Combining Visible–Near-Infrared and Pedotransfer Functions for Parameterization of Tile Drain Flow Simulations

Ioannis Varvaris,* Zampela Pittaki-Chrysodonta, Per Moldrup, Lis Wollesen de Jonge, and Bo V. Iversen

Estimation of soil hydraulic parameters is essential when generating a hydrogeological model for simulating water flow dynamics in an agricultural field. However, estimation of the input parameters through direct measurements is time consuming and costly, and the spatial variability presents an uncertainty. Therefore, we proposed a rapid and inexpensive concept (integration of visible–near-infrared spectroscopy [vis-NIR] and a pedotransfer function [PTF]) to estimate hydraulic properties considering catchment scale. An existing vis-NIR–predicted Campbell retention function was used for estimating the Campbell \(b\) parameter and the water content at \(-1000\) cm \(H_2O\) soil–water matric potential \((\log|−1000| = pF 3)\). A PTF was developed for predicting the saturated hydraulic conductivities using the vis-NIR–predicted Campbell \(b\) and the effective porosity, defined as the difference in volumetric water contents at \(pF 0.3\) and \(3\). The concept was evaluated by developing a hydrogeological model in HYDRUS-2D software for simulating the tile drainage discharge from a clayey agricultural subcatchment in Denmark, using as input hydraulic parameters the output from the suggested approach. The suggested approach simulated the main attributes of the flow hydrograph with a reasonable degree of accuracy \((R^2\) and \(RMSE\) values of 0.86 and 1.25 L s\(^{-1}\), respectively). A sensitivity analysis was performed to determine the response of the model to changes in values of predicted parameters when predicting the drainage discharge, and it showed that small variations (<10%) would not affect the predictive ability of the model.

Abbreviations: EM, electromagnetic; NSE, Nash–Sutcliffe model efficiency coefficient; OM, organic matter content; PTF, pedotransfer function; REV, representative elementary volume; SWRC, soil water retention curve; vG, van Genuchten; vis-NIR, visible–near-infrared.

Fertilizers and pesticides are intensively used in agriculture in Denmark to increase productivity. However, these practices are increasingly causing concern. According to the Danish National Groundwater Monitoring Program, pesticides have been detected in 48% of all screens monitored and in 55% of the screens placed in the upper groundwater (Kjær et al., 2009). In artificial subsurface drained agricultural fields, pesticide and nitrification-derived nitrate transport to tile drains may present rapid and uneven breakthrough curves because of preferential flow, which leads to short residence times in the porous medium and direct discharge to surface waters during intensive rainfall events (Branger et al., 2009; Jørgensen et al., 2004; Larsson and Jarvis, 1999; Mohanty et al., 1998; Norgaard et al., 2013; Paradelo et al., 2016).

Vanclooster and Boesten (2000) suggested that pesticide risk assessment should be performed using validated solute transport models. However, an accurate model for simulating the fate of pesticides and nutrients in the vadose and saturated zone requires a water flow model with good predictive ability. The main challenge with respect to water flow model parameterization is how to develop a simple model with enough physical basis that represents the actual internal flow and transport pathways within a field. The main obstacle to this is the lack of sufficient field data on soil hydraulic properties. Determination of hydraulic properties through direct measurements is time consuming and costly, and often the lack of high-quality measured soil physical data is a source of uncertainty. The
spatial variability of the hydraulic properties within a field adds significant uncertainty in relation to the modeling results (Gómez-Hernández and Gorelick, 1989; Nielsen et al., 1973; Peck et al., 1977; Wood et al., 1988). The spatial variability of the hydraulic properties is linked to the soil physical properties, which at field scale is related to the soil-forming factors where meteorological conditions, organisms, topography, and parent material are dominant factors (Jenny, 1941). Therefore, prediction models that incorporate the spatial variability in soil physical properties within a field are considered important hydrogeological tools that cover the absence of knowledge regarding the parameterization of water flow models.

Many researches have shown the potential of using pedotransfer functions (PTFs) as viable tools for predicting hydraulic parameters from available directly or indirectly measured soil properties, such as clay-size fraction and organic matter (OM) content (Borgesen et al., 2008; Iversen et al., 2011; Stolf et al., 2011). Additionally, several studies have investigated the ability of PTFs to represent the spatial variability of soil hydraulic properties (da Silva et al., 2017). Pedotransfer functions have shown good predictive ability for estimating saturated hydraulic conductivity ($K_s$) and the field-average potential water flow through the porous medium (Gärdenäs et al., 2006; Iversen et al., 2011; Nielsen et al., 2018). Iversen et al. (2011) developed reasonably accurate PTFs for predicting $K_s$ and near-saturated hydraulic conductivity at a matric potential of $-10 \text{ cm } H_2O$ using artificial neural networks. The predictors used were particle size classes, OM, bulk density ($\rho_b$), information from the pedological description of the different soil horizons, large soil pores parameter (amount of pores drained at $-10 \text{ cm } H_2O$), and soil structure grade. Furthermore, Nielsen et al. (2018) developed a PTF for predicting the $K_s$ for sandy agricultural fields in Denmark, and a linear function of the logarithm of $K_s$ and effective porosity was derived.

