Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Research on the prediction of dangerous goods accidents during highway transportation based on the ARMA model

Xiao Li\textsuperscript{a}, Yong Liu\textsuperscript{a,b,\*}, Linsheng Fan\textsuperscript{a}, Shiliang Shi\textsuperscript{a,\**}, Tao Zhang\textsuperscript{a}, Minghui Qi\textsuperscript{a}

\textsuperscript{a} School of Resource Environment and Safety Engineering, Hunan University of Science and Technology, Xiangtan, Hunan, 411201, China
\textsuperscript{b} Work Safety Key Lab on Prevention and Control of Gas and Roof Disasters for Southern Coal Mines, Hunan University of Science and Technology, Xiangtan, Hunan, 411201, China

\textbf{ARTICLE INFO}

\textbf{Keywords:}
Dangerous goods
Highway transportation
Time series
ARMA model
Accident prediction

\textbf{ABSTRACT}

The COVID-19 epidemic has caused a lack of data on highway transportation accidents involving dangerous goods in China in the first quarter of 2020, and this lack of data has seriously affected research on highway transportation accidents involving dangerous goods. This study strives to compensate for this lack to a certain extent and reduce the impact of missing data on research of dangerous goods transportation accidents. Data pertaining to 2340 dangerous goods accidents in the process of highway transportation in China from 2013 to 2019 are obtained with webpage crawling software. In this paper, the number of monthly highway transportation accidents involving dangerous goods from 2013 to 2019 is determined, and the time series of transportation accidents and an autoregressive moving average (ARMA) prediction model are established. The prediction accuracy of the model is evaluated based on the actual number of dangerous goods highway transportation accidents in China from 2017 to 2019. The results indicate that the mean absolute percentage error (MAPE) between the actual and predicted values of dangerous goods highway transportation accidents from 2017 to 2019 is 0.147, 0.315 and 0.29. Therefore, the model meets the prediction accuracy requirements. Then, the prediction model is applied to predict the number of dangerous goods transportation accidents in the first quarter of 2020 in China. Twenty-two accidents are predicted in January, 23 accidents in February and 27 accidents in March. The results provide a reference for the study of dangerous goods transportation accidents and the formulation of accident prevention and emergency measures.

1. Introduction

The chemical industry plays an important role in the economic and social development of China. The highway transportation link of chemically dangerous goods (hereinafter referred to as dangerous goods) constitutes an important part of the circulation process (Liu et al., 2018). Every year, more than 1 billion tons of dangerous goods is transported via highways in China, accounting for more than 70% of the total transport volume of dangerous goods, and more than 30% of the total transport volume of dangerous goods via highways, which reveals an increasing trend (http://www.xinhuanet.com/fortune/2018-12/03/c_1123797125.html). Dangerous goods transportation vehicles are typically loaded with dangerous goods exhibiting different hazard characteristics and drive at high speeds on highways, which readily results in traffic accidents such as collisions, roll-overs and rear-end collisions (Zhao et al., 2012), thus causing dangerous goods leakage, fire, explosion, poisoning and other disasters. At approximately 16:40 on June 13, 2020, an explosion accident involving a tank truck occurred in the Daxi Section of the Shenhua Expressway, Wenling, G15, causing 20 deaths and 175 injuries and the collapse of several houses (http://yjglj.zjj.gov.cn/art/2020/7/15/art_1229052179_50738478.html), which aroused wide public attention. Accidents caused by the transportation of dangerous goods have become one of the important restricting factors of the development of the chemical industry and have yielded great threats to lives of people and property safety. Therefore, the effective prediction, prevention and control of dangerous goods transportation accidents have remained major issues for experts and scholars in the field of safety. Within this context, the research and prediction of highway transportation accidents involving dangerous goods provide an important reference for accident prevention and emergency disposal.

