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.
The effects of Wuhan highway lockdown measures on the spread of COVID-19 in China

Xin Meng a, Mingxue Guo a, Ziyou Gao a,*, Zhenzhen Yang a,**, Zhilu Yuan b,c, Liujiang Kang a,***

a School of Traffic and Transportation, Beijing Jiaotong University, Beijing, 100044, China
b School of Architecture and Urban Planning, Research Institute for Smart Cities, Shenzhen University, Shenzhen, 518052, China
c Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources, Shenzhen, China

ARTICLE INFO
Keywords:
COVID-19
Human mobility
Epidemic control and prevention
Lockdown measures
Highway toll system

ABSTRACT
To verify the effects of Wuhan highway lockdown measures on the spread of COVID-19 across China cities, we extracted the vehicle outflow from Wuhan to 245 cities from the Chinese highway toll system. A dynamic exponential risk model that considered the vehicle outflow, city gross domestic product, city population, and distance between two cities was established to characterize the spread of pandemics and quantify the blocking effects. Results showed that an early highway lockdown measure could indeed reduce the confirmed cases and vehicles with 1–9 seats played a leading role. The confirmed cases in Guangxi, Henan, and Shanxi could be reduced by more than 50%, as well as Hubei by 20% if the highway was closed 3 days in advance. The blocking effects on Fujian, Jiangxi, Guangdong, Hunan, and Shandong were not obvious, where the number of confirmed cases only decreased by a small proportion (below 10%). The findings could be used to help each provincial government to adjust policies properly and improve the effectiveness of epidemic control and prevention. Moreover, the proposed method could also be applied to various countries or regions affected by COVID-19, as well as other similar pandemics.

1. Introduction
In the early 2020, the coronavirus disease 2019 (COVID-19) has swept the globe. Globally, as of December 16, 2021, there have been 271,376,643 confirmed cases of COVID-19, including 5,324,969 deaths (WHO, 2021).

Previously, a study proved that if the epidemic prevention measures had been taken one week, two weeks, and three weeks earlier, the infection cases in China could be reduced by 66%, 86%, and 95%, respectively. However, if the measures had been delayed by one week, two weeks, or three weeks, the number of infections in China would increase by 3 times, 7 times, and 18 times (Lai et al., 2020). Zhong Nanshan’s team also mentioned in their study that postponing the implementation of mandatory measures such as “closing the city” for five days would triple the scale of the epidemic in China. However, if the measures had been taken five days earlier, the number of confirmed cases in Hubei would not exceed 25,000 (Yang et al., 2020). Most of these studies focused on the scale of the whole country and rarely conducted specific studies on smaller spatial scales (Kraemer et al., 2020; Li et al., 2021; Zhang et al., 2020). Studies on the whole scale of the country might only capture the macroscopic transmission mode, but different impacts of epidemic situations on provinces and regions were usually ignored. In addition, other scholars also used data reflecting population movements (such as mobile device data, high-speed rail network data, and air transportation system data) to analyze the effects of lockdown measures (Hu et al., 2021; Nouvellet et al., 2021), but the highway traffic data has not been used yet. The highway is an important part of China’s road traffic system, which bears most of human mobility, especially in the period approaching the Chinese Spring Festival. If all the vehicles are checked, efficiency is low and it is easy to cause traffic congestion. If we understand the effects of different vehicle types on the spread of pandemics and check one or two types, not only can we improve the working efficiency and avoid traffic congestion, but also reduce the working hours of medical staff. In response to the above problems, we plan to start the study from two aspects. One is to study whether the effects of Wuhan highway lockdown measures on different

* Corresponding author.
** Corresponding author.
*** Corresponding author.
E-mail addresses: zygao@bjtu.edu.cn (Z. Gao), yangzhenzhen@bjtu.edu.cn (Z. Yang), kanglj@bjtu.edu.cn (L. Kang).

https://doi.org/10.1016/j.tranpol.2022.01.011
Received 31 May 2021; Received in revised form 27 December 2021; Accepted 17 January 2022
Available online 20 January 2022
0967-070X/© 2022 Elsevier Ltd. All rights reserved.
provinces/cities are different. The other one is to classify the highway traffic data according to vehicle types and explore the impact of different vehicle types on the spread of COVID-19.

In this paper, we obtain 139 million pieces of real-time data from the Chinese highway toll system and screen out 443,310 records of the vehicle outflow from Wuhan before the lockdown measures. We categorize the data according to the type of vehicles, set different weights for each type, and aggregate them together. After studying the association of the vehicle outflow, city gross domestic product (GDP), city population, and distance between two cities with the confirmed cases, we take them as the factors to model the confirmed case of 245 cities in China and try to capture the trend of the epidemic over time. We improve the risk model proposed by Jin et al. (2020) to make it more suitable for our problem. Results show that there is a positive correlation between the confirmed cases and the vehicle outflow from Wuhan. Vehicles with 1–9 seats are the most relevant to the spread of COVID-19. For a national perspective, the highway lockdown measure is effective. However, due to the different key factors in provinces and regions, the measure has different effects. These results have instructive significance for formulating specific intervention policies and quickly assessing their effectiveness. At the same time, it can be used by policymakers to adopt measures according to local conditions, avoiding overreaction, policy simplification, and a one-size-fits-all phenomenon.

The remainder of the paper is structured as follows. Section 2 introduces the source of the data used in this paper. Section 3 analyzes the correlation between factors and confirmed cases. The importance of various factors is also evaluated. Section 4 constructs the improved risk model reflecting the relationship between the spread of COVID-19 and factors to fit the distribution of confirmed cases and then establishes the dynamic exponential risk model to describe the change of confirmed cases over time, observing the characteristics in different provinces and regions of China. Section 5 assesses the implementation effects of the highway lockdown measure. Section 6 proposes some policy suggestions. Section 7 discusses the results and concludes this paper.

