Soil Classification Principles

ARNOLD Richard W.

National Cooperative Soil Survey, Washington, 9311 Coronado Terrace
Fairfax, VA 22031-3835, USA
Email: ct9311@aol.com

Pedology deals in decisions; probability devises many choices.

Purpose	Prioritisation
Domain	Diagnostics
Identity	Membership
Differentiation	Certainty

Abstract

The sources – the foundations – the causes – from which soil classification has proceeded have been stated as three principles for setup, four principles for organization, and one principle for the future. Their sequence provides a mode of operation for the design and evaluation of a classification system.

Setup.

1. Principle of **Purpose.** The reasons for wanting to organize soil knowledge.
2. Principle of **Domain.** The universe of objects relevant to the purpose.
3. Principle of **Identity**. The individual members of the domain are defined and named.

Organization

4. Principle of **Differentiation.** The protocol-guided hierarchical structure of a system with categories, and classes within categories.
5. Principle of **Prioritization.** The priority of knowledge by sequencing categories and sequencing classes within categories.
6. Principle of **Diagnostics**. The quantification and use of soil properties, sets of properties, and selected features (diagnostics) that provide objectivity.
7. Principle of **Membership.** Class membership for individuals based on quantified class limits and described central tendencies.

Future

8. Principle of **Certainty.** The recognition that change is inevitable and the driving force continual testing of a system.

Introduction

"You must tell someone what you know before you can show him what you see."
N.R. Hanson (1969,p 108), cognitive specialist

Soil classification is a window into knowledge about soils at a given moment in time. It is like a snapshot that reveals the current condition without clear glimpses of the past or hints of the future. Although science continues to advance at a rapid pace, a certain state-of-the-art is encapsulated in its classification. The more popular soil classification schemes of the past century have been morphogenetic ones built around the hypotheses and theories of soil development and pedological transformations of materials at or near the lithosphere's contact with the biosphere and atmosphere.

Before the creation of Pedology in the 19th century, most geologists believed soils were weathered or transported products of rocks. For agriculturalists soils supported and seemed to be well suited to certain crops and plants. For urban people and others whose contact with soils was minimal, soils were dirty and

those who tilled the land were almost non-existent social beings. Thus for many people soil was considered to be part of something else, or something to be left alone.

Soils were finally recognized as independent natural bodies worthy of scientific study. Pedology was originally conceived by Russian scientists as genetic soil science whereby soils were recognized as cause-and-effect results of processes that were influenced by natural environmental factors and conditions. This has been the most fundamental change in the concept of soil in history (Cline, 1961). This new paradigm spread throughout the scientific community gaining its own adherents in the process. Soil properties were described and genetic hypotheses proposed to explain the presence and spatial occurrence of soils. Over the years the concept of soil as the outer layer of the earth's crust, the pedosphere, has vacillated back and forth between soil as a continuum to soil as a collection of natural bodies capable of supporting plants (Kellogg, 1959). As predictions of soil behaviour assume increasing significance in environmental responses, soils are often treated as continua of data sets best understood by mathematical manipulations.

Technology has provided opportunities to make more precise measurements, causing a major shift from qualitative definitions to quantitative ones to take place (Simonson, 1962). Field studies have revealed the complexity and spatial intricacies of pedi-sediment formation and interruption in unconsolidated materials, and concepts of soil genesis have been modified. Such evidence has had a tremendous influence on research and models of soils as landscapes. Time has taken on fresh connotations as more details of paleopedology have been obtained. Sequential development and polygenetic cycles challenge previous theories of when and how soils form and develop. Even soil forming processes have been revisited, dissected, and combined in interesting ways.

The biggest change in soil classification has been the quantification of facts and definitions. Current systems of soil classification will continue to change as new facts, correlations, interpretations, concepts, and hypotheses are developed.

The class limits in newer classifications are direct consequences of theories and models, however the definitions are coldly factual and make departure of theory from fact readily apparent (Cline, 1961). Some of people's most cherished concepts have received harsh treatment at the hands of precise definition.

