IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 21, 2020, accepted April 15, 2020, date of publication April 30, 2020, date of current version May 14, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2990751

Short-Term Truckload Spot Rates' Prediction in Consideration of Temporal and Between-Route Correlations

WEI XIAO[1,2,3], CHUAN XU[1,2,3], HONGLING LIU[1], HONG YANG[4], AND XIAOBO LIU[1,2,3]

[1]School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, China
[2]National United Engineering Laboratory of Integrated and Intelligent Transportation, Southwest Jiaotong University, Chengdu 610031, China
[3]National Engineering Laboratory of Integrated Transportation Big Data Application Technology, Chengdu 611756, China
[4]Department of Computational Modeling and Simulation Engineering, Old Dominion University, Norfolk, VA 23529, USA

Corresponding author: Chuan Xu (xuchuan@swjtu.edu.cn)

This work was supported in part by the Key Research and Development Plan of Ministry of Science and Technology, China, under Grant 2018YFB1601402, in part by the Energy Foundation through the Project: Preliminary Investigation and Research on Freight Industry under Grant VQ99QT1136Y17005, and in part by the Open Research Fund for National Engineering Laboratory of Integrated Transportation Big Data Application Technology under Grant CTBDAT201912.

ABSTRACT Truckload spot rate (TSR), defined as a price offered on the spot to transport a certain cargo by using an entire truck on a target transportation line, usually price per kilometer-ton, is a key factor in shaping the freight market. In particular, the prediction of short-term TSR is of great importance to the daily operations of the trucking industry. However, existing predictive practices have been limited largely by the availability of multilateral information, such as detailed intraday TSR information. Fortunately, the emerging online freight exchange (OFEX) platforms provide unique opportunities to access and fuse more data for probing the trucking industry. As such, this paper aims to leverage the high-resolution trucking data from an OFEX platform to forecast short-term TSR. Specifically, a lagged coefficient weighted matrix-based multiple linear regression modeling (Lag-WMR) is proposed, and exogenous variables are selected by the light gradient boosting (LGB) method. This model simultaneously incorporates the dependency between historical and current TSR (temporal correlation) and correlations between the rates on alternative routes (between-route correlation). In addition, the effects of incorporating temporal and between-route correlations, time-lagged correlation and exogenous variable selection in modeling are emphasized and assessed through a case study on short-term TSR in Southwest China. The comparative results show that the proposed Lag-WMR model outperforms autoregressive integrated moving average (ARIMA) model and LGB in terms of model fitting and the quality and stability of predictions. Further research could focus on rates' standardization, to define a practical freight index for the trucking industry. Although our results are specific to the Chinese trucking market, the method of analysis serves as a general model for similar international studies.

INDEX TERMS Freight transportation, truckload spot rates, lagged weighted matrix, short-term prediction, weighted multiple regression, trucking economy.

I. INTRODUCTION

Trucking is a critical component of freight transportation in many countries. In the United States (US), for example, the trucking industry generated $700 billion in economic activity in 2017 [1]. Likewise, China's trucking market is estimated at more than $750 billion [2]. A salient feature of the trucking industry is its high fragmentation. For example, China gradually dismantled its state-owned trucking companies since opening up its markets after 1978 [3]. China's trucking industry has been largely privatized and has become highly fragmented and competitive. For example, most trucking companies in the country today are owner-operated, with 70% owning just one truck [4]. The consolidation of many small owner-operators reduces bargaining power and limits access to information (e.g., about where to find high-value loads). The rise of online freight exchange (OFEX) platforms, such as Uber Freight, Truck Alliance, and Convoy,

The associate editor coordinating the review of this manuscript and approving it for publication was Anand Paul.

offers a promising solution to improve the efficiency of the industry through information sharing and consolidation. Such a system can bring shippers and carriers together to exchange freight information, and provide value-added services, such as designing shipping tours for carriers [5]. Currently, shippers cover transportation orders either through long-term contracts or on-the-spot markets [6]. Shippers, in most cases, post transportation requirements on an online platform (e.g., OFEX). To guide truckers to high-value loads, predicting truckload spot rate (TSR), a price offered on the spot to transport a certain cargo by using an entire truck on a target transportation line (i.e., a route connecting two transshipment points). The question that we aim to answer in this paper is how well short-term route-specific TSR can be predicted.

This question is related to understanding not only the short-term forecasting model but also the behavior of the trucking industry, which has become a hot topic of research in recent years [7]–[11]. The driver of this enthusiasm is the availability of an unprecedented amount of data, such as the type of truck required, truck size, the type of goods, deal prices, the order-posted time, and truckers' current locations—made available through online trading through OFEX platforms.

This paper aims to leverage real-world data and advanced predictive models to offer a new means of modeling short-term TSR. Specifically, a lagged coefficient weighted matrix-based multiple linear regression (Lag-WMR) modeling approach with the selection of exogenous variables by the light gradient boosting (LGB) method is proposed to predict the short-term TSR. Its predictive performance is compared with the traditional time series forecasting approaches and machine learning methods.

This investigation is distinct from the literature because it contributes to the multifaceted investigations into TSR forecasting and freight econometrics. First, this paper demonstrates the existence of correlations between the rates on alternative routes (between-route correlation) and presents a temporal approach for addressing the issue in the short-term predictive analysis by considering the between-route correlation and the dependency between historical and current TSR (temporal correlation). Second, the conditions that augment the performance of the proposed model are discussed. An in-depth analysis of the exogenous factors that influence predictive performance can help practitioners choose appropriate variables and predictive models. Third, we detail the specification of a lagged weighting matrix for forecasting modeling (e.g., how routes are inter-related). Instead of using traditional weighting matrices, a lagged weighting matrix is proposed and compared with the matrix calculated based on the Pearson correlation coefficient. The proposed weighting matrix helps account for more temporal characteristics of the between-route correlation of the TSR.

II. LITERATURE REVIEW

In freight rate forecasting, the objects are typically classified into two categories: spot freight rate and contract freight rate. A contract rate is the price that a carrier and the third-party logistic (3PL) agree on to move a shipper's freight in a set lane over a set period, and a spot rate is a price that a shipper offers on the spot to move a load from point A to point B. It can be easily understood that a contract rate is a long-term route-based price with non-binding, whereas a spot rate is more a concern of an immediate shipment order [6], which is so dynamically based on market conditions that they can change over the course of a day. When addressing a forecasting issue, it is easier to reach much higher accuracy and less forecasting variability for a contract than spot rates. Additionally, when referring to improving the accuracy of prediction, for a contract rate, we could increase the model's performance by adding volume, rates of adjacent routes, and retraining of the model, whereas for spot rate, there is no general method to improve forecasting accuracy. For example, if adding the past value of the contract rate, the model updated with the new information may not improve yet because of barely short-term information transmission between the spot and contract rate.

One set of studies has investigated the pricing of less-than-truckload (LTL) shipments, which are less than 10,000 pounds. Baker [12] examines LTL carriers' practices for 64 city-pairs based on published freight rates collected from 24 large LTL carriers concerning the exploration of the variance of net rates that carriers charge for transportation between the same two cities, the extent of discounts that carriers offer for shipments, and practices regarding carries' class rate. Smith *et al.* [13] use data from large US LTL carriers to estimate a regression model to predict revenue for different customers in different lanes with transportation characteristics of shipment information (e.g., cargo density, shipped weight) Kay and Wasring [10] use publicly available data to estimate a nonlinear regression model concerning an investigation of tariff-based rates. Özkaya *et al.* [14] use LTL market data from Schneider Logistics to develop a regression model to estimate LTL market rates.

A second set of studies has examined factors that affect the pricing on the full truckload (also known as truckload) spot market. Caldwell [9] develops a regression model to estimate the extent that lead time and transportation factors (e.g., distance, origin, and destination of load) affect the full truckload shipment price. Lindsey *et al.* [15] estimate a regression model with spot market transaction data provided by a third-party carrier to predict lane/shipment-level spot prices. This model considers factors including lane (e.g., traffic congestion along the route, lane miles) and shipment (e.g., distance, type of freight, national transport) predictors. Scott [16] explores how shipment participants' (shippers and carriers) factors influence the price premium for a spot market shipment by using a private transaction dataset from a large national shipper. Budak *et al.* [8] implement an artificial neural network and quantile regression to forecast a TSR based on Turkey highway data. They apply the model in a route-based approach and general approach to forecast spot marketing rates that incorporate both shipment (e.g., load, freight type), carriers (e.g., vehicle type), and market (e.g., unit price of fuel, month variable). Bai [7] develops a nonlinear regression model that incorporates predictor variables to forecast lane-level spot and contract rates, with a comparison that uses the autoregressive integrated moving average (ARIMA) model and conventional nonlinear regression model. Miller [17] develops an ARIMA framework to develop forecasts for three time series of monthly archival truckload trucking prices by using two public data sources, which also examines the evolution of price based on the temporal dynamics of the freight market.

This paper's contributions differ from the aforementioned studies. The first contribution shifts from considering the same predictor variables set for each lane, to incorporate the importance of variable assessment in a regression model to select the true factors that actually influence the route-specific TSR, because the market rate of each lane may differ. Second, unlike the aforementioned studies that have forecast prices by using only the conventional macro-factors (e.g., distance, weight of shipment), this manuscript focuses on developing a framework to predict short-term route-level TSR rates using intraday dynamics of transaction data with consideration of both temporal and between-route correlations. As before the arising of online freight exchange platforms, shippers and truckers mainly determine the TSR based on the empirical rates information, e.g., the past rates on a specific route and the latest TSR of the neighboring routes. In this case, we not only consider the temporal correlation as typical time series forecasting models do but also incorporate the between-route correlation in our new model.

Developing such a model to predict the short-term route-level price is valuable for carriers and shippers to capture the current temporal dynamics of TSR and bargain effectively. For example, if today's spot market rate is higher than yesterday's, shippers and carriers may reach a higher negotiation price, and more carriers tend to be more motivated to share capacities for more revenue. This model is established on the basis of the multiple regression model because it allows us to find the functional relationships among the dependent variables and explanatory variables; however, they are appreciated because of practical advantages, such as being relatively easy to implement, the requirement of less computational power than other statistical methods (e.g., neural network, genetic algorithm), satisfactory prediction ability, and increased availability of data through smart metering [18]. Multiple regression presents satisfactory predictability in short-term forecasting. For example, Saber and Alam [19] use big data in a power system to estimate a multiple regression model for short-term load forecasting and explore how different components of weather (e.g., humidity, temperature) influence load demand.

