ground water

Multiple Well-Shutdown Tests and Site-Scale Flow Simulation in Fractured Rocks

by Claire R. Tiedeman[1], Pierre J. Lacombe[2], and Daniel J. Goode[3]

Abstract

A new method was developed for conducting aquifer tests in fractured-rock flow systems that have a pump-and-treat (P&T) operation for containing and removing groundwater contaminants. The method involves temporary shutdown of individual pumps in wells of the P&T system. Conducting aquifer tests in this manner has several advantages, including (1) no additional contaminated water is withdrawn, and (2) hydraulic containment of contaminants remains largely intact because pumping continues at most wells. The well-shutdown test method was applied at the former Naval Air Warfare Center (NAWC), West Trenton, New Jersey, where a P&T operation is designed to contain and remove trichloroethene and its daughter products in the dipping fractured sedimentary rocks underlying the site. The detailed site-scale subsurface geologic stratigraphy, a three-dimensional MODFLOW model, and inverse methods in UCODE_2005 were used to analyze the shutdown tests. In the model, a deterministic method was used for representing the highly heterogeneous hydraulic conductivity distribution and simulations were conducted using an equivalent porous media method. This approach was very successful for simulating the shutdown tests, contrary to a common perception that flow in fractured rocks must be simulated using a stochastic or discrete fracture representation of heterogeneity. Use of inverse methods to simultaneously calibrate the model to the multiple shutdown tests was integral to the effectiveness of the approach.

Introduction

In fractured-rock aquifers, characterizing site-scale groundwater flow is challenging because the complex geometry of fracture networks causes extreme spatial variability of hydraulic properties. Aquifer testing continues to be one of the most useful methods for determining site-scale (here, defined as distances of about 10 to 100 m) heterogeneity and hydraulic properties in fractured rocks, as has been the case for decades. However, aquifer testing is a challenge at contaminated fractured-rock sites, where detailed characterization of the heterogeneity, including identification of highly permeable fractures or fracture zones, often is needed for remediation design and monitoring. At such sites where pump-and-treat (P&T) operations are underway, it is difficult to design classical aquifer tests, which have a single pumping well and multiple observation wells. Conducting such a test would require shutting down the entire P&T operation, allowing water levels to recover (potentially for several days), and then initiating the aquifer test by turning on a pump in one well. Shutdown of the P&T system would compromise the hydraulic containment of contaminated groundwater that is typically achieved by such a system, and potentially yield off-site contaminant migration.

In this work, we report on a novel method for conducting and interpreting multiple aquifer tests at a contaminated fractured-rock site undergoing P&T operations. The aquifer testing approach involves multiple shutdown tests conducted sequentially in several pumping wells of the P&T system. For each test, the pump in one well of the system is turned off, and pumps in other wells of the

[1]Corresponding author: U.S. Geological Survey, 345 Middlefield Road MS 496, Menlo Park, CA 94025; (650) 329-4583; fax: (650) 329-4463; tiedeman@usgs.gov
[2]U.S. Geological Survey, West Trenton, NJ 08628; placombe@usgs.gov
[3]U.S. Geological Survey, Exton, PA 19341; djgoode@usgs.gov
Received April 2009, accepted October 2009.

doi: 10.1111/j.1745-6584.2009.00651.x

system remain on. The test is monitored in observation wells by measuring water level rises produced by turning off the pump. The test ends when the pump is restarted, and water levels then return to their previous state with the P&T system fully operational. Following adequate stabilization of the water levels, another test is performed by turning off the pump in a different well of the P&T system. Because each shutdown test is a type of aquifer test, the two terms are used interchangeably in this paper to describe the tests we conducted.

To our knowledge, the work described here is the first application reported in the literature of this aquifer testing approach at a contaminated fractured-rock site. The method allows aquifer properties to be estimated at sites where characterization activities are constrained by the requirement that contaminated groundwater be hydraulically contained. In addition, the method uses remediation activities to facilitate site characterization. Remediation often must be implemented prior to having carried out substantial subsurface characterization, because of limited financial resources or the immediate need to halt off-site migration of contaminants. Our aquifer testing method enables characterization to continue during remediation, which can potentially lead to refinement and improvement of the existing remediation strategy. Furthermore, it easily enables multiple aquifer tests to be conducted, as P&T systems often include more than one pumping well.

We analyze the multiple shutdown tests simultaneously using a numerical model with heterogeneity represented deterministically and inverse methods. The extreme variability of the fractured-rock hydraulic properties is modeled as extreme variability in hydraulic conductivity, using the equivalent porous media approach. In the following discussion that places this analysis method in the context of previous work, we focus on applications involving multiple aquifer tests. Inclusion of data from multiple tests is key to our work, and analyses that simultaneously consider more than one test are far less common than those involving only a single aquifer test.

One of the most prevalent approaches for analyzing multiple aquifer tests is by application of analytical models. This usually involves analyzing each pumping well—observation well pair individually to estimate hydraulic properties attributed to some region between the pumping and observation wells. For fractured media, analytical solutions are available for a variety of conceptual models of flow to a pumping well (e.g., Neuman 2005). Many of these solutions are based on the concept of a dual continuum (rock matrix and fracture network) model. Application of these solutions to data from multiple aquifer tests has provided insightful results about the variation of hydraulic properties with length scale in a number of fractured-rock settings (e.g., Maréchal et al. 2004; Nastev et al. 2004; Bernard et al. 2006; Kaczmaryk and Delay 2007a,b). However, because application of analytical models to aquifer test data usually yields different estimates of hydraulic properties for different well pairs, it can be difficult to use these estimates to develop a consistent conceptual model of heterogeneity in aquifer properties at the well-field scale (e.g., Raghavan 2004; Neuman 2005; Yeh and Lee 2007).

Because of these limitations, numerical modeling is becoming more prevalent for simultaneously evaluating observation well data from multiple aquifer tests. Geostatistical inverse methods are the most common approaches reported for developing and calibrating these models. The majority of applications have been for unconsolidated aquifers, but a few studies have considered fractured rocks (e.g., LaVenue et al. 1995; LaVenue and de Marsily 2001; Meier et al. 2001; Hendricks Franssen and Gómez-Hernández 2002).

In recent years, a class of methods referred to as hydraulic tomography (HT) has emerged for using multiple aquifer tests to estimate well-field scale heterogeneity in hydraulic and storage properties. Most HT methods use geostatistically based inverse procedures to estimate aquifer properties on a cell-by-cell basis, by sequentially including data from successive aquifer tests (e.g., Yeh and Liu 2000; Zhu and Yeh 2005, 2006; Li et al. 2007). This type of HT has primarily been applied to unconsolidated aquifers (e.g., Li et al. 2007; Straface et al. 2007; Cardiff et al. 2009), but two recent studies have shown that the method also has promise for fractured rocks. Hao et al. (2008) found that when applied to synthetic fractured-rock aquifers, geostatistically based HT could detect the locations of connected, high-K fracture zones, but the estimated K field is smoother than the synthetically generated field. Illman et al. (2009) used this type of HT to delineate large-scale heterogeneities in hydraulic properties of a fractured granite and found that the estimated features are consistent with independent hydrogeologic data.

An alternative to geostatistical inverse methods for simultaneously analyzing multiple aquifer tests is inverse modeling using a numerical model in which the heterogeneity structure is defined deterministically on the basis of hydrogeologic information. To our knowledge, there have been few studies that have used a deterministic approach to simultaneously analyze multiple aquifer tests in fractured rocks at the well-field scale. Allen and Michel (1999) used a trial-and-error method to calibrate a two-dimensional numerical model of a faulted and fractured carbonate aquifer to observation well drawdowns from six aquifer tests. Goode and Senior (2000) used nonlinear regression to calibrate a well-field-scale flow model to two aquifer tests conducted in different wells that were open to the same fracture zones in the sedimentary rock aquifer. Martinez-Landa and Carrera (2006) calibrated a model of five cross-hole aquifer tests in fractured granite in which the dominant fractures controlling flow were identified deterministically and a background network of less important fractures was generated stochastically. Halford and Yobbi (2006) showed that using nonlinear regression and data from six aquifer tests to calibrate a cross-sectional model of a layered aquifer composed of clastic sediments and carbonates produced reasonable estimates of hydraulic parameters, in contrast to the estimates from analytical solutions.

In this paper, we demonstrate that the deterministic approach to heterogeneity definition, combined with an equivalent porous media simulation approach and use of nonlinear regression for calibration, can be used to produce a model of three-dimensional site-scale flow in dipping fractured mudstones that is a reasonable representation of the true system and provides a good fit to data from our novel method of conducting multiple aquifer tests using an existing P&T system. We believe that the deterministic approach to aquifer test analysis in fractured rocks is underutilized. A common view appears to be that representing the hydraulic conductivity heterogeneity stochastically or as a discrete fracture network is most appropriate for simulating flow in fractured rocks (e.g., Neuman 2005 and references therein). This view arises because of the complex geometry and connectivity of fracture networks and the uncertainty associated with characterizing the heterogeneity of these systems. However, we demonstrate that a deterministic approach can be a viable alternative to the stochastic or discrete fracture representation of heterogeneity, particularly for simulating flow. For our site, synergistic use of (1) detailed site-scale subsurface geologic stratigraphy, (2) qualitative and quantitative information from multiple short-duration shutdown tests, and (3) inverse modeling produces conceptual and numerical models that delineate the major hydraulic conductivity variations that are most important for reproducing the aquifer test results.

