Soil saturated hydraulic conductivity assessment from expert evaluation of field characteristics using an ordered logistic regression model

Florencio Ingelmo¹, ², Mª José Molina¹, José Miguel de Paz², Fernando Visconti¹, ²*

¹Centro de Investigaciones sobre Desertificacion-CIDE (CSIC, UVEG, GV), Crta. Moncada-Nàquera Km 4,5, 46113 Moncada, València, Spain.

²Instituto Valenciano de Investigaciones Agrarias-IVIA, Centro para el Desarrollo de la Agricultura Sostenible-CDAS. Crta. Moncada-Nàquera Km 4,5, 46113 Moncada, València, Spain.

*Corresponding author: Tel.: (+34) 963 424 000; fax: (+34) 963 424 001. E-mail address: fernando.visconti@uv.es (F. Visconti).

Abstract

The knowledge of the soil saturated hydraulic conductivity (Ks) is essential for irrigation management purposes and for hydrological modelling. Several attempts have been done to estimate Ks in base of a number of soil parameters. However, a reliable enough model for qualitative Ks estimation based on the expert assessment of field characteristics had not been developed up to date. Five field characteristics, namely macroporosity (M), stoniness (S), texture (T), compaction (C) and sealing (L), in addition to tillage (G) were carefully assessed according to three classes each, in 202 sites in an agricultural irrigated area in Eastern Mediterranean Spain. After the evaluation of field characteristics, a single ring infiltrometer was used to determine the Ks value as the solution of the infiltration equation when the steady state was reached. The distribution of the Ks was assessed and five classes with 10-fold separations in class limits were defined accordingly. The relationships amongst site characteristics and Ks were analyzed through a correspondence analysis (CA). Next, an ordered logistic regression model (OLRM) for the prediction of the Ks class was developed. The CA revealed that, though tightly related, the set of six site characteristics should not be simplified into a smaller set, because each characteristic explains a significantly different aspect of Ks. Consequently, the OLRM was based on the six characteristics, which presented the following order of importance: L > M > G > T > C > S. According to the cross-validation of the OLRM the hit probability for the prediction of the Ks class attained an average value of 50%, which increased to 63% for the highest class of Ks. Moreover, wrong estimation of the Ks class exceeded the ± 1 range only in 3% of sites.
Therefore, a reliable enough assessment of Ks can be based on the expert assessment of field characteristics in combination with an OLRM.
1. Introduction

Saturated hydraulic conductivity (Ks), as a measure of the ability of soil to transmit water, is essential in infiltration-related applications such as irrigation and drainage management (Wu et al., 1999; Radcliffe and Rasmussen, 2002) and for modelling the hydrology of the landscape. This parameter is obviously related to the hazard of ponding and to the potential of soils for tile drainage, which can affect the production of certain crops (McKeague et al., 1982).

Ring infiltrometers are often used for measuring the water intake rate at the soil surface. The total flow rate into the soil from a single-ring infiltrometer is a combination of both vertical and horizontal flow. Wu et al. (1997) found that the infiltration rate of a single-ring infiltrometer was related to the one-dimensional (1-D) infiltration rate for the same soil. For a relatively small ponded head, the 1-D final infiltration rate of a field soil is approximately equal to the field Ks, which is valuable information for computer modelling and irrigation management.

Even with improved equipment, the Ks measurement is time consuming, and thus, models are recommended. Several attempts (Rawls et al., 1982; Tietje and Hennings, 1996; Rawls et al., 1998; Dexter and Richard, 2009) have been made to estimate the Ks from readily available analytical soil data such as particle size distribution, bulk density and organic matter content by means of pedotransfer functions or by physical modelling of the pore size distributions. However, all these estimation methods exhibit large differences between predictions and measurements of Ks (Tietje and Hennings, 1996; Landini et al., 2007), or the hydraulic conductivity close to water saturation could not be estimated based only on the usually available estimators (Weynants et al., 2009).

Models based on soil characteristics such as bulk density and pore size distribution give better predictions as shown by Mbagwu (1995), who estimated Ks from bulk density, macroporosity, mesoporosity and microporosity. Since these models are generally high data demanding and need cumbersome laboratory determinations, the applicability for farmers in irrigation management is reduced. To avoid this, qualitative models based on the expert assessment of morphological characteristics of soil could be an alternative approach to model the Ks. This idea of qualitatively describing water flow through soils has been credited to the Soil Conservation Survey (Norton, 1939). Since then, several models for the qualitative classification of soil ease to permit water flow have been
developed. Mason et al. (1957) developed such a model based on the expert assessment of 14 soil morphologic characteristics in order to classify 900 soils in an ordinal scale of seven permeability classes, defined as the ease with which pores of a saturated soil permit water movement. They attained a hit probability of 30%, and suggested that 95% probability of making a correct prediction could be achieved by using only three to five permeability classes.

McKeague et al. (1982) developed guidelines for estimating the class of saturated hydraulic conductivity of soil horizons from observations of soil morphology in 78 soil horizons ranging in texture from sandy to clayey. The major factors contributing to high Ks values were abundant biopores, textures coarser than loamy fine sand, and strong, fine to medium blocky structure. The lowest values were associated with clayey horizons that had been compressed or puddled by cultivation. The guidelines presented, though incomplete and subjective to some degree, improved the estimates of Ks in limited testing by pedologists. The results also indicated that it was not feasible to assign a unique Ks estimate to near-surface horizons of cultivated soils of a particular series. Tillage practices and current land use have a major effect on soil structure, porosity and density, and hence on Ks.

