

## Describing the soil physical characteristics of soil samples with cubical splines

Jan G. Wesseling · Coen J. Ritsema · Jannes Stolte ·  
Klaas Oostindie · Louis W. Dekker

Received: 28 September 2006 / Accepted: 13 March 2007 / Published online: 17 April 2007  
© Springer Science+Business Media B.V. 2007

**Abstract** The Mualem-Van Genuchten equations have become very popular in recent decades. Problems were encountered fitting the equations' parameters through sets of data measured in the laboratory; parameters were found which yielded results that were not monotonic increasing or decreasing. Due to the interaction between the soil moisture retention and the hydraulic conductivity relationship, some data sets yield a fit that seems not to be optimal. So the search for alternatives started. We ended with the cubical spline approximation of the soil physical characteristics. Software was developed to fit the spline-based curves to sets of measured data. Five different objective functions are tested and their results are compared for four different data sets. It is shown that the well-known least-square approximation does not always perform best. The distance between the measured points and the fitted curve, as can be evaluated numerically in a simple way, appears to yield good fits when applied as a criterion in the optimization procedure. Despite an increase in computational effort, this method is recommended over the least square method.

**Keywords** Splines · Soil moisture retention curve · Hydraulic conductivity curve · Curve-fitting · Optimization · Pressure head · Moisture content · Objective function

---

J. G. Wesseling (✉) · C. J. Ritsema · J. Stolte · K. Oostindie · L. W. Dekker  
Alterra, Soil Science Centre, P.O. Box 47, 6700 AA Wageningen, The Netherlands  
e-mail: jan.wesseling@wur.nl

C. J. Ritsema  
Wageningen University, Erosion and Soil Water Conservation Group,  
P.O. Box 47, 6700 AA Wageningen, The Netherlands  
e-mail: coen.ritsema@wur.nl

J. Stolte  
e-mail: jannes.stolte@wur.nl

K. Oostindie  
e-mail: klaas.oostindie@wur.nl

L. W. Dekker  
e-mail: louis.dekker@wur.nl

## 1 Introduction

The capacity of digital computers has increased enormously in recent years. This means more computations can be performed in less time. Now it becomes possible to design and run programs on a personal computer or minicomputer that previously could run only on a mainframe. In the field of unsaturated soil moisture flow this has resulted in models getting more and more detailed.

Up to a decade ago the majority of numerical models simulated one-dimensional (i.e. vertical) moisture flow only (see e.g. Hansen et al. 1990; Simunek et al. 1998; Van Dam et al. 1997a, b). Nowadays more and more two- (e.g. Vogel 1997; Heinen 1997) and three-dimensional models (e.g. Russo et al. 1998) are available for potential users. See the website of the International Groundwater Modeling Center (<http://www.mines.edu/igwmc/software/>) for more models on groundwater flow and their availability. Most of these models simulate transient moisture flow based upon the Richards' equation.

Whatever number of dimensions is considered and whatever equation is applied, some soil physical relationships should be known. In case of the Richards' equation, these relationships are the soil moisture retention curve (a relationship between pressure head and moisture content) and the hydraulic conductivity function (a relationship between the hydraulic conductivity and the pressure head). Many researchers do not take soil samples themselves but depend on data found in literature or available databases such as the Staringseries (Wösten 1987; Wösten et al. 2001b; Stolte et al. 2007, in preparation), the UNSODA database (Nemes et al. 2001) or the Hypres database (see <http://www.macaulay.ac.uk/hypres/>, Wösten et al. 1998). Therefore it is of the greatest necessity to describe the measured relationships the best way possible.

A lot of attempts have been made to describe the soil physical relationships with analytical equations. Nowadays one of the most widely used methods is the so-called Mualem-Van Genuchten equation (Van Genuchten 1980). It describes both the soil moisture retention curve and the hydraulic conductivity relationship using only 6 parameters. This simplicity is at the same time the weakness of the method: not all combinations of curves can be approximated adequately. Software is available to estimate the 6 parameters of the equations using the measured points (Van Genuchten et al. 1991; Stolte 1997). This software is usually based upon the least square approximation and some second order optimization technique. The main disadvantage of the least square approximation is that it usually takes into account the vertical distance between the measured point and the fitted curve. The second order optimization techniques yield a fast way to a (possibly local) minimum. This requires quite some numerical calculations since the second order derivative of a function, the Hessian matrix, should be evaluated numerically. Still, it often appears that not all combinations of soil water retention curves and the hydraulic conductivity relationship can be adequately approximated.

Therefore, several authors have presented the approximation of the relationships using polynomials or cubical splines (e.g. Erh 1972; Kastanek and Nielsen 2001; Prunty and Casey 2002; Bitterlich et al. 2004). Despite the fact that these publications show the power and capability of the spline functions, application has not yet been very popular in research. Therefore this study shows another spline approximation of the soil physical relationships. For this purpose software was written to fit curves through the points measured in the laboratory. This software is based upon some simple direct-search optimization techniques. In most publications the spline-functions are fitted using the least square of deviations criterion. Despite the usefulness of this method, there are other and better criterions known in mathematics. In this paper five different objective functions are described and tested on four

soil samples. Results are also compared to approximations obtained with the commonly used RETC-program.

## 2 Materials and methods

To understand the importance of correct soil physical relationships, one should be familiar with the equations governing soil moisture flow in the unsaturated zone of the soil. This partial differential equation is presented in the next paragraph. The individual points of these relationships are usually measured in the field or the laboratory, yielding a collection of points. Plotting these points in a graph usually results in some 'clouds' of points. Several attempts have been made to develop equations that are generally applicable to describe the relationships. Some of these are mentioned in the following paragraph. The parameters of these relationships should be estimated in such a way that the curve fits the measured points in the best possible way. We choose to describe the curve by cubical splines because it can be expected that the accuracy of the derived approximation would increase. These splines are described in detail. In order to obtain the best fit of the spline-approximation through the measured points, some optimization procedure is required. The general theory of these procedures is presented. An optimization procedure can work only if there is some function it has to optimize, or in this case, minimize. Five different types of these functions are described. The optimal points of the spline curves are calculated by using these functions.

### 2.1 Soil moisture flow

#### 2.1.1 The governing equations

Soil moisture flow in a  $N$ -dimensional system can be described by the following partial differential equation

$$\frac{\partial \theta}{\partial t} = \nabla \cdot (K \nabla H) - S \quad (1)$$

where

$$\nabla = \sum_{j=1}^N \frac{\partial}{\partial x_j} \quad (2)$$

and  $H$ =total head [L];  $x_i$ =coordinate in  $i$ th direction;  $\theta$ =volumetric moisture content [ $L^3L^{-3}$ ];  $t$ =time [T];  $K$ =prevailing hydraulic conductivity [ $LT^{-1}$ ];  $S$ =sink term representing drainage flow or root extraction [ $T^{-1}$ ].

Then introduce anisotropy (for three-dimensional flow)

$$K = [K_{ij}] = \begin{pmatrix} K_{11} & K_{12} & K_{13} \\ K_{21} & K_{22} & K_{23} \\ K_{31} & K_{32} & K_{33} \end{pmatrix} \quad (3)$$

and write the total head  $H$  as

$$H = h + x_3 \quad (4)$$

where  $h$ =pressure head [L] and  $x_3$ =vertical position [L].

Introducing the differential moisture capacity  $C [L^{-1}]$  then yields

$$C \frac{\partial H}{\partial t} = \sum_{i=1}^3 \frac{\partial}{\partial x_i} \sum_{j=1}^3 K_{ij} \frac{\partial H}{\partial x_j} - S \tag{5}$$

with

$$C = \frac{d\theta}{dH} = \frac{d\theta}{dh} \tag{6}$$

Both  $K$  and  $C$ -values depend on the prevailing pressure head  $h$ . To solve Eq. 5, the  $C(h)$ ,  $K(h)$  and  $h(\theta)$ -relationships should be known.

