A Combined Prediction Model Composed of the GM (1,1) Model and the BP Neural Network for Major Road Traffic Accidents in China

Qingwen Guo,1 Baohua Guo,1,2 Yugang Wang,3 Shixuan Tian,1 and Yan Chen1,2

1School of Energy Science and Engineering, Henan Polytechnic University, Jiaozuo 454000, China
2Jiaozuo Engineering Research Center of Road Traffic and Transportation, Henan Polytechnic University, Jiaozuo 454000, China
3Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan 430063, China

Correspondence should be addressed to Baohua Guo; guobh@hpu.edu.cn

Received 27 January 2022; Revised 7 March 2022; Accepted 29 March 2022; Published 15 April 2022

1.Introduction

China’s economy is growing at a fast pace, and as a result of the use of scientific knowledge in the transportation field, China’s road traffic industry has made great achievements. The appearance of automobiles has made a significant contribution to human society and economic prosperity; however, the caused road traffic accidents have also harmed people’s production, lives, and property. Besides, a traffic accident is one of the most important factors that affect road operation efficiency. China has entered the peak period of road traffic accidents, especially the major traffic accidents that caused heavy casualties and huge property losses [1]. In this paper, the major traffic accident refers to a road traffic accident with 10 or more fatalities at one time.

The fatality rate of a major traffic accident is about 50%, which is very high compared with the national road traffic accident fatality rate of about 20%. The consequences of a major traffic accident are even worse, seriously affecting the social peace and harmony, even the social stability, and as a result, it’s critical to look at the big development laws traffic accidents correctly using some prediction methods. The objective of major traffic accident expectations is to assess the possibility of major traffic accidents in the future and analyze the level of danger and the pattern of significant traffic accidents so that appropriate steps can be made to prevent them as soon as feasible. Additionally, it has an important practical significance for road traffic safety evaluation, road planning, and decision-making [2]. Few special research on major road traffic accidents have been
carried out in developed countries and regions because the major road traffic accidents with 10 or more fatalities at one time are rare due to differences in the road traffic environment and other aspects. However, major road traffic accidents, affected by passengers, vehicles, roads, and the environment, cannot be eradicated thoroughly. Because of China’s current economic development, it is critical to reduce the severity of serious traffic accidents by analyzing their frequency.

Over the past several decades, scholars at home and overseas have done a lot of research work on road traffic accident prediction and summarized several prediction methods including regression analysis method, empirical model method, gray model prediction method, time series prediction method, neural network prediction method, and so on [3]. The linear regression prediction model is widely used because of its simplicity and efficiency, and according to the basic principle of multiple linear regression analysis, Qiu et al. [4] established a mathematical model for multiple linear regression analysis of road traffic accidents, which provided a scientific basis for reasonable and effective analysis and prediction of road traffic accidents. When more information is collected, the time series prediction model can use several event attribute variables for regression prediction; Parvareh et al. [5] used the timing chain inquiry method to describe and forecast the frequency of damages caused by traffic accidents in Kurdistan province. These findings require adjacent watching of accidents to regulate and decrease the injury rate during high-risk situations. Chen et al. [6] deeply analyzed the complex coupling relationships between accident factors causing single-vehicle and multi-vehicle accidents in the streets; the results suggest that the Bayesian routing algorithm can describe the genuine link between various components and may be used as the targeted system of significant road accidents in China when utilizing the Bayesian collision intensity model. Chuan et al. [7] established a GM (1, 1) gray prediction model for military vehicle traffic accidents based on the original data of military vehicle traffic accidents from 2005 to 2009 and made short-term predictions determined by the number of collisions and casualties. Qiu et al. [8] established a GM (1, 1) gray prediction model for road traffic accidents using gray theory methods and verified the applicability of the prediction model through several examples. Wang et al. [9] established a gray GM (1, 1) model of traffic accidents created on the amount of road traffic accidents and mortalities in China from 2009 to 2013. Meanwhile, neural networks have been widely used in prediction problems. For example, Qian et al. [10] used an Elman regression neural network model to calculate the number of people killed in automobile accidents in China and expect the number of short-range road traffic mortalities. The findings suggest that the Elman method is used to forecast and anticipate short-term patterns as well as assess the impact of highway safety planning based on historical road traffic fatality data. Singh et al. [11] used a deep neural network (DNN) model to predict road traffic accidents, and the performance of DNN in road accident prediction has been improved. Tang and Ying [12] established a traffic accident prediction model by using BP neural network and verified its accuracy with MATLAB software. Zhu et al. [13] concluded that the BP neural network model has a good effect on the prediction of traffic accidents. Deng et al. [14] established a BP neural network-based accident forecasting model and verified the applicability of the model to specific expressway sections.

