



An Algorithm for Hydraulic Tomography Based on a Mixture Model

Carlos Minutti¹✉, Walter A. Illman², and Susana Gomez¹

¹ Institute of Research in Applied Mathematics and Systems,
National University of Mexico, Mexico City, Mexico
`carlos.minutti@iimas.unam.mx`

² Department of Earth and Environmental Sciences, University of Waterloo,
Waterloo, ON, Canada

Abstract. Hydraulic Tomography (HT) has become one of the most robust methods to characterize the heterogeneity in hydraulic parameters such as hydraulic conductivity and specific storage. However, in order to obtain high resolution hydraulic parameter estimates, several pumping/injection tests with sufficient monitoring data are necessary. In highly heterogeneous media, even with large numbers of measurements, the resolution may not be enough for predicting contaminant transport behavior. In addition, during inverse modeling, the groundwater flow equation is solved numerous times, thus the computational burden could be large, especially for a large, three-dimensional, transient model.

In this work we present a new approach to model aquifer heterogeneity, based on a Gaussian Mixture Model (GMM) to parameterize the K field, which significantly reduces the number of parameters to be estimated during the inversion process. In addition, a new objective function based on the spatial derivatives of hydraulic heads is introduced.

The developed approach is tested with synthetic data and data from a previously conducted sandbox experiments. Results indicate that the new approach improves the accuracy of the K heterogeneity map produced through HT and reduces the computational effort. For two dimensional synthetic experiments, this approach was able to achieve a significant reduction in the error for K field estimation as well as computational time compared to a geostatistical inversion approach. Similar results were also achieved when the approach was tested using pumping test data conducted in a synthetic aquifer constructed in the laboratory.

Keywords: Hydraulic tomography · Gaussian mixture ·
Aquifer heterogeneity

1 Introduction

During a hydraulic tomography experiment, water is sequentially pumped from or injected into an aquifer at different intervals of the aquifer. During each pumping/injection event, hydraulic head responses of the aquifer at different intervals

are monitored, yielding a set of head/discharge (or recharge) data. By sequentially pumping/injecting water at one interval and monitoring the steady state head responses at others, many head/discharge (recharge) data sets are obtained (Yeh and Liu [11]).

These head values can be compared with results given by the groundwater flow equation (the forward model) that describes the head response taking into account the hydraulic conductivity (K) values and the pumping/injection rate. The goal of hydraulic tomography (HT) is to find the K values throughout the aquifer that minimizes the difference between observed and simulated head values and this process is also known as the inverse problem.

In order to estimate the K values, many approaches have been developed. For example, Hoeksema and Kitanidis [4] used cokriging as an approach to model the K heterogeneity of the aquifer, but this approach relies on the availability of enough head and K measurements. Also, because cokriging is a linear estimator, its application and accuracy is limited due to the non-linear nature of the problem.

Gottlieb and Dietrich [3] minimized the quadratic error between the measured and the simulated head values using Tikhonov regularization to penalize how much the solution deviates from a geological *a priori* guess. The K value of each element in a finite element methods is optimized. The smoothness and accuracy of the solution can be affected by the amount of available data and how good the prior guess is. Also, the optimization of each K value is computationally expensive.

Yeh *et al.* [10] developed an iterative stochastic estimation approach called the Successive Linear Estimator (SLE) where the cokriged log- K field is updated at each iteration with new covariances and cross covariance estimates, in order to try to capture the non-linear relationship between the head values and K . Yeh and Liu [11] then extended the SLE approach by sequentially incorporating data sets from multiple pumping/injection tests in the inverse model, which resulted in a Sequential Successive Linear Estimator (SSLE).

Illman *et al.* [5] reported that the order in which data is included into the SSLE algorithm has a significant effect in the K estimates. Later, based on SSLE, Xiang *et al.* [9] presented the Simultaneous Successive Linear Estimator (SimSLE) which uses all data sets simultaneously in the inversion algorithm.

A number of studies on HT have been published which demonstrate the robust performance of SSLE and SimSLE. In particular, Illman *et al.* [6,7], and Berg and Illman [1] validated steady state hydraulic tomography (SSHT) based on SSLE with data collected from laboratory sandbox experiments. Xiang *et al.* [9] used SimSLE to reproduce the K field of synthetic data as well as data from an actual sandbox experiment. Zhao *et al.* [14] improved the results of Illman *et al.* [6,7], and Berg and Illman [1] by including geological data as a prior distribution of K in SimSLE.

Berg and Illman [2] compared different methods of modeling aquifer heterogeneity at a highly heterogeneous field site through data collected at the North Campus Research Site (NCRS) on the University of Waterloo campus,

and showed that transient hydraulic tomography (THT) analysis with SSLE performed the best. Subsequently, Zhao and Illman [12, 13] performed SSHT and THT analyses of pumping tests at the NCRS and improved the results of Berg and Illman [2] by including additional pumping tests and geological data as the initial distribution of K in the inversion process. They compared these results with the effective parameter and geological zonation approaches, showing that HT with SimSLE performs the best in terms of model calibration and validation. The inclusion of geological data was useful in producing less smooth hydraulic parameter distributions and also showed how the initial distribution of K produces different solutions of the K field, suggesting the necessity of taking into account the problem of local optimal solutions.

Regardless of the good results in HT that have been achieved with geostatistical inversion approaches such as SSLE and SimSLE, computational time could be an issue in the case of large, three-dimensional, transient models. In addition, geostatistical inverse approaches can depend significantly on the prior distribution of K , and the algorithms may suffer from slow convergence, especially for highly heterogeneous aquifers, because the hydraulic parameters are continuously updated to capture the non-linear dependence of hydraulic parameters and head.

