
Capture zone, travel time, and solute-transport predictions using inverse modeling and different geological models

William G. Harrar · Torben Obel Sonnenborg ·
Hans Jørgen Henriksen

Abstract Six regional-scale flow models are compared to gain insight into how different representations of hydraulic-conductivity distributions affect model calibration and predictions. Deterministic geological models were used to define hydraulic-conductivity distributions in two steady-state flow models that were calibrated to heads and baseflow estimates using inverse techniques. Optimized hydraulic-conductivity estimates from the two models were used to calculate layer and model mean hydraulic-conductivity values. Despite differences in the two geological models, inverse calibration produced mean hydraulic-conductivity values for the entire model domain that are quite similar. The layer and model mean hydraulic-conductivity values were used to generate four additional flow models and forward runs were performed. All of the models adequately simulate the observed heads and total baseflow. The six flow models were used to predict the steady-state impact of a proposed well field, and the flow solutions were used in simulating particle tracking and solute transport. Results of the predictive simulations show that, for this example, simple models of heterogeneity produce capture zones similar to more complex models, but with very different travel times and breakthroughs. Inverse modeling combined with different geological models can provide a measure of capture zone and breakthrough reliability.

Résumé Six modèles d'écoulement à l'échelle régionale sont comparés afin d'avoir un aperçu de la manière dont les différentes représentations de la distribution de la conductivité hydraulique affectent la calibration et les prédictions de modèles. Des modèles géologiques déter-

ministes ont été utilisés pour définir les distributions de la conductivité hydraulique dans deux modèles d'écoulement en régime permanent qui ont été calibrés avec des estimations des charges et des écoulements de base faites par des techniques inverses. Les estimations optimisées de la conductivité hydraulique de ces deux modèles ont servi à calculer les valeurs de conductivité hydraulique moyenne des couches et du modèle. Malgré des différences entre les deux modèles géologiques, la calibration inverse a donné des valeurs de conductivité hydraulique moyenne pour le domaine complet du modèle qui sont complètement semblables. Les valeurs de la conductivité moyenne des couches et du modèle ont été utilisées pour générer quatre modèles d'écoulement supplémentaires et des traitements ont été effectués. Tous les modèles simulent correctement les charges observées et l'écoulement de base total. Les six modèles ont servi à prédire l'impact en régime permanent d'un champ captant projeté et les solutions d'écoulement ont été utilisées dans une simulation par suivi de particules et de transport de soluté. Les résultats de simulations prédictives montrent que, pour cet exemple, de simples modèles d'hétérogénéité fournissent des zones de capture semblables aux modèles plus complexes, mais pour des temps de parcours et des restitutions très différents. Une modélisation inverse combinée à différents modèles géologiques peut assurer une mesure de la zone de capture et une fiabilité de la restitution.

Resumen Se compara seis modelos de flujo a escala regional para conocer cómo afecta a la calibración y a la predicción del modelo diversas representaciones de la distribución de la conductividad hidráulica. Se ha utilizado modelos geológicos deterministas para definir las distribuciones de la conductividad hidráulica en dos modelos de flujo permanente, calibrados mediante técnicas inversas con niveles piezométricos y estimaciones del flujo de base. Se ha adoptado estimaciones optimizadas de la conductividad hidráulica de los dos modelos para calcular las cotas de las capas y sus conductividades hidráulicas medias. A pesar de las diferencias entre ambos modelos geológicos, con la calibración inversa se obtiene valores similares de conductividad hidráulica en todo el dominio. Estos valores de las capas y de las conductividades hidráulicas medias han servido para generar cuatro modelos adicionales de flujo y realizar predicciones.

Received: 7 January 2002 / Accepted: 13 May 2003
Published online: 12 July 2003

© Springer-Verlag 2003

W. G. Harrar (✉) · H. J. Henriksen
Geological Survey of Denmark and Greenland,
Øster Voldgade 10, 1350 Copenhagen K, Denmark
e-mail: bwh@geus.dk
Fax: +45-38-142050

T. O. Sonnenborg
Environment & Resources DTU, Technical University of Denmark,
2800 Kgs. Lyngby, Denmark

Todos los modelos simulan de forma adecuada los niveles observados y los caudales de base. Los seis modelos han sido aplicados a la predicción del impacto estacionario de un campo de pozos, y las soluciones del flujo permiten simular el transporte de partículas y de solutos. Los resultados de estas predicciones muestran que, para este ejemplo, los modelos sencillos de la heterogeneidad dan lugar a zonas de captura similares a las generadas por modelos más complejos, pero aparecen grandes diferencias en los tiempos de tránsito y en las curvas de llegada. Una combinación de modelación inversa y de modelos geológicos diferentes puede proporcionar una medida de la fiabilidad de la zona de captura y de las curvas de llegada.

Keywords Sedimentary heterogeneity · Geological models · Zonation · Capture zone · Inverse modeling · Predictions

Introduction

The drinking water supply in Denmark is based exclusively on groundwater. In recent years, over 500 water-supply wells have been forced to close due to water quality problems; the majority of these were shallow wells contaminated with nitrates and pesticides (Stockmarr 2001). The decrease in the quality of shallow groundwater has increased the interest in using deep aquifers for water supply. In Denmark, permission for the establishment of a new well field is granted based partly upon predictions of well-field capture zones. Accurate delineation of capture zones is important with regard to potential contamination risks and land-use restrictions. Methods used for capture-zone prediction range from simple analytical expressions, where hydraulic properties are described by a single average value, to numerical models that more accurately represent aquifer heterogeneity. Regardless of the method used, a common concern is the effect of averaging of hydraulic-conductivity values on the accuracy or uncertainty of predicted capture zones.

Bhatt (1993) evaluated the effect of data uncertainty on simulated capture zones and concluded that the precision of aquifer parameters is the most important factor in delineating capture zones. Cole and Silliman (2000) examined the utility of simple models for capture-zone prediction in heterogeneous aquifers and applied safety factors to model predictions in order to account for uncertainty. Capture-zone predictions from analytical and semianalytical models are often misleading due to oversimplification of complex hydrogeological settings, and therefore the use of numerical models is preferable (Springer and Bair 1992). However, uncertainty of parameter values, zonation, and boundary conditions often result in predictions from numerical models that are not unique, and it is difficult to evaluate how model structure and averaging of hydraulic-conductivity values affect these predictions. Rayne et al. (2001) used a three-dimensional deterministic flow model to delineate capture

zones in a complex sedimentary aquifer system, and imply that the predictions are reliable because of consistency between field observations and model results. Evers and Lerner (1998) present a method to determine the range of capture-zone predictions arising from alternative calibrations of a numerical model and, from these, identify zones of certainty and uncertainty. Stochastic models have been used to quantify prediction uncertainty by evaluating an ensemble of capture zones reflecting the variability and uncertainty of hydraulic property estimates (Feyen et al. 2001; van Leeuwen et al. 2000; Vassolo et al. 1998; Varljen and Shafer 1991; Bair et al. 1991). However, stochastic models are rarely used outside of research applications due to extensive data requirements and lack of standardized methodology.

The hydrogeologist interested in predicting capture zones is faced with a dilemma. Simple models tend to provide unreliable predictions, deterministic numerical models provide a single prediction the uncertainty of which is not readily quantifiable and, while prediction uncertainty is quantifiable with stochastic models, their application is generally limited to field sites with large datasets. The most common approach is to use deterministic numerical models with the hope that the model represents both mean hydraulic-conductivity values and aquifer heterogeneity in sufficient detail to produce reliable capture-zone predictions.

In this study, two different deterministic geological models were used to define hydraulic-conductivity distributions in two steady-state flow models that were calibrated to hydraulic heads and flows using inverse techniques. Optimized hydraulic-conductivity values from each calibrated model were used to calculate mean hydraulic-conductivity values for each model layer and the model as a whole. Forward runs were then performed using the layer and model mean hydraulic-conductivity values, and results were compared with hydraulic-head measurements and baseflow estimates. Simulations of capture zones, particle-travel times and solute transport were performed with six flow models with the objectives of comparing predictions from optimized models constructed from different geological models, and to determine the effects of zonation and averaging of hydraulic conductivity on model predictions.

