Proposal for a National Database of Soil Hydraulic Functions in Hungary

[1] A. NEMES, [2] J. H. M. WÖSTEN and [3] A. LILLY

[1] Research Institute for Soil Science and Agricultural Chemistry (RISSAC) of the Hungarian Academy of Sciences, Budapest; [2] DLO Winand Staring Centre for Integrated Land Soil and Water Research, (SC-DLO), Wageningen (The Netherlands), and [3] Macaulay Land Use Research Institute, Craigiebuckler (Scotland)

Introduction

Agricultural and industrial activities are increasingly causing the quality of our soils and waters to deteriorate. The use of fertilizers, pesticides, inorganic and organic chemicals has already resulted in considerable environmental damage in many parts of Europe (EIJSACKERS & HAMERS, 1993).

In order to partly control and ultimately rectify this damage, scientists have developed increasingly complex computer models that simulate water and solute movement in the unsaturated zone of the earth's crust. These models have now become indispensable in research aiming to quantify and integrate the most important physical, chemical and biological processes active in the unsaturated zone of agricultural soils (ADDISCOTT & WAGENET, 1985). Simultaneously, models ranging from very simple to highly complex, are being used to evaluate the effects of alternative management practices on crop yield and groundwater quality (TENG & PENNING DE VRIES, 1992).

The use of models for research and management has shown that many input data have to be quantified in order to be able to make reliable predictions on regional, national or supra-national scale. At the same time these data are usually fragmented, of different degree of detail, of varying reliability and are held in different databases or in different institutes. In Hungary a comprehensive database was developed for the main physical and hydrophysical characteristics (SP, particle-size distribution, bulk density, pF curves, saturated and unsaturated hydraulic conductivity) during the last decade (RAJKAI, 1984; RAJKAI & VÁRALLYAY, 1989; RAJKAI et al., 1981; VÁRALLYAY et al., 1979, 1980).

Nowadays, in various fields of model application (like hydrology, environmental risk analysis or global change assessment), the lack of accurate parameters (particularly on soil hydraulics) is considered to be a major obstacle to progress. The unsaturated soil hydraulic parameters (water retention and hydraulic conductivity curves) are key parameters in this respect (VAN GENUCHTEN & LEIJ, 1992). The problem is also aggravated by the awareness of the significance of effects of temporal and spatial variability in the hydraulic parameters on model results which means that many more samples are needed than previously thought, to properly characterize a given field (WARRICK & MYERS, 1987).

Because most models greatly depend on accurate information on these parameters much work has been done in the measurement area. KLUTE & DIRKSEN (1986) and GREEN et al. (1986) reviewed the available laboratory and field measurement techniques, respectively. Despite these efforts, the unfortunate fact remains that the hydraulic parameters are still always notoriously difficult to measure, especially for undisturbed field soils. A possible alternative for these costly and difficult direct measurements is estimation of hydraulic parameters with pedotransfer functions (BOUMA, 1989). Pedotransfer functions relate hydraulic parameters with more easily measured soil data, such as soil texture, organic matter content and/or other data routinely measured in soil surveys. However, before the transfer functions can be derived it is necessary to create a comprehensive database of soil hydrological and pedological data. Moreover, since environmental changes are not restricted by national boundaries, there is a general consensus that these issues should be studied on an European scale and also on a worldwide scale. Therefore, databases designed to provide the necessary information to address these issues must include data from a wide range of soils and environmental conditions. As a consequence, the EU funded a project aimed at bringing together the available soil hydraulic data held by different institutes in the European Union, into one central database. Thus the HYPRES (HYdraulic PRoperties of European Soils) database was developed. This database holds measured and derived soil hydraulic parameters as well as related pedological and environmental data.

It was recognized that similar soil data has also been collected in Hungary as well as in many other countries under the framework of various different research projects and sometimes by different institutes. In Hungary these hydraulic data have not been gathered and stored in a national database (VÁRALLYAY et al., 1979, 1980). Collecting all information into a central, national database with the agreement of the possible contributing parties, would already widen possibilities in this respect. Attempting to do so, it is of a major concern to create a database which is not only compatible with existing soils databases in Hungary, but that it can also be linked to ongoing and future European and worldwide research projects. This compatibility enhances the potential for world-wide future co-operation.

