The Romanian PROFISOL Database

Introduction

The use of databases is becoming more and more frequent in soil science, as in many other scientific fields. Databases referring to soil profile description and analysis are examples of this trend (Van Waveren and Bos, 1988, 1989; Kimble *et al*., 1990).

An earlier soil profile database was developed in Romania in the 1970s (Mielcescu *et al.,* 1977; Canarache *et al*., 1981). A Felix-256 computer with FORTRAN and COBOL software was used. The database referred only to soil physical properties.

A. Canarache
V. Vlad
I. Munteanu
N. Florea
Anisoara Rasnoveanu
Daniela Popa

Research Institute for Soil Science and Agrochemistry
Bd. Marasti 61,
RO-71331 Bucharest 32,
ROMANIA

Data for pre-determined soil layers 20, 50, 100 and 150 cm deep on 1,650 soil profiles were stored. Processing consisted of calculating average soil properties for selected groups of profiles, and these averages were used in a preliminary synthesis on physical properties of soils in Romania.

Based on this early version, a new database, PROFISOL, was later developed. Details, especially on the computer-science procedures, are presented elsewhere (Vlad *et al.,* 1996). A general description of the content and facilities offered by PROFISOL, as well as on possible future developments, is presented in this paper.

Structure and functions of the PROFISOL database

Hardware and software

Two versions of the PROFISOL database are now in use:

- A minicomputer database was developed at an earlier stage. It made use of a DEC PDP-11 compatible minicomputer (CORAL-5030/4021) running FORTRAN (F77) and COBOL 81 under RSX-11M system software. At this stage, data for 120 profiles with complete data, 350 profiles with only physical and chemical data, and 3,400 profiles with only physical data were stored.
- A PC database was later developed using an IBM compatible 486 and Paradox 3.4 running under MS-DOS. Data from the minicomputer database were transferred here, and additional data were stored, bringing the total to 170 profiles with complete data, 450 profiles with physical and chemical data and 4,200 profiles with physical data only.

Input

Representative soil profiles described and analysed during current soil survey work or in any other similar activities are the main input in the PROFISOL database. Each profile is identified by a numeric chronological number. Most of the input data refer to genetic soil horizons, a maximum of 10 horizons being possible for each profile. All input is provided according to current Romanian soil survey methodology (ICPA, 1987).

Three forms for the input data have been developed:

1. M form, 4 sheets, for location and site description (96 items), profile description (24 items) and profile morphology (13 items for each horizon);
2. F form, 2 sheets, for soil physical properties (40 items for each horizon);
3. C form, 2 sheets, for soil chemical properties (34 items for each horizon) and groundwater properties (13 items).

In addition, information on analytical methods used (74

items), methods of drying the soil samples (30 items) and origin of soil physical data (1 item for each horizon) are included on the input forms. Altogether as many as 2167 data items can be stored for a single soil profile. The possibilities of including in the database at a later stage other kinds of soil data, such as microbiology, trace elements or clay mineralogy, are considered.

The existing forms are well suited to soil profiles where all or most properties are available. As in many cases a somewhat less complete description of the soil profile is used, the preparation of an alternative, less detailed, series of input forms is foreseen.

The M form contains mainly coded data. It was designed in such a way as to enable direct filling in during field work. In most cases the "menu" technique was used. The F and C forms contain only numerical data. They repeat some of the basic site and profile data, and the F form also repeats some of the basic chemical data. This way it is possible, when required, to use only one of these forms for several profiles.

Data are introduced to the computer interactively, using a selection of "video-maquettes", screen images designed to be very similar in format to the input forms. Each form, and various data within the forms, may be entered at any time. Several validation procedures make checking of selected input data possible. Video-maquettes may be displayed on the computer screen for examination and/or validation. All input data are stored. Corrections, changes and updating of data are possible after storage.

The software was developed in a modular way, with independent modules using communication between modules through common, permanent or temporary files. Linkage of software and modules was obtained through a system of hierarchical menus. This resulted in several advantages:

- management of the high complexity of the software system
- flexibility;
- a relatively easy extension of the software system;
- increase in work performance of teams;
- decrease in size and complexity of the individual software components and relative ease of their modification;
- protection in inter-conditional processing;
- better adjustment to the need to perform various operations at different time.