Methodology

The study area is situated within a dissected glacial-till plain on the Jutland peninsula in western Denmark and covers an area of 41 by 25 km, as shown in Fig. 1. The topography varies from over 90-m elevation above mean sea level in the northeastern corner of the study area to sea level in the southwest corner. Land use is predominantly agricultural followed by wetlands, forests and towns.

In 1990 the Geological Survey of Denmark and Greenland (GEUS) was contracted to locate a new groundwater source with a capacity of up to 2.8×10^6 m³/year for the city of Esbjerg in western

Fig. 1 Location of the study area. *Inset* shows the topography and streams. Location of cross sections (Figs. 2 and 4) indicated by line W–E

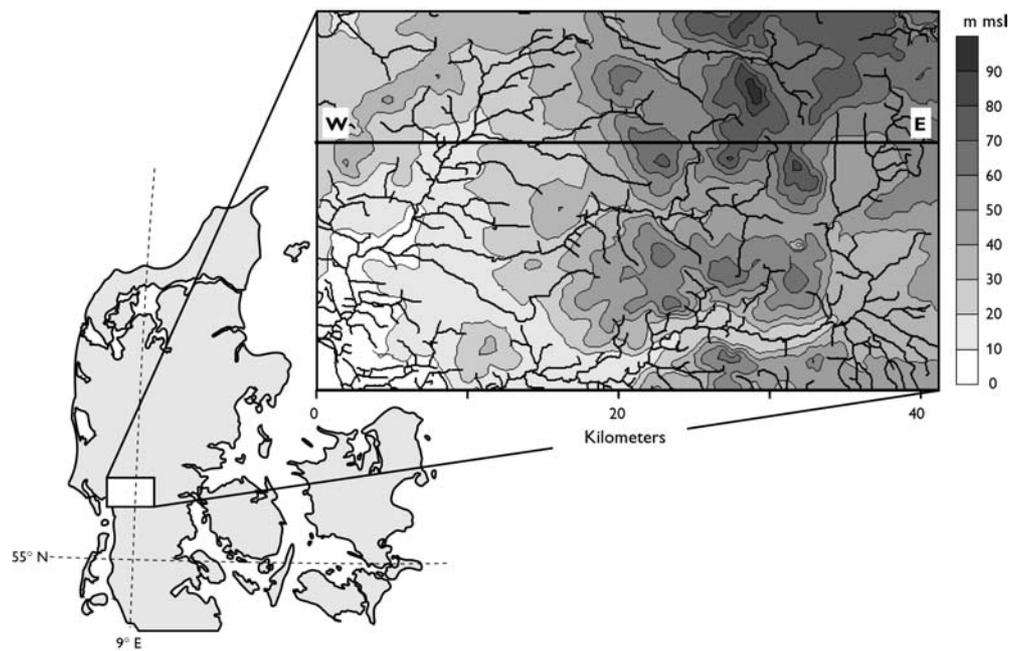
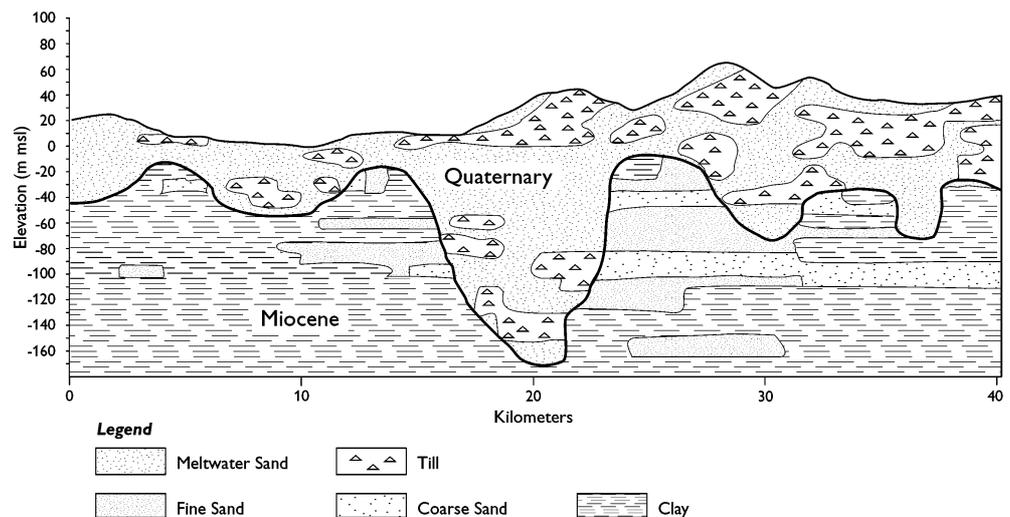


Fig. 2 Cross-section showing the geology of study area



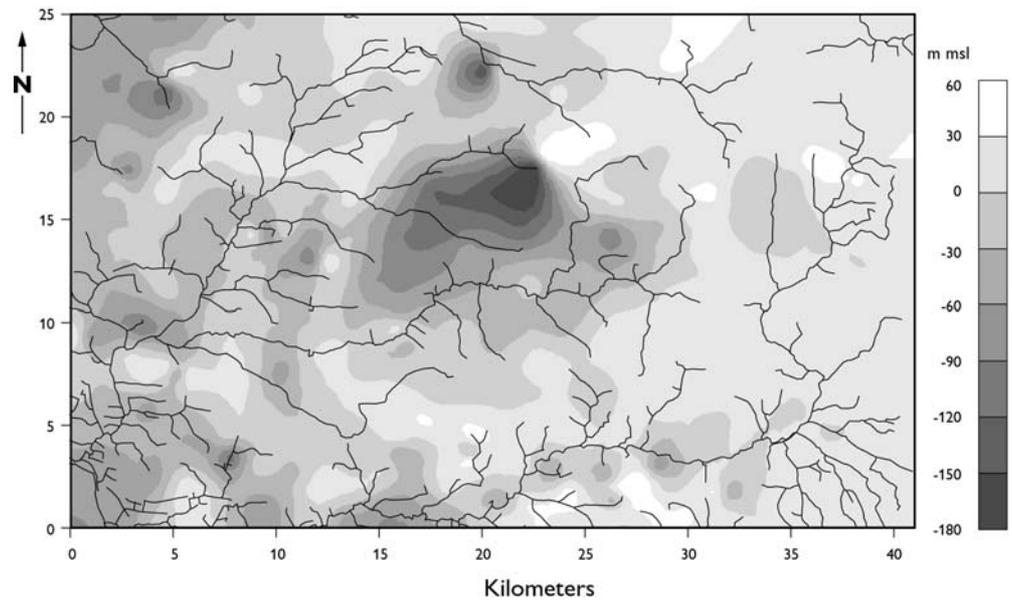
Denmark. The resource was investigated over a five-year period by conducting geophysical surveys, drilling of deep wells, borehole logging, chemical analysis, pumping tests, and monitoring of groundwater levels and stream discharge (Klitten et al. 1995). As part of the investigation, a steady-state flow model was developed and calibrated through trial-and-error to predict the response of the flow system to pumping from the proposed well field (Henriksen et al. 1995). The flow model was modified and used as the basis for this study.

