Estimation and mapping of surface soil properties in the Caucasus Mountains, Azerbaijan using high-resolution remote sensing data

Elton Mammadov[1*]

[1*]Institute of Soil Science and Agrochemistry of Azerbaijan National Academy of Sciences, 5 M.Rahim, AZ1073, Baku, Azerbaijan. Phone: (+99470) 7877474; (+99455) 7877474; E-mail: elton.eldaroglu@gmail.com

Jakub Nowosad[2]

[2]Institute of Geoecology and Geoinformation, Adam Mickiewicz University in Poznan, Krygowskiego 10, 61-680, Poznan, Poland. E-mail: nowosad@amu.edu.pl

Cornelia Glaesser[3]

[3]Department of Remote Sensing and Cartography, Martin Luther University Halle-Wittenberg, Von-Seckendorff Platz 4, 06120 Halle (Saale), Germany. E-mail: cornelia.glaesser@geo.uni-halle.de

Abstract

Soil surveys and mapping with traditional methods are time-consuming and expensive especially in mountainous areas while demand for detailed soil information is steadily increasing. This study tested two spatial hybrid approaches to predict and map basic soil properties using high resolution digital elevation model (DEM) and multispectral satellite imagery in a study area located in the Caucasus Mountains, Azerbaijan. Terrain attributes and spectral indices extracted from DEM with 12.5m spatial resolution and Pléiades-1 data were used as auxiliary variables. A total of 115 soil samples were collected from the surface layer of 423 ha area and tested for soil organic carbon, soil reaction (pH in H_2O and KCl solutions), calcium carbonate ($CaCO_3$), sand, silt, clay and hygroscopic water content. The predictive capability of Universal Kriging (UK) and Random Forest Kriging (RFK) was evaluated using spatial cross-validation technique. To model and quantify the associated uncertainty of these models a probabilistic framework, kriging variance approach was applied. The uncertainty models were validated using independent and randomly selected control points (20% of the reference samples). For this, the actual fraction of true values falling within symmetric prediction intervals was calculated and visualized known as accuracy plot. Although the performances of the tested models were similar, RFK was superior in view of both accuracy and computed biases. The models were capable of delineating spatial pattern, mostly elevation dependent as well as the local patterns attributed by e.g., variations in vegetation, land use and soil erosion. UK model produced a few local erratic spatial patterns (e.g., in the case of pH) corresponding to the artifacts such as roads and houses in the image that should be considered in future applications. When comparing the uncertainties, both the models produced considerable underestimations and overestimations depending on soil property. RFK provided better uncertainty estimation for the most of soil properties than UK, the latter technique was more appropriate for the clay and pH_{KCl} prediction. This case study confirmed the importance of assumptions made in uncertainty modelling and quantification. Those soil properties were therefore reliably predicted that their residuals were compatible with the normality assumption and showed apparent spatial correlation, e.g., both the models severely overestimated uncertainty of $CaCO_3$ due to lack of normality assumption and low spatial correlation. This study showed that high resolution remote sensing data are promising, and the procedure presented in this study can be reliably used to map the studied soil properties and extended to partially larger adjacent areas characterized by similar environmental conditions in the Caucasus Mountains. However, with respect to future digital soil mapping, we assume that it is

important to consider sampling design, testing other modelling approaches their uncertainties and multi-scale digital terrain analysis as well.

Key words: soil properties, terrain attributes, spectral indices, hybrid spatial models, uncertainty modeling, Kastanozems, the Caucasus Mountains

1. Introduction

The soil, particularly in the mountainous regions is an important controlling factor of many environmental processes such as forest and crop growth, hydro-geochemical cycle, nutrient leaching and greenhouse gas emission (Lal, 2004; Wilding, 1985). The quantitative estimation of spatial variability of soil is important in a better understanding of complex relationships between soil properties and environmental factors. This complexity includes intrinsic and extrinsic factors (Heuvelink and Webster, 2001) varying greatly depending on topography, climate, vegetation and anthropogenic activity all of which significantly affect spatial variability of soil properties (Shi et al., 2009).

Availability of numerous digital covariates as well new types of space-born remote sensors and advances in computers and information technologies have created a quantitative approach in soil mapping, termed digital soil mapping (DSM) (McBratney et al., 2003). By making use of large diversity of auxiliary data sources and modeling approaches from simple linear models to complex machine learning techniques, DSM has shifted from a research phase into operational use (Minasny and McBratney, 2016) and become comparable, more accurate and an indispensable approach compared to conventional mapping (Bazaglia Filho et al., 2013; Collard et al., 2014).

Overall, DSM especially in mountainous areas, is a complex task due to heterogeneity, the nature of soil-forming factors, and the close relationships that occurred between them. This is especially complex in the Caucasus Mountains where the local communities host traditional agriculture, and typical land ownership is characterized by small patches (Salukvadze and Medvedkov, 2011) when high resolution soil mapping is considered. On the other hand, soil surveys in mountainous areas can be costly, and DSM usually faces a general lack of available observations that often limits the quantity and quality of the samples on which the mapping procedure is based. In this sense, high resolution DSM in mountainous areas can contribute to the proper evaluation and sustainable management of soil resources as well different environmental modelling applications.

In mountainous areas, the topography plays a crucial role and influences significantly most of the other soil-forming factors such as e.g., climate and vegetation cover (Pendleton and Jenny, 1945). The character of such close relations between soil and environmental factors encourages considerably linear regression technique (McKenzie and Ryan, 1999), whereas non-linear relationships between soil properties and environmental factors usually occur and reduce the applicability of the linear regression to limited areas (Lian et al., 2009). These problems can be partially solved by non-parametric machine learning techniques and hybrid spatial models as well auxiliary data used to compensate for the data shortage.

A variety of approaches, such as traditional statistical (Lian et al., 2009), geostatistical and machine learning techniques have been extensively applied in DSM (Keskin and Grunwald, 2018; Veronesi and Schillaci, 2019; Wiesmeier et al., 2011). The geostatistical approach is one of the best spatial interpolation techniques (Webster and Oliver, 2008) and its reliable performance owing to e.g., sampling design, the nature of the study area was presented in many studies in particular with main focus on soil organic carbon mapping (Cambule et al., 2014; Pouladi et al., 2019; Veronesi and Schillaci, 2019). However, it was stressed that kriging becomes uncertain with a small number of samples and unable to model local variations (stationarity assumption) of soil properties caused by differences in e.g., topography, climate and vegetation (Elbasiouny et al., 2014; Eldeiry and Garcia, 2010).

The application of machine learning techniques and hybrid spatial models is steadily increasing in environmental and soil science (Keskin and Grunwald, 2018; Li et al., 2011; Veronesi and Schillaci, 2019). Random Forest (RF) (Breiman, 2001) is one of the most applied machine learning technique quickly gained popularity due to its ability to model non-linear relationships on high dimensional data and the possibility to provide variable importance (Liaw and Wiener, 2002; Myles et al., 2004). Therefore, its superiority over other geostatistical and machine learning techniques was well presented in several studies (Subburayalu and Slater, 2013; Wiesmeier et al., 2011) as well as in trend estimation when hybrid approach was considered (Hengl et al., 2015). Hybrid spatial models employ simultaneously both kriging interpolation and auxiliary variables in order to increase prediction accuracy (Odeh et al., 1995; Phachomphon et al., 2010). Recent studies have emphasized outperformances of the hybrid models over other approaches in combination with different scales, usage of varying resolution auxiliary and response variables (Chen et al., 2019; Tziachris et al., 2019).

As regards auxiliary data, digital elevation models (DEM) and multispectral satellite images (MI) have been widely used in DSM (McBratney et al., 2003). Majority of DSM discussions preferably used freely available

DEMs and MIs. Together with the modeling approach and auxiliary data, the scale is an essential consideration in DSM. As the soil-forming and environmental factors vary and respond at different scales, selecting DEM and MI resolution is a noteworthy phase. The terrain features can influence the prediction performance of soil properties, the scale at which pedogenic processes occur in the landscape, and whether the DEMs can represent the features in the terrain (McBratney et al., 2003). More recently, there has been growing interest in evaluating how the characteristics of auxiliary variables contribute to the success of DSM. It is commonly accepted that more spatial information the auxiliary variables gain the more accurately they describe environmental conditions (Hengl et al., 2013). Accordingly, several studies focusing on the effect of multiscale terrain analysis (Behrens et al., 2010; Cavazzi et al., 2013), spatial resolution of DEM (Smith et al., 2006) and MI (Xu et al., 2018) confirmed that more detailed covariates lead to more accurate predictions. The overall accuracy of finer resolution DEM is better than that of medium resolution such as SRTM and ASTER DEM, and the rougher and steeper topography is satisfyingly depicted with high resolution DEM (Kramm and Ho, 2019). However, Kim and Zheng, (2011) found inverse results that fine-scale topographic information is not always optimal for understanding soil spatial variability. Likewise, the effect of fine and high-resolution MIs delivered by commercial satellites such as SPOT 6, Rapid Eye, Pléiades-1, World View-2/3 have commonly less studied. Sumfleth and Duttmann, (2008) and Xu et al., (2018) confirmed advantages of using higher spatial resolution images for predicting different soil properties in terms of accuracy and error assessment, while others (Samuel-Rosa et al., 2015; Xu et al., 2018) pointed out that the auxiliary variables extracted from higher spatial resolution images may not always produce the most accurate soil prediction.

In this context, discussions for addressing high resolution auxiliary data are limitedly available in mountainous regions. Especially, no study was concerned with DSM using high resolution data in the Caucasus Mountains. Therefore, this study specifically focused on evaluating the potential of high-resolution DEM and MI (Pléiades-1) in predicting soil properties in a local test area in the Caucasus Mountains. The DEM used in this study was a product of the Advanced Land Observing Satellite (ALOS) that has been made publicly available recent years (Alaska Satellite Facility). To produce a model possibly higher accuracy, high-resolution auxiliary data was combined with modelling approaches. Therefore, we tested two spatial hybrid models, Universal Kriging (UK) and Random Forest Kriging (RFK) as well as their associated uncertainty to evaluate and map basic soil properties, soil organic carbon (SOC), soil reaction (pH), particle sizes (sand, silt, and clay), calcium carbonate ($CaCO_3$) and hygroscopic water content (WC).

2. Material and methods

2.1. Description of the study area

The study was performed in the Caucasus Mountains, in the administrative area of the Tovuz district, west part of Azerbaijan (Fig. 1). The test area is in the foothill belt, between 40.84°N and 40.87°N latitudes, 45.62°E and 45.64°E longitudes. The foothill belt is a transition zone between semi-arid (semi-desert and dry steppe climate zone) and a moderate climate. Despite being a transition zone, its landscapes are characterized by considerable variability. The elevation ranges from 700 to 1100 m above sea level along the 423 hectares area of the test site. The morphology of the study area is characterized by gentle and dominantly west and east-facing slopes. Pastures, hayfields and arable lands are typical land uses, while natural shrubbery is locally typical due to extensive agricultural use. In combination with herbaceous vegetation, the shrub association is dominant in north and north-west face slopes at higher elevations. The lower timberline of deciduous forest passes through the highest south part of the study area.

The climate is characterized by a moderate with mild-dry winters. An average air temperature and annual precipitation is 10.2°C and 570 mm, respectively. Rainfall mostly occurs maximum in spring with frequent high-intensity rainstorms and minimum in winter with short-term (a few days) persistence of snow cover. As typical to Mediterranean climate, winters are mild and dry, and summers are mild and warm. The soil moisture regime is xeric, and the soil moisture control section is moist for ½ to ¾ of time, moist for more than 45 consecutive days in winter, or dry in summer.

The geological setting of the study area comprises of volcanic, volcanic-sedimentary and sedimentary rocks. The typical rocks are quartz-diorites, tonalites of the upper Jurassic lower Cretaceous system (Bayramov et. al., 2008). However,

Fig. 1. Location of the study area (DEM source: Alaska Satellite Facility, https://asf.alaska.edu/sar-data-sets/alos-palsar/).

sedimentary rocks, especially diluvial sediments prevail as parent material at lower elevations or foot slopes.

Before this study, no extensive soil survey had been performed in this area. The only existing largest scaled soil map (1:100 000) of the area is the official state soil map of Azerbaijan (Salayev et. al., 1990), which is practically unsuitable for characterizing soils in areas smaller than 1000 hectares. In the study area, the

accommodating soil group is Kastanozems (Babayev et. al., 2017). The soils in the south and south-east face slopes, especially at higher elevations are shallow in depth and skeletal with sandy-loam and loam in texture. As the foothill belt is susceptible to agricultural use and soil erosion, outcrops of bedrocks and severe erosion is commonly typical.

2.2. Soil sampling and laboratory analyses

A total of 115 soil samples were collected from the surface layer (0-15 cm) resulting in 1 sample per 3.7 hectare. The sample locations were placed irregularly taking soil-landscape features, such as topography, land use/cover, geological substrate and the erosive state into account. Geographical coordinates of the sampling locations were determined using a Garmin GPS Map 62s. The samples were air-dried, milled and sieved through a 2 mm sieve then analyzed in a laboratory. SOC content was determined using the Walkley-Black method (Sparks et al., 1996), soil reaction in 1:1 soil to H_2O (pH_{H2O}) and KCl (pH_{KCl}) solutions (ISO 10390, 1997) using pH-meter. Particle sizes (sand, silt, and clay) were determined by the hydrometer method (Klute, 1988) and classified according to the textural classes of International Society of Soil Science. $CaCO_3$ content was determined using the pressure calcimeter method and, WC of the samples was determined by drying at a temperature of 105°C for 12 hours.