In order to model the spatial distribution of K , in this work we propose a new algorithm, the GM inversion algorithm, to solve the inverse problem for HT applications. This algorithm is based on the statistical concept of a mixture model, specifically a Gaussian Mixture Model (GMM), which will be used to parameterize the K field.

In the proposed inversion algorithm, we introduce a new objective function, based on the spatial derivatives of head values. This function allows the inversion algorithm to increase the accuracy of K estimates and it makes use of a regularization term to address the ill-posed problem to estimate more parameters than available observations.

The results of the new inversion algorithm are compared with the results of SSLE using synthetic experiments as well as a laboratory sandbox aquifer. With this new approach, we were able to achieve good estimates of the K field, reducing the computational time as well as the amount of data needed to perform HT, even if the spatial distribution of K is complex or very heterogeneous.

2 The GM Inversion Algorithm

2.1 Parameterization of the Conductivity Field

The parameterization of the K field refers to the process in which a set of values, *i.e.* the parameters, models the K field of the aquifer, with each K field represented with a different set of parameter values.

Although the value of these parameters are not known *a priori*, it is possible to estimate them by comparing the different values of pressure (head) generated using a forward model (the groundwater flow equation), against the observed

data measured during the pumping/injection tests, *i.e.*, through the minimization of the error defined in an objective function.

The usual and the simplest selection for the parameterization of the K field is $G(\boldsymbol{\theta}) = \mathbf{K}$, *i.e.*, the parameterization is exactly \mathbf{K} , consisting of a vector of local K estimates. However, to optimize each K value for a K field can be computationally expensive, and possibly an ill-posed problem, thus we propose to use a parameterization function G , based on a Gaussian Mixture Model.

Gaussian Mixture Model. In statistics, a mixture model is a probabilistic model used to represent different sub-populations within a larger population, through the combination of many probability density functions where each component of the function could be interpreted as a “probability” or the likelihood of the existence of each sub-population in each spatial coordinate.

In this case, the main population would be the aquifer material in the porous media that has a larger volume or the one that is more spread out and the sub-populations would be any other materials found among the main population.

Although a mixture describes a probability distribution, its value can be re-scaled to describe different attributes, as a density or K values. In this approach, it describes the logarithm of K ($\log\text{-}K$).

If Gaussian distributions are used for the mixture model (GMM), this is known as the Gaussian Mixture Model, which is represented by:

$$G(\boldsymbol{\theta}) = G(X, \mu_i, \Sigma_i) = \sum_{i=1}^{N_k} \omega_i N(X, \mu_i, \Sigma_i) \quad (1)$$

where ω_i is a weight for the i -th Gaussian distribution with mean μ_i and covariance matrix Σ_i . N_k is the number of components in the mixture, and X is the vector of spatial coordinates.

If the K field has a complex spatial distribution K that does not resemble Gaussian kernels, it can be approximated by increasing the number of components in the mixture: as higher the number of components, the higher the complexity of the K field that can be modeled.

In order to model the spatial distribution of K with a GMM, we propose to approximate the true K field, \mathbf{K} , in the GM algorithm by:

$$\log(\mathbf{K}) \approx G(\boldsymbol{\theta}^*) - \log(\bar{K}) \quad (2)$$

where \bar{K} is the effective or mean K value of the aquifer and $G(\boldsymbol{\theta})$ is the GMM that minimizes the proposed objective function.

If Eq. (2) is used to approximate the spatial distribution of a two-dimensional K field, a set of bi-variate Gaussian kernels are needed to generate the GMM. For the GMM, the parameter N_k is the most important, as it is the number of components/materials in the mixture and the higher its value the greater the level of complexity that the GMM can model. Under certain assumptions, each component in the mixture could be considered as a layer of the material.

The rest of parameters in the GMM are estimated by the optimization of an objective function, and these parameters control the spatial correlation of the heterogeneity in the estimated K field. In particular, Σ_i , the covariance matrix, contains the variance in the spatial (horizontal and vertical) directions (σ_x^2 , σ_z^2), and the correlation between them (ρ_{xz}). In the case of layered materials, one can expect to have a larger horizontal variance (this constraint can also be included in the optimization process), because the layer is typically longer in the horizontal direction rather than the vertical. μ_i is the spatial location or *centroid* for each component in the horizontal and vertical directions, while ω_i is the weight of each component in the mixture. This value is directly related with the K value of each component/material.

2.2 Objective Function

Even though the proposed approach is closely related to cluster analysis and statistical classification, the main difference is that the K field to be approximated with a GMM is not known *a priori*, so it is necessary to estimate it, by using hydraulic head as a proxy variable of K .

Typically the error to be minimized in the inverse problem of HT would be:

$$F_{LS}(\boldsymbol{\theta}) = \sum_{i=1}^n (h_i(\boldsymbol{\theta}) - h_i^*)^2 \quad (3)$$

where h_i^* are the observed head values for all the pumping/injection tests and $h_i(\boldsymbol{\theta}) = H_i(G(\boldsymbol{\theta}))$ are the head values simulated using the forward model, where $G(\boldsymbol{\theta})$ is the estimated log- K field.

