RAIRO-Oper. Res. 57 (2023) 1745–1765
https://doi.org/10.1051/ro/2023090

RAIRO Operations Research
www.rairo-ro.org

MULTI-OBJECTIVE OPTIMIZATION FOR MULTIMODAL TRANSPORTATION ROUTING PROBLEM WITH STOCHASTIC TRANSPORTATION TIME BASED ON DATA-DRIVEN APPROACHES

Yong Peng[1], Shu Han Gao[1], Dennis Yu[2,*], Yun Peng Xiao[1] and Yi Juan Luo[1]

Abstract. We study a multi-objective optimization model of a stochastic multimodal transportation network considering key impact factors such as transit cost, time, and transport mode schedule while minimizing total transportation cost and transportation time. In this study, we apply the Monte Carlo simulation to deal with the stochastic transportation time in the network and propose a data-driven approach that combines historical data and the dataset generated by the data mining algorithm to accelerate the search for the nondominated solution in the simulation. To validate the effectiveness of the proposed Data-Driven Multi-Objective Simulation Ant Colony (DD-MSAC) algorithm, we compare the optimum-seeking performance and the running time consumption of the Nondominated Sorting Genetic Algorithm-II (NSGA-II) and the Multi-Objective Simulation Ant Colony (MSAC) algorithm. Then, the MSAC algorithm is adopted as the benchmark for the comparison study on the solving performance of the proposed DD-MSAC algorithm. We conducted 30 times simulation run under different network scales in our numerical examples to show that the DD-MSAC algorithm can be equally effective as the non-data-driven MSAC algorithm in finding a nondominated solution as the average error does not exceed 5%. Meanwhile, we analyze the impact of different data-driven approaches, including data pool and support vector machine, on the solution quality and the running time. Finally, we use an example of China's Belt Road Initiative to verify the effectiveness of the proposed algorithm.

Mathematics Subject Classification. 90B06.

Received June 28, 2022. Accepted June 8, 2023.

1. Introduction

With the development of economic globalization and information technology, it has become much easier to communicate with each other in the world. A single transportation mode can no longer meet the needs of the transportation market, so multimodal transportation arises at the historic moment. A typical example of the multimodal transportation, which resembles advanced development of transportation, is the intermodal cargo transport which represents the transportation of cargo containers from the origin to the customer point *via* combined different transport modes such as sea, road, rail, and air [1]. Apparently, multimodal transporta-

Keywords. Stochastic transportation time, multimodal transportation, data-driven, ant colony algorithm, simulation optimization.

[1] School of Traffic and Transportation, Chongqing Jiaotong University, Chongqing 400074, P.R. China.
[2] The David D. Reh School of Business, Clarkson University, Potsdam, NY 13699, USA.
*Corresponding author: dyu@clarkson.edu

tion takes advantage of different transport modes to improve cargo transportation efficiency [2]. The optimal selection of transportation routes with multiple transport modes has become a focal point of the research in multimodal transportation due to the rapid growth of the container transportation industry [3]. In contrast, routing optimization and the concept of synchromodality are among the important research directions in multimodal transportation [4–6].

As the core problem of optimizing multimodal transport routes, the objective of route selection is to search for a route that meets the multi-needs and multi-objectives of customers or merchants under various uncertain conditions that can be represented by the uncertain transport duration between nodes of the network. From the perspective of optimization models, multimodal route optimization can be considered as an extension of the classic shortest path problem [7,8]. The nuance is that the multimodal transportation route optimization takes into account the actual transportation environment and innovative transportation methods while reducing transportation costs, increasing utilization, providing environmentally friendly transportation, and meeting customer needs.

However, with the continuously increasing scale of the multimodal transport network and the development of more accurate individual transport demand, the path choice problem calls for a dynamic real-time solution set with consideration of uncertain conditions and multiple objective functions, which leads to increased difficulty in solving the model. Therefore, the conventional algorithms cannot live up to the challenge of solving such problems because of their insufficient solving speed and low accuracy. There is an urgent need for a new effective algorithm to solve the stochastic multimodal transport route selection problem.

2. Literature review

We categorize the relevant literature into two research streams: the literature on multimodal transportation studies in different scenarios and the literature on algorithms for solving multimodal transportation models. We review the representative studies of these two respective categories in what follows.

2.1. Multimodal transportation study in different scenarios

In the existing literature on the route optimization of multimodal transportation, the optimization models are often developed with a single objective function under deterministic conditions [9,10]. However, the requirements of customers and transportation enterprises and the constraints of emergencies are often diverse, even conflicting in reality. For example, Zhang *et al.* [11] argue that decision-makers may have different preferences for cost, time, and emission targets in the decision-making process. Some customers prefer the lowest freight rates while others would rather pay more for faster delivery. Alternatively, a particular customer may want to reduce both transportation cost and time simultaneously [12], which implies that the route optimization of the multimodal transportation should not be treated as a single objective optimization problem but as a multi-objective optimization problem [13,14]. In general, the algorithms for solving multi-objective optimization problems can be classified as traditional optimization algorithms (such as weighting method and constraint method) and intelligent optimization algorithms (such as evolutionary algorithm and particle swarm algorithm), while the weighting method and constraint method are more commonly used in the existing study [15,16].

In practice, transportation operations are negatively influenced by uncertainties that cause ε-vulnerability and reduced serviceability of transportation networks. The impact of uncertain factors on the multimodal transport network is substantial in addition to the consideration of multi-objective issues in the multimodal transport network. Li *et al.* [17] showed that a fixed departure schedule may cause the detention time of cargo in railway and waterway transportation when the cargo arrives at the transshipment nodes [18]. Liu *et al.* [19] further indicated that the timetable restriction may affect the selection of transportation routes and modes. Zhang *et al.* [20] established a multimodal transport route selection model of railway transportation with the consideration of external interference. In a study of the rail-truck intermodal transportation system optimization, Wang *et al.* [21] also introduced uncertain information in the modeling process. Abbassi *et al.* [22] considered using cost uncertainties in their study on the capacities of terminals and transportation costs.

Some studies focus on the uncertainties of multi-objective multimodal transport route optimization in the existing literature. Lu *et al.* [23] studied the multi-mode route problems with uncertainties in time and capacity of intercontinental trains. They solved a fuzzy mixed-integer linear programming model under uncertain conditions using the defuzzification approach. Demir *et al.* [24] considered the design of green intermodal transport service networks with transportation duration and demand uncertainties for intermodal routing decisions. They used a sample average approximation method to generate robust transportation plans. Fazayeli *et al.* [25] used time window constraints and fuzzy numbers to represent the uncertainties of transportation time and demand in their model to study the location routing decision-making issue of the multimodal transport network. Ziaei and Jabbarzadeh [26] proposed a robust counterpart optimization approach to deal with the uncertainties of accident probabilities, emission factors, and costs of establishing transfer points with polyhedral uncertainty sets in their study on the similar issue of location routing decision-making. Monte Carlo simulation is often used to study the transportation time uncertainty in the multimodal transportation problem, but it may significantly increase calculation time [27, 41].

2.2. Algorithms for solving multimodal transportation models

Among the existing approaches of solving multimodal transportation problems, the effective ones are typically meta-heuristic algorithms because of their efficacy in dealing with NP-hard combinatorial optimization problems, for example, multimodal transportation route optimization problems. Meta-heuristic algorithms can be classified into the following four categories: Genetic Algorithm (GA) and Differential Evolution (DE), which simulate the survival of the fittest in nature; Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO), which simulate swarm intelligence to achieve global optimal solutions; Harmony Search (HS) and Firework Algorithm (FA), which are inspired by human behavior; Simulated Particle Algorithm (SPA) and Gravitational Local Search Algorithm (GLSA), which summarize the physical rules and chemical reactions of the universe. Recently, meta-heuristic algorithms have been improved based on different application scenarios. Guo *et al.* [28] proposed a location selection method for emergency rescue facilities from the perspective of global optimization by designing a fast Nondominated Sorting Genetic Algorithm with Cellular Automata (NSGA-II-CA) to solve a multi-objective nonlinear programming model. Cintrano *et al.* [29] solved a bi-objective shortest path problemby using a fast NSGA-II algorithm and a Multi-Objective Evolutionary Algorithm (MOEA). The NSGA-II algorithm with the directional mutation can be used to solve multi-objective optimal path planning problems through adaptive adjustment of crossover probability and mutation probability [30] while improving the convergence rate and avoiding local optimal to some extent. Li *et al.* [31] proposed an effective hybrid Heuristic Kernel Search and Dynamic Programming (HKSDP) for multimodal transportation problems with time windows and network capacity constraints. Peng *et al.* [32] studied the path optimization of multimodal transport networks by using the NSGA-II algorithm that combines the $(\mu+\lambda)$ selection method and the external archiving method to solve a multi-objective robust optimization model that includes transportation time and cost with intervals representing uncertainties.

There are studies that combine different meta-heuristic algorithms, which may significantly improve the computational efficiency, but the quality of the solution sets may not be substantially different from each other. For example, Cheng and Dai [33] proposed a fusion method of Ant Colony Algorithm (ACA) and GA to obtain the optimal solution with fewer iterations with a significant reduction in the calculation time. Among others, the ACA is often selected as the basic algorithm for improvement because of its high solving speed and few iterations. As an example, Zhang *et al.* [34] considered a crossover, recombination, mutation, and other related operations of a GA over ACA to address the lack of Pareto non-dominant solutions generated when solving multi-objective problems. Xie *et al.* [35] proposed a multi-objective optimal routing algorithm for satellite networks based on ACA, which considered the bandwidth as the direction guiding factor during initialization pheromone to reduce suboptimal solutions and speed up the convergence speed of the algorithm.