2.3. Statistical analysis of the tested soil properties

To summarize the characteristics of the response variables (soil properties), the results of chemical tests were described in terms of minimum (Min), maximum (Max), mean, median, standard deviation (SD), coefficient of variation (CV), skewness and kurtosis. Initially, the results of analyses of soil properties were checked for normality test. As the soil properties were non-normally distributed, Spearman's Rho correlation coefficients between soil properties were calculated to find out the relationships between them in view of modelling results. The statistics resolve whether the response variables are complex to model using linear or non-parametric techniques. The correlation coefficients were interpreted based on Wuensch and Evans, (1996) classification (r = 0.00-0.19-very weak; 0.20-0.39-weak; 0.40-0.59-moderate; 0.60-0.79-strong and 0.80-1.00-very strong). The CV values were interpreted according to Wilding, (1985), the CV values of >15%, 15 < CV < 35%, and > 35% represent low, moderate, and high variability, respectively. The descriptive statistics were calculated in the R software environment (R Development Core Team, 2019).

2.4. Remote sensing data

2.4.1. Digital elevation model and extracting terrain attributes

The DEM used in this study was a product collected with the Phased Array Type L-Band Synthetic Aperture Radar instrument of the Advanced Land Observing Satellite (ALOS-PALSAR, Japan Aerospace Exploration Agency, 2012). The ALOS-PALSAR data was processed within a project of the Alaska Satellite Facility (ASF, Alaska Satellite Facility, 2019) that produced PALSAR data at 12.5 and 30 m resolutions, both of which are freely accessible to a broader community of users. The very-high-resolution radiometric- and terrain-corrected DEM with a spatial resolution of 12.5 m was downloaded from the ASF's archive (Laurencelle et al., 2015). The terrain attributes possibly representing environmental factors in the mountainous condition were extracted from the DEM (Table 1) using R packages (Brenning, 2008; Conrad et al., 2015; Hijmans, 2016; Leutner and Horning, 2017) and GRASS GIS software (Neteler and Mitasova, 2008).

2.4.2. Multispectral satellite data and extracting spectral indices

In this study, high-resolution atmospheric- and radiometric-corrected MI from the Pléiades satellite constellation was used. The Pléiades constellation delivers panchromatic product with 0.5 m resolution and four-band multispectral product with 2 m pixel resolution. The image acquisition date (20 September 2015) was selected to be as close to the soil sampling campaign (4 September 2015) as possible. This period was selected to coincide with the higher sun angle and peak of the dry season (end of summer). Because the soil is bare and dry after cereals and potato harvesting in arable lands as well pastures and hayfields are represented with scarce vegetation cover that is convenient for satellite sensors to collect information from the soil surface. The reflectance values of the four bands, and a set of spectral indices (Table 1) were calculated in R (Hijmans, 2016; Leutner and Horning, 2017; R Development Core Team, 2019) and GRASS GIS software (Neteler and Mitasova, 2008). Soil Adjusted Vegetation Index (SAVI) was computed with considering a value of constant variable, L equal to 0.5. In order to integrate auxiliary variables and facilitate the mapping of target soil properties, the spectral indices (2 m) were resampled to the original resolution of the terrain attributes (12.5 m).

Table 1 Auxiliary variables extracted from the DEM and MS.

2.5. Hybrid spatial prediction and mapping

Spatial prediction is the process of estimating the values of a target quantity at unvisited locations and its application to a whole study area is referred to as spatial mapping. The spatial prediction and mapping with hybrid techniques combines kriging and use of auxiliary information. Thus, hybrid spatial techniques combine regression

between the target soil variable and secondary parameters, with geostatistical methods on the residuals of the regression that optimizes the prediction of target soil properties in unsampled locations (Hengl et al., 2007; McBratney et al., 2000). The prediction by hybrid spatial model is the sum of the prediction of regression and the prediction of the kriged residuals of the regression that can be described and modelled in terms of a deterministic component and a stochastic component:

$$Z(u) = m(u) + \varepsilon(u) \quad [1]$$

where Z is the soil property, m is the deterministic part describing structural variation, ε is the stochastic part (error) consisting of random variation and u is the vector of the geographical coordinates (Ahmed and De Marsily, 1987). The predictions of any given model are always different from true values termed as error which is unavoidable (Heuvelink, 2014). The error is defined as the difference between the true and predicted value that could be spatially correlated. In fact, error is not known spatially exhaustively, and we often do not know it because true values are difficult to measure everywhere in the study area.

Therefore, uncertainty refers to a state of our knowledge that is uncertain about the given soil property. As the soil map predicted based on Eq. [1] is simplified representations of complex and partially unknown patterns of soil variations, any predictions derived from this map have an irreducible uncertainty (Heuvelink, 2014). This distinction between the two parts of Eq. [1] is important because it allows the separate interpolation of the two components and the application of different regression and machine learning techniques (Hengl et al., 2007). In this regard, an emphasis was given to testing performances of two hybrid spatial models. In this study, UK and RFK were compared not only for their ability to predict the soil properties of interest but also for their ability to deliver an accurate estimate of the associated uncertainty. Before modelling, the Spearman's Rho correlation coefficients were calculated between all auxiliary variables to identify the relationships between them. Very strong correlation coefficients were found between SAVI and NDVI (r = 0.99), TRI and SL (r = 0.99), B, G, R and TGSI (r > 0.95). Therefore, one of the highly correlated auxiliary variables (B, G, R, NDVI, and TRI) were removed from the predictor data.

2.5.1. Universal Kriging (UK)

UK, regression-kriging and kriging with external drift belong to the group of the so-called 'hybrid' geostatistical approaches (McBratney et al., 2000). These approaches are practically the same generic methods whereas differ in view of methodological steps used. UK, first introduced by Matheron (1969), combines regression

of the target soil variable on environmental covariates with kriging of the regression residuals (Hengl et al., 2007). According to Wackernagel (1998), the term UK is referred to a case where the drift is modelled as a function of geographical coordinates only. Therefore, this approach is ideal from the perspective that both the correlated variables and model residuals are handled simultaneously. In terms of Eq. (1) the deterministic component is modelled by a multiple linear regression whereas the stochastic part of the variation is modelled by kriging using the regression residuals that are assumed as multivariate normal. This model automatically generates prediction uncertainty via the kriging variance that is the sum of the estimation variance of the deterministic component and the prediction variance of the kriged residuals (Hengl et al., 2007). Therefore, it reflects the position of unsampled locations in both geographic and feature space. By means of UK prediction and its variance a parametric model of uncertainty can be produced with the assumption of normality. It is a full and mathematically concise model since a normally distributed random variable is fully determined by its mean and variance.

2.5.2. Random Forest Kriging (RFK)

RFK is considered a type of regression kriging and regression kriging uses regression on auxiliary information and ordinary kriging to interpolate the residuals from the regression model (Hengl et al., 2007). The prediction of RFK consists of two phases in terms of Eq. [1], the deterministic component is modelled by RF in the first phase, and in the second phase the stochastic part of variation is modelled by kriging using the computed residuals and then added to the predictions at corresponding locations. The variogram of the stochastic component is estimated according to Matheron (1963). RF is an extension of classification and regression tree models. As a non-parametric method, RF is capable of modeling non-linear relationship on high dimensional data. In more detail, RF consists of a large number of individual tree models trained from bootstrap samples of the data (Breiman, 2001). The results of all individual trees are aggregated to make a single prediction. This method can also rank the predictor variable's relative importance based on the regression prediction error of out-of-bag (OOB) predictions. RF perturbs each variable and calculates the importance of the resulting change in the OOB error. Two basic parameters in the RF method are the number of trees (in our case ntree = 500) and the number of variables available for selection in each split (mtry = 4) (Houborg and McCabe, 2018).

The general effect of auxiliary variables on RF model was determined using the actual impurity reduction importance metric (Nembrini et al., 2018; Sandri and Zuccolotto, 2008). Additionally, partial dependence plots (Friedman, 2001) were created to determine the most important covariates for the predicted soil properties using the

DALEX R package (Biecek, 2018; Greenwell, 2017). Further R packages used for creating models were gstat (Pebesma, 2004), GSIF (Heuvelink and Malone, 2020), and ranger (Wright and Ziegler, 2017).

2.6. Model validation

The performance of each model was evaluated based on the spatial cross-validation technique proposed by Brenning, (2012). Based on this procedure, the reference sampling points were divided into six spatial partitions (folds) to validate the considered models. Fig. 1 shows the sampling locations represented in different spatial partitions. Each model was created based on the five folds and evaluated on the remaining fold. This procedure was successively repeated until each fold was used as a testing set in view to cope with possible variations of performances when the soil properties are predicted at the field scale in a heterogeneous mountainous area. The quality of each model was determined using root mean square error (RMSE), mean error (ME), and residual prediction deviation (RPD). The standard deviation of each metric (SD_{RMSE}, SD_{ME}, and SD_{RPD}) was calculated across all repetitions and then averaged. A reasonable goal is to produce map with RMSE as low as possible and ME close to zero. The validation procedure was performed using the mlr3spatiotemporal package by Schratz, (2019).

The spatial cross-validation technique allowed computing residuals for all the spatial folds used as a validation subset. Then a nested variogram model was fitted for the variogram of the residuals of each soil property. A nugget model was fist fitted to describe spatial discontinuity at the origin and isotropic exponential model was used for each soil property.

2.7. Uncertainty quantification and validation

To model and quantify the uncertainty, a probabilistic framework was applied in this study. This is by far the most accepted approach and has the important advantage that it relies on solid statistical theory and that it can characterize uncertainty in great detail, including uncertainty dependencies in space, time and between soil properties (Heuvelink, 2014).

In the probabilistic framework to model and quantify the uncertainty at a prediction location, where the unknown value z(u) is considered as a realization of a random variable Z(u). The (cumulative) distribution function of the random variable Z(u) fully models the uncertainty because it gives the probability that the unknown is no greater than any given threshold z, that is

$$F(\mathbf{u}; z) = \text{Prob}\{Z(\mathbf{u}) \leq z\}$$

This equation is considered as the model of uncertainty at the prediction location u.

To model and quantify the uncertainty of the prediction, kriging variance approach was applied (Hengl et al., 2007). A parametric uncertainty model of UK was produced by means of UK prediction and its variance with the assumption of normality. For RFK, the kriging variance of the stochastic part of the variation and the RFK prediction described above are used to model the uncertainty as in the case of UK (Vaysse and Lagacherie, 2017). However, this uncertainty model does not account for the uncertainty in estimating the deterministic component. Therefore, there is a need to assume that the varying local mean is exactly equal to the RF prediction. In addition, we must assume normality to be able to construct a parametric model of uncertainty (Szatmári and Pásztor, 2019).

In order to spatially explicitly visualize the uncertainty of a spatial prediction, map of the upper and lower limit of the 90% prediction interval (PI) is considered (Arrouays et al., 2014; Heuvelink, 2014). This PI reports the range of values within which the true value is expected to occur 9 times out of 10. If the uncertainty model is parametric and normally distributed, the lower and upper limit of the 90% PI can be readily computed by subtracting and adding 1.64 times the prediction standard deviation to the prediction.

The uncertainty models were validated using independent and randomly selected control points (20% of the reference samples). For this, the actual fraction of true values falling within symmetric PIs of varying width p was calculated. The PIs can be readily derived by (1-p)/2 and (1+p)/2 quantiles of the distribution function (Coburn et al., 2020). For a set of control data, the fraction is computed by

$$\bar{\xi}(p)=\frac{1}{m}\sum_{i=1}^{m}\xi(u_i;p)\;\forall\, p\in[0,1]$$

$$\xi(u_i;p)=\begin{cases}1, if\ z(u_i)\in[p_{lower},p_{upper}]\\ 0, otherwise\end{cases}$$

where is the fraction for the PI of width p, $\xi(u_i;p)$ is the indicator function, $z(u_i)$ is the true value, p_{upper} is the upper limit of PI, p_{lower} is the lower limit of PI and m is the number of control points. A graphical way to check the performance of the uncertainty models is to plot against p that is frequently referred to as accuracy plot (Deutsch, 1999), but also known as prediction interval coverage probability (PICP) plot (Malone et al., 2011). Ideally, the observed fractions are equal to the expected fractions. If they are lower than the expectations, then the uncertainty has been underestimated. If they are higher, the uncertainty has been too liberally estimated (i.e. overestimated).

3. Results

3.1. Descriptive statistics of soil properties

The descriptive statistics of the soil properties and Spearman's Rho correlation coefficients between them are reported in Table 2 and Fig. 2, respectively.

In respect to the variation in topography, the SOC content showed moderate variability (CV = 34.8%), ranging from 1.24 to 6.75%. Its mean and median values were similar (3.30 and 3.08, respectively), and skewness value (0.67) indicated that the values are relatively evenly distributed on both sides of the mean, typically but not necessarily implying a symmetric distribution (Table 2). The sand content showed a wide range (from 9 to 78%) as opposed to the silt and clay contents. Therefore, its SD (±17) and CV (58.6%) values were considerably higher than the corresponding values of the silt and clay. The mean values of the sand, silt, and clay contents were 29, 50, and 20%, respectively that apparently presented the predominance of silt fraction. Accordingly, the soils were prevalently silty-clay, silty-clay-loam, and partially clay-loam in texture. The $CaCO_3$ content is largely variable ranging from 0% to 8.72% with a mean value of 2.88%. The high variability of $CaCO_3$ content (CV = 83.6%) is attributed by the topography and distribution the carbonate-rich parent material in the study area. The soil reaction (pH_{H2O}) is dominantly neutral, with a mean value of 7.25. The values of pH_{H2O} >7.0 account for 71% of all samples in the soils of the test site. The pH level of soils in both H_2O and KCl solution showed low variability. Similar to the SOC, the WC ranged from 1.70 to 6.67%, with a mean value of 3.73 and moderate variability (CV = 26.5%). The skewness and kurtosis values of the tested soil properties imply approximately symmetric and light-tailed distributions.