Outside the scope of this article, there are several applications for multiple regression models: Silva *et al.* [20] implement a regression model to estimate the area at risk and environmental variables that best relate to the disease to municipalities in the state of Rio Grande do Sul and demonstrate that the model had certain general employing; Seo *et al.* [21] present a multiregression-based framework to efficiently and accurately determine the load rating of complex steel bridges, and the findings display accurate and rapid prediction of a load rating, given unknown truck data.

The remainder of this paper is organized as follows. Section III describes the proposed methodology and explains how the multiple regression model is optimized. Section IV contains the application of the framework and presents a preliminary analysis of the data. Section V displays related results and discussions. Section VI summarizes the paper and suggests topics for further research.

III. METHODOLOGY

This section aims to provide insights into the general concept of the proposed model, to better understanding how the devising framework is established. Three conventional models are used in the framework: simple linear regression model, multiple regression model, and machine learning. The multiple regression model is used as the basis of the proposed model, and the simple linear regression model and machine learning are used to optimize the multiple regression model. The proposed framework for creating a regression model to perform short-term TSR forecasting is described herein.

The relevant framework comprises three pieces (Figure 1):

The Online_1 component (Step 1) is used to collect field data. A database is derived from the OFEX platform, which provides trucker data (e.g., types of truck, length of truck), shipper data (e.g., shipper ID, registration place of shipper), order data (e.g., specified type of truck, type of cargo), and TSR data (e.g., price, trucker ID, the closed order date). The integration of the aforementioned data is further used for the preliminary analysis, which helps develop a thorough investigation of TSR that makes the best use of the micro trucking data.

The Offline component is used to illustrate the proposed model in detail and comprises three steps (Step 2 to Step 4):

In Step 2, we compute the lagged coefficient-based weighted matrix. Before the emergence of the OFEX platform, truckers and shippers referred to historical rates when bargaining; thus, we propose that the weight matrix should include the time-lagged correlation between route-specific rates. Inspired by Nimon [22] summary of the nature and meaningfulness of variables in a linear regression model, we establish a time-lagged matrix as weights in proposed model to describe the influences from regional rates, and the element of this matrix is defined as the lagged coefficients, which is calculated as follows:

$$\mathrm{y}_i^t = \alpha_0 + \alpha_{ij}\mathrm{y}_j^{t-1}, \quad \forall i, j \in \{1, \ldots, n, i \neq j\} \tag{1}$$

where the dependent variable y_i^t represents the vector of any one route-specific rate on day t, y_j^{t-1} represents other rates on day *t-1*, and the constant α_0 and coefficients α_{ij} are to be estimated from the data. To distinguish it from other matrices in the following analysis, we define this time-lagged matrix as W_r, and W_r is the asymmetric matrix.

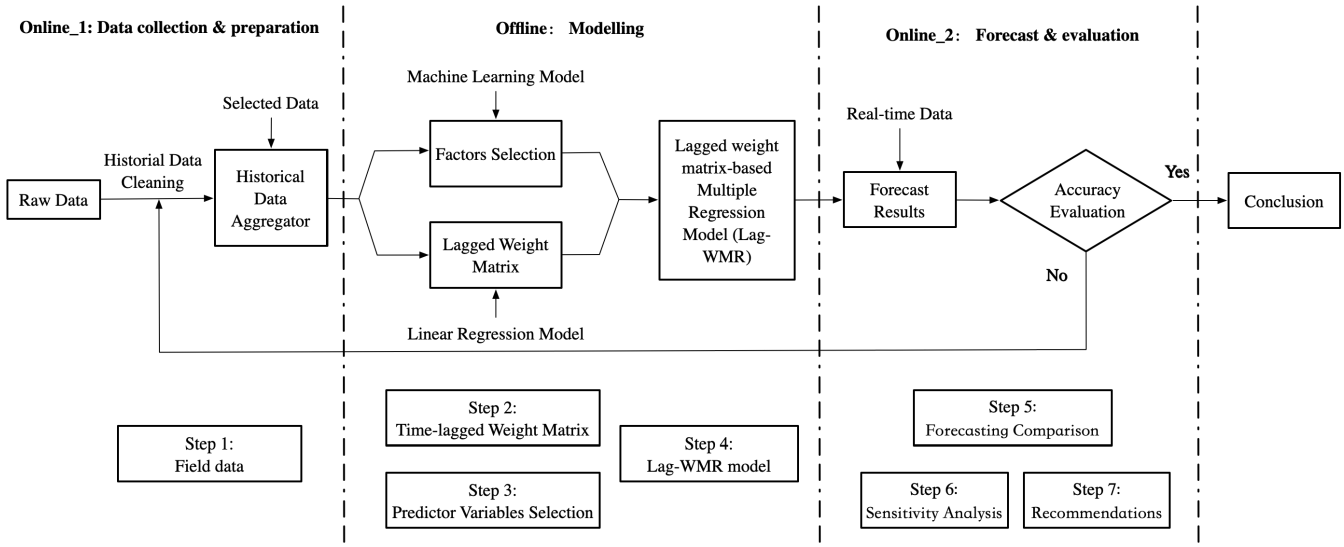


FIGURE 1. Structure of the proposed framework.

In Step 3, we determine which factors actually influence route-specific rates. According to the data information in Step 1, a basic variable set is initially established using information collected in Step 1. To identify the different variables' combinations for each route, an ensemble method LGB is implemented. An LGB is a novel gradient boosting decision tree algorithm that Microsoft proposed in 2017 [23]. The importance of each variable is calculated explicitly for each attribute in the dataset, namely, calculated for a single decision tree by the amount that each attribute split point improves the performance measure, and weighted by the number of observations the node is responsible for. The variables importances are then averaged across all the decision trees within the model. Multiple importances ranking plots are created from the algorithm to show graphically how the variables' combinations differ by routes.

In Step 4, we create the Lag-WMR model. The Lag-WMR model is constructed based on multiple regression (or multiple linear regression). The conventional multiple regression model is formulated as follows:

$$y = \alpha + X\beta + \varepsilon \tag{2}$$

Two modifications are proposed to optimize Equation (2). First, influences from adjacent route-specific TSR are considered (defined as the WY term). Second, the differences in the predictor variables at the lane level is considered. Hence, the variate multiple regression is formulated as follows:

$$y = \alpha + \rho WY + X\beta + \varepsilon \tag{3}$$

where y denotes the dependent variables; X denotes the independent variables; α is the constant;ρ and β are the regression parameters; and ε is the error term. Specifically, W is the one-day lag n × n weighting matrix W_r calculated in Step 2, and the X term is the result of variable selection in Step 3. We digress here to note that Equation (3) is very similar to the spatial econometric model (e.g., spatial autoregression model) [24], but the use of a spatial econometric model here is inappropriate because defining "spatial dependency" is difficult. Spatial dependency is "the propensity for nearby locations to influence each other and to possess similar attributes" [25]; in other words, things that are closer together tend to be more related to each other than are things that are far apart. However, for route-specific rates, although we can say that they may have correlations, we cannot assert the existence of spatial dependency without grounds to do so. In one region, any route may intersect at least two other routes; thus, it is difficult to define the concept of "contiguity" or "distance" among route-specific rates.

This optimized model is used for short-term TSR forecasting in the Online_2 part (Step 5 to Step 7):

In Step 5, we compare the Lag-WMR model with conventional models. The time-lagged coefficient-based regression model is compared with the traditional time series model and machine learning approach. The comparison can be statistically investigated, namely, how well each model is capable of predicting short-term TSR. The prediction model with the smallest values for the evaluation criteria is considered the best fit model.

In Step 6, we conduct sensitivity analysis using the Lag-WMR model to determine how temporal and between-route correlations, different weighted matrices and variable selection affect the rates' predictions. Statistical information helps us draw a conclusion regarding whether the consideration of different variable combinations and time-lagged matrices affect the predicted rates.

In Step 7, we propose recommendations based on the aforementioned analyses that reflect some significant predictor variables and types of weight matrices that affect short-term TSR prediction.

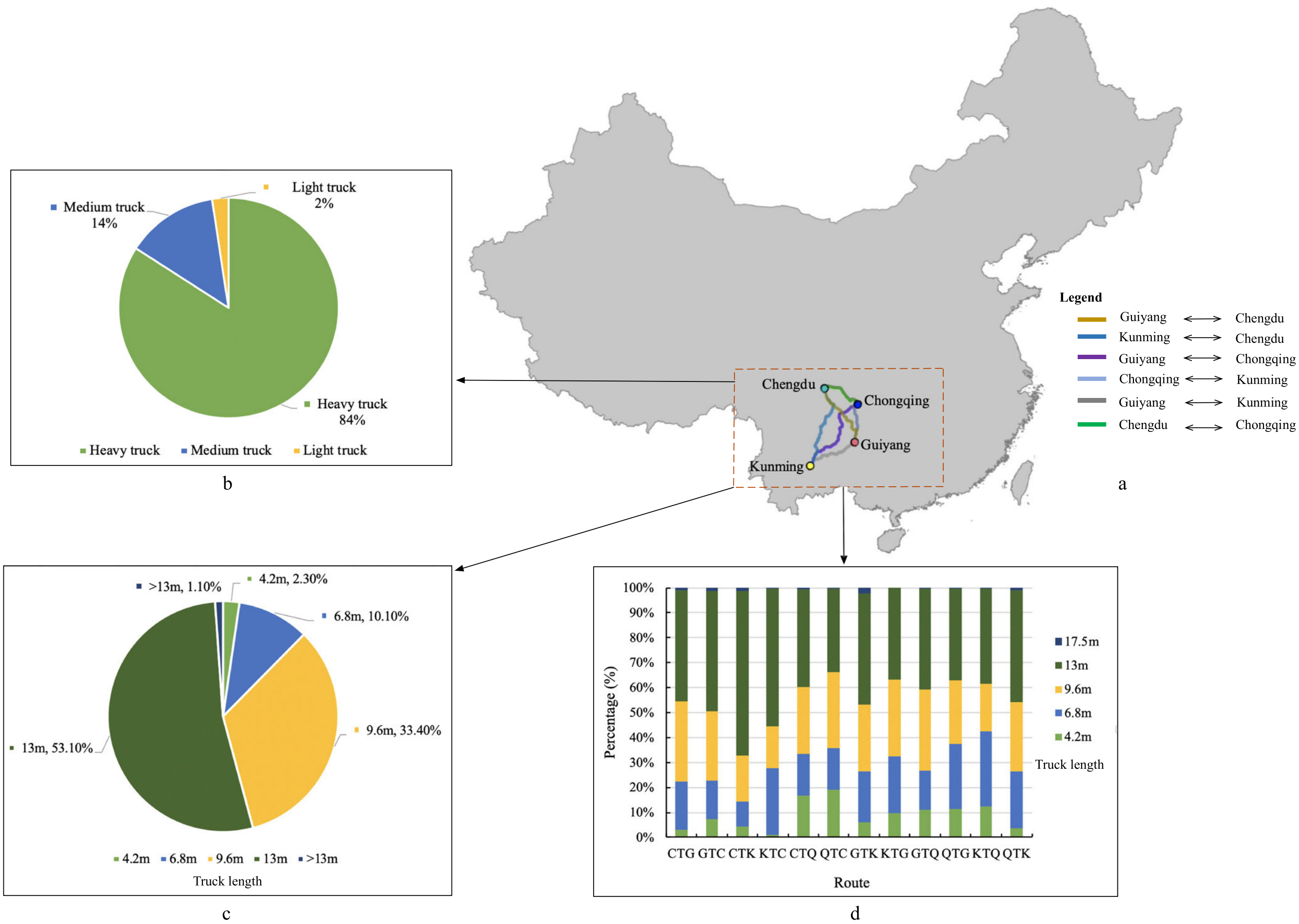


FIGURE 2. (a) Map of the study area; (b) market share of truck classification; overall (c) and route-specific (d) distribution of truck length of heavy trucks.[1]

IV. APPLICATION OF PROPOSED FRAMEWORK

A sample application in the trucking industry of this framework is detailed in this section by following the steps of the framework.

