



Analyzing the effects of geological and parameter uncertainty on prediction of groundwater head and travel time

X. He¹, T. O. Sonnenborg², F. Jørgensen², A.-S. Høyer^{2,3}, R. R. Møller², and K. H. Jensen¹

¹Department of Geosciences and Natural Resource Management, Copenhagen University, Copenhagen, Denmark

²Geological Survey of Denmark and Greenland (GEUS), Copenhagen, Denmark

³Department of Earth Sciences, Aarhus University, Aarhus, Denmark

Correspondence to: X. He (xih@geo.ku.dk)

Received: 14 February 2013 – Published in Hydrol. Earth Syst. Sci. Discuss.: 6 March 2013

Revised: 11 June 2013 – Accepted: 8 July 2013 – Published: 16 August 2013

Abstract. Uncertainty of groundwater model predictions has in the past mostly been related to uncertainty in the hydraulic parameters, whereas uncertainty in the geological structure has not been considered to the same extent. Recent developments in theoretical methods for quantifying geological uncertainty have made it possible to consider this factor in groundwater modeling. In this study we have applied the multiple-point geostatistical method (MPS) integrated in the Stanford Geostatistical Modeling Software (SGeMS) for exploring the impact of geological uncertainty on groundwater flow patterns for a site in Denmark. Realizations from the geostatistical model were used as input to a groundwater model developed from Modular three-dimensional finite-difference ground-water model (MODFLOW) within the Groundwater Modeling System (GMS) modeling environment. The uncertainty analysis was carried out in three scenarios involving simulation of groundwater head distribution and travel time. The first scenario implied 100 stochastic geological models all assigning the same hydraulic parameters for the same geological units. In the second scenario the same 100 geological models were subjected to model optimization, where the hydraulic parameters for each of them were estimated by calibration against observations of hydraulic head and stream discharge. In the third scenario each geological model was run with 216 randomized sets of parameters. The analysis documented that the uncertainty on the conceptual geological model was as significant as the uncertainty related to the embedded hydraulic parameters.

1 Introduction

With the prevalent application of groundwater modeling, the inherent uncertainties associated with model simulation are well acknowledged (Delhomme, 1979; Beven and Binley, 1992; Feyen et al., 2001; Hassan et al., 2008). Dettinger and Wilson (1981) divided uncertainty in groundwater systems into two classes: intrinsic uncertainty and information uncertainty. A subsurface property, such as hydraulic conductivity, is considered as intrinsic uncertainty due to its spatial stochastic variations. The spatial variation is affected by differences in large-scale parameter values between geological units superimposed by smaller-scale variation within the individual units. Webb and Anderson (1996), Journel and Alabert (1990), Carle et al. (1998) and Mariethoz et al. (2009) have suggested to incorporate these two contributions in a modeling framework by first defining the overall geological structure and then incorporating the smaller-scale spatial heterogeneity of the hydrogeological parameters within the units.

A common approach to simulate spatial heterogeneity in hydrogeology is to use geostatistics, and the traditional method is to employ variogram-based techniques (Delhomme, 1979; Wingle and Poeter, 1993; Johnson, 1995; Klise et al., 2009). Despite that these traditional methods have been applied extensively during the last three decades, they only consider correlation between two spatial locations, which often fails to depict distinct largely connected geological structures. Further, due to mathematical simplifications these methods can only capture a limited number of data types (Caers and Zhang, 2004; Journel, 2005). Renard (2007)

proposed three alternative techniques: the truncated pluri-Gaussian method, the continuous-lag Markov chain, and the multiple-point geostatistical approach (MPS), where the latter was viewed as the most appealing method. The advantages of MPS have been demonstrated by, e.g., Strebelle (2002), Caers and Zhang (2004), Liu et al. (2005), Journel and Zhang (2006), Liu (2006), and Strebelle (2006). Even though the ability of using MPS for 3-D simulations was acknowledged, only 2-D simulations were presented in these studies. Fundamental to the application of MPS is the definition of an appropriate training image (TI), which is a 2-D or 3-D image of the geological characteristics or patterns of the model area (Caers and Zhang, 2004). In the study of Lake Chad Basin, Le Coz et al. (2011) made a 3-D TI by replicating a single horizontal 2-D TI. However, they also realized that in this way the degree of similarity between two successive layers is maximized vertically, resulting in a very high nugget effect from 3-D simulations. Comunian et al. (2011) demonstrated a hierarchical multiple-point method with a complicated framework, which includes breaking large-scale structures into a number of subregions and application of 3-D MPS to each heterogeneous subregion. Overall, their three-dimensional TI was generated by object-based techniques for which the parameters were estimated manually by trial and error from 2-D observed sections. Honarkhah and Caers (2012) demonstrated a new MPS modeling paradigm with 3-D TI, but the data source and method of generating the 3-D TI was not mentioned. Multiple-point geostatistics is indeed appealing in stochastic hydrogeology, but so far the application to 3-D problems seems to be constrained by the difficulty of acquiring a full 3-D TI.

Parameter uncertainty is usually analyzed by applying Monte Carlo-based stochastic approaches, and the Generalized Likelihood Uncertainty Estimation (GLUE) (Beven and Binley, 1992) is one of the most popular methodologies due to its conceptual simplicity and ease of implementation. However, the method has been subject to criticism as the likelihood measure is “less formal” (Mantovan and Todini, 2006; Mantovan et al., 2007), the choice of likelihood measure is subjective and the updating process requires abundant observations (Feyen et al., 2001). Recently, several formal Bayesian methods for highly parameterized groundwater models have emerged (Hendricks Franssen et al., 2004; Tonkin and Doherty, 2005; Vrugt et al., 2008). These global search methods have the advantage of providing a more robust uncertainty analysis but they are also highly-CPU demanding and time-consuming.

Several studies have been reported in the literature on the impact of parameter and geological uncertainty on groundwater modeling. Freeze et al. (1990) were among the first to jointly consider both geological and parameter uncertainty. Abbaspou et al. (1998) proposed the uncertainty analysis framework BUDA (Bayesian Uncertainty Development Algorithm). Despite its ability to quantify uncertainty by conditioning on both hard and soft data, the approach for analyz-

ing uncertainty is based on two-point-based kriging. Feyen et al. (2001) presented a study on stochastic capture zone delineation by using the GLUE method. The stochastic hydraulic conductivity distribution was generated by sequential Gaussian simulation, but the distribution was still governed by a two-point covariance-based function. Harrar et al. (2003) also conducted flow model uncertainty analysis considering both geological and parameter uncertainty, but only two deterministic geological models were used. The two models were defined from maps of cross sections and geologists’ experience without applying any geostatistical method. Fleckenstein et al. (2006) used the transition probability-based geostatistic simulation TPROGS (Carle, 1999) to generate geological models. Each model was calibrated individually to elucidate the impact of hydrogeofacies attributes on river–aquifer exchange processes. However, only six models were analyzed since their attempt was not to carry out a full stochastic uncertainty analysis. Feyen and Caers (2006) presented a study in which the uncertainty on flow and transport modeling was quantified. For a relatively small two-dimensional synthetic fluvial system they generated 6500 geological realizations using MPS and 286 000 realizations of intrafacies hydraulic conductivity distribution using a bimodal distribution. In their analysis the effect of uncertainty on large-scale or effective parameters was not addressed. This factor may be a significant source of uncertainty; see, e.g., Refsgaard et al. (2012).

The objective of this study is to apply the multiple-point geostatistical approach to analyze the contribution from the two sources of uncertainty in groundwater modeling: geological model uncertainty and effective parameter uncertainty. The parameter uncertainty considered here is related to the estimation uncertainty of hydraulic conductivity of the considered geological units. These units are assumed homogeneous and not subject to intrafacies variability. We apply the MPS method to a real field system and use a full 3-D TI based on field measurements from an airborne geophysical campaign, which provides detailed 3-D data on the geological composition. Using such data for defining the TI introduces more field evidence in the geological characterization than in most other MPS studies, which is fundamental to the credibility of this particular stochastic approach.

