General Introduction

Problem definition

One of the global concerns in the twenty-first century is food security (Greenland et al., 1997). The United Nations estimate a 65% increase in the world’s population between 1995 and 2050, with the largest growth mainly in the developing countries. The necessary increase in food production has to be supplied from higher yields through growing more crops and cultivating more land. This increased food production is required concurrently with an accelerating threat to land resources through land degradation and soil and water pollution.

Responses to such issues require *inter alia*, the generation of scenarios constructed with detailed biophysical data coupled to process-simulation models. Process-simulation modelling of soil-crop-water processes has become an important tool in agriculture and environmental management. 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. Verhagen (1997) demonstrated the successful application of such models. He used a coupled soil-water, heat and solute transport, and crop growth model to predict the nitrogen fertilizer management effect on nitrate leaching and potato production on a Dutch arable farm.

Process-simulation models have been widely developed, such as APSIM (Agricultural Production Systems iMulator) in Australia (McCown et al., 1996), which integrates biological, environmental, managerial and economic processes. But the use of these models is limited, as they need a large number of soil properties as input.

The essential inputs to drive a soil-water model are soil hydraulic properties. Soil hydraulic properties are physical properties that describe the soil-water relationship. The most important properties are the water retention and hydraulic conductivity characteristics. The water retention curve of a soil describes the relationship between soil-water potential and its volumetric water content, which is a characteristic for different soil types. It is used to predict the soil water storage, water supply to the
plants and soil aggregate stability (Collis-George and Figueroa, 1984). Due to the hysteretic effect of water filling and draining the pores, different wetting and drying curves may be distinguished. Hydraulic conductivity describes the ease of water flow in the soil. It is a constant of proportionality between the flux of water and its driving force (hydraulic gradient). In saturated conditions the saturated hydraulic conductivity reflects the number of pores and their arrangement. Hydraulic conductivity also depends on the water content and potential of the soil. If the soil is unsaturated, part of the pores with water in them will be empty, hence hydraulic conductivity decreases as water content decreases. The hydraulic conductivity – water content /potential relationship is also subject to hysteretic effect (Figure 1).

Figure 1. (a) The water retention curve, and (b) hydraulic conductivity characteristic both showing the effect of hysteresis.

Soil hydraulic properties are known to vary in space (Nielsen et al., 1973), hence to simulate realistic field conditions, a large number of samples is required. However, 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. Wösten (1990) postulated that the use of indirect methods is acceptable as long as the uncertainty of the estimates is provided.
Furthermore, the use of data with zero error is not necessary because many problems do not require exact solutions.

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 is generally known as a pedotransfer function (PTF) (Bouma, 1989). Indirect methods also encompass another solution known as the inverse method. This denotes the observation of transient soil-water flow, and estimating the hydraulic properties by back-matching the observed data with simulation results.

In response to an existing paucity of information, the United States Salinity Laboratory held two workshops addressing the use of indirect methods for predicting soil hydraulic properties in 1989 and 1997 (van Genuchten et al., 1992; van Genuchten et al., 1999). Many international papers were presented, offering different approaches and uses of pedotransfer functions and inverse methods. The unique environmental condition and distinctive soil properties of Australia (Williams, 1983) require extra consideration and the methods developed elsewhere may not always be suitable for immediate use.

**Efficient methods**

*Efficient* indirect methods for estimating soil hydraulic properties form the main thrust of this study. What do we mean by efficient methods? How may soil hydraulic properties be estimated efficiently? According to Chambers’ Dictionary (Macdonald, 1977), efficient means *capable of doing what may be required or effective*. The Merriam-Webster dictionary (1999) defined efficient as: (1) *being or involving the immediate agent in producing an effect*, or (2) *productive of desired effects; especially productive without waste*. The Infoplease Dictionary (2000) defined it as *performing or functioning in the best possible manner with the least waste of time and effort; having and using requisite knowledge, skill, and industry.*

The first two dictionary definitions do not encompass the whole the description of efficiency, thus further elaboration can be suggested. The Infoplease Dictionary definition suggests that efficiency relates to using all means of information to predict the hydraulic properties with the most certainty in order to minimise time and cost. Efficiency can therefore be associated with the trade-off between the uncertainty and
cost of measurements. It is the objective of this work to develop efficient methods and
to determine the relative efficiency of the methods for predicting soil hydraulic
properties.

**Aims**

The aims of the study presented here, are:

1. To develop pedotransfer functions for estimating water retention curves and
   hydraulic conductivity characteristics for Australian soil.
2. To evaluate the uncertainty of the predicted hydraulic properties from the
   pedotransfer functions.
3. To evaluate the use and efficiency of the disc permeameter in characterizing soil
   hydraulic properties in the field.
4. To develop an inverse solution to estimate soil hydraulic properties from disc
   permeameter measurements.
5. To compare the efficiency of hydraulic properties determined by different
   methods.
References


CAB International, Wallingford, UK.


I. Pedotransfer Functions, particularly hydraulic ones

‘You can’t make a silk purse out of a sow’s ear’

Old adage

1.1. Overview of pedotransfer functions

1.1.1. An introduction to pedotransfer functions

The concern on soil degradation and sustainable agriculture has prompted the development of the soil quality concept. Soil quality is defined as the capacity of a specific kind of soil to function, within natural or managed ecosystem boundaries, to sustain plant and animal productivity, maintain or enhance water and air quality, and support human health and habitation (Karlen et al., 1997). Soil quality may serve as a central link between agricultural practices and sustainability through the soil's function as a growth medium in food production, and is controlled by interactions of physical, chemical and biological properties.

Process-simulation models are useful in predicting the outcome of agricultural management on soil quality. The most difficult and expensive step towards the process of modelling is collection of data. Soil properties can be highly variable spatially and temporally and measuring these properties is time consuming and expensive. Depending on the resolution where the model is applied, the finer the resolution, the more data and measurements are needed. However, the soil properties needed as inputs are affected by other physical, chemical and biological properties, therefore it is possible to develop empirical relationships to predict them.

The data that are mostly available come from soil survey, but survey data only contain basic soil properties such as field morphology, texture, structure, and pH. Therefore it is essential to derive relationships that link the basic soil properties to the properties that are more difficult to measure. Bouma (1989) defined this concept with the term pedotransfer function (PTF), which he described as translating data we have into what we need, or predictive function of certain soil properties from other easily, routinely, or cheaply measured properties. Pedotransfer functions allow basic
information from soil surveys or geographic information systems (GIS) databases to be translated into other more laborious and expensively determined soil properties.

Most pedotransfer functions (PTFs) have been developed to predict soil hydraulic properties (water retention or hydraulic conductivity), nevertheless PTFs are not restricted to hydraulic properties. PTFs for estimating soil chemical, biological and mechanical properties have been developed (Table 1.1.1).

Figure 1.1.1. Pedotransfer functions translate soil survey data in other more expensively determined soil properties (after Edye and McBratney, 2000).

There are two different approaches in using pedotransfer functions for estimating soil quality indicators (Figure 1.1.2). The first approach is a static one, where pedotransfer functions are used to stipulate soil quality indicators. The second, dynamic approach, predicts those soil properties which will be used as inputs into a process-simulation model. This model predicts effects of agricultural management scenarios on soil quality.

Wösten et al. (1995) recognized two types of PTFs based on the amount of available information: class PTFs and continuous PTFs. Class PTFs predict certain soil properties based on the class to which the soil belongs. For example, prediction of available water content based on the soil texture class or soil groups. Continuous
PTFs predict certain soil properties as a continuous function of measured properties, for example prediction of field capacity from bulk density and clay content.

**Table 1.1.1. Example of pedotransfer functions**

<table>
<thead>
<tr>
<th>Predicted soil properties</th>
<th>Predicting properties</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Physical properties</strong></td>
<td>水 retained, hydraulic conductivity</td>
<td>Rawls and Brakensiek (1985)</td>
</tr>
<tr>
<td></td>
<td>infiltration rate after certain period</td>
<td>Canarache et al. (1968)</td>
</tr>
<tr>
<td></td>
<td>infiltration parameters</td>
<td>van de Genachte et al. (1996)</td>
</tr>
<tr>
<td></td>
<td>bulk density</td>
<td>Rawls (1983)</td>
</tr>
<tr>
<td></td>
<td>soil thermal conductivity</td>
<td>De Vries (1966)</td>
</tr>
<tr>
<td></td>
<td>time for initial water breakthrough</td>
<td>Booltink et al. (1993)</td>
</tr>
<tr>
<td></td>
<td>soil erodibility factor</td>
<td>Römkens et al. (1986)</td>
</tr>
<tr>
<td></td>
<td>soil mechanical resistance</td>
<td>da Silva and Kay (1997)</td>
</tr>
<tr>
<td></td>
<td>rate of structural change</td>
<td>Rasiah and Kay (1994)</td>
</tr>
<tr>
<td></td>
<td>volumetric shrinkage, liquid limit, plastic limit, plasticity index</td>
<td>Mbagwu and Abe (1998)</td>
</tr>
<tr>
<td></td>
<td>degree of overconsolidation</td>
<td>McBride and Joose (1996)</td>
</tr>
<tr>
<td><strong>Chemical properties</strong></td>
<td>cation exchange capacity (CEC)</td>
<td>Bell and van Keulen (1995)</td>
</tr>
<tr>
<td></td>
<td>Phosphorous (P) adsorption</td>
<td>Sheinost and Schwertmann (1995)</td>
</tr>
<tr>
<td></td>
<td>P saturation</td>
<td>Kleinman et al. (1999)</td>
</tr>
<tr>
<td></td>
<td>critical P level, P buffer coefficient</td>
<td>Cox (1994)</td>
</tr>
<tr>
<td></td>
<td>Cd sorption coefficient</td>
<td>Springob et al. (1998)</td>
</tr>
<tr>
<td></td>
<td>pH buffering capacity</td>
<td>Helyar et al. (1990)</td>
</tr>
<tr>
<td></td>
<td>K/Ca exchange</td>
<td>Scheinost et al. (1997a)</td>
</tr>
</tbody>
</table>
nitrogen-mineralization parameters
Al saturation
Haematite content
soil organic matter

CEC, total N, organic carbon, silt, clay
base saturation, organic carbon, pH
soil colour
soil colour

Rasiah (1995)
Jones (1984)
Torrent et al. (1983)
Fernandez et al. (1988)

Figure 1.1.2. (a) Static vs. (b) dynamic approaches using pedotransfer functions for predicting soil quality.
A special case of class PTFs is defined as pedotransfer rules (Batjes, 1996; 1997). These rules comprised of various ‘if-then’ conditions stratifying soil types and available information to arrive at a prediction. Batjes (1996) applied these rules to estimate available water content for the FAO world map. Batjes (1997) also derived the median of soil physical and chemical properties for 106 soil units of the FAO-UNESCO world soil map. Assimakopoulos et al. (1999) created pedotransfer rules for the purpose of classifying Greek agricultural soil in suitability classes for urea application.

The decision to use PTFs does not only depend on the amount of information we have, but also on the amount of effort needed to make the measurements, relating to the associated improvement of the results (Bouma, 1989). An example of the increasing level of information required to predict hydraulic properties is illustrated in Figure 1.1.3. The most basic information from a soil survey is field morphology and field texture. If we were to use only this information to predict hydraulic properties, the prediction would then be most uncertain. Predictions can be improved by extending the input data to properties such as particle-size distribution, bulk density, porosity (Rawls et al., 1982) or organic matter content (Vereecken et al., 1989). Additional improvements are possible by including one or more water retention data points in the models (Rawls et al., 1982; Schaap et al., 1998). With increasing laboratory measurements, the prediction is more accurate, but the cost of obtaining the additional information is also higher. Thus, there is a balance between the cost and quality of prediction. Hierarchical approaches, where PTFs are based on the level of available information, have been introduced (Schaap et al., 1998) to permit more flexibility in using the available data.

<table>
<thead>
<tr>
<th>Field Morphology</th>
<th>Field texture</th>
<th>Particle size analysis</th>
<th>Organic C</th>
<th>ρ₀</th>
<th>ϕ</th>
<th>θ₀</th>
<th>θ₀.33</th>
<th>θ₀.1500</th>
<th>θ(h)</th>
<th>Kₛ</th>
<th>K(θ)</th>
</tr>
</thead>
</table>

**Increasing laboratory measurements & cost**

**Improving predictions, decreasing uncertainty**

**Decreasing use of PTFs**

Figure 1.1.3. The level of available information for hydraulic PTFs.
Bouma and Hoosbeek (1996) further differentiate methods of estimating soil properties based on the knowledge level and complexity of the models. The five levels (Figure 1.1.4) distinguish the approaches in modelling: qualitative and quantitative on the horizontal axis and empirical and mechanistic on the vertical axis. The levels are user’s expertise (level K1), expert knowledge (level K2), simple models (level K3), complicated, comprehensive models (level K4), and complicated, detailed models (level K5).

Bouma and Droogers (1999) illustrated this classification in estimating soil moisture supply capacity to plants (MSC). Level K1 uses the farmers’ experience on soil condition in the field. Level K2 uses the estimation of available water content from pedotransfer functions. Level K3 uses simple empirical dynamic simulation, such as the tipping bucket model (Burns, 1974). Level K4 uses complex mechanistic model, which solves the Richards’ equation, coupled with a crop-growth model. Level K5 is the extension of K4 with detailed sub-models, which take into account the occurrence of macropores, accessibility of water in peds by plant roots. They found that the Level K2 could indicate the soil’s relative water supply capacity, but cannot measure MSC. Level K3 did not give a realistic value because of the simplification of the soil-water system. Level K4, which assumed a homogenous and isotropic soil, overestimate soil water content. Realistic estimates can only be obtained by level K5. They further conclude that, important processes such as bypass flow, internal catchment and water accessibility for roots are more important than hydraulic properties.

Figure 1.1.4. Knowledge levels to predict of soil properties (Bouma and Droogers, 1999).
A new PTF is established by first characterizing basic and specific soil properties. Using these data, which potentially contribute to a soil database, statistical relationships relating basic to specific soil properties are derived. Once the PTF is formulated, an existing soil survey map can be used to extrapolate the specific soil properties. This framework was proposed by Wösten and Nugroho (1993) for the establishment of a regional soil hydraulic map in Indonesia.

Developing new PTFs is an arduous task, as it requires a large soil database containing many soil measurements. It will be wise, in the first instance, to utilise the functions that have already been developed. But the validity of a given PTF should not be interpolated or extrapolated beyond the pedological origin or soil type on which they are developed. The distinct properties of Australian soil (Williams, 1983) means that PTFs developed elsewhere cannot be directly applied without testing. Testing is required so that the most suitable PTFs to use can be identified. Stratification and calibration of PTFs are also essential.

Stratification is needed to establish separate PTFs based on soil type and input information. Stratification has been made variously according to soil horizons (Hall et al., 1977; Lamp and Kneib, 1981); genetic horizons (Pachepsky et al., 1992); hydraulic-functional horizons (Wösten et al., 1986); numerical soil classification (Williams et al., 1983); FAO soil classes (Batjes, 1996), textural classes (Tietje and Hennings, 1996); parent material and horizon morphology/lithomorphic (Franzmeier, 1991); management units/phenof orm (Droogers and Bouma, 1997); great soil groups, temperature regime, moisture regime (Pachepsky and Rawls, 1999).

Calibration is required to adjust properties measured to the one required as input. Calibration is needed because of the differences in criteria and measurements from existing pedotransfer functions. For example, sand fractions are different according to the ISSS/Australian classification (particle diameter 20-2000µm) and the FAO/USDA criteria (particle diameter 50-2000µm). Other calibrations are listed in Table 1.1.2.

The use of PTFs is not limited to predicting soil properties. PTFs can be utilised to evaluate the changes in soil porosity caused by different types of tillage (Rawls et al., 1983), effect of crusting on infiltration parameters (Brakensiek and Rawls, 1983), sensitivity of soil structure due to changes in organic carbon content (Kay et al., 1997). Further applications include assessing regional water flow and pesticide leaching (Petach et al., 1991), predicting pesticide leaching to groundwater on a
watershed (de Jong and Reynolds, 1995), developing regional groundwater vulnerability maps that indicate the impact of leaching (Soutter and Pannatier, 1996), assessing the magnitude of cadmium accumulation on a regional scale (Titak et al., 1999), estimating soybean yields (Timlin et al., 1996), estimating regional crop yields (Haskett et al., 1996), and evaluating crop yield and nitrate leaching as a result of different management practices (Droogers and Bouma, 1997).

Table 1.1.2. Calibration of measurement techniques.

<table>
<thead>
<tr>
<th>Measured properties</th>
<th>Predicted properties</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silt fraction</td>
<td>Silt Fraction</td>
<td>Minasny et al. (1999)</td>
</tr>
<tr>
<td>ISSS (2-20mm)</td>
<td>FAO/USDA (2-50 mm)</td>
<td></td>
</tr>
<tr>
<td>Oven-dried Bulk density</td>
<td>Bulk density at field capacity (-33kPa)</td>
<td>Kern (1995)</td>
</tr>
<tr>
<td>Organic carbon</td>
<td>Total carbon</td>
<td>Merry and Spouncer (1988)</td>
</tr>
<tr>
<td>(Walkley and Black method)</td>
<td>(Leco dry combustion)</td>
<td></td>
</tr>
<tr>
<td>pH in water</td>
<td>pH in CaCl₂</td>
<td>Little (1992)</td>
</tr>
</tbody>
</table>

1.1.2. *A brief history of the hydraulic pedotransfer functions*

Although not formally named until 1989, the concept of pedotransfer function has long been applied to estimate soil properties that are difficult to determine. Probably because of the difficulty and cost of measurement, the most comprehensive research in developing PTFs has been for the estimation of water retention. The first attempt to use such predictions came from the study of Briggs and McLane (1907), later refined by Briggs and Shantz (1912). They determined the wilting coefficient, which is defined as *percentage water content of a soil when the plants growing in that soil are first reduced to a wilted condition from which they cannot recover in an approximately saturated atmosphere without the addition of water to the soil*, as a function of particle-size:

\[
\text{Wilting coefficient} = \frac{0.01 \text{sand} + 0.12 \text{silt} + 0.57 \text{clay}}{1 \pm 0.025}. \quad [1.1.1]
\]

In their studies, the sand fraction is defined as particles within diameter 50-2000 µm, silt: 5 – 50 µm, and clay < 5µm.
With the introduction of the field capacity (FC) and permanent wilting point (PWP) concepts by Veihmeyer and Hendricksen (1927), research in the period 1950-1980 attempted to correlate particle size distribution, bulk density and organic matter with water content at FC(-33 kPa), PWP(-1500 kPa), and available water content (AWC). From a study in North Queensland, Stirk (1957) suggested an estimate of water content at -1500kPa for soil up to 60% clay as:

\[ PWP = \frac{2}{5} \text{ clay}. \]  \[1.1.2\]

Burrows and Kirkham (1958) estimate field capacity as a function of clay fraction and bulk density \( (\rho_b \text{ Mg m}^{-3}) \) from 4 soil types:

\[ \text{FC} = 6.69 + 0.637 \text{ clay} + 1.67 \rho_b \quad (R = 0.932). \]  \[1.1.3\]

Nielsen and Shaw (1958) presented a parabolic relationship between clay content and PWP from 730 Iowa Soils.

In the 1960’s various papers dealt with the estimation of FC, PWP, and AWC, particularly in a series of papers by Salter and Williams (1965, 1966, 1967, 1969). They explored relationships between texture classes and available water capacity, which are now known as class PTFs. They also developed functions relating particle-size distribution to AWC, now known as continuous PTFs. They asserted that their functions could predict AWC to a mean accuracy of 16%.

In the 1970’s more comprehensive research using large databases were developed. Of particular is the study of Hall et al. (1977) from soil in England and
Wales, they established field capacity, permanent wilting point, available water content, air capacity as a function of textural class, and also derived continuous functions estimating these soil-water properties. In the USA, Gupta and Larson (1979) developed 12 functions relating particle-size distribution and organic matter to water content at potentials ranging from -4 kPa to -1500 kPa.

With the flourishing development of hydraulic models (van Genuchten, 1980) and computer modelling of soil-water and solute transport (De Wit and van Keulen, 1975), the need for hydraulic properties as input to these models becomes more evident. Clapp and Hornberger (1978) derived the average values for the parameters of a power function water retention curve, sorptivity and \( K_s \) for different texture classes.

In probably the first research of its kind, Bloemen (1980) derived the relationship between parameters of hydraulic models and particle size distribution. He derived the Brooks-Corey parameters:

\[
\begin{align*}
    h_b &= 2914 \ f^{-0.79} \ M_d^{0.96} \text{ (cm)} \\
    \lambda &= 1.512(\exp(0.3f)-1) \\
    K_s &= 0.2 \ M_d^{1.93} \ f^{-0.74} \text{ (mm/d)}
\end{align*}
\]  

[1.1.4]

where \( f \) is the grain size distribution index and \( M_d \) is the median grain size (mm).

Lamp and Kneib (1981) formally introduced the term ‘pedofunction’, while Bouma and van Lanen (1987) used the term ‘transfer function’. To avoid confusion with the terminology ‘transfer function’ used in soil physics and many other disciplines with many other meanings, Bouma (1989) later called it a ‘pedotransfer function’.

Since then, development of hydraulic PTFs has become a boom industry mostly in the US and Europe. Results of the research have been reported in the USA (Rawls et al., 1982), UK (Mayr and Jarvis, 1999), Canada (de Jong et al., 1983), Netherlands (Wösten et al., 1995), Belgia (Vereecken et al., 1988), Germany (Scheinost et al., 1997), Portugal (Goncalves et al., 1997), and Sweden (Rajkaj et al., 1996).

In Australia, the first comprehensive attempt appears to come from the study of Williams et al. (1983). They classified water retention based on textural class, and provide average parameter of the power function model for the water-retention curve based on texture class. Further studies by Williams et al. (1992) estimated Campbell’s parameters from texture and structure information.
More recently, many hydraulic PTFs have been developed in Australia, such as those by Cresswell and Paydar (1996), McKenzie and Jacquier (1997), Smettem and Gregory (1994), Smettem et al., (1999), and Smettem and Bristow (1999).

1.2. Water retention curve PTFs

1.2.1. Type of water retention PTFs

Soil hydraulic properties are physical properties that describe the soil-water relationship. The most important properties are the water retention $\theta(h)$ and hydraulic conductivity $K(h)$ curves. The first use of these properties is to estimate the available water for plants. Developments in computer modelling of water and solute transport in soil is advancing rapidly, as speed of computation increases and complexity of models expand. The models are used to describe environmental problems. With this advance, the need for $\theta(h)$ and $K(h)$ relationships as an input to the models is becoming more apparent. Direct measurements of these properties are time consuming and expensive (Klute, 1986; Dirksen, 1991). Since hydraulic properties are affected by other physical properties (Table 1.2.1), it is possible to develop empirical relationships (PTFs) to predict them.

Table 1.2.1. Properties that affect water retention in soil (Rawls et al., 1991).

<table>
<thead>
<tr>
<th>Particle size properties</th>
<th>Hydraulic properties</th>
<th>Morphological properties</th>
<th>Chemical properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sand</td>
<td>-10 kPa water retention</td>
<td>Bulk density</td>
<td>Organic Carbon/organic matter</td>
</tr>
<tr>
<td>Silt</td>
<td>-33 kPa water retention</td>
<td>Porosity</td>
<td>CEC</td>
</tr>
<tr>
<td>Clay</td>
<td>-1500 kPa water retention</td>
<td>Horizon</td>
<td>Sodium adsorption ratio</td>
</tr>
<tr>
<td>Fine sand</td>
<td></td>
<td>Structure</td>
<td>CaCO$_3$</td>
</tr>
<tr>
<td>Very coarse sand</td>
<td></td>
<td>Field Texture</td>
<td>Fe</td>
</tr>
<tr>
<td>Coarse sand</td>
<td></td>
<td>Color</td>
<td>Al</td>
</tr>
<tr>
<td>Coarse fragments</td>
<td></td>
<td>Clay type</td>
<td></td>
</tr>
<tr>
<td>Particle size distribution</td>
<td></td>
<td>Consistence</td>
<td></td>
</tr>
<tr>
<td>Median particle size</td>
<td></td>
<td>Pedality</td>
<td></td>
</tr>
<tr>
<td>Geometric mean particle size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard deviation of mean particle size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water stable aggregates</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface area</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The predicted hydraulic properties can be used to evaluate the effect of agricultural management on crop production. This has been demonstrated by Smettem et al. (1999) where the hydraulic properties estimated by PTFs were used in a mechanistic model for spatial modelling of water balance to predict wheat yield.

Pedotransfer functions for predicting the water-retention curve can be divided into 3 types: (1) point estimation, (2) parametric PTFs, and (3) physico-empirical models.

Point estimation PTFs are empirical functions that predict the water retention at a pre-defined potential. The most frequently estimated $\theta$ are at -10, -33 kPa (corresponding to field capacity) and at -1500kPa (corresponding to permanent wilting point), which is commonly measured to predict available water content.

In parametric pedotransfer functions, the assumption was made that $\theta(h)$ and $K(h)$ relationships can be described adequately by a hydraulic model that is a closed-form equation with a certain number of parameters, e.g. Brooks and Corey (1964); Campbell (1974) and van Genuchten (1980) (Table 1.2.2).

Rawls et al. (1982) listed the average parameter values for the Brooks-Corey model and subsequently Rawls and Brakensiek (1985) derived the parameters as function of clay, sand, and porosity for soil in the US. Schaap et al. (1998) derived the average parameters for the van Genuchten hydraulic models based on the US textural classes (Table 1.2.3).

A parametric approach is usually preferred as it yields a continuous function of $\theta(h)$ relationship. Water retention at any potentials can be estimated and it also ensures that the water content predicted at lower potential will be smaller than the one at higher potential. The estimated parameters can be used to predict the unsaturated hydraulic conductivity based on the hydraulic models (Mualem, 1976). Soil-water transport models usually only require the parameters of the hydraulic functions, thus the predicted parameters can be used directly in the models.
### Table 1.2.2. Hydraulic properties models

<table>
<thead>
<tr>
<th>Model</th>
<th>Function</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brooks-Corey</td>
<td>[ \theta(h) = \begin{cases} \theta_s &amp; \text{if } h \geq h_b \ \theta_s + (\theta_s - \theta_r) \left( \frac{h_b}{h} \right)^{\lambda} &amp; \text{if } h &lt; h_b \end{cases} ]</td>
<td></td>
</tr>
<tr>
<td>(1964)</td>
<td></td>
<td>( \theta_r ) = residual water content ( \theta_s ) = saturated water content ( h_b ) = air entry potential ( \lambda ) = pore size index ( S_e ) = degree of saturation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( S_e = \frac{\theta - \theta_r}{\theta_s - \theta_r} )</td>
</tr>
<tr>
<td></td>
<td>( K(S_e) = \begin{cases} K_s &amp; \text{if } h \geq h_b \ K_s S_e^{2b+3} &amp; \text{if } h &lt; h_b \end{cases} )</td>
<td></td>
</tr>
<tr>
<td>Campbell</td>
<td>[ \theta(h) = \begin{cases} \theta_s &amp; \text{if } h \geq h_b \ \left( \frac{h_b}{h} \right)^{1/b} &amp; \text{if } h &lt; h_b \end{cases} ]</td>
<td></td>
</tr>
<tr>
<td>(1974)</td>
<td></td>
<td>( b = ) fractal coefficient</td>
</tr>
<tr>
<td></td>
<td>[ K(\theta) = \begin{cases} K_s &amp; \text{if } h \geq h_b \ K_s \left( \frac{\theta}{\theta_s} \right)^{2b+3} &amp; \text{if } h &lt; h_b \end{cases} )</td>
<td></td>
</tr>
<tr>
<td>van Genuchten</td>
<td>[ \theta(h) = \theta_r + \frac{\theta_s - \theta_r}{\left(1 + (\alpha h)^n \right)^m} ]</td>
<td></td>
</tr>
<tr>
<td>(1980)</td>
<td></td>
<td>( \alpha = ) scaling parameter ( n = ) curve shape parameter ( m = 1 - 1/n ) ( l = ) empirical parameter describing pore tortuosity (usually = 0.5)</td>
</tr>
<tr>
<td></td>
<td>[ K(S_e) = K_s S_e^{l} \left[1 - (1 - S_e^{1/m})^m \right]^l ]</td>
<td></td>
</tr>
</tbody>
</table>
Table 1.2.3. Average van Genuchten parameters according to US textural classes (Schaap et al., 1998).

<table>
<thead>
<tr>
<th>Texture Class</th>
<th>N</th>
<th>$\theta_i$ (m$^3$ m$^{-3}$) $\mu$</th>
<th>$\sigma$</th>
<th>$\theta_s$ (m$^3$ m$^{-3}$) $\mu$</th>
<th>$\sigma$</th>
<th>$\alpha$ (m$^{-1}$) $\mu$</th>
<th>$\sigma$</th>
<th>n $\mu$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clay</td>
<td>84</td>
<td>0.098 0.107</td>
<td>0.459</td>
<td>0.079 2.55</td>
<td>0.19</td>
<td>1.26</td>
<td>0.041</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C Loam</td>
<td>140</td>
<td>0.079 0.076</td>
<td>0.442</td>
<td>0.079 2.74</td>
<td>0.22</td>
<td>1.44</td>
<td>0.018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loam</td>
<td>242</td>
<td>0.061 0.073</td>
<td>0.399</td>
<td>0.098 2.05</td>
<td>0.14</td>
<td>1.50</td>
<td>0.013</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Sand</td>
<td>201</td>
<td>0.049 0.042</td>
<td>0.39</td>
<td>0.07 4.48</td>
<td>0.33</td>
<td>1.80</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sand</td>
<td>308</td>
<td>0.053 0.029</td>
<td>0.375</td>
<td>0.055 3.79</td>
<td>0.17</td>
<td>3.30</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S Clay</td>
<td>11</td>
<td>0.117 0.114</td>
<td>0.385</td>
<td>0.046 4.86</td>
<td>0.50</td>
<td>1.21</td>
<td>0.050</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S C Loam</td>
<td>87</td>
<td>0.063 0.078</td>
<td>0.384</td>
<td>0.061 3.77</td>
<td>0.45</td>
<td>1.35</td>
<td>0.028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S Loam</td>
<td>476</td>
<td>0.039 0.054</td>
<td>0.387</td>
<td>0.085 3.83</td>
<td>0.30</td>
<td>1.47</td>
<td>0.015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silt</td>
<td>6</td>
<td>0.05 0.041</td>
<td>0.489</td>
<td>0.078 0.73</td>
<td>0.01</td>
<td>1.71</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Si Clay</td>
<td>28</td>
<td>0.111 0.119</td>
<td>0.481</td>
<td>0.08 2.60</td>
<td>0.17</td>
<td>1.34</td>
<td>0.030</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Si C Loam</td>
<td>172</td>
<td>0.09 0.082</td>
<td>0.482</td>
<td>0.086 1.25</td>
<td>0.03</td>
<td>1.55</td>
<td>0.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Si Loam</td>
<td>330</td>
<td>0.065 0.073</td>
<td>0.439</td>
<td>0.093 0.74</td>
<td>0.01</td>
<td>1.70</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Texture Class</th>
<th>$K_s$ (mm h$^{-1}$)* $\mu$</th>
<th>$\sigma$</th>
<th>l $\mu$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clay</td>
<td>1.24</td>
<td>0.76</td>
<td>-1.561</td>
<td>1.39</td>
</tr>
<tr>
<td>C Loam</td>
<td>2.08</td>
<td>0.71</td>
<td>-0.763</td>
<td>0.90</td>
</tr>
<tr>
<td>Loam</td>
<td>1.54</td>
<td>0.68</td>
<td>-0.371</td>
<td>0.84</td>
</tr>
<tr>
<td>L Sand</td>
<td>10.13</td>
<td>0.72</td>
<td>-0.874</td>
<td>0.59</td>
</tr>
<tr>
<td>Sand</td>
<td>10.20</td>
<td>0.72</td>
<td>-0.930</td>
<td>0.49</td>
</tr>
<tr>
<td>S Clay</td>
<td>1.81</td>
<td>0.91</td>
<td>-3.665</td>
<td>1.80</td>
</tr>
<tr>
<td>S C Loam</td>
<td>2.89</td>
<td>0.72</td>
<td>-1.28</td>
<td>0.99</td>
</tr>
<tr>
<td>S Loam</td>
<td>6.45</td>
<td>0.68</td>
<td>-0.861</td>
<td>0.73</td>
</tr>
<tr>
<td>Silt</td>
<td>1.39</td>
<td>0.87</td>
<td>0.624</td>
<td>1.57</td>
</tr>
<tr>
<td>Si Clay</td>
<td>1.32</td>
<td>0.78</td>
<td>-1.287</td>
<td>1.23</td>
</tr>
<tr>
<td>Si C Loam</td>
<td>0.93</td>
<td>0.76</td>
<td>-0.156</td>
<td>1.23</td>
</tr>
<tr>
<td>Si Loam</td>
<td>0.73</td>
<td>0.76</td>
<td>0.365</td>
<td>1.42</td>
</tr>
</tbody>
</table>

* $K_s$ for Mualem-van Genuchten model

In a physico-empirical approach, hydraulic properties are derived based on physical attributes (Figure 1.2.1). In water retention curve modelling, Arya and Paris (1981) translated the particle-size distribution into a water retention curve by converting solid mass fractions to water content, and pore-size distribution into hydraulic potential from by means of capillary equation. The main obstacle is the need to predict a parameter that characterizes the packing of the soil particles. Tyler and Wheatecraft (1989) proposed the packing (scaling) parameter as a fractal dimension. Arya et al. (1999a) improved the model by incorporating textural dependency on the scaling parameter. Furthermore the model was extended to predict the $K(\theta)$ relationship via the Hagen-Pouiselle’s law (Arya et al., 1999b).
Figure 1.2.1. Physico-empirical model translates particle-size distribution data into water retention curve.

Other physico-empirical models include prediction of the water retention curve from fractal models of soil structure (Anderson and McBratney, 1995; Bird et al., 2000)

With the large number of PTFs published, care must be taken when using them to predict the water retention, Bastet et al. (1999) deduced that:

- the performance of published PTFs when applied to other soil is highly variable, ranging from acceptable to extremely bad;
- the performance varied according to the pedological origin of the soil on which they are tested, consequently the validity of a given PTF should not be considered as general;
- the difference in performance is not related to the number or kind of variables used as predictors;
- the performance of the PTFs is better for water retention at low matric potentials, which is largely due to the influence of soil texture on the distribution of the soil microporosity.
1.2.2. Methods to fit water-retention PTFs

1.2.2.1. Multiple linear regression

Different methods are utilised to derive the empirical relationship for PTFs. The most common method used in point estimation PTFs is to employ multiple linear regression. For example:

\[ \theta_p = c_1 \text{sand} + c_2 \text{silt} + c_3 \text{clay} + c_4 \text{organic matter} + c_5 \text{bulk density} \]  \[1.2.1\]

where \( \theta_p \) is the water content (m\(^3\)m\(^{-3}\)) at potential p and \( c_1, c_2, c_3, c_4, c_5 \) are regression coefficients (Gupta and Larson, 1979).

Multiple linear regression is also used in parametric PTFs. Parameters of the hydraulic model are first estimated by fitting the model to water-retention data. Then empirical relationships between basic soil properties and parameters are formed.

Williams et al. (1992) assumed a power-function model for the water-retention curve and developed PTFs from texture, structure classes (massive or structured) and organic matter. The prediction variables were inserted into the water-retention model and the parameters were estimated using multiple linear regression. Their model is:

\[ \ln \theta(h) = \left[ a_0 + a_1 v_1 + a_2 v_2 + ... \right] + \left[ c_0 + c_1 v_1 + c_2 v_2 + ... \right] \ln h \]  \[1.2.2\]

where \( a, c \) are parameters and \( v \) are the predictor variables.

The hydraulic model parameters sometimes exhibit skewed distributions and logarithmic transformations are usually made on the parameters e.g. \( h_b \) and \( \lambda \) for Brooks-Corey model (Rawls and Brakensiek, 1985). Wösten et al. (1995) applied the transformations: \( \ln(\alpha) \) and \( \ln(n - 1) \) to the van Genuchten parameters to normalize their distribution. Transformation is also needed in the case of interdependency amongst the variables. Vereecken et al. (1989) transformed the particle-size distribution into principal factors in order to reduce the linear dependency among the variables.

1.2.2.2. Extended nonlinear regression

A problem in parametric PTF is the interdependency amongst the hydraulic model parameters. To overcome this problem, van den Berg et al. (1997) suggested the following approach:

1) fit the model to observed data,
2) apply regression analysis to one of the parameters,
Scheinost et al. (1997b) reported difficulties in estimating the scaling and shape parameters $\alpha, n$ of the van Genuchten equation using the multiple linear regression approach. Realizing the over-parametization (too many adjustable parameters relative to number of data points) of the van Genuchten equation, they proposed the following approach:

1. set-up the expected relationship between the parameters of the hydraulic model and soil properties,
2. insert the relationship into the model and estimate the parameters of the relationship by fitting the extended model using nonlinear regression.

### 1.2.2.3. Artificial neural networks

A more recent approach for fitting PTFs is to use artificial neural networks (ANN) (Tamari et al., 1996; Pachepsky et al., 1996; Schaap and Bouten, 1996). Neural networks attempt to build a mathematical model that supposedly works in an analogous way to human brain. ANN is simply a sophisticated regression, which has a network of many simple elements or processors or ‘neurons’. The elements are connected by communication channels or ‘connectors’ which usually carry numeric data, encoded by a variety of means, and often organized into subgroups or layers. A neural network can perform a particular function when certain values are assigned to the connections or ‘weights’ between elements. To describe a system, there is no assumed structure of the model, instead the networks are adjusted, or ‘trained’, so that a particular input leads to a specific target output, which is called supervised learning (Demuth and Beale, 1998). Tamari and Wösten (1999) gave a review of ANN methods and their application to hydraulic PTFs.

The mathematical model of ANN comprises of a set of simple functions linked together by weights $w$ (Figure 1.2.2). The network consists of a set of input units, a set of output units and a set of hidden units, which link the inputs to outputs. The hidden units extract useful information from input units and use them to predict the output units. The type of ANN model considered here is called the ‘multilayer perceptron’ (MLP).
If a network has a single input $x$, then it is transmitted through a connection that is multiplied by weight $w$ to form $wx$ (Figure 1.2.2a). The weighted input and a bias of one represent the argument to the transfer function $F$ to produce estimated output $\hat{y}$. A bias is simply a weight with constant input of 1 that also functions as a constant added to the weight. The model can be written as:

$$\hat{y} = F(wx + w_0)$$  \[1.2.3\]

Figure 1.2.2. Structure of a neural network.

Similar principles applies to MLP with multiple inputs $x$ and outputs $y$. If the MLP has only one hidden layer (Figure 2.7b) then the model can be represented as:

$$\hat{y}_i(w,W) = F_i \left[ \sum_{j=1}^{nh} W_{ij} f_j \left( \sum_{l=1}^{ni} W_{lj} x_l + w_{j0} \right) + w_{i0} \right]$$  \[1.2.4\]

where $x$ is the input units; $\hat{y}$ is the output units; $w$ and $W$ are the elements of the weight matrices for the input and hidden layer; $f$ and $F$ are the activation or transfer functions; $nh$ is the number of hidden units and $ni$ is the number of input units.

The weights are the adjustable parameters of the network and are determined from a set of data through the process of training (Nørgaard, 1997). The objective of the training is to minimize the residual sum of squares between the measured and predicted output:
\[ O(w, W, x) = \sum_{i=1}^{N} (\hat{y}(i) - \hat{\hat{y}}(i))^2 \]  

[1.2.5]

The objective function is minimized through an iterative process, e.g. the Levenberg-Marquardt algorithm.

Pachepsky et al. (1996) used ANN to estimate water content at 8 water potentials and also van Genuchten parameters from particle-size distribution and bulk density of 230 soil data. They found that for point estimation PTFs, ANN was better than the regression method, but for parametric PTFs the performance of both approaches was comparable. Schaap et al. (1998) estimated van Genuchten parameters for 1209 soil samples from the US using ANN. They distinguished their PTFs based on available information level: texture class; clay, silt and sand; texture + bulk density; texture + bulk density + measured \( \theta \) at -33 and -1500 kPa. They found that ANN performed better than four published PTFs and accuracy of prediction generally increased if more input data are used, but there was always a considerable difference between predicted and measured values. Koekkoek and Booltink (1999) applied similar approach to estimate water retention at different potentials from Dutch and Scottish databases. They found that ANN performed somewhat better than PTFs of Gupta and Larson (1979), but the improvements were not significant.

1.2.2.4. Other methods

There is room for different approaches to derive water retention PTFs. Various other approaches has been adopted, e.g. logarithmic interpolation. McQueen and Miller (1974) presented a method to derive water retention on a log potential \((pF = \log_{10}[\text{-}h/cm])\) versus \( \theta \) graph. They recognized three segments in the curve: (1) capillary segment (from saturation to -3.3 kPa); (2) absorbed film segment (-3.3 to -1000 m); and (3) tightly absorbed segment (-1000 to –10000 m). Using only one point, the whole curve can be generated (Figure 1.2.3).

Ahuja et al. (1985) introduced the log-log interpolation/extrapolation method, which assumes a power function or Campbell (1974) model for water retention lower than air-entry value \( h_b \). By fitting a straight line through \( \theta \) at -3.3 and -1500 kPa on a log-log scale, the water retention at other potentials can be interpolated or even extrapolated. Cresswell and Paydar (1996) used this approach to predict Campbell’s parameters.
Gregson et al. (1987) and Ahuja and Williams (1991) further modified this relationship into a one-parameter model:

$$\ln |h| = a + b \ln (\theta - \theta_r),$$  \[1.2.6\]

where $a$ and $b$ has a linear relationship:

$$a = p + q b.$$  \[1.2.7\]

Parameters $p$ and $q$ are related to textural group and Ahuja and Williams (1991) provided the mean values for different soil textures. From soil textural information and one measured $(\theta, h)$ point, the $h(\theta)$ relationship can be constructed.

Various types of modern regression techniques have been developed in recent years and can be used to develop PTFs. Alternative methods can include generalized additive models (Hastie, 1992) or regression trees (Clark and Pregibon, 1992), which have been implemented in statistical program such as S-Plus (Mathsoft, 1999). McKenzie and Jacquier (1997) used a regression tree model to predict $K_s$ from soil morphology data.

Pachepsky et al. (1998) and Pachepsky and Rawls (1999) applied the so-called group method of data handling (GMDH) (Farrow, 1984) to estimate water retention curve. The GMDH, which is generally used for data-mining purposes, provides an automated selection of essential input variables and builds hierarchical polynomial
regressions of necessary complexity. They found that GMDH is more useful than ANN as it automates finding essential input variables to be included in PTFs and presents an explicit form of the equations.

1.3. Hydraulic conductivity PTFs

1.3.1. Development of hydraulic conductivity PTFs

Saturated hydraulic conductivity \((K_s)\) is an important physical factor determining water and solute transport in soil. Knowledge of \(K_s\) is useful in managing irrigation and drainage problems and in environmental planning.

Early research attempted to find relationships between soil properties and \(K_s\). In civil and petroleum engineering much research has been focused on obtaining relationships between grain-size distribution and \(K_s\), such as Hazen (1893) and Krumbein and Monk (1942). In soil science, Baver (1938) proposed \(K_s\) as a function of the non-capillary porosity: \(K_s = f(\text{porosity factor})\), where the porosity factor is non-capillary porosity at the air-entry potential divided by the air-entry potential. Aronovici (1946) related \(K_s\) to soil texture from lighter-textured sediments of the Imperial Valley, California, while O'Neal (1949) indicated that \(K_s\) is significantly influenced by soil structure.

Because of the supposed distinctive characteristics of Australian soil, soil hydraulic properties may well be different (Williams, 1983). Talsma and Flint (1958) found a curvilinear relationship between clay or (clay + silt, or particles < 20 µm) and \(K_s\) of subsoil in the Murrumbidgee Irrigation Area of New South Wales. They also derived relationships between \(K_s\) at different depths and percentage of (clay + silt) and used it for tile-drainage problems. They found no relationship between \(K_s\) and pH, electrical conductivity, and total chloride of subsoil samples. Loveday and Pyle (1973) found a correlation between the dispersion index and \(K_s\) for soil samples from the Riverina and Namoi regions in New South Wales.

These findings clearly demonstrated the strong relationship between \(K_s\) and particle- and pore-size distribution and suggest a possible way of estimating \(K_s\) via PTFs. Indeed many PTFs have been developed in Australia. Attempts have been made to estimate \(K_s\) from field morphology (McKenzie et al., 1991; McKenzie and Jacquier, 1997), and basic soil properties (Bui et al., 1996; Bristow et al., 1999; Smettem and Bristow, 1999).
Pedotransfer functions for predicting $K_s$ can be divided into 2 categories: empirical and physico-empirical relationships.

### 1.3.2. Empirical models

Empirical relationships relate the $K_s$ values to basic soil properties, usually by using multiple linear regression. Examples include the models by Brakensiek et al. (1984), Cosby et al. (1984), and Saxton et al. (1986). Estimating $K_s$ from qualitative soil morphology has also been made. For example, McKeague et al. (1982) developed guidelines for estimating the $K_s$ class from soil field morphology. McKenzie et al. (1991) found that this method provides poor predictions when applied to soil from southern Australia, but there are indications for improvement when using the morphological description of Hollis and Woods (1989). Realizing the insufficiency of conventional morphological description, McKenzie and Jacquier (1997) developed a functional morphology description and applied it to 99 soil horizons from 36 sites across south-eastern Australia. They used regression-tree analysis to predict the log transformed $K_s$ value $\log_{10}(K_s+5)$. Soil pore structure, as characterised in the field by areal porosity, and other properties such as bulk density, and index of clay dispersion were found to be useful in estimating $K_s$.

Tamari et al. (1996) used ANN to predict $K$ at various potentials using horizon, textural class, organic matter content, bulk density, and water content at a particular potential. They concluded that ANN is more efficient than multilinear regression. However, they also noted that ANN could only be useful if a large database with accurate measurements was available. Schaap et al. (1998) used 620 samples from the USA to predict $K_s$ from basic soil properties with ANN, and showed that ANN gave better prediction than all previously published PTFs.

Rawls et al. (1998) derived the average $K_s$ set according to the US soil textural classes from large data set ($N = 953$).
Table 1.3.1. Average $K_s$ values according to US textural classes (Rawl et al., 1998).

<table>
<thead>
<tr>
<th>USDA Texture</th>
<th>$N$</th>
<th>$K_{s*}$ (mm/h)</th>
<th>$K_s^\phi$ 25% (mm/h)</th>
<th>$K_s^\phi$ 75% (mm/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sand</td>
<td>69</td>
<td>267.79</td>
<td>134.9</td>
<td>92.5</td>
</tr>
<tr>
<td>Fine Sand</td>
<td>23</td>
<td>123.0</td>
<td>106.6</td>
<td>92.5</td>
</tr>
<tr>
<td>Loamy Sand</td>
<td>47</td>
<td>43.83</td>
<td>65.0</td>
<td>36.3</td>
</tr>
<tr>
<td>Loam Fine Sand</td>
<td>14</td>
<td>57.5</td>
<td>35.6</td>
<td>122.0</td>
</tr>
<tr>
<td>Sandy loam</td>
<td>187</td>
<td>15.95</td>
<td>22.9</td>
<td>10.2</td>
</tr>
<tr>
<td>Fine Sandy Loam</td>
<td>60</td>
<td>12.2</td>
<td>5.1</td>
<td>29.2</td>
</tr>
<tr>
<td>Loam</td>
<td>109</td>
<td>5.02</td>
<td>5.2</td>
<td>2.8</td>
</tr>
<tr>
<td>Silt</td>
<td>107</td>
<td>7.7</td>
<td>3.6</td>
<td>22.9</td>
</tr>
<tr>
<td>Sandy Clay Loam</td>
<td>130</td>
<td>5.49</td>
<td>3.5</td>
<td>1.0</td>
</tr>
<tr>
<td>Clay loam</td>
<td>73</td>
<td>3.41</td>
<td>1.2</td>
<td>3.0</td>
</tr>
<tr>
<td>Silty Clay Loam</td>
<td>59</td>
<td>4.63</td>
<td>4.3</td>
<td>13.0</td>
</tr>
<tr>
<td>Sandy Clay</td>
<td>17</td>
<td>4.73</td>
<td>1.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Silty Clay</td>
<td>17</td>
<td>4.01</td>
<td>1.6</td>
<td>7.1</td>
</tr>
<tr>
<td>Clay</td>
<td>41</td>
<td>7.60</td>
<td>1.9</td>
<td>6.8</td>
</tr>
</tbody>
</table>

* Geometric mean for $K_s$ according to Schaap et al. (1998)
‡ Geometric mean for $K_s$ according to Rawls et al. (1998), with 25 and 75 percentiles

<table>
<thead>
<tr>
<th>USDA Texture</th>
<th>$\phi$ (m$^3$ m$^{-3}$)</th>
<th>$\theta_{33}$ (m$^3$ m$^{-3}$)</th>
<th>$\theta_{1500}$ (m$^3$ m$^{-3}$)</th>
<th>$P_{50-2000}$ (dag kg$^{-1}$)</th>
<th>$P_{c2}$ (dag kg$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sand</td>
<td>0.42</td>
<td>0.08</td>
<td>0.03</td>
<td>92</td>
<td>4</td>
</tr>
<tr>
<td>Fine Sand</td>
<td>0.45</td>
<td>0.07</td>
<td>0.03</td>
<td>90</td>
<td>3</td>
</tr>
<tr>
<td>Loamy Sand</td>
<td>0.40</td>
<td>0.12</td>
<td>0.05</td>
<td>82</td>
<td>6</td>
</tr>
<tr>
<td>Loam Fine Sand</td>
<td>0.44</td>
<td>0.12</td>
<td>0.04</td>
<td>82</td>
<td>6</td>
</tr>
<tr>
<td>Sandy loam</td>
<td>0.41</td>
<td>0.21</td>
<td>0.11</td>
<td>67</td>
<td>12</td>
</tr>
<tr>
<td>Fine Sandy Loam</td>
<td>0.40</td>
<td>0.22</td>
<td>0.11</td>
<td>69</td>
<td>14</td>
</tr>
<tr>
<td>Loam</td>
<td>0.42</td>
<td>0.29</td>
<td>0.15</td>
<td>41</td>
<td>22</td>
</tr>
<tr>
<td>Silt</td>
<td>0.45</td>
<td>0.32</td>
<td>0.14</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>Sandy Clay Loam</td>
<td>0.39</td>
<td>0.29</td>
<td>0.21</td>
<td>58</td>
<td>27</td>
</tr>
<tr>
<td>Clay loam</td>
<td>0.42</td>
<td>0.34</td>
<td>0.24</td>
<td>34</td>
<td>35</td>
</tr>
<tr>
<td>Silty Clay Loam</td>
<td>0.46</td>
<td>0.36</td>
<td>0.23</td>
<td>10</td>
<td>33</td>
</tr>
<tr>
<td>Sandy Clay</td>
<td>0.40</td>
<td>0.31</td>
<td>0.23</td>
<td>49</td>
<td>36</td>
</tr>
<tr>
<td>Silty Clay</td>
<td>0.48</td>
<td>0.39</td>
<td>0.27</td>
<td>6</td>
<td>47</td>
</tr>
<tr>
<td>Clay</td>
<td>0.44</td>
<td>0.38</td>
<td>0.30</td>
<td>22</td>
<td>51</td>
</tr>
</tbody>
</table>

1.3.3. Physico-empirical models

The second approach for developing PTFs attempts to establish a physico-empirical relationship between $K_s$ and water-retention curves or particle-size distribution. Based upon a pore-size distribution model and certain connectivity assumption (Childs and Collis-George, 1950; Mualem, 1976) hydraulic conductivity can been derived from the water retention curve. For example Laliberte et al. (1966) derived the expressions for predicting $K_s$ from the Brooks-Corey model’s parameters. Mishra
and Parker (1990) derived the relationship from the van Genuchten model based on Mualem (1976) model:

\[ K_s (\text{mm s}^{-1}) = 1080 (\theta_s - \theta_r)^{5/2} \alpha^2, \quad [1.3.1] \]

where \( \alpha \) is in cm\(^{-1} \), for the Brooks-Corey model:

\[ K_s (\text{mm s}^{-1}) = 1080 (\theta_s - \theta_r)^{5/2} \left[ \lambda/(\lambda + 1) \right]^2 h_b^{-2}. \quad [1.3.2] \]

Since the water-retention curves are usually not measured in routine soil survey, a relationship is needed to estimate \( K_s \) from basic soil properties. By inserting the parametric PTF into the capillary models, \( K_s \) then can be predicted from basic soil properties (Campbell, 1985). This model also allows prediction of the \( K(\theta) \) relationship.

Bloemen (1980) related \( K_s \) to the median particle size and particle-size distribution index. Ahuja et al. (1984) associated \( K_s \) to the effective porosity (\( \phi_e \)) based on the simplified Carman-Kozeny power-function model:

\[ K_s = B \phi_e^m \quad [1.3.3] \]

where \( \phi_e \) is defined as the total porosity (\( \phi \)) minus the water content (\( \theta \)) at a potential of -33 kPa, \( B \) is a constant relating to soil type and the exponent \( m \) is an empirical value. This relationship takes into account the pore-size distribution within a single characteristic \( \phi_e \). Parameter \( B \) is related to the hydraulic radii of the flow channels and \( m \) is usually equal to 4 or 5 (Ahuja et al., 1984). Due to its empirical nature, \( B \) and \( m \) have been found to vary between data sets (Table 1.3.2). Instead of estimating the parameters empirically, Pachepsky et al. (1999) related the \( B \) and \( m \) to \( h_b \) and \( \lambda \) of the Brooks-Corey parameters. Utilising ANN, their model explained 72% of the variance for 450 samples.

<table>
<thead>
<tr>
<th>Soil</th>
<th>Author</th>
<th>( B ) (mm/h)</th>
<th>( m )</th>
<th>( R^2 )</th>
<th>( N )</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 soil profiles from US</td>
<td>Ahuja et al. (1989)</td>
<td>7645</td>
<td>3.29</td>
<td>0.67</td>
<td>473</td>
</tr>
<tr>
<td>Indiana soil</td>
<td>Franzmeir (1991)</td>
<td>18036</td>
<td>3.21</td>
<td>0.86</td>
<td>14</td>
</tr>
<tr>
<td>Swedish soil:</td>
<td>Messing (1989)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15-25% clay</td>
<td></td>
<td>4680</td>
<td>2.24</td>
<td>0.44</td>
<td>69</td>
</tr>
<tr>
<td>25-40% clay</td>
<td></td>
<td>4900</td>
<td>1.81</td>
<td>0.36</td>
<td>155</td>
</tr>
<tr>
<td>40-60% clay</td>
<td></td>
<td>3720</td>
<td>1.44</td>
<td>0.31</td>
<td>185</td>
</tr>
<tr>
<td>&gt;60% clay</td>
<td></td>
<td>19500</td>
<td>2.05</td>
<td>0.28</td>
<td>92</td>
</tr>
<tr>
<td>Australian soil</td>
<td>Minasny and McBratney</td>
<td>23190</td>
<td>3.66</td>
<td>0.44</td>
<td>428</td>
</tr>
</tbody>
</table>

Table 1.3.2. Parameters of Kozeny-Carman model for different soil.
The fractal dimension of the particle-size distribution has been related to soil hydraulic properties (Rieu and Perrier, 1998). This includes estimation of the soil water retention curve and hydraulic conductivity (Rieu and Sposito, 1991). Rawls et al. (1993) showed that by using Marshall's model for hydraulic conductivity and fractal principles, \( K_s \) (in mm/h) can be formulated as:

\[
K_s = 4.41 \times 10^4 \left( \frac{\phi^{\tau}}{nc^2} \right) R_1^2
\]

[1.3.4]

where \( \phi \) is the total porosity, \( nc \) is total number of pore-size classes, \( \tau \) is the pore interaction exponent which can be taken as 4/3, and \( R_1 \) is the largest equivalent pore radius. They further derived the relationship:

\[ nc = 1.86 D_{pore}^{5.34} \]

where \( D_{pore} \) is the pore fractal dimension and \( R_1 \) is calculated as:

\[ R_1 = \frac{0.148}{h_b} \]

with \( h_b \) the mean air entry potential (cm).

1.4. Pedotransfer functions validation

1.4.1. Statistical and functional validation

The main objective of developing PTFs is to be able to predict soil properties that are difficult to measure. How well PTFs predict certain soil properties can be evaluated by comparing the observed/measured data with the predicted one. The predicability of the PTFs is usually evaluated on a set of data not used in generating the PTFs (usually called the validation set).

There are several statistical measures that are used to assess the performance or predicability of the PTFs, or known as ‘goodness-of-fit’. Imam et al. (1999) distinguished two main categories of goodness of fit measures i.e. the residual-based, and the statistical-association-based approaches. Residual-based measures provide quantitative estimates of the deviation of PTFs predictions from measured data. The indicators used usually are the mean error (ME), the mean absolute error (MAE), the sum of squared error (SSE), root mean squared error (RMSE), and Akaike’s Information Criterion (AIC). Statistical-association-based measures provide quantitative estimates of the statistical covariation between observed and predicted values, such as the correlation coefficient \( R \), coefficient of determination \( R^2 \).
The residual-based measure is useful to compare the magnitude of errors by different models. Although the coefficient of determination ($R^2$) has been proven to be a poor measure of model performance (Willmott, 1981), it is still widely used to assess prediction (e.g. Vereecken et al., 1989). The $R^2$ value is generally interpreted as ‘the proportion of the variance explained’ and offers value between 0 (worst fit) and 1 (perfect fit) which is easy to interpret. But this measure only evaluates the linear relationship between the variables, and is insensitive to additive and proportional differences. For example, if prediction $\hat{y}$ that has a linear relationship with the observed data $y$: $\hat{y} = c_1 y + c_2$, $R^2$ will always give 1 whatever the value of $c_1$ and $c_2$ are. It represents the proportion of the variation explained by the model only if the model is linear with a constant term (intercept) (Ratkowsky, 1990). Alternative measures to overcome the insensitivity have been proposed, such as the coefficient of efficiency (Nash and Sutcliffe, 1970) and the index of agreement (Willmott et al., 1985).

Tietje and Tapkenhinrichs (1993) proposed the use of Mean Deviations (MD) and Root Mean Squared Deviations (RMSD) as a measure of how well the PTFs fit to the retention curve. It is the sum of the area difference between the observed and predicted water retention curves. A single number can represent how well the PTFs fit to the whole water retention curve. MD indicate whether the PTFs over or underestimated the observed data, while RMSD measure the absolute deviation from the observed data.

Statistical validation only assesses how well the PTFs describe the data. As the main or final objective of generating PTFs is to serve as input for simulation models, the validation should be evaluated in the final application. Vereecken et al. (1990) termed it functional validation, i.e. the evaluation of the performance of PTFs at the context of specific applications. Thus, validation depends on the final application of interests. Wösten et al. (1986) proposed a similar concept called functional criteria. The properties evaluated require calculations using both water retention and hydraulic conductivity curves. They evaluated the following properties: travel times of water from the soil surface to a defined water table, water tables which allow a defined upward-flux density to a defined level, and downward flux densities that correspond with a defined air content of the soil.
Vereecken et al. (1992) performed functional validation of hydraulic PTFs to evaluate land qualities in Belgium. Using soil-plant water simulation, the parameters evaluated are: moisture supply capacity (ratio of the actual to the potential rate of transpiration), and drainability (cumulative amount of drainage from the soil profile). The effect of uncertainty of the input variables of the PTFs (e.g. bulk density or clay content) on the error of the estimated variables was evaluated by means of Monte Carlo simulation (Vereecken et al., 1992).

Wösten et al. (1990) compared class and continuous PTFs in relation to certain aspects of land management. They concluded that the results from functional validation can be different and ambiguous depending on the aspects evaluated. Wösten et al. (1990) evaluated four different methods to generate soil hydraulic properties in characterizing soil water profile. The methods used in generating the soil hydraulic properties are: laboratory measurements, hydraulic properties averaged on a regional scale, hydraulic properties averaged on a national scale and continuous PTFs. These hydraulic properties were used as input into soil-water simulation model. The soil water storage in the upper 0.5 m of the soil profile was evaluated in the simulation for a seven-year period. They showed that there are no significant differences between the direct measurements, class PTFs and continuous PTFs. Nevertheless, direct measurements provide the closest estimate of water storage as measured by neutron probe. A subsequent study by van Alphen et al. (2000) showed that continuous PTFs with measured saturated water content provide the closest estimate of soil water content with least uncertainty. Loage (1992) found that the prediction of stormflow for a small rangeland catchment were significantly reduced when using $K_s$ estimated by PTFs rather than from field infiltration data.

Timlin et al. (1996) used published water retention and $K_s$ PTFs as input to soyabean growth and development model. They simulated soyabean yields for seven locations in Mississippi for a seven-year period. Water retention predicted by PTFs correlated well with measured values while $K_s$ estimation were poor. Simulated yields were affected more by the method of estimating water retention rather than $K_s$. Crop yields simulated using PTFs predicted properties differ significantly from the yields simulated using measured properties. The average simulated yields from PTFs were higher than those obtained using measured properties because PTFs gave higher estimates of available water.
Bond *et al* (1999) compared direct measurements and PTF estimates of hydraulic properties for a Red Kandosol (Isbell, 1996) in Wagga Wagga. The hydraulic properties were also used to simulate bromide transport and water balance in the field. They found that there are large differences in hydraulic properties between the methods, but only small differences in the simulation result. Simulated results showed good agreement with field measurements, indicating the applicability of PTFs in this scenario.

Espino *et al.* (1995) evaluated the performance of published water retention and hydraulic conductivity PTFs to predict soil water contents, pressure heads and drainage fluxes for a layered profile. Simulations using PTFs as inputs over-predict the actual moisture content throughout the soil profile, but predict pressure heads near the soil surface quite well. The drainage fluxes were four times higher compared to the values calculated using measured hydraulic properties. They presented some cautionary notes on the use of PTFs:

- PTFs were derived from limited data, as a result, they will never completely capture the hydraulic processes of interest;
- the applicability of PTFs should be defined;
- functional validation of the PTFs should be evaluated first on the soil series where the functions were derived;
- different types of hydraulic models and PTFs should also be evaluated;
- despite the large uncertainties, they are confident that the use of PTFs will not jeopardize the modelling results.

### 1.4.2. The use of PTFs in quantifying spatial variability

Simulation of soil-water systems is not limited to a single site (point), usually simulations over space are required. Modelling hydrological processes on a regional scale not only require hydraulic properties characterization but also the description of spatial variability.

Springer and Cundy (1987) compared the parameters of Green-Ampt infiltration equation from field measurements and those calculated from hydraulic PTFs. They showed that the mean and variance of the parameters when estimated by PTFs were not preserved, the variances are always lower. The spatial trends and cross-
correlations amongst the parameters were also induced. They further used the PTFs to simulate overland flow and found that the results were significantly different when using field measured parameters. Mulla (1988) showed that matric potential from laboratory measurements and PTFs showed similar mean and spatial patterns, the PTF estimates were less accurate (larger variance) however.

Romano and Santini (1997) evaluated the performance of PTFs in quantifying the spatial variability of water retention of top soil within an area sampled 50 m apart along a 5 km transect. Using point regression and parametric PTFs, they evaluated the semivariogram of water content at -1, -10 and -100 kPa. They found that most of the PTFs were adequate in preserving the variance within the field. These studies seem to be unclear as to whether using PTFs can preserve the spatial patterns or underestimate/overestimate the variances.

Measured properties are usually limited in a field, consequently spatial prediction is required to generate continuous soil properties. A robust spatial interpolation method in geostatistics known as kriging can be applied. Combination of kriging and PTFs can generate a continuous map, and there are two possibilities to combine them. The first approach is first to interpolate related soil properties at unvisited location using kriging, then apply PTFs to the interpolated variables. The second approach applies PTFs to point measurements in the field and then interpolates the predicted results. Bocneau (1998) compared these approaches to estimate CEC in West-Flanders province, Belgium and found that the performance of both methods is almost equal. Sinowski et al. (1997) compared these approaches in estimating water retention curve and found that the first approach yields better prediction. Nevertheless, they found that the interpolated $\theta$ (from both methods) deviate from the measured values up to 0.1 m$^3$ m$^{-3}$.

1.5. Conclusions

The concept of PTFs has long been applied in soil science and has been revived in the past decades to accommodate the need for input data to process-simulation models evaluating agronomic and environmental problems. This has been aided by the development of large electronic soil databases and computational methods. PTFs are useful tools for assessing the static and dynamic soil quality indicators but special care must be taken when applying published PTFs developed elsewhere.
The main problem developing PTFs in Australia is the limited database available. Extensive soil physical and chemical measurements have rarely accompanied routine soil survey in Australia. Currently there are only few published data sets that contain such information, e.g. Prebble (1970), Colwell (1977), Forrest et al. (1985), McGarry et al. (1989) and Geeves et al. (1995).

Different authors have made various conclusions whether properties predicted from PTFs could yield reliable estimates when used in process-simulation models. It may be suggested that that hydraulic properties estimated through PTFs are useful for simulating large areas, but may not be accurate enough to predict the properties for site-specific purposes. When applying the PTFs within a farm-scale, the performance of PTFs may be questioned. As the nature of regression by its very definition is bringing the variation into the mean, the expected variation can be lost. The usefulness of PTFs will clearly depend on the scale and objective of the study.

PTFs are still being developed, tackling issues such as searching for (a) better mathematical expressions of PTF equations, (b) the most influential soil basic parameters to be used as PTF inputs, (c) ways to group soils, to have more accurate PTFs for each group (Pachepsky and Rawls, 1999), and (d) alternative methods to derive or fit the PTFs. There is little effort in integrating and applying the available functions and resources. An optimum pedotransfer system is envisaged to provide the best estimates of soil properties from the available information as well as to report the associated uncertainties.
References


Chapter I – Pedotransfer functions


II. The Disc Permeameter

Soil hydraulic properties measured in the field are important parameters describing water and solute transport dynamics. The disc permeameter, because of its relative ease of use, is one of the most popular devices for characterizing in situ saturated and unsaturated surface soil hydraulic properties. This chapter reviews the application of disc permeameter in soil science, the analysis of the data, and models describing water flow from the disc.

2.1. The disc permeameter and its application

Conventional techniques for measuring in-situ infiltration include the use of a single or double ring infiltrometer. The cumbersome double ring only measures flow under ponded (saturated) conditions, and when used in soil with distinct macropores, preferential flow will dominate. This does not reflect the rainfall or sprinkler irrigation. Many authors attempted to create a negative potential (tension) on the water flow. This is to exclude macropores in the flow process, hence only measuring the soil matrix flow. According to White et al. (1992), Willard Gardner and Walter Gardner developed a negative head permeameter as early as 1939. Dixon (1975) developed a closed-top ring infiltrometer to quantify macropores. Water is applied to a closed-top system, which permits the imposition of negative head or pressure on the ponded water surface. Negative tension can be considered as simulating a positive soil air pressure, created by a negative air pressure above ponded surface water. A simplification was made by Topp and Zebchuk (1985). The limitation of this device is the infiltration has to be started by ponding the closed-top infiltrometer (applying a positive head), then adjusted to a negative pressure. Little research effort was continued in this area, instead attention has been given mainly to the sorptivity apparatus of Dirksen (1975) which used a ceramic plate as a base. Based on this design, Clothier and White (1981) developed the ‘sorptivity tube’ which can provide a constant negative potential (tension) on the soil surface. However, the ‘sorptivity tube’ had many shortcomings, hence modifications to the design led to the development of the disc permeameter, also known as the tension infiltrometer (Perroux and White, 1988).
B. Minasny – Efficient methods predicting soil hydraulic properties

Gardner’s 1959 tension permeameter

Sorptivity apparatus (Dirksen, 1975)

Closed-top infiltrometer (Topp and Zebchuk, 1985)

Sorptivity tube (Clothier and White, 1981)

Disc permeameter (Perroux and White, 1988)

Detachable base tension infiltrometer (Soil Measurement System)

Mini disk infiltrometer (Decagon devices)

Automated tension infiltrometer (Ankeny et al. 1988)
Chapter II – The Disc Permeameter

Hood infiltrometer (UGT)

1. fine sand sealing
2. hood
3. overflow chamber
4. stand pipe
5. infiltration reservoir
6. bubble tower
7. air pipe
8. U-tube manometer
9. connection tube
10. air outlet tube
11. infiltration tube
12. tripod

Figure 2.1.1. Various designs of the field permeameter.

TRIMS Triple Ring Infiltrometer at Multiple Suction (Angulo-Jamarillo et al., 2000)

Double-disk tension infiltrometer (Smettem et al., 1995)

Double ring tension infiltrometer (UGT)

Pressure ring infiltrometer (Angulo-Jamarillo et al., 2000)
The disc permeameter comprises of a nylon mesh supply membrane, a water reservoir and a bubbling tower. The bubbling tower is connected to the reservoir and controls the potential applied to the membrane. It can be used to supply potential ranging \(-200 \text{ mm} \leq h_0 \leq 0 \text{ mm}\), effectively excluding pores with diameter bigger than 0.075 mm. Many different designs have evolved (Figure 2.1.1), including automated recording tension infiltrometer (Ankeny et al., 1988), a detachable disc (Soil Moisture Measurement http://www.soilmeasurement.com), a double ring disk permeameter (Smettem et al., 1995), mini-disc infiltrometer (Decagon devices, http://www.decagon.com), and the hood infiltrometer (Umwelt-Geräte-Technik, http://www.ugt-online.de) which is based on the closed-top infiltrometer.

The disc has been used extensively in characterizing in situ soil hydraulic properties as well as in various other applications in soil science research as summarised in Table 2.1. The design has also been used to provide a constant head of water flowing through soil cores in laboratory (Clothier et al., 1995). Furthermore, White et al. (1992) suggested that the disc could be used for measuring liquid flow in porous materials other than soil. For example, ‘it may be used for testing the absorbency of diapers, paper products, cloth and towelling’.

The disc has been found useful in evaluating tillage effects on soil hydraulic properties. An example is shown in Figure 2.1.2, which shows previously unpublished disc permeameter measurements made by the author along a transect under two land management regimes in a Black Vertosol from Narrabri, NSW. Sorptivity is higher under cultivated site due to break-up of aggregates in smaller sizes allowing more water sorptivity in the big pores. The pasture site which has not been cultivated for the past 5 years show relatively low conductivity than the cultivated site.

