



OPEN ACCESS

EDITED BY

Wei Guo,
Tianjin University, China

REVIEWED BY

Weibin Chen,
University of Macau, China
Binhui Ma,
Hunan University of Science and
Technology, China

*CORRESPONDENCE

Beite Li,
✉ 2410302028@st.gxu.edu.cn

RECEIVED 10 September 2025

REVISED 17 October 2025

ACCEPTED 21 October 2025

PUBLISHED 02 December 2025

CITATION

Wang J, Liao G, Shen A, Wang S, Li B and
Yan T (2025) Study on settlement prediction
methods for high embankments in
mountainous areas.
Front. Built Environ. 11:1702533.
doi: 10.3389/fbuil.2025.1702533

COPYRIGHT

© 2025 Wang, Liao, Shen, Wang, Li and Yan.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Study on settlement prediction methods for high embankments in mountainous areas

Jinhui Wang¹, Guangjian Liao¹, Anbin Shen², Shengli Wang²,
Beite Li^{3*} and Tianyi Yan³

¹China Railway Communications Investment Group Co., Ltd., Nanning, China, ²Guangxi China Railway Nanheng Expressway Co., Ltd., Nanning, China, ³College of Civil Engineering and Architecture, Guangxi University, Nanning, China

Due to the influence of special geological conditions and economic factors, the construction of highways in the mountainous areas of Guangxi has employed a large number of high embankment structures. Characterized by significant settlement amounts and long settlement cycles, these high embankments are prone to issues such as cracking and uneven settlement. This study utilizes three prediction methods—exponential curve model fitting, grey system theory, and backpropagation neural network (BPNN)—to forecast the settlement of a high embankment section of a highway in Guangxi. The prediction results from these three methods are compared with actual measured values. The findings indicate that among the three methods, the BPNN achieves the best overall fitting performance; the grey system theory meets accuracy requirements while being the least affected by different spatial locations; and the exponential curve fitting method involves lower computational costs but shows greater dependence on parameter selection, yet its accuracy improves significantly as the time interval increases. The results provide a methodological basis for the settlement prediction and control of high embankments in mountainous expressways, offering practical significance for ensuring their long-term operational safety.

KEYWORDS

high embankment, settlement prediction, curve prediction method, grey system theory, BP neural network

Highlights

- Conducted a comparative analysis of settlement predictions using multiple methods.
- Explored suitable models for predicting settlement in high embankment sections of highways.
- Innovatively applied grey system theory to the settlement prediction of high embankments.

1 Introduction

China's vast territory means that its expressway networks inevitably traverse mountainous terrain. High embankments have become a common subgrade form in the construction of expressways in mountainous areas. Notably, extensive regions of Guangxi are characterized by karst topography. (Zhang et al., 2023), High embankments have

been extensively employed in the construction of expressways in Guangxi. Currently, high embankment structures with different construction schemes commonly face challenging issues such as instability and cracking. (Zhu and Li, 2020; Xia et al., 2025; Zhang et al., 2025), therefore, effectively addressing these challenges, reducing repair costs, and ensuring road safety and smooth traffic flow have become critical issues that require urgent resolution. Accurate prediction of embankment settlement is a prerequisite that must be met to tackle this problem.

In recent years, the application of BP neural networks in the field of settlement prediction in geotechnical engineering has been continuously deepened, showing strong vitality. Its core advantage lies in the ability to adaptively adjust weights and thresholds through the backpropagation algorithm, without the need to explicitly define the constitutive relationship of the soil, thus achieving high-precision approximation of the highly nonlinear settlement-time relationship (Li et al., 2021). Especially in dealing with practical engineering problems with parameter uncertainty and complex boundary conditions, BP neural networks have demonstrated flexibility that is incomparable to traditional analytical and numerical simulation methods (Ding et al., 2023; Wei, 2023). The BP neural network can effectively calibrate soil parameters during the prediction process (Wang et al., 2007), based on *in-situ* measurement data from single-point settlement gauges embedded at the shoulder and center of high embankments, Yang proposed a settlement prediction method utilizing nonlinear fitting models and neural network algorithms, achieving favorable predictive outcomes (Yang et al., 2025). The use of Finite Element Models (FEM) for settlement prediction offers high accuracy and is a widely employed method. However, its drawback lies in the high computational cost involved (Huang et al., 2025). Tian integrated the use of Response Surface Methodology, Bayesian updating, and Finite Element Models to predict embankment settlement, which partially mitigated the limitations associated with Finite Element Models (Tian et al., 2023). Some scholars have also utilized 2D finite element models to investigate the settlement behavior of soft structural soils (Müthing et al., 2018). Curve fitting is also an effective method for predicting embankment settlement. This approach is straightforward to apply and requires relatively minimal data (Koslanant et al., 2025). Furthermore, Chunlin Li proposed a simplified method for predicting settlement on clay embankments in his research (Li, 2014), offering a valuable optimization to the curve fitting approach for settlement forecasting. Moreover, Liu innovatively proposed an evolutionary support vector regression approach based on an improved grey wolf optimizer for predicting high embankment settlement (Li et al., 2014). Additionally, grey system theory serves as a viable method for settlement prediction. Mao employed this theory to develop a strain prediction model for Pisha sandstone, achieving results with high accuracy and reliability. This success provides a validation for its application in predicting settlement of high embankments (Mao et al., 2025). In addition to the methods mentioned above, some scholars have also employed the Ensemble Kalman Filter (EnKF) (Tao et al., 2020) for soil settlement prediction, which also represents a potential approach for forecasting embankment settlement.

In summary, existing research has made numerous attempts in predicting embankment settlement and soil deformation, yet there remains a lack of systematic comparison among different methods. This paper takes a high embankment section of a highway in Guangxi as the study object, employing exponential curve fitting, grey system theory, and BP neural network as research methods to analyze and compare the predictive effectiveness of these approaches for high embankment settlement.

2 Methods

The road section studied in this paper is part of Guangxi's first eight-lane, dual-direction expressway, which constitutes a key segment of the autonomous region's planned "Horizontal 9" route. It serves as a critical link in establishing a rapid intercity corridor between Nanning and Yulin, boosting development in important towns along the route such as Hengxian, and also forms a major part of the new interprovincial passage between Guangxi and Guangdong. This study focuses on section K129 + 130 to K129 + 250 of Segment 6 of this expressway. The embankment slope at this site reaches a fill height of 40 m, meeting the definition of a high embankment. The slope angle ranges between 25° and 30°. The embankment is designed with a stepped slope, each tier having a height of 8 m. The field conditions are shown in Figure 1.

2.1 Settlement monitoring scheme for high embankments

The monitoring content of this study primarily includes layered settlement monitoring: the elevation of the borehole orifice at each monitoring point is set as 0 m, and the depth of each settlement ring within the settlement point is measured. Based on the initial depth of the settlement rings and their subsequent readings, the settlement amount of each ring can be calculated. To comprehensively observe the settlement behavior, two settlement monitoring points were installed at the left shoulder, and the right shoulder, labeled as G129-1 and G129-2. The distribution of the monitoring points is illustrated in Figure 2.

The layout of the sensing settlement rings is determined based on the length of the settlement pipe at each monitoring point. The strategy adopted is to install one sensing settlement ring every 4 m. Specifically, four settlement rings were deployed at point G129-1, and five at point G129-2. The detailed arrangement of the settlement rings is illustrated in Figure 3 (Schematic Diagram of Settlement Pipe and Sensing Settlement Rings). The parameters of each settlement monitoring point are provided in Table 1 (Description of Layered Settlement Monitoring Point Information).

The layered settlement monitoring adopted in this study employed the YGK-1900 electromagnetic settlement system manufactured by Nanjing Yousi Engineering Instruments Co., Ltd. With a minimum resolution of 1 mm and a maximum measuring depth of 50 m, the instrument was produced and calibrated in accordance with the relevant provisions for deformation-monitoring instruments in the PRC standards GB/T 15406-2007 "Basic parameters and general technical conditions for geotechnical



FIGURE 1
High embankment slope and pavement surface on site.

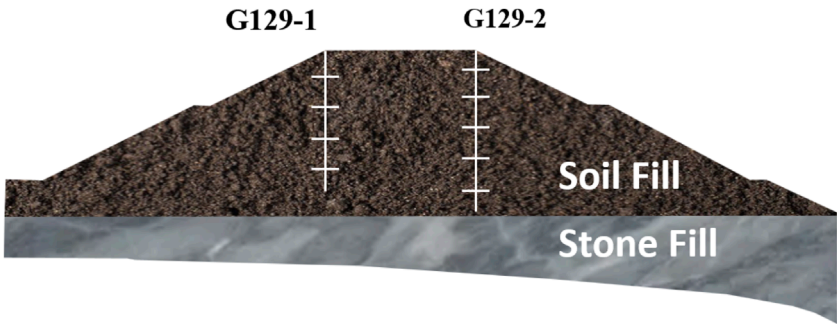


FIGURE 2
Layout of layered settlement monitoring points.

