An Overview of Soil Survey and Classification as a Source of Secondary Information

Guug S. Samuel^{1, 2} Kugbe X. Joseph¹

1. Faculty of Agriculture, University for Development Studies, Tamale-Ghana

2. West African Science Service Center on Climate Change and Adapted Land use (WASCAL), Bolgatanga-

Ghana

Abstract

The extraction of information from surveyed and classified soil maps by desired end-users has increased in recent times due to the high cost involved in the classification and mapping out of such soils by the individual users, and is also a time consuming process. In some instances, the end users do not even understand the methods that were used in producing the maps, the errors associated with it and the potential limitations of usage. Knowledge over the years has shown that estimates contained in some soil maps are not perfect because they are typically based on limited data and limited information. To acknowledge that soil maps are not free of errors, the uncertainty in the estimates may be represented with error bounds that characterize the accuracy of the map. In recent times more advanced methods have been introduced with high precision of accuracy, including the use of artificial neural networks, remote sensing and photogrammetric procedures, combined with geographic information system (GIS). However, research has shown that, traditional soil survey persists as the most popular form of soil mapping and inventory. The need therefore arises to integrate and update rather than ignore the traditional soil survey techniques in favour of modern interpolation techniques.

Keywords: survey, classification, soil maps, secondary information, desired user

1. INTRODUCTION

Soil, can be seen universally as a weathered and fragmented outer layer of the earth's terrestrial surface. However, it can also be defined or viewed based on an individual's or group's perspective: such as agricultural, engineering and geological, etc. It takes nature thousands of years to create a unit soil out of sterile bedrock, but a few decades for the activities of mankind to degrade productive soils (*Hillel, 2003*).

As human life and cyclic productivity depends mostly on the limited soil resources (*Plaster, 2009*), people have varied concepts of what the soil is; which makes this resource a universal entity. The farmer is concerned with ways in which soils influence crop growth and productivity. Often, the farmers' interests do not extend below the depth of soil disturbed by a plough (*White, 2006*). The civil engineer is interested in the composition, strength and consolidation of the soil and the use of hydraulic principles to deal with issues concerning construction. The hydrologist will study the soil and its importance in partitioning water between infiltration and runoff, storage, filtering, physical and chemical support to vegetation, etc. (*Dunne, 1978*).

The different uses of soil have required mankind to critically observe and describe, collect, establish and systematically group facts, principles and methods in order to acquire an in- depth knowledge of the soils, their properties and potential for production and conservation (*Verheye, 2007*). Over the years, these have resulted into the creation of many disciplines in this field, including: pedology, survey or mapping, and classification. Through trial and error, mankind has observed spatial variation and differences in soils, where users can tell where a certain type of soil is coming from mostly based on physical attributes like texture and colour. Through similar observations, one can find a reasonable enough solution to a known soil problem.

Spatially demarcated soils on the field are now produced in world maps, where the entire Earth's surface has been projected in millions of two-dimensional images at a wide range of scales (*Hartemink et al., 2013*). However, Soil maps also depict the distribution of soils on the earth's surface and have an important role in aggregating current knowledge of soil resources and their distribution. In recent years, digital soil mapping activities have strongly increased all over the world and, as a result, soil maps are produced by quantitative and statistical methods (*Brus at el., 2011*); serving as important sources of information on land resources. Use of digital soil mapping has increased in recent years; the extraction of information from these published maps by the desired user (e.g. researchers and students) has also increased in recent times (*Brus et al., 2011*), mostly due to high costs involved in the classification and mapping out of individual soils, which is in itself time consuming. The applicability of soil maps, however, depends on the quality of survey and the scale of mapping (*Oberthur et al. 1999*), which may not be appropriate for management and decision making in some fields.

In most cases, however, the overall aims of soil survey are to produce soil maps, incorporating reports on and interpretations of the soil mapping units or series as stated noted by *McBratney et al. (2000)*. Traditionally, soil management and land-use planning have been the predominant reasons for the creation of soil maps and the conduction of surveys at all scales. However, with increasing concerns on the degradation of the environment and the unsustainable use of natural resources, soil survey has moved from its traditional subjective position to a more

quantitative modelling with increases in accompanying accuracy and decline in issues of uncertainty, *Hennings* and Boess (2000). Others include: identification of spatial variability of soil properties based on discriminant analysis of terrain attributes (*Sinowski and Auerswald*, 1999).

