

Research Article

Variability-Weighted Interpolation Algorithm Based on Fixed-Frame Sampling of Data

Bo Cheng,^{1,2} Mingnan Zhang ,³ and Xiaomei Hu³

¹School of Computer Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China

²The Sixth Institute, 601 Branch of China Aeronautical Science and Technology Corporation, Hohhot 010076, China

³The Key Laboratory of Intelligent Manufacturing and Robotics, School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200072, China

Correspondence should be addressed to Mingnan Zhang; mingnanzhang@shu.edu.cn

Received 5 April 2022; Revised 17 May 2022; Accepted 23 May 2022; Published 11 June 2022

Academic Editor: Gopal Chaudhary

Copyright © 2022 Bo Cheng 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.

A spatial data interpolation algorithm is a method to transform scattered and sparse measurement data into regular and continuous applicable data. This algorithm can make the display results smooth and continuous without changing the data characteristics, which is closer to the real acquisition object. In response to the low interpolation accuracy of the inverse distance weighting (IDW) algorithm and the poor interpolation efficiency of the ordinary kriging (OK) algorithm, this study proposes a variational-weighted (VW) interpolation algorithm based on data fixed-frame sampling. By introducing the variational function of OK into IDW and taking different weight coefficients from OK, a new computational model is constructed to improve the interpolation accuracy and adapt to different types of data characteristics. In the whole interpolation process, the data are sampled in a fixed frame and a new sampling point search method is proposed to improve the computational efficiency. The study not only compares the accuracy and efficiency of the algorithm using two types of data but also tests the stability of the new algorithm for different data volumes. It is shown that VW based on a fixed-grid sampling of the data is more accurate than the inverse distance interpolation algorithm for both data types. The VW interpolation algorithm based on data lattice sampling improves the computational efficiency for featureless data and exhibits better performance in terms of accuracy and computational efficiency for data with features, compared to the common kriging interpolation algorithm. In addition, the VW interpolation algorithm based on a fixed-grid sampling of data has a more stable performance for different amounts of data.

1. Introduction

Visualization technology is a technique for transforming collected data into images and graphics that are intuitively visible to humans, through which one can better analyze, monitor, and understand the properties and characteristics of the visible object [1]. As an important branch of visualization technology, data visualization is a technology that uses computer graphics, computer vision, image processing techniques, and user interfaces to transform data into displayable graphics or images with interactive, multidimensional, and visual characteristics [2]. However, discrete, sparse data sets are common due to the difficulty of the acquisition process and the sheer volume of data collected. To achieve a realistic and complete visualization of such data,

it is necessary to supplement the missing data with interpolation algorithms and to adjust the data transition for the visualization part, so that the displayed picture is more in line with the real situation while meeting people's requirements for a clear and concise visualization [3]. Interpolation algorithms are therefore commonly used in the fields of meteorological visualization, terrain visualization, the spatial distribution of soil properties, and image processing [4–8]. In recent years, many researchers have conducted extensive research on interpolation algorithms, and [9] proposed an adaptive inverse distance weighting (AIDW) method, which uses the distribution density of sampling points to adjust the distance decay parameters, thus giving the interpolation method better flexibility to adapt to heterogeneous distance decay parameters in space.

[10] proposed an improved kriging method based on a message passing interface model to improve the computational efficiency of the spatial interpolation process. [11] proposed a common kriging interpolation algorithm (wind-OK) combining the influence of wind direction on spatial and temporal distribution of ground PM concentration in a given area and proved through experiments that the performance of the new algorithm was improved compared with the common kriging algorithm.

Spatial data interpolation algorithms can convert sparse, irregular, and scattered data into continuous, regular, and rich data sets by interpolation methods from discrete data and sparse data [12]. In this article, based on IDW and OK, a variational-weighted interpolation algorithm based on data fixation sampling is proposed, which combines the respective advantages of IDW and OK, so that the interpolation results of the original data have a high interpolation efficiency while ensuring a certain interpolation accuracy.

2. Research on Sampling Point Search Method and Interpolation Algorithm

Sampling point search is a method of determining the known data points to be involved in the interpolation of the points to be interpolated through a certain search method. Since it is very computationally intensive and inefficient to use all known data points directly for interpolation, sampling point search is an important step before the interpolation operation, and the goodness of the sampling point search method also affects the overall interpolation results to a certain extent. The traditional sampling point search methods are distance search and direction search, both of which determine the sampling points to be involved in the interpolation operation by measuring the distance between the known data points and the points to be interpolated before the interpolation operation is performed [13, 14].

The essence of the distance search method is to select the N nearest known data points to the point to be interpolated as the sampling points. When the known original data point is closer to the point to be interpolated, the greater the influence on the interpolation result, and vice versa, the smaller the influence on the interpolation result [15]. As shown in Figure 1, for the point Q to be interpolated, the search radius R is first determined and the known data points within the search radius are identified. Next, the distance between each data point and the point to be interpolated is calculated, and the four closest points are determined. As shown in Figure 1, $Q1$, $Q2$, $Q3$, and $Q4$, these four points are used as the sampling points involved in the interpolation operation.

The disadvantage of this method is that it is based purely on distance and does not place restrictions on the location of the sampling points. When the distribution of known data points is extremely irregular, all the sampling points may come from the same direction. As can be seen from Figure 1, $Q1$, $Q2$, $Q3$, and $Q4$ are all on the lower side of point Q . This results in no data points on the upper side being involved in the interpolation operation. Although the distance between the upper side data points and point Q is not as close as the lower side data points, they have a certain influence on the

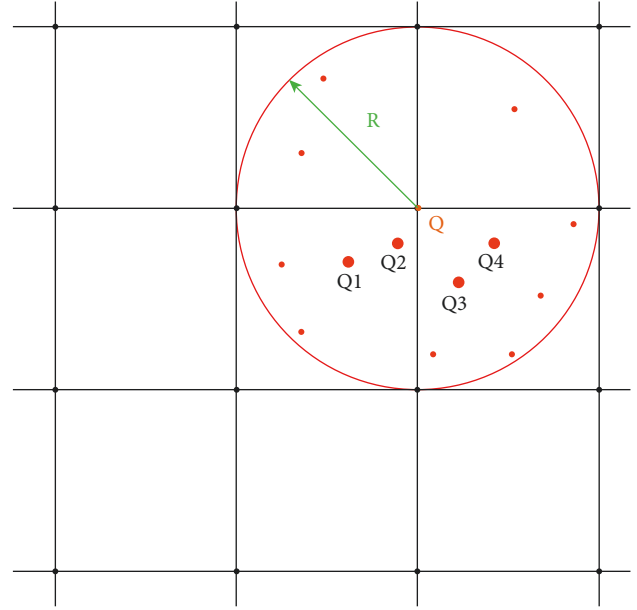


FIGURE 1: Schematic diagram of the distance search method.

interpolation results, and not using them at all will result in a large error between the interpolation results and the actual situation.

The directional search method is a method of improving interpolation accuracy by using sampling points in different directions on the basis of distance search, which makes up for the possible disadvantages of the distance search method of taking points in one direction. The essence of the directional search method is to select the nearest known data points in different directions to be interpolated as sampling points for interpolation. Directional search methods mainly include the four-way search method, the six-way search method, and the eight-way search method [16]. The eight-way search method is used as an example to briefly describe the search process. As shown in Figure 2, the entire original data are first searched to determine the known data points in eight directions within the interpolation radius R . Next the distance between each data point and point Q within each direction is calculated separately. Finally, the closest point within each direction is determined as the sampling point to be involved in the interpolation operation. The seven points $Q1$, $Q2$, $Q3$, $Q4$, $Q5$, $Q7$, and $Q8$ in the figure are used as the sampling points for the interpolation operation of point Q .

