A new seasonal cycle GM (1,1) model and its application in railway passenger volume forecasting

Huan Wang, Yuhong Wang and Dongdong Wu
School of Business, Jiangnan University, Wuxi, China

Received 14 November 2020
Revised 18 January 2021
8 March 2021
Accepted 16 March 2021

Abstract

Purpose – To predict the passenger volume reasonably and accurately, this paper fills the gap in the research of quarterly data forecast of railway passenger volume. The research results can also provide references for railway departments to plan railway operation lines reasonably and efficiently.

Design/methodology/approach – This paper intends to establish a seasonal cycle first order univariate grey model (GM(1,1) model) combing with a seasonal index. GM (1,1) is termed as the trend equation to fit the railway passenger volume in China from 2014 to 2018. The railway passenger volume in 2019 is used as the experimental data to verify the forecasting effect of the proposed model. The forecasting results of the seasonal cycle GM (1,1) model are compared with the traditional GM (1,1) model, seasonal grey model (SGM(1,1)), Seasonal Autoregressive Integrated Moving Average (SARIMA) model, moving average method and exponential smoothing method. Finally, the authors forecast the railway passenger volume from 2020 to 2022.

Findings – The quarterly data of national railway passenger volume have a clear tendency of cyclical fluctuations and show an annual growth trend. According to the comparison of the modeling results, the authors know that the seasonal cycle GM (1,1) model has the best prediction effect with the mean absolute percentage error of 1.32%. It is much better than the other models, reflecting the feasibility of the proposed model.

Originality/value – As the previous grey prediction model could not solve the series prediction problem with seasonal fluctuation, and there are few research studies on quarterly railway passenger volume forecasting, GM (1,1) model is taken as the trend equation and combined with the seasonal index to construct a combination forecasting model for accurate forecasting results in this study. Besides, considering the impact of the epidemic on passenger volume, the authors introduce a disturbance factor to deal with the forecasting results in 2020, making the modeling results more scientific, practical and referential.

Keywords GM (1,1) model, Trend equation, Seasonal index, Railway passenger volume, Combination forecasting

Paper type Research paper

1. Introduction

With the continuous development of China's economy, more attentions are paid to railway infrastructure construction. More and more investments have been achieved in railway constructions and operations, and railway transportations have also been continuously developed. Besides, the traveling demands by railway transportations are also increasing. In the long run, the research on the railway passenger volume forecasting is beneficial for the government to carry out macroeconomic regulation and railway construction control. It is also beneficial for the railway departments to plan the railway passenger traffic routes rationally. Short-term railway passenger volume forecasting is critical to real-time management, timely response, public transportation service efficiency and the

The research reported in this paper was partially supported by the National Natural Science Foundation of China (71871106) and the Fundamental Research Funds for the Central Universities (JUSRP1809ZD; 2019JDZD06). The work was also sponsored by the major projects of Educational Science Fund of Jiangsu Province in 13th Five-Year Plan (A/2016/01); the key project of philosophy and social science research in Universities of Jiangsu Province (2018SJZDI051) and the Grey System Theme Innovation Zone Project (GS2019005).

Grey Systems: Theory and Application
Vol. 12 No. 2, 2022
pp. 293-317

2043-9377
DOI 10.1108/GS-11-2020-0146

improvement of medium- and long-term forecasting precision. Therefore, it is of profound significance to analyze and forecast railway passenger volume.

With the growth of economy and population, the continuous development of transportation industry and the expand of railway carrying capacity, the railway passenger volume has an overall upward trend. In addition, due to climate conditions, holidays, cultural customs, industrial and agricultural production activities and other factors, railway passenger volume has periodic fluctuations. For example, by the effect of holiday travel, back-to-school season, Spring Festival travel season and other factors, the passenger traffic shows the character of cyclical changes. The presence of unexpected events will also cause significant random fluctuations in passenger traffic, such as the outbreak of COVID-19. The rapid response of local governments and the adoption of measures such as limiting passenger flow and closing down passenger stations have resulted in large fluctuations in passenger traffic.

Many scholars have paid much attention to passenger traffic volume forecasting in recent years. They proposed different methods to revise the existing forecasting model of passenger traffic volume and traffic flow and study the emergency plan of an emergent situation. Their research results can help predict passenger transportation volume accurately, achieve the rational use of resources and improve service efficiency.

Scholars studied the influencing factors of passenger volume. Pan *et al.* took passenger volumes of each station as a dependent variable, and the explaining variables included employment and population of the station area, the commuting distance of residents, subway web accessibility, station status of transfer stations, coupling with commercial activity centers, etc. (Pan *et al.*, 2017). Banerjee summarized the research on transportation planning from the model principle, forecasting method, demand influence factor, etc. The results indicated a lack of standardization in method description and testing for requirements prediction. It is suggested to facilitate benchmarking of new models by using open-source testing platforms (Banerjee *et al.*, 2020). Li *et al.* identified the main factors affecting passenger volume through factor analysis. And then they reduced the dimensions of the factors, considering the specificity of the selected time series (the data in 2003 were seriously affected by SARS), and they studied them separately as a whole and two time periods. Based on the factor analysis and logistic model, the passenger volume was predicted, and the results were analyzed (Li *et al.*, 2014). Based on the data of different regions in China from 2001 to 2010, Wang *et al.* calculated the elastic coefficients between passenger volume, freight volume and gross domestic product (GDP). They analyzed the relationship between regional traffic volume change and regional economic development. The combination model was used to optimize the prediction model (Wang *et al.*, 2013).

Because of the nonstability and nonlinear characteristics of time series, many scholars analyze it with different methods. Sun *et al.* proposed a new wavelet-support vector machine hybrid model that combines the advantages of wavelet analysis and support vector machine. Wavelet transform was used to decompose the passenger flow data, support vector machine (SVM) was used to predict the different frequency data and then they used the wavelet transform to reconstruct the prediction sequence. The results showed that this method has a good prediction effect, strong robustness and has a strong application prospect (Sun *et al.*, 2015). To better describe the nonstationary, complex and spatial correlation of traffic volume, Peng *et al.* proposed a traffic prediction model based on wavelet denoising and phase space reconstruction (WDSR-GA-BP). The traffic volume was pretreated by the wavelet denoising method, and then the prediction model was established by Back-Propagation (BP) neural network with an optimized genetic algorithm. The prediction error is less than 0.006%, and the precision of the prediction model is improved. It provides a reasonable basis for the planning and development of the transportation system (Peng and Xiang, 2020). Wang *et al.* considered the carbon emission and its decoupling research of transportation (Wang *et al.*, 2017a, b). Zhao *et al.* proposed a nonlinear model to predict the public transport system's passenger flow rate and observed the model's chaotic characteristics. The wavelet analysis method was adopted, and the passenger flow data in a day were decomposed into multi-scale, and then the decomposition sequences were obtained. Subsequently, the neural network method was used to predict sequences. Finally, the predicted value of passenger flow can be predicted by reconstructing the prediction sequence. The results showed that this method is a feasible passenger flow forecasting method (Zhao *et al.*, 2011). Liang *et al.* put forward an integrated forecasting model based on singular spectrum analysis (SSA), considering the noise of time series and its influence on forecasting validity, and they used it to forecast China's air passenger volume (Liang *et al.*, 2017).

Also, many scholars combined deep learning, intelligent algorithm and other methods to conduct in-depth research on passenger volume-related issues. Zhang *et al.* adopted the averaging time-varying model to forecast the passenger volume of the top five airports in China (Zhang *et al.*, 2020). Liu *et al.* proposed a new depth-learning-based hourly traffic prediction model, which can extract the nonlinear features embedded in the input in-depth and abstractly without any labels (Liu *et al.*, 2017). Yang *et al.* developed an improved space-time long-term and short-term memory model (SP-LSTM) to predict the short-term outbound passenger flow (Yang *et al.*, 2021). Li *et al.* proposed a new dynamic radial basis function (RBF) neural network to predict outgoing passenger flow and improve passenger flow control (Li *et al.*, 2019). Hao *et al.* proposed an end-to-end deep learning framework for simultaneous multi-step prediction of all stations in a large metro system (Hao *et al.*, 2019). Li *et al.* proposed a new chaotic accelerating genetic algorithm to predict the urban passenger traffic, which improved the prediction precision of the original model (Li *et al.*, 2012). Wang *et al.* developed a new Bayesian combination method (BCM) to improve the performance of traditional BCM. They linearly combined three single predictors into BCM to make full use of the advantages of each method (Wang *et al.*, 2014). Liu *et al.* designed a passenger counting system based on a convolutional neural network and spatio-temporal context (STC) model (Liu *et al.*, 2017). Tsai *et al.* proposed two new neural network structures for short-term railway passenger demand forecasting (Tsai *et al.*, 2009). Deng *et al.* adopted a hybrid optimization algorithm combining computational intelligence technology to solve highway passenger volume's multi-factor forecasting problem (Deng *et al.*, 2011). Qin *et al.* proposed two kinds of hybrid forecasting methods. It is combined with the echo state network (ESN) based on a grasshopper optimization algorithm (GOA) and adaptive enhancement (Adaboost algorithm) to predict monthly passenger flow in China (Qin *et al.*, 2019). Cheng *et al.* proposed an extended grey model GM (1,1, ebk) for forecasting air passenger volume (Cheng *et al.*, 2019). Zhou *et al.* revised the traditional time series decomposition method for combination forecasting and compared the results with the traditional regression model (Zhou and Wang, 2019).

For the prediction of time series with periodic changes, many scholars have studied the data processing method, the calculation of seasonal index, the combination model prediction and so on. Moreover, the grey seasonal model played an important role.

