Hindawi
Advances in Meteorology
Volume 2020, Article ID 9185283, 14 pages
https://doi.org/10.1155/2020/9185283

Research Article

A Spatial Interpolation of Meteorological Parameters considering Geographic Semantics

Wenjun Wu,[1,2] Ruijie Gan,[1,3] Junli Li ,[1,4] Xiu Cao,[1] Xinxin Ye,[4] Jie Zhang,[1] and Hongjiao Qu[1]

[1]*School of Resources and Environment, Anhui Agricultural University, Hefei 230036, China*
[2]*Huangshan Meteorology Bureau, Huangshan, Anhui 245000, China*
[3]*Nanchang Meteorological Administration, Nanchang, Jiangxi 330038, China*
[4]*Anhui Province Key Lab of Farmland Ecological Conservation and Pollution Prevention, Hefei 230036, China*

Correspondence should be addressed to Junli Li; 85401500@qq.com

Received 27 August 2019; Revised 26 March 2020; Accepted 12 August 2020; Published 2 September 2020

Academic Editor: Maria Ángeles García

Spatial interpolation of meteorological parameters, closely related to the earth surface, plays important roles in climatological study. However, most of traditional spatial interpolation methods ignore the geographic semantics of interpolation sample points in practical application. This paper attempts to propose an improved inverse-distance weighting interpolation algorithm considering geographic semantics (S-IDW), which adds geographic semantic similarity to the traditional IDW formula and adjusts weight coefficient. In the interpolation process, the geographic semantic differences between sample points and estimation points are considered comprehensively. In this study, 3 groups of land surface temperature data from 2 different areas were selected for experiments, and several commonly used spatial interpolation methods were compared. Experimental results indicated that S-IDW outperformed IDW and several existing spatial interpolation methods, but there were also some abnormal value and interpolation outliers. This method provides a new insight toward the estimation accuracy, data missing, and error correction of spatial attributes related to meteorological parameters.

1. Introduction

Spatial interpolation of meteorological parameters is to obtain relatively accurate descriptions of spatial attributes related to climatological dynamics and weather patterns by using some reasonably located samples [1]. Traditionally, sampling observation is the best way to obtain the regional mean conditions in order to ensure equal sampling opportunities for each location in the region. However, the observation sampling points are sparse and of random distribution in practical application [1]. For example, the location of the sample points is systematic and changes smoothly. Furthermore, most meteorological models are obtained by sampling from observation stations at present. Spatial interpolation method is widely used to transform discrete observation data into continuous surface so as to better measure the spatial distribution pattern of data elements [2]. Currently, familiar spatial interpolation methods, such as IDW, Kriging, Spline, and trend surface method, have been widely used in different fields. Most of them have some limitations in application, such as distance weighting method with some problems, which affects calculation results due to distance, being not suitable for a large range [3]. Kriging method can adopt different variogram forms and parameters for different sampling data points, with certain flexibility. However, it loses the high efficiency of the original inverse-distance weighting method by first determining the variogram form and fitting the parameters of variogram. Kriging variograms require artificial selection, and there exists the problem that computation increases sharply when there are too many combinations of variograms [4]. Spline method is not suitable for sparse and finite sampling points and is often used for high-density sample point interpolation [5]. The trend surface method relies more on the existing spatial distribution trend of interpolation elements [6]. Consequently, many authors have carried on the continuous exploration and improvement to the spatial interpolation method [7]. For instance, the complexity of terrain and elevation factor were introduced by some researchers into the inverse-distance weighting [8, 9], and Li et al. brought the harmonic weighting coefficient of azimuth into the distance weighting interpolation [2]. The natural neighborhood relationship was led into the distance weighting interpolation [10], and some authors introduced the fuzzy trigonometric function into the distance weighting interpolation [3]. Others took into account the spatiotemporal variation characteristics of geographical factors and introduced time-series data to remove some numerical fluctuations in time, such as spatiotemporal weighting Kriging and spatiotemporal inverse-distance weighting interpolation [9]. The succession of methods proposed by the above-mentioned authors had achieved remarkable academic impact and showed high spatial autocorrelation, but most of them were based on numerical interpolation methods, without considering geographic semantics.

Inspired by the gradient theory in the field of image processing, the gradient is the first-order differential of gray value, reflecting the change rate between adjacent pixels in the direction of X and Y [11]. Where the gradient change rate of image is larger in the region, the types of land cover tend to change, such as the boundary between land and water in the image. Existing research based on remote sensing image inversion, such as land surface temperature (LST), vegetation index, and moisture index, is to some extent a model to describe the relationship between remote sensing signals or remote sensing data and surface applications [12]. For example, the temperature nearby residential buildings is quite different from forest land or water body. The air temperature of some exposed land surfaces, like build roofs and pavement, is hotter than that of the shades of forests. Therefore, geographic semantics are indispensable to exploring geospatial description of surface remote sensing pixel information. Currently, some authors have put forward semantic Kriging method, which has achieved excellent research results, but there are still some problems such as the complexity of calculating the variogram imported by the semantic similarity [13–15]. In addition, the prediction of multivariable meteorological factors by embedding geographic semantics into Bayesian networks weakens the influence of parameter uncertainty but lacks the knowledge of meteorological modeling [16]. Although the aforementioned spatial interpolation methods show performance in different applications, there still exists scope of improvements by introducing geographic semantics into spatial interpolation process. Furthermore, information semantics are growing in the field of spatial statistics and environmental modeling [17, 18].

This paper attempts to introduce the geographic semantics into inverse weighting spatial interpolation by embedding hierarchical geographic semantics into spatial interpolation model and using semantic similarity to measure factor weight. The following analyses were carried out in this study: (1) the S-IDW methods used in this study are explained in the next section; (2) the findings are discussed in the Experimental Results and Comparison section; (3) finally, our conclusions and subsequent research are drawn in the Conclusions section.

2. Methodology

The S-IDW integrates the geographic semantic knowledge into the inverse-distance weighting interpolation method. Considering the effect of distance on interpolation results, the influence of land-use type on land surface temperature interpolation is added. The S-IDW reconsiders the interpolation weight, increases the weight of the same land-use type, and reduces the weight of different land-use types on the basis of distance, constructing the S-IDW method [19].

In the S-IDW, the first step is to calculate the semantic similarity of geographic entities. The formula is as follows:

$$T = \sum_{i=1}^{m} \varphi_i T_i, \tag{1}$$

$$\varphi_i = \frac{\left(\kappa_j/d_j\right)^n}{\sum_{i=1}^{m}\left(\kappa_j/d_j\right)^n}, \tag{2}$$

$$\kappa_j = \begin{cases} 1\,(\text{land use types are the same}), \\ 0 \sim 1\,(\text{land use types are different}), \end{cases} \tag{3}$$

$$d_i = \sqrt{\left(x_i - x\right)^2 + \left(y_i - y\right)^2}. \tag{4}$$

In equations (1)–(4), T is the estimated value of the ith point to be interpolated;T_i is the measured value of the ith discrete point; d_i is the distance between the ith discrete point and the ith point to be interpolated; x_i is the latitude of the point to be interpolated; x is latitude of discrete points; y_i is the longitude of the point to be interpolated; y is the longitude of the discrete point; m is the number of measured sample points participating in the interpolation; n is the power exponent, which controls the degree to which the weight coefficient decreases with the increase of the distance between the point to be interpolated and the sample point. When n is larger, the closer sample point is endowed with higher weight; when n is smaller, the weight is more evenly distributed to all sample points. When $n = 1$, it is called inverse-distance weighting method, which is a common and simple spatial interpolation method. When $n = 2$, it is called inverse-distance squared method, which is often used in practical application. In this study, $n = 2$ is taken.