2.1.2 Analytical approximations of soil moisture retention and hydraulic conductivity functions

The relationships between moisture content, pressure head and hydraulic conductivity are very important to obtain a correct description of unsaturated moisture flow. They can be measured in the field or in the laboratory using different methods (e.g. Ahuja et al. 1980; Bresler et al. 1978; Wind 1966; Boels et al. 1978; Bouma et al. 1971; Bouma and Denning 1972; Stolte and Veerman 1990; Kool and Parker 1987; Kool et al. 1987; Van Dam et al. 1990). Several authors try to derive these relationships from other known soil physical data, e.g. particle size distribution and organic matter content (Ahuja et al. 1988; Alexander and Skaggs 1987; Bloemen 1980; Bruce 1972; Schuh and Bauder 1986). Several analytical functions have been developed to describe the soil physical relationships. See Leij et al. (1997) and Leong and Rahardjo (1997) for an excellent overview and comparison of a lot of closed-form expressions.

One of the most frequently used closed-form descriptions of the soil physical relationships is the one introduced by Van Genuchten (1980) who describes the  $K(h)$ - and  $h(\theta)$ -relationships as S-shaped curves with only 6 parameters. It appears, however, that this closed-form approximation can still be improved (Fuentes et al. 1991, 1992). These equations state

$$\theta = \theta_r + \frac{\theta_s - \theta_r}{(1 + |\alpha h|^n)^m} \tag{7}$$

$$h = - \frac{\left( \left( \frac{\theta_s - \theta_r}{\theta - \theta_r} \right)^{\frac{1}{m}} - 1 \right)^{\frac{1}{n}}}{\alpha} \tag{8}$$

$$K = K_s \frac{((1 + |\alpha h|^n)^m - |\alpha h|^{n-1})^2}{(1 + |\alpha h|^n)^{m(l+2)}} \tag{9}$$

with

$$m = 1 - \frac{1}{n} \tag{10}$$

Differentiation also yields an equation for the differential moisture capacity:

$$C_\theta = \frac{\partial \theta}{\partial h} = n \cdot m \cdot \alpha \cdot |\alpha h|^{n-1} \cdot \frac{\theta_s - \theta_r}{(1 + |\alpha h|^n)^{m+1}} \tag{11}$$

where  $h$  = pressure head (cm),  $\theta$  = moisture content ( $\text{cm}^3 \text{cm}^{-3}$ ),  $K$  = hydraulic conductivity ( $\text{cm d}^{-1}$ ) and  $C$  = differential moisture capacity ( $\text{cm}^{-1}$ );  $\alpha$  ( $> 0$ , in  $\text{cm}^{-1}$ ) is related to the inverse of the air entry suction;  $n$  ( $> 1$ ) is a measure of the pore-size distribution;  $K_s$  is a fitted matching point at saturation ( $\text{cm d}^{-1}$ ) and  $l$  is an empirical parameter that is normally assumed to be 0.5.

Parameters for the closed-form expressions can sometimes be determined by inverse modeling (Yeh 1986; Jhorar et al. 2001), but the majority is still derived from laboratory data. Relationships between soil physical characteristics and other soil properties (e.g. organic matter content, particle size distribution, texture etc.) are getting more and more popular as well (Van Alphen et al. 2001; Elsenbeer 2001; Minasny et al. 1999; Schaap and Leij 1998; Schaap et al. 2001; Stolte et al. 1996; Nemes 2003; Wösten et al. 2001a). These are the pedotransfer functions.

Instead of using a closed-form description of the soil physical relationships, sometimes the soil physical characteristics in numerical models are described by tables with data-couples, e.g.  $(\theta, h)$ - and  $(h, K)$ -points. To obtain the corresponding values for data between the specified nodes, linear interpolation is applied (sometimes after taking the  $^{10}\log$  of the values. This may be the cause of some unnecessary iterations because these functions may cause 'jumps': the value of the differential moisture capacity of a point may not be continuous across such a point.

In the following sections it is assumed that we have a number of measured points of a soil physical relationship. At this point it does not matter how these data are obtained, they are just available and assumed to be correct. They can be presented as a cloud of points in a graph. To be able to use the relationship in numerical modeling, it should be described by some function.

## 2.2 Splines

One of the most powerful mathematical methods of describing a function is the one applying splines. These are piecewise-continuous polynomials. Several types of splines have been applied in soil science, e.g. cubic splines (Hampton 1990; Kastanek and Nielsen 2001; Prunty and Casey 2002) and quadratic B-splines (Bitterlich et al. 2004). Though nowadays spline interpolation is included in mathematical and statistical packages like MathLab (<http://www.mathworks.com/products/curvefitting/description5.html>) and Statistica (<http://www.statsoft.com>), these packages are not generally available for researchers. Beside that, it appears that most researchers are unfamiliar with the theory of splines. Therefore this section will describe the theory of cubical splines in a general way. Splines are described by defining only a few data points. These points are sometimes called Virtual Data Points (see Kastanek and Nielsen 2001). The function-values for points located between these Virtual Data Points are then calculated by computing the value of the spline function. Now suppose we have a series of (virtual) data pairs  $(x_i, y_i)$ ,  $i = 1 \dots N$ . We assume the series is ordered by  $x$ , i.e.  $x_i < x_{i+1}$  and we focus on the interval  $[x_i, x_{i+1}]$ . Then linear interpolation in that interval gives the interpolation formula

$$y = \alpha y_i + \beta y_{i+1} \quad (12)$$

where

$$\alpha = \frac{x_{i+1} - x}{x_{i+1} - x_i} \quad (13)$$

$$\beta = 1 - \alpha = \frac{x - x_i}{x_{i+1} - x_i} \tag{14}$$

Since it is piecewise linear, the equation above has a second derivative of zero in the interior of each interval and an undefined second derivative at the abscissas  $x_j$ . Now the goal of cubic spline interpolation is to get an interpolation formula that is smooth in the first derivative and continuous in the second derivative, both within an interval and at its boundaries. It is required that the second derivative varies linearly from a value  $y''_i$  on the left to a value of  $y''_{i+1}$  on the right. In order to fulfill all these requirements, we have to introduce a cubic polynomial instead of a linear relationship. Doing so, we will have the desired continuous derivative. The polynomial should be constructed in such a way it has zero values at  $x_i$  and  $y_{i+1}$  so adding it will not spoil the agreement with the values  $y_i$  and  $y_{i+1}$  at  $x_i$  and  $x_{i+1}$ . Some elementary calculus now yields the new interpolation formula:

$$y = \alpha y_i + \beta y_{i+1} + \gamma y''_i + \eta y''_{i+1} \tag{15}$$

where

$$\gamma = \frac{1}{6} (\alpha^3 - \alpha) (x_{i+1} - x_i)^2 \tag{16}$$

and

$$\eta = \frac{1}{6} (\beta^3 - \beta) (x_{i+1} - x_i)^2 \tag{17}$$

It is easy to check that  $y''$  is the second derivative of the interpolating polynomial. Remembering that  $\alpha, \beta, \gamma$  and  $\eta$  are functions of  $x$ , it is possible to find the derivative of  $y$  with respect to  $x$ :

$$\frac{dy}{dx} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} = \frac{3\alpha^2 - 1}{6} (x_{i+1} - x_i) y''_i + \frac{3\beta^2 - 1}{6} (x_{i+1} - x_i) y''_{i+1} \tag{18}$$

and the second derivative

$$\frac{d^2y}{dx^2} = \alpha y''_i + \beta y''_{i+1} \tag{19}$$

The problem now is that the values of the  $y''_i$ 's are supposed to be known but are not. However, up to now it was not required that the first derivative be continuous across the boundary between two intervals. The key idea about cubic splines is to require this continuity and to use it to get equations for the second derivatives.

The required equations are obtained by setting the first derivative for  $x = x_i$  in  $[x_{i-1}, x_i]$  equal to the same equation for  $x = x_i$  in  $[x_i, x_{i+1}]$ . After some rearrangement this gives for each  $i \in [2 \dots N - 1]$ :

$$\frac{x_i - x_{i-1}}{6} y''_{i-1} + \frac{x_{i+1} - x_{i-1}}{3} y''_i + \frac{x_{i+1} - x_i}{6} y''_{i+1} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} - \frac{y_i - y_{i-1}}{x_i - x_{i-1}} \tag{20}$$

These are  $N - 2$  linear equations in the  $N$  unknowns  $y''_i, i = 1 \dots N$ . For a unique solution we need to specify two further conditions, typically taken as boundary conditions at  $x_1$  and  $x_N$ . The most common ways to do this are either to set one or both of the boundaries to zero or calculate the first derivative to have a specified value on either or both boundaries.