Using a cycle truncation accumulated generating operation, Xiao *et al.* proposed the seasonal rolling grey forecasting model for traffic flow forecasting (Xiao *et al.*, 2017). Wang *et al.* and Jiang *et al.* combined seasonal adjustment model with the support vector regression (SVR) for forecasting of electricity demand (Wang *et al.*, 2009; Jiang *et al.*, 2020). Zhu *et al.* predicted the first-order moving average (MA) sequence as a trend item, and then made seasonal adjustment compensation weight distribution to form a combined model to predict (Zhu *et al.*, 2011). Holt *et al.* provided a systematic development of the forecasting expressions for exponential weighted moving averages (Holt, 2004). Wang *et al.* proposed a new grey model based on data grouping, and buffer operator is to forecast the residential solar energy consumption in the United States (Wang *et al.*, 2019).

Wang *et al.* optimized the GM (1,1) and SGM (1,1) models, and the calculation method of seasonal index was improved for higher forecast precision of grey model (Wang *et al.*, 2012a, b). Carmona *et al.* improved the damp trend grey model (DTGM) with a dynamic seasonal damping factor to forecast routes passenger demand (Carmona-Bentez and Nieto, 2020). Wu *et al.* establish hybrid models by combining seasonal exponential adjustment method (SEAM) with the regression methods to handle the seasonal time series (Wu *et al.*, 2013). Li *et al.* combined the grey seasonal model with particle swarm optimization to improve the forecasting accuracy of the grey model (Li *et al.*, 2021). Based on the Hodrick-Prescott filter (HP filter), Qian and Wang proposed an improved seasonal GM (1,1) model to forecast the wind power generation in China (Qian and Wang, 2020). Zhou *et al.* added seasonal index to the nonlinear grey Bernoulli model (NGBM) to construct the seasonal nonlinear grey Bernoulli model nonlinear grey Bernoulli model (SNGBM), which obtained significant effect on air quality prediction of the Yangtze river (Zhou *et al.*, 2020a, b, c and d). Wang *et al.* used seasonal adjustment and adaptive particle swarm optimization in a chaotic system for forecasting of short-term electricity price in the Australian power market (Wang *et al.*, 2012a, b). To overcome the problems of seasonality and limited data, Xia and Wong (2014) proposed a new seasonal discrete grey forecasting model based on cycle truncation accumulation to improve sales forecasting accuracy (Xia and Wong, 2014). Incorporating the seasonal dummy variables into the conventional model, Zhou *et al.* put forward a novel discrete grey seasonal model (DGM (1,1)) to deal with data with monthly and quarterly seasonal fluctuations (Zhou *et al.*, 2020a). Using a novel seasonal GM (1,1) model with dynamic seasonal adjustment factors, Wang *et al.* forecasted the industrial solar energy consumption (Wang *et al.*, 2020). Fang *et al.* combined the GM (1,1) grey forecasting model and the ratio-to-moving-average deseasonalization method to forecast time series with seasonality characteristics (Tseng *et al.*, 2001). Zhou *et al.* proposed a grey seasonal least square support vector regression, which combined the dummy variables, the framework of the LSSVR model and grey accumulation generation operation to reflect seasonal variations in functional forms, variables and parameters (Zhou *et al.*, 2020a, b, c and d).

Through reviewing the relevant literature of the grey seasonal mode, we can know that in the aspect of model optimization, most of the existing grey seasonal prediction models are improved by combining with the algorithm to improve the calculation precision, or the combined models are constructed to optimize the models. However, there are still few studies on the improvement of seasonal index or the characteristics of data. In the practical application of the model, the existing grey seasonal prediction models are mostly used in energy prediction but lack of other research on application of grey prediction model. Through reviewing the relevant literature of railway passenger volume forecasting, it can be seen that the research of railway passenger volume forecasting is mostly based on the combination forecasting. Usually, two or more models are combined, or by adjusting the weight of the internal model of the combination forecasting model to achieve the optimal combination forecasting model of related data, but rarely considering the characteristics of the data itself. Based on the analysis of the historical data of railway passenger volume, it is found that the data on railway passenger volume in China have a periodic changing trend and an increasing trend. In this paper, the growth trend is fitted by the GM (1,1) model, and the quarterly data of railway passenger volume from 2014 to 2018 are used as the training data to fit the model. The seasonal cycle GM (1,1) model is established to forecast the quarterly data of railway passenger volume in 2019. The forecasting results of seasonal cycle GM (1,1) model are compared with other forecasting models. Finally, this paper adopted the proposed method to forecast the later period railway quarterly passenger volume.

The remainder of the paper is structured as follows. Section 2 presents the methodology. Section 3 presents the results of empirical analysis. Section 4 concludes this paper.

2. Methodology

2.1 The traditional GM (1,1) model

Since Professor Deng put forward the grey system theory (Deng, 1982), it has been developing continuously and continually optimized and widely used. Ye *et al.* optimized interval grey number, combined with the discrete grey GM (1,1) model, and the improved model was used to forecast new energy consumption. Based on multivariable dynamic optimization (Ye *et al.*, 2019), Wang *et al.* improved the initial condition to optimize the GM (1,1) model (Wang *et al.*, 2010) and proposed a kind of nonlinear strengthening operators for predicting the output value of China's marine electric power industry (Wang *et al.*, 2016), and Wang and Lu improved GM (1,1) model and used it to predict the natural gas consumption (Wang and Lu, 2020). Wang *et al.* predicted China's hydroelectric power consumption with a grey forecasting model based on data grouping (Wang *et al.*, 2017a, b). Wang *et al.* predicted electricity consumption of primary economic sectors with seasonal GM (1,1) model (Wang *et al.*, 2018a, b). Xiong *et al.* constructed a grey extension model based on iterative reweighted least squares method (IRLS) and applied it to smog pollution, which proved the superiority of the GM (1,1, U (*t*)) model in simulation and prediction under iterative least squares (Xiong *et al.*, 2019). Wang *et al.* combined the ARIMA model with a nonlinear grey forecasting model considering information updating to forecast American oil production (Wang *et al.*, 2018a, b).

Assuming that the initial data sequence for modeling is $X^{(0)}$, $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)\}$.

The first-order cumulative generated operator is $X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(n)\}$,

$$X^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), \quad k = 1, 2, \ldots, n. \tag{1}$$

The original GM (1,1) model is

$$x^{(0)}(k) + ax^{(1)}(k) = b \tag{2}$$

Its whiten differential equation is

$$\frac{dx^{(1)}}{dt} + ax^{(1)}(t) = b \tag{3}$$

The background value sequence is $Z^{(1)} = \{z^{(1)}(2), z^{(1)}(3), \ldots, z^{(1)}(n)\}$, then

$$z^{(1)}(k) = 0.5\left(x^{(1)}(k) + x^{(1)}(k-1)\right), \quad k = 2, 3, \ldots n \tag{4}$$

The parameter series $[a, b]^T$ in GM (1,1) model is carried out by the least square estimation.

$$[a, b]^T = \left(B^T B\right)^{-1} B^T C \tag{5}$$

where $B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \ldots & \ldots \\ -z^{(1)}(n) & 1 \end{bmatrix}$, $C = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \ldots \\ x^{(0)}(n) \end{bmatrix}$.

The time response to the whiten differential equation is

$$\hat{x}^{(1)}(k) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{-a(k-1)} + \frac{b}{a}, \quad k = 2, 3, .., n. \tag{6}$$

The cumulative reduction formula is

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1) = (1-e^{a})\left(x^{(0)}(1) - \frac{b}{a}\right)e^{-a(k-1)}, \quad k = 2,3,..,n. \tag{7}$$

2.2 The seasonal cycle GM (1,1) model

2.2.1 Choice of trend equation. The ordinary trend equation includes a linear equation, exponential equation, power equation and so on. The grey prediction model has the advantage of dealing with small sample data, and it is suitable for forecasting the data with the development trend characteristic (Deng, 1989). The grey univariate model is chosen as the trend equation, the trend value is T_t and the actual value is y_t.

The GM (1,1) model is taken as the trend term equation of the seasonal cycle GM (1,1) model. So, the trend equation is

$$T_t = (1-e^{a})\left(x^{(0)}(1) - \frac{b}{a}\right)e^{-a(t-1)}, \quad t = 2,3,..,n. \tag{8}$$

2.2.2 The calculation of the seasonal index. The seasonal index reflects the average degree that the actual value deviates from the trend value due to seasonal effects in the current period (Wang *et al.*, 2018a, b). The calculation method of seasonal index can be divided into two ways according to the sequence of solving seasonal index (Fan *et al.*, 2012).

According to $I_t = \frac{y_t}{T_t}$, T_t is the trend value, y_t is the actual value, $t = 1,2,......,n$. In the traditional method, the trend equation is not determined first, and the mean of each period is used to replace the trend value, and then the seasonal index is obtained. But the original data are not processed by the deseasonal trend and has seasonal characteristics. The seasonal index obtained in this way is not accurate enough. Thus, this paper adopts the method based on Fan's paper. (Fan *et al.*, 2012). First, the long-term trend without seasonal periodic fluctuation is determined, then the seasonal index is calculated and finally the forecasting model is established.

First, the trend equation is determined, then the initial value of the seasonal index is calculated according to $I_t = \frac{y_t}{T_t}$, then the average value of the seasonal index is normalized and finally the final seasonal index is obtained.

The steps for calculating the seasonal index are as follows:

Step 1: calculate the preliminary seasonal index: $\tilde{I}_t$

$$\tilde{I}_t = \frac{y_t}{T_t} \tag{9}$$

T_t is the trend value, y_t is the actual value, $t = 1,2,......,n$.

Step 2: calculate the value $\bar{I}_i$, the mean of $\tilde{I}_t$ in the same season.

$$\bar{I}_i = \frac{\sum_{i}^{i+m-1}\tilde{I}_i}{m}, i = 1,2,......,L. \tag{10}$$

m is the number of cycles, and L is the numbers of data in a cycle.

Step 3: calculate the mean of the overall seasonal index: I

$$I = \frac{\sum_{i=1}^{L}\bar{I}_i}{L}, \quad i = 1,2,......L. \tag{11}$$

Step 4: calculate the final value of the seasonal index: I_i

$$I_i = \frac{\bar{I}_i}{\bar{I}}, \quad i = 1, 2, \ldots\ldots L. \tag{12}$$

2.2.3 The construction of the seasonal cycle GM (1,1) model. Once the trend equation is determined, the seasonal index can be calculated according to the trend equation, and then the seasonal cycle GM (1,1) model can be constructed.

