Vidyodaya J. Arts, Sci. Lett., Silver Jubilee pp. 501-520

PRINCIPLES OF SOIL TAXONOMY

by

P. WICKRAMAGAMAGE

1.0 The definition of soil

The study of soils as an independent science is mainly due to the works of the Russian School of Pedology led by Dukuchaev. The concept that soils were independent natural bodies was first introduced by Dukuchaev and they were conceived as products of a combination of processes and factors namely climate, living matter, parent material, relief and age. This was a revolutionary idea (Soil Survey Staff, 1960; p.1), which made it possible to investigate soils themselves rather than infering from other factors. This Russian concept of soil, according to Marbut "established the study of soils firmly as an independent science with criteria, point of view, method of approach, process of development applicable to soils alone and inapplicable to any other series of natural bodies" (quoted by Basinski, 1959). These developments were not known to the rest of the world until the Glinka's famous text on world soils was translated into English (Soil Survey Staff, 1951, p. 3). The early American and W. European pedologists did not examine the soil in depth, often confined their investigations to the plough layer and soils were *studied* as a part of related sciences (e.g. geology). Soils were sometimes treated as static storage bins for plant nutrients (Soil Survey Staff, 1951, p.1).

Having identified soils as independent natural bodies, Dukuchaev tried to present a definition of soil. As quoted by Glinka (1928, p. 2), the soil was defined by Dukuchaev as "the layers of materials lying on the surface of the earth or near it which have been changed by natural processes under the influence of water, air and living and dead organic matter." In his definition of soil the presence of genetic horizons with properties reflecting the effects of local and zonal soil forming processes led to the exclusion of those soils which have no genetic horizons or are not thick enough (USDA, 1975). On the other hand it is not possible to distinguish between soil and parent materials. The definition of soil adopted by the U.S. Soil Survey can be treated as a logical development of the Russian concepts of soils and the contribution of Marbut and others in the 1920s and 1930s. In presenting the 7th Approximation System, the Soil Survey Staff (1960, p. 1) defined the soils as a "collection of natural bodies on the earth's surface containing living matter and supporting or capable of supporting plants. At its upper limit is air or water. At its lateral margins it grades to deep water or barren areas of rock, ice, salt or shifting desert sand dunes. Its lower limit is perhaps the most difficult to define.................The lower limit of soil, therefore, is the lower limit of the common rooting of the perennial plants." In this definition the genetic horizons are not included and therefore soils of recent origin were also included.

1.1 Principles of soil taxonomy

According to Gilmour (1936) "classification is primarily utilitarian. It is a tool by the aid of which the human mind can deal effectively with the almost infinity variety of the universe. It is not something inherent in the universe but it is a conceptual order imposed on it by man for his own purposes." Classification therefore, is essentially utilitarian in that its uses could be preconceived and *limited* or could be left undefined. As has been pointed out by Gilmour (1937) it is stated in logic that more propositions could be made regarding the constituent members of the natural classification than about the population as a whole. Broadly speaking there are two types of classifications, namely.

(1) classification for a pre-defined purpose or purposes.

(2) classification, of which uses may not have been defined *a priori*.

The former classification is described by various terms such as extrinsic, artificial or special purpose. It best serves the purpose or the purposes defined and is based on the characteristics (attributes) specially relevant to the proposed need. The classification of soils for agricultural use or for engineering purposes are examples of special purpose classifications.

The second type is generally known as natural, general purpose, intrinsic or taxonomic classifications. The soil taxonomist who is interested in soil classification needs such a system to serve as a frame of reference for soil mapping and other soil studies. However, there are conflicting views regarding the nature of a natural classification and the natural group or the taxon. Early ideas of a natural system were based on Aristotelian logic. As has been pointed out by Sneath and Sokal (1973, p. 19) the purpose of the Aristototelian system as applied to taxonomy is to discover the essence of a taxonomic group (natural taxon) in such a way that essence is expressed in axioms that give rise to properties which are inevitable consequences. They have illustrated this logical system by the example of a triangle on a plane surface. The essence of a triangle is expressed by its definition as a figure bounded by three straight sides, and an inevitable consequence is that any two sides together are longer than the third. This logical system is described as a "system of analysed entities" which is not suitable to classify natural entities which represent a"system of unanalysed entities" (Sneath and Sokal, 1973, p. 19–20). It is not possible to define natural groups in such a way that many consequences follow from the definition without exception (Sneath, 1964). Although these ideas have been expressed in relation to biological taxonomy, they also have relevance to soil classification. Soil groups cannot be defined using "essential characteristics" since they are not known to the soil taxonomist prior to classification. The 'a priori' weighting of characteristics of soils would lead to a special purpose classification (Gilmour, 1936; Kubiena, 1958; Basinski, 1959). Therefore, the Aristotelian system cannot be applied to soil taxonomy.

