Comparison of spatial interpolation methods based on ArcGIS

Zenglu Liu\textsuperscript{1,2} and Tianfeng Yan\textsuperscript{1,2*}

\textsuperscript{1} School of electronic and information engineering, Lanzhou Jiaotong University, Lanzhou, Gansu, 730070, China
\textsuperscript{2} Gansu Radio Monitoring and Positioning Industry Technology Center, Lanzhou, Gansu, 730070, China
\textsuperscript{*}hetc3@greatqcc.com

Abstract. Spatial interpolation algorithm is based on known data to predict the study area, its accuracy is of great significance to research. In different application scenarios, the interpolation method with higher accuracy should be selected. Based on ArcGIS, this paper conducts experimental analysis on inverse distance weighting method and Kriging method, so as to provide a reference for the selection of interpolation method in different application scenarios.

1. Introduction
Spatial interpolation analysis algorithm is a kind of algorithm applied to transform the measurement data of discrete points into continuous data surface. It can compare the distribution of continuous data surfaces with other spatial phenomena. It has a wide range of application scenarios in spatial information, especially in geographic information.\cite{1} Its main research object is the human living environment, and its main application scenarios include Remote Sensing (RS), Global Positioning System (GPS), Geographic Information System (GIS), and other technologies.

The spatial interpolation algorithm can predict and analyze the spatial characteristics of unknown areas according to the known data of the research area. It plays an important role in the research of geographic information. In practical applications, the spatial distribution model and spatial interpolation method are usually used to calculate the raster value in the spatial area, which greatly simplifies the calculation program and improves the efficiency of information prediction in the spatial area studied\cite{2}. Dirks summed up spatial interpolation methods as global interpolation and local interpolation\cite{3}. Vicente-Serrano subdivided the interpolation methods into global interpolation, local interpolation, geostatistical method, and mixed interpolation\cite{4}.

The mean absolute error (MAE), mean relative error (MRE) and root mean square error (RMSE) were used to evaluate the error in this paper. MAE reflects the margin of error of the estimate. MRE reflects the accuracy of the estimate against the observed value. RMSE reflects the sensitivity and extremum of the estimates\cite{5,6}. The expression is as follows.

\begin{align}
    MAE &= n^{-1} \sum_{i=1}^{n} |ABS(Z_{a,i} - Z_{e,i})| \tag{1} \\
    MRE &= n^{-1} \sum_{i=1}^{n} \frac{ABS(Z_{a,i} - Z_{e,i})}{Z_{e,i}} \tag{2} \\
    RMSE &= n^{-1} \sum_{i=1}^{n} \frac{ABS(Z_{a,i} - Z_{e,i})}{n} \tag{3}
\end{align}

$Z_{a,i}, Z_{e,i}$ are the actual observed values and interpolation predicted values respectively.
2. Inverse distance weighting algorithm (IDW)

IDW was first proposed by Shepard and has been continuously improved and developed since then. One of its most important assumptions is that the observation point has a local effect on the interpolation point. The influence of the value of any observational point on the interpolation point decreases with the increase of the distance. The advantage of IDW is that it is intuitive and efficient. This interpolation works best with evenly distributed points. IDW is sensitive to outliers. Furthermore, unevenly distributed data clusters result in introduced errors. The formula of IDW model can be expressed as follows:

\[ \hat{Z}_0 = \sum_{i=0}^{n} (Z_i, Q_i) \]  (4)

\(\hat{Z}_0\) is the estimate at the point \((x_0, y_0)\). \(Q_i\) is the weight coefficient of the interpolation point corresponding to the observational point. \(N\) represents the number of interpolation points.

As a global interpolation algorithm, inverse distance weighted interpolation involves all discrete observation points in calculating the value of each interpolation point. At the same time, it is also a kind of accurate interpolation, and the predicted observed values in the generated surfaces are completely consistent with the measured ones. It combines the advantages of the Natural Neighbor based on Tyson polygon and multiple regression gradient methods. It not only takes the distance factor into account, but assigns weights to discrete observation points adjacent to the interpolation points according to the distance. The weight of direction is also taken into account when anisotropy occurs. The distance weight function is inversely proportional to the power of the distance from the interpolation point to the observation point. With the continuous expansion of the distance between the observation point and the interpolation point, the weight presents a decreasing trend of power function.

