4-16-2014

Data Processing for Oscillatory Pumping Tests

Tania Bakhos
Jen-Hsun Huang Engineering Center

Michael C. Cardiff
University of Wisconsin-Madison

Warren Barrash
Boise State University

Peter K. Kitanidis
Jen-Hsun Huang Engineering Center

Data Processing for Oscillatory Pumping Tests

Tania Bakhos[a], Michael C. Cardiff[b], Warren Barrash[c], Peter K. Kitanidis[d,a]

[a] *Institute for Computational and Mathematical Engineering, Jen-Hsun Huang Engineering Center, Stanford CA*
[b] *Department of Geoscience, University of Wisconsin-Madison, Madison, WI*
[c] *Center for Geophysical Investigation of the Shallow Subsurface Department of Geosciences, Boise State University, Boise, ID*
[d] *Department of Civil and Environmental Engineering, Yang and Yamazaki Environment and Energy Building, Stanford, CA*

Abstract

Characterizing the subsurface is important for many hydrogeologic projects such as site remediation and groundwater resource exploration. Methods based on the analysis of conventional pumping tests have the notable disadvantage that at a certain distance, the signal is small relative to the noise due to the effects of recharge, pumping in neighboring wells, change in the level or adjacent streams, and other common disturbances. This work focuses on oscillatory pumping tests in which fluid is extracted for half a period, then reinjected. We discuss a major advantage of oscillatory pumping tests: small amplitude signals can be recovered from noisy data measured at observation wells and quantify the uncertainties in the estimates. We demonstrate results from a joint inversion of storativity and transmissivity. We conclude with an analysis of the duration of the initial transient, providing lower bounds on the length of elapsed time until the effects of the transient can be neglected.

Keywords: oscillatory pumping tests, data processing

1. Introduction

Subsurface imaging, or determining important hydraulic parameters such as spatially-distributed hydraulic conductivities (K) and specific storage (S_s),

Email address: taniab@stanford.edu, phone number: 650-724-3313 (Tania Bakhos)

remains an important challenge in hydrology. Various pressure-based methods, i.e., methods that use changes in head or flow rate as the primary source of measurements, have been used to obtain an image of the 3-D heterogeneity of the flow parameters. Examples of such methods include partially penetrating slug tests (e.g Bouwer and Rice (1976); Butler (1998); Cardiff et al. (2011); Zlotnik and McGuire (1998)), direct push methods (e.g Dietrich and Leven (2009); Butler et al. (2002)) and borehole flow meters (e.g. Hess (1986); Paillet (1998)).

Hydraulic tomography (Hao et al., 2007; Illman et al., 2009; Yeh and Liu, 2000) is an imaging method that uses data from aquifer tests in which the pressure is changed at several distinct locations and the measurements of pressure responses at many locations in the aquifer are recorded. Inversion of the resulting data set provides an estimate of 3-D spatially heterogeneous flow parameters (Gottlieb and Dietrich, 1995). One example of such a method is transient hydraulic tomography (Zhu and Yeh, 2005; Cardiff et al., 2012; Berg and Illman, 2011; Xiang et al., 2009). A more comprehensive review of publications on research related to hydraulic tomography is offered by Cardiff and Barrash (2011).

A difficulty associated with traditional pumping and slug tests and also hydraulic tomography based on these tests is that the signal weakens with distance and, after a certain point becomes submerged in the ambient noise. The hydraulic head is sensitive to external changes, such as changes in the level of rivers adjacent to the field area, pumping or irrigation in close proximity to the observation well, tidal effects, barometric pressure, changes in overburden, etc. Noise from these sources may affect results in a variety of ways (Spane and Mackley, 2011). A disadvantage of hydraulic tomography using constant-rate pumping tests is that the signal associated with hydraulic tomography may not be easily distinguishable from these noises and trends.

Oscillatory hydraulic tomography is a subsurface imaging method that employs a tomographic analysis of oscillatory signals. In oscillatory signal tests, a periodic pressure signal can be imposed at one or more stimulation points, and the transmitted effects of this signal are recorded at monitoring wells. The idea of harmonic testing was first proposed in the petroleum literature by Kuo (1972) as an extension to pulse testing (Johnson et al., 1966; McKinley et al., 1968). More recent publications on reservoir characterization using harmonic tests include Fokker et al. (2012); Fokker and Verga (2011); Ahn and Horne (2011). Oscillatory aquifer tests have similarly been used to estimate aquifer hydraulic parameters (Engard et al., 2005; Wachter et al., 2008; Becker and Guiltinan, 2010).

Oscillatory pumping tests have several advantages over traditional pumping tests including (1) a reduction in the cost of disposing of contaminated water because there is no net extraction or injection into the aquifer, (2) a reduced computational cost through use of a steady-periodic model and (3) an ability to distinguish the signal from the background noise. Disadvantages of oscillatory pumping tests may include (1) the need for potentially different field equipment to generate a periodic stimulation and (2) the amplitude of signals at the observation locations may be much smaller than those of signals generated by constant-rate pumping.

As a modification to oscillatory pumping test analysis, multi-frequency oscillatory hydraulic imaging was proposed by Cardiff et al. (2013) in which multiple signals of different frequencies are used as a stimulation to obtain information on the aquifer heterogeneity. The authors use a "steady-periodic" model formulation to analyse the head responses to the stimulation, which allows for a reduced computational cost in numerically solving the fully-transient model. This formulation assumes that the signal has reached a steady periodic state and assumes that the initial transient effects are negligible. An analysis of when this assumption can accurately be made is an important question that, to the best of our knowledge, has not yet been addressed. Black and Kipp Jr (1981) first introduced an analytic solution for the steady-periodic response of the signal to a line-source oscillatory stimulation for a homogeneous isotropic aquifer that is effectively laterally unbounded. This approach provided an estimate of the hydraulic diffusivity using the ratio of the amplitude or phase shift from two observations wells. Rasmussen et al. (2003) derived the leaky and partially penetrating analytic solution for transmissivity and storativity in a confined aquifer. They also provide expressions for the transient solution that decays with time.

We use the analytic expressions to show that the duration of the initial transient (i.e. number of periods required for the signal to achieve a steady-periodic response) is a function of a non-dimensional quantity. The non-dimensional expression depends on the following physical parameters: the frequency of oscillations, the radial distance from the source, and the hydraulic diffusivity. We extend the analysis to more general heterogeneous aquifers and derive bounds for the time required for the signal to reach a steady-periodic response.

The existence of signal processing routines for signal extraction and denoising for oscillatory signals was briefly discussed in Cardiff et al. (2013). To denoise an oscillatory signal, methods such as the discrete Fourier transform (Renner and Messar, 2006; Hollaender et al., 2002) and ordinary least squares (Rasmussen et al., 2003; Toll and Rasmussen, 2007) are commonly and successfully used. We assume the frequency of oscillations is known and demonstrate the effectiveness of ordinary least squares in recovering the signal in the presence of common sources of noise. We quantify the uncertainties in the estimates and show that the errors in estimating the components (phase and amplitude) of a signal decay with time. Using regression for denoising and using the results of the covariance of the estimator, we present a joint inversion of storativity and transmissivity of a synthetic 2-D example.

