ARTICLE

Geospatial modeling approaches for mapping topsoil organic carbon stock in northern part of Mongolia

Samdandorj M.* and Purevdorj Ts.

*Division of Soil Science, Institute of Geography and Geoecology,
Mongolian Academy of Sciences, Ulaanbaatar, Mongolia*

Abstract: *Soil organic carbon (SOC) is one of the most important indicators of soil quality and agricultural productivity. This paper presents the application of Regression Kriging (RK), Geographically Weighted Regression (GWR) and Geographically Weighted Regression Kriging (GWRK) for prediction of topsoil organic carbon stock in Tarialan. A total of 25 topsoil (0-30 cm) samples were collected from Tarialan soum of Khuvsgul aimag in Mongolia. In this study, seven independent variables were used including normalised difference vegetation index (NDVI), soil adjusted vegetation index (SAVI), normalised difference moisture index (NDMI), land surface temperature (LST) and terrain factors (DEM, Slope, Aspect). We used root-mean-square error (RMSE), mean error (ME) and determination coefficient (R^2) to evaluate the performance of these methods. Validation results showed that performance of the GWRK, GWR, and RK approaches were good with not only low values of root-mean-square error (1.38 kg/m², 1.48 kg/m², 0.69 kg/m²), mean error (0.28 kg/m², -0.22 kg/m², 0.17 kg/m²) but also high values of R^2 (0.76, 0.72, 0.94). The estimated SOC stock values ranged from 0.28-16.26 kg/m², 0.72–15.24 kg/m², 0.16–15.83 kg/m² using GWRK, GWR, RK approaches in the study area. The highest average SOC stock value was in the wetland (6.47 kg/m², 6.08 kg/m², 6.44 kg/m²) and the lowest was in cropland (1.63 kg/m², 1.48 kg/m², 1.80 kg/m²) using these approaches. According to the validation, GWRK, GWR, and RK approaches produced satisfactory results for estimating and mapping SOC stock. However, Regression Kriging was the best model, followed by GWRK and GWR to predict topsoil organic carbon stock in Tarialan.*

Keywords: *Mongolia; Soil organic carbon stock; Regression Kriging; Geographically Weighted Regression; Geographically Weighted Regression Kriging; Spatial distribution;*

INTRODUCTION

Soil organic carbon (SOC) is a major component of the terrestrial carbon (C) reservoir [1; 2]. According to Lal (2004), 62% (1550 Pg) of the global soil carbon pool of 2500 Pg (1 Pg=10^{15} g) is SOC [3], which is approximately 2 times more than the quantity of C (760 Pg) in the atmosphere [3] and 1.5-3 times larger than the amount of C stored in vegetation [4]. The terrestrial SOC stock in the top 1 m of soil has been estimated to be 1462-

*corresponding author: samdandorj.tes@gmail.com   **https://orcid.org/0000-0003-2007-3197**

1548 Pg C, about half (684-724 Pg C) of which is stored in top 30 cm of soil [1]. Therefore, even a slight change in SOC contents or stocks due to changes in land use, soil management or rates of soil erosion can considerably increase atmospheric CO_2 concentrations [5; 6]. SOC is also an important component of soil structure, fertility, and water holding capacities that play a significant role in agricultural productivity [6; 7]. Information on the spatial distribution of SOC is required for assessing soil ecological functions, understanding soil sequestration processes and agricultural management [8; 9]. In addition, high quality maps of SOC not only provide guidance for soil management practices but also enable more accurate estimations of C stocks [10].

Soil properties, including SOC, were highly spatially variable and dependent [8; 11]. In order to estimate and map SOC, digital soil mapping uses field and laboratory measurements of soil carbon and combines them with environmental data (climate, satellite imagery, digital elevation model) [12]. Many studies have shown relationship between different environmental factors and soil properties at the regional and national scales [8; 13; 14; 10]. But these factors that affect the SOC in different geographical locations are difficult to determine because of the complex and uncertain spatial patterns of SOC [15]. Furthermore, notable uncertainties associated with estimating SOC are due to natural soil variability and limited availability of reliable soil sampling data and measurement [1; 8]. A broad range of methods have been used for modeling and mapping SOC concentration or stock over the past years at different scales. However, no single best method has been developed to model and predict the spatial distribution of SOC because of spatial variability of soil properties, climate, and land use management [15; 16; 7].

Spatial prediction is the process of estimating a target variable at unsampled locations using a wide range of models, including the very popular kriging model, which is still commonly applied in digital soil mapping [15; 17]. In the soil science discipline, it is essential to estimate the SOC stock regionally or globally due to its impact on the global carbon cycle [15]. The key advantages of the techniques - geostatistical interpolation, and, ordinary and universal kriging, are included in multiple software applications and are relatively easy to use [17]. Numerous studies have shown the feasibility of mapping soil properties, especially SOC via different approaches, such as ordinary kriging, cokriging, regression kriging and multiple linear regression [18; 19; 20; 15]. Multiple linear regression (MLR), regression kriging (RK) are widely applied methods for mapping the SOC stock. Multiple linear regression is used to explain the linear relationship between a target variable and independent variables. Regression kriging is a hybrid approach that combines drift and simple kriging of the regression residuals. Given some limitations of both approaches, MLR and RK assumes a stationary relationship between a target variable and independent variable over the study area. Furthermore, numerous studies have demonstrated that RK is relatively easy to perform and its accuracy often outperforms multiple linear regression [21; 15], ordinary linear regression, ordinary kriging [18; 8; 20], and cokriging [13; 19].

