Journal of Hydrology 627 (2023) 130468

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Dominant factors determining the hydraulic conductivity of sedimentary aquitards: A random forest approach

Martijn D. van Leer [a,*], Willem Jan Zaadnoordijk [b,c], Alraune Zech [d], Jelle Buma [b], Ronald Harting [b], Marc F.P. Bierkens [a,e], Jasper Griffioen [b,f]

[a] Department of Physical Geography, Utrecht University, Princetonlaan 8a, 3584 CB Utrecht, the Netherlands
[b] TNO—Geological Survey of the Netherlands, P.O. Box 80015, 3508 TA, Utrecht, Netherlands
[c] Water Resources Section, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, Netherlands
[d] Department of Earth Sciences, Utrecht University, Princetonlaan 8a, 3584CB Utrecht, the Netherlands
[e] Unit Subsurface and Groundwater Systems, Deltares, Daltonlaan 600, 3584 BK Utrecht, the Netherlands
[f] Copernicus Institute of Sustainable Development, Utrecht University, Princetonlaan 8a 3584 CB Utrecht, the Netherlands

ARTICLE INFO

This manuscript was handled by Corrado Corradini, Editor-in-Chief, with the assistance of William Payton Gardner, Associate Editor

Keywords:
Aquitards
Hydraulic conductivity
Parameterisation
Machine learning
Groundwater
the Netherlands

ABSTRACT

Aquitards are common hydrogeological features and their hydraulic conductivity is an important property for various groundwater management issues. Predicting their hydraulic conductivity proves challenging, given its dependence on numerous variables. In this study, the dominant factors for predicting aquitard hydraulic conductivity are identified. To this end, a random forest model is trained on a dataset consisting of more than 1000 hydraulic conductivity measurements of core-scale sediment samples from a wide range of stratigraphic units and depths in the Netherlands. The dataset contains textural properties, such as the grain size distribution and porosity, as well as structural data, such as location, sampling depth, stratigraphical unit, lithofacies, organic carbon content, carbonate content and groundwater chloride concentration. Results show that clay fraction, stratigraphic unit, depth, lithofacies and x-coordinate are the most important features for predicting the hydraulic conductivity. Here, x-coordinate is presumably a proxy for distance from marine influence. Using a more detailed grain size distribution or using derived parameters such as the grain size percentiles does not improve the model any further. Our findings indicate that structural properties play a significant role in predicting aquitard conductivity, as they serve as indicators of processes such as compaction and soft-sediment deformation. The model is furthermore an effective method to estimate hydraulic conductivity for sediment samples without conducting costly and time-consuming hydraulic conductivity measurements.

1. Introduction

Aquitards play an important role in hydrogeological studies, for e.g. water resource management (Gurwin and Lubczynski, 2005), land subsidence (Zhuang et al., 2017), contaminant transport (Ponzini et al., 1989), subsurface energy storage (Sommer et al., 2015) and radioactive waste disposal (Hendry et al., 2015). Although the importance of aquitards is widely recognized (Hart et al., 2006; Keller et al., 1989), many hydrogeological field studies focus on the characterization of aquifers and ignore aquitards (Fogg and Zhang, 2016). However, in recent years there has been a growing interest in the characterization of aquitard properties (e.g. Bense et al., 2022; Ferris et al., 2020; Zhao and Illman, 2018; Zhuang et al., 2020, 2019).

We focus on aquitards which consist mostly of unconsolidated sediment with a large fraction of clay, silt and organic matter. The hydraulic conductivity of such matrices can vary by several orders of magnitude (Neuzil, 1994). This underlines the importance of having accurate estimates regarding the hydraulic conductivity. However, collecting samples and measuring the hydraulic conductivity of aquitard material is difficult, as the samples are easily disturbed during collection, transport and lab analysis (Clark, 1998). Also, the pressure and flow conditions of the sample in the field have to be reproduced in the lab to obtain representative hydraulic conductivity values (Boynton and Daniel, 1985; Dafalla et al., 2015). In addition the hydraulic conductivity measurements themselves can be time consuming for low hydraulic conductivity values (e.g. Yu et al., 2013).

* Corresponding author.
E-mail address: m.d.vanleer@uu.nl (M.D. van Leer).

https://doi.org/10.1016/j.jhydrol.2023.130468
Received 3 August 2023; Received in revised form 3 October 2023; Accepted 28 October 2023
Available online 11 November 2023

For these reasons, alternative methods to estimate hydraulic conductivity values based on material properties have been studied for a long time. Hydraulic conductivity in any medium is primarily dependent on the number, size distribution and connectivity of pores (Alaoui et al., 2011; Bittelli et al., 2015; Nielsen et al., 2018). These factors are related to the texture and structure of the material. The texture determines the distribution of the pore size, while the structure determines how pores are connected. Diagenetic cement may also influence these properties but this lies beyond the scope of this manuscript. As the texture can easily be measured in soils, parameters such as the grain size distribution and porosity have been used widely in pedotransfer functions. Initially, pedotransfer functions were physically based, such as the Carman-Kozeny equation (Carman, 1937; Kozeny, 1927) and have been extended with empirical relationships from regression (Jabro, 1992; Puckett et al., 1985). Van Looy et al. (2017) and Zhang & Schaap (2019) reviewed pedotransfer functions extensively.

Although pedotransfer functions are useful due to their simplicity in relating soil properties to conductivity, they have limitations with regards to the impact of soil and sediment structural properties on hydraulic conductivity. Sedimentary aquitards generally contain significant clay fractions, which are more prone to structural alteration than sands due to the less rigid chemical and physical properties of clay. They, therefore, have a wide range of hydraulic conductivity values for similar textures (Neuzil, 1994). In addition, it is difficult to find quantitative measures for these structural properties that can be universally applied, due to highly non-linear effects of these types of properties on the hydraulic conductivity (Díaz-Zorita et al., 2002).

To incorporate the non-linear control of material properties on hydraulic conductivity, machine learning models have been used increasingly. These have mainly been applied towards conductivity values for (unsaturated) soils, where other parameters impact flow behaviour compared to aquifers and aquitards. The first model using a simple neural network for soils was Rosetta, which applied multiple pedotransfer functions to predict hydraulic conductivity and the water retention curve for soils (Schaap et al., 2001) considering only textural properties for the prediction. In many studies since then, various algorithms have been used to predict hydraulic conductivity based on textural properties of the solid matrix, such as k-nearest neighbours (Botula et al., 2013; Nemes et al., 2006), support vector regression (Kotlar et al., 2019; Mady and Shein, 2018; Singh et al., 2021) and tree-based regression (Granata et al., 2022; Sihag et al., 2019; Szabó et al., 2019; Tilahun and Korus, 2023). Structural properties were used in the training of machine learning models in the form of land use and bulk density (Jorda et al., 2015) and showed to be the most important predictor for hydraulic conductivity for that dataset. Araya and Ghezzehei (2019) used bulk density and organic carbon content as indicators for soil structural parameters on a large dataset, which also included textural parameters, and used several algorithms to train the model. The model predicted the hydraulic conductivity better than any pedotransfer function model, but the dataset consisted mostly of sandy soils, and had a relatively low number of silt and clay samples.

