Applied Energy 404 (2026) 127170

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

An integrated architecture for fixed truck assignment optimization in open-pit mines: Synergistic optimization, hybrid modeling, and multi-objective optimization

Hongyang Xu [a,b,*], Hongze Zhao [a,e,**], Junyu Lu [a], Hui Zuo [b], Wen Zheng [a], Pei Guo [d], Di Yang [c], Bo Zhang [c], Qipei Mei [b,***]

[a] School of Energy and Mining Engineering, China University of Mining & Technology-Beijing, Beijing 100083, China
[b] Department of Civil and Environmental Engineering, University of Alberta, Edmonton T6G 2W2, Alberta, Canada
[c] School of Mining and Petroleum Engineering, University of Alberta, Edmonton T6G 2W2, Alberta, Canada
[d] Heidaigou Open-Pit Mine, China National Energy, Erdos 010399, China
[e] State Key Laboratory for Geomechanics and Deep Underground Engineering, China University of Mining & Technology-Beijing, Beijing 100083, China

HIGHLIGHTS

- Optimization of transport grouping and fleet speed for varying haul distances.
- Integration of Kernel PCA-ANN into the Mixed Integer Programming model.
- Adaptive reference point-based NSGA-III incorporates letter and real-valued encoding.
- Nonworking time for shovels and trucks is reduced by 679 h and 1.65×10^5 h.
- Fuel consumption is reduced by 6.2×10^6 L, equivalent to 1.65×10^4 tons (8.3 %) of CO_2.

ARTICLE INFO

Keywords:
Fixed truck assignment
Synergistic optimization
KPCA-ANN
Adaptive reference point-based NSGA-III
Equipment utilization
Energy consumption
CO_2 emission

ABSTRACT

Fixed truck assignment (FTA) is a prevalent method of truck dispatching in open-pit mining. To improve equipment utilization and reduce energy consumption, this paper proposes an integrated architecture for optimizing FTA. The innovations of this architecture include three aspects: (1) The concept of synergistic optimization is introduced into FTA for the first time, with a proposed approach of "haul distance–transport grouping–fleet speed." This approach establishes the relationship between equipment utilization, energy consumption, and their influencing factors. (2) A hybrid modeling method is designed, integrating Kernel Principal Component Analysis and Artificial Neural Network (KPCA-ANN) into the programming model to represent the NP-hard nature of the FTA optimization. (3) An adaptive reference point-based Nondominated Sorting Genetic Algorithm III (ARP-NSGA-III) is introduced, incorporating both letter and real-valued encoding to solve the $3 \times b$-objective ($\forall b \in Z^+$) optimization problem across various route topologies. Finally, the integrated architecture is evaluated using historical data from a case mine. The results show that the optimal FTA solution reduces the shovel idle time (SIT) by 679 h, truck waiting time (TWT) by 1.65×10^5 h, and truck fuel consumption (TFC) by 6.2×10^6 L. The reduction in fuel consumption (FC) is equivalent to reducing 1.65×10^4 tons of CO_2 emissions, accounting for 8.3 % of the total CO_2 emissions from haul activities.

* Corresponding author at: School of Energy and Mining Engineering, China University of Mining & Technology-Beijing, Beijing 100083, China; Department of Civil and Environmental Engineering, University of Alberta, Edmonton T6G 2W2, Alberta, Canada
** Corresponding author at: School of Energy and Mining Engineering, China University of Mining & Technology-Beijing, Beijing 100083, China; State Key Laboratory for Geomechanics and Deep Underground Engineering, China University of Mining & Technology-Beijing, Beijing 100083, China
*** Corresponding author at: Department of Civil and Environmental Engineering, University of Alberta, Edmonton T6G 2W2, Alberta, Canada
E-mail addresses: hyangxucumtb@126.com (H. Xu), 201223@cumtb.edu.cn (H. Zhao), qipei.mei@ualberta.ca (Q. Mei).

https://doi.org/10.1016/j.apenergy.2025.127170
Received 25 February 2025; Received in revised form 15 October 2025; Accepted 25 November 2025
Available online 7 December 2025

Nomenclature

FTA	Fixed truck assignment
KPCA-ANN	Kernel Principal Component Analysis and Artificial Neural Network
ARP-NSGA-III	Adaptive reference point-based Nondominated Sorting Genetic Algorithm III
SIT	Shovel idle time
TWT	Truck waiting time
TFC	Truck fuel consumption
FC	Fuel consumption
MOO	Multi-objective optimization
IP	Integer Programming
MIP	Mixed Integer Programming
NSGA-II/III	Nondominated Sorting Genetic Algorithm II/III
MOEA/D	Multi-Objective Evolutionary Algorithm based on Decomposition
SPEA2	Strength Pareto Evolutionary Algorithm 2
ILP	Integer Linear Programming
PCA-ANN	Principal Component Analysis and Artificial Neural Network
LHS	Loaded hauling speed
EHS	Empty hauling speed
EVD	Eigenvalue Decomposition
MAPE	Mean Absolute Percentage Error
POS	Pareto optimal solution
GPS	Global Positioning System
Leaky ReLU	Leaky Rectified Linear Unit
RBF	Radial Basis Function
HV	Hypervolume
NSGA-III UCAM	NSGA-III with uniform crossover and adaptive mutation
RVEA	Reference Vector-guided Evolutionary Algorithm
AGE-II	Approximation-Guided Evolution II

1. Introduction

Fixed truck assignment (FTA) is a widely used method for dispatching trucks in open-pit mining. Its core concept is to assign a specific type and number of trucks to each haul route based on production demands and operational constraints [1,2]. As shown in Fig. 1, FTA effectively addresses the resource allocation problem in transportation tasks by providing an initial assignment scheme consisting of a heterogeneous fleet. Unlike dynamic scheduling methods that rely on real-time communication, FTA does not require real-time data support, which results in lower system development costs and higher stability [3,4]. Especially in large-scale and long-term transportation tasks, FTA remains the most common and reliable dispatching method. Therefore, choosing FTA as the optimization object holds significant theoretical importance and has a substantial impact on practical applications.

Improving equipment utilization and reducing energy consumption are key to optimizing FTA [1,4], which is a typical multi-objective optimization (MOO) problem. The first step in achieving MOO is to establish the relationship between influencing factors and optimization objectives. Existing FTA optimization approaches mainly focus on a single factor, adjusting decision variables related to this factor to find the optimal solution [3,5]. However, studies [2,6] and production practices have shown that considering only a single factor makes it difficult to generate a reasonable initial truck assignment scheme. This challenge arises from the complexity of the open-pit mining system, where multiple factors (such as transport grouping and fleet speed) impact both equipment utilization and energy consumption [3,7].

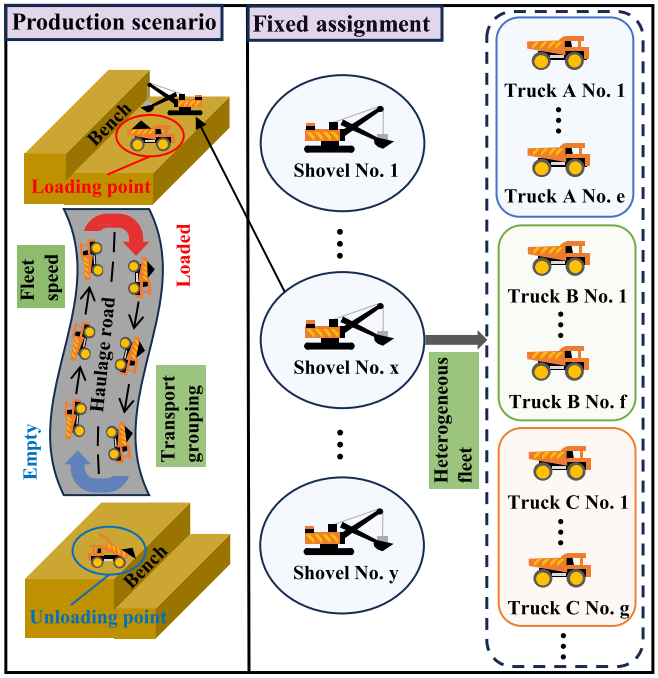

Fig. 1. FTA schematic diagram.

Constructing a mathematical model to quantify the optimization objective is the second step in achieving MOO. FTA modeling relies on programming models, such as Integer Programming (IP) [3,6] and Mixed Integer Programming (MIP) [1,2]. These models typically combine knowledge-based and mechanism-based approaches to describe FTAs of different scales. However, knowledge-based modeling depends on expert knowledge, rules, and experience [8,9], and inaccurate or incomplete information can reduce model accuracy. Mechanism-based modeling describes system interactions and dynamic processes through mathematical or physical models [10,11]. When the system's internal mechanisms are highly complex, the modeling process becomes cumbersome, which may result in insufficient model generalization.

Selecting the optimization technique to solve the mathematical model is the third step in achieving MOO. MOO algorithms are mainly divided into two categories: traditional and heuristic algorithms. Traditional MOO algorithms, such as the weighted sum aggregation method [12] and ideal point method [13], derive optimal solutions based on mathematical optimization theory. However, the FTA optimization problem is a typically high-dimensional, multi-objective, and nonlinear NP-hard problem [14,15], increasing the computational complexity of the traditional MOO algorithm. Heuristic MOO algorithms, such as Nondominated Sorting Genetic Algorithm II/III (NSGA-II/III) [16,17], Multi-Objective Evolutionary Algorithm based on Decomposition (MOEA/D) [18], and Strength Pareto Evolutionary Algorithm 2 (SPEA2) [19], along with their improved versions [15,20,21], are based on heuristic search and simulate the natural evolution process. However, these algorithms have limitations when dealing with irregular, non-uniform, and non-continuous Pareto frontier problems. In particular, the Pareto front obtained from FTA optimization usually has an unknown shape, which presents higher requirements for the search ability of heuristic MOO algorithms in multi-dimensional spaces.

This paper proposes a synergistic optimization approach for FTA and applies it to hybrid modeling. The MOO technique is then used to solve the hybrid model. The contributions of this study are as follows:

(1) A "haul distance–transport grouping–fleet speed" synergistic optimization approach is proposed, providing new theoretical guidance for subsequent FTA multi-objective modeling.
(2) A hybrid model combining mechanism-based and data-driven models is designed to represent the objective functions related to time and fuel consumption (FC).
(3) An adaptive reference point-based Nondominated Sorting Genetic Algorithm III (ARP-NSGA-III) is introduced, using both letter and real-value encoding to solve MOO problems with heterogeneous parameters.
(4) In the case study, the designed integrated architecture significantly improves equipment utilization and reduces energy consumption. The optimal FTA solution reduces the shovel idle time (SIT) by 679 h, truck waiting time (TWT) by 1.65×10^5 h, and truck fuel consumption (TFC) by 6.2×10^6 L, compared to the reference value.

This paper is organized as follows: Section 2 reviews the research progress on FTA optimization; Section 3 defines the FTA optimization problem using a mathematical model; Section 4 describes the various components of the integrated architecture; Section 5 presents the current status of the case mine and the parameter design of the integrated architecture; Section 6 evaluates the performance of the integrated architecture and demonstrates its application effect; Section 7 summarizes the paper and discusses future research directions.

2. Related work

2.1. Modeling methods

Optimizing FTA is primarily achieved by constructing objective functions through programming models, which can be categorized into two types: single-objective and multi-objective. Single-objective models are widely used due to their simplicity and efficiency. For example, Rodrigo et al. (2013) [1] proposed an MIP model aimed at maximizing fleet productivity, taking into account the reliability, availability, and maintainability of equipment. Zhang et al. (2015) [5] developed an IP model to minimize operating costs of trucks, establishing the relationship between fleet size and operating expenses. Similarly, Ristoski et al. (2017) [3] constructed an IP model designed to minimize number of trucks, applicable to both homogeneous and heterogeneous fleets. Zhang et al. (2022) [4] proposed an MIP model focused on reducing the energy consumption of autonomous truck fleets.

Although single-objective modeling can represent objective functions in many cases, it struggles to capture the multidimensional complexity of FTA optimization. As a result, multi-objective modeling has become a more applicable method, as it can simultaneously address multiple conflicting or competing objectives. For example, Konur et al. (2013) [2] proposed a bi-objective MIP model aimed at minimizing service cost and cost range. Similarly, Vandani et al. (2019) [14] designed a bi-objective MIP model to minimize operating costs while maximizing workload. Some scholars have extended this method by incorporating more objectives into their models. Bakhtavar et al. (2020) [6] constructed a three-objective IP model aimed at minimizing truck operation cycle, assignment cost, and number of trucks. Mohtashami et al. (2020) [7] proposed a three-objective MIP model focused on reducing operating time, product delivery lateness, and earliness.

The multi-objective modeling method used in dynamic scheduling optimization provides valuable insights for FTA optimization. For instance, Afrapoli et al. (2019) [22] proposed a three-objective mixed Integer Linear Programming (ILP) model, aiming to minimize SIT, TWT, and deviations in flow rates. Zhang et al. (2020) [15] developed a three-objective ILP model, focused on minimizing haul cost, TWT, and ore grade deviation. However, the objective functions in these models rely on knowledge-based and mechanism-based approaches, which can limit their generalization. This issue becomes more pronounced when modeling dynamic characteristics, such as FC [23,24].

2.2. MOO technologies

MOO technology can be divided into two categories: traditional and heuristic algorithms. Traditional MOO algorithms, based on mathematical optimization theory, can ensure the optimality of the solution. For instance, Chou et al. (2015) [12] achieved a trade-off between the information rate and harvested power in the system through semi-definite relaxation-based weighted aggregation. Fukumoto et al. (2019) [13] explored how estimation methods for ideal and nadir points impact the optimization scheme. Shahabi et al. (2021) [25] applied an augmented quadratic epsilon constraint to solve a three-objective MIP model and addressed the assignment problem of heterogeneous fleets in cross-dock tasks.

The traditional MOO algorithm is widely applicable to continuous, linear, and convex problems. However, for high-dimensional and nonlinear NP-hard problems, its computational complexity increases sharply, and it is prone to falling into local Pareto optima. To address this issue, heuristic MOO algorithms use a heuristic search mechanism to avoid local optima and reduce computational complexity, making them more suitable for MOO tasks. For example, Vandani et al. (2019) [14] employed the NSGA-II and multi-objective particle swarm optimization to solve a bi-objective MIP model, achieving a balanced assignment of trucks and cranes in container terminals. Fard et al. (2019) [26] applied the multi-objective imperialist competitive algorithm and multi-objective grey wolf optimizer to solve a bi-objective MIP model, optimizing the sequencing and assignment of inbound and outbound trucks at a multi-door cross-dock. Zhang et al. (2020) [15] proposed a decomposition-based constrained dominance NSGA-II to solve a three-objective ILP model, which reassigned optimal paths for each truck and generated a Gantt chart for vehicle scheduling, effectively supporting visual management of production. Gao et al. (2025) [20] proposed a probability distribution-based MOEA/D to solve a three-objective programming model, achieving a balance among power output, land usage, and costs.

In recent years, heuristic MOO algorithms have been widely applied to engineering optimization problems of higher dimensions, providing valuable insights for FTA optimization. For example, Hakanen et al. (2016) [27] applied the MOEA to a multiple-disk clutch brake design problem with five objectives. In building optimization design, Razmi et al. (2022) [28] used the NSGA-III to solve a five-objective model constructed by Principal Component Analysis and Artificial Neural Network (PCA-ANN), achieving a balanced optimization of energy efficiency, daylight, and thermal comfort. In fog computing task scheduling, Altin et al. (2024) [21] employed the SPEA2 to solve a five-objective model. The evaluation results show that integrated expansion plays a significant role in reducing data transmission costs.

3. Problem definition

Assuming that b represents the number of haulage roads and considering the operating characteristics of typical open-pit mines, two FTA modes are defined.

