Forecasting the multifactorial interval grey number sequences using grey relational model and GM (1, N) model based on effective information transformation

Jing Ye[1*], *Yaoguo Dang*[2], *Yingjie Yang*[3]

1 *School of Management Science and Engineering, Nanjing University of Finance & Economics, Nanjing, Jiangsu, 210023, PR China*

2 College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu, 210016, PR China

3 Centre for Computational Intelligence, De Montfort University, Leicester, LE1 9BH, UK

Abstract： In the context of data eruption, the data often shows a short-term pattern and changes rapidly which makes it difficult to use a single real value to express. For this kind of small-sample and interval data, how to analyze and predict muti-factor sequences efficiently becomes a problem. By this means, grey system theory (GST) is developed in which the interval grey numbers, as a typical object of GST, characterize the range of data and the grey relational and prediction models analyze the relations of multiple grey numbers and forecast the future. However, traditional grey relative relational model has some limitations: the results obtained always show low resolution and there are no extractions for the interval feature information from the interval grey number sequence. In this paper, the grey relational analysis model (GRA) based on effective information transformation of interval grey numbers is established, which contains comprehensive information of area differences and slope variances and optimizes the resolution of traditional grey degree. Then, according to the relational results, the multivariable GM model (GM(1,N)) is proposed to forecast the interval grey number sequence. To verify the effectiveness of this novel model, it is established to analyze the relationship between the degree of traffic congestion and its relevant factors in the Yangtze River Delta of China and predict the development of urban traffic congestion degrees in this area over the next five years. In addition, some traditional statistical methods (principal component analysis, multiple linear regression models and curve regression models) are established for comparisons. The results show high performances of the novel GRA model and GM(1,N) model, which means the models proposed in this paper are suitable for interval grey numbers from regional data. The strengths which recommend the use of this novel method lie in its high recognition mechanism and muti-angle information transformation for interval grey numbers as well as its characteristic of timeliness in information processing.

Key Words: Grey numbers; Grey system theory; Grey relational analysis; GM(1,N); traffic congestion; China

1. Introduction

* Corresponding author.
E-mail: yejingjenson@163.com (J Ye).

For most traditional relational and prediction methods of multivariate sequences, data analysis often needs to be built on the basis of statistical data which refers to long term data (Rehborn et al., 2011; Shankar et al., 2012; Xu et al., 2013; Younes and Boukerche, 2015). The large volume of historical data can easily ignore the small amount of the latest data which leads to its insensitivity to the latest change. In the era of big data, although the total amount of information is rapidly expanding, the relationships between the relevant data will change with their surrounding environment as well. Furthermore, with the accelerated pace of social developments, numerical values of the same factor or index often changes its order of magnitude over a short time which may lead to qualitative changes. So, the latest data is usually more significant. In reality, it is quite possible that the effects of long-term old data on the status quo are less than that of short-term recent data and this conclusion has been proved in many papers (Wu et al., 2013a; Ujjwal and Jain, 2010; Erdal et al., 2010). In a word, time limitations should be considered. In this regard, grey system theory is established to solve this problem.

Grey system theory (Liu and Lin, 2010) has been proposed to take full advantage of the latest short data information. To be specific, grey models are designed to characterize uncertain systems with two main characteristics: latest small samples and limited information. That is to say, grey models are suitable for cases of ‘small samples’ -only 4 discrete data samples are sufficient, which is one of the considerable advantages of grey forecasting modeling over other methods. So far, many investigations have demonstrated the effectiveness of grey models compared with classical mathematical and statistical models (e.g. ARIMA, exponential smoothing, Holt-Winters method, linear method) and modern heuristic methods (e.g. ANN, fuzzy systems) (Lin et al., 2012; Xia and Wong, 2014; Wu et al., 2015a; Wu et al., 2016). Grey modeling is an alternative tool for those systems whose structure is complex, uncertain and chaotic, where ‘limited latest information’ is the most valuable information (Wu et al., 2015b).

As a result, the scholars have carried out much research using grey system models. For factor analysis and selection, the grey relational analysis model has been established, and researchers have obtained a series of academic achievements: Wang et al.(2015) established a mathematical model for a synthesized evaluation according to theories of grey relational analysis (GRA) and the analytic hierarchy process (AHP) to select a biomass briquette fuel (BBF) system scheme. Grey relational analysis was exploited to evaluate the hydrogen evolution performance of eight different non-precious metal alloy cathodes by Kadier et al. (2015). Wei(2011) discussed the dynamic hybrid multiple attribute decision making problems, where the decision making information was expressed in real numbers, interval numbers or linguistic labels, by utilizing three different GRA (grey relational analysis, real-valued GRA, interval-valued GRA and fuzzy-valued GRA) methods. Mohammadi and Makui (2017) used grey relational coefficients in muti-attribute group decision making based on interval-valued intuitionistic fuzzy sets and evidential reasoning methodology to establish a new approach for supporting decisions in intuitionistic fuzzy environments. For grey prediction models which are used to forecast future trends, fruitful outcomes are also produced by scholars; different grey prediction models have been applied to a variety of fields and results have been compared with other models. Among them, Hsu and Wang (2009) proposed a new prediction approach using the multivariate grey model combined with grey relational analysis to forecast integrated circuit outputs. Pao et al. (2012) employed the nonlinear grey Bernoulli model (NGBM) to predict three indicators (carbon emissions, energy consumption and real outputs) to analyze the relationships between them. Bahrami et al. (2014) presented a new model based on a combination of the wavelet transform and grey model for short term electric load forecasting, and it was

improved by a PSO (particle swarm optimization) algorithm. Kayacan et al. (2010) investigated the accuracy of different grey models such as GM(1,1), Grey Verhulst model, and modified grey models using the Fourier Series which showed higher performances not only in model fitting but also in forecasting. Li et al. (2012) confirmed that the forecasting performances and results of AGM(1,1), based on grey theory, were better when compared with those obtained from back propagation neural networks (BPN) and support vector regression (SVR), by using the Asia-Pacific economic cooperation energy database to deal with the problem of forecasting electricity consumption, especially when the sample size was limited. Evans (2014) put forward a flexible generalisation of the Grey-Verhulst model and applied it to forecast the intensity of the UK steel usage and produced very reliable muti-step predictions. Wang et al. (2010) proposed a novel approach to improve prediction accuracy of the GM(1,1) model through optimization of the initial condition, which was comprised of the first item and the last item of a sequence generated from applying the first-order accumulative generation operator on the sequence of raw data. This could express the principle of new information priority emphasized in grey systems theory fully. Wu et al. (2013b) put forward a new Grey System model with the fractional order accumulation, where the priority of new information could be better reflected when the accumulation order number became smaller in the in-sample model. Liu et al. (2014) discussed the relative error between the solutions of a whitenization GM(1,1) model, called GM(1,1,W), and a connotation GM(1,1) model (GM(1,1, C)). Chen and Huang (2013) demonstrated scientific and effective procedures to solve a singular phenomenon of the GM(1,1) prediction model with practical cases which occurred when computers were used to calculate matrix values to obtain the grey development coefficient, and showed its application in forecasting the moving path of the typhoon MORAKOT. Wu et al. (2015c) proved that the errors from the inverse accumulating generation operator are affected by the order number of accumulating generation operator which was used to smooth the randomness in grey forecasting model and proposed fractional order AGO method to optimize the model's performance. All these papers show that grey system theory (GST) is an emerging and booming research field and there is much work to be done.

Turning to the reality, with the increasingly prominent contradiction between supply and demand of urban transportation, urban traffic congestion is becoming more serious than ever before. Considering the short-term and regional characteristics of traffic congestion data, it is suitable to be used as the case to test models. However, the related research findings regarding traffic problems by using GST are limited. Specifically, Li et al. (2015) established a novel grey generation relational analysis model based on the grey exponential law, in which the dynamic change trend similarity of the original time series was characterized by the proximity of the generation rate sequence to explore traffic congestion. However, this model only considered a one aspect-generation rate and the traffic congestion degree was characterized as real numbers which could not exactly reflect its uncertain feature in the real world. Guo et al. (2013) forecasted the short-term traffic flow based on the GM(1,1|τ,r) concerning the delay and nonlinear properties of traffic flow in urban road systems. In this paper, the authors only considered one characteristic of traffic flow.

In order to study the relevant factors affecting traffic congestion and predict the future trend of traffic congestion in the Yangtze River Delta region, a novel grey relational model based on effective information transformation of interval grey numbers is established in this paper, and the degree of traffic congestion is predicted by adding the relevant factor information through the GM (1, N) model. This article differs from the existing studies in a number of ways. Firstly, interval

grey numbers are introduced to represent regional characteristics of traffic congestion in the Yangtze River Delta area, which not only objectively expresses regional differences, but also the superiority of grey modeling. Secondly, in order to reflect influencing factors of urban traffic congestion comprehensively and effectively, four shout-term related factor index sequences(consumption level of urban residents, urban population density, public transport vehicles per million people and urban road area per ca pita) have been selected from the aspects of driving factors, demand factors, supply factors, etc. Thirdly, according to the related interval grey numbers' sequences, a novel GRA model based on effective information transformation is proposed to recognize key influencing factors of traffic congestion. Fourthly, by using these key factors, multivariate factors model- GM (1, N) is set up to predict the trend of traffic congestion in the Yangtze River Delta region in the next 5 years. Finally, by comparing with other traditional statistical methods, some conclusions have been drawn.