Using some of the advantages of the recently established HYPRES database and following the structure and classification systems used within that EU database, it is intended to establish a national soil hydraulic database for Hungary. This database will be compatible to the EU database allowing the sharing of data at a later stage. This paper describes the structure and contents of the HYPRES database and makes suggestions for the development, potential structure and some of the possible uses of an updated national database for Hungary.

The Hydraulic Properties of European Soils (HYPRES) Database

In this section a description of the contents and structure of the HYPRES database is given. This database has been developed and populated with over 5500 soil samples, and has undergone a process of data standardization prior to the successful derivation of pedotransfer functions to predict the hydraulic properties of many European soils. The flexibility of HYPRES meant that it could remain compatible with UNSODA (LEIJ et al., 1994) an existing world-wide database, while introducing key features that allowed the data to be linked with existing European databases. The process of the development of HYPRES, its structure and data standardization provides a useful template for the development of a database for the storage and manipulation of Hungarian soil hydrological data.

Database contents and structure

A prerequisite for deriving pedotransfer functions which can be applied at a European scale was the availability of the basic soil data along with soil hydraulic properties from a wide range of soils across Europe. These data were fragmented and held in various institutions scattered throughout Europe. However a group of 20 institutions from 12 European countries recently collaborated to bring their available measured soil hydraulic data together into one central database. The development of the database was not a final product itself but was meant to be a starting point for various future research projects with the aims of, for example, the development of pedotransfer functions, or the creation of maps showing the hydrological aspects of European soils through the linkage with the existing 1:1,000,000 scale Soil Geographical Data Base of Europe (JAMAGNE et al., 1994).

Since its data would likely be in demand from a wide range of users for future soil, environment and climate research within Europe, it was important to have a database with a relational structure which allowed flexibility in data extraction, for example, using a variety of fields or by a combination of fields. Therefore HYPRES was developed within the Oracle Relational Database Management System™ . As Oracle uses SQL as its query language, it is compatible with many other database systems and so the use of Oracle does not exclude the transfer of data to other database formats. The desire to have compatibility with existing EU-wide soils databases led directly to the selection of the key identifying parameters used throughout HYPRES as well as many of the attributes stored.

The HYPRES database comprises six separate tables (see Table 1) each of which uses a *geo-reference* as the primary key and, where appropriate, also the *horizon* notation as the secondary key. The attributes stored in each table closely resemble those stored in the UNSODA database to avoid incompatibility between already existing international databases. These attributes were confirmed by the project participants as being the most important for improving existing models of soil hydrological processes and for deriving new pedotransfer functions (VAN GENUCHTEN & LEIJ, 1992). However, additional attributes needed to link with existing EU databases were also collected.

The BASICDATA table contains the „descriptor" data, for example, information on the soil type, where the soil profile was located and a description of the site and other environmental conditions. The unique primary key field is the geo-reference which is used both as a unique identifier and to ensure that referential integrity is maintained within the subsequent tables so that all data can be related to these descriptor data. This field also provides a link between the database and other European soil databases and allows the data to be related to other spatially referenced geo-physical factors, such as climate or land use. Before the data could be manipulated within EU-wide Geographic Information Systems (GIS), national geo-referencing had to be converted to a common system. To allow linkages to be made with the 1:1,000,000 scale Soil Geographical Data Base of Europe (JAMAGNE et al., 1994) each soil sample was named according to the modified FAO soil legend at the soil unit level wherever possible (CEC, 1985).

The five remaining tables are linked by both the *geo-reference* and the *horizon* notation as each soil profile generally contains more than one sampled horizon. The horizon designation is the secondary key which, when used in conjunction with the geo-reference, uniquely identifies each sample. It follows the FAO system (FAO, 1990) and carries the main pedological information. It is formatted according to a set of rules in order to simplify the selection of data and to allow greater refinement in this selection process. As the *horizon* field is a key identifying attribute which links tables, it follows that there must also be a unique combination of *geo-reference* and *horizon* for each sample in the database. Possible sample replicates are marked with an additional alphabetic character in the *horizon* notation.