After calculating the second derivatives once, their values should be stored and the interpolation between points can rapidly be performed according to Eq. 15. See also Press et al. (1986) and Erh (1972).

### 2.3 Parameter estimation

In the previous section the spline approximation has been described. But how can we obtain the values of the Virtual Data Points  $(x_i, y_i)$  that fit the measured data in the best possible way? For that purpose we need the parameter estimation or optimization techniques. These will be presented in general in this section.

#### 2.3.1 The problem

Most systems and relationships in soil physics can be described by a mathematical function  $f$ , having some parameters  $p$ . Depending on the value of the parameter set  $\underline{p}$ , a series of input data (state variables)  $\underline{x}$  (a vector), applied to the system described by  $f$ , and assuming some initial and boundary conditions, produces a series of output data  $\hat{\underline{y}}$  (a vector as well). Written in a mathematical form

$$\hat{\underline{y}} = f(\underline{p}, \underline{x}) \quad (21)$$

where  $\underline{x}$  = vector with independent (state) variables;  $\underline{p}$  = parameter vector;  $\hat{\underline{y}}$  = vector with output data obtained from the combination of the vector with state variables  $\underline{x}$  and the vector with parameters  $\underline{p}$ ;  $N_x$  = number of state variables (dimension of  $\underline{x}$ );  $N_y$  = number of output data (dimension of vector  $\underline{y}$ );  $N_p$  = number of parameters (dimension of vector  $\underline{p}$ );  $f$  = mathematical function that describes the conversion of input data  $\underline{x}$  to output data  $\underline{y}$ , depending on the parameter set  $\underline{p}$ .

In this context the function  $f$  may vary from a simple relation to a very complex set of equations. See e.g. [Vrugt et al. \(2002\)](#) and [Mertens \(2003\)](#) for more details about parameter identifiability and parameter estimation strategies in subsurface hydrology.

In some cases the parameter vector  $\underline{p}$  is known beforehand. Then the output vector  $\hat{\underline{y}}$  of the function  $f$  can be seen as an accurate response to the data in the input vector  $\underline{x}$ , presuming the system is described correctly by  $f$ . However, this is not always the case. Presuming again that  $f$  presents a correct description of the system under consideration, one has to vary the values in the parameter vector  $\underline{p}$  to obtain an output vector  $\hat{\underline{y}}$  that approximates the vector  $\underline{y}$  with measured values as closely as possible (calibration of parameter values). Often the parameter vector  $\underline{p}$  represents physical characteristics which may be measured in laboratories or in the field, e.g. saturated hydraulic conductivity or the slope of the soil water retention curve. Usually the measurements of physical characteristics are time consuming and expensive. When  $f$  describes the system under consideration correctly, the parameter vector  $\underline{p}$  can sometimes be derived from simple input–output measurements, (e.g. the One-step outflow method ([Kool et al. 1985](#)) or the Multi-step method ([Van Dam et al. 1990](#)) in case of soil physical parameters). In these cases the input to the system is fully controlled and the output is measured. Knowing the input- and output-vectors of the system by measurement, an attempt can be made to find the vector  $\underline{p}$ . The values of the parameter vector  $\underline{p}$  should be set in such a way that the calculated values in the output vector  $\hat{\underline{y}}$  approximate the measured values in vector  $\underline{y}$  as well as possible. This is called an indirect inverse problem. The solution of the problem depends on several factors, e.g. the type of function  $f$  and the amount of parameters  $N_p$ . Several authors presented solutions of the inverse problems ([Baker 2006](#); [Tarantola 1987](#); [Spriet and Vansteenkiste 1982](#); [Press et al. 1986](#)). The inverse problem can be solved by all known optimization techniques (e.g. [Cesari 1983](#); [Gill et al. 1981](#); [Luenberger 1973](#); [Press et al. 1986](#); [Rust and Burrus 1972](#); [Van Beek and Hendriks 1983](#); [Baker 2006](#)). The problem discussed above can be stated as follows:

$$\underset{\underline{p}}{\text{Minimize}} \ g \left( \underline{p}, \underline{y}, \widehat{\underline{y}} \right) \tag{22}$$

where  $g$  is the objective function. The value of  $g$  is not only depending on the measured ( $\underline{y}$ ) and the calculated ( $\widehat{\underline{y}}$ ) output values, but one can implicitly take care of constrains as well. If only positive values are allowed for a parameter, negative values will yield a high value of the function (in case of a minimization procedure). Stating the problem this way, the function  $g$  should be chosen such that its value gets smaller when the calculated values approximate the measured values.

Translating the theory above to our spline problem we can consider the vector  $\underline{p}$  as consisting of the  $x$ - and  $y$ -coordinates of the Virtual Data Points and the function  $g$  as being some function considering the goodness of fit, the objective function. This function will be described later in this article.

Sometimes one is forced to apply weight functions to obtain reasonable results. Then each point is assigned a value  $\omega$  and the function is calculated as

$$\gamma \left( \underline{y}, \widehat{\underline{y}} \right) = \sum_{i=1}^{N_y} \omega_i g \left( y_i, \widehat{y}_i \right) \tag{23}$$

where  $\gamma$  is the resulting objective function. Usually  $\omega_i$  is determined by expert judgment or by trial and error. To find the optimum values of the weight function, a separate optimization procedure should be developed. In the remainder of this paper the values of the weight functions will be set to 1 (all points are equally important).

### 2.3.2 The minimization procedure

From literature many optimization procedures are known. These procedures can be divided into two main classes:

1. Direct search methods. When applying these methods only the function  $g(y, \widehat{y})$  has to be evaluated.
2. Line search methods. In all of the methods belonging to this group a search for the minimum is performed where the minimum of  $g$  is expected. In most cases this direction is found by calculating the derivative  $g'(y, \widehat{y})$ , where

$$\underline{g}' \left( \underline{y}, \widehat{\underline{y}} \right) = \left( \frac{\partial g}{\partial p_1}, \frac{\partial g}{\partial p_2}, \dots, \frac{\partial g}{\partial p_{N_p}} \right)^T \tag{24}$$

Some line-search methods take into consideration the matrix with second derivatives (the Hessian matrix)  $H$  as well. In most cases the function  $g$  cannot be described analytically, so  $g'$  and  $H$  have to be evaluated numerically, requiring a large amount of function evaluations.

To obtain the results presented in the remainder of this paper, two different direct search optimization techniques were investigated: the Simplex method (Nelder and Mead 1965; Press et al. 1986) and the Controlled Random Search algorithm or CRS (Price 1979). The Simplex method was chosen because it is a well-known method that has been applied in a wide range of problems. The CRS-method is known for its simplicity and has been applied with good results (Metselaar 1999). As both methods yielded very similar results, only the results with the CRS-method are discussed.



### 2.3.3 The objective functions

In the previous sections the optimization problem has been stated and some optimization procedures have been mentioned. But what function do we have to optimize? And how do we calculate this function? Some of the many possible objective functions will be described below.

It is assumed that a number of measured data  $y(x, p)$  has to be approximated by  $\hat{y} = f(x, p)$ . The function  $y_i = f(x_i, p)$  is graphically represented by a geometrical figure  $\Omega$  in space  $R^2$ . This may be either a straight line or some kind of complex curve (See Fig. 1).

1. The simplest objective function is the sum of the absolute deviation between the measured value  $y_i$  and the calculated value  $\hat{y}_i$  for a specified  $x_i$ . This is written as

$$g(\underline{y}, \underline{\hat{y}}) = \sum_{i=1}^{N_y} |y_i - \hat{y}_i| \tag{25}$$

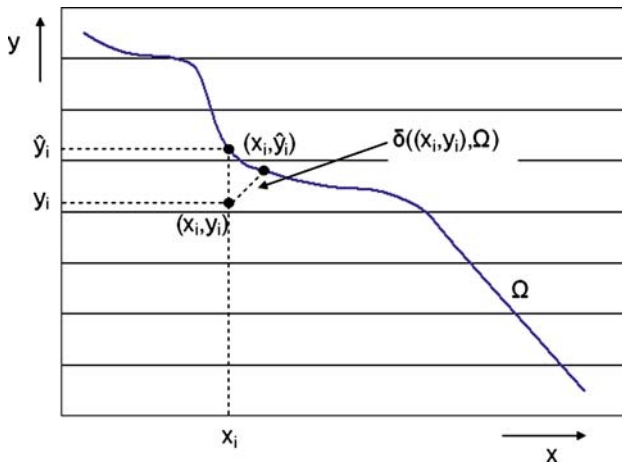
The main advantage of this equation is its simplicity, while one of its drawbacks is that large deviations have an extremely big influence on the results.

