



PREDICTING WATER RETENTION PROPERTIES OF DRYLAND SOILS

Muhammed Khlosi

Promoters: Prof. dr. ir. Wim M. Cornelis
UNESCO Chair of Eremology
Department of Soil Management
Faculty of Bioscience Engineering, Ghent University, Belgium

Em. Prof. dr. ir. Donald Gabriels
UNESCO Chair of Eremology
Department of Soil Management
Faculty of Bioscience Engineering, Ghent University, Belgium

Dean: Prof. dr. ir. Guido Van Huylenbroeck

Rector: Prof. dr. Anne De Paepe

Muhammed Khlosi

**PREDICTING WATER RETENTION PROPERTIES OF
DRYLAND SOILS**

Thesis submitted in fulfilment of the requirements for the degree of Doctor
(PhD) in Applied Biological Sciences: Land and Forest Management

Dedication

This thesis is dedicated to my father, Ibrahim Mustafa Khlosi, who taught me that even the largest task can be accomplished if it is done one step at a time. He inspired me to continue my study. However, he did not live long enough to see this work, which is a fruit of his long time care for me.

Acknowledgments

First of all, I would like to express my gratitude to my promoters, Prof. Dr. Ir. Donald Gabriels and Prof. Dr. Ir. Wim Cornelis, for their guidance and valuable advice. Without their generous support, I could not have carried out this work.

I would also like to thank Prof. Dr. Farouk Fares for helping in the selection of the topic of this dissertation. He was always there whenever I needed him.

This work would not have been possible without the cooperation of Prof. Dr. Ken'ichiro Kosugi (Division of Forest Science, Graduate School of Agriculture, Kyoto University, Kyoto, Japan), Prof. Dr. Gürkan Sin (Technical University of Denmark), Prof. Dr. Ir. Marc Van Meirvenne, Prof. Dr. Ir. Rien van Genuchten (Federal University of Rio de Janeiro), Ir. Bruno De Vos (The Research Institute for Nature and Forest, Ministry of the Flemish Community, Gaverstraat 4, B-9500 Geraardsbergen, Belgium), Dr. Ir. Nathalie Cools (Research Institute for Nature and Forest) and MSc. Kari A. Winfield (U.S. Geological Survey).

My special appreciation goes to Dr. Ir. Koen Verbist (UNESCO), Ir. Maarten De Boever (UGent), Dr. Ir. Liesbeth Bouckaert (UGent), Dr. Ahmed Douaik (Morocco), Dr. Ir. Jan Vermang (LNE), Mansonia Alejandra Pulido Moncada (UC Venezuela), and MSc. Ibrahim Aburas (Syria). I would like to thank all of them; they were always a source of help and practical suggestions. Thanks are further due to Patrick Dossche, Anita Lostrie, Valentijn Van Parys, Maarten Volckaert, Luc De Boosere, Marie Therese Buyens, Eric Delmulle, all from the Department of Soil Management, Ghent University.

Last, but certainly not least, my loving mother has made all the difficulties endurable through her support, sacrifices, prayers and patience. I would like also to express my love to my kids Ibrahim and Fatima.

Thanks !

List of abbreviations

AIC	Akaike Information Criterion
ANN	Artificial Neural Networks
AWC	Available Water Content
CEC	Cation Exchange Capacity
FAO	Food and Agriculture Organisation
FIM	Fisher Information Matrix
ME	Mean Error
MLR	Multiple Linear Regression
MSE	Mean Square Error
OC	Organic Carbon
O.M	Organic Matter
PCA	Principal Component Analysis
PL	Plastic Limit
PTFs	Pedotransfer Functions
RMSD	Root Mean Square Difference
RMSE	Root Mean Squared Error
SC	Soil Carbonate
SSA	Specific Surface Area
SSE	Sum of Squared Errors
SVMs	Support Vector Machines
SVR	Support Vector Regression
SWRC	Soil-Water Retention Curve
UNSODA	Unsaturated Soil Hydraulic Database

Table of contents

Dedication	i
Acknowledgments	ii
List of abbreviations	iii
Table of contents	iv
List of tables	vii
List of figures	viii
Chapter 1 General introduction	1
1.1 Problem statement	3
1.1.1 Soil water retention curve equations	6
1.1.2 Water retention PTFs	8
1.2 Study objectives	10
1.3 Outline of the dissertation	11
Chapter 2 Comparison of unimodal analytical expressions for the soil-water retention curve	15
2.1 Introduction	17
2.2 Review of some soil-water retention curve approaches	18
2.3 Materials and methods	23
2.3.1 Evaluation of the data set and soil sample analysis	23
2.3.2 Parameter estimation	26
2.3.3 Evaluation methods	27
2.4 Results and discussion	29
2.4.1 Evaluation of the models	29
2.4.2 Behaviour of the models	32
2.4.3 Effect of soil properties on the model performance	36
2.5 Conclusions	38
Chapter 3 Simple modification to describe the soil-water retention curve between saturation and oven-dryness	41
3.1 Introduction	43
3.2 Approach	44
3.3 Materials and methods	46
3.3.1 Experimental data	46
3.3.2 Model analysis	46
3.4 Results and discussion	48
3.5 Conclusions	52

Chapter 4 Performance evaluation of models that describe the soil-water retention curve between saturation and oven dryness	55
4.1 Introduction	57
4.2 Available Soil Water Retention Models	58
4.3 Materials and methods	64
4.3.1 Sources of SWRC data	64
4.3.2 Models for the data fit	66
4.3.3 Comparison methods	66
4.3.3.1 Goodness of fit statistics	66
4.3.3.2 Comparing model parameters with basic soil properties	69
4.4 Results and discussion	69
4.4.1 Evaluation of the models	69
4.4.2 Performance of the models	76
4.5 Conclusions	78
Chapter 5 Impact of particle-size distribution changes associated with carbonates on the predicted soil-water retention curve	79
5.1 Introduction	81
5.2 Materials and methods	82
5.2.1 Study area and soils	82
5.2.2 Field and laboratory measurements	85
5.2.3 Multiple linear regression analysis	86
5.2.4 Evaluation criteria	87
5.3 Results and discussion	89
5.3.1 Stepwise regression results	91
5.3.1.1 PSD methods and developing new PTF	91
5.3.1.2 PSD methods and the PTF of Vereecken et al. (1989)	96
5.3.1.3 Comparing the new PTF with the Ghorbani Dashtaki et al. (2010) PTF	98
5.4 Conclusions	100
Chapter 6 Exploration of interaction between hydraulic and physico-chemical properties of Syrian soils	103
6.1 Introduction	105
6.2 Materials and methods	106
6.2.1 Area description and soil sampling	106
6.2.2 Soil analysis	106
6.2.3 Statistical analysis	107
6.3 Results and discussion	108
6.3.1 Exploratory statistical analysis	108
6.3.2 Relationships between soil-water content and other soil properties	110
6.3.2.1 Soil texture	110
6.3.2.2 Organic matter	111
6.3.2.3 Bulk density	112
6.3.2.4 Plastic limit	114
6.3.2.5 Soil carbonate and gravel content	114
6.3.2.6 Cation exchange capacity, specific surface area and hygroscopic water content	116
6.3.3 Principal Component Analysis	117

6.3.3.1 PC loadings and communalities	117
6.3.3.2 Loading plot: PCs – soil properties relationships	120
6.3.3.3 Score plot: PCs – soil samples relationships	121
6.4 Conclusions	123
Chapter 7 Support vector machines to enhance the performance of pedotransfer functions for predicting the water retention properties of calcareous soils	125
7.1 Introduction	127
7.2 Materials and methods	128
7.2.1 Data description	128
7.2.2 Support vector regression	129
7.2.3 Artificial neural networks	131
7.2.4 Evaluation criteria	132
7.3 Results and discussion	133
7.3.1 Performance comparison of MLR, ANN and SVR	133
7.3.2 Including additional soil properties for improving PTFs accuracy	134
7.4 Conclusions	137
Chapter 8 General conclusions and future research	139
8.1 Introduction	140
8.2 General conclusions	140
8.2.1 Soil water retention curve equations	140
8.2.2 Pedotransfer functions for predicting SWRC of Syrian soils	141
8.3 Future research	143
8.3.1 Soil water retention curve equations	143
8.3.2 Pedotransfer functions for predicting SWRC	143
Summary	145
Samenvatting	148
References	151
Curriculum vitae	165

List of tables

Table 2.1 Statistical indices of the models.	30
Table 2.2 Pearson correlation coefficient between model parameters and basic soil properties.	31
Table 2.3 Average model parameters for different soil textural classes.	36
Table 2.4 Pearson correlation coefficient between SSE and basic soil properties for each model.	37
Table 3.1 Optimized Parameters, Confidence Intervals, and Correlation Coefficient between Parameters for the KCGS Model.	49
Table 4.1 Average model parameter values for different soil textural classes.	70
Table 4.2 Statistical measures of the models for 11 textural classes and all data sets.	71
Table 4.3 Pearson correlation coefficient between model parameters and basic soil properties.	75
Table 5.1 Pearson correlation coefficients between soil-water content at different matric potentials and clay, silt and sand of the two PSD methods.	91
Table 5.2 Pedotransfer function coefficients and their confidence interval, the <i>t</i> -statistic and its <i>p</i> -value.	93
Table 5.3 Comparison of the validation indices of the predicted SWRC by the PTF of Vereecken et al. (1989) using two different PSD inputs.	96
Table 5.4 Comparison of PTFs performance.	99
Table 6.1 List of soil properties used in predictive procedures.	107
Table 6.2 Descriptive statistics of soil properties used in predictive procedures.	109
Table 6.3 Correlation matrix of predictor and response variables.	113
Table 6.4 Loadings of 14 variables on significant principal components.	119
Table 6.5 Squares of loadings of 14 variables on significant PCs.	119
Table 7.1 Prediction performance comparison of MLR, ANN and SVR.	133
Table 7.2 Pedotransfer function coefficients and their confidence interval, the <i>t</i> -statistic and its <i>p</i> -value.	134
Table 7.3 Comparison of the prediction performance of ANN and SVR.	137

List of figures

Fig. 1.1 Four general categories of SWRC-PTFs.....	8
Fig. 1.2 Overview of the structure of this dissertation.	12
Fig. 2.1 Variation of clay, silt and sand content in the dataset.	25
Fig. 2.2 Variation of bulk density and organic matter content in the dataset.	26
Fig. 2.3 Observed soil-water retention curves grouped per soil textural class and VG1 model fitted to the pooled data sets.	34
Fig. 2.4 Observed and fitted soil-water retention curves for sand and silt loam.	35
Fig. 3.1 Output sensitivities of water content to the parameters in the KCGS.	51
Fig. 3.2 Comparison of measured and fitted soil-water retention curves using FS (a and c) and KCGS (b and d) models for Salkum with 95% uncertainty bands.....	52
Fig. 4.1 Texture distribution of 137 soils from UNSODA (Nemes et al., 2001) used in model development.	66
Fig. 4.2 Variation of bulk density and organic matter content in the dataset.	67
Fig. 4.3 Boxplot of (a) RMSE, (b) R-square (R^2) and (c) adjusted R-square (R^2_{adj}) values for different SWRC models.	74
Fig. 4.4 Observed and fitted soil-water retention curves for a loamy sand, silt loam and clay.....	78
Fig. 5.1 Map of study area showing the distribution of sampling sites.	85
Fig. 5.2 Variation of clay, silt and sand content in the dataset for the two PSD methods.....	91
Fig. 5.3 Effect of clay content on soil-water content at -1 and -1500 kPa.....	94
Fig. 5.4 Effect of clay content on available water content (AWC).....	96
Fig. 5.5 Effect of soil carbonates on available water content (AWC).	96
Fig. 5.6 Measured and predicted soil-water retention curves for a calcareous soil.	98
Fig. 5.7 Measured versus predicted soil-water contents at matric potential ψ of -1500 kPa for the two PSD methods.	99
Fig. 5.8 Measured versus predicted soil-water contents at matric potential ψ of -1, -33, -100, and -1500 kPa.	101
Fig. 6.1 Boxplot of soil hydraulic properties from the four zones.....	111
Fig. 6.2 Effect of plastic limit on soil-water content at -1, -33, -100, and -1500 kPa.	116
Fig. 6.3 The scree plot for the principal component analysis.....	119
Fig. 6.4 Loading Plot: soil properties in the PC1-PC2 plan.....	121
Fig. 6.5 Score Plot: soil samples in the PC1-PC2 plan.....	122
Fig. 7.1 Architecture of artificial neural network.....	133
Fig. 7.2 Measured and predicted soil-water retention curves for a calcareous soil.	137

Chapter 1 General introduction

1.1 Problem statement

There are many debates concerning the quality and quantity of natural resources required for sustaining human life (Pretty *et al.*, 2010; Zhao and Running, 2010). In dry areas, water resources are limited and the share available to agriculture is decreasing at a time when more food is needed (Oweis, 1997). Furthermore, food security is one of the global concerns in the twenty-first century (Godfray *et al.*, 2010). According to the 2010 revision of the official United Nations population estimates and projections, the world population is projected to reach 9.3 billion persons by 2050, that is, 2.3 billion more than in 2011, an increase close to the combined populations of China and India today. Most of this growth will be absorbed by developing countries (UN, 2010).

The necessary increase in food production has to be supplied from higher yields through growing more crops and cultivating more land. At the same time, production increases from fertile lands are known to be declining, forcing people to use also marginal lands. Thus, both marginal and fertile lands are currently suffering from various forms of degradation, including nutrient depletion, soil erosion, soil and water pollution and reduced soil water retention. The shortage of soil water endangers both human life and natural environment. A century ago, this list would have been quite short, because human use of the planet's resources was much less, and not perceived as damaging (Rost *et al.*, 2008; Lambin and Meyfroidt, 2011)

Soils are particularly important natural resources in Syria. This appears from the fact that the national gross income depends on agricultural production, as more than 65% of the Syrian population relies on agriculture and land exploitation to earn their living (Al-Khaier, 2003). Earlier studies showed that while agricultural production, human use and industry require an increasing amount of water, water reserves that can be devoted to these do not increase sustainably (Ragab and Prudhomme, 2002). Due to its climatic regime, Syria faces a crisis of water resources scarcity. Moreover, the overgrowing population and the droughts in the country are putting water resources under pressure and compelling people to use poor-quality water resources for irrigation (Hazzouri and Khlosi, 1998; Möller *et al.*, 2005). Until any

alternative solution is discovered, optimizing water-use efficiency - and thus using soil water as well - may be the tool to resolve the problem of the difference between *needs* and *resources* of water (Gregory *et al.*, 2000; Oweis *et al.*, 2004). To accomplish that, we need to have proper understanding of soil hydraulic properties and their incorporation in cropwater and hydrological models. Following are some examples that illustrate the complicated situation soil water management experts and farmers have to face in Syria during a growing season.

Syria has a Mediterranean climate with rainy winters and hot rainless summers. Moving south-eastwards, the climate is becoming more arid (precipitation becomes less than 200 mm) and vegetation is becoming less abundant. However, even that amount of precipitation is not evenly distributed, neither spatially nor temporally. The country therefore has a high degree of aridity in large parts of its territories and is therefore highly vulnerable to drought (De Pauw *et al.*, 2000). In the summer period, the amount of precipitation is mostly zero, which along with high temperatures and a high proportion of sunny days results in plants being stressed by drought. The irrigation potential is limited by many factors, among others the limited amount of suitable irrigation water (Hazzouri and Khlosi, 1998; Möller *et al.*, 2005). Moreover, dry conditions, heat and the generally sparse vegetation cover may cause wind erosion in the marginal drylands. Results of four seasons of wind erosion research in Khanasser Valley indicated that wind erosion is a serious problem in sparsely covered dryland environments (Massri *et al.*, 2002). Since there are few feasible options to manage this situation, the main focus should be on prevention and preparation for this situation. In early spring, the problem is the opposite. As winter precipitation is much higher, rainwater is collected and stored in and on top of the soil. In this case the amount of water that exceeds the infiltration rate and the water holding capacity of the soils is substantial which causes severe water erosion (Shinjo *et al.*, 2000).

Soils of a considerable part of Syria have a high carbonate content (>15%) in the surface and/or in deeper horizons due to low rainfall which does not allow leaching of carbonates where the parent material of soil is derived from calcareous sedimentary rocks such as limestone, dolomite and marl (Tavernier *et al.*, 1981). Therefore, with calcareous parent material, calcification and movement of carbonate within the soil profile is the most important pedogenic process in arid and semi arid regions. Moreover, about 40% of all arable land is too saline to sustain plant growth (IAEA, 2003). The salinization processes have been

remarkably accelerated due to the introduction of rice cultivation. It has been estimated that in Syria, every year about 3,000 to 5,000 ha of irrigated lands became unsuitable for agricultural use due to extreme salinization (THF, 1994). The threat of recent salinization processes is of major concern and should be studied carefully to have a better understanding of environmental degradation and to avoid undesirable changes.

Regarding environmental concerns, the shortage of water may also cause different hazards. Soil pollution of agricultural areas surrounding big cities (such as Damascus and Aleppo) is a major environmental problem. Based on the water shortage, remarkable parts of these areas are irrigated with treated but also untreated sewage effluents from big cities. Möller *et al.* (2005) assessed the present degree and spatial distribution of heavy metal concentrations in 51 soil profiles and in 22 topsoil samples in the Damascus Ghouta. Direct ingestion of soil, e.g., by children and inhalation of dust may contribute largely to the accumulation of heavy metal in human bodies and livestock.

Avoiding or preparing to any of the aforementioned situations requires among other things, the generation of a number of hydrological models that rely on numerical techniques to simulate heat, water, and solute fluxes in the vadose zone and which can be used to simulate situations ranging from field-scale water flow to global climate change. Dynamic simulation of soil physical, chemical and biological processes can be used to predict the soil-water regime for irrigation purposes, movement of chemicals for environmental monitoring and crop yield for agronomic management. A large number of transient flow and transport models have been developed to simulate integrated effects of climate, soil, and plants. The Food and Agriculture Organization of the United Nations (FAO) has developed AquaCrop, a field-crop-water-productivity simulation model for use as a decision-support tool in planning and analysis (Steduto *et al.*, 2009; Hsiao *et al.*, 2009). This model simulates attainable yields of major herbaceous crops as a function of water consumption under rainfed, supplemental, deficit, and full irrigation conditions. Another model here is HYDRUS (Šimůnek *et al.*, 2006) which is a physically based mechanistic model that solves the Richards equation for water flow and a convection–dispersion equation for solute transport (Šimůnek *et al.*, 2008). This model has been successfully used during the last few years to simulate the transport of soil water, salts, nitrates, microorganisms, and organic contaminants in variably-saturated soil formations in a variety of soil geometries and irrigation systems. Such an understanding can aid the development of best management practices, e.g. for optimizing the design of water

harvesting techniques (Verbist et al., 2012). Developments in computer modelling of water and solute transport in soil are advancing rapidly, as speed of computation increases and complexity of models expands. However, the use of these models in low-income countries such as Syria is limited, because they need many soil property values as input.

To formulate soil-water relationships, soil hydraulic properties are required as essential inputs. The most important properties are the soil-water retention curve (SWRC) and hydraulic conductivity characteristics. Measuring hydraulic properties in the soil is difficult, time-consuming and expensive. The cost-effectiveness of obtaining soil hydraulic properties can be improved by using indirect methods, which pertain to the prediction of hydraulic properties from more easily measured procedures. Furthermore, the use of data with zero error is not necessary because many problems do not require exact solutions. It is, however, better to use much data (spatially distributed) with a high degree of accuracy rather than have only data from a few points sources. Indirect methods for estimating soil hydraulic properties are based on deriving the hydraulic properties from more easily, widely available, routinely, or cheaply measured properties. This concept lies behind the development and use of *pedotransfer functions* (PTFs), a term which was first coined by Bouma (1989). Pedotransfer functions can serve as a useful means of parameterizing complex models, providing that the level of accuracy achieved is adequate in a functional sense, and that the range of applicability of the functional relationships is known and respected (Wösten et al., 1990; Mayr and Jarvis, 1999). Pedotransfer functions are particularly useful for catchment and regional scale applications of models since the availability of measured hydraulic properties is inevitably limited across large areas. However, the capability to derive pedotransfer functions requires the establishment and use of a comprehensive database of soil hydrological and pedological data. In dryland areas, such as Syria, the availability of reliable data for water retention in relation to soil type, texture, and soil carbonate content is low. It is therefore desirable to explore the interaction between soil hydraulic properties and other physical and chemical properties in order to estimate the soil-water retention curve (SWRC) from easily measured soil parameters.

1.1.1 Soil water retention curve equations

The SWRC describes the relationship between its matric potential ψ and its water content θ (i.e., it describes the soil's ability to store or release water). It is an important soil property

related to the distribution of pore space (sizes, inter connectedness), which is strongly affected by texture and structure, as well as related factors including organic matter content. Water content and the potential energy of soil water are not uniquely related because the amount of water present at a given matric potential is dependent on the pore-size distribution and the properties of air-water-solid interfaces. It further depends on whether a soil experiences a drying or a wetting process. This latter phenomenon is known as hysteresis, and is more evident in soils such as sands having a large proportion of larger pores.

Discrete (θ , ψ) data sets can be either obtained from laboratory or field measurements, or predicted from other soil properties using pedotransfer functions, PTFs, or other approaches. Both methods yield discontinuous sets of θ - ψ data pairs within the range of matric potentials used for the measurements. For modeling purposes a continuous and smooth representation of the SWRC is preferred, which can be obtained by fitting a closed-form analytical expression to a discrete data set. To date, a large number of functions have been developed to describe the SWRC. The most commonly employed classical retention models are the two unimodal functions presented by Brooks and Corey (1964) and by van Genuchten (1980). However, these models are successful in the wet part of the SWRC where water is mainly held by capillary forces and are known to give poor results at low water contents where isothermal liquid film flow induced by adsorption forces dominates (Tuller and Or, 2001). Several investigators have reported difficulty in describing the dry end of the SWRC using these models. The residual soil water content, being the water content when the matric potential goes to infinity, has been found to be an ill-defined parameter, with values often becoming negative during optimization unless special precautions are taken. The dry part of the SWRC, however, is equally important in a number of water-related processes affected by water contents well below the residual value, such as deflation of soil particles by wind, the infiltration process into a soil with initial water content below the water residual content value, the evaporation process with a dry layer of soil at the ground surface, the soil desiccation process with a heat source or dry gas flow, microbial activity and N mineralization in soils, methane oxidation in soils, and applications in colloid science. There is hence a pressing need to accurately represent the SWRC for all matric potentials. In the last years and decades, several attempts have been made to account for this fact (Ross et al., 1991; Campbell and Shiozawa, 1992; Fredlund and Xing, 1994; Fayer and Simmons, 1995; Groenevelt and Grant, 2004; Lebeau and Konrad, 2010; Zhang, 2011; Peters, 2013).

1.1.2 Water retention PTFs

Many attempts have been made to indirectly estimate the SWRC from more easily measured and readily available data such as soil texture and bulk density. These pedotransfer functions have been categorized by Botula (2013) as follows (Figure 1.1):

(1) class PTFs (Wösten et al., 1999; Al Majou et al., 2008) and continuous PTFs (Rawls and Brakensiek, 1985; Schaap et al., 1998);

(2) point-based PTFs (Gupta and Larson, 1979; Ghorbani Dashtaki et al., 2010), parameter-based PTFs (Vereecken et al., 1989; Minasny et al., 1999), and pseudo-continuous PTFs (Haghverdi et al., 2012);

(3) PTFs based on a specific approach such as semi-physical approach (Arya and Paris, 1981; Nasta et al., 2009) and empirical approach (Vereecken et al., 1989; Lamorski et al., 2008; Nemes et al., 2009);

(4) equation-based PTFs and pattern-recognition PTFs (Twarakavi et al., 2009).

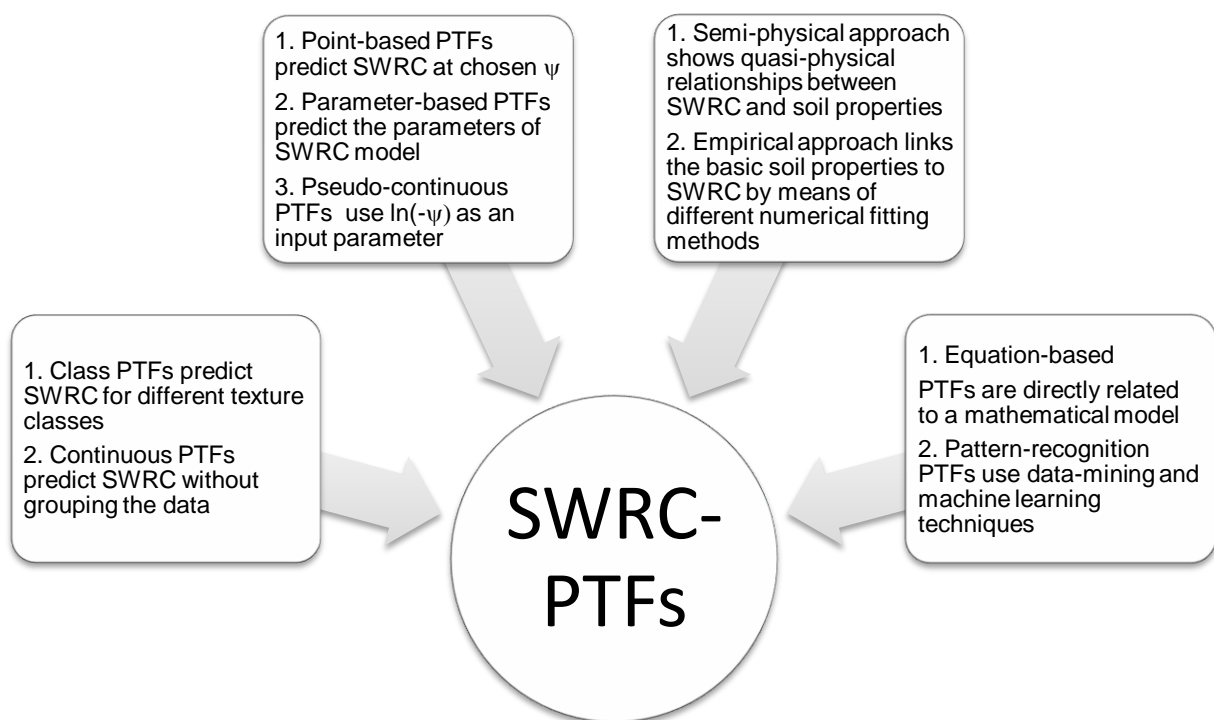


Figure 1.1. Four general categories of SWRC-PTFs

In a number of countries, different sizes of databases on soil hydraulic properties have been established and their analysis has resulted in a number of different pedotransfer functions (PTFs). A database dependency was found, so the application of these functions was reliable only for the area that the data source database represents (Schaap and Leij, 1998). The above may recommend that each interested country or region has to establish its own database to be able to develop reliable PTFs. Comprehensive soil hydraulic properties databases have long been developed in the US (Holtan et al., 1968; Rawls et al., 1982). The European Union has established its own data base HYPRES (Wösten et al., 1999) which was revisited in 2013 through the European Soil Hydropedological Data Inventory (EU-HYDI) project (Weynants et al., 2013). More general computerized databases such as the Unsaturated Soil Hydraulic Database (UNSODA) (Leij *et al.*, 1996; Nemes et al., 1999) and Grenoble Catalogue of Soils (GRIZZLY) (Haverkamp et al., 1997) have been released publicly. Most of the UNSODA and GRIZZLY soils came either from Europe or from North America. The International Soil Reference and Information Centre (ISRIC) has also made available different versions of a global soil profile database that was developed in the framework of a project entitled “World Inventory of Soil Emission Potentials” (WISE, Batjes, 2002a; Batjes, 2002b; Batjes, 2009). Botula et al. (2013) utilized a selected subset (534 soil samples) of IGBP-DIS international database from ISRIC (Tempel et al., 1996) to predict the water retention of soils from the Lower Congo in Central Africa. Other examples of data collections of soil hydraulic properties in particular countries include the databases of Australia (McKenzie et al., 2008), Belgium (Vereecken et al., 1989; Cornelis et al., 2001), Brazil (Tomasella et al., 2000), Hungary (Nemes, 2002), the Netherlands (Wösten et al., 2001). The research group “Soil Physics” of the Department of Soil Management (Ghent University) is currently establishing a large dataset of soils with water retention data from almost 1000 soil horizons distributed over the tropics, including data from countries like Chile, Cuba, D.R. Congo, Ethiopia, Kenya, Tanzania and Vietnam (W. Cornelis, personal communication; see also Botula et al., 2013; Phuong et al., 2014).

In fact, many PTFs have been developed to estimate SWRC using data from soils of temperate regions. However, a few studies have developed PTFs for prediction of SWRC of humid tropical soils and particularly of calcareous soils, and this is probably due to the relatively small amount of published data for such soils. Schaap (2004) noticed that current international databases have a serious bias towards soils from temperate regions. Tomasella

and Hodnett (1998) showed that, in many cases, the textures of many tropical soils, particularly oxisols such as those of Brazilian Amazonia, are outside the range of validity of these PTFs.

The traditional statistical techniques commonly used to develop PTFs are multiple linear regressions (MLR) and artificial neural network (ANN). Both methods are widely applied since their methodology has been established very well. Recently, support vector machines (SVMs) have gained popularity in many traditionally ANN-dominated fields due to their ability to tackle complex, highly nonlinear problems in a consistent, structured manner, while simultaneously avoiding problems of over-fitting on simpler problems (Lamorski et al., 2008; Twarakavi et al., 2009).

1.2 Study objectives

Presently in Syria, there are no published soil hydraulic data available and the collation of a national data base is not complete. The principal objective of this research was therefore to statistically explore the relationships of soil hydraulic properties to physical and chemical properties of dryland soils in order to identify which soil properties are most relevant for deriving pedotransfer functions for such areas. The primary focus was to understand how hydraulic properties of a large set of soils from a dryland area are related to the soil properties other than texture (such as organic matter content, particle density, CEC, pH, lime content, plasticity index, and specific surface area). Disturbed and undisturbed soil samples have been collected in Syria in 2005 from 18 profiles including 72 horizons. In selecting sampling sites, major emphasis was placed on covering all the agro-climatic zones of Syria (five zones). However, for logistical reasons, one zone was excluded in this study.

The specific objectives can be summarized as follows:

1. To develop a new model that describes the soil-water retention curve from saturation to oven-dryness, enabling to accurately describe the soil-water retention curve at very low saturations which are often encountered in dryland regions, and to represent a realistic fit in cases where data in the dry range are missing.

2. To investigate the influence of selected pretreatment procedures on soil texture and soil water retention curve of dryland soils.
3. To introduce additional predictor variables which are easily and cheaply determined.
4. To develop pedotransfer functions for predicting soil-water retention curve of dryland soils.

1.3 Outline of the dissertation

The dissertation is structured around the above described specific objectives. Chapters 2, 3 and 4 are devoted to preliminary testing and developing of closed-form analytical expressions to describe the water retention curve. Because of practical reasons, this preliminary testing was performed using international datasets from Belgium and USA. Chapters 5, 6 and 7 go into the development of the PTFs for dryland soils using data collected in Syria. A schematic overview of the structure of the dissertation, with indication of the overall and specific research questions, which are addressed in the various chapters, is presented in Figure 1.2.

Chapter 2 evaluates ten closed-form unimodal analytical expressions to describe the soil-water retention curve, in terms of their accuracy, linearity, Aikake Information Criterion (AIC), parameter uniqueness and parameter identifiability. This preliminary study was carried out on soil samples taken in duplicate from 48 horizons of 24 soil series in Flanders, Belgium.

Chapter 3 presents an alternative closed-form analytical expression that describes the soil-water retention curve over the complete range of soil-water contents (from saturation to oven dryness), which is of particular interest for dryland conditions. The model combines the best expression from Chapter 2 with the adsorption equation developed by Campbell & Shiozawa (1992). The model is tested against data taken from literature that cover the complete range of water contents, from saturation to almost oven-dryness. It was also tested to what extent the model represents a realistic fit in cases where data from the dry end of the water retention curve are not available.

Chapter 4 compares the new developed model with seven other closed-form unimodal analytical expressions that describe the soil-water retention curve across the complete range

of soil water contents. Retention data for 137 undisturbed soils from the Unsaturated Soil Hydraulic Database (UNSODA) were used for the model comparison. The eight models were compared in terms of their accuracy, linearity and prediction potential.

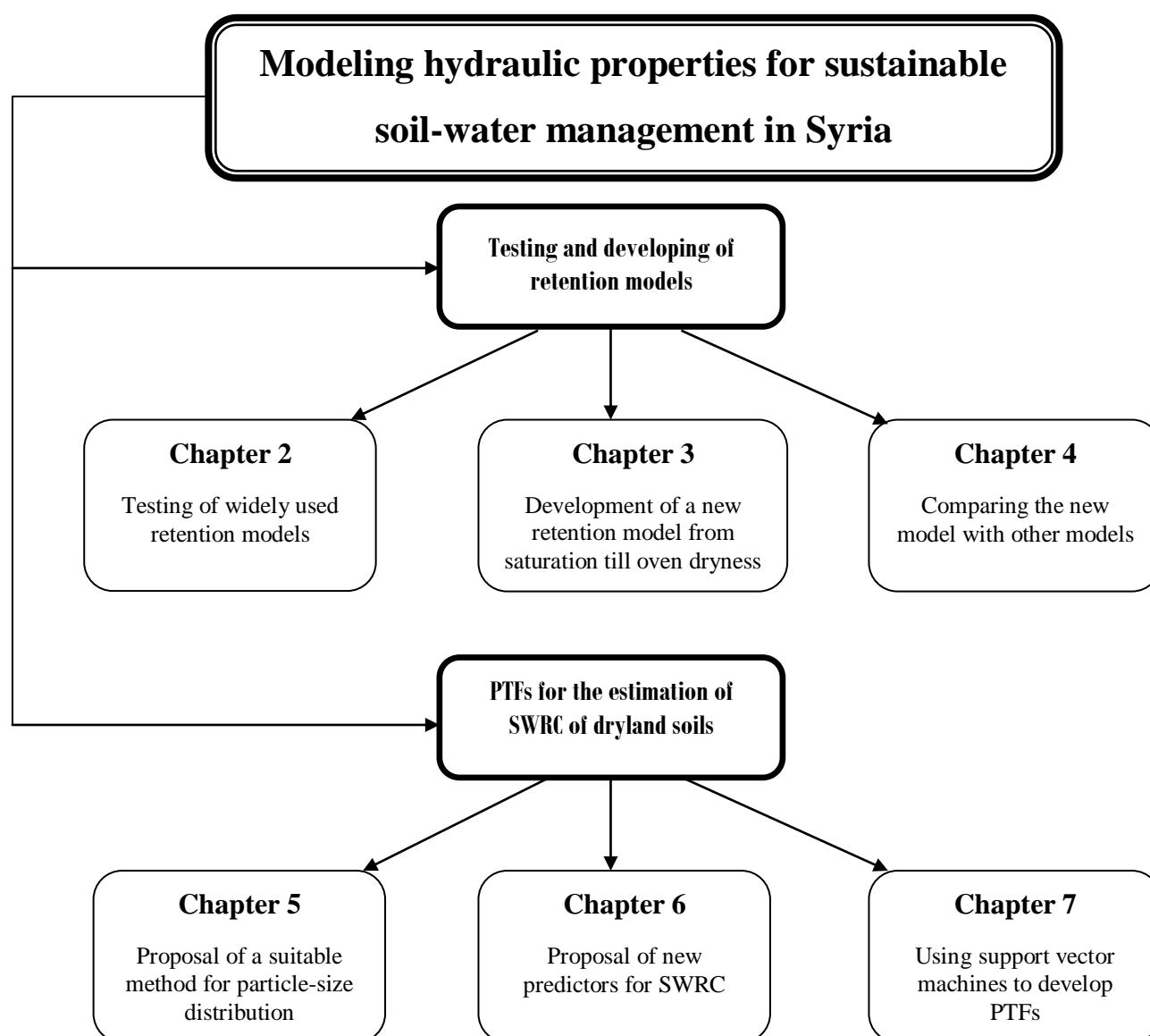


Figure 1.2. Overview of the structure of this dissertation

Chapter 5 investigates the influence of sample pre-treatment on sand, silt, and clay fractions (texture) and hence on the predictability of the soil-water retention curve of dryland soils. 72 soil samples have been collected from Syria. Two procedures with different pre-treatment

were applied. In a first one, carbonates were removed by hydrochloric acid as is typically done in particle size analysis worldwide. In the alternative one, carbonates were not removed, a procedure more typical for calcareous soils.

Chapter 6 discusses the interaction between soil hydraulic properties and other physical and chemical properties of selected dryland in Syria. The approach in this chapter is to investigate the possible use of new basic soil properties as predictors of soil hydraulic properties, which can be easily or cheaply measured.

Chapter 7 utilizes the support vector machine technique to enhance the performance of pedotransfer functions for predicting the water retention properties. The accuracy of this technique is compared with artificial neural networks (ANN) and multiple linear regression (MLR) in predicting the soil water retention curve of dryland soils.

Finally chapter 8 summarizes the main findings and major conclusions of the thesis and gives some recommendations for future research.

Chapter 2 Comparison of unimodal analytical expressions for the soil-water retention curve

Based on: Cornelis W.M., Khlosi M., Hartmann R., Van Meirvenne M. and De Vos B. (2005). Comparison of Unimodal Analytical Expressions for the Soil-Water Retention Curve. *Soil Science Society of America Journal* 69:1902–1911.

Abstract

This study was conducted to evaluate ten closed-form unimodal analytical expressions to describe the soil-water retention curve, in terms of their accuracy, linearity, Akaike Information Criterion (AIC), and prediction potential. The latter was evaluated by correlating the model parameters to basic soil properties. Soil samples were taken in duplicate from 48 horizons of 24 soil series in Flanders, Belgium. All sample locations were under forest and hence the samples had, besides their difference in texture, a high variety in bulk density (ρ_b) and organic matter content (OM). The van Genuchten model with m as a free parameter showed the highest overall performance in terms of goodness-of-fit. It had the highest accuracy, the highest degree of linearity, and the lowest AIC value. However, it had a low prediction potential. Imposing the constraint $m = 1 - 1/n$ and hence reducing the number of model parameters by one, increased the prediction potential of the model significantly, without losing much of the model's accuracy and linearity. A high degree of accuracy and linearity was also observed for the two Kosugi models tested. Restricting the bubbling pressure to be equal to zero resulted in a rather high prediction potential, which was not the case when keeping the bubbling pressure as a free parameter. A major drawback of van Genuchten and Kosugi type models is that they do not define the soil-water retention curve beyond the residual water content. We further demonstrated that the performance of all but one model in terms of their match to the data increased with increasing clay content and decreasing sand content, which is contradictory to the deterministic character of these models. Bulk density and OM did not have a significant effect on the accuracy of most models.

2.1 Introduction

Water relations are among the most important physical phenomena that affect the use of soils for agricultural, ecological, environmental, and engineering purposes. To formulate soil-water relationships, soil hydraulic properties are required as essential inputs. The most important hydraulic properties are the soil-water retention curve (SWRC) and the hydraulic conductivity. The SWRC describes the relationship between the soil's matric potential ψ and its water content θ .

To be useful in modeling processes depending on soil-water relationships, a continuous representation of the SWRC is required and needs to be incorporated in predictive models. One of the most manifest examples is the use of the SWRC to indirectly determine the unsaturated hydraulic conductivity, using statistical pore-size distribution models (see e.g., Mualem, 1986, for a review), which are well represented by the SWRC. Such SWRCs can be obtained by fitting closed-form analytical expressions containing several parameters to discrete (θ, ψ) data sets, which can be obtained through laboratory experiments or from pedotransfer functions (PTFs) that estimate distinct SWRC data pairs. The most widely adopted and best-performing PTFs enable, however, to directly predict the parameters of some closed-form analytical expressions (Cornelis et al., 2001). Applications of closed-form analytical expressions can also be attractive for other reasons, apart from their incorporation in predictive models. Van Genuchten et al. (1991) mention their applicability in more efficiently representing and comparing hydraulic properties of different soils and soil horizons, in scaling procedures for characterizing the spatial variability of soil hydraulic properties across the landscape and in interpolating and extrapolating to parts of the soil-water retention or hydraulic conductivity curves for which little or no data are available.

The objective of our study was to evaluate ten closed-form unimodal analytical expressions, including those reported by Brooks and Corey (1964) (BC), van Genuchten (1980) (VG1 and VG2), Tani (1982) (T), Russo (1988) (R), Rossi and Nimmo (1994) (RN), Kosugi

(1994)([K1], 1996, 1997 [K2]) and Assouline et al. (1998) (A1 and A2), in terms of their accuracy, linearity, AIC, and prediction potential. These models were retained in this study because they are widely adopted and cited, and because of their relative simplicity, which is needed to be easily incorporated into predictive pore-size distribution models for the hydraulic conductivity.

2.2 Review of some soil-water retention curve approaches

Many functions to represent the SWRC have been proposed for modeling purposes. In this section, we will try to give an overview of the different expressions for the SWRC that have been reported in literature, with special attention to the expressions that are evaluated in this study. However, this review does not pretend to be complete and focuses only on unimodal expressions. One of the first expressions for the SWRC was the still widely used four-parameter power function presented by Brooks and Corey (1964) (BC model):

$$\begin{aligned} \theta &= \theta_s && \text{for } \psi \geq \psi_b \\ \theta &= \theta_r + (\theta_s - \theta_r) \left(\frac{\psi_b}{\psi} \right)^\lambda && \text{for } \psi < \psi_b \end{aligned} \quad [2.1]$$

where θ_s and θ_r are the soil-water content at saturation and the residual soil-water content respectively, ψ_b the bubbling pressure or air-entry value, and λ is a pore-size distribution factor affecting the slope of the curve. The residual water content has been generally defined as the water content at which water movement ceases (Nitao and Bear, 1996), as the air-dry water content (Shao, 2000), as the water content close to the permanent wilting point of most plants, that is, at $\psi = -1.5$ MPa (van Genuchten, 1980), or simply as a fitting parameter equal to the water content where the differential soil-water capacity $d\theta/d\psi$ becomes zero (van Genuchten and Nielsen, 1985). The parameter ψ_b is assumed to be related to the maximum size of the pores forming a continuous network of flow channels within a soil. The discontinuous character of Eq. [2.1] is generally considered as a disadvantage, particularly in describing the SWRC near saturation (van Genuchten and Nielsen, 1985). Nevertheless, Eq. [2.1] is historically one of the most-widely used functions by soil scientists, hydrologists, and engineers.

Brutsaert (1966) evaluated several distribution functions to describe the soil's pore-size distribution, which can then be converted to a SWRC using the Young-Laplace equation. Ahuja and Swartzendruber (1972) inserted the power form of the hydraulic conductivity-soil-water content function suggested by Brooks and Corey (1964) and Brutsaert (1967)(1968) into the basic form of the diffusivity function (Bruce and Klute, 1956) to obtain their SWRC. Campbell (1974) presented a SWRC similar to Eq. [2.1], but with $\theta_r = 0$. Clapp and Hornberger (1978) and Hutson and Cass (1987) suggested replacing the sharp corner of Eq. [2.1] with a parabolic curve, leading to a smoothly joined two-part SWRC. Other expressions that are often cited are those presented by Visser (1966), Laliberte (1969), Gardner et al. (1970), White et al. (1970), and Su and Brooks (1975).

The most-widely adopted alternative for the BC model is the expression introduced by van Genuchten (1980). Originally, the model contained five parameters:

$$\theta = \theta_r + (\theta_s - \theta_r) \left(\frac{1}{1 + (\alpha |\psi|)^n} \right)^m \quad [2.2]$$

where α , and n and m are parameters respectively related to ψ^{-1} and the curve's slope at its inflection point. These parameters all depend on the pore-size distribution as was the case with the parameters ψ_b and λ in the BC model. Although van Genuchten and Nielsen (1985) found the five-parameter form of Eq. [2.2] superior to the four-parameter form with $m = 1 - 1/n$, the latter form might be recommended when only a limited range of retention data (usually in the wet range) is available, since keeping both n and m independent may lead to uniqueness problems in the parameter estimation process and consequently a less accurate description of the SWRC in the dry range (van Genuchten et al., 1991). In our study, both the five-parameter form and the four-parameter form, with $m = 1 - 1/n$, of Eq. [2.2] will be evaluated and are denoted as VG1 and VG2, respectively. Compared with the BC model, the van Genuchten (1980) model has a continuous character due to its inflection point. Note that Eq. [2.2] with $m = 1$ was earlier used by Ahuja and Swartzendruber (1972), Endelman et al. (1974), and Varallyay and Mironenko (1979).

Tani (1982) (T model) introduced a three-parameter expression:

$$\theta = \theta_r + (\theta_s - \theta_r) \left(1 + \frac{\psi}{\psi_{ip}} \right) \exp \left(- \frac{\psi}{\psi_{ip}} \right) \quad [2.3]$$

where ψ_{ip} is the soil-water potential at the inflection point. Because of its simplicity, this model has been widely used for modeling water movement in soils (e.g., Suzuki, 1984).

Russo (1988) proposed a four-parameter model (R model), which produces Gardner's (1958) exponential model for the water conductivity-capillary potential relationship when incorporated into Mualem's (1976) model for the relative hydraulic conductivity:

$$\theta = \theta_r + (\theta_s - \theta_r) \left[(1 + 0.5 \alpha' |\psi|) \exp(-0.5 \alpha' |\psi|) \right]^{2/(m'+2)} \quad [2.4]$$

where m' is a parameter which accounts for the dependence of the tortuosity and the correlation factors on the water content, and α' is related to the width of the pore-size distribution. Note that m' corresponds to the shape factor in Mualem's expression for the relative hydraulic conductivity and α' to the slope of Gardner's (1958) exponential equation. The reciprocal of α' can be interpreted as the air-entry value. Equation [2.4] is further similar to Tani's (1982) expression where $2/(m'+2) = 1$ and $0.5 \alpha' = 1/\psi_{ip}$.