Another approach to natural classification is one which is based on the phylogenetic relationships and can only be applied to the products of organic evolution (Crowson, 1970, p. 95–114). The biological taxonomists who have adopted this approach stress that the natural taxa should be constructed to reflect evolutionary relationships. This concept has no parallel in pedology, and cannot be treated as an adequate basis for all natural taxa.

Adanson rejected ideas of 'a *prior*' assumptions and pointed out that natural taxa should be based on the similarity measured taking all characteristics into consideration (Sneath and Sokal, 1973, p. 5). Therefore, a natural taxonomic classification should ideally be based on the intrinsic properties of the objects to be classified. According to Gilmour (1936) the terms 'artificial' and 'natural' are relative and no classification can be based on all attributes for reasons given below. Therefore a given classification can be more natural than another depending on the range of attributes used in the definition of groups. A classification based on a large number of attributes is more useful as a general reference system.

Soil taxonomists are not in agreement as to what a true 'taxonomic' classification is. Soils have developed in diverse parent materials and environments, and therefore are not related in biological sense. However, concepts of soil genesis have entered into soil classification directly or indirectly (Avery, 1968). A natural soil classification was conceived by many as one which would reflect genetic relationships. This concept of natural classification was derived from biology (Gilmour, 1961). The most important characteristics, which were considered as the basis of classification, were derived by circular argument by inspection of natural groups already recognized empirically (Cain, 1958). However, the Russian pedologists always recognized the 'genetic soil type' as the basic unit of soil classification (Basinski, 1959). Even the most recent soil classifications in the USSR have paid more attention to pedogenetic factors and the geographical environment than to soil properties. This emphasis on pedogenesis can be found in soil classification systems elsewhere as well. For example, the current USDA system has chosen some differentia which reflect soil genetic processes. The use of genetic homogeneity as an objective of soil classification runs into trouble for three reasons.

(1) The genesis of most soils is not known at the time of the classification or is controversial.

(2) The apparently genetically homogeneous soils may not necessarily be homogeneous in their intrinsic properties. Therefore, such a classification is no more than a special purpose classification of soil genetic factors and processes.

(3) The pedogenetic environments have changed over time and consequently some soils have developed under more than one genetic environment.

It may not be possible to determine the genesis of soils before classification but a successful soil classification could lead to better understanding of their genesis.

The natural soil classification defined by Kubiena (1953) was one which, it was claimed, was based on all characteristics of soils. The use of all characteristics of soil, however, is not possible (Gibbons, 1968) for three reasons :

(1) some attributes may not be known at the time of classification.

(2) the attributes that are known but not evaluated cannot be used. This is a particular problem when soil survey data are used to classify soils.

(3) the constraints of data manipulation. This constraint has been remarkably reduced by the development of electronic computers. But certain mathematical techniques do require the number of attributes to be less than the number of individuals.

Apart from these constraints some attributes could be excluded from the classification strategy when there is a strong inter-attribute correlation (Sarkar *et al.* 1966). Thus those attributes which have the maximum variance and the largest number of accessory properties should be used to classify soils.

Kubiena's concept of natural classification is not accepted by Leeper (1956), who suggests that a classification could be simply good or bad but not natural or artificial as claimed by the former. This is contrary to the general principles of taxonomy and cannot be accepted because this attitude would only lead to a special purpose classification. Muir (1962) has used the periodic classification of elements to illustrate the concept of an ideal natural classification and suggests that the soil taxonomists should try to produce a comparable classification system. But the task of the soil taxonomist is much more difficult. There was a unifying theory regarding the nature of the elements of the periodic table prior to the classification and the laws of chemistry were established unlike those of soil science. As far as soil is concerned there is no such theory that could help define and interpret soil groups. Also soil individuals are not discrete entities like elements in the periodic table. The soil universe is made up of an infinite number of individuals which are not mutually exclusive. Kubiena's approach to soil classification is based on Adansonian principles of natural taxonomy. Although it is not possible to give equal weight to all attributes for the reasons given above, equal weight can be given to all those attributes which are used in classification. When data is available on a large number of soil properties, a representative set of attributes could be chosen to characterize soil individuals.

1.2 Soil universe

Classification of soils involves the definition of the soil universe. The soil universe is made up of all soil individuals and classes, and therefore is a super class (Knox, 1965). The nature of the soil universe is fundamental to the understanding of the soil classification. Gibbons (1968) has suggested that there are three models of the nature of the soil universe.