Table 2 Descriptive statistics of tested soil properties.

Spearman's Rho correlation coefficients showed significant relationships between the soil properties. Both positive and negative correlation coefficients were found between the tested soil properties except the WC that was significantly correlated with only the SOC ($r = 0.46$) and $CaCO_3$ ($r = -0.25$) (Fig. 2). The SOC content was negatively correlated with all other soil properties except sand. As expected, the sand content was very strongly and inversely correlated with the silt ($r = -0.91$) and clay content ($r = -0.80$). However, the silt, clay and $CaCO_3$ contents as well pH_{H2O} and pH_{KCl} were positively correlated with each other. These relations are probably attributed to the predominating carbonate-rich parent material from which the silt and partially clay fractions originate. Therefore, the $CaCO_3$ content was strongly correlated with the silt content ($r = 0.63$). However, the correlation between $CaCO_3$ and clay content was weak ($r = 0.37$). As the $CaCO_3$ content dramatically affect soil reaction, the former was very

strongly correlated with pH_{H2O} (r = 0.75) and pH_{KCl} (r = 0.85). In addition, the relation between the silt and clay, pH_{H2O} and pH_{KCl} was expected.

Fig. 2. Spearman's Rho correlation coefficients between tested soil properties. White background cells donate insignificant correlations, P > 0.01.

The relationship between SOC and sand, SOC and clay is attributed by the characteristic distribution of carbonate-rich parent material at low elevations and quartz-rich bedrock at higher elevations. For this reason, both sand and SOC contents increase with elevation, while the $CaCO_3$ content decreases in the study area. Accordingly, the $CaCO_3$ content was found negatively correlated (r =−0.60) with SOC (Fig. 2).

3.2. Relation between soil properties and auxiliary variables

Fig. 3 summarizes the relationships between soil properties and auxiliary variables (terrain attributes and spectral indices) using Spearman's Rho correlation coefficients. For the SOC content, the highest positive correlation was found with TGSI (r = 0.57) followed by EL (r = 0.50). Furthermore, its positive relation was also found with SL (r = 0.37) while the highest negative correlation coefficient was found with NIR band (r =−0.52) followed by TWI (r =−0.41). Other auxiliary variables showed very weak relationships with SOC.

Fig. 3. Spearman's Rho correlation coefficients between tested soil properties and auxiliary variables.

The relationship between the sand content and auxiliary variables was slightly weaker but very similar to those of with the SOC, that was positive with EL (r = 0.48), SL (r = 0.41) and TGSI (r = 0.37), and negative with TWI (r =−0.34) and NIR (r =−0.38) (Fig. 3). Unlike the sand content, the silt and clay contents showed mostly similar and negative relationships with auxiliary variables. For the silt content, the highest correlation coefficient was found with EL (r =−0.56) followed by SL (r =−0.37), while its content was positively correlated with TWI (r = 0.32) and NIR (r = 0.32). Likewise, the clay content showed negative relation with EL (r =−0.41), SL (r =−0.33) and TGSI (r =−0.28), and positive relation with TWI (r = 0.28) and NIR (r = 0.26).

The $CaCO_3$ content showed the highest negative correlation coefficient with EL (r =−0.60) followed by TGSI (r =−0.30), while its considerable positive relation was found with NIR band (Fig. 3). Since the soil reaction (both pH_{H2O} and pH_{KCl}) is mostly dependent on the $CaCO_3$ content (as described in the previous § 3.1), the behavior

of the former with auxiliary variables was similar. Among the tested soil properties, the WC showed the lowest relation with auxiliary variables. Overall, its content showed weak dependency on SL and TGSI.

3.3. Performances of prediction models

The quality parameters of the tested models are presented in Table 3. Based on the model accuracy, UK was negligibly better than RFK for the SOC, pH_{H2O} and pH_{KCl} prediction. However, in terms of bias RFK was more reliable for the mentioned soil properties with the ME values of -0.02, 0.03 and 0.04, respectively. RFK model produced slightly more accurate prediction for the sand, clay, $CaCO_3$ and WC with RMSE values of 15.0, 6.5, 2.00 and 0.95, respectively. For the silt content, although both UK and RFK produced the same prediction accuracy (RMSE = 10.0) as well lower ME value of UK model, RFK was more stable model based on other metrics (SD_{RMSE}, SD_{ME} and SD_{RPD}).

Overall, the performance of both UK and RFK was very similar for the corresponding soil properties. Among the tested soil properties, the highest RPD value was found for the SOC prediction model (Table 3). The computed biases of both the models were positive for the silt, $CaCO_3$, pH_{H2O} and pH_{KCl} that indicated the soil properties were slightly overestimated. Contrarily, both the models underestimated clay and WC. The SOC and sand content was slightly overestimated by UK while underestimated by RFK.

Table 3 Performances of UK and RFK models of the tested soil properties.

The histograms of the computed residuals of UK and RF are presented in Fig. 4 and Fig. 5, respectively. The residuals showed no outlier values which is appropriate for the required normality assumption. Regardless of the model, the residuals of SOC, sand, clay, pH_{H2O}, pH_{KCl} and WC apparently showed symmetric distribution while others, such as silt and $CaCO_3$ showed negatively and positively skewed distribution, respectively. This may appear inappropriate for the normality assumption and affect uncertainty models.

Fig. 6 and Fig. 7 present omnidirectional variograms of the residuals of UK and RF and their corresponding fitted models, respectively. In addition, the parameters of the fitted models for the residuals of UK and RF are presented in Table 4. The range values considerably varied with the predicted soil properties as well depending on the deterministic model. In general, the range values of the fitted models of all soil properties (except sand and pH_{H2O}) for UK were significantly lower than that of RF. Correspondingly, the nugget to sill ratio values (except the case for the sand and $CaCO_3$) for UK were considerably higher than the corresponding values of RF model. In fact, such

high values are often observed in DSM (Hengl et al., 2015; Szatmári and Pásztor, 2019; Vaysse and Lagacherie, 2017). The higher range values of the residuals of RF confirm spatial continuity at longer distances that may appropriately affect uncertainty modeling.

Finally, for both UK and RFK the spatial prediction map of each soil property as well as lower and upper limit of the 90% PI was presented in Fig. 8. For all predicted soil properties, both UK and RFK was capable of depicting spatial pattern, in particular elevation dependent pattern as well as local patterns attributed by the variations in vegetation and land use. However, in terms of lower and upper prediction limits RFK model was found more stable in all cases.

Table 4 The semivariogram parameters of the best fitted models for the residuals of UK and RF.

3.4. Quantification and validation of uncertainty models

The PICP plots of the predicted soil properties were presented in Fig. 9 for UK and Fig. 10 for RFK. Uncertainty model for UK substantially overestimated the SOC and $CaCO_3$ whereas the uncertainty of clay content and partially pH_{H2O} was underestimated. However, the models for sand, silt, pH_{KCl} and WC were comparatively better. Similarly, the uncertainty of RFK severely overestimated $CaCO_3$ and underestimated clay content. Furthermore, slight underestimations were observed

Fig. 4. Histograms of the residuals of soil properties calculated based on UK.

Fig. 5. Histograms of the residuals of soil properties calculated based on RF.

Fig. 6. Variograms and best fitted models for the UK residuals.

Fig. 7. Variograms and best fitted models for the RF residuals.

for the uncertainties of sand, silt and pH_{H2O} whereas we assume that their effect on model quality was negligible. In some cases, unusual patterns in the PICP plots were found e.g., for the silt and pH_{KCl}. In the case of RFK the pH_{KCl} was slightly overestimated in lower confidence levels and underestimated in higher confidence levels. An inverse case was observed for the silt content when its uncertainty was explained by UK. Another unusual pattern arisen that the computed fractions revealed the same expectation values in different confidence levels depending on the soil properties and model applied.

Overall, the uncertainty models of RFK were better than those of UK for most of the soil properties except the cases for $CaCO_3$, pH_{KCl} and partially clay, of which the computed expectations are closer to x=y line than in RFK.

3.5. Important variables

As described above, RFK model was more reliable especially in terms of uncertainty quantification and validation for predicting most of the soil properties (six out of eight soil properties). Therefore, the actual impurity reduction importance metric was used to explain the impact of each auxiliary variable on the models' outputs. Fig. 11 shows the contribution of the auxiliary variables to RF model predictions.

As the important variables, both the terrain attributes and spectral indices contributed differently to the model output. Generally, from the terrain attributes EL, SL, TWI and ASc, whereas from the spectral indices TGSI and NIR were found to be most important variables. As regard the detailed interpretation of the important variables for each soil property, those considerably changed with the predicted soil property. For the SOC prediction model, the spectral indices such as TGSI and NIR contributed more than the terrain attributes that were EL, SL and TWI (Fig. 11). Contrarily, for the sand, silt and clay prediction EL and SL were the most important predictors followed by TGSI, TWI and NIR band with varying sequences depending on the particle size.

A different sequence of important variables appeared for the $CaCO_3$, pH_{H2O} and pH_{KCl}. Hence, EL was the most important variable with quite high partial importance score values while other auxiliary variables showed very low and close score values (Fig. 11). In addition, ASc was found to be more important variable (3^{rd} important variable for $CaCO_3$ and 2^{nd} for pH) than others while its importance for predicting the other soil properties was rather low. Unlike all the soil properties, for the WC prediction SL and TGSI were the most important predictors with quite close scores.

Overall, EL, SL, TGSI, NIR band, TWI and partially ASc and SAVI were the predictors contributing most to the prediction of all tested soil properties, particularly EL was the most important predictor for 6 out of 8 predicted soil properties. Therefore, the SOC and sand content showed a positive trend with EL while $CaCO_3$, silt, clay and pH showed a negative trend.

Fig. 8. Predicted map of the soil properties and their corresponding upper and lower limit of the 90% prediction interval. The unit of the maps is [%] for the all soil properties except pH_{H2O} and pH_{KCl}

Fig. 9. PICP plots of the soil properties predicted with UK.

Fig. 10. PICP plots of the soil properties predicted with RFK.

Fig. 11. Importance of covariates in RFK model (vertical axis: predictor variables and horizontal axis: partial importance score).

4. Discussions

4.1. Relationship between soil properties and auxiliary variables

The Spearman's Rho correlation coefficients between the tested soil properties and auxiliary variables showed that considerable correlations existed with EL, SL, TWI, TGSI and NIR band (Fig. 3). Furthermore, the actual impurity reduction importance metric of RF also showed that those auxiliary variables contributed most to the model outputs.

As typical to the Mediterranean climate, the precipitation increases, and air temperature decreases with EL both of which determine soil moisture and temperature regime in the study area. Therefore, EL was found as the most important terrain attribute for the prediction of most soil properties (Fig. 11). Among the terrain attributes, SL is one of the main predictors representing erosion, accumulation processes, water movement and soil moisture content in the landscapes. In our study area, the sand content increases while the clay and silt content decrease with SL gradient. Correspondingly, SL was apparently more important variable for predicting the sand, silt and clay content as well as the most important variable for the WC. In association with SL, aspect (ASc and ASs) affects insolation and evapotranspiration and consequently water balance in the soil (Böhner and Antonić, 2009). Among the tested soil properties, the spatial variation of $CaCO_3$ content and pH was significantly controlled by aspect, particularly the soil of south, south-east and south-west faced slopes is characterized by higher $CaCO_3$ content and alkaline environment. This was also confirmed by the relationships between $CaCO_3$ and pH_{KCl}, pH_{H2O} (Fig. 2). Likewise, the importance scores of ASc for predicting these soil properties were considerably higher than the other auxiliary variables. One of the top 5 important auxiliary variables for predicting most of the soil properties was TWI. This represents the tendency of water and other materials to accumulate in lower areas, driven by the combined effects of the specific catchment area and slope (Gallant and Austin, 2015). Although TWI was originally formulated as an index of depth to water table, it has been found to be an effective predictor of different soil properties and soil depth (Moore et al., 1993; Sumfleth and Duttmann, 2008) in supply-limited landscapes where the capacity to transport sediment downslope exceeds the soil production rate (Gessler et al., 1995). In these landscapes,

the soil depth tends to increase systematically down the slope, with thin soils in the upper slopes and thick soils in the lower slopes, a pattern that is well represented by TWI. The correlation coefficients between TWI and soil properties (Fig. 3) indicate that SOC, sand and WC tend to be high on the upper catchment positions and low on foot slopes whereas the $CaCO_3$, silt and clay content was contrariwise.

From spectral indices TGSI and NIR band considerably contributed to the prediction of the soil properties. TGSI is associated with physical properties (mechanical composition) of topsoil, indicating coarsening of topsoil grain size. Xiao et al., (2006) proposed TGSI to evaluate actual degree of desertification. A high TGSI value corresponds to the high content of fine sand in the topsoil and contrarily low content of clay and silt. Correspondingly, in our study case the sand content was positively, silt and clay content were negatively correlated with TGSI. Apart from the contribution of TGSI to the prediction of particle sizes, that was found to be one of the most important predictors particularly for the SOC, WC, $CaCO_3$ and pH_{H2O}. As regard NIR band, the soil properties showed both negative and positive correlation coefficients with NIR band (Fig. 3). According to Ben-Dor, (1999) soil reflectance decreases with increasing soil organic matter content in the visible near-infrared (350-2500 nm) wavelength range. In consistency with this, the highest correlation coefficient between NIR band and the soil properties was found for the SOC ($r = -0.52$). Therefore, NIR band contributed to the prediction of SOC more than its contribution to the other soil properties. The reflectance values of NIR band in arable lands of our study area were partially high due to bare and dry soil surface, as well as more optimal illumination of arable lands that are mostly allocated in gentle slopes. Furthermore, the reflectance of pastures that are characterized by sparse and dry grass (at the time of image acquisition and sampling) was also lower in NIR band in association with the fragmented shrubbery. The negative relationship between NIR band and SOC content indicates that arable lands and less vegetated areas tend to have lower SOC than the pastures shrubby lands.