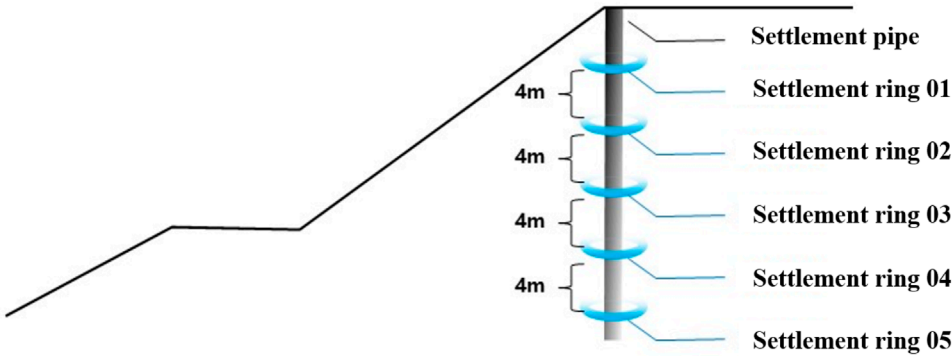


FIGURE 3
Schematic diagram of settlement pipes and settlement inducing settlement rings.

TABLE 1 Information description of layered settlement monitoring points.

Monitoring point	Location	Burial depth/m	Effective monitoring depth/m	Effective settlement rings/Units
G129-1	On the left shoulder of the road	20.0	21.0	4
G129-2	On the right shoulder of the road	24.3	23.0	5

TABLE 2 Water content of soil layers at different depths of embankments.

Depth(m)	−1	−3	−5	−7	−9	−11	−13	−15	−17	−19
Water Content (%)	9.6	13.3	14.1	13.2	15.8	15.3	13.4	21.8	17.8	19.5

TABLE 3 Basic physical performance indicators of fill soil.

Natural water content %	Natural density g/cm ³	Particle density of soil g/cm ³	Liquid limit %	Pplastic limit %	Coefficient of permeability cm/s	Maximum dry density g/cm ³	Optimum water content %
19.5	1.97	2.46	21.6	14.1	2.12×10^{-6}	1.776	15.7

engineering instruments” and JGJ 8-2016 “Code for deformation measurement of buildings and structures.” Prior to delivery, both the probe sensitivity and the length accuracy of the measuring cable were calibrated. During field monitoring, readings were verified against the benchmark at the borehole mouth before and after every measurement to ensure the long-term stability of the measuring system.

Regarding data processing, the study followed the relevant provisions in JGJ 8-2016. Dixon’s test was applied at the 95% confidence level to identify potential outliers. Each statistically flagged value was cross-checked against the trend of adjacent data points and concurrent construction activities (e.g., rainfall, fill placement). An outlier was cautiously discarded only when it could not be explained by any known construction event and clearly violated the physical laws of soil settlement. No such outliers requiring removal were found in the present data set.

2.2 Testing of physical properties of fillers

In this study, the fill soil required for high embankments is sourced from borrow pits near high embankments in the mountainous areas of Guangxi. The main component is the weathered layer of argillaceous siltstone. In its natural state, this fill soil exhibits a typical reddish-brown color, with significant viscosity and semi-hard plastic characteristics. The fill soil also contains a small amount of gravel and rubble, accounting for approximately 14%. These fragments are mainly composed of highly weathered argillaceous siltstone, with diameters ranging from 11 to 61 mm. When the fill soil is submerged in water, it exhibits pronounced cohesive properties, its color changes to a bright red, and its cohesiveness significantly increases when the moisture content reaches a certain level.

This study evaluated the physical properties of representative fill soil in accordance with the specifications of GB/T 50123-2019 “Standard for Geotechnical Testing Methods”, and the results are shown in the Tables 2–4 below.

Meanwhile, this study also remolded these fill materials to prepare specimens with varying degrees of compaction and initial water contents for triaxial testing. The test type was consolidated-undrained (CU), and the shearing rate was 0.075 mm/min. Details and results of the test programme can be found in Table 5.

2.3 Exponential curve fitting method

When studying the prediction of soil settlement, it is necessary to conduct an in-depth analysis of the pattern of changes in the degree of soil consolidation over time. In the prediction of settlement, a time variable t is introduced, and the corresponding settlement at time t is denoted as S_t (Equation 1 illustrates the definition of S_t). The initial settlement is represented as S_0 . It is assumed that there is an exponential functional relationship between the average degree of consolidation and time, with the coefficients in this exponential function being denoted as a and B . Typically, a and B are coefficients influenced by the specific conditions of the engineering site, representing the impact of the on-site engineering environment on the final settlement. Based on this assumption, a mathematical formula describing the settlement at any given time t has been further derived:

$$S_t = S_{\infty} - ae^{-Bt}S_0 \quad (1)$$

Since a and S_0 are both definite coefficients, we can simplify aS_0 as A :

$$S_t = S_{\infty} - Ae^{-Bt} \quad (2)$$

In Equation 2, there are three unknowns, S_{∞} , A , and B . To accurately predict the settlement at a certain moment, a modeling method called the three-point method is adopted. The specific implementation steps of the three-point method are as follows:

In the process of constructing the settlement prediction model, three sets of on-site settlement monitoring data with equal intervals were selected. They are marked as data points (t_1, S_1) , (t_2, S_2) , and (t_3, S_3) , with a time interval of $\Delta t = t_2 - t_1 = t_3 - t_2$. To enhance the accuracy of the model, the selected time interval Δt should be as long as possible. This data processing method helps to reduce data redundancy. Equations 3–5 demonstrate the relationship between the three time data points and their corresponding settlement values.

$$S_1 = S_{\infty} - Ae^{-B(t_2 - \Delta t)} \quad (3)$$

$$S_2 = S_{\infty} - Ae^{-Bt_2} \quad (4)$$

$$S_3 = S_{\infty} - Ae^{-B(t_2 + \Delta t)} \quad (5)$$

TABLE 4 Compression coefficient and compression modulus under different moisture contents.

Pressure (MPa)	Compression coefficient (MPa-1)			Compression modulus (MPa)		
	Water content 13%	Water content 16%	Water content 19%	Water content 13%	Water content 16%	Water content 19%
0.025~0.05	0.197	0.253	0.430	6.157	4.569	3.001
0.05~0.1	0.108	0.111	0.210	11.458	12.674	6.786
0.1~0.2	0.063	0.069	0.163	15.103	18.216	8.223
0.2~0.4	0.043	0.043	0.070	26.234	27.457	19.105

TABLE 5 Strength index of high embankment filler in triaxial test.

Specimen number	Degree of compaction (%)	Initial water Content (%)	Strength parameters	
			Cohesion c (kPa)	Angle of internal friction φ (°)
A-1	90	15.70	39.4	13.6
A-2	90	17.72	32.1	14.9
A-3	90	19.50	18.3	14.5
A-4	92	15.70	41.8	25.1
A-5	92	17.72	35.9	31.2
A-6	92	19.50	20.9	28.5
A-7	94	17.72	44.7	33.8
A-8	96	15.70	51.1	33.7
A-9	96	17.72	40.9	38.5
A-10	96	19.50	37.2	36.5

TABLE 6 Spline interpolation error analysis (partial).

Time (Day)	G129-1		G129-2	
	Spline interpolation (mm)	Measured settlement (mm)	Spline interpolation (mm)	Measured settlement (mm)
14	15.2	17	16.5	14
21	23.1	22	20.8	23
29	31.7	31	32.6	31
36	40.5	42	39.8	41
43	53.66	52	53.6	52
50	65.2	66	64.6	66
57	79	79	76.2	76

TABLE 7 Settlement data table for monitoring points G129-1 and G129-2.

Monitoring duration/d	G129-1		G129-2	
	Cumulative settlement/mm	Settlement Rate/mm·d ⁻¹	Cumulative settlement/mm	Settlement Rate/mm·d ⁻¹
0	0	0	0	0
7	8	1.20	9	1.29
14	17	1.20	14	0.71
21	22	0.68	23	1.29
29	31	1.20	31	1.00
36	42	1.55	41	1.43
43	52	1.37	52	1.57
50	66	2.05	66	2.00
57	79	1.89	76	1.43
70	95	1.20	91	1.15
84	103	0.60	112	1.50
99	116	0.88	116	0.27
113	127	0.77	124	0.57
129	136	0.53	130	0.38
143	145	0.68	141	0.79
157	154	0.60	154	0.93
171	160	0.43	160	0.43
185	167	0.52	166	0.47
196	173	0.35	176	0.64
212	177	0.25	181	0.50
226	180	0.25	189	0.43
254	182	0.17	190	0.21

To determine the specific values of the unknowns S_{∞} , A, and B, Equations 9–11 are solved simultaneously.