In the last few decades, many studies have attempted to predict soil properties using visible near-infrared (vis-NIR) spectroscopy (400–2500 nm), which is a fast, nondestructive, and relatively low-cost method using small sample amounts (Stenberg et al., 2010). The required preparation for this method is simple, involving only air-drying of the soil and sieving down to 2 mm. Visible near-infrared spectroscopy has been successfully used to predict soil properties such as the complete soil texture distribution (Hermansen et al., 2017), soil structure (Katuwal et al., 2018), and soil specific surface area (Ben-Dor et al., 2008; Knadel et al., 2018). Furthermore, studies have shown its ability to predict the wet part of the soil-water retention curve (Babaeani et al., 2015; Pittaki-Chrysodonta et al., 2018; Santra et al., 2009). Despite the attempts that have been made in recent years to develop models using vis-NIR spectroscopy for fast predictions of soil properties or soil functions, the application of these models is still not widespread.

The objective of this study was to develop a hydrogeological model for simulating the actual and cumulative tile-drainage dynamics in an agricultural landscape, using as input hydraulic parameters (i) the output from integrating an existing vis-NIR model and (ii) a developed PTF model for estimating the water release characteristics and $K_s$. Thus, a fast and inexpensive hydrogeological tool is proposed to cover the absence of hydraulic data and deal with the uncertainty derived from the spatial variability of the hydraulic properties within a field. For the evaluation of this suggested approach, a hydrogeological model was developed in HYDRUS-2D using a single-porosity model (Šimůnek et al., 2006), and a comparison between the simulated drainage dynamics and the observed data was performed. Finally, a sensitivity analysis was implemented to determine the sensitivity of predicted parameters and how their variation affected model performance.

Materials and Methods

Study Area

A tile-drained agricultural field located in the Norsminde catchment in eastern Jutland, Denmark, was selected for simulating the soil water regime and water balance ($55^\circ59' \text{ N, } 10^\circ4' \text{ E}$) (Fig. 1). The surface elevation varies from 95 to 72 m above mean sea level, and clayey till unit is located down to 3 m depth. In the northeastern part of the study field, the area is covered by forest. Clayey and sandy glacial deposits are found at larger depths (De Schepper et al., 2017; He et al., 2014). An electromagnetic (EM) survey was conducted in the study field using a sensor (DUALEM-21S, DUALEM Inc.). Little variation in electrical conductivity (EC) was observed across the field for the 0- to 25-cm depth, indicating a homogeneous pedological unit. The EC measurements varied for larger depths (25–130 cm), with the highest values (32.9 mS m$^{-1}$) identified close to the two depression zones (Fig. 1b and 1c). The study area was classified as sandy loam (0–25-cm depth, Ap horizon) according to Varvaris et al. (2018). Below plow depth (25–130 cm), stratified heterogeneity was revealed based on borehole descriptions and soil texture analysis, and the soil texture ranged from sandy loam to clayey loam, reflecting the pedological clay illuviation processes enhanced by the intense hydro-topographical gradients in the area (Varvaris et al., 2018). According to Varvaris et al. (2018), the soil profile below 25-cm depth was divided into three distinct horizons: E, Btg1, and Btg2; the average clay, silt and sand fraction varied from 0.19 to 0.21 kg kg$^{-1}$, from 0.28 to 0.32 kg kg$^{-1}$, and from 0.46 to 0.50 kg kg$^{-1}$, respectively. The field had 54 partially penetrating piezometers installed at various depths (20–180 cm) for monitoring the piezometric head of the more conductive layers. During the installation of the piezometers, 76 bulk soil samples from different depths and 30 undisturbed 100-cm$^3$ (3.5-cm height, 6.1-cm diameter) soil cores from 0 to 20 cm were collected (Fig. 1). Five bulk soil samples were collected from the locations where the undisturbed soil cores were sampled. The bulk soil samples were initially air dried and sieved at 2 mm, and the undisturbed soil cores were used for soil water retention analysis. The soil cores were placed in a sandbox and then slowly saturated with tap water from the bottom. The columns were then drained to matric potentials of $-30$ and $-100 \text{ cm } H_2O$ and moved to a ceramic pressure plate apparatus, where the samples were further drained to different matric potentials down to $-1000 \text{ cm } H_2O$. Dry bulk density was determined for the
same soil cores by oven-drying at 105 °C for 24 h. The study area was subdivided to take account of the spatial variability of the soil texture distribution into three subareas based on the EM survey (Fig. 1d) using iso-cluster unsupervised classification in ArcGIS 10.4 (ESRI). The EM measurements can be used as an indicator of the variation in surface elevation because the topography affects the soil characteristics and the formation of the pedological units (e.g., soil texture and thickness of horizons); consequently, these changes are reflected in electrical conductivity (EC) measurements (Serrano et al., 2014). The optimum number of clusters (subareas) was determined using as criterion the existence of a minimum number of bulk and undisturbed soil samples that could be used to characterize each subarea. Because of infeasibility of conducting EM measurements in the forest area, the forest had been considered part of Area 3. That assumption was based on the high organic matter contents in that area, which would be comparable to those in the defined depression zones (Ambus and Christensen, 1995).

Predicting Soil-Water Retention Curve Using Visible Near-Infrared Spectroscopy

Visible near-infrared spectroscopy measurements were performed on ~50 g of a representative air-dried bulk soil sample using a vis-NIR sensor (NIRS DS2500 spectrometer, FOSS) at 0.5-nm spectral resolution. The soil was placed in a 60-nm sample cup, and the reflectance was measured at seven positions in a controlled near-infrared laboratory (temperature, 23 °C; humidity, 48%). The averaged spectrum was extracted for each soil, and the reflectance was converted into absorbance by log(反射ance⁻¹).