In addition, here, we include the spatial derivatives of head values to the objective function:

$$F_d(\boldsymbol{\theta}) = \phi_f F_{LS}(\boldsymbol{\theta}) + \phi_x \sum_{i=1}^n \left(\frac{d}{dx} h_i(\boldsymbol{\theta}) - \frac{d}{dx} h_i^* \right)^2 + \phi_z \sum_{i=1}^n \left(\frac{d}{dz} h_i(\boldsymbol{\theta}) - \frac{d}{dz} h_i^* \right)^2 \quad (4)$$

where ϕ_i are weights such that $\sum_i \phi_i = 1$, and x , z are the horizontal and vertical spatial directions.

The spatial derivatives are especially useful, because they are more sensitive to the locations of the Gaussian components, being easier for the optimization algorithm to place each component depending on the head change between two observed head values.

A downside for the use of the derivatives is that the noise level in the observed head values could be increased, but this problem is addressed by choosing a small value of N_k , or by using a noise reduction method such as splines or the one proposed in Minutti *et al.* [8], which is designed to address noisy data in inverse problem when the derivative is required.

Regularization. A measure of smoothness of the K field, is added as a regularization term in the objective function. This term can lead to a faster optimization process avoiding to “visit” many local optima that only fit parts of the calibration data. The smoothness of the solution is measured by:

$$S(\boldsymbol{\theta}) = \sum_{i=1}^m \left(\frac{d}{dx} G(\boldsymbol{\theta}) \right)_i^2 + \sum_{i=1}^m \left(\frac{d}{dy} G(\boldsymbol{\theta}) \right)_i^2 \quad (5)$$

where m is the number of K values used for the solution of the forward model.

Finally, the objective function of the optimization problem can be written as:

$$F(\boldsymbol{\theta}) = (1 - \lambda)F_d(\boldsymbol{\theta}) + \lambda S(\boldsymbol{\theta}) \quad (6)$$

where λ ($0 \leq \lambda < 1$), the regularization parameter is especially useful if there are few data measurements and/or there is noisy data.

Even though λ and N_k can be used as smoothing parameters, these two values represent different types of smoothness. High values of λ enforces K fields to have a low variance, *i.e.* more homogeneous aquifers, while small N_k values tend to group different layers of materials with similar K , but it allows high variance between groups. In other words, λ can be used to describe intralayer heterogeneity and N_k for interlayer heterogeneity.

3 Experiments

In order to test the capabilities of the GM algorithm, data from a sandbox and some synthetic experiments are used. These experiments intend to measure (1) the effect of the amount of data used; (2) the effect of noisy data; (3) the speed and accuracy in the estimation of the K field and (4) the performance when the spatial distribution of the K fields are Non-Gaussian. We utilize VSافت (Xiang *et al.* [9]), a previously utilized geostatistical inversion model, to solve the same problems for the purpose of comparing the results obtained through the GM algorithm.

For all experiments we use a two-dimensional heterogeneous aquifer that is based on the structure of the sandbox experiment presented by Illman *et al.* [7] (Fig. 1). This sandbox consists of 18 layers of sandy material, that measures 193.0 cm in length, 82.6 cm in height, and is 10.2 cm thick.

In order to simulate a synthetic scenario of this sandbox experiment, the aquifer was discretized into 741 elements (a grid of 19 by 39) with element dimensions of 4.1 cm by 4.1 cm. In terms of boundary conditions, the top, left and right boundaries of the aquifer are set as constant head boundaries, while the bottom is treated to be a no-flow boundary.

For each element the K values originates from the measured values presented in Table 1 of Berg and Illman [1], where the log-mean K value of its corresponding layer, is used. For the cases of an element that corresponds to more than one layer of material, the log-mean K value of those layers was used.

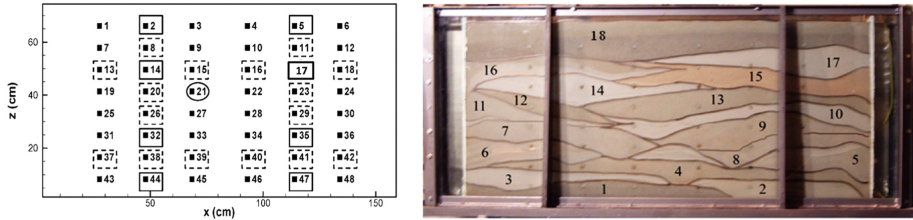


Fig. 1. Schematic diagram of the synthetic heterogeneous aquifer and Photograph of the synthetic heterogeneous aquifer from Illman *et al.* [7].

The above process results in a K field with a variance of $\sigma_{\log(K)}^2 = 0.1361$ (low-heterogeneity case). A high-heterogeneity case was simulated by re-scaling each K value (keeping the same mean $\log(K)$ value), resulting in a variance of $\sigma_{\log(K)}^2 = 0.8508$.

For the forward simulation and inverse modeling, a saturated and isotropic steady state forward model is used. Synthetic data are generated by simulating, nine pumping tests on the computer, with 48 observed head values for each one (8 rows and 6 columns), using a constant pumping rate of $Q = 1.25 \text{ cm}^3/\text{s}$. The data from the first 4 pumping tests are used as input for inverse modeling (this data is called the HT data set) and the remaining data from tests 5 to 9 (validation data set) were used for validation of the estimated field.

Finally, for the sandbox experiment, eight pumping tests were used as input to the inversion algorithms and the results are compared with the ones reported by Illman *et al.* [5].