Summarily, the correlative relations between the soil properties and auxiliary variables and the actual impurity reduction importance metric of RF model showed that terrain attributes contribute to the model outputs more than spectral indices. Many studies pointed out that most soil properties are correlated with terrain attributes (Bagheri Bodaghabadi et al., 2015; Sumfleth and Duttmann, 2008; Taylor et al., 2013). Dobos and Hengl, (2009) summarizes several previous studies using terrain parameters to predict soil properties noting the utility of TWI, SL, curvature and flow accumulation as commonly used covariates.

4.2. Performances of prediction and uncertainty models

The performances of the tested models were very similar especially in terms of prediction accuracy (Table 3). The accuracy of UK was slightly higher than RFK for SOC, pH_{H2O} and pH_{KCl} while RFK was more accurate for the other soil properties. Furthermore, when considering the computed biases, RFK resulted in the ME values close to zero for the most of soil properties except the silt content. These indicate that RFK can give better spatial prediction according to error measure and prediction accuracy and a suitable technique to estimate trend. This is also in agreement with the studies by Hengl et al., (2015) and Szatmári and Pásztor, (2019) that stated RF often outperforms the most commonly applied trend estimation techniques. The reason for RFK provides the most accurate prediction at unsampled location is that local error variance is minimized when the residuals are modelled by kriging (Webster and Oliver, 2008). Likewise, many study examples reported advantages of using machine learning and hybrid spatial models integrated with high-resolution remote sensing data. Tziachris et al., (2019) assessed predictive capability of different models for predicting soil organic matter and found that RF in combination with kriging of the residuals substantially increased the accuracy of prediction. Similarly, Chen et al., (2019) compared geostatistical, machine learning and hybrid approaches to predict SOC over an area of 3,600 km^2. They found that machine learning approaches overall outperformed geostatistical methods, and the performances of two-step hybrid models were more accurate than their corresponding model. Moreover, Xu et al., (2018) evaluated the spatial resolution effect of different satellite images to predict total nitrogen content in two test areas using regression kriging. They found a similar spatial pattern of total nitrogen regardless of the spatial resolution of the images. However, the highest prediction accuracy was obtained with the use of higher spatial resolution images (WorldView-2, Pléiades-1, and GeoEye-1) in both study areas. WorldView-2 particularly outperformed others due to its more spectral bands and finer spatial resolution. The spectral indices based on new bands of WorldView-2 such as coastal, yellow, red edge, and new near infrared bands showed relatively strong correlations with soil total nitrogen content.

As regard the validation results of the uncertainty models, both UK and RFK approach provided fairly contrasting results showing significant underestimation and overestimation especially for the residuals of SOC (in the case of UK), clay and $CaCO_3$ content (Fig. 9 and Fig. 10). The uncertainty model created based on the residuals of UK for the clay, pH_{KCl} and $CaCO_3$ was better than that of RF which gave better estimation for the rest of soil properties. Aside from underestimation and overestimation, the validation results of residuals showed slightly changeable pattern with confidence levels. Not only underestimated and overestimated residuals, but also irregular

patterns noted arise from different factors including residuals in trend estimation. We assume that the number of samples and the sampling design (randomly collected in our case) can be importantly included to the main factors affecting the uncertainty models of the soil properties. In particularly, when the uncertainty model was validated with the 20% of the sampling points, the kriging of the residuals with 80% of sampling points most probably led to the underestimations or overestimations in view of lower number of samples representing heterogeneous mountainous area. This assumption maybe also supported by the semivariogram parameters that confirm short distance variations (Table 4, Fig. 6 and Fig. 7). It is especially important to emphasize that spatial continuity of $CaCO_3$ content in the surface layer is characterized with short distance variations (in some places hotspot pattern distributions) as well its absence at elevations approximately higher than 850 m above sea level. The short distance variations are also typical to the spatial distribution of clay content which arisen from e.g., slope, aspect and erosion rate. Therefore, the sampling density used in this may not properly capture the spatial pattern of those soil properties.

The uncertainty of the predicted soil properties of which the residuals showed compatibility with the normality assumption (Fig. 4 and Fig. 5) to a certain degree were reliably estimated by both the models (Fig. 9 and Fig. 10). On the other hand, the uncertainty of those soil properties was better modelled when the residuals showed apparent spatial correlation (Table 4, Fig. 6 and Fig. 7). Thus, the validation results of uncertainty models were more reliable when the residuals meet simultaneously normality assumption and show apparent spatial correlation.

On the other hand, as the kriging variance represents a general measure of local accuracy which is data-value independent, some studies do not recommend to apply this approach (Goovaerts, 2001; Szatmári and Pásztor, 2019). Because, the kriging variance reflects the position of unsampled locations in geographical space without any reflection about their position in feature space (Szatmári and Pásztor, 2019). Therefore, with respect to future digital soil mapping studies in this region, we assume that it is important to consider i) sampling design depending on the mapping scale, ii) testing other modelling approaches to predict soil properties and quantify their uncertainties, and iii) multi-scale digital terrain analysis as it can help increase prediction accuracy compared to standard digital terrain analysis (Smith et al. 2006, Behrens et al. 2010). As the validation of the uncertainty models showed significant sensitivity to the number of samples used in the validation subset, a special emphasis should be given to the sampling density/design to enhance model quality and reduce resources for field sampling and chemical tests.

4. Conclusions

In this case study, basic soil properties were predicted and mapped using high resolution DEM (ALOS-PALSAR 12.5 m) and MI (Pléiades-1) based on two hybrid spatial models, UK and RFK. Although the accuracy of the tested models was similar, RFK was superior in view of both accuracy and computed biases. When comparing the uncertainty of the prediction models, both the models produced considerable underestimations and overestimations depending on the soil property. RFK provided better uncertainty estimation for the most of soil properties than UK of which the uncertainty estimation was more appropriate for the clay and pH_{KCl}. As the residuals of $CaCO_3$ prediction models showed very low spatial correlation, both UK and RFK severely overestimated its content, whereas UK was considerably better. While the residuals were compatible with the normality assumption and showed apparent spatial correlation, the uncertainty models provided more reliable prediction intervals that confirmed the importance of assumptions made in uncertainty modelling and quantification in agreement with previous studies.

The procedure presented in this study can be reliably used to map the studied soil properties and extended to fairly larger adjacent areas characterized by similar environmental condition. Moreover, this study is a contribution to decision making and land management in the mountainous area where high resolution soil information is in demand. However, with respect to future digital soil mapping studies in this region, we assume that it is important to consider i) sampling design depending on the mapping scale, ii) testing other modelling approaches to predict soil properties and quantify their uncertainties, and iii) multi-scale digital terrain analysis as it can help increase prediction accuracy compared to standard digital terrain analysis.

The tested hybrid spatial models showed that high resolution remote sensing data are promising for mapping soil properties in the given sampling density in the foothill belt of the Caucasus Mountains. Particularly, RFK was capable of delineating elevation-dependent spatial pattern as well as the local patterns of the soil properties attributed by variations in vegetation, land use and soil erosion. UK model produced a few local erratic spatial patterns (e.g., in the case of pH) corresponding to the artifacts such as roads and houses in the image that should be considered in future applications.

Acknowledgements

Funding: This study was supported by the Islamic Development Bank, Merit Scholarship Program for High Technology [36/11209317] and National Academy of Sciences of the Republic of Azerbaijan.

The authors are especially thankful to all members of the Department of Remote Sensing of Environment for supporting laboratory analyses of soils at Adam Mickiewicz University in Poznan, Poland.

References

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. https://doi.org/10.1029/WR023i009p01717

Alaska-Satellite-Facility, 2019. ALOS-PALSAR [WWW Document]. URL https://asf.alaska.edu/sar-data-sets/alos-palsar/

Arrouays, D., McBratney, A.B., Minasny, B., Hempel, J.W., Heuvelink, G.B.M., MacMillan, R.A., Hartemink, A.E., Lagacherie, P., McKenzie, N.J., 2014. The GlobalSoilMap project specifications, in: GlobalSoilMap: Basis of the Global Spatial Soil Information System - Proceedings of the 1st GlobalSoilMap Conference. https://doi.org/10.1201/b16500-4

Babayev et. al., 2017. An integration of the national soil classification of Azerbaijan to Worls Reference Base (in Azerb.).

Bagheri Bodaghabadi, M., MartÍnez-Casasnovas, J., Salehi, M.H., Mohammadi, J., Esfandiarpoor Borujeni, I., Toomanian, N., Gandomkar, A., 2015. Digital Soil Mapping Using Artificial Neural Networks and Terrain-Related Attributes. Pedosphere 25, 580–591. https://doi.org/10.1016/S1002-0160(15)30038-2

Bayramov et. al., 2008. Geological map of Azerbaijan Republic (1:500,000).

Bazaglia Filho, O., Rizzo, R., Lepsch, I.F., Prado, H. do, Gomes, F.H., Mazza, J.A., Demattê, J.A.M., 2013. Comparison between detailed digital and conventional soil maps of an area with complex geology. Rev. Bras. Ciência do Solo 37, 1136–1148. https://doi.org/10.1590/s0100-06832013000500003

Behrens, T., Zhu, A.X., Schmidt, K., Scholten, T., 2010. Multi-scale digital terrain analysis and feature selection for digital soil mapping. Geoderma 155, 175–185. https://doi.org/10.1016/j.geoderma.2009.07.010

Ben-Dor, E., 1999. Soil reflectance, in: Rencz, A.N. (Ed.), Remote Sensing for the Earth Sciences: Manual of Remote Sensing. Wiley, New York, pp. 111–188.

Biecek, P., 2018. Dalex: Explainers for complex predictive models in R. J. Mach. Learn. Res. 19, 1–5.

Böhner, J., Antonić, O., 2009. Land-surface parameters specific to topo-climatology, in: Heng, T., Reuter, H. (Ed.), Developments in Soil Science. Elsevier, Amsterdam, pp. 195–226. https://doi.org/10.1016/S0166-2481(08)00008-1

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324

Brenning, A., 2012. Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: The R package sperrorest. Int. Geosci. Remote Sens. Symp. https://doi.org/10.1109/IGARSS.2012.6352393

Brenning, A., 2008. Statistical geocomputing combining R and SAGA: The example of landslide susceptibility analysis with generalized additive models. SAGA--Seconds Out (hamburg. Beiträge zur Phys. Geogr. und Landschaftsökologie, vol. 19).

Cambule, A.H., Rossiter, D.G., Stoorvogel, J.J., Smaling, E.M.A., 2014. Soil organic carbon stocks in the limpopo national park, mozambique: Amount, spatial distribution and uncertainty. Geoderma 213, 46–56. https://doi.org/10.1016/j.geoderma.2013.07.015

Cavazzi, S., Corstanje, R., Mayr, T., Hannam, J., Fealy, R., 2013. Are fine resolution digital elevation models always the best choice in digital soil mapping? Geoderma 195–196, 111–121. https://doi.org/10.1016/j.geoderma.2012.11.020

Chen, L., Ren, C., Li, Lin, Wang, Y., Zhang, B., Wang, Z., Li, Linfeng, 2019. A comparative assessment of geostatistical, machine learning, and hybrid approaches for mapping topsoil organic carbon content. ISPRS Int. J. Geo-Information 8, 1–18. https://doi.org/10.3390/ijgi8040174

Coburn, T.C., Yarus, J.M., Chambers, R.L., 2020. Geostatistical Modeling of the Spaces of Local, Spatial, and

Response Uncertainty for Continuous Petrophysical Properties. Stoch. Model. Geostatistics 59–79. https://doi.org/10.1306/1063807ca53229

Collard, F., Kempen, B., Heuvelink, G.B.M., Saby, N.P.A., Richer De Forges, A.C., Lehmann, S., Nehlig, P., Arrouays, D., 2014. Refining a reconnaissance soil map by calibrating regression models with data from the same map (Normandy, France). Geoderma Reg. 1, 21–30. https://doi.org/10.1016/j.geodrs.2014.07.001

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Böhner, J., 2015. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. Geosci. Model Dev. https://doi.org/10.5194/gmd-8-1991-2015

Deutsch, C. V., 1999. Direct assessment of local accuracy and precision. Geostatistics Wollongong 96 - Proc. Fifth Int. Geostatistics Congr. Wollongong, Aust. Sept. 1996.

