ORIGINAL RESEARCH

Open Access

Investigation of distinctive physico-chemical soil correlations for Kirkuk city using spatial analysis technique incorporated with statistical modeling

Aram Mohammed Raheem* and Najat Qader Omar

*Correspondence:
engaram@yahoo.com
Civil Engineering
Department, University
of Kirkuk, Kirkuk, Iraq

Abstract

In the sense of challenging economic situation, it is difficult to perform laboratory tests in a whole intended area to define the soil characteristics for whatever task or situation within an entire city. Hence, this study has used the spatial analysis through applying the inverse distance weighted (IDW) technique to integrate with the available data for 56 different soil samples collected from various depths of Kirkuk city. The studied characteristics were mainly classified to physical soil properties included gravel, sand, silt, and clay contents with chemical soil properties included organic and gypsum contents, pH content, sulfur trioxide, and total dissolved solids (TDS) content. Moreover, statistical investigations such as physico-chemical correlation, linear single and linear multi regression models have been applied. The results of this study have focused on constructing thematic digital maps for visualizing different physical and chemical characteristics of Kirkuk soil. In statistical analysis, negative linear correlations are obtained from the relation between the values of the pH, gypsum, organic, and SO₃ contents with the increase of the gravel content, whereas a positive linear correlation is attained for the values of TDS content with the increase of the gravel content. Furthermore, the proposed linear multi regression model predicts the pH values reasonably where most of the values were close to the equality line. The obtained digital maps accompanied with the statistical models will provide comprehensive spatial, agricultural, chemical and mechanical visualizations of the nature and morphology of the soils in Kirkuk city.

Keywords: Digital maps, Soil characteristics, Inverse distance weighted, Physico-chemical correlation, Single and multi linear regression models, PH, Sulfur trioxide, Total dissolved solids

Introduction

Producing a digital soil map, which is a fundamental constituent used for geovisualization and geospatial mapping, has an extraordinary effect on the research in terms of production time for districts and whole countries. Maps can be used to characterize the lands and assess their agricultural capability depending on the strategies of ground treatment. Classifying the soil zones in any city is necessary for determining

the landfill areas to avoid the population extents, prevent contamination of any nearby drinking and irrigation groundwater, and prohibit the threatening of the human health [4]. To prepare a simple digital soil map, there is a need for more than millions of pixels or cells. A GIS or “Geographical information system” is an assembly of automated functions and records that are coupled simultaneously for the principle of retrieving, storing, managing, and analyzing information [12]. The GIS procedure can represent maps coinciding different aspects such as road, industry, prospective pollution sources, crop supplies and degradation distributions [20]. Currently, the soil scientists provide a huge database of boreholes data such as borehole logs, borehole plans, borehole tests, surface and subsurface soil properties. Most of these soil data are kept as hard copies in boxes, drawers, and file cabinets. The processes of finding such information or soil data are very time consuming. Always, it is very challenging and costly to identify the soil properties for any agricultural, civil or geotechnical engineering purposes. The GIS applications were introduced as an advanced technology to classify the potential of geological hazards, arranging and tracking fieldwork, generating maps and images, and expanding communications [1, 25]. The land use regression model incorporated with geographical information analysis was implemented for measuring different weather seasons in Iran [22].

It is recognized that the soil plays a vital role in miscellaneous activities all over the world. It is proved that the soil has a major role on the biological environment accompanied with the human health since it is a chemical source for plants [36]. The process of sorption of insecticides in any soil relies on the several significant characteristics of soil such as pH, organic content and clay content [19]. Progressively, consideration has developed the focus on the character and function of large voids in the soil depending on both chemical and water movements. The hypothesis of avoid movement through macro pores is an ideal path and the reality reveals that such voids allow fast travel of both chemicals and water [7]. The controlling factor during the process of infiltration is the soil permeability rather than the effect of capillary forces. In addition, the highly permeable soils represented by gravels and sands can expose the unconfined groundwater aquifer to contaminations from the surface sources [8, 33]. The ordinary clay minerals such as kaolinite, illite and smectite have great impacts on several soil characteristics represented by organic substance, nature, cation replacement capability and dispersity [31]. Clay minerals are increasingly known as an essential portion in soils affecting the holding capability and bioavailability and mobility of different heavy metals in certain environments [34]. However, clayey soils can be considered as fundamental barriers against liquid contaminants. Thus, having information on clayey soil distribution provides a thorough understanding in opposition to contamination spreading [5]. The contaminated soil can cause miscellaneous engineering problems represented by differential settlement, damage to structures due to soil leachate, cracking and inclination in the ground surface [15]. In addition, the locations and volumes of contaminated soils are necessary for remediation process. However, problematic engineering soils could be fertilized lands for agricultural uses. Thus, collecting various available agricultural, geotechnical, chemical and physical soil characteristics in several digital maps are essential specifically in Kirkuk city where such maps are necessary to identify the right places for various applications.

The physico-chemical soil analysis has been applied in the agricultural process in several conducts. Initially, the physico-chemical soil characterization has been used for determining the biological properties of soils in different agricultural locations in Italy [21]. The physico-chemical analysis for spiroxamine hydrolysis has been inspected under various pH values to enhance the knowledge about the effect of water and its degradation impacts [9]. Moreover, the physico-chemical investigations for both wastewater and ground water have been evaluated to check the quality of such water for drinking and agricultural purposes where such quantifications included measuring pH, temperature and conductivity [6, 10]. On the other hand, the sorption of ten agricultural soils collected from various places has been tested where pH, organic content, electrical conductivity, clay content, gravel content, and silt content have been measured [37]. Furthermore, the physico-chemical characteristics of both corn and rice with the advancement of thin flaked rice have been inspected [3, 18]. Likewise, the physico-chemical features have been applied in different agricultural products such pomegranate [16], the powder of papaya leave [32], and Indian wheat [30]. Thus, it is essential to identify the distribution of any soil type to characterize the compatibility of the physico-chemical properties of the soil and the planted agricultural products.

