
Flow and transport predictions during multi-borehole tests in fractured chalk using discrete fracture network models

Daniel Kurtzman · Ronit Nativ · Eilon M. Adar

Abstract Multi-borehole pumping and tracer tests on the 10 to 100-m scale were conducted in a fractured chalk aquitard in the Negev Desert, Israel. Outcrop and core fracture surveys, as well as slug tests in packed-off intervals, were carried out at this site to obtain the parameters needed for construction of a stochastic discrete fracture network (DFN). Calibration of stochastic DFNs directly to the multiple borehole test data was inadequate. Instead, two equivalent deterministic DFN flow models were used: the vertical-fractures (VF) model, consisting of only vertical fractures, and the fractures' intersections (INT) model, consisting of vertical and horizontal fractures with enhanced transmissivity at their intersections. Both models were calibrated against the multi-borehole response of one pumping test and their predictions were tested against three other independent pumping tests. The average accuracies of all transient drawdown predictions

of the VF and INT models were 65 and 66%, respectively. In contrast to this equality in average drawdown predictions of both models, the INT model predicted better important breakthrough curve features (e.g., first and peak arrival times), than the VF model. This result is in line with previously assumed channeled flow, derived from analytical analysis of these pumping and tracer tests.

Résumé Des tests de pompage et de traçage sur plusieurs puits à l'échelle de 10 à 100 m ont été réalisés dans un aquitard crayeux fracturé du Désert du Négev, Israël. Les levés des affleurements et des fractures traversés par les carottages, ainsi que des slug tests réalisés sur des intervalles isolés par des packers, ont été réalisés sur le site pour obtenir les paramètres nécessaires à la construction d'un réseau stochastique discret de fractures (DFN, en anglais). Une calibration des DFNs stochastiques, directement sur les données des forages, était inadéquate. A la place de cette approche, deux modèles d'écoulement de DFN, équivalents et déterministes, ont été utilisés : le modèle de fractures verticales (VF) qui comme son nom l'indique ne comprend que des fractures verticales, et le modèle à intersections de fracture (INT), comprenant des fractures verticales et horizontales et des transmissivités augmentées aux intersections. Les deux modèles ont été calibrés suivant la réponse d'un test de pompage suivi sur plusieurs puits, et leurs prédictions ont été testées suivant trois autres essais de pompage. L'exactitude moyenne de toutes les prédictions transitoires de rabattement des modèles VF et INT était de 65 et 66%, respectivement. Contrairement à la coïncidence dans les prédictions de rabattement moyen des deux modèles, le modèle INT prédit mieux les caractéristiques importantes des courbes de restitution (e.g., temps de la première arrivée et du pic de concentration) comparé au modèle VF. Le résultat est consistant avec des écoulements de chenaux supposés précédemment, dérivés d'analyses analytiques de ces tests de pompage et de traçage.

Resumen Se realizaron pruebas de bombeo en varios sondeos y pruebas con trazadores en una escala que varió de 10 a 100m en un acuitardo fracturado calcáreo en el Desierto Negev, Israel. Se llevaron a cabo en el sitio levantamientos en afloramientos y fracturas de núcleo, así como pruebas con remoción/adición rápida de agua en

Received: 29 January 2007 / Accepted: 21 June 2007
Published online: 24 July 2007

© Springer-Verlag 2007

Ronit Nativ, deceased, may her memory be blessed.

D. Kurtzman (✉)
Bureau of Economic Geology, Jackson School of Geosciences, The University of Texas at Austin,
10100 Burnet Rd.
Austin, TX 78758-4445, USA
e-mail: daniel.kurtzman@beg.utexas.edu
Tel.: +1-512-4759527
Fax: +1-512-4710140

R. Nativ
The Seagram Center for Soil and Water Sciences, Faculty of Agricultural, Food and Environmental Quality Sciences,
The Hebrew University of Jerusalem,
P.O. Box 12,
Rehovot, 76100, Israel

E. M. Adar
The J. Blaustein Institutes for Desert Research,
Zuckerberg Institute for Water Research and Department of Environmental Geological Sciences,
Ben-Gurion University of the Negev,
Sede Boker Campus
Negev, 84990, Israel

intervalos sin empaque, para obtener los parámetros necesarios en la construcción de una red de fracturas discreta estocástica (RFD). La calibración de RFDs estocásticas directamente con los datos de pruebas de sondeo múltiples fue inadecuada. En vez de esto, se utilizaron dos modelos de flujo RFD determinísticos equivalentes: el modelo de fracturas verticales (FV), que consiste únicamente de fracturas verticales, y el modelo de intersección de fracturas (INT), que consiste de fracturas horizontales y verticales con transmisividad realzada en sus intersecciones. Ambos modelos fueron calibrados en relación con la respuesta de una prueba de bombeo en sondeos múltiples y sus pronósticos fueron evaluados en relación a otras tres pruebas de bombeo independientes. Las imprecisiones promedio de todas las predicciones transitorias de descenso de los modelos INT y FV fueron 65 y 66%, respectivamente. En contraste con esta igualdad en las predicciones promedio de descenso de ambos modelos, el modelo INT pronosticó mejor las características de curvas de avance importantes (por ejemplo, tiempos de llegada inicial y tiempos pico) que el modelo FV. Este modelo es congruente con el flujo de canales que se asumió previamente derivado del análisis analítico de las pruebas con trazadores y pruebas de bombeo.

Keywords Fractured rocks · Discrete fracture network (DFN) · Numerical modeling · Israel · Solute transport

Introduction

Since the early 1960s, hydrogeologists and petroleum engineers have invested significant effort in better understanding flow and transport through fractured porous media. Double-porosity models such as the one proposed by Warren and Root (1963), in which the matrix serves as the storage volume and the fractures as the flow paths, were considered suitable for solving problems associated with water and oil production. Since the 1980s, however, when the isolation of radioactive waste in low-permeability but inevitably fractured formations became a pressing environmental problem, the matrix permeability has come to be considered less important in flow and transport problems. Consequently, flow and transport models based only on the discrete fracture network (DFN) were developed (Andersson and Dverstorp 1987; Long and Billaux 1987; Cacas et al. 1990a, b), requiring realistic description of the geometry of the fracture network. Because knowledge of this network's three-dimensional geometry is always limited, stochastic methods were adopted for its description (Chile's and de Marsily 1993). The importance of a discrete representation of the fracture network was also recognized by continuum modelers, who incorporated discrete fractures into the continuous matrix, constructing hybrid models (Therrien and Sudicky 1996; Lee et al. 2001). The general theory and methodologies of DFNs are discussed at length by Chile's and de Marsily (1993), Priest (1993), the National

Research Council (1996) and Dershowitz et al. (1998), among others, and will not be discussed here. Instead, a site-specific implementation of these methodologies, in which DFN models were used to predict the results of multi-borehole pumping and tracer experiments, will be described in detail.

Stochastic DFN models are computationally intensive and thus limited to applications involving small volumes of rock. Selroos et al. (2002) suggested that regional-scale flow and transport in a fractured domain should be modeled with the stochastic continuum approach (Neuman 1987; Tsang et al. 1996; Ando et al. 2003), while the site scale should be modeled by stochastic DFN. Many studies discussed the up-scaling of DFN models to a continuum representation (Long and Billaux 1987; Cacas et al. 1990a; Niemi et al. 2000; Benke and Painter 2003; Öhman and Niemi 2003; Min et al. 2004). Stochastic DFN provided background permeability, complementing the few deterministic, site-scale, conductive, structural features (Andersson et al. 2002). Cvetkovic et al. 2004 used stochastic DFN models to calculate transport from one side of a simulation box to the opposite side. In the present study, an attempt was made to calibrate stochastic DFNs directly to head data from several boreholes. Computation intensity, fracture-connectivity and non-deterministic calibration issues that arose during this calibration were resolved by moving from stochastic DFN realizations to equivalent deterministic DFNs, for simulating the observations of multi-borehole pumping and tracer tests.

The objective of this study is to test the predictions obtained by DFN models. Namely, the predictions of two deterministic DFN flow models which represent two different conceptual understandings of the dominant conduit system in the fractured chalk, were examined. The two models were tested both for predictions of transient-head drawdowns during multi-borehole pumping tests and break-through curves (BTCs) from tracer tests.