I find it difficult to improve on what has been written about the fundamentals of soil classification, consequently I have taken the liberty to restate some of those ideas (Cline, 1949) as eight priority principles that have influenced soil classification during the past five decades. Here I refer to a principle as a source, or cause, from which a thing proceeds. In addition to being a fundamental cause, it often is an established mode of operation.

SETUP: The First Three Principles

"A paradigm that is not expressed linguistically can be learned only with great inefficiency and at great expense".

B. Hudson (1992, p 838), pedologist

1. **Principle of Purpose.** Usually there is an overarching reason for wanting to organize knowledge about soils. A scientist, or group of specialists, may decide that they want to show genetic relationships in space and time. Others may want to predict soil behaviour when used and managed in different ways. Applied uses and scientific knowledge have both been major purposes of soil classification. The rationale of why a particular purpose is chosen supports this crucial first principle.
2. **Principle of Domain**. The realm of soils has different connotations, therefore it is necessary to specify what is to be included in soil classification and what is not. One can consider geographic bodies of soils, the pedosphere, or a broader spectrum of surficial materials, the geosphere, or abstract soil space that is based on concepts derived from small representative volumes. As techniques for examining relationships have evolved, it is now possible to consider a domain of subjective functional properties of soils and even include environmental parameters where relevant (Van Alphen and Stoorvogel, 2000). The choice of a domain depends upon the purpose that gives rise to a soil classification.
3. **Principle of Identity.** A domain indicates the population, or universe, that is being included in a soil classification scheme. It does not, however define the entities or members that will be the

source of data for the classification. Identity is a means of providing a name for a non-divisible component or individual that would not otherwise be recognizable. In this sense, such an entity if it is divided, is destroyed. There have been a number of objects employed as members of interest such as polypedon, pedon, profile, arbitrary body, soil landscape unit, and continuum segment (Arnold, 1983). For some interpretive classifications, functional soil units characterized by functional properties are suitable entities, for example, when considered as management areas for precision agriculture (Van Alphen and Stoorvogel, 2000). Arbitrary soil individuals can serve as reference bodies between different classification systems if their boundaries are independent of soil properties and class limits (Van Wambeke, 1966). Precise definitions of the desired individuals may be difficult and rather cumbersome, however, it is nearly impossible to proceed without resolving this principle of identity.

Equally important to identity is the nomenclature used for the classes within a system. One of the functions of a classification is to assist in remembering important properties of the objects. A systematic nomenclature that indicates location within the system and also provides mnemonic links to important properties is extremely useful. Links with prior names is often desired as a way of preserving history. The use of connotative elements has been successfully demonstrated in Soil Taxonomy.

At this point the reasons for wanting to classify have been agreed on, the population or domain of interest has been specified, and the individuals from which data will be obtained have been defined. The first three principles are also significant for any evaluation of a system because a system need only satisfy its internal specifications and definitions to be an adequate system. When disagreements about different systems occur it commonly is the result of applying external expectations and assumptions rather than viewing the system as a stand-alone integral whole.

A hierarchical system of classification is a separation of a domain into successively more specific classes. It divides the domain into smaller groups of members that are also mutually exclusive classes. Consequently the differentiae that separate classes (groups) focus on the limits or boundaries of the classes, and subordinate the central concept of each class. This is done for operational reasons, nevertheless, the essence of a class is its central concept and its description serves as the focus and image of the class.

An important feature of a hierarchy is that successively lower level classes accumulate all of the limits and descriptions that apply to the classes above it (Cline, 1949). A class at a lower categorical level cannot have limits that exceed those established by the higher-level classes of which it is a subdivision. This amounts to an accumulation of defining characteristics. In addition, it is appropriate to also make statements about those properties and features that are correlated with the defining characteristics, thus each class has a set of defining characteristics and a set of associated, or accessory, characteristics which permit many useful statements to be made about the classes in the lower categories.

Due to the many areas of uncertainty in the measurement of properties and the relevance of different combinations of sets of properties, in addition to the lack of precision of concepts and relationships, it is obvious that classification is not a truth that can be discovered. I am not aware of a way to determine the correct details of structure of a hierarchical system. The number of categories required to adequately represent a specific domain of millions of individuals is not readily known, and whether mutually exclusive classes and those that have overlapping boundaries are equally relevant and proper is still unclear.