Nevertheless, the heuristic algorithm is still very computationally expensive in terms of solving efficiency. There are proposed data-driven approaches in the literature that may attempt to help with the inefficiency. Wang *et al.* [36] introduced a data-driven approach to the simulation optimization algorithm that can effectively reduce

the computational complexity. Chao *et al.* [37] proposed data-driven methods to predict energy consumption and optimize modeling using an adaptive neural fuzzy inference system based on an ACA. Gu *et al.* [38] solved the data-driven constrained multi-objective mixed-integer optimization problem with a Random Forest (RF) classifier being a proxy model to approximate targets and constraints. When the objective and constraint can only be calculated based on a large amount of data, Wang and Jin [39] proposed using RFs and radial basis function network as substitutes to approximate objective and constraint functions. They showed that the data-driven approach might be feasible for solving constrained multi-objective combinatorial optimization problems. Lou *et al.* [40] combined Tabu Search Algorithm (TSA) with Support Vector Machine (SVM) and used the data-driven optimization process to increase the algorithm efficiency. Peng *et al.* [41] derived a simple data-driven approach to solving the stochastic multi-objective multimodal transportation routing problem with a schedule. They applied historical data to the iterative process to reduce the calculation times, which overcomes the time-consuming problem of the Monte Carlo simulation.

2.3. Contributions

The main contributions presented in this paper are:

(a) A calculation framework of the multi-objective Nondominated ACA is proposed to address the issue of the long calculation time of the NSGA-II algorithm [32].
(b) An alternative SVM approach is derived based on the existing data pool acceleration approach in the literature [41].
(c) We analyze different data-driven approaches and compare them with the ACA and GA-based approaches.

In contrast to traditional deterministic models, we build a multi-objective optimization model to minimize the total transportation cost and time of a stochastic multimodal transportation network with uncertain transport duration between nodes while considering transport schedules. We implement Monte Carlo simulation to deal with the stochastic transportation time of the multimodal transportation network under the schedule while applying the Pareto theory to handle multiple mutually exclusive optimization objectives. Based on the meta-heuristic algorithm, we develop a data-driven approach by combining the data pool composed of historical data and the data mining of the machine learning algorithm to solve for the nondominated solution. Eventually, we compare the time validity under various data-driven approaches and verify their effectiveness using actual computational cases.

3. Problem description

As our modeling setup is an extension based on two previous papers [32, 41], we adopt similar assumptions of their studies, which are specified as follows:

(1) One cargo is only allowed to be transported as a whole, and no sub-packing of the cargo is allowed during transportation.
(2) The arrival time of cargos is the start time of transshipping.
(3) The latest departure time of the timetable for the selected transport mode is the time that a cargo leaves for the next node after unloading and loading are completed.
(4) Transport and transit facilities have sufficient capacity.

3.1. Index set

The multimodal transportation network is defined as $G = (V, E, M)$, where V is the node set (node s and node d are the starting and ending nodes of cargo transportation respectively, $s, d \in V$), E is the edge set, and M is a collection of transportation method, with $E = \{E_{ij}^m | i, j \in V, m \in M\}$. Among them, uncertain variables are identified by the superscript " $\sim$ ".

3.1.1. Notations

C	The total transportation cost
$\widetilde{T}$	The total transportation time
Q	The total cargo weight
c_{ij}^{m}	The unit cost of a cargo transported from node i to node j through transport mode m
c_{jk}^{mn}	The cost of transporting cargos at node j from transportation mode m to node k through transport mode n
$\widetilde{w_j^m}$	The waiting time for the cargos to leave at node j in transport mode m
$\widetilde{t_{ij}^m}$	The time for the cargos to be transported from node i to node j by transport mode m
t_{jk}^{mn}	The time it takes for the cargos to be transferred from transport mode m to node k through transport mode n at node j
$\widetilde{T_j^A}$	The moment when the cargos arrive at node j
$\widetilde{T_j^L}$	The departure time of the cargos leaving node j
$\widetilde{\delta_j}$	The moment when the cargos is finished loading, unloading, and transshipping at node j
$T_{ij}^{m,r}$	The departure time of the rth shift in the direction of node i to node j for transport mode m
ϖ_{ij}^{m}	The period of the dispatch time of transport mode m in the direction of node i to node j

3.1.2. Decision variables

$$x_{ij}^{m} = \begin{cases} 1, & \text{if the cargos are from node } i \text{ to node } j \text{ choose the} \\ & \text{transportation mode } m \\ 0, & \text{otherwise} \end{cases}$$

$$y_{ik}^{mn} = \begin{cases} 1, & \text{if the cargos are transferred at node } i \text{ form transport mode } m \\ & \text{to node } j \text{ direction transport mode } n \\ 0, & \text{otherwise.} \end{cases}$$

3.2. Multi-objective optimization model of the stochastic multimodal transportation network

In the constructed model, the objective function is to minimize the value of both transportation cost and transportation time, the Z can be expressed as follows:

$$\min Z = \left[C, \widetilde{T}\right]. \tag{1}$$

The C can be formulated as follows:

$$C = Q\left(\sum_{m\in M}\sum_{i\in V}\sum_{j\in V} c_{ij}^{m}x_{ij}^{m} + \sum_{m\in M}\sum_{n\in M}\sum_{i\in V}\sum_{j\in V} c_{ij}^{mn}y_{ij}^{mn}\right). \tag{2}$$

The $\widetilde{T}$ can be expressed as follows:

$$\widetilde{T} = \sum_{m\in M}\sum_{i\in V}\widetilde{w_i^m} + \sum_{m\in M}\sum_{i\in V}\sum_{j\in V}\widetilde{t_{ij}^m}x_{ij}^{m} + \sum_{m\in M}\sum_{n\in M}\sum_{i\in V}\sum_{j\in V} t_{ij}^{mn}y_{ij}^{mn} \tag{3}$$

s.t.

$$\sum_{m\in M}\sum_{j\in V} x_{ij}^{m} = 1, \qquad i = s, \quad i \neq j \tag{4}$$

$$\sum_{m\in M} x_{ij}^{m} \leq 1, \qquad \forall i, \quad j \in V \tag{5}$$

$$\sum_{m\in M}\sum_{n\in M}\sum_{k\in V} y_{jk}^{mn} \le 1, \qquad \forall j \in V/\{s,d\} \tag{6}$$

$$\sum_{j\in V} x_{ij}^{m} + \sum_{j\in V} x_{ji}^{n} \le 2\sum_{j\in V} y_{ij}^{nm}, \qquad \forall i \in V/\{s,d\}, \quad \forall m,n \in M \tag{7}$$

$$\sum_{m\in M}\sum_{i\in V} x_{ij}^{m} = 1, \qquad j = d, \quad i \ne j \tag{8}$$

$$\sum_{m\in M}\sum_{j\in V} x_{ij}^{m} - \sum_{m\in M}\sum_{j\in V} x_{ji}^{m} = 0, \qquad \forall i \in V/\{s,d\}, \quad j \ne i \tag{9}$$

$$\widetilde{T_j^A} = \begin{cases} 0, & j = s \\ \sum_{m\in M}\sum_{i\in V}\left(\widetilde{T_i^L} + \widetilde{t_{ij}^m}\right)x_{ij}^m, & \forall j \in V/\{s\}. \end{cases} \tag{10}$$

Constraint (4) implies that the first node of the path is the starting point; Constraint (5) ensures that the transport mode selected between nodes is unique; Constraint (6) limits node transfer times; Constraint (7) shows that the transport mode exists and continues after transshipment; Constraint (8) ensures that the last node of the path is the endpoint; Constraint (9) represents that the traffic on each node in the path is balanced and continuous; Constraint (10) calculates the time when a node is reached.

$$\widetilde{\delta_j} = \widetilde{T_j^A} + \sum_{m\in M}\sum_{n\in M}\sum_{k\in V} t_{jk}^{mn} y_{jk}^{mn}, \qquad \forall j \in V/\{d\}. \tag{11}$$

Constraint (11) represents the time when the transport is complete on the node.

$$\widetilde{w_j^m} = \begin{cases} \sum_{m\in M}\sum_{k\in V}\left[T_j^{m,1} - \mathrm{mod}\left(\widetilde{\delta_j}, \varpi_{jk}^m\right)\right]x_{jk}^m, & \mathrm{mod}\left(\widetilde{\delta_j}, \varpi_{jk}^m\right) \le T_j^{m,1} \\ \sum_{m\in M}\sum_{k\in V}\left[T_j^{m,2} - \mathrm{mod}\left(\widetilde{\delta_j}, \varpi_{jk}^m\right)\right]x_{jk}^m, & T_j^{m,1} < \mathrm{mod}\left(\widetilde{\delta_j}, \varpi_{jk}^m\right) \le T_j^{m,2} \\ \cdots\cdots & \\ \sum_{m\in M}\sum_{k\in V}\left[T_j^{m,r} - \mathrm{mod}\left(\widetilde{\delta_j}, \varpi_{jk}^m\right)\right]x_{jk}^m, & T_j^{m,r-1} < \mathrm{mod}\left(\widetilde{\delta_j}, \varpi_{jk}^m\right) \le T_j^{m,r} \\ \sum_{m\in M}\sum_{k\in V}\left[T_j^{m,1} + \varpi_{jk}^m - \mathrm{mod}\left(\widetilde{\delta_j}, \varpi_{jk}^m\right)\right]x_{jk}^m, & \mathrm{mod}\left(\widetilde{\delta_j}, \varpi_{jk}^m\right) \ge T_j^{m,r} \end{cases} \quad \begin{matrix} \delta_j \ne 0, \\ \forall j \in V/\{d\}. \end{matrix} \tag{12}$$

Constraint (12) indicates the time to wait for the transport mode to depart at node j.

$$\widetilde{T_j^L} = \widetilde{\delta_j} + \widetilde{w_j^m}, \qquad \forall j \in V, \quad \forall m \in M. \tag{13}$$

Constraint (13) represents the time to leave the node.

