Unsaturated zone modelling:
the role of soil database classification

Z. Kozma, T. Ács & L. Koncsos
Department of Sanitary and Environmental Engineering,
Budapest University of Technology and Economics, Hungary

Abstract

Physical and hydraulic soil properties are essential input parameters for models from different sciences (e.g. hydrology, agriculture, water management, nature preservation). Generally texture composition, porosity and other easily measurable physical properties of soils are known. However, saturated hydraulic conductivity and characteristic values of the water retention curve are usually missing information. Therefore, based on the physical similarity of soils (classes), they are substituted by data derived from soil databases. The aim of this study was to assess the currently unknown uncertainties of such classified databases. To do so, a large variety of tests were carried out: (i) static and dynamic, (ii) 1D and 3D (iii) hydraulic and hydrologic applied tests, (iv) real and synthetic soils, parameterized accordingly, and (v) HUNSODA and/or HYPRES databases. The results were sorted with respect to FAO and USDA classification systems. Soil class overlapping was evaluated through the statistics of basic hydraulic parameters (retention curve, hydraulic conductivity). Indicators related to hydrologic extremities (excess water and drought) were used to quantify the uncertainties of soil texture based on parameter substitution. It was concluded that the two evaluated soil classification systems did not sort soils reliably from the hydrologic and hydraulic viewpoint: the test results of classes showed major overlaps. Moreover, in most cases class synthetic parameter combinations poorly represented real soils. As a general consequence the results based on classified soil databases should be accepted only with reservation.

Keywords: unsaturated zone, soil database, soil classification, hydrological modelling, agricultural yield estimation.

1   Introduction

The hydraulic behaviour of unsaturated soils is an important factor in a number of environmental phenomena. The rate of water and solute transport in the unsaturated zone directly or indirectly governs most hydrologic processes, as well as erosion, plant growth, ecosystem functioning, nutrient cycling and diffuse pollution.

   According to its importance, unsaturated zone is assessed in the frame of many sciences at various scales. Among others hydrologic, water management, agricultural (e.g. crop yield – McCown *et al.* [1], van Ittersum and Rabbinge [2]) and climate (e.g. regional atmospheric) models must handle the impacts of hydraulic processes in the unsaturated zone to a certain extent. The topic has a great significance in Hungary as well, since the country has to face a unique combination of extreme hydrologic symptoms: the extensive, regular appearance of flood, excess water (see definition in e.g. van Leeuwen *et al.* [3]) and drought.

   As a result of increased computational capacity the theoretical insight is broadening remarkably in the topic. Up-to-date one- or multidimensional solvers as well as integrated hydrologic models attain great complexity and became essential parts of research (Harter and Hopmans [4]). However, there are still many open questions and problems in relation with understanding and predicting the role of vadose zone in landscape scale hydrology (Pachepsky *et al*. [5]). Probably the most important of these issues is data reliability. Considering the quality of the simulation results, usually the bottleneck is the input data. Soil maps and databases (Wösten *et al.* [6], Soil Survey Staff [7]) are used expansively to provide information about the 3D spatial distribution of various soil types and their hydraulic parameterization (e.g. water retention and hydraulic conductivity). Such databases are often developed by using soil classification. Though it is a common practice, even the magnitude of uncertainty associated to soil categorization is practically unknown.

   Concerning the aspects above, recently a research was started with the aim to (i) quantify the uncertainty related to classified soil data used for hydrologic modelling, and to (ii) derive useful information and hydrologic indicators for soil samples from Hungary. In the first phase of the study (Kozma *et al.* [8]) the well-known FAO classification system (FAO [9]) was analysed by using it on the Unsaturated Soil Hydraulic Database of Hungary (HUNSODA – Nemes [10]). Three possible directions were outlined for the second phase of the research:

  (a) the dynamic analysis of all soils in the database instead of using only parameter class envelopes,
  (b) the evaluation of the USDA classification system (USDA [11]) with the same methodology, and
  (c) the analysis of the FAO and/or USDA methods using the Database of Hydraulic Properties of European Soils (HYPRES).

In this paper task (a) and (b) will be discussed in detail. Since the complete HYPRES database is not accessible for the broad scientific community, task (c) is limited to the published FAO class averages derived for HYPRES.

