American Journal of Innovative Research and Applied Sciences. ISSN 2429-5396 I www.american-jiras.com

ORIGINAL ARTICLE

DIGITAL SOIL NUTRIENTS MAPPING IN SEBEYA CATCHMENT AGRICULTURAL LAND IN RWANDA

| Félicien Majoro [1] * | Albert Niyonzima [1] | Concilie Mukamwambali [2] | Adrien Mugabushaka [3] | and | Athanase Twagirimana [4] |

[1]University of Rwanda |College of Science and Technology | Department of Civil | Environmental and Geomatics Engineering | P.O. Box 3900 | Kigali| Rwanda |
[2]University of Rwanda | College of Education |Department of Mathematics | Science and Physical Education | P.O. Box 55, Rwamagana | Rwanda |
[3] Integrated Polytechnic Regional Center | IPRC Ngoma | P.O. Box 35 | Kibungo | Rwanda |
[4]University of Rwanda |College of Science and Technology | Department of Geography | P.O. Box 3900 | Kigali | Rwanda |

| Received August 27, 2020 | | Accepted September, 2020 | | Published October 06, 2020 | | ID Article| Majoro–Ref1-ajira270920 |

ABSTRACT

Background: Soil is a natural three-dimensional body, which varies in space and time. Expanding environmental concern has increased the interest for regional land use analysis and particular in Rwanda. The rapid growth of population has testified the effect of soil degradation through overexploitation of arable lands which has been accompanied by a decline in soil productivity due to the depletion of soil nutrients by soil erosion and insufficient fertilizers application. Digital soil mapping is a method used to create relationships between observed soil properties and environmental variables. **Objective**: This study has focused mainly on spatial analysis and interpolation of soil properties in Sebeya catchment located in Western Province of Rwanda. **Methods**: The availability of soil properties at 12 locations in Sebeya catchment were clipped from 2017 Rwanda soil nutrients map. **Results**: The selected soil nutrients to be mapped were five (Aluminium (Al), Calcium (Ca), Copper (Cu), Phosphorus (P) and pH) while Regression Kriging (RK) and Random Forest Kriging (RF) were 2 methods used for interpretation. The accuracy of these both methods was assessed by using three distinct validation indices: the mean absolute estimation error (MAEE), the root mean square error (RMSE) and the coefficient of determination (R^2). The best method for predicting all the 5 selected soil nutrients was the random forest kriging. **Conclusion**: More accurate digital soil maps (DSMs) will be obtained if much denser soil profile data and detailed soil information from legacy soil maps are used.
Keywords: Digital soil mapping, Sebeya catchment, Rwanda

1. INTRODUCTION

In Rwanda, land degradation often reveals the effect of quick population growth through land use and management [1]. Besides anthropogenic factors, the spatial variation of soil properties is likewise influenced by natural factors such as climate, topography, vegetation and parent material [2, 3, 4]. Thus, mapping of soil properties which takes into account the previously stated factors influencing the variation of soil nutrient can be one way to deal with soil productivity management in Sebeya catchment.

Cultivation on steep slopes without proper erosion control can also lead to soil nutrient losses from Sebeya catchment agricultural highlands. Detailed legacy soil maps are more than 20 years old and inadequately represent the actual conditions as a result of the land use/land use changes [1]. Amongst these changes is land fertility degradation, which has turned into a remarkable issue for agricultural management in Sebeya catchment as well. Therefore, there is a need to produce maps of soil fertility to deal with fertility management related issues. In order to respond to this need, interpolation methods are needed to convert isolated point observations to a full-cover map. Geostatistics is a branch of science in which such interpolation methods are developed and has provided reliable approaches to analyse the spatial distribution of soil attributes and the pattern of variation. One method, the kriging interpolation algorithm, is often applied in geostatistics. Nowadays, quantitative approaches are more used than qualitative ones [5, 6, 7].

By applying regression kriging and random forest kriging, this research aims to establish a spatial distribution and analysis of soil properties by interpolation based on existing measured soil properties. The output maps are still expected to be indicative of nutrient deficiencies and hence fertilizer recommendations could be refined.

The regression kriging, referred simply as RK in this study, is a generally used in DSM mapping [8], which involves linear regression of soil attributes with explanatory covariates variables, followed by kriging of regression residuals. The regression predictions are eventually summed up with the kriged residuals to produce a final prediction map. This method was used as a reference model and further comparison with the machine learning method (Random Forest regression kriging (RF)) was made to inspect which method is more accurate for DSM of the soil properties under consideration in this study.

2. METHODOLOGY

2.1. Study area

2.1.1 Site localization

This study is entirely focused on Sebeya catchment (363.3 km^2), located in the Western Province of Rwanda as presented in figure1. The main river flowing in this catchment is Sebeya, which originates in the mountains of Rutsiro District. The catchment is shared by four administrative units namely Rubavu, Nyabihu, Rutsiro and Ngororero (figure1).