We developed a machine learning model for predicting hydraulic conductivity that incorporates an extensive list of textural and structural properties with a focus on aquitard material.. Given the general importance of aquitards in shallow sedimentary basins, we fill a research gap as previous work has considered either soils (e.g. Araya and Ghezzehei, 2019a; Zhang and Schaap, 2017) or (sandy) aquifer materials (e.g. Rogiers et al., 2012). The model will be trained on data from the terrestrial part of the Netherlands using additional features compared to pedotransfer functions, as improvement suggested by Zhang and Schaap (2019).

2. Methods

2.1. General approach

We use the random forest regression model for predicting saturated hydraulic conductivity (target variable) from textural and structural sediment properties (model features). The model data stems from the TOPINTEGRAAL dataset, where we make use of a subset containing aquitard sample data. A variety of properties and information is available for each sample. We will only outline those used during modelling, focusing on particle size distributions and derived quantities thereof, such as sand/silt/clay content, diameter percentiles (such as D10) as well as stratigraphic unit and lithofacies.

2.2. Data

The TOPINTEGRAAL dataset (Vernes et al., 2020) stems from a data collection program that currently covers about 60 % of the terrestrial part of the Netherlands and provides detailed information on sedimentological, physical and geochemical properties of the upper 30–50 m of the subsurface. The drilling locations are not selected randomly, but at sites where certain stratigraphic units and lithofacies are expected. This way the diverse dataset covers all lithostratigraphic units that have been identified in the shallow subsurface of the Netherlands (TNO-GDN, 2023). Having started in 2006, the data collection is actively progressing with new data being continuously added. We use the TOPINTEGRAAL dataset version of the 1st of June 2023. The compiled dataset as used for training the model is publicly available (Harting et al., 2023).

The data set used in this study contains 1033 samples. From the complete TOPINTEGRAAL dataset, we selected samples that originate from aquitards, such as clay, loam and peat. The locations where the aquitard samples were collected are shown in Fig. 1. The properties of the samples we used during modelling are: x and y coordinates, depth, particle size distribution (PSD, fraction up to 2000 μm), porosity, organic carbon content (OC), calcium carbonate content ($CaCO_3$), estimated chloride concentration of the pore water at the sampled location, stratigraphic unit, lithofacies and the measured hydraulic conductivity.

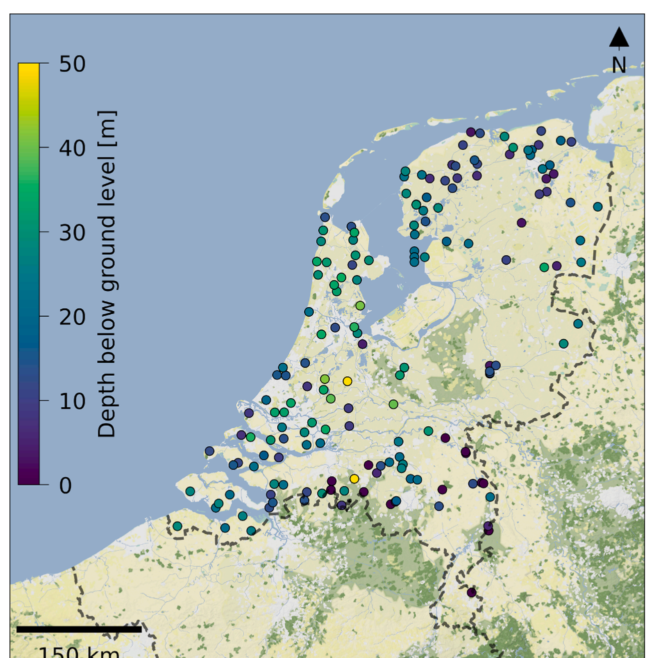

Fig. 1. Locations of samples of the TOPINTEGRAAL dataset. The colour indicates the depth of the deepest collected sample. Samples from sediment cores at different depths have been collected per sampling site.

The samples from the TOPINTEGRAAL dataset stem from the Pleistocene and Holocene deposits. In these Epochs sedimentation occurred at the border of the North Sea sedimentary basin, resulting in marine and deltaic sediments in the western part and mainly fluvial deposits in the eastern and central part of the country. On multiple occasions the fluvial environments were interrupted by glacial environments in the north. Due to this, the samples in the dataset belong to three main depositional environments: marine, fluvial and glacial. Furthermore, the stratigraphic units are defined by the different rivers where the sediment originates and geological age, following the stratigraphic nomenclator of the Netherlands (TNO-GDN, 2023). Each sample is assigned to a formation, some of which contain members. If samples belong to a member with more than 10 samples, they are assigned to the member instead of the formation for training the model. This results in 41 different stratigraphic units in the dataset. Of the 1033 samples 461 belong to fluvial deposits, 312 to marine deposits and 260 to glacial and periglacial deposits. The lithofacies as used in this study are units with specific lithological, textural and structural characteristics that distinguish them from other lithofacies units within a sedimentary environment (i.e. marine, fluvial, eolian, glacial and organic environments, and other). These units are interpreted based on their characteristics from the cores. A lithofacies unit can be present in multiple lithostratigraphic units, as opposed to being a subdivision within a unit. In the TOPINTEGRAAL programme, undisturbed cores are collected in PVC liners with a diameter of 100 or 110 mm and a length of 1 m, covering the entire depth of the drilling. The cores are sawed into two parts asymmetrically, resulting in a thicker and a thinner core segment. The thinner segment is used for detailed lithological description and stratigraphic interpretation from which a stratigraphic unit and lithofacies are assigned.

Hydraulic conductivity is measured in samples taken from the thicker part of the cores. Samples of aquitard material are tested for the vertical hydraulic conductivity only. If the sample interval in the core is visually homogeneous, vertical hydraulic conductivity is measured in an adapted falling head oedometer containing a sample with thickness of 2 cm. If the sample interval is heterogeneous, e.g. due to sedimentological layering, vertical hydraulic conductivity is measured in a constant head triaxial setup, using a sample with thickness of 10 cm. During the drilling, transport, storage and opening stages, the core material can be expected to deform to some extent due to abnormalities from the in-situ ground pressure at the depth of origin. Therefore, the sample is subjected again to the in-situ ground pressure. This pressure is estimated by overburden calculations based on the lithological characteristics of the overlying sedimentary succession, and standard values for volumetric weight for each lithology.

