JOURNAL OF DECISION MAKING AND HEALTHCARE
Volume 1 (2024), No. 1, 16–29
https://doi.org/10.69829/jdmh-024-0101-ta02
Tulipa Opera Scholarum

RESEARCH ON THE PROMOTION AND OPTIMIZATION OF BICYCLE SHARING IN THE POST-COVID-19 EPIDEMIC ECONOMY

RONG WANG[1], KENAN YE[1], YANGYANG YANG[1], XIAOHUI LI[2], HUILIN CHEN[2], AND SHENG-YA FENG[1,*]

[1] *School of Mathematics, East China University of Science and Technology, Shanghai 200237, P.R. China*
[2] *School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, P.R. China*

Abstract. In the post-pandemic era, people's lifestyles have changed accordingly. Bicycle sharing has become a common means of transportation in people's daily lives. In this article, we consider the factors that affect people's choice to travel and analyze the supply and demand relationship of city bicycles. To this end, we use the analytic hierarchy process (AHP) to infer the conclusion that people prefer to use city bicycles as a means of travel during rush hour commuting and short-distance travel. In addition, we further propose a discrete gray model (DGM) to predict the development trend of city bicycles. The results show that the demand for city bicycles is steadily increasing. The research in this article can help governments and organizations plan urban transportation more rationally and maximize the use of urban transportation resources.

Keywords. Bicycle sharing, Analytic Hierarchy Process (AHP), Discrete Gray Model (DGM), COVID-19 epidemic, Economic recovery

1. Introduction

With the continuous improvement of global environmental awareness, more and more people are beginning to pay attention to the impact of transportation on the environment. Green travel can not only reduce carbon emissions, but also reduce air pollution, which is of great benefit to protecting the environment.

1.1. Background.

The development of science and technology has also provided more possibilities for green travel, such as electric vehicles and shared bicycles. In the future, we may see the emergence of more environmentally friendly and intelligent means of transportation. At the same time, the government and enterprises are also increasing their investment and support in green travel, which will further promote its development.

Bicycles have gradually faded out of people's lives in the past, mainly due to the expansion of the city, the increase in people's travel distance, and the safety of storing bicycles. In addition, the popularity of automobiles and electric vehicles has also gradually reduced the use of bicycles. However, with the continuous highlighting of environmental and energy issues, shared bicycles have emerged, allowing bicycles to return to people's lives. The emergence of bicycle sharing has solved the "last kilometer" problem of people's travel, provided a more convenient and environmentally friendly way of travel. It also promotes the development of cycling culture, allowing people to feel the charm of bicycles again.

[*]Corresponding author.
E-mail address: 21012399@mail.ecust.edu.cn (R. Wang), 21012429@mail.ecust.edu.cn (K. Ye), 21012404@mail.ecust.edu.cn (Y. Yang), 19001827@mail.ecust.edu.cn (X. Li), 19003924@mail.ecust.edu.cn (H. Chen), and s.y.feng@ecust.edu.cn (S.-Y. Feng)
Accepted: March 12, 2024.

1.2. **Solutions.** In order to predict people's travel patterns, some scholars have proposed some mathematical models to explain it. Cantelmo et al [3] produced a low-dimensional model to predict the demand pattern for bicycles in New York City by using three-level data clustering. All these predictive models can serve urban planning well and have a positive impact on the development of urban transportation networks. However, some predictive models are only applicable to certain cities. Therefore, how to analyze people's travel patterns from multiple dimensions will become a technical issue that we need to emphasize. In this article, we will construct other mathematical models to optimize our design based on the previous research. First, we used the analytic hierarchy process (AHP) to analyze the use of bicycle sharing in the city. The analysis process mainly considers factors such as weather, the duration of the trip, the distance of the trip, and the expenses that may be required to use the shared bicycle. Secondly, combined with the discrete GM(1,1) gray prediction model, we collect data on the use of shared bicycles in recent years and predict the use in the next five years. After analyzing and evaluating the above results, we will give objective and reasonable optimization suggestions for the shared plan. Finally, considering that the reliability of the results of the model is related to the time span (the longer the time, the greater the error), we propose an optimization method for building a model based on exponential growth series.

2. Literature Review

As a popular, healthy and environmentally friendly mode of transportation, city bicycles have become an important part of the transportation network. The establishment of an effective urban bicycle management system will be of great significance to improve the urban environment and improve the quality of life of the people. At present, many scholars have used different methods to explore and study the parking location, balance usage and operating costs of bicycle sharing.

García-Palomares et al [12] studied the characteristics and utility of shared bicycle stations using GIS as well as location-allocation models to determine the optimal station locations for shared bicycles, and their results showed that the increase in the number of stations did not improve the accessibility and coverage very much. In fact, using minimum impedance and maximum coverage, they found that with the increase of sites, the growth rate of coverage demand and the increased rate of reachability decrease, and the law of diminishing returns. The study also showed that central sites as well as high-attraction stations have higher levels of accessibility (areas that attract more trips than they generate, with the proportion of generator and hybrid stations increasing as the number of sites increases).

In order to study the impact of bicycle sharing on the use of motor vehicles, Fishman et al [9] conducted a survey and data analysis related to bicycle sharing programs in Melbourne and other areas. The results show that when considering switching from cars to shared bicycles, the latter has great potential in reducing the use of motor vehicles. On the one hand, the increase in the popularity of shared bicycles will greatly reduce the use of motor vehicles, while on the other hand, the impact of shared bicycles on the use of motor vehicles is largely determined by the replacement rate of automobiles.