\* Corresponding author. School of Resource Environment and Safety Engineering, Hunan University of Science and Technology, Xiangtan, Hunan, 411201, China.
\** Corresponding author.
E-mail address: 278541005@qq.com (Y. Liu).

https://doi.org/10.1016/j.jlp.2021.104583
Received 20 December 2020; Received in revised form 13 May 2021; Accepted 16 June 2021
Available online 28 June 2021
0950-4250/© 2021 Elsevier Ltd. All rights reserved.
Many scholars have studied highway transportation accidents of dangerous goods. A study of 322 accidents that occurred during the transportation of hazardous materials (hazmat) via highways in China from 2000 to 2008 was carried out (Yang et al., 2010). A total of 490 hazardous material transportation accidents was analyzed in China from 2006 to 2010 based on the types of materials, regions, months and economic sectors involved in the accidents (Wu et al., 2011). The theories and methods for historical analysis of accidents during the transportation of hazardous materials in China and abroad over the last twenty years were systematically reviewed (Xin and Wang, 2012). A total of 886 transportation accidents in China involving hazmat tank trucks was analyzed in terms of the time of occurrence, location of the accident, accident form, type of highway and accident cause (Shen et al., 2013). Petroleum and chemical accidents were analyzed in terms of the accident level, accident type, occurrence time, chemicals involved, operation type and country (in China home or abroad) during the period between 2011 and 2012 (Du et al., 2013). Regarding the statistical scope of hazardous material accidents (HMAs) during highway transportation, statistics of hazardous material accidents were analyzed in China from 2011 to 2013, and a research countermeasure was proposed (Li et al., 2014). Data on highway transportation accidents involving dangerous goods in Korea from 2008 to 2017 were considered to study the influencing factors of vehicle collision accidents involving dangerous goods (Jianluo and Tansel, 2016). A total of 562 HMAs was analyzed from 2013 to 2014 in China in 3 aspects (Chen et al., 2017). Data on dangerous goods highway transportation accidents in China from 2013 to 2018 were collected and studied to summarize the statistical characteristics of these accidents (Liu et al., 2020). The above studies have provided a certain guiding significance in regard to emergency response and safety management of accidents. Moreover, these studies have a commonality, namely, they all require research data and achieve a high integrity and continuity. When data are affected by external factors and are missing or damaged, similar research is difficult to conduct. In 2020, the COVID-19 epidemic raged around the world. The Chinese government and people alike adopted unprecedented strict isolation measures, and the dangerous goods transportation industry stagnated for a long time. Therefore, the data integrity and continuity of highway transportation of hazardous chemicals were greatly affected, especially in the first quarter of 2020, and the data were almost entirely distorted. The use of the actual data to conduct research on dangerous goods accidents could affect the research results. Therefore, it is highly important to study the evolution of hazardous chemical accidents to predict and replace any missing data.

In the past, time series research methods have been widely applied in the fields of finance (Ding, 2017) and artificial intelligence (Cools et al., 2007). Time series research methods exhibit a notable advantage in the study of time-ordered data series (Bouarouk and Djeddour, 2015). Moreover, although single highway transportation accidents of dangerous goods exhibit a high uncertainty, the features of these accidents are consistent with the characteristics of time series from the point of view of the multi-year occurrence regularity (Zhu and Li, 2015). Currently, researchers apply time series research methods in accident prediction, highway transportation analysis, etc. (Yagi et al., 2018).

In view of the above, information on 2340 dangerous goods accidents during highway transportation in China from 2013 to 2019 was obtained for analysis purposes. In this study, a time series analysis method was applied to study the acquired accident data, and an autoregressive moving average (ARMA) prediction model was established to predict the number of accidents in China in the first quarter of 2020. This study provides a reference for research on future dangerous goods accidents and the planning of emergency management measures. In the meanwhile, this study can also provide reference for the impact of other potential global epidemics or other public events on the safety of dangerous goods transportation.

2. Methods

2.1. Data acquisition

As one of the important steps in the study of dangerous goods accidents, the acquisition of accident data guarantees the smooth development of the study and prediction of accident statistics (Yoon and Craig, 2015). The source and method of accident data acquisition affect the authenticity and integrity of the acquired data (Wang et al., 2018).