Although this method compensates for the potential disadvantages of distance search, for each point to be interpolated, the entire original data have to be searched to determine the data points within the search radius R , and the nearest point to the point to be interpolated in each direction has to be determined, which is a large search effort and also reduces the efficiency of the entire interpolation process to a certain extent. The maximum number of eight points to be interpolated also limits the interpolation accuracy to a certain extent. The fixed nature of the search radius makes it impossible to adapt to the characteristics of different parts of the data, and the search radius may be too large or too small in some parts.

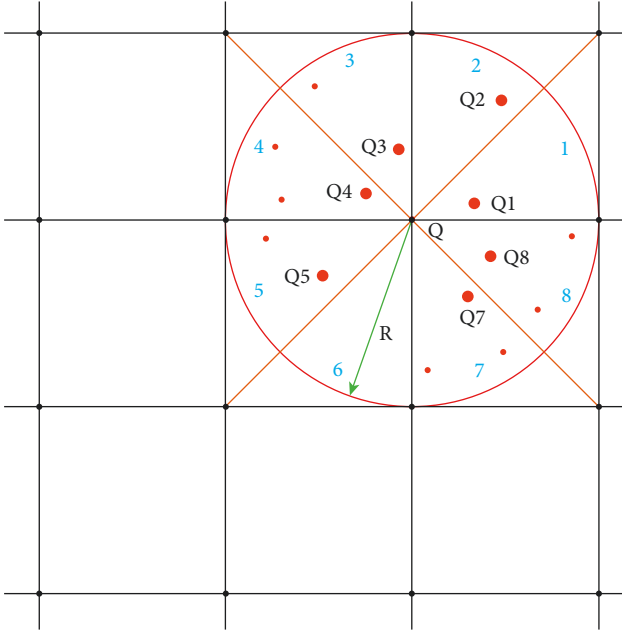


FIGURE 2: Schematic diagram of the eight-way search method.

Spatial interpolation algorithms based on the basic assumption of the “First Law of Geography” mainly rely on the collected data to perform interpolation calculation to fill in the unknown data in the original data. The closer the data points are to each other, the greater the relative influence, and conversely, the further the distance, the smaller the relative influence [17]. This study focuses on the inverse distance weighting method, the spline function method (SF), the kriging interpolation method, and the trend surface method (TS).

The inverse distance weighting method is an interpolation method based on the principle of similarity. Its essence is to fit a smooth mathematical plane equation with the known point data and then calculate the data values of the unknown points according to the equation, that is, to perform multiple regression analysis to obtain the smooth mathematical plane equation based on the relationship between the attribute data and the geographical coordinates of the sampled points. If the coordinates of two points P_i and P_j in space are (x_i, y_i) and (x_j, y_j) , then the value of the distance d_{ij} between the two points P_iP_j is as follows [18]:

$$d_{ij} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}. \quad (1)$$

Let P_j be the interpolation point, P_i be the sample point, n is the number of known sample points around P_j , and the attribute value of P_i be Z_i , the attribute value Z_j of the interpolation point P_j is found. The value Z_j calculated using the inverse distance-weighted interpolation method is as follows:

$$Z_j = \sum_{i=1}^N (w_{ij} Z_i), \quad (2)$$

where w_{ij} represents the weight, which is inversely proportional to the distance d_{ij} from a known point in the vicinity of P_j . To normalize the weights w_{ij} :

$$w_{ij} = \frac{(1/d_{ij}^k)}{\sum_{i=1}^N (1/d_{ij}^k)}, \quad (3)$$

where k represents the value of the distance attenuation parameter. When the value of k is large, the closer the sample point to be interpolated, the greater the influence of the sample point. When the value of k is small, the more distant the sampling point is from the point to be interpolated, the greater the influence on it. Substituting into equation (2) gives the attribute value Z_j for point P_j as

$$Z_j = \frac{\sum_{i=1}^N (Z_i/d_{ij}^k)}{\sum_{i=1}^N (1/d_{ij}^k)}. \quad (4)$$

The spline function method is a method that uses functions for interpolation calculations, often using segmented polynomials with both first-order and second-order derivatives that are continuous to approximate known data points, thus generating smooth interpolation curves. This method is simple to operate, not too computationally intensive, fast, suitable for interpolating contours based on dense points, and suitable for surfaces with progressive changes. However, this method is more difficult to estimate the interpolation error and has poor interpolation results when the points are not dense [19].

The kriging interpolation method is an interpolation method based on the theory of variational functions for the unbiased optimal estimation of regionalized variables in a finite region [20]. This algorithm pays more attention to the relative position relationship between each sampling point while considering the relative position relationship between the sampling points and the interpolation points. The estimates of kriging interpolation are obtained by weighting a linear combination of existing data points, which means that kriging interpolation is linear. Kriging interpolation brings the mean residual or error close to zero, meaning that kriging interpolation is unbiased. Kriging interpolation, therefore, gives results that are closer to the actual situation. However, this algorithm is computationally inefficient and consumes computer resources. The calculation is also more complicated with large amounts of data.

The kriging method of interpolation assumes that the region under study is D , point $x_i \in D (i = 1, \dots, N)$. Let the regionalized variable under study on the region be $Z(x)$ and the attribute value at point x_i be $Z(x_i)$. Based on the kriging method of interpolation, the attribute valuation $Z^*(x_0)$ at interpolation point $x_0 (x_0 \in D)$ is the weighted sum of the attribute values $Z(x_i) (i = 1, \dots, n)$ of the n known points involved in interpolation around the interpolation point x_0 , as shown in

$$Z^*(x_0) = \sum_{i=1}^n \lambda_i Z(x_i), \quad (5)$$

where $\lambda_i (i = 1, \dots, n)$ is the weighting factor to be determined. According to the unbiased condition of kriging interpolation, $\lambda_i (i = 1, \dots, n)$ must satisfy the relation:

$$\sum_{i=1}^n \lambda_i = 1. \quad (6)$$

Then the system of equations for solving the kriging method for the coefficient λ_i to be determined is obtained as

$$\begin{cases} \sum_{i=1}^n \lambda_i = 1 \\ \sum_{i=1}^n \lambda_i \text{Cov}(x_i, x_j) + \mu = \text{Cov}(x_0, x_j) \quad (j = 1, \dots, n), \end{cases} \quad (7)$$

where $\text{Cov}(x_i, x_j)$ is the covariance function of $Z(x_i)$ and $Z(x_j)$ and μ is the Lagrange multiplier. From the above system of equations, λ_i is solved and substituted to obtain an estimate $Z(x_i)$. There is a certain correlation between $Z(x_i)$, which is related not only to the distance but also to its relative direction.

The trend surface method is similar to the inverse distance weighting method in that it also uses known data points to fit a smoothed mathematical plane equation and then calculates the attribute values of the unknown points from the equation [21]. It uses the relationship between data attributes and geographic coordinates of the original data to perform a multiple regression analysis to obtain a smoothed trend surface function. The trend surface method is an approximate interpolation method because, in practice, it is generally difficult for the trend surface to pass exactly through the original data points, except in special cases such as when there are few data points and they happen to be passed by a surface. The trend surface method is a statistical method of interpolation that presupposes a series of interrelated spatial data where the trend and period of the interpolated point is a function of the other variables associated with it. It is also one of the more commonly used overall interpolation methods, allowing a smooth mathematical plane to be used to describe some geographic property that varies continuously in space.

Traditional spatial interpolation methods are widely used in the corresponding fields with their respective characteristics, so it is necessary to choose a reasonable and applicable interpolation algorithm according to the characteristics of the data spatial distribution [22, 23]. Through the analysis of the purpose, accuracy, and operation speed of interpolation, the above interpolation aspects are compared from the perspectives of interpolation accuracy, operation efficiency, extrapolation capability, and data distribution requirements, as shown in Table 1:

As can be seen from Table 1, the inverse distance weighting method has high operational efficiency, but the interpolation accuracy is lower compared to the kriging interpolation algorithm, while the kriging interpolation method has high operational accuracy and a wide range of applications, but low efficiency. Compared with the spline function method and the trend surface method kriging, the advantages of these two interpolation algorithms are more obvious.