In recent years, Geographically Weighted Regression (GWR) approach has received much attention, which allows for different relationships at different locations in space [22; 23]. GWR approach is described in great detail by Brunsdon et al. [22]; Fotheringham et al. [23]. This approach is an extension of the traditional regression framework; in the model, coefficients are specific to a location rather than being global estimates [22]. GWR could consider spatial heterogeneity and variability of relationships between target and independent variables over space by local regression fitting. Geographically Weighted Regression Kriging (GWRK) is an extension of the GWR approach that combines GWR

with kriging of the regression residuals [24]. GWR and GWRK approaches have been used for estimating the spatial distribution of SOC stock at different spatial scales [25; 15; 21; 26; 10].

Mongolia is one of the least researched areas in mapping the terrestrial SOC pool. The main objectives of this study were to predict and map the spatial distribution of topsoil (0-30 cm) organic carbon stock for Tarialan and to compare the performance of GWR, GWRK and RK modeling approaches.

MATERIALS AND METHODS

Descriptions of the study area

The study area, Tarialan soum (49.28°-50.07°N, 101.64°-102.77°E) is located in Khövsgöl aimag, north Mongolia (Fig. 1). It covers an area about of 3407.3 km². The altitude of the study area ranges from 936 m to 2052 m above sea level. The climate in the area is extremely continental and shows a high range of temperature with the coldest temperature at −40°C in January and the hottest temperature at 36°C in July. The mean annual temperature is between -4 and -2°C and the mean annual precipitation ranges from 250 mm to 400 mm, the majority of which falls from July to September.

Due to complex topography, the main soils in the area are *Umbrisols, Leptic Chernozems, Mollic Leptosols, Kastanozems* and *Gelic Histosols* [27]. Vegetation types consist of forest-steppe vegetation and main tree species such as *Larix sibirica, Pinus sibirica,* and *Betula platyphylla*, and the predominant crop in the arable lands is wheat. Fig. 1 shows the geographical location of sampling sites, land use land cover (LULC) classes across the study area. LULC map was generated from Landsat OLI images (Path 135/Row 25, 26) based on maximum likelihood supervised classification techniques in ArcGIS 10.2. (Esri Inc., USA).
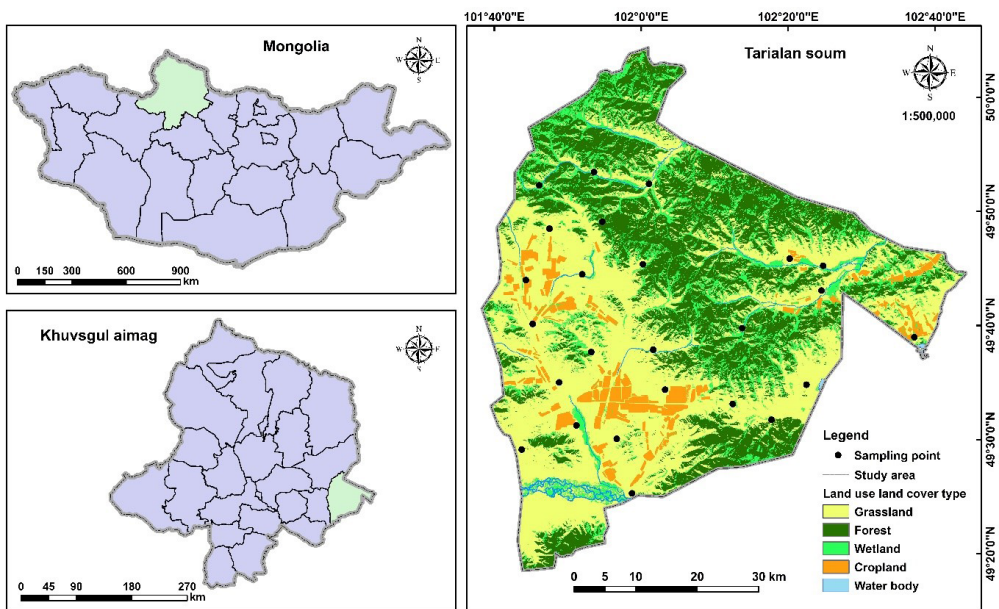

Figure-1. Location of study area and soil sampling sites

Soil sampling and analysis

The soil survey was conducted in July 2018 and a total of 25 topsoil (0-30 cm depth) samples were collected from the entire study area based on random sampling method. The soil samples were air-dried and then passed through a 2 mm sieve for laboratory analysis. Gravels (>2 mm) were sorted out and weighed. To estimate bulk density, 95 cm^3 of undisturbed soil cores at two depths of 10 cm and 20 cm were collected from each sampling sites, they were then dried for 6 hours at 105°C for bulk density measurement. Soil organic carbon content was measured using dichromate oxidation method same as Walkley-Black method [28] at the Soil Science Laboratory, Institute of Geography and Geoecology, Mongolian Academy of Sciences.

