A Time-Dependent Fuzzy Programming Approach for the Green Multimodal Routing Problem with Rail Service Capacity Uncertainty and Road Traffic Congestion

Yan Sun, Martin Hrušovský, Chen Zhang, and Maoxiang Lang

1 School of Management Science and Engineering, Shandong University of Finance and Economics, No. 7366, Second Ring East Road, Jinan, Shandong Province 250014, China
2 Institute for Production Management, WU Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna, Austria
3 Unit of Logistics and Informatics, KTH Royal Institute of Technology, Teknringen 10, 10044 Stockholm, Sweden
4 School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Yan Sun; sunyanbjtu@163.com

Received 30 October 2017; Revised 7 April 2018; Accepted 2 May 2018; Published 24 June 2018

Academic Editor: Lucia Valentina Gambuzza

Copyright © 2018 Yan Sun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study explores an operational-level container routing problem in the road-rail multimodal service network. In response to the demand for an environmentally friendly transportation, we extend the problem into a green version by using both emission charging method and bi-objective optimization to optimize the CO₂ emissions in the routing. Two uncertain factors, including capacity uncertainty of rail services and travel time uncertainty of road services, are formulated in order to improve the reliability of the routes. By using the triangular fuzzy numbers and time-dependent travel time to separately model the capacity uncertainty and travel time uncertainty, we establish a fuzzy chance-constrained mixed integer nonlinear programming model. A linearization-based exact solution strategy is designed, so that the problem can be effectively solved by any exact solution algorithm on any mathematical programming software. An empirical case is presented to demonstrate the feasibility of the proposed methods. In the case discussion, sensitivity analysis and bi-objective optimization analysis are used to find that the bi-objective optimization method is more effective than the emission charging method in lowering the CO₂ emissions for the given case. Then, we combine sensitivity analysis and fuzzy simulation to identify the best confidence value in the fuzzy chance constraint. All the discussion will help decision makers to better organize the green multimodal transportation.

1. Introduction

With the rapid development of globalization, companies are seeking for specialized partners all over the world to outsource their businesses and also for new markets to make more profit. International trade and global commodity circulation thus get significantly motivated during the last decades [1]. Supported by the transportation industry, this prosperity results in a remarkable expansion of the transportation network from a limited region to the entire world. The expanded transportation network extends the distribution channels and enhances the difficulty of transportation organization from the viewpoint of economy, efficiency, reliability, and environmental protection. The traditional unimodal transportation is no longer suitable for this situation. Meanwhile, an advanced transportation mode, namely, the multimodal transportation, emerges in the transportation industry. It combines different transportation modes to generate origin-to-destination routes and integrates their respective advantages in the transportation process. Furthermore, by using standardized containers, multimodal transportation can be highly mechanized and the efficiency of the operations at terminals is greatly improved [2]. Many empirical case studies, for example, Bookbinder and Fox [3], Janic [4], and Liao et al. [5], have demonstrated the superiority of the multimodal transportation in the long-haul transportation.
setting. Nowadays, multimodal transportation is getting more and more popular and plays a crucial role in international trade. Since the transportation chains in this area are usually long combining different actors, effective solutions for their coordination are necessary in order to ensure that the goods are delivered according to the customer's requirements. In this context, the multimodal routing problem has gained interest of both researchers and practitioners [6, 7]. This operational problem aims at selecting the best routes to move containers from their origins to destinations through the multimodal service network according to customer demands [2]. In order to achieve this objective, a number of different factors have to be considered.

First of all, multimodal transportation combines different modes with different operational constraints that have to be represented in the mathematical model. In general, it can be distinguished between time-flexible modes (e.g., road) that can be used whenever they are needed and schedule-based modes (e.g., rail) that operate according to fixed schedules planned in advance. These schedules regulate the route as well as the (un-)loading, departure, and arrival times of the transportation service. Since these schedules cannot be easily changed, they should be considered when planning the routes of the containers and their transshipment between different transportation services; otherwise, the multimodal plans might be infeasible in reality.

Secondly, the actors and stakeholders involved in multimodal transportation systems have usually different interests which are often in conflict. Whereas transportation operators usually tend to minimize only the transportation costs, customers are also interested in the reliability of the routes in order to avoid situations where a plan becomes infeasible due to late arrival of a service or insufficient capacity. Since the multimodal routing is done before the start of transportation, when the information about the transportation services (e.g., traffic situation, travel time, and capacity) is not completely known, possible sources of uncertainty should be considered already during the planning [8]. As an example, the scheduled rail services have usually limited capacity due to the limited length of the tracks in the railway stations and limited locomotive power [9]. However, the available capacities of rail services are uncertain since they are also dependent on other transportation plans that are difficult to predict [10]. If only deterministic capacities are considered in the model, two consequences may happen in practice: the planned routes are either infeasible if the capacities are estimated to be too large, or the planned routes are not the actual best ones if the capacity estimation is too small. Therefore, capacity uncertainty is a factor that should be considered in the decision-making process. In addition to that, road services frequently suffer from various disruptions, such as natural disasters (e.g., bad weather) and manmade disruptions (e.g., traffic accidents and congestion) [11]. Among these disruptions, traffic congestion is very common due to the rapidly increasing traffic volume in recent years [12]. Traffic congestion results in the travel time uncertainty and delayed arrivals of the containers at the nodes, which disrupts the transshipment if the containers have to take another scheduled transportation service at the same node and also possibly violates the due date constraint at the destination. Thus, congestion is another factor that should be also considered in the decision-making process.

Thirdly, the environmental aspects of transportation have drawn considerable attention from the public and the government [13]. Among the various environmental issues, global warming caused by greenhouse gas emissions has been considered as a tremendous threat to human sustainable development. CO₂ accounts for approximately 80% of the entire greenhouse gas emissions [5]. During recent decades, large numbers of countries have proposed regulations to ease the global warming by reducing CO₂ emissions. For example, in 2012, the National Development and Reform Commission of China proposed a pilot project on trading carbon emissions [2]. Since the transportation sector is one of the biggest contributors to the CO₂ emissions [14, 15], the transportation industry should accept this responsibility and consider controlling and optimizing the CO₂ emissions during the transportation planning stage to promote the development of an environmentally friendly transportation. In order to respond to these environmental requirements, in this study, we extend the multimodal routing problem into a green version by considering the CO₂ emissions.

The inclusion of uncertain factors and the combination of multiple objectives (i.e., economic and environmental) in multimodal transport planning is an emerging field which has been studied in recent years [13–15]. However, the authors usually consider only one source of uncertainty or model the CO₂ emissions in form of charges which limits their influence on the resulting routes. Therefore, we present a multimodal routing planning approach which deals with the following question:

How can the sustainability and reliability of multimodal routing be improved by considering environmental factors and uncertainty?

The contributions of the paper are fivefold:

(1) A mixed integer nonlinear optimization model for multimodal routing including multiple objectives (i.e., economic and environmental) and multiple uncertainty factors (i.e., road traffic congestion and rail service capacity) is defined.

(2) A fuzzy logic is proposed to model the capacity uncertainty, and a piecewise linear function is used to represent the road traffic congestion.

(3) The influence of CO₂ emissions on the routes is analyzed comparing the CO₂ emissions charging method and a bi-objective approach in which minimizing CO₂ emissions is set as an independent objective.

(4) Since the proposed optimization model is nonlinear, a linearization approach is proposed in order to reduce the complexity of the model.

(5) The planning approach is applied to a real-world case study from China with specific network design and parameters. In this case, the network is limited to road and rail services; however, other modes with fixed schedules (e.g., inland waterway and deep sea
The multimodal transportation. Finally, conclusions of this section are presented, which can further help the decision makers to better organize the transportation orders characterized by their origin, destination, volume of containers, release time at origin, and due date at destination. Whereas the release time is usually considered as fixed, the due date can be either considered as a hard constraint (see, e.g., Verma et al. [17], Sun and Lang [2, 18], and Sun et al. [19]) or delayed deliveries can be accepted causing additional penalty costs (see, e.g., Demir et al. [20], Hrušovský et al. [14]).

Moreover, considering only late arrivals might not be sufficient since early deliveries can also cause problems with the storage of the goods until they are needed. Therefore, various studies consider due date time windows instead of fixed schedules (see, e.g., Demir et al. [20], Hrušovský et al. [14], and Zhang et al. [21]).

However, considering only late arrivals might not be sufficient since early deliveries can also cause problems with the storage of the goods until they are needed. Therefore, various studies consider due date time windows instead of fixed schedules (see, e.g., Demir et al. [20], Hrušovský et al. [14], and Zhang et al. [21]).