IDW is simple and easy to operate, and there will be no unexplained meaningless results. A reasonable result can be obtained even if the data set of observation points fluctuates greatly. But IDW is particularly sensitive to weight functions. Subtle differences in the weight function can have a big impact and are easily influenced by the data set of the observation points.

3. Kriging interpolation algorithm

Kriging interpolation algorithm is also known as spatial self-covariance optimal interpolation method. It is an optimal interpolation method named after South African mining engineer Krige and is based on variogram theory and structural analysis. It is suitable for regionalization and exists spatial correlation. It is assumed that all random errors are spatially correlated and all random errors have second-order stability. Its expression is as follows:

\[ Z^*_k = \sum_{i=1}^{n} \lambda_i Z_i \]  (5)

Where \(Z^*_k\) is the estimated value of Kriging interpolation at point \((x_0, y_0)\). Here \(\lambda_0\) is the weight coefficient. The weight coefficient is a set of optimal coefficients, which meets the requirement of minimum variance unbiased estimation.

From this point of view, the quality of the interpolation point value depends entirely on the weight coefficient. The weight coefficients of all types of kriging interpolation methods must meet the conditions of optimality and unbiasedness. The weight of the Kriging interpolation method is determined by the semivariogram. According to statistically unbiased and optimal requirements, the semivariogram can be obtained by the principle of Lagrangian minimization. It can be calculated with the following formula:

\[ \gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \]  (6)

\(\gamma(h)\) is the semivariogram. \(h\) is the lag distance or step length. \(N(h)\) is the number of sample points whose distance is equal to \(h\). \(Z(x_i)\) and \(Z(x_i + h)\) are respectively regionalized variables. The measured semivariogram of \(Z(x)\) at positions \(x_i\) and \(x_i + h\) can be fitted with theoretical models, such as the Nugget model, linear model, spherical model, exponential model, Gaussian model, etc. The
fitted variogram model reflects the structure (such as directionality, spatial autocorrelation distance, etc.) and randomness (such as random error, etc.) of regionalized variables. At the same time, it also provides useful information for ordinary Kriging interpolation. Appropriate parameters and function models are conducive to improving the accuracy of the difference results [10].

Kriging interpolation is also in constant development. Chen Guang uses the NM simplex algorithm to improve and optimize the Kriging variogram model [20] to improve the performance of wireless sensor networks. Gu Junhua used the VIRE algorithm and Kriging interpolation to achieve accurate indoor positioning [21]. Deng Yuechuan uses the Kriging interpolation method to construct a multi-path error model from the perspective of spatial distribution. To realize the exploration of the spatial distribution characteristics of multi-path errors in the specified test area [22].

Kriging interpolation algorithm is applied to the random and structural characteristics of the sample data. It is also very important to select the model and parameters for Kriging interpolation [23].

4. Data analysis and processing
This experiment uses ArcGIS software as the analysis tool, and the version is ArcGIS 10.0. ArcGIS is a GIS desktop software that can provide geographic data display, mapping, management, analysis, creation, and editing. The ArcView part has hundreds of tools that can perform spatial analysis and geoprocessing tasks. Geoprocessing tasks include conventional GIS operations such as layer overlay, buffer analysis, and data conversion. The ArcCatalog part can be used to organize, manage and create GIS data. The ArcToolbox toolbox in ArcView contains the inverse distance weight interpolation and Kriging interpolation tools needed for the experiment. The experimental data in this paper are from Digital Elevation Model (DEM) data of Huang-Huai-Hai Region, an international scientific data service platform. First, the data is processed, as shown in the following figure:

![Figure 1. Raw data (left) and elevation point data (right)](image)

The first step is to process the data with the inverse distance weighted interpolation algorithm. As shown in Figure 2. The second step uses the Kriging interpolation algorithm to process the experimental data. As shown in Figure 3.
Figure 2. Inverse distance weighted interpolation algorithm stretches histogram.

Figure 3. Kriging interpolation algorithm stretches the histogram.

The comparison shows that the overall data of inverse distance weighted interpolation is more stable than the overall data of kriging interpolation. It is closer to the original data in the stretch area. However, the variation of kriging interpolation data is greater than that of inverse distance weighted interpolation.