In terms of data acquisition methods, the data of hazardous material accidents considered in this study are obtained from certain websites, for example, from the Petrochemical Accident Analysis Platform (PAAP) website. The PAAP, as an information publishing platform of Sinopec accidents, is operated by the National Registration Center for Chemicals (NRCC) and the Sinopec Research Institute of Safety Engineering. The PAAP is an online statistical platform and continually reports Sinopec accidents.

In regard to data acquisition tool Octoparse, web crawling software (provided by Skieer Co. Ltd., Shenzhen, PR China) is adopted to obtain data on dangerous goods transportation accidents. The software collects data on dangerous goods accidents in China from 2013 to 2019 by establishing a user-defined web information collection model.

2.2. Data processing

The original data obtained from the information platform include various accidents in the production, storage and use of dangerous goods and traffic accidents related to the transportation of non-dangerous goods. These data are unrelated to the research objective, so further data processing is needed. The method of manual data cleaning was employed to eliminate irrelevant data, and 2340 valid data points were obtained.

A time series is a series formed by sorting the values of a certain statistical indicator in chronological order. Time series forecasting analyzes a given time series based on the development process, direction and trend reflected by the time series, analogy or extent, and predicts the level that may be reached during the next period or in the next few years. In this article, the monthly number of highway transportation accidents involving dangerous goods in China from 2013 to 2019 is adopted as a data sample for research and prediction. Table 1 presents the time series of monthly accidents of highway transportation involving dangerous goods in China from 2013 to 2019 obtained after data processing.

As indicated in Table 1, correspond to January to December 2013, X1 – X12 correspond to January to December 2014, X13 – X24 correspond to January to December 2015, X25 – X36 correspond to January to December 2016, X37 – X48 correspond to January to December 2017, X49 – X60 correspond to January to December 2018, and X61 – X72 correspond to January to December 2019.

According to the data in Table 1, the maximum number of monthly accidents during highway transportation of dangerous goods in China from 2013 to 2019 was 64, which occurred in December 2013. The month with the smallest number was February 2018, at nine incidents. The mean value of the data series was approximately 27.86, with a median of 28. The total number of accident data points in each year remained basically stable.

3. Characteristics of the ARMA model

3.1. Overview of the ARMA model

The ARMA model is a common stochastic time series prediction model proposed by Box and Jenkins, also referred to as the B-J method (Wang et al., 2016). The model is a short-term forecasting method of time series with a high accuracy. Its basic idea is that certain time series comprise a group of random variables dependent on the time. Although a single sequence value retrieved from a time series contains...
In practical use, there are three basic types of ARMA models: autoregressive (AR), moving average (MA) and autoregressive moving average (ARMA) models. The ARMA model is a mixed model, and the structure and characteristics of time series can be understood more fundamentally, and optimal prediction with a minimum variance can be achieved.

In general, the modelling steps of the time series with the ARMA model are shown in Fig. 1 below.

In this paper, the establishment and prediction of the ARMA model are realized in Econometric Views (Eviews) software, which is Windows-based software developed by QMS (Quantitative Micro Software Co., USA). QMS specializes in data analysis, regression analysis, and prediction. Eviews is characterized by a simple operation and powerful functions and is widely applied in the fields of data analysis, evaluation and prediction.

### 3.2. Stationary test of the time series

The use of the ARMA model for prediction requires that the time series is stationary, so stationarity testing of the data series should be carried out before the construction of the prediction model.

The time series diagram of the exported data in Eviews software is shown in Fig. 2. Fig. 2 shows that the observed monthly number of dangerous goods highway transportation accidents from 2013 to 2019 fluctuates around the mean value, and the mean value does not change over time. Therefore, it is preliminarily determined that the data sequence reveals no obvious trend and basically remains stable.

To ensure the stability of the data series studied, the augmented Dickey-Fuller (ADF) test was further performed to assess the unit root of the sequence. The sequence was evaluated in Eviews software, and the test results are shown in Fig. 3.

