

Research Article

Demand Prediction of Emergency Supplies under Fuzzy and Missing Partial Data

Ming Zhang , Hanlin Wu, Zhifeng Qiu, Yifan Zhang, and Boquan Li

College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Correspondence should be addressed to Ming Zhang; zhangming_nuaa@126.com

Received 13 December 2018; Revised 31 March 2019; Accepted 24 April 2019; Published 9 May 2019

Academic Editor: Luca Pancioni

Copyright © 2019 Ming Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

An accurate demand prediction of emergency supplies according to disaster information and historical data is an important research subject in emergency rescue. This study aims at improving supplies demand prediction accuracy under partial data fuzziness and missing. The main contributions of this study are summarized as follows. (1) In view that it is difficult for the turning point of the whitenization weight function to determine fuzzy data, two computational formulas solving “core” of fuzzy interval grey numbers were proposed, and the obtained “core” replaced primary fuzzy information so as to reach the goal of transforming uncertain information into certain information. (2) For partial data missing, the improved grey k-nearest neighbor (GKNN) algorithm was put forward based on grey relation degree and K-nearest neighbor (KNN) algorithm. Weights were introduced in the filling and logic test conditions were added after filling so that filling results were of higher truthfulness and accuracy. (3) The preprocessed data are input into the improved algorithm based on the genetic algorithm and BP neural networks (GABP) to obtain the demand prediction model. Finally the calculation presents that the prediction accuracy and its stability are improved at the five-group comparative tests of calculated examples of actual disasters. The experiments indicated that the supplies demand prediction model under data fuzziness and missing proposed in this study was of higher prediction accuracy.

1. Introduction

During an earthquake disaster, adverse situations occur, such as urgent rescue time, limited manpower and material resources, and chaotic scenes. This situation requires timely demand prediction of emergency rescue supplies with strong phase features. However, this task is affected by various factors, such as social and environmental factors. The fuzziness and uncertainty of information in this situation intensify the difficulty of demand prediction and greatly affect the accuracy and truthfulness of the prediction result. An accurate demand prediction of rescue supplies can save precious time for follow-up rescue work (such as rescue scheduling) [1]. This critical problem needs urgent solutions from emergency management department and planning researchers and has important research values.

The supplies demand prediction problem under partial data fuzziness and missing is studied. In view that it is difficult for the turning point of the whitenization weight function to determine fuzzy data, two computational formulas solving “core” of fuzzy interval grey numbers are proposed, and

the obtained “core” replaced primary fuzzy information so as to reach the goal of transforming uncertain information into certain information; for partial data missing, the improved GKNN algorithm is put forward based on grey relation degree and K-nearest neighbor algorithm. Weights are introduced in the filling and logic test conditions are added after filling so that filling results were of higher truthfulness and accuracy. The preprocessed data are input into the neural network model optimized through the improved genetic algorithm to obtain the trained emergency supplies demand prediction model which is then tested and became optimal after adjustment of model parameters. Therefore, this study was aimed to investigate the most effective method to be used in emergency supplies demand prediction. The main contributions and innovative points of this study are summarized as follows:

- (1) Grey system theory is introduced for fuzzy information. Given that the turning points of the whitenization weight function cannot be determined directly, two computing formulas for solving the “core” of fuzzy interval grey numbers are established. The

obtained “core” is used as a replacement to the primary fuzzy information to transform uncertain information into certain information.

- (2) Because the performance of the traditional GKNN algorithm is greatly affected by the initial K value, resulting in large error fluctuations in the filling results, this study is to improve the GKNN algorithm by adding an innovative weight function to the algorithm and reducing the dependence of the algorithm on the initial K value for missing information. Compared with the traditional GKNN algorithm, the improved GKNN algorithm can reduce the maximum average relative error by more than 20%.
- (3) This study is to modify the adaptive function of crossover probability and mutation probability of the genetic algorithm in GABP model, so that the model can better find the initial weight and threshold. Finally the calculation presents that the prediction accuracy is improved by 20.31% ~ 69%, and its stability is improved by 24.1% ~ 88.21% at the five-group comparative tests of calculated examples of actual disasters.

The remainder of the study is organized as follows. In Section 2, we review the related literature of solutions to the problem of fuzzy and missing disaster information and optimization method of the demand prediction. Section 3 presents improved models and algorithm for demand prediction of emergency supplies under fuzzy and missing partial data. Section 4 establishes five groups of experiment through actual calculated examples for determining the performance of the demand prediction model. Section 5 elaborates the conclusions and expectations of this study.

2. Related Work

We survey the related literature from three aspects in this section. Section 2.1 presents a literature review on solving the problem of fuzzy partial data. Section 2.2 reviews main existing methods for solving the problem of missing partial data. Section 2.3 details the related research on GA algorithms for optimization of demand prediction.

2.1. Solving the Problem of Fuzzy Partial Data. Grey system theory [2] focuses on the system containing certain and uncertain information. It aims at studying the development of uncertain and unknown information in the system into certain and known information, and the effect it reaches accords considerably with the goal of this study.

Existing related studies have used fuzzy information processing using grey system theory. Liu et al. [3] established operation axiom and rules of interval grey numbers and transformed grey number operation into real number operation. For the convenience of discussion, whitenization number of mean grey number was taken as the core of grey number in the current study. Whereafter interval grey numeral sequence prediction to a certain degree [4] and multiattribute-group decision-making method [5] were put

forward. On the basis of traditional heterogeneous discrete grey prediction model, Xie and Liu [6] redefined interval grey number and its algebraic operation and constructed a new interval grey numeral sequence prediction method combined with the nonhomogenous discrete grey forecasting model.

Meanwhile, Zeng et al. [7], Dang et al. [8], Ye and Dang [9], and Wang [10] defined the standard form of interval grey numbers and established an extended correlation analysis method for incomplete information sequences.

In the present study, whitenization weight function of interval grey numbers is determined to realize whitenization processing of fuzzy interval grey numbers and solve the core of interval grey numbers. The whitenization weight function type is mainly decided by location and number of turning points. No unified definite method for selection of turning points of the whitenization weight function exists at present, and turning points are mainly subjectively determined by researchers depending on related grey value distribution information under the research background. For example, Sun and Li [11] studied the whitenization problem in multicriteria decision making and pointed out that turning points of the whitenization weight function were determined by researchers through their subjective judgment based on objective information. No concrete and definite computational formula is available to help researchers determine turning points. Most methods use equal-weighted mean value whitenization method or avoid concrete selection process of turning points while directly taking the whitenization weight function as a known condition for the research. The core formula of interval grey numbers should be solved on the condition that the locations of turning points of the whitenization weight function are uncertain to solve whitenization processing problems of fuzzy interval grey numbers under insufficient grey value distribution information.

2.2. Solving the Problem of Missing Partial Data. Main existing methods for solving the problem of missing emergency supply demand data are single-imputation, multiple-imputation, and modeling methods. Among single-imputation methods, random imputation, and mean value imputation have simple operations but neglect information correlation and excessively simplified problems. Thus, they are inapplicable to completion of missing disaster information. Another representative single-imputation method is regressive imputation [2].

Given that each disaster case has different environmental backgrounds, a universal regression model for multiple disaster case properties is unsuitable. Multiple-imputation method is developed on the basis of single-imputation method, and its main idea is to repeatedly execute the same operation for several times to obtain multiple and complete candidate data sets. However, which method should be selected as the repeatedly executed operation needs specific determination. Modeling methods must analyze the mechanism of missing data and then establish a model for properties that require filling. Nevertheless, these methods are inapplicable to small sample sizes.

Related works have studied missing emergency supply demand information. For example, Mohammed et al. [12]

used KNN algorithm to seek for optimal results of vehicle routing problems for realizing optimized solution of the capacitated vehicle routing problem model while reducing route distance and arrival time and accelerating the client's arrival at the destination. Suguna and Thanushkodi [13] combined genetic algorithm (GA) and KNN to improve classification performance of the hybrid algorithm. On the basis of KNN, Meng et al. [14] sorted missing rates of data in an ascending order to complete the filling work and expanded the filled data into the adjacent selection set. Žukovič and Hristopoulos [15] put forward a Directional Gradient-Curvature method based on an objective function. Zhang et al. [16] indicated that grey relation degree was more suitable than Euclidean distance or other distances to calculate similarities between two samples.

Grey relational analysis (GRA) [17–19] could reflect the tendency relation between the alternative and ideal alternative solutions effectively but could not reflect the situational relation between them. In the present study, values of distinguishing coefficients are redefined to ensure that they sufficiently reflect system completeness while reserving the interfering effect. However, the computing method [20, 21] for distinguishing coefficients in this study may be inapplicable to other fields.

In summary, the similarities of grey relation degrees between measuring objects are reasonable. However, GRA method should be improved in terms of selection of distinguishing coefficients and analysis sorting angle. Combining GRA and KNN as a means of improving the algorithm performance is interesting. Thus, the improved GKNN algorithm is proposed in this study.

