

Multi-scale aquifer characterization and groundwater flow model parameterization using direct push technologies

Bart Rogiers¹, Thomas Vienken², Matej Gedeon¹, Okke Batelaan^{3,4,5}, Dirk Mallants⁶, Marijke Huysmans^{3,4} & Alain Dassargues⁷

¹ Institute for Environment, Health and Safety, Belgian Nuclear Research Centre (SCK•CEN), Boeretang 200, BE-2400 Mol, Belgium.

² Dept. Monitoring and Exploration Technologies, UFZ - Helmholtz Centre for Environmental Research, Permoserstr. 15, 04318 Leipzig, Germany.

³ Dept. of Earth and Environmental Sciences, KU Leuven, Celestijnenlaan 200e - bus 2410, BE-3001 Heverlee, Belgium.

⁴ Dept. of Hydrology and Hydraulic Engineering, Vrije Universiteit Brussel, Pleinlaan 2, BE-1050 Brussels, Belgium.

⁵ School of the Environment, Flinders University, GPO Box 2100, Adelaide SA 5001, Australia.

⁶ Groundwater Hydrology Program, CSIRO Land and Water, Waite Road - Gate 4, Glen Osmond SA 5064, Australia.

⁷ Hydrogeology and Environmental Geology, Dept. of Architecture, Geology, Environment and Civil Engineering (ArGenCo) and Aquapole, Université de Liège, B.52/3 Sart-Tilman, BE-4000 Liège, Belgium.

Corresponding author: Bart Rogiers; brogiers@sckcen.be; +32 14 33 31 23

Abstract

Direct push (DP) technologies are typically used for cost-effective geotechnical characterization of unconsolidated soils and sediments. In more recent developments, DP technologies have been used for efficient hydraulic conductivity (K) characterization along vertical profiles with sampling resolutions of up to a few centimetres. Until date, however, only a limited number of studies document high-resolution *in situ* DP data for three-dimensional conceptual hydrogeological model development and groundwater flow model parameterization. This study demonstrates how DP technologies improve building of a conceptual hydrogeological model. We further evaluate the degree to which the DP-derived hydrogeological parameter K , measured across different spatial scales, improves performance of a regional groundwater flow model. The study area covers an area of ~ 60 km² with two overlying, mainly unconsolidated sand, aquifers separated by a 5-7 m thick highly heterogeneous clay layer (in north-eastern Belgium). The hydrostratigraphy was obtained from an analysis of cored boreholes and about 265 cone penetration tests (CPTs). The hydrogeological parameter K was derived from a combined analysis of core and CPT data and also from hydraulic direct push tests. A total of 50 three-dimensional realizations of K were generated using a non-stationary multivariate geostatistical approach. To preserve the measured K values in the stochastic realizations, the groundwater model K realizations were conditioned on the borehole and direct push data. Optimization was performed to select the best performing model parameterization out of the 50 realizations. This model outperformed a previously developed reference model with homogeneous K fields for all hydrogeological layers. Comparison of particle tracking simulations, based either on the optimal heterogeneous or reference homogeneous groundwater model flow fields, demonstrate the impact DP-derived subsurface heterogeneity in K can have on groundwater flow and solute transport. We demonstrated that DP technologies, especially when calibrated with site-specific data, provide high-resolution 3D subsurface data for building more reliable conceptual models and increasing groundwater flow model performance.

Keywords

Direct push technologies, cone penetration testing, injection logging, hydraulic profiling, heterogeneity, conditional simulation, non-stationary geostatistics

1. Introduction

Groundwater flow and solute transport models are commonly used to support decision making regarding waste disposal options (Patyn *et al.* 1989; Gedeon *et al.* 2011; Gedeon and Mallants 2012; Selroos and Painter 2012), sites contaminated by surface or subsurface sources (*e.g.* Hammond and Lichtner 2010; Huysmans *et al.* 2006, Battle-Aguilar *et al.* 2009), or to develop and test cost-effective groundwater remediation or monitoring schemes (*e.g.* Saito and Goovaerts 2003; Bierkens 2005). Such models are influenced by different sources of uncertainty (Nilsson *et al.* 2007; Christian *et al.* 2006, 2007; Rojas *et al.* 2008), including those due to spatial variability in aquifer and aquitard properties as hydraulic conductivity (K ; a list of symbols and abbreviations is provided in Table 1). Quantifying spatial variability of K remains challenging. Classical drilling techniques for shallow heterogeneous unconsolidated sedimentary deposits involving continuous coring are expensive and time-consuming. This is especially true when the area of interest exceeds several tens of km².

Alternative techniques such as direct push technologies use hydraulic rams, supplemented with vehicle weight, or high-frequency hammering, to advance small-diameter measuring devices into the subsurface (*e.g.* Dietrich and Leven 2006; McCall *et al.* 2006). These devices are typically used for cost-effective geotechnical characterization of unconsolidated deposits using cone penetration tests (CPTs; Lunne *et al.* 1997), with typically a vertical measurement spacing of ~ 2 cm. In hydrogeological studies, such cone penetration testing data is mainly used for indirect lithostratigraphic profiling, but some studies report the use of those data to obtain K

estimates. The latter is mainly achieved by using soil behavior types (SBTs; *e.g.* Robertson 2010; Rogiers *et al.* 2010) or indirectly through other variables like fines content or grain size distributions (*e.g.* Flach *et al.* 2005; Tillmann *et al.* 2008), yielding a set of discrete K estimates. Predicting continuous K estimates has been done for piezocone sounding (CPTu; *e.g.* Elsworth and Lee 2007; Lee *et al.* 2008; Chai *et al.* 2011) and the corresponding dissipation testing (*e.g.* Parez and Fauriel 1988; Teh and Houlsby 1991; Van Baars and Van De Graaf 2007; Robertson 2010; Chai *et al.* 2012), but is seldom performed for standard CPT data. A comparison between standard CPT parameters and K was made by Vienken (2010).

Other typical uses of direct push technologies are soil or groundwater sampling (*e.g.* Johnson and Carpenter 2013; Rein *et al.* 2011). More recent but less established developments like direct push injection logging (DPIL; Butler and Dietrich 2004; Dietrich *et al.* 2008; Liu *et al.* 2009; Lessof *et al.* 2010; Liu *et al.* 2011) and the hydraulic profiling tool (HPT; GEOPROBE 2007; McCall *et al.* 2009; Köber *et al.* 2009) also allow for high-resolution hydraulic characterization with direct push equipment, by which relative K values are obtained. The more standard direct push slug testing (DPST; Butler 2002; Butler *et al.* 2002) or the direct push permeameter (Butler *et al.* 2007) are often used for calibration of the obtained high-resolution relative K data (*e.g.* Dietrich *et al.* 2008; Liu *et al.* 2009; Lessof *et al.* 2010).

The sampling depth of these direct push tools depends on the equipment used (*i.e.* applied load) and sediment properties (friction). In sandy unconsolidated sediments, depths of ~40 m can be reached. Up to now however, most of these methods are applied at small test sites. In only a limited number of studies this type of data has been applied on a regional scale, *e.g.* to parameterize regional groundwater flow models. To the best of our knowledge, the study by Flach *et al.* (2005) represents the largest area (8 km²) where direct push data has been applied for groundwater flow model parameterization. These authors successfully used CPTu data for deriving four discrete K classes based on fines content, estimated from the CPTu data, and normalized pore pressure. A threshold for the latter parameter was used to define an extra low K class. Because classic CPT tests do not have provisions for measuring pore pressure data, the parameterization of an aquitard using this type of data is particularly challenging (as high pore pressures may be induced by the cone penetration under undrained conditions, and cannot be corrected for). Moreover, as uptake of the hydraulic direct push tools by the environmental industry only slowly increases, demonstration of their usefulness, effectiveness, and complementarity with geotechnical and borehole data is much needed to improve, in a cost-effective manner, data availability in data poor areas.

The objectives of this study are to parameterize at an unprecedented spatial scale, a regional groundwater flow model using data from various types of direct push technologies, and determine the improvement in model performance by explicitly accounting for spatial variability in K . To achieve this, the usefulness of CPTs for continuous K estimation along vertical profiles is explored, and a geostatistical methodology is used to integrate the secondary CPT data with the primary calibrated hydraulic direct push and borehole data. The approach is applied to part of the Neogene aquifer and covers an area of ~60 km² in north-eastern Belgium. Data were collected in the framework of the ONDRAF/NIRAS (2011) cAt project for the disposal of low- and intermediate-level short-lived (category A) nuclear waste.

This work is a follow-up of the study by Gedeon and Mallants (2012) and Rogiers *et al.* (2012b). In the first study, using a groundwater flow model very similar to this study, it was demonstrated that head predictions are most sensitive to uncertainty in the following parameters (from highest to lowest): recharge, K values of an upper aquifer, K values and their anisotropy of a 5-7 m thick clay aquitard, and K values of a lower aquifer. Rogiers *et al.* (2012b) subsequently addressed the uncertainty about K values for the aquitard. Their study demonstrated that groundwater flow model calibration significantly improved by optimizing the 5-7 thick clay aquitard heterogeneous K field using CPT data. This resulted in a decrease of the sum of squared errors to 35% of the reference value obtained with a spatially homogeneous K for the aquitard and aquifer layers. In the current study the uncertainty about K values for the two aquifers is addressed, in an attempt to obtain more reliable groundwater head and flow predictions. The approach of using CPT data for K determination is now extended from a relatively short section associated with the clay aquitard to the entire upper aquifer and the top of the lower aquifer. Approximately 40 m of the uppermost litho-stratigraphical layers will be characterized with absolute CPT-based K estimates, additional borehole core data (Rogiers *et al.* 2012c, 2013b) and hydraulic direct push data (Vienken *et al.* 2012; Tinter 2012).

2. Methods

In this study different sets of independent data are combined to inform conceptual model building and to constrain a numerical groundwater flow model. An overview of all different datasets used in this work is provided in Table 2.

At the scale of the entire model domain (~60 km²), 265 standard CPT logs were gathered. The advantage of this type of data is that it is not expensive nor time-consuming compared to conventional drilling if the goal is to obtain logs that can be related to lithology. The disadvantage is that it can be difficult to relate the obtained geotechnical properties to K without any calibration data, and this is often associated with a high degree of

uncertainty. The CPT data were used for derivation of continuous K depth profiles. Detailed laboratory K and air permeability measurements on seven cored boreholes served as calibration data (Rogiers et al. 2013b). A smaller number of hydraulic direct push tests were performed on the future disposal site (~ 0.02 km²). The hydraulic direct push tests provided both relative (Direct Push Injection Logging - DPIL, Hydraulic Profiling Tool - HPT) and absolute (Direct Push Slug Tests - DPST) K information. The absolute measurements allow calibration of the relative data, thus producing additional high-resolution vertical K profiles. These tests, however, are slightly more time-consuming and laborious to perform than the CPT tests.

We first discuss the main hydrogeological layers that make up the groundwater flow model. Then we describe the three hydraulic direct push testing methodologies (DPIL, HPT and DPST) and their interrelationships. Next, we demonstrate the conversion of the geotechnical direct push tests into continuous K profiles using borehole core K measurements for calibration. Finally, we describe the development of the groundwater flow model and the conditioning of the shallow aquifer parameters on the borehole and direct push data. For the deeper parts of the aquifer, we use K estimates previously derived from grain-size data (Rogiers et al. 2012a).

2.1. Site hydrostratigraphy

The Mio-Pliocene geological formations characterized by the different direct push campaigns are the Mol, Kasterlee and Diest Formations. These are covered by Quaternary sediments of a few meter thickness. The main lithologies of these formations include fine- to medium-grained, glauconitic, micaceous sands with iron sandstone and lignite layers. A varying clay content is found in certain units (up to 50% in the 5-7 m thick Kasterlee Clay aquitard), while basal gravels are present between the units (Laga et al. 2001). Deposition took place in a shallow marine to perimarine environment at the southern margin of the North Sea Basin (Louwye et al. 2007; Louwye and Laga 2008; Louwye and De Schepper 2010). A lithostratigraphical profile through the Neogene aquifer in the study area is shown in Figure 2. Hydrogeologically, the Neogene aquifer is divided in an upper aquifer containing the Quaternary, the Mol Upper, Mol Lower and Kasterlee Sands, an aquitard corresponding to the Kasterlee Clay, and a lower aquifer containing the clayey top of the Diest Formation, and the Diest, Dessel, Berchem and Voort Sands (approximately 84 m thick at the location of the future disposal site; Beerten *et al.* 2010). The lower boundary of the Neogene aquifer is defined by the top of the Rupelian Boom Clay aquitard (Yu *et al.* 2013).

2.2. Hydraulic direct push tests

The hydraulic direct push tests were targeting the uppermost lithostratigraphical units of the upper aquifer (Figure 1C, D). High spatial resolution K profiles are needed at this site to develop reliable models of groundwater flow and solute transport as part of the performance assessment of the future disposal facility (ONDRAF/NIRAS, 2010). In total 17 Direct Push Injection Loggings (DPIL), 6 Hydraulic Profiling Tool (HPT) loggings and 19 Direct Push Slug Tests (DPST) were performed along two quasi orthogonal, 180- and 40 m-long transects (Figure 1D). The vertical spacing of DPIL tests was about 0.3 m, whereas the quasi-continuous HPT logs resulted in a data point about each 0.015 m (see Table 2). The 19 slug tests were performed at 6 different locations, with vertical distances between tests from 1 to 7 m. Application of the DPST involves application of compressed air to generate a near instantaneous change in head, followed by a monitoring of the recovery of the groundwater level. The DPSTs were performed as rising-head tests in temporary DP monitoring wells with a screened interval of 1.06 m length. The DPST-based K values were derived using the method of Springer and Gelhar (1991) with the correction of Butler (2002) for small-diameter wells (Tinter 2012). Examples of the data obtained at five locations up to depths of approximately 15 m are displayed in Figure 3. Data beyond the lower limit of the tool ranges were discarded. The relative K values (K_{rel}) obtained by DPIL and HPT were subsequently transformed into absolute K values using the DPST (K_{DPST}) data for calibration (see section 2.3).