Geology

The study area is underlain by Tertiary and Quaternary sedimentary strata, as shown in Fig. 2. The Tertiary is comprised of Miocene age sediments deposited in alter-

nating terrestrial/nearshore and shelf environments (Dybkjaer et al. 2001). The Tertiary sediments consist of a thick sequence of clay with some interbeds of micaceous sand in the western part of the study area, and become coarser-grained to the east. The top surface of the Tertiary sediments, shown in Fig. 3, is dissected by channels and buried valleys, the deepest of which is present in the central part of the model area and extends to depths exceeding -150 m s.l. (Klitten et al. 1995). The channels and buried valleys are filled with Quaternary glaciofluvial sands with some thick meltwater clay and till bodies. The upper part of the Quaternary consists of glacial till and meltwater sand. The geology of the Quaternary sequence is complex and, despite the high density of subsurface data (average of >5 borehole logs/ km^2), the interconnectivity of sands and the dimensions of

Fig. 3 Map of the top surface of the Tertiary sequence



till and clay deposits are not well defined. The Quaternary sediments range in thickness from 10 m in the western part of the study area to over 250 m in the central erosional feature.

Water Balance

A mass-balance statement for the model area was compiled using data from 60 precipitation stations, 100 stream gauging stations, groundwater extraction records from over 900 production wells and soil maps (Henriksen et al. 1995). Precipitation varies spatially throughout the model area, primarily as a function of elevation, and values range from 900 mm/year in low-lying areas to 1,200 mm/year in uplands. The average precipitation over a 10-year period is 950 mm/year. Average evapotranspiration for the model area, estimated using a root-zone model and 10-year records of measured potential evapotranspiration, is 450 mm/year. The summation of overland flow, discharge to drains, and shallow interflow is estimated to be 260 mm/year based upon stream-discharge records from a 224-km² catchment located in the central part of the study area. The average recharge to the flow system, defined here as the precipitation minus actual evapotranspiration minus surface/near-surface flow, is estimated to be 240 mm/year. An average baseflow of 195 mm/year within a total catchment area of 1,963 km² was calculated from synchronous low-flow stream measurements from 109 stations made during August in the period 1989–1994. It is assumed that late-summer stream flow is comprised solely of influent groundwater. Total groundwater extraction (waterworks, irrigation and domestic wells) for 1994 was reported at 24 mm/year within an area of 1,963 km². The sum of the baseflow and groundwater extraction is 219 mm/year or about 90% of the estimated recharge, indicating that groundwater discharge to

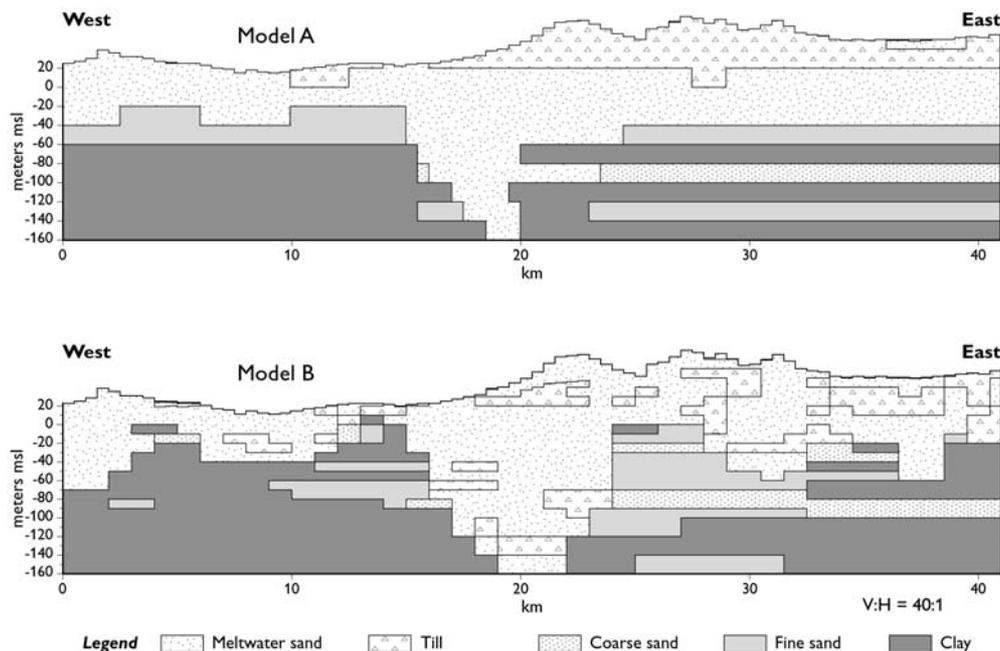
streams and pumping wells represents the major discharge from the study area.

Hydraulic Parameters

Transmissivity values, calculated from the results of 28 pumping tests conducted in and adjacent to the central buried valley, vary in the range $7 \cdot 10^{-4}$ – $1.5 \cdot 10^{-2}$ m²/s for the Quaternary, and $2.1 \cdot 10^{-3}$ – $5.8 \cdot 10^{-2}$ m²/s for the Miocene deposits (Klitten et al. 1995). Hydraulic-conductivity values were estimated at pumping wells by dividing the transmissivity by the filter length, and are in the range $5.8 \cdot 10^{-5}$ – $8.1 \cdot 10^{-4}$ m/s for the Quaternary sands, and $1.7 \cdot 10^{-4}$ – $9.3 \cdot 10^{-3}$ m/s for the Miocene sands and gravels. Transmissivity values were also calculated using specific-capacity data from an additional 43 wells, and were assigned qualitative estimates of high, medium, and low. The transmissivity estimates (absolute and relative) were used in model parameterization. Mean hydraulic-conductivity values were used as initial values for the inverse models.

Transmissivity values (absolute and relative) were plotted on geological cross sections and maps, and transmissivity trends were identified. In the central and eastern parts of the study area, the Quaternary sequence was assigned transmissivity estimates of low–medium for elevations above sea level, due to the predominant occurrence of glacial tills, and high to depths below sea level where meltwater sands are predominant. West of the central buried valley, the Quaternary sequence was assigned transmissivity estimates of low–medium, reflecting both the local presence of fine-grained deposits and a decrease in overall thickness. The Miocene was assigned transmissivity estimates of low in the western part, where the stratigraphic sequence is predominantly clay, and high–medium in the eastern part of the model area where the sediments become more coarse. The

Fig. 4 Cross sections showing the geological models used to construct the optimized flow model



location of the contact between the low- and high-transmissivity Miocene deposits is imprecise and designated along a north–south line coinciding with the western wall of the central erosional feature.

Numerical Modeling

Two groundwater-flow models, constructed using two different geological models, were calibrated to steady-state conditions against observed heads and baseflow estimates using MODFLOW (McDonald and Harbaugh 1989) and UCODE (Poeter and Hill 1998). The following sections describe the model design and calibration.

Discretization

The model area is 41 by 25 km and was discretized into 82 columns and 50 rows with a constant grid-cell dimension of 500×500 m. A digital-terrain map was used to define the upper surface of each model, and the lower surface was assigned a constant elevation of –160 m m.s.l. Model layers were assigned constant top and bottom elevations with a vertical discretization of 20 m in model A, and 10 m in model B (Fig. 4). The top and bottom elevations of layers above sea level were locally adjusted to accommodate for variations in topography. The number of layers in models A and B are 11 and 21, respectively.