Dobos and Hengl, 2009. Soil mapping applications, in: Hengl, T. and Reuter, H.I. (Ed.), Geomorphometry: Concepts, Software, Applications. Amsterdam, pp. 461–479. https://doi.org/10.1016/S0166-2481(08)00036-6

Elbasiouny, H., Abowaly, M., Abu-Alkheir, A., Gad, A.A., 2014. Spatial variation of soil carbon and nitrogen pools by using ordinary Kriging method in an area of north Nile Delta, Egypt. Catena 113, 70–78. https://doi.org/10.1016/j.catena.2013.09.008

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. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000208

Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. Ann. Stat. 11, 12–32. https://doi.org/10.2307/2699986

Gallant, J.C., Austin, J.M., 2015. Derivation of terrain covariates for digital soil mapping in Australia. Soil Res. 53, 895–906. https://doi.org/10.1071/SR14271

Gessler, P.E., Moore, I.D., McKenzie, N.J., Ryan, P.J., 1995. Soil-landscape modelling and spatial prediction of soil attributes. Int. J. Geogr. Inf. Syst. 9, 421–432. https://doi.org/10.1080/02693799508902047

Goovaerts, P., 2001. Geostatistical modelling of uncertainty in soil science. Geoderma 103, 3–26. https://doi.org/10.1016/S0016-7061(01)00067-2

Greenwell, B.M., 2017. pdp: An R package for constructing partial dependence plots. R J. 9, 21–36. https://doi.org/10.32614/rj-2017-016

Hengl, T., Heuvelink, G.B.M., Kempen, B., Leenaars, J.G.B., Walsh, M.G., Shepherd, K.D., Sila, A., MacMillan, R.A., De Jesus, J.M., Tamene, L., Tondoh, J.E., 2015. Mapping soil properties of Africa at 250 m resolution: Random forests significantly improve current predictions. PLoS One 10, 1–26. https://doi.org/10.1371/journal.pone.0125814

Hengl, T., Heuvelink, G.B.M., Rossiter, D.G., 2007. About regression-kriging: From equations to case studies. Comput. Geosci. 33, 1–15. https://doi.org/10.1016/j.cageo.2007.05.001

Hengl, T., MacMillan, R.A., Nikolić, M., 2013. Mapping efficiency and information content. Int. J. Appl. Earth Obs. Geoinf. 22, 127–138. https://doi.org/10.1016/j.jag.2012.02.005

Heuvelink, G., Malone, B., 2020. Package ‘ GSIF .’

Heuvelink, G.B.M., 2014. Uncertainty quantification of globalsoilmap products. Glob. Basis Glob. Spat. Soil Inf. Syst. - Proc. 1st Glob. Conf. 335–340. https://doi.org/10.1201/b16500-62

Heuvelink, G.B.M., Webster, R., 2001. Modelling soil variation: Past, present, and future. Geoderma 100, 269–301. https://doi.org/10.1016/S0016-7061(01)00025-8

Hijmans, R.J., 2016. Raster: Geographic Data Analysis and modeling, R package version 2.5-8. https://CRAN.R-project.org/package=raster. https://doi.org/10.1375/bech.25.1.1

Horn, B.K.P., 1981. Hill Shading and the Reflectance Map. Proc. IEEE 69, 14–47. https://doi.org/10.1109/PROC.1981.11918

Houborg, R., McCabe, M.F., 2018. A hybrid training approach for leaf area index estimation via Cubist and random forests machine-learning. ISPRS J. Photogramm. Remote Sens. 135, 173–188. https://doi.org/10.1016/j.isprsjprs.2017.10.004

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309. https://doi.org/10.1016/0034-4257(88)90106-X

Japan Aerospace Exploration Agency, 2012. JAXA Advanced Land Observing Satellite DAICHI (ALOS) [WWW Document]. Japan Aerosp. Explor. Agency Website.

Keskin, H., Grunwald, S., 2018. Regression kriging as a workhorse in the digital soil mapper’s toolbox. Geoderma 326, 22–41. https://doi.org/10.1016/j.geoderma.2018.04.004

Kim, D., Zheng, Y., 2011. Scale-dependent predictability of DEM-based landform attributes for soil spatial variability in a coastal dune system. Geoderma 164, 181–194. https://doi.org/10.1016/j.geoderma.2011.06.002

Klute, A., 1988. Methods of Soil Analysis 2d ed., pt. 1; Physical and Mineralogical Methods. Soil Sci. 383–411. https://doi.org/10.1097/00010694-198808000-00014

Kramm, T., Ho, D., 2019. A Relief Dependent Evaluation of Digital Elevation Models on Different Scales for Northern Chile. ISPRS Int. J. Geo-Information 8, 1–25. https://doi.org/10.3390/ijgi8100430

Lal, R., 2004. Soil carbon sequestration to mitigate climate change. Geoderma 123, 1–22. https://doi.org/10.1016/j.geoderma.2004.01.032

Laurencelle, J., Logan, T., Gens, R., 2015. ASF Radiometrically Terrain Corrected ALOS PALSAR products. Product guide, revision 1.2. Alaska Satell. Facil.

Leutner, B., Horning, N., 2017. Package 'RStoolbox.' R Found. Stat. Comput.

Li, J., Heap, A.D., Potter, A., Daniell, J.J., 2011. Application of machine learning methods to spatial interpolation of environmental variables. Environ. Model. Softw. 26, 1647–1659. https://doi.org/10.1016/j.envsoft.2011.07.004

Lian G, Guo XD, Fu BJ, H.C., 2009. Prediction of the spatial distribution of soil properties based on environmental correlation and geostatistics. Trans. Chin. Soc. Agric. Eng 25, 112–122. https://doi.org/10.3969/j.issn.1002-6819.2009.07.043

Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R News 23, 18–22.

Malone, B.P., McBratney, A.B., Minasny, B., 2011. Empirical estimates of uncertainty for mapping continuous depth functions of soil attributes. Geoderma. https://doi.org/10.1016/j.geoderma.2010.11.013

Matheron, G., 1963. Principles of geostatistics. Econ. Geol. 58, 1246–1266. https://doi.org/10.2113/gsecongeo.58.8.1246

McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping, Geoderma. https://doi.org/10.1016/S0016-7061(03)00223-4

McBratney, A.B., Odeh, I.O.A., Bishop, T.F.A., Dunbar, M.S., Shatar, T.M., 2000. An overview of pedometric techniques for use in soil survey. Geoderma 97, 293–327. https://doi.org/10.1016/S0016-7061(00)00043-4

McKenzie, N.J., Ryan, P.J., 1999. Spatial prediction of soil properties using environmental correlation. Geoderma 89, 67–94. https://doi.org/10.1016/S0016-7061(98)00137-2

Minasny, B., McBratney, A.B., 2016. Digital soil mapping: A brief history and some lessons. Geoderma 264, 301–311. https://doi.org/10.1016/j.geoderma.2015.07.017

Moore, I.D., Gessler, P.E., Nielsen, G.A., Peterson, G.A., 1993. Soil attribute prediction using terrain analysis. Soil Sci. Soc. Am. J. 57, 443–452. https://doi.org/10.2136/sssaj1993.03615995005700020026x

Myles, A.J., Feudale, R.N., Liu, Y., Woody, N.A., Brown, S.D., 2004. An introduction to decision tree modeling. J. Chemom. https://doi.org/10.1002/cem.873

Nembrini, S., König, I.R., Wright, M.N., 2018. The revival of the Gini importance? Bioinformatics 34, 3711–3718. https://doi.org/10.1093/bioinformatics/bty373

Neteler, M., Mitasova, H., 2008. Open source GIS: A GRASS GIS approach, Open Source GIS: A GRASS GIS Approach. https://doi.org/10.1007/978-0-387-68574-8

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. https://doi.org/10.1016/0016-7061(95)00007-B

Pebesma, E.J., 2004. Multivariable geostatistics in S: The gstat package. Comput. Geosci. 30, 83–91. https://doi.org/10.1016/j.cageo.2004.03.012

Pendleton, R.L., Jenny, H., 1945. Factors of Soil Formation: A System of Quantitative Pedology, Geographical Review. Dover Publications. https://doi.org/10.2307/211491

Phachomphon, K., Dlamini, P., Chaplot, V., 2010. Estimating carbon stocks at a regional level using soil information and easily accessible auxiliary variables. Geoderma 155, 372–380. https://doi.org/10.1016/j.geoderma.2009.12.020

Pouladi, N., Møller, A.B., Tabatabai, S., Greve, M.H., 2019. Mapping soil organic matter contents at field level with Cubist, Random Forest and kriging. Geoderma 342, 85–92. https://doi.org/10.1016/j.geoderma.2019.02.019

R Development Core Team, 2019. : A language and environment for statistical computing. R Found. Stat. Comput. Vienna, Austria. URL http//www.R-project.org/.

Salayev et. al., 1990. State Soil Map of Azerbaijan (1:100,000).

Salukvadze, J., Medvedkov, O., 2011. Land governance in the South Caucasus Region: Comparative Study of Georgia, Armenia and Azerbaijan 18–22.

Samuel-Rosa, A., Heuvelink, G.B.M., Vasques, G.M., Anjos, L.H.C., 2015. Do more detailed environmental covariates deliver more accurate soil maps? Geoderma 243–244, 214–227. https://doi.org/10.1016/j.geoderma.2014.12.017

Sandri, M., Zuccolotto, P., 2008. A bias correction algorithm for the gini variable importance measure in classification trees. J. Comput. Graph. Stat. 17, 11–28. https://doi.org/10.1198/106186008X344522

Schratz, P., 2019. mlr3spatiotempcv-package: mlr3spatiotempcv: Spatial support for mlr3.

Shi, W., Liu, J., Du, Z., Song, Y., Chen, C., Yue, T., 2009. Surface modelling of soil pH. Geoderma 150, 113–119. https://doi.org/10.1016/j.geoderma.2009.01.020

Smith, M.P., Zhu, A.X., Burt, J.E., Stiles, C., 2006. The effects of DEM resolution and neighborhood size on digital soil survey. Geoderma 137, 58–69. https://doi.org/10.1016/j.geoderma.2006.07.002

Sparks, D.L., Page, A.L., Helmke, P.A., Loeppert, R.H., Nelson, D.W., Sommers, L.E., 1996. Total Carbon, Organic Carbon, and Organic Matter. https://doi.org/10.2136/sssabookser5.3.c34

Subburayalu, S.K., Slater, B.K., 2013. Soil series mapping by knowledge discovery from an Ohio county soil map. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj2012.0321

Sumfleth, K., Duttmann, R., 2008. Prediction of soil property distribution in paddy soil landscapes using terrain data and satellite information as indicators. Ecol. Indic. 8, 485–501. https://doi.org/10.1016/j.ecolind.2007.05.005

Szatmári, G., Pásztor, L., 2019. Comparison of various uncertainty modelling approaches based on geostatistics and machine learning algorithms. Geoderma 337, 1329–1340. https://doi.org/10.1016/j.geoderma.2018.09.008

Taylor, J.A., Jacob, F., Galleguillos, M., Prévot, L., Guix, N., Lagacherie, P., 2013. The utility of remotely-sensed vegetative and terrain covariates at different spatial resolutions in modelling soil and watertable depth (for digital soil mapping). Geoderma 193–194, 83–93. https://doi.org/10.1016/j.geoderma.2012.09.009

Tziachris, P., Aschonitis, V., Chatzistathis, T., Papadopoulou, M., 2019. Assessment of spatial hybrid methods for predicting soil organic matter using DEM derivatives and soil parameters. Catena 174, 206–216. https://doi.org/10.1016/j.catena.2018.11.010

Vaysse, K., Lagacherie, P., 2017. Using quantile regression forest to estimate uncertainty of digital soil mapping products. Geoderma 291, 55–64. https://doi.org/10.1016/j.geoderma.2016.12.017

Veronesi, F., Schillaci, C., 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. https://doi.org/10.1016/j.ecolind.2019.02.026

Wackernagel, H., 1998. Multivariate Geostatistics: An Introduction With Applications, 2nd ed. Springer, Berlin.

Webster, R., Oliver, M.A., 2008. Geostatistics for Environmental Scientists: Second Edition, Geostatistics for

Environmental Scientists: Second Edition. https://doi.org/10.1002/9780470517277

Wiesmeier, M., Barthold, F., Blank, B., Kögel-Knabner, I., 2011. Digital mapping of soil organic matter stocks using Random Forest modeling in a semi-arid steppe ecosystem. Plant Soil 340, 7–24. https://doi.org/10.1007/s11104-010-0425-z

Wilding, L.P., 1985. Spatial variability: its documentation, accommodation and implication to soil surveys, in: Nielsen, D.R. (Ed.), Soil Spatial Variability. Wageningen, pp. 166–194.

Wilson, M.F.J., O'Connell, B., Brown, C., Guinan, J.C., Grehan, A.J., 2007. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. Mar. Geod. 30, 3–35. https://doi.org/10.1080/01490410701295962

Wright, M.N., Ziegler, A., 2017. Ranger: A fast implementation of random forests for high dimensional data in C++ and R. J. Stat. Softw. 77, 1–17. https://doi.org/10.18637/jss.v077.i01

Wuensch, K.L., Evans, J.D., 1996. Straightforward Statistics for the Behavioral Sciences., Journal of the American Statistical Association. Brooks/Cole Publishing, Calif. https://doi.org/10.2307/2291607

Xiao, J., Shen, Y., Tateishi, R., Bayaer, W., 2006. Development of topsoil grain size index for monitoring desertification in arid land using remote sensing. Int. J. Remote Sens. 27, 2411–2422. https://doi.org/10.1080/01431160600554363

Xu, Y., Smith, S.E., Grunwald, S., Abd-Elrahman, A., Wani, S.P., Nair, V.D., 2018. Estimating soil total nitrogen in smallholder farm settings using remote sensing spectral indices and regression kriging. Catena 163, 111–122. https://doi.org/10.1016/j.catena.2017.12.011

Table 1 Auxiliary variables extracted from the DEM and MS.