All the above models can achieve the prediction of future traffic accidents, but each prediction model has its own shortcomings and defects. Especially, the gray prediction model is mainly used for short-term predicting problems with strong tendency and little fluctuation; relatively, satisfactory prediction results can be obtained even based on less data, but the predicting ability is weak when there are regular fluctuations or sudden changes in the observed sequence. However, though the modeling process of the BP neural network model is relatively complicated and requires more training samples, it has a good processing ability for data with regular fluctuations or sudden changes. Therefore, their combined model may gain benefits and avoid drawbacks of the above two prediction models, enhance predicting capabilities, and improve predicting accuracy. By applying this combined model, Gao et al. [15] predicted the elastic strain helps to set in the conventional drying process of wood accurately. Ming et al. [16] predicted the mechanical properties of mulberry silk which compared the two single predictive models above and found that the gray neural network model outperformed the black neural network model single gray GM (1, 1) model. Liu et al. [17] predicted the market demand after traffic interruption with an improved combination model and proved that the enhanced combination model is superior to the standard gray model. Therefore, their combined model composed of the GM(1,1) model and the BP neural network model may gain benefits and avoid drawbacks of the above two prediction models, enhance predicting capabilities, and improve predicting accuracy. In this research, the GMBP model will be used to forecast the number of serious traffic accidents, the number of mortalities, and the number of grievances in major traffic accidents, and the prediction accuracy of these two prediction models will be analyzed. The importance of improving road traffic safety management cannot be overstated level and reduces the number of fatalities and property losses by analyzing the rising movement of the future major traffic accidents.

A paper’s organizational structure is as follows: the first section of the paper provides a thorough overview. The approaches utilized in the report to conduct solid research are defined in Section 2. Section 3: read the case study and collect data in order to come up with conclusions. The paper comes to a close with Section 4.
2. Methodology

2.1. GM(1,1) Principle. Deng and Hong et al. proposed the gray system theory in 1982, which is a method to study small-sample problems or uncertain problems with weak information [18, 19]. In cybernetics, the depth of color is frequently used to indicate the degree of clarity of information. For example, “black” denotes that all information is unknown, “white” denotes that the information is totally clear, and “gray” denotes that only some of the information is clear. As a result, a system with entirely clear information in a network with totally unknown knowledge is alluded to as a black system, a system with entirely unknown knowledge is referred to as a gray system [20, 21]. The gray GM (1, 1) model, a first-order univariate gray prediction model, is the main model of the gray system in gray system theory. Based on a little amount of data, it can be used to anticipate the changing law of the data sequence.

Assume that

\[ X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)\}. \]  

(1)

Is there an existent nonnegative information pattern or first total combined creation operator on \( X^{(0)} \) has the following sequence:

\[ X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(n)\}, \]  

(2)

where

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

(3)

We may derive the sequences of the created mean value of subsequent neighbor in the following from the sequence \( X^{(1)} \) obtained by applying the first-order exponentially increasing generation algorithm to sequence \( X^{(0)} \)

\[ Z^{(1)} = \{z^{(1)}(2), z^{(1)}(3), \ldots, z^{(1)}(n)\}, \]  