Ross et al. (1991) modified Campbell's equation (1974) to force the SWRC to predict zero soil-water content at oven dryness. Combining the Ross et al. (1991) correction, which includes the Campbell model (1974) with the parabolic correction near saturation proposed by Hutson and Cass (1987), Rossi and Nimmo (1994) developed a four-parameter sum model and a three-parameter junction model that covers the entire range from saturation to oven-dryness. Their sum model (RN model), which originally included seven parameters and showed a higher accuracy in their study compared to their junction model, was written as:

$$\begin{aligned} \theta = \theta_s \theta_I = \theta_s \left[1 - \beta \left(\frac{\psi}{\psi_b} \right)^2 \right] & \quad \text{for } 0 < \psi < \psi_i \\ \theta = \theta_s \theta_{II} = \theta_s \left[\left(\frac{\psi_b}{\psi} \right)^\lambda - \left(\frac{\psi_b}{\psi_o} \right)^\lambda + \gamma \ln \left(\frac{\psi_o}{\psi} \right) \right] & \quad \text{for } \psi_i < \psi < \psi_o \end{aligned} \quad [2.5]$$

where ψ_i is the soil-matric potential at the junction point where the two curves join, ψ_o is the soil-matric potential at oven dryness, and β and γ are shape parameters. The term θ_I represents the Hutson and Cass (1987) parabolic curve that joints the Campbell function

(1974) at the junction point ψ_i . The Ross et al. (1991) correction is included in the expression for θ_{II} . Further, using data sets from Schofield (1935) and Campbell and Shiozawa (1992), Rossi and Nimmo (1994) showed that at very low soil-water content, the latter becomes proportional to the logarithm of the soil-matric potential, as can be recognized as well in θ_{II} . Equation [2.5] contains seven parameters. However, two of them can be determined by the conditions that ensure the continuity of both Eq. [2.5] and its first derivative to ψ_i . Here we have chosen to explicitly determine β and γ as analytical functions of ψ_b , ψ_i , ψ_o and λ :

$$\beta = \psi_b^2 \frac{1 - \left(\frac{\psi_b}{\psi_i}\right)^\lambda + \left(\frac{\psi_b}{\psi_o}\right)^\lambda + \ln\left(\frac{\psi_o}{\psi_i}\right) \left(\frac{\psi_b}{\psi_i}\right)^\lambda}{\left[1 + 2 \ln\left(\frac{\psi_o}{\psi_i}\right)\right] \psi_i^2} \lambda \quad [2.6]$$

$$\gamma = \frac{2 - 2\left(\frac{\psi_b}{\psi_i}\right)^\lambda + 2\left(\frac{\psi_b}{\psi_o}\right)^\lambda - \left(\frac{\psi_b}{\psi_i}\right)^\lambda}{1 + 2 \ln\left(\frac{\psi_o}{\psi_i}\right)} \lambda \quad [2.7]$$

Since the measured water contents are usually based on oven drying, it is more appropriate to assign a finite matric potential, at which the water content becomes zero, to a value corresponding to oven dry conditions at 105°C. This yields a matric potential of $\approx -10^6$ kPa depending on laboratory conditions. When setting ψ_o arbitrarily at -10^6 kPa (Ross et al., 1991; Rossi and Nimmo, 1994), and with Eqs. [2.6] and [2.7], the number of model parameters can be reduced to four.

Kosugi (1994) proposed a five-parameter expression (K1 model) that resulted from applying three-parameter lognormal distribution laws to the pore-size distribution function and to the pore capillary pressure potential distribution function. The resulting expression for the SWRC is:

$$\begin{aligned} \theta &= \theta_s && \text{for } \psi \geq \psi_b \\ \theta &= \theta_r + \frac{1}{2}(\theta_s - \theta_r) \operatorname{erfc} \left[\frac{\ln\left(\frac{\psi_b - \psi}{\psi_b - \psi_{md}}\right) - \sigma^2}{\sqrt{2} \sigma} \right] && \text{for } \psi < \psi_b \end{aligned} \quad [2.8]$$

where “erfc” denotes the complementary error function defined by:

$$\text{erfc} = 1 - \text{erf}(x) \quad [2.9]$$

in which “erf” is given by:

$$\text{erf} = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt \quad [2.10]$$

In Eq. [2.8], ψ_{md} is the matric potential corresponding to the median pore radius, and σ is the variance of the distribution of $\ln(r/r_{\text{max}} - r)$, in which r is the pore radius and r_{max} the maximum pore radius. Later, Kosugi (1996; 1997) modified Eq. [2.8] to have a relatively simpler functional form introducing the restriction that $\psi_b = 0$ (K2 model). The new expression hence becomes:

$$\theta = \theta_r + \frac{1}{2}(\theta_s - \theta_r) \text{erfc} \left[\frac{\ln\left(\frac{\psi}{\psi_{\text{md}}}\right)}{\sqrt{2}\sigma} \right] \quad [2.11]$$

Assouline et al. (1998) proposed a conceptual model, which is based on the assumption that the soil structure results from an uniform random fragmentation process where the probability of fragmentation of an aggregate is proportional to its size, and that a power function relates the volume of the aggregates to the corresponding pore volume. The fragmentation process determines the particle-size distribution of the soil. The transformation of the particle volumes into pore volumes via a power function and the adoption of the capillarity equation leads to the following expression:

$$\theta = \theta_L + (\theta_s - \theta_L) \left\{ 1 - \exp \left[\xi (\psi^{-1} - \psi_L^{-1})^\eta \right] \right\} \quad [2.12]$$

where ψ_L is the soil-matric potential limit of the domain of interest of the SWC under study corresponding to θ_L , and ξ and η are parameters depending on the packing and shape of the particles and hence on the pore-size distribution. The parameter ξ further depends on the bubbling pressure potential ψ_b . Equation [2.12] contains five parameters and will be referred to as the A1 model. However, Assouline et al. (1998) suggested reducing the number of parameters to three by choosing the value of (θ_L, ψ_L) according to the specific soil type under

consideration since “there is no need to consider the water retention curve beyond a capillary head, ψ_L , that corresponds to a very low water content, θ_L , at which the hydraulic conductivity is negligible.” It should be noted that strictly speaking in case of the Assouline et al. (1998) model, ψ is the capillary potential and hence does not include adsorption forces. This means that Eq. [2.12] is not defined in the very low matric potential range where adsorption forces become dominant, although mathematically speaking, ψ_L can tend to -10^6 kPa (or even $-\infty$) at zero water content. Since choosing an exact value of (θ_L, ψ_L) is not evident, Assouline et al. (1998) proposed to truncate ψ_L at -1.5 MPa, as was suggested by van Genuchten (1980). The latter alternative was also evaluated here with, θ_L as a free parameter. The number of parameters in Eq. [2.12] hence reduces to four (A2 model). It should further be noted that when applying the A1 or A2 model as described above, Eq. [2.12] is not defined for soil-water contents below θ_L .

The expressions described above were, although they are in fact simple curve-fitting equations, mainly based on pore-size distribution functions in combination with the bundle-of-capillaries concept, in which the pores are represented by cylindrical capillary tubes obeying the Young-Laplace equation. Recently, new theories have been developed, including a pore-scale network theory (Reeves and Celia, 1996; Fischer and Celia, 1999; Held and Celia, 2001a, 2001b), and a theory first presented by Tuller et al. (1999) in which (1) a pore is represented as being composed of an angular pore cross-section connected to slit-shaped spaces, and (2) the soil-matric potential is related not only to capillary forces, but also to adsorptive forces (see e.g., also Or and Tuller, 1999, 2002; Tuller and Or, 2001).

2.3 Materials and methods

2.3.1 Evaluation of the data set and soil sample analysis

The study was based on soil samples taken in duplicate from 48 horizons of 24 soil series in Flanders, Belgium. They were collected in the context of assessing the predictive quality and usefulness of the Belgian soil map and historical forest soil profile data for mapping purposes. All sample locations were under forest and at each location the samples were taken from Ah and E horizons down to a depth of 30 cm. The soils used in this study cover a wide

range of textures within Flanders (Figure 2.1). They were classified according to soil taxonomy (Soil Survey Staff, 2003) as Spodosols, Entisols (suborders Psamments, Fluvents, and Aquents), Alfisols, and Inceptisols.

Undisturbed soil samples were taken using the core method. A Riverside auger was used to prepare a flat sampling platform at a predetermined depth within a specific horizon after which standard sharpened steel 100-cm³ Kopecky rings (height = 5 cm, diameter = 5.3 cm) were driven into the soil using a dedicated ring holder (Eijkelkamp Agrisearch Equipment, Giesbeek, the Netherlands). In hard layers, a percussion-free hammer was applied for hammering the ring holder with a minimum of vibration into the soil. The soil-filled cylinder was carefully removed from the ring holder and the oversized sample was trimmed flush using a sharp knife. Cylinders with stones, charcoal, or roots larger than 2 mm in diameter were rejected and resampled in the same horizon. The samples were then covered with plastic lids which prevented them from drying out and transported in special carrying cases to the laboratory to minimize disturbance (De Vos et al., 2005).

The particle-size distributions were determined on disturbed samples using a Coulter LS200 laser diffractometer (Beckman Coulter, Fullerton, CA). The results were then calibrated and validated using standard pipetting and sieving procedures (ISO 11277) after application of the same pretreatment. Organic matter content ranged from 2.3 to 130.0 g kg⁻¹ and was determined by means of the Walkley and Black (1934) method. Bulk densities ρ_b varied from 0.76 to 1.78 Mg m⁻³. They were measured by weighing the 100-cm³ sized undisturbed soil samples at -10 kPa and subtracting the corresponding mass of water measured on a 25-cm³ sized subsample. The spread of both OM and ρ_b is illustrated in Figure 2.2.

The samples' SWRC was constructed by measuring soil-water content at nine soil-matric potentials using the undisturbed soil samples. For the pressure potentials ranging from -1 to -10 kPa, the sand box apparatus (Eijkelkamp Agrisearch Equipment, Giesbeek, the Netherlands) was used. Each sample, which was covered with a nylon cloth at its cutting edge, was placed on the sandbox in 1 mm of water and gently pressed downward to create a good contact between the sample and the sand. To saturate the samples by capillary rise, the water level on top of the sand was raised until 2.5 cm (halfway the sample height). Once the samples were saturated, a suction was applied by adjusting the suction regulator of the sandbox apparatus. After having reached equilibrium between the applied pressure and the

quantity of water in the sample, the samples were removed from the sandbox, weighed, placed back on the sandbox, and the suction applied on the sample was increased. After having determined the sample weight at -10 kPa, a subsample was taken, it was weighed, placed in the oven at 105°C for 24 h, and weighed again to determine the water contents at pressures between -1 and -10 kPa. This also allowed calculating ρ_b . The sandbox was thus used to determine five (θ, ψ) data pairs on one single soil sample. This sample was further divided into two undisturbed subsamples using sharpened steel 20-cm^3 cylinders and into two disturbed subsamples. The undisturbed subsamples were used to determine water content at -20 and -33 kPa and the disturbed subsamples for water content determination at -100 and -1500 kPa using pressure chambers (Soilmoisture Equipment, Santa Barbara, CA). After having obtained equilibrium between the applied pressure and the quantity of water in the sample, the samples were weighed and placed in the oven at 105°C for 24 h. Then they were weighed again and water content was calculated.

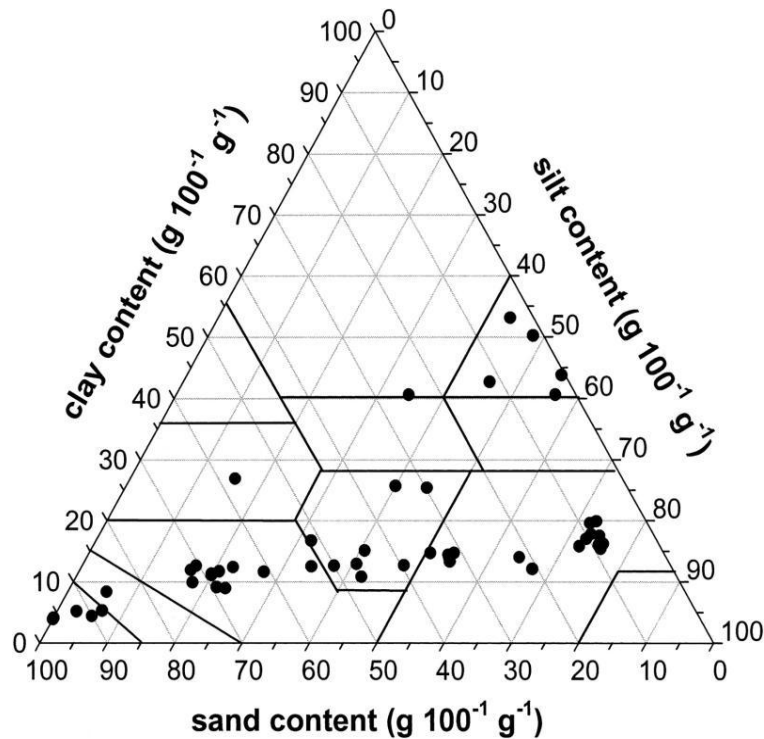


Figure 2.1. Variation of clay, silt and sand content in the dataset

our study, θ_j is the soil-water content corresponding to the j th data pair for each soil, and *obs* and *fit* denote observed and fitted values, respectively. The quasi-Newton routine was performed employing the mathematical software program MathCad (Mathsoft, Cambridge, MA). It resulted in slightly better fits, that is, lower SSE values, compared with the conjugant-gradient method (Press et al., 1992) and Levenberg-Marquardt's maximum neighborhood method modified by More et al. (1980). In selecting values for the initial estimates of the model parameters in the iterative procedure, data of fitted parameter values for different soils reported in literature were, if available, considered. When not available, routinely rerunning the program with different initial parameter estimates was performed. This should have prevented convergence of SSE in local minima in the objective function. To avoid negative θ_r or θ_L values, we introduced the constraint θ_r or $\theta_L \geq 0$, except for the RN model. The constraint $\theta_s = \theta_{-1\text{kPa}}$ was used to keep the θ_s parameter close to the near saturation value at -1 kPa, which reduces the number of parameters of each expression with one. We did not use the porosity calculated from ρ_b and particle density for this purpose, since the latter was not determined in our study. Finally, we introduced the constraint $\psi_L < -1500$ kPa in case of the A1 model, which was the lower limit of our database. Otherwise unrealistic fits were produced for those data sets where ψ_L was calculated to be larger than -1500 kPa. Furthermore, it reduced the dependency of the model to initial estimates of its parameters considerably.

2.3.3 Evaluation methods

Several statistical indices can be applied to assess the 'goodness-of-fit' of a given model. In this study, the fitting accuracy of the different models was determined by using the root of the mean of squared errors, that is, RMSE, the coefficient of determination R^2 , and the AIC, which were calculated for each soil sample.

The RMSE ($\text{m}^3 \text{ m}^{-3}$) is an indication for the overall error of the evaluated function and should approach zero for best model performance. The R^2 is a measure for the linearity between observed and fitted data. An R^2 value that approaches unity, means that the measured and fitted data pairs are linearly located around the line of perfect agreement (or 1:1 line) or that the fitted curve is of comparable shape as the measured discrete curve. The mean square error and root mean square error were derived from the SSE using:

$$\text{MSE} = \frac{1}{N} \text{SSE} \quad [2.14]$$

$$\text{RMSE} = \sqrt{\text{MSE}} \quad [2.15]$$

R^2 reflects the proportion of the total sum of squares (SST) that is partitioned into the model sum of squares (SSM) since SST is equal to SSM plus SSE:

$$R^2 = \frac{\text{SSM}}{\text{SST}} = \frac{\text{SST}-\text{SSE}}{\text{SST}} = 1 - \frac{\text{SSE}}{\text{SST}} \quad [2.16]$$

The AIC index, which is often used for model-discrimination tests (Akaike, 1974), was computed as

$$\text{AIC} = N \left[\log(2\pi) + \log\left(\frac{\text{SSE}}{N-p}\right) + 1 \right] + p \quad [2.17]$$

where p is the number of model parameters. The “best” model is the one that minimizes AIC, or in other words, which combines the lowest SSE value with the lowest number of model parameters. Although computers can nowadays easily handle models with many parameters, overparameterization should be avoided as it results in a non-identifiable model, that is, a model leading to sample configuration probabilities identical to those of a simpler model with fewer parameters, in large variances of the estimated model parameters for similar soils, and in a high degree of correlation between the parameters (or low parameter uniqueness) if the number of observations is limited as is often the case with laboratory-determined SWRCs. Further, it is advantageous to minimize the number of model parameters when attempting to predict the SWRC in terms of parameters of closed-form analytical expressions from readily available data using PTFs. To facilitate the comparison between the different expressions, the mean of RMSE, of R^2 , and of AIC was calculated for each expression.

We further computed the Pearson coefficient r_{sp} of correlation between model parameters and soil properties, including ρ_b , OM, and sand, silt, and clay content. This index was used as a measure for the prediction potential of the model, in that the higher the correlation, the higher becomes the prediction potential of the parameters in the model.

2.4 Results and discussion

2.4.1 Evaluation of the models

Table 2.1 shows the values of the statistical indices, which were computed to evaluate the ten closed-form analytical expressions. When considering the mean of RMSE, the VG1 model showed the lowest values, meaning that the fitted curve produced the highest match with the measured SWRC. The VG1 model led to the best fit in 67% of the soil samples. Second best was the K1 model, followed by the K2 model, the VG2 model, the A1 model, and the A2 model. The worst models were the RN model and the T model. Intermediate results were obtained with the R model and the BC model. As regards the mean of R^2 , a similar trend could be observed, with VG1 as the best model in terms of linearity, closely followed by K1, K2, VG2, A1, and A2. The mean of AIC again showed a similar trend. VG1 resulted in the lowest AIC value, but now followed by K2 and then K1, VG2, A2, and A1. In the case of for example, the VG1 model, the positive effect of a reduced SSE was higher than the negative effect associated with an increased number of model parameters. The fact that it has one parameter more compared with most other models did not counterbalance its high performance. Overall, the VG1 with five model parameters scored best, followed by the K1, K2, VG2, A1, and A2 model. Intermediate results were observed for the R and BC model. The least performing were the T and RN model.

Table 2.1. Statistical indices of the models. †

model	mean RMSE	mean R^2	mean AIC
	$\text{m}^3 \text{ m}^{-3}$	-	-
BC	0.0141	0.965	-13.60
VG1	0.0072	0.990	-17.27
VG2	0.0101	0.982	-16.17
T	0.0211	0.915	-11.42
R	0.0128	0.966	-14.18
RN	0.0186	0.876	-11.16
K1	0.0081	0.988	-16.42
K2	0.0094	0.985	-16.83
A1	0.0104	0.982	-14.38
A2	0.0108	0.980	-15.73

† RMSE, root of mean squared errors; AIC, Akaike Information Criterion; BC, Brooks and Corey (1964) model; A1, Assouline et al. (1998) with five free parameters; A2, Assouline et al. (1998) with four free parameters; K1, Kosugi (1994); K2, Kosugi (1996)(1997); R, Russo (1988); RN, Rossi and Nimmo (1994); T, Tani (1982); VG1, van Genuchten (1980) with five free parameters; VG2, van Genuchten (1980) with four free parameters.

Table 2.2 summarizes the Pearson correlation coefficients r_{sp} computed between all parameters on the one hand and the two most significant soil properties on the other hand. As could be expected, θ_s , which was constrained at $\theta_{-1 \text{ kPa}}$, was highly correlated to ρ_b and to a lesser extent to clay content. This was also concluded by Vereecken et al. (1989) based on their principal component analysis. The variation in θ_r (or θ_L) was for all models to a relatively high extent explained by ρ_b and clay content as well. All models showed comparable prediction potential for θ_r (or θ_L), except the BC model, which showed significant lower r_{sp} values. The other parameters, which mainly determine the specific shape of the SWRC, showed lower correlations. Highest r_{sp} values were observed for the T model, which only has one additional parameter. Unfortunately, this model performed rather poorly in terms of goodness-of-fit to SWRC data. Relatively high values were also observed for the two additional parameters of the K2 and VG2 models. The A2 model, which also has two additional parameters, showed a relatively low correlation for its ξ parameter. The lowest

values were computed for the R and RN models. Further, the models with three additional parameters, such as VG1, K1 and A1, also showed an overall low correlation. In the case of the VG1 model, this was merely due to a low correlation with the parameter m . The K1 and A1 models showed a relatively high correlation with only one of the additional parameters.

It should be noted that the above conclusions were drawn on a limited data set representing 48 horizons of 24 soil series (from soils which were under forest) and covering eight soil textural classes. As an illustration for the variability of the data set, all original SWRC points are depicted in Figure 2.3.

Table 2.2. Pearson correlation coefficient between model parameters and basic soil properties.

model†	θ_s vs. ρ_b	θ_s vs. clay	θ_s/θ_L vs. ρ_b	θ_s/θ_L vs. clay	$a_{\ddagger}^{\#}$ vs. sand/clay¶	$a_{\ddagger}^{\#}$ vs. silt/clay#	$b^{\dagger\dagger}$ vs. sand/OM ‡‡	$b^{\dagger\dagger}$ vs. silt/clay/ ρ_b §§	$c^{\ \ }$ vs. sand	$c^{\ \ }$ vs. silt/clay##
BC	-0.94	0.65	-0.52	0.51	-0.23§	0.34	0.65	-0.60	-	-
VG1	-0.94	0.65	-0.75	0.72	0.45	-0.50	0.50	-0.49	-0.18§	0.29§
VG2	-0.94	0.65	-0.75	0.72	0.46	-0.51	0.54	-0.49	-	-
T	-0.94	0.65	-0.72	0.83	-0.63	0.74	-	-	-	-
R	-0.94	0.65	-0.72	0.82	0.10§	-0.09§	-0.49	0.57	-	-
RN	-0.94	0.65	-	-	-0.17§	0.19§	0.30	-0.29	-0.22§	0.46
K1	-0.94	0.65	-0.74	0.81	0.31	-0.19§	<u>-0.32</u>	<u>0.27§</u>	0.50	0.55
K2	-0.94	0.65	-0.74	0.81	-0.47	0.55	-0.53	0.48	-	-
A1	-0.94	0.65	-0.69	0.78	0.27§	-0.25§	0.56	0.52	0.21§	0.39
A2	-0.94	0.65	-0.70	0.82	0.26§	-0.24§	0.54	-0.50	-	-

† BC, Brooks and Corey (1964) model; A1, Assouline et al. (1998) with five free parameters; A2, Assouline et al. (1998) with four free parameters; K1, Kosugi (1994); K2, Kosugi (1996)(1997); R, Russo (1988); RN, Rossi and Nimmo (1994); T, Tani (1982); VG1, van Genuchten (1980) with five free parameters; VG2, van Genuchten (1980) with four free parameters.

‡ a corresponds to $\psi_b, \alpha, \alpha, \psi_{ip}, \alpha', \psi_b, \psi_{md}, \psi_{md}, \zeta,$ and ζ in the BC, VG1, VG2, T, R, RN, K1, K2, A1 and A2 model, respectively.

§ Not significant at the 0.05 level.

¶ All a values are correlated to sand content, except for the value in italics, which is correlated to clay content.

All a values are correlated to silt content, except for the value in italics, which is correlated to clay content.

†† b corresponds to $\lambda, n, n, m', \lambda, \sigma, \sigma, \eta$ and η in the BC, VG1, VG2, R, RN, K1, K2, A1 and A2 model, respectively.

‡‡ All b values are correlated to sand content, except for the value underlined, which is correlated to OM.

§§ All b values are correlated to silt content, except for the value in italics, which is correlated to clay content, and the value underlined, which is correlated to ρ_b

¶¶ c corresponds to m, ψ_i, ψ_b and ψ_L in the VG1, RN, K1, and A1 model, respectively.

All c values are correlated to silt content, except for the values in italics, which is correlated to clay content.

2.4.2 Behaviour of the models

To illustrate the behaviour of the ten closed-form analytical expressions compared in this study when fitted to soil-water retention data of a relative coarse-textured soil (sand, $\rho_b = 1.668 \text{ Mg m}^{-3}$, $\text{OM} = 37.98 \text{ g kg}^{-1}$) and fine-textured soil (silt loam, $\rho_b = 1.682 \text{ Mg m}^{-3}$, $\text{OM} = 2.71 \text{ g kg}^{-1}$), observed and fitted data are compared in Figure 2.4. When considering the sand, all models gave relatively good and realistic fits, at least when the soil-matric potential remains higher than the lower limit of the data sets, that is, higher than -1500 kPa . The only model that showed a reliable behaviour beyond the driest measured point (i.e. accounting also for adsorptive water retention and reaching a value of zero water content) is the RN model, which is not surprisingly as it was developed for that purpose. All other models resulted in a SWRC that is undefined for soil-water contents below θ_r or θ_L . This is a serious drawback since many water related processes such as deflation of soil particles by wind (Cornelis et al., 2004), microbial activity, and N mineralization in soils (De Neve and Hofman, 2002), methane oxidation in soils (De Visscher and Van Cleemput, 2003), and applications in for example, colloid science (Blunt, 2001) and food technology (Weerts et al., 2003) are affected by soil-water contents well below residual. On the other hand, the discontinuous character of the BC model and the K1 model did not seem to be problematic for sand, at least in comparison with our limited number of observations near saturation.

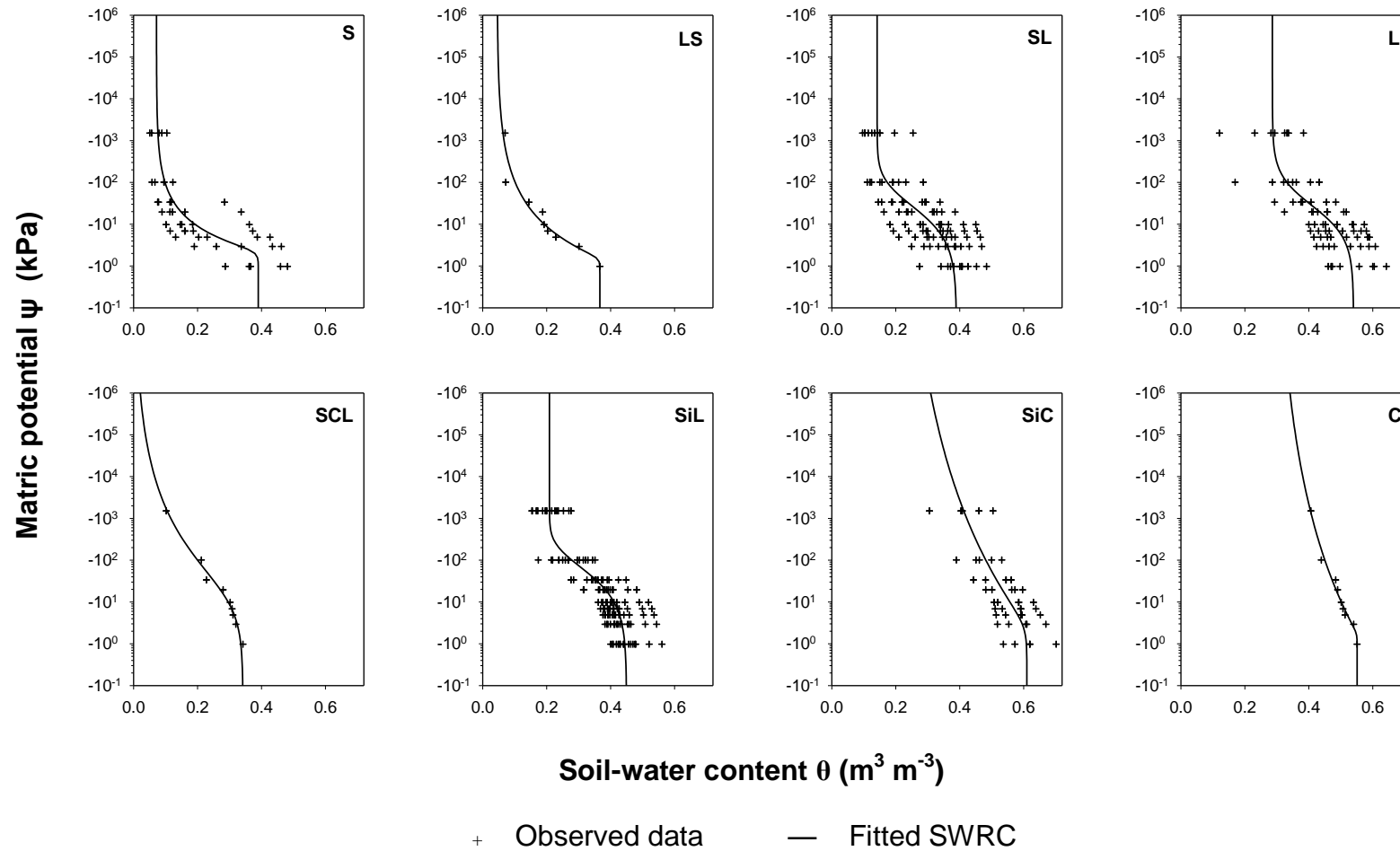


Figure 2.3. Observed soil-water retention curves grouped per soil textural class and VG1 model fitted to the pooled data sets.

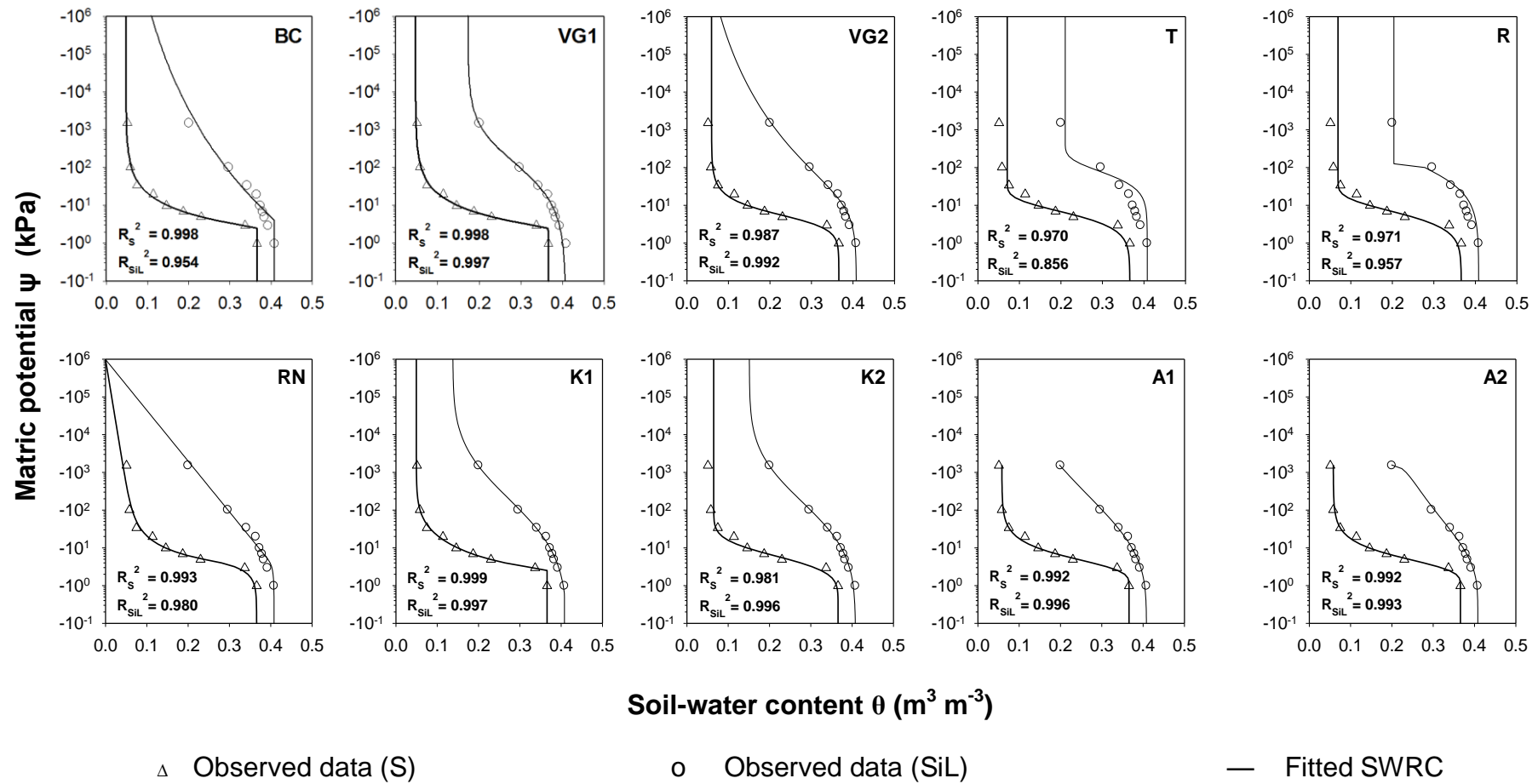


Figure 2.4. Observed and fitted soil-water retention curves for sand and silt loam. The subscripts ‘S’ and ‘SiL’ denote sand and silt loam, respectively.

With respect to the silt loam example, the performance of the BC, T, R, and RN model seemed to be reduced compared with the sand example. The BC model produced a relatively poor match near saturation, due to its discontinuous character and its unrealistic high estimated bubbling pressure value. The T model and the R model seemed to be unrealistic over the whole data range. This is in the case of the T model due to the exponential term, which when indexed to the inflection point ψ_{ip} , shows a typical sigmoid shape. When multiplying the exponential term with $1 + \psi/\psi_{ip}$, the sigmoid shape becomes even more pronounced, for the effect of this term is that it increases θ when ψ decreases. The R model even showed a discontinuity at the inflection point $0.5 \alpha'$ or $1/\psi_{ip}$, which occurs at relatively high m' values. The higher the inflection point (which is the case as the soil texture becomes finer), the lower the α' value, and hence the higher the m' value should be to keep the curve straight near saturation. Compared with the T model, the $1 + \psi/\psi_{ip}$ term is here augmented with a power $2/(2 + m')$, and hence its effect becomes more pronounced as m' decreases. The RN model showed a poor fit near saturation. This is because the inflection point ψ_i should be high enough to ensure an acceptable fit in the logarithmic part of the SWRC (including the point at oven dryness). It further performed rather poor in the dry range, due to its low flexibility in the shape of the curve. This is associated with a lower degree of freedom in the dry range compared with the other models where θ_r or θ_L are free parameters. Finally, both forms of the van Genuchten (1980) model, VG1 and VG2, and the Kosugi (1994, 1996, 1997) models, K1 and K2, showed very good fits to the silt loam data, but have still the drawback of an undefined SWRC for soil-water contents below θ_r . Also the two Assouline et al. (1998) functions, A1 and A2, described the SWRC rather well for the silty loam. As both are mathematically not defined at soil-water contents lower than θ_L , the curve was not drawn beyond that point (which is only apparent for A2 in Fig. 2.4).

In Table 2.3, parameter values are given for the different models and for different soil textural classes. They were obtained by curve fitting the models to the whole data set for each soil textural class. These data can be useful to the reader as initial estimates when attempting to use one of the evaluated expressions. In the case of the BC model and the van Genuchten (1980) model, existing PTFs that are widely reported can also be used for that purpose. Table 2.3 further illustrates that the parameter values of n of the VG2 model follows a more

pronounced trend compared with n calculated for the VG1 model, in that for example, the curves become steeper (lower n) as the soils become finer in texture.

Table 2.3. Average model parameters for different soil textural classes. †

model	par	S	LS	SL	L	SCL	SiL	SiC	C
BC	θ_r	0.071	0.045	0.043	0.131	0.000	0.000	0.091	0.284
	θ_s	0.400	0.366	0.390	0.540	0.342	0.450	0.610	0.551
	ψ_b	2.098	1.538	2.386	2.607	3.917	4.031	2.240	1.742
	λ	0.644	0.416	0.212	0.161	0.171	0.128	0.074	0.117
VG1	θ_r	0.072	0.045	0.142	0.286	0.000	0.209	0.203	0.285
	θ_s	0.390	0.366	0.390	0.540	0.342	0.450	0.610	0.551
	α	0.447	0.650	0.013	0.030	0.075	0.002	0.373	0.573
	n	7.073	9.304	0.921	1.063	0.969	0.877	2.299	8.283
	m	0.092	0.045	2.504	1.203	0.256	6.230	0.045	0.014
VG2	θ_r	0.083	0.059	0.120	0.267	0.000	0.161	0.371	0.355
	θ_s	0.390	0.366	0.390	0.540	0.342	0.450	0.610	0.551
	α	0.279	0.398	0.147	0.113	0.115	0.076	0.136	0.273
	n	1.940	1.559	1.521	1.540	1.222	1.393	1.321	1.225
T	θ_r	0.113	0.119	0.166	0.314	0.156	0.234	0.444	0.434
	θ_s	0.390	0.366	0.390	0.540	0.342	0.450	0.610	0.551
	ψ_{ip}	3.440	3.174	10.72	12.54	15.63	21.61	14.27	9.729
R	θ_r	0.104	0.100	0.155	0.302	0.117	0.214	0.437	0.423
	θ_s	0.390	0.366	0.390	0.540	0.342	0.450	0.608	0.551
	α'	3.740	43.65	14.71	14.64	14.66	14.61	0.536	14.71
	m'	24.07	402.5	404.1	492.7	1013	955.4	12.57	424.6
RN	θ_s	0.390	0.366	0.390	0.540	0.342	0.450	0.610	0.551
	ψ_b	1.715	0.855	0.362	2.978	6.093	2.758	0.107	0.107
	ψ_o	3.934	2.590	4.577	2.980	6.093	5.796	9.726	15.623
	λ	1.036	0.551	0.413	0.031	0.136	0.076	0.000	0.000
K1	θ_r	0.074	0.062	0.139	0.286	0.060	0.202	0.411	0.392
	θ_s	0.390	0.366	0.390	0.540	0.342	0.450	0.610	0.551
	ψ_b	2.491	1.984	0.000	0.000	0.917	0.000	0.127	1.796
	ψ_{md}	2.520	2.002	1.981	2.692	1.049	3.343	1.341	1.811
	σ	2.209	2.393	1.544	1.513	2.585	1.643	1.849	2.778
	ψ_L	1.519	2.443	840.464	1.616	1.519	330.563	107.621	2901.038
K2	θ_r	0.090	0.071	0.139	0.286	0.074	0.202	0.412	0.401
	θ_s	0.390	0.366	0.390	0.540	0.342	0.450	0.610	0.551
	ψ_{md}	6.560	7.963	21.47	26.54	86.79	49.69	37.16	30.27
	σ	1.154	1.647	1.544	1.513	2.312	1.643	1.827	2.122
	ψ_L	1.519	2.443	840.464	1.616	1.519	330.563	107.621	2901.038
	ψ_L	1.519	2.443	840.464	1.616	1.519	330.563	107.621	2901.038
A1	θ_r	0.083	0.066	0.123	0.286	0.102	0.167	0.386	0.382
	θ_s	0.390	0.366	0.390	0.534	0.342	0.450	0.610	0.551
	ξ	4.572	2.890	4.997	7.372	4.383	6.102	4.626	3.447
	η	1.009	0.706	0.626	0.728	0.437	0.509	0.481	0.428
	ψ_L	1.519	2.443	840.464	1.616	1.519	330.563	107.621	2901.038
	ψ_L	1.519	2.443	840.464	1.616	1.519	330.563	107.621	2901.038
A2	θ_r	0.083	0.068	0.136	0.283	0.102	0.206	0.414	0.406
	θ_s	0.390	0.366	0.390	0.540	0.342	0.450	0.610	0.551
	ξ	4.570	2.904	5.259	6.243	4.383	6.733	4.917	3.522
	η	1.009	0.715	0.668	0.675	0.437	0.583	0.550	0.495
	ψ_L	1.519	2.443	840.464	1.616	1.519	330.563	107.621	2901.038

† The units of the parameters are based on soil-water content θ expressed in $m^3 m^{-3}$ and ψ in kPa, except when otherwise mentioned; S = sand, LS = loamy sand, SL = sandy loam, L = loam, SCL = sandy clay loam, SiL = silt loam, SiC = silty clay, C = clay. BC, Brooks and Corey (1964) model; A1, Assouline et al. (1998) with five free parameters; A2, Assouline et al. (1998) with four free parameters; K1, Kosugi (1994); K2, Kosugi (1996, 1997); R, Russo (1988); RN, Rossi and Nimmo (1994); T, Tani (1982); VG1, van Genuchten (1980) with five free parameters; VG2, van Genuchten (1980) with four free parameters.

‡ ψ_L in MPa.

2.4.3 Effect of soil properties on the model performance

To assess the dependency of the model performance on soil properties, the SSE computed per soil sample for each model was correlated to ρ_b , OM, and sand, silt and clay content (see Table 2.4) The Pearson correlation coefficient between SSE and ρ_b was not significant at the 0.05 level for all models except the RN model. When correlating SSE and OM no

significance was found at the 0.05 level for all models. The performance of the models, except the RN model, was thus not affected by ρ_b and OM, which can vary substantially in forest soils. When relating the model's SSE values to soil texture, the performance of the models appeared to increase as the soils became finer in texture, that is, higher in clay and silt content (negative correlation), except for the RN model. The opposite was true when considering sand content. This demonstrates once more that it is simply the specific mathematical form of the models that determines their performance, rather than the physical meaning of their parameters or their conceptual background. Such deterministic models were derived by applying distribution laws to pore-size distribution functions, in combination with capillarity laws. The lower the dominance of the capillary forces over the adhesive and osmotic forces in retaining water to the soil matrix, as is the case when soils become higher in clay and OM, the lower the performance of such deterministic models is expected to be, whereas in our study, the opposite was observed.

Table 2.4. Pearson correlation coefficient between SSE and basic soil properties for each model.

model	ρ_b	O.M	sand	silt	clay
BC	-0.14†	0.19†	0.13†	-0.01†	-0.31
VG1	0.08†	0.02†	0.50	-0.45	-0.35
VG2	0.05†	0.05†	0.46	-0.38	-0.40
T	-0.04†	0.09†	0.30	-0.27†	-0.22†
R	0.15†	-0.02†	0.50	-0.49	-0.28†
RN	-0.58	0.11†	-0.31	0.06†	0.66
K1	-0.01†	0.08†	0.30	-0.22†	-0.31
K2	0.14†	0.01†	0.56	-0.49	-0.43
A1	-0.01†	0.09†	0.38	-0.29	-0.39
A2	-0.02†	0.09†	0.35	-0.26†	-0.38

† not significant at the 0.05 level.

The poor overall performance of the RN model can mainly be attributed to its poor fits when textures became relatively fine. So, the RN model is, although it has a more realistic shape than all the other models evaluated here and describes the SWRC over the complete range of soil-water contents (from saturation to oven dryness), not a reliable alternative for the superiorly performing VG1, VG2, K1, or K2 model, at least when using data sets with a limited number of data pairs, as is most often the case in practice. If more data are available, then the RN model could perhaps perform better as was demonstrated by Rossi and Nimmo (1994) for seven soils.

2.5 Conclusions

Using five a limited data set taken from 48 horizons of forest soils in Flanders, Belgium and representing eight soil-textural classes, we have evaluated ten closed-form unimodal analytical expressions for the SWRC. It was shown that the van Genuchten (1980) model with five model parameters had the highest performance in terms of the RMSE, R^2 and AIC. However, its prediction potential was rather poor, due to the low correlation between the m parameter and basic soil properties. Reducing the number of parameters to four, increased the prediction potential of the model significantly, without losing much of its performance. A high performance was also observed for the five-parameter and four-parameter Kosugi models (1994, 1996, 1997) and for the five-parameter and four-parameter Assouline et al. (1998) models. Yet, these models had, except for the four-parameter Kosugi (1996, 1997) model, a low prediction potential.

A major drawback of these models is that they do not define the soil-water content vs. soil-matric potential relationship beyond the residual water content. The only model we evaluated that is able in doing so is the Rossi and Nimmo (1994) model. However, it showed the lowest performance in terms of goodness-of-fit, at least when using a limited number of nine data pairs as was the case in our study. It further showed a low prediction potential. Therefore, more recently developed expressions for the SWRC between saturation and oven dryness need to be evaluated or new expressions should be developed.

Finally, it was shown that the performance of all models in terms of their match to the data, increased with increasing clay content and decreasing sand content, except for the Rossi and

Nimmo (1994) model, which is contradictory to the deterministic character of these models. Furthermore, it was shown that ρ_b and OM, at least within the range of our data set, did not have a significant effect on the accuracy of all models, except the least performing ones.

Acknowledgment

The database used in this work was setup in the framework of project 00/05 of the VLINA research program (1995-2001), which was funded by the Ministry of the Flemish Community and conducted by Ghent University and the Institute for Forestry and Game Management. The authors thank Maarten Vanoverbeke, Koen Willems, Ann Capieau and Jan Restiaen for their assistance in sampling, analysis and data-reporting. The helpful suggestions of two anonymous reviewers, and of Wolfgang Durner and Horst Gerke are also greatly acknowledged.

Chapter 3 Simple modification to describe the soil-water retention curve between saturation and oven-dryness

Based on: Khlosi M., Cornelis W.M., Gabriels D., and Sin G. (2006). Simple modification to describe the soil water retention curve between saturation and oven dryness. *Water Resour. Res.* 42: Art. No. 11501.

Abstract

Prediction of water and vapor flow in porous media requires an accurate estimation of the soil water retention curve describing the relation between matric potential and the respective soil water content from saturation to oven dryness. In this study, we modified the Kosugi (1999) function to represent soil water retention at all matric potentials. This modification retains the form of the original Kosugi function in the wet range and transforms to an adsorption equation in the dry range. Following a systems identification approach, the extended function was tested against observed data taken from literature that cover the complete range of water contents from saturation to almost oven dryness with textures ranging from sand to silty clay. The uncertainty of parameter estimates (confidence intervals) as well as the correlation between parameters was studied. The predictive capability of the extended model was evaluated under two reduced sets of data that do not contain observations below a matric potential of -1500 and -100 kPa. This evaluation showed that the extended model successfully predicted the water content with acceptable uncertainty. These results add confidence into the proposed modification and suggest that it can be used to better predict the soil water retention curve, particularly under reduced data sets.