The seasonal cycle forecast model equation is written as $\hat{y}_t$, $t = 2, 3, \ldots\ldots, n$,

$$\hat{y}_t = \hat{T}_t \cdot \hat{I}_t \tag{13}$$

where $\hat{T}_t$ is the forecast value of the trend item, and $\hat{I}_t$ is the forecast value of seasonal index.

We can make a further prediction according to the derived seasonal index.

The equation for the seasonal cycle prediction model is $\hat{y}_t$, $t = 2, 3, \ldots\ldots, n$, $\hat{y}_t = \hat{T}_t \cdot \hat{I}_t$, GM(1,1) is the trend equation of the model.

$$\hat{T}_t = (1 - e^a)\left(x^{(0)}(1) - \frac{b}{a}\right)e^{-a(t-1)}, \quad t = 2, 3, .., n \tag{14}$$

So, the seasonal cycle forecasting model equation are denoted as

$$\hat{y}_t = (1 - e^a)\left(x^{(0)}(1) - \frac{b}{a}\right)e^{-a(t-1)} \cdot I_i, \quad t = 2, 3, .., n; i = 1, 2, \ldots\ldots, L \tag{15}$$

The quarterly periodic GM (1,1) forecasting model is obtained, which can be used to forecast the national passenger traffic data.

2.3 The standard for error evaluation

To evaluate the validity of the model, we need to determine the error evaluation criteria. We have three evaluation criteria: root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE). Here, we take the MAPE as the evaluation standard of the fitting and prediction error.

The error is calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e^{(2)}(i)}, \tag{16}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |e(i)|, \tag{17}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{e(i)}{x^{(0)}(i)}\right| \times 100\%, \tag{18}$$

$$e(i) = \hat{x}^{(0)}(i) - x^{(0)}(i). \tag{19}$$

We think that the MAPE in the $[1\%, 5\%]$ interval is less error, which means the model fit or forecast effect is better.

2.4 Article flow chart

The flow chart of this article is shown in Figure 1. First, we have preliminary literature review and research, data acquisition, primary data processing. And then, selecting the trend

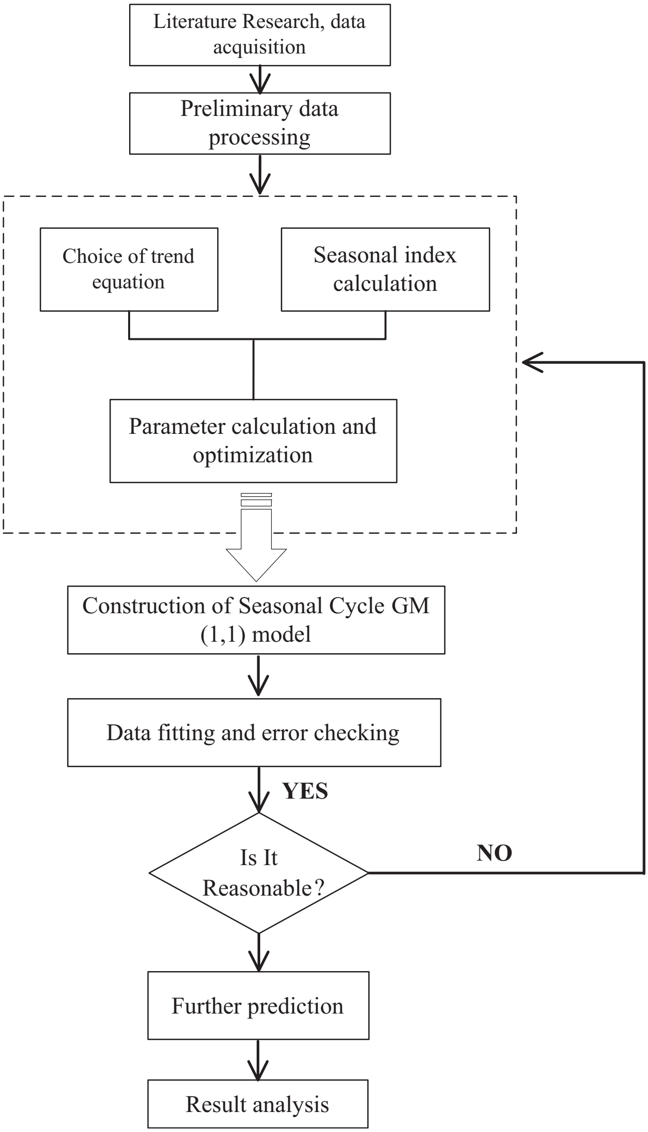

Figure 1.
Flow chart of the paper

equation of the model, calculating the seasonal index, calculating and optimizing the parameters. The seasonal cycle GM (1,1) model is constructed to fit the data, detect the error and judge the accuracy of the model. Finally, we use it to make a further prediction and analyze the results.

3. Empirical analysis

This article intends to model the quarterly data of national railway passenger volume from 2014 to 2018 as fitting sets. The quarterly data of national railway passenger volume in 2019 are taken as experimental data and then compared the predicted data with the real data. We also forecast passenger traffic for 2020–2022. (Data source: the National Bureau of Statistics of the People's Republic of China)

Quarterly data on national passenger volume for 2014–2019 are shown in Table 1:

As Figure 2 shows, the quarterly data of national passenger volume clearly show cyclical fluctuations and show an annual growth trend.

As shown in Figure 3, the national passenger traffic shows an increasing trend in a single quarter. We can know that the national passenger traffic data has both the growth trend and the characteristics of cyclical fluctuations.

The quarterly data of national passenger volume from 2014 to 2018 is considered fitting sets to construct the seasonal cycle GM (1,1) model. The quarterly data of national passenger traffic volume in 2019 are taken as testing data to test the forecast precision and accuracy of the proposed model.

Quarter/Year	2014	2015	2016	2017	2018	2019
*Q*1	53,079	58,693	66,515	72,905	78,257	85,314
*Q*2	58,336	62,925	69,969	76,978	83,561	92,072
*Q*3	66,887	72,116	78,743	84,953	94,869	103,327
*Q*4	57,402	59,766	66,178	73,543	80,808	85,289

Table 1. 2014–2019 quarterly data of national passenger volume (unit: ten thousand)

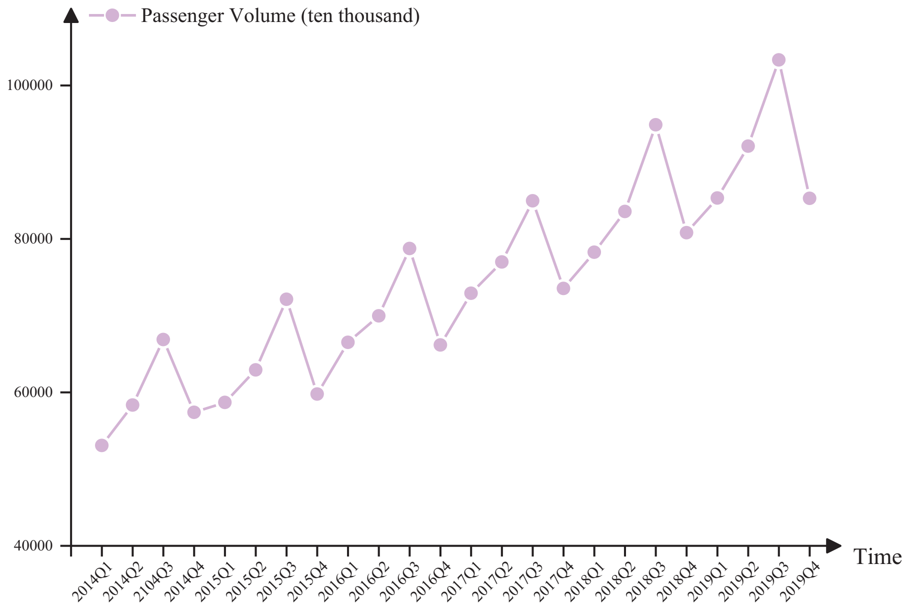

Figure 2. 2014–2019 quarterly data of national railway passenger volume

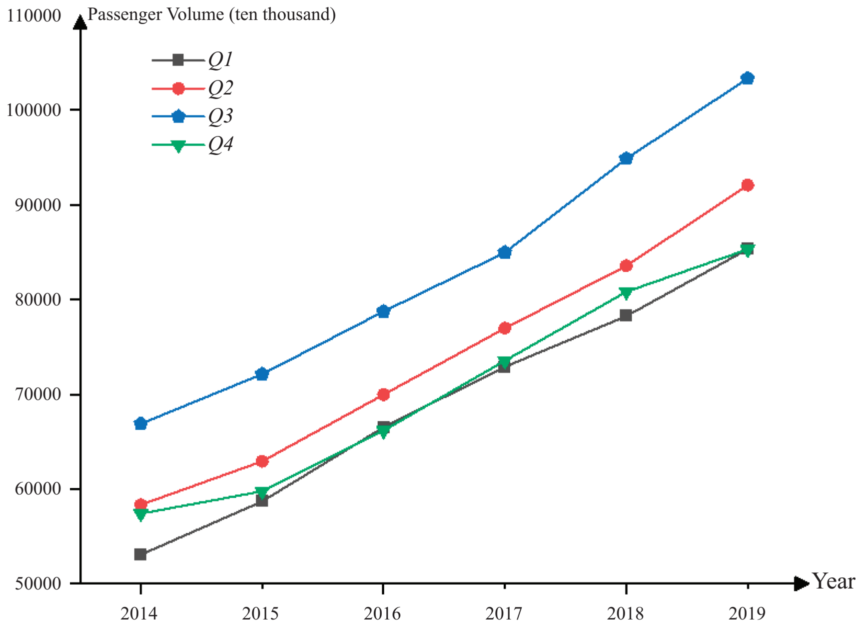

Figure 3.
2014–2019 quarterly trend chart of national railway passenger volume

The original sequence $X^{(0)}$	$X^{(1)}$	$Z^{(1)}$
53,079	53,079	–
58,336	111,415	82,247
66,887	178,302	144858.5
57,402	235,704	207,003
58,693	294,397	265050.5
62,925	357,322	325859.5
72,116	429,438	393,380
59,766	489,204	459,321
66,515	555,719	522461.5
69,969	625,688	590703.5
78,743	704,431	665059.5
66,178	770,609	737,520
72,905	843,514	807061.5
76,978	920,492	882,003
84,953	1005,445	962968.5
73,543	1078,988	1042216.5
78,257	1157,245	1118116.5
83,561	1240,806	1199025.5
94,869	1335,675	1288240.5
80,808	1416,483	1376,079

Table 2.
The calculation result of $X^{(1)}$ and $Z^{(1)}$

3.1 Model establishment

3.1.1 The seasonal cycle GM (1,1) model. First, we calculate parameters of trend equation.

The original sequence is $X^{(0)}$, the first-order accumulation sequence $X^{(1)}$ and the sequence of the nearest mean of $X^{(1)}$ is $Z^{(1)}$.