(4)

where

\[ z^{(1)}(k) = \frac{1}{2}(x^{(1)}(k) + x^{(1)}(k-1)), \quad k = 2, 3, \ldots, n. \]  

(5)

The following formula is a gray differential equation, also known as the GM (1, 1) model, for a nonnegative pattern of raw data \( X^{(0)} \), and \( X^{(1)} \) is a freshly created arrangement with the implementation of the first-order cumulated generation operator on \( X^{(0)} \). \( Z^{(1)} \) is a new pattern when the produced mean value of sequential neighbors’ operator is used on \( X^{(1)} \)

\[ x^{(0)}(k) + az^{(1)}(k) = b, \]  

(6)

And the equation

\[ \frac{dx^{(1)}(k)}{dt} + ax^{(1)}(k) = b, \]  

(7)

is the whitened equation of the GM (1, 1) model. If we let

\[ \tilde{a} = [a, b]^{T}, \]  

(8)

\[ Y = \left[ \begin{array}{c} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{array} \right], \]  

\[ B = \left[ \begin{array}{c} -z^{(1)}(2) \\ -z^{(1)}(3) \\ \vdots \\ -z^{(1)}(n) \end{array} \right], \]  

the parameter predictions for the gray GM (1, 1) model can be obtained using the least square approximation method as follows:

\[ \tilde{a} = [a, b]^{T} = (B^{T}B)^{-1}B^{T}Y, \]  

(9)

The pearly white equation’s time function produces

\[ x^{(1)}(t) = \left(x^{(1)}(1) - \frac{b}{a}\right)e^{-a(1-t)} + \frac{b}{a}, \]  

(10)

The temporal response equation for GM (1, 1) is as follows:

\[ \tilde{x}^{(1)}(k + 1) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{-ak} + \frac{b}{a}, \quad k = 1, 2, 3, \ldots, n. \]  

(11)

The restored values of raw data are given below,

\[ \tilde{x}^{(0)}(k + 1) = \tilde{x}^{(1)}(k + 1) - \tilde{x}^{(1)}(k) \]

\[ = \left(1 - e^{a}\right)\left(x^{(0)}(1) - \frac{b}{a}\right)e^{-ak}, \quad k = 1, 2, 3, \ldots, n. \]  

(12)

2.2. BP Neural Network. Based on the imperfect back-propagation method, a layered feed-forward neural network is known as a BP neural network training [22]. It uses a linear arrangement of nonlinear kernel function to perform nonlinear translation from input to output space, and approximate arbitrary nonlinear mapping through training, and it has a better processing ability for nonlinear time series predicting problems. An input data, one may be more hidden units, and output units make up a BP neural network, with the input element’s neuron representing information compiled from external sources and serving as an interpreter or contributor and the obscured layer’s neurons transferring substantiation from the input layer to the output layer [23]. Each layer has numerous neurons, and the neurons between the levels are connected to each other, while the neurons within the same layer are not. Figure 1 depicts the fundamental network structure diagram.

The simple unit of the BP neural network is a neuron. Each neuron takes the output of the previous layer of
neurons as input and output after passing through the activation function. The hidden layer neurons generally use nonlinear functions such as Tanh (tangent hyperbolic) function as the activation function, so as to realize the nonlinear mapping from input to output. The connection between each neuron corresponds to a weight. The input of the stimulation function is the sum value of each input value of the neuron multiplied by the corresponding weight and the bias value of the neuron itself, and then, the corresponding output can be got. The aim of BP neural network training is to use the error back-propagation algorithm to modify each weight and each bias value to make the network converge. The specific implementation process of the back-propagation algorithm is as follows [24):

(1) Assign initial values for each weight and bias of the network randomly and determine the activation function;

(2) Input the training sample and calculate the output of each layer of the network according to the following formula:

\[ Z = \sigma (WX + B), \]

where \( Z \) is the actual output vector of each layer of the neural network, \( W \) is the weight matrix, \( X \) is the input vector, \( B \) is the bias vector, and \( \sigma \) is the activation function.