4. Data-driven simheuristic approach

4.1. The design method

The multimodal transportation routing problem is NP-hard, and its solution set space is discontinuous. The ACA is widely used to solve this type of problem [43] because it can deal with the discontinuous solution sets with reasonable solving time. We apply the Pareto optimization theory to take into account of the diversity and mutual exclusion of optimization objectives. A meta-heuristic algorithm combined with Monte Carlo simulation has been proved to be an effective way to solve stochastic combinatorial optimization problems for the transportation network with highly uncertain transportation time [41, 42]. However, the MSAC algorithm is very computation time consuming, and the simulation also leads to a surge in computing time. To address the time consuming issue associated with the algorithms, we derive a data-driven approach by fully mining the available information behind the acquired data. The data-driven approach can reduce the computing time that is needed

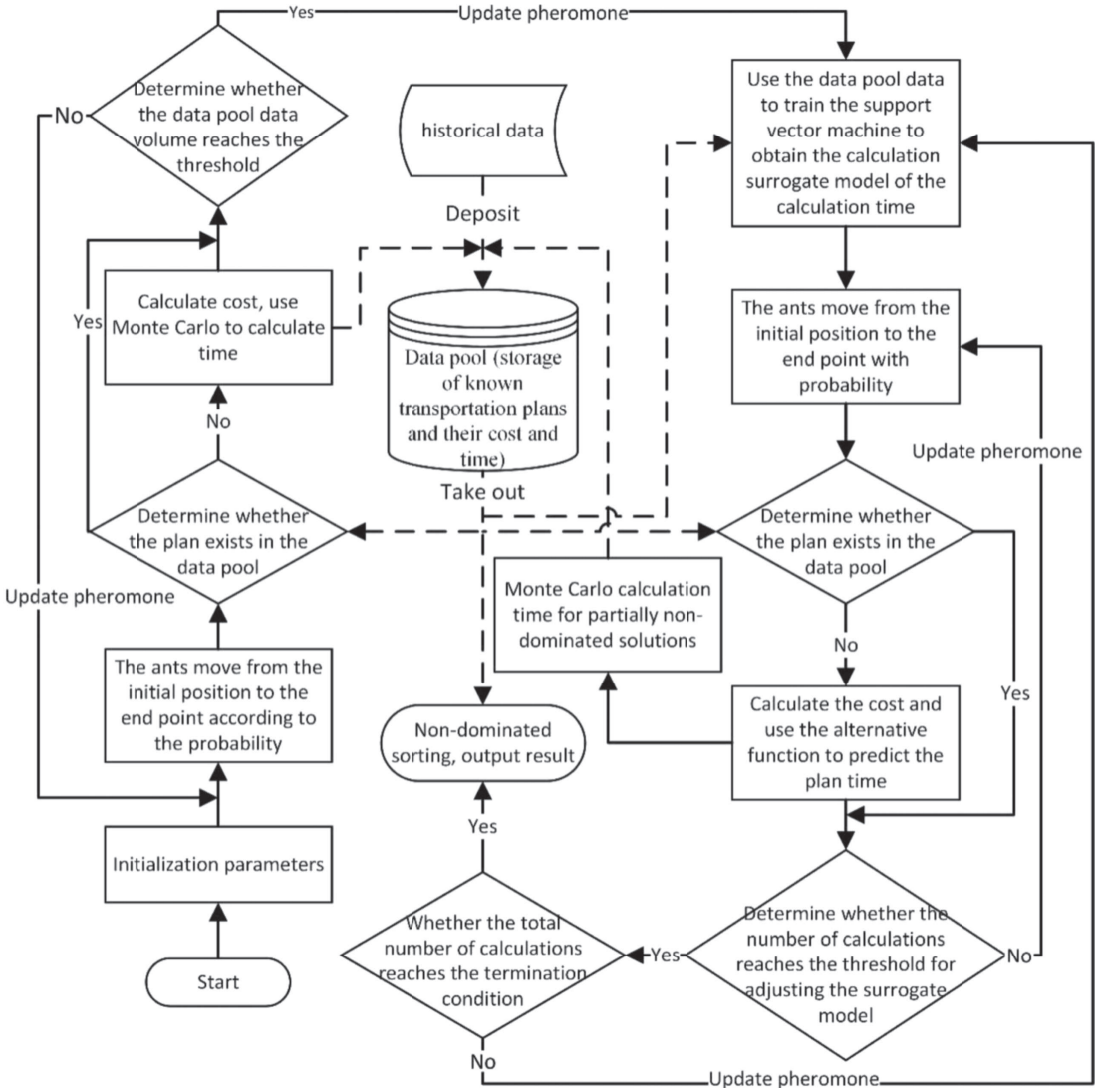

FIGURE 1. DD-MSAC algorithm flow chart.

to evaluate fitness values while improving the iteration efficiency due to the reduced number of evaluations. Based on the data-driven approach combined with the simulation, we design a Data-driven Multi-objective Ant Colony (DD-MSAC) algorithm. The flow chart of the algorithm is shown in Figure 1.

The algorithm steps are specified in the following:

(1) Parameter initialization

Initialize the ACA parameters, all ants in the starting point.

(2) Each ant moves to find its way

Each ant starts from the starting point, determines the feasible path for the next step and saves it in allowed_k. If allowed_k is empty, the ant returns to the starting point and restarts the path selection. If allowed_k is not

empty, the transfer probability p_{ij}^k is determined by formula (14), then the ant selects the next node. The ant completes the path until it reaches the end [43].

$$p_{ij}^k = \begin{cases} \frac{[\tau_{ij}]^\alpha * [\eta_{ij}]^\beta}{\sum_{j \in \text{allowed}_k} [\tau_{ij}]^\alpha * [\eta_{ij}]^\beta} & j \in \text{allowed}_k \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

In (14), τ_{ij} and α are the pheromone concentrations and heuristic factors between nodes i and j respectively. η_{ij} and β represent the visibility between nodes i and j, and their heuristic factors. The path visibility is $\eta_{ij} = 1/d_{ij}$, where d_{ij} represents the physical distance between node i and node j. allowed_k is the next set of paths the ant can take.

(3) Objective function adaptation calculation

After all ants arrive at the destination, the adaptive value of the objective function of the evaluation scheme is the total transportation cost C and total transportation time $\widetilde{T}$. The time $\widetilde{T}$ can be calculated by Monte Carlo simulation and Support Vector Regression (SVR) fitting function. In the initial stage, the total time $\widetilde{T}$ is mainly calculated by the simulation. When the data in the data pool reaches the set threshold, the SVR fitting alternative model is selected for calculation.

(4) Data archiving and SVR model training

We update the data in the data pool and use the updated data to train the SVR regression model. The first step is to construct an SVR feature vector matrix (*i.e.*, transportation path scheme), whose length is determined by the scale of the transportation network. The column number corresponds to the node number, and different values represent different transport modes. The disconnected nodes in the path are added to the matrix with zero completions. Secondly, we need to train surrogate models $y = f(x)$. The step is to enter a set of training set data in the data pool $\{(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\}$, where $\{x_i\}$ is the transportation route scheme and $\{y_i\}$ corresponds to transportation time T. The output is the surrogate model $y = f(x)$.

(5) Solve the nondominated solution

A fast nondominated sorting algorithm is used to solve the nondominated solutionof all feasible solutions up to this iteration [44]. In multi-objective programming, if all the targets of solution x_A are better than the targets of solution x_B, then x_A dominates x_B. If no solution governs x_A, x_A is nondominated solution. Multi-objective ACA is combined with the nondominated solution to update pheromone, and the path pheromone is determined by its retained information and the ant pheromone corresponding to the nondominated solution.

(6) Pheromone renewal

Pheromone τ_{ij} was updated between nodes i and j in nondominated solution path, and the update strategy is the following.

$$\tau_{ij}(u+1) = (1-\rho) * \tau_{ij}(u) + \Delta\tau_{ij}(u+1) \tag{15}$$

$$\Delta\tau_{ij}(u+1) = \frac{\sum_{k=1}^{p} \vartheta_{ij}^k(u+1)}{\sum_{k=1}^{p} \psi_k(u+1)}. \tag{16}$$

In particular

$$\psi_k(u+1) = \frac{1}{2}\left(1 - \frac{c_k(u+1)}{\max_\phi c_k(u+1)}\right) + \frac{1}{2}\left(1 - \frac{t_k(u+1)}{\max_\iota t_k(u+1)}\right) \tag{17}$$

$$\vartheta_{ij}^k(u+1) = \begin{cases} \psi_k(u+1) & \text{If the } k\text{th ant passes } ij \\ 0 & \text{else} \end{cases} \tag{18}$$

where c and t represent the time and cost in the nondominated solution of the Kth iteration. ϕ and ι represent the time and cost solution sets of the Kth iteration respectively. p represents the number of nondominated solutions that the Kth iteration belongs to the nondominated solution set. $\tau_{ij}(u)$ represents the pheromone of the $(k-1)$th path; $\tau_{ij}(u+1)$ is the updated pheromone value; $\Delta\tau_{ij}(u+1)$ represents pheromone increment; ρ is pheromone volatilization coefficient. $\vartheta_{ij}^{k}(u+1)$ represents the pheromone that ants pass through ij, $\psi_k(u+1)$ represents the pheromone of each iteration of the Kth ant.

(7) SVM regression model was updated

Extracting the path scheme partly calculated by surrogate model $y = f(x)$, simulation is used to calculate the transportation time $\widetilde{T}$ instead of the original scheme. The following is to update the data pool data on time and train the SVR with the updated data pool data to update the surrogate model.

4.2. Data-driven approach

(1) The data pool

In addition to the computational framework of the ant colony algorithm, a data pool is designed to read and store optimization objectives. The data pool is filled with offline historical data and an online heuristic algorithm to avoid repeated simulation and information redundancy. In addition, a machine learning algorithm is used for data mining in the data pool to find the surrogate model of optimization objective function to make the best use of data information.