Site Description

Site History

The multiple aquifer tests evaluated in this paper were conducted at the former Naval Air Warfare Center (NAWC) in West Trenton, New Jersey (Figure 1), which overlies dipping, bedded sedimentary rocks of the Newark Basin. The U.S. Navy tested jet engines at the NAWC from the mid-1950s until the late 1990s. The fractured bedrock was contaminated with accidental releases of trichloroethene (TCE) used during the engine testing operations. The TCE has microbially degraded by reductive dechlorination to form dichloroethene (DCE) and vinyl chloride (VC). A groundwater extraction and treatment system has been in operation since the mid-1990s and has limited off-site migration of contaminants, but groundwater concentrations of TCE, DCE, and VC remain very high. The U.S. Geological Survey (USGS) and cooperators are investigating the fate, transport, and remediation of these chlorinated ethenes at the NAWC (Lacombe 2000; Goode et al. 2007; Tiedeman et al. 2008; Lacombe and Burton, in press).

Geology and Hydrogeology

The NAWC is underlain by lacustrine mudstones of the Lockatong Formation (Fm) and sandstones and shales of the Stockton Fm. These rocks are Triassic in age and were deposited during rifting that created the Newark Basin. Following deposition, the rocks were compacted, tilted, and faulted. A broad, steeply dipping fault

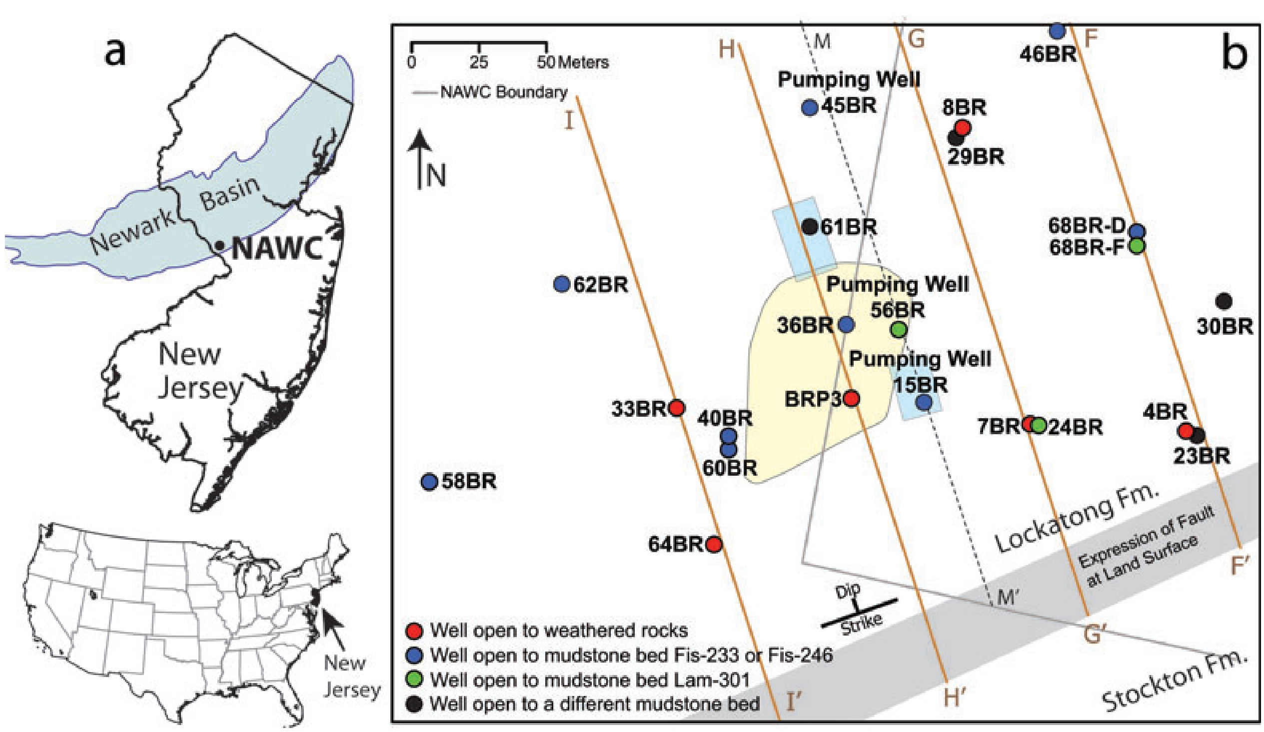

Figure 1. (a) Location of the former Naval Air Warfare Center (NAWC) site in New Jersey and the Newark Basin. (b) Plan view of the NAWC site showing locations of pumping and monitoring wells for aquifer tests conducted in August 2006. Yellow shaded region is the area in which the hydraulic conductivity (*K*) of model layer 14 (representing Fis-233) is lowered by 2 orders of magnitude. Blue shaded regions are areas in which the parameter *Vani_Xbed* applies to allow for the possibility of enhanced vertical hydraulic communication across low-*K* mudstone beds.

zone separates the heavily contaminated Lockatong Fm from the mostly uncontaminated Stockton Fm (Figure 1). Geologic and hydrologic investigations have focused primarily on the Lockatong Fm. At the NAWC, fill and weathered saprolite are present from land surface to depths of about 1 to 8 m. The saprolite is underlain by highly fractured weathered mudstones to depths of about 8 to 30 m. Moderately dipping, competent, variably fractured mudstone beds underlie the weathered rocks.

Lacombe (2000) developed a geologic framework for the NAWC that focuses on the Lockatong Fm and is based largely on interpretation of borehole natural gamma-ray logs. This framework consists of a series of layers formed by connecting borehole intervals that have similar gamma-ray signatures, and accounting for the strike and dip of the mudstone beds. Subsequently, a high-resolution stratigraphic framework consisting of 41 individual dipping mudstone beds was produced by combining the gamma-ray information with geologic data from 20 rock cores (Lacombe and Burton, in press). This framework is depicted as a series of cross sections (Figure 2a-d). The 42 beds have been classified into three types: black fissile mudstones, gray laminated mudstones, and gray or reddish massive mudstones. The geologic framework includes the saprolite and highly fractured rocks that overlie these dipping beds.

The saprolite consists largely of silt and silty clay, and its geometric mean K from slug-test measurements is about 6×10^{-6} m/s (Lewis-Brown and Rice 2002). The underlying weathered rocks are fractured to the extent that individual mudstone beds are difficult to distinguish, and have geometric mean K of about 4×10^{-5} m/s (Lewis-Brown and Rice 2002). This unit acts like an unconsolidated media, in which groundwater can readily flow in horizontal and vertical directions across former bedding planes. This unit is denoted the upper weathered zone, for reasons discussed later. Within the dipping competent mudstones, the black fissile mudstone beds are very thin ($\sim$0.3-m thick), tend to be highly fractured, and for a given depth typically have the highest K among the three mudstone types. The gray laminated mudstones are thicker (0.3 m to 5.5 m) and less fractured than the fissile mudstones. Some fissile and laminated beds contain high-K bedding-plane fractures that are hydraulically connected over distances of up to 100 m. Fissile beds Fis-233 and Fis-246 and laminated bed Lam-301 (Figure 2a-d) are three such units, and are the primary mudstone beds stressed by the aquifer tests discussed in this paper (the numeric part of each of these names refers to the depth, in feet below land surface [b.l.s.], of these beds in a 123-m continuously cored borehole at the NAWC). There also is heterogeneity within the fissile and laminated beds, in that fracturing tends to decrease with depth, and lower-K features can be present at shallow locations. The massive mudstones can be very thick (up to 8 m), typically have very low K, and tend to act as flow barriers.

High-angle, bed-limited, fractures are observed in all mudstone types in outcrop, and in rock core or televiewer logs from some wells. Their presence also has been inferred from single-well hydraulic tests in some packed-off borehole intervals. However, little is known about their distribution in the subsurface.

The depth to the water table varies between about 1 and 5 m b.l.s. Recharge occurs mainly by infiltration of precipitation, and groundwater discharges to pumping wells, culverts and intermittent streams that drain the site, and a small stream adjacent to the site. Groundwater movement is controlled by the highly heterogeneous K distribution caused by variations in rock type and characteristics.

There are approximately 100 monitoring wells completed in the saprolite and rock underlying the former NAWC, completed to depths of 3.5 to 130 m and with open intervals in a wide range of mudstone beds (Figures 1 and 2). Most wells are in the Lockatong Fm. Almost all bedrock wells have open intervals that are 4.5 to 7.5 m in length. One exception is borehole 68BR, which is an open 52-m deep borehole. This well is instrumented with pneumatic packers to hydraulically isolate six intervals of the borehole (Figure 2d), four of which were monitored during the aquifer tests.

Methods

Short-Term Shutdown Tests

In August 2006, six short-term aquifer tests were conducted during a 6-d period by temporarily shutting down pumps in individual recovery wells of the P&T system at the NAWC. This system consists of eight extraction wells with a total pumping rate of about 2.5×10^{-3} m^3/s (40 gpm) and individual well pumping rates ranging from about 6.3×10^{-5} to 6.3×10^{-4} m^3/s (1 to 10 gpm). Pumped water is conveyed to an on-site treatment plant. Prior to starting the aquifer tests, the pumps in all eight recovery wells had been continuously on for several days. The aquifer testing included seven of the eight pumping wells; one well is very shallow and was not tested. Five of the tests involved shutting down an individual pumping well. For the sixth test, two closely spaced pumping wells open to the same mudstone beds were simultaneously shut down. A large rainfall event substantially affected water levels for two of the six aquifer tests and precluded determining observation well responses solely due to pump shutdown. One test was conducted in the Stockton Fm. and did not produce water level changes in observation wells because (1) there are few wells in the Stockton Fm. at the NAWC and (2) the fault zone limits transmission of hydraulic stresses from the Stockton Fm. to the Lockatong Fm. Therefore, of the six tests conducted, three aquifer tests yielded responses amenable to analysis. These tests were conducted using pumping wells 45BR, 56BR, and 15BR (Figures 1 and 2b), and are the focus of this paper.