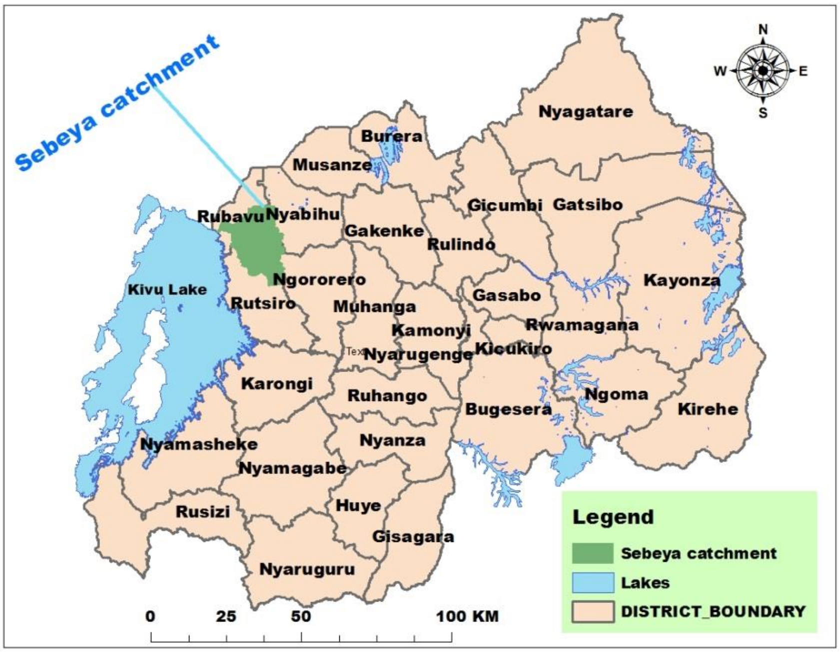

Figure 1: Location of Sebeya catchment

2.1.2 Population density

The average density of Sebeya catchment is estimated to 644 hab/Km^2 while the average population density of Rwanda is about 415 habitants per km^2 [9]. The Northern part of Sebeya catchment in Rubavu District (Nyakiriba, Rugerero and Gisenyi sectors) is occupied by a significant urban population. The sectors along the shores of Lake Kivu and along the main road from Rubavu to Musanze are very highly populous with more than 1000 hab/km^2 while the sectors in the highlands of the South-East show the lowest population density fluctuating from 250 to 500 hab/km^2 [10]. High demographic pressure is one the indirect factors accelerating soil erosion in Sebeya catchment.

2.1.3 Soil characteristics

The soil in this catchment favors agriculture due to its high infiltration rates and its high minerals content except for the case of clay soils on flat topography encountered in the catchment. The combination of the geological formation and soil data characterize Sebeya catchment as a fragile ecosystem susceptible to heavy erosion [11].

2.1.4 Site topography and rainfall

Sebeya catchment is located in the high elevation region of the country with altitude varying between 1,462 m to 2,979 m a.b.s.l. (meters above sea level). This catchment is also characterised by steep slopes (varying from 0% to 90%) and abundant rainfall varying between 1,200 mm to 1,700 mm per year [12] revealing that a great part of this catchment falls in medium risk to very high risk of erosion according to the classification of MoE in 2018.

2.2 Data collection

2.2.1 Auxiliary data

Several sources of auxiliary data can be retrieved and analysed so as to capture the spatial variation of the soil forming factors in the area of study. Various data related to climate, organisms (fauna and flora), topography and parent material were used to assess the relationship between the soil forming factors and the targeted soil properties.

Climate refers to the meteorological records including precipitation, temperature, and wind prevailing commonly in a particular region over long periods of time. The information on climate of an area may exist as point data or as raster files. Climate data have been obtained at Rwanda Meteorology Agency from different weather stations set up in various regions of the country. The main task was to convert local data and climate maps to a full-cover covariate map of climatic variables such as precipitation, evaporation, and transpiration. Topography is characterised by the land surface elevation, shape, and features of the surface. A general term for digital dataset of topography is known as Digital Elevation Model (DEM). DEM is often described as a grid of pixels representing the land surface elevation. Digital elevation models created from remote sensing data often have problems that should be overwhelmed before they are suitable for soil mapping and a lot of software and algorithms are available [13]. In this research, the terrain parameters needed to map the soil attributes will be derived from DEMs drawn from the existing legacy maps. Here the task will be to derive DEM and its derivatives (eg. Slope, aspect, wetness, hill shade, etc.) as full-cover covariate maps.

Parent material indicates the underlying geologic material from which soil is formed. Parent material, from which soil develops, originates from different sources and this is because parent material is not spatially homogeneous. DSM data on soil parent material can be seized from existing geological maps. Effort were made to extract the parent material from geologic information (soil map), but this task was only to be performed as long as the parent material is considered an important factor in this study.

2.2.2 Statistical modelling of DSM

The whole set of explanatory variables derived from the auxiliary data with regard to climate, land use, topography and parent material can be used for the spatial prediction of soil nutrient by means of regression kriging [8]. In this case the models for the chosen soil attributes were obtained with step wise multiple regression analysis [14].
However, the statistical analysis for digital soil mapping (DSM) was performed using regression kriging method (RK) and random forest kriging method.

2.2.3 GIS tool and statistical software

In this study, ArcGIS which is a geographical information system (GIS) package that works with maps and geographic information, has been a useful tool in extracting the environmental covariates maps employed for predicting target soil properties. The version ArcMap10.4.1 helped in processing and generating all covariates maps to the same extent and resolution as well as transformation of the generated maps into a common projection (conversion into same coordinate reference system).

2.2.4 Spatial prediction and simulations

According to Hengl et al. (2004), the main concept of any spatial model is to provide best estimate at some un-sampled locations (s_0), given the sampled values ($z_1, \ldots\ldots, z_n \text{ or } \mathbf{z}$) and some auxiliary predictors ($\mathbf{q}$), which are available over the whole area of interest. This may be expressed as:

$$\hat{z}(s_0) = f(z, \mathrm{q}); \qquad \hat{\sigma}(s_0) = \mathrm{Var}[z(s_0) - \hat{z}(s_0)] \tag{1}$$

where $\hat{z}(s_0)$, is the estimated or predicted value at unobserved location and $\widehat{\sigma}(s_0)$ is the error estimated during prediction, that is an estimated difference between the observed and predicted values.