The table SOIL_PROPS stores most of the data essential to the derivation of pedotransfer functions, such as particle size distribution data, organic matter contents and bulk densities as well as additional pedological information. The HYDRAULIC_PROPS table holds only derived or standardized data such as the derived Mualem – van Genuchten parameters (VAN GENUCHTEN et al., 1991) and calculated soil moisture retention and hydraulic conductivities at 14 pre-determined pressure heads. The „RAW" tables, that is RAWRET, RAWK and RAWPSD, store the original data on moisture retention, conductivity and particle size distributions, respectively, which were contributed by the network partners and is in its „raw" state, that is, prior to any standardization. The RAWRET and RAWK tables are very large, containing numerous pairs of data which are not readily usable, however, it is important to continue to store these „raw" data for a variety of reasons, for example, if new or improved parameterization methods become available or for testing and comparing novel methods of analysis.

For more detailed description of the database we refer to WÖSTEN et al. (1998) and LILLY et al. (1998).

Data standardization

As the data were collected from over 10 countries, it was inevitable that different methods had been used in their derivation, for example, moisture retention data were derived by both evaporation methods or by desorption which gave rise to an imbalance in the number of θ(h) data pairs throughout the database. Similarly, different class intervals were used in describing the soil particle size distribution. Therefore, in order to utilize these disparate data, it was necessary to standardize both the particle size and the hydraulic data.

The particle size data within the Soil Geographical Data Base of Europe follows the FAO system and in order to make HYPRES compatible with this database it was decided to standardize the HYPRES texture data on that system. Standardization of particle size data, where it was necessary, included applying an interpolation technique to estimate intermediate points on the cumulative particle size distribution curve of the soil, using techniques developed by NEMES et al. (1998). Texture classes defined by FAO were used while grouping the soils and those were further stratified as topsoils or subsoils. This resulted in 11 texture classes, that is 5 topsoil classes, 5 subsoil classes plus an organic class. Definition of FAO soil texture classes and the organic (Histic) layer can be found in FAO (1990). Soil samples with the FAO horizon notation *A* were classified as topsoils, all others were classified as subsoils (apart from the histic horizons). There was no further stratification within the organic class. It is important that the original particle size data for each soil sample are stored in the RAWPSD table, so those are available for any further research.

Before being able to develop pedotransfer functions it was necessary to standardize the soil hydraulic data as well to reduce the possibility of statistical bias caused by the imbalanced number of datapairs available for each sample. The volumetric water content, θ, and hydraulic conductivity, K, as functions of pressure head, h, were parameterized with the equations derived by VAN GENUCHTEN (1980). The modified version of the non-linear least-squares optimization program RETC (VAN GENUCHTEN et al., 1991) was used to predict the unknown Mualem–van Genuchten parameters (θ_r, θ_s, K_s, α, l and n) simultaneously from measured water retention and hydraulic conductivity data. The subscripts r and s refer to residual and saturated values and α, l, and n are the parameters that determine the shape of the curves.

Once the parameterization was completed, the optimized Mualem–van Genuchten model parameters were used to generate water content and hydraulic conductivity values for the following 14 selected pressure values: 0, -10, -20, -50, -100, -200, -250, -500, -1000, -2000, -5000, -10000, -15000, -16000 cms, which were selected as the most commonly used pressure head values in research throughout Europe. In this way all soil horizons, regardless of the number of measured data points could be represented by an equal weight in the process of developing pedotransfer functions The derived data are stored in the HYDRAULIC_PROPS table (see Table 1). In the meantime the original measured data are kept as those were