![Figure 2.1.2. Effect of different land management on sorptivity and steady-state infiltration rate as measured by the disc permeameter.](image-url)
Table 2.1.1. Some applications of the disc permeameter

<table>
<thead>
<tr>
<th>Applications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tillage effects</strong></td>
<td></td>
</tr>
<tr>
<td>tillage effect and wheel-traffic effect on infiltration</td>
<td>Ankeny <em>et al.</em> (1991)</td>
</tr>
<tr>
<td>temporal effect on hydraulic properties</td>
<td>Messing and Jarvis (1993), Logsdon (1993)</td>
</tr>
<tr>
<td>determination of soil structure parameters</td>
<td>White <em>et al.</em> (1992)</td>
</tr>
<tr>
<td>flow properties under different tillage treatments</td>
<td>Smettem and Ross (1992)</td>
</tr>
<tr>
<td>hydraulic properties, water repellance, macroporosity in a red earth</td>
<td>Chan and Heenan (1993)</td>
</tr>
<tr>
<td>spatial variability of surface hydraulic properties associated with traffic and cropping systems</td>
<td>Mohanty <em>et al.</em> (1994)</td>
</tr>
<tr>
<td><strong>Soil physical characterisation</strong></td>
<td></td>
</tr>
<tr>
<td>near saturation hydraulic properties of aggregated soil</td>
<td>Smettem and Kirkby (1990)</td>
</tr>
<tr>
<td>soil water repellency assessment</td>
<td>Wallis <em>et al.</em> (1991)</td>
</tr>
<tr>
<td>effect of raindrop induced aggregate breakdown on changes to soil hydraulic properties</td>
<td>Somaratne and Smettem (1993)</td>
</tr>
<tr>
<td>infiltration measurement on sports turf</td>
<td>Gibbs (1993)</td>
</tr>
<tr>
<td>soil structure and texture effect on infiltration</td>
<td>Jarvis and Messing (1995)</td>
</tr>
<tr>
<td>estimating mobile-immobile water content</td>
<td>Clothier <em>et al.</em> (1992)</td>
</tr>
<tr>
<td>estimating solute adsorption isotherm</td>
<td>Clothier <em>et al.</em> (1996)</td>
</tr>
<tr>
<td>quantifying and delineating spatial variations in surface infiltration</td>
<td>Sullivan <em>et al.</em> (1996)</td>
</tr>
<tr>
<td>impact of ant burrows on hydraulic properties of a sand</td>
<td>Wang <em>et al.</em> (1996)</td>
</tr>
<tr>
<td>determination of hydraulic properties of surface crusts</td>
<td>Vandervaere <em>et al.</em> (1997)</td>
</tr>
<tr>
<td>soil structure changes during growing season</td>
<td>Angulo-Jaramillo <em>et al.</em> (1997)</td>
</tr>
<tr>
<td>estimating hydraulic properties via inverse method</td>
<td>Šimunek <em>et al.</em> (1998)</td>
</tr>
<tr>
<td>scaling near saturation hydraulic conductivity</td>
<td>Shouse and Mohanty (1998)</td>
</tr>
<tr>
<td>topographic and seasonal variation in hydraulic conductivity</td>
<td>Hedddadj and Gascuel-Odoux (1999)</td>
</tr>
</tbody>
</table>
A recent review on the application of disc permeameter in field measurement of surface hydraulic properties and solute transport parameters was presented by Angulo-Jamarillo et al. (2000).

2.2. Analysis of disc permeameter data

2.2.1. Steady-state infiltration analysis

Due to the three-dimensional water flow from the disc, a special formulation is needed to take into account the lateral absorption of water. The analyses are derived from ‘the beautifully simple, steady-state analysis’ of Wooding (1968). For ‘steady infiltration from a circular, shallow, inundated area’, Wooding found that ‘a remarkable feature of this curve is the fact that it never departs far from the straight line’:

\[ Q^* = 2\pi a + 4 \]  

[2.2.1]

where \( Q^* \) is the dimensionless flux, \( a = \alpha_g r_0^2/2. r_0 \) is the radius of the disc [L] and \( \alpha_g \) [L\(^{-1}\)] is the sorptive number (White and Sully, 1987) or the parameter of Gardner’s (1958) hydraulic conductivity function:

\[ K(h) = K_s \exp(\alpha_g h) \]  

[2.2.2]

where \( K \) is the hydraulic conductivity [L T\(^{-1}\)], \( K_s \) is saturated \( K \) and \( h \) is soil water potential [L]. In terms of the actual steady-state infiltration rate \( q_\infty \) [L T\(^{-1}\)]:

\[ q_\infty = \alpha_g \phi_0 + \frac{4\phi_0}{\pi r_0} \]  

[2.2.3]

where \( \phi_0 \) is the matrix flux potential [L\(^2\) T\(^{-1}\)]. The matrix flux potential is defined as:

\[ \phi_0 = \int_{h_n}^{h_0} K(h)dh \]  

[2.2.4]

where \( h_0 \) is the applied potential, and \( h_n \) is the soil’s initial potential. From Gardner’s hydraulic conductivity relationship, the matric flux potential can be defined as:

\[ \phi_0 = \alpha_g^{-1} [K(h_0) - K(h_n)] \]  

[2.2.5]

Substituting this relationship into Equation [2.2.3], we arrive at the more general relationship:

\[ q_\infty = \Delta K + \frac{4\phi_0}{\pi r_0} \]  

[2.2.6]
where $\Delta K = K(h_0) - K(h_n)$. When the soil is initially dry, $K(h_n)$ is very small and can be neglected. The first term of the equation accounts for the vertical flow beneath the disc due to gravitational flow (as in one-dimensional infiltration) and the second term takes into account the capillary absorption.

Various analyses derived from Wooding’s solution for estimating hydraulic conductivity from disc permeameter measurements have been proposed. These include: single infiltration from a disc based on microscopic capillary length theory (White and Sully, 1987), multiple tension measurement from a disc (Ankeny et al., 1987; Reynolds and Elrick, 1991), and single tension with different discs radii (Smettem and Clothier, 1989). These equations are summarised in Figure 2.2.1.
Figure 2.2.1. Different analysis for disc permeameter derived from Wooding’s solution.
2.2.2. Hydraulic characteristics derived from disc permeameter

The main hydraulic properties that can be estimated are sorptivity ($S_0$ [LT$^{-1/2}$]), and hydraulic conductivity ($K_0$). Other time characteristics and length scales can also be derived. These derivations are based on the ‘small-times’ and ‘large-times’ approximations. The ‘small-times’ approximation assumes the capillary force dominates over the gravitational force at the early stages of infiltration:

$$\lim_{t \to 0} \frac{dI}{dt} \approx \frac{1}{2} S_0 t^{-1/2}.$$ 

Philip (1969) introduced $t_{\text{geom}}$ the time after which the geometric effect dominates over the soil capillary forces in the absorptive flow at very short-times:

$$t_{\text{geom}} = \left[ \frac{r_0 \left( \theta_0 - \theta_a \right)}{S_0} \right]^2 \quad [2.2.7]$$

For ‘large-times’, when the gravitational force dominates, the infiltration rate is assumed to reach steady-state:

$$\lim_{t \to \infty} \frac{dI}{dt} \approx q_\infty.$$ 

Philip (1969) derived $t_{\text{grav}}$ which indicates the time at which gravity effect is equal to the capillary absorption:

$$t_{\text{grav}} = \left( \frac{S_0}{K_0} \right)^2. \quad [2.2.8]$$

This is originally derived for one-dimensional infiltration, but can be used to calculate the time to reach steady-state.

The importance of capillary sorption to three-dimensional flow can be expressed by a parameter called the macroscopic capillary length $\lambda_c$ [L] (White and Sully, 1987). Mathematically it is defined as:

$$\lambda_c = \Delta K^{-1} \phi_0 \quad [2.2.9]$$

It is a flow-weighted mean soil-water potential, which can be used to scale distance and soil-water potential. For Gardner’s hydraulic conductivity function (Equation [2.2.2], $\lambda_c = \alpha_a^{-1}$. Using this parameter, we can define Wooding’s equation as:

$$q_\infty = \Delta K \left( 1 + \frac{4 \lambda_c}{\pi r_0} \right). \quad [2.2.10]$$
The parameter $\lambda_c$ can be used to derive a measure of soil structure through its relationship with a characteristic mean pore size $\lambda_m$ [L] by means of capillary rise equation (White and Sully, 1987):

$$\lambda_m = \frac{\sigma_w}{\rho_w g \lambda_c}$$  \[2.2.11\]

where $\sigma_w$ [MT$^{-2}$] is the soil-water surface tension, $\rho_w$ is the density of water and $g$ [LT$^{-2}$] is the gravity acceleration. For pure water at 20°C, equation [2.2.11] reduces to:

$$\lambda_m = 7.4/\lambda_c.$$  \[2.2.12\]

where $\lambda_m$ and $\lambda_c$ is in mm.

White and Sully (1987) derived $\lambda_c$ in terms of readily measured soil properties:

$$\lambda_c = \frac{b S_0^2}{\Delta \theta \Delta K}$$  \[2.2.13\]

where $b$ is a parameter which describe the shape of soil-water diffusivity function and constrained to $0.5 \leq b \leq \pi/4$, with a ‘typical’ average value of 0.55. The associated capillary time scale $t_c$, which estimates the time when the infiltration is dominated by capillary absorption can be calculated as:

$$t_c = b (S_0/K_0)^2$$  \[2.2.14\]

White et al. (1992) showed some applications of the length scales in characterizing structural changes.

Smettem et al. (1994) presented an estimate of experimental time limit for the disk to reach a wetting depth of $z_1$ [L]:

$$t_{exp} = \left[ \frac{z_1(\theta_0 - \theta_n)}{S_0} \right]^2$$  \[2.2.15\]

2.2.3. Some limitations

Limitations of these analyses arise from the estimates of $S_0$ and $q_\infty$. The estimates of $S_0$ from small-times measurement commonly exhibit large error. Warrick (1992) showed that for two soil with $t_{geom}$ less than 1 and 30 hours, one-dimensional flow is only dominant for $< 20$ and 100 seconds respectively. Haerkamp et al. (1994) questioned the estimation of $S_0$ because it only applied at a very short time.

Calculation of $t_{grav}$ for different soil by Zhang (1998) showed the time to reach steady-state for ponded infiltration is: 3.6 hours for a silt loam, 9.7 hours for a loam
and 18.5 days for a clay. Hence Zhang (1998) proposed some procedures for estimating steady-state rate from the infiltration rate measured at certain times.

White et al. (1992) realized some limitations of the disc permeameter:

- the analyses assume that the soil is uniform and homogeneous, and non-swelling,
- Wooding’s solution is based on a simplified model of the $K(h)$ relationship,
- the time to reach steady-state can be very long,
- the need for close contact between the disc and the soil,
- the permeameter must be level and the soil strength has to be able to support the weight of the disc. For unstable soil the collapse of the soil structure could reduce the infiltration rate.

Wooding’s equation is not valid for a small pond or small disk diameters, therefore Weir (1987) provided a solution for small disks ($r/\lambda_c < 0.8$).

Some advantages (+) and disadvantages (−) of the various analyses may be summarised as follows:

**Single disc, single tension of White and Sully (1987):**

+ Provide estimates of $K_0$ at specified potential from single measurements.
- Need to measure $\theta$ beneath the disc (at supply potential), negative $K_0$ can be obtained due to overestimation of $S_0$ from early-time infiltration or underestimation of $\Delta \theta$.

**Single tension with discs having different radii:**

+ Error in estimates of $S_0$ from early-time measurements can be avoided because only steady-state rate is measured.
- Two measurements required to give one estimate of $K_0$, small-scale spatial variability could introduce error.

**Single disc, multiple tension:**

+ Error in estimates of $S_0$ from early-time measurements can be avoided because only steady-state rate is measured, can provide $K$ at different potentials at one location, small scale variability can be avoided.
- Two measurements required to give one estimate of $K_0$, hysteresis, difficulty in identifying the change in $q_\infty$ when modifying small changes in potential.
2.3. Multiple tension infiltration measurements

Ankeny et al. (1991) first introduced measurements using multiple tensions with the disc permeameter on a single spot. The method only requires measurement of the steady-state infiltration rate, which is a more stable estimate compared to sorptivity, and avoids the measurement of initial and final water content. This method is also more desirable compared to using discs with different radii. Ankeny’s method used measurement at pairs of applied tensions, and assumes an exponential hydraulic conductivity function.

Logsdon and Jaynes (1993) assume Gardner’s $K(h)$ relationship (equation 2.2.1) and use nonlinear fit to estimate the $K_s$ and $\alpha$ parameters from different $h_0$ and $q$ data. Reynolds and Elrick (1991) proposed the use of a piecewise exponential function between two consecutive applied tensions. Interpolation using a log-linear relationship enables the estimate of $K$ at the mid-point between two consecutive tensions.

The piecewise exponential functions appear to give more reasonable results than the nonlinear fit approach. Numerical calculations by Reynolds and Elrick (1991) showed that this method had an accuracy of ±7%. The single exponential function may not always fit the data (Logsdon, 1999). The nonlinear fit may require more than 3 data points to achieve a good fit and reasonable parameter estimates. Joel and Messing (2000) found that the piecewise method generally gave smaller estimates of $K$ than the single tension.

A previously unpublished comparison made by the author on the single exponential, piecewise exponential and Ankeny’s method to estimate $K$ for a silt loam from Mt. Annan Botanic Garden, near Sydney is shown in Figure 2.3.1. The piecewise exponential and Ankeny’s method show similar estimates, while the single exponential does not allow for the large conductivity jump with increasing potential.

Limitations of the multiple-tension analysis arise from the assumption that $K(h_n) \ll K(h_0)$. This simplification applies when the initial soil is relatively dry, but does not hold when applying consecutive tension. Nevertheless, Reynolds and Elrick (1991) found that this assumption does not limit the accuracy of this method. They further indicated that when $K(h)$ is very steep (as in structured soil) a small change in tension $\Delta h$ (~10 mm) is required to estimate $K$ accurately. On the other hand, when $K(h)$ is flat (as in clay soil) small changes in $\Delta h$ can not show the differences in $q$. In
many instances the steady state rate $q$ estimated at lower potential can be greater than
the higher potential. Such an example is found in infiltration study made by the author
on a silt loam from Mt. Annan Botanic Garden (Figure 2.3.2).

![Graph showing different estimates of hydraulic conductivity from multiple tension flow.](image)

**Figure 2.3.1.** Different estimate of hydraulic conductivity from multiple tension flow.

![Graph showing infiltration with multiple tension application in a silt loam from Mt. Annan.](image)

**Figure 2.3.2.** Infiltration with multiple tension application in a silt loam from Mt. Annan.
The supply potential has been applied both in a decreasing manner (from more negative to less negative or more positive potential) and in an increasing manner (from positive to negative potential). Early applications of the multiple tension started with an increasing tension (Watson and Luxmore, 1986; Wilson and Luxmore, 1988; Ankeny et al., 1991). Realising the hysteresis effect, Jarvis and Messing (1993) and Logsdon (1993) employed a decreasing tension. Decreasing the tension will give higher values of $K$ compared with the increasing ones due to the hysteretic effect and simultaneous wetting and draining. (Figure 2.3.3)

![Figure 2.3.3. Hysteresis effect of changing supply potential from disc permeameter measurement of hydraulic conductivity on Waukegan silt loam (data from Logsdon, 1999).](image)

### 2.4. Algebraic models for the disc permeameter

Algebraic models relating cumulative infiltration $I$ [L] with time $t$ [T] have been developed. These models are useful for capturing the whole infiltration curve, rather than only analysing $I$ at early-times and large-times. The empirical or semi-analytical models are based on the one-dimensional infiltration model which satisfies the condition that the cumulative infiltration is proportional to the square-root of time at short times and reaches a steady state infiltration rate at long times (Collis-George, 1977). The available algebraic models are as follows:

1. **Two-term model**

   The two-term Philip’s equation (1957) for one-dimensional water infiltration has been adapted for the infiltration from a disc:

   $I = S_0 t^{1/2} + At$  \[2.4.1\]
Zhang (1997a, 1997b) developed empirical relationship between $S$ and $A$ to the parameters of hydraulic model.

(2) Modified Horton

Warrick et al. (1992) developed the empirical model:

$$I = S_0 t^{1/2} + q_\infty \{ t + c_1/c_2 [1-\exp(-c_2 t)] \} \quad [2.4.2]$$

where $a$ and $c$ are empirical constants.

(3) Linear Diffusion Model

For water absorption (neglecting gravity) in a soil having a constant diffusivity ($D$), Warrick (1992) developed an approximation for linear diffusion from a disc:

$$I = 1.27 \Delta \theta r_0 \{ T + 1/2 (\pi T)^{1/2} - 0.054 [1-\exp(-4.01 T)] \} \quad [2.4.3]$$

where $T = D t / r_0^2$

(4) Buried source model

The approximation is based on the ‘linearized, time-dependent’ infiltration from a buried spherical cavity (Philip, 1986). Pullan and Collins (1992) adapted the solution for a disc:

$$I = q_\infty t + B f(Ct) \quad [2.4.4]$$

where

$$B = 2 \lambda_c^2 K_0 / D$$

$$C = D / 4 \lambda_c^2$$

$$f(u) = 1/2 - (u+1/2) \text{erfc}(u^{1/2}) + (u/\pi)^{1/2} \exp(-u)$$

(5) Scaling model

Youngs (1987) developed the model for water flow from a ring infiltrometer based on the concept of scaling. For dimensionless cumulative infiltration $\bar{I} < 1.5$:

$$\bar{I} = 0.73 t^{1/2} + t^* \quad [2.4.5]$$

where

$$\bar{I} = I \pi / (\Delta \theta r_0)$$

$$t^* = \lambda_k \sigma_w t / (\eta_w r_0^2 \Delta \theta)$$

$\lambda_k$ is the macroscopic capillary length, $\sigma_w$ and $\eta_w$ are the surface tension and viscosity of water, respectively.

(6) Turner and Parlange model

Turner and Parlange (1972) developed an approximation for two-dimensional water infiltration for short times (neglecting gravity).
\[ I = S_0 t^{1/2} + \left[ \frac{1}{3} K_0 + \frac{0.48 S_0^2}{r_0 \Delta \theta} \right] t \]  \hspace{1cm} [2.4.6]

(7) Haverkamp model

Haverkamp et al. (1994) developed the model for disc permeameter based on the model of Turner and Parlange (1972):

\[ I = S_0 t^{1/2} + \left[ K_i + \frac{\gamma S_0^2}{r_0 \Delta \theta} + \frac{1}{3} (K_0 - K_i)(2 - \kappa) \right] t \]  \hspace{1cm} [2.4.7]

where \( \gamma \) is a proportionality constant (0.6 < \( \gamma \) < 0.8) to take into account for the gravity effect at the water profile edges, \( \kappa \) is a shape factor (0 < \( \kappa \) < 1) related to:

\[
\kappa = 2 - 2 \int_{\theta_0}^{\theta_n} \left( \frac{K - K_n}{K_0 - K_n} \right) \frac{\theta - \theta_n}{\theta_0 - \theta_n} \frac{D(\theta)}{D(\theta) d\theta} \int_{\theta_0}^{\theta_n} D(\theta) d\theta
\]  \hspace{1cm} [2.4.8]

When initially the soil is dry \( K_n << K_0 \) and taking an average value for \( \kappa = 0.6 \) and \( \gamma = 0.75 \), the equation 6.2.5 becomes (Vandervaere et al., 2000):

\[ I = S_0 t^{1/2} + \frac{7}{15} K_0 t + \frac{0.75 S_0^2}{r_0 (\theta_0 - \theta_n)} t \]  \hspace{1cm} [2.4.9]

Fitting algebraic models to experimental results should not be justified simply by convenience in representing data, rather the parameters should have physical significance. Hussen and Warrick (1993a, 1993b), and Cook and Broeren (1994) evaluated different methods for estimating hydraulic properties from disc permeameter measurements. Hussen and Warrick (1993a) compared the modified Horton, linear diffusion and buried-source model. They found that the modified Horton and linear diffusion model gave similar estimates of \( S_0 \), while the buried source model generally underestimates it, but all give similar estimates of the steady-state rate. Comparison of single tension, double disc and multiple tension by Hussen and Warrick (1993b) showed that the single disc with multiple tensions (more than three supply potentials) gave the most accurate and repeatable estimates of \( K_0 \). They also found no difference in the \( K_0 \) between large \((r_0 = 118 \text{ mm})\) and small \((r_0 = 52 \text{ mm})\) disc.

Cook and Broeren (1994) found that the methods based on Wooding’s equation (White-Sully, twin rings and Ankeny) gave similar estimates of \( S_0 \) and \( K_0 \). The double
tension method of White and Perroux (1989) also gave consistent values of $K_0$. Compared with the other methods, the scaling method of Youngs (1987) gave very low and inconsistent values of $K_0$. They suggested that this method was developed for calculating $K$, and cannot be used for calculating unsaturated $K$.

The algebraic models only provide the estimates of sorptivity and hydraulic conductivity from the applied tension. Only the model by Haverkamp may provide some estimates of the hydraulic relationship of the soil ($\theta(h)$ and $K(h)$). The parameters in the equation are directly related to the hydraulic properties.

Smettem et al. (1995) used a double-disc tension infiltrometer that simultaneously measures one- and three-dimensional infiltration. They estimated parameter $\gamma$ from the difference between the three- and one-dimensional infiltration. Jacques et al. (1999) attempted to use this model in an inverse method context to estimate the parameters of the van Genuchten hydraulic model. They found that their inverse method requires a good initial start, at least two infiltration curves (infiltration from different tension) and some $\theta(h)$ measurements. The estimate of the proportionality constant $\gamma$ remains a difficult problem.

2.5. A numerical model for the disc permeameter

Although many solutions and algebraic models for water flow from the disc permeameter have been proposed, there is still no general analytical solution. A numerical solution is needed to simulate water flow in the soil underneath the disc permeameter. Numerical models also have the advantage of being flexible, in defining the initial and boundary conditions, and also soil types. Numerical models have been applied for comparing and checking the validity of empirical or analytical solution (Reynolds and Elrick, 1991; Warrick, 1992), developing algebraic models (Zhang, 1997a, 1997b), simulating application of water and solute in the soil (Quadri et al., 1994), evaluating the effect of sand placement underneath the disc (Reynolds and Zebchuk, 1996), and providing a model for the inverse solution (Šimunek et al., 1997).

Water flow from the disc permeameter is three-dimensional, but numerical solution for three-dimensional systems can be complicated and requires heavy computation. However, an assumption can be made that horizontal flow normal to the radial ($r$) axis does not occur. This allows the problem to be formulated in an axially-
symmetric form (Felton and Reddel, 1992). The pseudo three-dimensional flow can be solved using the axisymmetric two-dimensional formulation of Richards’ equation.

\[
\begin{align*}
\frac{\partial \theta}{\partial t} &= \frac{1}{r} \frac{\partial}{\partial r} (r \theta u) + \frac{\partial}{\partial z} (\theta v) \\
&= -K \frac{\partial H}{\partial r} \\
&= -K \frac{\partial h}{\partial r} \\
&= -D \frac{\partial \theta}{\partial r}
\end{align*}
\]

Figure 2.5.1. Three-dimensional water flow from disc permeameter (left) approximated by two-dimensional axisymmetric model (right).

For steady, axially-symmetric flow of water in unsaturated soil with the vertical direction coinciding with the axis of symmetry, the mass balance equation is (Raats, 1971):

\[
\frac{\partial \theta}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} (r \theta u) + \frac{\partial}{\partial z} (\theta v)
\]

where \( r \) and \( z \) is the radial (horizontal) and vertical (depth) coordinates [L], \( \theta \) is the volumetric water content [L\(^3\) L\(^{-3}\)], \( u \) and \( v \) are the radial and vertical components of water velocity. Expressions for the radial and vertical components are given by Darcy’s law:

\[
\begin{align*}
\theta u &= -K \frac{\partial H}{\partial r} \\
&= -K \frac{\partial h}{\partial r} \\
&= -D \frac{\partial \theta}{\partial r}
\end{align*}
\]
where $K$ is the hydraulic conductivity [L T$^{-1}$], $H$ is the total potential [L] (the sum of pressure head $h$ and gravitational head $z$ ($H = h - z$), $D (= K \frac{dh}{d\theta}$) is the diffusivity [L$^2$ T$^{-1}$]. Substituting [2.5.2] into [2.5.1] gives:

$$\frac{\partial \theta}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left( rK \frac{\partial h}{\partial r} \right) \varphi + \frac{\partial}{\partial z} \left( K \frac{\partial h}{\partial z} \right) - \frac{\partial K}{\partial z} \tag{2.5.3}$$

The initial and boundary conditions for water flow into the soil under the disc with radius $r_0$ is given by (Warrick, 1992):

$$\begin{align*}
h(r, z, t) &= h_n \quad r \geq 0, z \geq 0, t = 0 \quad \text{- initial condition in terms of } \theta \text{ or } h \\
\theta(r, z, t) &= \theta_n \quad r \geq 0, z \geq 0, t = 0 \quad \text{- initial condition in terms of } \theta \text{ or } h \\
h(r, z, t) &= h_0(t) \quad 0 < r < r_0, z \geq 0, t > 0 \quad \text{- boundary condition for the disc} \\
\frac{\partial h}{\partial r} = 1 \quad z = 0, \quad r > r_0 \quad \text{- no flow at surface} \\
h(r, z, t) &= h_n \quad r^2 + z^2 \rightarrow \infty \quad \text{- no flow at the flow domain edges} \tag{2.5.4}
\end{align*}$$

The initial conditions specify the initial soil’s water status in terms of either pressure head or water content. The boundary condition prescribes the applied tension from the disc (Figure 2.5.2). The other boundaries are: no flow (zero flux) on the remainder of the soil surface ($r_0 < r < r_{\text{max}}$), bottom of profile ($z_{\text{max}}$), x-axis boundaries ($r_{\text{max}}$) and axis of the symmetry ($r = 0$). The last condition also ensures that the boundaries in $r$ and $z$ direction are sufficiently distant from the disc so they do not influence the flow process.

![Figure 2.5.2. Two-dimensional model water for water flow from the disc permeameter.](image-url)
Equations [2.5.3] and [2.5.4] can be solved numerically using the finite-difference or finite-element approach. Felton (1992) simulated water flow beneath a double ring infiltrometer, where a negative head was applied in the inner ring and positive head in the outer ring. He used a two-phase (air and water), two-dimensional finite-element model to evaluate the potential gradient beneath the tension infiltrometer. Warrick (1992) developed a finite-element program called DISC, which was used to formulate an algebraic model for the infiltration from the disc. Quadri et al. (1994) developed a finite-difference solution simulating water and solute transport from a disc. They verified their simulated results with a laboratory experiment using a box of repacked sand. Using a quarter of the disc, they found agreement between the simulated flux density, water content, pressure head and bromide concentration distribution. Šimunek et al. (1998) used their finite-element program SWMS_2D (Šimunek et al., 1994) in an inverse solution. Using numerical examples, they predicted the hydraulic properties from cumulative infiltration, soil water content and pressure head data (more details in Chapter 3).

2.6. Conclusions

The disc permeameter has many practical uses in characterizing in situ soil hydraulic properties. The main problem is the analysis of the results. Most of the analyses rely on Wooding’s steady-state rate approximation. Many assumptions lie behind the analyses, and can only provide the hydraulic properties at the applied tensions. The algebraic models remain empirical and so far have not provided additional merits compared with the steady-state analysis. A numerical model coupled with an inverse solution may be a promising model to provide estimates of hydraulic properties of the soil.
References


Chapter II – The disc permeameter


III. Inverse methods predicting soil hydraulic properties

3.1. Introduction

One of main uses of computers and computational methods science is modelling. Based on physical, chemical or biological principles, models are designed to mimic real-world phenomena. Models provide an effective, and sometimes the only means of understanding the process of a system, and/or evaluating the outcome of a system as a result of specific scenario. They are also used to investigate relationships among the parameters that affect the functioning of the system. Models are usually not developed as an end in themselves. At the next level, the computer can be used to examine the results of modelling and the model is rerun to adjust the parameters so it can provide the best description of real-life situation. This is called fitting or inversion (Gill et al., 1981; Claerbout, 1992).

In soil physical modelling, the main problem is defining water, solute, heat, or gas transport process in the soil. If we know the soil transport properties, we can predict the outcome of specified input to the soil system. This is recognized as the forward problem. On the other hand, we can observe the output or response of the system from a prescribed input and work out the transport properties of the system. This is called the inverse problem (Figure 3.1.1).
The forward method involves prediction of the behaviour of a system given its inherent properties, whereas the inverse method determines the properties of a system from its response to a known stimulus (Sposito et al., 1995). The solution of an inverse problem entails determining unknown causes based on observation of their effects, while the direct problem involves finding effects based on a description of their causes (Hopmans and Šimunek, 1999). The word ‘inverse’ is derived from ‘matrix inversion’ (Claerbout, 1992). Despite its association with a well-known and well-defined mathematical solution, inverse solutions are not simple. When the properties to be estimated are characterized by a number of parameters, the method is referred to as parameter identification (DuChateau, 1996).

One of the most difficult and time-consuming properties to be measured in soil science is unsaturated hydraulic properties. The conventional method for determining the water-retention curve involves equilibrating soil under certain potential or simultaneous measurements of water content and potential during infiltration or drainage (Klute, 1986; Reeve and Carter, 1991). Hydraulic conductivity is measured by imposing specified upper and lower boundary conditions in laboratory soil columns or soil profiles in the field and measuring the steady-state water flow (Klute and Dirksen, 1986; Dirksen, 1991). Water retention and hydraulic conductivity curves are usually parameterised (represented by model/s with certain number of parameters that adequately characterize them) and used as inputs to soil-water simulation models. Under defined boundary conditions (such as irrigation, drainage, precipitation, potential evapotranspiration) dynamic simulation provides water-balance calculations, such as infiltration, actual evapotranspiration, change in soil water storage, runoff, and net flux of water in the soil. The water balance enables us to evaluate the effect of water management on agricultural outputs and environmental factors, such as crop yield and groundwater pollution. Examples of such applications have been given by Feddes et al. (1988).

It is usually easier to measure the response of the flow processes (e.g. cumulative infiltration or water content changes) rather than measuring the hydraulic properties themselves. The inverse method enables us to estimate hydraulic properties from flow processes, offers a flexible experimental setup and provides simultaneous estimation of water retention and hydraulic conductivity. As long as the upper and lower boundary conditions are well defined, an inverse solution can be formulated. In a transient water flow system, flow attributes such as water content, matric potential,
cumulative inflow or outflow is recorded. A parametric model is then assumed to represent the soil’s hydraulic properties. Starting with an initial guess of the parameters, a numerical model describing the flow calculates the theoretical flow attributes. An optimization procedure is then used to adjust the parameters to minimize the differences between the observed and calculated attributes (Figure 3.1.2).

The application of the inverse method in saturated groundwater flow for predicting $K_s$ has been studied extensively and has been reviewed by Yeh (1986) and McLaughin and Townley (1996). Reviews of the application of the inverse method predicting unsaturated soil hydraulic properties have been made by Kool et al. (1987) and Hopmans and Šimunek (1999).

In the tradition of tracing the history of the application of inverse methods in soil science, the subsequent sections review the application of inverse methods in soil physics, followed by the mathematical formulation, and some definitions and problems found in inverse methods.
Figure 3.1.2. (a) Forward problem, involves determination of soil hydraulic properties and prediction of soil-water behaviour under certain boundary conditions (b) Inverse problem, estimation of hydraulic properties from flow response under prescribed boundary conditions.
3.2. Inverse methods for predicting soil hydraulic properties

Inverse methodology or parameter identification has become a popular subject in soil physics methodology. A sub-section is given in the review of unsaturated hydraulic conductivity determination by Dirksen (1991). The third edition of Methods of Soil Analysis by the Soil Science Society of America also assigned a special section for this method. Although most inverse methods developed recently have used a numerical solution coupled with an optimization procedure, techniques for deriving unsaturated hydraulic properties from flow measurement have long been investigated. As an example, Doering (1965) provided an analytical solution for transient outflow which permitted the estimation of soil water diffusivity from water outflow measurements. Peck (1968) deduced soil water diffusivity from a similar outflow experiment by matching the experimental water outflow curve with the one derived analytically. Although many analytical solutions provided estimates of hydraulic properties at certain potential (e.g. Philip, 1990), only methods that estimate the whole water retention and/or hydraulic conductivity curve will be discussed.

The uses of inverse methods have been expedited by the development of numerical solutions of soil-water flow problems facilitated by computers. Bouwer and Jackson (1974) reviewed the available methods for measurement of hydraulic conductivity at that time by assigning a section called ‘computer techniques’, which they described as ‘techniques that utilize computers to arrive at conductivity-water content and conductivity-pressure head relations’. The first use of such a method was by Richards and Weeks (1967), where transient water flow in horizontal soil column was monitored by tensiometers located at four points along the column. The solution, based on Richards and Weeks (1953), was programmed in Fortran IV, analysed the water content and pressure head data and fitted a power function to the $\theta(h)$ and $K(h)$ data.

Whisler and Watson (1968) proposed for the first time, an inverse procedure that utilised a numerical method to estimate $K(\theta)$ from the column drainage experiment. They formulated a numerical solution for one-dimensional gravity drainage and investigated the sensitivity of the numerical results to the change in $K(\theta)$ relationships. If $K_s$ is determined accurately, slight changes in the shape of the $K(\theta)$ curve will only cause small changes in the cumulative water outflow vs. time curve $Q(t)$ compared to the curve determined using a true $K(\theta)$ relationship. But a 10%
variation in $K_s$ will cause significant variation in the $Q(t)$. They proposed the use of a computer to calculate different $Q(t)$ from a column with fixed $\theta(h)$ and $K_s$ but with varying shapes of $K(\theta)$ curve. The calculated $Q(t)$ is then matched with the experimental curve to determine the appropriate $K(\theta)$ relationship. Subsequently, Watson and Whisler (1968) conducted an experiment on two types of sand to verify their procedures. Generating three different shaped $K(\theta)$ curves, they found the best curve that matched the $Q(t)$ (Figure 3.2.1).

Figure 3.2.1. Three assumed $K(\theta)$ curves for the sand and computed outflow using the relationships, circles represent experimental values (Watson and Whisler, 1968)

Skaggs et al. (1971) described a method for estimating $K(h)$ from ponded infiltration in a soil column. Assuming that $\theta(h)$ is known and a Gardner-type $K(h)$ relationship:

$$K(h) = \left( \frac{h}{h_b} \right)^{c_1} + c_2 \right]^{-1}.$$

[3.1.1]

They estimated parameters $h_b$ and $c_1$ by matching the infiltration rate with time $i(t)$ calculated from numerical model to measured values. Unlike the curve matching technique of Whisler and Watson, they defined an objective function to be minimised, i.e. the difference in area between the observed and predicted infiltration rate (which is equivalent to minimising mean absolute difference or $L_1$ norm):

$$O = \int_0^\infty \left| i(t) - \hat{i}(t) \right| dt.$$

[3.1.2]
They plotted the response surface of the objective function as a function of $h_b$ and $c_1$, thereby searching for a minimum where the solution exists (the response surface method is used widely in today’s inverse method, see section 3.5.1). They also presented an alternative method for estimating $h_b$ using the one-dimensional golden-section search method. Afterwards, Skaggs et al. (1973) successfully applied their method in laboratory columns of a sand and sandy loam.

There was not much further research in this area until Zachmann et al. (1981, 1982) presented a numerical study which considered the drainage from a saturated soil column. They compared four measurements for predicting $\theta(h)$ and $K(\theta)$ relationships: cumulative outflow with time $Q(t)$; maximum cumulative outflow $Q_\infty$; pressure head at a point with time; and water content at a point with time. Utilizing nonlinear least squares, they matched the simulated variables with true values. They found that $Q(t)$ measurements provide the best estimate of hydraulic properties.

Stimulated by this finding, Hornung and Messing (1982) and Hornung (1983) presented a numerical study that assumed a soil column under a constant infiltration rate at the surface. Assuming $K_s$, $\theta_s$, and $\theta_i$ of the van Genuchten equation were known, parameters $\alpha$ and $n$ were estimated by matching the outflow rate from the bottom of the column calculated by numerical solution to the ‘measured’ values. He questioned whether the infiltration experiment provided enough information for estimating soil hydraulic properties and introduced the term ‘uniqueness’ of solution. No unique solution could be obtained when only using outflow data, but could be resolved when additional information such as the final steady pressure head at a known position in the column was included.

Dane and Hruska (1983) took the challenge further by analysing the gravity drainage data from field experiment on a clay loam. Water contents at different depths were monitored from an initially saturated profile at 0, 7 and 25 days after drainage. These data were used as matching variables in an inverse solution to predict $\alpha$ and $n$ of the van Genuchten hydraulic model, assuming $K_s$, $\theta_s$, and $\theta_i$ were known. They obtained good agreement between the measured $\theta(h)$, but $K(\theta)$ were highly overestimated compared with the independently determined values. Better estimates were found when fixing $K_s$ at a much smaller value. The use of a lower $K_s$ value was justified by the possibility of macropore occurrence during the saturation of the soil profile.
Kool et al. (1985a) formulated the inverse problem from a transient one-step pressure outflow experiment (Gardner, 1956; Doering, 1965). A soil core is placed in a cell with a porous plate at the lower end (Figure 3.2.2). After the core is saturated with water, an air pressure is applied at the top of the core and outflow from the bottom is recorded with time. They successfully applied this method to estimate parameters $\theta_r$, $\alpha$ and $n$ from cumulative outflow measurements $Q(t)$ supplemented with steady-state outflow $Q_\infty$. Nonetheless, they remarked that in order to give reliable estimates, the initial parameter guess must be reasonably close to their true values, and the outflow measurement errors must be small. Kool et al. (1985b) provided the first public-domain code, ONESTEP, for the analysis of the inverse problem.

Parker et al. (1985) subsequently applied the one-step outflow procedure to four different soils ranging from a sandy loam to clay. They concluded that $\theta(h)$ and $K(\theta)$ can be optimized simultaneously by using $Q(t)$ supplemented with $Q_\infty$. However, $\theta(h)$ were overestimated at lower potential, nevertheless it could be alleviated by including a direct measurement of $\theta$ at -1500 kPa in the objective function.

Despite the successful application of the one-step outflow method by Parker et al. (1985), it was not successful when applied by other laboratories (van Dam et al., 1990, 1992) suggesting the non-uniqueness of the solution. An early study by Collis-George and Rosenthal (1966) had shown that unrealistic $K$ values at large moisture contents could be obtained from the one-step outflow experiments and hysteresis was also present in $K(\theta)$ and $D(\theta)$ relationships. Poor contact between the soil sample and
porous plate could reduce the water outflow from the column (Peck, 1968), and small leaks of gas or water in the pressure cell also cause large errors (van Dam et al., 1990). A numerical study by Russo (1988) investigated the influence of the parametric form of the hydraulic functions on the one-step outflow optimization. He concluded that the van Genuchten model performed best compared with the Brooks–Corey and Gardner–Russo (Russo, 1988) hydraulic models. Subsequently, however, Russo et al. (1991) pointed out that the larger number of parameters in the van Genuchten model may enhance the likelihood of non-uniqueness and instability in the inverse solution. Toorman et al. (1992) analysed the response surface of the one-step outflow objective function from outflow data supplemented with various combinations of water content and pressure head in the soil core. They indicated that uniqueness could be met if pressure head data at a point some distance from the outflow were supplemented in the objective function. From an x-ray tomography study of one-step outflow, Hopmans et al. (1992) found that a sudden pressure change in an initially saturated soil caused a non-uniform flow, which violates the supposed uniform Darcian flow.

van Dam et al. (1992) ascertained the need to independently measure soil water retention in the one-step outflow inverse problem. Optimization using outflow data alone is insufficient, and supplementary $\theta(h)$ data are necessary to give reliable $K(\theta)$ estimates. To circumvent the need for additional water retention and pressure head measurements in the outflow experiment, van Dam et al. (1994) proposed the multi-step outflow method, in which the air pressure was increased in several smaller steps. They also argued that a sudden change in the applied pressure does not represent natural conditions and could cause non-uniform flow. They estimated $\alpha, n, K_s$, and $l$ of the Mualem–van Genuchten model for a loam and modified the ONESTEP program into MULSTP (van Dam et al., 1990). Results showed that the outflow data from a multi-step experiment provide sufficient information to yield a unique solution.

Eching and Hopmans (1993a) and Eching et al. (1994) showed that the multi-step method combined with an automated pressure head measurement in the core during outflow could provide a unique solution for the inverse problem. Good agreement was found between optimized and independently measured $\theta(h)$ and $K(\theta)$ for four different soil textures. Eching and Hopmans (1993b) presented a modified ONESTEP and MULSTP program called MLTSPM. Further modifications, SF-OPT (Chen et al.,
1997), included measurement of time-dependent lower boundary conditions (pressure head at the bottom of the cores), and allowing for the estimation of two-fluid capillary pressure-saturation and permeability functions (Chen et al., 1999).

Durner et al. (1996) further modified the multi-step method by introducing a continuous smooth change in pressure. The method was also used to estimate the bimodal hydraulic functions (Zurmuhl and Durner, 1998). A review of the revolution from one-step to multi-step and continuous-step methods has been presented by Durner et al. (1999).

(a)

(b)

(c)

Figure 3.2.3. (a) Cumulative outflow from multi-step experiment and the fitted values using inverse method, measured and predicted water retention curve (b), hydraulic conductivity (c).
Although Hopmans et al. (1992) had shown the invalidity of the one-step method, the multi-step method became a popular inverse method as it is easy to implement (Crescimanno and Iovino, 1995; Vereecken et al., 1997). An unpublished multi-step outflow experiment on a Mount Annan silty-loam (Figure 3.2.3) showed that although the outflow is fitted well, the water retention and hydraulic conductivity do not correspond well with independent measurements.

Hollenbeck and Jensen (1998a) studied the reproducibility of the one-step and multi-step experimental procedures on a sand. They revealed that experiments with small step pressure changes were not reproducible, hence the inverse problem becomes ill-defined (Figure 3.2.4). Large step pressure changes are reproducible but the outflow was virtually the same for different pressure changes inducing a non-unique system. The irreproducibility of the onestep outflow method had already been reported by Butijn and Wesseling (1959) forty years ago. They took core samples from field profiles of a dune sand and sandy loam. They showed that predicted $K$ could vary 100–1000 times between replicates at the same pressure increment. These imply experimental limitations for the outflow procedures. Although theoretical analysis showed that the multi-step method is well-posed for the inverse solution, the experimental limitation restricts its application.

Figure 3.2.4. (a) Observed outflow variability (PV= pore volume) from one-step and multi-step outflow under small pressure steps and (b) outflow under large pressure step changes, all three pressure changes yield similar outflow responses. Numbers beside pressure $\psi$ indicate the pressure (in cm H$_2$O) applied (Hollenbeck and Jensen, 1998a)

While most research on the inverse method has focussed on the one-step outflow method, other researches seek for alternative measurements. Various experimental
possibilities have been explored in numerical studies, and laboratory and field experiments (Table 3.2.1).

Table 3.2.1. Studies involving inverse methods for predicting hydraulic properties from transient flow experiments.

<table>
<thead>
<tr>
<th>Flow processes</th>
<th>Measurement</th>
<th>Predicted properties &amp; hydraulic models</th>
<th>Method/soil</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Numerical studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>water absorption by soil through an impeding surface layer</td>
<td>position of the wetting front, cumulative inflow data</td>
<td>$\theta(h)$, $K(\theta)$: BC</td>
<td>quasi-analytical solution</td>
<td>Zayani <em>et al</em> (1991), Zayani <em>et al</em> (1992)</td>
</tr>
<tr>
<td>water redistribution</td>
<td>profile $h$</td>
<td>$\theta(h)$: Kosugi $K(\theta)$: Mualem</td>
<td>numerical study, layered profile</td>
<td>Kosugi &amp; Nakayama (1997)</td>
</tr>
<tr>
<td>radial multi-step water extraction</td>
<td>cumulative outflow &amp; $h$</td>
<td>$\theta(h)$, $K(\theta)$: BC &amp; VGM</td>
<td>numerical study</td>
<td>Finsterle &amp; Faybishenko (1999)</td>
</tr>
<tr>
<td>Ponded infiltration</td>
<td>cumulative infiltration &amp; $h$</td>
<td>$\theta(h)$, $K(\theta)$: VGM</td>
<td>numerical study</td>
<td>Vogel <em>et al</em>. (1999)</td>
</tr>
<tr>
<td>centrifuge</td>
<td>$h$ distribution</td>
<td>$K(h)$: BC, exponential</td>
<td>numerical study with analytical solution</td>
<td>Basha &amp; Mina (1999)</td>
</tr>
<tr>
<td>drainage</td>
<td>profile $\theta$</td>
<td>$\theta(h)$, $K(\theta)$: VGM</td>
<td>numerical study, layered profile</td>
<td>Zijlstra &amp; Dane (1996)</td>
</tr>
<tr>
<td><strong>Field studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>drainage from initially saturated profile</td>
<td>profile $h$ &amp; $\theta$</td>
<td>$\theta(h)$: VG $K(\theta)$: exponential</td>
<td>field, volcanic soil</td>
<td>Romano (1993)</td>
</tr>
<tr>
<td>drainage from initially saturated profile</td>
<td>profile $h$ &amp; $\theta$</td>
<td>$\theta(h)$, $K(\theta)$: VGM</td>
<td>lysimeter: sand &amp; loam</td>
<td>Dikiniya (1993)</td>
</tr>
<tr>
<td>drainage from initially saturated profile</td>
<td>profile $\theta$</td>
<td>$\theta(h)$, $K(\theta)$: Campbell</td>
<td>field data</td>
<td>Ross (1993); Ross and Parlange (1994)</td>
</tr>
<tr>
<td>ponded infiltration</td>
<td>cumulative infiltration &amp; $\theta(h)$</td>
<td>$\theta(h)$: VG $K(\theta)$: Gardner</td>
<td>field, clay soil</td>
<td>Bohne <em>et al</em>. (1992)</td>
</tr>
<tr>
<td>ponded infiltration</td>
<td>cumulative infiltration &amp; $\theta(h)$</td>
<td>$\theta(h)$, $K(\theta)$: VGM</td>
<td>numerical study &amp; field soil</td>
<td>Bohne <em>et al</em>. (1999)</td>
</tr>
<tr>
<td>precipitation and evaporation</td>
<td>$h$ &amp; $\theta$ for different depths</td>
<td>$\theta(h)$, $K(\theta)$: VGM</td>
<td>field study: loamy soil (5 horizons) sandy soil (4 horizons)</td>
<td>Arning (1994)</td>
</tr>
<tr>
<td>Field Water Regime</td>
<td>( \theta )</td>
<td>( K(h) ):</td>
<td>Field Data (Numerical Study) With Unknown Boundary Condition</td>
<td>Author(s)</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>--------------</td>
<td>-------------------------</td>
<td>---------------------------------------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Field Water Regime</td>
<td>( \theta )</td>
<td>Power Function</td>
<td>Numerical Study &amp; Field, Layered Profile</td>
<td>Lehmann and Ackerer (1997)</td>
</tr>
<tr>
<td>Precipitation and Evapotranspiration</td>
<td>Profile ( h )</td>
<td>( \theta(h), K(\theta) ): VGM</td>
<td>Field, Silt Loam</td>
<td>Inoue et al. (1998)</td>
</tr>
<tr>
<td>Multiple Step Soil Water Extraction</td>
<td>Cumulative Water Extracted &amp; ( h )</td>
<td>( \theta(h), K(\theta) ): VGM</td>
<td>Laboratory &amp; Field, Sandy Soil</td>
<td>Kodešová et al. (1998, 1999)</td>
</tr>
<tr>
<td>Cone Permeameter</td>
<td>Cumulative Infiltration &amp; ( h )</td>
<td>( \theta(h), K(\theta) ): VGM</td>
<td>Laboratory &amp; Field, Sandy Soil</td>
<td></td>
</tr>
<tr>
<td>Water Flux into Tensiometer</td>
<td>Flux into Tensiometer</td>
<td>( \theta(h), K(\theta) ): Campbell</td>
<td>Field, Loamy Sand</td>
<td>Timlin &amp; Pachepsky (1998)</td>
</tr>
<tr>
<td>Multi-Step Drainage in Lysimeter</td>
<td>Profile ( h ) &amp; ( \theta )</td>
<td>( \theta(h), K(\theta) ): VGM</td>
<td>Field</td>
<td>Abbaspour et al. (1999)</td>
</tr>
<tr>
<td>Infiltration &amp; Drainage</td>
<td>Profile ( h ) &amp; ( \theta )</td>
<td>( \theta(h), K(\theta) ): VGM</td>
<td>Field, Profile with 4 Layers</td>
<td>Abbaspour et al. (2000)</td>
</tr>
<tr>
<td>Constant Flux Infiltration</td>
<td>Soil Water Storage</td>
<td>( K(\theta), D(\theta) ): Broadbridge &amp; White</td>
<td>Numerical &amp; Field Study</td>
<td>Si &amp; Kachanoski (2000)</td>
</tr>
</tbody>
</table>

**Laboratory Studies**

| Upward Infiltration in Soil Cores   | \( h \) & \( \theta \) | \( \theta(h), K(\theta) \): VGM | Laboratory, Loamy Fine Sand                                   | Hudson et al. (1996)              |
| Constant Flux Infiltration          | \( \theta \) from Gamma-Ray Attenuation | \( \theta(h) \): VG, D(\( \theta \)): Exponential | Laboratory Column of a Loam                                   | Barataud et al. (1996)            |
| Syringe Pump Water Extraction & One-Step Outflow | Cumulative Outflow | \( \theta(h), K(\theta) \): VGM | Laboratory Cores, Coarse to Fine Sands                       | Windenschild et al. (1997)        |
| Evaporation                         | \( h \) & \( \theta \) | \( \theta(h) \): VG, K(\( \theta \)): Exponential | Laboratory Cores                                              | Ciollaro and Romano (1995), Santini et al. (1995) |
| Evaporation & Multi-Step Outflow    | \( h \) & \( \theta \) | \( \theta(h), K(\theta) \): VG, MVG | Lab, Swelling Clay                                            | Garnier et al. (1997)             |
| Multi-Step Outflow                  | \( h \) & \( \theta \) | \( \theta(h), K(\theta) \): VGM | Glass Beads & Sand                                            | Nützmann et al. (1998)            |
| Horizontal Infiltration             | \( \theta \) | \( \theta(h), K(\theta), D(\theta) \): VGM | Columbia Silt Loam                                            | Simunek et al. (2000)             |

^1: For hydraulic models VG = van Genuchten, MVG = Mualem-van Genuchten, BC = Brooks-Corey.

A promising laboratory procedure is the evaporation method, where initially saturated soil core is allowed to dry and the mass loss and matric potential is
monitored periodically (Wind, 1963). Feddes et al. (1988) employed the inverse method to estimate \( K(\theta) \) from an evaporation experiment and compared it with the standard analysis of Wind’s method (Wind, 1968). Santini et al. (1995) used the inverse method in evaporation experiments and showed good comparison with independently measured retention data and \( K_r \). Although Halbertsma (1996) suggested that the evaporation method may not be suited for inverse analysis due to a non-unique solution, Šimunek et al. (1998c) obtained good agreement between \( \theta(h) \) and \( K(\theta) \) estimated by the inverse method and the Wind’s method. They also showed that similar results could be obtained with only one set of tensiometer readings contrary to Wind’s method, which requires pressure head measurements at several locations. Garnier et al. (1997) applied this method to a swelling soil with measurements of linear and horizontal shrinkage during evaporation. An inverse solution was applied to the data considering anisotropic deformation of the soil. Romano and Santini (1999) presented a numerical analysis on the sensitivity and stability of the evaporation inverse solution.

Although laboratory experiments have the advantage of being easy to implement, they could produce hydraulic properties that are not necessarily representative in the field (Kasteel et al., 1999). Hence, considerable attention has also been given to estimate in situ soil hydraulic properties. Many studies are still based on numerical experiments as presented by Kool and Parker (1988). Bohne et al. (1992) investigated the plausibility of obtaining hydraulic properties from ponded infiltration using Philip’s two-term equation. Utilising the analytical solution of Barry et al. (1995), Bohne et al. (1999) successfully predicted \( K(\theta) \) from cumulative infiltration and measured \( \theta(h) \). Feddes et al. (1993) analysed the possibility of obtaining effective hydraulic properties of a catchment. In a numerical study, they generated 32 soil samples from a catchment having different hydraulic properties and performed soil water simulations on the 32 samples. The simulated water content of the profiles and the surface fluxes (precipitation and evaporation) were averaged and fed into an inverse model to predict the effective hydraulic parameters. They concluded that it is possible to derive the average hydraulic parameters by an inverse method using areal evapotranspiration and surface soil moisture.

Special instruments have been devised to obtain transient flow measurements for inverse procedures, such as the cone permeameter (Kodešová et al., 1998) and soil
water extraction instrument (Inoue et al., 1998). Inverse methods have also been applied to estimate parameters from solute transport (Parker and van Genuchten, 1984). Mishra and Parker (1989) also explored the possibility of estimating simultaneously hydraulic and solute transport properties from transient flow with applied tracer. Inoue et al. (2000) investigated an infiltration using NaCl solution in a soil column where transient pressure head, apparent electrical conductivity, and effluent flux were monitored continuously. They used this information to simultaneously estimate water retention, unsaturated hydraulic conductivity, and solute dispersion coefficient.

Computer programs for computing the inverse solution have been developed. Aside from the one-step to multi-step program, Kool and Parker (1987) developed a general inverse program, SFIT, to estimate the van Genuchten parameters from one-dimensional transient soil-water flow experiments, which also considered hysteresis. Šimunek et al. (1999b) extended their one-and two-dimensional finite element soil-water flow into HYDRUS-1D and HYDRUS-2D and incorporated inverse procedures. Šimunek and van Genuchten (2000) recently presented a program called DISC for the inverse solution of disc permeameter measurement.

3.3. Inverse method from disc permeameter measurements

Šimunek and associates from the US Salinity Laboratory proposed inversion of field disc permeameter measurements for estimating \( \theta(h) \) and \( K(\theta) \) relationships. In a purely numerical study, Šimunek and van Genuchten (1996) first explored the feasibility of performing the inverse solution from single tension disc infiltration data. Coupling the numerical two-dimensional axi-symmetric solution of Richards’ equation with a nonlinear least-squares procedure, they found that cumulative infiltration data from single tension did not provide enough information to estimate the hydraulic parameters. Additional information such as water content and pressure head measurements over time are also needed. To avoid additional measurements, Šimunek and van Genuchten (1997) proposed the use of multiple tension infiltration data. Nevertheless, cumulative infiltration data only may not produce a well-posed inverse solution. Additional measurements collected in a typical disc infiltration study can provide useful information, such as: final water content underneath the disc, unsaturated hydraulic conductivity calculated from Wooding’s analysis, and defining
the initial condition in terms of water content. They applied this method on the disc permeameter data collected from the soil hydrology program of the Hydrologic Atmospheric Pilot Experiment (HAPEX)-Sahel regional-scale experiment (Šimunek et al., 1998a). They found ‘very close agreement’ between the measured and calculated cumulative infiltration data. Although the true $\theta(h)$ and $K(\theta)$ were never actually measured, they obtained ‘relatively good agreement’ between the predicted curves with the neural network PTFs of Schaap et al. (1998).

Two disc permeameter measurements were carried out on a fine sandy loam using consecutive tensions of -200, -100 and -30 mm (Šimunek et al., 1998b). ‘Excellent agreement’ was found between the measured and simulated cumulative infiltration using the inverse method. Hydraulic conductivity determined by the inverse method corresponded well with those calculated from Wooding’s solution. However water retention estimated from the inverse solution showed very poor agreement with those measured in the laboratory. The predicted retention curve was about 10% lower than those measured in laboratory. The infiltration measurement from the first run was also higher than the second one. Simulations using the soil hydraulic parameters as determined in the laboratory did not correspond with the field infiltration data. They indicated that different laboratory and field measurements may give different results for the hydraulic properties.

Šimunek et al. (1999a) used the multiple tension disc on a repacked loamy soil equipped with sixteen tensiometers installed at eight positions underneath the disc and a TDR probe ~2 cm beneath the disc. The objective function for the inverse method was defined using different combinations of the cumulative infiltration, water content, and pressure head measurements. Similar to previous results, hydraulic conductivity from the inverse method corresponded well with Wooding’s analysis. The drying branches of the $K(h)$ function determined by inverse method also correspond relatively well with those obtained by the evaporation method. However, retention curves estimated from the inverse method are still different from those determined using the evaporation and pressure chamber method (although the laboratory methods represent a drying curve, while the infiltration represents a wetting curve). Nevertheless, they argued that the properties predicted by the inverse method may be more useful in describing infiltration and the transport of contaminants through the unsaturated zone.
Jacques et al. (1999) took a different approach, instead of using a numerical model, they utilised the semi-analytical solution of Haverkamp et al. (1994):

\[
I = S_0 t^{\alpha} \left[ K_i + \frac{\gamma S_0^2}{r_0 \Delta \theta} + \frac{1}{3} (K_0 - K_i) (2\kappa) \right] t \tag{3.3.1}
\]

Sorptivity and \( \kappa \) are related to \( D(\theta) \) and \( K(\theta) \) relationship and the proportionality factor \( \gamma \) represent the gravity effects at the profile edges and calculated from the difference between infiltration in three- and one-dimension:

\[
I_{3D} - I_{1D} = \frac{\gamma S_0^2}{r_0 \Delta \theta} t \tag{3.3.2}
\]

In a numerical study they found that \( \gamma \) values increased as the supply potential \( h_0 \) is increased. Difficulties are encountered when optimising the parameters and calculating the \( \gamma \) factor. When applied on multiple tension measurements, \( \gamma \) was found to increase with decreasing \( h_0 \). Therefore a penalty factor is introduced in the objective function, whenever \( \gamma \) values violate the true relationship, the objective function is increased by a factor of 10. They concluded that, in order to obtain a well-posed solution at least two infiltration curves should be used with some \( \theta(h) \) measurements, parameter space with penalty factor added to the objective function is screened a priori to obtain good initial estimates.

These studies have shown that inverse methods work perfectly well when applied to numerically generated data. However when applied to real measurements, large discrepancies are found in the water retention curves. The infiltration curve is also not reproducible suggesting experimental shortcomings. The studies by Šimunek and associates have not satisfactorily provided an efficient method (installing sixteen tensiometers beneath the disc is as complicated and time consuming as laboratory measurements).

### 3.4. Formulation and solution of inverse problem

The inverse problem for estimating hydraulic properties requires the following steps: 
1. parameterization of the hydraulic properties; 
2. solution of the forward problem (soil-water flow processes); 
3. formulation of a performance criterion and 
4. optimization procedure to estimate the parameters. Parameterization of water retention and hydraulic conductivity curve have been proposed in closed-form
equations by many authors and will not be discussed here. The second and third point will be discussed in section 3.4.1, and the fourth point in section 3.4.2.

### 3.4.1. Mathematical formulation

The general water transport equation in the soil according to Richards (1931) is:

\[
\frac{\partial \theta}{\partial t} = \nabla (K(\theta) \nabla H)
\]  

where \( \frac{\partial \theta}{\partial t} \) is the change of water content over time, \( \nabla \) (del) is the gradient operator or derivative with respect to space \((x, y, z)\), \( K \) is the hydraulic conductivity, \( H \) is the total soil water potential. For one-dimensional water flow on \( x \) direction:

\[
\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial x} \left[ K(\theta) \left( \frac{\partial h}{\partial x} - 1 \right) \right]
\]

Under certain initial and boundary conditions, this equation can be solved numerically using finite-difference or finite-element approaches (Remson et al., 1971). The \( \theta(h) \) and \( K(h) \) relationships can be represented by closed-form equations such as those of van Genuchten (1980). Parameters of the model are represented by vector \( \beta \), for van Genuchten equation \( \beta = [\theta_s, \theta_r, \alpha, n, K_s, l] \).

Suppose the response of a flow process is measured at \( N \) time steps, the observed inputs to the system at time \( t \) is represented by \( X_t \) and observed responses by \( q_t \), a model could be written as:

\[
q_t = \hat{q}(X_t, \beta) + \varepsilon_t, \quad \text{for } t = 1, 2, ..., N.
\]

The input \( X \) is fed into a model which produces a predicted response \( \hat{q}(X, \beta) \) with random errors \( \varepsilon \) added to the observed response. The errors are assumed to be independent, normally distributed with the expected value zero and constant variance \( \sigma^2 \). The maximum likelihood of the function could be formulated as:

\[
L = \frac{1}{2} \ln(2\pi\sigma^2) + \frac{1}{2} \left( \frac{q - \hat{q}}{\sigma} \right)^2
\]

where \( L \) is the negative log-likelihood (Gershenfeld, 1999). The problem is to find values of parameter \( \beta \) that minimise \( L \). The last term on the right hand side measures the distance between the data and the model, while the first term ensures that a model with huge variance will not dominate. The first term can be dropped since it is not relevant to \( \beta \). Therefore Equation [3.4.4] reduced into:
\[ O(\beta, \hat{q}) = \sum_{i=1}^{N} \left[ w_i \left( q_i - \hat{q}(x_i, \beta) \right) \right]^2 \]  

where the \( w \) is the weight given to the residuals \( \sim 1/\sigma_e \). If we do not know what the variance should be, we assume that \( \sigma_e \) is constant at unity and independent of each other, which reduces to the ordinary least-squares error model (Gershenfeld, 1999). Equation [3.4.5] is usually called weighted least-squares. The objective function is the criterion to be minimised in the inverse procedure.

The objective function in transient-flow experiments usually contains different sets of measurements (e.g. cumulative infiltration with water content and pressure head), and different weights should be assigned to data. Therefore a general objective function commonly used in inverse estimation of soil hydraulic properties are (Šimunek and van Genuchten, 1997):

\[ O(\beta, \hat{q}_k) = \sum_{j=1}^{k} \left[ v_j \sum_{i=1}^{n_j} \left[ w_i \left( q_j(t_i) - \hat{q}_j(t_i, \beta) \right) \right]^2 \right] \]  

where \( k \) represents different sets or types of measurements, \( n_j \) is number of measurements in a particular set; \( q_j(t_i) \) are measurements at specified time \( t_i \); \( \hat{q}_j(t_i, \beta) \) are the model predictions, \( v_j \) and \( w_i \) are the weights associated with a particular measurement set or point.

Hollenbeck and Jensen (1998b) criticized the lack of statistical basis when applying the weighting scheme in the inverse solution. Often \( w \) is set to unity and \( v \) is to some constant that is related to the magnitude of the data for each set (van Genuchten et al., 1991). The weight usually used for \( v \) for a particular data set is:

\[ v_j = \left[ n_j \sigma_j^2 \right]^{1/2} \]  

which normalizes the measurement by its variance \( \sigma_j^2 \) (Clausnitzer and Hopmans, 1995). When transient flow data alone are not sufficient to yield a well-posed solution, they are usually combined with other information, such as water retention. For example a combined objective function for outflow data \( O_\delta(\beta) \) is coupled with the objective function of water retention curve \( O_\theta(\beta) \) weighted by \( v \):

\[ O(\beta) = O_\delta(\beta) + v O_\theta(\beta). \]  

It is only meaningful to include more than one type of data in the objective function when all the estimators ‘point’ at the same estimates. For example, the \( \theta(h) \) and \( K(\theta) \) of the van Genuchten model share common parameters \( \alpha \) and \( n \). If the same parameter...
were valid for both functions, the estimates for Equation [3.4.8] should yield the same estimates for any \( v \). However, Abbaspour et al. (1997) reported different estimates can be obtained when different values of \( v \) were assigned to different data sets. The addition of the \( v \) weighing factor is superfluous and can distort the estimation results (Hollenbeck et al., 2000).

Different criteria applied to the objective function will influence the estimated results. In flow experiments, usually time lags can occur between the observed and simulated time courses (Richter et al., 1996). In this case, the application of the least-squares criterion will result in large residual and bad performance of optimization procedures. Richter et al. (1996) suggested alternative measures of the objective function to be minimised. One possibility is to use the Euclidean distance between measurements and simulated curve.

\[
O(\beta, \hat{q}) = \sum_{i=1}^{N} w_i \left( t_i - \hat{t}(\beta) \right)^2 + (q_i - \hat{q}(\beta))^2 \]  

[3.4.9]

where \( \hat{t}(\beta) \) is the predicted time. The disadvantage of this method lies in the case where large amplitudes at small time scale are measured and can minimise the wrong deviation (the first term on the right hand side dominates the second term). Another criterion is the difference of the areas between the simulated and measured curves,

\[
O(\beta, \hat{q}) = \sum_{i=1}^{N} w_i \left[ \frac{1}{t_N - t_1} \int_{t_1}^{t_N} |q_i - \hat{q}(\beta)| \, dt \right] 

[3.4.10]

Figure 3.4.1. Comparison of different objective function criteria (Richter et al., 1996)

It can be recognized that this criterion (Equation [3.4.10]) is the same measure as proposed by Skaggs et al. (1971) (Equation [3.1.1]) and also the RMSD of water
Chapter III – Inverse methods

Retention PTFs. Arning (1994) as quoted by Richter et al. (1996) found that this criterion gives the best results when used in problems with time lags.

Two categories of inverse solution can be distinguished based on their solution (DuChateau, 1996):

1. Methods which focus on the input/output mapping and attempt to formulate inversion from this map. Or inversion with an analytical solution. When the direct problem can be solved analytically, the parameters of the unknown properties are sought and formulated in explicit or implicit form. The solution could be linear or linearized (solved by matrix inversion) or nonlinear (solved by algebra or calculus). Examples of such methods are provided in DuChateau (1996).

2. Inversion via optimization. The solution of the forward problem could be an analytical (such as infiltration solution by Barry et al., 1995) or numerical model. The unknown parameters are sought by optimization procedures, which minimised the difference between the observed and predicted values. This method is generally employed in soil science.

3.4.2. Optimization procedures

The objective function is usually minimised by the nonlinear optimization algorithm. The most frequently used is the nonlinear least-squares estimation by Levenberg-Marquardt algorithm (Marquardt, 1963). The mathematical description and algorithm is presented in Appendix 3.1. Other nonlinear least-squares methods include: Gauss-Newton (Gill et al., 1981), and quasi-Newton method (Zijlstra and Dane, 1996).

The nonlinear least-squares approach assumes that it is possible to calculate the first derivative of the function to be minimised, the starting parameter values are near the desired function, and the function is reasonably smooth. In practice, these often do not apply, therefore an alternative approach known as global optimisation is introduced. Global optimization attempts to find the best parameters to achieve the objective function and can deal with problems with (possibly) several local minima, but usually this approach requires more model evaluation. Various optimization methods differ in deciding which direction the parameters have to go in order to reach an optimum (minimum) value. The methods include downhill simplex (Nelder and Mead, 1965), tabu search (Glover and Laguna, 1997) and methods that are based on natural phenomena. Simulated annealing (Kirkpatrick et al., 1983) simulates the annealing process where a substance is heated then cooled gradually to produce
crystal, which minimises its energy probability distribution. The genetic algorithm (Goldberg, 1994) models natural selection where a population is evolved under certain selection rules to maximise the fitness. Abbaspour et al. (1997) proposed a sequential uncertainty domain inverse procedure, which searches for the best parameter by sampling the respective parameter space. Further mathematical description of the optimization methods can be found in the world-wide-web mathematical programming glossary: http://www.cudenver.edu/~hgreenbe/glossary/glossary.html.


3.5. Some definitions

In the course of reviewing the application and results of the inverse methods, many obscure terms have been noted. The following are the description of the terms normally quoted in the ‘inverse’ literature.

3.5.1. Response surface of parameter space

The parameter space is a subspace of $np$-dimensional parameter space consisting of the set of $\beta$ values which is usually plotted in two-dimensional axis. The response surface refers to the relationship between parameters and the objective function. Each pair of parameters are discretized around their true values, while the other parameters are kept constant at the true value. Numerical simulation is performed for each combination of parameters, and the objective function is calculated. The objective function is then projected onto a contour plot.

The response surface indicates the points where the objective function is the smallest, hence the optimum values for the parameters. The response surface is useful in identifying the existence and uniqueness of the solution, whether multiple minima exist, and the sensitivity of the parameters to the objective function. Stable parameters are justified by compact elliptical contours appearance (Figure 3.5.1a). The response surface provides insight on the shape of the parameters’ distribution, correlation and confidence regions of the parameters (Press et al., 1992). Examples of their use are

![Figure 3.5.1. Examples of response surface of objective function as a function of two parameters (a) stable solution with compact elliptical contours and well-defined minimum (b) a non-stable solution with many local minima and banana shaped contours](image)

### 3.5.2. Ill-posedness

Given a functional relationship between the response or measurement of the flow system \( R \) and the parameters describing the hydraulic properties \( \beta \), \( R = F(\beta) \), we wish to determine an inverse relationship of \( \beta \) that describe \( R \), \( \beta = F^{-1}(R) \). The inverse problem is said to be well posed if it has all the properties of (Carrera and Neuman, 1986):

- **existence**: the solution exists,
- **uniqueness**: there is only one solution,
- **stability**: the solution does not change with slight modification of data.

The terms identifiablity, uniqueness and stability are usually quoted. *Identifiability* refers to the forward relationship \( F \). If more than one parameter set \( \beta \) leads to a similar response, the parameters are unidentifiable. *Uniqueness* refers to the inverse relationship \( F^{-1} \). In the optimization procedure which minimised an estimation criterion, the solution is non-unique if it has local minimum or more than one global minimum in the parameter space. In other words, a given response leads to more than one set of parameter values. *Stability* means that small errors in the response must not result in large changes in the estimated parameters.
To illustrate this concept, consider Figure 3.5.2, which is a simplification of the analyses presented by Zachmann and White (1991). The figures show the plot of parameter $\beta$ on the x-axis and the response $Y$ on the y-axis, we seek for the value of $\beta$ that correspond to two responses $F(\beta)$ and $G(\beta)$. The solution exists when $F(\beta) = G(\beta)$ i.e., when the two curves intercept.

Figure 3.5.2. Identification of $\beta$ from two type of responses $F(\beta)$ and $G(\beta)$.

Figure 3.5.2a shows the case where there is no solution and the response to the system is virtually the same for different values of $\beta$ (unidentifiable). In the case of Figure 3.5.2b although the system responds distinctly for different $\beta$ (identifiable) there is still no solution. The solution is unique when there is only one intercept (Figure 3.5.2c). If the responses are as in Figure 3.5.2d, the solution is non-unique because three intercepts can be found and the solution is not stable, or very sensitive to the change in the response curve.

Non-identifiability relates to the conditions where different parameter sets lead to a similar flow response, while non-uniqueness is concerned with conditions where different parameters originate from a given flow measurement. Identifiability can be analysed by examining the forward relationship, but uniqueness cannot be fully resolved without first speculating on the inverse relationship. A mathematical analysis on the identifiability is provided by Mous (1993). A non-identifiable system is shown
in Figure 3.2.4b, where different boundary conditions applied to the system created indistinguishable responses.