TABLE 1: Comparison between several commonly used interpolation methods.

	IDW	SF	OK	TS
Interpolation accuracy	Higher	Low	High	Low
Operational efficiency	High	High	Low	Higher
Extrapolating ability	Poor	Strong	Stronger	Strong
Data requirements	Uniformity	Dense	All available	Limited

3. The Principle of Data Fixed-Frame Sampling and Variability-Weighted Interpolation

3.1. Data Framing Sampling. As the original data have sparse and irregular characteristics, the sampling point search and interpolation methods are very important aspects for the accuracy of the interpolated data and the subsequent application effect of the data. Among the traditional sampling point search methods, the distance search method has no restriction on the orientation of the sampling points, which reduces the interpolation accuracy; the direction search method has lower search efficiency. Among the traditional interpolation algorithms, the kriging method has high interpolation accuracy and good effect, but the algorithm is complex and time-consuming to compute; the inverse distance interpolation algorithm has high computational efficiency, but poor computational accuracy [24]. Combining the characteristics of the original data and the respective advantages and disadvantages of the traditional sampling point search method and the traditional interpolation algorithm, this study proposes a variability-weighted interpolation algorithm based on the fixed-grid sampling of the data, which is divided into two parts: a fixed-grid sampling of the original data and variability-weighted interpolation of the points to be interpolated.

Compared with the distance search method, the directional search method adds a limitation on the sampling orientation of the sampling points, which improves the interpolation accuracy to a certain extent and presents a better interpolation effect. However, for the directional search method, a circle is drawn with the search radius to search for sampling points, which limits the search range to a certain extent and may result in a certain degree of missing sampling points. When the data sparsity is relatively high, the octagonal search method can only determine a relatively small number of sampling points, and if the search range is expanded at this time, a new data point search is required, which is a more complex and tedious process. In this study, we propose data grid sampling to ensure that a certain number of sampling points are involved in the interpolation operation by searching the data on the grid around the interpolation point and simplifying the search process by searching the grid data in a single pass.

In all traditional methods of searching for sampling points, there is a problem of inefficient searching. The main reason: when each point to be interpolated for interpolation

calculation, the entire original data have to be searched and then the sampling points involved in the operation are determined. This process is undoubtedly more tedious, so this study adopts a single search of the fixed-grid data, where the original data are first gridded, and then a single full search of the original data is used to determine the location of the grid where the data are located. This allows the data points within the grid to be extracted directly when applying the data gridding sampling method to search for sampling points, improving the efficiency of the search.

Assume that the original data study area D , which will first be gridded, is processed to satisfy the following process: each grid after division is a square with h as the side length, and attention should be paid to the division process to ensure that the data points are evenly distributed within the grid as far as possible; the grid coordinates after division are the grid coordinates for interpolation, and all grid points should have feature values present after interpolation is completed. Suppose any point Q is located on grid coordinate (X, Y) . In order to facilitate the positioning and extraction of the original data points, each grid will be represented by its lower left coordinate, then the grids around the point $Q(X, Y)$ are (X, Y) , $(X - 1, Y)$, $(X - 1, Y - 1)$, and $(X, Y - 1)$, as shown in Figure 3.

Assume that there are N original data points (x_i, y_i, z_i) , $(i = 1, 2, \dots, N)$ within the study area D , and record all data points within grid (X, Y) using $V_{X,Y}$ and $V_{X,Y,j}$. Prior to a single search of the fixed-grid data, it is necessary to:

- (1) Set initial values for $V_{X,Y}$ and $V_{X,Y,j}$:

$$\begin{cases} V_{X,Y} = 0, & (X = 1, 2, \dots, L_X; Y = 1, 2, \dots, L_Y), \\ V_{X,Y,j} = 0, & (j = 1, 2, \dots, N), \end{cases} \quad (8)$$

where L_X represents the maximum value of the grid's horizontal coordinates and L_Y represents the maximum value of the grid's vertical coordinates.

- (2) For any data point (x_i, y_i, z_i) , $(i = 1, 2, \dots, N)$, locate the coordinates of the grid in which it is located as

$$\begin{cases} X = \text{floor}\left[\frac{(x_i - x_1)}{h}\right] + 1, \\ Y = \text{floor}\left[\frac{(y_i - y_1)}{h}\right] + 1, \end{cases} \quad (9)$$

where (x_1, y_1) are the coordinates of the leftmost lower grid point in the study area D , h is the side length of the small square grid, and floor denotes the leftward rounding function.

The steps for a single search of the fixed-grid data are shown in Figure 4.

When data points within any grid need to be extracted during the interpolation process, proceed as follows:

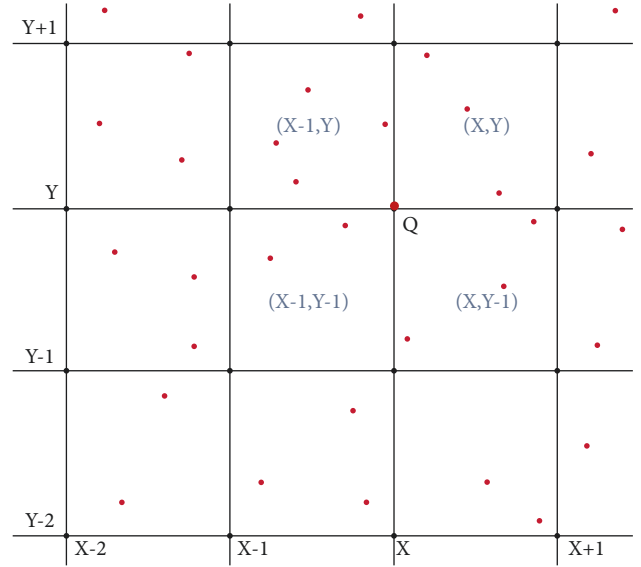


FIGURE 3: Gridding of raw data and representation of grid coordinates.

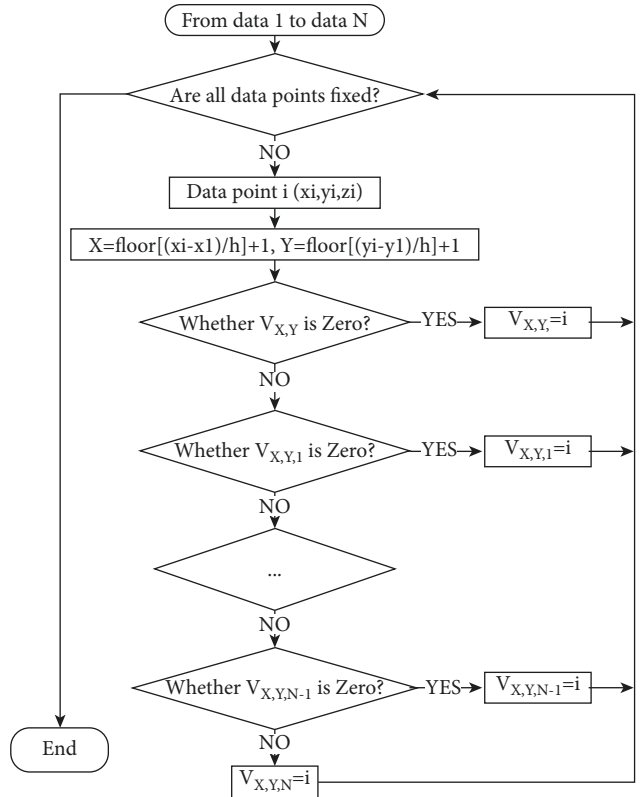


FIGURE 4: Single search for fixed-grid data.