2.3. Optimization Problems Using GA for Demand Prediction. GA has been demonstrated to be efficient and robust at global searching [22] and thus to accelerate the training rate and improve the performance of the algorithm.

Some researchers focus on the optimization algorithm using GA. For example, Liu et al. [23] proposed an improved GA related to simulated annealing concept to optimize the weight of the pressure vessel under a sudden pressure constraint. Bu et al. [24] applied the improved GA to pollution source governance. Chen et al. [25] used the improved GA to study dynamic airspace sectorization problem. Qin et al. [26] put forward an improved GA optimizer, which was used to solve key technical problems in conceptual design stage of automobiles. Lyu et al. [27] and Wang et al. [28] proposed an improved GA and optimize the BPNN model. Su and Lin [29] designed an algorithm combining GA and BPNN to realize maximum power point tracking in photovoltaic power generation. Fang et al. [30] established a BPNN improved on the basis of GA to simulate external-phase concentration and extraction efficiency of 1-PHE under different operating conditions.

Therefore, it can be seen that combining GA and BPNN can obtain ideal optimizing result. However, the development of GA is insufficiently perfect and “premature” phenomenon or evolutionary stagnation problems will easily occur. Thus, improved GABP algorithm for demand prediction is proposed in this study.

3. Methodology

In the case of incomplete and uncertain disaster information, the collected information needs to be preprocessed in order to more accurately predict the demand for disaster relief materials. With the help of grey theory, uncertain information is regarded as grey information, and two formulas for calculating the kernel of interval grey number are proposed, and the specific value after whitening is used to replace the uncertain interval value, so as to achieve the purpose of preprocessing uncertain information. For incomplete information, an improved GKNN algorithm is proposed to fill in the missing value. This method combines the grey relational degree and the k-nearest neighbor algorithm, and adds weight function and logic test on this basis. Finally, the improved GABP model was used to predict the material demand of the preprocessed disaster information. The flow chart of the method is shown in Figure 1

3.1. Fuzzy Data Processing of the Improved Whitening Weight Function. Grey value distribution information must be first obtained prior to the construction of whitening weight function of interval grey numbers. The function can be used to solve the core of the interval grey number, and the core can be used to approximately represent this interval grey number. The whitening weight function expresses the inclination degree embodied by different values taken for this interval grey number within its interval range. Solving the core of the interval grey number using the whitening weight function is comparable to solving the value that can mostly embody its inclination degree. An interval grey number $\otimes \in [a_1, b_1]$ ($a_1, b_1 \in R$ and $a_1 < b_1$) is set, and general form of its whitening weight function can be expressed as follows:

$$f_{\otimes}(x) = \begin{cases} 0, & x \notin [a_1, b_1], \quad a_1, b_1 \in R, a_1 < b_1; \\ \frac{x - a_1}{c_1 - a_1}, & x \in [a_1, c_1], \quad c_1 \in R, a_1 < c_1; \\ 1, & x \in [c_1, d_1], \quad d_1 \in R, c_1 < d_1; \\ \frac{b_1 - x}{b_1 - d_1}, & x \in [d_1, b_1], \quad d_1 < b_1. \end{cases} \quad (1)$$

Then, the core \otimes^h of the interval grey number can be expressed as follows:

$$\begin{aligned} \otimes^h &= \frac{\int_{a_1}^{b_1} x f_{\otimes}(x) dx}{\int_{a_1}^{b_1} f_{\otimes}(x) dx} \\ &= \frac{b_1^2 + b_1 d_1 + d_1^2 - a_1^2 - a_1 c_1 - c_1^2}{3(b_1 - a_1 + d_1 - c_1)}. \end{aligned} \quad (2)$$

The whitening weight function above is typical, and the geometric figure comprising turning points of the interval grey number and endpoints is a trapezoid. Apart from the typical whitening weight function, triangular whitening weight function is also a commonly used function. This function can be obtained by setting $c_1 = d_1$ in the

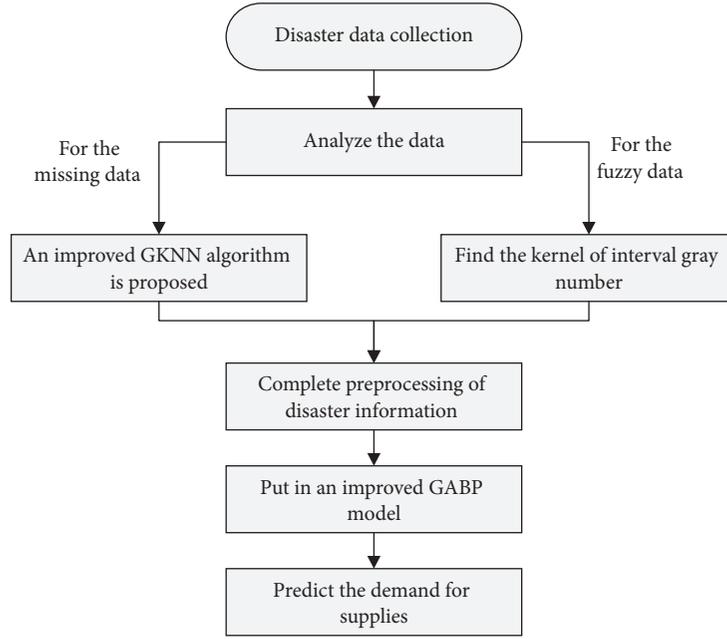


FIGURE 1: Disaster information preprocessing and material demand forecasting.

typical whitenization weight function. Thus, the triangular whitenization weight function can be expressed as follows:

$$f_{\otimes}(x) = \begin{cases} 0, & x \notin [a_1, b_1], \quad a_1, b_1 \in R, a_1 < b_1; \\ \frac{x - a_1}{c_1 - a_1}, & x \in [a_1, c_1], \quad c_1 \in R, a_1 < c_1; \\ 1, & x = c_1; \\ \frac{b_1 - x}{b_1 - c_1}, & x \in [c_1, b_1], \quad c_1 < b_1. \end{cases} \quad (3)$$

The core \otimes^h of the interval grey number can be expressed as follows:

$$\otimes^h = \frac{\int_{a_1}^{b_1} x f_{\otimes}(x) dx}{\int_{a_1}^{b_1} f_{\otimes}(x) dx} = \frac{b_1^2 + b_1 c_1 - a_1^2 - a_1 c_1}{3(b_1 - a_1)}. \quad (4)$$

The condition above is obtained on the basis of the already known whitenization weight function. However, a certain information quantity is needed under actual environment to determine the whitenization weight function. Determining turning points of the whitenization weight function in the absence of information is difficult. Thus, a concrete formula to express the core of the interval grey number will be discussed under uncertain turning points of the whitenization weight function.

The whitenization weight function of the K (th) interval grey number \otimes_k is set as $f_{\otimes_k}(x)$, upper and lower limits of each interval grey number are known, and $\otimes_k \in [a_k, b_k]$. As for the K (th) interval grey number, a maximum of two turning points Z_K^L and Z_K^R and a minimum of one turning point Z_K^M exist. $a_k \leq Z_K^L < Z_K^R \leq b_k$ when two turning points

exist, and $a_k \leq Z_K^M \leq b_k$ when only one turning point exists. According to formula (1), the whitenization weight function of the interval grey number can be expressed as follows:

$$f_{\otimes_k}(x) = \begin{cases} 0, & x \notin [a_k, b_k], \quad a_k, b_k \in R, a_k < b_k; \\ \frac{x - a_k}{Z_K^L - a_k}, & x \in [a_k, Z_K^L], \quad Z_K^L \in R, a_k < Z_K^L; \\ 1, & x \in [Z_K^L, Z_K^R], \quad Z_K^R \in R, Z_K^L < Z_K^R; \\ \frac{b_k - x}{b_k - Z_K^R}, & x \in [Z_K^R, b_k], \quad Z_K^R < b_k. \end{cases} \quad (5)$$

If two turning points $Z_K^L \in (a_k, (a_k + b_k)/2)$ and $Z_K^R \in ((a_k + b_k)/2, b_k)$ exist, then they can be expressed as follows:

$$Z_K^L = \left(\frac{a_k + b_k}{2} - a_k \right) \alpha + a_k = \frac{b_k - a_k}{2} \alpha + a_k, \quad (6)$$

$$\alpha \in (0, 1);$$

$$Z_K^R = \left(b_k - \frac{a_k + b_k}{2} \right) \beta + \frac{a_k + b_k}{2} \quad (7)$$

$$= \frac{b_k - a_k}{2} \beta + \frac{a_k + b_k}{2}, \quad \beta \in (0, 1).$$

α and β are coefficients selected by the researcher on the basis of the only existing grey value distribution information.

