Yanbu Journal of Engineering and Science

Volume 23 | Issue 1 Article 4

2026

Evaluating the Interpolation Techniques for Completing the Soil Data and Mapping Jazan City, Saudi Arabia, using GIS

Ehab Sabi
Assistant Professor, Civil and Architectural Engineering Department, College of Engineering and Computer Science, Jazan University, Jazan, 45142, Saudi Arabia, esabi@jazanu.edu.sa

Mahmoud Abdelrahim Abdelgiom
Assistant Professor, Civil and Architectural Engineering Department, College of Engineering and Computer Science, Jazan University, Jazan, 45142, Saudi Arabia, mabdelmahmoud@jazanu.edu.sa

Follow this and additional works at: https://yjes.researchcommons.org/yjes

The journal in which this article appears is hosted on Digital Commons, an Elsevier platform.

Recommended Citation

Sabi, Ehab and Abdelgiom, Mahmoud Abdelrahim (2026), Evaluating the Interpolation Techniques for Completing the Soil Data and Mapping Jazan City, Saudi Arabia, using GIS, *Yanbu Journal of Engineering and Science*: Vol. 23: Iss. 1, 62-84.
DOI: https://doi.org/10.53370/3122-3303.1210
Available at: https://yjes.researchcommons.org/yjes/vol23/iss1/4

This Research Article is brought to you for free and open access by Yanbu Journal of Engineering and Science. It has been accepted for inclusion in Yanbu Journal of Engineering and Science by an authorized editor of Yanbu Journal of Engineering and Science.

YANBU JOURNAL OF ENGINEERING AND SCIENCE 2026;23:62–84

Scan the QR to view the full-text article on the journal website

RESEARCH ARTICLE

Evaluating the Interpolation Techniques for Completing the Soil Data and Mapping Jazan City, Saudi Arabia, using GIS

Ehab Sabi, Mahmoud Abdelrahim Abdelgiom *

Assistant Professor, Civil and Architectural Engineering Department, College of Engineering and Computer Science, Jazan University, Jazan, 45142, Saudi Arabia

ABSTRACT

The goal of the study is to determine which of the chosen interpolation techniques is optimal for forecasting soil characteristics in unmeasured areas for infrastructure and building projects. The study also incorporates geotechnical data from 23 regional boreholes with the UNESCO soil classification system developed by the Food and Agriculture Organization (FAO). The mapping and accuracy of different interpolation techniques were assessed using ArcGIS 10.4. Additionally, Spline, Inverse Distance Weighting (IDW), Kriging, Natural Neighbor, and Radial Basis Functions are used to assess the precision and efficacy of several interpolation techniques. The investigation's conclusions showed that the "natural neighbor" strategy was the most reliable way to gauge Jazan's soil quality. With 98% accuracy, this approach yielded the lowest RMSE value error rate among the others. Compared to the other methods, this one yielded the lowest RMSE value error rate, and its 98% accuracy rate showed that the R-square validation closely matched the real data. In regions outside the Red Sea, the groundwater table is likewise much lower, at about 0.5 meters below the surface, according to the findings of soil testing. According to the study, sand, clay, and silt make up the majority of the soil in Jazan.

Keywords: Geographic information systems (GIS), Borehole, Soil, Interpolation methods, Soil tests, ArcGIS10.4

1. Introduction

1.1. Background of interpolation techniques

Finding the unknown values that fall between the known data points is known as interpolation. Its primary function is to forecast unknown values for any geographically connected data points, like elevation, rainfall, noise level, and so forth. Numerous interpolation and approximation techniques have been created to address particular scattered data modeling issues. Numerous studies [1–3] and textbooks, e.g., [4], have addressed these techniques. Different geodetic interpolation problems are solved using radial basis function-based techniques, including kriging and linear prediction interpolation, in addition to the well-known polynomial and spline interpolation. The five interpolation techniques used in the study are "Spline, Inverse Distance Weighting (IDW), Kriging, Natural Neighbor, and Radial Basis Functions". Moreover, two other methods, "spline interpolation and thin plate spline interpolation," represent a wide use in geology and geophysics. Using a mathematical technique called spline interpolation, new data points can be created within the range of a discrete set of known data points. A smooth curve traveling across these locations is produced using piecewise polynomial functions or splines. The linear, cubic, and B-splines are common forms of spline interpolation. There are several uses for this technique, including :(a) Digital Terrain Models: To comprehend topography and aquatic characteristics,

Received 9 October 2025; revised 6 January 2026; accepted 12 February 2026.
Available online 3 March 2026

* Corresponding author.
E-mail addresses: esabi@jazanu.edu.sa (E. Sabi), mabdelmahmoud@jazanu.edu.sa, mahmoudabdelgiom@gmail.com (M. A. Abdelgiom).

https://doi.org/10.53370/3122-3303.1210
3122-3303/© 2026 The Author(s). Yanbu Journal of Engineering and Science. This is an open access article under the CC BY 4.0 Licence (https://creativecommons.org/licenses/by/4.0/).
Published by Yanbu Industrial College, Saudi Arabia. The journal is hosted on Digital Commons, an Elsevier Platform.

detailed digital terrain models are created by interpolating elevation or bathymetric data.

(b) Potential Field Data Mapping: Visualizing and analyzing subsurface geological features by interpolating magnetic and gravity data. (c) Borehole Data Analysis: By interpolating data from boreholes, continuous 3D models of subsurface features are created, aiding in environmental research and resource development [4]. One type of spline interpolation that works especially well for managing unevenly spaced data points in two or three dimensions is Thin Plate Spline (TPS) interpolation. It produces a smooth and realistic-looking representation of the data by reducing the bending energy of the interpolating surface. Thin plate spline interpolation has several uses in geophysics and geology, such as: (i) High-Resolution Digital Elevation Models: These models are essential for various geological investigations and are created by interpolating topographic or bathymetric data. (ii)Subsurface Structure Modeling: This technique helps to comprehend the geological formations beneath the Earth's surface by interpolating possible field data using TPS. (iii)Geochemical and Petro physical Data Modeling: This process is essential for resource management and environmental evaluations and involves interpolating geochemical or petro physical data to create continuous 3D models of subsurface features.(iv)Geological Structure Fitting: This technique improves our understanding of geological processes and formations by fitting mathematical surfaces to geological features like fault planes or folds [5].

The following describes, with citations, the relationships between the particular properties of soil data and the chosen interpolation techniques (Inverse Distance Weighting (IDW), Kriging, Natural Neighbor, and Radial Basis Functions):

a. The fundamental tenet of IDW is that close data points have the greatest impact on values at unsampled places. This is consistent with the fact that many soil qualities are localized, with substantial correlations between conditions in one location and those nearby. The inverse distance power directly controls the resulting surface's smoothness, enabling some tweaking to match the anticipated rate of change in soil properties.
b. Kriging, which has its roots in geostatistics, models the spatial autocorrelation in the data using variograms. Because of this, it can be applied to soil parameters that show structured spatial dependence, in which values at one place are statistically related to values at nearby locations. While Universal Kriging (UK) can account for trends, Ordinary Kriging's (OK) stationarity assumption suggests a rather consistent mean throughout the region.
c. iii. The Natural Neighbor approach uses Voronoi diagrams to identify a point's "natural" neighbors and weights them according to proportionate areas. Its ability to adapt to sporadic sampling and maintain local detail is what connects it to soil data. It is suitable when soil qualities change suddenly or unevenly because it does not impose a worldwide trend.
d. A surface that precisely passes through the observed data points is produced by RBFs. When the objective is to produce a surface that respects the high accuracy of the input soil data, this "exact interpolation" is helpful. The smoothness and general form of the interpolated surface are influenced by the radial basis function selection (e.g., Gaussian, multiquadric).

In this study, the cross-validation method was used with GIS and mathematical techniques. The purpose of this study is to determine the spatial variability of various soil parameters at the study area and to assess the accuracy of these interpolation for determining the values related to these parameters because of the insufficient sample points available around the study area. This is because the variability of the results depends on the accuracy of soil laboratory results, density and distribution of sample points, spatial correlation of the sampled data, and controls the smoothness of the interpolated surface. This should make it possible to identify the regions that need remediation in order to improve crop development and implement the proper land management procedures.

1.2. Soil of Jazan city

Situated on a salt dome that rises around 50 meters above sea level and covers an area of 4 square kilometers, Jazan City is a fascinating natural feature. Around 190 kilometers of Sabkha flats and extensive stretches of wind-deposited sand encircle this highland area. The incredibly diverse soil composition of these Sabkha deposits poses significant geotechnical challenges [6]. Previous studies have highlighted the low bearing capacity and high compressibility of Jazan's Sabkha soil, making it crucial to consider these factors in construction and infrastructure projects [7–10].

Jazan has environmental and geotechnical issues due to its unique geological heritage [11]. With a shallow water table only a few meters below the surface, the ground surface mainly comprises layers of loose sand and salty clay. Given Jazan's location inside the seismically active southern Red Sea zone,

this composition makes the soil extremely vulnerable to liquefaction during seismic events. The fact that the soil profile differs throughout the area adds to the difficulty. While very soft clay and silt deposits rich in organic matter are found in inland zones, loose, fine sand predominates in coastal locations. Jizan's location on a salt dome encircled by low-lying terrain is mostly responsible for this unpredictability. "Sabkha" soil, a salty crust formed by evaporation, appears as a result of its proximity to the Red Sea. The prevalence of salt domes and salt-rich soil significantly impacts foundation stability and construction practices [12]. Furthermore, the entire Jazan region is covered by alluvial sediments deposited by the Red Sea. These sediments, overlaying Tertiary deposits from 5 to10 km wide, have experienced various tectonic movements, contributing to the region's susceptibility to earthquakes [13–16].

This study's novelty reflects the sophisticated integration of regional climate variables. Local microclimate effects: The investigation showed that local microclimatics brought about by topographical features like hills and valleys have a significant impact on the soil qualities in Jazan City. Published research, which frequently rely on broad climate data, do not sufficiently address this precise understanding of how local climate influences soil qualities. Additionally, the accuracy of common interpolation methods (such Spline and the natural neighbor) varied depending on the type of soil and the location. Notably, the application of machine learning-based methods such as various Random soils and Gradient Boosting yielded better outcomes in areas with complex soil variability a comparison that had not been previously documented in the specific context of Jazan.

1.3. Soil classification systems by the food and agriculture organization (FAO)

The Food and Agriculture Organization's (FAO) soil classification system was developed in three main phases. The legend first appeared in the 1974 FAO–UNESCO Soil Map of the World, which was completed between 1960 and 1981. The FAO was the main organization in charge of this project. The Revised Legend of the Soil Map of the World was the outcome of the second phase, which involved refining the previous legend into a more thorough system between 1981 and 1990. Ultimately, the FAO updated legend served as the foundation for the International Society of Soil Science (ISSS), which had been developing a global soil nomenclature since 1981 [17]. According to the FAO Legend of 1974, the lowest category identified on the global soil map was 26 Great Soil Groups (world classes) divided into 106 Soil Units (world subclasses).The information of soil associations, which made up the majority of soil mapping units, were listed on the back of each printed map sheet [18–23].

1.4. GIS and open source soil data for soil mapping

Soil mapping is essential for understanding land capabilities, agricultural productivity, and environmental management. Combining open-source soil data with GIS has enabled the optimization of soil mapping processes in Jazan, Saudi Arabia. Many benefits of using GIS will be achieved in this work:

a. GIS provides strong capabilities for mapping soil data, helping stakeholders understand spatial patterns and relationships;
b. GIS supports sophisticated interpolation techniques like Kriging and Inverse Distance Weighting, increasing the accuracy of soil property estimates across the region;
c. These techniques close data collection gaps by forecasting soil properties in unsampled regions;
d. GIS systems effectively manage large datasets, speeding up the data retrieval and analysis process; e. Open source data is publicly available, which reduces the cost of data acquisition and promotes greater information accessibility;
e. Open source data is publicly accessible, which promotes greater information accessibility and lowers data acquisition costs;
f. Open data enhances the breadth and depth of soil databases by encouraging contributions from academics and experts; and
g. Local researchers in Jazan can use global soil databases to verify and compare their results with international standards.

2. Materials and methods

In this study, the methodology carried out focuses on literature review, collecting the soil samples from 23 boreholes, digitizing a base map, downloading FAO soil data, carrying out the experimental test, classifying soils, choosing and examining the interpolation methods, combining the FAO data, creating a digital soil data model, and updating the digital model of Jazan City. The following is a list of the actions that will be taken to complete each phase in the approach (see Fig. 1).