Environmental covariates

Various environmental covariates have been used to estimate SOC concentrations or stocks in different geographical locations [15; 16]. Common environmental covariates are remote sensing images, geological, soil and land use maps [14]. Terrain factors, including elevation, slope, and aspect, affected the spatial distribution of SOC [8]. In the present study, environmental covariates, including terrain factors (elevation, slope, aspect), spectral indices (normalised difference vegetation index (NDVI), normalised difference moisture index (NDMI), soil adjusted vegetation index (SAVI)), and land surface temperature (LST), were used to predict SOC stock (kg m^{-2}). Two scenes of Landsat 8 OLI/TIRS images (Path 135/Row 25, 26), acquired in July 2017, were used to estimate spectral indices and LST. The slope gradient and aspect were generated from the ASTER GDEM v2 using ArcGIS 10.2 (Esri Inc., USA). All covariate layers were clipped by region of interest and projected to UTM Zone 48N.

NDVI is the most extensively used general indicator in environmental and vegetation monitoring studies [29]. In order to reduce soil background effect, Huete (1988) used a soil-adjustment factor L to account for first-order soil background variations and obtained a soil-adjusted vegetation index (SAVI) [30]. The NDVI, NDMI, and SAVI are described as:

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)} \tag{1}$$

$$NDMI = \frac{(NIR - MIR)}{(NIR + MIR)} \tag{2}$$

$$SAVI = (1 + L) * \frac{(NIR - Red)}{(NIR + Red + L)} \tag{3}$$

where *NIR* is the near-infrared band, *Red* is the red band [30] and *MIR* is the middle-infrared band. *NIR*, *Red*, and *MIR* correspond to band 4, 5 and 6 in Landsat OLI imagery, respectively. Also, *L* is a soil adjustment factor.

The land surface temperature was derived from Landsat 8 satellite images thermal bands. To estimate land surface temperature, first, digital numbers were converted into top of atmospheric (TOA) spectral radiance values using the following equation.

$$L_\lambda = M_L \times Q_{cal} + A_L \tag{4}$$

where L_λ is TOA spectral radiance ((watts/ (m^2 × srad × μm)), M_L is band-specific multiplicative rescaling factor, Q_{cal} is quantized and calibrated standard product pixel values (DN), A_L - band-specific additive rescaling factor.

Secondly, Eq. (5) was used to convert the TOA spectral radiance values into brightness temperature.

$$T = \frac{K_2}{In\left(\frac{K_1}{L_\lambda} + 1\right)} \tag{5}$$

where, T is the brightness temperature (K), K_1 and K_2 are thermal conversion constant.

Calculation of soil organic carbon stock

The SOC stock was calculated as follows:

$$SOC_s = \sum_{i=1}^{n} S_i \times BD_i \times D_i \times (1 - G_i\%) \times 10^{-2} \tag{6}$$

where, SOC_S is the SOC stock (kg/m²) of topsoil (0-30 cm), i is the soil horizon (1, 2, 3 . . . n), S_i is the SOC concentration (g/kg), BD_i is the bulk density (g/cm³), D_i is the soil thickness (cm) and $G_i\%$ is the gravel concentration (>2 mm).

Mapping approaches

In this study, we used 3 different approaches (GWR, GWRK, and RK) for spatial distribution of soil organic carbon stock.

Ordinary kriging (OK)

Kriging makes predictions of a variable at unsampled locations using weighted average values of variables at sampled locations. Ordinary kriging is the most common form of kriging because it is a simple method to use and explain [17]. The weights are chosen to minimise the prediction error variance. OK is an optimal method, in this work, which was used to interpolate the residuals in RK, GWRK and is as follows:

$$\hat{Z}(S_0) = \sum_{i=1}^{n} \lambda_i \times Z(S_i) \tag{7}$$

where $\hat{Z}(S_0)$ is the estimated value of the target variable at location S_0, $Z(S_i)$ is the measured value of target variable at sampling site S_i, λ_i are the weights associated with the each measured values, and n is the number of observed values.

Geographically weighted regression (GWR)

The GWR considers relationships between target variable (SOC stock) and predictors (environmental variables) at different locations [22; 23]. This is an advanced approach for modeling spatially heterogeneous process. The GWR model used in the present study is defined in the equation (8) below:

$$\hat{C}_{GWR}(S_0) = \beta_0 + \sum_{k=0}^{p} \beta_k(S_0) \times X_k(S_0) + \varepsilon_0 \tag{8}$$

where $\hat{C}_{GWR}$ is the estimated value of SOC stock using GWR approach at location S_0, β_0 is the intercept, $\beta_\kappa(S_0)$ are the regression coefficients at location S_0, $X_\kappa(S_0)$ are the environmental variables at location S_0, p is the number of predictors or environmental variables, and ε_0 is the error term.