The time-invariant and dynamic tests applied so far consider only a single homogeneous soil column (pedon), therefore their ability to analyze complex landscape scale hydrologic behaviour is limited. For this reason, another type of test, a simple sensitivity analysis was carried out as well: hydrologic processes and excess water were simulated with the WateRisk integrated hydrologic model (Kozma and Koncsos [12]) at a relatively humid Hungarian lowland watershed.

2 Theoretical background

2.1 Classified soil databases

As the direct determination of soil properties is expensive and time-consuming, such data is often provided by pre-parameterized spatial soil databases. These contain the 2D/3D distribution of soil types and physical parameters associated to them. The magnitude of uncertainty related to database elaboration is usually unknown, but probably there are two major sources of error: (i) spatial interpolation between sample locations, and (ii) soil classification. As stated above, in this research only (ii) is analyzed.

The aim of classification is to synthesize field data, so that predictive estimates/concepts can be set up about soil behaviour. Samples are categorized into soil types with respect to easily definable properties (e.g. genesis, morphology, texture). Classes are then characterized with geometric average values for also such properties that are more difficult to quantify or expensive to measure and can be missing for many soils (water retention, saturated hydraulic conductivity, etc.). This way classification allows the extrapolation of scarcely known soil information.
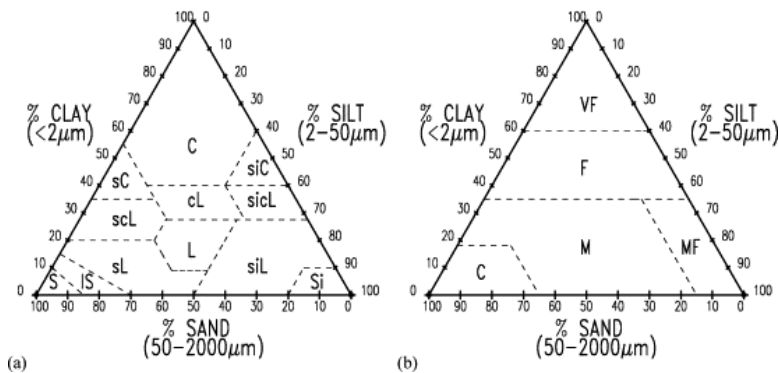

Figure 1:   The soil texture triangle showing class limits for (a) the USDA method and for (b) the FAO method.

Two of the most widespread systems are the FAO [9] and the USDA [11] method. Both robust approaches sort mineral soils into classes with respect to their textural composition (sand, silt and clay content). The applied limit values are shown on Figure 1. The USDA method uses 12 classes (class names are created from the words loam, sand, silt and clay, e.g.: sicL – silty clay Loam), while the FAO system uses only five types (C – coarse; M – medium; MF – medium-fine; F – fine; VF – very fine).

One would expect that the twelve-class USDA method provides more detailed and sound representation of soils. Contrary, Wösten *et al.* [6] and Nemes [10] suggest that the estimation of hydraulic behaviour is more reliable in case of the FAO method. This might be a reason, why this method is more accepted in Europe: among others the HYPRES as well as the HUNSODA was developed with the FAO method.

2.2 Soil hydraulic functions

Water flow and solute transport in various soils can differ significantly. To describe hydraulic behaviour, generally two characteristic functions are used: the WRC and the hydraulic conductivity function (HCF). The WRC provides the water content (θ [m^3 m^{-3}]) that a soil can hold at various matric pressures (ψ [cm]). Figure 2 illustrates the physically interpretable content of the WRC for an arbitrary medium textured soil (note that $pF = \log(-\psi)$).

Experimentally the WRC is defined by measuring θ–ψ data pairs. Various parametric analytical formulas can be used to express the WRC. These functions have to be fitted to the measured discrete data points. Most accepted among these is the function introduced by van Genuchten (vG) [13], which was used in this study as well. In eqn. (1) α and n are fitted parameters, while $m = 1 – 1/n$.
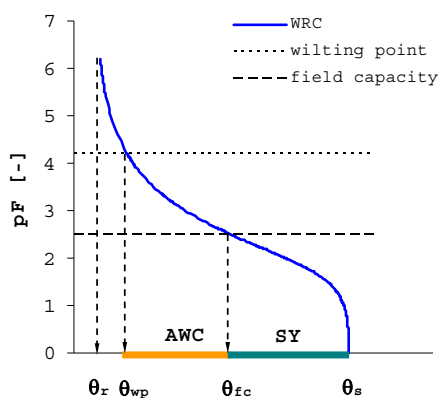
residual water content:
$$\theta_r = \theta(pF\ 6.2);$$
water content at wilting point:
$$\theta_{wp} = \theta(pF\ 4.2);$$
water content at field capacity:
$$\theta_{fc} = \theta(pF\ 2.5);$$
saturated water content:
$$\theta_s = \theta(pF\ 0);$$
available water content:
$$AWC = \theta_{fc} - \theta_{wp};$$
specific yield:
$$SY = \theta_s - \theta_{fc}.$$

