



## The Dutch soil physical units map: BOFEK

M. Heinen<sup>\*</sup>, H.M. Mulder, G. Bakker, J.H.M. Wösten<sup>1</sup>, F. Brouwer, K. Teuling, D.J.J. Walvoort

Wageningen Environmental Research, P.O. Box 47, 6700 AA Wageningen, The Netherlands

### ARTICLE INFO

Handling Editor: Budiman Minasny

#### Keywords:

Soil map clustering  
Hydraulic conductivity characteristic  
Simulation  
Soil hydraulic properties  
Transpiration reduction  
Water retention characteristic

### ABSTRACT

Soils and their properties play an important role in land evaluation studies. Often such studies focus on larger scales ranging from watersheds up to the national scale or even larger. Soil properties are often known at smaller scales, sometimes at the level of individual soil samples. The aim of this study is to show how point information on soil hydraulic properties, *i.e.*, water retention and hydraulic conductivity characteristics, can be upscaled via soil textural classes to a soil physical units map of a region or nation. Base information is the Dutch soil map (1:50,000) and the hydraulic properties of individual soil samples. All individual soil samples for which hydraulic properties were measured were divided based on their texture into eighteen top-soils and eighteen sub-soils. For each of these thirty-six groups geometric average water retention and hydraulic conductivity characteristics were derived. In total 368 derived soil profiles are distinguished in the Dutch soil map consisting of soil layers that are linked to the thirty-six texture groups. For each soil profile eight static hydraulic properties were calculated. The soil profiles were then clustered based on these properties into seventy-nine clusters or units, which then make up a soil physical units map for the Netherlands. It has been demonstrated that dynamically simulated transpiration reduction for the clustered situation is similar to obtained for all individual soil profiles. At the Dutch national scale, the difference in simulated transpiration reduction between runs using all 368 soil profiles or the 79 soil physical units was less than 2.5 % percentage-points in 96 % of all plots or less than 5 % percentage-points in 99 % of all plots. Similar good correspondence was obtained for other water balance terms as well, including actual transpiration, actual evaporation at the soil surface, degree of saturation at 15 and 30 cm depth, the integrated water flux at 100 cm depth and surface runoff.

### 1. Introduction

Soils and soil functions contribute to land-related ecosystem services in line with the various Sustainable Development Goals (SDGs) and the EU Green Deal (e.g., Bouma, 2020; Bouma et al., 2021; Keesstra et al., 2016; Veerman et al., 2020). Soil types function as carriers of information either for soil survey interpretations, input for dynamic simulation models, spatial analysis, digital mapping techniques, and for communication purposes in the policy arena (Bouma et al., 2022). The degree of detail about this information must match the questions under investigation in land (soil) evaluation studies, such as studies considering the impact of hydrological interventions (e.g., groundwater pumping; surface water level regulation) on agricultural crop production via its influence on (evapo)transpiration (e.g., Hack-ten Broeke et al., 2016; 2019). Agrohydrological simulation models are used in land evaluation studies in which the fate of water, its availability for evapotranspiration, and its impact on the environment are investigated. For example, Hack-

ten Broeke et al. (2019) presented the Watervision Agriculture program for land evaluation in relation to water management. In this program the one-dimensional agrohydrological model SWAP (Kroes et al., 2017) is used, which describes the water movement in soils. The main determining soil properties in SWAP, or any other soil water simulation model, are the soil hydraulic properties, *i.e.*, the water retention and hydraulic conductivity characteristics. The question then arises how to obtain regional or national information regarding these properties, knowing that in most cases these are measured in a relatively small amount of soil samples? One way could be to select from existing measured hydraulic characteristics those that are close to the soils under investigation. Such information can be obtained from existing databases, such as UNSODA (Nemes et al., 2001), HYPRES (Wösten et al., 1999) or its successor (Tóth et al., 2014), or more local data bases available at regional or national level, or even at the scale of the study (e.g., Defterdarović et al., 2021). As an alternative, much research has been performed in developing so-called pedotransfer functions that can

<sup>\*</sup> Corresponding author.

E-mail address: [marius.heinen@wur.nl](mailto:marius.heinen@wur.nl) (M. Heinen).

<sup>1</sup> Retired.

estimate those characteristics based on more easily available/measurable soil texture information (e.g., ROSETTA (Schaap et al., 2001), HYPRES or its successor (Wösten et al., 1999)). The disadvantage of the use of pedotransfer functions is that they apply only to the population it was developed upon, i.e., extrapolation to soils that have non-similar basic properties should be done with care. Moreover, the explained variance by such pedotransfer functions may not always be high. For example, the pedotransfer functions in the HYPRES database predict parameter values for the Mualem (1976) – Van Genuchten (1980) functions with some very low fitting correlation coefficients. This poses questions about the prediction quality for the full curves.

It is not realistic to expect that one will be able to have detailed, measured soil hydraulic characteristics for all soils present in a certain region or nation. Can we think of an alternative way to use the measured hydraulic properties in extrapolating these to other regions? More specifically, can we focus on the possibility to reduce the number of unique soil profiles based on their physical behavior? This is typically useful when land evaluation is based on many scenario simulations, so that a reduction in number of physically unique soil profiles minimizes the total number of simulation runs without losing too much information. The main goal of this paper is to show how point information on soil hydraulic properties can be upscaled via soil textural classes to a soil physical units map of a region or nation. Base information is the Dutch soil map (1:50,000) and the hydraulic properties of individual soil samples. These samples can be allocated to soil textural classes with average hydraulic properties, the so-called Staring series. Combining the soil profiles in the soil map with the Staring series finally will result in a soil physical units map. The performance of this procedure will be checked by comparing simulated transpiration reduction for situations where all soil profiles are considered versus the outcome based on the clustered soil profiles (verification).

2. Material and methods

2.1. Dutch soil map 1:50,000

Soil classification in the Netherlands is mainly based on soil texture. Soil profiles are described based on soil augerings up to 120 cm below soil surface. In total about 350,000 soil profiles have been described over the last several decades in the Netherlands (total surface area just over 41,500 km<sup>2</sup> of which 18.4 % is water; in total the soil map covers 30,686 km<sup>2</sup>). A total of 368 derived (standard) soil profiles are distinguished in the Netherlands (de Vries, 1999). These soil profiles consist of layers with different soil textural properties. These layers are classified based on soil texture: eighteen classes for top-soils (mainly the A-horizon), and eighteen classes for sub-soils (Table 1).

2.2. Individual water retention and hydraulic conductivity characteristics

Soil physical hydraulic properties are often summarized in functional relationships. Here we used the well-known relationships of Mualem (1976) and Van Genuchten (1980) (further denoted by MvG) for the water retention and hydraulic conductivity characteristics, given by.

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

and, provided that  $m = 1-1/n$ ,

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

where  $\theta$  is the volumetric water content (cm<sup>3</sup> cm<sup>-3</sup>),  $h$  is the pressure head (cm),  $K$  is the hydraulic conductivity (cm d<sup>-1</sup>). These relationships are characterized by the following six parameters:  $\theta_r$  is the (asymptotic) residual  $\theta$  (cm<sup>3</sup> cm<sup>-3</sup>),  $\theta_s$  is  $\theta$  at saturation (cm<sup>3</sup> cm<sup>-3</sup>),  $\alpha$  (cm<sup>-1</sup>),  $n$

Table 1

The soil texture-based division of the Dutch soil horizons into eighteen top-soils (B01 – B18) and eighteen sub-soils (O01 – O18). The sand soils are classified based on fraction < 50 μm and the median of the sand grain size (M50), the loam and clay soils are classified based on fraction < 2 μm, the peat(y) soils are classified based on organic matter content (OM) and (for top-soils) on fraction < 2 μm, and the loess soils are classified based on fraction < 50 μm.

| Code         | Description  | < 2 μm (%) | < 50 μm (%) | OM (%) | M50 (μm) |
|--------------|--|------------|-------------|--------|----------|
| <i>Sand</i>  |  |            |             |        |          |
| B01          | poor loamy, very fine to moderate fine sand        |            | 0–10        | 0–15   | 105–210  |
| B02          | weak loamy, very fine to moderate fine sand        |            | 10–18       | 0–15   | 105–210  |
| B03          | strong loamy, very fine to moderate fine sand      |            | 18–33       | 0–15   | 105–210  |
| B04          | very strong loamy, very fine to moderate fine sand |            | 33–50       | 0–15   | 105–210  |
| B05          | coarse sand  |            |             | 0–15   | 210–2000 |
| B06          | boulder clay sand                                  |            | 0–50        | 0–15   | 50–2000  |
| <i>Loam</i>  |  |            |             |        |          |
| B07          | (very) light loam                                  | 8–12       |             | 0–15   |          |
| B08          | moderate light loam                                | 12–18      |             | 0–15   |          |
| B09          | heavy loam   | 18–25      |             | 0–15   |          |
| <i>Clay</i>  |  |            |             |        |          |
| B10          | light clay   | 25–35      |             | 0–15   |          |
| B11          | moderate heavy clay                                | 35–50      |             | 0–15   |          |
| B12          | (very) heavy clay                                  | 50–100     |             | 0–15   |          |
| <i>Loess</i> |  |            |             |        |          |
| B13          | sandy loess  |            | 50–85       | 0–15   |          |
| B14          | silty loess  |            | 85–100      | 0–15   |          |
| <i>Peaty</i> |  |            |             |        |          |
| B15          | peaty sand   | 0–8        |             | 15–25  |          |
| B16          | (sandy) peat                                       | 0–8        |             | 25–100 |          |
| B17          | peaty clay   | 8–100      |             | 16–45  |          |
| B18          | clayey peat  | 8–100      |             | 25–70  |          |
| <i>Sand</i>  |  |            |             |        |          |
| O01          | poor loamy, very fine to moderate fine sand        |            | 0–10        | 0–3    | 105–210  |
| O02          | weak loamy, very fine to moderate fine sand        |            | 10–18       | 0–3    | 105–210  |
| O03          | strong loamy, very fine to moderate fine sand      |            | 18–33       | 0–3    | 105–210  |
| O04          | very strong loamy, very fine to moderate fine sand |            | 33–50       | 0–3    | 105–210  |
| O05          | coarse sand  |            |             | 0–3    | 210–2000 |
| O06          | boulder clay sand                                  |            | 0–50        | 0–3    | 50–2000  |
| O07          | brook loam sand                                    |            | 33–50       | 0–3    | 50–150   |
| <i>Loam</i>  |  |            |             |        |          |
| O08          | (very) light loam                                  | 8–12       |             | 0–3    |          |
| O09          | moderate light loam                                | 12–18      |             | 0–3    |          |
| O10          | heavy loam   | 18–25      |             | 0–3    |          |
| <i>Clay</i>  |  |            |             |        |          |
| O11          | light clay   | 25–35      |             | 0–3    |          |
| O12          | moderate heavy clay                                | 35–50      |             | 0–3    |          |
| O13          | (very) heavy clay                                  | 50–100     |             | 0–3    |          |
| <i>Loess</i> |  |            |             |        |          |
| O14          | sandy loess  |            | 50–85       | 0–3    |          |
| O15          | silty loess  |            | 85–100      | 0–3    |          |
| <i>Peat</i>  |  |            |             |        |          |
| O16          | oligotrophic peat                                  |            |             | 35–100 |          |
| O17          | mesotrophic and eutrophic peat                     |            |             | 35–100 |          |
| O18          | peaty interlayer                                   |            |             | 15–35  |          |