ORGANIZATION: the Next Four Principles

4. **Principle of Differentiation**. If one is aware of a domain and a purpose for classifying its members, it is possible to start with the domain and divide it and subdivide it and so on. It is also possible to group the individuals, then group the groups, and so on. What sometimes if forgotten is that once organized into a hierarchy, it is only possible to employ the system from the top down. The "rules of engagement" have only been devised, or clearly stated, for the processes of separation.

   The first separation of a domain into classes creates the highest category. The definition of this particular set of classes must be quite abstract to capture the intent of the classification. It may be genetic, geographic, interpretive, or functional but it must apply to all members of the population

(a rule sometimes forgotten). The definition of the category suggests possible indicators that are consistent with, and satisfy, the definition. Such marks or evidence are soil properties assumed to be closely related to the category definition. The number of classes of the highest category is subjective, that is, it depends on the designer of the system.

The separation of each class at the next lower level is guided by the definition of this next lower category. It must be less abstract than the definition of the category above. For example, difference of kind among soils is more abstract than difference in degree. The classes being divided at this level are already bound by the definitions at the higher category, thus with each successive lower category the groups become smaller and more specific.

Two important aspects of a hierarchy are (a) that all members of the domain being considered are included in each categorical level, and (b) that differentiating criteria (derived from class limits) accumulate in each lower set of classes.

Defining categories is a difficult challenge. The set of category definitions may refer to morphogenesis of soil profiles, geographic associations of kinds of soils, or potentials for use but they must be derived from the purpose intended for the classification. The set of category definitions establishes a priority among the concepts and theories of the paradigm. It is easy to produce confusion when the definitions are violated.

Evaluating the adequacy of a system relies heavily on understanding the definitions of categories and their classes. This is where most "pet theories and sacred cows" are led to slaughter.

5. **Principle of Prioritization**. It is assumed, or clearly stated, that measurable soil properties and features are the data sources, and not the concepts or theories themselves. The main reason is objectivity. If the values and the methods of measurement are specified then other scientists can repeat the procedure and observe the same, or very similar, values. This removes part of the bias of the classification architects. The properties selected to be measured are obviously subject to the understanding of relationships between concepts and soil property data. These connections give substance to the framework and meaning to the pattern of order that is displayed.

   A category definition leads to differentiae among the classes, and soil properties are selected that are correlated with the differentiae. Another difficult task is defining sets of properties or features that support the meaning and intent of the classes being designed. Because many soil classifications are morphogenetically based, the relationships and concepts guiding the selection of properties are major cause-and-effect ones.

   The sequence of the classes in the highest category establishes the tone for the remainder of the framework, and also serves as the entry key to the system. The sequence of classes within each category is also a priority scheme that promotes consistency and facilitates using the system as an identification key.

6. **Principle of Diagnostics**. The definitions of soil properties and sets of properties must use values based on the methods of measurement. Obviously there are many choices, such as color by Munsell color charts or by spectrometers. And the readings may vary according to conditions outdoors or in a laboratory. General availability of methods and equipment often restrict widespread acceptance of the parameters selected. It is thought best to present some information as ratios or percents, others as weights or concentrations. Sets of properties may define horizon sets and be given specific names, such as a mollic epipedon, or a kandic horizon. Depths of occurrence of features may be diagnostic for some features. Quantification facilitates consistency and "diagnostics" represent a short hand way to aid recognition and placement of individuals. The principle of diagnostics refers to quantified soil properties and features and not to generalities or abstractions. The presence or absence of features may be diagnostic insofar as the features are defined and are relevant to the differentiae of the classes.

7. **Principle of Membership**. Regardless of how the individuals are defined, their membership into classes is characterized by two concepts. One is the central tendency that is like a mean, mode, or idealized abstract entity. The other is the boundary of the class with other classes. The limits of a

class include properties with adjacent classes in the same category and with classes in other categories that share a common property. Although to have mutually exclusive groups of members is desirable and theoretically possible, the uncertainties of measuring properties and the relevance of precise limits indicate that actual membership acceptance may be more probabilistic than deterministic.