Through simultaneous formulas (2), (6), and (7), the core of the interval grey number \otimes_k can be expressed as follows:

$$\begin{aligned} \otimes_k^h &= \frac{b_k^2 + b_k Z_K^R - (Z_K^R)^2 - a_k^2 - a_k Z_K^L - (Z_K^L)^2}{3(b_k - a_k + Z_K^R - Z_K^L)} \\ &= \frac{A_k + B_k + C_k - D_k}{E_k}. \end{aligned} \tag{8}$$

In the formula,

$$A_k = \frac{1}{4}(7b_k^2 + 4a_k b_k - 11a_k^2); \tag{9}$$

$$B_k = \frac{1}{4}(b_k - a_k)^2(\beta^2 - \alpha^2); \tag{10}$$

$$C_k = \frac{1}{2}(2b_k^2 + a_k b_k - a_k^2)\beta; \tag{11}$$

$$D_k = \frac{3}{2}(a_k b_k - a_k^2)\alpha; \tag{12}$$

$$E_k = \frac{3}{2}(b_k - a_k)(3 + \beta - \alpha). \tag{13}$$

When only one turning point exists, the whitenization weight function of the K (th) interval grey number can be expressed as follows according to formula (3):

$$f_{\otimes_k}(x) = \begin{cases} 0, & x \notin [a_k, b_k], \quad a_k, b_k \in R, a_k < b_k; \\ \frac{x - a_k}{Z_K^M - a_k}, & x \in [a_k, Z_K^M], \quad a_k, Z_K^M \in R, a_k < Z_K^M; \\ 1, & x = Z_K^M; \\ \frac{b_k - x}{b_k - Z_K^M}, & x \in [Z_K^M, b_k], \quad Z_K^M, b_k \in R, Z_K^M < b_k. \end{cases} \tag{14}$$

where the turning point can be expressed as follows:

$$Z_K^M = (b_k - a_k)\gamma + a_k, \quad \gamma \in (0, 1). \tag{15}$$

Through the simultaneous formulas (4) and (15), the core of the interval grey number \otimes_k can be expressed as follows:

$$\begin{aligned} \otimes_k^h &= \frac{b_k^2 + b_k Z_K^M - a_k^2 - a_k Z_K^M}{3(b_k - a_k)} \\ &= \frac{2a_k + b_k}{3} + \frac{(b_k - a_k)}{3}\gamma, \quad \gamma \in (0, 1). \end{aligned} \tag{16}$$

In accordance with the grey value distribution information provided by the historical disaster data base, the corresponding whitenization weight function is constructed for the fuzzy interval grey number and the core of each interval grey number is solved. Each piece of fuzzy information is turned into directly usable data value for follow-up prediction work. Under difficult acquisition of distribution information, double-turning-point whitenization weight function inferred in this section is used to solve the core formula (8), whereas single-turning-point whitenization weight function is used to solve the core formula (16). Thus, rapid and convenient whitenization processing of interval grey numbers can be realized.

3.2. Missing Data Filling Algorithm Based on the Improved GKNN. The proposed missing data filling algorithm is improved on the basis of GKNN algorithm, which is a combination of GRA and KNN. KNN uses Euclidean distance to express similarities between samples, whereas GKNN uses grey relation degree as a replacement to Euclidean distance. The use of grey relation degree to express sample similarities can consider the relations of the entire data set fully but not those of two samples.

When using GKNN algorithm, we need to consider three aspects. In fact, these three aspects are the specific settings of the use conditions and parameters of the algorithm. First, how to set the proper initial K value of the algorithm? In this study, the value of K is randomly set within the range of $(2, \sqrt{m})$ (where \sqrt{m} represents the square root value of the number of samples), and the logic test condition test algorithm is set to fill in the rationality of the result. If the test condition is not met, the value of K is reselected until the result is reasonable. Second, what distance formula should be used to measure the distance of the algorithm? This study adopts the grey relational degree. The use of grey relation degree to express sample similarities can consider the relations of the entire data set fully but not those of two samples. Third, how to consider the weight of the algorithm's measurement distance? This problem is directly related to the accuracy of the filling results. The weight formula proposed in this study can well solve this problem and reduce the dependence of algorithm performance on the initial K value.

All algorithms improved on the basis of KNN algorithm have one feature, namely, accuracy of algorithm filling result due to K value. No definite and concrete computing formula for K value exists. Thus, cut-and-trial method is commonly used when faced with this problem. For example, Mohammed et al. [12] considered K values within a certain range for multiple repeated experiences and then used mean value of MSEs of the experiments and finally selected K value corresponding to a low error level as the parameter used in the experiment. This method could solve a K value with favorable algorithm filling performance. However, multiple cut-and-trial values needed multiple repeated experiments and calculations, which increased time cost and energy input unavoidably. Research conclusions of some scholars indicated that selection of K value is related to sample size of the study object. For example, Lall and Sharma [31] and Ennenbach et al. [32] believed that K value was not greater than square root of the sample size. Although this conclusion has not achieved a consensus among numerous scholars, it provides other researchers with a reference for K value selection.

The first improvement made for the proposed GKNN algorithm is to reduce its dependence on K value.

First, K value is randomly initialized, and then relation degree sorting of K candidate samples is conducted. The sample with great relation degree is similar to the target sample to be filled. Thus, a large weight is set for this sample. The sample with small relation degree has low similarity with the target sample. Therefore, a small weight is set for this sample. The weight of each candidate sample is the percentage of the grey relation degree between the corresponding candidate and target samples in the sum of grey relation degrees between

all candidate samples and the target sample. Given that the principle of the traditional KNN algorithm is to fill by taking the mean value of K candidate samples, the filling effect is restricted by K value selection. Assigning weights to candidate samples but not taking the mean value can strip off the dependence of KNN algorithm on K value to a great degree. Accordingly, the range of K value should be determined, and an ideal filling effect can be obtained by randomly taking a K value in this range. Following the conclusions of [31], the range of K value can be determined as $2 \leq K \leq \sqrt{n}$.

The second improvement made for the proposed GKNN algorithm is to add a test link to its filling result.

During missing information filling of the algorithm, a common mistake is assigning an impossible value as missing value of the target sample. To avoid this problem in the filling algorithm, whether the filling result is reasonable must be tested after filling of each piece of missing information is completed.

Test conditions should be set in accordance with the research contents. In this study, attributes within the target sample should be compared for the sake of completing filling. For example, attributes of one sample s are number A_s of households, population B_s , and number I_s of people losing residence. Test conditions can be different depending on dissimilar filling attributes, and test conditions under five circumstances are provided.

- (1) When filling attribute is A_s while other attributes are complete, the filling result should satisfy $0.5 B_s > A_s > 0.25 B_s$.
- (2) When filling attribute is A_s while B_s is vacant and other attributes are complete, the filling result should satisfy $5 I_s > A_s > 1.25 I_s$.
- (3) When filling attribute is B_s while other attributes are complete, the filling result should satisfy $4 A_s > B_s > 2 A_s$.
- (4) When filling attribute is B_s , A_s is vacant, $I_s > 1000$, and other attributes are complete, the filling result should satisfy $10 I_s > B_s > 6 I_s$.
- (5) When filling attribute is B_s , A_s is vacant, $I_s < 1000$, and other attributes are complete, the filling result should satisfy $20 I_s > B_s > 10 I_s$.

If the filling result does not conform to the set conditions, then return to determination phase of K value. K value will be reselected within the range, and valuation method is $K = K - 1$; the filling will be continued after K value is determined. If the filling result still does not meet the set conditions, then return to selection phase of K value for reselection; when $K = 1$, $K = \sqrt{n}$ is reset. Adding the test conditions can avoid logistic mistakes of the filling result, thereby making this result reasonable.

The logic flow of the improved GKNN algorithm is given as in Algorithm 1.

3.3. Improved GABP Algorithm. A self-adaptive GA was introduced here to realize weight value initialization and

threshold optimization of the BPNN model. The proposed improved GABP algorithm exerts not only fast global search advantage of GA but also extensive mapping advantage of BPNN.

The algorithm flow is given as follows.

Step 1. The samples are input to train the BPNN model. The number of iterations of network training is set as 2,000 times, the training accuracy is 0.001, the learning rate is 0.1, and L-M optimization algorithm (trainlm) is selected as the training function [33]. After the training, the initial network structure is obtained, and then the initial weight value and threshold are coded to randomly generate an initial population.

Step 2. Parameter initialization of GA is conducted [34]. The maximum evolution frequency of GA is set as 40, the population size is 15, the crossover rate is 0.5, the mutation probability is 0.01, the collected training samples are input, the error of the network prediction result is calculated, and the reciprocal of error sum of squares is taken as individual fitness. If individual fitness value is great, then the corresponding prediction error is small; otherwise, the prediction error is large. The fitness values of the individuals should be as great as possible.