(1) The soil universe is essentially particulate (made up of discrete natural individuals). The classes are identified by peaks in the distribution curves of attributes (Kubiena, 1958); individuals and classes are found but not constructed. Soil individuals do not have clearly defined boundaries (section 1.3).

(2) The soil universe is essentially continuous, both classes and individuals are constructed not found (Knox, 1965). It is important to recognize the continuous nature of the soil universe but it is possible, for practical purposes, to describe soil individuals in the field. The continuous nature of soils may hide the differences between soils and the task of the taxonomist is to uncover the hidden differences.

(3) The soil universe is continuous, the individuals are *real* and sound but the classes are abstract (Soil Survey Staff, 1960, p. 1–11).

The third view was adopted here and the classification of soils is considered as the identification of those soil groups whose members are similar in their intrinsic properties. Therefore, soil groups are not truths to be discovered by man but they are constructed for practical purposes and such a classification is essentially utilitarian. All soil groups are associated with a certain degree of variation in their properties. A successful classification would minimize this variation.

1.3 Soil individual

The main difficulty in the definition of the soil individual is due to the fact that it cannot be identified as an exclusive entity like biological organisms or the elements in the periodic table. But a soil body is a natural entity, the properties of which can be observed in the field. Knox (1965) suggests that only a particulate universe has natural individuals and the individuals in a continuous universe are artificial individuals created in the absence of natural individuals for the purpose of classification. This distinction between the artificial and natural individuals may be considered as an over simplification. As has been suggested by Soil Survey Staff (1960) "a soil individual is not found as a distinct entity clearly separated from all others, but grades on its margins to other soil individuals with unlike properties." Therefore, the soil individual is a natural individual which can be found in the field.

Knox (1965) has described eight soil bodies which have been used in of which Soil Studies six soil bodies can be identified as forming the concept of soil individuals for various purposes.

(1) *Primary particles*, such as crystals or crystal fragments. These particles do not form a soil and, therefore, cannot be used as the soil individuals in the classification of soils.

(2) *Hand specimen*. These are samples of soil materials used for laboratory determinations. A given hand specimen is treated as a homogeneous sample and is only a part of a soil body and does not represent the soil body as a whole.

(3) *The soil horizon*. The soil horizon has all the characteristics of the hand specimen plus the thickness. A soil horizon can be observed in the field (where present) and is a layer of relatively homgeneous soil material. The soil horizon has been used in soil classification as the soil individual by Rayner (1956). Fitzpatrick (1967) has proposed a system of soil classification using the soil horizon sequences to determine the similarity between soils. However, the soil horizon cannot be used as a soil individual, as it is only a part of the soil (section 1.0).

(4) *The soil profile.* This is a vertical cross-section of soil without the lateral dimension. It represents the vertical variation of soil properties and the horizonation can also be observed where present. The importance of the soil profile as a unit of classification was introduced to the U.S. Soil Survey by Marbut (Simonson, 1952). This concept had been described by Dukuchaev much earlier than Marbut but was not known to those outside Russia. Simonson (1952) claims that the introduction of the soil profile to soil science was comparable to the introduction of anatomy to biology some centuries ago. Marbut believed that all soils at maturity developed a soil profile and the features of the soil were expressed as the features of the soil profile. This concept has influenced the methods of the modern soil surveys in the world. Both in the USA and the UK representative soil profiles are used to charactcrice the Soil Series. Therefore, collection of soil data have been done using the soil profile as the basic unit of sampling. For practical reasons the soil profile has an area of about $1m^2$ and the depth is left undetermined. However, in practice the depth of soils is considered to be the rooting depth of the perennial plants (Soil Survey Staff, 1960, p. 1). The soil profile is the most effective unit of soil which could be used as the soil individual. The main criticism of the use of the soil profile as the basic unit of classification has been that it does not represent soil in three dimensions, but it is not possible to collect data on larger soil bodies efficiently.

(5) *Pedon.* The pedon is the smallest three dimensional unit that could be called as soil (Soil Survey Staff, 1960, pp. 2–4). It includes all the characteristics of the soil profile plus lateral variation. The area of a pedon varies from $1m^2$ to $10m^2$ depending on the variability of the soil horizons.

(6) *The soil landscape unit.* This is a geographical body of soil which is a mappable unit unlike the previous five soil bodies. The soil series, which is the lowest category of the soil survey of England and Wales classification belongs to this. The soil landscape unit cannot be considered as the soil individual because of the high degree of heterogeneity involved in the definition of such landscape units. For detailed soil surveys they are useful as mapping units, the identification of which should be done after the classification of soils.

The other two soil bodies, the delineated soil body and the soil type, described by Knox (1965) are also soil landscape units which have been used for mapping purposes rather than for soil classification.