Field site and fracture characterization

Site description and supportive observations

The field site is located in the Negev Desert, Israel, hosting a large industrial complex (Fig. 1) and underlain by densely fractured Eocene chalk (100–280 m thick) with low matrix hydraulic conductivity ($\sim 1 \times 10^{-3}$ m/day; Nativ et al. 1999, 2003). Groundwater contamination by the industrial complex and the natural lateral groundwater drainage to the Coastal Plane Aquifer in the west are sources of concern. The field site is located east of the confluence of the Naim and Hovav ephemeral washes (Fig. 1). Two perpendicularly oriented, natural outcrops, located nearby, enabled one-dimensional and two-dimensional fracture surveys. Four slanted open boreholes—RH11, RH11a, RH11b and RH11c—served for the various hydraulic tests (see borehole drilling data in Table 1). Three additional, shorter boreholes—RH11d, RH11e and RH11f—served as hydraulic-head controls for

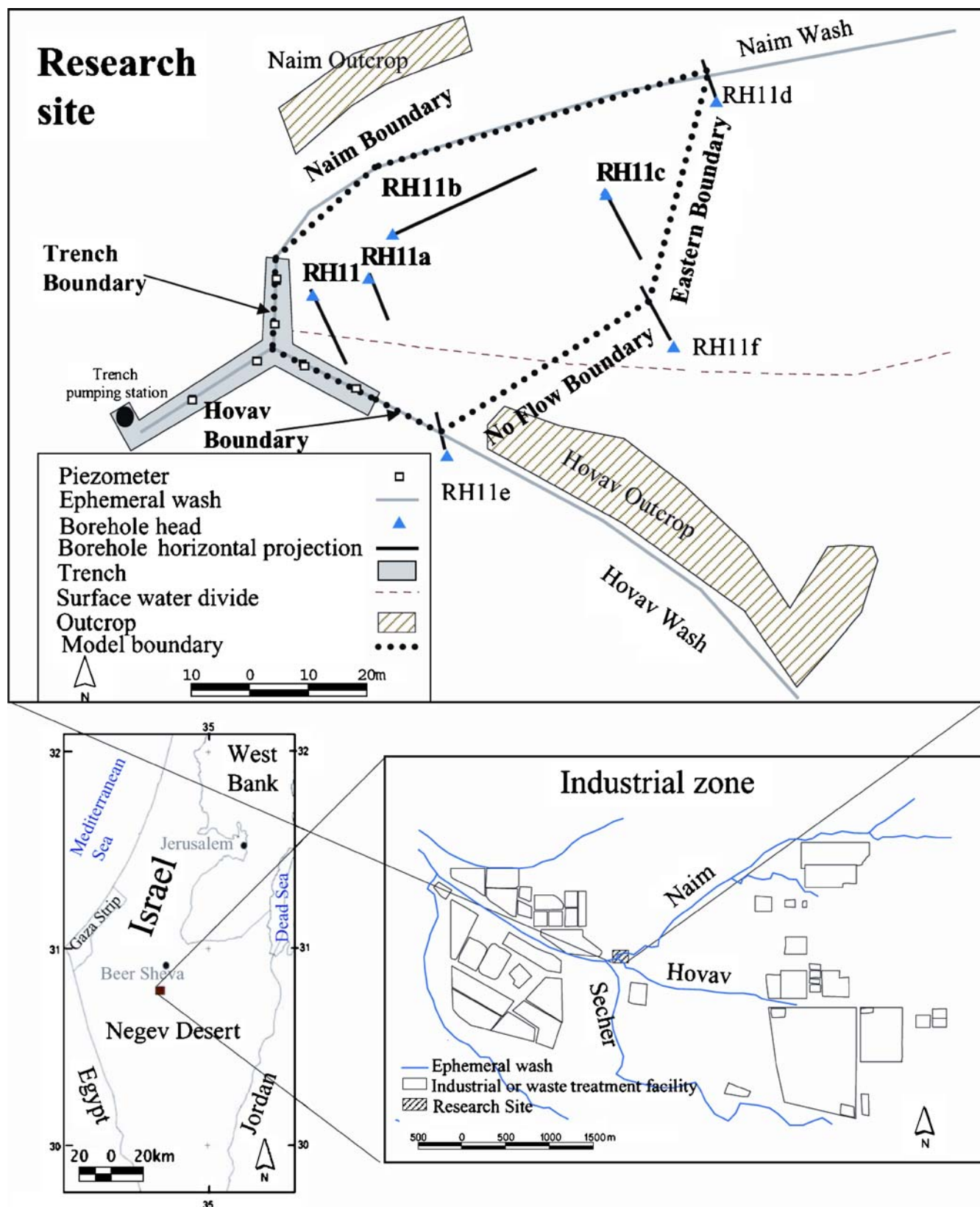


Fig. 1 The research site in the northern Negev Desert, Israel (physical entities and model boundaries)

Table 1 Details of the inclined boreholes used for hydraulic tests

Borehole name	Drilling method	Length (m)	Plunge (degrees from horizontal)	Trend (degrees from north)	Borehole radius (cm)
RH11	Coring	40	68	155	4.3
RH11a	Coring	25	68	155	4.3
RH11b	Dry augering	40	45	65	6.9
RH11c	Dry augering	40	68	155	6.9

the numerical model's boundary nodes. All the inclined boreholes were cased in the upper few meters against the loose, granular overburden, but no casing was installed against the fractured chalk. An excavated trench, located at the intersection of the washes and used to pump and treat the seeping groundwater, drains this field site. Four piezometers placed in the trench enabled monitoring of the hydraulic head along this boundary (Fig. 1).

Fractures observed in cores and borehole logs (e.g., caliper, video, and heat-pulse logs) of boreholes RH11 and RH11a and an identically inclined, 100-m-deep borehole located in between these two boreholes (not drawn in Fig. 1), were investigated by Nativ et al. (2003) with respect to their ability to conduct flow. The main results of their study, of relevance to the present multi-borehole flow and transport work were: (1) conductive fractures lay within the upper 30 m of the chalk; (2) conductive borehole intervals always contained fractures striking 048–060°; (3) intersections of sub-vertical fractures with a bedding plane were observed in some of the most conductive intervals. The first observation supported the decision to perform the multi-borehole tests in relatively shallow boreholes and assume that hydraulic connections to deeper layers can be neglected. The second observation supported the drilling of most boreholes at the same orientation (Fig. 1), thereby increasing the chances of intersecting conductive fractures striking at ~060°. The third observation coincides with an outcrop observation reported by Kurtzman et al. (2003), describing the apparent aperture of horizontal fractures along bedding planes at this site as varying frequently from 0 (completely clogged) to centimeter-scale wide at some intersections with vertical fractures. Kurtzman et al. (2003) also reported the following similar one-dimensional fracture densities (number of fractures per scanline or core length) for the ~060°-striking fracture set: 1.4, 1.4, and 1.3 m⁻¹ measured in the Hovav outcrop, and RH11 and RH11a boreholes, respectively.

The lateral and vertical proximity of the outcrops and boreholes at this site (Fig. 1) and the aforementioned observations, strongly support the assumption that fractures surveyed on outcrops and fractures, in which flow and transport between boreholes occur, are from the same fracture populations, thereby making the subsequent fracture characterization worthwhile.

Geometrical and hydraulic characterization of fractures

Scanline surveys were carried out at the Hovav and Naim outcrops (Fig. 1). Strike and dip angles of all fracture

traces longer than 0.2 m intersecting the scanlines were measured by a geological compass. The letters W (west) or E (east) were appended to the dip angle to signify the dip's direction. Orientations of all fractures observed in the RH11 and RH11a (inclined) cores were obtained by means described in Nativ et al. (2003). These fracture-surface orientations were converted to pole-to-plane trends and plunges (Priest 1993). Mean angle and Fisher distribution for each set were calculated using the interactive set identification system (ISIS) module included in the FracMan package (Dershowitz et al. 1998).

Fracture traces were mapped using large rectangular windows (30×11.5 m and 22.4×6.5 m on the Hovav and Naim outcrops, respectively) by drawing trace intersections or the intersections of their extensions with four measuring tapes nailed to the chalk outcrop. The resulting maps were designed to best represent the fracture traces intersecting a planar outcrop (the outcrops are not planar in the field). In the FracMan package (Dershowitz et al. 1998), a large number of fractures were (forward) simulated with a known orientation distribution and an assumed radius distribution. A virtual outcrop, identical to the one surveyed in the field (in terms of orientation, width and length), was then inserted into the same simulation box together with the simulated fractures. The trace-length distribution left by the simulated fractures on the virtual outcrop was compared with the trace-length distribution measured in the field and the assumed radius was adjusted until a satisfactory fit was achieved.

Three-dimensional fracture densities (total fracture area/rock volume) were estimated by forward-simulating the three-dimensional densities (following the fitting of orientation and radius distributions) until a good fit between the two-dimensional densities (total traces length/outcrop area) of the simulated and actual trace maps was achieved.

Interval transmissivities were estimated using 35 slug tests carried out in 2.0 or 2.5-m packed-off intervals in the RH11, RH11b and RH11c boreholes (Kurtzman et al. 2003). Head responses were interpreted using the Bouwer and Rice (1976) method (Brown et al. 1995). Fracture-transmissivity distributions were calculated with the OXFILET procedure (Dershowitz et al. 1998), which uses the interval transmissivities and one-dimensional fracture densities (obtained in this case from outcrop scanlines and core analysis).

A semivariogram of the fracture areal density (number of fractures/area) showed almost pure nugget for one fracture set, implying a uniform spatial distribution of fractures in this set. The semivariogram of the second fracture set had two minima, one at the shortest distance-lag and one at the distance-lag equal to the distance

between the shear zones, implying a “shear zone” spatial distribution model (Kurtzman et al. 2003).

Table 2 summarizes the results of these characterization surveys and analysis for the two sub-vertical fracture sets found on site (the dominant set and the secondary set). The results were summarized in three probability density functions (PDF) for (1) fracture orientation (Fisher distribution), (2) fracture length (exponential distribution) and (3) fracture transmissivity (lognormal distribution), a constant three-dimensional fracture density and a fracture spatial distribution model. Fracture storativity was assumed constant and equal to $1e-4$. These parameters enabled the construction of stochastic DFN realizations.

