Development of Soil Hydraulic Pedotransfer Functions on a European scale: Their Usefulness in the Assessment of Soil Quality

A. Nemes*, J.H.M. Wösten, and A. Lilly

ABSTRACT

Many simulation models concerned with water and solute movement in the vadose zone require water retention and unsaturated hydraulic conductivity data. However, these properties are often difficult and time consuming to measure. One method for overcoming this lack of data is through the development of pedotransfer functions, which predict soil hydraulic properties from other, more easily measured soil properties. Recently the European Union funded a project (entitled: Using existing soil data to derive hydraulic parameters for simulation modeling in environmental studies and in land use planning ~ CHRX-CT94-0639) in which 20 Institutions from 12 European countries collaborated to develop the HYPRES (Hydraulic Properties of European Soils) database [Wösten et al., 1999]. This database draws together the basic soil information and soil hydraulic data, which existed at the collaborating Institutions and from which pedotransfer functions (PTFs) applicable to Europe can be derived.

The project had five distinct phases: the first was to collect all the data from the participating Institutions. Many of the soil properties collected are similar to those in the UNSODA database [Leij et al., 1994]. In the second phase, a database structure had to be designed and populated with the relevant soil properties in such a way as to remain compatible with existing EU-wide soil databases. Data standardization took place in the third phase where it was necessary, and the fourth phase was to derive continuous and class pedotransfer functions. These class pedotransfer functions were determined for each of the five soil texture classes and for two pedological classes within them (topsoils and subsoils) plus an additional class, which encompassed the organic soil horizons giving a total of 11 classes. This classification was adopted from the 1:1,000,000 scale Soil Geographical Database of Europe [Jamagne et al., 1994]. The final phase was to demonstrate some of the capabilities and possible uses of the database by linking the class pedotransfer functions with the Soil Geographical Database of Europe to provide a spatial distribution of the soil hydraulic properties. These can be viewed as baseline data on particular soil quality parameters under existing land use systems.

INTRODUCTION

Soil hydraulic properties are key aspects in determining soil quality and soil function. Many soil water and solute transport models are currently applied for the investigation and prediction of a wide range of complex environmental processes, which are indicative of soil quality, for example, infiltration capacity. These models require data on soil water retention and hydraulic conductivity characteristics. However, collection of these data is difficult, time consuming and sometimes rather costly, so there is a continued interest in the establishment of pedotransfer functions, which predict soil hydraulic properties from other - more easily measured - soil properties. The European Union (EU) funded a project (entitled: Using existing soil data to derive hydraulic parameters for simulation modeling in environmental studies and in land use planning ~ CHRX-CT94-0639) in which 20 Institutions from 12 European countries collaborated to develop the HYPRES (Hydraulic Properties of European Soils) database [Wösten et al., 1999]. This database draws together the basic soil information and soil hydraulic data, which existed at the collaborating Institutions and from which pedotransfer functions (PTFs) applicable to Europe can be derived.

The project had five distinct phases: the first was to collect all the data from the participating Institutions. Many of the soil properties collected are similar to those in the UNSODA database [Leij et al., 1994]. In the second phase, a database structure had to be designed and populated with the relevant soil properties in such a way as to remain compatible with existing EU-wide soil databases. Data standardization took place in the third phase where it was necessary, and the fourth phase was to derive continuous and class pedotransfer functions. These class pedotransfer functions were determined for each of the five soil texture classes and for two pedological classes within them (topsoils and subsoils) plus an additional class, which encompassed the organic soil horizons giving a total of 11 classes. This classification was adopted from the 1:1,000,000 scale Soil Geographical Database of Europe [Jamagne et al., 1994]. The final phase was to demonstrate some of the capabilities and possible uses of the database by linking the class pedotransfer functions with the Soil Geographical Database of Europe to provide a spatial distribution of the soil hydraulic properties. These can be viewed as baseline data on particular soil quality parameters under existing land use systems.