Step 3. Population crossover and mutation probabilities will change with the fitness during the iteration process, and appropriate change can give the following features to the population evolution process. When individual fitness is small, crossover and mutation probabilities should be properly increased to intensify population diversity for preventing the algorithm from being caught in locally optimal solution. When individual fitness is large, crossover and mutation probabilities should be properly reduced to elevate optimizing and convergence rate for preventing the algorithm from being caught in random walk. Crossover P_c and mutation P_m probabilities of the self-adaptive GA are adjusted using the following formula:

$$P_c = \begin{cases} P_c + \frac{P(f_{avg} - f_1)}{f_{avg} - f_{min}}, & f_1 < f_{avg} \\ P_c - \frac{P(f_1 - f_{avg})}{f_{max} - f_{avg}}, & f_1 \geq f_{avg}; \end{cases} \quad (17)$$

$$P_m = \begin{cases} P_m + \frac{P(f_{avg} - f_2)}{f_{avg} - f_{min}}, & f_2 < f_{avg} \\ P_m - \frac{P(f_2 - f_{avg})}{f_{max} - f_{avg}}, & f_2 \geq f_{avg}; \end{cases}$$

where f_{max} is the maximum fitness value in the population; f_{avg} is the mean fitness value of each population; f_{min} is the minimum fitness value in the population; f_1 is the maximum fitness value between two crossover individuals; f_2 is the fitness value of the mutant individual; P is a value between (0, 1).

Step 4. Step 3 is repeated to evolve the weight and threshold of the neural network continuously until the error of network

```

Input: Disaster information data set  $X$  containing missing values
Output: Complete disaster information data set
(1) Initialization  $K$ 
(2) Set test conditions
(3) FOR each target sample in  $X$ 
(4)   FOR each candidate sample
(5)     Calculate grey relation degree  $G_i$ 
(6)     For  $G_i$  sorting,  $K$  candidate samples are selected
(7)     Fill missing information for the candidate samples by combining weight values
(8)     IF the result conforms to the test conditions
(9)       THEN fill the next target sample
(10)    ELSE change  $K$  value and re-fill.
(11)   END FOR
(12) END FOR

```

ALGORITHM 1

prediction result reaches the training goal or the number of iterations of the self-adaptive GA reaches the maximum value. Then, GA ends, and optimal initial weight and threshold are output.

Step 5. The obtained optimal initial weight and threshold are assigned to the BPNN model for simulated prediction.

4. Experiments and Results

We analyze the collected historical disaster information and preprocess the data according to different situations, so as to serve as the data basis for the subsequent prediction model. For the range data in disaster information, we regard it as the interval grey number in grey number theory and calculate the kernel of grey number. The obtained nuclear substitution has only a general range of grey number, and the whitening of grey number is completed in Section 4.1. We propose an improved GKNN algorithm to fill the gaps in disaster information in Section 4.2. Then we used the improved adaptive genetic algorithm to optimize the initial weight and threshold of BP neural network, compared the improved model with the traditional GABP model, and verified the stability and superiority of the model. Finally, the preprocessed disaster information was used as the training data of the improved GABP model, and the demand for emergency supplies after the disaster was predicted in Section 4.3.

4.1. Whitening Processing of Interval Grey Numbers. This study took the earthquake disaster loss data set in mainland China during the period of 2006~2010 as an example. Incomplete information in the data set was analyzed, and fuzzy information and missing information were preprocessed. Each row of the data set expresses one disaster case. For example, the first row in Table 1 represents the 5.0-magnitude earthquake in Mojiang, Yunnan Province on Jan. 12, 2006. Each column of the data set represents detailed event information, including number of households, population, earthquake magnitude, focal depth, epicentral intensity, disaster-stricken area, death toll, number of injuries, and number of people losing houses. The void item in each row

indicates that this data item has missing information. If one item in the sample row is interval number, then this item is interval fuzzy number, namely, interval grey number.

A total of 14 interval grey numbers were selected. Double-turning-point and single-turning-point whitening weight functions proposed in Section 3 were used to solve the core formulas (8) and (16), respectively. A total of 20 groups of experiment were carried out, and the experiment was divided into two major groups. The experimental background was that interval grey number was already known. Given that grey value distribution information was insufficient, number of turning points of the whitening weight function of each interval grey number was unknown. The discussion was conducted under different circumstances: two turning points of the whitening weight function and one turning point. Under the two circumstances, auxiliary variables influencing turning points randomly took values within the interval (0, 1). Numbers of turning points in different experiments were different because values were randomly selected. As a result, the solved core results of interval grey numbers were different. Therefore, core solving process of each interval grey number was repeated 20 times, and relative error between each result and the reference value was recorded. The reference value here was calculated through the definition of interval grey number “core” with lack of grey value distribution information following [3]. The principle was as follows: all interval grey numbers were put under mean value whitening processing due to the lack of grey value distribution information. In other words, the whitening weight function of each interval grey number was regarded as mean value distribution with the interval range, and the embodied grey value distribution probabilities were equal.

Core solving process of the first interval grey number is taken as an example (see the second line in the Table 1; that is, $k = 1$).

Under insufficient grey value information distribution, the following condition is obtained when the core of the interval grey number is solved using formula (8):

$$\otimes_1^h = \frac{A_1 + B_1 + C_1 - D_1}{E_1}, \quad (18)$$

TABLE 1: Partial data sets of earthquake disaster losses in mainland China during the period of 2006~2010.

Number of households	Population	Earthquake magnitude	Focal depth/km	Epicentral intensity	Area of the disaster-stricken area/km ²	Death toll	Number of injuries	Number of people losing houses
14,062	60,273	5	16	6	725	0	1	7,900
	[83,000, 86,500]	5	15	6	1,565	0	2	1,045
14,814	69,054	5	15	6	737	1	19	6,643
[7,830, 8,023]	[29,315, 31,442]	5.1	20	5	58	0	0	0
14,062	60,273	5	16	6	725	0	0	1,145
35,428	151,168	5.1	9	6	890	22	114	20,606
61,170	262,538	5.1	7	7	1,350	2	67	18,509
351	423	5.1	17	5	200	0	0	0
	[246,598, 248,931]	4.9	6	6	578	0	0	0
94,286	403,128	6.4	5	8	167	3	419	61,780
30,311	115,845	5.7	25	7	4,224.2	0	0	19,383
21,023	74,419	4.6	6	6	150	0	1	370
46,562	61,880	7.3	33	7	12,940	0	0	12,081
14,319	66,824	5	11	6	423	0	0	11,060
1,650	6,500	4.1	10	6	35	0	0	0
		5	33	6		0	0	1,061
8,861	29,669	5.1	22	6	480	0	0	7,393
	42,131	5.2	14	6	[9,948, 11,024]	0	0	0
8,1299	355,395	5.9	7	7	4,511	5	130	99,480
365,189	1,261,823	6.1	10	8	9,634	35	638	121,756

where

$$\begin{aligned}
 A_1 &= \frac{1}{4} (7b_1^2 + 4a_1b_1 - 11a_1^2) = \frac{1}{4} (7 \times 86500^2 + 4 \\
 &\quad \times 83000 \times 86500 - 11 \times 83000^2); \\
 B_1 &= \frac{1}{4} (b_1 - a_1)^2 (\beta_1^2 - \alpha_1^2) = \frac{1}{4} (86500 - 83000)^2 \\
 &\quad \cdot (\beta_1^2 - \alpha_1^2); \\
 C_1 &= \frac{1}{2} (2b_1^2 + a_1b_1 - a_1^2) \beta_1 = \frac{1}{2} (2 \times 86500^2 \\
 &\quad + 83000 \times 86500 - 83000^2) \beta_1; \\
 D_1 &= \frac{3}{2} (a_1b_1 - a_1^2) \alpha_1 = \frac{3}{2} (83000 \times 86500 \\
 &\quad - 83000^2) \alpha_1; \\
 E_1 &= \frac{3}{2} (b_1 - a_1) (3 + \beta_1 - \alpha_1) = \frac{3}{2} (86500 - 83000) \\
 &\quad \cdot (3 + \beta_1 - \alpha_1).
 \end{aligned} \tag{19}$$

In the formulas, α_1 and β_1 randomly take values within (0, 1), and value assigning results are $\alpha_1 = 0.8003$ and $\beta_1 = 0.1419$. Then, the core of the interval grey number based on double-turning-point whitenization weight function is calculated as $\otimes_1^h = 84731$.

When the core the interval grey number is solved using formula (16), the following condition is obtained:

$$\begin{aligned}
 \otimes_1^h &= \frac{2a_1 + b_1}{3} + \frac{(b_1 - a_1)}{3} \gamma_1 \\
 &= \frac{2 \times 83000 + 86500}{3} + \frac{(86500 - 83000)}{3} \gamma_1.
 \end{aligned} \tag{20}$$

In the formula, γ_1 randomly takes a value within (0, 1), and it is $\gamma_1 = 0.7922$ here. Consequently, the core of the interval grey number based on single-turning-point whitenization weight function is obtained as $\otimes_1^h = 85091$.