The remainder of this paper is organized as follows: Section 2 outlines the models and both the grey relational analysis and GM approaches are presented; Section 3 presents the data used and empirical findings. Section 4 summarizes and concludes the paper.

2. Methodology

The formation and development of system reference sequences are often influenced by the sequences of one or more behavioral factors. How to identify the key influencing factors is the basis for judging the future trend of the reference sequences accurately. In this section, we firstly review the modeling steps of the traditional relative grey relational model and put forward the existing problems in Section 2.1. In Section 2.2, on the basis of improving the traditional grey relational model, a novel GRA model based on effective information transformation is proposed, which is mainly used to identify the key factors from related factors that affect future traffic congestion. Then, in Section 2.3, the nonlinear grey prediction model with N variables (GM (1, N)) is established, based on the sequences of identified key factors and the traffic congestion sequence data together, to predict the future trend of traffic congestion and verify the validity of the key factor selection from the quantitative point of view. Finally, for the practical purpose of applying to interval grey numbers, the modeling steps of the novel grey relational model, based on effective information transformation proposed in this paper and GM (1, N) prediction model are summarized.

2.1 Traditional relative grey relational analysis model

Grey relational modeling is used to explore the strength of the relations between relevant factors under the circumstances of small samples or non-significant regularity. It is an important part of grey system theory and provides the basis for grey clustering, grey system modeling and decision making. The basic idea of grey relational analysis is to judge whether relevant factors associate closely with each other according to the similarity of the geometric shapes of the related sequences and the closeness of their quantities (Liu et al., 2006).

Let $X_0 = (x_0(1), x_0(2), \ldots, x_0(n))$ and $X_1 = (x_1(1), x_1(2), \ldots, x_1(n))$ as the behavior sequences, the form of traditional relative grey relational analysis model is denoted as

$\varepsilon_{0i} = \dfrac{1+|s_0|+|s_i|}{1+|s_0|+|s_i|+|s_i-s_0|}$ and the modeling process is as follows (Liu et al., 2010):

First, normalize data by Operator D_1 and Operator D_2. To be specific, the initial process of $X_i = (x_i(1), x_i(2), \ldots, x_i(n)), i = 0,1$ by using D_1 can be implemented as follows.

$$x_i(k)d_1 = x_i(k)/x_i(1), x_i(1) \neq 0, k = 1,2,\ldots,n, i = 0,1 \quad (1)$$

And this process can be denoted as

$X_i' = (x_i'(1), x_i'(2), \ldots, x_i'(n)) = X_i D_1 = (x_i(1)d_1, x_i(2)d_1, \ldots, x_i(n)d_1), i = 0,1$.

Then, the start zeroized process of $X_i' = (x_i'(1), x_i'(2), \ldots, x_i'(n)), i = 0,1$ by using D_2 can be implemented as follows.

$$x_i'(k)d_2 = x_i'(k) - x_i'(1), k = 1,2,\ldots,n, i = 0,1 \quad (2)$$

And this process can be denoted as

$X_i'^0 = (x_i'^0(1), x_i'^0(2), \ldots, x_i'^0(n)) = X_i' D_2 = (x_i'(1)d_2, x_i'(2)d_2, \ldots, x_i'(n)d_2), i = 1,2$.

Calculate $|s_0|$, $|s_1|$ and $|s_0 - s_1|$.

$$|s_0| = \left|\sum_{k=2}^{n-1} x_0'^0(k) + \frac{1}{2}x_0'^0(n)\right| \quad (3)$$

$$|s_i| = \left|\sum_{k=2}^{n-1} x_i'^0(k) + \frac{1}{2}x_i'^0(n)\right| \quad (4)$$

$$|s_i - s_0| = \left|\sum_{k=2}^{n-1} (x_i'^0(k) - x_0'^0(k)) + \frac{1}{2}(x_i'^0(n) + x_0'^0(n))\right| \quad (5)$$

Finally, the relative degree of grey relational analysis is:

$$\varepsilon_{0i} = \frac{1+|s_0|+|s_i|}{1+|s_0|+|s_i|+|s_i-s_0|} \quad (6)$$

The range of relative degree of grey relational model should be [0, 1]. The change rates of these two sequences relative to the starting point are more consistent, while the value of the relative degree of grey relational model is larger. However, in the traditional relative grey relational analysis model, here is $\varepsilon_{0i} = \dfrac{1+|s_0|+|s_i|}{1+|s_0|+|s_i|+|s_i-s_0|} = 1 - \dfrac{|s_i-s_0|}{1+|s_0|+|s_i|+|s_i-s_0|}$.

Because $\dfrac{|s_i-s_0|}{1+|s_0|+|s_i|+|s_i-s_0|} \geq 0$, then $\varepsilon_{0i} \leq 1$. While considering the extreme situation

($s_0 \to 0$ and $s_i \to \infty$), $\frac{|s_i - s_0|}{1+|s_0|+|s_i|+|s_i - s_0|} \approx 0.5$ and the value of this degree is $\varepsilon_{0i} = \frac{1+|s_0|+|s_i|}{1+|s_0|+|s_i|+|s_i - s_0|} = 1 - \frac{|s_i - s_0|}{1+|s_0|+|s_i|+|s_i - s_0|} > 0.5$. That means even if there is a huge difference between X_0 and X_i, the result of the traditional degree can still reach above 0.5 which is invalid or distorted. So the range of the traditional degree becomes (0.5, 1] which cannot cover the range of [0, 1]. This probably leads to the reduction of resolution of this grey relational model and the misunderstanding of the relational results. Furthermore, the traditional relative grey relational analysis model does not make the usage of the information between two sequences effectively as it only considers the change rates of these two sequences.

2.2 A novel grey relational analysis model based on effective information transformation

To improve the limitations and enhance accuracy of the traditional relative grey relational model, a novel grey relational model based on effective information transformation is proposed. For the effective information transformation, the proposed model contains two aspects.

Assume the original reference sequence to be $X_0^{(0)} = (x_0^{(0)}(1), x_0^{(0)}(2), \ldots, x_0^{(0)}(n))$, and the behavior sequences to be

$$X_1^{(0)} = (x_1^{(0)}(1), x_1^{(0)}(2), \ldots, x_1^{(0)}(n))$$

$$X_2^{(0)} = (x_2^{(0)}(1), x_2^{(0)}(2), \ldots, x_2^{(0)}(n))$$

$$\vdots$$

$$X_N^{(0)} = (x_N^{(0)}(1), x_N^{(0)}(2), \ldots, x_N^{(0)}(n)), \quad k = 1,2,\ldots,n.$$

Here, the grey relational degree between $X_0^{(0)}$ and $X_i^{(0)}$ $(i = 1,2,\ldots,n)$ based on the effective information transformation can be derived as $\gamma(X_0^{(0)}, X_i^{(0)})$:

$$\gamma(X_0^{(0)}, X_i^{(0)}) = 0.5 \cdot (\varepsilon'_{0i} + \delta_{0i}), i = 1,2,\ldots,n \tag{7}$$

Among them, ε'_{0i} is the grey relational degree based on area differences and δ_{0i} is the grey relational degree based on slope variances. They represent the differences of relative amounts and growth rates respectively which cover more available information.

The modeling processes of the two parts are as follows:

(1) The novel degrees of grey relation models based on area differences

According to the traditional model, the first step is also to normalize original data. But in

different ways, here data are dealt with the average operator D_3 (Liu et al., 2010).

$$x_i^{(0)}(k)d_3 = \frac{x_i^{(0)}(k)}{\overline{X}_i^{(0)}}, \overline{X}_i^{(0)} = \frac{1}{n}\sum_{k=1}^{n} x_i^{(0)}(k), k = 1,2,\ldots,n, i = 0,1,2,\ldots,n \quad (8)$$

And this process can be denoted as

$$X_i'^{(0)} = (x_i'^{(0)}(1), x_i'^{(0)}(2), \ldots, x_i'^{(0)}(n)) = X_i^{(0)} D_3 = (x_i^{(0)}(1)d_3, x_i^{(0)}(2)d_3, \ldots, x_i^{(0)}(n)d_3)$$

$, i = 0,1,2,\ldots,n$.

Instead of using the initial operator D_1, the main reason is the consideration of the magnitude of data. With developments in all aspects of society, data explosion often show that earlier numerical values in a sequence are extremely small compared with the latter values. In this situation, the effect of data normalizing would be weakened by using the initial operator that cannot effectively explore system characteristics, resulting in the distortion of the incidence degree. Furthermore, it is obvious that data compression using the average operator processes must be more powerful than the initial operator that can better show the changing rates of the sequences.

Then, deal with the start zeroized process of $X_i'^{(0)} = (x_i'^{(0)}(1), x_i'^{(0)}(2), \ldots, x_i'^{(0)}(n))$ by using the operator D_2 which is the same as the traditional model.

$$x_i'^{(0)}(k)d_2 = x_i'^{(0)}(k) - x_i'^{(0)}(1), k = 1,2,\ldots,n, i = 0,1,2,\ldots,n \quad (9)$$

And this process can be denoted as

$$X_i''^{(0)} = (x_i''^{(0)}(1), x_i''^{(0)}(2), \ldots, x_i''^{(0)}(n)) = X_i'^{(0)} D_2 = (x_i'^{(0)}(1)d_2, x_i'^{(0)}(2)d_2, \ldots, x_i'^{(0)}(n)d_2)$$

$, i = 0,1,2,\ldots,n$.