3.1 Introduction

Understanding the behaviour of water in unsaturated porous media is a challenge for scientists. However, an adequate description of water behaviour in the unsaturated soils relies mainly on accurate estimates of the soil water retention and the unsaturated soil hydraulic functions. The soil water retention curve (SWRC) describes the relationship between matric potential, ψ , and soil-water content, θ . Several analytical functions for describing the SWRC can be found in literature (e.g., Brooks and Corey, 1964; van Genuchten, 1980). Most of these retention models are successful in the wet region of the SWRC. Cornelis et al. (2005) compared ten closed-form unimodal analytical functions to describe the SWRC. In their study, the van Genuchten (1980) and Kosugi (1994) models showed good fits to the observed data, specifically at high and medium water content. In the dry region of the SWRC as water content goes to zero, however, these models often fail to describe the observed trend.

During the last two decades, several attempts have been made to obtain the complete retention curve (Ross et al., 1991; Campbell and Shiozawa, 1992; Rossi and Nimmo, 1994; Fayer and Simmons, 1995; Morel-Seytoux and Nimmo, 1999; Webb, 2000; Tuller and Or, 2002; Groenevelt and Grant, 2004). In this paper, an alternative closed-form analytical expression for the SWRC is proposed and evaluated following a systems identification approach. The model is tested against data taken from literature that cover the complete range of water contents, from saturation to almost oven-dryness. To evaluate the predictive capability of the model at the dry region of the SWRC, two reduced sets of data that do not contain observations beyond a matric potential of -1.5×10^3 kPa and -100 kPa, respectively, will be used. Our motivation for using these reduced data sets is that (1) constructing the SWRC beyond -1.5×10^3 kPa does not belong to the standard procedure in most labs. Some labs have even difficulties in maintaining a pressure of -1.5×10^3 kPa. And (2), in lysimeter or field studies where water content and matric potential are measured simultaneously, which allows constructing an in situ SWRC, the range of matric potentials is limited to -10^3 kPa when using the heat dissipation method or gypsum blocks, to -200 kPa when using the electrical moisture method, and to -85 kPa when using tensiometry.

3.2 Approach

In recent years, considerable attention has been given towards extending classical retention models to oven-dryness. The most commonly employed classical retention models are the two unimodal functions presented by Brooks and Corey (1964) and by van Genuchten (1980). Recently, the Kosugi (1999) model has gained popularity. These functions are respectively written as:

$$\begin{aligned} \theta &= \theta_s && \text{for } \psi \geq \psi_b \\ \theta &= \theta_r + (\theta_s - \theta_r) \left(\frac{\psi_b}{\psi} \right)^\lambda && \text{for } \psi < \psi_b, \end{aligned} \quad [3.1]$$

$$\theta = \theta_r + (\theta_s - \theta_r) \left(\frac{1}{1 + (\alpha |\psi|)^n} \right)^{\left(1 - \frac{1}{n}\right)} \quad [3.2]$$

$$\theta = \theta_r + \frac{1}{2} (\theta_s - \theta_r) \operatorname{erfc} \left[\frac{\ln \left(\frac{\psi}{\psi_{md}} \right)}{\sqrt{2} \sigma} \right] \quad [3.3]$$

where θ_s and θ_r are the saturated and residual soil-water content, respectively, ψ_b is the bubbling pressure potential or air entry value, ψ_{md} is the matric potential corresponding to the median pore radius, λ , α and n are curve-fitting parameters related to the pore-size distribution, σ is a dimensionless parameter to characterize the width of the pore-size distribution, and “erfc” denotes the complementary error function. Since (3.1) and (3.2) are historically the most-widely used functions by soil scientists, hydrologists and engineers, various attempts have been conducted to extend these models for all matric potentials. Attempts to extent the Kosugi (1999) model have not been reported yet.

The problem with (3.1), (3.2) and (3.3) is that they, mathematically, define the residual water content as the water content where $d\theta/d\psi$ becomes zero, or at $\psi = -\infty$ MPa, which is physically not realistic. Further, θ_r often becomes negative in the curve-fitting procedure. As negative water content is undefined, θ_r is then forced to converge to zero, and this result as

well is an unrealistic path of the retention curve at low water contents (Cornelis et al., 2005). Moreover, as mentioned in Chapter 2, this region is critical in a number of water related processes which are affected by soil-water contents well below residual.

Campbell and Shiozawa (1992) used a modified form of the van Genuchten (1980) equation for improving fits to dry data, in which adsorption of water on soil was described with a semi-logarithmic expression. Fayer and Simmons (1995) further modified the van Genuchten retention function by replacing the residual water content with the adsorption equation of Campbell and Shiozawa (1992) as:

$$\theta = \theta_a \left(1 - \frac{\ln(\psi)}{\ln(\psi_o)} \right) + \left(\theta_s - \theta_a \left(1 - \frac{\ln(\psi)}{\ln(\psi_o)} \right) \right) \left[\frac{1}{1 + (\alpha \psi)^n} \right]^{\left(1 - \frac{1}{n} \right)} \quad [3.4]$$

where θ_a is a curve-fitting parameter representing the soil-water content at $\psi = 1$ kPa, and ψ_o is the matric potential at oven dryness. This expression is denoted here as FS model. Since Cornelis et al. (2005) demonstrated that the four-parameter Kosugi model performed slightly better than four-parameter van Genuchten (1980) model, we replace the residual water content in the Kosugi (1999) model by the adsorption equation of Campbell and Shiozawa (1992) (denoted here as KCGS model):

$$\theta = \theta_a \left(1 - \frac{\ln(\psi)}{\ln(\psi_o)} \right) + \left(\theta_s - \theta_a \left(1 - \frac{\ln(\psi)}{\ln(\psi_o)} \right) \right) \frac{1}{2} \operatorname{erfc} \left[\frac{\ln \left(\frac{\psi}{\psi_{md}} \right)}{\sqrt{2} \sigma} \right] \quad [3.5]$$

3.3 Materials and methods

3.3.1 Experimental data

Three soils having retention properties measured and reported by Campbell and Shiozawa (1992) are used to test the modified functions. These soils include L-soil, Salkum, Palouse B, and have textures of sand, silt loam, and silty clay (sandy, loamy and clayey soils) respectively. The data sets have been chosen based on their wide range of potentials, from -3 kPa to -3.2×10^5 kPa, which is from nearly saturation to nearly oven-dryness. This wide range, almost six orders of magnitude in potential and referred to as data set #1, enabled us to validate our proposed model (KCGS) and identify its parameters. The KCGS function was further tested against reduced data sets, i.e. containing matric potentials $\geq -1.5 \times 10^3$ kPa (data set #2) and matric potentials ≥ -100 kPa (data set #3), and was compared to the FS model.

3.3.2 Model analysis

The fitting procedure to all three data sets was performed as previously described in Chapter 2 (see chapter 2, section 2.3.2). To gain a better insight into the proposed model structure and to better judge the credibility of the model on commonly agreed scientific grounds (Dekker et al., 2001, Dochain and Vanrolleghem, 2001), the identifiability of parameters (sensitivity analysis, confidence intervals, parameter correlation) and uncertainty analysis of model outputs were also studied.

A local sensitivity analysis approach was used to study the quantitative relationship between the model parameters and the output (in this case y stands for θ) (Saltelli et al., 2005):

$$S = \frac{\partial y}{\partial \beta} \quad [3.6]$$

where S is the output sensitivity function of the model output, y to the parameter, β . The sensitivity functions were derived analytically using MathCad[®].

To evaluate the accuracy of the estimated parameters, the covariance matrix of the estimated parameters, $\text{COV}(\beta)$ was approximated using the inverse of the Fisher Information Matrix (FIM) (see Dochain and Vanrolleghem, 2001; Omlin and Reichert, 1999):

$$FIM = \sum_{i=1}^N \left(\frac{\partial y_i}{\partial \beta_j}(\beta) \right)^T Q_i^{-1} \left(\frac{\partial y_i}{\partial \beta_j}(\beta) \right) \quad [3.7]$$

$$COV(\beta) \approx \frac{SSE}{N-m} FIM^{-1} \quad [3.8]$$

where $\frac{\partial y_i}{\partial \beta_j}$ is the absolute sensitivity of the model output y_i to model parameter β_j and Q is the matrix of the variance of measurement errors. $SSE/(N-m)$ is equal to s^2 , which is an estimate of the variance of model fits $\hat{\sigma}^2$. SSE is the sum of squared errors corresponding to the minimum cost function. N is the total number of measurements and m is the total number of parameters estimated. The confidence interval of parameter $\beta_j, \delta\beta_j$, is:

$$\delta\beta_j = \pm \sqrt{COV(\beta)_{jj}} t(N-m; \hat{\alpha}) \quad [3.9]$$

where $t(n-m; \hat{\alpha})$ is the value of two-tailed student t-distribution at $1-\hat{\alpha}$ the confidence level with $n-m$ degrees of freedom. $\hat{\alpha}$ is typically taken as 0.05 (5%) which means that 95 times of the cases the estimated value of the parameter will lie within the given confidence interval. The linear correlation between two parameters R_{ij} is:

$$R_{ij} = \frac{COV(\beta_i, \beta_j)}{\sqrt{\sigma_{\beta_i}^2 \hat{\sigma}_{\beta_j}^2}} \quad [3.10]$$

and can range between -1 and +1. Values of R_{ij} close to zero imply zero or no correlation.

The covariance matrix of the model output was approximated by first order, linear propagation of the uncertainty of parameter estimates through the model structure (Omlin and Reichert, 1999):

$$COV(y_i) = \sum_{i=1}^N \left(\frac{\partial y_i}{\partial \beta_j}(\beta) \right)^T COV(\beta) \left(\frac{\partial y_i}{\partial \beta_j}(\beta) \right) \quad (3.11)$$

$$\delta y_i = \pm \sqrt{COV(y_i)_{ii}} t(N-m; \hat{\alpha}) \quad (3.12)$$

where δy_j indicates the uncertainty of the i^{th} model output at the confidence level $(1 - \hat{\alpha})$. The term $\sqrt{COV(y_i)_{ii}}$ indicates the diagonal elements of the covariance matrix which is the variance of i^{th} model output.

Finally, the prediction accuracy of the FS and KCGS models are quantified using mean error (ME) and root mean square error (RMSE) criteria:

$$ME = \frac{1}{N} \sum_{i=1}^N [\hat{\theta}_i - \theta_i] \quad (3.13)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N [\hat{\theta}_i - \theta_i]^2} \quad (3.14)$$

where N is the number of $\theta(\psi)$ pairs in a SWRC, $\hat{\theta}_i$ and θ_i are estimated and measured water contents, respectively. The value of $\hat{\theta}_i$ is computed by evaluating the appropriate retention function at the observed matric potential with the estimated retention parameters. ME measures the bias of the predicted variable and should be as small as possible. RMSE is a measure of the precision of the predicted variable and should be as small as possible for unbiased precise prediction.

3.4 Results and discussion

Table 3.1 summarizes the estimated parameters along with their 95% confidence intervals and the correlation coefficients between parameters as computed for the three soils using the KCGS model. For reliable and accurate parameter estimation, the confidence interval of the parameter should be as low as possible, which indicates low uncertainty on the estimated parameter value. When this low parameter uncertainty propagates properly to model-outputs, e.g. first order error propagation, the uncertainty band around the model output will also be small (see below). That said, it appears that confidence intervals of the parameter estimates of the KCGS are low (see Table 3.1).

Table 3.1. Optimized Parameters, Confidence Intervals, and Correlation Coefficient between Parameters for the KCGS Model. †

	L-soil				Salkum				Palouse B			
	θ_a	θ_s	ψ_m	σ	θ_a	θ_s	ψ_m	σ	θ_a	θ_s	ψ_m	σ
Optimized	0.07	0.16	6.47	0.35	0.20	0.48	77.36	1.69	0.36	0.55	16.90	1.90
$\delta\beta_j$	0.00	0.01	0.47	0.09	0.01	0.01	9.71	0.15	0.01	0.04	7.62	0.42
θ_a	1	-0.08	-0.17	-0.25	1	-0.28	-0.28	-0.59	1	-0.38	0.18	-0.63
θ_s	-	1	-0.68	0.57	-	1	-0.70	0.74	-	1	-0.92	0.84
ψ_{md}	-	-	1	-0.37	-	-	1	-0.40	-	-	1	-0.80
σ	-	-	-	1	-	-	-	1	-	-	-	1

† The units of the parameters are based on soil-water content θ expressed in $m^3 m^{-3}$ and ψ in kPa.

Figure 3.1 shows the output sensitivity functions of water content to the parameters of the KCGS model as a function of matric potential. One observes that a significant correlation exists between model parameters. The degree of correlation appears to be dependent on the matric potential, e.g. the parameters θ_a and θ_s are inversely proportional between -10 and -10^3 kPa whereas the correlation breaks down beyond -10^3 kPa (see Figure 3.2). Although the correlation coefficients are relatively low in the KCGS model, still they cannot be ignored. Therefore the parameter estimates should be interpreted with care considering the existing correlation.

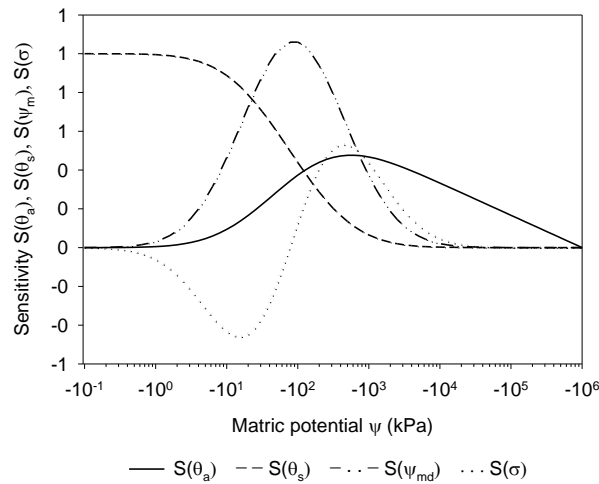


Figure 3.1. Output sensitivities of water content to the parameters in the KCGS. Note that $S(\psi_{md})$ was multiplied by 10^3 and σ by 10^6 for clarity of representation on one single scale.

In comparing the KCGS and FS models, they were first calibrated using a reduced number of data pairs, datasets #2 and #3 as explained above, and then used to predict the entire (original) data sets which were not used in the calibration. In Figure 3.2, the SWRCs and their uncertainty bands established by using both datasets of the Salkum soil, as explained above, are plotted for the KCGS and FS models. When considering dataset #2, both models were able to satisfactorily reproduce the data in the extrapolation region with very narrow 95% uncertainty bands (i.e. those data not used in calibration) (see Figures 3.2a and 3.2b) although the fit of the KCGS model was relatively better: the ME and RMSE values associated with KCGS ranged from -0.0009 to 0.0008 and from 0.0040 to 0.0079 respectively, whereas for the FS model, they varied between -0.0007 and 0.0055 , and 0.0036 and 0.0105 respectively.

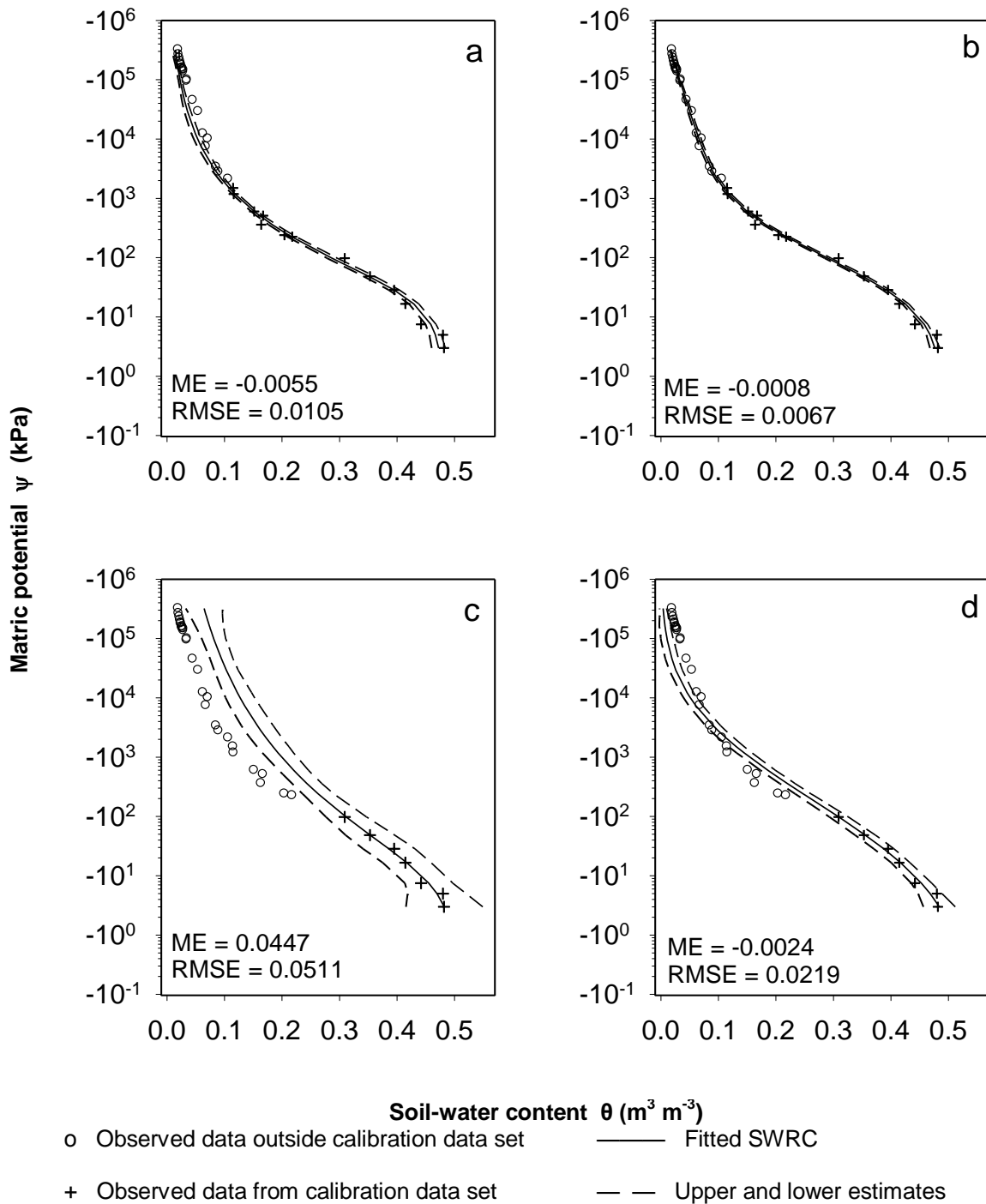


Figure 3.2. Comparison of measured and fitted soil-water retention curves using FS (a and c) and KCGS (b and d) models for Salkum with 95% uncertainty bands. The fits were obtained by calibrating the models using two data sets of (θ, ψ) : data #2, and #3 include ψ values up to -1.5×10^3 kPa (a and b), and -100 kPa (c and d), respectively.

The difference between the two models becomes pronounced when a more reduced dataset, i.e, #3, was used for calibration (see Figures 3.2c and 3.2d). The predictions of the FS model were accompanied with a large uncertainty band, and most important, the predictions deviated significantly from the data in the extrapolation region. The KCGS model, however, was still able to follow the observed measurements in the extrapolation region with acceptable uncertainty bands. This means that the KCGS model is able to describe the SWRC between saturation and oven-dryness, without having (θ, ψ) data beyond a matric potential of -100 kPa. This would eliminate the need to experimentally determine the SWRC beyond -1.5×10^3 kPa or even beyond -100 kPa. This is an important improvement compared to e.g. the Rossi and Nimmo model (1994) for the SWRC between saturation and oven-dryness. Cornelis et al. (2005) have demonstrated that the latter model behaves rather poorly when using data sets with a limited number of data pairs (such as data set #2). The ME and RMSE associated with the KCGS model using dataset #3 are well below those of the FS model: the ME and RMSE values associated with KCGS ranged from -0.0054 to 0.0024 and from 0.0049 to 0.0219 respectively, whereas for the FS model, they varied between -0.0447 and -0.0015 , and 0.0045 and 0.0511 respectively.

Overall, these preliminary results are promising and suggest that the SWRC can be reproduced by KCGS model with acceptable accuracy under a limited range of (θ, ψ) data, such as in lysimeter and field studies, or when maintaining a pressure of -1.5×10^3 kPa is cumbersome as is often observed in many under-developed labs. Nevertheless, additional research on testing the model performance under a wide range of data sets from diverse soils and textures will be needed to confirm these promising results.

3.5 Conclusions

In this paper we modified the Kosugi (1999) model to improve the description of water retention across the entire range of soil-water content from saturation to oven-dryness. We have replaced the residual water content in the four-parameter form of Kosugi (1999) model by the adsorption equation of Campbell and Shiozawa (1992). Tested on three different soils having retention properties measured and reported by Campbell and Shiozawa (1992), the new model reproduced well the observed data in the whole range of soil-water contents. The

model parameters could be identified uniquely using a limited matric potential range. However, the inherent correlation existing between model parameters could not be totally overcome in the new model.

The major achievement of the model is its ability to predict the entire region of SWRC when calibrated using a limited dataset that includes only those measurements of water content beyond -100 kPa matric potential. These promising results suggesting that the model remains largely valid when calibrated with a limited data of SWRC need to be confirmed with a wider range of soils and texture data.

Acknowledgment

Authors gratefully acknowledge Kari A. Winfield (U.S. Geological Survey) for providing the soil-water retention data.

Chapter 4 Performance evaluation of models that describe the soil-water retention curve between saturation and oven dryness

Based on: Khlosi M., Cornelis W.M., Douaik A., van Genuchten M.Th., and Gabriels D. (2008). Performance evaluation of models that describe the soil water retention curve between saturation and oven dryness. *Vadose Zone J.* 7:87–96.

Abstract

The objective of this work was to evaluate eight closed-form unimodal analytical expressions that describe the soil-water retention curve over the complete range of soil-water contents. To meet this objective, the eight models were compared in terms of their accuracy (root mean square error, RMSE), linearity (coefficient of determination R^2 and adjusted coefficient of determination R^2_{adj}), and prediction potential. The latter was evaluated by correlating the model parameters to basic soil properties. Retention data for 137 undisturbed soils from the Unsaturated Soil Hydraulic Database (UNSODA) were used for the model comparison. The samples showed considerable differences in texture, bulk density, and organic matter content. All functions were found to provide relatively realistic fits and anchored the curve at zero soil water content for the coarse-textured soils. The performance criteria were similar when averaged across all data sets. The criteria were found to be statistically different between the eight models only for the sandy clay loam soil textural class. An analysis of the individual data sets separately showed that the performance criteria were statistically different between the models for 17 data sets belonging to six different textural classes. We found that the Khlosi model with four parameters was the most consistent among different soils. Its prediction potential was also relatively good due to significant correlation between its parameters and basic soil properties.

4.1 Introduction

The unsaturated soil hydraulic properties are key factors governing the partitioning of rainfall and irrigation into soil water storage, evapotranspiration and deep drainage. The hydraulic properties involve the soil-water retention curve (SWRC), which relates the matric potential ψ with the soil-water content θ , and the hydraulic conductivity function. Discrete (θ , ψ) data sets can be either obtained from laboratory or field measurements, or predicted from other soil properties using pedotransfer functions, PTFs, or other approaches. Both methods yield discontinuous sets of θ - ψ data pairs within the range of matric potentials used for the measurements. For modeling purposes a continuous and smooth representation of the SWRC is preferred, which can be obtained by fitting a closed-form analytical expression to a discrete data set. To date various expressions appear in the literature to represent the SWRC (e.g., Brooks and Corey, 1964; van Genuchten, 1980; Kosugi, 1999). Most of the retention models are successful in the wet part of the SWRC. However, the dry part of the SWRC is equally important in different water related processes as mentioned earlier in Chapter 2.

There is hence a pressing need to accurately represent the SWRC for all matric potentials. In the last few years and decades, several attempts have been made to represent the complete retention curve (Ross et al., 1991; Campbell and Shiozawa, 1992; Rossi and Nimmo, 1994; Fayer and Simmons, 1995; Morel-Seytoux and Nimmo, 1999; Webb, 2000; Groenevelt and Grant, 2004, Khlosi et al., 2006). All of these models, except the equation by Groenevelt and Grant (2004), were tested on data reported by Campbell and Shiozawa (1992) and Schofield (1935), who measured water contents far below -1500 kPa. All models performed relatively well. However, when testing the Rossi and Nimmo (1994) sum model against data sets in which ψ ranged between -1 and -1500 kPa, Cornelis et al. (2005) found that this model behaved rather poorly compared to the van Genuchten (1980) and Kosugi (1999) models, despite of its physically realistic shape. In this paper we therefore compare eight closed-form unimodal analytical expressions to describe the SWRC over the complete range of soil-water contents. The comparison includes expressions by Campbell and Shiozawa (1992), Rossi and Nimmo (1994), Fayer and Simmons (1995), Webb (2000), Groenevelt and Grant (2004) and Khlosi et al. (2006), which were tested using a limited number of data pairs (e.g., UNSODA), as is most often the case in practice. Three statistical criteria were considered to define the

best models: accuracy (RMSE), linearity (R^2 and R^2_{adj}), and prediction potential (correlation between model parameters and basic soil properties).

4.2 Available Soil Water Retention Models

A large number of functions have been proposed over the years to describe the SWRC over the complete range of soil-water contents. Some of these functions are new while others are extensions of existing models. The extended functions are mostly modifications of the popularly used models by Brooks and Corey (1964), van Genuchten (1980) and Kosugi (1999). These functions, which will be referred to in this study as the BC, VG and K_{LN} models, respectively, were given earlier in Chapter 3 (see chapter 3, Eqs. [3.1], [3.2] and [3.3]).

Unfortunately, equations [3.1], [3.2] and [3.3] have considerable difficulty in representing the retention of water as the degree of saturation approaches zero, often giving an unrealistic path of the retention curve. To overcome this problem, various improvements have appeared in the literature. A first attempt to cover the complete retention curve was made by Ross et al. (1991). They modified Campbell's (1974) equation, which is identical to the power function of the BC model with the residual water content taken as zero, to extend the SWRC to oven dryness. Campbell and Shiozawa (1992) and Schofield (1935) measured water contents of soils ranging from sand to silty clay at matric potentials far below $\psi = -1500$ kPa. Inspection of their data suggests a log-linear relationship between the matric potential and the water content for matric potentials less than approximately -30 kPa and -1000 kPa for sand and silt loam respectively (the limiting values in their data set). Their silty clay soil showed an intermediate value. Based on these observations, Campbell and Shiozawa (1992) expressed the (θ, ψ) relationship in the low potential range as:

$$\theta = \theta_a \left(\frac{\ln(\psi_o) - \ln(\psi)}{\ln(\psi_o) - \ln(\psi_a)} \right) \quad [4.1]$$

or simply as:

$$\theta = \theta_a \left(1 - \frac{\ln(\psi)}{\ln(\psi_o)} \right) \quad [4.2]$$

where θ_a is a curve-fitting parameter representing the soil-water content at $\psi = -1$ m and ψ_o is the matric potential at oven-dryness. Note that Campbell and Shiozawa (1992) expressed their matric potentials in units of m and consequently θ_a in their equations corresponds to the soil-water content at -10 kPa or -1 m with a $\ln|\psi_a|$ value equal to zero. The matric potential at oven-dryness ψ_o depends on the temperature, pressure, and humidity at which the soil is dried. Assuming a logarithmic behaviour in the very dry range of the SWRC is consistent with the adsorption theory of Bradley (1936), which considers adsorbed molecules to build up in a layered film in which the net force of electrical attraction diminishes with increasing distance from the soil particle (Rossi and Nimmo, 1994). Incorporating [4.2] in a VG type model, Campbell and Shiozawa (1992) described the SWRC from saturation to oven-dryness as:

$$\theta = \theta_a \left(1 - \frac{\ln(\psi)}{\ln(\psi_o)} \right) + A \left(\frac{1}{1 + (\alpha \psi)^4} \right)^{1/m} \quad [4.3]$$

where A and m are curve-fitting parameters.

Rossi and Nimmo (1994) created a four-parameter sum model (RN1 model) and a three-parameter junction model (RN2 model) to represent the SWRC over the entire range from saturation to oven-dryness. Both models are based on the Campbell (1974) model with the residual water content taken as zero. Their four-parameter sum model (RN1), which consists of two functions joined at one point, was written as:

$$\theta = \theta_s \theta_I = \theta_s \left\{ 1 - \beta \left(\frac{\psi}{\psi_b} \right)^2 \right\} \quad \text{for } 0 \geq \psi \geq \psi_i$$

$$\theta = \theta_s \theta_{II} = \theta_s \left\{ \left(\frac{\psi_b}{\psi} \right)^\lambda - \left(\frac{\psi_b}{\psi_o} \right)^\lambda + \gamma \ln \left(\frac{\psi_o}{\psi} \right) \right\} \quad \text{for } \psi_i \geq \psi \geq \psi_o \quad [4.4]$$

where ψ_i is the soil-matric potential at the junction point where the two curves join, and β and γ are shape parameters. The term θ_I represents the Hutson and Cass (1987) parabolic curve that joints the Campbell function (1974) at the junction point ψ_i . The Ross et al. (1991)

correction is included in the expression for θ_{II} . Further, using data sets from Schofield (1935) and Campbell and Shiozawa (1992), Rossi and Nimmo (1994) showed that at very low soil-water content, the latter becomes proportional to the logarithm of the soil-matric potential, as can be recognized as well in θ_{II} . Equation [4.4] contains seven parameters. However, two of them can be determined from conditions that ensure continuity of both Eq. [4.4] and its first derivative with respect to ψ_i . Here we have chosen to explicitly determine β and γ as analytical functions of ψ_b , ψ_i , ψ_o and λ (Cornelis et al., 2005):

$$\beta = \psi_b^2 \frac{1 - \left(\frac{\psi_b}{\psi_i}\right)^\lambda + \left(\frac{\psi_b}{\psi_o}\right)^\lambda + \ln\left(\frac{\psi_o}{\psi_i}\right) \left(\frac{\psi_b}{\psi_i}\right)^\lambda}{\left[1 + 2 \ln\left(\frac{\psi_o}{\psi_i}\right)\right] \psi_i^2} \quad [4.5]$$

$$\gamma = \frac{2 - 2\left(\frac{\psi_b}{\psi_i}\right)^\lambda + 2\left(\frac{\psi_b}{\psi_o}\right)^\lambda - \left(\frac{\psi_b}{\psi_i}\right)^\lambda \lambda}{1 + 2 \ln\left(\frac{\psi_o}{\psi_i}\right)} \quad [4.6]$$

When setting ψ_o arbitrarily at -10^6 kPa (Ross et al., 1991; Rossi and Nimmo, 1994), and with Eqs. [4.5] and [4.6], the number of model parameters can be reduced to four.

The three-parameter junction model of Rossi and Nimmo (1994) (RN2 model) consists of three functions, which are continuous at the two points where the functions are joined:

$$\begin{aligned} \theta = \theta_s \theta_I = \theta_s \left\{ 1 - \beta' \left(\frac{\psi}{\psi_b} \right)^2 \right\} & \quad \text{for } 0 \geq \psi \geq \psi_i \\ \theta = \theta_s \theta_{II} = \theta_s \left(\frac{\psi_b}{\psi} \right)^\lambda & \quad \text{for } \psi_i \geq \psi \geq \psi_j \\ \theta = \theta_s \theta_{III} = \theta_s \gamma' \ln \left(\frac{\psi_o}{\psi} \right) & \quad \text{for } \psi_j \geq \psi \geq \psi_o \end{aligned} \quad [4.7]$$

where ψ_i and ψ_j are the soil-matric potential at the two junction points, and β' and γ' are shape parameters. To describe the shape of the SWRC near saturation, Rossi and Nimmo combined

the parabolic equation proposed by Hutson and Cass (1987) with the BC model (as described by θ_{II}). The equation for θ_{II} is a power law for ψ smaller than the air entry value ψ_b . The simple power law overestimates the water content at very low matric potentials. For this reason, a third part, θ_{III} , as proposed by Ross et al. (1991), was added to obtain water content of zero at ψ_o . In this case there are six parameters other than θ_s , as well as four conditions by imposing continuity of the global function and its first derivative at the two junction points. Four parameters (β' , ψ_i , ψ_j , and γ') can be calculated from analytical functions of the remaining two fitted parameters ψ_b and λ :

$$\beta' = 0.5 \lambda \left(\frac{2}{2+\lambda} \right)^{\left(\frac{2+\lambda}{\lambda} \right)} \quad [4.8]$$

$$\psi_i = \psi_b \left(\frac{2}{2+\lambda} \right)^{\left(\frac{-1}{\lambda} \right)} \quad [4.9]$$

$$\psi_j = \psi_o \exp\left(\frac{-1}{\lambda} \right) \quad [4.10]$$

$$\gamma' = \lambda \exp(1) \left(\frac{\psi_b}{\psi_o} \right)^{\lambda} \quad [4.11]$$

The three free parameters of the RN2 model are then θ_s , ψ_b and λ' . However, Rossi and Nimmo (1994) obtained better accuracy with their four-parameter sum model compared to their three-parameter junction model.

Fayer and Simmons (1995) further modified the Brooks-Corey and van Genuchten functions by replacing the residual water content with the adsorption equation of Campbell and Shiozawa (1992) to obtain:

Brooks-Corey:

$$\theta = \begin{cases} \theta_s & \text{for } \psi \geq \psi_b \\ \theta_a \left\{ 1 - \frac{\ln(\psi)}{\ln(\psi_o)} \right\} + \left[\theta_s - \theta_a \left\{ 1 - \frac{\ln(\psi)}{\ln(\psi_o)} \right\} \right] \left\{ \frac{\psi_b}{\psi} \right\}^{\lambda} & \text{for } \psi < \psi_b \end{cases} \quad [4.12]$$

van Genuchten:

$$\theta = \theta_a \left\{ 1 - \frac{\ln(\psi)}{\ln(\psi_o)} \right\} + \left[\theta_s - \theta_a \left\{ 1 - \frac{\ln(\psi)}{\ln(\psi_o)} \right\} \right] \left\{ \frac{1}{1 + (\alpha \psi)^n} \right\}^{\left(\frac{1-1}{n} \right)} \quad [4.13]$$

where θ_a is a curve-fitting parameter representing the soil-water content at $\psi = -1$ kPa, and ψ_o is the matric potential at oven dryness. Equations [4.12] and [4.13] are denoted here as FS1 and FS2, respectively.

Morel-Seytoux and Nimmo (1999) extended the BC model to oven-dryness using the three-parameter junction model (RN2). They divided the matric potential values into three levels: a low-potential level (from oven dryness to near field capacity), a middle level (field capacity to about air-entry matric potential), and a high level (air-entry matric potential to zero suction). For the low-potential level a slightly modified form of the RN2 model (θ_{III}) was selected. For the high-potential level the following algebraic relation was adopted:

$$\psi = \psi_m \left[1 - \frac{M(S_e - S_{em})}{S_{em}} - \frac{\{(M+1)S_{em} - M\}(S_e - S_{em})^2}{S_{em}(1 - S_{em})^2} + a_{MS}(S_e - S_{em})^2(1 - S_{em}) \right] \quad [4.14]$$

where S_e is effective saturation, $S_e = (\theta - \theta_r)/(\theta_s - \theta_r)$, S_{em} is effective saturation at the matching point, ψ_m is the corresponding matric potential, $M = 1/\lambda$ and a_{MS} is defined as:

$$a_{MS} = \frac{1}{S_{em}(1 - S_{em})} \left[\frac{M(M+1)}{2S_{em}} + \frac{\{(M+1)S_{em} - M\}}{(1 - S_{em})^2} \right] \quad [4.15]$$

In this case there are certain conditions of continuity and smoothness to be satisfied to reattach Eq. [4.14] with the traditional BC model for the middle potential range.

Webb (2000) proposed a new approach (W model) to combine the VG model with a dry region expression. This model does not necessitate refitting of experimental data and consists of two regions. Region 1 is an adsorption region described with a linear function on a semi-log plot of $\log(\psi)$ vs. soil-water content (θ). Region 2 is a capillary flow region described with the VG model (or any other desired function) in which any previous fitting parameter is retained. The linear relation between θ and ψ on a semi-log plot for the dry region was expressed by Webb (2000) as:

$$\theta = \frac{\log(\psi) - \log(\psi_m)}{\mu} + \theta_m \quad [4.16]$$

where

$$\mu = -\frac{\log(\psi_o) - \log(\psi_m)}{\theta_m} \quad [4.17]$$

and ψ_m and θ_m are, respectively, the matric potential and water content at the matching point, and μ is the slope of Eq. [4.16]. Different steps are required for determining the water content at matching point θ_m . First, we rewrite Eq. [4.16] as:

$$\psi = 10^{\mu(\theta - \theta_m) + \log(\psi_m)} \quad [4.18]$$

Second, the VG model is formulated in terms of θ and its slope is calculated. Third, the slope of the VG model is combined with Eq. [4.18] to give the intercept at ψ_o as:

$$\psi_o = 10^{\left(\log(\psi) - \theta_m \frac{d(\log(\psi))}{d\theta_m} \right)} \quad [4.19]$$

where

$$\frac{d(\log(\psi))}{d\theta_m} = -\log \left[\frac{1}{(\theta_m - \theta_r)(n-1) \left\{ 1 - \left(\frac{(\theta_m - \theta_r)}{(\theta_s - \theta_r)} \right)^{n/(n-1)} \right\}} \exp(1) \right] \quad [4.20]$$

Finally, the value of θ_m is determined from Eq. [4.19] by iteration. Since the value of ψ_o in Eq. [4.19] should equal -10^6 kPa one can give an initial estimate to θ_m and adjust this estimate until the proper value of ψ_o is obtained.

Recently, a new three-parameter model for the SWRC was developed by Groenevelt and Grant (2004) (GG model). This model anchors the curve at zero soil-water content using the log scale in a model for which the pF, defined as $\log(-\psi)$ with ψ expressed in unit of cm (Schofield, 1935), is the independent variable. They found the equation for the model to be capable of fitting pF curves with remarkable success over the complete range from saturation to oven-dryness. The soil-water content hence would be a function of pF as:

$$\theta(\text{pF}) = k_1 \left\{ \exp\left(\frac{-k_0}{6.9^\eta}\right) - \exp\left(\frac{-k_0}{\text{pF}^\eta}\right) \right\} \quad \text{for } \text{pF} > 0$$

$$\theta = k_1 \exp\left(\frac{-k_0}{6.9^\eta}\right) \quad \text{for } \text{pF} = 0 \quad [4.21]$$

where k_0 , k_1 and η are the three dimensionless parameters, and 6.9 is the pF value at oven-dryness according to the Schofield equation (1935).

Some of the functions described above utilize existing fitted curves. The Fayer and Simmons (1995) approach uses existing fitted curves to estimate the parameters in their modified SWRC expressions. Morel-Seytoux and Nimmo (1999) link up existing SWRCs with a dry region expression such that existing fitted curves can be directly employed. To utilize other existing fitted curves, Khlosi et al. (2006) used the adsorption equation of Campbell and Shiozawa (1992) to modify the Kosugi (1999) model. The new expression (KCGS model) hence combines the adsorption equation of the CS model with the K_{LN} model as:

$$\theta = \theta_a \left\{ 1 - \frac{\ln(\psi)}{\ln(\psi_o)} \right\} + \left[\theta_s - \theta_a \left\{ 1 - \frac{\ln(\psi)}{\ln(\psi_o)} \right\} \right] \frac{1}{2} \operatorname{erfc} \left\{ \frac{\ln\left(\frac{\psi}{\psi_{md}}\right)}{\sqrt{2} \sigma} \right\} \quad [4.22]$$

respectively, where ψ_{md} is the matric potential corresponding to the median pore radius, σ is a dimensionless parameter to characterize the width of the pore-size distribution, and "erfc" denotes the complementary error function.

4.3 Materials and methods

4.3.1 Sources of SWRC data

SWRC measurements for a selected set of undisturbed soils from different parts of the world for various soil types were used in this study. The set consisted of 137 undisturbed soils selected from the UNSODA database (Nemes et al., 2001). The 137 soils were selected using

the following criteria: (i) SWRC data were available from at least near saturation to, when possible, near oven dryness (some of the soils had measurements at matric potentials far below $\psi = -1500$ kPa), (ii) nearly all soil texture classes were represented (Figure 4.1), and (iii) their basic soil properties were known (notably clay, silt, sand, organic matter and bulk density). The organic matter content ranged from 0.7 to 214.0 g kg⁻¹, while bulk densities varied from 0.59 to 1.76 Mg m⁻³ (Figure 4.2).

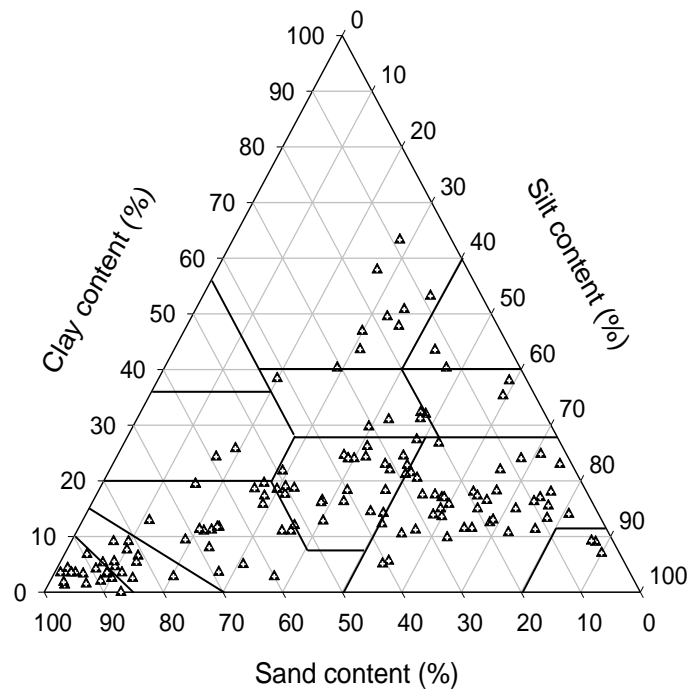


Figure 4.1. Texture distribution of 137 soils from UNSODA (Nemes et al., 2001) used in model development: sandy (sand, loamy sand), loamy (sandy loam, loam, silt loam, and silt), and clayey (sandy clay loam, silty clay loam, clay loam, sandy clay, silty clay, and clay) soils.

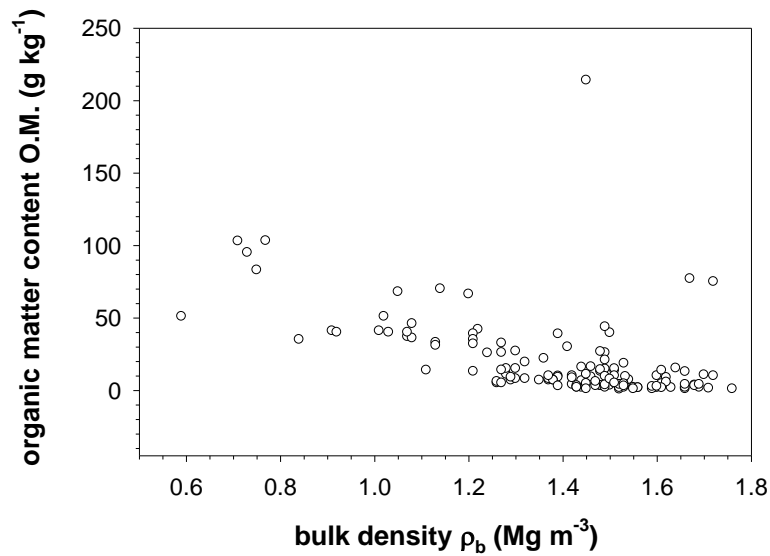


Figure 4.2. Variation of bulk density and organic matter content in the dataset.

4.3.2 Models for the data fit

From available SWRC models, we selected eight models: those proposed by Campbell and Shiozawa (1992), Rossi and Nimmo (1994) (RN1 and RN2 models), Fayer and Simmons (1995) (FS1 and FS2), Webb (2000), Groenevelt and Grant (2004), and Khlosi et al. (2006).

The parameters of the chosen models were obtained through the same procedures considered in Chapter 2 (see chapter 2, section 2.3.2). In order to avoid unrealistically large positive (or even negative) values for θ_a , θ_r , α , ψ_b in the CS, RN1, FS1, FS2, W and KCGS models, we constrained their parameters to $\alpha > 0$ for the CS model, $\psi_b > \psi_i$ for the RN1 model, $\theta_r \geq 0$ for the W model, and θ_a by the range of values found for the same texture class for the CS, FS2 and KCGS models.

4.3.3 Comparison methods

4.3.3.1 Goodness of fit statistics

Various statistical measures can be employed to compare the fitting accuracy of the SWRC models. In this study, we used as measures the root of the mean of squared errors (RMSE,

Eq. [2.15]), the coefficient of determination (R^2 , Eq. [2.16]), and the adjusted coefficient of determination (R^2_{adj}), which were calculated for each soil sample.

An additional measure of fit is the adjusted R-square, R^2_{adj} , (Neter et al., 1996), which is designed to take into account the number of parameters in the model. R^2_{adj} better reflects how the degree of correlation between observed and fitted data will change as additional parameters are added to or deleted from the model. R^2_{adj} is defined by:

$$R^2_{adj} = \frac{(N-1)R^2 - (p-1)}{N-p} \quad [4.23]$$

where p is the number of model parameters. The R^2_{adj} statistic can take on any value less than or equal to 1, with a value closer to 1 indicating a better fit.

In general, a model with more parameters can fit the observational data better. Although R^2_{adj} is generally one of the best indicators of the quality of the fit when adding additional parameters to the SWRC model, overparameterization should be avoided since it results in a non-identifiable model (i.e. a model leading to sample configuration probabilities identical to those of a simpler model with fewer parameters), in large variances of the estimated model parameters for similar soils, or in a high degree of correlation between the parameters (or low parameter uniqueness) if the number of observations is limited as is often the case with laboratory-determined SWRCs (Cornelis et al., 2005). Moreover, it is advantageous to minimize the number of model parameters when attempting to predict the SWRC from readily available data using pedotransfer functions.