2.2.5 Combination – *clorp(t)* and kriging

McBratney et al (2003) denotes that both kriging and the environmental correlation approach are combined in what is commonly known as "regression kriging". In this respect 'clorp (t)' is used to predict the soil property of interest from environmental variables and kriging is used on the residuals. Regression-kriging (RK) is a statistical technique which allows predictions using the sampled data and auxiliary predictors, as described in Eq. (1). In the case of RK, the effect of both environmental predictors and spatial location of observations are modelled using the (additive) universal model of spatial variation [8]:

$$z = q \,.\, \hat{\beta} + e; \qquad E(e) = 0 \tag{2}$$

Where $\boldsymbol{q}$ represents predictors at sampled locations are $\hat{\boldsymbol{\beta}}$ is the vector of the fitted regression coefficients and $\mathbf{e}$ are the residuals.

2.2.6 Validation

The prediction accuracy of the resulting maps can be evaluated using n-fold cross validation. The leave-one-out cross validation is the form of validation known for its advantage over the independent validation that, the splitting is repeated, which explains its efficiency compared to data-splitting in which the data set is split into two [16]. In leave-one-out cross validation the data set is split into (n-1) times observations for calibration and one for validation. For each observation, the model is refitted leaving that observation out of the calibration dataset and estimated using the remaining observations. The ignored observation is returned to the data set and another is removed and estimated. This process is repeated until all observations are estimated. The differences between observed and predicted values can thus be computed with the Mean Estimation Error (MEE), Root Mean Square Error (RMSE) and the standardized Root Mean Square (SRMSE) using Eq.(3), (4) and (5) respectively [16].

$$\mathrm{MEE} = \frac{1}{\mathrm{n}} \sum_{\mathrm{i=1}}^{\mathrm{n}} (\mathrm{obs_i} - \mathrm{pred_i}) \tag{3}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\mathrm{obs}_i - \mathrm{pred}_i)^2} \quad \textbf{(4)}$$

$$\mathrm{SRMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\mathrm{obs}_i - \mathrm{pred}_i)^2 / \,\mathrm{pred\ var}_i} \quad \textbf{(5)}$$

Where **n** is the number of sample observations; obs_i = concentration of observed point (or measured point); pred_i = predicted value; pred var_i = predicted variance.

There are various other indices that can be used in conjunction with the aforementioned ones to assess the performance of an interpolation method such as; the mean absolute estimation error (MAEE) which differs to RMSE in that it is less sensitive to outliers, and the Pearson correlation coefficient (R^2) which indicates the strength of the linear relationship between predictions and observations.

3. RESULTS AND DISCUSSIONS

3.1 Lognormal transformation

The lognormal transformation was obtained by calculating the natural logarithm of the skewed data sets. For the transformed soil properties, it was necessary to make a back transformation after the analysis to get the final prediction maps that can be correctly interpreted Transformation and back transformation task was performed automatically in software.

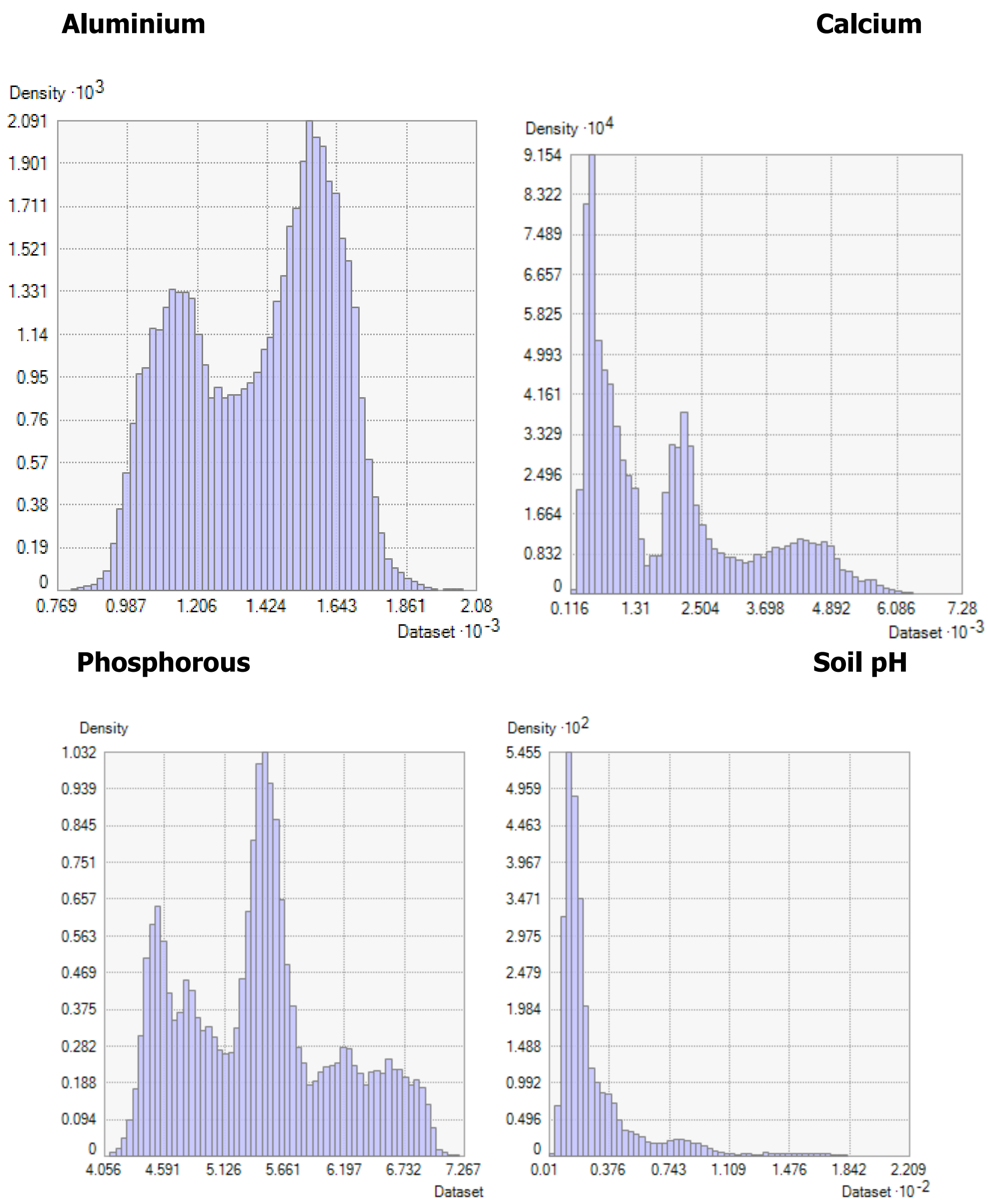

Figure 2: The histograms showing the lognormally transformed soil properties.