Multivariate Data Analysis

To obtain a soil-water retention curve (SWRC) based solely on the vis-NIR measurements, a preexisting model predicting the SWRC was used (Pittaki-Chrysodonta et al., 2018). Pittaki-Chrysodonta et al. (2018) anchored the Campbell soil water retention function at a matric potential of −1000 cm H₂O (pF 3 = log|−1000 cm H₂O|) instead of saturated water content as a reference point. The modification of the traditional Campbell function was made because it has reference point at saturated water content and would typically be weakly related to soil texture properties. For that reason, anchoring the function at pF 3, the soil-water content at this dryer condition is related to the basic soil characteristics, and it was successfully predicted by vis-NIR spectroscopy (Pittaki-Chrysodonta et al., 2018). The derived function was expressed as:

---

Legend (a)
- 100-cm³ core soil samples
- Bulk soil samples
- Drainage Wells
- Drainage Discharge Station
- Tile Drainage Network
- Drainage subcatchment boundary

Legend (b-d)
- High : 32.9
- Low : 4.6
- EC (mS/m)
- Area 1
- Area 2
- Area 3
- Forest

Fig. 1. (a) Drainage subcatchment with the mapped tile drainage network, the drainage discharge station, and the collected soil samples, (b,c) electrical conductivity (EC) maps at two depths (0–25 and 25–130 cm), and (d) drainage subcatchment subdivided into the three areas (Areas 1–3).
\[
\psi = -1000 \left( \frac{\theta}{\theta_{pF3}} \right)^{-b}
\]

where \( \theta \) is the volumetric water content (cm\(^3\) cm\(^{-3}\)), and \( \theta_{pF3} \) is the volumetric water content at pF 3. Campbell \( b \) is regarded as a pore size distribution index (Moldrup et al., 2001) and is equal to the slope of the SWRC on a log-log system.

The vis-NIR modeled by Pittaki-Chrysodonta et al. (2018) predicted the Campbell \( b \) and the \( \theta_{pF3} \). Spectra from Danish agricultural soil samples with a wide textural range from sand to loam were included in their study. Campbell \( b \) and \( \theta_{pF3} \) were predicted, with \( R^2 \) values of 0.86 and 0.92, respectively, and RMSE values of 1.52 (dimensionless) and 0.002 (cm\(^3\) cm\(^{-3}\)).

**Pedotransfer Function**

A PTF was developed to predict \( K_s \) derived from pores with an equivalent pore diameter larger than 3 \( \mu \text{m} \). Specifically, for the \( K_s \) prediction, vis-NIR–predicted Campbell \( b \) and effective porosity \( \phi_e \) were used as predictors. The value for \( \phi_e \) is given by

\[
\phi_e = \phi_{\text{sat}} - \phi_{pF3}
\]

where \( \phi_{\text{sat}} \) and \( \phi_{pF3} \) (L\(^3\) L\(^{-3}\)) are the water contents corresponding to saturation and −1000 cm H\(_2\)O matric potential. It is assumed that the difference represents the active porosity, which is the pore space that actively transports water (Rezanehzad et al., 2009). According to Lin et al. (1996), pores with equivalent diameters >60 \( \mu \text{m} \) contribute ~99% of the total water transport.

In this study, 194 soil measurements of \( K_s \) from undisturbed 100-cm\(^3\) soil core samples from three agricultural fields (Karup et al., 2016) were included in the development of the PTF for predicting \( K_s \): Estrup (44 soil samples), Silstrup (65 samples), and Jyndevad (87 samples). Detailed soil and site information for Estrup, Silstrup, and Jyndevad can be found in Lindhardt et al. (2001), Masis-Melendez et al. (2014), Norgaard et al. (2013), Paradelo et al. (2015). Thus, the suggested PTF for predicting the \( K_s \) (\( \mu \text{m} \text{s}^{-1} \)) is

\[
\log K_s = (A\phi_e - C)b
\]

where \( b \) is the negative slope of the SWRC on a log-log scale [log(\( \psi \)] vs. log(\( \theta \) system], and \( A \) and \( C \) are regression coefficients. However, Messing and Jarvis (1995) found that an underestimation of the hydraulic conductivity in the suction range −15 cm H\(_2\)O to saturation may occur when measuring on small core samples (10-cm height, 7.2-cm diameter) compared with larger soil samples (~45-cm height, 30-cm diameter). This can be explained by the fact that the aperture and spacing of macropores may be representative over a representative elementary volume (REV), for which that the soil structure would be successfully represented (Bachmat and Bear, 1986). Iversen et al. (2001) also found a relatively large scaling exponent between small (100 cm\(^3\)) and large (6280 cm\(^3\)) soil samples. They studied the scaling behavior and estimated the scaling exponent for each of the studied fields (Tylstrup, Silstrup, Estrup, Jyndevad). To upscale the measured \( K_s \) to a value related to a desired REV for the studied soils, the equation from Schulze-Makuch et al. (1999) was used describing the scaling behavior:

\[
K_s = c (V')^m
\]

where \( c \) is the y intercept of the regression line, \( V \) is the volume of the tested material, and \( m \) is the scaling exponent. In this study, the agricultural field presented similar soil texture distributions and crop management to Estrup, and the values for \( m \) and \( V' \) were therefore set according to Iversen et al. (2001).