Definition 1. Single independent operated haulage road ($b = 1$)

Based on the production plan, a fixed type and number of shovels are deployed at the loading point for mining (or stripping) operations. Following the industry practice of "big trucks with big shovels, small trucks with small shovels," a transport grouping, composed of either homogeneous or heterogeneous fleets, is assigned to the haulage road to perform transport tasks. After completing loading and loaded haulage, the trucks unload at the unloading point and return empty to the loading point to begin the next cycle, as shown in the "production scenario" in Fig. 1. To prevent overtaking or traffic congestion, it is assumed that all trucks in the transport grouping haul at the same loaded (or empty)

speed.

Definition 2. Multiple haulage roads with overlapping segments ($b \geq 2, \forall b \in Z^+$)

Each haulage road has an independent transport grouping, and the loading points (or unloading points) of different haulage roads may be the same. The operation of each transport grouping follows the same process as described in "Definition 1" and is independent of the others. To ensure the efficient and safe passage of trucks on overlapping segments, it is assumed that all transport groupings haul at the same loaded (or empty) speed in these segments.

Therefore, the FTA optimization problem can be considered as the synergistic optimization of various truck-related parameters in two modes, aiming to maximize equipment utilization while minimizing energy consumption. Its mathematical expression is shown in Eq. (1).

$$Optimize[\theta_1, \ldots, \theta_b] \rightarrow \left\{ \left[\max F_{1,1}(\theta_1), \min_{\forall b \in Z^+} F_{2,1}(\theta_1) \right], \ldots, \left[\max F_{1,b}(\theta_b), \min F_{2,b}(\theta_b) \right] \right\} \tag{1}$$

θ_i represents the set of truck parameter types related to the i-th transport grouping, $i \in (1, b)$, $F_{1,i}$ and $F_{2,i}$ represent the equipment utilization and energy consumption in the i-th transport grouping, respectively.

When $b = 1$, Eq. (1) represents a three-objective optimization problem. Let $s_{truck,r}$ represent the loaded (or empty) hauling speed of the r-th truck, and n be the number of trucks. The loaded (or empty) hauling speed for all trucks ($\forall r \in \{1, \ldots, n\}$) is assumed to be the same within the transport grouping, i.e., $s_{truck,1} = s_{truck,2} = \ldots = s_{truck,n}$, $s_{truck,r} \in \theta$. When $b \geq 2$, Eq. (1) expands to a $3 \times b$-objective optimization problem, where $s_{formation,i}$ represents the loaded (or empty) hauling speed of the i-th transport grouping. The loaded (or empty) hauling speed for all transport groupings ($\forall i \in \{1, \ldots, b\}$) is assumed to be the same, i.e., $s_{grouping,1} = s_{grouping,2} = \ldots = s_{grouping,b}$, $s_{grouping,i} \in \theta_i$.

4. Methodology

4.1. Synergistic optimization

4.1.1. Approach introduction

This subsection considers the impact of the three factors of haul distance, transport grouping, and fleet speed on two nonlinear dependent variables of equipment utilization and energy consumption, and establishes a connection between them, as shown in Fig. 2. On this basis, a unique "haul distance–transport grouping–fleet speed" synergistic optimization approach is proposed to improve the FTA strategy in open-pit mines. The approach aims to create balanced transport groupings and optimal fleet speeds for different haul distances by adjusting the types and numbers of trucks, as well as the loaded hauling speeds (LHSs) and empty hauling speeds (EHSs), thereby minimizing the SIT, TWT, and TFC.

4.1.2. Decision variables and parameters

The decision variables and parameters affecting the SIT, TWT, and TFC are defined in this subsection. The decision variables related to transport grouping include the type and number of trucks, with the truck identifier as the parameter.

The round-trip process of a truck is shown in Fig. 3. During the haul from loading point A to unloading point B, the parameters involved in the "Loading" phase include the payload and loading time. The decision variable in the "Loaded" phase is the LHS, with parameters including the loaded haul distance and lifting height (elevation difference from A to B). The parameter for the "Wait to unload" phase is the waiting time, and the parameter for the "unloading" phase is the unloading time. During the return haul from unloading point B to loading point A, the decision variable in the "Empty" phase is the EHS, with parameters including the empty haul distance and lifting height (elevation difference from B to A). The parameter for the "Wait to load" phase is the waiting time.

Parameters related to production demand include the mining (or stripping) volume at the loading point and the duration of a production shift. Parameters related to haul routes include the number of routes with the same overlapping segments.

Notably, this study adopted the method in reference [23, 29, 30] to obtain the speed value, specifically by using the statistical haul distance and haulage time, and calculated the average speed through Eq. (2) to guide onsite work. Among them, $D_{haul\ distance}$ represents the haul distance, $T_{haulage\ time}$ represents the haulage time, and $S_{hauling\ speed}$ represents the hauling speed. This decision stems from the following considerations: time and distance data collected by sensors are more readily available and accurate than speed data. Therefore, the decision variables EHS and LHS can be replaced by loaded hauling time and empty hauling time, respectively, through Eq. (2). Table 1 summarizes these decision variables and parameters.

$$S_{hauling\ speed} = \frac{D_{haul\ distance}}{T_{haulage\ time}} \tag{2}$$

4.2. Hybrid modeling

4.2.1. Objective functions

4.2.1.1. Time models. This subsection establishes two time-related objective functions to represent the SIT and TWT, inspired by the objective functions (Eqs. (5) and (6)) outlined by Afrapoli et al. (2019) [22]. The two differ in application scenarios and expressions:

(1) Afrapoli et al. (2019) [22] studied dynamic scheduling with a

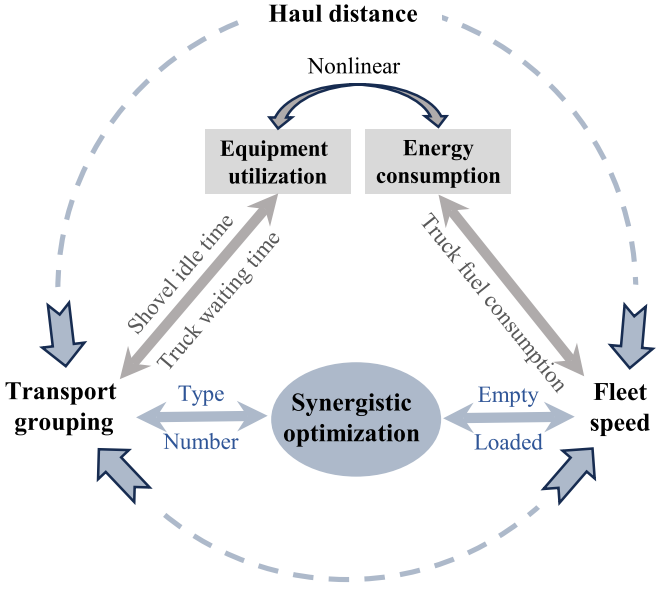

Fig. 2. Connections between the factors of haul distance, transport grouping, and fleet speed and the nonlinear dependent variables of equipment utilization and energy consumption.

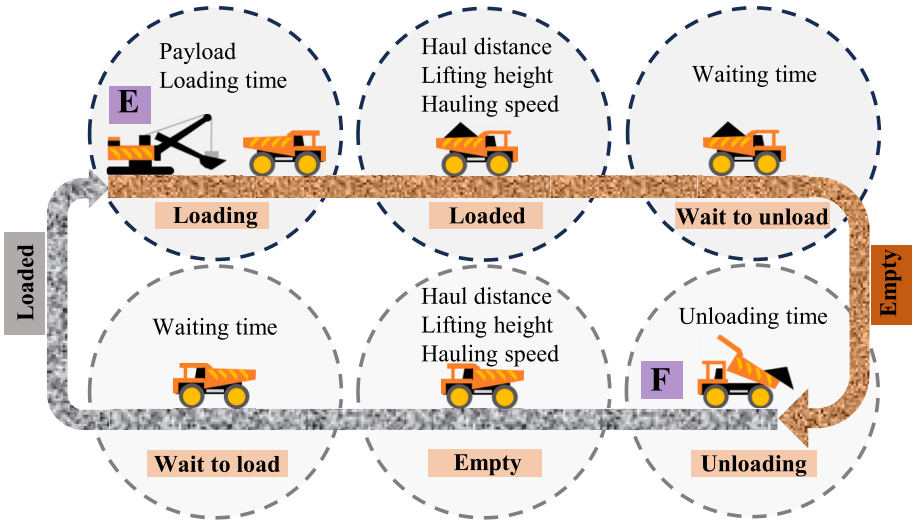

Fig. 3. Decision variables and parameters involved in the round trip of a truck.

Table 1
Summary of decision variables and parameters.

	Transport grouping	Symbol
Decision variable	Type of trucks	The set L_{type}, $L_{type} = \{A, B, C\ldots\}$, A, B, C … represent different types of trucks
	Number of trucks	$n, \forall n \in Z^+$
Parameter	Truck identifier	$r, \forall r \in Z^+, \forall r \in \{1, \ldots, n\}$
	The round trip of a truck	**Symbol**
Decision variable	Loaded haulage time (s)	t_1
	Empty haulage time (s)	t_2
Parameter	Waiting time (s)	t_3
	Loading time (s)	t_4
	Unloading time (s)	t_5
	Loaded haul distance (km)	d_1
	Empty haul distance (km)	d_2
	Payload (ton)	p
	Lifting height (m)	h
	Production demand	**Symbol**
Parameter	The mining (or stripping) volume at loading points (ton)	Q
	The duration of a production shift (h)	T_s
	Haul routes	**Symbol**
Parameter	Number of routes with the same overlapping segments	$b, b \in Z^+$

focus on the dynamic variation in the type and number of trucks within a transport grouping, adhering to a "n trucks for m shovels" strategy. Conversely, this paper examines the FTA approach, ensuring a constant type and number of trucks within a transport grouping and adopting a "n trucks for 1 shovel" strategy.

(2) Afrapoli et al. (2019) [22] defined equipment nonworking time as the cumulative duration of equipment idle time during a truck's round trip. This study categorizes equipment into working and nonworking states, with the SIT being the time of a production shift minus the shovel working time, and the TWT being the time of a production shift minus the truck working time. This approach simplifies the objective function through a decrement method rather than the accumulation method used by Afrapoli et al. (2019) [22]. It also introduces decision variables related to equipment working status for future FTA optimization.

The construction process of these two time-related objective functions is as follows:

a) The number of round trips (N_{cycle}) for the transport grouping can be calculated via Eq. (3), which represents the ratio of the mining (or stripping) volume (Q) at the loading point to the sum of the single haul capacities of the n trucks ($\sum_{r=1}^{n} p_r$) in the transport grouping. Here, $\lceil \rceil$ represents the ceiling function, which rounds up any value with a decimal part to the next integer.

$$N_{cycle} = \lceil \frac{Q}{\sum_{r=1}^{n} p_r} \rceil \qquad \forall n \in Z^+, \forall r \in \{1, \ldots, n\} \tag{3}$$

b) If the duration of a production shift is T_s, the SIT, denoted as f_1, at that loading point can be expressed by Eq. (4). This expression represents the duration of a production shift minus the cumulative loading time of all the trucks within the transport grouping (where the cumulative loading time of the trucks corresponds to the effective working time of the shovels).

$$f_1 = T_s - \sum_{r=1}^{n} t_{4,r} N_{cycle} \qquad \forall n \in Z^+, \forall r \in \{1, \ldots, n\} \tag{4}$$

c) The TWT within the transport grouping, denoted as f_2, can be calculated via Eq. (5). In Eq. (5), f_2 equals the duration of a production shift minus the summation of working times for each truck within the transport grouping, which encompasses loaded haulage, empty haulage, loading, and unloading stages.

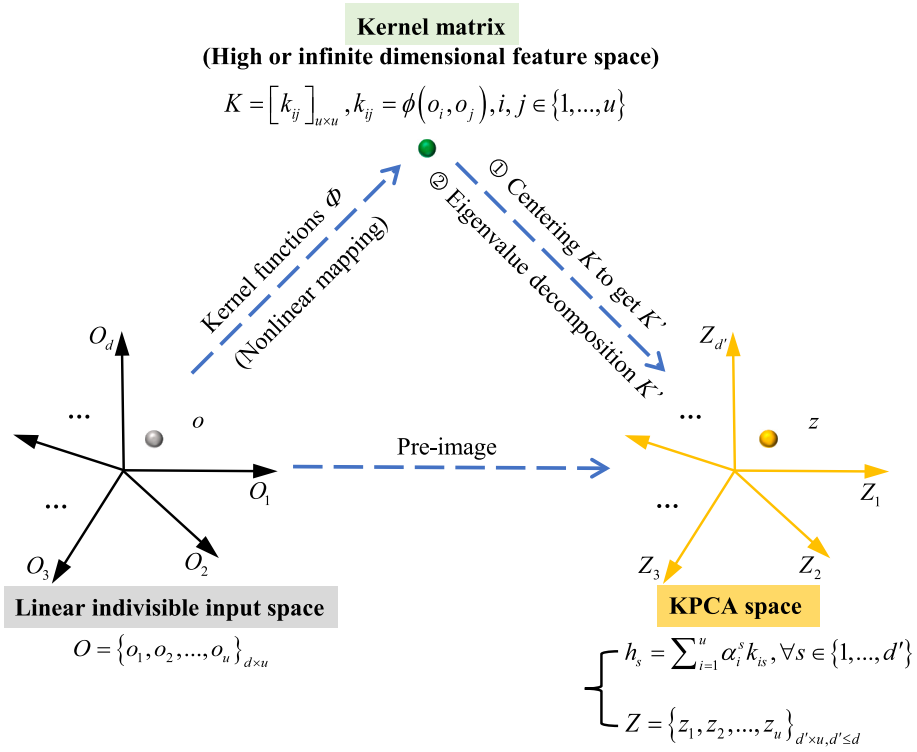

Fig. 4. Schematic representation of the KPCA principle.

$$f_2=\sum_{r=1}^{n}\left[T_s-\left(t_{1,r}+t_{2,r}+t_{4,r}+t_{5,r}\right)N_{cycle}\right] \qquad \forall n\in Z^+, \forall r\in\{1,\ldots,n\} \tag{5}$$

4.2.1.2. FC model. A data-driven model is established in this subsection based on Kernel Principal Component Analysis and Artificial Neural Network (KPCA-ANN) [31,32] to represent the FC. In contrast to knowledge- and mechanism-based models, data-driven models excel in scalability and real-time performance [33,34]. Especially when a large amount of data needs to be processed or the internal mechanism of the system is unclear, the accuracy of data-driven models tends to be superior [35].

In the KPCA-ANN model, KPCA [36,37] is used for nonlinear dimensionality reduction, as shown in Fig. 4. The specific operation steps are as follows:

a) In the linearly indivisible input space O ($O=\{o_1,o_2,\ldots,o_u\}_{d\times u}$), the kernel function ($\phi$) is used to calculate the kernel matrix K ($K=\left[k_{ij}\right]_{u\times u}, k_{ij}=\phi(o_i,o_j), i,j\in\{1,\ldots,u\}$), where d is the sample dimension, and u is the number of samples.

b) The centering of the kernel matrix K yields K', which is then subjected to Eigenvalue Decomposition (EVD). The centering process can be found in reference [36, 38], and the EVD constitutes the PCA dimensionality reduction calculation.

c) The eigenvectors $[\alpha_1,\alpha_2,\ldots,\alpha_{d'}]$ corresponding to the largest d' eigenvalues are selected.

d) On the basis of the feature vectors, the coordinates after dimensionality reduction are calculated to form the KPCA space Z. Here, the s-th-dimensional coordinate is h_s ($h_s=\sum_{i=1}^{u}\alpha_i^s k_{is}, \forall s\in\{1,\ldots,d'\}$), and the space is Z ($Z=\{z_1,z_2,\ldots,z_u\}_{d'\times u, d'\le d}$).