2. Data description

In this study, we use the whole quantity data from the Chinese highway toll system of all municipalities and some prefecture-level cities in China, from January 18 to January 30, 2020. Since the data includes all vehicle inflows and outflows of various types in each city, it can reflect the distribution of the highway traffic completely and accurately when compared with other types of data. Thereby it reveals human mobility between China cities. Note that Wuhan was locked down on January 23, 2020, the vehicle outflow dropped rapidly after January 23 (see Fig. 1). We pick out the data for six days from January 18 to January 23, 2020, including the vehicle outflow from Wuhan to four municipalities (Beijing, Tianjin, Shanghai, and Chongqing), eighteen provinces (Hebei, Shanxi, Shaanxi, Liaoning, Heilongjiang, Jiangsu, Anhui, Fujian, Guangdong, Guizhou, Yunnan, Shanxi, Henan, Hunan, Hubei, Jiangxi, Sichuan, and Gansu), and four autonomous regions (Inner Mongolia, Guangxi, Ningxia, and Qinghai), a total of 240 prefecture-level cities and one autonomous prefecture (Enshi, one prefecture in Hubei) without the traffic to Jilin, Zhejiang, Xinjiang, Hainan, Tibet, Hong Kong, Macau, and Taiwan (Since the main contribution of the vehicle mobility to the spread of pandemics is the movement of the infected population, we excluded the truck and special vehicle data and only kept the intercity coach and car data).

The city gross domestic product (GDP) and the city population data sets we used are from the 2019 City/Prefectures Statistical Year Book of China. From the Baidu Epidemic Real-Time Big Data Reporting Platform, we get the confirmed case data of each city from January 24 to February 20, 2020.

3. Correlation test

3.1. Correlation between factors and confirmed cases

The transmission process of COVID-19 is that virus is excreted from the infection, through close contact, respiratory droplets, and other transmission channels passed into the susceptible population to cause new infections (Peng et al., 2020). The highway vehicle outflow can reflect the number of infected from Wuhan to other cities corresponding-ly. Since highway travel is greatly limited by distance, we believe that the vehicle outflow from Wuhan is affected by distance. The number of susceptible individuals in each region is mainly related to the city population. A remarkable thing is the opportunity and frequency of contact between susceptible and infected populations. GDP can reflect the level of social activity according to the level of regional economic development. The higher the economic level, the more frequent the movement and contact of people (Yan et al., 2017), which will accelerate the spread of pandemics. Therefore, we regard the vehicle outflow, population, GDP, and distance as factors that affect the spread of COVID-19. Next, we will analyze the importance of these factors from a quantitative perspective.

Firstly, we display the confirmed cases and the number of vehicle outflows from Wuhan on the map of China. Same as the qualitative analysis, if a city received more vehicles, it would have more confirmed cases (see Fig. 2). In fact, according to the result after taking the logarithm, we find that there is a certain linear correlation between the vehicle outflow and confirmed cases (see Fig. 3). Then, we use Pearson’s correlation coefficient to verify the influence of different factors. Among these factors, the outflow has a higher correlation with cases, whose Pearson’s correlation coefficient is more than 0.9. While the coefficient between GDP and confirmed cases and that between population and confirmed cases, gradually decrease over time, which suggests that community transmission begins to predominate. Besides, the coefficient with distance is always negative, and the value of the correlation coefficient increases over time, which verifies that distance is also an important factor (see Fig. 4).

According to the Vehicle Classification of the Toll for Highway (People’s Republic of China Transportation Industry Standards, JT/T 489-2019),

![Fig. 1. (a) The number of vehicle outflows from Wuhan before and after the lockdown measures. (b) The number of vehicle outflows from Wuhan to other provinces.](image-url)
vehicles are classified into four types: 1–9 seats, 10–19 seats, 20–39 seats, and above 40 seats respectively. In our data, the vehicles are divided into these four types, too. To analyze the importance of the above four vehicle types, we calculate the Pearson’s correlation coefficient between each vehicle type and the confirmed cases over one month (see Fig. 5). Among the four types, vehicles with 1–9 seats have the strongest correlation, while vehicles with 20–39 seats are the weakest.

3.2. Analysis of the importance of factors

In addition to vehicle types, we analyze the importance of various factors using the random forest. Let $x_1$, $x_2$, $x_3$, and $x_4$ be the numbers of vehicles with 1–9 seats, 10–19 seats, 20–39 seats, and above 40 seats from Wuhan to a city during Jan. 18–23, respectively. Moreover, $x_5$ represents city population. $x_6$ represents city GDP. $x_7$ represents the distance between two cities. $y$ represents the number of confirmed cases in a certain city. To reduce the order of magnitude for $x_1 \sim x_7$, we use the following vector, $[\ln x_1, \ln x_2, \ln x_3, \ln x_4, \ln x_5, \ln x_6, \ln x_7]$, as variables. Similarly, $\ln y$ is considered as the target value. For accuracy, we delete the results of $\ln y = 0$, accounting for 4.49% of the 245 cities. In the random forest model, we set 500 decision trees. 80% of the data is taken as the training set and 20% as the test set. The Mean Absolute Percentage Error (MAPE) is used to measure the accuracy of prediction. As a result, $\text{MAPE} = 81.63\%$ and the normalized importance of various factors is shown in Table 1.
As we can see, the vehicles with 1–9 seats are the key factor to the confirmed cases, followed by GDP, population, and the vehicles above 40 seats. Vehicles with 10–19 seats and 20–39 seats are less important as they account for a small proportion among the 443310 vehicles outflows from Wuhan (1.08%: 4806 vehicles in total). In this study, for a comprehensive analysis, we take all the above factors into account and model the confirmed cases based on the above 7 factors.