The containers from transportation orders need to be transported in the multimodal service network by various transportation services with their specific characteristics. One of these characteristics is the usually fixed schedule of the high-capacity and environmentally friendly transportation modes (e.g., railway or inland waterway) in contrast to the time-flexible (nonscheduled) road service. However, the majority of the related articles neglects the schedules and considers rail services to be also time-flexible. This facilitates the modelling since only transshipment times between services in terminals need to be considered (Sun et al. [23], Jiang and Lu [24], Cai et al. [25], Zhang et al. [21], and Xiong and Wang [26]). However, this does not correspond to reality where railway services are usually planned well in advance [2, 19]. Therefore, the schedules are usually fixed in the tactical network design phase (see, e.g., Wieberneit [27], Andersen et al. [28], and SteadieSeif et al. [29] for this area) and have to be accepted in the operative multimodal routing considered in this paper (Liu et al. [30], Lin [31], Demir et al. [20], and Hrušovský et al. [14]).

The routes can be optimized according to different criteria. Traditionally, the optimization models concentrate on costs or time [32]. However, various surveys show that multiple criteria, among others also sustainability and reliability, are important for transportation planners [33, 34]. In case of sustainability, vast majority of available literature only considers economic and environmental criteria whereas social aspects of transportation are neglected [35, 36]. As a result, the comparison of transportation costs and CO2 emissions is decisive for finding the optimal routes. The CO2 emissions can be either integrated into the cost objective function as emission charges or modelled as an independent objective in multiobjective optimization. The majority of the articles on optimizing CO2 emissions in transportation adopts the first approach (e.g., Chang et al. [37], Zhang et al. [21], Wen et al. [38], Sun and Lang [2, 18], Chen et al. [39], Demir et al. [20], and Hrušovský et al. [14]), whereas only a few studies used the second approach (e.g., Sun and Chen [40] and Qu et al. [41]). However, it is difficult to decide which method is better to use since the charging method is easier to implement but might not always be applicable, especially in cases where the emission charges are too low in comparison to transportation costs and therefore do not have any influence on the resulting routes [22].

Another important decision is the choice of the emission calculation method since most of the models require a lot of parameters (e.g., fuel type, vehicle and route characteristics, and traffic condition) that cannot be always estimated in advance [20, 42–45]. As an alternative, the activity-based method that multiplies activity intensity (unit: TEU-km) with the CO2 emission factor (unit: g/TEU-km) to calculate the total CO2 emissions [5] yields a better feasibility in transportation practice and has been already applied in various studies [2, 5, 37, 39, 41, 46].

The reliability of transportation is influenced by various sources of uncertainty (e.g., weather, accidents, congestion, and technical issues) that can cause disruptions in transportation processes. Despite this fact, most of the existing models consider all factors as deterministic and do not take into account uncertainty [47]. In order to improve the reliability of transportation plans, the uncertainty should be considered during the planning phase. Despite the variety of possible reasons for disruption, their effects influence either the demand, the travel time, or the supply (capacity) of transportation. Whereas extensive literature about demand uncertainty in multimodal routing is available (see, e.g., Lium et al. [48] and Ho et al. [49]), the consideration of travel time and capacity uncertainty is rather limited [50, 51]. Moreover, usually the models deal with only one source of uncertainty. However, the consideration of multiple uncertainties at the same time might lead to improved routes and therefore in
our study the travel time uncertainty for road services and capacity uncertainty for rail services should be combined.

Road traffic congestion leads to travel time uncertainty of road services. Only a few studies, for example, Demir et al. [20], Sun and Chen [40], and Uddin and Huynh [11], adopted stochastic programming to formulate the multimodal service network problem under disruptions. However, in practical planning, due to the lack of reliable historical data [52], it is quite difficult to fit the probability distributions to the travel times of all road services in the multimodal service network. Such disadvantage of stochastic programming in addressing the uncertainty is emphasized by Zarandi et al. [52] in a location-routing problem. Furthermore, the fuzzy programming that uses fuzzy variables (usually triangular or trapezoidal numbers) to evaluate the uncertainty fails to represent the variation of the travel time of a road service en route with respect to the time of the day [53]. Therefore, it is worthwhile to find an applicable approach (e.g., time-dependent travel time that is widely discussed in the vehicle routing problems [53, 54]) to formulate the road traffic congestion and introduce it to the multimodal routing modelling.

The capacity uncertainty has only been considered by a few studies in different transportation planning problems, for example, multiobjective solid minimal cost flow problem by Wang [55], multiobjective multi-item solid transportation problem by Kaur and Kumar [56], and resource allocation problem for containerized cargo transportation networks by Kundu et al. [57]. Whereas Wang [55] used stochastic programming to address this issue, Kaur and Kumar [56] and Kundu et al. [57] formulated the capacity uncertainty from a fuzzy programming viewpoint. The fuzzy programming approach has been also used in the context of vehicle routing problem [52, 58] and shortest path problem [59], since it can avoid the infeasibility of the stochastic programming in case of lack of historical data [58]. However, to the best of our knowledge, there exists no study which would consider capacity uncertainty in multimodal routing problem.

The considered aspects of the multimodal routing approach increase the complexity of the problem. The modelling of constraints for schedule-based services leads to the nonlinearity of the model, which can be seen in the models proposed by Liu et al. [30], Lin [31], and Chang [60]. Moreover, the modelling of uncertainty sources might also lead to nonlinearity of the model. As for the problem formulated by a nonlinear model, Xie et al. [61] pointed out that the nonlinearity makes the problem very difficult to solve, while using linearization technique to get the equivalent linear model can enable the problem to be relatively easily solved. The linearization technique should be designed according to the specific nonlinearity of the nonlinear model. Its main idea is to introduce new linear forms to replace the nonlinear components in the model equivalently.

As the literature review shows, the interest of researchers in multimodal routing has been increasing during the last years, considering different aspects of the studied planning problem. However, there still exists potential for improvement especially in the area of sustainability and reliability of transportation. Therefore, our paper contributes to these aspects of the multimodal routing since it introduces and compares two different methods for modelling CO₂ emissions. Moreover, two different sources of uncertainty (travel time and capacity) are considered within the planning process which should reduce the influence of potential disruptions on created transportation plans and improve the reliability. In case of rail service capacity, the fuzzy programming approach is used in order to better represent the possible capacity variations. Although the application of the model to the Chinese case study requires some specific assumptions, the general model can be used for any multimodal service network.

In order to give an overview of the current state of the literature, the most important articles with the covered aspects are shown in Table 1.

3. Modelling Methodology

As already mentioned, the objective of the multimodal routing problem in this study is to select the best green routes to move containers from multiple transportation orders through the multimodal service network with road traffic congestion and rail service capacity uncertainty. The multimodal service network is composed of nodes and arcs, where nodes represent the terminals serving as origins and destinations of transportation orders as well as transshipment points between rail and road services. The arcs represent the services connecting the nodes. Related information on road and rail services, such as schedules, transportation distance, and capacity, is regarded as the parameters of the nodes and arcs.

In the studied network, rail services are the backbone for multimodal container transportation, while road services play an important role in origin pickup and destination delivery. In some cases, road services also take direct origin-to-destination transportation. Figure 1 shows the basic topological structure of the multimodal service network. Overall, the network structure is similar to a hub-and-spoke network with multiple hubs where origins and destinations represent spokes while railway container stations represent hubs.

3.1. Modelling Transportation Services. Service characteristics are dependent on the chosen transportation mode. In case of road service, the trucks are assumed to be fully flexible which means that they are available every time they are needed and it is possible to assign as many trucks as necessary to carry the containers. As a result, the road services are uncapacitated and their departure time can be adjusted according to the customer requirements and the current state of the network. This means that the departure time can be either immediately after the containers are loaded on the vehicle, but can be also postponed in order to avoid traffic congestion. However, it has to be ensured that the truck arrives to its destination before the loading cutoff time of the next planned service in order to avoid infeasibility of the transportation plan. Besides that, too early arrivals should be also avoided as inventory costs can be charged for containers waiting for the next service at a terminal. Therefore, it is important to
determine the planned departure time of a truck from the current node.

In contrast to that, rail services are modelled as block container trains driving between two nodes according to fixed schedules with the following characteristics:

(1) Operation time window at the loading/unloading organization station that regulates the operation start time and cutoff time regarding loading/unloading containers on/off the train

(2) Scheduled departure/arrival time from/at the loading/unloading organization station

(3) Operation period. For the convenience of modelling, same rail services in different operation periods are considered as different services.