By applying this equation, the settlement process can be predicted and analyzed more accurately.

$$S_{\infty} = \frac{S_3(S_2 - S_1) - S_2(S_3 - S_2)}{(S_2 - S_1)(S_3 - S_2)}$$

(6)

$$A = \left[(S_{\infty} - S_1)e^{Bt_1} + (S_{\infty} - S_2)e^{Bt_2} + (S_{\infty} - S_3)e^{Bt_3} \right] / 3$$

(7)

$$B = [\ln(S_2 - S_1) / (S_3 - S_2)] / \Delta t$$

(8)

Based on Equations 6–8, the specific values of the unknowns S_{∞} , A, and B are calculated, thereby constructing the complete equation. The settlement corresponding to any time t can be solved.

2.4 Grey system theory

2.4.1 Data preprocessing

The GM(1,1) model, namely, the first-order single-variable prediction model in grey theory, serves as the foundation and core in the field of grey prediction (Deng and Dong, 2019; Wang et al., 2020). Its wide application goes without saying. When applying this model, the first step is to perform cumulative addition on the original data. This step not only provides key intermediate information

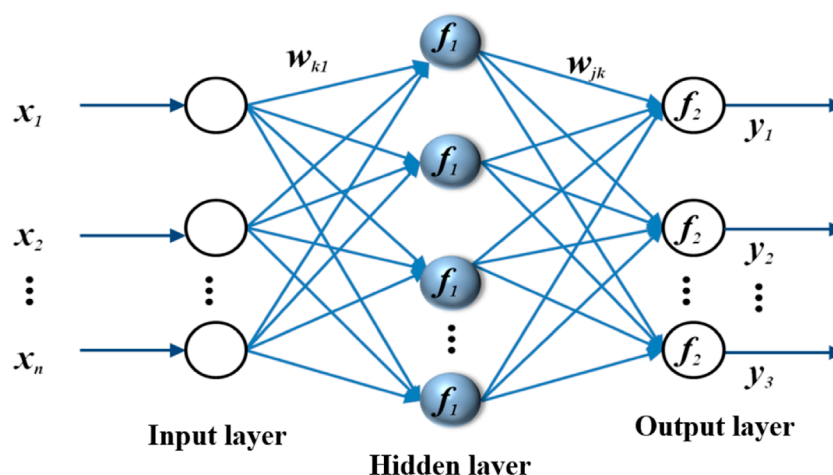


FIGURE 4
Topological structure of a three-layer BP network.

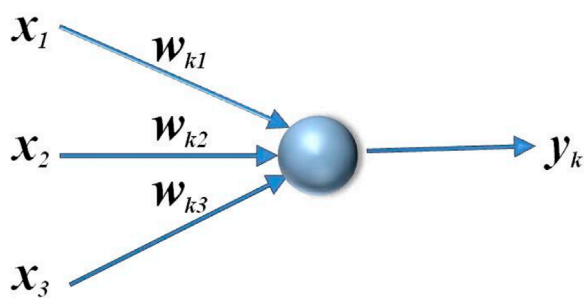


FIGURE 5
Artificial neuron model.

for the modeling process but also effectively reduces the random fluctuations in the original data, thereby enhancing the stability and accuracy of the prediction results (Liu et al., 2014; Yu et al., 2015).

Suppose there are two sets of data, $X^{(0)}$ and $X^{(1)}$:

$$X^{(0)} = (X^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)) \quad (9)$$

$$X^{(1)} = (X^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)) \quad (10)$$

The basic expression form of the GM(1,1) model is:

$$X^{(0)}(k) + ax^{(1)}(k) = b \quad (11)$$

Let the sequence $Z^{(1)}$ be shown in Equation 12:

$$Z^{(1)} = (z^{(1)}(1), z^{(2)}(2), \dots, z^{(1)}(n)) \quad (12)$$

Where, $z^{(1)}(k)$ is shown in Equation 13,

$$z^{(1)}(k) = \frac{1}{2}(x^{(1)}(k) + x^{(1)}(k-1)) \quad (13)$$

Then the basic form of the GM(1,1) model is shown in Equation 14:

$$X^{(0)}(k) + az^{(1)}(k) = b \quad (14)$$

In this paper, to accurately predict the development of post-construction settlement of high embankments, the NDGM(1,1) model in grey models is selected (Xu et al., 2025). Since the monitoring data of post-construction settlement of high embankments are often affected by various factors such as on-site conditions, it is difficult to ensure the data are read at equal time intervals, resulting in sequences with unequal time intervals. The NDGM(1,1) model is particularly good at dealing with such non-equidistant data sequences. Therefore, it can be closer to the actual engineering situation and provide more accurate predictions of high embankment settlement. Based on the GM(1,1) model, the original differential NDGM(1,1) model with unequal time intervals is further developed in this paper, in the hope of achieving more scientific and precise analysis and prediction in the field of high embankment settlement prediction.

The process of transforming a non-equidistant sequence into an equidistant sequence can be described as follows:

1. When processing the original monitoring data sequence, let the sequence be $X_0^{(0)}$, which is shown in Equation 15:

$$X_0^{(0)} = \{x_0^{(0)}(1), x_0^{(0)}(2), \dots, x_0^{(0)}(k), \dots, x_0^{(0)}(n)\} \quad (15)$$

The time interval expression for this data sequence is:

$$X_0^{(0)} = \{x_0^{(0)}(1), x_0^{(0)}(2), \dots, x_0^{(0)}(k), \dots, x_0^{(0)}(n)\} \quad (16)$$

The average time interval is expressed as:

$$\Delta T = \frac{1}{n-1} \sum_{k=1}^{n-1} T_k = \frac{1}{n-1} (t_n - t_1) \quad (17)$$

Based on the time interval T_k defined in Equation 16, a cubic spline interpolation function was used to fit the original data and generate an equidistant sequence for ΔT . $X_1^{(0)}$ is defined in Equation 18.

$$X_1^{(0)} = \{x_1^{(0)}(1), x_1^{(0)}(2), \dots, x_1^{(0)}(k), \dots, x_1^{(0)}(n)\} \quad (18)$$

TABLE 8 Isochronal settlement sequence of post-construction settlement for G129-1.

Post-construction duration/d	Settlement amount/mm	Post-construction duration/d	Settlement amount/mm	Post-construction duration/d	Settlement amount/mm
0	0.0	50	65.9	100	117.0
5	5.3	55	75.5
10	12.3	60	83.6	155	152.0
15	18.0	65	90.1	160	155.5
20	21.4	70	94.9
25	25.9	75	98.1	235	178
30	32.6	80	100.6	240	179.4
35	40.6	85	103.6	245	180.8
40	47.6	90	107.6	250	181.7
45	55.8	95	112.2	255	182.0

TABLE 9 G129-1 Normalization results of post-construction settlement data sequence.

Post-construction duration/d	Membership degree	Post-construction duration/d	Membership degree	Post-construction duration/d	Membership degree
0	0.000	50	0.366	100	0.646
5	0.030	55	0.419
10	0.069	60	0.463	155	0.846
15	0.100	65	0.499	160	0.860
20	0.119	70	0.526
25	0.143	75	0.544	235	0.979
30	0.181	80	0.557	240	0.987
35	0.225	85	0.574	245	0.994
40	0.263	90	0.596	250	0.999
45	0.308	95	0.621	255	1.000

To establish the DGM(1,1), the cumulative addition operation was performed on Equation 17, thereby obtaining the 1-AGO (first-order accumulated generating) time series (see Equations 19, 20):

$$X_1^{(1)} = \{x_1^{(1)}(1), x_1^{(1)}(2), \dots, x_1^{(1)}(k), \dots, x_1^{(1)}(n)\} \quad (19)$$

$$x_1^{(1)}(k) = \sum_{i=1}^k x_1^{(1)}(i), k = 1, 2, \dots, n \quad (20)$$

The method of adjacent mean (MEAN) was adopted to generate a specific sequence. This sequence is obtained by calculating the

average value of adjacent data points in the original data sequence (We present this sequence in Equation 21).

$$Z_1^{(1)} = \{z_1^{(1)}(1), \dots, z_1^{(1)}(k), \dots, z_1^{(1)}(n-1)\} \quad (21)$$

Where $z_1^{(1)}(k)$ is shown in Equation 22:

$$z_1^{(1)}(k) = \frac{1}{2} \left(x_1^{(1)}(k) + x_1^{(1)}(k+1) \right), k = 1, 2, \dots, n-1 \quad (22)$$

The core of the GM(1,1) model lies in its grey differential equation, which is expressed in Equation 23:

$$x_1^{(0)}(k) = a z_1^{(1)}(k) = b \quad (23)$$

TABLE 10 G129-1 forecast samples.

