Evaluating Spatial Interpolation Maps of the Age Structure of the Population of Thi Qar Governorate Using Geographic Information Systems (GIS) Technologies

Nariman Jamal Kazem, Wissam Ahmed Rashid

University of Baghdad, College of Arts, Department of Geography and Geographic Information Systems, Iraq.
Email: neriman.Jamal2207p@coart.uobaghdad.edu.iq

Abstract

The research aims to harness spatial interpolation techniques to produce maps with a high level of perceptual accuracy in representing the population data of the study area. This is achieved after exploring the statistical and spatial nature of the databases used, analyzing them, and determining their distribution using a variety of spatial data exploration tools available within the GIS (Geographic Information Systems) environment. These tools contribute to evaluating the characteristics, distribution, and analysis of data, including testing data distribution, identifying its direction, and uncovering its spatial correlations. This highlights the importance of the study and understanding the distribution pattern of the population data in the study area, thereby facilitating the preparation of future that serve spatial organization by building spatial models for the numerical distribution of the data for age groups: young, middle-aged, and elderly, at the level of administrative units of the study area by relying on the inductive approach and the quantitative analysis approach. This is done using Geostatistical Analysis methods within the concept of spatial interpolation in GIS. The research also seeks to predict the population maps for Thi Qar Governorate by comparing spatial interpolation methods, namely Inverse Distance Weighting (IDW) and Simple Kriging (SK). The study concluded, after verifying the accuracy of the results using the Cross-Validation curve, that the Simple Kriging (SK) method is the most accurate model in spatial prediction of population data for representing the data of the young age group in the study area, as it had the lowest Root Mean Square Error (RMSE) of 10.012. It was followed by the Inverse Distance Weighting (IDW) method for representing the data of the elderly age group with an RMSE of 350.48. This indicates that the use of statistical criteria improves the accuracy of spatial interpolation and that the spatial prediction of the population data for the study area using the Simple Kriging (SK) method was more accurate than the Inverse Distance Weighting (IDW) method based on statistical indicators.

Keywords: Spatial Interpolation, Spatial Statistical Analysis, (IDW), (SK), Spatial Correlation, Population Maps, Population Data, Visual Analysis.

1. Introduction

Contemporary geography has evolved to go beyond the current state of geographical phenomena, focusing on predicting the future using representation and simulation to create an outlook. This is achieved through spatial statistical analysis within the environment of Geographic Information Systems (GIS), which is the optimal tool for handling geographic databases and performing spatial analysis of geographical phenomena. The study applied spatial interpolation techniques within the GIS environment to analyze the population data of the study area. Spatial statistical analysis techniques are widely used for spatial interpolation of data, relying on mathematical and statistical methods to fill in missing data based on the input data. These techniques help evaluate the accuracy of predictions when preparing population maps. They are used to predict values at unknown locations by referencing known values at other locations. These techniques involve collecting data from specific locations within the study area and then using this data to predict values in unmeasured areas through mathematical processes known as Geostatistical Analysis. This method includes creating a mathematical model that represents the relationship between known and unknown variables in the study area. Population data for the study area from 2010 to 2023 were collected, and a database was created in the ArcGIS 10.6 program. The data was explored to verify statistical properties and determine its suitability for statistical representation. Spatial statistical analysis was then conducted using geostatistical techniques in ArcGIS 10.6.

The study compared the Inverse Distance Weighting (IDW) method and the Simple Kriging (SK) method to predict the spatial statistics of the population data for the study area. The final model was selected based on the results of the cross-validation curve to test the validity of the chosen model. The Root Mean Square Error (RMSE) of all models was used to provide the best representation for creating maps of the study area's population, closely approximating reality using one of the interpolation techniques.

The research problem lies in whether spatial interpolation techniques can be employed within the GIS environment in preparing population maps, and whether they can replace the methods often used by researchers in representing thematic maps in general and population maps in particular, with a high level of visual perception.

Spatial interpolation technologies can simulate reality mathematically and statistically and can also contribute to creating maps that are an effective means of analyzing the variation of population data and providing results that are no less valuable than the results obtained using thematic population maps that are perceived and understood by the reader and recipient of the map.

2. Importance of research

The importance of the research lies in highlighting the importance of spatial interpolation techniques in predicting and analyzing the spatial distribution of population data in the study area. The research aims to compare spatial interpolation methods, the (IDW) method and the (SK) method, and their spatial accuracy in preparing spatial prediction maps for the population of the study area.

The spatial boundaries of the research include the southern part of Iraq, extending between latitudes (32°00-36°30) north and longitudes (47°12 - 45°36) east. As for its administrative boundaries, it is bordered to the north by Wasit Governorate, to the east by Maysan Governorate, to the west by Qadisiyah and Muthanna Governorates, and to the south by Basra and Muthanna Governorates. The city of Nasiriyah is the administrative center of the governorate, as its total area is (12900) km2 and represents (3.0%) of the total area of Iraq, which is ((435052. The governorate includes fifteen districts and twenty administrative units until the year 2023. See Map (1). As for the temporal boundaries of the research, it was based on population data according to population estimates for the year 2023.

Map (1) Location of the study area

Source: Based on: Ministry of Water Resources, General Authority for Survey, Map Production Department, Thi Qar Governorate, 1:250,000 scale map, 2011.

3. Results and discussion:

1- Data analysis: Age structure of the population of Thi Qar Governorate

First: The numerical and relative distribution of the young population of Thi Qar Governorate:

It is clear from Table (1) that the number of young people within the age group (0-14) years reached 1,028,293 people, thus constituting 43% of the total population of Thi Qar Governorate for the year 2023, which indicates an increase in fertility rates considering the economic and social developments witnessed by Thi Qar Governorate, which was reflected in women's reproductive behavior.

Table (1): Numerical and relative distribution of the young age group in the population of Thi Qar Governorate for the year 2023

Administrative units	young	%
Nasiriyah District	259723	25.3
Ur District	33815	3.3
Al-Rafai District	78694	7.7
Souq Al-Shuyukh District	60530	5.9
Akika District	24985	2.4
Karmat Bani Saeed District	30484	3.0
Al-Fadhiliyah District	28984	2.8
Al-Tar District	11743	1.1
Al-Jabaish District	25052	2.4
Al-Hammar District	5666	0.6
Al-Shatra District	119145	11.6
Al-Dawaya District	49355	4.8
Al-Islah District	23100	2.2
Sayed Dakhil District	33023	3.2
Qalaat Sukkar District	50491	4.9
Al-Fuhoud District	25690	2.5
Al-Gharraf District	62446	6.1
Al-Nasr District	48799	4.7
Al-Fajr District	32653	3.2
Al-Batha District	23918	2.3
Total	1028293	100

Source: Based on the appendix, based on the following equation:

Since: PN = population in the target year PO = population in the base year r = growth rate n = number of years between the base year and the target year.

For more, see: Abdul Rahim Bouadakji, Issam Khoury, Demography: Theories and Concepts, 1st ed., Dar Al-Rida Publishing, Damascus, 2002, p. 135.

At the administrative unit level, Al-Nasiriyah district ranked first in the number of young people, with a total of 259,723 individuals, making up 25.3% of the total young population in the governorate. This is attributed to the high fertility rate in the district. Al-Shatrah district came in second place at the young age group, with 119,145 individuals, accounting for 11.6% of the total young population in Thi Qar Governorate. The smallest young populations were recorded in the Al-Hammar and Al-Tar sub-districts, with 11,743 and 5,666 individuals respectively, making up

1.1% and 0.6% of the total, respectively. In the remaining administrative units, the percentage of the young age group ranged from 2.2% to 7.7%. This indicates a variation in the young population across the study area.