Geological Models

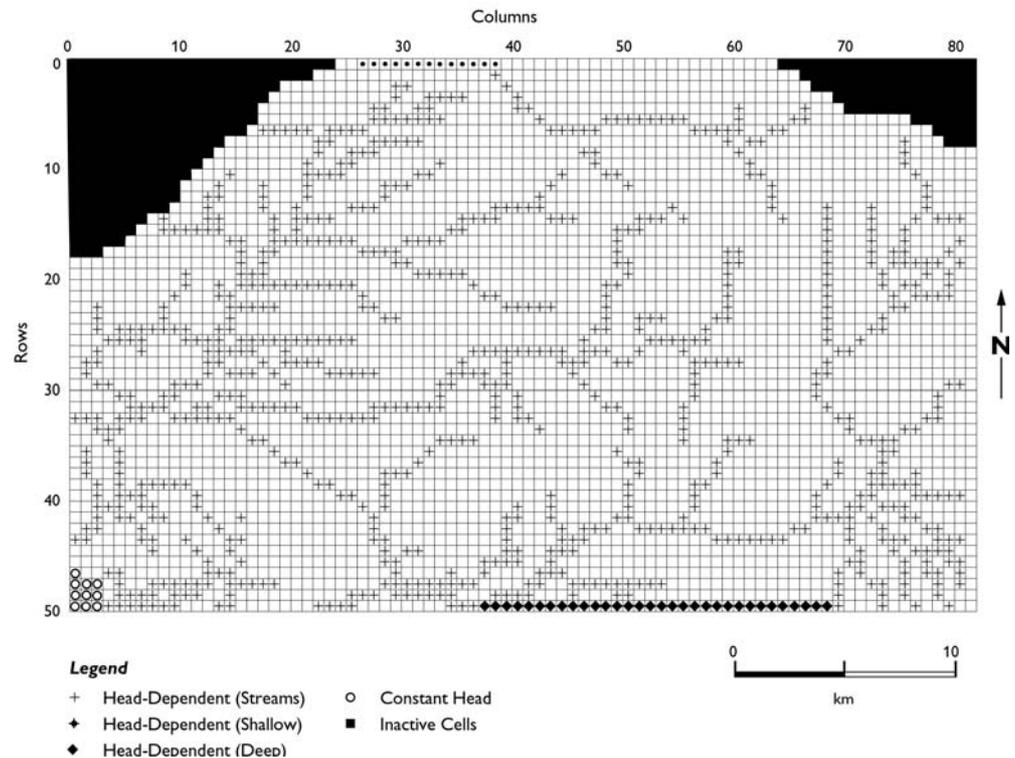
Two geological models were constructed using over 5,400 borehole logs archived at the Geological Survey of Denmark and Greenland. The geology in the two models is represented by the following five sediment types: (1)

Quaternary meltwater sands (meltwater sand), (2) Quaternary glacial till (glacial till), (3) Tertiary quartz sand (coarse sand), (4) Tertiary micaceous sand (fine sand), (5) Tertiary clay (clay).

Model A was constructed using 18 east–west geological cross sections with a spacing of 2 km between the cross sections, transmissivity maps, and a digital map of the top of the Tertiary surface. First, the digital map of the top of the Tertiary surface was used to define the distribution of Quaternary and Tertiary sediments in each model layer. The resulting maps were used as templates for delineation of geological zones by manually assigning the sediment type to each cell, using the geological cross sections and transmissivity maps. The borehole data were averaged during generation of the cross sections, and additional averaging was performed when the distribution of sediment types was delineated in each model layer based upon the cross sections and transmissivity maps.

Whereas model A was interpreted in vertical cross section, model B was interpreted and correlated in horizontal layers. The stratigraphic succession in model B was interpreted in 10-m-thick layers using a horizontal grid size of 1,000×1,000 m. The borehole data (sediment type) were plotted in each of the 21 model layers, and the dominant sediment type in each grid cell was determined. The cells containing borehole information were then coded exclusively according to this sediment type. Sediment types for the remaining cells were then delineated manually in each layer by interpolation between cells containing borehole information. In contrast to model A, the borehole data were used directly in constructing model B, and the distribution of sediment types assigned to each layer in the model adheres explicitly to the borehole data.

Fig. 5 Model grid and boundary conditions



In general, the major differences between the two geological models are the degree of heterogeneity represented in the Quaternary deposits, and the continuity of the Miocene sediments. As shown in Fig. 4, model A resembles a stratified aquifer system, while model B more closely reflects the observed heterogeneity. The same borehole data were used in constructing the two models but, in order to obtain the highest degree of independence between the two models, the interpretations were carried out by two different geologists. Differences between the models are partly due to how the data were interpolated, and partly to the interpretation carried out by the individual geologist. Sediment zones in the geological models were used in defining hydraulic-conductivity distributions in the two flow models.

Boundary Conditions

Boundary conditions were assigned based largely upon regional groundwater levels and the mass-balance statement. The model grid and boundary conditions are shown in Fig. 5. The lateral boundaries coincide with regional groundwater divides and are simulated as no-flow boundaries with the following exceptions. The southwest corner of the model area coincides with the North Sea and is simulated as a constant-head boundary. A head-dependent boundary was assigned at depth along a segment of the southern boundary, allowing for discharge where high-permeability Miocene sands extend out of the model area. During initial trial-and-error calibration, it was found necessary to allow discharge from the model at shallow depth along a segment of the northern boundary.

This corresponds to where coarse-grained valley-fill sediments extend out of the model area, and is simulated as a head-dependent boundary. Streams were simulated as head-dependent boundaries and divided into 36 reaches, based upon baseflow estimates from 109 monitoring locations and maps of the surface geology. The downstream endpoint of the majority of the stream reaches corresponds to monitoring stations where both stage and discharge were measured. Stream stage was assigned in each cell along the stream reaches by linear interpolation between measured values of stage.

The upper surface of the model was simulated as a specified flux boundary representing recharge to the water table from the infiltration of precipitation. It is assumed that recharge represents the portion of the infiltrating precipitation that contributes to the regional groundwater budget and excludes drain flow and shallow interflow. Recharge was distributed based upon an analysis of precipitation records from 60 monitoring stations, topography, evapotranspiration estimates, land use, and surficial geology (Henriksen et al. 1995). The initial recharge was applied at a total rate of 240 mm/year distributed over five zones with values varying in the range 40–320 mm/year. The distribution of the recharge zones and their initial rates differ somewhat between the two models, due to differences in the interpretation of the near-surface geology, whereas the initial estimates of the total recharge, applied to both models, are the same. The zonal values of recharge were optimized using a multiplication factor, and therefore the total recharge rate was ultimately a result of model calibration, as described below.

Calibration Data

The steady-state models were calibrated to baseflow estimates and hydraulic-head measurements representative of low-flow conditions. Baseflow estimates for 36 stream reaches were used for comparison with the simulated groundwater discharge to streams. The baseflow estimates were calculated from synchronous low-flow stream measurements made during August in the period 1989–1994 at 109 monitoring stations. Several of the stream reaches are within subcatchments of larger catchments, and the discharge values will therefore be correlated. However, the monitoring stations from which the 36 reaches were identified are distributed relatively uniformly throughout the stream system, and the correlation is therefore assumed to be relatively small and it is not considered explicitly. The magnitude of the estimated baseflow for the 36 stream reaches varies in the range 15–820 L/s. The uncertainty of the baseflow estimates is assumed to be dependent upon the absolute magnitude of the measured stream flow. The uncertainty is represented by four classes: 0–20, 20–100, 100–500, and above 500 L/s, where the coefficients of variation are estimated at 1.0, 0.5, 0.2, and 0.1, respectively. This means that the uncertainty of a measurement of, for example, 50 L/s is assumed to equal $s_Q=25$ L/s. Analysis of annual minimum daily discharge from nine gauging stations in the model area indicates that the specified uncertainties are reasonable.