(dimensionless),  $m = 1-1/n$  (dimensionless) and  $\lambda$  (dimensionless) are curve shape parameters, and  $K_s$  is  $K$  at saturation (cm d<sup>-1</sup>). In simulation models that describe water movement in unsaturated soils the derivative of the water retention characteristic is often needed, which is known as the differential moisture capacity,  $C$  (cm<sup>-1</sup>), here given by.

$$C(h) = \frac{d\theta}{dh} = \alpha n m (\theta_s - \theta_r) |\alpha h|^{n-1} (1 + |\alpha h|^n)^{-1-m} \tag{3}$$

Over the last six decades 999 soil samples were analyzed on water retention and conductivity. Based on their soil texture, these samples can be assigned to one of the eighteen top-soils or eighteen sub-soils from Table 1. Historic data (until 2001; Wösten et al., 2001) were often obtained from sites where experiments were carried out and were analyzed with a variety of methods (sand box, hot air, evaporation, sorptivity; see, e.g., Stolte et al., 1994). Since 2001 the database is being extended by specifically sampling soil layers that were not well represented in the database at that time, by getting a better spread of the total area in the Netherlands, and by standardizing the laboratory and data analyses methods. All samples are now analyzed using the evaporation method to obtain  $\theta(h)$  and  $K(h)$  data for  $-800 < h < 0$  cm (using cylindrical samples with diameter 10 cm and height 8 cm with tensiometers installed at four depths; based on Wind, 1966; Boels et al., 1978), pressure plate  $\theta(h)$  for  $h = -1,000$  and  $-16,000$  cm (e.g., Dane & Hopmans, 2002), air dry  $\theta(h)$  for  $h$  is approximately  $-10^6$  cm depending on relative vapor pressure (e.g., Koorevaar et al., 1983), and hydraulic conductivity at saturation  $K(h = 0)$  via the constant head method using samples with diameter 20 cm and height 10 cm (e.g., Reynolds & Elrick, 2002; de Vos, 1997). This results in measured  $\theta(h)$  and  $K(h)$  data which then are used to get estimates for the MvG parameters by using the RETC parameter optimization program (van Genuchten et al., 1991). Details can be found in Bakker et al. (2019).

### 2.3. Average water retention and hydraulic conductivity characteristics: The Staring series

For each of the thirty-six soil texture classes from Table 1 several individual soil samples are present, each having their own parameterized MvG relationships. Within a class, for all individual members the water contents and conductivities at thirteen pressure heads (0, -10, -20, -31, -50, -100, -250, -500, -1000, -2500, -5000, -10,000, -16,000 cm) were calculated. Then, for each pressure head the water contents and conductivities were geometrically averaged. Geometric averaging was used since in many cases both the water contents and conductivities thus obtained showed skewed distributions (data not shown). This is in line with Wösten and van Genuchten (1988) and Wösten et al. (1999). The resulting thirteen  $\theta(h)$  and thirteen  $K(h)$  data pairs per texture class were again parameterized according to the MvG-model using RETC (Van Genuchten et al., 1991). This procedure was used in earlier versions of the Staring series (Wösten et al., 2001) as well as in a European database HYPRES (Wösten et al., 1999). Here we fixed  $\theta_r$  to the same values as was done in Wösten et al. (2001). As some of the samples were more than 30 years old, and these samples were partly measured with methods that are out of use at present, we introduced a weighting factor in favor of newer samples. In case more than ten samples were available that were analyzed since 2001, all old samples in the same class were not considered. Otherwise, a weighting based on the ratio new samples versus old samples was employed (details in Heinen et al., 2020). Fig. 1 provides an overview of the sampling locations in The Netherlands.

This results in sets of MvG parameters for all thirty-six soil textural classes: the so-called Staring series. Each soil layer of all 368 derived soil profiles of the Dutch soil map can be linked to one of the thirty-six hydraulic properties of the Staring series.

### 2.4. Soil physical units map: BOFEK

In deriving the soil physical units map (BOFEK, a Dutch acronym for soil physical units map) the following procedure was performed. Although a total of 368 derived soil profiles are present (de Vries, 1999), some of these soil profiles have similar sequences of soil layers, or soil layers which might behave similar in terms of water retention and conductivity. In that case a further aggregation can be considered.

To determine which of the 368 derived soil profiles can be aggregated in a soil hydrological sense, the hydrological behavior of each derived soil profile needs to be determined; such was previously

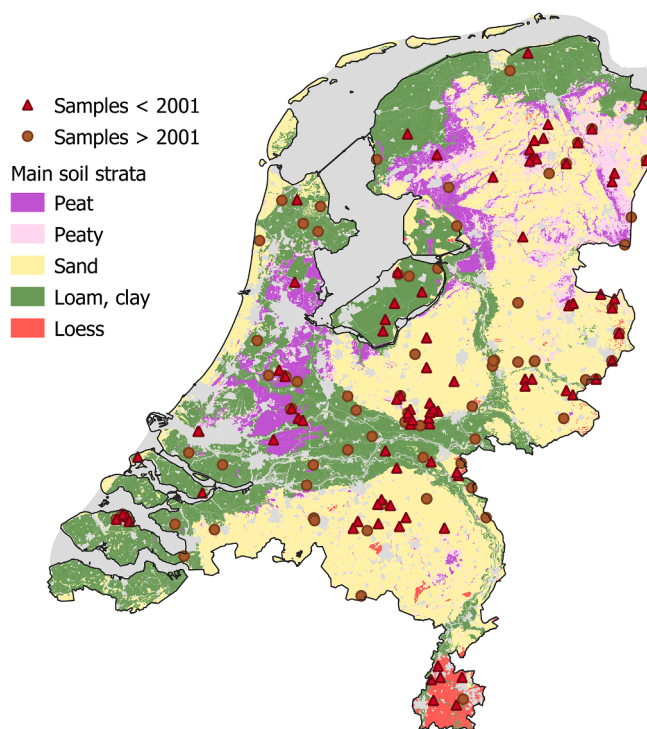


Fig. 1. Soil sample locations (often in twofold, often multiple soil layers) split in old (<2001) and new (>2001) years, projected on the main soil strata in the Netherlands (as described later in text).

suggested by Wösten et al. (1986) and Wösten and van Genuchten (1988). For this purpose, for each of the 368 individual soil profiles eight static, hydrological properties were calculated. These eight variables were selected from a total set of twenty-three, for which in many cases high correlations were obtained (details in Heinen et al., 2021) making most of them redundant. The eight selected properties will be described below and are visualized in Fig. 2.

The hydraulic resistance,  $c$  (d), of the soil profile is given by.

$$c = \sum_{i=1}^N \frac{\Delta z_i}{K_{s,i}} \quad (4)$$

where  $\Delta z$  is the thickness of the soil layer (cm),  $K_s$  is the hydraulic conductivity at saturation of the soil layer ( $\text{cm d}^{-1}$ ), and  $N$  is the total number of soil layers within the soil profile.

The total available water content in the rootzone is often defined as the difference in water content between pressure heads of  $-100$  and  $-16,000$  cm. Here we split this between a part that is readily available, without stress ( $W_{av,1}$ , cm) and another part that is available under stress ( $W_{av,2}$ , cm), separated at the pressure head of  $-400$  cm.

$$W_{av,1} = \sum_{i=1}^M (\theta(-100)_i - \theta(-400)_i) \Delta z_i \quad (5)$$

$$W_{av,2} = \sum_{i=1}^M (\theta(-400)_i - \theta(-16,000)_i) \Delta z_i \quad (6)$$

where  $M$  is the number of soil layers in the root zone, here taken as the layer 0–30 cm. Here the available water content is multiplied by the thickness of each layer so that a thickness corresponding to water storage (in cm) is obtained. In case the bottom of the root zone falls inside a soil layer, the last value for  $\Delta z_M$  is truncated accordingly.

Besides the potentially available water considered above, the ability of the subsoil to provide water from the subsoil via capillary rise is another property of interest. Here we consider the maximum depth of

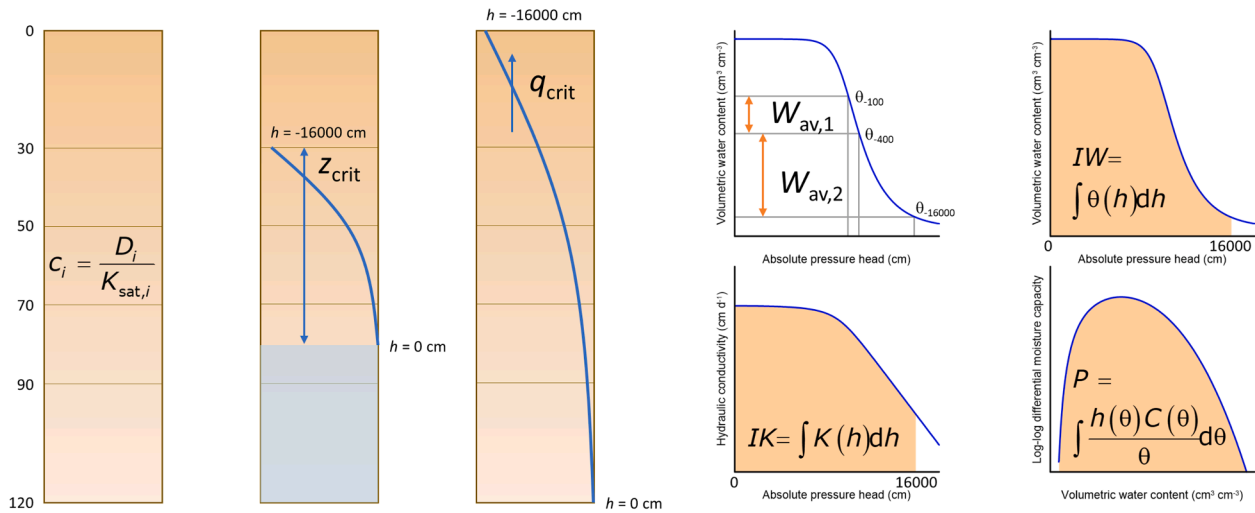


Fig. 2. Schematic visualization of the eight, static soil properties that were calculated for all 368 derived soil profiles.

the water table below the bottom of the root zone such that this part of the soil can supply a certain water flux when the pressure head at the bottom of the root zone is at wilting point ( $h = -16,000$  cm). This distance or critical capillary height,  $z_{crit}$  (cm), follows from the Darcy expression for water flux density integrated over depth for a given constant flux. The depth-pressure head profile then follows from

$$z_{crit} = \int_{-16,000}^0 \frac{K(h)_i}{K(h)_i + q_{0.2}} dh \quad (7)$$

where  $q_{0.2}$  is the desired water flux of  $0.2 \text{ cm d}^{-1}$ . The value of  $z_{crit}$  is obtained via numerical integration. This is done stepwise, starting at the bottom of the root zone ( $z = 30$  cm) where  $h = -16,000$  cm with step size  $dh = 1$  cm and calculating the corresponding depth. When a boundary between two layers is encountered (and thus the MvG parameters change), the step size  $dh$  is repeatedly halved until the exact depth ( $\pm 10^{-12}$  cm) of this boundary is reached; thereafter, the step size is reset to  $dh = 1$  cm.