2. The most widely used objective function is the sum of the squares of deviations:

$$g(\underline{y}, \underline{\hat{y}}) = \sum_{i=1}^{N_y} (y_i - \hat{y}_i)^2 \tag{26}$$

Complete optimization theories have been developed based upon this objective function, e.g. by Marquardt (1963). The disadvantage mentioned before is even stronger in this case. This least square function has been applied regularly for use in describing soil physical properties (e.g. Stephens and Rehfeldt 1985; Kool and Parker 1988; Bruckler et al. 1987).

3. If the values of  $y$  differ some magnitudes it is advisable to consider the relative error instead of the absolute error that was taken in the previous 2 functions. Then  $g$  can be written as in



**Fig. 1** A schematic representation of a measured point  $(x_i, y_i)$ , a calculated point  $(x_i, \hat{y}_i)$  and the distance  $\delta$  between the measured point and the calculated line  $\Omega$

$$g(\underline{y}, \underline{\hat{y}}) = \sum_{i=1}^{N_y} \frac{|y_i - \hat{y}_i|}{|y_i|} \tag{27}$$

A disadvantage of this method is the fact that problems arise when  $y_i$  approaches 0. In that case an exception should be made and a different function should be taken for this point.

4. Another possibility is to calculate the difference between the absolute values of the quotients of measured and calculated values

$$g(\underline{y}, \underline{\hat{y}}) = \sum_{i=1}^{N_y} \frac{|\hat{y}_i|}{|y_i|} + \frac{|y_i|}{|\hat{y}_i|} - 2 \tag{28}$$

The main advantage of this method is that it does not matter if the calculated value is a certain fraction higher or lower than the measured value: the contribution to the objective function is the same. This can easily be seen when the value of  $g$  is plotted against the percentage of deviation. There are two disadvantages to this method: the first one occurs when either  $y_i$  or  $\hat{y}_i$  approaches 0 (division by zero), the second problem is that the possibility of different signs of  $y_i$  and  $\hat{y}_i$  is not taken into account. If this sign difference occurs, an extra (large) value has to be added to the objective function to avoid this sign difference in the final solution.

5. The objective function that is intuitively seen as the most correct one represents the sum of the distances from the point  $(x_i, y_i)$  to  $\Omega$ .

$$g(\underline{y}, \underline{\hat{y}}) = \sum_{i=1}^{N_y} \delta[\Omega, (x_i, y_i)] \tag{29}$$

This would yield the best approximation of a line through a series of points. In practice it is hardly applied, due to the fact that it is difficult to calculate the distance between a point and a line if the line is described by a complex equation. If this is the case, calculating the distance can be viewed as another problem of optimization: find the coordinates of the element of  $\Omega$  where the distance to the point under consideration is at minimum. Application of this criterion may take a lot of extra programming. It can be considered as a Total Least Squares approach used when both  $x$  and  $y$  contain errors in a regression problem.

### 2.3.4 Goodness of fit

The most widely spread indicator for the goodness of fit is the root-mean-square error (see e.g. Prunty and Casey 2002). This value is based upon the sum of the squares of deviations between measured and calculated values at a given  $x$ -value. Or mathematically:

$$SSQ(\underline{p}) = \sum_{i=1}^{N_y} (y_i - f(\underline{p}, x_i))^2 \tag{30}$$

and

$$RMSE = \sqrt{\frac{SSQ(\underline{p})}{N_y}} \tag{31}$$

As we stated before, the sum of the distances from the measured point to the calculated line is a better criterion for the goodness of fit than the root of squares of deviations. Therefore we want to introduce a new criterion for the goodness of fit, the Mean Distance of Points to Line (MDPL). This can simply be calculated by

$$\text{MDPL} = \frac{1}{N_y} \sum_{i=1}^{N_y} \delta(\Omega, (x_i, y_i)) \quad (32)$$

### 2.3.5 Calculating the distance between a point and a line

From mathematics many equations are known to calculate the distance from a point to a straight line. Equations for some simple other line types can be found as well. But what to do with a set of spline functions fitted through a number of Virtual Data Points? As that was not so simple, we solved the problem numerically:

- Suppose we have data points in the range  $x_1$  to  $x_{N_x}$ . Each  $x$ -value has a corresponding  $y$ -value.
- Divide the range into a large number ( $N$ , say 1,000) intervals. In this way we obtain  $N + 1$  values on the  $x$ -axis, say  $\rho_j$ . For each of these values we can calculate a corresponding function value  $\sigma_j$  by spline-interpolation.
- For each measured point  $(x_i, y_i)$  the distance to the line can now be found as the minimum value of the measured point to one of the generated points, or in an equation:

$$\delta[\Omega, (x_i, y_i)] = \min_j (\delta((x_i, y_i), (\rho_j, \sigma_j))) = \min_j \left( \sqrt{(x_i - \rho_j)^2 + (y_i - \sigma_j)^2} \right) \quad (33)$$

## 2.4 Software to estimate the points of the optimal spline functions

In this paper methods are developed to approximate the soil physical characteristics by means of a number of cubical splines. Only a few points of the line have to be found. Values for other points can be calculated by means of a simple cubical interpolation. Assuming the splines are described by Eq. 15, then we may consider this equation as a special form of the general equation (21). From the laboratory we have a number of combinations of  $x$ - and  $y$ -values ( $\theta$  and  $h$  in case of soil moisture retention,  $h$  and  $K$  in case of hydraulic conductivity). The optimization techniques can now be applied to this equation to find the optimal fit for the data. Any of the objective functions described above (Eq. 25–29) may be used.

The equations and methods described earlier are translated into some computer programs (all developed in Delphi 7 and running under Windows XP):

- Optim.exe is a program to
  1. fit a polynomial of an arbitrary degree through a number of points
  2. fit the Van Genuchten parameters for a set of  $K(h)$  and  $h(\theta)$  data
- Spline.exe is a program to manually fit cubical splines through a dataset. The spline-points have to be entered and the program shows the interpolated values as a graph.
- Splop.exe automatically fits a spline function through a set of data.

It is essential that the fitted line is monotonic. See e.g. Hampton (1990) and Bitterlich et al. (2004) who applied piecewise cubic Hermite interpolation to insure monocity. In general, one can indicate the requirement of a monotonic increasing or monotonic decreasing line.

The check of being monotonic is performed by dividing the range of the  $x$ -values into a large number of intervals and checking if the corresponding  $y$ -value is higher (resp. lower) than the one corresponding to the previous  $x$ -value.

It would go too far to describe the programs in more detail here. Only results of applying Splop.exe will be discussed here.

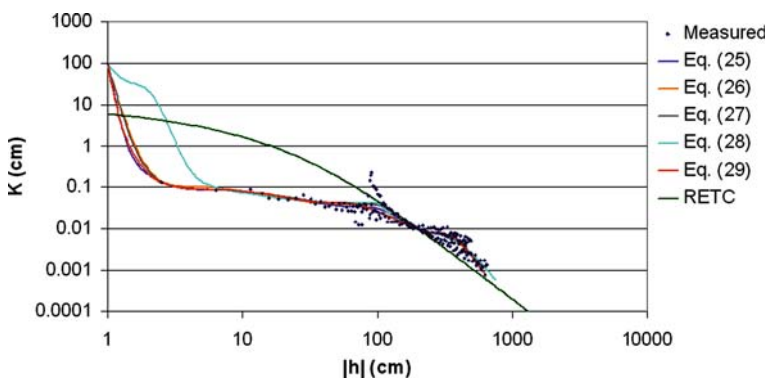
### 3 Results and discussion

We tested the five objective functions on five different data sets. We applied the methods described above on the hydraulic conductivity relationship of two soil samples. We approximated both the hydraulic conductivity function and the soil moisture retention curve of a soil sample. This sample was also used to show the influence of the number of Virtual Data Points on the optimal value of the objective function. Finally a comparison of the results obtained with the different objective functions is presented. Wherever applicable, the Mualem-Van Genuchten approximation of the data obtained by the program RETC (Van Genuchten et al. 1991) is shown as well. All runs with RETC were made with the same initial values (Stolte et al. 2007, in preparation) and the measured points of both the soil moisture retention curve and the hydraulic conductivity characteristic were taken into account, implying that the points of the soil moisture retention curve may influence the fit of the hydraulic conductivity curve and vice versa.