2.1. The study area

The capital and administrative center of the Jazan Municipality is Jazan City, which is situated in the extreme southwest of Saudi Arabia. This

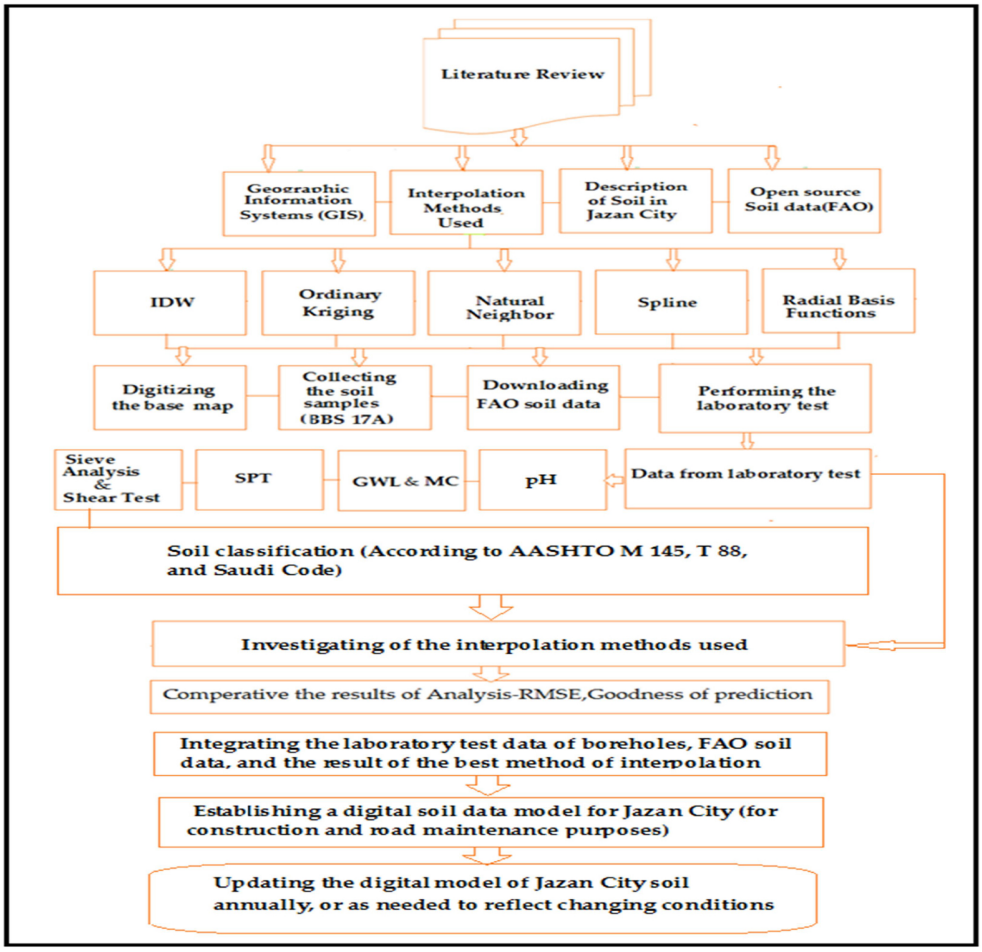

Fig. 1. Methodology flow chart.

region is strategically situated between longitudes 16.4–18.33°N and 41.4–43.4°E, with a 330-kilometer shoreline along the Red Sea to the west, the Asir region to the north and east, and the Republic of Yemen to the south and southeast (see Fig. 2).. Although the Jazan Municipality only occupies about 477.8 square kilometers of land, it controls about 80 Red Sea islands. This archipelago includes Farasan Island, the largest among them, boasting an area of about 702 square kilometers. Despite its expansive maritime territory, the Jazan Municipality represents a mere 0.7% of Saudi Arabia's total landmass, making it one of the Kingdom's smallest regions. In November 2006, Saudi Arabia welcomed its fourth economic city, Jazan Economic City, situated in Baish, approximately 60 kilometers northwest of Jazan City. This ambitious project spans 113 square kilometers along a 12-kilometer stretch of the Red Sea coastline.

One significant national economic project in Saudi Arabia's Jazan Province is the Jazan Economic City (JEC). The City represents a pivotal moment in the Kingdom's economy's ongoing growth and development. Under Saudi Aramco's supervision, preliminary construction on a pioneer port, refinery, mineral processing plant, roadways, and other supporting infrastructure has been completed [24].

2.2. Soil data collection

Using data from 23 strategically placed boreholes across the study region, researchers used an interpolation technique to evaluate the accuracy of soil forecasts throughout Jazan City. The following factors were probably taken into consideration when choosing the 23 borehole locations in the Jazan region:

a. The locations of the boreholes were selected to guarantee a representative distribution across the region's various soil types and landforms.
b. Infrastructure and closeness to highways or existing facilities were taken into consideration when choosing sites for drilling and data gathering.

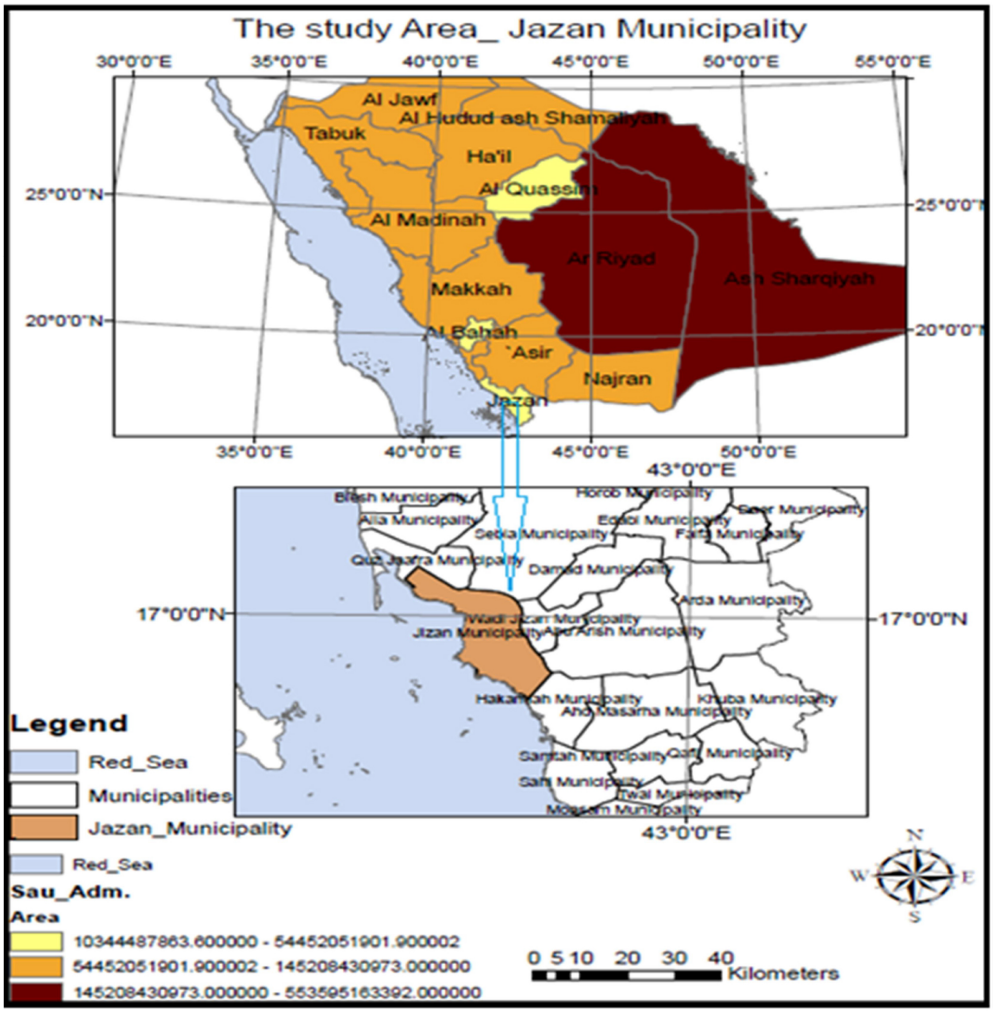

Fig. 2. The base map of the study area(derived from AcGIS10.4).

c. To capture a wide range of properties and improve the overall quality of soil data, areas with a variety of soil characteristics were given priority.
d. The choice to evaluate soil characteristics pertinent to particular uses may have been impacted by different land use patterns, such as agricultural, urban, or natural areas.
e. The placements of boreholes may have been chosen for their significance to the area hydrology, such as being close to erosion and flooding.

A soil core sampling device was used to collect soil samples, and a global positioning system was used to record the samples' exact locations. Additionally, using total station Topcon version 7500, ten randomly selected test points of elevation data are observed; these points are utilized to examine the prediction's cross-validation. A thorough spatial analysis was made possible by the use of a GIS technique. The sampling sites were Hai Alshati, Hai Alrawabi, King Abdullah Suburb, Hai Alsafa, ALmaragan, The Industrial Area, Alarak, Anaseem, Hai Alshamia, Alrawabi Embankment Road, and Jazan. The sampling sites, which comprised Hai Alshati, Hai Alrawabi, King Abdullah Suburb, Hai Alsafa, ALmaragan, The Industrial Area, Alarak, Anaseem, Hai Alshamia, Alrawabi Embankment Road, and Jazan Sea Port, represented a variety of locations within Jazan City. Technology and geographic information science (GIScience) have the ability to greatly raise the standard and efficacy of techniques used to collect spatial soil data [25, 26]. Because scientists are using soil data more and more elaborately, GIS-based predictive soil mapping is essential [27–32]. For instance, chemical and physical soil data are needed for ecological and hydrological process models that assess worldwide change [33–36]. A Boyles (BBS 17A) core borehole rotating rig was used to gather soil data samples from 23 boreholes for foundation reasons. These samples

Table 1. Summary of data collected and the experimental soil tests of 23 boreholes.

Borehole ID	Easting (m)	Northing (m)	Elevation (m)	Soil Type	Moisture Content (%)	Chlorides (CL) mg/kg	SPT	GWL
BH-1	245142.0	1880239.0	23	Clay	15	218	20.2	7.5
BH-2	243116.0	1880979.0	9	Clay	16	218	20	4.5
BH-3	238669.8	1877571.0	2	Silty sand	11	2626	2.6	1.5
BH-4	244620.0	1873842.0	12	Silty sand	11	2626	57.8	1.5
BH-5	243774.0	1873517.0	7	Silty sand	11	2626	38.2	1.5
BH-6	244237.0	1870045.0	8	Silty sand	11	1806.6	16.6	3.5
BH-7	245431.0	1873993.0	15	Clay	16	7894	54.6	1.5
BH-8	247781.1	1866891.4	10	Silty sand	11	17820	21.4	4.5
BH-9	239790.2	1869156.6	5	Silty sand	11	11100	20.2	1.6
BH-10	238197.4	1869214.4	9	Silty gravel	10	10750	13	0.8
BH-11	237795.2	1869171.6	4	Silty gravel	10	3583	4	2.5
BH-12	249345.7	1874161.2	20	Clay	17	300	41.4	6.5
BH-13	248160.1	1862348.9	6	Silty sand	11	10750	14.6	4.8
BH-14	244557.3	1862781.2	2	Silty gravel	10	10750	33	1.0
BH-15	242470.5	1863409.0	1	Silty gravel	10	10750	4.4	0.0
BH-16	242580.0	1877037.0	9	Silty sand	11	10750	24.6	4.2
BH-17	242638.0	1874438.0	4	Silty sand	11	8000	17.4	1
BH-18	243867.0	1874386.0	3	Silty sand	11	6500	19	1
BH-19	242562.0	1876048.0	6	Silty sand	11	8500	20.2	8.0
BH-20	242740.0	1878205.0	5	Silty sand	11	9000	21.2	2.6
BH-21	241178.0	1879306.0	6	Silty sand	11	9000	13.4	3.1
BH-22	240958.0	1875991.0	4	Silty sand	11	7500	5.8	1
BH-23	246692.0	1882534.0	9.5	Clay	18	300	20	8

were then evaluated in the geotechnical laboratory of the geotechnical and material testing facility in Jazan City. These samples had a core test diameter of 7.6 cm and were taken from a depth of 3-6 m. Each borehole's geographic location (coordinates), soil types at five depths every two meters, groundwater level (GWL), water content, elevation of the borehole above mean sea level, technical office name, and more were all documented in the output test paper. These soil data samples were subjected to tests, including bearing capacity, Procter test, Standard Penetration Test (SPT), and sieve analysis. The gathered and tested data, along with their locations within the study region, are displayed in Table 1 and Fig. 3.

2.3. Soil samples laboratory tests and mapping

The majority of site studies for building projects or assessing possibly contaminated or polluted land involve very shallow drilling and sampling [37]. Sieve analysis, direct shear, and compaction testing were among the laboratory experiments that were performed. The sieve analysis determines the gradation of the distribution of aggregate particles by size inside a particular sample in order to evaluate compliance with design, production control standards, and verification requirements [38]. The Direct Shear Test is used to determine a soil's consolidated drained shear strength. A specimen is put under controlled deformation on or close to a single shear plane in order to perform the test [39, 40]. Over the past few decades, a great deal of research has been done on the shear of rock fractures using laboratory direct shear testing on individual rock fractures. The joint mechanical properties, such as the peak and residual shear strengths and the friction coefficient, or the joint hydro-mechanical properties, such as the stress-dependent permeability, with additional coupled or uncoupled fluid flow tests, have been the primary focus of this research [41–45].