**Flow Model**

**Hydraulic Model Description**

The HYDRUS-2D finite element code (Šimůnek et al., 1999, 2003) was used for simulating the drainage dynamics in the studied agricultural field. A multiregion model (e.g., dual-porosity, dual-permeability model) would be able to simulate the tile drainage dynamic patterns more accurately. However, the large number of parameters to be defined would result in a more complex model, and their estimation would be problematic because of parameter non-uniqueness and stability (Hopmans and Šimůnek, 1997). Moreover, Varvaris et al. (2018) examined the capabilities and weaknesses of three different model approaches for modeling the water flow dynamics in the same field as for this study and that the single-porosity approach gave an acceptable accuracy with significantly fewer parameters to define compared with dual-permeability and dual-porosity models, which are more complex approaches. Thus, a simpler, conceptual approach will be favorable to a more sophisticated physical model; therefore, an equilibrium-flow, single-porosity model was used in which the air entry value was set to a matric potential of −2 cm H\(_2\)O, corresponding to a situation where pores with equivalent diameters >1500 \( \mu \text{m} \) are drained. The variably saturated flow is described using Richards’ equation:

\[
\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial x_i} \left( K \left( \frac{\partial h}{\partial x_i} + K^{\alpha} \right) \right) - S
\]

where \( \theta \) (L\(^3\) L\(^{-3}\)) is the volumetric water content, \( t \) [T] is the time, \( b \) [L] is the suction, \( S \) [T\(^{-1}\)] is a sink term, \( x_i \) (\( i = 1, 2 \) [L]) are the spatial coordinates, and \( K^{\alpha} \) are components of a dimensionless anisotropy tensor \( K^\alpha \). The \( K \) [L \ T\(^{-1}\)] denotes the hydraulic conductivity function. The van Genuchten–Mualem model (Mualem, 1976; van Genuchten, 1980) was used to characterize the soil hydraulic properties:

\[
\theta(b) = \begin{cases} 
\frac{\theta_{s} - \theta_{r}}{1 + |b|^{1/m}} b < 0 \\
\theta_{s} b \geq 0
\end{cases}
\]

\[
K(b) = \begin{cases} 
K_s \sqrt[1/m]{1 - \left(1 - S_{e}^{1/m} \right)^2} b < 0 \\
K_s b \geq 0
\end{cases}
\]
where $\theta_r$ and $\theta_s$ [L$^3$ L$^{-3}$] denote the residual and saturated water contents, respectively; $\alpha$ [L$^{-1}$] is related to the inverse of the air entry suction; $n$ (dimensionless) is a measure of the pore-size distribution index; $m = 1 - 1/n$ (dimensionless); $K_s$ [L$^0$ T$^{-1}$], $l$ (dimensionless) is the pore-connectivity parameter; and $S_e$ (dimensionless) is the effective saturation given by

$$S_e = \frac{\theta - \theta_r}{\theta_s - \theta_r}$$

### Drainage Data and Input Data

The tile drainage discharge measurements from 9 May 2013 to 18 June 2014 (9744 hourly values) were used as observed data for the comparison with the simulated drainage discharge. Hourly precipitation data were extracted from two nearby meteorological stations (3 and 6 km from the studied field). The climate in the study field as well as for the areas where the meteorological stations are located is characterized as coastal temperate (mean daily temperature, 10°C; annual precipitation, 671 mm). The daily potential evaporation (PE) and potential transpiration (PT) were estimated using the water balance model EVACROP (Olesen and Heidmann, 1990), in which data from the meteorological station in Foulum, Denmark (56°29′ N, 9°35′ E) (distance 63 km) were used as input variables. To match the hourly discharge values, the daily data were transformed into hourly data distributed between the hours of 5:00 AM and 8:00 PM using a Gaussian distribution (Bakhtiari et al., 2006). A more detailed description regarding the input variables (PE and PT) is provided in Varvaris et al. (2018).

### Conceptual Model and Boundary Conditions

The simulated domain had a width of 20 m and a depth of 1.3 m, corresponding to the domain used by Varvaris et al. (2018). The domain was subdivided into two distinct layers because the average clay, silt, and sand fraction for the three horizons below topsoil were similar. The soil texture analysis for the two defined layers is presented in Table 1. The thickness for the first layer was 0.25 m, and one single layer was selected to represent the entire subsurface. An impermeable layer was set below the 1.3-m depth. The tile drain was placed in the middle of the domain at a depth of 1.2 m (Fig. 2).

Atmospheric boundary conditions at the soil surface were used for the top of the domain, Neumann boundary conditions were used for the bottom, and lateral no-flux boundaries were used at the sides of the domain. The tile drain was implemented using seepage face boundary conditions and functioned only when the suction along the drain was equal to zero and the surrounding soil was saturated. When the effective saturation of the surrounding soil was $<1$, the tile drain discharge was equal to zero, and the tile drain was treated as a nodal sink (Šimůnek et al., 2006). More information about the conceptual model and the boundary conditions can be found in Varvaris et al. (2018).

### Root Water Uptake

Root water uptake was simulated using the plant water stress function of Feddes et al. (1978) with model uptake parameters for winter wheat ($Triticum aestivum$ L.) for the entire simulation. The maximum rooting depth was set to 1.2 m according to (Palosuo et al., 2011), with a depth of maximum intensity of 0.7 m (Gärdenäs et al., 2006). For optimum root water uptake, all input values originated from Varvaris et al. (2018): the threshold value was set to $h_{opt} = -0.1$ m, and the critical pressure heads for the drought stress ($h_3$) and the wilting point ($h_4$) were set to $-15$ and $-160$ m, respectively.