As shown in Fig. 3, the t-statistic value is less than its critical value at the 1%, 5% and 10% test levels. Hence, the null hypothesis, i.e., the
should be employed to identify the most suitable model type for the correlation function (ACF) and partial autocorrelation function (PACF).

3.3. Model form recognition

When the B-J method is applied to model a time series, the autocorrelation function (ACF) and partial autocorrelation function (PACF) should be employed to identify the most suitable model type for the sequence. The autocorrelation function is defined as:

\[ r_k = \frac{\sum_{t=1}^{n} (X_t - \bar{X})(X_{t+k} - \bar{X})}{\sum_{t=1}^{n} (X_t - \bar{X})^2} \]

Multiplying both sides of the above equation by \( X_{t+k} \) yields:

\[ X_{t+k}X_t = X_{t+k} \sum_{i=1}^{n} \rho_i X_{t+i} - X_{t+k} \sum_{j=1}^{n} \phi_j u_{t+j} + X_{t}u_t \]  

Then, by determining the expectation of both sides, the k-order co-variance function of the ARMA model can be obtained:

\[ \gamma_k = \sum_{i=1}^{p} \phi_i \gamma_{k-i} - \sum_{j=1}^{q} \theta_j E(X_{t+i}u_{t+j}) + E(X_{t}u_t) \]

Moreover, \( E(X_{t+i}u_{t+j}) = 0, k > I, \) and thus:

\[ \gamma_k = \sum_{i=1}^{p} \phi_i \gamma_{k-i}, k \geq (q+1) \]  

Therefore, the k-order autocorrelation function of the ARMA model can be obtained as follows:

\[ \rho_k = \phi_1 \rho_{k-1} + \cdots + \phi_p \rho_{k-p}, k \geq (q+1) \]  

The partial autocorrelation function is another statistical feature of the ARMA model, which is defined as:

\[ \varphi_{ik} = \frac{\rho_k - \sum_{j=1}^{k-1} \rho_{k-j} \rho_j}{1 - \sum_{j=1}^{k-1} \rho_{k-j} \rho_j}, k = 2, 3, 4\ldots \]  

In the above equation, \( k \) denotes hysteresis, \( \varphi_{ij} = \varphi_{k-i-j} - \varphi_{k-1-k-j} \)  

The characteristics of the autocorrelation and partial autocorrelation functions indicate that in the process of the autocorrelation function of the sequence, at \( k > p, \rho_k \) is always 0, suggesting p-order truncation. If \( \rho_k \) is always a non-zero value regardless of \( k \) and the partial correlation function decays exponentially, this property is referred to as trailing.

The time series subject to the ARMA\((p, q)\) model exhibits obvious statistical characteristics, and its autocorrelation function and partial autocorrelation function reveal the phenomenon of trailing or truncating. Therefore, model identification can be accomplished based on the autocorrelation or partial autocorrelation phenomenon of trailing or truncating, respectively. The specific determination method is summarized in Table 2.

Table 2 Characteristics of the autocorrelation function of the ARMA model.

| Model category | Autocorrelation function | Partial autocorrelation function |
|---------------|--------------------------|---------------------------------|
| AR(p)         | trailing                 | p-order truncation               |
| MA(q)         | q-order truncation       | trailing                         |
| ARMA(p, q)    | trailing                 | trailing                         |
Based on the ACF and PACF diagrams in Fig. 4, it is also determined that the partial autocorrelation function decreases after lagging 2 steps, while the autocorrelation coefficient of the time series occurs at the edge of the random interval when lagging 4 steps. The models that may be suitable for preliminary assessment are AR(1), AR(2), MA(3), MA(4), ARMA(1,3), ARMA(2,3), ARMA(1,4) and ARMA(2, 4).

In the process of determining the order of the ARMA model, the adjusted R² value, Akaike information criterion (AIC) and Schwarz criterion (SC) are important criteria for the selection of the model. When choosing the best-fitting model, it is necessary to comprehensively examine the values of these criteria.

