Article | Received 23 December 2024; Accepted 18 March 2025; Published date 31 March, 2025
https://doi.org/10.55092/sc20250005

A back-analysis method of deep excavation in soft soil based on BIM-NS-ML integrated technology

Shu Jiang[1], Ziqian Li[1], Rongjun Zhang[1,2,3,*], Junjie Zheng[1] and Shaoyan Zhou[4]

[1] School of Civil Engineering, Wuhan University, Wuhan 430072, China
[2] State Key Laboratory of Water Resources Engineering and Management, Wuhan University, Wuhan 430072, China
[3] Key Laboratory of Rock Mechanics in Hydraulic Structural Engineering of the Ministry of Education, Wuhan University, Wuhan 430072, China
[4] School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology, Wuhan 430072, China

* Correspondence author; E-mail: ce_zhangrj@whu.edu.cn.

Highlights:

- Integration of BIM and numerical simulation improves the efficiency of building calculation models.
- A back-analysis method integrates BIM-numerical simulation and machine learning with monitoring data.
- Validated in deep excavation, the framework identifies parameters and aids deformation prediction.

Abstract: It is witnessed that building information modeling (BIM) technology has shown its capabilities in data integration in the construction industry. Incorporating innovative geotechnical theories into BIM helps to further develop its application potential. In the practice of deep excavation engineering, obtaining accurate soil parameters is the key to preventing deep excavation accidents and reducing construction costs. Aiming at the complexity of soil properties in soft soil deep excavations, an intelligent inversion framework for soil parameters in deep excavations is established by using BIM technology, finite difference method (FDM), and nondominated sorting genetic algorithm II (NSGA-II). Firstly, a building information modeling-numerical simulation (BIM-NS) integrated component is implemented based on a transformation interface, including geometric meshing processing and controlled script automated execution. Then, a back-analysis component based on NSGA-II optimization is attached to the BIM-NS processing to improve the accuracy of soil parameters. Subsequently, a framework of the building information modeling- numerical simulation-machine learning (BIM-NS-ML) integrated technology is established, enabling the usage of optimal soil parameters for automatic deep excavation simulation. Finally, the integrated framework is applied to a subway deep excavation project, which verifies that the proposed intelligent integrated framework can accurately identify soil parameters in an efficient manner. The BIM-NS-ML integrated technology significantly improves the efficiency of

modeling and calculating. The multi-objective optimization algorithm effectively addresses the problem of parameter complexity in soft soil. In addition, the intuitiveness of parameter inversion results is further enhanced to provide support for construction management and decision-making.

Keywords: deep excavation; back-analysis; building information modeling; numerical simulation; machine learning

Nomenclature

Abbreviations	
ANN	Artificial neural network
ACO	Ant colony optimization
BIM	Building information modeling
BPNN	Backpropagation neural network
ESA	Evolution strategy algorithm
FDM	Finite difference method
FEM	Finite element method
FVM	Finite volume method
ML	Machine learning
NSGA	Nondominated sorting genetic algorithm
OD	Orthogonal design
PSO	Particle swarm optimization
RF	Random field
UUID	Universally Unique Identifier
Symbols	
E	Young’s modulus, MPa
c	Cohesion, kPa
φ	Friction angle
μ	Poisson’s ratio
γ	Unit weight
φ_i	Friction angle for interface elements
c_i	Cohesion for interface elements
K_n	Normal stiffness for interface elements
K_s	Shear stiffness for interface elements

1. Introduction

With the rapid development of the urban economy, the large-scale construction of subway transportation networks has become the key to alleviating traffic pressure [1]. However, subway construction typically involves deep excavation, and most stations are located in busy residential and commercial areas. The surrounding buildings are densely populated, the pedestrian flow is large, and the underground pipe network is complex. Consequently, the construction is significantly restricted by the environment [3]. This requires stricter deformation control of subway deep excavation projects, especially in soft soil areas. Soft soil has low strength, high compressibility, and strong sensitivity, which can easily lead to uneven foundation settlement and cause engineering accidents [4]. Therefore, it is particularly important to accurately predict the ground deformation caused by deep excavation construction in soft soil and implement safety control based on the prediction results for accident prevention.

It is of great significance to use monitored data to back-analyze the soil parameters of deep excavation and guide construction [5]. To improve the computing speed and accuracy, optimization algorithms are widely used, such as the particle swarm optimization (PSO) algorithm [7], evolution strategy algorithm (ESA) [10], ant colony optimization (ACO) [12], *etc.*, which are combined with numerical calculations to perform parameter inverse analysis and optimization. The nondominated sorting genetic algorithm II (NSGA-II) [14] is proposed based on the NSGA to solve multi-objective optimization problems. It has the advantages of fast running speed and good convergence, making it highly suitable for application in the fast and highly automated feedback analysis of deep excavation construction in soft soil. Traditional back analysis methods require manual modeling and repetitive complex calculations [17,18], which consume a lot of time and affect the construction progress of deep excavation projects.

Building information modeling (BIM) is not only capable of generating three-dimensional (3D) models of the design but also serves as a comprehensive information management platform tool [19]. Its advantages in information integration and management provide a potential solution for the realization of automatic modeling technology. However, in subway construction, engineering design and numerical calculation are often two independent parts, which are difficult to integrate. Some scholars have made efforts to develop the integrated BIM-FEM/FDM approach for tunnel excavation [20]. In deep excavation projects, Xie *et al.* proposed a BIM-RF integrated technology to implement the probability analysis based on random field theory [24]. Nevertheless, research on parameter back analysis based on the building information modeling-numerical simulation (BIM-NS) integrated technology is relatively limited. The complexity and variability of soil parameters, especially in soft soil environments, pose significant challenges for accurate geotechnical modeling. Current BIM-FEM/FDM integration approaches often lack precision in soil parameter evaluation, limiting their ability to adapt to complex subsurface conditions. To address this gap, a BIM-based back analysis method could seamlessly integrate numerical simulations (e.g., FEM/FDM) with real-time project data, enabling iterative calibration of soil properties during deep excavation.

Above all, the following existing gaps in deep excavation engineering need to be addressed: (1) In soft soil deep excavation projects, it is difficult to obtain accurate soil parameters, resulting in large deviations between numerical simulation results and measured values. (2) In conventional parameter back analysis workflows, iterative modeling processes often require semi-manual interventions (e.g., manual data conversion between BIM and numerical tools), which can lead to inefficiency in sample construction. Despite the use of computational algorithms, the absence of fully automated integration frameworks remains a challenge, particularly in scenarios requiring multi-software collaboration or multi-objective optimization.

In response to the above problems, this paper proposes a back analysis framework for soil parameters of deep excavation in soft soil based on the combination of BIM technology and ML algorithm. This framework effectively integrates the information integration function of BIM technology with the parameter inversion function of backpropagation neural network-nondominated sorting genetic algorithm II (BPNN-NSGA-II). The specific research content includes the following two aspects: (1) A BIM-NS calculation integrated technology is adopted, which can automatically construct numerical calculation samples of deep excavation in soft soil; (2) A multi-objective optimization inversion method of soil parameters is designed for deep excavation in soft soil.

2. Framework of BIM-NS-ML integrated technology

This study integrates the BIM-NS technology with the ML method to establish the framework for soil parameter identification and excavation response prediction (e.g. deformation) in deep excavation in soft soil. Figure 1 illustrates the schematic diagram of the proposed analysis framework, which comprises two main phases:

(1) The database of soil parameters and corresponding excavation responses is efficiently established based on BIM-NS technology;

(2) The mapping relation between soil parameters and excavation responses is constructed by the BPNN model, and then parameter inverse analysis is carried out by the multi-objective optimization algorithm.