### Effective Parameterization of Properties by Integrating Visible–Near-Infrared and a Pedotransfer Function

Figure 3 illustrates the concept for predicting the actual and cumulative tile drainage discharge in the study area using vis-NIR spectroscopy and the developed PTF. Having predicted the parameters of the anchored Campbell SWRC ($b$ and $\theta_{pF3}$) for the 81 bulk soil samples using the vis-NIR models by Pittaki-Chrysodonta et al. (2018), the average Campbell $b$ and $\theta_{pF3}$ for each subarea were estimated. Specifically, 17 (four for the first layer, 13 for the second layer), four (two for both layers), and 56 (12 for

---

### Table 1. Soil texture analysis according to the USDA.

<table>
<thead>
<tr>
<th>Depth (cm)</th>
<th>Clay (kg kg$^{-1}$)</th>
<th>Silt (kg kg$^{-1}$)</th>
<th>Sand (kg kg$^{-1}$)</th>
<th>Organic C (kg kg$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–25</td>
<td>0.13 (0.11–0.15)†</td>
<td>0.24 (0.22–0.26)</td>
<td>0.63 (0.59–0.67)</td>
<td>0.011 (0.008–0.015)</td>
</tr>
<tr>
<td>25–130</td>
<td>0.22 (0.09–0.41)</td>
<td>0.32 (0.24–0.45)</td>
<td>0.46 (0.20–0.68)</td>
<td>0.059 (0.001–0.221)</td>
</tr>
</tbody>
</table>

† Means, with range in parentheses.
the first layer, 44 for the second layer) soil samples were used for defined Subareas 1, 2, and 3, respectively (Fig. 1d). Afterward, the weighted-average Campbell SWRC for the entire drainage subcatchment was estimated considering the weighted factors ($w_{1,2,3}$), and the weighted-average van Genuchten (vG) parameters ($\alpha$ and $n$) for each subarea were then estimated using the software RETC (van Genuchten et al., 1991). The weighted factors were determined based on the area of each subarea. However, weighted factors that represent the contribution of each subarea to the total drainage output would lead to a more accurate average Campbell SWRC. Furthermore, $K_s$ was predicted by using the developed PTF and assuming that the saturated water content ($pF_0$) is equal to the defined air-entry pressure at $-2$ cm H$_2$O in the described water flow model ($\theta_{sat} \approx \theta_{pF0.3}$).

### Sensitivity Analysis

A sensitivity analysis was performed to examine the effect on drainage output by varying the parameters Campbell $b$, $\theta_{pF3}$, and $K_s$ on the simulation results and in this way determine the “sensitive” and “insensitive” parameters. The selected range in the percentage change in parameter Campbell $b$ and $\theta_{pF3}$ was set based on Moldrup et al. (2001) and Pittaki-Chrysodonta et al. (2018) to preserve the physical meaning of these parameters in relation to the soil texture analysis from the studied field. The sensitivity analysis for $K_s$ was selected to be within one magnitude.

### Statistical Analysis

The evaluation of the water flow model was performed using the $R^2$ value, the Nash–Sutcliffe model efficiency coefficient (NSE), and the RMSE. The RMSE is a measure of the differences between predicted values and observed data and is defined as

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\bar{x}_i - x_i)^2}$$  \hspace{1cm} [9]$$

where $N$ is the number of samples, $\bar{x}_i$ is the predicted values, and $x_i$ are the observed data.

The NSE equation (Nash and Sutcliffe, 1970) was also used to evaluate the agreement between measured and simulated flow hydrographs and is defined as

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N} (x_i - \bar{x}_i)^2}{\sum_{i=1}^{N} (x_i - x_{mean})^2}$$ \hspace{1cm} [10]$$

where $x_{mean}$ is the mean value of the observed data. The NSE coefficient can range from $-\infty$ to 1.0, and it is equal to 1.0 in the case of a perfect agreement. The accuracy of prediction is considered acceptable for values between 0.0 and 1.0.

### Results and Discussion

#### Predicted Soil-Water Retention Curve and Macroporosity

The SWRCs of the three defined subareas were predicted fairly well by the vis-NIR model (Fig. 4). Although the NIR-predicted water retention curves were systematically underpredicted compared with the measured data, the slopes (changes of water content with magnitude of matric potential, pF) were well predicted, making the prediction of effective porosity fairly good. For Area 1, predicted Campbell $b$ and $\theta_{pF3}$ were underestimated by 14.2 and 19.0%, respectively, compared with measured data (Fig. 4a). For Areas 2 and 3, predicted water contents at a matric potential of $-1000$ cm H$_2$O presented similar differences from the obtained data for
the soil samples collected from the studied field. The Campbell parameter compared well with measured values (Fig. 4b and 4c), particularly for Areas 2 and 3 where estimated values were 11.87 and 13.86 compared with measured values of 12.52 and 15.00 and of 0.27 and 0.31, respectively (Table 2). These discrepancies may be because no spectral measurements from the collected soil samples in the study field were included in the spectral library of the vis-NIR SWRC model, which would have extended its ability to more accurately predict the two investigated parameters. The limited number of soil-water retention data led to the decision to use them only for the validation of the predicted SWRCs rather than to further calibrate the vis-NIR SWRC model. Additionally, the offset of the predicted values from the measured values might be due to the slightly higher clay content in the Fensholt subcatchment than in the soil samples used in the study by Pittaki-Chrysodonta et al. (2018).

The effective porosity for all of the subareas was estimated at 0.15 cm $^3$ cm$^{-3}$, and the water content at pF 0.3 (log$[-2 \text{ cm H}_2\text{O}]$) was 0.40, 0.37, and 0.41 cm$^3$ cm$^{-3}$ for Areas 1, 2, and 3, respectively. The lower $\theta_s$ and $\theta_{pF3}$ values for Area 2 reflected the higher content of fine particles because the clay illuviation and erosion processes were more intensive in this part of the field due to the steep topographical gradients. In contrast, Area 3 had the highest $\theta_s$ and $\theta_{pF3}$, reflecting the two identified depression zones and, consequently, the higher contents in clay and OM. Figure 4d depicts the Campbell area-weighted average SWRC based on predicted data for the three subareas using the weighted factors in Table 2.