The drawback of  $z_{crit}$  is that it not always includes the whole soil profile: the upper 30 cm (root zone) is excluded, and sometimes  $z_{crit} < 90$  cm, so that the lower part of the soil profile is not included. Therefore, it was decided to determine the critical flux that this profile can provide when the groundwater level is located at its bottom ( $h = 0$  cm at  $z = 120$  cm), and when the soil surface is at wilting point ( $h = -16,000$  cm at  $z = 0$  cm). This critical flux,  $q_{crit}$  ( $\text{cm d}^{-1}$ ), follows from

$$\int_{-16,000}^0 \frac{K(h)_i}{K(h)_i + q_{crit}} dh = 120 \quad (8)$$

where  $q_{crit}$  is obtained iteratively (using Brent's method, Press et al., 1992) by repeating the solution procedure as described above for obtaining  $z_{crit}$ .

Whereas water availability is based on only part of the water retention characteristic, we also considered the integral of the complete water retention characteristic according to

$$IW = \sum_{i=1}^N \left[ \int_{-16,000}^0 \theta(h)_i dh \right] \frac{\Delta z_i}{L} \quad (9)$$

where  $L = \sum_{i=1}^N \Delta z_i$  is the total length of the soil profile (120 cm). Note that this integral property is a weighted sum based on the relative thickness of the soil layers.

The critical capillary height and critical capillary upward flux are based on the integral of the hydraulic conductivity curves, and here we also consider the actual integral of these curves (a weighted sum based on the relative thickness of the soil layers)

$$IK = \sum_{i=1}^N \left[ \int_{-16,000}^0 K(h)_i dh \right] \frac{\Delta z_i}{L} \quad (10)$$

Note that  $IK$  is also known as the matric flux potential (Raats, 1970), a useful property in solving specific flow problems, like in some root water uptake models (e.g., De Jong van Lier et al., 2008, 2013; De Willigen et al., 2012).

Besides the soil water retention and hydraulic conductivity relationships, also the differential moisture capacity, the derivative of the water retention characteristic, is an important quantity in many soil water movement simulation models. Therefore, it was decided to include the integral property of this differential moisture capacity. Here we used the  $P$  index of Haverkamp et al. (2005) given by (a weighted sum based on the relative thickness of the soil layers)

$$P = \sum_{i=1}^N \left[ \frac{1}{\theta_{s,i} - \theta_{r,i}} \int_{\theta_{r,i}}^{\theta_{s,i}} \frac{d(\ln(\theta)_i)}{d(\ln(h)_i)} d\theta \right] \frac{\Delta z_i}{L} \\ = \sum_{i=1}^N \left[ \frac{1}{(\theta_{s,i} - \theta_{r,i})} \int_{\theta_{r,i}}^{\theta_{s,i}} \frac{h(\theta)_i C(\theta)_i}{\theta_i} d\theta \right] \frac{\Delta z_i}{L} \quad (11)$$

Note that the  $P$  index originally refers to the slope of the In-transformed water retention curve, which after rewriting (see Appendix A) equals the product of the water retention curve and the differential moisture capacity curve divided by  $\theta$ . According to Haverkamp et al. (2005)  $P$  is in the range  $[0,3]$  and according to their analyses  $P$  is often less than 1.

As stated before, the integrals in the expressions for  $z_{crit}$  and  $q_{crit}$  need to be solved numerically since no analytical expressions are available. The integral properties  $IW$ ,  $IK$  and  $P$ , however, can be expressed analytically in terms of the Gauss hypergeometric function (Appendix A) for which a fast computational solution is available (Michel and Stoitsov, 2008).

## 2.5. Clustering

For all 368 soil profiles the eight parameters were calculated. One way to determine which profiles have similar (or identical) properties is by clustering. Several clustering techniques are available in the literature, and several performance criteria can be used. Here we used the well-known k-means clustering technique (Hartigan and Wong, 1979). Often the number of clusters is determined based on the elbow method (Thorndike, 1953) which shows how the summed variation within the clusters decreases when the total number of clusters increases. This elbow method can be extended by looking at the summed internal



variations relative to the total variation of all data. This measure is comparable to the percentage of variation accounted for in classical statistics and is more intuitive to judge. In this way one can prescribe a certain percentage variation accounted for, and from that determine the number of clusters to choose. In order to prevent the possibility that clusters would arise where sand, clay or peat soils were mixed, first a stratification was performed. We distinguished five major soil types: peat soils (more than 50 % is peat in 0–80 cm; 44 profiles), peaty soils (less than 50 % is peat in 0–80 cm; 22 profiles), sand (less than 8 % <2  $\mu\text{m}$  and less than 50 % <50  $\mu\text{m}$  and less than 15 % organic matter; 135 profiles), loam (8–25 % <2  $\mu\text{m}$ ) and clay soils (25–100 % <2  $\mu\text{m}$ ; 151 profiles), and loess (50–100 % <50  $\mu\text{m}$ ; 16 profiles) soils. The 368 soil profiles were divided in these five strata, and clustering was performed for each of these strata. Fig. 1 shows the geographical spread of these five main soil strata in The Netherlands.

Clustering was performed in R (R Core Team, 2020) with the function ‘kmeans’ of the stats-package, using the default algorithm ‘Hartigan-Wong’ with 100 random sets (starts) and 100 maximum iterations. Prior to the clustering, all data were standardized by subtracting the mean and dividing this result by the standard deviation for each of the eight properties.

Each cluster thus obtained represents a soil physical unit, consisting of soil profiles that have similar physical properties. For future calculations one can use the hydraulic properties of one of these members and set the outcome thus obtained also for the other soil profiles. Here we choose the profile with the largest area according to the soil map as the representative for the specific unit.

## 2.6. Verification procedure

The construction of the soil physical units (map), as described above, relies on the similarity in the eight static properties. The question then arises: how is the behaviour within and between units for transient situations? The total agricultural area in the Netherlands is divided in 331,109 plots with a resolution of 250 m, each having its own characteristics regarding crop type, soil type, meteorology and groundwater levels. The groundwater levels were imposed and were taken from the Dutch national MODFLOW-based groundwater model (de Lange et al., 2014). Two simulation runs were performed: i) each plot was assigned to one of the 368 derived soil profiles and ii) each plot was assigned to one of the BOFEK units. Simulations were performed for a period of thirty years (1991–2020). For crop production related evaluation studies the reduction in transpiration (transpiration ratio), which results in reduction of crop growth, can be regarded as one of the major items to consider in soil water balance simulations. Differences in the thirty-year average transpiration ratios between the two simulation runs were presented. In addition, results were provided for thirty-year average relative differences in actual transpiration, actual soil evaporation at the soil surface, the degree of saturation (volumetric water content divided by volumetric water content at saturation) at depths of 15 and 30 cm, the integrated water flux at depth 100 cm and surface runoff (overland flow). Relative here means relative to the 30-year average results obtained from the runs with the 368 soil profiles.

Simulations were performed with the SWAP-WOFOST model (Kroes et al., 2017). The crop growth model WOFOST (Boogaard et al., 2014) was integrated in the SWAP model to simulate the potential and actual crop development of arable crops. For grassland an attuned version of WOFOST is embedded in SWAP (Kroes and Supit, 2011) and for other crops a simple crop growth module is used assuming prescribed crop development as a function of time. SWAP simulates water movement in unsaturated–saturated soils and calculates the actual root water uptake. Actual root water uptake becomes less than the potential demand for water when the soil dries out. This is described by the Feddes reduction function (Feddes et al., 1978). When (part of) the root zone becomes too wet, roots suffer from oxygen deficiency and reduction in root water uptake will occur. The oxygen stress is described according to

Bartholomeus et al. (2008). In the Supplemental Material details are provided for the root water uptake reduction functions for drought and oxygen stress.

## 3. Results and discussion

### 3.1. Staring series

The Mualem – Van Genuchten parameters for the thirty-six Staring series soil classes are given in Table 2.

The thirty-six soil textural classes are characterized by band widths of soil textures (see Table 1). This alone may result in quite some differences in soil hydraulic properties measured on different soil samples that fall within the same class. This was also observed in our analysis (data not shown). But also the dry bulk density might influence water retention and conductivity. This was studied by Assouline (2006a,b) and further elaborated upon based on the MvG parameterization by Tian et al. (2019) and Kool et al. (2019). This would suggest that these characteristics best be normalized to the same dry bulk density. However, according to the studies by Tian et al. (2019) and Kool et al. (2019) there is still some debate on the best way how to transform the  $n$  parameter as a function of dry bulk density. We thus believe that more research in this area is needed (globally and locally) before this can be accounted for.