Inspired by the operation of shared bicycles, and taking the truck rebalancing (move and take away) as an example and motivation, Chemla et al [4] used the slitting algorithm to solve the relaxation problem of the many-to-many pickup and delivery problem. The vertices in the model can be accessed multiple times, and they use forbidden search to calculate the upper bound of the optimal value. When the vehicle capacity is constant and the average number of bicycles increases, the optimal solution will visit the apex multiple times. If the capacity is smaller, the more difficult the problem will be to solve.

Dell'amico et al [6] also discussed the static (night) rebalancing of bicycle sharing, using integer linear programming to reduce fleet operating costs, and combining it with a branch-and-cut algorithm to solve the problem. Based on the computational results, it is also realized that the computational difficulty will increase with the increase in the number of vertices and the decrease in the vehicle capacity, which is in line with the performance of the results of Chemla et al. and also with the expectations.

Forma et al [11] proposed a three-step heuristic algorithm to solve the static relocation (rebalancing) problem, introducing the second method of clustering preferred routes, which aggregates the set of points that may be accessed consecutively, reducing the size of the routing problem. Stations are clustered using geography and inventory, routing vehicles centrally and taking into account temporary inventory decisions made by vehicles for each station, imposing restrictions on vehicle traversals, and creating mixed-integer linear programming to solve the relocation problem. Unlike others, the number of bicycles required at a station constitutes a soft constraint.

Pal and Zhang [15], contributed to the mathematical planning formulation by optimizing a solution to handle the static rebalancing problems of Free Floating Bike Share (FFBS) and Station Based Bike Share Vehicle (SBBS) through a hybrid algorithm of nested large neighborhood search and variable neighborhood descent. They proposed a simple and effective solution that enables the existing formulas for the traveling salesman problem and the vehicle route problem to be used to formulate the complete rebalancing problem of bicycles and multi-vehicles. They also proposed a hybrid nested large neighborhood search method based on variable neighborhood descent algorithm (NLNS+VND), which can more effectively solve the static rebalancing problem of FFBS and SBBS, with faster speed and higher accuracy.

Schuijbroek et al. [17] provide a more effective and accurate algorithm for the inventory rebalancing problem and vehicle route planning problem in the bicycle sharing system. The authors introduced a new method for approximating the cost of the maximum span star routing of polynomial size. This method is based on the clustering problem of mixed integer planning with service-level feasibility constraints to achieve rapid synchronization. The authors propose an improvement scheme based on reduction and elimination to reduce approximation errors, and give a constrained planning formula for the rebalancing problem of bicycle sharing. Algorithmic analysis shows that when the number of trucks transporting shared bicycles is sufficient to allocate the workload, the hybrid integer program of the clustering problem can better solve the inventory rebalancing problem and vehicle route planning problem in the shared bicycle system. When the number of trucks is small and the transportation route is long, the constrained planning (CP) formulation is more conducive to achieving the problem solution.

Caggiani et al. [1] developed a decision support system for dynamic free-floating bicycle redistribution to solve the imbalance of bicycles between regions caused by one-way travel. They concluded that the bicycle sharing system can be enhanced through a dynamic migration process. The high degree of aggregation of shared bicycle areas is conducive to the management effect of operating companies. Operators can redeploy the necessary number of bicycles according to the maximum value of bicycles available within a given interval, thereby reducing costs. Based on these results, they summarized the percentage of improvement in unfavorable time, lost users, and satisfied users. In most of the literature studying BSS-related activity patterns, the comprehensive method proposed by the authors considers the high dynamics of BSS, unpredictable fluctuations in bicycle trends, and changes in user choices, thereby optimizing the shortcomings of spatiotemporal clustering. With a wide range of application possibilities, setting all the relevant parameters based on a case study allows to find a good application for different real-world solutions.

Fishma et al [10] used probability tests and logistic regression models to analyze the factors affecting people's use of bike sharing and the degree of influence of various factors for two cities, Melbourne and Brisbane. The results showed that convenience, age, safety awareness, the travel patterns of people around them, and the government's mandatory helmet-wearing policy were the main factors influencing people to become bike-sharing members, with convenience being an important predictor of the number of bike-sharing members in a given area, and the distance between bike-sharing docking stations and destinations becoming an important predictor. Differences in safety perceptions between bike-sharing members and non-members provide insights for bicycle infrastructure planners and bike-sharing companies, where bike-sharing non-members feel less safe traveling by bicycle, and therefore the number of bike-sharing users can be increased by, for example, expanding the bicycle infrastructure network.

Campbell et al [2] investigated the factors affecting the choice of shared bikes and the interaction between the shared bike system and the existing transportation system in Beijing by building a multinomial logit model (MNL) and combining it with a SP pivot design. The study shows that the choice of shared bikes is not only influenced by travel attributes and environmental conditions but also by traditional travel behaviors and user demographic data. Weather and air quality are the most significant factors influencing bike-sharing choices, while traffic congestion and license plate restrictions have a lesser impact on bike-sharing choices. To more accurately predict the impact of these factors on demand for bike sharing, further research is needed to quantify and validate the relationship between public perceptions of temperature, precipitation, air quality, and objective measures.

El-Assi et al [7] conducted regression analyses at three different levels of trip generation, trip attractiveness, and trip distance with distributed lag models, multilevel/linear mixed-effects models (random intercepts and slopes), and utilized year-round historical travel data from Toronto to analyze the factors affecting bike-share ridership in Toronto. The study shows that roadway network configuration and bicycle infrastructure have a significant impact on bike share demand and that there is a significant correlation between temperature, land use, and bike share travel activity. Higher temperatures, lower humidity levels, and lower snowpack on the ground were positively correlated with bike-share ridership, temperature, station capacity, regional population, and employment density, and transportation near stations were positively correlated with bike-share ridership, and bicycle distance was negatively correlated with bike-share ridership.