After the first two steps, data normalization of the original behavior sequences has been completed. The sequences have been transformed into $X_i''^{(0)} = (x_i''^{(0)}(1), x_i''^{(0)}(2), \ldots, x_i''^{(0)}(n)), i = 0,1,2,\ldots,n$.

Third, calculate s_0', s_i', $s_i' - s_0'$. Due to the different processing method in the first step compared with the traditional model, the results of s_0', s_i', $s_i' - s_0'$ may totally different. There are two situations.

To be specific, **for** s_0' **and** s_i', s_0' is the area surrounded by the sequence $X_0''^{(0)} = (x_0''^{(0)}(1), x_0''^{(0)}(2), \ldots, x_0''^{(0)}(n))$ and the horizontal axis and s_i' is the area surrounded

by the sequence $X_i''^{(0)} = (x_i''^{(0)}(1), x_i''^{(0)}(2), \ldots, x_i''^{(0)}(n)), i = 1,2,\ldots,n$ and the horizontal axis. As the calculation processes of s_0' and s_i' are the same, the following calculation steps take s_0' as the example.

The first situation is all the items to be calculated are on the same side of the horizontal axis. In this situation, the results of s_0', s_i', $s_i' - s_0'$ can refer to the traditional model seen in Equation (3)-(5).

For the second situation, items to be calculated are not always on the same side of the horizontal axis and the results can be more complex:

The first item $x_0''^{(0)}(1)$ and the second item $x_0''^{(0)}(2)$ of the sequence $X_0''^{(0)} = (x_0''^{(0)}(1), x_0''^{(0)}(2), \ldots, x_0''^{(0)}(n)), i = 1,2,\ldots,n$, the shape of this area is a triangle because of the start zeroized process, so

$$|s_0'|(1) = \left|\frac{1}{2} x_0''^{(0)}(2)\right| \tag{10}$$

From the second item, for the items $x_0''^{(0)}(k)$ and $x_0''^{(0)}(k+1)$, if the two items are on the same side of the horizontal axis, then here is

$$|s_0'|(k) = \left|\frac{1}{2}(x_0''^{(0)}(k) + x_0''^{(0)}(k+1)\right| \tag{11}$$

While if the two items are on the opposite sides of the horizontal axis, then here is

$$|s_0'|(k) = \frac{1}{2}(x_0''^{(0)}(k) * -k)\left|x_0''^{(0)}(k)\right| + \frac{1}{2}(k+1-x_0''^{(0)}(k)*)\left|x_0''^{(0)}(k+1)\right| \tag{12}$$

Among them, $x_i''^{0}(k)*$ is the crossover point of the connection line of the two items with the horizontal axis and here is

$$x_0''^{(0)}(k)* = \frac{x_0''^{(0)}(k)}{x_0''^{(0)}(k) - x_0''^{(0)}(k+1)} + k \tag{13}$$

Thus, the total value of s_0' can be obtained:

$$|s_0'| = \sum_{k=1}^{n-1} |s_0'|(k) \tag{14}$$

Similarly, s_i' can be obtained:

$$|s_i'| = \sum_{k=1}^{n-1} |s_i'|(k) \tag{15}$$

For $s_i' - s_0'$, it expresses the difference area between s_0' and s_i'. For discrete time series, here is $s_i' - s_0' = \int_1^n |x_i''^{(0)}(t) - x_0''^{(0)}(t)|dt$ which are the area surrounded by $X_0''^{(0)} = (x_0''^{(0)}(1), x_0''^{(0)}(2), \ldots, x_0''^{(0)}(n))$ and $X_i''^{(0)} = (x_i''^{(0)}(1), x_i''^{(0)}(2), \ldots, x_i''^{(0)}(n))$ (the shadow part in Fig.1).

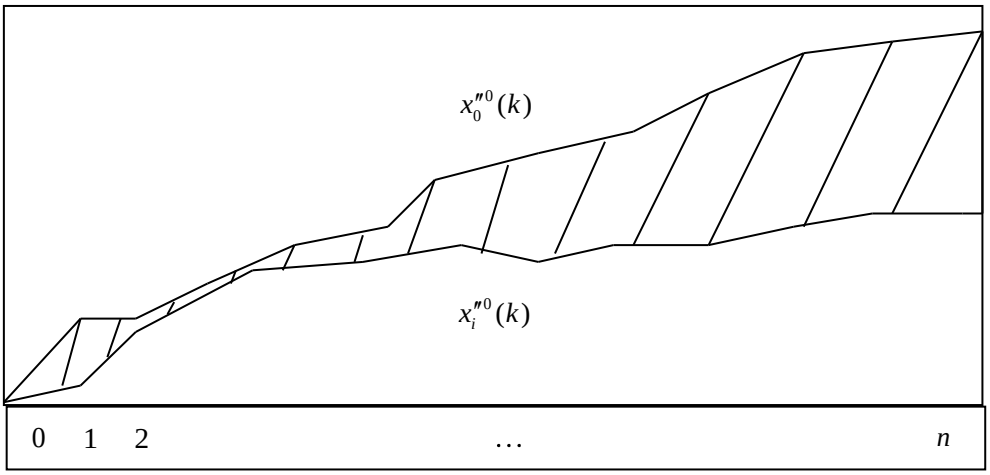

Fig.1. The diagrammatic sketch of $s_0 - s_i$ by $X_0''^{(0)}$ and $X_i''^{(0)}$

For the first item and the second item of $X_0''^{(0)} = (x_0''^{(0)}(1), x_0''^{(0)}(2), \ldots, x_0''^{(0)}(n))$ and $X_i''^{(0)} = (x_i''^{(0)}(1), x_i''^{(0)}(2), \ldots, x_i''^{(0)}(n)), i = 1,2,\ldots,n$ respectively, the shape of this area is a triangle, so

$$s_i' - s_0'|_1^2 = \int_1^2 |x_i''^{(0)}(t) - x_0''^{(0)}(t)|dt = \frac{1}{2}|x_i''^{(0)}(2) - x_0''^{(0)}(2)| \tag{16}$$

From the second item and the third item, **if** $(x_i''^{(0)}(k) - x_o''^{(0)}(k)) \cdot (x_i''^{(0)}(k+1) - x_o''^{(0)}(k+1)) > 0$, the shape of this area is a trapezoid in the interval of $[k, k+1]$, and

$$s_i' - s_0'|_k^{k+1} = \int_k^{k+1} |x_i''^{(0)}(t) - x_o''^{(0)}(t)|dt = \frac{1}{2}(|x_i''^{(0)}(k) - x_o''^{(0)}(k)| + |x_i''^{(0)}(k+1) - x_0''^{(0)}(k+1)|) \tag{17}$$

While, **if** $(x_i''^{(0)}(k) - x_o''^{(0)}(k)) \cdot (x_i''^{(0)}(k+1) - x_o''^{(0)}(k+1)) < 0$, the shape of this area are

two triangles in the interval of $[k, k+1]$. The crossover point ($x_{0i}^{\prime\prime(0)}(k)*$) should be obtained first. Through the horizontal coordinate of this crossover point, made up of the line segment linked by $(k, x_o^{\prime\prime(0)}(k))$, $(k, x_o^{\prime\prime(0)}(k))$ and the line segment linked by $(k, x_o^{\prime\prime(0)}(k))$, $(k, x_o^{\prime\prime(0)}(k))$, the value is

$$x_{0i}^{\prime\prime(0)}(k)* = \frac{x_0^{\prime\prime(0)}(k) - x_i^{\prime\prime(0)}(k)}{x_i^{\prime\prime(0)}(k+1) - x_i^{\prime\prime(0)}(k) - x_0^{\prime\prime(0)}(k+1) + x_0^{\prime\prime(0)}(k)} + k \tag{18}$$

Then, the area of two triangles can be obtained:

$$s_i' - s_0'\Big|_k^{k+1} = \frac{1}{2}[(x_{0i}^{\prime\prime(0)}(k)* - k)\left|x_i^{\prime\prime(0)}(k) - x_o^{\prime\prime(0)}(k)\right| + (k+1-x_{0i}^{\prime\prime(0)}(k)*)\left|x_i^{\prime\prime(0)}(k+1) - x_0^{\prime\prime(0)}(k+1)\right|] \tag{19}$$

Finally, considering the range limit of the traditional relative grey relational model, a novel degree of grey relational model based on area differences is proposed:

$$\varepsilon_{0i}' = 1 - \frac{|s_i' - s_0'|}{|s_0'| + |s_i'|} \tag{20}$$

As $\varepsilon_{0i}' = 1 - \frac{|s_i' - s_0'|}{|s_0'| + |s_i'|}$, here is $0 \le |s_i' - s_0'| \le |s_0'| + |s_i'|$, then $0 \le \frac{|s_i' - s_0'|}{|s_0'| + |s_i'|} \le 1$, so $0 \le \varepsilon_{0i}' = 1 - \frac{|s_i' - s_0'|}{|s_0'| + |s_i'|} \le 1$, the novel incidence coefficient can cover the range of [0, 1] which presents higher resolution than the traditional degree. For example, if $s_0 = 1$ and $s_i = 10$, the traditional degree is $\varepsilon_{0i} = 0.57$ and the novel degree of grey incidence based on area differences is $\varepsilon_{0i}' = 0.18$; if $s_0 = 3$ and $s_i = 7$, the traditional degree is $\varepsilon_{0i} = 0.73$ and the novel degree is $\varepsilon_{0i}' = 0.6$. It is obvious that the novel degree can provide clearer representation of the relationship between two sequences.