Saturated hydraulic conductivity can also be related to soil morphological criteria based on the expert assessment and the classes of the Factual Key (McKenzie et al., 2000).

Lin et al. (2006) presented a vision that advocates hydropedology as an advantageous integration of pedology and hydrology for studying the intimate relationships between soil, landscape, and hydrology. Landscape water flux is suggested as a unifying precept for hydropedology, through which pedologic and hydrologic expertise can be better integrated. The discretization of continuous field measurements such as the Ks, is usually of high practical value to perform this integration. The indication of a class of Ks is, on the one hand, more informative, and on the other hand, more stable in space and time than the indication of an X% confidence interval derived from an ordinary least squares regression model.

Given a saturated hydraulic conductivity expressed in an ordinal scale, the datum to predict is not longer the actual value of Ks, but the probability of an observation to belong to a certain class of Ks. This can be adequately performed using logistic regression models (LRM). Logistic regression modelling has been previously used in soil research to assess water erosion from expert evaluation of site characteristics (Sonneveld and Albersen, 1999). The development of a LRM appears as an adequate
methodology for predicting an ordinal variable from other ordinal variables, which to our knowledge has not been carried out up to date for the Ks assessment.

The objective of the present study was to develop a methodology for the estimation of the class of soil saturated hydraulic conductivity based on several field characteristics such as tillage, macroporosity, stoniness, texture, compaction and sealing. This main objective was split into two partial objectives: i) the development of a methodology for the expert evaluation of the soil characteristics, and ii) the development of an ordered logistic regression model for the Ks prediction on basis the six field characteristics.

2. Materials and methods

2.1. Study area

The study area (Figure 1) has 12400 ha, of which approximately 6300 ha are agricultural irrigated lands. Citrus is the main crop with 53% of the irrigated area, followed by vegetables (mainly melon and watermelon) with 14%. Rice crop in lands with shallow watertables accounts for no more than 3% of the area. Citrus orchards and some vegetables are generally drip irrigated. Drip irrigation is used on 65% of the total irrigated area. The climate can be considered as semiarid following the UNESCO classification (De Paw et al. 2000), with annual rainfall of 500 mm and reference evapotranspiration (ET₀) of 1000 mm.

In regard on landscape features, three main areas, associated with soil types, can be distinguished: i) the colluvial and glaci area, with soil materials moved, accumulated, removed or even replaced, especially when calcareous duricrusts can limit the effective soil depth for citrus crops; typical soils there are, respectively, aric Anthrosols and petric Calcisols; ii) the floodplains and alluvial area, with more fertile finer-textured soils such as Luvisols, which are typically cultivated for citrus; iii) the third area located near the coast, characterized by lacustrine fine-grained deposits from infilling of lakes and early used for rice and horticultural crops. In this last area, a watertable is present seasonally, and subsurface horizons of lacustrine soils often show visible greyish colours and prominent reddish mottles when oxidising conditions occur periodically by tile drainage.

2.2. Soil survey
The surveyed plots were selected according to a combined systematic and random point selection in agreement with de Paz et al. (2011), in order to have all soil types and crops represented according to their predominance in the area. The study was carried out during the irrigation season, and in days when the soil water content was close to field capacity, i.e. between 1 and 3 days after irrigation. In each of the 101 plots (Figure 2), two separated points were selected for a total of 202 survey sites. In each site several 10 x 10 cm areas with no vegetation were delimited for the expert assessment of field characteristics, thereafter infiltration was measured in one of the areas.

2.3. Saturated hydraulic conductivity determination

The soil water infiltration rate was measured using a single head ring infiltrometer according to the method by Wu et al. (1999). The single-ring infiltrometer consisted of an infiltration ring 12 cm in diameter and 6.5 cm in height with a calibrated water supply column that maintains a constant water pressure head of 1 cm on the soil inside the ring (Figure 3). The ring insertion depth in the soil was 5 cm. The cumulative infiltration at different times was measured by annotating the height of water in the water supply column. The time of measurement was sufficiently long to achieve a steady-state infiltration rate, which was usually 20 minutes. Calculations of saturated hydraulic conductivity were obtained from Equation 1, Equation 2 and Equation 3 (Wu et al., 1999),

\[ K_s = \frac{A}{(a f)} \quad \text{[1]} \]

\[ f \approx \frac{H+1/a}{G^*} + 1 \quad \text{[2]} \]

\[ G^* = d + r / 2 \quad \text{[3]} \]

where \( A \) is the slope of the cumulative infiltration (cm) vs. time (min.) curve, \( a \) is a constant with value close to 1, \( H \) is the constant ponded head of the infiltrometer (1 cm), \( a \) values were assumed to be 0.36 cm for coarse textured soils, 0.12 cm for loamy soils, and 0.04 cm for clayey soils (Elrick and Reynolds, 1992), \( d \) is the depth of insertion of the ring (5 cm), and \( r \) is the radius of the ring (6 cm).
2.4. Assessment of the field characteristics

The field characteristics credited to be most related to the hydraulic behavior of soils according to Hillel (1998), Porta et al. (1994) and Radcliffe and Rasmussen (2002), include tillage (G), macroporosity (M), texture (T), compaction (C), sealing (L), and stoniness (S). These characteristics were assessed as qualitative parameters by visual (G, M, L, S) and feeling to touch (T) methods (Porta et al., 1994; Milford et al., 2001; Ball and Douglas, 2003; Mueller et al., 2009), and by the resistance to the insertion of the ring of the infiltrometer (C). The results of the visually evaluated characteristics macroporosity and stoniness could be validated by image processing. Although this validation was not formally carried out, the expert classification of macroporosity and stoniness was found to be in accordance to the results of the image processing performed on the zenithal photographs of the soil surface taken in some fields.