Figure 2:   The water retention curve of a medium textured soil, showing characteristic water content values.

$$\theta(\psi) = \theta_r + (\theta_s - \theta_r)\left[1 + |\alpha\psi|^n\right]^{-m} \tag{1}$$

Hydraulic conductivity function describes the variation of the $k(\theta)/K_s$ [-] ratio with respect to water content (where $k(\theta)$ [cm day^{-1}] is the unsaturated hydraulic conductivity and K_s [cm day^{-1}] is the saturated hydraulic conductivity). Since K_s values can vary in a wide range within soil classes, usually this is the dominant source of uncertainty. Moreover, the HCF is strongly non-linear, also ranging over several orders of magnitude (OMs), and its measurement is heavily biased with error. In this research for the HCF – as van Genuchten suggested [13] – the Mualem model was used.

3   Time-invariant analysis

In the first phase of this research FAO classes of the HUNSODA database showed significant uncertainties in soil hydraulic parameters (Kozma *et al.* [8]). Therefore the database has been categorized and analyzed also with the USDA method.
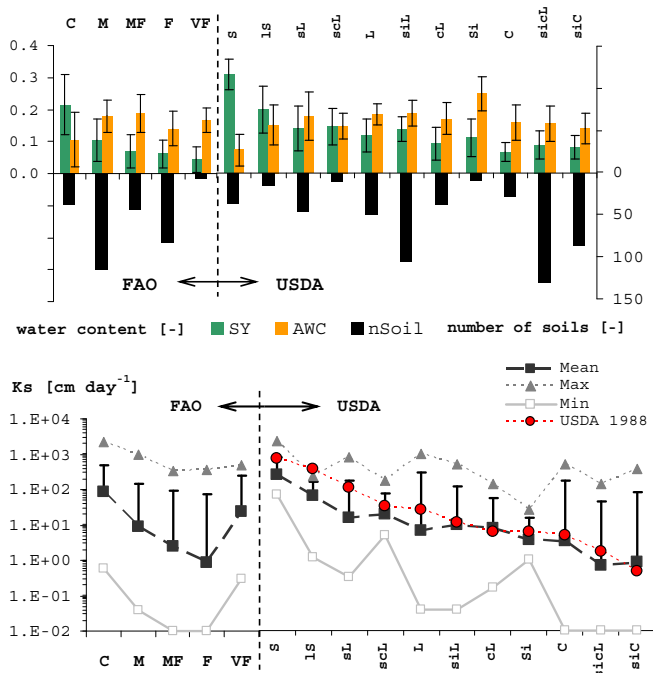

Figure 3:   Comparison of (a) FAO and (b) USDA class statistics for specific yield and available water content and for saturated hydraulic conductivity.

Figure 3 illustrates the number of soils in classes, class averages and their variation for AWC, SY and K_s values. The AWC and SY are derived from θ values at two pF level, thus they provide the best single value aggregated information about the WRC shape.

Three remarks can be taken: (i) the number of soils in certain classes is diverse. The VF class of FAO and four USDA classes contain less than 20 samples, which questions the reliability of any derived statistical measure for these classes (the sC USDA class contains no soils, so it is not included on the graph). (ii) The relative values of class variances are between 20–80% (FAO) and 15–72% (USDA) of the average, indicating major diversity of the WRC shapes. And finally (iii) variances often exceed the change between class averages – especially in case of mixed, medium textured soil types.

The saturated hydraulic conductivity varies within a wide range: values in certain classes cover 3 to 5 OMs for both methods. The only exceptions are the S, scL and Si USDA classes with only 2 OMs intervals. However scL and Si contain only 11 and 9 soils respectively, which in fact emphasizes the diversity of K_s. Large values of class variances (displayed with error bars) indicate great scatter in data as well.