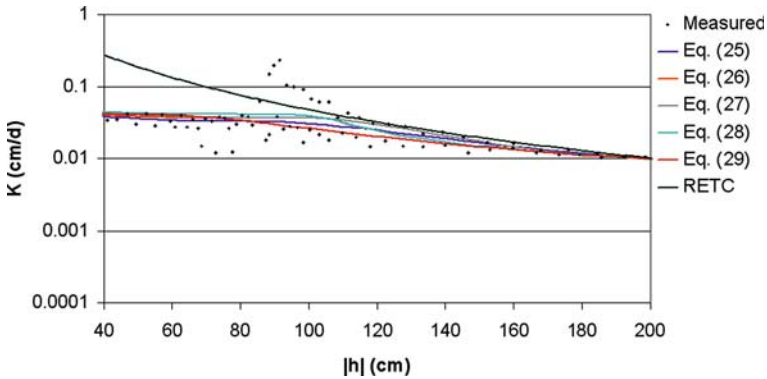
#### 3.1 Fitting $K(h)$ -relationships

From a site in The Netherlands, a sandy soil sample (here referred to as ‘Tox’) was collected at 30–38 cm depth for soil physical determination in the laboratory. Using the evaporation method (Wind 1966; Halbertsma and Veerman 1994), a series of  $K(h)$ -points was obtained. These data are used here to fit an approximation with splines. All objective functions were applied and the resulting curves are presented in Fig. 2.

It can be seen that on the left-hand side of the figure the objective function with the distance between point and line yields the best result. The quotient objective function yields a rather strange exception. After examining the data the cause became clear. Before optimizing the curve, first the  $^{10}\log$  of the  $y$ -values was taken. In this case the value is 0.1, yielding a



**Fig. 2** Measured hydraulic conductivity  $K$  ( $\text{cm d}^{-1}$ ) as a function of the pressure head  $h$  (cm) and the optimal fits for the Tox-sample



**Fig. 3** Measured hydraulic conductivity  $K$  ( $\text{cm d}^{-1}$ ) as a function of the pressure head  $h$  (cm) and the optimal fits for the Tox-sample in the range of  $-200$  to  $-40$  cm pressure head

log-value of  $-1$ . The program gets a  $y$ -value of  $10$ , yielding a log-value of  $1$ . Inserting these values in Eq. 29 yields an exact equality as the absolute values are considered. At the area with many points, the different object functions yield slightly different results again. It is also clear that the RETC-fit is nowhere near the measured point for values of  $h > -100$  cm. The conductivity at saturation ( $h = 0$ ) is not well approximated either. When zooming in on the range between  $40$  and  $200$  cm, the differences become visible (Fig. 3).

During the past decades the soil physical characteristics of quite a lot of soil samples have been determined at Alterra, Wageningen, The Netherlands. All these data were collected into the Priapus database and software was developed to make the data accessible via internet (Stolte et al. 2007, in preparation). From this database another soil sample was selected to use in this study. It was a sample from Glanerbrug in The Netherlands. The sample will be referred to as kh61. We selected this sample because, at first sight, both the soil moisture retention curve and the hydraulic conductivity curve seemed to show a good fit through the measured points (Fig. 4). However, when you plot only the values in the range of pressure head values between  $-200$  and  $-40$  cm (Fig. 5), it becomes clear there is a difference between the lines and that the RETC-fit will produce lower  $K$ -values than the other fits.

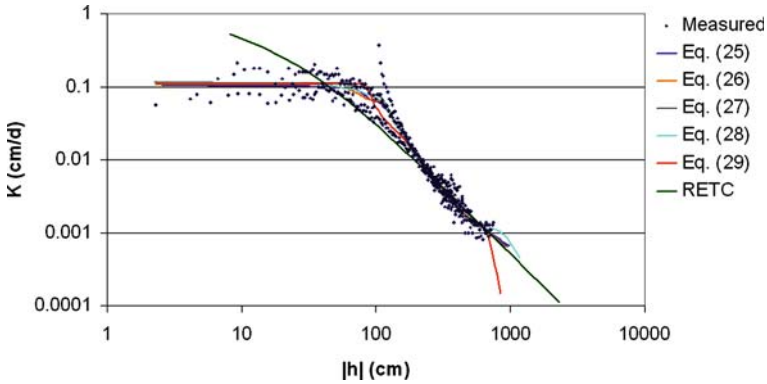
The optimization was performed for 10 spline points. The values between these points were interpolated applying Eqs. 25–29. Now an interesting question is how do these spline points differ for each objective function? In all cases we used 10 points (0 . . . 9). The values for the points of the optimal curves are presented in Table 1. The first column presents the point number. Then the values of the  $(h, k)$  point are presented where the subscript indicated the applied objective function (deviation, squared, relative, quotient and distance, respectively). From this table it can be seen that, though the curves look similar, both the  $h$  and the  $k$ -values differ quite a lot between the objective functions.

### 3.2 Fitting $h(\theta)$ and $K(h)$

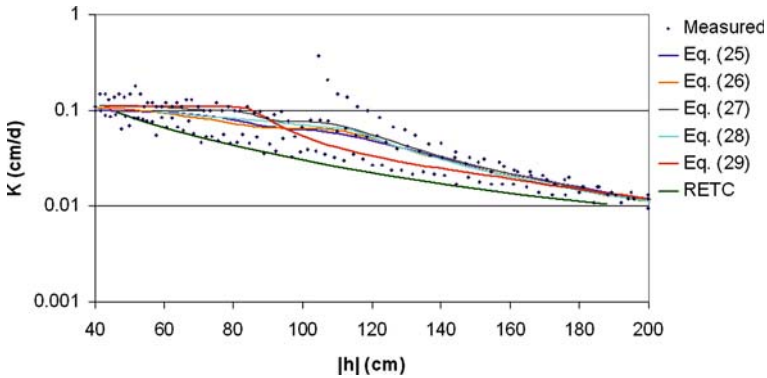
The sample considered next is one from a loamy soil near Catsop in the south-eastern part of The Netherlands (Code Cat3\_0-8B, nr. 809 in the Priapus database (Stolte et al. 2007, in preparation)). The hydraulic conductivity fit and the fit of soil moisture retention curve will be called kh809 en pF809 in this article. The Van Genuchten parameters for this sample can be obtained from the database as well. These parameters are calculated with the default settings of the computer program RETC (Van Genuchten et al. 1991). The curves obtained with RETC do not fit the data very well. Therefore this was an excellent example to read into

**Table 1** ( $K-h$ )-points of the spline-interpolation function for each objective function with  $h$  in cm,  $K$  in ( $\text{cm d}^{-1}$ ). The subscript of the  $h$ 's and  $K$ 's refer to the different objective functions