The uniqueness of the inverse solution depends on the structure, extent and quality of the data (Zilstra and Dane, 1996). The structure of data refers to specific positions and times at which the data are collected. Knopman and Voss (1987) pointed out that the best parameter estimates result from sampling over the intervals where the data are most susceptible to changes in parameter values (further discussion of this can be found in sensitivity analysis section). The extent of the data reflects the range of the flow response. For example, data that represent a narrow range of change in water content or potential, such as drainage in a short column, may lead to ill-posedness and possibly non-uniqueness of the solution. The quality of data is reflected by measurement error. If the measurement error is known, the weighting of data based on the variance of the measurement may help to overcome the effects of data error. If no prior information about the parameters is available, the solution is non-unique. The existence of prior information will, in many cases, promote uniqueness.

Figure 3.5.3. Response surface of the objective function for the Gardner-Russo model for parameters $K_s$ and $\alpha_g$ (a) small error in infiltration data and no prior information on $K_s$ and $\alpha_g$ (b) moderate error and no prior information (c) moderate error and prior information on $K_s$ only (d) moderate error and prior information $\alpha_g$ only.
As an example consider the inverse solution of the Gardner-Russo hydraulic model from infiltration measurement (Russo et al., 1991) in Figure 3.5.3. If no prior information on the parameters on the model is available, relatively small error introduced to the data cause the inverse solution to be non-unique as shown by the multiple minima (Figure 3.5.3a & b). But when prior information on one of the parameter exists, the inverse solution is well behaved (Figures 3.5.3c and d).


It should be noted that ill-posedness does not imply that an inverse problem is meaningless. It merely indicates that the problem formulation must be modified (Mclaughlin and Townley, 1996). Ill-posedness can also be reduced by limiting the number of optimized parameters. Non-uniqueness can be determined posteriori by solving the inverse problem repeatedly using different initial parameter estimates. From their study, Russo et al. (1991) concluded that:
- identifiability depends on the structure of the hydraulic models;
- uniqueness and stability depend on the structure of the hydraulic model, quality of prior information on model parameters and magnitude of measurement error;
- despite the usefulness of prior information inclusion, it is not sufficient to guarantee uniqueness and stability, or it is impossible to determine a priori whether the inverse problem is well posed or not.

3.5.3. Sensitivity analysis

The experimental design via sensitivity analysis as proposed by Knopman and Voss (1987) is intended to choose the time, location, type and frequency of measurements to maximise the value of data collected. This analysis is usually carried out on numerical examples prior to experiments. The analysis provides insight on the sensitivity of each observation with respect to the parameter. An experiment designed for parameter optimization should include measurements that are most sensitive to changes in the parameters. Šimunek and van Genuchten (1996) have carried out such analysis for disc permeameter measurement, Šimunek et al. (1998c) for evaporation method, and Inoue et al (1998) for multi-step extraction.
Sensitivity is evaluated from the response of the system due to the change in the parameter values, the sensitivity of measurement at time $q_i$ with respect to parameter $\beta_j$ can be evaluated from the parameter’s partial derivative:

$$\frac{\partial q}{\partial \beta} = \frac{q(\beta + \Delta \beta) - q(\beta)}{\Delta \beta}.$$ \[3.5.1\]

where $\Delta \beta$ is a small change in parameter $\beta$. Šimunek and van Genuchten (1996) suggested that sensitivity coefficient is evaluated as:

$$S_{i,j} = q_i(\beta_j + \Delta \beta_j) - q_i(\beta_j).$$ \[3.5.2\]

where $S_{i,j}$ represents the change of the measurement variable ($q_i$) relative to a small change of the parameter $\beta_j$. It should be noted that high sensitivity is necessary but not a guarantee to provide a well-posed problem.

As an example, Figure 3.5.4 shows the sensitivity analysis of the infiltration from the disc permeameter by Šimunek and van Genuchten (1996). Water content at a point 10 cm beneath the disc is monitored during infiltration. During movement of the wetting front to the measurement point, water content is very sensitive to $\alpha$ and $K_s$ but not to $n$ and $\theta_s$. After the wetting front has passed this point, the sensitivity to $\alpha$ and $K_s$ decreases gradually, in contrast the sensitivity to $\theta_s$ increases. This is because the soil water content is approaching saturation.

Figure 3.5.4. Sensitivity of the water content at observation 10 cm beneath the disc to a 1% change in hydraulic parameters $\alpha$, $n$, $\theta_s$, and $K_s$ (Šimunek and van Genuchten, 1996).

The sensitivity analysis informs us that (Šimunek and van Genuchten, 1996) observed data should be taken at points in time and space which show the highest
sensitivities to the hydraulic parameters and which are accessible to measure. The inverse solution will not be improved by including in the objective function data which are insensitive to the parameters to be optimised. For example, including pressure head data measured before the arrival or after passage of the moisture front will not improve the identifiability of the hydraulic parameters since they will lead to similar predicted pressure heads during those times.

3.5.4. Influential observations

The sensitivity analysis provides insight on the influence of observations in time and space on the parameters sensitivity. In an experiment some unusual observations can unduly influence the results of the inverse solution. The effects of such observations, called influential observations, cannot be overlooked. An outlier is an observation whose dependent or independent variable is unusual compared to other values, but an outlier is not necessary an influential observation. Consider the infiltration rate curve in Figure 3.5.5 where Philip’s two term model \( i = \frac{1}{2} S_0 t^{-1/2} + A \) was fitted to the data. In Figure 3.5.5a, the outlier is at the middle of observation; as a consequence, deletion or inclusion of this point has small effect on the estimated \( S_0 \) and \( A \). Because the unusual observation is at the end of the measurement period (Figure 3.5.5b), it has a strong leverage on the parameters. This situation may or may not be desirable, if the outlier in 3.5.5b is an error its presence could bias the estimated parameters, but if it is valid deleting this point would lower the estimate of steady-state infiltration rate \( A \). The inclusion or deletion of influential observations could affect the accuracy of estimated parameters.

![Outlier and influence in infiltration rate curve](image-url)
Statistical regression diagnostics developed for linear regression can be used to detect influential observations in nonlinear regression if the models are approximately linear. The influential data can be measured by examining the impact of deleting each observation in turn on the parameter estimates. One such measure is Cook’s $D$ (Cook and Weisberg, 1982), which calculates the influential observation as:

$$\text{Influence on parameter} = \text{Discrepancy} \times \text{Leverage}$$

The influence is not only measured by the residual (discrepancy), but also the extent to which observation $q_i$ can affect the estimated $\hat{q}_i$. A more detailed mathematical description can be found in Cook and Weisberg (1982).

Identification of influential observations is important in detecting which measurements are important or significant for the inverse solution, whether the observations suffer from error and providing guidance for reducing number of observations. Yager (1998) applied Cook’s $D$ in groundwater inverse modelling, although this measure is designed for a linear model, it is useful to detect influential observations.

### 3.5.5. Parameters correlation

The variance-Covariance matrix of the parameters $\mathbf{V}$ is defined as:

$$\mathbf{V} = E \left[ \left( \mathbf{\bar{b}} - \hat{\mathbf{b}} \right)^{\top} \left( \mathbf{\bar{b}} - \hat{\mathbf{b}} \right) \right]$$  \hspace{1cm} [3.5.3]

where $\mathbf{\bar{b}}$ is the ‘true’ parameters, $\hat{\mathbf{b}}$ is the estimated parameters and $E$ is the mathematical expectation. Approximation for nonlinear regression is represented as:

$$\mathbf{V} \approx \frac{\text{SSR}}{N - np} \left( \mathbf{J}^\top \mathbf{J} \right)^{-1}$$  \hspace{1cm} [3.5.4]

where SSR is the sum-of-squared residuals and $\mathbf{J}$ is the Jacobian (parameters’ partial derivative) matrix. The covariance-matrix provides information regarding the reliability of the estimated parameters. A well-estimated parameter is characterized by a small variance. The correlation matrix can be obtained by dividing the covariance between two variables by the standard deviation of each variable.

$$\text{Corr}(\beta_1, \beta_2) = \frac{\text{Cov}(\beta_1, \beta_2)}{\sqrt{\text{Var}(\beta_1) \text{Var}(\beta_2)}}$$  \hspace{1cm} [3.5.5]

It is generally believed that strongly correlated parameters can cause difficulty in optimization convergence and lead to an ill-posed problem (Yeh, 1986; Finsterle and...
But this is not true (Ratkowsky, 1990), convergence and ill-posedness are related to the shape of the response surface. The closer the shape to ellipsoidal, the better the posedness and faster the convergence. Convergence and ill-posedness are not related to the orientation of the ellipse, which is the factor that determines parameter correlation. However, high parameter correlation is sometimes indicative of over-parametization (too many parameters to be optimized), which may cause a convergence problem. It is indeed desirable to obtain a good estimation and also low parameter correlations, however this cannot always be accomplished.

### 3.5.6. Parameter confidence intervals

It is usually desirable to obtain a measure of the precision of the parameters that were estimated. Statistical techniques for obtaining these measures include the use of confidence regions and confidence limits. Confidence regions are used to calculate a joint probability of the set of parameters, while confidence intervals apply to individual parameter. Techniques for obtaining confidence regions are discussed by Hollenbeck and Jensen (1998b) and Hollenbeck et al. (2000). Cooley (1997) compared different methods for obtaining confidence intervals for ground-water inverse models. Reliable estimates of parameter confidence intervals may indicate which model corresponds better to the observations and which parameters can be estimated (van Dam, 1996).

The approximate standard error (SE) of the estimated parameters can be calculated based on the linear least-squares method:

$$ SE(\hat{\beta}_j) = (V_{jj})^{1/2} \quad [3.5.6] $$

where $V_{jj}$ is the diagonal component of the variance-covariance matrix of the parameters. The confidence intervals for each parameter can be calculated as:

$$ \hat{\beta}_j \pm t_{(N-np)}SE(\hat{\beta}_j) \quad [3.5.7] $$

where $t$ is the t-value with $N-np$ degrees of freedom at significance level $\alpha$. This method is simple and is provided in many optimization programs, but unless the model is close to linear the limits may be incorrect.

Better methods, requiring more computation are the log-likelihood method (Cooley, 1997) and profile log-likelihood method (Cook and Weisberg, 1990). Intensive computational methods such as the bootstrap can also be utilised (Efron, 1987). This requires sampling with replacement of some percentage the original data.
and refitting the model. A large number of runs are required to obtain a good estimate of the parameters distribution. For inverse methods that require heavy computation, this method is limited. van Dam (1996) evaluated the linearized method to estimate the confidence intervals for van Genuchten parameters from a multi-step outflow experiment. The linearized method provides larger confidence intervals but show no significant differences compared to the Monte Carlo simulation.

3.5.7. Prediction confidence intervals

Similarly if we wish to estimate the accuracy of the predicted values, we need to evaluate the confidence limit of the prediction. The most common method employs the first-order error analysis. The predictive model is linearised to give the variance of the prediction:

$$\text{Var}(\hat{q}) = \sum_{j=1}^{np} \sum_{k=1}^{np} \frac{\partial^2 \hat{q}}{\partial \beta_j \partial \beta_k} \text{Cov}(\beta_j, \beta_k)$$  \[3.5.8\]

The confidence limits of the prediction can be calculated: $$\hat{q} \pm t_{\alpha/2} (N - np) \text{SE}(\hat{q})$$.

A more accurate, but more computationally intensive method is the Monte-Carlo simulation (Kuczera, 1994), where the parameters are assumed to have a multivariate normal distribution with variance-covariance matrix V. The parameters are sampled \(ns\) times, and an error was randomly sampled from the responses. The model is evaluated using the sampled parameters and random error is added to the estimated responses. Statistics such as \(100(1-\alpha)\) percentiles can be evaluated from the \(ns\) predicted responses \(\hat{q}\). If the random error is not sampled, the responses only reflect the uncertainty in parameter.

3.6. Conclusions

The inverse methods have gained popular attention for the estimation of soil water retention and unsaturated hydraulic conductivity characteristics. The methods have been explored in many possible ways to provide estimates of hydraulic properties from transient-flow experiments. The one-step outflow method has provided an impetus to the inverse method. The lessons that were learned is to consider the ill-posedness, to characterize the uniqueness of the solution via the response surface of parameters’ space, to provide additional information that may enhance the solution, to look for alternative experimental set-ups, and to investigate the experimental
limitations. Inversion applied to perfect data can give a perfect result, but usually experimental measured data suffer from errors, inaccuracy and incompleteness. This sometimes makes inversion more appealing academically than practically (Claerbout, 1992). Although numerical and statistical analyses have proved that the modifications of one-step into multi-step and inclusion of additional measurements provide a well-posed problem, the experimental limitation of the method is often rarely considered.

Different kinds of measurements sometimes yield different estimates, therefore which estimates provide a ‘true’ representation of the soil in the field has to be scrutinized. Despite these shortcomings, careful measurements and error considerations may yet provide a robust technique.

Some factors that may affect the inverse solutions are: limited measurements in space and time, measurement error, experimental limitations, the flexibility of the hydraulic model in describing both water retention and hydraulic conductivity, numerical discretization of the differential equations, stability of the numerical solution and the ill-posedness of the inverse solution (van Dam, 1996). The advantages of inverse method compared with conventional method are: no simplified assumption in the flow process, high flexibility in defining the initial and boundary conditions, simultaneous estimation of water retention and hydraulic conductivity, estimates of the prediction quality and estimated parameters confidence limits. These provide an insight into the shortcomings of measurements (van Dam, 1996; Poeter and Hill, 1997).
Chapter III – Inverse methods

References


Chapter III – Inverse methods


Chapter III – Inverse methods


Appendix 3.1.

Levenberg-Marquardt nonlinear least-squares

The objective function of nonlinear least-squares method is to minimise the objective function $O$, i.e. sum of squared residuals between the observed data and predicted models:

$$O(\beta, \hat{q}) = \sum_{i=1}^{N} \left[w_i(q_i - \hat{q}(x, \beta))\right]^2$$

where $q = \text{observed data}; \hat{q}$ is the predicted value; $\beta$ is parameters of $j = 1, \ldots, np$; $w$ is weighing factor, $N$ is the number of observed data, and $np$ is the number of parameters. Due to the nonlinearity of the function, there is no single solution for estimation of the parameters. The solution has to be found by an iterative process. Marquardt (1963) proposed an efficient method for parameter estimation, which he termed the maximum neighbourhood method. The method was also proposed earlier by Levenberg (1944) and therefore termed as the Levenberg-Marquardt method.

The Levenberg-Marquardt method estimates the parameters $\beta$ using iteration by solving the following normal equation

$$(H + \lambda I) \delta = g$$

where

- $H = J^T J$ = Hessian or curvature matrix
- $g = J^T r$ = gradient vector
- $r = w(q - \hat{q})$ = weighted residual vector
- $J = w \frac{\partial \hat{q}(\beta)}{\partial \beta}$ = Jacobian matrix
- $\delta$ = parameter correction vector
- $\lambda$ = damping factor which is a positive scalar
- $I$ = identity matrix

Or in mathematical form (Press et al., 1992):

$$H_{jk} = \sum_{i=1}^{N} w_i \frac{\partial \hat{q}_i(\beta)}{\partial \beta_j} \frac{\partial \hat{q}_i(\beta)}{\partial \beta_k} \approx \frac{1}{2} \frac{\partial^2 \hat{q}(\beta)}{\partial \beta_j \partial \beta_k} \quad j = k = 1, \ldots, np$$

$$g_j = \sum_{i=1}^{N} w_i \frac{\partial \hat{q}_i(\beta)}{\partial \beta_j} [q_i - \hat{q}_i(\beta)] \approx -\frac{1}{2} \frac{\partial^2 \hat{q}(\beta)}{\partial \beta_j^2}$$
The method is supposed to be a compromise between two well-known least-squares solutions:

1. The gradient method (also called Steepest descent),
   where the optimum value of parameters $\beta$ are obtained by minimizing $O$ with respect to each parameters. The estimation of the parameter at each iteration is then approximated by the gradient of the function:
   $$\delta = -\left(\frac{\partial O}{\partial \beta}\right).$$
   The method is known to be convergent, but need many iterations.

2. Taylor series method (also called Gauss-Newton),
   by approximating the objective function in a first order of Taylor's series, the minimum of the function is found at $\partial O/\partial \delta = 0$. Thus, the correction vector is found by solving the normal equation: $H \delta = g$. The method converges rapidly once the estimated parameter is 'close' to its optimum.

To visualize it geometrically consider Figure A3.1.1, if we start from point $p$ the algorithm calculates the gradient of the response surface, and using this information it guesses where the optimum parameters lie. The steepest-descent follows the gradient down the objective function surface, while the Gauss-Newton method's direction is almost orthogonal to steepest-descent (Davis, 1983).

The search direction in Levenberg-Marquardt is controlled by $\lambda$. If $O$ is found to decrease at the end of an iteration $\lambda$ is decreased. When $\lambda$ becomes small the method transforms into Gauss-Newton method. If the objective function $O$ does not decrease $\lambda$ is increased and when $\lambda$ has become large the damped curvature matrix becomes diagonally dominant. The equation becomes:

$$\delta \approx \frac{g}{H_{\beta}(1 + \lambda)}$$

which is identical to the gradient method. Increasing $\lambda$ also ensures that eventually the equations become non-singular and also contracting the step making $\delta$ small enough that a new-lower function value can be found (Nash and Walker-Smith, 1984).

If the correlation among the parameters is high, $\lambda$ can be increased to an unreasonably large number. To prevent that, step-factor $\eta$ is introduced. By evaluating the cosine of angle between the correction vector and gradient, we define
the angle as $\phi$. If $\phi < \text{critical angle}(\phi_0)$ (Marquardt suggested $\phi_0 = \pi/4$), do not increase $\lambda$ further but decrease $\eta$ by factor of 2 so that $O^* < O$.

As noted by Marquardt (1963), this method performs an interpolation between the Taylor series method and the gradient method. The algorithm shares with:

- the gradient method, the ability to converge from an initial guess which may be outside the region of other methods;
- the Taylor series method, the ability to close in on the converged values rapidly after the vicinity of the converged values has been reached.

The algorithm can be outlined as follows:

1. Set initial $\lambda$, estimate $\beta$ and calculate objective function $O(\beta)$.
2. Decrease $\lambda$ by the factor DROP.
3. Calculate Jacobian matrix $J$ by using forward difference approximation, form matrix $H$ and $g$ and scale matrix $H$ and $g$ by $\sqrt{H_{\beta\beta}}$.
4. Form the normal equation, $(H + \lambda \ I) \delta = g$, solve for $\delta$ by using matrix inversion procedure and define $\eta = 1$.
5. New parameter: $\beta^* = \beta + \eta \delta$, compute the new objective function $O^*(\beta + \eta \delta)$.
(4) If $O^* > O$, evaluate $\cos\varphi$ (angle between the correction vector $\delta$ and $g$). If angle $\varphi < \text{critical angle}$ (or $\cos\varphi > \text{COSCRIT}$), decrease $\eta$ by 0.5, increase $\lambda$ by factor $\text{BOOST}$ and return to (5), if $\cos\varphi < \text{COSCRIT}$ increase $\lambda$ by factor $\text{BOOST}$ and return to (3).

(5) If $O^* \leq O$, test for convergence. If relative change of parameters $\beta^*$ over $\beta > \text{STOPCR}$, then solution has not converged, return to (2). If $\beta^*/\beta < \text{STOPSQ}$, solution has converged.

(6) Print out results and proceed with post analysis.

The algorithm outlined above is based on the original algorithm as proposed in Marquardt's paper (1963). Marquardt noted the need to scale matrix $H$ and $g$ due to the invariant properties of the normal equations under linear transformation. He suggested an initial $\lambda = 0.01$. Values of DROP = 0.1 and BOOST = 10 were found to be a good choice. However, Lampton (1997) found that additive damping $(H_j + \lambda)$ with a small increment (BOOST=1.5) and a large damping decrement (DROP=0.1) is more efficient for solving difficult problems.

There are many proposed modifications and improvements to the algorithm. Realizing that this is a standard least-squares problem, Osborne (1972) solve the normal equation through matrix transformation (QR decomposition):

$$
\begin{pmatrix}
J \\
\lambda^{1/2} I
\end{pmatrix}
\delta 
\approx
\begin{pmatrix}
q \\
0
\end{pmatrix}
$$

This is implemented in MINPACK of Moré et al. (1980). Another modification is based on the criterion for selection of $\lambda$ at each iteration. Fletcher (1971) proposed heuristic rules for estimating $\lambda$ based on the ratio of actual and predicted reduction of the objective function. Moré et al. (1980) and Clausnitzer and Hopmans (1995) also proposed their modification on determination of $\lambda$ based on similar criterion.

The above algorithm is implemented in FORTRAN which is used in this study, the code is a modification of the nonlinear least-squares program NONLINWOOD presented in Daniel and Wood (1971), as manifested by van Genuchten and associates in programs such as ONESTEP (Kool et al., 1985b), RETC (van Genuchten et al., 1991), CXTFIT (Toride et al., 1995), and others.

The modifications from the original code are:
• Addition of masks and constraints to parameters. Constraints (bounds) are added to ensure that parameters do not lie beyond a reasonable value. During iteration, when a parameter violates the prescribed bound, its value is set to the upper/lower bound. A parameter can be masked (fixed) and not altered in the program by zeroing the elements of masked parameter in matrix $H$ and $g$.

• BOOST (=1.5) and DROP (=0.1) factor for $\lambda$ increment/decrement are added to improve the optimization method.

• The decision to accept the solution from a new objective function $O^*$ compared to the previous one $O$, can be based on simulated annealing (Metropolis et al., 1953):

  \[
  p = \exp\left[-\frac{(O^* - O)}{kT}\right]
  \]

where $k$ is the Boltzmann's constant and $T$ is the temperature which is decreased at each iterations.

• The iteration has to pass 2 consecutive convergence test before termination/converging, this is to prevent early termination or falling into local minima.
\[ \beta^* = \beta + \mu \eta \]

Compute \( \Omega^* \)

\[ \text{rand} < \exp(-\Delta \Omega/kT) \]

Decrease \( T \)

\[ \beta^*/\beta \leq \text{CRIT} \]

Converged

\[ \eta = \eta/2 \]

Figure A3.1.2. Flow Chart for Levenberg-Marquardt Algorithm
IV. Development and evaluation of hydraulic pedotransfer functions for Australian soil

4.1. Soil hydraulic database for Australian soil

4.1.1. Introduction

The first step in developing hydraulic PTFs is establishing a soil hydraulic database, which is a collation of existing soil data that consists of basic soil properties such as particle-size distribution, bulk density, organic matter content, and measurement of water retention and hydraulic conductivity.

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 database HYPRES (Wösten et al., 1999). More general computerized database such as UNSODA (Leij et al., 1996; Nemes et al., 1999b), and GRIZZLY (Haverkamp et al., 1997) has been released publicly. The USDA-NRCS Soil Survey Laboratory has provided a world-wide-web access to their soil survey data (USDA-NRCS, 1997), which contains morphological descriptions, and physical and chemical measurements for more than 23000 profiles in the US.

Meanwhile in Australia there are little published data available and the collation of a national database has just begun (Cresswell et al., 1999; Cresswell and McKenzie, 2000). The databases containing measured water retention and hydraulic conductivity values from Australian soil profiles are also very moderately populated. Few published databases include $K_s$ measurements, for example; Forrest et al. (1985) measured $K_s$ in the laboratory from undisturbed core samples; Geeves et al. (1995) measured near-saturated $K$ in the field using the disc permeameter and predicted $K_s$ from the Mualem-van Genuchten model. Moreover, only a few published Australian papers have reported measurements of unsaturated hydraulic conductivity (Rose et al., 1965; Ollson and Rose, 1978). Unsaturated hydraulic conductivity is usually predicted using the pore-size distribution model (Greacen et al., 1974) or Campbell’s hydraulic model (Bond et al., 1996).
4.1.2. Particle-size conversion

The different classification of particle-size fractions used in Australia compared with other countries presents a problem for the immediate adoption of exotic pedotransfer functions. Nemes et al. (1999a) encountered similar problems when evaluating different textural classes used in European country (Figure 4.1.1).

There are two major textural classifications used in the world, i.e. the International and the USDA/FAO system. The particle-size limits are:

<table>
<thead>
<tr>
<th></th>
<th>International</th>
<th>USDA/FAO</th>
</tr>
</thead>
<tbody>
<tr>
<td>clay</td>
<td>&lt; 2 µm</td>
<td>&lt; 2 µm</td>
</tr>
<tr>
<td>silt</td>
<td>2–20 µm</td>
<td>2–50 µm</td>
</tr>
<tr>
<td>sand</td>
<td>20–2000 µm</td>
<td>50–2000 µm</td>
</tr>
</tbody>
</table>

The first classification, termed the International system, was first proposed by Atterberg (1905), and was based on his studies in southern Sweden. Atterberg (1908) chose 20 µm for the upper limit of silt fraction because particles smaller than that size were not visible to the naked eye, the suspension could be coagulated by salts, capillary rise within 24 hours was most rapid in this fraction, and the pores between compacted particles were so small as to prevent the entry of root hairs. Commission One of the International Society of Soil Science (ISSS) recommended its use at the first International Congress of Soil Science in Washington in 1927 (International Society of Soil Science, 1929). Australia took on this system, and according to Marshall (1947) its equal logarithmic intervals are an attractive feature which is worth maintaining. The USDA adopted its own system in 1938 (Knight, 1937), and the FAO used the USDA system in the FAO-UNESCO world soil map (FAO-UNESCO, 1974) and recommended its use (FAO, 1990).

Generally log-linear interpolation on the cumulative particle-size distribution is used to estimate missing particle-size classes for a given classification (Tietje and Hennings, 1996). This method assumes that particle-size is lognormally distributed (Shirazi and Boersma, 1984). Shirazi et al. (1988) established a conversion table between the USDA texture classifications and the International system. Their conversion assumed a lognormal particle-size distribution assumption within each size fraction (sand, silt, and clay). Buchan (1989) also proposed conversion of the USDA texture to the International system based on a lognormal particle-size distribution model. Rousseva (1997) defined several closed-form exponential and
power functions and investigated the suitability of these models in describing cumulative particle-size distributions. Suitability of the models appeared to be influenced by textural type rather than number of measured points and size ranges. Nemes et al. (1999a) evaluated four different interpolation methods (log-linear interpolation, fitting a Gompertz curve, spline interpolation and similarity method) in order to achieve compatibility of particle-size distributions within the European soil hydraulic database HYPRES. They introduced the similarity procedure, which uses an external reference data set that contains a range of soil materials with seven or eight measured particle-size fractions. The procedure involves searching for soil samples in the external reference that match the particle-size distribution the soil to be interpolated. This procedure was found to give the most accurate interpolations.

<table>
<thead>
<tr>
<th>Country</th>
<th>Particle-size limits (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>2, 20, 200, 2000</td>
</tr>
<tr>
<td>Belgium</td>
<td>2, 10, 50, 100, 200, 500, 1000, 2000</td>
</tr>
<tr>
<td>Denmark</td>
<td>2, 20, 63, 125, 200, 500, 2000</td>
</tr>
<tr>
<td>France</td>
<td>2, 50, 2000</td>
</tr>
<tr>
<td>Germany</td>
<td>2, (6.3), 20, 63, (125), 200, 500, 2000</td>
</tr>
<tr>
<td>Greece</td>
<td>2, 6, 20, 60, 200, 600, 2000</td>
</tr>
<tr>
<td>Italy</td>
<td>2, 20, 60, 200, 600, 2000</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>2, 16, 50, 105, 150, 210, 300, 2000</td>
</tr>
<tr>
<td>Portugal</td>
<td>2, 20, 50, 105, 150, 210, 300, 2000</td>
</tr>
<tr>
<td>Slovac Republic</td>
<td>2, 20, 200, 2000</td>
</tr>
<tr>
<td>Spain</td>
<td>2, 20, 250, 2000</td>
</tr>
<tr>
<td>Sweden</td>
<td>2, 6, 20, 60, 200, 600, 2000</td>
</tr>
<tr>
<td>England &amp; Wales</td>
<td>2, 20, 60, 100, 200, 600, 2000</td>
</tr>
<tr>
<td>Northern Ireland</td>
<td>2, 60, 2000</td>
</tr>
<tr>
<td>Scotland</td>
<td>2, 60, 2000</td>
</tr>
<tr>
<td>USA</td>
<td>2, 50, 100, 250, 500, 1000, 2000</td>
</tr>
</tbody>
</table>

*Figure 4.1.1. Particle-size limits used in different countries (adapted from Nemes et al., 1999)*
Because in routine soil survey data particle size is measured only on 3 to 4 fractions (clay, silt, fine sand, coarse sand), it is not appropriate to fit a model that has three to four parameters. An empirical regression seems more suitable, e.g. Marshall (1947) approximated the conversion as:

\[ P_{2-20} = 0.5 \pm 0.13 P_{2-50}. \]

Minasny et al. (1999) converted the 2-20 µm fraction into a 2-50 µm fraction based on the Geeves and Smettem data sets that provide measurements for both fractions. The model calculated the increment of particles when converting 2-20 µm to 2-50 µm fractions \( P_{20-50} \):

\[
P_{20-50} (\text{dag/kg}) = 48.4593 - 0.2225 P_{20-2000} - 0.0029 (P_{20-2000})^2 - 0.6952 P_{<2} + 0.0018 (P_{<2})^2 \quad (R^2 = 0.76, N = 290).
\]

[4.1.1]

In order to achieve better prediction, a larger data set was compiled. The data set is derived from databases which contain measurement of both fractions, i.e. UNSODA \((N = 194)\), GRIZZLY \((N = 488)\), and the USDA-NRCS Soil Survey Laboratory data (USDA-NRCS, 1997). The Australian + UNSODA + GRIZZLY + NRCS data were combined and randomly split into two: a prediction set \((N = 1210)\) and a validation set \((N = 400)\). The particle-size conversion was modelled using a multiple linear regression. The model is:

\[
\hat{P}_{2-50} = -18.3914 + 2.0971 (P_{2-20}) + 0.6726 (P_{20-2000}) - 0.0142 (P_{2-20})^2 - 0.0049 (P_{20-2000})^2 \quad (R^2 = 0.8235)
\]

If \( \hat{P}_{2-50} < 0 \) then \( \hat{P}_{2-50} = 0.8289 (P_{2-20}) + 0.0198 (P_{20-2000}). \) [4.1.2]

Conversely, if we had measurements based on the FAO classification, a model was derived to predict \( P_{2-20} \):

\[
\hat{P}_{2-20} = -0.4070 - 0.1271 (P_{<2}) + 0.5527 (P_{2-50}) + 0.0017 (P_{<2})^2 - 0.0019 (P_{2-50})^2 + 0.0059 (P_{<2}) (P_{2-50}) \quad (R^2 = 0.8179)
\]

If \( \hat{P}_{2-20} < 0 \) then \( \hat{P}_{2-20} = 0.1154 (P_{<2}) + 0.2212 (P_{2-50}) \) [4.1.3]

A summary of the model parameters is given in Table 4.1.1.

Table 4.1.2 presents the mean error (ME) and root mean squared error (RMSE) between the predictions and measured values. A log-linear interpolation was also used by way of comparison:

\[
CP_i = CP_{i-1} + \frac{CP_{i+1} - CP_{i-1}}{\log(PS_{i+1}) - \log(PS_{i-1})} \left[ \log(PS_i) - \log(PS_{i-1}) \right] \quad [4.1.4]
\]
where \( CP \) is the cumulative amount of particles mass on \( PS \) particle-size limit, \( i, i-1, i+1 \) represent the point to interpolate, and the preceding and succeeding neighbour limits (Nemes et al., 1999a). The results clearly showed that the regression model is better than the log-linear interpolation. Predicting \( P_{2.20} \) is more accurate (± 5%) than \( P_{2.50} \) (± 7%).

Table 4.1.1. Parameters for conversion between particle-size classes.

<table>
<thead>
<tr>
<th>Term</th>
<th>Estimate</th>
<th>Std Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_{2.20} ) model:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-18.3914</td>
<td>1.7653</td>
</tr>
<tr>
<td>( P_{2.20} )</td>
<td>2.0971</td>
<td>0.0931</td>
</tr>
<tr>
<td>( P_{20.2000} )</td>
<td>0.6726</td>
<td>0.0552</td>
</tr>
<tr>
<td>( (P_{2.20})^2 )</td>
<td>-0.0142</td>
<td>0.0018</td>
</tr>
<tr>
<td>( (P_{20.2000})^2 )</td>
<td>-0.0049</td>
<td>0.0005</td>
</tr>
<tr>
<td>Otherwise term:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( P_{2.20} )</td>
<td>0.8289</td>
<td>0.1147</td>
</tr>
<tr>
<td>( P_{20.2000} )</td>
<td>0.0198</td>
<td>0.0020</td>
</tr>
<tr>
<td>( P_{2.20} ) model:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.4070</td>
<td>0.4539</td>
</tr>
<tr>
<td>( P&lt;2 )</td>
<td>-0.1271</td>
<td>0.0412</td>
</tr>
<tr>
<td>( P_{2.50} )</td>
<td>0.5527</td>
<td>0.0263</td>
</tr>
<tr>
<td>( (P&lt;2)^2 )</td>
<td>0.0017</td>
<td>0.0005</td>
</tr>
<tr>
<td>( (P_{2.50})^2 )</td>
<td>-0.0019</td>
<td>0.0004</td>
</tr>
<tr>
<td>( (P&lt;2)(P_{2.50}) )</td>
<td>0.0059</td>
<td>0.0007</td>
</tr>
<tr>
<td>Otherwise term:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( P&lt;2 )</td>
<td>0.1154</td>
<td>0.0160</td>
</tr>
<tr>
<td>( P_{2.50} )</td>
<td>0.2212</td>
<td>0.0132</td>
</tr>
</tbody>
</table>

Table 4.1.2. Statistics of the error in particle size classes conversion.

<table>
<thead>
<tr>
<th>Method</th>
<th>Predicted particle size</th>
<th>Equation</th>
<th>Prediction set ((N = 1210))</th>
<th>Validation set ((N = 400))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>( P_{2.20} )</td>
<td>4.1.3</td>
<td>-0.01 4.63 0.26 5.38</td>
<td></td>
</tr>
<tr>
<td>Log-linear</td>
<td>( P_{2.20} )</td>
<td>4.1.4</td>
<td>-6.06 8.73 6.10 9.08</td>
<td></td>
</tr>
<tr>
<td>Regression</td>
<td>( P_{2.50} )</td>
<td>4.1.2</td>
<td>0.00 7.48 0.39 8.01</td>
<td></td>
</tr>
<tr>
<td>Log-linear</td>
<td>( P_{2.50} )</td>
<td>4.1.4</td>
<td>1.60 11.96 -1.90 13.10</td>
<td></td>
</tr>
</tbody>
</table>

4.1.3. Textural classes conversion

The particle-size distribution is usually represented in a texture diagram, relating the percentages of sand, silt and clay to a texture class. According to Marshall (1947) the first texture diagram relating the particle-size distribution to textural classes was drawn in 1911 by Whitney (1911) where the clay and silt percentages were
represented on a right-angled triangle. Davis and Bennett (1927) replaced it by a more
difficult to read ternary diagram which shows percentages of clay, silt, and sand. In
1945, the USDA revised its texture triangle to accommodate changes in analytical
procedure and particle-size limits.

Prescott et al. (1934) devised a texture diagram for the International system based
on mechanical analysis and field descriptions made by the CSIR (Council for
Scientific and Industrial Research) Division of Soils at that time. This was later
revised by Marshall (1947) and has been used in Australia until the present
(McDonald et al. 1990). Toogood (1958) presented a ‘simplified’ Canadian texture
diagram (which is similar to the USDA) where percentages of sand and clay were
plotted on a right triangle. McBratney suggested this type of diagram to be used in

The small range of silt in the International system and the non-existence of the
silt texture in the Australian texture triangle posed a question how the textural classes
of the USDA/FAO system correspond to those of Australia. Shaw (1935) compared
the textural classes on the USDA and International system, and found it difficult to
harmonise the significance of the figures. Marshall (1947) indicated that it was
impossible to make direct comparison between texture on the two systems. Marshall
(1947) presented a ‘rough conversion’ of the USDA textural classes into the
International system coordinates. Since it is based on a linear transformation ($P_{2.20} =
0.5 P_{2.50}$) the texture classes were shown to occupy exactly one half of the
International system’s triangle. Using Equation [4.1.3], the vertices from each of the
textural class from the USDA/FAO soil texture triangle (Soil Survey Staff, 1975)
were converted to the International system. Similarly, the Australian texture classes
(Marshall, 1947) were converted to the USDA/FAO system using Equation [4.1.2].

Figure 4.1.2a shows that the USDA/FAO textural class is approximately a
‘boomerang’ shape in the International system. Table 4.1.3 shows the centroid of the
polygon for each textural class and the percentage of the area occupied on the triangle
and the ‘boomerang’. The total area is about 60% of the triangle, 10% more than
predicted by Marshall (1947). Figure 4.1.3 shows the USDA texture classes plotted
together with the Australian system. There are overlaps between the two textural
classes, especially for the sandy clay and sandy clay loam textures.
Figure 4.1.2. (a) The USDA texture classes plotted in the axes of international system. The centroids of the classes are indicated by dark dots. (b) The Australian texture classes plotted in the axes of the USDA/FAO system.
<table>
<thead>
<tr>
<th>Texture class</th>
<th>Centroid</th>
<th>% area triangle</th>
<th>% area boomerang</th>
</tr>
</thead>
<tbody>
<tr>
<td>sand</td>
<td>94.1</td>
<td>2.6</td>
<td>3.3</td>
</tr>
<tr>
<td>loamy sand</td>
<td>87.8</td>
<td>6.3</td>
<td>5.9</td>
</tr>
<tr>
<td>sandy loam</td>
<td>77.1</td>
<td>12.3</td>
<td>10.7</td>
</tr>
<tr>
<td>loam</td>
<td>59.9</td>
<td>21.4</td>
<td>18.7</td>
</tr>
<tr>
<td>silt</td>
<td>53.8</td>
<td>32.1</td>
<td>14.0</td>
</tr>
<tr>
<td>sandy clay loam</td>
<td>57.4</td>
<td>37.1</td>
<td>5.5</td>
</tr>
<tr>
<td>clay loam</td>
<td>65.8</td>
<td>7.0</td>
<td>27.1</td>
</tr>
<tr>
<td>silty clay loam</td>
<td>46.9</td>
<td>19.3</td>
<td>33.8</td>
</tr>
<tr>
<td>sandy clay</td>
<td>31.5</td>
<td>34.4</td>
<td>34.0</td>
</tr>
<tr>
<td>silty clay</td>
<td>54.7</td>
<td>3.6</td>
<td>41.7</td>
</tr>
<tr>
<td>clay</td>
<td>20.1</td>
<td>33.3</td>
<td>46.7</td>
</tr>
<tr>
<td>Total</td>
<td>59.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.1.3. The USDA (itals black font) and Marshall’s textural classes (red font) plotted on the axes of the International system.
The particle-size of 2430 soil samples that contain measurements on both systems are plotted on the two texture triangles (Figure 4.1.4). The data are more evenly distributed in the USDA/FAO system, while in the International system most of the data is within the ‘boomerang’. Very few data fell outside the boomerang, in this case only 5%. This ‘blank’ region represent the soil that has $P_{2-20} > 40\%$. It is quite infrequent to find a soil which has such a large mass in the $2 – 20 \mu m$ range. The non-parametric quantile density plots show that the data are distributed in the middle of the USDA/FAO texture triangle, while in the International triangle the data is only within the ‘boomerang’.

A textural class contingency table of the USDA/FAO and Australian systems (Table 4.1.4) shows that the USDA classes overlap to a fair degree with the Australian classes, which confirms the representation in Figure 4.1.3. Sandy loam (USDA/FAO) can translate into sand, loamy sand, sandy loam, loam, silt loam and sandy clay loam in the Australian system. Silt loam (USDA) can also shift into eight other classes in the Australian system.

Some overlaps should not have occurred according to Figure 4.1.3 (e.g. silt loam to loam). This is because the conversion is based on an approximate regression and in reality the particle-size distributions can be multi-modal (Walker and Chittleborough, 1986) and the data are also bound to contain a degree of measurement error.

Table 4.1.4. Contingency table showing the distribution of the USDA texture classes with the international system, numbers are in percentage of the number of data.

<table>
<thead>
<tr>
<th></th>
<th>S</th>
<th>LS</th>
<th>SL</th>
<th>L</th>
<th>ZL</th>
<th>Z</th>
<th>SCL</th>
<th>CL</th>
<th>ZCL</th>
<th>SC</th>
<th>ZC</th>
<th>C</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>USDA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>12.2</td>
<td>2.2</td>
<td>0.2</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14.7</td>
</tr>
<tr>
<td>LS</td>
<td>0.5</td>
<td>6.6</td>
<td>14.8</td>
<td>1.8</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23.9</td>
</tr>
<tr>
<td>SL</td>
<td>0.0</td>
<td>0.6</td>
<td>4.6</td>
<td>0.2</td>
<td>0.1</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.6</td>
</tr>
<tr>
<td>L</td>
<td>4.4</td>
<td>5.9</td>
<td>2.4</td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13.2</td>
</tr>
<tr>
<td>ZL</td>
<td>1.2</td>
<td>2.5</td>
<td>7.7</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11.5</td>
</tr>
<tr>
<td>SCL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.5</td>
</tr>
<tr>
<td>CL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1</td>
<td></td>
<td>6.5</td>
</tr>
<tr>
<td>ZCL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.2</td>
</tr>
<tr>
<td>SC</td>
<td>0.1</td>
<td>0.5</td>
<td>0.1</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.2</td>
</tr>
<tr>
<td>ZC</td>
<td>0.2</td>
<td>0.1</td>
<td>1.4</td>
<td>0.2</td>
<td>0.3</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11.7</td>
</tr>
<tr>
<td>Sum</td>
<td>12.8</td>
<td>9.4</td>
<td>25.3</td>
<td>12.1</td>
<td>11.4</td>
<td>1.9</td>
<td>5.7</td>
<td>5.2</td>
<td>0.3</td>
<td>4.1</td>
<td>9.7</td>
<td>12.6</td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.1.4. Distribution of particle size data in the (a) USDA/FAO system, and (b) International system. Classes are the USDA/FAO system. The contour lines represent the quantile density, with 5% interval between each line. The outer contour line represents 95% of the data density.
4.1.4. The hydraulic database

Previously published water-retention and saturated hydraulic conductivity data sets from across Australia were compiled. The data consists of measurements made by Bridge (1968), Prebble (1970a, 1970b), Forrest et al. (1985), Geeves et al. (1995), Bristow et al. (1999), various studies from the University of Sydney (USyd), and a published paper (Olsson and Rose, 1978). The description and number of data are summarised in Table 4.1.5 and 4.1.6.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Water Retention</th>
<th>$K_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Profiles</td>
<td>Samples</td>
</tr>
<tr>
<td>Prebble</td>
<td>17</td>
<td>78</td>
</tr>
<tr>
<td>Forrest</td>
<td>60</td>
<td>118</td>
</tr>
<tr>
<td>Geeves</td>
<td>74</td>
<td>146</td>
</tr>
<tr>
<td>Smettem</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>Bristow</td>
<td>46</td>
<td>450</td>
</tr>
<tr>
<td>SydU</td>
<td>17</td>
<td>35</td>
</tr>
<tr>
<td>Bridge</td>
<td>59</td>
<td>104</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>270</strong></td>
<td><strong>946</strong></td>
</tr>
</tbody>
</table>

The particle-size distributions are plotted in Figure 4.1.5. As can be seen from Figure 4.1.1a the texture range for the Australian data is quite similar to that for European data contained in HYPRES (Nemes et al., 1999a), except that there are more points in the sandy loam – sandy clay loam region but a distinct lack of silt – silt loam in the Australian data.
## Table 4.1.6. Description of Australian hydraulic data sets.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Location</th>
<th>Samples</th>
<th>Particle size ranges (µm)</th>
<th>Water potential (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prebble (1970a)</td>
<td>Queensland</td>
<td>undisturbed cores (0.06m diameter, 0.034m deep) from various depths</td>
<td>&lt;2, 2-20, 20-2000 (data from Stace et al., 1968)</td>
<td>0, -0.4, -1, -3.3, -6.8, -20, -70, -150</td>
</tr>
<tr>
<td>Forrest et al. (1985)</td>
<td>wheat-lands of eastern Australia (Southern Queensland to South Australia)</td>
<td>Undisturbed cores (0.076m diameter, 0.050m deep) from top- &amp; subsoil [water retention]</td>
<td>&lt;2, 2-20, 20-200, 200-2000</td>
<td>0, -0.5, -1, -10, -150</td>
</tr>
<tr>
<td>Geeves et al. (1995)</td>
<td>wheat-belt, southern New South Wales and northern Victoria</td>
<td>undisturbed cores (0.098m diameter, 0.075m deep) from top- &amp; subsoil [water retention]</td>
<td>&lt;2, 2-20, 20-63, 63-212, 212-425, 425-1000, 1000-2000, &gt;2000</td>
<td>-0.1, -0.3, 0.5, -1, -3.3, -6.6, -10, -30, -50, -150</td>
</tr>
<tr>
<td>Smettem and Gregory (1996)</td>
<td>northern and central wheatbelt, Western Australia</td>
<td>undisturbed cores (0.047m diameter, 0.03m deep) from various depths [water retention]</td>
<td>&lt;2, 2-20, 20-53, 53-125, 125-180, 180-250, 250-355, 355-500, 500-1000, 1000-1400, 1400-2000</td>
<td>0, -0.15, -0.4, -0.9, -3, -10, -30, -150</td>
</tr>
<tr>
<td>Bristow et al. (1999)</td>
<td>tropical north Queensland</td>
<td>Undisturbed cores (0.075 m diameter, 0.05 m deep) from various depths [water retention &amp; Ks]</td>
<td>&lt;2, 2-20, 20-200, 200-2000, &gt;2000</td>
<td>0, -0.1, -0.3, -0.5, -1, -3, -10, -30, -150</td>
</tr>
<tr>
<td>Bridge (1968)</td>
<td>Conargo sandy loam, Purdana sandy loam, Billabong clay, NSW</td>
<td>Undisturbed cores from various depths [water retention]</td>
<td>&lt;2, 2-20, 20-200, 200-2000</td>
<td>0, -0.01, -0.1, -0.2, -0.5, -1, -3, -6, -30, -150, -1500</td>
</tr>
<tr>
<td>SydU (Department of Soil Science, The University of Sydney)</td>
<td>various locations in New South Wales</td>
<td>Disturbed samples [water retention]</td>
<td>&lt;2, 2-20, 20-200, 200-2000</td>
<td>0, -1, -3.3, -10, -32, -100, -150, -1000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Undisturbed cores [Ks] from various depths</td>
<td>&lt;2, 2-20, 20-200, 200-2000</td>
<td>0, -1, -3.3, -10, -32, -100, -150, -1000</td>
</tr>
</tbody>
</table>
Figure 4.1.5. Particle size distribution of soil database plotted in (a) USDA/FAO system, shaded area represents texture range from European database (HYPRES); (b) Australian system... continued
Figure 4.1.5. Particle size distribution of soil database plotted in (c) International system with USDA textural classes.

The distribution of the data according to textural class is given in Table 4.1.7. Most data within the clay and loam-loamy sand texture range and there are no data on silt. When using the USDA/FAO system, the texture classes shift from loamy sand into sandy loam.

Table 4.1.7. Distribution of texture classes in the database

<table>
<thead>
<tr>
<th>Texture class</th>
<th>$N_{\text{Australian}}$</th>
<th>$N_{\text{USDA/FAO}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>clay</td>
<td>233</td>
<td>191</td>
</tr>
<tr>
<td>clay loam</td>
<td>131</td>
<td>107</td>
</tr>
<tr>
<td>loam</td>
<td>170</td>
<td>172</td>
</tr>
<tr>
<td>loamy sand</td>
<td>108</td>
<td>24</td>
</tr>
<tr>
<td>sand</td>
<td>35</td>
<td>18</td>
</tr>
<tr>
<td>sandy clay</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>sandy clay loam</td>
<td>15</td>
<td>41</td>
</tr>
<tr>
<td>sandy loam</td>
<td>67</td>
<td>211</td>
</tr>
<tr>
<td>silt loam</td>
<td>75</td>
<td>43</td>
</tr>
<tr>
<td>silty clay</td>
<td>62</td>
<td>80</td>
</tr>
<tr>
<td>silty clay loam</td>
<td>47</td>
<td>58</td>
</tr>
</tbody>
</table>
The soil physical properties common to all data sets were tabulated and will be used to develop hydraulic pedotransfer functions. The properties are:

1. Bulk density \( (\rho_b) \) in Mg m\(^{-3}\).
2. Porosity \( (\phi) \) in m\(^3\)m\(^{-3}\),
   calculated as \( \phi = 1 - \rho_b/\rho_s \), where \( \rho_s \) is the soil particle density; if no measurement was available 2.65 Mg m\(^{-3}\) was used.
3. Saturated water content \( (\theta_s) \), with the exception of the Bridge, Prebble, Forrest, Smettem and Bristow data sets, which contain measured values of \( \theta_s \), the empirical relationship of Williams et al. (1992) was used: \( \theta_s = 0.93 \phi \).
4. The mass of particles <2 \( \mu \)m (P <2), 2-20 \( \mu \)m (P 2-20) and 20-2000 \( \mu \)m (P 20-2000), which were normalized to sum to 100 dag/kg.
5. Geometric mean particle-size diameter \( (d_g \text{ in mm}) \) and geometric standard deviation \( (\sigma_g \text{ in mm}) \). These were calculated according to Shirazi and Boersma (1984) as:
   \[ d_g = \exp c_1; \quad \sigma_g = \exp c_2, \]
   \[ c_1 = 0.01 \sum_{i=1}^{nf} f_i \ln M_i ; \quad c_2 = \sqrt{0.01 \sum_{i=1}^{nf} f_i (\ln M_i)^2 - a^2} \]
   where \( nf \) is the number of particle fractions, and \( f_i \) is the percent of total soil mass having a diameter equal to or less than \( M_i \). The number of particle fractions used for calculation depends on the availability of particle-size distribution data.

Organic carbon was not included as a predictive variable for developing the PTFs because of the unavailability of this measurement in some data sets, different laboratory methods among data sets, and the relatively small amount of organic matter content in some Australian soil materials (Bristow et al., 1999).

The water retention data is fitted to the van Genuchten and Campbell’s equation (see Table 1.2.2) using nonlinear regression with the Levenberg-Marquardt algorithm (Marquardt, 1963). The \( \theta_r \), \( \alpha \) and \( n \) parameters of van Genuchten were estimated with the following constraints imposed: \( \theta_r \geq 0 \text{ m}^3\text{m}^{-3} \), \( 0.01 \leq \alpha \leq 100 \text{ m}^{-1} \) and \( 1.01 \leq n \leq 10 \). If \( \theta_r < 0.0001 \) then its value was fixed at 0.
### Table 4.1.8. Statistics of the soil properties from Australian database*

<table>
<thead>
<tr>
<th>Properties</th>
<th>Units</th>
<th>Min.</th>
<th>Max.</th>
<th>Median</th>
<th>$\mu$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_b$</td>
<td>Mg m$^{-3}$</td>
<td>0.57</td>
<td>1.95</td>
<td>1.49</td>
<td>1.45</td>
<td>0.21</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Mg m$^{-3}$</td>
<td>0.26</td>
<td>0.78</td>
<td>0.44</td>
<td>0.46</td>
<td>0.08</td>
</tr>
<tr>
<td>$P_{&lt;2}$</td>
<td>dag kg$^{-1}$</td>
<td>1.0</td>
<td>75.0</td>
<td>25.9</td>
<td>29.1</td>
<td>17.3</td>
</tr>
<tr>
<td>$P_{2-20}$</td>
<td>dag kg$^{-1}$</td>
<td>0.6</td>
<td>77.8</td>
<td>14.9</td>
<td>17.3</td>
<td>10.3</td>
</tr>
<tr>
<td>$P_{20-2000}$</td>
<td>dag kg$^{-1}$</td>
<td>7.1</td>
<td>96.0</td>
<td>56.2</td>
<td>53.5</td>
<td>21.5</td>
</tr>
<tr>
<td>$\ln(d_g)$</td>
<td>ln (mm)</td>
<td>-5.804</td>
<td>-3.512</td>
<td>-3.390</td>
<td>1.267</td>
<td></td>
</tr>
<tr>
<td>$\ln(\sigma_g)$</td>
<td>ln (mm)</td>
<td>1.080</td>
<td>2.258</td>
<td>2.274</td>
<td>0.385</td>
<td></td>
</tr>
<tr>
<td>$\theta_l$</td>
<td>m$^3$m$^{-3}$</td>
<td>0.000</td>
<td>0.346</td>
<td>0.013</td>
<td>0.056</td>
<td>0.078</td>
</tr>
<tr>
<td>$\theta_b$</td>
<td>m$^3$m$^{-3}$</td>
<td>0.220</td>
<td>0.401</td>
<td>0.418</td>
<td>0.080</td>
<td></td>
</tr>
<tr>
<td>$\ln(\alpha)$</td>
<td>ln ($\sigma$)</td>
<td>-4.605</td>
<td>1.030</td>
<td>1.016</td>
<td>1.485</td>
<td></td>
</tr>
<tr>
<td>$\ln(n-1)$</td>
<td>-3.992</td>
<td>2.171</td>
<td>1.514</td>
<td>1.537</td>
<td>0.788</td>
<td></td>
</tr>
<tr>
<td>$\ln(h_o)$</td>
<td>ln (m)</td>
<td>-9.721</td>
<td>-1.709</td>
<td>-1.817</td>
<td>1.332</td>
<td></td>
</tr>
<tr>
<td>$\ln(b)$</td>
<td>-2.637</td>
<td>4.100</td>
<td>1.952</td>
<td>1.649</td>
<td>1.394</td>
<td></td>
</tr>
<tr>
<td>$\theta$(-10)</td>
<td>m$^3$m$^{-3}$</td>
<td>0.073</td>
<td>0.321</td>
<td>0.330</td>
<td>0.102</td>
<td></td>
</tr>
<tr>
<td>$\theta$(-33)</td>
<td>m$^3$m$^{-3}$</td>
<td>0.045</td>
<td>0.281</td>
<td>0.293</td>
<td>0.107</td>
<td></td>
</tr>
<tr>
<td>$\theta$(-1500)</td>
<td>m$^3$m$^{-3}$</td>
<td>0.018</td>
<td>0.164</td>
<td>0.175</td>
<td>0.090</td>
<td></td>
</tr>
<tr>
<td>AWC</td>
<td>m$^3$m$^{-3}$</td>
<td>0.000</td>
<td>0.110</td>
<td>0.117</td>
<td>0.053</td>
<td></td>
</tr>
<tr>
<td>$\ln(K_s)$</td>
<td>ln (mm/h)</td>
<td>-5.169</td>
<td>2.788</td>
<td>2.543</td>
<td>2.595</td>
<td></td>
</tr>
</tbody>
</table>

*$N = 946$, except for $K_s$, $N = 661$

### 4.1.5. Textural stratification

For stratification purposes, the particle-size data is classified into textural classes based on the percentage of $P_{<2}$, $P_{2-20}$ and $P_{20-2000}$ by the fuzzy $k$-means algorithm (McBratney and DeGruijter, 1992). Because the data do not cover all the texture range, they were grouped into 3 classes (Figure 4.1.6):

- sandy or coarser texture (centroid: 12.2% $P_{<2}$, 11.8% $P_{2-20}$, 76.0% $P_{20-2000}$)
- loamy or medium texture (centroid: 26.6% $P_{<2}$, 18.2% $P_{2-20}$, 55.1% $P_{20-2000}$)
- clayey or lighter texture (centroid: 47.2% $P_{<2}$, 25.4% $P_{2-20}$, 27.4% $P_{20-2000}$).

The membership to a particular texture class $u$ is calculated as:

$$u_c = \frac{d_c^{-2(\phi-1)}}{\sum_{j=1}^{k} d_j^{-2(\phi-1)}} \tag{4.1.5}$$

where $u_c$ is the membership at class $c$, $k$ is the number of classes = 3, $d$ is the Euclidean distance between the sample and centroid and $\Phi$ is the fuzzy exponent = 1.5.
Figure 4.1.6. Stratification of texture into three classes plotted in the International system, the centroid for each class is represented by the star symbol.
4.2. Water retention curve PTFs

The water-retention curve, which defines the relationship between soil water content ($\theta$) and hydraulic potential ($h$) is an important physical property of soil material. This study compared different approaches for deriving point estimation and parametric PTFs from particle-size and bulk density data. The performances of developed PTFs for estimating water retention for Australian soil were also compared.

### 4.2.1. Methods

#### 4.2.1.1. The data

The water-retention from the Australian database was used to develop the PTFs and a subset of the data was randomly selected and used as a validation set (Table 4.2.1). A summary of the soil properties is given on Table 4.2.2.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>All data Profiles</th>
<th>All data Samples</th>
<th>Prediction set Profiles</th>
<th>Prediction set Samples</th>
<th>Validation Set Profiles</th>
<th>Validation Set Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prebble</td>
<td>17</td>
<td>78</td>
<td>15</td>
<td>71</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Forrest</td>
<td>60</td>
<td>118</td>
<td>54</td>
<td>106</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Geeves</td>
<td>74</td>
<td>146</td>
<td>60</td>
<td>132</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Smettem</td>
<td>4</td>
<td>15</td>
<td>3</td>
<td>10</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Bristow</td>
<td>46</td>
<td>450</td>
<td>42</td>
<td>414</td>
<td>4</td>
<td>36</td>
</tr>
<tr>
<td>SydU</td>
<td>17</td>
<td>35</td>
<td>17</td>
<td>35</td>
<td>42</td>
<td>109</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>218</strong></td>
<td><strong>842</strong></td>
<td><strong>174</strong></td>
<td><strong>733</strong></td>
<td><strong>42</strong></td>
<td><strong>109</strong></td>
</tr>
</tbody>
</table>

To test the applicability of the developed PTFs, published databases available from the world-wide web were used. The first is GRIZZLY (Haverkamp et al., 1997) from Laboratoire d'Étude des Transfers en Hydrologie et Environnement (LTHE), Grenoble, which consists of 660 soil water retention from different countries. This database has been used by Bouraoui et al. (1999) to develop a physically-based PTF. The mass of particles $P_{2-20}$ and $P_{20-2000}$ were calculated from the van Genuchten type cumulative particle-size distribution function provided in the program. The second database UNSODA (Nemes et al. 1999) is from the US Salinity Laboratory. This data set has been used extensively by researchers in US to develop PTFs (Schaap et al., 1998; Arya et al., 1999). 533 water retention sets are used and the sand, silt fraction of
the USDA classification were converted to the international system according to equation 4.1.3.

Table 4.2.2. Means and standard deviations of soil properties for the prediction and validation sets

<table>
<thead>
<tr>
<th>Properties</th>
<th>Units</th>
<th>Prediction set $N = 733$</th>
<th>Validation set $N = 109$</th>
<th>prob $t$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\mu$</td>
<td>$\sigma$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>$\rho_b$</td>
<td>Mg m$^{-3}$</td>
<td>1.43</td>
<td>0.22</td>
<td>1.42</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Mg m$^{-3}$</td>
<td>0.46</td>
<td>0.08</td>
<td>0.47</td>
</tr>
<tr>
<td>$P_{&lt;2}$</td>
<td>dag kg$^{-1}$</td>
<td>27.41</td>
<td>16.19</td>
<td>32.78</td>
</tr>
<tr>
<td>$P_{2-20}$</td>
<td>dag kg$^{-1}$</td>
<td>18.11</td>
<td>10.25</td>
<td>19.88</td>
</tr>
<tr>
<td>$P_{20-2000}$</td>
<td>dag kg$^{-1}$</td>
<td>54.47</td>
<td>21.36</td>
<td>47.33</td>
</tr>
<tr>
<td>ln($d_r$)</td>
<td>ln (mm)</td>
<td>-3.42</td>
<td>1.16</td>
<td>-2.88</td>
</tr>
<tr>
<td>ln($\sigma_g$)</td>
<td>ln (mm)</td>
<td>2.28</td>
<td>0.39</td>
<td>2.11</td>
</tr>
<tr>
<td>$\theta_{(-10)}$</td>
<td>m$^3$m$^{-3}$</td>
<td>0.33</td>
<td>0.10</td>
<td>0.36</td>
</tr>
<tr>
<td>$\theta_{(-33)}$</td>
<td>m$^3$m$^{-3}$</td>
<td>0.29</td>
<td>0.10</td>
<td>0.33</td>
</tr>
<tr>
<td>$\theta_{(-1500)}$</td>
<td>m$^3$m$^{-3}$</td>
<td>0.18</td>
<td>0.09</td>
<td>0.19</td>
</tr>
<tr>
<td>AWC</td>
<td>m$^3$m$^{-3}$</td>
<td>0.12</td>
<td>0.05</td>
<td>0.14</td>
</tr>
</tbody>
</table>

* probability associated with $t$ test for two samples come from distributions with the same means.

* samples significantly come from distributions with the same means.

The third data set came from the study of Ratcliffe et al. (1983) where plant-available soil water limits were determined. The measured properties are:

- drainage upper limit (DUL) – the largest field-measured soil water content after it had been saturated and allowed to drain until drainage became negligible, or equivalent to $\theta$ at -33 kPa (FC);
- lower limit (LOL) – the smallest field measured soil water content after plants has stopped extracting water and were at or near water stress condition, or equivalent to $\theta$ at -1500 kPa (PWP);
- potential extractable soil water (PLEXW) – the difference between DUL and LOL, or equivalent to AWC.

The data were derived from 61 horizons across the US, and 312 data that contain particle size distribution and bulk density were selected. Sand and silt fractions were converted to the ISSS system by using equation 4.1.3.
4.2.1.2. Pedotransfer function development

Two types of PTF were considered:

- **point estimation**, where water contents were estimated at potentials of -1, -33 and -1500 kPa. Water content at -1500 kPa $\theta_{-1500}$ is referred to as the Permanent Wilting Point (PWP). Available water content (AWC) is defined as:
  \[
  AWC = \theta_{-33} - \theta_{-1500};
  \]

- **parametric estimation**, where the parameters of the van Genuchten equation: $\theta_r$, $\theta_s$, $\alpha$ and $n$ are estimated.

Four different approaches were used to produce the PTFs:

1. **Multiple linear regression**
   Multiple linear regression was used to develop PTFs for single-point estimation (MLR) and also parametric estimation (MRP). Variables used to predict $\theta$ were selected using stepwise regression.

2. **Extended nonlinear regression (ENR)**
   The extended nonlinear regression (ENR) method of Scheinost *et al.* (1997) was used to develop 6 different combinations of variables for predicting the van Genuchten parameters (Table 4.2.3) based on the following assumptions:
   - $\theta_l$ is predicted from $P_{<2}$ and PWP, as $\theta_l$ is related to the asymptotic $\theta$ at the dry end ($<-1500kPa$) which in turn is related to $P_{<2}$ (Scheinost *et al.*, 1997). When PWP data are not available, $\theta_l$ is predicted from $P_{2-20}$ and $P_{<2}$.
   - $\theta_s$ is predicted from $P_{<2}$ and $\phi$ (Scheinost *et al.*, 1997).
   - Campbell (1985) found that $h_b (\approx 1/\alpha)$ is a function of $d_g^{-0.5}$ and $b$ is a function of $h_b$ and $\sigma_g$. Scheinost *et al.* (1997) proposed that $\alpha$ is related to $d_g$ and $n$ related to $1/\sigma_g$ and Bouraoui *et al.* (1997) found that $n$ is related to $\ln(P_{<2} / (P_{<2} + P_{2-50}))$.

PTFs were also developed by stratifying the texture into 3 classes (sandy, loamy, clayey) as described in 4.1.5. PTFs were developed for each class by holding $\theta_l$ and $\theta_s$ constant and varying $\alpha$ and $n$. The class PTF for each soil depends solely on the class to which it belongs, while the continuous PTF is calculated from

\[
PTF = u_{\text{sandy}} PTF_{\text{sandy}} + u_{\text{loamy}} PTF_{\text{loamy}} + u_{\text{clayey}} PTF_{\text{clayey}}
\]

where $u$ is the membership of the particular class.
Table 4.2.3. Extended nonlinear regression (ENR) models for parametric PTFs

<table>
<thead>
<tr>
<th>PTF</th>
<th>np</th>
<th>$\theta_1$</th>
<th>$\theta_2$</th>
<th>$\alpha$</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENR1</td>
<td>11</td>
<td>$r_0 + r_1 P &lt;2 + r_2 PWP$</td>
<td>$s_0 + s_1 P &lt;2 + s_2 \phi$</td>
<td>$\alpha_0 + \alpha_1 d_g$</td>
<td>$n_0 + n_1 \sigma_\theta + n_2 \sigma_g$</td>
</tr>
<tr>
<td>ENR2</td>
<td>8</td>
<td>$r_1 P &lt;2 + r_2 PWP$</td>
<td>$s_0 P &lt;2 + s_1 \phi$</td>
<td>$\alpha_0 + \alpha_1 d_g$</td>
<td>$n_0 + n_1 \sigma_\theta$</td>
</tr>
<tr>
<td>ENR3</td>
<td>8</td>
<td>$r_1 P &lt;2 + r_2 PWP$</td>
<td>$s_0 P &lt;2 + s_1 \phi$</td>
<td>$\alpha_0 + \alpha_1 d_g^{1/2}$</td>
<td>$n_0 + n_1 / \sigma_\theta$</td>
</tr>
<tr>
<td>ENR4</td>
<td>8</td>
<td>$r_1 P &lt;2 + r_2 PWP$</td>
<td>$s_0 P &lt;2 + s_1 \phi$</td>
<td>$\alpha_0 + \alpha_1 d_g$</td>
<td>$n_0 + n_1$</td>
</tr>
<tr>
<td>ENR5</td>
<td>8</td>
<td>$r_1 P &lt;2 + r_2 PWP$</td>
<td>$s_0 P &lt;2 + s_1 \phi$</td>
<td>$\alpha_0 + \alpha_1 d_g$</td>
<td>$n_0 + n_1 / \sigma_\theta$</td>
</tr>
<tr>
<td>ENR6</td>
<td>8</td>
<td>$r_1 P &lt;2 + 0.7 PWP$</td>
<td>$s_0 P &lt;2 + s_1 \phi$</td>
<td>$\alpha_0 + \alpha_1 d_g$</td>
<td>$n_0 + n_1 \sigma_\theta + n_2 \rho_b$</td>
</tr>
<tr>
<td>ENR7</td>
<td>9</td>
<td>$r_0 + r_1 P &lt;2$</td>
<td>$s_0 P &lt;2 + s_1 \phi$</td>
<td>$\alpha_0 + \alpha_1 d_g$</td>
<td>$n_0 + n_1 \sigma_\theta$</td>
</tr>
<tr>
<td>ENR8</td>
<td>16</td>
<td>Continuous function of ENR2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENR9</td>
<td>16</td>
<td>Class function of ENR2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The extended van Genuchten model was fitted to all water retention data using nonlinear regression with the Levenberg-Marquardt algorithm (Marquardt, 1963) which minimises the sum of squared residuals (SSR):

$$SSR = \sum_{i=1}^{NZ} \sum_{j=1}^{Z(i)} \left( \theta_j - \hat{\theta}_j(b) \right)^2$$  \[4.2.1\]

where $b$ is the parameter vector = \{r, s, a, n\} that replaces \{$\theta_0$, $\theta_1$, $\alpha$, n\} in the van Genuchten equation, $Z(i)$ is the number of water retention data for soil sample $i$ and $NW$ is the number of soil samples ($NZ = 733$). To test and compare the performance of the PTFs, ‘leave-one-out’ cross validation (Efron and Tibshirani, 1993) was performed. For each soil sample $i$ of measured water retention the model was fitted leaving out that sample, and then the predicted value for the sample as $\sum_{j=1}^{Z(i)} \hat{\theta}_j$ was computed. The estimated cross-validation prediction error (QCV) is calculated as:

$$QCV = \frac{1}{N} \sum_{i=1}^{NZ} \sum_{j=1}^{Z(i)} \left( \theta_j - \hat{\theta}_j \right)^2$$  \[4.2.2\]

where $N$ is the total number of fitted points ($N = 6356$). The average parameter vector is computed from the average parameters estimated from the cross validation procedure:

$$\bar{b} = \frac{1}{NZ} \sum_{i=1}^{SZ} b_i.$$  \[4.2.3\]
(3) Artificial neural networks

For point estimations, 7 input variables $P_{<2}$, $P_{2-20}$, $P_{20-2000}$, $\rho_b$, $\phi$, $\ln(d_g)$, and $\ln(\sigma_g)$ were used to predict 4 output variables, namely $\theta$ at -10, -33 and -1500kPa, and AWC. The network model consists of 5 hidden layers and using hyperbolic tangent activation function for hidden and output layers (see §1.2.2.3). For parametric estimations (ANP), the same input variables were used to predict the van Genuchten parameters. Parameters $\alpha$ and $n$ were transformed to $\ln(\alpha)$ and $\ln(n-1)$. It was found that the input and output variables for training had to be normalized in order to produce reliable results. The following normalization was used for each variable $v$: $(v - \bar{v})/\sigma_v$, where $\bar{v}$ and $\sigma_v$ are the mean and standard deviation of the variable. The network model consisted of 5 hidden units with the hyperbolic tangent activation function for the hidden layer and linear activation function for the output layer. The resulting weights after training were then rescaled and the predicted outputs were calculated from the rescaled weight matrices. The calculations were carried out using program MATLAB ver. 5.0 with freeware toolbox NNSYSID v1.1 (Nørgaard, 1997).

(4) Published PTFs

Published PTFs were selected from those that only use particle-size distribution and bulk density. Particle-size from the data were converted to the FAO system according to the empirical equation 4.1.2. The following published PTFs were used:

- **RNB** (Rawls and Brakensiek, 1985) which predicts Brooks-Corey parameters from $P_{50-2000}$, $P_{<2}$ and $\phi$.
  \[
  \theta_r = -0.0182 + 0.000873 P_{50-2000} + 0.00513 P_{<2} + 0.0294 \phi - 0.000154 (P_{<2})^2
  \]
  \[
  - 0.00108 P_{50-2000} \phi - 0.000182 (P_{<2} \phi)^2 + 0.000307 (P_{<2})^2 \phi
  \]
  \[
  + 0.00236 P_{<2} \phi + 0.0000054 P_{50-2000} (P_{<2})^2 - 0.00000235 (P_{<2} \phi)^2
  \]
  \[
  - 0.00236 P_{<2} \phi
  \]

- **CAM** (Campbell, 1985) which predicts the Campbell parameters from $d_g$ and $\sigma_g$.
  \[
  h_b (\text{in cm}) = -0.05 d_g^{-1/2}
  \]
  \[
  \ln(h_b) = \frac{5.340 + 0.184 P_{<2} - 2.484 \phi - 0.00214 (P_{<2})^2 - 0.0436 P_{50-2000} \phi}{0.617 P_{<2} \phi + 0.00144 (P_{50-2000} \phi)^2 - 0.00855 (P_{<2} \phi)^2 - 0.0000128 (P_{50-2000})^2 P_{<2} + 0.00895 (P_{50-2000})^2 \phi - 0.000724 (P_{<2} \phi)^2 - 0.5 \phi^2 P_{<2} + 0.0000054 P_{50-2000} (P_{<2})^2 - 0.5 \phi^2 P_{<2}}
  \]

- **CAM** (Campbell, 1985) which predicts the Campbell parameters from $d_g$ and $\sigma_g$.
  \[
  h_b (\text{in cm}) = -0.05 d_g^{-1/2}
  \]

\[\text{[4.2.4]}\]
Chapter IV – Hydraulic pedotransfer functions for Australian soil

\[ b = -20 h_b + 0.2 \sigma_g \] \hspace{1cm} [4.2.5]

- HAL (Hall et al., 1977) which predicts \( \theta \) at \(-1500\) kPa as a quadratic function of \( P<2 \).
  