The data after dimensionality reduction by KPCA will be used as input for the ANN [23], with FC as the output. According to Table 2 in reference [23], there is a strong nonlinear relationship between the variables that affect TFC. Therefore, compared with ANN [23] and PCA-ANN [28], using KPCA-ANN to model FC offers theoretical advantages.

Determining the appropriate input variables is a prerequisite for establishing the FC model. Siami et al. (2015) [23] determined the input variables of the ANN model to be payload, loaded haulage time, empty haulage time, loading time, idle time while loaded, and idle time while empty. On this basis, Wang et al. (2021) [29] incorporated mining area geological factors, such as road length and lifting height, into the

Table 2
Details of each type of shovel and truck.

	Type	Number	Weight (ton)	Monthly mining (or stripping) volume ($\times10^4$ m^3)	Maximum load (ton)	Fuel tank capacity (L)	Maximum speed (km/h)
Shovel	WK20	3	731	20	–	–	–
	WK395	2	937	45	–	–	–
	WK35	2	1062	65	–	–	–
	WK55	7	1458	85	–	–	–
Truck	630E	6	108	–	154	3217	55
	730E	5	138	–	185	3217	55
	SF33900	19	164	–	220	4520	65
	830E	20	166	–	220	4542	65
	930E	18	210	–	290	4542	65
	XDE300	10	210	–	300	4542	60
	HT3363	22	237	–	360	5400	64

eXtreme Gradient Boosting model, which reduced the Mean Absolute Percentage Error (MAPE) to 8.78 %. Drawing upon the studies mentioned above and the information from Fig. 3 and Table 1, the final input variables for the KPCA-ANN model are determined, comprising loaded haulage time (t_1), empty haulage time (t_2), loading time (t_4), unloading time (t_5), loaded haul distance (d_1), empty haul distance (d_2), payload (p), and lifting height (h).

The architecture of the KPCA-ANN model is shown in Fig. 5. The input layer consists of the eight parameters mentioned in the previous paragraph. In the KPCA layer, principal components are extracted from the input variables. In the ANN hidden layer, recursion drives forward propagation, while gradient descent enables backpropagation. The output value is then obtained through the activation function. For detailed parameter settings, refer to Section 5.4.2. The functional relationship established by the KPCA-ANN model is denoted as δ, and the TFC (f_3) can be expressed by Eq. (6).

$$f_3 = \sum_{r=1}^{n} \delta(t_{1,r}, t_{2,r}, t_{4,r}, t_{5,r}, d_{1,r}, d_{2,r}, p_r, h_r) N_{cycle} \qquad \forall n \in Z^+, \forall r \in \{1, \ldots, n\} \tag{6}$$

4.2.2. Constraint conditions

The constraint conditions are given by Eqs. (7) to (17).

$$S_{min} \le d_{1,r}/t_{1,r}, d_{2,r}/t_{2,r} \le S_{max} \qquad \forall n \in Z^+, \forall r \in \{1, \ldots, n\} \tag{7}$$

$$d_{1,r=1,i}/t_{1,r=1,i} = d_{1,r=2,i}/t_{1,r=2,i} = \ldots = d_{1,r=n,i}/t_{1,r=n,i} \forall n \in Z^+, \forall r \in \{1, \ldots, n\}, \forall b \in Z^+, \forall i \in \{1, \ldots, b\} \tag{8}$$

$$d_{2,r=1,i}/t_{2,r=1,i} = d_{2,r=2,i}/t_{2,r=2,i} = \ldots = d_{2,r=n,i}/t_{2,r=n,i} \quad \forall n \in Z^+, \forall r \in \{1, \ldots, n\}, \forall b \in Z^+, \forall i \in \{1, \ldots, b\} \tag{9}$$

$$d_{1,r,1}/t_{1,r,1} = d_{1,r,2}/t_{1,r,2} = \ldots = d_{1,r,b}/t_{1,r,b} \qquad \forall n \in Z^+, \forall r \in \{1, \ldots, n\}, \forall b \in Z^+ \tag{10}$$

$$d_{2,r,1}/t_{2,r,1} = d_{2,r,2}/t_{2,r,2} = \ldots = d_{2,r,b}/t_{2,r,b} \qquad \forall n \in Z^+, \forall r \in \{1, \ldots, n\}, \forall b \in Z^+ \tag{11}$$

$$(n_i - 1) \times 50 \le D_i \qquad \forall n \in Z^+, \forall b \in Z^+, \forall i \in \{1, \ldots, b\} \tag{12}$$

$$n_{each\ type} \le N_{total} \qquad \forall n_{each\ type} \in Z^+ \tag{13}$$

$$L_{type,after\ mutation} \subseteq L_{type} \tag{14}$$

$$Q_u \le Q \le Q_l \tag{15}$$

$$p_r \le P_r \qquad \forall n \in Z^+, \forall r \in \{1, \ldots, n\} \tag{16}$$

$$f_1, f_2 \ge 0 \tag{17}$$

Eq. (7) ensures that the LHS (d_1/t_1) and EHS (d_2/t_2) of each truck within the transport grouping stay within the speed limits of the mainline road. Specifically, the speeds must not exceed the maximum limit (S_{max}), set at 40 km/h, and must not fall below the minimum limit (S_{min}), set at 15 km/h. Eqs. (2) and (7) indicate that when the maximum speed limit (S_{max}) and the minimum speed limit (S_{min}) are known, once the loaded haul distance (d_1) and the empty haul distance (d_2) are known, the value ranges of the loaded haulage time (t_1) and empty haulage time (t_2) can be easily determined. Eqs. (8) and (9) ensure that the LHS (d_1/t_1) and EHS (d_2/t_2) of each truck in the same transport grouping remain consistent on any given haul route, thereby guaranteeing smooth operation within each transport grouping and preventing overtaking or congestion. Eqs. (10) and (11) ensure that each transport grouping hauls at the same LHS (d_1/t_1) and EHS (d_2/t_2) on routes with the same overlapping segments, thereby ensuring smooth operation between the transport groupings and preventing overtaking or congestion.

Eq. (12) ensures that the length of the i-th transport grouping (n_i) does not exceed the length of the haul route (D_i). The minimum safe distance between dump trucks is typically 50 m. Eq. (12) indicates that when the length of the haul route (D_i) is known, it becomes straightforward to ascertain the maximum number of trucks that can fit into the transport grouping. Eq. (13) ensures that the number of trucks of each type ($n_{each\ type}$) in the generated transport groupings does not exceed the number of corresponding truck types (N_{total}) owned by the mining company. Eq. (14) ensures that in the transport grouping represented by letter encoding, when the letter at any position is mutated, the resulting letter still belongs to the original letter set. Eq. (15) ensures that the mining (or stripping) volume (Q) does not exceed the maximum production capacity (Q_l) of shovels at the loading point while meeting the production demand (Q_u) of the crushing station (the parameter (Q_u) is not required if it is bound for the dumping site). Eq. (16) ensures that the payload of each truck (p_r) does not exceed its maximum payload (P_r). Eq. (17) ensures that the time value of the solution is non-negative.

4.3. Moo

4.3.1. Definition of the MOO problem

The FTA optimization problem can be further expressed as follows: Under the two FTA modes defined in Section 3, the SIT (f_1), TWT (f_2), and TFC (f_3) of each transport grouping are minimized by synergistically optimizing the types (L_{type}) and number (n) of trucks, as well as the loaded haulage time (t_1) and empty haulage time (t_2) of the fleet. Its mathematical form is further expressed from Eq. (1) as Eq. (18).

$$Optimize\left\{\left(L_{type,1}, n_1, t_{1,1}, t_{2,1}\right), \ldots, \left(L_{type,b}, n_b, t_{1,b}, t_{2,b}\right)\right\} \rightarrow Min\left\{\left[f_{1,1}, f_{2,1}, f_{3,1}\right], \ldots, \left[f_{1,b}, f_{2,b}, f_{3,b}\right]\right\} \quad \forall n \in Z^+, \forall b \in Z^+ \tag{18}$$

Consistent with the definition in Section 3, when $b = 1$, Eq. (18) describes a three-objective optimization problem, assuming that the loaded (or empty) haulage time is the same for all trucks within the transport grouping. When $b \ge 2$, Eq. (18) expands into a $3 \times b$-objective optimization problem, with the assumption that the loaded (or empty) haulage time is the same across all transport groupings.

4.3.2. ARP-NSGA-III with hybrid encoding

The NSGA-III algorithm [16,17] directs the population to explore the Pareto frontier using a set of uniformly distributed reference points. Jain et al. [39] discovered that some reference points can never be linked to well-distributed Pareto optimal solutions (POSs), rendering them useless, whereas others can correspond to multiple optimal solutions. The presence of useless reference points may reduce the effectiveness of the NSGA-III in screening diverse solutions during population search.

To address the issue of redundant reference points, this subsection introduces a density-based probabilistic model for adaptive generation

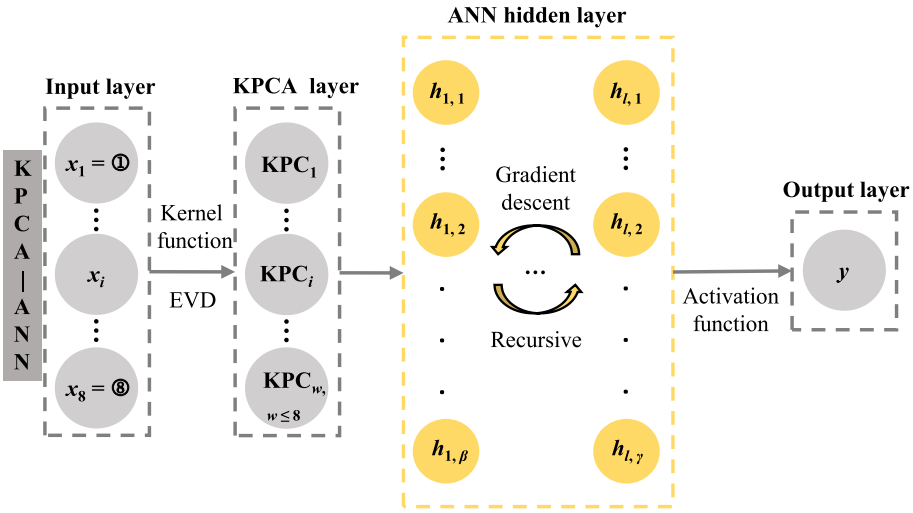

Fig. 5. Architecture design of the KPCA-ANN.

of reference points [40]. Algorithm 1 presents the fundamental framework of the ARP-NSGA-III. In the first line of Algorithm 1, the ARP-NSGA-III employs a hybrid encoding approach, using letter encoding to represent the type (L_{type}) and number (n) of trucks and real-valued encoding for $X_{round\ trip}$. This hybrid encoding approach alters the generation, crossbreeding, and mutation processes of individuals, which are discussed in detail in Section 5.4.3. The subsequent focus will be on the steps involved in generating adaptive reference points:

a) Following Das and Dennis's systematic approach [41], the entire objective space is evenly partitioned into several sub-locations ($\widehat{w}_1, \widehat{w}_2, \widehat{w}_3, \ldots, \widehat{w}_k \in W$) (Algorithm 1, Line 4).

b) The number of solutions associated with each sub-location ($\widehat{w}$) is documented in the archive ($E(w)$). A solution ($\widehat{s}$) can be defined as associated with $\widehat{w}$ when it has the minimum perpendicular distance to $\widehat{w}$ (Algorithm 1, Line 13). $d_{s,w}{}^{\perp}(s,w)$ represents the vertical distance between s and w.

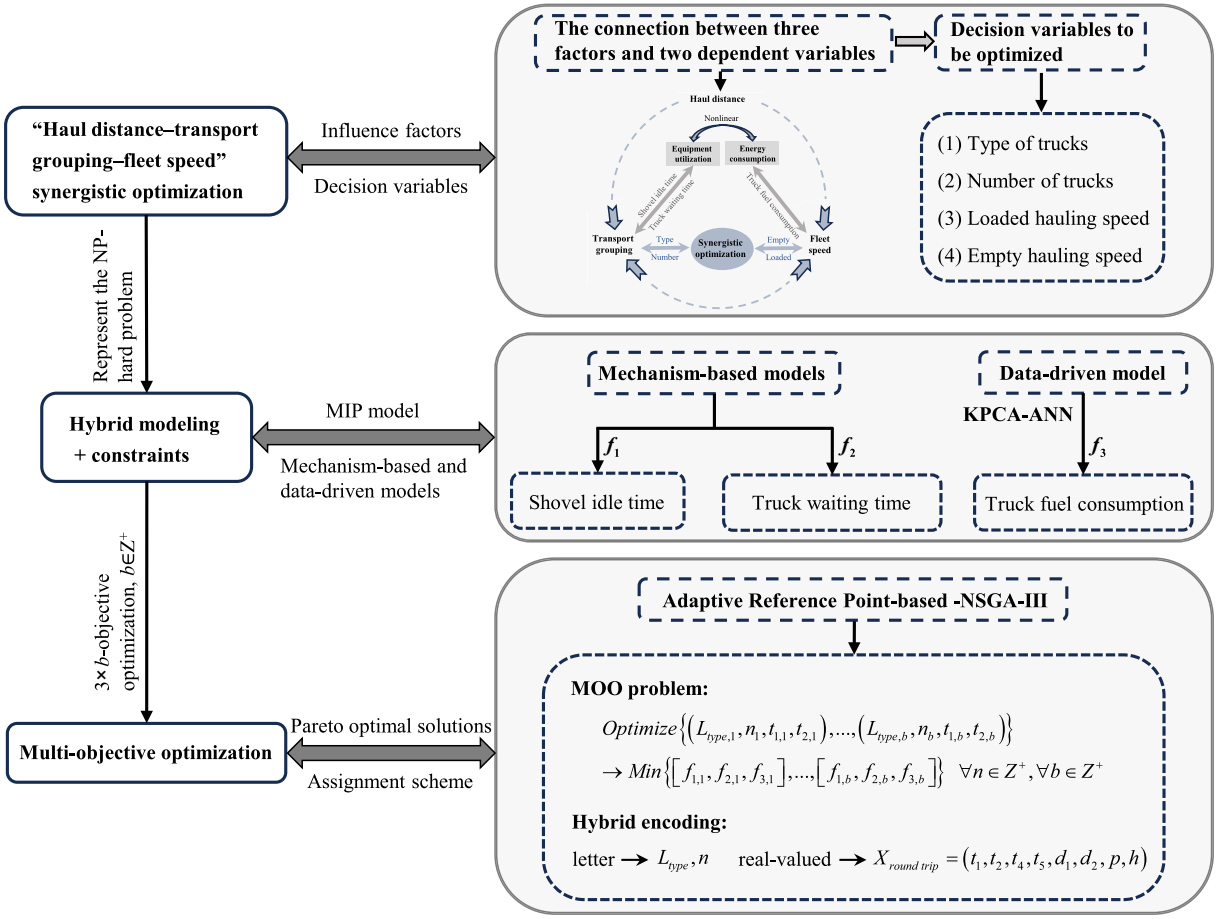

Fig. 6. Development of the integrated architecture.

Algorithm 1. The framework of the ARP-NSGA-III.

Algorithm 1. The framework of the ARP-NSGA-III.