A number of factors are identified as being important for the creation of soil maps. The correctness of soil classification and mapping plays an important role in evaluating the purpose of the end user. For that reason, *Robinson (1949)* singled out classification of soils as one of the most important branches of pedology. The efficacy of soil classification also depends on the environment. *Voltz and Webster (1990)* showed that classifications are performed satisfactorily where the soil changes abruptly. In situations where soils change gradually, classifications are not performed less satisfactorily. In addition, conventional soil maps take no account of the within-class heterogeneity, a factor that is considered in recent soil mapping technologies. *Geib (1923)* gave an example of how within-class heterogeneity received attention over the years. He noted of an earlier classification in which Miami series included practically all material in the Glacial and Loessial Province; which was upland and light coloured. However, a later classification presents that there are at least eight series where previously only one was recognized. This recognition of a great variety of differences in what was previously mapped as one series means that at present we are able to distinguish differences in texture, colour, structure, topography, chemical composition etc, and that we can interpret these to an even finer degree.

After mapping and classification of soils, the extraction of information and applicability will however, depend on the end user (*White, 2006; Dunne, 1978*). Factors to be considered in using such information include the age of that classification, the method used for the classification and more importantly how frequently the information is updated. In general, terms, the quality of produced maps has been directly related to measures of uncertainty (*Bishop et al. 1975*), which include stochastic, deterministic and semantic uncertainties as explained by *McBratney (1992)*. Today, validation of soil maps before usage is very essential as temporal errors associated with unvalidated maps increases with time (*Grunwald, 2009; Webster et al., 1968; Brus et al., 2011; Bishop et al 1975*). In some cases and over time, a large part of the area depicted on the map does not correspond with the true soil type (*Brus et al., 2011*).

The present advances in soil sciences require that a universally accepted system of soil classification be introduced. Such a system should replace or at least complement the numerous systems now in operation worldwide. While this should be the case, such universal definitions come with associated problems. There is the need to look beyond characteristics defined in such systems. It has been a common practice that most scientists look up for defined diagnostic horizons and do not go further to look for other unique variables that can occur concurrently.

The purpose of this paper is therefore to review work done with the various methods, validations and qualities of soil classification and mapping, both conventional and those integrated by digital soil mapping procedures and how they can be usefully employed. A discussion on the errors, uncertainties associated with classifications, potentials, and the limitations of such classifications, including the extent to which the desired user can rely on such data are explored.

2. CLASSIFICATION TECHNIQUES

Soil classification techniques during the last few decades have created a tremendous potential for improvement in the way that soil maps are produced. The field study of soils presents a multitude of soil profile and site data, which are worth reliable and trusted unless fitted into a well-classified technique to which generalizations, relationships and specific categorizations can be developed.

Soil classification systems have been developed to provide engineers, scientists and resource managers with generalized information about the nature of a soil found in a particular location. Several classification systems have been developed for soils, which are in use worldwide. These systems categorize soils according to their general behaviour under given physical conditions.

Traditional soil survey persists as the most popular form of soil mapping and inventory, and in many cases is the only manner in which the highly variable nature of the soil landscape is catalogued. However, artificial intelligence methods such as the neural networks (NN) and fuzzy system (FS) *(Lameck et al., 2002, Brus at el., 2011)*, remote sensing and photogrammetric combined with geographic information systems have recently emerged as promising alternatives to various conventional methods of pattern recognition and classification.

2.1 Local/Indigenous Knowledge

Indigenous knowledge refers to institutionalized local knowledge that has been built upon and passed from one generation to another (*Braimoh*, 2006). Some of these knowledge has been modified and well documented while others are still in the consideration stages as studies are ongoing comparing these indigenous knowledge with scientific models. Soil survey and classifications by this method has caught lots of attention in that sense. This kind of descriptive information is referred to as 'soft' information because it relates indirectly to the primary attribute (*Oberthur et al. 1996*).

