

Article

Transient Hydraulic Tomography Analysis of Fourteen Pumping Tests at a Highly Heterogeneous Multiple Aquifer–Aquitard System

Zhanfeng Zhao ^{1,2,*}, Walter A. Illman ², Yuanyuan Zha ^{3,4}, Tian-Chyi Jim Yeh ³, Chin Man Bill Mok ⁵, Steven J. Berg ^{2,6} and Dongmei Han ^{1,*}

- Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
- ² Department of Earth and Environmental Sciences, University of Waterloo, Waterloo, ON N2L 3G1, Canada
- ³ Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ 85721, USA
- ⁴ State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China
- ⁵ GSI Environmental Inc., Oakland, CA 94612, USA
- ⁶ Aquanty Inc., Waterloo, ON N2L 5C6, Canada
- * Correspondence: zhaozhanfeng@igsnrr.ac.cn (Z.Z.); handm@igsnrr.ac.cn (D.H.)

Received: 17 July 2019; Accepted: 5 September 2019; Published: 8 September 2019



Abstract: Hydraulic tomography based on geostatistics has proven to be robust in characterizing subsurface heterogeneity in hydraulic conductivity (K) and specific storage (S_s) through the joint inversion of drawdown records from multiple pumping tests. However, the spatially variable estimates can be smooth or even erroneous for areas where pumping/observation data densities are not high. Previous hydraulic tomography surveys conducted at the North Campus Research Site (NCRS) on the University of Waterloo campus in Waterloo, Canada, revealed that the estimated hydraulic parameters were smooth and the known aquitard was erroneously identified as a high K zone. This was likely the consequence of the site being highly heterogeneous, while only utilizing four pumping tests and not having measurable drawdowns in the low K aquitard for inverse modeling. Here, we investigate whether improved K and S_s estimates could be obtained through the inclusion of additional pumping test data by stressing both aquifer and aquitard zones for a sufficiently long period. Specifically, six additional pumping/injection tests were conducted at the site, and a transient hydraulic tomography analysis with 14 tests was completed. Results reveal that there is a significant improvement to the K and S_s tomograms in terms of the visual correspondence with various geologic units, including its connectivity. More importantly, with the availability of additional data, we found that the inverse model now can better capture the high and low K features for nine boreholes when compared with K values obtained from permeameter tests. The estimated K and S_s tomograms are then used for the forward simulation of one additional pumping test not used for model calibration, revealing reasonable predictions. While encouraging results are obtained by including a large number of pumping tests to the transient hydraulic tomography analysis, stratigraphic boundaries are still smoothed, which is a direct consequence of utilizing a geostatistics-based inversion approach that assumes stationarity in statistical properties. To capture such sharp boundaries, incorporation of additional data types, such as geological and geophysical information, may be necessary when data densities are not sufficiently high.

Keywords: hydraulic tomography; subsurface heterogeneity; hydraulic conductivity; specific storage; geostatistical inverse modeling; connectivity; aquitards



1. Introduction

The subsurface distributions of hydraulic conductivity (K) and specific storage (S_s) control fluid, solute, and energy transport in geologic media. Conventional hydrogeological characterization techniques, such as slug and single-hole tests, as well as laboratory permeameter tests and grain size analyses (e.g., [1,2]), are invasive and cost-intensive to obtain high-resolution information. Direct push techniques can provide high-resolution information on K during each push (e.g., [3]), but without interpolation, there is no information of K between push locations. In addition, while direct push techniques are useful, they do not provide information on S_s and also have difficulty penetrating unconsolidated media when there are cobble size particles and hard clays. Traditional pumping tests commonly treat the aquifer as homogeneous and yield non-unique spatially averaged properties when a Theis or Cooper–Jacob method of analysis is used.

Recently, hydraulic tomography (HT) has been proposed as an alternative site characterization method to obtain information on heterogeneity in groundwater flow parameters between boreholes. In particular, HT is a method for determining the spatial distribution of K and S_s in heterogeneous aquifers through the use of cross-well interference tests. HT involves stressing an aquifer by pumping water from or injecting water into a well, and the aquifer's response is monitored simultaneously at other wells. Sequentially switching the pumping or injection location, without installing additional wells, results in a large number of aquifer responses caused by stresses induced at different locations. These pairs of pumping and drawdown data sets at different locations result in multiple snapshots of the aquifer heterogeneity. Subsequent processing of the data sets from all tests through a consistent mathematical model provides the detailed spatial distribution of hydraulic properties of the aquifer, patterns of connectivity of highly conductive zones, locations of low conductive aquitards, and the uncertainties associated with the spatial distribution.

The operational methodology of HT is similar to a computerized axial tomography (CAT) scan technology employed in medical sciences; sequential tests in HT induce propagation of drawdown at multiple locations throughout the aquifer, and the data sets, therefore, contain more information than a pumping test conducted within a single well. At sites where ongoing remedial operations include pump-and-treat systems, HT surveys can be conducted by simply modifying the pumping rates or by taking advantage of the pumping shut-off and commencement operations. The interpretation methodology for HT data has been developed specifically to account for aquifer heterogeneities so that the final interpretation is consistent with all the observation well data collected during all the pumping tests, a result that cannot be reproduced using conventional approaches that assume aquifer homogeneity.

The power of HT has been recognized after the work by Yeh and Liu [4] and Liu et al. [5]. Initially, HT research consisted largely of numerical studies [4,6–11], followed by laboratory sandbox studies [5,12–14], leading to much optimism in imaging heterogeneity between boreholes. Both the numerical and laboratory studies on HT were critical in validating the method under controlled conditions and laid the foundation for its field application.