They were calculated and shown in Table 2.

The grey development coefficient a and the grey action quantity b can be obtained by grey modeling software, according to the least squares:

$$a = -0.022 \quad b = 56660.062$$

According to the calculation steps and principles in Section 2.1, the time response sequence of the trend equation is obtained:

$$\hat{x}^{(1)}(k) = 2628533.5455e^{0.022(k-1)} - 2575454.5455, \quad k = 1, 2, .., n$$

so, the trend equation is $T_t = 2628533.5455e^{0.022(t-1)} - 2575454.5455, \quad t = 1, 2, .., n.$

Then, we can calculate the seasonal index.

Follow the steps in Section 2.2, we can calculate the seasonal index:

Step 1: calculate the preliminary seasonal index: $\tilde{I}_t$.

Step 2: calculate the value $\bar{I}_i$ – the mean of $\tilde{I}_t$ in the same season.

Step 3: calculate the mean of the overall seasonal index: $I = 1.0002$.

Step 4: calculate the final value of the seasonal index: I_i.

The results are shown in Table 3:

Finally, we can get the forecasting function: $\hat{y}_t = (2628533.5455e^{0.022(t-1)} - 2575454.5455) \cdot I_i, \quad t = 2, 3, .., n$, and I_i can be obtained according to Table 3.

Thus,

$$\hat{y}_t = \begin{cases} (2628533.5455e^{0.022(t-1)} - 2575454.5455) \times 0.971434, & t = 4k-3, \quad k = 1, 2, \cdots n \\ (2628533.5455e^{0.022(t-1)} - 2575454.5455) \times 1.000887, & t = 4k-2, \quad k = 1, 2, \cdots n \\ (2628533.5455e^{0.022(t-1)} - 2575454.5455) \times 1.107702, & t = 4k-1, \quad k = 1, 2, \cdots n \\ (2628533.5455e^{0.022(t-1)} - 2575454.5455) \times 0.9199775, & t = 4k, \quad k = 1, 2, \cdots n \end{cases}$$

and the training and testing results are shown in Tables 4 and 5, respectively.

3.1.2 The traditional GM (1,1) model. According to Section 2.1 and using MATLAB toolbox, we can get the parameter vector $\hat{\alpha}$ of the traditional GM (1,1) model.

Where $\hat{\alpha} = \begin{bmatrix} a \\ b \end{bmatrix}^T = [-0.022, 56660.062]$.

The time response function is $\hat{x}^{(1)}(k) = 2628533.5455e^{0.022(t-1)} - 2575454.5455$, and the training and testing results are shown in Tables 4 and 5, respectively.

3.2 Comparison of fitting and prediction results

According to Section 2.3 and 3.1.1, we can forecast the passenger volume data in 2019, and $\hat{T}_t$ is the trend item forecast value and $\hat{I}_t$ is the seasonal index forecast value.

	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Preliminary seasonal index $\tilde{I}_t$	1	0.997804	1.119225	0.939659033
	0.939933	0.985828	1.105292	0.896121999
	0.975663	1.004046	1.10542	0.9088585
	0.979507	1.011775	1.092355	0.925110609
	0.963037	1.005983	1.117323	0.931057201
Mean values $\bar{I}_i$	0.971628	1.001087	1.107923	0.920161468
Final values I_i	0.971434	1.000887	1.107702	0.919977512

Table 3. Seasonal index

Table 4.
Comparison of fitting models during 2014–2018

Time	The real value	Seasonal cycle GM (1,1) model Fitting value	Absolute error (%)	SGM(1,1) model Fitting value	Absolute error (%)	SARIMA model Fitting value	Absolute error (%)	GM(1,1) model Fitting value	Absolute error (%)	Exponential smoothing method Fitting value	Absolute error (%)	Moving average method Fitting value	Absolute error (%)
2014Q1	53,079	51562.73	2.86	53079.00	0.00	–	–	53079.00	0.00	–	–	–	–
2014Q2	58,336	58516.28	0.31	57847.27	0.84	–	–	58464.41	0.22	53079.00	9.01	–	–
2104Q3	66,887	66198.31	1.03	66839.72	0.07	–	–	59761.87	10.65	62744.60	6.19	–	–
2014Q4	57,402	56199.69	2.09	58042.80	1.12	–	–	61088.12	6.42	59004.78	2.79	–	–
2015Q1	58,693	60660.01	3.35	57890.49	1.37	–	–	62443.80	6.39	58786.53	0.16	58926.00	0.40
2015Q2	62,925	63886.19	1.53	63193.81	0.43	–	–	63829.57	1.44	61683.46	1.97	60329.50	4.12
2015Q3	72,116	72273.19	0.22	73017.38	1.25	72069.32	0.06	65246.09	9.53	68986.24	4.34	61476.75	14.75
2015Q4	59,766	61357.02	2.66	63407.40	6.09	63053.72	5.50	66694.04	11.59	62532.07	4.63	62784.00	5.05
2016Q1	66,515	66226.66	0.43	63241.02	4.92	63700.88	4.23	68174.13	2.49	65320.12	1.80	63375.00	4.72
2016Q2	69,969	69748.89	0.31	69034.49	1.34	68618.64	1.93	69687.07	0.40	68574.34	1.99	65330.50	6.63
2016Q3	78,743	78905.55	0.21	79766.02	1.30	78035.51	0.90	71233.58	9.54	75692.40	3.87	67091.50	14.80
2016Q4	66,178	66987.62	1.22	69267.83	4.67	66317.72	0.21	72814.42	10.03	69032.32	4.31	68748.25	3.88
2017Q1	72,905	72304.14	0.82	69086.07	5.24	73110.36	0.28	74430.33	2.09	71743.20	1.59	70351.25	3.50
2017Q2	76,978	76149.60	1.08	75415.01	2.03	76684.56	0.38	76082.11	1.16	75407.56	2.04	71948.75	6.53
2017Q3	84,953	86146.55	1.40	87138.39	2.57	85502.14	0.65	77770.54	8.45	82089.37	3.37	73701.00	13.24
2017Q4	73,543	73134.94	0.55	75669.91	2.89	73425.98	0.16	79496.44	8.10	76106.91	3.49	75253.50	2.33
2018Q1	78,257	78939.33	0.87	75471.35	3.56	80228.95	2.52	81260.64	3.84	77611.97	0.82	77094.75	1.49
2018Q2	83,561	83137.69	0.51	82385.24	1.41	84279.41	0.86	83064.00	0.59	81776.29	2.14	78432.75	6.14
2018Q3	94,869	94052.03	0.86	95192.16	0.34	92237.47	2.77	84907.38	10.50	90941.19	4.14	80078.50	15.59
2018Q4	80,808	79846.38	1.19	82663.70	2.30	81520.03	0.88	86791.66	7.40	83847.96	3.76	82557.50	2.17
MAPE			1.18		2.19		1.52		5.54		3.29		6.58

Table 5. Comparison of forecasting results of different models in 2019

Time	The real value	Seasonal cycle GM (1,1) model		SGM(1,1) model		SARIMA model		GM(1,1) model		Exponential smoothing method		Moving average method	
		Predicted value	Absolute error (%)	Predicted value	Absolute error (%)	Predicted value	Absolute error (%)	Predicted value	Absolute error (%)	Predicted value	Absolute error (%)	Predicted value	Absolute error (%)
2019Q1	85,314	86183.43	1.02	82446.79	3.36	86308.58	1.17	88717.76	3.95	84874.19	0.52	84373.75	1.10
2019Q2	92,072	90767.06	1.42	89999.69	2.25	91718.72	0.38	90686.61	1.53	89912.66	2.35	86412.67	6.15
2019Q3	103327	102682.99	0.62	103990.29	0.64	102986.38	0.33	92699.15	10.35	99302.70	3.89	87838.50	14.99
2019Q4	85,289	87173.71	2.21	90303.90	5.88	89510.04	4.95	94756.35	10.86	89493.11	4.93	80808.00	5.25
MAPE			1.32		3.03		1.71		6.67		2.92		6.87

Through the trend equation of GM (1,1), the quarterly data of national passenger traffic volume in 2019 are preliminarily forecasted as follows: 88,717.761, 90,686.607, 92,699.145, 94,756.347

According to the corresponding seasonal index for 2019 national passenger volume quarterly data, the final forecast values are 86,183.4283, 90,767.05922, 102682.9894, 87,173.70839

The predicted results are compared with the actual values, as shown in Table 6:

It can be seen from Table 6, the absolute percentage error of seasonal cycle GM (1,1) is 1.02, 1.42, 0.62, 2.21% for four quarters of 2019. The MAPE of the forecast result is 1.32%.

