A SYSTEM OF RESEARCH INFORMATION RETRIEVAL BY PUNCHED FEATURE CARDS

1969

John Makin

National Agricultural Laboratory, Kenya

Introduction.

Over the years a considerable amount of agronomic research has been carried out in Kenya. It is however often difficult to track down experimental data on individual research stations , and indeed there has been more than one case of duplicated experimentation.

It is suggested that some system of recording the continuing inflow of data should be devised: a simple system is here proposed for retrieving and correlating this information in a rational manner. The ultimate objective is to answer the type of question: "Is there any information concerning the cultivation of cotton on sandy soils in the Eastern Province?; and if so where may this information be located?."

The proposed arrangement is for a simple Punch Card System relating to a Master Ledger which acts as a guide to the information. Tentative examples are then provided to illustrate the system in operation in the field of Soil Research.

Feature Cards

There are really two basic forms of punched card-- the Item Card, with every single 'unit of information' recorded on separate cards; and the Feature Card (Jolley, 1959), wherein each card merely represents a particular property or characteristic of the information which is recorded in its entirety elsewhere. Since a feature card system employs only a fraction of the cards that would be required for recording every item, it is the one selected for research purposes. It is, moreover, a simple, speedy and inexpensive system for retrieving data, requiring no mechanical aid other than a fine punch.

In the proposed system each feature card would have facilities for, say, up to 10,000 punched holes (these may be conveniently punched into a card measuring 10 by 12 inches). A typical card is covered by a lattice of small squares, the position of each being described by means of a numerical grid code (i. e. from 1 to 99) arranged along each edge of the card, just as grids enumerated on map margins enable one to indicate a geographical position through a Grid Reference. These feature card codes relate in turn to entries in a master ledger: these indicate where the information is actually located (e. g. a particular file). In cases where only brief research results were obtained, a summary of the experimental result might be included in the master ledger.

The method of operation is best described by means of a simple example:--We have a piece of information regarding the behaviour of cotton on sandy soils. This involves 2 feature cards: a crop feature card for cotton, and one for soils having sandy textures. An unallocated position on the 'Sandy Soil' feature card is punched to indicate the existence of this information. Let us suppose that the "grid reference" of this position is 96/25. This entry might well indicate the year the information was obtained, followed by reference either to the file containing the original details or to a compiled summary of research information in a particular field (e.g. a compendium on cotton research). Thus, in the above example, the master ledger might read:

	96/25	Report No. 18 p. 32 1956
or	96/25	File XYZ Soil Chemistry Section p. 8 1956

When it comes to retrieving all information regarding cotton on sandy soils, the two relevant feature cards are placed one upon the other when all through-going holes serve to indicate the existence of the required information. Armed with the grid references of these coincidental holes, the enquirer then looks up the relevant code in the master ledger. One of these entries would carry the code 96/25.

Thus in the example given, the feature card system used 2 cards whereas an item card system would have employed as many cards as there were units of information. The stacking of feature cards to determine the through-going holes is therefore a rapid means of assessing, from many items, those that will fulfill a given set of requirements. Jolley (1959) has also pointed out that there may be an advantage in punching the holes on the cards in a special sequence (e.g. to represent the chronological order of experimentation) so that the positions of the holes themselves have some significance.

This relatively simple system is suitable for the individual research station, or for a group of research units based in a similar ecological zone. Provided that some central coordination were effected in the initial stages, it would at a later date be relatively easy to computerise the entire set-up for the retrieval of research information on a national basis.

The remainder of this paper is devoted to reflections on the devising of meaningful criteria for use as features within such a system as applied to soil research.

General Features

It is important that the system should accommodate feature cards to enable soil properties to be related to the general environment. Some obvious examples spring to mind:

(a). Location Feature Cards

It is clearly valuable for the soil to be related to its geographic situation. Existing

District boundaries would make suitable location sub-divisions and, where feasible, each District would be allocated a separate feature card.

(b). Climatic Zone

From the point of view both of Soil Development and of Soil/Crop Relations, a suitable and relatively simple system may be devised to take into account the most significant properties of climate viz. rainfall probability and distribution, total rainfall and temperature. Each climatic zone to be awarded a feature card would be formally defined on the basis of rainfall and the appropriate altitudinal range.

(c). Crops

It would be ideal for each crop to be allocated a feature card. Some crops, of lesser importance but similar function (e.g. essential oil crops) might be grouped together, though this could jeopardise the effectiveness of the system.

(d). Type of Experiment

There are numerous kinds of agronomic trials that might require feature cards. Below is a tentative list of useful features for experiments concerned with Soil Improvement:

i. Fertilizers and Manures

ii. Trace Elements

iii. Liming Trials

iv. Amendment by Gypsum and/or Leaching

v. Irrigation

vi. Drainage

vii. Nematode Control

Soil Features

The Kenya Soil Survey has emphasised the concept of the 'Soil Type' in the belief that the morphometric character of a soil is fundamental in defining the entity. It is recognised, however, that the soil type merely represents the modal concept: many individual profiles within the type concept may in fact be somewhat atypical. This is because soil exists as a universal continuum with each soil series grading into the next as external circumstances (e.g. drainage) change. What one attempts to achieve is a definition of the type species at a characteristic point on the continuum. In this way individual soils in both Kenya and Uganda have been related to specific positions on recurring landscape patterns.