Training samples	Input vector				Expected output
1	0.000	0.03	0.069	0.100	20 days settlement
2	0.030	0.069	0.100	0.119	25 days settlement
3	0.069	0.100	0.119	0.143	30 days settlement
—	—	—	—	—	—
40	0.909	0.923	0.934	0.944	220 days settlement
41	0.923	0.934	0.944	0.951	225 days settlement
Verification samples	Input vector				Expected output
1	0.944	0.954	0.961	0.968	230 days settlement
2	0.954	0.961	0.968	0.971	235 days settlement
—	—	—	—	—	—
5	0.971	0.979	0.987	0.994	250 days settlement
6	0.979	0.987	0.994	0.999	255 days settlement

In the GM(1,1) model, the development coefficient is a key parameter that reflects the speed and trend of system development. Additionally, the grey action quantity is also an important variable that reflects the influence of internal grey information in the system. To determine the specific values of these two parameters, the least squares method is used for precise calculation, as shown in Equations 24, 25:

$$[a, b]^T = (B^T B)^{-1} B^T Y \quad (24)$$

$$B = \begin{bmatrix} -z^{(1)}(1) & 1 \\ -z^{(1)}(2) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} Y = \begin{bmatrix} x_1^{(0)}(2) \\ x_1^{(0)}(3) \\ \vdots \\ x_1^{(0)}(n) \end{bmatrix} \quad (25)$$

From the original form of the GM(1,1) model (Equation 11), we have Equation 26.

$$x^{(1)}(k+1) - x^{(1)}(k) + ax^{(1)}(k+1) = b \quad (26)$$

After rearranging, we get Equations 27–29.

$$x^{(1)}(k+1) = \left(\frac{1}{1+a} \right) x^{(1)}(k) + \frac{b}{1+a} \quad (27)$$

$$x^{(1)}(k) = C \left(\frac{1}{1+a} \right)^k + \frac{b}{a} \quad (28)$$

$$x_1^{(0)}(k) = \begin{cases} x_1^{(0)}(1), k=1 \\ \left(x^{(1)}(1) - \frac{b}{a} \right) \left(\frac{1}{1+a} \right)^k + \frac{b}{a}, k=2, 3, \dots, n \end{cases} \quad (29)$$

When constructing the original differential DGM(1,1) model, a specific prediction equation was obtained, namely, Equation 29. This

equation provides a means to predict unknown data points based on model parameters and data characteristics, and it is a key component of the model application.

2. Restore $\hat{x}_1^{(0)}$ to the original non-equidistant time interval sequence:

On the basis of obtaining the prediction values of the GM(1,1) model, an exponential function is used for fitting. Through the characteristics of the function, the prediction values corresponding to the original time nodes are calculated more precisely, which are the prediction results of the NGM(1,1) model, which is shown in Equation 30.

$$x_0^{(0)} = \{x_0^{(0)}(1), x_0^{(0)}(2), \dots, x_0^{(0)}(k), \dots, x_0^{(0)}(n)\} \quad (30)$$

2.4.2 Interpolation error analysis

To verify the rationality of the data preprocessing steps, this paper conducted a systematic error analysis of the cubic spline interpolation method. The results show that there are reasonable differences in interpolation accuracy among different monitoring points: The Root Mean Square Error (RMSE) at point G129-1 is 1.211 mm, and the mean absolute error (MAE) is 1.035 mm; the corresponding values at point G129-2 are 3.404 mm and 2.767 mm, respectively. These differences mainly stem from the inherent characteristics of the settlement data at the two monitoring points—point G129-2 exhibits more pronounced nonlinear fluctuations during the period from day 70 to day 129, reflecting the impact of local soil inhomogeneity on the settlement process in actual engineering.

Despite the relatively higher interpolation error at point G129-2, its relative error rate of 1.792% is still far below the threshold commonly accepted in engineering practice (2%–5%). The absolute

TABLE 11 G129-1 Original data of stratified settlement monitoring.

Time	Settlement ring 01	Settlement ring 02	Settlement ring 03	Settlement ring 04
	Cumulative settlement	Cumulative settlement	Cumulative settlement	Cumulative settlement
3/6	0.000	0.000	0.005	0.01
3/13	0.008	0.011	0.009	0.012
3/20	0.017	0.015	0.012	0.015
3/27	0.022	0.022	0.017	0.02
4/7	0.031	0.029	0.023	0.023
4/17	0.042	0.035	0.03	0.028
5/12	0.095	0.085	0.082	0.063
5/27	0.103	0.100	0.093	0.076
6/12	0.116	0.111	0.1	0.081
6/26	0.127	0.114	0.105	0.086
7/11	0.136	0.121	0.112	0.089
7/25	0.145	0.128	0.117	0.094
8/8	0.154	0.136	0.119	0.095
8/23	0.160	0.144	0.125	0.098
9/4	0.167	0.15	0.131	0.098
9/20	0.172	0.156	0.136	0.100
10/6	0.175	0.161	0.141	0.104
10/19	0.179	0.164	0.144	0.105
11/15	0.181	0.167	0.146	0.105

error values at both monitoring points are much smaller than the total settlement of the high embankment (178–190 mm), proving that the errors introduced by cubic spline interpolation have a negligible impact on the prediction accuracy of the subsequent NDGM(1,1) model. The error analysis of partial spline interpolation is shown in Table 6.

2.4.3 Establishment of the NDGM(1,1) model

Settlement observation points G129-1 and G129-2 are located at the highest shoulder of the embankment, and their total settlement represents the overall settlement of the high embankment. This study uses the data from these two monitoring points to construct the NDGM(1,1) model for prediction calculations, where the settlement data from 0 to 180 days is used as the basis for modeling, and the settlement data from 196 to 254 days is used to verify the accuracy of the predictions. The settlement data obtained from points G129-1 and G129-2 are detailed in Table 7.

The monitoring data of G129-1 and G129-2 are integrated and introduced to complete the corresponding construction of the NDGM(1,1) model. The calculation process for building the model

is relatively cumbersome; Section 3.3 of Chapter 2 of this paper describes the steps for constructing this model, and this section will not elaborate on it in detail.

2.5 BP neural network method

In the structure of a neural network, there are typically input layer nodes, output layer nodes, and one or more hidden layer nodes located between them (Ji et al., 2025; Jie et al., 2025). As shown in Figure 4, this is a schematic diagram of a three-layer network architecture, which includes the input layer, hidden layer, and output layer. In this architecture, the arrows represent the paths of information flow. Each neuron is connected to all neurons in the next layer, but there are no direct connections between neurons within the same layer (Figure 5). This design ensures the layer-by-layer transmission and processing of information, enhancing the flexibility and expressive power of the network (Wang et al., 2007).

Relying on the BP neural network model, this study constructs a model for predicting the post-construction

TABLE 12 G129-2 stratified settlement monitoring raw data.

Time	Settlement ring 01	Settlement ring 02	Settlement ring 03	Settlement ring 04	Settlement ring 05
	Cumulative settlement	Cumulative settlement	Cumulative settlement	Cumulative settlement	Cumulative settlement
3/6	0.000	0.000	0.000	−0.000	0.000
3/13	0.014	0.012	0.009	−0.003	0.009
3/20	0.019	0.015	0.014	0.001	0.009
3/27	0.028	0.022	0.02	0.005	0.012
4/7	0.036	0.028	0.025	0.008	0.016
4/17	0.046	0.036	0.034	0.014	0.018
5/12	0.096	0.087	0.061	0.055	0.042
5/27	0.117	0.101	0.073	0.066	0.052
6/12	0.121	0.11	0.085	0.072	0.057
6/26	0.129	0.122	0.099	0.077	0.06
7/11	0.135	0.13	0.106	0.083	0.063
7/25	0.146	0.134	0.111	0.086	0.066
8/8	0.159	0.14	0.118	0.087	0.068
8/23	0.165	0.145	0.123	0.091	0.07
9/4	0.171	0.151	0.13	0.093	0.07
9/20	0.18	0.155	0.133	0.094	0.071
10/6	0.187	0.158	0.136	0.097	0.072
10/19	0.193	0.161	0.138	0.097	0.074
11/15	0.191	0.163	0.140	0.098	0.074

settlement of high embankments using MATLAB software, thereby achieving precise prediction of high embankment settlement. Taking monitoring point G129-1 as an example, before constructing the post-construction settlement prediction model, the monitoring data samples need to be converted into equidistant settlement sequences to ensure that the time intervals between each data point are consistent. This study uses interpolation and writes corresponding programs to convert the original monitoring data into equidistant sequences with a 5-day interval. A total of 51 sets of data were generated in this conversion, and the detailed data are shown in Table 8.

Given the large volume and wide range of changes in the monitoring data of G129-1, normalization of the data is required to ensure the smooth progress of subsequent research. Normalization adjusts the data to the interval [0,1] to reduce the differences between data points.

In this paper, the min-max standardization method is used to normalize the sample data, with the calculation formula shown in Equation 31.

$$k = \frac{s_k - s_{\min}}{s_{\max} - s_{\min}} \quad (31)$$

In the processing of settlement monitoring data, the concept of membership degree is introduced, which represents the result after data normalization. According to Formula 37, the monitoring data of G129-1 is normalized, and the results are shown in Table 9.

To construct the BP network prediction model, the following steps of data selection and processing were carried out:

1. Data partitioning. The first 45 sets of data were selected from the total dataset for training the BP network prediction model; the remaining 6 sets of data were designated as validation samples to test the predictive performance of the model.