4. Modeling confirmed cases

4.1. Model description

There is a strong correlation between confirmed cases and human mobility. Therefore, it is meaningful to analyze the trend of human mobility and depict the spread of pandemics. The gravity model is classic in the field of human mobility, which is verified by various data in many studies. In terms of COVID-19, Jia et al. (2020) used an improved gravity model to establish a risk model to study the spread of COVID-19 nationwide based on the mobile device data. However, the mobile device data is single dimension data without weights. In this paper, we aim to study both the spread of COVID-19 and the correlation between different vehicle types with confirmed cases. So, we use the risk model developed by Jia et al. (2020) in Equation (1) and improve it by considering different vehicle types.

\[ y_i = c \prod_{j=1}^{k} e^{x_i \lambda_{kj}} \]  

(1)

where \( y_i \) is the number of confirmed cases in administrative region \( i \), \( x_{ij} \) is the population outflow from Wuhan to \( i \), \( \lambda_{kj} \) is the city population of \( i \), \( x_{ij} \) is the GDP of \( i \), \( \beta_j \), and \( \lambda_i \) are parameters that need to be estimated, where \( \lambda_k \) is the fixed effect of province \( k \). And \( n \) is the number of provinces. \( \lambda_k \) is equal to 1 if administrative region \( i \) belongs to the province \( k \), otherwise, \( \lambda_k = 0 \).

Note that different vehicle types have different carrying capacities, and the impacts on the spread of COVID-19 are also different. We give each vehicle type a weight, donated by \( w_k \) and the numbers of vehicles in different types are \( n_{z_i} \), where \( z = 1, 2, 3, 4 \). According to the toll standard of the highway vehicles, the value of \( w_{1i}, w_{2i}, w_{3i}, \) and \( w_{4i} \) are 1, 1.2, 2, and 4. Therefore, \( x_{ij}' \) can be captured by Equation (2).

### Table 1

| Variable | Parameter | Importance (sum to 1) |
|----------|-----------|-----------------------|
| Vehicles with 1–9 seats | \( x_1 \) | 0.364 |
| City GDP | \( x_6 \) | 0.269 |
| City population | \( x_5 \) | 0.119 |
| Vehicles with 40 seats and above | \( x_4 \) | 0.114 |
| Distance between two cities | \( x_7 \) | 0.081 |
| Vehicles with 10–19 seats | \( x_2 \) | 0.042 |
| Vehicles with 20–39 seats | \( x_3 \) | 0.011 |

### Table 2

Results for the risk model on each date.

| Date | \( R^2 \) | \( C \) | \( \lambda_1 \) | \( \lambda_2 \) | \( \lambda_3 \) |
|------|--------|--------|-------------|-------------|-------------|
| 1.24 | 0.823  | 0.002  | 5.293       | 2.277       | 1.061       |
| 1.25 | 0.796  | 1.553  | 1.285       | 0.679       | 0.514       |
| 1.26 | 0.858  | 4.989  | 1.247       | 0.186       | 0.170       |
| 1.27 | 0.903  | 3.153  | 1.339       | 0.087       | 0.206       |
| 1.28 | 0.914  | 3.240  | 1.437       | 0.114       | 0.214       |
| 1.29 | 0.909  | 3.425  | 1.562       | 0.169       | 0.193       |
| 1.30 | 0.935  | 3.940  | 1.495       | 0.041       | 0.349       |
| 1.31 | 0.932  | 4.705  | 1.374       | 0.066       | 0.357       |
| 2.1  | 0.918  | 6.553  | 1.362       | 0.124       | 0.327       |
| 2.2  | 0.933  | 13.905 | 1.344       | 0.218       | 0.368       |
| 2.3  | 0.972  | 6.268  | 1.349       | 0.202       | 0.240       |
| 2.4  | 0.934  | 5.691  | 1.539       | 0.175       | 0.264       |
| 2.5  | 0.928  | 8.289  | 1.802       | 0.095       | 0.312       |
| 2.6  | 0.927  | 4.062  | 1.879       | 0.055       | 0.332       |
| 2.7  | 0.926  | 7.934  | 1.879       | 0.055       | 0.333       |
| 2.8  | 0.928  | 3.892  | 1.927       | 0.027       | 0.351       |
| 2.9  | 0.928  | 4.495  | 1.882       | 0.011       | 0.379       |
| 2.10 | 0.926  | 5.969  | 1.894       | 0.004       | 0.364       |
| 2.11 | 0.925  | 6.949  | 1.911       | –0.010      | 0.370       |
| 2.12 | 0.924  | 4.209  | 1.932       | –0.027      | 0.371       |
| 2.13 | 0.929  | 8.198  | 1.828       | 0.005       | 0.290       |
| 2.14 | 0.946  | 4.944  | 1.847       | –0.005      | 0.264       |
| 2.15 | 0.949  | 8.887  | 1.868       | –0.029      | 0.265       |
| 2.16 | 0.949  | 4.030  | 1.917       | –0.049      | 0.281       |
| 2.17 | 0.950  | 3.683  | 1.969       | –0.069      | 0.296       |
| 2.18 | 0.950  | 5.187  | 1.969       | –0.087      | 0.305       |
| 2.19 | 0.948  | 5.368  | 2.008       | –0.088      | 0.312       |
| 2.20 | 0.944  | 3.332  | 2.063       | –0.092      | 0.327       |
\begin{table}
\centering
\caption{Results for the improved risk model on each date.}
\begin{tabular}{ccccccc}
\hline
Date & \(R^2\) & \(C\) & \(\beta_1\) & \(\beta_2\) & \(\beta_3\) & \(\beta_4\) \\
\hline
1.24 & 0.715 & 0.116 & 0.744 & 0.914 & 0.977 & -0.853 \\
1.25 & 0.753 & 0.628 & 0.928 & 0.661 & 0.604 & -0.506 \\
1.26 & 0.754 & 1.332 & 1.141 & 0.339 & 0.713 & -0.292 \\
1.27 & 0.799 & 2.161 & 1.177 & 0.193 & 0.702 & -0.275 \\
1.28 & 0.854 & 2.565 & 1.292 & 0.129 & 0.794 & -0.272 \\
1.29 & 0.878 & 3.010 & 1.313 & 0.158 & 0.679 & -0.316 \\
1.30 & 0.893 & 3.906 & 1.502 & 0.000 & 0.714 & -0.191 \\
1.31 & 0.887 & 4.953 & 1.525 & -0.016 & 0.704 & -0.165 \\
2.1 & 0.880 & 5.386 & 1.550 & 0.014 & 0.703 & -0.189 \\
2.2 & 0.896 & 5.794 & 1.637 & 0.058 & 0.646 & -0.196 \\
2.3 & 0.893 & 7.147 & 1.665 & 0.048 & 0.591 & -0.122 \\
2.4 & 0.906 & 6.738 & 1.717 & 0.055 & 0.618 & -0.152 \\
2.5 & 0.908 & 6.472 & 1.776 & 0.029 & 0.657 & -0.167 \\
2.6 & 0.906 & 6.905 & 1.797 & 0.006 & 0.659 & -0.162 \\
2.7 & 0.906 & 6.933 & 1.795 & 0.006 & 0.661 & -0.162 \\
2.8 & 0.910 & 7.404 & 1.785 & -0.007 & 0.668 & -0.175 \\
2.9 & 0.911 & 7.874 & 1.800 & -0.032 & 0.685 & -0.165 \\
2.10 & 0.912 & 8.546 & 1.768 & -0.030 & 0.677 & -0.180 \\
2.11 & 0.912 & 8.864 & 1.764 & -0.038 & 0.681 & -0.184 \\
2.12 & 0.912 & 9.325 & 1.754 & -0.047 & 0.681 & -0.189 \\
2.13 & 0.922 & 10.758 & 1.792 & -0.051 & 0.581 & -0.144 \\
2.14 & 0.920 & 11.546 & 1.769 & -0.051 & 0.557 & -0.156 \\
2.15 & 0.934 & 12.068 & 1.765 & -0.065 & 0.548 & -0.155 \\
2.16 & 0.934 & 11.990 & 1.772 & -0.074 & 0.556 & -0.157 \\
2.17 & 0.935 & 11.801 & 1.781 & -0.083 & 0.564 & -0.160 \\
2.18 & 0.936 & 11.896 & 1.785 & -0.095 & 0.565 & -0.160 \\
2.19 & 0.935 & 11.717 & 1.787 & -0.093 & 0.574 & -0.163 \\
2.20 & 0.932 & 11.344 & 1.768 & -0.084 & 0.601 & -0.184 \\
\hline
\end{tabular}
\end{table}