To date, relatively few field studies of HT have been published [7,15-21]. In particular, Bohling et al. [16] assessed the steady shape hydraulic tomography using 23 pumping tests conducted in an alluvial aquifer at the Geohydrologic Experimental and Monitoring Site (GEMS) of the Kansas Geological Survey. Steady shape data were selected from a pumping test during the period when the hydraulic gradients had reached steady state, while hydraulic heads had not [16]. Their tomographic analysis produced 1-D vertical profiles of *K* between pumping wells that matched reasonably well with *K* values obtained from a forced gradient tracer test and from direct push tests.

Straface et al. [15] performed the first field transient hydraulic tomography (THT) study using the geostatistics-based inverse code developed by Zhu and Yeh [22]. The heterogeneous distributions of estimated transmissivity (T) and storage coefficient (S) in two-dimensions reasonably represented the geological setting of the site. Later, Cardiff et al. [7] used steady-state hydraulic tomography (SSHT) of nine pumping tests to estimate the distributions of depth-averaged K in an unconfined aquifer

at the Boise Hydrogeophysical Research Site (BHRS). They obtained a *K* tomogram that contained site-specific geological features. At the Mizunami site in Japan operated by the Japan Atomic Energy Agency (JAEA), Illman et al. [21] demonstrated a large-scale application of THT in fractured rock using two cross-hole pumping tests, which was the first 3-D field application of transient hydraulic tomography. The estimated *K* and S_s distributions in 3-D showed several continuous high *K* and low S_s zones, indicating possible geological fault zones. However, the validity of *K* and S_s tomograms was not rigorously evaluated through predictions of independent drawdowns not used in model calibration effort due to the lack of additional pumping test data, until a recent study by Zha et al. [23].

More recently, Berg and Illman [18] conducted nine pumping tests at the North Campus Research Site (NCRS) site on the University of Waterloo campus in Waterloo, Canada. They performed geostatistical inverse analyses of individual pumping tests as well as a three-dimensional THT study using four pumping tests, while the other five tests were reserved for validation purposes. They found that HT yielded better results when compared to geostatistical inverse modeling of individual pumping tests. Then, Berg and Illman [24] compared THT with traditional characterization and modeling approaches and found THT to be superior in terms of predicting independent pumping tests not used in the model calibration effort. However, in the work of Berg and Illman [18,24], the inversion of four pumping tests yielded: (1) relatively smooth hydraulic parameter estimates; (2) less satisfying model validation results when compared to the laboratory validation of HT [25,26]; and (3) misidentification of a low *K* aquitard at the bottom of the investigated area, when in fact, the unit was found to be low *K* through core sampling. Since no significant drawdowns were obtained by Berg and Illman [18] from aquitard layers, it is possible that HT results might be improved by including a larger number of pumping tests and extending their durations to sufficiently stress the aquifer-aquitard system.

Most recently, Zhao and Illman [27] conducted an SSHT study of eight pumping tests at the NCRS and investigated the use of geological information as prior information. This study revealed improved model calibration and validation results. While the results have been found to be improved, the model still identified a high *K* zone where core sampling indicated a low *K* zone. In addition, the study did not utilize a larger number of pumping tests to investigate the aforementioned issues encountered by Berg and Illman [18]. The inclusion of other data types (e.g., geological, geophysical, small-scale data, flowmeter, etc.) into HT may be desirable when data densities are low and when only a limited number of pumping test data are available. However, there is a question about whether the inversion results could be further improved when additional pumping test data are included in the inversion.

Six additional pumping/injection test data have become available since the initial field studies of HT by Berg and Illman [18]. Therefore, the objective of this paper is to examine the impact of including additional tests into the THT survey and to gauge its improvement in terms of the issues encountered by Berg and Illman [18].

2. Experimental Design

2.1. Site Description

Based on previous site investigations by Alexander et al. [2] and Berg and Illman [18], the North Campus Research Site (NCRS) of the University of Waterloo (UW) in Waterloo, Ontario, Canada, consists mainly of highly heterogeneous glaciofluvial deposits and the aquifer zone is located approximately 8 to 13 m below ground surface (mbgs). Borehole logs and core sample analyses indicated that the aquifer zone consists of two high *K* units separated by a discontinuous low *K* layer, with the upper unit consisting of sand and sandy silt and the lower unit consisting of sandy gravel [2]. Below the main aquifer zone, the dense Catfish Creek Till, which acts as a hydraulic barrier, is found at approximately 18 mbgs and treated as the bottom boundary for this study. Near the surface, the aquifer zone is generally confined by a laterally extensive aquitard layer, but is known to contain stratigraphic windows [28]. Previous pumping tests conducted at the site revealed that the aquifer is confined [2]. The groundwater

flows generally towards the south-east as indicated by the water level measurements near the well nests at the site.

A total of seven wells and two well nests were instrumented within a 15 × 15 m square area of the site for pumping tests and data collection purposes, and the well layout is shown in Figure 1. Four continuous multichannel tubing wells (CMT1, CMT2, CMT3, and CMT4), each containing seven screen intervals spaced 2-m apart, were installed in between the corners of the square area to monitor groundwater level changes in 3-D. The uppermost screens of the CMT wells were located between 4.5 to 5.5 mbgs, and the lowermost ones between 16.5 to 17.5 mbgs. Three multi-screen pumping wells (PW1, PW3, and PW5) were constructed with diameters of 10 cm and 1 m length screens. For each multi-screen well, the screen interval was 1 m. PW1, PW3, and PW5 were extended to around 18 mbgs, 12 mbgs, and 12 mbgs, respectively. PW2 and PW4 well nests consist of three separate wells each and have a diameter of 5 cm. Each well in the nests has a 1-m long screen with mid-point elevations for PW2 at 4, 7, and 8 mbgs, and mid-point elevations for PW4 at 5, 8.5, and 11.5 mbgs. During the installation of these wells, bentonite was placed between adjacent sand packs around the well screens to isolate potential hydraulic connections, except for the well nests (PW2 and PW4), which were directly installed in contact with the native formation. A three-dimensional perspective view of the wells is shown in Figure 2, indicating the locations of various screens and bentonite seals at the site.