2 Study area

The study area is situated near Ølgod in the western part of Jutland, Denmark, and covers an area of about 14.5 km by 13.9 km (Fig. 1). The regional topography is gently flat, ranging from 63.4 m above mean sea level (m a.s.l.) at Bavnsbjerg in the northwestern part of the area to 17.4 m a.s.l. along stream valleys. Seven streams originate in this region, which drains to Skjern River to the north and Varde River to the south. Groundwater is the main source of domestic and agricultural water supply in the area. The predominant land use

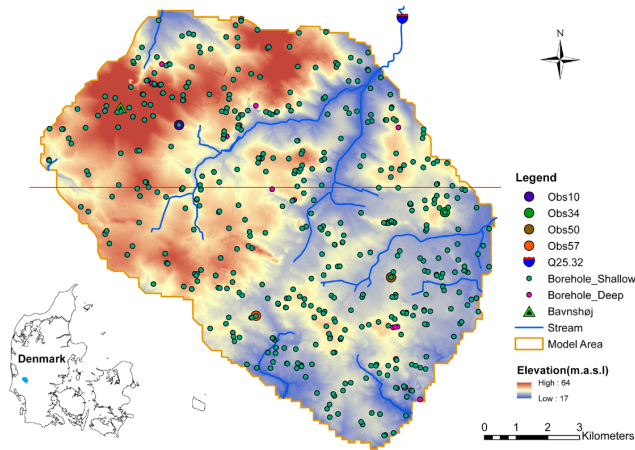


Fig. 1. Topography, stream network and observation stations around Ølgod. “Borehole.Shallow” are boreholes with bottom elevation higher than -70 m a.s.l., while “Borehole_Deep” are boreholes with bottom elevation lower than -70 m a.s.l. The red line indicates the location of the cross section in Fig. 2.

is agriculture, which is featured by a well-developed sub-surface drainage system. Weather in this area is influenced by both the North Sea and the European continent. The average annual temperature is 8.2°C , with a maximum of 16.5°C in August and a minimum of 1.4°C in January. The mean annual precipitation in the area is approximately 1050 mm yr^{-1} , with frequent intense rain in autumn and winter and less rain in spring (Stisen et al., 2011).

2.1 Geology

The lower boundary of the aquifers in the area is constituted by thick impermeable Paleogene clay deposited in hemipelagic environments and located at depths of 260–320 m (Høyer et al., 2013). The Paleogene clay is covered by Miocene clay, silt and sand deposits mainly originating from deltaic and shallow marine environments (Rasmussen et al., 2010). The thickness of these Miocene deposits increases from east to west in the study area and attains thicknesses of up to approximately 150 m. During the Pleistocene, phases of erosion, deposition and deformation resulted in a highly heterogeneous glaciotectionic complex above the Paleogene and Miocene. Buried tunnel valleys were deeply incised and subsequently filled, glacial and interglacial sediments were deposited and the entire setting was finally heavily deformed by one or more glaciers in the Late Pleistocene (Jørgensen and Sandersen, 2006). The resulting Pleistocene sequence is therefore highly variable in thickness and ranges from 0 to more than 300 m. Borehole data and seismic data collected in the study area confirm a highly heterogeneous setting of the Pleistocene sediments above -100 m a.s.l. (Høyer et al., 2011). Groundwater reservoirs in this region are often found

Table 1. Observations of groundwater head and river flow.

| Observations | Number of observations | Standard deviation | Observation period |
|--------------|------------------------|--------------------------------|--------------------|
| Head Obs. 1 | 68 | 2 m | 2009–2010 |
| Head Obs. 2 | 151 | 4 m | 1960–2009 |
| River Flow | 1 | $812\text{ m}^3\text{ d}^{-1}$ | 1990–2005 |

in the Miocene sand deposits or within the heterogeneous Pleistocene setting.

2.2 Data collection and processing

2.2.1 Geological and geophysical data

The three-dimensional geostatistical models developed in the study are based on multiple data sources, including borehole description, seismic data and SkyTEM (airborne transient electromagnetic method) data (Høyer et al., 2011). Borehole data are obtained from the Danish national geological database JUPITER (<http://www.geus.dk/jupiter/index-dk.htm>). JUPITER contains over 240 000 boreholes with information regarding geology, geography, hydrogeology, and groundwater chemistry. 525 boreholes with geological description are located within the study area; however, 96 % of them are shallow wells with bottom elevation above -70 m a.s.l. (Fig. 1), and only 22 boreholes have data deeper than -70 m a.s.l. In the subsequent analysis the geological descriptions from JUPITER have been simplified and categorized into four main geological units: Quaternary sand, Quaternary clay, pre-Quaternary sand and pre-Quaternary clay.

2.2.2 Construction of training image (TI)

The training image was indirectly constructed from SkyTEM data collected in the study area. The SkyTEM system (Sørensen and Auken, 2004) measures the electrical resistivity of the subsurface down to about -250 m, and this parameter can be utilized for determining the clay content if groundwater salinity does not change considerably across the survey area. The data were collected with a flight line spacing of only 125 to 270 m (Høyer et al., 2011), providing a dense data coverage. Data processing and inversion is described in Høyer et al. (2011). For the conversion from electrical resistivity to clay content a geostatistical estimation concept involving borehole information and SkyTEM data was developed (Jørgensen et al., 2012). This concept uses non-linear inversion to estimate clay content in all cells of a regular 3-D grid by optimizing a translator function for the conversion. The final result was a binary sand–clay model discretized in a regular 3-D grid covering the entire study area (Jørgensen et al., 2012).

2.2.3 Hydrological data

Two groups of hydraulic head data are available (Table 1). Group 1 includes 68 observations, which were obtained during the latest sampling campaigns in 2009 and 2010 (Alectia, 2011), and Group 2 includes 151 observations retrieved from the JUPITER database. Group 1 data were considered to be more accurate than Group 2 data, hence the standard deviation for the uncertainty of the two types of data was assumed 2 and 4 m, respectively.

Stream discharge is only available from one station, Q25.32, located slightly outside the study area in the north-eastern part (Fig. 1). The topographical catchment area to the station is around 64 km². Based on the historical daily data from 1990–2005, river discharge ranges from less than 0.1 m³ s⁻¹ during summer to more than 5.0 m³ s⁻¹ during late autumn, and the 16 yr daily average discharge is 0.8 m³ s⁻¹ (69 000 m³ d⁻¹).

Estimates of groundwater recharge are extracted from the Danish national water resources model, the DK-model (Henriksen et al., 2003). The net recharge (precipitation minus actual evapotranspiration) varies between 0 to 2.5 mm d⁻¹, and on average the net groundwater recharge in the model area is computed to 611 mm yr⁻¹.

Groundwater abstraction data are extracted from the JUPITER database. 165 pumping wells for domestic use, irrigation, fish farms and industry consumption are located in the study area, with a total abstraction of 3.2×10^6 m³ yr⁻¹ for the period 2000 to 2010.

3 Modeling methodology

3.1 Geostatistical simulation

An important part of this study is the generation of stochastic geological models using the multiple-point geostatistical approach (MPS) (Guardiano and Srivastava, 1993; Strebelle, 2002; Journel and Zhang, 2006), which has been integrated in the Stanford Geostatistical Modeling Software (SGeMS) (Remy et al., 2009). We apply the single normal equation simulation algorithm (SNESIM) (Strebelle, 2002) of MPS, which has reconciled the strengths of traditional object-based and pixel-based simulation algorithms (Liu et al., 2005). Object-based geostatistics is advantageous in capturing spatial pattern and structure but encounters difficulty when conditioned to large amounts of local data. In contrast, traditional pixel-based algorithms honor local data by reproducing a variogram or covariance model with two-point statistics, and in this process it eludes the geostatistical information concealed in large-scale patterns. The newly developed pixel-based algorithm, SNESIM, maintains the flexibility of data conditioning and directly collects the probability distribution from a training image (TI) with multiple points, hereby overcoming the limitation of two-point geostatistics.

Geostatistical simulations primarily depend on the conditional probability distribution function (cpdf). In SNESIM, the cpdf is set equal to the corresponding training image proportions (Strebelle, 2002). This training proportion is obtained by scanning one or several training images (TI) with a multiple-node training template. The classical sequential simulation paradigm (Goovaerts, 1997, p. 376) is used to draw the stochastic images. The main step is to first assign hard data to the closest node of the simulation grid, while all the unknown nodes are visited once and only once by a random path. At each unknown node, all the surrounding hard data presented within a certain search template are retained as conditioning data events, with which the corresponding conditional probability is computed.

The crucial requisite of applying SNESIM is to provide a TI which represents the geological characteristics of the study area. As the TI is the primary source of uncertainty when using MPS, an important step is to obtain the appropriate TI (Comunian et al., 2011). Previous applications of MPS have mostly been constrained to 2-D TIs (Caers, 2001; Strebelle, 2002; Comunian et al., 2012). In this study we use the 3-D training image mainly based on the detailed geophysical data collected during the flight campaign (Jørgensen et al., 2012).

3.2 Groundwater and optimization models

A steady-state groundwater flow model was constructed using the groundwater modeling computer code MODFLOW-2000 (Harbaugh et al., 2000) within the framework of the Groundwater Modeling System (GMS). To enable an accurate representation of the geological heterogeneity in the stochastic realizations, the Hydrogeologic-Unit Flow (HUF) package (Evan and Mary, 2000) was used. The HUF package allows the vertical stratigraphy to be defined independently of the numerical model layers by using hydrogeologic units. For the groundwater flow process, the HUF package computes the hydraulic properties of the model grid according to the hydrogeologic units within the model grid. The cell-to-cell flow conductance in horizontal and vertical directions is calculated as the arithmetic and harmonic mean, respectively, as discussed by McDonald and Harbaugh (1988). Other activated packages are specified head, recharge (RCH), river (RIV), drain (DRT) and well (WEL).