3.2 Model comparaison

Table 1: The table presents the comparison of independent validation results for both regression kriging and random forest kriging methods by use of root mean square error as validation index.

Soil property	Interpolation algorithme	Validation index (RMSE)	MAEE	R^2	Min	Max
Alminium	Regression kriging	201.6319	153.857	0.346883	599.75	2095.72
	Random forst kriging	165.3363	146.4039	0.267326	768.91	2079.72
Calcium	Regression kriging	628.2151	613.1421	0.371943	-1126.66	6576.43
	Random forest kriging	471.268	433.582	0.263203	115.66	7280.36
Copper	Regression kriging	1.040593	1.136428	0.89399	-0.6	5.1
	Random forst kriging	0.699545	0.696984	0.916653	0.3	6.4
Phosphorous	Regression kriging	83.47951	72.55144	0.110605	0	97.96
	Random forest kriging	56.4457	48.01328	0.070863	1.006	220.85
pH	Regression kriginging	0.738401	0.623938	0.065059	3.2	8.1
	Random forest kriging	0.591112	0.490341	0.125096	4.05	7.26

The above table represent the summary of results after validation, in this study the random forest regression kinging were found to be good technique for Sebeya Digital Soil nutrients Mapping, by comparing their RMSE. This means that the lower the Root Mean Square Error the high the accuracy.

Based on the results presented above, it has been clearly noticed that the soil forming factors did not show strong dependency or relationships with soil properties as would be expected. This is not very surprising for counting on the limitations with respect to the used data, especially the auxiliary data or soil covariates, since not all relevant data were available for use. The output maps are still expected to be indicative of nutrient deficiencies, and hence fertilizer recommendations could be refined, even though the added value of geospatial analysis for improving nutrients recommendations seemed to be not very impressive. It is also important to note that Sebeya catchment soils are highly variable and consequently, general conclusions about soil fertility recommendations should be strained with much care

3.3 Final prediction map of soil nutrient

Prediction of the five soil targeted properties namely; Aluminium (Al), Calcium (Ca), Copper (Cu), phosphorus (P), and soil pH was completed by applying one of the two tested models, relying on the one that demonstrated relatively good performance in terms of prediction accuracy after comparison. To predict a given soil property at sampled locations in the study area, the model was based on the relationships established from 14 observations (points data) sampled across Sebeya catchment at a depth of 20 cm and a combination of soil covariates that reflect basic soil formation processes. Nine (9) soil predictors were used of which 5 are continuous variables and the remaining 4 are categorical variables. Numerical variables used are: digital elevation model (DEM), slope, Topographic Wetness Index (TWI), Annual precipitation (Pann) and precipitation surplus (Psur) while the categorical ones are soil types, soil texture degree of soil profile and land use. The resulting maps of soil properties distribution were based on the results of the run model generated by randomly selecting 64% of points for calibration and 36% for validation splits of the input point data. To evaluate the performance of the final interpolation method, suitably designed and unbiased testing procedures were adopted basing on the separate controlled data set (5 observations) at different locations used to validate the final model by computing the root mean square error (RMSE which is the measure of accuracy of prediction), mean absolute estimated error (MAEE which is a measure of the bias of prediction model), and the coefficient of determination (R2 which accounts for the amount of variation explained by the model). Below are the output full cover maps of soil nutrients produced at a resolution of 30 m (pixel size = 30 m x 30 m) along with the interpretation of the patterns of soil properties variation throughout the study area. From the maps, it is important to note that areas shown in white colour represent cities, mines, lakes and other surface waters. Consequently, there are no quantitative predictions assigned to these non-soil parts of the study boundary.

3.3.1 Prediction and mapping of Al content

It has been so far noticed that pH is a major factor influencing the availability of potentiol nutrients for the plants growth and solubility of toxic as well. Acid toxicity is commonly brougth about by the lack of essential nutrients in the soil and the excess of toxic metals within the plant root zone. Alminium (Al) is the most important acid cation and most toxic. In this present study it is found that, the factor controlling the Al levels in soil were expressed by estabilishing dependency of soil property on available environment factors. They were found to be elevation(Elv), profile development stage, land use (Lu), annual precipitation (Pann) and precipitation suplus (psur). As suggested, regression model, altitude and annual distribution of rainfall and weathering have shown great importance in variation of aluminum saturation. From figure3, the prediction map shows a range between 768.91 and 2079.72 ppm for Aluminum. The map indicates that Al is abundant in North East part and Southern part of the Sebeya catchment. The justification behind the high Al saturation in those areas is exposed by high rainfall amount giving rise to leaching power of basic cations. As long as basic cations are leached out, Al becomes the dominant cation in soil and thus leading to toxicity. Al toxicity is considerably low in the North West zones of Sebeya catchment and where the climate does not favor leaching because most of this area is built up area, thus keeping the base saturation (BS) more or less stable. In general, Al toxicity occurs readily under acidic conditions and soil pH is always influential. If the soil pH decreases (at a pH less than 5.4), Al becomes soluble. Even a small drop in soil pH can further result in large increase in soluble aluminum. Consequently, Al in this soluble form hampers organic matter (OM) decomposition, which can affect the plant growth due to unavailability of phosphorus (P) and nitrogen (N), reduced soil Cation Exchange Capacity (CEC) and base saturation (BS). Al variance map shows a range from 0.016 to 0.019.