In fuzzy c-means classification membership in a class can be expressed on a continuous scale from 0 to 1, thus partial class membership is usual. Membership values are interpolated with kriging techniques for functionally defined classes and they can be combined into a single confusion index (Van Alphen and Stoorvogel, 2000). Thresholds are commonly set for class boundary detection.

The principle of membership is fraught with doubt due to the usual 'yes or no' answers without any options for 'maybe', 'almost but not quite', or other near misses. There appear to be opportunities of innovativeness yet to be taken (Galbraith and Bryant, 1998).

The four principles of organization are differentiation, prioritisation, diagnostics, and membership. They permit relationships to be displayed and allow the concepts of order to be seen. As often has been said, the devil is in the details, and during the fleshing out of definitions, setting priorities, selecting properties, and determining placement there are numerous demons at work.

It is easy to lose sight of the objectives, to forget that the purpose is the driving force, and that this individual and not that one is being considered. As we all know, writing good definitions is really hard to do. Ambiguity seems to run rampant and at one point we want to say, "I just don't care anymore. It doesn't matter." The saving grace is that none of us are in this game alone. There are teams, and groups, and concerned organizations that do care and are willing to work together for the common good. And this brings us to the last principle, that of certainty.

ENSURING THE FUTURE

1. **Principle of Certainty.** New facts and new relationships are a given. They are a certainty; they will happen. Paradigms must constantly be tested for their validity and ability to accommodate additional information. Classifications need to be tested and evaluated for their capacity to be modified without self-destructing. This can be considered as system flexibility. Without a mechanism to ensure continual testing a system will stagnate, or even worse, it might start to be accepted as truth.

   Soil genesis operates with a 'factor-process-property' paradigm and soil survey operates with a 'soil-landscape' paradigm (Hudson, 1992). Soil correlation appears to operate with an 'arbitrary volume-property-matching' paradigm and several applied interpretations may even operate with a 'soil data-function-functional unit' type of paradigm. Paradigms and the classifications they spawn require inquiry and evaluation by dedicated soil scientists for there to be meaningful tomorrows in soil classification.

"A paradigm that is not expressed linguistically can be learned only with great inefficiency and at great expense"

B .D. Hudson (1992, p 838), pedologist

Postscript

I hope you like one of these phrases – because each is a mnemonic aid for these eight principles of soil classification.

" Pedology discovers important details; pedologists discuss major changes."
" Pedons described in detail; profile discontinuities marked clearly"
" Pedology deals in decisions; probability devises many choices"

References

Arnold, R. W. 1983. Concepts of soils and pedology. p. 1-21, In L.P. Wilding, N. E. Smeck, and G. F. Hall (eds) Pedogensis and Soil Taxonomy. I. Concepts and interactions. Elsevier Sci. Publ., Amsterdam.

Cline, M. G. 1949. Principles of soil classification. Soil Science 67: 81-91.

Cline, M. G. 1961. The changing model of soil. Soil Sci. Soc. Amer. Proc. 25 (6): 442-446.

Galbraith, J. M., and R. B. Bryant. 1998. A functional analysis of Soil Taxonomy in relation to expert system techniques. Soil Science 163 (9): 739-747.

Hanson, N. R. 1969. Perception and discovery –an introduction to scientific inquiry. (posthumous edition edited by W. C. Humphreys). Freeman, Cooper and Company, San Francisco. pp. 108

Hudson, B. D. 1992. The soil survey as paradigm-based science. Soil Sci. Soc. Am. J. 56: 836-841.

Kellogg, C. E. 1959. Soil classification and correlation in the soil survey. USDA-SCS. Washington. 1-17.

Simonson, R. W. 1962. Soil classification in the United States. Science 137: 1027-1034.

Van Alphen, B. J. and J. J. Stoorvogel. 2000. A functional approach to soil characterization in support of precision agriculture. Soil Sci. Soc. Am. J. 64: 1706-1713.

Van Wambeke, A. 1966. Soil bodies and soil classification. Soils and Fertilizers 29: 507-510.