Investigations using spatial analysis in the presence of heterogeneity in the soil characteristics play essential functions in improving soil mapping and surveys, developing powerful information for precise fertilization and enhancing the nutrient management in the soils [23]. The spatial adaptability for soil physico-chemical classification has been used in order to predict various Iranian maps with distinctive soil properties [24]. In addition, it was proved that the phosphorous variations in soils require detailed temporal and spatial investigations for better understanding of using phosphoric management systems [11].

Based on ESRI specifications [13], GIS applications have been customized for natural landslide searching, habitat mapping, agricultural reclamation and air pollutant monitoring. In addition, an expert based scheme was developed to provide two-dimensional visualizations for the ground circumstances [2]. An integrated geographical information system and data based Geo-statistical classification has been advanced to consider protecting the subsurface records. The Arc GIS system can connect the soil formation with their properties at any spot within the subject region.

The main objective of this study is to provide a full database for the available soil information obtained from 56 different soil samples for various depths of Kirkuk city using spatial analysis technique tools with inverse distance weighted method. The specific objectives can be summarized as digital maps for visualizing the physical and chemical soil characteristics of Kirkuk city, including the percentages of gravel, sand, silt, clay, pH, gypsum, total dissolved solids, organic and sulfur trioxide distributions. In addition, predicting the physico-chemical behavior of Kirkuk soils using single and multi linear regression models. Such obtained digital maps and statistical models can provide physical, agricultural, chemical and mechanical detailed visualizations about the behavior and features of the soils in Kirkuk city.

The study area

Kirkuk city is considered to be one of the Iraq's governorates where its heritage belonged to 5000 years ago [28]. Kirkuk is located on "latitude $35^{\circ} 28' 5''$ north" and "longitude $44^{\circ} 23' 31''$ east", and about 350 m above the sea level based on "UTM" reference project zone 38 N. The Kirkuk area's is about "9679 km²" where this area represents 2.2% of the entire area of Iraq [29]. Kirkuk contains 13 governmental offices forming four districts where these districts are Kirkuk, Al-Hawiga, Daquq, and Maxmur shown in Fig. 1.

Methodology

Assembling geographic information is considered to have the greatest cost and most time-consuming process of developing a database for GIS application. The process of collecting and rearranging data in GIS is split into two groups of processes: assembling and integrating information. Data collection contains entering new data into GIS, whereas data integration includes converting data and preserving its integrity. If the spatial distribution of any studied characteristics is random, global autocorrelation analysis is permitted to be used in GIS [27]. Moreover, the spatial analysis can be used for standardized spatial weights matrix [14]. It is meaningful to mention that the progress and use of spatial method has speeded up the development of landform arrangement [35].

The main source of the data is the Kirkuk constructional laboratory for material testing. In addition, fieldwork observations for site investigation have been used in which soil data for 56 different samples from distinctive depths with physical and chemical properties including latitude and longitude Global position system data by (GS 20 Leica) for each point have been collected. Examples of geographical, physical and chemical properties of soil samples from Kirkuk city have been summarized in Table 1.

Every source data were typed in the excel sheet with latitude and longitude for each investigated point as well as for the laboratory data. The thematic map or the shape file was created depending on the tools in the Arc Map. The interpolation maps were created by interpolating the surface from the points using an inverse distance weighted (IDW) technique as shown in Fig. 2.

The procedure of inverse distance weight is simple, wide, and one of the utmost vital methods for interpolation. The inverse distance weighted algorithm is assumed that the average values of the neighborhood points cannot be used to estimate the value for the sampled point, however; the accuracy of the value can be increased using the average values for the non-surrounding points.

$$Y_o = \sum_{i=1}^N \frac{y_i d_1^{-n}}{d_1^{-n}} \quad (1)$$

Y_o = Estimation value of variable z in point I ; y_i = Sample value in point I ; d_1 = Euclidian distance between observe sample point to estimated point; N = A coefficient of weigh based on a distance; n = inverse distance weighing power.

The used interpolation method is based on a distance-weighted average method or a potential model method. In all these models, the exponent related to the distance is defined by the user. In the created digital maps, a surface calculated by IDW is a function of both sample value and strategy of neighborhood search. Both maximum and minimum values of any surface can be exactly interpreted using IDW. The developed digital maps are highly affected by any clustering or outliers. IDW supposes that the modeled phenomenon is determined by local variation and it can be captured using defined acceptable neighborhood search.