Due to the large number and varied character of the data included in the database, a data dictionary had to be prepared for each entry. It includes names, codes, measuring units, implicit character (when needed), possible range of values (in some cases), formulae used for calculation or estimation. More than 600 different data types are included in the dictionary.

Processing possibilities

The existing software makes a series of calculations which provides more complete information on the stored soil profiles. These are:

1. *The identification and/or calculation of new properties not included in the input.* Some of these new properties refer to global site or morphological indicators such as physio-geographic region, texture of specific layers, degree of gleying, degree of salinisation, etc. They are obtained using specific algorithms. Other properties are soil physical and chemical properties resulting from classical formulae, such as total porosity, various pore size classes, available water capacity and C/N ratio.
2. *Calculation of certain data missing from the input, but which may be obtained from classical formulae.* Examples are some particle size classes, contraction index, etc.
3. *Estimation of missing data using pedotransfer functions.* Such estimation procedures are now available in Romania for:

- transformation of Atterberg silt (0.002-0.02 mm) and fine sand (0.02-0.2 mm) into USDA silt (0.002-0.05 mm) and fine sand (0.05-0.2 mm) or vice-versa;
- the parameters of Van Genuchten closed-form water retention capacity (Simota, 1993);
- several soil moisture constants such as wilting point or field capacity (Canarache, 1986);
- saturated hydraulic conductivity (Canarache, 1987) or for resistance to penetration (Canarache, 1990).

They are all continuous pedotransfer functions,. The pedotransfer functions use as input mainly clay content and bulk density, and most of them are valid only for mineral soils. For the soil hydraulic properties, the existing pedotransfer functions are prepared as graphs from which spline functions have

been obtained and included in the database. Work is now being done to improve these estimation procedures, to use for estimation a larger variety of input data, to develop multiple regression equations instead of the spline functions, to extend pedotransfer functions for organic soils too, and to develop also class pedotransfer functions estimating texture and bulk density from information on physio-geographic region, soil map legend and soil profile morphology.

4. *Transformation of measuring units.* This type of calculation is being done for soil moisture constants to enable presentation of results as either weight (mass) percentage, volume percentage, or water depth (mm/ha).
5. *Recalculation of data for layers of pre-determined depth of the soil profile.* For many users of soil properties, such as designers of irrigation and drainage systems, there is a need for data to be referred not to soil horizons, but to specified standard soil layers. The PROFISOL software enables calculation of such data for most soil physical properties. The resulting data represent averages weighted according to the thickness and bulk density of the horizons.
6. *Selection of groups of soil profiles*. At present, selection may be done using a listing of soil profiles given by the user. It is foreseen that such selection will also be done automatically, using such criteria as soil classification, soil texture, physiographic region, depth of groundwater, etc.
7. *Statistical treatment of data for groups of soil profiles.* Once profiles of a specific soil group are selected as described above, data for each horizon or layer may be processed according to classical statistical procedures. Means, standard errors, medians, modes, standard deviations, sample variances, kurtosis, skewness, ranges, minimum and maximum values, and confidence levels are calculated. Various regressions may also be calculated. Possibilities exist to use numerical taxonomy techniques. For the next step, adding data treatment for kriging, linkage of the database to the geographical information system and other procedures are envisaged.

As described above, most of the processing possibilities refer to soil physical properties. It is also planned to extend some of these procedures to chemical properties. Automatic diagnosis of soil classification and of soil/ land classes is also considered.

Various reports are provided. They include:

- reports with input data, one each for each of the three input data forms;
- report with soil profile morphology description, including also the formula of the land mapping unit;
- reports with physical data for soil horizons, different reports referring to a more or less complete selection of properties (Table 1);
- reports with physical data for soil layers of pre-determined depth, different reports being provided for different methods of subdividing the profile into layers;
- a report with both soil physical and chemical properties;
- a listing of profiles included in a specific group of profiles, including some soil properties of each profile;
- reports with statistical parameters of the main physical properties for a group of soil profiles, with different reports for various selections of statistical parameters.