In this study, spatial adaptive kernel is a critical setting for GWR model and this is calculated with the adaptive bi-square function:

$$w_{ij} = \left[1 - \left(\frac{d_{ij}}{b}\right)^2\right]^2 \text{If } j \text{ is one of the } n^{th}$$

nearest neighbors of i, and b is the distance of the nth nearest neighbor. =0 otherwise

Akaike Information Criterion (AIC) was used to determine the optimal bandwidth [23]. In the GWR model, densely sampling data reduce the bandwidth and sparsely sampling data has a larger bandwidth.

Hurvich et al. (1998) proposed a corrected version, the AICc, which was used to select smoothing parameters in nonparametric regressions and which is defined as below [32]:

$$AIC_c = 2n \log_e(\hat{\sigma}) + n \log_e(2\pi) + n \left(\frac{n + tr(S)}{n - 2 - tr(S)} \right) \quad (9)$$

where *n* is sample size, $\hat{\sigma}$ is the estimated standard deviation of the error term, and *tr(S)* is the trace of the hat matrix, which is a function of bandwidth.

Geographically weighted regression kriging (GWRK)

Residuals from the GWR were interpolated by ordinary kriging, after which the kriged residual map was added to the regression predicted map to obtain GWRK map. The equation used to perform GWRK approach is given below (Eq. 10):

$$\hat{C}_{GWRK}(S_0) = \hat{C}_{GWR}(S_0) + \hat{\varepsilon}_{OK}(S_0) \quad (10)$$

where $\hat{C}_{GWRK}$ is the estimated SOC stock using GWRK approach at location S_0, $\hat{C}_{GWR}(S_0)$ is the drift fitted by GWR using Eq. (8), and $\hat{\varepsilon}_{OK}(S_0)$ is the residual values interpolated by OK.

Regression kriging (RK)

This model was first introduced by Ahmed and de Marsely, [33] and Odeh et al. [34; 13] and later named it "Regression kriging", while Goovaerts, [35] uses the term "kriging" after detrending. Regression kriging is a hybrid method that combines regression model with simple or ordinary kriging of the regression residuals where regression is used to fit the explanatory variation and simple kriging with expected value 0 is used to fit the residuals [13, 36]. In other words, the residuals and drifts are fitted separately and then results were summed to obtain the final regression kriging map and the equation used to perform this approach is shown below:

$$\hat{C}_{RK}(S_0) = \hat{C}_{MLR}(S_0) + \hat{\varepsilon}_{OK}(S_0) \quad (11)$$

$$\hat{C}_{MLR}(S_0) = \sum_{k=0}^{p} \hat{\beta}_k \times q_k(S_0) \quad (12)$$

$$\hat{C}_{RK}(S_0) = \sum_{k=0}^{p} \hat{\beta}_k \times q_k(S_0) + \sum_{i=1}^{n} \lambda_i \times e(S_i) \quad (13)$$

where $\hat{C}_{RK}(S_0)$ is the estimated SOC stock using regression kriging approach at location S_0, $\hat{C}_{MLR}(S_0)$ is the estimated SOC stock using multiple linear regression approach at location S_0, $\hat{\varepsilon}_{OK}(S_0)$, are residuals interpolated with ordinary kriging at location S_0, $q_k(S_0)$ are the independent variables, $\hat{\beta}_k$ are estimated drift model coefficients, λ_i are the kriging weights determined by the spatial dependence structure of the residual, and $e(S_i)$ are the regression residuals at location .

Model validation

The accuracy of each approach, Geographically Weighted Regression, Geographically Weighted Regression Kriging, and Regression Kriging approach was evaluated by the mean error (ME), the root mean squared errors (RMSE) and coefficient of determination (R^2). These accuracy indices are calculated as follows:

$$ME = \frac{1}{n}\sum_{i=1}^{n}(P_i - M_i) \qquad (14)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - M_i)^2} \qquad (15)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(P_i - M_i)^2}{\sum_{i=1}^{n}(P_i - \overline{M})^2} \qquad (16)$$

where P_i, M_i and $\overline{M}$ are predicted, measured SOC stock values and the mean of measured value at location i.

The lower values of RMSE, ME and the higher values of R^2 indicates that better prediction model.

RESULTS AND DISCUSSION

Descriptive statistics

The descriptive statistical results of measured SOC stock (dependent variable) and environmental covariates used as independent variables are illustrated in Table 1. The measured SOC stock ranged from 0.80 to 12.14 kg/m² with standard deviation and mean value of 2.76 kg/m² and 5.16 kg/m², respectively.