At the same time, class averages show surprisingly moderate trends: in case of the FAO method 2 OMs are barely covered and the physically expected monotonous decline is biased by the known anomaly of the VF class (Nemes [10], Wösten *et al.* [6]). For the USDA method the trend is more obvious: when moving from coarser to finer textured soils, average K_s decreases over 3 OMs. Only the sicL and siC mixed classes serve as exceptions: their lower conductivity is usually attributed to their special textural composition.

The very fine FAO class contains only 13 WRCs and 6 K_s values. Such small sample demonstrates well, how uncertainties of class averaging arise. As an obvious consequence, class statistics are rather obscure. A more complex effect can also be observed: for soils with measured K_s values the WRC decreases relatively fast: average SY is 0.085, indicating that the theoretical capillary fringe would reach up only to 100-200 cm above the water table. Thus these soils – regardless of their K_s – have limited water conveying capacity under unsaturated conditions. However, for the 7 soils without K_s the lower portion of the WRC is steep, the average SY is 0.012, and the capillary fringe would extend to more than 10 meters. By averaging all 13 WRCs (determining geometrical means of θs at all pF levels), the summation curve has a SY of 0.010 (note that the resultant SY is not a linear combination of individual SY values), which would lead to great capillary rise. As a result, with its large averaged K_s value the synthetic soil representing the VF class average shows unexpected and rather unrealistic hydrologic behaviour (see results of VF class in Figure 5).

4 Dynamic numerical analysis

Beside the time-invariant analyses two transient numerical tests were applied on soils of HUNSODA database to assess how the revealed uncertainties would affect the dynamic behaviour of soils and soil classes. In the first one 1D unsaturated flow was simulated in soil columns with Hydrus-1D (Simunek *et al.* [14]), focusing mainly on the hydraulic responses of certain soils to real meteorological conditions. In the second test the WateRisk model was used in order to evaluate the hydrologic behaviour of soil types at heterogeneous regional scale. The second test is discussed under section 4.2.

Due to the climatic, topographic and hydrographical conditions of the Great Hungarian Plain, excess water and drought occur regularly, even in the same year. Therefore, assessment of soils was carried out with emphasis on these hydrologic extremities. Model input data (meteorology for the 1D analysis; topology, hydrometeorology and HUNSODA based soil map for the 3D analysis) was selected from affected lowland regions.

4.1 1D unsaturated flow test

The HUNSODA database contains 840 soil samples, of which 576 have ten-point WRC and 252 have measured K_s values. Due to lack of measured K_s or measured $\theta-\psi$ data pairs, 588 records were neglected from further assessment. Additionally, the test was applied on two other datasets as well, namely the collection of fictive soils characterized by the five-five FAO class average parameter sets derived from the (i) HYPRES and (ii) HUNSODA databases.

Atmospheric boundary condition with 100 cm allowed maximum water cover depth was set at ground surface of the soil column. Daily precipitation (P) and temperature time series recorded at Napkor weather station for the period 2010–2012 were used as input. Potential evapotranspiration PET was estimated by the robust Varga-Haszonits equation (Varga-Haszonits [15]). The evaluated time period is relevant for both excess water and drought. 2010 was the wettest year of the last century (P ≈ 900 mm and PET ≈ 740 mm) while 2011 and 2012 were extreme dry years (P ≈ 730 mm and PET ≈ 2180 mm for the two years). Vegetation with 50 cm root depth was assumed. Plant rooting distribution decreased linearly with depth from 1 to 0. For the separation of evapotranspiration to evaporation and transpiration, LAI for crops was used.

Soil columns were homogeneous and depth was set to 20 m. The possibly finest (1001 nodes), uneven discretization was used to minimize numerical problems. Constant flux (Q = 0 m^{3}day^{-1}) was used as lower boundary condition.

Considering drought, one of the most important question from agricultural and nature preservation viewpoint is the ratio of water uptake and water demand of plants. As a rather practical and simple approach, this ratio was assessed through transpiration. Among other characteristics current excess water management practice focuses mainly on spatial and temporal extent of water coverage. In case of a 1D approach only the latter can be analyzed (see the 3D analyses for

evaluation of other factors). Both extremities are affected by the status of the groundwater. Considering the above mentioned aspects, evaluation of soils was carried out through three hydrologic indicators. These are:

(a) $\Sigma T / \Sigma T_P$, where T [cm] stands for actual transpiration and T_P [cm] is potential transpiration;
(b) ExWat cover – total duration of excess water [days];
(c) Avg GW level – average groundwater level [m below/above surface].