In order to check if R^2 , R^2_{adj} , and RMSE of the 8 models are similar or different, a statistical test of significance is needed. Regarding the coefficients of determination or, equivalently, their corresponding coefficients of correlation, the initial hypothesis is that all 8 correlation coefficients are equal (Steel and Torrie, 1980):

$$H_0 : r_1 = r_2 = \dots = r_8 \quad [4.24]$$

with r being either R or R^2_{adj} , and 1 to 8 referring to the 8 models. The first step is then to transform the correlation coefficients into a new variable z such that:

$$z = \frac{1}{2} \log_e \left(\frac{1+r}{1-r} \right) \quad [4.25]$$

with \log_e being the natural logarithm. Since we have the same number of $\theta(\psi)$ pairs for the 8 models for a given data set (N ranging between 5 and 27), the variable χ^2_{obs} is next calculated as:

$$\chi^2_{obs} = (N - 3) \sum_{i=1}^p (z_i - \bar{z})^2 \quad [4.26]$$

where \bar{z} is the arithmetic mean of z_i , p is the number of models. This variable is subsequently compared to a theoretical value (χ^2_{theor}) at the 95 % confidence level having $(p - 1)$ degrees of freedom. Hypothesis [4.24] is rejected, meaning that the 8 correlation coefficients and thus the coefficients of determination are statistically different, when:

$$\chi^2_{obs} \geq \chi^2_{theor} \quad [4.27]$$

For RMSE, the test of significance is possible for variances. For this purpose we first square RMSE in order to calculate MSE, which represents variances. The hypothesis for homogeneity of variances, also called homoscedasticity (Hartley, 1950), is then:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_8^2 \quad [4.28]$$

where σ^2 is the MSEs, with 1 to 8 referring to the 8 models. The Hartley test is now applied by first computing the observed value:

$$H_{obs} = \frac{\sigma_{\max}^2}{\sigma_{\min}^2} \quad [4.29]$$

in which σ_{\max}^2 and σ_{\min}^2 are, respectively, the maximum and minimum values among the 8 MSEs. This observed value is compared to a theoretical one (H_{theor}) at the 95 % confidence level with p and $(N - 1)$ degrees of freedom. Hypothesis [4.28] is rejected, meaning that the 8 MSEs and thus the RMSEs are statistically different, when:

$$H_{obs} \geq H_{theor} \quad [4.30]$$

To conveniently compare the goodness of fit of the 8 models, RMSE, R^2 and R^2_{adj} were calculated separately for each model and for each data set. Next we calculated mean values of

these performance criteria for each of the 11 soil textural classes, as well as mean values for the whole data sets for the same performance criteria.

4.3.3.2 Comparing model parameters with basic soil properties

To provide further insight, the r_{sp} Pearson coefficient of correlation between the model parameters and several basic soil properties was computed for each soil sample. Mean values were calculated also for each model. The soil properties considered here were bulk density, ρ_b , organic matter content, OM, and sand, silt and clay content. The Pearson coefficient of correlation was used as a measure for the prediction potential of a model in that the closer r_{sp} is to either 1 or -1, the higher the prediction potential of the parameters in the model. High correlations between model parameters and basic soil properties are useful keys in developing reliable pedotransfer functions for the model parameters. The statistical significance of the correlation coefficients was tested using the same procedure as for R and R^2_{adj} (equations [4.24] to [4.27]). Results are reported only for the mean correlation coefficients over the whole data sets.

4.4 Results and discussion

4.4.1 Evaluation of the models

Parameter values for the different models and for different soil textural classes (Table 4.1) were obtained by curve fitting the models to the entire data set for each soil textural class. These values can serve as useful initial estimates when attempting to use one of the evaluated expressions.

Table 4.1. Average model parameter values for different soil textural classes. †

Model ‡	Par§	S	LS	SL	SiL	Si	L	SCL	SiCL	CL	SiC	C
No. of samples¶		18	11	19	41	3	23	4	1	5	2	10
No. of data pairs#		9	10	11	11	25	12	10	27	9	10	9
CS	θ_a	0.058	0.165	0.157	0.199	0.216	0.208	0.268	0.122	0.310	0.400	0.351
	A	0.312	0.238	0.247	0.212	0.189	0.268	0.177	0.327	0.274	0.143	0.170
	α	0.253	0.528	0.193	0.084	0.010	0.628	0.179	0.084	0.473	0.020	0.173
	m	5.446	4.887	10.090	8.265	0.964	17.537	21.072	11.887	9.675	6.813	6.123
RN1	θ_s	0.377	0.423	0.414	0.427	0.402	0.501	0.458	0.447	0.639	0.521	0.548
	ψ_b	3.718	1.741	7.734	11.552	27.977	0.433	0.609	0.245	0.886	1.272	0.739
	ψ_i	5.403	3.455	17.612	16.318	33.337	2.636	3.167	21.399	2.460	19.389	32.616
RN2	λ	1.248	1.064	0.468	0.239	2.173	10.073	8.019	1.939	1.644	0.026	0.956
	θ_s	0.378	0.426	0.415	0.425	0.403	0.498	0.454	0.447	0.639	0.521	0.552
	ψ_b	3.100	1.672	9.004	9.442	15.279	3.636	2.150	12.991	0.537	11.980	19.479
FS1	λ	0.653	0.289	0.229	0.209	0.336	0.130	0.111	0.137	0.138	0.106	0.101
	θ_a	0.041	0.127	0.118	0.160	0.040	0.290	0.283	0.189	0.123	0.470	0.304
	θ_s	0.373	0.419	0.417	0.430	0.401	0.499	0.471	0.456	0.638	0.527	0.569
FS2	ψ_b	3.266	2.095	4.439	4.988	13.741	0.723	0.553	2.802	0.497	2.642	0.163
	λ	0.945	0.620	0.288	0.279	0.384	0.279	0.189	0.135	0.173	0.268	0.137
	θ_a	0.066	0.137	0.169	0.145	0.078	0.300	0.213	0.347	0.376	0.328	0.390
W	θ_s	0.377	0.428	0.421	0.429	0.403	0.506	0.473	0.460	0.647	0.528	0.586
	α	0.227	1.564	0.350	0.414	0.039	11.053	1.707	1.968	2.690	0.183	503.170
	n	2.534	1.863	1.484	1.372	2.835	1.394	1.170	1.406	1.815	1.208	1.450
	θ_s	0.029	0.047	0.041	0.013	0.028	0.011	0.000	0.063	0.091	0.055	0.056
	θ_s	0.379	0.430	0.421	0.428	0.408	0.506	0.466	0.461	0.650	0.527	0.564
GG	α	0.232	0.772	0.236	0.183	0.041	0.977	0.420	0.170	1.386	0.098	1.472
	n	2.286	1.599	1.375	1.293	1.627	1.166	1.132	1.205	1.251	1.174	1.159
	s	112.482	52.626	27.580	26.285	52.909	17.046	15.604	17.792	14.197	11.218	14.531
	k_0	40.192	5.612	9.075	12.907	26.723	5.594	7.224	8.756	4.337	7.326	8.201
	k_i	0.385	9.684	0.834	9.269	0.434	22.130	104.901	41.471	56.121	2.742	126.554
KCGS	η	4.202	1.872	1.771	1.805	3.147	0.760	0.497	0.962	0.497	0.808	0.390
	θ_a	0.082	0.148	0.203	0.175	0.164	0.285	0.182	0.374	0.441	0.430	0.350
	θ_s	0.379	0.449	0.429	0.447	0.406	0.542	0.550	0.463	0.648	0.540	0.590
	ψ_{md}	10.386	7.045	54.609	108.626	59.118	47.248	54.908	177.609	2.503	187.658	947.421
	σ	1.086	2.427	2.148	2.855	1.291	3.890	6.901	4.588	1.432	3.621	4.973

† S, sand; LS, loamy sand; SL, sandy loam; SiL, silt loam; Si, silt; L, loam; SCL, sandy clay loam; SiCL, silty clay loam; CL, clay loam; SiC, silty clay; C, clay.

‡ CS, Campbell and Shiozawa (1992) model; RN1, Rossi and Nimmo sum model (1994); RN2, Rossi and Nimmo junction model (1994); FS1, Fayer and Simmons (1995) modified Brooks-Corey model; Fayer and Simmons (1995) modified van Genuchten model; W, Webb (2000); GG, Groenevelt and Grant (2004); KCGS, Khlosi et al. (2006).

§ Curve-fitting parameters: the units of the parameters are based on soil water content θ expressed in $m^3 m^{-3}$ and ψ in kPa.

¶ Number of data set in each soil textural class.

Average number of θ - ψ data pairs in each soil textural class.

Table 4.2 contains results of the statistical measures computed for each expression to compare their goodness of fit at three levels: each separate data set (137 values for each statistical measure), each soil textural class (11 values), and all of the data sets combined (1 value).

Table 4.2. Statistical measures of the models for 11 textural classes and all data sets.†

model‡	SI§	S	LS	SL	SiL	Si	L	SCL	SiCL	CL	SiC	C	TM
CS	RMSE	0.0092	0.0095	0.0092	0.0099	0.0100	0.0100	0.0100	0.0039	0.0100	0.0101	0.0105	0.0105
	R ²	0.993	0.991	0.992	0.990	0.989	0.990	0.989	0.999	0.989	0.988	0.986	0.985
	R ² _{adj}	0.989	0.986	0.987	0.984	0.984	0.985	0.984	0.998	0.983	0.982	0.978	0.977
RN1	RMSE	0.0096	0.0097	0.0111	0.0139	0.0141	0.0140	0.0140	0.0043	0.0138	0.0138	0.0140	0.0140
	R ²	0.993	0.992	0.988	0.979	0.979	0.979	0.978	0.998	0.978	0.978	0.975	0.974
	R ² _{adj}	0.989	0.987	0.981	0.968	0.968	0.968	0.966	0.998	0.967	0.967	0.962	0.961
RN2	RMSE	0.0147	0.0150	0.0149	0.0165	0.0167	0.0162	0.0162	0.0041	0.0160	0.0159	0.0160	0.0161
	R ²	0.987	0.986	0.983	0.975	0.974	0.975	0.974	0.998	0.975	0.975	0.971	0.971
	R ² _{adj}	0.983	0.981	0.978	0.968	0.967	0.968	0.967	0.998	0.968	0.968	0.963	0.963
FS1	RMSE	0.0108	0.0108	0.0127	0.0156	0.0159	0.0154	0.0151	0.0050	0.0149	0.0148	0.0146	0.0146
	R ²	0.992	0.991	0.985	0.976	0.975	0.976	0.976	0.998	0.977	0.977	0.978	0.978
	R ² _{adj}	0.987	0.986	0.976	0.963	0.963	0.964	0.965	0.997	0.966	0.966	0.966	0.966
FS2	RMSE	0.0097	0.0097	0.0104	0.0126	0.0128	0.0125	0.0123	0.0039	0.0122	0.0121	0.0120	0.0120
	R ²	0.993	0.992	0.990	0.984	0.983	0.983	0.984	0.999	0.984	0.984	0.984	0.984
	R ² _{adj}	0.988	0.988	0.983	0.975	0.975	0.975	0.976	0.998	0.976	0.976	0.976	0.976
W	RMSE	0.0104	0.0106	0.0109	0.0128	0.0130	0.0130	0.0130	0.0047	0.0129	0.0127	0.0127	0.0128
	R ²	0.992	0.991	0.989	0.983	0.983	0.982	0.982	0.998	0.983	0.983	0.983	0.982
	R ² _{adj}	0.988	0.986	0.983	0.975	0.975	0.974	0.974	0.998	0.974	0.974	0.973	0.973
GG	RMSE	0.0146	0.0149	0.0143	0.0149	0.0152	0.0146	0.0144	0.0086	0.0144	0.0142	0.0143	0.0142
	R ²	0.988	0.986	0.985	0.981	0.980	0.981	0.981	0.993	0.981	0.982	0.980	0.980
	R ² _{adj}	0.984	0.982	0.980	0.975	0.975	0.976	0.976	0.992	0.976	0.976	0.975	0.975
KCGS	RMSE	0.0105	0.0106	0.0105	0.0113	0.0115	0.0112	0.0111	0.0047	0.0110	0.0109	0.0109	0.0109
	R ²	0.992	0.992	0.990	0.988	0.988	0.988	0.988	0.998	0.988	0.988	0.988	0.988
	R ² _{adj}	0.987	0.986	0.985	0.982	0.981	0.982	0.982	0.998	0.982	0.982	0.981	0.981
Sig.¶	RMSE	No	No	No	No	No	No	No	Yes	No	No	No	No
	R ²	No	No	No	No	No	No	Yes	No	No	No	No	No
	R ² _{adj}	No	No	No	No	No	No	Yes	No	No	No	No	No

† S, sand; LS, loamy sand; SL, sandy loam; SiL, silt loam; Si, silt; L, loam; SCL, sandy clay loam; SiCL, silty clay loam; CL; clay loam; SiC, silty clay; C, clay; TM, mean values across the whole data sets.

‡ CS, Campbell and Shiozawa (1992) model; RN1, Rossi and Nimmo sum model (1994); RN2, Rossi and Nimmo junction model (1994); FS1, Fayer and Simmons (1995) modified Brooks-Corey model; Fayer and Simmons (1995) modified van Genuchten model; W, Webb (2000); GG, Groenevelt and Grant (2004); KCGS, Khlosi et al. (2006).

§ SI, statistical indices (mean texture class values).

¶ Test of the equality of the three performance criteria for the eight models: No, no statistical difference between the models; Yes, the performance criteria are statistically different for the eight models.

Based on the mean value of the three performance criteria over all data sets, Table 4.2 (last column) and Figure 4.3 show that the mean RMSE values varied between 0.0105 and 0.0161. The statistical analysis shows that the 8 models do not differ between each other regarding their RMSEs. The same table and figure also show that relatively high values were found for the mean R² (more than 0.97) and mean R²_{adj} (more than 0.96) for all models, thus indicating that all equations can be considered valid. Indeed, the test of significance for both R² and R²_{adj} further suggest that the 8 models do not differ. A more detailed analysis was done by considering each textural class separately (Table 4.2, columns 3 to 13). The test of significance shows that the 8 models do differ in terms RMSE only for the silty clay loam, but not for the remaining 10 textural classes.

Regarding R^2 and R^2_{adj} , the 8 models differ statistically for the class sandy clay loam, while they can be considered to give similar results for 10 remaining textural classes. Because of the differences in results between all data sets combined and those from each textural class, we further analyzed our data by considering each data set separately and focusing mainly on R^2_{adj} . The results show that for 17 out of the 137 data sets, the 8 models were statistically different, while they gave similar R^2_{adj} coefficients for the remaining 120 data sets classified as silty clay loam (1 data set), silty clay (2), silt (3), loamy sand (11), and clay loam (5). The 17 data sets were classified as sand (3 data sets out of 18), sandy loam (1 from 19), silty loam (2 from 41), loam (5 from 23), sandy clay loam (3 from 4), and clay (3 from 10). These results reflect the fact that mean values sometimes will not reveal differences when individual data sets are used. To further examine the goodness of fit of the various models, we plotted the distributions of RMSE, R^2 and R^2_{adj} for all soil samples in Figure 4.3. Each boxplot shows the median (solid line), mean (dotted line), the 25 and 75% percentiles (top and bottom of the box), the 10 and 90% percentiles (whiskers), and the outliers (circles). Notice that the fitting errors for FS2, W and KCGS models were smaller than those for the other models. This is because of the better representation of these models for all soil samples. Results suggest that the KCGS model is the most suitable for describing the observed data.

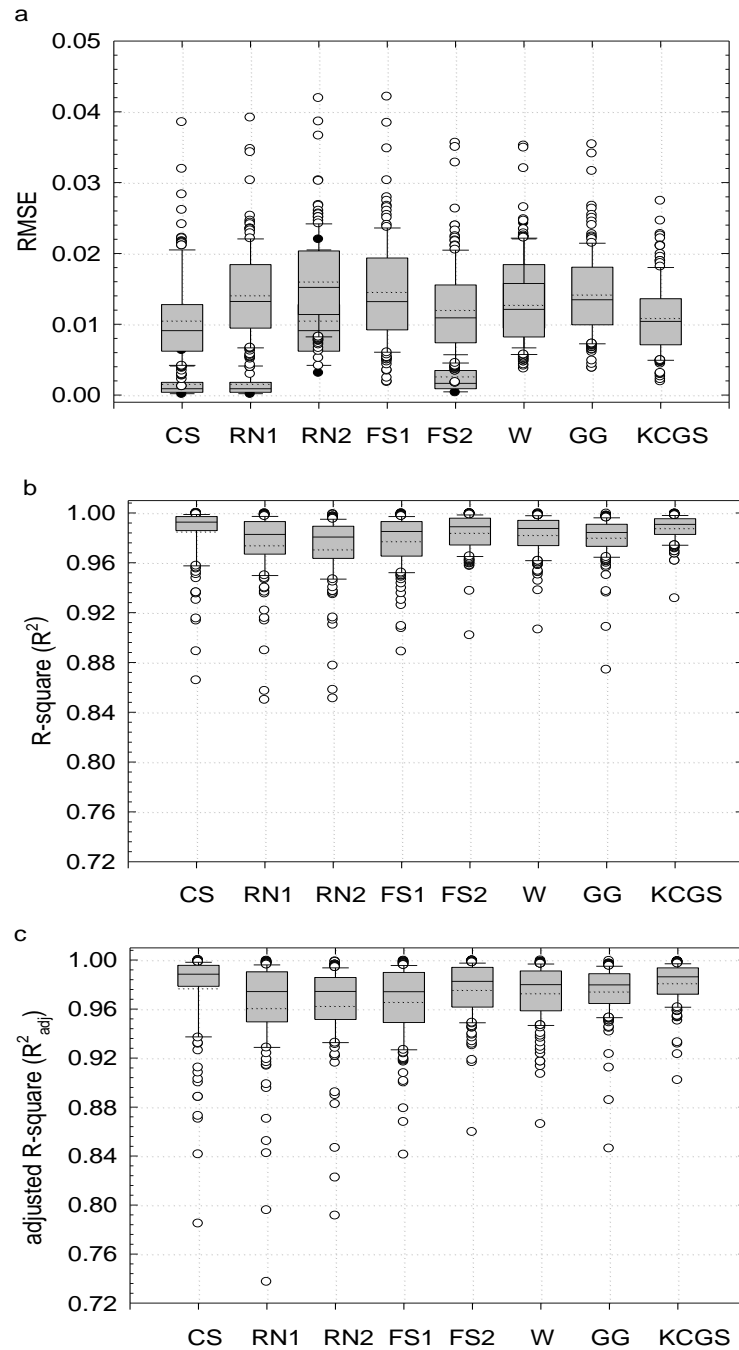


Figure 4.3. Boxplots of (a) RMSE, (b) R^2 , and (c) adjusted R^2 (R^2_{adj}) values for different soil-water retention curve models: Campbell and Shiozawa (CS); Rossi and Nimmo four-parameter sum (RN1); Rossi and Nimmo three-parameter junction (RN2); Fayer and Simmons modified Brooks–Corey (FS1); Fayer and Simmons modified van Genuchten (FD2); Webb (W); Groenevelt and Grant (GG); and Khlosi et al. (KCGS). The box plots summarize the distribution of RMSE, R^2 , and R^2_{adj} . The horizontal full line in each box signifies the median value and the mean in a dotted line, whereas the bottom and top of the box represent the 25th and 75th percentile. The whiskers display the 10th and 90th percentile, while the points indicate the outliers.

Table 4.3 lists the Pearson correlation coefficients, r_{sp} , between all parameters and the basic soil properties. The KCGS and FS2 models produced the highest correlation values between θ_r (or θ_a), and ρ_b and clay content, while most models showed high correlation between θ_s , and ρ_b and clay content. Notice that θ_s was highly correlated to ρ_b and to a lesser extent to clay content while the opposite was true for θ_a . The other parameters, which mostly affect the shape of the SWRC, showed much lower correlations, except for the KCGS, W, FS1, RN2 and GG models. This was to be expected for the RN2 and GG models since RN2 does not contain θ_r (or θ_a), while GG does not have θ_r (or θ_a) and θ_s . Relatively high values were also obtained for the FS1 and W models. The KCGS model showed high correlations for its additional parameters, as well as for θ_a and θ_s . Other models such as CS, RN1, and FS2 showed good correlation with at least one additional parameter.

The results above indicate that the KCGS model performed best. This model performed as good as the other models in terms of the goodness-of-fit, and showed significant correlation between all of its parameters and the basic soil properties. Compared with the KCGS model, CS expression exhibited two main disadvantages. One is that the correlation between model parameters and basic soil properties was far less than that of the KCGS model for all parameters. A second disadvantage is related to the fact that the boxplot of the RMSE, R^2 and R^2_{adj} values showed many outliers for CS as compared to KCGS. This suggests that CS is less consistent when applied to different soils. The FS2 model, which showed the same fitting performance as the KCGS and CS models, did not represent the dry range well, without having (θ, ψ) data below a matric potential of -100 kPa (Khlosi et al., 2006). Although the W model does not necessitate the refitting of observed data, this model does not require many iterations to find the matching point.

Table 4.3. Pearson correlation coefficient between model parameters and basic soil properties. §

model†	θ_r/θ_a vs ρ_b	θ_r/θ_a vs clay	θ_s/A vs ρ_b	θ_s/A vs clay	$a‡$ sand/clay¶	$a‡$ silt/clay#	$b††$ sand/OM ‡‡	$b††$ silt/clay/ ρ_b §§	$c¶¶$ sand	$c¶¶$ silt/clay##
CS	<u>-0.28</u>	<u>0.64</u>	<u>-0.37</u>	<u>-0.24</u>	0.12	-0.12	<u>0.17</u>	-0.33		
RN1	-	-	<u>-0.79</u>	<u>0.50</u>	<u>-0.21</u>	<u>0.33</u>	<i>-0.04</i>	0.07	<u>-0.26</u>	<u>0.23</u>
RN2	-	-	<u>-0.79</u>	<u>0.51</u>	<u>0.27</u>	<u>0.27</u>	<u>0.52</u>	<u>-0.48</u>	-	-
FS1	<u>-0.26</u>	<u>0.50</u>	<u>-0.79</u>	<u>0.56</u>	<u>-0.28</u>	<u>0.34</u>	<u>0.51</u>	<u>-0.42</u>	-	-
FS2	<u>-0.37</u>	<u>0.69</u>	<u>-0.78</u>	<u>0.55</u>	-0.09	<u>0.30</u>	<u>0.35</u>	<u>-0.29</u>		
W	0.05	0.11	<u>-0.79</u>	<u>0.52</u>	<u>0.24</u>	-0.12	<u>0.50</u>	<u>-0.42</u>	<u>0.60</u>	<u>-0.52</u>
GG	-	-	-	-	0.16	<i>-0.15</i>	<u>0.27</u>	<u>0.43</u>	<u>0.45</u>	<u>-0.56</u>
KCGS	<u>-0.39</u>	<u>0.67</u>	<u>-0.73</u>	<u>0.49</u>	<u>-0.18</u>	<u>0.39</u>	<u>-0.31</u>	<u>0.41</u>	-	-

† CS, Campbell and Shiozawa (1992) model; RN1, Rossi and Nimmo sum model (1994); RN2, Rossi and Nimmo junction model (1994); FS1, Fayer and Simmons (1995) modified Brooks-Corey model; Fayer and Simmons (1995) modified van Genuchten model; W, Webb (2000); GG, Groenevelt and Grant (2004); KCGS, Khlosi et al. (2006).

‡ a corresponds to α , ψ_b , ψ_{b_1} , ψ_{b_2} , α , α , k_0 and ψ_b in the CS, RN1, RN2, FS1, FS2, W, GG and KCGS model, respectively.

§ All values that are underlined indicate a significant correlation at the 0.05 level.

¶ All a values are correlated to sand content, except for the value in italics, which is correlated to clay content.

All a values are correlated to silt content, except for the value in italics, which is correlated to clay content.

†† b corresponds to m , λ , λ , λ , n , n , k , and σ in the CS, RN1, RN2, FS1, FS2, W, GG and KCGS model, respectively.

‡‡ All b values are correlated to sand content, except for the value in italics, which is correlated to OM.

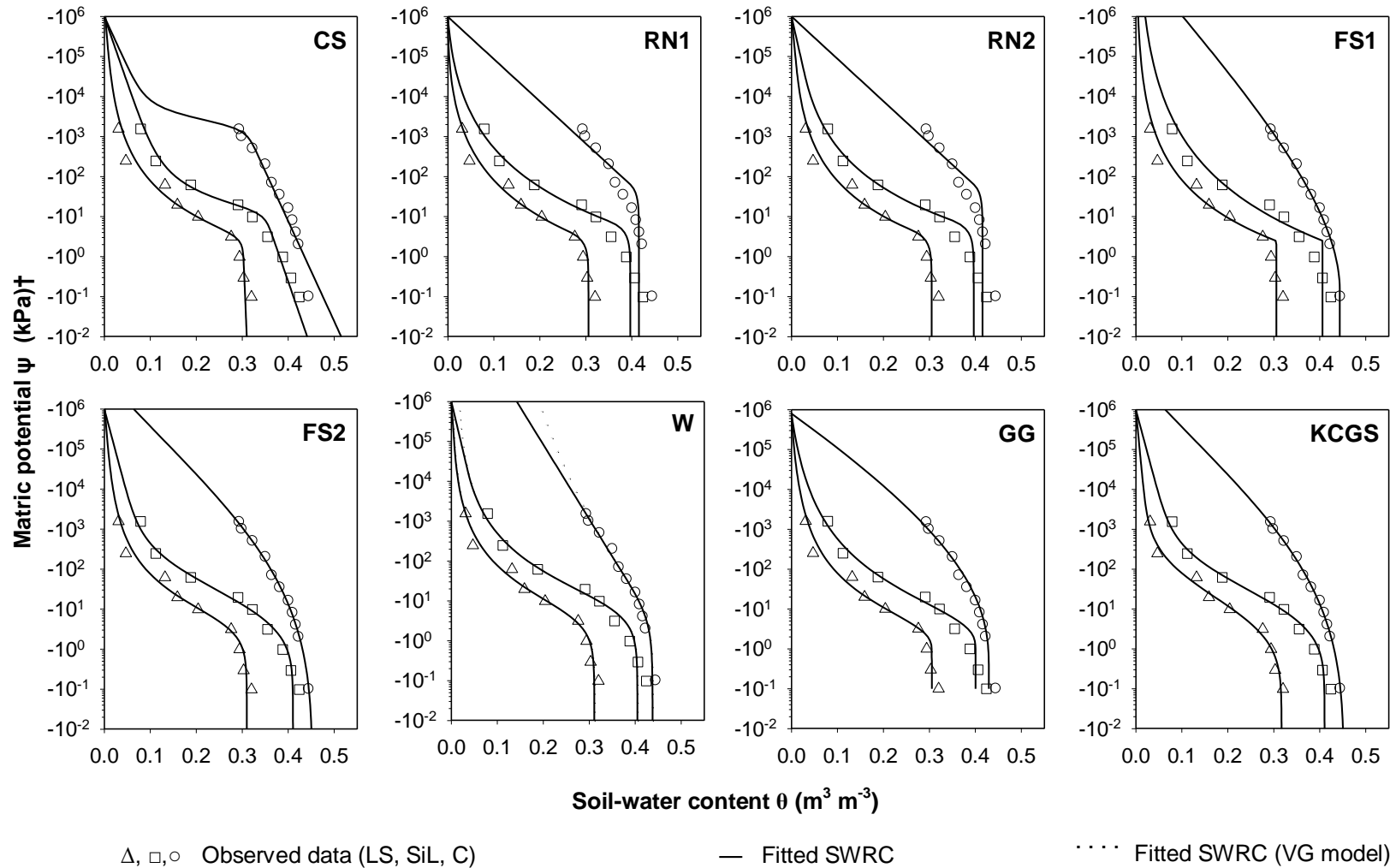
§§ All b values are correlated to silt content, except for the value in italics, which is correlated to clay content, and the value in bold, which is correlated to ρ_b .

¶¶ c corresponds to ψ_i , s and η in the RN1, W and GG model, respectively.

All c values are correlated to silt content, except for the values in italics, which is correlated to clay content.

4.4.2 Performance of the models

The performance of the eight closed-form analytical expressions is further demonstrated below by fitting them to soil-water retention data of a relatively coarse-textured soil (loamy sand, $\rho_b = 1.68 \text{ g m}^{-3}$, O.M. = 3 g kg^{-1}), a medium-textured soil (silt loam, $\rho_b = 1.51 \text{ g m}^{-3}$, O.M. = 5 g kg^{-1}), and a fine-textured soil (clay, $\rho_b = 1.47 \text{ g m}^{-3}$, O.M. = 6 g kg^{-1}). Figure 4.4 indicates that all models performed very well for the loamy sand in that they produced comparatively realistic fits and anchored the curves at zero soil-water content. We can notice that still, the KCGS model showed the best fit for all matric potentials. The correspondence between observed and fitted SWRCs for the silt loam exhibited deviations for the RN1, RN2, FS1, W and GG models. The RN1 and RN2 models showed a poor match near saturation. This is because both models use the Hutson and Cass expression (1987) at high matric potentials, which is less flexible because of its parabolic shape. The RN1 and RN2 models further performed rather poorly in the dry range since these functions require many data points to provide a smooth match at the junction points. The FS1 model mostly missed the shape of the data near saturation due to its discontinuous character, which is an inherent feature of the original BC model. Figure 4.4 additionally shows that the GG model did not accurately match several points near saturation not unlike some of the other models. This can be attributed to the discontinuity at $\psi = 1 \text{ cm}$ ($\text{pF} = 0$) which decreases the flexibility of the curve in that region. By contrast, CS, FS2, W and KCGS showed very good fits to the silt loam data although CS seems to have a less realistic shape (linear) in the wet region. Slightly different results were obtained for the clay soil.



† The curve in case of GG model is only defined for potential values lower than -0.1 kPa (or -1 cm)

Figure 4.4. Observed and fitted soil-water retention curves for a loamy sand, silt loam and clay.

The RN1, RN2 here again showed poor fits as explained earlier. By keeping the oven-dryness pressure (ψ_o) as a free parameter one can improve their fits but with the possibility of producing oven-dryness values lower than -10^6 in fine-textured soils. The GG model now showed a better match to the data, because it is more flexible when less data are presented near saturation. For the same reason, the discontinuous character of the FS1 model did not seem to be problematic for the clay soil. On the other hand, the CS model showed a biased curve since θ_s was not considered here. The FS2, W, and KCGS models showed excellent fits also for the clay example.

4.5 Conclusions

Using 137 soil samples from the UNSODA database, we compared eight closed-form unimodal analytical expressions for the soil-water retention curve. The expressions were evaluated in terms of their goodness-of-fit using different statistical indices. All eight models defined the soil-water content vs. soil-matric potential relationship below the residual water content. The performance of the models in terms of matching the data varied greatly depending upon with the degree of aggregation or desegregation of the data: individual SWRC data sets, averages over individual soil textural classes, or using the overall mean. Our results show that lumping of the data without considering their textural class provided similar results for the 8 models in terms of their ability to fit observed data. However, when the textural class was taken into account, the 8 models were found to perform differently for the sandy clay loam class, while being not statistically different for the remaining 10 textural classes. An analysis for each data set separately showed that the 8 models behaved differently for 17 of the individual 137 data sets, representing 6 different textural classes. The Khlosi et al. model with four parameters was found to be the most consistent for the different soils. Moreover, its prediction potential was relatively good because of significant correlation between its parameters and basic soil properties. We hence do recommend this analytical formula for reliable modeling of the whole range of soil-water contents of unsaturated soils, which can vary substantially in bulk density, soil texture, and organic matter content.

Chapter 5 Impact of particle-size distribution changes associated with carbonates on the predicted soil-water retention curve

Based on:

Khlosi M., Cornelis W.M., and Gabriels D. (2008). Analyzing the effects of particle-size distribution changes associated with carbonates on the predicted soil-water retention curve. In D. Gabriels et al. (2008). Proceedings of the Conference on Desertification (UGent), Combating desertification: assessment, adaptation and mitigation strategies. International Centre for Eremology – Belgian Development Cooperation. P. 100-105.

Abstract

Particle-size distribution (PSD) is an essential soil property which correlates well with many other soil properties. Accurate determination of PSD is needed to predict more difficult-to-measure soil hydraulic properties such as the soil-water retention curve and hydraulic conductivity. Differences in methodologies for PSD assessment include removal or non removal of cementing materials such as calcium carbonate. The objective of this study was to investigate the influence of sample pre-treatment on sand, silt, and clay fractions (texture) and hence on the predictability of the soil-water retention curve of dryland soils. 72 soil samples have been collected from most of the agro-climatic zones of Syria. The pipette method was used for determination of particle fractions. Two procedures with different pre-treatment were applied. In a first one, carbonates were removed by hydrochloric acid as is typically done in particle size analysis worldwide. In the alternative one, carbonates were not removed, a procedure more typical for calcareous soils. Great variability between both procedures was found in the sand, silt, and clay fractions. Using the basic concept of shape similarity between the soil-water retention curve and the cumulative PSD function, we tested the prediction capability of two methods. The results indicate that for our soils only texture without removing carbonates can be translated into the soil-water retention curve. When carbonates were removed from the soil, contents of soil textural fractions showed no link to the soil-water retention curve. Reasonable results in terms of R^2 and root mean square error of multiple regression equations for predicting water content at various matric potentials were only achieved when carbonates were not removed. The importance of not removing carbonates was further supported when applying the PTF of Vereecken and coworkers. Non removal of carbonates showed much better prediction accuracy of the soil-water retention curve than removing it. The advantages associated with non-removal of carbonates include decrease in time and work, and allowing accurate measurements of calcareous soil samples. Further, the derived point PTFs provided better accuracy than those reported by Ghorbani Dashtaki and coworkers. This study clearly shows the risks of removing carbonates when using or developing PTFs for calcareous soils.

5.1 Introduction

Numerical models for simulating water flow and solute transport in unsaturated–saturated soil systems are enjoying considerable popularity as a tool for soil survey interpretations. Their success and reliability, however, are critically dependent on accurate information of soil hydraulic properties. The most important properties are the soil-water retention curve (SWRC) and hydraulic conductivity characteristics. Direct measurement of soil hydraulic properties is difficult, tedious to accomplish and expensive by currently available methods. When such data are not available, pedotransfer functions (PTFs) (Bouma, 1989) which utilize physical or empirical relations between soil hydraulic properties and other easily and cheaply measured properties can be used as alternative method. In this context, particle-size distribution (PSD) is the most important key predictor to most soil hydraulic PTFs. Other commonly used soil physical properties for PTFs prediction are organic carbon (OC) and bulk density (ρ_b). Precise and accurate determination of PSD is, therefore, needed and required to provide good representation of soil hydraulic properties.

Several methods exist to determine PSD. Methods not only differ in how pretreated soil samples are analysed (e.g., sieve-pipette method, hydrometer method, laser diffraction method), but also in how samples are pretreated. An important difference in pretreating samples is the removal or non-removal of cementing materials such as calcium carbonate. Soil carbonates are a very common mineral in soils of dryland areas. The carbonates may have been present originally in the soil's parent material; they may have formed in the upper horizons, or in the whole profile. Because of low rainfall, percolating water through the profile is not sufficient to remove the carbonates present in the parent material, or that is produced by reaction between carbonic acid and the calcium hydrolyzed from mineral material (Harper, 1957). The use of PSD as the first and most basic input parameter of PTFs raises the question of how different pre-treatments affect their prediction quality. Soil carbonates usually act as binding agent, and it has therefore been general practice to remove all carbonates by HCl treatment. Francis and Aguilar (1995) recommended that calcareous rich soil should be pre-treated for CaCO_3 removal prior to particle-size analysis. However, when carbonate content is substantial in soils of dryland areas (up to values of 60% of the total soil in our case) not all of it acts as cementing agent. Most of it is actually present in the soil as distinct non-clay minerals such as calcite, dolomite, magnesite, aragonite and vaterite

particles, possessing their own inherent PSD, ranging from very fine clay-like powder to coarser, fine silt-like deposit.

For predicting SWRC using PTFs, a multitude of methods can be used, including Artificial Neural Networks (Pachepsky et al., 1996), Support Vector Machines (Lamorski et al., 2008), Nearest Neighbours (Nemes et al., 2009), Regression Trees (Pachepsky and Rawls, 2003) among others. At present, the most widely adopted method is, however, still Multiple Linear Regression (MLR) which has provided good predictive models for the SWRC formulations (Gupta and Larson, 1979; Rawls and Brakensiek, 1982; Vereecken et al., 1989; Scheinost et al., 1997; Wösten et al., 1999, 2001; Merdun et al., 2006; Ghorbani Dashtaki et al., 2010; Minasny and Hartemink, 2011; among others). It is worth mentioning here that PTFs which predict soil-water content at specific matric potential (point PTFs) are more accurate than parametric ones because the estimated water retention from single-point regression has less uncertainty compared with using parametric PTFs (McBratney et al., 2002).

The objective of this study was to see whether soil texture with non-removal of carbonates may have some advantages in predicting the SWRC of calcareous soils. Our hypothesis was that non-removal of carbonates and thus considering them as inherent to the soil's mineralogy improves the predictive power of PTFs. To test this hypothesis, we tried to develop point PTFs based on two PSD procedures (with and without removing carbonates) using MLR. We also used the PTF of Vereecken et al. (1989), which have shown to be applicable in environments other than those in which they were calibrated (Romano and Santini, 1997) to validate the two PSD procedures. The derived point PTFs were finally compared with those reported by Ghorbani Dashtaki et al. (2010).

5.2 Materials and methods

5.2.1 Study area and soils

All soil samples have been collected in Syria, Western Asia, on the eastern shore of the Mediterranean Sea between latitudes 32-37° north and longitude lines 35-42° east. Syria has a characteristically Mediterranean climate with rainy winters and hot rainless summers. One of the major tasks in developing hydraulic PTFs is establishing a soil hydraulic database,

which is a collation of data that are obtained during soil survey such as soil texture, bulk density, organic matter, and soil-water retention curve. In Syria there are no published data available and the collation of a national database is not complete. Farmers on the other hand have developed over time their own methods of water use estimation. Their knowledge is based on experience, going back for ages. Because the farmers' approach to soil and water management is an empirical one, there are no examples available to guide their decisions. Since it was impossible within the framework of this study to measure basic soil properties and soil hydraulic properties for all types and regions, a selection had to be made from the beginning. Two criteria have been emphasized to obtain a relevant set of soils. First, the selected sampling sites should cover most of the agro-climatic zones of Syria. Secondly, the selected set of soil types should contain as much as possible of the variance of the texture and profile developments. The study area was, therefore, located in the rural areas of northwest Syria between the parallels 35°28' and 36°53' North and meridians 36°19' and 37°54' East, covering a surface of 7126 km² (Figure 5.1) (UTM coordinates: between 290,000 East and 4,084,000 North UTM zone 37 and 402,000 East and 3,934,000 North). It includes the Kurd Dagh block mountains in the northwest, the gently undulating plains in the central part and alluvial-colluvial plains and basalt plateaux with the Salt Lake 'Jabboul' in the southeast. The climate of the study area is typical East-Mediterranean, exhibiting hot, dry summers with an average temperature of 28.2°C in July, and cool, wet winters with an average temperature of 5.6°C in January. Zone 1 in Figure 5.1 has an annual rainfall greater than 350 mm. It is characterized by rolling landscape with shrubland, degraded forest, rock outcrops, and tree crops on the slopes and hilltops and cereal cultivation in the valleys. Olive is the main agricultural product of the region. Other perennials are apricot, almond, cherry and peach. Zone 2 has an annual rainfall between 250 and 350 mm. It is characterized by tree crops (olive, fig, pistachio and grapes) and rainfed field crops (wheat, chickpea and lentil) with irrigated crops (cucumber, potato, watermelon and eggplant). Soils in zones 1 and 2 were classified as Cambisols according to World Reference Base for Soil Resources (WRB) (FAO, 1998) and/or Inceptisols according to U.S. Soil Taxonomy (Soil Survey Staff, 1999). Cambisols are developed in medium and fine-textured materials derived from a wide range of rocks, mostly in alluvial, colluvial and aeolian deposits. Zone 3 has an annual rainfall of 250 mm. It is characterized by rainfed field crops (wheat and barley) with irrigated crops (tomato,

zucchini, and melon) and livestock. Zone 4 has an annual rainfall between 200 and 250 mm. It is characterized by rainfed field crops (mainly barley) and livestock.

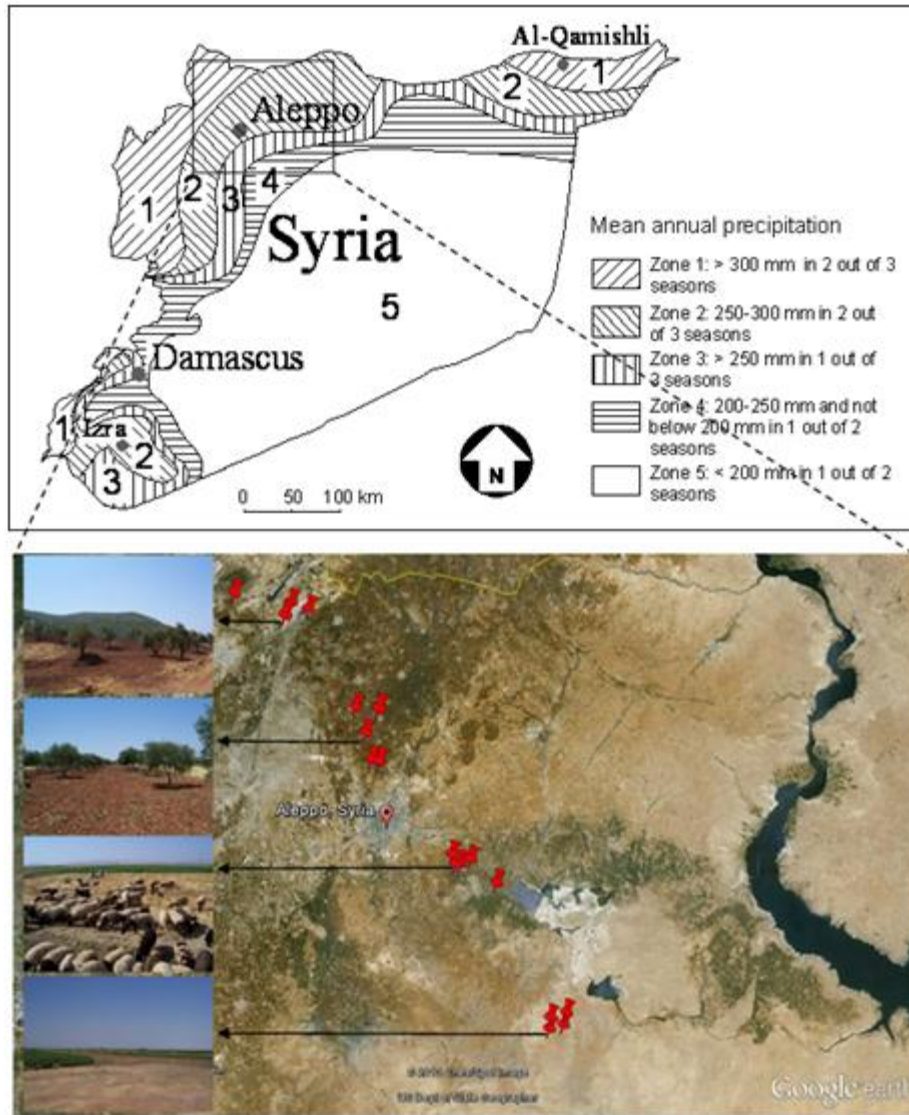


Figure. 5.1. Map of study area showing the distribution of sampling sites. Zone 1 is characterized by shrubland, degraded forest, rock outcrops and tree crops; zone 2 is characterized by tree crops and rainfed field crops with irrigated crops; zone 3 is characterized by rainfed field crops with irrigated crops and livestock; zone 4 is characterized by rainfed field crops and livestock.

The dominant soil types in zone 3 and 4 are Calcisols and Gypsisols according to World Reference Base for Soil Resources (WRB) (FAO, 1998) and/or Aridisols according to U.S. Soil Taxonomy (Soil Survey Staff, 1999). Calcisols are developed in mostly alluvial,

colluvial and aeolian deposits of calcareous weathering material. Gypsisols are developed in mostly unconsolidated alluvial, colluvial and aeolian deposits of base-rich weathering material. Most profiles have free carbonates.

5.2.2 Field and laboratory measurements

A total of 72 undisturbed soil samples (Kopecky rings, 5 cm height, 5.3 cm diameter, 100 cm³ volume) were collected from these four agro-climatic zones which cover a wide range of soil types (Figure 5.2). The Kopecky rings were pressed gradually, slowly and evenly into the soil. Special attention was paid to avoid soil compaction. The soil around the ring was then removed and the ring containing the soil sample was withdrawn. The end of the sample was not cut with a knife but was removed with gently brushing to conserve the natural structure of the soil. Soils were first examined immediately on site for texture using the texture-by-feel technique (Thien, 1979). Accordingly, soil texture ranged from sandy loam to clay, with loam, clay loam and clay as the most common textural classes.

The undisturbed soil samples were used first to determine the SWRC and bulk density following the procedures outlined in Chapter 2. Bulk densities ρ_b varied from 1 to 1.8 Mg m⁻³. The samples' SWRC was determined at eight matric potentials. To reduce possible unresolved variance, the other basic soil properties were determined directly on the same samples after the SWRC was established using disturbed samples. Organic matter content ranged from 0 to 3.5% and was determined by means of the Walkley and Black (1934) method. Soil carbonate content ranged from 2.3 to 64.5% with a mean value of 27.6% and was determined by back-titration approach (Nelson, 1982). Determination of the PSD was undertaken only on the fine earth fraction (< 2 mm) using the sieve-pipette method (Gee and Bauder, 1986) which is the standard analytical method (ISO 11277). As a first step, the soil samples were dispersed in an aqueous suspension. The organic matter was removed by oxidizing it with hydrogen peroxide, a strong oxidizing agent. Then two procedures with a different pre-treatment process were applied. In a first one, carbonates were removed by hydrochloric acid as is typically done in pretreating samples for PSD analysis (the PSD_{-C} method). In an alternative one, carbonates were not removed (the PSD_{+C} method). The results of these two methods were used to predict the SWRC. Bulk density and organic matter content of the real soil with carbonates are considered in this study (they were taken as independent variable).