Compaction of soil samples at specified water content in a conventional mold with standard compaction energy constitutes the Proctor compaction test. In the traditional Proctor test, three distinct soil layers are compacted using a 4-inch-diameter mold and 25 blows from a 5.5-pound hammer that falls 12 inches [46, 47]. For sustainable land use and management, Jazan soil mapping using Geographic Information Systems (GIS) and interpolation techniques is crucial. Stakeholders may make educated decisions that increase agricultural productivity, protect the environment, and promote sustainable development by employing cutting-edge approaches. Experiences from throughout the world provide insightful viewpoints and methods that can be adapted for local use. Applications of GIS and soil mapping have been employed worldwide for similar purposes in a number of places. The following are included in the international context:

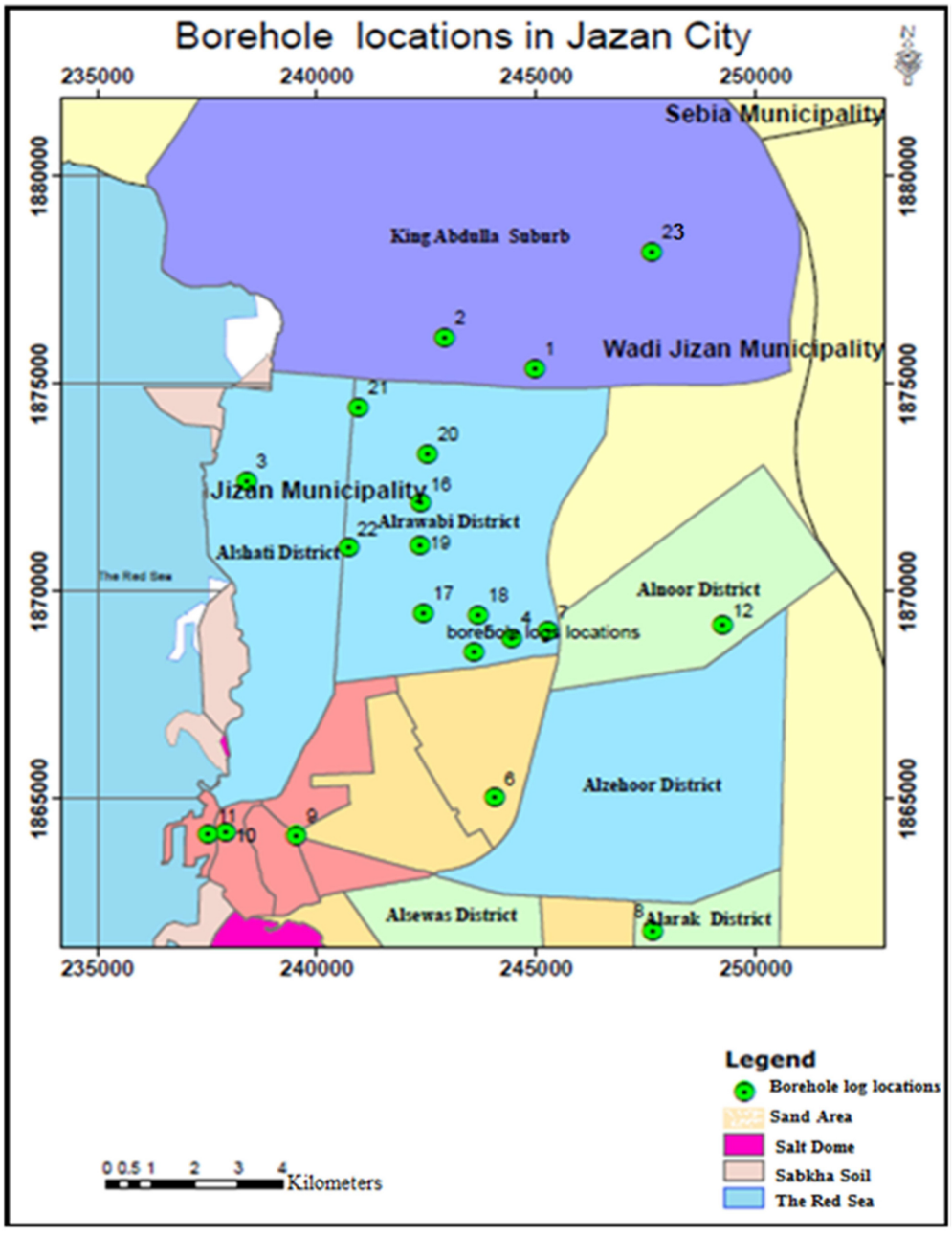

Fig. 3. Sample of boreholes locations in the study area (Presented by ArcGIS10.4).

- **Europe:** Advanced soil mapping initiatives using GIS have been implemented to address climate change impacts on agriculture.
- **United States:** The USDA utilizes soil surveys and GIS for land use planning and conservation efforts.
- **Australia:** Soil mapping is crucial for managing arid landscapes, where soil properties significantly influence land management decisions.

2.4. Interpolation methods

Estimates of the physical and chemical properties of soil are obtained by interpolating data from geo-information systems. The inverse distance weighting (IDW) approach and the Kriging method the natural neighbor approach and the spline interpolation method are the two most used techniques for this

purpose [47]. The study region is where the modeling accuracy is calculated. Data analysis reveals that the most effective methods for assessing the correctness of a digital relief model made with the ArcGIS 10.4 geo-information system are spline algorithms and inverse distance weighting. To estimate points surrounding the sample data, each interpolation method employs a distinct statistical method. The particular statistical studies for each interpolation method are listed below. The inverse distance weight (IDW), the Kriging, the natural neighbor, and the Spline are the first four. Finding the experimental structure of the observed data and creating a theoretical model for it form the basis of geostatistical research [48–50].

2.4.1. Inverse distance weight (IDW)

IDW is a popular interpolation technique that uses measurements of the surrounding values of the projected sites to predict the values of unmeasured locations [51]. The foundation of IDW is the notion that the impact of a phenomenon decreases with increasing distance. Compared to samples put farther away, those closer should contribute more to assessing the unknown regions [52]. The interpolating function is as follows:

$$\mathbf{Z}(\mathbf{x}) = \frac{\sum_{i=1}^{n} \mathbf{w_i z_i}}{\sum_{i=1}^{n} \mathbf{w_i}} \tag{1}$$

$$\mathbf{W_i} = \mathbf{d_i}^{-\mathbf{u}} = \frac{1}{\mathbf{d_i}^{\mathbf{u}}} \tag{2}$$

Where $\mathbf{w_i}$ indicates the weight assigned to spot i, di indicates the distance between spot i and the prediction point, n is the total number of known locations utilized in interpolation, and Z(x) is the estimated value at an interpolated place. In closer proximity to the interpolated spot, values receive higher weighting values. Weight drops as distance rises, and u, the weighing power, determines the weight reduction method as the distance grows [44]. The distance measured from the interpolation point which is defined as:

$$\sqrt{(\mathbf{x} - \mathbf{x_i})^2 + (\mathbf{y} - \mathbf{y_i})^2} \tag{3}$$

The coordinates of the default data point and the interpolation point are (x, y) and (xi, yi), respectively.

2.4.2. Kriging

The theory behind Kriging is that the interpolated parameter could function as a regionalized variable. Regression Kriging performs better than other statistical and geostatistical models, as Bishop and McBratney have shown [53]. Regression Kriging is less effective than multiple linear regression at the landscape scale, according to Scull et al. [54], if soil is not tested at distances closer than the typical range of spatial dependency. The following is the interpolating function:

$$\mathbf{Z}^* = \sum_{i=1}^{n} \mathbf{w_i z_{(X_i)}} \tag{4}$$

Z* displays the interpolated value, w_i is the statistical weight for sample xi, and Z(xi) is the observed value at location xi [55].

2.4.3. Natural neighbor

Robin Sibson developed the “natural neighbor” (or “Sibson”) interpolation technique for spatial interpolation [56]. The Voronoi tessellation of a discrete collection of spatial places serves as the foundation for the technique. Compared to simpler interpolation techniques like nearest-neighbor interpolation, this approach offers a smoother approximation to the underlying “true” function.

The natural neighbor interpolation approach uses a weighted average of the nearby observations, with the weights corresponding to the “borrowed area [57–60]. The fundamental formula is:

$$\mathbf{G}(\mathbf{x}) = \sum_{i=1}^{n} \mathbf{w_i}(\mathbf{x})\, \mathbf{f}(\mathbf{x_i}) \tag{5}$$

Where:

G(x) : Represents the estimate at x,$\mathbf{w_i}$, **$\mathbf{f(x_i)}$ are the known points of porhole at $\mathbf{x_i}$.**

$\mathbf{w_i}$: the weight of sandi scalculated by the following formula

$$\mathbf{w_i}(\mathbf{x}) = \frac{\mathbf{A}(\mathbf{x_i})}{\mathbf{A}(\mathbf{x})} \tag{6}$$

Where A(x): Is the volume of the new cell centered in x.

$\mathbf{A(x_i)}$:The old cell centered in $\mathbf{x_i}$and the new cells centered in x intersect at this volume.

2.4.4. Radial basis functions (RBFs)

Radial basis functions (RBFs) are a collection of exact interpolation techniques that demand that the surface cross each observed sample value. The value of a radial basis function (RBF) is determined by the distance (often the Euclidean distance) to a center (xc) in the input space. The Gaussian RBF is the most widely used RBF. It is defined as follows [61] and has the same form as the Gaussian probability density function’s kernel.Radial basis functions (RBFs) can be

predicted using the following formula:

$$\varphi(\mathbf{x}, \mathbf{x_c}) = \mathbf{exp}\left(-\frac{\|\mathbf{x} - \mathbf{x_x}\|}{2\sigma^2}\right) \quad (7)$$

Where:

$\|\mathbf{x} - \mathbf{x_x}\|$:Represents the Euclidean distance (the length) between two points.

(x) and (x_c): Complex numbers in the complex plane.

The value of RBF decreases with increasing distance. Therefore, the local properties of a nonlinear function at x_c can be approximated using an RBF with center x_c. How big σ is determines how smooth a Gaussian RBF is [62].

2.4.5. Spline

Smooth trajectories, like those in computer graphics or robot motion, are commonly produced using spline interpolation. In general, these splines have to meet four mathematical requirements:

- They must be (at least) C1 continuous; all differentiable functions with a continuous derivative belong to class C1, which is also known as continuously differentiable functions;
- They must be interpolating Splines, meaning that they pass through all of their control points;
- They must be affine invariant, meaning that the sum of all blending polynomials is always 1,and
- They must allow local control, meaning each control point affects only a tiny, compact portion of the generated curve [63].

The surface interpolation formula utilized by the Spline tool algorithm is as follows:

$$\mathbf{S}(\mathbf{x}, \mathbf{y}) = \mathbf{T}(\mathbf{x}, \mathbf{y}) + \sum_{\mathbf{j}=1}^{\mathbf{N}} \lambda_{\mathbf{j}} \mathbf{R}(\mathbf{r_j}) \quad (8)$$

Where:

j= 1,2,3,. . ...n, n represents the number of points, λ_j the are coefficients that are discovered while solving a set of linear equations,$\mathbf{r_j}$is the separation between the point and the point (x,y).

Depending on which option is chosen, T(x,y) and R(r) have distinct definitions [64].

2.5. Root mean square error (RMSE)

The standard deviation of the residuals or prediction mistakes is known as the root mean square error, or RMSE. The residuals' spread, or the separation between the data points and the regression line, is measured by RMSE. To put it another way, it displays the percentage of the data that is centered on the line of greatest fit. Because it enables us to compute the average difference between the actual and expected data values, this number is useful. This is in contrast to the model's R-squared, which displays the percentage variation in the response [4]. This is how the RMSE is computed:

$$\mathbf{RMSE} = \sqrt{\frac{\sum_{\mathbf{i}=1}^{\mathbf{n}} [\mathbf{Z}^*(\mathbf{x_i}) - \mathbf{Z}(\mathbf{x_i})]^2}{\mathbf{n}}} \quad (9)$$

Where:

Z (xi) represents the observed value at site $\mathbf{x_i}$, Z* (xi) displays the interpolated value at site xi,

2.6. Assessing interpolation results with cross-validation techniques

This study examines the model's ability to anticipate values at new locations before selecting an interpolation model for decision-making. However, how can you determine whether your model accurately and consistently forecasts the values if you don't know the actual values at the places between the measured points? You must be aware of the values of the locations you haven't sampled in order to respond to this question. Nonetheless, cross-validation is a popular and extensively used technique for evaluating the accuracy and dependability of interpolation. Many statistics are computed for each point during cross-validation [65]. These statistics can be viewed on the Table tab of the cross-validation dialog box or saved in a feature class using the Cross Validation tool. The following fields are produced for each point:

a. Measured: The measured value of the hidden point; b. Anticipated: The values predicted by cross-validation at the hidden point location;
b. Error: The discrepancy between the observed and projected values (predicted minus measured). If the error was positive, the prediction was more significant than the measured value; if the error was negative, the opposite is true;
c. Standard Error: The standard error of the estimated value. If the errors are regularly distributed, almost two-thirds of the measured values will fall within one standard error of the predicted value, and around 95% will fall within two standard errors;
d. Standardized error: The error divided by the standard error; in order to utilize the probability output kinds, standardized error values must follow a standard normal distribution (mean equal to zero);
e. The procedures for cross-validation using ArcGIS10.4 in this paper are as follows:

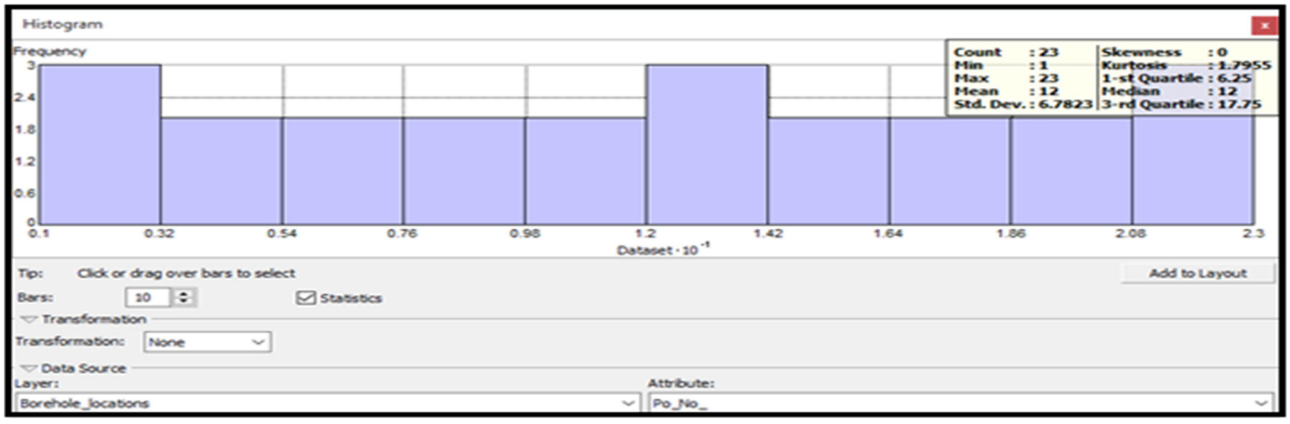

Fig. 4. Histograms and QQPlot for borehole distributions in the study area.