Utilizing parameters obtained from part of the input data, the outliers have been evaluated where the IDW interpolation clearly renders the inferences that are connected to each other in which they are more comparable than those that are quite away. IDW applies the measured data accompanying the simulation place to estimate a value for

Table 1 Geographical, physical and chemical properties of examples of studied Kirkuk soil samples

FID	North	East	Depth (m)	Gravel (%)	Sand (%)	Silt (%)	Clay (%)	pH	Gypsum (%)	TDS (%)	Organic (%)	SO ₃ (%)
0	3,923,377.94	444,192.197	1	31	35	26	8	8	1.38	4.42	0.32	0.64
1	3,923,386.27	444,189.98	1	53	23	18	6	8	1.38	1.96	0.054	0.64
2	3,929,536.3	442,401.13	1	68	20	6	6	8.62	1.2	0.22	0.16	0.38
3	3,929,541.1	442,400.09	1.5	52	15	22	11	8.82	0.15	3.4	0.36	0.4
4	3,926,057.9	442,431.98	1.5	0	8	72	20	8	1.36	0.11	0.25	0.05
5	3,919,590.35	445,747.869	1	2	11	47	40	9	0.14	2.3	0.29	0.065
6	3,919,584.76	445,755.401	1.5	3	14	57	26	9	0.47	0.82	0.097	0.22
7	3,917,744.09	442,661.82	2.5	0	12	45	43	9	0.795	0.21	0.15	0.053
8	3,917,729.03	442,655.165	3	0	7	50	43	9	0.177	0.47	0.29	0.082
9	3,932,270.6	442,220.16	1.5	0	12	46	42	8.82	0.49	0.69	0.033	0.23
10	3,932,291	442,221.14	1.5	0	10	60	30	8	3.83	1.6	0.065	1.78
11	3,917,799.91	442,604.916	2.5	0	10	55	35	9	0.17	0.854	0.879	0.074

every unknown point. The calculated values nearest to the forecast place have larger implications than those further out on the expected value. IDW maintains that each determined point has a local impact that declines with time. It offers greater weights to places nearest to the forecast position, and as a function of distance, the weights decrease, thus the name of inverse distance weighted.

Results

Physical soil distribution

The gravel distribution of Kirkuk map is analyzed using the IDW interpolation method. The studied area can be classified into five main classes in terms of gravel content, which are very low, low, moderate, high, and very high or numerically as demonstrated in Fig. 3a. The first class starts from 0.007 to 9.54%, whereas the range of the fifth class is 32.07 to 75.21%, which represents the highest stable area identified by the green color and distributes in the center of the studied area. The second class is starting from 9.54 to 17.71%, which is the light brown color and distributes in the eastern and northeastern parts of the studied area. The third class is starting from 17.71 to 25.67% that illustrates the moderate class represented by light yellow colors, and distributes in the central, eastern, western south, and northern parts of the studied area. It is clearly predicted that the gravel distribution has the highest percentage in the middle and western zones of Kirkuk city with the lowest percentage in the

northern and middle southern zones. It should be noted that areas with high gravel content own several characteristics such as high bearing capacity, high permeability, high water and chemical infiltration capability. In addition, the high gravel content areas are more susceptible to the influence of soil boiling under the impact of any expected earthquakes.

Similarly, the sand distribution of Kirkuk map is analyzed using Arc GIS technique as shown in Fig. 3b. The sand and gravel distribution has almost similar distribution on Kirkuk map where the highest portion of the sand was also in the middle and western zones of Kirkuk city with the lowest percentage in the northern and middle southern zones. The ranges of the sand content start from 1.20 to 15.01% for very low zone, 15.01 to 20.6% for low zone, 20.6 to 26.78% for moderate zone, 26.78 to

32.71% for high zone, and 32.71 to 64.03% for the very high zone. In a similar manner, the high sand content areas are considered as coarse grain zones with relatively high bearing capacity, high permeability, high water and chemical infiltration potentials. Moreover, such high sand area zones are exposed to the impact of soil boiling under the effect of any anticipated earthquakes. Furthermore, the zones of high gravel and sand contents are not suitable for agricultural purposes since the water retain capability for such soils are low.

In addition, the Arc GIS technique is used to check down the silt distribution of the Kirkuk map as shown in Fig. 3c. It is obviously indicated that the silt distribution has the highest percentage in the center of the northern zones of Kirkuk city with the lowest percentage in the middle central to middle eastern zones. The ranges of the silt content start from 1.46 to 24.92% for very low zone, 24.92 to 32.81% for low zone, 32.81 to 40.38% for moderate zone, 40.38 to 47.57% for high zone, and 47.57 to 73.73% for the very high zone. It is worthwhile to mention that the zones with high silt contents are the most appropriate areas for agricultural purposes since such soils can retain the water and necessary chemical additives properly.

The clay distribution of Kirkuk map is examined using Arc GIS technique as demonstrated in Fig. 3d. It is noticeably appeared that the clay distribution has the highest share in the middle southern and upper northern zones of Kirkuk city with the lowest percentage in the western zone of the metropolis. The ranges of the clay content start from 0.02 to 9.2% for very low zone, 9.2 to 14.91% for low zone, 14.91 to 21.35% for moderate zone, 21.35 to 28.39% for high zone, and 28.39 to 52.93% for the very high zone. It should be observed that the areas with high clay contents maintain specific features such as low permeability, low water and chemical infiltration potentials. In addition, the high clay content zones exhibit a high expected consolidation settlement that might require a long time to vanish with the dissipation of any generated excess pore water pressure. Moreover, such high clay content areas can operate as a barrier against heavy metal contamination.

Chemical soil distribution

The Arc GIS technique was used to evaluate the pH distribution of the Kirkuk map as shown in Fig. 4a. It is clearly shown that the pH distribution has the highest percentage in the southeastern zone of the city while the lowest percentage in the southwestern zone of the city. The ranges of the pH content start from 6.13 to 7.51% for very low zone, 7.51 to 7.88% for low zone, 7.88 to 8.25% for moderate zone, 8.25 to 8.59% for high zone, and 8.59 to 9.69% for the very high zone. The zones with low pH content might have negative consequences on the future of any organic or inorganic chemicals, erosion, and leaching losses.