And we compared fitting and prediction results of the seasonal cycle GM (1,1) model with the traditional GM (1,1) model (Deng, 1989), SGM (1,1) model (Wang *et al.*, 2018a), SARIMA model (Zhou *et al.*, 2020c), moving average method and exponential smoothing method. We can get the results as shown in Tables 4 and 5. Also, charts of prediction results and error about these models' quarterly national railway passenger volume are compared for further analysis.

As shown in Table 4, the fitting result of the seasonal cycle GM (1,1) model is the best. It can be seen that the MAPE of quarterly data fitting from 2014 to 2018 is 1.18%, followed by the SARIMA model with a MAPE value of 1.52%. The SGM (1,1) model also has a good fitting accuracy with a MAPE value of 2.19%. The traditional GM (1,1) model, the exponential method and the moving average method have fitting effect with a MAPE value of 5.54, 3.29 and 6.58%, respectively. By comparing the fitting results of the six models, it can be seen that the seasonal cycle GM (1,1) model has apparent advantages in fitting the quarterly data of national passenger volume. Its MAPE values ranging from 0.21 to 3.35%, indicating that the fitting error fluctuates slightly. Therefore, it is feasible and credible to use the seasonal cycle GM (1,1) model to predict the national passenger traffic quarterly data in 2019.

The fitting effect of the SGM (1,1) model, the SARIMA model, exponential smoothing method, traditional GM (1,1) model and the moving average method are not as good as those of the seasonal cycle GM (1,1) model. The MAPE values of the SGM (1,1) model, the SARIMA model and the exponential smoothing method fitting are within 5%, while the MAPE value of the traditional GM (1,1) model and the moving average method is within 6%. They can be used as a control group with the seasonal cycle GM (1,1) model to forecast the national passenger volume. According to the results, we can compare each model's forecast accuracy and analyze the reasons for the results.

It can be seen from Table 5 that the forecast value of the national railway passenger volume quarterly data in 2019. The seasonal cycle GM (1,1) model has the best prediction effect, with the MAPE value of 1.32%. The SARIMA model has a prediction accuracy with a MAPE value of 1.71%. The prediction errors of SGM (1,1) model, the exponential smoothing method, the GM (1,1) model and the moving average method are larger, which are 3.03, 2.92, 6.67 and 6.87%, respectively. The forecast errors of the seasonal cycle GM (1,1) model are relatively small, and the values of forecast error are all less than 2.5%. Besides, the MAPE value of the exponential smoothing method and SARIMA model also fluctuate slightly; both are less than 5%. While the traditional GM (1,1) model, SGM (1,1) model and the moving average method have more massive prediction error, the absolute percentage error of a single

Table 6.
Comparison of predictive value and actual value

Time	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Predictive value	86183.43	90767.06	102682.99	87173.71
The real value	85,314	92,072	103,327	85,289
Absolute error (%)	1.02%	1.42%	0.62%	2.21%

prediction value of GM (1,1) model and the moving average method exceed 10%. It also shows that the prediction results of traditional GM (1,1) and the moving average method are not stable enough. The prediction accuracy and stability of national passenger traffic volume are not as good as the seasonal cycle GM (1,1) model.

To have a better understanding of the models and methods, the fitting and prediction results will be further analyzed.

To have a clear comparison of the prediction effect of these models in the railway passenger volume forecasting, we compared the fitting and prediction results of these models.

As shown in Figures 4 and 5, data fitting and prediction results in each model are compared with the real value. Figures 4 and 5 show that the fitting curves of seasonal cycle model are very close to the actual data for the fitting of quarterly data from 2014 to 2018 and the prediction of quarterly data of national passenger traffic in 2019. From Figure 4, we can know that the SGM (1,1) model and the SARIMA models also have close curves with the actual data. But the SGM (1,1) model has significant deviation in the first quarter and the fourth quarter, and the SARIMA model has initial deviation from the actual data and gets closer and closer to the true value curve gradually. The traditional GM (1,1) model only has a linear growth trend without periodic fluctuation characteristics.

From Figure 5, we can know that the exponential smoothing method shows periodic fluctuations and trend growth, which is also close to the actual value curve. The simple GM (1,1) model and the moving average method have no obvious periodic fluctuations, so they have larger deviation than other models from the actual data curve.

Combining the comparison of Figures 4 and 5, we can know that the seasonal cycle model, the SGM (1,1) model, the SARIMA model and exponential smoothing method show periodic fluctuations and trend growth similar to the real value. In contrast, the fitting and prediction curves corresponding to the traditional GM (1,1) model and the moving average method only have a linear growth trend but no periodic fluctuation characteristics. It shows that the seasonal cycle model, the SGM (1,1) model, the SARIMA model and exponential smoothing method are more suitable for forecasting the national passenger traffic quarterly data than GM (1,1) model and the moving average method.

To further analyze each model's fitting and prediction error, we did some work, as shown in Figures 6 and 7.

According to Figures 6 and 7, we can see that the seasonal cycle GM (1,1) model not only has the smallest fitting and prediction error but also has the slightest fluctuations in fitting and prediction error. It means the fitted value and predicted value deviate from the real data to a smaller extent. The effect of the SARIMA model, the SGM (1,1) model and the exponential smoothing method is good, while the simple GM (1,1) model and the moving average method have relatively large errors and large fluctuations. It shows the deviation among the fitting value, the predicted value and the actual value is large.

Figure 8 compares the MAPE fitting and prediction of each model. It can be further learned from Figure 8 that the seasonal cycle GM (1,1) has the smallest MAPE for fitting value and prediction value among the four models. And the MAPE of fitting value is less than 1.5%. The MAPE of fitting and forecasting by SARIMA model and SGM (1,1) model and the exponential smoothing method is less than 4%. In comparison, the GM (1,1) model and the moving average method are more significant than 5%.

Comparing the prediction results, it can be seen that the national passenger traffic data have both a periodic fluctuation trend and a linear growth trend. The results predicted by the traditional GM (1,1) model and the moving average method are not optimal because the conventional GM (1,1) model, and the moving average method can effectively predict data

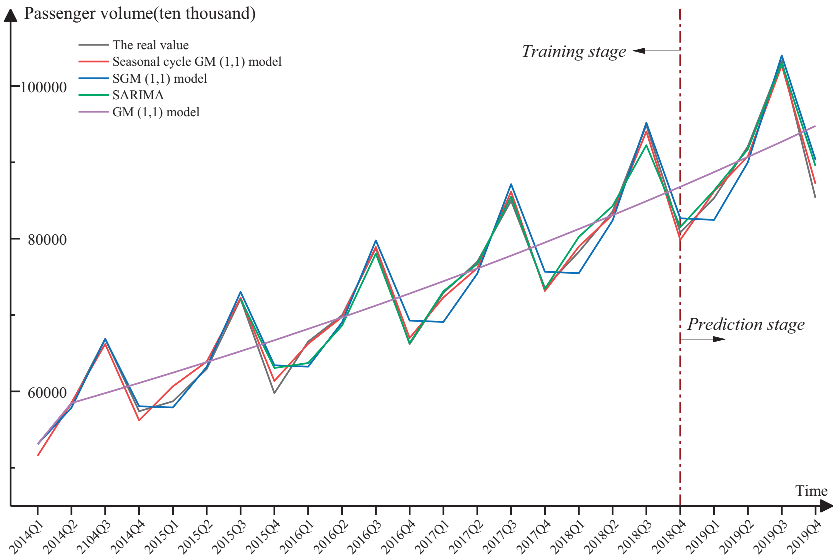

Figure 4.
Comparison chart of fitting and forecasting results9(a)

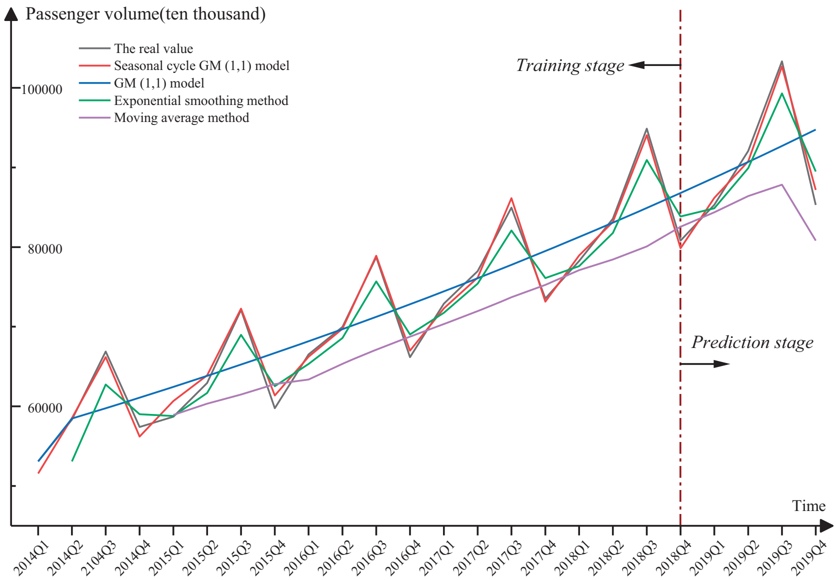

Figure 5.
Comparison chart of fitting and forecasting results(b)

that only have a growth trend. Still, they have no advantage in processing data with periodic fluctuations.