A. FIELD DATA

1) PRELIMINARY DESCRIPTION

This paper uses online trading data obtained from an OFEX platform based in Southwest China. The OFEX platform uses an online mechanism to allow carriers and shippers to bid competitively on a load. The detailed transaction data were obtained from the Truck Alliance [26]. These data reflect the truckload, long-distance, general freight trucking activities, and concern the intraday detailed transaction prices shippers paid for the transportation of freight transported by different trucks (e.g., a heavy, medium, or light truck), in which a heavy truck is the primary type of equipment used to haul most freight. The truck classification is determined based on the vehicle's gross vehicle weight rating, grouped broadly as light trucks, medium trucks, and heavy trucks. A heavy truck refers to tractors, trailers, and straight trucks with a gross vehicle weight of more than 3500 kg. By 2017, the Truck Alliance had already signed up over one-third of China's 6 million truck drivers, and currently, it hosts 70,000 transactions—equivalent to $110 million in shipping costs—each day through its app and website (IFC, 2017). In this study, daily data with a continuous time span of six months, starting from March 1, 2018 are assembled. These data are the full data of the OFEX platform and are not seasonally adjusted. Given the coverage of freight on Truck Alliance.com, we can reasonably view these data as having an acceptable degree of validity in capturing the southwestern spot market full truckload prices in the respective sectors.

The data covers the southwestern part of China (Figure 2) and the colored routes represent the shortest routes between pairwise trips from Google Maps. Southwest China

[1]For brevity, trips from city i to city j are defined as 'iTj', where $ij \in$ [Chengdu (C) , Guiyuang (G) , Chongqing (Q) , Kunming (K)], and $i \neq j$. For example, "CTG" means a trip from Chengdu to Guiyang; Gan et al. [28] map three typical operational models of truckers in the southwestern part of China indicating that the main corridors between capitals are unique, because most truckers in China prefer the highway. Therefore, although there are many realized trips between a given origin–destination (OD), the routes between capitals are considered identical, that is, highway-orientated transportation routes.

TABLE 1. Illustration of sample data.

Description	Example
Order ID	1_5201265
Shipper ID	1_3023651
Shipper registration province	Chengdu
Shipper registration city	Sichuan
Creation time	2018/3/11
Deal time	2018/3/12
Deal price(CNY)	5,320
Order departure	Chongqing
Order destination	Guiyang
Type of truck	Heavy
Length of truck (m)	13
Type of cargo	Vegetable

contains three provincial capitals—Chengdu, Kunming, and Guiyang—and a centrally administered city (Chongqing), which is one of the most promising areas with innovation and technology. Southwest China's combined nominal Gross Domestic Product (GDP) in 2016 was US$ 1.15 trillion, making the four provincial economies equivalent to the then 14th largest economy in the world and with a total of over 190 million population in an area as big as Western Europe and its advanced engineering sectors are automotive, aerospace, advanced manufacturing and marine [27]. In addition, Southwest China is piloting the opening up of the world's largest electricity market (worth approx £300 billion a year); The pilot markets in Chongqing, Guizhou and Yunnan are worth £20 billion a year [29]. In the aspect of technology, 70% of all iPads and 50% of CPUs in the world are made in Southwest [30].

Each TSR record includes information on the truckers, shippers, types of transported goods, and other relevant information. In all, we obtain 179,623 records, with a default record if the status of an order is "Closed/Deal." An illustration of the sample data is presented in Table 1. Such data were rarely available in early studies of China's trucking industry. A data cleaning procedure is implemented to exclude records with null values and outliers. We first remove null values and then observed that the distribution of TSR is close to a Gaussian distribution. Next, we implement a two- boundary outlier elimination to reject abnormal rates with different standard deviation boundaries. After several experiments on the combination of limits and standard deviation, 1.65 standard deviations and 1.96 standard deviations of TSR are selected as the first and second limits for elimination boundaries, respectively, and the final data set retains more than 90% of the records.

Figure 2 displays the truck distribution of the trucking market in Southwest China. Figure 2b shows the market shares of the types of truck, and Figure 2c and Figure 2d show the overall and route-specific distribution of length of heavy trucks, respectively. Figure 2b indicates that heavy trucks

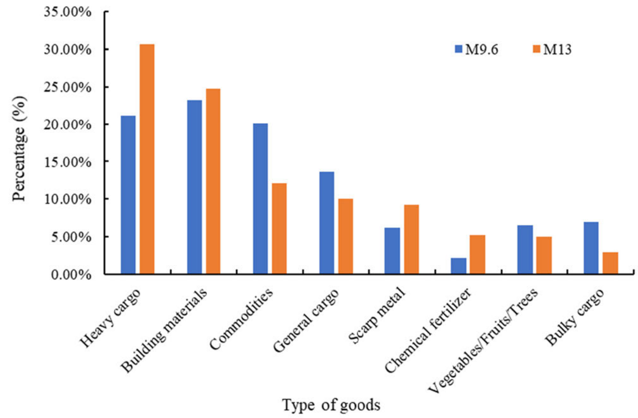


FIGURE 3. Type of transported goods of M9.6 and M13.

are the key type of truck in the study area, with a market ratio reaching up to 84%. At the heavy truck branch level, generally, trucks with a length of 9.6 m and 13 m account for 86.5% of the total (Figure 2c). Figure 2d also demonstrates the big market ratios of trucks of 9.6 m and 13 m in length from a route-specific view. Hence, we target heavy trucks with a length of either 13 m or 9.6 m.

Next, we discuss the data selection procedures in this study.

2) SELECTION OF STUDIED OBJECT

We first investigate quantitatively whether there are any relationships between the rates of the two truck length subsets: 9.6 m (denoted as M9.6) and 13 m (denoted as M13). Given a specific origin-destination (OD), the corresponding Pearson correlation coefficient (PCC) [31] is calculated as follows:

$$\rho_{X,Y} = \frac{Cov(X,Y)}{Var(X)Var(Y)} = \frac{E[(X-E(X))(X-E(Y))]}{Var(X)Var(Y)} \tag{4}$$

where, $Var\,(X) = E(X^2) - [E(X)]^2$, $Var\,(Y) = E(Y^2) - [E(Y)]^2$, and X and Y are the intraday average rates of M13 and M9.6, respectively. The intraday average rate $\bar{p}_t$ is calculated as follows:

$$\bar{p}_t = \frac{\sum_{t=1}^{m_t}(P_t)}{m_t} \tag{5}$$

where m_t is the number of TSR at day t (t = 1,2,…,T), and P_t is the price at day t. Results of the PCC values (0.2 - 0.7) are not strong correlations, indicating that M13 and M9.6 are not necessarily correlated; however, they share the same combination of freight type (Figure 3). For both M13 and M9.6, the top four types of cargo are heavy cargo, building, materials, commodities, and general cargo; thus, given a specific route, a truck with a length of 13 m can sometimes replace a truck with a length of 9.6 m, and vice versa. As M13 and M9.6 are different objects, we pick the object of the study by applying several statistical criteria.

From a general perspective, we consider (i) sample size, (ii) counts of missing values in cargo information, and (iii) coefficient of variation (CV) as our data selection criteria.