2.4.1. Tillage

Management and tillage practices have significant influence on different hydraulic properties, because under agricultural land use, the properties of the macropore system strongly depend on them among others factors (Wahl et al., 2003). Particularly, the saturated hydraulic conductivity of soils is dominated by the micromorphology of soil pores rather than by the total porosity. Micromorphology of pores in topsoils is subjected to continuous disturbance by frequent tillage, while subsoils tend to be compacted without serious changes of micromorphology of soil pores (Nakano and Miyazaki, 2005). Two classes of tillage were defined:

- Class 1 or tilled soils. The soil surface clearly shows tillage practices (Figure 4 left). This always occurs in vegetable gardens and in some citrus orchards with tillage for weeds control.
- Class 0 or non-tilled soils. This kind of soil management occurs for most soils under citrus, either drip or surface irrigated (Figure 4 right).

2.4.2. Texture

Texture was evaluated in the field according to the feeling to touch, the stickiness and the internal cohesion of the soil particles according to the method described by Milford et al. (2001), e.g. a soil sample of about 25 g was held on the palm of the hand, then it
was moldered to form a ball or bolus, adding soil or water until the bolus failed to stick to the fingers. The behavior of the bolus and of the ribbon produced by pressing out between thumb and forefinger and the feel of the material between both fingers were used to classify texture. Three textural classes were used, each one representing a general increment from more to less ease for soil tillage: coarse (1), medium (2) and fine (3). In Table 1 the correspondences between the textural classes obtained with the Milford et al. (2001) method and the USDA classes are shown.

2.4.3. Compaction

Soil compaction affects hydraulic properties, and thus can lead to soil degradation and other adverse effects on environmental quality (Zhang et al., 2006). Soil compaction changes the ability of soil to hold water, decreases infiltration rate and saturated hydraulic conductivity, and increases penetration resistance (Shafiq et al., 1994). Given that the penetration resistance decreases with water content and increases with bulk density, a qualitative field evaluation of soil compaction was assessed during the insertion of the ring of the infiltrometer in a fresh vertical exposure when the soil was near field capacity. The strength of the soil was judged from the time required to insert the ring just 5 cm into the soil when the compressive force of a nylon hammer applied at a height of 50 cm strikes on the coverlid of the ring, which is equal in diameter. This estimate of compaction could be given from measurement of the bulk density, but could not be substituted by a penetrometer resistance measurement. The use of the ring was preferred over a penetrometer because the former gives averaged information of the area subsequently used for the infiltration assay, while the latter gives just point information, which is very variable. Depending on the time needed to insert the cylinder the compaction class was estimated as follows:

- Class 1 or low compaction: the time was less than 1 minute.
- Class 2 or medium compaction: the time was within 1 to 3 minutes.
- Class 3 or high compaction: the time was longer than 3 minutes.

2.4.4. Macroporosity

Macroporosity is defined as the abundance of voids with diameters between 1 and 5 mm. It includes the modification of the soil surface morphology, in particular structure, either by soil fauna, mainly microarthropods and earthworms (Weiler and Naef, 2003),
plant roots (Holden, 2005) or by permanent cracks between peds that usually do not close on wetting. Despite their low abundance, macropores can have a significant impact on infiltration and runoff (Jarvis, 2007), and play an important role in the rapid transmission of water, known as preferential or bypass flow, in both the topsoil and the subsoil (Franklin et al., 2007). It may result in the lowering of the bulk density and subsequent increasing of the water infiltration rate. Therefore, the visible abundance of both biota-made macropores (circular voids, Figure 5 top row) and cracks (planar voids, Figure 5 bottom row) is jointly described (Dexter and Richard, 2009). Three classes of macroporosity in the site areas of 10 x 10 cm were evaluated according to Figure 6, which is based on several results from the literature (Lachnicht et al., 1997; Buczko et al., 2006; Holden and Gell, 2009):

- Class 1 or low macroporosity: 0 to 3 voids per dm².
- Class 2 or medium macroporosity: 3 to 6 voids per dm².
- Class 3 or high macroporosity: more than 6 voids per dm².