Input:

$Optimize\left\{\left(L_{type,1},n_1,t_{1,1},t_{2,1}\right),...,\left(L_{type,b},n_b,t_{1,b},t_{2,b}\right)\right\}\rightarrow Min\left\{\left[f_{1,1},f_{2,1},f_{3,1}\right],...,\left[f_{1,b},f_{2,b},f_{3,b}\right]\right\}$

Set the transport grouping to $\left[A_{r=1},...,A_{r=i},B_{r=i+1},...,B_{r=j},C_{r=j+1}...\right]_n$, *where* A, B, C... *denote the type of trucks, n denotes the number of trucks, and r denotes the identifier of trucks*

$X_{round\ trip}=\left(t_1,t_2,t_4,t_5,d_1,d_2,p,h\right)$

P = *parent population*
N = *number of individuals*
T = *number of iterations*
H = *initial number and initial position of reference points on the hyperplane* (the "Das and Dennis" [41] approach is chosen for H)

Output: A set of nondominated solutions

Letter encoding generates L_{type} and n, and real-valued encoding generates $X_{round\ trip}$;
Initialize the population P_0;
Evaluate the population P_0;
Generate the W that partitions the objective space into sub-locations ($\hat{w}$);
Set $g\leftarrow 0$;
while $t<t_{\max}$ **do**
    Generate the offspring population Q_t using crossover, mutation, and reproduction;
    foreach $Q^*\in Q_t$ **do** Evaluate Q^*;
    $R_t\leftarrow P_t\bigcup Q_t$;
    Apply nondominated sorting on R_t and find ($F_1^*,F_2^*...$);
    Normalize the population members: $S_t'=ObjectiveNormalization(S_t)$;
    foreach $w\in W$ **do**
        Identify members of S_t' associated with w;
        $Assign\left(E\left(\hat{w}\right),D\left(\hat{w}\right)\right)=Associatew\left(S_t',W\right)$; (Density model, as detailed in Algorithm 2 in reference [40])
    End
    $Assign\ Z_g^*=Generate\left(E\left(\hat{w}\right),D\left(\hat{w}\right),S_t',W\right)$; (The adaptive reference point generation process, as detailed in Algorithm 3 in reference [40])
    Construct the new population P_{t+1} by the NSGA-III association operation and niche-preservation operation;
    $t\leftarrow t+1$;
End
return The nondominated individuals $p^*\subseteq p_{tmax}$;

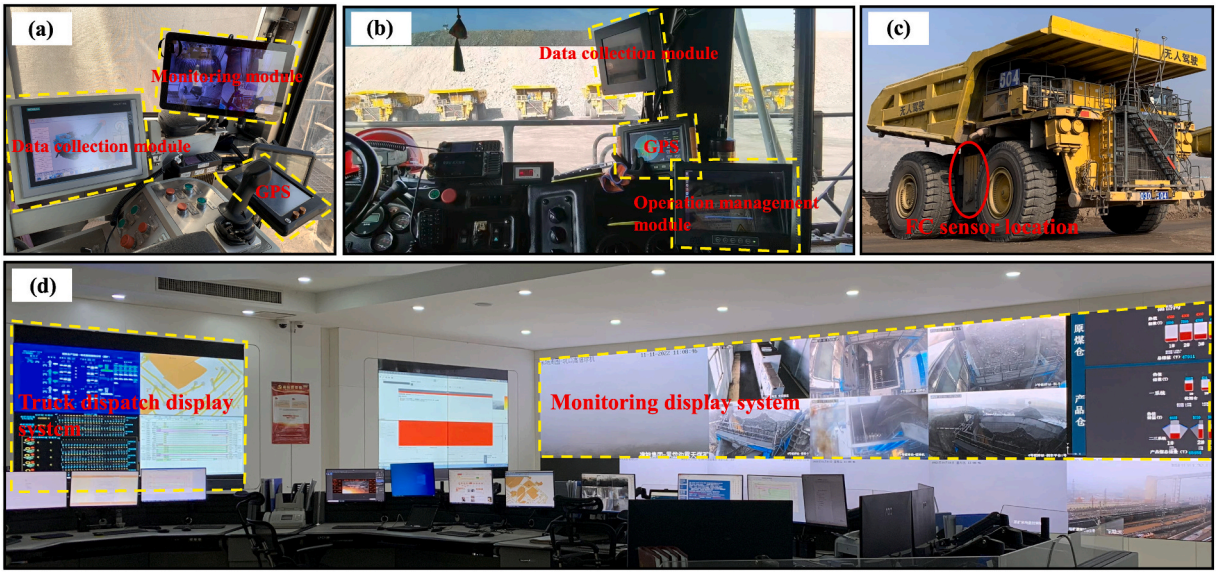

Fig. 7. Data collection devices for time and FC: (a) inside the shovel, (b) inside the truck, (c) FC sensor location, (d) tuck dispatch center.

Parameter	Value
Testing range (m)	0.05 ~ 5
Accuracy (grade)	0.1
Pressure-bearing range (MPa)	-0.1 ~ 32
Temperature range (℃)	-50 ~ 250

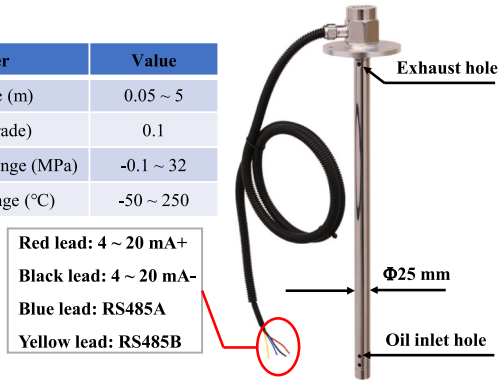

Fig. 8. Appearance and main parameters of the FC sensor.

c) It is necessary to calculate the density probability of solutions in each sub-location ($\widehat{w}$), which is defined by Eq. (19):

$$P_D(D|\widehat{w} \in W) = \frac{\left\| \sum \left(\text{argmin}_{s \in S} d_{s,w}{}^{\perp}(s,w) \right) \right\|}{\|S\|} \tag{19}$$

where $\left\| \sum \left(\text{argmin}_{s \in S} d_{s,w}{}^{\perp}(s,w) \right) \right\|$ represents the number of solutions associated with each sub-location ($\widehat{w}$) and $\|S\|$ is the number of non-dominated solutions obtained thus far. The normalized individuals ($\overline{S}_g$) are associated with sub-location (W), and the number and indices of solutions associated with each $\widehat{w}$ are recorded in the archive ($E(w)$). The density of solutions at each $\widehat{w}$ is then computed (Algorithm 1, Line 14), with the density calculation Eq. (20) as follows:

$$D(\widehat{w}) = \|P_D(D|\widehat{w})\| \times N_{number\ of\ reference\ points} \tag{20}$$

The detailed implementation process of Line 14 can be found in Algorithm 2 presented in reference [40].

d) An adaptive process for generating reference points is described in Line 16 of Algorithm 1. The number of reference points at the sub-location ($\widehat{w}$) is defined by Eq. (21):

$$set\ nref = \|D(\widehat{w})\| \tag{21}$$

Clearly, the greater the density is, the more reference points emerge at the sub-location ($\widehat{w}$). Once the number of reference points at the sub-location ($\widehat{w}$) is determined, the positions of these reference points need to be defined. The sub-location ($\widehat{w}$) is set as the position of the first reference point. If this position coincides with a vertex, the vertex method from reference [40] is employed. If this position is not a vertex, the midpoint method from reference [40] is applied. The detailed implementation of Line 16 can be found in Algorithms 3 and 4, as shown in reference [40].

4.4. Integrated architecture

This study develops an integrated architecture to optimize FTA strategies in open-pit mines, consisting of three main components, as illustrated in Fig. 6.

The first part introduces the "haul distance–transport grouping–fleet speed" synergistic optimization approach. By adjusting truck types and numbers, along with LHSs and EHSs, the approach creates balanced transport groupings and optimal fleet speeds for different haul distances, minimizing the SIT, TWT, and TFC.

The second part presents a hybrid modeling method that integrates both mechanism-based and data-driven models into a constrained MIP model to represent the NP-hard FTA optimization problem. The mechanism-based model represents the SIT and TWT, while the KPCA-ANN-based data-driven model represents the TFC.

The third part defines a three-objective optimization problem for the single independently haul route ($b = 1$), and a $3 \times b$-objective optimization problem for the multiple haul routes ($b \geq 2$) with overlapping road segments. Using ARP-NSGA-III with letter and real-valued encoding, an approximate POS set is generated.

5. Case description

This section uses the Heidaigou open-pit coal mine in Ordos city, Inner Mongolia, China, as a case study to construct the integrated architecture shown in Fig. 6. The case mine employs intermittent mining methods, and its annual raw coal production has exceeded 30 million tons. As one of the most technologically advanced open-pit mines in China, the mine has established a comprehensive system that includes a surface control system, a vehicle-to-infrastructure wireless communication system, and an onboard control system. This system facilitates information transmission among the dispatch center, shovels, and trucks, providing the data foundation for FTA optimization.

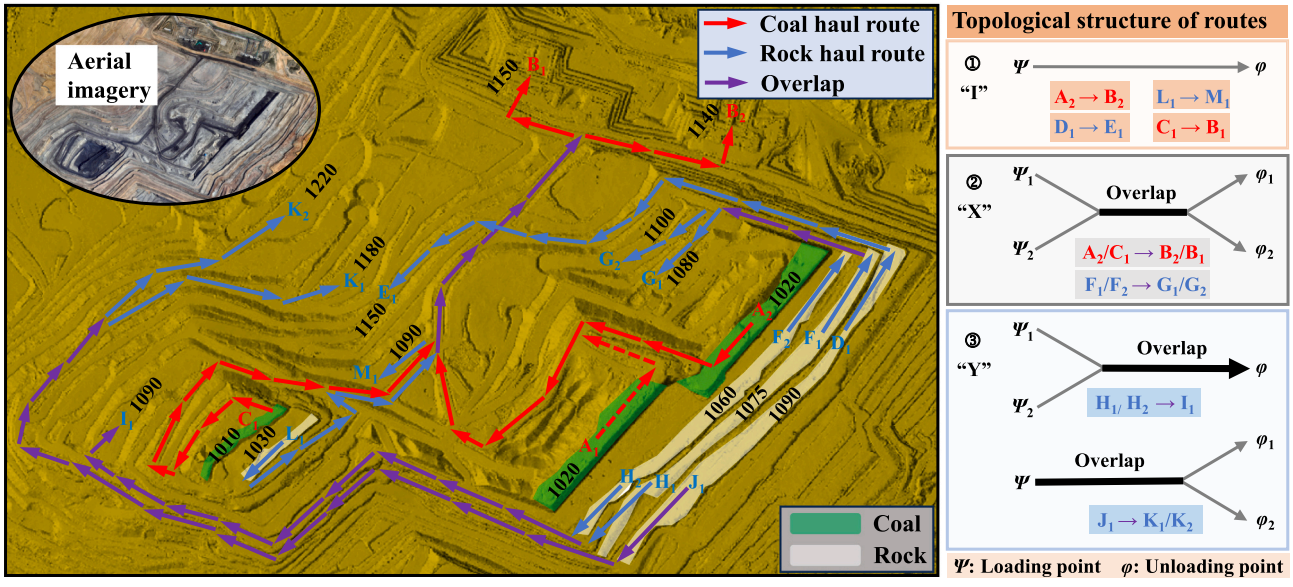

Fig. 9. On-site information of the case mine (left) and topological structures of the route network (right).

Table 3
Basic information about each route.

Route purpose	Route	Shovel	Truck type	Haul distance (km)	Lifting height (m)	Lifting height per kilometer (m/km)
		Type and number				
Coal haul	$A_2 \rightarrow B_2$	WK35 (2)	630E 730E SF33900	5.7–6.7	120	21.1–17.9
	$C_1 \rightarrow B_1$	WK20 (3)	630E 730E SF33900	5.8–6.4	140	24.1–21.9
	$D_1 \rightarrow E_1$	WK55 (1)	830E 930E XDE300	3.8–5.3	60	15.8–11.3
	$F_1 \rightarrow G_1$	WK55 (1)	830E 930E XDE300	1.6–2.8	5	3.1–1.8
	$F_2 \rightarrow G_2$	WK55 (1)	830E 930E XDE300	1.6–2.7	40	25.0–14.8
Rock haul	$H_1 \rightarrow I_1$	WK55 (1)	830E 930E HT3363	4.3–6.2	15	3.5–2.4
	$H_2 \rightarrow I_1$	WK55 (1)	830E 930E HT3363	4.1–6.0	30	7.3–5.0
	$J_1 \rightarrow K_1$	WK55 (1)	830E 930E HT3363	7.0–8.8	90	12.9–10.2
	$J_1 \rightarrow K_2$	WK55 (1)	830E 930E HT3363	7.0–8.7	130	18.6–14.9
	$L_1 \rightarrow M_1$	WK395 (2)	830E 930E HT3363	2.8–3.1	45	16.1–14.5

5.1. Experimental equipment

5.1.1. Excavation and transportation

The excavation equipment includes four types of shovels, while the transportation equipment includes seven types of trucks. These shovels and trucks work together to complete loading, transport, and unloading operations. Table 2 summarizes the equipment details.

5.1.2. Data collection

Time and FC data are collected via sensors. Fig. 7(a) shows the three modules of Global Positioning System (GPS), monitoring, and time data collection for the shovel, while Fig. 7(b) shows the corresponding modules for the truck, including GPS, operation management, and time data collection. These modules work together to obtain the position and status information of both the shovel and the truck in real-time.

As shown in Fig. 8, the industrial-grade FC sensor (model CR-6063), designed to withstand high temperatures and pressures, is used for FC data collection. Fig. 7(c) shows the sensor installation location. The operational principle of the CR-6063 involves a coaxial container as the sensing component, where the introduction of oil causes a change in capacitance between the sensor housing and the induction electrode. The circuit converts this change, applies linear correction and temperature compensation, and outputs a 4–20 mA standard signal to the display instrument. On-site demonstrations have shown that this equipment meets the requirements for a wide range, high precision, and strong anti-interference capabilities, particularly in unstructured roads of open-pit mines. Fig. 7(d) represents the truck dispatch center, which includes the truck dispatch display system and monitoring display system, providing an intuitive visualization of the data collected by the sensors.

5.2. Experimental site

Fig. 9 shows the on-site information at the case mine site (left) and the topological structure of the route network (right). On the left, the haul routes, elevations, loading points, and unloading points are marked. The red arrows represent coal haul routes, the blue arrows represent rock haul routes, the purple arrows represent overlapping segments, the green areas denote coal mining benches, and the white areas denote rock stripping benches. The mine has two main coal mining benches: A_1 (or A_2) and C_1. The A_1 and A_2 benches utilize throw blasting, and when one is in operation, the other is stopped. Among these routes, A_2, L_1, D_1, C_1, F_1, F_2, J_1, H_1, and H_2 are loading points, whereas B_2, M_1, E_1, B_1, G_1, G_2, K_1, K_2, and I_1 are unloading points.

The right panel of Fig. 9 shows that the topological structure of the route network can be categorized into three types. The first type is the

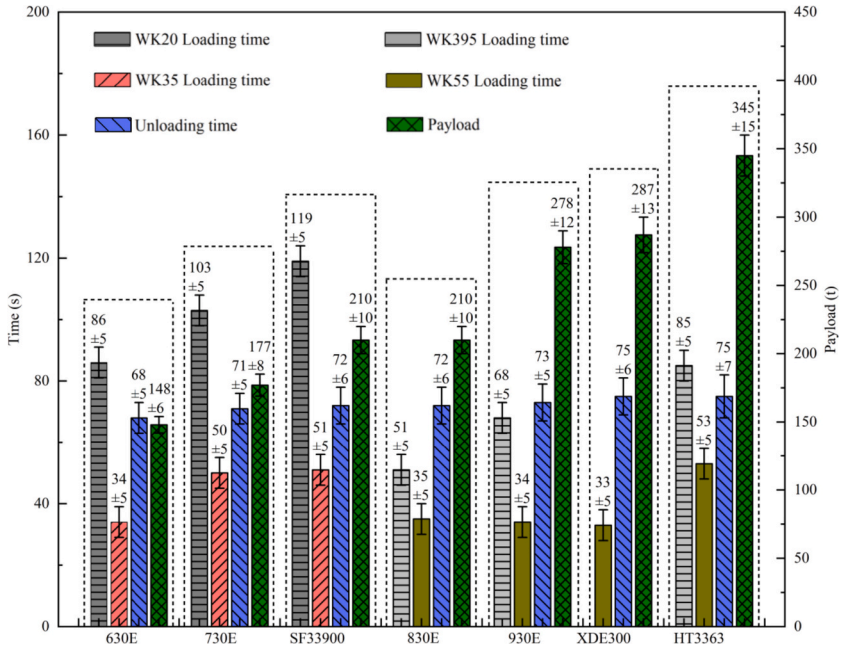

Fig. 10. Average and standard deviation of the loading time, payload, and unloading time for each truck type.