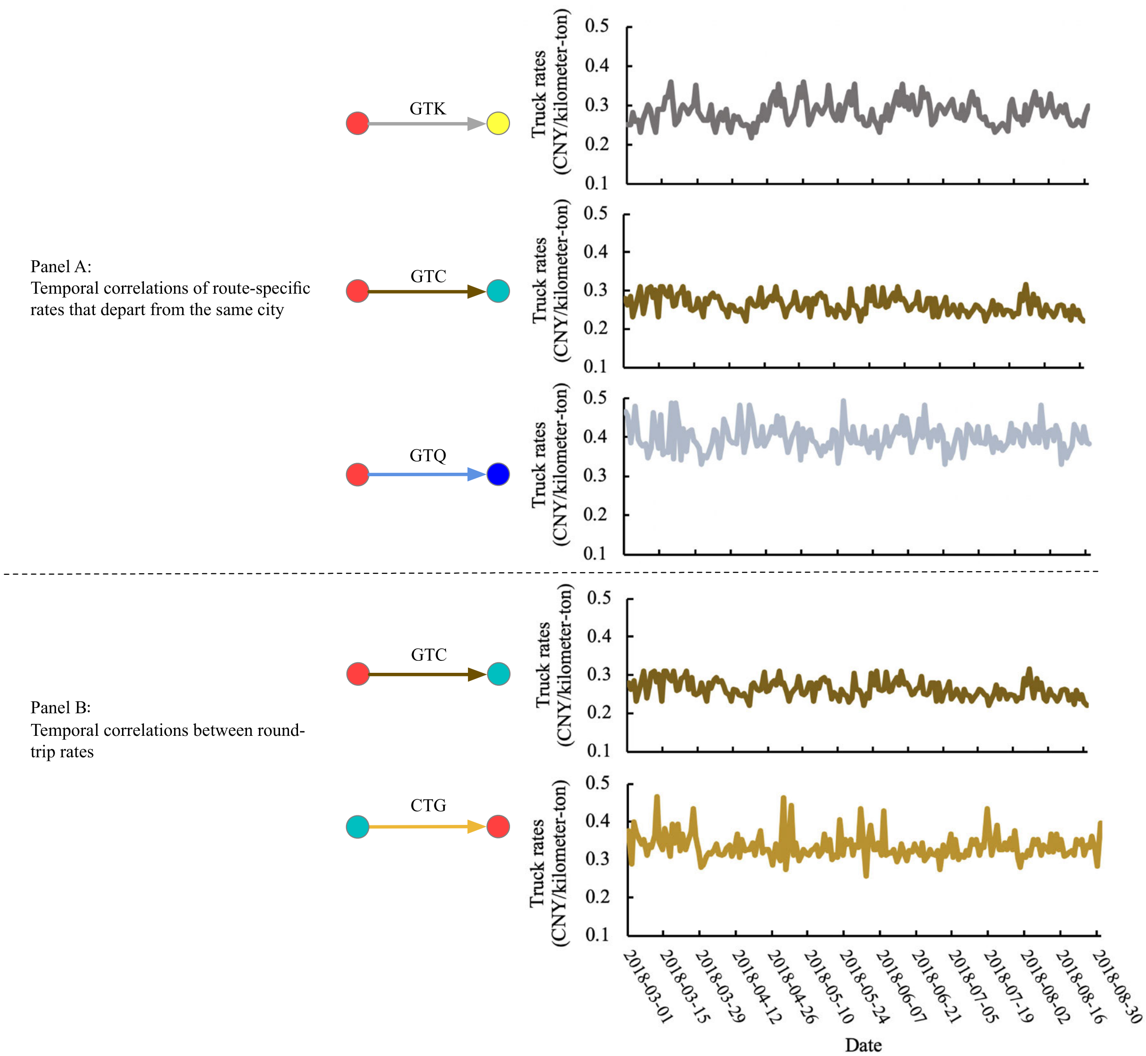


FIGURE 4. **Geographic information of the studied region.**

The reasons why we use the three criteria are as follows: (i) Sample size: As larger sample sizes have the obvious advantage of providing more data for researchers to work with, we prefer the sub-market with a larger sample size. (ii) Missing values in cargo information: As we observed that the fluctuation of rates/trip is influenced by combinations of different types of transported goods, consideration of the cargo information when discussing the rates changes is essential. (iii) CV: CV is calculated for two main factors that most influence rates, namely, posted orders and deal orders, and computed as the standard deviation over the average. The results are shown in Appendix A and Appendix B.

In terms of "null counts" and "sample size" in Appendix A, the sample size of M13 is almost twice that of M19, and M13 has a smaller number of null values. As the greater the number of null values, the more unusable the data; thus, it is preferable to select samples with fewer null values. Obviously, M13 has a much smaller null ratio (3.60%) than M9.6 (18.20%). Appendix B shows that M13 has a lower average CV in both posted orders and closed orders (0.192 and 0.195, respectively). The smaller the CV, the more stable the market is likely to be. Based on a combination of the aforementioned results, this paper focuses on the more stable M13 in the following analysis.

B. LAGGED WEIGHTED MATRIX

Instead of presenting the time-lagged weight matrix directly, we illustrate why we consider the devising matrix first. We first visually display the temporal characteristics among route-specific TSR as time series figures. In brief, we take rates related to the city of Guiyang as examples to illustrate the temporal correlations (Figure 4).

In general, Panel A illustrates the temporal correlations between the rates of routes that share the same city of departure. The rates of GTC and GTK show similar dynamic changes, beginning with a decline from March 18 to April 19, then fluctuating until June 10, after which there is a downtrend from June 20 to July 31, another uptrend at the

TABLE 2. PCC results between rates of routes in M13.

OD		CTG	CTK	CTQ	GTC	GTK	GTQ	KTC	KTG	KTQ	QTC	QTG	QTK
OD	CTG	1	0.211	0.114	-0.276	-0.178	-0.301	0.111	-0.132	0.189	0.673	0.341	0.154
	CTK	0.211	1	0.311	0.232	-0.145	-0.319	0.224	0.328	-0.167	0.107	0.129	-0.151
	CTQ	0.114	0.311	1	-0.208	-0.376	-0.212	0.171	0.134	0.351	0.121	-0.091	0.263
	GTC	-0.276	0.232	-0.208	1	0.215	0.298	0.161	0.412	-0.241	0.265	0.284	-0.132
	GTK	-0.178	-0.145	-0.376	0.215	1	0.234	-0.154	0.214	-0.146	0.198	0.212	0.104
	GTQ	-0.301	-0.319	-0.212	0.298	0.234	1	0.127	-0.211	0.343	-0.213	0.241	-0.144
	KTC	0.111	0.224	0.171	0.161	-0.154	0.127	1	-0.433	0.321	0.269	0.209	-0.187
	KTG	-0.132	0.328	0.134	0.412	0.214	-0.211	-0.433	1	-0.263	0.32	-0.118	0.243
	KTQ	0.189	-0.167	0.351	-0.241	-0.146	0.343	0.321	-0.263	1	0.249	0.092	-0.189
	QTC	0.673	0.107	0.121	0.265	0.198	-0.213	0.269	0.32	0.249	1	-0.082	0.424
	QTG	0.341	0.129	-0.091	0.284	0.212	0.241	0.209	-0.118	0.092	-0.082	1	0.186
	QTK	0.154	-0.151	0.263	-0.132	0.104	-0.144	-0.187	0.243	-0.189	0.424	0.186	1

beginning of the August, and a subsequent downtrend. Panel B shows the temporal correlations between the rates of round trips. Visually, the trend of rates of GTC and CTG are likely to be opposite to each other. Comparing the changing pattern in the period studied, we demonstrated that from March 31 to May 1 and from June 30 to July 31, the rates of GTC display a concave shape, followed by a downtrend at the end of August, whereas those of CTG display a convex shape, followed by an uptrend along the tail.

As for the interval between the lowest points of each series, for the rates of GTK, GTQ, and GTC, there are approximately 50 days between the lowest points, whereas for CTG, the next lowest rate is every 30 days. Moreover, the GTC and GTQ rates are more stable than the other two rates, because the oscillatory intervals for GTC and GTQ are between 150 CNY/ton to 200 CNY/ton and 125 CNY/ton to 185 CNY/ton, respectively.

Inspecting Figure 4, The rates changes during the festival appear to be larger than usual, which arouses our interests to consider whether the festival should be considered as a key factor for rates fluctuation. Generally, national festivals (usually three to seven days in China) can increase traffic demand. In order to stimulate the holiday economy, the Chinese government proposed the toll-free policy for small passenger cars on the national festivals since October 2012, this toll-free policy encourages more people travel on the highways, and thus causes severe highway congestions [32], which also increase travel time and fuel consumption of the trucks and thus result in higher roadway transportation prices. Furthermore, Chinese labor laws mandate that, for work during the statutory holidays, the employer should pay the employee no less than three times the normal wage rate [33], which leads to rises in the transportation prices during the festivals as well. Thus, it is necessary to observe the movements of TSR before and after festivals. The period under study includes three national festivals: the Sweeping Tomb Festival (April 5–April 7), Labor Day (May 1–May 4), and the Dragon Boat Festival (June 7–June 9). The four rates show significant fluctuation during Labor Day: a brief rally is observed on the first day of the Labor Day Festival (May 1), the trucking market erases the gains on May 2, and then the rates return to normal. The above dynamic fluctuations of TSR time series indicate that the TSR may display a boom-bust pattern before and after the festivals but return to normal soon. Thus, we consider the national festivals as one of the explanatory variables in our initial feature set.

Furthermore, we use PCC (Equation 4) to quantify the correlation of rates, in which X and Y are the rates of different routes in M13. The results are reported in Table 2 and presented as a heat map for ease of reading.

The cell color in Table 2 provides a perceptual intuition of the correlation between rates of different routes. Remarkably, the rates of CTG show an obviously positive correlation with those of QTC (0.673), QTG (0.341), and GTQ (0.301) but also a negative correlation with the rates of GTC (−0.276). Additionally, the greatest degree of correlation with the rates of GTC is observed in the rates of KTG (0.412), but these rates also show correlations with the rates of GTQ (0.298), QTG (0.284), CTG (−0.276), QTC, (0.265), and so forth. We observe that the rates of a given route show different degrees of correlation with other route-specific rates. However, the range of PCC, 0.3(-0.3) – 0.7, does not show strong correlations; thus, using the PCC values as the weights may be inappropriate. Therefore, we consider different transfer correlations as a weighted matrix calculated in Equation (1). The results are displayed in Table 3. Each element in Table 3 means the change in the value of one route-specific rate corresponding to the unit change in the values of other route-specific rates, which will be used in the forecasting model.

C. PREDICTOR VARIABLES' SELECTION

We believe propose that the more effective information we use, the more accurate our prediction will be. In the literature, most of the explanatory variables have been chosen empirically and used one predictor set by default for every object. However, in reality, the variable combinations may vary by route. To solve this issue, we initialize a basic predictor

TABLE 3. Time-lagged weighted matrix W_r.

		OD											
		CTG	CTK	CTQ	GTC	GTK	GTQ	KTC	KTG	KTQ	QTC	QTG	QTK
OD	CTG	-	0.108	0.201	0.615	-0.065	-0.101	-0.045	0.030	-0.031	0.433	0.087	0.319
	CTK	-0.111	-	-0.367	0.702	0.301	-0.144	-0.076	0.276	0.132	0.128	0.114	0.063
	CTQ	-0.282	-0.352	-	-0.140	-0.410	0.131	-0.065	0.076	0.201	0.028	-0.060	-0.043
	GTC	-0.983	0.346	0.132	-	0.093	-0.232	0.012	0.389	0.043	0.034	-0.102	0.194
	GTK	0.212	0.361	-0.076	0.142	-	-0.087	-0.076	0.108	0.043	-0.083	0.218	0.199
	GTQ	0.382	-0.018	0.104	-0.149	-0.063	-	0.259	-0.093	-0.138	0.427	0.032	0.174
	KTC	-0.319	-0.673	-0.753	0.427	0.467	0.798	-	-0.429	-0.124	-0.163	0.294	0.043
	KTG	-0.043	0.341	0.136	0.392	-0.132	0.432	-0.084	-	0.103	0.202	-0.043	0.293
	KTQ	0.498	0.034	0.241	-0.239	-0.673	0.613	0.127	-0.076	-	0.482	0.193	-0.052
	QTC	-0.026	0.048	0.166	0.345	-0.028	-0.065	-0.034	0.081	-0.012	-	0.293	0.026
	QTG	0.482	0.012	0.361	0.561	0.853	-0.031	0.032	-0.008	0.154	0.285	-	0.265
	QTK	0.003	-0.329	-0.351	0.901	0.205	0.074	0.063	0.210	-0.023	0.342	-0.038	-

TABLE 4. Lists of features per trip.