Data collection

The data were collected from Institutions in Belgium, Denmark, France, Germany, Greece, Italy, The Netherlands, Portugal, Scotland, England, Wales and Northern Ireland (all UK), Slovakia, Spain and Sweden. The data were...
transferred to the co-coordinators via electronic media (floppy disk and email) and as hard copy. There are 5521 soil samples (including replicates) from 4,486 soil horizons 1,777 soil profiles in version V1.0 of the HYPRES database. These were taken from 1,777 individual locations scattered throughout the collaborating nations and represent 95 different soil types according to the modified FAO Soil Legend [CEC, 1985]. A total of 197,844 data pairs of water retention *versus* pressure head and 120,419 data pairs of unsaturated hydraulic conductivity *versus* pressure head were collected. There were 1,136 soil horizons with both measured water retention and unsaturated hydraulic conductivity data.

**Database design and structure**

Although the derivation of the pedotransfer functions was the main objective of the project, the database can also be seen as one of the main products, which may have a wide application in future research projects examining the effects of different land use systems on soil hydraulic properties as well as in soil, environmental and climate research. Therefore it was important to develop a database with a relational structure which allowed a high degree of flexibility in data extraction by a variety of fields or a combination of fields, in data manipulation and which could accommodate the great diversity and volume of data collected. The database was developed within the Oracle Relational Database Management System™, which uses SQL as its query language. Data can be entered and retrieved in a variety of formats, the whole database can easily be updated with a few simple commands, and new fields or tables can be added any time and data can easily be transferred to other database systems.

Version V1.0 of the HYPRES database comprises six separate tables (Figure 1), which are linked by the field 'geo-reference' and, where appropriate, also by the horizon notation ('horizon'). The attributes stored in each table closely resemble those stored in UNSODA database as these attributes were deemed by the delegates of the International Workshop on "Indirect methods for estimating the hydraulic properties of unsaturated soils" to be the most important for improving existing parametric and physico-empirical models and for deriving new pedotransfer functions [van Genuchten and Leij, 1992]. However, additional attributes, which were needed to link with existing EU databases, were also collected.

The BASICDATA table (Figure 2) contains the 'descriptor' data, for example, information on the soil type, source information (the original source of the sample), location of the soil profile and a description of the site and some of its environmental conditions. The unique primary key field is '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 the descriptor data. This field also provides a link between HYPRES and other European soil datasets and allows the data to be related to other spatially referenced geo-physical factors such as climate or land use without the need to collect and store these additional data in the database. The geo-references were all converted to a common EU-wide system (wherever possible) from a variety of different co-ordinate systems, which are in current use throughout Europe. This table also stores comments on methods used for sampling, data collection and the land use at the time the sample was taken. Where the BASICDATA table is linked to all five other tables by the geo-reference, those tables are linked to each other by both 'geo-reference' and the 'horizon' notation as each soil profile generally contains more than one sampled horizon. The horizon symbols follow the FAO system [FAO, 1990] and are spaced according to a set of rules in order to make the selection of data easier and to allow greater refinement.
in this selection process, for example, certain character spaces within the string are reserved for Master horizons only while others contain Sub-horizon symbols. 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. Where a horizon was sampled in replicate, an additional alphabetic character was added to the end of
the horizon notations. Information on the soil horizon relative to soil hydraulic properties provides a link between soil forming factors and indicators of soil quality.

The table SOIL_PROPS stores most of the data likely to be used in deriving the pedotransfer functions for example particle size distribution data, organic material contents, bulk density and porosity. This table also holds the Mualem-van Genuchten parameters where those were contributed rather than paired $\theta(h)$ or $K(h)$ data. Where a system of particle size distribution other than the FAO/USDA was used, these data had to be standardized and so the proportions of the various size fractions according to the FAO system were estimated where necessary. These estimates were stored in this table. Just as the BASICDATA table, this table also has comments on the methods used to collect the stored data.

The 'RAW' tables, that is RAWRET, RAWK and RAWPSD store the data on moisture retention, hydraulic conductivity and particle size distributions consecutively. These data are in their 'raw' state (that is, prior to any standardization) as contributed by the network partners. Although these tables are very large, containing numerous pairs of data and are not readily usable, it is important 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. As the data stored in these tables are the actual measurements of such properties as saturated hydraulic conductivity and saturated moisture contents, they can be viewed as an important resource in defining limits to these properties as indicators of soil quality and soil function.