Although several advanced inverse modeling methods based on global search algorithms have been presented recently, they are mostly quite time-consuming. Since we only focus on hydraulic conductivities for steady-state flow conditions, advanced algorithms designed for multi-objective optimizations are not required. Instead, we use the Model-independent parameter estimation & uncertainty analysis (PEST) inversion code (Doherty, 2005) for parameter estimation and for parameter sensitivity assessment. PEST uses a local search algorithm and has the advantage of fast convergence. Additionally, besides estimating the optimum

parameter (Brauchler et al., 2007; Cheng and Chen, 2007; Baalousha, 2012), we also considered the equifinality of different parameter sets by sampling within parameters' distribution space.

Based on the flow solutions generated by MODFLOW, the particle-tracking post-processing code MODPATH (David, 1994) was used to simulate particle travel time distributions. The backward particle tracking option simulates the pathways of particles from a certain cell to the possible corresponding groundwater recharge points reversely along the flow direction. By sampling the group of particles in a cell, one can compute the probability distribution of travel time using basic statistical analysis.

3.3 Uncertainty analysis

The uncertainty associated with geological structure was analyzed by utilizing 100 stochastic geological realizations generated by SNESIM as input to MODFLOW. To distinguish between parameter and geological uncertainty, three scenarios were defined:

Scenario 1: GeoModel-I: In the first scenario fixed hydraulic parameters were specified for the individual geological units for all MODFLOW models. The parameter values were taken as the median of the values obtained by calibration of each individual model.

Scenario 2: GeoModel-II: In this scenario each model was calibrated by PEST. The parameters included in the optimization were horizontal hydraulic conductivity of Quaternary sand, Quaternary clay and pre-Quaternary sand.

Scenario 3: GeoParModel: In the third scenario parameter uncertainty was added. Each MODFLOW model based on GeoModel-II was subject to a stochastic parameter analysis. Only the uncertainty of the hydraulic conductivity of Quaternary sand, Quaternary clay and pre-Quaternary sand was considered, since the sensitivity analysis documented that the groundwater flow simulations were mostly sensitive to these parameters.

In the analysis, the uncertainty of hydraulic conductivity was assumed to be log-normally distributed (Freeze, 1975; Hoeksema and Kitanidis, 1985). The variation of hydraulic conductivity between the realizations was generated based on the optimization results, where the optimized values represent the mean values and the standard deviations are derived from the 95 % confidence intervals (CI) estimated by PEST.

Random sampling in the parameter space was carried out using the Latin hypercube approach (McKay et al., 1979). Latin Hypercube Sampling (LHS) is a stratified sampling method where the parameter space S is divided into N segments of equal marginal probability $1/N$ and sampled values from different parameters are randomly combined. In

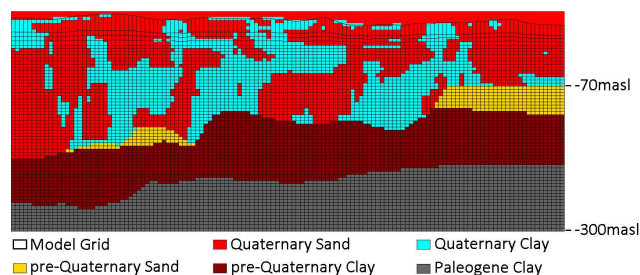


Fig. 2. Cross section with model grid and training image geological structure. Location on plain view is shown in Fig. 1. Vertical exaggeration is 15.

this way LHS captures the full parameter space in a simplified manner and requires much fewer model runs compared to normal random sampling. In our study, the distribution space for each of the three hydraulic parameters was divided into six segments of equal probability, and therefore the total number of simulations for each geological realization is $6^3 = 216$. We refer to the set of simulations for each geological realization based on sampling in the parameter space as ParModel, while the ensemble of simulations for all geological realizations and all parameter variations (90×216 simulations in total) is referred to as GeoParModel.

For all scenarios backward particle tracking was applied to four selected wells using MODPATH.

4 Groundwater model setup

4.1 Geological structure

The groundwater model extends from land surface to -300 m a.s.l., where the Paleogene clay occupies the whole area and is therefore taken as lower boundary of the model (Fig. 2). The corresponding geological setup consists of two parts. The upper part includes layers from the land surface to -70 m a.s.l. This part contains only heterogeneous Quaternary sediments and has abundant borehole data for conditioning simulation; hence this part is subject to the geostatistical simulations by SNESIM on a $100 \text{ m} \times 100 \text{ m} \times 5 \text{ m}$ grid. The lower part is generally dominated by comparably more homogeneous pre-Quaternary sediments. The geological structure of the pre-Quaternary sediments is described by a manual interpretation of mainly seismic data (Høyer et al., 2011; Jørgensen et al., 2012), since only few boreholes reach this deeper part and the SkyTEM data show limited resolution capability here (Høyer et al., 2011). At the places where Quaternary sediments are located between the pre-Quaternary surface and -70 m a.s.l., the geological model is defined by the SkyTEM-based training image (Fig. 2).

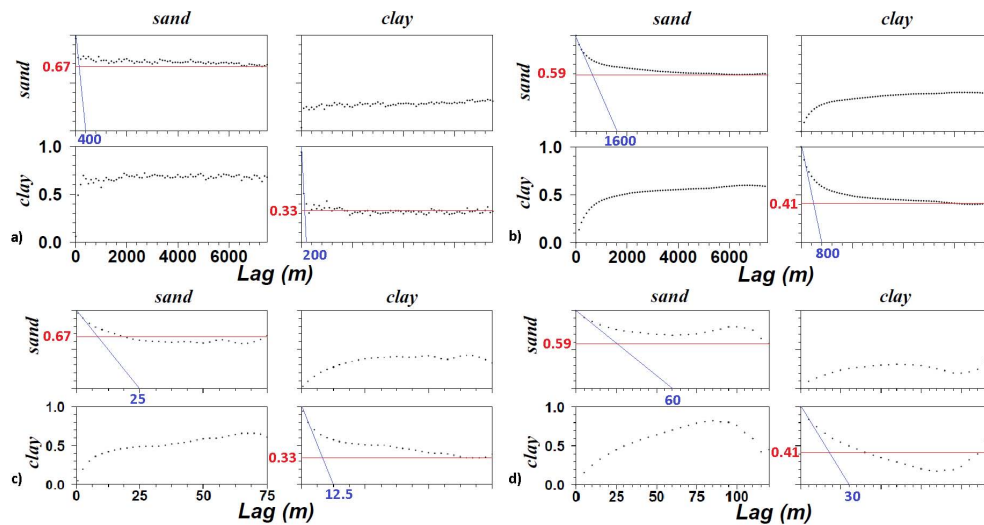


Fig. 3. Transition probability of Quaternary sediments. (a) Borehole data in lateral direction; (b) data from 3-D training image (TI) in lateral direction; (c) borehole data in vertical direction; (d) data from 3-D TI in vertical direction.

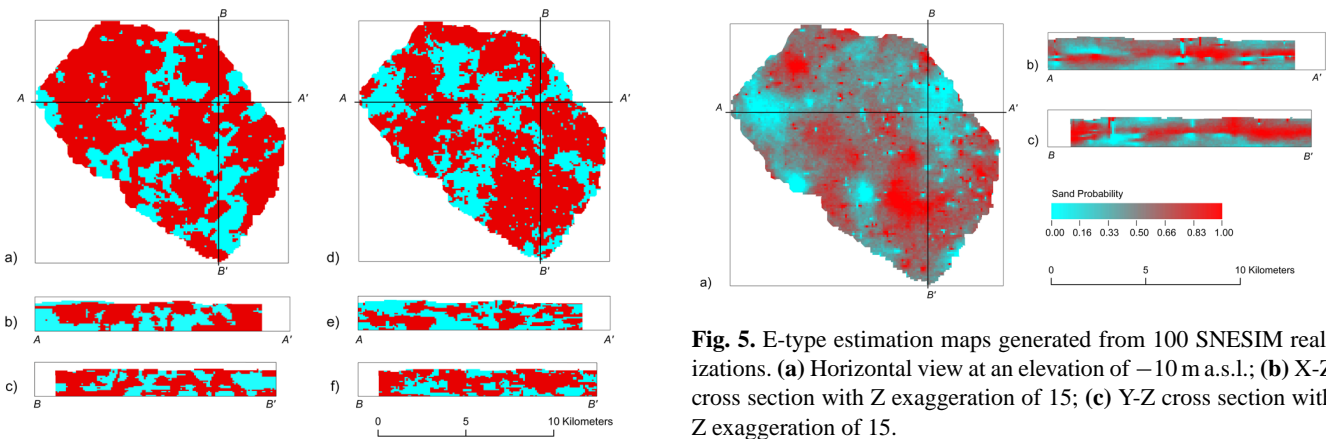


Fig. 4. 3-D training image (left) and an example of a geological realization (right). Red and blue indicate Quaternary sand and clay, respectively. (a) Horizontal view of 3-D training image at elevation of -10 m a.s.l.; (b) X-Z cross section of 3-D TI with Z exaggeration of 15; (c) Y-Z cross section of 3-D TI with Z exaggeration of 15; (d) horizontal view of realization at elevation of -10 m a.s.l.; (e) X-Z cross section of realization with Z exaggeration of 15; (f) Y-Z cross section of realization with Z exaggeration of 15.