4 Results

For each experiment, the Mean Squared Error (MSE) is calculated for the HT data set ($E(H_{HT})$) and for the validation data set ($E(H_V)$). For the synthetic experiments, the MSE of the $\log(K)$ is also estimated ($E(K)$) as well as a Relative MSE measure ($E_\sigma(K)$), which takes into account the variance of the $\log(K)$ values, being $E_\sigma(K) = E(K)/\sigma_{\log(K)}^2$. This allows one to compare the errors between the low and the high heterogeneity cases.

4.1 Synthetic Experiments

Low-Heterogeneity Case: Table 1 presents the results of the estimated K fields for the case $\sigma_{\log(K)}^2 = 0.1361$, with VSAFT2 and GM as inversion methods, using head values from 1 to 4 pumping tests as input data.

If only one pumping test is used, it can be seen that the error measurements for K and for the head values in the HT data set, are very similar. However the K field estimated with GM, has a higher variance, more similar to the true variance, and a smaller error of the head values in the validation data set.

Table 1. CPU-Time and different error measurements for the HT problem, in a synthetic aquifer with a low-heterogeneity K field ($\sigma_{\log(K)}^2 = 0.1361$).

N_{tests}	Method	$\hat{\sigma}_{\log(K)}^2$	$E(H_{HT})$	$E(H_V)$	$E(K)$	$E_{\sigma}(K)$	CPU-Time
1	VSAFT	0.0254	0.0018	0.0471	0.1012	0.7436	0.62
	GM	0.0632	0.0012	0.0178	0.0926	0.6804	0.17
2	VSAFT	0.0201	0.0058	0.0534	0.0950	0.6980	1.06
	GM	0.0799	0.0015	0.0085	0.0359	0.2638	0.23
3	VSAFT	0.0305	0.0032	0.0263	0.0711	0.5224	1.84
	GM	0.0997	0.0012	0.0018	0.0156	0.1146	0.25
4	VSAFT	0.0456	0.0017	0.0094	0.0481	0.3534	2.84
	GM	0.1165	0.0012	0.0011	0.0096	0.0705	0.53

N_{tests} = Number of pumping tests used, Method = Inversion method, $\hat{\sigma}_K^2$ = Variance of the estimated K field, $E(H_{HT})$ = MSE of the Head values used in the inversion process, $E(H_V)$ = MSE of the Head values used for validation purposes, $E(K)$ = MSE of the log- K values, $E_{\sigma}(K) = E(K)/\sigma_{\log(K)}^2$, CPU-Time = Minutes used to estimate the K field.

As the number of pumping tests is increased, the errors of K and of the head values, tend to be more significant between both methods. When 4 pumping tests are used, GM only has one fifth of the error of K and uses one fifth for the CPU-Time, in contrast with VSAFT.

Figure 2 presents the K tomograms generated with each inversion method, the scatter plot of the observed vs estimated K values and the drawdowns when 4 pumping test were used. Although both methods detect the main features of heterogeneity in K , the level of resolution obtained with GM is higher than the one obtained with VSAFT, suggesting that the K tomogram generated with GM better approximates the true K field. It can be seen that GM reproduces better the low and high K areas.

High-Heterogeneity Case: Table 2 presents the results of the estimation of the K field when $\sigma_{\log(K)}^2 = 0.8508$ with VSAFT and GM as inversion methods, using 1 to 4 pumping tests as input data.

Similar to the low-heterogeneity case, when only one pumping test was used for inverse modeling, the MSE of both inversion approaches are similar, but when the number of pumping tests increases, the differences between the two methods become more significant.

Examination of Table 2 reveals that when 4 pumping tests are used for inverse modeling, the GM algorithm has less than one fifth of the error of K than VSAFT, using less than one third of the computational time. The K tomograms for this case are presented in Fig. 3.

Mean Performance: Using the results of Tables 1 and 2, we compare the performance of both inversion algorithms for the low and high heterogeneity

Table 2. CPU-Time and error measurements for the HT problem in a synthetic aquifer with a high-heterogeneity K field ($\sigma_{\log(K)}^2 = 0.8508$).

N_{tests}	Method	$\hat{\sigma}_{\log(K)}^2$	$E(H_{HT})$	$E(H_V)$	$E(K)$	$E_{\sigma}(K)$	CPU-Time
1	VSAFT	0.1053	0.0243	0.1621	0.6848	0.8049	0.62
	GM	0.5253	0.0047	0.1853	0.5138	0.6039	0.37
2	VSAFT	0.0827	0.0758	0.2145	0.6350	0.7464	1.06
	GM	0.3857	0.0479	0.2544	0.3452	0.4057	0.28
3	VSAFT	0.1358	0.0391	0.0975	0.5331	0.6266	1.80
	GM	0.6654	0.0170	0.0155	0.1096	0.1288	0.87
4	VSAFT	0.2043	0.0215	0.0414	0.4079	0.4794	2.88
	GM	0.6647	0.0080	0.0023	0.0702	0.0825	0.82

N_{tests} = Number of pumping tests used, Method = Inversion method, $\hat{\sigma}_K^2$ = Variance of the estimated K field, $E(H_{HT})$ = MSE of the Head values used in the inversion process, $E(H_V)$ = MSE of the Head values used for validation purposes, $E(K)$ = MSE of the log- K values, $E_{\sigma}(K) = E(K)/\sigma_{\log(K)}^2$, CPU-Time = Minutes used to estimate the K field.

cases, where $E_{\sigma}(K)$ is the main measure for the error of K . Figure 4 presents the average $E_{\sigma}(K)$ value and the average CPU-Time by the number of pumping tests used for each inversion algorithm.