Figure 4 illustrates the variation of the indicators as a function of time for a randomly chosen soil profile. Note that evaporation is not displayed on the figure. Results of converged simulations are summarized on Figure 5. Coloured stripes represent the ± one sigma interval while vertical black lines show the minimum and maximum values of the indicators for each class. To emphasize the trends related to classification, medians were also displayed on the figures.

The median of the $\Sigma T / \Sigma T_P$ and ExWat cover indicators denote moderate trends that meet the general expectations: the finer the soil (i) the less water it can transport to the root zone and (ii) the longer it takes to the water to infiltrate from the surface. However, considering the other statistical measures there are remarkable overlaps between the soil classes. More homogeneous is the picture in case of Avg GW level where even the median does not show significant difference between classes. Simulation results of HYPRES and HUNSODA class average parameterization show unexpected anomalies when compared with real soils.

It is to be noted that the evaluation of the results with USDA classification was also carried out. In harmony with the findings of Wösten *et al.* [6] and Nemes [10] it was concluded that even though USDA is a more detailed classification system, it holds uncertainties at least at the same magnitude.
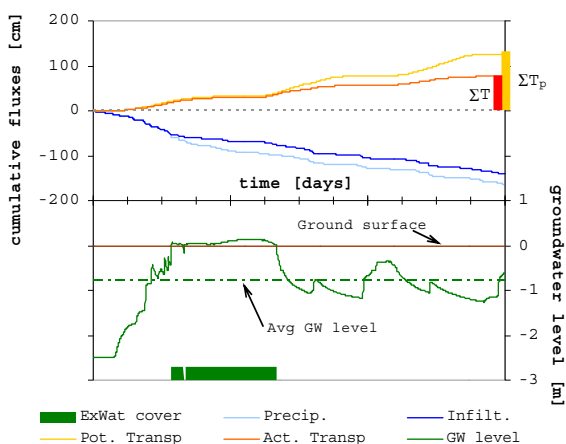

Figure 4:   Simulated boundary water fluxes and groundwater level for an arbitrary soil.
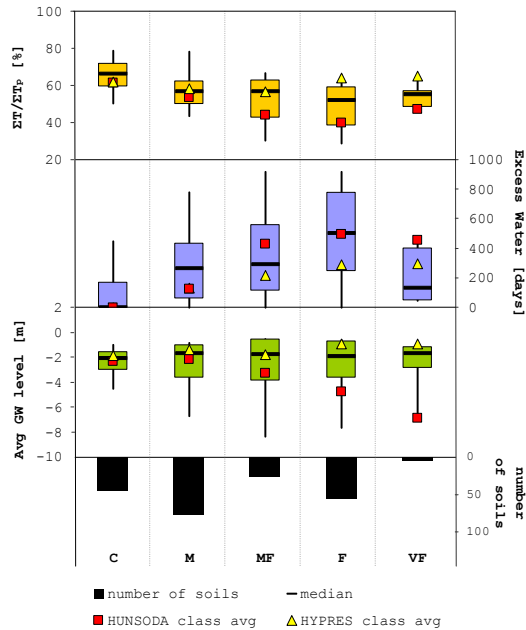
Figure 5: Summarized results for the three hydrologic indicator.

4.2 3D hydrologic sensitivity analysis

WateRisk (Kozma and Koncsos [12]) is a physically based, distributed parameter, fully coupled hydrologic model. It simulates the major processes of the hydrologic cycle, including evapotranspiration, channel and overland flow as well as unsaturated zone and shallow groundwater movement.

To analyze soil data uncertainty, hydrologic processes were simulated at the Szamos-Kraszna Interfluve, a 457 km^2 large sub-catchment in north-eastern Hungary. The area had been a marshland before it was drained off to be adapted to arable crop production. Ever since, an extremely long and costly drainage channel network is maintained on the territory to mitigate excess water. However the area suffers from this phenomenon almost every year. The heterogeneous but mostly finer soil composition of the region is a defining factor in this issue.