(3) According to the following error function formula, find the performance error of the system:

\[ E = \frac{1}{2} \sum_i (Y_i - Z_i) \]

where \( Y_i \) is the expected output value of the \( i \)th neuron in the output layer and \( Z_i \) is the actual output value of the \( i \)th neuron in the output layer.

(4) Find the error of each layer:

\[ e^l = \nabla E \odot \sigma ' (W^l X^l + B^l), \]

where \( \odot \) is the Hadamard product operator of the matrix and \( W^l, X^l, B^l \) are the weight matrix, input vector, and bias vector of the last layer, respectively.

For the error of other layers, \( e^l \) have:

\[ e^l = (W^{l+1})^T e^{l+1} \odot \sigma ' (W^l X^l + B^l), \]  

where \( l \) is the current layer and \( l + 1 \) is the next layer.

(5) Adjust weights and bias:

\[ W_{ij}^{l+1} = W_{ij}^{l} - \alpha e_{ij}^{l+1}, \]

where \( \alpha \) is the learning rate, that is, the step length; generally, a constant between 0.01 and 0.8 is selected.

(6) Repeat the process from step (2) to step (5) until the system performance error drops to the preset value or the number of training times reaches the preset value, and the trained BP neural network can be used for data prediction after the training ends.

2.3. Combined Prediction Model. In the late 1960s, Bates and Granger proposed the combined predicting theory firstly [25]. Different prediction models were combined to establish an appropriate combined prediction model for achieving the best predicted results according to the characteristics of the prediction object. In the early period, the combination prediction method has been applied to traffic flow and road accidents in the traffic research field. Gu et al. [26] proposed an enhanced Bayesian grouping model founded on deep learning for traffic flow prediction using the real traffic flow data collected by microwave sensors of the Beijing expressway and found that IBCM-DL was superior to other advanced methods in accuracy and stability. Khouban et al. [27] proposed an expert system for road traffic noise modeling based on an artificial neural network and found that the hybrid model had more advantages in improving the performance prediction and reducing the measurement cost. Raza and Zhong [28] used a genetic algorithm, a neural network, and locally weighted regression to get the best prediction under numerous input and traffic conditions, concluding that the combined model’s error was lower than the previous model. Zhang and Zhang [29] suggested a quick macroscopic traffic flow approach based on the regularity, stationarity, and abnormality of data series in order to assess and solve the LSTM-XG Boost model and found that not only may a combined prediction model enhance forecast accuracy but it can also enhance the model's viability, real-time capability, and scalability. Zhong et al. [30] combined the Verhulst model and multiple linear regression model into a combined prediction model with weight coefficient, which can not only characterize the mortality toll data but also measure the sort of authority of various influencing factors on fatality toll, with high accuracy and strong practicability. A combination prediction model composed of the GM (1, 1) model and BP neural network, which is simplified as the GMBP model, will be proposed in this paper for major traffic accident prediction to improve the prediction accuracy and increase the prediction reliability. The prediction steps of the GMBP model are as follows:
2.4. Concrete Realization of Combined Model

(1) Obtaining the predicted data of the gray GM (1, 1) model by inputting the original data into the gray GM (1, 1) model.

(2) Under the condition of using the gray GM (1, 1) model’s information amassed as input data and the actual data as output data, the BP neural network was trained.

(3) Obtaining the predicted data of the GMBP model by inputting the predicted data of the gray GM (1, 1) model into the trained BP neural network model.

3. Case Study

3.1. Original Data. The major traffic accidents data with 10 or more fatalities at one time from 2008 to 2020 in China as shown in Figure 2, collected from the website of the Ministry of Transport of the People’s Republic of China, are used to verify the prediction effect of the GMBP model, and then the predicted results of the following two prediction methods including the single gray GM (1, 1) model and a GMBP model are compared and analyzed.