2.4.5. Sealing

Unstable materials high in silt and fine sand, or fine textured well-structured soils may slake, swell and disperse. Under such surface conditions, porosity and water infiltration decrease (Laj et al., 2001). The sealing variable is a simple field assessment of the effects of aggregate instability and dispersibility induced by texture and structure conditions of soil as well as by irrigation water quality resulting in the accumulation of fine silt and fine sand or dispersed clay and salts on the soil surface forming seals and crusts of mechanical origin (Figure 7 left). Under shaded areas of drip-fertirrigated citrus, this kind of seals and crusts enhance algae, mosses and lichens colonization, and contribute to form biological crusts that usually are water repellent (Figure 7, center and right). Both types of seals have been jointly evaluated. Three sealing classes were defined according to seal thickness and the percentage of covered soil surface (Figure 8) in the site areas of 10 x 10 cm:

- Class 1 or low sealing: the seal is < 1 mm thick and covers less than 10% of the soil surface.
- Class 2 or medium sealing: the seal is 1-5 mm thick and covers between 10 and 60% of the soil surface.
2.4.6. Stoniness

Three classes of stoniness were estimated from the observation of the percentage of the site areas of 10 x 10 cm covered by rock fragments according to Figure 8 (Daniels et al., 1968; Poesen et al., 1990):

- Class 1 or low stoniness: 0 - 25%
- Class 2 or medium stoniness: 25-50%
- Class 3 or high stoniness: 50-75%

Field examples of the three classes of stoniness are shown in Figure 9.

2.5. Correspondence analysis (CA)

The field characteristics are related to each other as has been stated when describing their estimation in the previous section, e.g. tillage, macropororosity and compaction. Therefore, a correspondence analysis (CA) was applied to find out to what extent the characteristics were associated and if the likely information redundancies could be reduced.

An exploratory study of the multiple dimensional contingency table of the \( n = 6 \) independent variables and Ks was carried out by means of CA. Particularly, the matrix used for such analysis was not the contingency table itself but the matrix of scores of \( s \) sites x \( n \) field characteristics. Similarly to principal components analysis (PCA) the eigenvectors of the matrix were obtained in order, from the one that accounted for the highest variance, or inertia in CA terms, to the lowest. Then, the coordinates of the different classes of the field characteristics and those of the Ks classes were calculated, which allowed to project them on a Euclidean space and interpret the relationships among them.

The CA was carried out using the statistical package R (The R Development Core Team, 1999) and graphics drawn with a Microsoft Excel© spreadsheet.

2.6. The ordered logistic regression model (OLRM)

There are several approaches to logistic modelling. However, when the outcome represents an underlying continuous scale subdivided in several categories, as is the
case with Ks, the most adequate modeling framework is a cumulative approach (Fullerton, 2009). The generalized ordered logistic model for the cumulative approach, when modeling a variable $z$ split in $m + 1$ categories from $n$ independent variables, is given by the following equations (Equation 4 and Equation 5).

$$y = \alpha + \beta x \quad [4]$$

$$y_i = \log\left(\frac{\Pr(z \geq \omega_i)}{\Pr(z < \omega_i)}\right) \quad [5]$$

In Equation 4 $y$ is a $m \times 1$ vector of logits, each one given by Equation 5. A logit is defined as the logarithm of the quotient of the probability of being in a class with $z$ equal or higher than a cut value $\omega_i$ to the probability of being in a category with $z$ lower than $\omega_i$. This quotient is called odds for short. In Equation 4 $\alpha$ is a $m \times 1$ vector of intercepts, $x$ is the vector of $n$ independent variables, and $\beta$ is a $m \times n$ matrix of regression coefficients to be estimated. The marginal probability ($\Pr$) of being in each category is given by a set of $m + 1$ equations with the following structure.

$$\Pr(z < \omega_1) = 1 - \Pr(z \geq \omega_1) = 1 - F(\alpha_1 - \sum_{i=1}^{n} \beta_{i,1} x_i) \quad [6]$$

$$\Pr(\omega_j \leq z < \omega_{j+1}) = \Pr(z \geq \omega_j) - \Pr(z \geq \omega_{j+1}) =$$

$$F(\alpha_j - \sum_{i=1}^{n} \beta_{i,j} x_i) - F(\alpha_{j+1} - \sum_{i=1}^{n} \beta_{i,j+1} x_i) \quad [7]$$

$$\Pr(z \geq \omega_m) = F(\alpha_m - \sum_{i=1}^{n} \beta_{i,m} x_i) \quad [8]$$

Where the first and last equations (Equation 6 and Equation 8) are for calculating the probability of being in the extreme categories, lowest and highest respectively, and Equation 7 is for the intermediate categories with $j$ between 1 and $m - 1$. In all equations $F$ is the logistic cumulative density function (cdf).

The identification of significant variables for each logit of $y$ was carried out by using a stepwise selection procedure. The decision to include a variable was based on the significance of the log-likelihood ratio of the estimation and the $\chi^2$ statistical test of the variables. In each selection round, the variable that gave the most significant improvement of the log-likelihood was included in the model. After a variable was
included, it was tested to see whether the exclusion of a variable included at an earlier stage gave a significant decline of the log-likelihood. This process was terminated when, on the one hand, the inclusion of an extra variable did not lead to a significant improvement of the model, and on the other, exclusion of an included variable did not lead to a significant worsening. The level of significance for acceptance and rejection in the stepwise selection was 0.05 and 0.10, respectively.

The statistical package SPSS 17.0 (SPSS Inc. An IBM Company, Chicago, Illinois) was used for carrying out the logistic regression analyses.