Since the loading/unloading operations of both road and rail services are highly mechanized, the time needed for loading/unloading is relatively low. Consequently, we assume that these times can be neglected in the multimodal routing model. Besides that, the capacities of rail services are limited. Taking into account these assumptions, the origin-to-destination transportation process combining road and rail services is depicted in Figure 2 and can be described as follows:

### Table 1: Literature review of the articles in the multimodal routing problem.

<table>
<thead>
<tr>
<th>Articles</th>
<th>Due date consideration</th>
<th>Rail schedule constraints</th>
<th>Rail service capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chang (2008) [60]</td>
<td>Time point</td>
<td>American scenario</td>
<td>Deterministic</td>
</tr>
<tr>
<td>Sun et al. (2008) [23]</td>
<td>Time window</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jiang and Lu (2009) [24]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Çakir (2009) [62]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cai et al. (2010) [25]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chang et al. (2010) [37]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liu et al. (2011) [30]</td>
<td></td>
<td>Scheduled departure time</td>
<td></td>
</tr>
<tr>
<td>Zhang et al. (2011) [21]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moccia et al. (2011) [63]</td>
<td></td>
<td>Italian scenario</td>
<td></td>
</tr>
<tr>
<td>Ayar and Yaman (2012) [64]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verma et al. (2012) [17]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sun and Chen (2013) [40]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wen et al. (2013) [38]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xiong and Wang (2014) [26]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demir et al. (2015) [20]</td>
<td></td>
<td>Scheduled departure time</td>
<td></td>
</tr>
<tr>
<td>Lin (2015) [31]</td>
<td></td>
<td>Scheduled departure time</td>
<td></td>
</tr>
<tr>
<td>Uddin and Huynh (2016) [11]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hrušovský et al. [14]</td>
<td></td>
<td>Scheduled departure time</td>
<td></td>
</tr>
<tr>
<td>Qu et al. (2016) [41]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sun et al. (2016) [19]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhang et al. (2017) [22]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The containers will first depart from the origin (node 1) at the planned departure time that is not earlier than the earliest release time defined by the customer, and then arrive at the loading organization station (node 2).

In order to be able to load the containers on the train operated from node 2 to node 3, their arrival time at node 2 should be no later than the operation start time. If the arrival time is earlier than the operation start time, there will be an inventory period from the arrival time to the operation start time at node 2, and the containers should wait until the operation start time and then get loaded. Moreover, there exists a free-of-charge period for inventory costs [64]. Only when the inventory period is longer than this period will inventory costs be created.

When loaded on the train, the containers will wait until the scheduled departure time and then leave node 2 along with the train.

The containers will arrive at the unloading organization station (node 3) at the scheduled arrival time of the train. However, they should wait until the operation start time and then get unloaded. Therefore,
the effective arrival time of the containers at node 3 is the operation start time instead of the scheduled arrival time.

(5) After unloading from the train, the transportation organizer can select a planned departure time that is no earlier than the operation start time to move containers to their destination (node 4). In case of an early arrival to the destination, the containers have to be stored again until the planned delivery date in a warehouse provided by the third-party inventory companies that can have different conditions in comparison to warehouses used in intermodal terminals [65].

3.2. Modelling Uncertainty. Considering the uncertainty factors, traffic congestion is an important factor in road transportation that influences the travel times. Congestion can occur either regularly due to insufficient road capacity or unexpectedly due to, for example, an accident. Since the latter is hard to predict, only regularly occurring congestion is considered in this study. However, this congestion can vary with the time of the day or the section considered, therefore piecewise linear functions can be used to model such situations [53, 54]. By combining limited historical data and experiences from the experts, such travel time functions can be established. Figure 3 shows an example of the time-dependent travel time function. There are two peak hours from 9 am to 10 am and from 6 pm to 8 pm on the arc served by road services.

Besides the road congestion, the available capacity of a rail service can be also uncertain due to various reasons (e.g., order cancellation, vehicle breakdown, and lack of available vehicles). Since the reasons are manifold and the historical data about capacity variation might not be available, using the stochastic programming for capacity uncertainty might not be feasible. As a consequence, capacity uncertainty can be effectively expressed in a fuzzy way by using the triangular fuzzy variable represented by the minimum of the possible capacities, most possible capacity and maximum of the possible capacities. Similar to the traffic congestion issue, fuzzy capacity can be attained by limited historical data and expert experiences.

3.3. Modelling and Optimizing the CO₂ Emissions. As stated in Section 2, there are two approaches to optimize the CO₂ emissions, including the CO₂ emission charging method and the bi-objective optimization method. The optimization model is firstly constructed in Section 4 using the charging method, and the bi-objective method is then introduced in Section 6 where both methods are compared. For calculating the CO₂ emissions, the activity-based method is used. This method calculates the total CO₂ emissions for a transportation service using the formula \((EM \cdot Q \cdot L)\) where \(EM\) is the CO₂ emission factor of the service (unit: g/TEU-km), \(Q\) is the number of the containers carried by the transportation service (unit: TEU), and \(L\) is the corresponding transportation distance (unit: km).

3.4. Modelling the Customer Demands. Transportation orders are considered to be unsplittable which means that all containers of an order have to use the same route. The release times of orders are fixed whereas the delivery can be done within a due date time window limited by a lower and upper bound. For each transportation order, it is best if the arrival time of the containers falls within the time window. In this case, there will be no extra costs at the destination. However, if the containers arrive at the destination earlier than the lower bound, inventory costs will be created, because the goods need to be stored before being processed at the destination. If the arrival time is later than the upper bound, penalty costs should be charged for the delayed delivery.

4. Mathematical Model for the Multimodal Routing Problem

4.1. Notations. In this study, \(G = (N, A, S)\) represents a road-rail multimodal service network in form of a directed graph, where \(N, A,\) and \(S\) separately denote the node set, arc set, and transportation service set in the network. For each transportation order \(k (\forall k \in K\) where \(K\) is the transportation order set), its five attributes are represented as origin \(o_k\), destination \(d_k\), volume \(q_k\) (unit: TEU), earliest release time \(t^r_k\) at origin, and due date \(t^d_k\) where \(t^r_k\) and \(t^d_k\) are separately the lower bound and upper bound of the due date time window, respectively. The rest of the symbols in the mathematical model and their respective representations are summarized as follows in Table 2.

4.2. A Fuzzy Chance-Constrained MINLP Model. The framework of the mathematical model is derived from the classical arc-node-based formulation that has been widely used in the routing problems of various research fields, for example, telecommunications, manufacturing systems, and internet service systems [1]. The mathematical model for the multimodal routing problem explored in this study is presented as follows. The proposed model is a combination of the fuzzy chance-constrained programming and the mixed integer nonlinear programming (MINLP).

Objective function:

\[
\text{minimize } \sum_{k \in K} \sum_{(i,j) \in A \in S_i} c_{ij}^k q_k x_{ij}^k, \quad (1)
\]

\[
+ \sum_{k \in K} \sum_{(i,j) \in A \in S_i} \left( \sum_{h \in \Delta} \sum_{(i,j) \in S_h} c_{ij}^h q_k x_{ij}^h + \sum_{j \in \Delta} \sum_{(i,j) \in S_j} c_{ij}^j q_k x_{ij}^j \right), \quad (2)
\]

\[
+ \sum_{k \in K} \sum_{(i,j) \in A \in S_i} c_{ij}^k q_k w_{ij}^k + \sum_{k \in K} c_{d_k}^k q_k \max \left\{ t^d_k - y_{d_k}^k, 0 \right\}, \quad (3)
\]

\[
+ \sum_{k \in K} \text{pen} \cdot \max \left\{ y_{d_k}^k - t^U_k, 0 \right\}, \quad (4)
\]

\[
+ \sum_{k \in K} \sum_{(i,j) \in A \in S_i} c_{co_2} \cdot \text{em}_i q_k d_{ij} x_{ij}^k, \quad (5)
\]
Eqs. (1), (2), (3), (4), and (5) are successively the transportation costs en route, loading and unloading costs at the nodes, inventory costs, penalty costs caused by delayed deliveries at destinations, and CO2 emission costs. Their linear summation represents the generalized costs. The objective function minimizes the generalized costs. The objective function minimizes the generalized costs. The objective function minimizes the generalized costs. The objective function minimizes the generalized costs. The objective function minimizes the generalized costs.