Feature number	Feature description
F1	Whether the day is a weekend: 0 for weekends, and 1 for weekdays
F2	Whether the day is a state holiday: 0 for state holidays, and 1 for weekdays
F3	Number of unoccupied truckers at departure city
F4	Number of unoccupied truckers at departure city with 'Heavy truck' at a length of 13 m
F5	Total order inflow into departure city
F6	Total order outflow into destination city
F7	Sum of posted orders
F8	Sum of closed orders
F9	Sum of posted orders with 'Heavy truck' at a length of 13 m
F10	Sum of closed orders with 'Heavy truck' at a length of 13 m
F11	1- Intraday median value of TSR 2- Weekly median value of TSR
F12	1- Intraday 75th percentile of TSR 2- Weekly 75th percentile of TSR
F13	1- Intraday mean value of TSR 2- Weekly mean value of TSR
F14	1- Intraday minimal value of TSR 2- Weekly minimal value of TSR
F15	1- Intraday maximal value of TSR 2- Weekly maximal value of TSR
F16	1- Intraday range of TSR 2- Weekly range of TSR
F17	1- Intraday standard deviation of TSR 2- Weekly standard deviation of TSR

set according to the data structure and empirical experience and use variable assessment tools to select the appropriate predictor variables for every single route, to collect the predictor variables that affect the route-specific TSR. By contrast, it is challenging to consider a large number of potential input combinations, the potential correlations between inputs, very weak correlations between future dependent variables and current and past inputs, and time-varying structures. In particular, a large number of possible input combinations can lead to the risk of overfitting. To solve this problem, LGB is used to select explanatory variables. Table 4 summarizes the candidate list of initial predictor variables $\boldsymbol{X_i}$, in which variables F11−F17 represent the statistical characteristics of temporal correlation. The importances of the ranking plots are displayed in Figure 5.

Figure 5 shows the relative importance of the variables based on the feature importance of the LGB. Obviously, each OD has different sets of important variables, which implies the inappropriateness of the default identical exogenous variables set for each route. We observe that the historical statistics of rates F11–F17 all rank at the top in most cases. For example, GTK, GTQ, and KTQ are much more sensitive to the Intraday 75th percentile of TSR (F12-1), indicating carriers and shippers on these routes pay more attention to the intraday higher TSR to offer a higher price when making decisions. By contrast, carriers and shippers located in KTC, CTG, and CTQ care much more about the range of TSR (F17). In addition, we observe that the number of posted/closed orders and unoccupied truckers remain critical variables for some routes, such as QTG and QTC. The above aforementioned results demonstrate the existence of the diversity of variable combinations of different routes, and we demonstrate that choosing appropriate input variables is likely to help improve the accuracy of predictions. According to the results in Figure 5, features with relative importance higher than 0.1 are selected as the exogenous variables of the Lag-WMR model (Table 5).

D. LAG-WMR MODEL

According to Equation (3), to create a Lag-WMR model, we must calculate a weighting matrix and variable combinations from the weights and variables identification. The analytical results of the time-lagged matrix and predictor selections are combined into a Lag-WMR model. As aforementioned, each route-specific set of variable combinations is presented in Table 5, and the weights assigned to rates of adja-

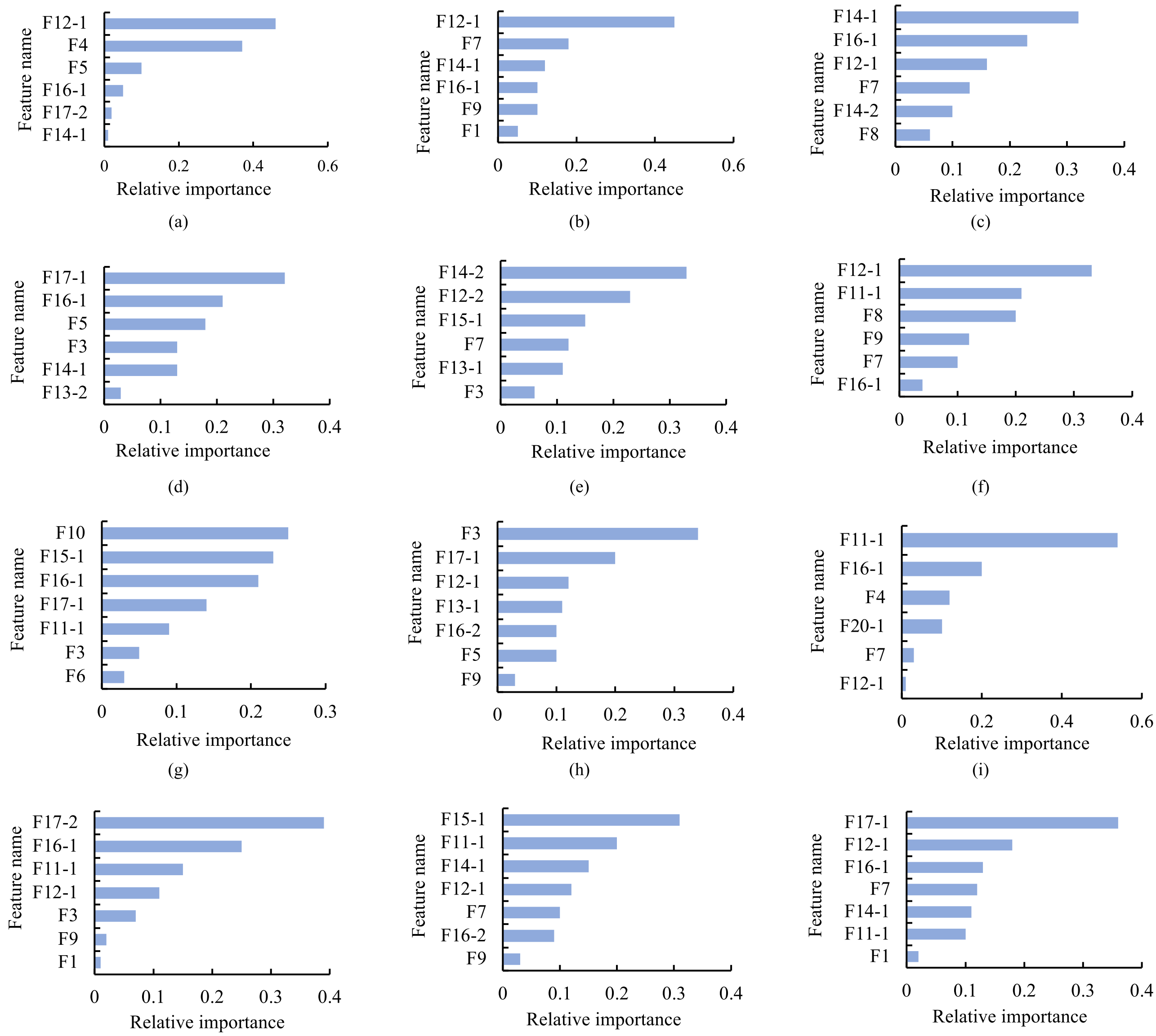


FIGURE 5. Relative importance of route-specific variables by LGB: (a) GTK. (b) GTQ. (c) GTC. (d) KTC. (e) KTG. (f) KTQ. (g) QTC. (h) QTG. (i) QTK. (j) CTG. (k) CTK. (l) CTQ.

cent routes are presented in Table 3. To solve the Lag-WMR model, an ordinary least square (ols) method is implemented. To calculate the optimal estimate for α, β, and ρ, a choice-criterion is necessary; in the case of OLS, the criterion is the sum of the squared residuals. We calculate α, β, and ρ for the case in which the sum of all squared residuals (ε) is minimal, that is,

$$\min_{\hat{\alpha},\hat{\beta},\hat{\rho}} \sum (\varepsilon)^2 \tag{6}$$

The residual ε is named as the difference between dependent variable y and the estimated systematic influence on X and WY on y:

$$\varepsilon = y - \alpha - \beta x - \rho WY \tag{7}$$

where y denotes the dependent variables; X denotes the independent variables, WY represents influences from adjacent route-specific TSR; α is the constant;ρ and β are the regression parameters; and ε is the error term. Combining Equation (6) and Equation (7), we obtain the function as follows:

$$\min_{\hat{\alpha},\hat{\beta},\hat{\rho}} \sum (y - \alpha - \beta x - \rho WY)^2 = S(\hat{\alpha}, \hat{\beta}, \hat{\rho}) \tag{8}$$

The above solving process is achieved through Python 3.7, and the one-step-ahead TSR predicted from the Lag-WMR model are compared with the conventional models in a statistical manner, which we show in Section V.

V. RESULTS AND DISCUSSION

A. FORECASTING COMPARISON

The proposed Lag-WMR model is compared with two benchmark methods, a typical time series prediction model ARIMA and a machine learning approach LGB model. ARIMA is a traditional time series forecasting model that has been widely

TABLE 5. Selected variables for each route.

OD	Selected variables
CTG	F17-2, F16-1, F11-1, F12-1
CTK	F15-1, F11-1, F14-1, F12-1, F7
CTQ	F17-1, F12-1, F16-1, F7, F14-1, F11-1
GTC	F14-1, F16-1, F12-1, F7, F14-2
GTK	F12-1, F4, F5
GTQ	F12-1, F7, F14-1, F16-1, F9
KTC	F17-1, F16-1, F5, F3, F14-1
KTG	F14-2, F12-2, F15-1, F7, F13-1
KTQ	F12-1, F11-1, F8, F9, F7
QTC	F10, F15-1, F16-1, F17-1
QTG	F3, F17-1, F12-1, F13-1, F16-2, F5
QTK	F11-1, F16-1, F4, F20-1

applied in many fields of study such as finance [34], shipping [35], logistics [17], and electric power [36] and recently ARIMA model has been proved to be superior to the artificial neural networks in the short-term forecasting [37], [38]. And LGB is a promising ensemble machine learning method with a gradient boosting framework and a tree-based learning algorithm that is used for ranking, classification and prediction [39], [40]. The training- (in-sample) and test- (out-of-sample) samples are considered for all employed models and we consider two performance measures: mean absolute error (MAE) and mean absolute percentage error (MAPE), to assess the forecast accuracy. The performances measures are calculated based on the following equations:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \tag{9}$$

$$\mathrm{MAPE} = 100 \times \frac{1}{N}\sum_{i=1}^{N}\frac{|\hat{y}_i - y_i|}{y_i} \tag{10}$$

First, the appropriate ARIMA model is to be selected to forecast the group of the TSR time series. ARIMA model is determined by the lowest Akaike information criterion (AIC) [41] and the stationarity is examined by Phillips-Perron (PP) test [42] using the auto.arima function in the Forecast package provided by programming language R. According to the AIC, the selected model is ARIMA (2, 1, 2).