Among the soil bodies considered, the soil profile is the most convenient unit which could be used as a soil individual for the collection of data and classification. The other units described above are either not representative soil bodies or they are too large to be homogeneous enough to be used in soil classification. Therefore, the classification of soils by numerical methods, will involve the classification of soil profiles. Identification of soil mapping units similar to soil profile groups is the task of the soil surveyor.

1.4 Taxa

The meaning of the taxonomic group or the taxon was given earlier but it is necessary to define it in relation to soil classification. Classification of soil individuals (profiles) involves the determination of affinity (taxonomic similarity defined by Sneath and Sokal, 1973, p. 31–40) between individuals with respect to all, or an adequate number of soil properties. This has been done by the traditional taxonomists by defining a set of diagnostic features. In the USDA system, the soil orders are defined by using a small number of diagnostic features. In the Soil Survey of England and Wales Classification System (Avery, 1973, 1980), the higher categories are identified using diagnostic features termed as 'keys.'

The taxonomic class (taxon) is necessarily a polythetic class based on all characteristics or an adequate number of characteristics. The objective of numerical taxonomy is to define such groups. The USDA (1975, p. 9–10) considers that the choice of the attributes (characteristics) should be made in such a way that the chosen attributes should have the greatest number of accessory attributes. Such attributes have been generally considered as related to the genesis of soils. It has been pointed out by Webster (1977) that the soil properties do not *covary* as had been expected earlier. This may be due to the fact that such correlations may not exist between soil properties observed for different soil populations. The nature of correlation between soil properties may vary with soil groups. But Sarkar *et al.* (1966) demonstrated that the sixty-one soil attributes used in their classification of Kansas soils were correlated and when the number of attributes was reduced to twenty-two, the same results could be obtained. Because of such disagreements it is necessary to classify soil attributes prior to the classification of soils.

1.5 Historical development of soil classification

Prior to the works of Dukuchaev and his colleagues, soil was studied as a part of other sciences such as geology and agriculture. There was no clear definition of soil although the importance of it had been recognised. Some early attempts to classify soils were made in Western Europe. Thaer in 1853 proposed a soil classification based on textural properties at the primary level and agricultural properties at lower categorical levels. In 1886 Richthofen proposed a classification with the emphasis on geological properties and nomenclature. These early classifications reflect the state of contemporary pedological knowledge. Buol *et al.* (1973, p. 174) describe such classifications as technical classifications.

Dukuchaev (1846–1903) recognized soils as independent natural bodies and attempted to classify them, in part, with respect to their properties. Dukuchaev's work on the Russian Chernozem states that "soils must be classified according to their *Properties*" (quoted by Buol *et al.* 1973, p. 175). However, in practice he used soil properties only at the lower categorical level and the highest categories were separated using environmental factors on the assumption that they were related to the broad climatic and vegetational zones. Dukuchaev and his colleagues (specially Sibirtsev and Glinka) gave a great importance to soil genesis and soil properties were chosen in such a way that they would reflect the genetic environment and the factors of soil formation.

These ideas are comparable with the views of the early biological taxonomists that a taxonomic classification should be based on the Aristotleian logic. This trend continued in the USSR. The soil classification proposed by Kovda *et al.* (Fitzpatrick, 1980, p. 174) is claimed to be an historical genetic classification using properties which reflect the evolution of soil in time. This system breaks with the old Russian tradition by using soil properties rather than environmental factors as a basis for soil classification but it is identical with the old system in that it was the interpretation of facts (data) rather than facts themselves which were used. Another recent Soviet soil classification proposed by Rozov and Ivanova is described by Avery (1968) as a coordinate system. In this system the categories below 'Types' are based on their relationship with three groups of soil properties (Coordinate axes).

Axis 1 The properties of soil and environment that can be little changed by man.

Axis 2 The moisture characteristics of soil.

Axis 3 Bio-physico-chemical soil ranges

(i) peculiarities of organic matter decomposition

(ii) saturation and absorption complex and cation exchange complex

(iii) general structure of the soil profile and the presence or absence of carbonates, gypsum and soluble salts.

This system is comparable with that of Avery (1968) in principle but the categories above Types have not been worked out. There has not been a substantial change of emphasis in the tradition of soil classification in the USSR.

The concepts of soil classification worked out by Dukuchaev and subsequently developed by Sibirtsev and Glinka were introduced to the West by Marbut by translating the German edition of Glinka's text on the world soils. The US soil classifications used during the period 1899–1922 have been described by Buol et al (1973, p. 176–177) as single factor soil classifications with a bias towards geological techniques and the nomenclature. In Western Europe and in America the concept of soil geology (geological derivation of soils) prevailed during the 19th century (Cruickshank 1972, 1972, p. 13–31). Hilgard (1833–1906) was the first in the USA to recognize soils as independent natural bodies (Buol *et al*, 1973, p. 176), though his ideas were not applied in operational soil surveys in the USA. After him Coffey suggested in 1912 that soils were natural bodies, the classification of which should be based on soil properties. These ideas did not have an impact on the pedological thinking of the USA until Marbut introduced the Russian concepts.