(2) The novel degrees of grey relational model based on slope variances

The novel degrees of grey incidence based on difference areas between grey numbers' sequences only reflect the level of overlap of different interval grey number sequences on the whole. To make the effective information transformation, the trend of each sequence has to be taken into account. For this purpose, novel degrees of grey relational model based on slope variances are established.

First, normalize original data by dealing with the average increment operator D_4 which is a new way of processing data.

$$x_i^{(0)}(k)d_4 = \frac{x_i^{(0)}(k)}{\Delta\overline{X}_i^{(0)}}, \Delta\overline{X}_i^{(0)} = \frac{1}{n-1}\sum_{k=2}^{n}\left|x_i^{(0)}(k) - x_i^{(0)}(k-1)\right| \tag{21}$$

Among them, $k = 2,3,\ldots,n, i = 0,1,2,\ldots,n$.

And this process can be denoted as

$\dot{X}_i^{(0)} = (\dot{x}_i^{(0)}(1), \dot{x}_i^{(0)}(2), \ldots, \dot{x}_i^{(0)}(n)) = X_i^{(0)}D_4 = (x_i^{(0)}(1)d_4, x_i^{(0)}(2)d_4, \ldots, x_i^{(0)}(n)d_4)$,

$i = 0,1,2,\ldots,n$.

To explain the feature of D_4, the nature of consistency and parallelism should be introduced first. For the sequences of $X_0^{(0)} = (x_0^{(0)}(1), x_0^{(0)}(2), \ldots, x_0^{(0)}(n))$ and $X_i^{(0)} = (x_i^{(0)}(1), x_i^{(0)}(2), \ldots, x_i^{(0)}(n))$, $i = 1,2,\ldots,n$. If

$$x_i^{(0)}(k) = Ax_0^{(0)}(k), A = const, k = 1,2,\ldots,n \tag{22}$$

It is called consistent transformation and when the above two sequences are considered, meets the nature of consistency (Xie and Liu, 2007).

If

$$x_i^{(0)}(k) = x_0^{(0)}(k) + c, c = const, k = 1,2,\ldots,n \tag{23}$$

It is called parallel transformation and when the above two sequences are considered, meets the nature of parallelism (Xie and Liu, 2007).

To clearly reflect the similarity of sequences' trends and avoid the impact of orders of magnitude, the average increment operator D_4 is adopted instead of the average operator D_3 which fulfills the requirements of consistency and parallelism. For instance, there are two data sequence $X_1 = (100,150,200,250)$ and $X_2 = (0,50,100,150)$. It is obvious that the increments are equal. In other words, the growth trends of these two sequences are equal. After D_4 processing, increments are still equal, while after D_3 processing, equal increments cannot be obtained that shows that the former operator D_4 can retain the characteristics of the original data better to value the trends of sequences

Then, calculate the increments.

$$\Delta\dot{x}_i^{(0)}(k) = \dot{x}_i^{(0)}(k) - \dot{x}_i^{(0)}(k-1), k = 2,3,\ldots,n, i = 0,1,\ldots,n \quad (24)$$

This step reflects the main idea of slope variances which is the core perspective in this Section to measure the degree of grey correlation.

Third, calculate the incidence coefficient of the interval $[k, k+1]$ which is similar with the degree of grey relational model based on area differences proposed above:

$$\varsigma_{0i}(k) = 1 - \frac{\left|\Delta\dot{x}_i^{(0)}(k) - \Delta\dot{x}_0^{(0)}(k)\right|}{\left|\Delta\dot{x}_i^{(0)}(k)\right| + \left|\Delta\dot{x}_0^{(0)}(k)\right|}, k = 2,3,\ldots,n, i = 1,2,\ldots,n \quad (25)$$

Similarly, $0 \le \varsigma_{0i}(k) = 1 - \frac{\left|\Delta\dot{x}_i^{(0)}(k) - \Delta\dot{x}_0^{(0)}(k)\right|}{\left|\Delta\dot{x}_i^{(0)}(k)\right| + \left|\Delta\dot{x}_0^{(0)}(k)\right|} \le 1$ which means the range of $\varsigma_{0i}(k)$ can cover [0, 1].

Finally, calculate the novel degrees of grey relational model based on slope variances.

$$\delta_{0i} = \frac{1}{n-1}\sum_{k=2}^{n}\varsigma_{0i}(k), k = 2,3,\ldots,n, \ i = 1,2,\ldots,n \quad (26)$$

(3) The properties of the novel grey relational model

After the two parts are obtained, the novel grey relational model can be integrated through the results from the degrees of grey relational model based on area differences and slope variances: $\gamma(X_0^{(0)}, X_i^{(0)}) = 0.5 \cdot (\varepsilon'_{0i} + \delta_{0i}), i = 1,2,\ldots,n$ which has been shown in the beginning of Section 2.2. To prove the effectiveness and practicality of the novel GRA model based on effective information transformation, some properties should be discussed below.

Normativity: $0 \le \gamma(X_0^{(0)}, X_i^{(0)}) \le 1$. As discussed above, the ranges of ε'_{0i} and δ_{0i} are both [0, 1]. Therefore, $0 \le \gamma(X_0^{(0)}, X_i^{(0)}) = 0.5 \cdot (\varepsilon'_{0i} + \delta_{0i}) \le 1$.

Proximity: The smaller $|s_i - s_0|$ and $\left|\Delta\dot{x}_i^{(0)}(k) - \Delta\dot{x}_0^{(0)}(k)\right|$ are, the bigger $\gamma(X_0^{(0)}, X_i^{(0)}) = 0.5 \cdot (\varepsilon'_{0i} + \delta_{0i}), i = 1,2,\ldots,n$ is.

Symmetry: $\gamma(X_0^{(0)}, X_i^{(0)}) = \gamma(X_i^{(0)}, X_0^{(0)})$.

For $\varepsilon'_{0i} = 1 - \frac{|s'_i - s'_0|}{|s'_0| + |s'_i|}$, here is $|s_i - s_0| = |s_0 - s_i|$; for

$\varsigma_{0i}(k) = 1 - \dfrac{\left|\Delta\dot{x}_i^{(0)}(k) - \Delta\dot{x}_0^{(0)}(k)\right|}{\left|\Delta\dot{x}_i^{(0)}(k)\right| + \left|\Delta\dot{x}_0^{(0)}(k)\right|}$, here is $\left|\Delta\dot{x}_i^{(0)}(k) - \Delta\dot{x}_0^{(0)}(k)\right| = \left|\Delta\dot{x}_0^{(0)}(k) - \Delta\dot{x}_i^{(0)}(k)\right|$.

Therefore, $\gamma(X_0^{(0)}, X_i^{(0)}) = \gamma(X_i^{(0)}, X_0^{(0)})$.

By proposing the novel GRA model based on effective information transformation, the aim is to find key factors that relate to the reference sequence by a developed standard which will be discussed in case study. If the degree of GRA is above the standard, the corresponding factor will be selected, otherwise, it will not be chosen. Thus, it is a vital step to pave the way for the next step of quantifying the developing relationships between the reference time sequence and key factors' time sequences.

2.3 The GM(1,N) model

Once the key factors are determined, the GM(1,N) model is introduced to forecast the reference sequence by considering the trends of key factors and the relationships between them. Because it is usually tricky to judge the trend of the reference sequence in the future directly, and through these predictable key factors, the results can be obtained reasonably.

The grey model (GM) (Liu et al., 2010) is an important forecasting method of grey system theory. GM(1, N) is one of the typical models in the grey model group which is represented as GM(1,N) for dealing with 1, the order of a differential equation with N variables. The modeling procedures of GM(1, N) can be carried out as follows.

Assume the original behavior sequences to be

$X_0^{(0)} = (x_0^{(0)}(1), x_0^{(0)}(2), \ldots, x_0^{(0)}(n))$

$X_1^{(0)} = (x_1^{(0)}(1), x_1^{(0)}(2), \ldots, x_1^{(0)}(n))$

$X_2^{(0)} = (x_2^{(0)}(1), x_2^{(0)}(2), \ldots, x_2^{(0)}(n))$

$\vdots$

$X_N^{(0)} = (x_N^{(0)}(1), x_N^{(0)}(2), \ldots, x_N^{(0)}(n))$, $k = 1,2,\ldots,n$.

Based on the initial series, AGO is defined as:

$X_0^{(1)} = (x_0^{(1)}(1), x_0^{(1)}(2), \ldots, x_0^{(1)}(n))$

$X_1^{(1)} = (x_1^{(1)}(1), x_1^{(1)}(2), \ldots, x_1^{(1)}(n))$

$X_2^{(1)} = (x_2^{(1)}(1), x_2^{(1)}(2), \ldots, x_2^{(1)}(n))$

$\vdots$

$X_N^{(1)} = (x_N^{(1)}(1), x_N^{(1)}(2), \ldots, x_N^{(1)}(n))$, $k = 1,2,\ldots,n$,

where $x_i^{(1)}(k) = \sum_{j=1}^{k} x_i^{(0)}(j), i = 1,2,\ldots,N$.