It can be seen from Figure 2(a) that the number of China’s major road traffic accidents with 10 or more fatalities at one time, the number of the fatalities, and the injuries in China’s major road traffic accidents are on the decline from 2008 to 2020. It can be seen from Figure 2(b) that March, August, and October are three peaks in the lot of large road traffic accidents in China as well as the fatality rate and damages. The number of accidents and fatalities, in particular, and injuries in March was 28, 415, and 364, respectively, and the number of accidents, fatalities, and injuries in August was 26, 382, and 268, respectively. February and March are usually the Spring Festival holidays and the winter holiday for the students in China, so the number of accidents is larger as the road traffic volume is much larger than in the other months. August is the summer vacation for students, and the number of road traffic accidents is also significantly larger.

Major traffic accidents and economic and demographic data for each Chinese province from 2008 to 2020 are provided in Table 1 to explore the links among the number of road traffic accidents and GDP or population. From Table 1, we can see that Guizhou, Hebei, Hunan, and Yunnan provinces have more major traffic accidents than others, and the number of accidents ranges from 12 to 17.

Figure 3 explains the traffic accidents in China with detail of accidents, fatalities, and injuries included also. The relationships between the number of major traffic accidents, fatalities, and injuries and the population, GDP, and per capita GDP are shown in Figure 3. The frequency of major traffic accidents, fatalities, and injuries is favorably connected with each province’s population and GDP but adversely correlated with each province’s per capita GDP. To put it another way, the number of major traffic accidents and fatalities is related to the population and level of economic development of each province, particularly the per capita GDP. Some academics have also spoken out about this phenomenon. Smeed [31] thought that a huge number of traffic accidents happened in the initial stage of national motorization, and when the level of economic development and the degree of motorization reach a certain level, the level of traffic safety will be improved. Ren et al. [32] concluded that the number of traffic accidents first increases and then decreases with the increasing per capita GDP, and the association between the number of traffic accidents and the level of economic development presents an inverted U-shaped curve.

3.2. Predicted Results and Analysis. After establishing the GMBP combined prediction model by composing the GM (1, 1) model and BP neural network model, three sets of predicted data of major traffic accidents, victims, and injuries were, respectively, obtained as shown in Figure 4.

It can be seen from Figure 4 that the predicted data of the gray GM (1, 1) model and the GMBP model show a downward trend, on the whole; however, for the number of major traffic accidents, fatalities, and injuries, the predicted data of the GMBP model are obviously more consistent with the original data of major traffic accidents.

From the development trend of the above data, with the increase in China’s economic development and traffic safety investment, China’s major traffic accidents will be significantly reduced. It should be noted that relatively severe novel coronavirus pneumonia outbreak occurred in China in the first quarter of 2020, which has a certain impact on China’s economy, and the number of major traffic accidents, fatalities, and injuries may be relatively low compared with the normal years, so it will have a certain impact on the accuracy of data prediction in the future years. In 2020, the predicted number of major traffic incidents, fatalities, and injuries in the GM (1, 1) and GMBP model is not less than its original number, and in particular, the predicted number of fatalities and injuries is bigger obviously.

The quadratic formula of the ratio of the squared sum of the variance between the expected values and the original value and the data set $n$ is the RMSE (root mean square error), which can be used to evaluate the prediction accuracy of the GM (1, 1) and GMBP model. The calculation formula is as follows:
Figure 2: Original data of major traffic accidents from 2008 to 2020 in China. (a) Major traffic accident data by year and (b) major traffic accident data by month.