Hydraulic tests and modeling methods

Pumping tests

Four pumping tests were performed in the four boreholes RH11, RH11a, RH11b and RH11c (Fig. 1). During these tests, one borehole at a time was pumped, while drawdown was monitored in the other observation boreholes as well as in the pumping borehole. A 2" (5.08 cm) submersible pump was used, positioned against the most permeable zone in each borehole, as indicated by the results of slug tests in packed-off intervals in each of these boreholes. A constant pumping rate was maintained throughout the tests. The pumping rate and duration in the RH11, RH11a, RH11b and RH11c boreholes were 6.8 L/min for 4:10 h, 6.3 L/min for 5:20 h, 6.6 L/min for 4:05 h and 11.3 L/min for 5:40 h, respectively. Water pumped from the RH11, RH11a and RH11b boreholes was discharged into the trench and water pumped from the RH11c borehole was discharged into the Naim Wash (Fig. 1). Water levels were recorded with pressure transducer-data logger systems throughout the tests.

Tracer tests

Rhodamine WT and uranine solutions were injected into the RH11b and RH11c boreholes, respectively. Injection was performed after a steady-state forced-gradient was established by pumping the RH11a borehole at a rate of 8 L/min. A double-excitation, double-detection field fluorometer was installed above the pump in RH11a, and

was initially meant to measure the uranine and rhodamine WT concentrations at 2-min intervals. However, the uranine and rhodamine WT concentrations in the pumping borehole exceeded the range in which the fluorescence of the solutions can be calibrated to uranine/rhodamine WT concentrations, consequently the breakthrough curve (BTC) concentrations were based on diluted samples, analyzed in the laboratory with a spectrofluorometer. In this study, the first 86 h after the tracer's injection was analyzed. During this time, the pumping borehole (RH11a) was sampled at least every 2 h (subsequent to this period, samples were collected twice a day).

Tracers were released into the injection borehole by slowly pulling out a hose, filled with tracer solution, ensuring that all the borehole intervals were injected uniformly. The injection boreholes were constantly mixed and sampled at different depths to confirm that a constant tracer concentration existed throughout the water column and to infer the concentration decay rate at the injection boreholes. Further details concerning this test can be found in Kurtzman et al. (2005).

Flow and transport simulations

Flow simulations were carried out with MAFIC (Miller et al. 1999), the flow simulator of the FracMan package (Dershowitz et al. 1998). Flow and transport were restricted to the fractures (no flow in the matrix). In this finite-element code, the flow in the three-dimensional fracture network is simulated as flow in a network of confined aquifers, i.e. the two-dimensional flow in each fracture is described by Eq. 1

$$T\nabla^2 H = S \frac{\partial H}{\partial t} + q \quad (1)$$

where T (L^2/T) is the fracture transmissivity, S (dimensionless) is the fracture storativity, H (L) is the hydraulic head, ∇^2 ($1/L^2$) is the two-dimensional Laplace operator and q (L/T) is either a source term accounting for flow into the fracture from adjacent (connected) fractures, or a sink term accounting for flow out of the fracture to adjacent (connected) fractures. Transport was simulated with a particle-tracking approach. Particle movement was proportional and parallel to the velocity vector at its location. Two dispersive processes were simulated: the

Table 2 Geometrical and hydraulic properties of the two vertical fracture sets documented in this study

Fracture set (strike, dip in degrees)	Fisher dispersion coefficient	Mean fracture radius (m) (exponential distribution)	Three-dimensional density (m^2/m^3)	Spatial distribution pattern	Fracture transmissivity distribution (mean, standard deviation) of log10 transmissivity (m^2/s)
Dominant (058, 88E)	30	1.9	1.0	Uniform	(−6.8, 1.2)
Secondary (331, 83W)	15	1.1	0.4	Shear zones	(−7.2, 0.7)

first was multiple pathways connecting a source and a sink (“heterogeneous advection”; Becker and Shapiro 2003); the second was advective tracer dilution at the injection borehole. No attempt was made to fit coefficients of random dispersive movement of particles within the fractures or diffusion into the matrix, and these processes were purposely neglected. The technique in which the BTCs were predicted will be described in detail with the prediction results later.

All simulations were transient and natural steady-state conditions were achieved following a few time steps. At that point, the pumping simulation was initiated. The simulation domain was 40 m high, to include the entire length of the boreholes. A group-flux boundary condition was assigned to the four test boreholes (RH11, RH11a, RH11b, and RH11c), which means that heads in all the fracture nodes connected to the borehole are equal to the borehole’s head and that the sum of their fluxes is equal to the borehole’s flux. A no-flow boundary was assigned at the bottom. Six vertical rectangles made up the side external boundaries. Figure 1 displays these boundaries from a top view. Hydraulic heads at the piezometers in the trench were essentially the same and did not respond to pumping in the RH11, RH11a, RH11b and RH11c boreholes. Consequently, a constant-head boundary condition was assigned for the trench boundary (Fig. 1). The hydraulic head in the RH11e borehole was constant and did not respond to pumping, resulting in a boundary condition of $H=aX+b$ assigned to the Hovav boundary (X being the east coordinate, and the coefficients a and b were determined according to the heads in the trench and RH11e). The hydraulic head in the RH11d borehole hardly responded to pumping in the RH11a borehole, resulting in a similarly assigned boundary condition of the form $H=aX+b$ for the Naim boundary, according to the heads at the trench and at the RH11d borehole. The hydraulic head response to pumping observed in the RH11f borehole was larger than that of the RH11d borehole, but smaller than that of the RH11c borehole. Because MAFIC does not allow the head boundary condition to vary in both space and time, the boundary condition at the eastern boundary had the form of $H=H(t)$, where t is time. The head values were between those measured at the RH11d and RH11f boreholes and the transient response to pumping was about half the magnitude of that in RH11c. The boundary connecting the RH11f and RH11e boreholes was assigned as a no-flow boundary, thereby avoiding the head discontinuity that might have formed near the RH11f borehole had a head boundary condition been assigned there.

The DFNs and their predictions

Stochastic DFN realizations

Stochastic DFN realizations with the determined radii distributions and densities (Table 2) could not be realized in the modeled aquitard volume due to computational

limitations (too many fractures). Instead, realizations of fractures with larger radii and lower densities ($\langle r \rangle = 6$ m; three-dimensional density $= 0.3 \text{ m}^{-1}$ and $\langle r \rangle = 4$ m; three-dimensional density $= 0.15 \text{ m}^{-1}$, for the dominant and secondary sets, respectively), were possible. Twelve such realizations were carried out and the transient pumping test (in which the RH11a borehole was pumped) was simulated through them. Whereas a large and rapid drawdown was recorded in the pumping borehole in all realizations, drawdowns in observation boreholes were only observed in the RH11 borehole, in only two out of the twelve realizations. Such results typically suggest the lack of a hydraulic connection between boreholes, as was confirmed by a careful examination of the connected fractures and boreholes in these 12 realizations. Although, in two out of the twelve realizations, the RH11a and RH11 boreholes were connected, calibration of the model to the head response in these boreholes through the change of fractures-sets’ transmissivity distributions, could not be efficiently improved. In fact, a good calibration could only be achieved through a deterministic transmissivity adjustment of the connecting fractures, thereby switching the focus to deterministic DFN models.

The deterministic vertical-fractures (VF) model drawdown predictions

Head-recovery rates observed in inclined boreholes that were significantly larger than those observed in vertical boreholes drilled in this Eocene chalk (Nativ et al. 1999) and the two-dimensional radial responses derived from flow dimension analysis of pumping tests (Kurtzman et al. 2005) led to a simple conceptual model in which vertical fractures are the major flow conduits. The VF model is a deterministic DFN implementation of this conceptual model. The 12 deterministic fractures of this model (Fig. 2) were embedded in the modeled volume such that they intersected the boreholes at relatively transmissive intervals, as indicated by the results of the slug tests in the packed-off intervals carried out in these boreholes. The orientations of these fractures were fixed to approximate the mean orientations of the two fracture sets on site (Table 2), and their transmissivity values ranged between $1\text{E-}6$ and $6\text{E-}5 \text{ m}^2/\text{s}$. The model was calibrated to simulate the results of the pumping test in which the RH11a borehole was pumped and the drawdown was observed in the RH11, RH11a, RH11b and RH11c boreholes. Figure 3a shows the fit of the calibrated model to the transient head data in the four boreholes. Figure 3b–d presents three predictions of the calibrated model in Fig. 3a, through simulated water-level responses in the four boreholes to independent pumping in the RH11b, RH11c and RH11 boreholes, respectively. A goodness-of-fit grade (GFG) was calculated for the calibration and predictions as follows:

$$DR(bh,t) = \text{model drawdown}(bh,t) / \text{data drawdown}(bh,t);$$

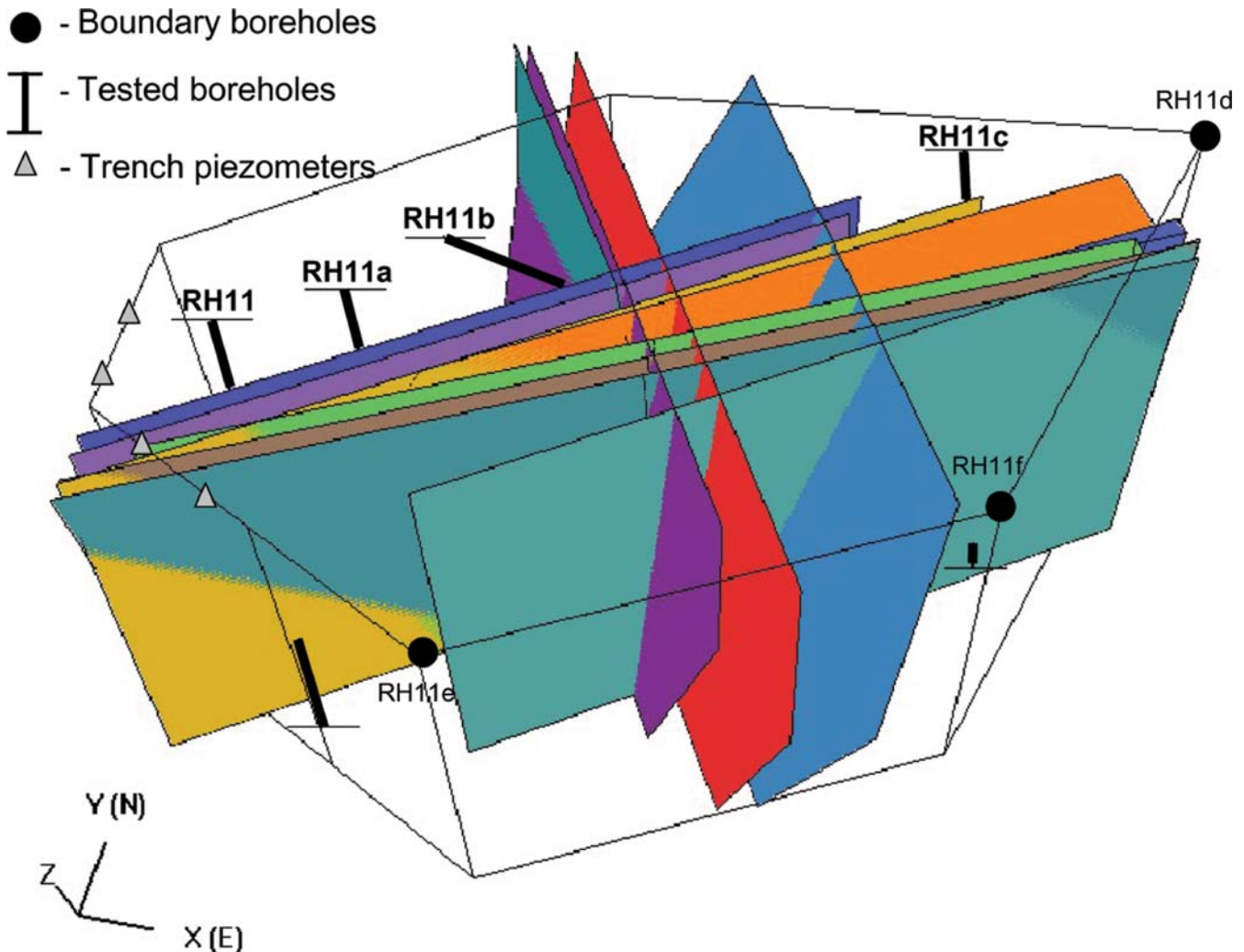


Fig. 2 The vertical-fractures (VF) deterministic DFN model

where $DR(bh, t)$ is the ratio between the simulated and observed drawdown in borehole bh during time t . The GFG for borehole bh at time t is:

if $(DR(bh, t) < 1)$ then $(GFG(bh, t) = DR(bh, t))$;
otherwise $(GFG(bh, t) = 1/DR(bh, t))$.

The overall GFG at a specific borehole is:

$$GFG(bh) = \text{Average}(GFG(bh, t)) \times 100\%$$

The overall GFG (considering all boreholes) is:

$$GFG = \text{Average}(GFG(bh))$$

A high overall GFG was the target of the trial-and-error calibration performed with the deterministic DFN models. The steady-state goodness-of-fit grades (SSGFG) are the GFG when only the latest time data and simulated drawdowns are used.

As expected, calibration of the VF model to the responses to pumping in the RH11a borehole (Fig. 3a)

resulted in a higher GFG than predictions of this model through the simulated and observed head responses in the three other pumping tests (Fig. 3b–d). The observed drawdowns in the RH11 borehole appeared the most difficult to predict by this model, as reflected by an average $GFG(RH11)$ of 47%, calculated from the three values of the predictions (Fig. 3b–d). This difficulty may be related to the proximity of the RH11 borehole to the low, constant-head boundary at the trench (Fig. 1), where the hydraulic gradient becomes steeper. The observed drawdowns in the RH11a borehole were best predicted (with an average $GFG(RH11a)$ of 77%), possibly because it is located relatively far from the model's boundaries. In addition, drawdowns in boreholes closer to the pumping borehole were typically (but not always) better predicted than those in boreholes further away. The calibration and predictions of the steady-state drawdown were 2.5% better, on average, than the average transient drawdowns. The better fit can be explained by the extra attention paid to the transmissivity of the fractures (controlling the steady-state drawdown) vs. that given to their storativity (controlling the transient response) during the calibration process (Fig. 3; Table 3).

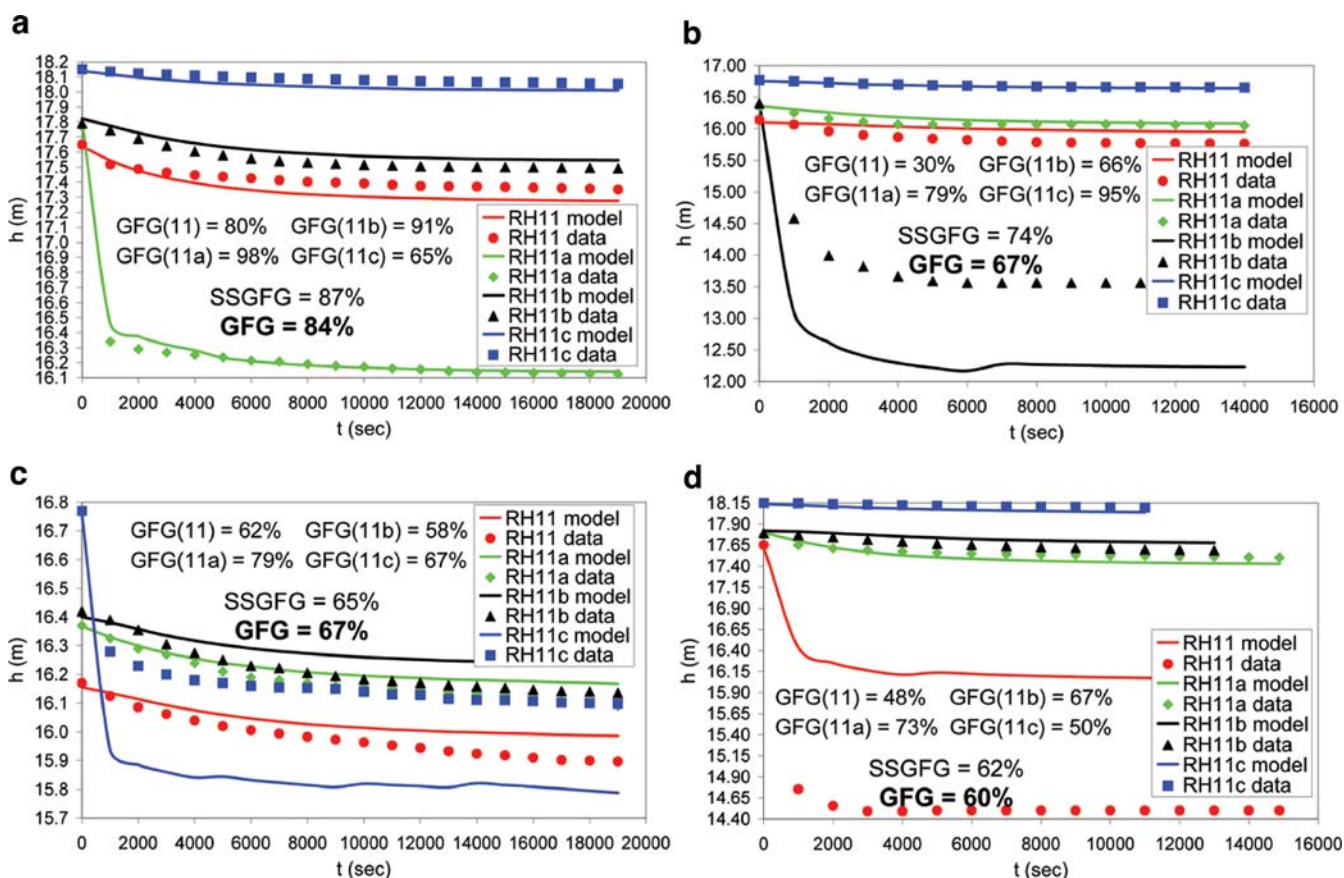


Fig. 3 Actual groundwater heads during the four pumping tests and those simulated by the vertical-fractures (VF) model, including their goodness-of-fit grades (GFG): **a** calibration to the head responses to pumping in the *RH11a* borehole; **b** prediction of the responses to pumping in the *RH11b* borehole; **c** prediction of the responses to pumping in the *RH11c* borehole; **d** prediction of the responses to pumping in the *RH11* borehole

The deterministic fractures- intersections (INT) model drawdown predictions

Macroscopic analysis of these pumping and tracer tests (rather than the discrete-fracture approach of the present interpretation) led to the conclusion that the dominant flow features at this site are channels formed at intersections between bedding planes and vertical fractures (Kurtzman et al. 2005). This conceptual understanding was also supported by some previous observations reported by Dahan et al. (2000) and Nativ et al. (2003). The fractures' intersections conceptual model was implemented in a deterministic DFN model consisting of 12 vertical and 11 horizontal fractures (Fig. 4). The transmissivity values of the vertical fractures in the calibrated INT model ranged between $1\text{E-}7$ and $9\text{E-}6$ m^2/s , higher than those of the horizontal fractures which were between $1\text{E-}9$ and $1\text{E-}7$ m^2/s , following conclusions made by Kurtzman et al. (2005). Transmissivity values of fracture

elements located near horizontal and vertical fractures' intersections (within a radius of 10 cm) were enhanced through multiplication by a calibrated factor of 33. Having three types of water conduits (vertical and horizontal fractures and their intersections) vs. one in the VF model made the calibration process of the INT model five times longer. The high connectivity of this model and the transmissive intersections caused any change in fracture transmissivity during calibration to have more negative effects beside the desired one. The calibration fit was therefore surprisingly lower than that of the VF model (Figs. 3a and 5a). The usefulness of this model will nevertheless be revealed in transport predictions.