Indigenous soil knowledge has been much studied by anthropologists (Sandor and Furbee, 1996), but few soil scientists have used this knowledge for quantitative resource assessment (Arrouavs, 1987), mainly because the integration of local knowledge and modern soil science information posed difficulties. Even though, this technique is considered as the cheaper way of classifying soil, other techniques have recently been tested for their utility, which have improved their accuracy at acceptable costs.

Generally, indigenous knowledge of field survey is highly subjective, however McBratney et al. (2000), stated that this involves the development of a mental model, which relates the soil with landform conditions, followed by formulation of hypotheses, which are then tested by ground-truth survey. In most cases, it normally ends by a mere generalization for more than 100s of square kilometres. Aside that, some soil scientists have also found that use of indigenous knowledge facilitates soil survey for agricultural development and increases the probability that resulting projects will meet community needs and respect cultural values (Tabor, 1992). Though such are the cases, indigenous knowledge on soil classification has seen a tremendous decline for the past few decades because of the increased use of computerised and cost effective soil information systems. Consequently, the interpolation does not take full advantage of the available knowledge about the geological differences in the landscape, natural boundaries, and relief (Heuvelink and Bierkens, 1992). This is unfortunate because the resultant map could undoubtedly benefit from the additional information contained in the indigenous soil map. This point was also pointed out by Bouma (1985), when he stated that care should be taken when traditional soil survey techniques are ignored in favour of modern interpolation techniques and where he advocated the integration of both techniques.

Many developing countries still rely on mostly the traditional soil information for agricultural and engineering purposes (Braimoh, 2006; Ogunkunle 1991; Braimoh and Ogunkunle, 1992). On the other hand, Owing to the importance of soil classification, both in experimental and practical agriculture, most developed nations of the world have developed their own national soil classification systems. Over the years, a number of anticipatory and observatory trends have been use to model the best possible classification of soils based on merged properties of primary importance.

2.2 Artificial neural network technique

Artificial neural networks (ANNs) are chains of statistical learning algorithms with generally, systems of interconnected "neurons" which can compute values from inputs, and are capable of machine learning as well as pattern recognition and are used to estimate or approximate functions that can depend on a large number of inputs and are generally unknown. According to Freire et al. (2013), ANN modelling is able to provide the means to predict soil types at locations where there are no current soil maps, by combining soil map data from other areas with landscape features. Artificial Neural Networks (ANN), has enabled and eases the mapping of spatial distribution of soils in a cheap, more accurate, reproducible, and flexible way in terms of data storage and visualization, using surrogate landscape data that are easy to obtain. The basic element of a neural network is the processing node (Figure 1). Each processing node mimics the biological neuron and performs two functions. First it sums the values of its inputs. This sum is then passed through an activation function to produce the node's output value (Paola and Schowengerdt, 1994). The ANN changes its structure based on external or internal information that flows through the network during the learning phase. A basic ANN consists of three layers, the input, hidden and the output layers otherwise referred to as the back propagation algorithm (Figure 1).





directly use the results produced by the Neural Networks, instead, these results have to be interpreted to user understandable forms (*Baskar et al., 2013*). The interpretation to user understandable forms was an earlier setback in the use of such technology. Today however, this setback has been eliminated by the usage of fuzzy rules which directly point towards easily comprehensible results. *Baskar et al. (2013)* in their work concluded that fuzzy rules provide better accuracy than only performing the process using a Neural Network. In that regard, the performance of artificial neural networks are promising and capable of identifying positive results in soil classification with a high degree of success and accuracy, especially when the distances between the sampled points are reasonable. This may be used as a good source of data generation for soils, but should be calibrated with supervised models when dealing with few data or large distances to obtain high efficiency results. However, there exist different ANN model architectures and there is a shortage of studies that seek to test and compare their performance for various problems, including the spatial prediction of soil classes at a regional scale. Because of the complication of the soil behaviour, a combination of more than one model will lead to more possibility of accurate results and should be exploited. Where such is unattainable more data should be used for the given models.