Figure 5 presents the area-weighted average SWRC transformed from Campbell function anchored at pF 3 to vG function. The near-saturation water contents ($\theta_{pF1}$) were well predicted; the greatest difference was at $\theta_{pF3}$. The predicted values at $\theta_{pF1}$ show an absolute difference of 7.5% from the measured data, whereas for $\theta_{pF3}$ this was 31%. The fitting of the vG function to the measured data for all of the areas was well predicted.

Table 2. Campbell soil-water retention function parameters (Campbell $b$, $\theta_{pF3}$) and effective porosity ($\phi_e$) for the three subareas and the area-weighted average.

<table>
<thead>
<tr>
<th>Depth</th>
<th>Parameter</th>
<th>Area 1 (wft: 0.35)</th>
<th>Area 2 (wft: 0.30)</th>
<th>Area 3 (wft: 0.35)</th>
<th>Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>cm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–25</td>
<td>Campbell $b$</td>
<td>16.01</td>
<td>12.52</td>
<td>15.00</td>
<td>14.45</td>
</tr>
<tr>
<td></td>
<td>predicted</td>
<td>13.73</td>
<td>11.87</td>
<td>13.86</td>
<td>13.08</td>
</tr>
<tr>
<td></td>
<td>$\theta_{pF3}$, cm$^3$ cm$^{-3}$ measured</td>
<td>0.31</td>
<td>0.27</td>
<td>0.31</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>predicted</td>
<td>0.25</td>
<td>0.22</td>
<td>0.26</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>$\phi_e$, cm$^3$ cm$^{-3}$ measured</td>
<td>0.11</td>
<td>0.17</td>
<td>0.13</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>predicted</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>25–130</td>
<td>Campbell $b$</td>
<td>21.68</td>
<td>23.79</td>
<td>18.47</td>
<td>20.92</td>
</tr>
<tr>
<td></td>
<td>predicted</td>
<td>21.68</td>
<td>23.79</td>
<td>18.47</td>
<td>20.92</td>
</tr>
<tr>
<td></td>
<td>$\theta_{pF3}$, cm$^3$ cm$^{-3}$ measured</td>
<td>0.37</td>
<td>0.37</td>
<td>0.34</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>predicted</td>
<td>0.12</td>
<td>0.11</td>
<td>0.14</td>
<td>0.12</td>
</tr>
</tbody>
</table>

† Area-weighted factor.
data also showed a large discrepancy at $\theta_{pF3}$ (17%). The vG shape parameters ($a$ and $n$) from the measured data were 0.018 cm$^{-1}$ and 1.20, respectively, whereas for the predicted model they were 0.054 cm$^{-1}$ and 1.17. Despite the observed difference, the behaviors of the predicted and observed vG SWRCs were almost identical, leading to a relatively good prediction of effective porosity and to a relatively good description of the SWRC that is used as input in HYDRUS-2D.

**Pedotransfer Function for Predicting Saturated Hydraulic Conductivity**

The developed PTF for predicting $K_s$ is shown in Fig. 6 together with the soil texture triangle (Fig. 6a) showing the origin of the data used. An $R^2$ value of 0.93 was obtained, indicating that the $K_s$ was strongly correlated with effective porosity and Campbell $b$. Specifically, the logarithm of $K_s$ divided by Campbell $b$ presented a linear relation to $f_e$ (Fig. 6b).

Dividing the parameter of log $K_s$ by Campbell $b$ takes into account the pore size distribution, which is related to the soil texture class and OM. These parameters are indirectly related to the structure of the soil, which is correlated with $K_s$ (Dexter et al., 2004; Jarvis et al., 2002). By inserting the parameter Campbell $b$, the log $K_s$ measured data were evenly distributed over $\phi_e$ and led the PTF to a linear fit (Fig. 6b). The linear regression analysis showed standard error of estimate equal to 0.051 and $p < 0.001$. The estimated $K_s$ was subsequently upscaled using Eq. [3], where the parameter $m$ was set to 0.62 as determined by Iversen et al. (2001) for the corresponding Estrup soil. The parameter $V$ in Eq. [3] was set to 5670 cm$^3$ (equivalent to soil cores with an inner diameter of 19 cm and a height of 20 cm). This was the same volume as the cylindrical soil columns used at the Estrup field site for studying macroporosity (defined as the pores with equivalent diameters $>1.2$ mm) using X-ray computed tomography (Katuwal et al., 2018). The optimum REV for medium to coarse soils should be larger than 2500 cm$^3$ according to Bouma (1985) and Mallants et al. (1997). The predicted $K_s$ values for the first and second layer were 23.6 and 5.4 cm h$^{-1}$, respectively. The estimated $K_s$ values were in agreement with the range of the measured $K_s$ values from Estrup (Lindhardt et al., 2001; Paradelo et al., 2016).

**Simulation of Actual and Cumulative Drainage Discharge**

Regarding the input variables in HYDRUS-2D, the Campbell function anchored at $pF_3$ was used for predicting the parameters of the developed PTF ($\phi_e$, Campbell $b$) for the average-weighted area for the two distinct layers. Having predicted the vG shape parameters for each subarea, the predicted Campbell $b$ and $\theta_{pF3}$ were inserted into the Campbell function to estimate the volumetric water content at $pF_0.3$ and, subsequently, $\phi_e$. Using Eq. [8], $K_s$ was determined (Table 3). The approach used was able to simulate fairly well the main attributes of the flow hydrograph; specifically, the general shape, timing, and peak discharge rates were accurately predicted (Fig. 7).