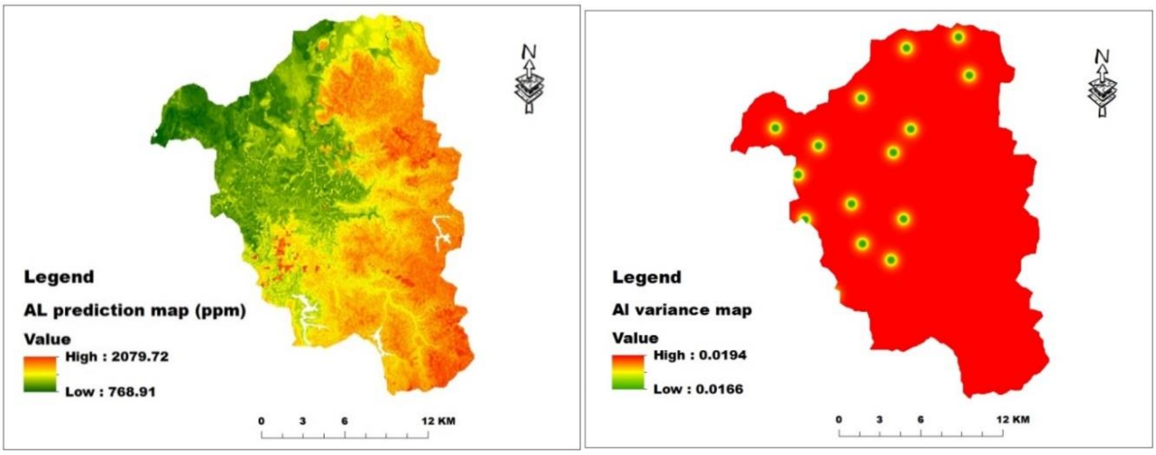

Figure 3: Alminium prediction map (left), alminium variance map_random forest kriging.

3.3.2 Prediction and mapping of Ca content

Calcium (Ca) is deliberated as a secondary important plant nutrient after nitrogen and potassium that are also required by plants in large amounts. In the framework to map available Ca distribution in Sebeya catchment soils, the relationships between soil Ca and environmental predictors were explored by means of 'random forest' model. Those environmental predictors for soil Ca were found to be: the elevation, slope, topographic wetness index, land use, soil types, annual precipitation and precipitation surplus. Except Psur, the prediction map from figure4, displays a large prediction range between 115.66 and 7280.36 ppm for Ca and the variance ranges from 0.005 to 0.007, which is exceptionally low variation throughout the catchment. From the map, it is understood that controls on soil calcium distribution are driven by parent material and climate factors, especially altitude and rainfall patterns. The map indicates that high calcium concentrations to the North parts of the Sebeya catchment which is influenced by volcanic parent materials from which soils are derived and topography. Low calcium concentrations over the Central and Southern parts of the catchment. On the map, Ca distribution can also be explained by interactions with other soil properties (eg. soil nutrients mapped earlier). In acid soils, low Ca concentrations are encountered, and high pH soils usually contain more Ca. If the soil pH increases (at pH greater than 7.0), free Ca starts to accumulate in the soil and this Ca is available to interact with other nutrients. However, much of the free Ca react with phosphorus (existing as an anion) to form insoluble or very slowly soluble Ca-P compounds that are not readily available to plant, thus rendering P less available as there is normally much more available Ca in soil than P. Conversely, as soil pH decreases, more Al^{3+} and Fe^{3+} become soluble and complex Ca to form insoluble compounds as well.

Calcium is basically found in most primary and secondary minerals in soils, although it is, in this state, essentially insoluble for agricultural considerations. These components of primary and secondary minerals are the original sources of available forms of Ca (soluble forms). It becomes in fact, available in soluble forms as a positively charged cation (Ca^{++}), which is adsorbed to the exchange complex of soil colloids. It is therefore this ionic form of Ca that is regarded to be generally, soils derived from calcium minerals such as limestone, marl, etc. will tend to have high calcium levels, while soils derived from sandstone or shale will tend to have low Ca levels. Although calcium is considered not to be toxic to plants, excess levels of Ca in the soil is liable for reduced plant uptake of other nutrients such as phosphorus (P), potassium (K), magnesium (Mg), boron (B), copper (Cu), iron (Fe), and zinc (Zn), thus making these nutrients deficient.