2.3 Remote sensing and photogrammetric techniques

In recent years thematic mapping has undergone a revolution as the result of advances in geographic information science and remote sensing (*Ali and Kotb, 2010*). Remote sensing and photogrammetric techniques provide spatially explicit, digital data representations of the Earth's surface that can be combined with digitized paper maps in geographic information systems (GIS) to allow efficient characterization and analysis of vast amounts of data. In the modern field of science and remote sensing technology, satellite or aerial imagery contains a large amount of information, for instance the use of Digital Elevation Model (DEM) to derive landscape attributes that are utilized in land forms characterization where important features or data can be extracted and processed for the classification of soils. It is also proven by *Moore et al., (1993)* that linear regression can be used with terrain variables derived from DEM to predict soil attributes (organic matter content, extractable phosphorous, pH and texture) at unvisited sites. However, this requires efficient and intelligent analytical techniques as found in the algorithm of the artificial neural network. There are many satellites, onboard spacecrafts (figure 2) as found in the Ground-Penetrating Radar (GPR), which is able to penetrate into the soil without any disturbance and supply high -resolution subsurface imagery of soil horizon profiles which can then be interpreted by Neural networks (NN). Detailed work about image analysis and classification cited by *Zhang (2000)*, *Paola and Schowengerdt, (1995)* illustrates this in detail.



Figure 2: Satellite and Ground-Penetrating Radar

The developments till date has shown that soil scientists have done a lot of work on soil classification using the GPR technology and has proved the ability of this technology to map soil textural variations, organic matter content, thickness and depth of soil horizons (*Johnson et al., 1982; Doolittle, 1982; Doolittle, 1983; Olson and Doolittle, 1985; Collins and Doolittle, 1987; Schellentrager et al., 1988; Truman et al., 1988; Mokma et al., 1990; Raper et al., 1990; Doolittle and Amussen, 1992; Collins, 1992; Doolittle and Collins, 1995, Freeland et al., 1998)*. Many studies such as Stoner *et al.* (1980), Henderson *et al.* (1992) and Csillag *et al.* (1993) have come out with comprehensive soil surveys that have shown the potential benefits of using remote sensing for soil identification and mapping. Others including *Agbu et al.* (1990), *Coleman et al.* (1993), *Seyler et al.* (1998) also carried out remote sensing studies using spectral reflectance in the survey of soils, based on broad-band sensors such as Landsat TM.

3. CONCLUSION

Although soil surveys have been carried out in many countries, the scale and area coverage of the resulting soil maps are not ideal for planning and applications at national levels. Additionally, there is a lack of consistency

between soil classifications and legends across countries, which contributes towards a slow progression in integrating soil datasets. Globally, about two-thirds of the countries have soil maps at a 1:1 million scale or finer, but over two thirds of the total land area has yet to be mapped even at a 1:1 million scale. The artificial neural network method is an effective tool in solving complex, nonlinear and causal problems. Neural networks have been applied to solve a wide variety of problems. The consistency, accuracy, and the volume of learning data, are enabling factors that strengthen this method. In *Zhang et al, 1993* using image processing and GIS in their work, found a closed match in the soil distribution and the land use types and concluded that GIS is a powerful tool to improve soil and land use mapping but largely based on the remote sensing data. Though such elite methods exist, most countries still rely on the traditional soil information system for many purposes. Since that is a norm, care should be taken to integrate rather than ignore the traditional soil survey techniques in favour of modern interpolation techniques. There is therefore the need to always update existing soil maps to the benefits of the end user.

Reference

Agbu, P.A., Fehrenbacher, D.J., Jansen, I.J., (1990): Soil property relationships with SPOT satellite digital data in east central Illinois. *Soil Science Society of America Journal* 54, 807–812

Arrouays, D. (1987). Farmers' knowledge and soil survey: a multivariate treatment of agropedological data. *Soil Survey and Land Evaluation*, 7(2), 81-86.

Baskar S.S., Arockiam L., and Charles S., (2013): Enhancing the Interpretability of Soil Classification Data from a Neural Network System by Using Fuzzy Approaches. *International Journal of Advanced Research in Computer Science and Software Engineering*, 3(10)

Bishop, Y.M.M., Fienberg, S.E. and Holland, P.W. (1975): Discrete Multivariate Analysis: Theory and Practice. MIT Press, Cambridge.

Brus, D. J., De Gruijter, J. J., Marsman, B. A., Visschers, R., Bregt, A. K., Breeuwsma, A., and Bouma, J. (1996). The performance of spatial interpolation methods and choropleth maps to estimate properties at points: a soil survey case study. *Environmetrics*, 7(1), 1-16.