The first aquifer test discussed here involved shutting down the pump in well 45BR, which is open to mudstone beds Fis-233 and Fis-246 (Figure 2b), for a duration of 8.6 h (Table 1). Water levels in 48 wells, or packed-off intervals of wells, were monitored using pressure

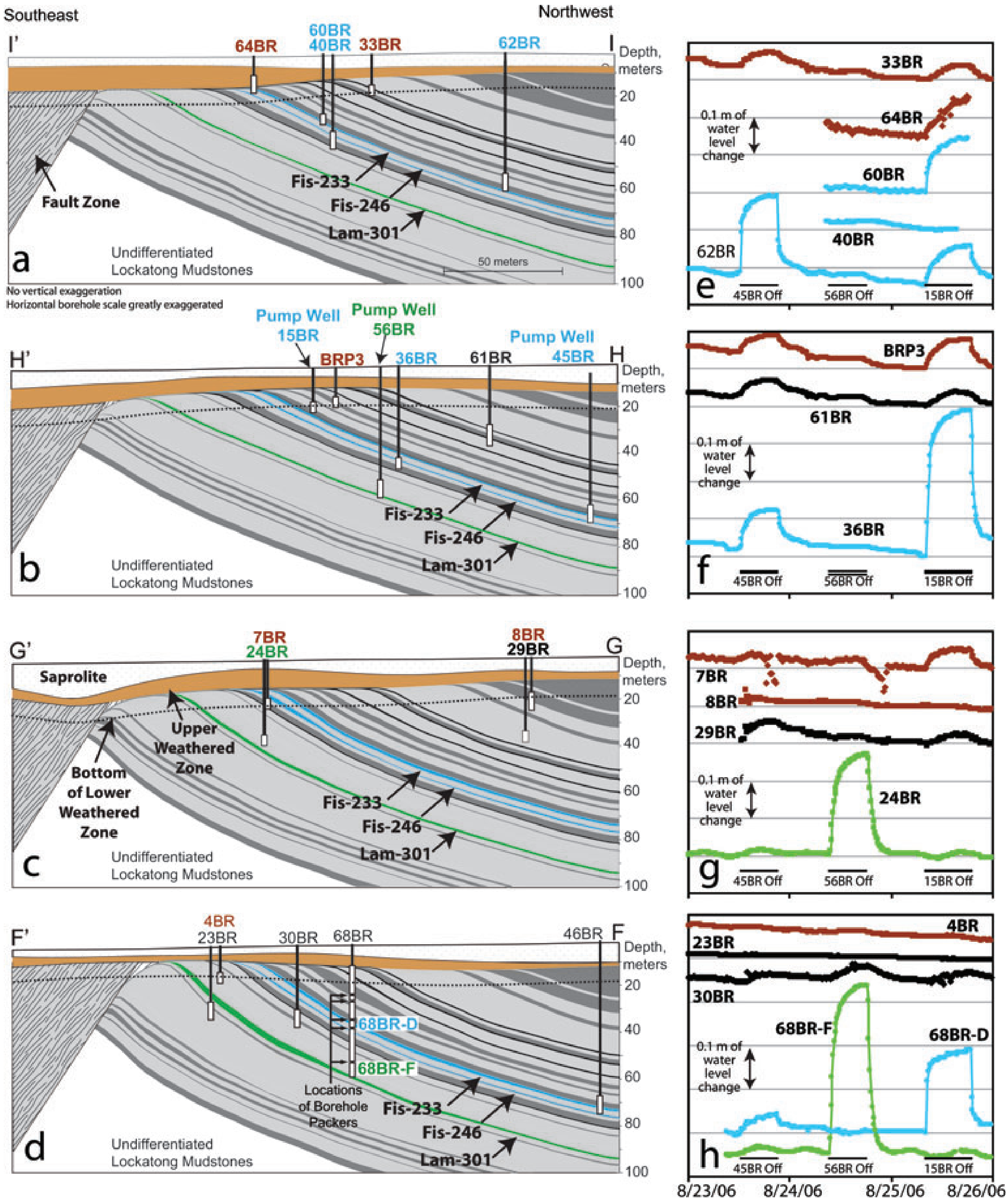

Figure 2. (a-d) Interpretive geologic cross sections showing distribution of mudstone beds in the Lockatong Fm and well open intervals (white rectangles). Undifferentiated mudstones shown in white, massive mudstones shown in medium gray, laminated mudstones shown in dark gray or green, and fissile mudstones shown in black or blue. See Figure 1 for section locations. For 68BR, a total of five inflatable packers separate the well into six open intervals. For nested wells, the label for the shallower well is above that for the deeper well. (e-h) Water level hydrographs, showing water level changes over the 3-d period of aquifer testing in wells 45BR, 56BR, and 15BR. For each panel, the position of each hydrograph with respect to the vertical axis is arbitrary; the panels illustrate the relative water level changes at the different observation locations.

transducers and water level floats, and recorded on data loggers. These wells are located about 50 to 500 m from 45BR. The water level monitoring interval varied from 2 s to 2 min, depending on the anticipated strength of hydraulic connection between the observation well and 45BR and on the time elapsed since the beginning of the test. Following the end of the aquifer test in 45BR (the time at which the pump was turned back on), the pumps

Table 1
Summary of Shutdown Tests and Water Level Observations

Date	Well Turned Off	High-K Mudstone Bed(s) to Which Well is Open	Q^1 (gal/min)	Duration of Aquifer Test (min)	Duration of Recovery (min)	Number of Wells or Intervals Monitored	Locations with Nonzero Water Level Rise Used in Calibration	Observations Used in Calibration
8/23/06	45BR	Fis-233, Fis-246	4.0	520	735	48	7	137
8/24/06	56BR	Lam-301	1.9	540	840	44	3	51
8/25/06	15BR	Fis-233	7.3	630	865	45	12	210

[1]When pump is on.

in all eight wells of the P&T system remained on for 12.25 h. The second aquifer test involved shutting down the pump in well 56BR, which is open to mudstone bed Lam-301 (Figure 2b), for 9 h (Table 1). Water levels were monitored in 44 wells or packed-off intervals located 20 to 470 m from 56BR. The third aquifer test involved shutting down pumping well 15BR, which is open to mudstone bed Fis-233 (Figure 2b), for 10.5 h (Table 1). Water levels were monitored in 45 wells or intervals located 10 to 470 m from 15BR.

In summary, taking advantage of the existing P&T system at the NAWC enabled six separate aquifer tests to be conducted in this flow system over a span of just 6 d. The short duration of each test, and the continued operation of most wells in the P&T system, resulted in little disruption to the hydraulic containment of contaminated groundwater. In addition, because the tests were conducted by shutting down pumps in the existing P&T system, no additional contaminated groundwater was withdrawn during the tests, alleviating the need to store and treat such water. Finally, as discussed in the Results section, the 45BR, 56BR, and 15BR tests yielded water level responses that are highly informative about both the connectivity and properties of the fractured mudstones underlying the NAWC.

Aquifer Test Analysis

Modeling Strategy

We use a deterministic numerical modeling approach, the principle of superposition, and inverse methods to analyze the multiple aquifer tests. As described above, each aquifer test was conducted by shutting down the pump in one extraction well while pumps in the other extraction wells remained on. The superposition model for a test simulates fluid injection at the well in which the pump was shut down and no pumping in the other extraction wells. During model calibration, simulated water level rises caused by the injection are compared to observed water level rises caused by turning off a pump. The superposition approach is reasonable because the saprolite, represented by the top model layer, is thought to act as a semiconfining unit. Water levels in this unit do not respond strongly to pumping in any of the three extraction wells used to conduct the aquifer tests. The tests are simulated with MODFLOW-2000 (Harbaugh et al. 2000; Hill et al. 2000) with all layers represented as confined. The model is oriented with rows parallel to the strike direction and columns parallel to the dip direction. There are 142 rows, 175 columns, and 33 layers. The model uses telescoping grid refinement in the areal dimension, with a minimum cell size of 0.5 $\times$ 0.5 m (at pumping well locations), and cells that gradually increase in dimension up to a maximum size of 20 m $\times$ 20 m. The top of the model domain is the potentiometric surface in the saprolite, and the bottom of the domain is an elevation of approximately 150 m b.l.s. At the top of the model domain, horizontal model layers represent the saprolite and underlying weathered rocks (Figure 3). Details of the vertical discretization are discussed later. Inclined model layers represent the competent dipping mudstone beds. These beds are absent up-dip of where they abut the weathered rocks (Figure 2a-d). In the model, the inclined layers become horizontal and very thin (0.1 m) where the dipping beds are absent (Figure 3). These thin cells are assigned a very large value of vertical K and a very small value of horizontal K. They do not represent a geologic unit; rather, they simply allow groundwater to move directly from cells representing the weathered rocks to the underlying cells representing the dipping beds. This strategy for representing inclined layers beneath a horizontal layer is similar to that used by Lewis-Brown and Rice (2002), Risser and Bird (2003), and Eaton et al. (2007). Details of the discretization of model layers are provided in Supporting Information Figure S1.