In addition, the Arc GIS technique is also used to investigate the gypsum distribution of the Kirkuk map as shown in Fig. 4b. It is clearly shown that the gypsum distribution has the highest percentage in the western zone of the city while the lowest percentage extended from the middle to the southeastern zone of the city. The ranges of the gypsum content start from 0.22 to 0.82% for very low zone, 0.82 to 1.19% for low zone, 1.19 to 1.66% for moderate zone, 1.66 to 2.36% for high zone, and 2.36 to 5.75% for the very high zone. The places with high gypsum content are more susceptible to differential

Fig. 4 **a** pH distribution in Kirkuk city, **b** gypsum distribution in Kirkuk city, **c** Total Dissolved Solid (TDS) distribution in Kirkuk city, **d** Organic distribution in Kirkuk city, and **e** Sulfur Trioxide (SO₃) distribution in Kirkuk city

settlement specifically in rainy weathers since gypsum dissolves in water and can be washed away producing uneven large voids in the soil. Thus, a rigorous caution should be considered in designing the sewage system in such places to prevent any geotechnical problems under the foundation of the built structures.

The Arc GIS or Arc map with Arc tool map technique are used to study the total dissolved solid (TDS) distribution of the Kirkuk map as shown in Fig. 4c. It is clearly shown that the TDS distribution has the highest percentage in the southwestern zone of the city while the lowest percentage in the middle western zone of the city. The ranges of the TDS content start from 0.02 to 1.08% for very low zone, 1.08 to 1.79% for low zone, 1.79 to 3.01% for moderate zone, 3.01 to 4.49% for high zone, and 4.49 to 15.18% for the very high zone. The zones with high TDS content are more toxic due to the increase in the salinity substance, alterations in the ionic water configuration, and poisonousness of discrete ions. Thus, such high TDS areas might have constraint biodiversity, changes in biotic groups, and enduring consequences at certain life phases.

The organic distribution of Kirkuk map is considered using an Arc map with Arc tool map technique as demonstrated in Fig. 4d. It is clearly shown that the organic distribution has the highest percentage in the southwestern zone of the city while the lowest percentage extended from the northeastern to southeastern zone of the city. The ranges of the organic content start from 0.045 to 0.36% for very low zone, 0.36 to 0.63% for low zone, 0.63 to 1% for moderate zone, 1 to 1.44% for high zone, and 1.44 to 3.49% for the very high zone. The places with high organic content could be beneficial for agricultural purposes. In addition, the high organic content has a positive impact on the process of sorption of insecticides that represent biochemical resources of the plants.

Finally, the sulfur trioxide (SO_3) distribution of Kirkuk map is analyzed using Arc GIS technique as shown in Fig. 4e. It is distinctly established that the SO_3 distribution has the highest percentage in the southwestern zone of the city while the lowest percentage in the southeastern zone of the metropolis. The ranges of the SO_3 content start from 0.06 to 0.34% for very low zone, 0.34 to 0.54% for low zone, 0.54 to 0.79% for moderate zone, 0.79 to 1.13% for high zone, and 1.13 to 3.22% for the very high zone. It is very significant to state that the SO_3 is very toxic and corrosive material and its poisonous effect exceeds the SO_2 by more than ten times [17]. In addition, it is verified that SO_3 can cause health problem to human beings represented by asthma. Hence, the zones with high SO_3 content represent places of extreme risk against human health.

Analysis

Physico-chemical correlation

The correlations between physico-chemical soil characteristics such as gravel content, sand content, silt content and clay content with distinctive contents of pH, gypsum, TDS, organic, and SO_3 have been shown Table 2. It is clearly shown that there are significant positive correlations between the physical properties of the Kirkuk soil with the highest degree of correlation between gravel and clay contents up to 0.99. In addition, it can be noticed that there are positive correlations between the contents of gravel, sand, silt and clay with the pH content with the highest degree of correlation between the clay and pH content up to 0.55. However, negative correlations can be observed between the contents of gravel, sand, silt and clay with the gypsum content where the highest negative

Table 2 Correlations between physico-chemical soil characteristics in Kirkuk city

	Gravel	Sand	Silt	Clay	pH	Gypsum	TDS	Organic	SO ₃
Gravel	1								
Sand	0.98	1							
Silt	0.90	0.91	1						
Clay	0.99	0.97	0.92	1					
pH	0.51	0.53	0.45	0.55	1				
Gypsum	−0.19	−0.16	−0.06	−0.19	−0.24	1			
TDS	−0.30	−0.25	−0.16	−0.27	−0.30	0.70	1		
Organic	−0.13	−0.08	−0.07	−0.13	−0.29	0.56	0.74	1	
SO ₃	−0.19	−0.15	−0.04	−0.18	−0.29	0.84	0.76	0.47	1

degree of correlation was obtained between gravel and clay contents with the gypsum content and up to -0.19 . Similarly, negative correlations are obtained between all the studied physical soil properties and the TDS content with the highest negative degree of correlation of -0.30 between gravel and TDS contents. Moreover, negative correlations are obtained between all the studied physical soil properties and the organic and SO₃ contents with the highest negative degree of correlation of -0.19 between gravel and SO₃ contents. It is worthwhile to mention that the correlations between the chemical soil properties vary between positive and negative correlations with the highest positive correlation between gypsum and SO₃ contents of 0.84 and highest negative correlation between pH and TDS contents of -0.30 .