(2) Support Vector Machine (SVM)

SVM [45] is a supervised machine learning algorithm based on statistical learning theory with excellent generalization ability. SVR is a generalization of SVM, which is widely used in the regression of traffic problems.

The objective of the SVR is to accurately fit the regression function based on the given training data set $\{(X_1, T_1), (X_2, T_2), \cdots, (X_N, T_N)\}$ to accurately predict the corresponding target $\{T_i\}$ of a set of inputs $\{X_i\}$, where $X_i \in R^n$ is the input and $T_i \in R$ is the output.

For nonlinear problems, the nonlinear mapping $g_i(x)$ can be used to map samples to a high-dimensional feature space to build a linear model:

$$y = f(x) = \omega^T g_i(x) + b. \tag{19}$$

To minimize the true error of equation (19), the minimum value of the following structural risk function is required:

$$R(\omega, b) = \frac{1}{2}\|\omega\|^2 + C\frac{1}{N}\sum_{i=1}^{N} L_\varepsilon(T_i - f(x_i), x_i). \tag{20}$$

The structural risk function is the summation of the two terms. $\frac{1}{2}\|\omega\|^2$ is the function smoothness related to model complexity and $C\frac{1}{N}\sum_{i=1}^{N} L_\varepsilon(T_i - f(x_i), x_i)$ is an empirical error function. The penalty coefficient δ is used to control the weight of two items in the structural risk function. $L_\varepsilon(T_i - f(x_i), x_i)$ is the ε insensitive loss function, which is defined in the following.

$$L_\varepsilon(y_i - f(x_i), x_i) = |T - f(x)|_\varepsilon = \begin{cases} 0, & |T - f(x)| \le \varepsilon \\ |T - f(x)| - \varepsilon, & |T - f(x)| > \varepsilon. \end{cases} \tag{21}$$

To solve the problem that some outliers deviate from ε insensitive region, ξ_i, ξ_i^* relaxation variables are introduced to transform the SVM regression problem into the following optimization problem.

$$\min\left[\frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*)\right] \tag{22}$$

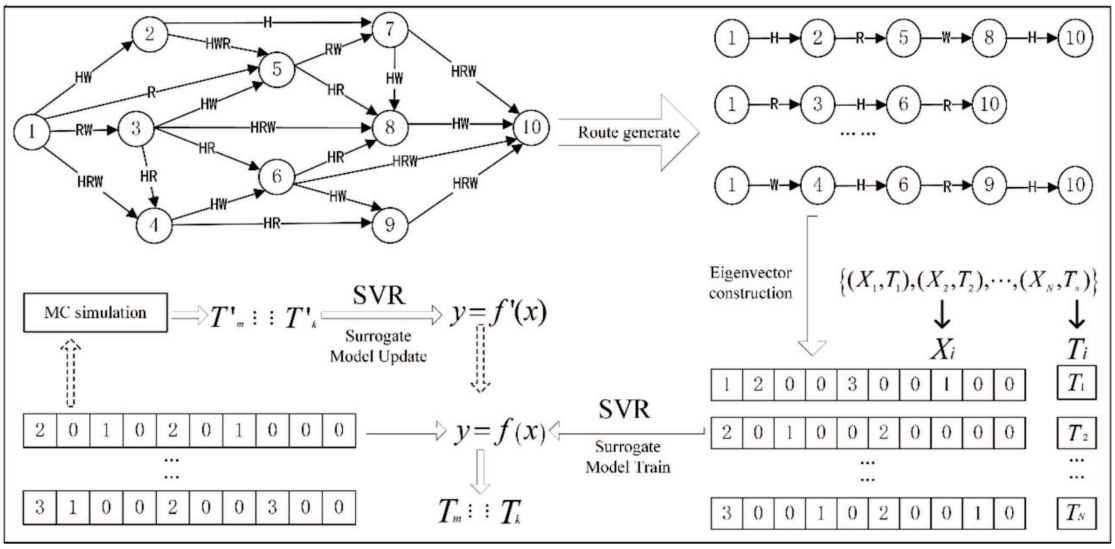

FIGURE 2. Data archiving, SVR regression model training and update.

s.t.

$$f(x_i) - T_i \le \varepsilon + \xi_i \tag{23}$$

$$T_i - f(x_i) \le \varepsilon + \xi_i^* \tag{24}$$

$$\xi_i \ge 0, \quad \xi_i^* \ge 0, \qquad i = 1, 2, \cdots, N. \tag{25}$$

To solve the above optimization problem, the optimal solution of the original problem is obtained by solving its Lagrangian dual problem:

$$f(x) = \sum_{i=1}^{N} (\alpha_i^* - \alpha_i) K(x_i, x) + b. \tag{26}$$

α_i and α_i^* are the Lagrangian multipliers. $K(x_i, x)$ is a kernel function that meets Mercer's conditions and is used to solve the problem of "dimension disaster" in the operation of high-dimensional feature space. The radial basis kernel function is widely used among many kernel functions because of its good performance [46]. This paper uses the radial basis kernel function to construct SVR, which is given in the following.

$$K(x_i, x) = \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right). \tag{27}$$

In this section, the penalty coefficient C and radial basis kernel function parameter σ to be determined are optimized through the grid search developed by Ahmad *et al.* [46].

Data archiving, SVR regression model training and update process are shown in Figure 2 (the illustrated numerical example is solved by using the statistics and machine learning toolbox of the MATLAB with default parameters).

5. Numerical experiments and analysis

5.1. Experimental data and setting

The experiment platform is a notebook computer with Intel Core i5 CPU and 16 GB memory, and the programming software is MATLAB 2014b. To the best of our knowledge, there is no standard benchmark dataset

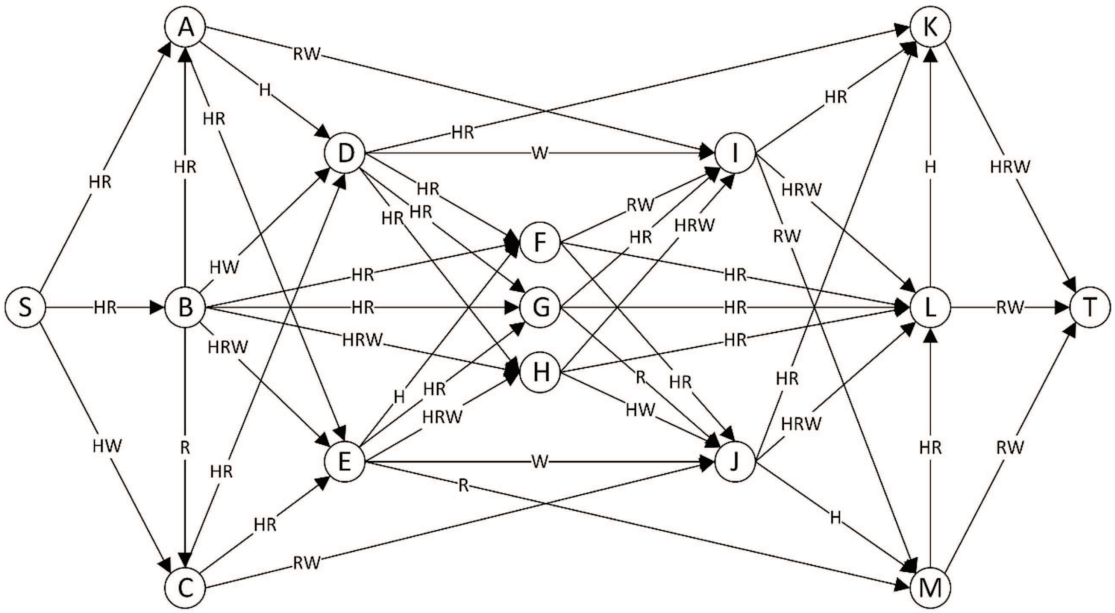

FIGURE 3. Multimodal transportation network.

TABLE 1. Parameters of transfer.

	Unit transportation time between different transport modes (h)			Unit transportation cost between different transport modes (RMB)		
	Highway	Railway	Waterway	Highway	Railway	Waterway
Highway	1	1.5	2	30	30	40
Railway	1.5	2	4	30	40	50
Waterway	2	4	3	40	50	50

for the extent of the stochastic optimization problem of this study as of now. Accordingly, three multimodal transportation networks are randomly constructed (H: highway; R: railway; W: waterway), including a small (15 nodes), a medium (30 nodes), and a large (50 nodes) scale network. Figure 3 is a multimodal transportation network consisting of 15 nodes. It is assumed that the random transportation duration of each road section follows a normal distribution. Parameter of the transportation network are provided in Table 1. The transfer data of each transportation mode are shown in Table 2 and the timetables of railway and waterway for each node are assumed to be the same. The time when the cargos are shipped from the starting point is set to be 7:30 am.

5.2. Comparison of algorithms MSAC and NSGA-II

5.2.1. Algorithm evaluation indicator

Since this paper compares the algorithms developed by Peng *et al.* [41], the measurement standards are the same, and the specific contents are as follows:

We employ three performance metrics to evaluate the performance of the proposed algorithm. "NNS", "DUN" and "ET" represent the "Number of nondominated solution", "Distribution uniformity of nondominated solution" and "Every running time" respectively. The function $F(A)$ as defined in equation (29) represents the

TABLE 2. Multimodal transportation network data.