  Top soils: \( PWP = 2.94 + 0.83 P<2 - 0.0054 (P<2)^2 \)
  
  Sub soils: \( PWP = 1.48 + 0.84 P<2 - 0.0054 (P<2)^2 \) \hspace{1cm} [4.2.6]

- SCH (Scheinost et al., 1997) which predicts the van Genuchten parameters using the extended nonlinear regression from \( \phi, P<2, d_g, \sigma_g \), and organic carbon. Since the data do not contain complete organic carbon measurement and considering their small content in Australian soil, the value was set to 0.
  
  \[ \begin{align*}
  \theta_i & = 0.0051 P<2 \\
  \theta_s & = 0.85 \phi + 0.0013 P<2 \\
  \alpha & = 0.023 + 0.7 d_g \text{ (in m)} \\
  n & = 0.33 + 2.6/\sigma_g \\
  m & = -1
  \end{align*} \] \hspace{1cm} [4.2.7]

- ROS (Schaap et al., 1998), which was developed using the neural-network approach. The model used six hidden units and calculation were done using the program ‘Rosetta’ available from the world-wide-web (Schaap, 1999). Two levels of input information were used:
  
  ROS4: \( \rho_b, P_{50-2000}, P_{50-2}, P<2 \)
  
  ROS6: \( \rho_b, P_{50-2000}, P_{50-2}, P<2, \theta \) at -33 and -1500 kPa.

4.2.1.3 Evaluation Criteria

All methods for estimating the PTF parameters were based on minimizing the sum of squared residuals (SSR) of measured \( \theta \) and predicted \( \hat{\theta} \). Mean residual (MR) indicated whether the prediction underestimates or overestimates:

\[ \text{MR} = \frac{1}{N} \sum_{i=1}^{N} (\theta_i - \hat{\theta}_i) \] \hspace{1cm} [4.2.8]

where \( N \) is the number of data points. The Root Mean Square of Residuals (RMSR) was used to calculate the mean accuracy of prediction:

\[ \text{RMSR} = \left( \frac{1}{N} \sum_{i=1}^{N} (\theta_i - \hat{\theta}_i)^2 \right)^{1/2}. \] \hspace{1cm} [4.2.9]

Because the water content depends on water potential, relative residuals was evaluated to predict the magnitude of the residuals:

\[ \text{RR} = \left( \frac{1}{N} \sum_{i=1}^{N} \left( \frac{\theta_i - \hat{\theta}_i}{\theta_i} \right)^2 \right)^{1/2} \times 100\%. \] \hspace{1cm} [4.2.10]
The mean deviation (MD) and root mean squared deviation (RMSD) of Tietje and Tapkenhinrichs (1993) was used to evaluate the goodness of fit for the curve.

\[
MD = \frac{1}{NK} \sum_{i=1}^{NK} \frac{1}{b-a} \int_{a}^{b} \left( \hat{\theta} - \theta \right) d(\log_{10} h)
\]  \[4.2.11\]

negative values indicate that the PTFs underestimate the prediction and positive values indicate overestimation.

\[
RMSD = \frac{1}{NK} \sum_{i=1}^{NK} \left[ \frac{1}{b-a} \int_{a}^{b} \left( \theta - \hat{\theta} \right)^2 d(\log_{10} h) \right]^{1/2}
\]  \[4.2.12\]

where \( a \) and \( b \) are the integration boundaries set at \( a = 0.1 \) and \( b \) is set according the smallest (most negative) potentials measured in the sample. Calculation of MD and RMSD was done on the logarithmic of the potential and the integration is done numerically by calculating the area under the curves within the range of each measured point. Akaike’s Information Criterion/ AIC (Akaike, 1973), a statistical Ockham’s razor:

\[
AIC = N \ln \left( \sum_{i=1}^{N} (\theta_i - \hat{\theta}_i)^2 \right) + 2np
\]  \[4.2.13\]

was used to account for the number of parameters in the model. The best model is the one that has the smallest AIC. For single-point PTF prediction, \( N \) is the number of soil samples and for parametric PTF \( N \) is the total number of water retention data, \( np \) is the number of PTF parameters. For ANN \( np \) is the number of weights used =

\[
np = (\text{No. inputs} + 1) \times \text{No. hidden units} + (\text{No. hidden units} + 1) \times \text{No. outputs}
\]

where 1 is due to bias.

4.2.2. Results and Discussion

4.2.2.1. Parametric PTFs

Selecting the best ENR

The results of different PTFs' performance for predicting \( \theta \) for the prediction and validation data sets are shown in Table 4.2.4 and 4.2.5. Using the cross-validation prediction error (QCV) to compare the performance of the extended nonlinear regression (ENR) models for the prediction set, it was found that most of the models performed similarly except for ENR6 which does not use PWP to predict \( \theta_r \). Similarly, RMSR for all ENR models have values less than 0.04 m$^3$ m$^{-3}$ except for
ENR6. This indicates that prior information on the smaller potential $\theta$ is needed to obtain reasonable predictions of $\theta$, as $\theta$ also represents $\theta$ at small potentials where $d\theta/dh \approx 0$. When this information is not available ENR7 can be used but it gives an inferior estimate.

Classifying the soil texture into three classes and extending into three functions provides better prediction of $\theta$ but the improvement is not significant. Both class and continuous PTFs (ENR7 and 8) did no better than a single PTF. The RMSR for both functions are similar i.e. 0.037 m$^3$m$^{-3}$. Tietje and Tapkenhinrichs (1993) suggested that establishment of separate PTFs for different soil types can yield good results. The results shown here do not seem to confirm this, possibly because the database here does not cover the full range of texture classes. Pachepsky et al. (1999) found that stratifying PTFs made better predictions but do not significantly improve the prediction.

Using RMSR and RMSD to compare PTFs for the validation set, it can be seen that the ENR overall gave a good performance (RMSD < 0.04 m$^3$ m$^{-3}$, accuracy ~20%) except for ENR7. By way of comparison, the PTFs of Scheinost et al. (1997) for German soil gave a RMSD of 0.035 m$^3$ m$^{-3}$ and RMSR of 0.028 m$^3$ m$^{-3}$ which predicts better than the one presented here. This may suggest that the inclusion of organic matter as a predictive variable might improve the prediction.
### Table 4.2.4. Comparison of parametric PTFs performance.

<table>
<thead>
<tr>
<th>PTF</th>
<th>np</th>
<th>Prediction set (N=6356)</th>
<th>Validation Set (N=880)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSR</td>
<td>AIC</td>
<td>QCV</td>
</tr>
<tr>
<td>ENR1</td>
<td>11</td>
<td>0.0359</td>
<td>13388.64</td>
</tr>
<tr>
<td>ENR2</td>
<td>8</td>
<td>0.0375</td>
<td>13911.83</td>
</tr>
<tr>
<td>ENR3</td>
<td>8</td>
<td>0.0375</td>
<td>13918.05</td>
</tr>
<tr>
<td>ENR4</td>
<td>8</td>
<td>0.0380</td>
<td>14081.10</td>
</tr>
<tr>
<td>ENR5</td>
<td>8</td>
<td>0.0374</td>
<td>13876.57</td>
</tr>
<tr>
<td>ENR6</td>
<td>8</td>
<td>0.0672</td>
<td>21432.69</td>
</tr>
<tr>
<td>ENR7</td>
<td>16</td>
<td>0.0471</td>
<td>16817.74</td>
</tr>
<tr>
<td>ENR8</td>
<td>16</td>
<td>0.0366</td>
<td>13611.79</td>
</tr>
<tr>
<td>ENR9</td>
<td>16</td>
<td>0.0365</td>
<td>13603.23</td>
</tr>
<tr>
<td>MRP</td>
<td>11</td>
<td>0.1102</td>
<td>27645.08</td>
</tr>
<tr>
<td>ANP</td>
<td>64</td>
<td>0.0358</td>
<td>13469.64</td>
</tr>
<tr>
<td>CAM</td>
<td>3</td>
<td>0.0660</td>
<td>21894.28</td>
</tr>
<tr>
<td>RNB</td>
<td>34</td>
<td>0.0510</td>
<td>18585.93</td>
</tr>
<tr>
<td>SCH</td>
<td>7</td>
<td>0.0530</td>
<td>19034.77</td>
</tr>
<tr>
<td>ROS4</td>
<td>52</td>
<td>0.0620</td>
<td>21187.011</td>
</tr>
<tr>
<td>ROS6</td>
<td>64</td>
<td>0.0400</td>
<td>15482.142</td>
</tr>
</tbody>
</table>

### Table 4.2.5. The average of mean deviations, root mean squared deviations and relative residuals of water retention.

<table>
<thead>
<tr>
<th>PTF</th>
<th>MD (m$^3$ m$^{-3}$)</th>
<th>RMSD (m$^3$ m$^{-3}$)</th>
<th>RR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prediction</td>
<td>Validation</td>
<td>Prediction</td>
</tr>
<tr>
<td></td>
<td>NK = 733</td>
<td>NK = 108</td>
<td>NK = 733</td>
</tr>
<tr>
<td>ANP</td>
<td>0.0054</td>
<td>-0.0003</td>
<td>0.030</td>
</tr>
<tr>
<td>ENR2</td>
<td>0.0020</td>
<td>0.0136</td>
<td>0.031</td>
</tr>
<tr>
<td>ENR6</td>
<td>0.0025</td>
<td>0.0124</td>
<td>0.033</td>
</tr>
<tr>
<td>ENR8</td>
<td>0.0008</td>
<td>0.0116</td>
<td>0.031</td>
</tr>
<tr>
<td>RNB</td>
<td>0.0115</td>
<td>-0.0030</td>
<td>0.039</td>
</tr>
<tr>
<td>SCH</td>
<td>-0.0308</td>
<td>-0.0176</td>
<td>0.051</td>
</tr>
<tr>
<td>ROS4</td>
<td>0.0156</td>
<td>0.0184</td>
<td>0.054</td>
</tr>
<tr>
<td>ROS6</td>
<td>0.0071</td>
<td>0.0130</td>
<td>0.036</td>
</tr>
</tbody>
</table>
### Table 4.2.6. Parametric PTFs developed in this study.

<table>
<thead>
<tr>
<th>PTF</th>
<th>( \theta_r ) m(^3) m(^{-3})</th>
<th>( \theta_s ) m(^3) m(^{-3})</th>
<th>( \alpha ) m(^{-1})</th>
<th>( n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENR2</td>
<td>-0.00092 P(_{&lt;2}) + 1.17748 PWP</td>
<td>0.00112 P(_{&lt;2}) + 0.8331 ( \phi )</td>
<td>1.561 + 17.046 ( d_g )</td>
<td>1.3978 + 0.0027 ( \sigma_g )</td>
</tr>
<tr>
<td>ENR6</td>
<td>0.00172 P(_{&lt;2}) + 0.7 PWP</td>
<td>0.00973 P(_{&lt;2}) + 0.842 ( \phi )</td>
<td>0.147 + 1.919 ( d_g )</td>
<td>1.3716 + 0.003077 ( \sigma_g ) - 0.01 ( \rho_b )</td>
</tr>
<tr>
<td>ENR7</td>
<td>-0.00733 + 0.00427 P(<em>{&lt;2}) + 0.00267 P(</em>{&lt;20})</td>
<td>0.00110 P(_{&lt;2}) + 0.82607 ( \phi )</td>
<td>1.361 + 16.929 ( d_g )</td>
<td>1.4062 - 0.0050 ( \sigma_g )</td>
</tr>
<tr>
<td>ENR8</td>
<td>( \theta_r ) = -0.00156 P(_{&lt;2}) + 1.22333 PWP</td>
<td>( \theta_s ) = 0.00149 P(_{&lt;2}) + 0.81851 ( \phi )</td>
<td>Sandy: ( \alpha ) = 0.984 + 15.607 ( d_g )</td>
<td>( n ) = 1.4990 - 0.0024 ( \sigma_g )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Loamy: ( \alpha ) = 1.787 + 4.152 ( d_g )</td>
<td>( n ) = 1.3373 + 0.0053 ( \sigma_g )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Clayey: ( \alpha ) = 3.186 + 181.363 ( d_g )</td>
<td>( n ) = 1.2433 + 0.00156 ( \sigma_g )</td>
</tr>
</tbody>
</table>

**Multiple regression PTFs**

Using multiple-linear regression to estimate the van Genuchten parameters (MRP method) and then estimating \( \theta_r \), gave the worst result (RMSR = 0.11 m\(^3\) m\(^{-3}\)). It generally underestimated water content (Figure 4.2.1), which shows that this method is not satisfactory when used for parametric PTF. Usually it is difficult to find the relationship between soil properties and the scaling and shape parameters \( \alpha \) and \( n \).

Tietje and Tapkenhinrichs (1993) found no correlation between \( \alpha \) and \( P_{<2} \), \( P_{2-50} \) and \( P_{50-2000} \). One explanation is that when parameters of the hydraulic model are estimated from nonlinear regression some error in prediction has already been embedded, this could result in large uncertainty when correlating the parameters with predictive variables. Further the parameters of the van Genuchten equation are highly correlated, such a large parameter correlation usually is an indicator of over-parameterization (Ratkowsky, 1990). This may explain why the published PTF RNB (Rawls and Brakensiek, 1985), which estimates Brooks-Corey parameters predicts better than MRP.
Figure 4.2.1. Measured vs. predicted water content by using (a) parametric multiple regression (MRP), (b) neural networks (ANP), and (c) extended non-linear regression (ENR).

**Artificial neural networks**

For the prediction set, neural networks’ (ANP) performance is comparable with that of ENR2 (RMSR = 0.036 m$^3$m$^{-3}$, RR = 14%). By using only bulk density and particle-size distribution data, neural networks are able to find a relationship between basic soil properties and the van Genuchten parameters. However, for the validation data set it gave poorer prediction than ENR.
Figure 4.2.2. Measured and PTFs predicted water retention curve for four different type of soil.

According to Schaap and Bouten (1996), neural networks are useful in developing PTFs because they do not require an \textit{a priori} model. In our case the neural networks approach did not improve the accuracy of $\theta$ prediction compared with ENR. This result was also recognised by Tamari and Wösten (1999). Their study showed that neural networks were less accurate than MLR when a small database was used (< 1000 samples), but can be useful if a large soil database with good quality data is available.
Based on the AIC, the best PTFs are ENR2 and ENR7 (continuous PTF). Considering the number of parameters used and the fact that the data set does not cover the full texture range, the single PTF (ENR2) is considered as the most adequate PTF. Comparison of MRP, ENR and ANN on \( \theta \) prediction can be seen in Figure 4.2.1 and 4.2.2.

**Comparing published PTFs**

Published PTFs of RNB gave a good prediction with RMSD 0.04 m\(^3\)m\(^{-3}\). Tietje and Tapkenhinrichs (1993) found that RNB gave an average RMSD 0.075 m\(^3\)m\(^{-3}\) when applied to German soil. Meanwhile, Campbell’s PTF displays the least ability with RMSR = 0.07 m\(^3\)m\(^{-3}\). PTFs of SCH showed similar predicability as RNB. The neural-networks ROS4 that only use bulk density and particle size data performed badly but when inclusion of FC and PWP as input variables (ROS6) the performance prevails over other published PTFs and is comparable with the PTFs developed in this study. It is not surprising that the more relevant inputs are fed into the PTF, the better the prediction is. As deduced by Bastet *et al.* (1997), the general applicability of published PTFs should not be assumed without question.

**4.2.2.2 Point Estimation PTF**

Comparison between different PTFs developed to estimate \( \theta \) at -10, -33 and -1500 kPa and AWC are given in Table 4.2.7. ANN gave the least RMSR compared with other methods, followed by MLR and ENR except for \( \theta_{-1500} \) and AWC. MLR and ANN gave better predictions for the point estimation. van den Berg *et al.* (1997) also reported similar results. ENR gave better prediction at \( \theta_{-1500} \) because PWP was included as a variable in estimating \( \theta \).

\[
\theta_{-10} = 0.1728 + 0.64389 \phi - 0.24146 \phi^2 + 0.001735 P_{2-20} - 0.00003 (P_{2-20})^2 \\
- 0.00102 P_{20-2000} - 0.00001 (P_{20-2000})^2 - 0.14684 d_g
\]  
\[4.2.14\]

\[
\theta_{-33} = -0.39350 + 0.29300 \phi + 0.00855 P_{<2} - 0.00003 (P_{<2})^2 + 0.00754 P_{2-20} \\
- 0.00003 (P_{2-20})^2 + 0.00476 P_{20-2000} - 0.00001 (P_{20-2000})^2 - 0.05326 d_g
\]  
\[4.2.15\]

\[
\theta_{-1500} = -2.41449 + 0.49579 \phi - 0.53838 \phi^2 + 0.03059 P_{<2} - 0.00005 (P_{<2})^2 + 0.02506 P_{2-20} + 0.02151 P_{20-2000} + 0.00001 (P_{20-2000})^2 + 0.01217 d_g^2 + 0.0016727 \sigma_g
\]  
\[4.2.16\]
AWC = 1.97571 – 0.19669 φ + 0.52921 φ^2 - 0.02170 P<2 + 0.00003 (P<2)^2 – 0.0172 P2-20 – 0.00002 (P2-20)^2 – 0.01639 P20-2000 – 0.0000252 (P20-2000)^2 – 0.07054 dg – 0.00131 σ_g

[4.2.17]

### Table 4.2.7. Comparison of PTFs performance in predicting θ at -10, -33, -1500 kPa and AWC on the validation set

<table>
<thead>
<tr>
<th>PTF</th>
<th>-10 kPa RMSR</th>
<th>-10 kPa AIC</th>
<th>-33 kPa RMSR</th>
<th>-33 kPa AIC</th>
<th>-1500 kPa RMSR</th>
<th>-1500 kPa AIC</th>
<th>AWC RMSR</th>
<th>AWC AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENR2</td>
<td>0.096</td>
<td>-19.59</td>
<td>0.091</td>
<td>-17.12</td>
<td>0.023</td>
<td>-277.44</td>
<td>0.095</td>
<td>116.23</td>
</tr>
<tr>
<td>MLR</td>
<td>0.091</td>
<td>-31.62</td>
<td>0.090</td>
<td>-54.23</td>
<td>0.051</td>
<td>-52.23</td>
<td>0.115</td>
<td>0.08</td>
</tr>
<tr>
<td>ANN</td>
<td>0.072</td>
<td>30.79</td>
<td>0.072</td>
<td>30.79</td>
<td>0.072</td>
<td>30.79</td>
<td>0.062</td>
<td>-0.90</td>
</tr>
<tr>
<td>ANP</td>
<td>0.075</td>
<td>74.62</td>
<td>0.065</td>
<td>43.19</td>
<td>0.047</td>
<td>-25.69</td>
<td>0.084</td>
<td>99.52</td>
</tr>
<tr>
<td>ROS4</td>
<td>0.079</td>
<td>61.71</td>
<td>0.100</td>
<td>115.23</td>
<td>0.096</td>
<td>104.13</td>
<td>0.061</td>
<td>4.68</td>
</tr>
<tr>
<td>HALL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.063</td>
<td>-85.62</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 4.2.8. Comparison of MLR and ANN performance on the prediction set

<table>
<thead>
<tr>
<th>PTF</th>
<th>-10 kPa RMSR</th>
<th>-10 kPa AIC</th>
<th>-33 kPa RMSR</th>
<th>-33 kPa AIC</th>
<th>-1500 kPa RMSR</th>
<th>-1500 kPa AIC</th>
<th>AWC RMSR</th>
<th>AWC AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLR</td>
<td>0.124</td>
<td>71.72</td>
<td>0.129</td>
<td>82.15</td>
<td>0.106</td>
<td>42.22</td>
<td>0.103</td>
<td>38.88</td>
</tr>
<tr>
<td>ANN</td>
<td>0.123</td>
<td>146.95</td>
<td>0.126</td>
<td>152.29</td>
<td>0.105</td>
<td>112.05</td>
<td>0.101</td>
<td>104.74</td>
</tr>
<tr>
<td>ROS4</td>
<td>0.130</td>
<td>170.96</td>
<td>0.176</td>
<td>236.64</td>
<td>0.221</td>
<td>286.02</td>
<td>0.133</td>
<td>175.50</td>
</tr>
</tbody>
</table>

The PTF of Hall et al. (1977) gave reasonable estimates of θ_{-1500} (RMSR = 0.063 m^3 m^{-3}) although it was developed for soil in England and Wales. It even predicts better than ANN. HALL predicts θ_{-1500} as a quadratic function of P<2. The strong relationship with P<2 has been previously recognised (e.g. Nielsen and Shaw, 1958). Hall et al. (1977) explained that as pore size becomes finer, potential decreases and the influence of clay content on θ increases. van den Berg et al. (1997) also noted that PTFs developed for tropical soils showed a consistent relationship for PWP and P<2 content, but not for θ_{-33} and AWC. It can be seen that θ at -10, -33 and -1500 kPa had an exponentially increasing trend with clay (P<2) and linearly decreasing trend with P_{20-2000} (Figure 4.6). Exponential function can be fitted to the data:

\[
\begin{align*}
θ_{-10} &= 0.3064 \left(1 - \exp(-0.0406 P_{<2})\right) + 0.1452 \\
θ_{-33} &= 0.3543 \left(1 - \exp(-0.0385 P_{<2})\right) + 0.0830 \\
θ_{-1500} &= 0.4016 \left(1 - \exp(-0.0230 P_{<2})\right) + 0.0027
\end{align*}
\]

[4.2.18]
Figure 4.2.3. Relationship between clay and water content at –33 and –1500kPa (b) relationship between sand and water content at –33 and –1500kPa.

The scatter in points around the regressions in Figures 4.2.3 is due no doubt to variations in structure, organic carbon and measurement errors. Consequently, by taking the difference between the function of $\theta_{-33}$ and $\theta_{-1500}$, we can derive the relationship for AWC as a function of $P_{<2}$. From Figure 4.2.4 it can be seen that $P_{<2}$ has little effect on available water content, but there’s a trend that AWC increases with increasing $P_{<2}$ content until about 25% before it slowly decreases. This is because at low $P_{<2}$ content a relatively small amount of water is held at lower
potentials and higher amount at high potentials. But as \( P_{<2} \) increases, more water is held at a lower potential as explained earlier resulting in less available water.

Table 4.2.8 shows the comparison between MLR and ANN for point estimation on the prediction set. ANN’s performance is only slight better than MLR. When taking into account the number of parameters used ANN always gave a higher AIC value because of the high number of parameters.

Figure 4.2.4. Relationship between clay and available water content.

### 4.2.2.3. Application to other dataset

To test the whether the PTFs developed can be applied to data from other areas, public-domain water retention data were used. The data sets are GRIZZLY and UNSODA, which is a compilation of numerous mostly laboratory water retention from US, Europe and other countries. The PTFs used are ENR2 developed in this paper and the neural-network ROS4 and ROS6. This neural-network PTF was found to be better than other published PTFs (Imam et al., 1999). To make fairer comparison the PWP for use in ENR2 is predicted from PTF (Equation 4.2.16).

The MR and RMSR for both data sets are given in Table 4.2.9. As can be seen the PTFs performed similarly for both data sets following the same trend as before ROS4 < ENR2 < ROS6. For the UNSODA data set ROS4 perform slightly better than ENR2, as the UNSODA were used to develop the PTFs (Schaap et al., 1998). ENR2 predicts water retention with the accuracy of 19%.
Table 4.2.9. Mean residuals (MR) and root mean squared residuals (RMSR) for water retention PTFs on the GRIZZLY and UNSODA database.

<table>
<thead>
<tr>
<th>Data set</th>
<th>ENR2</th>
<th>ROS4</th>
<th>ROS6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MR</td>
<td>RMSR</td>
<td>MR</td>
</tr>
<tr>
<td>GRIZZLY</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min</td>
<td>-0.287</td>
<td>0.008</td>
<td>-0.308</td>
</tr>
<tr>
<td>Max</td>
<td>0.058</td>
<td>0.102</td>
<td>0.104</td>
</tr>
<tr>
<td>Average</td>
<td>-0.053</td>
<td>0.037</td>
<td>-0.050</td>
</tr>
<tr>
<td>RR</td>
<td>19%</td>
<td>20%</td>
<td>12%</td>
</tr>
<tr>
<td>UNSODA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min</td>
<td>-0.235</td>
<td>0.003</td>
<td>-0.175</td>
</tr>
<tr>
<td>Max</td>
<td>0.168</td>
<td>0.188</td>
<td>0.171</td>
</tr>
<tr>
<td>Average</td>
<td>-0.024</td>
<td>0.037</td>
<td>0.026</td>
</tr>
<tr>
<td>RR</td>
<td>18%</td>
<td>18%</td>
<td>13%</td>
</tr>
</tbody>
</table>

When applying the prediction to the soil-water limits for plants, parametric ENR and point estimation MLR predict as well (or as badly) as each other (Table 4.2.10). The prediction is slightly better than ROS4. As seen in Figure 4.2.5 the drainage upper limit DUL is predicted quite well by ENR, slight overestimation is observed in the lower-limit water content LOL. Good correlation is achieved when predicting individual points but when applying the difference (Figure 4.2.5c) the prediction is quite bad and underestimates potential extractable water PLEXW. When taking the difference between two predictions, error propagation is bound to happen.

Theoretically the variance of error in prediction of PLEXW is:

\[
\sigma_{\epsilon}^2_{\text{PLEXW}} = \sigma_{\epsilon}^2_{\text{DUL}} + \sigma_{\epsilon}^2_{\text{LOL}} - 2 \text{Cov}(\epsilon_{\text{DUL}}, \epsilon_{\text{LOL}}).
\]  [4.2.19]

The variance of the error does not only depend on the error in DUL and LOL but also the covariance between the two variables. The standard deviation of error in DUL \((\sigma_{\epsilon_{\text{DUL}}}) = 0.05\), \(\sigma_{\epsilon_{\text{LOL}}} = 0.06\) and correlation between the two errors is 0.76, resulting in the standard deviation error of 0.04 m^3 m^{-3}. The correlation between DUL and LOL is also high (0.89), using ENR2 the correlation between predicted DUL and LOL becomes much larger (0.98).

Table 4.2.10. Mean residuals and root mean squared residuals for soil-water limits.

<table>
<thead>
<tr>
<th></th>
<th>DUL</th>
<th>LOL</th>
<th>PLEXW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ENR2</td>
<td>ROS4</td>
<td>MLR</td>
</tr>
<tr>
<td>MR</td>
<td>0.017</td>
<td>0.034</td>
<td>-0.011</td>
</tr>
<tr>
<td>RMSR</td>
<td>0.056</td>
<td>0.070</td>
<td>0.058</td>
</tr>
</tbody>
</table>
4.2.2.4. Some afterthoughts

So far it has been shown that ENR predicts water retention quite well compared with other methods. But when relating predicted parameters $\alpha$ and $n$ with calculated values, ENR does not seem to produce a good estimate (Figure 4.2.6). It can be argued that the predicted water content depends on the conjunction of all the parameters. Wösten and van Genuchten (1988) showed the sensitivity of the water-retention curve by varying one parameter and holding the others constant. In Figure 4.2.7 it can be shown that varying more than two parameters simultaneously could well produce a similar curve which fit the data. The parameters should be interpreted empirically rather than physically.

Figure 4.2.5. Soil-plant water limits as predicted by PTFs (a) drainage upper limit, (b) lower limit, and (c) plant extractable water.
Figure 4.2.6. Observed and ENR predicted shape factor $\alpha$ and $n$ of the van Genuchten parameters.

The distribution of the van Genuchten parameters was evaluated for different PTFs (Table 4.2.11). For residual water content $\theta_r$, ENR2 and ENR8 seem to overestimate its value, as evidenced by the coefficient 1.17 applied in the equation. ANP, ENR6, ROS4 and ROS6 gave reasonable ranges. A limitation of ENR2 is that at very low clay content, $\theta_r$ could be negative. If this is encountered, the value is set to 0. For saturated water content, similar prediction range was found by all the methods because it mostly depends on porosity. The ENR methods produce a larger range for the scaling parameter $\alpha$ compared with ROS4 and ROS6. ANP gave an $\alpha$ range which is too high.
Chapter IV – Hydraulic pedotransfer functions for Australian soil

Figure 4.2.7. Varying the van Genuchten parameters may produce similar fit to water retention data.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>ANP</th>
<th>ENR2</th>
<th>ENR6</th>
<th>ENR8</th>
<th>ROS4</th>
<th>ROS6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_r$</td>
<td>Median</td>
<td>0.060</td>
<td>0.176</td>
<td>0.159</td>
<td>0.163</td>
<td>0.066</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.000</td>
<td>0.018</td>
<td>0.017</td>
<td>0.016</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.122</td>
<td>0.400</td>
<td>0.367</td>
<td>0.391</td>
<td>0.122</td>
</tr>
<tr>
<td></td>
<td>80% range*</td>
<td>0.080</td>
<td>0.253</td>
<td>0.236</td>
<td>0.240</td>
<td>0.061</td>
</tr>
<tr>
<td>$\theta_s$</td>
<td>Median</td>
<td>0.399</td>
<td>0.397</td>
<td>0.394</td>
<td>0.399</td>
<td>0.400</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.243</td>
<td>0.231</td>
<td>0.232</td>
<td>0.232</td>
<td>0.275</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.655</td>
<td>0.691</td>
<td>0.693</td>
<td>0.694</td>
<td>0.741</td>
</tr>
<tr>
<td></td>
<td>80% range*</td>
<td>0.176</td>
<td>0.178</td>
<td>0.167</td>
<td>0.188</td>
<td>0.176</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Median</td>
<td>2.61</td>
<td>2.09</td>
<td>2.10</td>
<td>2.81</td>
<td>1.68</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.44</td>
<td>1.61</td>
<td>1.50</td>
<td>1.47</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>419.99</td>
<td>19.66</td>
<td>14.90</td>
<td>177.68</td>
<td>5.45</td>
</tr>
<tr>
<td></td>
<td>80% range*</td>
<td>7.44</td>
<td>3.47</td>
<td>2.50</td>
<td>3.95</td>
<td>2.13</td>
</tr>
<tr>
<td>$n$</td>
<td>Median</td>
<td>1.21</td>
<td>1.42</td>
<td>1.39</td>
<td>1.40</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>1.02</td>
<td>1.41</td>
<td>1.36</td>
<td>1.25</td>
<td>1.16</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>1.75</td>
<td>1.47</td>
<td>1.44</td>
<td>1.49</td>
<td>3.22</td>
</tr>
<tr>
<td></td>
<td>80% range*</td>
<td>0.32</td>
<td>0.03</td>
<td>0.03</td>
<td>0.22</td>
<td>0.30</td>
</tr>
</tbody>
</table>

*80% prediction limits around the median, calculated from 90 percentiles minus 10 percentiles.
Table 4.2.12. Distribution of the van Genuchten parameters predicted by ENR2.

<table>
<thead>
<tr>
<th>Database</th>
<th>$\theta_r$</th>
<th>$\theta_s$</th>
<th>$\alpha$</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRIZZLY Median</td>
<td>0.096</td>
<td>0.386</td>
<td>3.46</td>
<td>1.431</td>
</tr>
<tr>
<td>NK = 660 Min</td>
<td>0.000</td>
<td>0.210</td>
<td>1.59</td>
<td>1.400</td>
</tr>
<tr>
<td>NK = 660 Max</td>
<td>0.334</td>
<td>0.687</td>
<td>18.78</td>
<td>1.464</td>
</tr>
<tr>
<td>NK = 660 80% range</td>
<td>0.293</td>
<td>0.194</td>
<td>14.79</td>
<td>0.044</td>
</tr>
<tr>
<td>UNSODA Median</td>
<td>0.096</td>
<td>0.380</td>
<td>3.18</td>
<td>1.421</td>
</tr>
<tr>
<td>NK = 533 Min</td>
<td>0.000</td>
<td>0.220</td>
<td>1.63</td>
<td>1.402</td>
</tr>
<tr>
<td>NK = 533 Max</td>
<td>0.331</td>
<td>0.789</td>
<td>17.54</td>
<td>1.469</td>
</tr>
<tr>
<td>NK = 533 80% range</td>
<td>0.202</td>
<td>0.178</td>
<td>11.19</td>
<td>0.035</td>
</tr>
</tbody>
</table>

For shape parameter $n$, the ENR gave a very narrow range around 1.4. ENR tends to scale the shape factor around a constant value. This is also observed in other data sets (Table 4.2.12). As can be seen in Figure 4.2.8b, the scaled water content ($S_e$) and potential ($\alpha h$) using ENR produce a mean curve. The PTFs significantly reduce the variations and bring them to the mean. This is because in ENR we are essentially fitting one equation to the whole data set. Larger variation is observed when using the neural-networks PTF Figure 4.2.8c and d. The predictability of ENR is quite similar for different textures (Figure 4.2.9) and also comparable with ROS6. This implies that for large-scale prediction the ENR can be used quite well as demonstrated in the large data sets derived from different locations and countries. If it was going to be used for fine-scale spatial prediction, it may pose a problem, because it reduces the variation.
Figure 4.2.8. Scaled water content and water potential for (a) original data (b) predicted by ENR (c) ANP and, (d) ROS6.
4.2.3. Conclusions

Point estimation and parametric pedotransfer functions have been developed for Australian soil. Because of the unique soil properties, different cutoffs used in particle-size classification, and unavailability of data on some soil properties, most PTFs developed elsewhere cannot be applied directly. Using particle-size distribution and bulk density data, the PTFs developed predict water content at different potentials with reasonable accuracy. Difficulties have been encountered when relating the scaling and shape parameters of the van Genuchten equation to soil properties, because of over-parameterization. This can be countered by using extended nonlinear regression as discussed in this paper.

Parametric estimation using ENR predicts water content better than ANN and MRP. This is because ENR is developed by fitting simultaneously all the soil water retention data as opposed to the ANN and MRP, which predict the parameters obtained from individual water-retention curve fits. The performance of neural networks is however comparable with that of extended nonlinear regression. Since the latter is considered as a black-box approach, the former method with a more physical basis is preferred at the current time.

Figure 4.2.9. Median of water retention curve RMSD based on texture for (a) ENR2 and (b) ROS6. The error bar represent 10 and 90 percentiles.
4.2. Water retention curve PTFs

The water-retention curve, which defines the relationship between soil water content ($\theta$) and hydraulic potential ($h$) is an important physical property of soil material. This study compared different approaches for deriving point estimation and parametric PTFs from particle-size and bulk density data. The performances of developed PTFs for estimating water retention for Australian soil were also compared.

4.2.1. Methods

4.2.1.1. The data

The water-retention from the Australian database was used to develop the PTFs and a subset of the data was randomly selected and used as a validation set (Table 4.2.1). A summary of the soil properties is given on Table 4.2.2.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>All data</th>
<th>Prediction set</th>
<th>Validation Set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Profiles</td>
<td>Profiles</td>
<td>Profiles</td>
</tr>
<tr>
<td></td>
<td>Samples</td>
<td>Samples</td>
<td>Samples</td>
</tr>
<tr>
<td>Prebble</td>
<td>17</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>Forrest</td>
<td>60</td>
<td>54</td>
<td>6</td>
</tr>
<tr>
<td>Geeves</td>
<td>74</td>
<td>60</td>
<td>14</td>
</tr>
<tr>
<td>Smettem</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Bristow</td>
<td>46</td>
<td>42</td>
<td>4</td>
</tr>
<tr>
<td>SydU</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>218</td>
<td>174</td>
<td>42</td>
</tr>
</tbody>
</table>

To test the applicability of the developed PTFs, published databases available from the world-wide web were used. The first is GRIZZLY (Haverkamp et al., 1997) from Laboratoire d’Etude des Transfers en Hydrologie et Environnement (LTHE), Grenoble, which consists of 660 soil water retention from different countries. This database has been used by Bouraoui et al. (1999) to develop a physically-based PTF. The mass of particles $P_{2-20}$ and $P_{20-2000}$ were calculated from the van Genuchten type cumulative particle-size distribution function provided in the program. The second database UNSODA (Nemes et al. 1999) is from the US Salinity Laboratory. This data set has been used extensively by researchers in US to develop PTFs (Schaap et al., 1998; Arya et al., 1999). 533 water retention sets are used and the sand, silt fraction of
the USDA classification were converted to the international system according to equation 4.1.3.

<table>
<thead>
<tr>
<th>Properties</th>
<th>Units</th>
<th>Prediction set</th>
<th>Validation set</th>
<th>prob $t^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_b$</td>
<td>Mg m$^{-3}$</td>
<td>1.43</td>
<td>0.22</td>
<td>1.42</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Mg m$^{-3}$</td>
<td>0.46</td>
<td>0.08</td>
<td>0.47</td>
</tr>
<tr>
<td>$P_{&lt;2}$</td>
<td>dag kg$^{-1}$</td>
<td>27.41</td>
<td>16.19</td>
<td>32.78</td>
</tr>
<tr>
<td>$P_{2-20}$</td>
<td>dag kg$^{-1}$</td>
<td>18.11</td>
<td>10.25</td>
<td>19.88</td>
</tr>
<tr>
<td>$P_{20-2000}$</td>
<td>dag kg$^{-1}$</td>
<td>54.47</td>
<td>21.36</td>
<td>47.33</td>
</tr>
<tr>
<td>$\ln(d_g)$</td>
<td>ln (mm)</td>
<td>-3.42</td>
<td>1.16</td>
<td>-2.88</td>
</tr>
<tr>
<td>$\ln(\sigma_g)$</td>
<td>ln (mm)</td>
<td>2.28</td>
<td>0.39</td>
<td>2.11</td>
</tr>
<tr>
<td>$\theta_{(-10)}$</td>
<td>m$^3$m$^{-3}$</td>
<td>0.33</td>
<td>0.10</td>
<td>0.36</td>
</tr>
<tr>
<td>$\theta_{(-33)}$</td>
<td>m$^3$m$^{-3}$</td>
<td>0.29</td>
<td>0.10</td>
<td>0.33</td>
</tr>
<tr>
<td>$\theta_{(-1500)}$</td>
<td>m$^3$m$^{-3}$</td>
<td>0.18</td>
<td>0.09</td>
<td>0.19</td>
</tr>
<tr>
<td>AWC</td>
<td>m$^3$m$^{-3}$</td>
<td>0.12</td>
<td>0.05</td>
<td>0.14</td>
</tr>
</tbody>
</table>

* probability associated with $t$ test for two samples come from distributions with the same means.

* samples significantly come from distributions with the same means.

The third data set came from the study of Ratcliffe et al. (1983) where plant-available soil water limits were determined. The measured properties are:

- drainage upper limit (DUL) – the largest field-measured soil water content after it had been saturated and allowed to drain until drainage became negligible, or equivalent to $\theta$ at -33 kPa (FC);
- lower limit (LOL) – the smallest field measured soil water content after plants has stopped extracting water and were at or near water stress condition, or equivalent to $\theta$ at -1500 kPa (PWP);
- potential extractable soil water (PLEXW) – the difference between DUL and LOL, or equivalent to AWC.

The data were derived from 61 horizons across the US, and 312 data that contain particle size distribution and bulk density were selected. Sand and silt fractions were converted to the ISSS system by using equation 4.1.3.
4.2.1.2. Pedotransfer function development

Two types of PTF were considered:

- **point estimation**, where water contents were estimated at potentials of -1, -33 and -1500 kPa. Water content at -1500 kPa $\theta_{-1500}$ is referred to as the Permanent Wilting Point (PWP). Available water content (AWC) is defined as:
  \[ AWC = \theta_{-33} - \theta_{-1500}; \]
- **parametric estimation**, where the parameters of the van Genuchten equation: $\theta_r$, $\theta_s$, $\alpha$ and $n$ are estimated.

Four different approaches were used to produce the PTFs:

1. **Multiple linear regression**
   Multiple linear regression was used to develop PTFs for single-point estimation (MLR) and also parametric estimation (MRP). Variables used to predict $\theta$ were selected using stepwise regression.

2. **Extended nonlinear regression (ENR)**
   The extended nonlinear regression (ENR) method of Scheinost et al. (1997) was used to develop 6 different combinations of variables for predicting the van Genuchten parameters (Table 4.2.3) based on the following assumptions:
   - $\theta_i$ is predicted from $P_{<2}^2$ and PWP, as $\theta_i$ is related to the asymptotic $\theta$ at the dry end ($< -1500$ kPa) which in turn is related to $P_{<2}$ (Scheinost et al., 1997). When PWP data are not available, $\theta_i$ is predicted from $P_{2-20}$ and $P_{<2}$.
   - $\theta_s$ is predicted from $P_{<2}$ and $\phi$ (Scheinost et al., 1997).
   - Campbell (1985) found that $h_b$ ($\approx 1/\alpha$) is a function of $d_g^{-0.5}$ and $b$ is a function of $h_b$ and $\sigma_g$. Scheinost et al. (1997) proposed that $\alpha$ is related to $d_g$ and $n$ related to $1/\sigma_g$ and Bouraoui et al. (1997) found that $n$ is related to $\ln(P_{<2}/(P_{<2} + P_{2.50}))$.

   PTFs were also developed by stratifying the texture into 3 classes (sandy, loamy, clayey) as described in 4.1.5. PTFs were developed for each class by holding $\theta_r$ and $\theta_s$ constant and varying $\alpha$ and $n$. The class PTF for each soil depends solely on the class to which it belongs, while the continuous PTF is calculated from
   \[ \text{PTF} = u_{\text{sandy}} \text{PTF}_{\text{sandy}} + u_{\text{loamy}} \text{PTF}_{\text{loamy}} + u_{\text{clayey}} \text{PTF}_{\text{clayey}} \]
   where $u$ is the membership of the particular class.
Table 4.2.3. Extended nonlinear regression (ENR) models for parametric PTFs

<table>
<thead>
<tr>
<th>PTF</th>
<th>np</th>
<th>( \theta_i )</th>
<th>( \theta_j )</th>
<th>( \alpha )</th>
<th>( n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENR1</td>
<td>11</td>
<td>( r_0 + r_1 ) P&lt;2 ( + ) s0 + s1 P&lt;2 + s2</td>
<td>( \phi ) a0 + a1 d&lt;sub&gt;g&lt;/sub&gt;</td>
<td>n0 + n1 ( \sigma_g ) + n2 d&lt;sub&gt;g&lt;/sub&gt;</td>
<td></td>
</tr>
<tr>
<td>ENR2</td>
<td>8</td>
<td>( r_1 ) P&lt;2 ( + r_2 ) PWP</td>
<td>s1 P&lt;2 + s2 ( \phi )</td>
<td>a0 + a1 d&lt;sub&gt;g&lt;/sub&gt;</td>
<td>n0 + n1 ( \sigma_g )</td>
</tr>
<tr>
<td>ENR3</td>
<td>8</td>
<td>( r_1 ) P&lt;2 ( + r_2 ) PWP</td>
<td>s1 P&lt;2 + s2 ( \phi )</td>
<td>a0 + a1 d&lt;sub&gt;g&lt;/sub&gt; ( ^{1/2} )</td>
<td>n0 + n1/( \sigma_g )</td>
</tr>
<tr>
<td>ENR4</td>
<td>8</td>
<td>( r_1 ) P&lt;2 ( + r_2 ) PWP</td>
<td>s1 P&lt;2 + s2 ( \phi )</td>
<td>a0 + a1 d&lt;sub&gt;g&lt;/sub&gt;</td>
<td>n0 + n1</td>
</tr>
<tr>
<td>ENR5</td>
<td>8</td>
<td>( r_1 ) P&lt;2 + 0.7 PWP</td>
<td>s1 P&lt;2 + s2 ( \phi )</td>
<td>a0 + a1 d&lt;sub&gt;g&lt;/sub&gt; ( + n_2 \rho_b )</td>
<td></td>
</tr>
<tr>
<td>ENR6</td>
<td>8</td>
<td>( r_0 + r_1 ) P&lt;2 + ( r_2 ) P&lt;2 ( - 20 )</td>
<td>s1 P&lt;2 + s2 ( \phi )</td>
<td>a0 + a1 d&lt;sub&gt;g&lt;/sub&gt;</td>
<td>n0 + n1 ( \sigma_g ) + n2 ( \rho_b )</td>
</tr>
<tr>
<td>ENR7</td>
<td>9</td>
<td>Continuous function of ENR2</td>
<td>PTF = ( u_{sandy} ) PTF&lt;sub&gt;sandy&lt;/sub&gt; ( + u_{loamy} ) PTF&lt;sub&gt;loamy&lt;/sub&gt; ( + u_{clayey} ) PTF&lt;sub&gt;clayey&lt;/sub&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENR8</td>
<td>16</td>
<td>Class function of ENR2</td>
<td>If sandy then PTF&lt;sub&gt;sandy&lt;/sub&gt; else if loamy then PTF&lt;sub&gt;loamy&lt;/sub&gt; else if clayey then PTF&lt;sub&gt;clayey&lt;/sub&gt;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The extended van Genuchten model was fitted to all water retention data using nonlinear regression with the Levenberg-Marquardt algorithm (Marquardt, 1963) which minimises the sum of squared residuals (SSR):

\[
\text{SSR} = \sum_{i=1}^{NZ} \sum_{j=1}^{Z(i)} (\theta_j - \hat{\theta}_j(b))^2
\]  

where \( \mathbf{b} \) is the parameter vector = \{r, s, a, n\} that replaces \{\( \theta_r \), \( \theta_s \), \( \alpha \), n\} in the van Genuchten equation, \( Z(i) \) is the number of water retention data for soil sample \( i \) and \( NW \) is the number of soil samples (\( NZ = 733 \)). To test and compare the performance of the PTFs, ‘leave-one-out’ cross validation (Efron and Tibshirani, 1993) was performed. For each soil sample \( i \) of measured water retention the model was fitted leaving out that sample, and then the predicted value for the sample as \( \sum_{j=1}^{Z(i)} \hat{\theta}_j \) was computed. The estimated cross-validation prediction error (QCV) is calculated as:

\[
\text{QCV} = \frac{1}{N} \sum_{i=1}^{NZ} \sum_{j=1}^{Z(i)} (\theta_j - \hat{\theta}_j)^2
\]  

where \( N \) is the total number of fitted points (\( N = 6356 \)). The average parameter vector is computed from the average parameters estimated from the cross validation procedure:

\[
\overline{\mathbf{b}} = \frac{1}{NZ} \sum_{i=1}^{NZ} \mathbf{b}_i
\]
(3) Artificial neural networks

For point estimations, 7 input variables $P_{<2}$, $P_{2-20}$, $P_{20-2000}$, $\rho_b$, $\phi$, $\ln(d_g)$, and $\ln(\sigma_g)$ were used to predict 4 output variables, namely $\theta$ at -10, -33 and -1500kPa, and AWC. The network model consists of 5 hidden layers and using hyperbolic tangent activation function for hidden and output layers (see §1.2.2.3). For parametric estimations (ANP), the same input variables were used to predict the van Genuchten parameters. Parameters $\alpha$ and $n$ were transformed to $\ln(\alpha)$ and $\ln(n-1)$. It was found that the input and output variables for training had to be normalized in order to produce reliable results. The following normalization was used for each variable $v$: $(v - \bar{v}) / \sigma_v$, where $\bar{v}$ and $\sigma_v$ are the mean and standard deviation of the variable. The network model consisted of 5 hidden units with the hyperbolic tangent activation function for the hidden layer and linear activation function for the output layer. The resulting weights after training were then rescaled and the predicted outputs were calculated from the rescaled weight matrices. The calculations were carried out using program MATLAB ver. 5.0 with freeware toolbox NNSYSID v1.1 (Nørgaard, 1997).

(4) Published PTFs

Published PTFs were selected from those that only use particle-size distribution and bulk density. Particle-size from the data were converted to the FAO system according to the empirical equation 4.1.2. The following published PTFs were used:

- **RNB** (Rawls and Brakensiek, 1985) which predicts Brooks-Corey parameters from $P_{50-2000}$, $P_{<2}$ and $\phi$.

  \[
  \theta_i = -0.0182 + 0.000873 P_{50-2000} + 0.00513 P_{<2} + 0.0294 \phi - 0.000154 (P_{<2})^2 \\
  - 0.00108 P_{50-2000} \phi - 0.000182 (P_{<2})^2 \phi \\
  + 0.000307 (P_{<2})^3 \phi - 0.00236 P_{<2} \phi + 0.000236 P_{<2} \phi^2 \\
  - 0.00154 (P_{<2})^2 \phi - 0.00108 P_{50-2000} \phi + 0.000054 P_{50-2000} (P_{<2})^2 - 0.5 \phi^2 P_{<2}
  \]

  \[
  \ln(h_b) = 5.340 + 0.184 P_{<2} - 2.484 \phi - 0.00214 (P_{<2})^2 - 0.0436 P_{50-2000} \phi - \\
  0.617 P_{<2} \phi + 0.00144 (P_{50-2000} \phi)^2 - 0.0855 (P_{<2} \phi)^2 - \\
  0.000128 (P_{50-2000})^2 P_{<2} + 0.00895 (P_{<2})^2 \phi - 0.00724 (P_{50-2000})^2 \phi + \\
  0.0000054 P_{50-2000} (P_{<2})^2 - 0.5 \phi^2 P_{<2}
  \]

  \[
  \ln(\lambda) = -0.784 + 0.0177 P_{50-2000} - 1.062 \phi - 0.000053 (P_{50-2000})^2 - 0.00273 (P_{<2})^2 + 1.11 \phi - 0.0309 P_{50-2000} \phi + 0.000266 (P_{50-2000} \phi)^2 - 0.0061 (P_{<2} \phi)^2 - 0.00000235 (P_{50-2000})^2 \phi + 0.00799 (P_{<2})^2 \phi - 0.00674 P_{<2} \phi^2
  \]

- **CAM** (Campbell, 1985) which predicts the Campbell parameters from $d_g$ and $\sigma_g$.

  \[
  h_b \text{ (in cm)} = -0.05 d_g^{-1/2}
  \]
\[ b = -20 \ h_b + 0.2 \ \sigma_g \quad [4.2.5] \]

- **HAL** (Hall et al., 1977) which predicts \( \theta \) at \(-1500\) kPa as a quadratic function of \( P_{<2} \).
  
  Top soils: \[ PWP = 2.94 + 0.83 \ P_{<2} - 0.0054 \ (P_{<2})^2 \]
  
  Sub soils: \[ PWP = 1.48 + 0.84 \ P_{<2} - 0.0054 \ (P_{<2})^2 \quad [4.2.6] \]

- **SCH** (Scheinost et al., 1997) which predicts the van Genuchten parameters using the extended nonlinear regression from \( \phi, P_{<2}, d_g, \sigma_g \), and organic carbon. Since the data do not contain complete organic carbon measurement and considering their small content in Australian soil, the value was set to 0.
  
  \[
  \begin{align*}
  \theta_i &= 0.0051 \ P_{<2} \\
  \theta_s &= 0.85 \ \phi + 0.0013 \ P_{<2} \\
  \alpha &= 0.023 + 0.7 \ d_g \ \text{(in m)} \\
  n &= 0.33 + 2.6/\sigma_g \\
  m &= -1
  \end{align*}
  \quad [4.2.7] \]

- **ROS** (Schaap et al., 1998), which was developed using the neural-network approach. The model used six hidden units and calculation were done using the program ‘Rosetta’ available from the world-wide-web (Schaap, 1999). Two levels of input information were used:
  
  ROS4: \( \rho_b, P_{50-2000}, P_{50-2}, P_{<2} \)
  
  ROS6: \( \rho_b, P_{50-2000}, P_{50-2}, P_{<2}, \theta \) at -33 and -1500 kPa.

### 4.2.1.3. Evaluation Criteria

All methods for estimating the PTF parameters were based on minimizing the sum of squared residuals (SSR) of measured \( \theta \) and predicted \( \hat{\theta} \). Mean residual (MR) indicated whether the prediction underestimates or overestimates:

\[
\text{MR} = \frac{1}{N} \sum_{i=1}^{N} (\theta_i - \hat{\theta}_i) \quad [4.2.8]
\]

where \( N \) is the number of data points. The Root Mean Square of Residuals (RMSR) was used to calculate the mean accuracy of prediction:

\[
\text{RMSR} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\theta_i - \hat{\theta}_i)^2} .
\quad [4.2.9]
\]

Because the water content depends on water potential, relative residuals was evaluated to predict the magnitude of the residuals:

\[
\text{RR} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left( \frac{\theta_i - \hat{\theta}_i}{\theta_i} \right)^2} \times 100\% .
\quad [4.2.10]
The mean deviation (MD) and root mean squared deviation (RMSD) of Tietje and Tapkenhinrichs (1993) was used to evaluate the goodness of fit for the curve.

\[
\text{MD} = \frac{1}{NK} \sum_{i=1}^{NK} \frac{1}{b-a} \int_{b}^{a} (\theta - \hat{\theta}) d(\log_{10} h)
\]

[4.2.11]

negative values indicate that the PTFs underestimate the prediction and positive values indicate overestimation.

\[
\text{RMSD} = \frac{1}{NK} \sum_{i=1}^{NK} \left[ \frac{1}{b-a} \int_{b}^{a} (\theta - \hat{\theta})^2 d(\log_{10} h) \right]^{1/2}
\]

[4.2.12]

where \(a\) and \(b\) are the integration boundaries set at \(a = 0.1\) and \(b\) is set according the smallest (most negative) potentials measured in the sample. Calculation of MD and RMSD was done on the logarithmic of the potential and the integration is done numerically by calculating the area under the curves within the range of each measured point. Akaike’s Information Criterion/ AIC (Akaike, 1973), a statistical Ockham’s razor:

\[
\text{AIC} = N \ln \left( \sum_{i=1}^{N} (\theta_i - \hat{\theta}_i)^2 \right) + 2np
\]

[4.2.13]

was used to account for the number of parameters in the model. The best model is the one that has the smallest AIC. For single-point PTF prediction, \(N\) is the number of soil samples and for parametric PTF \(N\) is the total number of water retention data, \(np\) is the number of PTF parameters. For ANN \(np\) is the number of weights used =

\[
np = (\text{No. inputs} + 1) \times \text{No. hidden units} + (\text{No. hidden units} + 1) \times \text{No. outputs}
\]

where 1 is due to bias.

4.2.2. Results and Discussion

4.2.2.1. Parametric PTFs

Selecting the best ENR

The results of different PTFs' performance for predicting \(\theta\) for the prediction and validation data sets are shown in Table 4.2.4 and 4.2.5. Using the cross-validation prediction error (QCV) to compare the performance of the extended nonlinear regression (ENR) models for the prediction set, it was found that most of the models performed similarly except for ENR6 which does not use PWP to predict \(\theta_r\). Similarly, RMSR for all ENR models have values less than 0.04 m\(^3\) m\(^{-3}\) except for
ENR6. This indicates that prior information on the smaller potential $\theta$ is needed to obtain reasonable predictions of $\theta$, as $\theta$ also represents $\theta$ at small potentials where $d\theta/dh \approx 0$. When this information is not available ENR7 can be used but it gives an inferior estimate.

Classifying the soil texture into three classes and extending into three functions provides better prediction of $\theta$ but the improvement is not significant. Both class and continuous PTFs (ENR7 and 8) did no better than a single PTF. The RMSR for both functions are similar i.e. 0.037 m$^3$ m$^{-3}$. Tietje and Tapkenhinrichs (1993) suggested that establishment of separate PTFs for different soil types can yield good results. The results shown here do not seem to confirm this, possibly because the database here does not cover the full range of texture classes. Pachepsky et al. (1999) found that stratifying PTFs made better predictions but do not significantly improve the prediction.

Using RMSR and RMSD to compare PTFs for the validation set, it can be seen that the ENR overall gave a good performance (RMSD < 0.04 m$^3$ m$^{-3}$, accuracy ~20%) except for ENR7. By way of comparison, the PTFs of Scheinost et al. (1997) for German soil gave a RMSD of 0.035 m$^3$ m$^{-3}$ and RMSR of 0.028 m$^3$ m$^{-3}$ which predicts better than the one presented here. This may suggest that the inclusion of organic matter as a predictive variable might improve the prediction.
Table 4.2.4. Comparison of parametric PTFs performance.

<table>
<thead>
<tr>
<th>PTF</th>
<th>np</th>
<th>Prediction set (N= 6356)</th>
<th>Validation Set (N=880)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RMSR</td>
<td>AIC</td>
</tr>
<tr>
<td>ENR1</td>
<td>11</td>
<td>0.0359</td>
<td>13388.64</td>
</tr>
<tr>
<td>ENR2</td>
<td>8</td>
<td>0.0375</td>
<td>13911.83</td>
</tr>
<tr>
<td>ENR3</td>
<td>8</td>
<td>0.0375</td>
<td>13918.05</td>
</tr>
<tr>
<td>ENR4</td>
<td>8</td>
<td>0.0380</td>
<td>14081.10</td>
</tr>
<tr>
<td>ENR5</td>
<td>8</td>
<td>0.0374</td>
<td>13876.57</td>
</tr>
<tr>
<td>ENR6</td>
<td>8</td>
<td>0.0672</td>
<td>21432.69</td>
</tr>
<tr>
<td>ENR7</td>
<td>16</td>
<td>0.0471</td>
<td>16817.74</td>
</tr>
<tr>
<td>ENR8</td>
<td>16</td>
<td>0.0366</td>
<td>13611.79</td>
</tr>
<tr>
<td>ENR9</td>
<td>16</td>
<td>0.0365</td>
<td>13603.23</td>
</tr>
<tr>
<td>MRP</td>
<td>11</td>
<td>0.1102</td>
<td>27645.08</td>
</tr>
<tr>
<td>ANP</td>
<td>64</td>
<td>0.0358</td>
<td>13469.64</td>
</tr>
<tr>
<td>CAM</td>
<td>3</td>
<td>0.0660</td>
<td>21894.28</td>
</tr>
<tr>
<td>RNB</td>
<td>34</td>
<td>0.0510</td>
<td>18585.93</td>
</tr>
<tr>
<td>SCH</td>
<td>7</td>
<td>0.0530</td>
<td>19034.77</td>
</tr>
<tr>
<td>ROS4</td>
<td>52</td>
<td>0.0620</td>
<td>21187.01</td>
</tr>
<tr>
<td>ROS6</td>
<td>64</td>
<td>0.0400</td>
<td>15482.14</td>
</tr>
</tbody>
</table>

Table 4.2.5. The average of mean deviations, root mean squared deviations and relative residuals of water retention.

<table>
<thead>
<tr>
<th>PTF</th>
<th>MD (m³ m⁻³)</th>
<th>RMSD (m³ m⁻³)</th>
<th>RR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prediction</td>
<td>Validation</td>
<td>Prediction</td>
</tr>
<tr>
<td></td>
<td>NK = 733</td>
<td>NK = 108</td>
<td>NK = 733</td>
</tr>
<tr>
<td>ANP</td>
<td>0.0054</td>
<td>-0.0003</td>
<td>0.030</td>
</tr>
<tr>
<td>ENR2</td>
<td>0.0020</td>
<td>0.0136</td>
<td>0.031</td>
</tr>
<tr>
<td>ENR6</td>
<td>0.0025</td>
<td>0.0124</td>
<td>0.033</td>
</tr>
<tr>
<td>ENR8</td>
<td>0.0008</td>
<td>0.0116</td>
<td>0.031</td>
</tr>
<tr>
<td>RNB</td>
<td>0.0115</td>
<td>-0.0030</td>
<td>0.039</td>
</tr>
<tr>
<td>SCH</td>
<td>-0.0308</td>
<td>-0.0176</td>
<td>0.051</td>
</tr>
<tr>
<td>ROS4</td>
<td>0.0156</td>
<td>0.0184</td>
<td>0.054</td>
</tr>
<tr>
<td>ROS6</td>
<td>0.0071</td>
<td>0.0130</td>
<td>0.036</td>
</tr>
</tbody>
</table>
Table 4.2.6. Parametric PTFs developed in this study.

<table>
<thead>
<tr>
<th>PTF</th>
<th>$\theta_r$</th>
<th>$\theta_s$</th>
<th>$\alpha$</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>m$^3$ m$^{-3}$</td>
<td>m$^3$ m$^{-3}$</td>
<td>m$^{-1}$</td>
<td></td>
</tr>
<tr>
<td>ENR2</td>
<td>-0.00092 $P_{&lt;2}$ + 1.17748 PWP</td>
<td>0.00112 $P_{&lt;2}$ + 0.8331 $\phi$</td>
<td>1.561 + 17.046 $d_g$</td>
<td>1.3978 + 0.0027 $\sigma_g$</td>
</tr>
<tr>
<td>ENR6</td>
<td>0.00172 $P_{&lt;2}$ + 0.7 PWP</td>
<td>0.00973 $P_{&lt;2}$ + 0.842 $\phi$</td>
<td>0.147 + 1.919 $d_g$</td>
<td>1.3716 + 0.003077 $\sigma_g$ - 0.01 $\rho_b$</td>
</tr>
<tr>
<td>ENR7</td>
<td>-0.00733 + 0.00427 $P_{&lt;2}$ + 0.00267 $P_{2:20}$</td>
<td>0.00110 $P_{&lt;2}$ + 0.82607 $\phi$</td>
<td>1.361 + 16.929 $d_g$</td>
<td>1.4062 - 0.0050 $\sigma_g$</td>
</tr>
<tr>
<td>ENR8</td>
<td>$\theta_r$ = -0.00156 $P_{&lt;2}$ + 1.22333 PWP</td>
<td>$\theta_s$ = 0.00149 $P_{&lt;2}$ + 0.81851 $\phi$</td>
<td>Sandy: $\alpha$ = 0.984 + 15.607 $d_g$</td>
<td>n = 1.4990 - 0.0024 $\sigma_g$</td>
</tr>
<tr>
<td></td>
<td>Loamy: $\alpha$ = 1.787 + 4.152 $d_g$</td>
<td>n = 1.3373 + 0.0053 $\sigma_g$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Clayey: $\alpha$ = 3.186 + 181.363 $d_g$</td>
<td>n = 1.2433 + 0.00156 $\sigma_g$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Multiple regression PTFs**

Using multiple-linear regression to estimate the van Genuchten parameters (MRP method) and then estimating $\theta$, gave the worst result (RMSR = 0.11 m$^3$m$^{-3}$). It generally underestimated water content (Figure 4.2.1), which shows that this method is not satisfactory when used for parametric PTF. Usually it is difficult to find the relationship between soil properties and the scaling and shape parameters $\alpha$ and $n$. Tietje and Tapkenhinrichs (1993) found no correlation between $\alpha$ and $P_{<2}$, $P_{2:50}$ and $P_{50:2000}$. One explanation is that when parameters of the hydraulic model are estimated from nonlinear regression some error in prediction has already been embedded, this could result in large uncertainty when correlating the parameters with predictive variables. Further the parameters of the van Genuchten equation are highly correlated, such a large parameter correlation usually is an indicator of over-parameterization (Ratkowsky, 1990). This may explain why the published PTF RNB (Rawls and Brakensiek, 1985), which estimates Brooks-Corey parameters predicts better than MRP.
Figure 4.2.1. Measured vs. predicted water content by using (a) parametric multiple regression (MRP), (b) neural networks (ANP), and (c) extended non-linear regression (ENR).

**Artificial neural networks**

For the prediction set, neural networks’ (ANP) performance is comparable with that of ENR2 (RMSR = 0.036 m³m⁻³, RR = 14%). By using only bulk density and particle-size distribution data, neural networks are able to find a relationship between basic soil properties and the van Genuchten parameters. However, for the validation data set it gave poorer prediction than ENR.
Figure 4.2.2. Measured and PTFs predicted water retention curve for four different type of soil.

According to Schaap and Bouten (1996), neural networks are useful in developing PTFs because they do not require an a priori model. In our case the neural networks approach did not improve the accuracy of \( \theta \) prediction compared with ENR. This result was also recognised by Tamari and Wösten (1999). Their study showed that neural networks were less accurate than MLR when a small database was used (< 1000 samples), but can be useful if a large soil database with good quality data is available.
Based on the AIC, the best PTFs are ENR2 and ENR7 (continuous PTF). Considering the number of parameters used and the fact that the data set does not cover the full texture range, the single PTF (ENR2) is considered as the most adequate PTF. Comparison of MRP, ENR and ANN on $\theta$ prediction can be seen in Figure 4.2.1 and 4.2.2.

Comparing published PTFs

Published PTFs of RNB gave a good prediction with RMSD 0.04 m$^3$m$^{-3}$. Tietje and Tapkenhinrichs (1993) found that RNB gave an average RMSD 0.075 m$^3$m$^{-3}$ when applied to German soil. Meanwhile, Campbell’s PTF displays the least ability with RMSR = 0.07 m$^3$m$^{-3}$. PTFs of SCH showed similar predictability as RNB. The neural-networks ROS4 that only use bulk density and particle size data performed badly but when inclusion of FC and PWP as input variables (ROS6) the performance prevails over other published PTFs and is comparable with the PTFs developed in this study. It is not surprising that the more relevant inputs are fed into the PTF, the better the prediction is. As deduced by Bastet et al. (1997), the general applicability of published PTFs should not be assumed without question.

4.2.2.2 Point Estimation PTF

Comparison between different PTFs developed to estimate $\theta$ at -10, -33 and -1500 kPa and AWC are given in Table 4.2.7. ANN gave the least RMSR compared with other methods, followed by MLR and ENR except for $\theta_{-1500}$ and AWC. MLR and ANN gave better predictions for the point estimation. van den Berg et al. (1997) also reported similar results. ENR gave better prediction at $\theta_{-1500}$ because PWP was included as a variable in estimating $\theta$.

$$\theta_{-10} = 0.1728 + 0.64389 \phi - 0.24146 \phi^2 + 0.001735 P_{20-20} - 0.00003 (P_{20-20})^2 - 0.00102 P_{20-2000} - 0.00001 (P_{20-2000})^2 - 0.14684 d_g$$

[4.2.14]

$$\theta_{-33} = -0.39350 + 0.29300 \phi + 0.00855 P_{<2} - 0.00003 (P_{<2})^2 + 0.00754 P_{20-20} - 0.00003 (P_{20-20})^2 + 0.00476 P_{20-2000} - 0.00001 (P_{20-2000})^2 - 0.05326 d_g$$

[4.2.15]

$$\theta_{-1500} = -2.41449 + 0.49579 \phi - 0.53838 \phi^2 + 0.03059 P_{<2} - 0.00005 (P_{<2})^2 + 0.02506 P_{20-20} + 0.02151 P_{20-2000} + 0.00001 (P_{20-2000})^2 + 0.01217 d_g + 0.0016727 \sigma_g$$

[4.2.16]
\[ AWC = 1.97571 - 0.19669 \phi + 0.52921 \phi^2 - 0.02170 P_{<2} + 0.00003 (P_{<2})^2 - 0.0172 P_{2-20} - 0.00002 (P_{2-20})^2 - 0.01639 P_{20-2000} - 0.0000252 (P_{20-2000})^2 - 0.07054 d_g - 0.00131 \sigma_g \]

Table 4.2.7. Comparison of PTFs performance in predicting \( \theta \) at -10, -33, -1500 kPa and AWC on the validation set

<table>
<thead>
<tr>
<th>PTF</th>
<th>-10 kPa RMSR</th>
<th>AIC</th>
<th>-33 kPa RMSR</th>
<th>AIC</th>
<th>-1500 kPa RMSR</th>
<th>AIC</th>
<th>AWC RMSR</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENR2</td>
<td>0.096</td>
<td>-19.59</td>
<td>0.091</td>
<td>-17.12</td>
<td>0.023</td>
<td>-277.44</td>
<td>0.095</td>
<td>116.23</td>
</tr>
<tr>
<td>MLR</td>
<td>0.091</td>
<td>-31.62</td>
<td>0.090</td>
<td>-54.23</td>
<td>0.051</td>
<td>-52.23</td>
<td>0.115</td>
<td>0.08</td>
</tr>
<tr>
<td>ANN</td>
<td>0.072</td>
<td>30.79</td>
<td>0.072</td>
<td>30.79</td>
<td>0.072</td>
<td>30.79</td>
<td>0.062</td>
<td>-0.90</td>
</tr>
<tr>
<td>ANP</td>
<td>0.075</td>
<td>74.62</td>
<td>0.065</td>
<td>43.19</td>
<td>0.047</td>
<td>-25.69</td>
<td>0.084</td>
<td>99.52</td>
</tr>
<tr>
<td>ROS4</td>
<td>0.079</td>
<td>61.71</td>
<td>0.100</td>
<td>115.23</td>
<td>0.096</td>
<td>104.13</td>
<td>0.061</td>
<td>4.68</td>
</tr>
<tr>
<td>HALL</td>
<td>0.063</td>
<td>-85.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2.8. Comparison of MLR and ANN performance on the prediction set

<table>
<thead>
<tr>
<th>PTF</th>
<th>-10 kPa RMSR</th>
<th>AIC</th>
<th>-33 kPa RMSR</th>
<th>AIC</th>
<th>-1500 kPa RMSR</th>
<th>AIC</th>
<th>AWC RMSR</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLR</td>
<td>0.124</td>
<td>71.72</td>
<td>0.129</td>
<td>82.15</td>
<td>0.106</td>
<td>42.22</td>
<td>0.103</td>
<td>38.88</td>
</tr>
<tr>
<td>ANN</td>
<td>0.123</td>
<td>146.95</td>
<td>0.126</td>
<td>152.29</td>
<td>0.105</td>
<td>112.05</td>
<td>0.101</td>
<td>104.74</td>
</tr>
<tr>
<td>ROS4</td>
<td>0.130</td>
<td>170.96</td>
<td>0.176</td>
<td>236.64</td>
<td>0.221</td>
<td>286.02</td>
<td>0.133</td>
<td>175.50</td>
</tr>
</tbody>
</table>

The PTF of Hall et al. (1977) gave reasonable estimates of \( \theta_{-1500} \) (RMSR = 0.063 m\(^3\)m\(^{-3}\)) although it was developed for soil in England and Wales. It even predicts better than ANN. HALL predicts \( \theta_{-1500} \) as a quadratic function of \( P_{<2} \). The strong relationship with \( P_{<2} \) has been previously recognised (e.g. Nielsen and Shaw, 1958). Hall et al. (1977) explained that as pore size becomes finer, potential decreases and the influence of clay content on \( \theta \) increases. van den Berg et al. (1997) also noted that PTFs developed for tropical soils showed a consistent relationship for PWP and \( P_{<2} \) content, but not for \( \theta_{33} \) and AWC. It can be seen that \( \theta \) at -10, -33 and -1500 kPa had an exponentially increasing trend with clay (\( P_{<2} \)) and linearly decreasing trend with \( P_{20-2000} \) (Figure 4.6). Exponential function can be fitted to the data:

\[
\begin{align*}
\theta_{-10} &= 0.3064 \left(1 - \exp(-0.0406 P_{<2})\right) + 0.1452 \\
\theta_{-33} &= 0.3543 \left(1 - \exp(-0.0385 P_{<2})\right) + 0.0830 \\
\theta_{-1500} &= 0.4016 \left(1 - \exp(-0.0230 P_{<2})\right) + 0.0027
\end{align*}
\]
Figure 4.2.3. Relationship between clay and water content at –33 and –1500 kPa (b) relationship between sand and water content at –33 and –1500 kPa.