Table 1

Structure of the HYPRES Database

BASICDATA	SOIL_PROPS	HYDRAULIC_PROPS	RAWPSD	RAWRET	RAWK
geo-reference	geo-reference	geo-reference	geo-reference	geo-reference	Geo-reference
local sitename	horizon	horizon	horizon	horizon	horizon
soil name (in FAO class.)	upper sample depth	dMVG param. θsat.	particle size range	flag lab. / field	flag laboratory / field
country of origin	lower sample depth	dMVG param. θres.	percent. of particles	presure head (h)	indicator if k(h) / k(θ)
local geo-reference	primary structure (FAO)	dMVG param. alpha		θ at pres. head value	value of indicator variable
local soil name	second. structure (FAO)	dMVG param. n			hydr. cond. (k) at ind. value
local soil series	percent. of clay (FAO)	dMVG param. m			
highest groundwater depth	percent. of silt (FAO)	dMVG param. l			
lowest groundwater depth	percent. of sand (FAO)	dMVG param. Ksat.			
site description	Ksaturated	theta(1)			
sampling date	saturated water-content	theta(2)			
annual rainfall	bulk density				
average temp. in January	particle density	…theta(14)			
average temp. in July	porosity	conductivity(1)			
name of contact person	organic material	conductivity(2)			
address of contact person	MVG param. θsat.				
email of contact person	MVG param. θres.	…conductivity(14)			
relevant publication	MVG param. alpha				
comments1	MVG param. n				
comments2	MVG param. m				
keywords of used methods	MVG param. l				
No. of horizons in profile	MVG param. Ksat.				
rating of data quality	flag if estim. PSD data				
data quality rated by	comments				
	keywords of methods				

Remarks: Each column represents a separate table linked by geo-reference and horizon. Notation MVG refers to the equations of Mualem and van Genuchten, dMVG referes to derived parameters of the same equations contributed to allow any further users to use original data for their projects. These original soil hydraulic data can be found in the RAWRET and RAWK tables (see Table 1).

Applications for the database

Class pedotransfer functions for the soil texture classes and continuous pedotransfer functions were then developed for the database according to WÖSTEN et al. (1998). This procedure of deriving pedotransfer functions has proven to be successful in a number of cases with the effect that the movement of water and solutes could be simulated more accurately (WÖSTEN et al., 1990; WÖSTEN & VAN DER ZEE, 1993). The transfer functions and the data stored in the database can be used as direct inputs into simulation models describing water and solute transport in soils for a certain region or to predict yield of crops.

Throughout the study great care was taken to ensure that the HYPRES database and the derived products were compatible with existing EU soil databases, like the Soil Geographical Data Base of Europe (JAMAGNE et al., 1994) behind the 1:1,000,000 Soil Map of Europe. The class pedotransfer functions comprise geometric mean water retention and hydraulic conductivity properties for the 11 soil texture classes which accord with those used in the 1:1,000,000 Soil Geographical Data Base of Europe. Each Soil Typological Unit (STU) of the Soil Geographical Database was characterized by its topsoil and subsoil textures, soil depths and horizon thickness (KING et al., 1995). Using the same texture classes allowed a soil physical interpretation of the existing 1:1,000,000 Soil Map of Europe and thus it is possible to generate information on the soil physical composition of the unsaturated zone for areas of land within Europe. These data can also be combined within a GIS with other biophysical or administrative region data to derive more complex predictions of a wide range of environmental processes and land suitabilities.

Recommendations for a National Database in Hungary

It is of a great interest to develop a similar database of soil hydraulic data in Hungary by bringing together existing data from different sources (VÁRALLYAY et al., 1979, 1980). In order to allow the sharing of data and expertise, it would be desirable to develop such a database along similar lines to HYPRES.

What advantages are there for constructing such a national database and making it compatible with a major European database?

– Researchers could get information about the spatial distribution of soil hydrological properties within the country.

– Users of such a systematic database can have access to data from a wider range of soils to extend their research beyond a national level.

– Linking to other national or international databases will allow various new environment related spatial information to be generated.

– By sharing information or even by joining the European database in the future, will allow more meaningful pedotransfer functions to be developed for all European soil types and allow their more widespread use for agricultural and environmental studies in Europe.