It can be seen from Figure 1 that the parameter inversion and excavation response prediction method based on building information modeling-numerical simulation-machine learning (BIM-NS-ML) integrated technology involves two critical points: (1) How to combine BIM and numerical simulation to efficiently establish the database; (2) How to implement the multi-objective inversion procedure based on BPNN-NSGA-II.

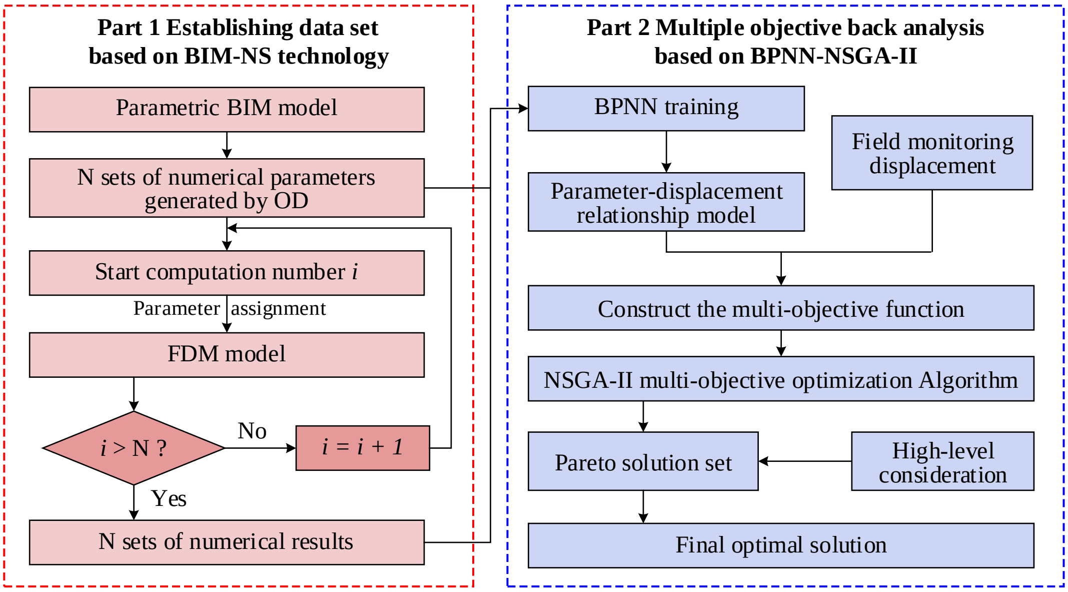

Figure 1. The framework of inversion procedure.

2.1. BIM-based solution for deep excavation in soft soils

Based on the work of XIE [24], the automatic execution of BIM-NS mainly requires two steps. Firstly, it is necessary to mesh the geometry of the BIM model to meet the computational requirements for numerical simulation; Secondly, it is necessary to automatically generate and execute control scripts for numerical simulation software that match the BIM model.

2.1.1. Object identification in deep excavation BIM model

In numerical simulations, struts within the excavation support system are generally simulated using structural elements, while soil and retaining structures are generally simulated using solid elements. Before conducting numerical simulation and analysis, solid elements require meshing, whereas structural elements need to be created based on the coordinates of specific key nodes. Given these differences, it is necessary to classify engineering objects into solid and non-solid objects when extracting them.

This paper involves three types of engineering objects: soil mass, diaphragm walls, and struts. Among the properties of these three types of engineering objects, the parameter that best reflects their specificity is the constitutive model. During the excavation, if appropriately designed, diaphragm walls are generally in an elastic state, hence diaphragm wall objects are labeled with the elastic constitutive model (elastic). Struts in the excavation area are commonly simulated with beam elements within structural elements, therefore strut objects are labeled with the beam element (beam) model. Among these, the constitutive model of soil mass is relatively complex; different soil constitutive models describe the stress-strain relationships of different soils. This study employs the Mohr-Coulomb model to describe the deformation characteristics of soil during the excavation. Therefore, it is necessary to pre-agree on the definitions of the various parameters of the Mohr-Coulomb model in the parsing program to facilitate the acquisition of data tag content. By traversing the constitutive model tags of each engineering object, soil objects and diaphragm wall objects can be identified, facilitating subsequent meshing.

In BIM, the vertical excavation sequence can be described using the storey that is inherited from IfcBuildingStorey class with a keyword *exc* tagged in the storey name. To accurately obtain the construction information of the excavation section, the key-value pair format is used to store and retrieve excavation information. All storeys with the keyword *exc* are parsed and extracted as construction information during the numerical simulation process.

In this paper, all attributes of engineering objects in the BIM model are labeled within their property sets in BIM. The properties are parsed and serialized in JSON format.

2.1.2. Geometric meshing processing

In this paper, tetrahedral mesh elements are used for mesh generation. Initially, the established BIM model is exported as an IFC standard file. Subsequently, the IFCOpenShell tool is employed to parse the existing IFC file and produce an OBJ geometric file corresponding to the BIM model. To ensure that each engineering object matches its attribute parameters, a UUID must be appended to the ‘g’ attribute in the OBJ file during the parsing process to generate the OBJ file. Then, the OBJ file is serialized into the Constructive Solid Geometry format within the NETGEN software. The geometric data is meshed using Delaunay and advancing front mesh generation algorithms, producing tetrahedral elements compatible with the finite volume method (FVM) spatial discretization scheme. Additionally, the meshed entities are named with their UUIDs.

Although UUID is a common technical detail in computer-aided design software, it is important to note that in this paper, the following additional explanations are necessary:

(1) During the process of traversing and parsing BIM primitives, the geometric data in the BIM model becomes separated from the parameter set it is attached to, in order to facilitate further

data processing. The geometric data needs to undergo meshing operations, while the parameter set requires further parsing to obtain specific parameter values. To accurately assign values to various entity objects in the subsequent computational software scripts (the parameter sets in the original BIM model), it is essential to bind the corresponding geometric entities with the parameter sets in the database using UUIDs.

(2) Taking the excavation project in this paper as an example, the numerical calculation process involves different types of engineering objects, including the diaphragm walls, soil, and struts. In the automated execution program, various engineering objects need to be categorized and labeled for processing. A prefix is added to the UUID, such as "1 –", to identify the category of the engineering object.

2.1.3. Automatic calculation script execution

To achieve a complete numerical simulation processing, primarily the initialization script and the construction scripts are required. The FLAC^{3D} is chosen as the target simulation software for illustration in this paper, which employs an explicit finite difference scheme for time discretization combined with a finite volume approach for spatial discretization [25]. The workflow of the BIM-NS processing is shown in Figure 2.

(1) Initialization script

The initialization script for a deep excavation simulation includes the following contents:

(a) Import the mesh geometry generated in Subsection 2.1.2.
(b) Assign the computational parameters. The UUID is adapted to match each solid zone with its computational parameters listed in the BIM model. The calculation parameters are assigned to the solid zone based on the parameter template of the corresponding constitutive model (e.g. Poisson's ratio, Young's modulus for the elastic model and cohesion, friction angle for the Mohr-Coulomb model, *etc.*).
(c) Determine the boundary conditions. Search for the boundary coordinate values of all entities and sort them to determine the overall model's boundaries. The zone face apply velocity command is used to constrain all displacements at the bottom of the model, as well as the normal displacements of the four vertical face boundaries.
(d) Solve the initial stress field. Use the solve command in FLAC^{3D} to reach the initial state of the numerical model.
(e) Clear the self-weight consolidation settlement. Use the zone gridpoint initialize command in FLAC^{3D} software to reset the settlement of the model consolidated under self-weight stress to zero.