5.2.3 Multiple linear regression analysis

Multiple linear regression analysis was first introduced by Sir Francis Galton in the latter part of the 19th century. MLR is a method used to model the linear relationship between a dependent variable (or predictand) and one or more independent variables (or predictor). MLR is based on least squares: the model is fit such that the sum-of-squares of differences of observed and predicted values is minimized. Nowadays, the methodology is still widely used for developing PTFs to predict the SWRC from other soil properties. We selected the soil-water content at different matric potentials $\theta(\psi)$ as the predictand and the basic soil property variables as predictors (i.e., sand, silt, clay content, organic carbon content and bulk density). The model is fit to a set of soils (the calibration set, $n = 49$) for which $\theta(\psi)$ and basic soil properties data overlap. In the process of fitting, or estimating, the model, statistics are computed that summarize the accuracy of the regression model for the calibration set. The performance of the model on data not used to fit the model is checked in some way by a validation process ($n = 23$). Basic soil properties data from the calibration set are substituted into the prediction equation to get a *reconstruction* of the predictand. The reconstruction is a “prediction” in the sense that the regression model is applied to generate estimates of the predictand variable outside the set used to fit the data.

The model expresses the value of a predictand variable as a linear function of one or more predictor variables and an error term, and takes the form

$$y_i = b_0 + b_1x_{i,1} + b_2x_{i,2} + \dots + b_Kx_{i,K} + e_i \quad [5.1]$$

where y_i is the predictand for soil i , $x_{i,k}$ is the value k^{th} of predictor in soil i , b_0 is the regression constant, b_k is the coefficient on the k^{th} predictor, K is the total number of predictors, and e_i is the error term. The model (5.1) is estimated by least squares, which yields parameter estimates such that the sum of squares of errors is minimized. The resulting prediction equation is

$$\hat{y}_i = \hat{b}_0 + \hat{b}_1x_{i,1} + \hat{b}_2x_{i,2} + \dots + \hat{b}_Kx_{i,K} \quad [5.2]$$

where the variables are defined as in (5.1) except that “ $\hat{}$ ” denotes estimated values. As few variables as possible should be chosen because of the risk of overfitting the data. Hence, stepwise multiple regression with maximum R^2 improvement was invoked to select the

variables which correlated best with $\theta(\psi)$. According to Muller and Fetterman (2002), using these criteria for a stepwise modeling strategy provides a better approximation to the all-possible-regressions strategy, which requires fitting all possible models. The stepwise regression model starts without any independent variables and chooses the best predictors one by one, on the basis of statistical significance. For each coefficient, the T -ratio tests whether the value of the coefficient is zero, and if its p -value is less than 0.05, the calculated value is considered statistically significant. A large T -ratio implies small p -value. If the p -value was larger than 0.05 for the T -statistic, it was concluded that the independent variable was not statistically different from zero and the variable was excluded from the regression equation; variables with p -values less than or equal to 0.05 were kept. Similarly, the F -statistic was used to test whether the values of the coefficients for the entire equation are equal to zero, and if the p -value for the F -statistic is less than 0.05, the multiple regression equation is statistically significant. While T -statistic examines the significance of individual independent variables, the F -statistic examines the significance of all independent variables collectively. To assess statistical validity of the predictive equation, we also computed the root mean squared error (RMSE) and coefficient of multiple determination (R^2). The latter provides the amount of variability in observations that is explained by the independent variables. The stepwise MLR was accomplished in MATLAB (Mathworks, 2007).

5.2.4 Evaluation criteria

Various pedotransfer functions appear in the literature to predict the SWRC (Saxton and Rawls, 2006). In an earlier study of Cornelis et al. (2001) in which nine PTFs to predict the soil-water retention curve were compared, it was shown that the PTF of Vereecken et al. (1989) was the most accurate one for arable soils in Belgium. It was shown in other studies to be rather accurate for other regions as well. Tietje and Tapkenhinrichs (1993) and Romano and Santini (1997) reported an overall satisfactorily behaviour of the pedotransfer published by Vereecken et al. (1989). Kern (1995) showed that it performed the best among six other PTFs for a large number of different types of soils sampled throughout the USA. Therefore, the Vereecken et al. PTF (1989) was used to test the hypothesis in that carbonates should not be removed. This PTF predicts parameters of the van Genuchten equation: residual (θ_r) and saturated soil-water content (θ_s), a scaling parameter (α) related to ψ^{-1} , a parameter related to the curve's slope at its inflection point (n), and an empirical constant (m).

In order to quantify the prediction accuracy of the two methods for a given soil, the estimated SWRCs were compared with the experimental ones using three complementary indices: the mean difference MD ($\text{m}^3 \text{ m}^{-3}$), the root of the mean squared difference RMSD ($\text{m}^3 \text{ m}^{-3}$) between the measured and estimated SWRC, and the Pearson correlation coefficient r . Assume the measured moisture retention function to be $\theta(\psi)_{m_i}$ for soil i (i.e., a continuous van Genuchten was fitted to the discrete set of measured $\theta(\psi)$ values), and the predicted moisture retention function to be $\theta(\psi)_{p_i}$ for soil i (i.e., a continuous van Genuchten curve as predicted by the Vereecken PTF), where $i = 1, 2, \dots, N$, with N the total number of soils in the evaluation data set. Consequently, the MD and the RMSD ($\text{m}^3 \text{ m}^{-3}$) for soil i were calculated by:

$$\text{MD}_i = \frac{1}{b-a} \int_a^b [\theta(\psi)_{p_i} - \theta(\psi)_{m_i}]^2 d\log|\psi| \quad [5.3]$$

$$\text{RMSD}_i = \sqrt{\frac{1}{b-a} \int_a^b [\theta(\psi)_{p_i} - \theta(\psi)_{m_i}]^2 d\log|\psi|} \quad [5.4]$$

where a and b are values defining the range of the experimental SWRC (in our case $a = \log 1$ kPa and $b = \log 1500$ kPa corresponding to the lowest and highest $|\psi|$ values applied in the experiment). In computing Eq. [5.3] and [5.4], $\log|\psi|$ was preferred over $|\psi|$ to avoid assigning too much weight to more negative soil water matric potentials (Tietje and Tapkenhinrichs, 1993). The use of the two indices is necessary because we are not comparing single values of water content but several values within a specific soil water matric potential range. MD indicates whether the PTFs overestimate or underestimate the measured data, while RMSD measures the absolute deviation from the measured data. The absolute value of MD should be as small as possible. Nevertheless, MD allows the overestimation (the positive difference) and the underestimation (the negative difference) to cancel out. Therefore, MD was used in our case (with range from 1 to 4.18) only to indicate whether a PTF overestimates ($\text{MD} > 0$) or underestimates ($\text{MD} < 0$) the water content, while RMSD, which is always positive, can be viewed as the continuous analogue of the standard deviation over the whole SWRC, providing therefore an absolute error index.

The Pearson correlation coefficient r (dimensionless) for soil i was calculated by:

$$r_i = \frac{\int_a^b [\theta(\psi)_{m_i} - \bar{\theta}_{m_i}] [\theta(\psi)_{p_i} - \bar{\theta}_{p_i}] d \log |\psi|}{\sqrt{\int_a^b [\theta(\psi)_{m_i} - \bar{\theta}_{m_i}]^2 d \log |\psi| \int_a^b [\theta(\psi)_{p_i} - \bar{\theta}_{p_i}]^2 d \log |\psi|}} \quad [5.5]$$

where $\bar{\theta}_{m_i}$ is the mean moisture content of the measured SWRC for soil i , and $\bar{\theta}_{p_i}$ is the mean moisture content of the predicted SWRC for soil i . The index r is an expression of the linearity between the measurements and predictions. An r value approaching 1 indicates that measured and predicted data pairs are linearly located around the trend line with perfect agreement (or 1:1 line). The van Genuchten parameters in each of $\theta(\psi)_{p_i}$ were obtained in the same way as $\theta(\psi)_{m_i}$. Also here, $\log|\psi|$ was used rather than $|\psi|$ in calculating the moisture content.

Using the validation data set, the new developed PTF was compared with that reported by Ghorbani Dashtaki et al. (2010). The main reason for choosing this PTF is because it has been also developed for dryland soils in Iran. Comparison was made by calculating means of mean error (ME), root mean square error (RMSE), and the coefficient of determination (R^2), which were calculated for each soil sample.

5.3 Results and discussion

Figure 5.2 shows the textural distribution of the soil samples according to the two methods of determining PSD. Great variability can be noticed in the sand, silt, and clay fractions. Clay fraction is the most affected after destroying carbonates with hydrochloric acid. The texture triangles clearly depict a strong decrease in clay content. It is clear that carbonates content seems to be the soil characteristic that highly influences the PSD of calcareous soils.

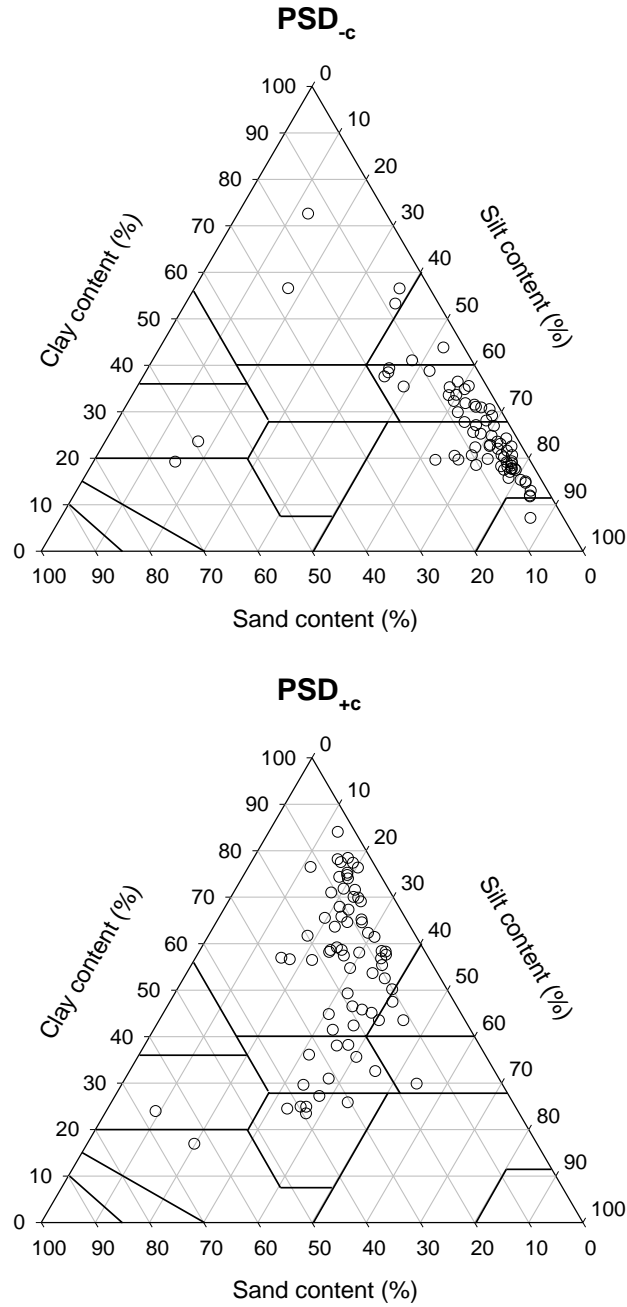


Figure. 5.2. Variation of clay, silt and sand content in the dataset for the two PSD methods. PSD_{-c} is texture with destroying carbonates and PSD_{+c} is texture without removing them (clay fraction is the most affected by the removal of carbonates).

5.3.1 Stepwise regression results

5.3.1.1 PSD methods and developing new PTF

We performed stepwise regression using two independent variable groups related to PSD_{-C} and PSD_{+C} methods. Both have the same values of bulk density and organic carbon, while they are different in the percent sand, silt and clay. When considering the first group (PSD_{-C}), no regression equation could be established. Its variables did not show a significant correlation with the dependent variables (soil-water content at matric potential points). In a multivariable problem, the regression equation is arrived at in a sequence of multiple linear regression equations in a stepwise manner. At each step of the sequence, one variable is added to the regression equation. However, none of the five PSD_{-C} predictors themselves was significant at the 95 percent level and consequently no one could be featured in the regression equation. Table 5.1 shows the Pearson correlation coefficients (*r*) between clay, silt, and sand contents obtained by the two methods and soil-water content at different matric potentials. The correlations for the PSD_{-C} and $\theta(\psi)$ were not significant at the 0.05 level and soil texture could in this case not be translated into a soil-water retention curve. Consequently, PTFs should not be developed from data sets that contain soils with large CaCO₃ content mixed with others and in which carbonates were removed prior to PSD analysis.

Table 5.1. Pearson correlation coefficients between soil-water content at different matric potentials and clay, silt and sand of the two PSD methods. PSD_{-C} is texture with destroying carbonate and PSD_{+C} is texture without removing it. §

ψ (kPa)	PSD _{-c}			PSD _{+c}		
	clay	silt	sand	clay	silt	sand
-1	0.14	-0.13	0.05	<u>0.48</u>	<u>-0.41</u>	<u>-0.40</u>
-3	0.06	-0.07	0.05	<u>0.50</u>	<u>-0.40</u>	<u>-0.44</u>
-5	0.03	-0.05	0.05	<u>0.50</u>	<u>-0.37</u>	<u>-0.46</u>
-7	0.03	-0.06	0.06	<u>0.55</u>	<u>-0.41</u>	<u>-0.50</u>
-10	0.01	-0.05	0.07	<u>0.56</u>	<u>-0.42</u>	<u>-0.52</u>
-33	-0.02	0.00	0.02	<u>0.64</u>	<u>-0.46</u>	<u>-0.60</u>
-100	-0.10	0.13	-0.10	<u>0.71</u>	<u>-0.49</u>	<u>-0.68</u>
-1500	-0.14	0.12	-0.03	<u>0.76</u>	<u>-0.51</u>	<u>-0.74</u>

§ All values that are underlined indicate a significant correlation at the 0.05 level.

We found a very different behaviour as regards the second group (PSD_{+C}) of predictors. Looking at the correlation across the whole range of measured $\theta(\psi)$, the soil texture here seemed to be translated into a soil-water retention curve as the correlation coefficients increased significantly. After running stepwise MLR, the predictors and their coefficients retained for each soil-water matric potential in the new PTF (denoted as PTF1) are tabulated in Table 5.2.

Results show that clay content is a basic predictor of $\theta(\psi)$ at all matric potentials indicating a significant influence of the clay fraction on the physical and chemical state of water in soil. This is primarily because the small particles have such a large and reactive surface area (Jury and Horton, 2004). From Figure 5.3 it can be seen that θ had a linearly increasing trend for the soil matric potential range from near saturation to wilting point with increasing clay content. Hall et al. (1977) explained that as matric potential decreases and pore size becomes finer, the proportion of water retained increases. The soil-water content at -1 kPa (near saturation) was not only predicted by clay but also by bulk density. Since bulk density (ρ_b) is directly related to soil porosity ϕ by the relation $\phi = 1 - \rho_b / \rho_s$ (where ρ_s is the density of the solid particles of the soil), it could be expected, as we observed, that the regression coefficient associated with bulk density will be negative (Scheinost et al., 1997; Weynants et al., 2009) as soils with high porosity tend to retain more water near saturation. The soil-water content at -1500 kPa (wilting point) was predicted with silt and clay. Table 5.1 shows that clay and silt contents were positively and negatively correlated with θ_{-1500} kPa, respectively, indicating that samples with more clay and less silt retained more water. One could speculate that the cause of this is related to soil water retention at low matric potential values being attributed to adsorption on soil particles, which is strongly dependent on soil surface area (Gardner, 1968).

Table 5.2. Pedotransfer function coefficients and their confidence interval, the *t*-statistic and its *p*-value.

ψ	Predictor	Coefficient	Value	t- statistic	p-value	Confidence Interval	
						Lower Bound	Upper Bound
-1	intercept	a_1	0.6193	10.0462	0.0000	0.4952	0.7434
	bulk density	b_1	-0.1398	-3.4320	0.0013	-0.2218	-0.0578
	clay	c_1	0.0014	3.9818	0.0002	0.0007	0.0022
-3	intercept	a_2	0.3750	15.1374	0.0000	0.3251	0.4248
	clay	b_2	0.0019	4.4247	0.0001	0.0011	0.0028
-5	intercept	a_3	0.3522	13.8577	0.0000	0.3010	0.4033
	clay	b_3	0.0020	4.5458	0.0000	0.0011	0.0029
-7	intercept	a_4	0.3220	12.6246	0.0000	0.2707	0.3733
	clay	b_4	0.0023	5.0898	0.0000	0.0014	0.0032
-10	intercept	a_5	0.3048	11.8715	0.0000	0.2532	0.3565
	clay	b_5	0.0024	5.2670	0.0000	0.0015	0.0033
-33	intercept	a_6	0.2579	10.1044	0.0000	0.2066	0.3093
	clay	b_6	0.0028	6.2589	0.0000	0.0019	0.0037
-100	intercept	a_7	0.2015	8.0385	0.0000	0.1511	0.2519
	clay	b_7	0.0032	7.1822	0.0000	0.0023	0.0041
-1500	intercept	a_8	-0.0127	-0.1987	0.8434	-0.1409	0.1156
	clay	b_8	0.0047	7.2514	0.0000	0.0034	0.0059
	silt	c_8	0.0026	2.2725	0.0278	0.0003	0.0048

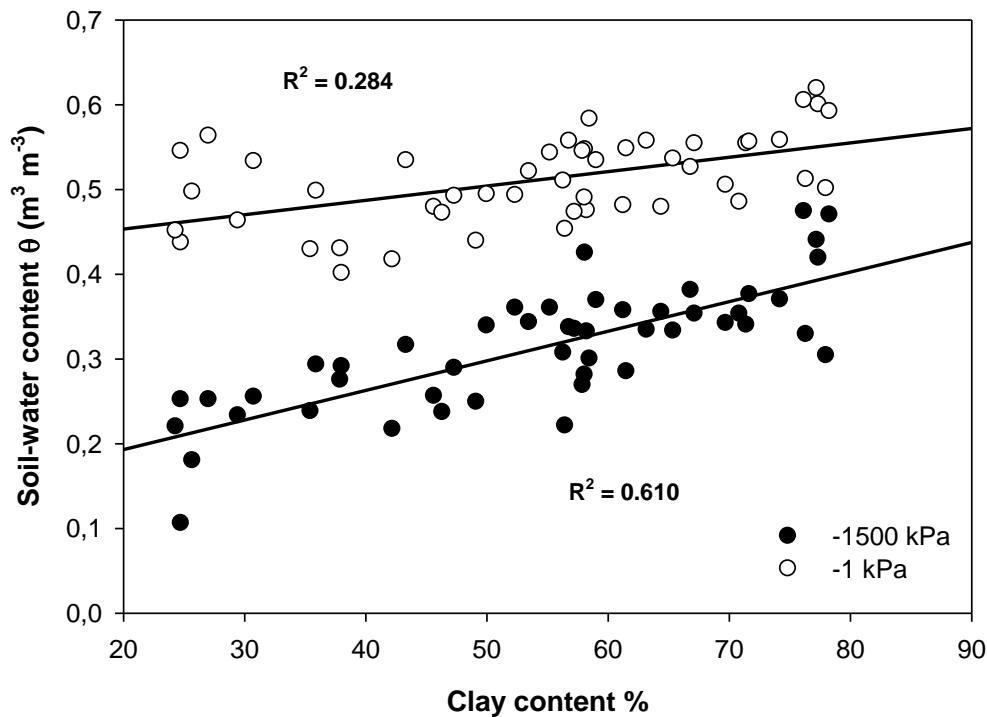


Figure 5.3. Effect of clay content on soil-water content at -1 and -1500 kPa. Solid line represents a linear relationship between clay content and soil-water content at -1 and -1500 kPa.

An interesting relationship can be made between the available water content (AWC) and clay content. The AWC is basically calculated as the difference between the soil-water contents at field capacity (-10 kPa for coarse-textured soils or -33 kPa for medium- to fine-textured soils) and at permanent wilting point (-1500 kPa) (Jury et al., 1991). Since our data set comprised mainly fine-textured soils we considered field capacity at -33 kPa. Figure 5.4 shows that the AWC decreased linearly with clay content. This finding reconfirms the previous observation of Jamison and Kroth (1958) who examined the interaction of AWC with texture and organic matter content. They concluded that increased AWC was due to the textural change associated with increased coarse silt and decreased clay. At low matric potential water adsorption is common and increases with clay content. The water, in this case, is strongly held in the fine pores by the London–van der Waals forces. In our case, the clay fractions are expected to contain chiefly carbonates with dolomite ($\text{CaMg}(\text{CO}_3)_2$), magnesite (MgCO_3) and calcite (CaCO_3) as non-clay minerals (Fares, 1991). This will lead to different behaviour in terms of water adsorption as compared to non-calcareous soils. Figure 5.5 indicates that the available water content is positively correlated with soil carbonates ($p < 0.05$). It is interesting to note that carbonates appear to increase the available water content, in contrast with findings of Khodaverdiloo and Homae (2004). A number of reasons explain this result. First, very fine carbonate particles can coat clay particles and reduce their surface tension with water (McCauley et al., 2005). In this case the particles are larger and behave as silt particles from a water retention point of view. Secondly, the soil pores are filled by carbonates which appears to reduce the amount of water retained at wilting point and thereby increasing the AWC (Duniway et al., 2007). This means that minute silicate clay particles are accompanied by a tremendous number of adsorbed cations and water molecules, while CaCO_3 particles of the same size are not. In non-calcareous soils, the high clay content increases the storage capacity of soils for water and minerals but decreases the aeration so essential for good root growth and functioning

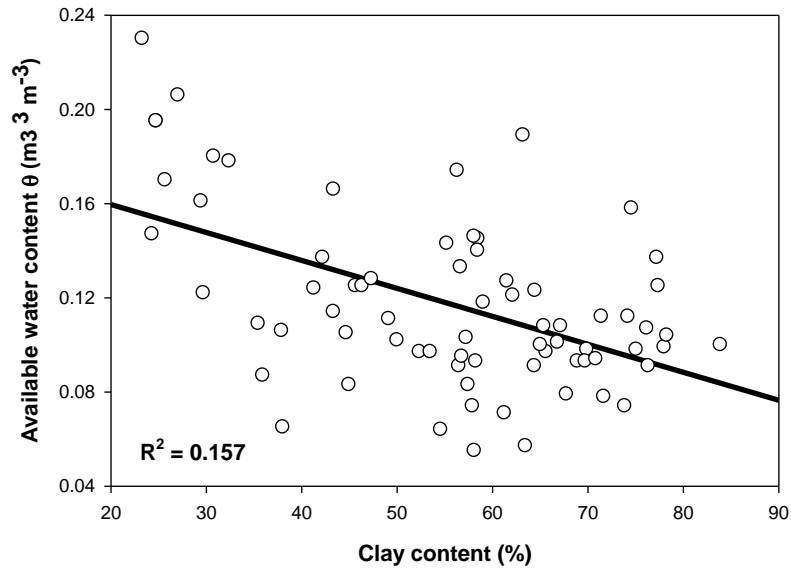


Figure 5.4. Effect of clay content on available water content (AWC). Solid line represents a linear relationship between clay content and AWC.

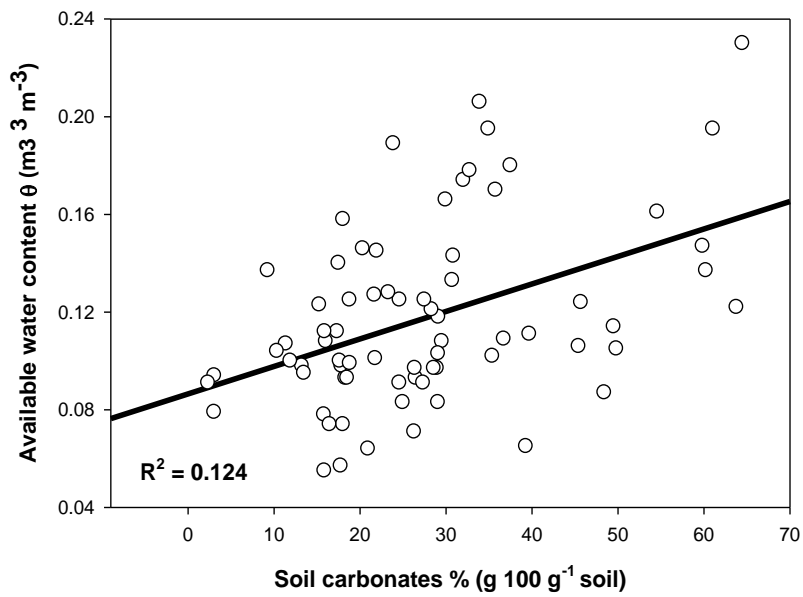


Figure 5.5. Effect of soil carbonates on available water content (AWC). Solid line represents a linear relationship between carbonates content and AWC.

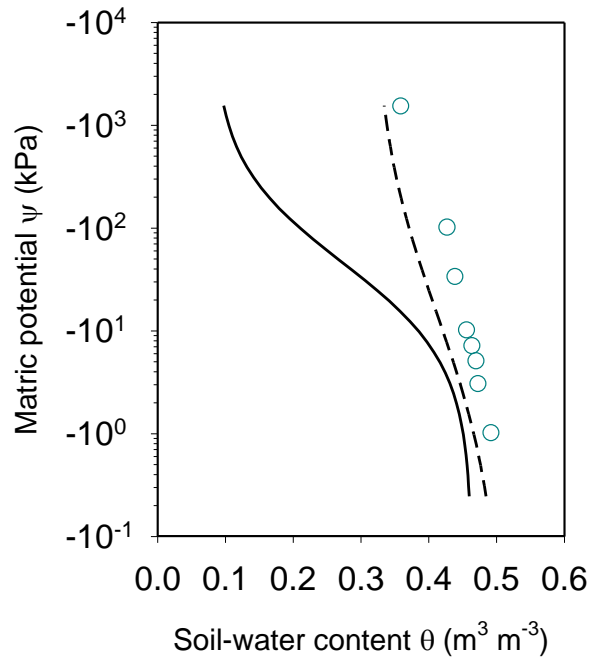
Organic carbon (OC) is an important variable used in PTFs (Wösten et al., 1999), but the contribution of OC content in estimating θ (Ψ) was very low and it was not significantly different from zero ($p > 0.05$). This was not surprising, since in naturally low organic matter soils common in dryland areas, OC is of limited influence. The low organic matter content is a general problem met in dryland environments. Organic matter contributes to soil fertility in three ways. It accounts for a large portion of the cation exchange capacity, it supplies energy and body-building constituents for most of the micro-organisms and it is responsible for the stability of soil aggregates (Brady, 1990), although calcium carbonate, as discussed above, will also contribute to structure stability.

5.3.1.2 PSD methods and the PTF of Vereecken et al. (1989)

The importance of not removing carbonates was further supported by applying the Vereecken et al. PTF to compare the prediction accuracy when using two different PSD inputs. Table 5.3 contains results of the different validation indices calculated for each PSD inputs. The PSD_{-C} method showed higher values of the mean of the absolute values of MD and the mean of RMSD (Eq. [5.3] and [5.4]) values. When considering the mean of MDs, it can be noticed that the PSD_{-C} tend to underestimate the SWRC (Table 5.3 and Figure 5.6). This underestimation occurs mainly at water content below -10 kPa.

Table 5.3. Comparison of the validation indices of the predicted SWRC by the PTF of Vereecken et al. (1989) using two different PSD inputs.

PSD methods	mean MD	mean abs. MD	mean RMSD	mean r	SD RMSD
	$\text{m}^3 \text{ m}^{-3}$			-	$\text{m}^3 \text{ m}^{-3}$
PSD _{-C}	-0.0778	0.0949	0.1073	0.9742	0.0494
PSD _{+C}	-0.0400	0.0620	0.0698	0.9820	0.0574



◦ Measured data — Predicted SWRC PSD_C - - Predicted SWRC PSD_{+C}

Figure 5.6. Measured and predicted soil-water retention curves for a calcareous soil. The curves correspond to the same soil with OC = 0.02%, $\rho_b = 1.21 \text{ Mg m}^{-3}$, $\text{CaCO}_3 = 28.60\%$, but with different PSD results. The solid line corresponds to the PSD_C with clay = 15.18%, silt = 81.02%, sand = 3.80%, and the dashed line corresponds to the PSD_{+C} with clay = 52.36%, silt = 37.37%, sand = 10.27.

The above findings are also supported by Figure 5.7 in which measured water content values are plotted against predicted water content values at -1500 kPa matric potential. The PSD_{+C} method particularly performs relatively well at the dry end of the SWRCs ($\psi = -1500 \text{ kPa}$). Finally, it can be concluded from here that the PSD_{+C} method (without removal of carbonate) is more adequate to classify soil texture. An additional advantage associated with this method is that it reduces time and work in determining PSD.

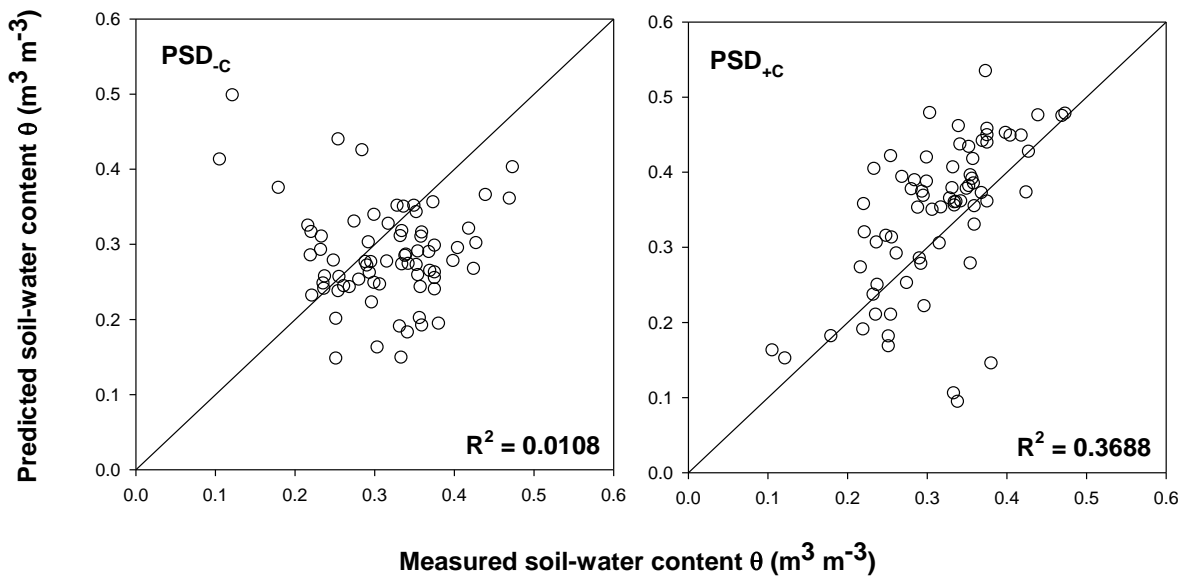


Figure 5.7. Measured vs. predicted soil-water contents at matric potential ψ of -1500 kPa for the two PSD methods.

5.3.1.3 Comparing the new PTF with the Ghorbani Dashtaki et al. (2010) PTF

The derived PTF was compared with the Ghorbani Dashtaki et al. (2010) PTF. The results of its performance for predicting θ using the validation data sets are summarized in Table 5.4. The Ghorbani Dashtaki et al. (2010) PTF shows a general underestimation of the soil-water content (Figure 5.8) with a considerable bias compared with the proposed PTF. As regards the RMSE, the proposed PTF has smaller values than those for the Ghorbani Dashtaki et al. (2010) PTF, which is mainly due to a low prediction error. As concerns the coefficient of determination (R^2), again the same trend can be perceived: the proposed PTF performs better. Table 5.4 shows clearly that the performance of the proposed PTF was better than the Ghorbani Dashtaki et al. (2010) PTF in terms of the ME, RMSE, and R^2 . Ghorbani Dashtaki et al. (2010) have shown earlier that their PTF performed better than Rosetta package (Schaap et al., 2001). However, the ranges of validity of the latter PTF do not cover the whole range of calcareous texture commonly encountered in soils of the dryland area. Although the comparison of PTFs inside its range of validity is not questionable, the lack of a specific PTF for dryland soils has resulted in applying PTFs derived from temperate regions. Lacking hydraulic information, Ouessar et al. (2009) e.g. used the PTF of Saxton et al. (2005)

developed for USA soils in modifying the SWAT model for evaluating the hydraulic impact of water harvesting techniques in Southern Tunisia. The limitation of the Ghorbani Dashtaki et al. (2010) PTF when applied to Syrian soils was evident, even within the range of validity (textures) of our dryland dataset.

The above findings are also supported by Figure 5.8 in which measured water content values are plotted against predicted water content values at matric potential of -10, -33, -100, and -1500 kPa. From Figure 5.8 it can be deduced again that the PSD_{+C} soil texture resulted in a decrease in prediction error for different matric potentials.

Table 5.4. Comparison of PTFs performance.

ψ	PTF presented in this study (Table 5.2)			Ghorbani Dashtaki et al. (2010) PTF		
	ME	RMSE	R^2	ME	RMSE	R^2
-10	0.0173	0.0683	0.9643	0.0270	0.0709	0.9615
-33	0.0160	0.0653	0.9692	-0.0760	0.1022	0.9246
-100	0.0188	0.0629	0.9729	-0.0992	0.1200	0.9015
-1500	0.0199	0.0539	0.9796	-0.0594	0.0810	0.9538

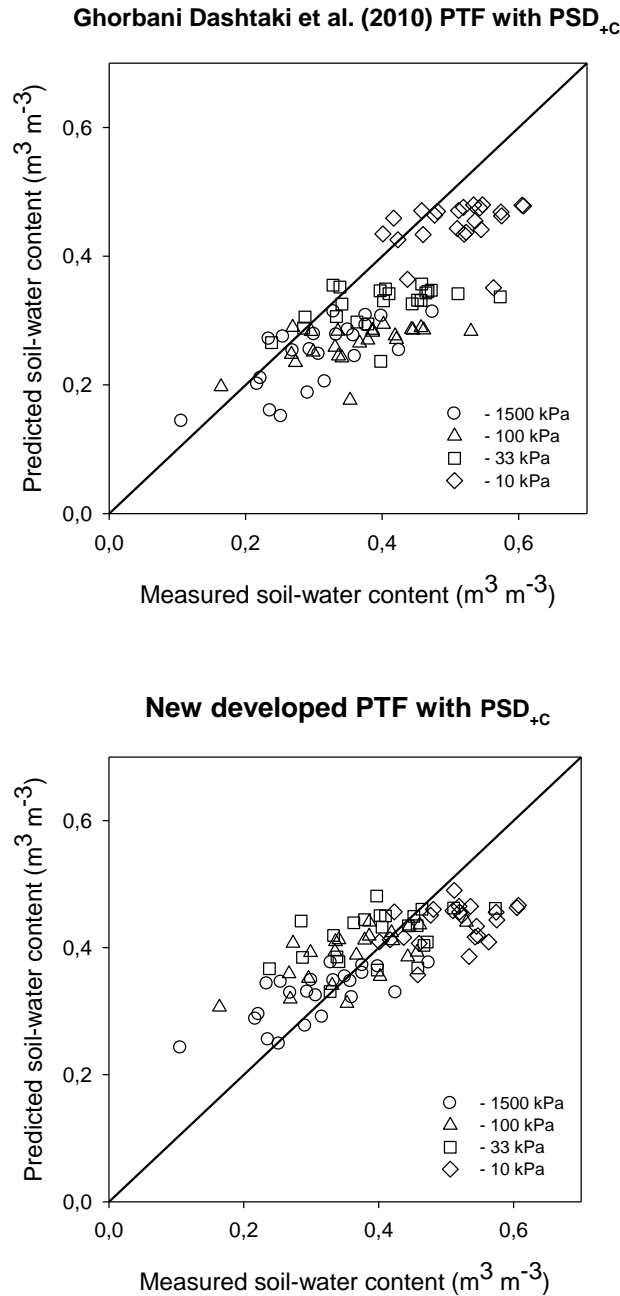


Figure 5.8. Measured vs. predicted soil-water contents at matric potential ψ of -10, -33, -100, and -1500 kPa.

5.4 Conclusions

Pedotransfer functions are very useful tools for obtaining the soil hydraulic properties. All published studies to date use soil texture as essential key predictor to most soil hydraulic

PTFs. The majority of used information on texture was obtained from a methodology that considers removing calcium carbonates as a pretreatment. Our results show, however, that in case of calcareous soils which are abundant in dryland environments, non-removal of soil carbonates considerably affected sand, silt, and clay fractions to the extent that PTFs could not be established when carbonates were not removed. Only texture that forgoes the pretreatment of destroying carbonates can be translated into the soil-water retention curve. The PTF established in this study allows the prediction of the soil-water contents at different matric potentials using soil texture (based on PSD_{+C}) and bulk density only. Further, the derived point PTFs provided better accuracy than those reported by Ghorbani Dashtaki et al. (2010) for dryland soils of Iran in terms of the ME, RMSE, and E.

It can be concluded from this study that the PSD_{+C} method (without removal of carbonate) is more adequate to classify soil texture and hence to predict the SWRC. Besides resulting in better predictions, the advantage of this method is a reduction in time and work when determining PSD.

Chapter 6 Exploration of interaction between hydraulic and physico-chemical properties of Syrian soils

Based on: Khlosi, M., W.M. Cornelis, A. Douaik, A. Hazzouri, H. Habib, D. Gabriels (2013). Exploration of the interaction between hydraulic and physicochemical properties of Syrian soils *Vadose Zone J.* 2013. 12:–. doi:10.2136/vzj2012.0209.

Abstract

In dryland areas, the availability of reliable data for water retention in relation to soil type, texture, and soil carbonate content is low. It is therefore desirable to explore the interaction between soil hydraulic properties and other physical and chemical properties in order to estimate the soil-water retention curve (SWRC) from easily measured soil parameters. In the present study, 72 soil samples were collected from rural areas throughout northwest Syria, covering most of its agro-climatic zones and soil types. Soil water content at different matric potentials and 11 chemical and physical soil properties were determined. Then, a Pearson correlation matrix was computed on which principal component analysis was applied to three soil-water contents, namely at -1 , -33 and -1500 kPa, and the 11 soil properties. Four principal components (PC) explained 77% of the variation in the data set. The three soil-water contents were highly linked to PC1 which is correlated to plastic limit, texture, soil carbonate, and specific surface area. In addition, soil-water content at -1 kPa was also linked to PC4 which is correlated to bulk density. Therefore, from the initial 11 soil properties, seven contribute to the three soil-water contents (plastic limit, texture, soil carbonate, specific surface area, and bulk density); the remaining four others (organic matter, gravel, CEC, and hygroscopic water content) have a negligible influence. Consequently, pedotransfer functions might be estimated using the original seven, from the initial 11, soil properties or their corresponding PCs in order to estimate the SWRC.

6.1 Introduction

The unsaturated soil hydraulic properties are key factors for land management in dryland areas. The hydraulic properties involve the soil water retention curve (SWRC), which relates the matric potential (ψ) with the soil water content (θ) and the hydraulic conductivity function. Traditional methods to determine soil hydraulic properties are difficult, tedious to accomplish and expensive. Hence, pedotransfer functions (PTFs) (Bouma, 1989) are being increasingly applied as a cost-effective way to estimate unsaturated soil hydraulic properties from easily measurable or already available soil data. Subsequently, many PTFs appeared in the literature for estimating water retention properties.

To date, numerous attempts have been made to relate soil physical and chemical properties to the SWRC. The most common predictors of SWRC are soil texture, bulk density and organic carbon. However, dryland soils have low organic carbon content. Other easily measured soil properties, such as plastic limit, require less time and are less demanding than organic carbon measurement. Furthermore, PTFs developed at one spatial extent are generally not suited for other spatial extent (Nemes et al., 2003). Pertinent to this, Bastet et al. (1997) found that performance of PTFs varies with the pedological origin of the soils on which they were developed. Consequently, the validity of any given PTF should be considered appropriately with caution before extrapolation beyond their geographical training area. Meanwhile, in Syria there is no published work available on using pedotransfer functions for predicting the SWRC.

From the above, it is clear that there is a need to explore the interaction between soil hydraulic properties and other physical and chemical properties. The approach in this study is to investigate the possible use of new basic soil properties as predictors of soil hydraulic properties, which can be easily or cheaply measured. The main objective of this study is thus to explore the interaction between SWRC and other physical and chemical properties and to gain insight into the experience and views of SWRC-basic soil properties interaction in dryland soils. Another objective is to improve our knowledge on the hydraulic properties of typical soils from Syria. We utilized a multivariate approach, principal component analysis (PCA), to provide a comprehensive evaluation of all data and a holistic comparison between

soil properties. PCA was used to examine multivariate relationships between soil-water contents and other physical and chemical properties.

6.2 Materials and methods

6.2.1 Area description and soil sampling

The same data sets considered in the previous chapter are used here (see chapter 5, section 5.2.1).

6.2.2 Soil analysis

The undisturbed soil samples were used to determine the SWRC and bulk density following the procedure described in Cornelis et al. (2005). In brief, the undisturbed soil cores were saturated from the base upward with distilled water. The samples' SWRC was determined at eight matric potentials. This was done with a sand box apparatus (Eijkelkamp Agrisearch Equipment, Giesbeek, the Netherlands) for matric potentials between -1 and -10 kPa, and with pressure chambers (Soilmoisture Equipment, Santa Barbara, CA) for matric potentials between -20 kPa and -1500 kPa. Bulk density was determined before the samples were brought to the pressure plates. The samples were oven dried (105°C) until constant weight (>24 h). To reduce possible unresolved variance, the basic soil properties were determined directly on the same samples after the SWRC was established. All properties that have been analysed in this study are listed in 6.1.

The cation-exchange capacity (CEC) was determined by the method of Polemio and Rhoades (1977) which is particularly suited to dryland soils, including those containing carbonates (Fares et al., 2005). It is a simple and much less laborious method compared to other methods (Misopolinos and Kalovoulos, 1984). Organic matter was determined by means of the Walkley and Black (1934) method. The specific surface area of oven-dry (105°C), 2-mm sieved soil was estimated using a Ströhlein Areameter II apparatus with N_2 adsorbate in conjunction with the Brunauer, Emmett, and Teller (BET) equation. The particle fractions were determined using the sieve-pipette method without removal of carbonate (the PSD_{+C} method) which is more adequate to classify soil texture and hence to predict the SWRC. The plastic limit was determined as the gravimetric water content at which a soil sample could be

rolled by hand into a thread of 3.2 mm diameter without breaking (ASTM, 1998). We further measured soil carbonates by back-titration (with 0.5 M NaOH) of an excess 0.25 M H₂SO₄ added to 1 g of soil (Nelson, 1982), the gravel percentage (> 2 mm) by sieving, and the hygroscopic water content by weight differences after drying the samples at 105°C during 24 hours. Dirksen and Dasberg (1993) defined the hygroscopic water content as the one corresponding to a matric potential of -10⁵ kPa. When soil is exposed to atmospheric air, it will dry or wet, until a thermalisation is established, according to the potentials on both sides of the liquid-vapor interface.

Table 6.1. List of soil properties used in predictive procedures.

Symbol	Variable information
$\theta(\psi)$	Soil water retention curve (SWRC).
θ	Soil-water content (m ³ m ⁻³) determined at eight matric potentials.
ψ	Matric potential (kPa).
ρ_b	Soil bulk density in weight per volume (Mg m ⁻³).
O.M.	Organic matter content (g kg ⁻¹) determined by means of the Walkley and Black (1934) method.
Clay	Clay content (0–2 μm), after dispersion with sodium hexametaphosphate (% [w/w] of the soil fraction <2 mm; Day, 1965).
Silt	Silt content (2–50 μm), (% [w/w] of the soil fraction <2 mm; Day, 1965).
Sand	Sand content (50–2000 μm), (% [w/w] of the soil fraction <2 mm; Day, 1965).
CEC	Cation-exchange capacity (cmol kg ⁻¹), measured by Polemio and Rhoades (1977) method.
SC	Soil carbonates content (%) determined by back-titration approach (Nelson, 1982).
GRAVEL	Fraction of the bulk soil >2 mm (% [w/w] of the bulk soil).
SSA	Specific surface area (m ² g ⁻¹), measured by nitrogen adsorption with BET determination.
PL	Plastic limit criterion (m ³ m ⁻³), a soil-water content equal to that of the plastic limit.
θ_{hy}	The hygroscopic water content was determined by weight differences after drying the samples at 105°C during 24 hours.

6.2.3 Statistical analysis

First of all, a matrix of Pearson correlation coefficients between any two variables was computed and the coefficients were tested for their statistical significance. Then, we used principal component analysis (PCA) (Jolliffe, 2002); a multivariate analysis technique which

provides an excellent means for gaining useful information from data sets with many variables (Vereecken and Herbst, 2004). In particular, PCA can aid in the compression and classification of data. The purpose is to reduce the dimensionality of a data set by finding a new set of variables, smaller than the original set of variables, which nonetheless retains most of the sample's variance. Success relies upon the presence of correlations among at least some of the original variables; otherwise the number of new variables will be almost the same as the number of original variables. The new variables, called principal components, are uncorrelated, and are ordered by the fraction of the total variance each retains. The PCA technique generates interesting hypotheses for predicting the SWRC. The PCA was carried out in this study with SPSS 15.0 (SPSS, 2006).

6.3 Results and discussion

6.3.1 Exploratory statistical analysis

Descriptive statistics as minimum, maximum, mean, median, and standard deviation of the soil properties considered in this study are presented in Table 6.2. In addition, to get an idea about the distribution of the soil hydraulic properties and to assess the presence of eventually any outliers, a box plot is presented (Figure 6.1). The statistical parameters will be discussed in the next section. Regarding the boxplots, it is clear that most of the medians are in the middle of the different boxes corresponding to the three soil hydraulic properties from the four zones with the exception of $\theta_{-33\text{kPa}}$ from the second and third zones. This indicates that most of the soil properties follow approximately a normal distribution. Moreover, there are clearly some outliers, which are unexpectedly very high or very low compared to the statistical distributions. The only case for which the value was abnormally high was $\theta_{-1\text{kPa}}$ for sample 11 from zone 1. In all the other cases, values were unexpectedly low: it is the case of sample 18 from zone 2 for the three soil hydraulic properties, sample 19 again from zone 2 for $\theta_{-33\text{kPa}}$ and $\theta_{-1500\text{kPa}}$, and samples 60 and 72 from zone 4 for $\theta_{-33\text{kPa}}$.