Existing applications

The main use of the PROFISOL database since it was started in the 1970s has been to provide data on soil hydraulic properties needed in design of irrigation and drainage projects. Most of the processing facilities offered – listings of soil profiles, or of average data for groups of profiles were currently used.

The database has also been used to improve existing pedotransfer functions and to develop new ones. Various regression calculation were used to this end.

A synthesis of most data included in the database was performed to characterise the eco-regions into which Romania has recently been subdivided. Soil profiles included in the database were grouped by ecoregions and according to soil classification, and the average of each soil property was calculated.

More recently, a similar approach was used to fulfil requirements of the EU Soil Profile Analytical Database. About a third of the total number of profiles existing in the PROFISOL database was selected for processing, classified according to the FAO/UNESCO World Soil Map Legend and grouped by major soil groupings. Subgroups had to be established to take into account the

significant differences between soil profiles within each major grouping. This subgrouping referred to soil texture in the case of Fluvisols and Gleysols, to the depth of bedrock and of groundwater when present, in some cases to soil units with specific soil horizons, and to land use. The number of soil profiles within each soil group, and the number of soil subgroups, are shown in Table 2. Subsequently, averages and statistical parameters of each soil property and each subgroup were calculated using the facilities offered by the PROFISOL database.

Table 1: Physical and chemical properties of soil profile 006334

Location: Segarcea (Dolj) typical argilluvial chernozem

Horizon	MEAS.	Ap	Am	A/B	Bt1	Bt2	Cc
Horiz.depth	UNIT	0-18	18-41	42-73	73-105	103-140	140-160
Sampl.depth		0-18	20-30	50-60	80-90	110-120	145-155
Phys.prop.							
Coarse frag.	%W/W	0	0	0	0	0	0
Coarse sand	%W/W	0.3	0.6	0.6	0.5	0.6	0.5
Fine sand	%W/W	32.4	29.1	26.2	27.5	30.0	31.3
Silt	&W/W	26.9	22.9	22.8	24.7	23.6	23.2
Clay	%W/W	40.4	47.4	50.5	47.3	44.8	45.0
Phys.clay	&W/W	54.6	62.1	66.6	62.2	58.8	60.0
Bulk density	G/CM3	1.28	1.33	1.40	1.51	1.43	1.40
Total poros.	% V/V	51.6	49.8	47.3	43.6	46.5	48.1
Air porosity	% V/V	15.9	14.4	12.2	8.8	14.5	16.5
Comp.degree	% V/V	0.0	5.5	11.1	17.2	11.0	8.1
Hygr.coeff.	% W/W	9.0	11.1	12.6	11.5	11.5	11.1
Wilting point	% W/W	13.5	16.6	18.6	17.2	17.5	16.6
Field capacity	% W/W	27.9	26.6	25.1	23.0	22.7	22.6
Tot.wat. Cap.	% W/W	40.3	37.4	33.8	28.8	32.5	34.4
Av.moist.cap.	% W/W	14.4	9.9	6.2	5.7	5.1	5.9
Water yield	% W/W	12.4	10.8	8.7	5.8	10.7	11.8
Hydr.conduc	MM/H	1.9	0.8	0.4	0.2	0.5	0.5
Chem.prop.							
Humus	% W/W	4.0	2.5	1.9	1.2	0.7	0.8
Total n	% W/W	0.180	1.120	0.089	0.055	0.033	0.037
C/N	-	15.0	1.5	14.4	14.6	14.1	14.6
CaCO3	% W/W	0.0	0.0	0.0	0.0	0.0	8.2
pH in H2O	-	6.4	6.3	6.3	6.5	6.9	8.1
Exch.bases	ME/100 G	30.4	30.2	31.5	30.3	nd	27.6
Exch.Ca	ME/100 G	24.8	24.1	24.8	24.0	nd	22.7
Exch.Mg	ME/100 G	4.7	5.4	5.9	5.4	nd	4.0
Exch.K	ME/100 G	0.6	0.5	0.6	0.6	nd	0.5
Exch.Na	ME/100 G	0.3	0.2	0.3	0.3	nd	0.4
Exch.H	ME/100 G	6.8	6.7	5.3	4.7	nd	nd