This setting reflects the objective of satisfying customer demands and minimizing monetary expenditures. Therefore, the multimodal routing in this study is a kind of service-oriented planning.

Subject to

\[
\begin{align*}
\sum_{j \in S(j)} \sum_{i \in S_i} x_{ij}^k - \sum_{h \in S(h)} \sum_{i \in S_h} x_{hi}^k &= \\
&= \begin{cases} 
1 & i = o_k \\
0 & \forall i \in N \setminus \{o_k, d_k\} \forall k \in K, \forall i \in N, \\
-1 & i = d_k
\end{cases} \\
\sum_{s \in S_i} x_{ij}^k &\leq 1 \forall k \in K, \forall (i, j) \in A.
\end{align*}
\]

Eq. (6) is the flow conservation constraint which ensures the continuity of the nodes and arcs on the planned routes, which is a common constraint in the node-arc-based formulation [1]. Eq. (7) means that no more than one transportation service on one arc can be utilized to serve a transportation order. The combination of the two equations above ensures that containers in a transportation order are unsplittable, which is claimed in Section 3.4.

\[
\text{Cr}\left\{ \sum_{k \in K} q_k \cdot x_{ij}^k \leq \bar{v}_{ij} \right\} \geq \alpha \forall (i, j) \in A, \forall s \in \Gamma_{ij}.
\]

Eq. (8) is the fuzzy chance constraint of the rail service capacity based on the fuzzy credibility measure [66]. In this equation, \(\text{Cr}\{\cdot\}\) represents the credibility of the event in \{\cdot\}. Eq. (8) ensures that the credibility of the event that the entire loaded container volume on a rail service does not exceed its available capacity should not be lower than a predetermined confidence \(\alpha \in [0, 1]\). The value of \(\alpha\) indicates the decision makers’ preference for addressing the corresponding decision-making.

\[
y_{ijs}^k = t_{\text{release}}^k \forall k \in K.
\]

Eq. (9) assumes that the arrival time of the containers of transportation order \(k\) at the origin equals its earliest release time.

\[
y_i^k \leq z_i^k \forall k \in K, \forall i \in N \setminus \{d_k\},
\]

\[
n_i^k = \left[ \frac{z_i^k}{24} \right] \forall k \in K, \forall i \in N \setminus \{d_k\},
\]

\[
t_{ij}^k = f_{ij}\left(z_i^k - 24 \cdot n_i^k\right) \forall k \in K, (i, j) \in A, \forall s \in \Omega_{ij},
\]

\[
\left(z_i^k + \bar{v}_{ij}^k - y_{ijs}^k\right) \cdot x_{ij}^k = 0 \forall k \in K, (i, j) \in A, \forall s \in \Omega_{ij}.
\]

Eqs. (10), (11), (12), and (13) compute the arrival times of the containers of transportation order \(k\) at the nodes on the route by road services. First, we determine the planned departure time of the containers at the current node \(i\) by (10). Then, we normalize such departure time into range [0, 24] by using the floor integral function (11) and normalization formula \((z_i^k - 24 \cdot n_i^k)\), so that we can further generate the corresponding travel time \(t_{ij}^k\) of the containers on arc \((i, j)\) by road service \(s\) according to the piecewise linear function (12). Finally, we can gain the arrival time of container flow \(k\) at the following node \(j\) by (13).

\[
\left(t_{ij}^k - y_{ijs}^k\right) \cdot x_{ij}^k = 0 \forall k \in K, (h, i) \in A, \forall r \in \Gamma_{hi}.
\]

Eq. (14) computes the arrival times of the containers of transportation order \(k\) at the nodes on the route by rail services. As claimed in Section 3.1, \(t_{ij}^k\) is the effective arrival time of containers at node \(i\). Because we assume that
loading/unloading time can be neglected, \( y^k_i \) is also the earliest departure time of containers at transshipping node.

\[
y^k_i \leq u^i_s \cdot x^h_{js} + M \cdot (1 - x^h_{js}) \quad \forall k \in K, (i, j) \in A, \forall s \in \Gamma_{ij}, \quad (15)
\]

Eq. (15) is the rail service operation cutoff time constraint. It ensures that the arrival time of the containers of transportation order \( k \) at node \( i \) should not exceed the operation cutoff time of rail service \( s \) at the same node, if this rail service has to be used in order to move these containers on arc \((i, j)\).

\[
\left( \max \left\{ p^r_i - f^i_{js} - r^s, 0 \right\} - \omega^i_{js} \right) \cdot x^h_{js} = 0 \quad \forall k \in K, (i, j) \in A, \forall s \in \Gamma_{ij}, \quad (16)
\]
\[
\begin{align*}
\left( \max \left\{ x^k_i - y^k_i - r^k_i, 0 \right\} - w^k_{ij} \right) \cdot x^k_{ij} \\
= 0 & \quad \forall k \in K, \forall (i, j) \in A, \forall s \in \Omega_{ij}.
\end{align*}
\]  

(17)

Eqs. (16) and (17) compute the charged inventory time of the containers of transportation order \( k \) at node \( i \) before being moved on arc \((i, j)\) by a rail service and a road service, respectively. In (17), especially if \( i = o_k \) and \( s \in \Omega_{ij} \), \( r^k_i = 0 \).

\[ n^k_i \in \{0, 1, 2, \ldots \} \quad \forall k \in K, \forall i \in N, \]  

(18)

\[ t^k_{ij} \geq 0 \quad \forall k \in K, (i, j) \in A, \forall s \in \Omega_{ij}, \]  

(19)

\[ w^k_{ij} \geq 0 \quad \forall k \in K, \forall (i, j) \in A, \forall s \in S_{ij}, \]  

(20)

\[ x^k_{ij} \in \{0, 1\} \quad \forall k \in K, \forall (i, j) \in A, \forall s \in S_{ij}, \]  

(21)

\[ y^k \geq 0 \quad \forall k \in K, \forall i \in N, \]  

(22)

\[ z^k \geq 0 \quad \forall k \in K, \forall i \in N \setminus \{d_k\}. \]  

(23)

Eqs. (18), (19), (20), (21), (22), and (23) are the variable domain constraints.

The above mathematical model for the multimodal routing problem is summarized in Figure 4. Figure 4 classifies all equations (excluding variable domain constraints) into three categories, including "Modelling the network" which realizes the methodology proposed in Section 3. "Setting the objective" which reflects that satisfying customer demands is the core of the routing problem, and "Basic assumptions" which are the basic conditions of the model. This should help the readers to better understand the functions and roles of various equations in the mathematical model.

By using the solution strategy designed in the next section, the proposed model can be used to optimize any green multimodal routing problems that are composed of scheduled rail services with uncertain capacities and time-flexible road services with uncertain travel times. It can plan the best multimodal routes to accomplish multiple transportation orders with soft due date time windows.

5. An Exact Solution Strategy

5.1. Determining the Solution Strategy. The proposed mathematical model in Section 4.2 is easily understood, because it is established straightforwardly according to the settings of the specific multimodal routing problem. However, the mathematical model contains three categories of nonlinearity, including (1) fuzzy chance constraint as (8), (2) piecewise linear constraint as (12), and (3) general nonlinear components as (3), (4), (13), (14), (16), and (17). Such nonlinearity increases the difficulty of the problem to be effectively solved by the exact solution algorithms (e.g., branch-and-bound algorithm and simplex algorithm). Especially for the large-scale real-world problem, due to the nonlinearity, the solution to the problem will usually get into the local optimum, and massive computational time will be consumed to generate the solution. Therefore, if we can eliminate the nonlinearity of the model by designing the equivalent linear reformulations to the nonlinear equations in the model, then the problem can be solved with the help of the mathematical programming software (e.g., LINGO, CPLEX, and GAMS) where exact solution algorithms can be implemented efficiently.

We devote this study to developing a linearization-based exact solution strategy to address the multimodal routing problem. Although it is widely acknowledged that heuristic algorithms have better performance than the exact solution algorithms in dealing with the multimodal routing, which is known as an NP-hard problem, exact solution algorithms have two significant advantages. The first one is that they can test if the model observes the mathematical logic. The second is that they can provide an exact benchmark to verify the efficiency and accuracy of the heuristic algorithms [67]. Fazayeli et al. [68] also highlighted that the development of an exact algorithm and the comparison between the algorithms will be their future work for the location-routing problem on a multimodal transportation network.