- (1) Determine if $V_{X,Y}$ is zero, when $V_{X,Y} = 0$, there is no known data point in grid (X, Y) and no subsequent steps are necessary; when $V_{X,Y} = a (\neq 0)$, data point (x_a, y_a, z_a) is the first known data point in grid (X, Y) , continue with step (2).

- (2) When $V_{X,Y,1} = b (\neq 0)$, data point (x_b, y_b, z_b) is the second known data point within grid (X, Y) .
- (3) When $V_{X,Y,2} = c (\neq 0)$, data point (x_c, y_c, z_c) is the third known data point within grid (X, Y) .
- ...
- (4) Until $V_{X,Y,j} = 0 (j = 1, 2, \dots, N)$, then the above data points are all known data points located within grid (X, Y) .

After all the original data have been located by searching and stored, the interpolation of the missing points can be performed. Before the points to be interpolated can be subjected to the interpolation operation, the sampling points to be involved in the interpolation operation need to be determined. Data framing sampling ensures the effectiveness of the sampling point search method by looking at the amount of data to be sampled, the interpolation efficiency, and the sampling direction of the sampled points.

When searching for sampling points, the search first covers the first layer of the grid around the point to be interpolated. If there are less than 8 data points in the first layer of the grid, the second layer of the grid is added and so on. This way of searching ensures that the number of participating interpolation points is not less than 8, without wasting other data points within the grid, and improves the interpolation accuracy. In terms of efficiency, it is not necessary to take only 8 data points in this section to cut down on time as the search is already significantly more efficient by searching for location storage data. The data fixing sampling method also focuses on the sampling direction of the sampling points while ensuring the amount of data to be sampled and the efficiency of the interpolation. As the grid is divided with appropriate edge lengths so that the original data are distributed as evenly as possible across the grid, and no data points are deleted from the grid during the interpolation process, this ensures to a certain extent that the sampling direction of the sampling points involved in the interpolation is guaranteed. In addition, when the data points in the first search grid are less than 8, the second layer is allowed to participate in the interpolation operation, which also solves the problem of missing data in some directions in the first layer. At the edges of the grid, the same rules are applied until the number of sampling points exceeds 8. The search method is shown in Figure 5.

Point $Q(X, Y)$ to be interpolated follows the following steps when conducting a search for sampling point data:

- (1) Firstly, the data points in grids 1, 2, 3, and 4 are extracted from the graph and the number of data points is judged, if the amount of data is less than 8, then step (2) is performed. If the amount of data is greater than or equal to 8, all the data in grids 1, 2, 3, and 4 are used as interpolation operations, and the attribute value of the point $Q(X, Y)$ to be interpolated is calculated.

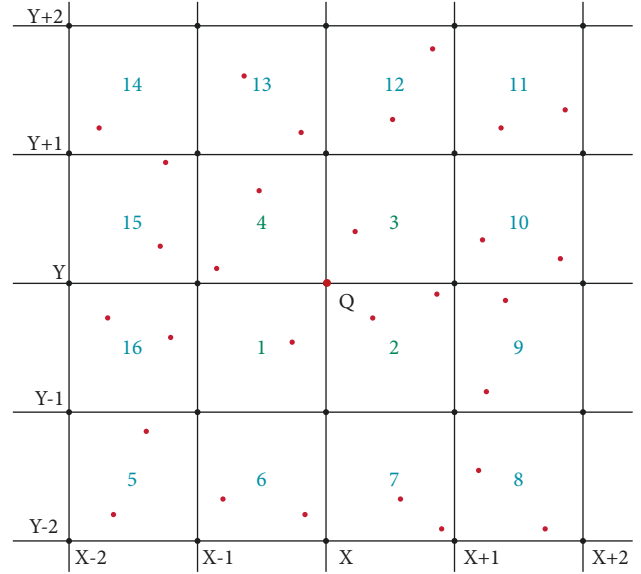


FIGURE 5: Data grid sampling search method.

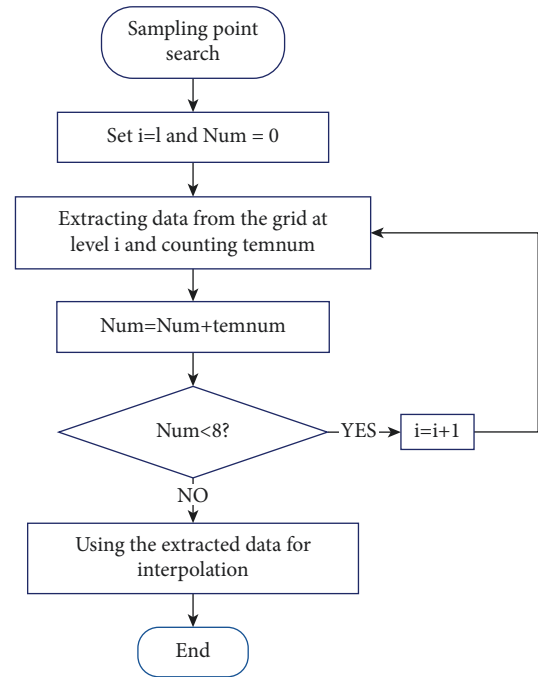


FIGURE 6: Flow chart of data framing sampling.

- (2) The data points in grids 5, 6, ..., 15, 16 are extracted and the total number of data points extracted is calculated to be less than 8. If the amount of data is less than 8, the data points in the third level of the grid are extracted, and so on. If the amount of data is greater than or equal to 8, all the data in grids 1, 2, 3, ..., 15, 16 are used for interpolation, and the attribute value of the point $Q(X, Y)$ to be interpolated is calculated.

The flow chart for data framing sampling is shown in Figure 6.

3.2. Variability-Weighted Interpolation Algorithm. The variability-weighted interpolation algorithm combines the high efficiency of the inverse distance weighting algorithm with the high accuracy of the ordinary kriging interpolation algorithm. The inverse distance weighting formula and the variability weighting formula are then weighted separately to determine the interpolation interval for the high accuracy interpolation results by comparing the interpolation results for different weights.

If the original data study area D , the N original data points $x_i (x_i \in D, i = 1, 2, \dots, N)$ corresponding to the attribute value $Z(x_i)$ and the point to be interpolated is $x_0 (x_0 \in D)$. The weighting factor of the inverse distance weighting algorithm can be expressed as follows:

$$w_i = \frac{(1/d_{i0}^k)}{\sum_{i=1}^N (1/d_{i0}^k)}, \quad (10)$$

where w_i denotes the weight coefficient, $\sum_{i=1}^N w_i = 1$, d_{i0} indicates the distance between the point to be interpolated and the original data point number i , and k denotes the distance decay parameter value.

The above formula shows that the weight coefficient is only related to the distance d_{i0} between the two points, and the variation function in ordinary kriging is also a quantity related only to the distance between the two points, so the distance d_{i0} in formula (11) can be replaced by the various function of ordinary kriging, and the n data points involved in interpolation are selected in the experimental process using data fixation sampling, and the weight coefficient obtained can be expressed as follows:

$$\lambda_i = \frac{(1/\gamma(x_i, x_0)^k)}{\sum_{i=1}^n (1/\gamma(x_i, x_0)^k)}, \quad (11)$$

where the variance function $\gamma(x_i, x_0) = \gamma(x_i - x_0) = 0.5E[Z(x_i) - Z(x_0)]^2$.

$$Z(x_0) = a \sum_{i=1}^N w_i Z(x_i) + b \sum_{i=1}^n \lambda_i Z(x_i), \quad (12)$$

where a and b are weighting factors and $a + b = 1, 0 < a < 1, 0 < b < 1$.

4. Experimental Process

Through the analysis of the data fixation sampling method and the variability-weighted interpolation algorithm, the general steps of the experiment were obtained as follows:

- (1) Experimental data information is obtained.
- (2) For interpolation gridding, the original data are located and stored by a single search of the fixed-grid data.
- (3) The variability function is calculated based on the original data.