Table 2: Number of profiles with stored soil physical properties; selected from the PROFISOL database and classified according to the FAO/UNESCO World Soil Map Legend

Major soil grouping	Number of profiles	Number of subgroups
Fluvisols	692	72
Gleysols	50	7
Regosols	16	2
Leptosols	20	3
Arenosols	72	5
Andosols	19	2
Vertisols	33	4
Cambisols	69	17
Solonetz	19	4
Solonchaks	7	2
Kastanozems	11	1
Chernozems	218	9
Phaeozems	277	15
Greyzems	16	2
Luvisols	166	19
Podzols	19	3
Histosols	2	1
Total	**1602**	**8 16**

Conclusions

1. The PROFISOL data-base is a complex tool, including site and soil morphological, physical and chemical characteristics, pedotransfer functions and various processing facilities in an integrated form.

2. At present PROFISOL includes a large number of soil profile characteristics for most agricultural and non-agricultural areas of Romania, especially physical soil properties.

3. The complex software that was developed for PROFISOL achieved a good balance between software modularity, data structure and efficiency. It proved to be well adjusted to processing large amounts of complex and often asynchronous data.

4. PROFISOL has been used in various research projects, significantly contributing to a better knowledge and utilisation of Romanian soils and environment

5. Future development of the PROFISOL database is envisaged which will include continuous storage of existing and new data; preparation of a simplified version for current use; complete transfer to the PC platform; addition of new and improved pedotransfer functions; new processing procedures; extension to other kinds of data such as trace elements, clay mineralogy, micro-biology; introduction of compatibility with various simulation models, expert systems, decision support systems and with the Romanian soil geographic information system.

References

CANARACHE, A. (1986). Improved procedure for estimation of field capacity (in Romanian). *Stiinta Solului*, No. 1, pp. 43-48.

CANARACHE, A. (1987). Improved procedure for estimation of permeability (in Romanian). *Stiinta Solului*, No. 1, pp. 14-23.

CANARACHE, A. (1990). PENETR, a generalized semi-empirical model estimating soil resistance to penetration. *Soil & Tillage Research,* 16: 51-70.

CANARACHE, A., POPESCU, M. and MIELCESCU, A. (1981). Physical characterization of soil using computer storage and processing of data. In: Traitement Informatique des Données de Sol (ed. M. C. Giraud), 3-ème Colloque AISS, Paris, 14-17 Sept. 1981. Institut National Agronomique, Paris-Grignon, Departement des Sols, No. 4, Tome 1, pp. 39-48.

ICPA (1987). Soil survey methodology (in Romanian). Research Institute for Soil Science and Agrochemistry. Method, Reports, Guidebooks, No.20, 3 vol., 766 pp.

KIMBLE J.M., ESWARAN H. and COOK, T., (1990). Organic carbon on a volume basis in tropical and temperate soils. Transactions of the 14th International Congress of Soil Science, Kyoto, Japan, August 1990, 5: 248-25.

MADSEN, H. Breuning. and JONES, R. J. A. (1995). Soil profile analytical database for European Union. *Danish Journal of Geography,* 95: 49-58.

NORR, A. (ed.) (1989). Computerized land evaluation databases. Catalogue of a questionnaire survey. 2nd. edition. CEC, Agricultural Series, Luxembourg, 520 pp.

SIMOTA, C. (1993). Note on semi-empirical models for the estimation of pedotransfer functions used in simulation models for the soil water dynamics and crop yield formation (in Romanian). Technical report, Research Institute of Soil Science and Agrochemistry, Bucharest, 13 pp.

VAN WAVEREN, E. J. and BOS, A. B. (1988). ISRIC Soil Information System. Technical Papers, No.14-15, ISRIC, 48+68 pp.

VLAD, V., TARHOACA, E., POPA, D., ALBU, V., IANCU R., BALUTA, M., TAPALAGA, M., CANARACHE, A., MUNTEANU, I., FLOREA, N., RASNOVEANU, A., VLAD, L. and NACHE, M. (1996). Software system for the PROFISOL database of soil profiles (in Romanian). In: Lucrarile Simposionului National de Informatica in Cercetarea Agricola, Bucharest (in print).