Figure 1. Plan view of the 15 m by 15 m study area showing the well locations. The circles indicate continuous multichannel tubing (CMT) wells, the stars indicate four-inch multiple screen pumping wells (PW), and the crosses indicate two-inch well nests (from [18]).



Figure 2. Three-dimensional perspective view of various wells and pumped locations. The red squares are pumped locations to generate drawdown data for model calibration, and the green square indicates the same, but the data are used for model validation (modified after [18]).

2.2. Sensors and Data Collection

A groundwater level monitoring network consisting of various types of transducers was installed at the NCRS to collect pressure head changes at up to 44 observation ports depending on the particular pumping location. For CMT wells, each channel was instrumented with a 0 to 15 psig (model MP100: Micron Systems, Micron Instruments, Simi Valley, CA, USA) pressure transducer. Water level changes in PW2 and PW4 nest wells were monitored with 0 to 5 or 0 to 10 psig pressure transducers (model 3001 LT Leveloggers Junior, Solinst Canada Ltd., Georgetown, ON, Canada). Since PW1, PW3, and PW5 wells were designed to stress the aquifer–aquitard systems at different elevations, an inflated double-packer pumping system was used when pumping from one of the multi-screened wells, and FLUTe (FLUTe Ltd., Albuquerque, NM, USA) liner systems were installed in the other two wells, to prevent hydraulic short-circuiting between adjacent screens. Each of the two liners contained five vented pressure transducers (Level Troll, In Situ Inc., Fort Collins, CO, USA) placed at depths corresponding to the pumping well (PW) screens. These liners could be moved among PW1, PW3, and PW5 due to their similar well constructions. When pumping from PW2 or PW4 wells, a blank FLUTe liner was added to one of the multi-screened wells.

2.3. Pumping Test Description and Data Selection

A total of 15 pumping/injection tests were performed in a tomographic fashion at the NCRS and are summarized in Table 1. Nine tests were conducted previously by [18]. Of the 15 tests, 12 were conducted by pumping water, while three were conducted through water injection. The tests ranged in duration from 4.4 h to 26.5 h. Pumping tests in PW1, PW3, and PW5 wells were performed through a submersible pump (Model SQE05, Grundfos Canada Inc., Oakville, ON, Canada) while pumping tests in PW4 and PW2 well nests were performed using a surface lift pump. Although site investigation

5 of 18

revealed that the aquifer is generally confined, pumping tests were intentionally performed during the dry periods of the summer of 2009 and 2014 to avoid groundwater recharge.

Well Location	Pumping/Injection Rate (L/min) ^c	Duration (hour)	Maximum Pressure Head Change (m)	Port of Maximum Pressure Head Change
PW1-1 ^b	-1.89	4.5	0.55	CMT2-2
PW1–3 ^a	10.50	6	1.10	CMT4-3
PW1–4 ^a	6.30	8.5	0.98	CMT2-4
PW1-5 ^a	4.40	22.5	0.18	CMT2-4
PW1-6 ^b	0.95	6.5	0.74	CMT3-6
PW1-7 ^b	1.05	26.5	0.42	CMT3-6
PW2-3 ^b	1.91	7	0.30	CMT4-3
PW3-1 ^b	-0.94	4.4	0.21	CMT1-2
PW3–3 ^a	2.10	22	0.57	CMT1-3
PW3–4 ^a	1.50	22	0.08	CMT1-5
PW4–3 ^a	30.20	22.5	0.88	CMT3-4
PW5-1 ^b	-0.85	4.52	0.12	CMT2-1
PW5–3 ^a	7.80	22	0.92	CMT2-3
PW5-4 ^a	7.80	8.5	0.45	CMT2-4
PW5–5 ^a	8.10	22	0.32	CMT2-3

Table 1. Summary of Pumping/Injection Tests Performed at the North Campus Research Site.

^a—conducted by [18] (Summer, 2009); ^b—conducted for this study (Summer, 2014); ^c—positive values represent pumping tests, negative values represent injection tests.

Data from 14 pumping tests are selected at various times for HT interpretation, as early as 1 min and as late as 1600 min depending on the pumping duration and hydraulic responses of each observation ports. For each drawdown curve, one to three points were selected to define the curve. During some pumping tests, some ports showed no obvious response. These data were also included in the inversion as they may provide information on connectivity between the pumped and observation ports. In addition, some deep ports (mostly at ports 6 and 7) of CMT wells recorded negative drawdowns, known as the Noordbergum effect [29,30]. For these ports, no data were selected since the groundwater flow model used for this study does not consider poroelastic processes. In total, 813 data points were selected from 14 pumping/injection tests for THT analysis.

3. Inverse Modeling

3.1. Inverse Modeling Approach

We performed the geostatistical inverse modeling of 14 tests using the simultaneous successive linear estimator (SimSLE), which was developed by Xiang et al. [31] based on the sequential successive linear estimator (SSLE) [4,22]. Unlike SSLE, which analyzes data sequentially, SimSLE can simultaneously include all data sets from multiple pumping tests. This can improve computational efficiency and also avoid the issue of reaching a slightly different final estimate by providing more constraints to the inverse problem than SSLE [31]. A brief description of the method is provided below.

The inverse model of SimSLE assumes that the natural logarithm of K (ln K) and S_s (ln S_s) can be treated as stationary stochastic processes, and the mean values, variances, and correlation structure of the K and S_s fields are known *a priori*. The inversion process of SimSLE starts by creating the cokriged mean removed ln K and ln S_s (f, i.e., perturbations of ln K and ln S_s) fields, based on the initial estimates of effective K (K_{eff}) and effective S_s (S_{seff}) as well as selected pressure head (h) data from all

pumping tests. During this step, the exponential model is used for the covariance function as required in cokriging, and the sensitivity matrix is evaluated using the adjoint-state approach [22,31]. Then, the cokriged perturbation fields are added back to the logarithm of K_{eff} and S_{seff} , back-transformed, and used to solve the governing flow equation.