The AR(1), AR(2), MA(3) MA(4), ARMA(1,3) ARMA(2,3), ARMA(1,4) and ARMA(2, 4) models are calculated in Eviews software, and the adjusted R², AIC and SC values of these eight groups of models are determined. The results are listed in Table 3.

By comparing the adjusted R², AIC and SC values of the eight groups of models, the data indicated that MA (4) model yields the lowest AIC and SC values, while the adjusted R² value was high. Therefore, MA (4) model could be considered as the best model.

### 3.4. Model parameter estimation

After the best-fitting model is selected, the least-squares estimation method is applied to estimate the value of the unknown parameters in the model. The estimation results for this model in Eviews software are shown in Fig. 5.

The above table indicates that the least-squares estimation result of the model is:

$$X_t = 0.381 + 0.391 E_{t-2} + 0.2805 E_{t-3} + 0.3737 E_{t-4} + 0.3012 E_{t-5} + 0.3491 E_{t-6} + 0.163 E_{t-7} + 0.018 E_{t-8} - 0.098 E_{t-9} - 0.178 E_{t-10} + 0.092 E_{t-11} - 0.067 E_{t-12}$$

The estimated value of the error term variance is:

$$\hat{\sigma}^2 = 8.254781$$

### 3.5. Model test

After completing the above steps, the ARMA prediction model is basically determined, after which adaptability testing of the ARMA model is carried out. In essence, the model residual sequence is tested in regard to white noise. If the residual sequence is not white noise, this indicates that there remains useful information in the residual sequence that has not been extracted, so the model should be further improved.

The autocorrelation function of the residual sequence is:

$$r_k = \sum_{n=1}^{\infty} e_t e_{t-k} \div \sum_{n=1}^{\infty} e_t^2$$

In the above equation, n is the sequential observation quantity to calculate and m is the lag period. The test statistic is:

$$Q = n(n+2) \sum_{k=1}^{m} \frac{r_k^2}{n-k}$$

Under the null hypothesis, Q obeys the distribution of χ²(m – p – q).

Based on the autocorrelation analysis results depicted in Fig. 6, it is determined that the autocorrelation coefficients of the residual series fall within the random interval, indicating that the residual series is purely random. Hence, the MA(4) model achieves a good fitting effect on the time series.

### Table 3

Comparison of the model parameters.

| Model       | Adjusted R² | AIC      | SC        |
|-------------|-------------|----------|-----------|
| AR(1)       | 0.125982    | 7.248002 | 7.334817  |
| AR(2)       | 0.211316    | 7.159373 | 7.375126  |
| MA(3)       | 0.181183    | 7.206588 | 7.351279  |
| MA(4)       | 0.242495    | 7.145744 | 7.319374  |
| ARMA(1,3)   | 0.187075    | 7.211502 | 7.385131  |
| ARMA(1,4)   | 0.248935    | 7.152597 | 7.354665  |
| ARMA(2,3)   | 0.214968    | 7.190128 | 7.392696  |
| ARMA(2,4)   | 0.246414    | 7.167819 | 7.399326  |

### Table 4

Comparison of the model parameters.

| Variable | Coefficient | Std. Error | t-Statistic | Prob.   |
|----------|-------------|------------|-------------|---------|
| C        | 27.65419    | 2.047461   | 13.50658    | 0.0000  |
| MA(1)    | 0.397444    | 0.116483   | 3.41027     | 0.0010  |
| MA(2)    | 0.441044    | 0.155228   | 2.841272    | 0.0057  |
| MA(3)    | 0.085528    | 0.141358   | 0.605043    | 0.5469  |
| MA(4)    | 0.317902    | 0.138652   | 2.239369    | 0.0245  |
| SKMASQ   | 63.7787     | 8.261044   | 7.721303    | 0.0000  |

### Fig. 4. Time series ACF diagram.

### Fig. 5. ARMA model estimation results.

### Fig. 6. Test result chart of model residual sequence.
4. Model accuracy test and prediction

4.1. Accuracy test

Before predicting the accident data in 2020, the number of dangerous goods highway transportation accidents in each month from 2017 to 2019 is first predicted, and the predicted results are compared to the actual observed values from 2017 to 2019 to evaluate the prediction accuracy of the model.