Figure 5 presents both the observed and simulated (by the INT model) heads, similar to Fig. 3 for the VF model. Predictions made by the INT model varied between 19% lower to surprisingly 7% higher than the calibration accuracy (the predictions of the responses to pumping in

Table 3 A comparison between the drawdown prediction results from the VF and INT deterministic DFN flow models

Model	Calibration GFG	Average transient prediction GFG	Average steady-state prediction	Range of prediction GFGs	Average worst predicted borehole, GFG	Average best predicted borehole, GFG
VF	84%	65%	67%	60–67%	RH11, 47%	RH11a, 77%
INT	74%	66%	70%	55–81%	RH11, 63%	RH11c, 71%

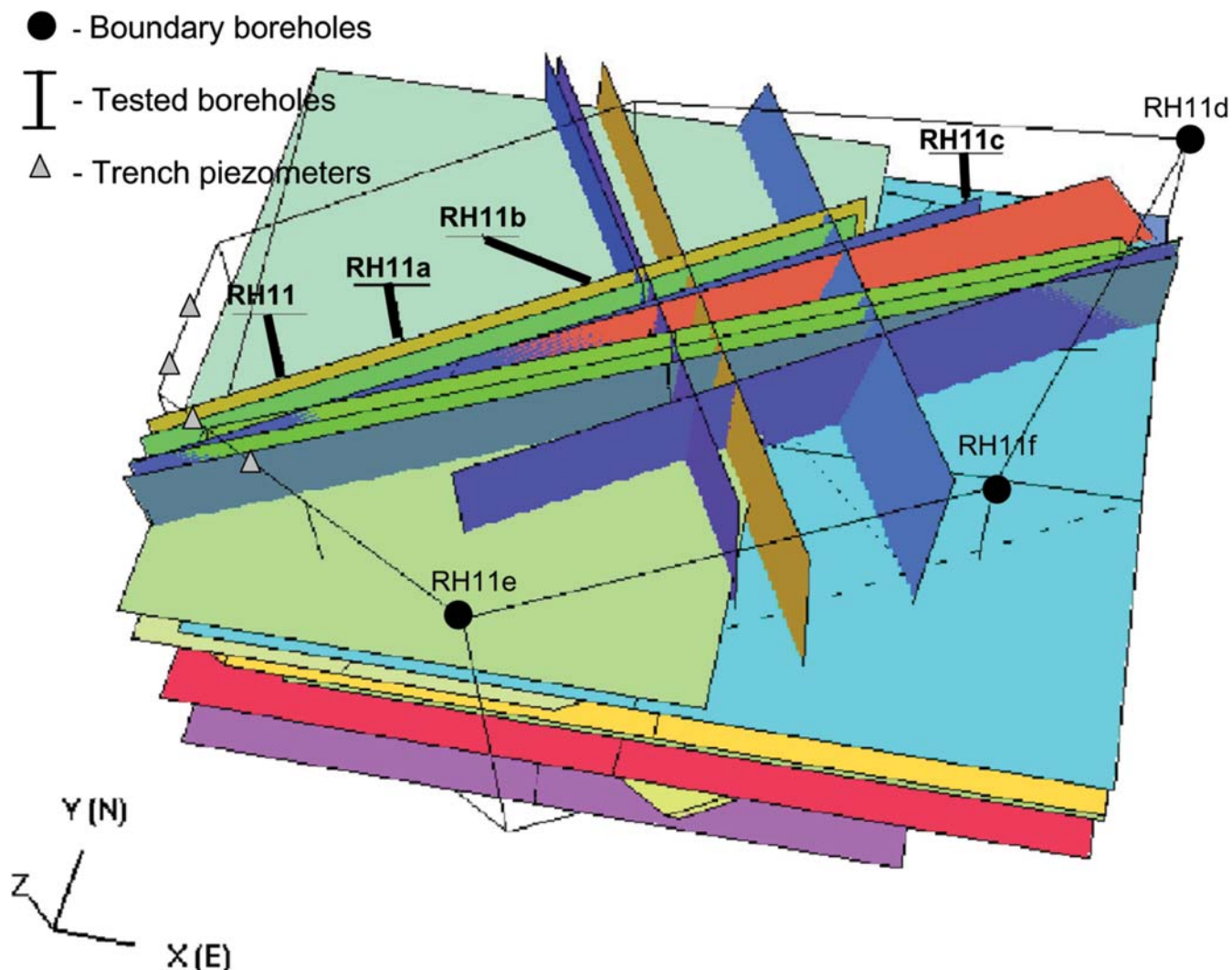


Fig. 4 The fractures' intersections (INT) deterministic DFN model

RH11b and RH11c, respectively, Fig. 5). As opposed to the VF model, drawdowns in any one borehole were not predicted significantly better or worse than those in the other boreholes in the three prediction tests, and no correlation was observed between a borehole's GFG and its distance to the pumping borehole. Table 3, summarizing the results presented in Figs 3 and 5, shows that the GFG of the calibration fit obtained through the INT model was lower than that achieved with the VF model. The average temporal GFG and the SSGFG predictions were slightly higher, in the much more detailed INT model. The average prediction-GFG of both models was about 66%.

Tracer breakthrough predictions of the VF and INT models

Particle-tracking simulations were carried out with the VF model as part of the planning of the tracer test. Through these simulations, the predicted first tracer arrival times in the pumping RH11a borehole were 6 and 7 h with BTC peaks at 6 and 16 h, following their injection (as a short pulse) in the RH11b and RH11c boreholes, respectively.

In reality, observed peak arrival times were 11.3 and 13 h, respectively. Although these results were very useful in planning the time frame for the tracer test, they were misleading with respect of the tracer's first arrival times (40 and 28 min from RH11b and RH11c boreholes, respectively). The boundary conditions used in these simulations differed from the test conditions due to unexpected rain, resulting in significant changes in water levels at the trench.

Figure 6 presents the simulation results which incorporate the hydraulic boundary conditions that existed during the tracer test. It shows the percentage of tracer particles recovered in 2-h intervals during the pumping of RH11a borehole, out of all the particles recovered there. These particles were injected as a short pulse between $0 < t < 2$ h. With the VF model, the main particle mass injected into the RH11c borehole arrived at the RH11a borehole 2–4 h after the injection (Fig. 6a), the last particle arrived at 42 h. When using the INT model (Fig. 6c), early arrival (90 min) was followed by a peak at 6–8 h and a long tail with late particles arriving after more than 100 h. The observed difference between the two