The sole attempt at producing a broad soil classification for Kenya was originally published (1959) as a legend to a soil map of Kenya at a scale of 1:3,000,000. In this legend, Scott and Gethin Jones defined 38 major soil groups related by means of catenary associations; these soils were summarised by Bellis (1964). There may now be a case for re-classifying Kenya soils employing a more generally acceptable and comprehensive international system (e.g. U. S. Dept. Agric. 1960, 1964).

However, a debate concerning the desirable form of classification serves to obscure the objective of defining suitable soil parameters for use as features. Now there is a fundamental objection against all attempts to categorise soils, in that Classification per se, whether it be founded upon a morphometric or a genetic bias, leads inevitably by its very nature to the suppression of valuable soil parameters. Sneath (1957) maintains that a classification based on as many features as possible is the most "natural", and that such a system is appropriate for general taxonomy, especially when grouping objects in which specific characters tend to be so fluid that several relationships are feasible. A further argument against the use of soil individuals as features is that soil classification in East Africa has followed several differing systems, and in some cases no system at all! Indeed, on certain of the recent ad hoc development projects in Kenya, soil types and even soil series have been grouped together into 'Land Management Classes' (e.g. on the basis of suitability for irrigation).

One must conclude that, compared with Soil Individuals (series), Soil Properties embody a sounder basis for the present purpose. Yet Soil Properties themselves exist as continua, and to reduce them to a manageable form requires the division of information into categories which may then be allocated a feature card. This grouping into preconceived divisions of data which would otherwise vary continuously, provides a restriction on the value of the information ; yet it is an essential prerequisite for efficient retrieval.

Below are listed what are considered to be the most important soil properties in Kenya:

(i). Morphological. Drainage, Texture, Structure, Soil Depth, Distribution of Stones and Gravel, Horizon Rooting Densities.

(ii). Analytical. Bulk Density, Permeability, Available Moisture, Clay Mineral Composition, Organic Content, Acidity, Alkalinity, Cation Exchange Capacity, Content of Electrolyte.

These may be too numerous to allocate amongst individual feature cards. Moreover, some properties are mutually related (e. g. Texture with Available Moisture).

It is tentatively proposed that a simple grouping of soils could be made on the basis of Drainage and Texture.

Drainage features could include:

1. Well Drained
2. Seasonally Poorly Drained
3. Very Poorly Drained

The textural features might well be based on a modified version of the system described by the U. S. D. A. (1964).

Features	U. S. Soil Survey Manual (1951) Approximate Equivalents
Sandy	Sand and Loamy Sand
Silty	Silt, Silt Loam, Silty Clay Loam
Loamy	Loam, Sandy Loam, Sandy Clay Loam
Coarse Clay	Clay Loam, Sandy Clay, Silty Clay, Light Clay
Fine Clay	More than 60% Clay

Here one comes up against the snag that soil textures (and indeed most soil properties) vary according to the depth at which they are assessed. There appear to be three possible approaches to this problem:

(a). Average the values obtained for a particular property throughout the entire profile. This could only be valid for rather uniform profiles.

(b). Single out the horizon that demonstrates attributes most likely to be limiting from the point of view of plant growth (e.g. the horizon with the slowest permeability in the solum).

(c). Select two horizons as representative of 'Topsoil' and 'Subsoil' . Topsoil could be taken as the surface horizon, provided this exceeds 3 inches in depth: Subsoil as that horizon most nearly approximating to a hypothetical layer between 24 and 36 inches in depth. This is the course favoured by the Soil Survey of Kenya, though it implies separate parameters for topsoil and subsoil, thereby increasing the required number of feature cards.

There are other useful parameters besides texture and drainage and some examples are presented below. These have been defined with a view to their being meaningful under Kenya conditions.

Soil Depth

1. 0 - 1 metre
2. More than 1 metre

Subsoil Permeability

1.	Very Slow	Less than 25 m.m./ 24 hrs.	
2.	Slow	25-100	"
3.	Moderate	100 - 1000	"
4.	Rapid	1000 - 5000	"
5.	Excessive	More than 5000	"

Clay Mineral Composition

1. More than 50% 2:1 Expanding Clays
2. More than 75% Kaolinite or Halloysite
3. More than 75% Gibbsite
4. Amorphous materials predominating
5. Mixed Clays (normally including Illites)

Analytical Data

Three examples of possible parameters are given below:

1. Permanent Charge Exchange Acidity exceeds 2 m. e. %; pH less than 5.5
2. Alkalinity: More than 12% Exchangeable Sodium; or more than 35% Exchangeable Magnesium
3. Electrical Conductivity in excess of 6 m. mhos/ cm.

REFERENCES

BELLIS, E. 1964 "Soil Surveys in Kenya" African Soils IX. No. 1 p. 137

JOLLEY, J. L. 1959 "Punched Feature Cards" Data Processing, April-June 1959 p. 4

SNEATH, P. H. A. 1957 "The Application of Computers to Taxonomy" J. gen. Microbiol. 17 p. 201

U. S. SOIL SURVEY 1951 "Soil Survey Manual" Agric. Handbook No. 18 U. S. Dept. Agric.

U. S. SOIL SURVEY 1960 "Soil Classification, A Comprehensive System; Seventh Approximation" U. S. Dept. Agric.

U. S. SOIL SURVEY 1964 'Supplement to above U. S. Dept. Agric.