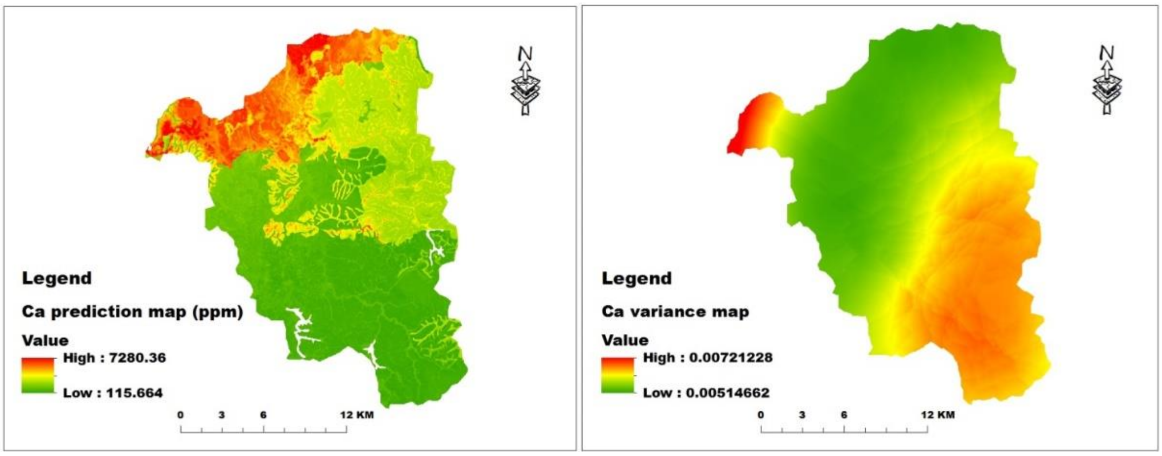

Figure 4: Calcium prediction map (left), calcium variance map (right)_random forest kriging.

3.3.3 Prediction and mapping of Cu content

Copper is the only soil micronutrient considered in this present study as it is one of the crucial micronutrients needed for plant growth after the major soil macronutrients. Copper (Cu) availability was modelled by the random forest kriging method and from the available environmental factors, altitude, land use and soil types have shown relevance during the interpolation of copper nutrient to other predictors. The map on the left side of figure5, shows that Cu has small prediction range (0.3to 6.4 ppm) and its variance ranges from 0.13 to 0.3 the spatial patterns of Cu show irregular and a wide range of variation throughout the study area. Low spatial patterns for Cu are mainly spread along areas covered by peat and organic soils (parent material H). The distribution and availability of Cu in Sebeya catchment soils may also be partly explained by the correlation or interaction with some soil properties such as soil pH, soil organic matter clay and cation exchange capacity. Cu in soils is significantly negatively correlated with soil pH since copper (like other micronutrient cations) availability decreases as pH in soil increases, as a result of decreased solubility of Cu minerals (Cu-hydroxides or oxides). The presence of organic matter in soil inhibits Cu availability since copper is readily and tightly complexed by OM. As far as the interaction with other nutrients is concerned, copper availability may be decreased as zinc (Zn), iron (Fe) and phosphorus (P) contents are high in soil solution.

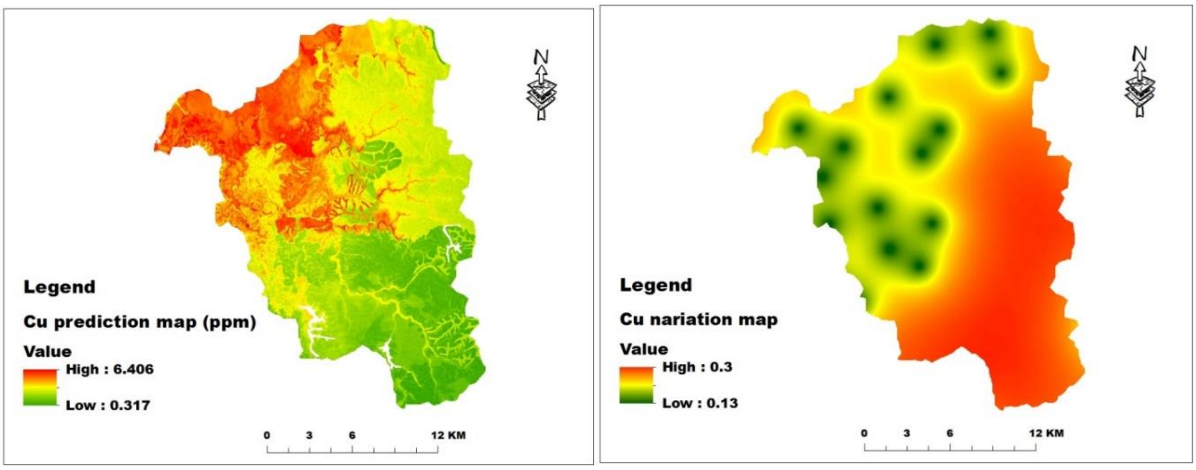

Figure 5: Copper prediction map (left), copper variance map (right) _random forest kriging.

3.3.4 Prediction and mapping of P content

Phosphorus nutrient availability is generally attributed to some inherent soil properties and climate. Those basic factors include soil age, parent material, topographic position, etc. but their importance on soil P status may be highly limited. The regression model has attributed the relevance of soil P variation to only two predictors; parent material (soil types) and climate (precipitation surplus). From the prediction map on the left side of figure6, soil P was predicted between 1.0006 and 220.85 ppm. It is noticed that P availability is limited to those areas where soil pH is slightly between 6 and 7.5. In locations where soil pH is lower than 5.5 and higher than 7.5, P availability is restricted as a result of fixation by aluminium (Al), iron (Fe), or calcium (Ca).

This indicates that there is a positive correction between pH and available of phosphorous. As pH increases there is perceptible increase in phosphorous availability within finest range. In acidic condition, pH is very low and exchangeable Fe and Al became available in soluble form and ready for complexing P to be available. This results in insoluble P compounds that impede the variability of P for plants.

Availability of soil P in some parts towards the North West is commonly associated with soil parent material. Soils derived from parent material B (basic rocks) and V (volcanic materials) contain relatively significant available P, that is about two times P content of the soils derived from granitic (G) and quartzite (Q) materials. P variance ranges from 48.3 to 59.5 which is relatively a very high variation throughout the catchment.