- Open the geostatistical analyst extension;
- Load your dataset as a point feature class;
- Select Interpolate Point from the geostatistical analyst tools menu;
- Choose your preferred interpolation method, such as Kriging and
- Use Leave-One-Out Cross-Validation (LOOCV) by looking at the Cross-Validation option.

A validation method called Leave-One-Out Cross-Validation (LOOCV) involves removing one observation from the dataset and training the model on the remaining data, repeating this process for every observation. A metric called Root Mean Square Error (RMSE) measures the discrepancies between actual and predicted values and offers information about the correctness of the model. Standard references like "The Elements of Statistical Learning" by Hastie et al. and "Pattern Recognition and Machine Learning" by Bishop are recommended for those seeking a deeper comprehension of these ideas. After each data point is successfully removed from the dataset, the model is trained on the remaining data points using the LOOCV technique [66].

3. Results and discussion

3.1. Investigative statistics for soil properties

In this study, Fig. 4 displays summary data and the frequency distributions of soil properties as a histogram. A unimodel shape can be seen in the histograms. The data diverged from a normal distribution, as indicated by the positive and small values of skewness (0.5-0) and kurtosis values (2.6–3.1) near 3.0 [67].

Moreover, the context of soil tests, linkage analysis usually entails looking for patterns or correlations in the interactions between different soil parameters. Fig. 5 shows the results of normal QQ plot of soil data using GIS.

3.2. Results of analysis using ArcGIS10.4 and laboratory tests

The analysis used to generate the paper's findings is broken down into the following phases:

3.2.1. The ideal method of soil data interpolation results

The optimum method for predicting the missing value of soil data in soil sample collections from 23 boreholes in Jazan City was identified by comparing several interpolation strategies using the geo-statistical expression of root mean square error (RMSE). The RMSE values for heights, pH, and GWL prediction data were shown in Table 2 and Fig. 6.

The results of the analysis were obtained using ten control points. For elevation, pH, and groundwater level (GWL), the Natural Neighbor approach produced the best RMSE values for the application, which are 0.171, 0.081, and 0.362, respectively. A low RMSE denotes better performance and a smaller difference between prediction and actual borehole data. However, 0.081 is the application's best RMSE result. This value is shown following the application of the natural neighbor interpolation approach.

Moreover, the result of spatial analysis (interpolation of heights in Jazan City is shown in Fig. 7)

3.2.2. Results of soil data laboratory test

This work employs a GIS-based approach to manage and assess soil data, which comprises essential procedures including data storage, interpolation, classification, and map creation. This approach creates a reliable GIS soil model and makes it easier to understand soil properties across the study area. This model can be a useful tool for a variety of

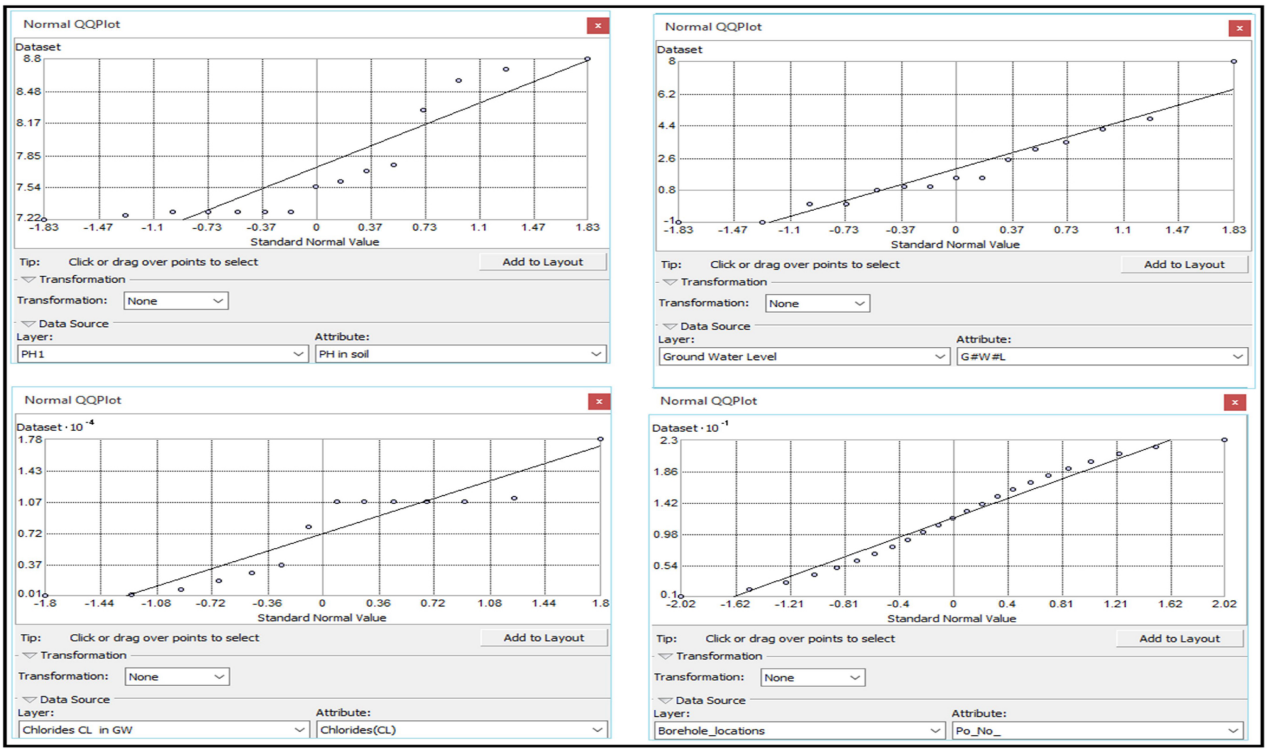

Fig. 5. Normal QQPlot for soil properties.

Table 2. Compare the Root mean square error (RMSE) values using the difference of interpolation methods.

Interpolation Model (I.M)	IDW	Kriging	Natural Neighbor	Spline	RBF
Mathematical models (M.Mo)	1.0469	1.0729	0.171	0.9172	1.822
RMSE _Elevation Data					
RMSE _ pH Data	0.5294	0.579	0.081	0.1734	0.4955
RMSE _GWL Data	2.7795	2.2986	0.3619	0.4361	2.7216

infrastructure projects, including construction, utility services, and road development, by providing important insights on soil qualities and their geographical distribution.

When assessing land for pollution or construction, investigations may involve collecting soil samples and shallow drilling. These samples are then examined in a laboratory to determine their qualities and suitability for building. The study's samples were collected in the Hyderabad area. Twenty-three (23) samples were collected between two and four feet below the surface. Sieve analysis, the direct Shear Test, chemical features, organic matter content, chloride content, sulfate content, and PH value were among the laboratory tests performed on the materials at the Geotechnical Laboratory. Department of Architecture and Civil Engineering in accordance with AASHTO and ASTM guidelines. The AASHTO approach has been used to classify the soils in these samples. The Proctor Compaction Test, Direct Shear Test, and Sieve Analysis are the tests used as an example of this excremental job.

This study focuses on shallow drilling and sample methods that are frequently used in contaminated soil assessments or building site investigations. A number of laboratory tests were carried out to characterize the soil samples that were obtained, including:

-Sieve Analysis: This test establishes a soil sample's particle size distribution (grading), which is crucial for evaluating the soil's suitability for different construction uses and guaranteeing adherence to design guidelines. The sieve analysis test graph for the sample taken from the ALrawabi embankment road borehole in Jazan City is displayed in Table 3 and Fig. 8.

-Direct Shear Test: Consolidated drained shear strength, a critical metric for comprehending soil behavior under load and guaranteeing the stability of slopes and foundations, is measured by this test. The hydro-mechanical characteristics of rock cracks have

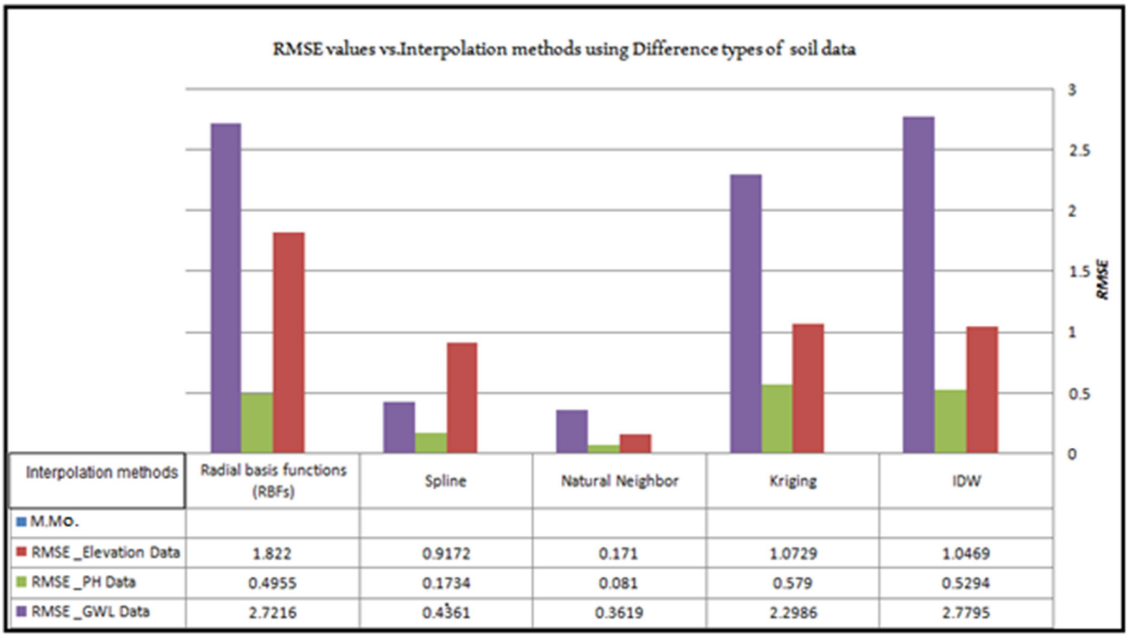

Fig. 6. Graph of comparing the RMSE vs. interpolation Methods using difference soil data.

Table 3. Sieve analysis test.

Sieve Size(mm)	Weight of soil Retained (gm)	Weight of soil Pass(gm)	Cumulative of Retained%	Total Cumulative of Retained%	pass on each Sieve%
31.5	570.3	2429.7	19.01	19.01	80.99
22.5	339.2	2090.5	11.31	30.32	69.68
19.0	197.3	1893.2	6.58	36.9	63.1
12.5	246.8	1646.4	8.23	45.13	54.87
9.5	175.6	1470.8	5.85	50.98	49.02
4.75	264.1	1206.7	8.80	59.78	40.22
2	589.7	617	19.67	79.45	20.55
0.425	548.8	68.2	18.29	97.74	2.26
0.250	50.4	17.8	1.68	99.42	0.58
0.075	14.2	3.6	0.47	99.89	0.11
pan	0	0			
$\sum$	2996.4		99.89		

Table 4. The laboratory results in values of normal stress and shear stress.

Normal Stress (Kg/Cm^2)	Shear Stress(Kg/Cm^2)
0.25	0.5
0.4	0.8
0.6	1.1
0.75	1.4
1	1.74

also been thoroughly studied using the direct shear test, especially with regard to stress-dependent permeability and shear strength traits. The direct shear graph of the borehole in the Alshawajrah area is shown in Table 4 and Fig. 9. The friction angle (Φ) equals 29° as a result.