The discrete equation of GM(1, N) can be written as $x_0^{(0)}(k) + az_0^{(1)}(k) = \sum_{i=1}^{N} b_i x_i^{(1)}(k)$, where $k \geq 2$, z is background value, $z_0^{(1)}(k) = 0.5x_0^{(1)}(k) + 0.5x_0^{(1)}(k-1)$, a is a developing coefficient, and b is a control variable. These parameters can be estimated by the following matrix:

$$\hat{a} = \begin{bmatrix} a \\ b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} = (B^T B)^{-1} B^T Y \quad , \quad \text{where} \quad B = \begin{bmatrix} -z_0^{(1)}(2) & x_1^{(1)}(2) & \cdots & x_N^{(1)}(2) \\ -z_0^{(1)}(3) & x_1^{(1)}(3) & \cdots & x_N^{(1)}(3) \\ \vdots & \vdots & \vdots & \vdots \\ -z_0^{(1)}(N) & x_1^{(1)}(N) & \cdots & x_N^{(1)}(N) \end{bmatrix} \quad \text{and}$$

$$Y = \begin{bmatrix} x_0^{(0)}(2) \\ x_0^{(0)}(3) \\ \vdots \\ x_0^{(0)}(N) \end{bmatrix}.$$

The forecasting equation of GM(1, N) is denoted as follows:

$$\hat{x}_0^{(1)}(k+1) = (x_0^{(0)}(1) - \frac{1}{a}\sum_{i=1}^{N} b_i x_i^{(1)}(k+1))e^{-ak} + \frac{1}{a}\sum_{i=1}^{N} b_i x_i^{(1)}(k+1), k = 1,2,\ldots,N-1 \qquad (27)$$

Finally, subtract $\hat{x}_0^{(1)}(k+1)$ consecutively, the forecasting value is :

$$\hat{x}_0^{(0)}(k+1) = \hat{x}_0^{(1)}(k+1) - \hat{x}_0^{(1)}(k) \qquad (28)$$

When N = 1, then GM (1, N) is transformed into a uni-variate forecasting model GM(1, 1), is used to forecast the time series with only one variable. The first order differential equation of GM(1, 1) is: $x_0^{(0)}(k) + a'z_0^{(1)}(k) = b$, where $k \geq 2$. Similarly, z is background value, $z_0^{(1)}(k) = 0.5x_0^{(1)}(k) + 0.5x_0^{(1)}(k-1)$, a' is a developing coefficient, and b' is a control variable. a' and b' are parameters, they are estimated using OLS.

$$\hat{a}' = \begin{bmatrix} a' \\ b' \end{bmatrix} = (B'^T B')^{-1} B'^T Y', \text{ where } B' = \begin{bmatrix} -z_0^{(1)}(2) & 1 \\ -z_0^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z_0^{(1)}(N) & 1 \end{bmatrix} \text{ and } Y' = \begin{bmatrix} x_0^{(0)}(2) \\ x_0^{(0)}(3) \\ \vdots \\ x_0^{(0)}(N) \end{bmatrix}.$$

The forecasting value of GM(1, 1) is denoted as follows:

$$\hat{x}_0^{(0)}(k+1) = (1-e^{a'})(x_0^{(0)}(1) - \frac{b'}{a'})e^{-a'k}, \quad k = 1,2,\ldots,N-1. \tag{29}$$

Both of the GM(1,N) model and the GM(1, 1) model will be applied in case study in Section 3 to forecast the trend of the traffic congestion degree.

For the models' error analysis, the relative percentage error (RPE) is introduced which describes the percentage of difference between the real and the fitting or forecasting values to evaluate the precision at a certain time instance k. The RPE can be defined as

$$RPE(k) = \left| \frac{\hat{x}_0^{(0)}(k) - x_0^{(0)}(k)}{x_0^{(0)}(k)} \right| * 100\%$$

The total precision of a model can be described by calculating the average relative percentage error (ARPE):

$$ARPE = \frac{1}{m-k+1} \sum_{i=k}^{m} RPE(i), m \geq k$$

Note: Given θ, when $ARPE < \theta$ and $RPE(k) < \theta$, the model can be seen as a satisfactory residual model. Usually, less than 5% is considered to be the first level of accuracy.

2.4 The novel grey relational analysis model and GM(1,N) model of interval grey numbers

As the novel GRA model based on effective information transformation and the forecasting model of GM(1,N) and GM(1, 1) have been proposed, this paper extend them to the application field of interval grey numbers to describe the regional character of the data. The modeling process is shown below:

Let $\otimes \in [a,b]$, $a < b, a, b \in R$, then $\otimes$ is called interval grey number; if $a = b$, then $\otimes$ is real number. For interval grey number sequences $X^{(0)}(\otimes_i) = (\otimes_i^{(0)}(1), \otimes_i^{(0)}(2), \ldots, \otimes_i^{(0)}(n)), i = 0,1,\ldots,n$, $\otimes_i(k) \in [a_i(k), b_i(k)]$, its upper bound sequences and lower bound sequences are $B_i^{(0)} = (b_i^{(0)}(1), b_i^{(0)}(2), \ldots, b_i^{(0)}(n))$ and $A_i^{(0)} = (a_i^{(0)}(1), a_i^{(0)}(2), \ldots, a_i^{(0)}(n))$ $, i = 0,1,\ldots,n$ respectively. Among them, $B_0^{(0)} = (b_0^{(0)}(1), b_0^{(0)}(2), \ldots, b_0^{(0)}(n))$ and $A_0^{(0)} = (a_0^{(0)}(1), a_0^{(0)}(2), \ldots, a_0^{(0)}(n))$ are reference

sequences, while $B_i^{(0)} = (b_i^{(0)}(1), b_i^{(0)}(2), \ldots, b_i^{(0)}(n))$ and $A_i^{(0)} = (a_i^{(0)}(1), a_i^{(0)}(2), \ldots, a_i^{(0)}(n))$ $, i = 1,2,\ldots,n$ are behavior sequences. For interval grey numbers, the upper and lower bound sequences have different trends, so they need to be modeled respectively. The modeling process of the novel GRA model and GM(1,N) model is as follows:

Step 1: calculate the degrees of GRA based on area differences between the reference sequence and the behavior sequences. For the upper bound sequences, $\varepsilon_{0i}^U = 1 - \frac{\left|s_0^U - s_i^U\right|}{\left|s_0^U\right| + \left|s_i^U\right|}$, $, i = 1,2,\ldots,n$. For the lower bound sequences, $\varepsilon_{0i}^L = 1 - \frac{\left|s_0^L - s_i^L\right|}{\left|s_0^L\right| + \left|s_i^L\right|}, i = 1,2,\ldots,n$.

Step 2: calculate the degrees of GRA based on slope variances between the reference sequence and the behavior sequences. For the upper bound sequences, $\delta_{0i}^U = \frac{1}{n-1}\sum_{k=2}^{n} \varsigma_{0i}^U(k)$, $, i = 1,2,\ldots,n$. For the lower bound sequences, $\delta_{0i}^L = \frac{1}{n-1}\sum_{k=2}^{n} \varsigma_{0i}^L(k)$, $i = 1,2,\ldots,n$.

Step 3: Integrate the novel grey relational analysis model based on effective information transformation between the reference sequence and the behavior sequences. For the upper bound sequences, $\gamma(B_0^{(0)}, B_i^{(0)}) = 0.5 \cdot (\varepsilon_{0i}^U + \delta_{0i}^U)$, $i = 1,2,\ldots,n$. For the lower bound sequences, $\gamma(A_0^{(0)}, A_i^{(0)}) = 0.5 \cdot (\varepsilon_{0i}^L + \delta_{0i}^L)$, $i = 1,2,\ldots,n$.

Step 4: Determine key factors. Through the values of $\gamma(B_0^{(0)}, B_i^{(0)}) = 0.5 \cdot (\varepsilon_{0i}^U + \delta_{0i}^U)$ and $\gamma(A_0^{(0)}, A_i^{(0)}) = 0.5 \cdot (\varepsilon_{0i}^L + \delta_{0i}^L)$, a standard value should be set. If the values in $\gamma(B_0^{(0)}, B_i^{(0)}) = 0.5 \cdot (\varepsilon_{0i}^U + \delta_{0i}^U)$ and $\gamma(A_0^{(0)}, A_i^{(0)}) = 0.5 \cdot (\varepsilon_{0i}^L + \delta_{0i}^L)$ are larger than the standard value, the corresponding behavior sequences should be chosen as key factors.

Step 5: Establish GM(1,1) models of the sequences from key factors respectively to obtain the modeling values of GM(1,N) model in the next step.

Step 6: Based on the original reference sequence and the sequences from key factors, the

GM(1,N) model should be proposed to forecast the trend of the reference sequence in the future.