The paper is organized as follows. In section 2 we review the governing equations. In section 3, we discuss denoising the signal under various types of noise, which is followed by a joint inversion of storativity and transmissivity in section 4. In section 5, we analyze the behavior of the initial transient and follow with concluding remarks in section 6.

2. Governing Equations

In this section, we review the governing equations. This closely follows the notation and presentation of Cardiff et al. (2013). Groundwater flow through a 2-D depth-averaged confined aquifer with horizontal confining layers for a domain Ω and boundary $\partial\Omega$ is described by the following equations,

$$
\begin{aligned}
S(x)\frac{\partial h(x,t)}{\partial t} - \nabla \cdot (T(x)\nabla h(x,t)) &= q(x,t), \qquad x \in \Omega && (1)\\
h(x,t) &= 0, \qquad x \in \partial\Omega_D && (2)\\
\nabla h(x,t) \cdot \mathrm{n} &= 0, \qquad x \in \partial\Omega_N && (3)
\end{aligned}
$$

where n is the normal vector, $x \in \mathbb{R}^2$ (L) denotes the position vector, h (L) represents the hydraulic head, $S(x)$ (-) represents the storativity and $T(x)$ (L^2/T) represents the transmissivity. Ω_D and Ω_N refer to Dirichlet (constant head) and Neumann boundary conditions (constant flux) respectively.

Using Euler's formula, we represent the oscillator as an exponential function. For the the case of one source at position x_s oscillating at a fixed frequency ω (radians/T), $q(x,t)$ is given by

$$q(x,t) = Q_0\delta(x - x_s)e^{i\omega t} \tag{4}$$

Because the solution is linear in time, the signal (after some initial time has elapsed) achieves a steady-periodic response and can be represented as,

$$h(x,t) = \Phi(x)e^{i\omega t} \tag{5}$$

where $\Phi(x)$, known as the phasor, carries information about the amplitude and phase of the signal. Plugging these definitions into (1) results in the more computationally efficient form,

$$
\begin{aligned}
i\omega S(x)\Phi(x) - \nabla \cdot (T(x)\nabla\Phi(x)) &= Q_0\delta(x - x_s), \qquad x \in \Omega && (6)\\
\Phi(x) &= 0, \qquad x \in \partial\Omega_D && (7)\\
\nabla\Phi(x) \cdot \mathrm{n} &= 0, \qquad x \in \partial\Omega_N && (8)
\end{aligned}
$$

The hydraulic head is given by (5) once Φ is known. Note that the steady-periodic formulation, i.e. equations (6)-(8), only holds if we are able to neglect the initial transient.

3. Signal Denoising

In this section, we will assume that the effects of the transient can be neglected and that the solution to the groundwater equations is a sinusoid of known frequency. Even though the solution is a sinusoid of known frequency, in practice, the measurement signals are corrupted by noise. In this section, we address how to recover the signal from a set of noisy measurements. We demonstrate the effectiveness of linear regression on four common types of noise: white noise, white noise with a jump in the signal, white noise with a linear drift and correlated noise, and quantify the errors in the estimates. This analysis hinges on the fact that the frequency is known however if the frequency is unknown, one can extract the frequency of the sinusoid by using the discrete Fourier transform and then proceed with this analysis.

Consider the measurement time series at a given point,

$$\Phi(\bar{x}, t_i) = \beta_1 \cos(\omega t_i) + \beta_2 \sin(\omega t_i) + \epsilon(t_i) \tag{9}$$

where $\epsilon(t_i)$ is the residual or error term. We assume ϵ has zero mean. If ϵ has known mean μ, it can be detrended by subtracting it from (9). If μ is not known, it will be shown that the following analysis holds true provided the time between measurements is small enough. Rewrite Φ as

$$\Phi = X\beta + \epsilon, \qquad X = \begin{pmatrix} \cos(\omega t_1) & \sin(\omega t_1) \\ \cos(\omega t_2) & \sin(\omega t_2) \\ \vdots & \vdots \\ \cos(\omega t_m) & \sin(\omega t_m) \end{pmatrix}, \qquad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \tag{10}$$

Note that if the signal was not perfectly a single sinusoid but instead a sum of several sinusoids oscillating at distinct frequencies then the columns of X would be extended to incorporate the additional frequencies. For this analysis however we limit ourselves to the case of a single sinusoid. The solution to the least-squares problem for $\hat{\beta} = [\hat{\beta}_1, \hat{\beta}_2]^T$ is given by

$$\hat{\beta} = \left(X^T X\right)^{-1} X^T \Phi \tag{11}$$

Estimating for β_1 and β_2 is equivalent to regressing on the phase and amplitude of the signal however it circumvents the problem of non-uniqueness of the phase. The covariance of the estimates is given by

$$\mathrm{Cov}(\hat{\beta}) = (X^T X)^{-1} X^T E[\epsilon\epsilon^T] X (X^T X)^{-1} \tag{12}$$

whe E[ ] denotes the expected value. Expression (12) depends on the covariance matrix of ϵ, and can be simplified under certain assumptions of the noise.

For our numerical results, all of our examples are synthetic and we consider the signal

$$\Phi(x,t) = 0.02\cos(\omega t) + 0.05\sin(\omega t) \quad (m) \tag{13}$$

with $\omega = 2\pi/40$ (1/s). Assume the data is being collected for a total of 30 periods (i.e. 20 minutes) at sampling intervals of 0.1 seconds. We present results for four distinct types of noise.

1. First we consider the case of white noise (figure 1). Suppose $\epsilon_i \sim \mathcal{N}(0,\sigma^2)$. Then, $\mathrm{E}[\epsilon\epsilon^T] = \sigma^2 I$ and expression (12) simplifies to,

$$Cov(\hat{\beta}) = \sigma^2 (X^T X)^{-1} \tag{14}$$

$$= \sigma^2 \begin{pmatrix} \sum_{i=1}^{m} \cos^2(\omega t_i) & \sum_{i=1}^{m} \cos(\omega t_i)\sin(\omega t_i) \\ \sum_{i=1}^{m} \sin(\omega t_i)\cos(\omega t_i) & \sum_{i=1}^{m} \sin^2(\omega t_i) \end{pmatrix}^{-1} \tag{15}$$

Each of the sums in (15) can be viewed as a product of $1/\Delta t$ and the left Riemann sum of their respective functions. If the interval of time between measurements Δt is small and the total sampling time, T_s, is a multiple of the period of the signal,

$$Cov(\hat{\beta}) \approx 2\sigma^2 \frac{\Delta t}{T_s} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{16}$$

The covariance decreases with an increase in $T_s/\Delta t$, the number of data measurements. The result (16) indicates that there is no posterior covariance between the two estimates, i.e the errors in estimating β_1 and β_2 are uncorrelated. We can thus write the error in the estimates as,

$$|\hat{\beta} - \beta| \approx \frac{2\Delta t}{T_s} \begin{pmatrix} \sum \epsilon_i \cos(\omega t_i) \\ \sum \epsilon_i \sin(\omega t_i) \end{pmatrix} \tag{17}$$

In the case where ϵ has a nonzero mean μ, the estimates will not be affected provided the data is being collected for a multiple of the period. This is because the solution is given by,

$$\hat{\beta} = \left(X^T X\right)^{-1} X^T (\Phi - \mu) \tag{18}$$

and if the data is being collected for a multiple of the period, $X^T \mu = 0$.