As the soil hydraulic data were derived by many different methods, they were standardized prior to the development of the pedotransfer functions by deriving a set of Mualem-van Genuchten parameters for each individual sample (where possible). These fitted Mualem-van Genuchten parameters were subsequently used to calculate moisture retention and hydraulic conductivity data at 14 different pre-determined pressure heads for each soil horizon. These derived hydraulic data are stored in table HYDRAULIC_PROPS.

**Data standardization**

As the data were collected from 12 different countries, it was inevitable that different class definitions and intervals were used in describing the soil particle size distribution. Therefore, in order to be able to utilize most of these disparate data in the database, it was necessary to find a solution to standardize the particle size distribution. Similarly, it is important for simulation modeling to have a uniform dataset. As different methods had been used in the derivation of soil hydraulic data, for example, water retention data were derived by both evaporation methods and by desorption techniques, the hydraulic properties were standardized using the Mualem-van Genuchten parameters.

**Standardization of the particle size distribution**

As the particle size classification within the Soil Geographical Database of Europe follows the FAO system (FAO, 1990), it was decided to standardize the texture data within HYPRES on the same particle size classes. The FAO system describes clays as particles <2 µm, silts between 2 and 50 µm and sands between 50 and 2000 µm, however, many countries in Europe use different class intervals. Therefore, in order to derive European-wide continuous or class pedotransfer functions there was a need to standardize the soil texture data. In general, this involved estimating the proportions of particles in the 2-50 µm range from datasets where, for example, silt was defined as the fraction between 20 and 60 µm or 20 and 63 µm or where sand was defined as the fraction between 20 and 2000 µm. Approximately half of the horizons required the particle size data to be standardized in this way.

Four methods for estimating the particle size distribution as defined by the FAO were tested against each other [Nemes et al., 1999]. Two of the most successful methods were then applied to interpolate intermediate points of each individual cumulative particle size distribution curve. One procedure (fitting an open form spline function to the measured data) was used where the measured points in the interpolation were in close proximity to the point to be estimated, for example, where the silt fraction was taken as being less than 60 or 63 µm. The second interpolation procedure involved comparing the cumulative particle size distribution with a well quantified reference dataset. An iteration procedure was used to select at least 10 cumulative particle size curves from the reference set which were most similar to that with the missing 50 µm point. The proportion of particles smaller than 50 µm were then read from these 10 selected curves. The average value was then applied to the curve with the missing 50 µm point as the estimated value. The standardized data will allow more direct comparisons of the measured hydraulic properties from different locations throughout Europe based on particle size class and soil texture.

**Standardization of the soil hydraulic properties**

The hydraulic properties of some 3,891 soil horizons were parameterized using the Mualem-van Genuchten equations for water retention and for unsaturated hydraulic conductivity where applicable. The remainder did not have sufficient soil hydraulic data to be parameterized and were eliminated from further analysis. Parameterization was necessary as in this way, soil horizons with data derived from the evaporation method and those where the data were from simple desorption techniques - which generally generates much less data points - could be successfully combined with a minimal risk of statistical bias in the data, regardless of the number of measured points and the applied pressure head values. The data were parameterized using a modified version of the RETC code [van Genuchten et al., 1991]. This semi-automated procedure involved visual
inspection of the goodness-of-fit. When the parameterization was complete, the parameters were then used to generate soil water retention and unsaturated hydraulic conductivity points for 14 pre-determined pressure head values. These 14 data pairs were then used to determine the class pedotransfer functions. Standardization of the soil hydraulic properties is important for retaining a degree of integrity when using these data for simulation modeling at regional scales.

**Pedotransfer functions**

The preceding chapters described the steps that were taken in order to establish a suitable dataset primarily for the derivation of pedotransfer functions. Two different sets of pedotransfer functions were derived. Class pedotransfer functions predict the hydraulic characteristics for each soil texture class and continuous pedotransfer functions can predict hydraulic properties from individual measurements of sand, silt, clay, organic matter and bulk density.

**Class pedotransfer functions**

The class pedotransfer functions are so called as they predict the average hydraulic characteristics for a soil texture/pedological class. These texture classes and pedological classes (topsoil or subsoil) are those contained within the 1:1,000,000 Soil Geographical Database of Europe (Figure 3). Firstly, the optimized Mualem-van Genuchten parameters were used to determine the moisture contents and conductivities at 14 pressure heads for each soil horizon. Next, the horizons were grouped by the 11 texture/pedological classes referred to above. The individual water retention and hydraulic conductivity curves derived for each soil horizon were plotted in soil textural/pedological groups and any outliers were identified and removed from subsequent calculations. The geometric means and standard deviations of the water content and hydraulic conductivity were then calculated for each of the 14 generated points within each texture/pedological class. This allowed the derivation of average curves and a measure of their variability around the geometric mean (Figure 4) for each soil texture/pedological class.