The determined MA(4) model is applied in Eviews software to predict the number of accidents in each month from 2017 to 2019, and a comparison of the predicted and actual values is shown in Fig. 7. Fig. 7 reveals that the trend of the predicted monthly accidents during dangerous goods highway transportation in China from 2017 to 2019 is basically consistent with the trend of the actual values.

To evaluate the prediction accuracy, we calculated the mean absolute percentage error (MAPE) in the prediction results from 2017 to 2019. The equation of the MAPE is:

\[ MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \tag{12} \]

After calculation, the MAPE between the predicted and actual values from 2017 to 2019 is summarized in the table below.

Table 4

| Year | MAPE  |
|------|-------|
| 2017 | 0.147 |
| 2018 | 0.315 |
| 2019 | 0.29  |

The data in Table 4 indicate that the predicted MAPE values from 2017 to 2019 are low, which suggests that the model achieves a suitable prediction accuracy and meets expectations.

4.2. Prediction results

After verifying that the prediction accuracy of the ARMA model meets the short-term prediction demand, the MA(4) model is employed to predict the number of dangerous goods highway transportation accidents in each month from 2017 to 2019.

Fig. 7. Comparison of the predicted and actual numbers of accidents from 2017 to 2019.
accidents in each month of the first quarter of 2020. Since the number of accidents should be an integer, the predicted values obtained with Eviews software are rounded up. The predicted results are listed in Table 5.

The predicted data indicate that the number of highway transportation accidents involving dangerous goods in the first quarter of 2020 in China is determined to be 22 incidents in January, 23 incidents in February and 27 incidents in March. The predicted values are significantly higher than the actual values.

The predicted number of dangerous goods highway transportation accidents in the first quarter of 2020 is compared to the accident data pertaining to the first quarter of the last three years. The comparison results are listed in Table 6.

Table 6 presents a comparison of the actual number of accidents in each month of the first quarter from 2017 to 2019 to the predicted number of accidents during the same period in 2020. Obviously, although there were many accidents in 2017, the number of accidents stabilizes over the past two years. The predicted number of accidents in the first quarter of 2020 is also basically the same as that over the past two years. Compared to the actual number of accidents in the first quarter of 2020, the predicted number is significantly smaller. This suggests that the actual number of accidents in the first quarter of 2020 has indeed been affected, and the predicted value is highly credible.

4.3. Analysis and discussion

1) It has been proven that it is very important to obtain statistics and conduct research on highway transportation accidents involving dangerous goods. However, previous studies have often ignored the serious impact of missing data on the obtained research findings. This paper aimed to implement a time series method to study the monthly number of dangerous goods highway transportation accidents in China from 2013 to 2019 and establish a prediction model. The established model was then applied to predict the number of highway transportation accidents involving dangerous goods in the first quarter of 2020.

2) The occurrence of highway transportation accidents involving dangerous goods is typically considered to exhibit notable randomness and uncertainty, so the difficulty of accident prediction greatly increases. However, by analysing the time series of monthly accidents during dangerous goods highway transportation in China from 2013 to 2019, it was observed that although a single dangerous goods highway transportation accident is highly uncertain, the change in the whole sequence exhibits a certain regularity. The above description is applicable to the study of the time series method of the monthly accidents during dangerous goods highway transportation in China from 2013 to 2019.

3) According to the calculation data provided in Table 1, the average monthly number of dangerous goods highway transportation accidents in China from 2013 to 2019 is approximately 27.86, and the total number of accidents in each year remains basically stable. The monthly average number of accidents from 2013 to 2019 was calculated, which was approximately 28. This indicates that the time series of the monthly number of accidents during highway transportation of dangerous goods in China from 2013 to 2019 remains stable, which is suitable for ARMA model research and prediction.