In order to be consistent with the conditions represented by the baseflow estimates, hydraulic heads measured during the late summer (August 1995) were used for calibration. An analysis of climate data collected over a 10-year period shows that 1995 is an average year with respect to precipitation and evapotranspiration. Monthly observations of hydraulic head over an eight-year period in a well located in the central part of the model area show that the August 1995 head is lower than the long-term average head, but representative of the average summer hydraulic head (Henriksen et al. 1995). The hydraulic-head data used for calibration consist of 48 observations collected during a synchronous sampling round, thus eliminating uncertainty with respect to temporal fluctuations in groundwater levels. The data are considered to be very accurate, as the location and elevation of the monitoring wells were measured precisely. The standard deviation of the head data, $s_{h,i}$, is estimated to be 1 m. The density of the measurements is relatively uniform over depth, with observations in layers 2 through 10 in model A (11-layer model) and layers 5 through 20 in model B (21-layer model). The August 1995 head measurements are all from the central part of the model area and the range in observed heads is 33 m. A second set of hydraulic-head measurements that covers the entire model area was used for model validation and is described below. These observations were not used for calibration because they are considered to represent average annual conditions, and possess a

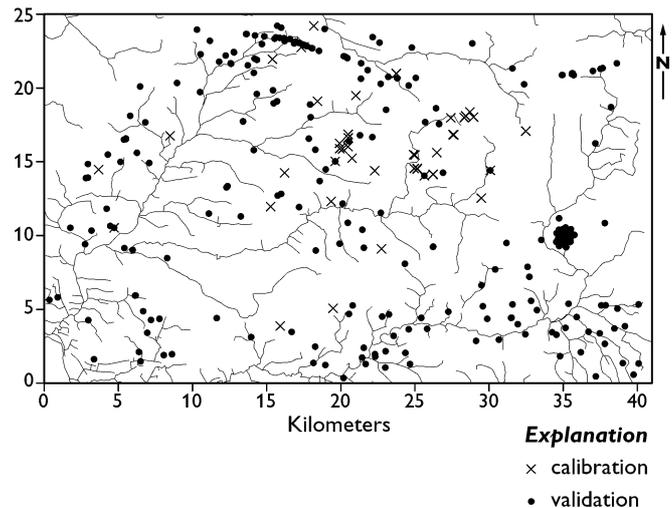


Fig. 6 Location of the monitoring wells used for model calibration and validation

relatively high observation uncertainty. The locations of the wells used for calibration and validation are shown in Fig. 6.

Optimization

Optimization of the flow models was performed using UCODE (Poeter and Hill 1998). The weighted least-squares objective function used in this study is expressed as

$$G(\underline{b}) = \sum_{i=1}^{N_h} \left(\frac{h_{obs,i} - h_i(\underline{b})}{s_{h,i}} \right)^2 + \sum_{j=1}^{N_Q} \left(\frac{Q_{obs,j} - Q_j(\underline{b})}{s_{Q,j}} \right)^2 \quad (1)$$

where \underline{b} is the vector of parameters to be estimated, $h_{obs,i}$ (L) and $h_i(\underline{b})$ (L) are the i th observed and simulated values of hydraulic head, respectively, N_h is the number of head observations, $s_{h,i}$ (L) is the standard deviation of the i th hydraulic-head measurement, $Q_{obs,j}$ ($L^3 T^{-1}$) and $Q_j(\underline{b})$ ($L^3 T^{-1}$) are the j th observed and simulated stream discharge, respectively, N_Q is the number of discharge observations, and $s_{Q,j}$ ($L^3 T^{-1}$) is the standard deviation of the j th discharge observation. The optimal parameter set \underline{b}^* is found by solving the optimization problem

$$\underline{b}^* = \text{Min}\{G(\underline{b})\} \quad (2)$$

A modified Gauss-Newton method is used for the iterative solution of the problem. The iteration process converges when (1) the relative change in parameter values is less than β_1 or (2) the value of the objective function G changes less than β_2 for three sequential iterations. A value of 0.01 was assigned to β_1 and β_2 . In the present study, all parameters of \underline{b} are log-transformed when estimated. Unconstrained optimization is performed, i.e. no upper and lower limits are imposed on the parameters. Prior information about the parameter

Table 1 Parameters included in the optimization, their initial estimates, and their corresponding field measurements. K_h and K_v represent horizontal and vertical hydraulic conductivity, respectively

No.	Description of parameter	Initial estimate	Field estimates		
			Observed range	Mean	95% Confidence intervals
1	K_h of Quaternary meltwater sands (m/s)	$5 \cdot 10^{-4}$	$5.7 \cdot 10^{-5}$ – $8.4 \cdot 10^{-4a}$	$4.7 \cdot 10^{-4}$	$2.4 \cdot 10^{-4}$ – $7.1 \cdot 10^{-4}$
2	K_v of Quaternary glacial till (m/s)	$2.5 \cdot 10^{-7}$	$5.0 \cdot 10^{-9}$ – $1.0 \cdot 10^{-6b}$	–	–
3	K_h of Tertiary quartz sand (m/s)	$1 \cdot 10^{-3}$	$1.7 \cdot 10^{-4}$ – $9.0 \cdot 10^{-3a}$	$1.5 \cdot 10^{-3}$	$1.5 \cdot 10^{-5}$ – $3.0 \cdot 10^{-3}$
4	K_v of Tertiary clay (m/s)	$5 \cdot 10^{-9c}$	–	–	–
5	Bed conductance of stream system (s^{-1})	$3 \cdot 10^{-3c}$	–	–	–
6	Multiplication factor for recharge	1.0^c	–	–	–

^a K estimated from pumping tests (Klitten et al. 1995)

^b K estimated from slug and infiltration tests (Nilsson et al. 2001)

^c The initial parameter estimate is from trial-and-error calibration results

values is used only to specify initial parameter estimates for the optimization process.

Calibration Parameters

Parameters considered for estimation included hydraulic-conductivity values for all of the sediment zones defined in the geological models, streambed conductance distributed on the basis of surficial geology, and recharge. Composite scaled sensitivities, calculated from a sensitivity analysis (Hill 1998) in which hydraulic-head and baseflow data were considered, showed that the model is sensitive to the hydraulic conductivity of four of the five sediment zones, streambed conductance, and a multiplication factor for recharge. Experience with trial-error-model calibration of the models showed that the fit to baseflow estimates for individual stream reaches could be improved by distributing streambed conductance with as few as two zones. However, the sensitivity analysis showed that regression was incapable of producing distributed streambed conductance values that were statistically significant. Thus, a single value of streambed conductance was considered for optimization in both models. The six parameters and their initial estimates are presented in Table 1.

Only the sensitive components of hydraulic conductivity are estimated. These are the horizontal hydraulic conductivity of the relatively high hydraulic-conductivity sediments (Quaternary meltwater sand and Tertiary quartz sand) and the vertical hydraulic conductivity of the relatively low hydraulic-conductivity sediments (Quaternary glacial till and Tertiary clay). The horizontal and vertical hydraulic conductivity are subsequently linked through the anisotropy ratio, $\alpha = K_v/K_h$. A value of $\alpha = 0.1$ is used for all hydrofacies except for the Quaternary glacial till, where a value of $\alpha = 0.5$ is specified. The anisotropy ratios were assigned based upon the results of trial-and-error calibration by Henriksen et al. (1995).

Performance Criteria

The reliability of the estimated parameters is quantified by UCODE assuming model linearity. Here, the calculated standard deviation, s_b , of the estimated parameter,

b^* , is used to calculate linear 95% confidence intervals for the estimates, assuming normally distributed errors.

The optimized steady-state models are evaluated with respect to the match between observed and simulated dependant variables. The root mean squared weighted residuals ($RMSW$) with respect to hydraulic head is calculated as

$$RMSW_h = \sqrt{\frac{1}{N_h} \sum_{i=1}^{N_h} \left(\frac{h_{obs,i} - h_i(b^*)}{s_{h,i}} \right)^2} \quad (3)$$

A similar measure is used for the stream-discharge residuals

$$RMSW_Q = \sqrt{\frac{1}{N_Q} \sum_{j=1}^{N_Q} \left(\frac{Q_{obs,j} - Q_j(b^*)}{s_{Q,j}} \right)^2} \quad (4)$$

Calibration Results

The optimized parameter estimates for the two models are presented in Table 2. The optimized hydraulic-conductivity values for the Quaternary sands, till, and Tertiary coarse sand are within the range of field estimates in both models, with the exception of the Tertiary coarse sand in model B, which is a factor of four less than the lowest conductivity estimated for this zone from pumping-test results. In both models, the upper 95% confidence interval of the optimized hydraulic conductivity for the Quaternary sands overlaps with the lower 95% confidence interval of the mean field estimate. The upper 95% confidence interval of the optimized estimate for the Tertiary coarse sand overlaps with the mean field estimate in model A, and with the lower range of field estimates in model B. The uncertainty of the estimated parameters is generally low. However, the confidence intervals for the estimated vertical hydraulic conductivity of the glacial till and Tertiary clay (parameters 2 and 4) in model B range over four orders of magnitude, indicating that reliability of these parameters is relatively poor. The maximum correlation coefficients (Table 2) indicate that correlation between parameters is of no concern. According to Hill