Mualem-van Genuchten parameters were then derived for these average curves in order to reduce each to a set of simple parameters, which could be used as input into soil water simulation models. In most cases these curves were closely aligned when plotted. Table 1 summarizes the Mualem-van Genuchten parameters of the class pedotransfer functions.

![Figure 3. The FAO/USDA soil texture triangle and the 5 texture classes acknowledged by the 1:1,000,000 scale Soil Geographical Database and Soil Map of Europe.](image-url)
Figure 4. Individually measured and calculated average water retention and hydraulic conductivity characteristics.
Continuous pedotransfer functions
Unlike early forms of continuous pedotransfer functions [for example, Ahuja et al., 1985; Rawls and Brakensiek, 1985], regression equations were developed to predict Mualem-van Genuchten parameters instead of specific individual points of the water retention and/or the hydraulic conductivity characteristics following those described in Wösten et al. (1994). By predicting the Mualem-van Genuchten parameters from soil particle size distribution, it is then possible to predict both the moisture content and hydraulic conductivity for a wider range of pressure heads giving greater flexibility than methods which predict tabular values only.

Multivariate linear regression was used to investigate the dependency of each Mualem-van Genuchten model parameter on the more easily measured, basic soil properties. To comply with a number of physical boundary conditions, transformed model parameters were used in the regression analysis rather than the original Mualem-van Genuchten parameters. Basic soil properties like percent clay, percent silt, organic matter content, bulk density were used as regression variables. The qualitative variables of topsoil and subsoil were converted to binary data (0 or 1) for
Table 1. The Mualem-van Genuchten parameters of the class pedotransfer functions.

<table>
<thead>
<tr>
<th></th>
<th>θr</th>
<th>θs</th>
<th>α</th>
<th>n</th>
<th>m</th>
<th>l</th>
<th>Ks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topsoils</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coarse</td>
<td>0.025</td>
<td>0.403</td>
<td>0.0383</td>
<td>1.3774</td>
<td>0.2740</td>
<td>1.2500</td>
<td>60.000</td>
</tr>
<tr>
<td>Medium</td>
<td>0.010</td>
<td>0.439</td>
<td>0.0314</td>
<td>1.1804</td>
<td>0.1528</td>
<td>-2.3421</td>
<td>12.061</td>
</tr>
<tr>
<td>Medium Fine</td>
<td>0.010</td>
<td>0.430</td>
<td>0.0083</td>
<td>1.2539</td>
<td>0.2025</td>
<td>-0.5884</td>
<td>2.272</td>
</tr>
<tr>
<td>Fine</td>
<td>0.010</td>
<td>0.520</td>
<td>0.0367</td>
<td>1.1012</td>
<td>0.0919</td>
<td>-1.9772</td>
<td>24.800</td>
</tr>
<tr>
<td>Very Fine</td>
<td>0.010</td>
<td>0.614</td>
<td>0.0265</td>
<td>1.1033</td>
<td>0.0936</td>
<td>2.5000</td>
<td>15.000</td>
</tr>
<tr>
<td>Subsoils</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coarse</td>
<td>0.025</td>
<td>0.366</td>
<td>0.0430</td>
<td>1.5206</td>
<td>0.3424</td>
<td>1.2500</td>
<td>70.000</td>
</tr>
<tr>
<td>Medium</td>
<td>0.010</td>
<td>0.392</td>
<td>0.0249</td>
<td>1.1689</td>
<td>0.1445</td>
<td>-0.7437</td>
<td>10.755</td>
</tr>
<tr>
<td>Medium Fine</td>
<td>0.010</td>
<td>0.412</td>
<td>0.0082</td>
<td>1.0861</td>
<td>0.0793</td>
<td>-3.7124</td>
<td>8.500</td>
</tr>
<tr>
<td>Fine</td>
<td>0.010</td>
<td>0.481</td>
<td>0.0198</td>
<td>1.0861</td>
<td>0.0793</td>
<td>-3.7124</td>
<td>8.500</td>
</tr>
<tr>
<td>Very Fine</td>
<td>0.010</td>
<td>0.538</td>
<td>0.0168</td>
<td>1.0730</td>
<td>0.0001</td>
<td>8.235</td>
<td></td>
</tr>
<tr>
<td>Organic*</td>
<td>0.010</td>
<td>0.766</td>
<td>0.0130</td>
<td>1.2039</td>
<td>0.1694</td>
<td>0.4000</td>
<td>8.000</td>
</tr>
</tbody>
</table>