Next, SimSLE uses the successive linear estimator (SLE) [32] to update the above estimated *K* and S_s fields, based on the differences between the simulated and observed *h* values. Assuming the model domain is discretized into *N* elements, *Y* is an *N* × 1 vector containing *K* or S_s estimates. The estimator can be generally described by Equation (1).

$$\chi_{c}^{r+1} = Y_{c}^{r} + \chi^{r}(\psi_{obs} - \psi^{r}),$$
(1)

where Y_c^r is the estimated parameter of Y at iteration r (r > 1). When r = 1, Y_c^1 is the cokriged K or S_s field described above. ψ_{obs} is the vector of observed pressure head, ψ^r is the simulated pressure head with hydraulic parameters estimated at the r^{th} iteration, χ^r is the weight matrix calculated using the covariance of observed pressure heads, the cross-covariance between observed pressure heads and parameters to be estimated. This step is carried out in an iterative fashion by SimSLE until the following convergence criteria are met: (1) the spatial variances of the K and S_s estimates stabilize or (2) the changes in simulated h between successive iterations are smaller than the prescribed tolerance.

3.2. Inverse Model Setup

The model domain used for all inversions and forward simulations was 90 m by 90 m by 15 m using an adaptive mesh. Generally, the size of the element decreases from the model boundary to the central portion of the model domain. In total, the model domain consists of 40,424 elements and 37,664 nodes. The boundary conditions were set as no flow for the top and bottom boundaries, and constant head boundaries for the remaining outer boundaries and the system was assumed to be at steady state before pumping.

The stochastic inversion of THT analysis of 14 pumping tests was performed on a PC-cluster consisting (of 1 head node and 2 compute nodes) at the University of Waterloo using 32 processors and a maximum RAM of 130 GB. The system of the cluster was CentOS 5.3 based on a 64-bit system, running the Rocks 6.1 Linux Cluster Distribution from San Diego Supercomputer Center, and the total computational time was approximately seven days for the inversion of all 14 pumping tests.

SimSLE needs initial estimates of the hydrogeologic structure information of the aquifer, such as the correlation lengths (λ_x , λ_y , λ_z) and variances (σ^2_{lnK} , σ^2_{lnSs}) to provide as model inputs. Since estimates of σ^2_{lnK} have been shown to have a negligible effect on the resulting tomograms [4], we used a value of 5.0 as an initial value for both *K* and *S*_s. We selected an initial horizontal correlation length of 10 m for both *K* and *S*_s and the vertical correlation length of 0.5 m, which is estimated as the thickness of one element block for the model. Additionally, convergence criteria were set as that the maximum head or variance change in the estimated fields is less than 0.01.

4. Results

4.1. K and S_s Tomograms

Figure 3a,c show the final *K* and S_s tomograms, respectively, after including the drawdown data from 14 pumping tests, while Figure 3b,d show the corresponding variance maps. In general, the THT analysis of 14 pumping tests captures the alternating layers of high and low *K* zones in the central area of the study area. Meanwhile, the upper and lower portions of the area are recognized as low *K* zones, which correspond well with the borehole logs that are known to be silt/clay. Examination of the variance maps (Figure 3b,d) indicates that the estimates in this region have a low degree of uncertainty within the central 15 m by 15 m area compared to the outer area toward the boundaries. The uncertainty was also lower near the top and the bottom of the domain, which was not found by Berg and Illman [18] who inverted only 4 pumping tests for their study. Such reductions of uncertainties indicate the positive impact of using additional pumping and observation data for areas that previously had no observable drawdowns by Berg and Illman [18].



Figure 3. The estimated tomograms of (**a**) K and (**c**) S_s and corresponding variances of (**b**) ln-K and (**d**) ln- S_s obtained from transient hydraulic tomography (THT) analysis of 14 pumping tests.

4.2. Model Calibration

The quality of model calibration was first examined by comparing the simulated drawdowns using the estimated *K*- and S_s -tomograms to the observed drawdowns of the 14 pumping tests. Forward simulations were performed using the groundwater model *MMOC3* [33]. Figure 4 is a scatterplot comparing simulated and observed drawdowns at all ports for all 14 pumping tests at the same time points included in the THT code for the creation of these tomograms. In Figure 4, the black solid line indicates a perfect 1:1 correlation, while the red dashed line is the linear model fit to the data, the mean absolute error (*L*1) norm, the mean square error (*L*2) norm, and the coefficient of determination (R^2). Those quantities are computed as

$$L1 = \frac{1}{n} \sum_{i=1}^{n} |\chi_i - \hat{\chi}_i|,$$
(2)

$$L2 = \frac{1}{n} \sum_{i=1}^{n} (\chi_i - \hat{\chi}_i)^2,$$
(3)

$$R^{2} = \left[\frac{\frac{1}{n}\sum_{i=1}^{n}(\chi_{i}-\mu_{\chi})(\hat{\chi}_{i}-\mu_{\hat{\chi}})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\chi_{i}-\mu_{\chi})^{2}\cdot\frac{1}{n}\sum_{i=1}^{n}(\hat{\chi}_{i}-\mu_{\hat{\chi}})^{2}}}\right]^{2},$$
(4)

where *n* is the total number of input observation data, *i* indicates the data number, and χ_i and $\hat{\chi}_i$ represent the simulated and measured drawdowns, respectively, and μ_{χ} and $\mu_{\hat{\chi}}$ are corresponding means of the simulated and measured drawdowns, respectively. Overall, the drawdowns were evenly

distributed along the 1:1 line with minimal bias, with a coefficient of determination (R^2) of 0.88 and low *L*1 and *L*2, indicating a good match for the drawdown for both pumping and injection tests.