Table 5
Forecast for the first quarter of 2020.

| Month   | Actual value | Predicted value |
|---------|--------------|-----------------|
| 2020M01 | 17           | 23              |
| 2020M02 | 9            | 24              |
| 2020M03 | 16           | 29              |
| 2020Q1  | 42           | 76              |

4) According to the data listed in Table 6, the predicted number of highway transportation accidents involving dangerous goods in each month of the first quarter in 2020 fluctuates to a certain extent over the actual number of accidents during the same period from 2017 to 2019. However, within the acceptable range, the prediction results provide a reference value. By comparing the actual number of accidents in the first quarter from 2017 to 2019 to the predicted number of accidents in 2020, it is determined that the number of accidents in the first quarter over the past two years generally remains stable and fluctuates near 70. The monthly average number of accidents reaches approximately 25, slightly lower than the overall monthly average value. The first quarter is not the high-incidence period in terms of accidents throughout the year, which is consistent with the statistical characteristics of accident data from 2013 to 2019.

5) The forecast data reveal that there is a slight upward trend in the number of accidents in the first quarter over the last three years. Regulatory authorities should pay attention to this phenomenon, further consolidate the main responsibility of the safe operation of dangerous goods highway transportation enterprises and improve the safety management level of these enterprises. Conversely, measures targeted at the rectification of highway transportation safety of dangerous goods should be formulated to effectively curb the rising trend of accidents.

5. Conclusions

1) In this paper, the time series analysis method is employed to study highway transportation accident data involving dangerous goods. Although the occurrence of highway transportation accidents involving dangerous goods is highly random and uncertain, the rules of accident development and change can still be studied and predicted with the time series method, which lays a foundation for the research and prediction of the evolution of highway transportation accidents involving dangerous goods.

2) In this paper, the monthly number of dangerous goods highway transportation accidents in China from 2013 to 2019 is determined, a corresponding time series is constructed, and its stability is tested. The ARMA model is applied to fit the constructed time series. The established model is used to predict the number of accidents during dangerous goods highway transportation in each month from 2017 to 2019. The predicted results are compared to the actual values. By analysing the prediction error, the predicted MAPE values from 2017 to 2019 are 0.147, 0.315 and 0.29. Therefore, it is determined that the sequence fitting degree is suitable, and the prediction accuracy meets the research needs.

3) The ARMA model is adopted to predict the number of highway transportation accidents involving dangerous goods in each month of the first quarter in 2020. The prediction results suggest that the predicted number of accidents in the first quarter of 2020 is 22 in January, with 23 incidents predicted in February and 27 incidents predicted in March.

4) The actual number of dangerous goods highway transportation accidents in the first quarter of 2017, 2018 and 2019 is compared to the predicted values in 2020. The results indicate that the number of accidents in the first quarter of recent years exhibits a slight upward
trend. This provides a reference for safety supervision departments to suitably manage dangerous goods transportation enterprises.

Author statement

Xiao LI: Investigation, Data processing, Modification. Yong LIU: Conceptualization, Methodology, Investigation, Formal analysis, Project administration, Data curation, Writing - Original Draft. Lin-sheng FAN: Investigation, Data processing. Shi-liang SHE: Resources, Writing - Review & Editing, Supervision. Ming-hui QI: Writing -Review & Editing. Tao ZHANG: Writing -Review & Editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos.52074118, 51974120 and 51974119), Hunan Provincial Natural Science Foundation of China (Grant No. 2019JJ50152), Scientific Research Foundation of Hunan Provincial Education Department (Grant Nos. 18B210, 19A167 and 20A192).

References

Boularouk, Y., Djeddour, K., 2015. New approximation for ARMA parameters estimate. Math. Comput. Simulat. 118, 116–122. https://doi.org/10.1016/j.matcom.2015.01.004.

Chen, X.Y., Shi, S.L., Ren, J.Z., et al., 2017. Statistical analysis and countermeasure research on the traffic accident of highway dangerous chemicals from 2013 to 2014 in country. J. Hunan Univ. Sci. Technol. (Nat. Sci. Ed.) 3, 91–95. https://doi.org/10.13582/j.cnki.1672-9102.2017.03.016 (in Chinese).