Based on this synthetic study, the results for the inversion of 4 pumping tests reveal that GM has on average one fifth the estimation error of VSAFT, using one fourth of the computation time. It is also important to note that the error achieved with GM using 2 pumping tests, is lower than the one with VSAFT using 4 pumping tests, *i.e.*, GM was able to achieve a similar error than VSAFT using only 2 pumping tests instead of 4 tests, with less than one tenth of the computational time.

4.2 Sandbox Case

Table 3 presents the results of the estimation of the K field for the sandbox experiment conducted by Illman *et al.* [7] using VSAFT and GM as inversion methods.

It can be seen that with 8 pumping tests the amount of CPU-Time required with GM is almost one fifth of the time used with VSAFT and the variance of the estimated K field is similar in both cases, as well as the error of the head values.

Even though it is not possible to calculate the estimation error of the K field, there are some indications of how good the estimation is. For example, the mean value of log- K is reported in Table 3 as $\log(K_G)$, and in both cases, the estimations are similar to the value of -2.56 , reported in Illman *et al.* [7] using 48 core samples. Although the variances of log- K are slightly different to the value estimated using core samples ($\sigma_{\log(K)}^2 = 0.868$), the variance obtained with the GM algorithm is closer to that value.

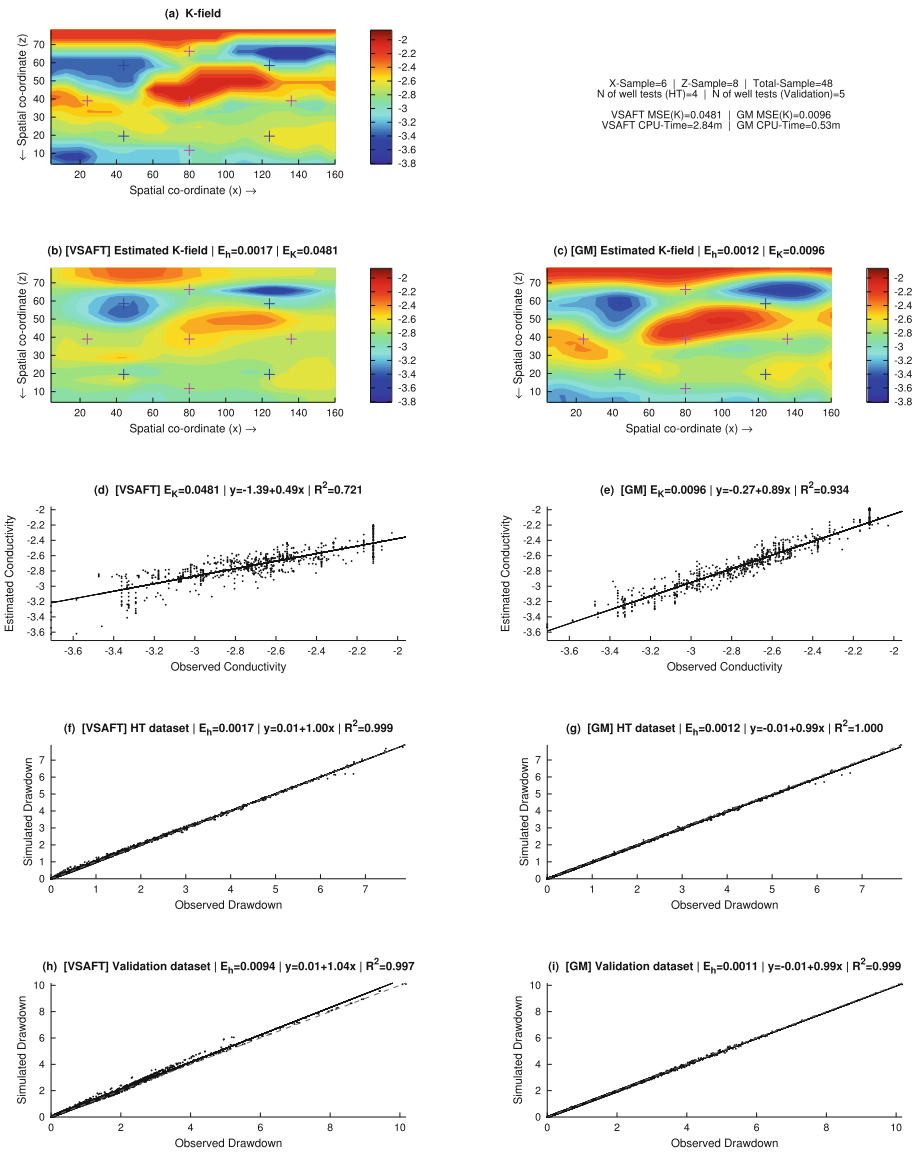


Fig. 2. (a) Synthetic K field of **low-heterogeneity** and the log- K tomograms obtained with (b) VSAFT and (c) GM as inversion methods, **using 4 injection tests** with 48 head values per test. The blue crosses mark the location of the injection wells and the red crosses the location of the injection tests used for validation (h)-(i). The dashed line is the 45-degree line, representing a perfect match, the solid line is a linear model fit with the slope, intercept, and the coefficient of determination (R^2) provided.