In Kirkuk city, a strong positive degree of correlation between gravel and clay (0.99) has been noticed whereas a weak positive degree of correlation between gravel and clay (0.384) has been reported in the literature [26]. This disparity in the degree of correlation is primarily related to the soil topographic distribution where a large range of soil samples has been obtained all over the Kirkuk area with no restriction to particular soil type. In addition, a strong positive degree of correlation between sand and clay (0.97) has been observed in Kirkuk city, whereas a strong negative degree of correlation between sand and clay (-0.866) has been stated in the literature [24]. The variation in the degree of the correlation is attributed to the difference in the geological soil characteristics between the studied and compared areas. Moreover, a significant positive degree of correlation between sand and silt (0.91) has been achieved while a good negative degree of correlation between sand and silt (-0.79) has been specified in the literature [24]. In a similar manner, the distinction in the degree of the correlation is attributed to the divergence in the geological soil characteristics between the examined and compared areas. Furthermore, a high positive degree of correlation between silt and clay (0.92) has been obtained whilst a weak negative degree of correlation between silt and clay (-0.35) has been determined in the literature [24]. The key explanation could be due to the distinctive definition for grain sizes between silt and clay in various soil classification schemes accompanied with different geological distribution between the inspected and compared areas.

There was a reasonably positive correlation between pH and sand (0.53), while a strong positive correlation between pH and sand (0.778) was found in the literature [37]. The key discrepancy between the degrees of correlation might be attributed to the size of

samples in which ten soil samples were collected from one agricultural area throughout the comparative region, whilst the analyzed region has included a wide variety of soil samples distributed all over the Kirkuk city. In comparison, a reasonably positive degree of pH-silt correlation (0.45) was observed, whereas a strong negative degree of pH-silt correlation (-0.814) was recorded in the literature [37]. In a similar manner, the contrast between the degrees of correlation might be due to the difference in the size of soil samples and the composition of the environment surveyed for both inspected and compared regions. In addition, a moderate degree of positive correlation between pH and clay (0.55) has been obtained while a weak degree of negative correlation between pH and clay (-0.027) has been reported in the literature [24]. The variation in the degree of correlation is due to the difference in geological soil characteristics between the examined and compared areas.

Linear regression model

A linear regression model has been proposed to predict the chemical soil properties based on the observed physical soil properties as follows:

$$\text{Chemical Soil Property} = A * \text{Physical Soil Property}(\%) + B \quad (2)$$

where A and B are model parameters.

A summary for all the details of the proposed model (Eq. 2) has been shown in Table 3. The least square method has been used to solve the linear regression model. In Table 3, all the obtained model equations with their parameters (A and B) accompanied with the values of R^2 are presented. The A and B values range from -0.0004 to 0.030 and 0.39 to 8.36 respectively. In addition, the R^2 values range from 0.0001 to 0.3 . The variations of

Table 3 Linear regression model analysis for different soil properties in Kirkuk city

Chemical property	Physical property	A	B	Equation	R^2
pH	Gravel	-0.006	8.11	$\text{pH} = -0.006 * (\text{Gravel}(\%)) + 8.11$	0.043
Gypsum		-0.005	1.91	$\text{Gypsum}(\%) = -0.005 * (\text{Gravel}(\%)) + 1.91$	0.0026
TDS		0.0133	2.47	$\text{TDS}(\%) = 0.0133 * (\text{Gravel}(\%)) + 2.47$	0.009
Organic		-0.002	0.84	$\text{Organic}(\%) = -0.002 * (\text{Gravel}(\%)) + 0.84$	0.002
SO_3		-0.0004	0.86	$\text{SO}_3(\%) = -0.0004 * (\text{Gravel}(\%)) + 0.86$	0.0001
pH	Sand	-0.016	8.36	$\text{pH} = -0.016 * (\text{Sand}(\%)) + 8.36$	0.14
Gypsum		-0.012	2.09	$\text{Gypsum}(\%) = -0.012 * (\text{Sand}(\%)) + 2.09$	0.009
TDS		-0.020	3.25	$\text{TDS}(\%) = -0.02 * (\text{Silt}(\%)) + 3.25$	0.011
Organic		-0.007	0.98	$\text{Organic}(\%) = -0.007 * (\text{Silt}(\%)) + 0.98$	0.017
SO_3		-0.01	1.09	$\text{SO}_3(\%) = -0.01 * (\text{Sand}(\%)) + 1.09$	0.026
pH	Silt	0.005	7.79	$\text{pH} = 0.005 * (\text{Silt}(\%)) + 7.79$	0.02
Gypsum		0.026	0.85	$\text{Gypsum}(\%) = 0.026 * (\text{Silt}(\%)) + 0.85$	0.07
TDS		0.026	1.83	$\text{TDS}(\%) = 0.026 * (\text{Silt}(\%)) + 1.83$	0.03
Organic		0.011	0.41	$\text{Organic}(\%) = 0.011 * (\text{Silt}(\%)) + 0.41$	0.06
SO_3		0.013	0.39	$\text{SO}_3(\%) = 0.013 * (\text{Silt}(\%)) + 0.39$	0.07
pH	Clay	0.030	7.52	$\text{pH} = 0.03 * (\text{Clay}(\%)) + 7.52$	0.3
Gypsum		-0.027	2.26	$\text{Gypsum}(\%) = -0.027 * (\text{Clay}(\%)) + 2.26$	0.04
TDS		-0.061	3.78	$\text{TDS}(\%) = -0.061 * (\text{Clay}(\%)) + 3.78$	0.08
Organic		-0.009	0.95	$\text{Organic}(\%) = -0.009 * (\text{Clay}(\%)) + 0.95$	0.02
SO_3		-0.0124	1.06	$\text{SO}_3(\%) = -0.0124 * (\text{Clay}(\%)) + 1.06$	0.03