Lin et al [13] proposed a novel graph convolutional neural network (GCNN) model with a data-driven graph filter (DDGF) that can capture hidden heterogeneous pairwise correlations between sites to predict hourly demand in a large bike-sharing network. They found that the GCNN-DDGF structure implements a recurrent block based on the Long Short-Term Memory (LSTM) neural network architecture to learn the temporal dependence in the bike-sharing demand system, in addition to the heterogeneous pairwise correlations that are automatically acquired to improve the prediction capability. In the model comparison, the authors revealed hidden heterogeneous pairwise correlations between stations not found in any other matrix. The development and use of the GCNN model provide a new approach for future research on situations where bike-sharing demand prediction models need to take into account more variables (e.g., social events), and other transportation problems that can be represented graphically.

To fill a gap in the public's knowledge of the rapidly emerging dockless electric scooter industry, McKenzie et al [14] identified similarities and differences between the new scooters and bicycle-shared service platforms by comparing their spatial and temporal activity patterns. The results show that recreational shared bikes and new shared scooter sharing have similarities in time, but there are major differences in spatial distribution. On the other hand, membership shared bikes are generally used differently from shared scooters in these two areas, so it can be concluded that the two shares serve activities for which people have different purposes. The authors analyzed the majority of audiences and deficiencies of the new scooter service by comparing it with the more mature bike-sharing system, efficiently providing practical guidance and suggestions for the dockless electric scooter industry. Besides, the analysis of the data identifies the bias factors in some cities due to the lack of statistical information on the demographics of the transportation service and the room for improvement in the resolution of the data collection, which is of a strong objective nature.

Through an in-depth discussion of the various factors affecting bicycle-sharing, Eren and Uz [8] aim to analyze how these factors affect the demand for bicycle-sharing trips and to identify the most common factors influencing the demand for trips. They concluded that poor weather conditions reduce the demand for bike sharing, that dedicated traffic lanes, suitable cycling environments, extensive station stops, and reliable security promote demand for cars, and that males are more likely to use bike sharing than females and the young are more likely to use bike sharing than the elderly. It is unclear what factors have the dominant impact on bike sharing, but the most unfavorable weather condition affecting demand for bike sharing is rain. The authors studied the six major categories of elements covered by most of the bicycle sharing literature, established a framework for comprehensively demonstrating the factors affecting the demand for bicycle sharing, and conducted a comprehensive analysis from multiple perspectives.

3. Proposed method

In order to overcome the limitations of low volume and incomplete collection of statistical data, we use analytic hierarchy process and a discrete GM(1,1) gray model for derivation and prediction, followed by a pure exponential growth series model to intensify accuracy .

3.1. **Analytic Hierarchy Process.**

The Analytic Hierarchy Process (AHP) is a decision analysis method that combines qualitative and quantitative methods to solve complex multi-objective problems. This method combines quantitative analysis with qualitative analysis, and uses the experience of decision makers to judge the relative importance of the criteria for measuring whether the goals can be achieved. It rationally gives the weights of each standard of each decision-making scheme, uses the weights to obtain the order of advantages and disadvantages of each scheme, and is more effectively applied to those problems that are difficult to solve by quantitative methods.

Step 1: The establishment of a hierarchical structure model.

In a hierarchical structure, complex problems are decomposed into constituent elements based on the considered attributes of the problem and their interrelationships to form a hierarchy. The constituent elements in each layer are criteria that control the effect on the relevant elements in the next layer. These layers are divided into three categories. The highest level is the goal layer, where only one element is usually a predetermined goal or a problem to be solved. The middle level is the criteria level, which includes intermediate elements involved in achieving the target. The bottom level is the program level, including various measures, decision options, etc.

Step 2: Structure of judgment matrix.

As the weights between different factors at different levels are only qualitative, it is unconvincing to accept the results. Therefore, Saaty et al [16] proposed a consistent matrix method, in which all factors are not compared, but paired with each other, and relative scales are used to compare factors of different properties to improve accuracy. For example, for a given criterion, options are compared two by two and graded according to their importance, where a_{ij} is the result of the comparison between element i and element j. The matrix formed by the result of a two-by-two comparison is called the judgment matrix. This judgment matrix has the following properties

$$a_i j = \frac{1}{a_j i}.$$

The specific criteria are shown in Table 1.

Step 3: Hierarchical single ordering and its consistency test.

The eigenvector corresponding to the maximum eigenvalue of the judgment matrix is normalized and represented by w (the sum of all components is 1). The elements of w are the ranking weights of the factors at the same level concerning the relative importance of the factors at the previous level. This process is known as hierarchical single ranking. A consistency test is necessary to determine the hierarchical single ordering. Consistency test means determining the permissible range of inconsistency for matrix A, where matrix A is an $n \times n$ matrix consisting of a_{ij} as matrix elements, namely $A = (a_{ij})$. The unique non-zero characteristic root of a consistent $n \times n$ matrix is n and the largest characteristic root of a positive reciprocal matrix A of order n satisfies $\lambda \geq n$. A is a consistent matrix

Table 1. Criteria for the values of elements in the judgment matrix

Scale	Meaning
1	Indicates that two factors are equally important compared to each other.
3	Indicates that the former is slightly more important than the latter.
5	Indicates that the former is significantly more important than the latter.
7	Indicates that the former is far more important than the latter.
9	Indicates that the former is extremely more important than the latter.
2, 4, 6, 8	The median of the above two neighboring judgments.