κ_j is the semantic similarity between the jth point to be interpolated and the ith discrete point, and the value range is $0 < \kappa_1 \leq 1$. Semantic similarity refers to the degree to which two concepts can replace each other in the same context without changing the semantic structure of the text [19]. The larger the change of semantic structure, the smaller the similarity; the smaller the change of semantic structure, the greater the similarity. In this study, a comprehensive semantic similarity algorithm for geographic ontology is adopted. On the basis of analyzing the influencing factors of semantic distance similarity, the weighted sum method is used to calculate semantic distance similarity, concept attribute similarity, and information similarity. The calculation formula is as follows [20]:

$$\kappa_j = \lambda_1 * \kappa_j^{\text{dist}} + \lambda_2 * \kappa_j^{\text{att}} + \lambda_3 * \kappa_j^{\text{ic}}, \tag{5}$$

$$\kappa_j^{\text{dist}} = \frac{\mu}{\text{Dist}(a,b) + \mu}, \tag{6}$$

$$\kappa_j^{\text{att}} = \frac{\text{count}[\text{att}(a) \cap \text{att}(b)]}{\text{count}[\text{att}(a) \cup \text{att}(b)]}, \tag{7}$$

$$\kappa_j^{\text{ic}} = \frac{2 * \text{ic}[\text{LCA}(a,b)] + \chi}{\text{ic}(a) + \text{ic}(b) + \chi}, \tag{8}$$

$$\text{ic}(a) = -\log p(a). \tag{9}$$

The semantic similarity is calculated by referring to the hierarchical structure of geographical entities in Table 1 and Figure 1. In formulas (5)–(8), $\text{dist}(a,b)$ is the semantic distance, which refers to the shortest path between any two concept nodes a and b in the ontology hierarchy, and η is the regulating factor. In this paper, $\eta = 8$. κ_j^{att} represents the similarity of concept attributes between concept nodes a and b. The function att is the set of entity attributes, and count is the number of attributes. In addition, $\chi > 0$, χ is a real number, and its value is controlled at [0, 2max (IC (a, b))], where $\chi = 1$. The information quantity $\text{ic}(a)$ is defined as the −log function of the occurrence probability of concept a. In equation (5), when the land-use types of the jth point to be interpolated and the jth discrete point are equivalent,$\kappa_j = 1$; when the land-use types of the jth point to be interpolated and the jth discrete point are not equivalent, $0 < \kappa_j < 1$.

In equation (5), $\lambda_1 = 0.1$, $\lambda_2 = 0.8$, and $\lambda_3 = 0.1$; λ_1, λ_2, and λ_3 are the adjustment coefficients of semantic distance similarity, concept attribute similarity, and information similarity, respectively; $\lambda_1 + \lambda_2 + \lambda_3 = 1$. a and b are geographical entities. For the convenience of calculation, GB/T21010-2017 classification of land use in China and its meaning are used to extract geographical entities. The semantic attributes of geographical entities are shown in Table 1, and the ontological hierarchical network structure of land-use status classification is shown in Figure 1. Based on equation (5), the semantic similarity for some geographical entities of land-use type ontology is calculated as shown in Table 2.

3. Experimental Results and Comparison

3.1. Experimental Design and Error Metric. The experimental study has been carried out using land surface temperature (LST) data from Landsat 8 OLI-TIRS satellite. Due to the complex and changeable surface environment, LST shows different characteristics in different surface environments. In order to explore the spatial interpolation accuracy in different areas and different land surface temperatures, LST at diverse time intervals were selected in the 2 study areas, and 3 distinct LST conditions of high temperature, low temperature, and normal temperature were used to carry out experiments. The interpolation accuracy of traditional numerical interpolation methods is often closely related to the density and sparsity of the discrete points. In this paper, the discrete points and the points to be valued are selected randomly and distributed evenly. 15 points to be valued and 60 discrete points were randomly selected in the experiment. Assuming that the LST values of the 15 points are missing or abnormal, we use 60 discrete points of known LST values to interpolate the 15 points in order to compensate for and correct the missing or abnormal values. The popular approaches for spatial interpolation include Kriging, IDW, Natural, Spline, and S-IDW. On this basis, we compared and analyzed the results of 5 interpolation methods with the original LST values of 15 points to be valued. As shown in Figure 2, the experimental flow chart of semantic inverse-distance weighting interpolation is shown.

In the experiment, the accuracy of the estimated value is evaluated by means of root mean squared of errors (RMSE) [21], mean absolute error (MAE), mean absolute percentage error (MAPE) [16], and the ratio of variance of the estimated values to variance of the observed values (RVAR) [22, 23]. The formal definition of each indicator is given as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(T_i - T)^2}, \tag{10}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|T_i - T\right|, \tag{11}$$

$$\text{MAPE} = \frac{\left|M_{T_i}^s - M_T^s\right|}{\left|M_{T_i}^s\right|}, \tag{12}$$

$$\text{RVAR} = \frac{\text{var}(T^S)}{\text{var}(T_i^S)}. \tag{13}$$

In formulas (10)–(13), n is the total number of measured values; T_i is the measured value of the ith discrete point; T is the estimated value of the ith point to be interpolated; $M_{T_i}^s$ is the average value of measured value at discrete point; M_T^s is the average value of the estimated value of the point to be evaluated; $\text{var}(T^S)$ is the variance of the estimated value of the point to be interpolated; $\text{var}(T_i^s)$ is the variance of measured value at discrete point. The best fitting between measured value and estimated value under ideal conditions can be obtained as follows: RMSE≈0, MAE≈0, MAPE≈0, and RVAR≈1.

3.2. Analysis and Discussion of the Interpolation Results in Study Area-1. In order to verify the interpolation effect of S-IDW under three temperature environments, the imaging dates of the remote sensing images in study area-1 are January 11, 2018, April 17, 2018, and August 9, 2013. The corresponding image clouds are 0.54%, 0.05%, and 4.76%,

Table 1: Part attributes of geographical entity in land-use type ontology.

Geographic entities	Parent class	Parent use	Specific uses	Cover (or location)	Operational characteristics
01 Farmland	Land	Agricultural production	Planting crops	Crops	
02 Garden plot	Land	Agricultural production	Planting and collecting fruit, leaf, and root crops	Perennial woody plants and herbs	Intensive management
03 Woodland	Land	Agricultural production	Growing trees, bamboos, shrubs, and coastal mangroves	Woody plants	Production management, ecological management
04 Grassland	Land	Agricultural production	Growing herbs	Herbs	
05 Commercial land	Land	Building construction	Constructing business and service industry	Houses, buildings	
06 Industrial mining warehouse land	Land	Building construction	Constructing industrial production and material storage sites	Houses, buildings	
07 Residential land	Land	Building construction	Constructing living places	Houses, buildings	
08 Public management and service land	Land	Building construction	Constructing public management and public service places	Houses, buildings	
09 Special land	Land	Building construction	Building construction	Houses, buildings	
10 Transport land	Land	Building construction	Constructing ground routes and stations for transport		State-owned
11 Waters and water conservancy facility land	Land		Hydraulic structures of land waters, shallows, ditches, and marshes		
12 Other lands	Land				

respectively. Under these conditions, LST inversion is carried out and the inversion data are extracted and processed. The distribution of 60 discrete points and 15 points to be valued in study area-1 is shown in Figure 3.

Interpolation data results and interpolation accuracy of five methods under low-temperature conditions in study area-1 can be seen in Tables 3–5. As shown in Table 3, the interpolation results of S-IDW for 8 of the 15 points to be valued are closer to the land surface temperature than those of the other 4 interpolation methods. Generally, through the mathematical statistics analysis and Pearson correlation analysis of the five interpolation methods, it is indicated that the MAE, MAPE, and RMSE of S-IDW are closer to the best fitting values between measured and estimated values under ideal conditions than the other 4 interpolation methods. As far as RVAR is concerned, Natural and Spline are better than S-IDW, but S-IDW is better than Kriging and IDW. In terms of Pearson correlation, the results of S-IDW, Kriging, IDW, and Natural interpolation are significantly correlated with LST at 0.01 level (two-tailed), of which the correlation coefficient r between S-IDW interpolation results and LST is 0.959, with the strongest correlation, and the significant correlation coefficient r between Spline interpolation results and LST was 0.616 at 0.05 level (two-tailed), with the weakest correlation.