### 3.2. BOFEK

The selected eight hydrological properties are not fully uncorrelated (Fig. 3). Since they are all based on the MvG parameterization, it is likely

**Table 2**

The average Mualem – Van Genuchten parameters for the eighteen top-soils (B01–B18) and eighteen subsoils (O01–O18) ( $m = 1 - 1/n$ ).

| Name | $\theta_r$ (cm <sup>3</sup> cm <sup>-3</sup> ) | $\theta_s$ (cm <sup>3</sup> cm <sup>-3</sup> ) | $\alpha$ (cm <sup>-1</sup> ) | $n$ (-) | $\lambda$ (-) | $K_s$ (cm d <sup>-1</sup> ) |
|------|--|--|------------------------------|---------|---------------|-----------------------------|
| B01  | 0.02   | 0.427  | 0.0217                       | 1.735   | 0.981         | 31.23                       |
| B02  | 0.02   | 0.434  | 0.0216                       | 1.349   | 7.202         | 83.24                       |
| B03  | 0.02   | 0.443  | 0.0150                       | 1.505   | 0.139         | 19.08                       |
| B04  | 0.02   | 0.462  | 0.0149                       | 1.397   | 0.295         | 34.88                       |
| B05  | 0.01   | 0.381  | 0.0428                       | 1.808   | 0.024         | 63.65                       |
| B06  | 0.01   | 0.385  | 0.0209                       | 1.242   | -1.200        | 104.10                      |
| B07  | 0  | 0.401  | 0.0183                       | 1.248   | 0.952         | 14.58                       |
| B08  | 0.01   | 0.433  | 0.0105                       | 1.278   | -1.919        | 3.00                        |
| B09  | 0  | 0.430  | 0.0070                       | 1.267   | -2.387        | 1.75                        |
| B10  | 0.01   | 0.448  | 0.0128                       | 1.135   | 4.581         | 3.83                        |
| B11  | 0.01   | 0.591  | 0.0216                       | 1.107   | -5.549        | 6.31                        |
| B12  | 0.01   | 0.530  | 0.0166                       | 1.091   | -4.494        | 2.25                        |
| B13  | 0.01   | 0.416  | 0.0084                       | 1.437   | -1.357        | 29.83                       |
| B14  | 0.01   | 0.417  | 0.0054                       | 1.302   | -0.335        | 0.90                        |
| B15  | 0.01   | 0.528  | 0.0237                       | 1.282   | -1.478        | 87.45                       |
| B16  | 0.01   | 0.786  | 0.0211                       | 1.279   | -1.221        | 12.36                       |
| B17  | 0  | 0.719  | 0.0191                       | 1.137   | 0.000         | 4.48                        |
| B18  | 0  | 0.765  | 0.0205                       | 1.151   | 0.000         | 13.14                       |
| O01  | 0.01   | 0.366  | 0.0160                       | 2.163   | 2.868         | 22.32                       |
| O02  | 0.02   | 0.387  | 0.0161                       | 1.524   | 2.440         | 22.76                       |
| O03  | 0.01   | 0.340  | 0.0172                       | 1.703   | 0.000         | 12.37                       |
| O04  | 0.01   | 0.364  | 0.0136                       | 1.488   | 2.179         | 25.81                       |
| O05  | 0.01   | 0.337  | 0.0303                       | 2.888   | 0.074         | 17.42                       |
| O06  | 0.01   | 0.333  | 0.0160                       | 1.289   | -1.010        | 32.83                       |
| O07  | 0.01   | 0.513  | 0.0120                       | 1.153   | -2.013        | 37.55                       |
| O08  | 0  | 0.454  | 0.0113                       | 1.346   | -0.904        | 8.64                        |
| O09  | 0  | 0.458  | 0.0097                       | 1.376   | -1.013        | 3.77                        |
| O10  | 0.01   | 0.472  | 0.0100                       | 1.246   | -0.793        | 2.30                        |
| O11  | 0  | 0.444  | 0.0143                       | 1.126   | 2.357         | 2.12                        |
| O12  | 0.01   | 0.561  | 0.0088                       | 1.158   | -3.172        | 1.08                        |
| O13  | 0.01   | 0.573  | 0.0279                       | 1.080   | -6.091        | 9.69                        |
| O14  | 0.01   | 0.394  | 0.0033                       | 1.617   | 0.514         | 2.50                        |
| O15  | 0.01   | 0.410  | 0.0078                       | 1.287   | 0.000         | 2.79                        |
| O16  | 0  | 0.889  | 0.0097                       | 1.364   | -0.665        | 1.46                        |
| O17  | 0.01   | 0.849  | 0.0119                       | 1.272   | -1.249        | 3.40                        |
| O18  | 0.01   | 0.580  | 0.0127                       | 1.316   | -0.786        | 35.95                       |

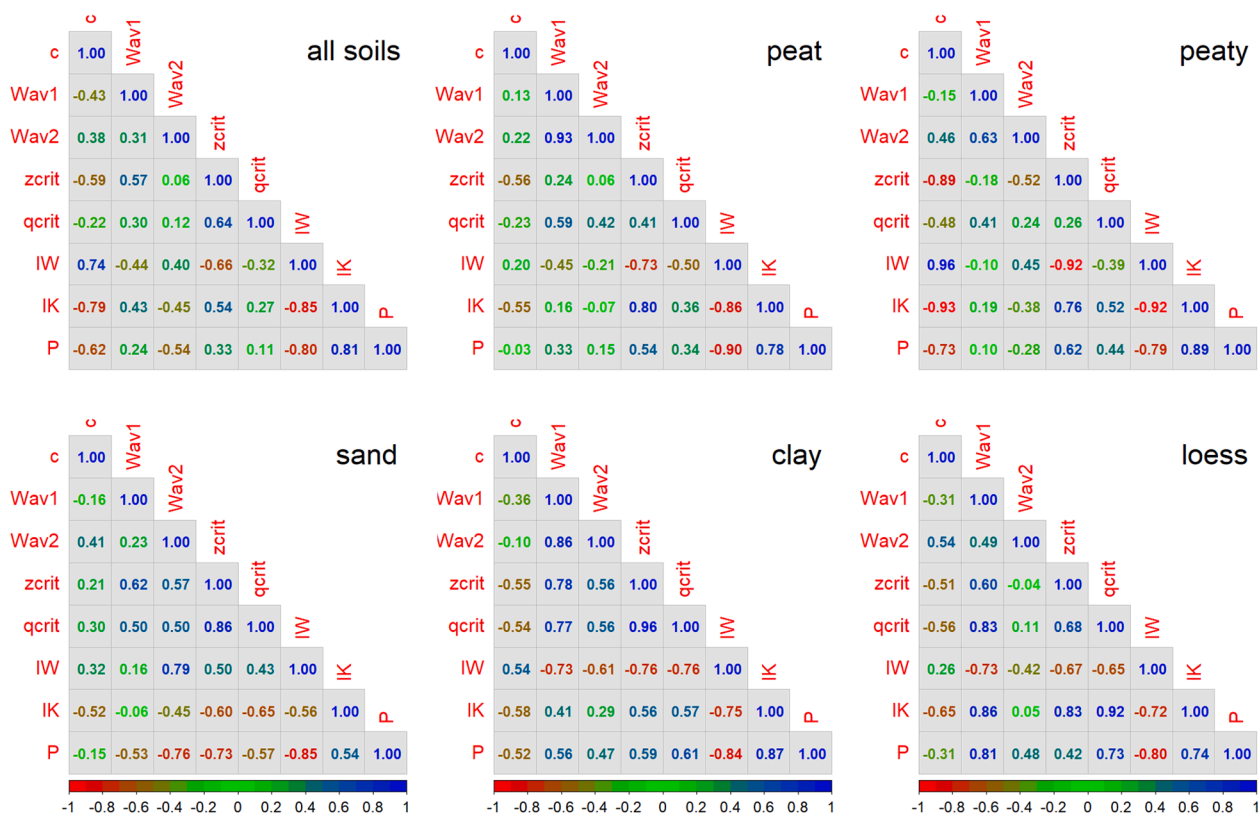


Fig. 3. Correlation (Pearson) matrices for the eight physical properties for all 368 soils profiles and for soils within the five soil strata.

that some of these are still correlated to some extent. The current choice consists of rather well-known quantities that are often used in (agro) hydrology (i.e.,  $c$ ,  $W_{av}$ ,  $z_{crit}$ ). New properties are the critical flux, and the integral properties of the retention, conductivity and differential moisture capacity characteristics. We do observe some high correlations, but often these are not consistent among the five major soil types. The only exception is the relatively high (negative) correlation between the  $P$  index and the integral of the water retention characteristic ( $IW$ ). This is, of course, because the expression of  $P$  contains the integral of the water retention characteristic. However, since the  $P$  index also contains information about  $C$  it was decided to keep  $P$  in the further analysis.

The relationship between the percentage of variance accounted for (PVA) and the number of clusters or units ( $k$ ) is shown in Fig. 4 for the

five main soil types (strata). In all cases PVA increases monotonically with increasing  $k$ . Of course, PVA becomes 100 % when  $k$  is equal to the total number of soil profiles in a stratum. When the desired PVA is set at 95 %, the total number of units is 79: 18 peat, 7 peaty, 23 sand, 24 loam/clay and 7 loess soil physical units. When the 95 % level occurred between two integer values of  $k$  the largest of these two values was chosen.

When a less strict PVA criterion was used (e.g., 92.5 %) the reduction in units was about 25 % (59 units), but the memberships of soil profiles within each unit was judged as less good than in case  $PVA = 95\%$ .

As an example, Table 3 provides the resulting division into units for the loess profiles based on the eight properties. The sixteen soil profiles are clustered into seven units (5001–5007; the numbering is arbitrary). Some units contain only a single soil profile, while other units contain

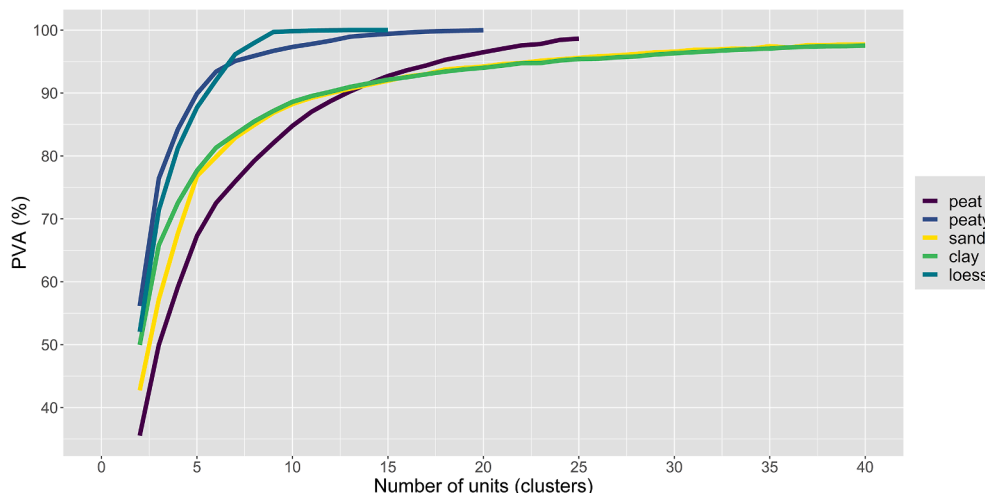


Fig. 4. Relationship between percentage of variance accounted for (PVA) and the number of units (clusters) for the five major soil strata.