if and only if $\lambda = n$. Since λ is continuously dependent on a_{ij}, the greater λ is more than n, the more serious the inconsistency of A is. The consistency index is calculated by CI. The smaller CI is, the greater the consistency. Using the eigenvector corresponding to the largest eigenvalue as the weight vector of the degree of influencing the compared factor on the superior factor, the greater the degree of inconsistency, the greater the judgment error. Therefore, the degree of inconsistency of A can be expressed in terms of the magnitude of the value of $\lambda - n$. Coherence indicators are defined as

$$CI = \frac{\lambda - n}{n-1}.$$

$CI = 0$ means that the consistency is perfect, and CI is close to 0, which means that the consistency is satisfactory. The larger the CI, the more serious the inconsistency. In order to measure the magnitude of CI, a random consistency index RI is introduced by

$$RI = \frac{CI_1 + CI_2 + \cdots + CI_n}{n}.$$

The random consistency index RI is related to the order of the judgment matrix. Typically, the greater the order of the matrix, the greater the likelihood of a random deviation in consistency. The corresponding relationship is shown in Figure 1.

The matrix order	1	2	3	4	5	6	7
R.I.	0	0	0.52	0.89	1.12	1.26	1.36
The matrix order	8	9	10	11	12	13	14
R.I.	1.41	1.46	1.49	1.52	1.54	1.56	1.58

Figure 1. Values of the random consistency index RI

Considering that the deviation that affects the consistency of the judgment matrix may be caused by random reasons, it is necessary to compare the consistency index CI with the random consistency index RI. The calculation formula is as follows,

$$CR = \frac{CI}{RI}.$$

If $CR < 0.1$, the judgment matrix is considered to pass the consistency test, otherwise the consistency is unsatisfactory.

3.2. **Discrete GM(1,1) model.** In recent decades, gray system theory has continued to develop. It is widely used in many fields such as economy, education, meteorology, etc. This theory believes that the system is time-varying, and a large amount of data is not essential to obtain good prediction results. Importantly, its prediction results generally have high accuracy. However, in some cases with rich structural characteristics, the possibility of adjustment is not ruled out to make the prediction result ideal and stable.

3.2.1. *Relevant definition.* For the observation value of the behavioral feature sequence of the system

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$$

with $x^{(0)}(k) \geq 0$, $k = 1, 2, \cdots, n$, one defines the first-order *accumulating generation* by

$$X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\right), \tag{3.1}$$

where $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i)$, $k = 1, 2, \cdots, n$.

Moreover, one defines the *mean generation* by

$$Z^{(1)} = \left(x^{(1)}(2), x^{(1)}(3), \cdots, x^{(1)}(n)\right)$$

with $z^{(1)}(k) = \frac{1}{2}\left(x^{(1)}(k-1) + x^{(1)}(k)\right)$, $k = 2, 3, \cdots, n$ and calls that $x^{(0)}(k) + az^{(1)}(k) = b$ the *grey differential equation*, which is also commonly referred to as the GM(1, 1) model.

The first-order ordinary differential equation (ODE)

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

is called a whitenization (or image) equation of the GM(1,1) model. The parameter $-a$ is the development coefficient, reflecting the development trend, and b is the gray role amount. Empirically, GM makes long-term forecasts as $-a \leq 0.3$, short-term forecasts as $0.3 < -a \leq 0.8$. As long as $0.8 < -a \leq 1$, the residual correction GM model works better than the GM model, while $-a > 1$, the GM model will no longer be applicable.

The discrete GM (1, 1) model (DGM(1,1)) follows that

$$x^{(1)}(k+1) = \beta_1 x^{(1)}(k) + \beta_2, \tag{3.2}$$

where the estimation of the parameter vector

$$\hat{\beta} = [\beta_1, \beta_2]^T$$

is determined by the least squares estimation

$$\hat{\beta} = \left(B^T B\right)^{-1} B^T Y$$

with

$$Y = \begin{bmatrix} x^{(1)}(2) \\ x^{(1)}(3) \\ \vdots \\ x^{(1)}(n) \end{bmatrix}, \qquad B = \begin{bmatrix} x^{(1)}(1) & 1 \\ x^{(1)}(2) & 1 \\ \vdots & \vdots \\ x^{(1)}(n-1) & 1 \end{bmatrix}.$$

To end this subsection, we suppose that $x^{(1)}(1) = x^{(0)}(1)$ and make use of Eq. (3.2) to recursively obtain that

$$\hat{x}^{(1)}(k+1) = \beta_1^k \hat{x}^{(1)}(1) + \left(\beta_1^{k-1} + \beta_1^{k-2} + \cdots + \beta_1 + 1\right)\beta_2,$$

where $\hat{x}$ denotes the predicted value of the original x. It finally leads to the time response equation for the discrete GM(1,1) model as

$$\hat{x}^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{\beta_2}{1-\beta_1}\right]\beta_1^k + \frac{\beta_2}{1-\beta_1}.$$

According to accumulating generation, one similarly has

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k), \qquad k = 1, 2, \cdots, n-1.$$

3.2.2. *Model testing.* We first introduce residual test

$$e(k) = x^{(0)}(k) - \hat{x}^{(0)}(k)$$

and error

$$Q(k) = \frac{e(k)}{x^{(0)}(k)}.$$

The reliability of the model is tested using the posterior error test, where the accuracy of the model is determined by the small error probability P and posterior error ratio C. The posteriori error ratio C is formulated by

$$C = \frac{\sqrt{\frac{1}{n}\sum_{k=1}^{n}\left[x^{(0)}(k) - \bar{x}^{(0)}\right]^2}}{\sqrt{\frac{1}{n}\sum_{k=1}^{n}\left[e(k) - \bar{e}\right]^2}},$$

where $\bar{x}^{(0)}$ and $\bar{e}$ denote respectively the mean values of the original sequence $X^{(0)}$ and the residual sequence E.