After testing, quality checks are performed with respect to drift of the hydraulic conductivity during the test, and the extent to which the sample is deformed when subjected to the estimated in situ pressure in the test setup. The measurements of samples that showed deformation are not taken into account in this study. A histogram of the measured hydraulic conductivity values (being the target variable in the machine learning model) of the 1033 accepted aquitard samples is shown in Fig. 2.

For textural properties, subsamples were sieved over a 2000 μm sieve. One such subsample was analysed with thermal graphimetric analysis (TGA) to obtain organic carbon content and carbonate content. The particle size distribution (PSD) was performed on a similarly sieved sample, after removal of the organic matter fraction with H_2O_2 15 % treatment as well as removal of the carbonate fraction using 0.5 % HCl treatment and suspension in 1 % $Na_4P_2O_7{\cdot}10H_2O$ solution.

The particle size distribution (PSD) of each sample is given in 32 fractions in the 0 tot 2000 μm range. The gravel fraction (greater than 2 mm) was excluded in our dataset. For the machine learning model, we used different groupings of the PSD-data:

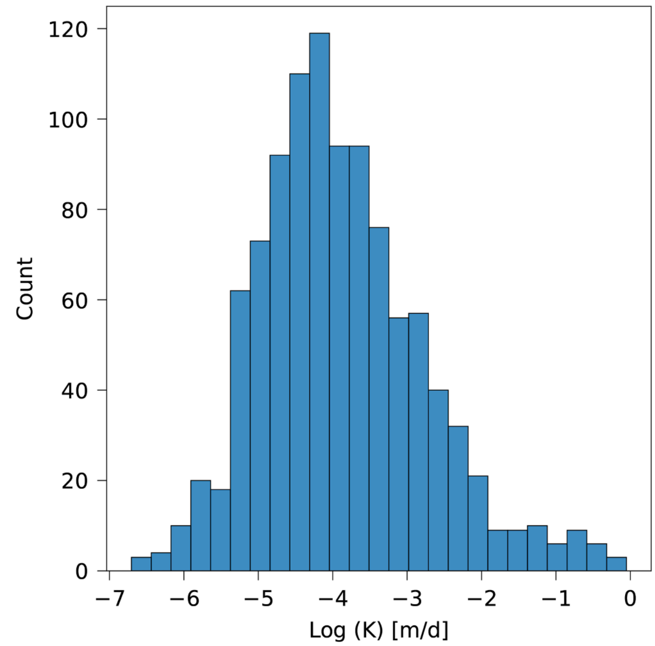

Fig. 2. Histogram of the measured hydraulic conductivity values for the entire dataset.

- 3 textural classes: clay (<8 μm), silt (8–63 μm) and sand (63–2000 μm) fractions
- 8 fractions [in micron]: 0.1–0.5, 0.5–2, 2–25, 25–63, 63–150, 150–300, 300–600, 600 – 2000
- derived descriptive values: D10, D50, D60, D70, D90 and D10/D60 where D refers to the percentile.

One may note that the laser diffraction fraction of <8 μm (that we employed) is equivalent to the classic clay fraction <2 μm found by the pipette method (Konert and Vandenberghe, 1997). Further, 63 μm is more often used as boundary between silt and sand than 50 μm in the Netherlands (NEN 5104, 1989). The median grain size of the sand fraction was used as a feature, too. The grain size distributions only relate to the clastic fraction. The textures of the 1033 aquitard samples are shown in Fig. 3.

We use the groundwater chloride concentration to identify whether salinization/freshening of clays impact the hydraulic conductivity as fresh to saline groundwater is present at shallow depth in the Western and Northern part of the Netherlands. This property is not provided by the TOPINTEGRAAL data set. Instead, we used chloride concentrations from the nationwide LHM fresh-salt groundwater model (Delsman et al., 2023).

2.3. Machine learning model

We used the random forest regression model (Breiman, 2001) as it is well suited for categorical data, detects highly non-linear patterns and is robust to outliers. Tree based regression models have shown to perform well in various studies on hydraulic conductivity predictions (Araya and Ghezzehei, 2019; Gupta et al., 2021; Jorda et al., 2015; Tilahun and Korus, 2023). The model is developed with the sklearn library in Python (Pedregosa et al., 2011).

We tested the random forest model with three combinations of sample properties as model features, summarized in Table 1. The target variable is the logarithm of the measured hydraulic conductivity in all cases. Categorical data, such as stratigraphical units and lithofacies were transformed using target encoding. Here, the categories were replaced with the mean hydraulic conductivity values of those samples within the categorical unit at hand. The encoder is trained on the training set and

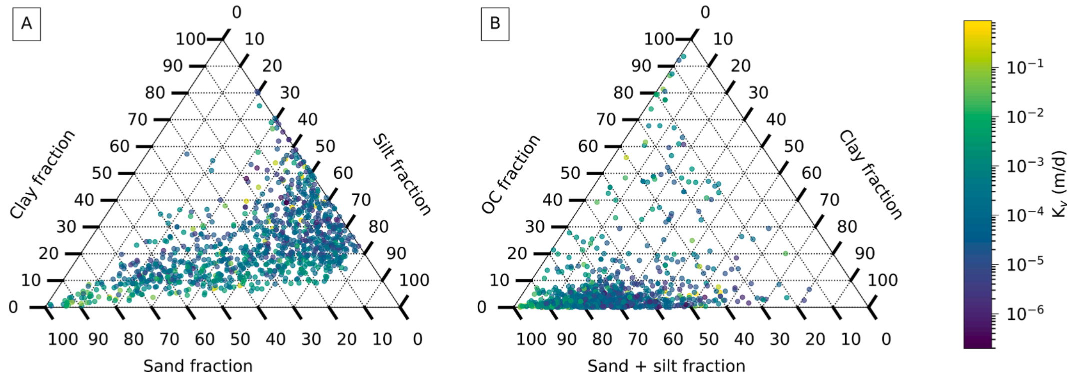

Fig. 3. Composition of the samples. A) a texture triangle based on the NEN5104 in which clay (<8 μm), silt (8–63 μm) and sand (63–2000 μm) fractions add up to 100 % and B) triangle where organic matter (OC), clay, and silt + sand fractions add up to 100 %. The colour indicates the hydraulic conductivity.

Table 1
Random forest models with different features and their performance results (R^2 – coefficient of variation and RMSE - root mean square error).