3. Case study

The Yangtze River Delta includes Shanghai City, Jiangsu Province and Zhejiang Province which is the largest economic zone in Mainland China and has been internationally recognized as one of the six world-class city groups. As the largest economic circle in Mainland China, the Yangtze River Delta region has created the equivalent of 20% of Mainland China's GDP, even though it accounts for just 1% of Mainland China's land area. Thus, the Yangtze River Delta region plays a decisive role in China's social and economic development process. With the development of urbanization and the continuous expansion of urban scale in the Yangtze River Delta, traffic congestion has become increasingly prominent and has become an 'urban disease' affecting the healthy development of cities. The study of urban traffic congestion in the Yangtze River Delta region, not only has a clear practical significance at present, but it also has a long term positive effect on government policy adjustment and urban development planning.

Applying the methodologies described in Section 2, the key factors for traffic congestion of the Yangtze River Delta is analyzed and the traffic performance in the near future is predicted in this section. This is organized as follows: firstly, we introduce the variables and list each variable's data. We secondly discuss the relationship between the degree of traffic congestion and influencing factors by using the novel GRA model proposed in this paper to find the key factors for forecasting, and then compare the results with those of traditional grey relational analysis and principal component analysis. Thirdly, we establish GM(1, N) by using the results from this paper's method to predict the degree of traffic congestion in the Yangtze River Delta. To make a comparison, some traditional regression models are discussed.

3.1 Data description

The data of the Yangtze River Delta during 2007-2015, which come from the National Statistics Yearbook of China (2008-2016), Statistical Yearbooks of Shanghai, Jiangsu and Zhejiang (2008-2016), are chosen for this research. These data (variables) include: Traffic Congestion Degree in urban areas (TCD), Consumption Level of Urban Residents (CLUR), Urban Population Density (UPD), Public Transport vehicles per million people (PTP) and urban Road Area per ca pita (RAP). The interval grey number sequences for Shanghai, Jiangsu and Zhejiang cover the following variables: TCD (vehicle number per kilometer) refers to traffic density which is equal to the number of vehicles owned by civilians divided by distance travelled at the end of each year; CLUR(CNY per ca pita) is calculated from the total consumption of urban residents in GDP of the reporting period, divided by the annual average population of the reporting period; UPD (people per square kilometer) is obtained from the population, which is the sum of urban population and urban temporary resident population, divided by urban areas; PTP (standard unit) is the ratio of public transport vehicles and population, which is the same as the population in UPD; RAP (square meter) is calculated with urban road areas divided by the sum of the urban population and temporary urban residents.

3.2 The novel grey relational analysis between TCD and CLUR, UPD, PTP, RAP

Taking the related data of the Yangtze River Delta, $X_0^{(0)}$ denotes the TCD which is the reference sequence; $X_1^{(0)}$ denotes the CLUR which is one of the comparison sequences; $X_2^{(0)}$

denotes the UPD which is one of the comparison sequences; $X_3^{(0)}$ denotes the PTP which is one of the comparison sequences; $X_4^{(0)}$ denotes the RAP which is one of the comparison sequences. The historical data are shown in Table 1.

Table 1

The historical data of all variables (L for the lower bound, U for the upper bound)

Year	**TCD($X_0^{(0)}$)**		**CLUR($X_1^{(0)}$)**		**UPD($X_2^{(0)}$)**		**PTP($X_3^{(0)}$)**		**RAP($X_4^{(0)}$)**	
	L	U	L	U	L	U	L	U	L	U
2007	429.21	637.53	13165	25919	1748	2930	11.34	12.88	4.5	19.28
2008	469.29	657.65	14930	29250	1757	2978	12.41	13.24	4.63	20.28
2009	456.64	692.79	15965	31608	1742	3030	12.75581	13.70479	4.48	20.42
2010	433.20	735.36	18243	34588	1773	3630	12.27407	13.62774	4.04	21.26
2011	449.08	734.52	21598	37558	1741	3702	11.79232	13.55069	4.04	21.86
2012	437.03	740.47	24101	39095	1786	3754	11.91	13.96	4.08	22.35
2013	446.05	757.88	28753	41464	1818	3809	12.11	14.64	4.11	23.22
2014	429.70	793.92	32186	45352	1828	3826	11.97	15.46	4.11	23.89
2015	417.05	788.28	33358	48750	1914	3809	12.36	15.99	4.27	24.42

First of all, the reference series $X_0^{(0)}$ is determined and $X_1^{(0)}$, $X_2^{(0)}$, $X_3^{(0)}$, $X_4^{(0)}$ are used as behavior series. According to the GRA model based on effective information transformation proposed in Section 2.2, the following results are obtained.

The grey relational degrees of GRA model based on area differences are $\varepsilon_{01}^{L}=0.1443$, $\varepsilon_{02}^{L}=0.3210$, $\varepsilon_{03}^{L}=0.6517$, $\varepsilon_{04}^{L}=0.1002$; $\varepsilon_{01}^{U}=0.5878$, $\varepsilon_{02}^{U}=0.8158$, $\varepsilon_{03}^{U}=0.8071$, $\varepsilon_{04}^{U}=0.9325$. And the grey relational degrees of grey relational model based on slope variances are $\delta_{01}^{L}=0.7078$, $\delta_{02}^{L}=0.6323$, $\delta_{03}^{L}=0.8272$, $\delta_{04}^{L}=0.4899$; $\delta_{01}^{U}=0.6558$, $\delta_{02}^{U}=0.5085$, $\delta_{03}^{U}=0.5892$, $\delta_{04}^{U}=0.5531$. Then, synthesize the degrees of GRA model based on effective information transformation of the lower bound sequence and the upper bound sequence respectively and the results are shown in Table 2. The relationships of determinants in terms of the synthesized degrees are ranked by: $\gamma(B_0^{(0)},B_4^{(0)})>\gamma(B_0^{(0)},B_3^{(0)})>\gamma(B_0^{(0)},B_1^{(0)})>\gamma(B_0^{(0)},B_2^{(0)})$ (for the upper bounds of sequences) and $\gamma(A_0^{(0)},A_3^{(0)})>\gamma(A_0^{(0)},A_2^{(0)})>\gamma(A_0^{(0)},A_1^{(0)})>\gamma(A_0^{(0)},A_4^{(0)})$ (for the lower bounds of sequences). The results mean that the important sequence of the determinants are ranked by RAP>PTP>CLUR>UPD (for the upper bounds of sequences) and

PTP>UPD>CLUR>RAP (for the lower bounds of sequences).

To compare with the novel grey relational method, the results of traditional relative grey degrees (Section 2.1) and principal component analysis are also calculated (shown in Table 2).

Table 2

The relational result comparisons with TCD

Method	**CLUR**	**UPD**	**PTP**	**RAP**
The upper bound	U	U	U	U
The traditional relative grey relational model	0.7242	0.7057	0.7407	0.8957
The grey relational model based on effective information	0.6433	0.6362	0.6853	0.7406
Principal component analysis	0.977	0.973	0.886	0.960
The lower bound	L	L	L	L
The traditional relative grey relational model	0.5836	0.8520	0.8182	0.6527
The grey relational model based on effective information	0.4407	0.5340	0.7734	0.2722
Principal component analysis	-0.344	-0.406	0.637	0.532

It is obvious that the results of the traditional relative grey degrees are all above 0.5, even in lower bound sequences between TCD and CLUR (shown as extremely different in Fig.2), which means the traditional model has low resolution. In addition, the traditional relative grey degrees only consider the area differences, while the slope and area differences are all taken into account in the novel method proposed in this paper. For instance, see the Fig.2 below, the UPD line and the PTP line are all very close to the TCD line. If we only consider the area differences, the UPD line is closer that is why the result of UPD based on traditional relative grey degrees is a little bigger (0.8520>0.8182). However, if the slope differences are also considered which represent the development trends, the PTP line is more similar with the TCD line, so the result of UPD based on the novel grey relational method is much bigger (0.7734>0.5340). To sum up, the novel grey relational analysis method is superior to the traditional relative grey degree model both in resolution and comprehensive view. From a broader point of view of traditional statistical methods, principal component analysis is commonly used to explore the relationship between factors. From the results in Table 2, the associated values of the upper bound and lower bound data don't pass the KMO test whose KMO values are 0.622 and 0.424. Thus, the results are ineffective. It can be explained that the traditional statistical methods are typically based on a certain sample size which are not suitable for small samples.

Next, the vital step is to choose the key factors based on the values of $\gamma(A_0^{(0)}, A_i^{(0)})$ and $\gamma(B_0^{(0)}, B_i^{(0)})$, $i = 1,2,3,4$. However, there is no certain standard to determine distinguishing

coefficients, in order to help choose a factor. Kuo et al. (2008) discussed the differences by choosing different distinguishing coefficients, while Hsu (2009) chose 0.89 as the distinguishing coefficient. In this paper, we do not try to give the standard which may depend on each specific case, but usually the distinguishing values should be above 0.6. Therefore, according to the results of the novel grey relational model in Table 2, the grey variables are chosen to construct the grey forecasting model (GM (1, 2) and GM (1, 5)). For the lower bound sequences, **PTP** is selected (when $\gamma(A_0^{(0)}, A_i^{(0)}) > 0.6$). For the upper bound sequences, **RAP, PTP, CLUR and UPD** are selected (when $\gamma(B_0^{(0)}, B_i^{(0)}) > 0.6$).