2. Consider the case where in addition to white noise, there is an abrupt shift in the hydraulic head at some time in the time series. If the shift occurs for exactly a multiple of the period (figure 2), it will not affect the least squares estimates because of its orthogonality with X^T. The worst case would be when it happens for an additional half period (figure 3). While the error due to the non-orthogonal components will remain present, the overall error can be reduced by taking a longer measurement collecting interval.
3. Consider the case where there is a linear drift in addition to white noise such that the measured signal is

$$\Phi(x,t) = 0.02\cos(\omega t) + 0.05\sin(\omega t) + \epsilon \quad (m) \tag{19}$$

where $\epsilon_i = \alpha t_i + n_i$, $n_i \sim \mathcal{N}(0, \sigma^2)$ and α (m/s) is the drift coefficient. We consider two cases: (1) where the presence of the drift is unknown and too small to be visible in the raw data, and (2) when the presence of a linear drift is known or visible. In the former case (see figure 4) and by keeping the same regressors, the errors in the estimate of β are given by,

$$\hat{\beta} - \beta = (X^T X)^{-1} X^T (\alpha t + \epsilon) \tag{20}$$

NOTICE: this is the author's version of a work that was accepted for publication in Journal of Hydrology. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Journal of Hydrology, (2014)] DOI: 10.1016/j.jhydrol.2014.01.007

If the sampling time Δt is small enough and that data is being collected for a multiple of the period, then

$$\left|\hat{\beta} - \beta\right| \approx \left|\frac{2\alpha}{\omega}\begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{2\Delta t}{T_s} X^T n\right| \tag{21}$$

Note that the second term is precisely the error that results from having pure white noise. The additional errors incurred by the presence of a linear drift do not affect the estimates $\hat{\beta}_1$. The estimates for $\hat{\beta}_2$ depend on both α and ω and do not decrease with the sampling time, however, if there is a constant linear drift present, a longer sampling time will increase the likelihood of the detection of the drift by looking at the measured signal. If the presence of the drift is known or can be detected by looking at the measured signal, the regressors can be modified and the estimates improved.

$$\Phi = X\beta + n, \qquad X = \begin{pmatrix} \cos(\omega t_1) & \sin(\omega t_1) & t_1 \\ \cos(\omega t_2) & \sin(\omega t_2) & t_2 \\ \vdots & \vdots & \vdots \\ \cos(\omega t_m) & \sin(\omega t_m) & t_m \end{pmatrix}, \qquad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \alpha \end{pmatrix} \tag{22}$$

By regressing for the drift coefficient, this allows for more accurate results (see figure 5). In particular, the error of using the new regressors results in an error,

$$|\hat{\beta} - \beta| \approx \frac{2\Delta t}{T_s}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 + \frac{12}{-12+2T_s^2\omega^2} & \frac{6\omega}{-12+T_s^2\omega^2} \\ 0 & \frac{6\omega}{-12+T_s^2\omega^2} & \frac{3\omega^2}{-12+2T_s^2\omega^2} \end{pmatrix}\begin{pmatrix} \sum n_i \cos(\omega t_i) \\ \sum n_i \sin(\omega t_i) \\ \sum n_i t_i \end{pmatrix} \tag{23}$$

Note that the additional errors incurred by assuming drift behave as $\Delta t/T_s^2$ and thus their effects are negligible if the sampling time is long enough.

4. Consider the presence of a stationary $AR(1)$, or first-order autoregressive, noise (figure 6). Such a process has the property that the output depends

on the value at the previous time. It can be written as

$$\Phi(\bar{x}, t_i) = \beta_1 \cos(\omega t_i) + \beta_2 \sin(\omega t_i) + \epsilon_i \tag{24}$$

where $\epsilon_i = c\epsilon_{i-1} + n_i$ and $n_i \sim \mathcal{N}(0, \sigma^2)$, $|c| < 1$.

In all four of the cases discussed, we have shown that using linear regression allows us to recover the signal from a set of noisy measurements.

4. Inversion by Geostatistical Approach

4.1. Geostatistical Approach

The next section will briefly describe the geostatistical method for inversion and demonstrate examples from synthetic cases of single frequency oscillatory hydraulic imaging, with the signal denoising done by least squares as described in the previous section. The geostatistical method for inversion is one of the prevalent methods to solve stochastic inverse problems (Kitanidis, 1995, 2010, 2007). We closely follow the algorithm discussed in (Li et al., 2005) for joint inversion. The idea of the geostatistical method for inversion is to represent the unknown field as the sum of a deterministic term and a stochastic term that models small-scale variability. Inference of the parameters is made through the posterior probability distribution function by using information from the prior combined with the likelihood of the measurements. The measurement equation can be written as,

$$y = h(s) + v, \qquad v \sim \mathcal{N}(0, R) \tag{25}$$

where y represents the noisy measurements and v is a random vector corresponding to observation error with mean zero and covariance matrix R. Let $s = [s_k^T, s_s^T]^T$ be the function to be estimated where s_k and s_s correspond to the log transmissivity and log storativity fields respectively.

$$s_k \sim \mathcal{N}(X_k\beta_k, Q_k), \qquad s_s \sim \mathcal{N}(X_s\beta_s, Q_s) \tag{26}$$

where, X_k and X_s are matrices of known base functions and β_s and β_k are a set of drift coefficients to be determined. The log-transformation was used to ensure that the forward problem is well-posed since the fields need to be positive. Denote the full quantities,

$$X = \begin{pmatrix} X_k & 0 \\ 0 & X_s \end{pmatrix}, \qquad \beta = \begin{pmatrix} \beta_k \\ \beta_s \end{pmatrix}, \qquad Q = \begin{pmatrix} Q_k & 0 \\ 0 & Q_s \end{pmatrix} \tag{27}$$

The expression for Q requires the assumption that log transmissivity and log storativity are uncorrelated. More detail on how to choose the modeling parameters Q and X can be found in Kitanidis (1995). To choose R, we use the covariance of the least-squares estimates as a lower bound. Following the geostatistical method for quasi-linear inversion, we compute $\hat{s}$ and $\hat{\beta}$ corresponding to the maximum-a-posteriori probability. To solve the optimization problem, the Gauss-Newton algorithm is used. Starting with an initial estimate for the field s_0, the procedure is described in algorithm 1.