Point pair	Highway data	Railway data	Waterway data	Point pair	Highway data	Railway data	Waterway data
SA	[26, 7], 6240	[41, 3], 9020	∼	EH	[31, 5], 7440	[43, 3], 9460	[102, 7], 3927
SB	[26, 7], 6240	[41, 3], 9020	∼	EJ	∼	∼	[97, 8], 3734.5
SC	[29, 4], 6960	∼	[71, 7], 2733.5	EM	∼	[32, 2], 7040	∼
AD	[47, 6], 11280	∼	∼	FI	∼	[42, 3], 9240	[102, 9], 3927
AE	[25, 7], 6000	[40, 3], 8800	∼	FJ	[26, 7], 6240	[41, 3], 9020	∼
AI	∼	[42, 2], 9240	[102, 9], 3927	FL	[25, 7], 6000	[41, 3], 9020	∼
BA	[26, 7], 6240	[41, 3], 9020	∼	GI	[25, 7], 6000	[40, 3], 8800	∼
BC	∼	[32, 2], 7040	∼	GJ	∼	[32, 2], 7040	∼
BD	[29, 4], 6960	∼	[70, 7], 2695	GL	[25, 7], 6000	[41, 3], 9020	∼
BE	[31, 5], 7440	[43, 3], 9460	[101, 7], 3888.5	HI	[30, 5], 7200	[43, 3], 9460	[102, 7], 3927
BF	[26, 7], 6240	[42, 3], 9240	∼	HJ	[29, 5], 6960	∼	[70, 7], 2695
BG	[26, 7], 6240	[41, 4], 9020	∼	HL	[26, 7], 6240	[41, 3], 9020	∼
BH	[30, 5], 7200	[43, 3], 9460	[102, 7], 3927	IK	[26, 7], 6240	[41, 3], 9020	∼
CD	[26, 7]; 6240	[41, 3]; 9020	∼	IL	[30, 6], 7200	[43, 3], 9460	[102, 7], 3927
CE	[26, 7], 6240	[41, 3], 9020	∼	IM	∼	[42, 3], 9240	[101, 9], 3888.5
CJ	∼	[42, 3], 9240	[102, 9], 3927	JK	[26, 7], 6240	[41, 3], 9020	∼
DF	[26, 7], 6240	[42, 3], 9240	∼	JL	[31, 6], 7440	[43, 3], 9460	[102, 7], 3927
DG	[25, 7], 6000	[41, 3], 9020	∼	JM	[47, 6], 11280	∼	∼
DH	[26, 7], 6240	[41, 3], 9020	∼	KT	[31, 5], 7440	[43, 3], 9460	[101, 7], 3888.5
DI	∼	∼	[96, 8], 3696	LK	[46, 6], 11040	∼	∼
DK	[26, 7], 6240	[41, 3], 9020	∼	LT	∼	[42, 3], 9240	[102, 9], 3927
EF	[47, 6], 11280	∼	∼	ML	[26, 7], 6240	[40, 3], 8800	∼
EG	[26, 7], 6240	[40, 3], 8800	∼	MT	∼	[42, 3], 9240	[102, 9], 3927

Notes. (1) data format: Pont pair; Highway data; Railway data; waterway data; (2) the data format of Highway data: [Mean transit time, standard deviation (h)], cost (RMB/T). The other two groups have the same format; (3) Data is empty: ∼. Taking any two nodes "FL" in Figure 4 as an example "FL; [25, 7], 6000; [41, 3], 9020; ∼" means there are two transport modes from F node to L node including highway and railway, and ∼ means there is no waterway. "FL" is a point pair and represents the section name.

distribution degree of solutions in the nondominated solution" set A, where:

$$d_i^* = \min_{j \in A \setminus \{j \neq i\}} \sum_{x=1}^{l} \left| f_x^i - f_x^j \right|, \bar{d} = \left(\sum_{i=1}^{\|A\|} d_i^* / \|A\| \right), \quad i, j \in A \tag{28}$$

and the number of the objective functions is l. The function value of $F(A)$ decreases as the solutions in set A are distributed more uniformly. If the solutions are equally spaced in the objective function value space, the value of $F(A)$ is zero.

$$F(A) = \sqrt{\frac{1}{\|A\| - 1} \sum_{i=1}^{\|A\|} \left(d_i^* - \bar{d} \right)^2}. \tag{29}$$

The convergence of the solution set can be explained as follows. Let u and v denote two solutions in the nondominated solution set. The dominance relation ship between u and v. Are defined in formula (30). The function $S(A, B)$ in formula (31) calculates the number of times in which the solution in A dominates that in B. Formula (32) provides the relative convergence rate of solution set A to B, which is proposed by Zitzler [47].

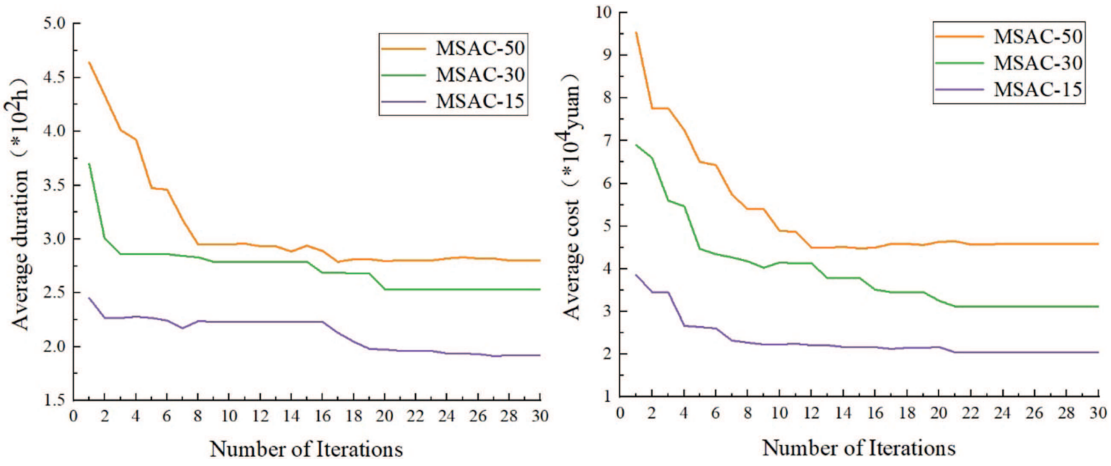

FIGURE 4. The change of the mean value of the target value under three network sizes.

When A has better convergence than B, we have $r(A,B) > r(B,A)$.

$$S_{u,v} = \begin{cases} 1, & \text{if } u \text{ dominates } v \\ 0, & \text{if } v \text{ dominates } u \end{cases} \tag{30}$$

$$S(A,B) = \sum_{u \in A} \sum_{v \in B} S(u,v) \tag{31}$$

$$r(A,B) = \frac{\sum_{u \in A} \sum_{v \in B} S(u,v)}{\sum_{u \in A} \sum_{v \in B} S(u,v) + \sum_{u \in B} \sum_{v \in A} S(u,v)}. \tag{32}$$

As an extension of the classical shortest path problem, various heuristic algorithms have proved feasible and effective in solving multimodal transportation routing optimization. For example, Peng *et al.* [41]. adopted a fast NSGA-II to solve the multimodal transportation route optimization problem by considering the timetable limitation under uncertain environments. These algorithms have advantages in optimum-seeking performances and solving time consumption by using different optimization techniques. To validate the effectiveness of the proposed DD-MSAC algorithm, we first compare the optimum-seeking performances and solving time consumption of the NSGA-II designed by Peng *et al.* [41] and MSAC. In the NSGA-II, the population parameter setting, the chromosome coding, the crossover operator and the mutation operator, and other steps are adapted from Peng *et al.* [40, 41]. In the MSAC, the algorithm parameters are set to be the following. The number of ants is 20, the number of iterations is 30, and the trials per simulation are 500, particularly $\beta = 0.2, \rho = 0.3$, and $\alpha = 0.8$. .

Figure 4 depicts the change in the average value of each objective value of nondominated solutions obtained by MASC algorithm during each iteration under three network scales. The two objective function values evolve to the smallest direction during iteration, which preliminarily verifies the validity of the designed algorithm under this set of algorithm parameters.

To reduce the randomness of iterations in the two optimization processes, we repeatedly apply the two algorithms 30 times to obtain the statistical results of each objective function in each experiment. Table 3 shows the comparison results based on the solutions in three network scales generated by the two Pareto optimization approaches. Columns of "Obj.1" and "Obj.2" represent the "Total transportation cost" and the "Total transportation time," respectively. Rows of "Max", "Min" include the maximum, the minimum of the various optimization results obtained, respectively. According to Table 3, the distribution ranges of the maximum

TABLE 3. The comparison values of the two algorithms under different network scales.

Network size		MSAC		NSGA-II	
		Obj.1 (RMB)	Obj.2 (h)	Obj.1 (RMB)	Obj.2 (h)
15	Max	27 232.0	363.8	26 970.0	395.7
	Min	15 080.4	113.0	15 463.7	112.3
30	Max	41 391.3	390.9	37 749.7	425.9
	Min	22 121.6	173.1	21 837.5	158.0
50	Max	59 845.4	458.1	49 030.0	490.5
	Min	36 845.9	206.5	33 496.7	214.9

TABLE 4. The comparison values of the two algorithms under different network scales.

Network size	MSAC			NSGA-II		
	NNS	DUN	ET	NNS	DUN	ET
15	5.9	1938.5	33.7	5.3	897.7	106.6
30	5.9	2677.9	95.7	6.4	2145.4	276.6
50	15.2	1480.8	203.7	6.7	2290.1	372.2

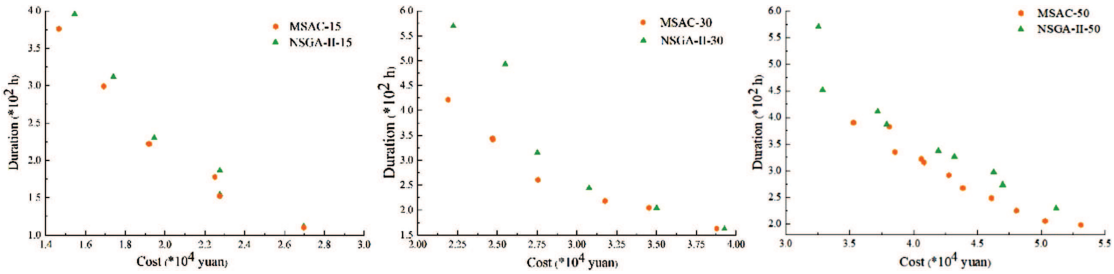

FIGURE 5. Distribution of nondominated solution sets.

and minimum objective function values solved by the two algorithms are not very different, but they reflect boundaries of the nondominated solutions obtained by the two algorithms.