(2) Construction script

After getting the initial stress field in the numerical model, the process of the excavation is simulated by executing the construction script. Notably, the number of excavation steps consists with the storeys named with the keyword exc in Subsection 2.1.1. The main contents are as follows:

(a) Delete the soil elements to be excavated. In each excavation process, soil objects above the current excavation elevation are deleted.

(b) Create the strut beam elements. The struts are created above the current excavation elevation with two key nodes extracted in Subsection 2.1.1. The properties of the struts follow the same data flow with solid zones in the initialization script.
(c) The solve command is used to calculate the state of the model after excavation.
(d) Execute steps 1–3 in a loop until all excavation steps are completed.

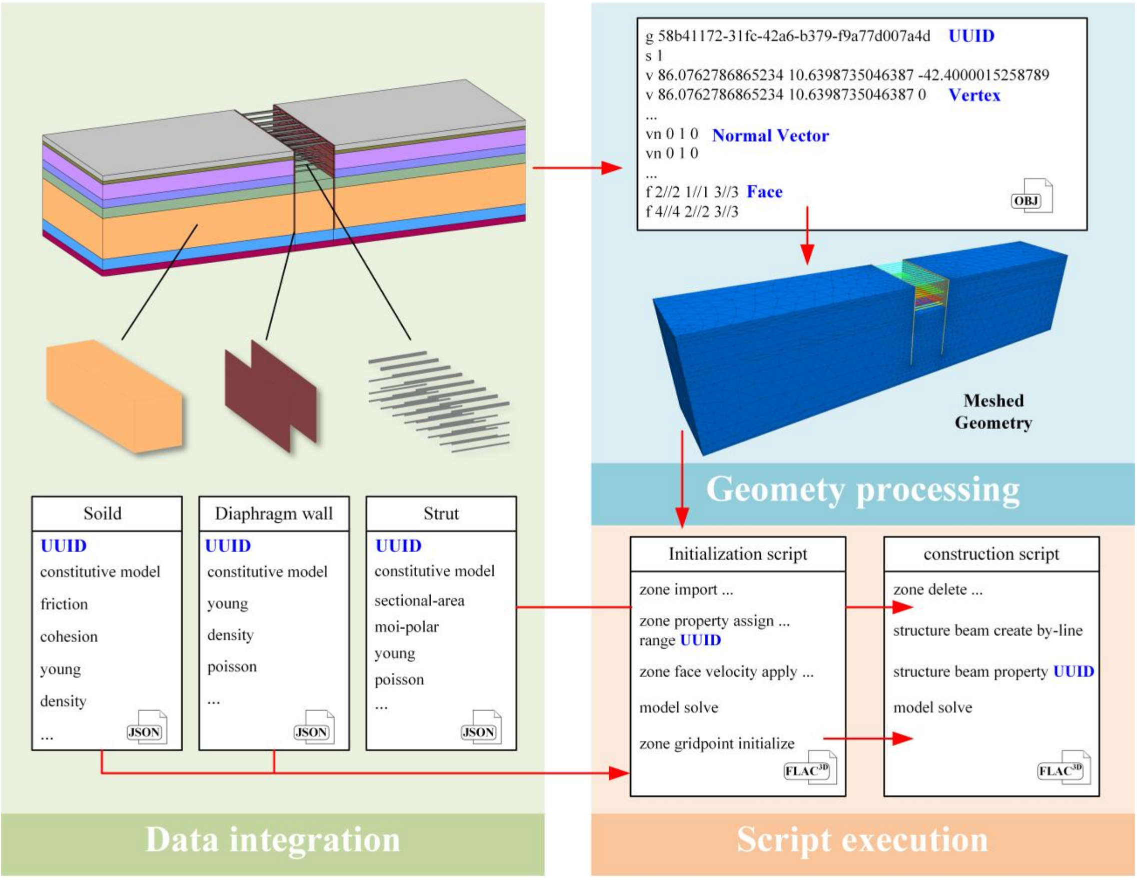

Figure 2. Workflow of BIM-NS processing.

2.2. Back-analysis method using multi-objective optimization

This section will comprehensively introduce the back-analysis method based on BPNN-NSGA-II. The BPNN model is used to replace the time-consuming numerical calculation process, and then the NSGA-II is used to obtain the Pareto solution set of soil parameters.

2.2.1. BPNN model

In the BIM-NS integrated technology introduced in Subsection 2.1, the excavation responses of deep excavation under different soil parameters are obtained. However, it is extremely time-consuming for a large number of numerical calculations to be called in the inverse analysis procedure. Because of its strong ability of nonlinear regression, the artificial neural network (ANN) is well-suited for modeling the complex relation between multi-dimensional soil parameters and soil responses [26]. Therefore, the ANN is introduced in this paper to replace the numerical calculations in the subsequent inverse analysis. This paper adopts the widely used BPNN model. As shown in Figure 3, the BPNN model is a neuron-like structure composed of one input layer, one or more hidden layers, and one output layer, which completes the nonlinear mapping between input and output through the information interaction among neurons. The architecture of the Backpropagation Neural Network (BPNN) in this study employs a three-layer structure with specifically designed activation functions. Between the input and hidden layers, a hyperbolic tangent sigmoid (*tansig*) function is implemented to introduce nonlinear transformation capabilities, while a linear transfer function (*purelin*) is adopted at the output layer. The number of neurons in the input and hidden layers is determined by the number of inversion parameters and monitoring points, respectively. The number of neurons in the hidden layer is determined by a trial-and-error procedure.

The process of constructing the sample database for the BPNN model based on BIM-NS calculation integrated technology is shown in Figure 3. Parametric BIM modeling is carried out based on Section 2.1 after determining the geometric parameters, some soil parameters (excluding soil parameters to be inversed), *etc*. Then, the computational command containing the information of inversion parameters is generated automatically. The inversion soil parameters are obtained using the OD (orthogonal design) method. The OD method selects some representative points that are uniformly dispersed and neatly comparable. Its superiority in multi-factor and multi-level experiments has been fully proved [29]. The corresponding mechanical responses of deep excavation can be obtained by numerical calculation of various parameters in batches. Inversion parameters are utilized as inputs for BPNN. Concurrently, the mechanical responses serve as outputs for constructing the prediction model. The trained model can replace the complex and time-consuming numerical calculation of deep excavation.