Following [3], the core of the interval grey number obtained through calculation is given as follows:

$$\otimes_1^h = \frac{1}{2} (b_1 + a_1) = \frac{1}{2} (86500 + 83000) = 87450. \tag{21}$$

Two calculation results obtained in this study were compared with the reference value calculated in [3]. Notably, the relative error of the result obtained through the calculation with the double-turning-point whitenization weight function is 0.02%, whereas that with the single-turning-point whitenization weight function is 0.40%. The results obtained through the two core solving formulas are approximate to the reference value, and the slight difference is actually embodiment of simulated random distribution of grey value information under lack of grey value information distribution. In practical application, information data should not only be accurate but also be accompanied with a certain

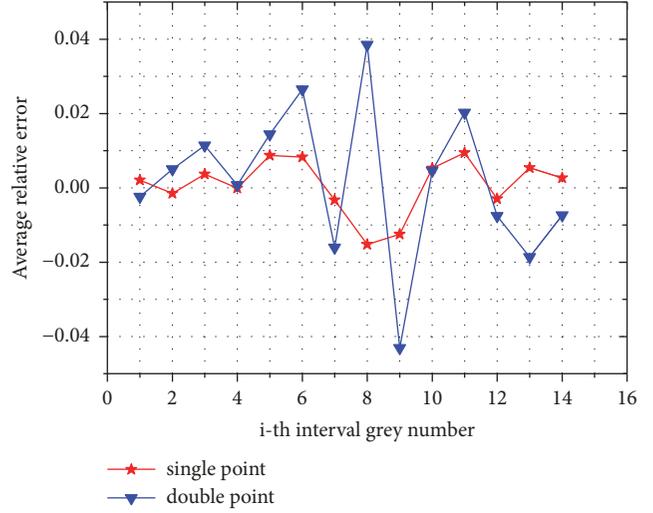


FIGURE 2: Comparison between results obtained through two core solving formulas and the reference value.

“noise,” which reflects information uncertainty. The two core solving formulas proposed in this study can embody this uncertainty of information distribution without affecting data accuracy. Error ranges of the results of 14 interval grey numbers calculated through the two core solving formulas are given in Figure 2.

Figure 2 shows the comparison of results obtained through two core solving formulas and the reference value in [3]. The x-coordinate expresses one interval grey number, and y-coordinate represents average relative error (ARE) of the calculated result based on the reference value. For example, mean results of four interval grey numbers through 20 experiments indicate that the results calculated through the two core solving formulas are approximate to the reference value. The overall mean error of the results obtained through core solving formula (6) of double-turning-point whitenization weight function is 1.19%, and the mean used time is 1.23 s. The two figures under single-turning-point whitenization weight function are 1.19% and 0.2 s. The two methods are the same in the calculation process. However, the core of interval fuzzy numbers calculated using formula (8) is more accurate than that calculated using formula (16).

In summary, if grey value distribution information is insufficient and the researcher must transform interval grey numbers into real numbers for the convenience of the follow-up use demand, then the proposed core solving formula can be used for whitenization processing of interval grey numbers. When the researcher deems that most interval grey numbers in the studied target samples are partial to typically measured whitenization weight function (that is, standard trapezoidal whitenization weight function), the computing formula for the core of the interval grey number based on the double-turning-point whitenization weight function is recommended. If most interval grey numbers in the studied target samples are partial to moderately measured whitenization weight function (that is, isosceles triangular whitenization weight function), then the computing formula for the core of the interval weight function based on single-turning-point

whitenization weight function is recommended. If many interval grey numbers need whitenization processing (that is, the data size is large), then the computing formula for the core of the interval weight function based on single-turning-point whitenization weight function can be selected in consideration of time cost. If accuracy of the whitenization processing result is pursued, then the computing formula for the core of the interval grey number based on the double-turning-point whitenization weight function can be selected. The advantage of the proposed core computing formula lies in that whitenization processing can reflect uncertainty of information distribution while guaranteeing result accuracy. The computing method is also simple and easy to use.

4.2. Completion Processing of Missing Information. After interval grey numbers are processed, missing values will be filled. The filling algorithm proposed in this study combines GKNN and conditional screening (GKNN-CS hereinafter) and is based on certain screening conditions. The filling effect of this algorithm will be compared with that of GKNN algorithm in [16]. The two algorithms are used to fill the data set, and ARE and filling effective rate of filling results are taken for comparison. ARE is calculated as follows:

$$ARE = \frac{1}{k} \sum_{i=1}^k \frac{|x_i - x_{iz}|}{x_i}, \quad (22)$$

where x_i is a filling value and x_{iz} is the true value at the location corresponding to x_i .

Then, the calculated example is used to present the steps of the proposed algorithm.

Step 1. The second row in Table 1 is selected as target samples to be filled. The algorithm checks the number of missing items in the target samples first and then finds only one missing item, which is recorded as Y_T . The missing attribute is “number of households,” and it is selected as the filling target.

Step 2. The matrix consisting of all samples including missing and complete data sets is made dimensionless through standard mapstd method with a mapping mean value of 0 and a mapping variance of 1.

Step 3. On the basis of the target samples, difference values between all residual and target samples are calculated. In this process, if surplus missing items exist in addition to the filling target in the target samples, then these missing items are temporarily replaced by zero in case that all difference sequences in the column where void values are located are void in the follow-up steps. After the calculation of difference sequences, the values are restored.

Step 4. Minimum and maximum elements in the obtained difference sequence matrix X are obtained and recorded as u_{\min} and u_{\max} , respectively.

Step 5. Relational coefficient matrix of each sample relative to the target samples is calculated, and the calculation formula is given as follows:

$$R(A, B) = \frac{(u_{\min} + p \times u_{\max})}{p \times u_{\max} + X(A, B)}, \quad (23)$$

where $R(A, B)$ expresses row A and column B in the relation coefficient matrix and $X(A, B)$ expresses row A and column B in the difference sequence. A small distinguishing coefficient p indicates great distinguishing capacity. This coefficient generally takes a value within $(0, 1)$, and it is taken as $p = 0.5$ here.

Step 6. The sum of rows of the relation coefficients matrix is solved, each row is divided by number n of columns in the overall data set matrix (that is, number of sample attribute types) to obtain relation degree matrix r , and the elements in the matrix r are sorted.

Step 7. K value, which can be randomly selected within the range $(2, \sqrt{m})$, is initialized. The initial K value is integrated into K_0 , where m represents number of rows of the overall data set matrix (that is, sample size).

Step 8. K numbers, which are great in the sorted matrix r , are selected and recorded as g_i , $i \in (1, 2, \dots, K)$. Attribute values of the candidate samples corresponding to the K numbers are extracted and recorded as Y_i , $i \in (1, 2, \dots, K)$. A value is assigned to Y_T to complete the filling process.

Step 9.

$$Y_T = Q_1 Y_1 + Q_2 Y_2 + \dots + Q_K Y_K, \quad (24)$$

where Q_i , $i \in (1, 2, \dots, K)$ is the weight of the attribute value of each candidate sample, and its concrete computing formula is given as follows:

$$Q_i = \frac{g_i}{g_1 + g_2 + \dots + g_i + \dots + g_K}. \quad (25)$$

Step 10. Whether the filling result of Y_T is reasonable is checked. Under this circumstance, Y_T represents the number of households, which cannot exceed four times the “number of people” attribute value in the sample where it is located and cannot be lower than the “population” attribute value. If the filling result does not meet this condition, then Step 11 is entered; otherwise, Step 12 is entered.

Step 11. $K = K + 1$, and Steps 8 to 10 are repeated. If $K > \sqrt{m}$, then $K = 2$, and Steps 8 to 10 are repeated. In this step, whether K is equal to K_0 is checked when K value changes. If yes, then Step 12 is directly entered.

Step 12. After the filling for this target is finished, then the next missing item is filled. If no missing item exists in this target sample, then the next target sample is determined. A new missing item is searched by repeating Step 1.

Step 13. After all missing items of the entire data set is completed, the algorithm is ended.

To verify the algorithm performance in the processing of samples of different missing degrees, the same data set was artificially deleted, the deletion quantity was controlled to simulate data sets of different missing degrees, and missing proportion ranged from 5% to 50%. Figure 3 shows that sample missing rate increases, and relative errors of both filling algorithms increase. When sample missing rate is within 5% ~ 10%, the two algorithms have favorable filling performance. When sample missing rate increases to approximately 25%, it considerably influences the accuracy of the filling result. Under the same missing rate, GKNN-CS is more approximate to the true value than the traditional GKNN. This advantage is obvious when filling samples with a large missing rate. This result indirectly indicates the reduced dependence of GKNN-CS on K value after the weight is added and its better accuracy than the traditional GKNN algorithm.