4.2 Model discretization

The geological model was imported to the groundwater model using the HUF package. As discussed above the HUF package computes the hydraulic properties of a numerical grid by averaging hydraulic parameters of all hydrogeologic units present in that cell. Therefore, in order to avoid extensive averaging of geological properties from the geostatistical simulations, the numerical groundwater model grid was

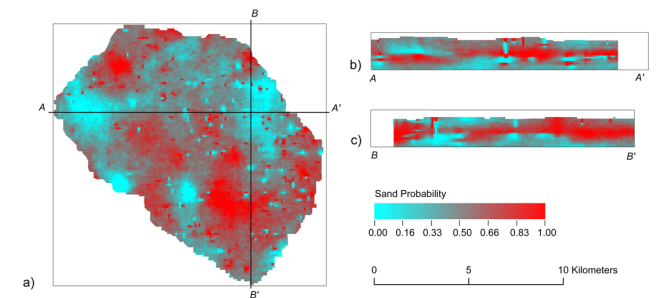


Fig. 5. E-type estimation maps generated from 100 SNESIM realizations. (a) Horizontal view at an elevation of -10 m a.s.l.; (b) X-Z cross section with Z exaggeration of 15; (c) Y-Z cross section with Z exaggeration of 15.

created to match the discretization of the geological model in nearly all layers expect for the upper five layers (Fig. 2).

The horizontal discretization was specified to $100\text{ m} \times 100\text{ m}$ while 63 layers were used in vertical direction. A cross section showing the vertical discretization is presented in Fig. 2. To avoid the problem of dry cells in the numerical simulations, the top layer was set to be thicker than any other layer; on average the depth of the top layer was 13 m. Average thickness of layers 2 to 5 was 4 m, while layers 6 to 63 had thicknesses of 5 m each in correspondence with the geological model. In total there were 792 603 active cells in the MODFLOW setup.

4.3 Boundary conditions

No natural hydrological boundaries could be identified for the groundwater model. Instead, the available observations of groundwater head were used to interpolate the groundwater

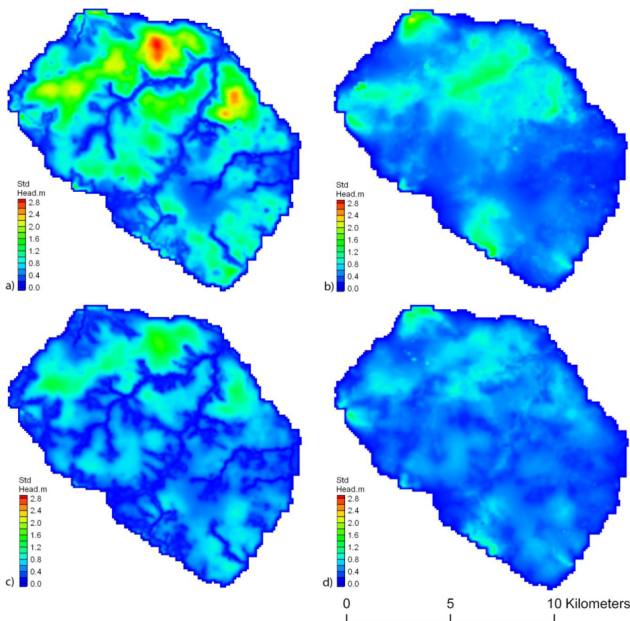


Fig. 6. Standard deviation of groundwater head distribution derived from 90 GeoModel-I models (first row) and from 90 GeoModel-II models (second row). (a) and (c) 1st layer, (b) and (d) 4th layer.

head variation along the boundary. The interpolated heads were specified as fixed head boundaries to the MODFLOW model, and the lower boundary was assigned as no-flow boundary.

4.4 Sources and sinks

Groundwater recharge from the DK-model (Henriksen et al., 2003) was used as input to the local model. The grid size of the DK-model is $500\text{ m} \times 500\text{ m}$. In order to fit the model, which has a discretization of $100\text{ m} \times 100\text{ m}$, the recharge data were downscaled by linear interpolation.

Large parts of the model area are drained by subsurface tile drains (Henriksen et al., 2003). Therefore, drains were specified in the entire model area using the drain package (DRT) (Harbaugh et al., 2000). Drain elevation was taken as 1 m below land surface, and the drain time constant was set to $1.7 \times 10^{-2}\text{ s}^{-1}$ in accordance with the DK-model (Henriksen et al., 2003). The river package (RIV) (Harbaugh et al., 2000) in MODFLOW was used to simulate the discharge in streams. The river network was determined from Geographic information system (GIS) data, while the river level was taken from the national digital elevation model with a resolution of 1.6 m, and the river bottom elevation was assumed to be 1 m below river level. The river conductance was set to 16.8 m d^{-1} according to the DK-model (Henriksen et al., 2003). River discharge station Q25.32 was used for model calibration. The average discharge for the period 1990 to 2005 was $69\,000\text{ m}^3\text{ d}^{-1}$. Since only part of the catchment area to the discharge station was included in the model,

Table 2. Hydraulic conductivity of the geological units. Values for Quaternary sand, Quaternary clay and pre-Quaternary sand are found as the medians of 90 estimated parameter values from PEST and used in GeoModel-I. Values for pre-Quaternary clay and Paleogene clay are inferred from the DK-model.

| Geological unit | Hydraulic conductivity (K) (m d^{-1}) | Log10 (K) | Standard deviation of Log10 (K) |
|---------------------|--|-----------|---------------------------------|
| Quaternary sand | 2.0 | 0.3 | 0.1 |
| Quaternary clay | 0.09 | -1.0 | 0.5 |
| Pre-Quaternary sand | 4.6 | 0.7 | 0.4 |
| Pre-Quaternary clay | 0.005 | -5.3 | - |
| Paleogene clay | 8.6×10^{-4} | -7.1 | - |

the discharge data were downscaled according to the model area, leading to a value of $59\,900\text{ m}^3\text{ d}^{-1}$. The well package (WEL) (Harbaugh et al., 2000) was used to simulate groundwater abstraction with a total pumping rate of $8900\text{ m}^3\text{ d}^{-1}$.

5 Results and discussion

5.1 Geostatistical analysis and geological realizations

The geostatistical analysis presented in this section refers to Quaternary sediments from the land surface to -70 m a.s.l. The primary geostatistical information required in SNESIM is the radius of the search template and target proportion, which is related to the mean length, and the proportion of each sedimentary unit. These values can be inferred from transition probabilities, which represent the spatial variability and structure of geological units in terms of conditional probabilities of occurrence. The transition probability model has the ability to quantitatively translate concepts and subjective observations into a spatial variability model with information of categorical variables' proportion, mean length and juxtapositional tendency (Carle and Fogg, 1996). The "sill" of a transition probability curve implies the proportion of the category, while the distance where the slope line intersects with the lag axis corresponds to the mean length of the category.

Figure 3a shows the lateral transition probabilities of the borehole data. The mean lengths of the two categories can be inferred from the intersections of the slope lines at lag distance 0 with the x axis. The mean length of Quaternary sand bodies is estimated to be around 400 m in lateral direction, while that of Quaternary clay bodies is around 200 m. Based on borehole data the proportions of Quaternary sand and clay are around 67 and 33 %, respectively, which is confirmed by Fig. 3a.

Figure 3b shows the lateral transition probability based on the 3-D training image. The estimated mean lengths of Quaternary sand and clay bodies are around 1600 and 800 m, respectively, and the proportions of these units are 59 and

Table 3. Travel time based on backward particle tracking for four wells for GeoModel-I, GeoModel-II, GeoParModel, and median of ParModels. *R* denotes the fraction of the 95 % confidence interval to the one from GeoParModel. Location of the four wells is shown in Fig. 8.

| | | W1 | W2 | W3 | W4 |
|------------------------|----------------------|-------|-------|-------|-------|
| GeoModel-I | Median (year) | 50 | 29 | 46 | 31 |
| | 95 % Interval (year) | 29 | 34 | 171 | 14 |
| | <i>R</i> | 91 % | 83 % | 87 % | 70 % |
| | Skewness | 0.3 | 1.6 | 2.3 | 6.0 |
| GeoModel-II | Median (year) | 51 | 31 | 47 | 31 |
| | 95 % Interval (year) | 33 | 41 | 282 | 36 |
| | <i>R</i> | 103 % | 98 % | 143 % | 183 % |
| | Skewness | 0.7 | 1.2 | 3.5 | 3.8 |
| GeoParModel | Median (year) | 52 | 31 | 47 | 31 |
| | 95 % Interval (year) | 32 | 41 | 196 | 19 |
| | <i>R</i> | 100 % | 100 % | 100 % | 100 % |
| | Skewness | 0.8 | 1.3 | 10.5 | 8.1 |
| Median of 90 ParModels | Median (year) | 52 | 31 | 48 | 31 |
| | 95 % Interval (year) | 6 | 5 | 21 | 5 |
| | <i>R</i> | 18 % | 12 % | 11 % | 26 % |
| | Skewness | 1.1 | 0.4 | -0.4 | -0.1 |

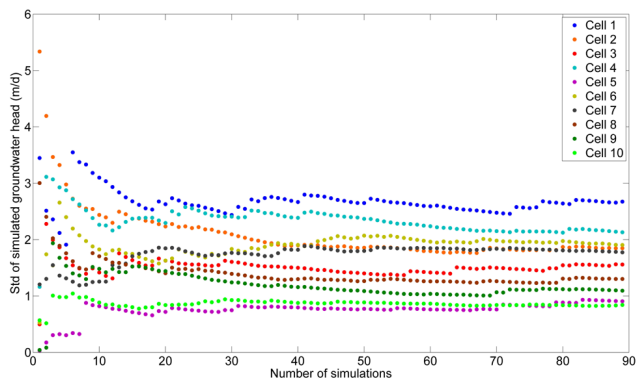


Fig. 7. Standard deviation of simulated groundwater head at selected cells (GeoModel-I).