**Table 3**

The eight soil hydrological properties ( $c$ ,  $W_{av,1}$ ,  $W_{av,2}$ ,  $z_{crit}$ ,  $q_{crit}$ ,  $IW$ ,  $IK$ ,  $P$ ) for the sixteen loess soil profiles and the division of these into seven BOFEK units. The profile codes refer to the Dutch soil map identifiers and the area refers to the total area for this profile in The Netherlands; note that the profiles within a unit are ordered from high-to-low area.

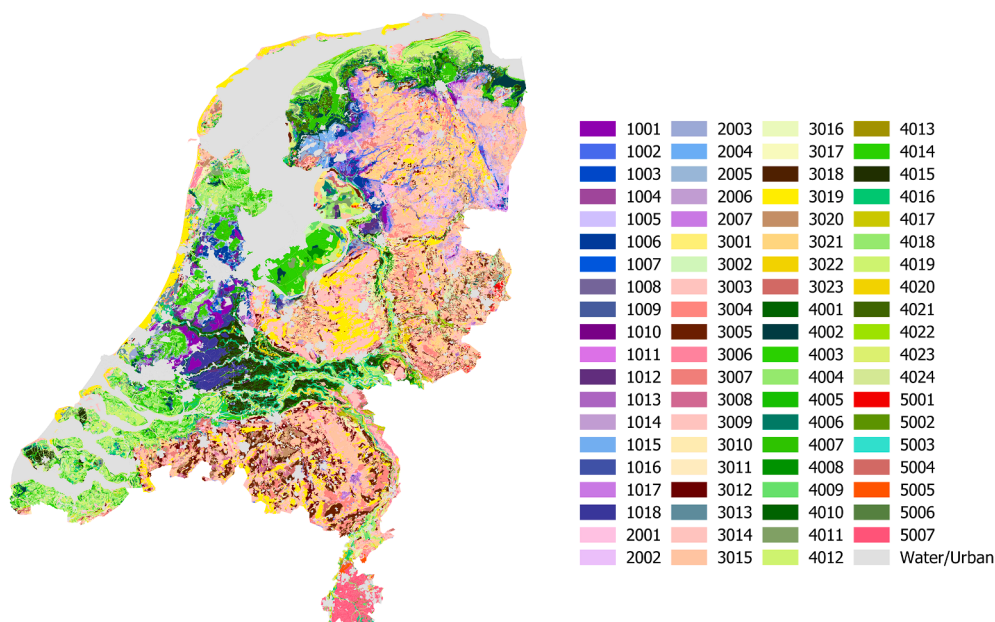
| Profile | Area (ha) | $c$ (d) | $W_{av,1}$ (cm) | $W_{av,2}$ (cm) | $z_{crit}$ (cm) | $q_{crit}$ (cm d <sup>-1</sup> ) | $IW$ (cm) | $IK$ (cm <sup>2</sup> d <sup>-1</sup> ) | $P$ (-) | Unit |
|---------|-----------|---------|-----------------|-----------------|-----------------|----------------------------------|-----------|---|---------|------|
| 22,010  | 6199.98   | 3.34    | 2.15            | 3.75            | 200.39          | 0.609                            | 1671.2    | 304.2                                   | 0.203   | 5001 |
| 22,020  | 423.22    | 13.77   | 1.72            | 4.24            | 22.82           | 0.022                            | 6075.7    | 7.9                                     | 0.073   | 5002 |
| 22,011  | 881.38    | 93.55   | 2.43            | 5.61            | 32.94           | 0.039                            | 4214.7    | 14.4                                    | 0.152   | 5003 |
| 18,020  | 3385.49   | 28.20   | 3.30            | 5.68            | 159.62          | 1.473                            | 1261.0    | 359.7                                   | 0.354   | 5004 |
| 18,010  | 3349.17   | 29.31   | 3.30            | 5.68            | 175.57          | 1.449                            | 1133.7    | 346.9                                   | 0.383   |      |
| 18,031  | 2172.32   | 22.91   | 3.30            | 5.68            | 156.62          | 1.250                            | 988.1     | 351.2                                   | 0.389   |      |
| 18,030  | 1304.42   | 28.60   | 3.30            | 5.68            | 119.81          | 1.504                            | 1066.3    | 376.3                                   | 0.519   |      |
| 18,021  | 561.15    | 28.60   | 3.30            | 5.68            | 119.81          | 1.504                            | 1066.3    | 376.3                                   | 0.519   |      |
| 5010    | 1647.80   | 39.46   | 3.20            | 5.96            | 148.67          | 0.389                            | 1607.2    | 272.8                                   | 0.332   | 5005 |
| 5011    | 865.53    | 36.03   | 3.20            | 5.96            | 136.97          | 0.397                            | 1512.0    | 284.8                                   | 0.380   |      |
| 9020    | 212.36    | 26.05   | 3.39            | 5.41            | 315.95          | 1.706                            | 1392.6    | 527.0                                   | 0.346   | 5006 |
| 18,050  | 13151.40  | 61.97   | 2.23            | 5.73            | 93.34           | 0.122                            | 2400.3    | 38.9                                    | 0.212   | 5007 |
| 5020    | 10760.84  | 61.97   | 2.23            | 5.73            | 93.34           | 0.122                            | 2400.3    | 38.9                                    | 0.212   |      |
| 5030    | 10405.31  | 58.17   | 2.25            | 5.64            | 93.34           | 0.124                            | 2393.6    | 39.8                                    | 0.211   |      |
| 18,040  | 2164.33   | 61.97   | 2.23            | 5.73            | 93.34           | 0.122                            | 2400.3    | 38.9                                    | 0.212   |      |
| 18,041  | 782.05    | 52.94   | 2.23            | 5.73            | 104.45          | 0.130                            | 1853.6    | 115.7                                   | 0.403   |      |

several profiles. Some variation in properties within a unit can be observed, but the variation between the units is more evident. Within units that encompass more than one soil profile, the one with the largest area is chosen as the representative soil profile. For units 5004, 5005 and 5007 these are the first ones mentioned, *i.e.*, 18020, 5010 and 18050. The complete table for all 368 soil profiles divided over the 79 BOFEK units is available at <https://www.wur.nl/nl/show/Bodemfysische-Eenhedenkaart-BOFEK2020.htm> (in Dutch).

Since each of the 368 soil profiles of the Dutch 1:50,000 soil map has been assigned to one of these 79 units, a soil physical units map can be constructed: BOFEK (Fig. 5). This map greatly resembles the Dutch soil map, but the legend now consists of 79 units versus 368 units in the soil

map. The great resemblance is also a result of the stratification employed.

The approach described in this study was based on an existing soil map, an existing soil textural classification and existing soil physical properties for the soil textural classes. The approach followed by Kozłowski and Komisarek (2017) shows some similarities with our approach. However, they did not make use of own measured hydraulic properties, but derived them from pedotransfer functions from the ROSETTA database. The number of units  $s$  they used was 16 without defining how this number was obtained. Likely this number was chosen in order to compare the results against their 16 Polish textural classes. Périard et al. (2017) performed a soil clustering based on both



**Fig. 5.** The soil physical units map of the Netherlands (BOFEK) with 79 units. The first digit of the four-digit numbers refer to the five main strata: 1xxx = peat, 2xxx = peaty, 3xxx = sand, 4xxx = loam/clay, 5xxx = loess.

physicochemical and hydraulic properties, of which one was the matric flux potential (cf. Eq. (10)). Their focus was on specific locations (soils with cranberry production) in Canada, and not yet nationwide. Jin et al., (2015) performed analyses where all measured information was used in a clustering analysis, from which unique units were derived.

### 3.3. Verification

Figure 6 shows the difference in percentage-points in transpiration ratio between the case where only BOFEK soil profiles were considered versus the case where all 368 soil profiles were considered. Small differences are seen, *i.e.*, 96 % of all plots fall in the range  $\pm 2.5$  % and 99 % of all plots fall in the range  $\pm 5$  % (see frequency distribution in Fig. 7). For many plots the difference is exactly zero, since for these plots the same soil profiles were used in the two runs (46 % of all plots). Within the class  $[-2.5$  %,  $+2.5$  %] 16 % of the total number of plots had an absolute difference of  $<0.1$  %. When performing the clustering, as discussed above, we accepted that not all variance can be accounted for. So, some variations in the outcome in this example could be expected. The observed differences are regarded as very small and negligible.

Table 4 presents the percentage distribution of the transpiration reduction ( $T_{red}$ ), actual transpiration ( $T_{act}$ ), actual evaporation at the soil surface ( $E_{act}$ ), degree of saturation at 15 ( $S_{15}$ ) and 30 ( $S_{30}$ ) cm depth, the integrated water flux at 100 cm depth ( $F_{100}$ ) and surface runoff (RO). The similarity between the simulations performed with the 368 soil profiles and with the 79 BOFEK profiles is large (the least for runoff): for  $>90$  % of all plots the absolute value of the deviation is less than 5 %; the best correspondence was obtained for  $T_{red}$  (99 %), followed by  $S_{15}$ ,  $E_{act}$ ,  $T_{act}$ ,  $S_{30}$  and  $F_{100}$  (90 %). On a relative basis the similarity for surface

runoff was somewhat less. The amount of runoff in the Netherlands is small, Van Bakel et al. (2008) estimated that long-term average runoff (1971–2000) ranges from nearly zero to  $10 \text{ mm yr}^{-1}$  for sandy soils and from 10 to  $50 \text{ mm yr}^{-1}$  for clay soils. In our simulations the average yearly surface runoff was 30 mm, which falls in this range. Because of this small reference value, each mm deviation in surface runoff means already  $>3$  % deviation.

Note that the verification method employed here is not meant to verify whether the simulated reduction in transpiration is correct in an absolute sense. It is only meant to show that, when using the limited number of soil profiles for covering the complete soil map, similar predictions in transpiration reduction were simulated compared to the situation where all 368 soil profiles are used and to show that within a unit the variation of predicted transpiration reduction is small. In the Supplemental Material it is shown for four BOFEK units to what extent the transpiration reduction varies between all soil profiles that belong to the same soil physical unit and how this differs from year-to-year.

### 3.4. General discussion

In our study we have used the classical uni-modal representations of Muallem – Van Genuchten. We are aware that several other as well as multi-modal functional relationships are available. The approach presented here could easily be repeated for any other functional description.