The fault separating the Lockatong and Stockton formations is explicitly represented in the model by a nearly vertical zone of low-K cells (Figure 3). The horizontal bottom model layer represents low-K rocks below a depth of 100 m, considered the lowermost depth of active groundwater flow in this aquifer system. Where the inclined model layers abut the bottom layer, they become horizontal and very thin, in a manner similar to their representation where they abut the weathered rocks. The lateral boundaries of the model domain are no flow, and are far enough away from the pumping wells that they

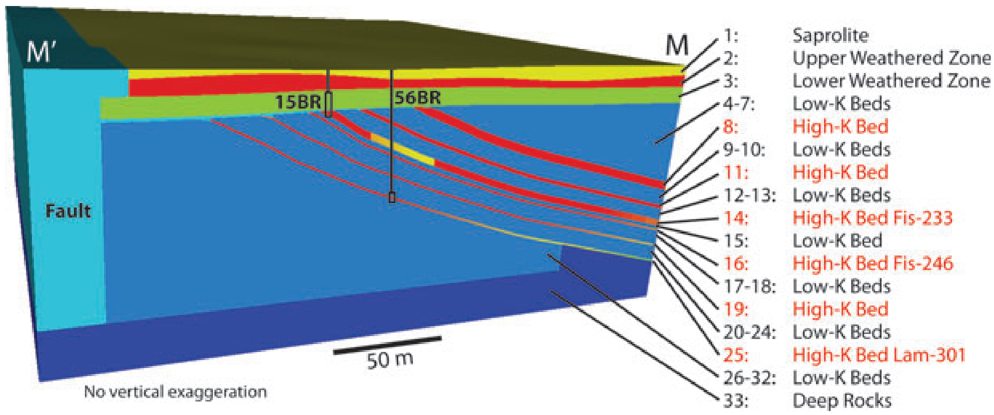

Figure 3. Block diagram of a portion of the model domain in the vicinity of wells 15BR and 56BR. Location of upper edge (M-M') of block face shown is given on Figure 1. Numbered labels identify model layers and the geologic feature they represent. Schematic depiction of wells shows that 15BR pumps from layers 3 and 14, and 56BR pumps from layer 25.

have no effect on simulated water levels at the monitoring wells.

Model Calibration

Calibration of the model involves estimating the distribution and values of hydraulic properties, including horizontal and vertical K and specific storage (Ss). Parameterization of hydraulic properties is guided by the geologic framework model, the short-term aquifer tests, and previous aquifer and slug tests.

The universal inverse modeling software UCODE_2005 (Poeter et al. 2005) is used to estimate hydraulic property values. UCODE_2005 uses a Gauss-Newton method to determine the optimal model parameter values that minimize an objective function S, defined as the sum of squared weighted residuals. The weighted residual for observation i is defined as:

$$e_i = \omega_i^{1/2}(y_i - y_i') \qquad (1)$$

where y_i and y_i' are, respectively, observed and simulated water level rises at location i. ω_i is the weight for observation i, calculated as:

$$\omega_i = 1/\sigma_i^2 \qquad (2)$$

where σ_i is the standard deviation of measurement error for observation y_i. σ_i can be expressed as $cv_i \times y_i$, where cv_i is the coefficient of variation of measurement error for y_i. To determine the optimal parameters for the NAWC model, MODFLOW is simulated once for each aquifer test and the weighted residuals from all three simulations are included in the single objective function value S.

Composite scaled sensitivities (css), which are a measure of the information the calibration observations provide about each model parameter, are used to guide which parameters can be estimated. The css for parameter b_j is calculated as (Hill and Tiedeman 2007, 50):

$$css_j = \left[\sum_{i=1}^{ND} \left[\left(\frac{\partial y_i'}{\partial b_j} \right) b_j \omega_i^{1/2} \right]^2 / ND \right]^{1/2} \qquad (3)$$

where ND is the number of observations, and $\partial y_i'/\partial b_j$ is the sensitivity of simulated value y_i' to parameter b_j.

Linear, individual, 95% confidence intervals are computed to assess parameter uncertainty. These intervals are calculated as (Hill and Tiedeman 2007, 138):

$$b_j \pm t(ND - NP, 0.975)\ s_{b_j} \qquad (4)$$

where $t(ND\text{-}NP, 0.975)$ is the Student's t-statistic for ND-NP degrees of freedom and a significance level of 0.05, NP is the number of model parameters, and s_{b_j} is the standard deviation of parameter b_j, which is computed by UCODE_2005.

Results

Water Level Responses, Heterogeneity, and Connectivity

The water level responses to each aquifer test varied dramatically over the network of monitoring wells (Figures 2e-h and 4). Interpreting these responses in the context of the detailed stratigraphic framework provides substantial insight about which mudstone beds have connected high-K fractures and serve as preferential flow paths, and which are lower in K and primarily serve as flow barriers. This process enables development of a conceptual model of groundwater flow through the system, which is the basis for a numerical model for testing inferences about heterogeneity and connectivity. The discussion below focuses on water level responses that provide information about the mudstone beds Fis-233, Fis-246, and Lam-301, and the weathered rocks.

When the pump in 15BR is shut down, water levels rise quickly in monitor wells 36BR, 58BR, 60BR, 62BR, and 68BR-D (Figure 2e,f,h; 58BR not shown). All these wells are open to mudstone bed Fis-233, with the exception of 36BR (Figure 2a,b,d). This suggests that bed Fis-233 contains connected bedding-plane fractures across a large portion of the area shown on Figure 1. Well 36BR is open to mudstone Fis-246, which is separated from Fis-233 by a 4-m thick massive mudstone bed. The response

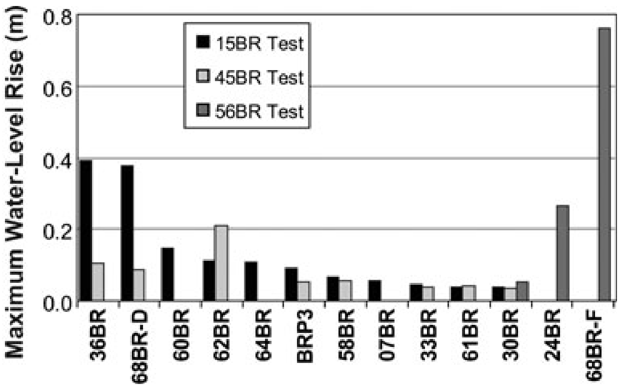

Figure 4. Maximum water level rise at wells or intervals with ≥0.03 m rise at the end of any test. Wells 60BR and 64BR had water level rises >0.03 m for the 45BR test, but the data sets for these wells are missing the maximum rises that occurred at the end of the test. Wells 15BR, 45BR, and 56BR tests are open to, respectively, mudstone beds Fis-233, Fis-233 and Fis-246, and Lam-301.

in 36BR suggests that there are relatively permeable cross-bed fractures in this massive bed that permit a rapid pressure response between Fis-233 and Fis-246. Recent drilling supports this hypothesis: high-angle strata-bound fractures have been identified in wells drilled between 15BR and 36BR. Detection of such fractures is rare in the vertical wells at the site.

Pumping well 45BR is open to mudstones Fis-233 and Fis-246, and the monitor wells that respond to the pump shutdown are generally the same as those that respond during the 15BR aquifer test (Figures 2 and 4). The responses tend to be smaller than for the 15BR test, because of the lower pumping rate in 45BR (Table 1) and the greater distance of 45BR from many of the monitor wells (Figure 1). The responses to the 45BR test further support the conclusion that mudstone Fis-233, and possibly Fis-246, are paths for active groundwater flow at the site scale.

Only two monitor wells, 24BR and 68BR-F, exhibit significant water level rises in response to shutdown of pumping in 56BR (Figures 2g, h and 4), which is open to mudstone Lam-301 (Figure 2b) and has a relatively small pumping rate (Table 1). The stratigraphic framework indicates that interval 68BR-F is open to this unit (Figure 2d), but that the open interval of 24BR lies slightly above Lam-301 (Figure 2c). However, the water level data clearly show that 24BR is hydraulically connected to Lam-301. This mudstone bed might actually intersect the interval, or the connection could be caused by hydraulically active cross-bed fractures linking Lam-301 with the overlying massive mudstone unit. Monitor well 30BR also responds to 56BR, but with a much smaller water level rise than at 24BR or 68BR-F (Figures 2h and 4). 30BR is open to beds that lie above Lam-301 (Figure 2d) and thus it likely has a weak connection to unit Lam-301 through cross-bed fractures in the massive mudstones.

The 15BR aquifer test also provides insight about groundwater flow in the overlying weathered rocks. Water level rises during this test indicate that shallow well 15BR (Figure 2b) is hydraulically connected to dipping mudstone bed Fis-233 as well as to rocks that surround 15BR horizontally. This is illustrated by the responses in 33BR (Figure 2a, e), BRP3 (Figure 2b, f), and 7BR (Figure 2c, g), which are open to dipping mudstones just below the upper weathered zone, the bottom depth of which is inferred from core and drilling logs. The responses in these shallow wells suggest the presence of a lower weathered zone in which water levels respond more like those in the upper weathered zone, where horizontal flow dominates, than like those in the dipping mudstones, where flow along bedding planes dominates. It suggests that although the rocks intersected by the open intervals of these wells are not visibly broken up, chemical weathering may have altered them in a way that has enhanced local fracturing and/or the hydraulic conductivity of the rock matrix.