the soil chemical properties with the gravel contents for Kirkuk city have been shown in Fig. 5a–e. Negative linear correlations are noticed for the relation between the values of the pH, gypsum, organic, and SO_3 contents with the increase of the gravel content, whereas a positive linear correlation is obtained for the values of TDS content with the increase of the gravel content. In addition, the variations of the soil chemical properties with the sand contents for Kirkuk city have been shown in Fig. 6a–e. Negative linear correlations are observed for the relation between the values of the pH, gypsum, TDS, organic, and SO_3 contents with the increase of the sand content. Similarly, the variations of the soil chemical properties with the silt contents for Kirkuk city have been shown in Fig. 7a–e. Positive linear correlations are attained for the relation between the values of the pH, gypsum, TDS, organic, and SO_3 contents with the increase of the silt

content. In comparative approach, the variations of the soil chemical properties with the clay contents for Kirkuk city have been shown in Fig. 8a–e. Negative linear correlations are noticed for the relation between the values of the gypsum, TDS, organic, and SO₃ contents with the increase of the clay content, whereas a positive linear correlation is obtained for the values of pH content with the increase of the clay content.

Linear multi regression model

A linear multi regression model has been proposed to predict the chemical soil properties based on the observed physical soil properties as follows:

$$\text{Chemical Soil Property} = C * \text{gravel}(\%) + D * \text{sand}(\%) + E * \text{silt}(\%) + F * \text{clay}(\%) + G \quad (3)$$

where C, D, E, F, and G are model parameters.

A summary for all the details of the linear multi regression model (Eq. 3) has been shown in Table 4. The least square method has been used to solve the linear multi regression model. In Table 4, all the proposed model parameters (C, D, E, F and G) with the obtained values of the model multiple R are displayed. The values of the C, D, E, F and G parameters range from -0.217 to 0.006 , 0.016 to 0.148 , -0.017 to 0.07 , -0.162 to 0.104 and 0.376 to 7.4 respectively. In addition, the values of the model multiple R range from 0.27 to 0.61 . The variations of the predicted versus observed soil chemical properties using the linear multi regression model (Eq. 3) have been shown in Fig. 9a–e. In the Fig. 9a, the proposed model predicts the pH values reasonably where most of the values are close to the equality line. In the Fig. 9b, the

Table 4 Linear multi regression analysis for chemical soil properties in Kirkuk city

Property	Linear multi regression coefficients					Multiple R
	C	D	E	F	G	
pH	−0.052	0.024	−0.017	0.104	7.40	0.61
Gypsum	0.006	0.031	0.07	−0.162	1.109	0.35
TDS	−0.217	0.148	0.068	0.029	1.007	0.42
Organic	−0.020	0.057	0.010	−0.053	0.376	0.27
SO_3	−0.014	0.016	0.037	−0.057	0.378	0.37

proposed model over predicts most of the gypsum content values, whereas the model underestimates the observed gypsum contents with values more than 2%. Similarly, in the Fig. 9c, the proposed model over predicts most of the TDS content values, whereas the model underestimates the observed TDS contents with values more than 4%. Likewise, in the Fig. 9d, the proposed model over predicts most of the organic content values, whereas the model underestimates the observed organic contents with values more than 1%. Comparably, in the Fig. 9e, the proposed model over predicts most of the SO_3 content values, whereas the model underestimates the observed SO_3 contents with values more than 1.5%. Overall, all the chemical properties were predicted reasonably using the proposed linear multi regression model.

Conclusions

In this study, the spatial analysis has been implemented using the inverse distance weighted (IDW) technique with the data of 56 different soil samples collected from distinctive depths of Kirkuk city. Both physical and chemical properties for collected soil samples have been investigated. Moreover, statistical investigations represented by physico-chemical correlation, linear single and linear multi regression models have been utilized. Based on the available data and the results of this study, the following conclusions can be summarized:

1. Nine digital maps for soil type distribution of Kirkuk city including the percentages of the square area of gravel, sand, silt and clay distributions have been obtained.
2. The gravel and sand are distributed similarly on the created digital maps where the highest percentages of both elements lie in the middle and western zones of Kirkuk city. These high gravel and sand content zones are more susceptible to the influence of soil boiling under the impact of any expected earthquakes.
3. The silt distribution is condensed in the center of the northern zones of Kirkuk city, whereas the clay distribution has the highest share in the middle southern and upper northern zones of Kirkuk city. The zones with high silt content are most suitable areas for agricultural purposes while the zones with high clay content are more appropriate to work as barriers against heavy metal contamination.
4. The gypsum and organic distributions on the produced digital maps have the highest percentages in the western and southwestern zones of the city respectively. The high gypsum content zones are susceptible to differential settlement while the high organic content zones can be classified as rich biochemical resources of the plants.
5. The pH distribution has the highest percentage in the southeastern zone of the city where such zones might affect the organic and inorganic chemicals negatively. However, both SO_3 and TDS distributions on the developed digital map have the highest percentages in the southwestern zone of the city where the areas with high SO_3 and TDS contents are very toxic and represent corrosive materials.
6. For Kirkuk soils, it has been proved that the gravel and clay contents were correlated strongly with a degree of correlation of 0.99 whereas positive correlations have been obtained between physical soil characteristics and pH content with the highest degree of correlation of 0.55.
7. Using the linear regression model for Kirkuk soils, negative linear correlations are proved for the relation between the values of the pH, gypsum, organic, and SO_3 contents with the increase of the gravel content, whereas a positive linear correlation is obtained for the values of TDS content with the increase of the gravel content.
8. Using the linear regression model for Kirkuk soils, negative linear correlations are noticed for the relation between the values of the gypsum, TDS, organic, and SO_3 contents with the increase of the clay content, whereas a positive linear correlation is obtained for the values of pH content with the increase of the clay content.
9. Using the linear multi regression model for Kirkuk soils, the proposed model predicts the pH values reasonably where most of the values were close to the equality line. However, the proposed model over predicted most of the values of gypsum, TDS, organic and SO_3 contents.