The variation coefficient values of elevation, LST, NDVI, and SAVI were 15.78%, 18.31%, 20.37% and 21.17% respectively, and these values can be considered moderate variability. The variation coefficient values of SOC stock, slope, aspect, and NDMI were 53.61%, 67.19%, 60.82% and 68.92% respectively, indicating higher spatial variability [37]. Also, the mean values were 21.66°C, 0.35 and 0.21 for LST, SAVI, and NDMI and standard deviation values were 3.96°C, 0.07 and 0.14 for these predictors, independently. The descriptive statistical values of other environmental variables are shown in Table 1.

Table 1. Descriptive statistics of dependent and independent variables used in this study

	Max[*]	Min	Mean	SD	CV (%)
Dependent variable					
SOC stock (kg/m²)	12.14	0.80	5.16	2.76	53.61
Independent variables					
Elevation (m)	2079.00	913.00	1386.21	218.79	15.78
Slope (%)	206.22	0.00	22.08	14.84	67.19
Aspect (°)	359.75	-1.00	175.39	106.68	60.82
LST (°C)	36.27	13.49	21.66	3.96	18.31
NDVI	0.83	0.13	0.61	0.12	20.37
SAVI	0.68	0.08	0.35	0.07	21.17
NDMI	0.52	-1.00	0.21	0.14	68.92

[*] Max, maximum; Min, minimum; SD, standard deviation; CV, coefficient of variation

The GWR parameters used for estimating the SOC stock are shown in Table 2. The R^2 and Adj R^2 values for GWR were 0.78 and 0.59 respectively, these values are closer to 1 indicating good explanatory performance.

Additionally, the SSR and AICc in GWR were 357.92 and 179.97, separately, smaller values of these parameters representing better model performance.

Table 2. Diagnostic information of the GWR model used for mapping SOC stock

Residual sum of squares (SSR)	357.92
Akaike Information Criterion (AICc)	179.97
Coefficient of determination (R^2)	0.78
Adjusted R^2 (Adj R^2)	0.59

Model performance

In this study, we compared the results of 3 different approaches for mapping soil organic carbon stock. Table 3 shows the model performance of GWRK, GWR, and RK using RMSE, ME and R^2. Validation results indicated that regression kriging had minimum prediction errors (RMSE of 0.69 kg/m², ME of 0.17 kg/m²) for estimating the SOC stock compared to GWR (RMSE=1.48 kg/m², ME= -0.22 kg/m²) and GWRK (RMSE=1.38 kg/m², ME=0.28 kg/m²) models. Here quite similar validation results were obtained from both the GWRK and GWR approach. Regression Kriging, Geographically Weighted Regression Kriging and Geographically Weighted Regression produced lower values of RMSE, ME and higher values of R^2 indicating that these approaches did not underestimate or overestimate the SOC stock in the entire study area. However, the R^2 value (0.94) of the Regression Kriging model was much higher than the GWR (R^2=0.72) and GWRK (R^2=0.76).

Table 3. Comparison of the performances of GWRK, GWR, RK approaches used in this study (n=25)

	GWRK	GWR	RK
RMSE (kg/m²)	1.38	1.48	0.69
ME (kg/m²)	0.28	-0.22	0.17
R^2	0.76	0.72	0.94

Spatial distribution of SOC stock

Across the study area, SOC stock estimated with GWRK and GWR approaches ranged from 0.28 to 16.26 kg/m² and from 0.72 to 15.24 kg/m², with a mean value of 4.99 kg/m² and 3.86 kg/m², with a standard deviation of 1.95 kg/m² and 2.10 kg/m², respectively. For the Regression Kriging approach, the estimated SOC stock varied from 0.16 to 15.83 kg/m², with mean and standard deviation of 3.93 kg/m² and 2.11 kg/m², separately (Table 4). Also, both the GWR and RK approaches yielded here almost similar estimation of average SOC stock (3.86 kg/m² and 3.93 kg/m², independently). For GWRK estimated average SOC stock of 4.99 kg/m² was substantially higher than those that are obtained by GWR and RK.

Table 4. Summary statistics of the estimated SOC stock (kg/m²) using GWRK, GWR and RK approach

	GWRK	GWR	RK
Max (kg/m²)	16.26	15.24	15.83
Mean (kg/m²)	4.99	3.86	3.93
Min (kg/m²)	0.28	0.72	0.16
SD (kg/m²)	1.95	2.10	2.11

Table 5. Estimated average SOC stock (kg/m²) using GWRK, GWR and RK approach for the main land use, land cover (LULC) classes

LULC types	Area	GWRK	GWR	RK
	km²	kg m⁻²		
Grassland	1599.03	4.01	3.43	4.19
Forest	1068.44	4.72	4.25	3.57
Wetland	551.51	6.47	6.08	6.44
Cropland	185.32	1.63	1.48	1.80
Water body	2.96	0.00	0.00	0.00