2.7. Cross-validation of the model

The OLRM was subjected to a 10-fold cross-validation analysis in the study area. The 202 sites were randomly allocated to ten groups of equal size, that is, 20 sites per group, which constituted ten validation data sets. For each validation data set there was a different training data set made up of 182 sites. One OLRM was developed on each training data set and the class of saturated hydraulic conductivity in the 20 sites of the corresponding validation data set was assessed. In total 200 (20 x 10) independent predictions of saturated hydraulic conductivity were obtained.

3. Results

3.1. Magnitude and distribution of the saturated hydraulic conductivity

The saturated hydraulic conductivity presents a minimum of 0.04 mm h$^{-1}$ and a maximum of 1374 mm h$^{-1}$ following an exponential distribution as shown in Figure 10 left. Taking the logarithm of Ks to base 10, a normal-like histogram is obtained as shown in Figure 10 right. The null hypothesis of a normal distribution for log Ks was tested using the Kolmogorov-Smirnov test. The maximum difference between the sample and the normal cumulative distribution frequencies is 0.89, with an associated p value of p = 0.41. According to this result a log-normal distribution can be assumed for Ks and, therefore, an ordinal scale of 5 levels each one representing a 10-fold increment in the saturated hydraulic conductivity can be adequately used: < 0.3, [0.3, 3.6[, [3.6, 35[, [35, 350[ and ≥ 350 mm h$^{-1}$. This ordinal scale for Ks classification is almost the same in use by the US National Soil Survey Staff (USDA, 2010) except for the two
lowest Ks classes (< 0.036 and [0.03, 0.3]) which have been merged. According to these Ks ranges, 15, 41, 90, 47 and 9 sites are respectively located within each order of Ks, giving a symmetrical graph (Figure 11). Saturated hydraulic conductivity classifications where each class limit represents a 10-fold increment regarding the previous were early proposed at the 1971 NCSS Conference (USDA, 1971). They proposed five classes separated by limits 0.4, 4.2, 42 and 420 mm h\(^{-1}\). Such log-like classifications are appropriate if Ks follows a log-normal distribution; this was observed in our study area.

3.2. Correspondence analysis of site characteristics

Each class of the field characteristics texture, compaction, macroporosity, stoniness and sealing in addition to tillage was observed in at least 15 sites and no more than 159 (Table 2), this latter corresponding to non-tilled soils. No-tillage is a common practice for citrus orchards, which predominate over vegetables in the study area. As five field characteristics were measured in an ordinal scale of three levels each (texture, compaction, macroporosity, stoniness and sealing), and tillage of only two, a total of 486 (3^5 x 2) different combinations were possible. However, 396 of these combinations were not observed (82%), whereas 90 were observed between one and eight times: 48 combinations were observed once, 23 were observed twice, 11 three times, 8 four times, 3 five times, 1 six times, other 2 seven times, and one combination was observed eight times. Thus, the levels of the field characteristics were somewhat associated to each other, i.e. the variables are dependent on one another.

The associations among the different classes of the explanatory variables were studied through correspondence analysis (CA) including also the dependent variable Ks class. The first two factors were the only having an associated eigenvalue higher than 0.5. Two coordinates for each class of the site characteristics were calculated to give a two dimensional map of variables (Figure 12). The first factor explained mainly the variability associated with tillage and texture, and to a lesser extent to low and medium compaction and medium and high macroporosity. Based on the cluster formed by G1, T1 and C1, the tilled soils are often light textured and low compacted and, given the proximity of S1 and L1, also usually have low stone content and low sealing. The soil surface of low compacted soils has usually continuous and clearly visible connecting pores and cracks that do not close on wetting. The structure presents strong, stable, cemented aggregates and nodules and gravels are scarce. Given the cluster formed by
G0, T2, T3, C2, C3, S2 and S3, the non-tilled soils are predominantly medium to heavy textured, medium to high compacted and medium to high stony. The structure is moderate and blocky, polyhedral peds are common. If coarser, the inter-particle voids are filled with fine minerals. There are few visible (hand lens) pores that conduct water when wet. The pores and channels which remain open when wet are visible. Rock fragments are commonly present.

On the other hand, the second factor explained the variability associated to stoniness, low and medium macroporosity and medium and high sealing. The variability of macroporosity, sealing and stoniness is interestingly associated with the second factor. Low macroporosity is associated with medium to high sealing, medium macroporosity with low to medium sealing and high macroporosity with low sealing. The association between macroporosity and stoniness, although weaker than the preceding one, is also present. High and medium stoniness is associated with high and medium macroporosity, whereas low stoniness is associated with low macroporosity.

Stone fragment in soils modifies the pore space in the contact with the surrounding fine-earth fraction (Sauer and Logsdon, 2002). This effect influences water flow by affecting hydraulic continuity near fragment surfaces. These relatively subtle morphological factors may have disproportionate impact on water flow under near-saturation conditions in these soils, and variability in space and time in water infiltration rates could be mainly attributed to the influence of the stone fragment content (Verbist et al., 2009). Ravina and Magier (1984) showed that the increase in coarse fragment content increased the resistance to compaction and, consequently, preserved lower soil bulk density. It was concluded that coarse fragments contributed to improved physical conditions by acting as a "skeleton," which resists soil compaction. As suggested by Ingelmo et al. (1994), larger volumes of large pores in soils with coarse fragments, can lead to an increase of the effective porosity of the fine earth fraction of the soil and to a highest hydraulic conductivity between fragments.