However, algorithm filling result does not always conform to the reality. When the same size is limited, K -neighbor optimization of GKNN algorithm cannot find sufficient and proper approximate samples to fill missing values. Effective rate of algorithm filling expresses the percentage of reasonable quantity of filling results in the total filling quantity. Figure 4 shows that, when the sample missing rate is not large, GKNN-CS can effectively fill all missing values. However, when missing rate further increases, filling effective rates of both algorithms gradually decline. In the samples with the same missing degree, filling effective rate of GKNN is not as good as that of GKNN-CS. This result fully certifies that, after the test conditions are added, GKNN-CS can acquire filling results in great agreement with reality.

4.3. Improved GABP Prediction Model. A complete data set was obtained after all missing items of the data set were completed. The complete data set was used to train basic GABP and the improved GABP based on MATLAB platform. The trained networks were used to predict the number of missing people. Each network repeated the prediction 50 times, and then result errors of GABP and improved GABP were compared. Y-coordinate represents relative error, which refers to the percentage of the true difference value between the predicted values obtained through GABP or the improved GABP and true value. The closer the relative error is to zero, the better. X-coordinate represents the forecasting times through the first to the 50th experiments. Figure 5 shows that error fluctuation range of GABP predicted values is large, whereas many errors of predicted values obtained through the improved GABP are distributed to nearly 0%. Through the calculation, average prediction error in 50 times of experiment carried out by GABP is approximately 18.17%, whereas that carried out by improved GABP is approximately 6.83%. Evidently, the initial weight and threshold of the improved GABP algorithm are better for the BPNN model to seek for global optimum during the iteration process.

The primary and processed data were then input into the improved GABP model. Primary data include fuzzy interval numbers and missing information, and these primary samples will affect prediction accuracy of the neural network. The processed data have become a complete data set, and

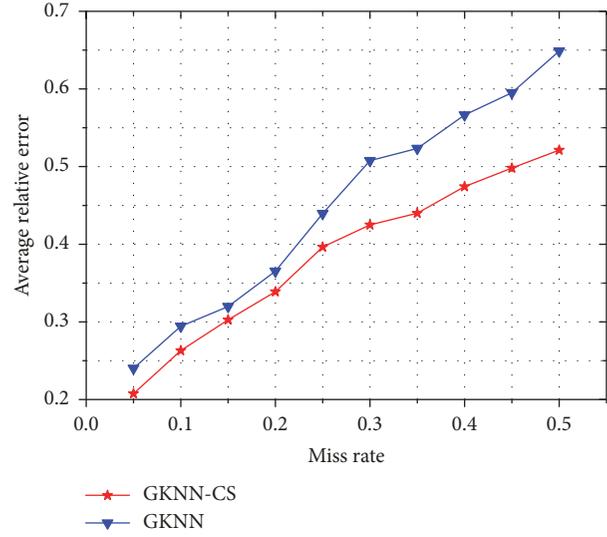


FIGURE 3: Comparison between algorithms in filling errors.

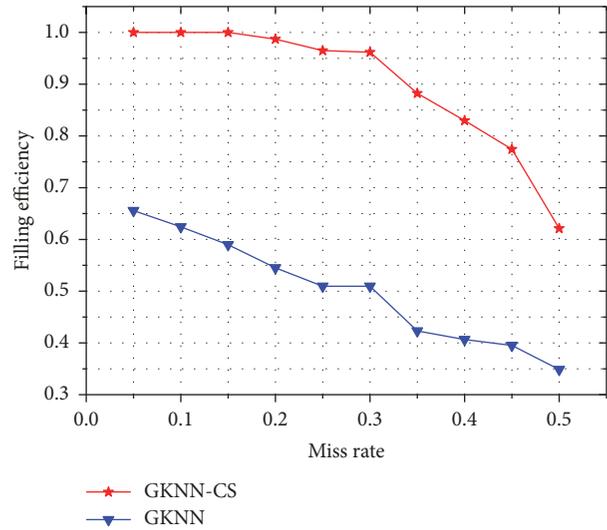


FIGURE 4: Comparison between algorithms in filling efficiency.

they are abbreviated as preprocessed samples (presamples). Complete and sufficient data contribute highly to prediction of the neural network. Prediction results in different training processes of the neural network are inconsistent. Thus, average value of the same target in repeated predictions must be obtained to determine the concrete predicted value. Accordingly, each group of experiment was repeated 20 times, each result was recorded, and relative error between predicted and true values was calculated.

The experiment was divided into five groups. Primary samples and presamples in each group were used to train the improved GABP model, and results were predicted. The first-group prediction target was number of reliefs needed in the Dangchang 4.3-magnitude earthquake in Gansu Province on March 27, 2006. The second-group prediction target was rice demand in the Yanjin 5.1-magnitude earthquake accompanied with 4.7-magnitude earthquake in Yunnan

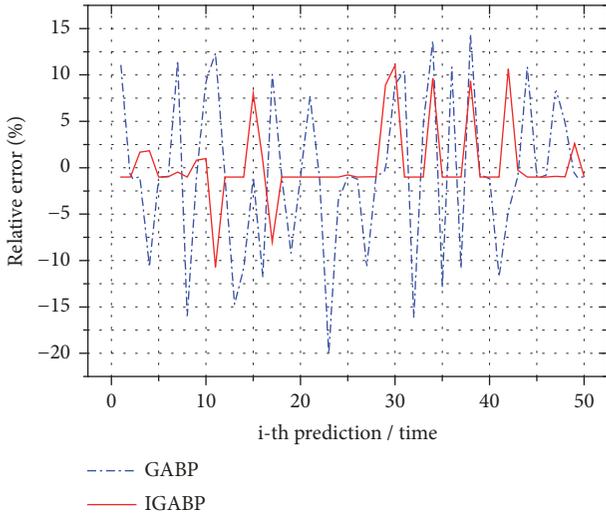


FIGURE 5: Performance comparison of GABP model before and after the improvement.

Province from August 25 to 29 in 2006. Fuzzy and missing data existed in primary samples in both groups. The third-group prediction target was freshwater demand in the Zushan (Hubei)-Baihe (Shanxi) 4.1-magnitude earthquake on March 24, 2008, and missing phenomenon existed in the primary data in this group. The fourth-group prediction target was freshwater demand in the Dangxiong 6.6-magnitude earthquake in the Tibet Autonomous Region on October 6, 2008, and fuzzy information existed in primary data in this group. Prediction target in the fifth group was cotton quilt demand in the Ulugqat 5.1-magnitude earthquake in the Xinjiang Uygur Autonomous Region on June 10, 2010.

Prediction results of the first and second groups are shown in Figures 6 and 7, respectively. In the first group, the prediction results of the primary samples show that the ARE and variance are 12.40% and 41.13, respectively. The prediction results of presamples indicate that the ARE and variance in this group are 5.07% and 4.85, respectively. In the second group, the prediction results of primary samples reveal that the ARE and variance are 14.55% and 39.16, respectively. The prediction results of presamples imply that the ARE and variance in this group are 4.51% and 7.02, respectively. Notably, the prediction result obtained by training the neural network using complete data set is more stable and has higher accuracy than the result obtained using primary data.

Prediction results of the third group are shown in Figure 8. In the third group, the prediction results of primary samples show that the ARE and variance are 15.47% and 64.31, respectively. The prediction results of presamples indicate that the ARE and variance in this group are 8.53% and 12.33, respectively. Similarly, the prediction result obtained by training the neural network using complete data set is more stable and accurate than the result obtained using primary data. However, the improvement degree of prediction accuracy by presamples is restricted compared with those in the first and second groups. In conclusion, the prediction stability is the poorest when fuzzy and missing data simultaneously exist

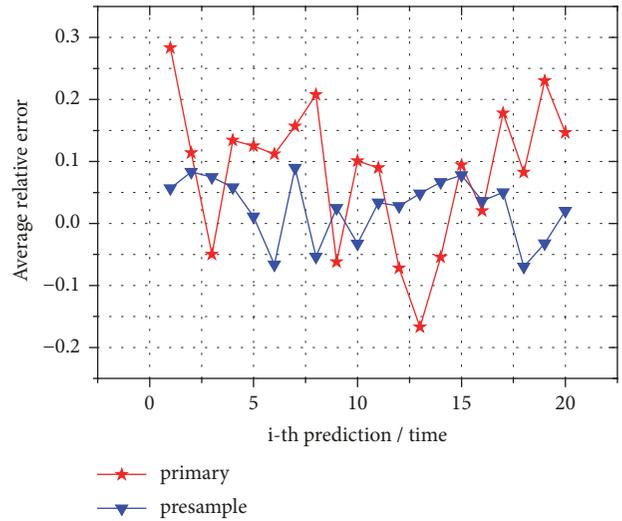


FIGURE 6: First group of experiment.