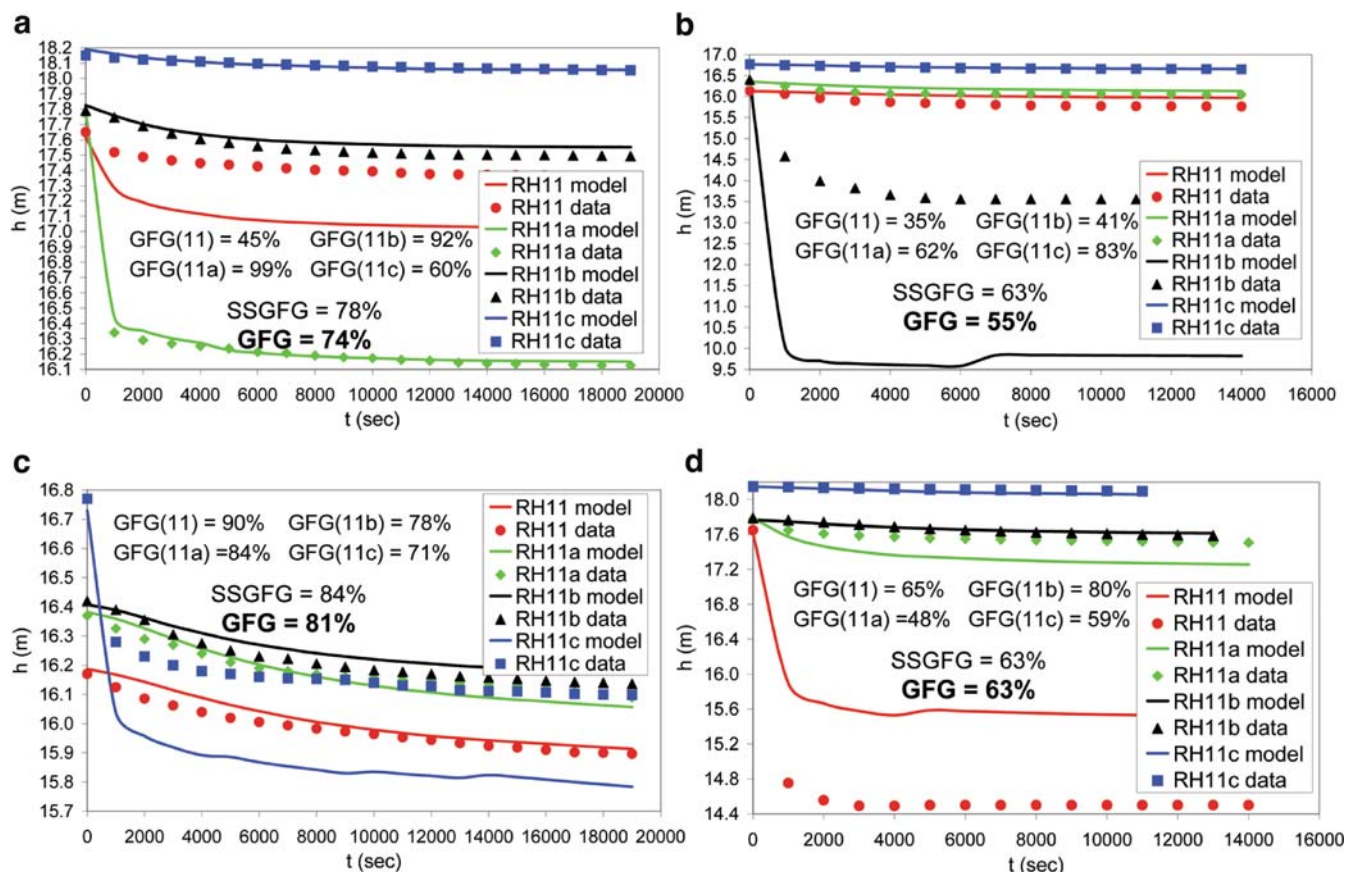


Fig. 5 Actual groundwater heads during the four pumping tests and those simulated by the fractures' intersections (INT) model, including their goodness-of-fit grades (GFG): **a** calibration to the head responses to pumping in the RH11a borehole; **b** prediction of the responses to pumping in the RH11b borehole; **c** prediction of the responses to pumping in the RH11c borehole; **d** prediction of the responses to pumping in the RH11 borehole

models' results was expected as the range of volume and transmissivity of the fluid conduits in the INT model is much larger than that in the VF model.

Particles injected into the RH11b borehole were recovered both at the Naim boundary and in the RH11a borehole. Indeed, during the tracer test, total rhodamine mass recovery from the RH11b borehole was much lower than that of uranine from the RH11c borehole (15 and 80%, respectively), suggesting the loss of a significant amount of rhodamine to the Naim boundary. The combination of a small head gradient near the RH11b borehole and the relatively low hydraulic conductivity of the fractures intersected by this borehole (secondary set; Table 2) resulted in a slow tracer release from this borehole (Kurtzman et al. 2005). Prior to this particle-tracking analysis, the authors were unaware of the small gradient near this borehole and related the slow concentration decrease solely to the low hydraulic conductivity of the secondary set. Because water from this area can flow towards either the pumping borehole or the Naim Wash, it is highly dependent on, and very sensitive to the Naim boundary conditions. Consequently, the breakthrough time distributions of the particle-tracking simulations are highly vulnerable (Fig. 6b and d) and the credibility of both models' predictions of the tracer's breakthrough from the

RH11b borehole is lower than that of the predictions for tracer injected in the RH11c borehole (Fig. 7b).

Instantaneous pulses are often wrongly assumed when modeling slug-injection tracer tests. Instead the advectively decaying concentration at the injection boreholes was assumed in this prediction model. Therefore, after the breakthrough distribution of the pulse injection over time was obtained through the particle-tracking simulation (Fig. 6), these distributions were superposed, as if pulses had been re-injected every 2 h (following the actual injection), to obtain predictions of the overall BTC at the pumping borehole (Fig. 7). The decreased concentration (or number of particles) of each successive injection pulse, $C_{inj}(t)$, were assigned according to a decay exponent calculated from the flow through the injection borehole (Q_{inj}) obtained by the flow model, the injection borehole volume (V_{inj}) and an arbitrary large number of particles injected in the first pulse (C_0 ; Eq. 2).

$$C_{inj}(t) = C_0 e^{-\frac{Q_{inj}t}{V_{inj}}} \quad (2)$$

The predicted BTCs shown in Fig. 7 are the result of the cumulative contributions of the individual decaying successive-pulses' breakthroughs (Fig. 6). All the curves

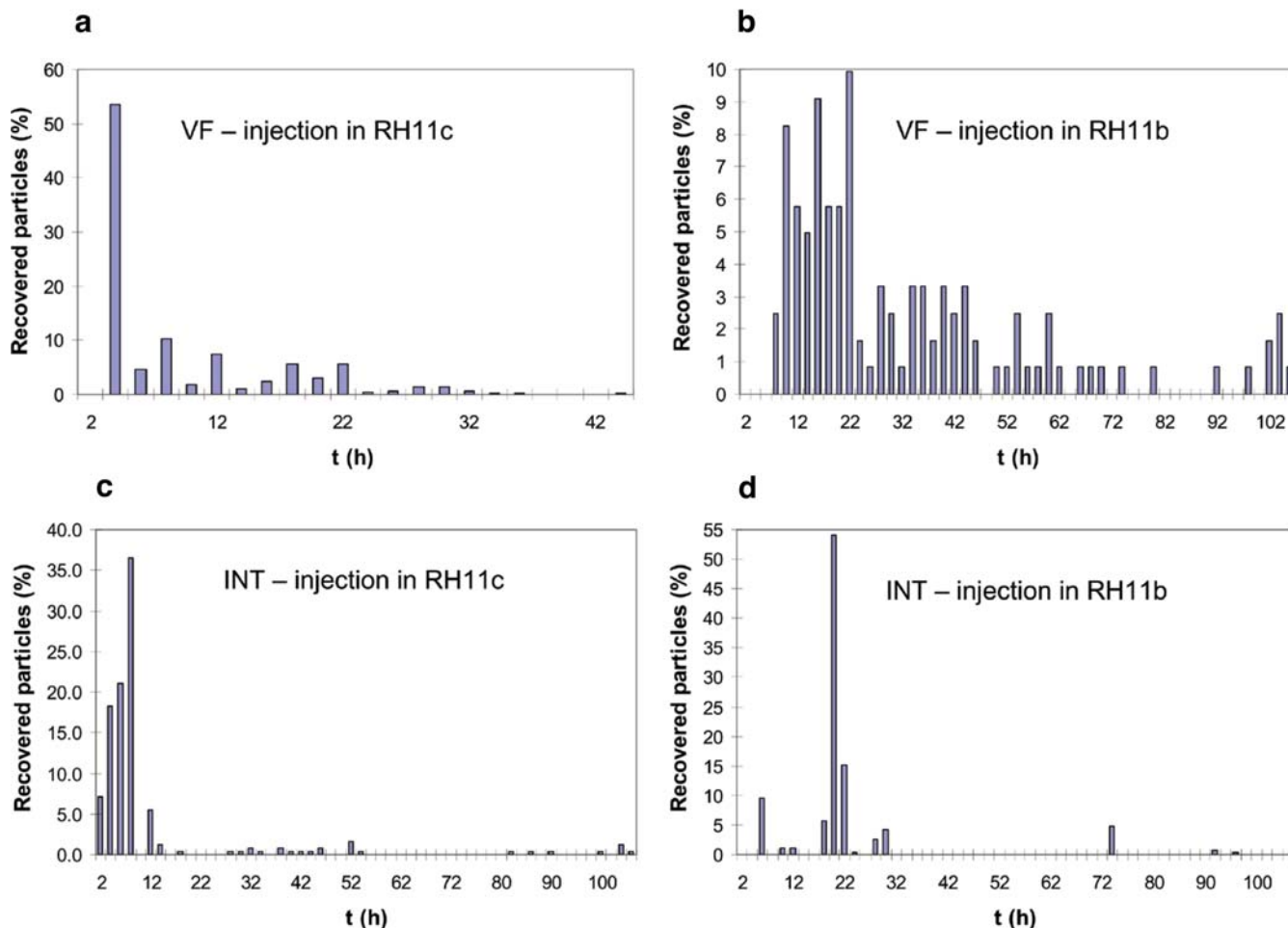


Fig. 6 Recovery-time distribution of simulated particles, injected into the *RH11c* and *RH11b* boreholes and tracked in the pumping *RH11a* borehole, as simulated by the *VF* model (a and b, respectively) and by the *INT* model (c and d, respectively)

in Fig. 7 were normalized, so that the area below them equals 1. The INT model captured the fast first arrival, the peak recovery time and the long tail of the tracer injected in the *RH11c* borehole better than the *VF* model (Fig. 7a). The low flow rate (Q_{inj}) through borehole *RH11c* estimated by the *VF* model, resulted in slow concentration decay at this injection borehole, and yielded a flat predicted BTC (Fig. 7a). The relatively poor fit between the predicted BTC of the INT model and the experimental BTC (Fig. 7a) would probably have been improved if in-fracture dispersion and/or matrix diffusion had been incorporated into the transport model: diffusion into and out of the matrix would have retarded the peak arrival time, reduced the peak concentration and raised the tail concentrations, thereby improving the overall fit. However, the transport prediction described here was designed to test the prediction of the investigated flow models rather than to obtain transport parameters by maximizing the fit to the observed BTC. It should be kept in mind that the nicer fits achieved in posterior parameter fitting of analytic transport models (e.g., Kurtzman et al. 2005, Fig. 5 for the same the same data discussed here) cannot be obtained before the tracer test is conducted, as opposed to the

models presented here, which were calibrated only from a pumping test data, with no assumptions concerning in-fracture dispersion and matrix diffusion.