The scatter in points around the regressions in Figures 4.2.3 is due no doubt to variations in structure, organic carbon and measurement errors. Consequently, by taking the difference between the function of $\theta_{-33}$ and $\theta_{-1500}$, we can derive the relationship for AWC as a function of $P_{<2}$. From Figure 4.2.4 it can be seen that $P_{<2}$ has little effect on available water content, but there’s a trend that AWC increases with increasing $P_{<2}$ content until about 25% before it slowly decreases. This is because at low $P_{<2}$ content a relatively small amount of water is held at lower
potentials and higher amount at high potentials. But as $P_{<2}$ increases, more water is held at a lower potential as explained earlier resulting in less available water.

Table 4.2.8 shows the comparison between MLR and ANN for point estimation on the prediction set. ANN’s performance is only slight better than MLR. When taking into account the number of parameters used ANN always gave a higher AIC value because of the high number of parameters.

![Figure 4.2.4. Relationship between clay and available water content.](image)

### 4.2.2.3. Application to other dataset

To test the whether the PTFs developed can be applied to data from other areas, public-domain water retention data were used. The data sets are GRIZZLY and UNSODA, which is a compilation of numerous mostly laboratory water retention from US, Europe and other countries. The PTFs used are ENR2 developed in this paper and the neural-network ROS4 and ROS6. This neural-network PTF was found to be better than other published PTFs (Imam et al., 1999). To make fairer comparison the PWP for use in ENR2 is predicted from PTF (Equation 4.2.16).

The MR and RMSR for both data sets are given in Table 4.2.9. As can be seen the PTFs performed similarly for both data sets following the same trend as before ROS4 < ENR2 < ROS6. For the UNSODA data set ROS4 perform slightly better than ENR2, as the UNSODA were used to develop the PTFs (Schaap et al., 1998). ENR2 predicts water retention with the accuracy of 19%.
Table 4.2.9. Mean residuals (MR) and root mean squared residuals (RMSR) for water retention PTFs on the GRIZZLY and UNSODA database.

<table>
<thead>
<tr>
<th>Data set</th>
<th>ENR2</th>
<th>ROS4</th>
<th>ROS6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MR</td>
<td>RMSR</td>
<td>MR</td>
</tr>
<tr>
<td>GRIZZLY</td>
<td>Min</td>
<td>-0.287</td>
<td>0.008</td>
</tr>
<tr>
<td>N = 660</td>
<td>Max</td>
<td>0.058</td>
<td>0.102</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>-0.053</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>RR</td>
<td>19%</td>
<td>20%</td>
</tr>
<tr>
<td>UNSODA</td>
<td>Min</td>
<td>-0.235</td>
<td>0.003</td>
</tr>
<tr>
<td>N = 533</td>
<td>Max</td>
<td>0.168</td>
<td>0.188</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>-0.024</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>RR</td>
<td>18%</td>
<td>18%</td>
</tr>
</tbody>
</table>

When applying the prediction to the soil-water limits for plants, parametric ENR and point estimation MLR predict as well (or as badly) as each other (Table 4.2.10). The prediction is slightly better than ROS4. As seen in Figure 4.2.5 the drainage upper limit DUL is predicted quite well by ENR, slight overestimation is observed in the lower-limit water content LOL. Good correlation is achieved when predicting individual points but when applying the difference (Figure 4.2.5c) the prediction is quite bad and underestimates potential extractable water PLEXW. When taking the difference between two predictions, error propagation is bound to happen.

Theoretically the variance of error in prediction of PLEXW is:

$$\sigma_{\text{PLEXW}}^2 = \sigma_{\text{DUL}}^2 + \sigma_{\text{LOL}}^2 - 2 \text{Cov}(\varepsilon_{\text{DUL}}, \varepsilon_{\text{LOL}}). \quad [4.2.19]$$

The variance of the error does not only depend on the error in DUL and LOL but also the covariance between the two variables. The standard deviation of error in DUL ($\sigma_{\varepsilon_{\text{DUL}}}$) = 0.05, $\sigma_{\varepsilon_{\text{LOL}}}$ = 0.06 and correlation between the two errors is 0.76, resulting in the standard deviation error of 0.04 m$^3$ m$^{-3}$. The correlation between DUL and LOL is also high (0.89), using ENR2 the correlation between predicted DUL and LOL becomes much larger (0.98).

Table 4.2.10. Mean residuals and root mean squared residuals for soil-water limits.

<table>
<thead>
<tr>
<th>DUL</th>
<th>LOL</th>
<th>PLEXW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ENR2</td>
<td>ROS4</td>
</tr>
<tr>
<td>MR</td>
<td>0.017</td>
<td>0.034</td>
</tr>
<tr>
<td>RMSR</td>
<td>0.056</td>
<td>0.070</td>
</tr>
</tbody>
</table>
Figure 4.2.5. Soil-plant water limits as predicted by PTFs (a) drainage upper limit, (b) lower limit, and (c) plant extractable water.

4.2.2.4. Some afterthoughts

So far it has been shown that ENR predicts water retention quite well compared with other methods. But when relating predicted parameters $\alpha$ and $n$ with calculated values, ENR does not seem to produce a good estimate (Figure 4.2.6). It can be argued that the predicted water content depends on the conjunction of all the parameters. Wösten and van Genuchten (1988) showed the sensitivity of the water-retention curve by varying one parameter and holding the others constant. In Figure 4.2.7 it can be shown that varying more than two parameters simultaneously could well produce a similar curve which fit the data. The parameters should be interpreted empirically rather than physically.
Figure 4.2.6. Observed and ENR predicted shape factor $\alpha$ and $n$ of the van Genuchten parameters.

The distribution of the van Genuchten parameters was evaluated for different PTFs (Table 4.2.11). For residual water content $\theta_r$, ENR2 and ENR8 seem to overestimate its value, as evidenced by the coefficient 1.17 applied in the equation. ANP, ENR6, ROS4 and ROS6 gave reasonable ranges. A limitation of ENR2 is that at very low clay content, $\theta_r$ could be negative. If this is encountered, the value is set to 0. For saturated water content, similar prediction range was found by all the methods because it mostly depends on porosity. The ENR methods produce a larger range for the scaling parameter $\alpha$ compared with ROS4 and ROS6. ANP gave an $\alpha$ range which is too high.
Chapter IV – Hydraulic pedotransfer functions for Australian soil

Figure 4.2.7. Varying the van Genuchten parameters may produce similar fit to water retention data.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>ANP</th>
<th>ENR2</th>
<th>ENR6</th>
<th>ENR8</th>
<th>ROS4</th>
<th>ROS6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_r$</td>
<td>Median</td>
<td>0.060</td>
<td>0.176</td>
<td>0.159</td>
<td>0.163</td>
<td>0.066</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.000</td>
<td>0.018</td>
<td>0.017</td>
<td>0.016</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.122</td>
<td>0.400</td>
<td>0.367</td>
<td>0.391</td>
<td>0.122</td>
</tr>
<tr>
<td>80% range*</td>
<td>0.080</td>
<td>0.253</td>
<td>0.236</td>
<td>0.240</td>
<td>0.061</td>
<td>0.089</td>
</tr>
<tr>
<td>$\theta_s$</td>
<td>Median</td>
<td>0.399</td>
<td>0.397</td>
<td>0.394</td>
<td>0.399</td>
<td>0.400</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.243</td>
<td>0.231</td>
<td>0.232</td>
<td>0.232</td>
<td>0.275</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.655</td>
<td>0.691</td>
<td>0.693</td>
<td>0.694</td>
<td>0.741</td>
</tr>
<tr>
<td>80% range*</td>
<td>0.176</td>
<td>0.178</td>
<td>0.167</td>
<td>0.188</td>
<td>0.176</td>
<td>0.191</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Median</td>
<td>2.61</td>
<td>2.09</td>
<td>2.10</td>
<td>2.81</td>
<td>1.68</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.44</td>
<td>1.61</td>
<td>1.50</td>
<td>1.47</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>419.99</td>
<td>19.66</td>
<td>14.90</td>
<td>177.68</td>
<td>5.45</td>
</tr>
<tr>
<td>80% range*</td>
<td>7.44</td>
<td>3.47</td>
<td>2.50</td>
<td>3.95</td>
<td>2.13</td>
<td>4.56</td>
</tr>
<tr>
<td>$n$</td>
<td>Median</td>
<td>1.21</td>
<td>1.42</td>
<td>1.39</td>
<td>1.40</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>1.02</td>
<td>1.41</td>
<td>1.36</td>
<td>1.25</td>
<td>1.16</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>1.75</td>
<td>1.47</td>
<td>1.44</td>
<td>1.49</td>
<td>3.22</td>
</tr>
<tr>
<td>80% range*</td>
<td>0.32</td>
<td>0.03</td>
<td>0.03</td>
<td>0.22</td>
<td>0.30</td>
<td>0.40</td>
</tr>
</tbody>
</table>

*80% prediction limits around the median, calculated from 90 percentiles minus 10 percentiles.
Table 4.2.12. Distribution of the van Genuchten parameters predicted by ENR2.

<table>
<thead>
<tr>
<th>Database</th>
<th>$\theta_c$</th>
<th>$\theta_s$</th>
<th>$\alpha$</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRIZZLY</td>
<td>Median</td>
<td>0.096</td>
<td>0.386</td>
<td>3.46</td>
</tr>
<tr>
<td>NK = 660</td>
<td>Min</td>
<td>0.000</td>
<td>0.210</td>
<td>1.59</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.334</td>
<td>0.687</td>
<td>18.78</td>
</tr>
<tr>
<td></td>
<td>80% range</td>
<td>0.293</td>
<td>0.194</td>
<td>14.79</td>
</tr>
<tr>
<td>UNSODA</td>
<td>Median</td>
<td>0.096</td>
<td>0.380</td>
<td>3.18</td>
</tr>
<tr>
<td>NK = 533</td>
<td>Min</td>
<td>0.000</td>
<td>0.220</td>
<td>1.63</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.331</td>
<td>0.789</td>
<td>17.54</td>
</tr>
<tr>
<td></td>
<td>80% range</td>
<td>0.202</td>
<td>0.178</td>
<td>11.19</td>
</tr>
</tbody>
</table>

For shape parameter $n$, the ENR gave a very narrow range around 1.4. ENR tends to scale the shape factor around a constant value. This is also observed in other data sets (Table 4.2.12). As can be seen in Figure 4.2.8b, the scaled water content ($S_c$) and potential ($\alpha h$) using ENR produce a mean curve. The PTFs significantly reduce the variations and bring them to the mean. This is because in ENR we are essentially fitting one equation to the whole data set. Larger variation is observed when using the neural-networks PTF Figure 4.2.8c and d. The predictability of ENR is quite similar for different textures (Figure 4.2.9) and also comparable with ROS6. This implies that for large-scale prediction the ENR can be used quite well as demonstrated in the large data sets derived from different locations and countries. If it was going to be used for fine-scale spatial prediction, it may pose a problem, because it reduces the variation.
Figure 4.2.8. Scaled water content and water potential for (a) original data (b) predicted by ENR (c) ANP and, (d) ROS6.
4.2.3. Conclusions

Point estimation and parametric pedotransfer functions have been developed for Australian soil. Because of the unique soil properties, different cutoffs used in particle-size classification, and unavailability of data on some soil properties, most PTFs developed elsewhere cannot be applied directly. Using particle-size distribution and bulk density data, the PTFs developed predict water content at different potentials with reasonable accuracy. Difficulties have been encountered when relating the scaling and shape parameters of the van Genuchten equation to soil properties, because of over-parameterization. This can be countered by using extended nonlinear regression as discussed in this paper.

Parametric estimation using ENR predicts water content better than ANN and MRP. This is because ENR is developed by fitting simultaneously all the soil water retention data as opposed to the ANN and MRP, which predict the parameters obtained from individual water-retention curve fits. The performance of neural networks is however comparable with that of extended nonlinear regression. Since the latter is considered as a black-box approach, the former method with a more physical basis is preferred at the current time.
4.3. Saturated hydraulic conductivity PTFs

Saturated hydraulic conductivity ($K_s$) is an important physical factor determining water transport in soil. Measurement of $K_s$ can be done in the field or laboratory. In the field, the disk permeameter has been used extensively for measuring saturated and near-saturated hydraulic conductivity ($K$). However, $K_s$ is rarely measured in routine soil survey because it is time-consuming and can be expensive. This study aim to evaluate the performance of published PTFs on the collected published data set of $K_s$ for Australian soil. Since the data set is derived from different studies and different areas, a large variation is expected. PTFs are also developed using different techniques and their predicability and implications are discussed.

4.3.1. Theory

The fractal dimension ($D$) of particle size-distribution (PSD) has been related to soil hydraulic properties (Rieu and Perrier, 1998). This includes estimation of the soil water retention curve and hydraulic conductivity (Rieu and Sposito, 1991). Tyler and Wheatcraft (1989) applied the fractal power law to estimate the fractal dimension $D_1$ of the soil PSD as:

\[ N_g = cr^{-D_1} \]  \[ 4.3.1 \]

where $c$ is an empirical constant, $r$ is the arithmetic mean of the radius between two successive particle sizes, and $N_g$ is the number of spherical particles calculated by :

\[ N_g = \frac{3w}{4\pi\rho_s r^3} \]  \[ 4.3.2 \]

$w$ is the fraction mass of particles within the size range and $\rho_s$ is the particle density. The application of Equation 4.3.1 has proven to give a good fit to data but the value estimated often exceeds 3, which is physically impossible (Rieu and Perrier, 1998). Tyler and Wheatcraft (1992) and Chang and Uehara (1994) then proposed estimating the fractal dimension $D_2$ by assuming a power function:

\[ W = cr^{3-D_2} \]  \[ 4.3.3 \]

where $W$ is the mass percentage of particles whose size is less than $r$. By taking the logarithm of the above equation:

\[ \ln W = \ln c + (3-D_2) \ln r \]  \[ 4.3.4 \]
$D_2$ is estimated from the slope of $\ln W$ and $\ln r$. Since the lower and upper limit used in particle-size analysis is $<2 \ \mu m$ and $2000 \ \mu m$, it allows determination of $D_2$ from mass percentage of clay (Bui et al. 1996):

$$D_2 = 0.1448 \ln (P_{<2}) + 2.333 \quad [4.3.5]$$

This relationship provides an estimate of fractal dimension in the range of $0 < D < 3$.

Kravchenko and Zhang (1998) proposed the relationship between $W$ and $r$ to estimate the fractal dimension $D_3$:

$$\ln W = \ln c + \left( \frac{3D_3^2 - 13D_3 + 14}{D_3^2 - 5D_3 + 4} + 1 \right) \ln r \quad [4.3.6]$$

To determine $D_3$ from the slope of PSD a nonlinear root-finding algorithm is needed. From equations 4.3.5 and 4.3.6 we see that $D_2$ has a curvilinear relationship with $D_3$. The relation between fractal dimension $D_1$ and $D_2$ is shown in Fig. 4.3.1. The differences illustrate the disagreement when estimating the number of particles. As noted by Rieu and Perrier (1998) the various estimates of fractal dimension using the above approaches are approximate and a number of assumptions must be considered. Therefore, the estimation should be interpreted as empirical.

Figure 4.3.1. Relationship between fractal dimension $D_1$ and $D_2$. 
4.3.2. Methods

4.3.2.1. Data

Previously published saturated hydraulic conductivity data sets from across Australia compiled (see § 4.1.2). This compilation, which will be called the ‘combined data set’, covers large areas with different soil types and different methods and sample volumes in the measurement of $K_s$. A subset of data from Bristow et al. (1999) was randomly selected and used as a validation set. A summary of the data used is given in Table 4.3.1.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Location (state)</th>
<th>All data</th>
<th>Prediction set</th>
<th>Validation set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forrest</td>
<td>NSW, Qld, Vic &amp; SA</td>
<td>60  118</td>
<td>60  118</td>
<td></td>
</tr>
<tr>
<td>Bristow</td>
<td>Qld</td>
<td>41  369</td>
<td>19  170</td>
<td>22  199</td>
</tr>
<tr>
<td>Bridge</td>
<td>NSW</td>
<td>59  104</td>
<td>59  104</td>
<td></td>
</tr>
<tr>
<td>USyd</td>
<td>NSW</td>
<td>11  60</td>
<td>11  60</td>
<td></td>
</tr>
<tr>
<td>Olsson and Rose</td>
<td>Qld</td>
<td>1  5</td>
<td>1  5</td>
<td></td>
</tr>
<tr>
<td>Prebble</td>
<td>Qld</td>
<td>1  5</td>
<td>1  5</td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td>173  661</td>
<td>151  462</td>
<td>22  199</td>
</tr>
</tbody>
</table>

The soil physical properties common to all data sets were used as the predictive variables to estimate hydraulic conductivity. The properties used were:

- bulk density ($\rho_b$) in Mg m$^{-3}$;
- porosity ($\phi$) in m$^3$ m$^{-3}$;
- the mass of particles <2 µm (P$_{<2}$), 2-20 µm (P$_{2-20}$) and 20-2000 µm (P$_{20-2000}$), which were normalized to sum to 100 dag kg$^{-1}$;
- geometric mean particle-size diameter ($d_g$ in mm) and geometric standard deviation ($\sigma_g$ in mm) calculated according to Shirazi and Boersma (1984);
- Particle-size distribution fractal dimensions: $D_1$ (Tyler and Wheatcraft 1990), $D_2$ (Chang and Uehara 1994), and $D_3$ (Kravchenko and Zhang 1998);
- effective porosity ($\phi_e$) calculated as:
  \[
  \phi_e = \phi - \theta \text{ at } -10 \text{ kPa}.
  \]

A summary of the soil properties is given in Table 4.3.2.
Table 4.3.2. Summary of physical properties of the soil used in this study

<table>
<thead>
<tr>
<th></th>
<th>Prediction (N = 462)</th>
<th>Validation (N = 199)</th>
<th>prob t*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min.</td>
<td>Max.</td>
<td>μ</td>
</tr>
<tr>
<td>ρ_b  Mg m^{-3}</td>
<td>0.57</td>
<td>1.80</td>
<td>1.45</td>
</tr>
<tr>
<td>ϕ    m^3 m^{-3}</td>
<td>0.32</td>
<td>0.78</td>
<td>0.45</td>
</tr>
<tr>
<td>P_{&lt;2} dag kg^{-1}</td>
<td>5.00</td>
<td>75.00</td>
<td>31.00</td>
</tr>
<tr>
<td>P_{20-2000} dag kg^{-1}</td>
<td>1.00</td>
<td>53.00</td>
<td>16.64</td>
</tr>
<tr>
<td>ln d_g ln (mm)</td>
<td>-5.74</td>
<td>0.06</td>
<td>-3.46</td>
</tr>
<tr>
<td>ln σ_g ln (mm)</td>
<td>1.52</td>
<td>3.00</td>
<td>2.36</td>
</tr>
<tr>
<td>D_1</td>
<td>2.63</td>
<td>3.56</td>
<td>3.04</td>
</tr>
<tr>
<td>D_2</td>
<td>2.51</td>
<td>2.96</td>
<td>2.80</td>
</tr>
<tr>
<td>D_3</td>
<td>2.79</td>
<td>2.99</td>
<td>2.93</td>
</tr>
<tr>
<td>ϕ_e m^3/m^3</td>
<td>0.00</td>
<td>0.34</td>
<td>0.12</td>
</tr>
<tr>
<td>ln K_s ln (mm/h)</td>
<td>-5.17</td>
<td>9.53</td>
<td>2.24</td>
</tr>
</tbody>
</table>

* probability associated with t test for two samples come from distributions with the same means.  
* samples significantly come from distributions with the same means.

4.3.2.2. Pedotransfer function evaluation

4.3.2.2.1. Published PTFs

Due to the difference in classification of particle-size fractions in Australia and other countries, conversion of the silt fraction from ISSS (2-20 µm) to FAO/USDA (2-50 µm) was done using equation [4.1.1].

The following published PTFs were used to predict $K_s$ (in mm/h):

- **Campbell (1985):**
  \[ K_s = 141.25 d_g \left( \frac{1.3}{\rho} \right)^{1.3b} \]  
  where $b = d_g^{-0.5} + 0.2 \sigma_g$.

- **Campbell and Shiozawa (1994):** subsequently suggested a simpler equation:
  \[ K_s = 54 \exp (-0.07 P_{50-2000} - 0.167 P_{<2}) . \]  

- **Cosby et al. (1984):**
  \[ K_s = 25.4 \cdot 10^{(-0.6+0.012P_{<2000}-0.0064P_{<2})} \]  
  \[ R^2 = 0.872, N = 1448 \]  

- **Puckett et al. (1984)** first developed:
  \[ K_s = 156.96 \exp[-0.1975 P_{<2}] \]  
  \[ R^2 = 0.77, N = 42, \]  
  using a larger data sets, Dane and Puckett (1994) established the following:
$K_s = 303.84 \exp[-0.144 \text{ P<2}] \ (R^2 = 0.453, N = 577). \ [4.3.11]$

- **Brakensiek et al. (1984):**

$$K_s = 10 \exp[19.52348 \: \phi - 8.96847 - 0.028212 \: \text{ P<2} + 0.00018107 \: (\text{ P}_{50-2000})^2 - 0.0094125 \: (\text{ P<2})^2 - 3.895215 \: \phi^2 + 0.077718 \: \text{ P}_{50-2000} \: \phi - 0.00298 \: (\text{ P}_{50-2000})^2 - 0.019492 \: (\text{ P<2})^2 \: \phi + 0.001434 \: (\text{ P}_{50-2000})^2 \: \phi - 0.0000035 \: (\text{ P<2}) \: \text{ P}_{50-2000}] \quad [4.3.12]$$

- **Saxton et al. (1986):**

$$K_s = 10 \exp[12.012 - 0.0755 \: \text{ P}_{50-2000} + (-3.895 + 0.03671 \: \text{ P}_{50-2000} - 0.1103 \: \text{ P<2} + 0.00087546 \: (\text{ P<2})^2)/ \: \theta_s]$$

$$(R^2 = 0.95, N = 230) \ [4.3.13]$$

where $\theta_s = 0.332 - 0.0007251 \: \text{ P}_{50-2000} + 0.1276 \: \log_{10}(\text{ P<2}).$

- **Schaap et al. (1998):**

The PTFs used the neural-network approach. The model used six hidden units and calculation were done using the program ‘Rosetta’ available from the world wide web (Schaap, 1999). Two levels of input information were used:

- **K07:** $\phi_{b}, \text{ P}_{50-2000}, \text{ P}_{50-2}, \text{ P<2}$ $(R^2 = 0.57, N = 620)$
- **K08:** $\phi_{b}, \text{ P}_{50-2000}, \text{ P}_{50-2}, \text{ P<2}, \theta$ at -33 and -1500 kPa $(R^2 = 0.68, N = 620)$

### 4.3.2.2.2. Development of PTFs

Following Ahuja *et al.* (1984) the empirical relationship:

$$K_s = B \: \phi_e^m \ [4.3.14]$$

was applied to the data set here. The parameters $B$ and $m$ are estimated using linear regression on ln $K_s$ versus ln $\phi_e$. A continuous PTF was also developed by using the membership of the textural class (see § 4.1.5). The model is:

$$K_s = u_{\text{sandy}} B_{\text{sandy}} \: \phi_e^m + u_{\text{loamy}} B_{\text{loamy}} \: \phi_e^m + u_{\text{clayey}} B_{\text{clayey}} \: \phi_e^m \ [4.3.15]$$

A multiple regression was used in order to relate basic soil properties to $K_s$:

$$\ln K_s = p_1 + p_2 \: \text{ P<2} + p_3 \: \text{ P}_{20-2000} + p_4 \: \ln d_e + p_5 \: \ln \phi_e + p_6 \: D \ [4.3.16]$$

The three different types of D ($D_1$, $D_2$ and $D_3$) were used to evaluate the predictive ability of each estimate. In soil surveys particle-size analysis is routinely measured in 3-4 size classes, therefore by plotting the log-log relationship between cumulative percentage of material and particle size will give an approximate estimate of $D$. 
The fractal model of Rawls et al. (1993) was also applied here. As we did not have the measurement of $h_b$ in the soil survey data and based on the water retention for the Van Genuchten (1980) parameter $\alpha \approx 1/h_b$, an empirical relationship is used:

$$1 / R_1 = a + b \, d_g$$

An empirical relationship was developed to derive the expression for $K_s$:

$$K_s = 4.41 \times 10^4 \left( \frac{\phi^{4/3}}{p_1 D_2 e_2} \right)^2 \left( \frac{1}{p_3 + p_4 d_g} \right)^2$$

[4.3.17]

The unknown parameters were estimated by fitting the above relationship to the measured values using the nonlinear least-squares method.

<table>
<thead>
<tr>
<th>Model</th>
<th>Equation</th>
<th>Training data</th>
<th>References/Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previously published</td>
<td>4.3.8</td>
<td>Various published data</td>
<td>Campbell (1985)</td>
</tr>
<tr>
<td>K02</td>
<td>4.3.9</td>
<td>Data of Puckett et al. (1985)</td>
<td>Campbell and Shiozawa (1994)</td>
</tr>
<tr>
<td>K03</td>
<td>4.3.10</td>
<td>Holtan et al. (1968), Rawls et al. (1982)</td>
<td>Cosby et al. (1984)</td>
</tr>
<tr>
<td>K04</td>
<td>4.3.11</td>
<td>Ultisols from Lower Coastal Plain of Alabama $N = 577$</td>
<td>Dane and Puckett (1994)</td>
</tr>
<tr>
<td>K05</td>
<td>4.3.12</td>
<td>Data from US</td>
<td>Brakensiek et al. (1984)</td>
</tr>
<tr>
<td>K06</td>
<td>4.3.13</td>
<td>Rawls et al. (1982) $N = 230$</td>
<td>Saxton et al. (1986)</td>
</tr>
<tr>
<td>K07</td>
<td></td>
<td>Data from US $N = 620$</td>
<td>Schaap et al. (1998)</td>
</tr>
<tr>
<td>K08</td>
<td></td>
<td>Data from US $N = 620$</td>
<td>Schaap et al. (1998)</td>
</tr>
<tr>
<td>Developed de novo</td>
<td>4.3.14</td>
<td>Current data set $N = 462$</td>
<td>Power function of $\phi_e$</td>
</tr>
<tr>
<td>K10</td>
<td>4.3.15</td>
<td>Current data set $N = 462$</td>
<td>Continuous function of $\phi_e$</td>
</tr>
<tr>
<td>K11</td>
<td>4.3.16</td>
<td>Current data set $N = 462$</td>
<td>Multiple Linear Regression</td>
</tr>
<tr>
<td>K12</td>
<td>4.3.17</td>
<td>Current data set $N = 462$</td>
<td>Fractal model using $D_2$</td>
</tr>
<tr>
<td>K13</td>
<td></td>
<td>Current data set $N = 462$</td>
<td>ANN 7 input</td>
</tr>
<tr>
<td>K14</td>
<td></td>
<td>Current data set $N = 462$</td>
<td>ANN 8 input</td>
</tr>
</tbody>
</table>

Artificial neural networks were also used to develop PTFs. The multilayer perceptron type of ANN (Demuth and Beale, 1998) was used in this study. Two input types were considered to predict the logarithmically transformed $K_s$ (ln $K_s$):

- K13: 7 input variables $P_{<2}$, $P_{2-20}$, $P_{20-2000}$, $\rho_b$, $\ln(d_g)$, $\sigma_g$, $D$
- K14: 8 input variables $P_{<2}$, $P_{2-20}$, $P_{20-2000}$, $\rho_b$, $\ln(d_g)$, $\sigma_g$, $D$, $\phi_e$

It was found that the input and output variables for the training set had to be normalized in order to produce reliable results. The following normalization was used
for each variable \( v \): \( (v - \bar{v})/\sigma_v \), where \( \bar{v} \) and \( \sigma_v \) are the mean and standard deviation of the variable. From numerical experimentation, 5 hidden units with the hyperbolic tangent activation function for the hidden layer and linear activation function for the output layer was found to provide a good estimate. The calculations were carried out using program MATLAB ver. 5.3 with the Neural Network Toolbox ver 3.0 (Demuth and Beale, 1998). The PTFs used are labelled K01 - K14 and the description of them is given in Table 4.3.3.

4.3.2.2.3. Evaluation Criteria

Since \( K_s \) exhibits a long-tailed distribution, usually considered to be lognormal (Nielsen \textit{et al.}, 1973), log transformation was used in developing the PTFs. The sum of squared residuals (SSR) of measured \( \ln K_s \) and predicted \( \hat{\ln K_s} \) were the basis of the PTF development. The Root Mean Squared Residual (RMSR) was used to calculate the mean prediction error. The Geometric Root Mean Squared Residuals (GRMSR), representing the absolute differences (in mm/h), was then calculated as:

\[
\text{GRMSR} = \exp[\text{RMSR}]
\]

Akaike’s Information Criterion, AIC (Akaike, 1973), was used to account for the number of parameters in the model.

Following Tietje and Hennings (1996) the error ratio was used to describe the ratio between the predicted and measured hydraulic conductivity.

\[
\varepsilon = \frac{\hat{K}_s}{K_s}
\]

the Geometric Mean of Error ratio (GMER) was calculated as:

\[
\text{GMER} = \exp\left(\frac{\sum_{i=1}^{n} \ln \varepsilon_i}{n}\right)
\]  \[4.3.18\]

and the Geometric Standard Deviation of the Error Ratio (GSDER) is calculated as:

\[
\text{GSDER} = \exp\left(\frac{1}{n-1}\sum_{i=1}^{n} (\ln \varepsilon_i - \ln \text{GMER})^2\right)
\]  \[4.3.19\]

GMER indicates the bias, whether the prediction systematically overestimates (GMER > 1) or underestimates (GMER < 1) the measured values. GSDER indicates the deviation around the GMER. The prediction is exact only if GSDER = 1.
4.3.3. **Results and discussion**

4.3.3.1. **Published PTFs**

The results for the PTF prediction of $K_s$ are presented in Table 4.3.4. Figure 4.3.2 clearly shows that prediction of $K_s$ showed large differences. The performance of published PTFs (K01 to K08) is comparable to the PTFs developed using the current data set (K09 to K14). As shown in Fig 4.3.2, most of the PTFs underestimate $K_s$ (GMER < 1). The ANN of Schaap et al. (1998), K07 and K08, performed best on the prediction set. The PTF K08 that used basic soil properties and $\theta$ at -33 and -1500 kPa gave the lowest error ratio ($0.96 \pm 1.51$) and residuals (10.6 mm/h). Obviously when more relevant inputs are provided, the prediction will be better. But the cost of obtaining the necessary data will be larger and in routine soil survey water retention is usually not measured. Regression PTF K03 (Cosby et al., 1984) also shows good prediction (error ratio = $1.46 \pm 1.72$, residuals = 13.2 mm/h). Meanwhile PTFs K02 (Campbell and Shiozawa, 1994) and K06 (Saxton et al., 1986) exhibited the largest error ratio and residuals. K02 underestimated $K_s$ about 300 times and K06 gave a ten-fold underestimation. By way of comparison, Tietje and Hennings (1996) found that the error ratio of K01 = $0.21 \pm 10.21$, K03 = $1.42 \pm 7.86$, K05 = $0.17 \pm 11.52$, K06 = $1.19 \pm 7.72$ for Northern German soil ($N = 1067$).
### Table 4.3.4. Evaluation of fit criteria for the prediction of $K_s$ using different PTFs

<table>
<thead>
<tr>
<th>Model</th>
<th>np</th>
<th>GMER</th>
<th>GSDER</th>
<th>RMSR</th>
<th>GRMSR</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prediction ($N = 462$)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K01</td>
<td>5</td>
<td>0.240</td>
<td>2.208</td>
<td>3.28</td>
<td>26.7</td>
<td>3943.0</td>
</tr>
<tr>
<td>K02</td>
<td>3</td>
<td>0.003</td>
<td>1.644</td>
<td>6.41</td>
<td>605.0</td>
<td>4556.6</td>
</tr>
<tr>
<td>K03</td>
<td>4</td>
<td>1.463</td>
<td>1.716</td>
<td>2.58</td>
<td>13.2</td>
<td>3719.4</td>
</tr>
<tr>
<td>K04</td>
<td>2</td>
<td>0.429</td>
<td>1.723</td>
<td>2.71</td>
<td>15.1</td>
<td>3761.2</td>
</tr>
<tr>
<td>K05</td>
<td>13</td>
<td>0.211</td>
<td>1.865</td>
<td>3.08</td>
<td>21.9</td>
<td>3901.4</td>
</tr>
<tr>
<td>K06</td>
<td>9</td>
<td>0.098</td>
<td>1.757</td>
<td>3.53</td>
<td>34.0</td>
<td>4016.9</td>
</tr>
<tr>
<td>K07</td>
<td>37</td>
<td>0.75</td>
<td>1.62</td>
<td>2.50</td>
<td>12.1</td>
<td>3745.0</td>
</tr>
<tr>
<td>K08</td>
<td>49</td>
<td>0.96</td>
<td>1.51</td>
<td>2.36</td>
<td>10.6</td>
<td>3716.6</td>
</tr>
<tr>
<td>K09</td>
<td>2</td>
<td>0.682</td>
<td>2.656</td>
<td>2.66</td>
<td>14.3</td>
<td>3743.7</td>
</tr>
<tr>
<td>K10</td>
<td>4</td>
<td>0.682</td>
<td>2.656</td>
<td>2.66</td>
<td>14.3</td>
<td>3747.7</td>
</tr>
<tr>
<td>K11</td>
<td>6</td>
<td>0.716</td>
<td>2.063</td>
<td>2.34</td>
<td>10.4</td>
<td>3631.4</td>
</tr>
<tr>
<td>K12</td>
<td>4</td>
<td>1.073</td>
<td>1.586</td>
<td>2.41</td>
<td>11.1</td>
<td>3654.0</td>
</tr>
<tr>
<td>K13</td>
<td>46</td>
<td>0.123</td>
<td>1.467</td>
<td>3.03</td>
<td>20.7</td>
<td>3950.8</td>
</tr>
<tr>
<td>K14</td>
<td>51</td>
<td>0.116</td>
<td>1.457</td>
<td>3.05</td>
<td>21.1</td>
<td>3967.5</td>
</tr>
<tr>
<td><strong>Validation ($N = 199$)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K01</td>
<td>5</td>
<td>0.240</td>
<td>1.127</td>
<td>2.90</td>
<td>18.1</td>
<td>1486.9</td>
</tr>
<tr>
<td>K02</td>
<td>3</td>
<td>0.003</td>
<td>1.176</td>
<td>6.98</td>
<td>1078.3</td>
<td>1832.9</td>
</tr>
<tr>
<td>K03</td>
<td>4</td>
<td>1.463</td>
<td>1.144</td>
<td>2.25</td>
<td>9.5</td>
<td>1384.5</td>
</tr>
<tr>
<td>K04</td>
<td>2</td>
<td>0.429</td>
<td>1.214</td>
<td>3.38</td>
<td>29.2</td>
<td>1541.5</td>
</tr>
<tr>
<td>K05</td>
<td>13</td>
<td>0.211</td>
<td>1.130</td>
<td>2.79</td>
<td>16.4</td>
<td>1488.4</td>
</tr>
<tr>
<td>K06</td>
<td>9</td>
<td>0.098</td>
<td>1.218</td>
<td>4.32</td>
<td>75.6</td>
<td>1654.2</td>
</tr>
<tr>
<td>K07</td>
<td>37</td>
<td>0.239</td>
<td>1.115</td>
<td>2.09</td>
<td>8.1</td>
<td>1421.4</td>
</tr>
<tr>
<td>K08</td>
<td>49</td>
<td>0.213</td>
<td>1.115</td>
<td>2.17</td>
<td>8.8</td>
<td>1459.9</td>
</tr>
<tr>
<td>K09</td>
<td>2</td>
<td>0.682</td>
<td>1.155</td>
<td>2.09</td>
<td>8.1</td>
<td>1351.4</td>
</tr>
<tr>
<td>K10</td>
<td>4</td>
<td>0.682</td>
<td>1.171</td>
<td>2.93</td>
<td>18.7</td>
<td>1489.0</td>
</tr>
<tr>
<td>K11</td>
<td>6</td>
<td>0.716</td>
<td>1.157</td>
<td>3.13</td>
<td>23.0</td>
<td>1520.0</td>
</tr>
<tr>
<td>K12</td>
<td>4</td>
<td>1.073</td>
<td>1.157</td>
<td>2.68</td>
<td>14.6</td>
<td>1453.7</td>
</tr>
<tr>
<td>K13</td>
<td>46</td>
<td>0.123</td>
<td>1.129</td>
<td>2.39</td>
<td>11.0</td>
<td>1492.8</td>
</tr>
<tr>
<td>K14</td>
<td>51</td>
<td>0.116</td>
<td>1.129</td>
<td>3.67</td>
<td>39.4</td>
<td>1673.4</td>
</tr>
</tbody>
</table>

### Table 4.3.5. Comparison of geometric root mean squared residuals using different predicting fractal dimension $D$

<table>
<thead>
<tr>
<th></th>
<th>GRMSR (mm/h)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prediction</td>
<td>Validation</td>
</tr>
<tr>
<td><strong>MLR $D_1$</strong></td>
<td>10.32</td>
<td>13.81</td>
</tr>
<tr>
<td><strong>MLR $D_2$</strong></td>
<td>10.36</td>
<td>14.59</td>
</tr>
<tr>
<td><strong>MLR $D_3$</strong></td>
<td>10.367</td>
<td>14.64</td>
</tr>
<tr>
<td><strong>ANN $D_1$</strong></td>
<td>5.35</td>
<td>7.56</td>
</tr>
<tr>
<td><strong>ANN $D_2$</strong></td>
<td>5.33</td>
<td>6.97</td>
</tr>
<tr>
<td><strong>ANN $D_3$</strong></td>
<td>5.22</td>
<td>9.50</td>
</tr>
</tbody>
</table>
Figure 4.3.2. Performance of pedotransfer functions on prediction and validation data (a) geometric mean error ratio, (b) Geometric standard deviation of error ratio, and (c) Geometric root mean squared residuals.
4.3.3.2. Development of PTFs

The assessment of developing PTFs using different methods for estimating the fractal dimension ($D$) of particle-size distribution is shown in Table 4.3.5. When comparing the predicability of $D$ using multiple linear regression (MLR) results show that MLR performed satisfactorily in predicting $K_s$. All 3 types of $D$ estimates gave similar performance. Therefore the PTF using $D_2$, which is easier to estimate and gives a higher correlation, was used as the prediction ($K_{11}$).

$$\ln K_s = 8.158 - 0.0832 P_{<2} - 0.0326 P_{20-2000} + 0.31632 \ln d_g + 3.0065 \ln \phi_e + 1.91 D_1 \quad (R^2 = 0.5470, N = 428)$$ \[4.3.20\]

$$\ln K_s = 0.6897 - 0.1029 P_{<2} - 0.043 P_{20-2000} + 0.31434 \ln d_g + 3.0434 \ln \phi_e + 5.1773 D_2 \quad (R^2 = 0.5482, N = 428)$$ \[4.3.21\]

$$\ln K_s = -22.0856 - 0.1017 P_{<2} - 0.0425 P_{20-2000} + 0.3169 \ln d_g + 3.0503 \ln \phi_e + 12.7254 D_3 \quad (R^2 = 0.5489, N = 428)$$ \[4.3.22\]

ANN results also show no significant differences when using $D_1$, $D_2$ or $D_3$ in predicting $K_s$. Therefore $D_2$ is used as a predictive variable for the artificial neural networks.

The power function model of Ahuja et al. (1984) developed here is given by:

$$K_s = 23190.55 \phi_e^{3.66}$$ \[4.3.23\]

Fig. 4.3.3a shows the prediction of $K_s$ as a function of effective porosity. Although there is a large scatter of data around the prediction, this equation ($K_{09}$) provides a good estimate compared with the other models (Table 4.3.4). This model suggests that additional information on water retention will improve the prediction. For textural stratification, 3 functions were developed:

Sandy: $K_s = 21088.86 \phi_e^{3.34}$

Loamy: $K_s = 15744.27 \phi_e^{3.34}$

Clayey: $K_s = 4303.27 \phi_e^{3.34}$ \[4.3.24\]

As shown in Fig. 4.3.3b the model separates the clayey soil from the sandy and loamy classes but the predictions for sandy and loamy materials seem to be similar. There is no significant improvement over the single model.
In order to evaluate whether the prediction depends on the calibration data, calibration of the power function model is made separately on the data of Forrest, Bridge, Bristow and Other (SydU + published papers). The model is fitted using linear regression on a log-log plot of $\phi_e$ and $K_s$ where dubious points (outliers) are excluded first. The results are:

Forrest: $\ln K_s = 10.8731 + 3.9140 \ln \phi_e$ ($R^2 = 0.6982, N = 118$)

Bristow: $\ln K_s = 10.4778 + 3.4106 \ln \phi_e$ ($R^2 = 0.5932, N = 149$)
Bridge: \( \ln K_s = 9.3413 + 4.3595 \ln \phi_e \) \( (R^2 = 0.6321, N = 103) \)

Other: \( \ln K_s = 8.7054 + 2.4901 \ln \phi_e \) \( (R^2 = 0.3736, N = 59) \)

Combined: \( \ln K_s = 10.0515 + 3.662 \ln \phi_e \) \( (R^2 = 0.4379, N = 428) \)

Figure 4.3.4. Saturated hydraulic conductivity as a function of effective porosity calibrated using different data sets: (a) Forrest, (b) Bristow, (c) Bridge, and (d) other data.

Clearly the linear trend is apparent in the individual data sets (Figure 4.3.4). Each data set exhibits a unique relationship with different \( B \) and \( m \), the exponent \( m \) ranges between 2.5 and 4.4. Although the data of Forrest comprises different types of soil (clay content varying from 5 - 69%), a strong linear relationship is present. The data of Bristow also consists of a wide range of texture with sample volume the same as Forrest et al. (1985) (0.076m diameter, 0.050m depth). The prediction is similar, but small values of \( \phi_e \) (< 0.05) are present. The data of Bridge consist only 2 major texture groups: sandy loam and clay. The sandy loams are characterised by the larger value of \( \phi_e \) as shown clustered near the top right hand side of Fig 4.3.4c. The other
data (SydU + published papers) show the largest variation. The latter 2 data sets show different results from the previous ones. We can suggest that the stratification based on texture does not improve the prediction because the variation is caused by some inconsistency within the combined data set rather than texture itself.

Using the fractal model, the model for \( K_s \) is given by

\[
K_s = 4.41 \times 10^8 \left( \frac{\phi^{4/3}}{0.00593 D^{1.83}} \right)^2 \left( \frac{1}{0.0467 - 0.0025 d_g} \right)^2
\]  

[4.3.25]

Although this is supposedly a physically-based model, the prediction is not significantly different from the empirical ones. This suggests the estimate of \( D \) from particle-size distribution is merely empirical and does not reflect the fractal dimension of the particle distribution. Using the single PTF also tends to scale the \( K_s \) values as demonstrated in Figure 4.3.5. The predicted \( \ln(K_s) \) from equation 4.3.25 ranges from 0.06 to 6.03, while the original data ranges from -5.17 to 9.53.

![Figure 4.3.5](image_url)

Figure 4.3.5. Comparison of the measured and predicted \( K_s \) using fractal model.

The ANN developed from the current data set (K13 and K14) does not provide better predictions than other models. Because of the large number of parameters used, the AIC is also the greatest. There is also no significant improvement when using more input data (K13 compared with K14). The prediction quality is also smaller compared to the one developed in the US (K07 and K08). As shown previously the combined data set used in this paper is somewhat inconsistent. Tamari and Wösten
(1999) noted that improvement of prediction using ANN may be achieved if a large soil database (> 1000 samples) with accurate data is used for training.

4.3.3.3. Evaluation of PTFs based on texture

To evaluate the predictability of the PTFs based on texture, the whole data set (prediction and validation) was combined and stratified into 3 textural classes (sandy, loamy and clayey). Results in Figure 4.3.6 show differences compared to previously presented results. Generally, the PTFs predicted better in sandy and loamy soil compared with clayey soil. As a general guide for published PTFs, PTF K04 (Dane and Puckett 1994) gave better predictions for sandy soil; while K03 (Cosby et al. 1984) is suited for loamy soil; and K07 and K08 (Schaap et al. 1998) for clayey soil.

Differences in the value of goodness-of-fit criteria are also apparent, for loamy soil the GMER for K07 is 0.37 but the GRMSR is the smallest amongst the PTFs (9.87 mm/h). GMER is the ratio between the measured and predicted, while GRMSR is the absolute differences (residuals) between the measured and predicted.

The single-power model (K09) showed good prediction for sandy soil, but the GRMSR for clayey soil is large (188 mm/h). In this case textural stratification of the function (K10) significantly decreased the uncertainty (29 mm/h). The continuous power function model (K10) and MLR model (K11) provided a good estimate for the 3 texture classes.

The use of PTFs for predicting $K_s$ is very problematic and carries with it a large degree of uncertainty. This is due to the inherent variability of $K_s$ itself, the measurement is highly sensitive to conditions such as sample volume, method of measurement, measurement error, and spatial variability. The variation between data sets here also added to the uncertainty.
Figure 4.3.6. Performance of pedotransfer functions based on texture class (a) geometric mean error ratio, (b) Geometric standard deviation of error ratio, and (c) Geometric root mean squared residuals.
4.3.4. Conclusions

Development and comparison of PTFs have been made on a combined Australian soil data set. Results show that published PTFs can be used depending on the accuracy expected from the prediction. Since the variables available in the data set are quite restrictive, increasing the number of predictor variables and stratification based on texture only showed slight improvement over prediction in the current study. Pedotransfer functions have been developed from the data set used in this paper. On a coarse-scale the PTFs can provide useful information. But because the data are derived from different types of studies over a wide-range of soil types, when using the function for site-specific prediction the results may have large uncertainties. The inconsistency of prediction between data sets emphasises the importance of the sample volume and method of measurement. Hence local calibration is needed and the use of standardised measurements on a representative volume is necessary (e.g. the procedures of McKenzie and Jacquier, 1996). Currently there are no large data sets with standard measurements, thus further investigation using standardised measurements over a region with additional information such as organic matter and pore structure will be valuable.
4.4. Functional examination of PTFs

Validation of hydraulic pedotransfer functions (PTFs) usually involves comparing the observed/measured hydraulic properties with those predicted. As the main objective of PTFs is to serve as input for simulation models, it seems wise that validation should be evaluated in the final application. This has been called functional validation (Vereckeen et al., 1990). The following study aims at evaluating the application and predictability of PTFs for specific applications.

4.4.1. Methods

4.4.1.1. Data and PTFs

Functional examination was used to evaluate the performance of parametric PTF. The PTFs used are the ones developed in this study (§ 4.2 & 4.3). Water retention was predicted using extended nonlinear regression (ENR) and artificial neural networks (ANP), while saturated hydraulic conductivity $K_s$ was predicted using the power function model (Equation 4.3.23). Effective porosity was evaluated from $\theta_{-10}$ predicted from water retention PTFs. The neural-networks PTF recently developed by the US Salinity Laboratory (Schapp et al., 1998) were used as a comparison. The model predicts the van Genuchten parameters and $K_s$. Since the input requires silt (P$_{2-50}$) and sand sized fractions (P$_{50-2000}$) according to the USDA, P$_{2-20}$ and P$_{20-2000}$ from the Australian data set were first transformed into the equivalent particle size classes by using Equation [4.1.2].

The data were those used in developing water retention PTFs (§ 4.2), i.e. the prediction and validation data set. An additional data set from the study of Bridge (1968) was added in this study (104 soil samples, 1072 water retention points). For functional validation that requires hydraulic conductivity, only data selected from Forrest (35 soil) and Bridge (49 soil) were used. Calculation of soil-water properties usually required $K(\theta)$ relationships. Since only $K_s$ are available, the unsaturated hydraulic conductivity is estimated from the Mualem-van Genuchten model (van Genuchten, 1980):

$$K(S_e) = K_s S_e^{1/2} \left[1 - \left(1 - S_e^{1/n}\right)^{\frac{1}{m}}\right]^\beta$$  \[4.4.1\]
4.4.1.2. Functional examination

Four functional properties were used to assess the applicability of PTFs. Since no real measurements are taken, the ‘true’ values of the functional properties were calculated using the measured water retention and hydraulic conductivity $K_s$. The functional properties used to evaluate PTFs are:

1. Available water content (AWC),
   defined as the difference in water content measured at $h = -33$ kPa and $-1500$ kPa:
   $$\theta_{-33} - \theta_{-1500}.$$  
   This calculation only uses 2 points from the water-retention curve.

2. Specific water-yield, defined as the amount of water that will drain out of the profile from an originally saturated soil (Hanks, 1992). Hanks (1992) shows that the amount of outflow as a function of time can be predicted by:
   $$Q_t = Q_\infty - Q_0 \exp\left(-\frac{Q_t}{Q_\infty} \right)$$  \[4.4.2\]
   where $Q_t$ is the cumulative outflow at time $t$, $Q_\infty$ is the total outflow at equilibrium and $Q_0$ is the initial flow rate. The amount of water drained from the profile at infinite time $Q_\infty$ is calculated from:
   $$Q_\infty = \theta_s L - W_L$$
   where $W_\infty$ is the water storage of the soil profile in equilibrium with a groundwater table $L$. The equilibrium pressure head-distribution above the water table is given by $h = -z$. Water storage in a soil profile of length $L$ at equilibrium is calculated from:
   $$W_L = \int_0^L \theta(h(z)) \, dz$$  \[4.4.3\]
   The integration is done numerically using the parameterized water retention curves.

Assuming the water table is 2 m below the soil surface and initial flow rate $Q_0$ 1200 mm day$^{-1}$, we estimate the amount of water draining out of the profile after 1 day ($Q$). This function employs the water retention curve data in the range of 0 to -20 kPa.

3. Cumulative infiltration under ponding, evaluated according to the physically-based model by Barry et al. (1995):
The dimensionless infiltration $I^*$ and dimensionless time $t^*$ is given by:

$I^* = t^* + 1 - \gamma - \exp\left[\frac{-2t^*}{1 + \frac{2t^*}{3}}\right] \left\{\begin{matrix} \frac{\gamma}{1 + t^*} \\
\exp\left[-\frac{2t^*}{3}\right]\left[1 - (1 - \gamma)^8 t^{n+1}\right] + (2\gamma + t^*)\ln\left(1 + \frac{t^*}{\gamma}\right)\end{matrix}\right\}$


\[4.4.4\]

$$I^* = \left(I - K_n t\right)\frac{2\Delta K}{S_0^2 + 2K_s h_{surf} \Delta \theta}$$

$$t^* = \frac{2t \Delta K^2}{S_0^2 + 2K_s h_{surf} \Delta \theta}$$

[4.4.5]

where $h_{surf}$ is depth of water ponded at the soil surface, $h_{str}=$minimum soil water pressure head at which there is a continuous gas phase within the soil as the soil is being wetted, $\Delta \theta = \theta_s - \theta_n$ and $\Delta K = K_s - K_n$. Subscript $s$ refers to the saturated condition and subscript $n$ refers to initial condition. $S_0$ is the sorptivity, which is calculated according to the approximation of Parlange (1975):

$$S_0^2 = \int_{h_n}^{0} [\theta_s + \theta(h) - 2\theta_n] K(h) \, dh$$

[4.4.6]

The term $\gamma$ is a dimensionless parameter defined as

$$\gamma = \frac{2K_s (h_{surf} + h_{str}) \Delta \theta}{S_0^2 + 2 K_s h_{surf} \Delta \theta}$$

[4.4.7]

$\gamma$ should be between 0 and 1. Parameter $h_{str}$ can be approximated by (Fuentes et al., 1992):

$$h_{str} = \lambda_s \frac{\Delta K}{K_s}, \text{ with } \lambda_s = \frac{1}{K_s} \int_{h_n}^{0} \frac{K(h) - K_n}{\Delta K S_s(h)} \, dh$$

[4.4.8]

The calculation involves using both $\theta(h)$ and $K(h)$ relationships. Calculations involving integration (Equations [4.4.6] and [4.4.8]) were computed numerically using adaptive-Gaussian-quadrature (Kahaner et al. 1987). The soil is assumed to be uniform with initial potential of -100 kPa and under 50 mm of ponded water. Cumulative infiltration ($I$) after 1 hour of ponding is calculated as the validation criterion.
(4) Dynamic simulation
Dynamic simulation of soil water status was evaluated using two approaches. The first model is the empirical ‘tipping-bucket’ model of Burns (1974). The model is based on the local mass balance of water (Figure 4.4.1). The soil profile is divided into layers or compartments of 1 cm. The amount of water infiltrating or evaporating from the soil is calculated from the difference between the amount of precipitation and evaporation. The maximum amount of water that can be held in the soil is defined at the upper limit of field capacity FC (-10 kPa). In the case of infiltration, the water will infiltrate into each compartment until it reaches field capacity, after that the excess water will be transported to the next underlying compartment. Evaporation is simulated by the loss of water until a critical limit, here defined at $\theta_r$. When $\theta_r$ is reached, no more water can evaporate from the soil compartment and the evaporation demand is met from the underlying compartment. The total water storage of the soil is calculated from the sum of water at each compartment in the profile.

![Figure 4.4.1. The empirical tipping-bucket model for soil-water flow.](image)

The second model is the mechanistic model SWAP (van Dam et al., 1997). The model is a finite-difference numerical solution of Richards’ equation. This model
requires both the water retention and the hydraulic conductivity curve. Weather data (solar radiation, temperature, wind speed, vapour pressure, and precipitation) from the cotton growing area of Bourke, NSW were used to drive the model (Figure 4.4.2). The data was obtained from the Australian Cotton Cooperative Research Centre website: http://www.cotton.pi.csiro.au/Tools/Weather. The soil was assumed to be under fallow and evaporation was calculated according to the Penman-Monteith approach (Smith, 1991). The soil profile was uniform to 1 m depth, with an initial matric potential at -33 kPa, and free drainage at the bottom of the profile. Simulation was carried out from April 1997 to December 1998. The first year (April – December 1997) was considered as a ‘warm-up’ period in order to minimise the effect of initial moisture conditions, only results from 1998 were analysed.

![Figure 4.4.2. Precipitation (bar chart) and evaporation (dotted lines) for 1998 at Bourke, NSW, used for simulation of water movement in the field.](image)

Results were recorded every 10 days, a total of 36 observations for 1 year. The outputs evaluated are:
• the water content at 3 depths: $\theta_1 = 50$ mm, $\theta_2 = 200$ mm, and $\theta_3 = 300$ mm below the soil surface;

• the water storage in the profile ($WS$) to a depth of 0.5 m:

$$WS = \int_{0}^{0.5} \theta(z) \, dz ;$$

• the soil moisture deficit ($SMD$), which is the difference between the amount of water at the upper limit of field capacity (-10 kPa) and the water storage in the field condition. If the soil is wetter than -10 kPa there will be no moisture deficit ($SMD = 0$), or

$$SMD = \int_{0}^{0.5} \Delta \theta(z) \, dz ;$$

where $\Delta \theta = (\theta - 10 - \theta)$ when $\theta < \theta_{10}$, and $\Delta \theta = 0$ when $\theta \geq \theta_{10}$.

4.4.1.3. Performance criteria

Statistical measures used to evaluate the goodness of fit are: bias, inaccuracy, imprecision, and the median of relative absolute error.

- **Bias** is calculated from mean error (ME):

$$ME = \frac{1}{N} \sum_{i=1}^{N} (Y_i - \hat{Y}_i)$$  \[4.4.9\]

where $Y$ is the property predicted using the actual hydraulic properties and $\hat{Y}$ is the property predicted from PTFs.

- **Inaccuracy** is calculated from root mean squared error (RMSE):

$$RMSE = \left[ \frac{1}{N} \sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2 \right]^{1/2}$$  \[4.4.10\]

- **Imprecision** is calculated according to the following relationship:

$$\text{Inaccuracy}^2 = \text{Bias}^2 + \text{Imprecision}^2$$  \[4.4.11\]

- To account for relative error, the median of the relative absolute error (RAE) was calculated:

$$\text{RAE} = \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \times 100\%$$
4.4.2. Results and discussion

4.4.2.1. Available water content

The performance of the parametric PTFs for predicting AWC is shown in Tables 4.4.1 and 4.4.2. Although ENR always performed better than ROS4 in point prediction (Table 4.4.3), it does not perform well when predicting AWC. Similar explanation can be given to that of §4.2.2.3. When evaluating AWC by taking the difference, the covariance between FC and PWP must be taken into account (Equation 4.2.19). In this case the correlation of the error in FC and PWP for ROS4 is larger than for ENR2 resulting in a smaller prediction error. It can also be seen that the correlation between the FC and PWP values predicted by ROS4 have values similar to the real data (Table 4.4.3), while ENR2 produced a larger correlation. ROS4 has a similar covariance structure compared with the real data and outperforms ENR in predicting AWC. A similar trend was observed when evaluating the performance based on texture (Figure 4.4.2b). As seen in Table 4.4.2 the error in prediction is larger in clayey soil compared with sandy and loamy materials.

<table>
<thead>
<tr>
<th>Table 4.4.1. Prediction of AWC (m³ m⁻³).</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bias</strong></td>
</tr>
<tr>
<td>Pred.</td>
</tr>
<tr>
<td>ROS4</td>
</tr>
<tr>
<td>ANP</td>
</tr>
<tr>
<td>ENR2</td>
</tr>
<tr>
<td>ENR8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4.4.2. Relative error in prediction of AWC (m³ m⁻³).</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Median RAE (%)</strong></td>
</tr>
<tr>
<td>ROS4</td>
</tr>
<tr>
<td>ANP</td>
</tr>
<tr>
<td>ENR2</td>
</tr>
<tr>
<td>ENR8</td>
</tr>
</tbody>
</table>
Table 4.4.3. Inaccuracy and covariance between FC and PWP as evaluated by PTFs.

<table>
<thead>
<tr>
<th></th>
<th>PTFs</th>
<th>True</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ROS4</td>
<td>ENR2</td>
</tr>
<tr>
<td>Prediction set</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inaccuracy FC</td>
<td>0.068</td>
<td>0.043</td>
</tr>
<tr>
<td>Inaccuracy PWP</td>
<td>0.085</td>
<td>0.026</td>
</tr>
<tr>
<td>Corr(ε_F, ε_PWP)</td>
<td>0.632</td>
<td>-0.113</td>
</tr>
<tr>
<td>Corr(FC, PWP)</td>
<td>0.900</td>
<td>0.949</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.856</td>
</tr>
<tr>
<td>Validation set</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inaccuracy FC</td>
<td>0.101</td>
<td>0.090</td>
</tr>
<tr>
<td>Inaccuracy PWP</td>
<td>0.096</td>
<td>0.024</td>
</tr>
<tr>
<td>Corr(ε_F, ε_PWP)</td>
<td>0.642</td>
<td>0.320</td>
</tr>
<tr>
<td>Corr(FC, PWP)</td>
<td>0.862</td>
<td>0.966</td>
</tr>
</tbody>
</table>

(a) Inaccuracy of predicted available water content based on (a) data sets, and (b) texture class.
4.4.2.2. Specific water yield

The prediction of specific water yield is shown in Tables 4.4.4 and 4.4.5 and Figure 4.4.4. When evaluating the prediction based on texture (Figure 4.4.4b), the performance of all PTFs is quite similar, except ROS4 with a higher imprecision in clayey soil. The imprecision increases with increasing clayeyness, although ROS6 used two points from the water retention, the relative error is larger compared to ENR and ANP. This showed that PTFs with good prediction in water retention may not always produce good functional estimates.

<table>
<thead>
<tr>
<th>PTFs</th>
<th>Bias Pred</th>
<th>Valid</th>
<th>Bridge</th>
<th>Inaccuracy Pred</th>
<th>Valid</th>
<th>Bridge</th>
<th>Imprecision Pred</th>
<th>Valid</th>
<th>Bridge</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROS4</td>
<td>28.8</td>
<td>35.8</td>
<td>-48.0</td>
<td>78.4</td>
<td>90.2</td>
<td>106.8</td>
<td>72.9</td>
<td>82.8</td>
<td>95.4</td>
</tr>
<tr>
<td>ROS6</td>
<td>22.1</td>
<td>14.4</td>
<td>-52.0</td>
<td>69.1</td>
<td>74.4</td>
<td>64.7</td>
<td>65.5</td>
<td>73.0</td>
<td>38.5</td>
</tr>
<tr>
<td>ANP</td>
<td>-10.0</td>
<td>29.0</td>
<td>-74.3</td>
<td>58.5</td>
<td>95.2</td>
<td>111.5</td>
<td>57.6</td>
<td>90.7</td>
<td>83.2</td>
</tr>
<tr>
<td>ENR2</td>
<td>-20.8</td>
<td>48.6</td>
<td>-53.5</td>
<td>61.5</td>
<td>114.2</td>
<td>71.4</td>
<td>57.9</td>
<td>103.4</td>
<td>47.3</td>
</tr>
<tr>
<td>ENR8</td>
<td>-21.6</td>
<td>68.5</td>
<td>-27.9</td>
<td>65.5</td>
<td>133.8</td>
<td>50.5</td>
<td>61.9</td>
<td>115.0</td>
<td>42.1</td>
</tr>
</tbody>
</table>

Table 4.4.5. Relative error in prediction of specific water-yield $Q$ (mm).

<table>
<thead>
<tr>
<th>PTFs</th>
<th>Median RAE(%)</th>
<th>Sandy</th>
<th>Loamy</th>
<th>Clayey</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROS4</td>
<td>24.9</td>
<td>41.1</td>
<td>79.0</td>
<td></td>
</tr>
<tr>
<td>ROS6</td>
<td>19.5</td>
<td>45.3</td>
<td>63.5</td>
<td></td>
</tr>
<tr>
<td>ANP</td>
<td>23.3</td>
<td>35.3</td>
<td>33.9</td>
<td></td>
</tr>
<tr>
<td>ENR2</td>
<td>22.3</td>
<td>32.8</td>
<td>38.2</td>
<td></td>
</tr>
<tr>
<td>ENR8</td>
<td>23.8</td>
<td>36.2</td>
<td>56.3</td>
<td></td>
</tr>
</tbody>
</table>
4.4.2.3. Cumulative infiltration under ponding

The data of Forrest (35 soil samples) and Bridge (49 soil samples) were used to calculate the cumulative infiltration under ponding. Comparisons were made for ENR: water retention by ENR2 and $K_s$ predicted from Equation 4.3.23; ENR+$K$: ENR2 with measured/true $K_s$; and ROS6. The ‘true’ infiltration values were calculated using measured water retention and $K_s$. Table 4.4.6 show that ENR performed better than RO6 but showed a large bias. When the true $K_s$ is used in conjunction with water retention PTFs the prediction increased significantly. A one-to-one line almost fits the prediction versus true values (Figure 4.4.6). Since the infiltration is under ponded condition, the water flow is mostly controlled by $K_s$ alone, and water retention only plays a small role.
Table 4.4.6. Prediction of cumulative infiltration ($I$) after 1 hour of ponding.

<table>
<thead>
<tr>
<th>Cumulative infiltration (mm)</th>
<th>ROS6</th>
<th>ENR</th>
<th>ENR+K</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forrest</td>
<td>-16.5</td>
<td>-11.8</td>
<td>5.4</td>
</tr>
<tr>
<td>($N=35$)</td>
<td>Inaccuracy 37.3</td>
<td>33.4</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td>Imprecision 33.4</td>
<td>31.2</td>
<td>5.2</td>
</tr>
<tr>
<td>Bridge</td>
<td>-15.1</td>
<td>5.1</td>
<td>2.3</td>
</tr>
<tr>
<td>($N=49$)</td>
<td>Inaccuracy 48.6</td>
<td>38.4</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td>Imprecision 46.1</td>
<td>38.1</td>
<td>4.7</td>
</tr>
</tbody>
</table>

Figure 4.4.5. Inaccuracy of predicted cumulative infiltration based on texture class.

Figure 4.4.6. Predicted versus true cumulative infiltration after one hour of ponding.
The inaccuracy also increased with increasing time, following the trend of infiltration curve (Figure 4.4.7a). When plotted on a relative scale (error divided by the ‘true’ value) the relative error is quite constant. When predicting ponding by only relying on PTFs a large inaccuracy can be expected (~580% for ROS6, ~270% for ENR).

![Figure 4.4.7. (a) Inaccuracy and (b) relative inaccuracy of predicted cumulative infiltration as a function of time.](image)

4.4.2.4. Dynamic simulation

4.4.2.4.1. Tipping bucket

The ‘tipping bucket’ model gives rough prediction of the water storage and soil moisture deficit, as it only assumes that water in the soil filled-up to field capacity before being transmitted to the bottom layer (Table 4.4.7). As the calculations rely heavily on FC and θ, ROS6 predicted better than ENR. The ENR poorly estimated water storage in the profile with an error up to 500%. This calculation might not reflect the true condition, but this functional criterion suggested that for simulation which simply uses few points in the retention curve, good estimates are needed to produce reasonable results and ENR cannot be applied in this situation.
Table 4.4.7. Prediction of water storage in 0.5 m soil profile using Burns’ model

<table>
<thead>
<tr>
<th></th>
<th>WS (mm)</th>
<th>SMD (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ENR</td>
<td>ROS6</td>
</tr>
<tr>
<td>Forrest</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias</td>
<td>52.6</td>
<td>8.5</td>
</tr>
<tr>
<td>Inaccuracy</td>
<td>117.2</td>
<td>25.0</td>
</tr>
<tr>
<td>Imprecision</td>
<td>104.8</td>
<td>23.5</td>
</tr>
<tr>
<td>Median RAE (%)</td>
<td>310</td>
<td>48</td>
</tr>
<tr>
<td>Bridge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias</td>
<td>48.7</td>
<td>20.0</td>
</tr>
<tr>
<td>Inaccuracy</td>
<td>69.0</td>
<td>27.0</td>
</tr>
<tr>
<td>Imprecision</td>
<td>48.9</td>
<td>18.1</td>
</tr>
<tr>
<td>Median RAE (%)</td>
<td>536</td>
<td>91</td>
</tr>
</tbody>
</table>

4.4.2.4.2. Mechanistic model

Tables 4.4.8 and 4.4.9 showed the error in prediction of soil moisture storage and deficit and water content for the Forrest and Bridge data set. Comparison with the average imprecision showed similar prediction for all the PTFs. There was no indication that inclusion of true $K_s$ would improve the prediction. But when we plotted the median of the relative error with time, some explanations can be given. Using ENR with true $K_s$ produced lower imprecision and bias in water storage and soil moisture deficit (Figures 4.4.8 and 4.4.9). Surface water content is also better predicted (Figure 4.4.8c) using true $K_s$ values. The errors also followed the precipitation and evaporation. The smallest error appeared to be between Julian days 200 and 240, after a high rainfall event (Figure 4.4.2). The largest imprecision and bias occurred during the period of low rainfall and high evaporation (day 2-100, and 250-360). Using hydraulic PTFs as input to the soil-water model, error in prediction of water storage and soil water content were quite large (within 20 – 80 %). This is illustrated using measured hydraulic properties as a reference. In the real situation, field measurements of water storage and water content may not well be reproduced when simulated using measured hydraulic properties. Depending on the scale of application, the simulation using PTF-predicted hydraulic properties may well be good enough to predict the soil-water status.
Table 4.4.8. Prediction of soil moisture storage and soil moisture deficit in 0.5 m soil profile (mm).

<table>
<thead>
<tr>
<th>Water storage (mm)</th>
<th>Soil moisture deficit (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROS6</td>
<td>ENR</td>
</tr>
</tbody>
</table>

Forrest $N = 1260$

<table>
<thead>
<tr>
<th>Bias</th>
<th>12.20</th>
<th>-10.81</th>
<th>-17.92</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inaccuracy</td>
<td>53.69</td>
<td>49.56</td>
<td>42.53</td>
</tr>
<tr>
<td>Imprecision</td>
<td>52.29</td>
<td>48.37</td>
<td>38.57</td>
</tr>
</tbody>
</table>

Bridge $N = 1764$

<table>
<thead>
<tr>
<th>Bias</th>
<th>2.68</th>
<th>20.83</th>
<th>18.20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inaccuracy</td>
<td>54.44</td>
<td>51.36</td>
<td>52.83</td>
</tr>
<tr>
<td>Imprecision</td>
<td>54.38</td>
<td>46.95</td>
<td>49.60</td>
</tr>
</tbody>
</table>

Table 4.4.9. Prediction of water content at 3 depth ($\theta_1 = 50$ mm, $\theta_2 = 200$ mm, $\theta_3 = 300$ mm below soil surface).

<table>
<thead>
<tr>
<th>$\theta_1$</th>
<th>$\theta_2$</th>
<th>$\theta_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROS6</td>
<td>ENR</td>
<td>ENR+K</td>
</tr>
</tbody>
</table>

Forrest $N = 1260$

<table>
<thead>
<tr>
<th>Bias</th>
<th>-0.009</th>
<th>0.026</th>
<th>0.021</th>
<th>0.006</th>
<th>0.038</th>
<th>0.032</th>
<th>0.010</th>
<th>0.043</th>
<th>0.037</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inacc.</td>
<td>0.103</td>
<td>0.095</td>
<td>0.102</td>
<td>0.115</td>
<td>0.102</td>
<td>0.103</td>
<td>0.114</td>
<td>0.107</td>
<td>0.107</td>
</tr>
<tr>
<td>Imprec.</td>
<td>0.103</td>
<td>0.091</td>
<td>0.099</td>
<td>0.115</td>
<td>0.095</td>
<td>0.098</td>
<td>0.113</td>
<td>0.098</td>
<td>0.100</td>
</tr>
</tbody>
</table>

Bridge $N = 1764$

<table>
<thead>
<tr>
<th>Bias</th>
<th>-0.058</th>
<th>-0.009</th>
<th>0.004</th>
<th>-0.047</th>
<th>-0.005</th>
<th>0.013</th>
<th>-0.039</th>
<th>0.001</th>
<th>0.019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inacc.</td>
<td>0.150</td>
<td>0.154</td>
<td>0.132</td>
<td>0.149</td>
<td>0.157</td>
<td>0.136</td>
<td>0.150</td>
<td>0.157</td>
<td>0.139</td>
</tr>
<tr>
<td>Imprec.</td>
<td>0.138</td>
<td>0.154</td>
<td>0.132</td>
<td>0.142</td>
<td>0.157</td>
<td>0.135</td>
<td>0.145</td>
<td>0.157</td>
<td>0.137</td>
</tr>
</tbody>
</table>

The application may also depend on the spatial and temporal scale. This study evaluated the prediction over the period of days. Comparison with true conditions in the field will be quite difficult to achieve. Cresswell and Paydar (2000) evaluated PTFs in terms of cumulative drainage and evapotranspiration averaged over a five-year period. They found that their PTFs predicted within a mean error of 3 – 4 mm/year or 3%. A relatively good prediction can be obtained when using cumulative values averaged over certain period. Therefore temporal scales determine the utility of PTFs in soil-water balance prediction. Wosten et al. (1990) found no significant difference when using direct measurements and PTFs predicted hydraulic properties in estimating evapotranspiration deficit and flux during seven years period. But
meteorological data strongly affected the calculation, when rainfall deficit was taken into account as a covariable, the methods were different significantly.