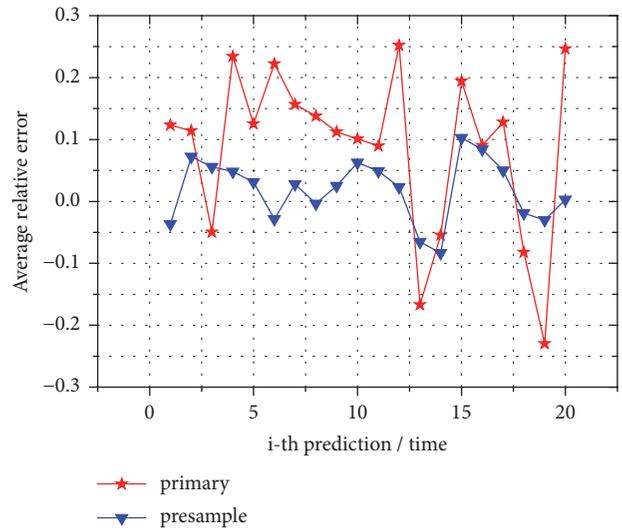


FIGURE 7: Second group of experiment.

in primary samples. When only fuzzy information exists in primary samples, the improvement in the model prediction performance is restricted.

Prediction results of the fourth group are shown in Figure 9. In the fourth group, the prediction results of primary samples imply that the ARE and variance are 12.21% and 61.09, respectively. The prediction results of presamples reveal that the ARE and variance in this group are 9.73% and 46.37, respectively. Although the prediction result obtained by training the neural network using complete data set has more advantages over that obtained using primary data, the advantages are inconsiderable. The third and fourth groups are similar because they have only one uncertain phenomenon, but their uncertain phenomena have different expression forms. Comparison of prediction results between the two groups indicates that fuzzy information contains the trace of finding the true information, and prediction accuracy

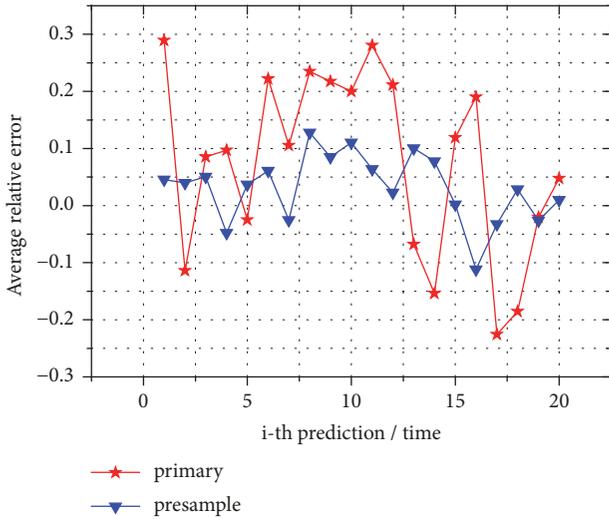


FIGURE 8: Third group of experiment.

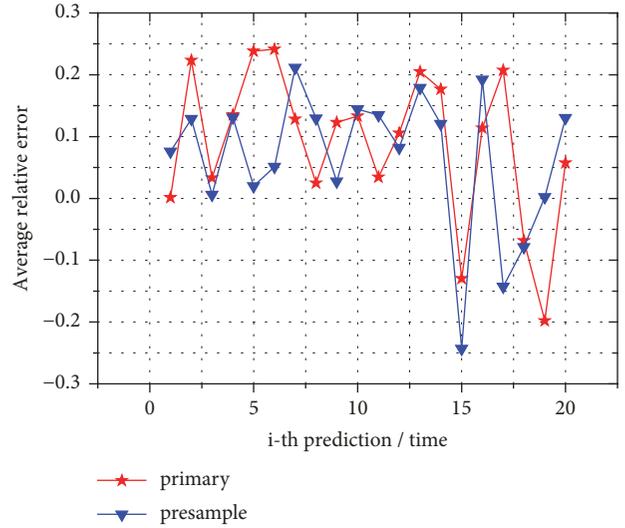


FIGURE 10: Fifth group of experiment.

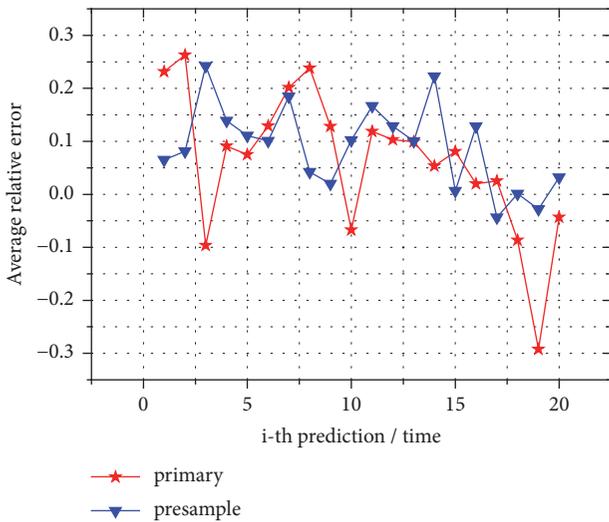


FIGURE 9: Fourth group of experiment.

can be greatly improved through whitening processing of fuzzy information. Missing phenomenon cannot provide any information, and filling algorithm can only use the complete data set to complete the missing items but results in biased estimation. Thus, the complete data set obtained through filling inconsiderably contributes in improving the prediction effect.

Prediction results of the fifth group are shown in Figure 10. In the fifth group, the prediction results of primary samples imply that the ARE and variance are 12.19% and 43.66, respectively. The prediction results of presamples indicate that the ARE and variance in this group are 11.92% and 43.98, respectively. No missing or fuzzy information exists in primary samples at the time. Thus, they are not different from presamples. The ARE and variance values imply that the difference between the primary samples and presamples is small in prediction result. From the comparison

of the experimental results in the fifth, first, and second groups, the following conclusion can be drawn. The complete data set obtained by preprocessing the primary data through the proposed data preprocessing method for fuzzy or missing information is better for accurate and effective prediction of the casualties in the earthquake disaster.

Under the reference effect of the fifth group, the experimental groups which can truly embody data completeness and effectiveness of the proposed data preprocessing method are the first, second, third, and fourth groups. The experiments show that the proposed preprocessing method not only can avoid missing information of potential values in the primary data but also can help improve stability and accuracy of the prediction result. The calculation indicates that the prediction accuracy is improved by 20.31% ~ 69%, and its stability is improved by 24.1% ~ 88.21%.

4.4. Experimental Results. Through five experiments, the pre-processed disaster information is compared with the original disaster information, and their contribution to the material demand forecast is analyzed. The experimental results show that if there are missing values and fuzzy values in the disaster information, using the method in this study to preprocess the disaster and then predict the material demand, the prediction accuracy can be significantly improved, at least 20% in the experiment. The proposed method effectively solved the prediction problem under incomplete information in an earthquake disaster, thereby providing information support and guarantee for demand prediction of emergency supplies.

It should be noted that the disaster information preprocessing method proposed in this study needs to meet certain conditions before it can be used. For example, the kernel formula of interval grey number requires that the whitening weight function of interval grey number is based on trapezoidal whitening weight function or triangular whitening weight function.

5. Conclusions and Future Work

The core solving concept of whitenization weight function of grey theory was applied to processing of fuzzy information in this study. Two simple and easy core solving formulas of fuzzy interval grey numbers were put forward for whitenization processing of fuzzy data under uncertain grey value information distribution. These formulas reserved uncertainty of information distribution objectively existing in the primary data while guaranteeing accuracy. It avoids the situation that it is difficult for researchers to use whitening weight function to calculate the interval grey number when they cannot determine the distribution of grey value information.

Grey relation degree and KNN algorithm were combined, logic test conditions were added, and weights were introduced in the filling link and used for completion of the missing information, which reduced the dependence degree of GKNN algorithm on K value. As a result, the improved GKNN algorithm had better performance, the improved GKNN algorithm can reduce the maximum average relative error by more than 20%, and its filling result conformed considerably to the actual logic.

GA was improved and combined with BPNN to predict the casualties in the earthquake, and a prediction model with accurate prediction effect was obtained. Experimental results showed that the modified GABP model reduced the average relative error by 11.34%.

Despite the above-mentioned innovative points, this study exhibited some deficiencies. (1) Only typically measured and triangular whitenization function types were considered in the derivation of core solving formulas based on the whitenization weight function. Other types of whitenization weight function were not involved. (2) During the optimization process of BPNN using self-adaptive GA, the simulation experiment was quite time consuming, and a large time cost should be input into the repeated simulation experiments. Therefore, future research can focus on shortening the time used in the operation process without sacrificing prediction accuracy.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (grants nos. U1633119, 71271113, and U1233101). The authors gratefully acknowledge China Earthquake Data Center for providing access to actual earthquake data.