Figure 4. Calibration scatterplot of simulated versus observed drawdowns for THT analysis.

4.3. Comparison of the THT Results to Permeameter Test K

Next, we compared the results from THT to the *K* values obtained from permeameter tests along nine wells (CMT1, CMT2, CMT3, CMT4, PW1, PW2, PW3, PW4, and PW5). The results are shown in Figure 5; the red lines are results from THT analysis of 14 pumping tests, and the black lines are results from permeameter tests. Blue lines are the results of the THT analysis of four pumping tests by Berg and Illman [18], which will be discussed later. Generally, the correspondences of the main high and low *K* features are very good for CMT wells, PW1, and PW2 wells along their entire depths. For PW3, PW4, and PW5 wells, although the matches are not as good as those in CMT wells, the general trends of varying *K* values are captured quite well.

4.4. Model Validation

To further evaluate the predictive capability of the estimated tomograms, the pumping test (PW1-5) not used in calibration effort was simulated at all observation plots and compared to the observed drawdown curves at available ports for both CMT and PW wells, as shown in Figure 6, together with the manually measured data with an electric sounder. The red lines are simulated drawdown curves using 14 tests for THT analysis, while the black dashed lines are simulated drawdown data by Berg and Illman [18]. The green lines represent the observed curves, and blue dots are manual measurements in each subplot. Visual comparisons reveal that the simulated drawdown curves generally capture the trend of the corresponding observed curves.



Figure 5. Vertical K profile along 9 boreholes for THT analysis with 14 pumping tests.

Additionally, drawdown data at several times (1, 2, 5, 50, 100, 300, 600, 1000 min) covering the early, intermediate, and late time of pumping test were selected from the ports with observed data to quantitatively evaluate the fit for the validation test, as shown in Figure 7. Data from different ports are represented with symbols of different type and color. Again, a solid black line indicating a perfect 1:1 correlation is plotted, and the dashed black line represents the best fit of observed and simulated values. In general, the points cluster along the 1:1 line with a minor bias, although the fit is not as good as the one shown in Figure 4 for model calibration. The overall fit has an R^2 value of 0.51 and a slope of 0.83, indicating that there is a fairly good correspondence between simulated and observed drawdowns.

Water 2019, 11, 1864



Figure 6. Cont.



Figure 6. Observed drawdown in (**a**) CMT and (**b**) PW ports during the pumping test at PW1–5. The blue dots are manual measured data with a water level tape, the green dots are the observed drawdown data, the red lines represent simulated drawdown data using 14 tests for THT analysis, and the black dashed lines are simulated drawdown data by Berg and Illman [18].



Figure 7. Validation scatterplot for different observation ports during pumping at PW1-5 using THT analysis results. The black solid line suggests a perfect 1:1 correlation, while the dashed line is the best-fit line. The linear fit results are also provided. L1 is the mean absolute error norm, and L2 is the mean square error norm, representing the correspondence between observed and simulated drawdowns.

5. Discussion

5.1. Comparison of Results to Those from Berg and Illman [18]

To further evaluate the improvements of the 14-test THT in our current study over the 4-test THT study by Berg and Illman [18], we extracted two cross-sections (CMT1 to CMT2 and CMT3 to CMT4) of *K* and S_s tomograms from Figure 3 and compared to those from Berg and Illman [18] along with a stratigraphy map, as shown in Figure 8. We also plotted the vertical *K* profiles along nine boreholes from the THT analysis of Berg and Illman [18], as shown in Figure 5, and provided transient drawdown predictions, as shown in Figure 6, for more detailed comparisons.

In Berg and Illman [18], a 45 m by 45 m by 15 m numerical model was built, and pumping tests conducted at ports PW1-3, PW3-3, PW4-3, and PW5-3 were used for their THT analysis. For this study, a larger modeling domain (90 m by 90 m by 15 m) was constructed, and more pumping tests were carried out at additional wells to obtain complementary response information from different locations of the aquifer–aquitard system. Thus, we first compare the *K* and *S*_s estimates in Figure 8g–j with the Figure 8c–f focusing on the central 45 m by 45 m to see the impact of utilizing additional pumping test data.



Figure 8. Comparison of stratigraphy map and the *K* and S_s tomograms: (**a**) stratigraphy section (CMT1 to CMT2); (**b**) stratigraphy section (CMT3 to CMT4); (**c**) *K* tomogram section (CMT1 to CMT2) from Berg and Illman [18]; (**d**) *K* tomogram section (CMT3 to CMT4) from Berg and Illman [18]; (**e**) S_s tomogram section (CMT1 to CMT2) from Berg and Illman [18]; (**f**) S_s tomogram section (CMT3 to CMT4) from Berg and Illman [18]; (**g**) *K* tomogram section (CMT1 to CMT2) from THT using 14 tests; (**h**) *K* tomogram section (CMT3 to CMT4) from THT using 14 tests; (**i**) S_s tomogram section (CMT1 to CMT2) from THT using 14 tests; (**i**) S_s tomogram section (CMT1 to CMT2) from THT using 14 tests; (**i**) S_s tomogram section (CMT1 to CMT2) from THT using 14 tests; (**i**) S_s tomogram section (CMT1 to CMT2) from THT using 14 tests; (**i**) S_s tomogram section (CMT1 to CMT2) from THT using 14 tests; (**i**) S_s tomogram section (CMT3 to CMT4) from THT using 14 tests; (**i**) S_s tomogram section (CMT3 to CMT4) from THT using 14 tests; (**i**) S_s tomogram section (CMT3 to CMT4) from THT using 14 tests; (**i**) S_s tomogram section (CMT4) from THT using 14 tests; (**i**) S_s tomogram section (CMT4) from THT using 14 tests.