Proctor Compaction Test: The ideal moisture content at which a soil can reach its maximum dry density under a certain compaction effort is determined by this test. For construction projects to achieve proper soil compaction and guarantee the long-term stability and functionality of foundations, pavements, and embankments, this knowledge is essential. The proctor test results at the ALshawajra borehole location are displayed in Fig. 10.

The soil of the study area is classified according to the study carried out by Ali and Mahmoud [68], furthermore the studies carried out by Mahmoud, et al.2024 [69, 70], Jazan has silt, clay, and sand soil. Jazan's soil has a cohesiveness of ($C = 0.2$ kg/cm^2) and an internal friction angle of ($\Phi = 31°$). Thus, the

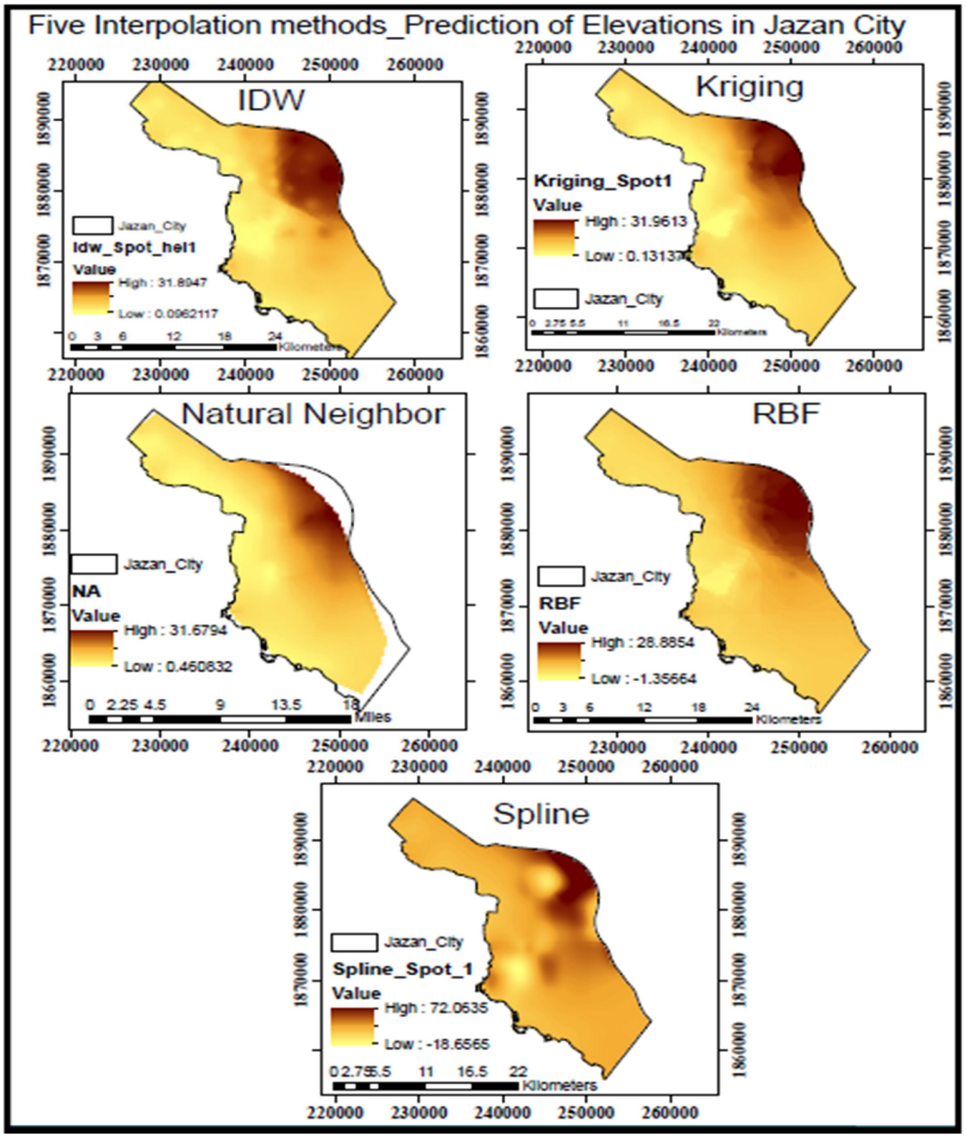

Fig. 7. Predicting heights in Jazan City by different interpolation methods: natural neighbor, Kriging, Natural neighbor, Radial basis functions (RBF), and Spline (plotted by ArcGIS10.4.)

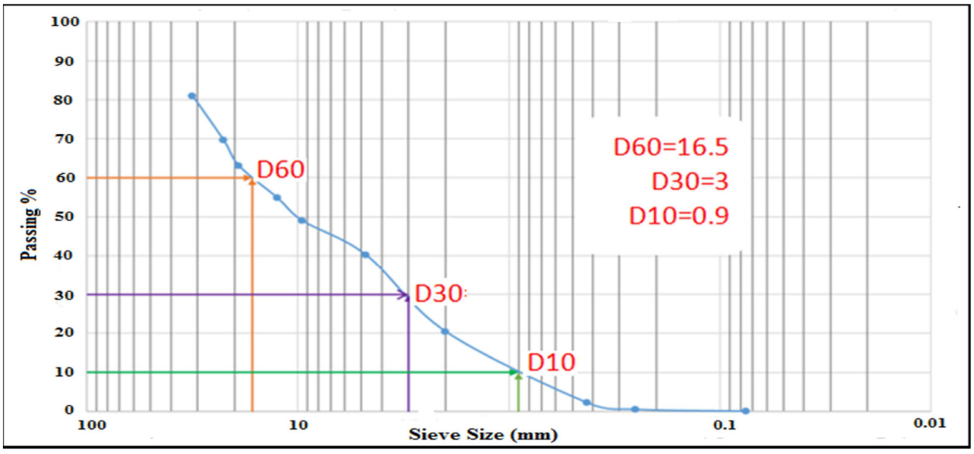

Fig. 8. The sieve analysis graph of the ALrawabi embankment road borehole (drawn by authors according to AASHTO T-88).

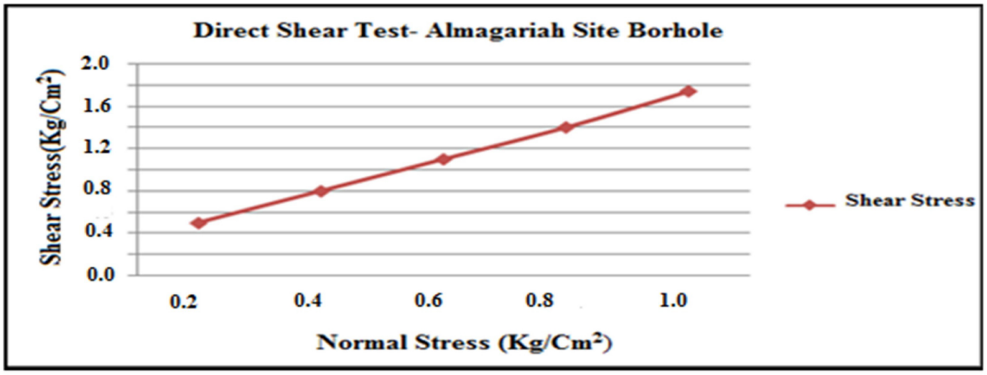

Fig. 9. The direct shear graph of the Alshawajrah area site borehole (drawn by authors).

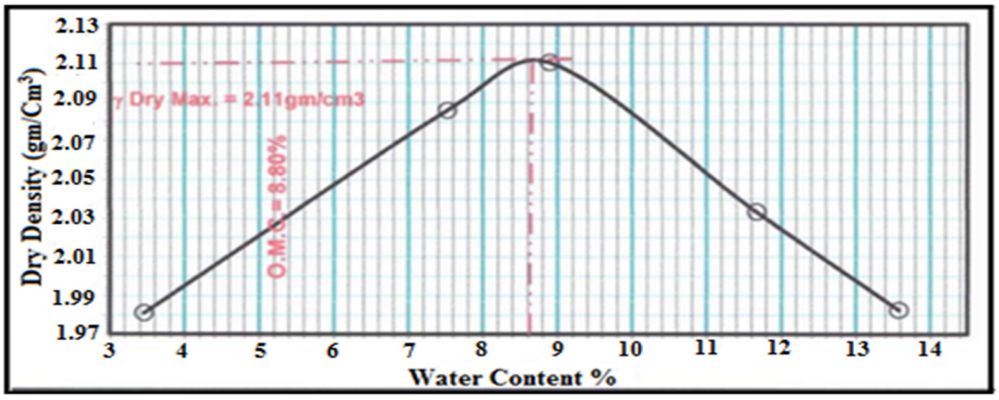

Fig. 10. The compaction test graph of the ALshawajra borehole location (drawn by authors According to AASHTOD T99).

unconfined compressive stress at failure is between 1.0 and 22.0 kg/cm^2 (qu).

3.2.3. Mapping layout of Jazan City for construction and road maintenance using ArcGIS10.4

The overlay analysis produced visual representations of the study area's contour maps, VAO soil classification map, and geotechnical map, all of which are crucial for evaluating the soil management practices employed there (Figs. 11 to 13). As an output isolation data set, the contour map displays Jazan City's various elevations.

3.2.4. Cross-validation of the results

Cross-validation is a powerful technique that is often used to evaluate the accuracy of interpolation methods [71]. This approach carefully removes data points, often one at a time, and then forecasts the value at the removed place using the remaining dataset. By comparing the expected and observed results, researchers can calculate the interpolation error. By utilizing the complete dataset for both model training and validation, cross-validation is a lovely method that addresses the issue of sparse data [72]. This approach, created by Stone and Geisser, provides a trustworthy assessment of an interpolation model's prediction ability. This study evaluated the performance of five different interpolation algorithms using a cross-validation methodology. The fit quality of the interpolation model was assessed using regression analysis, also known as the coefficient of determination. The degree of agreement between the observed values and the model's predictions is shown by this statistical indicator. A stronger fit, or how well the model represents the underlying patterns in the data, is indicated by a more significant coefficient of determination, or R^2. The robustness and reliability of the chosen interpolation technique for estimating soil parameters are ensured by this comprehensive evaluation process, which makes use of goodness-of-fit and cross-validation metrics [73]. ArcGIS10.4 is used in this work to summarize the cross-validation analysis of all soil data. Table 5 and Fig. 14 show the sample cross-validation summary utilizing the interpolation method of soil pH.

Investigating the Natural Neighbor interpolation approach delivers crucial insights that have broad uses in agriculture, land management, environmental protection, and urban planning. This approach encourages sustainable practices in various industries

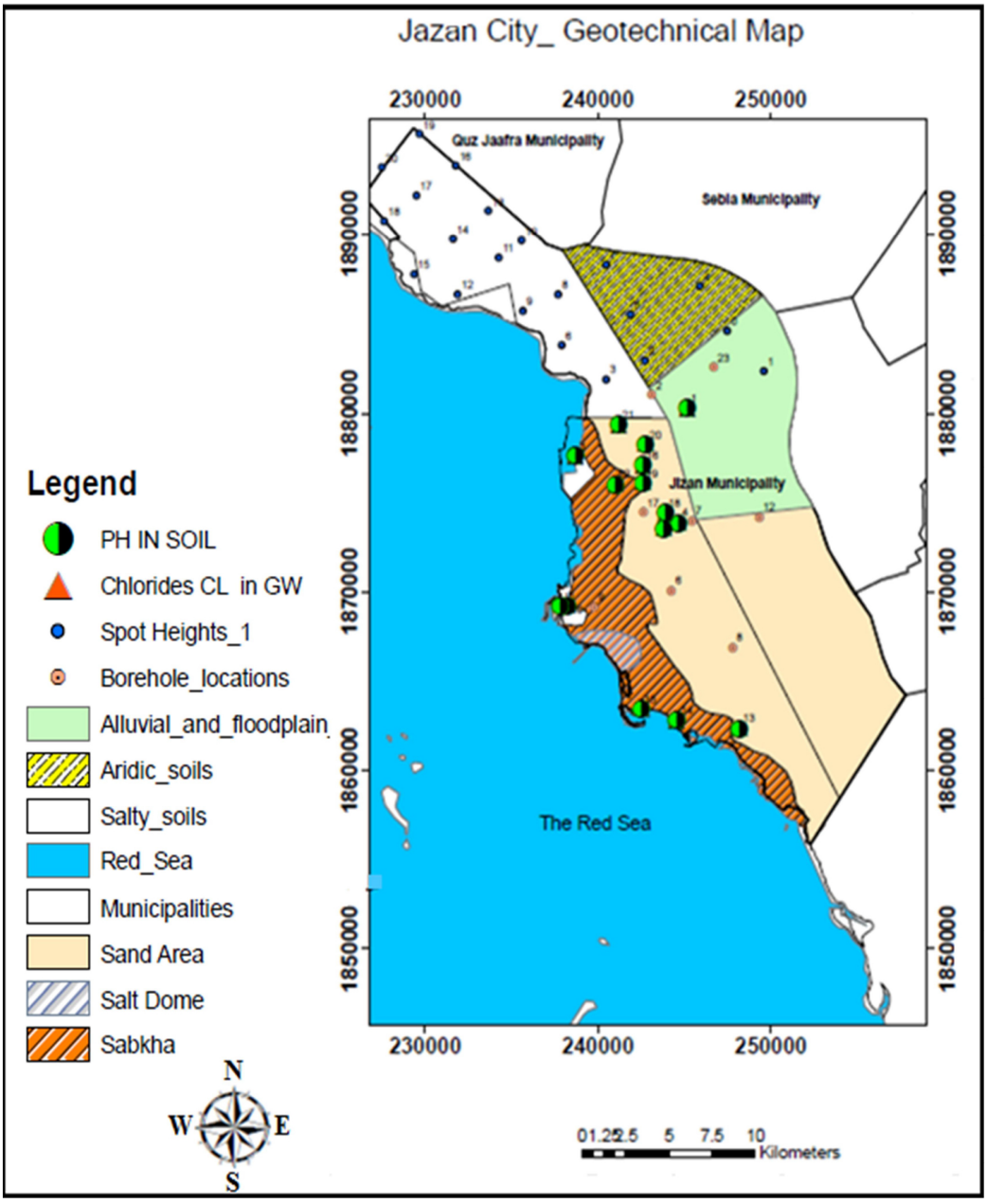

Fig. 11. Geotechnical Map and Soil classification (presented in ArcGIS10.4).