As the development of a European level database has already been proven, this would suggest that a similar approach to the development of a Hungarian database along similar lines may also be successful and that expertise exists to make its development and implementation as smooth as possible. Therefore, it is suggested that we follow the structure of HYPRES as described in this paper and in greater detail by WÖSTEN et al. (1998). The Research Institute for Soil Science and Agricultural Chemistry of the Hungarian Academy of Sciences is willing to coordinate such a project to establish the national database for measured soil hydraulic functions. In doing so, the institute will provide the data to establish the first version of the database. There are currently some 300-400 soil layers from about 100 soil profiles. These are collected from various regions of the country, and the samples cover a wide range of soil types, which include soils in agricultural use as well as soils under forests (VÁRALLYAY et al., 1980). To achieve compatibility with the European database multiple harmonization of data will be required, such as re-classification of soil profiles and horizons according to FAO (1990) (VÁRALLYAY, 1994; VÁRALLYAY et al., 1994). After completion of the classification, soil textural grouping and analysis of soil hydraulic data may begin. As is shown in HYPRES, this does not necessarily mean the loss of data as the database can allow existing soil classifications to be retained in addition to international classifications. The key element in the development of HYPRES was its flexibility, this can easily be extended to accommodate Hungarian (and other) data.

Summary

It is of a major concern to protect soils and water from being threatened and damaged by agricultural and industrial pollution. Simulation models which were developed to describe water and solute transport in the vadose zone can be used to help in the effective management of soils to reduce this potential damage. These models require a large variety of input data, in particular, soil hydraulic data. In the framework of an EU funded project, an international database has been successfully established with the participation of 12 countries. The HYdraulic PRoperties of European Soils (HYPRES) database holds both measured and predicted soil hydraulic data as well as environmental and pedological information originating from these 12 countries. This information can be used directly in these soil water simulation models and, as the database was designed to be compatible with other EU-wide soils databases and GIS, it is possible to extend the use of the database to a wide range of environmental and land related applications.

It is proposed to develop a similar, national database for soil hydraulic properties in Hungary. Developing this database will enhance the ability of Hungarian soil scientists to address many environmental issues of concern. In order to allow exchange of ideas and data, it makes sense to design this new database in a manner similar to HYPRES and so make it compatible.

References

ADDISCOTT, T. M. & WAGENET, R. J., 1985. Concepts of solute leaching in soils: A review of modeling approaches. J. Soil Sci. 36. 411-424.

BOUMA, J., 1989. Using soil survey data for quantitative land evaluation. In: Advances in Soil Science. Vol. 9. (Ed.: STEWART, B. A.) 117-123. Springer. Berlin.

COMMISSION OF THE EUROPEAN COMMUNITIES, 1985. Soil map of the European Communities, 1:1,000,000, Explanatory text. Directorate General Agriculture, Coordination of Agricultural Research. CEC. Luxemburg.

EIJSACKERS, H. J. P. & HAMERS, T. (Eds.), 1993. Integrated Soil and Sediment Research: A Basis for Proper Protection. Selected Proceedings of the First European Conference On Integrated Research for Soil and Sediment Protection and Remediation (EUROSOL), Maastricht, The Netherlands, 6-12 September 1992.

FAO, 1990. Guidelines for Soil Descriptions. 3rd ed. FAO. Rome.

GREEN, R. E., AHUJA, L. R. & CHONG, S. K., 1986. Hydraulic conductivity, diffusivity and sorptivity of unsaturated soils: field methods. In: Methods of Soil Analysis (Ed.: KLUTE, A.) Part 1. 771-798. 2nd ed. Agronomy Monogr. 9. ASA and SSSA, Madison, Wisconsin.

JAMAGNE, M. et al., 1994. Creation and use of a European Soil Geographic Database. Transactions of the 15th International Congress of Soil Science. 6a. 728-742. Acapulco, Mexico.

KING, D. et al., 1995. The EU Soil Geographical Database. In: European Land Information Systems for Agro-environmental Monitoring. (Eds.: KING, D., JONES, R. J. A. & THOMASSON, A. J.), Joint Research Centre. Ispra, Italy.

KLUTE, A. & DIRKSEN, C., 1986. Hydraulic conductivity and diffusivity: laboratory methods. In: Methods of Soil Analysis. (Ed.: KLUTE, A.) Part 1. 687-734. 2nd ed. Agronomy Monogr. 9. ASA and SSSA. Madison, Wisconsin.