Comparing the *K* and S_s tomogram sections of current study (Figure 8g–j) to those from Berg and Illman [18] (Figure 8c–f), we observe that: (1) the high *K* zones located at the top and bottom areas of Figure 8c,d are now properly mapped as low *K* zones in Figure 8g,h, which correspond well with the aquitard layers consisting of silt, silt and clay, and clay in Figure 8a,b; (2) the features of two high *K* aquifer layers as well as the aquitard layer in between are now more clearly captured in terms of their layer shapes and continuity in Figure 8g,h than those in Figure 8c,d; (3) high S_s values are assigned to the top and bottom areas and relatively low S_s values are estimated in the central domain of the Figure 8i,j. Such a distribution pattern generally corresponds with our geological knowledge of a double-layer aquifer zone overlain and underlain by the aquitard layers at the top and bottom of the modeling domain, whereas S_s distributions in Figure 8e,f are quite homogeneous and no visible geological features are revealed. These results clearly indicate that, with the inclusion of sufficient data from the area of interest, THT analysis is able to recover the *K* and S_s heterogeneity which is consistent with geological data. However, we should note that the specific shapes of high and low *K* layers are not exactly the same as the stratigraphy map, as shown in Figure 8a,b. Meanwhile, the locations of the upper high *K* zone in Figure 8c,d,g,h (z = 9 to 11 m) are not strictly consistent with the locations of the sand and gravel layer indicated in Figure 8a,b (z = 8 to 9 m). One potential explanation is the relatively coarse vertical observation interval (2 m) compared to the <1 m thickness of the "sand and gravel" layer. Another potential cause is that the stratigraphy is interpolated from borehole logs, thus only providing estimated locations of the layer boundaries. Interpolation of stratigraphy to construct the geological model was conducted using the commercial software Leapfrog Hydro (ARANZ Geo Limited, Vancouver, BC, Canada). Leapfrog Hydro utilizes the Fast Radial Basis Function method, which is an effective way of implementing dual kriging that interpolates the stratigraphy between boreholes based on the known geological information from available wells. A complete evaluation of the tomograms will require additional drilling efforts such as with direct push techniques or geophysical surveys.

Examination of Figure 5 reveals more details of the estimated *K* distributions along nine boreholes between the current study and those from Berg and Illman [18]. In Figure 5, the black, red, and blue lines are *K* values obtained from the permeameter tests, the 14-test THT, and 4-test THT, respectively. In general, the red lines show closer matches to the black lines, while some obvious biases are found among the fits between the blue lines with the black lines, especially in the ranges of 0 to 5 m and 12 to 15 m above the bottom boundary of the domain. Meanwhile, the red lines show a better fit with the black lines for the high *K* zones located in the range of 5 to 12 m in the subplots of wells PW1, PW2, PW4, CMT1, CMT3, and CMT4. In contrast, we see that the fits for well PW3 are poorer among the black, red, and blue lines, indicating that more complementary information (e.g., pressure head, flux, and geophysics data) are needed to improve the estimates at these well locations.

Meanwhile, drawdown predictions in Figure 6 show that the simulated drawdowns by 14-test THT results (red lines) have better captured the shapes and trends of the observed drawdowns (green and blue dots) than the drawdown simulated by Berg and Illman [18].

These results collectively suggest that, when pumping test data are collected from additional locations of the highly heterogeneous aquifer–aquitard system for the THT analysis, the estimated K and S_s tomograms have improved correspondence with the stratigraphic model and vertical K profiles along boreholes, which in turn lead to more accurate drawdown predictions not used in model calibration efforts.

5.2. Advantages and Limitations of the Technology

Prior research has shown that HT data inherently contain more information than single-well pumping tests, and the joint interpretation method is superior to conventional data analysis methods in delineating subsurface heterogeneities [18]. Conventional hydrogeological characterization techniques include the interpretation of borehole core or cuttings samples and aquifer pumping tests. Although high-resolution information may be obtained using borehole sampling, it is invasive and cost-intensive (especially for deep formations), and only provides geologic information at a finite number of discrete locations. Considerable uncertainty exists as to the geologic materials between boreholes, and spatial interpolation techniques may be misleading. Single-well pumping tests are commonly interpreted by assuming aquifer homogeneity and yield non-unique spatially averaged properties over the cone of depression when a Theis or Cooper–Jacob method of analysis is used (e.g., [15,34]). Geophysical methods have increasingly been used to supplement conventional characterization by producing a high-resolution image of the subsurface. Although these methods can be relatively quick and inexpensive to perform, they only provide a high-resolution image of geophysical properties instead of hydrogeologic properties. Site-specific petrophysical relationships may have to be developed to translate the geophysical properties to corresponding hydrogeologic properties which lead to considerable uncertainty. As a result, our ability to characterize the spatial distribution of hydrogeologic properties is still quite limited. This is one of the major causes of the failure of many remediation systems. Although HT technology infers from the test data potential low permeability areas that may be contaminant sources zones, HT technology does not actually identify contaminant source areas. A limitation of HT is that the resolution of results is dictated by the density of pumping wells and observation ports in wells. For example, Yeh and Liu [4] suggested that spacing of the observation ports in observation wells should be about the averaged thickness of the heterogeneity to be mapped in the vertical direction. Likewise, the spacing in the horizontal direction should be approximately the horizontal extent of the stratification. They also suggested pumping at four different locations (depths and directions) would be sufficient enough (i.e., the return of extensive pumping diminishes rapidly although it is still useful) through their synthetic study.