* There was no topsoil/subsoil distinction among organic soils.

Table 2. Continuous pedotransfer functions: a prediction of the Mualem-van Genuchten parameters for the individual soil horizons. The Mualem-van Genuchten parameters were transformed as \( K_s^* = \ln (K_s) \), \( a^* = \ln (a) \), \( n^* = \ln (n-1) \) and \( l^* = \ln ((l+10) / (10-l)) \) to comply with a number of physical boundary conditions. \( \theta_s \) was not transformed. The regressed variables were: \( C \) = percent clay, \( S \) = percent silt, \( OM \) = organic material content, \( D \) = bulk density; topsoil and subsoil were qualitative variables having the value of 1 or 0.

\[
\theta_s = 0.7919 + 0.001691* C - 0.29619* D - 0.000001491* S^2 + 0.0000821* OM^2 + 0.02427* C^{-1} + 0.01113* S^{-1} + 0.01472* \ln(S) - 0.0000733* OM* C - 0.000619* D* C - 0.001183* D* OM - 0.0001664* topsoil* S
\]

\[
\alpha^* = -14.96 + 0.03135* C + 0.0351* S + 0.646* OM + 15.29* D - 0.192* topsoil - 4.671* D^2 - 0.000781* C^2 - 0.00687* OM^2 + 0.0449* OM^{-1} + 0.0663* \ln(S) + 0.1482* \ln(OM) - 0.04546* D* S - 0.4852* D* OM + 0.00673* topsoil* C
\]

\[
n^* = -25.23 - 0.02195* C + 0.0074* S - 0.1940* OM + 45.5* D - 7.24* D^2 + 0.0003658* C^2 + 0.002885* OM^2 - 12.81* D^{-1} - 0.1524* S^{-1} - 0.19158* OM^{-1} - 0.2876* \ln(S) - 0.0709* \ln(OM) - 44.6* \ln(D) - 0.02264* D* C + 0.0896* D* OM + 0.00718* topsoil* C
\]

\[
l^* = 0.0202 + 0.0006193* C^2 - 0.001136* OM^2 - 0.2316* \ln(OM) - 0.03544* D* C + 0.00283* D* S + 0.0488* D* OM
\]

\[
K_s^* = 7.755 + 0.0352* C + 0.93* topsoil - 0.967* D^2 - 0.000484* C^2 - 0.000322* S^2 + 0.001* S^{-1} - 0.0748* OM^{-1} - 0.643* \ln(S) - 0.01398* D* C - 0.1673* D* OM + 0.02986* topsoil* C - 0.03305* topsoil* S
\]

The resulting continuous pedotransfer functions are presented in Table 2. The hydraulic characteristics can be obtained after the back-transformation to the original Mualem-van Genuchten model parameters.
Organic material content or changes of bulk density (tillage)

The continuous pedotransfer functions would allow researchers on soil quality to examine the effects of changes in organic matter content and bulk density (tillage or soil compaction) on the soil hydraulic properties. Such information may then be used in studies like vulnerability of soils to physical or chemical degradation. Future plans include the extension of the database to Central and Eastern European countries, and to apply and compare neural network models with the classical regression type models to derive more accurate pedotransfer functions.

ACKNOWLEDGEMENTS

The authors thank the European Soil Bureau for the permission to use the 1:1,000,000 Soil Geographical Database of Europe. The financial support of the Soros Foundation is gratefully acknowledged. Part of this work was funded by the Scottish Executive Rural Affairs Department.

REFERENCES


Commission of the European Communities (CEC), 1985. Soil map of the European Communities. Scale 1:1,000,000. Luxembourg.