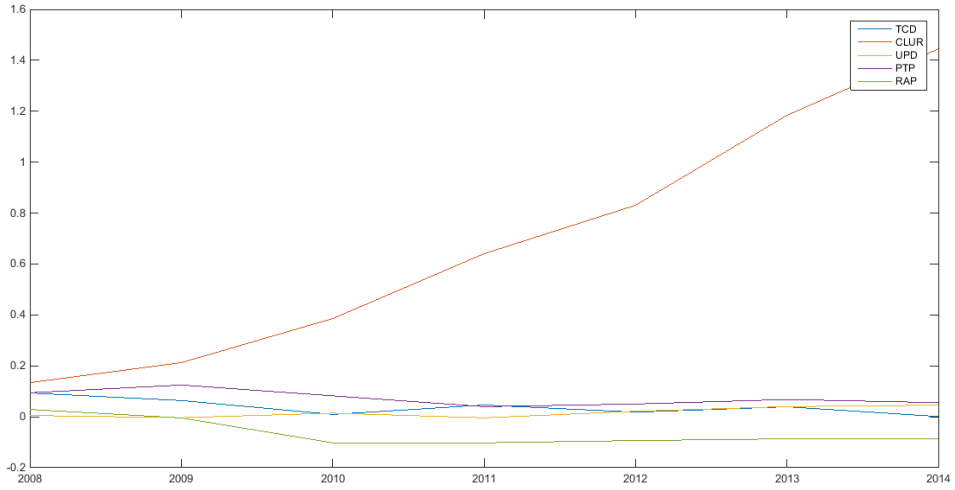

Fig. 2. The sequences of the lower bound after the initial operator process and the start zeroized operator process by the traditional relative grey degree

3.3 The forecast of traffic congestion degree by using GM(1,N)

Because the reference sequence is not easy to obtain directly, it is an effective way to make full use of relevant information of its influencing key factors to predict the reference sequence. In Section 3.2, the key factors of the reference sequences have been determined. Next, the GM(1,N) model should be established to forecast the sequence by using the key factors' sequences. According the model described in Section 2.3, the GM(1,N) model is built as follows.

Step 1**:** Predict the key factors' sequences respectively to obtain the values of 2015-2020 by using GM (1, 1). For the lower bound sequences, the data of PTP (2007-2014) are used. For the upper bound sequences, the data of RAP, PTP, CLUR and UPD (2007-2014) are used.

Step 2**:** Establish the GM (1,N) model. According to the results in Table 2, the GM (1,2) (2007-2014) is obtained for the lower bound sequences and the GM (1,5) (2007-2014) is obtained for the upper bound sequences.

Step 3**:** Predict the degrees of traffic congestion in 2015-2020. Among them, the data of 2007 – 2014 are for modeling, the data of 2015 is for model validation, and the data of 2016 – 2020 are for forecasting the future trends.

Combing the forecasting values in Step 2 and the GM (1,N) models in Step 3, the results are listed in Table 3 and Table 4. In addition, in order to illustrate the simulating results and future trends, the fitting sequences and the forecasting results by the GM (1,2) for the lower bound

sequence and the GM (1, 5) for the upper bound sequence are drawn in Fig. 3 and Fig. 4 respectively. Besides, multiple linear regression models and curve regression models (include Linear Type, Logarithmic Type, Quadratic Type, S Type, Exponential Type) are applied to compare with the results of GM models. However, the results obtained in these traditional statistical models are not effective. To be specific, for the lower and upper bound sequences, the results of binary linear regression, multiple linear regression model and curve regression models (include Linear Type, Logarithmic Type, Quadratic Type, S Type, Exponential Type) all failed in the corresponding tests. In binary and multiple linear regressions for the lower and upper bound respectively, the significant factors (sig. values) of the correlation key factors and the constant coefficient are all greater than the confidence level of 0.05, which means the established regression equations are invalid. In curve regressions for the lower bound sequences, the goodnesses of fit are all very small (around 0.15), which should be better when the value is closer to 1, which means the models are also invalid. To sum up, multiple linear regression models and curve regression models of the traditional statistical methods are all not applicable here in this case.

Table 3

The simulating values of TCD by the GM(1, N)

Year	The lower bound simulating values	The upper bound simulating values
	GM(1,2)	GM(1, 5)
2007	429.21	637.53
2008	390.64	583.07
2009	504.76	777.39
2010	458.69	774.76
2011	433.20	736.02
2012	436.12	745.36
2013	443.23	766.28
2014	438.07	790.79
The *ARPE* of fitting error	4.94%	3.91%

Table 4

The forecasting values of TCD by the GM(1, N)

Year	The lower bound forecasting values	The upper bound forecasting values
	GM(1,2)	GM(1, 5)
2015	425.24	868.18
2016	431.91	827.90
2017	428.58	854.01
2018	430.54	869.70
2019	431.23	890.04
2020	428.00	910.20

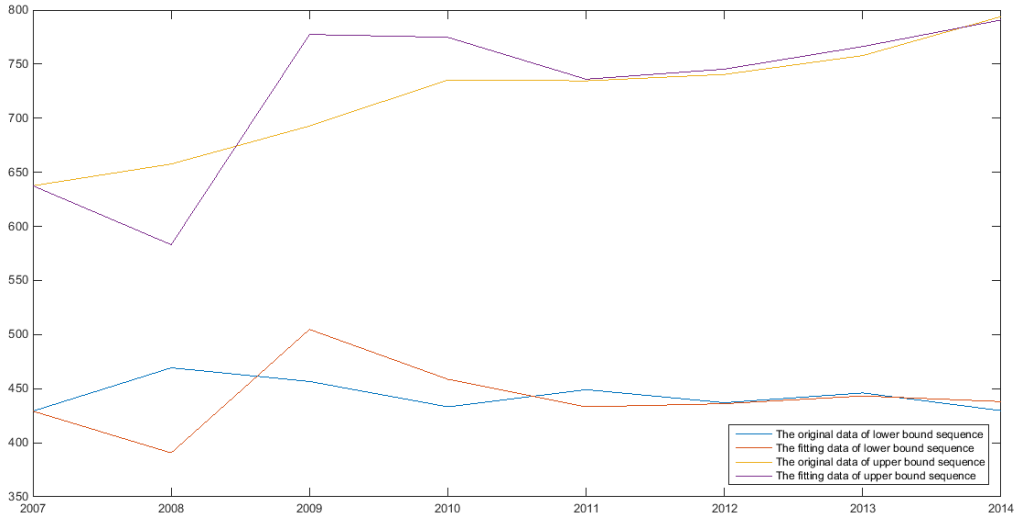

Fig. 3. The sequences of the original values and the fitting values

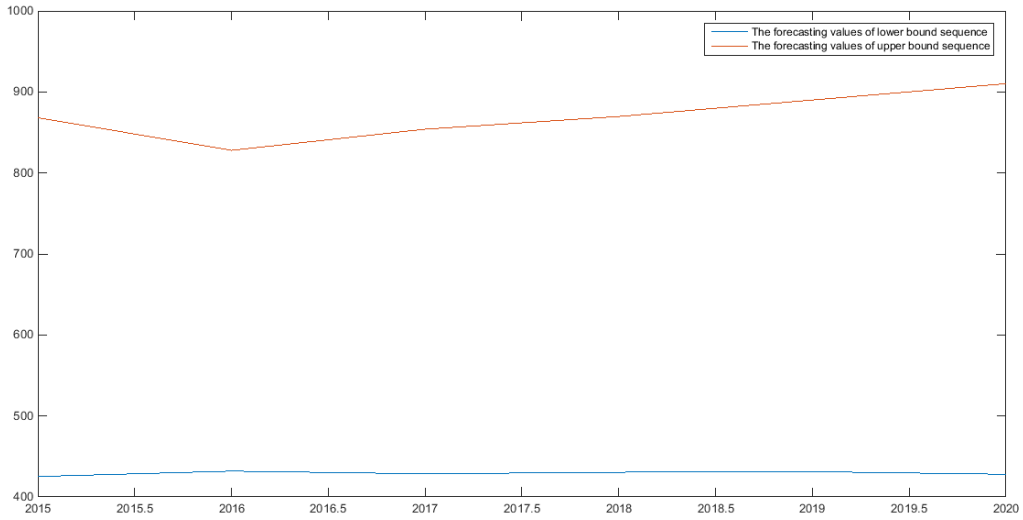

Fig. 4. The sequences of the forecasting values

Seen from Table 3, the average fitting errors are 4.94% and 3.91% (all below 5%), that shows excellent performance by using the GM (1, N). Specifically seen from Fig. 3, apart from fitting fluctuations in 2008 and 2009, the overall trends of the original data and the fitting data almost overlap whenever it comes to the lower bound sequence or the upper bound sequence. Therefore, the model proposed in this paper has verified its superb effectiveness and accuracy. Considering that the fitting results are great which reach the first level of accuracy, it is suitable for short-term interval grey nembers' processing by establishing the proposed model. However, the data are not suitable for classical mathematical and statistical models (requiring 15 observations or more).