Once the ARIMA model has been decided, the in-sample and out-of-sample predictions can be conducted. We divide the dataset into training- (in-sample) and test- (out-of-sample) samples. The in-samples are from March 1, 2018 to August 11, 2018, and the out-of-samples are from August 12, 2018 to August 31, 2018. In Table 6, we first present the in-sample forecast performances of Lag-WMR, LGB and ARIMA models by means of MAE and MAPE. Then in Table 7, the out-of-sample forecast performances of the employed models are displayed. According to Table 6, Lag-WMR model performs better than LGB and ARIMA models in most cases. According to Table 7, for all cases, Lag-WMR is superior to LGB and ARIMA in the

TABLE 6. Comparison results of the In-sample forecasts.

	In-sample period					
OD	Lag-WMR		LGB		ARIMA	
	MAE	MAPE	MAE	MAPE	MAE	MAPE
CTG	**0.015**	**0.064**	0.016	0.068	0.029	0.125
CTK	**0.016**	**0.040**	0.031	0.078	0.066	0.165
CTQ	**0.023**	**0.058**	0.046	0.115	0.061	0.155
GTC	**0.011**	**0.033**	0.025	0.074	0.046	0.138
GTK	**0.014**	**0.065**	**0.014**	**0.065**	0.038	0.177
GTQ	0.022	0.031	**0.021**	**0.029**	0.106	0.149
KTC	**0.013**	**0.035**	0.026	0.071	0.069	0.185
KTG	**0.017**	**0.051**	0.028	0.084	0.054	0.162
KTQ	**0.018**	**0.054**	0.062	0.186	0.055	0.165
QTC	**0.010**	**0.021**	0.030	0.064	0.094	0.198
QTG	**0.029**	**0.068**	0.036	0.085	0.078	0.184
QTK	**0.013**	**0.062**	0.025	0.120	0.031	0.146

Note: Values in bold indicate the best performance.

out-of-sample TSR forecasts. These low MAPEs reflect that the differences between the actual and predicted TSR are very small. However, the performances of different models presented above are judged solely in terms of MAPE and MAE. These comparisons give the ordinal rankings of the models but provide no evidence of whether the forecasts from one particular model are significantly better than those from another model in a statistical view. To address this issue, the Diebold-Mariano (DM) test [43] is implemented to compare the prediction results between the employed models. The DM test is widely used in determining whether the differences of time series predicting accuracy by different models are substantially crucial from a statistical perspective [44]. The results, presented in Table 8, show that it is possible to reject the null hypothesis that Lag-WMR and LGB (ARIMA) have the equal predictive capacity at the traditional levels of significance. In other words, the Lag-WMR forecast error is significantly different from that of LGB and ARIMA. Indeed, the results of DM test provide additional evidence to demonstrate the Lag-WMR model generates better accuracy in TSR

TABLE 7. Comparison results of the out-of-sample forecasts.

	Out-of-sample period					
OD	Lag-WMR		LGB		ARIMA	
	MAE	MAPE	MAE	MAPE	MAE	MAPE
CTG	**0.012**	**0.052**	0.020	0.088	0.030	0.131
CTK	**0.008**	**0.042**	0.020	0.103	0.031	0.158
CTQ	**0.023**	**0.053**	0.062	0.146	0.055	0.129
GTC	**0.008**	**0.036**	0.020	0.087	0.032	0.137
GTK	**0.015**	**0.058**	0.019	0.074	0.043	0.165
GTQ	**0.016**	**0.033**	0.027	0.056	0.073	0.151
KTC	**0.008**	**0.032**	0.021	0.088	0.043	0.178
KTG	**0.019**	**0.056**	0.033	0.098	0.052	0.154
KTQ	**0.010**	**0.065**	0.038	0.261	0.024	0.161
QTC	**0.007**	**0.023**	0.021	0.069	0.045	0.149
QTG	**0.025**	**0.064**	0.069	0.174	0.068	0.171
QTK	**0.013**	**0.059**	0.031	0.142	0.027	0.124

Note: Values in bold indicate the best performance.

TABLE 8. DM test for the out-of-sample forecasting results.

OD	DM test statistics	
	Lag-WMR vs. LGB	Lag-WMR vs. ARIMA
CTG	−1.912 (0.056) *	-2.412 (0.015) **
CTK	−1.826 (0.072) *	-2.268 (0.023) **
CTQ	−1.667 (0.097) *	-1.891(0.060) *
GTC	−1.715 (0.087) *	-1.812 (0.070) *
GTK	−2.281 (0.022) **	-1.756 (0.078) *
GTQ	−2.412 (0.016) **	-1.785(0.077) *
KTC	−1.981 (0.048) **	-1.815(0.070) *
KTG	−1.710 (0.087) *	-1.822 (0.068) *
KTQ	−2.445 (0.014) **	-1.756 (0.078) *
QTC	2.350 (0.021) **	-2.368 (0.018) **
QTG	2.114 (0.036) **	2.218 (0.025) **
QTK	−2.587 (0.010) **	-2.058 (0.039) **

Note: P-values are reported in the parentheses. ** refers to a significance at (5%), * refers to a significance at (10%).

forecast than LGB and ARIMA models and this improvement is statistically significant. This further strengthens the earlier conclusions derived from the MAPEs and MAEs and indicates that a machine learning approach might not be the best choice for the short-term TSR prediction.

Regarding the above analysis, while the average of MAPEs of Lag-WMR model (0.049) is lower than that of ARIMA (0.096) and the DM test statistics between Lag-WMR and ARIMA is statistically significant makes theoretical sense in that it suggests the Lag-WMR model captures the movement of TSR better, there is less explanation for why the Lag-WMR model plays such a strong predictive role. One potential explanation is that ARIMA modeling very simply makes use of data from either the recent or more distant past to model the existing data as well as to make predictions of future behavior, while the Lag-WMR model not only deeply investigates the temporal correlations among the recent and past TSR by calculating a set of rates statistics but also considering the shippers and truckers' on-the-spot behaviors and transforming them as critical terms in the model. Turning now to the spot market, the on-the-spot shippers and truckers mainly determine the rates based on the empirical rates information in one region, e.g., the past rates on the route and the latest spot rates of the neighboring routes. Before making decisions, truckers tend to first compare orders on the online exchange platform at the current city. This is because truckers are willing to haul high-value orders. And to make it a round trip, they also try to find a shipment back from the destination of the current shipment. Thus, truckers activated in one region prefer to pick the order that (i) the origin of the freight is at the same city as the truckers, (ii) sufficient orders back from destination and (iii) whose spot rate is the highest. Consequently, the route with more high-value orders will attract more truckers and thus results in the shortage of truckers on other routes, and then followed by the rises in spot rates on other routes. In this case, the behaviors of the truckers on selecting the on-the-spot orders could influence spot rates on alternative routes. These on-the-spot behaviors may be associated with between-route correlation among TSR in one region, which is neglected in the ARIMA model.

Another concern that often expressed by industry practitioners seeking to utilize the statistical model for applied uses, is about the choices of alternative routes when introducing the between-route correlations into the TSR forecasting. First, the routes between capitals undertake approximate 70% of cargos in the southwestern area and thus other non-capital cities are not considered. Second, the order choice behaviors of truckers show strong regional characters. Gan *et al.* [28] demonstrate that truckers in the southwestern area are preferred to operate in one region with a relatively small radius (about 300–400 km) as truckers tend to stick with the routes that they are familiar with. Similarly, the southwestern shippers hardly cooperate with the truckers outside this region. As a result, the routes between the four southwestern capital cities are included in the final model.

B. SENSITIVITY ANALYSIS

This paper aims to develop an accurate time series forecasting model to capture the movements of TSR and, in doing so, shed some lights on the underlying factors that give rise to the improvement in shaping and predicting TSR. To better address this regard, several sensitivity analyses are conducted in this section. The first analysis is to examine whether incorporating temporal and/or between-route correlations make improvements in TSR forecast. And the other two are to check the effectiveness of the time-lagged weight matrix and variable selection for short-term forecasting: (i) the extent of prediction power after wiping out the variable selection in the conventional multiple regression model, and (ii) the extent of prediction power after replacing the one-day lag weight matrix with conventional PCCs in the proposed model to evaluate in terms of MAE and MAPE. Results are presented in Table 9, Table 10 and Table 11, respectively.

TABLE 9. Comparison results of Out-of-sample forecast with and without temporal and between-route correlations.

	Out-of-sample period							
OD	Lag-WMR with considering temporal and between-route correlations		Lag-WMR without considering temporal correlation (wiping out $X\beta$ term from Equation (3))		Lag-WMR without considering between-route correlation (wiping out WY term from Equation (3))		Decreases in MAPE (%) after incorporating temporal correlation	Decreases in MAPE (%) after incorporating between-route correlation
	MAE_1	MAPE_1	MAE_2	MAPE_2	MAE_3	MAPE_3	$\mathrm{Diff}^t =$ (MAPE_2 − MAPE_1)%	$\mathrm{Diff}^b =$ (MAPE_3 − MAPE_1)%
CTG	0.012	0.052	0.026	0.112	0.034	0.147	6.0%	9.5%
CTK	0.008	0.042	0.021	0.108	0.029	0.151	6.6%	10.9%
CTQ	0.023	0.053	0.054	0.124	0.057	0.132	7.1%	7.9%
GTC	0.008	0.036	0.024	0.107	0.030	0.134	7.1%	9.8%
GTK	0.015	0.058	0.033	0.129	0.041	0.157	7.1%	9.9%
GTQ	0.016	0.033	0.057	0.117	0.072	0.148	8.4%	11.5%
KTC	0.008	0.032	0.026	0.103	0.035	0.138	7.1%	10.6%
KTG	0.019	0.056	0.043	0.126	0.045	0.132	7.0%	7.6%
KTQ	0.01	0.065	0.022	0.140	0.022	0.141	7.5%	7.6%
QTC	0.007	0.023	0.030	0.100	0.037	0.121	7.7%	9.8%
QTG	0.025	0.064	0.043	0.111	0.053	0.135	4.7%	7.1%
QTK	0.013	0.059	0.022	0.098	0.023	0.106	3.9%	4.7%

TABLE 10. Comparison of out-of-sample forecasting results with different weighted matrices.