$i$	$h_{25}$	$K_{25}$	$h_{26}$	$K_{26}$	$h_{27}$	$K_{27}$	$h_{28}$	$K_{28}$	$h_{29}$	$K_{29}$
0	2.29	$1.06 \times 10^{-1}$	2.27	$1.07 \times 10^{-1}$	2.25	$1.14 \times 10^{-1}$	2.31	$1.12 \times 10^{-1}$	2.29	$1.10 \times 10^{-1}$
1	12.5	$1.02 \times 10^{-1}$	3.15	$1.07 \times 10^{-1}$	1.06	$1.13 \times 10^{-1}$	6.44	$1.12 \times 10^{-1}$	8.14	$1.10 \times 10^{-1}$
2	53.7	$9.69 \times 10^{-2}$	52.7	$9.97 \times 10^{-2}$	47.1	$1.12 \times 10^{-1}$	17.3	$1.12 \times 10^{-1}$	58.4	$1.10 \times 10^{-1}$
3	83.1	$7.61 \times 10^{-2}$	77.7	$7.59 \times 10^{-2}$	63.8	$1.06 \times 10^{-1}$	112.0	$6.49 \times 10^{-2}$	80.1	$1.10 \times 10^{-1}$
4	96.0	$6.36 \times 10^{-2}$	89.5	$6.62 \times 10^{-2}$	73.8	$9.97 \times 10^{-2}$	113.0	$6.39 \times 10^{-2}$	83.1	$1.08 \times 10^{-1}$
5	102.0	$6.23 \times 10^{-2}$	99.8	$6.41 \times 10^{-2}$	84.9	$9.31 \times 10^{-2}$	116.0	$5.86 \times 10^{-2}$	90.2	$8.01 \times 10^{-2}$
6	111.0	$5.68 \times 10^{-2}$	111.0	$6.12 \times 10^{-2}$	91.8	$7.91 \times 10^{-2}$	509.0	$1.55 \times 10^{-3}$	151.0	$2.18 \times 10^{-2}$
7	136.0	$3.52 \times 10^{-2}$	128.0	$4.10 \times 10^{-2}$	105.0	$7.53 \times 10^{-2}$	566.0	$1.26 \times 10^{-3}$	592.0	$1.26 \times 10^{-3}$
8	249.0	$6.58 \times 10^{-3}$	213.0	$9.52 \times 10^{-3}$	128.0	$4.50 \times 10^{-2}$	826.0	$1.08 \times 10^{-3}$	654.0	$1.07 \times 10^{-3}$
9	980.0	$6.64 \times 10^{-4}$	916.0	$6.49 \times 10^{-4}$	962.0	$6.38 \times 10^{-4}$	1180.0	$4.53 \times 10^{-4}$	844.0	$1.48 \times 10^{-4}$

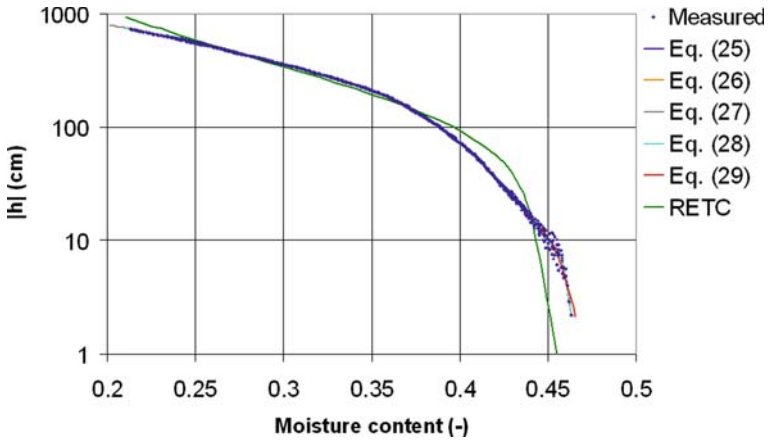


**Fig. 4** Measured hydraulic conductivity  $K$  ( $\text{cm d}^{-1}$ ) as a function of the pressure head  $h$  (cm) and the optimal fits for the kh61-sample

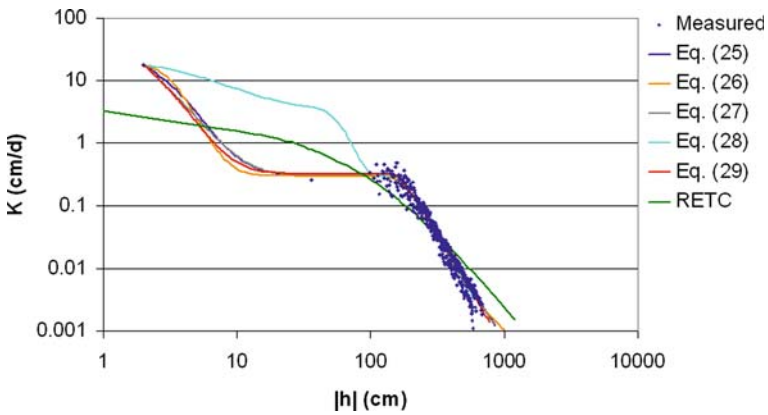


**Fig. 5** Part of the fitted curves through the measured hydraulic conductivity data of sample kh61 between pressure head values  $|h| = 40$  and  $|h| = 200$  cm

Splop and see what fits would be generated by that program. The results for the soil-moisture retention curve are shown in Fig. 6. From this figure it can be seen again that all objective functions yield nearly the same result. Only at the wet side of the curve the ‘distance’ function shows a slight difference. This is caused by the fact that the lines are almost vertical here. On the other hand the difference with the results of RETC is considerable. This difference is probably caused by taking the measured hydraulic conductivities into account as well. These measured hydraulic conductivity values are shown in Fig. 7, together with the computed results using the different objective functions. Considerable differences exist between the results of the various objective functions, due to the presence of only 2 points in the wet range, implying a lot of freedom in this zone. Again the objective function with the quotients is most divergent. As in the other example the values of 1 and  $-1$  are yielding the same value of the objective function. In this figure the difference with the RETC-solution is even more pronounced than in the previous figure.



**Fig. 6** Measured and fitted relationships between pressure head  $h$  (cm) and moisture content  $\theta$  ( $\text{cm}^3\text{cm}^{-3}$ ) for soil sample 809



**Fig. 7** Measured and fitted relationships between hydraulic conductivity  $K$  ( $\text{cm d}^{-1}$ ) and pressure head  $h$  (cm) for soil sample 809

### 3.3 How many points are required?

In the optimizations described in the previous sections we assumed the spline-functions were described with 10 Virtual Data Points. To investigate the influence of the number of spline-points on the minimum value of the objective function, a number of optimizations were performed with different objective functions and different number of points. The data were taken from the Glanerbrug example again. The minimum value that was reached is presented in Table 2. From this table it can be seen that for all objective functions a line-fit with 7 points is sufficient for the data presented here. When the number of points increases beyond 8, the minimum reached does not significantly decrease, indicating that the introduction of more spline points will not yield a better fit, only additional computational efforts. The same conclusion was made by [Bitterlich et al. \(2004\)](#).



**Table 2** Minima of objective functions  $G$  for different numbers of spline-points

Points	Eq. 25	Eq. 26	Eq. 27	Eq. 28	Eq. 29
3	105.17	26.55	82.25	20.80	24.06
4	96.34	18.44	72.34	12.93	21.75
5	88.26	16.09	65.10	12.08	19.51
6	88.22	16.00	64.70	11.70	19.47
7	87.96	15.49	65.07	11.45	19.47
8	87.90	15.49	64.78	11.44	19.12
9	87.86	15.48	64.63	11.44	19.10

### 3.4 The influence of the objective function

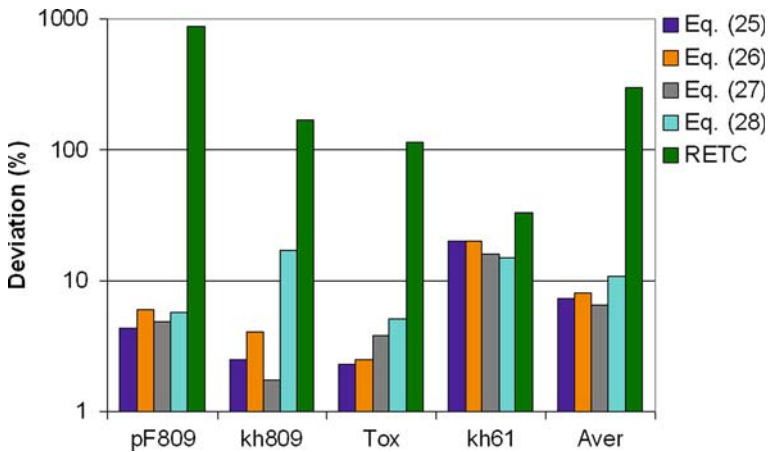
As was said before, the objective function with the squared deviation between measured and fitted values for a specified  $x$  (measured point) is most widely used. Intuitively however, the function that uses the distance between the (measured) point and the (calculated) line looks more reliable. Though at first sight all of the optimizations described in the previous sections yield good results, there are differences. These differences are hard to quantify however. One possible quantification is Mean Distance from Point to Line (MDPL, Eq. 32). This yields the function values in Table 3.