"I" shape, which includes four routes: $A_2 \rightarrow B_2$, $L_1 \rightarrow M_1$, $D_1 \rightarrow E_1$, and $C_1 \rightarrow B_1$. Taking the first two routes as examples, $A_2 \rightarrow B_2$ is a coal haul route: 2 WK35 shovels (at 1020 elevation) to the 1140 crushing station; $L_1 \rightarrow M_1$ is a rock haul route: 2 WK395 shovels (at 1030 elevation) to the 1190 dumping site.

The second type is the "X" shape, which includes two routes with overlapping segments: $A_2/C_1 \rightarrow B_2/B_1$ and $F_1/F_2 \rightarrow G_1/G_2$. Taking $A_2/C_1 \rightarrow B_2/B_1$ as an example: 2 WK35 shovels (at 1020 elevation)/3 WK20 shovels (at 1010 elevation) to the 1140/1150 crushing station. The defining feature of the "X" shape is that two routes converge at a point to form a new common route.

The third type is the "Y" shape, which also includes two routes with overlapping segments: $J_1 \rightarrow K_1/K_2$ and $H_1/H_2 \rightarrow I_1$. Taking $J_1 \rightarrow K_1/K_2$ as an example: 1 WK55 shovel (at 1090 elevation) to the 1180/1120 dumping site. Similar to the "X" shape, the "Y" shape involves multiple routes converging at a point to form a new common route. The basic information of each road, including the type and number of shovels, truck type, haul distance, lift height, and lift height per kilometer, is summarized in Table 3.

5.3. Dataset

The diversity and number of samples are crucial for enhancing model robustness and generalizability [42,43]. In this study, the samples are derived from a full year of truck operational data, spanning from December 2021 to November 2022, excluding extreme weather conditions such as rain, snow, and fog. For each route, 500 samples are randomly selected from the data in Table 3 to construct the time and FC models. Taking the $A_2 \rightarrow B_2$ route as an example, 500 random samples are chosen for each of the 630E, 730E, and SF33900 trucks, resulting in a total of 1500 samples. Each sample contains all the information listed in Table 1. Using this sampling method, a total of 15,000 samples are generated across the 10 haul routes shown in Table 3.

When the FC model is constructed, the input of Eq. (6) is $X_{round\ trip}$ ($X_{round\ trip} = (t_1, t_2, t_4, t_5, d_1, d_2, p, h)$), and the output is the FC of a truck for one haul cycle. The 15,000 samples are divided into training and validation sets at a ratio of 80 %:20 % for training the FC model. Section 6.1 presents data and verifies the reliability of the dataset.

5.4. Parameter design

5.4.1. Time model

For the case mine, the production shift time (T_s) is 8 h, and the mining (or stripping) volume (Q) is determined by the monthly mining (or stripping) volume of a single shovel, as specified in Table 2. Fig. 10 shows the average and standard deviation of the loading time (t_4), payload (p), and unloading time (t_5) for each truck type. The results indicate that the distribution of data points is within a reasonable range. Notably, the longest average loading time of 119 s occurred when the WK20 shovel loaded the SF33900 truck, while the shortest average loading time of 33 s occurred when the WK55 shovel loaded the XDE300 truck. The average payloads for the 630E, 730E, SF33900, 830E, 930E, XDE300, and HT3363 trucks are 148, 177, 210, 210, 278, 287, and 348 tons, respectively, and their average unloading times are 68, 71, 72, 72, 73, 75, and 75 s, respectively. These parameter values will be used to construct Eqs. (4) and (5).

5.4.2. FC model

Although the theoretical advantages of using the KPCA-ANN model have been discussed in Section 4.2.1.2, three different neural network models (ANN, PCA-ANN, and KPCA-ANN) are established for performance comparison. These models differ in that the ANN does not apply special treatment to the input parameters, the PCA-ANN uses EDV or singular value decomposition of the input parameters in the PCA layer to obtain the principal components, and the KPCA-ANN employs a kernel function and EVD for the input parameters in the KPCA layer to obtain the principal components. For all three models, the parameters to be determined include the learning rate, the number of network layers, the number of units per layer, the batch size, the number of epochs, and the activation function. Additionally, for PCA-ANN and KPCA-ANN, the number of principal components needs to be determined, with the latter also requiring the specification of the kernel function.

All three models employed the Adam optimization method [28] with a default learning rate of 0.001. In this investigation, a grid search

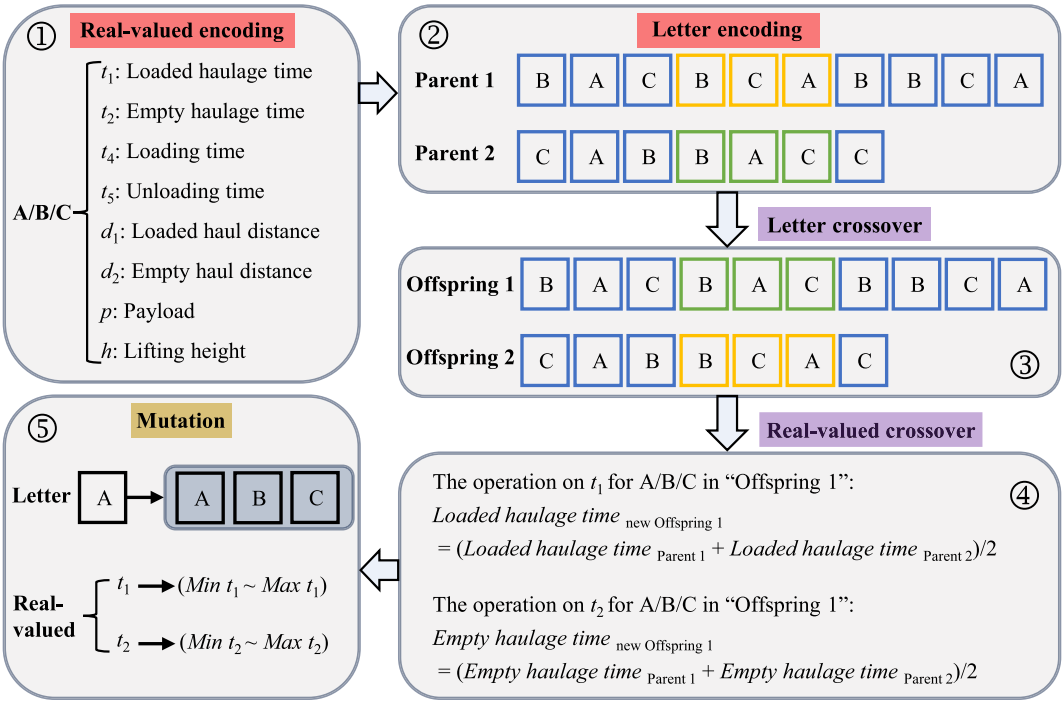

Fig. 11. Encoding, crossover, and mutation operations on letters and real values.

Table 4
Performance comparison of three neural networks based on R^2 score and MAPE.

Parameter	ANN	PCA-ANN			KPCA-ANN		
Kernel function	–	–			RBF		
Number of principal components	–	5	6	7	5	6	7
Explained variance ratio or kernel explained variance ratio	–	0.603	0.661	0.745	0.842	0.881	0.993
Number of hidden layers	3	2			2		
Number of neurons in layers	8–8–16-8-1	*x*-16-8-1			*x*-16-8-1		
Batch size	250	250			250		
Optimization algorithm	Adam	Adam			Adam		
Number of epochs	500	500			500		
Activation function	Leaky ReLU	Leaky ReLU			Leaky ReLU		
Indicator							
R^2 score	0.88	0.633	0.715	0.791	0.774	0.868	0.992
MAPE	10.6 %	28.7 %	23.4 %	19.3 %	21.8 %	11.4 %	2.8 %

method [44] is used to optimize the number of network layers, units per layer, batch size, and epochs, with the values of R^2 and MAPE optimization metrics presented in Section 6.1. Both the hidden and output layers utilize the Leaky Rectified Linear Unit (Leaky ReLU) [28,45] activation function.

For PCA-ANN and KPCA-ANN, R^2 and MAPE are computed for principal component numbers 5, 6, and 7, respectively. The "explained variance" metric is determined to be 0.99 for these transformations, indicating that the principal components account for 99 % of the original data. In KPCA-ANN, the kernel function uses a Radial Basis Function (RBF), which calculates the Gaussian similarity between input vectors and training samples, effectively handling cases of nonlinear separability [36,37].

In this study, KPCA-ANN, PCA-ANN, and ANN followed the method in the references [23, 29], performing supervised learning by randomly initializing the weights. The performances of these three neural network models are compared in detail in Section 6.1.

Table 5
Ground truth vs. KPCA-ANN prediction.

Route purpose	Route	Match	Input $X_{round\ trip} = (t_1, t_2, t_4, t_5, d_1, d_2, p, h)$	Truth TFC (L)	KPCA-ANN prediction (L)
Coal haul	$A_2 \rightarrow B_2$	WK35 + 630E	[852, 802, 36, 72, 5.8, 5.9, 147, 120]	90.5	87.4
		WK35 + 730E	[926, 933, 58, 65, 6.2, 6.4, 176, 120]	112.4	109.5
	$C_1 \rightarrow B_1$	WK20 + 730E	[885, 751, 50, 74, 5.9, 5.8, 177, 140]	111.7	113.1
		WK20 + SF33900	[875, 803, 54, 68, 6.1, 6.2, 215, 140]	156.1	160.2
	$D_1 \rightarrow E_1$	WK55 + 830E	[554, 536, 35, 67, 4.0, 4.2, 217, 60]	83.1	81.6
		WK55 + 930E	[653, 604, 32, 71, 4.7, 4.7, 286, 60]	117.5	119.5
	$F_1 \rightarrow G_1$	WK55 + 930E	[299, 319, 34, 68, 2.2, 2.4, 287, 5]	49.8	47.8
		WK55 + XDE300	[353, 373, 31, 72, 2.6, 2.8, 296, 5]	61.3	62.9
	$F_2 \rightarrow G_2$	WK55 + 830E	[308, 262, 31, 68, 2.2, 2.0, 217, 40]	48.4	50.7
		WK55 + XDE300	[330, 309, 36, 70, 2.3, 2.4, 298, 40]	64	67.7
Rock haul	$H_1 \rightarrow I_1$	WK55 + 930E	[679, 668, 34, 70, 5.0, 5.1, 288, 15]	110.7	108.1
		WK55 + HT3363	[867, 797, 52, 68, 5.9, 5.8, 357, 15]	161.4	167.3
	$H_2 \rightarrow I_1$	WK55 + 830E	[728, 636, 37, 72, 5.2, 5.0, 218, 30]	89.9	87.1
		WK55 + 930E	[834, 765, 32, 67, 5.7, 5.8, 287, 30]	131.8	134.6
	$J_1 \rightarrow K_1$	WK55 + 830E	[1171, 1057, 32, 67, 8.2, 8.1, 217, 90]	163.5	168.2
		WK55 + HT3363	[1145, 973, 53, 71, 7.7, 7.7, 356, 90]	238.7	232.3
	$J_1 \rightarrow K_2$	WK55 + 930E	[1133, 1017, 33, 68, 7.9, 7.8, 288, 130]	204.7	200.1
		WK55 + HT3363	[1272, 667, 51, 71, 8.2, 8.3, 358, 130]	283.8	277.6
	$L_1 \rightarrow M_1$	WK395 + 830E	[415, 360, 34, 69, 3.0, 2.8, 217, 45]	63.4	62.5
		WK395 + 930E	[400, 368, 36, 71, 2.8, 2.9, 287, 45]	76.3	74.3

5.4.3. Arp-NSGA-iii

This subsection outlines the individual encoding, crossover, and mutation methods, along with the design of key parameters, including population size, number of initial reference points, number of iterations, crossover rate, and mutation rate. The operation methods and parameter designs are described in detail below.

The population size is set to 200 individuals, with a set of 91 uniformly distributed initial reference points and a maximum of 3000 iterations. Individuals are generated using both real-valued and letter encoding.

(1) The real-valued encoding represents information related to the round trip of a truck in Table 1, as shown in Fig. 11 ①. Among them, t_1 and t_2 are the decision variables to be optimized, with value ranges determined by Eqs. (2) and (7). The remaining parameters are determined by specific haul routes and equipment.

(2) The letter encoding represents the type (L_{type}) and number (n) of trucks in the transport grouping, as shown in Fig. 11 ②. According to the industry experience of "big trucks with big shovels, small trucks with small shovels," there are usually no more than three types of trucks. The range of the number of trucks is determined by Eqs. (12) and (13). For example, "Parent 1" consists of three types of trucks, A, B, and C, with numbers 3, 4, and 3, respectively.

Different crossover operations and crossover rates are applied to the letter encoding and real-valued encoding, respectively.

(1) Letter crossover: Two individuals are randomly selected as "Parent 1" and "Parent 2". Then, a position from the shorter "Parent" is randomly chosen for crossover. Notably, when the number of letters of the shorter individuals in "Parent 1" and "Parent 2" is less than or equal to 6, two adjacent letters from the shorter individuals are randomly selected for crossover. Otherwise, if the number of letters exceeds 6, three adjacent positions are randomly selected for crossover. For example, in Fig. 11 ②, the orange and green positions are designated for crossover, resulting in "Offspring 1" and "Offspring 2" in Fig. 11 ③. The letter crossing rate is set to $P_{lc} = 1$.

(2) Real-valued crossover: Taking Fig. 11 ③ as an example, the averaging operation is performed in Fig. 11 ④ on the loaded haulage time (t_1) and empty haulage time (t_2) of "Offspring 1" to generate a "new Offspring 1." The same operation is applied with "Offspring 2" to produce a "new Offspring 2." The real-valued crossover rate is set to $P_{rc} = 0.8$.

Different mutation operations and mutation rates are applied to the letter encoding and real-valued encoding, respectively.