The accuracy model can be evaluated more accurately through monitoring points. The statistical results are shown in the following table:

| Serial number | Heigh | IDW   | KRINGING   |
|---------------|-------|-------|------------|
| 1             | 58    | 54.227564 | 58.958745  |
| 2             | 50    | 85.876532 | 93.534265  |
| 3             | 100   | 83.852435 | 84.325456  |
| 4             | 89    | 104.684322 | 108.231154 |
| 5             | 123   | 95.353483 | 87.635785  |
| 6             | 102   | 97.30123  | 92.52887   |
| 7             | 60    | 52.245698 | 59.201432  |

The statistical results can be used to compare the accuracy of inverse distance weighted interpolation and Kriging interpolation, as shown in the following table:

|                | RMSE | R-MSE | L-RMSE | ME      | SD       | AR       |
|----------------|------|-------|--------|---------|----------|----------|
| **IDW**        | 13.895 | 0.18913 | 0.2210 | 3.1824  | 15.1086  | 1.0124   |
| **KRINGING**   | 20.013 | 0.2582 | 0.2501 | 3.3508  | 20.0545  | 1.0113   |

It can be seen from the above table that the error and average error of the Kriging interpolation method are far greater than the distance weighted method. As mentioned above, the result data of kriging interpolation fluctuates greatly. It can be seen that, for this set of experimental data, the effect of Kriging interpolation is a little better.

The error and average error of the Kriging are much greater than that of the IDW method. This also proves that the results calculated by the Kriging interpolation mentioned in the previous article fluctuate greatly. Therefore, it is better to use Kriging interpolation for this group of elevation data.
5. Conclusion
In the literature\(^{24}\), it is shown that IDW has higher interpolation results for high-value areas and lower interpolation results for low-value areas. Literature\(^{25}\) found that 8 regional rainstorms occurred during the interpolation calculation of daily precipitation data of 5006 automatic stations in Sichuan Province. The calculation accuracy of the two methods is not much different. The literature\(^{26}\) shows that the ordinary Kriging interpolation method may obtain higher interpolation accuracy under the condition of the non-normal distribution of the original data. The cross-validation results show that the ordinary kriging method is better than IDW and TS methods in interpolation accuracy and interpolation error distribution range. In literature\(^{27}\), it is shown that the OK method is suitable for periods and regions with large precipitation, while the IDW method is suitable for periods and regions with small precipitation, and comprehensive consideration can be taken in the selection. Therefore, in different application scenarios, the selection of interpolation methods should be combined with actual conditions, and different interpolation methods may also be adopted in different geographical environments in the same scene.