Based on the results presented here, the HT analysis of four tests by Berg and Illman [18] yielded a coarser map of K and S_s heterogeneity. With the HT analysis of 14 pumping tests, this study showed significant improvement in the results. However, a further study is needed to examine the number of pumping tests and monitoring data that are required to "sufficiently" characterize a site to address various water quantity and quality issues faced by various parties.

6. Summary and Conclusions

In this study, 15 pumping/injection tests were performed in a 3D network of pumping and observation wells to evaluate Hydraulic Tomography (HT) using the code developed by Xiang et al. [31], based on a geostatistical inverse method. The collected pumping tests data was first used to estimate the heterogeneous distributions of K and S_s through selecting pumping tests at different locations covering the upper, central, and lower portions of the study area at the NCRS.

THT analysis of transient drawdown data from 14 pumping tests showed that the main features of the glaciofluvial aquifer–aquitard system can be captured very well based on a comparison to known geology and comparisons to permeameter *K* data, as well as the simulation of an independent pumping test not used in the calibration effort. In particular, this field THT analysis captured the presence of dual aquifers and the existence of low *K* zone at the bottom area and at the upper zone covering the study area.

Compared to the previous THT study by Berg and Illman [18] who used only four pumping tests for inverse modeling, our THT analysis using data from 14 pumping tests led to better characterization of K and S_s heterogeneity for the highly heterogeneous NCRS site, improved K correspondence with K profiles along nine boreholes, and more accurate drawdown predictions for a pumping test not used during model calibrations.

Based on the results of this study, we conclude that THT is a robust approach to characterize aquifer heterogeneity at the local scale. Due to the highly heterogeneous nature at the NCRS site, drawdown responses are monitored at a high density. However, recent studies [14,27] show that the inclusion of other information (i.e., geological data) can help THT to improve the estimates. Thus, more work needs to be done to validate THT at the site using more information other than pressure heads. However, these will be the topics for future investigations.

Author Contributions: Z.Z. conducted the additional pumping tests, interpreted the data and wrote the manuscript; W.A.I. directed the research and edited the manuscript as the Ph.D. supervisor of Z.Z.; Y.Z. modified the code and provided guidance on the analysis; T.-C.J.Y. and C.M.B.M. both provided guidance on the test design, its interpretation, and edited the manuscript; S.J.B. provided guidance on data collection and interpretation, and D.H. supervised the work of Z.Z. at his current place of employment.

Funding: This research was supported by the Environmental Security and Technology Certification Program (ESTCP) under grant ER201212 awarded to C.M.B.M, T.C.J.Y. and W.A.I. Additional support was provided by the National Natural Science Foundation of China (NSFC) under grant 41807202. The APC was funded by the National Natural Science Foundation of China.

Acknowledgments: Z.Z. acknowledges the support of "China Scholarship Council". W.A.I. acknowledges the additional support by the Discovery and Collaborative Research and Development Grants from the Natural Sciences and Engineering Research Council of Canada (NSERC), as well as grants from the Ontario Research Foundation (ORF) and Canada Foundation for Innovation (CFI). We thank two anonymous reviewers for their helpful comments.

Conflicts of Interest: The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- Sudicky, E.A. A natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity and its role in the dispersion process. *Water Resour. Res.* 1986, 22, 2069–2082. [CrossRef]
- 2. Alexander, M.; Berg, S.J.; Illman, W.A. Field Study of Hydrogeologic Characterization Methods in a Heterogeneous Aquifer. *Groundwater* **2011**, *49*, 365–382. [CrossRef]
- 3. Butler, J.J., Jr.; Dietrich, P.; Wittig, V.; Christy, T. Characterizing Hydraulic Conductivity with the Direct-Push Permeameter. *Groundwater* **2007**, *45*, 409–419. [CrossRef] [PubMed]
- 4. Yeh, T.C.; Liu, S. Hydraulic tomography: Development of a new aquifer test method. *Water Resour. Res.* **2000**, *36*, 2095. [CrossRef]
- 5. Liu, S.; Yeh, T.C.J.; Gardiner, R. Effectiveness of hydraulic tomography: Sandbox experiments. *Water Resour. Res.* **2002**, *38*, 2–10. [CrossRef]
- 6. Gottlieb, J.; Dietrich, P. Identification of the permeability distribution in soil by hydraulic tomography. *Inverse Probl.* **1995**, *11*, 353–360. [CrossRef]
- 7. Cardiff, M.; Barrash, W.; Kitanidis, P.K.; Malama, B.; Revil, A.; Straface, S.; Rizzo, E. A potential-based inversion of unconfined steady-state hydraulic tomography. *Groundwater* **2009**, *47*, 259–270. [CrossRef]
- 8. Hu, R.; Brauchler, R.; Herold, M.; Bayer, P. Hydraulic tomography analog outcrop study: Combining travel time and steady shape inversion. *J. Hydrol.* **2011**, *409*, 350–362. [CrossRef]
- 9. Jiménez, S.; Brauchler, R.; Bayer, P. A new sequential procedure for hydraulic tomographic inversion. *Adv. Water Resour.* **2013**, *62*, 59–70. [CrossRef]
- Soueid Ahmed, A.; Jardani, A.; Revil, A.; Dupont, J.P. Specific storage and hydraulic conductivity tomography through the joint inversion of hydraulic heads and self-potential data. *Adv. Water Resour.* 2016, *89*, 80–90. [CrossRef]
- 11. Soueid Ahmed, A.; Zhou, J.; Jardani, A.; Revil, A.; Dupont, J.P. Image-guided inversion in steady-state hydraulic tomography. *Adv. Water Resour.* **2015**, *82*, 83–97. [CrossRef]
- Illman, W.A.; Liu, X.; Craig, A. Steady-state hydraulic tomography in a laboratory aquifer with deterministic heterogeneity: Multi-method and multiscale validation of hydraulic conductivity tomograms. *J. Hydrol.* 2007, 341, 222–234. [CrossRef]
- 13. Brauchler, R.; Böhm, G.; Leven, C.; Dietrich, P.; Sauter, M. A laboratory study of tracer tomography. *Hydrogeol. J.* **2013**, *21*, 1265–1274. [CrossRef]
- 14. Zhao, Z.; Illman, W.A.; Berg, S.J. On the importance of geological data for hydraulic tomography analysis: Laboratory sandbox study. *J. Hydrol.* **2016**, *542*, 156–171. [CrossRef]
- 15. Straface, S.; Yeh, T.C.J.; Zhu, J.; Troisi, S.; Lee, C.H. Sequential aquifer tests at a well field, Montalto Uffugo Scalo, Italy. *Water Resour. Res.* 2007, 43, 1–13. [CrossRef]
- 16. Bohling, G.C.; Butler, J.J.; Zhan, X.; Knoll, M.D. A field assessment of the value of steady shape hydraulic tomography for characterization of aquifer heterogeneities. *Water Resour. Res.* **2007**, *43*, 1–23. [CrossRef]
- 17. Li, W.; Englert, A.; Cirpka, O.A.; Vereecken, H. Three-dimensional geostatistical inversion of flowmeter and pumping test data. *Groundwater* **2008**, *46*, 193–201. [CrossRef]
- 18. Berg, S.J.; Illman, W.A. Three-dimensional transient hydraulic tomography in a highly heterogeneous glaciofluvial aquifer-aquitard system. *Water Resour. Res.* **2011**, *47*, W10507. [CrossRef]
- 19. Hochstetler, D.L.; Barrash, W.; Leven, C.; Cardiff, M.; Chidichimo, F.; Kitanidis, P.K. Hydraulic Tomography: Continuity and Discontinuity of High-K and Low-K Zones. *Groundwater* **2016**, *54*, 171–185. [CrossRef]
- 20. Wang, Y.-L.; Yeh, T.-C.J.; Wen, J.-C.; Huang, S.-Y.; Zha, Y.; Tsai, J.-P.; Hao, Y.; Liang, Y. Characterizing subsurface hydraulic heterogeneity of alluvial fan using riverstage fluctuations. *J. Hydrol.* **2017**, *547*, 650–663. [CrossRef]
- 21. Illman, W.A.; Liu, X.; Takeuchi, S.; Yeh, T.C.J.; Ando, K.; Saegusa, H. Hydraulic tomography in fractured granite: Mizunami Underground Research site, Japan. *Water Resour. Res.* **2009**, *45*, 1–18. [CrossRef]
- 22. Zhu, J.; Yeh, T.C.J. Characterization of aquifer heterogeneity using transient hydraulic tomography. *Water Resour. Res.* **2005**, *41*, 1–10. [CrossRef]