Environmental variables	Abbreviation	References
Terrain attributes		
Elevation	EL	
Slope gradient	SL	
Sine of aspect	ASs	Horn, 1981
Cosine of aspect	ASc	Horn, 1981
Topographic wetness index	TWI	Horn, 1981
Topographic position index	TPI	Brenning, 2008; Conrad et al., 2015
Topographic ruggedness index	TRI	Wilson et al., 2007
Tangential curvature	TC	Wilson et al., 2007
Profile curvature	PC	
Spectral indices		
Near infrared band reflectance	NIR	
Soil adjusted vegetation index	SAVI	Huete, 1988
Normalized difference vegetation index	NDVI	Huete, 1988
Topsoil grain size index	TGSI	Xiao et al., 2006

Table 2 Descriptive statistics of tested soil properties.

Soil constituent	Min	Max	Mean	Median	SD	CV (%)	Skewness	Kurtosis
SOC (%)	1.24	6.75	3.30	3.08	1.15	34.8	0.67	0.01
Sand (%)	9	78	29	25	17	58.6	0.85	-0.26
Silt (%)	21	77	50	54	11	22.0	-0.64	-0.62
Clay (%)	1	36	20	22	7	35.0	-0.67	-0.26
$CaCO_3$ (%)	0.00	8.72	2.88	2.16	2.41	83.6	0.65	-0.83
pH_{H2O}	5.76	7.82	7.25	7.46	0.46	6.3	-0.90	-0.27
pH_{KCl}	5.04	6.94	6.28	6.57	0.52	8.3	-0.91	-0.77

WC (%)	1.70	6.67	3.73	3.66	0.99	26.5	0.49	0.28

Table 3 Performances of UK and RFK models of the tested soil properties.

Soil constituent	Models	RMSE	SD_{RMSE}	ME	SD_{ME}	RPD	SD_{RPD}
SOC	UK	0.94	0.26	0.04	0.23	1.20	0.22
	RFK	0.95	0.27	-0.02	0.35	1.20	0.24
Sand	UK	16.0	4.00	0.49	10.10	0.98	0.13
	RFK	15.0	3.50	-0.47	8.10	1.05	0.12
Silt	UK	10.0	1.60	0.24	5.80	1.00	0.22
	RFK	10.0	1.30	0.54	4.50	1.00	0.15
Clay	UK	7.1	1.80	-0.32	4.20	0.93	0.15
	RFK	6.5	2.00	-0.04	4.00	1.02	0.19
$CaCO_3$	UK	2.10	0.26	0.18	1.02	1.10	0.28
	RFK	2.00	0.42	0.11	0.79	1.10	0.30
pH_{H2O}	UK	0.36	0.08	0.06	0.18	1.10	0.34
	RFK	0.38	0.09	0.03	0.18	1.10	0.22
pH_{KCl}	UK	0.43	0.10	0.08	0.20	1.10	0.36
	RFK	0.45	0.12	0.04	0.20	1.10	0.22
WC	UK	1.00	0.12	-0.03	0.46	0.97	0.07
	RFK	0.95	0.12	-0.07	0.34	1.04	0.15

Table 4 The semivariogram parameters of the best fitted models for the residuals of UK and RF.

Prediction models	Soil properties	Nugget	Sill	Range (m)	Nugget/Sill	Model
Universal Kriging	SOC	0.77	0.09	107	0.89	Exponential
	Sand	0.00	284.25	137	0.00	Exponential
	Silt	77.12	48.56	690	0.61	Exponential
	Clay	24.89	34.97	287	0.42	Exponential
	$CaCO_3$	3.40	0.90	130	0.79	Exponential
	pH_{H2O}	0.10	0.04	346	0.71	Exponential
	pH_{KCl}	0.11	0.07	148	0.61	Exponential
	WC	0.63	0.46	259	0.58	Exponential
Random Forest	SOC	0.10	0.06	362	0.62	Exponential
	Sand	19.59	36.78	662	0.35	Exponential
	Silt	7.33	12.79	284	0.36	Exponential
	Clay	3.28	6.32	383	0.34	Exponential
	$CaCO_3$	0.83	0.04	259	0.95	Exponential
	pH_{H2O}	0.01	0.02	308	0.33	Exponential
	pH_{KCl}	0.01	0.03	239	0.25	Exponential
	WC	0.11	0.09	448	0.55	Exponential

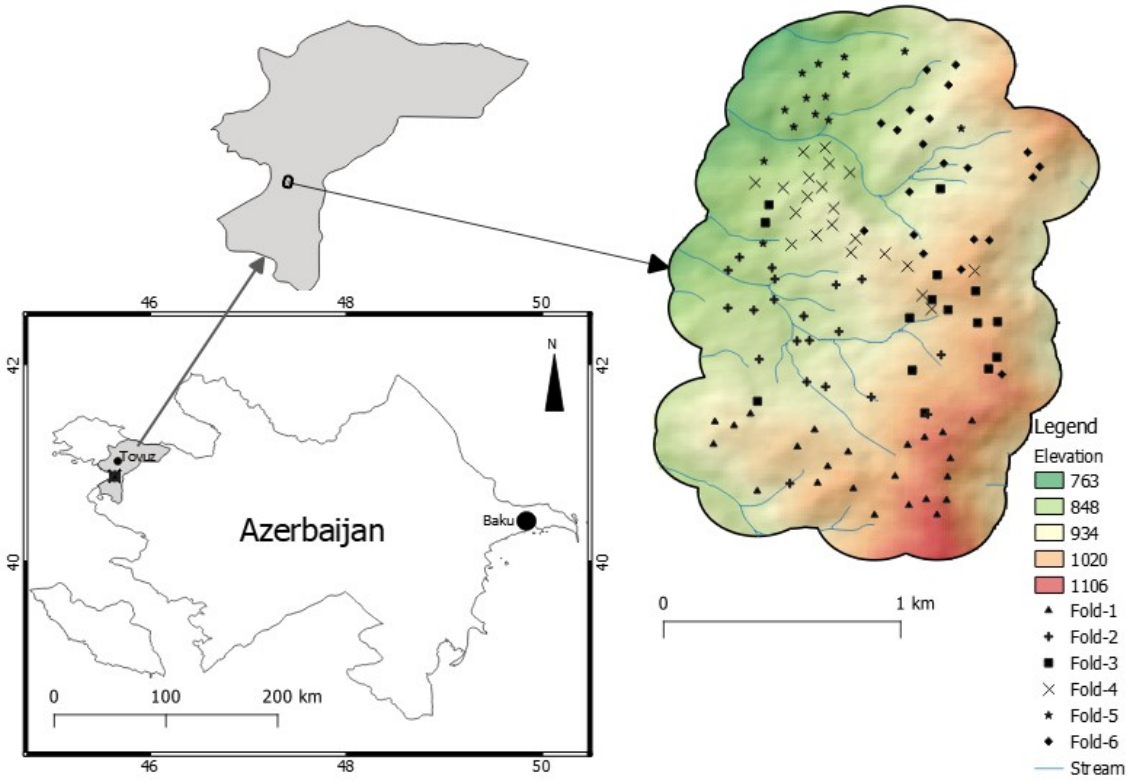


Fig. 1. Location of the study area (DEM source: Alaska Satellite Facility, https://asf.alaska.edu/sar-data-sets/alos-palsar/).

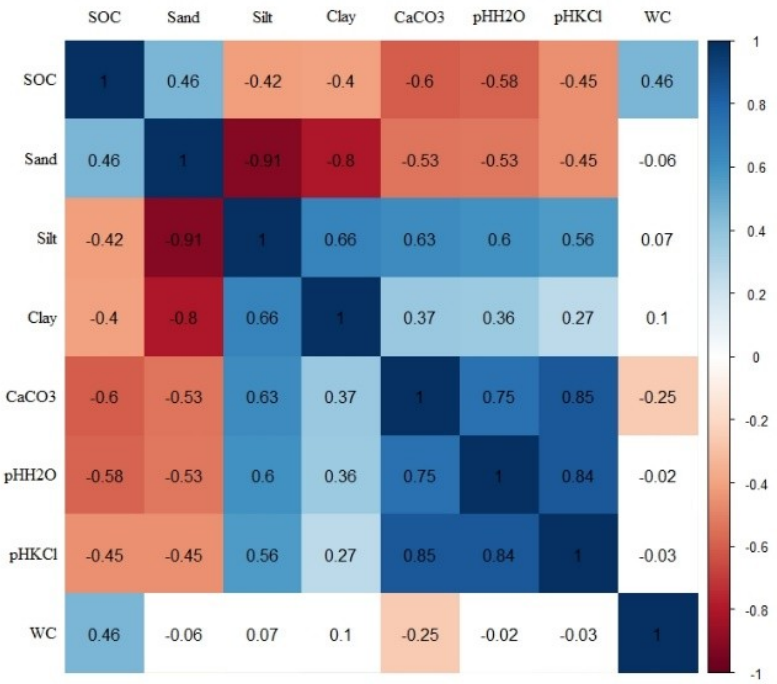


Fig. 2. Spearman's Rho correlation coefficients between tested soil properties. White background cells donate insignificant correlations, P > 0.01.

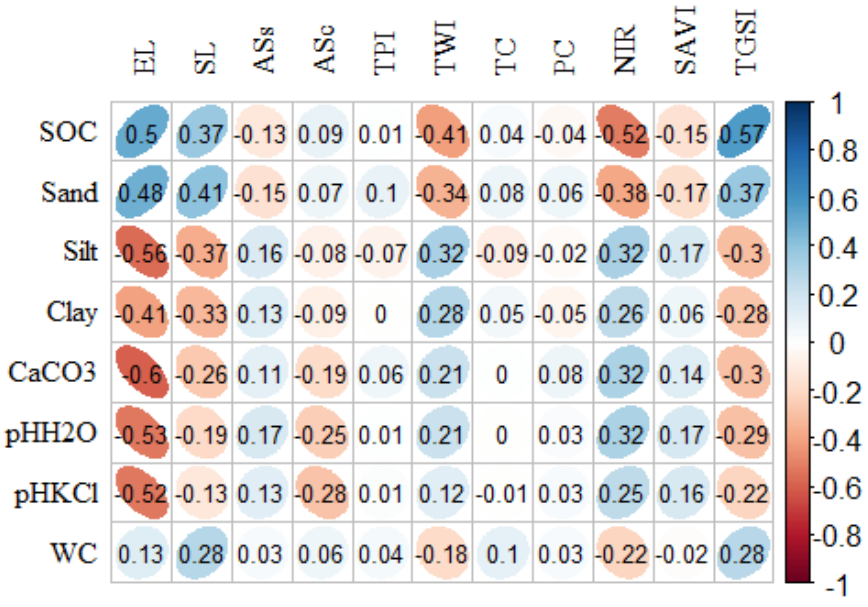
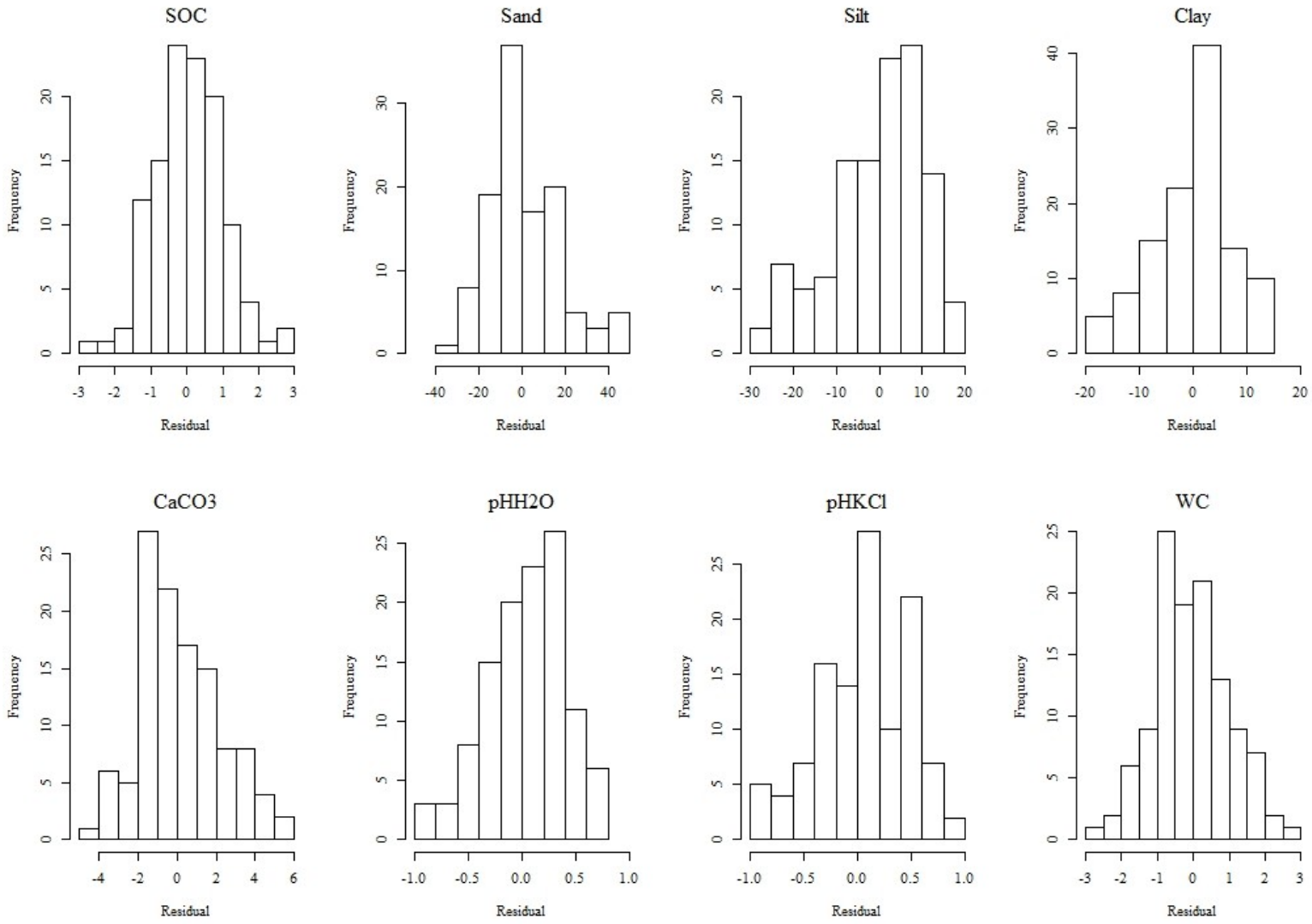


Fig. 3. Spearman's Rho correlation coefficients between tested soil properties and auxiliary variables.