Given the proximity among L3, M1, C3 and T3, high sealing is associated with low macroporosity, high compaction and heavy texture. Nevertheless, this association is weaker than the previous. In high compacted soils structural peds are indistinct and barely observable, or soil occurs as a coherent mass showing no evidence of any distinct arrangement of soil particles. There is absence of visible (hand lens) pores that could conduct water when wet. Cracks and macropores between peds, when present, close on wetting.
The saturated hydraulic conductivity variability is accounted for by both the first and the second factors, however, each factor accounts for the variability among different Ks classes. The first factor, which is mainly related to tillage, texture and compaction, accounts for the variability from the medium to the highest level of Ks (3, 4 and 5). The highest level of saturated hydraulic conductivity (5) is almost only found in tilled soils (8 out of 9 soils, see supplementary material 1). Nevertheless, since not all tilled soils (G1) exhibit an extremely high Ks (K5), the simultaneous presence of other field characteristics is decisive for the appearance of the highest level of Ks. These are light texture (T1) and low compaction (C1), which are typical of tilled soils but can also occur under no tillage.

The closeness between the high level of Ks (K4) and the highest level of macroporosity (M3) suggests this last as the most important characteristic for the appearance of high infiltration rates. The appearance of high macroporosity occurs in 52 sites and these are usually non-tilled (69%), however, independently of the tillage practices, the appearance of high macroporosity is almost always accompanied by medium to high levels of Ks (96%).

On the other hand, the second factor accounts for the variability from the medium to the lowest level of Ks (1, 2 and 3). Given the proximity between K1 and L3 and also M1, the lowest levels of Ks are mainly found in highly sealed soils with low macroporosity.

The six field characteristics, which have been sought to become explanatory variables for the class of Ks, are related to each other and with Ks as it has been shown. However, the cumulative percentage of inertia, i.e. the percent of variance explained by the first two factors was only 19% (Figure 13). It only reached more than 50% when considering seven factors, which is the number of variables used for the CA (6 explanatory + Ks).

This suggests that no variable could be suppressed a priori from the model estimation without loss of significant information. In the same way, no other set of variables derived from the six field characteristics by means of a mathematical transformation could be used instead of them without loss of significant information. Although tightly related, each field characteristic seems to explain a significant different part of Ks variability and, therefore, should be kept in the model.

3.3. Model estimation
The five classes of saturated hydraulic conductivity were expressed by means of four dummy variables as indicated in Table 3. Then, four logit regression equations were calibrated with the following structure (Equation 9).

$$\ln \frac{K_s}{1 - K_s} = a_i + b_{g_i}G + b_{c_i}C + b_{m_i}M + b_{s_i}S + b_{l_i}L$$  \[9\]

The estimation of the regression coefficients is shown in Table 4. As we expected from experience and the CA, tillage, macroporosity and stoniness have positive regression coefficients in every equation while the contrary occurs with texture, compaction and sealing. Accordingly, the higher classes of $K_s$ are more likely in tilled than non-tilled soils, and in soils with higher macroporosity and stoniness. On the other hand, lower levels of $K_s$ were observed with finer texture, and higher compaction and sealing. The magnitude, and where appropriate, the significance of the regression coefficients follow different trends for each characteristic.

The tillage variable is not significant for explaining the variability between the lowest class of hydraulic conductivity, which was found mainly under no-tillage, and the rest of classes. Nevertheless, the tillage is significant to distinguish between the lowest (1, 2, 3) and the highest levels (4, 5) of $K_s$. Moreover, its magnitude increases for the distinction between the extremely high and the rest of levels.

The texture is not significant to distinguish the extremely low (1) from the rest (2, 3, 4, 5), but it is significant to distinguish among these. The compaction is not significant to explain the variability between the extremely high and the rest of classes of hydraulic conductivity, but so is to explain the variability among these. Moreover, its magnitude increases as $K_s$ decreases. The macroporosity is not significant to distinguish the extremely low class of $K_s$ from the others. Nevertheless, its magnitude increases for the distinction among the higher classes of $K_s$. The stoniness is only significant for the distinction between the extremely low and low classes of $K_s$ from the others. Finally, the sealing is not significant for the distinction between the extremely high class of $K_s$ and the others. However so is to explain the variability among the lower classes, particularly to distinguish between the extremely low to medium (1, 2, 3) and high to extremely high (4, 5). The differences among the regression coefficients across the logit equations were found to be significant according to the test of parallel lines. Given this result a simplified model with a common set of regression coefficients ($b_{g_1}, b_{c_1}, b_{m_1}, b_{s_1}$,
which would be based on the so-called proportional odds hypothesis (parallel regression) should not be adequate. Such a model was, therefore, discarded. According to the magnitude of the regression coefficients the six field characteristics can be arranged from the most to the least important on determining the class of Ks: sealing > macroporosity > tillage > texture > compaction > stoniness.