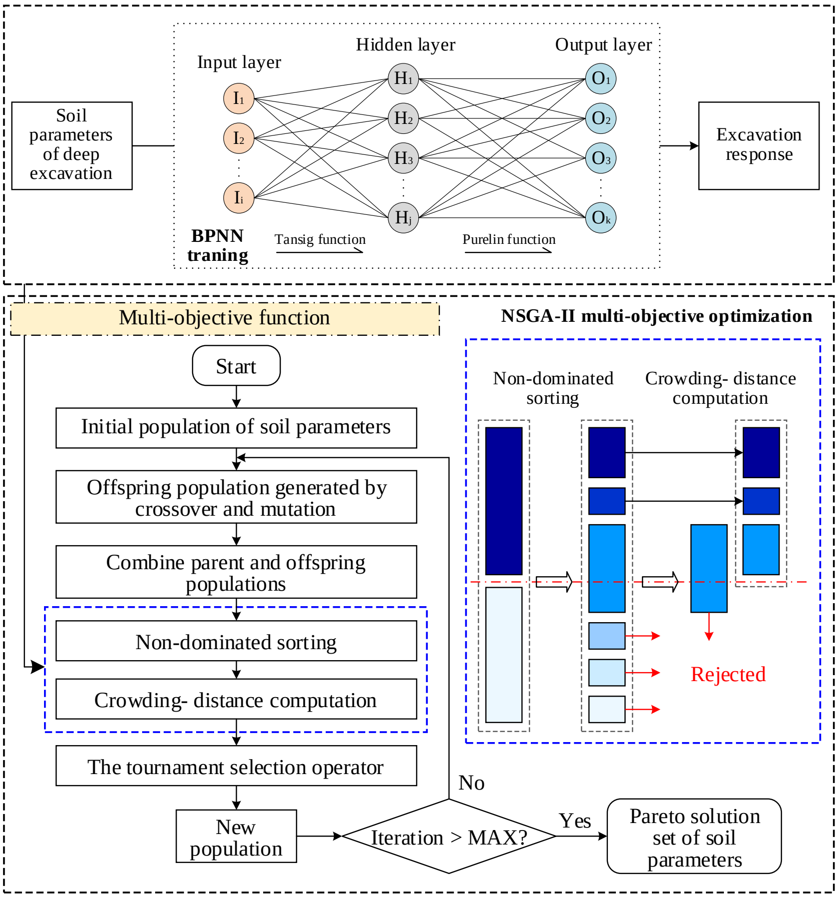

Figure 3. Workflow of BPNN-NSGA-II.

2.2.2. NSGA-II optimization algorithm

The NSGA-II optimization algorithm achieves parameter optimization through iterative updating of the population. It generates the Pareto optimal solution set by combining non-dominated sorting and crowding distance calculation [31]. The purpose of the multi-objective optimization method is to find the optimal solution set (*i.e.* Pareto front) within the possible range of soil parameters in each soil layer, so as to avoid the defect of local optimization of a single solution method (including the traditional NSGA method and the neural network method without considering the optimization steps).

The optimal solution of a multi-objective genetic algorithm optimization problem is in the form of a solution set, in which the objective function values of all solutions are not inferior to other individuals. As shown in Figure 4, when the number of objective functions is 2, the solution set is represented as a curve close to the two coordinate axes on the plane. When the number of objective functions is 3, the solution set is represented as a surface close to the coordinate axes in space. The multi-objective genetic algorithm has no specific requirements for the objective function, which can be determined according to the actual needs. In the back-analysis method, the purpose is to find a set of parameters to make the calculation values closest to the monitored values. The normalized mean squared error between calculated and monitored values is used as the objective function in this paper. Since the BPNN model is used instead of numerical calculation in Section 2.2.1, the objective function can be expressed as follows:

$$\begin{cases} \min F^1(\mathbf{X}) = \dfrac{1}{n_1}\sum_{i=1}^{n_1}[\dfrac{BPNN_i^1(\mathbf{X}) - \delta_i^1}{\delta_i^1}]^2 \\ \min F^2(\mathbf{X}) = \dfrac{1}{n_2}\sum_{i=1}^{n_2}[\dfrac{BPNN_i^2(\mathbf{X}) - \delta_i^2}{\delta_i^2}]^2 \\ \vdots \\ \min F^k(\mathbf{X}) = \dfrac{1}{n_k}\sum_{i=1}^{n_k}[\dfrac{BPNN_i^k(\mathbf{X}) - \delta_i^k}{\delta_i^k}]^2 \end{cases} \tag{1}$$

where, k is the number of the objective function; $\mathbf{X} = (x_1, x_2, ..., x_m)$ is a vector consisting of m inversion parameters; $BPNN_i^k(\mathbf{X})$ and δ_i^k represent the calculated value of the BPNN model and the field monitored value at the ith monitoring point in the kth objective function, respectively; n_k represents the number of monitoring points in the kth objective function.

The NSGA-II algorithm optimizes the objective function through the following steps [33]:

(1) The combination of soil parameters is randomly generated within the given range. Each set of parameters is called an individual. All individuals form a parent population P_0 with a population size of *N*.

(2) Generate the offspring population Qt (t = 0, 1, 2, …) based on selection, crossover, and mutation operations with a size of N. The parent and offspring populations of t generation are combined to form Rt with a population size of 2N.

(3) Perform the non-dominated sorting to generate the non-dominated set Fi (i denotes the ranking level). If the size of F1 is less than N, F2 is added to the next parent population Pt+1, and so on, until the population size is not less than N after the addition of generation Fm. If the population size is equal to N, stop. If the size is greater than N, the inferior individuals in Fm need to be eliminated based on crowding distance attributes, and the remaining individuals will serve as

the parent population of the next generation. Then, the next generation population Qt+1 is generated again using selection, crossover, and mutation operations, with a population size N.

(4) If the maximum number (Max) of generations is reached, all individuals in Pi+1 are output as the Pareto solution set. Otherwise, repeat steps (2) and (3).

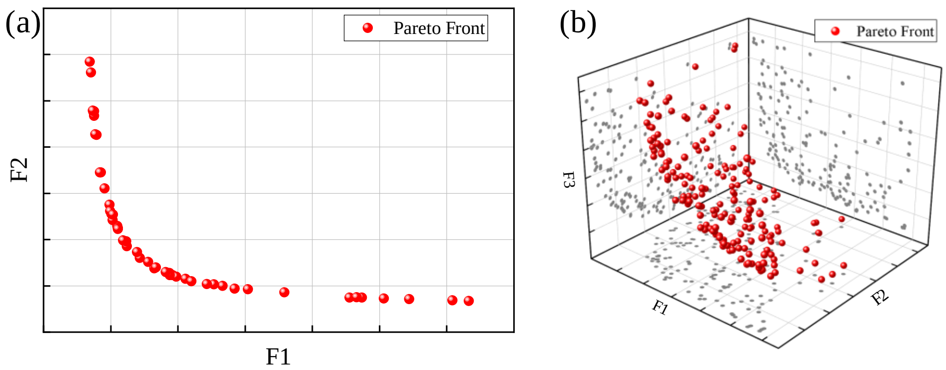

Figure 4. Schematic diagram of the Pareto solution: **(a)** 2 objective functions; **(b)** 3 objective functions.

2.3. Integration of BIM technology and ML-based inversion method

The inversion method based on BPNN-NSGA-II can effectively identify soil parameters but lacks the ability of numerical calculation to generate the required samples. However, BIM-NS technology provides an effective solution to this problem. This technology can integrate key engineering information into parametric models for batch numerical calculations. Therefore, this paper proposes an inversion method for deep excavation in soft soil based on BIM-NS-ML. In summary, this method includes the following stages. Firstly, sample data is generated based on BIM-NS integrated technology (Subsection 2.1) to construct the mapping relationship between soil parameters and excavation responses (Subsection 2.2.1). After ensuring the accuracy of the prediction model, the trained model is combined with NSGA-II for parameter inverse analysis. Finally, the inversion parameters are incorporated into the original BIM model, thereby enhancing the efficiency of project management.