Finally, several monitoring wells do not respond to shutdown of pumping during any of the three aquifer tests. These include 8BR (Figure 2c, g) and 4BR, 23BR, and 46BR (Figure 2d, h; 46BR water levels not shown). 46BR is open to Fis-233, and is a distance from 15BR at which other wells open to this bed do respond, suggesting that parts of Fis-233 are much less permeable, particularly at depth. Wells 4BR and 8BR are relatively shallow, and are probably beyond the radius of influence for 15BR in the weathered rocks. The open interval for 23BR nearly intersects Lam-301. The response at this well suggests that either the connected permeable fractures within bed Lam-301 do not extend to this location, or that the open interval of 23BR is actually isolated from this bed. This well has very low transmissivity, on the basis of slug testing (Lewis-Brown and Rice 2002).

Model Heterogeneity Distribution and Parameterization

Initial Model

We use the geologic framework together with information about heterogeneity and connectivity inferred from qualitative interpretation of the aquifer tests to delineate the initial distribution of hydraulic properties and to define parameters of the MODFLOW-2000 model (Table 2). We then refine the model on the basis of results from initial inverse model runs.

Three horizontal layers are defined at the top of the model. These represent (1) the saprolite, (2) the upper weathered zone, and (3) the lower weathered zone inferred from aquifer test responses at 33BR, BRP3, and 7BR (Figures 2a-d and 3). One K parameter applies to each of these layers (Table 2). Ss and horizontal anisotropy (*Hani*) are defined so that one parameter applies to the saprolite, and one parameter applies to the two weathered zones. For vertical anisotropy (*Vani*), one parameter applies to the saprolite and upper weathered zone, and a separate parameter applies to the lower weathered zone. The latter parameter represents the vertical

Table 2
Definition of Parameters Representing Hydraulic Properties in the NAWC Model

Geologic Unit	Model Layer(s)	Hydraulic Conductivity Parameter (m/s)	Storage Coefficient Parameter (m^{-1})	Horizontal Anisotropy Parameter[1] (–)	Vertical Anisotropy Parameter (–)
Saprolite	1	*K_Sap* (6×10^{-6})	*Ss_Sap* (1×10^{-4})	*Hani_Sap* (1.0)	*Vani_WthUp* (1.0)
Upper weathered zone (highly fractured)	2	***K_WthUp***	*Ss_Wth* (1×10^{-5})	*Hani_Wth* (1.0)	*Vani_WthUp* (1.0)
Lower weathered zone (less fractured)	3	***K_WthLow***	*Ss_Wth* (1×10^{-5})	*Hani_Wth* (1.0)	*Vani_WthLow* (2.0)
Competent dipping mudstones with high hydraulic conductivity	8, 11, 14, 16, 19	***K_HighK***[2]	***Ss_HighK***	***Hani_HighK***	*Vani_Rock*[3] (1.0) ***Vani_Xbed***[4]
Dipping mudstone bed Lam-301 with high hydraulic conductivity	25	***K_L301***[2]	***Ss_HighK***	***Hani_L301***	*Vani_Rock* (1.0)
Competent dipping mudstones with low hydraulic conductivity	4–7, 9–10, 12–13, 15, 17–18, 20–24, 26–32	***K_LowK***	***Ss_LowK***	*Hani_LowK* (1.0)	*Vani_Rock*[3] (1.0) ***Vani_Xbed***[4]
Competent dipping rocks below a depth of 100 m	4–33	*K_Deep* (1×10^{-9})	***Ss_LowK***	*Hani_LowK* (1.0)	*Vani_Rock* (1.0)
Fault	1–3, 17–32	*K_Fault* (1×10^{-8})	***Ss_LowK***	*Hani_LowK* (1.0)	*Vani_Rock* (1.0)

Note: Parameter names in bold type are estimated by nonlinear regression; those in normal type are specified at the values listed in parentheses. [1]Defined as the ratio of K in the dip direction to K in the strike direction. [2]Defined as the K of cells that are $\leq$15 m below land surface; see Equation 6 for full definition of K in these layers. [3]Applies to most cells. [4]Applies only to selected cells in two areas of the model, one near 61BR and one near 15BR.

communication between the horizontal layers and the underlying dipping beds.

Six high-K dipping mudstone beds are explicitly represented as inclined model layers (Figure 3) that range from 1 to 4 m thick. These include Fis-233, Fis-246, Lam-301 (layers 14, 16, and 25, respectively), and three additional beds (layers 8, 11, and 19) conceptualized as having high K on the basis of water level responses from previous hydraulic testing. In the initial model, a single parameter, *K_HighK*, was used to represent the K of all six of these layers (Table 2). Sets of beds that lie between the high-K beds are conceptualized as lower-K rocks, and are represented by a single K parameter, *K_LowK*. Supporting Information Figure S1 shows details of the discretization of the dipping layers. Below a depth of 100 m, $K = 1 \times 10^{-9}$ m/s in all model layers. Two Ss parameters are defined for the inclined layers, one associated with the thin high-K beds and one associated with the thicker low-K beds (Table 2). The true distribution of Ss is likely more heterogeneous than in the model, but the nature of its variability is unknown. Because of this uncertainty, it is common in fractured-rock flow models to represent the fractures with one Ss parameter and the rock matrix with another.

The high-K beds are conceptualized as anisotropic in the bedding-plane dimension (parameter *Hani_HighK*, Table 2), with K in the strike direction greater than K in the dip direction. Studies in the Newark Basin have found that groundwater flow systems typically have preferential flow along strike (e.g., Longwill and Wood 1965; Vecchioli 1967; Michalski and Britton 1997; Morin et al. 1997; Senior and Goode 1999; Risser and Bird 2003). The low-K beds are assigned a bedding-plane anisotropy of 1.0 (parameter *Hani_LowK*) because they are not significant pathways for flow along bedding planes.

The vertical anisotropy of most cells in the inclined model layers is specified as 1.0 (Table 2). The model construction itself, with alternating layers of high and low K, creates an anisotropic K distribution with preferential flow parallel to bedding. In two zones near 61BR and 15BR (Figure 1), a separate vertical anisotropy parameter, *Vani_Xbed*, is defined to represent enhanced vertical communication across low-K mudstone beds (Table 2). The response at 61BR during the 45BR aquifer test

(Figure 2f) suggests a moderate hydraulic connection between these two wells, possibly through cross-bed fracturing between the different mudstones to which these wells are open (Figure 2b). The locations of such fractures are unknown. To represent this possible feature in the model, a zone of enhanced vertical K across several beds is defined near 61BR. A similar zone is defined around 15BR, where recent drilling and hydraulic testing show the presence of hydraulically active cross-bed fractures.

Refined Model

Initial inverse model runs motivated three important refinements to the K distribution within the dipping beds. First, within each high-K model layer, K was modified to decrease exponentially with depth (z) b.l.s. according to the following relations:

$$K = K\_HighK; \ z < 15 \text{ m} \tag{5}$$

$$K = K\_HighK(10^{-a})$$
$$a = 3\left[\frac{15 - z}{15 - 100}\right]; \ 15 \text{ m} \leq z \leq 100 \text{ m} \tag{6}$$

By Equation 6, K at a depth of 100 m is 3 orders of magnitude smaller than K at a depth of 15 m. This change was prompted by inverse runs in which estimated parameter values that produced the best model fit caused water level rises at deeper observation locations to be significantly underestimated, suggesting that more resistance to flow, caused by lower K, is needed at depth. Slug test data from the NAWC (IT Corporation 1994; Lewis-Brown and Rice 2002) as well as heat-pulse flow logging data from other sites in the Newark Basin (Morin et al. 1997, 2000) show that K decreases with depth.

Second, a low-K feature in the vicinity of 36BR (Figures 1 and 3) was placed in layer 14 representing bed Fis-233. Without this feature, the model could not simultaneously match all observations for the 15BR test in this layer. Parameters that produced a match to the observed water level rise in 68BR-D severely overestimated water level rises in 58BR, 60BR, and 62BR. This problem did not occur for the 45BR test. This suggested that a reduced hydraulic connection was needed between 15BR and wells 58BR, 60BR, and 62BR. Tracer test data (Shapiro et al. 2008) also provide evidence of a low-K region around 36BR, but its extent is not known. In the model, the shape of the added low-K region was defined so that it serves as the necessary flow barrier between 15BR and the monitor wells to the west (Figure 1).

Third, separate K and *Hani* parameters, *K_L301* and *Hani_L301*, were defined for layer 15 representing bed Lam-301 (Table 2). When parameters *K_HighK* and *Hani_HighK* applied to all high-K layers, inverse model runs produced responses for the 56BR test (in Lam-301) at observation wells 24BR and 68BR-F that were too large at early time and too small at late time. This suggested that a lower-K, and possibly a different value of horizontal anisotropy, were needed in the layer representing Lam-301 than in the overlying high-K layers. Furthermore, *css* showed that the water level rise observations could support estimation of these two additional parameters. Parameter *K_L301* applies to this layer at depths ≤15 m, and K decreases within this layer as described in Equation 6.

Table 3
Coefficients of Variation Used to Calculate Weights for Observations with Nonzero Water Level Rises

Observed Water Level Rise	Coefficient of Variation
<0.01	1.00
0.01 to 0.05	0.50
0.05 to 0.10	0.25
>0.10	0.10

Model Calibration and Uncertainty

The model is simultaneously calibrated to 398 water level rise observations from the three shutdown tests, by including in the single objective function S the weighted residuals (Equation 1) from all three MODFLOW aquifer test simulations. For each test, there are 3 to 12 monitor wells or intervals that have nonzero water level rises (Table 1). For each of these locations, 15 to 25 observations are included in the calibration data set; these observations sample the temporal response over the full duration of the test and recovery. Locations with no detectable water level rise also are included, by specifying one observation with zero water level rise occurring at the end of the test.