Turning to the forecasting results (shown in Table 4 and Fig. 4), for the lower bound sequence, there is only one reverent factor (PTP) whose grey relational degrees is greater than 0.6, so the GM (1, 2) is proposed with this degree of traffic congestion. The forecasting error of 2015

is only 1.96% which means the proposed model effectively predicts the future trend of the lower bound. On the whole, the results reveals that the forecasting values of the smallest traffic congestion degree in the Yangtze River Delta shows slight fluctuations for 2015-2020. But the absolute values remain in the range of 425.24 and 431.91, which means the current traffic level will be basically stable in the near future.

For the upper bound sequence, the reverent factors (RAP, PTP, CLUR, UPD) whose grey relational degrees are greater than 0.6 ($\gamma(B_0^{(0)}, B_i^{(0)}) > 0.6$) are chosen together with the degree of traffic congestion to establish the GM (1, 5). Seen from the predicted results, the value of 2015 appears to have irregular volatility which is mainly due to the surge in 2013-2014. Thus, the forecasting error is a bit large, but it is acceptable (around 10%). After 2015, the trend presents a rising state that increases by about 20 vehicles per kilometer each year, and reaches 910.20 vehicles / km in 2020. These mean that the trend for the largest traffic congestion degree in the Yangtze River Delta will grow at a slightly faster rate in the next five years than the current expansion rate.

4. Conclusions

In this paper, the novel GRA model based on effective information transformation of interval grey numbers has been proposed and the GM(1,N) model has been established to forecast traffic congestion degree (TCD) by using key factors which are determined by the novel GRA model. This GRA model not only fully considers the information of area differences and slope variances, but also optimizes the resolutions of the traditional grey degrees. Furthermore, to characterize regional data, interval grey numbers are used in the whole process of modeling, which extends the application fields of the traditional grey models instead of focusing on the trend of real numbers in most of the existing papers using grey models. Finally, the traffic congestion degree of the Yangtze River Delta is discussed as a case in this paper. The results show high resolution of the proposed GRA model and describe the development trend within this area in the future. In general, the method proposed in this paper is very suitable for dealing with the recent information problem of small amounts of data that are more sensitive in capturing the characteristics of new information in the rapidly changing environment. Moreover, this novel GRA model based on effective information transformation is fast-processing which is suitable for promotion to the fields of real-time analysis. On the contrary, complex program algorithms often require large amounts of data to run the process which is very time-consuming. In summary, the method shown in this paper demonstrates the advantages in information trawling of small amounts data and to an extent, the processing of instant analysis.

5. Acknowledgments

The authors are grateful to anonymous referees for their helpful and constructive comments on this paper. This work was supported by a Marie Curie International Incoming Fellowship within the 7th European Community Framework Programme entitled “Grey Systems and Its Application to Data Mining and Decision Support” Grant No. FP7-PIIF-GA-2013-629051, a project of the Leverhulme Trust International Network entitled “Grey Systems and Its Applications”(IN-2014-020).The authors would also like to acknowledge the support of the National Natural Science Foundation of China (71771119) and Nanjing University of Finance & Economics. Jing Ye declares that she has no conflict of interest. Yaoguo Dang declares that he has

no conflict of interest. Yingjie Yang declares that he has no conflict of interest. This article does not contain any studies with human participants or animals performed by any of the authors.

References

Bahrami S, Hooshmand R, Parastegari M (2014) Short term electric load forecasting by wavelet transform and grey model improved by PSO (particle swarm optimization) algorithm. Energy 72: 434-442

Chen C, Huang S (2013) The necessary and sufficient condition for GM(1,1) grey prediction model. Applied Mathematics and Computation 219(11): 6152-6162

Erdal K, Baris U, Okyay K (2010) Grey system theory-based models in time series prediction. Expert Syst Appl 37(2): 1784–1789

Evans M (2014) An alternative approach to estimating the parameters of a generalised Grey Verhulst model: An application to steel intensity of use in the UK. Expert Systems with Applications 41(4): 1236-1244

Guo H, Xiao XP, Jeffrey F (2013) Urban Road Short-term Traffic Flow Forecasting Based on the Delay and Nonlinear Grey Model. Journal of Transportation Systems Engineering and Information Technology 13(6): 60-66

Hsu LC (2009) Forecasting the output of integrated circuit industry using genetic algorithm based multivariable grey optimization models. Expert Systems with Applications 36(4): 7898-7903

Hsu L, Wang C (2009) Forecasting integrated circuit output using multivariate grey model and grey relational analysis. Expert Systems with Applications 36(2): 1403-1409

Kadier A, Abdeshahianb P, Simayic Y, Ismaila M, Hamide AA, Kalila MS (2015) Grey relational analysis for comparative assessment of different cathode materials in microbial electrolysis cells. Energy 90: 1556-1562

Kayacan E, Ulutas B, Kaynak O (2010) Grey system theory-based models in time series prediction. Expert Systems with Applications 37(2): 1784-1789

Kuo Y, Yang T, Huang GW (2008) The use of grey relational analysis in solving multiple attribute decision-making problems. Computers & Industrial Engineering 55(1): 80-93

Li D, Chang C, Chen C, Chen W (2012) Forecasting short-term electricity consumption using the adaptive grey-based approach—An Asian case. Omega 40(6): 767-773

Li XM, Dang YG, Wang JJ (2015) Grey generation rate relational analysis model based on grey exponential law and its application. Control and Decision 30(7): 1245-1250

Lin YH, Chiu CC, Lee PC, Lin YJ (2012) Applying fuzzy grey modification model on inflow forecasting. Engineering Applications of Artificial Intelligence 25: 734-743

Liu J, Xiao X, Guo J, Mao S (2014) Error and its upper bound estimation between the solutions of GM(1,1) grey forecasting models. Appl Math Comput 246: 648-660

Liu SF, Dang YG, Fang ZG, Xie NM (2010) Grey system theory and application (the 5th edition). Beijing: Science Press

Liu SF, Fang ZG, Lin Y (2006) Study on A new Definition of Degree of Grey Incidence. Journal of Grey System 9(2): 115-122

Liu SF, Lin Y (2010) Grey systems: theory and applications. London: Springer-Verlag London Ltd

Mohammadi SE, Makui A (2017) Multi-attribute group decision making approach based on interval-valued intuitionistic fuzzy sets and evidential reasoning methodology. Soft Computing 21(17): 5061–5080

Pao H, Fu H, Tseng C (2012) Forecasting of CO_2 emissions, energy consumption and economic growth in China using an improved grey model. Energy 40(1): 400-409

Rehborn H, Klenov SL, Palmer J (2011) An empirical study of common traffic congestion features based on traffic data measured in the USA, the UK, and Germany. Physica A: Statistical Mechanics and its Applications 390(23-24): 4466-4485

Shankar H, Raju PLN, Rao KRM (2012) Multi model criteria for the estimation of road traffic congestion from

traffic flow information based on Fuzzy logic. J Transp Technol 2: 50-62

Ujjwal K, Jain VK (2010) Time series model(Grey–Markov, Grey model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India. Energy 35(4): 1709–1716

Wang YH, Dang YG, Li YQ, Liu SF (2010) An approach to increase prediction precision of GM(1,1) model based on optimization of the initial condition. Expert Systems with Applications 37(8): 5640-5644

Wang ZW, Lei TZ, Chang X, Shi XG, Xiao J, Li ZF, He XF, Zhu JL, Yang SH (2015) Optimization of a biomass briquette fuel system based on grey relational analysis and analytic hierarchy process: A study using cornstalks in China. Applied Energy 157: 523-532

Wei GW (2011) Grey relational analysis model for dynamic hybrid multiple attribute decision making. Knowledge-Based Systems 24(5): 672-679

Wu LF, Liu SF, Yao LG, Yan SL (2013a) The effect of sample size on the grey system model. Applied Mathematical Modelling 37: 6577–6583

Wu LF, Liu SF, Yao LG, Yan SL, Liu DL (2013b) Grey system model with the fractional order accumulation. Commun Nonlinear Sci 18(7): 1775-1785

Wu LF, Liu SF, Fang ZG, Xu HY (2015a) Properties of the GM(1,1) with fractional order accumulation. Applied Mathematics and Computation 252: 287-293

Wu LF, Liu SF, Liu DL, Fang ZG, Xu HY (2015b) Modelling and forecasting CO_2 emissions in the BRICS (Brazil, Russia, India, China, and South Africa) countries using a novel multi-variable grey model. Energy 79: 489-495

Wu LF, Liu SF, Yao LG, Xu RT, Lei XP (2015c) Using fractional order accumulation to reduce errors from inverse accumulated generating operator of grey model. Soft Comput 19: 483-488

Wu LF, Liu SF, Yang YJ (2016) Grey double exponential smoothing model and its application on pig price forecasting in China. Applied Soft Computing 39: 117-123

Xia M, Wong WK (2014) A seasonal discrete grey forecasting model for fashion retailing. Knowledge-Based Systems 57: 119-126

Xie NM, Liu SF (2007) The Parallel and Uniform Properties of Several Relational Models. System Engineering 25(8): 98-103

Xu L, Yue Y, Li QQ (2013) Identifying Urban Traffic Congestion Pattern from Historical Floating Car Data. Procedia-Social and Behavioral Sciences 96: 2084-2095

Younes MB, Boukerche A (2015) A performance evaluation of an efficient traffic congestion detection protocol (ECODE) for intelligent transportation systems. Ad Hoc Netw 24: 317-336