2

3 **Fig. 4**. Histograms of the residuals of soil properties calculated based on UK.

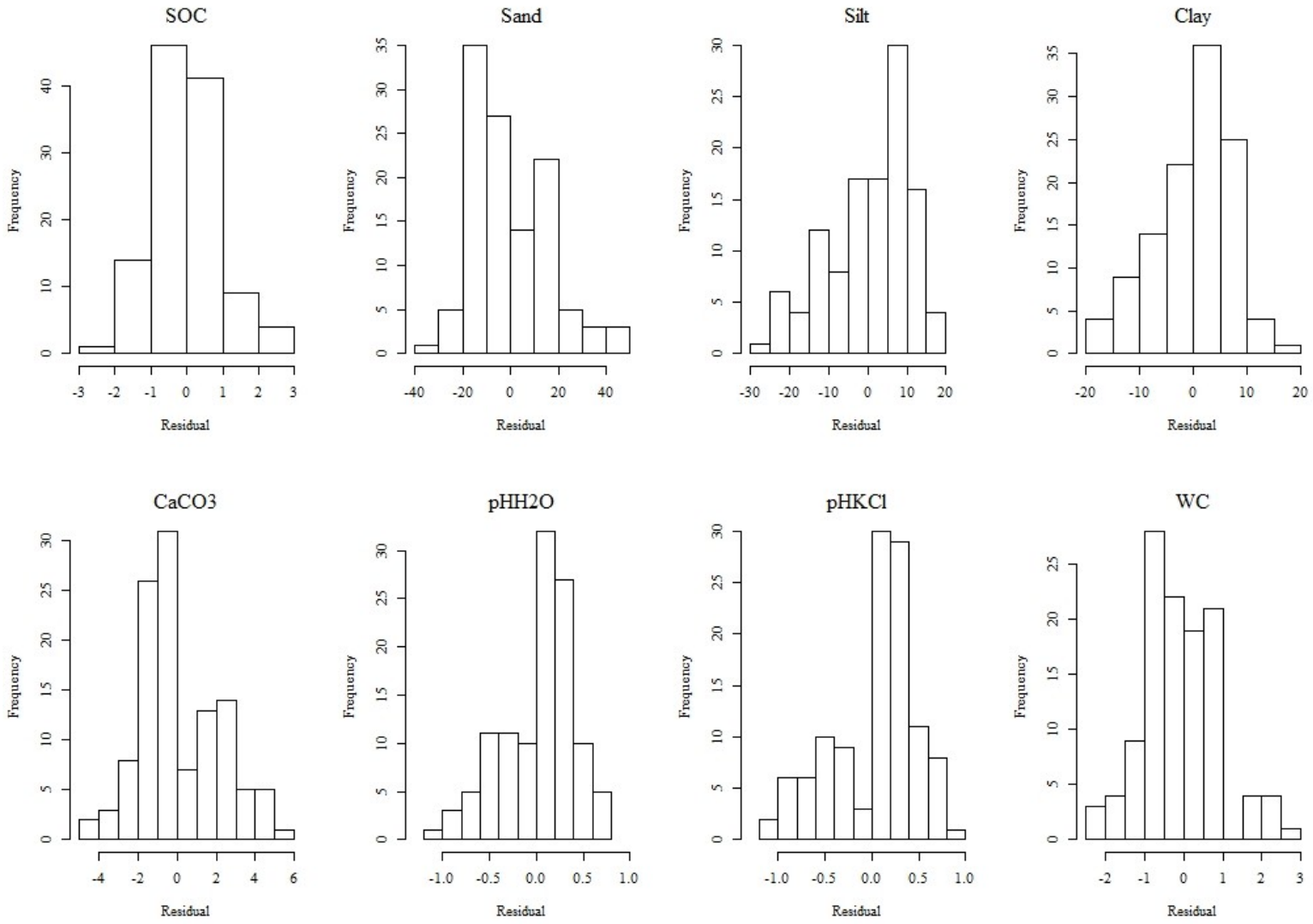


Fig. 5. Histograms of the residuals of soil properties calculated based on RF.

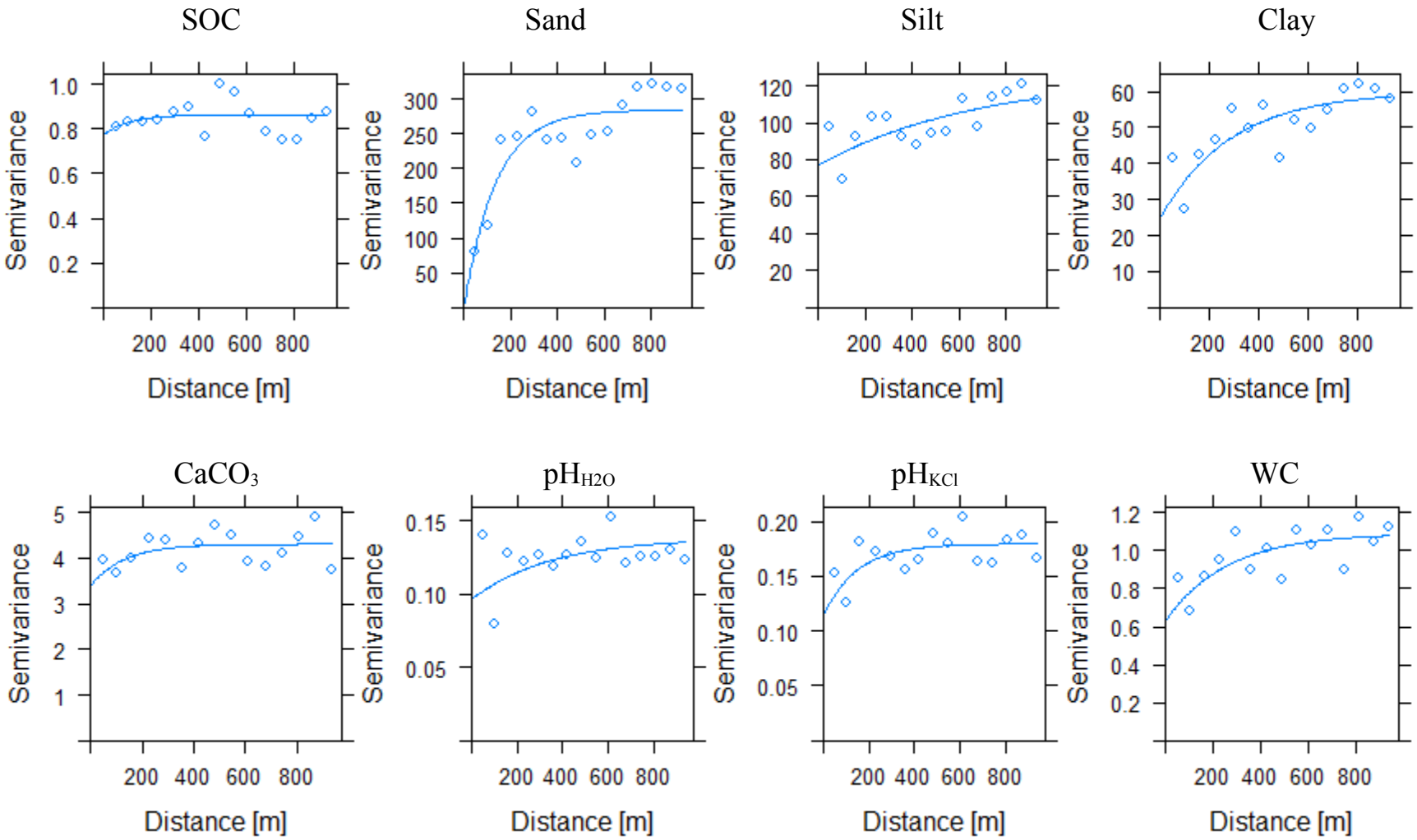


Fig. 6. Variograms and best fitted models for the UK residuals.

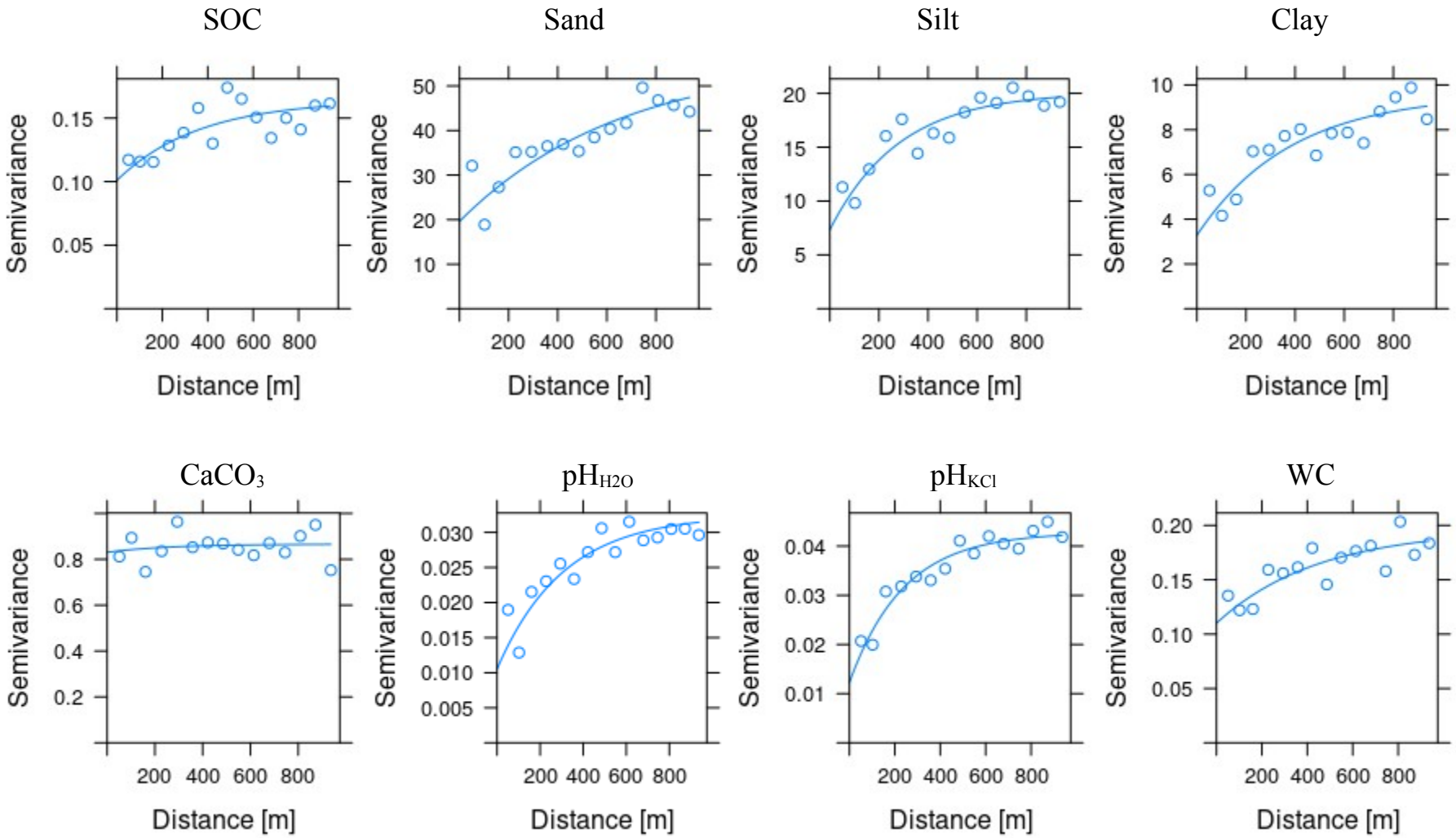


Fig. 7. Variograms and best fitted models for the RF residuals.