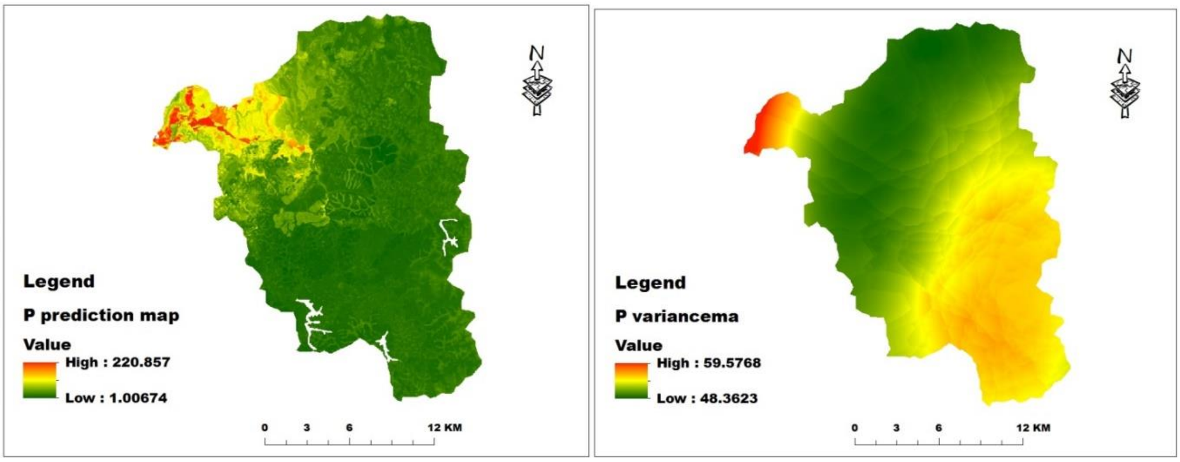

Figure 6: Phosphorous prediction map (left), phosphorous variance map (right)_ random forest kriging

3.3.5 Prediction and mapping of pH content

The relationship between pH and auxiliary predictors was modelled by the random forest kriging method method. The environmental factors controlling pH variation are mainly; elevation, slope, topographic wetness index, annual precipitation, precipitation surplus, land use and soil types (parent materials) as suggested by the output of the stepwise multiple regression analysis. All predictors showed relevance in predicting soil pH except texture (TXT) and the degree of profile development or development stage (DEVSTG). The latter variable has showed prediction importance only for aluminium (Al) from the remainder of the soil properties, while texture did not play any prediction role at all. At 5% level of significance, the p-value was used as a measure of significance or importance of covariates in predicting soil properties. For pH, continuous variables showing significance (p-value <0.05) include; Elev, slope, Pann, and Psur while TWI is slightly not significant (p-value >0.05). It is worth noting that all involved predictors have shown negative correlation with pH, but the elevation (altitude) and slope were the most landscape parameters explaining the variation of soil pH.

From figure7, the map to the left represents the spatial distribution of soil pH in Sebeya catchment. As shown by the map legend, soil pH was predicted within the range between 4.05 and 7.26 and the map on the right side shows the variation over the entire study area. The prediction map shows explicitly that the lowest pH values are found from Southern part of the catchment. While the highest pH values are distributed over North West. These patterns indicate that the higher the topography (altitude) and the steeper the slope, the lower the pH. The variance map shows small variation in soil pH with a range between 0.07 and 0.2. Some variations may be explained by large distance between sampling locations and the vicinities along these wetlands.

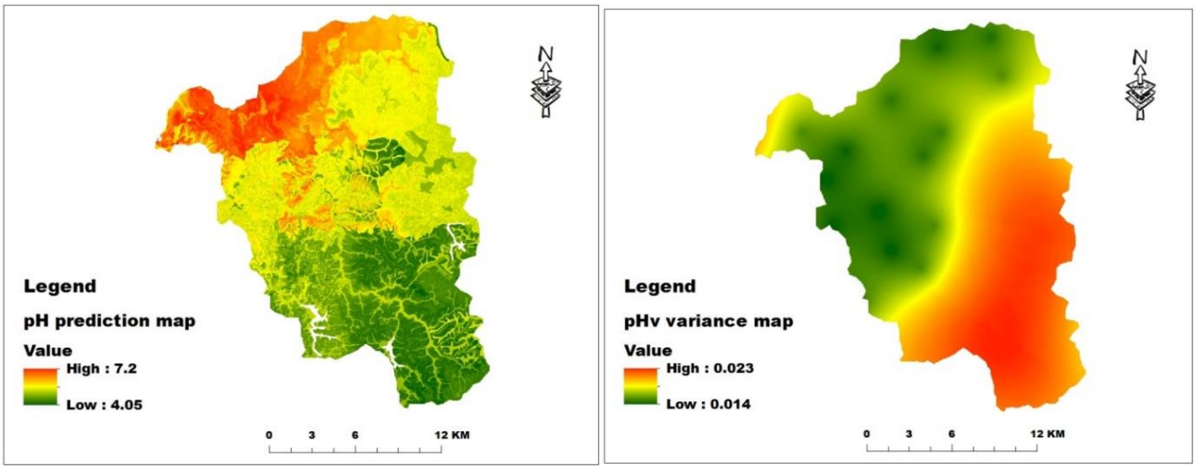

Figure 7: pH prediction map (left), pH variance map (right)_random forest kriging.

3.4 Output map quality and compared with ISRIC and VFRC maps

Much more effort has been devoted in this research for the purpose of making the comparison between the quality of output nutrients maps produced in this study and maps generated earlier by International Soil Reference and Information Centre (ISRIC)and the Virtual fertilizer Research Centre (VFRC) using either similar or different prediction and validation methods. The tables below summarize the comparison in terms of accuracy of the produced maps.

Table2: Output quality assessment using various validation indices.

Soil property	Depth	Model	R^2	RMSE	MAEE
Al	20 cm	RF	0.26	165.3	146.4
Ca	20 cm	RF	0.26	471.26	433.58
Cu	20 cm	RK	0.89	1.04	1.13
P	20 cm	RF	0.07	56.44	48.01
pH	20 cm	RK	0.06	0.73	0.62

Table3: ISRIC quality comparison using similar validation indices.