Interpolation data results and accuracy of five methods under normal temperature conditions in study area-1 can be seen in Tables 6–8. As shown in Table 6, the interpolation results of S-IDW for 8 of the 15 points to be estimated are closer to the land surface temperature than those of the other 4 interpolation methods. Generally, through the mathematical statistics analysis and Pearson correlation analysis of the 5 interpolation methods, it is found that the MAE, MAPE, and RMSE of S-IDW are closer to the best fitting values between measured and estimated values under ideal conditions than the other 4 interpolation methods. In terms of MAPE, Natural is better than S-IDW, but S-IDW is better than Kriging, IDW, and Spline. In terms of RVAR, Spline is better than S-IDW, but S-IDW is better than Kriging. In terms of Pearson correlation, the results of S-IDW, Kriging, IDW, Natural, and Spline are significantly correlated with LST at 0.01 level (two-tailed), of which the correlation coefficient r between S-IDW interpolation results and LST is 0.930, with the strongest correlation.

The interpolation data results and accuracy of five methods under high-temperature conditions in study area-1 can be seen in Tables 9–11. As shown in Table 9, the interpolation results of the S-IDW for 4 of the 15 points to be valued are closer to the land surface temperature than those of the other 4 interpolation methods. Generally, through the mathematical statistics analysis and Person correlation analysis of the 5 interpolation methods, it is indicated that the MAE and RMSE of S-IDW are closer to the best fitting values between measured and estimated values under ideal conditions than the other 4 interpolation methods. As far as MAPE is concerned, 1.252% of S-IDW is higher than 0.781% of IDW and 0.057% of Spline but lower than 1.818% of Kriging and 1.253% of Natural. In terms of RVAR, 0.983 of Natural is better than 0.827 of S-IDW, but S-IDW is better than Kriging and Natural. In terms of Person correlation, the interpolation results of S-IDW, Kriging, IDW, Natural, and Spline are significantly correlated with LST at 0.01 level (two-tailed), of which the correlation coefficient r between S-IDW and LST was 0.914, stronger than 0.843 of IDW, 0.794 of Kriging, 0.791 of Natural, and 0.669 of Spline.

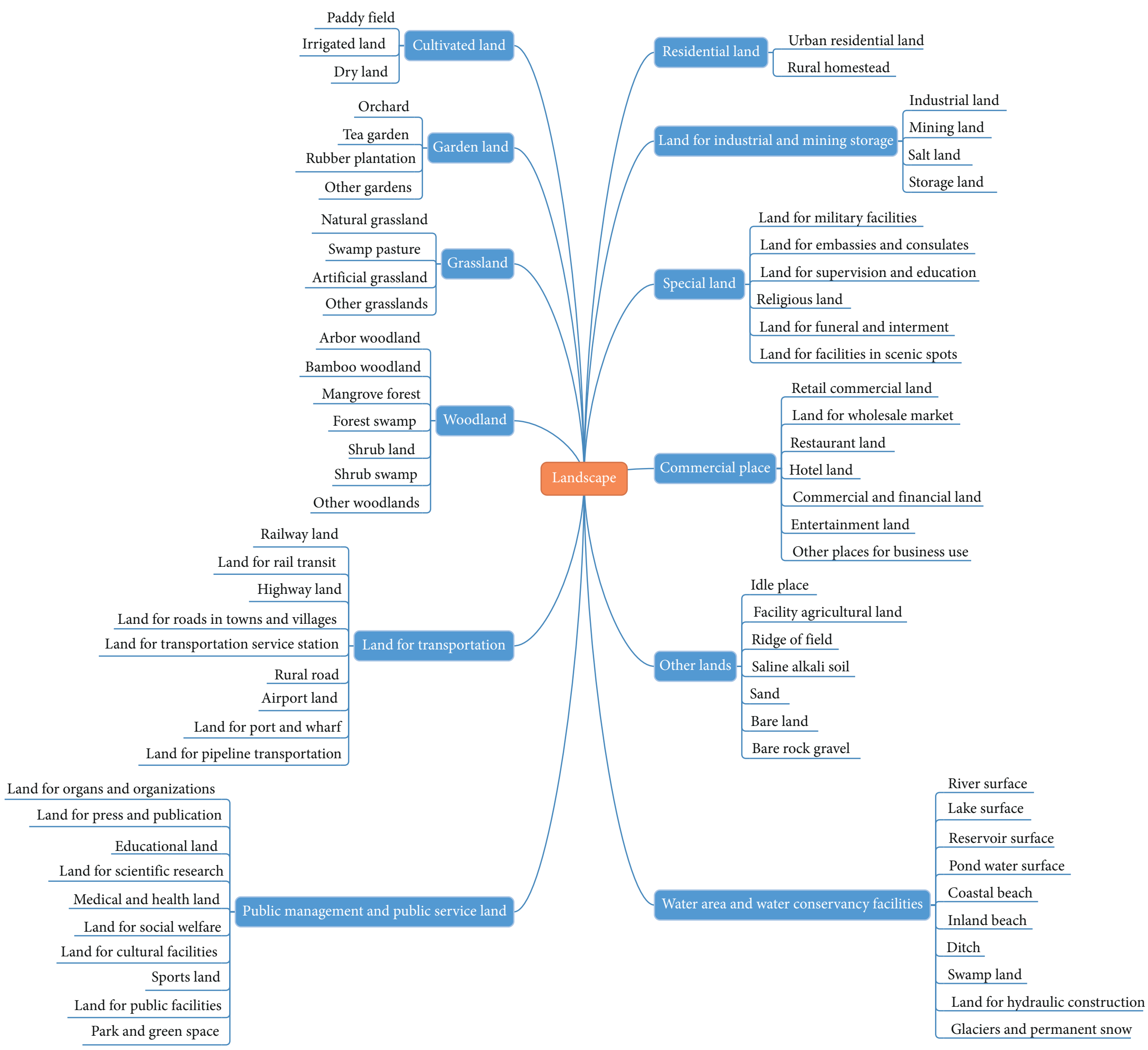

Figure 1: Part network hierarchy of land-use classification ontology.

Table 2: Semantic similarity for geographic entity of land-use type ontology.

Geographic entities	01 Farmland	02 Garden plot	03 Woodland	04 Grassland	07 Residential land	10 Transport land	11 Waters and water conservancy facility land
01 Farmland	1	0.4198	0.4218	0.4198	0.2576	0.2628	0.2632
02 Garden plot	0.4198	1	0.4229	0.4207	0.2585	0.264	0.2645
03 Woodland	0.4218	0.4229	1	0.4229	0.2597	0.2667	0.2665
04 Grassland	0.4198	0.4207	0.4229	1	0.2585	0.264	0.2645
07 Residential land	0.2576	0.2585	0.2597	0.2585	1	0.4212	0.2616
10 Transport land	0.2628	0.264	0.2667	0.264	0.4212	1	0.2685
11 Waters and water conservancy facility land	0.2632	0.2645	0.2665	0.2645	0.2616	0.2685	1

3.3. Analysis and Discussion of Interpolation Results in Study Area-2. In order to verify the interpolation effect of S-IDW under three temperature environments, the imaging dates of the remote sensing images in study area-1 are February 11, 2017, April 19, 2018, and July 10, 2013. The corresponding image clouds are 0.54%, 0.05%, and 4.76%, respectively. Under these conditions, LST inversion is carried out and the inversion data are extracted and processed. The distribution of 60 discrete points and 15 points to be valued in study area-2 is shown in Figure 4.