Observations are weighted in the nonlinear regression calibration procedure. For water level rise observations equal to zero, $\sigma = 0.025$ m is specified for calculating the weights (Equation 2). For nonzero water level rise observations, coefficients of variation are specified, and are assigned on the basis of the magnitude of observed water level rise (Table 3).

Composite scaled sensitivities (Equation 3) were used throughout calibration to inform which parameters could be estimated. The K parameters representing the weathered rocks and many of the parameters representing the dipping mudstone beds have relatively large *css* (Figure 5). Five of the six *Vani*, *Hani*, and *Ss* parameters that represent the saprolite and weathered rocks (*Vani_WthUp*, *Hani_Sap*, *Hani_Wth*, *Ss_Sap*, and *Ss_Wth*) have relatively small *css*, and their values are specified (Figure 5, Table 2). Although *Vani_WthLow* has much larger *css*, indicating that the observations are sensitive to this parameter, it could not be estimated, because it is highly correlated with parameter *K_WthLow*.

On the basis of the *css* and correlations, 10 of the 21 model parameters are estimated (Figure 5, Table 2). The largest absolute correlation for any parameter pair is 0.78, indicating no problems with estimating unique parameter values (Hill and Tiedeman 2007, 51). The K estimates vary over about 6 orders of magnitude

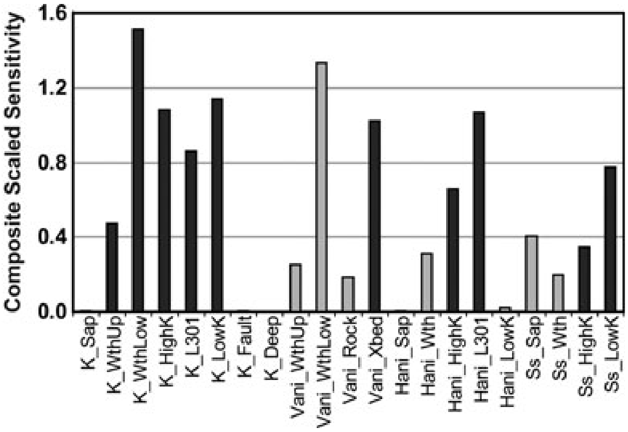

Figure 5. Composite scaled sensitivities (*css*) calculated at the optimal parameter estimates. Values of parameters with dark gray bars are estimated by nonlinear regression and values of parameters with light gray bars are specified.

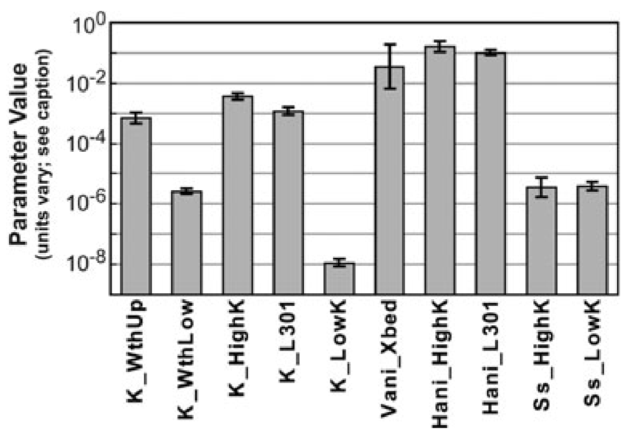

Figure 6. Optimal parameter estimates and individual, linear, 95% confidence intervals. *K* parameters have units of m/s, *Ss* parameters have units of s^{-1}, and *Vani* and *Hani* parameters are dimensionless.

(Figure 6). The relative magnitudes of the *K* parameters are as expected, the estimated values are reasonable on the basis of qualitative and quantitative information about the hydrogeology, and the estimates are generally consistent with previous flow modeling in rocks of the Lockatong Fm (Table 4). The *Ss* estimates for the high- and low-*K* beds are similar. This suggests that at the site scale, release of groundwater stored in the low-*K* beds is not a major process affecting the aquifer test responses. However, as discussed previously, the model representation of *Ss* is simplified, and there might be parts of the rock that locally have larger *Ss*. The uncertainty of most estimated parameters is small, as shown by the confidence intervals that are less than an order of magnitude for all parameters except *Vani_Xbed*.

The calibrated model generally provides a good fit to the observed water level rises (Figure 7), particularly to observations from wells open to beds Fis-233 and Lam-301. The match to observations in the weathered rocks is very good for the 15BR test, but is poor for the 45BR test. The model underestimates the water level rises produced by this test in the weathered rocks, particularly at BRP3. The poor fit may be caused by an unknown geologic feature that would cause a stress in 45BR to be transmitted through the dipping rocks and produce the observed response at BRP3. The overall fit to all observations is very good, as indicated by the standard error of regression (Hill and Tiedeman 2007, 95) equal to 1.6.

Discussion

Our well-shutdown method for conducting aquifer tests using a P&T system is potentially applicable at any contaminated site where extraction and treatment of groundwater is occurring as part of remediation, in unconsolidated formations as well as in fractured-rock environments. At other sites, the duration of each aquifer test (the length of time that an individual pump must be shut down) might need to be increased. At the NAWC, water level changes propagate rapidly through high-*K*, low-*Ss* zones of the mudstone beds that can extend over distances up to 100 m. At sites where the hydraulic diffusivity (K/Ss) is smaller, aquifer test responses will propagate more slowly toward observation wells. In addition, the utility of conducting an aquifer test in this manner could be limited if P&T system pumping rates are small and turning off a pump therefore affects water levels over only a small volume of aquifer. In this case, a test could be performed by increasing the pumping rate in one well, if allowed by treatment plant capacity, hydraulic conditions, and regulatory constraints.

Furthermore, the method is applicable at uncontaminated sites with multiple pumping wells, such as a well field used for water supply. At such sites, with high pumping rates, the method can potentially provide information about hydraulic connections over large distances. Water supply pumping wells usually have a long screen, so water level responses might only be attributable to a relatively thick geologic unit, but this is not necessarily a limitation. When characterizing the hydrogeology of water supply sites, the identification of detailed flow paths is typically not as important as at contaminated sites.

The deterministic equivalent porous media approach to heterogeneity representation at the NAWC is successful from the standpoint of developing a model of groundwater flow that reproduces aquifer test responses. The estimated *K* distribution also benefits development of solute transport models, because the *K* field controls groundwater velocity. The heterogeneity representation for the flow model focuses on major *K* contrasts, such as that between the mudstone beds that are flow paths and those that are flow barriers, and the presence of the low-*K* feature within high-*K* mudstone bed Fis-233. These major contrasts in *K* also are the dominant controls on the advective component of solute transport. Although additional smaller-scale heterogeneities need to be included in transport models to represent processes such as local-scale dispersion, the larger scale heterogeneities delineated during the aquifer test analyses and flow model calibration are a critical

Table 4
Comparison of Parameter Estimates for the NAWC Model with Estimates of Parameters Representing Similar Hydrogeologic Units in Other Groundwater Flow Models in the Lockatong Fm

Parameter	NAWC site-scale model (this paper)	NAWC regional model (Lewis-Brown and Rice 2002)[1]	Colmar, PA model (Risser and Bird 2003)[2]
K_WthUp (m/s)	6.5×10^{-4}	9×10^{-6}	3×10^{-5}
K_WthLow (m/s)	2.4×10^{-6}		
K of high-K Lockatong beds (m/s)	3.5×10^{-6} to 3.5×10^{-3}	1×10^{-6} to 3×10^{-5}	2×10^{-5}
K_LowK (m/s)	9.3×10^{-9}	n.a.	3×10^{-7}
Ss_HighK (m^{-1})	3.6×10^{-6}	4×10^{-5}	1×10^{-6}
Ss_LowK (m^{-1})	4.4×10^{-6}		1×10^{-5}

[1]One layer represents the weathered rocks, and one Ss parameter represents the dipping beds. [2]One layer represents the weathered rocks.