A comparison between the ‘true’ water storage computed with the mechanistic model SWAP and predicted using the Burns’ empirical model is shown in Table 4.4.10 and Figure 4.4.10. It is interesting to note that the predicted water storage from the empirical model correspond well with the mechanistic model. With a relative error of 25-50% the empirical model may be as valid to use as the mechanistic model. This is not a recommendation, but merely shows that prediction using PTFs can have a very large error at fine temporal and spatial scales. Since the results are mostly affected by meteorological data, both empirical and mechanistic model may show the trend in water balance, but to predict the water storage the error is quite large.

Table 4.4.10. Difference between water storage (mm) calculated by empirical model and the ‘true’ value calculated from mechanistic model.

<table>
<thead>
<tr>
<th></th>
<th>Water storage (mm)</th>
<th>Soil moisture deficit (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ENR</td>
<td>ROS6</td>
</tr>
<tr>
<td>Forrest  N = 1260</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias</td>
<td>-12.1</td>
<td>-46.0</td>
</tr>
<tr>
<td>Inaccuracy</td>
<td>40.9</td>
<td>68.7</td>
</tr>
<tr>
<td>Imprecision</td>
<td>39.1</td>
<td>51.0</td>
</tr>
<tr>
<td>Median RAE (%)</td>
<td>25.6</td>
<td>54.2</td>
</tr>
<tr>
<td>Bridge  N = 1764</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias</td>
<td>-34.2</td>
<td>-62.9</td>
</tr>
<tr>
<td>Inaccuracy</td>
<td>75.8</td>
<td>116.1</td>
</tr>
<tr>
<td>Imprecision</td>
<td>67.7</td>
<td>97.6</td>
</tr>
<tr>
<td>Median RAE (%)</td>
<td>50.0</td>
<td>63.6</td>
</tr>
</tbody>
</table>
Figure 4.4.8. Relative error of prediction of water storage and soil moisture deficit as a function of time for the Bridge data set.
Figure 4.4.9. Bias and relative error in prediction of water storage and soil water content as a function of time for the Forrest data set.
4.4.3. Conclusions

Functional examination has shown that functional criteria predicted by PTFs range from very poor to moderate. The following conclusions can be drawn for four kinds of functional criteria:

(1) Available water content can be predicted within 40 – 50% of error. The error is larger with increasing clay content.

(2) Specific water-yield, which is derived from the area of the water retention curve, PTFs produced an error of 30 - 40%. The error is also larger with increasing clay content.

(3) Ponded infiltration, cumulative water infiltration is more dependent on prediction of $K_s$ rather than water retention curve. The poor estimation of $K_s$ by PTFs therefore does not permit good predicts of infiltration.

(4) Dynamic simulation, PTFs can be used to predict dynamic water content and water storage with an error of 40-60%. Inclusion of true $K_s$ does not increase the predicability. The results are strongly dependent on meteorological data.
References


Prebble, R.E. 1970b. Soil physical measurements and grass roots distributions on a red podzolic at Samford, South East Queensland. CSIRO Division of Soils Technical Memorandum 13/70. CSIRO, Australia.


Tietje, O., Hennings, V., 1996. Accuracy of the saturated hydraulic conductivity prediction by pedo-transfer functions compared to the variability within FAO textural classes. *Geoderma* 69, 71-84.


V. Uncertainty analysis for pedotransfer functions

5.1. Introduction

Since the pioneering work by De Wit and van Keulen (1973), modelling soil processes has become an important tool in soil science, for example evaluating environmental quality and predicting crop yield. Pedotransfer functions have been developed particularly for translating readily-available information into the variables that are needed in simulation models. There are many uncertainties inherent in the input data and consequently the PTFs and simulation models can produce variable outputs. In many applications, it is no longer acceptable to only model a process and then present the best prediction from the output of the model (Iman, 1992). The uncertainty in the output must also be considered, and the main contributors to the uncertainty must be identified. Uncertainty analysis is required to gauge the reliability of the prediction, this is essential for ultimately assessing risks of applying certain input to the environment. However, uncertainty is seldom quantified in pedotransfer functions.

There are many factors that require consideration when using PTFs, such as the sensitivity of the model, uncertainties in parameters and data, and amount of effort needed to obtain the data. Vereecken et al. (1992) assessed the error in hydraulic PTFs and evaluated its effect on the output of a soil-water movement model. Finke et al., (1996) evaluated the effect of PTF uncertainty in simulating functional soil properties. They found that PTF uncertainty was an important factor when predicting the number of days that the soil is workable (moisture content at $h < -7$kPa) and days with sufficient aeration ($\phi_a \geq 0.1$ m$^3$ m$^{-3}$). Schaap and Leij (1998) considered the uncertainty in hydraulic PTFs based on the accuracy of the database where PTFs were generated.

The uncertainty of the output of a PTF or model can be written as:

$$\epsilon = \epsilon_m + \epsilon_p + \epsilon_v$$

[5.1.1]

where $\epsilon_m$ is the error due to the model bias, $\epsilon_p$ is the error due to parameter uncertainty and $\epsilon_v$ is the error due to the uncertainty in input variables (Leenhardt et al., 1994). A general method of quantifying uncertainty is the so-called Monte Carlo method. The
term ‘Monte Carlo’ was coined by Nicholas Metropolis during the Manhattan Project of World War II, because of the similarity of statistical simulation to games of chance, and the allusion to the capital of Monaco (Computational Science Education Project, 1995). The term Monte Carlo methods or simulations have been broadly used in soil science to refer to uncertainty and sensitivity analysis. However, according to the Encyclopedia of Statistical Sciences, Monte Carlo methods are those in which properties of the distributions of random variables are investigated by use of simulated random numbers (Gentle, 1982). The computational Science Education Project (1995) described it as statistical simulation methods, or any method that utilises sequences of random numbers to perform the simulation. Statistical simulation methods may be contrasted to conventional first-order error analysis, which is usually applied to mathematical equations that describe physical systems (e.g. Flühler et al., 1976). In the Monte Carlo method, the statistical model or physical process is simulated directly, and the only requirement is that the variables of the model can be described by probability density functions. By sampling repeatedly from the assumed probability density function of the input variables and evaluating the response of the model for each sample, the distribution of the results, along with the mean and other statistical measures can be estimated.

The conventional approach to sampling from the distribution function is by simple random sampling (Cochran, 1977). This approach yields reasonable estimates if the sample size drawn is quite large. However, in simulation models, which require a lot of computation time, using a large sample size is not always feasible. Although computation time will be largely reduced with increasing computer technology, it is still necessary to seek methods that reduce the sample size drawn from the distribution while also preserving the statistics. Methods proposed to reduce sample size include Latin hypercube sampling (McKay et al., 1979) and the sectioning method (Addiscott and Wagenet, 1985; Lark, 2000). These methods attempt to provide a full probabilistic coverage of the input variables. Since the final objective is to estimate the response/output of a model, a desired sampling method should provide an accurate (close to the ‘true’ value) and precise (small variance) estimate. The efficiency of sampling methods can be addressed by comparing the variability of the estimate obtained from each sampling method (Iman, 1992).

The objectives of the work presented in this chapter are (1) to present a modified Latin hypercube sampling strategy for sampling multivariate correlated distributions,
(2) to use this method to quantify the uncertainty in pedotransfer functions due to error in measurements of soil properties, and (3) to use the PTFs to evaluate the overall uncertainty in predictions from a soil-water model.

5.2. Sampling methods for uncertainty analysis

5.2.1. Simple random sampling

Simple random sampling involves repeatedly forming random vectors of parameters from prescribed probability distributions. A normally-distributed random variable \( x \) with mean \( \mu \) and standard deviation \( \sigma \) can be generated by:

\[
x^* = \sigma r_n + \mu \tag{5.2.1}
\]

where \( r_n \) are normally distributed random numbers with mean 0 and variance 1.

A multivariate normal distribution with variance-covariance matrix \( V \) can be sampled utilising the lower and upper triangular matrix (LU) decomposition method (Davis, 1987). The variance-covariance matrix \( V \) is first decomposed by Cholesky factorization:

\[
V = L L^T
\]

where \( L \) is the lower triangular matrix. To generate the random variables vector \( x \), matrix \( L \) is multiplied by vector, \( r_n \), of independent normal random numbers with mean 0 and variance 1:

\[
x = L r_n + \mu . \tag{5.2.2}
\]

The procedure is repeated for sample size \( n_s \), resulting in a set of variables with expected mean vector \( \mu \) and expected variance-covariance matrix: \( L \text{cov}(r_n) L^T \). Since the random numbers are independent, the covariance matrix \( \text{cov}(r_n) \) should equal \( I \) (the identity matrix),

\[
L \text{cov}(r_n) L^T = L I L^T = L L^T = V .
\]

5.2.2. The sectioning method

The sectioning method (Addiscott and Wagenet, 1985) uses systematic sampling of the input variables space. For \( k \) variables, the probability distribution of each variable is divided into \( s \) equiprobable sections. The section medians for all variables are then paired in all combinations with other variables. This method assumes independence among the variables. The method can be summarized as follows:
• divide the distribution of each variable into $s$ equiprobable sections;
• select the median of each section and transform the probability into a sampled value using the inverse of the distribution function;
• the $s$ sampled values for each variable are then paired in all combinations with other variables, generating a grid system of $s^k$ combinations.

Addiscott & Wagenet (1985) noted that the number of sections for each variable could be varied according to the goodness of fit of the measurements in the tails of the distribution. Alternatively a random sample can be taken for each section resulting in a stratified random sampling.

5.2.3. Latin hypercube sampling

Latin hypercube sampling (LHS), a stratified-random procedure, provides an efficient way of sampling variables from their distributions (Iman and Conover, 1980). It has been used in soil science and environmental studies recently, e.g. to quantify uncertainty in wheat-production functions (Viscarra Rossel et al., 2000), assessing the uncertainty in a soil-nitrogen model (Hansen et al., 1999), and the simulation of random fields (Pebesma and Heuvelink, 1999).

The LHS involves sampling $ns$ values from the prescribed distribution of each of $k$ variables $x_1, x_2, \ldots, x_k$. The cumulative distribution for each variable is divided into $ns$ equiprobable intervals. A value is selected randomly from each interval (Figure 5.2.1a). The $ns$ values obtained for each variable are paired randomly with the other variables. Unlike simple random sampling, this method ensures a full coverage of the range of each variable by maximally stratifying each marginal distribution (Figure 5.2.1b).

The LHS can be summarized as:
• divide the cumulative distribution of each variable into $ns$ equiprobable intervals;
• from each interval select a value randomly, for the $i$th interval, the sampled cumulative probability can be written as (Wyss and Jorgensen, 1998):
  \[
  \text{Prob}_{i} = \frac{1}{ns} r_u + \frac{(i - 1)}{ns}
  \]
  where $r_u$ is uniformly distributed random number ranging from 0 to 1;
• transform the probability values sampled into the value $x$ using the inverse of the distribution function $F^{-1}$:
  \[
  x = F^{-1}(\text{Prob})
  \]
• the $ns$ values obtained for each variable $x$ are paired randomly (equally likely combinations) with the $ns$ values of the other variables.

The method is based on the assumption that the variables are independent of each other, but in reality most of the input variables are correlated to some extent. Random pairing of correlated variables could result in impossible combinations (e.g. large clay content with small $\theta$), furthermore independent variables tend to bias the uncertainty.

Figure 5.2.1. (a) Random sampling from each probability interval as in Latin hypercube sampling (dark circles) and sampling from the median of each interval as in sectioning method (white triangle), (b) Latin hypercube sampling for two variables (dark dots) and simple random sampling (white rectangular).

5.2.4. Inducing correlation in Latin hypercube sampling

Iman and Conover (1982) proposed a method for inducing correlation among the variables by restricting the way the variables are paired based on the rank correlation of some target values. The method is based on the Cholesky decomposition of the correlation matrix. Suppose matrix $X$ is composed of independent random variables with correlation matrix $I$ and $C$ is the desired correlation matrix. Matrix $C$ can be written as $C = PP'$ where $P$ is the lower triangular matrix. Similar to the simple random sampling, multiplying vector $xP'$ yield random variables with correlation matrix $C$. Therefore the objective is to rearrange the input variables close to the target correlation matrix. The method is summarised as follows:

• generate matrix $R$ using Latin hypercube sampling of $k$ variables at sample size $ns$;
• calculate $T$, the correlation matrix of $R$;
• calculate the $P$ lower triangular matrix of the target correlation matrix $C$ using Cholesky factorization $C = PP^\top$ and also $Q$ the lower triangular matrix of $T$

\[ T = QQ^\top; \]

• solve to obtain matrix $S$ such that $STS^\top = C$, which is calculated from $S = P Q^{-1}$;

• calculate target correlation matrix $R^* = RS^\top$, which has a correlation matrix equal to $C$;

• rearrange the values of each variable in $R$ so they have the same rank (order) as the target matrix $R^*$.

Soutter and Musy (1999) have applied this method for sampling soil properties to perform sensitivity analyses of pesticide leaching models.

Stein (1987) also proposed a method for sampling dependent variables based on the rank of a target multivariate distribution.

• obtain matrix $R$ of $k$ variables at sample size $ns$ using simple random sampling;

• define $U$ the matrix with $k$ columns and $ns$ rows containing the order or rank corresponding to the target correlation matrix;

• obtain the Latin hypercube sample $x_{ij}$ ($i = 1, \ldots, ns; j = 1, \ldots, k$) by

$$x_{ij} = F^{-1}\left(u_{ij} - r_{ij}\right)$$

\[5.2.3\]

With this shifting (transformation), the sampled values yield an approximately joint distribution. Pebesma and Heuvelink (1999) have adapted it for simulation of spatially correlated Gaussian random fields.

The catch with these methods is that they do not necessarily produce the right correlation matrix. Correlations among the variables make some hypercubes more probable than the others. For example, if we wish to obtain a multi-normal distribution of soil properties:

\[
\begin{array}{cccccc}
\mu & \rho_b & \text{clay} & \text{sand} & \theta_{10} & \theta_{1500} \\
\sigma & 1.44 & 27.22 & 55.37 & 0.33 & 0.17 \\
\rho_b & 0.21 & 16.40 & 21.24 & 0.10 & 0.09 \\
\text{clay} & 1.00 & -0.23 & 0.47 & -0.56 & -0.32 \\
\text{sand} & -0.23 & 1.00 & -0.88 & 0.68 & 0.84 \\
\theta_{10} & 0.47 & -0.88 & 1.00 & -0.81 & -0.85 \\
\theta_{1500} & -0.56 & 0.68 & -0.81 & 1.00 & 0.79 \\
\end{array}
\]
Using the method by Iman and Connover, 30 runs of $ns = 10$, give an average correlation matrix of:

$$
\begin{array}{cccccc}
\rho_b & \text{clay} & \text{sand} & \theta_{10} & \theta_{1500} \\
1.000 & -0.192 & 0.406 & -0.531 & -0.302 \\
-0.192 & 1.000 & -0.831 & 0.598 & 0.752 \\
0.406 & -0.831 & 1.000 & -0.705 & -0.784 \\
-0.531 & 0.598 & -0.705 & 1.000 & 0.707 \\
-0.302 & 0.752 & -0.784 & 0.707 & 1.000 \\
\end{array}
$$

5.2.5. A modified sampling method

5.2.5.1. Orthogonal transformation

A modification to the Latin hypercube sampling to deal with correlation between variables is presented here. It was first described and applied by Lark (2000) for the sectioning method, and is based on the transformation of correlated variables into uncorrelated ones by means of principal component analysis. Principal component analysis transforms the correlated variables onto new axes, where the variables are independent of each other (uncorrelated) while also preserving the original information (Webster and Oliver, 1990). The principal components are linear combination of the original variables, which are calculated from the eigenvectors of the variance-covariance matrix.

The variance-covariance matrix $V$ has an eigenvector $e$ and corresponding eigenvalue $\lambda$ according to:

$$
V e = \lambda e
$$

[5.2.4]

The equation can be written as:

$$(V - \lambda I) e = 0$$

To find the eigenvalue $\lambda$ the above equation is solved by setting determinant $(V - \lambda I)$ to zero.

$$|V - \lambda I| = 0$$

This will produce a $k$th degree polynomial in $\lambda$ whose roots are the eigenvalues.

The principal component of vector $x$ with mean $\mu$ is defined as (Mardia et al., 1979):

$$
P = E^T (x - \mu)
$$

[5.2.5]
where $E$ is the matrix with $k$ columns of eigenvectors $e$. The $i$th principal component of $x$ may be written as:

$$p_i = e_i^T (x - \mu)$$

Latin hypercube sampling with random pairing then can be applied to the transformed variables $P$. The transformed coordinates are then transformed into real coordinates,

$$x = E P + \mu.$$  [5.2.6]

The method for sampling correlated variables via the LU decomposition is based on the assumption of multivariate normality, therefore the orthogonal transformation also yields a normal distribution. The back transformation assures a multivariate normal-distribution with variance-covariance matrix $V$.

The discrepancy of this method appears when transformation is made into the orthogonal space, sampled values from orthogonal space do not have the same probability as prescribed in the original space. Consider a normal distribution in Figure 5.2.2a, the dark dots represent five intervals of the probability 0.1, 0.3, 0.5, 0.7 and 0.9. If we generated a bivariate correlated normal distribution, sampled from the orthogonal space, and transformed back to the original scale, the probabilities sampled are 0.04, 0.24, 0.51, 0.77, and 0.96. If the sectioning method is adopted as per Lark (2000), the back transformation does not yield equiprobable intervals in its original space. If we sampled a value randomly from each interval, as in Latin hypercube sampling, in the orthogonal space, back transformation does not yield marginally maximally stratified samples. This sampling does not yield one value per equiprobable interval, some intervals are not sampled and some intervals are sampled more than once (Figure 5.2.2b).

The advantage of this method is that it reproduces a correct correlation matrix. Using the example as in §5.2.4, 30 runs of $ns = 10$, give an average correlation matrix:

<table>
<thead>
<tr>
<th></th>
<th>$\rho_b$</th>
<th>clay</th>
<th>sand</th>
<th>$\theta_{\tau 10}$</th>
<th>$\theta_{\tau 1500}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_b$</td>
<td>1.00</td>
<td>-0.226</td>
<td>0.468</td>
<td>-0.518</td>
<td>-0.278</td>
</tr>
<tr>
<td>clay</td>
<td>-0.226</td>
<td>1.000</td>
<td>-0.880</td>
<td>0.682</td>
<td>0.860</td>
</tr>
<tr>
<td>sand</td>
<td>0.468</td>
<td>-0.880</td>
<td>1.000</td>
<td>-0.803</td>
<td>-0.845</td>
</tr>
<tr>
<td>$\theta_{\tau 10}$</td>
<td>-0.518</td>
<td>0.682</td>
<td>-0.803</td>
<td>1.000</td>
<td>0.759</td>
</tr>
<tr>
<td>$\theta_{\tau 1500}$</td>
<td>-0.278</td>
<td>0.860</td>
<td>-0.845</td>
<td>0.759</td>
<td>1.000</td>
</tr>
</tbody>
</table>

There is a trade-off between having the correct marginal stratification (as the method of Iman and Conover) and correct correlation matrix. We wish to find out how many of the sampled values are correctly stratified, whether this method will yield reliable estimates and thereby estimate the efficiency of this method.
Chapter V - Uncertainty analysis for pedotransfer functions

5.2.5.2. Normality transform

The method for sampling correlated variables via the LU decomposition is based on the assumption of multivariate normality, but many soil properties, especially soil hydraulic properties and element concentrations, show positively-skewed distributions, e.g. Warrick et al. (1977) approximated hydraulic conductivity by a lognormal distribution. Normality (or approximate normality) can be achieved by a suitable transformation of scale. Carsel and Parrish (1988) applied the so-called Johnson system transformation to the van Genuchten parameters to achieve normality. Another simple transformation, not so frequently used in soil science, is the Box-Cox transformation (Box and Cox, 1964). This is a general method that finds the best transformation to use on a set of data to achieve normality. The Box-Cox transform of variable $x$ is defined as:

$$x_T = \frac{x^\beta - 1}{\beta}, \quad \text{for } \beta \neq 0$$

$$x_T = \ln|x|, \quad \text{for } \beta = 0,$$

[5.2.7]
where $x_T$ is the transformed value of $x$ and $\beta$ is the transformation parameter. When $\beta = 1$ there is no transformation (the data are normal), when $\beta = 0.5$ $x_T = \sqrt{x}$, when $\beta = -1$ $x_T = 1/x$, and when $\beta = 0$ $x_T = \ln|x|$. To find the best value of $\beta$ to achieve normality, we maximize the log-likelihood function, denoted $L(\beta)$:

$$L(\beta) = -\frac{N}{2} \ln(\sigma^2_T) + (\beta - 1) \sum_{j=1}^{N} \ln(x_j),$$  \[5.2.8\]

where $N$ is the number of data, and $\sigma^2_T$ is the variance of transformed $x$ values.

Equivalently, we can minimize the AIC (Akaike’s Information Criterion):

$$\text{AIC} = -2 \left( \max \text{imum log-likelihood of the model} \right) + 2 \left( \text{number of adjustable parameters of the model} \right) \quad \text{(Akaike, 1973)}.$$  \[5.2.9\]

For the univariate case, the number of adjustable parameters ($p$) is 3, i.e. $\mu_T$, $\sigma^2_T$ and $\beta$. Howarth and Earle (1979) proposed minimising the asymmetry of the distribution (skewness, Sk) and also reducing the peakedness (kurtosis, Ku). Their objective function was to minimise:

$$O = W|Sk| + |3 - Ku| \quad \text{[5.2.10]}$$

where $W$ is the weight given to reduce the skewness. Values $L$, AIC or $O$ can be optimized using the non-linear least-squares approach.

For multivariate data a joint distribution that incorporates the correlations among the variables is required. In this instance, joint Box–Cox transformations can be applied (Bozdogan & Ramirez, 1986). For $x_{ij}$, where $i = 1, ..., k$ variables of $j = 1, ..., N$ the transformation using $\beta = (\beta_1, \beta_2, ..., \beta_k)$ is defined as:

$$x_{ij_T} = \begin{cases} \frac{x_{ij}^{\beta_i} - 1}{\beta_i} & \text{for } \beta_i \neq 0 \\ \ln(x_{ij}) & \text{for } \beta_i = 0. \end{cases} \quad \text{[5.2.11]}$$

The transformed data are described by a $k$-variate normal model with variance–covariance matrix $V$. Parameter vector $\beta$ can be found by maximizing the log-likelihood:

$$L(\beta) = -\frac{Nk}{2} \ln(2\pi) - \frac{N}{2} \ln|V| + \sum_{i=1}^{k} (\beta_j - 1) \sum_{j=1}^{N} \ln(x_{ij}) + Nk \quad \text{[5.2.12]}$$

Similarly $\beta$ can also be found by minimizing the AIC:
\[
AIC(\beta) = -2L(\beta) + 2\left(2k + \frac{k(k + 1)}{2}\right). \tag{5.2.13}
\]

In order to obtain the true value from the transformed variables, the inverse Box–Cox transformation can be used:

\[
x = (1 + \beta x_T)^{(1/\beta)}, \quad \text{for } \beta \neq 0
\]
\[
x = \exp(x_T), \quad \text{for } \beta = 0. \tag{5.2.14}
\]

5.2.5.3. Modified sectioning method

Lark (2000) proposed a ‘modified’ sectioning method for sampling correlated variables. The method can be summarized as follows:

- generate random \( k \) multivariate normal correlated variables;
- transform the variables into \( k \) new uncorrelated variables via principal component analysis;
- apply the sectioning method to the distribution of the new variables, i.e. sample from the median of the sections for each variable;
- transform the combinations of the section medians to the original scales.

5.2.5.4. Modified Latin hypercube sampling

A modified Latin hypercube sampling for correlated multivariate data is proposed here as follows:

- plot the histogram of data for each of \( k \) variables, if the histograms appear skewed or peaked, the data may require transformation;
- transform the non-normal data using the joint Box-Cox transformations, find the optimal value of \( \beta \) for each variable;
- analyse the statistics and correlations amongst the transformed variables;
- define \( ns \) equiprobable intervals;
- calculate \( E \), matrix with \( k \) columns of eigenvectors \( e \);
- consider \( k \) principal components \( P \), take \( n \) observations from the distribution of each of the \( k \) components in \( P \) using Latin hypercube sampling; call this \( Q \), and randomly pair each of the \( k \) variables in \( Q \);
- each sampled variables of \( X \) is calculated as:
  \[
x = E^T q + \mu;
\]
• calculate the distribution of $X$, evaluate the percentage of distribution correctly sampled from the prescribed intervals $p_{\text{samp}}$;
• if the variables have been transformed to achieve normality, back-transform the sampled variables using the inverse Box-Cox transformation.

5.3. Methods

Uncertainty in the PTFs evaluated here is the uncertainty due to parameter uncertainty and the measurement error of input variables. The methods for sampling correlated variables: simple random sampling (SRS), Lark’s modified sectioning method (MSM) and the modified Latin hypercube sampling (MLH) were applied to quantify the uncertainty in the output of pedotransfer functions. This is illustrated in the second and third part of this section. Subsequently the PTFs evaluated from the sampling method are used to analyse the uncertainties in a field soil-water simulation model. The procedure is illustrated in Figure 5.3.1.

Figure 5.3.1. Flow chart for the analysis of the uncertainties in PTFs and modelling.
5.3.1. Calculation of the uncertainty in soil hydraulic properties

The uncertainty of the average water-retention curve from measurements at 24 sites in a field was analysed. The data were derived from the study of Bridge (1968) for hydraulic properties at a depth of 15–20 cm in a clayey soil. Water-retention data were fitted to the van Genuchten (1980) model and using the measured saturated hydraulic conductivity \( K_s \) values, unsaturated hydraulic conductivity \( K(h) \) was predicted according to the Mualem-van Genuchten model.

The parameter \( \theta_i \) is held constant at 0. The parameters \( \theta_s, \alpha, n \) and \( K_s \) appear to be non-normal, therefore the joint Box-Cox transformation was applied to them. The mean and standard deviation of the transformed variables were calculated, and correlations were found amongst the parameters (e.g. positive correlations between \( \theta_i \) and \( K_s \); and between \( 1/\alpha \) and \( K_s \)). These values were then sampled using MLH with \( ns = 100 \), then \( \theta(h) \) and \( K(h) \) were evaluated.

5.3.2. PTF to predict soil strength

A pedotransfer function (PTF) for soil mechanical resistance is given by (da Silva and Kay, 1997):

\[
R_s = c_1 \theta^{c_2} \rho^{c_3} \tag{5.3.1}
\]

with the following values for \( c_1, c_2, c_3 \)

\[
c_1 = \exp(-3.67+0.765 \, C_o -0.145 \, \text{clay})
\]

\[
c_2 = -0.481+0.208 \, C_o -0.124 \, \text{clay}
\]

\[
c_3 = 3.85 + 0.0963 \, \text{clay}. \tag{5.3.2}
\]

The function predicts soil penetration resistance \( R_s \) (in MPa) in terms of volumetric water content \( \theta \) and bulk density \( \rho \). The parameters of the function are predicted from percentage of clay and organic carbon (\( C_o \)). The variables \( \theta \) and \( \rho \) are assumed constant at 0.2 m\(^3\) m\(^{-3}\) and 1.35 Mg m\(^{-3}\), and the uncertainty of the input variables clay and organic carbon are evaluated. We assumed that clay and organic carbon are normally distributed and positively correlated (\( R = 0.5 \)), the mean \( \mu \) and standard deviation \( \sigma \) are listed in Table 5.4.1. Sampling from the bivariate normal distribution was done using simple random sampling, modified sectioning method, and modified Latin hypercube sampling.
5.3.3. Uncertainty in hydraulic PTFs

The PTFs for the van Genuchten model for water-retention curves and $K_s$ for Australian soil developed in Chapter IV is as follows:

$$\theta_r = r_1 \text{ clay} + r_2 \text{ PWP}$$

$$\theta_s = s_1 \text{ clay} + s_2 \phi$$

$$\alpha = a_0 + a_1 d_g$$

$$n = n_0 + n_1 \sigma_g$$

$$\ln(K_s) = k_0 + k_1 \ln(\phi_e)$$

[5.3.3]

Effective porosity $\phi_e$ is estimated from porosity minus the PTF water retention at -10 kPa. The $K(h)$ relationship is computed according to the Mualem-van Genuchten model. The effect of parameter uncertainty on the predicted water retention and hydraulic conductivity is analysed using the standard deviation and correlation coefficient obtained through least-squares analysis. The parameter value and its standard error with the correlation amongst the parameters are given in Table 5.3.1.

Table 5.3.1. Parameter for hydraulic PTFs along with standard error and correlation among the parameters.

<table>
<thead>
<tr>
<th>Param.</th>
<th>Value</th>
<th>Std.Err.</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r1</td>
<td>r2</td>
<td>s1</td>
</tr>
<tr>
<td>r1</td>
<td>-0.001</td>
<td>1.000</td>
<td>-0.902</td>
</tr>
<tr>
<td>r2</td>
<td>1.177</td>
<td>0.000</td>
<td>0.016</td>
</tr>
<tr>
<td>s1</td>
<td>0.001</td>
<td>1.000</td>
<td>0.014</td>
</tr>
<tr>
<td>a0</td>
<td>1.562</td>
<td>0.000</td>
<td>0.084</td>
</tr>
<tr>
<td>a1</td>
<td>17.022</td>
<td>1.000</td>
<td>0.024</td>
</tr>
<tr>
<td>n0</td>
<td>1.398</td>
<td>0.001</td>
<td>0.039</td>
</tr>
<tr>
<td>n1</td>
<td>0.003</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>k0</td>
<td>10.052</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>k1</td>
<td>3.662</td>
<td>0.001</td>
<td>0.000</td>
</tr>
</tbody>
</table>

The input variables required are percentage of clay and silt, bulk density ($\rho_b$) and permanent wilting point (PWP). The other variables ($\phi$ = porosity, $d_g$ = geometric mean of particle size, $\sigma_g$ = standard deviation of the mean of particle size) were calculated from the four input variables. Uncertainty in determining particle size is also incorporated with parameter uncertainty. Bulk density and PWP are fixed to a constant value. Two kinds of particle-size measurement are assumed: laboratory measurement and field estimates. The laboratory measurement is more precise than
Chapter V – Uncertainty analysis for pedotransfer functions

the field estimate. The uncertainties for the two methods are based on results from the literature. For example, in texture determination Gee and Bauder (1986) found that for the pipette method, clay could be determined within precision of 1%. Hodgson et al. (1976) found that for field estimates, the 80% confidence limit of clay is within ± 8% and silt is ± 12%. Semivariograms were used to account for spatial variability. The average variogram for clay (McBratney and Pringle, 1999) is:

$$\gamma(\text{lag}) = 6.4 + 17.8(1 - \exp(-\text{lag} / 68.8))$$  [5.3.4]

and for sand:

$$\gamma(\text{lag}) = 5.1 + 18.9(1 - \exp(-\text{lag} / 74.4))$$  [5.3.5]

A clay soil with 50% clay and 35% clay (correlation -0.70) is used for analysis. The standard deviation of clay and sand using laboratory measurement is 0.5%. Using field estimates, the standard deviation is 2.7 and 4.0% for clay and sand, respectively.

5.3.4. PTFs for calculating soil-water storage

The water-retention PTFs are used to calculate the soil-water storage of the soil profile in equilibrium with a groundwater table at depth \( L \) (Hanks, 1992). The value is calculated as:

$$W_L = \int_0^L \theta(h(z))dz$$  [5.3.6]

The depth \( L \) was chosen as 2 m, and the integral was computed numerically using Gaussian quadrature. The input variables required for PTFs are percentage of clay and silt, bulk density (\( \rho_b \)) and permanent wilting point (PWP). Correlations among the variables (Table 2) are based upon the data set used to generate the PTFs. Under the assumption of a normal distribution of the variables (Table 5.3.2), the uncertainty of the input variables was assessed. SRS, MSM and MLH were performed to produce the distribution of \( W_L \).

<table>
<thead>
<tr>
<th>Table 5.3.2. Statistics of the input variables for water retention PTFs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>mu (( \mu ))</td>
</tr>
<tr>
<td>sigma (( \sigma ))</td>
</tr>
<tr>
<td>Correlation matrix:</td>
</tr>
<tr>
<td>clay - silt</td>
</tr>
<tr>
<td>clay - ( \rho )</td>
</tr>
<tr>
<td>clay - PWP</td>
</tr>
<tr>
<td>silt - ( \rho )</td>
</tr>
</tbody>
</table>
5.3.4. Soil-water simulation

Simulations of the field soil-water regime were made using hydraulic PTFs (Equation 5.3.2) as inputs into a mechanistic model SWAP (van Dam et al., 1997). The model is a finite-difference numerical solution of Richards’ equation. Weather data (solar radiation, temperature, wind speed, vapour pressure, and precipitation) from the cotton growing area of Bourke, NSW were used to drive the model. The soil was assumed to be under fallow and evaporation was calculated according to the Penman-Monteith approach (Smith, 1991). The soil profile was assumed uniform to 1 m depth, with an initial matric potential at -33 kPa, and free drainage at the bottom of the profile. Simulation was carried out from April 1997 to December 1998. The first year (April – December 1997) was considered as a ‘warm-up’ period in order to minimise the effect of initial moisture conditions, only results from year 1998 were analysed. Results were recorded every 10 days, a total of 36 observations for 1 year. The outputs evaluated are:

- water content at 2 depths: \( \theta_1 = 50 \) and \( \theta_2 = 300 \) mm;
- water storage in the profile \( (WS) \) up to 0.5 m:
  \[
  WS = \int_{0}^{0.5} \theta(z) \, dz
  \]
- soil moisture deficit \( (SMD) \), which is the difference between the amount of water at the upper limit of field capacity (-10 kPa) and water storage in the field condition. If the soil is wetter than -10 kPa there will be no moisture deficit \( (SMD = 0) \), or
  \[
  SMD = \int_{0}^{0.5} \Delta \theta(z) \, dz, \quad \text{where} \quad \Delta \theta = (\theta_{10} - \theta) \quad \text{when} \quad \theta < \theta_{10}, \quad \text{and} \quad \Delta \theta = 0 \quad \text{when} \quad \theta \geq \theta_{10}
  \]

Uncertainties in the input variables: clay, silt, \( \rho \) and \( PWP \), were evaluated. Correlations among variables are the same as in Table 5.3.2. Both loam and clay soil were considered. Two kinds of uncertainty were evaluated: laboratory measurement and field estimates. The laboratory measurement is more precise than the field estimate (Table 5.3.3). The uncertainties for the two methods are based on results from the literature (the same as §5.4.3). For permanent wilting point measured with the pressure-plate extractor, the coefficient of variation is within 1–2% (Richards, 1965).
Table 5.3.3. Mean and standard deviation of input variables into hydraulic PTFs.

<table>
<thead>
<tr>
<th></th>
<th>clay (dag kg⁻¹)</th>
<th>silt (dag kg⁻¹)</th>
<th>ρᵥ (Mg m⁻³)</th>
<th>PWP (m³ m⁻³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loam</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>μ</td>
<td>20.0</td>
<td>35.0</td>
<td>1.45</td>
<td>0.170</td>
</tr>
<tr>
<td>σ (lab. measurement)</td>
<td>0.5</td>
<td>0.5</td>
<td>0.025</td>
<td>0.005</td>
</tr>
<tr>
<td>σ (field estimate)</td>
<td>2.7</td>
<td>4.0</td>
<td>0.200</td>
<td>0.010</td>
</tr>
<tr>
<td>Clay</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>μ</td>
<td>50.0</td>
<td>15.0</td>
<td>1.35</td>
<td>0.150</td>
</tr>
<tr>
<td>σ (lab. measurement)</td>
<td>0.5</td>
<td>0.5</td>
<td>0.025</td>
<td>0.005</td>
</tr>
<tr>
<td>σ (field estimate)</td>
<td>2.7</td>
<td>4.0</td>
<td>0.100</td>
<td>0.010</td>
</tr>
</tbody>
</table>

5.4. Results and discussion

5.4.1. Calculation of the uncertainty in soil hydraulic properties

Figure 5.4.1a shows the 24 water retention curves measured in a field (Bridge, 1968), and Figures 5.4.1b and 5.4.1c show the median of the water retention and hydraulic conductivity curve along with 10 and 90 percentiles. The area between 10 and 90 percentiles represent 80% prediction limits of the median. The results show the large uncertainties expected when averaging measurements from different sites within a field. The high variation of the hydraulic properties translates into a large uncertainty in the hydraulic conductivity curve.
Figure 5.4.1. (a) Water-retention curves of 24 sites within a field, (b) water-retention, and (b) hydraulic conductivity curve. Dark lines represent the median of prediction and the outer bands represent 10 and 90 percentiles.
5.4.2. Performance of the modified Latin hypercube sampling and uncertainty in soil strength PTF

This section demonstrates the application of the new Latin hypercube sampling method (MLH), and also compares it with the modified sectioning method (MSM), and simple random sampling (SRS). Table 5.4.1 summarises the mean ($\mu$) and standard deviation ($\sigma$) of clay and $C_o$ as sampled by the three methods. Table 5.4.2. shows the standard deviation of the variables using 100 replicates of MLH and SRS at different sample size. Figure 5.4.2 illustrates the sampling of 100 points using the MSM and MLH method. The MSM regularly stratifies the bivariate distribution into $10 \times 10$ sample grids. In the principal component axes (Figure 5.4.2a) and real coordinates (Figure 5.4.2b) the grid has a narrower spacing near the mean and progressively becomes wider as it moves towards the tail. MLH randomly samples from the distribution of the principal component axes (Figure 5.4.2c). MSM and MLH generate samples that are closer to the prescribed joint distribution ($\mu$, $\sigma$ and correlation coefficient, $R$) compared with SRS.

Table 5.4.1. Statistics of variables for predicting soil mechanical resistance.

<table>
<thead>
<tr>
<th>Variables</th>
<th>‘True’</th>
<th>$ns = 50$</th>
<th></th>
<th>$ns = 100$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SRS</td>
<td>MSM</td>
<td>MLH SRS</td>
<td>MSM</td>
</tr>
<tr>
<td>$\mu$ clay</td>
<td>dag kg$^{-1}$</td>
<td>35.0</td>
<td>34.34</td>
<td>35.00</td>
<td>35.00</td>
</tr>
<tr>
<td>$\mu$ $C_o$</td>
<td>dag kg$^{-1}$</td>
<td>1.2</td>
<td>1.175</td>
<td>1.200</td>
<td>1.201</td>
</tr>
<tr>
<td>$\sigma$ clay</td>
<td>dag kg$^{-1}$</td>
<td>2.0</td>
<td>2.111</td>
<td>1.842</td>
<td>1.835</td>
</tr>
<tr>
<td>$\sigma$ $C_o$</td>
<td>dag kg$^{-1}$</td>
<td>0.1</td>
<td>0.110</td>
<td>0.092</td>
<td>0.108</td>
</tr>
<tr>
<td>$R$ clay $\times$ $C_o$</td>
<td></td>
<td>0.5</td>
<td>0.506</td>
<td>0.500</td>
<td>0.501</td>
</tr>
</tbody>
</table>

Figure 5.4.3a shows the efficiency of the MLH method compared with SRS as measured by the variance of the mean predicted $R_s$ using 100 replicates ($\sigma^2 \overline{R_s}$, MLH)/($\sigma^2 \overline{R_s}$, SRS). Although the MLH does not marginally maximally stratify the probabilities of the input variables, it is much better than SRS. The efficiency appears to be highest at $n = 200$, MLH has a variance 120 times lower than SRS. The efficiency decreases exponentially as the sample size increases. The percentage of the intervals correctly sampled decreases with increasing sample size and reaches a plateau around 65%.
The ‘true’ distribution of soil mechanical resistance ($R_s$) was calculated from simple random sampling with 50000 realizations. The probability density of $R_s$ estimated by the three methods at $ns = 50$ is shown in Figure 5.4.4a. The non-parametric probability density function is generated using the density function of S-Plus program (MathSoft, 1999). MSM (using 7 sections, or $ns = 7^2 = 49$) shows the best distribution compared with the true distribution, followed by MLH and SRS.

The distribution of the median as predicted by 100 replicates of $ns = 50$ is shown in Figure 5.4.3b. The SRS has the least precision (wider spread or large variance) and accuracy, the MLH provides a more accurate prediction, but the variance is still large. MSM provides the most precise and accurate prediction.

Figure 5.4.2. Sampling from bivariate normal distribution of clay and organic carbon. Crosses represent 100 observations from joint distribution. (a) The sectioning method applied on the principal components, and (b) after transformation onto the original scale; (c) LHS from principal components, and (d) after transformation onto original scale.
Figure 5.4.3. (a) Efficiency of the modified Latin hypercube sampling compared to simple random sampling, (b) percentage of the intervals correctly sampled by the modified Latin hypercube sampling.

Table 5.4.2. Standard deviation of the mean variables from 100 replicates using simple random sampling and modified Latin hypercube sampling.

<table>
<thead>
<tr>
<th>Variables</th>
<th>MLH</th>
<th>SRS</th>
<th>MLH</th>
<th>SRS</th>
<th>MLH</th>
<th>SRS</th>
<th>MLH</th>
<th>SRS</th>
<th>MLH</th>
<th>SRS</th>
<th>MLH</th>
<th>SRS</th>
<th>MLH</th>
<th>SRS</th>
<th>MLH</th>
<th>SRS</th>
<th>MLH</th>
<th>SRS</th>
<th>MLH</th>
<th>SRS</th>
<th>MLH</th>
<th>SRS</th>
<th>MLH</th>
<th>SRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>clay</td>
<td>0.121</td>
<td>0.068</td>
<td>0.023</td>
<td>0.014</td>
<td>0.012</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
</tr>
<tr>
<td>SRS</td>
<td>0.639</td>
<td>0.471</td>
<td>0.265</td>
<td>0.173</td>
<td>0.157</td>
<td>0.088</td>
<td>0.068</td>
<td>0.060</td>
<td>0.048</td>
<td>0.038</td>
<td>0.025</td>
<td>0.030</td>
<td>0.019</td>
<td>0.014</td>
<td>0.010</td>
<td>0.007</td>
<td>0.005</td>
<td>0.004</td>
<td>0.003</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>$C_o$</td>
<td>0.006</td>
<td>0.003</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>SRS</td>
<td>0.030</td>
<td>0.019</td>
<td>0.014</td>
<td>0.010</td>
<td>0.007</td>
<td>0.005</td>
<td>0.004</td>
<td>0.003</td>
<td>0.003</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>$R$</td>
<td>0.055</td>
<td>0.029</td>
<td>0.012</td>
<td>0.007</td>
<td>0.006</td>
<td>0.005</td>
<td>0.005</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>SRS</td>
<td>0.276</td>
<td>0.180</td>
<td>0.100</td>
<td>0.069</td>
<td>0.054</td>
<td>0.033</td>
<td>0.029</td>
<td>0.022</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
</tr>
<tr>
<td>$R_s$</td>
<td>0.055</td>
<td>0.029</td>
<td>0.012</td>
<td>0.007</td>
<td>0.006</td>
<td>0.005</td>
<td>0.005</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>SRS</td>
<td>0.247</td>
<td>0.180</td>
<td>0.107</td>
<td>0.066</td>
<td>0.062</td>
<td>0.034</td>
<td>0.026</td>
<td>0.024</td>
<td>0.019</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
</tr>
</tbody>
</table>
Figure 5.4.4. (a) Distribution of predicted soil mechanical resistance, (b) Distribution of the median of soil resistance predicted from 100 runs of the three sampling methods. The ‘true’ value is represented by the solid bold lines (—), the value predicted by simple random sampling (dashed lines ----), the sectioning method (longer dashed lines ---) and Latin hypercube sampling (solid line —).
5.4.3. Uncertainty in hydraulic PTFs

The effect of parameter uncertainty on the prediction of water retention curve and hydraulic conductivity is shown in Figure 5.4.5. The parameters of the PTFs are quite certain, as shown by the small standard deviation estimates, thus the uncertainty of prediction in water retention is also quite small. With the parameter uncertainty and holding $\phi$ and PWP constant, results show that variation using laboratory measurement is not large, only a slight deviation from the parameter uncertainty. When using field texture, the estimated uncertainty is very large. The confidence limits using field texture are very large compared with the pipette method. When only field estimates of texture are provided we can expect a large uncertainty in the predicted water content. When using less effort and cost to obtain the information (e.g. using field estimate, or estimating $\phi$ and PWP from other properties), the accuracy of the prediction is sacrificed.

Incorporating parameter uncertainty, error in particle size measurement and spatial variability resulted in error propagation of estimated water retention. Figure 5.4.6. shows the 80% prediction range (90 percentile - 10 percentile) as a function of potential. The relative error is increasing with increasing potential, all the error seem to coalesce at potential range of -3 to -10 m. The uncertainty is quite small at the inflection part of the water retention curve. Results showed that even at short distance ($lag = 10$ m) the uncertainty is higher compared to field estimates. If the variation in field texture is combined with spatial variation, the error in prediction becomes very large. When using PTFs the trade-offs between the cost of obtaining the required input data and prediction quality needs careful consideration.
Figure 5.4.5. Uncertainty in (a) water retention and (b) hydraulic conductivity curves. The dark lines represent the median of predicted curves and the inner bands (dashed lines) are the 10 and 90 percentiles due to parameter uncertainty, second band (solid lines) when using more precise input data, and the outer bands (solid lines) are the confidence limits when using less precise data.
Figure 5.4.6. (a) Prediction range (90 minus 10 percentiles) and (b) relative prediction range for predicted water retention curve.
5.4.4. PTFs for calculating soil-water storage

The ‘true’ distribution of soil-water storage $W_L$, was calculated from simple random sampling of 50000 realizations. Table 5.4.3 shows the statistics of the sampled variables, and of the simulated $W_L$. Generally, as the sample size is increased the performance of SRS and MLH becomes comparable.

Table 5.4.3. Statistics of input properties of PTFs and subsequent prediction of soil-water storage.

<table>
<thead>
<tr>
<th>Variables</th>
<th>‘True’</th>
<th>SRS</th>
<th>MLH</th>
<th>MSM</th>
<th>SRS</th>
<th>MLH</th>
<th>MSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ns$</td>
<td>50000</td>
<td>50</td>
<td>50</td>
<td>16</td>
<td>100</td>
<td>100</td>
<td>81</td>
</tr>
<tr>
<td>$\mu_{\text{clay}}$</td>
<td>dag kg$^{-1}$</td>
<td>20.00</td>
<td>20.54</td>
<td>20.05</td>
<td>20.00</td>
<td>20.18</td>
<td>19.97</td>
</tr>
<tr>
<td>$\mu_{\text{silt}}$</td>
<td>dag kg$^{-1}$</td>
<td>35.00</td>
<td>34.96</td>
<td>34.93</td>
<td>35.00</td>
<td>35.11</td>
<td>35.01</td>
</tr>
<tr>
<td>$\mu_{\rho}$</td>
<td>Mg m$^{-3}$</td>
<td>1.45</td>
<td>1.46</td>
<td>1.45</td>
<td>1.45</td>
<td>1.47</td>
<td>1.45</td>
</tr>
<tr>
<td>$\mu_{\text{PWP}}$</td>
<td>m$^3$ m$^{-3}$</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>$\sigma_{\text{clay}}$</td>
<td>dag kg$^{-1}$</td>
<td>2.70</td>
<td>2.57</td>
<td>2.99</td>
<td>1.88</td>
<td>2.94</td>
<td>2.61</td>
</tr>
<tr>
<td>$\sigma_{\text{silt}}$</td>
<td>dag kg$^{-1}$</td>
<td>4.00</td>
<td>4.39</td>
<td>3.99</td>
<td>2.79</td>
<td>3.97</td>
<td>4.06</td>
</tr>
<tr>
<td>$\sigma_{\rho}$</td>
<td>dag kg$^{-1}$</td>
<td>0.10</td>
<td>0.11</td>
<td>0.10</td>
<td>0.07</td>
<td>0.10</td>
<td>0.09</td>
</tr>
<tr>
<td>$\sigma_{\text{PWP}}$</td>
<td>m$^3$ m$^{-3}$</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>$\mu W_L$</td>
<td>m</td>
<td>4.927</td>
<td>4.896</td>
<td>4.928</td>
<td>4.929</td>
<td>4.891</td>
<td>4.929</td>
</tr>
<tr>
<td>$\sigma W_L$</td>
<td>m</td>
<td>0.276</td>
<td>0.278</td>
<td>0.267</td>
<td>0.192</td>
<td>0.285</td>
<td>0.260</td>
</tr>
</tbody>
</table>

The MSM method is also evaluated in this case, using 2, 3, and 4 sections, resulting in $ns = 16, 81, $ and 256. The distribution of $W_L$ as predicted by SRS, MLH ($ns = 100$) and MSM ($ns = 81$) is shown in Figure 5.4.7a. The SRS and MLH predict similarly, while the values obtained using MSM have a narrower distribution than the
‘true’ values. For two variables, MSM estimates the distribution well. When the number of variables is increased, this method becomes increasingly inefficient as $n_s$ is raised to the power of the number of variables. Although MSM generate samples that have the mean values exactly as the prescribed, the standard deviation is usually smaller. The predicted mean of $W_L$ is close to its true value but the standard deviation is smaller.

To evaluate the accuracy of SRS and MLH in predicting the mean of $W_L$, 100 runs were performed with the 3 sampling methods and the distribution of the predicted mean is shown in Figure 5.4.7b. Although SRS performs similarly as MLH (Figure 5.4.7a), it has the least precision.

Figure 5.4.8 shows the decrease in the $RMSE$ as the sample size is increased. The $RMSE$ of percentile distribution is calculated as

$$RMSE = \sqrt{\frac{1}{1/101} \sum_{i=0}^{100} (q_i - \hat{q}_i)^2}$$

where $q_i$ is the ‘true’ $i$-th percentile and $\hat{q}_i$ is the percentile of the predicted distribution. For the current situation (4 input variables for PTFs) there is indication that $n_s$ should be greater or equal to 100 to give reasonable estimates. To achieve greater accuracy $n_s$ should be greater or equal to 800, but considering the small error in the prediction (0.04 mm) and the limited computation time, $n_s = 100$ could be accepted and further comparisons will be presented in the next section.
Figure 5.4.7. (a) Distribution of predicted soil-water storage, (b) Distribution of the mean of soil-water storage predicted from 100 runs of the three sampling methods. The ‘true’ value is represented by the solid bold lines (—), the value predicted by simple random sampling (dashed lines ----), the sectioning method (longer dashed lines ---) and Latin hypercube sampling (solid line —).
5.4.5. Soil-water simulation

First we look at the influence of sample size on the distribution of simulated results. Figure 5.4.9 shows the median of the simulation outputs on a loam with sample size of 2000, 100 and 50 observations. It shows that 100 observations provide close estimates and a similar trend for the median of water storage, and water content at the surface. Deviations were observed for soil moisture deficit and θ at 300 mm depth. The deviation is greater when there is little rainfall and large evaporation (between 1–120 Julian days and 240–360 Julian days). As will be shown later, the output variables appear to be sensitive to the change in climatic conditions.

Table 5.4.4 shows the RMSE between the simulated values using four sample sizes (50, 100, 200 and 500) with values calculated from 2000 samples. The median, 10 and 90 percentiles show that the errors using 100 samples are still within the acceptable range, the maximum difference of 3 mm water storage and 0.005 m³ m⁻³ θ are smaller than the errors in measurement.
Table 5.4.4. Root Mean Squared Error between the predicted and true parameters from soil-water simulations

<table>
<thead>
<tr>
<th>Sample size</th>
<th>( WS ) (mm)</th>
<th>( SMD ) (mm)</th>
<th>( \theta_1 ) (m(^3) m(^{-3}))</th>
<th>( \theta_2 ) (m(^3) m(^{-3}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>2.358</td>
<td>2.328</td>
<td>0.0015</td>
<td>0.0061</td>
</tr>
<tr>
<td>200</td>
<td>3.128</td>
<td>3.066</td>
<td>0.0014</td>
<td>0.0102</td>
</tr>
<tr>
<td>100</td>
<td>3.300</td>
<td>3.153</td>
<td>0.0030</td>
<td>0.0128</td>
</tr>
<tr>
<td>50</td>
<td>5.795</td>
<td>5.506</td>
<td>0.0049</td>
<td>0.0161</td>
</tr>
<tr>
<td>10% quantile</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>1.274</td>
<td>1.030</td>
<td>0.0010</td>
<td>0.0101</td>
</tr>
<tr>
<td>200</td>
<td>2.497</td>
<td>1.387</td>
<td>0.0019</td>
<td>0.0106</td>
</tr>
<tr>
<td>100</td>
<td>6.259</td>
<td>1.287</td>
<td>0.0162</td>
<td>0.0160</td>
</tr>
<tr>
<td>50</td>
<td>5.168</td>
<td>1.351</td>
<td>0.0137</td>
<td>0.0126</td>
</tr>
<tr>
<td>90% quantile</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>2.387</td>
<td>0.827</td>
<td>0.0125</td>
<td>0.0089</td>
</tr>
<tr>
<td>200</td>
<td>6.167</td>
<td>2.234</td>
<td>0.0072</td>
<td>0.0131</td>
</tr>
<tr>
<td>100</td>
<td>5.821</td>
<td>5.397</td>
<td>0.0123</td>
<td>0.0129</td>
</tr>
<tr>
<td>50</td>
<td>7.087</td>
<td>5.389</td>
<td>0.0183</td>
<td>0.0131</td>
</tr>
</tbody>
</table>
Figure 5.4.9. Median of simulated soil-water regimes: water storage, soil moisture deficit, $\theta$ at 30 mm and $\theta$ at 300 mm. Dark lines are the median using 2000 samples, dotted lines = 100 samples, and solid line = 50 samples.
Figure 5.4.10 shows the median of the estimated water-retention curves and its 10 and 90 percentiles when calculated using the uncertainty of laboratory measurement (inner band) and field estimation (outer band). The median, 10 and 90 percentiles of the hydraulic parameters are given in Table 5.4.5. The water retention and hydraulic conductivity curves show that the uncertainty is very large when using a less precise method for obtaining the data. When using less effort and cost to obtain the inputs (field estimate), accuracy of prediction is sacrificed. For instance, $\theta_s$ could vary from 0.3–0.5 m$^3$ m$^{-3}$ and $K_s$ could vary between 100–1000 mm day$^{-1}$.

Figure 5.4.10. (a) Water-retention, and (b) hydraulic conductivity curves predicted from PTFs. The dark lines represent the median of predicted curves and the inner bands (dashed lines) are the 10 and 90 percentiles when using more precise input data, and the outer bands (solid lines) are the confidence limits when using less precise data.
Table 5.4.5. Predicted hydraulic parameters from uncertainties in input to PTFs

<table>
<thead>
<tr>
<th></th>
<th>$\theta_r$ (m$^3$ m$^{-3}$)</th>
<th>$\theta_s$ (m$^3$ m$^{-3}$)</th>
<th>$\alpha$ (m$^{-1}$)</th>
<th>$n$</th>
<th>$K_s$ (mm day$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lab measurement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>0.182</td>
<td>0.400</td>
<td>2.45</td>
<td>1.443</td>
<td>366.33</td>
</tr>
<tr>
<td>10% quantile</td>
<td>0.174</td>
<td>0.390</td>
<td>2.41</td>
<td>1.442</td>
<td>297.46</td>
</tr>
<tr>
<td>90% quantile</td>
<td>0.189</td>
<td>0.409</td>
<td>2.48</td>
<td>1.444</td>
<td>432.04</td>
</tr>
<tr>
<td><strong>Field measurement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>0.181</td>
<td>0.394</td>
<td>2.43</td>
<td>1.443</td>
<td>366.37</td>
</tr>
<tr>
<td>10% quantile</td>
<td>0.169</td>
<td>0.361</td>
<td>2.25</td>
<td>1.438</td>
<td>190.29</td>
</tr>
<tr>
<td>90% quantile</td>
<td>0.196</td>
<td>0.445</td>
<td>2.69</td>
<td>1.447</td>
<td>645.43</td>
</tr>
</tbody>
</table>

Now we evaluate the effect of the uncertainties in hydraulic properties on soil-water models. Figure 5.4.11 shows the median of the prediction (solid line) along with the 10 and 90 percentiles (lower and upper lines) over time on a loam. The uncertainties in the prediction for field estimate show large uncertainties (wide prediction bands) around the median of prediction. Nevertheless, the laboratory measurement also show large uncertainties, especially for soil moisture deficit. The prediction is more certain between 200 and 240 days, after a large rainfall event and when evaporation was small.

The lower confidence (10 percentile) generally has wider difference from the median and less variable compared to the 90 percentile. This is because the values are bounded by the lower limit of the moisture retention ($\theta$). For SMD, the 90 percentile is very high and flat, this is because the value is bounded by the estimate of $\theta$ at -10 kPa. This demonstrates the sensitivity of the model to the change in hydraulic properties, the lower and upper limit of the moisture availability controls the limits of prediction.

The median of prediction from the field estimate is less wavering compared with the laboratory measurements. Although both methods have the same mean hydraulic properties, the less precise measurement in turn gives a less precise output.

Three days in the year were selected to reflect the distribution of soil-water parameters, namely day 105, the dry period when it has not rained for 65 days; day 203, during a period of large rainfall events; and day 205, 2 days after the large rainfall.

The probability density plots (Figures 5.4.12 to 5.4.14) show the distribution of the simulated parameters. The figures reveal that using more precise inputs yield more precise estimates, which do not appear in the median with confidence limits plots.
(Figure 5.4.11). The model using field estimates exhibits a large variance (spread) compared with the one based on laboratory measurements. During the dry period (Figure 5.4.12, day 105), the outputs show a bimodal distribution, i.e. storage at ~105 mm and ~130 mm. This suggests that the model output is quite uncertain in the dry period, and prone to changes in hydraulic properties. In the wet period (Figure 5.4.13, day 203), there is much less uncertainty with a marked peak at ~190 mm storage, although there is a small peak at ~100 mm. In the soil-water redistribution period (Figure 5.4.14, day 205), the distribution shows least variation, as shown by the single distinct peak at ~160 mm storage. The uncertainties arise due to the limits in evaporation and infiltration set in the model. Numerical solutions employed in SWAP define the weather as flux controlled top boundary condition. However, when the weather is too wet, the maximum infiltration rate is controlled by the height of water ponding. When the soil is under prolonged dry weather the maximum evaporation rate is controlled by the pressure-head of the soil surface in equilibrium with air humidity, and the maximum infiltration rate is controlled by the height of water ponding (van Dam and Feddes, 2000). This condition is set to ensure stability of the numerical solution.

Figure 5.4.15 shows the result of the simulation on a clay under the same climatic conditions. The clay stores more water than the loam, but all the output parameters exhibit similar trends to the loam. The upper boundary conditions for the simulation are the same therefore the trend reflects the climatic conditions. The lower prediction interval (10 percentile) in this instance showed more variation with time compared with the loam.
Figure 5.4.11. Results of the soil-water simulation over time on a loam for (a) water storage, and (b) soil moisture deficit. Continued next page
Figure 5.4.11. Results of the soil-water simulation over time on a loam for (c) $\theta$ at 30mm, (d) $\theta$ at 300 mm. The dark lines represent the median, the lower and upper dashed lines are the 10 and 90 percentiles. The graphs on the right-hand side are the results with more precise input data and the left-hand side is when using less precise data.
Figure 5.4.12. Probability density of simulated parameters during a dry period (Julian day 105). The solid line is the laboratory measurements and dotted lines are for field estimates.
Figure 5.4.13. Probability density of simulated parameters during a wet period (Julian day 203). The solid line is the laboratory measurements and dotted lines are for field estimates.
Figure 5.4.14. Probability density of simulated parameters during water redistibution period (Julian day 205). The solid line is the laboratory measurements and dotted lines are for field estimates.
Figure 5.4.15. Results of the soil-water simulation over time on a clay for (a) water storage, (b) soil moisture deficit. Continued next page.
Figure 5.4.15. Results of the soil-water simulation over time on a clay for (c) $\theta$ at 30mm, (d) $\theta$ at 300 mm. The dark lines represent the median, the lower and upper curves are the 10 and 90 percentiles. The graphs on the right-hand side are the results with more precise input data and the left-hand side is when using less precise data.
5.5. Conclusions

A modified Latin hypercube sampling method for multivariate correlated variables has been presented. The method was applied to characterize the uncertainty in the prediction of soil-water models. Although this modification does not marginally maximally stratified the probability distribution as required by Latin hypercube sampling, it is far more efficient than simple random sampling. The modified stratified sampling (Latin hypercube sampling and sectioning method) also performed better in predicting the distribution of the results and also show better precision. The sectioning method performs best when the number of variables is small (e.g. two variables). As the number of variables is increased, the required sample size is raised exponentially, and the method becomes inefficient. This is important because large number of inputs is usually needed to run simulation models. The MLH can cope with large number of variables and show high accuracy and precision.

Small magnitudes of uncertainty in the input data can translate into a large uncertainty in the prediction. Using more precise inputs also yield better prediction confidence. The boundary conditions and shape of the hydraulic curves has a great influence on the soil-water simulation models. The simulated model exhibited large uncertainties during dry periods and smaller uncertainties during the rainy periods. The soil-water redistribution period exhibits the more certain prediction.
References


Chapter VI – Estimation of Sorptivity from disc permeameter measurements

VI. Estimation of sorptivity from disc permeameter measurements

The disc permeameter, or tension infiltrometer, has been used extensively for the \textit{in-situ} characterization of soil hydraulic properties. Sorptivity ($S_0$) is generally derived from early-time infiltration because at this period, infiltration is dominated by capillary forces within the soil (Cook and Broeren, 1994). Sorptivity is a function of both soil capillarity and conductivity, thus can be related to the soil-water uptake capacity. Different hydraulic properties can be derived from the estimate of $S_0$ such as macroscopic capillary length (White and Sully, 1987) and hydraulic conductivity (White and Perroux, 1989). Determination of sorptivity is an important factor, as it will affect the estimation of other soil hydraulic properties.

The purposes of this study are to investigate the effect of placement of sand on the infiltration from the disc permeameter, as well as to compare different methods for estimating sorptivity and then to evaluate the effect of sand placement on the estimate. This study also investigated the possibility of estimating sorptivity using a small disk permeameter.

6.1. Effect of sand on sorptivity estimates

To assure good contact between the soil surface and the disc membrane, contact (capping) material is usually placed between the soil surface and the disc membrane (Perroux and White, 1988). Perroux and White (1988) recommended the use of porous materials between 3 and 5 mm thick with both high sorptivity and conductivity. The use of materials having the same hydraulic properties as the soil is recommended but usually this is difficult to obtain \textit{a priori} in the field. They suggested a suitable material of fine sand with $S_0 \approx 2$ mm s$^{-1/2}$ and saturated hydraulic conductivity $K_s \approx 3 \times 10^{-2}$ mm s$^{-1}$.

The influence of the layer of contact sand can usually be neglected in determining the steady-state flux as at large times the sand has small influence on the flux. But at early times, especially when the sand is dry, the high sorptivity of the sand may mask the true sorptivity of the soil. Although sand is generally used in disc permeameter
measurements in the field, and sorptivity is calculated from the early times of infiltration (Hussen and Warrick, 1993; Cook and Broeren, 1994), the effect of sand on the estimate has not been rigorously evaluated. Reynolds and Zebchuk (1996) did however find a discrepancy between the potential head applied on the membrane and the soil surface. They reported that the hydraulic gradient between the contact sand and soil surface could be greater than one, which creates a difference in potential head between the supply membrane and soil surface. The discrepancy depends on the thickness and hydraulic properties of the sand. They recommended the use of highly uniform glass sphere medium with $K_s = 0.11 \text{ mm s}^{-1}$ and presented a calculation procedure for correcting the supply potential on top of the soil surface. Close et al. (1998) found a non-uniform wetting pattern of the sand at the membrane interface. This resulted in differences in infiltration rate and variations in infiltrating area and depth of infiltration. This work stresses the need for good contact between the sand and soil.

Alternative methods for estimating $S_0$ include measurement of infiltration using discs with different radii (Scotter et al., 1982; Smettem and Clothier, 1989), or measurement using a single disc at more than one different supply potentials (Ankeny et al., 1991).

### 6.1.1. Deriving sorptivity from infiltration data

Cumulative absorption or desorption into or out of a horizontal column of soil with uniform properties and moisture content is proportional to the square root of time (Talsma, 1969). This proportionality coefficient has been termed sorptivity by Philip (1957), which he described as “a measure of the capacity of the medium to absorb or desorb liquid by capillarity”. Sorptivity is defined analytically as a function of soil water content and diffusivity (Philip and Knight, 1974). The calculation of the real sorptivity involves iterative numerical procedures, because of technical difficulties several approximations have been proposed (Elrick and Robin, 1981; Kutílek and Valentová, 1986). The approximation by Parlange (1975) has been found to give good results (Elrick and Robin, 1981):

$$S_0^2 = 2\sqrt{\theta_o - \theta_n} \int_{\theta_n}^{\theta_o} \sqrt{\theta - \theta_o} D(\theta) \, d\theta$$  \[6.1.1\]
where $\theta_0$ is the water content at applied potential head $h_0$, $\theta_t$ is the initial water content of the soil, and $D$ is the soil diffusivity.

In practice, however, Philip (1969) showed that sorptivity can be measured relatively easy from horizontal infiltration where water flow is only controlled by capillary absorption:

$$I = S_0 \sqrt{t}, \quad [6.1.2]$$

where $I$ is the cumulative infiltration at time $t$. For vertical 3-dimensional infiltration from the disc permeameter, sorptivity can also be determined from early stages of flow where capillary force dominate (White et al., 1992):

$$\lim_{t \to 0} \frac{dI}{d\sqrt{t}} \approx S_0. \quad [6.1.3]$$

Algebraic models for infiltration from the disc permeameter have been developed (see Chapter II). Most of the models were empirical and derived from the one-dimensional infiltration equation and satisfied the condition that cumulative infiltration is proportional to the square-root of time at short times and reaches a steady state rate at long times (Collis-George, 1977). Often models of infiltration have the form of Philip’s two-term equation (Philip, 1957). For one-dimensional infiltration, he proposed,

$$I_{1D} = S_0 \sqrt{t} + A t. \quad [6.1.4]$$

Here $I$ [L] is the cumulative infiltration, $t$ [T] is time, $S$ is the sorptivity [LT$^{-1/2}$] and $A$ [LT$^{-1}$] is a constant. Parameter $A$ can be viewed as the steady-state infiltration rate, as $t \to \infty : \frac{dI}{dt} \approx A$. This equation applies for all times except for very large $t$ (Philip, 1957). For three-dimensional infiltration from a disc permeameter, this equation has been modified as (Haverkamp et al., 1994):

$$I_{3D} = S_0 \sqrt{t} + (A_1 + A_2) t \quad [6.1.5]$$

with $A_1$ and $A_2$ defined as:

$$A_1 = K_n + \left(\frac{2 - \kappa}{3}\right)(K_0 - K_n), \quad A_2 = \frac{\gamma S_0^2}{r_0(\theta_0 - \theta_t)}$$

where $\kappa$ is a parameter related to capillary diffusivity, $\gamma$ is a proportionality constant and $r_0$ is the radius of the disc (see §2.4). From this relationship it can be observed that $S_0$ does not only affect the early stage of infiltration. By replacing $A'$ for $(A_1 + A_2)$ we define the three-dimensional cumulative infiltration as:
\[ I = S_0 \sqrt{t} + A' t \]  

[6.1.6]

and the infiltration rate \( dI / dt \) [LT\(^{-1}\)] is expressed as:

\[ \frac{dI}{dt} = \frac{1}{2} \frac{S_0}{\sqrt{t}} + A' \]  

[6.1.7]

To account for the use of contact sand on infiltration, Vandervaere et al. (1997) modified Equation (4):

\[ I = I_s + S_0 \sqrt{t-t_s} + A'(t-t_s) \]  

[6.1.8]

where \( I_s \) and \( t_s \) are the depth of water and time necessary to ‘wet’ the sand. Since the thickness of the sand is usually quite small, the infiltration into the sand can be approximated by the Green-Ampt infiltration model; \( I_s \) and \( t_s \) can be estimated by (Perroux and White, 1988):

\[ I_s = \Delta \theta_s L_s = S_s \sqrt{t_s} \]  

[6.1.9]

\[ t_s = (I_s / S_s)^2, \]

where \( L_s \), \( \Delta \theta_s \), and \( S_s \) are respectively, the change of thickness, water content and sorptivity of the sand. In order to eliminate the effect of the contact sand on the soil’s infiltration, Vandervaere et al. (1997) differentiated Equation [6.1.6] with respect to \( \sqrt{t} \):

\[ \frac{dI}{d\sqrt{t}} = S_0 \frac{t}{\sqrt{t-t_s}} + 2A' \sqrt{t} \]  

[6.1.10]

The differentiation eliminates the \( I_s \) factor, and as time increases the effect of \( t_s \) will also becomes small.

Sorptivity is usually estimated by fitting equation [6.1.3] or [6.1.4] to measured infiltration data. A variety of methods have been proposed to estimate the parameter \( S_0 \). These are:

(1) Estimate from early-time infiltration data

\[ I = S_0 \sqrt{t}, \]  

[6.1.11]

this involves plotting \( I \) against \( \sqrt{t} \), identifying portion of the graph with straight-line behaviour, and fitting a line (Cook and Broeren, 1994). \( S_0 \) is determined from the slope of the line.