(1) Letter mutation: Select any position of "new Offspring 1" for

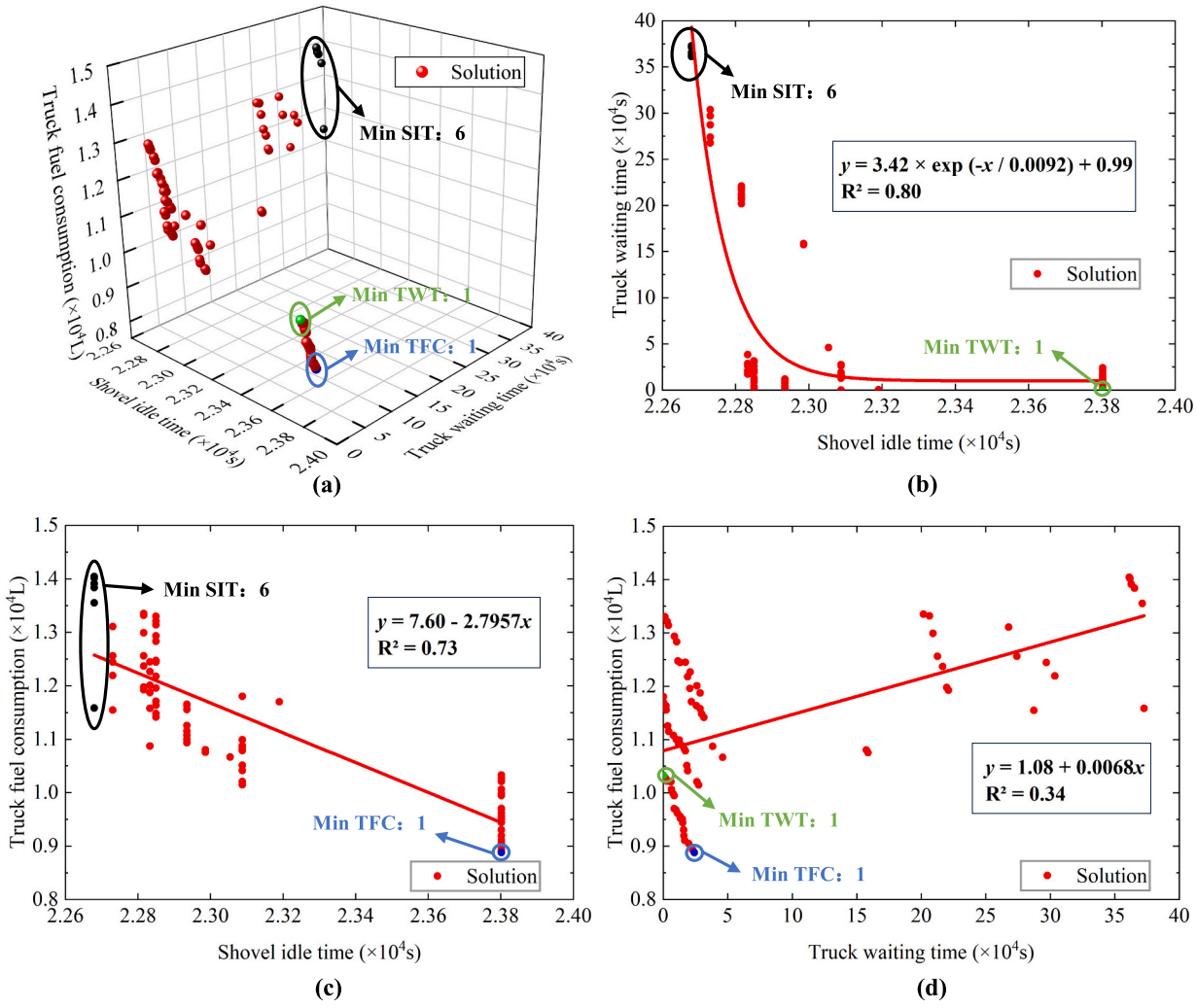

Fig. 12. Solutions obtained by the ARP-NSGA-III. (a) Three-dimensional Pareto front plot. (b) Relationship between the SIT and TWT. (c) Relationship between the SIT and TFC. (d) Relationship between the TWT and TFC.

Table 6
"Reference case" vs. four types of nondominated solutions ("Min SIT," "Min TWT," "Min TFC," and "overall best").

Type	Types and numbers of trucks	Hauling speed (km/h)		SIT (s)	TWT (s)	TFC (L)
		LHS	EHS			
Reference case	A:3, B:4, C:5	24.3	27.2	23,190	112,830	11,777
Min SIT	A:3, B:5, C:13	26.3	28.1	22,680	372,762	11,582
	A:6, B:5, C:11	22.7	27.9	22,680	371,268	13,550
	A:6, B:5, C:11	22.3	27.0	22,680	365,328	13,837
	A:6, B:5, C:11	22.3	26.4	22,680	362,820	13,909
	A:6, B:4, C:12	22.0	26.6	22,680	361,764	14,016
	A:6, B:4, C:12	22.0	26.5	22,680	361,236	14,044
Min TWT	A:0, B:0, C:7	22.3	26.4	23,802	112	10,326
Min TFC	A:0, B:0, C:7	26.5	29.0	23,802	24,122	8878
Overall best	A:2, B:5, C:2	26.3	27.7	22,935	7935	11,080

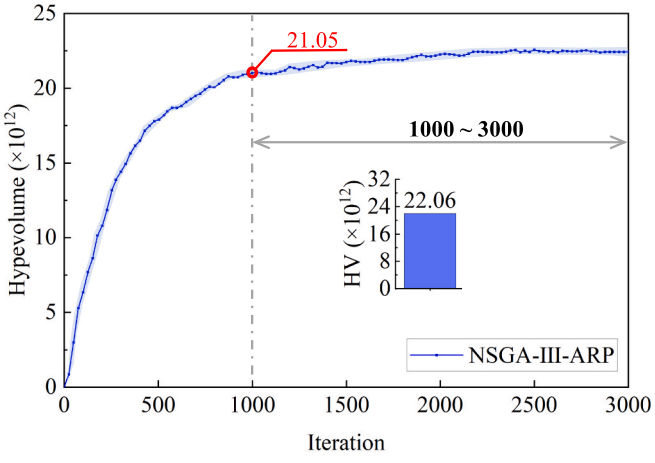

Fig. 13. Variation of HV in three-objective optimization with the number of iterations.

letter mutation. Using Fig. 11 ⑤ as an example, when the letter at the position to be mutated in "new Offspring 1" is A, the mutation range of the letter is (A ~ C).

(2) Real-valued mutation: Apply real-valued mutation to t_1 and t_2 of "new Offspring 1." Using Fig. 11 ⑤ as an example, the variation range of real values is determined by Eqs. (2) and (7). Both the letter and real-valued mutation rates are set to 0.3.

6. Result and discussion

This section employs Keras as the deep learning framework and TensorFlow as the backend engine to model the KPCA-ANN. Furthermore, the ARP-NSGA-III is modeled using the Distributed Evolutionary Algorithms in Python documentation. The experiments are conducted on a workstation equipped with a 4210R processor, 64 GB of RAM, and an A4000 GPU, with the software environment set to Anaconda Python 3.5.1.

6.1. Validation of KPCA-ANN

The optimal parameters and results obtained via the grid search method are summarized in Table 4. The results show that when the seven principal components obtained by KPCA are used as the input of the ANN, under the optimal model structure after adjustment, the R^2 score obtained by KPCA-ANN is 0.992, and the MAPE is 2.8 %. This performance outperforms both the ANN and PCA-ANN.

Table 5 compares the actual TFC with the predicted values. Two round trips from each haul route in Table 3 are randomly selected as inputs for the trained KPCA-ANN model. The results indicate that the prediction errors are within an acceptable range [23,29,30], providing a foundation for subsequent MOO.

6.2. Evaluation of ARP-NSGA-III

6.2.1. Case analysis

6.2.1.1. Case 1: Three-objective optimization. Topology 1 in Fig. 9 includes four routes: $A_2 \rightarrow B_2$, $L_1 \rightarrow M_1$, $D_1 \rightarrow E_1$, and $C_1 \rightarrow B_1$. Taking $A_2 \rightarrow B_2$ as an example, Eq. (18) is used to model this route, representing a three-objective optimization problem with $b = 1$.

In the operational scenario described in Section 5.2, two WK35 shovels operate continuously for 8 h at the A_2 loading point (one production shift lasts 8 h, with three shifts per working day). The mining volume (Q) of the two WK35 shovels during one shift is calculated to be 21,667 tons via Eq. (22). Here, q_i represents the monthly mining volume of a type-i shovel (information provided in Table 2), and n_{shovel} is the number of shovels. The coal density is 1.5 t/m^3. As shown in Table 4 and Table 5, the truck types and their respective maximum numbers on this route are: 6 units of 630E, 5 units of 730E, and 19 units of SF33900. These truck types and their numbers are represented using letter encoding described in Section 5.4.3, with 630E labeled as "A," 730E as "B," and SF33900 as "C." The haulage times for both loaded and empty trucks are encoded using real-valued encoding. Finally, the ARP-NSGA-III is used to solve Eq. (18).

$$Q = \frac{q_i \times n_{shovel}}{30 \times 3} \times 1.5t \Big/ m^3 \qquad \forall n_{shovel} \in Z^+ \tag{22}$$

The solution space of ARP-NSGA-III is shown in Fig. 12. Black points represent solutions minimizing the SIT, green points represent solutions minimizing the TWT, and blue points represent solutions minimizing the TFC. Fig. 12(a) demonstrates that the approximate POSs (79 solutions) are densely clustered in certain areas, covering a wide range, indicating that ARP-NSGA-III has excellent "local-global" search capabilities. Figs. 12(b) to 12(d) illustrate the relationships between the objectives: a negative correlation between the SIT and TWT, a negative correlation between the SIT and TFC, and a positive correlation between the TWT and TFC. These results highlight that finding the best solution to balance the SIT, TWT, and TFC is indeed a challenging task.

To identify and evaluate the overall best solution, a "reference case" is established, with LHS and EHS determined via Eq. (2). As shown in Table 6, the objective value for "Min SIT," "Min TWT," and "Min TFC" is reduced compared with the "reference case." For example, "Min SIT" shows a reduction of 510 s in the SIT, "Min TWT" shows a significant reduction of 112,718 s in the TWT, and "Min TFC" reveals a reduction of 2,899 L (24.6 %) in the TFC relative to the "reference case." In contrast, the "overall best" shows reductions in the SIT, TWT, and TFC by 255 s, 104,895 s, and 697 L (5.9 %), respectively, compared with the "reference case." Obviously, the "overall best" does not optimize any single objective as effectively as the "Min SIT," "Min TWT," or "Min TFC." The key advantage of the "best overall" lies in its ability to achieve balanced improvements across all three objectives: the SIT, TWT, and TFC. Notably, Table 6 displays the speed accuracy to one decimal place, but actual applications require these values to be represented as integers.

The hypervolume (HV) [28,46] is used to evaluate the quality of the solution set, with the results shown in Fig. 13. As the number of iterations increases, the HV rises rapidly and stabilizes after about 1000 iterations, with the curve remaining smooth throughout the process. Between 1000 and 3000 iterations, the average HV of ARP-NSGA-III is 22.06×10^{12}. This value will be compared with the baseline methods in Section 6.2.2 to further evaluate the optimization performance of ARP-NSGA-III.

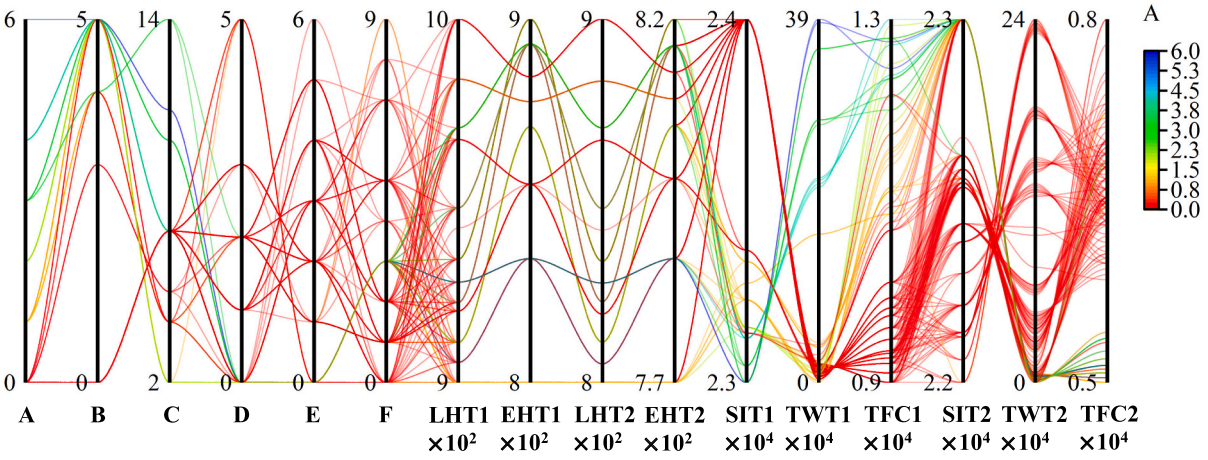

Fig. 14. Solutions obtained by the ARP-NSGA-III, with each polyline representing a truck assignment scheme.

Table 7
"Reference case" vs. seven types of nondominated solutions ("Min SIT1," "Min TWT1," "Min TFC1," "Min SIT2," "MinTWT2," "Min TFC2," and "overall best").

Type	Types and numbers of trucks	Hauling speed (km/h)		SIT1 (s)	TWT1 (s)	TFC1 (L)	SIT2 (s)	TWT2 (s)	TFC2 (L)
		LHS	EHS						
Reference case	A:3, B:3, C:5 D:3, E:2, F:2	24.4	26.8	23,190	81,048	11,893	23,224	97,408	6291
Min SIT1	A:6, B:5, C:11 D:0, E:0, F:3	26.2	27.4	22,680	387,636	12,249	23,445	5355	4752
	A:6, B:5, C:11 D:0, E:0, F:3	25.4	27.4	22,680	383,940	12,523	23,445	4185	4872
	A:3, B:4, C:14 D:0, E:0, F:3	24.0	26.6	22,680	356,130	12,561	23,445	855	5132
	A:3, B:4, C:14 D:2, E:1, F:1	24.0	26.6	22,680	356,130	12,561	22,935	2415	6842
Min SIT2	A:0, B:0, C:7 D:1, E:3, F:8	26.4	27.9	23,802	20,692	8997	22,085	239,165	6276
	A:0, B:0, C:7 D:2, E:1, F:9	26.0	26.9	23,802	16,282	9194	22,085	236,645	6403
	A:1, B:4, C:4 D:2, E:1, F:9	26.0	26.9	23,054	38,174	10,975	22,085	236,645	6403
Min TWT1	A:0, B:0, C:7 D:0, E:5, F:7	25.7	27.1	23,802	15,890	9252	22,085	236,465	6448
	A:1, B:5, C:2 D:0, E:0, F:3	25.6	26.6	22,935	855	11,419	23,445	3375	4873
Min TWT2	A:0, B:0, C:7 D:0, E:0, F:3	23.6	26.8	23,802	6972	9872	23,445	450	5194
	A:2, B:5, C:2 D:0, E:0, F:3	23.6	26.8	22,850	9440	12,693	23,445	450	5194
Min TFC1	A:0, B:0, C:7 D:0, E:0, F:3	26.4	27.9	23,802	20,692	8997	23,445	6255	4707
	A:0, B:0, C:7 D:1, E:3, F:0	26.4	27.9	23,802	20,692	8997	22,935	9615	6276
	A:0, B:0, C:7 D:0, E:3, F:2	26.4	27.9	23,802	20,692	8997	22,816	46,606	5753
	A:0, B:0, C:7 D:3, E:2, F:0	26.4	27.9	23,802	20,692	8997	22,833	30,003	6799
	A:0, B:0, C:7 D:1, E:4, F:0	26.4	27.9	23,802	20,692	8997	22,884	38,364	6276
	A:0, B:0, C:7 D:5, E:0, F:0	26.4	27.9	23,802	20,692	8997	22,850	21,710	7322
	A:0, B:0, C:7 D:1, E:2, F:4	26.4	27.9	23,802	20,692	8997	22,680	102,408	5858
	A:0, B:0, C:7 D:0, E:3, F:5	26.4	27.9	23,802	20,692	8997	22,493	131,021	5858
	A:0, B:0, C:7 D:2, E:5, F:2	26.4	27.9	23,802	20,692	8997	22,374	148,068	6590
	A:0, B:0, C:7 D:0, E:6, F:4	26.4	27.9	23,802	20,692	8997	22,272	181,752	6276
	A:0, B:0, C:7 D:1, E:3, F:8	26.4	27.9	23,802	20,692	8997	22,085	239,165	6276
	A:0, B:0, C:7 D:0, E:0, F:3	26.4	27.9	23,802	20,692	8997	23,445	6255	4707
Min TFC2	A:1, B:5, C:2 D:0, E:0, F:3	26.4	27.9	22,935	9135	11,017	23,445	6255	4707
	A:2, B:5, C:2 D:0, E:0, F:3	26.4	27.9	22,850	27,080	11,568	23,445	6255	4707
	A:1, B:4, C:4 D:0, E:0, F:3	26.4	27.9	23,054	43,439	10,741	23,445	6255	4707
	A:1, B:5, C:7 D:0, E:0, F:3	26.4	27.9	22,986	158,571	10,741	23,445	6255	4707
Overall best	A:1, B:5, C:2 D:0, E:0, F:3	26.4	27.9	22,935	9135	11,017	22,935	6255	4707

6.2.1.2. Case 2: Six-objective optimization. Topology 2 in Fig. 9 includes two routes: $A_2/C_1 \rightarrow B_2/B_1$ and $F_1/F_2 \rightarrow G_1/G_2$, and topology 3 includes two routes: $J_1 \rightarrow K_1/K_2$ and $H_1/H_2 \rightarrow I_1$. A key feature of these topologies is that multiple routes converge at certain nodes, forming new overlapping segments. Taking $A_2/C_1 \rightarrow B_2/B_1$ as an example, Eq. (18) is used to model this route, representing a six-objective optimization problem with $b = 2$.

