 1-15. [CrossRef]

[30] Wang, W., Cui, X., Qi, Y., Xue, K., Liang, R., & Bai, C. (2024). Prediction model of coal gas permeability based on improved DBO optimized BP neural network. *Sensors, 24*(9), 2873. [CrossRef]

[31] Chen, Q., & Wang, W. (2024). Prediction of shale gas horizontal well production using particle swarm optimisation-based BP neural network. *International Journal of Oil Gas and Coal Technology, 36*(4), 449-460. [CrossRef]

[32] Lin, Y. C., Chen, D. D., Chen, M. S., Chen, X. M., & Li, J. (2018). A precise BP neural network-based online model predictive control strategy for die forging hydraulic press machine. *Neural Computing and Applications, 29*(9), 585-596. [CrossRef]

[33] Chen, Y., Sang, Y., Wang, X., Ye, X., Shi, H., Wu, P., ... & Xiong, C. (2024). Study on Evaluation and Prediction for Shale Gas PDC Bit in Luzhou Block Sichuan Based on BP Neural Network and Bit Structure. *Applied Sciences, 14*(11), 4370. [CrossRef]

[34] Hecht-Nielsen, R. (1992). Theory of the backpropagation neural network. In *Neural networks for perception* (pp. 65-93). Academic Press. [CrossRef]

[35] Luo, S., Xu, T., & Wei, S. (2022). Prediction method and application of shale reservoirs core gas content based on machine learning. *Journal of Applied Geophysics, 204*, 104741. [CrossRef]

[36] Peng, Y., Chen, Z., Xie, L., Wang, Y., Zhang, X., Chen,

N., & Hu, Y. (2024). Prediction of Shale Gas Well Productivity Based on a Cuckoo-Optimized Neural Network. *Mathematics*, *12*(18), 2948. [CrossRef]

[37] Martinez, W. L. (2011). Graphical user interfaces. *Wiley Interdisciplinary Reviews: Computational Statistics*, *3*(2), 119-133. [CrossRef]

Geli Ma received the B.S. degree in Information Management and Information Systems from Chengdu College of Sichuan Normal University, in 2009. (Email: mageli@petrochina.com.cn)

Rong Huang received the B.S. degree in Engineering Costing Major from Xihua University, in 2013. (Email: 15608092928@163.com)

Yichen Dong received the B.S. degree in Petroleum Engineering from Chongqing University of Science and Technology in 2013. (Email: 992412729@qq.com)

Chao Lv received the B.S. degree in Southwest Petroleum University, in 2020. (Email: lc@kingray.cn)

Wenjing Lin received the M.S. degree in Southwest Petroleum University, Chengdu 610500, China, in 2024. (Email: 1377796963@qq.com)