Marbut can be considered as the founder of modern soil taxonomy in the USA. He not only introduced the ideas of Dukuchaev, Sibirtsev and Glinka, but developed his own ideas of soil classification and survey. The contribution made by Marbut has been summarized by Buol *et al* (1973, p. 171–181) under three headings.

(1) The establishment of the soil profile as the unit of study and the emphasis on soil properties (section 1.3).

(2) Establishment of criteria for soil series.

(3) The preparation of the first hierarchical multi-categoric system. This concept has been developed by the Soil Survey Staff (1960) to produce the 7th Approximation soil classification system.

Marbut put more emphasis on the soil properties than genesis in devising his system of soil classification. He divided soils into two major classes :

(i) Pedalfers

(ii) Pedocals

Pedalfers are those soils that accumulate sesquioxides while Pedocals have a horizon of carbonate accumulation. This system encountered problems when it tried to include Brownearths most of which accumulate neither iron nor carbonates (Fitzpatrick 1980, p. 126). However, this system can be treated as the true beginning of modern US soil classification. Along these lines a comprehensive multi-categoric system was proposed in 1938 by Baldwin, Kellog and Thorp (Buol *et al*, 1973, p. 179) but they included zonal concepts of Sibirtsev. The highest categories of this system, unlike those of Marbut, were defined in genetic terms.

All the soil classifications proposed in the USA prior to the 7th Approximation were qualitative to varying degrees. The classes were not defined using quantitative measurements and as a result the decisions on criteria were made subjectively. However, it can be seen over time there was more incorporation of quantitative data at various levels.

The 7th Approximation system (Soil Survey Staff, 1960) was proposed as a general purpose classification based on a large number of observed soil properties. The general purpose classification was conceived as a multicategoric system along the lines of the previous soil classifications of the USA, and such a system was conceived as hierarchical in organisation. The system could be put to a multitude of uses at its various categorical levels. The 7th approximation is an attempt to rationalize the criteria used to define various classes although the decisions on the choice of them have been made subjectively. Ragg and Clayden (1973, p. 12–13) have summarized the criticisms of the USDA system coming from various sources. These criticisms reflect the conflicting views held by soil taxonomists of different countries. For example, while some (Webster, 1968) object to the system for using genetically important properties, others have pointed out that no adequate consideration has been given to such properties (Duchaufour, 1963; Gerasimov *et al*, 1964.) However, the choice of the differentia has been subjectively made and, therefore, a large amount of information gathered on soils has not been used consistently and objectively. Webster (1968) claims that the fundamental fault of the system is its hierarchical organisation of the categories.

In the last century and in the early part of this century, important contributions were made towards the recognition of the study of soils as an independent science in Western Europe. In this respect, works of Muller (Sweden) and Ramann (Germany) are of great importance. In Great Britain the study of soils started formally with the establishment of a research institute at Rothamsted, England in 1843. Soil survey in Great Britain dates back to 1911 (Cruickshank, 1972, p. 23) when the first of special surveys were published. Robinson (1932) and Avery (1956) have been responsible for the development of the Soil Survey of England and Wales. A separate soil survey for Scotland was established in 1930. Three soil classification systems have been used in modern times in different parts of the British Isles. Since the beginning of modern British soil classification in the 1930s, morphological properties have been widely used. The soil series were defined by taking the nature and the sequence of soil horizons into consideration. The most recent soil classification proposed by Avery (1980) for England and Wales is a somewhat different approach, which originated as a coordinate system. This is also a multi-categoric system which is described as nonhierarchical by Avery (1968). The class differentia have been subjectively determined, as a result this system also has failed to eliminate the basic problem of subjectivity of all traditional classifications. The similarity between soil profiles has been determined on the basis of the presence or the absence of pre-selected diagnostic characteristics. In addition the existence of three soil classification systems for the British Isles may lead to inconsistencies and create difficulties in the communication of soil information among different authorities.

The majority of the traditional soil classifications have been devised to serve the needs of the country concerned and therefore they have taken local conditions into consideration. The USDA system was proposed as a comprehensive classification drawing samples from a wide range of geographical environments, in order to represent as many diverse soils as possible. But tropical soils were under-represented. The soil classification systems devised for individual national soil surveys cannot be applied to other areas successfully. Even the USDA system has been proved to require modification when it is tried in other countries (Ragg and Clayden, 1973; Kesseba *et al.* 1972). The soil classes defined in different countries according to different classification systems cannot be compared easily. Therefore, a great need is felt for an objective system of soil classification using as many properties as necessary.