Table 5. Cross-validation of errors of ten test points of pH in soil test data.

Interpo. Method(I.M) M.Mo	IDW	Kriging	Natural Neighbor	Radial basis functions (RBF)	Spline
Mean	–0.022	–0.0380	–0.005	–0.0540	0.061
Mean Absolute Error(MAE)	0.4380	0.4180	0.0651	0.4140	0.1450
Mean Square Error(MSE)	0.2833	0.2792	0.00651	0.2647	0.0465
Root Mean Square Error(RMSE)	0.5322	0.5284	0.08062	0.5145	0.2155
Regression function	y=-0.218 x + 5.91	y=-0.153 x + 6.51	y = 0.907x + 0.73	Y=0.507x + 3.7	y = 0.879x + 1.058
Value of R-Square %	24.2%	8.7%	98%	63.6%	92.8%

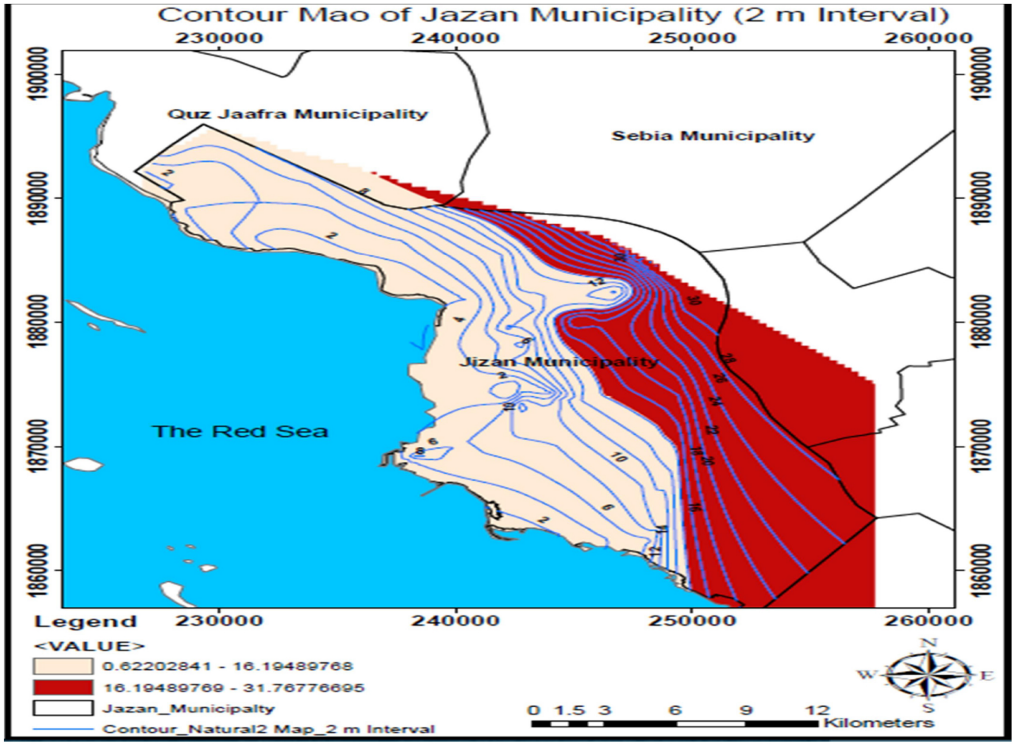

Fig. 12. The contour map of Jazan City_2 m Interval (is plotted by ArcGIS10.4 according to the Natural neighbor interpolation method (presented in ArcGIS10.4)

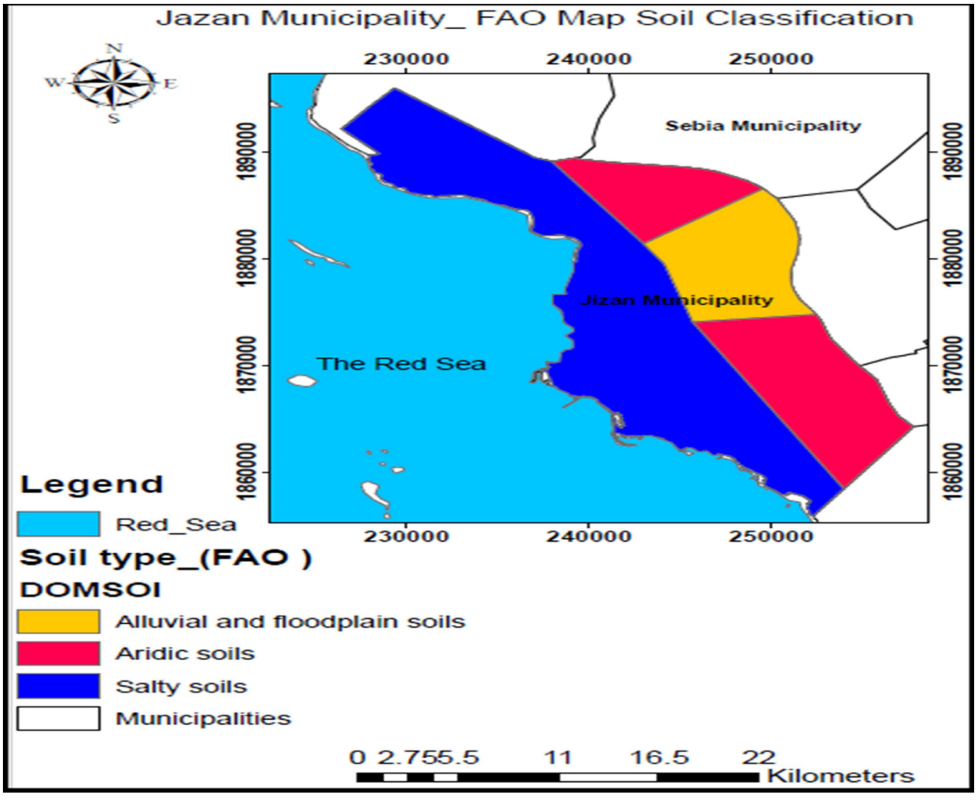

Fig. 13. Soil classification systems by FAO (Downloaded from FAO–UNESCO Soil Map of the World, and presented by authors in ArcGIS10.4).

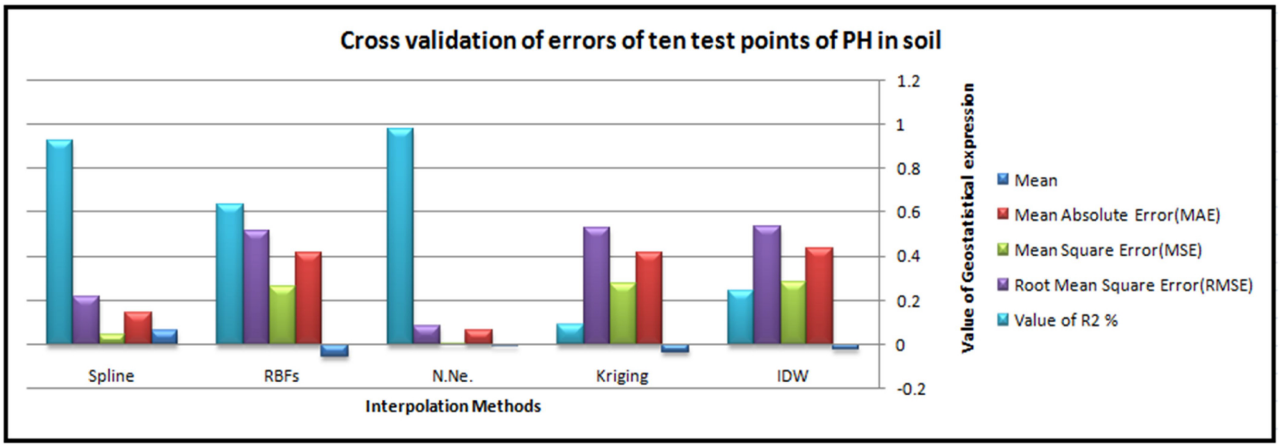

Fig. 14. Cross-validation of pH in soil (derived from ArcGIS10.4).

Table 6. Comparative Analysis of Interpolation Methods.

Interpolation Method	Root Mean Square Error(RMSE)-(prediction of Elevation data)	RMSE((prediction of Potential Hydrogen(PH) in Soil data))	RMSE((prediction of Ground Water Level (GWL)))	Average of RMSE	Stability Indicator	Comments
IDW	1.05	0.53	2.78	1.5	Low	Poor results on datasets with a lot of variation
Kriging	1.07	0.58	2.30	1.3	Low	Less stable; suitablefor local variances
Natural Neighbor	0.17	0.08	0.36	0.2	High	Ideal for soil data that is geographically linked
Spline	0.92	0.17	0.44	0.51	Moderate	Smooth surface; suitable for continuous data
RBF	1.82	0.50	2.72	1.7	Low	Poor results on datasets with a lot of variation

and facilitates well-informed decision-making by accurately representing soil qualities.

3.2.5. A comparative analysis of the calculation results of different interpolation methods

In this present study, Tables 6 and 7 show the structure for comparing the RMSE of different types of soil data and the validation of the pH soil data with the various interpolation techniques and according to stability and accuracy metrics. This comparative examination can facilitate the effective evaluation of each interpolation method's performance.

3.3. Discussion

The discussion of this paper is divided into three most important parts, which are highlighted in the following:

Part 1. The ideal method of interpolation of the soil data of Jazan City:

When measured values were not available, GIS interpolation techniques were employed in the current analysis to fill in the missing data values at such locations in all geotechnical maps. Spatial interpolation is a popular method for collecting continuous data from discrete locales. Table 2 displays the findings of this comparison analysis of the RMSE values for each interpolation technique. Among the data analyzed were soil pH, GWL, STP, and Jazan spot height. Furthermore, Fig. 4 showed examples of the spatial patterns that each approach expects. The inquiry was completed using the following five interpolation techniques: "IDW, OK, Natural Neighbor, Spline, and RBF." The natural neighbor approach is the best technique for interpolating soil data based on formulas 1 through 8, according to research on geostatistical analysis and prediction using ArcGIS 10.4. Even

Table 7. A comparative analysis of different interpolation methods' computational efficiency and complexity can be added to provide more reference for practical applications.

Interpolation Method	Mean Absolute Error(MAE)	Mean Square Error(MSE)	Root Mean Square Error(RMSE)	Value of R-Square %	Stability Indicator	Comments of pH in Soil Data
Inverse Distance Weight (IDW)	0.44	0.28	0.53	24.2%	Low	Poor results on datasets with a lot of variation
Kriging	0.42	0.28	0.53	8.7%	Low	Poor results on datasets with a lot of variation
Natural Neighbor	0.07	0.01	0.01	98%	High	Ideal in data that is geographically linked
Spline	0.41	0.27	0.52	63.6%	Moderate	Smooth surface; good for continuous data
RBF	0.15	0.05	0.22	92.8%	High	Very good for local variations; less stable

though the statistical analysis of the soil data was gathered from the various altitudes in Jazan City, the second approach, the spline method, produced appropriate interpolation findings. On the other hand, the IDW method produced subpar results. Additionally, Kriging is more efficient and yielded a reasonable soil data interpolation result for the well-distributed level area. These results emphasize how crucial it is to choose an interpolation technique depending on the features of the research region and the distribution of data points. Every sort of soil data is verified, and mapping is carried out to validate the interpolation outcomes.

Additionally, Table 5 and Fig. 14 show the validation sample of interpolation techniques using the pH sample data. The R-Square value of the Natural Neighbor interpolation method is certified to a higher percentage value of 98%.This number, which indicates the natural neighbor, is the most effective method for determining the pH levels in the soil. The neighbor method had the lowest RMSE value error rate in the investigation when compared to other approaches, and a 98% accuracy rate showed that the R-square validation closely matched the actual data.