(2) Using the linear least-squares method, i.e. fitting \( I \) as quadratic function of \( \sqrt{t} \)

\[ I = S_0 \sqrt{t} + A'(\sqrt{t})^2 \]  

[6.1.12]
(3) Square-root-of-time transformation on cumulative infiltration (Smiles and Knight, 1976):

\[
\frac{I}{\sqrt{t}} = S_0 + A'\sqrt{t}, \quad [6.1.13]
\]

This requires plotting \( I / \sqrt{t} \) against \( \sqrt{t} \), and identifying the portion of the graph with straight-line behaviour, and fitting a line. \( S_0 \) is determined from the intercept.

(4) Fitting the infiltration rate:

\[
\frac{dI}{dt} = \frac{1}{2} \frac{S_0}{\sqrt{t}} + A', \quad [6.1.14]
\]

Which can be done by plotting \( dI/dt \) against \( 1/\sqrt{t} \) or by using nonlinear least-squares.

(5) Differentiating cumulative infiltration with respect to the square-root of time (Vandervaere et al., 1997):

\[
\frac{dI}{d\sqrt{t}} = S_0 + 2A'\sqrt{t}. \quad [6.1.15]
\]

By plotting \( dI / d\sqrt{t} \) against \( \sqrt{t} \), and excluding the early time data which exhibit nonlinear behaviour, then fitting a line through the data with linear behaviour will give \( S_0 \) as the intercept of the line.

Bristow and Savage (1985) compared methods (2) and (3) to estimate the two-term Philip coefficients from field infiltration experiments. Applying a transformation to the original model introduces self-correlation between the dependent and independent variables which result in differences in parameter estimates. They also discussed the need to distinguish between fitting the data to empirical models, and deriving parameters from physically-based models. This implies that it is important not only to obtain a good fit to the data, but also a good estimate of the parameters.

6.1.2. Materials and Methods

6.1.2.1. Numerical study

To evaluate the effect of sand on infiltration from a disc permeameter, numerical simulation was performed. The water flow into the soil under the disc was simulated by axi-symmetrical two-dimensional water flow (see §2.5) using the finite-element
code SWMS_2D (Šimunek et al., 1994). The van Genuchten (1980) model was used to describe the soil hydraulic properties. The finite-element grid was generated on a 1000 x 1000 mm domain with 50 x 50 nodes in rectangular elements. Along the radial $r$ axis, 20 nodes were equally spaced within the disc radius of 100 mm and the remaining 30 nodes was set up in an arithmetic progression to 1000 mm (Warrick, 1992). A similar grid was arranged along the $z$ axis. The sand was virtually placed on top of the soil with the same radius as the disc having 20 x 2 nodes. The time-step used is 30 seconds and the infiltration is simulated for 3600 seconds. The finite-element discretisation, as shown in Figure 6.1.1, produced a water-balance error in the simulation of less than 0.05% for all time steps.

Two types of soil are considered; a loam (Carsel and Parrish, 1988) and a clay (Warrick, 1992). Following Reynolds and Zebchuk (1996), the contact material (sand) has a $K_s = 0.1$ mm s$^{-1}$. The hydraulic parameters for the contact material (sand), loam and clay are given in Table 6.1.1. The soil is assumed to be uniform with initial potential head ($h_n$) $-5000$ mm, and the sand has initial $\theta_n = 0.05$ m$^3$ m$^{-3}$. For disc infiltration, 3 potential heads ($h_0$) are considered: -20, -60 and -100 mm.

Figure 6.1.1. The finite-element discretisation for simulation of infiltration from the disc permeameter.
Section VI – Estimation of Sorptivity from disc permeameter measurements

Table 6.1.1. Hydraulic properties of the soil used for the numerical study.

<table>
<thead>
<tr>
<th></th>
<th>$\theta_l$</th>
<th>$\theta_s$</th>
<th>$\alpha$</th>
<th>n</th>
<th>$K_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>m$^3$m$^{-3}$</td>
<td>m$^3$m$^{-3}$</td>
<td>m$^{-1}$</td>
<td></td>
<td>mm m$^{-1}$</td>
</tr>
<tr>
<td>Sand</td>
<td>0.000</td>
<td>0.300</td>
<td>2.0</td>
<td>3.00</td>
<td>1.00 $\times 10^{-1}$</td>
</tr>
<tr>
<td>Loam</td>
<td>0.078</td>
<td>0.430</td>
<td>3.6</td>
<td>1.56</td>
<td>2.90 $\times 10^{-3}$</td>
</tr>
<tr>
<td>Clay</td>
<td>0.124</td>
<td>0.495</td>
<td>1.5</td>
<td>2.00</td>
<td>1.23 $\times 10^{-4}$</td>
</tr>
</tbody>
</table>

6.1.2.2. Field Study

An infiltration study using disc permeameter was conducted at the Australian Cotton Research Institute (149.59°E, 30.20°S), Narrabri, New South Wales. The soil is a Black Vertosol (Isbell, 1996) with heavy clay texture (51% clay, 27.5% silt, 21.5% sand) and $K_s$ of 0.001 mm m$^{-1}$. The part of the field under study had been cultivated periodically over the past 5 years. A plot of 1.6 m x 1.0 m was chosen, and tilled manually in order to produce uniform-sized aggregates and a smooth surface. The experiment setup is given in Figure 6.1.2.

The plot was setup as a randomized complete block experiment with 2 blocks. Each block consists of 3 replicates of 2 treatments: placement of contact sand and no contact sand. The initial water content of the soil was $\theta_n = 0.042 \pm 0.005$ m$^3$ m$^{-3}$ measured with the time-domain reflectometer (TDR).

![Figure 6.1.2. The randomized complete-block experimental setup in the field.](image-url)
The CSIRO disc permeameter with disc radius $r_0 = 100$ mm was used. For treatment with sand, a retaining ring of 200 mm diameter and 5 mm thick was placed on the soil surface and the sand with thickness around 5 mm was placed in the ring and leveled. The retaining ring was then removed. The sand used is a commercial fine-sand with $K_s = 0.05$ mm s$^{-1}$ and $S_0 = 2.23$ mm s$^{-1/2}$. The disc was set to supply potential of $-20$ mm and placed on the soil surface, the cumulative infiltration was then monitored at time steps of 10-30 seconds. The infiltration was carried out for 20 minutes. The final water content was $\theta_0 = 0.301 \pm 0.05$ m$^3$ m$^{-3}$ with bulk density $1.12 \pm 0.1$ Mg m$^{-3}$.

6.1.2.3. Laboratory determination of sorptivity

Soil aggregates from the experimental plot were collected for determination of sorptivity in the laboratory. Horizontal infiltration was carried out to determine sorptivity. The soil is packed into a perspex column which is comprised of small sections of 30 mm diameter tube joint by waterproof adhesive tape. A short column section, which was filled with sand retained by nylon mesh, was joined at the end of the column and connected to a mariotte bottle. The Mariotte bottle was set at a constant head of $-20$ mm relative to the axis of the column. The sorptivity was obtained by fitting the rate of infiltration to equation 6.1.8. The mean bulk density of the soil packed into the column was $1.11 \pm 0.08$ Mg m$^{-3}$.

6.1.2.4. Data Analysis

The cumulative infiltration data obtained from the numerical simulation and field study were analysed to obtain the estimate for sorptivity. Five methods of fitting the data were used, i.e.:

1. $I = S_0 \sqrt{t}$
2. $I = S_0 \sqrt{t} + A (\sqrt{t})^2$
3. $\frac{I}{\sqrt{t}} = S_0 + A' \sqrt{t}$
4. $\frac{dl}{dt} = \frac{1}{2} \frac{S_0}{\sqrt{t}} + A'$
5. $\frac{dl}{d\sqrt{t}} = S_0 + 2 A' \sqrt{t}$
Chapter VI – Estimation of Sorptivity from disc permeameter measurements

The infiltration rate \( \frac{dI}{dt} \) is approximated by central difference:

\[
\frac{dI}{dt} \approx \frac{\Delta I}{\Delta t} = \frac{I_{t+1} - I_{t-1}}{t_{t+1} - t_{t-1}}
\]

and the same approximation is also applied for \( dI / d\sqrt{t} \). The fitting of Methods (1), (3) and (5) were done by selecting part of the graph that is linear and fitting a straight line through the data. While methods (2) and (4) were done by fitting the equation to the whole data.

For numerical examples, the true sorptivity of the soil was calculated from soil water content and diffusivity using equation 6.1.1. The integration was numerically computed using the adaptive Gaussian quadrature algorithm (Kahaner et al., 1989).

**6.1.3. Results and discussions**

**6.1.3.1. Numerical study**

The effect of sand placement on the cumulative infiltration is shown in Figure 6.1.3a. The sand increased the cumulative amount of water infiltrating into the loam and clay. However, the sand has only a little effect on the infiltration rate (Figure 6.1.3b). Initially, as water infiltrates into the sand, the rate will be considerably higher compared to the soil without sand. As the water saturates the sand, \( t_s = 12 \) seconds for 5 mm of sand, the difference in rate is small and the sand does not affect the steady-state infiltration rate. But for the cumulative infiltration, because it is the sum of the amount of water infiltrating or the integral of infiltration rate, the effect will also be cumulative. Applying the correction for sand to the infiltration curve (Equations 6.1.8 and 6.1.9), it was found that the corrected infiltration curve did not always stay equidistant from the infiltration without the sand (control). The difference in infiltration between the control (without sand) and sand increased with time. Therefore the correction for sand as proposed by Vandervaere et al. (1997) only holds for early-time infiltration.

Results of sorptivity estimation using different methods (Table 6.1.2) showed that by placement of sand the estimate of \( S_0 \) was always higher than the control. Of all the methods, method 1 always overestimates the soil's sorptivity, either with or without placement of sand. It can be noted that for 3-dimensional infiltration, not only does
sorptivity influence infiltration at the early times. For the control treatment, fitting the cumulative infiltration curve as a quadratic function of the square-root of time, method (2), provides a close estimate to the real sorptivity of the soil. But with the placement of sand, the cumulative infiltration increases and therefore $S_0$ will also be overestimated. Dividing the cumulative infiltration by the square-root of time yields a linear relationship with the square-root of time. This transformation presents lower estimates than Methods (1) and (2), but does not improve the estimate of $S_0$ for treatment with sand.

Figure 6.1.3. Results of water infiltration from the numerical study: (a) cumulative infiltration, and (b) infiltration rate for loam and clay.
Table 6.1.2. Estimates of sorptivity using different methods, from the numerical study.

<table>
<thead>
<tr>
<th>Soil</th>
<th>Treatment</th>
<th>$h_0$ (mm)</th>
<th>True Sorptivity (mm s$^{-1/2}$)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>loam</td>
<td>no sand</td>
<td>-20</td>
<td>0.288</td>
<td>0.376</td>
<td>0.306</td>
<td>0.300</td>
<td>0.313</td>
<td>0.289</td>
</tr>
<tr>
<td>loam</td>
<td>no sand</td>
<td>-60</td>
<td>0.230</td>
<td>0.292</td>
<td>0.237</td>
<td>0.238</td>
<td>0.252</td>
<td>0.232</td>
</tr>
<tr>
<td>loam</td>
<td>no sand</td>
<td>-100</td>
<td>0.188</td>
<td>0.232</td>
<td>0.197</td>
<td>0.192</td>
<td>0.224</td>
<td>0.192</td>
</tr>
<tr>
<td>loam</td>
<td>sand 2mm</td>
<td>-20</td>
<td>0.288</td>
<td>0.387</td>
<td>0.354</td>
<td>0.346</td>
<td>0.344</td>
<td>0.298</td>
</tr>
<tr>
<td>loam</td>
<td>sand 2mm</td>
<td>-60</td>
<td>0.230</td>
<td>0.294</td>
<td>0.286</td>
<td>0.276</td>
<td>0.269</td>
<td>0.237</td>
</tr>
<tr>
<td>loam</td>
<td>sand 2mm</td>
<td>-100</td>
<td>0.188</td>
<td>0.234</td>
<td>0.244</td>
<td>0.229</td>
<td>0.226</td>
<td>0.193</td>
</tr>
<tr>
<td>loam</td>
<td>sand 5mm</td>
<td>-20</td>
<td>0.288</td>
<td>0.376</td>
<td>0.392</td>
<td>0.300</td>
<td>0.330</td>
<td>0.292</td>
</tr>
<tr>
<td>loam</td>
<td>sand 5mm</td>
<td>-60</td>
<td>0.230</td>
<td>0.307</td>
<td>0.328</td>
<td>0.305</td>
<td>0.281</td>
<td>0.247</td>
</tr>
<tr>
<td>loam</td>
<td>sand 5mm</td>
<td>-100</td>
<td>0.188</td>
<td>0.244</td>
<td>0.280</td>
<td>0.259</td>
<td>0.230</td>
<td>0.198</td>
</tr>
<tr>
<td>clay</td>
<td>no sand</td>
<td>-20</td>
<td>0.137</td>
<td>0.160</td>
<td>0.139</td>
<td>0.136</td>
<td>0.159</td>
<td>0.145</td>
</tr>
<tr>
<td>clay</td>
<td>no sand</td>
<td>-60</td>
<td>0.126</td>
<td>0.147</td>
<td>0.125</td>
<td>0.126</td>
<td>0.142</td>
<td>0.133</td>
</tr>
<tr>
<td>clay</td>
<td>no sand</td>
<td>-100</td>
<td>0.116</td>
<td>0.137</td>
<td>0.114</td>
<td>0.115</td>
<td>0.133</td>
<td>0.121</td>
</tr>
<tr>
<td>clay</td>
<td>sand 2mm</td>
<td>-20</td>
<td>0.137</td>
<td>0.163</td>
<td>0.190</td>
<td>0.184</td>
<td>0.163</td>
<td>0.142</td>
</tr>
<tr>
<td>clay</td>
<td>sand 2mm</td>
<td>-60</td>
<td>0.126</td>
<td>0.149</td>
<td>0.179</td>
<td>0.174</td>
<td>0.143</td>
<td>0.131</td>
</tr>
<tr>
<td>clay</td>
<td>sand 2mm</td>
<td>-100</td>
<td>0.116</td>
<td>0.139</td>
<td>0.170</td>
<td>0.161</td>
<td>0.140</td>
<td>0.124</td>
</tr>
<tr>
<td>clay</td>
<td>sand 5mm</td>
<td>-20</td>
<td>0.137</td>
<td>0.165</td>
<td>0.227</td>
<td>0.215</td>
<td>0.158</td>
<td>0.143</td>
</tr>
<tr>
<td>clay</td>
<td>sand 5mm</td>
<td>-60</td>
<td>0.126</td>
<td>0.151</td>
<td>0.215</td>
<td>0.202</td>
<td>0.144</td>
<td>0.133</td>
</tr>
<tr>
<td>clay</td>
<td>sand 5mm</td>
<td>-100</td>
<td>0.116</td>
<td>0.134</td>
<td>0.204</td>
<td>0.188</td>
<td>0.138</td>
<td>0.122</td>
</tr>
<tr>
<td>sand</td>
<td></td>
<td>-20</td>
<td>0.3639</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sand</td>
<td></td>
<td>-60</td>
<td>0.3358</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sand</td>
<td></td>
<td>-100</td>
<td>0.3059</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As the infiltration rate for the soil with or without placement of sand is similar, there is only a slight difference in the estimate of $S_0$. Differentiating the cumulative infiltration with respect to the square-root of time, Method (5), provides the closest estimate to the true sorptivity value. This method also produces consistent estimates with or without sand. The need to differentiate the cumulative infiltration in order to eliminate the effect of the initial water infiltrating into the sand is demonstrated in this case.

Vandervaere et al. (1997) suggested that the effect of the sand could result in the non-linear behaviour of the data when the infiltration data is transformed as in methods 3 and 5. While this transformation may reveal outliers from the early-time infiltration, this non-linear part may not always be attributed to the effect of sand.
Numerical results without placement of sand also revealed the non-linearity on the first few data points (Figure 6.1.4). This is because at early times the rate of infiltration does not follow the rate of square-root of time.

![Figure 6.1.4. The nonlinear behaviour of the first data points as a result of the square-root transformation of infiltration curves from the numerical study.](image)

6.1.3.2. Field experiment

The results of infiltration in the field are shown in Figures 6.1.5 – 6.1.8. The infiltration curves presented here are the average from 6 replicates. From Figures 6.1.4 we can see the behaviour of the infiltration in the field. The standard deviation ($\sigma$) of the cumulative infiltration is increasing with time following the pattern of the infiltration curve. This shows that the error follows the rate of infiltration. When infiltration starts, the error will be large initially as water level from the reservoir drops rapidly. If we assume that there is a measurement error during infiltration (e.g. error in reading the reservoir level), the variance should be constant with time. The cumulative effect of the error could be due to small scale spatial variability.

As expected, infiltration without sand is slower than with sand but the error is larger. The error is due to poor contact between the supply membrane and soil surface. The non-uniform wetting of the soil led to a high variation in infiltration curves. Close et al. (1998) reported as a result of non-uniform wetting for a sandy
clay loam, the maximum difference in infiltration is 80% and the infiltration area varied from 25 to 100%.

Figure 6.1.5. The cumulative infiltration for treatments with, and without sand, and associated standard deviation ($\sigma$) from the field experiment.
Figure 6.1.6. The infiltration rate for treatments with, and without sand, and associated standard deviation (σ) from the field experiment.
Figure 6.1.7. The square-root transformation of cumulative infiltration for treatments with, and without sand, and associated standard deviation ($\sigma$) from the field experiment.
Figure 6.1.8. The differentiation of infiltration with respect to the square-root of time for treatments with, and without sand, and associated standard deviation (σ) from the field experiment.
Table 6.1.3. Estimates of sorptivity using different methods, from the field experiment.

<table>
<thead>
<tr>
<th>Plot</th>
<th>Treatment</th>
<th>Sorptivity (mm s$^{-1/2}$)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>plot 1</td>
<td>sand</td>
<td>1.235</td>
<td>0.589</td>
<td>0.551</td>
<td>0.703</td>
<td>0.665</td>
<td></td>
</tr>
<tr>
<td>plot 4</td>
<td>sand</td>
<td>1.439</td>
<td>1.041</td>
<td>1.072</td>
<td>1.339</td>
<td>0.887</td>
<td></td>
</tr>
<tr>
<td>plot 6</td>
<td>sand</td>
<td>1.052</td>
<td>0.285</td>
<td>0.314</td>
<td>0.472</td>
<td>0.622</td>
<td></td>
</tr>
<tr>
<td>plot 7</td>
<td>sand</td>
<td>1.669</td>
<td>1.207</td>
<td>0.337</td>
<td>1.269</td>
<td>1.087</td>
<td></td>
</tr>
<tr>
<td>plot 8</td>
<td>sand</td>
<td>1.462</td>
<td>1.087</td>
<td>1.104</td>
<td>1.316</td>
<td>1.052</td>
<td></td>
</tr>
<tr>
<td>plot 9</td>
<td>sand</td>
<td>1.388</td>
<td>1.168</td>
<td>1.180</td>
<td>1.398</td>
<td>0.998</td>
<td></td>
</tr>
<tr>
<td>average</td>
<td></td>
<td>1.374</td>
<td>0.896</td>
<td>0.760</td>
<td>1.083</td>
<td>0.885</td>
<td></td>
</tr>
<tr>
<td>s.d.</td>
<td>(σ)</td>
<td>0.211</td>
<td>0.373</td>
<td>0.403</td>
<td>0.393</td>
<td>0.199</td>
<td></td>
</tr>
<tr>
<td>Weighted LS</td>
<td></td>
<td>1.372</td>
<td>0.866</td>
<td>0.905</td>
<td>0.817</td>
<td>1.050</td>
<td></td>
</tr>
<tr>
<td>s.e. WLS</td>
<td></td>
<td>0.016</td>
<td>0.016</td>
<td>0.006</td>
<td>0.077</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>plot 2</td>
<td>no sand</td>
<td>0.895</td>
<td>0.243</td>
<td>0.265</td>
<td>0.368</td>
<td>0.196</td>
<td></td>
</tr>
<tr>
<td>plot 3</td>
<td>no sand</td>
<td>1.385</td>
<td>0.594</td>
<td>0.647</td>
<td>0.976</td>
<td>0.955</td>
<td></td>
</tr>
<tr>
<td>plot 5</td>
<td>no sand</td>
<td>1.215</td>
<td>0.218</td>
<td>0.578</td>
<td>0.393</td>
<td>0.486</td>
<td></td>
</tr>
<tr>
<td>plot 10</td>
<td>no sand</td>
<td>1.245</td>
<td>0.780</td>
<td>0.646</td>
<td>0.864</td>
<td>1.019</td>
<td></td>
</tr>
<tr>
<td>plot 11</td>
<td>no sand</td>
<td>0.388</td>
<td>0.093</td>
<td>0.048</td>
<td>0.650</td>
<td>0.531</td>
<td></td>
</tr>
<tr>
<td>plot 12</td>
<td>no sand</td>
<td>0.703</td>
<td>0.152</td>
<td>0.196</td>
<td>0.181</td>
<td>0.186</td>
<td></td>
</tr>
<tr>
<td>average</td>
<td></td>
<td>0.972</td>
<td>0.347</td>
<td>0.396</td>
<td>0.572</td>
<td>0.562</td>
<td></td>
</tr>
<tr>
<td>s.d.</td>
<td>(σ)</td>
<td>0.380</td>
<td>0.275</td>
<td>0.260</td>
<td>0.310</td>
<td>0.359</td>
<td></td>
</tr>
<tr>
<td>Weighted LS</td>
<td></td>
<td>0.819</td>
<td>0.305</td>
<td>0.305</td>
<td>0.438</td>
<td>0.303</td>
<td></td>
</tr>
<tr>
<td>s.e. WLS</td>
<td></td>
<td>0.054</td>
<td>0.011</td>
<td>0.011</td>
<td>0.036</td>
<td>0.043</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.1.3 shows the estimate of $S_0$ using different methods. Analysis of variance (Figure 6.1.9) showed no significant differences between the methods for estimating $S_0$ except for Method (1) which apparently overestimates the value. Method (1) produces the largest error because water infiltrates rapidly into the soil in the first 100 seconds and the uncertainty is high. The estimate of $S_0$ from the field experiment showed different patterns from the numerical examples. The error in data is larger in the field, therefore when transforming or differentiating the infiltration data, the results may not be as consistent as the numerical examples. The standard deviation ($σ$) plot of methods (2) to (5) present insight on the error of infiltration with time. When using cumulative infiltration ($I$ and $I/\sqrt{t}$), the error increases with time, whilst the error in $dI/dt$ and $dI/d\sqrt{t}$ is large for the first 100 seconds and then becomes constant.
Figure 6.1.9. Analysis of variance for $S_0$ estimate from field experiment using different methods for treatments (a) without sand and (b) with sand. The means diamond illustrates the mean with 95% confidence interval. The circles represent the least significant differences for each pair of the methods.

Sorptivity obtained in the laboratory from the average from 4 replicates is 1.06 mm s$^{-1/2}$ with a standard deviation 0.87 mm s$^{-1/2}$. The high standard deviation is due to the use of soil aggregates in the column rather than sieved soil. The $S_0$ obtained in the lab is within the range of estimates from methods (2)-(5) for treatment with sand. The estimate of $S_0$ from the soil without sand is considerably smaller with larger error. Using the average curve we could estimate the most representative sorptivity by using a weighted least-squares, using reciprocal of the measurement variance $1/\sigma^2$ as the weight. The sorptivity of the soil according to method (2) is 0.866 mm s$^{-1/2}$, which is similar to the laboratory measurement. The method that provide close estimates are (2), (4) and (5).
For a single infiltration measurement, hydraulic conductivity \((K_0)\) is usually calculated based on Wooding's (1968) analysis combined with macroscopic capillary length theory of White and Sully (1987):

\[
K_0 = q_\infty - \frac{4bS^2}{(\theta_0 - \theta_s)\pi r_0}
\]  
[6.1.16]

where \(q_\infty\) is the steady-state infiltration rate, \(b\) is a shape factor for the soil-water diffusivity function which is usually taken as 0.55. Error analysis of the determination of \(K_0\) from the estimate \(S_0\) was performed. If \(S_0\) was estimated from method (5), which had a normal distribution with average and standard deviation of 0.885 and 0.2 mm s\(^{-1/2}\). Assuming the \(b\) factor is constant and equal to 0.55, the difference in water content was 0.26 m\(^3\) m\(^{-3}\) and the error in determination of \(\theta\) from TDR measurement was negligible compared to the other factors. The infiltration rate at the end of the experiment was taken as \(q_\infty\) and assumed to have a log-normal distribution with mean and standard deviation 0.05 and 0.01 mm s\(^{-1}\). Values of \(S_0\) and \(q_\infty\) from the prescribed mean and standard deviation were sampled 10000 times from their distribution and values of \(K_0\) were calculated from Equation [6.1.16]. It was found that the \(K_0\) calculated ranges from -0.029 to 0.089 mm s\(^{-1}\). The result indicated the high sensitivity of equation 6.4.1 to the change in \(S_0\).

6.1.4. Conclusions

The placement of sand has a large effect on the cumulative infiltration and therefore the estimate of \(S_0\). The effect of sand is usually neglected and conventionally \(S_0\) is estimated from the early-time infiltration. This can result in an overestimate of \(S_0\) and may explain why negative values of \(K_0\) can sometimes be obtained in practice when using equation 6.4.1. It is also shown that equation 6.4.1 is quite sensitive to a change in \(S_0\) hence the calculation of \(K_0\) needs to be assessed.

The placement of sand is required to ensure good contact between the supply membrane and the soil surface. The use of pre-wetting the sand may assist in minimising the effect of sand on the infiltration curves. However, in areas where the surface is rough or uneven or an undisturbed condition is desired, a good depth of sand is essential to provide a good contact. The estimation of sorptivity and hydraulic conductivity based on the methods discussed above is inadequate, the use of steady-state infiltration rate is indispensable in this case. Sorptivity and hydraulic
conductivity can be determined by analysis using the steady-state infiltration rate through methods of multiple tension (Ankeny et al., 1991) or multiple radii discs (Smettem and Clothier, 1989).

The results have an important implication in inverse modelling of cumulative infiltration data from disc permeameter measurement (Šimunek and van Genuchten, 1996). The placement of sand could result in inconsistent estimation of the hydraulic model parameters. The cumulative effect of the error in the cumulative infiltration can also result in high uncertainties in the parameter estimates.

6.2. Estimation of sorptivity using disk with small radius

6.2.1. Theory

The Mini Disk® infiltrometer (Figure 6.2.1) was developed by Decagon Devices (Pullman, Washington, USA, http://www.decagon.com). Using the same principle as the sorptivity tube of Clothier and White (1981), the applied tension is controlled by a small suction control tube near the base of the disc. The manufacturer recommends the use of the empirical relationship of Zhang (1997) to evaluate the hydraulic conductivity.

![Diagram of Mini Disk® infiltrometer](image)

Figure 6.2.1. The Mini Disk® infiltrometer.
Chapter VI – Estimation of Sorptivity from disc permeameter measurements

Gee and Ward (1999) tested the Mini Disk® to determine $K_s$ of a sand and a silt loam. From 9 to 12 minutes measurement they reckoned the estimated $K_s$ is comparable with independent laboratory measurements. They suggested that the speed and simplicity of this measurement could be used for intensive characterization of spatial variability of soil hydraulic properties in the field. The small area measured by the Mini Disk® could introduce more variability than using a disc with a much larger diameter however.

Talsma (1969) proposed measurement of sorptivity in the field from short-time ring infiltrometer data. Smith and Buchleiter (1999) measured the sorptivity in the field by using a method proposed by Smith (1999). The method consists of inserting a 100 mm diameter steel tube into the soil, pouring a volume of water equivalent to 1 cm depth ($I_v$) into the tube, and recording the time needed for water to imbibe until half of the soil surface is not ponded ($t_v$). Sorptivity is calculated from $S_0 = I_v / t_v^{1/2}$. This is a rough estimate and only measures sorptivity at saturation.

While studies have been done on the use of different-sized disc radii to calculate hydraulic conductivity (Wang et al., 1998), not much attention has been given to the early stage infiltration from a small-radius disc except for the theoretical analysis by Philip (1968, 1985).

Wooding’s analytical solution showed that the steady-state rate from a disk decreases with increasing disc radius:

$$ q^* = 1 + \frac{4}{\pi r^*} $$

[6.2.1]

where $q^* = q_\infty / (\pi r_0^2 K_0)$ is the dimensionless steady-state infiltration rate and $r^* = r / \lambda_c$ is the dimensionless pond radius (see § 2.2.2). For a small pond with $r^* < 0.8$, Weir (1987) refined the solution as:

$$ q^* = \frac{4 / r^* \sin^2 (r^*/2)}{(\pi / 2) r^* \sin (r^*/2) \cos (r^*/2) + r^* \sin^2 (r^*/2) \ln (r^*/2) - 1.073 (r^*/2)^2} $$

[6.2.2]

Wooding’s numerical analysis (Wooding, 1968) showed that using small disk radius the capillary effect towards the radial coordinate is bigger (Figure 6.2.3). As the radius increases the gravity influence increases and the direction of the flow tends to move downwards.

273
Figure 6.2.2. Relationship between dimensionless radius \( r^* \) and steady-state rate \( q^* \) according to Wooding’s solution.

Figure 6.2.3. Distribution of normalized flux density on the radial and vertical axis with varying disk size (Wooding, 1968).

Turner and Parlange (1974) analysed infiltration without gravity in a two-dimensional system (Figure 6.2.4) showing that \( z_1 \), the wetting front at which \( \theta \) changed from \( \theta_0 \) to \( \theta_n \) towards the \( z \) axis, can be described as:

\[
z_1 = \frac{S_0 \sqrt{t}}{(\theta_0 - \theta_n)}
\]  

[6.2.3]
Smettem et al. (1994) extended the analysis for the lateral two-dimensional wetting front \( r_1 \) as:

\[
r_1 = \frac{S_u \sqrt{t}}{4 \nu^2 \left( \theta_0 - \theta_s \right)^2},
\]

where \( \nu \) is a dimensionless scaling parameter.

### 6.2.2. Materials and methods

#### 6.2.2.1. Numerical study

A numerical study was performed to evaluate the effect of different disk radii on the estimation of sorptivity. The hypothetical loam (Table 6.1.1) was considered again under infiltration with disk of 15, 20, 30, 50 and 100 mm radii. The quasi three-dimensional axi-symmetric infiltration was calculated using the program SWMS_2D.

#### 6.2.2.2. Laboratory study

In the laboratory, a Lansdowne clay loam was packed to a density of 1.40 Mg m\(^{-3}\). The Mini Disk\textsuperscript{®} with 16 mm diameter and -20 mm supply potential was used. Sixteen replicates were made by recording the amount of water infiltrating every 20 seconds for six minutes. The lateral spreading of wetting front was measured using a ruler with 1 mm accuracy. Sorptivity was estimated by fitting the infiltration data to Philip’s two-term model, Equation [6.1.6]. Sorptivity was also determined using the horizontal infiltration procedure as described in §6.1.2.3.
6.2.3. Results and discussion

6.2.3.1. Numerical study

Figure 6.2.5 shows the water content distribution after one hour of infiltration under different disc radii. Using a small disc, the lateral spreading relative to the disc radii is large, with increasing disc radii the gravity effects dominate. The cumulative infiltration (Figure 6.2.6) shows that the small disc has much higher infiltration rate than the large disc. The large capillary effect increases the relative area wetted by the small disc.

Figure 6.2.5. Distribution of normalized water content $\theta/\theta_0$ after 1 hour of infiltration over the radial and vertical axis on a loam with different size disk radii. Numbers beside the curve represents the normalized water content $\theta/\theta_0$. 
By fitting the Philip’s two-term, Equation [6.1.6], sorptivity $S_0$ and steady-state infiltration rate $A'$ were estimated for the different disc radii (Table 6.2.1). As shown by Wooding’s solution infiltration rate decreases as disc radius increases (Figure 6.2.6). Sorptivity does not vary much with different disc radii. Although the value is higher than their true values it still provides a reasonable estimate. Thus a disc with a small radius can be used to estimate sorptivity.

Table 6.2.1. The effect of different disc radii on the estimation on sorptivity and steady-state infiltration rate.

<table>
<thead>
<tr>
<th>$r_0$ mm</th>
<th>-20 mm</th>
<th></th>
<th>-60 mm</th>
<th></th>
<th>-100 mm</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S_0$ mm s$^{-1/2}$</td>
<td>$A'$ mm s$^{-1}$</td>
<td>$S_0$ mm s$^{-1/2}$</td>
<td>$A'$ mm s$^{-1}$</td>
<td>$S_0$ mm s$^{-1/2}$</td>
<td>$A'$ mm s$^{-1}$</td>
</tr>
<tr>
<td>15</td>
<td>0.315</td>
<td>0.0140</td>
<td>0.252</td>
<td>0.0097</td>
<td>0.202</td>
<td>0.0064</td>
</tr>
<tr>
<td>20</td>
<td>0.313</td>
<td>0.0105</td>
<td>0.248</td>
<td>0.0073</td>
<td>0.202</td>
<td>0.0047</td>
</tr>
<tr>
<td>30</td>
<td>0.311</td>
<td>0.0074</td>
<td>0.249</td>
<td>0.0048</td>
<td>0.203</td>
<td>0.0032</td>
</tr>
<tr>
<td>50</td>
<td>0.315</td>
<td>0.0045</td>
<td>0.253</td>
<td>0.0029</td>
<td>0.207</td>
<td>0.0019</td>
</tr>
<tr>
<td>100</td>
<td>0.322</td>
<td>0.0024</td>
<td>0.258</td>
<td>0.0014</td>
<td>0.211</td>
<td>0.0009</td>
</tr>
<tr>
<td>True</td>
<td>0.288</td>
<td>0.230</td>
<td>0.188</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 6.2.6. Cumulative infiltration of the hypothetical loam under -20 mm tension with different disc radii.
6.2.3.2. Laboratory study

The sixteen infiltration curves on the Lansdowne clay loam using the mini disc with disk radius of 16 mm is shown in Figure 6.2.8. Under a uniform soil condition the infiltration shows a large variation. Because the area of the infiltration source is quite small, the small-scale (aggregates < 1 mm) soil variation caused different wetting pattern. This produced variable sorptivity estimates (Figure 6.2.9).

The mean $S_0$ from sixteen measurements is 2.03 mm s$^{-1/2}$ with standard deviation 0.67 mm s$^{-1/2}$. Fitting Philip’s two-term equation to the mean infiltration curves weighted by the inverse of the variance $1/\sigma^2$, the estimate is 1.91 mm s$^{-1/2}$ with standard error 0.029. The $S_0$ from the horizontal infiltration is 1.87 mm s$^{-1/2}$.

As a guide, based on Student’s $t$ distribution, to acquire a level $L$ of accuracy the number of samples required is $t_\alpha^2 / \sigma^2 / L^2$, with $t_\alpha$ is the Student’s $t$ at $\alpha$ probability level. For 95% confidence intervals 10 samples have the accuracy of 0.45 mm s$^{-1/2}$, 8 samples is 0.5, and 4 samples is 0.7 mm s$^{-1/2}$. Increasing the disk size may reduce the variance, four measurements of short-period infiltration (~5 minutes) may provide a good estimate of the soil’s sorptivity.
Chapter VI – Estimation of Sorptivity from disc permeameter measurements

Figure 6.2.8. Cumulative infiltration of Lansdowne clay loam measured by the mini disk.

Figure 6.2.9. Histogram of the sorptivity estimated using the mini disk.

The lateral wetting front as a function of time is given in Figure 6.2.10, as predicted by the theory it follows the square-root of time. An empirical equation is formulated for the lateral wetting front movement:

\[
\frac{r}{r_0} = 1 + \frac{\nu S_0}{(\theta_n - \theta_0)} \sqrt{t}
\]

where \( \nu \) [L\(^{-1}\)] is an empirical coefficient related to disc size. The larger the disk size, the smaller the value of \( \nu \). Equation [6.2.4] is not used as it predicts a sharp change of soil’s water content from \( \theta_n \) to \( \theta_0 \) in a two-dimensional system and does not take into
account the radius of the disc. From the experiment, the average \( \theta_n = 0.08 \), \( \theta_0 = 0.39 \) therefore \( \nu = 0.017 \text{ mm}^{-1} \).

\[
\frac{r}{r_0} = 1 + 0.108 \sqrt{t}
\]

Figure 6.2.10. Lateral movement of wetting front as a function of time.

6.2.4. Conclusions

A small radius disc permeameter can be used to measure sorptivity. Increasing disk radius enhances the gravity effect and reduces the infiltration rate into the soil. Lateral spreading of water from the disk follows the square root of time and is related to the radius of the disk. The smaller the disc size the wider the lateral water spreading. The effect of disc radius has to be taken into account.
Chapter VI – Estimation of Sorptivity from disc permeameter measurements

References


Appendix 6.1.

Infiltration data from the randomised block field experiment

<table>
<thead>
<tr>
<th>Plot Block</th>
<th>Treatment</th>
<th>p1</th>
<th>p4</th>
<th>p6</th>
<th>p7</th>
<th>p8</th>
<th>p9</th>
<th>μ</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>p1</td>
<td>two sand</td>
<td>0.140</td>
<td>0.358</td>
<td>0.067</td>
<td>0.358</td>
<td>0.325</td>
<td>0.319</td>
<td>0.261</td>
<td>0.125</td>
</tr>
<tr>
<td>p4</td>
<td>one sand</td>
<td>0.269</td>
<td>0.504</td>
<td>0.174</td>
<td>0.605</td>
<td>0.521</td>
<td>0.498</td>
<td>0.428</td>
<td>0.168</td>
</tr>
<tr>
<td>p6</td>
<td>two sand</td>
<td>0.392</td>
<td>0.594</td>
<td>0.258</td>
<td>0.762</td>
<td>0.672</td>
<td>0.650</td>
<td>0.554</td>
<td>0.191</td>
</tr>
<tr>
<td>p7</td>
<td>one sand</td>
<td>0.504</td>
<td>0.784</td>
<td>0.347</td>
<td>0.902</td>
<td>0.795</td>
<td>0.767</td>
<td>0.683</td>
<td>0.211</td>
</tr>
<tr>
<td>p8</td>
<td>one sand</td>
<td>0.588</td>
<td>0.896</td>
<td>0.414</td>
<td>1.042</td>
<td>0.896</td>
<td>0.862</td>
<td>0.783</td>
<td>0.233</td>
</tr>
<tr>
<td>p9</td>
<td>two sand</td>
<td>0.683</td>
<td>0.991</td>
<td>0.521</td>
<td>1.126</td>
<td>1.002</td>
<td>0.958</td>
<td>0.880</td>
<td>0.229</td>
</tr>
<tr>
<td>t (s)</td>
<td></td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>40</td>
<td>50</td>
<td>60</td>
<td>80</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>0.140</td>
<td>0.167</td>
<td>0.358</td>
<td>0.325</td>
<td>0.319</td>
<td>0.261</td>
<td>0.125</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.269</td>
<td>0.504</td>
<td>0.174</td>
<td>0.605</td>
<td>0.521</td>
<td>0.498</td>
<td>0.428</td>
<td>0.168</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.392</td>
<td>0.594</td>
<td>0.258</td>
<td>0.762</td>
<td>0.672</td>
<td>0.650</td>
<td>0.554</td>
<td>0.191</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>0.504</td>
<td>0.784</td>
<td>0.347</td>
<td>0.902</td>
<td>0.795</td>
<td>0.767</td>
<td>0.683</td>
<td>0.211</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.588</td>
<td>0.896</td>
<td>0.414</td>
<td>1.042</td>
<td>0.896</td>
<td>0.862</td>
<td>0.783</td>
<td>0.233</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>0.683</td>
<td>0.991</td>
<td>0.521</td>
<td>1.126</td>
<td>1.002</td>
<td>0.958</td>
<td>0.880</td>
<td>0.229</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>0.851</td>
<td>1.165</td>
<td>0.683</td>
<td>1.333</td>
<td>1.170</td>
<td>1.120</td>
<td>1.054</td>
<td>0.239</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.997</td>
<td>1.316</td>
<td>0.829</td>
<td>1.506</td>
<td>1.327</td>
<td>1.254</td>
<td>1.205</td>
<td>0.247</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>1.142</td>
<td>1.462</td>
<td>0.980</td>
<td>1.708</td>
<td>1.490</td>
<td>1.383</td>
<td>1.361</td>
<td>0.261</td>
<td></td>
</tr>
<tr>
<td>150</td>
<td>1.378</td>
<td>1.674</td>
<td>1.187</td>
<td>1.932</td>
<td>1.686</td>
<td>1.557</td>
<td>1.569</td>
<td>0.260</td>
<td></td>
</tr>
<tr>
<td>180</td>
<td>1.540</td>
<td>1.870</td>
<td>1.389</td>
<td>2.162</td>
<td>1.882</td>
<td>1.708</td>
<td>1.758</td>
<td>0.275</td>
<td></td>
</tr>
<tr>
<td>210</td>
<td>1.736</td>
<td>2.078</td>
<td>1.585</td>
<td>2.391</td>
<td>2.061</td>
<td>1.854</td>
<td>1.951</td>
<td>0.287</td>
<td></td>
</tr>
<tr>
<td>240</td>
<td>1.932</td>
<td>2.206</td>
<td>1.781</td>
<td>2.604</td>
<td>2.240</td>
<td>1.988</td>
<td>2.125</td>
<td>0.291</td>
<td></td>
</tr>
<tr>
<td>270</td>
<td>2.117</td>
<td>2.414</td>
<td>1.960</td>
<td>2.806</td>
<td>2.408</td>
<td>2.117</td>
<td>2.303</td>
<td>0.304</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>2.290</td>
<td>2.587</td>
<td>2.145</td>
<td>3.002</td>
<td>2.582</td>
<td>2.240</td>
<td>2.474</td>
<td>0.316</td>
<td></td>
</tr>
<tr>
<td>360</td>
<td>2.649</td>
<td>2.918</td>
<td>2.498</td>
<td>3.394</td>
<td>2.901</td>
<td>2.470</td>
<td>2.805</td>
<td>0.346</td>
<td></td>
</tr>
<tr>
<td>420</td>
<td>3.030</td>
<td>3.237</td>
<td>2.839</td>
<td>3.746</td>
<td>3.209</td>
<td>2.694</td>
<td>3.126</td>
<td>0.369</td>
<td></td>
</tr>
<tr>
<td>480</td>
<td>3.360</td>
<td>3.556</td>
<td>3.175</td>
<td>4.127</td>
<td>3.522</td>
<td>2.912</td>
<td>3.442</td>
<td>0.412</td>
<td></td>
</tr>
<tr>
<td>540</td>
<td>3.777</td>
<td>3.870</td>
<td>3.517</td>
<td>4.486</td>
<td>3.814</td>
<td>3.119</td>
<td>3.697</td>
<td>0.476</td>
<td></td>
</tr>
<tr>
<td>600</td>
<td>4.060</td>
<td>4.172</td>
<td>3.842</td>
<td>4.850</td>
<td>4.105</td>
<td>3.321</td>
<td>4.058</td>
<td>0.496</td>
<td></td>
</tr>
</tbody>
</table>
Infiltration data from the randomised block field experiment (continued)

<table>
<thead>
<tr>
<th>t (s)</th>
<th>Plot</th>
<th>p10</th>
<th>p11</th>
<th>p12</th>
<th>p2</th>
<th>p3</th>
<th>p5</th>
<th>μ</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>two</td>
<td>no sand</td>
<td>0.280</td>
<td>0.073</td>
<td>0.078</td>
<td>0.101</td>
<td>0.151</td>
<td>0.112</td>
<td>0.133</td>
</tr>
<tr>
<td>20</td>
<td>one</td>
<td>no sand</td>
<td>0.336</td>
<td>0.118</td>
<td>0.146</td>
<td>0.196</td>
<td>0.297</td>
<td>0.224</td>
<td>0.219</td>
</tr>
<tr>
<td>30</td>
<td>two</td>
<td>no sand</td>
<td>0.426</td>
<td>0.134</td>
<td>0.196</td>
<td>0.274</td>
<td>0.431</td>
<td>0.319</td>
<td>0.297</td>
</tr>
<tr>
<td>40</td>
<td>one</td>
<td>no sand</td>
<td>0.532</td>
<td>0.157</td>
<td>0.258</td>
<td>0.342</td>
<td>0.549</td>
<td>0.403</td>
<td>0.373</td>
</tr>
<tr>
<td>50</td>
<td>one</td>
<td>no sand</td>
<td>0.739</td>
<td>0.185</td>
<td>0.308</td>
<td>0.403</td>
<td>0.655</td>
<td>0.493</td>
<td>0.464</td>
</tr>
<tr>
<td>60</td>
<td>two</td>
<td>no sand</td>
<td>0.874</td>
<td>0.224</td>
<td>0.358</td>
<td>0.476</td>
<td>0.750</td>
<td>0.588</td>
<td>0.545</td>
</tr>
<tr>
<td>80</td>
<td>one</td>
<td>no sand</td>
<td>0.946</td>
<td>0.286</td>
<td>0.459</td>
<td>0.594</td>
<td>0.924</td>
<td>0.762</td>
<td>0.662</td>
</tr>
<tr>
<td>100</td>
<td>two</td>
<td>no sand</td>
<td>1.030</td>
<td>0.342</td>
<td>0.560</td>
<td>0.717</td>
<td>1.092</td>
<td>0.952</td>
<td>0.782</td>
</tr>
<tr>
<td>120</td>
<td>one</td>
<td>no sand</td>
<td>1.176</td>
<td>0.465</td>
<td>0.650</td>
<td>0.834</td>
<td>1.260</td>
<td>1.120</td>
<td>0.917</td>
</tr>
<tr>
<td>150</td>
<td>one</td>
<td>no sand</td>
<td>1.372</td>
<td>0.588</td>
<td>0.795</td>
<td>0.991</td>
<td>1.523</td>
<td>1.350</td>
<td>1.103</td>
</tr>
<tr>
<td>180</td>
<td>two</td>
<td>no sand</td>
<td>1.568</td>
<td>0.706</td>
<td>0.946</td>
<td>1.148</td>
<td>1.646</td>
<td>1.590</td>
<td>1.267</td>
</tr>
<tr>
<td>210</td>
<td>one</td>
<td>no sand</td>
<td>1.747</td>
<td>0.818</td>
<td>1.081</td>
<td>1.305</td>
<td>1.842</td>
<td>1.809</td>
<td>1.434</td>
</tr>
<tr>
<td>240</td>
<td>one</td>
<td>no sand</td>
<td>1.910</td>
<td>0.924</td>
<td>1.221</td>
<td>1.456</td>
<td>2.027</td>
<td>2.027</td>
<td>1.594</td>
</tr>
<tr>
<td>270</td>
<td>two</td>
<td>no sand</td>
<td>2.072</td>
<td>1.025</td>
<td>1.350</td>
<td>1.602</td>
<td>2.212</td>
<td>2.240</td>
<td>1.750</td>
</tr>
<tr>
<td>300</td>
<td>one</td>
<td>no sand</td>
<td>2.229</td>
<td>1.126</td>
<td>1.490</td>
<td>1.758</td>
<td>2.397</td>
<td>2.453</td>
<td>1.909</td>
</tr>
<tr>
<td>360</td>
<td>one</td>
<td>no sand</td>
<td>2.520</td>
<td>1.310</td>
<td>1.742</td>
<td>2.078</td>
<td>2.750</td>
<td>2.867</td>
<td>2.211</td>
</tr>
<tr>
<td>420</td>
<td>two</td>
<td>no sand</td>
<td>2.750</td>
<td>1.501</td>
<td>1.999</td>
<td>2.391</td>
<td>3.136</td>
<td>3.276</td>
<td>2.509</td>
</tr>
<tr>
<td>480</td>
<td>one</td>
<td>no sand</td>
<td>3.074</td>
<td>1.686</td>
<td>2.274</td>
<td>2.705</td>
<td>3.444</td>
<td>3.674</td>
<td>2.809</td>
</tr>
<tr>
<td>540</td>
<td>one</td>
<td>no sand</td>
<td>3.343</td>
<td>1.865</td>
<td>2.520</td>
<td>3.030</td>
<td>3.786</td>
<td>4.077</td>
<td>3.103</td>
</tr>
<tr>
<td>600</td>
<td>two</td>
<td>no sand</td>
<td>3.606</td>
<td>2.038</td>
<td>2.834</td>
<td>3.360</td>
<td>4.413</td>
<td>5.034</td>
<td>3.548</td>
</tr>
</tbody>
</table>
Appendix 6.2.

Particle size distribution of the Black Vertosol from Narrabri

<table>
<thead>
<tr>
<th>Particle size mm</th>
<th>Material dag/kg</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000-1000</td>
<td>0.38</td>
</tr>
<tr>
<td>1000-500</td>
<td>0.92</td>
</tr>
<tr>
<td>500-250</td>
<td>1.83</td>
</tr>
<tr>
<td>250-100</td>
<td>4.24</td>
</tr>
<tr>
<td>100-53</td>
<td>4.06</td>
</tr>
<tr>
<td>&lt;50</td>
<td>10.10</td>
</tr>
<tr>
<td>&lt;20</td>
<td>27.51</td>
</tr>
<tr>
<td>&lt;2</td>
<td>50.96</td>
</tr>
</tbody>
</table>
VII. Inverse method for predicting soil hydraulic properties from disc permeameter data

7.1. Introduction

The use of the inverse method for predicting soil hydraulic properties has become popular. Different kinds of laboratory and field measurements have been investigated for the possibility of applying the inverse solution. One of the most popular devices in measuring field infiltration is the disc permeameter. The data collected from the disc permeameter measurement are mainly used to derive near saturated hydraulic conductivity. Šimunek and associates (1996, 1997, 1998a, 1998b, 1999) have presented numerical studies and field experiments which suggest the possibility of applying an inverse solution to cumulative infiltration data. Recently, Šimunek and van Genuchten (2000) published a public domain program DISC for estimating water retention and hydraulic conductivity curves from disc permeameter data. Their numerical studies have suggested the type of data needed to provide a unique solution. The cumulative infiltration curve from a single tension infiltration experiment does not appear to hold enough information to yield a unique inverse solution. However, cumulative infiltration from consecutive supply tensions may enable a unique solution. They have also tested the predicted hydraulic properties from the inverse solution of field-infiltration data against laboratory measured properties but found that the water retention curve does not match at all. They argued that the properties estimated from the inverse solution may be more useful in describing field water infiltration, but did not resolve the problems satisfactorily (Šimunek et al., 1999).

There are still many questions regarding the plausibility of using an inverse solution on disc permeameter data. The field study reported in Chapter 6 showed that the error in infiltration data has a cumulative effect, and placement of sand could affect the infiltration data. This study was therefore initiated to evaluate the practicality of the inverse solution from disc permeameter data considering the error in field measurements and to seek a better solution than the one reported in the literature.
7.2. Materials and methods

7.2.1. Inverse solution

Water flow from the disc permeameter is computed using the axi-symmetric two-dimensional formulation of Richards’ equation:

\[
\frac{\partial \theta}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left( rK \frac{\partial h}{\partial r} \right) + \frac{\partial}{\partial z} \left( K \frac{\partial h}{\partial z} \right) - \frac{\partial K}{\partial z} \tag{7.2.1}
\]

where \( \theta \) is the volumetric water content \([L^3 L^{-3}]\), \( h \) is soil potential head \( K \) is the hydraulic conductivity \([L T^{-1}]\), \( r \) is the horizontal (radial) coordinate and \( z \) is depth (vertical coordinate) \([L]\). The initial and boundary conditions for water flow into the soil under the disc with radius \( r_0 \) is given by (Warrick, 1992):

\[
\begin{align*}
\theta(r, z, t) &= \theta_n & r \geq 0, z \geq 0, t = 0 & \text{- initial soil water content} \\
h(r, z, t) &= h_0(t) & 0 < r < r_0, z \geq 0, t > 0 & \text{- constant head at soil surface within the disc radius} \\
\frac{\partial h}{\partial r} &= 0 & z = 0, r > r_0 & \text{- no flow across the soil surface} \\
h(r, z, t) &= h_n & r^2 + z^2 \to \infty & \text{- no flow across the domain} \\
\end{align*}
\tag{7.2.2}
\]

Equations [7.2.1] and [7.2.2] can be solved using the finite-element method. The resulting solution is then coupled with an optimization procedure to yield an inverse solution.

Initially, the finite element Fortran code WARRICK (Warrick, 1992) (kindly supplied by Dr. Art. Warrick, University of Arizona) was combined with the Levenberg-Marquardt nonlinear least-squares algorithm to provide an inverse solution to disc permeameter infiltration data (description of the programs used in this chapter is given in Appendix 7.1). We called this program KCIRRAW, but no convergence could be achieved even when perfect data were fed into the optimization process. Analysis of the objective function response surface (Figure 7.2.1) revealed that the true minimum is hidden within the contours of large objective function values, suggesting, a stable forward solution is needed for a reliable inverse solution.

The two-dimensional unsaturated flow model SWMS_2D version 1.21 (Šimunek et al., 1994, downloaded from the US Salinity Laboratory web-site) was then used. The finite-element grid was generated on a 1000 x 1000 mm domain with 40 x 40 nodes in rectangular elements (similar setup as in Figure 6.3.1). Along the radial \( r \)
axis, 15 nodes were spaced equally within the disc radius $r_0$ and the remaining 30
nodes were placed on in arithmetic progression to 1000 mm (Warrick, 1992). A
similar grid was arranged along the $z$ axis. This element discretisation produced a
water-balance error in the simulation less than 0.1% for all times.

Figure 7.2.1 Response surface of the objective function as a function of parameter $\alpha$
and $n$ produced by the disc permeameter solution of WARRICK.

Infiltration rate was calculated as the sum of the fluxes in the finite-element
nodes associated with the prescribed constant head boundary conditions (Šimunek and
van Genuchten, 1997). The cumulative infiltration $I(t) [L]$ was calculated as the sum
of the fluxes ($q [L^3]$) over time normalised by the area of the disc:

$$I(t) = \frac{1}{\pi r_0^2} \int_{t_0}^{t} q(t) dt$$  [7.2.3]

The modified SWMS_2D code with rectangular element discretization was checked
against the triangular discretization of Warrick (1992) and Šimunek and van
Genuchten (2000). All three programs produced similar results for the cumulative
infiltration data.
The modified Fortran code SWMS_2D was merged with the Levenberg-Marquardt optimization routine (Marquardt, 1963) (see Appendix 3.1) to produce an inverse solution. We called this program DISC\(^\text{\textregistered}\). The objective function to be minimised is:

\[
O(\beta) = \sum_{k=1}^{s} \sum_{i=1}^{N_k} w_i [q(t_i) - \hat{q}(t_i, \beta)]^2
\]

[7.2.4]

where \(s\) is the number of measurement sets (e.g. cumulative infiltration, water content underneath the disc), \(w\) is the weight for data, \(q(t)\) is the measured infiltration data at time \(t\) and \(\hat{q}(t, \beta)\) is the predicted infiltration data under parameter vector \(\beta\). The infiltration can be fitted in terms of cumulative infiltration \(I\) [L], infiltration rate \(dI/dt\) [L T\(^{-1}\)] or water content beneath the disc \(\theta\) [L\(^3\) L\(^{-3}\)]. Cumulative infiltration is generally collected in the field, and infiltration rate can be calculated using finite-difference approximation (\(\Delta I/\Delta t\)) of cumulative infiltration data:

\[
\frac{dI}{dt} \approx \frac{\Delta I}{\Delta t} = \frac{I_{t+1} - I_{t-1}}{t_{t+1} - t_{t-1}}.
\]

The inverse solution of program DISC was compared (\(\ddot{\text{S}}\)imunek and van Genuchten, 2000) (kindly provided by Dr. Jirka \(\ddot{\text{S}}\)imunek, US Salinity Laboratory). Details of the programs used and developed are given in Appendix 7.1.

The sensitivity of the solution to a change in 1% of parameter values was evaluated as:

\[
S_{ij} = \left| q_i (\beta_j + \Delta \beta_j) - q_i (\beta_j) \right|
\]

[7.2.5]

\(S_{ij}\) represents the change of the measurement variable \(q_i\) relative to a small change \(\Delta\) of the parameter \(\beta_j\).

Cook's \(D\) (Cook and Weisberg, 1982) was calculated to identify influential observations during the infiltration procedure. Cook's \(D\) calculates the influential observation as (Fox, 1991):

Influence on parameter = Discrepancy \(\times\) Leverage

Influential observations are not only measured by the residual (discrepancy), but also the extent to which observation \(q_i\) can affect the estimated \(\hat{q}_i\). Detailed calculations can be found in Cook and Weisberg (1982). Although there are many threshold values proposed to identify the influential points, it is generally more useful to examine
observations with unusual values (Fox, 1991). Observation with a high Cook’s $D$ relative to the others reflects the relatively influential points.

### 7.2.2. Wooding’s solution for multiple tensions

Wooding’s solution for steady-state infiltration rate $q_{\infty}$ [L T$^{-1}$] under the disc is given by (Wooding, 1968):

$$q_{\infty} = \alpha_g \phi_0 + \frac{4 \phi_0}{\pi r_0}$$  \[7.2.6\]

where $\phi_0$ is the matrix flux potential and $\alpha_g$ is the sorptive number (White and Sully, 1987), or the parameter in Gardner’s hydraulic conductivity relationship (Gardner, 1958):

$$K(h) = K_s \exp(\alpha_g h)$$

$\phi_0$ can be written as:

$$\phi_0 = \alpha_g^{-1} [K(h_0) - K(h_n)].$$  \[7.2.7\]

When the soil is initially dry, $K(h_n)$ is very small and can be discarded.

If the multiple increasing applied tension were carried out, the hydraulic conductivity in the middle of two successively applied tensions can be estimated. Assuming a piecewise Gardner type $K(h)$ function between two consecutive tensions (Reynolds and Elrick, 1991; Jarvis and Messing, 1995). The average tensions between two consecutive applied tensions is defined as $\ddot{h}$:

$$\ddot{h} = \frac{h_i + h_{i+1}}{2} \text{ for } i = 1, ..., (p - 1) \text{ applied tensions.}$$

The geometric mean infiltration rate:

$$q(\ddot{h}) = \exp \left[ \frac{\ln(q_{i+1}) + \ln(q_i)}{2} \right] = \sqrt{q_{i+1} \times q_i}. $$  \[7.2.8\]

Parameter $\alpha_g$ for the interval between two-consecutive tensions is calculated as:

$$\alpha_g(\ddot{h}) = \frac{\ln(q_{i+1}/q_i)}{(h_{i+1} - h_i)}.$$  \[7.2.9\]

the hydraulic conductivity is given by:

$$K(\ddot{h}) = \frac{q(\ddot{h})}{1 + \frac{4}{\pi \sigma_0 \alpha_g(\ddot{h})}}.$$  \[7.2.10\]
The saturated hydraulic conductivity can be estimated from the calculated $\alpha_g$ and $K$ values:

$$K_s(h) = \frac{K(h)}{\exp(\alpha_g(h)h)}$$  \[7.2.11\]

### 7.2.3. Numerical study

#### 7.2.3.1. Hypothetical soil

Two types of soil are assumed for this study, a loam (Carsel and Parrish, 1988) and clay (Warrick, 1992). Water retention and hydraulic conductivity curves are represented by the van Genuchten model (Table 7.2.1). The loam has been used by Šimunek and van Genuchten (1997) to study the inverse method for disc permeameter. They showed that multiple tension data with initial condition defined in terms of water content provided enough information to yield a unique inverse solution.

<table>
<thead>
<tr>
<th></th>
<th>$\theta_i$</th>
<th>$\theta_s$</th>
<th>$\alpha$</th>
<th>$n$</th>
<th>$K_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loam</td>
<td>0.078</td>
<td>0.430</td>
<td>3.60</td>
<td>1.56</td>
<td>2.90 x 10^3</td>
</tr>
<tr>
<td>Clay</td>
<td>0.124</td>
<td>0.495</td>
<td>1.50</td>
<td>2.00</td>
<td>1.23 x 10^4</td>
</tr>
</tbody>
</table>

It was assumed that a disc with radius of 100 mm was used under three consecutive tensions during three hours of infiltration, at 0 to 3600 seconds: -200 mm; 3600 to 7200 seconds: -100 mm; and 7200 to 10800 seconds: -3.0 mm. Initial soil water content is defined in terms of water content and was set at potential of -500 kPa ($\theta = 0.148$ m$^3$ m$^{-3}$ for loam and 0.173 m$^3$ m$^{-3}$ for clay).

The ‘true’ infiltration curves were generated, then the data was selected at 60 seconds interval during initial infiltration and 100 to 200 seconds afterwards, resulting in 60 observation points within 3 hours infiltration. To account for measurement error, random numbers were added to the infiltration data. The CSIRO disc permeameter has a reservoir reading scale with 2 mm increments. Each mm of the reservoir scale represents 0.056 mm of water, therefore the reading of 2 mm has the error ± 0.11 mm of water. Normally distributed random numbers with mean zero and standard error of
0.11 mm were added to each time step of the cumulative infiltration data. Other sources of measurement error are measurement of initial soil water content; error in estimation of residual and saturated water content as well as saturated hydraulic conductivity; and error in applied tension from the disc permeameter. Water content was varied ± 0.05 m$^3$ m$^{-3}$ and applied tension was varied ± 5 mm. To further complicate the problem, a layer of 5 mm dry sand (hydraulic properties the same as given in Table 6.3.1 $\theta = 0.05$ m$^3$ m$^{-3}$) is placed between the disc and soil surface (similar setup as §6.3). Four types of error in input data was coded as follows:

(0) no error;
(1) error in infiltration data and initial water content;
(2) error in infiltration data, initial water content, and applied tension;
(3) placement of sand and error in infiltration data, initial water content, and applied tension.

7.2.3.2. The HAPEX-Sahel data

The Hydrologic Atmospheric Pilot EXperiment (HAPEX) in the Sahel (climatic region bordering the Southern Sahara from Senegal to Somalia) provided disc permeameter measurements at the Central Site East during summers of 1992 and 1993 (Cuenca et al., 1997). The first data, supplied with the program DISC, consist of multiple tension measurements from a 250 mm diameter disc, and a layer of sand (~30 mm) was placed between the disc and soil surface. The soil is a loamy sand with initial water content at 0.149 m$^3$ m$^{-3}$. The surface crust was removed and infiltration test was carried out under six consecutive tensions. Small breaks occurred during the 107.5 minutes infiltration due to refilling of the permeameter reservoir and adjustment of new supply potential. The final water content underneath the disc is 0.275 m$^3$ m$^{-3}$ with bulk density 1.794 Mg m$^{-3}$. Šimunek et al. (1998) analysed this data with hydraulic conductivity calculated from the steady-state infiltration rate using Wooding’s analysis with piecewise exponential model (Reynold and Elrick, 1990).

The second data set is taken from the HAPEX SAHEL Information System website: http://www.orstom.fr/hapex/. The data was collected at the WAB fallow Monteny location with five tensions applied -100,-70,-250,-20 and -10 mm during 70 minutes of infiltration with a short break during the change in applied tension due to refilling of the permeameter water reservoir. The soil is a sandy loam, similar to the
first data, with initial and final water content underneath the disc of 0.15 and 0.29 m$^3$ m$^{-3}$.

### 7.2.4. Actual infiltration data

#### 7.2.4.1. Mount Annan Silty loam

Field data from the soil from Mount Annan were analysed. The soil is a Red Chromosol (Isbell, 1996) with silty loam texture (clay = 13%, silt = 47%, sand = 40%). The infiltration was carried out using the CSIRO disc permeameter with a disc diameter of 200 mm. Three multiple tension infiltration experiments were carried out within 10 m radius. The average bulk density of the soil surface is 1.19 Mg m$^{-3}$ with initial moisture content 0.17 m$^3$ m$^{-3}$. The soil is under pasture which was trimmed prior to the experiment, a layer of sand was placed on the soil surface to aid contact between the disc membrane and soil surface.

The water retention curve was measured from samples collected in the field. For potential greater than -1 m, the air-dried samples was placed in a sintered funnel and water was imbibed from the base of the funnel. For potentials drier than -1 m, the filter paper method (Fawcett and Collis-George, 1967) was used. The sieved soil is gradually wetted to cover a range of moisture content, and three Whatman 42 filter papers were ‘sandwiched’ between the wetted soil. The filter papers were allowed to equilibrate with the soil for a minimum period of one week and the final water content of the filter paper and soil were determined gravimetrically. The potential of the soil is determined from the potential of the filter paper using the calibration curve of Deka et al. (1995)

#### 7.2.4.2. Marinya clay

Multiple tension disc permeameter measurements were also carried out in Marinya farm (Wee Waa, NSW). The soil, a Black Vertosol (Isbell, 1996), had been cultivated but was fallow at the time of measurements. Three multiple tension infiltrations were carried out within 10 meters radius. The average soil surface bulk density is 1.30 Mg m$^{-3}$ with initial water content 0.22 m$^3$ m$^{-3}$. The water retention curve was measured from samples collected in the field. For potential greater than -1 m, the sintered funnel method was used by wetting up air-dried samples. Potential less than -1 m was measured using suction table (Reeve and Carter, 1991).
7.2.4.3. Lansdowne clay loam

The Red Kandosol (Isbell, 1996) from Lansdowne, NSW, clay loam texture (sand: 22%, silt: 43% and clay 34%), was used in a laboratory study. The soil was ground and sieved to have particles < 2 mm. The soil was then packed in a container of area $0.35 \times 0.35 \, \text{m}^2$ up to a bulk density of 1.4 Mg m$^{-3}$. Two multiple tensions experiments were conducted. In experiment 1, the water content at 20 mm underneath the disc was monitored using a TDR with a probe of 80 mm length. The water retention curve of the sieved soil material was measured. For potential greater than -1 m, the air-dried samples was placed in a sintered funnel and water was imbibed from the base of the funnel. For potential less than -1 m, the filter paper method (Fawcett and Collis-George, 1957) was used.
7.3. Results and discussion

7.3.1. Numerical example

7.3.1.1. Hypothetical soil

7.3.1.1.1. Infiltration curve, sensitivity and influential observations

The cumulative infiltration curve for the loam and clay is shown in Figure 7.3.1. As seen on the graph, the infiltration in the loam is much larger than in the clay and the effect of change in supply tensions during infiltration are also more visible. The placement of 5 mm of sand between the disc membrane and soil surface increased the cumulative infiltration. As shown previously in Chapter VI, the effect is cumulative and is more prominent in clay than in loam.

The infiltration rate curve is shown in Figure 7.3.2. When an independent random effect is added to the cumulative infiltration, the rate is very noisy. Differentiation of the cumulative infiltration appears to alleviate the effect of sand, the rate with sand is quite similar to the data with random noise.

Sensitivity analysis evaluating the effect of 1% change in the hydraulic parameters \((\theta_r, \theta_s, \alpha, n, K_s)\) on the cumulative infiltration and infiltration rate curves are shown in Figure 7.3.3. These were derived using error-free data. The sensitivity increases with time following the infiltration pattern. Parameter \(\theta_r\) has the least effect on the infiltration curve, followed by \(\theta_s, K_s, \alpha\) and \(n\). Parameter \(\theta_s\) and \(K_s\) which can be estimated from TDR readings and Wooding’s analysis appear to be less sensitive than the unknown \(\alpha\) and \(n\). One percent change in the value of \(n\) produced 2% difference in the final cumulative infiltration. This shows that in the cumulative infiltration curve, large-time data are more important carriers of information compared with early-time data. For the infiltration rate, the early-time data also possess valuable information, however.
Figure 7.3.1. Cumulative infiltration for (a) loam and (b) clay. Smooth curves along the data points represent the fitted data from the inverse solution.
Figure 7.3.2. Infiltration rate for (a) loam and (b) clay.
Influential observations as measured by Cook’s $D$ were calculated for the non-error data. Figure 7.3.4a showed that the points where the applied tension changes are very influential in the inverse solution. The end points of the infiltration curve also have high influence, which determine the steady-state rate, in turn estimating $K_s$. Cook’s $D$ for the infiltration rate curve showed that the first point of the curve is the most influential point. Similar to the sensitivity analysis, the infiltration rate data show that the initial point is very influential. Points where supply tension changes are
also influential, but the final point in the curve is not as influential as in the cumulative infiltration case.

Figure 7.3.4. Cook’s $D$ identifying influential observation in (a) cumulative infiltration for the loam, and (b) infiltration rate for the clay. Dark square dots represent Cook’s $D$ and white dots represent measurement data.
7.3.1.1.2. Inverse analysis

Results for the parameter estimation are shown in Table 7.3.1 for the loam and Table 7.3.2 for the clay. As can be seen when no error is imposed on the cumulative infiltration data (L1a and C1a), all five parameters (\(\theta_r\), \(\theta_s\), \(\alpha\), \(n\), \(K_s\)) can be optimized and very good predictions are obtained. When errors in the data are imposed, the inverse solution cannot provide reliable estimates of the parameters (case L1e and C1e). For the loam (L1e) \(\theta_s\) and \(K_s\) were overestimated, while \(n\) was underestimated but \(\alpha\) was very close to actual values. In the case of the clay (C1e), all parameters were far from the actual values. In a real situation, measurement error is bound to occur and the estimates of \(\theta_s\) and \(K_s\) values are also not perfect. Fixing \(\theta_s\) and \(K_s\) close to their true values (L1b, L1c) provides good estimates of \(\alpha\) and \(n\) in the case of the loam. For the clay (C1b, C1c), the estimates were lower but still within an acceptable range. When error in supply tensions is added, the estimations tend to drift away from the true values. The case of L1d and C1d showed that fixing \(\theta_s\) and \(K_s\) close to their true values still provides a lower estimates of \(\alpha\) and \(n\).

When a layer of sand is added between the disc and soil surface, the error is larger. The effect of sand is more prominent on the clay compared to the loam (Figure 7.3.1). Difficulties in convergence were also experienced in the optimization process. For loam (L1f and L1g) \(\theta_s\) and \(K_s\) were highly overestimated, while in clay (C1f and C1g) fixing \(\theta_s\) and \(K_s\) close to their true values does not help to recover the reliability of \(\alpha\) and \(n\) estimates. Addition of a sand layer utterly alters estimates of the hydraulic parameters. Since the numerical model assumed a uniform homogenous soil medium, introduction of a layered system change the predicted hydraulic curves.

Since it has been shown that differentiating the cumulative infiltration with time in some way may alleviate the effect of sand, infiltration rate calculated from finite difference approximations were used as fitting criteria. Using infiltration rate is not as promising as the cumulative infiltration (case L2a – L2d for loam, C2a – C2d for clay). Using input of data without error could not recover the true parameter values. Fixing \(\theta_s\) and \(K_s\) values help directing the parameters closer to their true values.
Table 7.3.1. Parameter estimates from the inverse solution of disc permeameter measurement for the loam.