Several common features of the traditional soil classifications can be identified.

(1) Each of them has been devised to serve a national need and therefore the taxa of the *system* are limited to those which occur in the country for which the classification system was proposed. Although the USDA system is supposed to be a comprehensive system applicable to other countries, the attempts to use it elsewhere revealed the need for modifications and sometimes inadequacies (Ragg and Clayden, 1973; Kesseba *et al.* 1972).

(2) The use of 'a priori' assumptions on the diagnostic criteria. The diagnostic features have been defined prior to the classification even though they should have been discovered after the classification was devised.

(3) The affinity between soil profiles has been determined subjectively.

(4) The properties chosen were weighted without any empirical justification.

(5) Almost all the properties used were related to the soil morphology (Muir, *et al.* 1970).

Despite the drawbacks of the soil classification systems hitherto produced, a general improvement can be detected. This development is from early single factor classifications to more complex systems such as USDA system using a wide range of information on the soils themselves can be viewed as a considerable progress. Over the years the knowledge of the nature of soils has increased tremendously, so that a considerable amount of information about soils is available. The main problem of the classification of soils today is how best this information could be used to devise a taxonomic classification. Crowther (1953) conceived the problems arising from the multi-dimensional nature of soils and suggested that a coordinate system would be suitable to classify soils but the objective use of such a model was not then possible. The advent of the electronic computer made numerical taxonomy possible, consequently a whole range of techniques is available to handle a large amount of data. These methods could be used to replace old subjective methods of soil classification.

1.6 Numerical taxonomy of soils—a review of previous work

The term numerical taxonomy has been defined by Sneath and Sokal (1973, p. 4) as "the grouping by numerical methods of taxonomic units into taxa on the basis of their character states." The taxonomic units are the soil individuals in the form of soil profiles. The character states are the attributes which are presented in a numerical form. Therefore the phenetic similarity between individuals could be determined using a metric. Numerical taxonomy is a further development of Adansonian taxonomy.

Sneath and Sokal (1973, p. 11) claim that the principal aims of numerical taxonomy are repeatability and objectivity, which most soil taxonomies proposed earlier lack.

Prior to 1955 the use of numerical taxonomic methods was limited due to the fact that a large quantity of information could not be handled without the aid of a sufficiently powerful computer which was not available at the time. The selection of a set of attributes for classification and identification was done without appreciating the inter-attribute correlations and such methods are described by Arkley (1968) as suboptimal.

The traditional soil classifications have usually been hierarchical with a small number of differentia at each categorical level. This could distort the relative relationships between soil individuals, and groups can be constructed in such a way that soils similar in a few characteristics but dissimilar in all other characteristics could be grouped together.

The application of numerical methods to soil classification is to achieve objectivity and repeatability but it must be noted that the early use of numerical methods in the classification of soils encountered several problems, for example, the selection of a suitable similarity measure. A large number of similarity measures (similarity here refers to both similarity and dissimilarity measures) have been used to generate an inter-individual similarity matrix between soil individuals. At the early stages of numerical taxonomy product-moment correlation coefficient was a popular measure of the similarity. As has been pointed out by Sokal and Sneath (1973, p. 117), the choice of the similarity measure has been made without adequate theoretical justification. Moore and Russell (1967) demonstrated that five different similarity measures produced different classifications. There are some theoretical objections to the use of some similarity measures. Eades (1965) has objected to the use of product-moment correlation when the attributes are measured on different scales. The Euclidean distance metric has been used ignoring the correlation between attributes. This metric can be used only if the attribute vectors are mutually orthogonal. Kyuma and Kawaguchi (1975) used the Euclidean distance after orthogonalizing the attribute vectors by means of principal component analysis (PCA).

Several sorting strategies of agglomerative cluster analysis are available but the outcome of the strategy depends very much on the similarity measure. Also these strategies may be different from each other in terms of the clarity of the clusters in the dendrogram produced.

Both principal component analysis and principal coordinate analysis (PCO) have been (Rayner, 1966; Cuanalo and Webster, 1970; Webster and Burrough, 1972; *Campbell* et al 1970; Norris, 1971) applied to soil classification, but these methods have often failed to isolate clusters (Webster, 1976, 1979).

1.7 Aims and procedures of the present study

It has been found that the classifications obtained by numerical methods disagree not only with the traditional classifications but also among themselves (section 1.6). When the numerical taxonomic methods were first introduced to the soil taxonomy, no attempt was made to examine their properties or their behaviour in relation to soil data. This problem still retards the progress of numerical soil taxonomy. Therefore, it has become necessary to evaluate at least the widely available procedures in relation to soil classification.