Second: The numerical and relative distribution of the middle-aged population of Thi Qar Governorate:

From observing Table (2), the number of middle-aged people in Thi Qar Governorate reached 1,285,884 people, constituting 54% of the total population of the governorate in 2023. The increase in this group and its entry into the labor market leads to improving the standard of living and quality of life for the population. It is noted that the number of middle-aged people varies at the level of administrative units, as Nasiriyah District recorded the first place in the number of middle-aged people with a percentage of 27.7%, and the high percentage of this group is due to the youth of society, which positively reflects on fertility and stability of society in terms of internal or external migration. As for Al-Shatrah District, it came in second place with a percentage of 12.1%. As for the administrative units that came in last place, they are (Al-Hammar District and Al-Tar District) with a percentage of (0.4%-0.9%) respectively. As for the rest of the administrative units, the percentage of middle-aged people ranged between (2.0%-7.8%).

Table (2): Numerical and relative distribution of the middle-aged population of Thi Qar Governorate for the year 2023

Administrative units	Middle-aged	%
Nasiriyah District	356129	27.7
Ur District	38121	3.0
Al-Rafai District	100188	7.8
Souq Al-Shuyukh District	84679	6.6
Akika District	29134	2.3
Karmat Bani Saeed District	37185	2.9
Al-Fadhiliyah District	33532	2.6
Al-Tar District	11123	0.9
Al-Jabaish District	25149	2.0
Al-Hammar District	5289	0.4
Al-Shatra District	155907	12.1
Al-Dawaya District	49291	3.8
Al-Islah District	28318	2.2
Sayed Dakhil District	34145	2.7
Qalaat Sukkar District	64524	5.0
Al-Fuhoud District	27238	2.1
Al-Gharraf District	73409	5.7
Al-Nasr District	65881	5.1
Al-Fajr District	36779	2.9
Al-Batha District	29863	2.3
Total	1285884	100

Source: Based on the appendix and the previous equation

Third: The numerical and relative distribution of the elderly category in the population of Thi Qar Governorate:

It is noted through the analysis of Table (3) that this category constituted 66,766 people, representing 2.8% of the total population of the governorate. This indicates a low percentage of the elderly in Thi Qar Governorate, and this shows the short life span and the need for health, living and social care that increases life span and the percentage of this category.

Table (3): Numerical and relative distribution of the elderly category in the population of Thi Qar Governorate for the year 2023

Administrative units	The elderly	%
Nasiriyah District	18759	28.1
Ur District	2034	3.0
Al-Rafai District	5021	7.5
Souq Al-Shuyukh District	4313	6.5
Akika District	1647	2.5
Karmat Bani Saeed District	2251	3.4
Al-Fadhiliyah District	1801	2.7
Al-Tar District	643	1.0
Al-Jabaish District	1199	1.8
Al-Hammar District	321	0.5
Al-Shatra District	8176	12.2
Al-Dawaya District	2360	3.5
Al-Islah District	1080	1.6
Sayed Dakhil District	1711	2.6
Qalaat Sukkar District	3618	5.4
Al-Fuhoud District	1212	1.8
Al-Gharraf District	3845	5.8
Al-Nasr District	2958	4.4
Al-Fajr District	2147	3.2
Al-Batha District	1672	2.5
Total	66766	100

Source: Based on the appendix and the previous equation

As for the administrative units, Nasiriyah District took the first place in terms of the number of elderly people, as their number reached 18,759 people, or 28.1%, while Shatrah District took the second place with several 8,176 people, or 12.2%. The reason for this is the availability of health and social services, which are limited to the centers of the districts and sub-districts compared to the rest of the administrative units. Meanwhile, Al-Hammar District and Al-Tar District came with the lowest number among the administrative units, as the number of this category reached 321, 643 people, respectively, and they constitute 0.5%, 1.0%, respectively, of the total elderly category in the governorate. This is due to the lack of health care and the lack of services provided to this category, and the continuation of some of this category working, and the low standard of living, all of which are factors that contribute to shortening life spans. As for the rest of the administrative units, they ranged between the lowest and highest numbers mentioned above.

2- Exploring data on the population of the study area:

Spatial data exploration aims to better understand the data and detect significant errors in it that may significantly affect the expected results. To represent the population data of the study area on maps within geostatistical analysis techniques, an analysis of the spatial structure of that data should be conducted. Geostatistical Analyst techniques include a variety of spatial data exploration tools that help in evaluating the characteristics, distribution, and analysis of the data (Al-Azzawi, 2019, p. 232), including:

1-Histogram data distribution test:

2-Trend Analysis:

3-Spatial autocorrelation:

1- Histogram data distribution test: It is a tool used to understand the nature of data distribution, allowing for testing the distribution of data related to a specific phenomenon by directly observing the distribution and analyzing statistical indicators. Spatial interpolation methods used to represent data provide the best results when the data is normally distributed and moderately so. Important characteristics of the distribution include central tendency, where the mean and median should be close to achieve a normal distribution (Johnston, 2001, p.21).

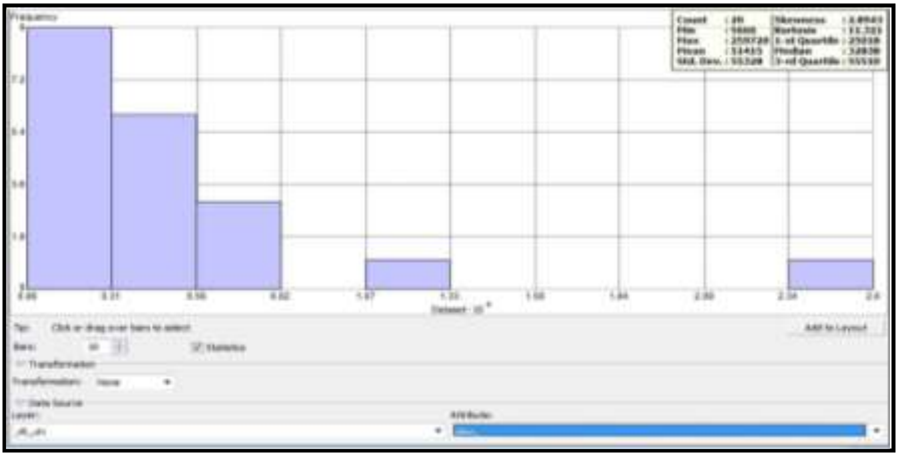
When the data has a unimodal normal distribution, the mean is approximately equal to the median, indicating a normal distribution shape, which results in a skewness value of zero. If there is skewness to the right, the mean becomes greater than the median, making the skewness coefficient positive. Conversely, if there is skewness to the left, the mean is less than the median, and the skewness coefficient becomes negative. It can be said that when the skewness value falls within the range of -1 to +1, the result is more accurate. Values outside this range indicate that the data distribution is skewed either to the right or left. This can be corrected using a logarithmic transformation to make the data more balanced. Therefore, it is essential to verify the normality of the data before representing it.