Table 5 shows the estimated average SOC stocks for the main land use and land cover classes of the study area. Land cover types give more indications of the effects of climate, vegetation, and human activities on SOC [4]. Among the land use classes, the highest average SOC stock was stored in wetland (6.47 kg/m², 6.08 kg/m² and 6.44 kg/m²), followed by forest (4.72 kg/m², 4.25 kg/m² and 3.57 kg/m²), grassland (4.01 kg/m², 3.43 kg/m² and 4.19 kg/m²), cropland (1.63 kg/m², 1.48 kg/m² and 1.8 kg/m²) according to the GWRK, GWR and RK approaches, respectively. However, for Regression Kriging approach, the estimated average SOC stock for the forest is 3.57 kg/m², which is slightly lower than 4.19 kg/m² for grassland (Table 5). The 3 major land cover classes: grassland, forest, and wetland that cover almost 95% of Tarialan soum, also grassland and forest had a similar average SOC stock of about 4 kg/m². Fig. 2 shows the spatial distribution of estimated SOC stock using GWRK, GWR and RK approaches across the study area. Generally, the higher SOC stocks are stored in wetland, forest and grassland, in the northern and eastern parts of the study area and lower SOC stock in cropland, in the central parts of Tarialan (Fig. 2). The three maps give similar spatial illustrations with relatively high SOC stock along the wetland, small river valleys in the eastern and northern parts of Tarialan. The predominant soils in these areas are represented by *Leptic Chernozems, Mollic Leptosols*, and *Gelic Histosols*. Furthermore, the areas with lower SOC stock were estimated in cropland, most of which lies in the central portions of the area, these low values can be affected by both natural and human-induced factors [38].

Due to the high SOC content of top 20 cm layer is likely to be affected by land-use change (cropping) and natural disturbances (wildfires) [4]. Spatial distribution of SOC is related to vegetation types and succession and human disturbance. As a result, topsoil organic carbon is greatly influenced by these factors [8]. In this study, environmental covariates related to vegetation, such as NDVI and SAVI, mainly affected the SOC stock across the study area.
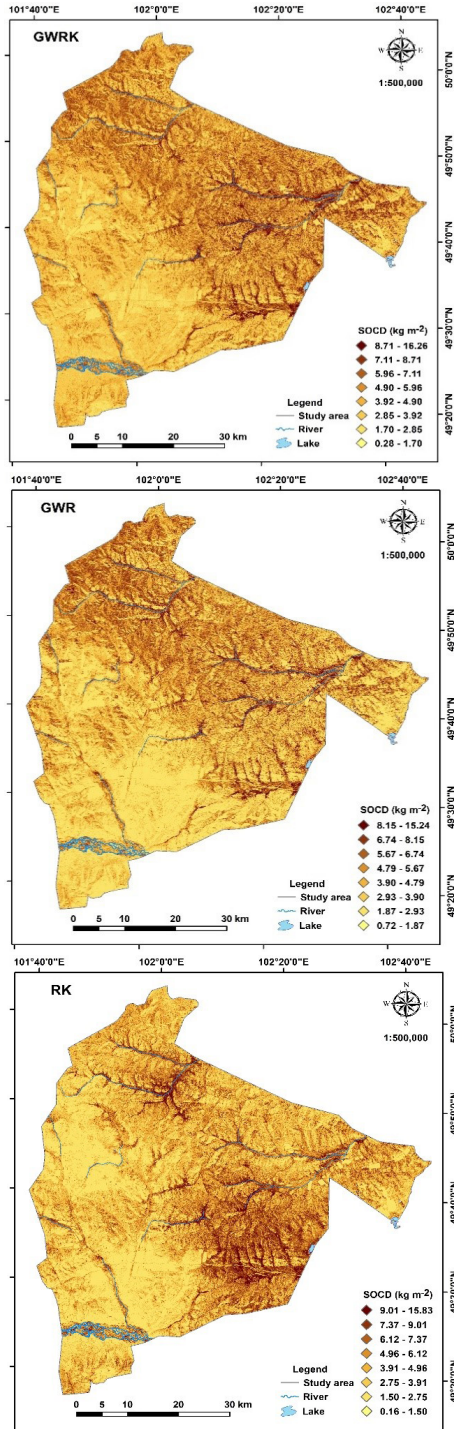
Figure- 2. Spatial distribution of estimated SOC density (SOCD) or stock (kg/m²) for the 0-30 cm depth using GWRK, GWR, and RK approaches

DISCUSSION

On a national scale, Zhang et al. (2011) investigated the use of GWR for spatial modeling of SOC in Ireland [10]. The results showed that the performance of GWR for spatial modeling of SOC and other geochemical parameters was good [10]. Our findings agree with their results. They also suggested that the selection of suitable environmental factors for GWR need to be considered. On a regional scale, Mishra et al. (2010) compared the GWR approach with RK and MLR methods for predicting the spatial variation of the SOC pool in the mid-western United States [25]. Their results demonstrated that GWR can produce more accurate SOC maps than MLR, with similar or better accuracy than the RK approach. On a state scale, a similar investigation was applied using GWRK and RK approaches for mapping SOC stock in the state of Pennsylvania, USA [21]. Their results showed that the performance of GWRK yielded fewer prediction errors, and outperformed the RK. Regression Kriging assumes the existence of constant physical relationship between the target variable (SOC stock) and environmental covariates in all parts of the study area [20]. However, the GWR approach considers there are different relationships between SOC stock and environmental covariates at different locations across the study area.