Interpolation data results and accuracy of 5 methods under low temperature conditions in study area-2 can be

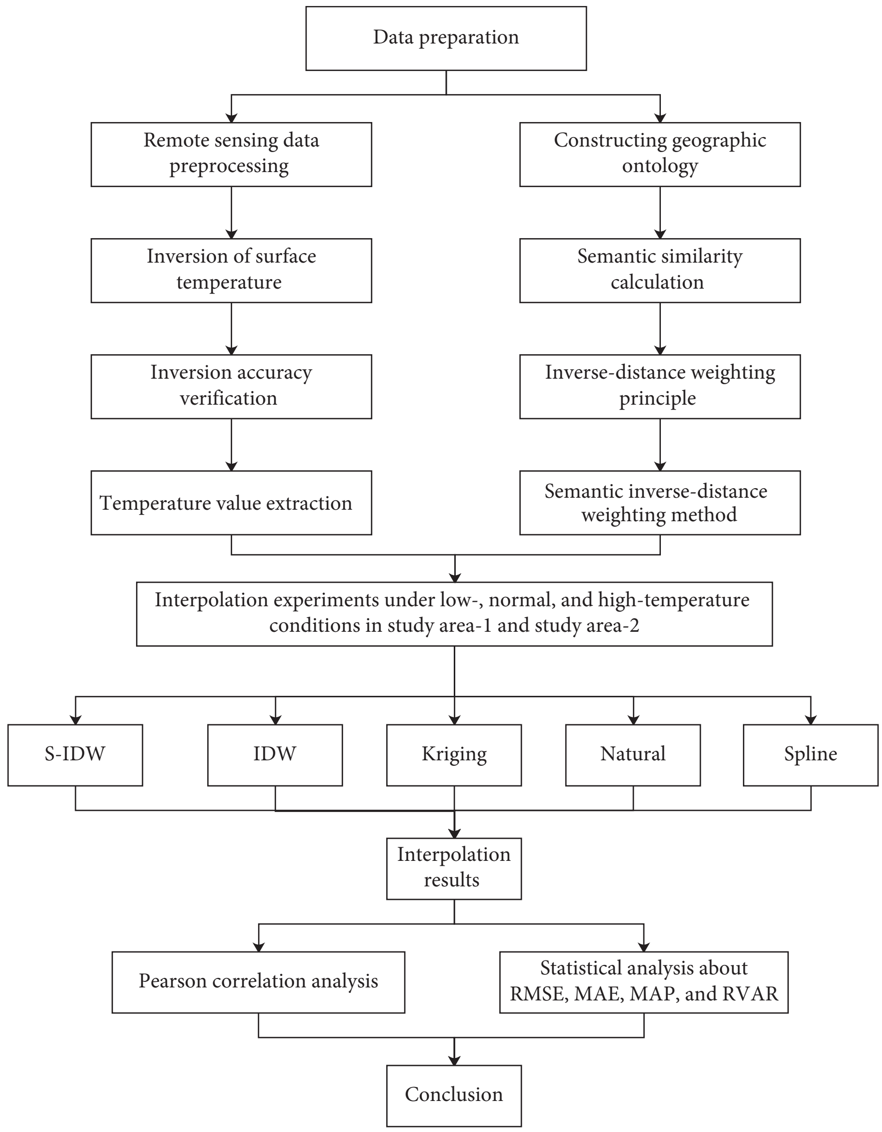

Figure 2: Experimental flow chart of semantic inverse-distance weighting interpolation.

seen in Tables 12–14. As shown in Table 12, the interpolation results of S-IDW for 9 of the 15 points to be valued are closer to the land surface temperature than those of the other 4 interpolation methods. The interpolation result of S-IDW at point 11 to be valued is 9.052°C, which deviates from the LST value more than the interpolation results of other 4 interpolation methods. Generally, through the mathematical statistics analysis and Pearson correlation analysis of the 5 interpolation methods, it is known that the MAE, MAPE, RVAR, and RMSE of S-IDW are closer to the best fitting values between measured and estimated values under ideal conditions than those of the other 4 interpolation methods. In terms of Pearson correlation, the results of S-IDW, Kriging, IDW, Natural, and Spline are significantly correlated with LST at 0.01 level (two-tailed), of which the correlation coefficient r between S-IDW interpolation results and LST is 0.890.

The interpolation data results and accuracy of 5 methods under normal temperature conditions in study area-2 can be seen in Tables 15–17. As shown in Table 15, the interpolation results of S-IDW for 6 of the 15 points to be valued are closer to the land surface temperature than those of the other 4 interpolation methods. The LST value of point 11 to be valued is 25.245°C, and the S-IDW interpolation result of this point to be valued is 30.266°C, which deviates from the LST value more than the interpolation results of the other 4 interpolation methods. Generally, through the mathematical statistics analysis and Pearson correlation analysis of the 5 interpolation methods, it is known that the MAE and RMSE of S-IDW are closer to the best fitting values between measured and estimated values under ideal conditions than those of the other 4 interpolation methods. In terms of MAPE, 1.856% of S-IDW is higher than 1.723% of Natural but lower than Kriging, IDW, and Spline. As far as RVAR is concerned, 1.051 of Natural and 0.844 of IDW are closer to the best fitting values between measured values and estimated values under ideal conditions than 0.806 of S-IDW. In terms of Pearson correlation, the results of S-IDW and

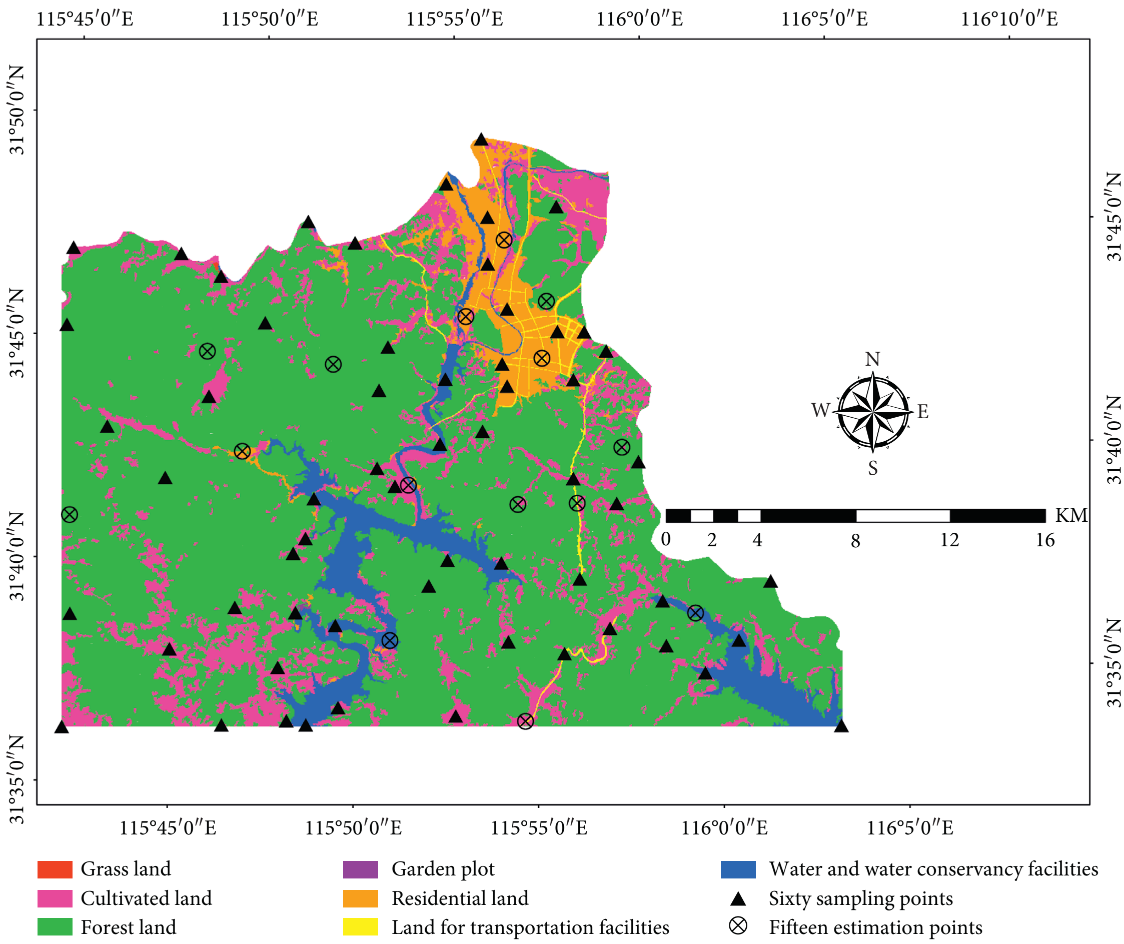

FIGURE 3: Distribution map of discrete points and pending valuation points in study area-1.