Set	Model Name	Features	$\overline{R^2}$	RMSE (log K) [m/d]
1	Main features	x, y, z, n, M_{sand}, OC, CO3, Cl, Strat, Lith, clay, silt, sand	0.551	0.734
2	Extended grain size fractions	x, y, z, n, M_{sand}, OC, CO3, Cl, Strat, Lith, 0.1–2, 2–8, 8–16, 16–35, 35–63, 63–210, 210–420, 420–2000	0.542	0.742
3	Grain size percentiles	x, y, z, n, M_{sand}, OC, CO3, Cl, Strat, Lith, D10, D50, D60, D70, D90, D60/D10	0.547	0.737

x: x-coordinate, y: y-coordinate, z: depth below ground level, clay, silt and sand are their respective fractions, n: porosity, M_{sand}: median grain size of sand fraction, OC: organic carbon content, CO3: carbonate content, Cl: chloride concentration, Strat: Stratigraphic unit, Lith: lithofacies, 0.1–2, 2–8,..: grain size fractions, D10, D50, …: grain size diameter percentiles.

applied to both the training and the test data set.

Fig. 4 shows the workflow of the model. The procedure is applied to each model feature set separately (Table 1). We used a data set split ratio of 80% and 20% for training and testing, respectively. We combined this with cross-validation to ensure that testing and training performance is representative of the entire data set and not impacted by the specific data set split. We, therefore, divided the entire data set into five parts. Four parts were used to train the model, and the fifth was used to test. This procedure was repeated 5 times (K-folds) resulting in 25 combinations of training and test sets applied to RF separately.

The sample data was not distributed completely random to the training and testing data sets, but stratified on the stratigraphic unit. So, we ensured that every split (of 20% data) contained a similar distribution of all stratigraphic units as the complete dataset. This is important to not miss out on this feature during training for stratigraphical units having a small number of samples.

We performed hyperparameter tuning on the depth of the trees and the considered features per split. These hyperparameters were optimized using (separate) cross validation for each of the 25 training sets (Fig. 4). The objective function we used for optimization as well as testing is the minimum mean squared error.

The random forest was trained for each of the 25 combinations using the optimized hyperparameters (Fig. 4). The permutation importance is calculated of each trained model for each feature, using 50 permutations. The trained model was then applied to the test data set for all 25 combinations. The model performance was evaluated using the coefficient of variation R^2 and the root mean square error (RMSE) of the modelled and measured log-conductivity values. Model performance results for a fixed set of model features is shown using the mean of the 25 combinations as well as statistical visualization in box plots.

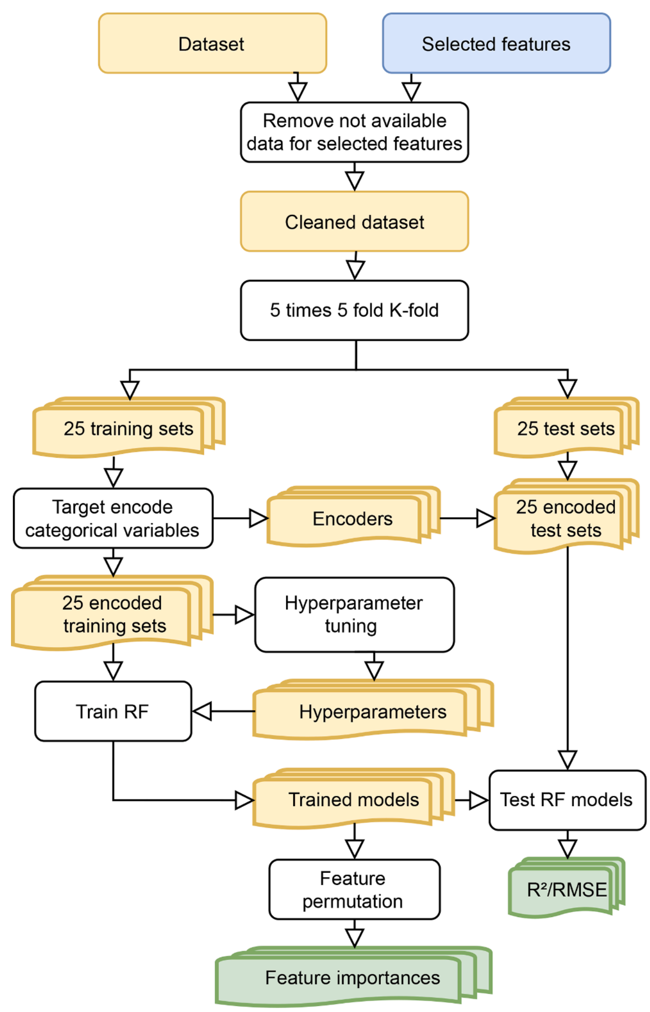

Fig. 4. Flow chart of the model workflow.

3. Results

3.1. Model performance

Table 1 shows the average performance of each of the three feature sets. The best performance was obtained by feature model #1. Only using the clay, silt and sand fractions from the PSD outperforms the models that contain more grain size descriptions. We partly link this to overfitting and increased complexity in the model which is not fully represented in the data, due to the limited amount of samples compared to the number of features. This result can also be explained by the fact that the hydraulic conductivity of the samples is largely dependent on structural parameters, which causes the grain size to be of lesser importance.

The mean predicted values versus measured values are shown in Fig. 5. The predictions are best in the range between 10^{-5} m/d and 10^{-3} m/d that also contains most measurements. They systematically deviate from the measured values towards both ends. Random forest does not extrapolate beyond the minimum and maximum values in the training set and will overestimate the lowest value and underestimate the highest value in the test set, if these are outside these limits. In addition these ranges have a limited number of samples, decreasing the prediction accuracy for these samples. This effect is also found in other hydraulic conductivity prediction models, especially for low values (Araya and Ghezzehei, 2019; Zhang and Schaap, 2019).

The residuals are larger for high hydraulic conductivity values as compared to low values. This is possibly caused by the fact that sample disturbance during collection, installation and measurement could cause preferential flow paths through cracks or along the wall. These paths carry more water than the matrix, which will likely result in overestimated hydraulic conductivity values, unrelated to the properties of the sample (Tokunaga, 1988). As all the samples have been measured once, the measurement error is not known, causing difficulty in separating model error versus measurement error.

4. Feature importance

The feature importances are shown in Fig. 6 for the three models. In all models, there is a clear distinction between the top five features i.e., stratigraphic unit, clay (or similar description of the fine grained fraction), depth, x-coordinate and lithofacies, which contribute the largest part and all other features. The importance of all features is addressed below.

4.1. Grain size

Fig. 5 shows the mean permutation importance and mean prediction results for the different models. For all models, the clay fraction is the most important textural feature. This is in line with the results of Bilardi et al. (2020) and Granata et al. (2022). In model set #2 that contains the extended grain size distribution, the 2–8 μm fraction is the most important textural feature and contains a large part of the clay fraction. Interestingly the contribution of the silt and sand fractions is minor. The model is only trained on aquitard material, causing most of the variation in sand fraction to be small compared to the silt and clay fractions. In model set #3 with the percentile description of the grain size distribution, D10 is the most important of the textural features. Araya & Ghezzehei (2019) also found that D10 was the most important feature in soils, although their study also contained sand samples without any clay fraction.