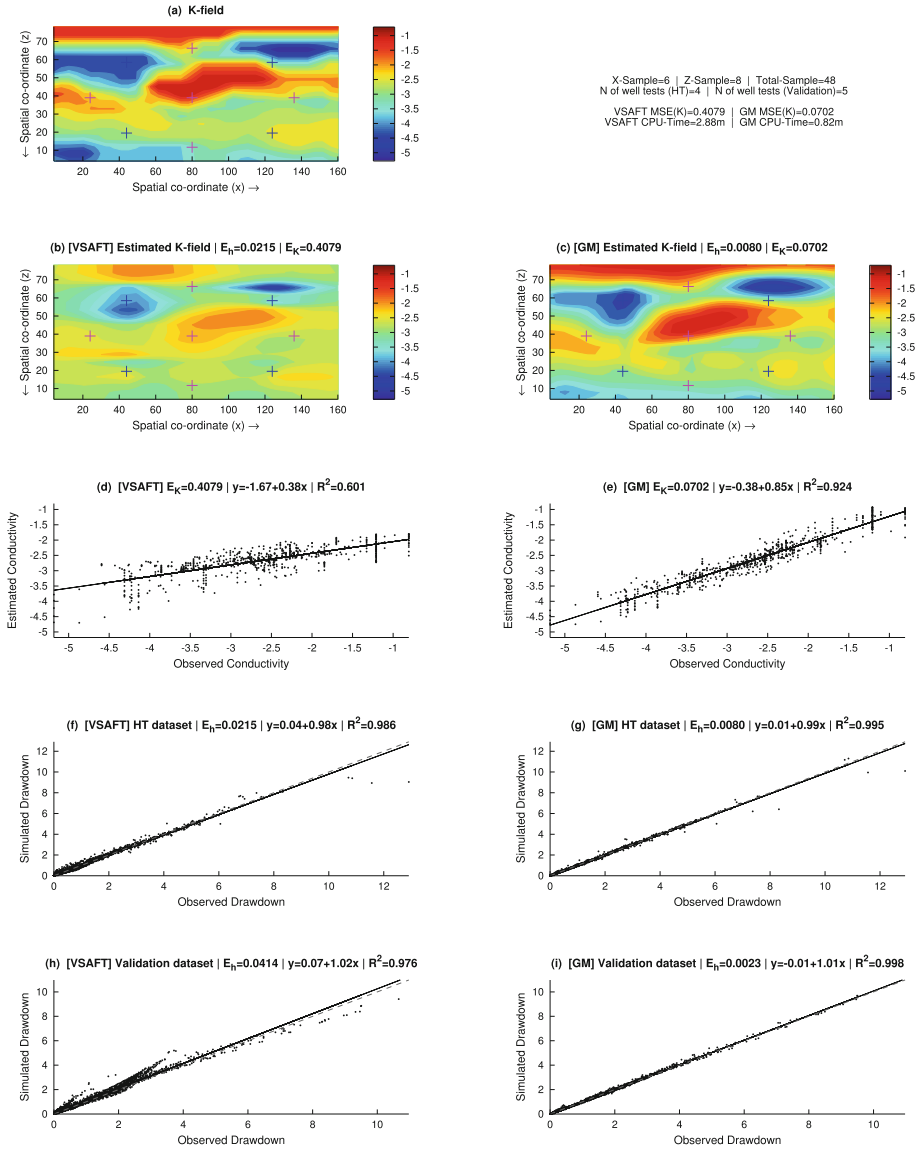


Fig. 3. (a) Synthetic K field of **high-heterogeneity** and the log- K tomograms obtained with (b) VSAFT and (c) GM as inversion methods, **using 4 injection tests** with 48 head values per test. The blue crosses mark the location of the injection wells and the red crosses the location of the injection tests used for validation (h)-(i). The dashed line is the 45-degree line, representing a perfect match, the solid line is a linear model fit with the slope, intercept, and the coefficient of determination (R^2) provided.

Table 3. CPU-Time and different error measurements for the HT problem in a sandbox experiment.

N_{tests}	Method	$\hat{\sigma}_{\log(K)}^2$	$E(H_{HT})$	$E(H_V)$	$\log(K_G)$	CPU-Time
2	VSAFT	0.7053	0.0102	0.0552	-2.28	2.58
	GM	0.7498	0.0088	0.0521	-2.34	0.92
4	VSAFT	0.8424	0.0066	0.0720	-2.32	5.70
	GM	0.7611	0.0105	0.0509	-2.28	1.18
8	VSAFT	1.1272	0.0081	0.0464	-2.31	10.83
	GM	1.0370	0.0054	0.0417	-2.35	2.40

N_{tests} = Number of different pumping tests used, Method = Inversion method that was used, $\hat{\sigma}_K^2$ = Variance of the estimated K field, $E(H_{HT})$ = MSE of the Head values that were used in the inversion process, $E(H_V)$ = MSE of the Head values that were left apart to validation purposes, $\log(K_G)$ is the mean value of the log- K field and CPU-Time = Minutes used to estimate the K field.