5.2. Model Linearization Technique

5.2.1. Crisp Equivalent and Linearization of the Fuzzy Chance Constraint. From the definition of the fuzzy credibility, (24) can be obtained where \( a \) is a deterministic number while \( \bar{b} \) is a triangular fuzzy number described by \((b_1, b_2, b_3)\) and \( b_1 < b_2 < b_3 \). [69]:

\[
\text{Cr}\left\{a \leq \bar{b}\right\} = \begin{cases} 
1, & \text{if } a \leq b_1 \\
\frac{2b_2 - b_1 - a}{2(b_2 - b_1)}, & \text{if } b_1 \leq a \leq b_2 \\
\frac{b_3 - a}{2(b_3 - b_2)}, & \text{if } b_2 \leq a \leq b_3 \\
0, & \text{if } a \geq b_3
\end{cases}.
\]  

(24)

According to (24), we can gain the crisp equivalency to the fuzzy chance constraint (8) by replacing \( a \) with \( \sum_{k \in K} q_k \cdot x^k_{ij} \) and \( \bar{b} \) with \( v_{ij} \). However, the crisp equivalency is a piecewise linear function which is still essentially a nonlinear formula. Thus, we conduct following modifications shown as (25) and (26) to transform it into a pure linear function. After such modification, during the simulation process, we can use (25) to substitute (8) if \( a \in [0.5, 1] \). (26) to substitute the same equation otherwise.

\[
2(1 - \alpha) \cdot v^M_{ij} + (2\alpha - 1) \cdot v_{ij}^{min} \geq \sum_{k \in K} q_k \cdot x^k_{ij} \quad \text{if } 0.5 \leq \alpha
\]  

\[
\leq 1, \forall (i, j) \in A, \forall s \in \Gamma_{ij},
\]  

(25)

\[
2\alpha \cdot v^M_{ij} - (2\alpha - 1) \cdot v_{ij}^{max} \geq \sum_{k \in K} q_k \cdot x^k_{ij} \quad \text{if } 0 \leq \alpha
\]  

\[
\leq 0.5, \forall (i, j) \in A, \forall s \in \Gamma_{ij},
\]  

(26)

5.2.2. Linearization of the Piecewise Linear Function. Similarly to (24), the piecewise linear function \( f_{ij}(\cdot) \) is also a nonlinear expression where \( \langle \cdot \rangle \) represents \( \langle \cdot \rangle - 24 \cdot n^k \) for the sake of conciseness. Let \( P_{ij} \) represent the set of the linear segments of \( f_{ij}(\cdot) \), \( p \) the index of the segments \( (p = 1, \ldots, |P_{ij}|) \),
and \([tt_{ij}^{kp}, tt_{ij}^{kp+}]\) the departure time interval corresponding to the linear segment \(\ell_{ij}^{kp}(\cdot)\). Then, \(f_{ij}(\cdot)\) has a universal formulation shown as (27).

\[
\begin{align*}
    f_{ij}(\cdot) &= \begin{cases}
        f_{ij}^{0}(\cdot), & \ell_{ij}^{0} \subseteq [tt_{ij}^{kp}, tt_{ij}^{kp+}] \\
        \cdots \\
        f_{ij}^{p}(\cdot), & \ell_{ij}^{p} \subseteq [tt_{ij}^{p-}, tt_{ij}^{p+}] \\
        \cdots \\
        f_{ij}^{s}(\cdot), & \ell_{ij}^{s} \subseteq [tt_{ij}^{s-}, tt_{ij}^{s+}] \\
    \end{cases} \\
    & \forall (i, j) \in A, \forall s \in \Omega_{ij}.
\end{align*}
\]

(27)

To linearize the piecewise linear function (27), the following procedures are conducted.

**Step 1.** Define a 0-1 variable \(\varphi_{ij}^{kp}\), \(\varphi_{ij}^{kp} = 1\), if \((z_{ij}^{k} - 24 \cdot n_{i}^{k})\) falls within \([tt_{ij}^{kp}, tt_{ij}^{kp+}]\), \(\varphi_{ij}^{kp} = 0\) otherwise. And \(\varphi_{ij}^{p}\) should observe the following constraint.

\[
\sum_{p \in P_{ij}} \varphi_{ij}^{p} = 1 \forall k \in K, \forall (i, j) \in A, \forall s \in \Omega_{ij}.
\]

(28)

**Step 2.** Define two nonnegative variables \(\lambda_{ij}^{kp}\) and \(\lambda_{ij}^{kp+}\), and distribute them to the lower bound \(tt_{ij}^{p}\) and upper bound \(tt_{ij}^{p+}\) of the corresponding departure time interval \([tt_{ij}^{kp}, tt_{ij}^{kp+}]\) as their weights. Hence, we have the following equations.

\[
\begin{align*}
    z_{ij}^{k} - 24 \cdot n_{i}^{k} &= \sum_{p \in P_{ij}} \left( \lambda_{ij}^{kp} \cdot tt_{ij}^{p} + \lambda_{ij}^{kp+} \cdot tt_{ij}^{p+} \right) \forall k \in K, \forall (i, j) \in A, \forall s \in \Omega_{ij}, \\
    \lambda_{ij}^{kp} + \lambda_{ij}^{kp+} &= \varphi_{ij}^{kp} \forall k \in K, \forall (i, j) \in A, \forall s \in \Omega_{ij}, \forall p \in P_{ij}, \\
    \lambda_{ij}^{kp} \geq 0 \forall k \in K, \forall (i, j) \in A, \forall s \in \Omega_{ij}, \forall p \in P_{ij}, \\
    \lambda_{ij}^{kp+} \geq 0 \forall k \in K, \forall (i, j) \in A, \forall s \in \Omega_{ij}, \forall p \in P_{ij}.
\end{align*}
\]

(29)

**Step 3.** Modify the piecewise linear function as follows.

\[
\begin{align*}
    f_{ij} \left( z_{ij}^{k} - 24 \cdot n_{i}^{k} \right) &= f_{ij} \left[ \sum_{p \in P_{ij}} \left( \lambda_{ij}^{kp} \cdot tt_{ij}^{p} + \lambda_{ij}^{kp+} \cdot tt_{ij}^{p+} \right) \right] \\
    &= \sum_{p \in P_{ij}} f_{ij} \left( \lambda_{ij}^{kp} \cdot tt_{ij}^{p} + \lambda_{ij}^{kp+} \cdot tt_{ij}^{p+} \right) \\
    &= \sum_{p \in P_{ij}} f_{ij} \left( \lambda_{ij}^{kp} \cdot tt_{ij}^{p} + \lambda_{ij}^{kp+} \cdot tt_{ij}^{p+} \right) \\
    &= \sum_{p \in P_{ij}} \lambda_{ij}^{kp} \cdot f_{ij} \left( tt_{ij}^{p} \right) + \lambda_{ij}^{kp+} \cdot f_{ij} \left( tt_{ij}^{p+} \right),
\end{align*}
\]

(30)
where \( f^p_{ij}(t^p_{ij}) \) and \( f^p_{ij}(r^p_{ij}) \) for \( V(i,j) \in A, \forall s \in \Omega_{ij}, \forall p \in P_{ij} \) are all known. Finally, the pure linear equivalency to (12) is shown as (31).
\[
\begin{align*}
t^k_{ij} &= \sum_{p \in P_{ij}} \left[ \lambda^p_{ij} \cdot f^p_{ij}(t^p_{ij}) + \lambda^p_{ij} \cdot f^p_{ij}(t^p_{ik}) \right] \\
&\quad \cdot \forall k \in K, V(i,j) \in A, \forall s \in \Omega_{ij}.
\end{align*}
\]
(31) 5.2.3. Linearization of Other Nonlinear Components

(1) Linear Reformulation 1. Eq. (3) with nonlinear component \( \sum_{k \in K} c^d_{dest} \cdot q_k \cdot \max \{ t^k_i - y_{d_i} , 0 \} \) can be linearized by (32), (33), and (34) where \( g_k \) is a newly introduced variable.
\[
\begin{align*}
\sum_{k \in K} \sum_{i \in S_j} c^j_{i} \cdot q_k \cdot w^j_{ij} + \sum_{k \in K} c^d_{dest} \cdot q_k \cdot g_k &= \sum_{k \in K} c^d_{dest} \cdot q_k \cdot g_k \\
g_k &\geq t^k_i - y_{d_i} \quad \forall k \in K, \\
g_k &\geq 0 \quad \forall k \in K.
\end{align*}
\]
(32) (33) (34)

If the containers arrive at the destination earlier than \( t^k_i \), \( t^k_i - y_{d_i} > 0 \). Consequently, \( \max \{ t^k_i - y_{d_i} , 0 \} = t^k_i - y_{d_i} \). According to (32) and (33), \( g_k \geq t^k_i - y_{d_i} \). Minimization of (32) will then lead to \( g_k = t^k_i - y_{d_i} \). In this case, \( \sum_{k \in K} c^d_{dest} \cdot q_k \cdot \max \{ t^k_i - y_{d_i} , 0 \} \) is equivalent to (32) and (33).