It can be seen that the optimization procedure using the sum of distances as the objective function yields the best result as expected and the RETC-fits yield the worse approximations. For the 'pF809'-case there are only slight differences between the objective functions. Here the order of accuracy is deviation, relative, quotient and squared. For the kh809-data the relative deviation and the deviation method yield almost the same result as the distance-method. The value of the squared method is a little higher and the one of quotient is much larger. For the Tox data the order is deviation, squared, relative and quotient. In case of kh61, the deviation and squared method do yield nearly equal results. So do the relative and the quotient method. The relatively large difference between the distance and the other methods was already shown in Fig. 3.

Figure 8 shows the percentages the MDPL is higher than the value obtained with the distance method for all other methods, as well as the average deviation. It can be seen that the deviations of the different spline approximations vary quite a bit. For the hk61-case all values are surprisingly high (15.07–20.12%). The deviations of the RETC-fits vary between 32.85 and 880.79%. The figure also shows the average deviation per object-function. These values are around 7%, except for the quotient method where it is 15.58 and 298.90% for RETC.

**Table 3** MDPL-values for the different optimization problems and with various objective functions

Case	Eq. 25	Eq. 26	Eq. 27	Eq. 28	Eq. 29	RETC
pF809	0.00257	0.00261	0.00256	0.00260	0.00246	0.02416
Kh809	0.01135	0.01152	0.01124	0.01298	0.01108	0.02967
Tox	0.00213	0.00213	0.00216	0.00218	0.00208	0.00446
hk61	0.02480	0.02481	0.02394	0.02375	0.02064	0.02743



**Fig. 8** The deviation of the results of each optimization case and objective function from the one calculated with the sum of distances

#### 4 Conclusions

- Using either the Simplex or the CRS optimization procedure, the objective functions ‘deviation’, ‘squared’, ‘relative’ and ‘distance’ achieved satisfactory results for pF and  $K(h)$ -data. The ‘quotient’ function in its basic form did not satisfy our requirements. It yields excellent results as long as the  $y$ -values remain positive all the time. The best results were obtained with the ‘distance’ objective function, however, this method requires more computer time.
- The spline approximation works very well to describe a relationship in hydrology. It has been applied and tested in a numerical model for one-dimensional moisture flow and appears to function easily and fast.
- The results of the different objective functions do yield differences in function values. They all yield some optimal curves, but the optima may differ from one another.
- The number of Virtual Data Points of a spline approximation is of some importance: if too few points are taken, the optimal value may not be approximated. If too many points are considered, there is no improvement of the optimal function compared to less points. There is only an increase in computational efforts.
- At this stage of the investigations it is not possible to make conclusions about the exact order of accuracy in fit for the objective functions. It will depend on the set of data points through which the curves have to be fitted.
- The frequently used RETC fitting code can be improved by using the ‘distance’ function instead of using the present ‘squared’ objective function.

**Acknowledgements** The authors wish to thank Cathelijne Stoof for testing the program splop.exe on her own datasets of  $K(h)$ - and  $h(\theta)$  relationships (Stoof et al. 2007, submitted). This resulted into the improvement of the software and the user-interface. We also wish to thank Demie Moore of Aquatrols for her excellent suggestions to improve the readability of the paper. Last but not least we would like to thank the reviewers of Transport in Porous Media who sent us a lot of questions, suggestions and remarks that helped tremendously in creating this final version.

## References

- Ahuja L.R., Green R.E., Chong S.-K. A simplified functions approach for determining soil hydraulic conductivities and water characteristics in situ. *Water Resour. Res.* **16**(5), 947–953 (1980)
- Ahuja L.R., Ross J.D., Bruce R.R., Cassel D.K. Determining unsaturated hydraulic conductivity from tensiometric data alone. *Soil Sci. Soc. Am. J.* **52**(1), 27–34 (1988)
- Alexander, L., Skaggs, R.W.: Predicting unsaturated hydraulic conductivity from texture. *J. Irrig. Drain. Eng.* **113**(2), 184–192 (1987)
- Baker, K.R.: *Optimization Modeling with Spreadsheets*. Thompson Brooks/Cole, Belmont, CA (2006)
- Bitterlich, S., Durner, W., Iden, S.C., Knabner, P.: Inverse estimation of the unsaturated soil hydraulic properties from column outflow experiments using free-form parameterizations. *Vadose Zone J.* **3**, 971–981 (2004)
- Bloemen, G.W.: Calculation of hydraulic conductivities from texture and organic matter content. *Z. Pflanz. Bodenkunde* **143**(5), 581–605 (1980)
- Boels, D., Gils, J.B.H.M. Van, Veerman, G.J., Wit, K.E.: Theory and systems of automatic determination of soil moisture characteristics and unsaturated hydraulic conductivities. *Soil Sci.* **126**(4), 191–199 (1978)
- Bouma, J., Denning, J.L.: Field measurement of unsaturated hydraulic conductivity by infiltration through gypsum crusts. *Soil Sci. Soc. Am. Proc.* **36**, 846–847 (1972)
- Bouma, J., Hille, D.I., Hole, F.D., Amerman, C.R.: Field measurement of unsaturated hydraulic conductivity by infiltration through artificial crusts. *Soil Sci. Soc. Am. Proc.* **35**, 362–364 (1971)
- Bresler, E., Russo, D., Miller, R.D.: Rapid estimate of unsaturated hydraulic conductivity function. *Soil Sci. Soc. Am. J.* **42**(1), 170–172 (1978)
- Bruce, R.R.: Hydraulic conductivity evaluation of the soil profile from soil water retention relations. *Soil Sci. Soc. Am. Proc.* **36**(4), 555–561 (1972)
- Bruckler, L., Renault, P., Aries, F.: Laboratory estimation of apparent soil thermal conductivity using a numerical approach. *Soil Sci.* **143**(6), 387–397 (1987)
- Cesari, L.: *Optimization – Theory and Applications*. Springer Verlag, New York, USA (1983)
- Elsenbeer, H.: Pedotransfer functions in hydrology. *J. Hydrol.* **251**, 121–122 (2001)
- Erh, K.T.: Application of spline functions to soil science. *Soil Sci.* **114**, 333–338 (1972)
- Fuentes, C., Haverkamp, R., Parlange, J.: Parameter constraints on closed-form soil water relationships. *J. Hydrol.* **134**(1–4), 117–142 (1992)
- Fuentes, C., Haverkamp, R., Parlange, J.-Y., Brutsaert, W., Zayani, K., Vachaud, G.: Constraints on parameters in three soil-water capillary retention equations. *Transport Porous Media* **6**(4), 445–449 (1991)
- Gill, P.E., Murray, W., Wright, M.H.: *Practical Optimization*. Academic Press, London, UK (1981)
- Halbertsma, J.M., Veerman G.J.: A new calculation procedure and simple set-up for the evaporation method to determine soil hydraulic functions. Report 88, DLO-Staring Centrum, Wageningen (1994)
- Hampton D.R. et al.: A method to fit soil hydraulic curves in models of subsurface flow in unsaturated soils. In: Gambolati, et al. (eds) *Computational methods in subsurface hydrology*. Computational mechanics Publ, Boston MA, pp 175–180
- Hansen, S., Jensen, H.E., Nielsen, N.E.: *DAISY: Soil Plant Atmosphere System Model*. Copenhagen, Denmark (1990)
- Heinen, M.: Dynamics of water and nutrients in closed, recirculating cropping systems in glasshouse horticulture. With special attention to lettuce grown in irrigated sand beds. Ph.D. Thesis, Wageningen Agricultural University (1997)
- Jhorar, R.K., Bastiaansen, W.G.M., Feddes, R.A., Dam, J.C. Van : Inversely estimation soil hydraulic functions using evapotranspiration fluxes. *J. Hydrol.* **258**(1–4), 198–213 (2001)
- Kastanek, F.J., Nielsen, D.R.: Description of soil water characteristics using cubic spline interpolation. *Soil Sci. Soc. Am. J.* **65**, 279–283 (2001)
- Kool, J.B., Parker, J.C.: Estimating soil hydraulic properties from transient flow experiments: SFIT user's guide. Report, Electric Power Research Institute, Palo Alto, California (1987)
- Kool, J.B., Parker, J.C.: Analysis of the inverse problem for transient unsaturated flow. *Water Resour. Res.* **24**(6), 817–830 (1988)
- Kool, J.B., Parker, J.C., Genuchten, M.Th. Van : Determining soil hydraulic properties from one-step outflow experiments by parameter estimation: I. Theory and numerical studies. *Soil Sci. Soc. Am. J.* **49**(6), 1348–1354 (1985)
- Kool, J.B., Parker, J.C., Genuchten, M.Th. Van : Parameter estimation for unsaturated flow and transport models – a review. *J. Hydrol.* **91**, 255–293 (1987)
- Leij, F.J., Russell, W.B., Lesch, S.M.: Closed-form expressions for water retention and conductivity data. *Ground Water* **35**(5), 848–858 (1997)
- Leong, E.C., Rahardjo, H.: Permeability functions for unsaturated soils. *J. Geotech. Geoenviron. Eng.* **123**(12), 1118–1126 (1997)