Table 6.2. Descriptive statistics of soil properties used in predictive procedures.

Variable	N	Min.	Max.	Mean	Median	SD
Clay (%)	72	23.30	83.90	55.40	58.00	16.06
Silt (%)	72	9.30	54.60	28.60	28.36	9.32
Sand (%)	72	3.20	42.30	15.99	13.07	9.98
OM (%)	72	0.00	3.50	1.24	1.02	0.78
ρ_b(Mg m⁻³)	72	0.98	1.80	1.33	1.33	0.16
PL (kg kg⁻¹)	72	0.23	0.55	0.39	0.38	0.08
SC (%)	72	2.30	64.50	27.63	25.64	14.50
Gravel (%)	72	0.10	45.60	7.14	3.37	9.15
CEC (cmol₍₊₎ kg⁻¹ soil)	72	2.00	58.90	16.97	16.01	10.11
SSA (m² g⁻¹)	72	7.12	87.11	39.29	39.10	16.80
θ_{hy} (m³ m⁻³)	72	0.03	0.09	0.07	58.00	0.01

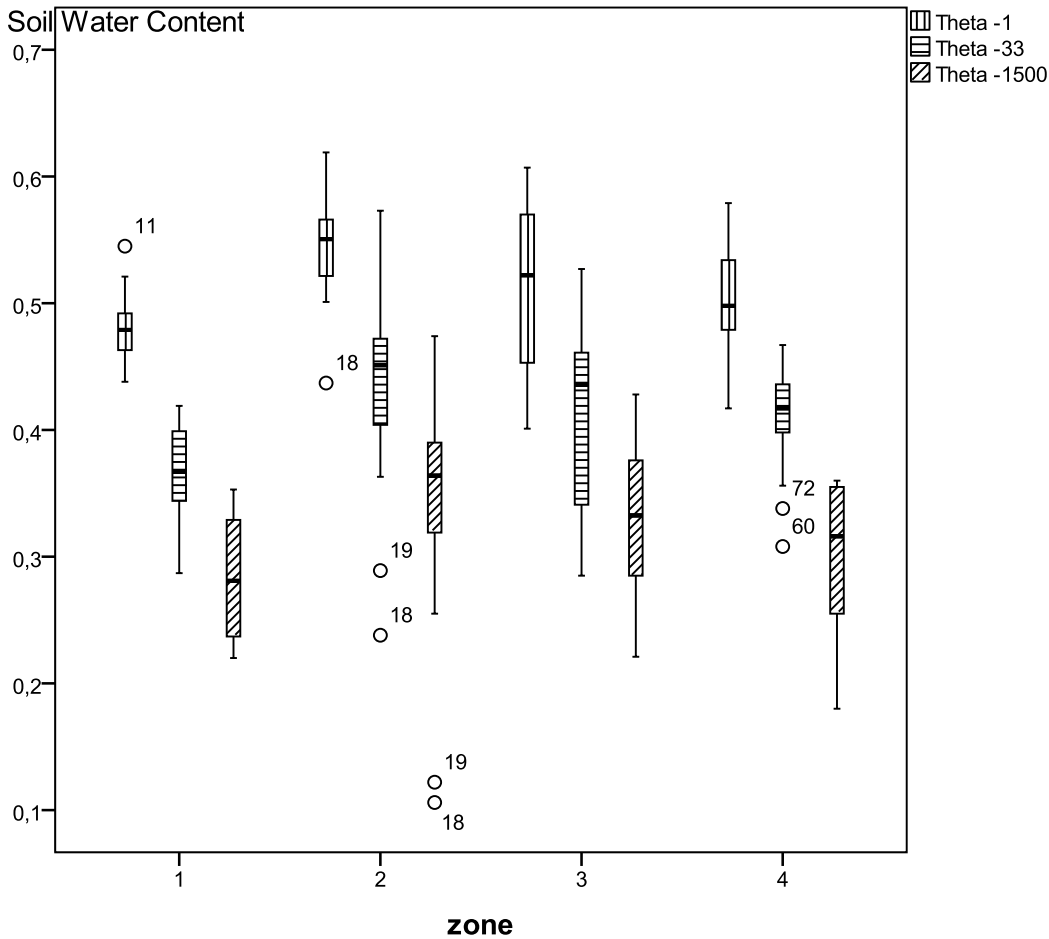


Figure 6.1. Boxplot of soil hydraulic properties from the four zones.

6.3.2 Relationships between soil-water content and other soil properties

6.3.2.1 Soil texture

Clay is dominant in our soils and its content had a mean value of 55.4% and ranged between 23.3 and 83.9% while silt content varied from 9.3 to 54.6% and had a mean value of 28.6%. Sand content was the least with a mean value of 16% and ranged between 3.2 and 42.3%. Correlation matrix (Table 6.3) is a useful tool in bringing important predictor variables to light. When a significant correlation was searched, at the 0.05 level, many strong and positive correlations were revealed between clay content and soil-water content at different matric

potentials. The slopes of these correlations increase with decreasing matric potential. As matric potential decreases, the quantity of water attached to the negatively charged clay particles increases relative to that retained in the soil pores by capillary forces. McBride and Mackintosh (1984) described the soil-water content at -1500 kPa as function of clay content. Such an interesting correlation has been previously recognized (e.g., Nielsen and Shaw, 1958). Van den Berg et al. (1997) and Botula et al. (2012) also provided a strong relationship between soil-water content at -1500 kPa and clay content for soils of the humid tropics. Minasny et al. (1999) found that soil-water content at -10 , -33 and -1500 kPa had an exponentially increasing trend with clay content and linearly decreasing trend with silt and sand content.

Our results showed also negative correlations between $\theta(\psi)$ and silt and sand contents. This confirms why most PTFs available in the literature use soil texture as the main predictor. Hillel (1998) defines soil texture as the permanent, natural attribute of the soil and the one most often used to characterize its physical makeup. In this context, Arya and Paris (1981) translated particle-size distribution data into a soil-water retention curve. Their PTF is based upon and capitalizes on the notable similarity between the nonlinear shapes of particle size distribution data and soil-water retention data. However, their model converts particle-size distribution to pore-size distribution considering only the capillary effect of the SWRC. The good correlation with clay is mainly caused by adsorption effects.

6.3.2.2 Organic matter

Our results show that organic matter had a mean value of 1.2%, ranged between 0 and 3.5%, and was not a major factor related to variations in SWRC across the entire range of soils we studied, even though correlation coefficients were significant between organic matter and cation exchange capacity, specific surface area and the hygroscopic water content (Table 6.3).

Soil organic matter is a source of food for soil fauna, and contributes to soil biodiversity by acting as a reservoir of soil nutrients such as nitrogen, phosphorus and sulphur; it is an important contributor to soil fertility. Organic matter absorbs water and causes soil particles to aggregate developing an open fabric that improves the physical environment for roots to penetrate through the soil. Soils containing organic matter have a better structure that improves soil hydraulic properties through bulk porosity, soil-water retention and hydraulic conductivity (Tisdale et al., 1993; Bossuyt et al., 2002). It supports water infiltration, and

reduces the soil's susceptibility to compaction, erosion, desertification and landslides (Schiettecatte et al., 2008). On the other hand, organic matter contributes to soil fertility in three ways. It accounts for a large portion of the cation exchange capacity, it supplies energy and body-building constituents for most of the micro-organisms and it is responsible for the stability of soil aggregates (Brady, 1990). High organic matter in the mineral soil efficiently retains water. The WRC of peat (e.g. Weiss et al. 1998) resembles the curve of clay soil; in the wet range of the curve, the water content becomes only slowly smaller with decreasing matric potential. The amount of organic matter is therefore an important factor of the shape of the WRC in peat soils. However, in naturally low organic matter soils common in dryland areas, organic matter has less influence.

6.3.2.3 Bulk density

Bulk density is one of the most important soil characteristics which describes the relative proportions of solid and void in a soil. Basically, it is required for predicting the soil-water retention curve, and is a necessary input parameter for water, sediment, and nutrient transport models. Recently, bulk density is increasingly used as a valuable indicator of soil physical quality (Wilson et al., 2013). Bulk density for soils from our study ranged from 1 to 1.8 Mg m⁻³ with a mean value of 1.3 Mg m⁻³. Concerning the SWRC, negative correlation coefficients were found between bulk density and soil-water content near saturation (θ_{-1} and θ_{-3}). Lower bulk densities increase the pore space and therefore, potentially, increase the conductive path for water. The bulk density was negatively correlated with plastic limit and positively correlated with soil carbonate indicating that bulk density increased while the plastic limit decreased with an increase in soil carbonate concentration (Table 6.3).

Table 6.3. Correlation matrix of predictor and response variables. §

Variable	$\theta_{.1}$	$\theta_{.3}$	$\theta_{.5}$	$\theta_{.7}$	$\theta_{.10}$	$\theta_{.33}$	$\theta_{.100}$	$\theta_{.1500}$	Clay	Silt	Sand	O.M.	ρ_b	PL	SC	Gravel	CEC	SSA	θ_{hy}	
$\theta_{.1}$	1.00																			
$\theta_{.3}$	0.92	1.00																		
$\theta_{.5}$	0.84	0.98	1.00																	
$\theta_{.7}$	0.79	0.95	0.99	1.00																
$\theta_{.10}$	0.78	0.94	0.98	0.99	1.00															
$\theta_{.33}$	0.76	0.91	0.95	0.98	0.98	1.00														
$\theta_{.100}$	0.68	0.84	0.87	0.90	0.90	0.95	1.00													
$\theta_{.1500}$	0.59	0.75	0.80	0.85	0.86	0.91	0.94	1.00												
Clay	0.48	0.50	0.50	0.55	0.56	0.64	0.71	0.76	1.00											
Silt	-0.41	-0.40	-0.37	-0.41	-0.42	-0.46	-0.49	-0.51	-0.82	1.00										
Sand	-0.40	-0.44	-0.46	-0.50	-0.52	-0.60	-0.68	-0.74	-0.84	0.38	1.00									
OM	-0.07	-0.07	-0.10	-0.12	-0.12	-0.08	-0.02	-0.02	-0.09	0.09	0.07	1.00								
ρ_b	-0.44	-0.20	-0.07	-0.02	-0.02	-0.05	-0.10	-0.01	-0.27	0.27	0.17	-0.06	1.00							
PL	0.68	0.65	0.62	0.63	0.64	0.71	0.75	0.73	0.80	-0.71	-0.63	-0.10	-0.49	1.00						
SC	-0.50	-0.47	-0.46	-0.47	-0.48	-0.54	-0.63	-0.64	-0.82	0.67	0.70	-0.09	0.39	-0.72	1.00					
Gravel	-0.40	-0.44	-0.46	-0.43	-0.42	-0.42	-0.44	-0.39	-0.20	-0.10	0.42	-0.05	0.07	-0.22	0.18	1.00				
CEC	0.35	0.38	0.36	0.40	0.40	0.42	0.38	0.38	0.45	-0.47	-0.30	-0.21	-0.19	0.39	-0.36	0.19	1.00			
SSA	0.35	0.41	0.45	0.46	0.48	0.50	0.56	0.56	0.50	-0.26	-0.57	-0.29	-0.16	0.51	-0.47	-0.33	0.17	1.00		
θ_{hy}	0.03	0.05	0.05	0.07	0.08	0.13	0.24	0.25	0.24	-0.15	-0.25	0.36	-0.19	0.11	-0.35	-0.09	0.05	0.14	1.00	

§ Italic numbers are correlation coefficients without significant correlation at the 0.05 level.

6.3.2.4 Plastic limit

The plastic limit, which is one of the so-called Atterberg limits, describes the soil-water content at which a soil begins to crumble when rolled into small threads, and divides the plastic and the semisolid states of a soil. In our study, the plastic limit had a mean value of $0.4 \text{ m}^3 \text{ m}^{-3}$ and varied from 0.2 to $0.6 \text{ m}^3 \text{ m}^{-3}$. It was strongly correlated with soil-water contents for all matric potentials indicating a promising predictor that can be generated to be used when developing a new PTF (Figure 6.2). Moreover, many positive and negative correlations were observed between plastic limit and other soil properties. Table 6.3 shows strong negative correlations between plastic limit and silt, sand, bulk density, and soil carbonate ($r = -0.49$ to -0.72). There were also some strong positive correlations between plastic limit and clay, cation exchange capacity, and specific surface area ($r = 0.39$ to 0.80). Hence plastic limit seems to be a powerful predictor of SWRC because it integrates many other soil properties. Previously, an attempt to use plastic limit in the water retention PTFs has shown some improvement in the PTF accuracy (Rawls and Pachepsky, 2002). Only organic matter was not correlated with plastic limit because our soils are poor in organic matter content. Blanco-Canqui et al. (2006) found that the plastic limit increased linearly with increasing soil organic carbon.

6.3.2.5 Soil carbonate and gravel content

Calcareous materials such as calcium carbonate and limestone are very common in soils of dryland areas. Soil carbonate content for soils from our study sites ranged from 2.3 to 64.5% with a mean value of 27.6%, whereas gravel ranged from 0.1 to 45.6% with a mean value of 7.1% (Table 6.2). These two variables showed similar behaviour from a hydraulic point of view. Negative correlation coefficients were found for soil carbonate and gravel with soil-water content at all matric potentials. Generally, the presence of gravel will increase the macroporosity and therefore decrease the soil water retention (van Wesemael et al., 1996; Schneider et al. 2010). According to Baetens et al. (2009), the soil water retention usually decreases when the rock fragments content increases. On the other hand, the negative correlation of soil carbonate with $\theta(\psi)$ across all matric potentials suggests that very fine carbonate particles, in the clay size range $< 2 \mu\text{m}$ diameter, can coat clay particles and reduce their surface tension with water (McCauley et al., 2005). In this context, Massoud (1972)

found when a large percentage of carbonate is present in the clay fraction (30% or higher), the soil's water holding capacity can be reduced.

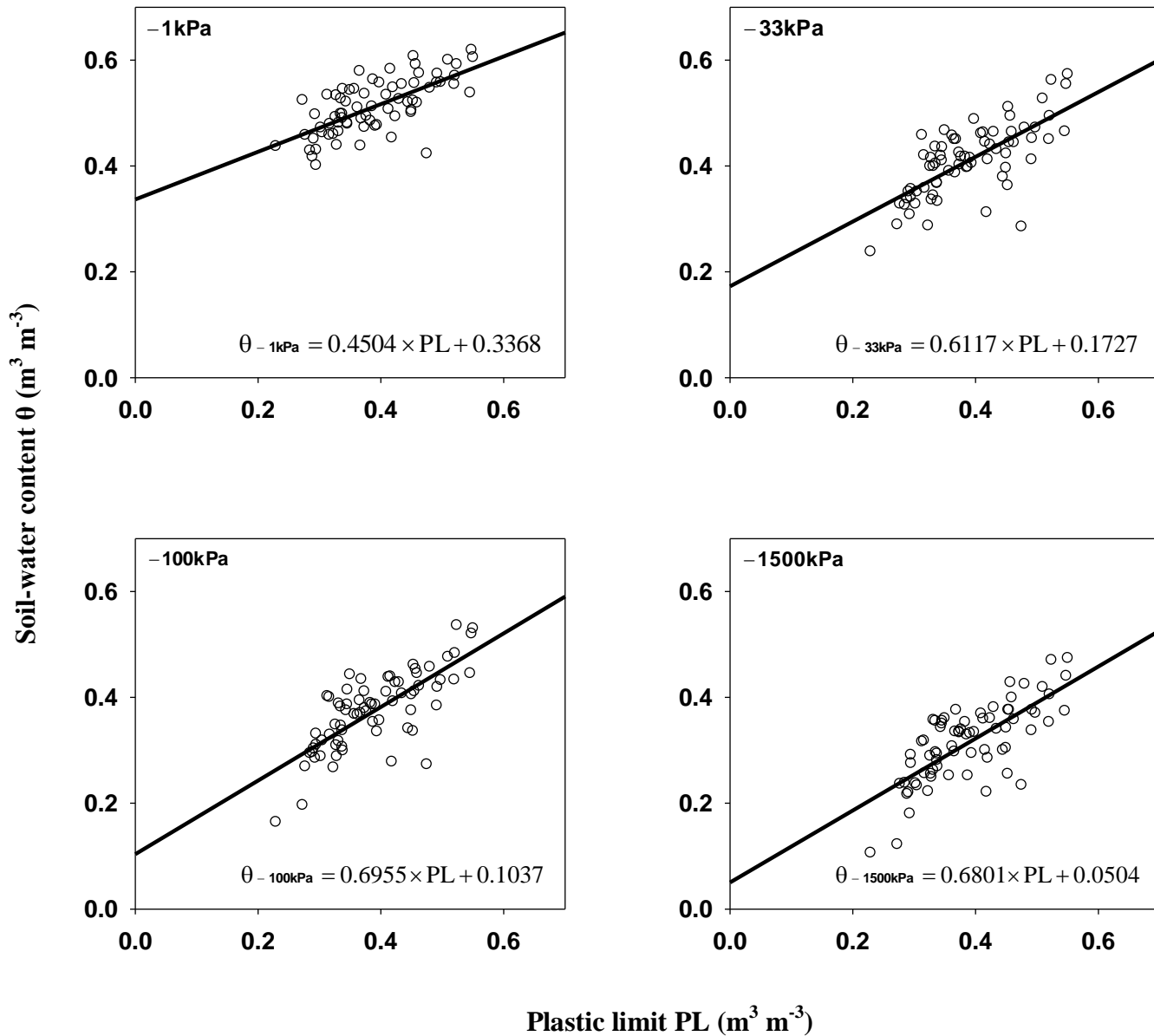


Figure 6.2. Effect of plastic limit on soil-water content at -1, -33, -100, and -1500 kPa. Solid line represents a linear relationship between plastic limit and soil-water content at -1, -33, -100, and -1500 kPa (Significant at $P < 0.001$).

6.3.2.6 Cation exchange capacity, specific surface area and hygroscopic water content

As cation exchange capacity, specific surface area and hygroscopic water content are dependent on texture as well as organic matter content of the soils, they might have a similar influence on the SWRC. The CEC ranged from 2 to 59 cmol kg^{-1} with a mean value of 17 cmol kg^{-1} while SSA ranged from 7.1 to 87.1 $\text{m}^2 \text{g}^{-1}$ with a mean value of 39.3 $\text{m}^2 \text{g}^{-1}$ (Table 6.2). Across all matric potentials, the CEC and SSA were positively correlated with the soil-water content indicating that soils with high cation exchange capacities and specific surface areas have high water holding capacities and greater swell potentials. Similar results were found by Campbell and Shiozawa (1992) who correlated SSA of six soils with measurements of the slope of a SWRC and found good correlation. Regarding the cation exchange capacity, similar results were reported by Gupta et al. (1983). They showed that CEC was positively correlated with soil-water retention at -33 and -1500 kPa in soils from northwest India.

The CEC and SSA were positively correlated to clay content and negatively correlated to silt and sand (Table 6.3). The smaller the particles, the greater the surface area per unit mass of soil and the higher the cation exchange capacity. In naturally low organic matter soils common in dryland areas, the CEC can be used as indicator of the clay mineralogy. Gaiser et al. (2000) investigated the influence of clay mineral composition on the SWRC and its contribution to the development of PTFs for soils from dryland regions. They demonstrated that PTFs for soils containing predominantly low activity clay ($\text{CEC} < 24 \text{ cmol/kg clay}$) differed considerably from those developed from non low activity clay ($\text{CEC} > 24 \text{ cmol/kg clay}$) soils.

Concerning the hygroscopic water content, it ranged in our study from 0.03 to 0.09 $\text{m}^3 \text{m}^{-3}$ with a mean value of 0.07 $\text{m}^3 \text{m}^{-3}$ (Table 6.2). Dirksen and Dasberg (1993) showed that θ_{hy} varied between 0.02 $\text{m}^3 \text{m}^{-3}$ for sandy soils with low SSA and 0.12 $\text{m}^3 \text{m}^{-3}$ for a Vertisol with high SSA. Therefore, we found that the θ_{hy} was only correlated to the soil-water content in the dry part of the SWRC (θ_{-100} and θ_{-1500}). At low matric potential water adsorption is common and increases with clay and organic matter contents. The water, in this case, is strongly held in the fine pores by the forces of cohesion and adhesion (Lal and Shukla, 2004). The latter explains the positive correlations (Table 6.3) between θ_{hy} , clay and organic matter.

6.3.3 Principal Component Analysis

6.3.3.1 PC loadings and communalities

As seen above, several correlations were found among the soil properties indicating that PCA can be carried out towards reducing the dimensionality of the original data set. Such correlations are useful because they numerically represented the similarity between the measures of two soil properties. We ran principal component analysis on 14 variables including soil-water contents at -1 , -33 , and -1500 kPa. The PCA was applied to the correlation matrix. Components with eigenvalue higher than one were retained (Kaiser and Rice, 1974). The Cattell (1978) scree plot (Figure 6.3) can be utilized as a visual tool to determine how many components are necessary to explain most of the original variation. Four principal components were selected; they explained 77% of the variation in the data set (Table 6.4). Loadings are the correlations between the PCs and the original variables: higher values mean closer relationship. The first component explained 46.2% of the data variability (an eigenvalue of 6.47 from a total of 14), and loadings were highest for $\theta_{-1\text{kPa}}$, $\theta_{-33\text{kPa}}$, $\theta_{-1500\text{kPa}}$, clay, silt, sand, PL, SSA, and SC. This means that these soil properties have a common ground of variance or they are inter-correlated. The second component, describing 11.4% of the total variance (an eigenvalue of 1.6 from a total of 14), was mainly correlated with gravel and CEC, whereas OM and θ_{hy} were highlighted in the third component (10.9%, an eigenvalue of 1.53 from a total of 14). The fourth component explained bulk density with data variability (8.1%, an eigenvalue of 1.13 from a total of 14). Communalities represent the proportion of the variance of each soil property that is accounted for by the principal component solution (the 4 PCs). They represent the total influence of the retained PCs on the original variables and are computed as the sum of the squares of the loadings. They can be interpreted as the coefficients of determination (the squares of the coefficients of correlation): the percentage of total variance in each soil property explained by the 4 PCs. The PCA should explain at least half of each characteristic's variance. The 4 PCs explained most of the variability of clay content, $\theta_{-1500\text{kPa}}$, bulk density, and plastic limit (90.5, 89.2, 88.2, and 85.7%, respectively) whereas they explained only around half of the variability of cation exchange capacity and specific surface area (57.5 and 52.4%, respectively). For the other soil properties, they explained between 66.9% (hygroscopic water content) and 82.2% (gravel) of their total variance. This means that the remaining percentage (to add to 100%), for each soil

property, is explained by the other 10 PCs not retained in the PCA following the eigenvalue higher than one criterion.

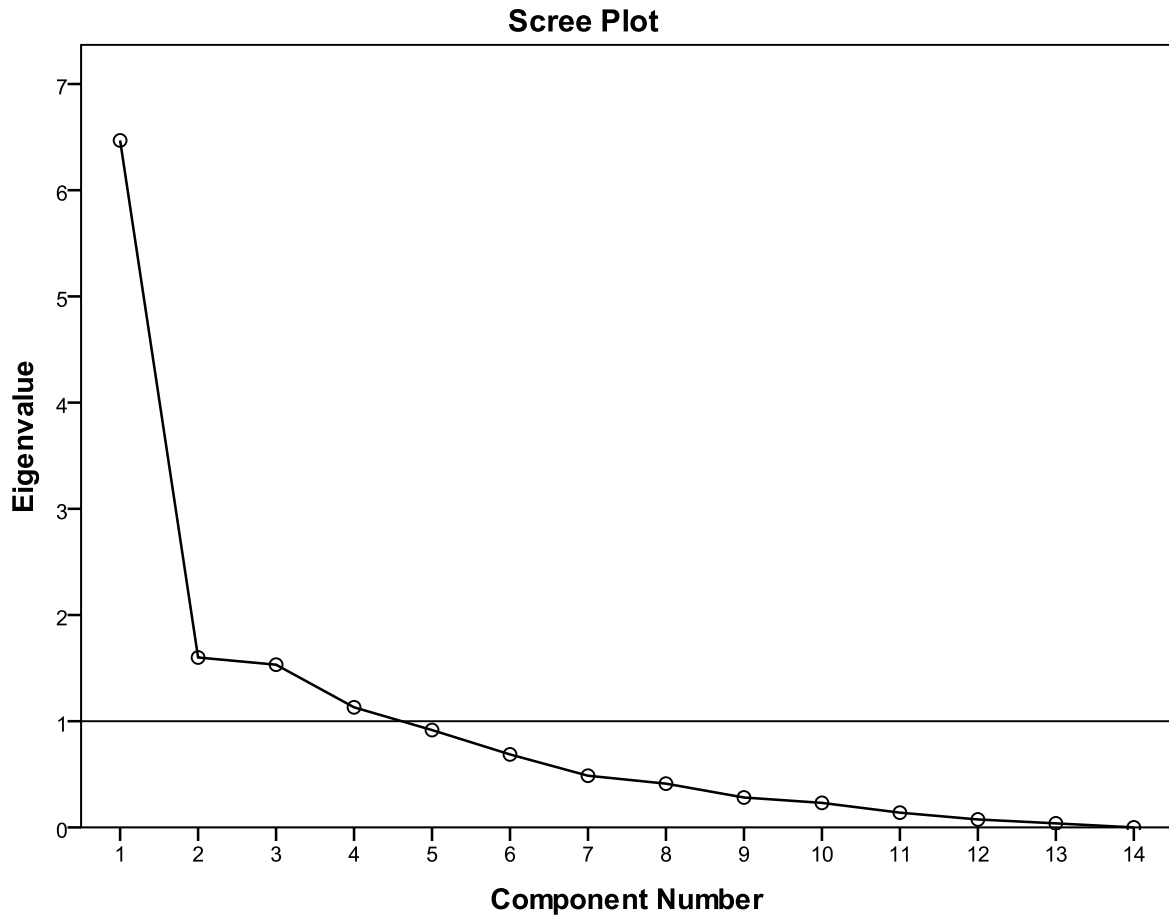


Figure 6.3. The scree plot for the principal component analysis. Four principal components were extracted, because their eigenvalues > 1 .

Communalities consider the 4 PCs together; however, one may be interested in the contribution of each PC to each original soil property separately. This contribution is the percentage of variance in a given soil property explained by a given PC and is computed as the squares of the PC loadings (Table 6.5); the sum of the contributions of the PCs is equal to the communality for each soil property (Table 6.4). PC1 explains most of the variation in clay, plastic limit, $\theta_{-1500\text{kPa}}$, soil carbonate, and $\theta_{-33\text{kPa}}$ (85, 79, 77, 71, and 69%) whereas it accounts for a very limited variation in organic matter, hygroscopic water content, gravel, and BD (1, 7, 12, and 13%, respectively). PC2 explains mainly gravel (62%) while PC3 explains OM and hygroscopic water content (64 and 52%, respectively) and PC4 explains BD (53%).

Regarding the soil water content, it is interesting to note that 52, 69, and 77% of variation at the -1kPa, -33kPa, and -1500kPa matric potentials is explained by PC1 while PC4 explains almost one quarter (24%) of variation in soil water content at the -1kPa matric potential. PC2 and PC3 explain a negligible amount of variation of soil water content (less than 6%).

Table 6.4. Loadings of 14 variables on significant principal components. The highest values (>0.5 threshold) are in bolds.

Variable ‡	Factor loadings and communality in PCA†				
	PC 1	PC 2	PC 3	PC 4	Com.¶
θ_{-1kPa}	0.72	0.07	-0.10	-0.49	0.779
θ_{-33kPa}	0.83	0.24	-0.21	0.04	0.797
$\theta_{-1500kPa}$	0.88	0.24	-0.09	0.24	0.892
Clay	0.92	-0.15	0.05	0.20	0.905
Silt	-0.71	0.50	-0.08	-0.13	0.783
Sand	-0.81	-0.23	-0.01	-0.20	0.748
OM	-0.09	0.22	0.80	0.05	0.698
BD	-0.36	0.30	-0.36	0.73	0.882
PL	0.89	-0.15	-0.01	-0.19	0.857
SC	-0.84	0.12	-0.28	-0.07	0.802
Gravel	-0.35	-0.79	0.05	0.28	0.822
CEC	0.50	-0.53	-0.17	0.11	0.575
SSA	0.63	0.25	-0.25	0.04	0.524
θ_{hy}	0.27	0.14	0.72	0.26	0.669
Eigenvalue	6.47	1.60	1.53	1.13	
Variance explained	46.20	11.43	10.93	8.08	
Cumulative percentage	46.20	57.63	68.56	76.64	

† PC1, principal component 1; PC2, principal component 2; PC3, principal component 3; PC4, principal component 4.
‡ See Table 1 for the abbreviation.
¶ Communality.

Table 6.5. Squares of loadings of 14 variables on significant PCs. The highest values are in bolds.

Variable ‡	Square of PC loadings (in %)†			
	PC1	PC2	PC3	PC4
θ_{-1kPa}	51.80	0.50	1.10	24.00
θ_{-33kPa}	68.90	5.80	4.40	0.20
$\theta_{-1500kPa}$	77.40	5.80	0.80	5.80
Clay	84.60	2.30	0.30	4.00
Silt	50.40	25.00	0.60	1.70
Sand	65.60	5.30	0.00	4.00
OM	0.80	5.80	64.00	0.30
BD	13.00	9.00	13.00	53.30
PL	79.20	2.30	0.00	3.60
SC	70.60	1.40	7.80	0.50
Gravel	12.30	62.40	0.30	7.80
CEC	25.00	28.10	2.90	1.20
SSA	39.70	6.30	6.30	0.20
θ_{hy}	7.30	2.00	51.80	6.80

† PC1, principal component 1; PC2, principal component 2; PC3, principal component 3; PC4, principal component 4.
‡ See Table 1 for the abbreviation.

6.3.3.2 Loading plot: PCs – soil properties relationships

Graphical representations are advocated for a better interpretation of PCA results. The loading plot presenting the original soil properties in the PC1-PC2 plan (Figure 6.4) confirms the results from Table 6.4 and indicate that $\theta_{-1\text{kPa}}$, $\theta_{-33\text{kPa}}$, $\theta_{-1500\text{kPa}}$ are highly positively related to clay, plastic limit, and specific surface area whereas they are highly negatively linked to soil carbonate, sand, and silt on the PC1. The other soil properties are near the origin of the PC1 axis and have a negligible influence on the soil water contents. Regarding PC2, soil water contents and the different soil properties, except gravel and CEC to a lesser extent, are again near the origin of the PC2 axis and indicate that there is a very limited influence of the soil properties on soil water contents. A similar plot in the PC1-PC4 plan, not shown, confirmed the strong negative relation between $\theta_{-1\text{kPa}}$ and bulk density.

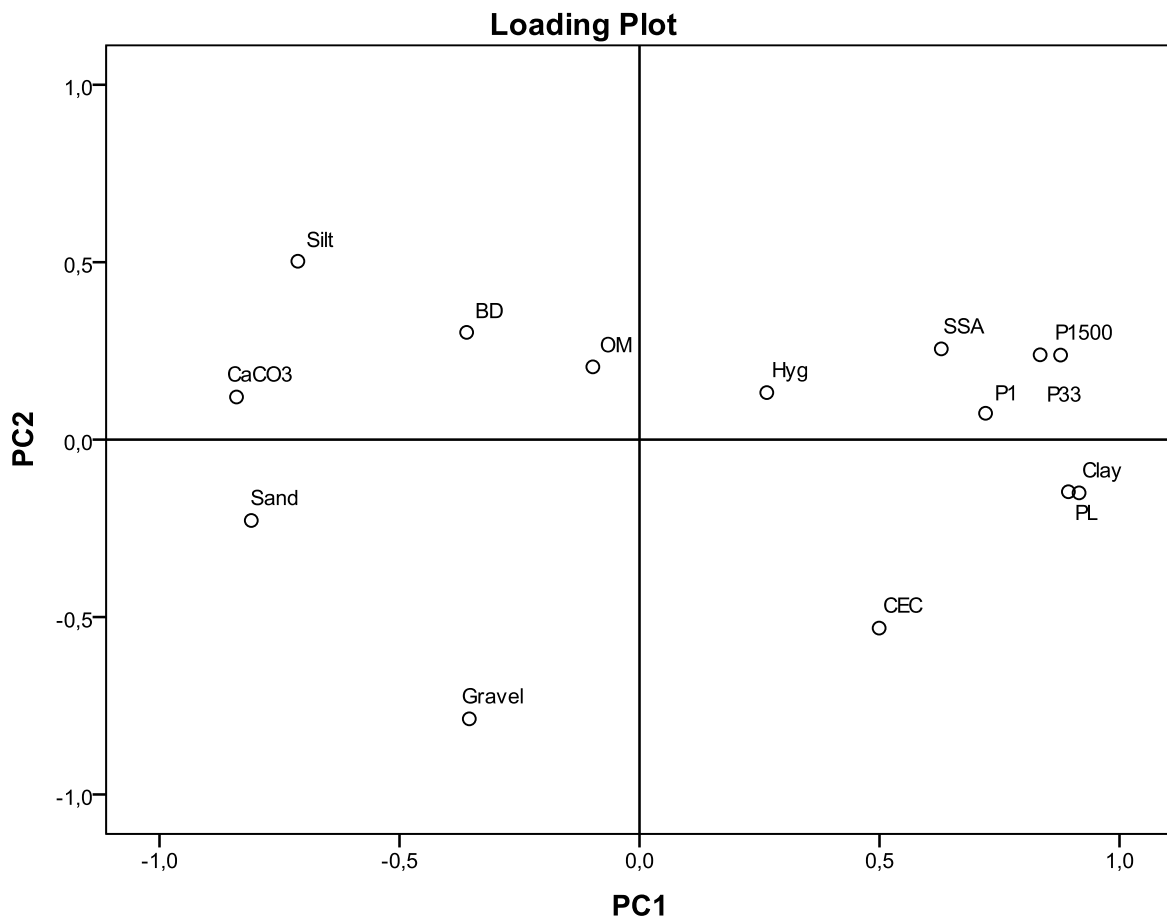


Figure 6.4. Loading Plot: soil properties in the PC1-PC2 plan.

6.3.3.3 Score plot: PCs – soil samples relationships

Another graphical representation similar to the loading plot can be done representing this time the observations or the samples, instead of the original soil properties, in two PCs plan. Figure 6.5 is an example of a score plot using the PC1-PC2 plan with different symbols showing the belonging of the 72 samples to the four different zones. Its interpretation should keep in mind what PC1 and PC2 represent (Table 6.4 and Figure 6.4) in terms of soil properties and soil water contents at the three matric potentials.

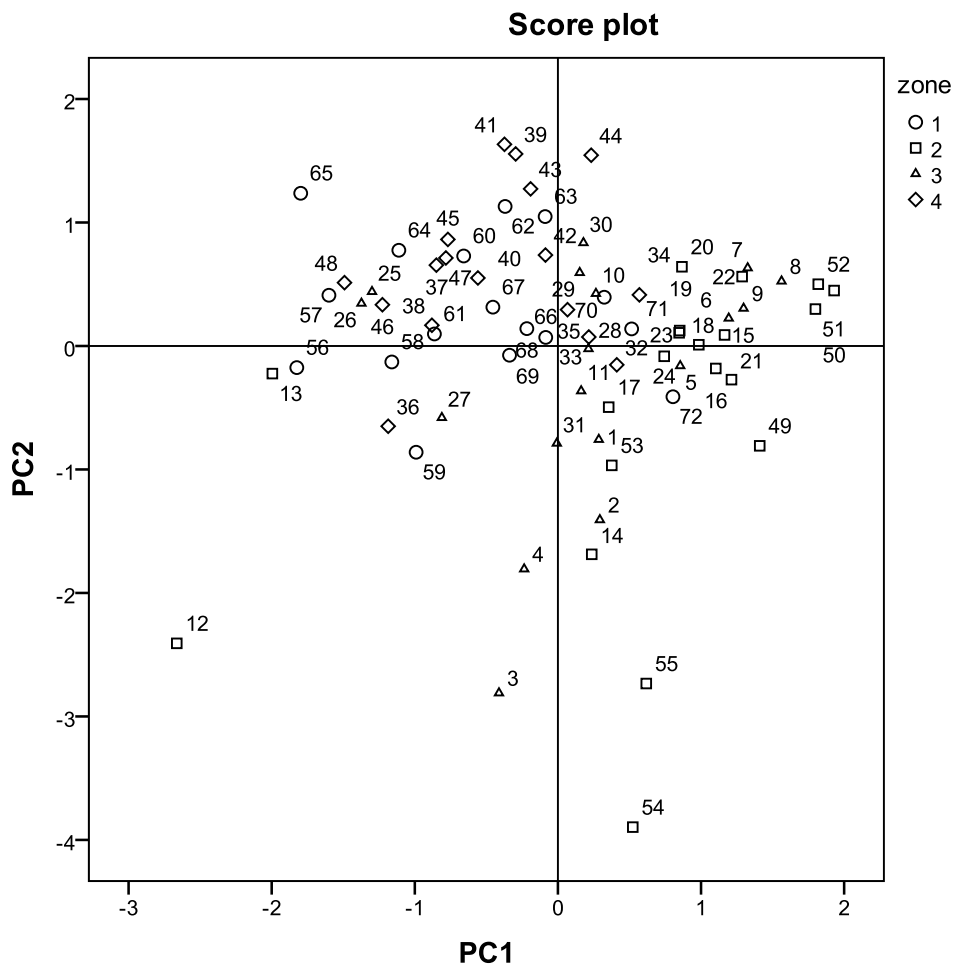


Figure 6.5. Score Plot: soil samples in the PC1-PC2 plan.

In respect to PC1, only 3 of the 17 soil samples (numbers 70, 71, and 72) from the first zone have weak positive scores indicating that they are characterized by relatively higher soil

water and clay contents, plastic limit, and specific surface area whereas the other 14 samples have either weak (samples 63 and 66) or strong (samples 56 and 65) negative scores implying their low water contents and their high sand, silt, and carbonate contents. In contrast, almost all of the soil samples from zone 2 (18 from 20) have intermediate to strong (samples 50, 51, and 52) positive scores indicating that they have high to very high soil water and clay contents, plastic limit, and specific surface area. Indeed, the maximal values for soil water contents and plastic limit correspond to one of these 3 soil samples while their clay content is near the maximal level. The remaining two samples from zone 2 (samples 12 and 13) have a strong negative scores thus low soil water contents and very high carbonate, sand, and silt contents. In fact, the minimal values for $\theta_{-33\text{kPa}}$, $\theta_{-1500\text{kPa}}$, clay, and plastic limit correspond to one of these samples while $\theta_{-1\text{kPa}}$ is not far from the minimal value. Concerning zone 3, 13 from the 18 soil samples have weak (samples 29, 30, and 31) to strong (samples 7, 8, and 9) positive scores while the remaining 5 samples have weak (samples 3 and 4) to strong (samples 25 and 26) negative scores. Finally, 5 soil samples from zone 4 (samples 35 and 44 among others) have weak positive scores; the 12 remaining soil samples have weak (samples 42 and 43) to strong (samples 46 and 48) negative scores.

Dealing with PC2, most of the soil samples from the four zones have weak to intermediate either positive or negative scores. However, some samples have particularly high to very high scores: they are exclusively from zones 2 (samples 12, 54, and 55) and 3 (sample 3). These samples are characterized by their very high content of gravel. Indeed, the maximal value for this soil property corresponds to that from sample 54 and the second highest corresponds to that from sample 12.

Regarding PC3, even though it is not correlated to soil water contents like PC2, it is informative about organic matter and hygroscopic water contents of soil samples since these two soil properties are linked to PC3. Soil samples 58, 62, and 66 (all from zone 1) had the highest positive scores and, consequently among the highest OM and hygroscopic water contents. In fact, the highest and the second highest OM contents correspond to samples 58 and 62 while the third highest hygroscopic water content corresponds to sample 66. In opposition, samples 8 (zone 3), 47 (zone 4), 55 (zone 2), and 65 (zone 1) had the highest negative scores and, consequently among the lowest OM and hygroscopic water contents. Indeed, sample 55 had the second lowest OM content while sample 65 had the third lowest hygroscopic water content.

An interesting feature may be inferred from PC4 since it explains 24% of the variance of $\theta_{-1\text{kPa}}$. From a graph with PC1-PC4 plan not shown here, soil samples 45 (zone 4), 54 (zone 1), and 71 (zone 2) showed the highest positive scores implying that they have among the highest bulk density and, as PC4 is negatively correlated with bulk density, lower $\theta_{-1\text{kPa}}$ values. In contrast, soil samples 38, 39, 40 (zone 4), and 13 (zone 2) had the highest negative scores indicating they have the lowest bulk density and higher $\theta_{-1\text{kPa}}$ values.

As mentioned earlier, the main idea behind PCA is to reduce the dimensionality of the data, so that their manipulation becomes easier. So far, we derived only few (new) variables (four components) which are uncorrelated and are still able to explain most of the information from the original data (14 soil properties).

6.4 Conclusions

This study was performed to explore the interaction between the SWRC and other physical and chemical properties of selected dryland soils in Syria. Specifically, we explored the usefulness of some attributes, like the plastic limit and SSA in addition to the common soil physical and chemical attributes that come from soil survey, in developing and improving PTFs. The correlation results indicate that soil water contents at -33 and -1500 kPa were strongly linked to 11 soil properties, all the properties investigated except organic matter, θ_{hy} , and bulk density; however, soil water content at the -1 kPa matric potential was also highly correlated with bulk density in addition to the same eight soil attributes found to be correlated with soil water contents at -33 and -1500 kPa. In fact, there is a distinct influence of plastic limit, texture, soil carbonate, and SSA on soil water retention. Neither plastic limit nor surface area measurements are routinely collected by soil surveys. The plastic limit is very easily measured as described above and integrates many other soil properties; however, the importance of including the plastic limit needs to be checked in a second step during the building of the PTFs.

Principal component analysis summarized the relationships between the soil characteristics and water retention. It identified four components that explained 77% of the total variance in the data. The water content at the three matric potentials (-1 , -33 , and -1500 kPa) corresponds to the first PC, which is correlated with texture (clay, silt, and sand), the plastic

limit, soil carbonates, and, to a lesser extent, SSA. In addition, a nonnegligible percentage (24%) of the variance in water content at the first matric potential (−1 kPa) is explained by PC4, which is highly negatively correlated with soil bulk density. Consequently, these seven soil properties have a strong (either positive or negative) impact on soil water content at the three matric potentials, whereas the remaining four soil properties (organic matter, gravel, CEC, and θ_{hy}) have a negligible influence on soil water contents: they explain only 0.8 to 5.8% of the variance in soil water contents. Therefore, PTFs can be estimated using either the four PCs or the original seven variables linked to the PCs and different statistical methods like multiple regression, artificial neural networks, etc

Chapter 7 Support vector machines to enhance the performance of pedotransfer functions for predicting the water retention properties of calcareous soils

Based on: Khlosi M., Cornelis W.M., Hazzouri A., Alhamdoosh M., and Gabriels D. Support vector machines to enhance the performance of pedotransfer functions for predicting the water retention properties of calcareous soils. *Submitted in 2014 for publication.*

Abstract

Knowledge of soil hydraulic properties is indispensable for land management in dryland areas. The most important properties are the soil-water retention curve (SWRC) and hydraulic conductivity characteristics. Direct measurement of the SWRC is time and cost prohibitive. Pedotransfer functions (PTFs) utilize data mining tools to predict SWRC. Nowadays, modern data mining techniques have become crucial in enabling high accuracy and good generalization to novel data. In this study we explore the use of Support Vector Machines (SVMs), a novel type of learning algorithm based on statistical theory, for predicting SWRC from more easily and cheaply measured properties. 72 undisturbed soil samples have been collected from different agro-climatic zones of Syria. The soil water contents at eight matric potentials were determined and selected as output variables. A brief overview of the theoretical background of this fairly new technique and the use of specific kernel functions are presented. Then, the model parameters were optimized with cross-validation and grid-search method. The performance of the SVM-based PTFs was analyzed using the coefficient of determination, root mean square error (RMSE) and mean error (ME). This study shows that SVMs have the potential to be a useful and practical tool for predicting the SWRC of calcareous soils of dryland areas. They support previous findings in that they perform better than ANN and MLR.

7.1 Introduction

Knowledge of the soil hydraulic properties, namely the soil water retention curve (SWRC) and hydraulic conductivity function, is essential for many agricultural, environmental and engineering applications. This knowledge is needed, for example, to describe and predict water and solute transport, as well as to model heat and mass transport in the unsaturated (vadose) zone between the soil surface and the groundwater table. The SWRC, which defines the relationship between the matric potential (ψ) and soil water content (θ), can be determined in the laboratory or in the field. However, current methods and experimental techniques are expensive, tedious and time-consuming to accomplish, especially for fine-textured soils. Therefore, indirect methods are based on translating the SWRC from more easily measurable and more readily available soil properties. Bouma (1989) introduced the name pedotransfer functions (PTFs) for such predictive functions. Enormous advances have been made during recent decades in developing different PTFs.

Support vector machines (SVMs), which are one of the modern data mining techniques, have gained much attention as a result of their strong theoretical background. SVMs are a class of machine learning algorithms that can perform pattern recognition and regression based on the theory of statistical learning and the principle of structural risk minimization (Vapnik, 1995; Müller et al., 2001). SVMs try to model the input variables by finding the separating boundary, called hyperplane, to reach classification of the input variables: if no separation is possible within a high number of input variables, the SVMs algorithm still finds a separation boundary for classification by mathematically transforming the input variables by increasing the dimensionality of the input variable space. SVMs are one of the most promising and powerful machine learning techniques that have been proved to be very successful in many applications. However, only few attempts have been made to utilize this technique to predict the SWRC (Lamorski et al., 2008; Twarakavi et al., 2009; Lamorski et al., 2013).