Algorithm 1 Quasi-linear Geostatistical Approach

1: Compute the $N_y \times N_s$ Jacobian J as,

$$J_i = \left.\frac{\partial h}{\partial s}\right|_{s=s_i} \tag{28}$$

2: Solve the system of equations,

$$\begin{pmatrix} J_i Q J_i^T + R & J_i X \\ (J_i X)^T & 0 \end{pmatrix} \begin{pmatrix} \xi_{i+1} \\ \beta_{i+1} \end{pmatrix} = \begin{pmatrix} y - h(s_i) + J_i s_i \\ 0 \end{pmatrix} \tag{29}$$

3: Update s_{i+1} by,

$$s_{i+1} = X\beta_{i+1} + Q J_i^T \xi_{i+1} \tag{30}$$

4: Add a line search if necessary. Repeat steps $1-3$ until the desired tolerance has been reached.

To construct the Jacobian, since the number of unknown parameters is generally larger than the number of measurements, the adjoint state method is used where by each row of the Jacobian is calculated by one adjoint 'run'. For a detailed derivation of the adjoint equations for oscillatory pumping tests refer to Cardiff et al. (2013). Note that if either log transmissivity or log storativity is known, it is treated as a normal random variable with known mean $X\beta$ and zero covariance and algorithm 1 remains unchanged. More details of the inversion can be found in Cardiff et al. (2013); Saibaba et al. (in press).

4.2. Numerical Results

Using the geostatistical method as discussed, we present inversion results for a synthetic example. Assuming a 2-D isotropic depth-averaged confined aquifer and given a set of discrete measurements of the hydraulic head our objective is to determine the random log conductivity field. We use FEniCS to discretize the governing equations using standard linear finite elements (Logg et al., 2012a,b; Logg and Wells, 2010) and use the Python interface. The modeling parameters are chosen to be $R = \tilde{\sigma}^2 I$, $X_s = X_k = [1, ..., 1]^T$. We choose the covariance matrices Q_k and Q_s to have entries $Q_k(i,j) = Q_s(i,j) = \kappa(x_i, y_j)$ corresponding to the exponential kernel,

$$\kappa(x, y) = \exp\left(-\|x - y\|_2/(L/5)\right) \tag{31}$$

such that the correlation length is $L/5 = 20$ [m]. where L is the length of the domain. (To reduce the computational and memory cost associated with forming these large covariance matrices, they are not formed explicitly and the fast Fourier transform (FFT) is used to accelerate the matrix-vector products.) The measurements were synthetically generated by adding noise $\nu \sim \mathcal{N}(0, \sigma^2)$, $\sigma = 0.01$ (m). The choice of $\tilde{\sigma}$ in the modeling parameter R was chosen based off of the covariance of the least squares estimator. The pumping volume was 1.4 (L/half cycle) and the pumping frequency was chosen to be $\omega = 2\pi/60$ (1/s). The pumping source is located at the center of the aquifer. We assume the signal has reached steady-periodic state and that data has been collected every 0.1 seconds for half an hour. The configuration for data aquisition is shown in figure 7, with the source in the center surrounded by 16 measurement locations. The system is discretized with 10201 points corresponding to a physical system of 100m x 100m with the area of interest being the 20m x 20m area centered at the origin. The boundary conditions are assumed to be Dirichlet and their effects

Definition	Parameters	Values
Aquifer length (m)	L	100
Mean storativity (-)	$\log_{10} S$	-4
Variance of storativity (first example)	$\sigma^2(\log_{10} S)$	0
Variance of storativity (second example)	$\sigma^2(\log_{10} S)$	0.11
Mean transmissivity (m^2/s)	$\mu(\log_{10} T)$	-5
Variance of transmissivity	$\sigma^2(\log_{10} T)$	0.12
Frequency (1/s)	ω	$\frac{2\pi}{60}$
Pumping volume (L/ half cycle)	Q	1.4

Table 1: Parameters Chosen For Test Problem

minimized by choosing the boundaries at a far enough distance from the source. At each measurement location we denoise the signal to get the two components of $\hat{\beta}$ which are then recorded. These components effectively correspond to the sine and cosine components of the signal and are both used in the inversion. The results are presented for known constant storativity ($S = 10^{-5}[-]$) (figure 8) and for the joint inversion case where both storativity and transmissivity are not known (figures 9 and 10). All true fields were considered to be Gaussian random fields generated using an exponential covariance kernel $\kappa(x, \mathbf{y}) = \exp(-\|x - y\|_2/(L/5))$ using the algorithm described in Dietrich and Newsam (1993). The parameters used in the generation of the numerical example are summarized in Table 1.

5. Analysis of the initial transient

5.1. Homogeneous Aquifers

We have so far considered the groundwater equations after the effects of the initial transient have subsided and can be neglected. In this section, we analyze the duration of this initial transient. Under the assumption of a homogeneous isotropic confined aquifer where the lateral extent of the aquifer is "infinite" compared to the measurement locations, the problem simplifies to the case of a penetrating line source of periodic flow for which the transient solution is known. An analytic solution to the steady periodic solution to this problem was introduced in (Black and Kipp Jr, 1981). A similar set of analytic solutions, including an expression for the initial transient, was derived in (Rasmussen et al., 2003) is

$$h(r,t) = \frac{Q_0 e^{i\omega t}}{2\pi T}\left(K_0\left(r\sqrt{\frac{i\omega}{D}}\right) - \int_0^\infty \frac{\lambda J_0(r\lambda)}{\frac{i\omega}{D}+\lambda^2} e^{-(i\omega + D\lambda^2)t} d\lambda\right) \tag{32}$$

where r (L) is the radial distance from the pumping source, $D = T/S$ (L^2/T) is the diffusivity and J_0 and K_0 are the zeroth-order modified Bessel functions of the first and second kind respectively. The first term corresponds to the steady periodic solution and the second term corresponds to the initial transient that decays with time. Equation (32) indicates that the duration of the transient depends on the parameters ω, D and r. Denote $\mathrm{T}_{5\%}$ and $\mathrm{NP}_{5\%}$ as the length of time and the number of periods respectively that is required for the magnitude of the transient solution to fall within 5% of the amplitude of the corresponding steady state solution. (The subscripts 1% and 10% correspond accordingly to the 1% and 10% marks - see figure 11).

To simulate realistic field conditions, we use an oscillating pumping stimulation that contains a period of "ramp-up".

$$q(x,t) = Q_0 \cos(\omega t)\left(1 - \exp(-(t/T)^2)\right)\delta(x - x_s) \tag{33}$$

where T, the time scale parameter is chosen to be the period of the oscillations. We use the adaptive Gauss-Konrad quadrature to numerically integrate the solution for a source term of the form (33) (Shampine, 2008). The duration of the initial transient increases as r and ω increase and decreases as D increases (figure 7 - top, middle). A natural non-dimensional scaling that combines the parameters of interest is

$$\gamma = \frac{\omega}{D} r^2 \tag{34}$$

The hypothesis that $\mathrm{NP}_{5\%}$ admits a scaling of this form is tested and we observe that the data collapses into a single curve (figure 7 - bottom). In other words, the number of periods is a self-similar solution with γ being the similarity variable. Figure 7 shows the behaviour of the initial transient for a specific range of γ and this range was chosen to be representative of the range of "measurable" signals, as demonstrated by figure 13 but it is not exhaustive. If the estimates of the aquifer parameters (storativity and transmissivity) are available, the curve in figure 7 provides a lower bound on the time needed to wait, depending on the desired error tolerance. If the values of interest do not fall within this range, these curves can be generated again as necessary.