Clustering is defined as the task of grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar (in some sense) to each other than to those in other groups (clusters). This implies that there is no unique division or best division. By dropping one

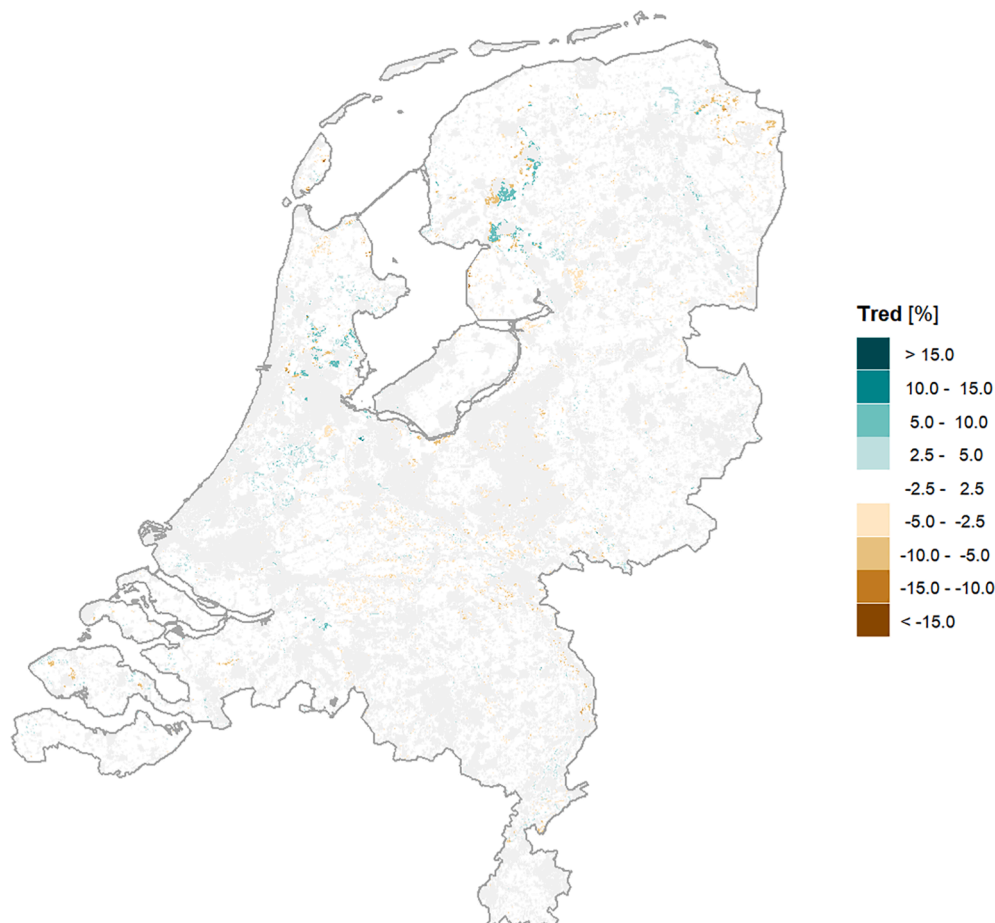
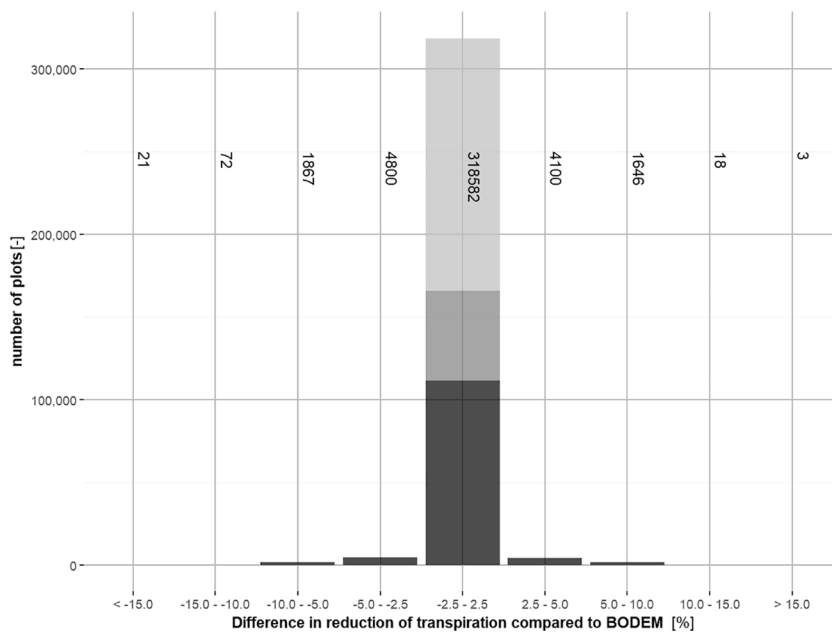


Fig. 6. Difference in transpiration reduction ( $T_{red}$ , %) between simulations based on the 79 BOFEK soil units and simulations based on all 368 derived soil profiles. The grey regions refer to urban areas, water bodies and nature reserves (which were all excluded in the simulations).





**Fig. 7.** Distribution of the difference in transpiration reduction between simulations based on the 79 BOFEK soil units and simulations based on all 368 derived soil profiles. The total number of plots was 331109. The central bar is divided into three parts: the upper light grey part shows the number of plots (152553) which were identical for the two runs, the middle dark grey part refers to the number of plots where the difference was in the range  $-0.1\%$  and  $+0.1\%$  (54564 plots), the lower black part refers to the remaining plots (166029 – 54654 = 111375).

**Table 4**

Distribution (in percentages) of the difference in transpiration reduction ( $T_{red}$ ), actual transpiration ( $T_{act}$ ), actual evaporation at the soil surface ( $E_{act}$ ), degree of saturation at 15 ( $S_{15}$ ) and 30 ( $S_{30}$ ) cm depth, the integrated water flux at 100 cm depth ( $F_{100}$ ) and surface runoff ( $RO$ ), between simulations based on the 79 BOFEK soil units and simulations based on all 368 derived soil profiles. Results pertain to the provided classes. The class  $-2.5..2.5\%$  includes the 46.1 % of the plots that resulted in identical results, and the values given between brackets refer to the sub-class  $-0.1..0.1\%$ .

|               | $T_{red}$   | $T_{act}$   | $E_{act}$   | $S_{15}$    | $S_{30}$    | $F_{100}$   | $RO$        |
|---------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| > 15 %        | 0.0         | 0.1         | 0.0         | 0.0         | 0.1         | 0.9         | 7.6         |
| 10..15 %      | 0.0         | 0.4         | 0.0         | 0.1         | 0.2         | 1.0         | 1.2         |
| 5..10 %       | 0.5         | 1.7         | 0.4         | 0.2         | 0.5         | 2.5         | 2.0         |
| 2.5..5 %      | 1.2         | 3.9         | 1.1         | 0.7         | 0.9         | 4.6         | 1.7         |
| $-2.5..2.5\%$ | 96.3 (16.6) | 88.2 (12.4) | 95.8 (18.4) | 96.6 (22.6) | 94.0 (21.9) | 79.7 (10.0) | 76.8 (16.1) |
| $-2.5..-5\%$  | 1.4         | 4.0         | 1.2         | 1.7         | 1.1         | 5.0         | 2.7         |
| $-5..-10\%$   | 0.5         | 1.3         | 1.0         | 0.5         | 1.8         | 3.7         | 2.9         |
| $-10..-15\%$  | 0.0         | 0.4         | 0.4         | 0.2         | 1.5         | 1.7         | 1.1         |
| < $-15\%$     | 0.0         | 0.1         | 0.0         | 0.0         | 0.0         | 1.0         | 4.1         |

or more of the eight properties considered in this study would result in less units; for instance, when only a single property of these eight was considered on average twenty units would result. When the original twenty-three properties would all be used, then ninety-nine units would result. Dropping one of the eight properties would result on average in 1.4 less units, whereas adding one of the fifteen non-considered properties would result on average in 3.5 more units.

#### 4. Conclusion

We have demonstrated how soil profiles from the Dutch soil map, and the hydraulic properties (water retention, hydraulic conductivity) measured on individual soil samples of common soil layers in the Netherlands, can be combined and clustered into a soil physical units map. These soil physical units include soil profiles that have the following comparable soil hydraulic properties: hydraulic resistance, readily available water and water available under stress, critical capillary rise, critical capillary flux, the integrals of the water retention and hydraulic conductivity characteristics, and  $P$ -index. We showed that dynamically simulated transpiration reduction for the 79 soil physical units is similar to transpiration reduction obtained for the 368 individual soil profiles. At the Dutch national scale, the difference in simulated long-term average transpiration reduction between runs using all 368 soil profiles or the 79 soil physical units was less than  $\pm 2.5\%$  percentage-points in 96 % of all plots or less than  $\pm 5\%$  percentage-

points in 99 % of all plots. Similar good correspondence was obtained for other water balance terms as well, including actual transpiration, actual evaporation at the soil surface, degree of saturation at 15 and 30 cm depth, the integrated water flux at 100 cm depth and surface runoff. This means that for future simulations for the Dutch soils it suffices to use the 79 soil physical units instead of all 368 derived soil profiles, at least for studies where the focus is on transpiration reduction.

Since the soil map may change in time (diminishing of peat and peaty areas and thus increasing areas of mineral soils; ripening of young clay soils) and the number of measured soil hydraulic properties on individual soil samples may increase over time and can change in time as a result of for example compaction, it is advised to repeat this procedure every several years to keep the soil physical units map up to date. This process can be easily automated, and, of course, be applied to other regions as well.

#### Funding

The derivation of the Staring series was financed by the Dutch Ministry of Agriculture, Nature and Food Quality: additional knowledge base funds for models (KB-33-001-021). The derivation of BOFEK was financed by NHI and STOWA united in the steering group “Regionale en Landelijke Modelinstrumentaria” under supervision of the NHI program-team Working Group Unsaturated Zone.

**Author contributions**

MH: projects initiation, analyses, writing, editing; HMM: describing/performing/analyzing verification tests; discussion; GB: data supply, internal review; JHMW: data supply, internal review; FB + KT: soil map description, check membership soil profiles within clusters; DJJW: discussion clustering, internal review. All authors have read the manuscript.

**Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Data availability**

Data will be made available on request.

**Appendix A**

The integral of the water retention characteristic is given by.

$$\int \theta_r + \frac{(\theta_s - \theta_r)\theta_s}{(1 + |\alpha h|^n)^m} dh = h \left( \theta_r + (\theta_s - \theta_r) {}_2F_1 \left[ m, \frac{1}{n}; 1 + \frac{1}{n}; -|\alpha h|^n \right] \right) \tag{A1}$$

where  ${}_2F_1[a,b,c;d]$  is the hypergeometric function. This function can be accurately computed according to the method of Michel and Stoitsov (2008; source code (Fortran, C++) available at <https://data.mendeley.com/datasets/76c3pp7rzm/1>). Here the complete integral is presented. In case the integral of part of the curve is needed, e.g., between  $h = 0$  cm and  $h = -16,000$  cm, then this follows from the difference of this expression at the two pressure heads considered.