3.4. Model self-evaluation

The class of saturated hydraulic conductivity in each site was assessed calculating the probability of each class of Ks (see supplementary material 2). The expected class of Ks was considered to be, therefore, the most likely one. According to the comparison of predictions and observations, the class of Ks was correctly calculated for 125 out of 202 sites, i.e. 62% of right calculations. The highest percentage of right calculations was attained for the medium class of Ks (71%), followed by the low class (68%), next the extremely high class (56%), and then the high class (51%). Conversely, the extremely low class of Ks was the only predicted with percentage less than 50%, to be precise 40%.

In 71 out of 77 sites (92%) where the class of Ks was not correctly calculated, the predicted Ks class was only one class higher or lower than the observed class. In the remaining six sites the predicted Ks was two classes higher than the observed one (extremely low and low Ks) or two classes lower (medium Ks). There were not differences higher than two between observed and predicted classes of Ks.

3.5. Cross-validation of the model

According to the cross-validation the class of hydraulic conductivity was independently assessed in 200 sites (see supplementary material 3). The class of Ks was correctly predicted in 99 sites, i.e. 50% of right calculations. As shown in Figure 11, the highest percentage of right calculations was attained in the highest and medium classes of Ks with 63 and 59% of right calculations, respectively. These were the only classes with hit percentages higher than 50%. The low, high, and lowest classes of Ks were predicted with hit percentages of 45, 38 and 33%, respectively.

In 95 out of the 101 sites where the class of Ks had not been correctly predicted (94%) the difference between the observed and the predicted class of Ks was one in absolute
value. In the remaining six sites the difference was two in absolute value. Differences in absolute value of three or larger did not appear (Table 5).

Finally, in 29 out of 200 sites the probability of the most likely class of Ks did not surpass 0.5. In order to have an individual hit probability of at least 0.5, more than one class of Ks had to be predicted for these 29 cases. Only two classes of Ks were enough, and the percentage of right predictions attained 57% (113 out of 200).

4. Conclusions

A methodology for the qualitative classification of the soil saturated hydraulic conductivity based on the evaluation of field characteristics has been developed. Based on the current knowledge on soil hydraulic conductivity, five field characteristics namely, macroporosity, sealing, texture, compaction and stoniness, in addition to tillage were selected, and a methodology for their field expert classification in three classes each was established. Particularly, a new methodology for compaction assessment was described. According to a correspondence analysis, several remarkable associations among the six field characteristics were revealed. However, the characteristics explained different aspects of soil saturated hydraulic conductivity, i.e. each class acts at different levels of the Ks, and simplification into a smaller set of variables would not have been adequate. Therefore, the six characteristics were subsequently used as a set of six independent variables to develop an ordered logistic regression model for the estimation of five classes of Ks with 10-fold separations in class limits. The cross-validation of the OLR model gave a hit probability of 50% with error estimations seldom outside the ± 1 range. Therefore, a reliable enough assessment of Ks can be based on the expert assessment of field characteristics plus the use of an ordered logistic regression model.

Supplementary material 1. Class of the field characteristics and Ks on each site.
Supplementary material 2. Calculated probability of each Ks class in each site according to the self-evaluation.
Supplementary material 3. Calculated probability of each Ks class in each site according to the cross-validation.

Acknowledgments
This work has been done in the framework of projects CGL2006-13233-CO2-01 and CGL2006-13233-CO2-02. The authors acknowledge the Ministerio de Educación y Ciencia from Spain their financial support. We would like to thank the two anonymous reviewers for their constructive comments.

References


The R Development Core Team, 1999. The Comprehensive R Archive Network.

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.


Captions of tables and figures

Table 1. Correspondence between the textural classes obtained with the Milford et al. (2001) method and the USDA texture classes.
Table 2. Number of observations for each class of the site characteristics.
Table 3. Codification algorithm for the saturated hydraulic conductivity as four dummy variables.
Table 4. Regression coefficients obtained for the four binary logistic regression models.
Table 5. Number of predictions against number of observations in each class of Ks according to the cross-validation.

Figure 1. Location of the study area.
Figure 2. Distribution of the 101 selected plots in the study area.
Figure 3. Scheme and field photograph of the constant head ring infiltrometer used to carry out the infiltration measurements.
Figure 4. Field examples of tillage in the study area.
Figure 5. Field examples of macroporosity evaluation in the study area.
Figure 6. Model sheets to evaluate the soil macroporosity: the black points represent the voids.
Figure 7. Field examples of sealing evaluation in the study area.
Figure 8. Model sheets to evaluate the surface cover of sealing and stone contents.
Figure 9. Field examples of stoniness evaluation in the study area.
Figure 10. Distribution of the saturated hydraulic conductivity in the study area.
Figure 11. Number of wrong and right predictions for each class of observed Ks according to the cross-validation.
Figure 12. Projection of the classes of the field characteristics on the plane given by the first and second factors extracted from the correspondence analysis. Successive Ks classes are linked by broken lines.
Figure 13. Cumulative percent of inertia and eigenvalue of the factors obtained in the correspondence analysis.
Figure 1. Location of the study area.
Figure 2. Distribution of the 101 selected plots in the study area.
Figure 3. Scheme and field photograph of the constant head ring infiltrometer used to carry out the infiltration measurements.
Tilled soil (class 1).