<table>
<thead>
<tr>
<th>Province</th>
<th>Accidents</th>
<th>Fatalities</th>
<th>Injuries</th>
<th>GDP/10^{12} yuan</th>
<th>Population/10^7</th>
<th>Per capita GDP/10^4 yuan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shanghai</td>
<td>1</td>
<td>12</td>
<td>12</td>
<td>3.82</td>
<td>2.42</td>
<td>15.78</td>
</tr>
<tr>
<td>Inner Mongolia</td>
<td>2</td>
<td>35</td>
<td>10</td>
<td>1.72</td>
<td>2.53</td>
<td>6.81</td>
</tr>
<tr>
<td>Qinghai</td>
<td>2</td>
<td>21</td>
<td>82</td>
<td>0.30</td>
<td>0.60</td>
<td>4.96</td>
</tr>
<tr>
<td>Liaoning</td>
<td>3</td>
<td>57</td>
<td>37</td>
<td>2.49</td>
<td>4.37</td>
<td>5.70</td>
</tr>
<tr>
<td>Shaanxi</td>
<td>3</td>
<td>43</td>
<td>39</td>
<td>2.58</td>
<td>3.84</td>
<td>6.72</td>
</tr>
<tr>
<td>Tianjin</td>
<td>3</td>
<td>46</td>
<td>26</td>
<td>1.41</td>
<td>1.56</td>
<td>9.06</td>
</tr>
<tr>
<td>Zhejiang</td>
<td>3</td>
<td>33</td>
<td>47</td>
<td>6.24</td>
<td>5.66</td>
<td>11.02</td>
</tr>
<tr>
<td>Chongqing</td>
<td>3</td>
<td>41</td>
<td>42</td>
<td>2.36</td>
<td>3.05</td>
<td>7.74</td>
</tr>
<tr>
<td>Guangxi</td>
<td>4</td>
<td>55</td>
<td>42</td>
<td>2.12</td>
<td>4.89</td>
<td>4.35</td>
</tr>
<tr>
<td>Ningxia</td>
<td>4</td>
<td>51</td>
<td>67</td>
<td>0.37</td>
<td>0.68</td>
<td>5.50</td>
</tr>
<tr>
<td>Fujian</td>
<td>5</td>
<td>91</td>
<td>94</td>
<td>4.24</td>
<td>3.91</td>
<td>10.84</td>
</tr>
<tr>
<td>Heilongjiang</td>
<td>5</td>
<td>65</td>
<td>109</td>
<td>1.36</td>
<td>3.79</td>
<td>3.59</td>
</tr>
<tr>
<td>Jilin</td>
<td>5</td>
<td>79</td>
<td>84</td>
<td>1.17</td>
<td>2.72</td>
<td>4.32</td>
</tr>
<tr>
<td>Hubei</td>
<td>6</td>
<td>76</td>
<td>67</td>
<td>4.58</td>
<td>5.90</td>
<td>7.77</td>
</tr>
<tr>
<td>Jiangsu</td>
<td>6</td>
<td>70</td>
<td>126</td>
<td>9.96</td>
<td>8.03</td>
<td>12.41</td>
</tr>
<tr>
<td>Anhui</td>
<td>7</td>
<td>85</td>
<td>132</td>
<td>3.71</td>
<td>6.25</td>
<td>5.93</td>
</tr>
<tr>
<td>Gansu</td>
<td>7</td>
<td>116</td>
<td>154</td>
<td>8.72</td>
<td>2.63</td>
<td>3.32</td>
</tr>
<tr>
<td>Tibet</td>
<td>7</td>
<td>124</td>
<td>86</td>
<td>0.17</td>
<td>0.34</td>
<td>5.04</td>
</tr>
<tr>
<td>Guangdong</td>
<td>8</td>
<td>116</td>
<td>142</td>
<td>10.77</td>
<td>11.17</td>
<td>9.64</td>
</tr>
<tr>
<td>Shandong</td>
<td>8</td>
<td>125</td>
<td>80</td>
<td>7.11</td>
<td>10.01</td>
<td>7.10</td>
</tr>
<tr>
<td>Xinjiang</td>
<td>8</td>
<td>112</td>
<td>199</td>
<td>1.36</td>
<td>2.44</td>
<td>0.56</td>
</tr>
<tr>
<td>Shanxi</td>
<td>9</td>
<td>138</td>
<td>146</td>
<td>1.70</td>
<td>3.70</td>
<td>4.60</td>
</tr>
<tr>
<td>Sichuan</td>
<td>9</td>
<td>133</td>
<td>43</td>
<td>5.67</td>
<td>8.30</td>
<td>6.82</td>
</tr>
<tr>
<td>Jiangxi</td>
<td>10</td>
<td>137</td>
<td>116</td>
<td>2.48</td>
<td>4.62</td>
<td>5.36</td>
</tr>
<tr>
<td>Henan</td>
<td>12</td>
<td>236</td>
<td>199</td>
<td>5.43</td>
<td>9.56</td>
<td>5.68</td>
</tr>
<tr>
<td>Guizhou</td>
<td>13</td>
<td>217</td>
<td>197</td>
<td>1.68</td>
<td>3.58</td>
<td>4.68</td>
</tr>
<tr>
<td>Hebei</td>
<td>14</td>
<td>240</td>
<td>36</td>
<td>3.51</td>
<td>7.52</td>
<td>4.67</td>
</tr>
<tr>
<td>Hunan</td>
<td>14</td>
<td>234</td>
<td>220</td>
<td>3.98</td>
<td>6.86</td>
<td>5.79</td>
</tr>
<tr>
<td>Yunnan</td>
<td>17</td>
<td>239</td>
<td>139</td>
<td>2.32</td>
<td>4.80</td>
<td>4.84</td>
</tr>
</tbody>
</table>
Figure 3: Relationship between major traffic accidents and economic or demographic data. (a) The number of accidents vs. population, (b) the number of fatalities and injuries vs. population, (c) the number of accidents vs. GDP, (d) the number of fatalities and injuries vs. GDP, (e) the number of accidents vs. per capita GDP, and (f) the number of fatalities and injuries vs. per capita GDP.
The RMSE (root mean square error) of the single gray GM(1,1) model and the combined GMBP model is listed in Table 2. It can be realized from Table 2 that when predicting the number of accidents, fatalities, and injuries, the RMSE of the single gray GM(1,1) model is 3.96, 53.44, and 45.76, respectively, while the RMSE of the GMBP model is 1.09, 16.37, and 24.63, respectively. Obviously, compared with the single gray GM(1,1) model, the prediction accuracy of the GMBP model is improved.