\[ x_{\text{ij}} = \sum_{j=1}^{k} w_{ij} R_j \]  

(2)

The values of \( \beta_i \) in the risk model change with different daily input data. It may be subject to one kind of distribution or conform to some growth curves. However, we can’t find a suitable distribution or curve to describe it. And it is also not mentioned in the article by Jia et al. (2020). So the fixed effects are not interpretable. To solve this problem, we remove the fixed effect item and add one factor, namely, distance into the model. Hence, \( x_{\text{ij}} \) is the distance from the administrative area \( i \) to Wuhan. To smooth the variables and make their fluctuations stable, while eliminating differences in orders of magnitude and dimension among the multiple regression variables to make them comparable, the z-score standardization is performed after taking the natural logarithm. So we get \( x_{ij}' = (\ln x_{ij} - \text{Mean}) / \text{Std} \) and the improved risk model in Equation (3).

\[ y_i = c \prod_{j=1}^{k} e^{R_j x_{ij}'} \]  

(3)

4.2. Parameter estimation and results

We fit the risk models (improved model and non-improved model) from January 24 to February 20. Table 2 shows the result of estimated parameters acquired by the risk model using the highway data, while the result of the improved risk model is shown in Table 3. The values of fixed effects in the risk model are displayed in Appendix (see Table A1).

In the early stage after the lockdown of Wuhan, the virus detection and statistics on the number of confirmed cases are not timely and accurate. So the number of confirmed cases reported in most cities at the beginning is less or even zero, which is different from the reality, leading to the fluctuation of parameters in Equations (1) and (3) unstable and poorly explanatory. With the passage of time, confirmed cases in regions gradually appear. According to Tables 2 and 3, among all parameters, the weight of the vehicle outflow from Wuhan, \( \beta_1 \) increases with time and is greater than that of the others. The weights of population and GDP gradually decrease with time, while the weight of distance is always negative.

We could also find that the \( R^2 \) fitted by the risk model and the improved risk model increased from 0.7 to 0.8 in the initial stage to 0.944 and 0.932 on February 20, respectively. Although no fixed effects are considered, the improved risk model still has a good fitting performance. In addition, the results show that parameter \( C \) in the improved risk model increases with time, which can reflect the changing trend of confirmed cases over time to a certain extent. However, the parameter \( C \) in the risk model increases quickly at first (From January 27 to February
2, it changed from 3.153 to 13.905, but fluctuates around 6 later. The parameter $C$ of the improved risk model changes more smoothly and is closer to the growth curve of confirmed cases (see Fig. 6). So the risk model is not suitable for our problem with the highway toll system data. The value of parameter $C$ in the improved risk model can improve the fitting accuracy of the growth curve.