The small error probability P is calculated by

$$P = P\left(|e(k) - \bar{e}| < 0.6745 S_0\right)$$

with S_0 standing for the raw data mean squared error and 0.6745 is the 75% quantile (or probable error) of the normal distribution.

The accuracy of the model is usually determined by the C and P values. Generally speaking, if the C-value is less than 0.35, the model accuracy level is satisfactory. If the C-value is less than 0.5, the accuracy of the model reaches the standard. If the C-value is less than 0.65, the model accuracy level is basically qualified. If the C-value is greater than 0.65, the model accuracy level is unqualified. The larger the P-value, the better the application of the model.

4. Results and Analysis

In this section, we use the model of the previous section to give the recent development trend of bicycle sharing, and propose an optimization scheme based on the polynomial growth sequence.

4.1. **Results of AHP.** In order to achieve the research goals, we constructed a hierarchical structure of the bicycle sharing problem, as shown in Figure 2. The standard layer has four evaluation indicators, and the program layer has five itinerary options.

Based on the nine important levels proposed by Saatty [16], we evaluated the weather (mainly temperature), travel time, travel distance (Euclidean distance) and expenditure as indicators, and then obtained five travel scenarios, namely bicycles, taxis, buses (including subways), private cars and walking (as a control group).

Through the literature [5], [18], [19], we obtained bicycle sharing usage analysis matrix A, weather (temperature) judgment matrix B_1, travel duration judgment matrix B_2, travel distance (Euclidean distance) judgment matrix B_3, and expenditure judgment matrix B_4 as follows,

$$A = \begin{bmatrix} 1 & 6 & 5 & 4 \\ 1/6 & 1 & 1/3 & 1/4 \\ 1/5 & 3 & 1 & 2 \\ 1/4 & 4 & 1/2 & 1 \end{bmatrix},$$

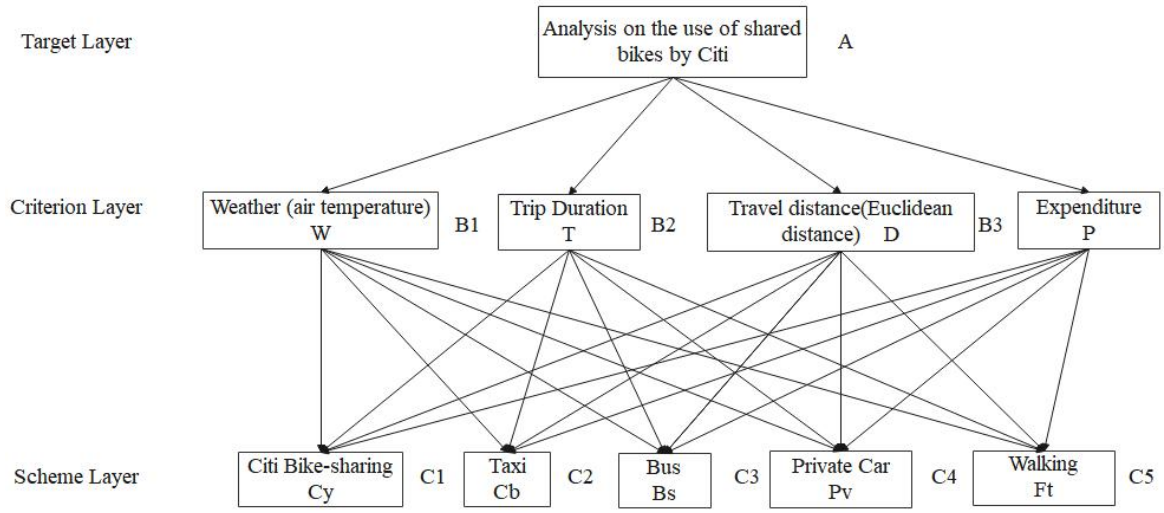

FIGURE 2. Hierarchy of travel modalities

$$B_1 = \begin{bmatrix} 1 & 1/5 & 1/4 & 1/6 & 3 \\ 5 & 1 & 3 & 1/3 & 5 \\ 4 & 1/3 & 1 & 1/5 & 4 \\ 6 & 3 & 5 & 1 & 6 \\ 1/3 & 1/5 & 1/4 & 1/6 & 1 \end{bmatrix}, \qquad B_2 = \begin{bmatrix} 1 & 1/5 & 1/5 & 1/8 & 3 \\ 5 & 1 & 4 & 1/3 & 5 \\ 5 & 1/4 & 1 & 1/5 & 5 \\ 8 & 3 & 5 & 1 & 9 \\ 1/3 & 1/5 & 1/5 & 1/9 & 1 \end{bmatrix},$$

$$B_3 = \begin{bmatrix} 1 & 1/4 & 1/3 & 1/6 & 3 \\ 4 & 1 & 5 & 1/3 & 6 \\ 3 & 1/5 & 1 & 1/5 & 4 \\ 6 & 3 & 5 & 1 & 7 \\ 1/3 & 1/6 & 1/4 & 1/7 & 1 \end{bmatrix}, \qquad B_4 = \begin{bmatrix} 1 & 5 & 3 & 8 & 1/2 \\ 1/5 & 1 & 1/3 & 3 & 1/6 \\ 1/3 & 3 & 1 & 5 & 1/3 \\ 1/8 & 1/3 & 1/5 & 1 & 1/8 \\ 2 & 6 & 3 & 8 & 1 \end{bmatrix}.$$

We find the corresponding weights through the special value method, as shown in Figure 3 and Figure 4.