\[
\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x}_i)^2}{n}},
\]  
(18)

where \(x_i\) is the true value, \(\bar{x}_i\) is the predicted value, and \(n\) is the number of observations.

The RMSE (root mean square error) of the single gray GM (1, 1) model and the combined GMBP model is listed in Table 2. It can be realized from Table 2 that when predicting the number of accidents, fatalities, and injuries, the RMSE of the single gray GM (1, 1) model is 3.96, 53.44, and 45.76, respectively, while the RMSE of the GMBP model is 1.09, 16.37, and 24.63, respectively. Obviously, compared with the single gray GM (1, 1) model, the prediction accuracy of the GMBP model is improved.

In Figure 4, two predicted curves indicate that the GMBP model has the benefit of the gray GM (1, 1) model in that it can deal with less sample data prediction well and the advantage of the BP neural network is that it can give an improved nonlinear approximation; thus, the prediction accuracy of the combined GMBP model is better than the single gray GM (1, 1) model, and the GMBP model’s forecasting curve is closer to the original data and in accordance with the current data’s changing trend. Figure 4 shows how the GM (1, 1) model and the GMBP model were used to forecast the number of accidents, fatalities, and injuries in China from 2021 to 2033. It can be seen from the above analysis that the number of accidents, fatalities, and injuries in major traffic accidents in China is on the decline.
As described above, the number of accidents, fatalities, and injuries will decrease gradually in the future. However, the task of preventing and reducing major road traffic accidents is still very arduous. On September 19, 2019, the CPC Central Committee and the State Council issued the “outline of building a transportation power,” in which the general goal, key tasks, and safeguard measures of building a transportation power are given. Refer to the “outline of building a transportation power,” the strategies to reduce major traffic accidents in China are given as follows.