If the containers arrive at the destination later than \( t^k_i \), \( t^k_i - y_{d_i} < 0 \). Consequently, \( \max \{ t^k_i - y_{d_i} , 0 \} = 0 \). According to (32) and (33), \( g_k \geq 0 \). Minimization of (32) will then lead to \( g_k = 0 \). In this case, \( \sum_{k \in K} c^d_{dest} \cdot q_k \cdot \max \{ t^k_i - y_{d_i} , 0 \} \) is also equivalent to (32) and (33).

Above all, the equivalency above is verified. Similarly, we have the following linear reformulation.

(2) Linear Reformulation 2. Nonlinear component \( c^p_{pen} \cdot \max \{ y_{d_i} - t^k_i , 0 \} \) in (4) can be linearized by (35), (36), and (37) where \( \sigma_k \) is also a newly introduced variable.
\[
\begin{align*}
\sum_{k \in K} c^p_{pen} \cdot \sigma_k &= \sum_{k \in K} c^p_{pen} \cdot \sigma_k, \\
\sigma_k &\geq y_{d_i} - t^k_i \quad \forall k \in K, \\
\sigma_k &\geq 0 \quad \forall k \in K.
\end{align*}
\]
(35) (36) (37)

(3) Linear Reformulation 3. Nonlinear constraint \( n^k = \lfloor x^k_i/24 \rfloor \forall k \in K, V(i,j) \in N \setminus \{ d_i \} \) (12) can be reformulated by linear constraints (38) and (39).
\[
\begin{align*}
n^k_i &\leq \lfloor s^k_i/24 \rfloor \forall k \in K, V(i,j) \in N \setminus \{ d_i \}, \\
n^k_i &\geq \lfloor s^k_i/24 \rfloor \forall k \in K, V(i,j) \in N \setminus \{ d_i \}.
\end{align*}
\]
(38) (39)

We take \( s^k_i = 60 \) as example to verify the equivalency above. According to (12), \( n^k_i = \lfloor 60/24 \rfloor = [2.5] = 2 \). Meanwhile, according to (45) and (46), there exists \( 1.5 < n^k_i < 2.5 \).

Because \( n^k_i \) is a nonnegative integer, finally \( n^k_i = 2 \). Consequently, the equivalency above is tenable.

(4) Linear Reformulation 4. By using the same linearization technique, nonlinear constraints \( (x^k + t^k_{ij} - y^k_i) \cdot x^k_{ij} = 0 \forall k \in K, V(i,j) \in A, \forall s \in \Omega_{ij}, \) (13) and \( (t^k_i - y^k_i) \cdot x^k_{hir} = 0 \forall k \in K, V(h,i) \in A, \forall r \in \Gamma_{hi} \) (14) can be reformulated by linear constraints (40) and (41) and (42) and (43), respectively.
\[
\begin{align*}
\sum_{k \in K} c^d_{dest} \cdot q_k \cdot g_k &= \sum_{k \in K} c^d_{dest} \cdot q_k \cdot g_k \\
g_k &\geq t^k_i - y_{d_i} \quad \forall k \in K, V(i,j) \in A, \forall s \in \Omega_{ij}, \\
g_k &\geq 0 \quad \forall k \in K,
\end{align*}
\]
(40) (41) (42)

(5) Linear Reformulation 5. By using the same linearization form, nonlinear constraints \( (\max \{ 0, x^k_i - y^k_i - \tau r \} - w^k_{ij}) \cdot x^k_{ij} = 0 \forall k \in K, V(i,j) \in A, \forall s \in \Gamma_{ij} \) (16) and \( (\max \{ 0, x^k_i - y^k_i - \tau s \} - w^k_{ij}) \cdot x^k_{ij} = 0 \forall k \in K, V(i,j) \in A, \forall s \in \Omega_{ij} \) (17) can be reformulated by linear constraints (44) and (45) and (46) and (47).
\[
\begin{align*}
w^k_{ij} &\geq M \cdot (x^k_{ij} - 1) + (t^k_i - y^k_i - \tau r) \\
&\quad \forall k \in K, V(i,j) \in A, \forall s \in \Gamma_{ij}, \\
w^k_{ij} &\leq M \cdot x^k_{ij} \forall k \in K, V(i,j) \in A, \forall s \in \Gamma_{ij}, \\
w^k_{ij} &\leq M \cdot (x^k_{ij} - 1) + (x^k_i - y^k_i - \tau r) \\
&\quad \forall k \in K, V(i,j) \in A, \forall s \in \Omega_{ij}, \\
w^k_{ij} &\leq M \cdot x^k_{ij} \forall k \in K, V(i,j) \in A, \forall s \in \Omega_{ij}.
\end{align*}
\]
(44) (45) (46) (47)

For detailed proofs of Linear Reformations 4 and 5, readers can refer to our previous study [2].

5.3. Improved Linear Reformulation. For the sake of readability, the improved linear reformulation derived from the mathematical model in Section 4.2.2 is summarized as follows.
Objective function:
Minimize (1)+(2)+(32)+(35)+(3)
Subject to:
Constraints (6), (7), (9), (10), (15), (18–23), (25), (26), (28), (29), (31), (33), (34), (36–47)
6. Empirical Case Study Based on the Chinese Scenario

6.1. Case Description. In this section, we present a real-world case to demonstrate the feasibility of the proposed model and exact solution strategy in optimizing the practical problem. This empirical case is oriented on a Chinese scenario. In this scenario, several consignments of containers need to be transported from the western inland cities (e.g., Lanzhou and Hohhot) to the eastern sea ports (e.g., Qingdao) through the road-rail multimodal service network.

In the multimodal service network, “road services” refer to the container trucks and “rail services” refer to the block container trains. The topological structure of the multimodal service network is shown in Figure 5. In the multimodal service network, there are 10 periodically operated road services (block container trains) and 15 road services (container trucks). The schedules of the rail services and the time-dependent travel times of the road services are separately presented in Appendix A (available here) and Appendix B (available here) in the supplementary material. Note that all the values in Appendix A (available here) and Appendix B (available here) are expressed as real numbers; for example, 10:30 is rewritten as 10.5 and 5:00 at the second day of the planning horizon is 29 (24+5). Other times in different days can be expressed in the same way.

There are 10 transportation orders within a short-term planning horizon. The containers in these orders need to be transported from Lanzhou and Hohhot (inland cities in Northwest China) to Jiaozhou and Huangdaoqang (sea ports in Qingdao, Shandong Province) and Lianyungang (a sea port in Jiangsu Province) through the multimodal service network shown in Figure 5. In the multimodal service network, there are 10 periodically operated road services (block container trains) and 15 road services (container trucks). The schedules of the rail services and the time-dependent travel times of the road services are separately presented in Appendix A (available here) and Appendix B (available here) in the supplementary material. Note that all the values in Appendix A (available here) and Appendix B (available here) are expressed as real numbers; for example, 10:30 is rewritten as 10.5 and 5:00 at the second day of the planning horizon is 29 (24+5). Other times in different days can be expressed in the same way.

There are 10 transportation orders within a short-term planning horizon. The containers in these orders need to be transported from Lanzhou and Hohhot (inland cities in Northwest China) to Jiaozhou and Huangdaogang (sea ports in Qingdao, Shandong Province) and Lianyungang (a sea port in Jiangsu Province) through the multimodal service network shown in Figure 5. In the multimodal service network, there are 10 periodically operated road services (block container trains) and 15 road services (container trucks). The schedules of the rail services and the time-dependent travel times of the road services are separately presented in Appendix A (available here) and Appendix B (available here) in the supplementary material. Note that all the values in Appendix A (available here) and Appendix B (available here) are expressed as real numbers; for example, 10:30 is rewritten as 10.5 and 5:00 at the second day of the planning horizon is 29 (24+5). Other times in different days can be expressed in the same way.