The parameter $C$ estimated by the improved risk model approximately shows a steady upward trend, which is quite like the growth curve that describes population growth. The most common growth curves, Logistic curve (Pearl and Reed, 1920), Gompertz curve (Gompertz, 1825), and Richards curve (Richards, 1959) are shown in Equations (4)–(6):

$$y = \frac{\alpha_1}{1 + ye^{-\beta t}}$$  \hspace{1cm} (4)

$$y = \alpha_2 \cdot b^t$$  \hspace{1cm} (5)

$$y = A(1 - b_2 e^{-kt})$$  \hspace{1cm} (6)

where $\alpha_1$, $\beta$, $\alpha_2$, $b_1$, $A$, $k$, $b_2$, and $m$ are parameters to be estimated. We use these three growth curves to fit the variation of the parameter $C$ with the LM algorithm. During the fitting process, it is discovered that on

![Fig. 7. The fitting result of the parameter C of each province and region.](image-url)
January 24, the number of confirmed cases in most provinces is small or even zero (still due to the incubation period of the COVID-19 virus, insufficient case detection capabilities, and statistical problems with confirmed case data). It has a certain influence on the fitting effect and the accuracy of parameters, so we use the data of confirmed cases from January 25 to February 20, 2020 for fitting. The $R^2$ of fitting Logistic curve, Gompertz curve and Richards curve are 0.958, 0.963 and 0.969 respectively. The estimated values of parameters in Equations (4)–(6) are given in Appendix (see Table A2).

The fitting results of the three curves are all good, indicating that...
they can represent the growth characteristics of the COVID-19 driven by the four elements as internal factors. When the factors are changed (such as using other data of population movement or adding other factors), the growth curve fitted by parameter $C$ will also change accordingly.

4.3. Results for each province and region

The parameters on each date in the improved risk model of provinces and regions are also estimated. Limited by the travel distance of vehicles, the closer the distance between Wuhan and city $i$, the more traffic flows. For simplifying the calculation, the four municipalities, Beijing, Tianjin, Chongqing, and Shanghai, are grouped into neighboring provinces. Eventually, the objects of the study included three regions, Beijing, Tianjin, and Hebei (referred to as BTH), Sichuan and Chongqing (referred to as SC), Jiangsu and Shanghai (referred to as JS), Hubei, Shandong, Shanxi, Henan, Guangdong, Guangxi, Anhui, Fujian, and Jiangxi, a total of ten provinces. In the fitting results, the value of $R^2$ in five provinces, Anhui, Shandong, Shanxi, Guangxi, and Fujian, are above 0.6, and the remaining eight provinces and regions exceeded 0.8 and 0.9 respectively.

We also find that the parameter $C$ in the improved risk model of each province and region increased steadily over time. The growth curves fitted with them are shown in Fig. 7, and the $R^2$ of fitting three growth curves are shown in Table 4.

Different from the results obtained from a national-wide perspective, the results of provinces and regions show different conclusions. Hubei, Henan, Guangxi, Fujian, and Shanxi have the largest weight $\beta_1$. The vehicle outflow from Wuhan is the key factor in the early stage in Hunan, while GDP becomes the main reason later. In Anhui, BTH, SC, and Shandong, the population is the primary element. And in Guangdong, Jiangxi, and JS, GDP is the crucial cause. These results indicate that although most provinces and regions have a relatively synchronized outbreak period, their spreads of pandemics are caused by different key factors.

4.4. Dynamic exponential risk model

The improved risk model can’t describe the changing trend of confirmed cases over time. In the fitting process, it is found that the change of parameter $C$ is very consistent with the growth curve, which is a function of time. Therefore, we consider replacing parameter $C$ in Equation (3) with the growth curve. However, we also find that using a Logistic curve to replace parameter $C$ can get a better fitting performance than the Gompertz curve and Richards curve. Therefore, the Logistic curve is selected to construct the dynamic exponential risk model, which is shown in Equation (7).

$$y_i = \frac{a}{1 + e^{\omega - \beta_j x_i}} \prod_{j=1}^{4} e^\delta i_j$$

The LM algorithm is still used for parameter estimation, and we obtain the dynamic exponential risk model of China and part of its provinces and regions as above. It can describe the growth law of the pandemic over time when incorporating the cases into a time-related model, and the fitting result is shown in Fig. 8 (a) and Fig. 9. The values of $R^2$ of the dynamic exponential risk model fitted to China and 13 provinces and regions are all above 0.9. That is to say, the model can capture the change of the epidemic over time. The Logistic curve reflects the change features of the epidemic in each region. The Logistic curve of Hubei province is significantly higher than that of other provinces and regions, so we compare the Logistic curves of other provinces and regions except for Hubei province. In order to observe the change of the Logistic curves more intuitively, the derivative of the curve is taken as the growth rate, and the result is shown in Fig. 8 (b).

Fig. 8 shows that the growth rate of all provinces and regions occurred around February 2. Except for Hubei province, the five provinces and regions with the highest growth rate are Jiangxi, Hunan, Anhui, and JS. Among them, Jiangxi, Hunan, and Anhui are directly adjacent to Hubei province, and these four provinces receive more vehicle outflows from Hubei (consistent with Fig. 1). The internal factors driving the spread of pandemics in provinces and regions determine the different characteristics of the growth curves.

5. Analysis of the effect of the highway lockdown measure

The lockdown measures imposed in Wuhan on January 23, 2020, has greatly restricted the outflow of the population (including various modes of transportation such as coach services, air, high-speed train, shipping, etc.). Thus, we try to quantify such impacts of highway lockdown measures (implemented in advance or not) on the spread of COVID-19 in this section. After estimating the parameters in the dynamic exponential risk model, we substitute the number of vehicle outflows from Wuhan on January 18 to January 20, January 21, and January 22 as the variable into the model (other variables remain being unchanged).