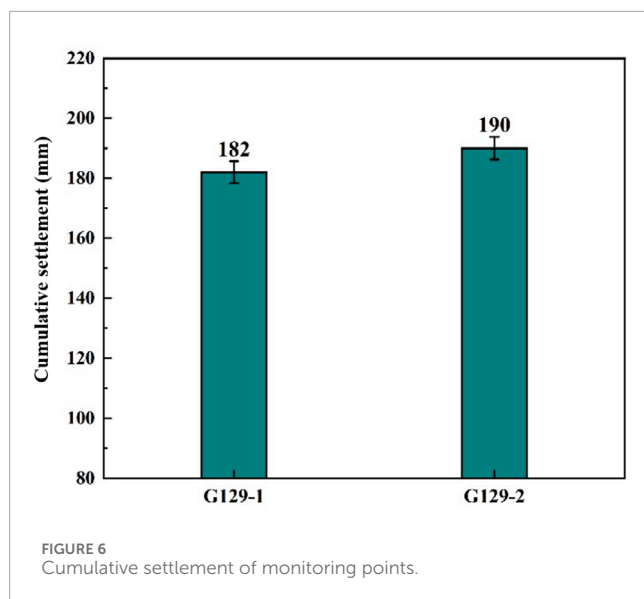


TABLE 13 Selection of exponential curve approximation samples.

Monitoring point	$\Delta \bar{t}$	Data 1	Data 2	Data 3
G129-1	43d	(38,47)	(79,98)	(124,131)
	57d	(38,47)	(94,116)	(152,149)
	71d	(38,47)	(108,127)	(180,162)
G129-2	43d	(38,53)	(79,108)	(124,130)
	57d	(38,53)	(94,111)	(152,149)
	71d	(38,53)	(108,119)	(180,161)

- Construction of learning samples. In accordance with the temporal sequence, five consecutive data points were selected from the training samples, with the first four data points forming the input sample and the last data point serving as the desired output. In this way, 41 effective learning samples can be constructed from the 45 training samples, as detailed in Table 10.

To construct the post-construction settlement prediction model for high embankments, this study, based on the obtained data samples and relevant literature, determined the following neural network structure parameters and training methods:

2.5.1 Network structure

Input layer: Contains 4 input vectors, responsible for receiving and processing raw data; Hidden layer: Composed of 6 neurons, used to capture and process the complex relationships of the input data; Output layer: Contains 1 neuron, responsible for producing the final prediction output.

The number of input-layer nodes is dictated by the underlying principle of the forecasting model. This study adopts a time-series approach: historical settlement records are used to predict

the settlement at the next time step. Specifically, five consecutive settlement readings are employed—the first four serve as inputs to capture the trend, while the fifth acts as the desired output. Consequently, the input layer is fixed at four nodes.

The number of hidden-layer nodes was chosen by combining an empirical heuristic with trial-and-error evaluation. Following the widely used architectural rule of thumb proposed by Yang (Yang et al., 2025) and Yuan (Yuan et al., 2023), $N = \sqrt{a + b} + c$, where $a = 4$ (input nodes), $b = 1$ (output node), and c is an integer from 1 to 10, we first narrowed the candidate range to $N = 3-12$. To refine the topology, we trained networks with every integer node count in this range and compared training error (MSE, MAPE) and goodness-of-fit (R^2). The architecture with six hidden neurons delivered the lowest error and the most stable performance on both training and validation sets, so the final hidden-layer size was fixed at six.

2.5.2 Transfer functions

The hidden layer uses the logsig logarithmic S-shaped transfer function, which helps the model deal with nonlinear relationships; The output layer uses the purelin linear transfer function to ensure the continuity and accuracy of the output.

2.5.3 Training function and parameters

The training function uses the trainlm function, which is suitable for medium and small networks and has a faster convergence speed. The maximum number of iterations for the network is set to 1,000 to ensure that the model has enough opportunities to find the optimal solution; The network training target error is set to 10^{-5} to ensure the prediction accuracy of the model.

2.5.4 Model training and verification

Train the network with the obtained input vectors until the training target is met or the maximum number of iterations is reached; After training, input all sample data into the network to evaluate the model's generalization ability; Input the defined verification vectors into the network to obtain the predicted values of the post-construction settlement of the high embankment.

3 Results

3.1 Settlement monitoring data analysis

Since the completion of the embankment filling project and the initiation of the monitoring program in April 2022, through to its conclusion in December of the same year, continuous monitoring was conducted for a period of 8 months. For the original data of stratified settlement monitoring at the settlement monitoring points G129-1 to G129-2, detailed information can be found in Tables 11, 12.

Since monitoring points G129-1 and G129-2 are both located at the top of the embankment, the cumulative settlement recorded by Settlement Ring 01 at points G129-1 and G129-2 can be considered as the total settlement of the entire high embankment. Figure 6 clearly shows the cumulative settlement data of each monitoring point at the end of the monitoring period.

According to the data shown in Figure 6, within the 8 months after construction, the cumulative total settlement of the high

TABLE 14 Parameters of the exponential curve forecasting model.

Monitoring point	$\Delta \bar{t}$	Parameter A	Parameter B	Model expression	Final settlement amount/mm
G129-1	43d	215.90	0.01251	$S_t = 181.1 - 215.90 \times e^{-0.01251t}$	181.1
	57d	233.03	0.00922	$S_t = 208.9 - 233.02 \times e^{-0.00922t}$	208.9
	71d	232.99	0.00938	$S_t = 210.1 - 232.99 \times e^{-0.00938t}$	210.1
G129-2	43d	298.01	0.02962	$S_t = 137.2 - 298.01 \times e^{-0.02962t}$	137.2
	57d	233.03	0.00921	$S_t = 210.5 - 233.03 \times e^{-0.00921t}$	210.5
	71d	239.11	0.00761	$S_t = 225.1 - 239.11 \times e^{-0.00761t}$	225.1

embankment road surface ranged from 178 to 189 mm, and the differential settlement of the road surface was 11 mm. This result complies with the settlement standards of relevant specifications.

3.2 Analysis of prediction accuracy of settlement using exponential curve fitting

When using the exponential curve method to predict the post-construction settlement of high embankments, it is necessary to collect three sets of data with equal time intervals, and the time interval is crucial for the accuracy of the prediction (Se et al., 2025). However, under on-site conditions, it is difficult to ensure exactly the same time intervals, and it is hard to find completely consistent data in practice. It is possible to approximately select three sets of data with similar time intervals as a substitute, and since the time intervals of these data are small, they are still suitable for prediction studies. This section will use $\Delta \bar{t}$ to replace t in the original model. Based on the analysis of the existing data, selection will be made according to Table 13.

The selected data samples were calculated and analyzed, and the parameters of the exponential curve prediction model are shown in Table 14.

Using the obtained exponential curve model, the settlement of the high embankment from 196 to 254 days after construction was predicted. A detailed analysis of the prediction error can be found in Table 15, and the prediction results are shown in Figure 7.

According to the data analysis in Table 15 and Figure 7, the exponential curve fitting method shows significant sensitivity to time intervals in the settlement prediction at monitoring points G129-1 and G129-2, and its prediction accuracy is significantly affected by the average time intervals of the data samples. The results clearly demonstrate that increasing the average time interval has a significantly positive impact on prediction accuracy. However, the marginal benefit of this improvement diminishes. Taking monitoring point G129-1 as an example, when the average time interval is set to 43 days, the prediction error is roughly stable within a range of -7% ; however, when the average time interval increases to 57 days and 71 days, its prediction curves almost completely coincide, showing excellent predictive performance, with errors reduced to about 2%, fully meeting the expected accuracy requirements. On the other hand, for monitoring point G129-2, with an average time interval of 43 days, the prediction error

exceeded -6.88% . But as the average time interval increased to 57 days and 71 days, the prediction curves also showed a high degree of coincidence, demonstrating excellent predictive effects, with errors stably controlled at around -2.5% , also achieving the expected accuracy standards.

3.3 Analysis of prediction accuracy of grey system theory

Using the established NDGM(1,1) model, the settlement of monitoring points G129-1 and G129-2 from day 196 to day 254 was predicted. The model's accuracy was evaluated against the criteria listed in Table 16; the prediction results and their precision are reported in Table 17, and the predictive performance is illustrated in Figure 8.

At monitoring point G129-1, the NDGM(1,1) model achieved an average relative error (a) of 0.62%, a mean-square-error ratio (C) of 0.13, and a small-error probability (P) of 0.97. The corresponding values for G129-2 are $a = 1.03\%$, $C = 0.17$, and $P = 0.95$. According to the grey-model accuracy standards, both points reach Grade I (excellent) precision ($C < 0.35$ and $P \geq 0.95$). It is noteworthy that G129-2 exhibits a peak prediction deviation on day 226 (measured 189 mm, predicted 180 mm, error 4.76%). This anomaly is likely caused by local soil heterogeneity or construction disturbances, underscoring the influence of dynamic engineering factors on the model. It may also be attributed to early mutation signals (e.g., a settlement rate >2.0 mm/d around day 50). Despite this outlier, the model overall demonstrates excellent robustness to initial conditions and reliable short-term extrapolation, still satisfying engineering accuracy requirements.