2.3. Geotechnical direct push tests

CPT data were gathered in two campaigns: in 2008 a total of 180 CPT tests were carried out in a study area of ~ 60 km² based on a grid with spacing of ~ 500 -800 m (see Figure 1B). In 2010 a smaller area of $\sim 200 \times 400$ m² was involved in collecting 85 CPT profiles (see Figure 1C). The cone area for all CPT tests was 1500 mm², and the vertical data resolution is 0.02 m (see Table 2).

In total seven continuously cored boreholes reaching depths between 40 and 50 m were also retrieved to calibrate co-located CPT data. The boreholes are named Dessel-2, Dessel-3, Dessel-4, Retie-1, Retie-2, Geel-1 and Kasterlee-1. CPT depth profiles ranged between 15 and 42 m, with 60% of the soundings over 30 m deep. Most of the 265 CPTs were standard CPT tests, without registration of pore pressure. The influence of correcting for pore pressure using a standard approach (Lunne *et al.* 1997) was tested for about 40 CPTu logs, with a net cone area ratio of 0.58, but was proven to be generally negligible (difference of < 0.1 MPa for 50% of the data and < 1 MPa for 90% of the data). For clay-dominated sediments, such as the Kasterlee Clay aquitard, pore pressure correction might be important. However, given the limited number of CPTu tests, such a correction was not performed.

The 5-7 m thick Kasterlee Clay aquitard consists of a heterogeneous alternation of medium to dense fine-grained silty or clayey sand layers and clay lenses. The aquitard was proven to be continuous within the study area, and thus separates the Neogene aquifer in an upper and lower aquifer unit (Beerten *et al.* 2010). The heterogeneous K field of this aquitard was parameterized by Rogiers *et al.* (2012b), using a CPT-based heterogeneous 2D parameter field (X-Y plane) involving optimization of the mean K and its variance. In this approach the vertical heterogeneity in K is represented by a single upscaled K value (harmonic mean of CPT-based K estimated every 0.02 m for K_v). The rationale for optimizing the mean K and its variance was based on the inability to capture the discrete thin clay lenses within the aquitard by means of CPT data. As a result, the very low K values typical for the clay lenses were not well represented, thus causing a bias (*i.e.* higher values than real) in the upscaled aquitard K estimates. The inability for the cone penetration device to detect thin clay lenses of a few up to ten centimetres is the consequence of a too large support volume of an individual CPT measurement compared to the limited thickness of the clay lenses. However, even though thin clay lenses may not be fully detectable, CPT measurements can be influenced by sediments up to several cone diameters distance from the tip of the CPT device (cone diameter of 4.4 cm was used here).

To overcome this limitation in achieving a high spatial granularity of the CPT measurement, Bhattacharya and Solomatine (2005) proposed to use a shape characteristic of the CPT log. This so-called boundary energy is used for translating the detailed CPT log characteristics into a soil classification using supervised (by expert interpretations) machine learning. Boundary energy is defined as the amount of energy required to modify the shape of a contour (*i.e.* the log of a CPT parameter) to a circle with the same perimeter as the original object. This circle is the shape which has the lowest energy level. While Bhattacharya and Solomatine (2005) used a multi-scale representation of boundary energy, we used a simplified finite difference approximation of the curvature and calculated the friction ratio boundary energy (BE_{Fr}) as

$$BE_{Fr} = \frac{1}{N} \sum_{n=1}^N \left(\frac{\left| \frac{d^2 F_{r,n}}{dz^2} \right|}{\left(1 + \left(\frac{dF_{r,n}}{dz} \right)^2 \right)^{3/2}} \right)^2 \quad [1]$$

with N points within a certain range z of the point where BE_{Fr} is calculated, and $F_{r,n}$ the friction ratio at point n . For the maximum z range we used 0.10 m, which corresponds to averaging the squared curvature over 11 points with a CPT data spacing of 0.02 m. Based on several tests, this range seemed most adequate for getting a robust signal, and still being able to detect thin stratigraphic features. The resulting BE_{Fr} values are displayed in Figure 4C for the Dessel-2 borehole, which clearly shows (at the 1 m depth or below -5 m) how such sudden but small excursions of the standard CPT logs (Figure 4B) are captured by the BE_{Fr} signal.

Rogiers *et al.* (2012b) tested the usefulness of this BE_{Fr} parameter to capture the Kasterlee Clay aquitard clay lenses by using a cutoff value for the BE_{Fr} signal to identify the two lithological classes sand and clay. After assigning representative K values to each of the two classes followed by vertical upscaling, the resulting groundwater flow model only minimally improved. The relative sum of squared errors decreased by 4% compared to a model with uniform aquitard K .

Rogiers *et al.* (2013b) further derived high-resolution K logs (*e.g.* Figure 4A) using air permeability measurements on the same borehole cores used in this study, in combination with laboratory constant head measurements on 100 cm³ steel ring samples from these cores. This allows a more data-driven approach to investigate the use of the boundary energy for capturing the effect of the clay lenses more carefully. Visual comparison of the high-resolution borehole log with the BE_{Fr} parameter in Figure 4 already indicates this parameter is a good indicator for thin clay lenses.

2.4. High resolution data calibration

Calibration of the geotechnical direct push and the hydraulic direct push test (DPIL and HPT) data occurs in different independent ways. The hydraulic direct push data calibration makes use of the Direct Push Slug Tests. For calibrating the geotechnical CPT data, the high-resolution K logs from the borehole cores are used.

Different approaches exist to calibrate the relative conductivities K_{rel} obtained from DPIL or HPT with the absolute conductivities (K_{DPST}) measured by Direct Push Slug Tests. Assuming the measurement support discrepancies between the absolute (*e.g.* due to the DPST screen length) and the relative K measurements can be neglected, the K_{rel} values are rescaled using a power law with standard regression techniques. This results in the following equation, as applied by Dietrich *et al.* (2008):

$$K_{DPST} = a \times K_{rel}^b \quad [2]$$

where a and b are calibration parameters. Liu *et al.* (2009) extended this approach by using a numerical model for simulating Direct Push permeameter test responses based on the rescaled relative K data, and used inverse

modelling to obtain the optimal a and b parameters. When the DPST screen lengths become much longer than the lengths over which measurements occur for DPIL and HPT, measurement support discrepancies cannot be neglected. In such case, upscaling of the relative K data can be performed prior to calibration if the relative K data spacing is small compared to the screen length used for the absolute K measurements, as proposed by Lessof *et al.* (2010). When vertical integral scales are only a fraction of the horizontal integral scales, arithmetic and harmonic means can be used for upscaling, respectively horizontal and vertical K values. If the number of relative K measurements within the length of the screen used for measuring calibration data is small, a better calibration outcome is pursued here by using ordinary kriging and sequential gaussian simulation along the screen prior to upscaling. This is especially useful for the DPIL logs, where a typical data spacing of 0.3 m is used.

We tested the following upscaling approach for DPIL and HPT measurements: 1) applying Eq. 2 with parameters a and b , 2) inter- or extrapolation with ordinary kriging or sequential gaussian simulation for estimating K at each cm along the DPST filter length (1.06 m in our case), and 3) upscaling by taking the arithmetic mean for obtaining a horizontal K value (K_h). This upscaled K_h can be compared to the corresponding horizontal K_{DPST} values. The mismatch between calculated and measured K_{DPST} values was minimized iteratively using the simplex algorithm of Nelder and Mead (1965). The use of sequential Gaussian simulation instead of interpolation by kriging for obtaining a DPST screen support estimate was tested but the gain proved to be negligible and was not further applied (results not shown). By using the power law prior to upscaling for calibrating DPIL and HPT measurements, we enable the use of harmonic means to obtain vertical K values (K_v) as well, at any given vertical scale, under the assumption of isotropic K_h . The obtained vertical anisotropy (K_h/K_v) is then an additional source of information for conditioning the groundwater flow model.

In calibrating the CPT data, the use of standard CPT data (Figure 4B) was tested first. As can be seen from Figure 4A, the Q_{in} signal (normalized cone resistance) is negatively correlated (convex vs concave sections in the logs) with the borehole K logs for the coarse white Upper Mol Sand. A positive correlation between Q_{in} and K is present for the remainder of the lithostratigraphical column. Furthermore, a non-linear data-driven modelling approach by Rogiers *et al.* (2012a) was not able to correct for this negative correlation as the high K Upper Mol Sand only occurs at the very top of one of the boreholes, and not enough data on this unit is available. Therefore, we tested the soil behaviour type index I_c by Robertson and Wride (1998)

$$I_c = ((3.47 - \log Q_{tn})^2 + (\log F_r + 1.22)^2)^{0.5} \quad [3]$$

with Q_{in} and F_r the normalized cone resistance and friction ratio, respectively, in this case calculated without correcting for pore pressure (as this data is not available for most tests, and in general, the correction proved to be negligible). I_c is essentially the radius of concentric circles that can be plotted in the soil behavior type diagram by Robertson (1990). A power law as in Eq. 2 was then applied to the I_c data obtained with Eq. 3, while adding an additional correction using BE_{Fr} :

$$K_{CPT} = \exp\left(\ln(a \times I_c^b) - c_{SBT} \times (\ln(BE_{Fr}) - d_{SBT})\right) \quad [4]$$

with K_{CPT} the estimated K . The additional parameters c_{SBT} and d_{SBT} depend on the soil behaviour type (SBT) obtained through site-specific SBT classification of the CPT data using the model-based clustering approach of Rogiers *et al.* (in prep.; Figure 4D). In the latter approach I_c and the stratigraphical depth are the considered variables. Including dependence on the SBTs is required as the CPT signal within the lower aquifer gives rise to BE_{Fr} peaks without the presence of clear discrete clay lenses. This could be caused by the presence of concretions as the Diest Formation is known for containing iron sandstones. After the corrections introduced via Eq. 4, both the CPT-based K estimates (K_{CPT}) and the high-resolution borehole K logs are upscaled for each section of one meter long with arithmetic (for K_h) and harmonic (for K_v) means. This is done to avoid strong influences of small errors in the vertical position of both logs. Finally, a linear model was derived for upscaled CPT-based K_h and K_v values using upscaled borehole K_h and K_v for calibration. The parameters a , b , c_{SBT} , d_{SBT} and upscaled linear model parameters are all iteratively optimized using the Nelder and Mead (1965) approach. To avoid overfitting in optimizing parameters for K_{DPST} and K_{CPT} , a leave-one-out (one DPST interval or one meter of borehole core log data) cross-validation is performed for all data calibration cases. As the measurement supports of the high-resolution data (see Table 2) are too small for conditioning the groundwater flow model, the calibrated upscaled metre-scale datasets are used further on.

2.5. Development and conditioning of the groundwater flow model

The groundwater flow model used in this paper is an adaptation of the steady-state MODFLOW-2005 (Harbaugh 2005) model developed by Gedeon *et al.* (2011). Sensitivity analysis for this model was performed by Gedeon and Mallants (2012). Rogiers *et al.* (2012b) used it for testing different conceptual models of the Kasterlee Clay aquitard parameterization based on the CPT dataset without any further model modifications. The original model

of Gedeon *et al.* (2011) will be modified here to evaluate the degree and nature of improvement in model performance when different additional model parameter sets are used.

Model boundary conditions in the upper model layer are unchanged in so far as river, drain, fixed and general head boundary conditions are concerned (Figure 5A). Groundwater recharge is modified from a spatially uniform value to a spatially variable recharge based on separate vertical flux modelling. For this purpose Leterme and Mallants (2012) applied the HYDRUS-1D simulator (Šimůnek 2005) and determined recharge for crops, meadow, deciduous and coniferous land cover types (Figure 5B). No recharge value for the built-up area was previously defined; here we calibrate it as a percentage of the meadow recharge. The boundary conditions at the model perimeter within the upper aquifer are applied to all numerical layers. Where the model perimeter coincides with a river boundary, no-flow conditions are assumed at the model boundary cell faces for all numerical layers belonging to the upper aquifer. The river itself is simulated with the MODFLOW River package (Harbaugh *et al.* 2000). In the second aquifer below the aquitard, no-flow conditions are imposed only under the main river (Kleine Nete), which is assumed to drain the entire Neogene aquifer. At other locations, fixed head conditions are used based on head values obtained from the regional scale Neogene Aquifer Model (Gedeon 2008) (Figure 5C). For the aquitard, no flow boundary conditions are assumed everywhere, as the main flow through this unit is vertical.

The 86 piezometers included in the Gedeon *et al.* (2011) model calibration are now supplemented by 44 additional piezometers with filter depths of maximum 21 m below surface (Figure 5D shows all 130 piezometers). The mean filter length is 4.8 m, while the minimum and maximum amount to 0.5 and 50 m respectively, with 85% of the filters having a length of < 5 m. Average head values for the 86 piezometers were calculated from monthly observations between 1990 and 2010; the same was done for the 44 additional piezometers using all available observations within the same period. Moreover, all piezometers are now implemented in MODFLOW as multi-layer observations, with weighted contributions of different cells to the mean piezometer head. The weights are defined by the transmissivities obtained from the Gedeon *et al.* (2011) model K_h values. This is necessary to include effects of small-scale heterogeneity on piezometer heads especially for a model with a large number of thin numerical layers. Applying this new approach with the reference parameterization resulted in a higher accuracy (*i.e.* smaller model error) for several of the piezometers, compared to the Gedeon *et al.* (2011) model. Two types of data were used for calibration of the steady-state model in this study: 1) 130 time-averaged head values for piezometers in upper and lower aquifers, and 2) 68 vertical head differences obtained from multi-level piezometers. Using the Preconditioned Conjugate-Gradient solver, the root mean squared error (RMSE) is calculated as objective function for all head measurements and head differences. No weighting factors were applied for these two data types in the objective function.