A well-known integral property of the hydraulic conductivity characteristic is the matrix flux potential  $\phi$  ( $\text{cm}^2 \text{d}^{-1}$ ) defined as (Raats, 1970).

$$\phi = \int_{h=-h_{ref}}^h K(h)dh = \int_{\theta=\theta_{ref}}^\theta D(\theta)d\theta \tag{A2}$$

The property  $\phi$  is used in studies where the Richards equation is linearized, and in some root water uptake models (De Jong van Lier et al., 2008; Heinen, 2001; De Willigen and van Noordwijk, 1987). De Jong van Lier et al. (2009), with a correction given by Heinen and Bakker (2016), presented the following expression for  $\phi$ .

$$\phi = \frac{K_s(1-m)}{\alpha(\nu-1)} \left( S^{1-\frac{1}{m}+\lambda} (f_1[S] + f_2[S] - 2) - S_{ref}^{1-\frac{1}{m}+\lambda} (f_1[S_{ref}] + f_2[S_{ref}] - 2) \right) \tag{A3}$$

with

$$f_1[x] = {}_2F_1[\nu-1, m, \nu; x] \tag{A4}$$

$$f_2[x] = {}_2F_1[\nu-1, -m, \nu; x] \tag{A5}$$

where  $\nu = m(1 + \lambda)$  and  ${}_2F_1[a,b,c;d]$  is the hypergeometric function. The calculation in obtaining  $IK$  involves the difference in  $\phi$  for the two pressure heads considered ( $h = 0$  cm and  $h = -16,000$  cm); the choice of the value for  $h_{ref}$  in that case is arbitrary and was here set at  $-10^8$  cm.

The original definition of the  $P$ -index is based on the derivative of log-log-transformed expression of the water retention characteristic.

$$P = \frac{1}{\theta_{s,i} - \theta_{r,i}} \int_{\theta_{r,i}}^{\theta_{s,i}} \frac{d(\ln(\theta))}{d(\ln(h))} d\theta \tag{A6}$$

Since  $d(\ln(\theta)) = d\theta/\theta$  and  $d(\ln(h)) = dh/h$  and  $d\theta/dh = C$ , with  $C$  the differential moisture capacity,  $P$  can be written as.

$$P = \frac{1}{(\theta_s - \theta_r)} \int_{\theta_r}^{\theta_s} \frac{d(\ln(\theta))}{d(\ln(h))} d\theta = \frac{1}{(\theta_s - \theta_r)} \int_{\theta_r}^{\theta_s} \frac{hd\theta}{\theta dh} d\theta = \frac{1}{(\theta_s - \theta_r)} \int_{\theta_r}^{\theta_s} \frac{h(\theta)C(\theta)}{\theta} d\theta \tag{A7}$$

Introducing  $S = (\theta - \theta_r)/(\theta_s - \theta_r)$  gives (NB:  $d\theta = (\theta_s - \theta_r)dS$ ).

$$P = \int_0^1 \frac{h(S)C(S)}{\theta_r + (\theta_s - \theta_r)S} dS \tag{A8}$$

With

$$h(S) = \frac{1}{\alpha} (S^{-1/m} - 1)^{1-m} \tag{A9}$$

and

$$C(S) = \frac{m}{1-m} \alpha (\theta_s - \theta_r) S^{1/m} (1 - S^{1/m})^m \tag{A10}$$

it follows that

$$P = \frac{m}{1-m} \int_0^1 \frac{(S^{-1/m} - 1)^{1-m} S^{1/m} (1 - S^{1/m})^m}{T + S} dS \quad (A11)$$

where  $T = \theta_r / (\theta_s - \theta_r)$ . With the transformation  $y = S^{1/m}$  it follows (NB:  $dS = my^{m-1}dy$ ).

$$P = \frac{m^2}{1-m} \int_0^1 \frac{(y^{-1} - 1)^{1-m} y^m (1 - y)^m}{T + y^m} dy = \frac{m^2}{1-m} \int_0^1 \frac{y^m y^{m-1} (1 - y)^m}{T + y^m} dy \quad (A12)$$

From this an expression for the  $P$  index can be given in terms of the hypergeometric function, as follows.

$$P = \begin{cases} \frac{m}{1-m^2} & \theta_r = 0 \\ \frac{m}{1-m^2} \left( (1+m)T \ln \left( 1 + \frac{1}{T} \right) - {}_2F_1 \left[ 1, 1 + \frac{1}{m}, 2 + \frac{1}{m}; -\frac{1}{T} \right] - 1 \right) & \theta_r > 0 \end{cases} \quad (A13)$$

All the above expressions of the integrals ( $IW$ ,  $IK$ ,  $P$ ) were derived/verified with the help of Mathematica (Wolfram Research, Inc., 2015). Note that the hypergeometric function can also be expressed in terms of the incomplete beta function in case the third argument of  ${}_2F_1$  equals either the first or the second argument +1, which is the case in our cases (<https://mathworld.wolfram.com/IncompleteBetaFunction.html>).

## Appendix B. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.geoderma.2022.116123>.

## References

- Bakker, G., Heinen, M., Gooren, H.P.A., de Groot, W.J.M., Peters, P.D., 2020. Hydrophysical properties of soil in the Database Registration Subsoils and in the Soil Information System, Update 2019. (In Dutch: Hydrofysische gegevens van de bodem in de Basisregistratie Ondergrond (BRO) en het Bodemkundig Informatie Systeem (BIS). Update 2019). WOt-technical report 186, Wettelijke Onderzoekstaken Natuur & Milieu, Wageningen, The Netherlands. doi: 10.18174/526509.
- Bartholomeus, R.P., Witte, J.P.M., van Bodegom, P.M., van Dam, J.C., Aerts, R., 2008. Critical soil conditions for oxygen stress to plant roots: substituting the Feddes-function by a process-based model. *J. Hydrol.* 360, 147–165. <https://doi.org/10.1016/j.jhydrol.2008.07.029>.
- Boels, D., van Gils, J.B.H.M., Veerman, G.J., Wit, K.W.E., 1978. Theory and system of automatic determination of soil moisture characteristics and unsaturated hydraulic conductivities. *Soil Sci.* 126 (4), 191–199.
- Boogaard, H.L., de Wit, A.J.W., te Roller, J.A., van Diepen, C.A., 2014. WOFOST Control Centre 2.1 and WOFOST 7.1.7. User's guide for the WOFOST Control Centre 2.1 and WOFOST 7.1.7 crop growth simulation model. Update of Technical Document 52, Alterra, WUR, Wageningen, The Netherlands. Available at: <https://www.wur.nl/en/Research-Results/Research-Institutes/Environmental-Research/Facilities-Tools/Software-models-and-databases/WOFOST/Documentation-WOFOST.htm>.
- Bouma, J., Pinto-Correia, T., Veerman, C., 2021. Assessing the role of soils when developing sustainable agricultural production systems focused on achieving the UN-SDGs and the EU Green Deal. *Soil Systems* 5, 56. <https://doi.org/10.3390/soilsystems5030056>.
- Bouma, J., Bonfante, A., Basile, A., van Tol, J., Hack-ten Broeke, M.J.D., Mulder, H.M., Heinen, M., Rossiter, D.G., Poggio, L., Hirmas, D., 2022. How can pedology and soil classification contribute towards sustainable development as a data source and information carrier? *Geoderma* 424, 115988. <https://doi.org/10.1016/j.geoderma.2022.115988>.
- Bouma, J., 2020. Contributing pedological expertise towards achieving the United Nations Sustainable Development Goals. *Geoderma* 375, 114508, doi: 10.1016/j.geoderma.2020.114508.
- Dane, J.H., Hopmans, J.W., 2002. Pressure plate extractor. In: J.H. Dane and G.C. Topp (eds.), *Methods of Soil Analysis. Part 4. Physical Methods*, section 3.3.2.4, pp. 688–690. Soil Science Society of America, Madison, WI.
- De Jong van Lier, Q., Dourado Neto, D., Metselaar, K., 2009. Modeling of transpiration reduction in van Genuchten-Mualem type soils. *Water Resour. Res.* 45 (2), W02422. <https://doi.org/10.1029/2008wr006938>.
- De Jong van Lier, Q., van Dam, J.C., Durigon, A., dos Santos, M.A., Metselaar, K., 2013. Modeling water potentials and flows in the soil-plant system comparing hydraulic resistances and transpiration reduction functions. *Vadose Zone J.* <https://doi.org/10.2136/vzj2013.02.0039>.
- De Jong van Lier, Q., van Dam, J.C., Metselaar, K., de Jong, R., Duijnsveld, W.H.M., 2008. Macroscopic root water uptake distribution using a matrix flux potential approach. *Vadose Zone J.* 7 (3), 1065–1078. <https://doi.org/10.2136/vzj2007.0083>.
- De Lange, W.J., Prinsen, G.F., Hoogewoud, J.C., Veldhuizen, A.A., Verkaik, J., Oude Essink, G.H.P., van Walsum, P.E.V., Delsman, J.R., Hunink, J.C., Massop, H.T.L., Kroon, T., 2014. An operational, multi-scale, multi-model system for consensus-based, integrated water management and policy analysis: The Netherlands Hydrological Instrument. *Environ. Modelling Software* 59, 98–108. <https://doi.org/10.1016/j.envsoft.2014.05.009>.
- De Vos, J.A., 1997. Water flow and nutrient transport in a layered silt loam soil. PhD Thesis, Wageningen Agricultural University, Wageningen, The Netherlands. Available at: <https://edepot.wur.nl/199876>.
- De Vries, F., 1999. Characterizing Dutch soils based on physical and chemical properties (In Dutch: Karakterisering van Nederlandse gronden naar fysisch-chemische kenmerken). Report 654, DLO-Staring Centre, Wageningen. Available at: <https://library.wur.nl/WebQuery/wurpubs/fulltext/298371>.
- De Willigen, P., van Noordwijk, M., 1987. Roots, plant production and nutrient use efficiency. Ph.D. Thesis. Wageningen Agricultural Univ., Wageningen, the Netherlands. Available at: <https://edepot.wur.nl/202228>.
- De Willigen, P., van Dam, J.C., Javaux, M., Heinen, M., 2012. Root water uptake as simulated by three soil water flow models. *Vadose Zone J.* <https://doi.org/10.2136/vzj2012.0018>.
- Defterdarović, J., Filipović, L., Kranjčec, F., Ondrašek, G., Kikić, D., Novosel, A., Mustač, I., Krevh, V., Magdić, I., Rubinić, V., Bogunović, I., Dugan, I., Čopeć, K., He, H., Filipović, V., 2021. Determination of soil hydraulic parameters and evaluation of water dynamics and nitrate leaching in the unsaturated layered zone: a modeling case study in Central Croatia. *Sustainability* 2021 (13), 6688. <https://doi.org/10.3390/su13126688>.
- Feddes, R.A., Kowalik, P.J., Zaradny, H., 1978. Simulation of field water use and crop yield. *Simulation Monographs*. Pudoc, Wageningen. Available at: <https://edepot.wur.nl/168026>.
- Hack-ten Broeke, M.J.D., Kroes, J.G., Bartholomeus, R.P., van Dam, J.C., de Wit, A.J.W., Supit, I., Walvoort, D.J.J., van Bakel, P.J.T., Ruijtenberg, R., 2016. Quantification of the impact of hydrology on agricultural production as a result of too dry, too wet or too saline conditions. *Soil* 2, 391–402. <https://doi.org/10.5194/soil-2-391-2016>.
- Hack-ten Broeke, M.J.D., Mulder, H.M., Bartholomeus, R.P., van Dam, J.C., Holshof, G., Hoving, I.E., Walvoort, D.J.J., Heinen, M., Kroes, J.G., van Bakel, P.J.T., Supit, I., de Wit, A.J.W., Ruijtenberg, R., 2019. Quantitative land evaluation implemented in Dutch water management. *Geoderma* 338, 536–545. <https://doi.org/10.1016/j.geoderma.2018.11.002>.
- Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-means clustering algorithm. *Appl. Stat.* 28, 100–108. <https://doi.org/10.2307/2346830>.
- Heinen, M., 2001. FUSSIM2: Brief description of the simulation model and application to fertigation scenarios. *Agronomie* 21, 285–296. <https://doi.org/10.1051/agro:2001124>.
- Heinen, M., Bakker, G., 2016. Implications and application of the Raats superclass of soils equations. *Vadose Zone J.* 15 (8) <https://doi.org/10.2136/vzj2016.02.0012>.
- Heinen, M., Brouwer, F., Teuling, K., Walvoort, D., 2021. BOFEK2020 – Soil physical schematization of The Netherlands. Update soil physical units map (In Dutch: BOFEK2020 – Bodemfysische schematisatie van Nederland. Update bodemfysische eenhedenkaart). Report 3056, Wageningen Environmental Research, Wageningen. doi.org/10.18174/541544.
- Heinen, M., Bakker, G., Wösten, H., 2020. Water retention and hydraulic conductivity characteristics of top- and sub-soils in The Netherlands: the Staring series. Update 2018 (In Dutch: Waterretentie- en doorlatendheidskarakteristieken van boven- en ondergronden in Nederland: de Staringreeks; update 2018). Report 2978, Wageningen Environmental Research, Wageningen. doi: 10.18174/512761.
- Jin, X., Zhang, L., Gu, J., Zhao, C., Tian, J., He, C., 2015. Modelling the impacts of spatial heterogeneity in soil hydraulic properties on hydrological process in the upper reach of the Heihe River in the Qilian Mountains, Northwest China. *Hydrology Processes* 29, 3318–3327. <https://doi.org/10.1002/hyp.10437>.
- Keesstra, S.D., Bouma, J., Wallinga, J., Titttonell, P., Smith, P., Cerda, A., Montanarella, L., Quinton, J., Pachepsky, Y., van der Putten, W.H., Bardgett, R.D., Moolenaar, S., Mol, G., Fresco, L.O., 2016. The significance of soils and soil science towards realization of the United Nations Sustainable Development Goals. *SOIL* 2, 111–128. <https://doi.org/10.5194/soil-2-111-2016>.
- Koorevaar, P., Menelik, G., Dirksen, C., 1983. Elements of Soil Physics. Developments in Soil Science 13, Elsevier, Amsterdam, The Netherlands.