<table>
<thead>
<tr>
<th>Case</th>
<th>Data Type</th>
<th>Initial estimates</th>
<th>Final estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\theta_i$ m$^{-3}$</td>
<td>$\theta_s$ m$^{-3}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>m$^{-3}$</td>
<td>m$^{-3}$</td>
</tr>
<tr>
<td>Cumulative infiltration data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1a</td>
<td>(0)</td>
<td>0.08</td>
<td>0.45</td>
</tr>
<tr>
<td>L1b</td>
<td>(1)</td>
<td>0.01</td>
<td>0.45</td>
</tr>
<tr>
<td>L1c</td>
<td>(1)</td>
<td>0.01</td>
<td>0.44</td>
</tr>
<tr>
<td>L1d</td>
<td>(2)</td>
<td>0.10</td>
<td>0.41</td>
</tr>
<tr>
<td>L1e</td>
<td>(2)</td>
<td>0.08</td>
<td>0.45</td>
</tr>
<tr>
<td>L1f</td>
<td>(3)</td>
<td>0.10</td>
<td>0.44</td>
</tr>
<tr>
<td>L1g</td>
<td>(3)</td>
<td>0.08</td>
<td>0.45</td>
</tr>
<tr>
<td>Infiltration rate data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2a</td>
<td>(0)</td>
<td>0.08</td>
<td>0.45</td>
</tr>
<tr>
<td>L2b</td>
<td>(0)</td>
<td>0.10</td>
<td>0.45</td>
</tr>
<tr>
<td>L2c</td>
<td>(2)</td>
<td>0.10</td>
<td>0.45</td>
</tr>
<tr>
<td>L2d</td>
<td>(3)</td>
<td>0.08</td>
<td>0.44</td>
</tr>
</tbody>
</table>

# maximum 30 iterations was reached but no convergence.

Table 7.3.2. Parameter estimates from the inverse solution of disc permeameter measurement for the clay.

<table>
<thead>
<tr>
<th>Case</th>
<th>Data Type</th>
<th>Initial estimates</th>
<th>Final estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\theta_i$ m$^{-3}$</td>
<td>$\theta_s$ m$^{-3}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>m$^{-3}$</td>
<td>m$^{-3}$</td>
</tr>
<tr>
<td>Cumulative infiltration data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1a</td>
<td>(0)</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>C1b</td>
<td>(1)</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>C1c</td>
<td>(1)</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>C1d</td>
<td>(2)</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>C1e</td>
<td>(2)</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>C1f</td>
<td>(3)</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>C1g</td>
<td>(3)</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Infiltration rate data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2a</td>
<td>(0)</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>C2b</td>
<td>(2)</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>C2c</td>
<td>(3)</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>C2d</td>
<td>(3)</td>
<td>0.08</td>
<td>0.44</td>
</tr>
</tbody>
</table>

# maximum 30 iterations was reached but no convergence.
Figure 7.3.5. True water retention and hydraulic curve along with inverse solution predictions for the loam.

Figure 7.3.6. True water retention and hydraulic curve along with inverse solution predictions for the clay.
When using infiltration rate as input to the inverse solution, the parameter estimates for the clay are more erroneous than for the loam. The infiltration rate is quite sensitive to the change in cumulative infiltration data, small error in the cumulative infiltration can translate into an erratic infiltration rate.

The water retention and hydraulic conductivity predicted from the inverse analysis is shown in Figure 7.3.5 for the loam and Figure 7.3.6 for the clay. When sand was added in the infiltration, the inverse solution over-predicted the water retention and hydraulic conductivity. Considering the error in measurement and field variability, the inverse solution may still provide sufficient information for soil-water characterization in the field.

To investigate whether small change in applied tension data may provide enough information to produce a good inverse solution, the loam is considered again with applied tensions: 0 to 3600 seconds: -100 mm, then an increment of -20 mm every 30 minutes up to three hours. The result with randomness added to the data is shown in Figure 7.3.7. The small change in tensions resulted in an unobservable change in cumulative infiltration curve. Parameters $\alpha$, $n$ and $K_s$ were optimized and the predicted values are 3.93 mm$^{-1}$, 1.85 and 0.0208 mm/s. The predictions are as good as the ones using large tension changes.
7.3.1.2. Analysis of HAPEX-Sahel data

The infiltration analysis on the first HAPEX-Sahel data is shown in Figure 7.3.8. Cook’s \( D \) indicated that influential observations occurred at the times where supply tensions were changing. The experimental conditions are given in Table 7.3.3. As per Šimunek \textit{et al.} (1998) near saturated hydraulic conductivity was calculated from the steady-state infiltration rate using Wooding’s analysis with a piecewise exponential model (Reynolds and Elrick, 1990). The analysis assumes a piecewise exponential hydraulic conductivity within adjacent applied tension, and a log-linear function is fitted between two points.

<table>
<thead>
<tr>
<th>( t_{\text{initial}} )</th>
<th>( t_{\text{final}} )</th>
<th>( h_0 )</th>
<th>Steady-state rate</th>
<th>( h' )</th>
<th>( K(h') )</th>
<th>( K_s(h') )</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>s</td>
<td>mm</td>
<td>mm s(^{-1})</td>
<td>mm</td>
<td>mm s(^{-1})</td>
<td>mm s(^{-1})</td>
</tr>
<tr>
<td>0</td>
<td>2250</td>
<td>-115</td>
<td>0.0260</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2430</td>
<td>3330</td>
<td>-90</td>
<td>0.0317</td>
<td>-102.5</td>
<td>0.0126</td>
<td>0.0283</td>
</tr>
<tr>
<td>3510</td>
<td>4290</td>
<td>-60</td>
<td>0.0373</td>
<td>-75.0</td>
<td>0.0120</td>
<td>0.0179</td>
</tr>
<tr>
<td>4470</td>
<td>5115</td>
<td>-30</td>
<td>0.0458</td>
<td>-45.0</td>
<td>0.0166</td>
<td>0.0226</td>
</tr>
<tr>
<td>5295</td>
<td>5835</td>
<td>-10</td>
<td>0.0529</td>
<td>-20.0</td>
<td>0.0204</td>
<td>0.0236</td>
</tr>
<tr>
<td>6015</td>
<td>6450</td>
<td>-1</td>
<td>0.0657</td>
<td>-5.5</td>
<td>0.0414</td>
<td>0.0473</td>
</tr>
</tbody>
</table>

Šimunek \textit{et al.’s} analysis (calculated from program DISC) optimised \( \theta_s \), \( \alpha \), \( n \), and \( K_s \). The weight set to the cumulative infiltration data (in cm\(^3\)) is the inverse of the data variance (0.000022), while final water content underneath the disc is arbitrarily weighted by 10. This forced the \( \theta_s \) to be estimated close to the final water content. The estimated saturated hydraulic conductivity is somewhat lower than the one obtained from Wooding’s analysis (Table 7.3.4).

Results of the analysis using \( \mathbf{T}^k \) is somewhat different from Šimunek’s (Table 7.3.4). Since the final water content is not used in the inverse solution and also not heavily weighted, optimizing \( \theta_s \) resulted in a small estimate (< 0.25 m\(^3\) m\(^{-3}\)). Therefore \( \theta_s \) had to be fixed close to the final water content. In the first estimate, \( \alpha \) and \( n \) are close to DISC’s answer but \( K_s \) is predicted too low (0.01 mm s\(^{-1}\)). According to Wooding’s analysis \( K_s \) should be 0.04 mm s\(^{-1}\), while Šimunek’s answer is 0.02 mm s\(^{-1}\). Therefore, in the second estimate \( K_s \) was fixed at 0.04 mm s\(^{-1}\), only \( \alpha \) and \( n \) were
optimized. Results showed that parameter $\alpha$ is quite close to DISC’s answer but $n$ is smaller.

![Figure 7.3.8. Cumulative infiltration from the HAPEX-Sahel data set 1. Dots represent measurement points and smooth line crossing the points are the fitted curve. The dotted line is Cook’s $D$ values identifying influential observation. Numbers beside the curve is the related applied tensions.](image)

**Table 7.3.4. Predicted hydraulic parameters using program DISC and $\beta$ for data set 1.**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>DISC</th>
<th>$\beta$ 1st estimate</th>
<th>$\beta$ 2nd estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>s.e.</td>
<td>Estimate</td>
</tr>
<tr>
<td>$\theta_r$ (m$^3$ m$^{-3}$)</td>
<td>0.010</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$\theta_s$ (m$^3$ m$^{-3}$)</td>
<td>0.275</td>
<td>0.0001</td>
<td>0.280</td>
</tr>
<tr>
<td>$\alpha$ (m$^{-1}$)</td>
<td>2.967</td>
<td>0.3300</td>
<td>3.573</td>
</tr>
<tr>
<td>$n$</td>
<td>1.973</td>
<td>0.0454</td>
<td>2.283</td>
</tr>
<tr>
<td>$K_s$ (mm s$^{-1}$)</td>
<td>0.0218</td>
<td>0.0070</td>
<td>0.0142</td>
</tr>
</tbody>
</table>
Since no true parameter values are available, a comparison with near-saturated \( K \) is given in Figure 7.3.9. Analysis from DISC fit better for \( K > -20 \) mm but not at saturation. Although fixing \( K_s \) to the ‘real’ value is not as good as DISC’s answer, it preserves the shape of the function at saturation. The difference in \( K \) estimates (~0.005 mm s\(^{-1}\)) is quite small compared with the accuracy of measurements.

The second data set (Table 7.3.5 and Figure 7.3.10) had a slower infiltration rate than the first set. Similar analysis was carried out on the data, and the hydraulic parameter estimates are shown in Table 7.3.6. As before if \( K_s \) was optimized, a small value (0.02 mm s\(^{-1}\)) was obtained. Based on Wooding’s analysis the value is fixed to 0.03 mm s\(^{-1}\).

**Table 7.3.5. Experimental state and Wooding’s analysis for data set 2.**

<table>
<thead>
<tr>
<th>( t_{\text{initial}} ) s</th>
<th>( t_{\text{final}} ) s</th>
<th>( h_0 ) mm</th>
<th>Steady-state rate ( h^* ) mm s(^{-1})</th>
<th>( h^* ) mm</th>
<th>( K(h^*) ) mm s(^{-1})</th>
<th>( K_s(h^*) ) mm s(^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1050</td>
<td>-100</td>
<td>0.0088</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1650</td>
<td>2623</td>
<td>-70</td>
<td>0.0138</td>
<td>-85</td>
<td>0.0071</td>
<td>0.0255</td>
</tr>
<tr>
<td>2923</td>
<td>3208</td>
<td>-50</td>
<td>0.0205</td>
<td>-60</td>
<td>0.0119</td>
<td>0.0391</td>
</tr>
<tr>
<td>3508</td>
<td>3731</td>
<td>-30</td>
<td>0.0272</td>
<td>-40</td>
<td>0.0150</td>
<td>0.0264</td>
</tr>
<tr>
<td>4045</td>
<td>4179</td>
<td>-10</td>
<td>0.0405</td>
<td>-20</td>
<td>0.0236</td>
<td>0.0351</td>
</tr>
</tbody>
</table>
Table 7.3.6. Predicted hydraulic parameters using program DISC and \( \mathbf{R} \) for data set 2.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>DISC</th>
<th>( \mathbf{R} )(^{1})st estimate</th>
<th>( \mathbf{R} )(^{2})nd estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>s.e.</td>
<td>Estimate</td>
</tr>
<tr>
<td>( \theta_r ) (m(^3) m(^{-3}))</td>
<td>0.010</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( \theta_s ) (m(^3) m(^{-3}))</td>
<td>0.291</td>
<td>0.0004</td>
<td>0.290</td>
</tr>
<tr>
<td>( \alpha ) (m(^{-1}))</td>
<td>7.173</td>
<td>0.1860</td>
<td>6.844</td>
</tr>
<tr>
<td>( n )</td>
<td>1.779</td>
<td>0.0732</td>
<td>1.603</td>
</tr>
<tr>
<td>( K_s ) (mm s(^{-1}))</td>
<td>0.030</td>
<td>0.0040</td>
<td>0.017</td>
</tr>
</tbody>
</table>

Figure 7.3.10. Cumulative infiltration from the HAPEX-Sahel data set 2. Dots represent measurement points and smooth line crossing the points is the fitted curve. Dotted lines are Cook’s \( D \) values identifying influential observation. Numbers beside the curve is the related applied tensions.

Results of the optimized \( \alpha \) and \( n \) is quite similar to DISC’s answers, only a higher \( \alpha \) was obtained. Figure 7.3.11 showed the predictions of near-saturated conductivity is quite similar. This exercise indicates that \( \theta_r \) as estimated from the final water content and \( K_s \), estimated from Wooding’s analysis, are useful in the inverse solution. These values should be fixed (not optimized) when unreasonable estimates are obtained.
7.3.2. Real infiltration data

7.3.2.1. Mount Annan silty loam

The infiltration of Mount Annan silty loam is shown in Figure 7.3.12. Although the three sites are close by (within 10 m range) the infiltration varies considerably. Although site 3 has the highest applied tension (-30 mm) the cumulative infiltration is lower than site 1 (-60 mm). Inverse analysis was carried out on the data with $\theta_r$, $\theta_s$ and $K_s$ fixed and only $\alpha$ and $n$ were optimized. Results in Table 7.3.7. show different parameter estimates for all three experiments. Experiment 1, which has higher infiltration, has a lower $\alpha$ estimate reflecting a higher air entry potential and thus higher moisture storage. While experiments 2 and 3 which has lower infiltration show large $\alpha$ values.

The highly variable infiltration data in the field produced spurious water retention and hydraulic conductivity estimates (Figures 7.3.13 and 7.3.14). Water retention was underpredicted, especially in Experiment 3. Field variability exists, from the highly variable hydraulic conductivity derived from Wooding’s analysis, this also suggests errors in the infiltration data.
Figure 7.3.12. Cumulative infiltration of Mount Annan silty loam. Numbers beside the curve is the applied tensions.

Table 7.3.7. Predicted hydraulic parameters for Mount Annan silty loam.

<table>
<thead>
<tr>
<th>Exp</th>
<th>$\theta_i$ m$^{-3}$</th>
<th>$\theta_s$ m$^{-3}$</th>
<th>$\alpha$ m$^{-1}$</th>
<th>$n$</th>
<th>$K_s$ mm s$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.10</td>
<td>0.46</td>
<td>6.85</td>
<td>1.76</td>
<td>0.002</td>
</tr>
<tr>
<td>2</td>
<td>0.10</td>
<td>0.46</td>
<td>11.95</td>
<td>2.97</td>
<td>0.001</td>
</tr>
<tr>
<td>3</td>
<td>0.00</td>
<td>0.46</td>
<td>15.84</td>
<td>5.50</td>
<td># 0.001</td>
</tr>
</tbody>
</table>

*Value reached maximum during optimization.
Figure 7.3.13. Water retention curve for Mount Annan silty loam. Dots represent laboratory measurements and solid line is the prediction using inverse analysis.

Figure 7.3.14. Near-saturation hydraulic conductivity analysed for Mount Annan silty loam. Dots represent calculation with Wooding’s analysis, solid line predicted from inverse solution.
7.3.2.2. Marinya clay

The infiltration in the Vertosol is shown in Figure 7.3.15. The infiltration is also quite variable, experiment 1 has the lowest cumulative infiltration although the initial supply potential is higher than Experiment 3. The experimental description and predicted near-saturation hydraulic conductivity were summarised in Table 7.3.8.

![Figure 7.3.15. Cumulative infiltration of Marinya clay. Numbers beside the curve is the applied tensions.](image)

<table>
<thead>
<tr>
<th>Exp.</th>
<th>t initial (s)</th>
<th>t final (s)</th>
<th>θ final (m$^3$ m$^{-3}$)</th>
<th>h (mm)</th>
<th>q (mm s$^{-1}$)</th>
<th>h$\ddot{\alpha}$ (mm)</th>
<th>K(h) (mm s$^{-1}$)</th>
<th>K(h$\ddot{\alpha}$) (mm s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>4500</td>
<td>-160</td>
<td>0.0005</td>
<td>-120</td>
<td>0.0006</td>
<td>0.0046</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4500</td>
<td>8340</td>
<td>-80</td>
<td>0.0020</td>
<td>-45</td>
<td>0.0043</td>
<td>0.0178</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8340</td>
<td>10200</td>
<td>0.47</td>
<td>-10</td>
<td>-45</td>
<td>0.0043</td>
<td>0.0178</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>2880</td>
<td>-90</td>
<td>0.0025</td>
<td>-80</td>
<td>0.0019</td>
<td>0.0102</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2880</td>
<td>5820</td>
<td>-70</td>
<td>0.0038</td>
<td>-50</td>
<td>0.0033</td>
<td>0.0083</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5820</td>
<td>9000</td>
<td>0.45</td>
<td>-30</td>
<td>-45</td>
<td>0.0043</td>
<td>0.0178</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>4680</td>
<td>-190</td>
<td>0.0012</td>
<td>-130</td>
<td>0.0019</td>
<td>0.0183</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4680</td>
<td>5940</td>
<td>0.46</td>
<td>-70</td>
<td>-50</td>
<td>0.0033</td>
<td>0.0083</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5940</td>
<td>9340</td>
<td>0.45</td>
<td>-30</td>
<td>-50</td>
<td>0.0043</td>
<td>0.0178</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1200</td>
<td>0.44</td>
<td>+10</td>
<td>+10</td>
<td>0.0439</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.3.8. Experimental state for Marinya clay.
Experiment 4 measured $K_s$ using the ponded version of the disc permeameter. The value was about two times higher than the ones predicted using the multiple tensions. This suggests the presence of macropores that accelerated the infiltration.

The hydraulic parameters predicted from the inverse solution are given in Table 7.3.9. Since $\theta_r$ was always over-predicted during the optimization process, the value is arbitrarily fixed at its air-dry moisture content $0.1 \text{ m}^3 \text{ m}^{-3}$. In experiment 1 if $K_s$ was optimized, the predicted value is ten times higher than values calculated from Wooding’s solution. If the pore connectivity factor in the hydraulic conductivity curve $l$ was also optimized, parameter $\alpha$ changed significantly in experiment 1, but not in experiment 2. Increasing the number of parameters produced better fit to the data, but did not necessarily give better estimates of hydraulic properties.

The predicted water retention and laboratory measured data is shown in Figure 7.3.16. Predicted water retention values do not match the laboratory measured data; they are mostly under-predicted. Experiments 1 and 2 yielded similar predictions, while experiment 3 which has higher infiltration rate produced closer prediction. Similar to the Mount Annan data, the higher the cumulative infiltration, the lower the $\alpha$ estimates and the closer to the laboratory measured water retention.

<table>
<thead>
<tr>
<th>Exp.</th>
<th>$\theta_i$ m$^3$ m$^{-3}$</th>
<th>$\theta_r$ m$^3$ m$^{-3}$</th>
<th>$\alpha$ m$^{-1}$</th>
<th>$n$</th>
<th>$K_s$ mm s$^{-1}$</th>
<th>$l$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1$^f$</td>
<td>0.482</td>
<td>12.16</td>
<td>2.03</td>
<td>0.018$^f$</td>
<td>0.5$^f$</td>
</tr>
<tr>
<td>1b</td>
<td>0.1$^f$</td>
<td>0.518</td>
<td>24.38</td>
<td>1.47</td>
<td>0.119$^f$</td>
<td>0.5$^f$</td>
</tr>
<tr>
<td>1c</td>
<td>0.1$^f$</td>
<td>0.485</td>
<td>6.322</td>
<td>1.54</td>
<td>0.033</td>
<td>13.72</td>
</tr>
<tr>
<td>2</td>
<td>0.1$^f$</td>
<td>0.462</td>
<td>9.520</td>
<td>2.15</td>
<td>0.008$^f$</td>
<td>0.5$^f$</td>
</tr>
<tr>
<td>2b</td>
<td>0.1$^f$</td>
<td>0.453</td>
<td>8.782</td>
<td>3.12</td>
<td>0.004$^f$</td>
<td>0.5$^f$</td>
</tr>
<tr>
<td>2c</td>
<td>0.1$^f$</td>
<td>0.455</td>
<td>9.053</td>
<td>2.90</td>
<td>0.004</td>
<td>0.2</td>
</tr>
<tr>
<td>3</td>
<td>0.1$^f$</td>
<td>0.524</td>
<td>8.896</td>
<td>1.99</td>
<td>0.018$^f$</td>
<td>0.5$^f$</td>
</tr>
<tr>
<td>3b</td>
<td>0.1$^f$</td>
<td>0.511</td>
<td>6.423</td>
<td>2.40</td>
<td>0.006</td>
<td>0.5$^f$</td>
</tr>
</tbody>
</table>

$^f$ Parameter value was fixed (not optimized) in the inverse solution.

The near-saturation hydraulic conductivity predicted from the inverse solution and calculated using Wooding’s analysis is shown in Figure 7.3.17. Although there is variation at saturation, the hydraulic conductivity curves estimated on the three experiments are quite similar. This data set is definitely more reliable than the Mount Annan data set. Hydraulic conductivity is also calculated from laboratory measured
water retention curve using the Mualem-van Genuchten model (MVG). The estimate is much lower than those of Wooding’s solution, which suggests that $K(h)$ from the inverse estimate is quite reliable.

Figure 7.3.16. Water retention curve for Marinya clay. Dots represent laboratory measurements and solid line is the prediction using inverse analysis.

Figure 7.3.17. Near-saturation hydraulic conductivity analysed for Marinya clay. Dots represent calculation with Wooding’s analysis, solid line predicted from inverse solution.
7.3.2.3. Lansdowne clay loam

Using a homogenous soil material, the experiment was carried out in the laboratory. Table 7.3.10 summarises the experimental conditions and calculated hydraulic conductivity. Figure 7.3.18 shows the two infiltration runs on the repacked soil. Under the same tension, run 2 has higher infiltration compared to run 1. Differences in packing may be the cause, run 1 has an average bulk density of 1.54 Mg m$^{-3}$, while run 2 was 1.50 Mg m$^{-3}$.

![Figure 7.3.18. Cumulative infiltration of Lansdowne clay loam. Numbers beside the curve is the applied tensions.](image)

<table>
<thead>
<tr>
<th>Exp.</th>
<th>t initial</th>
<th>t final</th>
<th>$\theta$ final</th>
<th>$h$</th>
<th>$q$</th>
<th>$i_\lambda$</th>
<th>$K(h)$</th>
<th>$K_s(h)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>s</td>
<td>s</td>
<td>m$^3$ m$^{-3}$</td>
<td>mm</td>
<td>mm s$^{-1}$</td>
<td>mm</td>
<td>mm s$^{-1}$</td>
<td>mm s$^{-1}$</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>3600</td>
<td>-100</td>
<td>0.0023</td>
<td>0.0062</td>
<td>-75</td>
<td>0.0023</td>
<td>0.0102</td>
</tr>
<tr>
<td></td>
<td>4500</td>
<td>6300</td>
<td>-50</td>
<td>0.0135</td>
<td>0.0076</td>
<td>-35</td>
<td>0.0061</td>
<td>0.0152</td>
</tr>
<tr>
<td></td>
<td>6300</td>
<td>7500</td>
<td>0.46</td>
<td>-20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>2880</td>
<td>-100</td>
<td>0.0046</td>
<td>0.0076</td>
<td>-75</td>
<td>0.0026</td>
<td>0.0055</td>
</tr>
<tr>
<td></td>
<td>4500</td>
<td>6300</td>
<td>0.43</td>
<td>-50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Results of the inverse solution are given in Table 7.3.11. Parameter $\theta_r$ was fixed at the air-dried water content of 0.06 m$^3$ m$^{-3}$. In run 1 different types of data
(cumulative infiltration, infiltration rate, water content underneath the disk) were used as the inputs to the inverse solution. Figure 7.3.19 show the observed and fitted infiltration rate, when taking the derivative of measured cumulative infiltration the rate is quite noisy. Figure 7.3.20 shows that the inverse solution matched the measured water content very well. The solutions all give similar parameter values. While in run 2, $K_s$ had to be fixed, otherwise a very high $n$ was estimated.

Table 7.3.11. Predicted hydraulic parameters for Lansdowne clay loam.

<table>
<thead>
<tr>
<th>Run</th>
<th>Data fitted</th>
<th>$\theta_r$ m$^{-3}$ m$^{-3}$</th>
<th>$\theta_i$ m$^{-3}$ m$^{-3}$</th>
<th>$\alpha$ m$^{-1}$</th>
<th>$n$</th>
<th>$K_s$ mm s$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$I$</td>
<td>0.06$^f$</td>
<td>0.436</td>
<td>10.20</td>
<td>2.10</td>
<td>0.0102</td>
</tr>
<tr>
<td>1b</td>
<td>$I$</td>
<td>0.06$^f$</td>
<td>0.440</td>
<td>11.40</td>
<td>1.89</td>
<td>0.0150$^f$</td>
</tr>
<tr>
<td>1c</td>
<td>$I$ &amp; $\theta$</td>
<td>0.06$^f$</td>
<td>0.411</td>
<td>11.71</td>
<td>1.90</td>
<td>0.0150$^f$</td>
</tr>
<tr>
<td>1d</td>
<td>$di/dt$</td>
<td>0.06$^f$</td>
<td>0.430</td>
<td>12.80</td>
<td>1.74</td>
<td>0.0190</td>
</tr>
<tr>
<td>1f</td>
<td>$\theta^g$</td>
<td>0.06$^f$</td>
<td>0.416</td>
<td>13.09</td>
<td>1.48</td>
<td>0.0150$^f$</td>
</tr>
<tr>
<td>2</td>
<td>$I$</td>
<td>0.06$^f$</td>
<td>0.431</td>
<td>6.03</td>
<td>4.50</td>
<td>0.0026</td>
</tr>
<tr>
<td>2b</td>
<td>$I$</td>
<td>0.06$^f$</td>
<td>0.481</td>
<td>12.32</td>
<td>1.49</td>
<td>0.0050$^f$</td>
</tr>
</tbody>
</table>

$^f$ Parameter value was fixed in the optimization,  
$^g$ Maximum 30 iterations but no convergence.

Figure 7.3.19. Infiltration rate for Lansdowne clay loam run 1. Dots represent values calculated from cumulative infiltration data and smooth curve is the fitted curve from the inverse solution. Numbers beside the curve is the prescribed tensions.
The predicted water retention is given in Figure 7.3.21. Run 1 gives a lower estimate of water retention. Unexpectedly, different types of data used as fitting criteria in the inverse solution provide similar retention curves. Run 1 predicts well near saturation, but as the potential becomes more negative the water content is predicted lower. Run 2 which has a higher infiltration produced closer fit to the water retention. Although the water content at the dry end of the curve \((h < -1 \text{ m})\) is predicted lower, this estimate is quite satisfactory. Under homogenous media in the laboratory, the inverse solution gave good prediction.

Near-saturation hydraulic conductivity from Wooding’s analysis and predicted from the inverse analysis shown in Figure 7.3.22. As before, \(K(h)\) from inverse analysis is quite consistent with Wooding’s analysis. However, \(K(h)\) calculated from laboratory measured water retention curve (MVG) do not match those of Wooding’s solution at all.
Figure 7.3.21. Water retention curve for Lansdowne clay loam. Dots represent laboratory measurements, curves below the dots are the predicted water retention from inverse solution for experiment 1 using different data as fitting criteria.

Figure 7.3.22. Near saturated hydraulic conductivity for Lansdowne clay loam. Dots represent Wooding’s analysis, and the smooth curves are the predicted using inverse solution.
7.3.3. **Overall discussion**

In the numerical example it was demonstrated that data with no error, provided a very good solution and all of the five parameters ($\theta_r$, $\theta_s$, $\alpha$, $n$ and $K_s$) can be determined simultaneously. Under real experimental conditions, perfect data cannot be obtained. Variable field conditions and experimental set-up can produce variable infiltration results. Placement of sand increases the cumulative infiltration. As a result, field infiltration data do not yield a good prediction of water retention curve as measured in laboratory. This problem was also found by Šimunek et al (1998b, 1999). Their results also showed that water retention predicted from the inverse solution were lower than those from laboratory measurements. They provide some rationalization:

- Soil hydraulic properties can be different as evaluated by different methods. Different laboratory methods may give different results compared to field methods.
- Local soil heterogeneity which create different water flow responses, resulting in different hydraulic properties estimates.
- The inappropriate representation of the hydraulic functions. Since $\theta(h)$ and $K(h)$ are estimated simultaneously by the van Genuchten equations that share common parameters, the functions may not properly fit the real hydraulic properties.

We add an additional point:

- Field measurements are conducted on the undisturbed samples, while transporting the column in laboratory create disturbance and the change in support volume.

The first and last points imply that field measurements cannot be replicated in the laboratory (Kasteel et al., 1999). Various authors (Topp et al., 1967; Smiles et al., 1971) have demonstrated that the water retention may not only depend on the wetting and drying characteristics but also on the dynamics of water flow. Topp et al. (1967) observed differences in the water retention characteristics of a fine sand obtained during transient flow compared to the ones measured under static equilibrium. Smiles et al. (1971) indicated that during a transient, non-hysteretic flow of water in a horizontal column, the water retention curve is unique during the wetting process. But during the drying process the water retention characteristics were not unique throughout the column. Plagge et al. (1999) suggested that the dynamics of the flow process also affect the hydraulic conductivity.
Wessolek et al. (1997) found that $\theta_s$ and $K_s$ derived from field measurements were higher than the ones measured in laboratory. Consequently, drainage calculated using laboratory measured hydraulic properties was lower than that calculated using field derived hydraulic properties. Mallants et al. (1997) found considerable differences in hydraulic properties measured using different techniques in the laboratory. Kasteel et al. (1999) compared the laboratory and field determined hydraulic parameters on simulations of water and solute transport in the field. The field hydraulic parameters were obtained by the inverse solution of constant irrigation experimental data. Simulations using field-determined hydraulic properties agreed well with field measured data, while laboratory determined properties overestimated water content of the whole profile.

The van Genuchten function only allows a smooth transition between potential and water content or hydraulic conductivity. Sometimes sharp changes in water content as potential becomes more negative can be found, especially the dry end. Ross and Smettem (1993) proposed the hydraulic curves as a sum of two or more hydraulic functions. When fitted to water retention data, parameter $\theta_1$ depends on the range of water content over which it is measured, therefore sometimes it cannot be treated physically. Ross et al. (1991) proposed an extension of the hydraulic model so the curve will terminate ($\theta = 0$) at potential of $-1000$ MPa. The use of separate water retention and hydraulic conductivity functions may force the fitted infiltration data and water retention to coincide as done by Šimunek et al. (1998b). Varying the pore connectivity factor $l$ may improve the fit but does not necessarily reflect the correct $K(h)$ relationship. At the present, we don’t see the benefit of adding more parameters, as it will create over-parameterization. According to Ockham’s razor non sunt multiplicanda entia praeter necessitatem (entities are not to be multiplied beyond necessity). Because the water retention and hydraulic conductivity curves share common parameters, the inverse solution will yield estimates which are compromise between both curves. The estimated parameters should be treated merely as empirical.

In their field experiments Šimunek et al. (1998b), observed two measurements within 1 m distance in a fine sandy loam gave different infiltration pattern. The infiltration of the first run was 50 to 100% higher than the second run and the final water content was also higher. This phenomena was also present in this study, which
raised the question how reproducible is the infiltration curve as measured by the disc permeameter.

Repeated disc permeameter measurements were conducted in the laboratory on sieved, air-dried Landsdowne clay loam, packed in a box of $0.25 \times 0.25 \, \text{m}^2$ area. The infiltration was carried out at supply potential of -80 mm for 15 minutes. Ten replicates were performed by recording the reservoir level of the disc permeameter every 20 seconds for the first five minutes and 30 seconds intervals afterwards. The average bulk density was $1.41 \pm 0.05 \, \text{Mg} \, \text{m}^{-3}$ with the final water content underneath the disk $0.38 \, \text{m}^3 \, \text{m}^{-3}$.

The cumulative infiltration results (Figure 7.3.23) are quite variable, the curves diverge with increasing time, while the infiltration rate showed a reverse trend. Initially the variation is quite large, but as the experiment continues the rate becomes stable and converges into similar estimates. The average infiltration rate of the ten infiltration curves during 12 – 15 minutes is $0.01 \pm 0.001 \, \text{mm/s}$. Standard deviation (Figure 7.3.21c) shows that the variance follows the infiltration pattern. This indicated that Wooding’s solution that utilises the steady-state infiltration rate is quite reproducible.

Figure 7.2.24 shows 16 measurements using the Mini Disk® on a repacked Lansdowne clay loam. Even larger variability is manifest when the sample volume is small. Figure 7.3.25 shows 20 measurements using the CSIRO disc permeameter along a transect on a Black Vertosol. The site is in the Australian Cotton Research Institute and the soil has been cultivated and the surface smoothed (same experimental site as in Chapter 6). Five mm of wet sand were placed between the disc membrane and soil surface. A tension of -20 mm was applied and measurements were made in 30-second intervals for 15 minutes. Highly variable infiltration curves are observed in the relatively small distances.
Figure 7.3.23. (a) Cumulative infiltration, (b) infiltration rate, and (c) the standard deviation of Lansdowne clay loam.
Figure 7.3.24. Cumulative infiltration of Lansdowne clay loam as measured by Mini Disk® at -20 mm tension.

Figure 7.3.25. Field infiltration of Narrabri clay at -20 mm tension using the CSIRO disc permeameter.

All the cumulative infiltration curves showed large variability and are irreproducible. The curves are reproducible only if there is a good contact between the disk membrane and soil surface (Close et al., 1998). In practice, it is difficult to ensure the same degree of contact for each measurement. Small differences in the
initial wetting condition can translate into a different cumulative infiltration pattern. As the infiltration proceeds, capillary force will transmit water to the non-wetted area, hence resulting in a similar infiltration rate. This sensitivity may become a problem in fitting the cumulative infiltration curve. Fitting the infiltration rate may be an alternative, but results showed that infiltration rate is not as robust as the cumulative infiltration.

There has not been many studies on the reproducibility of the infiltration curves. Quadri et al. (1994) conducted a laboratory experiment on a sand packed in a box. A quarter sector of the disc was used to infiltrate water and KBr into the soil. They reported the close agreement between the infiltration curve, water content, and water potential measured experimentally and generated by numerical model. Close et al. (1998) found irreproducible results due to poor contact between the contact sand and the disc membrane. Using dye as a tracer, the wetting pattern showed large variations in the infiltrating area, the non-uniform wetting caused variations in the infiltration rate. When there was a good contact between the sand and the membrane, the infiltration rate is reproducible.

In the field differences may also due to field heterogeneity, and in the laboratory different packing densities. The variation in cumulative infiltration maybe due to:

- Local scale variation, including the soil structural stability.
- The presence of entrapped air during infiltration process. Collis-George and Yates (1985) showed that encapsulated air can alter the outflow rates in laboratory cores under a constant head condition. Cislerova et al. (1988) found that the steady state ponded infiltration rate depended on the initial moisture content. The effect was ascribed to trapped air which alters the volume available for gravity-dominated flow.
- Experimental problems. The use of the disk permeameter requires a good contact between the membrane and soil surface. Small differences in the wetting condition may cause the infiltration curve to deviate with time. The use of contact sand in the field also promote the cumulative infiltration.
- The effect of structural stability during the infiltration processes. The slaking and/or dispersion of clay during the infiltration process can cause the structure to collapse. The integrity of the structure plays important role in estimating the hydraulic properties from infiltration processes.
7.4. Conclusions

The inverse solution is a robust method for simultaneously predicting water retention and hydraulic conductivity curves in the field. Numerical studies show that multiple-tension cumulative infiltration data provide enough information to yield a good prediction. The infiltration rate does not yield reliable estimates. When analysing field infiltration data, the water retention is always lower than the laboratory measured values. The differences may imply difficulties in obtaining representative undisturbed core samples (Field et al., 1984). Different runs within a close range show highly variable cumulative infiltration curves. The larger the cumulative infiltration, the higher is the predicted air-entry value and the closer to the laboratory measured values. Only in one case for the laboratory repacked clay loam, the inverse solution gave a close agreement between laboratory-measured and inverse-predicted water retention.

It can be argued that field-measured hydraulic properties are different from laboratory-measured values. The isotropic homogenous assumption of the axisymmetric disc permeameter model simplifies the soil system in the field. Thus it is a simplified representation of the soil system, which need not be true in reality (Konikow and Bredehoeft, 1992).

Prior to the performance evaluation of the hydraulic properties predicted by the inverse method in describing field water transport, only speculations can be made. The cumulative infiltration, which is used as the input data, is very sensitive to changes in initial condition. Differences in initial condition and area of wetting could produce different infiltration curves. The need for good contact between the disk membrane and soil surface restricts the reproducibility of the infiltration curves.

Recently, the ‘hood infiltrometer’ has been introduced by a German company, Umwelt-Geräte-Technik GmbH. The disc is replaced by a circular shaped hood filled with water directly on the soil surface. The pressure head in the water filled hood is regulated by a Mariotte water supply. The set-up requires no contact material between the soil and the constant head device therefore minimising the possibility of non-uniform wetting. This device may provide a more reliable infiltration data.

The current problem with inversion of disc permeameter data is the technical aspect. Peck (1968) stressed the importance of distinguishing the reproducibility of the method and reproducibility of soil samples. From a practical point of view, it can
be argued that if the sample reproducibility is poor, there is little reason to minimise the error in the method much beyond the degree of sample variation. However, accurate methods must be available to make assessment of sample variation.

Some recommendations for the inverse analysis of disc permeameter measurements can be made:

- Prewet the contact sand to reduce the high sorptivity by the sand, and ensure a good contact between the disc membrane and contact sand.
- Measure the initial soil water content and install a TDR waveguide approximately 2-3 cm underneath the soil surface.
- Perform multiple tension disk measurement, start with a low (more negative) tension, e.g. –100 mm; monitor water content and water reservoir periodically.
- After the flow rate is almost constant (depending on the soil type, usually 1 hour), increase the tension with a large step (e.g. –100 to –50 mm) and maintain the flow at least half an hour. Usually three tensions would provide enough data for the inverse analysis (e.g. -100 → -50 → -20 mm).
- Measure the final water content on the soil surface, collect the soil sample underneath the disc using bulk density cores.
- Perform replicates to ensure a representative characterization.
- Calculate the near saturation hydraulic conductivity using Wooding’s analysis with multiple tensions. The conductivity near saturation could be approximated as $K_s$ and the final water content as $\theta_s$.
- Analyse the cumulative infiltration and water content data using the inverse analysis. Fix $\theta_s$ and $K_s$ values, only optimize $\theta_r$, $\alpha$, and $n$. If $\theta_r$ drifts away to a high value during optimization, fixed the value at the initial water content or at air-dry condition.
References


Appendix 7.1. Description of the programs used in the inverse solution

**Forward-solution**

**WARRICK**
- two-dimensional axi-symmetric flow (for disc permeameter)
- water transport
- internal element generation
- finite-element – Galerkin method
- mass-conservative method (Celia et al., 1990)
- matrix solver: Choleski decomposition.

**SWMS_2D**
- general two-dimensional flow
- water, solute, and heat transport
- finite-element – Galerkin method
- mass-conservative method (Celia et al., 1990)
- matrix solver: preconditioned conjugate gradient.
In the context of predicting soil hydraulic properties, the document outlines an inverse method that involves solving for the unknown parameters. The approach employs numerical techniques for parameter estimation.

### Inverse solution

**KCIRRAW**
- Input data interface
- Modified Levenberg-Marquardt nonlinear least-squares
- Water flow subroutine: WARRICK

**DISC**
- Input data interface
- Levenberg-Marquardt nonlinear least-squares
- Disc water flow subroutine: HYDRUS2 (similar to SWMS_2D)
- Finite elements: triangular
- Input data: cumulative infiltration (cm³), initial water content, final water content
- Objective function:
  \[
  O = \sum_{i=1}^{N} \frac{w_1(I(t_i) - \dot{I}(t_i))^2}{\sigma_i^2} + w_2(\theta - \dot{\theta})^2
  \]
  
  where \( w_1 = \frac{N}{\sigma_i^2} \) and \( w_2 = 10 \).

Each infiltration data is weighted by the inverse of the data variance and final water content by 10.
- Initial condition: \( \theta \)

**Modified**
- Input data interface.
- Modified Levenberg-Marquardt nonlinear least-squares.
- Disc water flow subroutine: modified SWMS_2D.
- Input data: cumulative infiltration [L], infiltration rate [L T⁻¹], water content underneath disc [L³ L⁻³], initial water content [L³ L⁻³].
- Objective function:
  \[
  O = \sum_{j=1}^{k} \sum_{i=1}^{N_j} w_i [q_j(t_i) - \tilde{q}_j(t_i)]^2
  \]
  
  where \( k \) is the number of measurement sets, \( q_j(t_i) \) is the flow response of measurement set \( j \) at time \( t_i \), and \( w = \) weight for individual data point.
- Initial condition: \( \theta \)
B. Minasny – Efficient methods predicting soil hydraulic properties

<table>
<thead>
<tr>
<th></th>
<th>DISC</th>
<th>Program Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>finite element grid</td>
<td>fixed, 3 types: coarse, medium &amp; fine mesh</td>
<td>adjustable</td>
</tr>
<tr>
<td>optimization</td>
<td>Levenberg-Marquardt</td>
<td>Levenberg-Marquardt</td>
</tr>
<tr>
<td>input data</td>
<td>cumulative infiltration (cm³)</td>
<td>cumulative infiltration [L]</td>
</tr>
<tr>
<td></td>
<td>final water content</td>
<td>infiltration rate [L T⁻¹]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>water content underneath disc</td>
</tr>
<tr>
<td>weight of data</td>
<td>internal weighting</td>
<td>user-defined weights</td>
</tr>
<tr>
<td>output</td>
<td>fitted &amp; residuals</td>
<td>fitted, residuals &amp; influential observations (Cook’s D)</td>
</tr>
</tbody>
</table>


VIII. Discussion: - The efficiency of various approaches to obtaining estimates of soil hydraulic properties

8.1. Efficiency

Previous chapters have described some ‘efficient’ methods for predicting hydraulic properties and the measurement of uncertainty. It is the thesis of this work that the relative efficiency of methods for predicting soil hydraulic properties should be determined.

Marion et al. (1994) compared different methods for predicting the soil-water retention and hydraulic-conductivity curves. The methods included field measurements (instantaneous profile), laboratory measurements of soil cores (static equilibrium and the multi-step inverse method) and pedotransfer functions (Arya-Paris particle-size distribution model §1.2.1). Although the in situ instantaneous profile method is the most time consuming, the results were considered the most valid. The inverse solution of the multi-step outflow produced feasible soil-water retention estimates; and the time of measurement was rapid. They envisioned a combination of methods for characterizing soil hydraulic properties in the field, which would minimise effort and maintain accuracy. The multistep outflow method is proposed for providing a good description of field water retention curve, while determination of $K(\theta)$ required the use of instantaneous profile method. There is little research in comparing the efficiency of different methods for predicting hydraulic properties.

Various levels of information can be used to predict soil hydraulic properties and the more relevant information we acquire, the more certain our estimates, but this is usually accompanied by a higher effort and cost of obtaining the information. Efficiency may be defined by several ways:

- efficiency in terms of effort:
  
  \[ \text{Efficiency} = \frac{\text{quality of information}}{\text{effort}} \]

- efficiency in terms of cost:
Efficiency = quality of information / cost of information

- efficiency in terms of value of information

Efficiency = value of information – cost of information

The quality of information can be evaluated by the standard deviation ($\sigma$) of the predicted soil hydraulic properties as a result of the uncertainty in the measurements. 

Effort is described by the time to make the measurement or obtain the information, and cost is the cost of acquiring the measurement and/or information. The first two terms give the efficiency in terms of quality of the information. An alternative method should summarise the quality of the prediction, the cost of information, the application, and the effect of spatial variability. This can be achieved by evaluating the value of information. Value of information is not only evaluating the information itself but also evaluation of various contributing factors.

8.2. Efficiency 1 – The uncertainty and effort of various methods

Hydraulic properties can be predicted by various means depending on the amount of information we have, or are willing to obtain. The efficiency can be evaluated in terms of the quality of the prediction and the effort to obtain the data. Table 8.2.1 shows various measurements or information that can be used and its associated effort, cost and uncertainty in prediction of water retention and hydraulic conductivity.

The hydraulic properties are described by the van Genuchten equation, which is summarised in five parameters $\theta_r$, $\theta_s$, $\alpha$, $n$, $K_s$. The information starts at the coarsest level when minimum soil physical information is available, hence the hydraulic parameters were estimated from the average Australian data. The next level when hand texture is assessed therefore the average value for the textural class can be used to predict the hydraulic parameters (Appendix 4.1). The following levels include particle size analysis, measurement of bulk density, and two points in the retention curve and their associated pedotransfer functions. More information is obtained when measuring seven points on the retention curve, and fitting the van Genuchten equation. Saturated hydraulic conductivity can be estimated or measured, and $K(\theta)$ relationship is predicted using Mualem-van Genuchten model (van Genuchten, 1980). The most complete information is when measuring both water retention and hydraulic conductivity curves. The inverse method can also be used to predict both water
retention and hydraulic conductivity curves using disc permeameter data (Chapter VII).

Table 8.2.1. Various information on soil physical measurements predicting hydraulic properties and the associate effort and cost

<table>
<thead>
<tr>
<th>Information</th>
<th>Prediction method</th>
<th>time (hours)</th>
<th>cost (A$)</th>
<th>$\sigma \theta(h)$ (m$^3$ m$^{-3}$)</th>
<th>$\sigma K(\theta)$ log$_{10}$[cm day$^{-1}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>min. information</td>
<td>average Australian parameter [PTF]</td>
<td>0.2</td>
<td>0.5</td>
<td>0.15</td>
<td>2.00</td>
</tr>
<tr>
<td>hand texture</td>
<td>average texture class</td>
<td>0.25</td>
<td>5</td>
<td>0.09</td>
<td>1.60</td>
</tr>
<tr>
<td>particle size analysis (PSA)</td>
<td>$\rho_b$ – PTF</td>
<td>48</td>
<td>40</td>
<td>0.06</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>$K(\theta)$ – PTF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSA $\rho_b$</td>
<td>$\theta(h)$ – PTF</td>
<td>48</td>
<td>45</td>
<td>0.04</td>
<td>0.3</td>
</tr>
<tr>
<td>PSA $\rho_b + \theta_{1,10}$</td>
<td>$\theta(h)$ – PTF</td>
<td>336</td>
<td>75</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>PSA $\rho_b + \theta_{1500}$</td>
<td>$K(\theta)$ – PTF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>water retention curve</td>
<td>$\theta(h)$ – van Genuchten $K_c$ – PTF</td>
<td>504</td>
<td>100</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>$K(\theta)$ – Mualem’s model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>van Genuchten curve interpolation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>disc permeameter data</td>
<td>inverse method</td>
<td>24</td>
<td>90</td>
<td>0.03</td>
<td>0.03</td>
</tr>
</tbody>
</table>

The quality of the prediction is measured in terms of the uncertainty ($\sigma$) of the predicted hydraulic properties. For example, using particle size analysis and bulk density, hydraulic PTFs developed in Chapter IV were used to predict the van Genuchten parameters. The error associated with each measurement was incorporated and using the modified Latin hypercube sampling (§5.2.5), 100 values were sampled and the water retention and hydraulic conductivity curves were calculated. The standard deviation of the predicted hydraulic properties (from 100 samples) were then calculated. The efficiency of various methods calculated from the uncertainty ($\sigma$) of predicted $K(\theta)$ over the amount of time needed to obtain the information is summarised in Figure 8.2.1. An efficient method should have a small uncertainty and little effort, i.e. a high efficiency value. It appears that the highest efficiency is obtained by minimum information, followed by hand texture and inverse method. But it should be noted that this only measures the efficiency in terms of effort. The larger
the effort, the lower the efficiency despite the improvement of the uncertainty. To interpret the measure, we have to know the prediction range we want.

![Figure 8.2.1. Efficiency of different methods predicting hydraulic properties in terms of effort.](image)

**8.3. Efficiency 2 – The cost of various methods**

The acquisition of information subject to financial costs, therefore the efficiency can alternatively be represented as the quotient of uncertainty and cost. Figure 8.3.1 shows the cumulative cost of information and the expected uncertainty in predicted water retention and hydraulic conductivity curves. The cost of analysis (Table 8.2.1) is based on the current price charged by laboratories performing physical measurements in Australia. The scenario is a pecking order of information which can be used to estimate hydraulic properties. The coarsest level is when little information is available about the soil. This is followed by hand texture, particle-size analysis, bulk density, water content at -10 and -1500 kPa, water retention curve, measuring $K_s$ in laboratory, measuring near saturation $K$ in the field (using disc permeameter) and $K(\theta)$ curve. With the cumulative amount of information gathered, the cost is also building up as illustrated in Figure 8.3.1a for water retention and 8.1.1b for hydraulic conductivity characteristics.
The efficiency of various methods in terms of cost is given in Figure 8.3.2. As in the previous analysis, this only measures the cost of information. The order of efficiency follows the cost of analysis, the lower the cost the higher the efficiency. To elucidate the efficiency we should also look into the desired accuracy of prediction.

Figure 8.3.1. Relationship between cost of information and the uncertainty in prediction of (a) water retention, and (b) hydraulic conductivity curve.
8.4. Efficiency 3 – Efficiency evaluated by means of a value of information scenario

The two efficiency measures discussed above only deal with measurement. The spatial uncertainty should also be taken into account as the spatial and temporal scale defines the level in which the prediction can be applied. Since one of the objectives of predicting hydraulic properties are as inputs in a soil-water-crop model, the efficiency should be evaluated in terms of the final application. Combining the physical and economical disciplines (i.e. evaluating efficiency in terms of value) seems a powerful way of capturing this.

Information is said to be profitable to business when the value of information exceeds its cost (Swinton and Jones, 1999). Value of information also provides answers to scenarios such as, if we obtain an additional piece of information what is the increase in the expected utility?, or is it worth getting another piece of information?
8.4.1. Perfect irrigation, imperfect information scenario

8.4.1.1. Methods

To evaluate the value of information the mechanistic model, SWAP, simulating soil, water and plant relationship is used. A perfect irrigation (exact amount of water added when the soil is drier than certain potential) with imperfect information on soil hydraulic properties is assumed. The assumed soil is a clay ($P_{<2} = 52$, $P_{2-20} = 16$, $P_{20-2000} = 32$ dag kg$^{-1}$) with a bulk density of 1.35 Mg m$^{-3}$. The hydraulic properties are assumed to fit the van Genuchten function with parameters $\theta_r = 0.05$ m$^3$ m$^{-3}$, $\theta_s = 0.50$ m$^3$ m$^{-3}$, $\alpha = 1.1$ m$^{-1}$, $n = 1.13$ and $K_s = 30$ mm day$^{-1}$. The weather in Bourke, NSW (Figure 4.4.2) is considered again with the same initial and boundary conditions as in §4.4.1. The soil is assumed to be uniform to 1 m depth and under cotton. The simulation starts at 1st May 1997 and the virtual cotton is grown from 1st October 1997 and harvested on 20th April 1998. A simple crop growth (van Dam et al., 1997) model is used to simulate the growth of cotton, specifying the leaf area index, crop height and rooting depth as a function of crop development stage. Water uptake is a function of soil water potential and potential transpiration. At every stage of growth the model calculates the ratio between actual to potential yield from the potential and actual transpiration. The relative yield of the whole growing season is calculated as the product of the relative yields of each growing stage. The evapotranspiration is calculated according to the modified Penman-Monteith method (Smith, 1991).

A perfect irrigation scheme is assumed, when the soil at 100 mm depth is halfway to permanent wilting point (dries to a potential of -700kPa), an amount of water is added to bring the soil in the root zone back to field capacity (-33 kPa).

Different methods for obtaining the hydraulic properties were considered, which climb a ladder of increasing certainty of information. The scenario starts when we only have a minimum information about the soil, followed by hand texture, particle size analysis and bulk density, two points in the retention curve. These measurements require pedotransfer functions for estimating the hydraulic parameters. The next level of information measures the water retention curve, $K_s$ and $K(\theta)$ relationship. The inverse method is also employed to predict the hydraulic parameters. The costs and information are the same as those in Table 8.2.1. Hence each level of information does not necessary yield the same or correct hydraulic parameters.
The modified Latin hypercube method (§5.2.5) is then employed to sample the van Genuchten parameters from the associated uncertainty of each method. One hundred hydraulic parameters were sampled for each level of information and these are fed into the model SWAP. The total amount of water applied during the growing season and the relative yield are recorded for each simulation or realization.

The current cost of water irrigation in NSW cotton growing area is A$10 per ML, and the optimum yield of cotton is assumed to be 2.5 tonnes per ha. The price of cotton is assumed to be US$59 per lb, therefore an optimum yield has the value of A$5595 per ha. The gain per ha is obtained by subtracting the value of the cotton by the cost of the water.

8.4.1.2. Results and discussion

The mean and uncertainty of the gain obtained by different methods is given in Table 8.4.1. The gain is calculated by: \((\text{relative yield} \times \text{price of optimum yield}) - (\text{amount of irrigation} \times \text{cost of water})\), hence only calculates the benefit from water management.

The value of information is obtained by subtracting the gain for each method with the gain when predicted using minimum information. The certainty of predicted \(K(\theta)\) \((1/\sigma)\) plotted against the value of information is shown in Figure 8.4.1a.

<table>
<thead>
<tr>
<th>Information</th>
<th>(\sigma K(\theta)) log (_{10}) [cm day(^{-1})]</th>
<th>(\mu)</th>
<th>(\sigma)</th>
</tr>
</thead>
<tbody>
<tr>
<td>min. information</td>
<td>2.00</td>
<td>2001.568</td>
<td>1705.408</td>
</tr>
<tr>
<td>hand texture</td>
<td>1.60</td>
<td>2586.222</td>
<td>1526.656</td>
</tr>
<tr>
<td>particle size analysis (PSA)</td>
<td>0.70</td>
<td>3374.928</td>
<td>1511.242</td>
</tr>
<tr>
<td>PSA (\rho_b)</td>
<td>0.31</td>
<td>3695.642</td>
<td>884.808</td>
</tr>
<tr>
<td>PSA (\rho_b) + (\theta_{10}, \theta_{1500}) (\theta(h))</td>
<td>0.05</td>
<td>4303.355</td>
<td>174.272</td>
</tr>
<tr>
<td>(\theta(h) + K_s)</td>
<td>0.04</td>
<td>4312.816</td>
<td>174.409</td>
</tr>
<tr>
<td>(\theta(h) + K(\theta))</td>
<td>0.01</td>
<td>4284.99</td>
<td>168.162</td>
</tr>
<tr>
<td>Inverse</td>
<td>0.03</td>
<td>4280.833</td>
<td>204.797</td>
</tr>
</tbody>
</table>
Figure 8.4.1. (a) Value-of-information and its (b) uncertainty as a function of certainty in predicted hydraulic conductivity.
The value of information curve shows an exponential increase (Figure 8.4.1). Under the current water price of $10/ML the maximum value that can be obtained is about $2000 per ha. However as the price of water is bound to increase dramatically in the future, the value of information will be more precious (Figure 8.4.1a). The following analysis only uses current price of water, $10/ML.

For the highest uncertainty there is little value of information, but as we acquire more relevant information the value increases exponentially until reaching the point where we need to measure two points in the retention curve, then the value stabilizes. The economic efficiency as measured by (value – cost) is given in Figure 8.4.2. According to economic theory the optimum amount of information is at the maximum difference between value and cost. The most efficient method appears to be the inverse method and measuring the water retention curve. Measuring both $\theta(h)$ and $K(\theta)$ is too expensive hence not feasible. Efficient methods for a single measurement therefore can be achieved when using particle-size analysis with two points in the retention curve (pedotransfer functions), measuring the retention curve and inverse method.

Figure 8.4.2. Value – cost as a function of the certainty in predicted hydraulic conductivity.
8.4.2. Perfect irrigation, imperfect information scenario incorporating spatial variability

Soil hydraulic properties are known to vary in space (Nielsen et al., 1973). The effect of hydraulic spatial variability on soil-water transport in the field has also been studied (Russo and Bresler, 1981). Usually the spatial variability of soil properties is much larger than the variance of measurements (Viscarra Rossel and McBratney, 1998b). The following scenario attempts to incorporate the spatial variability in examining the efficiency of hydraulic properties.

A virtual field of 100 ha is considered with cotton growing on it under the same conditions as the previous section. The sampling method of McBratney et al. (1981) is employed here. Various measurements at different number of observations were made in the field, and we wish to predict for a grid of 5 m × 5 m. A double-spherical semivariogram is assumed to describe the spatial variability of hydraulic conductivity (Mohanty et al., 1994):

\[
\gamma(dis) = 3.1 \left[ \frac{3}{2} \frac{dis}{30} - \frac{1}{2} \left( \frac{dis}{30} \right)^3 \right] + 3.5 \left[ \frac{3}{2} \frac{dis}{60} - \frac{1}{2} \left( \frac{dis}{60} \right)^3 \right] \quad \text{for } dis \leq 30
\]

\[
\gamma(dis) = 3.1 + 3.5 \left[ \frac{3}{2} \frac{dis}{60} - \frac{1}{2} \left( \frac{dis}{60} \right)^3 \right] \quad \text{for } 30 < dis \leq 60
\]

\[
\gamma(dis) = 3.1 + 3.5 \quad \text{for } dis > 60
\]

where \( \gamma \) is the semivariance in \((\log_{10}[\text{cm day}^{-1}])^2\) and \( dis \) is the distance in m. The variance of each type of measurement was added to the semivariogram. Kriging was used to interpolate the hydraulic properties.

The average standard deviation of the predicted hydraulic conductivity for a point in the field (5 m × 5 m grid) is given in Table 8.4.2. It seems that the high spatial variability masks the uncertainty in predictions. Using hand texture is the least certain, followed by particle size analysis and measuring \( \rho_b \). As we include more relevant measurements (2 points in the retention curve) the uncertainty in the field remains the same despite more accurate measurements being taken.

The standard deviation of the predicted hydraulic conductivity is then used to calculate the value of information according to the relationship as in Figure 8.4.1. This enabled us to calculate the economic efficiency (value – cost) of using different methods for different number of measurements (Table 8.4.3).
Table 8.4.2. Average uncertainty of a point in the field for different methods predicting soil hydraulic properties.

<table>
<thead>
<tr>
<th>No. measurements</th>
<th>hand texture</th>
<th>PSA, $\rho_h$</th>
<th>PSA, $\rho_h + \theta_{10}, \theta_{1500}$</th>
<th>$\theta(h)$</th>
<th>$\theta(h), K_s$</th>
<th>$K(h), K(\theta)$</th>
<th>Inverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>6.37</td>
<td>6.00</td>
<td>5.91</td>
<td>5.91</td>
<td>5.91</td>
<td>5.91</td>
<td>5.91</td>
</tr>
<tr>
<td>42</td>
<td>5.45</td>
<td>5.33</td>
<td>5.30</td>
<td>5.30</td>
<td>5.30</td>
<td>5.30</td>
<td>5.30</td>
</tr>
<tr>
<td>90</td>
<td>5.31</td>
<td>5.19</td>
<td>4.18</td>
<td>5.14</td>
<td>5.14</td>
<td>5.14</td>
<td>5.14</td>
</tr>
<tr>
<td>380</td>
<td>4.73</td>
<td>4.34</td>
<td>4.17</td>
<td>4.18</td>
<td>4.18</td>
<td>4.18</td>
<td>4.18</td>
</tr>
<tr>
<td>900</td>
<td>3.16</td>
<td>2.27</td>
<td>1.91</td>
<td>2.90</td>
<td>2.89</td>
<td>2.89</td>
<td>2.89</td>
</tr>
<tr>
<td>1600</td>
<td>3.16</td>
<td>2.27</td>
<td>1.91</td>
<td>2.90</td>
<td>2.89</td>
<td>2.89</td>
<td>2.89</td>
</tr>
<tr>
<td>2500</td>
<td>2.59</td>
<td>1.66</td>
<td>1.28</td>
<td>1.28</td>
<td>1.28</td>
<td>1.28</td>
<td>1.28</td>
</tr>
<tr>
<td>3600</td>
<td>2.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8.4.3. Efficiency (value-cost) of different methods for a field. Numbers in the brackets represent the standard deviation of the efficiency.

<table>
<thead>
<tr>
<th>No. measurements</th>
<th>hand texture</th>
<th>PSA, $\rho_h$</th>
<th>PSA, $\rho_h + \theta_{10}, \theta_{1500}$</th>
<th>$\theta(h)$</th>
<th>$\theta(h), K_s$</th>
<th>$K(h), K(\theta)$</th>
<th>Inverse</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$5^*$</td>
<td>$45^*$</td>
<td>$75^*$</td>
<td>$100^*$</td>
<td>$130^*$</td>
<td>$300^*$</td>
</tr>
<tr>
<td>12</td>
<td>36896</td>
<td>(22284)</td>
<td>(22916)</td>
<td>(23096)</td>
<td>(23100)</td>
<td>(23099)</td>
<td>(23099)</td>
</tr>
<tr>
<td>42</td>
<td>(40289)</td>
<td>(23971)</td>
<td>(24287)</td>
<td>(24450)</td>
<td>(24450)</td>
<td>(24452)</td>
<td>(24451)</td>
</tr>
<tr>
<td>90</td>
<td>(40563)</td>
<td>(24281)</td>
<td>(24671)</td>
<td>(28032)</td>
<td>(24886)</td>
<td>(24887)</td>
<td>(24887)</td>
</tr>
<tr>
<td>380</td>
<td>(41484)</td>
<td>(25695)</td>
<td>(27124)</td>
<td>(28032)</td>
<td>(28042)</td>
<td>(28042)</td>
<td>(28042)</td>
</tr>
<tr>
<td>900</td>
<td>(42787)</td>
<td>(28202)</td>
<td>(31185)</td>
<td>(33276)</td>
<td>(33860)</td>
<td>(33867)</td>
<td>(33938)</td>
</tr>
<tr>
<td>1600</td>
<td>(43733)</td>
<td>(30709)</td>
<td>(36006)</td>
<td>(39655)</td>
<td>(39687)</td>
<td>(39697)</td>
<td>(39697)</td>
</tr>
<tr>
<td>2500</td>
<td>(43812)</td>
<td>(33470)</td>
<td>(40917)</td>
<td>(46499)</td>
<td>(46515)</td>
<td>(46532)</td>
<td>(46532)</td>
</tr>
<tr>
<td>3600</td>
<td>(42657)</td>
<td>(36099)</td>
<td>(40094)</td>
<td>(40094)</td>
<td>(40094)</td>
<td>(40094)</td>
<td>(40094)</td>
</tr>
<tr>
<td>5041</td>
<td>(38222)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Unit cost of observation

This analysis combines the accuracy of the prediction, cost of obtaining the information and the effect of spatial variability. At a small number of observations (Figure 8.4.3) (< 100 observations per 100 ha) most of the methods have similar efficiency, only complete $\theta(h)$ and $K(\theta)$ measurements are much less efficient. As we increase the number of observations the efficiency can increase or decrease depending on the relative cost of each method.
Chapter VIII – Efficiency of various approaches estimating hydraulic properties

Figure 8.4.3. Value – cost as a function of number of measurements.
8.5. Conclusions

As a summary, Table 8.5.1 shows the rank of the efficiency calculated by the different approaches. When only considering the amount of effort or cost we are willing to pay, hand texture seems to be the most efficient. If we wish to apply the method considering the uncertainty and the value we can get out of it in a single measurement, measuring the water retention curve or inverse method is the most efficient. But to apply the method in a field, the hand texture is the most efficient.

<table>
<thead>
<tr>
<th>Information</th>
<th>quality/effort</th>
<th>quality/effort</th>
<th>value-cost (single point)</th>
<th>value-cost (field)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hand texture [PTF]</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>PSA $\rho_b$ [PTF]</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>PSA $\rho_b + \theta_{10}, \theta_{1500}$ [PTF]</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>$\theta(h)$ [PTF]</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>$\theta(h) + K_s$</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>$\theta(h) + K(\theta)$</td>
<td>7</td>
<td>7</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Inverse</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

Given the large amount of spatial variation of soil hydraulic properties it is perhaps not surprising that we have found lots of cheap information and imprecise measurements are more efficient than a few more expensive precise ones. With increasing technology, such as development of the ‘on-the-go’ proximal soil sensing systems, it allows us to generate more information at lower costs. It has the implications of increased information (or reduce costs), and more efficient representation of the spatial variability. For example, on-the-go near infra-red (NIR) measurements of clay and sand (Viscarra Rossel and McBratney, 1998a) could be more certain than hand texture and potentially less expensive.

The analysis presented here assumes a perfect irrigation scenario, which is hard to find in real life. An imperfect irrigation scenario should also be investigated. A scenario is foreseen where a field with periodic neutron probe (moisture meter) monitoring. Irrigation is scheduled several times (~ 8 times) in a growing season, and when there is no rain for a period of time or water content in the profile falls below some critical value, an approximate amount of water is added to some arbitrary value. This scenario will give more realistic estimates of the value of information.
Chapter VIII – Efficiency of various approaches estimating hydraulic properties

References


IX. General discussion, conclusions and future research

9.1. General discussion and conclusions

*Hydraulic pedotransfer functions*

Pedotransfer functions have been developed for Australian soil to accommodate the supposedly unique properties and the exclusive particle-size classification. Using particle-size distribution and bulk density data, the PTFs can predict water content at different potentials with reasonable accuracy. Pedotransfer functions for saturated hydraulic conductivity have also been developed. Some shortcomings of the PTFs have been identified. The PTFs combined with van Genuchten-Mualem model provide the prediction of water retention and hydraulic conductivity curve. The application of the predicted hydraulic properties depends on the required scale and resolution. For site-specific application, the general PTFs cannot be used to accurately simulate soil-water dynamics. In this case, one might need local calibration.

In the process of establishing or calibrating PTFs, the data set is split into ‘prediction’ and ‘validation’ sets. There is no rule concerning the proportion of the data required in this splitting. To obtain a representative sample for developing PTFs, the ‘prediction’ data should have similar statistics to the ‘validation’ data. This can be tested using the $t$ statistics. In this study the data is arbitrary split into two and the ‘prediction’ and ‘validation’ set appeared not to come from the same population. Nevertheless, the application of the water retention PTFs still quite acceptable (§4.2.2).

*Uncertainty analysis*

A modified Latin hypercube for sampling multivariate correlated variables has been presented and tested to characterize the uncertainty in the prediction of soil-water models. Small error in measurement and spatial variability in input data could result in the propagation of error when estimating soil-water balance. The upper boundary conditions (weather) and the shape of the hydraulic functions have great influence on
simulated soil-water balance. The output of the model is more uncertain during a dry period. The prediction during the soil-water redistribution period is more certain.

*The use of the disc permeameter*

The use of the disc permeameter in the field is limited by the requirement of a good contact between the disc membrane and soil surface. This is usually accommodated by the use of contact sand. The placement of sand has a large effect on cumulative infiltration and estimate of sorptivity. Higher infiltration is observed, but this has a small influence on the steady-state infiltration rate. The use of Wooding’s equation that requires estimation of sorptivity is not recommended. The analysis using multiple tensions, which only requires the steady-state rate is better.

The need for intimate contact between the disc and soil surface resulted in the irreproducible cumulative infiltration curve obtained in the field. The uncertainty of the infiltration curve follows the infiltration pattern, i.e. increase in the cumulative infiltration and decrease in the infiltration rate.

The results have important implications for the inverse modelling of cumulative infiltration data from disc permeameter measurements. The placement of sand could result in inconsistency of the estimated hydraulic model parameters. The cumulative effect of the error in the cumulative infiltration can also result in high uncertainties in the parameter estimates.

*Inverse modelling of disc permeameter data*

Numerical studies have shown that the inverse method from multiple tension disc permeameter data is quite robust in identifying the hydraulic parameters. However application to field data showed that the estimated water retention curve is generally lower than the one obtained in laboratory measurements. Nevertheless the estimated near-saturated hydraulic conductivity matched with Wooding’s analysis. The inverse method can give reasonable estimate of the hydraulic parameters. However there are some experimental and theoretical problems:

- The difficulty in obtaining representative cumulative infiltration curve because of the use of contact sand, which enhances the water infiltration and the irreproducible cumulative infiltration curve as a result of uneven contact between the soil surface and the disc membrane.
• Hydraulic parameters estimated in the field are not necessarily the same as those obtained from laboratory measurements. Hydraulic properties estimated from field measurement should be more representative, as they are measured *in situ*.

• The van Genuchten equation does not always adequately describe both the water retention and hydraulic conductivity curves. Hence the optimised parameters from the inverse solution is a compromise between the two curves.

The most important point is to minimise the difference in infiltration curve, therefore the variation of the hydraulic properties obtained using the inverse method is attributed to the difference in the soil hydraulic properties rather than the due to experimental problems.

**Efficiency**

Most PTFs developed have been designed to obtain information we need from that which we have. Other factors that need to be considered are the analysis of accuracy of prediction, uncertainty in parameters and data, amount of effort or cost needed to obtain the information, and the efficiency of the prediction. The uncertainty analysis is extended to evaluate the efficiency of the different methods in predicting water retention and hydraulic conductivity. The analysis can identify the contribution of individual source of measurement errors to the overall uncertainty. This is beneficial to provide a feasibility analysis to answer questions such as does it pay off (in terms of reduced uncertainty) to use laboratory measurements rather than field estimates.

For single measurements, the inverse analysis is economically more efficient than using pedotransfer functions or increasing measurement of hydraulic properties in the laboratory. However, given the large amount of spatial variation of soil hydraulic properties it is perhaps not surprising that lots of cheap and imprecise measurements are more efficient than a few expensive precise ones.
9.2. Future work

Future work will test the application of hydraulic pedotransfer functions in predicting soil-water balance, crop production, and spatial variability of soil hydraulic properties in the field. Traditionally, soil-water-plant models have been used to predict an average soil-water balance or crop yield. With the recognition of spatial variability and the innovation of precision farming, simulation models can be used to predict site-specific soil-water status and crop yield. While many complex soil-water-plant models have been developed, there is not much effort to use them for predicting crop yield and environmental state, and managing agricultural production. Most research is still directed to developing empirical models relating soil properties to crop yield. This is mainly due to the difficulty in obtaining the parameters in the mechanistic model and also the lack of willingness and expertise to adopt them.

Most research is currently focussing on developing new transfer functions or testing published functions for a certain area or soil and there are no efforts to integrate these available resources in a knowledge-based system. Optimum (in the sense of minimum variance) pedotransfer functions are envisioned to provide the best estimates of soil properties and the associated uncertainties from the basic data that we have. This optimum pedotransfer function will provide information needed by policy makers concerned with an economically sound and environmentally safe agriculture. It is also necessary to supply the uncertainty of estimated properties, thus giving the confidence we can place in predictions. This system not only will apply for Australian soil but also can be adapted elsewhere in the world.

More field experimental work is needed to confirm the application of the inverse method from disc permeameter data. The first step is to perfect the disc permeameter method so that it can be reproducible. The second step is to confirm the applicability of the method, the predicted hydraulic properties should be able to estimate soil-water status in the field.

The efficiency analysis presented in Chapter VIII showed that hand texture combined with pedotransfer functions were quite efficient. This assumed each individual could perform this well. It is clear that texture carries a lot of information. So a cheap reproducible method needs to be developed. Near-infrared (NIR) methods should be further investigated.