3. Case study: a typical deep excavation in soft soil

3.1. Project background

In this paper, a subway station excavation project is considered for performing an inverse analysis based on BIM-NS-ML technology. The project is located in a soft soil area where the soil properties are complex. The input parameters are one of the key factors affecting the accuracy of numerical simulation, so it is necessary to clarify the soil parameters for the project. As shown in Figure 5 and Figure 6, the main dimensions of the parametric BIM model of this project are 76.5 m (*x* direction) × 184.8 m (*y* direction) × 55.4 m (*z* direction), the width of deep excavation is 21.1 m, and the depth of deep excavation is 15.7 m. This project is excavated in four stages, with excavation depths of −2 m, −8.7 m, −13 m, and −15.7 m. As shown in Figure 6, the parametric BIM model of this project involves 8 soil layers, and the basic properties of the soil layers based on the geological investigation report are shown in Table 1. In this paper, the Mohr-Coulomb constitutive model is used to describe the stress-strain behavior of soil.

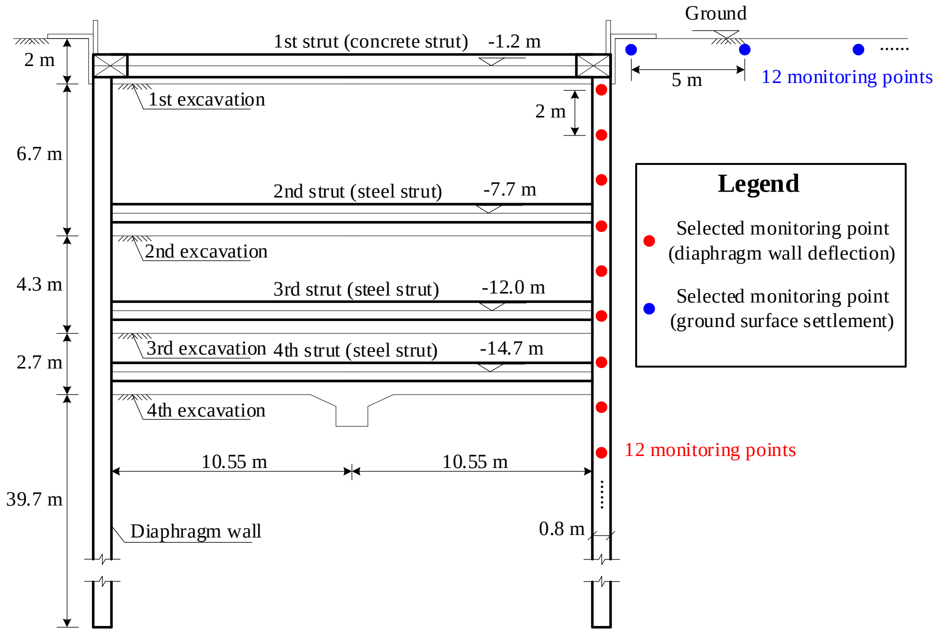

Figure 5. Cross-sectional view of the deep excavation (x = 35 m).

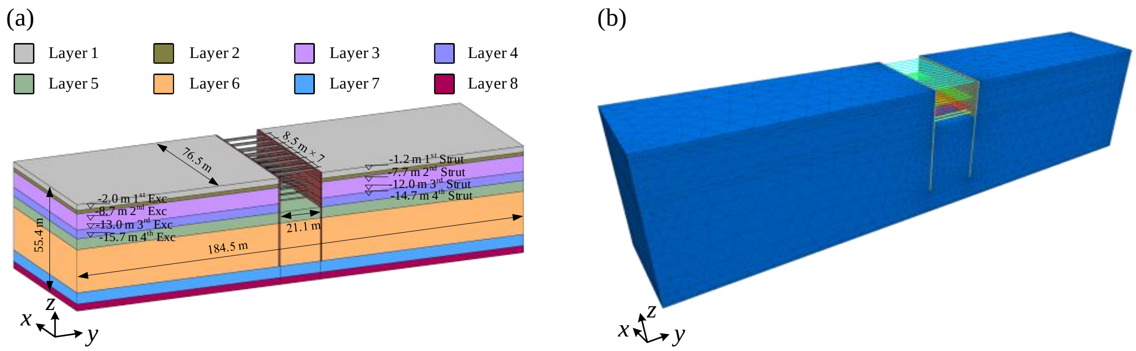

Figure 6. A typical case of subway station excavation: **(a)** BIM model; **(b)** FDM model.

The support system of this project comprises concrete diaphragm walls extending to a depth of 55.4 meters, complemented by a hybrid bracing system consisting of both concrete and steel struts. The struts in the first level are concrete square supports with cross-sectional dimensions of 800 mm (vertical) × 700 mm (horizontal). Subsequent support levels utilize structural steel pipe struts characterized by an 800 mm outer diameter and 16 mm wall thickness. The spacing between concrete support and steel support along the length of the deep excavation is 9 m and 3 m, respectively.

In this paper, the diaphragm walls are modeled as solid elements with a linear elastic material model, while the struts are modeled as beam structural elements. It is worth noting that no limiting plastic moment or plastic hinge is specified between beam elements, so beam elements behave as an isotropic, linearly elastic material with no failure limit [25]. The basic properties of the diaphragm walls and struts are presented in Table 1. To consider the contact behaviors between the soil layers and the diaphragm walls, interface elements are adopted, and the corresponding parameters are set as φ_i = 10° (friction angle), c_i = 10 kPa (cohesion), K_n = 40 MPa/m (normal stiffness), and K_s = 5 MPa/m (shear stiffness).

Table 1. Values of determinate material parameters.

Material	Thickness (m)	Young's modulus *E* (MPa)	Cohesion *c* (kPa)	Friction angle *φ* (°)	Poisson' ratio *μ*	Unit weight *γ* (kN/m³)
Soil layer 1	3.0	-	-	-	0.44	19.1
Soil layer 2	1.8	-	-	-	0.40	18.4
Soil layer 3	8.2	-	-	-	0.43	18.1
Soil layer 4	4.5	-	-	-	0.41	19.6
Soil layer 5	5.5	-	-	-	0.40	19.0
Soil layer 6	27.7	-	-	-	0.25	18.9
Soil layer 7	3.5	45	5	38	0.25	22.0
Soil layer 8	1.2	1200	50	20	0.35	27.0
Diaphragm wall	-	31000	-	-	0.2	25.0
Concrete strut	-	31000	-	-	0.2	25.0
Steel strut	-	210000	-	-	0.28	78.5

3.2. Implementation process of the BIM-NS-ML integrated framework

3.2.1. Preparation of BIM-NS calculation

Based on the engineering design documents, the BIM model is established accordingly. In this paper, a simple but effective entity inheritance is adopted, as shown in Figure 7. All the engineering objects, including soil mass, diaphragm wall, and struts are created from IfcBuildingElementProxy class with computational parameters tagged in the parameter set. The construction sequence represented as storeys are created from IfcBuildingStorey class with the keyword *exc* tagged in their names.

To generate samples for back analysis, the BIM-NS technology described in Subsection 2.1 is adopted to execute the excavation simulation automatically. The BIM model and FDM model are shown in Figure 6.

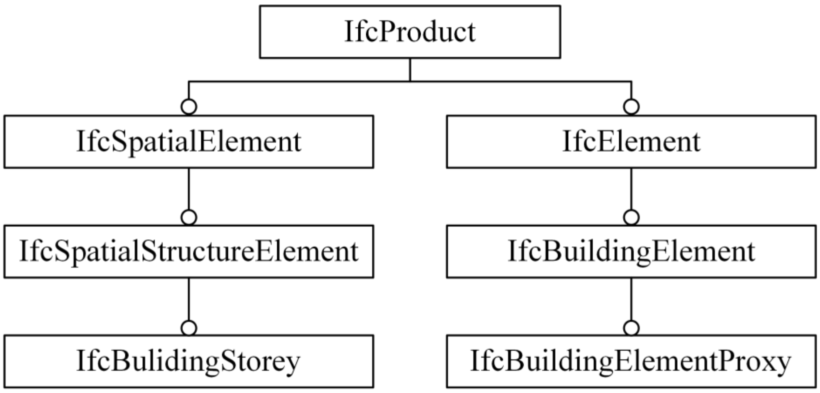

Figure 7. IFC standard description.