	CR	CI	Weather	Trip Duration	Travel Distance	Expenditure
A	0.0850	0.0765	0.5966	0.0621	0.1890	0.1523

FIGURE 3. CR, CI and Weights1

	CR	CI	Citi-bike	Taxi	Bus	Private Car	Walking
B1(Weather)	0.0926	0.1037	0.0676	0.2591	0.1422	0.4878	0.0433
B2(Trip Duration)	0.0903	0.1011	0.0540	0.2776	0.1348	0.5029	0.0308
B3(Travel Distance)	0.0898	0.1006	0.0700	0.2870	0.1179	0.4864	0.0386
B4 (Expenditure)	0.0354	0.0397	0.3133	0.0699	0.1532	0.0352	0.4292
Total impact weight			0.104629	0.236707	0.138823	0.419542	0.100408
Compared to walking			1.042040	2.357449	1.382589	4.178373	1.000000

FIGURE 4. CR, CI and Weights2

Under the above comprehensive conditions, bicycle sharing is the least affected mode of travel, followed by public transportation (including subways). Although weather is the biggest factor affecting all modes of travel, the most important factor affecting people's choice of bicycle sharing is expenditure.

For other unintuitive aspects, let's take the data from Open Data NYC in January 2014 as an example (the percentage of data for each month is roughly the same). About 60% of people use shared bicycles for less than 10 minutes (the driving distance is about 1.8 kilometers). 99.9% of people used shared bicycles within 15 minutes (about 2.7 kilometers). In addition, among users, 18.9% are female, 78.5% are male, and 2.6% are of unknown gender. In other words, male are more likely to use shared bicycles, which may be related to the fact that more office workers are male. On weekdays, the frequency of using shared bicycles near subway and bus stops is higher than in other regions. Therefore, adjusting the price and quantity of shared bicycles in these regions has the greatest impact on the overall operation of bicycle sharing.

4.2. **Results of DGM(1,1).** The data shown in Figure 5 is from Open data in New York City (https://data.cityofnewyork.us/NYC-BigApps/Citi-Bike-System-Data). The data for 2016 is not yet complete. In order to create a complete data series, we must use several methods to compensate for the lost data. Common methods include average generation and stepwise ratio generation.

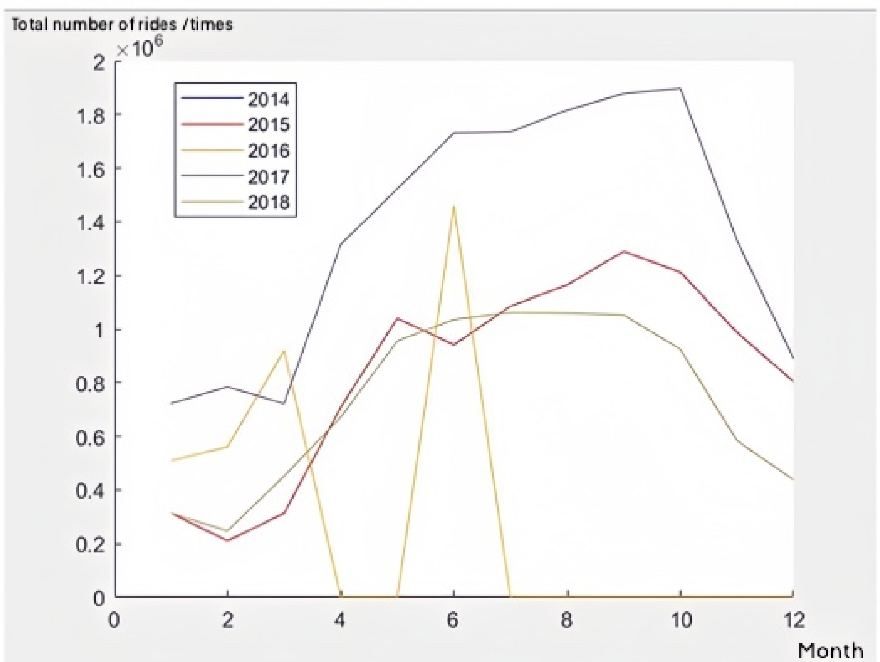

FIGURE 5. Total number of City Bike rides per year

(1) Average generation. For a sequence X which is blank $\phi(k)$ at location k, the value

$$x(k) = ax(k-1) + (1-a)x(k+1)$$

is called a *mean generated value of non-consecutive neighborhood* if $a = 1/2$. It is generally believed that the new data is more meaningful than the old data, so the value of a can be less than 1/2. For example, one takes $a = 0.4$ to get $x(k) = 0.4x(k-1) + 0.6x(k+1)$.

(2) Stepwise ratio generation. For a complete sequence of data, $\sigma(k) = \frac{x(k)}{x(k-1)}$ $(k = 2, 3, \cdots, n)$ is called stepwise ratios of the sequence X. While a sequence is blank at the two ends,

$$X = (\phi(1), x(2), \cdots, x(n-1), \phi(n)),$$

the stepwise ratio of the right-side neighbor of $\phi(1)$ is used to generate $x(1)$, and the stepwise ratio of the left-side neighbor of $\phi(n)$ is used to generate $x(n)$. In this way, $x(1)$ and $x(n)$ are said to be stepwise ratio generated, and specifically, one has

$$x(1) = \frac{x(2)}{\sigma(3)}, \qquad x(n) = x(n-1)\sigma(n-1).$$

For annual raw data

$$X^{(0)} = [8791987, 10068280, 3450557, 16347284, 17752410]\,,$$

the mean generation value is used to fill the missing data in 2016, and the following new sequence is obtained,

$$X'^{(0)} = [8791987, 10068280, 13835682, 16347284, 17752410]\,.$$

The results obtained by establishing a complete DGM(1,1) model are shown in Figure 6, as shown below.