Acknowledgements

The authors would like to express many thanks to the staff of the general commission of Kirkuk constructional laboratory and all the teams of the field work.

Received: 6 December 2020 Accepted: 11 March 2021

Published online: 02 July 2021

References

- Abdel-Kader FH (2011) Digital soil mapping at pilot sites in the northwest coast of Egypt: a multinomial logistic regression approach. *Egypt J Remote Sens Space Sci* 14:29–40. <https://doi.org/10.1016/j.ejrs.2011.04.001>
- Abdel-Kawy WA, Belal A (2011) Soil resilience mapping in selective wetlands, West Suez Canal, Egypt. *The Egyptian Journal of Remote Sensing and Space Sciences* 14:99–112. <https://doi.org/10.1016/j.ejrs.2011.11.001>
- Ali A, Wani TA, Wani IA, Masoodi FA (2016) Comparative study of the physico-chemical properties of rice and corn starches grown in Indian temperate climate. *J Saudi Soc Agric Sci* 15:75–82. <https://doi.org/10.1016/j.jssas.2014.04.002>
- Ancic M, Hudek A, Rihtaric A, Cazar M, Druzina VB, Kopjar N, Durgo K (2020) Physico chemical properties and toxicological effect of landfill groundwaters and leachates. *Chemosphere* 238:124574. <https://doi.org/10.1016/j.chemosphere.2019.124574>
- Baojun W, Bin S, Inyang HI (2008) GIS-based quantitative analysis of orientation anisotropy of contaminant barrier particles using standard deviational ellipse. *Soil Sediment Contam* 17(4):437–447. <https://doi.org/10.1080/15320380802146784>
- Bensabab F, Lamiri A, Naja J (2015) Effect of purified wastewater from the city of Settat (Morocco) on the quality of *Lippia citriodora* essential oil and infusion. *J Saudi Soc Agric Sci* 14:101–108. <https://doi.org/10.1016/j.jssas.2014.03.001>
- Bergstrom LF, Shirmohammadi A (1999) Areal extent of preferential flow with profile depth in sand and clay monoliths. *J Soil Contam* 8(6):637–651. <https://doi.org/10.1080/10588339991339513>
- Billah MM (2018) Mapping and monitoring erosion-accretion in an alluvial river using satellite imagery- the river bank changes of the Padma river in Bangladesh. *Quaestiones Geographicae* 37(3):87–95. <https://doi.org/10.2478/quageo-2018-0027>
- Conrado NR (2009) Hydrolysis study and extraction of spiroxamine from soils of different physico-chemical properties. *Chemosphere* 77:821–828. <https://doi.org/10.1016/j.chemosphere.2009.07.078>
- Dev R, Bali M (2019) Evaluation of groundwater quality and its suitability for drinking and agricultural use in district Kangra of Himachal Pradesh, India. *J Saudi Soc Agric Sci* 18:462–468. <https://doi.org/10.1016/j.jssas.2018.03.002>
- Dlugosz AP, Lemanowicz J, Dlugosz J (2020) The spatial pattern and seasonal changes in the soil phosphorus content in relation to the phosphatase activity: a case study of Luvisols. *Arch Agron Soil Sci*. <https://doi.org/10.1080/03650340.2020.1759798>
- ESRI. (2010). Manual of ARC GIS. version 9.3.
- ESRI. (2012). Environmental and geotechnical engineering applications of ARC GIS. Map Book Online Volume 24.
- Głębocki B, Kacprzak E, Kossowski T (2019) Multicriterion typology of agriculture: a spatial dependence approach. *Quaestiones Geographicae* 38(2):29–49. <https://doi.org/10.2478/quageo-2019-0021>
- Guo Y, Cao L, Feng X, Liu H (2019) Influence of leachate on properties and regions of compacted clay layer: a column experiment. *Soil Sediment Contam*. <https://doi.org/10.1080/15320383.2019.1650002>
- Hmid I, Hanine H, Elthmani D, Oukabli A (2018) The physico-chemical characteristics of Moroccan pomegranate and evaluation of the antioxidant activity for their juices. *J Saudi Soc Agric Sci* 17:302–309. <https://doi.org/10.1016/j.jssas.2016.06.002>
- Kikuchi R (2001) Environmental management of sulfur trioxide emission: impact of SO₃ on human health. *Environ Manage* 27:837–844. <https://doi.org/10.1007/s002670010192>
- Kumar S, Haq R, Prasad K (2018) Studies on physico-chemical, functional, pasting and morphological characteristics of developed extra thin flaked rice. *J Saudi Soc Agric Sci* 17:259–267. <https://doi.org/10.1016/j.jssas.2016.05.004>
- Kumari D, John S (2019) Sorption capacity of pesticides on soil in a predominant apple cultivation area. *Soil Sediment Contam*. <https://doi.org/10.1080/15320383.2019.1684876>
- Matarira D, Mutanga O, Dube T (2020) Landscape scale land degradation mapping in the semi-arid areas of the Save catchment. *South African Geographical Journal, Zimbabwe*. <https://doi.org/10.1080/03736245.2020.1717588>
- Minzi LR, Saviozzi A, Cardelli R, Riffaldi R (2002) Relationships between biological and physico-chemical soil characteristics in a Mediterranean agricultural area. *Arch Agron Soil Sci* 48(3):279–288. <https://doi.org/10.1080/03650340213843>
- Mohammadi A, Ghassoun Y, Lowner MO, Behmanesh M, Faraji M, Nemati S, Toolabi A, Abdollahnejad A, Panahi H, Heydari H, Miri M (2020) Spatial analysis and risk assessment of urban BTEX compounds in Urmia. *Iran Chemosphere* 246:125769. <https://doi.org/10.1016/j.chemosphere.2019.125769>
- Moosavi AK, Sepaskhah AR (2012) Spatial variability of physicochemical properties and hydraulic characteristics of a gravelly calcareous soil. *Arch Agron Soil Sci* 58(6):631–656. <https://doi.org/10.1080/03650340.2010.533659>
- Mousavifard SM, Momtaz H, Sepehr E, Davatgar N, Sadaghiani MHR (2012) Determining and mapping some soil physicochemical properties using geostatistical and GIS techniques in the Naqade region. *Iran Arch Agron Soil Sci*. <https://doi.org/10.1080/03650340.2012.740556>
- Nadal M, Kumar V, Schuhmacher M, Domingo JL (2006) Definition and GIS-based characterization of an integral risk index applied to a chemical/petrochemical area. *Chemosphere* 64:1526–1535. <https://doi.org/10.1016/j.chemosphere.2005.11.078>