The outstanding prediction results at both monitoring points fully confirm that the established NDGM(1,1) model is highly accurate and trustworthy. It not only closely matches the post-construction settlement characteristics of the high embankment in Section 6 of the Hengzhou-Nanning Expressway in Guangxi but also, by accounting for the region's unique geological conditions, effectively predicts future post-construction settlement of high embankments in this area, thereby providing strong technical support for related engineering projects.

In summary, the application of grey system theory to predicting post-construction settlement of high embankments highlights both its distinctive strengths and inherent limitations.

TABLE 15 Precision analysis of curve method forecast model.

Average time interval	Time	G129-1				G129-2			
		Measured value/mm	Predicted value/mm	Residual/mm	Error/%	Measured value/mm	Predicted value/mm	Residual/mm	Error/%
$\Delta \bar{t} = 43d$	196d	173	161	-12	-6.94	176	171	-5	-2.84
	212d	177	164	-13	-7.34	181	173	-8	-4.42
	226d	180	166	-14	-7.78	189	181	-7	-3.75
	254d	182	168	-14	-7.69	190	183	-7	-3.68
$\Delta \bar{t} = 57d$	196d	173	171	-2	-1.16	176	173	-3	-1.70
	212d	177	175	-2	-1.13	181	177	-4	-2.21
	226d	180	179	-1	-0.56	189	184	-5	-2.65
	254d	182	183	1	0.55	190	184	-6	-3.16
$\Delta \bar{t} = 71d$	196d	173	171	-2	-1.16	176	172	-4	-2.27
	212d	177	176	-1	-0.56	181	178	-3	-1.66
	226d	180	180	0	0.00	189	184	-5	-2.65
	254d	182	183	1	0.55	190	187	-3	-1.58

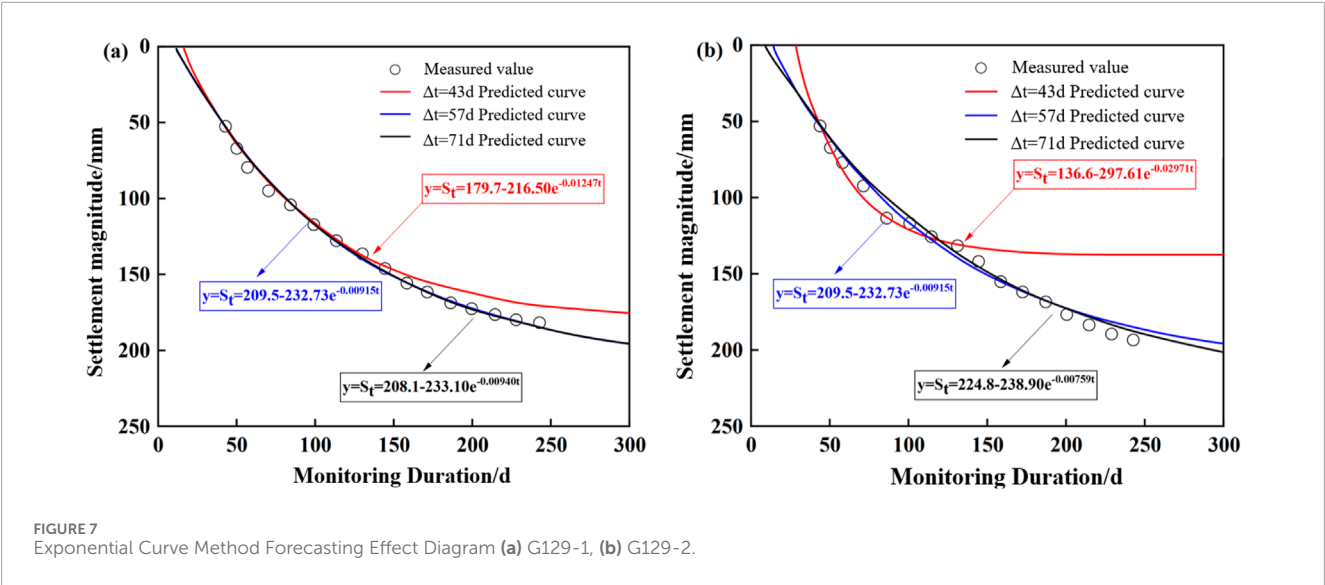


TABLE 16 Reference standards for model accuracy inspection.

Accuracy grade	Mean relative error a	Mean square error ratio C	Small-error probability P
I (Excellent)	$0.00 \leq a < 0.01$	$0.00 \leq C < 0.35$	$0.95 \leq P < 1$
II (Good)	$0.01 \leq a < 0.05$	$0.35 \leq C < 0.50$	$0.80 \leq P < 0.95$
III (Qualified)	$0.05 \leq a < 0.10$	$0.50 \leq C < 0.65$	$0.70 \leq P < 0.80$

TABLE 17 Analysis of prediction accuracy based on grey system theory.

Monitoring point	Monitoring duration	Measured value/mm	Predicted value/mm	Error analysis			Accuracy
				a (%)	C	P	
G129-1	196d	173	167	0.62	0.13	0.97	I
	212d	177	170				
	226d	180	173				
	254d	182	176				
G129-2	196d	176	168	1.03	0.17	0.95	II ~ I
	212d	181	175				
	226d	189	180				
	254d	190	184				

Its principal advantages lie in its ability to effectively handle incomplete and uncertain information, its suitability for small-sample datasets, its computational efficiency, and its generally high prediction accuracy—especially when the data exhibit monotonic trends. Nevertheless, the theory also suffers from drawbacks such as limited predictive performance for fluctuating or anomalous data, potential degradation of accuracy in medium- to long-term forecasting, and sensitivity to the quality of the original data.

3.4 Analysis of settlement prediction accuracy using BP neural network method

Following the procedure outlined in Section 3.4, this paper developed a post-construction settlement prediction model based on neural networks, which was subsequently trained and validated. The predicted post-construction settlement values for the high embankment are presented in Table 18, and the corresponding prediction curves are shown in Figure 9.

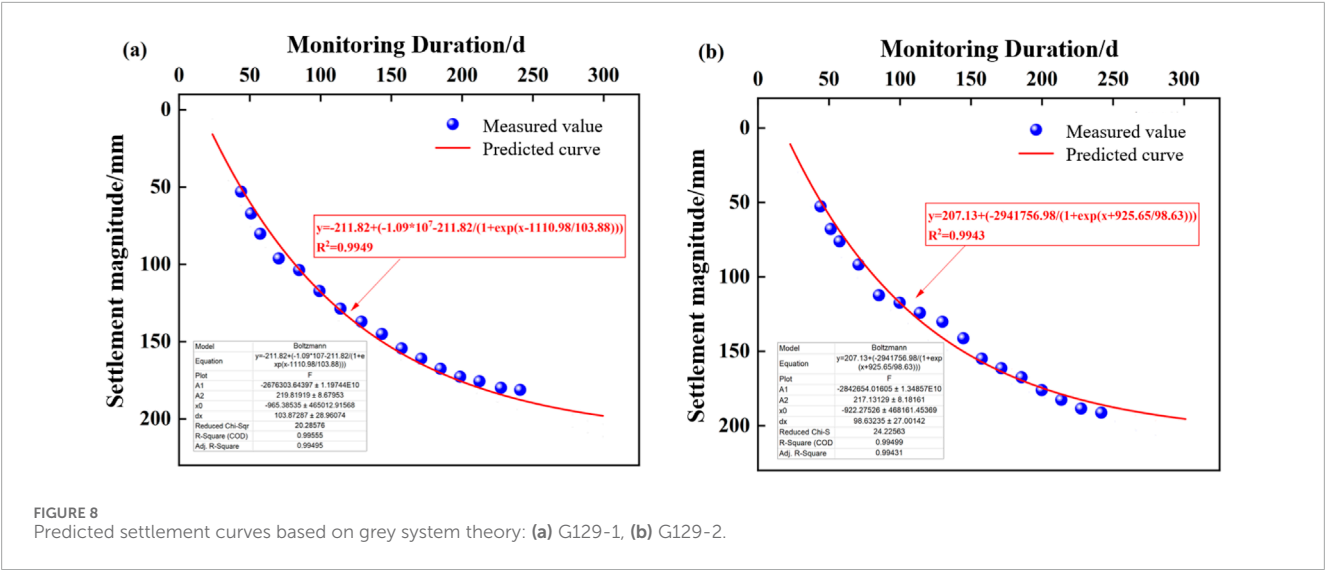


TABLE 18 Accuracy analysis of the BP neural network prediction model.