41 %. The diagonal graphs also indicate similar information, but the proportions are not strictly the same. It should be noticed that the relationship between proportions and the “sill” of the diagram is not rigorous, especially when the diagram is derived from raw data. Consider the expression of transition probability (Carle and Fogg, 1996):

$$t_{jk(x,h)} = \{\text{Pr category } k \text{ occurs at } (x+h) | \text{category } j \text{ occurs at } x\}. \quad (1)$$

When *h* approaches infinity, $t_{jk(x,h)}$ equals the proportion of category *k* for all categories *j*. Therefore, for the limited lag distance, this proportion–sill relationship is more a guide to fit the sill of a spatial variability model than an indicator of category proportion from sill of raw data.

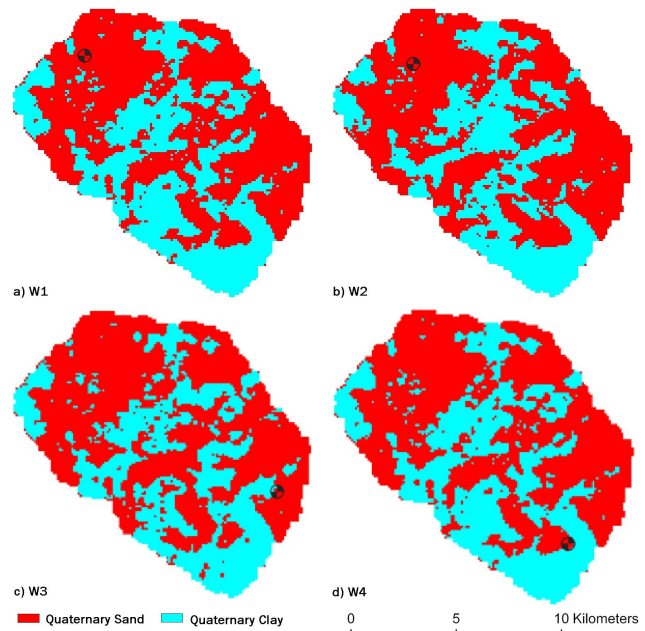


Fig. 8. Location of four wells for backward particle tracking (background shows geological structure of the training image at the depth of the well screen): (a) W1, well screen at 3.66 to -4.96 m a.s.l.; (b) W2, well screen at -4.98 to -10 m a.s.l.; (c) W3, well screen at 0.44 to -4.91 m a.s.l.; (d) W4, well screen at -4.93 to -10 m a.s.l.

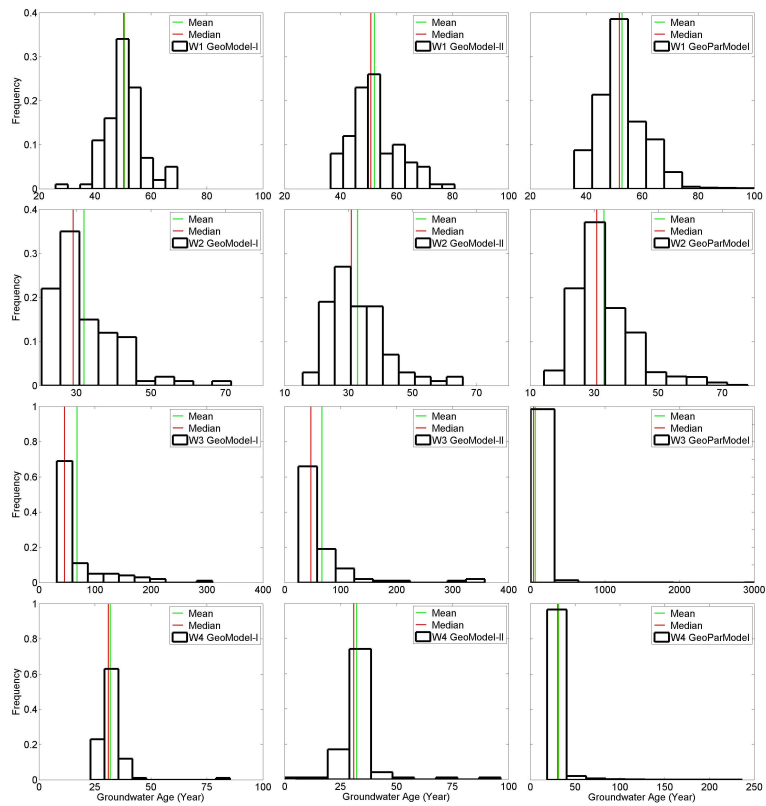


Fig. 9. Histogram of travel time of four wells for GeoModel-I (left column), GeoModel-II (middle column) and GeoParModel (right column). The green line indicates the mean value and the red line represents the median value.

Table 4. Simulated heads for four observation points for GeoModel-I, GeoModel-II, GeoParModel, and the median of 90 ParModels. R denotes the fraction of the 95 % confidence interval to the one from GeoParModel. Location of the four points is shown in Fig. 1.

| | | Obs 10 | Obs 34 | Obs 50 | Obs 57 |
|------------------------|-------------------|--------|--------|--------|--------|
| GeoModel-I | Median (m) | 36.4 | 26.6 | 27.9 | 30.0 |
| | 95 % Interval (m) | 2.5 | 2.3 | 1.0 | 1.7 |
| | R | 114 % | 70 % | 45 % | 121 % |
| GeoModel-II | Median (m) | 38.7 | 28.5 | 29.7 | 30.8 |
| | 95 % Interval (m) | 1.3 | 1.7 | 0.9 | 1.2 |
| | R | 59 % | 52 % | 41 % | 86 % |
| GeoParModel | Median (m) | 38.8 | 28.4 | 29.6 | 30.8 |
| | 95 % Interval (m) | 2.2 | 3.3 | 2.2 | 1.4 |
| | R | 100 % | 100 % | 100 % | 100 % |
| Median of 90 ParModels | Median (m) | 38.8 | 28.4 | 29.6 | 30.8 |
| | 95 % Interval (m) | 1.8 | 2.5 | 1.5 | 0.7 |
| | R | 82 % | 76 % | 68 % | 50 % |

Figure 3c and d show the vertical transition probability based on borehole data and TI, respectively. The mean length of Quaternary sand bodies is around 25 m as indicated by borehole data and around 60 m from 3-D TI data. For Quaternary clay bodies, the mean length is 12.5 and 30 m, correspondingly. Hence, both sources of data show that the mean

length of Quaternary sand bodies is twice the corresponding mean length of Quaternary clay, in both lateral and vertical directions.

Strebelle (2002) recommends that the global proportions of the training image should be similar to the desired proportion in the final model. In our case, borehole data show 67 %

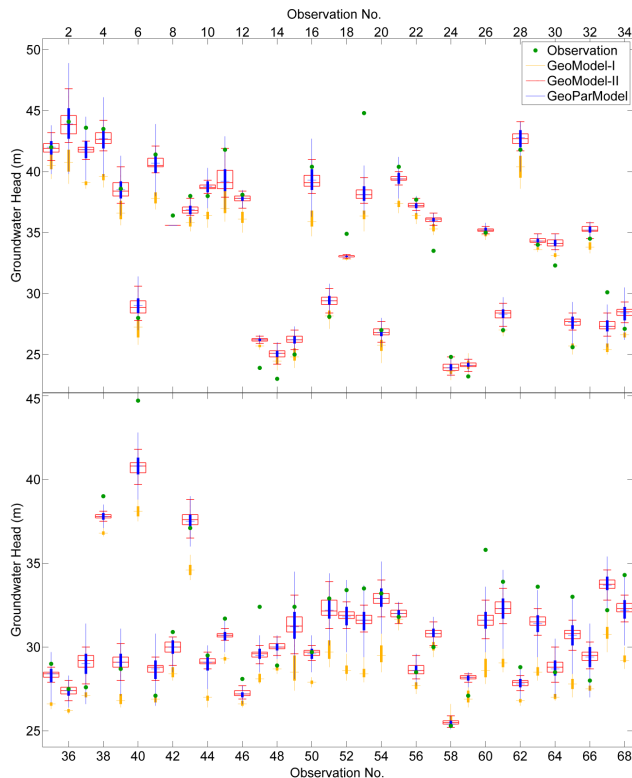


Fig. 10. Boxplot of simulated hydraulic heads against 68 observations from head obs. group 1. Green dots are observed groundwater head, yellow color plots represent simulations from 90 GeoModel-I, red color plots represent simulations from 90 GeoModel-II, and blue color plots represent simulations from GeoParModel. The edges of the boxes are the 25th and 75th percentiles, and the whiskers extend to the 2.5th and 97.5th percentiles.