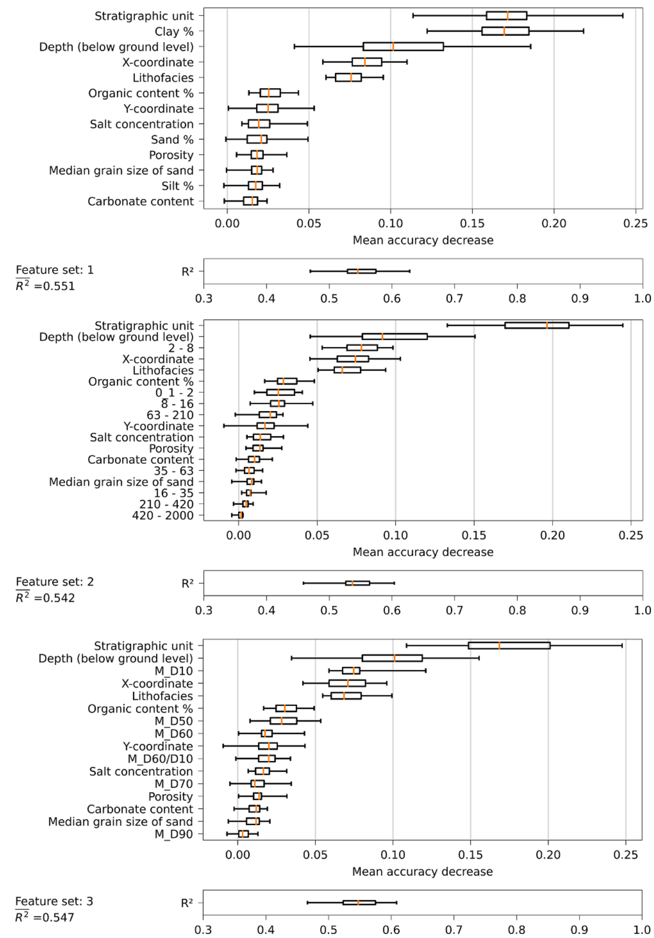

Fig. 6. Permutation feature importance and R^2 distribution for the three models with feature sets as listed in Table 1.

4.2. Geology

The stratigraphic unit is the most important general feature in all

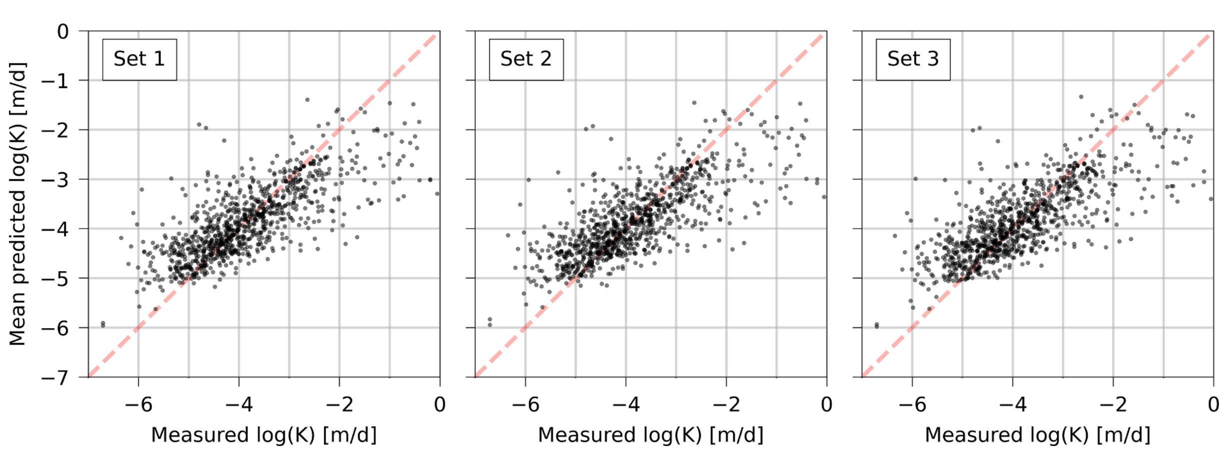

Fig. 5. Mean predicted vs measured hydraulic conductivity for the three models with feature sets as listed in Table 1.

three models. Lithofacies has a significant contribution, too. Although research on the determinants of hydraulic conductivity of aquitards has been limited, these findings are in line with a few studies on sands that incorporated e.g. depositional environment (Rosas et al., 2014) and river setting (Khaja et al., 2023) to determine hydraulic conductivity values. Stratigraphical units and lithofacies indirectly describe physical characteristics that are not covered by just the grain size distribution, such as sedimentological characteristics like layering, aggregating and sorting. Sorting is known to be an important factor to determine hydraulic conductivity as it may influence the existence and size of preferential pathways (Lind, 1999; Lopez et al., 2020). Within stratigraphic units it can also be expected that secondary processes after deposition, such as soil formation, chemical diagenesis and bioturbation are comparable.

4.3. Depth and location

The models show that accounting for depth of the samples is important. Depth dependence of hydraulic conductivity in aquitards has been demonstrated to be present in previous work (Ferris et al., 2020) and is an important consideration in groundwater resource management (Zhuang et al., 2022). Increasing compaction rates with depth cause a decrease in porosity, and reordering of clay particles, which will reduce the hydraulic conductivity. This effect is larger for samples with higher clay fractions, as sand is less sensitive to compaction due to the rigorous solid matrix (Revil et al., 2002). In addition, some stratigraphical units are present only at certain depth intervals, which cause variation in hydraulic conductivity not only caused by the depth itself, but also by other properties.

The location where the samples are taken is of minor importance in the models. Interestingly the x-coordinate contributes more to the performance than the y-coordinate. An explanation is that the sedimentary processes in the Netherlands occurred in direction of the major rivers flowing from east to west. Thus, depositional environments over time are mainly oriented from east to west, and less in north–south (y) direction. Exceptions to this are specific lithostratigraphical units which vary in the north–south direction, specifically units from glacial origin that are only found in the northern part of the Netherlands. This factor is, however, partly taken into account in the stratigraphic unit feature.

4.4. Chemical properties

Organic carbon content is of minor importance in predicting hydraulic conductivity compared to several other model features. This effect is in line with Araya & Ghezzehei, (2019) and Dexter et al., (2008) who showed that for clayey materials a higher organic carbon content correlates with higher hydraulic conductivity values. This effect however is limited, which might be caused by a skewed distribution of the organic carbon content within the samples, as the number of samples with high organic carbon content is limited (Fig. 3b).

Although no literature studies have shown that carbonate content has a significant impact on the hydraulic conductivity of unconsolidated clastic sediments, the data was included as it was available. In fact, the feature has a very limited effect. This can be partially caused by the distribution of the carbonate content in the dataset, but it is also likely that it has no effect on the hydraulic conductivity. Detrital calcium carbonate is typically found in the silt fraction of Dutch marine and fluvial sediments and shell fragments are usually larger than 200 μm (Griffioen et al., 2016). This explains why the hydraulic conductivity of poorly permeable clayey samples is invariant of carbonate content.