Initially, in the period from 15 to 30 May 2013 where the observed drain discharge peaked at 34.5 L s$^{-1}$, the model simulated accurately the observed drainage data and the maximum
differences varied from 1.4 to 1.9 L s⁻¹ (Fig. 7a). These differences between observed and modeled discharge may be explained by an underestimation of local precipitation by the nearby meteorological stations or a misfitting of initial conditions. Moreover, the horizontal shift of 0.04 cm³ cm⁻³ between the measured and predicted SWRC may be a reason for the presented difference because the higher downward force of gravity compared with the capillary forces will provoke a faster water transport toward the tile drains. Furthermore, the lower predicted \( q_s \) reduced the maximum available pore spaces to be filled with water and allowed us to quickly determine the infiltration capacity and, consequently, the \( K_s \). A dry period with low continuous discharge rates (~0.2 L s⁻¹) followed from 31 May 2013 to 27 Oct. 2013, and the model failed to simulate consistently this steady base outflow, but only with some small fluctuations compared with the observed values. An explanation for this discrepancy could be an unknown source that recharged the tile drainage system at a constant rate. The two identified depression zones, where the groundwater level was relatively high, may have contributed to this continuous, low-drainage discharge. The piezometric head of the more conductive layers in the two depression zones had higher values than the rest of the field during the monitoring of the installed piezometers in the relatively dry period of 2016 to 2017. Moreover, the higher \( \alpha \) predicted value in conjunction with the high evapotranspiration rates during that time period may have forced the soil to be drained faster, preventing the existence of base outflow. The period from 28 Oct. 2013 to 8 Apr. 2014 was characterized by intensive discharge rates, with drainage outputs varying from 5.2 to 41.6 L s⁻¹. From 18 Oct. 2013 to 1 Dec. 2013, the observed drainage rates were underpredicted. An explanation for this might be an overestimation of the local rainfall or an underestimation of the potential transpiration because of a lack of knowledge of the crop management in the study area. The model simulated the peaks from 2 Dec. 2013 to 8 Apr. 2013 with a relatively high degree of accuracy and was able to precisely match the recession limbs. This can be crucial in modeling of pesticide and nitrate transport because the mobility of pesticides and nitrate is significantly higher during high and continuous rainfall events than in a dry season characterized by a discontinuous regime of infiltration (Neal, 1991). The performance statistics of the model for the entire simulation were characterized fairly well (\( R^2 \), 0.86; NSE, 0.84; RMSE, 1.25 L s⁻¹) (Table 4) and compared well with results obtained by the dual-permeability model for the same time and field in Varvaris et al. (2018). In that study, three different models accounting for macropore flow were used to examine the ability of each approach to predict the drainage dynamics in the same drained agricultural field. The initial effective parameters for HYDRUS-2D were based on in situ observations of soil structure, soil hydraulic property maps of Denmark, and pedotransfer functions. According to the authors, the dual-permeability model gave the best agreement with the observed drainage data in the drainage subcatchment (\( R^2 \), 0.84; RMSE, 1.27 L s⁻¹). However, 17 parameters were calibrated for the dual-permeability model, whereas in this study the two-layer, single-porosity model successfully simulated the drainage dynamics without any calibration and relying only on the knowledge obtained by vis-NIR and the developed PTF.

### Table 3. Soil hydraulic input parameters for water flow models in HYDRUS-2D.

<table>
<thead>
<tr>
<th>Depth</th>
<th>Model</th>
<th>( \theta_r )</th>
<th>( \theta_s )</th>
<th>( \alpha )</th>
<th>( K_s )</th>
<th>( l )</th>
</tr>
</thead>
<tbody>
<tr>
<td>cm</td>
<td></td>
<td>cm⁻³ cm⁻³</td>
<td>cm⁻³ cm⁻³</td>
<td>cm h⁻¹</td>
<td>cm h⁻¹</td>
<td></td>
</tr>
<tr>
<td>0–25</td>
<td>predicted</td>
<td>0.39</td>
<td>0.054</td>
<td>1.17</td>
<td>23.6</td>
<td>-1</td>
</tr>
<tr>
<td>25–130</td>
<td>predicted</td>
<td>0.48</td>
<td>0.013</td>
<td>1.19</td>
<td>5.4</td>
<td>-1</td>
</tr>
<tr>
<td>0–25</td>
<td>calibrated</td>
<td>0.42</td>
<td>0.045</td>
<td>1.17</td>
<td>28.6</td>
<td>-1</td>
</tr>
<tr>
<td>25–130</td>
<td>calibrated</td>
<td>0.49</td>
<td>0.014</td>
<td>1.23</td>
<td>6.4</td>
<td>-1</td>
</tr>
</tbody>
</table>

Fig. 7. (a) Actual drainage discharge simulated using the visible near-infrared model and the pedotransfer function for estimating the input parameters for HYDRUS-2D, (b) calibrated model, and (c) cumulative drainage discharge for both models.

### Table 4. The \( R^2 \), root mean square error (RMSE), Nash–Sutcliffe efficiency (NSE), and cumulative bias statistics of the visible–near-infrared (vis-NIR) and calibrated models.

<table>
<thead>
<tr>
<th>Simulation</th>
<th>( R^2 )</th>
<th>RMSE</th>
<th>NSE</th>
<th>Cumulative bias</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L s⁻¹</td>
<td>m³ ha⁻¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>vis-NIR</td>
<td>0.86</td>
<td>1.25</td>
<td>0.84</td>
<td>4.16 × 10²</td>
</tr>
<tr>
<td>Calibrated</td>
<td>0.86</td>
<td>1.21</td>
<td>0.85</td>
<td>3.45 × 10²</td>
</tr>
</tbody>
</table>
The cumulative drainage discharge was overestimated by the model, with a cumulative bias of $4.16 \times 10^2$ m$^3$ ha$^{-1}$ (Fig. 7c). The difference may be related to the assumptions of a Neumann boundary condition (impermeable layer) as the bottom boundary and of the agricultural field as a systematically tile-drained field, whereas tile drains were in fact installed in only part of the field. Consequently, some of the percolated water could have bypassed the tile drain system, recharging a deeper aquifer (De Schepper et al., 2017; He et al., 2014).