After the z-score standardization process is performed on variables, the influence of the absolute changes in the value of variables on the last result is weakened. Therefore, the variables are only taken the natural logarithm in this step. In the multiple regression, the difference in the magnitude of variables makes the parameters no longer able to explain the importance of several variables, but the model still had a good fitting performance. As shown in Fig. 10, the impact of taking the highway lockdown measure in advance on the number of confirmed cases of China is analyzed first.

As can be seen from Fig. 10, if the highway lockdown measure is implemented one day in advance, the number of confirmed cases nationwide (except Wuhan) will drop to a certain extent. Implementing three days in advance (that is, the vehicle outflow from Wuhan ends on
January 20) makes the number drop by about 50% as of February 20. It can be inferred that if the measure is postponed or even not adopted, the total number of confirmed cases will be far more than the current number. That is, the highway lockdown measure taken in Wuhan has played an important role in suppressing the spread of COVID-19 (which successfully controls the epidemic by restricting the vehicle outflow). Based on the previous analysis, we conjecture that the implementing of the highway lockdown measure in advance has a more obvious effect on the provinces and regions where the number of vehicle outflows from Wuhan is the key factor to confirmed cases.

For provinces and regions perspectives, the weight of the vehicle outflow is 1.24 in Guangxi, 1.18 in Henan, and 1.05 in Shanxi, which are much higher than those of other factors. The spread of COVID-19 in these provinces is mainly affected by the vehicle outflow from Wuhan. Implementing the highway lockdown measure three days in advance can reduce confirmed cases in Henan and Shanxi by about 50%, and Guangxi even by about 60% (as shown in Fig. 11). We also find the top three weights of the population are BTH (1.11), SC (0.96), and Anhui (0.55). We can infer that population is the key factor in these three provinces. In Guangdong, Jiangxi, and JS, we get the top three weights of GDP, 1.24, 0.56, and 0.53, respectively. So, the pandemic is mainly affected by GDP in these provinces. If highway lockdown the measure is taken three days earlier, the number of cases in Hubei will be reduced by 20%. This may be because that a certain number of vehicle outflows from Wuhan have already flowed into other cities in Hubei province before January 18, which caused these cities to receive the infected individuals earlier. In addition, due to the lack of knowledge about the COVID-19 virus in the early days, the preventive measures adopted were not in place, and the frequent personnel activities and contacts intensified the spread of COVID-19 in Hubei province.

The result only represents the change in confirmed cases caused by the highway traffic flow. The provinces and regions, whose spread of COVID-19 with smaller change, are less affected by the flow of the inter-provincial traffic. In fact, it may be concerned with the internal population contacts and the movement under other modes of transportation. Therefore, it is necessary to take the targeted prevention and control measures to achieve the best epidemic prevention effect, according to the actual situation and the key factor deciding the spread of pandemics in each province and region. In addition, the confirmed cases in China (except Wuhan) and Hubei province (except Wuhan) show a similar trend. This is because about 60% of the cases in China are from Hubei province.

6. Policy suggestions

Based on the above model and analysis, several policy suggestions are given as follows.

For provinces that are more affected by the vehicle outflow from Wuhan, it is necessary to focus on the investigation and limitation of outsiders, especially those from the affected areas, to cut the source of infection. At the same time, according to the result of the correlation analysis, we should focus on vehicles with 1–9 seats and above 40 seats in the highway inspection process, because they have the highest correlation with confirmed cases.

In some provinces and regions, the key factor is population. Such areas may not be aware of the severity of the epidemic in the early stage, so they don’t take the preventive measures in time or implement the
home quarantine inadequately. Therefore, they should pay more attention to taking corresponding measures as soon as possible, and abide by the requirements of the epidemic protection strictly. For example, one can reduce the movement of population and crowd gathering, strengthen personal protection, and keep social distance to prevent close contact between the susceptible and infected individuals.

In addition, some provinces and regions are mainly affected by GDP. These areas are usually economically developed areas, such as Guangdong, Jiangsu and Shanghai. The more developed, the more foreign population and the higher frequency of social activities. In addition to controlling the domestic population movements, these cities should also screen people imported from abroad strictly. Moreover, the company should try to let employees work from home and make fewer trips outside. At the same time, public places (such as hotels, restaurants, and supermarkets) should be disinfected in time to cut off the transmission routes as much as possible.

Finally, the provinces and regions that are close to the epidemic area are severely affected and have more confirmed cases relatively. Therefore, they should maintain a higher degree of vigilance. Except for ensuring their protection against the epidemic, they should also control the outflow of the local population to prevent virus carriers from flowing into other provinces or regions, causing the second spread of pandemics.

7. Discussion and conclusion

The struggle between humans and pandemics never stops. Unlike the past, humans today can use many types of data to study the spread of pandemics from multiple angles, so as to respond to the next outbreak effectively. In this study, we proposed a data-driven method to analyze the trend of the spread of COVID-19 under the influence of intercity vehicle mobility. Based on more than 100 million pieces of data from the Chinese highway toll system, the traffic flow characterized by different vehicle types out from the Wuhan city was extracted. Moreover, we used the Pearson’s correlation coefficient test and the random forest method to check the correlation between the vehicle outflow (in various vehicle types), city population, city GDP, and distance between two cities with the cumulative confirmed cases in China cities. The result showed that the Pearson’s correlation coefficient of the vehicle outflow increased over time and exceeded 0.9 after February 2, 2020. But, the coefficients of the other factors were up to 0.5, which were much lower than the vehicle outflow. This indicated that the vehicle outflow from Wuhan was the main factor affecting the spread of COVID-19 from a national perspective. At the same time, the random forest model showed that the importance of the vehicles (1–9 seats) was 0.364, holding the closest relationship with the spread of COVID-19 among the other four vehicle types.