References
[1] Li Haitao, Shao Zedong. Review of spatial interpolation analysis algorithm [J]. Computer system application, 2019, 28 (07): 1-8.
[2] Guo Shidong, Liu Lan. Comparison of spatial interpolation methods based on ArcGIS [J]. Science and technology wind, 2019 (16): 227.
[3] Daly C. Guidelines for assessing the suitability of spatial climate data sets. International Journal of Climatology, 2006, 26(6): 707–721. [doi: 10.1002/(ISSN)1097-0088].
[4] DIRKS K N, HAY J E, STOW C D, et al. High-resolution studies of rainfall on Norfolk Island. Part II: Interpolation of rainfall data [J]. Hydrology, 1998, 208: 187-193.
[5] Vicente-Serrano S, M Saz-Sánchez, Cuadrat J. Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature [J]. Climate Research, 2003, 24(2): 161-180.
[6] Song L Q. Comparative analysis of spatial interpolation method and application of daily precipitation: A case study of Shenzhen City [J]. Geo-information Science, 2008, 10(5): 566-572.
[7] Gao Huaxi, Yin Kunlong. Correlation analysis of rainfall and landslide hazards and discussion on early warning and forecast thresholds [J]. Rock and Soil Mechanics, 2007(05): 1055-1060.
[8] Yu Xiaodong, Wu Ying, He Lamei. Improvement and comparison of inverse distance weighted mesh interpolation algorithm [J]. Chinese Journal of Engineering Geophysics, 2013, 10(06): 900-904.
[9] Lu GY, Wong DW. An adaptive inverse-distance weighting spatial interpolation technique. Computers & Geosciences, 2008, 34(9): 1044–1055.
[10] Zhang Jinming, Guo Liping, Zhang Xiaodan. Influence of interpolation parameters on DEM interpolation error in inverse distance weighted interpolation algorithm. Journal of Surveying and Mapping Science and Technology, 2012, 29(1): 51 – 56. [DOI: 10.3969/j.issn.1673-6338.2012.01.013]
[11] Gao Zhen, Ye Xueyi, Zhou Tianqi, et al. Underwater acoustic data visualization algorithm based on inverse distance weighted interpolation. Computer Engineering, 2015, 41 (9): 266 – 270, 275.
[12] Azpuru, Marco & Teixeira, Karina. (2010). A Comparison of Spatial Interpolation Methods for Estimation of Average Electromagnetic Field Magnitude. Progress In Electromagnetics Research M, 14. 135-145. 10.2528/PIERM10083103.
[13] Shepard, D., “A two-dimensional interpolation function for irregularly-spaced data,” ACM Annual Conference/Annual Meet-ing, 517–524, 1968.
[14] Wang Qi, Ji min, Cao pin Lian, sun Yong. Analysis and Research on spatial interpolation method of mining subsidence prediction [J]. Beijing surveying and mapping, 2014 (05): 1-5.
[15] Nalder IA, Wein RW., Spatial interpolation of climatic Normals: a test of a new method in the Canadian boreal forest. Agricultural and Forest Meteorology. http://dx.doi.org/10.1016/S0168-1923(98)00102-6
[16] Kidner DB. Higher-order interpolation of regular grid digital elevation models. International Journal of Remote Sensing, 2003, 24(14): 2981–2987. [doi: 10.1080/0143116031000086835]
[17] Marinoni O. Improving geological models using a combined ordinary-indicator Kriging approach. Engineering Geology, 2003, 69(1–2): 37–45.
[18] Shi Wenjiao, Yue Tianxiang, Shi Xiaoli, et al. Research progress on spatial interpolation methods and precision of soil continuous attributes. Journal of natural resources, 2012, 27 (1): 163 – 175
[19] Tian Lei, Dong Deming, Wei Qiang, et al. Comparison of three spatial interpolation methods in statistical processing of Pb monitoring data in road dust [J]. Journal of Jilin University (SCIENCE EDITION), 2011, 49 (5): 964 – 968
[20] Chen Guang. Study on Environmental Monitoring Method of WSNs Based on Improved Kriging Algorithm [Master's Thesis]. Nanjing: Nanjing University of Posts and Telecommunications, 2017
[21] Gu Junhua, Xu Peng, Dong Yao, et al. Research on Adaptive VIRE Indoor Location Algorithm Based on Kriging Interpolation. Computer Engineering and Applications, 2018, 54(12): 100 – 104. [doi: 10.3778/j.issn.1002-8331.1709-0242]
[22] Deng Yuechuan, Gu Shuangxi. Spatial distribution characteristics of multipath error in Kriging interpolation. Science of Surveying and Mapping, 2018, 43(4): 17 - 23.
[23] BAI Shen, Li Xiaomin. Spherical model of the optimal Parameters estimated [ M]. Geophysical and Geochemical Exploration Technology, 1998, 20(1):125-129
[24] Wang Rui, Deng Hai, Liang shaobiao, Dong Jinxiu, Wang Jiabin. Study on spatial interpolation method of soil Cd and Se contents in two different geomorphic and geological background areas of Chongqing [J]. Soil Bulletin, 2020,51 (06): 1332-1341
[25] Geng Wei, Zhang Liang, Yang Jin. Spatial interpolation method and area estimation of a regional rainstorm and precipitation distribution in Sichuan Province Based on GIS [J]. Plateau and mountain meteorological research, 2019,39 (03): 75-81
[26] Duan Zhipeng, Lin Jingying, Li Feng, Gan Huayang, Liang Kai, Wu Changhua, Zeng Xiangyun, Liu Xin, Huo Zhenhai. Selection of interpolation methods for spatial distribution characteristics of trace elements in offshore sediments based on GIS [J]. Marine environmental science, 2014,33 (02): 262-268
[27] Wu Lun, Wu Xiaojuan, Xiao Chenchao, Tian Yuan. Spatial and temporal distribution characteristics of errors in five commonly used precipitation interpolation methods: A case study of Shenzhen City [J]. Geography and Geo-Information Science,2010,26(03):19-24.