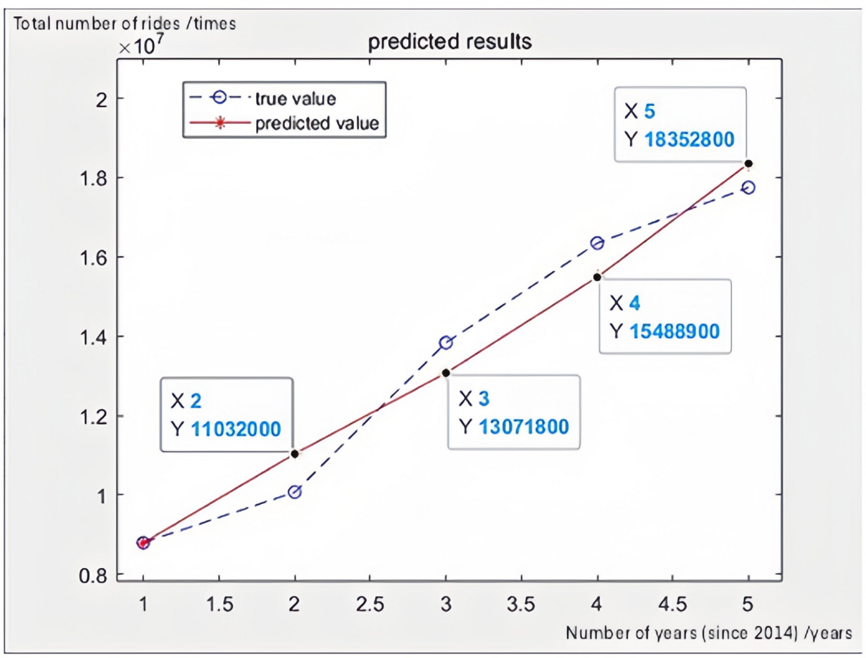

FIGURE 6. Predicted results

The posteriori error ratio C is 0.2083, and the small error probability P is 1, which shows that the model can obtain good results for this problem. Therefore, we use this model to predict the data from 2019 to 2023, and the forecast results follow that

$$X' = [21746234, 25767151, 30531543, 36176880, 42866049]\,.$$

The forecast results show that the total annual cycling of bicycle sharing is gradually increasing, and the demand for bicycle sharing is steadily increasing.

4.3. **Model optimization.** The GM(1,1) model provides good prediction results in a short period of time. When the time is long, the error of the model will become larger. In order to overcome such disadvantages, we introduced a model construction method based on polynomial growth sequence. It is assumed that the corresponding system is relatively stable over time, and the general term takes $x^{(0)}(k) = ac^k$, that is,

$$X^{(0)} = \left[ac^1, ac^2, ac^3, ac^4, ac^5\right].$$

The first-order accumulating generation reads

$$X^{(1)} = \left[ac^1, a\sum_{i=1}^{2} c^i, a\sum_{i=1}^{3} c^i, a\sum_{i=1}^{4} c^i, a\sum_{i=1}^{5} c^i\right].$$

Putting

$$Y = \begin{bmatrix} a\sum_{i=1}^{2} c^i \\ a\sum_{i=1}^{3} c^i \\ a\sum_{i=1}^{4} c^i \\ a\sum_{i=1}^{5} c^i \end{bmatrix}, \qquad B = \begin{bmatrix} ac^1 & 1 \\ a\sum_{i=1}^{2} c^i & 1 \\ a\sum_{i=1}^{3} c^i & 1 \\ a\sum_{i=1}^{4} c^i & 1 \end{bmatrix},$$

one computes that

$$\hat{\beta} = \left(B^T B\right)^{-1} B^T Y = \begin{bmatrix} c \\ ac \end{bmatrix},$$

and hence,

$$\hat{x}^{(1)}(k+1) = x^{(0)}(1)\beta_1^k + \frac{\beta_2}{1-\beta_1}\left(1-\beta_1^k\right) = ac^{k+1} + \frac{1-c^k}{1-c}ac = a\sum_{i=1}^{k+1} c^i. \tag{4.1}$$

Applying inverse accumulating generation to Eq. (4.1), one has

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1) = a\sum_{i=1}^{k} c^i - a\sum_{i=1}^{k-1} c^i = ac^k,$$

which shows that $\hat{x}^{(0)}(k)$ is equal to $x^{(0)}(k)$. In other words, it can accurately simulate the sequence of polynomial growth.

In the above derivation, the parameter c can take any value. Therefore, it can be seen that the DGM(1,1) model can be used to simulate and predict the original data series, as long as it has an approximate polynomial growth law. The results show that the deviation in the model prediction is caused by the deviation between the whitenization form e^{-a} and the discrete form β_1. When the forecast is short-term or the development coefficient $-a$ is small, the difference has little impact on the entire prediction model, so the prediction accuracy is high. However, when the prediction is long-term or when the development coefficient $-a$ is large, the impact of the deviation on the overall model increases sharply, and the prediction accuracy decreases, sometimes even leading to unacceptable results.