The raw data are mostly obtained by means of acquisition and measurement and is therefore discrete data. Discrete data are divided into data with a scattered distribution of coordinates and data with a

regular distribution of coordinates. For data with a discrete coordinate distribution, the variation function can be calculated by relaxing the distance and angle tolerances, so that data with a small degree of distance and angle deviation can be grouped into data with the same direction and step length for calculation. As the data used in this study are data with regular coordinate distribution and the coordinates are arranged at intervals of 1 along the x and y axes, the basic step length is directly set to 1 when calculating the variation function.

The specific steps for calculating the experimental variance function are as follows:

- (1) Initialize the data variables.
Define the variables: L is the basic step, which is taken to be 1 due to the nature of the experimental data. a is the number of directions of the variation function.
Define the array: RawData is used to store the original data. $D[a]$ (an integer array) holds the direction of each various function. $DataX(a)[N]$ and $DataY(a)[N]$ hold the data information for two points in each direction of variation. $V(a)[N]$ holds the value of the various function in each direction at the position of step $mL (0 < m \leq N)$. $PN(a)[N]$ holds the number of pairs of points corresponding to each direction.
- (2) Read N raw data and store it in array RawData. Determine the number of directions of the variational function a .
- (3) Select a pair of raw data in RawData and calculate the distance between the two points and the angle between them and the x axis. Use this to determine the direction of variation as $i (1 \leq i \leq N)$ and deposit it in the corresponding $DataX(i)[N]$ and $DataY(i)[N]$.
- (4) Repeat step (3) until all the data in the array RawData have been traversed.
- (5) Select two data points in each direction corresponding to arrays $DataX(a)[N]$ and $DataY(a)[N]$, and calculate the distance d between the two points. Since the basic step of the experiment L is 1, $dk = d/L$. The values of the attributes of the two points are subtracted and then squared, and then accumulated and counted in $V(i)[dk]$. $PN(i)[dk] + 1$.
- (6) Repeat step (5) until the arrays $DataX(a)[N]$ and $DataY(a)[N]$ are traversed.
- (7) Iterate through arrays $V(a)[N]$ and $PN(a)[N]$, dividing each element of array $V(a)[N]$ by two times the value of the corresponding element in array $PN(a)[N]$. The resulting result is returned to be assigned to $V(a)[N]$.

The value of the variance function of the experiment is stored in the array $V(a)[N]$. The flow chart for calculating the values of the experimental variation functions is shown in Figure 7.

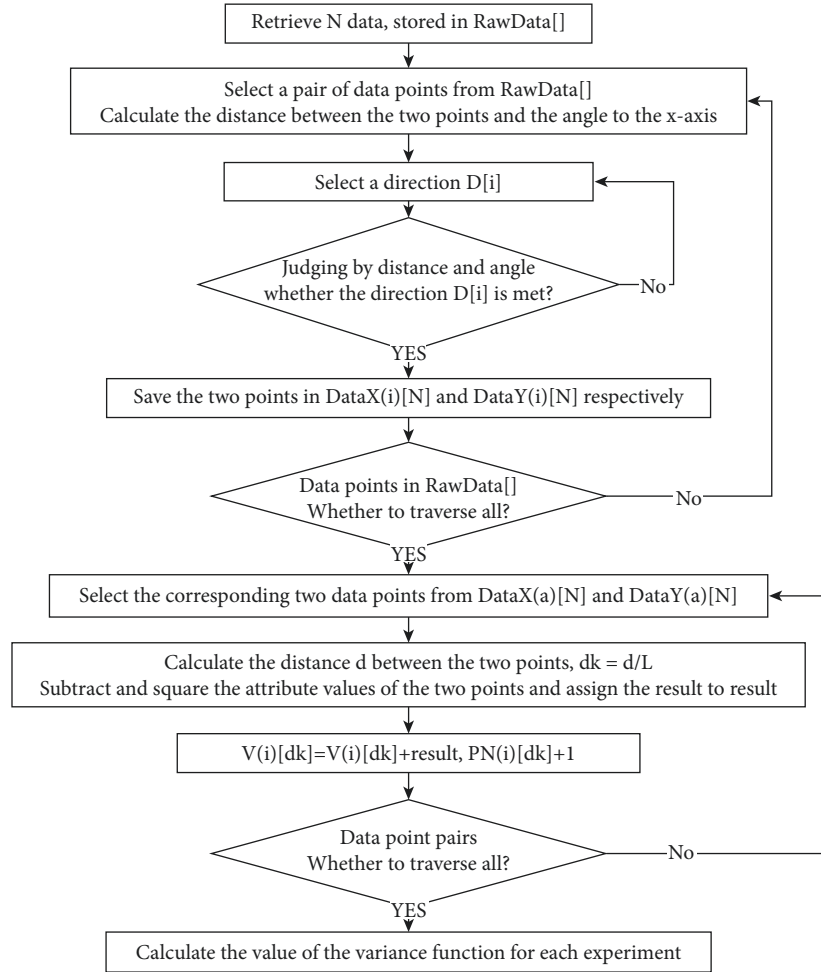


FIGURE 7: Flow chart for calculating experimental variogram values.

Experimental variation function values were obtained separately for smooth data without features and abrupt data with features by experimental calculations (Figures 8 and 9).

- (4) Select a suitable theoretical model of the various function according to the value of the various function for parameter fitting and determine the variation function formula.

The variation function is determined from the experimentally obtained values of the various function and fitted to the corresponding theoretical variation function model. The theoretical model of the various function is divided into two main types: models with abutment values and models without abutment values. Models with abutments include spherical models, exponential models, and Gaussian models. The models without abutments mainly include the power function model and the logarithmic function model. In practice, for most regionalized variables, models with abutments are chosen. Based on the values and distribution of the experimental variation functions obtained, the most suitable theoretical variation function model can be

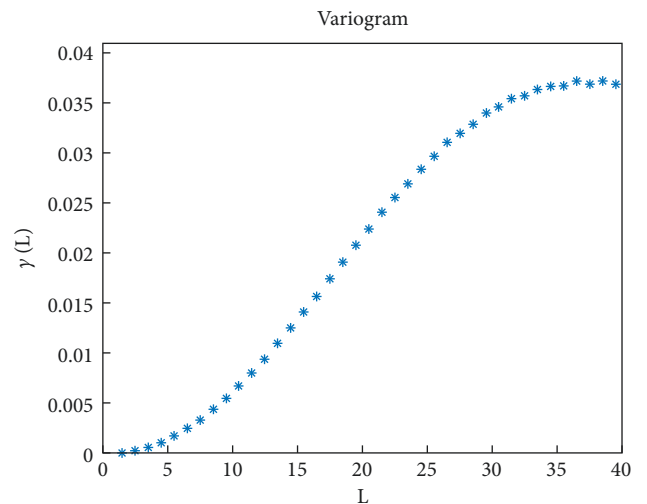


FIGURE 8: Variation function values for uncharacteristic smooth data.

determined as a Gaussian model. The theoretical formulation of the Gaussian model can be expressed as follows:

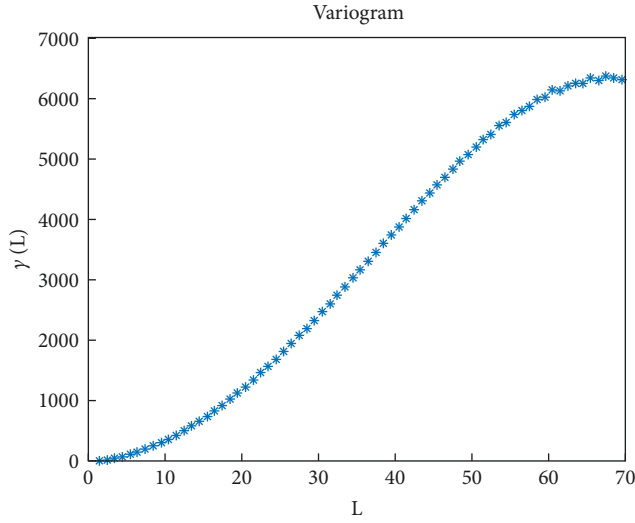


FIGURE 9: Variation function values for characteristic mutation data.