Therefore, Regression Kriging can be a much more appropriate approach for estimating the spatial distribution of SOC stock in the less heterogeneous environment, while GWR is a much more effective method for a more heterogeneous environment, with a similar purpose. In most cases, GWR and GWRK approaches have outperformed the RK, although RK in the present study provided superior results. In other words, the performance of RK provided relatively fewer prediction errors and outperformed the GWR and GWRK models on a local scale. The SOC concentrations or SOC stocks are affected by various environmental factors at different scales and different geographical locations. Most of the studies investigated the spatial distribution of SOC on a regional scale. Furthermore, there were large uncertainties in SOC stock estimations on regional scales because it is mainly derived from soil survey database and lacks detailed and reliable measurement [8]. Performances are assessed to identify the best model and they are highly dependent on not only the model but also the environment, data and predictors as well, that use used in the research [17]. Similarly, our study results can rely on the discrepancy in geographical locations, predictors, soil analysis and sampling techniques. However, no investigations have been made to estimate SOC stock, and to explore relationships between SOC and environmental factors in this region, which will need further research in Mongolia.

CONCLUSIONS

We mapped the spatial distribution of topsoil SOC stock for Tarialan with Regression Kriging, Geographically Weighted Regression Kriging, and Geographically Weighted Regression methods. The estimated SOC stock was highest in wetland, followed by forest and grassland in the eastern and northern parts of the area, while the lowest was in the cropland, majority of which are primarily located in the central parts of Tarialan. The largest average SOC stock was stored in wetland (6.47 kg/m^2, 6.08 kg/m^2 and 6.44 kg/m^2), followed by forest (4.72 kg/m^2, 4.25 kg/m^2 and 3.57 kg/m^2), grassland (4.01 kg/m^2, 3.43 kg/m^2 and 4.19 kg/m^2), cropland (1.63 kg/m^2, 1.48 kg/m^2 and 1.8 kg/m^2) according to the GWRK, GWR and RK approaches, respectively.

For GWR and GWRK, the estimated average SOC stocks were 3.86 kg/m^2 and 4.99 kg/m^2, independently; for RK the estimated

average SOC stock was 3.93 kg/m². The two modeling approaches for mapping SOC stock, GWR and GWRK, provided almost similar (model performances of GWRK was slightly better than those of the GWR) validation results. In conclusion, RK approach showed more accurate results than GWRK and GWR methods for estimating the spatial distribution of SOC stock on a local scale.

Conflicts of Interest. The authors declare that there is no conflict of interest.

References

1. Batjes, N.H., 1996. Total carbon and nitrogen in the soils of the world. Eur. J. Soil Sci. 47, 151–163.
2. Lal, R., 2002. Soil carbon dynamics in cropland and rangeland. Environ. Pollut. 116, 353–362.
3. Lal, R., 2004. Soil carbon sequestration impacts on global climate change and food security. Science 304, 1623–1627.
4. Wang, S., Huang, M., Shao, X., Mickler, R.A., Li, K., Ji, J., 2004. Vertical distribution of soil organic carbon in China. Environ. Manage. 33, S200–S209.
5. Lal, R., 2003. Global potential of soil carbon sequestration to mitigate the greenhouse effect. Critical Reviews in Plant Sciences 22 (2), 151–184.
6. Liu, Z., Shao, M., Wand, Y., 2011. Effect of environmental factors on regional soil organic carbon stocks across the Loess Plateau region, China. Agri. Ecosystems and Environment. 142, 184-194.
7. Mondal, A., Khare, D., Kundu, S., Mondal, S., Mukherjee, S., Mukhopadhyay, A., 2016. Spatial soil organic carbon (SOC) prediction by regression kriging using remote sensing data. Egypt. J. Remote Sensing Space Sci. 1-10.
8. Liu, L., Wang, H., Dai, W., Lei, X., Yang, X., Li, X., 2014. Spatial variability of soil organic carbon in the forestlands of northeast China. J. Forest. Res. 25 (4), 867–876.
9. Yang, L., Song, M., Zhu, A., Qin, C., Zhou, C., Qi, F., Li, X., Chen, Z., Gao, B., 2019. Predicting soil organic carbon content in croplands using crop rotation and Fourier transform decomposed variables. Geoderma 340, 289-302.
10. Zhang, C.S., Tang, Y., Xu, X.L., Kiely, G., 2011. Towards spatial geochemical modelling: use of geographically weighted regression for mapping soil organic carbon contents in Ireland. Appl. Geochem. 26, 1239–1248.
11. Cambardella, C., Moorman, T., Parkin, T., Karlen, D., Novak, J., Turco, R., Konopka, A., 1994. Field-scale variability of soil properties in central Iowa soils. Soil Sci. Soc. Am. J. 58, 1501–1511.
12. Adhikari, K., Owens, P.R., Libohova, Z., Miller, D.M., Wills, S.A., Nemecek, J., 2019. Assessing soil organic carbon stock of Wisconsin, USA and its fate under future land use and climate change. Science of the Total Environment 667, 833-845.
13. Odeh, I.O.A., McBratney, A.B., Chittleborough, D.J., 1995. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. Geoderma 67, 215–226.
14. McKenzie, N., Ryan, P., 1999. Spatial prediction of soil properties using environmental correlation. Geoderma 89 (1-2), 67-94.
15. Kumar, S., Lal, R., 2011. Mapping the organic carbon stocks of surface soils using local spatial interpolator. J. Environ. Monit. 13, 3128–3135.