TABLE 3: Comparison of interpolation results of 5 kinds of interpolation methods on Jan. 11, 2018.

Study area-1		Forecast date: January 11, 2018			Forecasting methods (unit of temperature: °C)				
Point to be valued	Land-use types	Lat	Lon	LST	S-IDW	IDW	Kriging	Natural	Spline
1	Farmland 1	31.743	115.91	1.428	1.76	1.552	1.572	1.577	1.356
2	Farmland 2	31.688	115.974	3.624	2.569	2.641	2.695	2.604	2.234
3	Farmland 3	31.682	115.876	2.394	1.924	1.909	2.305	2.294	2.028
4	Farmland 4	31.671	115.924	2.429	2.43	2.852	2.97	2.741	2.003
5	Woodland 1	31.745	115.947	2.186	1.840	1.729	1.544	1.507	1.306
6	Woodland 2	31.741	115.792	1.163	1.116	0.914	0.933	0.897	0.953
7	Woodland 3	31.731	115.848	1.226	1.533	1.483	1.323	1.296	0.573
8	Woodland 4	31.685	115.722	1.282	1.231	0.962	0.714	0.359	0.405
9	Residential land 1	31.77	115.931	1.703	1.168	1.143	1.095	1.049	0.774
10	Residential land 2	31.724	115.942	1.631	2.083	2.075	2.067	2.07	1.807
11	Residential land 3	31.702	115.803	2.113	2.108	1.515	1.365	1.475	0.299
12	Transport land 1	31.669	115.951	2.364	2.348	2.26	2.348	2.095	1.923
13	Transport land 2	31.59	115.917	2.324	2.366	2.232	1.555	5.592	5.458
14	Waters and water conservancy facility land 1	31.625	115.86	5.125	5.431	4.373	4.054	3.927	-0.376
15	Waters and water conservancy facility land 2	31.623	115.999	7.96	6.284	5.524	5.617	5.753	7.651

Table 4: Comparison of interpolation accuracy of 5 kinds of interpolation methods on Jan. 11, 2018.

Study area	Forecasting methods	Forecast date	January 11, 2018		
		RMSE	MAE	MAPE (%)	RVAR
Study area-1	S-IDW	0.581	0.376	7.091	0.674
	IDW	0.784	0.552	14.859	0.492
	Kriging	0.830	0.615	17.447	0.512
	Natural	1.173	0.813	9.544	0.788
	Spline	1.804	1.145	27.103	1.326

Table 5: Correlation analysis of interpolation results of 5 kinds of interpolation methods on Jan. 11, 2018.

Study area-1		LST	S-IDW	IDW	Kriging	Natural	Spline
LST	Pearson correlation	1	0.959**	0.954**	0.944**	0.763**	0.616*
	Significance (two-tailed)		0.000	0.000	0.000	0.001	0.015
	N	15	15	15	15	15	15

**Significant correlation at 0.01 level (two-tailed). *Significant correlation at 0.05 level (two-tailed).

Table 6: Comparison of interpolation results of 5 kinds of interpolation methods on Apr. 17, 2018.

Study area-1		Forecast date: January 11, 2018			Forecasting methods (unit of temperature: °C)				
Point to be valued	Land-use types	Lat	Lon	LST	S-IDW	IDW	Kriging	Natural	Spline
1	Farmland 1	31.743	115.91	28.266	26.247	27.014	26.981	25.915	28.132
2	Farmland 2	31.688	115.974	24.056	25.687	26.022	25.977	26.251	27.399
3	Farmland 3	31.682	115.876	27.657	26.708	27.342	24.66	26.983	28.251
4	Farmland 4	31.671	115.924	25.586	25.069	24.758	24.761	24.903	27.626
5	Woodland 1	31.745	115.947	27.254	27.542	28.959	28.502	28.852	28.134
6	Woodland 2	31.741	115.792	24.079	24.049	24.341	23.972	24.74	24.979
7	Woodland 3	31.731	115.848	22.263	23.859	24.202	24.382	23.472	23.559
8	Woodland 4	31.685	115.722	21.083	22.450	22.473	22.447	23.386	19.394
9	Residential land 1	31.77	115.931	30.309	29.533	29.472	28.717	29.137	28.786
10	Residential land 2	31.724	115.942	28.418	28.909	28.882	28.341	29.631	30.642
11	Residential land 3	31.702	115.803	26.475	24.428	23.534	23.277	23.506	26.003
12	Transport land 1	31.669	115.951	26.794	26.034	26.471	25.365	26.551	26.961
13	Transport land 2	31.59	115.917	26.191	24.145	23.216	23.356	21.925	22.816
14	Waters and water conservancy facility land 1	31.625	115.86	20.925	21.042	21.693	21.937	21.731	24.497
15	Waters and water conservancy facility land 2	31.623	115.999	17.673	19.391	20.451	20.954	18.998	17.321

Table 7: Comparison of interpolation accuracy of 5 kinds of interpolation methods on Apr. 17, 2018.

Study area	Forecasting methods	Forecast time	Apr. 17, 2018		
		RMSE	MAE	MAPE (%)	RVAR
Study area-1	S-IDW	1.295	1.090	0.514	0.640
	IDW	1.668	1.383	3.329	0.651
	Kriging	1.959	1.686	0.901	0.497
	Natural	1.886	1.578	0.278	0.763
	Spline	1.891	1.504	1.981	1.127

Table 8: Correlation analysis of interpolation results of 5 kinds of interpolation methods on Apr. 17, 2018.

Study area-1		LST	S-IDW	IDW	Kriging	Natural	Spline
LST	Person correlation	1	0.930**	0.867**	0.817**	0.824**	0.859**
	Significance (two-tailed)		0.000	0.000	0.000	0.000	0.000
	N	15	15	15	15	15	15

**Significant correlation at 0.01 level (two-tailed).

Table 9: Comparison of interpolation results of 5 kinds of interpolation methods on Aug. 9, 2013.

Study area-1		Forecast date: Aug. 9, 2013			Forecasting methods (unit of temperature: °C)				
Point to be valued	Land-use types	Lat	Lon	LST	S-IDW	IDW	Kriging	Natural	Spline
1	Farmland 1	31.743	115.910	44.086	44.572	45.299	44.968	43.377	41.169
2	Farmland 2	31.688	115.974	45.801	44.902	45.305	45.549	45.355	45.859
3	Farmland 3	31.682	115.876	44.473	45.434	45.592	42.255	45.385	46.336
4	Farmland 4	31.671	115.924	45.277	44.080	43.174	43.515	42.791	45.183
5	Woodland 1	31.745	115.947	44.369	45.916	47.383	46.870	47.569	48.472
6	Woodland 2	31.741	115.792	43.869	41.352	41.862	41.053	42.568	43.502
7	Woodland 3	31.731	115.848	39.674	41.126	40.758	41.101	40.362	40.240
8	Woodland 4	31.685	115.722	39.962	39.635	39.509	39.795	39.722	36.678
9	Residential land 1	31.770	115.931	49.092	48.774	48.794	47.180	48.849	49.691
10	Residential land 2	31.724	115.942	48.106	47.027	47.066	46.580	47.482	48.167
11	Residential land 3	31.702	115.803	44.395	42.426	41.770	41.326	42.269	43.038
12	Transport land 1	31.669	115.951	46.802	45.416	46.645	44.797	46.878	47.756
13	Transport land 2	31.590	115.917	45.973	42.301	41.063	40.912	38.718	39.513
14	Waters and water conservancy facility land 1	31.625	115.860	37.502	37.878	38.794	39.143	39.568	45.812
15	Waters and water conservancy facility land 2	31.623	115.999	36.887	37.213	38.130	39.294	37.152	35.224

Table 10: Comparison of interpolation accuracy of 5 kinds of interpolation methods on Aug. 9, 2013.