5.2. Heterogeneous Aquifers

Aquifers are, in general, not homogeneous and an analytic solution of the form (32) is not available. One approach is to use the analysis described for homoegeneous aquifers using effective parameters for storativity and transmissivity, if avalaible. Another approach for dealing with heterogeneous aquifers is to calculate a bound for which the time falls within some tolerance *tol* based on the eigenvalues of the discretization matrices. This can only be done if estimates for the fields are available. We semi-discretize the PDE (1),

$$Kh + M\frac{\partial h}{\partial t} = be^{i\omega t} \tag{35}$$

h and b are vectors corresponding to the spatial discretizations of the hydraulic head and the source term respectively. The time at which the solution falls within a given tolerance *tol* of the steady periodic solution (see the appendix for a derivation) is given by,

$$T = \frac{1}{\lambda_{min}} \log\left(\frac{\|\tilde{b}\|_2}{tol * (\sqrt{\lambda_{min}^2 + \omega^2})}\right) \tag{36}$$

where $\lambda_m in$ is the minimum eigenvalue of $M^{-1/2}KM^{-1/2}$ and $\tilde{b} = M^{-1/2}b$. Note that knowledge of λ_{min} requires estimates for the conductivity and storativity field to be known apriori. Also note that since this bound holds on the entire domain and we are only concerned about the behavior of the signal at specific locations, i.e. the measurement locations, it will be a loose upper bound. It will be a large overestimate of the time one has to wait, particularly if the domain is much larger than the measurement location area. While the method discussed has its limitations, it nonetheless provides a first analysis to estimate how long the effects of the initial transient persist.

6. Conclusions and Discussion

We have presented approaches to estimate the time needed until the signal reaches a steady-periodic response. When the noise level is low, the time at which the transient becomes insignificant is clear from the measurements. However, if the signal is submerged in the noise, it is difficult to distinguish the transient from the steady state. For the homogeneous case, we have shown how the number of periods scales with a non-dimensional scalar that depends on diffusivity, radius from the oscillating source, and frequency of oscillations. This analysis will be beneficial for those conducting field experiments as the analysis provided offers a lower bound for the duration during which the initial transient effects cannot be neglected. For heterogeneous aquifers and if estimates of the storativity and transmissivity fields are known, we suggested an alternate method however both methods discussed have their limitations and this question needs to be further investigated. One extension would be to consider a reduced order model for the groundwater equations.

A major benefit in oscillatory pumping tests is the ability to extract the signal from a variety of different types of noise, even when the signal is small compared to the level of noise provided the duration of the test is long enough. While we have focused our analysis on four different types of noise, the sinusoidal nature of the signal allows us to extract low magnitude signals from a wider variety of disturbances provided the time is long enough. In practice, there might be noise that has periodic components, such as daily tidal signals, however these can be identified prior to the actual test to ensure that the pumping frequency is unique in the sense that interference with such signals is minimized. We demonstrated the effectiveness of regression and concluded by presenting results for a joint inversion of storativity and transmissivity.

While we have only shown results of single frequency signals, multiple frequency signals can just as easily be denoised and the additional information obtained from the additional frequencies improves the resulting image reconstruction, as demonstrated by Cardiff et al. (2013). Instead of each test corresponding to a single-frequency oscillatory, pumping at multiple frequencies simultaneously would reduce the total time required to conduct a field test. This holds exciting prospects for oscillatory hydraulic tomography. In future studies, we will investigate which frequency, or range of frequencies, yields the best inversion results. There have been recent developments in efficient methods of solving the inverse problem using the geostatistical approach for oscillatory hydraulic imaging based on a Krylov subspace method for shifted systems (Saibaba et al., in press).

Our analysis was limited to the most basic two-parameter model. In many cases, a dual porosity model may be more appropriate. Additional questions of practical importance that we will investigate in future studies are the effects of leakage, boundaries and how the results from oscillatory hydraulic tomography compare with those resulting from transient and steady-state hydraulic tomography. It may be that combining these tests would provide more detail than a single test alone.

7. Acknowledgements

The research in this work was funded by NSF Award 1215742, "Collaborative Research: Fundamental Research on Oscillatory Flow in Hydrogeology." The authors would also like to thank Arvind Saibaba for useful discussions, and Todd Rasmussen and the anonymous reviewers for their insightful comments that helped improve this manuscript.

References

Ahn, S., Horne, R., 2011. The use of attenuation and phase shift to estimate permeability distributions from pulse tests, in: SPE Annual Technical Conference and Exhibition.

Becker, M., Guiltinan, E., 2010. Cross-hole periodic hydraulic testing of inter-well connectivity. Stanford Geothermal Workshop, vol. SGP-TR-188 .

Berg, S., Illman, W., 2011. Three-dimensional transient hydraulic tomography in a highly heterogeneous glaciofluvial aquifer-aquitard system. Water Resources Research 47, W10507.

Black, J., Kipp Jr, K., 1981. Determination of hydrogeological parameters using sinusoidal pressure tests: a theoretical appraisal. Water Resources Research 17, 686–692.

Bouwer, H., Rice, R., 1976. A slug test for determining hydraulic conductivity of unconfined aquifers with completely or partially penetrating wells. Water Resources Research 12, 423–428.

Butler, J.J., 1998. The design, performance and analysis of slug tests. CRC Press.

Butler, J.J., Healey, J.M., McCall, G., Garnett, E.J., Loheide, S.P., 2002. Hydraulic tests with direct-push equipment. Ground Water 40, 25–36.

Cardiff, M., Bakhos, T., Kitanidis, P.K., Barrash, W., 2013. Aquifer heterogeneity characterization with oscillatory pumping: Sensitivity analysis and imaging potential. Water Resources Research , n/a–n/aURL: `http://dx.doi.org/10.1002/wrcr.20356`, doi:10.1002/wrcr.20356.

Cardiff, M., Barrash, W., 2011. 3-d transient hydraulic tomography in unconfined aquifers with fast drainage response. Water Resources Research 47, W12518.

Cardiff, M., Barrash, W., Kitanidis, P., 2012. A field proof-of-concept of aquifer imaging using 3-d transient hydraulic tomography with modular, temporarily-emplaced equipment. Water Resources Research 48, W05531.

Cardiff, M., Barrash, W., Thoma, M., Malama, B., 2011. Information content of slug tests for estimating hydraulic properties in realistic, high-conductivity aquifer scenarios. Journal of Hydrology 403, 66–82.

Dietrich, C., Newsam, G., 1993. A fast and exact method for multidimensional gaussian stochastic simulations. Water Resources Research 29, 2861–2869.