As we can see from Table 4, the optimization performance of the two algorithms is close to each other, and both perform well. The metrics of NNS and DUN represent the diversity and uniformity of the nondominated solution set. The larger NNS value, the better diversity of the solution set. The smaller DUN value, the more uniform distribution of the solution set. ET is the average time calculated by the algorithm. The smaller ET is, the better time the algorithm's saving performance is. Among the three groups of experiments, The diversity of nondominated solution sets is not different between small and medium networks. However, the diversity of nondominated solution sets obtained by MSAC is better in large networks. The solution set obtained by NSGA-II has a lower degree of distribution and is more uniformly distributed.

By analyzing the computing time of the algorithms, we find that the MSAC is superior to the NSGA-II. The computing time of the MSAC approach is reduced to 3.16 times, 2.94 times, and 1.83 times in Experiments 1–3, respectively.

In the experiments, NSGA-II and MSAC are used to solve the distribution of the nondominated frontier solution set, as shown in Figure 5 (which illustrates the nondominated solution set of one experiment with a network size of 15, 30, and 50, respectively). It is clear that the MSAC algorithm generates better solutions

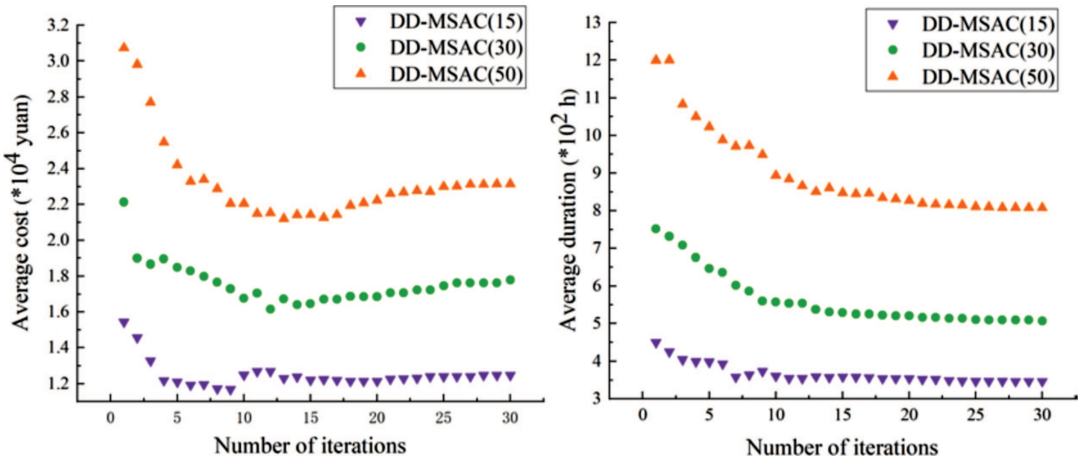

FIGURE 6. The changes of average time and cost with the number of iterations.

because the solution set extends the furthest towards the ends of the nondominated solution, which indicates that the MSAC approach has the capability of finding better solutions than the NSGA-II.

5.3. Algorithm analysis

The comparison results in the previous section preliminarily prove the validity of the MSAC algorithm. We introduce the MSAC algorithm based on the data-driven approach and derive the data-drive MSAC algorithm (DD-MSAC) accordingly. In order to further explore the algorithm modules of the DD-MSAC and the validity of the algorithm, we expand the algorithm by combining different algorithm modules.

According to different approaches, we derive the following four algorithm module combinations: DD-MSAC, the MSAC based on the data-pool approach (DP-MSAC), the MSAC based on the SVM (SVM-MSAC), and MSAC itself.

5.3.1. Algorithm validity

We use the same ACA parameters to run DD-MSAC and MSAC 30 times. The iterative average cost and the average duration of the nondominated solutions obtained by the DD-MSAC are shown in Figure 6. The quality of the algorithm is improved with the number of iterations, which shows the validity of the algorithm design.

Table 5 shows the error between the scheme target value obtained after 30 DD-MSAC runs and the MSAC scheme target value. The average error of all schemes is less than 5%, which shows the effectiveness of the data-driven approach. The reason for the scheme error fluctuation is speculated to be that most of the experimental SVM training sets are accumulated iteratively and randomly, and the quality fluctuation of the training set is transmitted to the scheme target value.

5.3.2. Comparison of different approaches

We keep the ACA parameters unchanged to run each of the algorithms DD-MSAC, DP-MSAC, SVM-MSAC, and MSAC 30 times. The results are shown in Table 6.

The average number and the average distribution of nondominated solution obtained by running different strategies 30 times under the three different network scales are shown in Table 6. By analyzing the network scales with 15, 30, and 50 nodes respectively, the fluctuation of NNS and DUN index values obtained by each algorithm is within 5%, which indicates that different approaches have little impact on the quality of solutions.

Figure 7 shows the bar chars of the coverage values under the three network scales. Along the horizontal axis, numbers 1 to 4 represent DD-MSAC, SVM-MSAC, DP-MSAC, and MSAC, respectively. Each bar plot

TABLE 5. Experimental scheme error.

Network size	Scheme number	Frequency	Error (%)	Scheme number	Frequency	Error (%)
15	1	13	0.013	6	9	1.137
	2	12	0.919	7	9	0.220
	3	12	0.866	8	9	0.022
	4	11	0.702	9	8	0.841
	5	10	0.345	10	8	0.485
30	1	11	0.095	6	9	0.198
	2	11	0.560	7	8	2.109
	3	10	1.218	8	7	0.488
	4	9	0.184	9	7	0.262
	5	9	0.723	10	7	0.637
50	1	9	0.021	6	7	4.984
	2	9	0.053	7	6	0.357
	3	8	0.436	8	6	0.296
	4	8	0.646	9	6	0.365
	5	7	3.612	10	6	1.501

TABLE 6. NNS and DUN values of different algorithms in different network sizes.

Network size	DD-MSAC		SVM-MSAC		DP-MSAC		MSAC	
	NNS	DUN	NNS	DUN	NNS	DUN	NNS	DUN
15	6.1	2370.2	6.3	2279.1	6	2062.9	5.9	2062.9
30	5.6	2683.7	6.3	2643.6	6.2	2698.4	5.9	2698.4
50	12.6	1692.7	12.3	1687.9	15.2	1668.8	15.7	1668.8

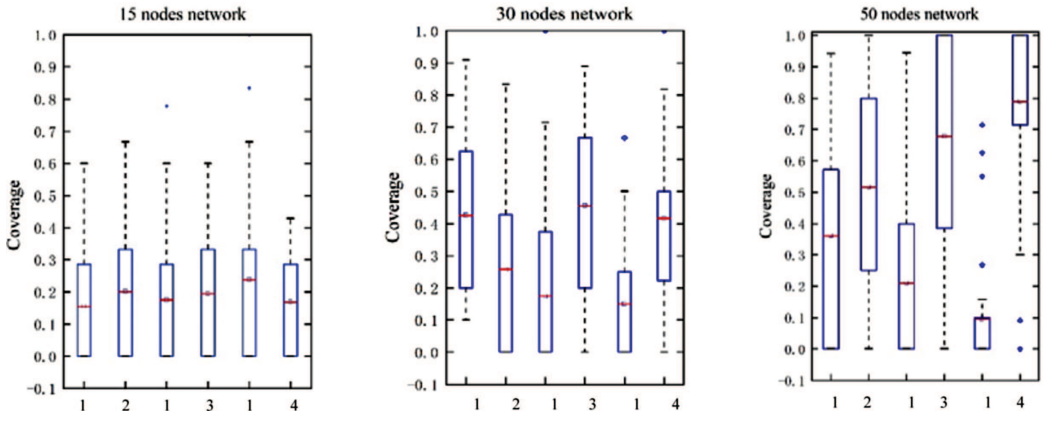

FIGURE 7. Box plots of coverage values for three network scales.