- Frequency histogram of data for the young age group:

It appears from the form of the frequency histogram (1) of the data of the young age group of the study area population and from the statistical summary that the data of the young age group are not normally distributed and that there are extreme values that affect the accuracy of the map representation. Therefore, the map representation process cannot be carried out using geostatistical analysis methods for this data unless the data is converted logarithmically.

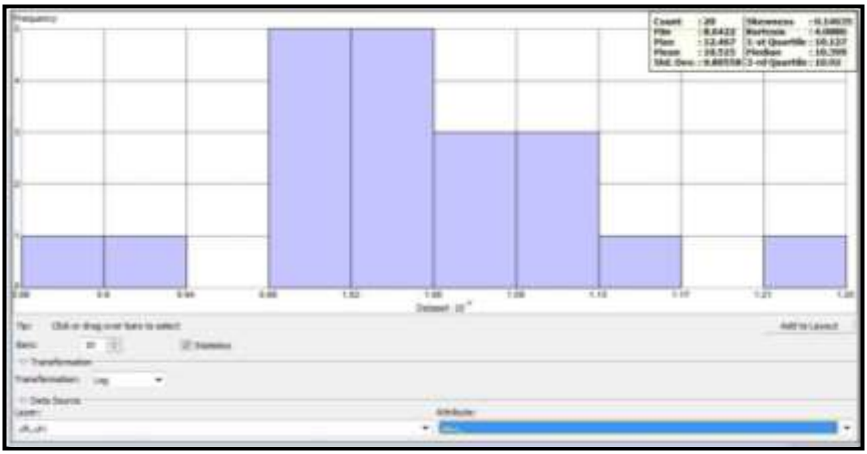
The arithmetic mean before the logarithmic conversion was (51415) and the median was (32838) and the value of the skewness coefficient was (2.8943), while after the logarithmic conversion process, there was a convergence between the values of the arithmetic mean, which was (10.515) and the median (10.399), while the value of the skewness coefficient was (0.14635), which is close to (1), which gives an impression of a deviation of the data distribution towards the right, indicative of the positive value of the skewness coefficient (0.14635) as in Figure (2), which indicates the possibility of representing the data of the Young people in the study area were identified using geostatistical analysis methods, consistent with the real distribution of data, and indicated by the low standard deviation, which reached a value of (0.80558) from the arithmetic mean.

Figure (1): Frequency histogram of data for the young age category before logarithmic transformation



Source: Based on GIS10.6 ARC

Figure (2): Frequency histogram of data for the young age group after logarithmic transformation

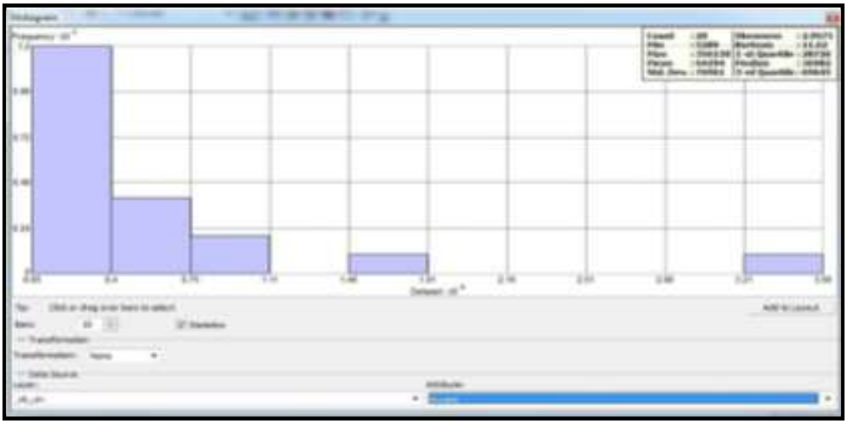


Source: Based on GIS10.6 ARC

-Frequent histogram of data for the middle-aged group:

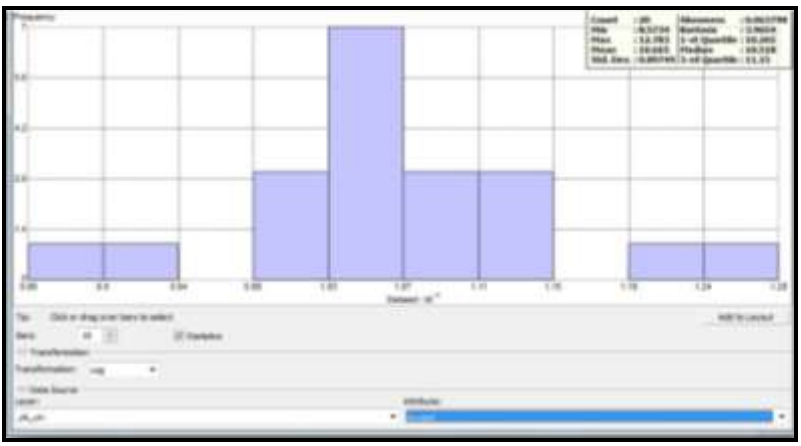
The histogram (Figure 3) and statistical summary indicate that the middle-aged group data for the study area's population does not follow a normal distribution, necessitating a logarithmic transformation of the data. There is a significant gap between the mean, which was 64,294, and the median, which was 36,982, while the skewness coefficient was 2.9571. After the logarithmic transformation, the mean became 10.665, and the median became 10.518, with the skewness coefficient reduced to 0.063798, which is close to 1. This positive skewness value (0.063798), shown in Figure 4, suggests a slight skew to the right. This transformation improves the accuracy of representing the middle-aged group data using geostatistical analysis methods. It also aligns with the actual distribution of the numerical data for the urban population across the administrative units in the study area, as indicated by the lower standard deviation value (0.89749) compared to the mean.

Figure (3): Frequency histogram of data for the middle-aged group before logarithmic transformation



Source: Based on GIS10.6 ARC

Figure (4): Frequency histogram of data for the middle-aged group after logarithmic transformation

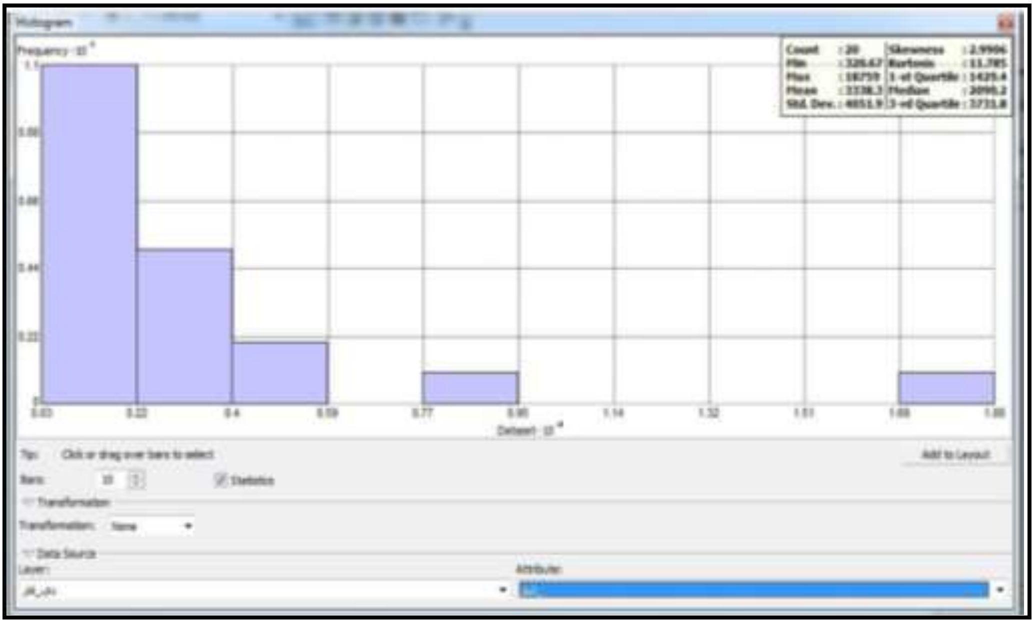


Source: Based on GIS 10.6 ARC

-Frequent histogram of data for the middle-aged group:

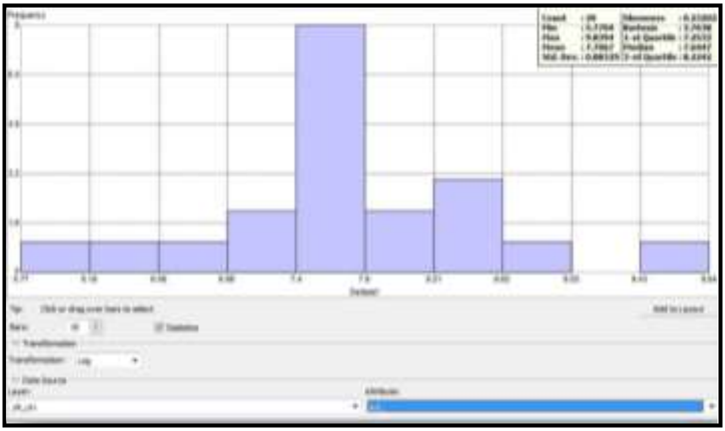
It is clear from the form of the frequency histogram (5) of the data of the elderly category for the study area population and from the statistical summary that the data of the elderly category are not normally distributed and that there are extreme values that affect the accuracy of the map representation. Therefore, the map representation process cannot be carried out using geostatistical analysis methods for this data unless the data is converted logarithmically. The arithmetic means before the logarithmic conversion was (3338.3) and the median (2090.2), and the value of the skewness coefficient was (2.9906), while after the logarithmic conversion process, there was a convergence between the values of the arithmetic mean, which was (7.7067) and the median (7.6447), while the value of the skewness coefficient was (0.23202), which is close to (1), which gives an impression of the deviation of the data distribution towards the right, indicative of the positive skewness coefficient value (0.23202), as in Figure (6), which indicates the possibility of representing the data of the elderly category for the study area population. By geostatistical analysis methods and in line with the true distribution of data and in terms of the standard deviation, whose value of (0.88329) decreased from the arithmetic mean.

Figure (5): Frequency histogram of the elderly category data before logarithmic transformation



Source: Based on GIS 10.6 ARC

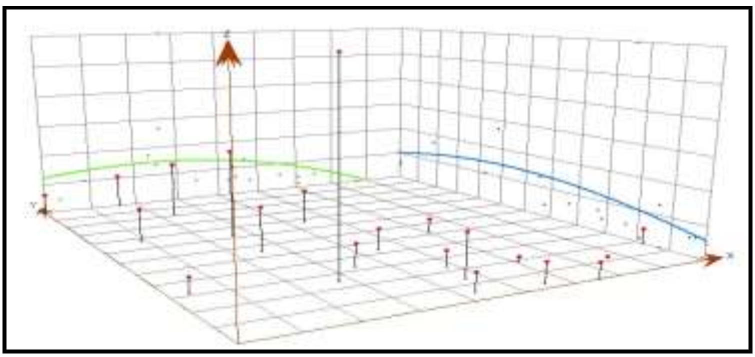
Figure (6): Frequency histogram of the elderly category data after logarithmic transformation



2- Trend Analysis:

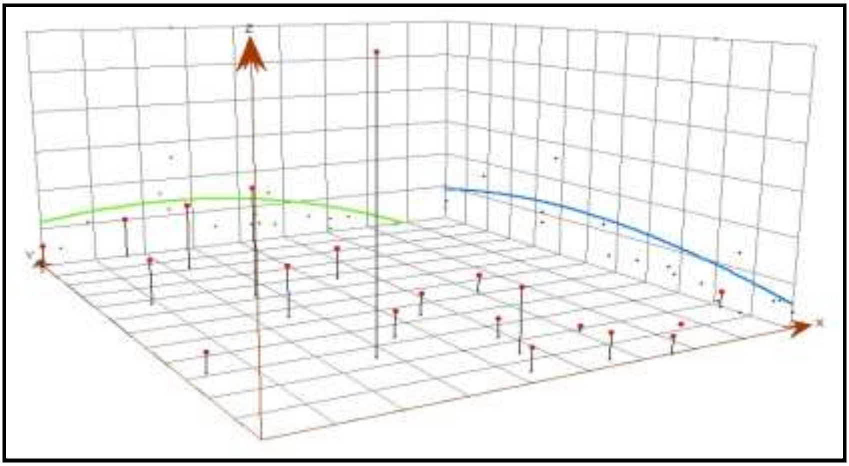
Understanding the trend of the phenomenon data is one of the basic steps before representing it on maps using geostatistical analysis techniques. If the data we are studying is not random, it is represented using specific mathematical formulas within these techniques. However, if the data is random, it is represented using statistical formulas because it provides an analysis of short-term variation, and the choice of one of these formulas depends mainly on the nature of the data in the study area (Johnston (2001). Op.cit. P.106). It is noted from Figure (7)(8)(9) which shows the trend of distribution of data for young and middle-aged people and data for the elderly that the blue line shows that there is a decrease from the northern parts and then gradually increases towards the southern parts of the study area, while the green line shows that there is a decrease in data from the eastern parts with a concentration of data in the central parts, which is consistent with the actual distribution of data within the database, which indicates the possibility of applying spatial interpolation techniques in representing the population data of the study area.

Figure (7): Distribution trends of age structure data in the study area



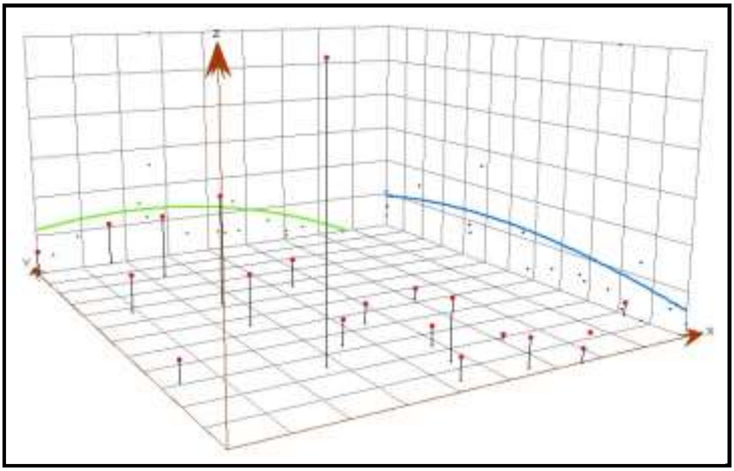
Source: Based on ARC GIS 10.6.

Figure (8): shows the trend of data for the middle-aged group.



Source: Based on ARC GIS 10.6.

Figure (9): shows the trend of the elderly category data.