Monitoring point	Time/d	Measured value/mm	Predicted value/mm	Residual	Error/%
G129-1	225	175.3	174.6	-0.7	-0.40
	230	176.5	175.7	-0.8	-0.45
	235	178.0	177.2	-0.8	-0.45
	240	179.4	178.5	-0.9	-0.50
	245	180.8	179.6	-1.2	-0.66
	250	181.7	180.0	-1.7	-0.94
	255	182.0	179.8	-2.2	-1.21
G129-2	225	179.6	178.4	-1.2	-0.67
	230	181.9	180.1	-1.8	-0.99
	235	184.2	181.3	-2.9	-1.57
	240	186.2	182	-4.2	-0.67
	245	188.0	183.5	-4.5	-2.39
	250	189.3	184.5	-4.8	-2.54
	255	190.0	184.9	-5.1	-2.68

As shown by the comparison in Table 18 and Figure 9, the BP neural network delivers excellent predictions at the beginning of the forecasting period; however, the model error accumulates markedly over time. By day 255, the predicted settlement at G129-1 deviates -1.21% from the measured value, while the deviation at G129-2 widens to -2.68%. This progressive error growth is closely linked to the soil behavior revealed by the multi-level settlement monitoring data. Beyond soil behavior, the error accumulation highlights a limitation in the network's extrapolation capability. Its architecture, optimized for the predominant trends in the training set, struggled to fully adapt to the asymptotic settlement

phase, which was less represented in the data. At G129-2, the deep settlement rings (rings 04 and 05) continue to compress, indicating ongoing delayed consolidation of the deeper strata; the BP model is insufficiently sensitive to this nonlinear deformation, resulting in a lagged prediction. In contrast, at G129-1 all soil layers have essentially stabilized (Table 11 shows that after 19 October the changes in rings 01–04 are ≤0.004 mm), so the model achieves higher accuracy at this point. Nevertheless, the overall performance of the BP neural network remains outstanding: for both G129-1 and G129-2, the residuals between predicted and measured settlements are within ±5 mm, and the relative errors are within ±3.0%.

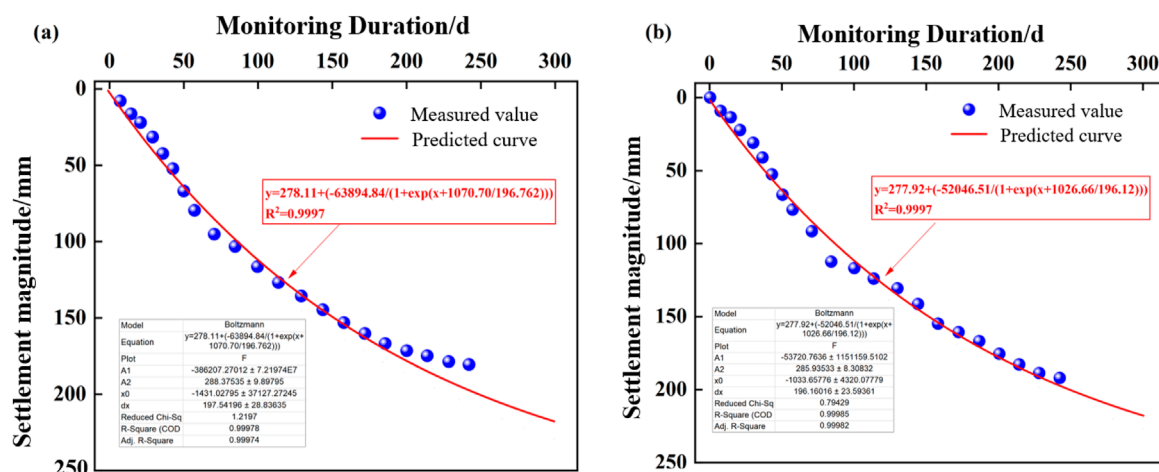


FIGURE 9
BP model prediction plots: (a) G129-1, (b) G129-2.

For monitoring point G129-1, the prediction residuals range from -0.7 mm to -2.2 mm, corresponding to relative errors of -0.40% to -1.21% . The BP model thus delivers excellent accuracy, fully meeting the required precision standards. At G129-2, residuals range from -1.2 mm to -5.1 mm, with relative errors from -0.67% to -2.68% . Although the accuracy is slightly lower than at G129-1, it still satisfies the established precision criteria.

Based on the preceding analysis, the BP neural network model demonstrates superior predictive capability for post-construction settlement of the high embankment in Section 6 of the Nanning–Hengxian Expressway, Guangxi. The model accurately reproduces the complex, dynamic characteristics of post-construction settlement in this region. Confronted with Guangxi's diverse and intricate geological conditions, the BP network's strength lies in its ability to directly handle highly variable settlement data, deliver high-accuracy forecasts (Liu et al., 2024), and exhibit excellent generalization, adapting readily to different geological settings and construction environments (Shen et al., 2016; Feng and Zhang, 2022). Nevertheless, its limitations must also be recognized: training is time-consuming; the model offers limited interpretability of the results; it imposes stringent requirements on the training dataset; and selection of network architecture and hyper-parameters remains largely empirical.

3.5 Comparison of settlement prediction accuracy and method evaluation

In the preceding sections, the exponential-curve fitting method, grey system theory, and BP neural network were employed—using field monitoring data—to predict the post-construction settlement of a high embankment on an expressway in Guangxi, and the prediction accuracy of each approach was thoroughly analyzed. To provide a more intuitive comparison and objective assessment of these methods, this study adopts the Root Mean Square Error (RMSE) and the mean absolute percentage error (MAPE) as quantitative metrics, while also summarizing the residual and error analyses of the aforementioned prediction methods in Table 19. The calculation formulas are presented in Equations 32, 33.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (32)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \bar{y}_i}{y_i} \right| \quad (33)$$

In the above formula, y_i represents the actual value of the i th observation, and \bar{y}_i represents the predicted value of the i th observation.

RMSE and MAPE are statistical metrics commonly used to quantify the accuracy of prediction models by measuring the discrepancy between predicted and observed values. As shown in the table above, the BP neural network demonstrates superior nonlinear modeling capability in complex scenarios, delivering overall satisfactory performance under both RMSE and MAPE criteria. Although the exponential-curve method performs well when $A = 71$ D, its predictive accuracy is highly sensitive to the chosen time interval; when the interval is short, its results become less reliable. While the exponential-curve results illustrate that parameter tuning can significantly improve traditional models, the practical value diminishes as A increases. Consequently, an appropriate time interval must be carefully selected when applying this method.

The data in the table also reveal notable spatial heterogeneity: predictions at monitoring point G129-1 are consistently more accurate than those at G129-2. This discrepancy may stem from differing topographic conditions or from difficulties in ensuring uniform compaction of the subgrade during construction. Despite these site-specific variations, the BP neural network performs well at both locations under the RMSE and MAPE criteria. Across all methods investigated, MAPE remained below 10%, whereas the RMSE for some approaches was relatively large. Among the prediction techniques examined, although grey system theory does not yield the smallest RMSE or MAPE values, the RMSE and MAPE figures it produces for different monitoring points exhibit strong consistency. This suggests that grey system theory is more robust against spatial variability and has considerable practical potential for settlement prediction in engineering applications.

TABLE 19 Accuracy comparison of different prediction methods.

Prediction method	G129-1		G129-2	
	RMSE/mm	MAPE/%	RMSE/mm	MAPE/%
Grey system theory	6.52	3.65	7.37	3.95
BP neural network	1.30	0.66	3.78	1.87
Exponential curve ($\Delta \bar{t} = 43\text{D}$)	13.28	7.44	7.11	3.79
Exponential curve ($\Delta \bar{t} = 57\text{D}$)	1.58	0.85	4.64	2.43
Exponential curve ($\Delta \bar{t} = 71\text{D}$)	1.22	0.57	3.84	2.04

4 Conclusion

In this study, the settlement of a high embankment was calculated using the exponential-curve fitting method, grey system theory, and BP neural network, and the results were compared with field monitoring data. The following conclusions are drawn:

1. The predictive performance of the exponential-curve method is closely tied to the length of the average time interval; extending this interval improves accuracy. The study shows that when the average interval reaches 71 days, the prediction error is confined to the range of 0.20 %–1.16%.
2. Predictions from the NDGM(1,1) model in grey system theory are slightly lower than the observed values. Residuals remain within 6–9 mm, and the relative error stays between 3.55% and 4.02%, demonstrating excellent predictive accuracy. Moreover, compared with other models, the grey-system fitting approach is less affected by spatial variability.
3. The BP neural-network model produced predictions that closely matched the measured data, with residuals confined to ± 5 mm and relative errors within $\pm 2.68\%$, demonstrating superior predictive performance. Compared with other methods, the BP neural network not only delivers higher accuracy but also operates under less restrictive conditions, offering broader applicability for effectively forecasting post-construction settlement of high embankments.

In summary, the practical application of these three settlement prediction methods can be strategically matched according to the data conditions and accuracy requirements of different engineering stages. The exponential curve method is easy to calculate and can serve as an effective tool for quick preliminary evaluation during construction. The grey system theory demonstrates remarkable robustness, especially suitable for the early operation stage when monitoring data is scarce, and it can provide reliable medium and long-term predictions with the minimum data input. After a large amount of settlement data has been accumulated in the operation stage, the BP neural network becomes the most powerful choice, providing the highest prediction accuracy for precise long-term settlement forecasting. Therefore, the selection of the prediction model can be based on the specific project stage, flexibly utilizing the unique advantages of each method.