Dietrich, P., Leven, C., 2009. Direct push-technologies, in: Groundwater Geophysics. Springer, pp. 347–366.

Engard, B., McElwee, C., Healey, J., Devlin, J., 2005. Hydraulic tomography and high-resolution slug testing to determine hydraulic conductivity distributions - year 1. Project Report to the Strategic Environmental Research and Development Program, U.S. DoD, EPA, and DOE, 81 pp., KGS Open-File Report no. 2005-36. .

Fokker, P., Renner, J., Verga, F., 2012. Applications of harmonic pulse testing to field cases, in: SPE Europec/EAGE Annual Conference.

Fokker, P., Verga, F., 2011. Application of harmonic pulse testing to water–oil displacement. Journal of Petroleum Science and Engineering 79, 125–134.

Gottlieb, J., Dietrich, P., 1995. Identification of the permeability distribution in soil by hydraulic tomography. Inverse Problems 11, 353.

Hao, Y., Yeh, T., Xiang, J., Illman, W., Ando, K., Hsu, K., Lee, C., 2007. Hydraulic tomography for detecting fracture zone connectivity. Ground Water 46, 183–192.

Hess, A.E., 1986. Identifying hydraulically conductive fractures with a slow-velocity borehole flowmeter. Canadian Geotechnical Journal 23, 69–78.

Hochbruck, M., Ostermann, A., 2010. Exponential integrators. Acta Numerica 19, 209–286.

Hollaender, F., Hammond, P., Gringarten, A., 2002. Harmonic testing for continuous well and reservoir monitoring. paper SPE 77692.

Illman, W., Liu, X., Takeuchi, S., Yeh, T., Ando, K., Saegusa, H., et al., 2009. Hydraulic tomography in fractured granite: Mizunami underground research site, japan. Water Resour. Res 45, W01406.

Johnson, C., Greenkorn, R., Woods, E., 1966. Pulse-testing: a new method for describing reservoir flow properties between wells. Journal of Petroleum Technology 18, 1599–1604.

Kitanidis, P.K., 1995. Quasilinear geostatistical theory for inversing. Water Resour. Res. 31, 2411–2419.

Kitanidis, P.K., 2007. On stochastic inverse modeling. AGU, Washington, D. C.. volume Geophysical Monograph 171. pp. 19–30.

Kitanidis, P.K., 2010. Bayesian and Geostatistical Approaches to Inverse Problems. John Wiley & Sons, Ltd. pp. 71–85. URL: http://dx.doi.org/10.1002/9780470685853.ch4, doi:10.1002/9780470685853.ch4.

Kuo, C., 1972. Determination of reservoir properties from sinusoidal and multirate flow tests in one or more wells. Old SPE Journal 12, 499–507.

Li, W., Nowak, W., Cirpka, O., 2005. Geostatistical inverse modeling of transient pumping tests using temporal moments of drawdown. Water resources research 41, W08403.

Logg, A., Mardal, K.A., Wells, G.N., et al., 2012a. Automated Solution of Differential Equations by the Finite Element Method. Springer. doi:10.1007/978-3-642-23099-8.

Logg, A., Wells, G.N., 2010. Dolfin: Automated finite element computing. ACM Transactions on Mathematical Software 37. doi:10.1145/1731022.1731030.

Logg, A., Wells, G.N., Hake, J., 2012b. DOLFIN: a C++/Python Finite Element Library. Springer. chapter 10.

McKinley, R., Vela, S., Carlton, L., 1968. A field application of pulse-testing for detailed reservoir description. Journal of Petroleum Technology 20, 313–321.

Paillet, F., 1998. Flow modeling and permeability estimation using borehole flow logs in heterogeneous fractured formations. Water Resources Research 34, 997–1010.

Rasmussen, T., Haborak, K., Young, M., 2003. Estimating aquifer hydraulic properties using sinusoidal pumping at the savannah river site, south carolina, usa. Hydrogeology Journal 11, 466–482.

Renner, J., Messar, M., 2006. Periodic pumping tests. Geophysical Journal International 167, 479–493.

Saibaba, A., Bakhos, T., Kitanidis, P., in press. A flexible krylov solver for shifted systems with application to oscillatory hydraulic tomography. SIAM Journal on Scientific Computing .

Shampine, L.F., 2008. Vectorized adaptive quadrature in matlab. J. Comput. Appl. Math. 211, 131–140. URL: `http://dx.doi.org/10.1016/j.cam.2006.11.021`, doi:10.1016/j.cam.2006.11.021.

Spane, F.A., Mackley, R.D., 2011. Removal of river-stage fluctuations from well response using multiple regression. Ground Water 49, 794–807.

Toll, N.J., Rasmussen, T.C., 2007. Removal of barometric pressure effects and earth tides from observed water levels. Ground water 45, 101–105.

Wachter, B., McElwee, C., Devlin, J., 2008. Hydraulic tomography and high-resolution slug testing to determine hydraulic conductivity distributions - year 4. Project Report to the Strategic Environmental Research and Development Program, U.S. DoD, EPA, and DOE, 74 pp., KGS Open-File Report no. 2008-23 .

Xiang, J., Yeh, T.C.J., Lee, C.H., Hsu, K.C., Wen, J.C., 2009. A simultaneous successive linear estimator and a guide for hydraulic tomography analysis. Water Resources Research 45.

Yeh, T.C.J., Liu, S., 2000. Hydraulic tomography: Development of a new aquifer test method. Water Resources Research 36, 2095–2105.

Zhu, J., Yeh, T., 2005. Characterization of aquifer heterogeneity using transient hydraulic tomography. Water Resources Research 41, W07028.

Zlotnik, V.A., McGuire, V.L., 1998. Multi-level slug tests in highly permeable formations: 1. modification of the springer-gelhar (sg) model. Journal of Hydrology 204, 271–282.

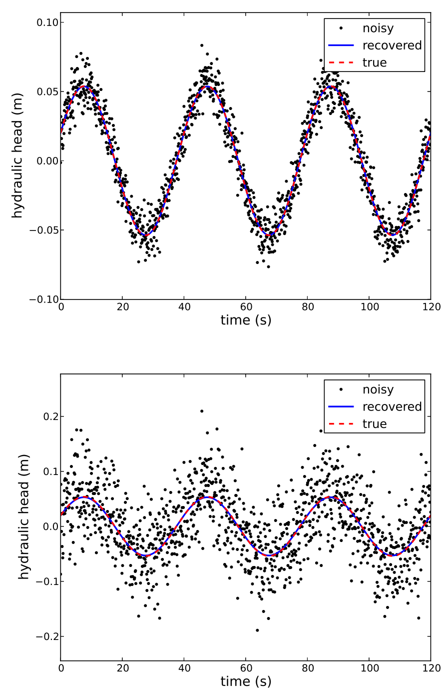

Figure 1: Hydraulic head at three periods, in the case of white noise. $\epsilon \sim \mathcal{N}(0, \sigma^2)$ with $\sigma = 1$ (cm) (top) and $\sigma = 5$ (cm) (bottom). The L^2 norm of the relative errors are respectively 0.36% and 1.7%. The root mean square errors in the estimates are respectively 0.01 (cm) and 0.06 (cm). The data is synthetic, with the true signal being that shown in (13).