Braimoh, A. K. (2006): Random and systematic land-cover transitions in northern Ghana, *Agriculture, ecosystems and environment*, 113(1), 254-263.

Bregt, A. K., Stoorvogel, J. J., Bouma, J., and Stein, A. (1992): Mapping ordinal data in soil survey: a Costa Rican example. *Soil Science Society of America Journal*, 56(2), 525-531.

Collins, M. E., and Doolittle J. A., (1987): Using ground penetrating radar to study soil microvariability. *Soil Science Society of America Journal*, 51(2): 491-493

Collins, M. E. (1992): Soil taxonomy: A useful guide for the application of ground penetrating radar, *In Fourth International Conference on Ground Penetrating Radar*.

Coleman, T.L., Agbu, P.A. and Montgomery, O.L. (1993): Spectral differentiation of surface soils and soil properties – is it possible from space platforms? *Soil Science* 155, 283–93

Csillag, F., Pasztor, L., Biehl, L.L. (1993): Spectral band selection for the characterization of salinity status of soils. *Remote Sensing of Environment*, 43, 231–42

Dobos, E., Carré, F., Hengl, T., Reuter, H. I., and Tóth, G. (2006): Digital soil mapping as a support to production of functional maps. Luxemburg: Office for Official Publications of the European Communities, EUR, 22123.

Doolittle, J. A. (1982): Characterizing soil map units with ground-penetrating radar. *Soil Survey Horizon*, 22(4) 3–10.

Doolittle, J. A. (1983): Investigating Histosols with ground-penetrating radar. *Soil Survey Horizons*, 23(3): 23-28. Doolittle, J. A. and Amussen L. E. (1992): Ten years of applications of ground-penetrating radar by the United States Department of Agriculture. In: Hanninen and Autio (editors), 4th Int. Conf. Ground Penetrating Radar, Geol. Surv. Finl. Spec. Paper 16: 139-147

Doolittle, J. A. and. Collins M. E. (1995): Use of soil information to determine application of ground penetrating radar. *Journal of Applied Geophysics*, 33(1-3). 101 - 108.

DuBose, P., Klimasauskas, C., (1989): Introduction to Neural Networks with Examples and Applications. NeuralWare Inc., Pittsburgh

Dunne, T. (1978): Field studies of hillslope flow processes. Hillslope hydrology, 227, 293

Freeland, R. S., Yoder, R. E., and Ammons, J. T. (1998): Mapping shallow underground features that influence site-specific agricultural production. *Journal of Applied Geophysics*, 40(1), 19-27.

Grunwald, S. (2009): Multi-criteria characterization of recent digital soil mapping and modeling approaches. *Geoderma*, 152, 195–207.

Hartemink, Alfred E., Pavel Krasilnikov, and Bockheim J. G. (2013): "Soil maps of the world." *Geoderma* 207, 256-267.

Henderson, T.L., Baumgardner, M.F., Franzmeier, D.P., Stott, D.E. and Coster, D.C. (1992): High dimensional reflectance analysis of soil organic matter. *Soil Science Society of America Journal* 56, 865–72

Hennings, V., and Boess, J. (2000): User-oriented concepts to assess the accuracy of nationwide land quality maps. In Proceedings of 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences (pp. 301-304)

Heuvelink, G.B.M. and Bierkens, M.F.P. (1992): Combining soil maps with interpolations from point observations to predict quantitative soil properties. *Geoderma*, 55(1), 1-15

Hillel, D. (2003): Introduction to environmental soil physics. Academic press

Johnson, R. W., Glasscum R., and Wojtasinski R., (1982): Application of ground penetrating radar to soil survey. *Soil Survey Horizons*, 23 (3), 17–25.

Lameck, O. O. Robert, S. F. Ronald, E. Y. Wesley Hines, J. (2002): Application of Fuzzy-Neural Network in Classification of Soils using Ground-penetrating Radar Imagery (No. 023097). ASAE Paper

McBratney, A. B., Odeh, I. O., Bishop, T. F., Dunbar, M. S., and Shatar, T. M., (2000): An overview of pedometric techniques for use in soil survey. *Geoderma*, 97(3), 293-327.