Table 2 Optimized parameter estimates, uncertainty, and maximum correlation

Parameter	Model A			Model B		
	Estimate	95% Confidence intervals	Maximum correlation coefficient ^a	Estimate	95% Confidence intervals	Maximum correlation coefficient ^a
K_h of Quaternary meltwater sands (m/s)	$7.2 \cdot 10^{-5}$	$2.9 \cdot 10^{-5}$ – $1.8 \cdot 10^{-4}$	–0.77 (4)	$1.8 \cdot 10^{-4}$	$1.3 \cdot 10^{-4}$ – $2.4 \cdot 10^{-4}$	0.80 (6)
K_v of Quaternary glacial till (m/s)	$1.2 \cdot 10^{-8}$	$3.9 \cdot 10^{-9}$ – $3.7 \cdot 10^{-8}$	0.36 (4)	$4.1 \cdot 10^{-7}$	$3.5 \cdot 10^{-9}$ – $4.8 \cdot 10^{-5}$	–0.56 (5)
K_h of Tertiary quartz sand (m/s)	$6.1 \cdot 10^{-4}$	$2.9 \cdot 10^{-4}$ – $1.2 \cdot 10^{-3}$	–0.73 (1)	$3.7 \cdot 10^{-5}$	$1.7 \cdot 10^{-6}$ – $8.1 \cdot 10^{-4}$	–0.34 (4)
K_v of Tertiary clay (m/s)	$4.5 \cdot 10^{-8}$	$2.7 \cdot 10^{-8}$ – $7.5 \cdot 10^{-8}$	–0.77 (1)	$2.7 \cdot 10^{-8}$	$4.3 \cdot 10^{-10}$ – $1.7 \cdot 10^{-6}$	–0.34 (3)
Bed conductance of stream system (s^{-1})	$2.8 \cdot 10^{-3}$	$1.7 \cdot 10^{-3}$ – $4.6 \cdot 10^{-3}$	–0.64 (1)	$1.9 \cdot 10^{-3}$	$1.4 \cdot 10^{-3}$ – $2.7 \cdot 10^{-3}$	–0.56 (2)
Multiplication factor for recharge	0.82	0.71–0.96	0.61 (3)	1.0	0.86–1.17	0.80 (1)

^a The number in parentheses indicates the parameter with which the maximum correlation coefficient was obtained

Table 3 Match between observed and simulated heads and stream discharge

	Calibration	
	$RMSW_h$ (-)	$RMSW_Q$ (-)
Model A	1.9	2.1
Model B	2.1	2.2

(1998), correlation coefficients less than 0.95 suggest that the estimated parameters are likely to be unique.

The match between observed and simulated heads and baseflow is presented in Table 3. Both models perform well with respect to the observed hydraulic heads and baseflow estimates. $RMSW_h$ values are 1.9 and 2.1 m for models A and B, respectively. The ratio of $RMSW_h$ values from both models to the range of head in the calibration dataset (33 m) is about 6% and errors are randomly distributed. $RMSW_Q$ values are 2.1 and 2.2 for models A and B, respectively. However, despite the good fit to the total baseflow, both models underpredict the baseflow by about 50% in one stream reach that accounts for 15% of the total baseflow. Models A and B underpredict the total estimated baseflow by 25 and 31%, respectively.

Layer and Model Average Hydraulic-Conductivity Values

Simpler hydraulic-conductivity distributions were calculated from the optimized parameter estimates with the objective of generating more uniform hydraulic-conductivity distributions that preserve the averages of the optimized values. Flow models constructed from the simpler hydraulic-conductivity distributions were then used to evaluate how the representation of geological complexity affects model predictions.

The optimized hydraulic-conductivity values from each of the two flow models were used to calculate the mean hydraulic conductivity (horizontal and vertical) for each model layer and for each model as a whole. The horizontal and vertical hydraulic-conductivity values were calculated as the arithmetic and harmonic mean, respectively. The layer mean horizontal hydraulic-conductivity values from model B decrease systematically with depth, while those from model A show considerable

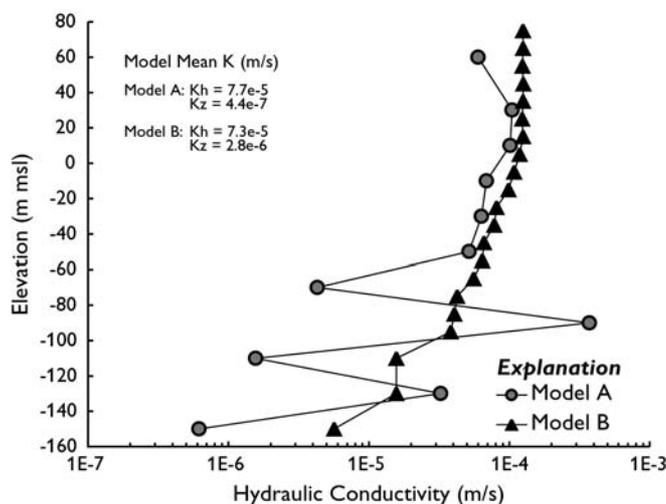


Fig. 7 Mean horizontal hydraulic-conductivity values for each model layer calculated using the optimized parameters from models A and B. Mean hydraulic-conductivity values for the entire model domain for both models are also shown

variation with depth, as shown in Fig. 7. The mean horizontal hydraulic conductivity calculated for the entire model domain is nearly identical for the two models ($K_h=7.7 \cdot 10^{-5}$ m/s for model A, and $K_h=7.3 \cdot 10^{-5}$ m/s for model B), while the model mean vertical hydraulic conductivity varies by a factor of six between the two models ($4.4 \cdot 10^{-7}$ for model A, and $2.8 \cdot 10^{-6}$ m/s for model B). It is important to note that the layer mean hydraulic-conductivity values represent the average of several different sedimentary zones, and therefore should not be compared to the field estimates of hydraulic conductivity for the individual sedimentary zones.

Model Validation and Comparison

Four additional flow models were constructed using the layer and model mean hydraulic-conductivity values from the two optimized models. These models are referred to as A_L , A_M , B_L and B_M , where the subscripts L and M refer to use of layer mean and model mean hydraulic-conductivity values, respectively. Forward runs were performed and the results from all six of the flow models (models A, A_L ,

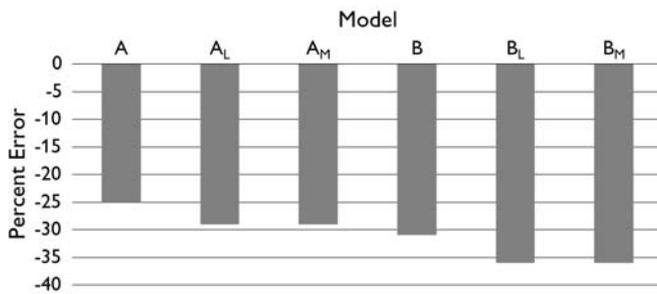


Fig. 8 Percent error of simulated versus observed total baseflow for the six flow models

A_M, B, B_L and B_M) were compared with 251 hydraulic-head measurements archived at the Geological Survey of Denmark and Greenland, and the baseflow estimates. The hydraulic-head observations range over 72 m and represent measurements made during different times of the year over a 30-year period. A comparison of these data with long-term hydrographs and head data from the synchronous sampling round (August 1995) showed that the 251 head observations represent a good approximation of average or steady-state conditions. However, there is a relatively high degree of uncertainty in the dataset due to seasonal and annual variations in the groundwater levels. It is estimated that these transient effects result in an uncertainty of $s_h=2$ m, where s_h is the standard deviation of the head data.