	OD											
	CTG	CTK	CTQ	GTC	GTK	GTQ	KTC	KTG	KTQ	QTC	QTG	QTK
Lag-WMR with time-lagged matrix												
MAE	**0.012**	**0.008**	**0.023**	**0.008**	**0.015**	**0.016**	**0.008**	**0.019**	**0.010**	**0.007**	**0.025**	**0.013**
MAPE	**0.052**	**0.042**	**0.053**	**0.036**	**0.058**	**0.033**	**0.032**	**0.056**	**0.065**	**0.023**	**0.064**	**0.059**
Lag-WMR with matrix defined by Pearson correlation												
MAE	0.047	0.020	0.080	0.033	0.028	0.088	0.027	0.026	0.036	0.030	0.056	0.051
MAPE	0.203	0.099	0.187	0.143	0.107	0.182	0.111	0.078	0.246	0.102	0.141	0.237

Note: Values in bold indicate the best performance.

Table 9 presents the results of either wiping out the temporal or between-route correlations for the out-of-sample TSR forecasting. The most salient result is that the performance improvement brought by the combination of temporal and between-route correlations is superior to that of adopting only one of them. The finding that the complete Lag-WMR model has the lowest MAPE values is consistent with the earlier augment that temporal and between-route correlations exist in TSR time series. Turning now to the two rightmost columns in Table 9, the Diff^t is larger than Diff^b for all cases. Regarding the interpretation of this finding, larger decreases in MAPE achieved by incorporating the between-route correlation indicates the on-the-spot shippers and truckers rely more on the TSR information of neighboring routes. This finding is consistent with the southwestern truckers' order choice behaviors – preferring to operate in one region and pick the best deal.

From the results of Table 10, the Lag-WMR with a time-lagged weighting matrix is found to achieve better forecasting results than those achieved using the Pearson correlation as a weight matrix. This finding may be attributed to the consideration of the lagged correlations of the past rates, and to some extent, this result corresponds to the industry behavior that truckers or shippers proposing prices for the next day will depend on recent historical rates for their decisions, which cannot be reflected in a Pearson correlation. Turning now to Table 11, to demonstrate the general practicability of the variable selecting process, the conventional multiple linear regression (MLR) is used as the base model instead of the Lag-WMR model. Regarding the values of performance measures, MLR model with the selected variables (by LGB) outperforms that with the full variables, which indicates that variable selection makes a great contribution to the accuracy of short-term rates' prediction. Additional explanatory

TABLE 11. Comparison of out-of-sample forecasting results with different independent variable sets.

	OD											
	CTG	CTK	CTQ	GTC	GTK	GTQ	KTC	KTG	KTQ	QTC	QTG	QTK
MLR with variables selected through LGB												
MAE	**0.020**	**0.019**	**0.036**	**0.019**	**0.019**	**0.023**	**0.024**	**0.031**	**0.020**	**0.009**	**0.048**	**0.027**
MAPE	**0.085**	**0.094**	**0.085**	**0.083**	**0.073**	**0.047**	**0.097**	**0.092**	**0.136**	**0.031**	**0.121**	**0.126**
MLR with full variables												
MAE	0.056	0.033	0.053	0.035	0.044	0.066	0.042	0.047	0.034	0.068	0.066	0.028
MAPE	0.241	0.166	0.125	0.150	0.170	0.137	0.173	0.139	0.230	0.226	0.167	0.129

Note: Values in bold indicate the best performance.

TABLE 12. Comparisons of Sample size and Null values counts.

	M9.6			M13		
OD	Null counts	Sample size	Null ratios	Null counts	Sample size	Null ratios
CTG	230	5871	3.90%	279	8183	3.40%
CTK	342	4377	7.80%	241	7561	3.20%
CTQ	320	1879	17.00%	210	6730	3.10%
GTC	441	2405	18.30%	278	7873	3.50%
GTK	1099	3840	28.60%	111	5603	2.00%
GTQ	1580	4666	33.90%	442	6729	6.60%
KTC	220	2106	10.40%	201	6382	3.10%
KTG	841	2154	39.00%	241	5390	4.50%
KTQ	422	4646	9.10%	398	7109	5.60%
QTG	597	2935	20.30%	201	6610	3.00%
QTK	611	4153	14.70%	332	8393	4.00%
QTC	781	5293	14.80%	136	8591	1.60%
Sum	6703	44324	-	3070	85154	-
Average	-	-	18.20%	-	-	3.60%

variables do not mean higher forecasting accuracy, but the correct variables do.

C. RECOMMENDATION

From the sensitivity results, we infer that different route-level TSRs are sensitive to different predictor variables. It is inappropriate to assume one same set of the predictors' combinations for all rates. Incorporating temporal and between-route correlations in TSR and identifying real predictor variables can positively affect short-term TSR. Based on the results, we recommend that carries and shippers may consider this model to forecast the future price to ensure the price with the least premium.

VI. CONCLUSION AND FUTURE STUDY

In this paper, we have proposed a Lag-WMR method to manage the issue of short-term TSR' forecasting, using data from the OFEX platform. The proposed approach is based on the multiple regression model, with two improvements for variables' selection and the weighted matrix. The first improvement is the implementation of a machine learning approach, LGB, for selecting the critical exogenous variables. The second is the use of a time-lagged weighted matrix established by the estimated coefficients of regression models. We compared the proposed model with two benchmark models: ARIMA and LGB. The comparative results show that the predictive performance of the Lag-WMR largely outperforms the benchmarks in terms of MAE and MAPE. Based on the results of DM test, empirical study shows that Lag-WMR generates better accuracy in out-of-sample TSR forecasts than LGB and ARIMA and the improvement is statistically significant. We have conducted further sensitivity analyses to verify whether the temporal and between-route correlations could improve forecasting accuracy, and check the effectiveness of the regression-based weight matrix and the selection of variables by LGB. Results demonstrate that the performance improvement brought by the combination of temporal and between-route correlations is superior to that of adopting only one of them. Meanwhile, Lag-WMR, with a lagged correlation-based weighting matrix, achieves better prediction accuracy than does prediction with the use of the Pearson correlation matrix, which should be attributed to the inclusion of a temporal lag correlation with the regression.

The proposed model framework can be applied to short-term route-level TSR prediction, which considering temporal and between-route correlations in TSR and determining the

TABLE 13. Comparison of CV for orders data.

M9.6								
OD	Posted Orders				Closed Orders			
	SD	Avg	CV	$\overline{CV}$	SD	Avg	CV	$\overline{CV}$
CTG	9.720	32.618	0.298	0.316	6.640	26.453	0.251	0.267
CTK	7.392	24.314	0.304		6.173	18.649	0.331	
CTQ	3.915	10.439	0.375		2.603	9.604	0.271	
GTC	5.317	13.360	0.398		4.859	11.623	0.418	
GTK	6.379	21.336	0.299		4.617	14.338	0.322	
GTQ	5.548	25.924	0.214		5.031	19.806	0.254	
KTC	4.528	11.699	0.387		1.600	8.988	0.178	
KTG	4.392	11.967	0.367		0.823	3.865	0.213	
KTQ	6.840	25.810	0.265		4.461	22.532	0.198	
QTC	3.913	16.304	0.240		4.122	14.364	0.287	
QTG	7.751	23.070	0.336		6.288	22.701	0.277	
QTK	8.998	29.404	0.306		4.959	24.670	0.201	

M13								
OD	Posted Orders				Closed Orders			
	SD	Avg	CV	$\overline{CV}$	SD	Avg	CV	$\overline{CV}$
CTG	9.820	45.461	0.216	0.192	7.716	40.824	0.189	0.195
CTK	4.285	42.006	0.102		7.106	33.520	0.212	
CTQ	11.778	37.389	0.315		5.583	36.492	0.153	
GTC	5.599	43.739	0.128		11.209	39.890	0.281	
GTK	3.113	31.128	0.100		3.364	23.688	0.142	
GTQ	10.505	37.383	0.281		6.695	33.641	0.199	
KTC	5.070	35.456	0.143		7.130	32.265	0.221	
KTG	6.558	29.944	0.219		4.163	23.656	0.176	
KTQ	10.940	39.494	0.277		10.801	37.243	0.290	
QTC	3.599	36.722	0.098		6.190	35.988	0.172	
QTG	11.237	46.628	0.241		5.097	45.509	0.112	
QTK	8.925	47.728	0.187		7.767	41.094	0.189	

Note: SD and Avg mean the standard deviation and the average, respectively, and the units are both orders/day. $\overline{CV}$ represent the average of CVs.

factors that actually influence route-level prices given an initial variables' set. Additionally, this proposed framework can be applied to forecast contract rates, for example, to examine which rates of adjacent routes actually improve the forecasting accuracy. As such, this model could also be used as a tool for shippers and carriers to use this model to develop their forward projections of the current trucking market in the following days.

Further research could focus on two topics. On the one hand, despite the improved performance, the approach we have presented requires additional verification if data from other regions become available, and there is a need to explore more exogenous variables that may affect the fluctuation of rates and investigate the regional differences among trucking markets. On the other hand, we consider a practical approach to implement the standardization of rates to eliminate the influences from different types of cargo, which is intended to create a general but efficient freight index for the southwestern part of China's trucking industry.

APPENDIX A
See Table 12.

APPENDIX B
See Table 13.

REFERENCES

[1] (2017). American Trucking Associations. *Reports, Trends & Statistics*. Accessed: Mar. 25, 2019. [Online]. Available: https://www.trucking.org/News_and_Information_Reports_Industry_Data.aspx

[2] (2017). Bloomberg. *China Transforms the Trucking Business*. Accessed: Mar. 20, 2019. [Online]. Available: https://www.bloomberg.com/opinion/articles/2017-11-30/china-transforms-the-trucking-business

[3] J. Xiong and D. Bensman, "The heart of the problem: Trucking in China's logistics sector," in *Proc. 62nd Annu. Labor Employment Relations Assoc. (LERA)*, Atlanta, GA, USA, 2018, pp. 273–279.

[4] W. B. Cassidy, "China's trucking sector takes on new life," *J. Commerce*, 2012. Accessed: Dec. 15, 2019. [Online]. Available: https://www.joc.com/chinas-trucking-sector-takes-on-new-life_20121001.html

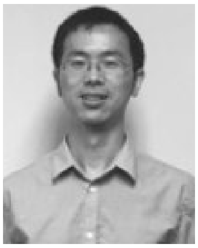
[5] J. Miller, Y. Nie, and A. Stathopoulos, "Crowdsourced urban package delivery: Modeling traveler willingness to work as crowdshippers," *Transp. Res. Record, J. Transp. Res. Board*, vol. 2610, no. 1, pp. 67–75, Jan. 2017.