of Quaternary sand and 33 % of Quaternary clay, while the proportions extracted from the 3-D training image are 59 and 41 %, respectively. Considering that borehole data are comparatively sparse in this area (approximately 0.04 boreholes per km^2) and 80 % of the boreholes do not penetrate through the whole model depth, the proportions from borehole data hold a higher uncertainty than those based on the 3-D training image. Therefore, 59 and 41 % were defined as proportions for Quaternary sand and clay, respectively, in the MPS simulations. Since both borehole data and 3-D TI show laterally isotropic geometry, the maximum and medium ranges for the search template ellipsoid were both set to 2000 m, and the minimum range in vertical direction was set to 80 m to capture the facies' heterogeneities. Figure 4 illustrates one of the MPS simulations. By visual comparison, the figures to the right (simulations) generally have similar patterns as the TI shown to the left, except that there appear to be slightly more small clay lenses in the simulation. However, Fig. 4 only shows three slices of one of the 100 3-D simulations. The 100 realizations are equally probable representations of the unknown reality. A broad overall view is shown in an

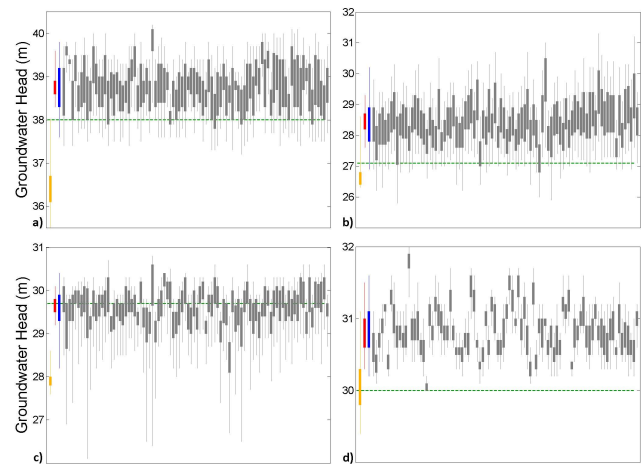


Fig. 11. Boxplot of simulated groundwater head at four selected observation points: (a) Obs10, (b) Obs34, (c) Obs50, (d) Obs57. The dashed green line is the observed groundwater head, the yellow color plots represent 90 simulations from GeoModel-I, the red color plots represent 90 simulations from GeoModel-II, the blue color plots represent GeoParModel, and the gray color plots represent simulations of the individual ParModels, each consisting of 216 simulations. The edges of the boxes are the 25th and 75th percentiles, and the whiskers extend to the 2.5th and 97.5th percentiles (values are listed in Table 4).

E-type map in Fig. 5. The E-type map is constructed as the arithmetic mean of all realizations, where E stands for “expected value” or precisely “conditional expectation” (Remy et al., 2009, p. 37). The heavy red or red dots are found at places where boreholes are located, and that indicates that the hard data have been honored in all 100 simulations. Except for those hard data spots, there are no remarkable scattered small size clay lenses on the E-type map, and the pattern of clay lenses in the E-type map shows a similar trend as in the training image.

5.2 Scenario I: geological uncertainty with constant parameter values (GeoModel-I)

The geological realizations simulated by SNESIM formed the basis for a series of MODFLOW models all having the same hydrogeological boundary conditions. In GeoModel-I the models were not calibrated and the same hydraulic parameters were specified to all models (only 90 geological models in total were considered, as 10 were judged non-behavioral and are listed with star marks in Table A1). The parameters used in this scenario were the median of the calibrated parameter values, see Table 2. Figure 6 (first row) shows the standard deviation of hydraulic head of the 1st and 4th numerical layer based on simulations of GeoModel-I. The variation in the top layer is affected by the external and internal boundary conditions, see Fig. 6a. As expected the variation of hydraulic head tends to be stable towards river courses and boundaries, while a maximum standard variation

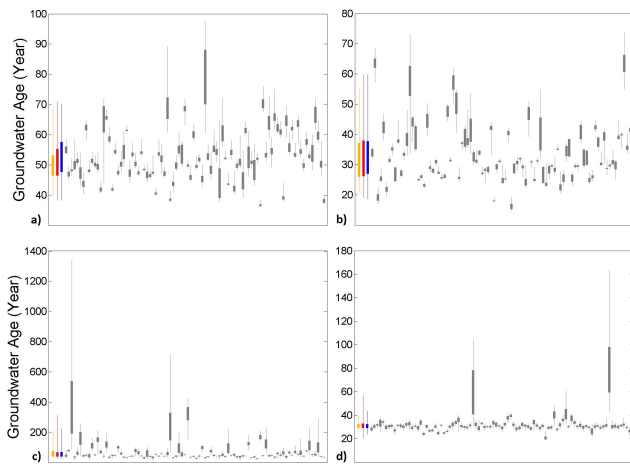


Fig. 12. Boxplot of simulated travel times for four wells: (a) W1, (b) W2, (c) W3, (d) W4. The yellow color plots represent 90 simulations from GeoModel-I, the red color plots represent 90 simulations from GeoModel-II, the blue color plots represent GeoParModel, and the gray color plots represent simulations of individual ParModels, each consisting of 216 simulations. The edges of the boxes are the 25th and 75th percentiles, and the whiskers extend to the 2.5th and 97.5th percentiles (values are listed in Table 3).

of nearly 3 m is found at the higher elevation in the northern part. For a further analysis of the convergence of the 90 models, standard deviation was computed for ten cells for different numbers of model runs. These ten cells were randomly selected in the top layer, since this layer has the highest hydraulic head variation. Further, the locations of the cells were in sufficient distances from the model boundary and river courses such that the standard deviation was not constrained. Figure 7 shows the development of the standard deviation of simulated groundwater head with the number of models. The standard deviation of groundwater head seems to stabilize after 30 model runs, and 90 simulations seem more than sufficient to collect the statistics information.

Backward particle tracking was applied to four abstraction wells having different positions and with the well screen at different levels (Fig. 8). Each well was assigned 100 particles, which were backtracked and travel time distributions computed using MODPATH. Table 3 lists median travel time, 95 % confidence interval, and skewness for these four wells. In Fig. 9 (left column) histograms of simulated particle travel time are shown. For wells located in the northwestern part, the distributions of simulated particle travel time are less skewed due to the primarily downward flow resulting in capture zones located close to the wells. For W1 the travel times are in the range of 25–70 yr with a median value of 50 yr and a skewness value of 0.3. For W2 the travel times are mostly in the range of 20–75 yr with a median value of 29 yr and a skewness value of 1.6. For the wells in the southeastern part, the distributions of travel time become more skewed, as this region is dominated by horizontal flow resulting in elon-

gated capture zones and therefore larger travel distances and spreading in travel times. For W3 the simulated travel time spans from 25 yr to 325 yr, with a median value of 46 yr and a skewness of 2.3. For W4 the range is 20–90 yr with a median value of 31 yr and a skewness of 6.0.

5.3 Scenario II: geological uncertainty with optimized parameter values (GeoModel-II)

In this scenario (GeoModel-II) we analyze to which extent parameter optimization affects the uncertainty of the model predictions. Each model was calibrated by PEST, and parameters included in the optimization were horizontal hydraulic conductivity of Quaternary sand, Quaternary clay and pre-Quaternary sand. The estimated values together with 95 % confidence limit of these 100 models are listed in Table A1. Abnormal parameter estimates as well as extreme confidence limits appear in 10 of the models (those marked with a star in the list). These models were taken out, and the analysis was applied on the remaining 90 models. Table 2 lists the median and the standard deviation of the 90 estimated parameter sets from the PEST analysis. As a result of the optimization the uncertainty is expected to decrease at least for hydraulic head, as each model is calibrated against head observations. The second row of Fig. 6 shows the results in the form of standard deviation of groundwater head for the same two computational layers as shown in the first row. Comparing the two sets of results it is apparent that for both layers the standard deviation has reduced for GeoModel-II, implying that the uncertainty reduces when the model parameters are conditioned on hydraulic head data. In Table 4 statistics of simulated groundwater head for four selected observation points are listed (see location in Fig. 1). As shown in Fig. 6, the uncertainty range differs at different layers, and to cover an appropriate range of groundwater head distributions, the observation points were selected to represent each of the layers 1 to 4. Using the 95 % confidence interval of the GeoParModel simulation as reference, it is apparent that GeoModel-I is subject to higher uncertainty than GeoModel-II as reflected by the *R* values. When calibrating against head measurements the simulations of heads obviously improve and the calibrated parameter values partially compensate for possible biases embedded in the different geological realizations.