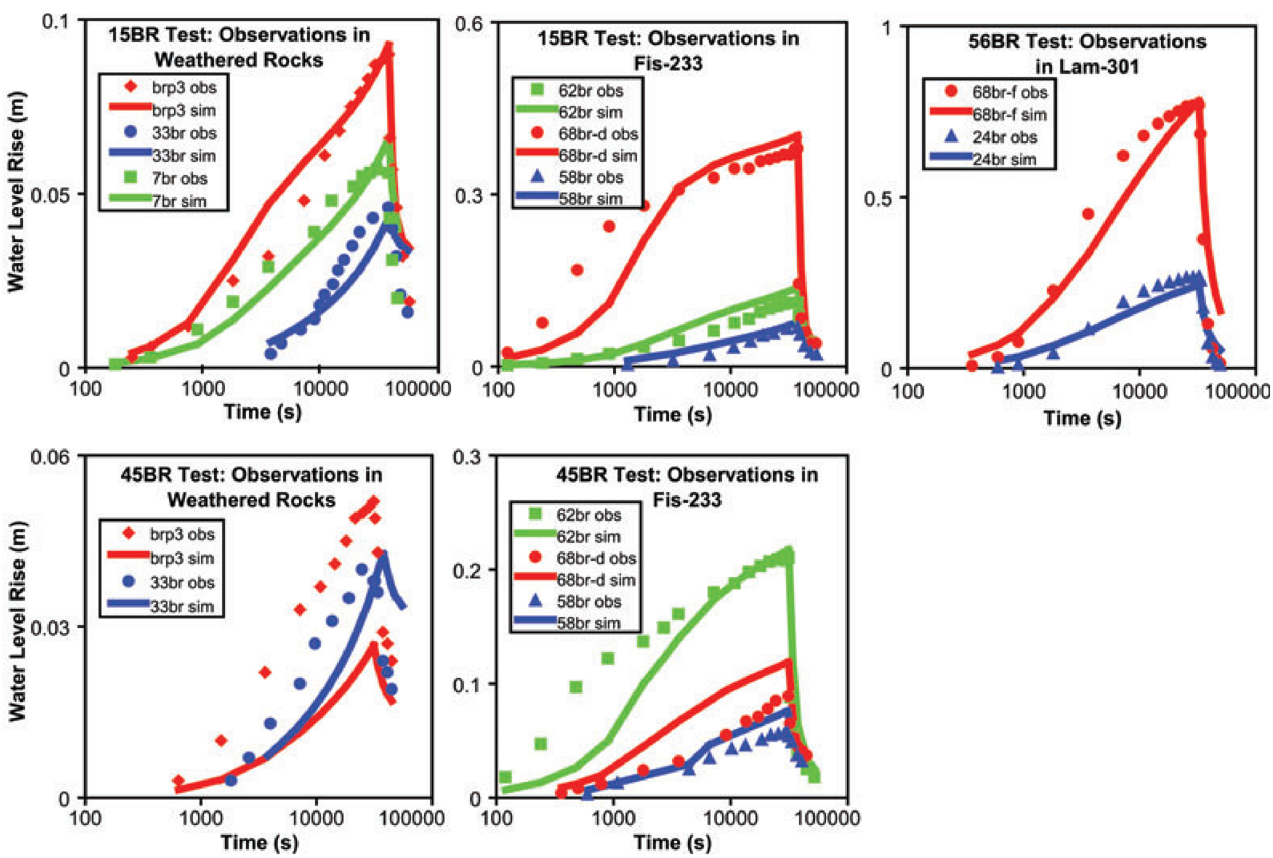

Figure 7. Semilog plots of observed and simulated water level rises at selected observation locations.

first step toward realistically simulating nonreactive solute transport, which is underway at the NAWC to analyze tracer tests conducted in the mudstones, and reactive transport, which is being conducted to evaluate remediation strategies.

The insights gained from the synergistic use of detailed site-scale geologic and aquifer test data to develop conceptual and numerical models of groundwater flow have broad relevance beyond the NAWC site. These insights include (1) the depth below land surface to which groundwater flows in a predominantly horizontal direction toward a shallow pumping well (as opposed to along dipping beds) extends beyond that inferred from inspection of visibly weathered rocks in core, possibly due to chemical weathering; (2) below the region of predominantly horizontal flow there are only a few mudstone beds, out of the 41 mapped beds that compose the geologic stratigraphy underlying the NAWC, that have high enough *K* to serve as major flow paths along bedding planes; and (3) the *K* of these beds decreases substantially with depth but is still higher than that of the surrounding lower-*K* beds. These findings are likely to be transferable to hydrogeologic model development at the more than 1000 sites in the fractured sedimentary rocks of the Newark Basin that are on the New Jersey Department of Environmental Protection's Site Remediation Program Comprehensive Site List, which contains sites with known groundwater contamination (New Jersey Department of Environmental Protection 2009). Beyond the Newark Basin, there are a number of other early Mesozoic basins in the eastern United States to which our findings are relevant, as these basins are characterized by similar sedimentary rocks and depositional styles (e.g., Trapp and Horn 1997). Even more broadly, the finding that very thin beds can be highly important hydraulically is potentially applicable to any groundwater flow system in layered sedimentary rocks.

Conclusions

Our results demonstrate several important findings about conducting and analyzing aquifer tests in the fractured sedimentary rocks underlying the NAWC.

Our new well-shutdown method for using an existing P&T system to conduct aquifer tests has several major advantages at contaminated sites. When an aquifer test is conducted by shutting down the pump in one well, the remainder of wells in the extraction network can remain on, and can continue to capture contaminated groundwater. If the test were conducted by turning off all pumping wells, letting the system recover, and then pumping from just one well, the groundwater containment ability of the extraction system is likely to be compromised. Also, when using wells in a P&T system, all extracted water is discharged to this system through existing plumbing. No additional contaminated water is withdrawn, alleviating the need for storage and treatment. Finally, conducting such tests at existing P&T sites is logistically simple and cost effective, and uses remediation activities to help characterize the site hydrogeology.

Short-term aquifer tests of only a few hours in duration provided definitive information about subsurface hydraulic properties and connections. The short-duration tests conducted at the NAWC were able to provide information about hydraulic connectivity over distances of >100 m because the higher-K bedding-plane fractures are characterized by small storativity (S, where $S = Ss \times b$, and b is thickness of the active flow zone), and thus large diffusivity, allowing the hydraulic responses to travel quickly through these fractures and to a lesser extent, through sparse high-K cross-bed fractures. Short-term shutdown tests are highly advantageous from a logistical standpoint, in that several tests can be conducted in the span of just a few days. These results are consistent with the work of Michalski and Britton (1997), who also demonstrated the utility of short-duration tests in fractured sedimentary rocks.

Use of detailed three-dimensional site-scale geologic stratigraphy to help interpret aquifer test data was critical for developing realistic conceptual and numerical models of site-scale groundwater flow. In sedimentary rocks, the magnitude and direction of flow are strongly controlled by the structure of the bedding. Preferential flow paths often are associated with bedding-plane partings or with particular beds that might be more fractured because of, for example, their depositional setting or tectonic history. Thin, highly fractured beds that could easily be overlooked in a cursory geologic investigation might play critical roles in groundwater flow and contaminant transport. At the NAWC, using aquifer test data in conjunction with the site-scale geologic framework made it possible to associate hydraulic connections inferred from aquifer testing with particular mudstone beds delineated in the framework, some of which are extremely thin. This allowed for identification of likely flow paths between open intervals of wells and for development of conceptual and numerical models of hydraulic conductivity heterogeneity and connectivity. Using multiple shutdown tests for the analysis provided information about hydraulic connections and barriers in a greater volume of the flow system than if fewer tests had been used.

A deterministic equivalent porous media approach to simulating the multiple well-shutdown tests, combined with inverse methods for model calibration, was successful for simulating site-scale flow in the highly heterogeneous fractured rocks. This approach produced a calibrated model that is a realistic representation of the true simulated flow system, as indicated by estimated parameter values that are reasonable compared to independent hydrogeologic information, small uncertainty in most parameter estimates, and a good fit to most water level rise observations. These results indicate that for simulating site-scale flow in fractured rocks, this approach is a viable alternative to delineating the heterogeneity using stochastic methods or discrete fracture networks.

Acknowledgments

We gratefully acknowledge funding support from the USGS Toxics Substance Hydrology Program; the U.S. Navy, Naval Facilities Engineering Command; the Department of Defense Strategic Environmental Research and Development Program (SERDP), project ER-1555; and the U.S. EPA Technology Innovation Program. We also thank ECOR Solutions, Inc. and the New Jersey Department of Environmental Protection for permission to temporarily turn off pumps in recovery wells at the NAWC and conduct the aquifer tests described in this paper. Reviews by Ward Sanford and Richard Yager of the USGS, AE Scott James, and two anonymous reviewers significantly improved the content of this paper.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Cross section showing model grid discretization for rows 7 to 129 and layers 1 to 33 along column 78.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

Allen, D.M., and F.A. Michel. 1999. Characterizing a faulted aquifer by field testing and numerical simulation. *Ground Water* 37, no. 5: 718–728.

Bernard, S., F. Delay, and G. Porel. 2006. A new method of data inversion for the identification of fractal characteristics and homogenization scale from hydraulic pumping tests in fractured aquifers. *Journal of Hydrology* 328, no. 3-4: 647–658.

Cardiff, M., W. Barrash, P.K. Kitanidis, B. Malama, A. Revil, S. Straface, and E. Rizzo. 2009. A potential-based inversion of unconfined steady-state hydraulic tomography. *Ground Water* 47, no. 2: 259–270.

Eaton, T.T., M.A. Anderson, and K.R. Bradbury. 2007. Fracture control of ground water flow and water chemistry in a rock aquitard. *Ground Water* 45, no. 5: 601–615.

Goode, D.J. and L.A. Senior. 2000. Simulation of aquifer tests and ground-water flowpaths at the local scale in fractured shales and sandstones of the Brunswick Group and Lockatong Formation, Lansdale, Montgomery County, Pennsylvania. U.S. Geological Survey Open-File Report 00-97, 46 p. http://pa.water.usgs.gov/reports/ofr00-97/.

Goode, D.J., C.R. Tiedeman, P.J. Lacombe, T.E. Imbrigiotta, A.M. Shapiro, and F.H. Chapelle. 2007. Contamination in Fractured-Rock Aquifers - Research at the former Naval Air Warfare Center, West Trenton, New Jersey. USGS Fact Sheet 2007-3074. http://pubs.usgs.gov/fs/2007/3074/.

Halford, K.J. and D. Yobbi. 2006. Estimating hydraulic properties using a moving-model approach and multiple aquifer tests. *Ground Water* 44, no. 2: 284–291.

Hao, Y., T.-C.J. Yeh, J. Xiang, W.A. Illman, K. Ando, K.-C. Hsu, and C.-H. Lee. 2008. Hydraulic tomography for detecting fracture zone connectivity. *Ground Water* 46, no. 2: 183–192.

Harbaugh, A.W., E.R. Banta, M.C. Hill, and M.G. McDonald. 2000. MODFLOW-2000, the U.S. Geological Survey modular ground-water model—User guide to modularization concepts and the Ground-Water Flow Process. U.S. Geological Survey Open-File Report 00-92, 121.