Sensitivity analysis (SA) was carried out for the five year period between the summers of 1995 and 2000 (average annual precipitation was 724 mm for the period). The model was previously calibrated for recent past conditions at this area (Jolánkai et al. [16]). 3D spatial soil data was based on the FAO method and the HUNSODA database (Bakacsi et al. [17]). During the SA all boundary conditions and parameters were unchanged except of porosity and specific yield values. Porosity represents total soil storage capacity, while specific yield determines mainly the intensity of groundwater table fluctuations under given boundary conditions. As physically these two properties are strongly related, their

simultaneous SA was preferable. The SA included a reference and two perturbed simulations: the reference (denoted 'mean') represented the calibrated model setup except of porosity and SY. For latter parameters the FAO class averages of the HUNSODA soils were used. For the two perturbed runs ('- sigma' and '+ sigma') class averages were reduced or increased with class variances (Table 1).

Table 1: Porosity and SY mean and variance values of FAO soil classes.

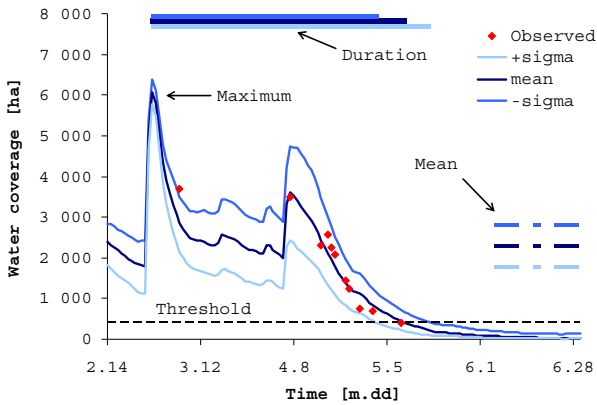| FAO class	Specific yield		Porosity	
	Mean [-]	Variance [%]	Mean [-]	Variance [%]
Coarse	0.215	44.1	0.419	15.7
Medium	0.104	62.0	0.458	14.1
Medium fine	0.069	74.1	0.477	9.9
Fine	0.062	69.9	0.486	9.9
Very fine	0.044	91.1	0.552	6.8

Figure 6:   Water coverage time series for the Szamos-Kraszna Interfluve in case of the three SA simulations.

Table 2: Sensitivity analysis results for aggregated hydrologic indicators.

indicator			- sigma	mean	+ sigma
	max	[ha]	+32 %	3596	-32 %
ExWat. coverage	mean	[ha]	+40 %	1547	-36 %
	dur.	[day]	+7	62	-8
ExWat. pumping		[1e6 m³]	+28 %	11.7	-11 %
Avr. annual ET		[mm]	-4	664	+12
GW depth	mean	[cm]	-7	158	+13
	var.	[cm]	+6	60	-7

Based on simulation results, Table 2 provides four aggregated indicators, which quantify (i) the severity of the excess water (denoted 'ExWat.') in 2000 and (ii) the hydrologic behaviour of the area for the 5 years period. These indicators also give information about model sensitivity to soil data. Excess water related indicators (maximum, mean and duration of water coverage that exceeds a certain threshold) are displayed in Figure 6.

Several conclusions can be drawn: (i) in accordance with theoretical knowledge about hydrologic processes and experiences with the WateRisk, the partly nonlinear model behaviour is reasonable, (ii) all excess water indicators behave as presumed: as soils get finer, the situation becomes more severe, (iii) as expected in case of excess water coverage, the variation of mean values exceeds the changes of maximums, (iv) groundwater response was also reasonable: depth and fluctuation correlated with porosity and SY.

5   Conclusions

The time-invariant analyses proved that both K_s values and the characteristic values of WRCs show considerable variances for each soil class. Furthermore, in harmony with the results of the 1D numerical test it was shown that in some cases variances exceed the differences between class averages. Also, it was concluded that class averaging may result in unrealistic parameter combination that poorly represents soils they stand for. Moreover, hydrologic responses of class average synthetic soils differed significantly from the majority of soils from the same class.

General consequences are: (i) texture based classification holds considerable uncertainties concerning hydraulic behaviour and (ii) it is not obvious that synthesized parameter combinations derived from soil databases appropriately characterize classes when modelling hydrological processes.