Backward particle tracking was applied to the same cells as in scenario GeoModel-I. Figure 9 (middle column) shows the histogram of travel time for 90 simulations of each well, and associated statistics are listed in Table 3. Compared to simulations of GeoModel-I the median travel times are quite similar for all four wells and the distributions for W3 and W4 are also more skewed than for W1 and W2. In contrast to the groundwater head simulations, the uncertainty ranges of travel time for GeoModel-I are smaller than for GeoModel-II as indicated by the *R* values. The higher variability in the simulated travel time of GeoModel-II is a result of the

calibration process. When calibrating against head measurements the resulting variation in parameter values gives rise to a higher variability in flow pathlines and velocities, and realizations with extreme travel time are introduced. Similar results were obtained by Refsgaard et al. (2012).

5.4 Scenario III: geological and parameter uncertainty (ParModel and GeoParModel)

The parameters included in the parameter uncertainty analysis were the same three parameters which were optimized by the PEST code in the previous scenarios. For each of the 90 models developed in the GeoModel-II scenario, 216 parameter sets were sampled and each sampled parameter set served as input to MODFLOW simulations. Figure 10 shows boxplots of the distribution of simulated heads based on GeoModel-I, GeoModel-II, and GeoParModel for each of the 68 observation points from Head Obs. Group 1 in Table 1. The edges of the boxes represent the 25th and 75th percentiles, while the whiskers extend to the 2.5th and 97.5th percentiles, respectively. The observed groundwater heads for each well are shown by green dots. The simulations based on GeoModel-I (yellow boxes) are generally lower than the corresponding observations. This bias is a result of the imposed constant parameter values. As expected, the results from GeoModel-II are in better agreement with observations and the associated uncertainties (red boxes) are less than for GeoModel-I. When parameter uncertainty is added represented by GeoParModel (blue boxes) the uncertainty of the simulated heads is increased and is higher compared to the two other scenarios.

To illustrate the impact of parameter uncertainty explicitly, boxplots for heads at four randomly selected observation points (same as in Table 4) are shown in Fig. 11. The green dashed lines in the figures represent the observed groundwater heads, while the yellow, red and blue boxplots represent the same scenarios as in Fig. 10. The gray boxplots show the results from the ParModels. As stated above, each ParModel represents the results of 216 simulations, which are generated from random sampling in the parameter spaces derived from the PEST analysis of each GeoModel-II. The median values and 95 % uncertainty ranges for all scenarios for these four points are listed in Table 4. The results listed in the last row of the table are the median values of the results of all the 90 ParModels, i.e., median of the medians and 95 % intervals. We take these results as a representation of the impact of parameter uncertainty. Using the uncertainty range of GeoParModel as reference, the R values for the GeoModel-II are much lower than 100 %, indicating that the uncertainty from GeoModel-II is less than that from GeoParModel. The difference between the two scenarios is that parameter uncertainty has been added to GeoParModel, suggesting that uncertainty on hydraulic head is increased when parameter uncertainty is considered. The median values of the 90 ParModels listed in Table 4 show that the median of the uncertainty intervals

of all 90 ParModels is slightly higher than the corresponding interval for GeoModel-II for three of the observation points. This suggests that the effect of parameter uncertainty is comparable to the effect of geological uncertainty with respect to hydraulic head uncertainty.

Backward particle tracking was applied to the 216 parameter realizations of each ParModel. The travel time distribution of 90×216 realizations is plotted in Fig. 9 (right column), and the corresponding values of median, 95 % interval and skewness are listed in Table 3. By introducing parameter uncertainty in the GeoParModel, the uncertainty ranges of travel time have increased somewhat, as more extreme travel times are found when comparing the right and middle columns of Fig. 9. However, this is not necessarily reflected in the R values. The R values are computed from the 95 % confidence intervals and when the skewness is not comparable between the scenarios as for W3 and W4, the R values may be misleading. Nevertheless, in contrast to hydraulic head simulation, parameter uncertainty has much less effect on travel time simulation. This can be inferred from the very low R values based on the median of the 90 ParModels and also evidenced from the boxplots of the individual ParModel simulations shown in Fig. 12. Although there are a few outliers, overall the uncertainty intervals of the individual ParModels are generally shorter than the corresponding intervals for GeoModel-I and GeoModel-II, suggesting that the effect of parameter uncertainty is relatively low with respect to travel time. In this regard the geological architecture is the most critical factor.

6 Conclusions

This study has examined the impact of geological and parameter uncertainty on real case simulations of groundwater heads and travel time using the multiple-point geostatistical method (MPS). A 3-D training image derived from geophysical data was used as basis for the MPS simulations. Usually geophysical data are used as soft data for conditioning; however, as used here it was possible to develop a reliable 3-D geological model as a training image input to the MPS.

Generally one would expect that the uncertainty range will decrease when calibrating models, but our results show that this is not always the case. Although the uncertainty of groundwater head simulations is reduced when the parameters are calibrated against hydraulic head observations, the opposite is the case for the uncertainty of travel time simulations. Calibration implies that biases in the realizations of geological architecture are compensated by errors in parameter values which lead to a larger range of variation in travel time. This also underpins previous findings that when using models outside the calibrated regime, which indeed is the case when simulating travel time, the prediction uncertainty is high.

- Beven, K. J. and Binley, A.: The future of distributed models: model calibration and uncertainty prediction, *Hydrol. Process.*, 6, 279–298, 1992.
- Brauchler, R., Cheng, J. T., Dietrich, P., Everett, M., Johnson, B., Liedl, R., and Sauter, M.: An inversion strategy for hydraulic tomography: coupling travel time and amplitude inversion, *J. Hydrol.*, 345, 184–198, 2007.
- Caers, J.: Geostatistical reservoir modeling using statistical pattern recognition, *J. Petrol. Sci. Eng.*, 29, 177–188, 2001.
- Caers, J. and Zhang, T.: Multiple-point geostatistics: a quantitative vehicle for integrating geologic analogs into multiple reservoir models, in: AAPG Special Volumes Memoir 80: Integration of Outcrop and Modern Analogs in Reservoir Modeling, edited by: Grammer, G. M., Harris, P. M., and Eberli, G. P., Tulsa, OK, USA, 383–394, 2004.
- Carle, S. F. and Fogg, G. E.: Transition Probability-Based Indicator Geostatistics, *Mathe. Geol.*, 28, 453–476, 1996.
- Carle, S. F.: TPROGS – Transition Probability Geostatistical Software, Version 2.1, User Manual, Hydrologic Sciences Graduate Group, University of California, Davis, California, 1999.
- Carle, S. F., LaBolle, E. M., Weissmann, G. S., VanBrocklin, D., and Fogg, G. E.: Conditional simulation of hydrofacies architecture: a transition probability/Markov approach, in: Hydrogeologic Models of Sedimentary Aquifers, Concepts in Hydrogeology and Environmental Geology No. 1, edited by: Fraser, G. S. and Davis, J. M., SEPM (Society for Sedimentary Geology) Special Publication, AAPG, Tulsa, OK, USA, 147–170, 1998.
- Cheng, C. and Chen, X. H.: Evaluation of methods for determination of hydraulic properties in an aquifer-aquitard system hydrologically connected to a river, *Hydrogeol. J.*, 15, 669–678, 2007.
- Comunian, A., Renard, P., Straubhaar, J., and Bayer, P.: Three-dimensional high resolution fluvio-glacial aquifer analog, Part 2: Geostatistical modelling, *J. Hydrol.*, 405, 10–23, 2011.
- Comunian, A., Renard, P., and Straubhaar, J.: 3-D multiple-point statistics simulation using 2-D training images, *Comput. Geosci.*, 40, 49–65, 2012.
- David, W. P.: User's Guide for MODPATH/MODPATH-PLOT, Version 3: a Particle Tracking Post-Processing Package for MODFLOW, the US Geological Survey Finite-Difference Ground-Water Flow Model, US Geological Survey Open File Rep. 94–464, Tulsa, OK, USA, 1994.
- Delhomme, J. P.: Spatial variability and uncertainty in groundwater flow parameters: a geostatistical approach, *Water Resour. Res.*, 15, 269–280, 1979.
- Dettinger, M. D. and Wilson, J. L.: First order analysis of uncertainty in numerical models of groundwater flow, *Water Resour. Res.*, 17, 149–161, 1981.
- Doherty, J.: PEST: Model Independent Parameter Estimation, 5th Edn. of user manual, Watermark Numerical Computing, Brisbane, Australia, 2005.
- Evan, R. A. and Mary, C. H.: Documentation of the hydrogeologic-unit flow (HUF) package, US Geological Survey Open File Rep. 00–342, US Geological Survey, Denver, CO, USA, 2000.
- Feyen, L. and Caers, J.: Quantifying geological uncertainty for flow and transport modeling in multi-modal heterogeneous formations, *Adv. Water Resour.*, 29, 912–929, 2006.
- Feyen, L., Beven, K. J., De Smedt, F., and Freer, J.: Stochastic capture zone delineation within the generalized likelihood uncertainty estimation methodology: conditioning on head observations, *Water Resour. Res.*, 37, 625–638, 2001.
- Fleckenstein, J. H., Niswonger, R. G., and Fogg, G. E.: River-aquifer interactions, geologic heterogeneity, and low-flow management, *Ground Water*, 44, 837–852, 2006.
- Freeze, R. A.: A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media, *Water Resour. Res.*, 11, 725–741, 1975.
- Freeze, R. A., Massmann, J., Smith, L., Sperling, T., and James, B.: Hydrogeological decision analysis: 1. A framework, *Ground Water*, 28, 738–766, 1990.
- Goovaerts, P.: Geostatistics for Natural Resources Evaluation, Oxford University Press, New York, 483 pp., 1997.
- Guardiano, F. and Srivastava, R. M.: Multivariate geostatistics: beyond bivariate moments: geostatistics-troia, in: Geostatistics-Troia, Vol. 1, edited by: Soares, A., Kluwer, Dordrecht, 133–144, 1993.
- Harbaugh, A. W., Banta, E. R., Hill, M. C., and McDonald, M. G.: MODFLOW-2000, the US Geological Survey Modular Ground-Water Model – User Guide to Modularization Concepts and the Ground-Water Flow Process, US Geological Survey Open-File Report 00–92, US Department of Interior, Reston, VA, USA, 121 pp., 2000.
- Harrar, W. G., Sonnenborg, T. O., and Henriksen, H. J.: Capture zone, travel time, and solute transport predictions using inverse modeling and different geological models, *J. Hydrol.*, 11, 536–548, 2003.
- Hassan, A.E., Bekhit, H. M., and Chapman, J. B.: Uncertainty assessment of a stochastic groundwater flow model using GLUE analysis, *J. Hydrol.*, 362, 89–109, 2008.
- Henriksen, H. J., Troldborg, L., Nyegaard, P., Sonnenborg, T. O., Refsgaard, J. C., and Madsen, B.: Methodology for construction, calibration and validation of a national hydrological model for Denmark, *J. Hydrol.*, 280, 52–71, 2003.
- Hendricks Franssen, H. J., Stauffer, F., and Kinzelbach, W.: Joint estimation of transmissivities and recharges—application: stochastic characterization of well capture zones, *J. Hydrol.*, 294, 87–102, 2004.
- Hoeksema, R. J. and Kitanidis, P. K.: Analysis of the spatial structure of properties of selected aquifers, *Water Resour. Res.*, 21, 563–572, 1985.
- Honarkhah, M. and Caers, J.: Direct Pattern-Based Simulation of Nonstationary Geostatistical Models, *Math. Geosci.*, 44, 651–672, 2012.
- Høyer, A. S., Lykke-Andersen, H., Jørgensen, F., and Auken, E.: Combined interpretation of SkyTEM and high-resolution seismic data, *Phys. Chem. Earth*, 36, 1386–1397, doi:10.1016/j.pce.2011.01.001, 2011.
- Høyer, A. S., Jørgensen, F., Lykke-Andersen, H., and Christiansen, A. V.: Iterative modelling of AEM data based on geological a priori information from seismic and borehole data, *Near Surf. Geophys.*, submitted, 2013.
- Johnson, N. M.: Characterization of alluvial hydrostratigraphy with indicator semivariograms, *Water Resour. Res.*, 31, 3205–3216, 1995.
- Journel, A. G.: Beyond covariance: the advent of multiple-point geostatistics, in: Geostatistics, Banff, edited by: Leuangthong, O. and Deutsch, C. V., Springer, Dordrecht, The Netherlands, 2004, 225–233, 2005.