3.2.2. Selection of monitoring points and determination of parameter space

As shown in Figure 5, 24 typical monitoring points in the subway station deep excavation are selected as inversion monitoring points. These monitoring points are located in the middle of the length direction of the deep excavation (x = 35 m). Among them, 12 monitoring points are located inside the diaphragm wall to monitor horizontal displacement, with a spacing of approximately 2 m. The other 12 monitoring points are located on the surface outside the deep excavation to monitor surface settlement, with a spacing of approximately 5 m.

The diaphragm walls extend deep into the sixth soil layer, and the parameters of the upper six soil layers have a significant impact on the subway station excavation. To reduce the inversion dimension, the modulus, cohesion, and friction angle of the upper six soil layers are taken as unknown inversion parameters. Two objective functions of horizontal displacement and settlement displacement are established to form multiple objective functions. Based on Equation (1), the multi-objective function applicable to this project can be constructed as follows:

$$\begin{cases} \min F^1(\boldsymbol{E},\boldsymbol{c},\boldsymbol{\varphi}) = \dfrac{1}{12}\sum_{i=1}^{12}[\dfrac{BPNN_i^1(\boldsymbol{E},\boldsymbol{c},\boldsymbol{\varphi}) - \delta_i^1}{\delta_i^1}]^2 \\ \min F^2(\boldsymbol{E},\boldsymbol{c},\boldsymbol{\varphi}) = \dfrac{1}{12}\sum_{i=1}^{12}[\dfrac{BPNN_i^2(\boldsymbol{E},\boldsymbol{c},\boldsymbol{\varphi}) - \delta_i^2}{\delta_i^2}]^2 \end{cases} \quad (2)$$

where, F^1 and F^2 represent the objective functions of horizontal displacement and surface settlement, respectively; $\boldsymbol{E} = (E_1, E_2, ..., E_6)$ is the six deformation modulus parameters to be inverted; $\boldsymbol{c} = (c_1, c_2, ..., c_6)$ is the six cohesion parameters to be inverted; $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, ..., \varphi_6)$ is the six friction angle parameters to be inverted.

As shown in Table 2, the ranges of 18 inversion parameters are determined based on the geological investigation report. Firstly, three levels are determined within the range of inversion parameters, and an OD method is used to generate 64 sets of parameters in Table 3, which are used as input data for training the BPNN model. Subsequently, 3D forward simulations using the BIM-NS technology are performed for each set of parameters. The displacement calculation results of monitoring points are used as the output data for training the prediction model. Afterward, following several trial calculations, the BPNN model is finally determined to contain three layers, with the neural architecture of 18-10-24, as shown in Figure 3. To verify the accuracy of the trained BPNN model, 20 combinations of the inversion parameters are randomly generated within the inversion parameter space, and the calculation results of the BPNN model and the BIM-NS are compared. As shown in Figure 8, the two calculation results are highly consistent. It indicates that the BPNN model can build a mapping relation between soil parameters and displacement of the deep excavation, and can replace the time-consuming numerical calculation process for subsequent inversion analysis.

Finally, the NSGA-II algorithm is applied to search for the Pareto solution set. After the NSGA-II algorithm completes the optimization process, the final optimal solution is selected from the Pareto solution set according to the designer's higher-level consideration to facilitate subsequent prediction analysis.

Table 2. The range of soil parameters.

Modulus (MPa)	$\boldsymbol{E_1}$	$\boldsymbol{E_2}$	$\boldsymbol{E_3}$	$\boldsymbol{E_4}$	$\boldsymbol{E_5}$	$\boldsymbol{E_6}$
min	2	6	6	13	12	22
max	6	15	10	21	19	44
Cohesion (kPa)	$\boldsymbol{c_1}$	$\boldsymbol{c_2}$	$\boldsymbol{c_3}$	$\boldsymbol{c_4}$	$\boldsymbol{c_5}$	$\boldsymbol{c_6}$
min	7	11	8	6	22	23
max	13	20	16	10	40	43
Friction angle (°)	$\boldsymbol{\varphi_1}$	$\boldsymbol{\varphi_2}$	$\boldsymbol{\varphi_3}$	$\boldsymbol{\varphi_4}$	$\boldsymbol{\varphi_5}$	$\boldsymbol{\varphi_6}$
min	6	6	5	12	17	18
max	10	12	9	22	33	35

Table 3. Training samples.

No.	Modulus (MPa)						Cohesion (kPa)						Friction angle (°)					
	E_1	E_2	E_3	E_4	E_5	E_6	c_1	c_2	c_3	c_4	c_5	c_6	φ_1	φ_2	φ_3	φ_4	φ_5	φ_6
1	2.0	6.0	6.0	17.0	19.0	33.0	13.0	11.0	8.0	10.0	31.0	33.0	6.0	9.0	5.0	22.0	17.0	35.0
2	2.0	6.0	6.0	21.0	12.0	44.0	10.0	15.5	12.0	6.0	40.0	43.0	6.0	6.0	9.0	12.0	25.0	18.0
3	4.0	6.0	6.0	13.0	12.0	33.0	10.0	21.0	8.0	10.0	22.0	43.0	8.0	6.0	7.0	12.0	17.0	26.5
4	2.0	10.5	6.0	21.0	12.0	22.0	10.0	11.0	8.0	6.0	22.0	33.0	10.0	12.0	5.0	22.0	25.0	26.5
5	4.0	6.0	6.0	17.0	12.0	22.0	7.0	11.0	16.0	8.0	40.0	23.0	8.0	6.0	5.0	12.0	25.0	35.0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
30	6.0	6.0	8.0	17.0	12.0	22.0	7.0	15.5	16.0	6.0	22.0	43.0	10.0	6.0	5.0	22.0	17.0	26.5
31	4.0	10.5	8.0	21.0	12.0	44.0	7.0	11.0	8.0	6.0	40.0	43.0	6.0	12.0	7.0	12.0	17.0	18.0
32	6.0	10.5	6.0	13.0	12.0	22.0	10.0	11.0	16.0	8.0	31.0	43.0	6.0	9.0	5.0	22.0	17.0	18.0
33	2.0	6.0	8.0	17.0	19.0	44.0	10.0	15.5	8.0	6.0	22.0	23.0	6.0	9.0	5.0	12.0	33.0	26.5
34	4.0	15.0	6.0	13.0	12.0	22.0	10.0	11.0	12.0	10.0	40.0	33.0	6.0	9.0	5.0	17.0	17.0	18.0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
60	2.0	6.0	6.0	21.0	15.5	44.0	13.0	11.0	8.0	8.0	40.0	43.0	6.0	9.0	5.0	17.0	17.0	26.5
61	2.0	15.0	6.0	13.0	15.5	33.0	7.0	11.0	8.0	10.0	40.0	23.0	10.0	6.0	5.0	12.0	25.0	18.0
62	2.0	6.0	6.0	13.0	19.0	22.0	7.0	15.5	12.0	10.0	22.0	23.0	6.0	12.0	7.0	22.0	25.0	35.0
63	6.0	10.5	6.0	13.0	19.0	22.0	13.0	21.0	8.0	6.0	40.0	33.0	6.0	6.0	7.0	12.0	25.0	26.5
64	6.0	10.5	6.0	17.0	15.5	44.0	7.0	11.0	12.0	6.0	22.0	23.0	6.0	6.0	5.0	12.0	17.0	35.0