Cools, M., Moons, E., Wets, G., 2007. Investigating effect of holidays on daily traffic counts: time series approach. Transport. Res. Rec. 2019 (1), 22–31. https://doi.org/10.3141/2019-08.

Ding, J., 2017. The filtering based multi-innovation stochastic gradient algorithm for multi-input output-error systems. In: 2017 36th Chinese Control Conference (CCC). Dalian, pp. 128–133. https://doi.org/10.23919/ChiCC.2017.8027332.

Du, H.Y., Wang, Y.P., Liu, J.C., 2013. Statistical analysis on petroleum and chemical accidents at home and abroad during the period of 2012. Journal of Safety Science and Technology 6, 184-188 (in Chinese).

Eran, B., Tanel, B., 2016. A transportation network assessment tool for hazardous material cargo routing: weighing exposure health risks, proximity to vulnerable areas, delay costs and trucking expenses. J. Loss Prev. Process. Ind. 40, 266–276. https://doi.org/10.1016/j.jlp.2016.01.002.

Li, J., Bai, X.Y., Ren, Z.Z., et al., 2014. Statistical analysis and countermeasures of hazardous chemicals accidents occurring in China during 2011-2013. Journal of Safety Science and Technology 6, 142-147 (in Chinese).

Liu, Y., Li, S., Li, R.Q., et al., 2018. Transport risk assessment model of hazardous materials in highway tunnel and its application. J. Hunan Univ. Sci. Technol. (Nat. Sci. Ed.) 33, 20–25. https://doi.org/10.13582/j.cnki.1672-9102.2018.04.004 (in Chinese).

Liu, Y., Fan, L.S., Li, X., et al., 2020. Trends of hazardous material accidents (HMAs) during highway transportation from 2013 to 2018 in China. J. Loss Prev. Process. Ind. 66, 104150. https://doi.org/10.1016/j.jlp.2020.104150.

Shen, X.Y., Li, X.N., Xie, P., et al., 2013. Statistical analysis on 866 highway HAZMAT transportation accidents by the tank truck. Journal of Safety Science and Technology 11, 43–48 (in Chinese).

Wang, Z., Huo, Y., Liu, Z., 2016. Prediction Model on Energy Consumption of Highway Transportation in Inner Mongolia Based on ARMA. https://doi.org/10.1007/978-981-10-3551-7_8.

Wang, B., Wu, C., Huang, L., et al., 2018. Prevention and control of major accidents (MAs) and particularly serious accidents (PSAs) in the industrial domain in China: current status, recent efforts and future prospects. Process Saf. Environ. Protect. 117, 254–266. https://doi.org/10.1016/j.psep.2018.04.025.

Wu, Z.Z., Zhang, S.Z., Zhang, Y., 2011. Statistical analysis of hazardous chemicals accidents occurring in China during 2006–2010. Journal of Safety Science and Technology 7, 5–9 (in Chinese).

Xin, C.L., Wang, J.L., 2012. Review on historical analysis of accidents in the transportation of hazardous materials. Chinese Journal of Safety Science 22, 89–94 (in Chinese).

Yagi, S., Nobel, D., Kawaguchi, H., 2018. Time series comparison of models of auto and motorcycle ownership and mode choice in a changing transportation environment. Transport. Res. Rec. 2317 (1), 40–50. https://doi.org/10.3141/2317-06.

Yang, J., Li, F.Y., Zhou, J.B., et al., 2010. A survey on hazardous materials accidents during highway transport in China from 2000 to 2008. J. Hazard Mater. 184 (1-3), 647–653. https://doi.org/10.1016/j.jhazmat.2010.08.085.

Yoon, Kyung Cha, Craig, A., 2015. Stow, Mining web-based data to assess public response to environmental events. Environ. Pollut. 198, 97–99. https://doi.org/10.1016/j.envpol.2014.12.027.

Zhu, K., Li, W.K., 2015. A bootstrapped spectral test for adequacy in weak ARMA models. J. Econom. 187, 113–130. https://doi.org/10.1016/j.jeconom.2015.02.005.