### 3.3.1. Improve the Intrinsic Safety Level of Traffic Infrastructure
To begin, it must strengthen technical standards and guidelines for traffic infrastructure safety, continue to boost infrastructure safety expenditure, and improve essential infrastructure proper safety capacity. Second, it must establish a modern engineering construction quality management system as well as support the development of high-quality products and good management. Thirdly, it needs to improve transportation management as well as monitoring and detection of infrastructure operations, improve the professional and information level of maintenance, and enhance the durability and reliability of facilities. Finally, it needs to strengthen the quality management of traffic means to ensure the safety of traffic equipment.

### 3.3.2. Improve the Traffic Safety Production System
To begin, it must strengthen the management system in compliance with the legislation as well as the production norms and criteria for traffic safety. Secondly, it is necessary to improve the safety responsibility system, reinforce enterprise primary responsibility, and clarify departmental regulatory obligations. Thirdly, the prevention and control system must be improved in order to effectively prevent and control systemic threats and establish a third-party certification system for traffic equipment and engineering. Fourthly, it needs to strengthen the investigation and evaluation of work safety accidents. Fifthly, it needs to improve the network security system, enhance the ability of science and technology, and strengthen the security protection of traffic information infrastructure. Sixthly, it needs to improve the support system and strengthen the construction of safety facilities. Seventhly, it needs to establish a natural disaster prevention and control system to improve the ability of traffic disaster prevention and resistance. Finally, it needs to strengthen the comprehensive management of traffic safety and effectively improve the traffic safety level.

### 3.3.3. Strengthen the Ability of Traffic Emergency Rescue
Firstly, it needs to establish and improve the comprehensive traffic emergency management system, regulations, and plans; strengthen the construction of emergency rescue professional equipment, facilities, and teams; and actively participate in international emergency rescue cooperation. Secondly, it needs to strengthen the social coordination ability of emergency rescue and improve the expropriation compensation mechanism.

### 3.3.4. Improve the Accident Prevention and Safety Management of Road Traffic Accident-Prone Period and Road Section
First of all, it needs to study and determine the road traffic accident-prone period and pay attention to the safety management and accident prevention during the major festivals and holidays. Secondly, it needs to increase the investment in traffic safety facilities and management in areas with low per capita GDP and improve the level of travelers’ traffic safety awareness, road safety facilities, and road traffic safety management; thirdly, it needs to improve road traffic safety management and accident prevention in densely populated areas. In densely populated areas, the per capita road resources are low, while the total number of motor vehicles is large, and the probability of road traffic accidents increases. Therefore, it needs to strengthen traffic safety management and accident prevention from the aspects of road infrastructure construction, vehicle ownership and use management, traffic safety facilities, and technical level.

### 4. Conclusions
Using a data set for significant road traffic incidents with 10 or more deaths at one time in China from 2008 to 2020, this article compared the forecast accuracy of the gray GM (1, 1) model and the combination GMBP model. The combined GMBP model has a higher prediction accuracy than the single gray GM (1, 1) model, and the predicted curve of the GMBP model is closer to the original data and more in accordance with the real data’s evolving trend. The predicted data of the number of major traffic accidents, fatalities, and injuries in China from 2021 to 2033 using the GM (1, 1) model and the GMBP model show that the number of accidents, fatalities, and injuries in major traffic accidents in China is on the decline, and China’s major traffic accidents will continue to decline in the future years. It is necessary to adopt the following steps to meet the goal of reducing the
number of significant traffic accidents and fatalities, including raising the intrinsic safety level of transportation facilities, improving the traffic safety production system, strengthening the emergency rescue capacity of traffic accidents, and improving the traffic management and accident prevention level in accident-prone time and road sections.

Data Availability
The data used to support the findings of this study are included in the article.

Conflicts of Interest
There are no competing interests declared by the authors.

Authors’ Contributions
Qingwen Guo is responsible for the investigation, methodology, and writing. Baohua Guo is responsible for the conceptualization, methodology, and project administration. Yugang Wang is responsible for the formal analysis and data curation. Shixuan Tian is responsible for the methodology, investigation, and formal analysis. Yan Chen is responsible for the data curation and writing.

References