26. Olaleye AO, Akinbola GE, Akintade B (2011) Gravel, soil organic matter, and texture in fallowed alfisols, entisols and ultisols: implications for root and tuber crops. *Commun Soil Sci Plant Anal.* 42:2624–2641. <https://doi.org/10.1080/00103624.2011.614038>
27. Olmo JC, Guervos RC, Rocha IM (2019) The spatial effects of violent political events on mortality in countries of Africa. *S Afr Geogr J.* <https://doi.org/10.1080/03736245.2019.1612770>
28. Omar NQ, Raheem AM (2016) Determining the suitability trends for settlement based on multi criteria in Kirkuk, Iraq. *Open Geosp Data Softw and Stand.* <https://doi.org/10.1186/s40965-016-0011-2>
29. Omar NQ, Ahamad MS, Hussin WMA, Samat N, Ahmed SZB (2014) Markov CA, multi regression, and multiple decision making for modeling historical changes in Kirkuk city, Iraq. *J Indian Soc Remote Sens* 42:165–178. <https://doi.org/10.1007/s12524-013-0311-2>
30. Panghal A, Chhikara N, Khatkar BS (2019) Characterisation of Indian wheat varieties for chapatti (flat bread) quality. *J Saudi Soc Agric Sci* 18:107–111. <https://doi.org/10.1016/j.jssas.2017.02.005>
31. Peinemann N, Amioti NM, Zalba P, Villamil MB (2000) Effect of clay minerals and organic matter on the cation exchange capacity of silt fraction. *J Plant Nutr Soil Soc* 163:47–52. [https://doi.org/10.1002/\(SICI\)1522-2624\(200002\)163:1%3c47::AID-JPLN47%3e3.0.CO;2-A](https://doi.org/10.1002/(SICI)1522-2624(200002)163:1%3c47::AID-JPLN47%3e3.0.CO;2-A)
32. Raja KS, Taip FS, Azmi MMZ, Shishir MRI (2019) Effect of pre-treatment and different drying methods on the physico-chemical properties of *Carica papaya* L. leaf powder. *J Saudi Soc Agric Sci* 18:150–156. <https://doi.org/10.1016/j.jssas.2017.04.001>
33. Shomar BH, Muller G, Yahya A (2005) Geochemical characterization of soil and water from a wastewater treatment plant in Gaza. *J Soil Contam* 14(4):309–327. <https://doi.org/10.1080/15320380590954042>
34. Sultan K (2006) Clay mineralogy of central victorian (Creswick) soils: clay mineral contents as a possible tool of environmental indicator. *J Soil Contam* 15(4):339–356. <https://doi.org/10.1080/15320380600751702>
35. Yulianto F, Suwarsono NUC, Nugroho NP, Sunarmodo W, Khomarudin MR (2020) Spatial-temporal dynamics land use/land cover change and flood hazard mapping in the upstream Citarum Watershed, West Java. *Indonesia Quaestiones Geographicae* 39(1):125–146. <https://doi.org/10.2478/quageo-2020-0010>
36. Zhang K, Zheng X, Li H, Zhao Z (2020) Human health risk assessment and early warning of heavy metal pollution in soil of a coal chemical plant in northwest China. *Soil Sediment Contam.* <https://doi.org/10.1080/15320383.2020.1746737>
37. Zhao W, Liu X, Huang Q, Cai P (2015) *Streptococcus suis* sorption on agricultural soils: role of soil physico-chemical properties. *Chemosphere* 119:52–58. <https://doi.org/10.1016/j.chemosphere.2014.05.060>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)