5. Conclusions

At present, the development of bicycle sharing is considered necessary for a livable living environment. With the development of technology, the bicycle sharing system has been continuously upgraded. Nowadays, they have become an indispensable new transportation option to solve problems such as excessive consumption of natural resources, traffic congestion and intensive lifestyles. In order to improve the quality of service, it is necessary to well understand the factors that affect bicycle sharing. In this work, we used an analytic hierarchy process to investigate the influence of four factors on people's travel choices, including weather (mainly temperature), travel duration, travel distance and expenditure. The results show that bicycle sharing is the most reliable way to travel, and among them, travel expenses are the biggest influence on people's choice of bicycle sharing. We use a discrete GM(1, 1) gray prediction model to confirm that the development of bicycle sharing will steadily improve in the next five years. The assessment can provide manufacturers with effective guidance to promote the upgrade of bicycle sharing systems, meet the needs of users, and help decision makers respond to the market to implement more effective business strategies. In addition, the research can also provide information to simplify and enhance the transportation distribution process, and reduce the management costs of governments and organizations.

Since the discrete GM(1,1) model uses older data from NYC open data, even if the model obtains a better solution in a shorter period of time, the longer the time, the greater the model error. To this end, in our paper, we propose a method for constructing a model based on the polynomial growth series, and the optimized model can qualitatively judge the prediction accuracy. However, considering the role of AHP, it is best to choose from alternatives. We cannot determine whether other factors have a greater impact on bike sharing than other factors (such as cycling environment and safety equipment). The follow-up work of the project is to compare these results with the spatial and temporal patterns of bicycle sharing services in other cities, and to further study the influence of multiple climatic factors, including rainfall.

Statements and Declarations

The authors declare that they have no conflict of interest, and the manuscript has no associated data.

Acknowledgements

This research is supported in part by the National Natural Science Foundation of China (Grant No. 11501203) as well as an Innovative Training Programs (Grant No. X202310251251, 202310251115).

References

[1] L. Caggiani, R. Camporeale, M. Ottomanelli, and W. Y. Szeto. A modeling framework for the dynamic management of free-floating bike-sharing systems. *Transportation Research Part C: Emerging Technologies*, 87:159-182, 2018.
[2] A. A. Campbell, C. R. Cherry, M. S. Ryerson, and X. Yang. Factors influencing the choice of shared bicycles and shared electric bikes in Beijing. *Transportation Research Part C: Emerging Technologies*, 67:399-414, 2016.
[3] G. Cantelmo, R. Kucharski, and C. Antoniou. Low-Dimensional Model for Bike-Sharing Demand Forecasting that Explicitly Accounts for Weather Data. *Transportation Research Record*, 2674(8):132-144, 2020.
[4] D. Chemla, F. Meunier, and R.W. Calvo. Bike sharing systems: Solving the static rebalancing problem. *Discrete Optimization*, 10(2):120-146,2013.
[5] S. Gössling, C. Neger, R. Steiger, and R. Bell. Weather, climate change, and transport: a review. *Nat Hazards*, 118:1341-1360, 2023.
[6] M. Dell'Amico, E. Hadjicostantinou, M. Iori, and S. Novellani. The bike sharing rebalancing problem: Mathematical formulations and benchmark instances. *Omega*, 45:7-19, 2014.
[7] W. El-Assi, M. S. Mahmoud, and K. N. Habib. Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto. *Transportation*, 44:589-613, 2017.
[8] E. Eren and V. E. Uz. A review on bike-sharing: The factors affecting bike-sharing demand. *Sustainable Cities and Society*, 54:101882, 2020.
[9] E. Fishman, S. Washington, and N. Haworth. Bike share's impact on car use: Evidence from the United States, Great Britain, and Australia. *Transportation Research Part D: Transport and Environment*, 31:13-20, 2014.
[10] E. Fishman, S. Washington, N. Haworth, and A. Watson. Factors influencing bike share membership: An analysis of Melbourne and Brisbane. *Transportation Research Part A: Policy and Practice*, 71:17-30, 2015.
[11] I. A. Forma, T. Raviv, and M. Tzur. A 3-step math heuristic for the static repositioning problem in bike-sharing systems. *Transportation Research Part B: Methodological*, 71:230-247, 2015.
[12] J. C. García-Palomares, J. Gutiérrez, and M. Latorre. Optimizing the location of stations in bike-sharing programs: A GIS approach. *Applied Geography*, 35(1-2):235-246, 2012.
[13] L. Lin, Z. He, and S. Peeta. Predicting station-level hourly demand in a large-scale bike-sharing network: A graph convolutional neural network approach. *Transportation Research Part C: Emerging Technologies*, 97:258-276, 2018.
[14] G. McKenzie. Spatiotemporal comparative analysis of scooter-share and bike-share usage patterns in Washington, D.C.. *Journal of Transport Geography*, 78:19-28, 2019.
[15] A. Pal and Y. Zhang. Free-floating bike sharing: Solving real-life large-scale static rebalancing problems. *Transportation Research Part C: Emerging Technologies*, 80:92-116, 2017.
[16] T.L. Saaty and L.G. Vargas. *Models, Methods, Concepts & Applications of the Analytic Hierarchy Process*, New York: Springer, 2012.
[17] J. Schuijbroek, R. C. Hampshire, and W.-J. van Hoeve. Inventory rebalancing and vehicle routing in bike sharing systems. *European Journal of Operational Research*, 257(3):992-1004, 2017.

[18] H. Lu and Q. Luo. Research on the impact of bike sharing on urban transportation-Take New York's citi bike as an example, Institute of Management Science and Industrial Engineering, In Proceedings of 2019 6th Asia-Pacific Conference on Social Sciences, Humanities (APSSH 2019), Francis Academic Press, UK: Computer Science and Electronic Technology International Society, 2019.

[19] J.F. Teixeira, C. Silva, and F.M. Sá. The motivations for using bike sharing during the COVID-19 pandemic: Insights from Lisbon. *Transportation Research Part F: Traffic Psychology and Behaviour*, 82:378-399, 2021.