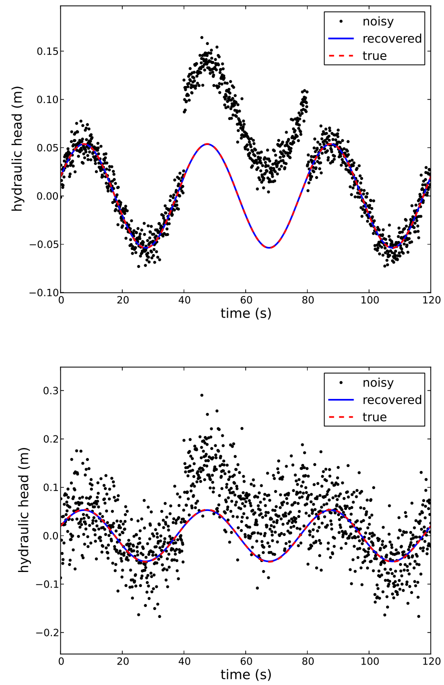

Figure 2: Hydraulic head at three periods, in the case of white noise with an abrupt shift of one period (The jump is exaggerated for illustration purposes). $\epsilon \sim \mathcal{N}(0, \sigma^2)$, $\sigma = 1$ (cm) (top) and $\sigma = 5$ (cm) (bottom). The L^2 norm of the relative errors are respectively 0.36% and 1.7%. The root mean square errors in the estimates are respectively 0.01 (cm) and 0.06 (cm). Note these errors are identical to the pure white noise case, because the disturbance occured for exactly a multiple of the period. The data is synthetic, with the true signal being that shown in (13).

Appendix A. Derivations

We derive bounds for which the solution of the groundwater equations is effectively steady-periodic. After semi-discretizing the partial differential equation (1),

$$Kh + M\frac{\partial h}{\partial t} = be^{i\omega t} \tag{A.1}$$

where K and M are the stiffness and mass matrix respectively, and b and h are now vectors corresponding to the discretization of the amplitude of the pumping

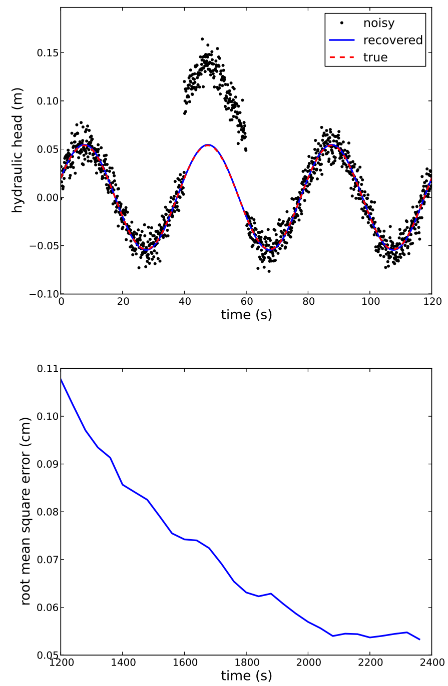

Figure 3: Hydraulic head at three periods, in the case of white noise with an abrupt shift of half a period (The jump is exaggerated for illustration purposes). $\epsilon \sim \mathcal{N}(0, \sigma^2)$, $\sigma = 1$ (cm) (top). The L^2 norm of the relative error is 2.8% and the root mean square error of the estimates is 0.1 (cm). and (bottom) plot of the root mean square error with time. The data is synthetic, with the true signal being that shown in (13).

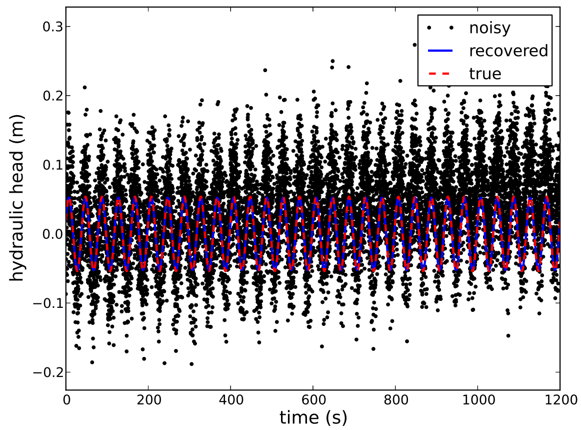

Figure 4: Hydraulic head at the entire sampling duration, in the case of white noise with a linear drift. $n \sim \mathcal{N}(0, \sigma^2)$ and a linear drift $\alpha = 0.005$ (cm/s). $\sigma = 5$ (cm) (top). The L^2 norm of the relative error is 2.9% and the root mean square error of the estimates is 0.1 (cm). and (bottom) plot of the root mean square error with time. The data is synthetic, with the true signal being that shown in (13).

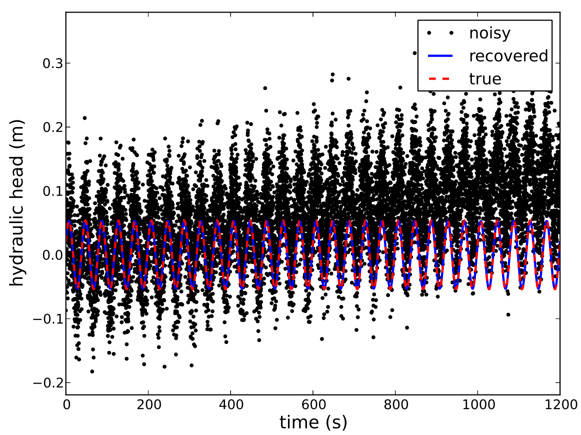

Figure 5: Hydraulic head at the entire sampling duration, in the case of white noise with a linear drift. $n \sim \mathcal{N}(0, \sigma^2)$ and a linear drift $\alpha = 0.01$ (cm/s). $\sigma = 5$ (cm) (left). The L^2 norm of the relative error is 1.7% and the root mean square error of the estimates is 0.06 (cm). and (right) plot of the root mean square error with time. The data is synthetic, with the true signal being that shown in (13).