When $A_2 \rightarrow B_2$ and $C_1 \rightarrow B_1$ operate simultaneously, according to Section 5.2 information, the A_2 loading point is equipped with two WK35 shovels, while the C_1 loading point, constrained by space, is equipped with three smaller WK20 shovels. Assuming each production shift lasts 8 h with three shifts per day, the mining volume (Q) for two WK35 shovels and three WK20 shovels in one shift is 21,667 tons and 10,000 tons, respectively, as calculated using Eq. (22). As shown in Table 4 and Table 5, the truck types and their maximum numbers on the $A_2/C_1 \rightarrow B_2/B_1$ route are as follows: 6 units of 630E, 5 units of 730E, and 19 units of SF33900. The truck types and numbers are represented using the letter encoding described in Section 5.4.3. In the $A_2 \rightarrow B_2$ route, 630E is labeled as "A," 730E as "B," and SF33900 as "C." In the $C_1 \rightarrow B_1$ route, 630E is labeled as "D," 730E as "E," and 830E as "F." The haulage times for loaded and empty trucks are encoded using real-valued encoding. The ARP-NSGA-III is then used to solve Eq. (18).

When six objective functions are involved, three-dimensional Pareto front plots can no longer effectively display the MOO results. Therefore,

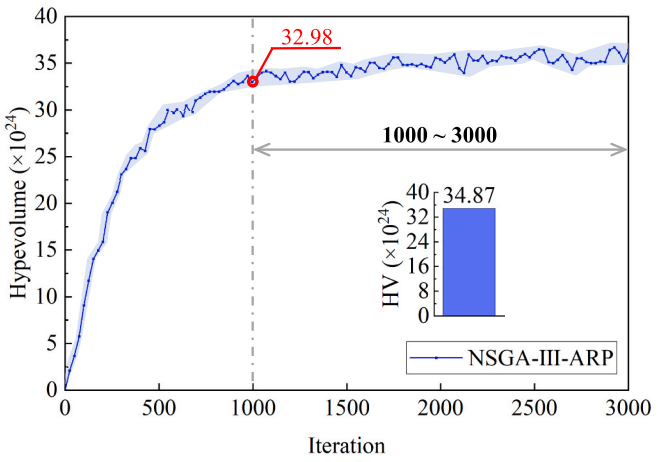

Fig. 15. Variation of HV in six-objective optimization with the number of iterations.

this study uses the parallel coordinates plot in Fig. 14 to visually display the solution results. Compared to Case 1, the number of solutions generated by the ARP-NSGA-III in Case 2 significantly increases, from 79 to 153. This indicates that as the number of overlapping routes increases, more solutions are generated.

Similar to Case 1, a "reference case" is established in Case 2 to select and evaluate the overall best solution. As shown in Table 7, the specified objective value for "Min SIT1," "Min SIT2," "Min TWT1," "Min TWT2," "Min TFC1," and "Min TFC2" is reduced compared with the "reference case." For example, compared with the "reference case," "Min SIT1" and "Min SIT2" show reductions in the SIT of 510 s and 1139 s, respectively; "Min TWT1" and "Min TWT2" show significant reductions in the TWT of 80,193 s and 96,958 s, respectively; and "Min TFC1" and "Min TFC2" reveal reductions in the TFC of 2,896 L (24.4 %) and 1,584 L (25.2 %), respectively. In contrast, the "overall best" shows reductions in the SIT1, TWT1, TFC1, SIT2, TWT2, and TFC2 by 255 s, 71,913 s, 876 L (7.4 %), 289 s, 91,153 s, and 1,584 L (25.2 %), respectively, compared with the "reference case." Same to Case 1, the "overall best" does not optimize any single objective as effectively as the specific objectives such as "Min SIT1," "Min SIT2," "Min TWT1," "Min TWT2," "Min TFC1," or "Min TFC2." The key advantage of the "best overall" is its ability to achieve balanced improvements across all six objectives. Notably, similar to Table 6, Table 7 displays speed accuracy to one decimal place, but actual applications require speed values to be represented as integers.

The HV is used to evaluate the quality of the solution set, with the results shown in Fig. 15. Compared to Case 1, the amplitude of the curve fluctuation is significantly larger. Between 1000 and 3000 iterations, the average HV of ARP-NSGA-III is 34.87 $\times$ 10^{24}. This value will be compared with the baseline methods in Section 6.2.2 to further evaluate the optimization performance of ARP-NSGA-III.

6.2.2. Comparison with baseline methods

In recent years, researchers have proposed new multi-objective heuristic optimization algorithms capable of handling complex problems with three or more objectives. These algorithms have made significant progress in balancing objectives, exploring the Pareto front, and maintaining solution diversity. To evaluate the performance of the ARP-NSGA-III, a comparison is made with the following five baseline methods.

(1) NSGA-III with uniform crossover and adaptive mutation (NSGA-III UCAM) [47], which combines uniform crossover and adaptive mutation based on the traditional NSGA-III.
(2) SPEA2 [19], which incorporates a fine-grained fitness assignment strategy, a density estimation technique, and an enhanced archive truncation method.
(3) MOEA/D [18], which decomposes the MOO problem into multiple scalar optimization subproblems, optimizing each using information from its neighboring subproblems.
(4) Reference Vector-guided Evolutionary Algorithm (RVEA) [48], which proposes an adaptation strategy to dynamically adjust the distribution of reference vectors based on the scales of the objective functions.
(5) Approximation-Guided Evolution II (AGE-II) [49], which stores all nondominated points in an archive to control its size and influence on runtime, allowing a trade-off between approximation quality and runtime, and enabling a faster approximation process.

Table 8 presents the optimization performance of the six methods on the two cases in Section 6.2.1, including the average HV during 1000 to 1500 iterations, the number of approximately POSs, and the optimized objective values in each dimension. The research results show that:

(1) The performance of the ARP-NSGA-III significantly outperforms that of other baseline methods. In both cases, the ARP-NSGA-III achieves the highest HV, generates the largest number of POSs, and attains the optimal value in each objective dimension.
(2) The RVEA and AGE-II have more obvious advantages in handling high-dimensional MOO problems. In case 1, the HV, the number of approximate POSs, and the optimal values of each objective dimension of the NSGA-III UCAM are significantly better than those of the RVEA and AGE-II; whereas in case 2, the HV and the number of approximate POSs of the RVEA and AGE-II are significantly better than those of the NSGA-III UCAM. In both research cases, the HV and the number of approximate POSs of the RVEA, AGE-II, and NSGA-III UCAM are better than those of the SPEA2 and MOEA/D.

6.3. Integration architecture application

6.3.1. FTA optimization

On the basis of the studies of Case 1 and Case 2, this subsection

Table 8
Comparison of optimization performance between the ARP-NSGA-III and five baseline methods.

Method	Three-objective optimization (Case 1)					Six-objective optimization (Case 2)							
	Average HV	Number of solutions	Min SIT (s)	Min TWT (s)	Min TFC (L)	Average HV	Number of solutions	Min SIT1 (s)	Min SIT2 (s)	Min TWT1 (s)	Min TWT2 (s)	Min TFC1 (L)	Min TFC2 (L)
ARP-NAGA-III	21.36	75	22,680	112	8878	33.78	148	22,680	22,085	855	450	8997	4707
NSGA-III UCAM [47]	19.20	33	22,680	2856	9225	30.36	75	22,731	22,170	2535	1275	8983	4751
SPEA2 [19]	16.58	20	23,802	9188	9840	26.22	53	23,802	23,445	6255	4215	8988	4727
MOEA/D [18]	17.14	22	23,088	8442	9843	27.26	50	23,802	23,445	5355	3375	8913	4655
RVEA [48]	18.73	28	23,054	6776	9426	30.92	86	22,833	22,850	1974	1491	8532	4428
AGE-II [49]	18.89	29	22,935	5992	9533	31.46	92	22,816	22,680	1875	854	8617	4531

Table 9
Optimization results of the integrated architecture on all routes vs. the reference cases.

Route purpose	Route	Match	Haul distance (km)	Lifting height (m)	Types and numbers of trucks (A: B: C)	Optimized LHS1 (km/h)	Optimized EHS1 (km/h)	Optimized objective values 1	Types and numbers of trucks (A: B: C: D: E: F)	Optimized LHS2 (km/h)	Optimized EHS2 (km/h)	Optimized objective values 2	Reference cases (A: B: C)	LHS (km/h)	EHS (km/h)	References values	Optimization 1	Optimization 2
Coal haul	$A_2 \rightarrow B_2$	WK35 630E 730E SF33900	6.5	120	1: 5: 3	26.4	28.6	23,054 s 46,013 s 10,662 L	1: 5: 2: 0: 0: 3	26.4	27.9	22,935 s 9135 s 11,017 L	3: 3: 5	23.4	26.7	23,190 s 75,724 s 12,273 L	136 s 29,711 s 1611 L (13.1 %)	255 s 66,589 s 1256 L (10.2 %)
	$C_1 \rightarrow B_1$	WK20 630E 730E SF33900	6	140	0: 0: 3	26.7	28.2	23,445 s 6975 s 4661 L				23,445 s 6255 s 4707 L	2: 2: 2	22.1	27.6	23,592 s 68,472 s 6543 L	147 s 55,242 s 1882 L (28.8 %)	147 s 62,217 s 1836 L (28.1 %)
	$D_1 \rightarrow E_1$	WK55 830E 930E XDE300	4.2	60	3: 2: 0	23.5	27.7	25,230 s 8235 s 4663 L	–	–	–	–	2: 2: 3	23.9	26.4	25,706 s 82,390 s 5072 L	476 s 74,155 s 409 L (8.1 %)	–
	$F_1 \rightarrow G_1$	WK55 830E 930E XDE300	2.6	5	2: 1: 0	27.3	29.6	25,230 s 6285 s 1101 L	2: 1: 0: 2: 1: 0	25.1	28.4	25,230 s 1665 s 1167 L	1: 2: 2	25.4	26	25,740 s 69,030 s 1266 L	510 s 62,745 s 165 L (13.0 %)	510 s 67,365 s 99 L (7.8 %)
	$F_2 \rightarrow G_2$	WK55 830E 930E XDE300	2.4	40	2: 1: 0	22.4	27.9	25,230 s 2400 s 1758 L				25,230 s 7440 s 1668 L	1: 2: 2	22.5	27.1	25,740 s 71,370 s 1966 L	510 s 68,970 s 208 L (10.6 %)	510 s 63,930 s 298 L (15.2 %)
Rock haul	$H_1 \rightarrow I_1$	WK55 830E 930E HT3363	5.3	15	3: 0: 2	28.1	29.6	24,924 s 7694 s 5200 L	3: 0: 2: 3: 0: 2	27.7	29.4	24,924 s 6364 s 5158 L	2: 2: 4	23.6	26.4	25,060 s 85,772 s 6231 L	136 s 78,078 s 1031 L (16.5 %)	136 s 79,408 s 1073 L (17.2 %)
	$H_2 \rightarrow I_1$	WK55 830E 930E HT3363	5.5	30	3: 0: 2	27.1	29.8	24,924 s 949 s 5567 L				24,924 s 1614 s 5635 L	2: 2: 4	25.2	26.3	25,060 s 84,540 s 5713 L	136 s 83,591 s 146 L (2.6 %)	136 s 82,926 s 78 L (1.4 %)
	$J_1 \rightarrow K_1$	WK55 830E 930E HT3363	7.2	90	1: 0: 6	23.5	26.8	25,060 s 33,080 s 10,465 L	1: 0: 6: 3: 1: 4	23.2	27.9	25,060 s 34,928 s 10,452 L	2: 2: 5	24.1	26.2	25,281 s 82,746 s 10,880 L	221 s 49,666 s 415 L (3.8 %)	221 s 47,818 s 428 L (3.9 %)
	$J_1 \rightarrow K_2$	WK55 830E 930E HT3363	7.5	130	3: 1: 4	22.1	28.3	25,060 s 29,012 s 16,992 L				25,060 s 32,884 s 16,328 L	2: 2: 5	22.6	26.8	25,281 s 71,649 s 17,413 L	221 s 42,637 s 421 L (2.4 %)	221 s 38,765 s 1085 L (6.2 %)
	$L_1 \rightarrow M_1$	WK395 830E 930E HT3363	3.0	45	1: 0: 2	23.1	27.3	25,128 s 7074 s 2377 L	–	–	–	–	2: 2: 2	23.1	26.3	25,230 s 83,820 s 2661 L	102 s 76,746 s 284 L (10.7 %)	–

Table 10
Practical solutions for three scenarios provided by the integrated architecture.

No.	Application scenario	Challenge	Solution	Expected outcomes
1	Reduce the intervention of "dynamic scheduling."	Dynamic scheduling heavily relies on real-time communication and data transmission, leading to high system hardware and algorithm maintenance costs.	Generate the optimal FTA alternative.	Preserve the inherent advantages of FTA's high adaptability and economic efficiency.
2	Guide the equipment procurement for new mines and equipment upgrading for existing mines.	Lack of precise "shovel-truck ratio" relationships may result in an overabundance or shortage of equipment.	Establish an accurate "shovel-truck ratio" relationship.	(1) Prevent equipment overabundance or shortage, effectively controlling procurement costs. (2) Improve equipment utilization, reducing operational costs.
3	Provide parameter support for autonomous truck fleet operations.	Lack of necessary parameters, such as appropriate speeds for loaded and empty conditions.	Establish an "optimal speed database."	(1) Enhance the operational efficiency of autonomous truck fleets and reduce FC. (2) Provide data support for the updating and iteration of autonomous driving technology.

adopts the integrated architecture developed in Section 4.4 to generate balanced transport groupings and optimal fleet speeds for all haul routes in the case mine.

Table 9 lists the FTA optimization solutions and their corresponding objective values. Specifically, it presents the three objective values ("Optimized objective values 1") for the independently operated routes and the six objective values ("Optimized objective values 2") for the routes with overlapping segments. All optimization solutions and their corresponding objective values are then compared with the "reference cases" and their corresponding "reference values." The results show that, compared with the "reference values," all objective values have been optimized to different extents. In "Optimization 1" (difference between "Reference values" and "Optimized objective values 1"), the SIT is reduced by up to 510 s, the TWT is reduced by up to 83,591 s, and the TFC is reduced by up to 1,882 L (28.8 %). In "Optimization 2" (difference between "Reference values" and "Optimized objective values 2"), the SIT is reduced by up to 510 s, the TWT is reduced by up to 82,926 s, and the TFC is reduced by up to 1,836 L (28.1 %).