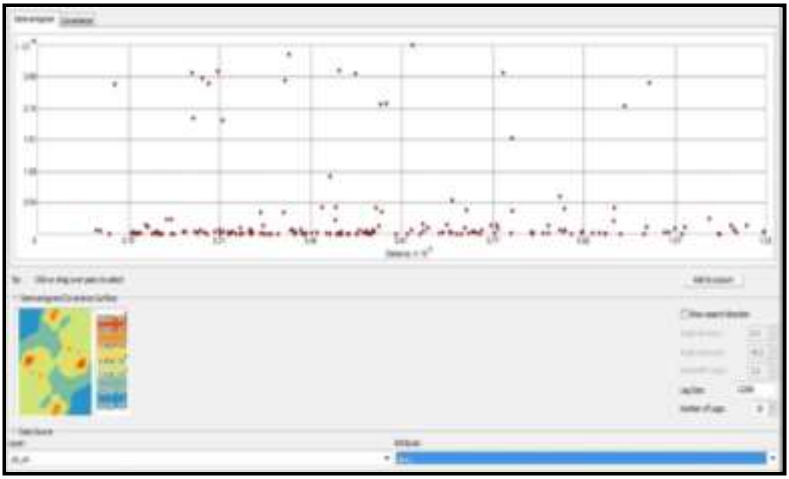
Source: Based on ARC GIS 10.6.

3-Semi variogram:

Spatial autocorrelation is a statistical tool used to explore the association between measurements of sample points in the same spatial domain. It contributes to making better decisions when choosing spatial prediction models. The semi-varogram tool tests the spatial autocorrelation between measurements of sample points and assumes that things that are close to each other are more similar. This tool helps test statistical relationships. (Al-Alam, 2020, p. 3) From Figures

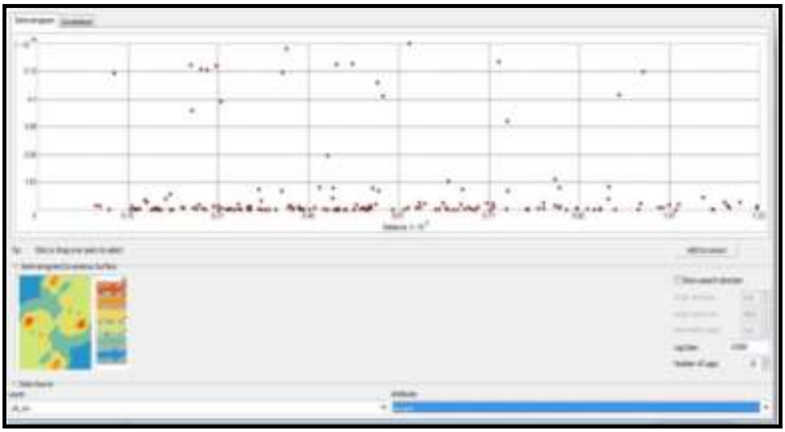
(12, 11, 10) respectively, the spatial autocorrelation of the data of the young, middle-aged, and elderly groups of the study area population is as follows: Most point pairs are close to each other on the (X) axis and move away as we move away from the (X) axis. We conclude from the above that most point pairs have a spatial relationship that they can use to represent the data.

Figure (10): Spatial autocorrelation of data for the young age group



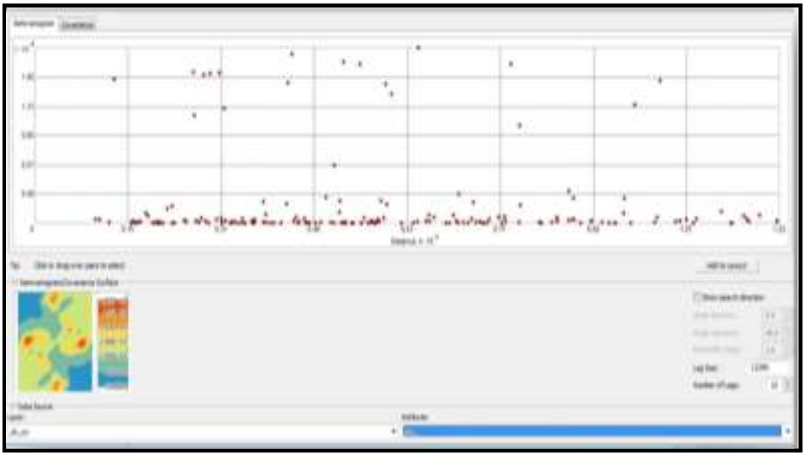
Source: Based on GIS10.6 ARC

Figure (5): Spatial autocorrelation of data for the middle-aged group



Source: Based on GIS10.6 ARC

Figure (11): Spatial autocorrelation of the elderly category data



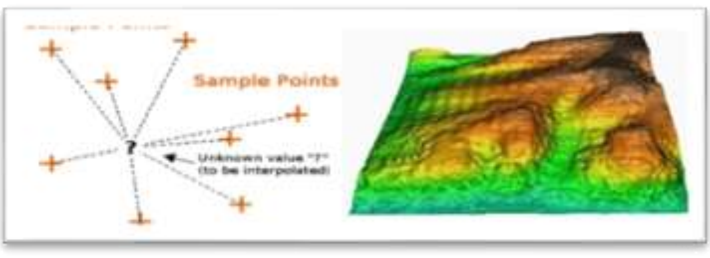
Source: Based on GIS 10.6 ARC

3- Spatial interpolation methods in Arc GIS 10.6 environment

First: Inverse Distance Weighting (IDW) method:

The inverse distance weight (IDW) method is used to create a surface that expresses a phenomenon by relying on the measured values at specific points on this surface. After that, a network of points is formed and the values of the phenomenon at these points are calculated using a special mathematical equation (Al-Azzawi, B. T., p. 485). This method relies on using measured data at specific points within the region to calculate the required information at unmeasured points, where the data at each known point has a stronger effect on the points close to it and its effect decreases the greater the distance between it and the unmeasured points. In this method, a specific weight is given to each known point for use in calculations (Al-Shaer, 2018, p. 28). One of the disadvantages of this method is that it gives the prediction either in the form of eyes or round lenses. Another disadvantage is that it is not suitable for prediction outside the sample limits or future prediction (Al-Sakani, 2020, p. 416).

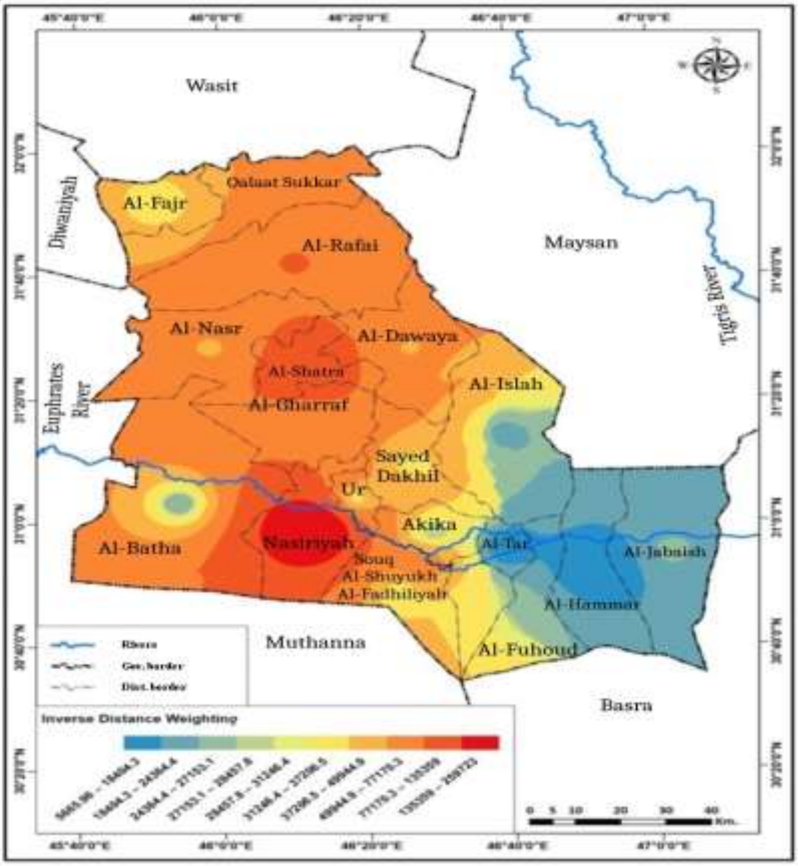
Image (1): Explains the mechanism of the inverse distance weight (IDW) method.