McBratney, A.B. (1992): On variation, uncertainty and informatics in environmental soil management. *Australian Journal of Soil Research* 30, 913–35

Mokma, D. L., Schaeltzl R. J., Johnson E. P., and Doolittle J. A., (1990): Assessing Bt horizon character in sandy soils using ground-penetrating radar: Implications for soil surveys. *Soil Survey Horizon*, Vol. 30(2): 1-8.

Olson, C. G., and J. A. Doolittle., (1985): "Geophysical techniques for reconnaissance investigations of soils and surficial deposits in mountainous terrain." *Soil Science Society of America Journal* 49 (6) 1490-1498.

Ogunkunle A. O. and A. K. Braimoh (1992): "Prospects of numerical correlation of Nigerian soils: A comparison of cluster and principal components analysis." *Journal of Agricultural Science and Technology*, 2(2): 101–110

Paola, J. D. and. Schowengerdt R. A. (1995): A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery. *International Journal of Remote Sensing*, 16(16): 3034-3058.

Plaster, E. J., (2009): Soil science and management. Cengage Learning.

Raper, R. L., L. E. Asmussen and Powell J. B. (1990): Sensing hard pan depth with ground-penetrating radar. Transaction of ASAE, 33(1): 41-46.

Robinson, G. W. (1949): Soils, their origin, constitution and classification, Murby, London, 3rd Edition.

Freire, S., de Lisboa, N., Fonseca, I., Brasil, R., Rocha, J., and Tenedório, J. A. (2013): Using Artificial Neural Networks for Digital Soil Mapping–a comparison of MLP and SOM approaches. AGILE.

Schellentrager, G. W., Doolittle, J. A., Calhoun, T. E., and Wettstein, C. A. (1988): Using ground-penetrating radar to update soil survey information. *Soil Science Society of America Journal*, 52(3), 746-752.

Seyler, F., Bernoux, M. and Cerri, C.C. (1998): Landsat TM image texture and moisture variations of the soil surface under the rainforest of Rondonia state, Brazil. *International Journal of Remote Sensing* 19, 1299–318.

Sandor, J. A., and Louanna Furbee (1996): "Indigenous knowledge and classification of soils in the Andes of Southern Peru." *Soil Science Society of America Journal* 60(5) 1502-1512.

Sinowski, W., and K. Auerswald., (1999): "Using relief parameters in a discriminant analysis to stratify geological areas with different spatial variability of soil properties." *Geoderma* 89(1) 113-128.

Stoner, E.R., Baumgardner, M.F., Weismiller, R.A., Biehl, L.L. and Robinson, B.F. (1980): Extension of laboratory-measured soil spectra to field conditions. *Soil Science Society of America Journal*, 44, 572–74

Oberthür, T., Goovaerts, P., and Dobermann, A. (1999): Mapping soil texture classes using field textuing, particle size distribution and local knowledge by both conventional and geostatistical methods. *European Journal of Soil Science*, *50(3)*, *457-479*.

Tabor, J. A. (1992): "Ethnopedological surveys: soil surveys that incorporate local systems of land classification." *Soil survey horizons (USA)*, Vol.30 (1)

Truman, C. C., Perkins, H. F., Asmussen, L. E., and Allison, H. D. (1988): Using ground-penetrating radar to investigate variability in selected soil properties. *Journal of soil and water conservation*, 43(4), 341-345

Verheye, Willy, (2007): Soils and soil sciences. *Land use, land cover and soil sciences, Vol.6* National Science Foundation, Flanders/Belgium, and Geography Department, University of Gent, Belgium

Voltz, M., and Webster, R. (1990): A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. *Journal of Soil Science*, *41*(3), 473-490.

Geib, W. J., Krusekopf, H. H., and Veatch, J. O. (1923): Report of the Committee on Improvement of Soil Maps and Reports. *Soil Science Society of America Journal*, 4(2), 102-103.

White, J. G., Welch, R. M., and Norvell, W. A. (1997): Soil zinc map of the USA using geostatistics and geographic information systems. *Soil Science Society of America Journal*, *61*(1), 185-194.

Zhang, G. P. (2000): Neural networks for classification: A survey. IEEE Trans. on Systems, Man, and Cybernetics, Part C: Applications and Reviews, Vol.30 (4). 451-462