LEIJ, F. J. et al., 1994. Unsaturated Soil Hydraulic Database, UNSODA 1.0 User's Manual. Office of Research and Development. United States Environmental Protection Agency. Ada, Oklahoma.

LILLY, A. et al., 1998. The development and use of the HYPRES Database in Europe. Proceedings of the International Workshop on Characterisation and Measurement of the Hydraulic Properties of Unsaturated Porous Media. Riverside, California. 22-24 October 1997. (In press).

NEMES, A. et al., 1998. Evaluation of different techniques for interpolation of the particle-size distribution. Proceedings of the International Workshop on Characterisation and Measurement of the Hydraulic Properties of Unsaturated Porous Media. Riverside, California. 22-24 October 1997. (In press).

RAJKAI, K., 1984. A method for calculating the capillary conductivity of the soil on the basis of the water retention (pF) curve. (In Hungarian) Agrokémia és Talajtan. 33. 409-438.

RAJKAI, K. & VÁRALLYAY, GY., 1989. Estimative calculation of hydrophysical parameters from simply measurable soil properties. Agrokémia és Talajtan. 38. 634-640.

RAJKAI, K. et al., 1981. Calculation of water retention data from the texture and the bulk density of soils. (In Hungarian) Agrokémia és Talajtan. 30. 409-438.

SOIL SURVEY STAFF., 1975. Soil Taxonomy. A Basic System of Soil Classification for Making and Interpreting Soil Surveys. USDA-SCS, Agric. Handbook No. 436. U.S. Government Printing Office. Washington D. C.

TENG, P. S. & PENNING DE VRIES, F. W. T. (Eds.), 1992. Systems Approaches for Agricultural Development. Proc. of the International Symposium held at the Asian Institute of Technology, Bangkok, Thailand, 2-6 December 1991.

VAN GENUCHTEN, M. TH., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J. 44. 892-898.

VAN GENUCHTEN, M. TH. & LEIJ, F. J., 1992. On estimating the hydraulic properties of unsaturated soils. In: Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils. (Eds.: VAN GENUCHTEN, M.TH. & LEIJ, F. J. & LUND, L. J.) 1-14. Riverside, California.

VAN GENUCHTEN, M. TH., LEIJ, F. J. & YATES, S. R., 1991. The RETC Code for Quantifying the Hydraulic Functions of Unsaturated Soils. USDA, US Salinity Laboratory, Riverside, California. United States EPA document EPA/600/2-91/065.

WARRICK, A. W. & MYERS, D. E., 1987. Calculations of error variances with standardized variograms. Soil Sci. Soc. Am. J. 51. 265-268.

VÁRALLYAY, GY., 1994. Harmonization of soil conservation monitoring systems. In: FAO/ECE Intern. Workshop on Harmonization of Soil Conservation Monitoring Systems (Budapest, 14-17 Sept., 1993). 11-16. RISSAC. Budapest.

VÁRALLYAY, GY. et al., 1979. Mathematical description of water retention curves. (In Hungarian) Agrokémia és Talajtan. 28. 15-38.

VÁRALLYAY, GY. et al., 1980. Soil water management categories of Hungarian soils and the map of soil water properties (1:100,000) (In Hungarian) Agrokémia és Talajtan. 29. 77-112.

VÁRALLYAY, GY. et al., 1994. SOTER (Soil and Terrain Digital Database) 1:500,000 and its application in Hungary. Agrokémia és Talajtan. 43. 87-108.

WÖSTEN, J. H. M. & VAN DER ZEE, S. E. A. T. M., 1993. Vulnerability of three spatially variable soil groups for solute leaching. Hydrological Processes. 7. 235-247.

WÖSTEN, J. H. M., LILLY, A. & NEMES, A., 1998. Using existing soil data to derive hydraulic parameters for simulation models in environmental studies and in land use planning. Staring Centre Report No. 156. The Winand Staring Centre. Wageningen.

WÖSTEN, J. H. M. et al., 1990. Functional sensitivity analysis of four methods to generate soil hydraulic functions. Soil Sci. Soc. Am. J. 54. 832-836.