The MODFLOW Hydrogeologic Unit Flow (HUF) package (Anderman and Hill 2000) was used for parameterization of the K_h and K_v data. The upper 38 m of the aquifer is treated differently from the remaining part of the model, as it includes most of the data for forward conditioning (*i.e.* for conditional geostatistical simulation) (Figure 6A). It is denoted as the HUF3D section (Figure 6B). The bottom of this section is obtained by smoothing the topography and subtracting 38 m. In the north-eastern part of the study area, this bottom is lowered by a few additional meters, to include the entire Kasterlee Clay unit within the HUF3D section. The HUF layers within this section are all parallel to the top of the Kasterlee Clay, which is the most pronounced lithostratigraphical boundary in the system. This horizon was mapped manually by Schiltz (2008, 2010; see also Wouters and Schiltz 2012), and validated by Rogiers *et al.* (in prep.) who showed that it is traceable in an automatic objective way using site-specific soil behaviour type classification. Height above this surface is referred to as stratigraphical depth z_{strat} further on. Due to the inclination of the geological layers, 39 inclined HUF layers of one meter thickness are needed above the Kasterlee Clay aquitard top, and 50 one-meter layers below it, to fill the entire HUF3D section. Where the layers extend beyond the HUF3D section, their thickness is set to zero. Parameterization of this section is conditioned on the metre-scale calibrated direct push and borehole datasets, in 3D. By setting the thickness of the grid cells in this section also to one metre, we avoid scaling issues, and the small measurement supports can be used for parameterizing the regional model.

The remaining part of the model includes most of the lower aquifer and contains part of the Diest, the Dessel, Berchem and Voort/Eigenbilzen Formations (Figure 6B, C). It is parameterized with 75 HUF layers parallel to the topography of the Formation boundaries (Figure 6B). Their varying thickness is based on the data density at the GT01, GT02 and Dessel-5 boreholes (Figure 6A). To parameterize this section, denoted as HUF1D, each of the 75 layers has a spatially uniform K_h and K_v value derived from the measurements from a single 180-m deep borehole across all relevant formations (Dessel-5) (Beerten *et al.* 2010; Labat *et al.* 2011), complemented with grain-size-based K predictions for two 90-m deep boreholes (GT01 and GT02), (Rogiers *et al.* 2012a). Because K values are available at only 3 spatial locations within the entire model domain, data conditioning within this HUF1D section is done in 1D, *i.e.* along z_{strat} .

Finally, the numerical discretization of the original model is adapted from 14 layers to 57 layers comprised of 50x50 m² cells (Figure 6C). The upper-most layer thickness is between 3 and 7.5 meter in order to have assurance the water table will be within the first layer for the entire model domain. This avoids numerical

instabilities due to dry cells resulting in poor convergence of the model. The following 46 layers all have a thickness of 1 m, and cover the entire HUF3D section, as well as the uppermost part of the HUF1D section in certain areas. The remaining 10 layers are used to fill up the remaining part of the model, with a linear increase in layer thickness. The maximum cell thickness is 14.3 m. Translation of the HUF parameters to the numerical grid is done automatically by MODFLOW. This enables running the model using bedding-parallel layers for the parameterization of the model, while keeping numerical stability.

To condition K within the HUF3D section on the direct push and borehole data, the following assumptions were made:

- 1) All K data obtained from laboratory and *in situ* hydraulic tests (borehole K logs and hydraulic direct push tests) are considered as the primary variable. The geotechnical direct push (CPT) data are used as secondary variable.
- 2) The amount of primary data is considerably less than that of the secondary data. Therefore, the latter will be used to derive reliable variograms which will then be used for both datasets, assuming their spatial statistics are similar. The cross-variogram is obtained by multiplying the direct variogram with the correlation between primary and secondary data, as calculated at the cored borehole locations.
- 3) Because the studied Neogene aquifer domain was shown to exhibit non-stationarity in K due to different K populations and regional trends associated with the different lithostratigraphical layers (Beerten *et al.* 2010; Rogiers *et al.* 2012b), a geostatistical approach that assumes reasonably stable (stationary) first and second-order moments cannot be used. A full 3D approach, accounting for non-stationarity of the mean, the correlation coefficient between primary and secondary data, and the secondary data variograms, is however computationally not feasible with the current implementation and processing capability. Therefore, we consider only non-stationarity in the vertical direction, obtained by using distance weighting (Harris *et al.* 2010; Machuca-Mory and Deutsch 2012). The data weights ω for this distance weighting at an anchor point o with n samples at locations u_α are obtained from the Gaussian kernel

$$\omega(u_\alpha; o) = \frac{\varepsilon + \exp\left(-\frac{(d(u_\alpha; o))^2}{2s^2}\right)}{n\varepsilon + \sum_{\alpha=1}^n \exp\left(-\frac{(d(u_\alpha; o))^2}{2s^2}\right)} \quad [5]$$

with vertical bandwidth s , background constant ε , and $d(u_\alpha; o)$ the vertical distance in our case. After testing different bandwidths, we decided to use a value of 8 m, as this resulted in the most diversity between the different locations within the lithostratigraphical column. For the background constant we used a value of 0.01. These weights are used for calculating local mean, correlations and direct variograms along the lithostratigraphical column. These geostatistical properties are then considered stationary within each of the bedding-parallel HUF layers.

- 4) To keep this multivariate non-stationary geostatistical simulation computationally feasible, we selected a subset of all target data locations (HUF cells) by only retaining each 5th row and column number. This effectively reduces the simulated data locations from 1,021,899 to 41,183. Inverse distance weighting with a power of 0.1 is used to extrapolate this data to the entire HUF3D section. Given the local stationarity assumption for all bedding-parallel HUF layers, a single layer can be simulated using standard sequential Gaussian simulation. Therefore, to speed up the simulation, the layers were simulated in a random order instead of using a random 3D path through all points. This simplification might induce slightly larger spatial correlation in the vertical direction, according to the observations of McLennan (2002). Finally, the subset of high-resolution CPT data from the $\sim 200 \times 400$ m² area is constrained to maximum one CPT location per 50x50 m² cell at the disposal site (*i.e.* 32 out of 85 CPTs are used). The secondary data at each of the 7 borehole locations is removed as it is redundant due to the presence of primary data. To limit the amount of correlated variables in the geostatistical simulation, the logarithmically-transformed K_h and K_v data were orthogonalized using principal component analyses. Geostatistical simulations were then done independently for the two resulting principal components (PCs). Creating one realization of the 41,183 HUF3D K_h and K_v parameters, by implementation of the above assumptions in the gstat R package (Pebesma 2004), requires about 70 minutes on a 2.00 GHz CPU.
- 5) We use direct simulation without a priori data transformation except for the use of logarithmic K values and orthogonalization. The marginal primary data histogram is approximated by post-processing while keeping the multivariate data conditioning intact, following the approach of Journel and Xu (1994; see also Goovaerts 1997):

$$z_{cc}(u_\alpha) = z(u_\alpha) + \left(\frac{\sigma(u_\alpha)}{\sigma_{max}}\right)^\varphi (z_c(u_\alpha) - z(u_\alpha)) \quad [6]$$

with $z_{cc}(u_\alpha)$ the final corrected value accounting for the data conditioning at location u_α , z_c the correction of the initial z value using rank transformation to obtain the target histogram, σ the kriging standard deviation obtained from all primary and secondary data, and σ_{max} the maximum of the latter. The parameter φ is the correction parameter and allows a balance between correction at all locations ($\varphi=0$), and no correction at all ($\varphi=\infty$). We use $\varphi=0.05$ after testing different values. Figure 7 provides a comparison between the histograms of the simulated data and the target primary data. Overall, the target histograms are well reproduced. Due to the fact that the primary data were collected in a particular section of the simulated domain and therefore do not provide a full spatial coverage, we decided not to perform hard conditioning on the target histograms. Such hard conditioning might lead to artificial spatial patterns with the rank transformation approach. This decision results in a small bias between the primary data and simulated data histograms, of which we believe it does not have a large impact on our model predictions.

The best performing (in terms of RMSE for head values and vertical head differences) out of 50 geostatistical K realizations was selected for the direct-push-conditioned HUF3D model section. Then optimization was performed for the following parameters: 1) a multiplication factor for the HUF1D K_h values, 2) a multiplication factor for the HUF1D vertical anisotropy values, 3) a multiplication factor for the overall groundwater recharge, and 4) the built area recharge percentage compared to the meadow recharge. The argument for optimising the HUF1D parameters is that we do expect systematic bias due to upscaling issues associated with our 1D approach. The 3D density of data for the HUF3D section is much higher (at least one CPT location on a grid with spacing of ~500-800 m), hence we expect less bias due to upscaling from point to model layers. Therefore we do not optimize the marginal distribution of that data; we only vary its spatial distribution in between the conditioning points. As we do not optimize the marginal distribution of the HUF3D data, we do calibrate the recharge. Starting values for three multiplication factors and built area recharge percentage were set at 1.5, 1.5, 0.6, and 70% respectively, after manual exploration of the parameter space. Automatic calibration was performed with the limited memory quasi-Newton algorithm of Byrd *et al.* (1995) for solving large nonlinear optimization problems with simple bounds on the variables.

To evaluate the relative model performance, it was compared with the calibrated reference simulation with homogeneous hydrogeological units as defined in Gedeon *et al.* (2011). In this case, the numerical grid and observation adaptations, which have a small but distinguishable effect on the model performance, are however included in the reference model, to make the comparison fair. One K value and vertical anisotropy are used per hydrogeological unit in this reference model. Three (grouped) K parameters were calibrated for obtaining the reference model parameterization (Gedeon *et al.* 2011).

2.6. Particle tracking

For illustrating the consequences of aquifer and aquitard heterogeneity on groundwater flow fields, we performed two particle tracking simulations. The particle tracking is performed with MODPATH version 5.0 (Pollock 1994). The particles are inserted at the future locations of the two disposal facilities for low-level radioactive waste (ONDRAF/NIRAS, 2010). Twelve numerical grid cells were selected for the source area covering an area of 30,000 m². Each grid cell was provided with a single particle at the water table in the centre of the cell.

3. Results and discussions

3.1. Calibration of DP data for K estimation

Table 3 summarizes the calibration approach and results for the different types of high-resolution direct push data. The calibrated hydraulic direct push K values (all K_h and upscaled to the meter-scale) are plotted in Figure 8 versus the K_{DPST} values used for calibration. The coefficient of determination R^2 is 0.82 for the HPT data, and 0.74 for the DPIL data. The higher R^2 value for HPT is probably a consequence of performing ordinary kriging on higher resolution data for the HPT logs compared to the lower resolution DPIL logs. Given these relatively high coefficients of determination, we decided to add this calibrated high-resolution hydraulic direct push data to the primary variable in the geostatistical simulation approach. The leave-one-out cross-validation results in Table 3 show that similar accuracy can be expected for the high-resolution hydraulic direct push log sections without corresponding K_{DPST} data.

The point-scale relationship between CPT-based I_c and the high-resolution borehole core K logs from Rogiers *et al.* (2013b) is displayed in Figure 9A. The coefficient of determination of 0.32 is rather low, which is probably due to the difficult matching (in terms of location) of borehole core and CPT logs at the cm scale, and the low sensitivity of CPT data to variations in K . The lowest K values, for instance, clearly remain undetected by the I_c parameter due to the relatively large support volume of the CPT measurement compared to the much smaller

spatial scale at which lithological heterogeneities occur. The results obtained after calibration with the BE_{Fr} approach, and upscaling to meter-scale values, are shown in Figure 9B. Coefficients of determination are 0.43 and 0.75 for respectively K_h and K_v . For the latter, this is similar to the hydraulic direct push data. The leave-one-out cross-validation in Table 3 again shows that similar accuracy is expected for the CPT-based K values for the entire dataset. Given the lower coefficient of determination for the CPT-based K_h and the fact that these values are derived from geotechnical characteristics of the subsurface, we decided to treat both CPT-based K_h and K_v data as a secondary variable in the geostatistical framework. For illustration of the CPT-based K estimates, the upscaled primary and calibrated secondary data at three cored boreholes are shown in Figure 10.

3.2. Hydraulic conductivity realizations

The non-stationary (along z_{strat}) means for logarithmically-transformed K_h and K_v and their first two PCs, and the correlation coefficients between primary and secondary data (for PC1 and PC2) are shown in Figure 11. The K means clearly illustrate the presence of the Kasterlee Clay aquitard, with K_v and K_h values as low as 10^{-6} and 10^{-5} m/s between 0 and -5 m stratigraphical depth, z_{strat} (Figure 11A). The lower aquifer also shows lower K values than the upper aquifer. For the uppermost (mainly Quaternary sediments) and lowest (Diest Formation) HUF layers, the mean values are put equal to the global mean, because little or no data are available at these locations. The same holds true for the PC1 and PC2 means in Figure 11B. The correlation coefficient between primary and secondary data is generally above 0.75, except for an interval around 25 m stratigraphical depth, which corresponds to the coarse white Mol Upper Sands. This is a consequence of the inability of the CPT test to detect this coarse, high K unit (Figure 4A,C), as it shows lower Q_{in} values where it is expected to yield higher Q_{in} values compared to the surrounding finer-textured sediments. Hence, this particular discrepancy between secondary and primary data is accounted for in the simulation by a lower correlation coefficient.