It was shown earlier that there is no common agreement among the soil taxonomists on the nature of the natural soil classification. However, a growing number of soil taxonomists tends to agree that such a classification should be based on the measured properties of soils. But this decision alone would not produce a unique classification as the methods used dictate the nature of a classification. It is necessary to determine the best method of soil description. A given soil property may be measured at a series of depth-levels and there are several ways of presenting such information (Lance and Williams, 1967*a*). Moore, Russell and Ward (1972) compared three soil profile models and concluded that both original data, weighted by an exponential function and the coefficients of depth functions fitted to all the attributes used in the study produced similar classifications. The use of all measurements of a given attribute may not be necessary, but the effect of the elimination of some depth levels and the nature and effect of inter-attribute correlations should be examined.

Although the hierarchical agglomerative strategies tend to produce similar classifications when the same inter-individual similarity matrix is used, the degree of distortion introduced by the clustering strategies varies. The dendrograms produced by them are different in terms of the clarity of the clusters. For example, the single linkage sort method is known to suffer from the chaining of individuals rather than producing clusters, when there are intermediate types of individuals in the sample.

Although the ultimate objective of a taxonomic classification is to discover 'natural' groups (as defined earlier) it is difficult to assess such classifications in mathematical terms. The artificial classifications can be evaluated in relation to the utility of the classification defined, the taxonomic classifications cannot be evaluated so easily. But it may be possible to define a statistical criterion to compare several classifications for numerical optimality.

It is possible to reallocate individuals until a measure defined achieves its optimum value. But in practice this is not feasible as the limit of computer time could impose a constraint. A possible way round of this problem is to use other classificatory strategies and then find the most optimum classification as determined by a suitable criterion and finally perform reallocation by an appropriate strategy. In this process both similarity measures and the classificatory strategies have been compard.

References

Arkley, R. J. (1968) "Statistical methods in soil classification." In Ninth Inter-national Congress of Soil Science (ed) *Transactions*, Vol. IV, p. 187–192., Adelaide.

Avery, B. W. (1956) "A classification of British Soils", In Sixth International Congress of Soil Science (ed) *Transactions*, E. p. 279–285., Paris.

(1968) "General soil classification: Hierarchical and co-ordinate systems.", In Ninth International Congress of Soil Science (ed) *Transactions*, Vol. IV, p. 169–175. Adelaide.

(1973) "Soil classification in the Soil Survey of England and Wales", *J. Soil Sci.*, 24, p. 324–338.

(1980) *System of Soil Classification for England and Wales, Higher Categories*, Technical Monograph No. 14, Soil Survey of Great Britain, Harpenden, England., 67 pp.

Basinski, J. J. (1959) "The Russian approach to soil classification and its recent development", *J. Soil Sci.*, 10, p. 14–26.

Buol, S. W. *et al.* (1973) *Soil Genesis and Classification*, The Iowa State Univ. Press. 360 pp.

Cain, A. J. (1958) "Logic and memory in Linnaeus's system of taxonomy", In *Proc. Linn. Soc. Lond.*, 169th Session, p. 144–163.

Campbell, N. A. *et al.* (1970) "Numerical classification of soil profiles on the basis of field morphological properties", *Aust. J. Soil Res.*, 8, p. 43–58.

Crowson, R. A. (1970) *Classification and Biology*. Anderton, New York, 350 pp.

Crowther, E M. (1953) "*The sceptical soil chemist.*" J. Soil Sci. 4, p. 1–22.

Cruickshank, J. C. (1972) *Soil Geography*, David and Charles : Newton Abbott, 256 pp.

Cuanalo de la, P. H., and Webster, R. (1970) "A comparative study of numerical classification and ordination of soil profiles in a locality near Oxford, part 1, Analysis of 85 sites", *J Soil Sci.*, 21, p. 340–352.

Duchaufour, P. (1963) "Soil classification : A comparison of the American and the French systems", *J. Soil Sci.*, p. 149–155.

Eades, D. C. (1965) "The inappropriateness of the correlation coefficient as a measure of taxonomic resemblance". *Syst. Zool.* 14, p. 98–100.

Fitzpatrick, E. A (1967) "Soil nomenclature and classification". *Geoderma*, 1, p. 91–105.

(1980) *Soils, Their Formation Classification and Distribution*, Longman, London, 353 pp.

Gerasimov, I. P. *et al.* (1964) "Discussion of the new American soil classification system". *Soviet Soil Sci.* p. 572–603.