$$\gamma(L) = \begin{cases} 0, & L = 0, \\ C\left(1 - e^{-(L^2/a^2)}\right), & L > 0, \end{cases} \quad (13)$$

where $\gamma(L)$ denotes the variation function. C denotes the arch height, representing the part of the parameter that varies structurally. L denotes the basic step. a denotes the variation range, representing the process by which the distance varies with respect to the spatially relevant differences. The parameters of the Gaussian model of the variance function for smooth data without features can be obtained by combining the values of the distribution of the variance function obtained: the value of C is 0.0372 and the value of a is 20.8604. Then, the expression of the variance function $\gamma(L)$ for smooth data without features is

$$\gamma(L) = \begin{cases} 0, & L = 0, \\ 0.0372\left(1 - e^{-(L^2/20.8604^2)}\right), & L > 0. \end{cases} \quad (14)$$

The fitted Gaussian model variance functions are shown in Figures 10 and 11.

- (5) Select point x_0 ($x_0 \in D$) to be interpolated, and select the sampling point x_i ($x_i \in D$) to be involved in the interpolation operation by using the data fixation sampling method.
- (6) Calculate the distance d_{i0} between the points to be interpolated and the sampling points.
- (7) Calculate the weight coefficient w_i based on the distance d_{i0} , and the weight coefficient λ_i based on the variation function formula.
- (8) Calculate the attribute value $Z(x_0)$ of the point x_0 to be interpolated.
- (9) Complete the calculation of all the points to be interpolated by cycling through steps 6 to 9.

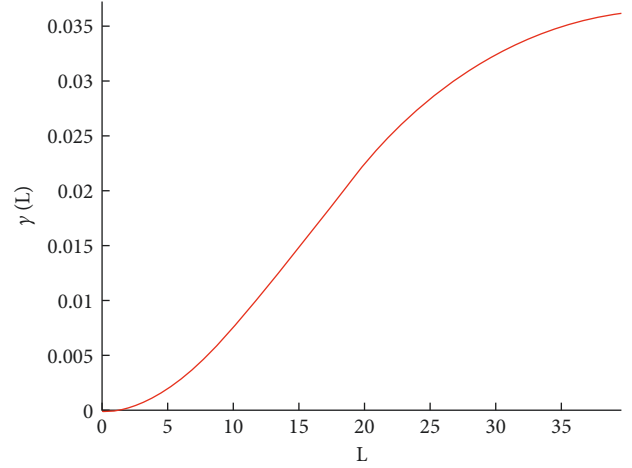


FIGURE 10: Fitted Gaussian model variance function for uncharacteristic smooth data.

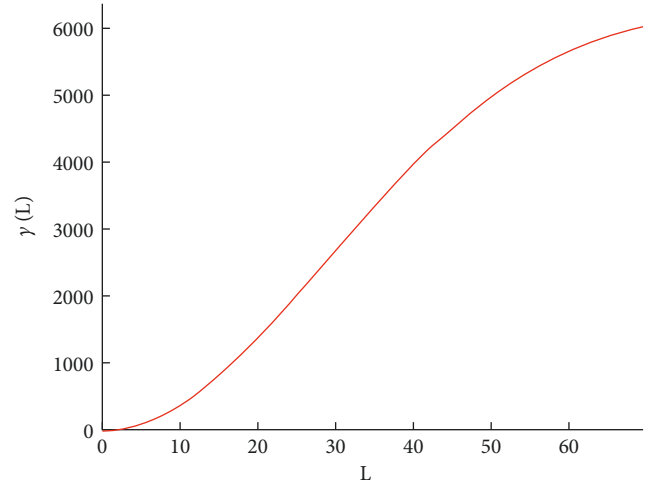


FIGURE 11: Fitted Gaussian model variance function for characteristic mutation data.

5. Comparison of Experimental Result Validation

In this article, two types of data are used, each of which is validated against the inverse distance-weighted interpolation algorithm, the ordinary kriging interpolation algorithm, and the variability-weighted interpolation algorithm based on a fixed-grid sampling of the data. Three tests, mean square error s , mean absolute deviation t , and standard deviation σ , were used to test the interpolation results.

(1) Featureless smooth data

For the featureless smooth data, the inverse distance-weighted interpolation algorithm, the ordinary kriging interpolation algorithm, and the variability-weighted interpolation algorithm based on data fixation sampling were used for interpolation calculation. A comparison of the interpolation results is shown in Table 2, and three-

TABLE 2: Comparison of interpolation results for smooth data without features.

Accuracy errors	IDW	OK	VW
Mean square error	0.0079	0.0026	0.0029
Mean deviation	0.0001910	0.0000015	0.0000538
Standard deviation	0.0138	0.0012	0.0073
Time (s)	11.0815	31.6114	13.8555

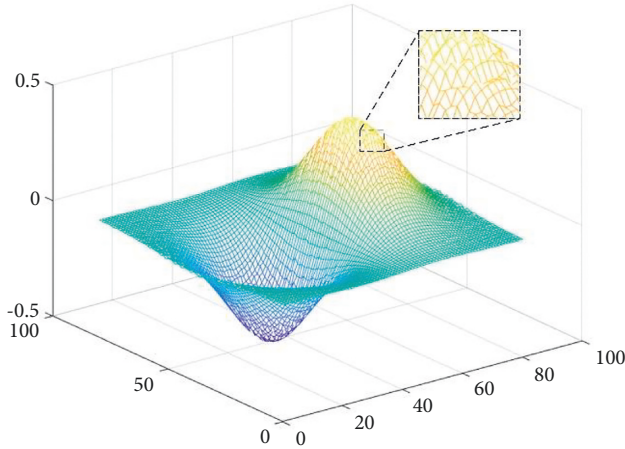


FIGURE 12: IDW interpolation algorithm applied to featureless data.

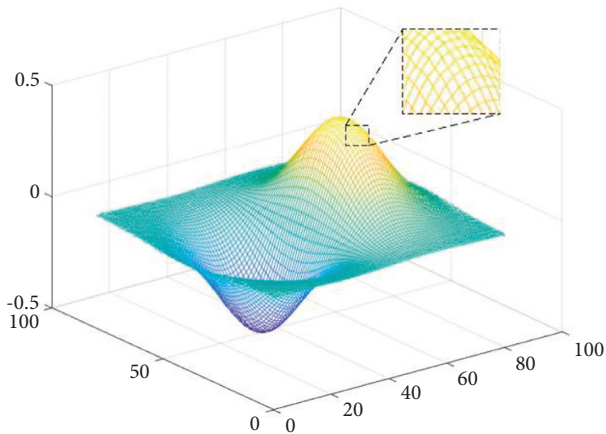


FIGURE 13: OK interpolation applied to featureless data.

dimensional stereograms of the interpolation results are shown in Figures 12–14.

(2) Characteristic mutation data

For the characteristic mutation data, the inverse distance-weighted interpolation algorithm, the ordinary kriging interpolation algorithm, and the variability-weighted interpolation algorithm based on data fixation sampling were used for the interpolation calculation. A comparison of the interpolation results is shown in Table 3 and three-dimensional stereograms of the interpolation results are shown in Figures 15–17.