- Luenberger, D.G.: Introduction to Linear and Non-linear Programming. Addison-Wesley Publishing Company (1973)
- Marquardt, D.W.: An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* **11**(2), 431–441 (1963)
- Mertens, J.: Parameter estimation strategies in unsaturated zone modelling. Ph.D. Thesis, Leuven Catholic University (2003)
- Metselaar, K.: 'Auditing predictive models : a case study in crop growth'. Ph.D. thesis, Wageningen Agricultural University (1999)
- Minasny, B., McBratney, A.B., Bristow, K.L.: Comparison of different approaches to the development of pedotransfer functions for water-retention curves. *Geoderma* **93**(3–4), 225–253 (1999)
- Nelder, J.A., Mead, R.: A simplex method for function minimization. *Comput. J.* **7**, 308–313 (1965)
- Nemes, A.: Multi-scale hydraulic transfer functions for Hungarian soils. Ph.D. Thesis, Wageningen Agricultural University (2003)
- Nemes, A., Schaap, M.G., Leij, F.J., Wösten, J.H.M.: Description of the unsaturated soil hydraulic database UNSODA Version 2.0. *J. Hydrol.* **251**(3–4), 151–162 (2001)
- Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T.: Numerical Recipes: The Art of Scientific Computing. Cambridge University Press, Cambridge, UK (1986)
- Price, W.L.: A controlled random search procedure for global optimization. *Comput. J.* **4**, 367–370 (1979)
- Prunty, L., Casey, F.X.M.: Soil water retention curve description using a flexible smooth function. *Vadose Zone J.* **1**, 179–185 (2002)
- Russo, D., Zaidel, J., Laufer, A.: Numerical analysis of flow and transport in a three-dimensional partially saturated heterogeneous soil. *Water Resour. Res.* **34**(8), 1451–1468 (1998)
- Rust, B.W., Burrus, W.R.: Mathematical Programming and the Numerical Solution of Linear Equations. American Elsevier Publishing Company Inc., New York, USA (1972)
- Schaap, M.G., Leij, F.J.: Using neural networks to predict soil water retention and soil hydraulic conductivity. *Soil Till. Res.* **47**(1/2), 37–42 (1998)
- Schaap, M.G., Leij, F.J., Genuchten, M.Th.: ROSETTA: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *J. Hydrol.* **251**, 163–176 (2001)
- Schuh, W.M., Bauder, J.W.: Effect of soil properties on hydraulic conductivity-moisture relationships. *Soil Sci. Soc. Am. J.* **50**(4), 848–855 (1986)
- Simunek, J., Sejna, M., Van Genuchten, M.Th.: The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat and multiple solutes in variably-saturated media, version 2.0. Technical Report IGWMC-TPS-70, Int. Ground Water Model. Cent., Colo. Sch. of Mines, Golden (1998)
- Spriet, J.A., Vansteenkiste, G.C.: Computer-aided Modeling and Simulation. Academic Press, London, UK (1982)
- Stephens, D.B., Rehfeldt, K.R.: Evaluation of closed-form analytical models to calculate conductivity in a fine sand. *Soil Sci. Soc. Am. J.* **49**(1), 12–19 (1985)
- Stolte, J.: Manual for soil physical measurements: version 3. Technical Document 37, DLO-Staring Centrum, Wageningen (1997)
- Stolte, J., Veerman, G.J.: Description of the evaporation method, crust method and hanging water column method to determine the hydraulic conductivity and water retention characteristics in Dutch. Rapport 74, DLO-Staring Centrum, Wageningen (1990)
- Stolte, J., Wesseling, J.G., Verzandvoordt-Van Dijk, S.: Soil physical characteristics made available for a large group of users. in preparation (2007)
- Stolte, J., Wesseling, J.G., Wösten, J.H.M.: Pedotransfer functions for hydraulic and thermal properties of soil and the tool HERCULES. Report 126, DLO-Staring Centrum, Wageningen (1996)
- Stoof, C., Wesseling, J.G., Ritsema, C.J.: Effect of texture on hydrological behaviour of putting green rootzone mixes. *Agron. J.* (2007, submitted)
- Tarantola, A.: Inverse Problem Theory. Methods for Data Fitting and Model Parameter Estimation. Elsevier Science Publishers B.V., Amsterdam, The Netherlands (1987)
- van Alphen, B.J., Bootink, H.W.G., Bouma, J.: Combining pedotransfer functions with physical measurements to improve the estimation of soil hydraulic properties. *Geoderma* **103**(1/2), 133–147 (2001)
- Van Beek, P., Hendriks, Th.H.B.: Optimization Techniques: Principles and Applications (in Dutch). Bohn, Sheltema en Holkema, Utrecht, The Netherlands (1983)
- Van Dam, J.C., Huygen, J., Wessling, J.G., Feddes, R., Kabat, P., Van Walsum, P.E.V., Groenendijk, P., Van Diepen, C.A.: Theory of SWAP version 2.0: simulation of water flow, solute transport and plant growth in the Soil-Water-Atmosphere-Plant environment. Technical Document 45, DLO-Staring Centrum, Wageningen (1997a)
- Van Dam, J.C., Huygen, J., Wesseling, J.G., Feddes, R.A., Kabat, P., Van Walsum, P.E.V., Groenendijk, P.: SWAP Users Manual, Version 2.0. Simulation of flow and transport processes in the Soil-Water-Air-Plant

- environment. Report 71, Dept. of Water Resources, Wag. Agr. University, Wageningen, The Netherlands (1997b)
- Van Dam, J.C., Stricker, J.N.M., Droogers, P.: From one-step to multi-step: determination of soil hydraulic functions by outflow experiments. Report 7, Dept. of Water Resources, Wag. Agr. University, Wageningen, The Netherlands (1990)
- Van Genuchten, M.Th.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* **44**(3), 892–898 (1980)
- Van Genuchten, M.Th., Leij, F.J., Yates, S.R.: The RETC code for quantifying the hydraulic functions of unsaturated soils. Technical report, EPA, Ada (1991)
- Vogel, T.: SWM 2 – numerical model of two-dimensional flow in a variably saturated porous medium. Research Report 87, Research report/Agricultural University, Department of Hydraulics and Catchment Hydrology (1997)
- Vrugt, J.A., Bouten, W., Gupta, H.V.: Toward improved identifiability of hydrologic model parameters: the information content of experimental data. *Water Resour. Res.* **38**(12), 48/1–13 (2002)
- Wind, G.P.: Capillary conductivity data estimated by a simple method. Proc. unesco/iash, Water in the unsaturated zone. Symp. 1966. Wageningen, The Netherlands (1966)
- Wösten, J.H.M.: Description of the soil moisture retention and hydraulic conductivity characteristics from the Staringsseries with analytical functions (in Dutch). Rapport 2019, STIBOKA Wageningen (1987)
- Wösten, J.H.M., Lilly, A., Nemes, A., LeBas, C.: Using existing soil data to derive hydraulic parameters for simulation models in environmental studies and in land use planning. Report 156, DLOSC (1998)
- Wösten, J.H.M., Pachepsky, Ya.A., Rawls, W.J.: Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *J. Hydrol.* **251**, 123–150 (2001a)
- Wösten, J.H.M., Veerman, G.J., DeGroot, W.J.M., Stolte, J.: Soil moisture retention and hydraulic conductivity characteristics of top- and sub-soils in The Netherlands: the Staringsseries (in Dutch). Rapport, Alterra, Wageningen (2001b)
- Yeh, W.W.-G.: Review of parameter identification procedures in groundwater hydrology: the inverse problem. *Water Resour. Res.* **22**(2), 95–108 (1986)