For a thorough evaluation of the performance of the model using as input parameters the output from integrating the vis-NIR models and the developed PTF, a further calibration was performed to determine the optimum predictive ability of the two-layer, single-porosity model (Fig. 7b, c). The effective parameters ($\theta_s$, $\alpha$, $n$, and $K_s$) from the vis-NIR SWRC models and the PTF were included as initial parameters, and further calibration was made. Using repeated stepwise manual calibration by setting the observed drainage discharge as the objective function, the calibrated model achieved an $R^2$ value of 0.86, NSE of 0.85, and RMSE of 1.21 L s$^{-1}$ (Table 4). Small deviations were seen between the calibrated and the vis-NIR model, specifically for $K_s$, for which the values for the first and the second layers in the calibrated model were 28.6 and 6.4 cm h$^{-1}$, respectively, whereas for the vis-NIR model they were 23.6 and 5.4 cm h$^{-1}$ (Table 3). The vG shape parameter $n$ for the first layer was the same in the calibrated and vis-NIR models, but the optimum $\alpha$ value in the calibrated model was 0.045 cm$^{-1}$, which was lower than in the vis-NIR model. For the second layer, the estimated $\alpha$ and $n$ parameters in the calibrated model were 0.014 cm$^{-1}$ and 1.23, respectively. For the volumetric water content at zero suction level, the optimum values were 0.42 and 0.49 cm$^3$ cm$^{-3}$ for the first and second layer, respectively, whereas for the vis-NIR model these values were 0.39 and 0.48 cm$^3$ cm$^{-3}$ (Table 3).

**Sensitivity Analysis**

Figure 8 illustrates the results from the sensitivity analysis and the resulting $R^2$ and RMSE by varying Campbell $b$ and $\theta_{pF3}$ from −20 to 20% and varying $K_s$ for both layers from −75 to 75%. Regarding $K_s$, the sensitivity analysis showed that variations within an order of magnitude would not significantly affect the results for the drainage discharge rates. Specifically, the $R^2$ varied from 0.85 to 0.73 and the RMSE from 1.63 to 1.45 L s$^{-1}$ for −75 and 75% changes in parameter $K_s$, respectively. Similarly, for Campbell $b$ and $\theta_{pF3}$, small variations (0–10%) did not affect the results significantly. For this range (−10 to 10%) in the parameters Campbell $b$ and $\theta_{pF3}$, the $R^2$ varied from 0.81 to 0.76 and from 0.76 to 0.81, respectively, and the RMSE varied from 1.67 to 1.56 and from 1.55 to 1.74 L s$^{-1}$, respectively. However, for larger uncertainties from 10 to 20%, the $R^2$ and RMSE values for Campbell $b$ decreased by 25 and 15%, respectively, whereas for $\theta_{pF3}$ the decreases were 20 and 28%, respectively. Despite these variations and the resulting influences on the predictive ability of the model, the average absolute variations for the prediction models of Campbell $b$ and $\theta_{pF3}$ parameters according to Pittaki-Chrysodonta et al. (2018) were 13 and 9%, respectively.

![Fig. 8. Parameter sensitivity analyses showing the effect of changing the three main soil hydraulic input parameters on the HYDRUS-2D simulated drainage discharge: (a,b) visible–near-infrared–predicted Campbell $b$ and volumetric water content at a matric potential of −1000 cm H$_2$O ($\theta_{pF3}$), and (c,d) saturated hydraulic conductivity ($K_s$) estimated from the pedotransfer function.](image-url)
Conclusions

This study presents a concept for developing a hydrogeological tool that can predict the effective input parameters to a model that simulates actual and cumulative drainage discharge in an agricultural field. Integrating an existing vis-NIR SWRC model with a developed PTF, the soil–water release characteristics from the Campbell soil–water retention function and the $K_s$ were estimated, and the output was used as input in HYDRUS-2D to simulate the drainage dynamics in a clayey tile-drained subcatchment in Denmark. The model was successfully evaluated by comparing the simulated drainage dynamics with the observed data. The tile drainage output from the vis-NIR model was also closely compared with that from the optimally calibrated model. The proposed approach can be used as an indirect method for covering the absence of measured data of soil hydraulic properties, taking into account the spatial variability of soil properties within a field using only vis-NIR measurements.

The predicted soil–water retention curves were offset from the measured data. However, the slope of the predicted curves was in fairly close agreement with the slope from the measured data, leading to a satisfactory estimation of effective porosity and, consequently, of saturated hydraulic conductivity. Including data from a wider range of textural classes will improve vis-NIR SWRC models. Therefore, the suggested approach can be applied also to hydrogeological models that aim to predict the actual and cumulative drainage discharge in spatially heterogeneous agricultural fields and with larger variations in soil texture distribution. The suggested approach needs to be further evaluated by simulating solute transport, and the model’s ability to represent the actual internal flow and transport pathways within a tile-drained agricultural subcatchment must be appraised.

Acknowledgments

This work was supported in part by the Innovation Fund Denmark, project Future Cropping (Grant no. AUFF-E-2016-9-36).

References


near infrared spectroscopy for prediction of soil structural properties. Vadose Zone J. 17:160054. doi:10.2136/vzj2016.06.0054