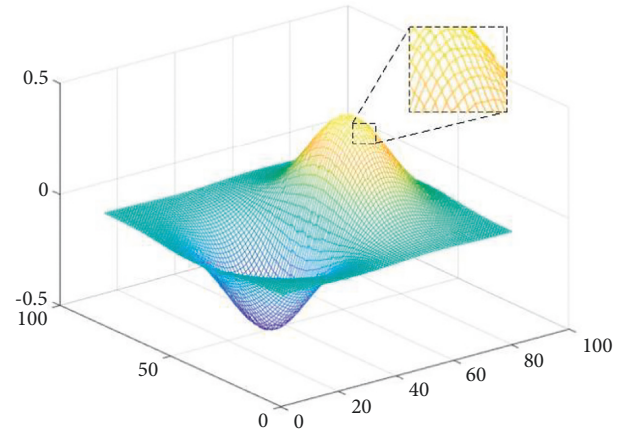


FIGURE 14: VW interpolation algorithm applied to featureless data.

TABLE 3: Comparison of interpolation results for mutation data with features.

Accuracy errors	IDW	OK	VW
Mean square error	3.8650	10.2072	3.1391
Mean deviation	0.0794	0.1860	0.0323
Standard deviation	0.2821	0.4317	0.1800
Time (s)	43.3303	326.7071	49.7731

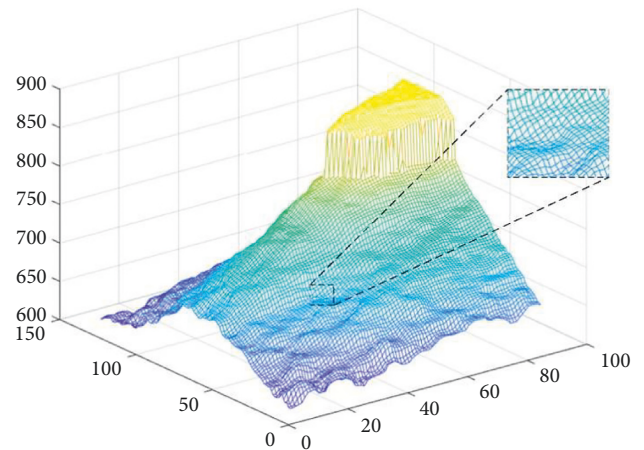


FIGURE 15: IDW interpolation algorithm applied to data with features.

From the above interpolation results, it can be seen that in the experiments on smooth data without features, VW based on data fixation sampling improved the mean square error, mean deviation, and standard deviation by 63.29%, 71.83%, and 47.1%, respectively, compared to VW and improved the efficiency by 56.17% compared to OK. In the experiments on the characterized mutant data, VW based on data fixation sampling improved by 18.78%, 59.32%, and 36.19% relative to IDW in terms of mean square error, mean deviation, and standard deviation, respectively, and relative to OK in terms of mean square error, mean deviation, standard deviation, and efficiency by 69.25%, 82.63%, 58.3%,

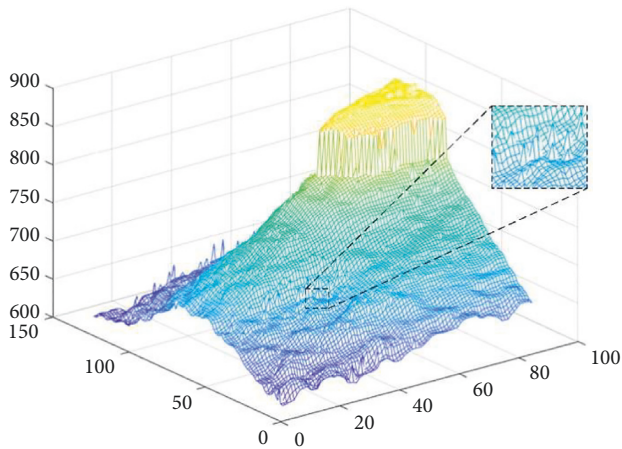


FIGURE 16: OK interpolation applied to data with features.

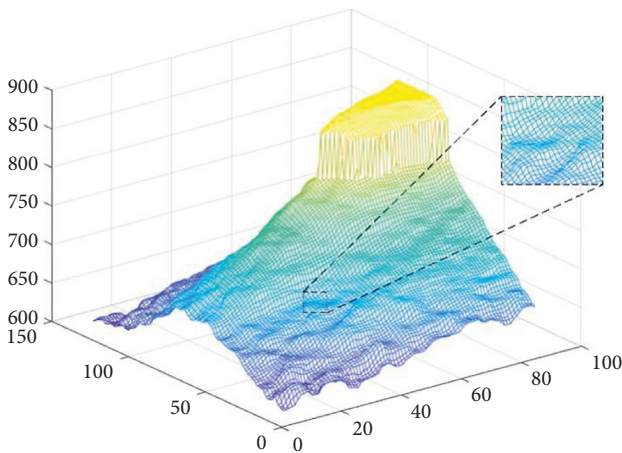


FIGURE 17: VW interpolation algorithm applied to data with features.

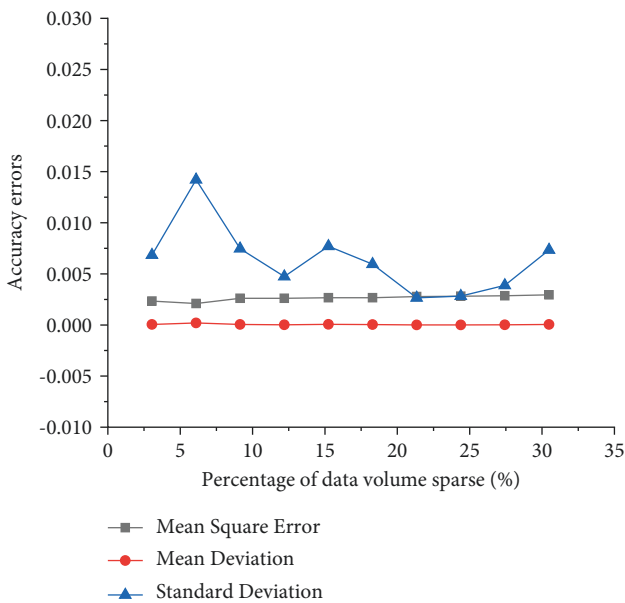


FIGURE 18: Comparison of stability of interpolation results for featureless data.

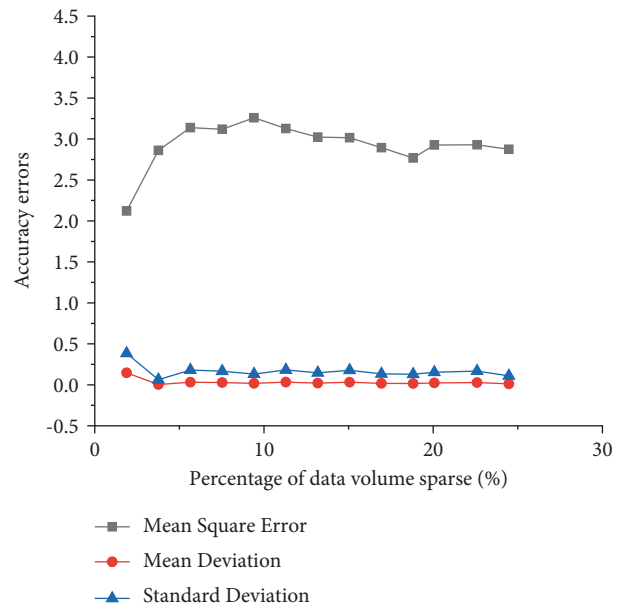


FIGURE 19: Comparison of stability of interpolation results for data with features.

and 84.77%, respectively. Thus, VW based on a fixed-grid sampling of the data can produce high interpolation accuracy and fast interpolation, and it is more suitable for both smooth data without features and mutation data with features. It is worth noting that, for data with features, OK can produce large data fluctuations, resulting in reduced interpolation accuracy. However, VW based on a fixed-grid sampling of the data solves this problem well and provides a significant improvement in efficiency.