Source: Based on GIS 10.6 ARC

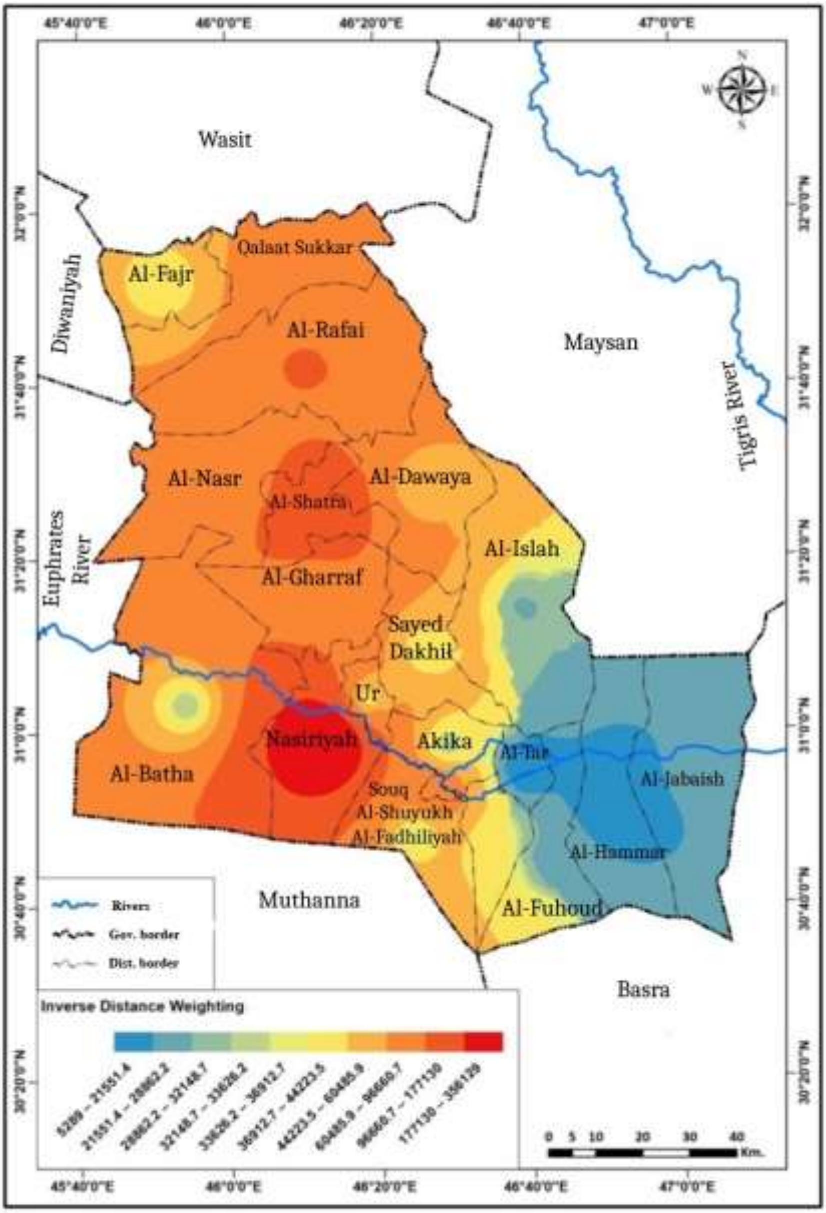
Through the visual interpretation of maps (2)(3)(4), it is clear from the dark red color variable that the southern parts represented by (Al-Nasiriyah District) recorded the highest concentration of young people, the red color ranges from dark to light in the northern parts represented by (Al-Shatra District) which recorded the second rank in terms of the number of young people, and the southeastern parts recorded the lowest concentration of young people according to the dark blue color variable which indicates the low concentration of young people in the following administrative units (Al-Hammar District and Al-Tar District). We conclude from the above that the spatial prediction of the data on the number of young people, middle-aged people, and the elderly using the inverse of the weight distance method is characterized by accuracy and perception when comparing the real values within the database and the level of administrative units with the values represented on the maps.

Map (2): Spatial interpolation of data for the young age group using the inverse distance weight method (IDW)



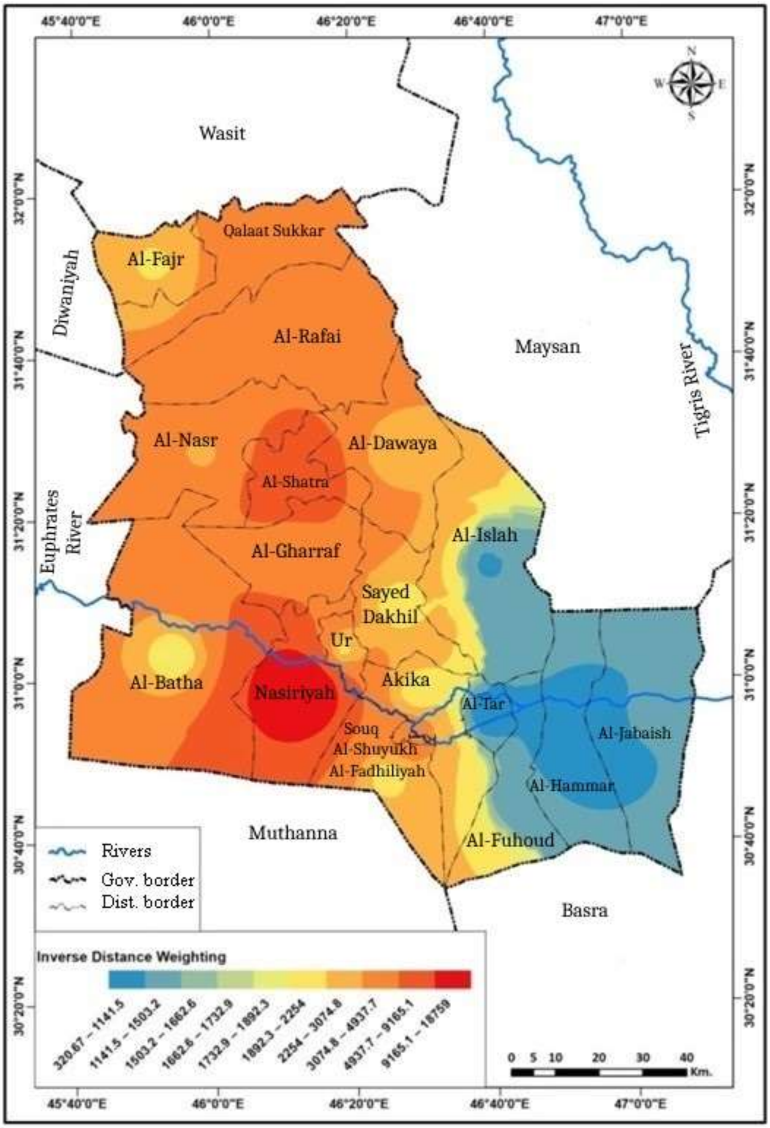
Source: Based on Table (1) in ARC GIS10.6

Map (3): Spatial interpolation of data for the middle-aged group using the inverse distance weighted (IDW) method



Source: Based on Table (2) in ARC GIS 10.6.