On the other hand, soil texture (sand, silt, and clay percentages) is one of the most frequently utilized soil properties for predicting the SWRC. The main assumption underlying most PTFs, especially in earlier studies, is that textural properties dominate the hydraulic behavior of soils (Lilly and Lin, 2004; Weynants et al., 2009). Other commonly used soil physical properties for PTFs prediction are organic carbon (OC), and bulk density. Additional readily

available soil properties were rarely used in developing PTFs (Wösten et al., 2001). Recently, Pachepsky et al. (2001), Leij et al. (2004), and Sharma et al. (2006) included certain available topographical and remotely sensed vegetation attributes, in addition to soil physical parameters, for developing PTFs. Zacharias and Wessolek (2007) developed a new PTF that forgoes the use of OC as a predictor.

The Atterberg limits, which describe steps of soil consistency such as the plastic limit (PL), can also be cheaply and easily determined. The PL is the minimum gravimetric water content at which the soil can be deformed without rupture (Dexter and Bird, 2001). As previously mentioned in Chapter 6, PL might be a good additional predictor. Odell et al. (1960) found a high correlation between Atterberg limits and three soil properties (percent of organic carbon, percent of clay, and percent of 2 : 1 clay minerals in the clay). Since 2 : 1 clay minerals are dominant in dryland areas (Jordán et al., 2003), the Atterberg limits can also be sensitive indicators of the type of clay. A novel feature of this study is to utilize additional readily available soil properties in predicting the SWRC at eight matric potentials using the SVMs technique.

The primary objective of the present study is to explore the accuracy of Support Vector Machines (SVMs) compared with artificial neural networks (ANN) and multiple linear regression (MLR) in predicting the soil water retention curve of calcareous soils. Secondary objective is to compare the performance of SVMs, ANN and MLR using different groups of readily available soil properties.

7.2 Materials and methods

7.2.1 Data description

The same data sets considered in chapter 5 are used here (see chapter 5, section 5.2.1). In determining the sand, silt and clay fraction, soil carbonates were not removed for reasons outlined in Chapter 5.

7.2.2 Support vector regression

The SVM algorithm is a powerful machine learning tool based on firm statistical and mathematical foundations concerning generalization and optimization theory. SVM was first introduced by Vapnik (1995) but soon it began to enjoy strong theoretical foundations and excellent empirical successes in many pattern-recognition applications. Because of its outstanding empirical performance, SVM has been well accepted by many scientific communities. Compared to artificial neural network, SVM has some advantages such as: it does not suffer from the local minima problem, it has fewer learning parameters to select, and it produces stable and reproducible results. The basic idea of SVM is to transform the samples into a high-dimensional feature space and construct an optimal separating hyperplane that maximizes its distance from the closest training samples.

SVMs were developed to solve the classification problem and later they have been extended to the domain of regression problems (Vapnik, 1998; Smola, 1998). In literature, the terminology for SVM is sometimes confusing. Gunn (1997) used the term SVM to refer to both classification and regression methods, and the terms support vector classification (SVC) and support vector regression (SVR) to the specific problems of classification and regression respectively. SVR is a novel machine-learning method that is receiving more and more attention and has been successfully applied in the prediction tasks of soil-water content (Wu et al. 2008), SWRC (Lamorski et al., 2008; Twarakavi et al., 2009), soil organic carbon normalized sorption coefficient (Wang et al., 2009), soil texture and pH (Kovačević et al., 2010).

The purpose of the regression model is to estimate an unknown continuous-valued function $y = f(x)$, which is based on a finite number of samples. In the present study, we want to investigate the relationship function between certain soil properties and soil-water content at given matric potentials. In order to achieve this, we apply ε -insensitive support vector regression (ε -SVR) (Vapnik, 2000). In an ordinary learning problem, a set of training data $(x_i, y_i), i = 1, \dots, N$ is given and is used to generate a dependency between the input x and output y . Hence, the expected form of support vector regression can be formulated as

$$f(x) = \sum_{i=1}^n w_i \phi(x_i) + b \quad [7.1]$$

where w is a weight vector defining the solution of the primal formulation, $\phi(x_i)$ is the point in feature space that is nonlinearly mapped from the input space x , and b is a scalar representing the bias. To estimate the function $f(x)$, the SVR framework can be adopted based on the ε -insensitive loss function (Vapnik, 1995). The objective is to find an $f(x)$ that has the most ε deviation from the observed soil-water content y_i for all the training data and at the same time is as flat as possible. Introducing slack variables ξ_i, ξ_i^* quantifying estimation errors greater than ε , the SWRC estimation task can be formulated as a constrained optimization problem: we search for w and b that minimize the regularized loss. Hence, the SVR problem can be formulated as:

$$\begin{aligned} & \text{minimize } \frac{1}{2} w^T w + C \sum_{i=1}^n (\xi_i + \xi_i^*) & [7.2] \\ & \text{subject to } \begin{cases} w^T \phi(x_i) + b - y_i \leq \varepsilon + \xi_i \\ y_i - w^T \phi(x_i) - b \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \quad i = 1, \dots, n \end{cases} \end{aligned}$$

where C is the regularization parameter that determines the trade-off between model complexity (flatness) and the degree to which deviations larger than ε are tolerated in the optimization formulation.

To solve this optimization problem, which is a convex quadratic program, two Lagrange multipliers were introduced for every constraint, and hence the final regression function can be formulated as:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad [7.3]$$

where α_i , and α_i^* are Lagrange multipliers for the first two constraints, and the kernel function $K(x_i, x)$ which is in our case a radial basis function (RBF kernel). The support vectors are those with corresponding non-zero values of the Lagrange multipliers. The RBF kernel was used to train and construct our SVR classifiers. Here we utilized the following RBF kernel:

$$K(x_i, x) = \exp(-\gamma \|x_i - x\|^2) \quad [7.4]$$

where γ is a positive free parameter defining the kernel width. In order to get a “good” model, the regularization parameter C and the kernel parameter γ need to be selected properly.

7.2.3 Artificial neural networks

ANN has been used extensively for predicting the SWRC. For process modeling, the commonly used network type is the feed-forward back-propagation network which usually consists of three layers. The layers are described as input, hidden, and output layers (Figure 7.1). Each neuron of the hidden layer performs a weighted sum on its input signals x_i and a bias term w_0 , and passes the result, $s = \sum_{i=1}^n w_i x_i + w_0$, through a nonlinear activation function $f(s)$. A commonly used activation function to introduce nonlinearity is the hyperbolic tangent function. A bias term (w_0) was added, serving as a constant added to the weight. Initially, the weights are chosen randomly. The same procedure is repeated in the output layer transforming the output from the hidden layer to the final output (Figure 7.1). The feed-forward process will stop once the output is predicted. Back propagation algorithms attempt to minimize the error of the mathematical system represented by neural network's weights and hence walk downhill to the optimum values for weights. Error is estimated as difference between actual and computed outputs. The error is back-propagated from the output nodes to the hidden and from the hidden nodes to the input nodes and the weights are altered according to the generalized delta rule. With several iterations, called training, the network outputs will eventually converge towards the desired outputs.

It is worthy to mention that SVMs have been developed in the reverse order to the development of neural networks (Wang, 2005). The development of ANNs followed a heuristic path, with applications and extensive experimentation preceding theory. In contrast, the development of SVMs involved sound theory first, then implementation and experimental.

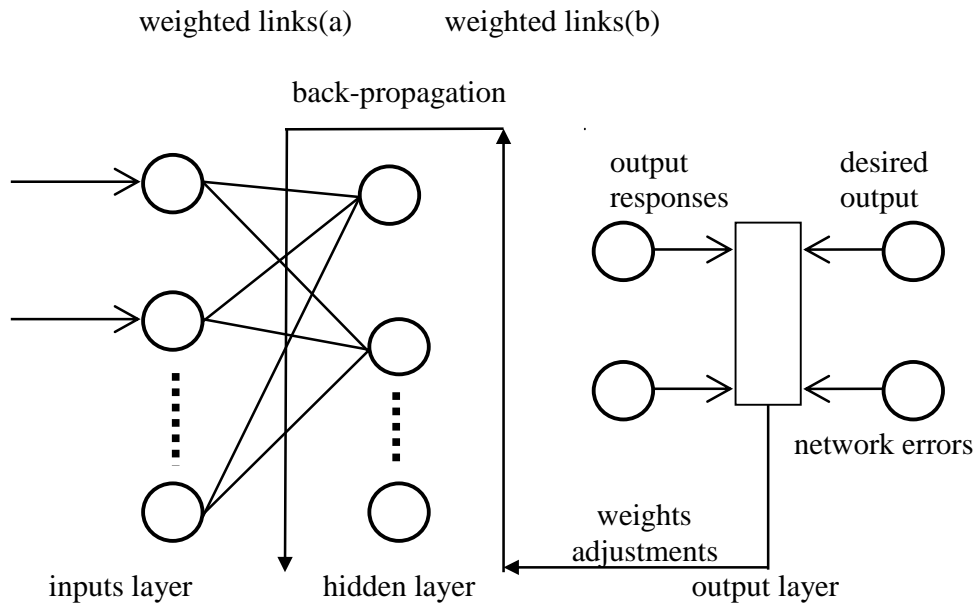


Figure 7.1. Architecture of artificial neural network.

7.2.4 Evaluation criteria

Various statistical criteria have been proposed in the literature for evaluating model predictions. We consider two criteria to compare the predictive performance of the MLR, ANN and SVR models: (1) root mean square error (RMSE); and (2) the coefficient of determination (R^2).

Leave-one-out cross-validation was used to validate the MLR, ANN and SVR models. One soil-water content observation was set aside at a time, and the remaining N-1 soil-water contents were used to build a model. The resulting model was then used to predict the outcome of the sample that was held out. This procedure was repeated N times, with each sample being held out exactly once. In this study, the ANN analysis was performed using the Neural Network Toolbox of MATLAB environment (The MathWorks, Natick, MA) and the SVR was performed under the Python environment.

7.3 Results and discussion

7.3.1 Performance comparison of MLR, ANN and SVR

Table 7.1 gives an overview of the prediction performance of MLR, ANN and SVR models determined by means of leave-one-out cross-validation using 72 training examples and five predictors (sand, silt, clay, OC, and bulk density) for predicting the SWRC at eight matric potentials. The RMSE values of MLR and ANN models varied from 0.052 to 0.074 $\text{m}^3 \text{m}^{-3}$ and from 0.042 to 0.054 $\text{m}^3 \text{m}^{-3}$, respectively, which were larger than those of SVR models, ranging from 0.039 to 0.051 $\text{m}^3 \text{m}^{-3}$. The R^2 of MLR and ANN models ranged from 0.06 to 0.54 and from 0.17 to 0.57, but the SVR models showed the best results (0.29 to 0.62). In terms of t-statistic ($P < 0.05$), the differences in performance of the three techniques were significant at matric potentials of -5, -7, -10, -100, and -1500 kPa. Hence, the predictions by the SVR-based PTFs showed considerable improvement over the MLR and ANN. Applying the same statistics, the accuracy of existing PTFs varies to a noticeable degree. When the same input predictors were used, the RMSE values for our SVM PTF are lower than those obtained by Twarakavi et al. (2009) when applying support vector machines (RMSE = 0.053) and ANN (using Rosetta) (RMSE = 0.068) for predicting soil-water content using the database from Schaap and Leij (1998) and Schaap et al. (2001). Typical values of RMSE achieved with PTFs to predict soil water retention range from 0.02 to 0.07 $\text{m}^3 \text{m}^{-3}$ (Pachepsky et al., 1999). Therefore, the SVM PTF developed in this study can be considered as having a moderate accuracy. Similarly, in comparing the ANN and SVR methods, Lamorski et al. (2008) found that SVR performed better at some matric potentials.

Table 7.1. Prediction performance comparison of MLR, ANN and SVR (PTF1).

ψ	MLR		ANN		SVR	
	RMSE	R^2	RMSE	R^2	RMSE	R^2
-1	0.052	0.210	0.042	0.334	0.039	0.415
-3	0.066	0.065	0.053	0.175	0.049	0.278
-5	0.070	0.177	0.054	0.205	0.051	0.294
-7	0.072	0.223	0.055	0.231	0.051	0.324
-10	0.074	0.233	0.054	0.252	0.050	0.352
-33	0.072	0.335	0.052	0.384	0.050	0.424
-100	0.065	0.463	0.051	0.476	0.047	0.558
-1500	0.059	0.544	0.047	0.576	0.044	0.624

7.3.2 Including additional soil properties for improving PTFs accuracy

Improving the accuracy and reliability of PTFs is not only influenced by applying data mining techniques but also by searching for additional soil properties as inputs in PTFs. As we found in Chapter 6, other easily determinable basic soil properties that might affect SWRC include soil carbonate content, specific surface area and plastic limit. To test their potential as prediction variables, they were additionally included on top of those properties that were selected in 7.3.1.. Water retention of soils with low organic matter, such as in our case, are not likely to be affected by organic matter. Then, OC as a predictor can be replaced by one of these variables. It can be deduced from Chapter 6 that the plastic limit seems the best choice. It is very easily measured and moreover was strongly correlated with soil water contents for all matric potentials as previously described in Chapter 6. To support this finding, a stepwise MLR was performed using a new independent variable group including clay, silt, sand, bulk density, soil carbonate, specific surface area and plastic limit. The new predictors and their coefficients retained for each soil-water matric potential in the new PTF (denoted as PTF2) are tabulated in Table 7.2.

Table 7.2. Pedotransfer function coefficients and their confidence interval, the *t*-statistic and its *p*-value.

ψ	Predictor	Coefficient	Value	<i>t</i> - statistic	p-value	Confidence Interval	
						Lower Bound	Upper Bound
-1	intercept	a_1	0.3164	12.9799	0.0000	0.2673	0.3654
	plastic limit	b_1	0.5086	8.1656	0.0000	0.3833	0.6339
-3	intercept	a_2	0.2644	9.5070	0.0000	0.2085	0.3204
	plastic limit	b_2	0.5620	7.9077	0.0000	0.4191	0.7050
-5	intercept	a_3	0.2474	8.2460	0.0000	0.1870	0.3077
	plastic limit	b_3	0.5618	7.3289	0.0000	0.4076	0.7160
-7	intercept	a_4	0.0166	0.2092	0.8352	-0.1429	0.1760
	plastic limit	b_4	0.6969	8.4316	0.0000	0.5305	0.8632
-10	bulk density	c_4	0.1221	2.7389	0.0087	0.0324	0.2118
	intercept	a_5	-0.0068	-0.0857	0.9321	-0.1665	0.1529
	plastic limit	b_5	0.7177	8.6696	0.0000	0.5511	0.8843
-33	bulk density	c_5	0.1246	2.7918	0.0076	0.0348	0.2145
	intercept	a_6	0.0134	0.1561	0.8766	-0.1590	0.1858
	plastic limit	b_6	0.6850	7.1216	0.0000	0.4912	0.8787
-100	bulk density	c_6	0.1187	2.8194	0.0071	0.0339	0.2036
	sand	d_6	-0.0013	-2.1061	0.0408	-0.0026	-0.0001
	intercept	a_7	0.0498	0.6134	0.5427	-0.1137	0.2133
	plastic limit	b_7	0.6572	7.2041	0.0000	0.4735	0.8410
-1500	bulk density	c_7	-0.0021	-3.4160	0.0014	-0.0033	-0.0008
	sand	d_7	0.0809	2.0263	0.0487	0.0005	0.1614
	intercept	a_8	-0.0511	-0.6858	0.4964	-0.2012	0.0990
	plastic limit	b_8	0.6306	7.5310	0.0000	0.4620	0.7993
-1500	bulk density	c_8	-0.0028	-4.9670	0.0000	-0.0039	-0.0016
	sand	d_8	0.1276	3.4804	0.0011	0.0538	0.2015

Results indicate, firstly, that soil carbonate and specific surface area were not significantly different from zero ($p > 0.05$). Secondly, the plastic limit once again appears to be a valuable predictor especially for a matric potential equal or higher than -10 kPa. This is because the matric potential at the plastic limit is the main cohesive stress and ranges between -63 and -200 kPa. The latter finding is also supported by Figure 7.2 in which, as an example, measured soil-water content values are plotted against fitted water content values at eight different matric potentials for a silt loam soil. The Khlosi et al. (2006) model was selected to be fitted to both the lab-measured and PTF-predicted SWRC. As shown in Figure 7.2, we can see that the SWRC of PTF1 (see Chapter 5, Table 5.2) deviates more from the real points compared to the PTF2.

Now we will compare the MLR, ANN and SVR methods using a new independent variable group including clay, silt, sand, bulk density and plastic limit. Table 7.3 shows the values of the statistical indices, which were computed to evaluate the prediction accuracy of these three techniques. When considering RMSE, the SVR showed again the lowest values, meaning that the curve produced the highest match with the measured SWRC. As regards the mean of R^2 , a similar trend could be observed, with SVR as the best prediction in terms of linearity followed by the ANN PTF. Including the plastic limit as a predictor provided better predictions of the SWRCs for the three statistical techniques as compared to section 7.3.1. Table 7.3 shows that plastic limit lowers the RMSE values of MLR, ANN and SVR. In the particular case of the SVR-based PTF, all RMSE values are now below $0.04 \text{ m}^3 \text{ m}^{-3}$. In comparison with typical RMSE values found in literature, this PTF performs very well. There is good evidence here that the plastic limit is a powerful predictor of SWRC. As described earlier in Chapter 6, the plastic limit was strongly correlated with soil water contents for all matric potentials. There were also many positive and negative correlations between the plastic limit and other soil properties. Only organic matter was not correlated with the plastic limit because our soils are poor in organic matter. Keller and Dexter (2012) found a positive correlation between the plastic limit and soil organic matter. They predicted that the clay content must be at least 10% for soils without organic matter to be plastic; however, soils with <10% clay can be plastic if organic matter is present. Rawls and Pachepsky (2002) stated that the soil structure and consistence properties, such as Atterberg limits, can serve as predictors of soil hydraulics properties. In their study, plasticity class, grade class, and dry consistency class were leading predictors of soil-water retention at both -33 kPa and -1500

kPa matric potentials. Thus, we conclude here that including PL with other basic soil properties in predicting the SWRC leads to better accuracy as it integrates many other soil properties.

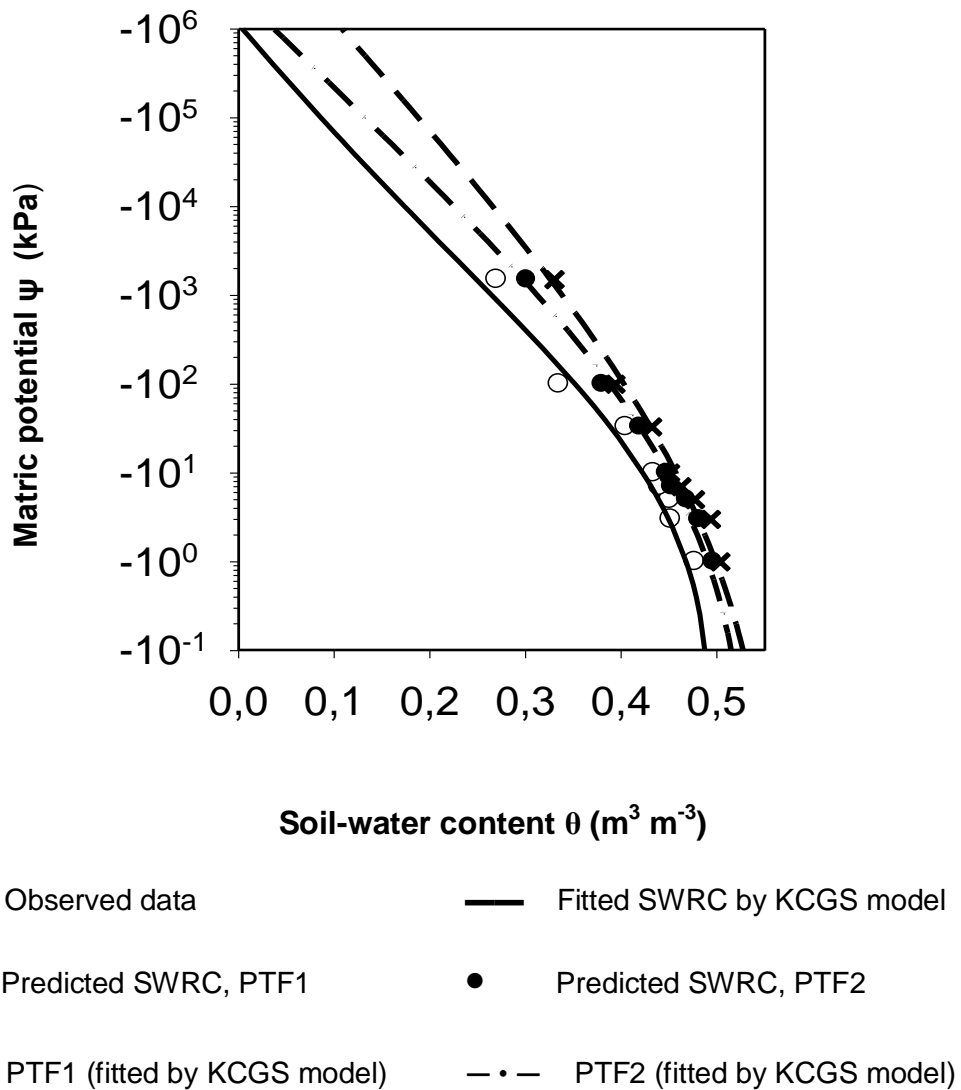


Figure 7.2. Measured and predicted soil-water retention curves for a silt loam calcareous soil.

Table 7.3. Comparison of the prediction performance of MLR, ANN and SVR (PTF2).

ψ	MLR		ANN		SVR	
	RMSE	R ²	RMSE	R ²	RMSE	R ²
-1	0.0465	0.2497	0.0334	0.5726	0.0318	0.6119
-3	0.0569	0.2101	0.0364	0.6052	0.0353	0.6304
-5	0.0613	0.1625	0.0404	0.5548	0.0366	0.6357
-7	0.0556	0.3244	0.0383	0.6189	0.0362	0.6602
-10	0.0541	0.3459	0.0394	0.6002	0.0356	0.6733
-33	0.0520	0.4690	0.0363	0.6958	0.0359	0.7032
-100	0.0511	0.5705	0.0411	0.6638	0.0394	0.6909
-1500	0.0428	0.6838	0.0358	0.7502	0.0346	0.7670

7.4 Conclusions

This study was carried out to compare support vector machines regression performance with the one of artificial neural networks and multiple linear regression techniques. Using a data set taken from 72 horizons of different soils in Syria and representing eight soil-textural classes, results showed that the performance of support vector machines was the best in terms of RMSE and R². On the other hand, improvements (of different statistical significance) were found for the SWRC by replacing OC with plastic limit which is easier and cheaper to measure. Our investigations show that plastic limit can be a suitable parameter to successfully represent the integrated effect of several soil properties on water retention. However, more research is needed to better quantify the effects of plastic limit on the SWRC for different regions and soils. We therefore recommend the use of support vector machines and plastic limit to further improve and develop PTFs.

Chapter 8 General conclusions and future research

8.1 Introduction

Modeling water flow and solute transport in soils are an essential means to address many problems in applied soil science, such as water, nutrient, and salinity management research. The research presented in this dissertation had four objectives: 1) to provide an alternative model to describe the soil-water retention curve over a range of water contents from saturation to oven dryness; 2) to investigate the effect soil carbonates on the predicted soil water retention curve of dryland soils; 3) to introduce additional predictor variables which are easily and cheaply determined; and 4) to develop pedotransfer functions for predicting soil-water retention curve of Syrian soils.

In this chapter, we briefly summarize our major findings and the contributions of this research work to the accomplishment of these objectives. Then, some recommendations for further research are mentioned.

8.2 General conclusions

8.2.1 Soil water retention curve equations

In Chapter 2, the commonly used retention models only account for capillary water retention. Adsorptive water retention is neglected. This leads to erroneous description of hydraulic properties in the dry range. In a first application (Chapter 2), we compared 10 SWRC models using data set taken from 48 horizons of forest soils in Flanders, Belgium. The van Genuchten (1980) and Kosugi (1994) models showed best fits to the observed data, specifically at high and medium water content. A high performance was also observed for the Kosugi (1999) model, which is a relatively simpler functional form of Kosugi (1994) model. Most of the compared models do not define the soil-water content vs. soil-matric potential relationship beyond the residual water content. The only model we evaluated that is able in doing so is the Rossi and Nimmo (1994) model. However, it showed the lowest performance

in terms of goodness-of-fit. It seems that more recently developed expressions for the SWRC between saturation and oven dryness need to be evaluated or new expressions should be developed. Therefore, in Chapter 3, we modified the Kosugi (1999) model to describe the SWRC between saturation and oven dryness. Our modification retains the form of the original Kosugi function in the wet range and transforms to an adsorption equation in the dry range. The predictive capability of our extended model was further evaluated under reduced sets of data that do not contain observations below a matric potential of -100 kPa. This evaluation showed that our model successfully predicted the water content with acceptable uncertainty, even when using the limited data set. In comparison with other models (Chapter 4), our expression was most consistent for different soils. Moreover, its prediction potential was relatively good as demonstrated by the significant correlation between its parameters and basic soil properties, which is promising for developing pedotransfer functions. Retention data were taken for 137 soils covering nearly all USDA textures from the UNSODA data base.

A major achievement of our model is its ability to predict the entire SWRC when calibrated using a limited data set that includes only those measurements of water content at matric potentials greater than -100 kPa. This was also confirmed by Lu et al. (2008) who compared our model with other models using a data set from saturation to oven dryness. They found that our model produced the best results when reduced data sets greater than -300 kPa were used for model establishment. Hence, our model is a continuous function from saturation to oven dryness and suitable for the different soils (Chapter 3 and 4).

8.2.2 Pedotransfer functions for predicting SWRC of Syrian soils

All published PTFs to date use soil texture as essential key predictor to most soil hydraulic PTFs. The majority of used information on texture was analyzed by a methodology that considers removing soil carbonates as a pretreatment. In our study (Chapter 5) we investigated the influence of pre-treatment on sand, silt, and clay fractions and compare their capability to predict the SWRC for calcareous soils. 72 soil samples were collected from rural areas throughout northwest Syria, covering most of its agro-climatic zones and soil types. Two procedures differing in the pre-treatment process were used. In the most widely used technique, carbonates were removed by hydrochloric acid, while in the alternative one, carbonates were not removed. Our results showed great variability in the sand, silt, and clay

fractions for both methods. Only texture that forgoes the pre-treatment can be translated into the soil-water retention curve. Also, it can be concluded from this study that the PSD_{+C} method (without removal of carbonate) is more adequate to classify soil texture and hence to predict the SWRC. The advantages associated with this method include decreasing the time and work and allowing accurate measurements of calcareous soils.

In our study on the other hand (Chapter 6), we investigated the interaction between SWRC and other physical and chemical properties of dryland soils. Specifically, we explored the usefulness of some attributes in addition to soil physical attributes that come from soil survey in developing and improving PTFs. We found a distinct influence of plastic limit and other attributes (such as specific surface area) on soil water retention. Neither plastic limit nor surface area measurements are routinely collected by soil surveys. Plastic limit is very easily measured as described previously and integrated many other soil properties. These results suggested including plastic limit in regular soil surveys. Principal component analysis summarized the relationships between the soil characteristics and water retention. It identified four components that explained 77% of the total variance in the data. The water content at the three matric potentials (-1, -33, and -1500 kPa) corresponds to the first PC, which is correlated with texture (clay, silt, and sand), the plastic limit, soil carbonates, and, to a lesser extent, SSA. In addition, a nonnegligible percentage (24%) of the variance in water content at the first matric potential (-1 kPa) is explained by PC4, which is highly negatively correlated with soil bulk density.

Finally (Chapter 7), three different techniques including support vector machines regression (SVR), artificial neural networks (ANN) and multiple linear regression (MLR) were used to predict the SWRC of dryland areas. It was determined that the SVR technique provided more accurate results than the other techniques that were tested. A comparison of results from the statistical performance demonstrates that the SVR technique predicts the SWRC more accurately than ANN and MLR techniques. This was evident from a lower RMSE and a higher R² value. On the other hand, improvements (of different statistical significance) were found for the SWRC by replacing OC with plastic limit which is easier and cheaper to measure. Our investigations confirm that plastic limit can be a suitable parameter to successfully represent the effects of the OC on water retention. By suggesting plastic limit as

a new predictor, the universality of the PTF has shown to be improved as it integrates other soil properties.

8.3 Future research

8.3.1 Soil water retention curve equations

Unsaturated flow is usually described with the Richards' equation. An accurate knowledge of the soil hydraulic functions is required to solve this equation, i.e., the soil water retention function and the hydraulic conductivity function. Conductivity measurements are indeed highly scale dependent and sensitive to the measurement technique used, especially close to saturation (Weynants et al., 2009). It remains extremely difficult to get reliable measured saturated hydraulic conductivity values given their extreme spatial variability (Verbist et al., 2013). In many cases, hence, measurements of unsaturated hydraulic conductivity are unavailable. As an alternative to direct measurements, models have been proposed to estimate hydraulic conductivity from water retention data, which is more easily measured. The selection of the correct model combination is of crucial importance. The well-established retention functions of Brooks and Corey (1964), van Genuchten (1980), or more recently Kosugi (1996) in combination with the capillary bundle models of Mualem (1976) or Burdine (1953) for conductivity prediction have opened new possibilities to extend the $\psi(\theta)$ relation to the oven-dry condition. Matthews et al. (2010) have proposed a simple model for saturated hydraulic conductivity using the van Genuchten function for the SWRC (van Genuchten, 1980). Nasta et al. (2013) used the water retention parameters of Brooks and Corey (1964) to predict the saturated hydraulic conductivity. The way forward is to extend the capillary models to account for film flow and hence improve hydraulic conductivity prediction. The SWRC equation presented in this dissertation consistently showed a good choice which can be incorporated in Mualem's equation.

8.3.2 Pedotransfer functions for predicting SWRC

This research showed that improvements in the prediction of SWRC were not only associated with applying data mining techniques but also with two additional effects. First, evaluation of available methods for soil texture measurement and selecting the most suitable one. Our

results show that in case of calcareous soils which are abundant in dryland environments, only texture that forgoes the pre-treatment of destroying carbonates can be translated into the soil-water retention curve. More studies should be dedicated in the mineralogy of soil carbonates to investigate the influence of their composition on the SWRC and their contribution to the development of PTFs for soils from dryland regions. Second, searching for additional soil properties as inputs in PTFs. Our results in this dissertation showed that plastic limit can be a suitable parameter to successfully represent the integrated effect of several soil properties on water retention. However, more research is needed to better quantify the effects of plastic limit on the SWRC for different regions and soils. We therefore recommend considering the use of support vector machines and plastic limit to further improve and develop PTFs.

Summary

Soil and water are scarce commodities in any part of the world, more being in dry regions where the amount of rainfall is quite small, the distribution is variable and the frequency is unpredictable. It happens with an erratic nature. All plants need water to grow and they take their water from the soil, which typically is a porous medium. The storage of water in the soil is therefore of crucial importance to plants. In dryland areas, such as Syria, evapotranspiration is often much greater than precipitation and the soil water storage decreases. The soil water flow processes for most months are dominated by the unsaturated hydraulic conductivity properties of the soil. The most important factors in this respect are the soil water retention and the hydraulic conductivity.

However, traditional methods to determine soil hydraulic properties are still difficult, time-consuming and expensive. When such data are not available, pedotransfer functions (PTFs) which utilize physical or empirical relations between soil hydraulic properties and other easily and cheaply measured properties can be used as alternative method. To date, various pedotransfer functions appear in the literature to predict the SWRC and only few PTFs have been developed for soils of the dryland areas. Meanwhile, in Syria published soil hydraulic data is lacking and the collection of a national database is not complete. It is therefore valuable to explore in this thesis the interaction between soil hydraulic properties and other physical and chemical properties in order to estimate the soil-water retention curve from easily measured soil parameters.

In this dissertation some possible improvements in modeling soil water retention were investigated with the goal of developing a practical model that can represent with the minimum possible number of parameters the SWRC over the entire range of saturation. Subsequently, the interaction between soil hydraulic properties and other physical and

chemical properties was studied to estimate the soil water retention curve from easily measured soil parameters.

In Chapter 2, ten closed-form unimodal analytical expressions were evaluated to describing the soil-water retention curve, in terms of their accuracy, linearity, Aikake Information Criterion (AIC), parameter uniqueness and parameter identifiability. Soil samples were taken in duplicate from 48 horizons of 24 soil series in Flanders, Belgium.

In Chapter 3, the Kosugi (1999) model was modified to describe the SWRC between saturation and oven dryness. Our modification retains the form of the original Kosugi function in the wet range and transforms to an adsorption equation in the dry range. The predictive capability of our extended model was further evaluated under reduced sets of data that do not contain observations below a matric potential of -100 kPa. It would be concluded that our model successfully predicted the water content with acceptable uncertainty, even when using the limited data set.

Chapter 4 showed that the proposed model was most consistent for different soils as compared with other models. Moreover, its prediction potential was good as demonstrated by the significant correlation between its parameters and basic soil properties, which is a basis for developing pedotransfer functions. Retention data were taken for 137 soils covering nearly all USDA textures from the UNSODA data base.

In Chapter 5, the influence of pre-treatment on sand, silt, and clay fractions was investigated and compared their capability to predict the SWRC for calcareous soils. 72 soil samples were collected from rural areas throughout northwest Syria, covering most of its agro-climatic zones and soil types. Two procedures differing in the pre-treatment process were used. In the most widely used technique, carbonates were removed by hydrochloric acid, while in the alternative one, carbonates were not removed. Our results showed great variability in the sand, silt, and clay fractions for both methods. Only texture that forgoes the pre-treatment can be translated into the soil-water retention curve. It is concluded in this study that the PSD without removal of carbonate is more adequate to classify soil texture and hence to predict the SWRC.

In Chapter 6, the interaction between SWRC and other physical and chemical properties of dryland soils was investigated. Specifically, this study explored the usefulness of some attributes in addition to soil physical attributes that come from soil survey in developing and

improving PTFs. a distinct influence of plastic limit and other attributes (such as specific surface area) on soil water retention was found. Neither plastic limit nor surface area measurements are routinely collected by soil surveys. Plastic limit is very easily measured as described previously and integrated many other soil properties. Principal component analysis summarized the relationships between soil characteristics influenced by water retention. It identified four factors which are uncorrelated and are ordered by the fraction of the total information each retains.

In Chapter 7, three different techniques including support vector machines regression (SVR), artificial neural networks (ANN) and multiple linear regression (MLR) were used to predict the SWRC of dryland areas. The SVR methodology was successfully applied and showed the best performance in terms of RMSE and R^2 . On the other hand, improvements (of different statistical significance) were found for the SWRC by replacing OC with plastic limit which is easier and cheaper to measure. Our investigations confirm that plastic limit can be a suitable parameter to successfully represent the effects of the OC on water retention. By suggesting the plastic limit as a new predictor, the universality of the PTF has shown to be improved as it integrates other soil properties.

Future research should incorporate the presented soil-water retention model in the hydraulic conductivity functions such as Mualem's equation (Mualem 1976). The use of support vector machines and plastic limit to further improve and develop PTFs that predict the SWRC for different regions and soils should be considered.

Samenvatting

In grote delen van de wereld zijn vruchtbare gronden schaars geworden. Dit fenomeen is nog meer uitgesproken in droge streken met een geringe en onregelmatig verdeelde neerslag. Daarenboven is ook de neerslag daar zeer moeilijk te voorspellen! Niettegenstaande dit alles hebben planten water nodig die ze via het wortelstelsel uit de poreuze bodem dienen te halen. Stockeren van voldoende water in de bodem is aldus noodzakelijk.

Evenwel is in ariede en semi-ariëde gebieden, zoals in Syrië, de hoeveelheid evapotranspiratie hoger dan de neerslag. Daardoor wordt de hoeveelheid gestockeerd water sterk verminderd en ontstaat er een deficit aan water voor de plantengroei.

In die poreuze onverzadigde bodem is de stroom van het water afhankelijk van een aantal bodemfysische parameters zoals de ‘onverzadigde’ hydraulische geleidbaarheid (permeabiliteit) en het waterhoudend vermogen (waterretentie). De bepaling of het begroten van de hydraulische bodemeigenschappen is niet alleen moeilijk en ook duur maar vraagt tevens veel uitvoeringstijd.

Wanneer nu die data en gegevens niet beschikbaar zijn dan kan worden gezocht naar alternatieve methodes zoals het opmaken van pedotranferfuncties (PTFs) die gebruik maken van fysische of empirische verbanden tussen de te bepalen hydraulische parameter en andere gemakkelijker te meten eigenschappen. Er bestaan reeds verschillende van die PTFs maar slechts enkele werden ontwikkeld voor bodems van ariede en semi-ariëde gebieden zoals Syrië. Er zijn in Syrië ook geen hydraulische bodemeigenschappen gepubliceerd en beschikbaar en daarenboven is het nationale databestand niet compleet. Daarom wordt in deze thesis aandacht besteed aan de relaties tussen hydraulische parameters en andere gemakkelijk te bepalen bodemchemische en bodemfysische karakteristieken. Zo werden enkele verbeteringen aangebracht bij het modelleren van de bodem-water-retentie curve (SWRC) of pF-curve. Het doel was een praktisch model te ontwikkelen dat, met een

minimum aantal parameters, de SWRC over het ‘verzadigd’ gedeelte kan bepalen. Daarvoor dienden uiteraard verbanden te worden gezocht tussen de hydraulische parameter en andere eenvoudig te bepalen fysische en chemische karakteristieken.

In hoofdstuk 2 werden tien ‘gesloten-vorm’ unimodale analytische vergelijkingen van de pF-curve geëvalueerd. De modellen vergeleken op hun nauwkeurigheid, lineariteit, Aikake Information Criterion (AIC), parameter eigenheid en identificatie. Om dit te verwezenlijken werden op 24 plaatsen in Vlaanderen (België) telkens 2 bodemstalen genomen.

In hoofdstuk 3 werd het Kosugi (1999) model gewijzigd om de pF curve tussen ‘bodemverzadiging’ en ‘ovendroogte’ te kunnen beschrijven. Evenwel behoudt het nieuwe model de vorm van de originele Kosugi functie in het ‘natte’ gedeelte en wordt getransformeerd naar een adsorptievergelijking in het ‘droge’ gedeelte. De graad van nauwkeurigheid van het voorspellen van het nieuwe model werd verder geëvalueerd door de data-set te verminderen door het niet-gebruiken van de data overeenkomstig een matrix potentiaal lager dan -100kPa.

In hoofdstuk 4 werd aangetoond dat het nieuwe model het best overeenkomt met de geobserveerde waarden. De nauwkeurigheid om de data te voorspellen was eveneens goed, hetgeen weerspiegeld wordt in de hoge correlatie tussen zijn parameters en de bodemkarakteristieken. Daarvoor werden de gegevens van de bodem-water-retentie van 137 bodems gehaald uit het UNSODA (Internationale Unsaturated Soil hydraulic DATA base).

In hoofdstuk 5 werd de invloed onderzocht van de ‘voorbehandeling’ van de bodemstalen op de hoeveelheid zand, leem en kleifraction en dit om de pF-curve van kalkrijke gronden te kunnen opmaken. Daarvoor werden 72 bodemmonsters verzameld uit verschillende velden in noordwest Syrië. Ofwel werd de klak niet vernietigd ofwel met zoutzuur vernietigd. Beide methoden toonden een grote variatie in zand, leem en klei aan. Enkel wanneer kalk niet werd verwijderd kon de pF-curve op basis van de textuur worden voorspeld.

In hoofdstuk 6 werd de interactie tussen de pF-curve en fysische, mechanische en chemische karakteristieken van ariëde en semi-ariëde bodems onderzocht. Er was een duidelijke invloed waar te nemen tussen de plasticiteitsgrens en eveneens ook het specifiek oppervlak op de bodem-water-retentie. Beide eigenschappen worden in een routine bodemonderzoek meestal niet bepaald. Een ‘Principal Component Analysis’ (PCA) geeft een overzicht van de bodemeigenschappen die gerelateerd zijn aan de waterretentie.

In hoofdstuk 7 werd de SWRC van ariede en semi-ariete voorspeld aan de hand van drie verschillende methodes namelijk de ‘support vector machines regression’ (SVR), de ‘artificial neural networks’ (ANN), en ‘multiple linear regression’ (MLR). De SVR methode voorspelde het best de SWRC weerspiegeld in RMSE en R^2 . Een statistische verbetering werd bekomen wanneer de ‘organische koolstof’ (OC) werd vervangen door de plasticiteitsgrens, gemakkelijker en goedkoper te bepalen. Met de introductie van de plasticiteitsgrens als nieuwe indicator of voorspeller wordt duidelijk het universele karakter van de PTF aangetoond. Verder onderzoek kan uitgevoerd worden om bijvoorbeeld het water-retentie-model in een hydraulische geleidbaarheidsfunctie zoals de Mualem vergelijking (Mualem, 1976) in te voeren.

References

- Ahuja, L.R., and D. Schwartzendruber. 1972. An improved form of the soil-water diffusivity function. *Soil Sci. Soc. Am. J.* 36:9-14.
- Akaike, H. 1974. New look at the statistical-model identification. *IEEE Trans. Autom. Control* AC19:667-73.
- Al-Khaier, F., 2003. Soil salinity detection using satellite remote sensing. MS. Dissertation, International Institute for Geo-information Science and Earth Observation and Utrecht University, Enschede.
- Al Majou, H., Bruand, A., Duval, O., 2008. The use of in situ volumetric water content at field capacity to improve the prediction of soil water retention properties. *Canadian Journal of Soil Science* 88, 533-541.
- Arya, L.M., and J.F. Paris. 1981. A physico-empirical approach to predict the soil water moisture characteristic from particle size distribution and bulk density data. *Soil Sci. Soc. Am. J.* 45:1023–1030.
- Assouline, A., D. Tessier, and A. Bruand. 1998. A conceptual model of the soil water retention curve. *Water Resour. Res.* 34:223-231.
- ASTM. 1998. ASTM D4318-10: Standard test method for liquid limit, plastic limit, and plasticity index of soils. ASTM Int., West Conshohocken, PA. doi:10.1520/D4318-10
- Baetens, J. M., K. Verbist, W. M. Cornelis, D. Gabriels, and G. Soto. 2009. On the influence of coarse fragments on soil water retention, *Water Resour. Res.*, 45, W07408, doi:10.1029/2008WR007402.
- Bastet, G., A. Bruand, M. Voltz, M. Bornand and P. Quélin. 1997. Performance of available pedotransfer functions for predicting the water retention properties of French soils, *Proc. Int. Workshop on the Characterization and Measurement of the Hydraulic Properties of Unsaturated Porous Media*. University of California, Riverside, CA.
- Batjes, N. H., 2002a. A homogenized soil profile data set for global and regional environmental research (WISE, Version 1.1) (available online at <http://www.isric.org>). Report 2002/01. Wageningen: International Soil Reference and Information Centre.
- Batjes, N. H., 2002b. Soil parameter estimates for the soil types of the world for use in global and regional modeling (Version 2.1; July 2002). *ISRIC Report 2002/02c* (Available online at <http://www.isric.org>). Wageningen: International Food Policy Research Institute (IFPRI) and International Soil Reference and Information Centre (ISRIC).
- Batjes, N.H., 2009. Harmonized soil profile data for applications at global and continental scales: updates to the WISE database. *Soil Use Manage.* 25, 124e127.