The six models all simulate the heads from the validation dataset reasonably well, as indicated by the statistics presented in Table 4. Error is randomly distributed in models A, B_L and B_M, while models A_L, A_M and B underpredict the heads at monitoring wells with observed values that are greater than 60 m m.s.l. Models A_L and A_M simulate the observed heads slightly better than model B. This is surprising in that it was expected that the models that more accurately represent the sedimentary heterogeneity would produce a better overall fit to the observed heads. The simulated total baseflow in all of the models is 25–35% less than the estimated baseflow, as shown in Fig. 8. The best fit to baseflow was obtained from model A, which also performed the best with respect to the observed heads. Differences between the models become more apparent upon comparison of simulated and estimated baseflow along the individual stream reaches. Field data show that there is a net gain of water along all of the stream reaches during low-flow conditions. This is well represented by the two optimized models (A and B), while 15–25% of the stream reaches in the layer and model mean hydraulic-conductivity models show a net loss of water.

Predictive Simulations

The six flow models were used to predict the steady-state impact of pumping 2.8×10^6 m³/year from a well field located within the central erosional feature. The flow

Table 4 Hydraulic-head statistics from the model validation and comparison. Head range is 72 m and $n=251$

	A	A _L	A _M	B	B _L	B _M
$RMSW_h$ (-)	1.6	1.75	1.74	1.84	1.92	1.92
RMS_h /head range (%) ^a	4.4	4.8	4.8	5.1	5.3	5.3

^a RMS_h indicates unweighted root mean squared value for hydraulic heads

Table 5 Predicted particle-travel times from the six flow models

Model	Median travel time (year)	Travel time (95% fractile) (year)
A	155	970
A _L	138	1,115
A _M	177	664
B	118	553
B _L	582	2,113
B _M	863	2,144

solution was used in the simulation of particle tracking and solute transport, and the results are presented below.

Capture Zones and Travel Times

The capture zone of the well field was simulated in each model by performing particle tracking using MODPATH (Pollock 1989). A total of 250 particles was released in each cell containing an extraction well and tracked backward through time. An effective porosity of 0.25 was assigned uniformly throughout the model. The resulting capture zones of the well field are shown in Fig. 9. They are similar, but do differ somewhat in shape and extent to the northeast far from the well field. Evers and Lerner (1998) show that simulated capture zones can vary significantly, given the uncertainties in parameter-value estimates and boundary conditions for any model. However, the results show that for this example, simple models of heterogeneity produce capture zones similar to more complex models if the mean hydraulic conductivity of the models is similar. This is consistent with the work of Feyen et al. (2001) who show that the mean hydraulic conductivity and variance are the most important parameters in the determination of stochastic capture zones.

Particle-travel times were evaluated by computing a median value and 95% fractile for the time taken for all particles released at the extraction wells to exit the flow system at the water table. The results are presented in Table 5 and display a wide range of values, showing that particle-travel times are sensitive to parameter values. The particle-tracking analysis shows that while predictions of capture zones are robust if the mean hydraulic conductivity of the flow system is well represented, travel times within the capture zone predicted by models calibrated to head and flow data possess a large degree of uncertainty due to aquifer heterogeneity.

Fig. 9 Capture zones predicted by the six models

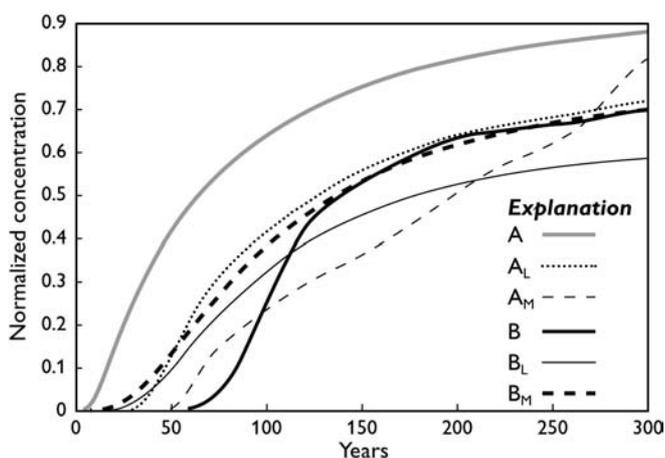
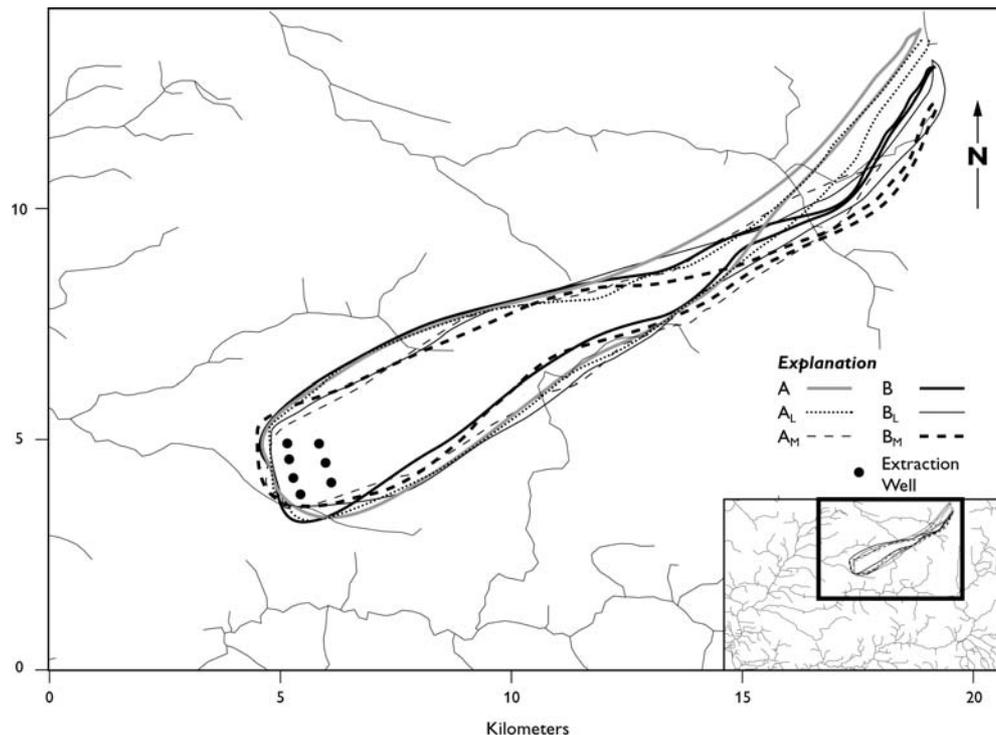


Fig. 10 Simulated breakthrough curves for the six models

Solute Transport

The effect of zonation on regional transport was evaluated by simulating the transport of a conservative solute using the steady-state well-field solution from the six models and MT3DMS (Zheng 1998). A constant concentration-boundary condition was applied to the uppermost active layer in each model over an area encompassing the simulated capture zones. Simulations were performed for a 300-year period during which time the concentrations were observed at each pumping well. A flux-weighted well-field concentration was calculated for each model and the resulting breakthrough curves (BTCs) are presented in Fig. 10. Predicted first arrivals, defined by $C/C_0=0.01$, from the optimized models (A and B) vary by

over an order of magnitude and represent the end members for all of the models. Predicted BTCs from all of the models vary significantly, and no distinguishable patterns related to different zonations are readily recognizable. The wide range of predicted first arrivals and concentrations from the six models highlights the uncertainty associated with predictions of solute transport from models that are calibrated using only hydraulic-head and flow data.