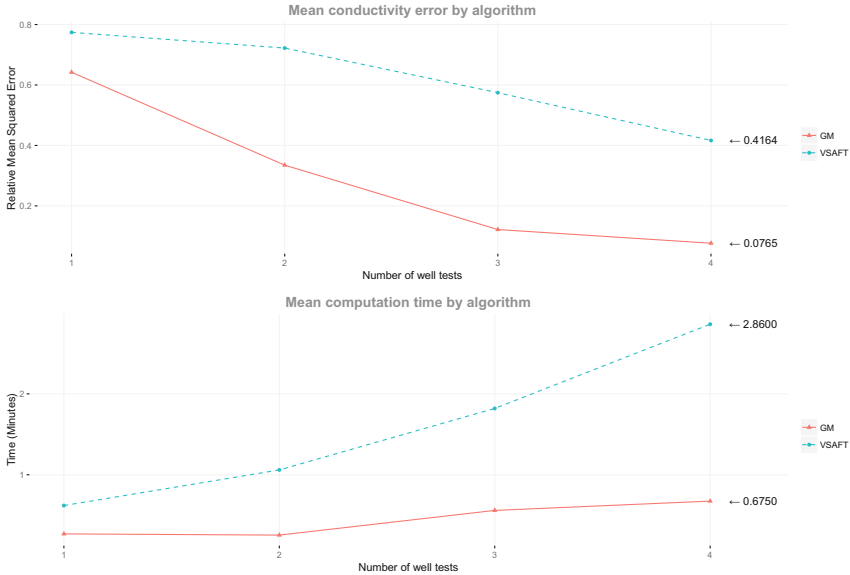


Fig. 4. Average estimation error and CPU-Time of the synthetic experiments by inversion algorithm

In addition, Fig. 5 shows the K tomograms generated with each inversion method and a photograph of the synthetic aquifer is overlain. The high K areas are indicated in red, where it is possible to see that each of these areas correspond to a different layer of sand in the sandbox. An estimated value of K of each sand type is presented in Illman *et al.* [7] (Table 1) with the four highest K sand types include (from highest to lower average K), the 16/30, 20/30, 20/40 and

X-Sample=8 | Z-Sample=6 | Total-Sample=48
 N of well tests (HT)=8 | N of well tests (Validation)=8
 VSAFT CPU-Time=10.83m | GM CPU-Time=2.40m

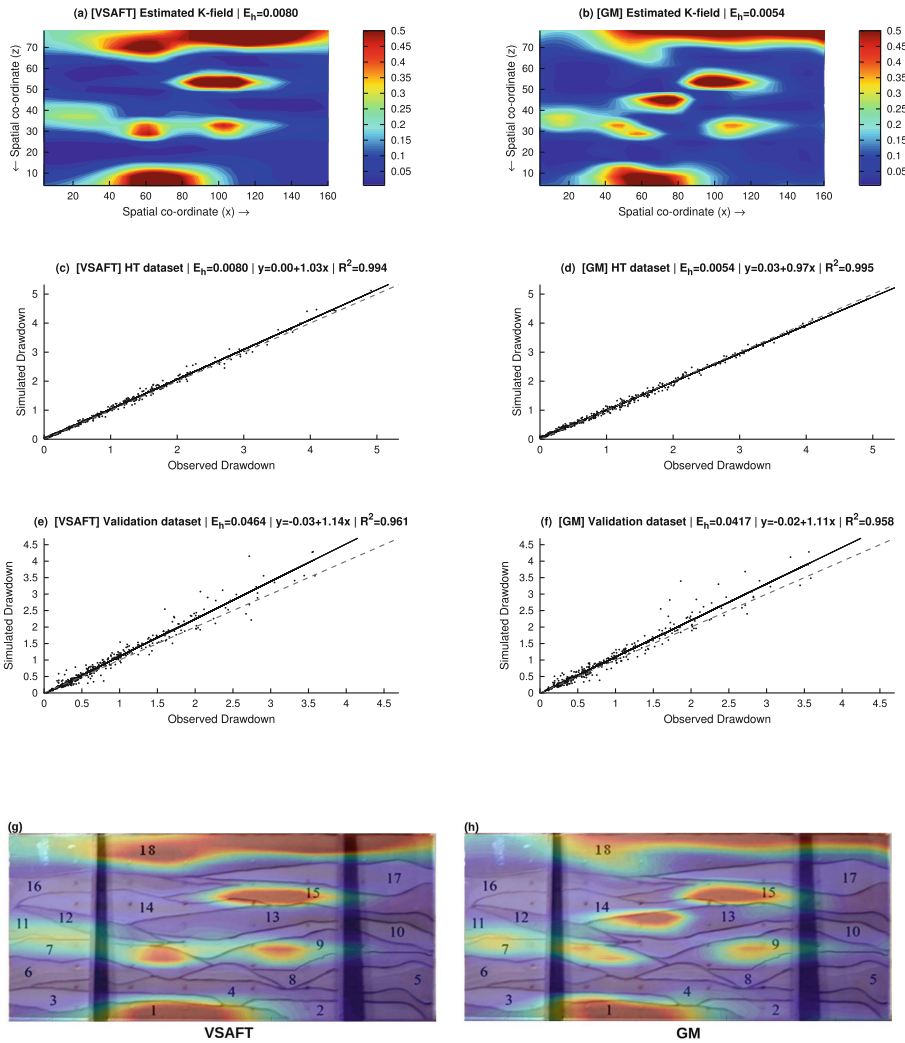


Fig. 5. K tomograms of the sandbox experiment obtained with (a) VSAFT and (b) GM as inversion methods, using 8 injection tests with 47 head values per test. Overlaid picture of the tomograms and the sandbox photograph (g)–(h). The dashed line is the 45-degree line, representing a perfect match.

#12 sands. Examination of Fig. 5 reveals that the 16/30 sand is the only type that is not shown in red in the K tomograms. This could be due to the fact this is the sand type with the second smallest volume in the sandbox.

VSAFT detected as high K areas, all layers with sand types 20/30, 20/40 and #12 with the only exception of layer 13, which is a sand 20/30. Nevertheless, this layer was detected as a high K area by the GM method. It is not until Berg and Illman [1] that layer 13 is detected as a high K area, using VSAFT transient hydraulic tomography, which requires more data and significantly more computational time.