- Journel, A. G. and Alabert, F.: New method for reservoir mapping, *J. Petrol. Technol.*, 42, 212–218, 1990.
- Journel, A. G. and Zhang, T.: The necessity of a multiple-point prior model, *Math. Geol.*, 38, 591–610, 2006.
- Jørgensen, F. and Sandersen, P. B. E.: Buried and open tunnel valleys in Denmark – erosion beneath multiple ice sheets, *Quaternary Sci. Rev.*, 25, 1339–1363, 2006.
- Jørgensen, F., Møller, R. R., Høyer, A. H., and Christiansen, A. V.: Geologisk model ved Ølgod og Skovlund – eksempel på effektiviseret modellering i et heterogent geologisk miljø, Danmarks og Grønlands Geologiske Undersøgelse Rapport 2012/82, GEUS, Copenhagen, Denmark, 83 pp., 2012.
- Klise, K. A., Weissmann, G. S., McKenna, S. A., Nichols, E. M., Frechette, J. D., Wawrzyniec, T. F., and Tidwell, V. C.: Exploring solute transport and streamline connectivity using Lidar-based outcrop images and geostatistical representations of heterogeneity, *Water Resour. Res.*, 45, W05413, doi:10.1029/2008WR007500, 2009.
- Le Coz, M., Genthon, P., and Adler, P. M.: Multiple-point statistics for modeling facies heterogeneities in a porous medium: the Komadugu-Yobe Alluvium, Lake Chad Basin, *Math. Geosci.*, 43, 861–878, 2011.
- Liu, Y.: Using the Snesim program for multiple-point statistical simulation, *Comput. Geosci.*, 32, 1544–1563, 2006.
- Liu, Y., Harding, A., Journel, A. G., and Gilbert, R.: A workflow for multiple-point geostatistical simulation, in: *Geostatistics Banff 2004*, Vol. 1, edited by: Leuangthong, O. and Deutsch, C. V., Springer, Dordrecht, 245–254, 2005.
- Mariethoz, G., Renard, P., Cornaton, F., and Jaquet, O.: Truncated Plurigaussian simulations to characterize aquifer heterogeneity, *Ground Water*, 47, 13–24, 2009.
- McDonald, M. G. and Harbaugh, A. W.: A Modular Three-Dimensional Finite-Difference Ground-Water Flow Model, US Geological Survey Open File Rep. 83–875, US GPO, Washington, USA, 1988.
- McKay, M. D., Conover, W. J., and Beckman, R. J.: A comparison of three methods for selection values of input variables in the analysis of output from a computer code, *Technometrics*, 2, 239–245, 1979.
- Mantovan, P. and Todini, E.: Hydrological forecasting uncertainty assessment: incoherence of the GLUE methodology, *J. Hydrol.*, 330, 368–381, 2006.
- Mantovan, P., Todini, E., and Martina, M. L. V.: Reply to comment by Keith Beven, Paul Smith and Jim Freer on “Hydrological forecasting uncertainty assessment: incoherence of the GLUE methodology”, *J. Hydrol.*, 228, 319–324, 2007.
- Rasmussen, E. S., Dybkjær, K., and Piasecki, S.: Lithostratigraphy of the Upper Oligocene Miocene Succession of Denmark, *Geological Survey of Denmark and Greenland Bulletin*, 22, GEUS, Copenhagen, Denmark, 92 pp., 2010.
- Refsgaard, J. C., Christensen, S., Sonnenborg, D. S., Højberg, A. L., and Trolborg, L.: Review of strategies for handling geological uncertainty in groundwater flow and transport modeling, *Adv. Water Resour.*, 36, 36–50, 2012.
- Remy, N., Boucher, A., and Wu, J.: *Applied Geostatistics with SGeMS*, Cambridge University Press, Cambridge, UK, 2009.
- Renard, P.: Stochastic hydrogeology: what professionals really need?, *Ground Water*, 45, 531–541, 2007.
- Sørensen, K. I. and Auken, E.: SkyTEM – a new high-resolution helicopter transient electromagnetic system, *Explor. Geophys.*, 35, 191–199, 2004.
- Stisen, S., Sonnenborg, T. O., Højberg, A. L., Trolborg, L., and Refsgaard, J. C.: Evaluation of climate input biases and water balance issues using a coupled surface-subsurface model, *Vadose Zone J.*, 10, 37–53, 2011.
- Strebelle, S. B.: Conditional simulation of complex geological structures using multiple-point statistics, *Math. Geol.*, 34, 1–21, 2002.
- Strebelle, S. B.: Sequential simulation for modeling geological structures from training images, in: *Stochastic Modeling and Geostatistics: Principles, Methods, and Case Studies*, vol. II: AAPG Computer Applications in Geology 5, edited by: Coburn, T. C., Yarus, J. M., and Chambers, R. L., AAPG, Tulsa, OK, USA, 139–149, 2006.
- Tonkin, M. and Doherty, J.: A hybrid regularized inversion methodology for highly parameterized environmental models, *Water Resour. Res.*, 41, W10412, doi:10.1029/2005WR003995, 2005.
- Vrugt, J. A., Stauffer, P. H., Wohling, T., Robinson, B. A., and Vesselinov, V. V.: Inverse modeling of subsurface flow and transport properties: a review with new developments, *Vadose Zone J.*, 7, 843–864, 2008.
- Webb, E. K. and Anderson, M. P.: Simulation of preferential flow in three-dimensional, heterogeneous conductivity fields with realistic internal architecture, *Water Resour. Res.*, 32, 533–545, 1996.
- Wingle, W. L. and Poeter, E. P.: Uncertainty associated with semi-variograms used for site simulation, *Ground Water*, 31, 725–734, 1993.