- Kozłowski, M., Komisarrek, J., 2017. Analysis of the suitability of Polish soil texture classification for estimating soil water retention and hydraulic properties. *Soil Sci. Annual* 68 (4), 197–204. <https://doi.org/10.1515/ssa-2017-0025>.
- Kroes, J.G., Supit, I., 2011. Impact analysis of drought, water stress and salinity on grassland production in the Netherlands using historical and future climate data. *Agric. Ecosyst. Environ.* 144 (1), 370–381. <https://doi.org/10.1016/j.agee.2011.09.008>.
- Kroes, J.G., van Dam, J.C., Bartholomeus, R.P., Groenendijk, P., Heinen, M., Hendriks, R. F.A., Mulder, H.M., Supit, I., van Walsum, P.E.V., 2017. SWAP version 4. Theory description and user manual. Report 2780, Wageningen Environmental Research, Wageningen, the Netherlands. Available at: <https://edepot.wur.nl/416321>, <https://swap.wur.nl/>.
- Michel, N., Stoitsov, M., 2008. Fast computation of the Gauss hypergeometric function with all its parameters complex with application to the Pöschl-Teller-Ginocchio potential wave functions. *Comput. Phys. Commun.* 178 (7), 535–551. <https://doi.org/10.1016/j.cpc.2007.11.007>.
- Mualem, Y., 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resour. Res.* 12, 513–522. <https://doi.org/10.1029/WR012i003p00513>.
- Nemes, A., Schaap, M.G., Leij, F.J., Wösten, J.H.M., 2001. Description of the unsaturated soil hydraulic database UNSODA version 2.0. *J. Hydrol.* 251, 151–162. [https://doi.org/10.1016/S0022-1694\(01\)00465-6](https://doi.org/10.1016/S0022-1694(01)00465-6).
- Périard, Y., José Gumiere, S., Rousseau, A.N., Caillier, M., Gallichand, J., Caron, J., 2017. Assessment of the drainage capacity of cranberry fields: Problem identification using soil clustering and development of a new drainage criterion. *Can. J. Soil Sci.* 97 (1), 56–70. <https://doi.org/10.1139/cjss-2016-0018>.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. *Numerical Recipes in Fortran 77. The art of scientific computing*, second edition. Cambridge University Press.
- R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Raats, P.A.C., 1970. Steady infiltration from line sources and furrows. *Soil Sci. Society Am. Proc.* 34, 709–714. <https://doi.org/10.2136/sssaj1970.03615995003400050015x>.
- Reynolds, W.D. and D.E. Elrick. 2002. Constant head soil core (tank) method. In: J.H. Dane and G.C. Topp (eds.), *Methods of Soil Analysis. Part 4. Physical Methods*, section 3.4.2.2, pp. 804–808. Soil Science Society of America, Madison, WI.
- Schaap, M.G., Leij, F.J., van Genuchten, M.T., 2001. ROSETTA: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *J. Hydrol.* 251, 163–176. [https://doi.org/10.1016/S0022-1694\(01\)00466-8](https://doi.org/10.1016/S0022-1694(01)00466-8).
- Stolte, J., Freijer, J.I., Bouten, W., Dirksen, C., Halbertsma, J.M., Van Dam, J.C., van den Berg, J.A., Veerman, G.J., Wösten, J.H.M., 1994. Comparison of Six Methods To Determine Unsaturated Soil Hydraulic Conductivity. *Soil Sci. Soc. Am. J.* 58, 1596–1603. <https://doi.org/10.2136/sssaj1994.03615995005800060002x>.
- Thorndike, R.L., 1953. Who belongs in the family? *Psychometrika* 18 (4), 267–276. <https://doi.org/10.1007/BF02289263>.
- Tóth, B., Weynants, M., Nemes, A., Makó, A., Bilas, G., Tóth, G., 2014. New generation of hydraulic pedotransfer functions for Europe. *Eur. J. Soil Sci.* <https://doi.org/10.1111/ejss.12192>.
- Van Bakel, P.J.T., Massop, H.Th.L., Kroes, J.G., Hoogewoud, J., Pastoors, M.J.H., Kroon, T., 2008. Actualisatie hydrologie voor STONE 2.3. Aanpassing randvoorwaarden en parameters, koppeling tussen NAGROM en SWAP, en plausibiliteitstoets. WOT-rapport 57. Wettelijke Onderzoekstaken Natuur en Milieu, Wageningen, The Netherlands. Available at: <https://library.wur.nl/WebQuery/wurpubs/fulltext/20745>.
- Van Genuchten, M.T., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* 44 (3), 892–898. <https://doi.org/10.2136/sssaj1980.03615995004400050002x>.
- Van Genuchten, M.Th., Leij, F.J., Yates, S.R., 1991. The RETC Code for Quantifying the Hydraulic Functions of Unsaturated Soils. Report EPA/600/2-91/065, Robert S. Kerr Environmental Research Laboratory, Office Of Research And Development, U. S. Environmental Protection Agency, Ada, Oklahoma 74820. Available at: <https://www.pc-progress.com/Documents/programs/retc.pdf>.
- Veerman, C., Bastioli, C., Biro, B., Bouma, J., Cienciala, E., Emmett, B. et al. 2020. Caring for soil is caring for life - Ensure 75% of soils are healthy by 2030 for food, people, nature and climate, Independent expert report, Eur. Comm. Publ. Office of the Eur. Union, Luxembourg.
- Wind, G.P. 1966. Capillary conductivity data estimated by a simple method. In: *Proceedings of the Wageningen Symposium Water in the unsaturated Zone*, June 1966, International Association of Scientific Hydrology. Wageningen, The Netherlands.
- Wolfram Research, Inc., 2015. *Mathematica*, Version 10.3, Champaign, IL. <https://www.wolfram.com/mathematica/>.
- Wösten, J.H.M., Bannink, M.H., de Gruijter, J.J., Bouma, J., 1986. A procedure to identify different groups of hydraulic-conductivity and moisture-retention curves for soil horizons. *J. Hydrol.* 86, 133–145. [https://doi.org/10.1016/0022-1694\(86\)90010-7](https://doi.org/10.1016/0022-1694(86)90010-7).
- Wösten, J.H.M., Lilly, A., Nemes, A., Bas, C.L., 1999. Development and use of a database of hydraulic properties of European soils. *Geoderma* 90, 169–185. [https://doi.org/10.1016/S0016-7061\(98\)00132-3](https://doi.org/10.1016/S0016-7061(98)00132-3).
- Wösten, J.H.M., van Genuchten, M.T., 1988. Using texture and other soil properties to predict the unsaturated soil hydraulic functions. *Soil Sci. Soc. Am. J.* 52, 1762–1770. <https://doi.org/10.2136/sssaj1988.03615995005200060045x>.
- Wösten, J.H.M., Veerman, G.J., de Groot, W.J.M., Stolte, J., 2001. Water retention and hydraulic conductivity characteristics of top-soils and sub-soils in The Netherlands: the Staring series. Update 2001 (In Dutch: Waterretentieve- en doorlatendheidskarakteristieken van boven- en ondergronden in Nederland: de Staringreeks Vernieuwde uitgave 2001). Alterra report 153, Alterra, Wageningen. Available at: <https://edepot.wur.nl/43272>.