The SARIMA model and SGM (1,1) model are great at dealing with data with periodic fluctuations; thus, they have distinct advantages to forecast the railway passenger volume. It can be seen from Figure 4 that the curve of the SGM (1,1) model closes to the real data curve, but it has significant deviation from the real data at the first quarter and the fourth quarter. It

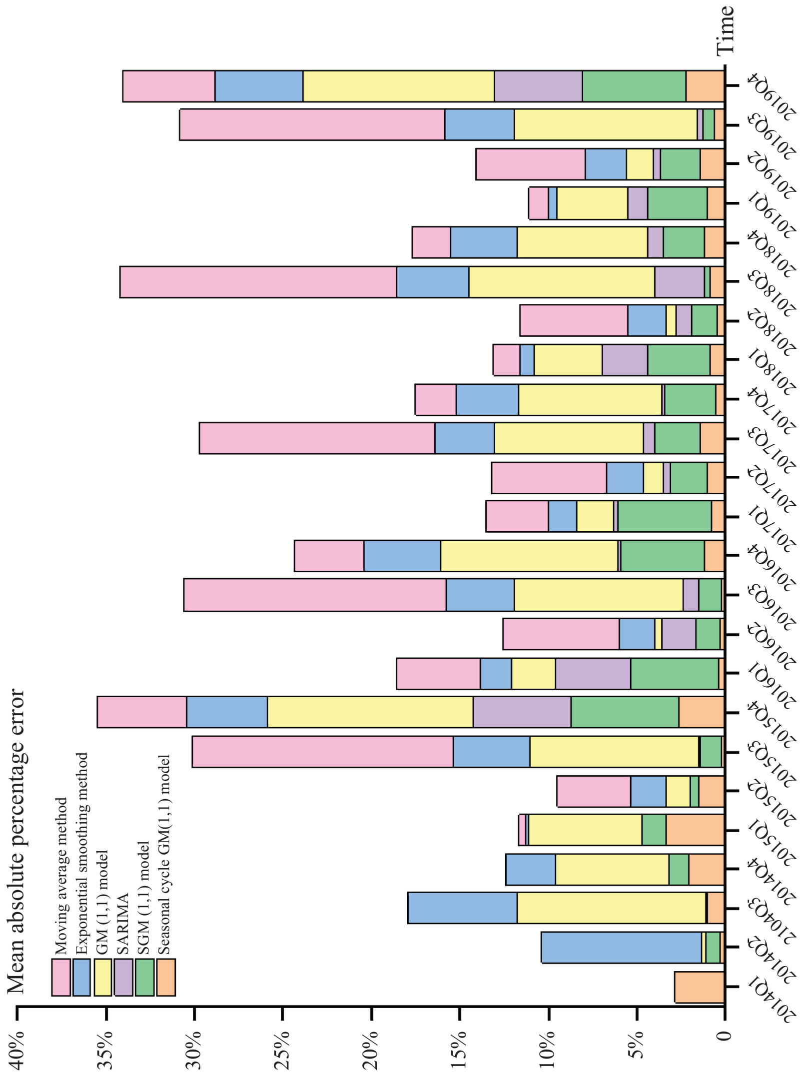

Figure 6.
Column stacking diagram of fitting and prediction errors of each model

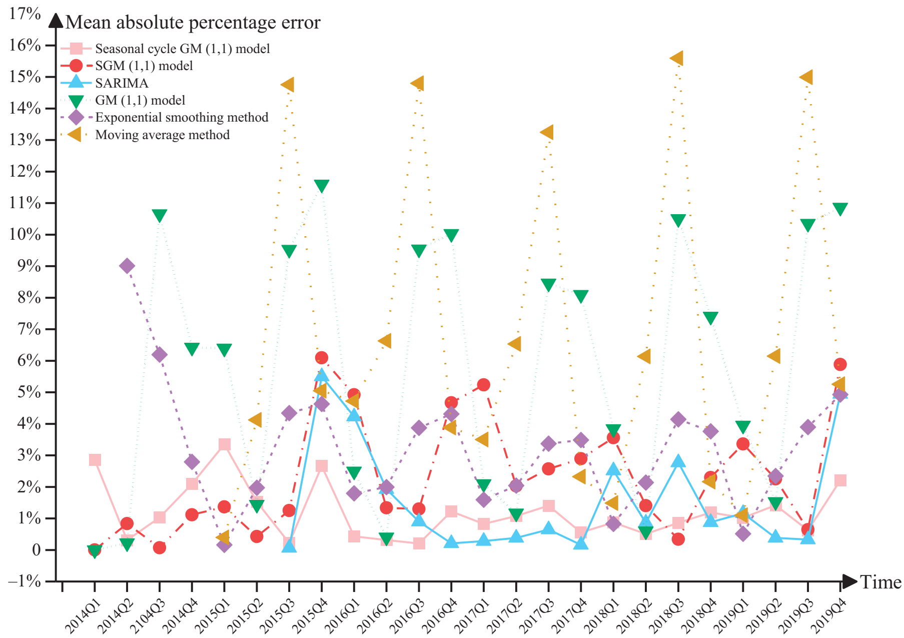

Figure 7.
Line chart of fitting and prediction error of each model

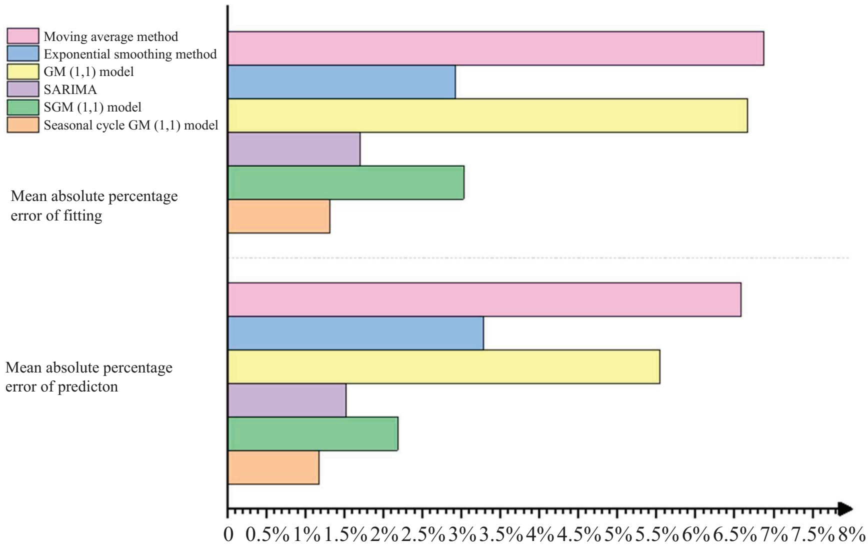

Figure 8.
Comparison chart of MAPE between models fitting and prediction

results from the deviation of the seasonal index in the first quarter and the fourth quarter. The SGM (1,1) model combines the seasonal index during the grey modeling process rather than modeling with the initial data, and the calculation of seasonal index only dependents on the real data can also produce error.

About the SARIMA model, although it has initial deviation from the real data in the first quarter and the fourth quarter of 2015, it closes to the curve of the real data gradually. Considering the trend, periodic change and random disturbance and the delay parameter, the SARIMA model can reflect the trend and change of the original time series and has great effect on prediction. The SARIMA model requires smoothing the data before modeling, but some information may be lost during the smoothing process. In the process of fitting the railway passenger volume, there is no fitting values of 2014Q1–2015Q2 because of the lag order, which has some influence on the comparison of the whole fitting effect.

With the help of exponential smoothing coefficients, the exponential smoothing method can have a particular buffer effect on the characteristics of periodic fluctuations. It reduces the prediction deviation caused by data fluctuations and make the prediction results more accurate. However, the exponential smoothing method is more dependent on recent data than the seasonal cycle GM (1,1) model. Besides, for data with cyclical fluctuation characteristics instead of only linear trends, the degree of dependence on recent data will affect the forecasting accuracy.

By analyzing the fitting results of the exponential smoothing method and the prediction results, it can be seen that the fitting error fluctuates normally. The initial fitting data error is large, while the prediction error gradually increases for the national quarterly passenger volume data in 2019, which has seen a significant increase and seasonal fluctuations. Although the exponential smoothing method can better deal with the data with a growth trend, it has some shortcomings in processing data with periodic change. For the quarterly data of national railway passenger volume in 2019, we use a specific proportion combination of recent data to forecast the next period of data and predicting the latest data. We find that the prediction results will gradually deviate from the actual value. As a result, the error will become larger and larger. Therefore, the increasing prediction error of the quarterly data of national passenger traffic in 2019 by the exponential smoothing method can be explained.

The seasonal cycle GM (1,1) model proposed by this paper takes advantage of the GM (1,1) model in forecasting data with trend change characteristics. Thus, we take GM (1,1) model as the trend equation of the forecasting model. At the same time, it combines with the features of seasonal index, and it can process the data with seasonal fluctuation characteristics. Thus, we construct a seasonal cycle GM (1,1) combination model combining with the characteristics of seasonal index, handling data with seasonal fluctuation characteristics and giving full play to the advantages of the models in the combination model. From the fitting results, it can be concluded that the seasonal cycle GM (1,1) model has an excellent fitting effect on the quarterly data of national passenger traffic from 2014 to 2018, and the fitting error is small and stable, which is suitable for further prediction. According to the forecast results, the seasonal cycle GM (1,1) model has an excellent forecast effect on national passenger traffic quarterly data in 2019. The forecast results are more accurate, and the forecasting error is stable.

Through the comparative analysis of the fitting results and prediction results, it is concluded that the seasonal cycle GM (1,1) model has certain accuracy in fitting and forecasting the national passenger traffic quarterly data and is applicable to the national passenger traffic quarterly data forecasting. Therefore, the model can be used to predict the national passenger traffic quarterly data in the future and obtain scientific and accurate forecasting results.

3.3 Forecast the national passenger traffic quarterly data

Based on the historical data from 2014 to 2019, the traditional GM (1,1) model is used as the trend term of seasonal model to predict national passenger traffic quarterly data in 2020. Simultaneously, the seasonal index of each quarter is calculated, and the predicted value of the trend term from 2020 to 2022 is multiplied by the seasonal index to obtain the final forecasting value of the national passenger volume quarterly data.

Due to the outbreak of COVID-19 at the end of 2019, transportation has been dramatically impacted, and the passenger traffic volume was significantly affected in 2020. To better forecast the passenger volume in 2020, this paper introduces the disturbance factor to consider the situation of the passenger volume under the COVID-19.