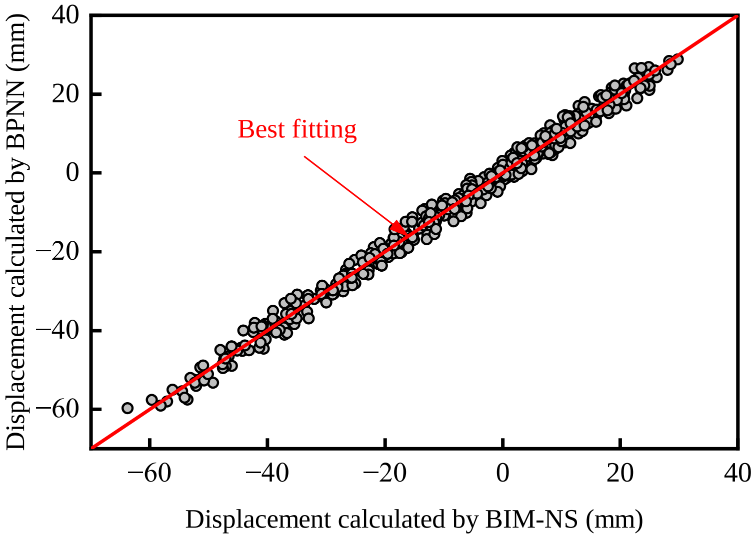

Figure 8. Comparison of displacement at the monitoring point calculated by the BPNN model and the BIM-NS technology.

3.3. Results of inverse analysis

3.3.1. Identification of inversion parameters

The population size, maximum number of iterations, probability of crossover, and probability of mutation for the NSGA-II algorithm are set as 100, 100, 0.7, and 0.4, respectively. When the iteration is complete, consider the following relative error function to select the final optimal solution from the Pareto solution set.

$$e = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{f_i - \delta_i}{\delta_i}\right)^2 \tag{3}$$

where, f_i represents the calculated displacement of the ith monitoring point based on the BIM-NS technology; δ_i represents the measured displacement of the ith monitoring point. In this project, n = 12 (12 diaphragm wall deflection monitoring points and 12 ground surface settlement monitoring points).

Table 4 shows the relative errors for the solutions in the Pareto-optimal set based on Equation (2). The average represents the average of relative errors for all solutions in the Pareto-optimal set. The best and the worst represent the lowest and highest relative errors in the Pareto-optimal set, respectively. As shown in Table 4, the lowest relative error is 3.2 %, corresponding to the best solution. Therefore, this solution is selected as the final solution, as shown in Table 5.

Table 4. The relative errors for the solution in the Pareto-optimal set.

	The best	**The worst**	**The average**
Relative errors (%)	3.2	12.8	8.9

Table 5. The parameter of inverse analysis.

Modulus (MPa)	E_1	E_2	E_3	E_4	E_5	E_6
value	4.47	12.07	8.02	16.38	13.88	32.84
Cohesion (kPa)	c_1	c_2	c_3	c_4	c_5	c_6
value	11.06	15.14	12.93	8.33	32.02	38.06
Friction angle (°)	φ_1	φ_2	φ_3	φ_4	φ_5	φ_6
value	5.12	7.66	11.08	20.63	26.46	31.00

3.3.2. Comparison of calculated displacements and monitored values

The comparison between the calculated displacement values and the monitored values of the two selected types of monitored data (diaphragm wall deflection and ground surface settlement) is shown in Figure 9(a) and Figure 9(b), respectively. The forward calculated displacement value through the final optimal solution obtained based on the multi-objective inversion method is very close to the corresponding monitored value. In particular, some displacements are almost the same as the monitored values, such as the maximum diaphragm wall deflection occurs near the surface and the fourth excavation step, and the maximum surface settlement occurs approximately 10.5 m from the edge of the deep excavation. In addition, the trend of calculated displacements with position generally matches the actual measurements. This indicates that the inverse analysis method for deep excavation in soft soil is reliable, laying the foundation for subsequent construction prediction.

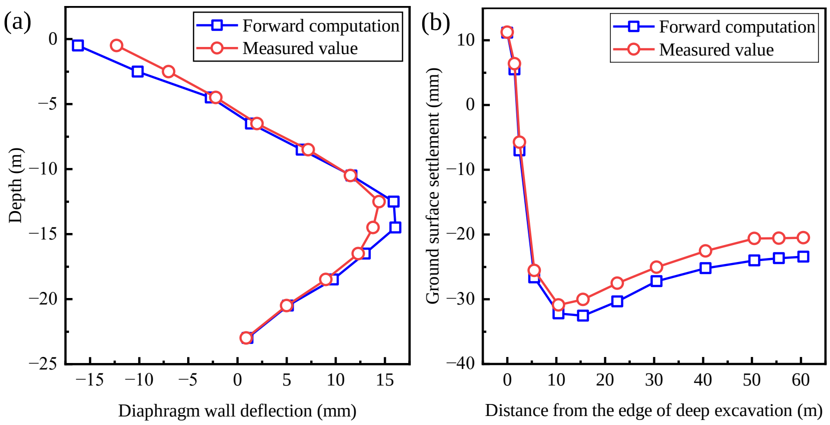

Figure 9. Comparison of monitored values and forward computation based on inversion parameters: **(a)** Diaphragm wall deflection; **(b)** Ground surface settlement.

3.3.3. Efficiency of inversion analysis

The efficiency of the proposed back-analysis framework is one of its key advantages. In the framework, a total of approximately 150 numerical simulations are required to construct the sample data for training the neural network model. Once the neural network is trained, subsequent back-analysis can be performed by simply calling the pre-trained neural network model, which significantly reduces the computation time. Compared to the time required for numerical simulations, the time needed to use the neural network model for parameter inversion is negligible. The numerical simulations for deep excavation in this paper require 5 minutes, using a desktop with an Intel CORE i7-13700. Therefore, a single inversion analysis takes about 12.5 hours.

In contrast, if the surrogate model is not used, the process would involve direct numerical calculations for each inversion iteration. This would require approximately 2500 numerical simulations, making inversion analysis considerably more time-consuming. A single inversion analysis takes about 7 days. Therefore, by utilizing the neural network-based approach, the proposed framework offers a substantial reduction in computational cost and time, making it a more efficient solution for deep excavation analysis.

4. Conclusion

In response to the issue of sample construction inefficiency in inversion methods based on multi-objective genetic algorithm optimization, this paper proposes a BIM-NS-based framework of BIM-NS-ML. In this framework, all calculation parameters are integrated into the BIM model of deep excavation. Through geometric transformation and automated script execution, the BIM model automatically executes numerical simulations, conveniently populating the database with the data samples needed for surrogate model construction. The inversion parameters are optimized and updated in the corresponding BIM model, achieving more efficient project management.

(1) A BIM-NS-ML framework is established to execute the back analysis, enhanced by automatic numerical simulation of the deep excavation. Repetitive numerical simulation modeling and modification processes are improved in BIM integration.

(2) A multi-objective back-analysis framework for deep excavation in soft soil is proposed. In this paper, the BIM-NS technology is integrated with the BPNN-NSGA-II inverse analysis method to achieve rapid identification of soil parameters based on multi-source monitored data. The framework has a high degree of versatility, theoretically allowing its application to various soil types and engineering projects of different scales.