Study area	Forecasting methods	Forecast time	Aug. 9, 2013		
		RMSE	MAE	MAPE (%)	RVAR
Study area-1	S-IDW	1.531	1.234	1.252	0.827
	IDW	1.955	1.537	0.781	0.870
	Kriging	2.290	1.977	1.818	0.605
	Natural	2.340	1.509	1.253	0.983
	Spline	3.233	2.177	0.057	1.450

Table 11: Correlation analysis of interpolation results of 5 kinds of interpolation methods on Aug. 9, 2013.

Study area-1		LST	S-IDW	IDW	Kriging	Natural	Spline
LST	Pearson correlation	1	0.914**	0.843**	0.794**	0.791**	0.669**
	Significance (two-tailed)		0.000	0.000	0.000	0.000	0.006
	N	15	15	15	15	15	15

**Significant correlation at 0.01 level (two-tailed).

Natural interpolation are significantly correlated with LST at 0.05 level (two-tailed), of which the correlation coefficient r between S-IDW interpolation results and LST is 0.620, with the strongest correlation.

The interpolation data results and accuracy of the 5 methods under high-temperature conditions in study area-2 can be seen in Tables 18–20. As shown in Table 18, the interpolation results of S-IDW for 5 of the 15 points to be valued are closer to the land surface temperature than those of the other 4 interpolation methods. The LST value of point 2 to be valued is 37.981°C and the interpolation result of S-IDW at point 2 is 35.398°C, which deviates from the LST value more than Spline interpolation but is better than Kriging, IDW, and Natural interpolation, similar to that of point 10. Generally, through the mathematical statistics analysis and Pearson correlation analysis of the 5 interpolation methods, it is found that the MAE, MAPE, and RMSE of S-IDW are closer to the best fitting values between measured and estimated values under ideal conditions than those of the other 4 interpolation methods. As far as RVAR is concerned, 0.870 of Kriging and 0.813 of Natural are closer to the best fitting values between measured values and estimated values under ideal conditions than 0.695 of S-IDW. In terms of Pearson correlation, the results of S-IDW, Kriging, IDW, Natural, and Spline interpolation are significantly correlated with LST at 0.01 level (two-tailed), of which the correlation coefficient r between S-IDW interpolation results and LST is 0.906, with the strongest correlation.

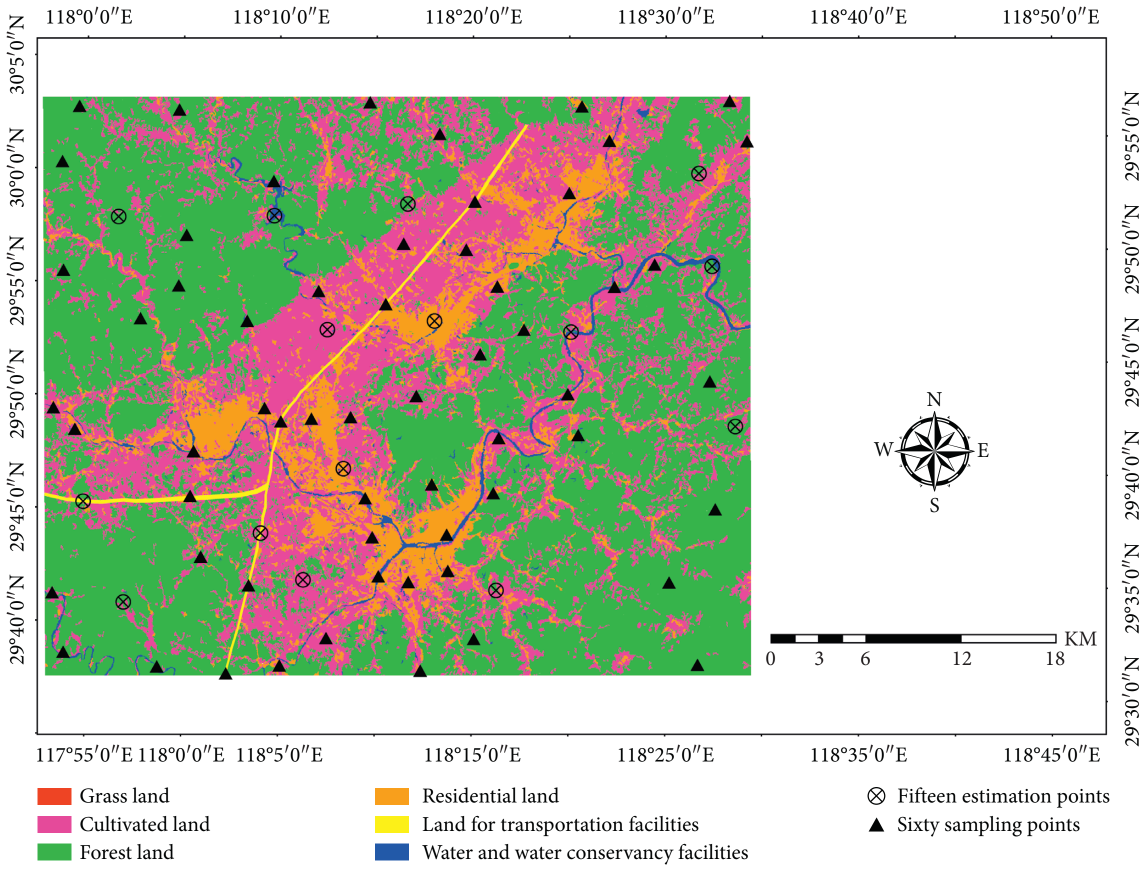

Figure 4: Distribution map of discrete points and pending valuation points in study area-2.

Table 12: Comparison of interpolation results of 5 kinds of interpolation methods on Feb. 11, 2017.

Study area-2		Forecast date: Feb. 11, 2017			Forecasting methods (unit of temperature °C)				
Point to be valued	Land-use types	Lat	Lon	LST	S-IDW	IDW	Kriging	Natural	Spline
1	Farmland 1	29.835	118.266	7.900	8.205	8.517	9.019	8.971	9.778
2	Farmland 2	29.697	118.228	6.902	8.867	8.106	8.125	8.254	8.812
3	Woodland 1	29.914	118.139	1.939	1.316	1.192	−0.869	−0.065	−1.176
4	Woodland 2	29.900	118.330	6.939	7.658	8.842	8.670	8.651	8.373
5	Woodland 3	29.894	118.524	7.147	7.798	8.637	8.982	9.025	10.682
6	Woodland 4	29.842	118.524	6.948	6.234	7.139	6.788	6.390	7.234
7	Woodland 5	29.750	118.525	3.188	2.662	2.873	1.054	2.014	0.001
8	Woodland 6	29.698	118.108	9.340	7.255	8.854	8.921	8.679	9.470
9	Woodland 7	29.677	118.353	5.960	5.508	7.421	5.451	4.136	2.635
10	Residential land 1	29.832	118.337	8.289	8.814	8.903	8.842	8.876	9.143
11	Residential land 2	29.756	118.264	7.201	9.052	8.589	8.172	8.283	8.133
12	Transport land 1	29.758	118.091	7.087	7.776	7.798	9.194	8.698	9.734
13	Transport land 2	29.727	118.204	8.000	8.102	8.697	8.538	8.509	9.173
14	Waters and water conservancy facility land 1	29.903	118.241	7.431	6.969	6.593	6.785	7.002	7.174
15	Waters and water conservancy facility land 2	29.816	118.425	7.803	7.933	8.714	8.676	8.636	8.850

Table 13: Comparison of interpolation accuracy of 5 kinds of interpolation methods on Feb. 11, 2017.