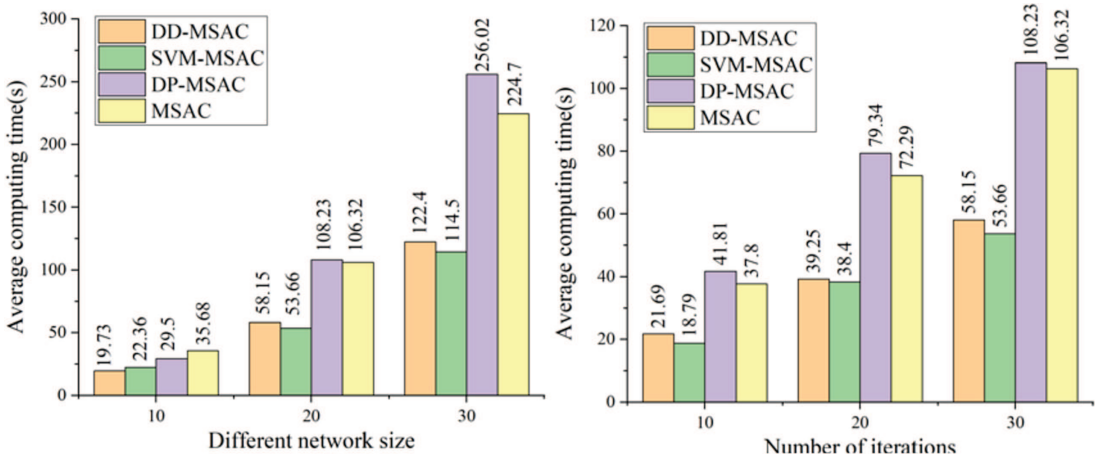

FIGURE 8. Average calculation time.

represents the distribution of the coverage values achieved in 30 repetitive comparisons. We use the experiment with the network of 15 nodes as an example to compare the pairwise strategies by showing the first, the second (the median), and third quartiles values of the coverage metric. With the comparison between the DD-MSAC and the SVM-MSAC, the three values of the DD-MSAC are 0, 0.16, and 0.29, while the three values of the SVM-MSAC are 0, 0.2, and 0.33, respectively. When comparing the DD-MSAC and the DP-MSAC, the three values of the DD-MSAC are 0, 0.21, and 0.29, while the three values of the DP-MSAC are 0, 0.21, and 0.33, respectively. By comparing the DD-MSAC and the MSAC, the three values of the coverage values of the DD-MSAC are 0, 0.26, and 0.34, while the three values of the MSAC are 0, 0.19, and 0.30, respectively. There is more than 20% (21% and 19%) of the solutions generated by the SVM-MSAC (DP-MSAC and MSAC) are dominated by (or equal to) the solutions generated by the DD-MSAC in more than 50% of the 30 comparisons, and about 29% (29% and 30%) of the solutions generated by the DD-MSAC are dominated by (or equal to) the solutions generated by the SVM-MSAC (DP-MSAC and MSAC) in more than 75% of the 30 comparisons. We find that the nondominated solution provided by the above four approaches in terms of the coverage metric are not significantly different from each other. Other network scale experiments exhibit similar results, which implies that the acceleration strategy had no significant impact on the quality of the solutions.

We further investigate the every running time (ET) of each strategy. Under the network scale of 15 nodes, the ET with the data-driven approach is less than those without any. The DD-MSAC has the most acceleration effect, saving 45% of the average computing time. As the network scale increases, the acceleration effect of DP-MSAC decreases. The reason may be that the increase of the network scale and complexity results in the increase of the amount of data in the data pool, but the number of repeated paths is small, which leads to the increase of computing time. The SVM-MSAC has the most stable acceleration effect, saving 37%, 50%, and 55% of the average computing time under the three network scales, respectively.

With a network of 30 nodes, we run each algorithm 10, 20, and 30 times respectively. The average calculation time is shown in Figure 8. The SVM-MSAC has the least average computing time, saving 51%, 47%, and 50%, respectively, compared with the MSAC. In summary, the DD-MSAC and the DP-MSAC have obvious advantages in computing time with small-scale networks. The SVM-MSAC has the most stable acceleration effect, with an average optimization range of about 50%.

5.4. A case analysis

China's "Belt and Road Initiative" is the abbreviation for the Silk Road Economic Belt and the 21st Century Maritime Silk Road, connecting 138 countries. Countries along the routes cooperate in the capital, production

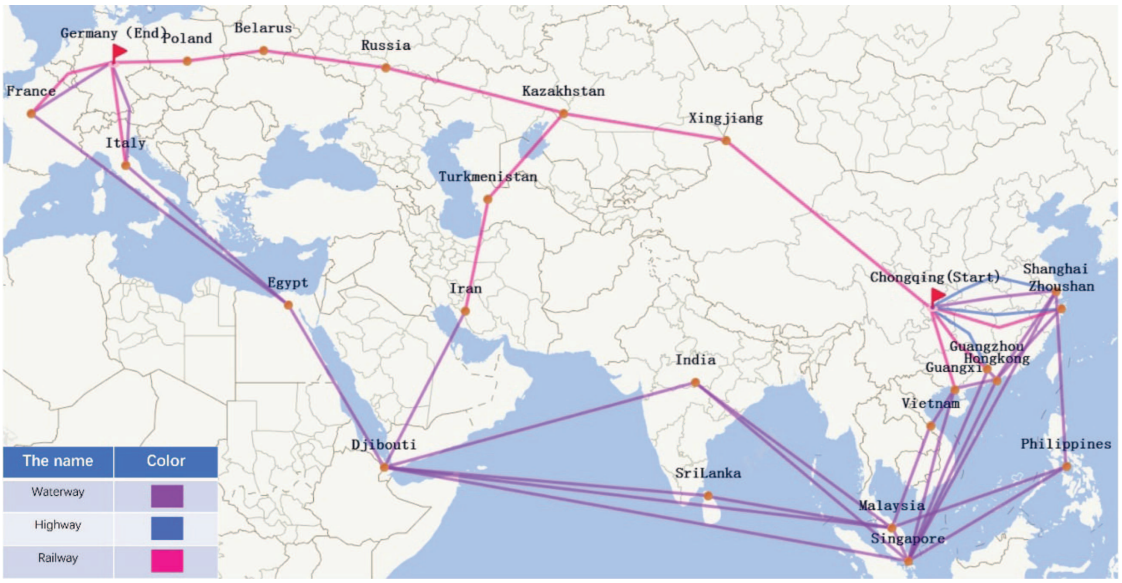

FIGURE 9. Belt and road initiative.

TABLE 7. Partial non-dominant transport scheme.

The specific path	Total cost (RMB)	Total time ($\widetilde{T}$)
1. Chongqing $\xrightarrow{W}$ Shanghai $\xrightarrow{W}$ Singapore $\xrightarrow{W}$ Srilanka $\xrightarrow{W}$ Djibouti $\xrightarrow{W}$ Egypt $\xrightarrow{W}$ France $\xrightarrow{W}$ Germany	28 787	44.95
2. Chongqing $\xrightarrow{W}$ Shanghai $\xrightarrow{W}$ Singapore $\xrightarrow{W}$ India $\xrightarrow{W}$ Djibouti $\xrightarrow{W}$ Egypt $\xrightarrow{W}$ Italian $\xrightarrow{R}$ Germany	30 338	41.59
3. Chongqing $\xrightarrow{R}$ Guangzhou $\xrightarrow{W}$ Singapore $\xrightarrow{W}$ Srilanka $\xrightarrow{W}$ Djibouti $\xrightarrow{W}$ Egypt $\xrightarrow{W}$ Italian $\xrightarrow{R}$ Germany	31 220	35.00
4. Chongqing $\xrightarrow{R}$ Guangxi $\xrightarrow{W}$ Singapore $\xrightarrow{W}$ India $\xrightarrow{W}$ Djibouti $\xrightarrow{W}$ Egypt $\xrightarrow{W}$ Italian $\xrightarrow{W}$ Germany	33 423	31.00
5. Chongqing $\xrightarrow{R}$ Xingjiang $\xrightarrow{R}$ Kazakhstan $\xrightarrow{R}$ Russia $\xrightarrow{R}$ Belarus $\xrightarrow{R}$ Poland $\xrightarrow{R}$ Germany	37 044	10.42

capacity, and technology, and the total trade volume has been increasing year after year, attracting attention from the international community [47]. In order to reduce the freight and transportation time of the cargos in transit, the path selection of the transport in the Belt and Road Initiative is used in this paper. By taking the path from Chongqing to Germany as an example, we choose 24 cities, three kinds of transportation cost and time to be included. The Belt and Road Roadmap of our example is shown in Figure 9.

Our results (Tab. 7) show that sea (water) transportation takes more time but costs less, while railway transportation is more flexible and faster but costs more. The critical impact factor of the route decision is the customer's sensitivity to the transportation cost and time. When transporting high value-added and time-sensitive cargos, customers are more sensitive to the total transportation time. Therefore, the route decision can include the rail-road combined transportation to take advantage of their flexible and fast characteristics to complete the transportation as soon as possible. On the other hand, one can choose sea transportation with a prolonged transportation cycle to save costs.

6. Conclusion

Adding the uncertainty of transportation time to the study of multimodal transportation route optimization may increase the applicability of the resulting selected transport mode schedule as well as the complexity of the optimization model. In our proposed multi-objective multimodal transportation route optimization model to determine the selected transport mode, we assume that the transportation time is a random variable that follows a specific probability distribution. According to our comparison study on the solution quality, we find that the MSAC algorithm performs better than the NSGA-II in terms of every running time. Furthermore, the computation time of the MSAC algorithm is reduced by 3.16 times, 2.94 times, and 1.83 times respectively. The solving time that the Monte Carlo simulation takes increases with the number of network nodes. The data set in the process of the evolutionary algorithm is added to the historical data as the SVR regression function is added, which helps reduce the number of calculations of the adaptive value function and improve the calculation time of the algorithm. In the 30 simulation experiments, the average error of DD-MSAC and MSAC is less than 5%, while the accuracy of the objective function value under the data-driven approach is higher. However, the acceleration of the data-driven approach has no significant impact on the solution quality. Further analysis of the every running time of different data-driven approaches suggests that DD-MSAC and DP-MSAC have apparent advantages in the calculation time of optimizing small-scale networks. The acceleration effect of the SVM-MSAC approach is the most stable, and the average optimization range is about 50%.

As a potential future research direction, it is essential to address environmental concerns in the study of multimodal transportation route optimization problems. For example, one possible approach is to add the reduction of carbon emissions to the multi-objective optimization framework of our study. Furthermore, the optimization model with multiple objectives including transportation time, cost, and carbon emission minimization together may need approaches of nondominated sorting for high-dimensional multi-objective optimization.

Acknowledgements. Authors thank the editor and two referees for their useful and valuable suggestions and comments. This work was supported in part by the Social Science Planning Project of Chongqing, China (No. 2019YBGL049) and Transportation Science and Technology Project of Chongqing Transportation Bureau, (No. 2022-17).

References

[1] T.G. Crainic and K.H. Kim, Intermodal transportation. *Handb. Oper. Res. Manage. Sci.* **14** (2007) 467–537.