(3) A multi-objective inverse analysis is applied to a real case of deep excavation, and the results show that the method can effectively identify parameters and provide an effective tool for deformation prediction and safety assessment.

Acknowledgments

This study has been supported by the National Natural Science Foundation of China (NSFC) (Grant Nos. 52122806, 52308372, and 52338007).

Conflicts of interests

The authors declare no conflicts of interest.

Authors' contribution

Shu Jiang: Methodology, Software, Writing – original draft. Ziqian Li: Validation, Writing – review & editing. Rongjun Zhang: Conceptualization, Funding acquisition. Junjie Zheng: Supervision, Methodology. Shaoyan Zhou: Data curation, Methodology. All authors reviewed the results and approved the final version of the manuscript.

References

[1] Lin D, Broere W, Cui J. Underground space utilisation and new town development: Experiences, lessons and implications. *Tunn. Undergr. Sp. Tech*. 2022, 119:104204.

[2] He R, Tiong RLK, Yuan Y, Zhang L. Enhancing resilience of urban underground space under floods: Current status and future directions. *Tunn. Undergr. Sp. Tech*. 2024, 147:105674.

[3] Pan Y, Qin J, Zhang L, Pan W, Chen JJ. A probabilistic deep reinforcement learning approach for optimal monitoring of a building adjacent to deep excavation. *Comput. Aided Civ. Inf. Eng.* 2024, 39(5):656–678.

[4] Gao Y, Ding Z. Case study of post uplift in deep excavation of a subway station in thick soft clay using long pile foundations. *Undergr. Sp.* 2022, 7(2):254–267.

[5] Wojciechowski M. Application of artificial neural network in soil parameter identification for deep excavation numerical model. *Comput. Assist. Methods Eng. Sci.* 2017, 18(4):303–311.

[6] Huang ZH, Zhang LL, Cheng SY, Zhang J, Xia XH. Back-analysis and parameter identification for deep excavation based on Pareto multiobjective optimization. *J. Aerosp. Eng.* 2015, 28(6):A4014007.

[7] Knabe T, Datcheva M, Lahmer T, Cotecchia F, Schanz T. Identification of constitutive parameters of soil using an optimization strategy and statistical analysis. *Comput. Geotech.* 2013, 49:143–157.

[8] Zhang Y, Gallipoli D, Augarde C. Parameter identification for elasto-plastic modelling of unsaturated soils from pressuremeter tests by parallel modified particle swarm optimization. *Comput. Geotech.* 2013, 48:293–303.

[9] Jin YF, Yin ZY, Riou Y, Hicher PY. Identifying creep and destructuration related soil parameters by optimization methods. *KSCE. J. Civ. Eng.* 2017, 21:1123–1134.

[10] Moreira N, Miranda T, Pinheiro M, Fernandes P, Dias D, *et al.* Back analysis of geomechanical parameters in underground works using an Evolution Strategy algorithm. *Tunn. Undergr. Sp. Tech*. 2013, 33:143–158.

[11] Li J, Chen C, Wu Z, Chen J. Multi-source data-driven unsaturated seepage parameter inversion: Application to a high core rockfill dam. *J. Hydrol.* 2023, 617:129171.

[12] Li S, Liu Y, Zhan S. Parameter estimate approach for groundwater pollution model with ant colony algorithm. *J. Comput. Inf. Syst.* 2008, 4(6):2803–2808.

[13] Patino-Ramirez F, Layhee C, Arson C. Horizontal directional drilling (HDD) alignment optimization using ant colony optimization. *Tunn. Undergr. Sp. Tech*. 2020, 103:103450.

[14] Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE T. Evolut. Comput.* 2002, 6(2):182–197.

[15] Xu Y, Wu Z. Parameter identification of unsaturated seepage model of core rockfill dams using principal component analysis and multi-objective optimization. *Structures*, 2022, 45:145–162.

[16] Jin YF, Yin ZY, Shen SL, Hicher PY. Selection of sand models and identification of parameters using an enhanced genetic algorithm. *Int. J. Numer. Anal. Met.* 2016, 40(8):1219–1240.

[17] Jia Y, Chi S. Back-analysis of soil parameters of the Malutang II concrete face rockfill dam using parallel mutation particle swarm optimization. Comput. Geotech. 2015, 65:87–96.

[18] Y. Sheu G. Direct back analysis by the meshless local Petrov-Galerkin method and Bayesian statistics. Int. J. Numer. Anal. Methods Geomech. 2006, 30(8):823–842.

[19] Campbell DA. Building information modeling: the Web3D application for AEC. *Proc. 12th Int. Conf. 3D Web Technol*. 2007:173–176.

[20] De Matteis F, Orci C, Bilosi S, Benedetti G. The 3D-BIM-FEM modeling of the mairie des Lilas Paris metro station line 11–from design to execution. *Tunnels and Underground Cities. Engineering and Innovation Meet Archaeology, Architecture and Art.* CRC Press, 2019. pp. 1998–2007.

[21] Sharafat A, Khan MS, Latif K, Seo J. BIM-based tunnel information modeling framework for visualization, management, and simulation of drill-and-blast tunneling projects. *J. Comput. Civ. Eng.* 2021, 35(2):04020068.

[22] Ninic J, Alsahly A, Vonthron A, Bui HG, Koch C, *et al.* From digital models to numerical analysis for mechanised tunnelling: A fully automated design-through-analysis workflow. *Tunn. Undergr. Sp. Tech*. 2021, 107:103622.

[23] Fabozzi S, Biancardo SA, Veropalumbo R, Bilotta E. I-BIM based approach for geotechnical and numerical modelling of a conventional tunnel excavation. *Tunn. Undergr. Sp. Tech*. 2021, 108:103723.

[24] Xie P, Zhang R, Zheng J, Li Z. Probabilistic analysis of subway station excavation based on BIM-RF integrated technology. *Autom. Constr.* 2022, 135:104114.

[25] Itasca. *FLAC 3D: Fast Lagrangian analysis of continua in 3 dimensions*, 6th ed. *Minneapolis*, MN: Itasca Consulting Group, 2019.

[26] Rafiai H, Jafari A, Mahmoudi A. Application of ANN-based failure criteria to rocks under polyaxial stress conditions. *Int J Rock Mech Min.* 2013, 59:42–49.

[27] Zhou CB, Liu W, Chen YF, Hu R, Wei K. Inverse modeling of leakage through a rockfill dam foundation during its construction stage using transient flow model, neural network and genetic algorithm. *Eng. Geol.* 2015, 187:183–195.

[28] Zhang M, Hu L, Yao L, Yin W. Surrogate models for sub-region groundwater management in the Beijing plain, China. *Water* 2017, 9(10):766.

[29] Gong W, Cai Z, Jiang L. Enhancing the performance of differential evolution using orthogonal design method. *Appl. Math. Comput.* 2008, 206(1):56–69.

[30] Ge Z, Gao Z, Sun R, Zheng L. Mix design of concrete with recycled clay-brick-powder using the orthogonal design method. *Constr. Build. Mater.* 2012, 31:289–293.

[31] Kao GK, Jacobson SH. Finding preferred subsets of Pareto optimal solutions. *Comput. Optim. Appl.* 2008, 40:73–95.

[32] Singh P, Kumar SD, Singh RK. Fuzzy efficient and Pareto-optimal solution for multi-objective linear fractional programming problems. *Int. J. Math. Oper. Res.* 2014, 6(3):357–376.

[33] Deep K, Thakur M. A new mutation operator for real coded genetic algorithms. *Appl. Math. Comput*. 2007, 193(1):211–230.