A step forward in the topic could be the establishment of a new soil classification method that focuses mainly on hydrologic characteristics of soils. However, it is questionable that such system can be built based on only generally available physical properties.

References

[1]   McCown, R.L., Hammer, G.L., Hargreaves, J.N.G., Holzworth, D.P. & Freebairn, D.M., APSIM: A novel software system for model development, model testing and simulation in agricultural systems research. *Agricultural Systems*, 50, pp. 255-271, 1996.

[2]   van Ittersum, M.K. & Rabbinge, R., Concepts in production ecology for analysis, quantification of agricultural input-output combinations. *Field Crops Research*, 52, pp. 197-208, 1997.

[3]   van Leeuwen, B., Tobak, Z. & Szatmári, J., Development of an integrated ANN-GIS framework for inland excess water monitoring. *Journal of Env. Geogr. I*, 3-4. pp. 1-6, 2008.

[4]   Harter, T. & Hopmans, J.W., Role of vadose-zone flow processes in regional-scale hydrology: review, opportunities and challenges. In: Feddes,

R.A., de Rooij, G.H. & van Dam J.C. (eds). *Unsaturated-zone Modeling*, Kluwer Academic Publishers, Netherlands, pp. 179-210, 2004.

[5]   Pachepsky, Y.A., Smettem, K.R.J., Vanderborght, J., Herbst, M., Vereecken, H. & Wösten, J.H.M., Reality and fiction of models and data in soil hydrology. In: Feddes, R.A., de Rooij, G.H. & van Dam J.C. (eds). *Unsaturated-zone Modeling*, Kluwer Academic Publishers, Netherlands, pp. 231-262, 2004.

[6]   Wösten, J.H.M., Lilly, A., Nemes, A. & C. Le Bas, Development and use of a database of hydraulic properties of European soils. *Geoderma*, 90, pp. 169-185, 1999.

[7]   Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. Available online at http://websoilsurvey.nrcs.usda.gov/. Accessed [03/23/2013].

[8]   Kozma, Zs., Ács, T. & Koncsos. L., Hydrological modelling of the unsaturated zone − evaluation of uncertainties related to the FAO soil classification system. Pollack *Periodica* (accepted), 2013.

[9]   Food and Agriculture Organisation (FAO), Guidelines for soil description. FAO/ISRIC, Rome. 1990. (3rd Ed.)

[10]   Nemes, A., Unsaturated Soil Hydraulic Database of Hungary: HUNSODA. *Agrokémia és Talajtan*, 51 (1-2), pp. 17-26, 2002.

[11]   United States Department of Agriculture (USDA). Soil Survey Manual. US Dept. Agriculture Handbook No. 18. Washington, DC. 1951.

[12]   Kozma, Zs. & Koncsos, L. Methodological Overview of a Coupled Water Resources Management Model System. *Proceedings of the Thirteenth International Conference on Civil, Structural and Environmental Engineering Computing*, B.H.V. Topping and Y. Tsompanakis, (Editors), Civil-Comp Press, Stirlingshire, United Kingdom, paper 157, 2011.

[13]   van Genuchten, M.T., A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Science Society of America Journal*, Vol. 44, pp. 892-898, 1980.

[14]   Simunek, J., Sejna, M. & van Genuchten, M.T., The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat and multiple solutes in variably-saturated media: version 2.0 IGWMC-TPS-70. International Groundwater Modeling Center, Colorado School of Miners, Golden, 1998.

[15]   Varga-Haszonits, Z. *Agro-meteorológia*. Pannon Agrártudományi Egyetem, Mezőgazdaságtudományi Kar (PATE), Mosonmagyaróvár, 1997.

[16]   Jolánkai, Zs., Kardos, M, Koncsos, L., Kozma, Zs. & Muzelák, B., Pilot Area Studies in Hungary with a Novel Integrated Hydrologic Model – WateRisk. *Proceedings of the 6th IWA International Conference for Young Water Professionals (IWA YWPC 2012)*, 10-13 July 2012, Budapest, Hungary, 2012, (electronic format on CD, id: IWA-9840).

[17]   Bakacsi, Zs., Pásztor L. & Szabó J., Compilation of a 3D soil physical database for the unsaturated zone, *19th World Congress of Soil Science, Soil Solutions for a Changing World*, Brisbane, Australia, 1-6 August 2010, Published on DVD.