[2] Z. He, K. Navneet, W. van Dam and P. Van Mieghem, Robustness assessment of multimodal freight transport networks. *Reliab. Eng. Syst. Saf.* **207** (2021) 107315.

[3] W. Zhu, H. Wang and X. Zhang, Synergy evaluation model of container multimodal transport based on BP neural network. *Neural Comput. App.* **33** (2021) 4087–4095.

[4] T. Yamada, B.F. Russ, J. Castro and E. Taniguchi, Designing multimodal freight transport networks: a heuristic approach and applications. *Transp. Sci.* **43** (2009) 129–143.

[5] H.G. Resat and M. Turkay, A discrete-continuous optimization approach for the design and operation of synchromodal transportation networks. *Comput. Ind. Eng.* **130** (2019) 512–525.

[6] W. Qu, J. Rezaei, Y. Maknoon and L. Tavasszy, Hinterland freight transportation replanning model under the framework of synchromodality. *Transp. Res.* **131** (2019) 308–328.

[7] Y. Sheng and Y. Gao, Shortest path problem of uncertain random network. *Comput. Ind. Eng.* **99** (2016) 97–105.

[8] L. Wang, L. Yang and Z.Y. Gao, The constrained shortest path problem with stochastic correlated link travel times. *Eur. J. Oper. Res.* **255** (2016) 43–57.

[9] X. Wang and Q. Meng, Discrete intermodal freight transportation network design with route choice behavior of intermodal operators. *Transp. Res. Part B Methodol.* **95** (2017) 76–104.

[10] R. Zhang, W.Y. Yun and I.K. Moon, Modeling and optimization of a container drayage problem with resource constraints. *Int. J. Prod. Econ.* **133** (2011) 351–359.

[11] Y. Zhang, B. Atasoy and R.R. Negenborn, Preference-based multi-objective optimization for synchromodal transport using adaptive large neighborhood search. *Transp. Res. Record* **2676** (2022) 71–87.

[12] H.G. Resat and M. Turkay, Design and operation of intermodal transportation network in the Marmara region of Turkey. *Transp. Res. Part E: Logistics Transp. Rev.* **83** (2015) 16–33.

[13] J.S.L. Lam and Y. Gu, A market-oriented approach for intermodal network optimization meeting cost, time and environmental requirements. *Int. J. Prod. Econ.* **171** (2016) 266–274.

[14] H. Wei and M. Dong, Import-export freight organization and optimization in the dry-portbased cross-border logistics network under the Belt and Road Initiative. *Comput. Ind. Eng.* **130** (2019) 472–484.

[15] G. Mavrotas and K. Florios, An improved version of the augmented ε-constraint method (AUGMECON2) for finding the exact pareto set in multi-objective integer programming problems. *Appl. Math. Comput.* **219** (2013) 9652–9669.

[16] M. Kalinina, L. Olsson and A. Larsson, A multi objective chance constrained programming model for intermodal logistics with uncertain time. *Int. J. Comput. Sci.* **10** (2013) 35–44.

[17] L. Li, R.R. Negenborn and B. De Schutter, Intermodal freight transport planning: a receding horizon control approach. *Transp. Res. Part C: Emerg. Technol.* **60** (2015) 77–95.

[18] Y. Sun, M. Lang and D. Wang, Bi-objective modelling for hazardous materials road-rail multimodal routing problem with railway schedule-based space-time constraints. *Int. J. Environ. Res. Publ. Health* **13** (2016) 762.

[19] S. Liu, Y. Peng, Q. Song and Y. Zhong, The robust shortest path problem for multimodal transportation considering timetable with interval data. *Syst. Sci. Control Eng.* **6** (2018) 68–78.

[20] H. Zhang, Y. Li, Q. Zhang and D. Chen, Route selection of multimodal transport based on china railway transportation. *J. Adv. Transp.* **2021** (2021) 9984659.

[21] R. Wang, K. Yang, L. Yang and Z.Y. Gao, Modeling and optimization of a road-rail intermodal transport system under uncertain information. *Eng. App. Artif. Intell.* **72** (2018) 423–436.

[22] A. Abbassi, E.H.A. Ahmed and J. Boukachour, Robust optimization of the intermodal freight transport problem: modeling and solving with an efficient. *J. Comput. Sci.* **30** (2019) 127–142.

[23] Y. Lu, M. Lang, Y. Sun and S. Li, A fuzzy intercontinental road-rail multimodal routing model with time and train capacity uncertainty and fuzzy programming approaches. *IEEE Access* **8** (2020) 27532–27548.

[24] E. Demir, W. Burgholzer, M. Hrusovsky, E. Arkan, W. Jammernegg and T.V. Woensel, A green intermodal service network design problem with travel time uncertainty. *Transp. Res. Part B: Methodol.* **93** (2016) 789–807.

[25] S. Fazayeli, A. Eydi and I.N. Kamalabadi, Location-routing problem in multimodal transportation network with time windows and fuzzy demands: presenting a two-part genetic algorithm. *Comput. Ind. Eng.* **119** (2018) 233–246.

[26] Z. Ziaei and A. Jabbarzadeh, A multi-objective robust optimization approach for green location-routing planning of multi-modal transportation systems under uncertainty. *J. Cleaner Prod.* **291** (2020) 125–293.

[27] P. Robbe, D. Nuyens and S. Vandewalle, Recycling samples in the multigrid multilevel (quasi-) Monte Carlo method. *SIAM J. Sci. Comput.* **4541** (2019) S37–S60.

[28] J. Guo, Q. Du and Z. He, A method to improve the resilience of multimodal transport network: location selection strategy of emergency rescue facilities. *Comput. Ind. Eng.* **161** (2021) 107678.

[29] C. Cintrano, F. Chicano and E. Alba, Facing robustness as a multi-objective problem: a biobjective shortest path problem in smart regions. *Inf. Sci.* **503** (2019) 255–273.

[30] H. Wang, L. Tan, J. Shi, X. Lv and X. Lian, An improved NSGA-II algorithm for UAV path planning problems. *J. Internet Technol.* **22** (2021) 583–592.

[31] Z. Li, Y. Liu and Z. Yang, An effective kernel search and dynamic programming hybrid heuristic for a multimodal transportation planning problem with order consolidation. *Transp. Res. Part E: Logistics Transp. Rev.* **152** (2021) 102408.

[32] Y. Peng, P. Yong and Y. Luo, The route problem of multimodal transportation with timetable under uncertainty: multi-objective robust optimization model and heuristic approach. *RAIRO: Oper. Res.* **55** (2021) S3035–S3050.

[33] X. Chen and Y. Dai, Research on an improved ant colony algorithm fusion with genetic algorithm for route planning, in IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). Vol. 1. IEEE (2020) 1273–1278.

[34] G. Zhang, H. Wang, W. Zhao, Z. Guan and P. Li, Application of improved multi-objective ant colony optimization algorithm in ship weather routing. *J. Ocean Univ. Chin.* **20** (2021) 45–55.

[35] F. Xie, J. Long, Z. Qian, Z. Ding and L. Liu, Multi-objective optimization routing for satellite network based on ant colony algorithm, in IEEE 13th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA). IEEE (2021) 353–356.

[36] H. Wang, Y. Jin, C. Sun and J. Doherty, Offline data-driven evolutionary optimization using selective surrogate ensembles. *IEEE Trans. Evol. Comput.* **23** (2018) 203–216.

[37] X. Chao, Y. Zhongqing, L. Jinhua and Y. Xixin, Application of data driven technology in wastewater treatment process, in IEEE 5th International Conference on Information Science, Computer Technology and Transportation (ISCTT). IEEE (2020) 48–55.

[38] Q. Gu, D. Wang, S. Jiang, N. Xiong and Y. Jin, An improved assisted evolutionary algorithm for data-driven mixed integer optimization based on Two_Arch. *Comput. Ind. Eng.* **159** (2021) 107463.

[39] H. Wang and Y. Jin, A random forest-assisted evolutionary algorithm for data-driven constrained multiobjective combinatorial optimization of trauma systems. *IEEE Trans. Cybern.* **50** (2018) 536–549.

[40] H. Lou, B. Gao, F. Jin, Y. Wan and Y. Wang, Shear wall layout optimization strategy for high-rise buildings based on conceptual design and data-driven tabu search. *Comput. Struct.* **250** (2021) 106546.

[41] Y. Peng, Y.J. Luo, P. Jiang and P.C. Yong, The route problem of multimodal transportation with timetable: stochastic multi-objective optimization model and data-driven simheuristic approach. *Eng. Comput.* **39** (2022) 587–608.

[42] A. Juan, J. Faulin, S. Grasman, D. Riera, J. Marull and C. Mendez, Using safety stocks and simulation to solve the vehicle routing problem with stochastic demands. *Transp. Res. Part C: Emerg. Technol.* **19** (2011) 751–765.

[43] M. Dorigo and L.M. Gambardella, Ant colony system: a cooperative learning approach to the traveling salesman problem. *IEEE Trans. Evol. Comput.* **1** (1997) 53–66.

[44] Z.Y. Duan, Z.X. Lei and S. Sun, Multi-objective robust optimisation method for stochastic time-dependent vehicle routing problem. *J. Southwest Jiaotong Univ.* **54** (2019) 565–572.

[45] V. Vapnik, The Nature of Statistical Learning Theory. Springer Science & Business Media (2013).

[46] M.S. Ahmad, S.M. Adnan, S. Zaidi and P. Bhargava, A novel support vector regression (SVR) model for the prediction of splice strength of the unconfined beam specimens. *Constr. Build. Mater.* **248** (2020) 118475.

[47] Y. Zhou, T. Kundu, M. Goh and J.B. Sheu, Multimodal transportation network centrality analysis for belt and road initiative. *Transp. Res. Part E: Logistics Transp. Rev.* **149** (2021) 102292.