16. Liu, Y., Guo, L., Jiang, Q., Zhang, H., Chen, Y., 2015. Comparing geospatial techniques to predict SOC stocks. Soil & Tillage Research 148, 46-58.

17. Veronesi, F., Calogero, S., 2019. Comparison between geostatistical and machine learning models as predictors of topsoil organic carbon with a focus on local uncertainty estimation. Ecol. Indic. 101, 1032-1044.

18. Minasny, B., McBratney, A.B., 2007. Spatial prediction of soil properties using EBLUP with Matern covariance function. Geoderma 140, 324–336.

19. Eldeiry, A.A., Garcia, L.A., 2010. Comparison of ordinary kriging, regression kriging and cokriging techniques to estimate soil salinity using Landsat images. J. Irrig. Drain. Eng. 136, 355–364.

20. Hengl, T., Heuvelink, G., Rossiter, D.G., 2007. About regression-kriging: from equations to case studies. Comput. Geosci. 33 (10), 1301–1315.

21. Kumar, S., Lal, R., Liu, D., 2012. A geographically weighted regression kriging approach for mapping soil organic carbon stock. Geoderma 189, 627–634.

22. Brunsdon, C., Fotheringham, A.S., Charlton, M.E., 1996. Geographically weighted regression: a method for exploring spatial nonstationarity. Geogr. Anal. 28, 281–298.

23. Fotheringham, A.S., Brunsdon, C., Charlton, M., 2002. Geographically weighted regression: the analysis of spatially varying relationships. John Wiley & Sons Ltd., Chichester, UK.

24. Harris, P., Fotheringham, A., Crespo, R., Charlton, M., 2010. The use of geographically weighted regression for spatial prediction: an evaluation of models using simulated data sets. Math. Geosci. 42, 657–680.

25. Mishra, U., Lal, R., Liu, D., Van Meirvenne, M., 2010. Predicting the spatial variation of the soil organic carbon pool at a regional scale. Soil Sci. Soc. Am. J. 74, 906-914.

26. Kumar, S., 2015. Estimating spatial distribution of soil organic carbon for the Midwestern United States using historical database. Chemosphere 127, 49-57.

27. Mongolian Academy of Sciences, 2009. National Atlas of Mongolia. Ulaanbaatar.

28. Walkley, A., Black, I.A., 1934. An examination of the Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. Soil Sci. 37 (1), 29-38.

29. Tucker, C.J., Vanpraet, C.L., Sharman, M.J., Van Ittersun, G., 1985. Satellite remote sensing of total herbaceous production in the Senegalese Sahel 1980–1984. Remote Sens. Environ. 17, 232–249.

30. Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309.

31. Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1973. Monitoring vegetation systems in the Great Plains with ERTS. In: Freden, S.C., Mercanti, E.P., Becker, M. (Eds.), Third Earth Resources Technology Satellite—1 Symposium—Technical Presentations, Section A 1. Goddart Space Flight Center, National Aeronautics and Space Administration, Washington, DC, 309–317.

32. Hurvich, C.M., Simonoff, J.S., Tsai, C.L., 1998. Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. J. R. Stat. Soc. Ser. B: Stat. Methodol. 60, 271–293.

33. Ahmed, S., De Marsily, G., 1987. Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. Water Resour. Res. 23, 1717–1737.

34. Odeh, I.O.A., McBratney, A.B., Chittleborough, D.J., 1994. Spatial prediction of soil

properties from landform attributes derived from a digital elevation model. Geoderma 63, 197–214.

35. Goovaerts, P., 1999. Using elevation to aid the geostatistical mapping of rainfall erosivity. Catena 34 (3–4), 227–242.

36. Hengl, T., Heuvelink, G., Stein, A., 2004. A generic framework for spatial prediction of soil variables based on regression kriging. Geoderma 122 (1–2), 75–93.

37. Wilding, L., 1985. Spatial variability: Its documentation, accommodation and implication to soil surveys. Soil spatial variability. Workshop, 166–194.

38. Kempen, B., Dalsgaard, S., Kaaya, A.K., Chamuya, N., Ruipérez-González, M., Pekkarinen, A., Walsh M.G., 2019. Mapping topsoil organic carbon concentrations and stocks for Tanzania. Geoderma 337, 164-180.