The results of the distance-weighted variography are shown in Figure 12. All variogram models consist of a nugget value and two superimposed spherical models. A first observation is the high semi-variances of the 0 and -10 m stratigraphical depth levels. The semi-variances at these depths are highly influenced by the heterogeneity within the Kasterlee Clay aquitard. A second consistency between all variograms is the low semi-variance values, especially at short lag distances, of the 20, 30 and 40 m stratigraphical depth locations. These correspond to the Mol Formation units, which are the most homogeneous sediments in the study area. The other variograms fall in between these two extreme cases, both with their total sill and relative nugget values.

The K_h and K_v data for three example HUF layers of the three best performing (in terms of RMSE) out of 50 realizations are shown in Figure 13. The contrast between the upper aquifer (HUF layer 25), Kasterlee Clay aquitard (layer 43) and lower aquifer (layer 50) is very clear. The highest K values are located within the upper aquifer, where only very little anisotropy seems to be present. The Kasterlee Clay aquitard clearly has low K values, but some of the K_h data are as large as those of the aquifer units. This results in a large range of vertical anisotropy values (up to 10^6). The lower aquifer K_h values are clearly somewhat lower than that of the upper aquifer, but the K_v values are clearly lower, which also results in a non-negligible vertical anisotropy (about $10^{1.5}$ for layer 50).

Figure 13 also illustrates that the presented approach is able to capture both large- and small scale heterogeneity. Different large-scale K field characteristics and different degrees of small-scale heterogeneity are clearly present in the different lithostratigraphical units. This is clearly a consequence of the multi-scale nature of the aquifer characterization (*i.e.* from centimetre-to regional scale), and the use of a small, but intermediate scale to put the characterization data in the model.

3.3. Groundwater flow model performance testing

The optimized, best performing realization out of a set of 50 realizations of the HUF3D K parameters was selected for comparison with the reference groundwater model of Gedeon *et al.* (2011). Figure 14 shows the steady-state model predicted heads versus the long-term mean observed head values, as well as the vertical head differences (inset graph) that were added to the objective function. The total RMSE decreased from 0.37 to 0.34 based on the model updates presented in this paper. The improvement is most easily noticeable for the vertical head differences, which reflects directly the more appropriate combination of recharge and K values in the model update. Many of the simulated head values are now also much closer to the observations, although a few anomalies occur in the updated model, with three misfits of more than one meter. These misfits are strongly influenced by local characteristics of the K field, as the misfits occur in the very sensitive southern part of the model (Gedeon and Mallants 2012), where the Kasterlee Clay aquitard is at its shallowest position. This is also evident from a comparison between the water table contour maps in the southern part of the model in Figure 15. To resolve the misfits, the use of a larger number of realizations to better capture the spatial heterogeneity, or the optimization of the spatially distributed parameters thereby honouring the geostatistical model, could be investigated. Moreover, nearby surface water boundary conditions might influence these outliers as well, but studying the effects of these boundary conditions is outside the scope of this study.

3.4. Illustration of differences in the flow field by particle tracking

The particle pathlines are shown in Figure 16. The general trend is similar in both cases, as the groundwater heads are similar and the sink locations (*i.e.* the groundwater discharges in the Witte Nete river at the eastern boundary of the model) are the same. There are however subtle differences, especially for the particles originating from the western source area (area 1 in Figure 16). The profile views indicate that the particles follow a shallower path through the aquifer in the heterogeneous model. This is a consequence of the presence of the shallow, coarse and high K Upper Mol Sands in the north-eastern part of the model. This unit is not treated separately and recognized as a separate unit in the homogeneous model. Because this layer could provide faster pathways for contaminants, it warrants further investigations.

As both models perform similar, the more data-based one is probably more realistic, but additional data like groundwater age tracer concentrations could potentially allow further discrimination between both cases. These results clearly illustrate the importance of accounting for subsurface heterogeneity, and the usefulness of direct push data for aquifer characterization and groundwater flow model parameterization.

4. Conclusions and perspectives

Direct push technologies are very well suited for aquifer characterization from the local to the regional scale. Moreover, we have illustrated that standard cone penetration test (CPT) data can be used to provide high-resolution hydraulic conductivity (K) estimates at regional scale, even if suitable calibration data is available only at a lower spatial granularity. Our novel approach allows mapping aquifer and aquitard properties in 3D for unconsolidated sediments for basins of several tens to hundreds of km^2 . Moreover, the use of hydraulic direct push tools such as direct push injection logging and the hydraulic profiling tool provides for an even higher accuracy in K compared to standard CPT when calibrated to the Direct Push Slug Tests. Both the geotechnical and hydraulic direct push K data were integrated together with borehole K data in a regional groundwater flow model. We considered all K data derived from laboratory and *in situ* hydraulic tests as the primary variable in a geostatistical simulation, while the geotechnical data (CPT) was considered as a secondary variable. A non-stationary (in the vertical direction) geostatistical approach was used to condition the model K parameter realizations. After optimization, and the selection of the best out of 50 realizations, model performance improved when compared to a previously developed groundwater model with spatially uniform K values for the different hydrogeological units (reference model). The model improvement was most noticeable in regards to the vertical head differences across the aquitard: the coefficient of determination R^2 increased from 0.79 to 0.88 while the RMSE decreased from 0.39 to 0.29 (all other goodness of fit indicators equally improved). Particle tracking simulations illustrated more tortuous flow paths for the groundwater flow model with heterogeneous K in aquifers and aquitards compared to the reference model.

We calibrated standard CPT data using cored borehole K data; a more cost-effective calibration could be achieved by replacing the more expensive coring with hydraulic direct push data. For example, a regional CPT campaign could then be complemented with a few well-chosen hydraulic direct push test sites, which would significantly reduce the need for expensive drilling. The systematic use of CPTu tests could also enhance the estimation of K from CPT data. Moreover, hydraulic tests can be carried out in piezometers or wells installed with direct push equipment; development of such wells is far less expensive than conventional drilling.

The present study demonstrates that the hydraulic direct push K data is compatible with the K data from the cored boreholes. Moreover, the hydraulic direct push methods were able to identify a high K sand layer that was previously unidentified by the CPT data. Even when calibrated, the standard CPT data did not really capture this high K zone. The discrepancy between primary (*i.e.* K from borehole cores) and secondary (CPT) data for this coarser layer was resolved in the geostatistical simulation by adopting a non-stationary geostatistical approach. The multi-scale nature of the aquifer characterization data, *i.e.* from centimetre-scale vertical data resolution to regional sampling using km-scale spatial intervals, allowed the reproduction of both small-scale vertical and large-scale horizontal heterogeneity. The systematic use of small support volumes facilitated the integration of the different data types considerably.

Based on the improvements made in the groundwater flow model, especially the ability to generate heterogeneous flow fields that better reproduce the real heterogeneity in subsurface flow and transport properties than a spatially uniform model, more realistic solute transport simulation can now be carried out that explicitly account for aquifer heterogeneity. The multi-scale nature of the site characterization data, and the derived K fields can be used for mimicking macroscopic dispersion. Finally, the use of groundwater geochemistry and age tracer data for calibrating or validating the groundwater flow and solute transport models is highly recommended.

Acknowledgements

The authors are grateful to ONDRAF/NIRAS, the Belgian Agency for Radioactive Waste and Enriched Fissile Materials, for providing the CPT data and supporting the UFZ hydraulic direct push campaign. Findings and

conclusions in this paper are those of the authors and do not necessarily represent the official position of ONDRAF/NIRAS.

The authors thank Markus Tinter, Manuel Kreck and Helko Kotas for performing and interpreting the HPT, DPIL and DPST measurements. The authors further wish to acknowledge the Fund for Scientific Research – Flanders for providing a Postdoctoral Fellowship to Marijke Huysmans. We also thank Luk Peeters for useful discussions concerning the groundwater flow model refinement.

References

- Anderman ER, Hill MC (2000) MODFLOW-2000, the U.S. Geological Survey Modular Ground-water model – documentation of the hydrogeologic-unit flow (HUF) package. USGS Open file report 00-342, Denver, Colorado
- Battle-Aguilar J, Brouyere S, Dassargues A, Morasch B, Hunkeler D, Hohener P, Diels L, Vanbroekhoeven K, Seuntjens P, Halen H (2009) Benzene dispersion and natural attenuation in an alluvial aquifer with strong interactions with surface water. *Journal of Hydrology* **361**: 305-317.
- Bhattacharya B, Solomatine DP (2005) Machine learning in soil classification. *Proceedings of International Joint Conference on Neural Networks* **19**: 2694 – 2699.
- Beerten K, Wemaere I, Gedeon M, Labat S, Rogiers B, Mallants D, Salah S, Leterme B (2010) Geological, hydrogeological and hydrological data for the Dessel disposal site. Project near surface disposal of category A waste at Dessel. STB-SIE(HYD) - Version 1, NIROND-TR 2009-05 E (p. 261).
- Bierkens MFP (2005) Designing a monitoring network for detecting groundwater pollution with stochastic simulation and a cost model. *Stochastic Environmental Research and Risk Assessment* **20**(5): 335–351. doi:10.1007/s00477-005-0025-2
- Bouwer H, Rice RC (1976) A slug test for determining hydraulic conductivity of unconfined aquifer with completely or partially penetrating wells. *Water Resources Research* **12**: 423-428.
- Butler JJ (2002) A simple correction for slug tests in small-diameter wells. *Ground Water* **40**(3): 303–308.
- Butler JJ, Healey JM, McCall GW, Garnett EJ, Loheide II SP (2002) Hydraulic tests with direct-push equipment. *Ground Water* **40**(1): 25–36.
- Butler JJ, Dietrich P (2004) New methods for high-resolution characterization of spatial variations in hydraulic conductivity. In: International Symposium on Hydrogeological Investigation and Remedial Technology, Jhongli, Taiwan, pp. 42-55.
- Butler JJ, Dietrich P, Wittig V, Christy T (2007) Characterizing hydraulic conductivity with the direct-push permeameter. *Ground Water* **45**(4): 409–419.
- Byrd RH, Lu P, Nocedal J, Zhu C (1995) A limited memory algorithm for bound constrained optimization. *SIAM J. Scientific Computing* **16**: 1190–1208.
- Chai J, Agung P, Hino T, Igaya Y, Carter J (2011) Estimating hydraulic conductivity from piezocone soundings. *Géotechnique* **61**(8) : 699–708. doi:10.1680/geot.10.P.009
- Chai J, Sheng D, Carter JP, Zhu H (2012) Coefficient of consolidation from non-standard piezocone dissipation curves. *Computers and Geotechnics* **41**: 13–22. doi:10.1016/j.compgeo.2011.11.005
- Christian J, Van Der Sluijs JP, Brown J, Van Der Keur P (2006) A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources* **29**: 1586–1597. doi:10.1016/j.advwatres.2005.11.013
- Christian J, Van Der Sluijs JP, Lajer A, Vanrolleghem PA (2007) Uncertainty in the environmental modelling process - A framework and guidance. *Water* **22**: 1543–1556. doi:10.1016/j.envsoft.2007.02.004
- Dietrich P, Leven C (2006) Direct Push-Technologies. In: Kirsch, R. (Ed.), *Groundwater Geophysics*. Springer, Berlin, pp. 321-340.
- Dietrich P, Butler JJ, Faiss K (2008) A rapid method for hydraulic profiling in unconsolidated formations. *Ground water* **46**(2): 323–8.
- Elsworth D, Lee DS (2007) Limits in determining permeability from on-the-fly uCPT sounding. *Géotechnique* **57**(8): 679–685. doi:10.1680/geot.2007.57.8.679
- Flach GP, Harris MK, Smits AD, Syms FH (2005) Modeling aquifer heterogeneity using cone penetration testing data and stochastic upscaling methods. *Environmental Geosciences* **12**(1): 1–15.
- Gedeon M (2008) Neogene Aquifer Model, SCK•CEN-ER-48 (p. 110). Mol.
- Gedeon M, Mallants D, Vandersteen K, Rogiers B, Laloy E (2011) Hydrogeological modelling of the Dessel site - Overview report (p. 227).
- Gedeon M, Mallants D (2012) Sensitivity Analysis of a Combined Groundwater Flow and Solute Transport Model Using Local-Grid Refinement: A Case Study. *Mathematical Geosciences* **44**(7): 881–899. doi:10.1007/s11004-012-9416-3
- GEOPROBE (2007) GEOPROBE hydraulic profiling tool (HPT) systems: Standard operating procedure. Technical Bulletin MK3137, GEOPROBE.