Then, we improved the risk model proposed by Jia et al. (2020), to aggregate different vehicle types into a unified vehicle flow and make the risk model more interpretable. The confirmed cases in some provinces and regions of China were fitted. The result showed that the improved model had a good fitting effect, with the $R^2$ equaling to 0.932. The value of parameter $C$ in the new risk model was close to the increasing trend of confirmed cases. We further calculated the value of parameter $C$ for each province and fitted them with three growth curves. We found that the goodness of fit is higher than 0.9. Then, to reflect the changing law of the COVID-19 over time, we established a dynamic exponential risk model.

The result of fitting the dynamic model with the daily confirmed cases showed that for most cities, the goodness of fit was higher than 0.97, while the goodness of fit in BTH, JS, and Fujian was higher than 0.95. It indicated that the dynamic exponential risk model could describe the spread of COVID-19 well in various provinces at the time dimension. Based on this dynamic model, we also analyzed the effect of the highway lockdown measure on the spread of COVID-19. Results showed that Guangxi was most affected by the vehicle outflow from Wuhan because implementing the highway lockdown measure 3 days in advance would reduce the number of confirmed cases to one-third. The same measure could reduce the number of confirmed cases in Henan and Shanxi by 50%. For Fujian, Jiangxi, Guangdong, Hunan, and the cities except Wuhan in Hubei, the same measure could only reduce the number of confirmed cases by 20% approximately. According to the fitted parameter values, we found that in addition to the vehicle outflow, other factors also played different roles in the spread of COVID-19 in different provinces. At last, we divided these provinces into different categories based on the characteristics of parameters and put forward corresponding policy suggestions for each category. The policy suggestions could not only provide the government departments with a basis for formulating policies of COVID-19 prevention and control but also accumulate experience to deal with similar pandemics.

Although we did some work as above, some improvements can be recognized in further research. First, we can use travel data of the railway, aviation, and other transportation modes for comparative analysis, and explore the impact of population movement under multiple transportation modes on the spread of COVID-19. Furthermore, we can distinguish the type of confirmed cases, that is, whether the confirmed cases are imported cases or secondary-generation cases. We will explore the relationship between population movement and the above two types of confirmed cases in the next step.

Author statement

Xin Meng: Conceptualization, Methodology, Validation, Formal analysis, Writing - Review & Editing. Mingxue Guo: Methodology, Software, Validation, Writing - Original Draft. Ziyou Gao: Term, Conceptualization, Funding acquisition. Zhenzhen Yang: Investigation, Resources. Liujiang Kang: Writing - Review & Editing, Supervision. Zhiwu Yuan: Visualization, Funding acquisition.

Declaration of competing interest

The authors state that there is no potential conflict of interest among the authors.

Acknowledgements

This work is supported by the Fundamental Research Funds for the Central Universities (No. 2021PT206) and the National Natural Science Foundation of China (72001017, 72091513, and 71901147).
Table A.1

The values of fixed effects with the estimation of the risk model.

| Date | $a_{1}$ | $a_{2}$ | $a_{3}$ | $a_{4}$ | $a_{5}$ | $a_{6}$ | $a_{7}$ | $a_{8}$ | $a_{9}$ | $a_{10}$ | $a_{11}$ | $a_{12}$ | $a_{13}$ | $a_{14}$ | $a_{15}$ | $a_{16}$ | $a_{17}$ | $a_{18}$ | $a_{19}$ | $a_{20}$ | $a_{21}$ | $a_{22}$ | $a_{23}$ | $a_{24}$ | $a_{25}$ | $a_{26}$ | $a_{27}$ | $a_{28}$ | $a_{29}$ | $a_{30}$ | $a_{31}$ | $a_{32}$ | $a_{33}$ | $a_{34}$ | $a_{35}$ | $a_{36}$ | $a_{37}$ | $a_{38}$ | $a_{39}$ | $a_{40}$ |
|------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 1.24 | -22.46  | -18.24  | -18.43  | -20.59  | -21.49  | -14.53  | -21.13  | -22.61  | -25.27  | 8.58    | -26.85  | -23.09  | -22.70  | 12.18   | -17.62  | -8.16   | -22.59  | 15.88   | -26.72  | -28.09  | -21.55  | -29.83  | -17.80  | -26.60  | -14.68  | -2.17   | -1.27   | -0.73   | -1.82   | -1.76   | -0.34   | -1.23   | -0.24   | -0.76   |
| 1.25 | 1.20    | 5.58    | 2.77    | 0.12    | 0.73    | 3.11    | 0.24    | 0.78    | 1.22    | 3.98    | 1.92    | 0.00    | 1.87    | 1.45    | 0.03    | 1.99    | 1.80    | 1.40    | 1.43    | 8.36    | 7.76    | 0.92    | 3.95    | -0.11   | 3.47    | -7.51   | -3.17   | -1.55   | -0.73   | -1.82   | -1.76   | -0.34   | -1.23   | -0.24   | -0.76   |
### Table A.2
The estimated values of parameters in Equations (4)–(6).