[6] R. A. Garrido, "Procurement of transportation services in spot markets under a double-auction scheme with elastic demand," *Transp. Res. Part B, Methodol.*, vol. 41, no. 9, pp. 1067–1078, Nov. 2007.
[7] X. Bai, "Forecasting short term trucking rates," M.S. thesis, Dept. Transp. Logistics, Massachusetts Inst. Technol., Cambridge, MA, USA, 2018.
[8] A. Budak, A. Ustundag, and B. Guloglu, "A forecasting approach for truckload spot market pricing," *Transp. Res. Part A, Policy Pract.*, vol. 97, pp. 55–68, Mar. 2017.
[9] B. C. Fisher and E. R. Caldwell, "The impact of lead time on truckload transportation rates," M.S. thesis, Dept. Eng. Syst., Massachusetts Inst. Technol., Cambridge, MA, USA, 2008.
[10] M. G. Kay and D. P. Warsing, "Estimating LTL rates using publicly available empirical data," *Int. J. Logistics Res. Appl.*, vol. 12, no. 3, pp. 165–193, Jun. 2009.
[11] A. Mendoza and J. A. Ventura, "Estimating freight rates in inventory replenishment and supplier selection decisions," *Logistics Res.*, vol. 1, nos. 3–4, pp. 185–196, Dec. 2009.
[12] J. A. Baker, "Emergent Pricing Structures in LTL Transportation," *J. Bus. Logistics*, vol. 12, no. 1, pp. 199–202, 1991.
[13] L. D. Smith, J. F. Campbell, and R. Mundy, "Modeling net rates for expedited freight services," *Transp. Res. Part E, Logistics Transp. Rev.*, vol. 43, no. 2, pp. 192–207, Mar. 2007.
[14] E. Özkaya, P. Keskinocak, V. R. Joseph, and R. Weight, "Estimating and benchmarking Less-than-truckload market rates," *Transp. Res. Part E, Logistics Transp. Rev.*, vol. 46, no. 5, pp. 667–682, Sep. 2010.
[15] C. Lindsey, A. Frei, H. Alibabai, H. S. Mahmassani, Y.-W. Park, D. Klabjan, M. Reed, G. Langheim, and T. Keating, "Modeling carrier truckload freight rates in spot markets," Northwestern Univ., Evanston, IL, USA, Tech. Rep., 2013. [Online]. Available: http://dynresmanagement.com/uploads/3/5/2/7/35274584/3pl_pricing_and_sourcing.pdf
[16] A. Scott, "The value of information sharing for truckload shippers," *Transp. Res. Part E, Logistics Transp. Rev.*, vol. 81, pp. 203–214, Sep. 2015.
[17] J. W. Miller, "ARIMA time series models for full truckload transportation prices," *Forecasting*, vol. 1, no. 1, pp. 121–134, 2019.
[18] N. Fumo and M. A. Rafe Biswas, "Regression analysis for prediction of residential energy consumption," *Renew. Sustain. Energy Rev.*, vol. 47, pp. 332–343, Jul. 2015.
[19] A. Y. Saber and A. K. M. R. Alam, "Short term load forecasting using multiple linear regression for big data," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Honolulu, HI, USA, Nov. 2017, pp. 1–6.
[20] A. E. P. Silva, C. D. C. Freitas, L. V. Dutra, and M. B. Molento, "Assessing the risk of bovine fasciolosis using linear regression analysis for the state of Rio Grande do Sul, Brazil," *Veterinary Parasitol.*, vol. 217, pp. 7–13, Feb. 2016.
[21] J. Seo, T. M. Czaplewski, J.-H. Kimn, and G. Hatfield, "Integrated structural health monitoring system and multi-regression models for determining load ratings for complex steel bridges," *Measurement*, vol. 75, pp. 308–319, Nov. 2015.
[22] K. F. Nimon and F. L. Oswald, "Understanding the results of multiple linear regression: Beyond standardized regression coefficients," *Organizational Res. Methods*, vol. 16, no. 4, pp. 650–674, Oct. 2013.
[23] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "LightGBM: A highly efficient gradient boosting decision tree," in *Proc. Adv. Neural Inf. Process.*, 2017, pp. 3146–3154.
[24] J. Paelinck, and L. Klaassen, *Spatial Econometrics*. Farnborough, U.K.: Saxon House, 1979.
[25] M. F. Goodchild, "Geographical information science," *Int. J. Geographical Inf. Syst.*, vol. 6, no. 1, pp. 31–45, Jan. 1992.
[26] *TruckAlliance*. Accessed: Aug. 29, 2019. [Online]. Available: https://truckingalliance.org/
[27] *Bricham_Marketing, British Consulate-General: Business is Great-Fact sheets for Southwest-Engineering*, Brit. Chamber Commerce Southwest China, Chengdu, China, 2017.
[28] M. Gan, Y. Nie, X. Liu, and D. Zhu, "Whereabouts of truckers: An empirical study of predictability," *Transp. Res. Part C, Emerg. Technol.*, vol. 104, pp. 184–195, Jul. 2019.
[29] *Bricham_Marketing, British Consulate-General: Business is Great-Fact sheets for Southwest-Energy*, Brit. Chamber Commerce Southwest China, Chengdu, China, 2017.
[30] *Bricham_Marketing, British Consulate-General: Business is Great-Fact sheets for Southwest-Technology*, Brit. Chamber Commerce Southwest China, Chengdu, China, 2017.
[31] D. C. LeBlanc, *Statistics: Concepts and Applications for Science*. Boston, MA, USA: Jones & Bartlett, 2004.
[32] Y. Li, Q. He, J. Li, H. Deng, and J. Shen, "Evaluation of toll-free policy on major holidays in China," in *Proc. Transp. Res. Board 95th Annu. Meeting*, Washington, DC, USA, 2016, pp. 3694–3703.
[33] Chinatax. *Labor Management-Wages, Working Hours, Holidays & Social Security Contribution*. Accessed: Feb. 28, 2020. [Online]. Available: https://www.china-tax.net/doing-business-in-china/labor-management-wages-working-hours-holidays.html
[34] A. A. Ariyo, A. O. Adewumi, and C. K. Ayo, "Stock price prediction using the ARIMA model," in *Proc. UKSim-AMSS 16th Int. Conf. Comput. Modeling Simulation*, Cambridge, U.K., Mar. 2014, pp. 106–112.
[35] Z. H. Munim and H.-J. Schramm, "Forecasting container shipping freight rates for the Far East–northern Europe trade lane," *Maritime Econ. Logistics*, vol. 19, no. 1, pp. 106–125, Mar. 2017.
[36] J. Contreras, R. Espinola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1014–1020, Aug. 2003.
[37] Z. H. Munim, M. H. Shakil, and I. Alon, "Next-day bitcoin price forecast," *J. Risk Financial Manage.*, vol. 12, no. 2, p. 103, 2019.
[38] Z. H. Munim and H. Schramm, "Forecasting container freight rates for major trade routes: A comparison of artificial neural networks and conventional models," *Maritime Econ. Logistics*, to be published, doi: 10.1057/s41278-020-00156-5.
[39] F. Li, L. Zhang, B. Chen, D. Gao, Y. Cheng, X. Zhang, Y. Yang, K. Gao, Z. Huang, and J. Peng, "A light gradient boosting machine for remainning useful life estimation of aircraft engines," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Maui, HI, USA, Nov. 2018, pp. 3562–3567.
[40] M. Ustuner and F. B, Sanli, "Polarimetric target decompositions and light gradient boosting machine for crop classification: A comparative evaluation," *ISPRS Int. J. Geo-Inf.*, vol. 8, no. 2, p. 97, 2019.
[41] D. G. Brooks, "Akaike information criterion statistics," *Technometrics*, vol. 31, no. 2, pp. 270–271, May 1989.
[42] P. C. B. Phillips and P. Perron, "Testing for a unit root in time series regression," *Biometrika*, vol. 75, no. 2, pp. 335–346, 1988.
[43] F. X. Diebold and R. S. Mariano, "Comparing predictive accuracy," *J. Bus. Econ. Statist.*, vol. 20, no. 1, pp. 134–144, 1995.
[44] G. Li, H. Song, and S. F. Witt, "Recent developments in econometric modeling and forecasting," *J. Travel Res.*, vol. 44, no. 1, pp. 82–99, Aug. 2005.

WEI XIAO received the B.S. degree in transport engineering from Southwest Jiaotong University, Chengdu, China, in 2013, and the M.S. degree in industry engineering from Texas A&M University, College Station, TX, USA. She is currently pursuing the Ph.D. degree in transportation planning and management with Southwest Jiaotong University. Her main expertise is in the area of econometrics and data science.

CHUAN XU received the B.S. degree in transport engineering from Southwest Jiaotong University, Chengdu, China, in 2009, and the Ph.D. degree from Tongji University, Shanghai, China, in 2014. From 2016 to 2018, he was a Postdoctoral Researcher with New York University. He has been working as an Associate Professor with the School of Transportation and Logistics, Southwest Jiaotong University, since 2019. His main expertise is in the area of traffic safety and data science.

HONGLING LIU received the B.S. degree in transportation from Southwest Jiaotong University, Chengdu, China, in 2019, where she is currently pursuing the M.E. degree in transportation engineering. Her current research interests include traffic planning and management, and traffic data science.



HONG YANG received the B.E. degree in transportation engineering from Southwest Jiaotong University, in 2004, the master's degree in transportation planning and management from Tongji University, in 2007, and the M.S. degree in applied and mathematical statistics and the Ph.D. degree in civil engineering from Rutgers, The State University of New Jersey, USA, in 2010 and 2012, respectively. He was a Postdoctoral Associate with the Department of Civil and Urban Engineering (CUE) and the Center for Urban Science and Progress (CUSP), New York University (NYU). He is currently an Assistant Professor with the Department of Computational Modeling and Simulation Engineering, Old Dominion University, USA. His research interests include smart mobility, transportation safety, urban informatics, data science, statistics, modeling and simulation, spatial data analytics, machine learning, traffic operations, transportation resilience, and emerging technologies in transportation.

XIAOBO LIU received the B.S. and M.E. degrees in transportation engineering from Southwest Jiaotong University, Chengdu, China, in 1999, and the Ph.D. degree in transportation engineering from the New Jersey Institute of Technology, ITE Metropolitan Section of NY&NJ, and the Most Outstanding Student Paper Award. He is currently the Dean of the School of Transportation and Logistics, Southwest Jiaotong University (SWJTU). At the School of Transportation and Logistics, he provides stewardship to faculty members, who secure $6 million in research funding annually. It is home to the National Engineering Laboratory for Transportation Big-Data Technology and Application. At the Laboratory, he leads interdisciplinary research groups working on research areas encompassing Integrated Transportation Planning, Advanced Network Operations, and Intelligent Big-Data-driven Logistics.