- Blanco-Canqui, H., R. Lal, W.M. Post, R.C. Izaurralde, and M.J. Shipitalo. 2006. Organic carbon influences on soil particle density and rheological properties. *Soil Sci. Soc. Am. J.* 70:1407–1414.
- Blunt, M.J., 2001. Flow in porous media - pore-network models and multiphase flow. *Curr. Opin. Colloid Interf. Sci.* 6:197-207.
- Bossuyt, H., J. Six, and P.F. Hendrix. 2002. Aggregate-protected carbon in no-tillage and conventional tillage agroecosystems using carbon-14 labeled plant residue. *Soil Sci. Soc. Am. J.* 66:1965–1973.
- Botula, Y.-D., W.M. Cornelis, G. Baert, and E. Van Ranst. 2012. Evaluation of pedotransfer functions for predicting water retention of soils in Lower Congo (D.R. Congo). *Agric. Water Manage.* 111:1-10.
- Botula, Y.-D., Nemes, A., Mafuka, P., Van Ranst, E., Cornelis, W.M., 2013. Prediction of water content of soils from the humid tropics by the non-parametric k-nearest neighbor approach. *Vadose Zone Journal* doi:10.2136/vzj2012.0123.
- Botula, Y.-D. 2013. Indirect methods to predict hydrophysical properties of soils of Lower Congo. PhD thesis, Ghent University.
- Bouma, J., 1989. Using soil survey data for quantitative land evaluation. *Advances in Soil Science* 9, 177-213.
- Bradley, R.S., 1936. Polymolecular adsorbed films: I. *J. Chem.Soc.* 139:1467-1474.
- Brady, NC 1990 *The nature and properties of soils* 10th ed Macmillan, New York.
- Brooks, R.H., and A.T. Corey. 1964. Hydraulic properties of porous media. *Hydrological Paper 3*, Civil Engineering Department., Colorado State University, Fort Collins, CO.
- Bruce, R.R., and A. Klute. 1956. The measurement of soil moisture diffusivity. *Soil Sci. Soc. Am. Proc.* 20:458-462.
- Brutsaert, W. 1966. Probability laws for pore-size distributions. *Soil Sci.* 101:85-92.
- Brutsaert, W. 1967. Some methods of calculating unsaturated permeability. *Trans. Am. Soc. Agr. Eng.* 10:400-404.
- Brutsaert, W., 1968. The permeability of a porous medium determined from certain probability laws for pore size distribution. *Water Resour. Res.* 4:425-434.
- Burdine, N.T., 1953. Relative permeability calculation from pore-size distribution data. *Trans. AIME* 198:71–78.
- Campbell, G.S., 1974. A simple method for determining unsaturated hydraulic conductivity from moisture retention data. *Soil Sci.* 117:311-314.
- Campbell, G.S., and S. Shiozawa. 1992. Prediction of hydraulic properties of soils using particle size distribution and bulk density data. In: *International workshop on indirect methods for estimating the hydraulic properties of unsaturated soils*. Univ. Calif. Press, Berkeley CA.
- Cattell, R.B., 1978. *The scientific use of factor analysis in behavioral and life science*. Plenum Press: New York.
- Chen, J., and H.S. Wheater. 1999. Identification and uncertainty analysis of soil water retention models using lysimeter data. *Water Resour. Res.* 35:2401-2414.
- Clapp, R.B., and G.M. Hornberger. 1978. Empirical equations for some soil hydraulic-properties. *Water Resour. Res.* 14:601-604

- Coppola, A., 2000. Unimodal and bimodal descriptions of hydraulic properties for aggregated soils. *Soil Sci. Soc. Am. J.* 64:1252–1262.
- Cornelis, W.M., J. Ronsyn, M. Van Meirvenne, and R. Hartmann. 2001. Evaluation of pedotransfer functions for predicting the soil moisture retention curve. *Soil Sci. Soc. Am. J.* 65:638–648.
- Cornelis, W.M., D. Gabriels, and R. Hartmann. 2004. A conceptual model to predict the deflation threshold shear velocity as affected by near-surface soil water: 1. Theory. *Soil Sci. Soc. Am. J.* 68:1154-1161.
- Cornelis, W. M., M. Khlosi, R. Hartmann, M. Van Meirvenne, and B. De Vos. 2005. Comparison of unimodal analytical expressions for the soil-water retention curve. *Soil Sci. Soc. Am. J.* 69: 1902–1911.
- Day, P.R. 1965. Particle fractionation and particle size analysis. In: Black, C.A., et al. (Ed.), *Methods of soil analysis, Part I: Agronomy*, 9, pp. 545-567.
- Dekker S.C., W. Bouten, F.C. Bosveld, 2001. On the information content of forest transpiration measurements for identifying canopy conductance model parameters, *Hydrological Processes* 15 , 2821-2832.
- De Neve, S., and G. Hofman. 2002. Quantifying soil water effects on nitrogen mineralization from soil organic matter and from fresh crop residues. *Biol. Fert. Soils* 35:379-386.
- De Pauw, E., Oberle A. and Zöbisch, M., 2000. An overview of land cover and land use in Syria. Research report Natural Resources Management Programme, ICARDA, Aleppo, Syria.
- De Visscher, A., and O. Van Cleemput. 2003. Simulation model for gas diffusion and methane oxidation in landfill cover soils. *Waste Management* 23:581-591.
- De Vos, B., M. Van Meirvenne, P. Quataert, J. Deckers, and B. Muys. 2005. Predictive quality of pedotransfer functions for estimating bulk density of forest soils. *Soil Sci. Soc. Am. J.* 69:500–510.
- Dexter, A.R., and N.R.A. Bird. 2001. Methods for predicting the optimum and the range of soil water contents for tillage based on the water retention curve. *Soil Tillage Res.* 57:203–212.
- Dirksen, C., and S. Dasberg. 1993. Improved calibration of time domain reflectometry soil water content measurements. *Soil Sci. Soc. Am. J.* 57: 660–667.
- Dochain D. and P.A.Vanrolleghem. 2001. *Dynamical Modelling and Estimation in Wastewater Treatment Processes*. IWA Publishing, London, UK. ISBN 1-900222-50-7. pp. 342.
- Duniway MC, Herrick JE, Monger HC. 2007. The high water-holding capacity of petrocalcic horizons. *Soil Science Society of America Journal* 71: 812–819.
- Endelman, F.J., G.E.P. Box, J.R. Boyle, R.R. Hughes, D.R. Keeney, M.L. Norhtrup, P.G. Saffigna. 1974. *The mathematical modeling of soil water-nitrogen phenomena*. Oak Ridge National Laboratory. EDFB-IBP-74-8.
- FAO. 1998. *World reference base for soil resources*. World Soil Resources Reports 84. FAO: Rome.
- Fares, F., 1991. *Fundamentals of soil science (in Arabic)*. Damascus University, Damascus, Syria, pp. 328-366.
- Fares, F., A.Albalkhi, J. Dec, , M. A. Bruns, and J. M. Bollag. 2005. Physicochemical characteristics of animal and municipal wastes decomposed in arid soils. *J. Environ. Qual.* 34 : 1392–1403.

- Fayer, M.J., and C.S. Simmons. 1995. Modified soil water retention functions for all matrix suctions. *Water Resour. Res.* 31:1233–1238.
- Fischer, U., and M.A. Celia. 1999. Prediction of relative and absolute permeabilities for gas and water from soil water retention curves using a pore-scale network model. *Water Resour. Res.* 35:1089-1100.
- Francis and Aguilar, 1995 R.E. Francis and R. Aguilar, Calcium-carbonate effects on soil textural class in semiarid wildland soils, *Arid Soil Research and Rehabilitation* 9 (1995), pp. 155–165.
- Fredlund, D. G., and A. Q. Xing. 1994. Equations for the soil-water characteristic curve, *Can. Geotech. J.*, 31(4), 521–532, doi:10.1139/t94-061.
- Gaiser T., F. Graef, and J. C. Cordeiro. 2000. Water retention characteristics of soils with contrasting clay mineral composition in semi-arid tropical regions. *Australian J. Soil Res.* 38: 523-36.
- Gardner, W.R. 1958. Some steady-state solutions of the unsaturated moisture flow equation with applications to evaporation from a water table. *Soil Sci.* 85:228-232.
- Gardner, W.H. 1968. Availability and measurement of soil water. p. 107– 135. In T.T. Kozlowski (ed.) *Water deficits and plant growth*. Vol. 1. Academic Press, New York.
- Gardner, W.R., D. Hillel, and Y. Benyamini. 1970. Post-irrigation movement of soil water. I. Redistribution. *Water Resour. Res.* 6:851-861.
- Gee, G.W., and J.W. Bauder. 1986. Particle-size analysis. p. 383–411. In A. Klute (ed.) *Methods of soil analysis*. Part 1. 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.
- Ghorbani Dashtaki, Sh., Homae, M., Khodaberdiloo, H., 2010. Derivation and validation of pedotransfer functions for estimating soil water retention curve using a variety of soil data. *Soil Use Manage.* 26, 68–74.
- Gregory PJ, Simmonds LP, Pilbeam CJ., 2000. Soil type, climatic regime, and the response of water use efficiency to crop management. *Agron J* 92:814-820.
- Godfray, C., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M., Toulmin, C., 2010. ‘Food security: the challenge of feeding 9 billion people’, *Science* 327, 812–818.
- Groenevelt P. H., and C. D. Grant. 2004. A new model for the soil water retention curve that solves the problem of residual water contents, *European Journal of Soil Science.* 55, 479-485.
- Gunn S., 1997. Support vector machines for classification and regression. ISIS technical report.
- Gupta, S.C., and W.E. Larson. 1979. Estimating soil water retention characteristics from particle size distribution, organic matter percent and bulk density. *Water Resour. Res.* 15:1633–1635.
- Gupta R.D., P.D. Sharma, C.L. Acharya, and B.R. Tripathi. 1983. Water retention characteristics of some soil profiles of north-west India in relation to soil properties under different bio- and climo-sequences. *J. Indian Soc. Soil Sci.* 31: 458–463.
- Haghverdi, A., Cornelis, W.M., Ghahraman, B., 2012. A pseudo-continuous neural network approach for developing water retention pedotransfer functions with limited data. *J. Hydrol.* 442, 46–54.

- Hall, D.G., Reeve, M.J., Thomasson, A.J. and Wright, V.F., 1977. Water retention, porosity and density of field soils. Soil Survey of England and Wales Harpenden Technical Monograph No. 9, 75 pp..
- Harper, W.G. 1957. Morphology and genesis of Calcisols. Soil Sci. Soc. Am. Proc. 21:420–424.
- Hartley, H. 1950. The maximum F-ratio as a short cut test for heterogeneity of variance. *Biometrika*, 37: 308-312.
- Haverkamp, R., Zammit, C., Bouraoui, F., Rajkai, K., Arrue, J.L., Heckman, N. 1997. GRIZZLY, Grenoble Soil Catalogue. Soil survey of field data and description of particle size, soil water retention and hydraulic conductivity functions. Laboratoire d'Étude des Transfers en Hydrologie et Environnement, LTHE, UMR5564, CNRS, INPG, ORSTOM, UJF, BP 53, 38041 Grenoble Cédex 09, France.
- Hillel, D. 1998. Environmental Soil Physics. Academic Press. San Diego, CA.
- Hazzouri, A., and Khlosi M., 1998. Effect of saline water on salt accumulation and the yield of different chickpea cultivars in the north of Syria. Pages 216-223 in Challenges Facing the Management Profession in the Middle East and European Countries to Manage Agricultural Resources Effectively: Proceedings, Bradford Management and Technology Centre Ltd/Ventures and Consultancy Bradford Ltd/University of Aleppo, 14-16 July 1998, Harrogate, United Kingdom. Bradford Management and Technology Centre, Bradford, United Kingdom.
- Held, R.J., and M.A. Celia. 2001a. Modeling support of functional relationships between capillary pressure, saturation, interfacial area and common lines. *Adv. Water Resour.* 24:325-343.
- Held, R.J., and M.A. Celia. 2001b. Pore-scale modeling extension of constitutive relationships in the range of residual saturations. *Water Resour. Res.* 37:165-170.
- Holtan, H.N., England, C.B., Lawless, G.P., Schumaker, G.A., 1968. Moisture-ension data for selected soils on experimental watersheds. Report ARS 41-144, Agricultural Research Services, Beltsville, Md..
- Hsiao, T.C., Heng, L., Steduto, P., Rojas-Lara, B., Raes, D., Fereres, E., 2009. AquaCrop – the FAO crop model to simulate yield response of water: III. Parameterization and testing for Maize. *Agronomy Journal* 101 (3), 448–459.
- Hutson, J.L., and A. Cass. 1987. A retentivity function for use in soil water simulation models. *J. Soil Sci.* 38:105–113.
- IAEA, Management of Crop Residues for Sustainable Crop Production, IAEA-TECDOC-1354, IAEA, Vienna (2003).
- Jamison and Kroth, 1958. V.C. Jamison and E.M. Kroth, Available moisture storage capacity in relation to texture composition and organic matter content of several Missouri soils. *Soil Sci. Soc. Am. Proc.* 22 (1958), pp. 189–192.
- Joliffe, I.T. 2002. Principal component analysis. 2nd ed. Springer: New York.
- Jordán M.M., J. Navarro-Pedreño, E. García-Sánchez, J. Mateu and P. Juan, 2003. Spatial dynamics of soil salinity under arid and semi-arid conditions: geological and environmental implications, *Environ. Geol.* 45 (4), pp. 448–456.
- Jury, W.A., Horton, R., 2004. Soil Physics, sixth ed. John Wiley & Sons, Inc., pp. 78–83.

- Jury, W.A., W.R. Gardner, and W.H. Gardner. 1991. *Soil physics*. 5th ed. John Wiley & Sons, New York.
- Kaiser, H.F., and J. Rice. 1974. Little jiffy, mark IV. *Educ. Psychol. Measurement*, 34 : 111–117.
- Keller, T., Dexter, A.R., 2012. Plastic limits of agricultural soils as functions of soil texture and organic matter content. *Soil Research* 50, 7–17.
- Kern, J.S., 1995. Evaluation of soil water retention models based on basic soil physical properties. *Soil Sci. Soc. Am. J.* 59, pp. 1134–1141
- Khlosi, M., W.M. Cornelis, D. Gabriels, and G. Sin. 2006. Simple modification to describe the soil water retention curve between saturation and oven dryness. *Water Resour. Res.* 42:W11501, doi:10.1029/2005WR004699.
- Khlosi, M., W.M. Cornelis, and D. Gabriels. 2008. Analyzing the effects of particle-size distribution changes associated with carbonates on the predicted soil-water retention curve. In D: Gabriels et al. (2008). *Combating desertification: assessment, adaptation and mitigation strategies*. International Centre for Eremology – Belgian Development Cooperation. P. 100-105.
- Khlosi, M., W.M. Cornelis, A. Douaik, A. Hazzouri, H. Habib, D. Gabriels, 2013. Exploration of the interaction between hydraulic and physicochemical properties of Syrian soils *Vadose Zone J.* 2013. 12:–. doi:10.2136/vzj2012.0209.
- Khodaverdiloo, H. and M. Homae. 2004. Pedotransfer Functions of some Calcareous Soils. In: Nicole Whrle and Maik Scheurer (Eds.). *EUROSOIL 2004*. 10: 27(1-11). September 4-12. Freiburg Germany.
- Kosugi, K. 1994. Three-parameter lognormal distribution model for soil water retention, *Water Resour. Res.* 30:891-901.
- Kosugi, K. 1996. Lognormal distribution model for unsaturated soil hydraulic properties. *Water Resour. Res.* 32:2697–2703.
- Kosugi, K. 1997. A new model to analyze water retention characteristics of forest soils based on soil pore radius distribution. *J. For. Res.* 2:1-8.
- Kosugi, K. 1999. General model for unsaturated hydraulic conductivity for soils with log-normal pore-size distribution, *Soil Sci. Soc. Am. J.* 63, 270–277.
- Kovačević M., Bajat B., and Gajić B., 2010. Soil type classification and estimation of soil properties using support vector machines. *Geoderma*, 154, 340-347.
- Kravchenko, A., Zhang, R., 1998. Estimating the soil water retention from particle-size distributions: A fractal approach. *Soil Science* 163, 171-179.
- Lal, R., and M.K. Shukla. 2004. *Principles of soil physics*. Dekker, New York.
- Laliberte, G.E., 1969. A mathematical function for describing capillary pressure-desaturation data. *Bull. Int. Assoc. Sci. Hydrol.* 142:131-149.
- Lambin, E.F. and P. Meyfroidt. 2011. Global land use change, economic globalization, and the looming land scarcity. *PNAS*.
- Lamorski, K., Ya.A. Pachepsky, C. Slawinski, and R.T. Walczak. 2008. Using support vector machines to develop pedotransfer functions for water retention of soils in Poland. *Soil Sci. Soc. Am. J.* 72:1243–1247.
- Lamorski, K., T. Pastuszka, J. Krzyszczak, C. Slawinski, and R.T. Walczak. 2013. Soil Water Dynamic Modeling Using the Physical and Support Vector Machine Methods *Vadose Zone J.* 2013. 12:–. doi:10.2136/vzj2013.05.0085.

- Lebeau, M., and J.-M. Konrad, 2010. A new capillary and thin film flow model for predicting the hydraulic conductivity of unsaturated porous media, *Water Resour. Res.*, 46, W12554, doi:10.1029/2010WR009092.
- Leij, F. J., Alves, W.J., van Genuchten, M.Th., Williams, J.R., 1996. *Unsaturated Soil Hydraulic Database, UNSODA 1.0 User's Manual*. Report EPA/600/R-96/095, U.S. Environmental Protection Agency, Ada, Oklahoma.
- Leij, F.J., Romano, N., Palladino, M., Schaap, M.G., Coppola, A., 2004. Topographical attributes to predict soil hydraulic properties along a hillslope transect. *Water Resources Research* 40, article W02407.
- Lilly, A., and H. Lin. 2004. Using soil morphological attributes and soil structure in pedotransfer functions. p. 115–141. In Ya. Pachepsky and W. Rawls (ed.) *Development of pedotransfer functions in soil hydrology*. Dev. Soil Sci. 30. Elsevier, Amsterdam.
- Lu, S., T. Reng, Y. Gong, and R. Horton. 2008. Evaluation of three models that describe soil water retention curves from saturation to oven dryness, *Soil Sci. Soc. Am. J.*, 72, 1542–1546, doi:10.2136/sssaj2007.0307N.
- Massri Z., M. Zöbisch, A. Bruggeman, P. Hayek, M. Kardous, 2002. Wind Erosion in Marginal Mediterranean Dryland Areas, Khanasser Valley- a Case Study. In : 2002, *Proceedings of ICAR5/GCTE-SEN Joint Conference*, International Center for Arid and Semiarid Lands Studies, eds Lee, Jeffrey A. and Zobeck, Ted M., Texas Tech University, Lubbock, Texas, USA Publication 02-2 p. 174.
- Massoud, F. 1972. Some physical properties of highly calcareous soils and their related management practices. *FAO/UNDP Regional Seminar on Reclamation and Management of Calcareous Soils*. Cairo, Egypt. November 27-December 2, 1972
- Mathworks. 2007. *MATLAB Release 14*. Mathworks, Natick, MA.
- Matthews, G. P., G. M. Laudone, A. S. Gregory, N. R. A. Bird, A. G. D. G. Matthews, and W. R. Whalley. 2010. Measurement and simulation of the effect of compaction on the pore structure and saturated hydraulic conductivity of grassland and arable soil, *Water Resour. Res.*, 46, W05501, doi:10.1029/2009WR007720.
- Mayr, T., Jarvis, N.J., 1999. Pedotransfer functions to estimate soil water retention parameters for a modified Brooks-Corey type model. *Geoderma*. 91, 1-9.
- McBride, R.A., and E.E. Mackintosh. 1984. Soil survey interpretations from water retention data: I. Development and validation of a water retention model. *Soil Sci. Soc. Am. J.* 48: 1338–1343.
- McCauley, A., C. Jones, and J. Jacobsen. 2005. *Basic Soil Properties*. Montana State University Extension Service, 1-12.
- McKenzie, N. J., Jacquier, D. W., and Gregory, L. J., 2008. Online soil information systems – recent Australian experience. In Hartemink, A. E., McBratney, A. B., de Lourdes Mendonça- Santos, M. (eds.), *Digital soil mapping with limited data*. Berlin: Springer Science+Business Media B.V. ISBN: 978-1-4020-8591-8. pp. 283–290.
- Merdun H., Cinar O., Meral R., Apan M., 2006. Comparison of artificial neural network and regression pedotransfer functions for prediction of soil water retention and saturated hydraulic conductivity. *Soil Till. Res.* 90, 108-116.

- McBratney, A.B., Minasny, B., Cattle, S.R., Vervoort, R.W., 2002. From pedotransfer function to soil inference system. *Geoderma* 109, 41–73.
- Minasny, B., A.B. McBratney, and K.L. Bristow. 1999. Comparison of different approaches to the development of pedotransfer functions for water-retention curves. *Geoderma* 93:225–253.
- Minasny, B., Hartemink, A.E., 2011. Predicting soil properties in the tropics. *Earth-Science Reviews* 106, 52–62.
- Misopolinos, N.D. and Kalovoulos, J.M. 1984. Determination of CEC and exchangeable Ca and Mg in non-saline calcareous soils. *J. Soil Sci.* 35: 93–98.
- Möller, A., Muller, H.W., Abdullah, A., Abdelgawad, G. and Utermann, J., 2005. Urban soil pollution in Damascus, Syria: concentrations and patterns of heavy metals in the soils of the Damascus Ghouta. *Geoderma*, 124(1-2): 63-71.
- More, J.J., B.S. Garbow, and K.E. Hillstrom. 1980. User's Guide to Minpack I. Argonne National Laboratory publication ANL-80-74.
- Morel-Seytoux, H. J., and J. R. Nimmo. 1999. Soil water retention and maximum capillary drive from saturation to oven dryness, *Water Resour. Res.* 35, 2031-2041.
- Mualem, Y., 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resour. Res.* 12:512-522.
- Mualem, Y. 1986. Hydraulic conductivity of unsaturated soils: Prediction and formulas. p. 799-823. In A. Klute (ed.) *Methods of Soil Analysis. Part 1. Physical and Mineralogical Methods.* 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.
- Müller, K.-R., Mika, S., Rätsch, G., Tsuda, K., and Schölkopf, B. 2001. An introduction to kernel-based learning algorithms. *IEEE Neural Networks*, 12(2):181–201.
- Muller, K.E., and B.A. Fetterman. 2002. *Regression and ANOVA: An integrated approach using SAS software.* SAS Institute, Cary, NC.
- Nasta, P., Kamai, T., Giovanni B., Chirico, Hopmans, J., Romano, R. 2009. Scaling soil water retention functions using particle-size. *Journal of Hydrology* 374: 223-234.
- Nasta, P., J. A. Vrugt, and N. Romano. 2013. Prediction of the saturated hydraulic conductivity from Brooks and Corey's water retention parameters, *Water Resour. Res.*, 49, 2918–2925, doi:10.1002/wrcr.20269.
- Nelson, R.E. 1982. Carbonate and gypsum. p. 181–197 In A.L. Page et al. (ed.) *Methods of soil analysis. Part 2.* 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.
- Nemes, A., Schaap, M.G., Leij, F.J., 1999. The UNSODA Unsaturated Soil Hydraulic Database Version 2.0. US Salinity laboratory, Riverside, CA.
- Nemes, A., M.G. Schaap, F.J. Leij, and J.H.M. Wösten. 2001. Description of the unsaturated soil hydraulic database UNSODA version 2.0. *J. Hydrol. (Amsterdam)* 251:151–162.
- Nemes, A., 2002. Unsaturated soil hydraulic database of Hungary: HUNSODA. *Agrokémia és Talajtan*, 51(1–2), 17–26.
- Nemes, A., M.G. Schaap, and J.H.M. Wösten. 2003. Functional evaluation of pedotransfer functions derived from different scales of data collection. *Soil Sci. Soc. Am. J.* 67:1093–1102.

- Nemes A., D. J. Timlin, Ya. A. Pachepsky, and W. J. Rawls, 2009. Evaluation of the Rawls et al. (1982) Pedotransfer Functions for their Applicability at the U.S. National Scale Soil Sci Soc Am J 2009 73: 1638-1645.
- Neter, J., M.H. Kutner, C.J. Nachtsheim, and W. Wasserman. 1996. Applied linear statistical methods. Irwin, Chicago, IL.
- Nielsen, D.R., and R.H. Shaw. 1958. Estimation of the 15-atmosphere moisture percentage from hydrometer data. Soil Sci. 86: 103–105.
- Nitao, J.J., and J. Bear. 1996. Potentials and their role in transport in porous media. Water Resour. Res. 32:225-250.
- Odell, R.T., T.H. Thornburn, and L.J. McKenzie. 1960. Relationships of Atterberg limits to some other properties of Illinois soils. Soil Sci. Soc. Am. Proc. 24:297–300.
- Omlin, M. and Reichert, P. 1999. A comparison of techniques for the estimation of model prediction uncertainty, Ecological Modelling 115, 45-59.
- Or, D., and M. Tuller. 1999. Liquid retention and interfacial area in variably saturated porous media: Upscaling from single-pore to sample-scale model. Water Resour. Res. 35:3591-3605.
- Or, D., and M. Tuller. 2002. Cavitation during desaturation of porous media under tension. Water Resour. Res. 38: Art. No. 1061
- Ouessar, M., A. Bruggeman, F. Abdelli, R.H. Mohtar, D. Gabriels, and W.M. Cornelis. 2009. Modelling water-harvesting systems in the arid south of Tunisia using SWAT. Hydrology and Earth System Sciences 13:2003-2021.
- Oweis, Th., 1997. Supplemental Irrigation: A Highly Efficient Water-Use Practice. ICARDA, Aleppo, Syria.
- Oweis, T., A. Hachum and M. Pala. 2004. Lentil production under supplemental irrigation in a Mediterranean environment. Agric. Water Manage., 68: 251-265.
- Pachepsky, Y.A., D. Timlin, and G. Varallyay. 1996. Artificial neural networks to estimate soil water retention from easily measurable data. Soil Sci. Soc. Am. J. 60:727–733.
- Pachepsky, Y., Rawls, W.J. & Giménez, D. 2001. Comparison of soil water retention at field and laboratory scales. Soil Science Society of America Journal, 65, 460–462.
- Pachepsky, Ya.A. and W.J. Rawls. 2003. Soil structure and pedotransfer functions. Eur. J. Soil Sci. 54:443–452.
- Peters, A. 2013, Simple consistent models for water retention and hydraulic conductivity in the complete moisture range, Water Resour. Res., 49, 6765–6780, doi:10.1002/wrcr.20548.
- Phuong N. M., Le Khoa Van, Cornelis Wim M. 2014. Using categorical soil structure information to improve soil water retention estimates of tropical delta soils. Soil Research 52, 443–452.
- Polemio, M. and J.D. Rhoades. 1977. Determining cation exchange capacity: A new procedure for calcareous and gypsiferous soils. Soil Sci. Soc. Am. J. 41: 524–527.
- Press, W. H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. 1992. Numerical recipes in C. Cambridge University Press, Cambridge.
- Pretty, J., Sutherland, W.J., Ashby, J., Auburn, J., Baulcombe, D., Bell, M., Bentley, J., Bickersteth, S., Brown, K., Burke, J., Campbell, H., Chen, K., Crowley, E., Crute, I., Dobbelaere, D., Edwards-Jones, G., Funes-Monzote, F., Godfray, H.C.J., Griffon, M.,

- Gypmantisiri, P., Haddad, L., Halavatau, S., Herren, H., Holderness, M., Izac, A.-M., Jones, M., Koochafkan, P., Lal, R., Lang, T., McNeely, J., Mueller, A., Nisbett, N., Noble, A., Pingali, P., Pinto, Y., Rabbinge, R., Ravindranath, N.H., Rola, A., Roling, N., Sage, C., Settle, W., Sha, J.M., Shiming, L., Simons, T., Smith, P., Strzepeck, K., Swaine, H., Terry, E., Tomich, T.P., Toulmin, C., Trigo, E., Twomlow, S., Vis, J.K., Wilson, J., Pilgrim, S., 2010. The top 100 questions of importance to the future of global agriculture. *International Journal of Agricultural Sustainability* 8, 219–236.
- Ragab R. and Prudhomme C. 2002 SW-Soil and Water: Climate Change and Water Resources Management in Arid and Semi-arid Regions: Prospective and Challenges for the 21st Century. *Biosystems Engineering*, Vol. 81, Issue 1, pp. 3-34.
- Rawls, W.J., and D.L. Brakensiek. 1982. Estimating soil water retention from soil properties. *J. Irrig. Drainage Div. ASCE* 108:166–171.
- Rawls, W.J., Brakensiek, D.L., Saxton, K.E., 1982. Estimation of soil-water properties. *Transactions of the ASAE* 25, 1316-1320.
- Rawls, W.J., and D.L. Brakensiek. 1985. Prediction of soil water properties for hydrologic Modeling. *Proceedings of Symposium on Watershed Management, ASCE*, pp: 293-299.
- Rawls, W.J. and Pachepsky, Y.A. 2002. Soil consistence and structure as predictors of water retention. *Soil Science Society of America Journal*, 66, 1115-1126.
- Reeves, P.C., and M.A. Celia. 1996. A functional relationship between capillary pressure, saturation, and interfacial area as revealed by a pore-scale network model. *Water Resour. Res.* 32:2345-2358.
- Romano, N., Santini, A., 1997. Effectiveness of using pedotransfer functions to quantify the spatial variability of soil water retention characteristics. *J. Hydrol.* 202, 137–157.
- Ross, P.J., J. Williams, and K.L. Bristow. 1991. Equation for extending water-retention curves to dryness. *Soil Sci. Soc. Am. J.* 55:923-927.
- Rossi, C., Nimmo, J.R., 1994. Modeling of soil water retention from saturation to oven dryness. *Water Resour. Res.* 30:701–708.
- Rost, S., D. Gerten, A. Bondeau, W. Lucht, J. Rohwer, and S. Schaphoff., 2008. Agricultural green and blue water consumption and its influence on the global water system, *Water Resour. Res.*, 44, W09405, doi:10.1029/2007WR006331
- Russo, D., 1988. Determining soil hydraulic properties by parameter estimation: On the selection of a model for the hydraulic properties. *Water Resour. Res.* 24:453-459.
- Saltelli, A., M. Ratto, S. Tarantola and F. Campolongo. 2005. Sensitivity Analysis for Chemical Models, *Chemical Reviews*, 105, 2811 - 2828.
- Saxton, K. E. 2005. Soil water characteristics hydraulic properties calculator, online available at: <http://hydrolab.arsusda.gov/soilwater/Index.htm>.
- Saxton, K.E., and W.J. Rawls. 2006. Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Sci. Soc. Am. J.* 70:1569–1578.
- Schaap, M.G., Leij, F.J., 1998. Database related accuracy and uncertainty of pedotransfer functions. *Soil Sci.* 163, 765–779.
- Schaap, M.G., F.J. Leij, and M.Th. van Genuchten. 2001. ROSETTA: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *J. Hydrol.* 251:163–176.

- Schaap, M.G., 2004. Accuracy and uncertainty in PTF predictions. In: Pachepsky, Y.A., Rawls, W.J. (Eds.), *Development of Pedotransfer functions in soil hydrology*. Elsevier, Amsterdam, pp. 33–43.
- Scheinost, A.C., W. Sinowski, and K. Auerswald. 1997. Regionalization of soil water retention curves in a highly variable soilscape: I. Developing a new pedotransfer function. *Geoderma* 78:129–143.
- Schiettecatte, W., D. Gabriels, W.M. Cornelis, and G. Hofman. 2008. Enrichment of organic carbon in sediment transport by interrill and rill erosion processes. *Soil Sci. Soc. Am. J.* 72: 50–55.
- Schneider, A., T. Baumgartl, D. Doley, and D. Mulligan. 2010. Evaluation of the heterogeneity of constructed landforms for rehabilitation using lysimeters. *Vadose Zone Journal* 9: 898–909.
- Schofield, R.K. 1935. The pF of the water in soil. p. 38-48. In *Trans. Int. Congr. Soil Sci.* 3rd, II.
- Shao, Y., 2000. *Physics and modelling of wind erosion*. Atmospheric and oceanographic sciences library, 23, Kluwer Academic Publishers, Dordrecht.
- Sharma, S.K., Mohanty, B.P., Zhu, J., 2006. Including topography and vegetation attributes for developing pedo transfer functions in southern great plains of USA. *Soil Sci. Soc. Am. J.* 70, 1430–1440.
- Šimůnek, J., van Genuchten, M.T., Šejna, M., 2006. The HYDRUS software package for simulating two- and three-dimensional movement of water, heat, and multiple solutes in variably-saturated media. *Technical Manual, Version 1.0*, PC Progress, Prague, Czech Republic, pp. 241.
- Šimůnek, J., van Genuchten, M. Th., Šejna, M., 2008. Development and applications of the HYDRUS and STANMOD software packages, and related codes. *Vadose Zone Journal* 7(2), 587-600.
- Shinjo H., H. Fujita, G. Gintzbuger, T. Kosaki. 2000. Impact of grazing and tillage on water erosion in northeastern Syria. *Soil Science and Plant Nutrition*, vol.46: p.151-162.
- Smola, A. J. 1998. *Learning with Kernels*. Ph.D. thesis, Technische Universität Berlin. GMD Research Series No. 25.
- Soil Survey Staff, 1999. *Soil Taxonomy. Handbook*, vol. 436. U.S. Department of Agriculture, pp. 869.
- Soil Survey Staff. 2003. *Keys to soil taxonomy*. 9th ed. USDA, Washington, DC.
- SPSS. 2006. *SPSS statistical software*. Release 15.0. SPSS: Chicago, IL.
- Steduto, P., Hsiao, T.C., Raes, D., Fereres, E., 2009. AquaCrop—The FAO crop model to simulate yield response to water. I. Concepts and underlying principles. *Agronomy Journal* 101, 426-437.
- Steel, R., and J. Torrie. 1980. *Principles and procedures of statistics: a biometrical approach*. 2nd Ed. McGraw Hill, 633 pp.
- Su, C., and R.H. Brooks. 1975. Soil hydraulic properties from infiltration tests. p. 516-542. In *Watershed Management Proceedings, Irrigation and Drainage Division*, Am. Soc. Civ. Eng., Logan, UT.
- Suzuki, M. 1984. The properties of a base-flow recession on small mountainous watersheds, I, Numerical analysis using the saturated-unsaturated flow model. *J. Jpn. For. Soc.* 66:174-182.
- Tani, M. 1982. The properties of water-table rise produced by a one-dimensional, vertical, unsaturated flow. *J. Jpn. For. Soc.* 64:409-418.

- Tani, M. 1985. Analysis of one-dimensional, vertical, unsaturated flow in consideration of runoff properties of a mountainous watershed. *J. Jpn. For. Soc.* 67:449-460.
- Tavernier, R., A. Osman, and M. Ilaoui. 1981. Soil Taxonomy and the soil map of Syria and Lebanon. p. 83–93. In *Proc. 3rd. Int. Soil Class. Workshop of the Soil Manag. Support Services, Damascus, Syria.* ACSAD, Damascus.
- Tempel, P., Batjes, N.H., van Engelen, V.W.P., 1996. IGBP-DIS soil data set for pedotransfer function development. *Work. Pap. 96/05.* ISRIC, Wageningen, the Netherlands.
- THF. The Syrian Syndicate of Agricultural Engineers. 1994. The annual agricultural report. Damascus, 76p. (in Arabic).
- Thien, S.J. 1979. A flow diagram for teaching texture-by-feel analysis. *J. Agr. Edu.* 8:54-55.
- Tietje, O., Tapkenhinrichs, M., 1993. Evaluation of pedo-transfer functions. *Soil Science Society of America Journal* 57, 1088-1095.
- Tisdale, S.L., W.L. Nelson, J.D. Beaton, and J.L. Havlin, 1993. *Soil fertility and fertilizers.* 5th ed. MacMillan Publishing Co., New York.
- Tomasella, J., Hodnett, M.G., 1998. Estimating soil water retention characteristics from limited data in Brazilian Amazonia. *Soil Science* 163, 190-202.
- Tomasella, J., Hodnett, M. G., and Rossato, L., 2000. Pedotransfer functions for the estimation of soil water retention in Brazilian soils. *Soil Science Society of America Journal*, 64, 327–338.
- Tuller, M., and D. Or. 1999. Adsorption and capillary condensation in porous media: Liquid retention and interfacial configurations in angular pores. *Water Resour. Res.* 35:1949-1964.
- Tuller, M., and D. Or. 2001. Hydraulic conductivity of variably saturated porous media: Film and corner flow in angular pore space. *Water Resour. Res.* 37:1257-1276.
- Tuller, M., and D. Or. 2002. Unsaturated hydraulic conductivity of structured porous media: A review of liquid configuration-based models, *Vadose Zone J.*, 1, 14–37, doi:10.2113/1.1.14.
- Twarakavi, N.K.C., J. Šimůnek and M.G. Schaap. 2009. Development of pedotransfer functions for estimation of soil hydraulic parameters using support vector machines. *Soil Sci. Soc. Am. J.* 73:1443–1452.
- UN, 2010. *World Population Prospects: The 2010 Revision.* Department of Economic and Social Affairs. Population Division, New York.
- van den Berg, M., E. Klant, L.P. van Reeuwijk, and G.Sombroek. 1997. Pedotransfer functions for the estimation of moisture retention characteristics of Ferralsols and related soils. *Geoderma*, 78: 161–180.
- 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., and D.R. Nielsen. 1985. On describing and predicting the hydraulic properties of unsaturated soils. *Ann. Geophys.* 3:615-628.
- van Genuchten, M. Th., F. J. Leij, and S. R. Yates. 1991. *The RETC Code for Quantifying the Hydraulic Functions of Unsaturated Soils, Version 1.0.* EPA Report 600/2-91/065, U.S. Salinity Laboratory, USDA, ARS, Riverside, CA.
- van Wesemael, B., J. Poesen, C.S. Kosmas, N.G. Danalatos, and J. Nachtergaele. 1996. Evaporation from cultivated soils containing rock fragments. *J. Hydrol.* 182:65–82.
- Vapnik, V.N. 1995. *The nature of statistical learning theory.* John John Wiley & Sons, New York.

- Vapnik, V., 1998. *Statistical Learning Theory*. Wiley, New York.
- Vapnik, V., 2000. *The nature of statistical learning theory*. Springer, New York.
- Varallayay, G., and E.V. Mironenko. 1979. Soil-water relationships in saline and alkali conditions. *Agrokem. Talajtan* 28 (Suppl.): 33-82.
- Verbist, K., Cornelis, W. M., Pierreux, S., McLaren, R., Gabriels, D., 2012. Parameterizing a coupled surface-subsurface 3D soil hydrological model to evaluate the efficiency of a runoff water harvesting technique. *Vadose Zone J.* Sub judice.
- Verbist, K.M.J., Cornelis, W.M., Torfs, S., Gabriels, D., 2013. Comparing methods to determine hydraulic conductivities on stony soils. *Soil Science Society of America Journal* 77, 25–42.
- Vereecken, H., J. Maes, J. Feyen, and P. Darius. 1989. Estimating the soil moisture retention characteristic from texture, bulk density and carbon content. *Soil Sci.* 148:389-403.
- Vereecken H., and M. Herbst. 2004. Statistical regression. In: Pachepsky Y.A., and W.J. Rawls (Eds). *Development of Pedotransfer Functions in Soil Hydrology*. *Developments in Soil Sci.*, 30: 3-18.
- Visser, W.C. 1966. Progress in the knowledge about the effect of soil moisture content in plant production. *Tech. Bulletin* 45. Inst. Land Water Management, Wageningen, Netherlands.
- Walkley, A., and I.A. Black. 1934. An examination of the Degtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. *Soil Sci.* 37:29-38.
- Wang, L. 2005. *Support Vector Machines: theory and applications*, volume 177. Springer, Berlin, Germany.
- Wang B, Chen J, Li X, Chen L, Zhu M, Yu H, Kühne R, Schüürmann G., 2009. Estimation of soil organic carbon normalized sorption coefficient (Koc) using least squares-support vector machine. *QSAR Comb. Sci.* 28(5): 561-567.
- Webb, S. W., 2000 A simple extension of two-phase characteristic curves to include the dry region, *Water Resour. Res.* 36, 1425–1430.
- Weerts, A.H., G. Lian, and D. Martin. 2003. Modeling rehydration of porous biomaterials: Anisotropy effects. *J. Food Science* 68:937-942.
- Weiss, R., J. Alm, R. Laiho, and J. Laine. 1998. Modeling moisture retention in peat soils. *Soil Sci. Soc. Am. J.*, 62 : 305–313.
- Weynants, M., H. Vereecken, and M. Javaux. 2009. Revisiting Vereecken pedotransfer functions: Introducing a closed form hydraulic model. *Vadose Zone J.* 8:86–95.
- Weynants, M., Montanarella, L., Tóth, G., Strauss, P., Feichtinger, F., Cornelis, W., Javaux, M., Matula, S., Daroussin, J., Hennings, V., Schindler, U., Bilas, G., Máki, A., Tóth, B., Romano, N., Iovino, M., Morari, F., Kværnø, S., Nemes, A., Børresen, T., Haugen, L.E., Slawinski, C., Lamorski, K., Gonçalves, M., Paryka, N.V., Shein, E., Houšková, B., Anaya-Romero, M., Kätterer, T., Wösten, H., Hannam, J., Keay, C., Lilly, A., Laktionova, T., 2013. *European HYdropedological Data Inventory (EU-HYDI)*. JRC Technical Reports, Report EUR 26053. Institute for Environment and Sustainability, Joint Research Centre, European Commission, Ispra, Italy, 167 pp.
- White, N.F., H.R. Duke, D.K. Sunada, and A.T. Corey. 1970. Physics of desaturation in porous materials. *J. Irr. Div., Am. Soc. Civ. Eng. Proc.* JR-2:165-191.
- Wilson, M.G., M. C. Sasal, and O. P. Caviglia. 2013. Critical bulk density for a Mollisol and a Vertisol using least limiting water range: Effect on early wheat growth. *Geoderma* 192 354–361.

- Wösten, J.H.M., Schuren, C.H.J.E., Bouma, J., Stein, A., 1990. Functional sensitivity analysis of four methods to generate soil hydraulic functions. *Soil Science Society of America Journal* 54, 832-836.
- Wösten, J.H.M., A. Lilly, A. Nemes, and C. Le Bas. 1999. Development and use of a database of hydraulic properties of European soils. *Geoderma* 90:169–185.
- Wösten, J.H.M., Pachepsky, Y.A., Rawls, W.J., 2001. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *J. Hydrol.* 251, 123–150.
- Wösten, J. H. M., Veerman, G. J., de Groot, W. J. M., and Stolte, J., 2001. Water retention and hydraulic conductivity characteristics of top and subsoils in the Netherlands: the Staring Series. Updated version 2001. Alterra report 153. Wageningen, The Netherlands. (in Dutch).
- Wu, W., Wang, X., Xie, D., and Liu, H., 2008. Soil Water Content Forecasting by Support Vector Machine in Purple Hilly Region, *Computer and Computing Technologies in Agriculture*, 1, 223–230.
- Zacharias, S., and G. Wessolek. 2007. Excluding organic matter content from pedotransfer predictors of soil water retention. *Soil Sci. Soc. Am. J.* 71:43–50
- Zhang, Z. F. 2011. Soil water retention and relative permeability for conditions from oven-dry to full saturation, *Vadose Zone J.*, 10(4), 1299–1308, doi:10.2136/vzj2011.0019.
- Zhao, M., Running, S.W., 2010. Drought-induced reduction in global terrestrial net primary production from 2000 through 2009. *Science* 329, 940–943

Curriculum vitae

Personal data

First name: Muhammed

Family name: Khlosi

Birthplace: Aleppo

Birthdate: 15 June 1971

Country of origin: Syria

E-mail: Muhammed.Khlosi@UGent.be

Academic Degrees

- M.Sc. in Physical Land Resources, 2003, Ghent University, Belgium
- M.Sc. in Agricultural Engineering, 1997, Aleppo University, Syria
- B.Sc. in Soil and Land Reclamation, 1996, Aleppo University, Syria

Positions

- 1998-2000: Research Assistant, Department Soil and Land Reclamation, Damascus University, Syria
- 1996-1998: Teaching Assistant, Department Soil and Land Reclamation, Aleppo University, Syria
- 2012-2013: Assistant, Department of Soil Management, Ghent University, Belgium

Awards and Honors

- 1996 Al-Bassel Award For Academic Excellence, Ministry of Higher Education, Damascus, Syria
- 2010 Best Student Paper Prize, A new model for describing the soil-water retention curve between saturation and oven dryness, National Union of Syrian Students in Belgium, the Netherlands, and Luxembourg, Brussels, Belgium

Scientific accomplishments

Articles in international journals with peer review, in ISI Web of Science (A1)

1. Cornelis, W.M., **M. Khlosi**, R. Hartmann, M. Van Meirvenne, and B. De Vos (2005). Comparison of unimodal analytical expressions for the soil-water retention curve. *Soil Sci. Soc. Am. J.*, 69:1902-1911.
2. **Khlosi, M.**, W.M. Cornelis, D. Gabriels, and G. Sin (2006). Simple modification to describe the soil water retention curve between saturation and oven dryness. *Water Resour. Res.* 42:W11501, doi:10.1029/2005WR004699.
3. **Khlosi, M.**, W.M. Cornelis, A. Douaik, M.Th. van Genuchten, and D. Gabriels (2008). Performance evaluation of models that describe the soil water retention curve between saturation and oven dryness. *Vadose Zone J.* 7:87–96.
4. **Khlosi, M.**, W.M. Cornelis, A. Douaik, A. Hazzouri, H. Habib, D. Gabriels, 2013. Exploration of the interaction between hydraulic and physicochemical properties of Syrian soils *Vadose Zone J.* 2013. 12:–. doi:10.2136/vzj2012.0209.

In preparation in international journals with peer review, in ISI Web of Science

1. **Khlosi, M.**, W.M. Cornelis, M. Hamdosh, A. Hazzouri, H. Habib, D. Gabriels. Support vector machines to enhance the performance of pedotransfer functions for predicting the water retention properties of calcareous soils. To be submitted to *Eur. J. Soil Sci.*
2. **Khlosi, M.**, W.M. Cornelis, A. Hazzouri, H. Habib, D. Gabriels. Impact of particle-size distribution changes associated with carbonates on the predicted soil-water retention curve. To be submitted to *J. Hydrology.*

Books (B3)

1. De Boever, **M**, **Khlosi, M.**, Delbecque, N., De Pue, J., Ryken, N., Verdoodt, A., Cornelis, W.M., Gabriels, D., 2013. Desertification and land degradation: processes and mitigation. Third conference on desertification and land degradation. Conference proceedings. Ghent University, UNESCO Chair on Eremology, 159 pp.

Articles in books (B2)

1. **Khlosi, M.**, W.M. Cornelis, D. Gabriels (2008). Analyzing the effects of particle-size distribution changes associated with carbonates on the predicted soil-water retention curve. In: Gabriels et al. (2008). Combating desertification. Assessment, adaptation and mitigation strategies. International Centre for Eremology - Belgian Development Cooperation. p. 100-105.

Poster presentations

1. Hazzouri, A., and **Khlosi M.** (1998). Effect of saline water on salt accumulation and the yield of different chickpea cultivars in the north of Syria. Pages 216-223 in Challenges Facing the Management Profession in the Middle East and European Countries to Manage Agricultural Resources Effectively: Proceedings, Bradford Management and Technology Centre Ltd/Ventures and Consultancy Bradford Ltd/University of Aleppo, 14-16 July 1998, Harrogate, United Kingdom. Bradford Management and Technology Centre, Bradford, United Kingdom.
2. **Khlosi, M.**, W.M. Cornelis and D. Gabriels (2008). Describing the soil-water retention curve between saturation and oven dryness. EGU General Assembly 2008, 13-18 April 2008, Vienna, Austria, Geophysical Research Abstracts, Vol. 10, EGU2008-A-08556.
3. **Khlosi, M.**, W.M. Cornelis, D. Gabriels (2008). Analysing the effects of particle-size distribution changes associated with carbonates on the predicted soil-water retention curve. In: D. Gabriels and W. Cornelis (Eds.). Conference on Desertification, 23 January 2008, Ghent, Belgium.
4. **Khlosi, M.**, W.M. Cornelis, D. Gabriels (2008). Assessment of hydraulic properties for sustainable soil and water management in northwest Syria. In: D. Gabriels and W. Cornelis (Eds.). Conference on Desertification, 23 January 2008, Ghent, Belgium.
5. **Khlosi, M.**, W.M. Cornelis, A. Douaik, A. Hazzouri, H. Habib, M. Hamdosh, D. Gabriels (2014). Framework for predicting hydraulic properties of calcareous arid lands. EGU General Assembly 2014, 27 April – 02 May 2014, Vienna, Austria, Geophysical Research Abstracts, Vol. 16, EGU2014-13846, 2014.