Gilmour, J. S. L. (1936) "Wither taxonomy?" A paper given by the author to Section K (Botany) of the British Association of Advancement of Science at its meeting Blackpool, *Classification Soc. Bull.*, Vol. 3, No. 4, 1976.

(1937) "Ataxonomic problem." *Nature*, London, 139, p. 1040–42.

(1961) "Taxonomy." In A. M. Macleod and L. S. Cobley (eds) *Contemporary Botanical Thought*, pp. 27–45, Oliver and Boyd, Edingburgh and Quadrangle Books, Chicago, 197 pp.

Glinka, K. D. (1928) *The Great Soil Groups of the World and their Development.* Edwards Bros., Ann Arbor, Mich., 255 pp.

Kesseba, A. *et al.* (1972) "Trends of soil classification in Tanzania— The experimental use of 7th Approximation. "*J. Soil Sci.* 23, p. 235–247.

Knox, E. G. (1965) "Soil individuals and soil classification". *Soil Sci. Soc. Am. Proc.* 29, p. 79–84.

Kubiena, W. L. (1953) *The Soils of Europe*, Murby, London, 317pp.

(1958) "The classification of soils". J. Soil Sci., 9, p. 9–19.

Kawaguchi, K. and Kyuma, K. (1977) *Paddy Soils of Tropical Asia. Their material nature and fertility.*, Honolulu, 258pp.

Lance, G. N. and Williams, W. T. (1967) "A note on the classification of multi-level data". *Comp. J.*, p. 381–383.

Leeper, G. W. (1956) "The soil classification." *J. Soil Sci.*, 7, p. 59–64.

Moore, A. W. and Russell, J. S. (1967) "Comparison of coefficient and grouping procedures in the numerical analysis of soil trace element data". *Geoderma*, 1, p. 139–158.

Moore, A. W., Russell, J. S. and Ward, W. T. (1972) "Numerical analysis of soils. A comparison of three soil profile models with field classification". *J. Soil Sci*, 23, 193–209

Muir, J. W. (1962) "The general principles of soil classification with reference to soils". *J. Soil Sci.*, 13, p. 22–30.

(19)

Muir, J. W. *et al.* (1970) "The classification of soil profiles by traditional and numerical methods". *Geoderma*, 4, 81–90.

Norris, J. M. (1971) "The application of multivariate methods to Soil studies"., *J. Soil Sci.*, 22, p. 69–80.

Ragg, J. M. and Clayden, B. (1973) *The Classification of Some British Soils According to the Comprehensive System of the United States.* Soil Survey Technical Monorgraph No. 3. Harpenden England.

Rayner, J. H. (1966) "Classification of soils by numerical methods". *J. Soil Sci.*, 17, 79–92.

Robinson, G. W. (1932) *Soils, their Origin, Constitution and Classification.*, Murby, London.

Sarkar, P. K. *et al* (1966) "Selection of characteristics for numerical classification soils"., *Soil Sci. Soc. Am. Proc.*, 30, p 269–272.

Simonson, R. W. (1952) "Lessons from the first half-century of soil survey, I classification of soils"., *Soil Sci.*, 74, p. 249–257.

Sneath, P. H. A. (1964) "Mathematics and classification from Adanson to present". In G. H. M. Lawrence (ed) *Adanson.*, The Bicentennial of Michael Adanson's "Familles des Plantes", part 2, p. 471–498. Hunt Botanical Library, Carnegie Inst. Technol., Pittsburgh, Penn. 635pp.

Sneath, P. H. A. and Sokal, R. R. (1973) *Numerical Taxonomy*, Freeman, San Francisco.

Soil Survey Staff (1951) *Soil Survey Manual*, U.S.D.A. Handbook, 18, U.S. Govt. Printing Office, Washington DC.

(1960) Soil classification, A Comprehensive System, 7th Approximation. U.S.D.A. Printing Office, Washington DC.

USDA (1975) *Soil Taxonomy*, A Basic System of Soil Classification for Making and Interpreting Soil Data. U S. Govt. Printing Office, Washington DC

Webster, R. (1968) "Fundamental objections to 7th Approximation". J Soil Sci., 19, p. 354–366.

(1976) "The nature of soil variation" , *Classification Soc Bull.*, 3, p. 43–55.

(1977)	*Numerical and Quantitative Methods in Soil Classification*, Oxford Univ. Press, 279 pp.
(1979)	"Exploratory uses of multivariate analysis in soil survey". In N. Wrigley (ed) *Statistical Application in the Spatial Sciences*. p. 286–306, Pion Ltd, London. 310 pp.
Webster, R. and Burrough, P. A. (1972)	"Computer based soil mapping of small areas from sample data, I Multivariate classification and ordination". J. Soil Sci. 23, p. 22–234.