In addition to comparing the interpolation results of the three interpolation methods to verify the interpolation accuracy and interpolation efficiency of the variability-weighted interpolation algorithm based on a fixed-grid sampling of the data, the variability-weighted interpolation algorithm was also applied to data with different degrees of sparsity to verify the stability of the interpolation results .

(1) Featureless smooth data

The interpolation results of the variability-weighted interpolation algorithm for featureless data with different sparsity levels are shown in Figure 18.

(2) Characteristic mutated data

The interpolation results of the variability-weighted interpolation algorithm for characteristic data with different sparsity levels are shown in Figure 19.

As can be seen from the stability comparison results, the interpolation accuracy results are relatively stable overall between 5% and 25% sparsity, although individual errors may fluctuate slightly when sparsity varies for both featureless smooth data and feature mutant data. Thus, the interpolation accuracy of the variability-weighted interpolation algorithm is relatively stable for data with varying degrees of sparsity.

6. Conclusions

This article proposes a variability-weighted interpolation algorithm based on data fixation sampling by combining the inverse distance-weighted interpolation algorithm and the ordinary kriging interpolation algorithm. The good performance and stability of the new method provide a good reference for the application of data interpolation in related fields and applications of different original data types. The interpolation efficiency is improved by data lattice sampling, and the interpolation accuracy is improved by variability-weighted interpolation. This article details the sampling search principle, the sampling search process, the interpolation algorithm principle and the interpolation algorithm process, and experimental verification of the proposed interpolation algorithm in terms of interpolation accuracy, efficiency, and interpolation stability using two different types of data. The experimental results show that the variational-weighted interpolation algorithm based on data fixation sampling has high interpolation accuracy and fast interpolation speed for both smooth data without obvious features and mutated data with obvious features. Regardless of the type of data, the variability-weighted interpolation algorithm is substantially faster than the ordinary kriging algorithm in terms of interpolation speed. The interpolation accuracy results are stable and scientifically valid when the sparsity of the original data varies.

The weighting coefficients a, b of the variational-weighted interpolation algorithm based on a fixed-grid sampling of the data used in this experiment have been continuously adjusted through experimentation so that subsequent research can be carried out in the area of weighting coefficient optimization. The adaptive optimization of the weighting coefficients for different types of data allows for the minimization of accuracy errors for all types of input data.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Defense Basic Research Program (JCKY2019204B020).

References

- [1] M. Böhlen, L. Bukauskas, P. Svante Eriksen et al., "3D visual data mining-goals and experiences," *Computational Statistics and Data Analysis*, vol. 43, no. 4, pp. 445–469, 2003.
- [2] R. M. Daniel and C. Dianne, "Visualization of data," *Current Opinion in Biotechnology*, vol. 11, no. 1, pp. 89–96, 2000.
- [3] C. Özkan, "Surface interpolation by adaptive neuro-fuzzy inference system based local ordinary kriging," *Computer Vision - ACCV*, pp. 196–205, 2006.
- [4] J. F. Chai and X. M. Hu, "Research of red tide visualization system based on OpenGL," in *Proceedings of the 2017 2nd International Conference on Image, Vision and Computing (ICIVC)*, pp. 729–733, IEEE, Chengdu, China, August 2017.
- [5] D.-W. Wang, L.-N. Li, C. Hu, Q. Li, X. Chen, and P.-W. Huang, "A modified inverse distance weighting method for interpolation in open public places based on wi-fi probe data," *Journal of Advanced Transportation*, vol. 2019, Article ID 7602792, 11 pages, 2019.
- [6] Z. Zhou, M. Zhang, Y. Wang, C. Wang, and M. Ma, "Application of kriging algorithm based on ACFPSO in geomagnetic data interpolation," *Mathematical Problems in Engineering*, vol. 2019, no. 1, 14 pages, Article ID 1574918, 2019.
- [7] M. Jing and J. Wu, "Fast image interpolation using directional inverse distance weighting for real-time applications," *Optics Communications*, vol. 286, no. Complete, pp. 111–116, 2013.
- [8] T. Zhang, X. Xu, and S. Xu, "Method of establishing an underwater digital elevation terrain based on kriging interpolation," *Measurement*, vol. 63, pp. 287–298, 2015.
- [9] G. Y. Lu and D. W. Wong, "An adaptive inverse-distance weighting spatial interpolation technique," *Computers and Geosciences*, vol. 34, no. 9, pp. 1044–1055, 2008.
- [10] H. Hu and H. Shu, "An improved coarse-grained parallel algorithm for computational acceleration of ordinary Kriging interpolation," *Computers and Geosciences*, vol. 78, no. C, pp. 44–52, 2015.
- [11] H. Zhang, Y. Zhan, J. Li et al., "Using Kriging incorporated with wind direction to investigate ground-level PM2.5 concentration," *The Science of the Total Environment*, vol. 751, Article ID 141813, 2021.
- [12] P. Chen, F. Deng, and S. Liu, "Research of three-dimensional space attribute interpolation methods," *Computer Knowledge and Technology*, vol. 11, no. 7, pp. 235–239, 2015.
- [13] J. Y. Chang and M. G. Kang, "Edge dependent interpolation based on adaptive search range," in *Proceedings of the IEEE Conference. The Institute of Electronics and Information Engineers*, pp. 803–804, Tokyo, Japan, March 2008.
- [14] H. Huang, C. Cui, L. Cheng, Q. Liu, and J. Wang, "Grid interpolation algorithm based on nearest neighbor fast search," *Earth Science India*, vol. 5, no. 3–4, pp. 181–187, 2012.
- [15] J. M. Zhang, L. P. Guo, and X. D. Zhang, "Effects of interpolation parameters in inverse distance weighted method on DEM accuracy," *Journal of Geomatics Science and Technology*, vol. 29, no. 1, pp. 51–56, 2012.
- [16] J. M. Zhang, X. You, and G. Wan, "Experimental research on optimization of DEM interpolation parameters," *Acta Geodaetica et Cartographica Sinica*, vol. 43, no. 2, pp. 178–185, 2014.
- [17] C. Wu, G. Liu, and C. Huang, "Prediction of soil salinity in the Yellow River Delta using geographically weighted regression," *Archives of Agronomy and Soil Science*, vol. 63, no. 7, pp. 928–941, 2017.
- [18] G. A. Achilleos, "The Inverse Distance Weighted interpolation method and error propagation mechanism - creating a DEM from an analogue topographical map," *Journal of Spatial Science*, vol. 56, no. 2, pp. 283–304, 2011.
- [19] M. F. Hutchinson, "Interpolating mean rainfall using thin plate smoothing splines," *International Journal of Geographical Information Systems*, vol. 9, no. 4, pp. 385–403, 1995.

- [20] A. S. Agung Setianto and T. T. Tamia Triandini, "Comparison of kriging and inverse distance weighted (IDW) interpolation methods in lineament extraction and analysis," *Journal of Southeast Asian Applied Geology*, vol. 5, no. 1, pp. 21–29, 2013.
- [21] H. L. Cho and J. C. Jeong, "Application of spatial interpolation to rainfall data," *Spatial Information Research*, vol. 14, no. 1, pp. 29–41, 2006.
- [22] S. Peng, "Developments of Spatio-Temporal Interpolation Methods for Meteorological Elements," Master Dissertation, Central South University, Changsha, China, 2010.
- [23] K. Y. Dai, G. R. Liu, K. M. Lim, and Y. T. Gu, "Comparison between the radial point interpolation and the Kriging interpolation used in meshfree methods," *Computational Mechanics*, vol. 32, no. 1-2, pp. 60–70, 2003.
- [24] N. Peng, J. Mo, and C. Han, "Volume rendering of spatial scattered data based on spatial data interpolation algorithms," *Journal of Computer Applications*, vol. 28, no. 07, pp. 1759-1760, 2008.