Slow decay in the *RH11b* borehole controls the shape of the experimental BTC's tail, as reflected in the observed curve in Fig. 7b. Both models underestimated the low flow rate through this injection borehole, thereby providing an even flatter tail than the experimental data (Fig. 7b). Both models failed in producing fast traveling particles (Fig. 6b and d) that would enable one to predict the close-to-peak concentrations after 6 h (Fig. 7b). Nevertheless, the INT model does show 10% of the particles with travel times of 4–6 h (Fig. 6d) better than the *VF* model (Fig. 6b), but these are only enough to form a small early step rather than the continuous early rising limb observed in the data (Fig. 7b).

Discussion

General DFN models of multi-borehole observations

In the DFN modeling approach used in this study, some issues were too complicated to address such as transmis-

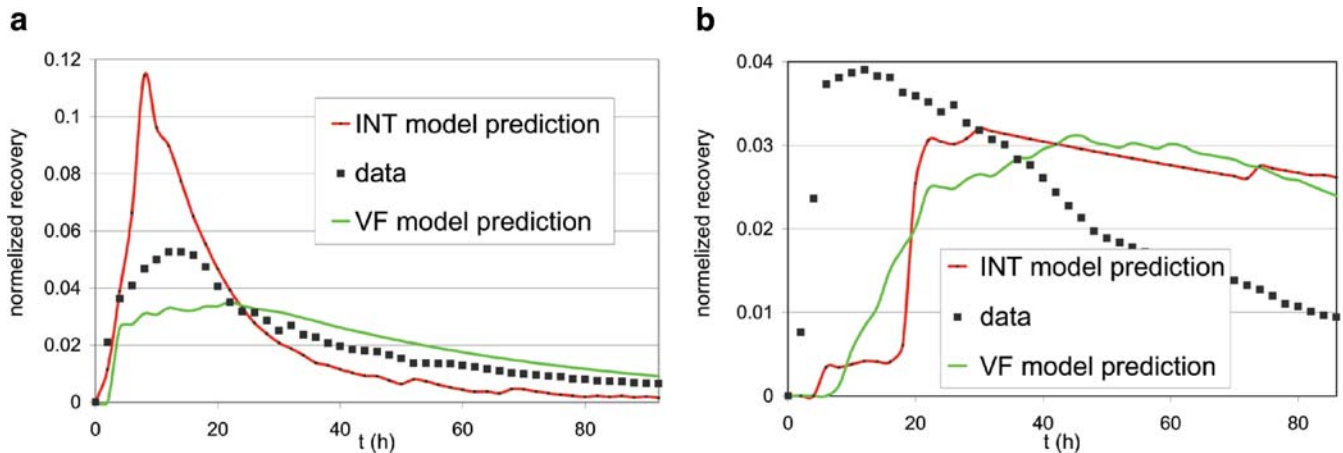


Fig. 7 Observed and predicted normalized BTCs of uranine and rhodamine WT injected in the *RH11c* borehole (a) and the *RH11b* borehole (b)

sivity-heterogeneity within a single fracture, fracture shape, fracture-plane curvature, and matrix processes. Other heterogeneities (i.e., in fracture orientations, radii, and transmissivity) were modeled stochastically. This modeling scheme is by no means sufficient for capturing the complex flow and transport in the fractured chalk field of the study. Nevertheless, it was a useful modeling tool for capturing important features observed at the site, including conductive and nonconductive intervals in the same borehole, anisotropy related to orientation of vertical fractures, and relatively fast pressure responses in observation wells.

Stochastic DFN models of multi-borehole observations

Computation capacity and connectivity

Purely stochastic DFNs (i.e., DFNs that have no deterministic fractures) are likely to have borehole to borehole connectivity problems similar to those described in this case. Outters (2003) reported that, whereas using a three-dimensional density of 10% of the calculated three-dimensional density of the fracture population seriously affects the connectivity and therefore the transport between two side walls of a box model, using 50% of the calculated three-dimensional density hardly affects the transport results. However, in the case reported here, the decrease in connectivity caused by using 30% of the calculated density proved to be fatal, when a simulation of actual borehole-to-borehole responses was attempted. Even the use of the calculated three-dimensional density does not confirm connections between boreholes in a purely stochastic model. Deterministic fractures (i.e., conditioning; Andersson et al. 1984) must be introduced to confirm the connectivity between boreholes. In this test area, the mean radius of the fractures is much smaller than the distance between boreholes; thus, the conditioning of stochastic realizations to the conductive intervals using the fracture radii realized from the revealed radii distributions (Table 2) did not result in multi-borehole connections.

Calibrating stochastic DFNs to borehole data

Cacas et al. (1990a) calibrated a three-dimensional network of pipes to single-borehole injection tests by adjusting the hydraulic conductivity PDF. Niemi et al. (2000) calibrated a stochastic model to single-well injection test data by adjusting the fractures' transmissivity PDFs, and concluded that numerous alternative PDFs can be used to simulate the observed data. In this study, attempts were made to calibrate stochastic DFNs to simultaneous multi-borehole head observations by adjusting the transmissivity PDFs of the two fracture sets. While this calibration procedure was found to be ineffective, model calibration through transmissivity adjustment of the individual fractures forming the hydraulic connection between the two boreholes was possible. These results put into question the usefulness of purely stochastic DFN models in the context of direct multi-borehole flow and transport modeling. Indeed, the use of a stochastic DFN model as the only tool for directly predicting flow and transport phenomena, based on multi-borehole field experiments, could not be found in the literature. In a review study, Neuman (2005) discusses at length why hydrogeologic models based on detailed fracture geometric models (i.e. stochastic DFNs) are less robust than those based on directly measurable flow and transport properties.

Deterministic equivalent DFN models of multi-borehole observations

Efficiency of equivalent deterministic DFN modeling

While stochastic DFN models try to capture a possible configuration of the real fracture network (a realization), the equivalent-deterministic models used in this study did not. Instead they were aimed at capturing the hydraulic behavior of the multi-borehole field using DFN modeling tools, limited outcrop observations, and fracture-flow conceptual models.

In contrast to the discrepancies associated with obtaining fracture radii PDFs from fracture surveys

(Bossart et al. 2001) and the questionable hydrological meaning of these radii, the directly measured fracture orientation and density in outcrops were found to be useful for hydrological characterization. The combination of these two parameters and hydraulic tests in packed-off intervals proved to be sufficient for constructing equivalent DFN models with reasonable predictions of flow and transport. This basic network of dominant structures can form the basis for modeling the observations from multi-borehole tests. An advantage of this procedure is its efficiency in field and computer work: two-dimensional fracture mapping is not needed, and neither are the determinations of PDFs or stochastic realizations. Anisotropy in the hydraulic conductivity is often closely related to the orientations of the fracture sets where two orthogonal fracture sets exist (National Research Council 1996). Therefore, these equivalent deterministic DFNs can closely capture the anisotropy, without spatial-correlation models.

Drawdown predictions

The predictions obtained by both the VF and the INT models suggest that both calibrated models can predict the drawdown in various distances from the pumping borehole (including the pumping borehole itself), orientations (in respect to fracture orientations), time from beginning of pumping, and pumping rates with a relative average goodness of fit of 66%. Steady state predictions were slightly better with 70% for the INT model (Table 3).

Adding complexity to the equivalent-DFN flow model, driven from a macroscopic analysis of the field tests (Kurtzman et al. 2005), has not improved the drawdown predictions (INT vs. VF model). The influence of the inaccuracy of the storage coefficients estimations in the models are likely to be smaller than the inaccuracies in transmissivities (see the small differences between the steady state and temporal predictions, Table 3). Probably, the models' equivalent bulk transmissivities (which are related to the harmonic mean permeability weighted by the flow paths lengths) that controls the observation boreholes' response to pumping has not improved on average between the models. Therefore the drawdowns, which are a measure of this averaged bulk-transmissivity, were predicted with approximately the same average accuracy with both models.

Tracer BTC prediction

Flow velocities rather than groundwater heads control transport phenomena, and predictions of these velocities differed between the two models. Therefore, in contrast to the similarity in average-quality drawdown predictions, quality of tracer BTC predictions differed significantly between the INT and VF models (Fig. 7a). The three categories of conduits in the INT model (low transmissivity horizontal fractures, medium transmissivity vertical fractures and concentrated high transmissivity in horizontal fracture-intersections) rather than the single type

conduit in the VF model, result in better chances to capture more features of the BTC with the INT model.

An example of differences between predictions of drawdown and flow in injection boreholes, which have an impact on BTCs, is presented below. The average drawdown predictions in RH11c borehole results in GFG (RH11c)=71% in both the INT and VF model (Figs. 3b–d and 5b–d, respectively). Yet the prediction of the flow through RH11c borehole in the steady-state phase is much smaller in the VF flow model than in the INT flow model, causing an extremely flat BTC to be predicted with the VF model rather than the time dependent BTC predicted by the INT model (Fig. 7a). The horizontal permeable channels modeled with the INT model also enable fast flow passages that make the early arrival more predictable.