- Hammond GE, Lichtner PC (2010) Field-scale model for the natural attenuation of uranium at the Hanford 300 Area using high-performance computing. *Water Resources Research* **46**(9): 1–31. doi:10.1029/2009WR008819
- Harbaugh AW, Banta ER, Hill MC, McDonald MG (2000) MODFLOW-2000, the U.S. Geological Survey modular ground-water model -- User guide to modularization concepts and the Ground-Water Flow Process: U.S. Geological Survey Open-File Report 00-92, 121 p.
- Harbaugh AW (2005) MODFLOW-2005, The US Geological Survey modular ground-water model—the Ground-Water Flow Process: US Geological Survey Techniques and Methods 6-A16, USGS, 2005.
- Harris P, Charlton M, Fotheringham A (2010) Moving window kriging with geographically weighted variograms. *Stoch Environ Res Risk Assess* **24**: 1193–1209. doi:10.1007/s00477-010-0391-2
- Huysmans M, Madarász T, Dassargues A (2006) Risk assessment of groundwater pollution using sensitivity analysis and a worst-case scenario analysis. *Environmental Geology* **50**(2): 180–193. doi:10.1007/s00254-006-0197-1
- Johnson BA, Carpenter PJ (2012) Geophysical response of slackwater and sandy terrace deposits near Savanna, Northwestern Illinois. *Environmental Earth Sciences* **68**(4): 955–964. doi:10.1007/s12665-012-1798-5
- Journel AG, Xu W (1994) Posterior Identification of histograms Conditional to Local Data. *Mathematical Geology* **22**: 323-359.
- Köber R, Hornbruch G, Leven C, Tischer L, Grossmann J, Dietrich P, Weiss H, Dahmke A (2009) Evaluation of combined direct-push methods used for aquifer model generation. *Ground Water* **47**(4): 536–46. doi:10.1111/j.1745-6584.2009.00554.x
- Labat S, Gedeon M, Beerten K, Maes T (2011) Dessel-5 borehole: technical aspects and hydrogeological investigations. SCK•CEN-ER-151 (p. 39).
- Laga P, Louwye S, Geets S (2001) Paleogene and Neogene lithostratigraphic units (Belgium). *Geologica Belgica* **4**(1-2): 135–152.
- Lee DS, Elsworth D, Hryciw R (2008) Hydraulic Conductivity Measurement from On-the-Fly uCPT Sounding and from VisCPT. *Journal of Geotechnical and Geoenvironmental Engineering* **134**(12): 1720-1729. doi:10.1061/(ASCE)1090-0241(2008)134:12(1720)
- Lessoff SC, Schneidewind U, Leven C, Blum P, Dietrich P, Dagan G (2010) Spatial characterization of the hydraulic conductivity using direct-push injection logging. *Water Resources Research* **46**(12): 1–9.
- Leterme B, Mallants D (2012) Simulation of evapotranspiration and groundwater recharge in the nete catchment accounting for different land cover types and for present and future climate conditions. Project near surface disposal of category A waste at Dessel. External report of the Belgian Nuclear Research Centre. SCK•CEN-ER-192.
- Liu G, Butler JJ, Bohling GC, Reboulet E, Knobbe S, Hyndman DW (2009) A new method for high-resolution characterization of hydraulic conductivity. *Water Resources Research* **45**(8): 1-6.
- Liu G, Butler JJ, Reboulet E, Knobbe S (2011) Hydraulic conductivity profiling with direct push methods. *Grundwasser*, **17**(1), 19–29. doi:10.1007/s00767-011-0182-9
- Louwye S, De Schepper S, Laga P, Vandenberghe N (2007) The Upper Miocene of the southern North Sea Basin (northern Belgium): a palaeoenvironmental and stratigraphical reconstruction using dinoflagellate cysts. *Geological Magazine* **144**(1): 33-52.
- Louwye S, Laga P (2008) Dinoflagellate cyst stratigraphy and palaeoenvironment of the marginal marine Middle and Upper Miocene of the eastern Campine area, northern Belgium (southern North Sea Basin). *Geological Journal* **43**(1): 75–94.
- Louwye S, De Schepper S (2010) The Miocene–Pliocene hiatus in the southern North Sea Basin (northern Belgium) revealed by dinoflagellate cysts. *Geological Magazine* **147**(05): 760-776.
- Lunne T, Robertson PK, Powell JJM (1997) Cone Penetration Testing in Geotechnical Practice. Blackie Academic/Routledge Publishing, New York.
- Machuca-Mory DF, Deutsch CV (2012) Non-stationary Geostatistical Modeling Based on Distance Weighted Statistics and Distributions. *Mathematical Geosciences* **45**(1): 31–48. doi:10.1007/s11004-012-9428-z
- Mallants D, Labat S, Gedeon M (2003) Bijkomende sitekarakterisatie voor de nucleaire zone Mol-Dessel : bepaling van de hydrogeologische parameters. SCK•CEN-R-3703 (p. 82). Mol.
- Mariethoz G, Renard P, Caers J (2009) Bayesian inverse problem and optimization with Iterative Spatial Resampling. *Water Resources Research* **46**(11): 1–17. doi:10.1029/2010WR009274
- McCall W, Nielsen DM, Farrington SP, Christy TM (2006) Use of Direct-Push Technologies in Environmental Site Characterization and Ground-Water Monitoring. In: Nielsen DM (ed), *Practical Handbook of Environmental Site Characterization and Ground-Water Monitoring*. CRC Press Taylor and Francis Group, Boca Raton, pp. 345-471.
- McCall W, Christy TM, Christopherson T, Issacs H (2009) Application of Direct Push Methods to Investigate Uranium Distribution in an Alluvial Aquifer. *Ground Water Monitoring & Remediation*, **29**(4): 65-76.

- McLennan JA (2002) The Effect of the Simulation Path in Sequential Gaussian Simulation. Centre for Computational Geostatistics report 2002-54 (p. 16).
- Nelder JA, Mead R (1965) A simplex algorithm for function minimization. *Computer Journal* **7**: 308–313.
- Nilsson B, Højberg A, Refsgaard J, Trolborg L, (2007) Uncertainty in geological and hydrogeological data. *Hydrology and Earth System Sciences* **11**, 1551–1561.
- ONDRAF/NIRAS (2010) Het cAt-project in Dessel. Een langetermijnoplossing voor het Belgische categorie A-afval. Retrieved from http://www.niras-cat.be/downloads/cAt_masterplan_NL_LOW.pdf on 07-12-2011.
- Parez L, Fauriel R (1988) Le piézocône améliorations apportées à la reconnaissance des sols. *Rev. Franc. Géotech.*(44): 13-27.
- Patyn J, Ledoux E, Bonne A (1989) Geohydrological research in relation to radioactive waste disposal in an argillaceous formation. *Journal of Hydrology* **109**(3-4): 267–285.
- Pebesma EJ (2004) Multivariable geostatistics in S: the gstat package. *Computers & Geosciences* **30**: 683-691.
- Pollock DW (1994) User's Guide for MODPATH/MODPATH-PLOT, Version 3: A particle tracking post-processing package for MODFLOW, the U.S. Geological Survey finite-difference ground-water flow model: U.S. Geological Survey Open-File Report 94-464, 6 ch.
- Rein A, Popp S, Zacharias S, Leven C, Bittens M, Dietrich P (2010) Comparison of approaches for the characterization of contamination at rural megasites. *Environmental Earth Sciences* **63**(6): 1239–1249. doi:10.1007/s12665-010-0797-7
- Robertson PK (1990) Soil classification using the cone penetration test. *Canadian Geotechnical Journal*, 27(1): 151-158.
- Robertson PK, Wride CE (1998) Evaluating cyclic liquefaction potential using the cone penetration test, *Canadian Geotechnical Journal* **35**: 442 – 459.
- Robertson P (2010) Estimating in-situ soil permeability from CPT & CPTu. CPT'10, 2nd international symposium on cone penetration testing **2**: 51.
- Rojas R, Feyen L, Dassargues A (2008) Conceptual model uncertainty in groundwater modeling: Combining generalized likelihood uncertainty estimation and Bayesian model averaging. *Water Resources Research* **44**(12): 1–16. doi:10.1029/2008WR006908
- Rogiers B, Schiltz M, Beerten K, Gedeon M, Mallants D, Batelaan O, Dassargues A, Huysmans M (2010) Groundwater model parameter identification using a combination of cone penetration tests and borehole data. Proceedings of the IAHR International Groundwater Symposium 2010, Valencia, 22-24 September 2010, 19 pp.
- Rogiers B, Mallants D, Batelaan O, Gedeon M, Huysmans M, Dassargues A (2011) Site-specific soil classification from cone penetration tests and borehole data: a multivariate statistical analysis. NovCare 2011, Cape Cod, 9-11 May 2011.
- Rogiers B, Mallants D, Batelaan O, Gedeon M, Huysmans M, Dassargues A (2012a) Estimation of hydraulic conductivity and its uncertainty from grain-size data using GLUE and artificial neural networks. *Mathematical Geosciences* **44**(6): 739–763.
- Rogiers B, Mallants D, Batelaan O, Gedeon M, Huysmans M, Dassargues A (2012b) The usefulness of CPTs for deterministic, spatially heterogeneous, large-scale aquitard parameterisation. In Oswald SE, Kolditz O, Attinger S (Eds.), *Models - Repositories of Knowledge, IAHS Publ. 355* (pp. 41–47).
- Rogiers B, Winters P, Huysmans M, Beerten K, Mallants D, Gedeon M, Batelaan O, Dassargues A (2012c) Centimeter-scale secondary information on hydraulic conductivity using a hand-held air permeameter on borehole cores. *Geophysical Research Abstracts* **14**: EGU2012-1794-1. EGU General Assembly 2012, Vienna, 22-27 April 2012.
- Rogiers B, Beerten K, Smeekens T, Mallants D, Gedeon M, Huysmans M, Batelaan O, Dassargues A (2013a) Derivation of flow and transport parameters from outcropping sediments of the Neogene aquifer, Belgium. *Geologica Belgica* **16**(3): 129-147.
- Rogiers B, Winters P, Huysmans M, Beerten K, Mallants D, Gedeon M, Batelaan O, Dassargues A (2013b) High resolution hydraulic conductivity logging of borehole cores using air permeability measurements. *Hydrogeology Journal*, DOI: 10.1007/s10040-014-1144-y
- Rogiers B, Mallants D, Batelaan O, Gedeon M, Huysmans M, Dassargues A (In prep.) Model-based classification and automated lithostratigraphical mapping of CPT data. In preparation.
- Saito H, Goovaerts P (2003) Selective remediation of contaminated sites using a two-level multiphase strategy and geostatistics. *Environmental science & technology* **37**(9): 1912–8.
- Schiltz M (2008) Lithological and Stratigraphical interpretation by means of cone penetration tests (CPT's) in the Dessel-Kasterlee-Geel-Mol area. Bvba SAMSUFFIT Geoservices, Fieldsurvey cAt 2008.
- Schiltz M (2010) Lithological and Stratigraphical interpretation of cone penetration tests (CPT's) executed for the first tumulus at the disposal site in Dessel and in the Dessel-Kasterlee-Geel-Mol area. Bvba SAMSUFFIT Geoservices, Fieldsurvey cAt 2010.

- Selroos J-O, Painter SL (2012) Effect of transport-pathway simplifications on projected releases of radionuclides from a nuclear waste repository (Sweden). *Hydrogeology Journal* **20**:1467–1481. doi:10.1007/s10040-012-0888-5
- Šimůnek J, Sejna M, van Genuchten MT (2005) HYDRUS-1D, version 4.14, code for simulating the one-dimensional movement of water, heat, and multiple solutes in variably saturated porous media, Tech. rep., University of California Riverside, 2005.
- Springer RK, Gelhar LW (1991) Characterization of large-scale aquifer heterogeneity in glacial outwash by analysis of slug tests with oscillatory response, Cape Cod, Massachusetts, U.S. Geol. Surv. Water Res. Invest. Rep. 91-4034, pp. 36-40.
- Teh C, Houlby G (1991) An analytical study of the cone penetration test in clay. *Geotechnique* **41**(1): 17-34.
- Tillmann A, Englert A, Nyari Z, Fejes I, Vanderborcht J, Vereecken H (2008) Characterization of subsoil heterogeneity, estimation of grain size distribution and hydraulic conductivity at the Krauthausen test site using Cone Penetration Test. *Journal of contaminant hydrology* **95**(1-2): 57–75.
- Tinter M (2012) Vergleich direkter und indirekter Verfahren zur Bestimmung des Durchlässigkeitsbeiwertes an klastischen Sedimenten der Mol-Formation am Standort Dessel in Belgien. Masterarbeit. Universität Leipzig.
- Van Baars S, Van De Graaf HC (2007) Determination of Organic Soil Permeability Using The Piezocone Dissipation Test. *Environmental and Engineering Geoscience* **13**(3): 197–203. doi:10.2113/gseegeosci.13.3.197
- Vienken T (2010) Critical evaluation of vertical high resolution methods for determining hydraulic conductivity. Dissertation der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen zur Erlangung des Grades eines Doktors der Naturwissenschaften.
- Vienken T, Tinter M, Rogiers B, Leven C, Dietrich P (2012) Evaluation of field methods for vertical high resolution aquifer characterization. Abstract ID 1494473, AGU Fall Meeting, San Francisco, USA, 3-7 December 2012.
- Wouters L, Schiltz M (2012) Overview of the field investigations in and around the nuclear site of Mol-Dessel, NIROND-TR 2011-42.
- Yu L, Rogiers B, Gedeon M, Marivoet J, De Craen M, Mallants D (2013) A critical review of laboratory and in-situ hydraulic conductivity measurements for the Boom Clay in Belgium. *Applied Clay Science* **75-76**: 1–12. doi:10.1016/j.clay.2013.02.018

Figure captions

Figure 1: A) Location of the study area within the hydrography of Flanders, Belgium. B) Overview of drilling and direct-push (DP) sites within the study area. C) Drilling and DP sites at the future low-level waste disposal site (ONDRAF/NIRAS, 2010). D) Location of the hydraulic direct push tests. E) Location of Flanders within Europe.

Figure 2: East-west lithostratigraphical profile through the study area. Location of the profile is indicated in Figure 1. Height exaggeration is 16x.

Figure 3: Examples of hydraulic direct push logs with approximate locations of the water table and the main lithostratigraphical boundaries estimated based on the Dessel-2 borehole interpretations (Beerten et al. 2010). K_{rel} = relative hydraulic conductivity (K) from Direct Push Injection Logging (DPIL; l/h/bar) or Hydraulic Profiling Tool (HPT; ml/min/Pa); K_{DPST} = absolute K (m/s) from Direct Push Slug Test (DPST). The DPIL data beyond the lower K limit of the method are not shown.