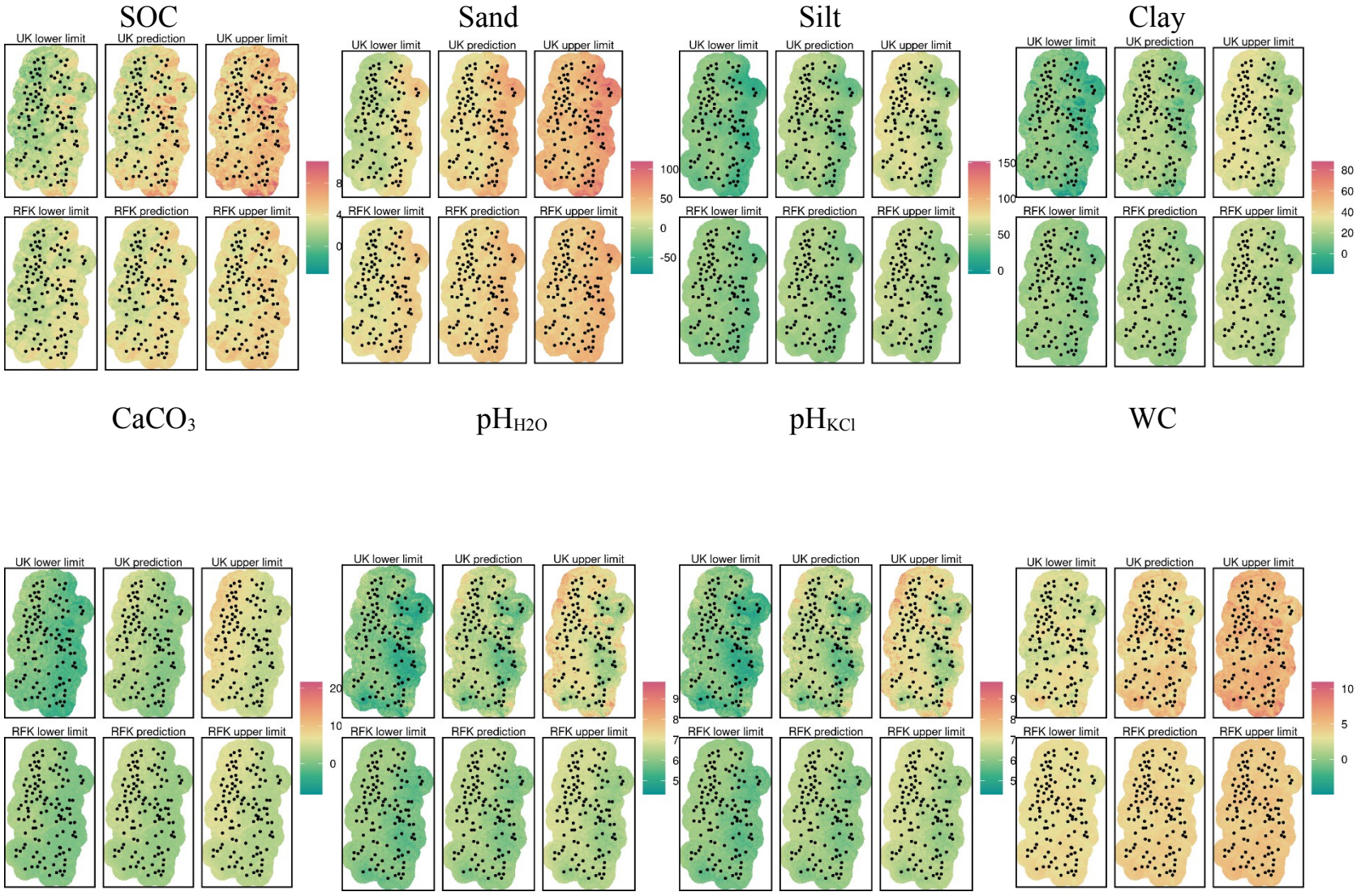


Fig. 8. Predicted map of the soil properties and their corresponding upper and lower limit of the 90% prediction interval. The units of the maps is [%] for the all soil properties except pH_{H2O} and pH_{KCl} and black points are sampling locations.

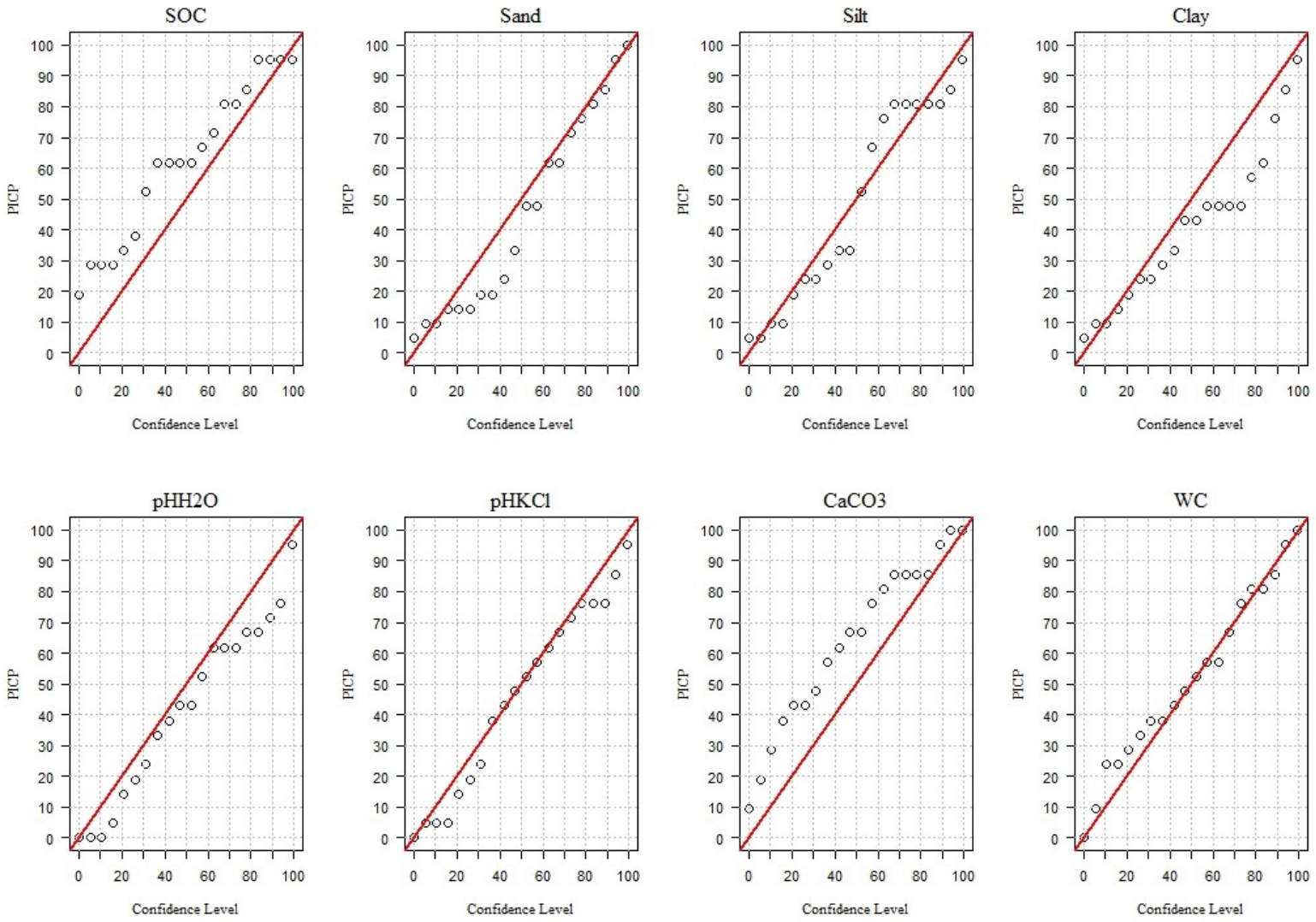


Fig. 9. PICP plots of the soil properties predicted with UK.

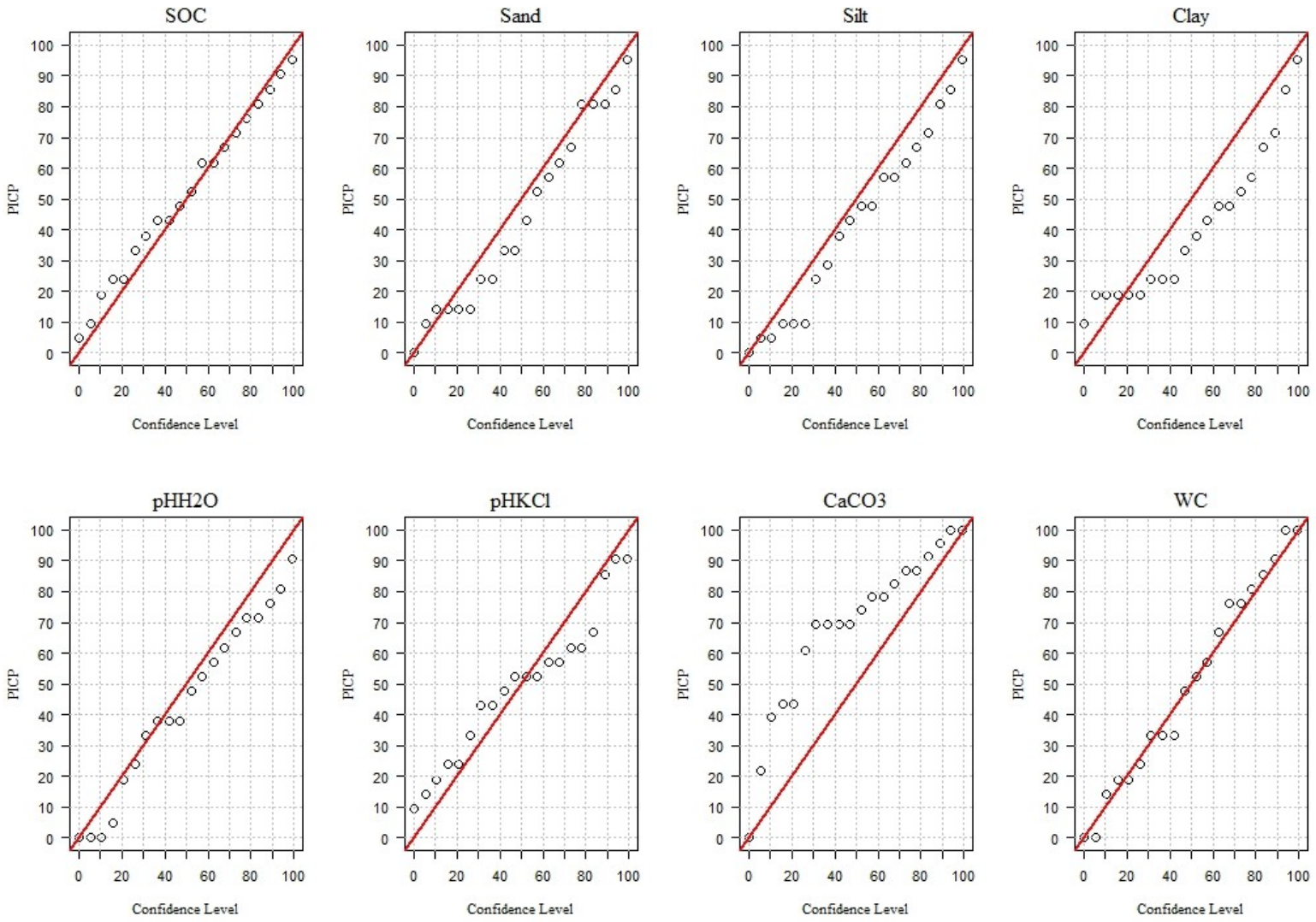


Fig. 10. PICP plots of the soil properties predicted with RFK.

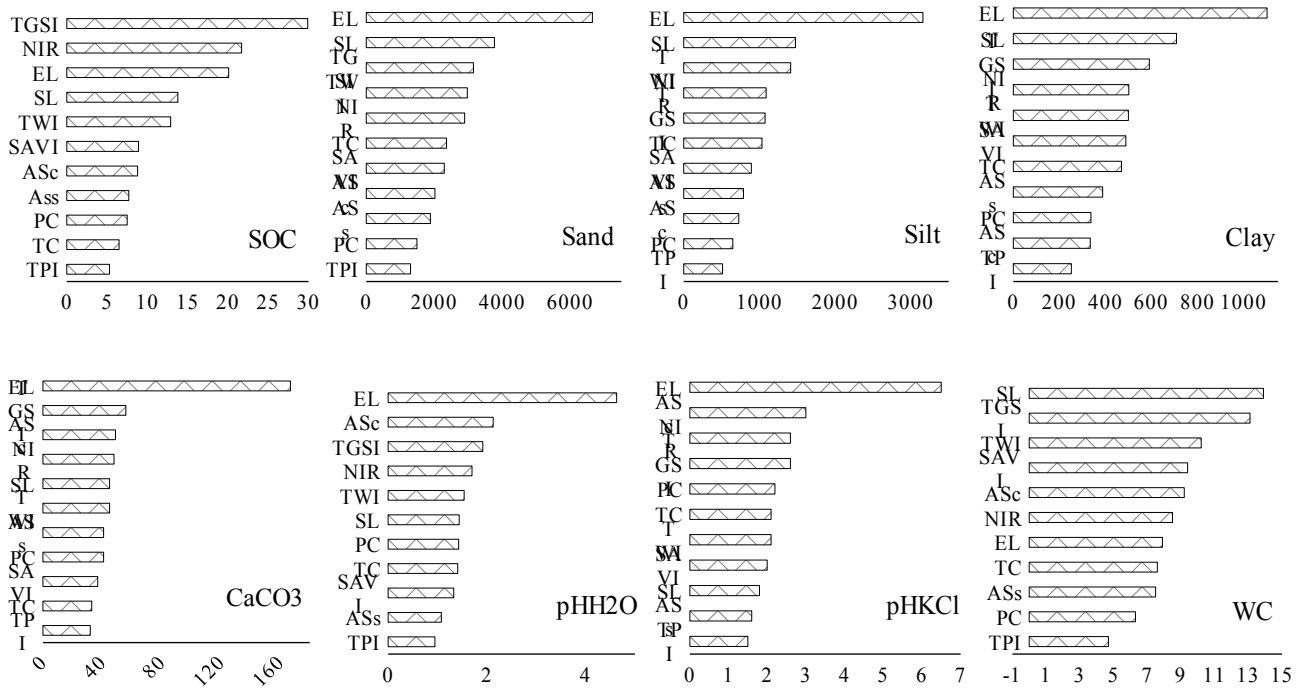


Fig. 11. Importance of covariates in RFK model (vertical axis: predictor variables and horizontal axis: partial importance score).