Figure 4. Field examples of tillage in the study area

Non-tilled soil (class 0)
Figure 5. Field examples of macroporosity evaluation in the study area.
Figure 6. Model sheets to evaluate the soil macroporosity: the black points represent the voids.
Figure 7. Field examples of sealing evaluation in the study area.
Figure 8. Model sheets to evaluate the surface cover of sealing and stone contents.
Figure 9. Field examples of stoniness evaluation in the study area.
Figure 10. Distribution of the saturated hydraulic conductivity in the study area
Figure 11. Number of wrong and right predictions for each class of observed $K_s$ according to the cross-validation.
Figure 12. Projection of the classes of the field characteristics on the plane given by the first and second factors extracted from the correspondence analysis. Successive Ks classes are linked by broken lines.
Figure 13. Cumulative percent of inertia and eigenvalue of the factors obtained in the correspondence analysis
Table 1. Correspondence between the textural classes obtained with the Milford et al. (2001) method and the USDA texture classes.

<table>
<thead>
<tr>
<th>Class with Milford et al. (2001) method</th>
<th>Equivalent USDA texture class</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sand, loamy sand, clayey sand, sandy loam</td>
</tr>
<tr>
<td>2</td>
<td>Loam, silty loam, sandy clay loam, silty clay loam</td>
</tr>
<tr>
<td>3</td>
<td>Sandy clay, silty clay, clay loam, clay</td>
</tr>
</tbody>
</table>
Table 2. Number of observations for each class of the site characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol</th>
<th>Code</th>
<th>Name (range if appropriate)</th>
<th>Number of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tillage</td>
<td>G</td>
<td>0</td>
<td>Non-tilled</td>
<td>159</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>Tilled</td>
<td>43</td>
</tr>
<tr>
<td>Texture</td>
<td>T</td>
<td>1</td>
<td>Light</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Medium</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Heavy</td>
<td>93</td>
</tr>
<tr>
<td>Compaction</td>
<td>P</td>
<td>1</td>
<td>Low</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Medium</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>High</td>
<td>50</td>
</tr>
<tr>
<td>Macroporosity</td>
<td>M</td>
<td>1</td>
<td>Low</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Medium</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>High</td>
<td>52</td>
</tr>
<tr>
<td>Stoniness</td>
<td>S</td>
<td>1</td>
<td>Low</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Medium</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>High</td>
<td>75</td>
</tr>
<tr>
<td>Sealing</td>
<td>L</td>
<td>1</td>
<td>Low</td>
<td>152</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Medium</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>High</td>
<td>15</td>
</tr>
<tr>
<td>Sat. Hyd. Cond.</td>
<td>K</td>
<td>1</td>
<td>Extremely low (Ks &lt; 0.3 mm h⁻¹)</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Low (0.3 ≤ Ks &lt; 3.6 mm h⁻¹)</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Medium (3.6 ≤ Ks &lt; 35 mm h⁻¹)</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>High (35 ≤ Ks &lt; 350 mm h⁻¹)</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>Extremely high (Ks ≥ 350 mm h⁻¹)</td>
<td>9</td>
</tr>
</tbody>
</table>
Table 3. Codification algorithm for the saturated hydraulic conductivity as four dummy variables

<table>
<thead>
<tr>
<th>Dummy variable</th>
<th>Condition for $K_s = 1$ / mm h⁻¹</th>
<th>Condition for $K_s = 0$ / mm h⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_{s1}$</td>
<td>$K_s \geq 0.3$</td>
<td>$K_s &lt; 0.3$</td>
</tr>
<tr>
<td>$K_{s2}$</td>
<td>$K_s \geq 3.6$</td>
<td>$K_s &lt; 3.6$</td>
</tr>
<tr>
<td>$K_{s3}$</td>
<td>$K_s \geq 35$</td>
<td>$K_s &lt; 35$</td>
</tr>
<tr>
<td>$K_{s4}$</td>
<td>$K_s \geq 350$</td>
<td>$K_s &lt; 350$</td>
</tr>
<tr>
<td>i</td>
<td>constant</td>
<td>Tillage</td>
</tr>
<tr>
<td>----</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>1</td>
<td>7.474</td>
<td>ns</td>
</tr>
<tr>
<td>2</td>
<td>5.194</td>
<td>2.122</td>
</tr>
<tr>
<td>3</td>
<td>18.132</td>
<td>1.096</td>
</tr>
<tr>
<td>4</td>
<td>-8.065</td>
<td>2.969</td>
</tr>
</tbody>
</table>

ns: not significantly different from zero
Table 5. Number of predictions against number of observations in each class of Ks according to the cross-validation.

<table>
<thead>
<tr>
<th>predicted</th>
<th>&lt; 0.3</th>
<th>[0.3, 3.6]</th>
<th>[3.6, 35]</th>
<th>[35, 350]</th>
<th>≥ 350</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0.3</td>
<td>5</td>
<td>7</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>[0.3, 3.6]</td>
<td>4</td>
<td>18</td>
<td>17</td>
<td>1</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td>[3.6, 35]</td>
<td>1</td>
<td>19</td>
<td>53</td>
<td>16</td>
<td>1</td>
<td>90</td>
</tr>
<tr>
<td>[35, 350]</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>18</td>
<td>5</td>
<td>47</td>
</tr>
<tr>
<td>≥ 350</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Totals</td>
<td>10</td>
<td>44</td>
<td>97</td>
<td>38</td>
<td>11</td>
<td>200</td>
</tr>
</tbody>
</table>