Suppose that the disturbance factor is $\lambda, \lambda \in (0, 1]$. While we considering the disturbance factor, the forecasting value of the passenger volume is $\widehat{y}_{\lambda t}$, $\widehat{y}_{\lambda t} = \lambda \cdot \widehat{y}_t$. And $\widehat{y}_t$ is the passenger volume forecasting value without the influence of the disturbance factor. It means that

$$\widehat{y}_{\lambda t} = \begin{cases} \widehat{y}_t, \lambda = 1 \\ \lambda \cdot \widehat{y}_t, \lambda \in (0, 1) \end{cases} \tag{28}$$

According to the official website of the China Statistics Bureau, the updated quarterly data of national railway passenger traffic in 2020 are compared with the forecasting results of the quarterly cycle model, as shown in Table 7.

Considering the development of the epidemic and the change of its impact on passenger volume, we will have a simple deal with the forecast volume in 2020. We will take the ratio of the real value to the predicted value as the disturbance factor, and its characteristics were analyzed to predict the disturbance factor in the fourth quarter of 2020.

Combing with the development of the epidemic situation, the impact of the epidemic situation on passenger volume and the recovery of passenger volume, the disturbance factor was analyzed by exponential modeling, and MATLAB was used for modeling calculation.

Suppose the exponential model is

$$f(k) = \alpha k^{\beta} + \delta, \quad k = 1, \ldots, n. \tag{29}$$

Through simple modeling, we can know that $\alpha = 0.079, \beta = 3.004, \delta = 0.3919$. That means $f(k) = 0.0079k^{3.004} + 0.3919, \quad k = 1, \ldots, n$. Through this model, the disturbance factor of the national passenger volume in the fourth quarter of 2020 is predicted, it is $f(k) = 0.9$. Therefore, we can know the disturbance factor of passenger traffic in the fourth quarter of 2020 is $\lambda = 0.9$, and the forecast value of railway passenger value in the fourth quarter of 2020 (considering the disturbance factor) is $\widehat{y}_{\lambda t} = \lambda \cdot \widehat{y}_t = 0.9 \times 94567.17 = 85110.46$.

With the gradual recovery of the economy, the epidemic's impact will gradually weaken. The estimated impact of the epidemic on railway passenger volume in 2021 and 2022 can be ignored. Therefore, we suppose that the data in 2021 and 2022 will not be affected by disturbance factors. Supposing that $\lambda = 1$, the prediction result is shown in Table 8.

It can be seen from Figure 9 that the quarterly data of national passenger volume from 2014 to 2019 show prominent cyclical fluctuation characteristics and growth trend. Due to the impact of the epidemic, railway operation has significantly been impacted (many cities have taken measures to close all railway passenger stations), and the quarterly data of national passenger volume in 2020 have dropped significantly, reaching the lowest level since 2014–2020. Thanks to the joint efforts of the people and the government, the epidemic situation in

Table 7.
Comparison of the latest data and forecasting data in 2020

Time	Actual value	Prediction value of seasonal cycle GM(1,1) model	Actual value/Predicted value
2020Q1	38,344	93708.45	0.409184
2020Q2	43,424	99059.69	0.438362
2020Q3	69,138	111912.1	0.617788

Table 8.
National passenger volume forecast results during 2020–2022

Time	Trend equation prediction value	Seasonal index	Disturbance factor	Final value
2020Q1	96595.55	0.970	0.3998	37464.64
2020Q2	98718.94	1.003	0.4571	45280.18
2020Q3	100889.00	1.109	0.6158	68915.47
2020Q4	103106.76	0.917	0.9	85110.46
2021Q1	105373.28	0.970	1	102223.81
2021Q2	107689.62	1.003	1	108061.33
2021Q3	110056.87	1.109	1	122081.65
2021Q4	112476.16	0.917	1	103160.58
2022Q1	114948.64	0.970	1	111512.98
2022Q2	117475.47	1.003	1	117880.96
2022Q3	120057.84	1.109	1	133175.32
2022Q4	122696.97	0.917	1	112534.87

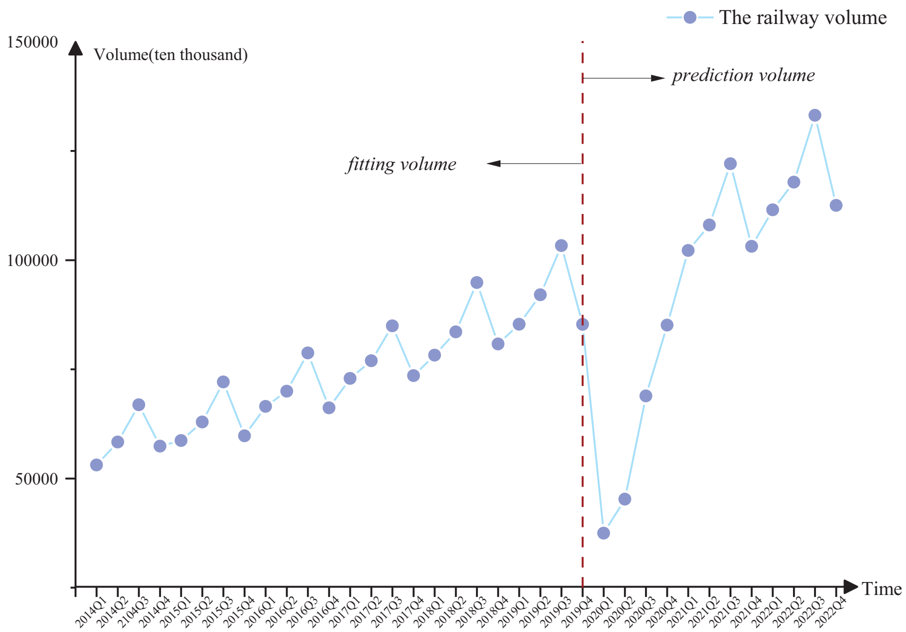

Figure 9.
The quarterly data graph of national railway passenger volume during 2014–2022

China gradually eased in the second and third quarter. Passenger stations are gradually returning to operation, and passenger volume is gradually returning to normal. It is expected that passenger volume will return to normal between 2021 and 2022. Therefore, railway passenger volume in 2021–2022 has the characteristics of trend fluctuation and trend growth, which also conforms to the overall characteristics.

4. Conclusion

In this paper, the quarterly data of national passenger volume from 2014 to 2019 are analyzed, and the seasonal cycle GM (1,1) model is constructed. Each model's applicability was compared with the quarterly data of national passenger volume from 2014 to 2019. On this basis, we forecast the national passenger volume quarterly data in 2020 further. The results suggest

(1) The data of national railway passenger volume have both trend fluctuation and periodic fluctuation. Through the chart analysis, it can be seen that the national data of the railway passenger volume's fluctuation cycle is 4. The data of the national volume showed the seasonal fluctuation characteristic, and in the third quarter of each year, passenger volume data reach the cycle peak and the peak value increases year by year. It means that the national passenger volume has a growing trend while it changes periodically with seasons' change. Therefore, it is necessary to consider whether the model can be applied to the data with periodic fluctuation and trend growth characteristics. The results show that the seasonal cycle GM (1,1) model can deal with the quarterly data of national passenger traffic with the characteristics of periodic fluctuation and trend growth.

(2) The quarterly data of national railway passenger volume from 2014 to 2018 are simulated by the seasonal cycle GM (1,1) model, SGM (1,1), SARIMA model, traditional GM (1,1) model, exponential smoothing method and the moving average method. We did prediction about 2019 as testing stage. According to the results, it can be known that the seasonal cycle GM (1,1) model has an excellent fitting and forecasting effect, the fitting MAPE is less than 2% and the fitting error is stable. The results show that the seasonal cycle GM (1,1) model is suitable for the quarterly railway passenger volume. Therefore, we can use the seasonal cycle GM (1,1) model to forecast the 2020 national passenger volume quarterly data and get more scientific and reliable data.

The SGM (1,1) and SARIMA model have advantage in forecasting seasonal time series, but they have less prediction accuracy of the railway passenger volume than the seasonal cycle GM (1,1) model. Combining trend equation and real data to get the seasonal index can make full use of real data information and get more reasonable seasonal index, so the forecasting result is more accurate.

(3) The seasonal cycle GM (1,1) model can effectively solve the time series data with periodic fluctuation and trend growth, which is proved to a certain extent. This paper analyzes and predicts passenger volume data (with periodic fluctuation and trend growth) by using the seasonal cycle GM (1,1) model. It can provide a practical reference for railway departments to make an operation plan and equipment management plan.

References

Banerjee, N., Morton, A. and Akartunal, K. (2020), "Passenger demand forecasting in scheduled transportation", *European Journal of Operational Research*, Vol. 286 No. 3, pp. 797-810, 7.

Carmona-Bentez, R.B. and Nieto, M.R. (2020), "SARIMA damp trend grey forecasting model for airline industry", *Journal of Air Transport Management*, Vol. 82.

Cheng, M., Shi, G. and Han, Y. (2019), "An extended grey model GM(1, 1, exp(bk)) and its application in Chinese civil air passenger volume prediction", *Journal of Systems Science and Information*, Vol. 7 No. 5, pp. 486-496.

Deng, J.L. (1982), "Control problems of grey systems", *Systems and Control Letters*, Vol. 1 No. 5, pp. 288-294.

Deng, J.L. (1989), "Introduction to grey system theory", *Journal of Grey System*, Vol. 1 No. 1, pp. 1-24.

Deng, W., Li, W. and Yang, X.H. (2011), "A novel hybrid optimization algorithm of computational intelligence techniques for highway passenger volume prediction", *Expert Systems with Applications*, Vol. 38, pp. 4198-4205.

Fan, D.C., Wang, S.H. and Zhang, W. (2012), "Study on application of seasonal cycle model in China's power load forecasting", *Power System Technology*, Vol. 36, pp. 106-110.

Hao, S., Lee, D.H. and Zhao, D. (2019), "Sequence to sequence learning with attention mechanism for short-term passenger flow prediction in large-scale metro system", *Transportation Research Part C: Emerging Technologies*, Vol. 107, pp. 287-300.