Comparing the computational complexity and effectiveness of different interpolation approaches may yield further references for practical applications, as Tables 6 and 7 demonstrate. Additionally, the RMSE analysis of the natural neighbor's accuracy and stability indicator revealed a high accuracy and stability indicator score for forecasting the soil data. This table layout makes it easy to compare different interpolation approaches using certain accuracy and stability parameters. You can determine which interpolation method is most effective for your particular dataset and application by entering the actual study results. Soil properties like pH, organic matter, and moisture can vary significantly over short distances. This geographic variety must be taken into account when selecting an interpolation method. The benefits of interpolation covered in this paper spatial data, visualization, assistance for informed decision-making, variety of methodologies, enabling spatial analysis, and completeness of data are reflected in future research objectives. Additionally, the interpolation's drawbacks include insufficient spatial coverage, non-uniform characteristics, sampling techniques, and failure to account for temporal changes.

The availability of a computerized model of Jazan soil attributes, which aids in borehole mapping plotting the locations of boreholes on a GIS map is represented by the future research directions. This spatial representation will make it easier to see how borehole data is distributed across the study area.

Visualization of Soil Classification: overlay the FAO soil classification map to see different soil types and their geographical distribution. Furthermore, by properly resolving these issues, interpolation findings' dependability can be increased. The Jazan region of Saudi Arabia presents a variety of geographic and environmental characteristics that can affect the performance of interpolation methods in ArcGIS 10.4. However, potential prejudices are covered, including bias in spatial dependency, selecting an interpolation method, sensitivity of parameters, boundary restrictions, density of data impact, data accuracy, and quality. This study validates earlier findings and presents an appropriate GIS-based technique for soil interpolation.

Part 2. Laboratory tests of soil borehole data:

To get the most accurate soil data, borehole locations should be dispersed throughout a wide area encircled by challenging terrain, lakes, and forests. But the cost of the lab test and the soil sample is high. Application design, the difficulty of drilling a borehole in specific geological formations, hole diameter, drilling depth, and casing requirements all have a major impact on borehole prices [74].

The optimal distribution of the soil sample holes determined how accurately the natural neighbor interpolated. Drilling holes has a lengthy history. For mining and other purposes, the Chinese used deep borehole drilling as early as the Han Dynasty (202 BC–220 AD). According to British sinologist and historian Michael Loewe [75], drilling sites may go as deep as 600 meters (2000 feet) at that time. However, deep boreholes were not commonly feasible until gasoline and diesel engines were invented.

In Fig. 8, D60 = 16.5, D30 = 3, and D10 = 0.9; from the graph, the soil employed in this study is described as gravelly sand in this location Embankment road in Alrawabi scheme. This result was compared with the Unified Classification System (USCS) used to classify the soil sample; the following can be said: The soil is coarse on a 0.075 mm sieve, retained >50 (99.89 > 50). The soil is gravel on a 4.75 mm sieve; the gravel is retained > 50 (59.78 > 50) and may be seen in the roads and road embankments. This combination of field and laboratory tests provides a comprehensive understanding of soil properties, essential for making informed decisions regarding foundation design, construction methods, and environmental management. The sample findings of the relative compaction and direct shear tests of the borehole at the Alshawajrah scheme location are displayed in Figs. 9 and 10. In addition, the research area's average dry density is 2.11 g/cm^3. Ground water levels can be as high as six meters above sea level in the western part of the country, and as low as 0.5 meters below sea level in areas that are isolated from the sea.

Part 3. Mapping of Geotechnical contours and FAO soil map of the study area:

Combining the FAO soil classification system with geotechnical borehole data in a Geographic Information System (GIS) framework can greatly enhance the connection between the FAO soil classification system and geotechnical borehole data using GIS soil analysis. Here's a step-by-step guide to completing this integration.

To ensure accuracy and consistency, the study's geographical data, including coordinates for maps and analysis, are based on the World Geodetic System 1984 and projected using the Universal Transverse Mercator system (UTM), Zone 38N. Ain El Abd UTM, Zone 38N is the precise datum for the study region.

Soil maps were created by spatially interpolating the geotechnical data gathered from 23 locations. The variance in soil type, SPT, pH in soil, chlorides (CL) in groundwater, and plasticity Index were displayed on the made maps. Fig. 11 depicts geotechnical maps of Jazan City that highlight the soil's chemical, physical, and mechanical characteristics to pinpoint locations with different soil environmental conditions, physical characteristics, and mechanical characteristics. This map can be used to introduce the results in various areas of the study area and display the data at the precise locations of the boreholes. Additionally, plotting and updating the digital geotechnical map anywhere in the study region is simple.

Fig. 12 presents a detailed contour map of Jazan City with a 2-meter interval, created using a computer model. This map is valuable for planning and executing construction and maintenance projects. Fig. 11 shows the soil classification for Jazan Municipality based on the FAO system. Most of Jazan City (56.12%) consists of saline soil, particularly along the Red Sea coast. Barren soil also makes up a significant portion, accounting for 28.66% of the municipality. The remaining 15.22% is classified as "remaining land," though the specific soil type is not identified in this classification. The study combined laboratory testing of borehole data with the FAO soil classification and created a robust model for predicting soil properties in Jazan City. In the present study, this digital model of soil allowed for the estimation of soil characteristics at any location within the study area using the techniques of ArcGIS 10.4 software. This model saves time and cost for the stockholders and investment companies and plays an important role in KSA's Vision 2030.

4. Conclusion

The study concluded that the researchers can spot patterns, support or refute accepted wisdom, and offer insightful information about the effectiveness of various interpolation techniques in difficult situations by placing the study's conclusions within the larger framework of the body of existing literature. Overall, a 98% accuracy rate showed that the Natural Neighbor approach yields very reliable findings that are useful in a variety of real-world situations. A solid basis for in-depth soil research is provided by integrating geotechnical borehole data with the FAO soil classification system using GIS. By encouraging a deeper understanding of soil qualities and their implications for a variety of uses, such as agriculture, building, and environmental management, this method improves decision-making. The availability of employing GIS in soil interpolation data was demonstrated by the outcomes of the five interpolation techniques. Although the minimum RMSE value displayed in Table 5 and Fig. 12 is reached, the cross-validation analysis of the results confirmed that the natural neighbor technique of interpolation is appropriate for the interpolation of soil data. Moreover,

this method is known for preserving data integrity and accurately representing original values in the created surface. Ensure a smooth and continuous surface with smooth transitions, including slopes. Effectively manage unstructured data in geophysical applications. These advantages make the natural neighbor method valuable for generating visually appealing and reliable surfaces from scattered data points. This study explores the application of natural neighbor interpolation for estimating geotechnical data. This approach is precious when data from laboratory tests and borehole observations are available. The Natural Neighbor method, combined with techniques like spline, kriging, and radial basis functions, can effectively address challenges such as determining the most suitable training datasets or parameter sets [76]. The analysis's findings and the results themselves can be applied to any usage of interpolation techniques, not just the creation of soil data but also other uses such as hydrology and environmental studies. This study utilizes the natural-neighbor method to estimate the missing data on soil bending in a real borehole data collection. Moreover, this interpolation method is a quick, durable, and accurate method for reassembling a surface from sample points that are erratically dispersed. The study is concluded by combining GIS, FAO's UNESCO soil classification system, sample soil tests, and interpolating soil value using the Natural Neighbor (NaN) method. This procedure is a very robust digital soil model that was built.

The results of this study demonstrated the wide range of applications of interpolation techniques, especially the natural neighbor method, which may be applied to a variety of tasks outside of the production of soil data. The investigation showed that the method was very well suited for estimating missing soil data from actual drilling samples due to its speed, accuracy with uneven data, and robustness to outliers. Furthermore, kriging regularly performed more accurately than IDW in this study. Its statistical method better captured the geographical relationships in the data, as seen by the lower error metrics it attained. Furthermore, the majority of Jazan's soil data was predicted with a reasonable inaccuracy, according to this analysis. This depended on how important it was to have accurate soil test findings.

5. Suggestion for future work

Authors suggest some creative ideas for future research on soil property interpolation using Kriging, Natural Neighbor, Inverse Distance Weighting (IDW), and Radial Basis Functions (RBF) methods:

a. The Natural Neighbor algorithm may efficiently handle irregularly dispersed sample points by concentrating on the local neighborhood without making rigid assumptions about global trends, providing precise and context-sensitive predictions, this method is suitable method for interpolating the soil properties.
b. In experimental Approach, to improve predictive skills, look at combining less popular interpolation techniques with IDW, Kriging, and others, such Bayesian methods.
c. To enhance model accuracy and supplement conventional soil property datasets, incorporate a variety of data types, such as satellite photography and remote sensing data.
d. To improve the datasets accessible for interpolation techniques, establish platforms for cooperative data sharing between researchers and practitioners.
e. Concentrate on creating open-source tools that give practitioners more functionality and freedom to apply different interpolation approaches.

Conflict of interest

The authors declare no conflict of interest.

Authors contributions

Dr. Ehab Sabi: Writing the introduction, revising the English language of the paper parts, and the methodology.

Dr. Mahmoud Abdelrahim Abdelgiom: Writing the draft of the paper, Methodology, digitization of base map, building geo-database, carrying out the spatial analysis, results and discussion, and conclusions.

Data availability

The authors confirm that the data supporting the findings of this study are available within the article.

References

1. Smolik M, Skala V. Large scattered data interpolation with radial basis functions and space subdivision. Integr Comput Aided Eng. 2017;25:49–62. https://doi.org/10.3233/ICA-170556.
2. Skala V, Karim SAA, Cervenka M. Finding points of importance for radial basis function approximation of large scattered data. In: 2020. p. 239–250. https://doi.org/10.1007/978-3-030-50433-5_19.