6.2. Parameter Setting. The unit transportation costs en route (unit: ¥/TEU) regarding the two transportation services are determined by (48).

\[ c_{ij}^k = \begin{cases} 
  c_{rail}^2 + c_{rail}^1 \cdot d_{ij} & s \in \Gamma_{ij} \\
  c_{road}^2 \cdot d_{ij} & s \in \Omega_{ij}
\end{cases} \quad \forall k \in K, \forall (i, j) \in A, \forall s \in S_{ij}. \tag{48} \]

As can be seen in (48), the unit transportation costs for rail services en route include two parameters: \( c_{rail}^1 \) which is the costs related to the container volume (unit: ¥/TEU) and \( c_{rail}^2 \) which is the costs related to the transportation distance, loading/unloading operation costs, inventory costs, penalty costs, and CO\(_2\) emission costs, which also correspond to (1) to (5) in the objective function.

Similarly, by using the data from 2015 China Transportation Industry Development Statistical Bulletin [73], we can also gain the CO\(_2\) emission factor of road services which is 1064 g/TEU-km. The values of all the parameters in the empirical case are summarized as shown in Table 3. In addition, the values of the unit inventory costs of transportation services at the nodes (\( c^i_s \)) and the free-of-charge inventory periods of services at the nodes (\( \tau^i_f \)) are all presented in Figure 5. When the nodes are specifically the railway stations, all the corresponding \( c^i_s = 3.125 ¥/TEU \) and \( \tau^i_f = 48 h \). When the nodes are the origins of the transportation orders, \( c^i_s \) varies based on the locations of the nodes and \( \tau^i_f = 0 h \) if \( s \in \Omega_{ij} \).

6.3. Simulation Environment. Described by the linearized programming model shown in Section 5.3, the empirical case can be solved by any exact solution algorithms, for example, branch-and-bound algorithm and simplex algorithm. In this study, we adopt the branch-and-bound algorithm to solve the problem. The algorithm is conducted by the mathematical programming software LINGO 12 designed by LINDO Systems Inc., Chicago, IL, USA [75], on a ThinkPad Laptop with Intel Core i5-5200U 2.20 GHz CPU 8 GB RAM. The scale of the empirical case is indicated by Table 4.

6.4. Optimization Result and Multimodal Route Illustration. The optimization result of the empirical case and the performance of the exact solution strategy are shown in Table 5.

The structure of the minimal generalized costs for the multiple transportation orders is illustrated by Figure 6 where C1 to C5 are successively the transportation costs en route, loading/unloading operation costs, inventory costs, penalty costs, and CO\(_2\) emission costs, which also correspond to (1) to (5) in the objective function.

The best routes for the containers in the multiple transportation orders are given in Table 6. According to the planned routes, the accomplishments of 4 transportation orders satisfy their due date time windows exactly.

6.5. Case Discussion

6.5.1. Sensitivity of the Multimodal Routing with respect to the CO\(_2\) Emissions. First of all, we analyze the effect of the CO\(_2\) emissions on the multimodal routing. Let the unit...
emission costs vary from 50 to 100 ¥/ton, the sensitivity of the multimodal routing result with respect to this variation can be seen in Figure 7.

Figure 7 shows that the CO₂ emission costs increase linearly due to the fact that the emission volume is a constant (87.1 tons), while the rest of the generalized costs stay constant with the unit emission costs increasing. Even though we continue the simulation until the unit emission costs reach up to 1000 ¥/ton, which is obviously infeasible in practice (according to the formulation of the Chinese Academy of Environmental Planning [73], the feasible unit emission costs should reach up to 100 ¥/ton at 2030), the sensitivity still keeps the same tendency as shown in Figure 6. It means that the best routes in the empirical case do not modify when the emission costs vary and are hence insensitive to this variation.

Therefore, charging for CO₃ emissions within a reasonable range is not particularly helpful to promote the green multimodal transportation, at least in this empirical case.

6.5.2. Bi-Objective Optimization regarding CO₂ Emissions. Furthermore, we modify the proposed model into a bi-objective optimization model with the objectives of minimizing the generalized costs (Z₁ = Eq.1 + Eq.2 + Eq.3 + Eq.4) and of minimizing the CO₂ emissions (Z₂ = Eq.51).

\[
\begin{align*}
\text{minimize} & \quad \sum_{k \in K} \sum_{(i,j) \in A} \sum_{s \in S_{ij}} \text{em}_s \cdot q_k \cdot d_{ij} \cdot x_{ij}^k, \\
\end{align*}
\]

Using the widely utilized weighted sum method [43] defined in (51) and (52), we can solve the bi-objective
optimization problem and generate its Pareto solutions shown in Figure 7.

$$
\text{minimize} \quad \omega \cdot Z_1 + (1 - \omega) \cdot Z_2, \quad (51)
$$

$$
0 < \omega \leq 1. \quad (52)
$$

Especially, in Figure 8, compared with the last Pareto solution, the second last one can lower the CO_2 emissions from 87.1 to 83.7 by 3.90% and cause only a slight increase of the generalized costs from 1571.2 to 1586.0 by 0.94%. By comparing Figure 8 with Figure 7, we can conclude that a bi-objective optimization approach is much more effective than charging for CO_2 emissions in helping decision makers to control and optimize the CO_2 emissions in the multimodal service network.

6.5.3. Sensitivity of the Multimodal Routing with respect to the Confidence. Confidence $\alpha$ in the fuzzy chance constraint (8) is set by the decision makers and reflects their preference for the reliability of the multimodal routing such that all the loaded volume of a rail service does not exceed its carrying capacity. Its value has a remarkable effect on the optimization result of the multimodal routing, which can be investigated by using sensitivity analysis. Let $\alpha$ vary from 0.1 to 1.0 with a step size of 0.1, we can obtain the sensitivity of the multimodal routing result (indexed by the minimal generalized costs) with respect to such variation, which can be seen in Figure 9.

We can also draw some helpful insights from Figure 9 summarized as follows.

(1) Overall, larger values of confidence $\alpha$ will lead to larger generalized costs for the best routes.

(2) The variation of the generalized costs with respect to the confidence is stepwise. In this case study, the multimodal routing is sensitive to the confidence when its value exceeds 0.3.

6.5.4. Fuzzy Simulation. During the decision-making process regarding the multimodal routing, it is challenging for decision makers to determine an objective confidence value. In order to help them select reasonable confidence values, fuzzy simulation is adopted in this study.

The fuzzy simulation simulates the actual deterministic transportation scenario by randomly generating the carrying capacities of rail services according to the membership degree of the triangular fuzzy numbers. The simulation process is shown in Figure 10 [8].

After the simulation, we can first gain deterministic carrying capacities $v_{ij}$ for $(i, j) \in A, s \in \Gamma_{ij}$, Eq. (8) can be consequently transformed into a deterministic carrying capacity constraint as (53).

$$
\sum_{k \in K} q_k \cdot x^k_{ij} \leq v_{ij} \quad \forall (i, j) \in A, \forall s \in \Gamma_{ij}. \quad (53)
$$

Then, we can check if the planned routes given by the fuzzy programming model satisfy (15) or not. If the planned routes satisfy the constraint, we define them as a successful plan; otherwise, a failed plan which means that the planned routes are infeasible due to the violation of the capacity constraint when adopted to move containers in the practical transportation. If the transportation starts and a plan is found to be infeasible due to the insufficient actual capacities of some rail services, an alternative plan should be generated according to the following principle in order to reduce the additional costs and CO_2 emissions caused by the alternative: the rail service with insufficient capacity should be preferentially used by the transportation orders with larger container volumes initially assigned to it as many as possible on the condition that (53) must be satisfied, that is, the transportation orders with larger container volumes should
be accomplished by the initial plan as far as possible, and the rest transportation orders usually with smaller container volumes assigned to the rail service should be accomplished by the origin-to-destination road services. The fuzzy simulation should be conducted several times. In this study, we ran the simulation 50 times, and the fuzzy simulation result is shown in Figure 11.

As we can see from Figure 11, the successful times of the planned routes increase remarkably when the confidence value changes from 0.7 to 0.8. Although the generalized costs increase approximately by 1.29%, the success ratio of the planned routes considerably increases from 26% to 100% by more than 3.8 times when the confidence value is set to 0.8 instead of 0.7.

Note that when $\alpha = 0.5$, according to (25) and (26), the fuzzy chance constraint can be rewritten as $\sum_{k \in K} q_k \cdot x^M_{ij} \leq v_{ij}^M$ which is the deterministic capacity constraint widely employed by the current literature (e.g., Sun and Lang [2] and Sun et al. [19]). So we know that the successful ratio of the multimodal routing with deterministic capacity consideration is only 12% in the 50 simulations for the specific case. As we can see from Figure 11, the successful ratio of the multimodal routing can be significantly improved by more than 8.3 times by considering the capacity uncertainty and setting the confidence to 0.8, 0.9, or 1.0. Therefore, it is of great value to formulate the capacity uncertainty when planning the multimodal routes, so that higher reliability of the multimodal routing decision-making can be guaranteed.