Future work will focus on developing intelligent prediction models that adequately account for typical geological features of karst areas, including groundwater fluctuations and underlying fissures. The ultimate goal is to advance settlement prediction from post-construction analysis to pre-construction guidance. By deeply integrating predictive outcomes with engineering practices such as dynamic design and ground improvement techniques like optimized drainage and targeted grouting, we aim to directly apply these results to guide construction decisions and risk control, thereby enhancing the precision and safety of engineering projects.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

JW: Investigation, Writing – original draft, Funding acquisition, Resources, Software. GL: Data curation, Formal Analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft. AS: Supervision, Validation, Visualization, Writing – review and editing, Funding acquisition, Resources, Software. SW: Funding acquisition, Resources, Software, Supervision, Validation, Visualization, Writing – review and editing. BL: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing. TY: Writing – original draft, Writing – review and editing, Conceptualization, Investigation, Software, Supervision, Visualization.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was supported by the National Natural Science Foundation of China (NO. 52368058).

Conflict of interest

Authors JW and GL were employed by China Railway Communications Investment Group Co., Ltd.

Authors AS and SW were employed by Guangxi China Railway Nanheng Expressway Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

References

- Deng, M. X., and Dong, Y. L. (2019). "Application of improved Grey GM (1,1) model in power prediction of wind farm," in *31st Chinese control and decision conference (CCDC)*, 3764–3769.
- Ding, Y., Hang, D., Wei, Y. J., Zhang, X. L., Ma, S. Y., Liu, Z. X., et al. (2023). Settlement prediction of existing metro induced by new metro construction with machine learning based on SHM data: a comparative study. *J. Civ. Struct. Health Monit.* 13 (6–7), 1447–1457. doi:10.1007/s13349-023-00714-4
- Feng, L., and Zhang, L. (2022). Enhanced prediction intervals of tunnel-induced settlement using the genetic algorithm and neural network. *Reliab. Eng. & Syst. Saf.* 223, 108439. doi:10.1016/j.res.2022.108439
- Huang, S., Huang, J., Kelly, R., Jonse, M., Wong, P., Yuen, S., et al. (2025). Efficient class C prediction of PVD-Improved soft clay settlements. *Eng. Geol.* 346, 107903. doi:10.1016/j.enggeo.2024.107903
- Ji, H., Zhang, X., Wang, T., Yang, K., Jiang, J., and Xing, Z. (2025). Oil spill area prediction model of submarine pipeline based on BP neural network and convolutional neural network. *Process Saf. Environ. Prot.* 199, 107264. doi:10.1016/j.psep.2025.107264
- Jie, P., Zhou, Y., Zhang, Z., and Wei, F. (2025). Heating energy consumption prediction based on improved GA-BP neural network model. *Energy* 328, 136392. doi:10.1016/j.energy.2025.136392
- Koslanant, S., Sakulsaksri, C., Boonyong, C., and Tanseng, P. (2025). Observational techniques by curve fitting for predicting radial consolidation settlement. *Proc. Institution Civ. Eng. - Ground Improv.* 178 (2), 163–176. doi:10.1680/jgrim.24.00047
- Li, C. (2014). A simplified method for prediction of embankment settlement in clays. *J. Rock Mech. Geotechnical Eng.* 6 (1), 61–66. doi:10.1016/j.jrmge.2013.12.002
- Li, P. P., Li, J. P., Liu, G. Y., and Zhou, P. (2014). "Prediction of dredged soil settlement based on improved BP neural network," in *5th GeoShanghai international conference*.
- Li, G. H., Han, C. Y., Mei, H., and Chen, S. (2021). Application of the WNN-based SCG optimization algorithm for predicting soft soil foundation engineering settlement. *Sci. Program. Appl.* 2021 (11), 1–13. doi:10.1155/2021/9936285
- Liu, J., Xiao, X., Guo, J., and Mao, S. (2014). Error and its upper bound estimation between the solutions of GM(1,1) grey forecasting models. *Appl. Math. Comput.* 246, 648–660. doi:10.1016/j.amc.2014.08.049
- Liu, D., Zhang, W., Tang, Y., Xie, B., Shi, Q., and Cao, K. (2024). Evolving support vector regression based on improved grey wolf optimization for predicting settlement during construction of high-filled roadbed. *Transp. Geotech.* 45, 101233. doi:10.1016/j.trgeo.2024.101233
- Mao, X., Hu, Y., Wu, Q., Cai, P., Wang, Y., and Lei, X. (2025). Study on prediction model of cumulative plastic strain of pisha sandstone based on grey systems theory. *Soil Dyn. Earthq. Eng.* 190, 109205. doi:10.1016/j.soildyn.2024.109205
- Müthing, N., Zhao, C., Hölter, R., and Schanz, T. (2018). Settlement prediction for an embankment on soft clay. *Comput. Geotechnics* 93, 87–103. doi:10.1016/j.compgeo.2017.06.002
- Se, H., Song, K., Jiang, J., Sun, C., Hu, Y., Wang, B., et al. (2025). A robust and fast odor detection method with remaining response curve forecasting. *Sensors Actuators A Phys.* 386, 116302. doi:10.1016/j.sna.2025.116302
- Shen, Q., Li, Y., and Peng, C. L. (2016). "Analysis and prediction of soft foundation settlement for expressway based on BP neural network," in *International conference on architectural engineering and civil engineering (AECE)*, 274–277.
- Tao, Y., Sun, H., and Cai, Y. (2020). Predicting soil settlement with quantified uncertainties by using ensemble Kalman filtering. *Eng. Geol.* 276, 105753. doi:10.1016/j.enggeo.2020.105753
- Tian, H., Li, D.-Q., Cao, Z.-J., and Du, W. (2023). Auxiliary Bayesian updating of embankment settlement based on finite element model and response surface method. *Eng. Geol.* 323, 107244. doi:10.1016/j.enggeo.2023.107244
- Wang, Z.-L., Li, Y.-c., and Shen, R. F. (2007). Correction of soil parameters in calculation of embankment settlement using a BP network back-analysis model. *Eng. Geol.* 91 (2), 168–177. doi:10.1016/j.enggeo.2007.01.007
- Wang, Z.-X., Wang, Z.-W., and Li, Q. (2020). Forecasting the industrial solar energy consumption using a novel seasonal GM(1,1) model with dynamic seasonal adjustment factors. *Energy* 200, 117460. doi:10.1016/j.energy.2020.117460
- Wei, L. Y. (2023). Research on settlement prediction of building foundation in smart city based on BP network. *Int. J. Adv. Comput. Sci. Appl.* 14 (6), 876–884. doi:10.14569/ijacsa.2023.0140693
- Xia, B., Zheng, G., Zhou, H., and Diao, Y. (2025). Failure mechanisms and characteristics of high-filled embankments with different types of basal reinforcement. *Transp. Geotech.* 51, 101528. doi:10.1016/j.trgeo.2025.101528
- Xu, Y., Wang, Y., and Pan, Y. (2025). Prediction of China's unconventional natural gas production based on grey waveform prediction model and non-homogeneous exponential discrete model. *Energy Rep.* 13, 2829–2843. doi:10.1016/j.egy.2025.02.012
- Yang, E., Wang, K., He, J., Liu, K., Wang, J., Zhang, H., et al. (2025). Settlement prediction of a high embankment based on non-linear regression and neural network algorithm. *Transp. Geotech.* 50, 101443. doi:10.1016/j.trgeo.2024.101443
- Yu, Z., Yang, C., Zhang, Z., and Jiao, J. (2015). Error correction method based on data transformational GM(1,1) and application on tax forecasting. *Appl. Soft Comput.* 37, 554–560. doi:10.1016/j.asoc.2015.09.001
- Yuan, Z., Gao, L., Chen, H. J., and Song, S. (2023). Study on settlement of self-compacting solidified soil in foundation pit backfilling based on GA-BP neural network model. *Buildings* 13 (8), 2014. doi:10.3390/buildings13082014
- Zhang, B., Hou, H., Liu, L., Huang, Z., and Zhao, L. (2023). Spatial prediction and influencing factors identification of potential toxic element contamination in soil of different karst landform regions using integration model. *Chemosphere* 327, 138404. doi:10.1016/j.chemosphere.2023.138404
- Zhang, S., Wang, Y., Zeng, L., Luo, J., Wang, J., Dong, T., et al. (2025). Long-term deformation rules of railway embankments in permafrost regions: classification and prediction. *Cold Regions Sci. Technol.* 231, 104425. doi:10.1016/j.coldregions.2025.104425
- Zhu, C., and Li, N. (2020). Ranking of influence factors and control technologies for the post-construction settlement of loess high-filling embankments. *Comput. Geotechnics* 118, 103320. doi:10.1016/j.compgeo.2019.103320