Study area	Forecasting methods	Forecast time		Feb. 11, 2017	
		RMSE	MAE	MAPE (%)	RVAR
Study area-2	S-IDW	1.00	0.786	2.033	1.389
	IDW	1.019	0.905	8.621	1.473
	Kriging	1.395	1.175	4.187	2.576
	Natural	1.267	1.152	3.904	2.205
	Spline	2.064	1.714	5.820	3.73

Table 14: Correlation analysis of interpolation results of 5 kinds of interpolation methods on Feb. 11, 2017.

Study area-2		LST	S-IDW	IDW	Kriging	Natural	Spline
LST	Pearson correlation	1	0.890**	0.934**	0.942**	0.925**	0.909**
	Significance (two-tailed)		0.000	0.000	0.000	0.000	0.000
	N	15	15	15	15	15	15

**Significant correlation at 0.01 level (two-tailed).

Table 15: Comparison of interpolation results of 5 kinds of interpolation methods on Apr. 19, 2018.

Study area-2		Forecast date: Apr. 19, 2018			Forecasting methods (unit of temperature: °C)				
Point to be valued	Land-use types	Lat	Lon	LST	S-IDW	IDW	Kriging	Natural	Spline
1	Farmland 1	29.835	118.266	28.319	29.329	30.205	28.295	31.402	33.163
2	Farmland 2	29.697	118.228	27.913	28.669	27.289	27.577	27.147	26.398
3	Woodland 1	29.914	118.139	26.573	25.333	25.081	25.916	24.636	24.906
4	Woodland 2	29.900	118.330	26.710	28.893	29.930	30.326	29.204	28.371
5	Woodland 3	29.894	118.524	27.722	28.717	29.123	28.665	29.350	32.286
6	Woodland 4	29.842	118.524	26.715	27.145	27.100	27.199	27.442	30.523
7	Woodland 5	29.750	118.525	26.039	25.032	24.704	26.027	24.679	25.179
8	Woodland 6	29.698	118.108	29.600	27.468	27.927	28.226	27.658	30.677
9	Woodland 7	29.677	118.353	26.205	27.298	29.026	28.417	26.651	18.758
10	Residential land 1	29.832	118.337	32.115	30.855	30.579	30.712	30.154	30.312
11	Residential land 2	29.756	118.264	25.245	30.204	29.724	29.506	29.511	28.246
12	Transport land 1	29.758	118.091	28.044	28.270	27.958	26.828	29.346	35.263
13	Transport land 2	29.727	118.204	28.170	28.289	28.016	28.063	28.075	30.576
14	Waters and water conservancy facility land 1	29.903	118.241	23.053	24.556	25.234	27.370	25.125	24.395
15	Waters and water conservancy facility land 2	29.816	118.425	26.727	26.616	27.856	28.520	25.823	24.521

Table 16: Comparison of interpolation accuracy of 5 kinds of interpolation methods on Apr. 19, 2018.

Study area	Forecasting methods	Forecast time	Apr. 19, 2018		
		RMSE	MAE	MAPE (%)	RVAR
Study area-2	S-IDW	1.719	1.268	1.839	0.805
	IDW	1.995	1.627	2.591	0.844
	Kriging	2.081	1.512	3.054	0.461
	Natural	1.965	1.666	1.723	1.051
	Spline	3.659	3.018	3.525	4.369

Table 17: Correlation analysis of interpolation results of 5 kinds of interpolation methods on Apr. 19, 2018.

Study area-2		LST	S-IDW	IDW	Kriging	Natural	Spline
LST	Pearson correlation	1	0.617*	0.514	0.382	0.541*	0.513
	Significance (two-tailed)		0.014	0.050	0.159	0.037	0.051
	N	15	15	15	15	15	15

*Significant correlation at 0.05 level (two-tailed).

Table 18: Comparison of interpolation results of 5 kinds of interpolation methods on Jul. 10, 2013.

Study area-2		Forecast date: Jul. 10, 2013			Forecasting methods (unit of temperature: °C)				
Point to be valued	Land-use types	Lat	Lon	LST	S-IDW	IDW	Kriging	Natural	Spline
1	Farmland 1	29.835	118.266	35.071	36.730	38.018	38.408	40.071	43.826
2	Farmland 2	29.697	118.228	37.981	35.398	34.928	34.230	34.662	35.745
3	Woodland 1	29.914	118.139	30.012	29.221	28.643	27.689	27.914	28.126
4	Woodland 2	29.900	118.330	34.162	35.963	37.472	36.833	36.132	34.809
5	Woodland 3	29.894	118.524	35.395	36.822	37.653	37.505	38.331	40.496
6	Woodland 4	29.842	118.524	36.442	35.110	35.729	35.586	35.827	39.280
7	Woodland 5	29.750	118.525	31.775	32.236	32.297	31.740	32.968	32.647

Table 18: Continued.

Study area-2		Forecast date: Jul. 10, 2013			Forecasting methods (unit of temperature: °C)				
Point to be valued	Land-use types	Lat	Lon	LST	S-IDW	IDW	Kriging	Natural	Spline
8	Woodland 6	29.698	118.108	33.573	34.817	34.728	35.205	35.340	32.472
9	Woodland 7	29.677	118.353	34.980	36.399	39.335	39.142	35.644	29.332
10	Residential land 1	29.832	118.337	41.655	39.583	39.386	39.255	39.308	39.936
11	Residential land 2	29.756	118.264	40.206	39.690	38.307	38.235	37.247	33.101
12	Transport land 1	29.758	118.091	36.348	35.363	34.768	36.077	36.815	40.110
13	Transport land 2	29.727	118.204	37.221	35.640	35.083	35.242	35.438	38.313
14	Waters and water conservancy facility land 1	29.903	118.241	28.349	29.977	30.234	30.705	30.449	30.414
15	Waters and water conservancy facility land 2	29.816	118.425	35.644	35.757	37.428	36.487	35.722	34.758

Table 19: Comparison of interpolation accuracy of 5 kinds of interpolation methods on Jul. 10, 2013.

Study area	Forecasting methods	Forecast time		Jul. 10, 2013	
		RMSE	MAE	MAPE (%)	RVAR
Study area-2	S-IDW	1.451	1.307	0.021	0.695
	IDW	2.303	2.083	0.982	0.835
	Kriging	2.345	2.047	0.667	0.870
	Natural	2.315	1.953	0.577	0.813
	Spline	3.894	3.048	0.861	1.753

Table 20: Correlation analysis of interpolation results of 5 kinds of interpolation methods on Jul. 10, 2013.

Study area-2		LST	S-IDW	IDW	Kriging	Natural	Spline
LST	Pearson correlation	1	0.906^{**}	0.755^{**}	0.746^{**}	0.747^{**}	0.541^{*}
	Significance (two-tailed)		0.000	0.001	0.001	0.001	0.037
	N	15	15	15	15	15	15

**Significant correlation at 0.01 level (two-tailed).

4. Conclusions

In this paper, S-IDW considering geographic semantics is proposed, which is a novel spatial interpolation algorithm of meteorological parameters. The geographical semantic similarity and weight between known observation points and estimated points are considered comprehensively, which makes the interpolation result of IDW more reasonable. We selected 2 research areas with abundant land-use types to analyze the interpolation under different temperature conditions and used 4 different statistical methods to evaluate the interpolation accuracy. At the same time, the interpolation results of 5 interpolation methods were analyzed and compared by Pearson correlation analysis. The experimental results show that the accuracy of S-IDW is generally higher than the inverse-distance weighting method, Kriging, natural neighbor interpolation, and spline function interpolation, but there are also some abnormal value and interpolation outliers. Comparing the interpolation results of five methods, it is found that the interpolation results of S-IDW are closer to the measured value of LST than those of four other interpolation methods. The MAE, MAPE, RVAR, and RMSE of S-IDW are closer to the best fitting value between the measured and estimated values under ideal conditions than those of the other 4 interpolation methods, and the correlation between the interpolation results of S-IDW and LST is also the strongest. Under the above experimental conditions, the interpolation results of S-IDW are more accurate and stable.