Hendricks Franssen, H.J.W.M. and J.J. Gómez-Hernández. 2002. 3D inverse modelling of groundwater flow at a fractured site using a stochastic continuum model with multiple statistical populations. *Stochastic Environmental Research and Risk Assessment* 16, no. 2: 155–174, doi 10.1007/s00477-002-0091-7.

Hill, M.C. and C.R. Tiedeman. 2007. *Effective Groundwater Model Calibration: With Analysis of Data, Sensitivities, Predictions, and Uncertainty*. New York: Wiley and Sons, 455.

Hill, M.C., E.R. Banta, A.W. Harbaugh, and E.R,. Anderman. 2000. MODFLOW-2000, the U.S. Geological Survey modular ground-water model—User guide to the Observation, Sensitivity, and Parameter-Estimation Processes and three post-processing programs. U.S. Geological Survey Open-File Report 00-184, 210.

Illman, W.A., X. Liu, S. Takeuchi, T.-C.J. Yeh, K. Ando, and H. Saegusa. 2009. Hydraulic tomography in fractured granite: Mizunami Underground Research site, Japan. *Water Resources Research* 45, W01406, doi:10.1029/2007WR006715.

IT (International Technology) Corporation. 1994. Remedial Investigation report, Installation Restoration Program, Naval Air Warfare Center, Trenton, New Jersey. Contract N624-86-c-1041, Project number 529538: Edison, N.J., 6 v.

Kaczmaryk, A. and F. Delay. 2007a. Interference pumping tests in a fractured limestone (Poitiers - France): Inversion of data by means of dual-medium approaches. *Journal of Hydrology* 337, no. 1-2: 133–146.

Kaczmaryk, A. and F. Delay. 2007b. Improving dual-porosity-medium approaches to account for karstic flow in a fractured limestone: Application to the automatic inversion of hydraulic interference tests (Hydrogeological Experimental Site, HES - Poitiers - France). *Journal of Hydrology* 347, no. 3-4: 391–403.

Lacombe, P.J. 2000. Hydrogeologic framework, water levels, and trichloroethylene contamination, Naval Air Warfare Center, West Trenton, New Jersey, 1993-97. U.S. Geological Survey Water-Resources Investigations Report 98-4167, 139.

Lacombe, P.J., and W.C. Burton. in press. Hydrogeologic framework of USGS research site for contaminated fractured rock. *Groundwater Monitoring and Remediation*. In press.

LaVenue, A.M., B.S. Ramarao, G. De Marsily, and M.G. Marietta. 1995. Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields 2. Application. *Water Resources Research* 31, no. 3: 495–516.

LaVenue, A.M., and G. de Marsily. 2001. Three-dimensional interference test interpretation in a fractured aquifer using the pilot point inverse method. *Water Resources Research* 37, no. 11: 2659–2675.

Lewis-Brown, J.C., and D.E. Rice. 2002. Simulated groundwater-flow, Naval Air Warfare Center, West Trenton, New Jersey. U.S. Geological Survey Water-Resources Investigations Report 02-4019, 44.

Li W., A. Englert, O.A. Cirpka, J. Vanderborght, and H. Vereecken. 2007. Two-dimensional characterization of hydraulic heterogeneity by multiple pumping tests. *Water Resources Research* 43, W04433, doi:10.1029/2006WR005333.

Longwill, S.M., and C.R. Wood. 1965. Ground-water resources of the Brunswick Formation in Montgomery and Berks Counties, Pennsylvania. Pennsylvania Geological Survey, 4th ser., Water Resource Report 22, 59 p., web edition. http://www.dcnr.state.pa.us/topogeo/pub/water/w022.aspx.

Maréchal, J.C., B. Dewandel, and K. Subrahmanyam. 2004. Use of hydraulic tests at different scales to characterize fracture network properties in the weathered-fractured layer of a hard rock aquifer. *Water Resources Research* 40, W11508, doi:10.1029/2004WR003137.

Martinez-Landa, L., and J. Carrerra. 2006. A methodology to interpret cross-hole tests in a granite block. *Journal of Hydrology* 325, no. 1-4: 222–240.

Meier, P.M., A. Medina, and J. Carrera. 2001. Geostatistical inversion of cross-hole pumping tests for identifying preferential flow channels within a shear zone. *Ground Water* 39, no. 1: 10–17.

Michalski, A. and R. Britton. 1997. The role of bedding fractures in the hydrogeology of sedimentary bedrock—evidence from the Newark Basin, New Jersey. *Ground Water* 35, no. 2: 318–327.

Morin, R.H., G.B. Carleton, and S. Poirier. 1997. Fractured-aquifer hydrogeology from geophysical logs; The Passaic Formation, New Jersey. *Ground Water* 35, no. 2: 328–338.

Morin, R.H., L.A. Senior, and E.R. Decker. 2000. Fractured-aquifer hydrogeology from geophysical logs: Brunswick Group and Lockatong Formation, Pennsylvania. *Ground Water* 38, no. 2: 182–192.

New Jersey Department of Environmental Protection (NJDEP). 2009. Digital Data Downloads in ArcView Shapefile Format, Groundwater Contamination Areas (CEA). http://www.nj.gov/dep/gis/stateshp.html#CEA (accessed July 30, 2009).

Nastev, M., M.M. Savard, P. Lapcevic, R. Lefebvre, and R. Martel. 2004. Hydraulic properties and scale effects investigation in regional rock aquifers, south-western Quebec, Canada. *Hydrogeology Journal* 12, no. 3: 257–269, doi:10.1007/s10040-004-0340-6.

Neuman, S.P. 2005. Trends, prospects, and challenges in quantifying flow and transport through fractured rocks. *Hydrogeology Journal* 13, no. 1: 124–147, doi: 10.1007/s10040-004-0397-2.

Poeter, E.P., M.C. Hill, E.R. Banta, S. Mehl, and S. Christensen. 2005. UCODE_2005 and Six Other Computer Codes for Universal Sensitivity Analysis, Calibration, and Uncertainty Evaluation. U.S. Geological Survey Techniques and Methods 6-A11, 283.

Raghavan, R. 2004. A review of applications to constrain pumping test responses to improve on geological description and uncertainty. *Reviews of Geophysics* 42, RG4001, doi:10.1029/2003RG000142.

Risser, D.W. and P.H. Bird. 2003. Aquifer tests and simulation of ground-water flow in triassic sedimentary rocks near Colmar, Bucks and Montgomery Counties, Pennsylvania. U.S. Geological Survey Water-Resources Investigations Report 03-4159, 73.

Senior, L.A., and D.J. Goode. 1999. Ground-water system, estimation of aquifer hydraulic properties, and effects of pumping on ground-water flow in Triassic sedimentary rocks in and near Lansdale, Pennsylvania. U.S. Geological Survey Water-Resources Investigations Report 99-4228, 1999, 112. http://pubs.water.usgs.gov/wri994228/.

Shapiro, A.M., C.R. Tiedeman, D.J. Goode, and P.A. Hsieh. 2008. Using in situ tracer experiments to identify pathways and residence times of ground water in fractured rock: Implications for designing ground-water remediation strategies. Proceedings of SERDP/ESTCP Partners in Environmental Technology Technical Symposium & Workshop, Dec 2-4, 2008, Washington, D.C., p. F-112. http://www.serdp-estcp.org/Symposium2008/Posters/upload/TuesdayPosters-2008.pdf.

Straface S., T.-C.J. Yeh, J. Zhu, S. Troisi, and C.H. Lee. 2007. Sequential aquifer tests at a well field, Montalto Uffugo Scalo, Italy. *Water Resources Research* 43, W07432, doi:10.1029/2006WR005287.

Tiedeman, C.R., D.J. Goode, A.M. Shapiro, P.J. Lacombe, F.H. Chapelle, P.M. Bradley, T.E. Imbrigiotta, J.H. Williams, G.P. Curtis, and P.A. Hsieh. 2008. Multidisciplinary investigation of the fate, transport, and remediation of chlorinated solvents in fractured rocks at the former Naval Air Warfare Center (NAWC): Scientific and management challenges, and strategies for a successful research program. *Eos Trans. AGU* 89, no. 53: Fall Meet. Suppl., Abstract H53E-1110.

Trapp, H. Jr., and M.A. Horn. 1997. Ground Water Atlas of the United States: Segment 11, Delaware, Maryland, New Jersey, North Carolina, Pennsylvania, Virginia, West Virginia. U.S. Geological Survey Hydrologic Atlas Report 730-L, 24.

Vecchioli, J. 1967. Directional hydraulic behavior of a fractured-shale aquifer in New Jersey. In International symposium on hydrology of fractured rocks, Dubrovnik, Yugoslavia, 1965. Proceedings: International Association of Scientific Hydrology, Publication 73, 1, 318–326.

Yeh, T.-C., and S. Liu. 2000. Hydraulic tomography: Development of a new aquifer test method. *Water Resources Research* 36, no. 8: 2095–2105.

Yeh, T.-C., and C.-H. Lee. 2007. Time to change the way we collect and analyze data for aquifer characterization. *Ground Water* 45, no. 2: 116–118.

Zhu, J. and T.-C.J. Yeh. 2005. Characterization of aquifer heterogeneity using transient hydraulic tomography. *Water Resources Research* 41, W07028, doi:10.1029/2004WR003790.

Zhu J. and T.-C.J. Yeh 2006. Analysis of hydraulic tomography using temporal moments of drawdown recovery data. *Water Resources Research* 42, W02403, doi:10.1029/2005WR004309.