5 Summary and Conclusions

It can be seen from the synthetic experiments that the GM algorithm achieved a higher accuracy with lower computational time than a geostatistical inversion approach. Furthermore, the K fields estimated with GM usually reconstructs the pattern of heterogeneity better, with values of $\sigma_{\log K}^2$ more similar to variance of the true K field, indicating that the GM algorithm is able to model complex K distributions with less computational effort.

Some parameters of the GMM describes the interlayer and intralayer heterogeneity and, in general, the GM inversion algorithm, depending on the value of N_k (and the availability of prior geological information), keeps a relationship between a geological and a geostatistical model. The multistart approach address the problem of multiple local/global optimal solutions, in addition these solutions are used to estimate a K field by consensus of all the solutions, the Bayes conductivity field. Furthermore, these solutions allow the estimation of an uncertainty map.

The proposed objective function uses the spatial derivatives of hydraulic head to detect more precisely, the spatial locations where changes in K lead to a better reproduction of the observed changes in head. Results suggest that the derivatives in the objective function are more sensitive to changes in the K values and reduce the correlation between parameters, which increases the accuracy of the estimates of K .

In synthetic and sandbox experiments, the effect of noisy data is reduced by changing the N_k parameter. The performance gap between the two inversion algorithms increases as the number of pumping test also increases. The GM inversion algorithm has shown a significant increase in the accuracy of the estimated K field using less computational time, compared with a geostatistical inversion approach, having up to one fifth of the error in one fourth of the computation time. Regarding the amount of data needed in the inversion process, the GM algorithm required half the number of pumping tests to achieve the same level of accuracy than SSLE/VSAFT, using one tenth of the computational time.

Due to the nature of the GM algorithm to group areas of similar K value, if the aquifer have very spread out areas of different materials, the algorithm could need a large number of components in order to capture the aquifer heterogeneity, resulting in a large number of parameters to be estimated. In addition, a mixture

model which includes different probability distributions (*e.g.* Gamma or Beta distributions), could increase the performance of the algorithm, making it able to reproduce even more complex spatial distributions of K , with a lower number parameters.

These results are promising but further testing of the GM algorithm is required. In particular, more extensive testing is necessary on non-Gaussian K fields, and under three-dimensional and transient flow conditions.

References

1. Berg, S.J., Illman, W.A.: Capturing aquifer heterogeneity: comparison of approaches through controlled sandbox experiments. *Water Resour. Res.* **47**(9) (2011). <https://doi.org/10.1029/2011WR010429>
2. Berg, S.J., Illman, W.A.: Comparison of hydraulic tomography with traditional methods at a highly heterogeneous site. *Groundwater* **53**(1), 71–89 (2015). <https://doi.org/10.1111/gwat.12159>
3. Gottlieb, J., Dietrich, P.: Identification of the permeability distribution in soil by hydraulic tomography. *Inverse Probl.* **11**(2), 353 (1995). <http://stacks.iop.org/0266-5611/11/i=2/a=005>
4. Hoeksema, R.J., Kitanidis, P.K.: An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling. *Water Resour. Res.* **20**(7), 1003–1020 (1984). <https://doi.org/10.1029/WR020i007p01003>
5. Illman, W.A., Craig, A.J., Liu, X.: Practical issues in imaging hydraulic conductivity through hydraulic tomography. *Groundwater* **46**(1), 120–132 (2008). <https://doi.org/10.1111/j.1745-6584.2007.00374.x>
6. 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.* **341**(3), 222–234 (2007). <https://doi.org/10.1016/j.jhydrol.2007.05.011>. <http://www.sciencedirect.com/science/article/pii/S0022169407002818>
7. 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.* **46**(4) (2010). <https://doi.org/10.1029/2009WR007745>
8. Minutti, C., Gomez, S., Ramos, G.: A machine-learning approach for noise reduction in parameter estimation inverse problems, applied to characterization of oil reservoirs. *J. Phys.: Conf. Ser.* **1047**, 012010 (2018). <https://doi.org/10.1088/1742-6596/1047/1/012010>
9. 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.* **45**(2) (2009). <https://doi.org/10.1029/2008WR007180>
10. Yeh, T.C.J., Jin, M., Hanna, S.: An iterative stochastic inverse method: conditional effective transmissivity and hydraulic head fields. *Water Resour. Res.* **32**(1), 85–92 (1996). <https://doi.org/10.1029/95WR02869>
11. Yeh, T.C.J., Liu, S.: Hydraulic tomography: development of a new aquifer test method. *Water Resour. Res.* **36**(8), 2095–2105 (2000). <https://doi.org/10.1029/2000WR900114>

12. 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.* **544**, 640–657 (2017). <https://doi.org/10.1016/j.jhydrol.2016.12.004>. <http://www.sciencedirect.com/science/article/pii/S002216941630796X>
13. Zhao, Z., Illman, W.A.: Three-dimensional imaging of aquifer and aquitard heterogeneity via transient hydraulic tomography at a highly heterogeneous field site. *J. Hydrol.* **559**, 392–410 (2018). <https://doi.org/10.1016/j.jhydrol.2018.02.024>. <http://www.sciencedirect.com/science/article/pii/S0022169418301008>
14. Zhao, Z., Illman, W.A., Berg, S.J.: On the importance of geological data for hydraulic tomography analysis: Laboratory sandbox study. *J. Hydrol.* **542**, 156–171 (2016). <https://doi.org/10.1016/j.jhydrol.2016.08.061>. <http://www.sciencedirect.com/science/article/pii/S0022169416305510>