The chloride concentration neither is of significant importance in predicting the hydraulic conductivity for this dataset, even though freshening of salt aquifers and salinization of fresh aquifers is known to alter hydraulic conductivity (Goldenberg et al., 1983). This can be partly due to the uncertainty in the chloride estimates, which do not come from the pore water of the samples themselves but have been obtained from a hydrological model. The model may have difficulty predicting the large variation of fresh, brackish and saline groundwater that is present at shallow depth in the western and northern part of the Netherlands.

5. Discussion

This study presents a machine learning model for predicting hydraulic conductivity of aquitard material from textural and structural parameters. Considering the challenges associated with measuring hydraulic conductivity in aquitard materials, we view the performance metrics of our random forest model as evidence of a successful application of AI for the prediction of this important hydrogeological parameter. Aquitard sediments cover a large range of hydraulic conductivity values and are sensitive to structural changes and therefore the hydraulic conductivity is difficult to predict accurately. In addition, the aquitard samples are more difficult to gather from larger depths and are more prone to disturbance during collection and measurement than sandy aquifer samples or soil samples, even with adequate quality control in place. Due to these issues the performance can not directly be compared to other hydraulic conductivity prediction models designed for soils or aquifers. Despite these challenges we deem the performance sufficient to provide useful indications for the feature importance. The random forest model developed in this study can be used to estimate hydraulic conductivity values for samples where the hydraulic conductivity testing failed, e.g. due to leakage along the wall of the sample container or disturbance during installation, while the analytical and other data are available. To negate the need for (costly) hydraulic conductivity measurements altogether, it is necessary to determine the measurement error and to compare it with the model error to be certain to what extent the model can replace the measurements (McIntyre et al., 1979; Tokunaga, 1988).

Although the random forest model is only applicable for data from The Netherlands, we deem the findings about the feature importance have more general validity. The models benefit greatly from additional data, such as the depth and geology, not using only textural measurements, which seems adequate for predicting hydraulic conductivities of sandy material. This suggests that structural properties are more important for aquitard material than for aquifers, as aquitards are more sensitive to compaction and soft sediment deformation.. The results also suggest that the sedimentological and geological properties of the samples cause variation in the hydrological properties of the samples that do not show in grain size distribution, but do affect the hydraulic conductivity, such as the degree of sorting and roundness of the grains. The geological properties as encountered in this study cannot be extrapolated to other regions, but it is promising to incorporate this type of information in hydraulic conductivity prediction models and to make them more useful not only for soils, but for all sorts of sediments.

6. Conclusions

Three random forest models have been developed to predict hydraulic conductivity of core-scale samples from sedimentary aquitard layers in The Netherlands, using textural and structural properties. The best fitting model has an R^2 of 0.56 and an RMSE of the log conductivity of 0.734. The five most important features are stratigraphic unit, clay fraction, depth, x-coordinate and lithofacies. Results thus show that the addition of geological properties to the model greatly improves the model performance, apart from the grain size distribution. Training the model using a grain size distribution with eight fractions and training the model with grain size percentiles did not yield better performances than the model with only clay, silt and sand fractions. Other features such as organic carbon content, carbonate content, porosity and groundwater chloride concentration are of minor importance in predicting the hydraulic conductivity of aquitard sediments.

CRediT authorship contribution statement

Martijn D. van Leer: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Willem Jan Zaadnoordijk:** Methodology, Formal analysis, Writing – review & editing, Funding acquisition. **Alraune Zech:** Methodology, Formal analysis, Writing – review & editing. **Jelle Buma:** Methodology, Validation, Investigation, Resources, Data curation, Writing – review & editing. **Ronald Harting:** Methodology, Validation, Investigation, Resources, Data curation, Writing – review & editing. **Marc F.P. Bierkens:** Methodology, Formal analysis, Writing – review & editing, Supervision, Project administration. **Jasper Griffioen:** Conceptualization, Methodology, Formal analysis, Writing – review & editing, Supervision, Funding acquisition.

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

The data used to train the model is available from https://data.mendeley.com/datasets/r2k54rr5p2/1

References

Alaoui, A., Lipiec, J., Gerke, H.H., 2011. A review of the changes in the soil pore system due to soil deformation: a hydrodynamic perspective. Soil Tillage Res 115-116, 1–15.

Araya, S.N., Ghezzehei, T.A., 2019. Using machine learning for prediction of saturated hydraulic conductivity and its sensitivity to soil structural perturbations. Water Resour. Res. 55, 5715–5737. https://doi.org/10.1029/2018WR024357.

Bense, V.F., Kruijssen, T., van der Ploeg, M.P., Kurylyk, B.L., 2022. Inferring aquitard hydraulic conductivity using transient temperature-depth profiles impacted by ground surface warming. Water Resour. Res. 58, e2021WR030586 https://doi.org/10.1029/2021WR030586.

Bilardi, S., Ielo, D., Moraci, N., 2020. Predicting the saturated hydraulic conductivity of clayey soils and clayey or silty sands. Geosci. 10, 1–16. https://doi.org/10.3390/geosciences10100393.

Bittelli, M., Campbell, G.S., Tomei, F., 2015. Soil Physics with Python. Oxford University Press, Soil Physics with Python, 10.1093/acprof:oso/9780199683093.001.0001.

Botula, Y.-D., Nemes, A., Mafuka, P., Van Ranst, E., Cornelis, W.M., 2013. Prediction of water retention of soils from the humid tropics by the nonparametric k -nearest neighbor approach. Vadose Zo. J. 12 (vzj2012), 0123. https://doi.org/10.2136/vzj2012.0123.

Boynton, S.S., Daniel, D.E., 1985. Hydraulic conductivity tests on compacted clay. J. Geotech. Eng. 111, 465–478. https://doi.org/10.1061/(ASCE)0733-9410(1985)111:4(465).

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324/METRICS.

Carman, P.C., 1937. Fluid flow through a granular bed. Trans. Inst. Chem. Eng. London 15, 150–156.

Clark, J.I., 1998. The settlement and bearing capacity of very large foundations on strong soils: 1996 R.M. Hardy keynote address. Can. Geotech. J. 35, 131–145. https://doi.org/10.1139/t97-070.

Dafalla, M., Shaker, A.A., Elkady, T., Al-Shamrani, M., Dhowian, A., 2015. Effects of confining pressure and effective stress on hydraulic conductivity of sand-clay mixtures. Arab. J. Geosci. 8, 9993–10001. https://doi.org/10.1007/s12517-015-1925-1.