Note that we check the sample points involved in the calculation and find that the semantic interpolation is a little less effective than the traditional numerical interpolation when there are many surface types of the same kind. When there are more homogeneous interpolation points, there is similarity to numerical interpolation. Other interpolation methods have obvious advantages in numerical interpolation. For example, Kriging interpolation method has a wide range of applicability, which can better reflect a variety of terrain changes. Spline interpolation method is suitable for gradually changing surfaces, such as temperature, elevation, groundwater level height, or pollution level. IDW interpolation is suitable for the data with large density and uniform distribution. In our experiment, when the type of interpolation point is single, the advantage of semantic interpolation is not obvious, even less than numerical interpolation. Meanwhile, when there are more types, the semantic interpolation method is obviously better.

However, there are still defects in our study, which need to be improved in further researches. First, for the future development framework of semantic interpolation, we hope to consider the continuity of time to make up for some missing data and combine the time factor [9, 16] with semantic interpolation method to study spatiotemporal semantic interpolation. In addition, we also attempt to integrate the density, direction, elevation, and other influencing factors [6–10] of interpolation points into semantic interpolation and develop multifactor semantic interpolation methods. Moreover, in order to handle the complexity and uncertainty of predicting spatial attributes in most real-world problems, deep learning and artificial intelligence technology [16, 24] including logical and statistical learning algorithms can be considered as future extension of the work in the age of Big Data.

Data Availability

The S-IDW data used to support the findings of this study are available from the corresponding author upon request via email.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Junli Li and Ruijie Gan conceived and designed the new analytical approach. Wenjun Wu, Ruijie Gan, Junli Li, and Xiu Cao wrote the paper. Xinxin Ye, Jie Zhang, and Hongjiao Qu advised on the methods applied in the study. Ruijie Gan performed the experimental analyses. All authors read and approved the final manuscript.

Acknowledgments

This research was financially supported by the National Natural Science Foundation of China (Grant no. 41571400) and supported in part by the Open Research Fund Program of Anhui Province Key Lab of Farmland Ecological Conservation and Pollution Prevention.

References

[1] M. J. De Smith, M. F. Goodchild, and P. Longley, *Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools*, Troubador publishing ltd, England, UK, 2007.

[2] Z. Li, Y. Wu, and C. C. Zhejiang, "Inverse distance weighted interpolation involving position shading," *Acta Geodaetica Et Cartographica Sinica*, vol. 44, no. 1, 2015.

[3] G. Y. Lu and D. W. Wong, "An adaptive inverse-distance weighting spatial interpolation technique," *Computers & Geosciences*, vol. 34, no. 9, pp. 1044–1055, 2008.

[4] M. Knotters, D. Brus, and J. O. Voshaar, "A comparison of kriging, Co-kriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations," *Geoderma*, vol. 67, no. 3-4, pp. 227–246, 1995.

[5] J. Wang, *Research on Spline Interpolation Poisedness and Interpolation Approach Problems*, Dalian university of technology, Dalian, China, 2004.

[6] J. Liu, H. Chen, and W. Yang, "Comparison of interpolation methods on annual mean precipitation in hebei Province," *Acta Ecologica Sinica*, vol. 29, no. 7, pp. 3493–3500, 2009.

[7] H. Zhang, L. Lu, Y. Liu, and W. Liu, "Spatial sampling strategies for the effect of interpolation accuracy," *ISPRS International Journal of Geo-Information*, vol. 4, no. 4, pp. 2742–2768, 2015.

[8] L. D. Fang, W. H. Liao, M. Y. Wang, and W. Z. Song, "Research of the inverse distance weighted interpolation method considering the elevation," *Yellow River*, vol. 37, no. 9, pp. 38–41, 2015.

[9] W. Binbin, S. Lihong, L. U. Yueming, and L. T. University, "A space-time reverse distance weighted interpolation method considering elevation," *Geomatics & Spatial Information Technology*, vol. 41, no. 10, pp. 82–85, 2018.

[10] P. Duan, Y. Sheng, L. I. Jia, L. V. Haiyang, and S. Zhang, "Adaptive idw interpolation method and its application in the temperature field," *Geographical Research*, vol. 196, no. 196, pp. 31–38, 2014.

[11] C. Solomon and T. Breckon, *Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab*, John Wiley & Sons, Hoboken, NJ, USA, 2011.

[12] S. Bhattacharjee, P. Mitra, and S. K. Ghosh, "Spatial interpolation to predict missing attributes in gis using semantic kriging," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 8, pp. 4771–4780, 2014.

[13] S. Bhattacharjee and S. K. Ghosh, *Spatio-Temporal Change Modeling of Lulc: A Semantic Kriging Approach*, Wiley, Hoboken, NJ, USA, 2015.

[14] S. Bhattacharjee, M. Das, S. K. Ghosh, and S. Shekhar, "Prediction of meteorological parameters: an a-posteriori probabilistic semantic kriging approach," in *Proceedings of the Acm Sigspatial International Conference on Advances in Geographic Information Systems*, Chicago, IL, USA, November 2016.

[15] S. Bhattacharjee, *Prediction of Meteorological Parameters: A Semantic Kriging Approach*, Acm Sigspatial, Seattle, DA, USA, Phd Workshop, 2016.

[16] M. Das and S. K. Ghosh, "Sembnet: a semantic bayesian network for multivariate prediction of meteorological time series data," *Pattern Recognition Letters*, vol. 93, p. S0167865517300028, 2017.

[17] C. Stasch, S. Scheider, E. Pebesma, and W. Kuhn, "Meaningful spatial prediction and aggregation," *Environmental Modelling & Software*, vol. 51, no. 1, pp. 149–165, 2014.

[18] F. Villa, I. N. Athanasiadis, and A. E. Rizzoli, "Modelling with knowledge: a review of emerging semantic approaches to environmental modelling," *Environmental Modelling & Software*, vol. 24, no. 5, pp. 577–587, 2009.

[19] R. Gan, *A Study of Meteorological Factor Interpolation Considering Geographical Semantics*, Anhui Agricultural University, Hefei, China, 2019.

[20] J. F. Suo, Y. Liu, and S. B. Zou, "Weighted semantic similarity algorithm based on geographic ontology," *Journal of Lanzhou University (Natural Sciences)*, vol. 53, no. 1, pp. 19–27, 2017.

[21] C. L. Shope and G. R. Maharjan, "Modeling spatiotemporal precipitation: effects of density, interpolation, and land use distribution," *Advances in Meteorology*, vol. 2015, pp. 1–16, Article ID 174196, 2015.

[22] A. Mellit, A. M. Pavan, and M. Benghanem, "Least squares support vector machine for short-term prediction of meteorological time series," *Theoretical & Applied Climatology*, vol. 111, no. 1-2, pp. 297–307, 2013.

[23] R. S. Teegavarapu, *Floods in a Changing Climate: Extreme Precipitation*, Cambridge University Press, Cambridge, UK, 2012.

[24] D. Zhu, X. Cheng, F. Zhang, X. Yao, Y. Gao, and Y. Liu, "Spatial interpolation using conditional generative adversarial neural networks," *International Journal of Geographical Information Science*, vol. 34, no. 4, pp. 735–758, 2019.