Map (4): Spatial interpolation of elderly data using the inverse distance weighted method (IDW).

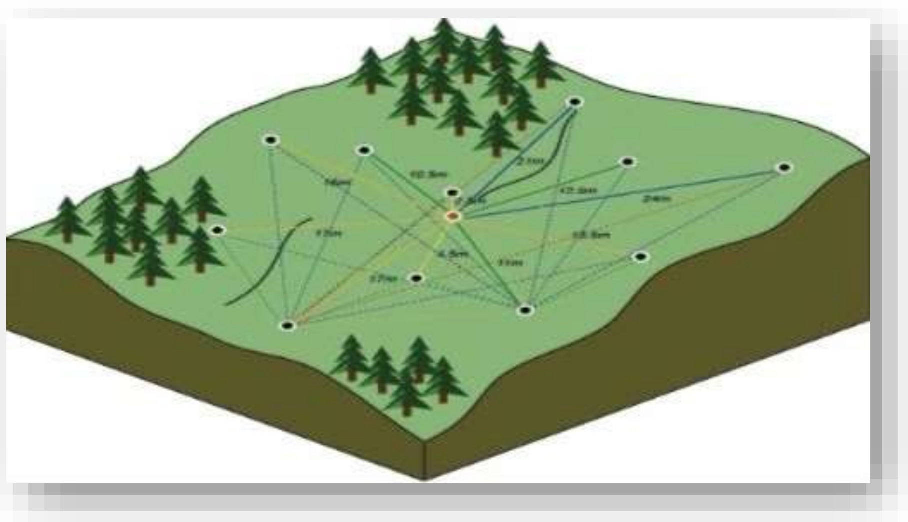


Source: Based on Table (3) in ARC GIS 10.6

Second: Simple Kriging (SK) method

Kringing is a geostatistical technique based on spatial prediction, as it uses statistical models that help produce a variety of maps, and the goal is to generate a surface containing the expected values (Njeban, p. 367). It is one of the most complex and powerful methods, as it uses advanced statistical methods and requires an understanding of spatial statistics due to the need to conduct an initial statistical examination of the data before implementation. Kriging is one of the statistical methods that deal with spatial inference and is used to measure the spatial correlation between phenomena to describe the difference and is calculated between all sites or between two sites within a specific area, then determining the inferred sites and the values they carry according to the correlation values between them. This method is compatible with maps that are related to spatial distribution such as the distribution of population data (Sharaf, 2011, p. 251).

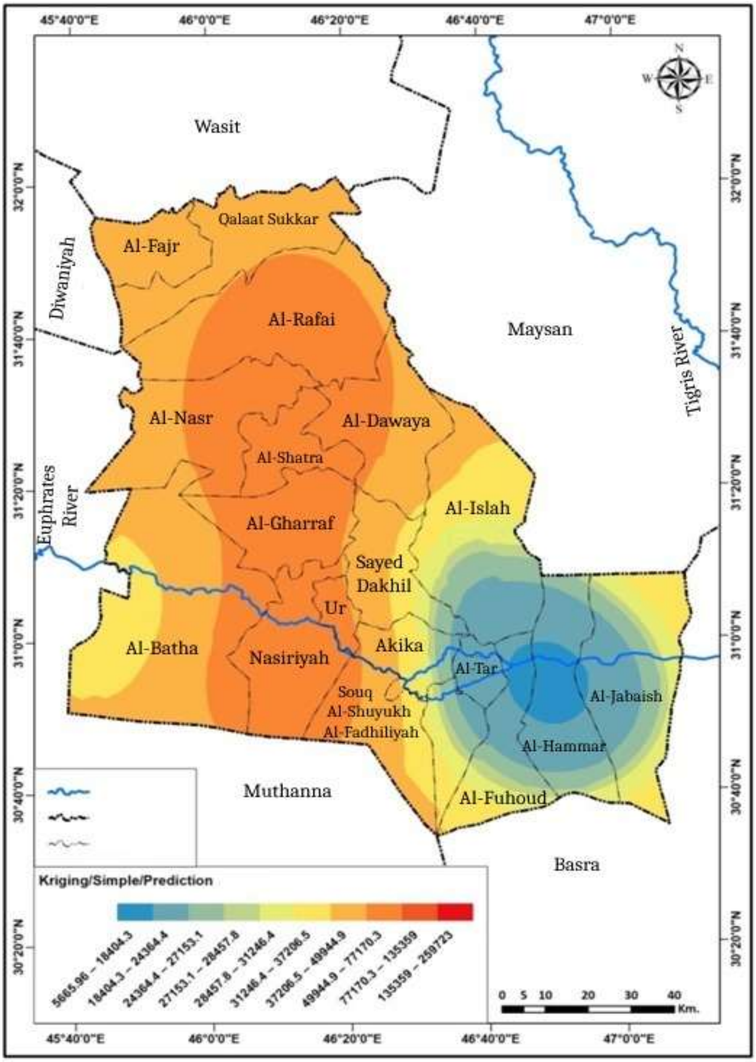
Image (2): shows how the simple kriking method (SK) works.



Source: Based on GIS10.6 ARC

It is clear from the visual perception of the map (5), that the southeastern parts of Thi Qar Governorate, represented by (Al-Hammar District), took the dark blue color variable, which represents the lowest concentration of the young, middle-aged, and elderly categories. However, the simple kriking method neglected the other color categories. It is clear from the above that the process of representing the numerical distribution data for the young, middle-aged, and elderly categories using the simple kriking method has weak visual perception and are incomprehensible maps, as the values represented on the maps and at the level of the administrative units of Thi Qar Governorate did not match when compared with the real values in the database.

Map (5): Spatial completion of data for the young age group using the simple kriging (SK) method.



Source: Based on Table (1) in ARC GIS 10.6

4. Statistical accuracy scale for comparing spatial interpolation methods:

Several statistical measures can be used to select the most accurate geostatistical analyst methods and evaluate the performance of the model, as the model that gives accurate estimates is as follows:

1- The mean error should be close to zero and is used to determine the model's validity.

2- Root-mean-square error and Average standard error should be as small as possible; these measures are important when comparing models.

3- Root-mean-square standard error should be close to one.