Soil property	Depth	Model	R^2	RMSE	MAEE
Al	20 cm	Mixed	0.55	312.00	240.91
Ca	20 cm	Mixed	0.45	695.73	500.89
Cu	20 cm	Mixed	0.37	200.45	196.75
pH	20 cm	Mixed	0.12	0.76	0.60

Table4: VFRC quality comparison using similar validation indices.

Soil property	Depth	Model	R^2	RMSE	MAEE
Ca	20 cm	RF	0.45	917	582
Cu	20 cm	RK	0.40	1.43	0.86
P	20 cm	RF	0.22	69.8	26.6
pH	20 cm	RF	0.40	0.68	0.54

4. CONCLUSION

This research has focused mainly on spatial analysis and interpolation of soil properties to establish their spatial distribution and the importance of landscape parameters on variation of soil properties. The mapped soil nutrients, which were five (Aluminium (Al), Calcium (Ca), Copper (Cu), Phosphorus (P) and pH), while Regression Kriging (RK) and Random Forest Kriging (RF) were 2 methods used for interpretation. Soil pH is regarded as a master variable in soils since it influences other nutrients availability for plant while rainfall is one of the greatest factors affecting soil acidity, and therefore some managerial practices are required to reduce soil acidity problems in Sebeya catchment. Additionally, Sebeya catchment acid soils are more likely associated with Al toxicity as demonstrated by the map patterns of Al distribution. Therefore, it would be recommended for farmers to apply alkaline materials such as lime to increase the soil pH thus eliminating Al toxicity, and apply P fertilizer to increase the availability of P in soils. Similarly, it has been noticed that acid soils are associated with deficiencies in Ca availability, as a result of pH decrease in these soils. The remediation of Ca problems requires nothing else than bringing and keeping soil pH to the desired optimal range for nutrient availability.

It has been found that Cu tends to accumulate in the top soil due to its high affinity with organic matter and consequently, the accumulation of heavy metal copper (Cu) in agriculture top soil will increase the risk of phytotoxicity (toxic effect by a compound on plant growth). Therefore, monitoring copper level in soils and its availability for crops within the optimal required quantities is of paramount in agriculture and environmental management.

Acknowledgements: The authors are grateful to the University of Rwanda (UR) for its support and wish to acknowledge the support of UR_Sweden Program in sponsoring this research and we cannot fail to commend and appreciate the work of various authors used for the present paper.

5. REFERENCES

[1] Mukashema A. *Mapping and modelling landscape-based soil fertility change in relation to human induction. Case study: Gishwati watershed of the Rwandan highlands*. [MSc thesis]. Enschede: International Institute of Geo-Information Science and Earth Observation; 2007.
[2] Jenny H. *Factors of soil formation, a system of quantitative pedology*. New York: McGraw-Hill; 1941.
[3] McKenzie NJ, Ryan PJ. Spatial prediction of soil properties using environmental correlation. *Geoderma. 1999; 89:* 67–94.
[4] Chen C, Condron L, Davis M, Sherlock R. Effects of afforestation on phosphorus dynamics and biological properties in a New Zealand grassland soil. *Plant and Soil. 2000; 220:*151-163.
[5] FAO. *A framework for land evaluation*. [Report]. Rome: Food and Agriculture Organization of the United Nations; 1976.
[6] Mora-Vallejo A, Claessens L., Stoorvogel J, Heuvelink GBM. Small scale digital soil mapping in Southeastern Kenya. *Journal of Catena. 2008; 76:* 44–53. doi: 10.1016/j.catena.2008.09.008
[7] McBratney AB, Odeh IOA, Bishop TFA, Dunbar MS, Shatar TM. An overview of pedometric techniques for use in soil survey. *Geoderma. 2000; 97:* 293–327.
[8] Hengl T, Heuvelink GMB, Stein A. A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma. 2004; 120:*75–93. doi: 10.1016/j.geodema.2003.08.018
[9] NISR. *Rwanda demographic and health survey 2014-15*. [Report]. Kigali: National Institute of Statistics of Rwanda; 2014.
[10] IWRM. *Sebeya catchment plan 2017-2023*. [Report]. Kigali: Ministry of Natural Resources; 2017.
[11] MoE. *Sebeya catchment management plan (2018-2024)*. [Report]. Kigali: Ministry of Environment; 2018.
[12] IWRM. *Sebeya catchment plan 2018-2024*. [Report]. Kigali: Ministry of Natural Resources; 2018.

[13] Lee HY, Kim T, Park W, Lee HK. Extraction of digital elevation models from satellite stereo images through stereo matching based on epipolarity and scene geometry. *Image and Vision Computing. 2003; 21(9):* 789–796.
[14] Stoorvogel JJ, Bakkenes M, Temme JAM, Batjes NH, Ben J. S-World: A global soil map for environmental modelling. *Land Degradation & Development. 2017; 28:* 22 –33, doi:10.1002/ldr.2656
[15] McBratney AB, Mendonça S, Minasny B. On digital soil mapping. *Geoderma. 2003; 117:* 3–52. doi:10.1016/S0016-7061(03)00223-4.
[16] Heuvelink GB, Webster R. Modelling soil variation: past, present, and future. *Geoderma. 2001; 100:*269–301.

Cite this article Félicien Majoro, Albert Niyonzima, Concilie Mukamwambali, Adrien Mugabushaka, and Athanase Twagirimana. DIGITAL SOIL NUTRIENTS MAPPING IN SEBEYA CATCHMENT AGRICULTURAL LAND IN RWANDA. *American Journal of Innovative Research and Applied Sciences. 2020; 11(4): 26-34*.