6.3.2. *Efficiency and energy*

Based on the calculation results in Table 9, the annual cumulative reductions in equipment nonworking time and FC are estimated after implementing the integrated architecture in the case mine.

When all haul routes (including $A_2/C_1 \rightarrow B_2/B_1$, $D_1 \rightarrow E_1$, $F_1/F_2 \rightarrow G_1/G_2$, $H_1/H_2 \rightarrow I_1$, $J_1 \rightarrow K_1/K_2$, $L_1 \rightarrow M_1$) operate normally, the statistical results of "Optimization 1" and "Optimization 2" show that, for each production shift, the SIT, TWT, and TFC are cumulatively reduced by 2714 s, 659,919 s, and 6,846 L, respectively. Assuming 300 operating days per year and three shifts per day, according to Eq. (23), the integrated architecture is estimated to reduce the annual SIT by approximately 679 h and TWT by approximately 1.65×10^5 h. Here, T_o represents the annual cumulative reduction in equipment nonworking time (unit: hours, h), and T_{total} denotes the reduction in nonworking time per production shift.

$$T_o = \frac{T_{total} \times 3 \times 300}{3600} \tag{23}$$

According to Eq. (24), the integrated architecture is estimated to reduce the annual TFC by approximately 6.2×10^6 L. Here, F_o represents the annual cumulative reduction in TFC (unit: iters, L), and F_{total} represents the reduction in TFC per production shift. According to Eq. (25), the integrated architecture is estimated to reduce annual CO_2 emissions by approximately 1.65×10^4 tons. In this equation, E_o represents the annual CO_2 reduction (unit: kilograms, kg), C_e is the CO_2 emission factor for heavy truck diesel, approximately 2.68 kg CO_2/L according to the "2019 Intergovernmental Panel on Climate Change Guidelines for National GHG Inventories [50]." Considering that the Heidaigou open-pit coal mine generates approximately 2.0×10^5 tons of CO_2 annually from haul activities, implementing this integrated architecture could result in an estimated 8.3 % reduction in emissions.

$$F_o = F_{total} \times 3 \times 300 \tag{24}$$

$$E_o = C_e \times F_o \tag{25}$$

6.3.3. *Practical application*

In addition to the theoretical optimization results, the developed integrated architecture provides practical and viable solutions for actual open-pit mining operations. Table 10 shows three potential application scenarios of the architecture, summarizing the challenges encountered in each scenario, along with the corresponding solutions and expected outcomes. This architecture supports the development of efficient, environmentally friendly, and intelligent open-pit mines, demonstrating significant market value for both manned mine operations and the future construction of unmanned mines.

7. Conclusion and future work

This study develops an integrated architecture for FTA optimization in open-pit mines, significantly enhancing equipment utilization while reducing energy consumption. Specifically, the architecture optimizes transport groupings (including truck types and numbers) and fleet speeds (including LHS and EHS) for various haul distances, thereby reducing the SIT, TWT, and TFC. The advantages of this integrated architecture manifest in the following three core components: (1) The "haul distance–transport grouping–fleet speed" synergistic optimization approach. Compared with single-factor optimization approaches, this approach, which considers multiple types of factors, offers greater adaptability for complex production systems such as open-pit mines. (2) The hybrid modeling method combining mechanism-based and data-driven models. The proposed decrement model simplifies the time function compared with accumulation models, and the KPCA-ANN model outperforms traditional ANN and PCA-ANN models in capturing strong nonlinear relationships among features, improving FC prediction accuracy and generalization. (3) The ARP-NSGA-III optimization technology incorporating letter and real-valued encoding. Compared with the baseline methods, the ARP-NSGA-III significantly increases the number and distribution range of approximate POSs, and the hybrid encoding allows the simultaneous representation of heterogeneous parameters, such as truck type, number, and speed, which single encoding methods cannot achieve.

FTA optimization across all routes in the case mine is estimated to reduce the SIT by 679 h, the TWT by 1.65×10^5 h, and the TFC by 6.2×10^6 L annually. The reduction in FC is equivalent to reducing 1.65×10^4

tons of CO_2 emissions, accounting for 8.3 % of the total CO_2 emissions from haul activities. The integrated architecture provides theoretical guidance for actual open-pit mining operations, supporting the development of efficient, environmentally friendly, and intelligent mines. Future research will focus on several directions: adopting pre-training methods and Transfer Learning to enhance the generalization of data-driven models; considering the impact of dynamic parameters such as adaptive speed on FTA modeling; and exploring the joint optimization of FTA and dynamic scheduling.

CRediT authorship contribution statement

Hongyang Xu: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Hongze Zhao:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition. **Junyu Lu:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Hui Zuo:** Writing – review & editing, Validation, Supervision. **Wen Zheng:** Writing – review & editing, Validation, Supervision, Software. **Pei Guo:** Supervision, Resources, Project administration, Methodology, Investigation, Data curation. **Di Yang:** Writing – review & editing, Supervision. **Bo Zhang:** Supervision, Resources, Methodology. **Qipei Mei:** Writing – review & editing, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. There is no conflict of interest in this report.

Acknowledgements

This research was funded by the National Key Research and Development Program of China (No. 2022YFB4703701) and the China Scholarship Council.

Data availability

Data will be made available on request.

References

[1] Rodrigo M, Enrico Z, Fredy K, Adolfo A. Availability-based simulation and optimization modeling framework for open-pit mine truck allocation under dynamic constraints. Int J Min Sci Technol 2013;23:113–9.
[2] Konur D, Golias MM. Cost-stable truck scheduling at a cross-dock facility with unknown truck arrivals: a meta-heuristic approach. Transp Res Pt e-Logist Transp Rev 2013;49:71–91.
[3] Ristovski K, Gupta C, Harada K, Tang HK, Acm.. Dispatch with confidence: Integration of machine learning, optimization and simulation for open pit mines. In: 23rd ACM SIGKDD international conference on knowledge discovery and data mining (KDD). Halifax, Canada: Assoc Computing Machinery; 2017. p. 1981–9.
[4] Zhang XT, Guo A, Ai YF, Tian B, Chen L. Real-time scheduling of autonomous mining trucks via flow allocation-accelerated tabu search. IEEE T Intell Veh 2022;7:466–79.
[5] Zhang LJ, Xia XH. An integer programming approach for truck-shovel dispatching problem in open-pit mines. In: 7th international conference on applied energy (ICAE). Abu Dhabi, U. Arab Emirates: Elsevier Science Bv; 2015. p. 1779–84.
[6] Bakhtavar E, Mahmoudi H. Development of a scenario-based robust model for the optimal truck-shovel allocation in open-pit mining. Comput Oper Res 2020;115:10.
[7] Mohtashami A. Scheduling different types of vehicles in distribution centers with fixed due dates and packed shipments. Appl Soft Comput 2020;94:22.
[8] Bimba AT, Idris N, Al-Hunaiyyan A, et al. Towards knowledge modeling and manipulation technologies: a survey. Int J Inf Manage 2016;36:857–71.
[9] Choi SJ, Song HJ, Yoon HG, Park SB, Park SY. A re-ranking model for accurate knowledge base completion with knowledge base schema and web statistic. In: IEEE congress on evolutionary computation (CEC) held as part of IEEE world congress on computational intelligence (IEEE WCCI). Vancouver, Canada: IEEE; 2016. p. 4958–64.
[10] Gorgogianni A, Eliás J, Le JL. Mesh objective stochastic simulations of quasibrittle fracture. J Mech Phys Solids 2022;159:16.
[11] Gu J, Song YP, Wang YY, Meng W, Shi L, Deng KY. Prediction of heat release and NOx emissions for direct-injection diesel engines using an innovative zero-dimensional multi-phase combustion model. Fuel 2022;329:20.
[12] Chou HJ, Chang RY, Wu JM. Multi-objective optimization of wireless information and power transfer in multiuser OFDMA systems. In: IEEE global telecommunications conference (GLOBECOM). San Diego, CA: IEEE; 2015.
[13] Fukumoto H, Oyama A. Impact of estimation method of ideal/nadir points on practically-constrained multi-objective optimization problems for decomposition-based multi-objective evolutionary algorithm. In: IEEE symposium series on computational intelligence (SSCI). Xiamen, China: IEEE; 2019. p. 2138–45.
[14] Vandani B, Mansour F, Soltani M, Veysmoradi D. Bi-objective optimization for integrating quay crane and internal truck assignment with challenges of trucks sharing. Knowledge-Based Syst 2019;163:675–92.
[15] Zhang S, Lu CW, Jiang S, Shan L, Xiong NN. An unmanned intelligent transportation scheduling system for open-pit mine vehicles based on 5G and big data. IEEE Access 2020;8:135524–39.
[16] Deb K, Jain H. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints. IEEE Trans Evol Comput. 2014;18:577–601.
[17] Jain H, Deb K. An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, part II: handling constraints and extending to an adaptive approach. IEEE Trans Evol Comput 2014;18:602–22.
[18] Zhang QF, Li H. MOEA/D: a multiobjective evolutionary algorithm based on decomposition. IEEE Trans Evol Comput. 2007;11:712–31.
[19] Zitzler E, Laumanns M, Thiele L. SPEA2: Improving the strength pareto evolutionary algorithm. Computer Engineering and Networks Laboratory: ETH Zurich; 2001.
[20] Zhang C, Li HT, Li XX, Cheng JJ, Lei ZY, Gao SC. Probabilistic bootstrap-based evolutionary algorithm for three-objective wind farm turbine position optimization. Swarm Evol Comput 2025;96:17.
[21] Altin L, Topcuoglu HR, Gürgen FS. Network congestion aware multiobjective task scheduling in heterogeneous fog environments. IEEE Trans Ind Inform. 2024;20:3015–24.
[22] Afrapoli AM, Tabesh M, Askari-Nasab H. A multiple objective transportation problem approach to dynamic truck dispatching in surface mines. Eur J Oper Res 2019;276:331–42.
[23] Siami-Irdemoosa E, Dindarloo SR. Prediction of fuel consumption of mining dump trucks: a neural networks approach. Appl Energy 2015;151:77–84.
[24] Dindarloo SR, Siami-Irdemoosa E. Determinants of fuel consumption in mining trucks. Energy 2016;112:232–40.
[25] Shahabi-Shahmiri R, Asian S, Tavakkoli-Moghaddam R, Mousavi SM, Rajabzadeh M. A routing and scheduling problem for cross-docking networks with perishable products, heterogeneous vehicles and split delivery. Comput Ind Eng 2021;157:21.
[26] Fard SS, Vandani B. Assignment and scheduling trucks in cross-docking system with energy consumption consideration and trucks queuing. J Clean Prod 2019;213:21–41.
[27] Hakanen J, Chugh T, Sindhya K, Jin YC, Miettinen K. Connections of reference vectors and different types of preference information in interactive multiobjective evolutionary algorithms. In: IEEE symposium series on computational intelligence (IEEE SSCI). Athens, Greece: IEEE; 2016.
[28] Razmi A, Rahbar M, Bemanian M. PCA-ANN integrated NSGA-III framework for dormitory building design optimization: energy efficiency, daylight, and thermal comfort. Appl Energy 2022;305:19.
[29] Wang Q, Zhang RX, Lv SK, Wang YT. Open-pit mine truck fuel consumption pattern and application based on multi-dimensional features and XGBoost. Sustain Energy Technol Assess 2021;43:10.
[30] Sahoo LK, Bandyopadhyay S, Banerjee R. Benchmarking energy consumption for dump trucks in mines. Appl Energy 2014;113:1382–96.
[31] Hao HM, Cao JA, Wang HL, Yu ZQ, Liu JH. Coupling kernel principal component analysis with ANN for improving analysis accuracy of seven-component alkane gaseous mixture. In: IEEE international conference on computational intelligence for measurement systems and applications. Istanbul, Turkey: IEEE; 2008. p. 9-+.
[32] Xu L, Guo DL. New methods of fracture effect forecast of CBM wells based on KPCA and artificial neural networks. In: 3rd international conference on computational modeling, simulation and applied mathematics (CMSAM). Wuhan, China: Destech Publications, Inc; 2018. p. 152–6.
[33] Chen Y, Yu H, Liu CJ, Xie J, Han J, Dai HD. Synergistic fusion of physical modeling and data-driven approaches for parameter inference to enzymatic biodiesel production system. Appl Energy 2024;373:15.
[34] Chu YH, Yang DZ, Yu HX, Zhao X, Li MY. Can end-to-end data-driven models outperform traditional semi-physical models in 1-min irradiance? Appl Energy 2024;356:14.
[35] Song HD, Liu XJ, Song MQ. Comparative study of data-driven and model-driven approaches in prediction of nuclear power plants operating parameters. Appl Energy 2023;341:14.
[36] Liu CJ. Gabor-based kernel PCA with fractional power polynomial models for face recognition. IEEE Trans Pattern Anal Mach Intell 2004;26:572–81.

[37] Kuang FJ, Xu WH, Zhang SY. A novel hybrid KPCA and SVM with GA model for intrusion detection. Appl Soft Comput 2014;18:178–84.
[38] Pandey A, De Meulemeester H, De Moor B, Suykens JAK. Multi-view kernel PCA for time series forecasting. Neurocomputing 2023;554:9.
[39] Jain H, Deb K. An improved adaptive approach for elitist nondominated sorting genetic algorithm for many-objective optimization. 7th International Conference on Evolutionary Multi-Criterion Optimization (EMO). Sheffield, United Kingdom: Springer-Verlag Berlin. 2013. p. 307–21.
[40] Masood A, Chen G, Mei Y, Zhang MJ. Adaptive reference point generation for many-objective optimization using NSGA-III. 31st Australasian Joint Conference on Artificial Intelligence (AI). Victoria Univ Wellington, Wellington, New Zealand: Springer International Publishing Ag. 2018. p. 358–70.
[41] Logist F, Van Erdeghem PMM, Van Impe JF. Efficient deterministic multiple objective optimal control of (bio)chemical processes. Chem Eng Sci 2009;64:2527–38.
[42] Valtierra-Rodriguez M, Romero-Troncoso RD, Osornio-Rios RA, Garcia-Perez A. Detection and classification of single and combined power quality disturbances using neural networks. IEEE Trans Ind Electron 2014;61:2473–82.
[43] Deng YP, Wang L, Jia H, Tong XQ, Li F. A sequence-to-sequence deep learning architecture based on bidirectional GRU for type recognition and time location of combined power quality disturbance. IEEE Trans Ind Inform 2019;15:4481–93.
[44] Wang ZZ, Dong YJ, Liu W, Ma Z. A novel fault diagnosis approach for chillers based on 1-D convolutional neural network and gated recurrent unit. Sensors 2020;20:20.
[45] Xu J, Li ZS, Du BW, Zhang MM, Liu J. Reluplex made more practical: Leaky ReLU. In: 25th IEEE symposium on computers and communications (ISCC). Rennes, France: IEEE; 2020. p. 703–9.
[46] Bader J, Zitzler E. HypE: an algorithm for fast hypervolume-based many-objective optimization. Evol Comput 2011;19:45–76.
[47] Yi JH, Deb S, Dong JY, Alavi AH, Wang GG. An improved NSGA-III algorithm with adaptive mutation operator for big data optimization problems. Futur Gener Comp Syst 2018;88:571–85.
[48] Cheng R, Jin YC, Olhofer M, Sendhoff B. A reference vector guided evolutionary algorithm for many-objective optimization. IEEE Trans Evol Comput. 2016;20:773–91.
[49] Wagner M, Neumann F. A fast approximation-guided evolutionary multi-objective algorithm. In: 15th genetic and evolutionary computation conference (GECCO). Amsterdam, Netherlands: Assoc Computing Machinery; 2013. p. 687–94.
[50] Gómez DRWJ, Americano BB, et al. IPCC guidelines for national greenhouse gas inventories. Kamiyamaguchi Hayama, Japan: Institute for Global Environmental Strategies; 2019.