Discussion and Conclusions

The large datasets available for model construction and calibration in this study provide the opportunity to compare the effects of different interpretations of hydraulic-conductivity distributions on the results of inverse calibration and predictions. Two different geological models were constructed using over 5,400 borehole logs. Water-balance estimates indicate that discharge to streams and extraction wells accounts for about 90% of the recharge to the model area. Furthermore, the dense network of stream-monitoring stations, from which base-flow estimates are derived, provides a detailed description of the flow distribution within the model area. These data, in combination with 48 head measurements used to constrain the inverse model and 251 head measurements used for model validation and comparison, allowed the examination of the effects of zonation on model calibration and predictions.

The two optimized flow models produced a similar quality of calibration, despite the differences in the geological models from which the models were con-

structed, showing that the solution to the inverse problem is non-unique. Forward runs with the flow models constructed using the layer mean and model mean hydraulic-conductivity values calculated from optimized parameters also produced acceptable fits to head and flow data, further highlighting the non-uniqueness of the problem. The lack of uniqueness may have been overcome through the use of prior information (Carrera and Neuman 1986), by using more sophisticated methods to delineate hydraulic-conductivity zones, such as those presented by Ritzi et al. (1994), or by collecting additional data for calibration. It is concluded from the work of Hill et al. (1998) and the results presented here that non-uniqueness is probably unavoidable in complex groundwater problems. Hill et al. (1998) also conclude that this lack of uniqueness does not indicate that models produce inaccurate results, as prediction accuracy depends upon the quality of available data and calibration methodology.

The optimized hydraulic-conductivity values for the Quaternary meltwater sands and Tertiary coarse sands in models A and B are in good overall agreement with field estimates. The 95% confidence intervals of the optimized and the mean field estimates overlap, with the exception of the Tertiary coarse sand in model B, where the upper 95% confidence interval of the optimized value overlaps with the lower range of the field estimates. This is in contrast to the results presented by Hill et al. (1998) and Barth et al. (2001), where optimized hydraulic-conductivity values differ from measured values due to differences between the scale of observations and the model grid size, and measurement inaccuracies. Sun and Yeh (1985) show that improper representation of parameter zones causes estimated parameter values to be incorrect. The results of the present study show that representative parameter estimates can be obtained from inverse models using different zonations. Furthermore, the mean horizontal hydraulic-conductivity values calculated for the entire model domain were essentially identical for the two optimized models, whereas the mean vertical hydraulic-conductivity values for the model domain differed by a factor of six. This indicates that while model structure does affect optimized values for individual zones, sufficient calibration data, especially flows, constrain inverse models so that mean hydraulic-conductivity values of the entire model domain are well represented.

Predictive simulations show that for this example, simple models of heterogeneity produce capture zones similar to more complex models. The results also show that different models of heterogeneity may produce very different predictions of travel time and solute breakthrough. This indicates that inverse modeling combined with different geological models can provide a measure of the reliability of capture zone, travel time, and solute-breakthrough predictions.

Acknowledgements This study was conducted as part of the project entitled "National Water Resources Model" that was cooperatively funded by the Danish Environmental Ministry and

the Geological Survey of Denmark and Greenland. We appreciate the comments and suggestions of John Barker, Jens Christian Refsgaard, and two anonymous reviewers. The authors thank Mette Dahl for assistance in analyzing the baseflow data, and Per Nygaard, Jørn Morthorst, and Bente Mörch for assistance with the geological models. We are also grateful to Kristian Rasmussen for preparing the technical illustrations.

References

- Bair ES, Safreed CM, Stasny EA (1991) A Monte Carlo based approach for determining travel time-related capture zones of wells using convex hulls as confidence regions. *Ground Water* 36:849–855
- Barth GR, Hill MC, Illangasekare TH, Rajaram H (2001) Predictive modeling of flow and transport in a two-dimensional intermediate-scale, heterogeneous porous medium. *Water Resources Res* 37:2503–2512
- Bhatt K (1993) Uncertainty in wellhead protection area delineation due to uncertainty in aquifer parameter values. *J Hydrol* 149:1–8
- Carrera J, Neuman S (1986) Estimation of aquifer parameters under transient and steady state conditions: 2. Uniqueness, stability and solution algorithms. *Water Resources Res* 22:211–227
- Cole E, Silliman SE (2000) Utility of simple models for capture zone delineation in heterogeneous unconfined aquifers. *Ground Water* 38:665–672
- Dybkjaer K, Rasmussen ES, Piasecki S (2001) Oligocene and Miocene stratigraphy in Vejle County (in Danish). *Geol Surv Denmark Greenland Rep* 2001/104
- Evers S, Lerner DN (1998) How uncertain is our estimate of a wellhead protection zone? *Ground Water* 36:49–57
- Feyen L, Beven KJ, De Smedt F, Freer J (2001) Stochastic capture zone delineation within the generalized likelihood uncertainty estimation methodology: conditioning on head observations. *Water Resources Res* 37:625–638
- Henriksen HJ, Harrar W, Morthorst J, Nyegaard P, Dahl M (1995) Groundwater exploration for Esbjerg Waterworks Phase IV—Model (in Danish). *Geol Surv Denmark Greenland Rep* 51/1995
- Hill MC (1998) Methods and guidelines for effective model calibration. *US Geol Surv Water Resources Inv Rep* 98-4005, Denver, Colorado, USA
- Hill MC, Cooley RL, Pollock DW (1998) A controlled experiment in ground water flow model calibration. *Ground Water* 36:520–535
- Klitten K, Morthorst J, Czako T, Henriksen HJ, Harrar WG (1995) Groundwater exploration for Esbjerg Waterworks Phase III (in Danish). *Geol Surv Denmark Greenland Rep* 25/1995
- McDonald MG, Harbaugh AW (1989) A modular three-dimensional finite-difference ground-water flow model. *US Geol Surv Open File Rep* 83-875
- Nilsson B, Sidle RC, Klint KE, Bøggild, CE, Broholm K (2001) Mass transport and scale-dependent hydraulic tests in a heterogeneous glacial till-sandy aquifer system. *J Hydrol* 243:162–179
- Poeter EP, Hill MC (1998) Documentation of UCODE: a computer code for universal inverse modelling. *US Geol Surv Water Resources Inv Rep* 98-4080, Denver, Colorado, USA
- Pollock DW (1989) Documentation of computer programs to compute and display pathlines using results from the U.S. Geological Survey modular three dimensional finite difference ground-water flow model. *US Geol Surv Open File Rep* 89-381
- Rayne TW, Bradbury KR, Muldoon MA (2001) Delineation of capture zones for municipal wells in fractured dolomite, Sturgeon Bay, Wisconsin USA. *Hydrogeol J* 9:432–450
- Ritzi RW, Jayne DF, Zahradnik AJ, Field AA, Fogg GE (1994) Geostatistical modeling of heterogeneity in glaciofluvial, buried-valley aquifers. *Ground Water* 32:666–674

- Springer AE, Bair ES (1992) Comparison of methods used to delineate capture zones of wells. 2. Stratified-drift buried-valley aquifer. *Ground Water* 30:908–917
- Stockmarr J (ed) (2001) Groundwater monitoring report (in Danish). Geol Surv Denmark Greenland Rep 87-7871-096-0
- Sun N-Z, Yeh WW-G (1985) Identification of parameter structure in groundwater inverse problem. *Water Resources Res* 21:869–883
- van Leeuwen M, Butler AP, Stroet CBMT, Tompkins JA (2000) Stochastic determination of well capture zones conditioned on regular grids of transmissivity measurements. *Water Resources Res* 36:949–957
- Varljen MD, Shafer JM (1991) Assessment of uncertainty in time-related capture zones using conditional simulation of hydraulic conductivity. *Ground Water* 29:737–748
- Vassolo S, Kinzelbach W, Schäfer W (1998) Determination of well head protection zone by stochastic inverse modelling. *J Hydrol* 206:268–280
- Zheng CM (1998) A modular three-dimensional multispecies transport (MT3DMS) model, release DoD_3.00.A. Waterw Exp Stn US Army Corps Eng, Vicksburg, Miss