Comparison of spatial interpolation methods and accuracy criteria of results:

To compare the methods of geostatistical analysis, Arc GIS 10.6 offers a set of spatial statistical tools, which include the Cross-Validation test to assess the accuracy of geostatistical analyses for the population data of the study area. This produces a cartographic generalization by comparing the actual data with expected values, leading to some degree of statistical error. The degree of error indicates the deviation between the real situation and the models used to represent the phenomenon. The geostatistical methods were tested using built-in statistical formulas in the software. When comparing geostatistical methods, the represented maps should give precise estimates. For the models to be accurate, the Root Mean Square Error (RMSE) and the Average Standard Error should be as small as possible, while the Root Mean Square Standard Error should be close to 1. Table (4) shows a significant variation in the expected average error between the models, indicating differences in their accuracy. Although the RMSE values rise and fall depending on the data in the database, the RMSE of all models was used to determine the most suitable method for creating maps of the study area's population. The table reveals a clear disparity between RMSE values for the different geostatistical methods, proving that the methods vary in accuracy. While the RMSE value depends on the nature of the data in the database, a smaller RMSE indicates higher prediction accuracy. Based on the lowest RMSE values from the statistical comparison between the IDW and SK methods and the validation results using Cross-Validation, it was determined that the SK method is the best for representing the spatial prediction map for the population of the study area's younger age group. The SK method had the smallest RMSE value of 10.012. In contrast, the IDW method had an RMSE of 350.48 for representing the older age group, as shown in Table (4), which provides the statistical values for the spatial interpolation methods.

Table (4): Cross-validation accuracy calculations for the models used Young people category data

Method	IDW	SK
Mean Average Error (MEAN)	531.923	0.044
Root Mean Square Error (RMSE)	1050.18	10.012
Average Standard Error	-	10.261
Mean Standardized	-	0.091
Root Mean Square Error (RMSSE)	-	0.971

Middle-age group data

Method	IDW	SK
Mean Error (MEAN)	639.247	200.83
Root Mean Square Error (RMSE)	2767.85	750.18
Average Standard Error	-	636.33
Mean Standardized	-	-0.061
Root Mean Square Error (RMSSE)	-	1.371

Seniors' category data

Method	IDW	SK
Mean Error (MEAN)	204.991	137.06
Root Mean Square Error (RMSE)	350.48	394.93
Average Standard Error	-	319.19
Mean Standardized	-	0.067
Root Mean Square Error (RMSSE)	-	1.42

Source: Based on GIS10.6 ARC outputs

4. Conclusions:

1- The research showed that the data is distributed abnormally, so the data was converted logarithmically to be dealt with in the mapping modeling process.

2- The efficiency of spatial interpolation techniques in producing high-resolution population maps according to accurate statistical standards.

3- The research showed, through analyzing the data direction, the existence of two main directions, the first is south-north and the other is west-east for the distribution of the population data of the study area.

4- The best method for producing spatial prediction maps for the study area population according to statistical standards is the simple kriking method (SK) compared to other methods after using the Cross-Validation test and obtaining the lowest estimate of the square root of the average standard error.

Table (5): Estimates of the population of Thi Qar Governorate according to five-year age groups, and gender environment for the year 2023

	Total			Rural			Urban		
Total	Females	Males	Total	Females	Males	Total	Females	Males	Age group
350,091	169,987	180,104	133,991	66,370	67,621	216,100	103,617	112,483	4-0
363,072	178,758	184,314	142,771	69,951	72,820	220,301	108,807	111,494	9-5
315,130	157,459	157,671	118,036	58,147	59,889	197,094	99,312	97,782	14-10
265,581	133,083	132,498	92,889	45,712	47,177	172,692	87,371	85,321	19-15
208,579	98,737	109,842	69,051	32,494	36,557	139,528	66,243	73,285	24-20
167,390	80,416	86,974	55,913	27,387	28,526	111,477	53,029	58,448	29-25
157,595	80,092	77,503	56,545	29,081	27,464	101,050	51,011	50,039	34-30
130,899	67,963	62,936	43,974	22,424	21,550	86,925	45,539	41,386	39-35
120,969	61,531	59,438	40,947	20,602	20,345	80,022	40,929	39,093	44-40
85,537	42,279	43,258	27,353	13,317	14,036	58,184	28,962	29,222	49-45
49,049	27,292	21,757	15,908	9,201	6,707	33,141	18,091	15,050	54-50
52,460	27,432	25,028	16,397	8,565	7,832	36,063	18,867	17,196	59-55
47,825	24,593	23,232	15,712	7,973	7,739	32,113	16,620	15,493	64-60
26,660	13,069	13,591	8,134	4,285	3,849	18,526	8,784	9,742	69-65
16,177	8,922	7,255	5,407	2,778	2,629	10,770	6,144	4,626	74-70
7,962	4,491	3,471	3,062	1,737	1,325	4,900	2,754	2,146	79-75
15,967	9,772	6,195	6,627	3,761	2,866	9,340	6,011	3,329	80+
2,380,943	1,185,876	1,195,067	852,717	423,785	428,932	1,528,226	762,091	766,135	Total

Source: Republic of Iraq, Ministry of Planning, Central Statistical Organization, Directorate of Population and Labor Force Statistics, Estimates of the population of Dhi Qar Governorate by five-year age groups, environment, and gender for the year 2023.

5. Recommendations:

1- The study recommends the necessity of employing spatial statistical analysis techniques and applying them in producing spatial interpolation maps to know the accuracy of the used map models due to their great importance, especially in modern scientific research that helps in analyzing, interpreting and processing data to reach acceptable results characterized by accuracy, clarity and sound decisions in decision-making.

2- Establishing a spatial and descriptive database for the population data of the study area to ensure the production and design of the best and most optimal digital maps.

3- The study recommends the necessity of working on studies like spatial statistical analysis techniques instead of traditional methods, especially in human studies, due to the high accuracy of these methods in representing data in a way close to the real distribution of population data at the level of administrative units of the study area.

4- The necessity of working on exploring data knowing the nature of its distribution and the direction of data distribution and knowing its spatial connections before representing it using the statistical indicators it provides to know its validity and accuracy in representing the real distribution of data.

WORKS CITED

Al-Sakani, A. Y. A. (2020). Using statistical analysis (Kriging method) to measure groundwater pollution in Baghdad and designing its maps. Journal of Arts, (134).

Al-Sha'er, A. A. H. (2018). Biochemical study of groundwater in Rafah Governorate: A hydro-environmental study using GIS (Unpublished master's thesis). Faculty of Arts and Humanities, Islamic University, Gaza.

Sharaf, M. I. M. (2011). Spatial analysis using geographic information systems. Alexandria, Egypt: Dar Al-Maaref Al-Jami'yah.

Al-Alam, M. M., Younes, D. Z., & Al-Qamati, A. A. (2020). Applying the inverse distance weighting (IDW) method in mapping some chemical properties of soil in Ain Hazam, Bata village, Takanis. Misurata University Journal of Agricultural Sciences, 2(1).

Al-Azzawi, A. A. A. (2019). Comparison of spatial interpolation for groundwater level maps in Tal Afar district using GIS. Kirkuk University Journal for Humanities Studies, 14(1).

Al-Azzawi, A. A. A. (n.d.). Modeling Iraq's rainfall maps using geostatistical techniques in GIS. College of Education Journal, Wasit University.

Johnston, Kevin, Vertloef, Jay M., KrivoruCH KO, Konstantin and Lucas Neil (2001). Using Arc GIS GeostatistiCal Analyst.Environmental systems Research Institute, Redlands, CA

Hassan Swadi Njeban, Comparison and Evaluation of GIS-Based Spatial Interpolation Methods for Estimation Groundwater Level in Al-Salman District –southwest Iraq, Journal of Geographic Information System.