Figure 4: Example of A) high-resolution borehole core K logs (horizontal (K_h) and vertical (K_v) K for the Dessel-2 borehole; Rogiers et al. 2013b), B) the corresponding standard normalized CPT data (cone resistance (Q_{in}) and friction ratio (F_r)), C) CPT- derived soil behaviour type index I_c and the friction ratio boundary energy BE_{Fr} and D) site-specific soil behaviour types (SBT) obtained using the model-based clustering approach of Rogiers et al. (in prep.) with I_c and stratigraphical depth z_{strat} as variables.

Figure 5: Boundary conditions (A-C) of the groundwater flow model, and location of the piezometers used in model calibration (D). A) upper aquifer boundary conditions, B) upper layer recharge in percentage of the total annual precipitation (899 mm; Leterme and Mallants 2012), and C) lower aquifer boundary conditions (applicable to all model layers). The built-up area recharge is taken as a parameter for optimization, and the lakes correspond to the general head conditions in A).

Figure 6: A) Litho-stratigraphy and projected borehole and direct push logs (HDP). B) Hydrogeological unit flow (HUF) package layers used for the parameterization of K . HUF3D indicates the section in which parameterization is done in 3D; HUF1D indicates all layers have a spatially uniform K value and parameterization is by layer (1D vertical direction). C) Vertical numerical model discretization. The K_h and K_v parameters provided for the HUF layers are mapped to the numerical layers using, respectively, arithmetic and harmonic means.

Figure 7: Comparison between simulated (A, B, C) and target primary (D, E, F) data histograms for horizontal hydraulic conductivity (K_h) (A, D), vertical hydraulic conductivity (K_v) (B, E), and vertical anisotropy (K_h/K_v) (C, F).

Figure 8: Calibration results for upscaled A) K_{HPT} and B) K_{DPIL} data, using the absolute K_{DPST} data. OK denotes the ordinary kriging that is performed for the DPIL data, to get a better DPST screen support estimate.

Figure 9: A) Scatterplot of the high-resolution (0.05 m) borehole core K_h log and the corresponding soil behaviour type index (I_c). B) Scatter plot of calibrated $K_{h,CPT}$ and $K_{v,CPT}$ versus upscaled (to 1 m) high-resolution K logs for the cored boreholes, obtained by Rogiers *et al.* (2013b).

Figure 10: Comparison between upscaled borehole core K logs and CPT-calibrated K_h (left) and K_v (right) logs for A) the Dessel-3 (D3), B) Dessel-4 (D4), and C) Retie-1 (R1) boreholes.

Figure 11: A) Non-stationary geometric hydraulic conductivity (K) means (from primary data) along the stratigraphical depth (z_{strat}). B) Non-stationary principal component (PC1 and PC2) means along z_{strat} . C) Non-stationary correlation coefficients between primary and secondary data PCs.

Figure 12: Overview of the distance-weighted variogram models for PC1 (A, B) and PC2 (C, D). Both horizontal (A, C) and vertical (B, D) variograms are shown. Different colour codes are for different stratigraphical depths between -50 and 50 m (with respect to the top of the aquitard), at 10 m intervals.

Figure 13: Example of K_h and K_v fields of the three best performing realizations out of 50. HUF layer 25 represents the upper aquifer, layer 43 the Kasterlee Clay aquitard, and layer 50 the lower aquifer.

Figure 14: Predicted versus observed heads for A) the reference case with homogeneous hydrogeological layers, and B) the best performing direct push-conditioned realization. The dashed lines represent a residual of 1 m. The inset shows the vertical head differences used in the objective function. RMSE = root mean squared error; MAE = mean absolute error; ME = mean error.

Figure 15: Hydraulic head in the first numerical model layer, representing the water table elevation, for the reference model with homogeneous hydrogeological units (A) and the heterogeneous improved model (B).

Figure 16: Particle pathlines for the reference model with homogeneous hydrogeological units (A, C) and the groundwater flow model with heterogeneous K in both aquifers (B, D). Both a map (A, B) and profile view (C, D) are shown.

Table captions

Table 1: List of symbols and abbreviations.

Table 2: Overview of the different data types used in this work, the amount of data, their support volume and vertical resolution.

Table 3: Overview of the calibration results for the high-resolution direct push data.

Table 1

Symbol	Explanation
K	Hydraulic conductivity
CPT	Cone Penetration Test
CPTu	CPT with pore pressure measurement (Piezocone test)
DPIL	Direct Push Injection Logging
HPT	Hydraulic Profiling Tool
DPST	Direct Push Slug Test
SBT	Soil behaviour type
Q_{tn}	Normalized cone resistance (MPa)
F_r	Normalized friction ratio (%)
z_{strat}	Stratigraphical depth
z_{masl}	Depth in meter above sea level
BE_{Fr}	Friction ratio boundary energy
I_c	Soil behaviour type index
PC	Principal Component
RMSE	Root mean squared error

Table 2

Data source	Number of logs	Number of measurements	Approximate measurement support	Vertical resolution	Reference
High resolution borehole core hydraulic conductivity logs	7	5,230 air permeameter measurements; 368 steel ring core sample lab analyses	100 cm ³	0.05 m	Rogiers et al. 2013b
Cone penetration test data (CPT)	265	~480,000	~500 cm ³	0.02 m	Schiltz 2008, 2010; Wouters & Schiltz 2012
Direct push injection logging (DPIL)	17	592	~4.2 dm ³	0.3 m	
Hydraulic profiling tool (HPT)	6	5818	~100 cm ³	0.015 m	Vienken et al. 2012; Tinter 2012
Direct push slug tests (DPST)	6	19	~900 dm ³ (based on Bouwer and Rice 1976)	-	
Grain size-based <i>K</i> estimates (for deeper part aquifer)	2	365	~100 cm ³	1 m	Rogiers <i>et al.</i> 2012a
Dessel 5 borehole core sample laboratory tests (for deeper part aquifer)	1	54	100 cm ³	2 m	Beerten <i>et al.</i> 2010 Labat <i>et al.</i> 2011

Table 3

Direct push technology	Calibration data	# (ups)	Calibration method	Calibration		Cross-validation		Log10(K _{ups})			
				R ²	RMSE	R ²	RMSE	min	Max		
DPIIL	DPST	13	LM, OK, Upscaling	0.74	0.18	0.68	0.19	-4.1	-3.2		
HPT	DPST	10	LM, Upscaling	0.82	0.16	0.74	0.19	-4.1	-3.2		
CPT	OCK constant head & air permeameter tests	201	LM, BE _{Fr} correction, LM, Upscaling	Kh		0.43	0.79	0.29	0.9	-7.5	-2.8
				Kv		0.75	0.95	0.68	1.07	-12	-3.6

OCK = ordinary co-kriging ; ups = upscaled data; OK = ordinary kriging; LM = linear model; BE_{Fr} = Friction ratio boundary energy