Holt, C.C. (2004), "Forecasting seasonals and trends by exponentially weighted moving averages", *International Journal of Forecasting*, Vol. 20, pp. 5-10.

Jiang, P., Li, R., Liu, N. and Gao, Y. (2020), "A novel composite electricity demand forecasting framework by data processing and optimized support vector machine", *Applied Energy*, Vol. 260, 114243.

Li, M.W., Kang, H.G. and Zhou, P.F. (2012), "Urban passenger prediction based on hybrid algorithm of new chaos accelerating genetic algorithm and PPPR model", *System Engineering Theory and Practice*, Vol. 32, pp. 903-910.

Li, X.H., Tan, Q.M. and Zhao, A.W. (2014), "China's passenger traffic based on factor analysis and Logistic model", *Journal of Systems Management*, Vol. 23 No. 3, pp. 444-450.

Li, H.Y., Wang, Y.T., Xu, X.Y., Qin, L.Q. and Zhang, H.Y. (2019), "Short-term passenger flow prediction under passenger flow control using a dynamic radial basis function network", *Applied Soft Computing*, Vol. 83, 105620.

Li, N., Wang, J., Wu, L. and Bentley, Y. (2021), "Predicting monthly natural gas production in China using a novel grey seasonal model with particle swarm optimization", *Energy*, Vol. 215 No. 4, 119118.

Liang, X.Z., Qiao, H., Wang, S.Y. and Zhang, X. (2017), "An integrate forecasting model for air passenger traffic in China based on singular spectrum analysis", *System Engineering Theory and Practice*, Vol. 37, pp. 1479-1488.

Liu, G., Yin, Z., Jia, Y. and Xie, Y. (2017), "Passenger flow estimation based on convolutional neural network in public transportation system", *Knowledge-Based Systems*, Vol. 123, pp. 102-115.

Pan, H., Li, J., Shen, Q. and Shi, C. (2017), "What determines rail transit passenger volume? Implications for transit oriented development planning", *Transportation Research Part D: Transport and Environment*, Vol. 57, pp. 52-63.

Peng, Y. and Xiang, W. (2020), "Short-term traffic volume prediction using GA-BP based on wavelet denoising and phase space reconstruction", *Physica A: Statistical Mechanics and Its Applications*, Vol. 549, 123913.

Qian, W.Y. and Wang, J. (2020), "An improved seasonal GM(1,1) model based on the HP filter for forecasting wind power generation in China", *Energy*, Vol. 209, 118499.

Qin, L., Li, W. and Li, S. (2019), "Effective passenger flow forecasting using STL and ESN based on two improvement strategies", *Neurocomputing*, Vol. 356, pp. 244-256.

Sun, Y., Leng, B. and Guan, W. (2015), "A novel wavelet-SVM short-time passenger flow prediction in Beijing subway system", *Neurocomputing*, Vol. 166, pp. 109-121.

Tsai, T.H., Lee, C.K. and Wei, C.H. (2009), "Neural network based temporal feature models for short-term railway passenger demand forecasting", *Expert Systems with Applications*, Vol. 36, pp. 3728-3736.

Tseng, F.M., Hsiao-Cheng, Y.U. and Tzeng, G.H. (2001), "Applied hybrid grey model to forecast seasonal time series", *Technological Forecasting and Social Change*, Vol. 67, pp. 291-302.

Wang, Y.H. and Lu, J. (2020), "Improvement and application of GM(1,1) model based on multivariable dynamic optimization", *Journal of Systems Engineering and Electronics*, Vol. 31 No. 3, pp. 593-601.

Wang, J., Zhu, W., Zhang, W. and Sun, D. (2009), “A trend fixed on firstly and seasonal adjustment model combined with the ε-SVR for short-term forecasting of electricity demand”, *Energy Policy*, Vol. 37, pp. 4901-4909.

Wang, Y.H., Dang, Y.G., Li, Y.Q. and Liu, S.F. (2010), “An approach to increase prediction precision of GM(1,1) model based on optimization of the initial condition”, *Expert Systems with Applications*, Vol. 37 No. 8, pp. 5640-5644.

Wang, J., Lu, H., Dong, Y. and Chi, D. (2012a), “The model of chaotic sequences based on adaptive particle swarm optimization arithmetic combined with seasonal term”, *Applied Mathematical Modelling*, Vol. 36, pp. 1184-1196.

Wang, J., Ma, X., Wu, J. and Dong, Y. (2012b), “Optimization models based on GM (1,1) and seasonal fluctuation for electricity demand forecasting”, *International Journal of Electrical Power and Energy Systems*, Vol. 43, pp. 109-117.

Wang, Y., Chen, X., Han, Y. and Guo, S. (2013), “Forecast of passenger and freight yraffic volume based on elasticity coefficient method and grey model”, *Procedia - Social and Behavioral Sciences*, Vol. 96, pp. 136-147.

Wang, J., Deng, W. and Guo, Y. (2014), “New Bayesian combination method for short-term traffic flow forecasting”, *Transportation Research Part C: Emerging Technologies*, Vol. 43, pp. 79-94.

Wang, Y.H., Qu, K. and Wang, Z. (2016), “A kind of nonlinear strengthening operators for predicting the putput value of China’s marine electric power industry”, *Journal of Grey System*, Vol. 28 No. 2, pp. 35-52.

Wang, Y.H., Xie, T.Y. and Yang, S.Li. (2017a), “Carbon emission and its decoupling research of transportation in Jiangsu Province”, *Journal of Cleaner Production*, Vol. 142 No. 2, pp. 907-914.

Wang, Z.X., Li, Q. and Pei, L.L. (2017b), “Grey forecasting method of quarterly hydropower production in China based on a data grouping approach”, *Applied Mathematical Modelling*, Vol. 51, pp. 302-316.

Wang, Q., Li, S., Li, R. and Ma, M. (2018a), “Forecasting US shale gas monthly production using a hybrid ARIMA and metabolic non-linear grey model”, *Energy*, Vol. 160, pp. 378-387.

Wang, Z.X., Li, Q. and Pei, L.L. (2018b), “A seasonal GM(1,1) model for forecasting the electricity consumption of the primary economic sectors”, *Energy*, Vol. 154, pp. 522-534.

Wang, Z.X., He, L.Y. and Zheng, H.H. (2019), “Forecasting the residential solar energy consumption of the United States”, *Energy*, Vol. 178, pp. 610-623.

Wang, Z.X., Wang, Z.W. and Li, Q. (2020), “Forecasting the industrial solar energy consumption using a novel seasonal GM(1,1) model with dynamic seasonal adjustment factors”, *Energy*, Vol. 200, 117460.

Wu, J., Wang, J., Lu, H., Dong, Y. and Lu, X. (2013), “Short term load forecasting technique based on the seasonal exponential adjustment method and the regression model”, *Energy Conversion and Management*, Vol. 70, pp. 1-9.

Xia, M. and Wong, W.K. (2014), “A seasonal discrete grey forecasting model for fashion retailing”, *Knowledge-Based Systems*, Vol. 57, pp. 119-126.

Xiao, X., Yang, J., Mao, S. and Wen, J. (2017), “An improved seasonal rolling grey forecasting model using a cycle truncation accumulated generating operation for traffic flow”, *Applied Mathematical Modelling*, Vol. 51, pp. 386-404.

Xiong, P.P., Yan, W.J., Wang, G.Z. and Pei, L.L. (2019), “Grey extended prediction model based on IRLS and its application on smog pollution”, *Applied Soft Computing*, Vol. 80, pp. 797-809.

Yang, X., Xue, Q., Ding, M., Wu, J. and Gao, Z. (2021), “Short-term prediction of passenger volume for urban rail systems: a deep learning approach based on smart-card data”, *International Journal of Production Economics*, Vol. 231, 107920.

Ye, J., Dang, Y., Ding, S. and Yang, Y. (2019), "A novel energy consumption forecasting model combining an optimized DGM (1, 1) model with interval grey numbers", *Journal of Cleaner Production*, Vol. 229, pp. 256-267.

Zhang, J., Sun, Y.Y., Zhang, X.Y. and Wang, S.Y. (2020), "Time-varying forecast averaging for air passengers in China", *System Engineering Theory and Practice*, Vol. 40, pp. 1509-1519.

Zhou, Z. and Wang, W.Q. (2019), "Modification and comparison of short-term forecasting models for railway passenger volume in China", *Statistics and Decision*, Vol. 35, pp. 66-71.

Zhao, S.Z., Ni, T.H., Wang, Y. and Gao, X.T. (2011), "A new approach to the prediction of passenger flow in a transit system", *Computers and Mathematics with Applications*, Vol. 61, pp. 1968-1974.

Zhou, W., Pan, J., Ding, S. and Wu, X. (2020a), "A novel discrete grey seasonal model and its applications", *Communications in Nonlinear Science and Numerical Simulation*, Vol. 93, 105493.

Zhou, W., Chen, Y., Ding, S., Chen, L. and Li, R. (2020b), "A grey seasonal least square support vector regression model for time series forecasting", *ISA Transactions*, Vol. 11, doi: 10.1016/j.isatra.2020.12.024.

Zhou, W., Wu, X., Ding, S. and Cheng, Y. (2020c), "Predictive analysis of the air quality indicators in the Yangtze River Delta in China: an application of a novel seasonal grey model", *The Science of the Total Environment*, Vol. 748, 141428.

Zhou, J., Ren, G.H., He, H.B., Hou, X.Y. and Deng, W.C. (2020d), "Application of the exponential smoothing model and ARIMA model in prediction of the endemic situation of schistosomiasis in Hunan Province", *Chinese Journal of Schistosomiasis Control*, Vol. 32 No. 3, pp. 236-241.

Zhu, S., Wang, J., Zhao, W. and Wang, J. (2011), "A seasonal hybrid procedure for electricity demand forecasting in China", *Applied Energy*, Vol. 88, pp. 3807-3815.

Corresponding author
Yuhong Wang can be contacted at: yuhongwang@jiangnan.edu.cn