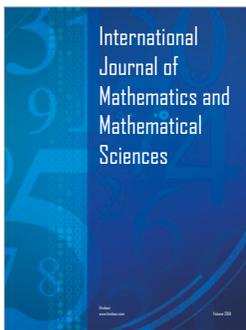
Supplementary Materials

Table 1: data sets of earthquake disaster losses in mainland China during the period of 2006–2010. Table 2: performance comparison of GABP model before and after the improvement. Table 3: the range of relative error based on the double-turning-point whitenization weight function. Table 4: the range of relative error based on the single-turning-point whitenization weight function. Table 5: comparison between algorithms in filling errors. Table 6: comparison between algorithms in filling efficiency. Table 7: first group of experiment. Table 8: second group of experiment. Table 9: third group of experiment. Table 10: fourth group of experiment. Table 11: fifth group of experiment. (*Supplementary Materials*)

References

- [1] M. Zhang, H. Yu, J. Yu, and Y. Zhang, "Dispatching plan based on route optimization model considering random wind for aviation emergency rescue," *Mathematical Problems in Engineering*, vol. 2016, Article ID 1395701, 11 pages, 2016.
- [2] J. I. L. Moreno, J. Latron, and A. Lehmann, "Effects of sample and grid size on the accuracy and stability of regression-based snow interpolation methods," *Hydrological Processes*, vol. 24, no. 14, pp. 1914–1928, 2010.
- [3] S. Liu, Z. Fang, and N. Xie, "Algorithm rules of interval grey numbers based on the 'Kernel' and the degree of greyness of grey numbers," *Systems Engineering & Electronics*, vol. 32, no. 2, pp. 313–316, 2010.
- [4] S.-F. Liu, Z.-G. Fang, and Y.-J. Yang, "Two stages decision model with grey synthetic measure and a betterment of triangular whitenization weight function," *Control & Decision*, vol. 29, no. 7, pp. 1232–1238, 2014.
- [5] P. Liu, "The multi-attribute group decision making method based on the interval grey linguistic variables weighted aggregation operator," *IOS Press*, vol. 24, no. 2, pp. 405–414, 2013.
- [6] N. Xie and S. Liu, "Interval grey number sequence prediction by using non-homogenous exponential discrete grey forecasting model," *Journal of Systems Engineering and Electronics*, vol. 26, no. 1, pp. 96–102, 2015.
- [7] B. Zeng, C. Li, G. Chen, and W. Zhang, "Verhulst model of interval grey number based on information decomposing and model combination," *Journal of Applied Mathematics*, vol. 4, pp. 1–8, 2013.
- [8] Y.-G. Dang, Y. Feng, S. Ding, and L. Wei, "Grey clustering model for interval grey numbers based on kernel and degree of greyness," *Control & Decision*, vol. 32, no. 10, pp. 1844–1848, 2017.
- [9] J. Ye and Y. Dang, "A novel grey fixed weight cluster model based on interval grey numbers," *Grey Systems: Theory and Application*, vol. 7, no. 2, pp. 156–167, 2017.
- [10] Z.-X. Wang, "Correlation analysis of sequences with interval grey numbers based on the kernel and greyness degree," *Kybernetes*, vol. 42, no. 2, pp. 309–317, 2013.
- [11] J. Sun and H. Li, "Financial distress early warning based on group decision making," *Computers & Operations Research*, vol. 36, no. 3, pp. 885–906, 2009.
- [12] M. A. Mohammed, M. K. A. Ghani, R. I. Hamed et al., "Solving vehicle routing problem by using improved K-nearest neighbor algorithm for best solution," *Journal of Computational Science*, vol. 21, pp. 232–240, 2017.

- [13] N. Suguna and K. Thanushkodi, "An improved k-nearest neighbor classification using genetic algorithm," *International Journal of Computer Science Issues*, vol. 7, no. 4, pp. 18–21, 2010.
- [14] F. Meng, C. Cai, and H. Yan, "A bicluster-based bayesian principal component analysis method for microarray missing value estimation," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 3, pp. 863–871, 2014.
- [15] M. Žukovič and D. T. Hristopulos, "A directional gradient-curvature method for gap filling of gridded environmental spatial data with potentially anisotropic correlations," *Atmospheric Environment*, vol. 77, no. 7, pp. 901–909, 2013.
- [16] S. Zhang, Z. Jin, and X. Zhu, "Missing data imputation by utilizing information within incomplete instances," *The Journal of Systems and Software*, vol. 84, no. 3, pp. 452–459, 2011.
- [17] B. Zhu, L. Yuan, and S. Ye, "Examining the multi-timescales of European carbon market with grey relational analysis and empirical mode decomposition," *Physica A: Statistical Mechanics and Its Applications*, vol. 517, pp. 392–399, 2019.
- [18] A. J. Moon-Grady, P. Moore, and A. Azakie, "Response surface method using grey relational analysis for decision making in weapon system selection," *Journal of Systems Engineering and Electronics*, vol. 25, no. 02, pp. 265–272, 2014.
- [19] L. Yan and H. Shan, "The fuzzy grey relational analysis of the factors influencing farm produce logistics," *Asian Agricultural Research*, vol. 6, no. 05, pp. 1–4, 2014.
- [20] B. Qiu, "Research on method of simulation model validation based on improved grey relational analysis," in *Proceedings of 2012 International Conference on Solid State Devices and Materials Science (SSDMS 2012 V25)*, Information Engineering Research Institute, 2012.
- [21] X. M. Li, K. W. Hipel, and Y. G. Dang, "An improved grey relational analysis approach for panel data clustering," *Expert Systems with Applications*, vol. 42, no. 23, pp. 9105–9116, 2015.
- [22] Y. Zhang and B. Pan, "Modeling batch and column phosphate removal by hydrated ferric oxide-based nanocomposite using response surface methodology and artificial neural network," *Chemical Engineering Journal*, vol. 249, pp. 111–120, 2014.
- [23] P. Liu, P. Xu, S. Han, and J. Zheng, "Optimal design of pressure vessel using an improved genetic algorithm," *Journal of Zhejiang University*, vol. 9, no. 9, pp. 1264–1269, 2008.
- [24] Q. Bu, Z. J. Wang, and X. Tong, "An improved genetic algorithm for searching for pollution sources," *Water Science and Engineering*, vol. 6, no. 04, pp. 392–401, 2013.
- [25] Y. Chen, H. Bi, D. Zhang, and Z. Song, "Dynamic airspace sectorization via improved genetic algorithm," *Journal of Modern Transportation*, vol. 21, no. 2, pp. 117–124, 2013.
- [26] H. Qin, Y. Guo, Z. Liu, Y. Liu, and H. Zhong, "Shape optimization of automotive body frame using an improved genetic algorithm optimizer," *Advances in Engineering Software*, vol. 121, pp. 235–249, 2018.
- [27] J. Lyu and J. Zhang, "BP neural network prediction model for suicide attempt among Chinese rural residents," *Journal of Affective Disorders*, vol. 246, pp. 465–473, 2019.
- [28] Y. Wang, C. Lu, and C. Zuo, "Coal mine safety production forewarning based on improved BP neural network," *International Journal of Mining Science and Technology*, vol. 25, no. 2, pp. 319–324, 2015.
- [29] Y. Su and X. Lin, "Research of PV power generation MPPT based on GABP neural network," *Journal of Physics*, vol. 1016, no. 1, pp. 1–6, 2018.
- [30] Z. Fang, X. Liu, M. Zhang et al., "A neural network approach to simulating the dynamic extraction process of l-phenylalanine from sodium chloride aqueous solutions by emulsion liquid membrane," *Chemical Engineering Research and Design*, vol. 105, pp. 188–199, 2016.
- [31] U. Lall and A. Sharma, "A nearest neighbor bootstrap for resampling hydrologic time series," *Water Resources Research*, vol. 32, no. 3, pp. 679–693, 1996.
- [32] M. W. Ennenbach, P. C. Larrauri, and U. Lall, "County-scale rainwater harvesting feasibility in the United States: climate, collection area, density, and reuse considerations," *Journal of the American Water Resources Association*, vol. 54, no. 1, pp. 255–274, 2017.
- [33] J. A. Prieto-Ruiz, R. Alis, S. García-Benlloch et al., "Expression of the human TIMM23 and TIMM23B genes is regulated by the GABP transcription factor," *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, vol. 1861, no. 2, pp. 80–94, 2018.
- [34] G. Sun, L. Qin, Z. Hou et al., "Feasibility analysis for acquiring visibility based on lidar signal using genetic algorithm-optimized back propagation algorithm," *Chinese Physics B*, vol. 28, no. 2, pp. 283–287, 2019.



Hindawi

Submit your manuscripts at
www.hindawi.com