| Region | Logistic $\alpha_1 \gamma \omega$ | Gompertz $\alpha_2 a b_1$ | Richards $A b_2 k m$ |
|--------|----------------------------------|--------------------------|----------------------|
| China  | 13.11 6.89 0.16                 | 14.40 0.08 0.90         | 33.14 1.01 0.01 0.80 |
| JS     | 50.11 12.96 0.23                | 54.14 0.03 0.87         | 58.75 0.90 0.10 2.27 |
| Fujian | 29.77 8.21 0.27                 | 30.69 0.05 0.83         | 32.76 1.07 0.11 1.25 |
| SC     | 18.30 9.05 0.26                 | 19.01 0.05 0.84         | 20.11 0.99 0.12 1.61 |
| Shandong| 32.99 9.67 0.23                | 35.10 0.05 0.87         | 35.11 0.00 0.14 1355.66 |
| BTH    | 28.94 34.75 0.23               | 33.87 0.01 0.88         | 35.43 0.57 0.11 7.24 |
| Jiangxi| 76.01 20.33 0.29               | 79.98 0.01 0.85         | 79.98 0.00 0.18 2076.60 |
| Guangdong| 26.32 8.94 0.18            | 29.59 0.06 0.90         | 29.59 0.00 0.11 1943.76 |
| Guangxi| 11.03 94.93 0.41              | 11.36 0.00 0.77         | 11.36 0.00 0.26 6556.84 |
| Anhui  | 49.66 17.22 0.27               | 52.85 0.02 0.85         | 52.86 0.00 0.16 5985.52 |
| Shanxi | 8.66 47.93 0.37                | 8.83 0.00 0.77         | 8.83 0.00 0.26 2960.85 |
| Hunan  | 46.67 20.47 0.31               | 48.75 0.01 0.82         | 48.75 0.00 0.20 5525.97 |
| Hupeh  | 1108.46 17.04 0.24             | 1199.56 0.02 0.87      | 1241.16 0.65 0.12 4.72 |

### References

- Gompertz, B., 1825. On the nature of the function expressive of the law of human mortality and on a new mode of determining life contingencies. Phil. Trans. Roy. Soc. Lond. 115, 513–585.
- Hu, S., Xiong, C., Yang, M., Younes, H., Luo, W., Zhang, L., 2021. A big-data driven approach to analyzing and modeling human mobility trend under non-pharmaceutical interventions during COVID-19 pandemic. Transport. Res. C Emerg. Technol. 124 (2021), 1–17.
- Jia, J., Lu, X., Yuan, Y., Xu, G., Jia, J., Christakis, N., 2020. Population flow drives spatio-temporal distribution of COVID-19 in China. Nature 582, 389–394.
- Kraemer, M.U.G., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D.M., du Plessis, L., Farra, N.R., Li, R., Hanage, W.P., Brownstein, J.S., Layton, M., Vespignani, A., Tian, H., Dye, C., Pybus, O.G., Scarpino, S.V., 2020. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science 368 (6490), 1–5.
- Lai, S., Ruktanonchai, N.W., Zhou, L., Prosper, O., Luo, W., Floyd, J.R., Wesolowski, A., Santillana, M., Zhang, C., Du, X., Yu, H., Tatem, A.J., 2020. Effect of non-pharmaceutical interventions to contain COVID-19 in China. Nature 585 (7825), 410–413.
- Li, T., Rong, L., Zhang, A., 2021. Assessing regional risk of COVID-19 infection from Wuhan via high-speed rail. Transport Pol. 106, 226–238.
- Nouvellet, P., Bhatia, S., Cori, A., Ainslie, K.E.C., Baguelin, M., Bhatt, S., Boonyasiri, A., Brazeau, N.F., Cattarino, L., Cooper, L.V., Coupland, H., Cucunuba, Z.M., Cuomo-Dannenburg, G., Dighe, A., Djadara, B.A., Dorigatti, I., Eales, O.D., van Elsland, S.L., Nascimiento, F.F., FitzJohn, R.G., Gaythorpe, R.A.M., Geidelberg, L., Green, W.D., Hamlet, A., Hauck, K., Hinsley, W., Imai, N., Jeffrey, B., Knock, E., Laydon, D.J., Lees, J.A., Mangal, T., Mellan, T.A., Nedjati-Gilani, G., Parag, K.V., Pons-Salort, M., Ragonnet-Cronin, M., Riley, S., Unwin, H.J.T., Verity, R., Vollmer, M.A.C., Volz, E., Walker, P.G.T., Walters, C.E., Wang, H., Watson, O.J., Whittaker, C., Whittle, L.K., Xi, X., Ferguson, N.M., Donnelly, C.A., 2021. Reduction in mobility and COVID-19 transmission. Nat. Commun. 2021 (1), 1–9.
- Pearl, R., Reed, L.J., 1920. On the rate of growth of the population of the United States since 1790 and its mathematical representation. Proc. Natl. Acad. Sci. Unit. States Am. 6 (6), 275–288.
- Peng, X., Xu, X., Li, Y., Cheng, L., Zhou, X., Ren, B., 2020. Transmission routes of 2019-nCoV and controls in dental practice. Int. J. Oral Sci. 12 (9), 1–6.
- WHO, 2021. Coronavirus Disease (COVID- 2019) Pandemic. World Health Organization. Retrieved from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019.
- Yan, X.-Y., Zhang, A., Wang, X.-W., Gao, Z.-Y., Lai, Y.-C., 2017. Universal model of individual and population mobility on diverse spatial scales. Nat. Commun. 8 (1), 1–9.
- Yang, Z., Zeng, Z., Wang, K., Song, S., Li, L., An, M., Liu, F., Pei, C., Bai, X., Gao, Z., Mai, Z., Jiang, J., Liu, X., Li, S., Li, Y., Ye, F., Guan, W., Yang, Y., Li, F., Luo, S., Xie, Y., Liu, B., Wang, Z., Zhang, S., Wang, Y., Zhong, N., He, J., 2020. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. J. Thorac. Dis. 12 (3), 165–174.
- Zhang, Y., Zhang, A., Wang, J., 2020. Exploring the roles of high-speed train, air and coach services in the spread of COVID-19 in China. Transport Pol. 94 (4), 34–42.