We can further achieve the optimization results of the 50 deterministic problems formulated by the linearized model with (25) and (26) replaced by (53). The optimization results are presented in Appendix D (available here) in the Table 5: Optimization result and strategy performance.

<table>
<thead>
<tr>
<th>Solution algorithm</th>
<th>Solver state</th>
<th>Best solution</th>
<th>Computational time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard branch-and-bound algorithm</td>
<td>Global optimum</td>
<td>¥ 1,575,559</td>
<td>1 min 06 sec*</td>
</tr>
</tbody>
</table>

*Average value of 10 times simulation.

Figure 6: Structure of the minimal generalized costs.

Table 6: Best routes in the empirical case.

<table>
<thead>
<tr>
<th>Number</th>
<th>Best multimodal routes</th>
<th>Arrival time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lanzhou [28] — (14) — Xinzhu — (4) — Jiaozhou</td>
<td>130.4</td>
</tr>
<tr>
<td>2</td>
<td>Lanzhou [37] — (14) — Xinzhu — (4) — Jiaozhou</td>
<td>130.4**</td>
</tr>
<tr>
<td>3</td>
<td>Lanzhou [111] — (14) — Xinzhu — (5) — Huangdaogang</td>
<td>214.6</td>
</tr>
<tr>
<td>5</td>
<td>Lanzhou [90] — (17) — Putian — (10) — Lianyungang</td>
<td>175**</td>
</tr>
<tr>
<td>6</td>
<td>Hohhot — (7) — Xingang [96] — (23) — Jiaozhou</td>
<td>106.9**</td>
</tr>
<tr>
<td>7</td>
<td>Hohhot — (7) — Xingang [101.5] — (23) — Jiaozhou</td>
<td>106.9</td>
</tr>
<tr>
<td>8</td>
<td>Hohhot — (7) — Xingang [148] — (23) — Jiaozhou</td>
<td>72*</td>
</tr>
<tr>
<td>9</td>
<td>Hohhot — (7) — Xingang [192] — (23) — Jiaozhou</td>
<td>72*</td>
</tr>
</tbody>
</table>

*Early delivery; **delayed delivery; numbers in {}: planned departure time.

Table 5: Optimization result and strategy performance.
Figure 7: Sensitivity of the multimodal routing with respect to the CO$_2$ emissions.

Figure 8: Pareto solutions to the bi-objective optimization problem.

Figure 9: Sensitivity of the multimodal routing with respect to the confidence.
supplementary material. Such results can be considered as the actual best routes. By comparing the actual best routes with the planned ones, we can find the confidence value that best matches the actual transportation.

In the 50 fuzzy simulations, the best actual routes whose generalized costs are equal to ¥1,575,559 emerge 38 times that account for 76% of the entire simulation times (see in Figure 12). That is to say, such routes are the most likely best ones in the practical transportation. The planned best routes given by the fuzzy programming model with confidence values of 0.8, 0.9, and 1.0 match the most likely best actual routes better than others. Above all, the best values of the confidence in this empirical case are recommended to be 0.8, 0.9, and 1.0.

**6.6. Managerial Implications.** It should be noted that the optimization results and conclusions in this section are drawn based on a specific case whose setting is shown in Sections 6.1 and 6.2. The results and conclusions are not universal and are sensitive to the case itself and might change if the case changes. Therefore, the above analysis and discussion in this section only provide a procedure and a demonstration on how to adopt the proposed methods to deal with the practical problems. Although the cases might be different, we can always use such procedure to address them to draw corresponding optimization results and conclusions.

Despite this fact, the results show important implications for the practical implementation of the model:

1. Although the multimodal service network is complex, the linearized optimization model can deliver the optimal solution in a very short time so that new plans for incoming orders can be found very fast.
2. The objective of the optimization is to minimize the total costs consisting of different cost categories
(i.e., transportation, transshipment, inventory, penalty, and emission costs). However, there is a clear dominance of transportation costs over all other cost categories which makes it difficult to influence the optimal solution by changing other cost factors than transportation costs.

(3) Based on the previous implication, it is clear that the charging method for CO$_2$ emissions does not improve the environmental sustainability of the multimodal routes if realistic values are used for emission costs. Therefore, rather a bi-objective approach should be used in order to be able to compare the cost-optimal and emission-optimal solution. In this case, often a solution can be found where one objective is significantly improved whereas there is only a small increase for the other one.

(4) The minimal values of total costs can be usually achieved under the deterministic setting since no buffer times or capacities need to be included. However, uncertainty plays an important role in practice and therefore should be considered also in planning. Although the total costs are increasing with the increasing confidence level, this also reduces the risk of disruptions during transportation which can be seen in a lower ratio of failed plans. As a result, more transportations can be performed as originally planned and the need for re-planning for failed plans is minimized. Since such re-planning usually leads to additional costs, reducing the share of failed plans is of high interest.

7. Conclusions

Nowadays, the multimodal transportation has been widely promoted all over the world, and its routing optimization has attached great importance by both researchers and practitioners. Meanwhile, the public concern on environmental protection raises, and environmentally friendly development gets tremendous highlights in almost every industry. Therefore, it is valuable to combine multimodal routing with the environmental views. In addition to that, different uncertainty factors can influence the performance of multimodal transportation and therefore should be considered already in the planning phase. However, all these factors increase the complexity of multimodal routing modelling and therefore make the transportation planning very challenging in practice.

In order to respond to these challenges, we developed a nonlinear mixed-integer optimization model covering multiple objectives (i.e., economic and environmental) and multiple uncertainty factors (i.e., traffic congestion and rail service capacity). Moreover, in order to facilitate the solution of this problem, a linearization approach was proposed. The application of the proposed model to a real-world case study showed that it can deliver optimal results in a relatively short time which is very helpful for the practical transportation planning. Besides that, it was shown that the inclusion of environmental criteria is meaningful only when the bi-objective method is used, since the weight of the emission costs is very low in comparison to transportation costs if the charging method is used. Regarding the reliability of transportation, including the uncertainty factors into planning and increasing confidentiality levels significantly reduce the risk of disruptions during the execution phase but also lead to increased total costs. However, the reduced risk of disruption also reduces the need for re-planning and therefore helps to avoid additional costs connected with this process.

Compared with the models proposed by other studies listed in the literature review section, our model can comprehensively formulate the issues of road traffic congestion and rail service capacity uncertainty that actually exist in the routing decision-making process and thus have better feasibility in dealing with the practical problem. Another improvement is that by modifying the objective function, our model can easily use two different approaches, including charging for CO$_2$ emissions and bi-objective optimization, to separately optimize the CO$_2$ emissions and identify which one is more effective in planning green routes. Although the model is nonlinear, it can be linearized, which enables the problem to be effectively solved by any exact solution algorithm on any mathematical programming software.

This study mainly focuses on the modelling and the utilization of a classical exact solution algorithm based on a linearized model and does not involve any intelligent algorithms. The exact solution strategy can efficiently optimize the presented case, but its performance is doubtful when the scale of the case is increasingly expanded. As a result, for further improvement to this study, we will test the feasibility of the exact solution strategy in dealing with a larger-scale empirical case. If the test result is not satisfactory, we will try to design some advanced intelligent solution algorithms (e.g., heuristic algorithms) with higher solution efficiency, so that the problem can be effectively optimized and routing suggestions can be provided to the decision makers in time when the problem scale gets larger.
Conflicts of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This study was supported by the Ernst Mach Scholarship financed by the Eurasia Pacific Uninet on behalf of the Austrian Federal Ministry of Science, Research and Economy (BMWFFW) under Reference no. ICM-2016-04319, the National Natural Science Foundation of China under Grant nos. 71390332-3 and 71501111, the Shandong Provincial Social Science Planning of China under Grant no. 16DGJL06, and the Shandong Provincial Natural Science Foundation of China under Grant no. ZR2017BG010.

Supplementary Materials

Appendix A: schedules of the rail services in the multimodal service network. Appendix B: travel distances and times of the road services in the multimodal service network. Appendix C: information on the transportation orders in the empirical case. Appendix D: optimization results of the fuzzy simulations. (Supplementary Materials)

References


[23] J. Jiang and J. Lu. “Research on Optimum Combination of Transportation Modes in the Container Multimodal