- Zha, Y.; Yeh, T.-C.J.; Illman, W.A.; Tanaka, T.; Bruines, P.; Onoe, H.; Saegusa, H.; Mao, D.; Takeuchi, S.; Wen, J.-C. An Application of Hydraulic Tomography to a Large-Scale Fractured Granite Site, Mizunami, Japan. *Groundwater* 2016, 54, 793–804. [CrossRef] [PubMed]
- 24. Berg, S.J.; Illman, W.A. Comparison of Hydraulic Tomography with Traditional Methods at a Highly Heterogeneous Site. *Groundwater* **2015**, *53*, 71–89. [CrossRef] [PubMed]
- 25. Illman, W.A.; Zhu, J.; Craig, A.J.; Yin, D. Comparison of aquifer characterization approaches through steady state groundwater model validation: A controlled laboratory sandbox study. *Water Resour. Res.* **2010**, *46*, 1–18. [CrossRef]
- 26. Berg, S.J.; Illman, W.A. Capturing aquifer heterogeneity: Comparison of approaches through controlled sandbox experiments. *Water Resour. Res.* **2011**, *47*, 1–17. [CrossRef]
- Zhao, Z.; Illman, W.A. On the importance of geological data for three-dimensional steady-state hydraulic tomography analysis at a highly heterogeneous aquifer-aquitard system. *J. Hydrol.* 2017, 544, 640–657. [CrossRef]
- Martin, P.J.; Frind, E.G. Modeling a Complex Multi-Aquifer System: The Waterloo Moraine. *Ground Water*. 1998, 36, 679–690. [CrossRef]
- 29. Verruijt, A. *Elastic Storage of Aquifers, Flow through Porous Media;* Academic Press: New York, NY, USA, 1969; pp. 331–376.
- 30. Hsieh, P.A. Deformation-Induced Changes in Hydraulic Head During Ground-Water Withdrawal. *Groundwater* **1996**, *34*, 1082–1089. [CrossRef]
- 31. Xiang, J.; Yeh, T.C.J.; Lee, C.H.; Hsu, K.C.; Wen, J.C. A simultaneous successive linear estimator and a guide for hydraulic tomography analysis. *Water Resour. Res.* **2009**, *45*, 1–14. [CrossRef]
- 32. Yeh, T.C.J.; Jin, M.; Hanna, S. An Iterative Stochastic Inverse Method: Conditional Effective Transmissivity and Hydraulic Head Fields. *Water Resour. Res.* **1996**, *32*, 85–92. [CrossRef]
- 33. Yeh, T.C.J.; Srivastava, R.; Guzman, A.; Harter, T. A Numerical Model for Water Flow and Chemical Transport in Variably Saturated Porous Media. *Groundwater* **1993**, *31*, 634–644. [CrossRef]
- 34. Wu, C.M.; Yeh, T.C.J.; Zhu, J.; Tim, H.L.; Hsu, N.S.; Chen, C.H.; Sancho, A.F. Traditional analysis of aquifer tests: Comparing apples to oranges? *Water Resour. Res.* **2005**, *41*, 1–12. [CrossRef]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).