Fracture-matrix diffusive fluxes, not accounted for in this analysis, are expected to retard the peak time, reduce its peak concentration and increase tail concentrations, thereby increasing the fit between predicted and observed BTCs (Fig. 7a). Although matrix exclusion from transport simulations in fractured chalk reduces the overall prediction's accuracy, the important parameters for test planning (e.g., approximated time of tracer's first arrival and peak and tail's significance) were captured here by a purely advective DFN flow model without the need of a priori assumptions regarding diffusion coefficients.

Summery and conclusions

Errors related to the connectivity of boreholes and fractures and the control of a relatively small number of fractures over the inter-borehole responses make it challenging to calibrate purely stochastic DFN models with observations from multiple boreholes. Alternatively, deterministic-equivalent DFN models were found to be an efficient tool for predicting head responses to pumping and important features of tracer BTCs of multi-borehole pumping and tracer tests, performed in fractured chalk. Equivalent-deterministic DFN modeling requires much simpler fracture surveying and analysis than stochastic modeling.

The predictions of drawdown in response to pumping had an average accuracy of ~66%, both in a simpler DFN consisting of vertical fractures only (the VF model) and in a more detailed DFN which included vertical and horizontal fractures and high permeability channels at their intersections (the INT model). The same level of average accuracy in both models can be explained by the fact that borehole-head responses are controlled by the averaged equivalent transmissivities, rather than by the detailed spatial distribution of transmissivities.

The detailed INT model provided better predictions of important BTC features such as first and peak arrival times, than the simpler VF model. These results were expected due to the higher variability of flow rates, velocities and flow paths in the INT model that make it a better fit for predicting the transient breakthrough concentrations. They are also in line with the conceptual

conclusion concerning the importance of flow in horizontal one-dimensional channels derived for this site by Kurtzman et al. (2005). Including in-fracture dispersivity and matrix diffusion in the transport modeling will probably increase the overall accuracy of the predicted BTCs. Yet, this work shows that a purely advective model, without any a-priori assumed transport coefficients, can be very helpful in predicting basic features of the BTC in fractured carbonates.

Acknowledgements The authors are thankful to Golder Associates, Inc., for permitting the use of FracMan and would like to specifically thank Dr. Bill Dershowitz for the many hours spent on training and discussions. The first author would also like to thank the staff and associates of the Reservoir Characterization Research Laboratory (RCRL) at the Bureau of Economic Geology, University of Texas at Austin, for the time given in finalizing this paper.

References

- Andersson J, Dverstorp B (1987) Conditional simulations of fluid flow in three-dimensional networks of discrete fractures. *Water Resour Res* 23:1876–1886
- Andersson J, Shapiro AM, Bear J (1984) A stochastic model of fractured rock conditioned by measured information. *Water Resour Res* 20:79–88
- Andersson P, Dershowitz B, Hermanson J, Meier P, Tullborg EL, Winberg A (2002) Final report of the TRUE block scale project: 1. characterization and model development. Tech. Rep. TR-02–13, Swedish Nuclear Fuel and Waste Management Co. (SKB), Stockholm
- Ando K, Kostner A, Neuman SP (2003) Stochastic continuum modeling of flow and transport in a crystalline rock mass: Fanay-Augères, France, revisited. *Hydrogeol J* 11:521–535
- Becker MW, Shapiro AM (2003) Interpreting tracer breakthrough tailing from different forced-gradient tracer experiment configurations in fractured bedrock. *Water Resour Res* 39(1), 1024. DOI 10.1029/2001 WR001190
- Benke R, Painter S (2003) Modeling conservative tracer transport in fracture networks with an approach based on the Boltzmann transport equation. *Water Resour Res* 39:1324–1334
- Bossart P, Hermanson J, Mazurek M (2001) Äspö Hard Rock Laboratory: analysis of fracture networks based on the integration of structural and hydrogeological observations on different scales. Tech. Rep. TR-01–21, Swedish Nuclear Fuel and Waste Management Co. (SKB), Stockholm
- Bouwer H, Rice RC (1976) A slug test for determining hydraulic conductivity of unconfined aquifers with completely or partially penetrating wells. *Water Resour Res* 12:423–428
- Brown DL, Demir Z, Narasimhan TN (1995) An evaluation of the Bouwer and Rice method of slug test analysis. *Water Resour Res* 31:1239–1246
- Cacas MC, Ledoux G, de Marsily G, Tillie B, Barbreau A, Durand E, Feuga B, Peaudecerf P (1990a) Modeling fracture flow with a stochastic discrete calibration fracture network and validation. 1. The flow model. *Water Resour Res* 26:479–489
- Cacas MC, Ledoux G, de Marsily G, Tillie B, Barbreau A, Durand E, Feuga B, Peaudecerf P (1990b) Modeling fracture flow with a stochastic discrete calibration fracture network and validation. 2. The transport model. *Water Resour Res* 26:491–500
- Chile's JP, de Marsily G (1993) Stochastic models of fracture systems and their use in flow and transport modeling. In: In Bear J, Tsang CF, De Marsily G (eds) *Flow and contaminant transport in fractured rock*. Academic Press, San Diego, CA, pp 169–236
- Cvetkovic V, Painter S, Outters N, Selroos JO (2004) Stochastic simulation of radionuclide migration in discretely fractured rock near the Äspö Hard Rock Laboratory. *Water Resour Res* 40, W02404. DOI 10.1029/2003WR002655
- Dahan O, Nativ R, Adar E, Berkowitz B, Weisbrod N (2000) On fracture structure and preferential flow in unsaturated chalk. *Ground Water* 38:444–451
- Dershowitz W, Lee G, Geier J, Foxford T, LaPointe P, Thoma A (1998) FRACMAN, interactive discrete feature data analysis, geometric modeling, and exploration simulation: user documentation. Golder Associates, Seattle, WA, USA
- Kurtzman D, Nativ R, Adar E (2003) Results from an intermediate-scale study area located at the confluence of the Naim and Hovav washes. In: *Contaminant transport, monitoring and remediation strategies in fractured chalk in the northern Negev*, FracFlux Project, Annual Report for the 2nd year, The Hebrew University of Jerusalem, Israel, pp 69–82
- Kurtzman D, Nativ R, Adar E (2005) The conceptualization of a channel network through macroscopic analysis of pumping and tracer tests in fractured chalk. *J Hydrol* 309:241–257
- Lee SH, Lough MF, Jensen CL (2001) Hierarchical modeling of flow in naturally fractured formations with multiple length scales. *Water Resour Res* 37:443–455
- Long JCS, Billau DM (1987) From field data to fracture network modeling: an example incorporating spatial structure. *Water Resour Res* 23:1201–1216
- Miller I, Lee G, Dershowitz W (1999) MAFIC, matrix/fracture interaction code with heat and solute transport user documentation. Golder Associates, Redmond, WA, USA
- Min KB, Jing L, Stephansson O (2004) Determining the equivalent permeability tensor for fractured rock masses using stochastic REV approach: method and application to the field data from Sellafield, UK. *Hydrogeol J* 12:497–510
- National Research Council (1996) *Rock fractures and fluid flow: contemporary understanding and application*. National Academic Press, Washington, DC
- Nativ R, Adar E, Becker A (1999) Designing a monitoring network for contaminated ground water in fractured chalk. *Ground Water* 37:38–47
- Nativ R, Adar E, Assaf L, Nygaard E (2003) Characterization of the hydraulic properties of fractured chalk. *Ground Water* 41:532–543
- Neuman SP (1987) Stochastic continuum representation of fractured rock permeability as an alternative to the REV and fracture network concepts. In: Farmer IW, Daemen JJK, Desai CS, Glass CE, Neuman SP (eds) *Rock mechanics: proceedings of the 28th US Symposium*, Tucson, Arizona. Balkema, Rotterdam, pp 533–561
- Neuman SP (2005) Trends, prospects and challenges in quantifying flow and transport through fractured rocks. *Hydrogeol J* 13:124–147
- Niemi A, Kontio K, Kuusela-Lahtinen A, Poteri A (2000) Hydraulic characterization and upscaling of fracture networks based on multiple-scale well test data. *Water Resour Res* 36:3481–3497
- Öhman J, Niemi A (2003) Upscaling of fracture hydraulics by means of an oriented correlated stochastic continuum model. *Water Resour Res* 39:1277–1289
- Outters N (2003) A generic study of discrete fracture network transport properties using FracMan/MAFIC. Tech. Rep. R-03–13, Swedish Nuclear Fuel and Waste Management Co. (SKB), Stockholm
- Priest SD (1993) *Discontinuity analysis for rock engineering*. Chapman and Hall, London
- Selroos JA, Walker DD, Ström A, Gylling B, Follin S (2002) Comparison of alternative modeling approaches for groundwater flow in fractured rock. *J Hydrol* 257:174–188
- Therrien R, Sudicky EA (1996) Three-dimensional analysis of variably-saturated flow and solute transport in discretely-fractured porous media. *J Contam Hydrol* 23:1–44
- Tsang YW, Tsang CF, Hale FV, Dverstorp B (1996) Tracer transport in a stochastic continuum model of fractured media. *Water Resour Res* 32:3077–3092
- Warren JE, Root PJ (1963) The behavior of naturally fractured reservoirs. *Soc Petrol Eng J* September:245–255