Figure 1

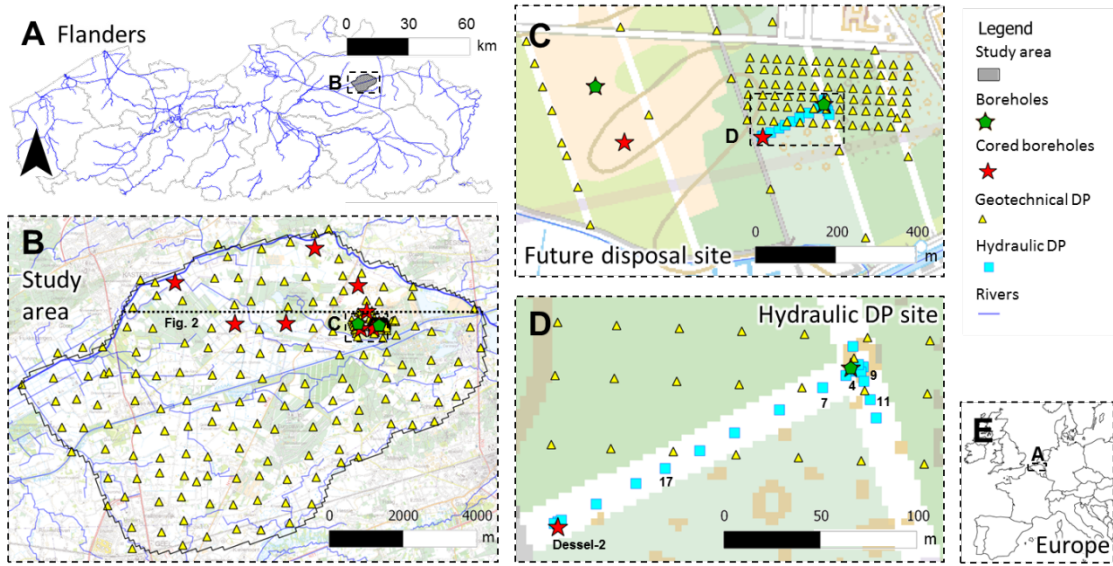


Figure 2

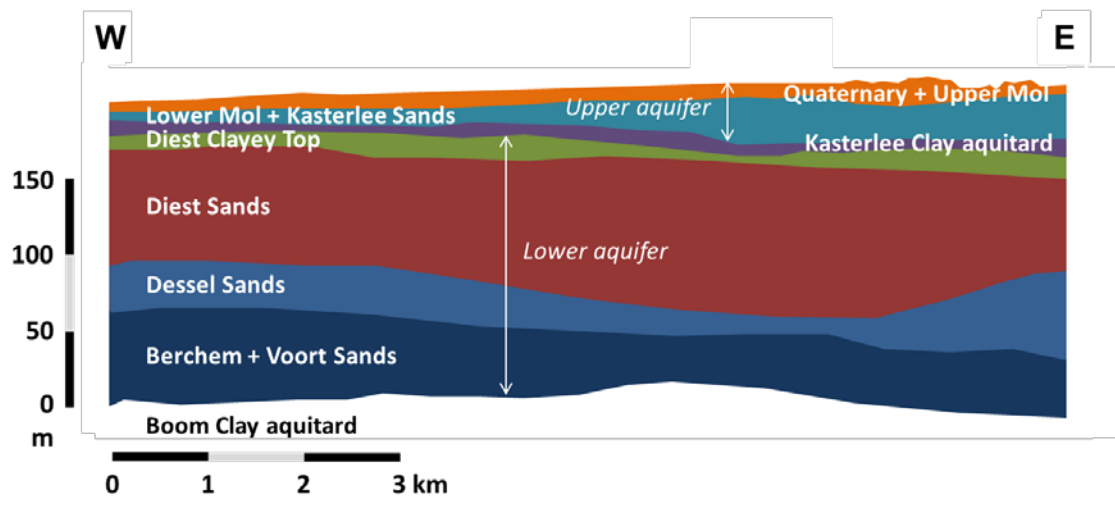


Figure 3

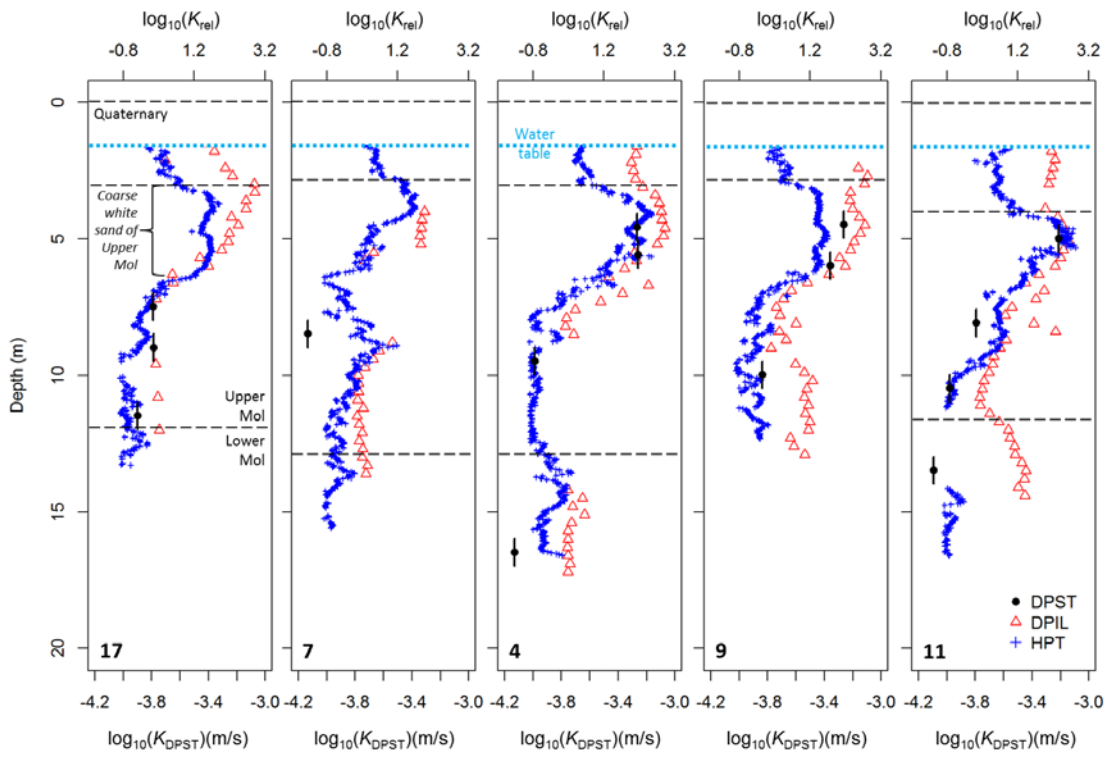


Figure 4

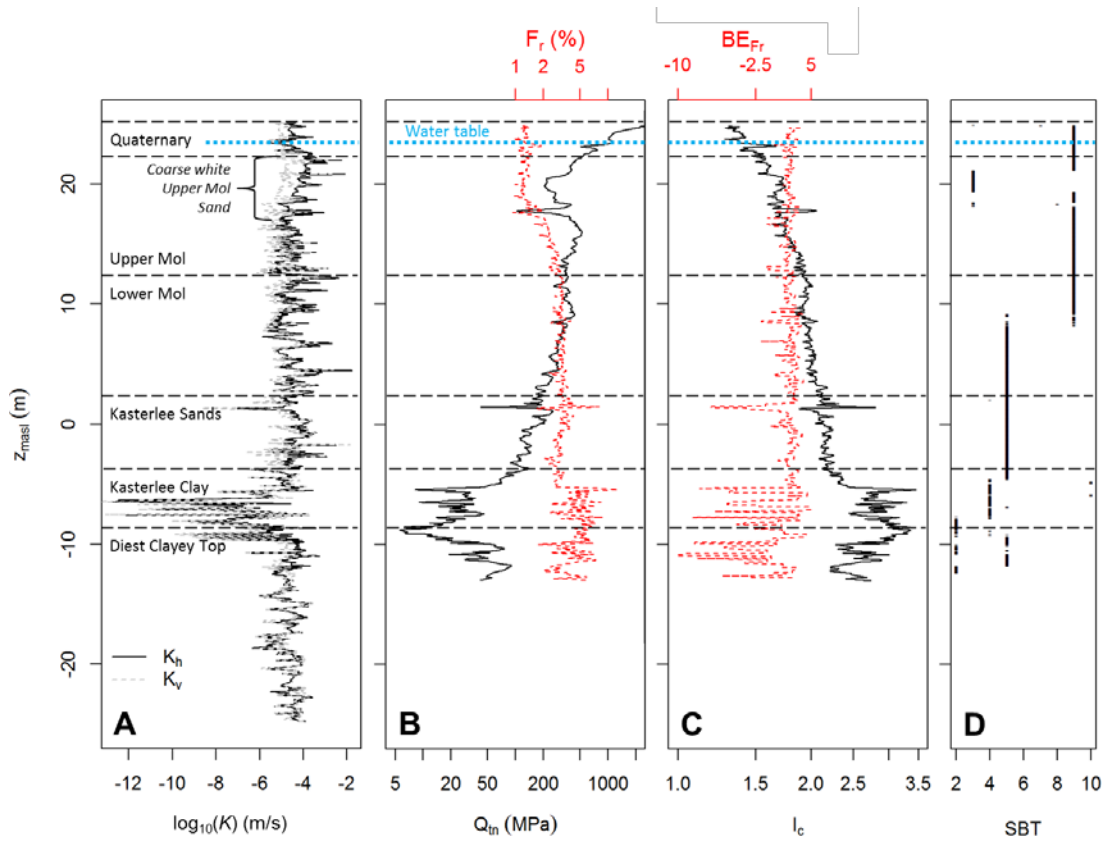


Figure 5

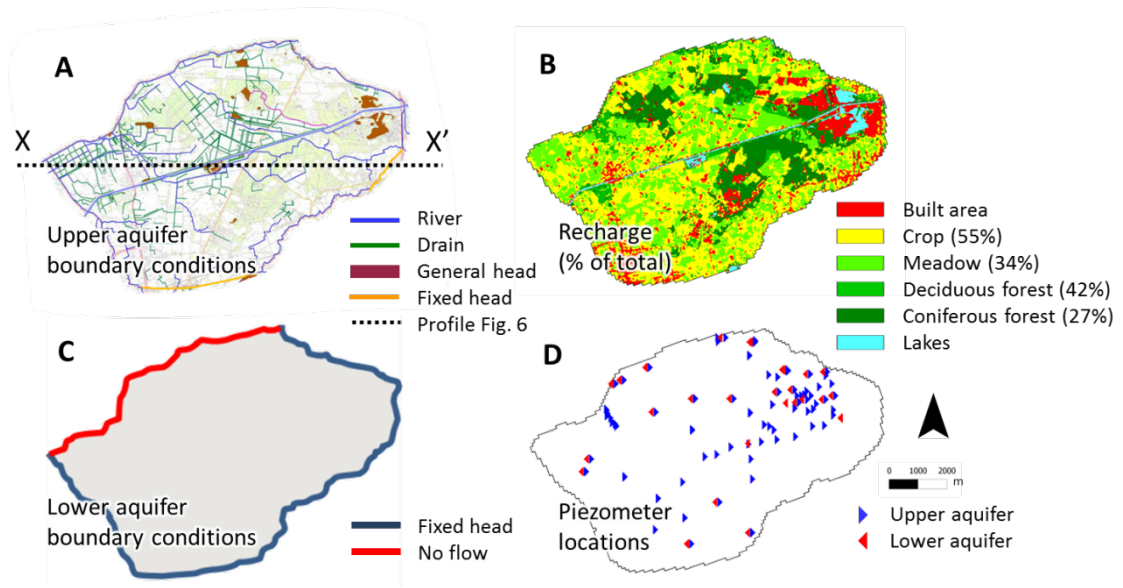


Figure 6

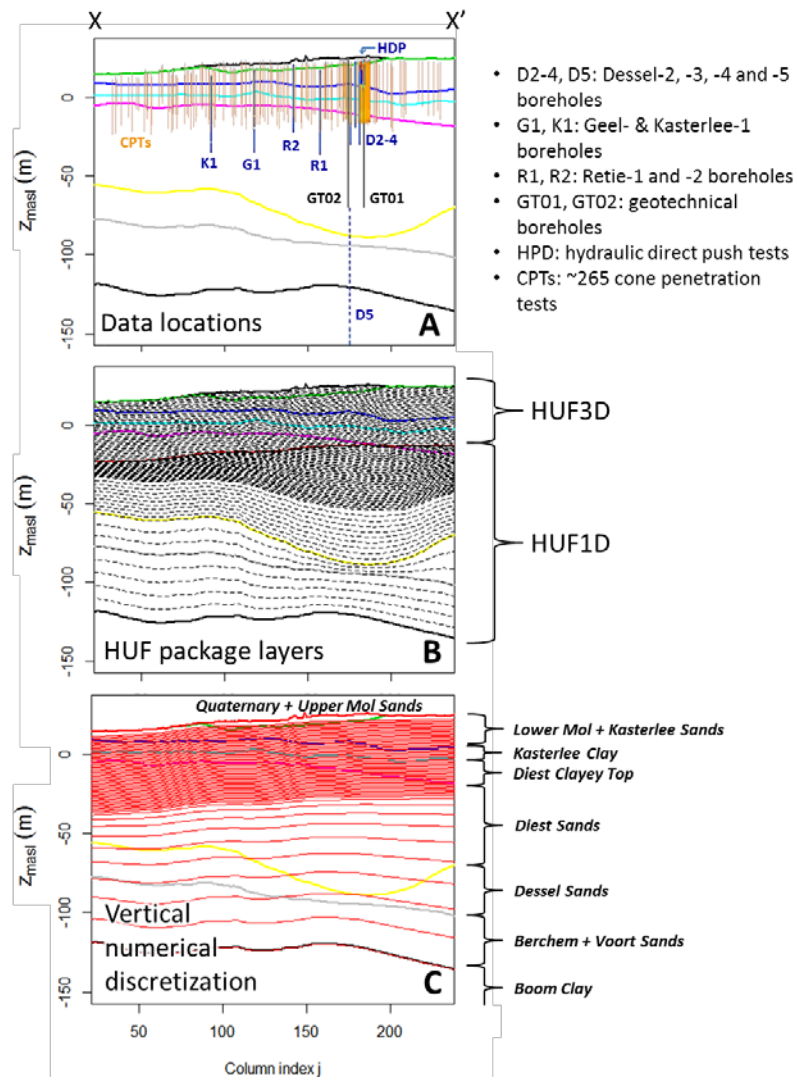


Figure 7

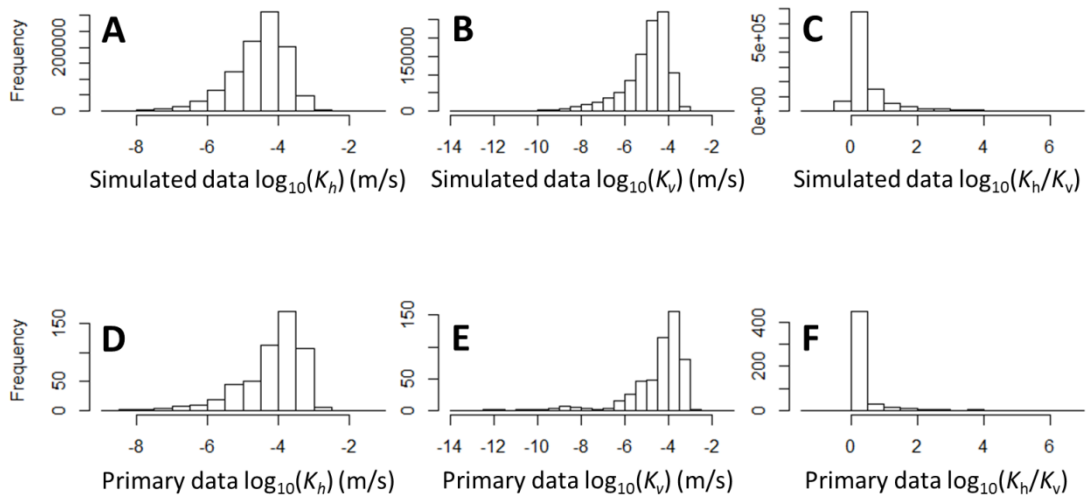


Figure 9