3. Hoschek J, Lasser D. Grundlagen der geometrischen Datenverarbeitung. Stuttgart: B.G. Teubner. 1992.
4. Keller W, Borkowski A. Thin plate spline interpolation. Journal of Geodesy. 2019;93:1251–1269.
5. Chan RTL, Hubbert S. A numerical study of radial basis function based methods for options pricing under the one dimension jump-diffusion model. 2010. Available from: http://arxiv.org/abs/1011.5650.
6. Dhowian AW. Laboratory simulation of field preloading on Jizan sabkha soil. J King Saud Univ - Eng Sci. 2017;29:12–21. https://doi.org/10.1016/j.jksues.2015.09.001.
7. Mahfouz AH. Geotechnical study of soil foundation and design of PCC piles to improve its bearing capacity, a case study: Ramps of Dawar El-Tawheed bridge in Jizan City - kSA. Int J For Soil Eros. 2016;4.
8. Arifuzzaman M, Habib MA, Al-Turki MK, Khan MI, Ali MM. Improvement and characterization of sabkha soil of Saudi Arabia: A review. J Teknol. 2016;78. https://doi.org/10.11113/jt.v78.5352.
9. Al-Homidy AA, Dahim MH, Abd El Aal AK. Improvement of geotechnical properties of sabkha soil utilizing cement kiln dust. J Rock Mech Geotech Eng. 2017;9:749–760. https://doi.org/10.1016/j.jrmge.2016.11.012.
10. Al-Boghdady AA, Hassanein KMA. Chemical analysis and environmental impact of heavy metals in soil of Wadi Jazan area, southwest of Saudi Arabia. Applied Ecology & Environmental Research. 2019;17:3.
11. Zanaty N, Mansour K, Fathi H. Satellite-based assessment of the anthropogenic impacts on environmental sustainability in Jazan region, Red Sea, Egypt. J Remote Sens Sp Sci. 2023;26:117–127. https://doi.org/10.1016/j.ejrs.2022.12.002.
12. Oliveira Junior JC, Furquim SAC, Nascimento AF, Beirigo RM, Barbiero L, Valles V, Couto EG, Vidal-Torrado P. Salt-affected soils on elevated landforms of an alluvial megafan, northern Pantanal, Brazil. Catena. 2019;172:819–830. https://doi.org/10.1016/j.catena.2018.09.041.
13. Alhumimidi MS. Geotechnical assessment of near-surface sediments and their hazardous impact: Case study of Jizan city, southwestern Saudi Arabia. J King Saud Univ - Sci. 2020;32:2195–2201. https://doi.org/10.1016/j.jksus.2020.02.031.
14. Pankratz HG, Sultan M, Abdelmohsen K, Sauck WA, Alsefry S, Alharbi H, Emil MK, Gebremichael E, Asaeidi A, Alshehri F, Hashim HI, Al-Shamrani HA, El-Sahly M. Use of Geophysical and radar interferometric techniques to monitor land deformation associated with the Jazan salt diapir, Jazan city, Saudi Arabia. Surv Geophys. 2021;42:177–200. https://doi.org/10.1007/s10712-020-09623-3.
15. Fatani MN, Khan AM. Foundations on salt bearing soils of Jizan. In: Third Int Conf Case Hist Geotech Eng. 1993;79–83.
16. Arrofi D, Abu-Mahfouz IS. Delineating reservoir area using geologic remote sensing analysis, Jizan (Saudi Arabia): Implications for geothermal exploration. Geoenergy Sci Eng. 2023;228:212033. https://doi.org/10.1016/j.geoen.2023.212033.
17. Nachtergaele FO. Classification systems | FAO. In: Encycl Soils Environ. Elsevier. 2005:216–222. https://doi.org/10.1016/B0-12-348530-4/00549-X.
18. Nachtergaele FO. The development of the FAO legend 1960 –. (n.d.).
19. Martinez-Villegas N. An overview of different soil classification systems una breve revisión acerca de los sistemas de clasificación de suelos utilizados en México. Terra Latinoam. 2007;25:357–362.
20. Fischer J. Dreiersysteme, Vierersysteme, Fünfersysteme und Größenlehre. Arch Für Elektrotechnik. 1960;45:225–232. https://doi.org/10.1007/BF01574372.
21. Selcer P. Fabricating unity: The FAO-UNESCO Soil Map of the World. Hist Soc Res. 2015;40:174–201. https://doi.org/10.12759/hsr.40.2015.2.174-201.
22. SpatialZhao Y-C, Shi X-Z. Prediction and Uncertainty Assessment of Soil Organic Carbon in Hebei Province, China. 2010. https://doi.org/10.1007/978-90-481-8863-5_19.
23. Islam S, Alam AKMR. Performance evaluation of FAO Penman-Monteith and best alternative models for estimating reference evapotranspiration in Bangladesh. Heliyon. 2021;7:e07487. https://doi.org/10.1016/j.heliyon.2021.e07487.
24. Arshad M, Eid EM, El-sheikh K. In 13 th Annual Research Day 13 th Annual Research Day. 2018:1–6.
25. Hewitt PL, Flett GL. Dimensions of perfectionism, daily stress, and depression: A test of the specific vulnerability hypothesis. J Abnorm Psychol. 1993;102:58–65. https://doi.org/10.1037/0021-843X.102.1.58.
26. McBratney AB, Odeh IOA. Application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions. Geoderma. 1997;77:85–113. https://doi.org/10.1016/S0016-7061(97)00017-7.
27. Nozari S, Pahlavan-Rad MR, Brungard C, Heung B, Borůvka L. Digital soil mapping using machine learning-based methods to predict soil organic carbon in two different districts in the Czech Republic. Soil Water Res. 2024;19:32–49. https://doi.org/10.17221/119/2023-SWR.
28. Mondal BP, Sahoo RN, Das B, Chakraborty D, Kumar N, Raj GB. Digital soil mapping: concepts, methods, and applications - remote sensing and GIS perspectives. In: Remote Sens Soils. Elsevier. 2024;195–212. https://doi.org/10.1016/B978-0-443-18773-5.00018-1.
29. Abdulraheem MI, Zhang W, Li S, Moshayedi AJ, Farooque AA, Hu J. Advancement of remote sensing for soil measurements and applications: A comprehensive review. Sustain. 2023;15. https://doi.org/10.3390/su152115444.
30. Chen S, Arrouays D, Mulder VL, Poggio L, Minasny B, Roudier P, Libohova Z, Lagacherie P, Shi Z, Hannam J, Meersmans J, Richer-de-Forges AC, Walter C. Digital mapping of GlobalSoilMap soil properties at a broad scale: A review. Geoderma. 2022;409:115567. https://doi.org/10.1016/j.geoderma.2021.115567.
31. Nelson L. Predictive digital soil mapping white paper prepared for LWS 548. 2021.
32. Fradelos E, Papathanasiou I, Mitsi D, Tsaras K, Kleisiaris C, Kourkouta L. Health based geographic information systems (GIS) and their applications. Acta Inform Medica. 2014;22:402. https://doi.org/10.5455/aim.2014.22.402-405.
33. Cone J. Principles of geographical information systems by peter A. N Z Geog. 1998;54:56–57. https://doi.org/10.1111/j.1745-7939.1998.tb02089.x.
34. Vereecken H, Amelung W, Bauke SL, Bogena H, Brüggemann N, Montzka C, Vanderborght J, Bechtold M, Blöschl G, Carminati M, Javaux M, Konings AG, Kusche J, Neuweiler I, Or D, Steele-Dunne S, Verhoef A, Young M, Zhang Y. Soil hydrology in the Earth system. Nat Rev Earth Environ. 2022;3:573–587. https://doi.org/10.1038/s43017-022-00324-6.
35. Keller AA, Garner K, Rao N, Enipping E, Thomas J. Hydrological models for climate-based assessments at the watershed.
36. Barton N. Barton-bandis criterion. 2020. https://doi.org/10.1007/978-3-319-12127-7_25-1.

37. Fernlund JMR. The effect of particle form on sieve analysis: a test by image analysis. Eng Geol. 1998;50:111–124. https://doi.org/10.1016/S0013-7952(98)00004-0.
38. MacDonald NR, Packulak TRM, Day JJ. A critical review of current states of practice in direct shear testing of unfilled rock fractures focused on multi-stage and boundary conditions. Geosci. 2023;13. https://doi.org/10.3390/geosciences13060172.
39. Zou L, Cvetkovic V. A new approach for predicting direct shear tests on rock fractures. Int J Rock Mech Min Sci. 2023;168:105408. https://doi.org/10.1016/j.ijrmms.2023.105408.
40. Bandis S, Lumsden AC, Barton NR. Experimental studies of scale effects on the shear behaviour of rock joints. Int J Rock Mech Min Sci Geomech Abstr. 1981;18:1–21. https://doi.org/10.1016/0148-9062(81)90262-X.
41. Bandis SC, Lumsden AC, Barton NR. Fundamentals of rock joint deformation. Int J Rock Mech Min Sci Geomech Abstr. 1983;20:249–268. https://doi.org/10.1016/0148-9062(83)90595-8.
42. Muralha J, Grasselli G, Tatone B, Blümel M, Chryssanthakis P, Yujing J. ISRM suggested method for laboratory determination of the shear strength of rock joints: Revised version. Rock Mech Rock Eng. 2014;47:291–302. https://doi.org/10.1007/s00603-013-0519-z.
43. Johansson F. Influence of scale and matedness on the peak shear strength of fresh, unweathered rock joints. Int J Rock Mech Min Sci. 2016;82:36–47. https://doi.org/10.1016/j.ijrmms.2015.11.010.
44. Li B, Ye X, Dou Z, Zhao Z, Li Y, Yang Q. Shear strength of rock fractures under dry, surface wet and saturated conditions. Rock Mech Rock Eng. 2020;53:2605–2622. https://doi.org/10.1007/s00603-020-02061-y.
45. Robinson DA, Campbell CS, Hopmans JW, Hornbuckle BK, Jones SB, Knight R, Ogden F, Selker J, Wendroth O. Soil moisture measurement for ecological and hydrological watershed-scale observatories: A review. Vadose Zo J. 2008;7:358–389. https://doi.org/10.2136/vzj2007.0143.
46. Janjić J, Tadić L. Fields of application of SWAT hydrological model—A review. Earth. 2023;4:331–344. https://doi.org/10.3390/earth4020018.
47. Alcaras E. Comparison of different interpolation methods for DEM production. Int J Adv Trends Comput Sci Eng. 2019;6:1654–1659. https://doi.org/10.30534/ijatcse/2019/91842019.
48. Matheron G. Principles of geostatistics. Econ Geol. 1963;58:1246–1266. https://doi.org/10.2113/gsecongeo.58.8.1246.
49. Xie Y, Chen TB, Lei M, Yang J, Guo QJ, Song B, Zhou XY. Spatial distribution of soil heavy metal pollution estimated by different interpolation methods: Accuracy and uncertainty analysis. Chemosphere. 2011;82:468–476. https://doi.org/10.1016/j.chemosphere.2010.09.053.
50. İmamoğlu MZ, Sertel E. Analysis of different interpolation methods for soil moisture mapping using field measurements and remotely sensed data. Int J Environ Geoinformatics. 2016;3:11–25. https://doi.org/10.30897/ijegeo.306477.
51. Harder C, Brown C. Mapping the Internet of Things: Harnessing the worldwide web of sensors. ArcGIS B. 10 Big Ideas about Appl Sci Where. 2017;1:133–146. www.TheArcGISBook.com.
52. Ajaj QM, Shareef MA, Hassan ND, Hasan SF, Noori AM. GIS based spatial modeling to mapping and estimation relative risk of different diseases using inverse distance weighting (IDW) interpolation algorithm and evidential belief function (EBF) (Case study: Minor Part of Kirkuk City, Iraq). Int J Eng Technol. 2018;7:185–191. https://doi.org/10.14419/ijet.v7i4.37.24098.
53. Bishop TFA, McBratney AB, Whelan BM. Measuring the quality of digital soil maps using information criteria. Geoderma. 2001;103:95–111. https://doi.org/10.1016/S0016-7061(01)00071-4.
54. Scull P, Franklin J, Chadwick OA. The application of classification tree analysis to soil type prediction in a desert landscape. Ecol Modell. 181 (2005) 1–15. https://doi.org/10.1016/j.ecolmodel.2004.06.036.
55. Arslan H. Estimation of spatial distrubition of groundwater level and risky areas of seawater intrusion on the coastal region in Çarşamba Plain, Turkey, using different interpolation methods. Environ Monit Assess. 2014;186:5123–5134. https://doi.org/10.1007/s10661-014-3764-z.
56. Musashi JP, Pramoedyo H, Fitriani R. Comparison of inverse distance weighted and natural neighbor interpolation method at air temperature data in malang region. CAUCHY J Mat Murni Dan Apl. 2018;5:48–54. https://doi.org/10.18860/ca.v5i2.4722.
57. Bardan V. Trace interpolation in seismic data processing *. Geophys Prospect. 1987;35:343–358. https://doi.org/10.1111/j.1365-2478.1987.tb00822.x.
58. Xhafa S, Kosovrasti A. Geographic information systems (GIS) in urban planning. Eur J Interdiscip Stud. 2015;1:85. https://doi.org/10.26417/ejis.v1i1.p85-92.
59. Fan Q, Efrat A, Koltun V, Krishnan S, Venkatasubramanian S. Hardware-assisted natural neighbor interpolation. In Proc Seventh Work Algorithm Eng Exp Second Work Anal Algorithms Comb. 2005:111–120.
60. Cueto E, Sukumar N, Calvo B, Martínez MA, Cegoñino J, Doblaré M. Overview and recent advances in natural neighbour galerkin methods. Arch Comput Methods Eng. 2003;10:307–384. https://doi.org/10.1007/BF02736253.
61. Wang H, Qin Q-H. Basics of fundamental solutions and radial basis functions. In: Methods Fundam Solut Solid Mech. Elsevier. 2019;91–105. https://doi.org/10.1016/B978-0-12-818283-3.00003-8.
62. Chen H, Bakshi BR. Linear approaches for nonlinear modeling. In: Compr Chemom. Elsevier. 2009;453–462. https://doi.org/10.1016/B978-044452701-1.00060-0.
63. Ogniewski J. Cubic spline interpolation in real-time applications using three control points. Comput Sci Res Notes. 2019;2901:1–10. https://doi.org/10.24132/CSRN.2019.2901.1.1.
64. Franke R. Smooth interpolation of scattered data by local thin plate splines. Comput Math with Appl. 1982;8:273–281. https://doi.org/10.1016/0898-1221(82)90009-8.
65. Wise S. Cross-validation as a means of investigating DEM interpolation error. Computers & Geosciences. 2011;37.8:978–991.
66. Joseph J, *et al.* Application of validation data for assessing spatial interpolation methods for 8-h ozone or other sparsely monitored constituents. Environmental Pollution. 2013;178:411–418.
67. Lopez-Granados F, Jurado-Exposito M, Pena- Barragan JM, Garcia-Torres L. Using geostatistical and remote sensing approaches for mapping soil properties. European Journal of Agronomy. 2005;23(3):279–289. doi:10.1016/j.eja.2004.12.003
68. Mahfouz AH, Abdelgiom MA. Geotechnical investigation and mapping of Jazan city, KSA USING GIS of Geological and Geophysical. 2020. https://doi.org/10.24941/ijcr.39562.09.2020.
69. Abdelgiom MA, Hassan M, Ali SEE. A biodiesel power plant, location, production and electricity demand of Jazan Municipalities, Saudi Arabia using GIS. Journal of Umm Al-Qura University for Engineering and Architecture. 2024:1–22.

70. Abdelgiom MA, Ali S. optimization area, simulation, and production of biodiesel fuel of jatropha in jazan municipalities, ksa. Jouf University Science and Engineering Journal (JUSEJ). 2022;9(2):31–49.
71. VOLTZ M, WEBSTER R. A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. J Soil Sci. 1990;41:473–490. https://doi.org/10.1111/j.1365-2389.1990.tb00080.x
72. Wise S. Cross-validation as a means of investigating DEM interpolation error. Computers & Geosciences. 2011;37(8):978–991.
73. Otto P, Fassò A, Maranzano P. A review of regularised estimation methods and cross-validation in spatiotemporal statistics. Statistic Surveys. 2024;18:299–340.
74. Bobrowsky PT, Marker B, eds. Encyclopedia of Engineering Geology. Cham: Springer International Publishing. 2018. https://doi.org/10.1007/978-3-319-73568-9.
75. Gaaloul N, Eslamian S, Ostad-Ali-Askari K. Boreholes. Encycl Earth Sci Ser. 2020:1–6. https://doi.org/10.1007/978-3-319-12127-7_31-1
76. Juda, Renard P, Straubhaar J. A framework for the cross-validation of categorical geostatistical simulations. Earth Sp Sci. 2020;7:1–17. https://doi.org/10.1029/2020EA001152.