Delsman, J.R., Mulder, T., Verastegui, B.R., Bootsma, H., Huizer, S., Essink, G.H.P.O., Zitman, P., Huizer, S., Oude Essink, G.H.P., 2023. Reproducible construction of a high-resolution national variable-density groundwater salinity model for the Netherlands. Environ. Model. Softw. 164, 105683 https://doi.org/10.1016/j.envsoft.2023.105683.

Dexter, A.R., Richard, G., Arrouays, D., Czyz, E.A., Jolivet, C., Duval, O., 2008. Complexed organic matter controls soil physical properties. Geoderma 144, 620–627. https://doi.org/10.1016/j.geoderma.2008.01.022.

Díaz-Zorita, M., Perfect, E., Grove, J.H., 2002. Disruptive methods for assessing soil structure. Soil Tillage Res. https://doi.org/10.1016/S0167-1987(01)00254-9.

Ferris, D.M., Potter, G., Ferguson, G., 2020. Characterization of the hydraulic conductivity of glacial till aquitards. Hydrogeol. J. 28 (5), 1827–1839.

Fogg, G.E., Zhang, Y., 2016. Debates—Stochastic subsurface hydrology from theory to practice: a geologic perspective. Water Resour. Res. 52, 9235–9245. https://doi.org/10.1002/2016WR019699.

Goldenberg, L.C., Magaritz, M., Mandel, S., 1983. Experimental investigation on irreversible changes of hydraulic conductivity on the seawater-freshwater interface in coastal aquifers. Water Resour. Res. 19, 77–85. https://doi.org/10.1029/WR019i001p00077.

Granata, F., Di Nunno, F., Modoni, G., 2022. Hybrid machine learning models for soil saturated conductivity prediction. Water (Switzerland) 14, 1729. https://doi.org/10.3390/w14111729.

Griffioen, J., Klaver, G., Westerhoff, W.E., 2016. The mineralogy of suspended matter, fresh and cenozoic sediments in the fluvio-deltaic rhine-meuse-scheldt-ems area, the netherlands: an overview and review. Netherlands Journal of Geosciences 95 (1), 23–107.

Gupta, S., Lehmann, P., Bonetti, S., Papritz, A., Or, D., 2021. Global prediction of soil saturated hydraulic conductivity using random forest in a covariate-based geotransfer function (CoGTF). Framework. J. Adv. Model. Earth Syst. 13, e2020MS002242 https://doi.org/10.1029/2020MS002242.

Gurwin, J., Lubczynski, M., 2005. Modeling of complex multi-aquifer systems for groundwater resources evaluation - Swidnica study case (Poland). Hydrogeol. J. 13, 627–639. https://doi.org/10.1007/S10040-004-0382-9/FIGURES/9.

Hart, D.J., Bradbury, K.R., Feinstein, D.T., 2006. The vertical hydraulic conductivity of an aquitard at two spatial scales. Ground Water 44 (2), 201–211.

Harting, R., Buma, J., van Leer, M., 2023. Hydraulic conductivity measurements of aquitard sediments in the Netherlands. Mendeley Data V1. https://doi.org/10.17632/R2K54RR5P2.1.

Hendry, M.J., Solomon, D.K., Person, M., Wassenaar, L.I., Gardner, W.P., Clark, I.D., Mayer, K.U., Kunimaru, T., Nakata, K., Hasegawa, T., 2015. Can argillaceous formations isolate nuclear waste? Insights from isotopic, noble gas, and geochemical profiles. Geofluids 15 (3), 381–386.

Jabro, J.D., 1992. Estimation of saturated hydraulic conductivity of soils from particle size distribution and bulk density data. Trans. Am. Soc. Agric. Eng. 35, 557–560. https://doi.org/10.13031/2013.28633.

Jorda, H., Bechtold, M., Jarvis, N., Koestel, J., 2015. Using boosted regression trees to explore key factors controlling saturated and near-saturated hydraulic conductivity. Eur. J. Soil Sci. 66, 744–756. https://doi.org/10.1111/ejss.12249.

Keller, C.K., Van Der Kamp, G., Cherry, J.A., 1989. A multiscale study of the permeability of a thick clayey till. Water Resour. Res. 25, 2299–2317. https://doi.org/10.1029/WR025I011P02299.

Khaja, M.A., Shah, S.R., Jha, R., 2023. Hydraulic conductivity estimation of sandy soils: a novel approach. ISH J. Hydraul. Eng. 1–13 https://doi.org/10.1080/09715010.2023.2187712.

Konert, MARTIN, Vandenberghe, JEF, 1997. Comparison of laser grain size analysis with pipette and sieve analysis: A solution for the underestimation of the clay fraction. Sedimentology 44 (3), 523–535.

Kotlar, A.M., Iversen, B.V., Jong van Lier, Q., 2019. Evaluation of parametric and nonparametric machine-learning techniques for prediction of saturated and near-saturated hydraulic conductivity. Vadose Zo. J. 18, 1–13. https://doi.org/10.2136/vzj2018.07.0141.

Kozeny, J., 1927. Uber kapillare Leitung des Wassers im Boden-Aufstieg, Versickerung und Anwendung auf die Bewasserung, Sitzungsberichte der Akademie der Wissenschaften Wien. Math. Naturwissenschaftliche Abteilung 136, 271–306.

Lind, B.B., 1999. The influence of grain-size and sediment sorting on the hydraulic conductivity of some Swedish tills. GFF 121, 107–111. https://doi.org/10.1080/11035899901212107.

Lopez, O.M., Hegy, M.C., Missimer, T.M., 2020. Statistical comparisons of grain size characteristics, hydraulic conductivity, and porosity of barchan desert dunes to coastal dunes. Aeolian Res. 43, 100576 https://doi.org/10.1016/j.aeolia.2020.100576.

Mady, A.Y., Shein, E.V., 2018. Support vector machine and nonlinear regression methods for estimating saturated hydraulic conductivity. Moscow Univ. Soil Sci. Bull. 73 (3), 129–133.

McIntyre, D.S., Cunningham, R.B., Vatanakul, S., G.a.,, 1979. Measuring hydraulic conductivity in clay soils: methods, techniques, and errors. Soil Sci. 128, 171–183. https://doi.org/10.1097/00010694-197909000-00006.

Nemes, A., Rawls, W.J., Pachepsky, Y.A., 2006. Use of the nonparametric nearest neighbor approach to estimate soil hydraulic properties. Soil Sci. Soc. Am. J. 70, 327–336. https://doi.org/10.2136/sssaj2005.0128.

NEN 5104, 1989.

Neuzil, C.E., 1994. How permeable are clays and shales? Water Resour Res. 30, 145–150. https://doi.org/10.1029/93WR02930.

Nielsen, J.E., Karup, D., de Jonge, L.W., Ahm, M., Bentzen, T.R., Rasmussen, M.R., Moldrup, P., 2018. Can the volume ratio of coarse to fine particles explain the hydraulic properties of sandy soil? Soil Sci. Soc. Am. J. 82, 1093–1100. https://doi.org/10.2136/sssaj2018.02.0083.