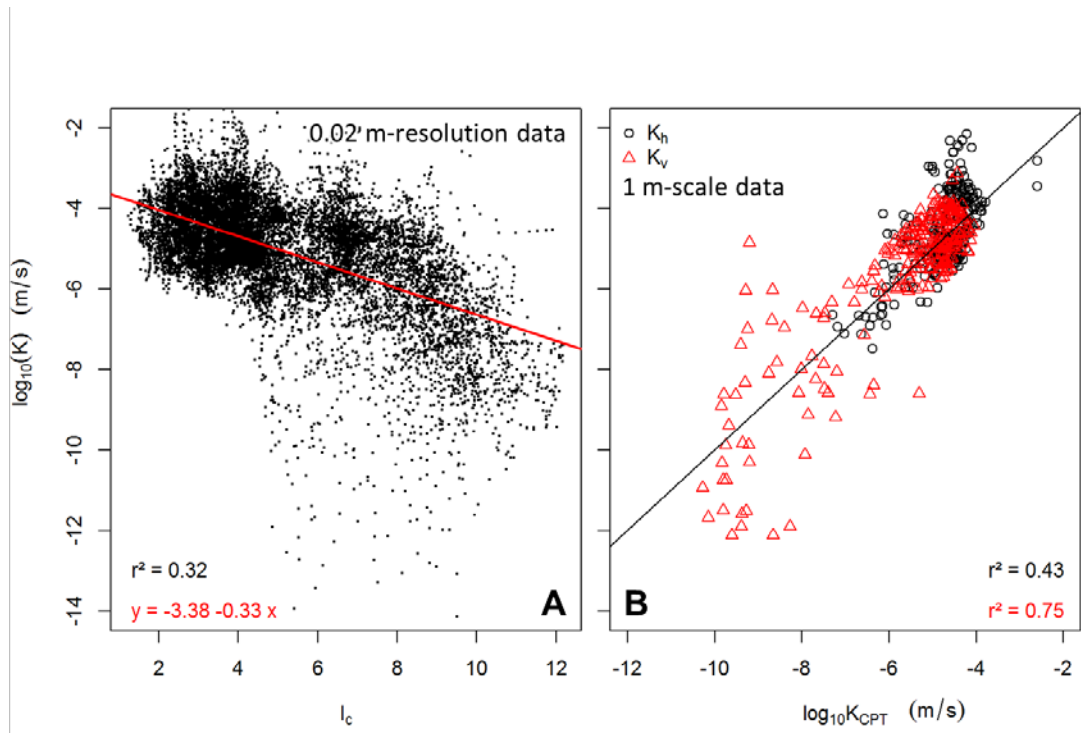


Figure 10

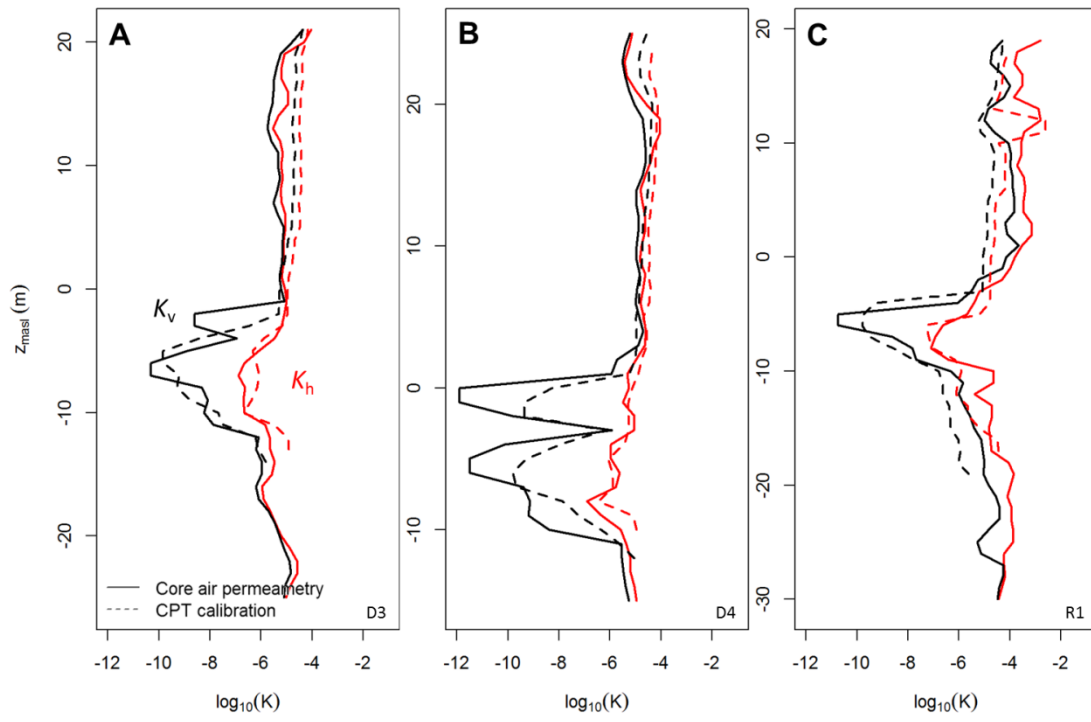


Figure 11

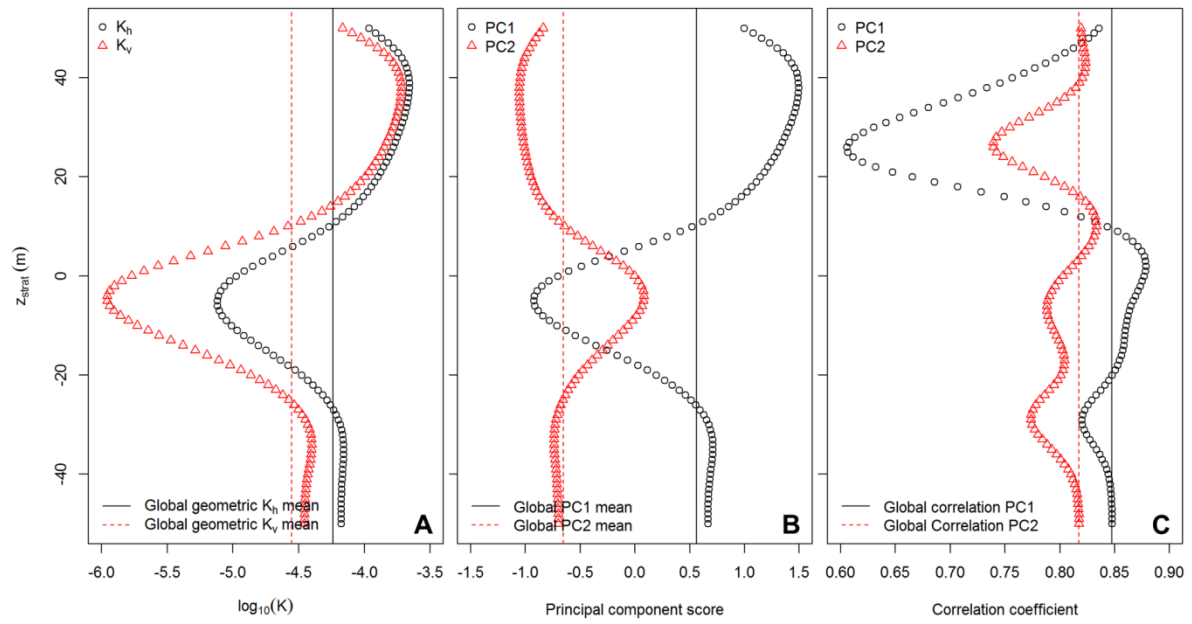


Figure 12

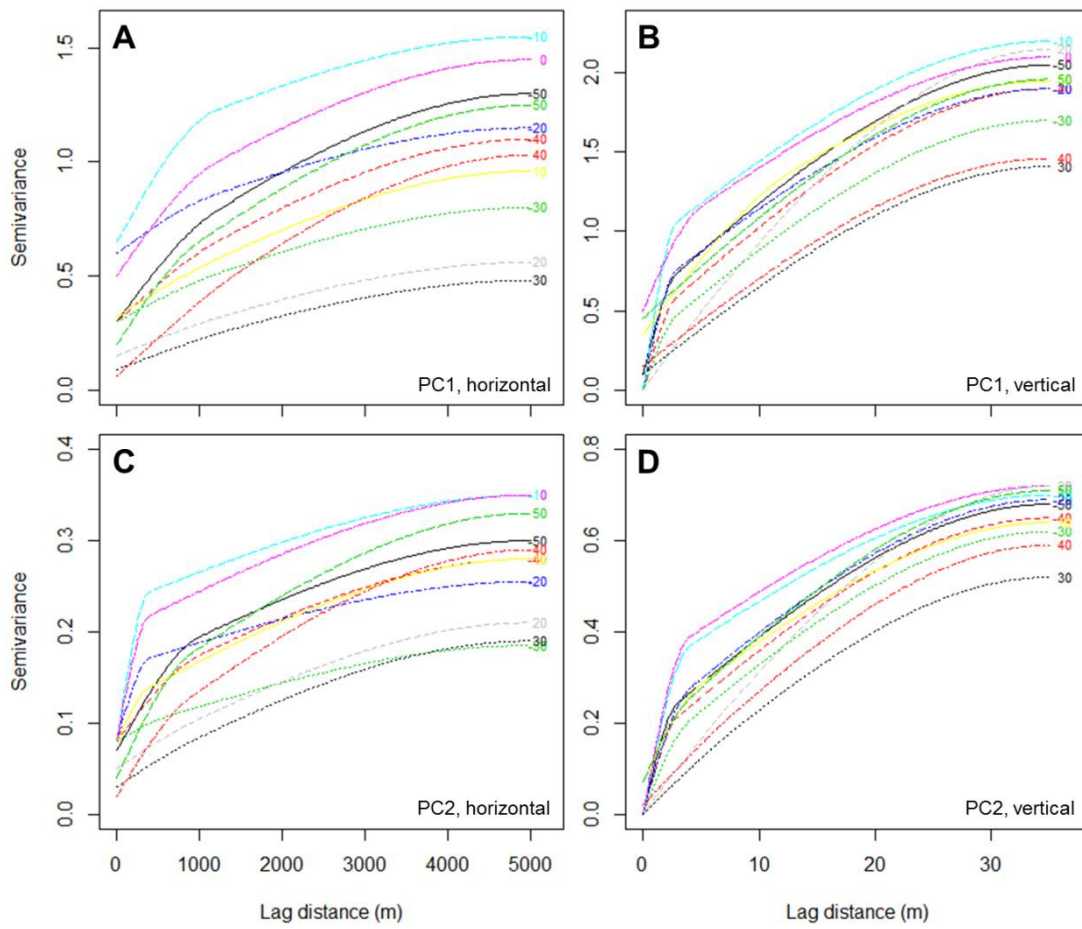


Figure 13

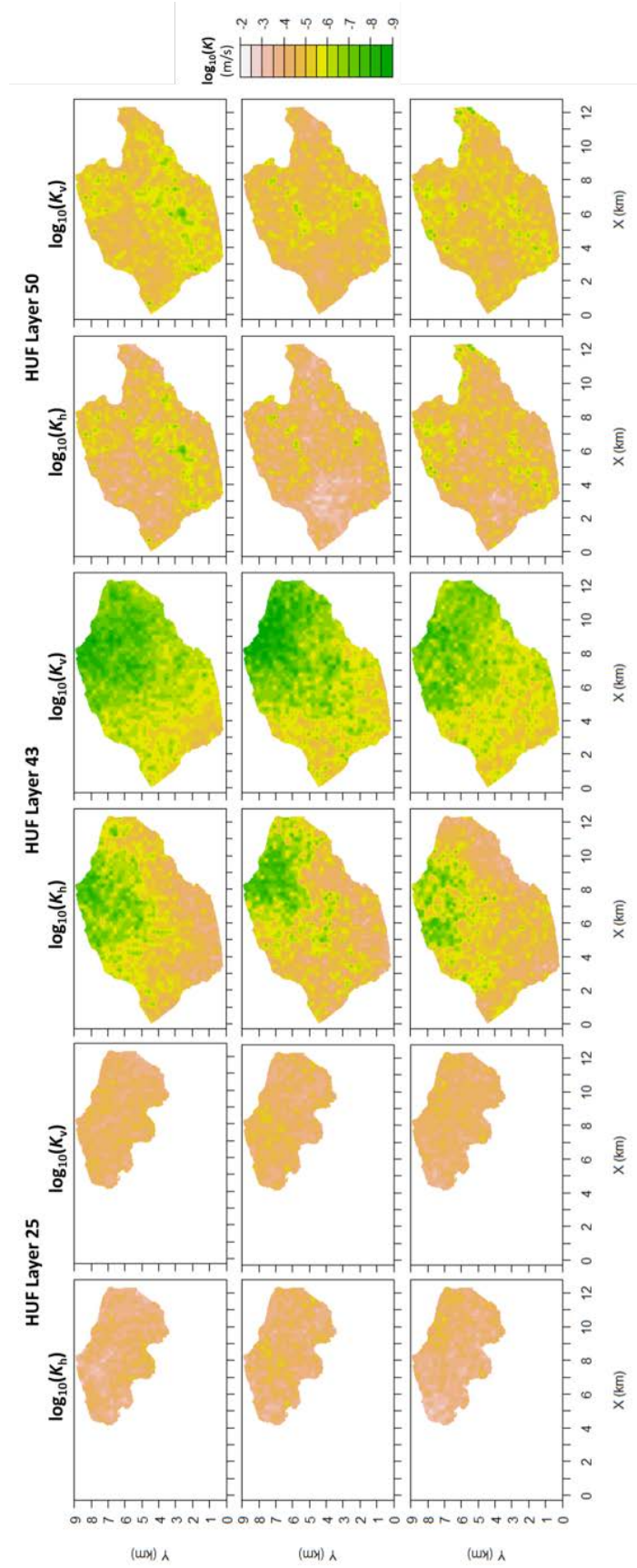


Figure 14

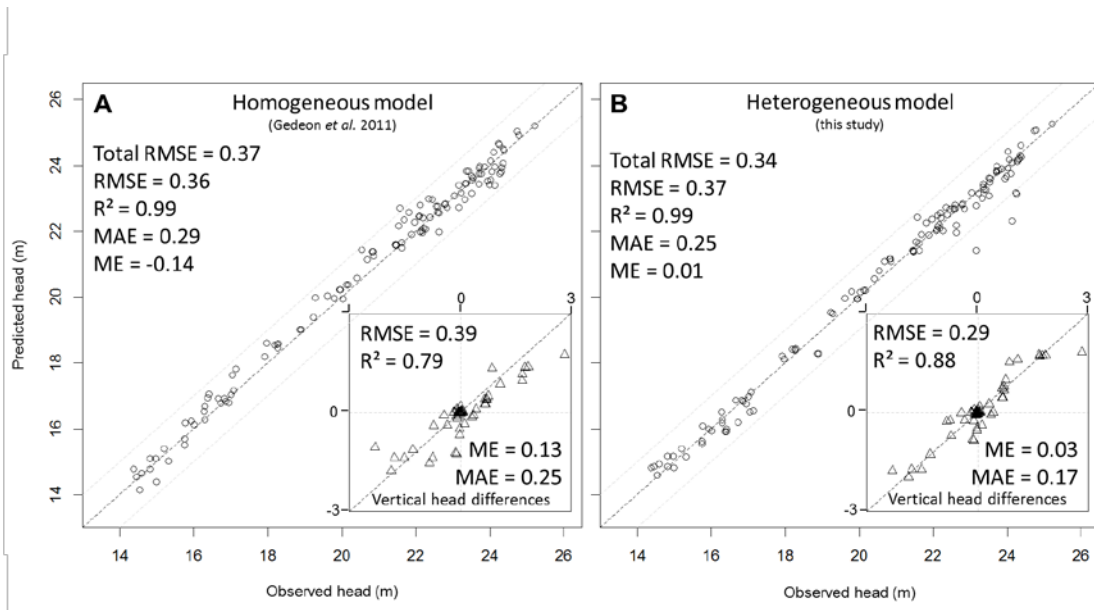


Figure 15

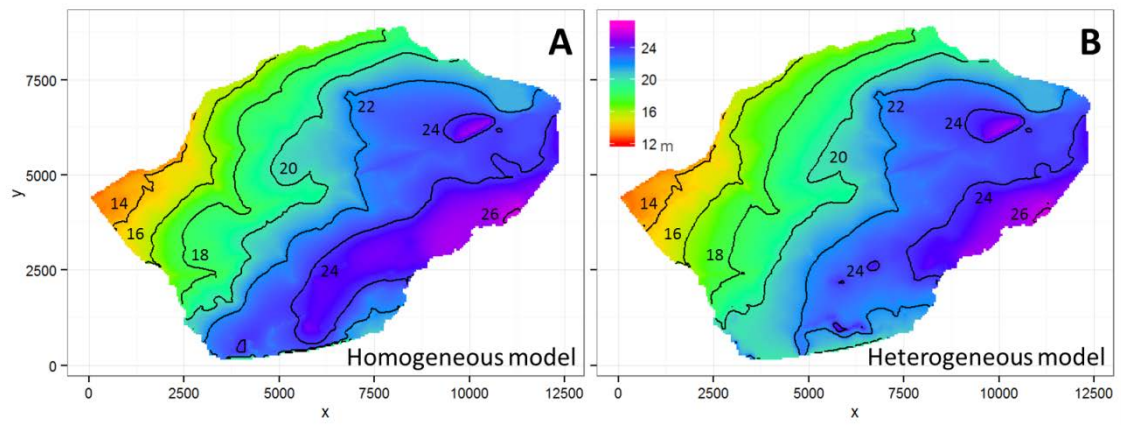


Figure 16