Pedregosa, F., Weiss, R., Brucher, M., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: machine learning in python. J. Mach. Learn. Res. 12, 2825–2830.

Ponzini, G., Crosta, G., Giudici, M., 1989. The hydrogeological role of an aquitard in preventing drinkable water well contamination: a case study. Environ. Health Perspect. 83, 77–95.

Puckett, W.E., Dane, J.H., Hajek, B.F., 1985. Physical and mineralogical data to determine soil hydraulic properties. Soil Sci. Soc. Am. J. 49, 831–836. https://doi.org/10.2136/sssaj1985.03615995004900040008x.

Revil, A., Grauls, D., Brévart, O., 2002. Mechanical compaction of sand/clay mixtures. ECV 11–1-ECV 11–15 J. Geophys. Res. Solid Earth 107. https://doi.org/10.1029/2001jb000318.

Rogiers, B., Mallants, D., Batelaan, O., Gedeon, M., Huysmans, M., Dassargues, A., 2012. Estimation of hydraulic conductivity and its uncertainty from grain-size data using

GLUE and artificial neural networks. Math. Geosci. 44, 739–763. https://doi.org/10.1007/s11004-012-9409-2.

Rosas, J., Lopez, O., Missimer, T.M., Coulibaly, K.M., Dehwah, A.H.A., Sesler, K., Lujan, L.R., Mantilla, D., 2014. Determination of hydraulic conductivity from grain-size distribution for different depositional environments. Groundwater 52, 399–413. https://doi.org/10.1111/gwat.12078.

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.

Sihag, P., Mohsenzadeh Karimi, S., Angelaki, A., 2019. Random forest, M5P and regression analysis to estimate the field unsaturated hydraulic conductivity. Appl. Water Sci. 9, 1–9. https://doi.org/10.1007/s13201-019-1007-8.

Singh, B., Sihag, P., Pandhiani, S.M., Debnath, S., Gautam, S., 2021. Estimation of permeability of soil using easy measured soil parameters: assessing the artificial intelligence-based models. ISH J. Hydraul. Eng. 27, 38–48. https://doi.org/10.1080/09715010.2019.1574615.

Sommer, W., Valstar, J., Leusbrock, I., Grotenhuis, T., Rijnaarts, H., 2015. Optimization and spatial pattern of large-scale aquifer thermal energy storage. Appl. Energy 137, 322–337. https://doi.org/10.1016/j.apenergy.2014.10.019.

Szabó, B., Szatmári, G., Takács, K., Laborczi, A., Makó, A., Rajkai, K., Pásztor, L., 2019. Mapping soil hydraulic properties using random-forest-based pedotransfer functions and geostatistics. Hydrol. Earth Syst. Sci. 23, 2615–2635. https://doi.org/10.5194/hess-23-2615-2019.

Tilahun, T., Korus, J., 2023. 3D hydrostratigraphic and hydraulic conductivity modelling using supervised machine learning. Appl. Comput. Geosci. 19, 100122 https://doi.org/10.1016/J.ACAGS.2023.100122.

Tno-gdn, Stratigraphic Nomenclature of the Netherlands [WWW Document] https://www.dinoloket.nl/en/stratigraphic-nomenclature 2023 accessed 6.1.23.

Tokunaga, T.K., 1988. Laboratory permeability errors from annular wall flow. Soil Sci. Soc. Am. J. 52, 24–27. https://doi.org/10.2136/sssaj1988.03615995005200010004x.

Van Looy, K., Bouma, J., Herbst, M., Koestel, J., Minasny, B., Mishra, U., Montzka, C., Nemes, A., Pachepsky, Y.A., Padarian, J., Schaap, M.G., Tóth, B., Verhoef, A., Vanderborght, J., van der Ploeg, M.J., Weihermüller, L., Zacharias, S., Zhang, Y., Vereecken, H., 2017. Pedotransfer functions in earth system science: challenges and perspectives. Rev. Geophys. 55 (4), 1199–1256.

Vernes, R.W., Bosch, A.J.H.A., Harting, R., de Heer, E., Griffioen, J., 2020. Towards a physical and chemical characterization of the shallow subsurface of the Netherlands. European Association of Geoscientists & Engineers, 10.3997/2214-4609-pdb.150.c03 p. cp-150-00020.

Yu, L., Rogiers, B., Gedeon, M., Marivoet, J., Craen, M.D., Mallants, D., 2013. A critical review of laboratory and in-situ hydraulic conductivity measurements for the boom clay in belgium. Appl. Clay Sci. 75–76, 1–12. https://doi.org/10.1016/j.clay.2013.02.018.

Zhang, Y., Schaap, M.G., 2017. Weighted recalibration of the Rosetta pedotransfer model with improved estimates of hydraulic parameter distributions and summary statistics (Rosetta3). J. Hydrol. 547, 39–53. https://doi.org/10.1016/J.JHYDROL.2017.01.004.

Zhang, Y., Schaap, M.G., 2019. Estimation of saturated hydraulic conductivity with pedotransfer functions: a review. J. Hydrol. 575, 1011–1030. https://doi.org/10.1016/j.jhydrol.2019.05.058.

Zhao, Z., Illman, W.A., 2018. Three-dimensional imaging of aquifer and aquitard heterogeneity via transient hydraulic tomography at a highly heterogeneous field site. J. Hydrol. 559, 392–410. https://doi.org/10.1016/J.JHYDROL.2018.02.024.

Zhuang, C., Zhou, Z.F., Li, Z.F., Guo, Q.N., 2017. A method for determining hydraulic parameters of an overconsolidated aquitard. Yantu Lixue/Rock Soil Mech. https://doi.org/10.16285/j.rsm.2017.01.008.

Zhuang, C., Zhou, Z., Illman, W.A., Wang, J., 2019. Geostatistical inverse modeling for the characterization of aquitard heterogeneity using long-term multi-extensometer data. J. Hydrol. 578, 124024 https://doi.org/10.1016/j.jhydrol.2019.124024.

Zhuang, C., Zhou, Z., Illman, W.A., Dou, Z., Wang, J., 2020. Parameter estimation of an overconsolidated aquitard subjected to periodic hydraulic head variations within adjacent aquifers. J. Hydrol. 583 https://doi.org/10.1016/j.jhydrol.2020.124555.

Zhuang, C., Li, Y., Zhou, Z., Illman, W.A., Dou, Z., Yang, Y., Wang, J., 2022. Effects of exponentially decaying aquitard hydraulic conductivity on well hydraulics and fractions of groundwater withdrawal in a leaky aquifer system. J. Hydrol. 607, 127439 https://doi.org/10.1016/j.jhydrol.2022.127439.