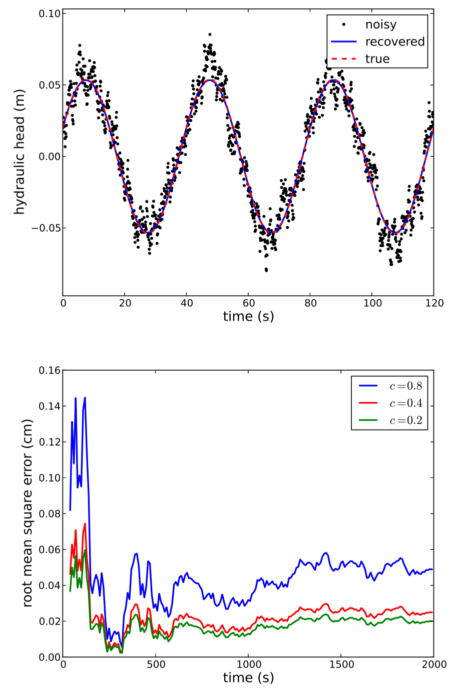

Figure 6: Hydraulic head at three periods, in the case of AR(1) noise . AR(1): $\epsilon_i = c\epsilon_{i-1} + n_i$, where $n_i \sim \mathcal{N}(0, \sigma^2(1-c^2))$, $\sigma = 1$ (cm), $c = 0.8$ (top). The L^2 norms of the relative error is 1.1% and the root mean square error is 0.04 (cm). and (bottom) plot of the root mean square error with time for various correlation coefficients. The data is synthetic, with the true signal being that shown in (13).

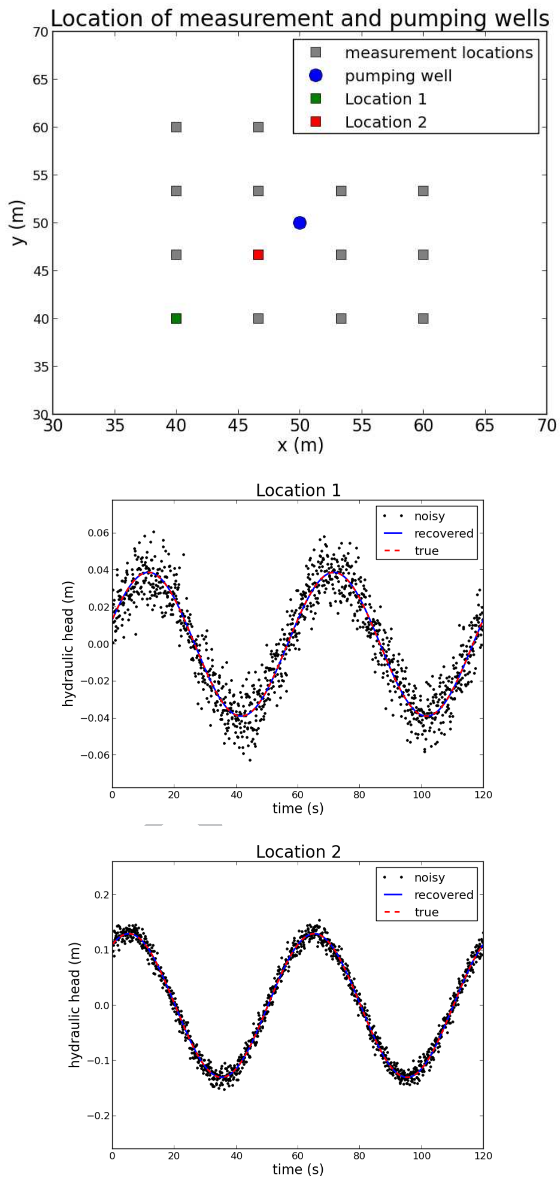

Figure 7: (top) The location of the pumping source and the measurement wells and (middle, bottom) the synthetic generated signal used for the inverse problem, noisy and denoised, at two locations.

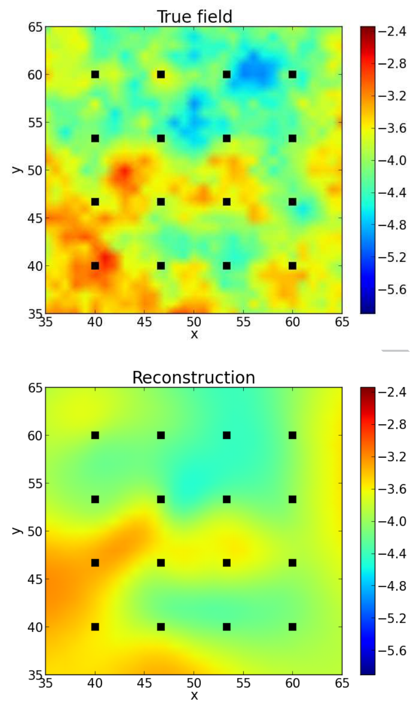

Figure 8: The true log transmissivity field (top) and the reconstructed log transmissivity field (bottom) the relative L^2 error within the area of measurements is 0.13 - for the inversion for transmissivity only. The plots are zoomed in so the area of measurements is more clearly visible.

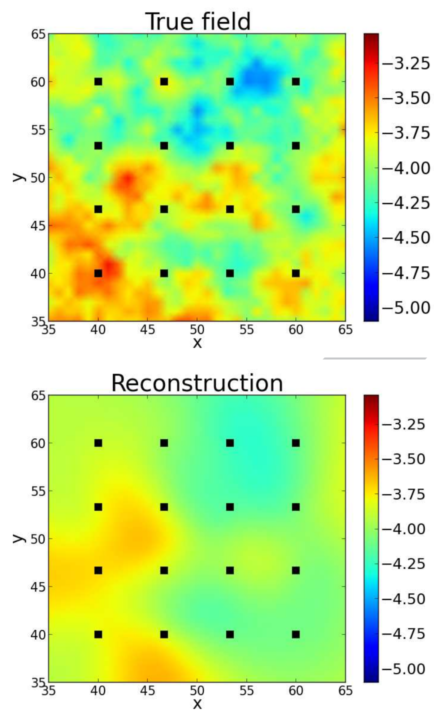

Figure 9: The true log transmissivity field (top) and the reconstructed log transmissivity field (bottom). The relative L^2 error within the area of measurements is 0.18 - for the joint inversion. The plots are zoomed in so the area of measurements is more clearly visible.

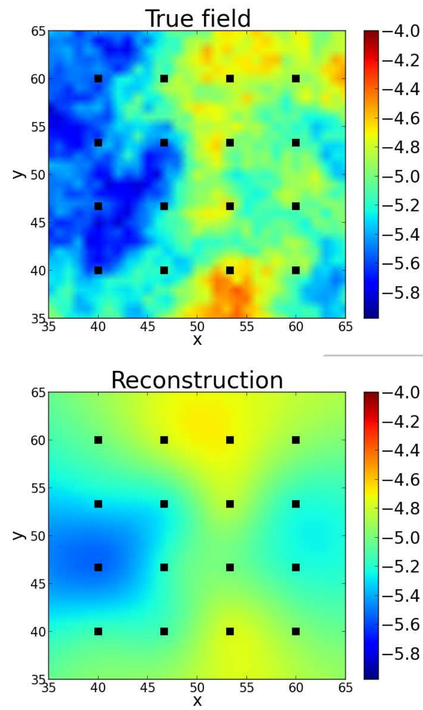

Figure 10: The true log storativity field (top) and the reconstructed log storativity field (bottom). The relative L^2 error within the area of measurements is 0.59 - for the joint inversion. The plots are zoomed in so the area of measurements is more clearly visible.

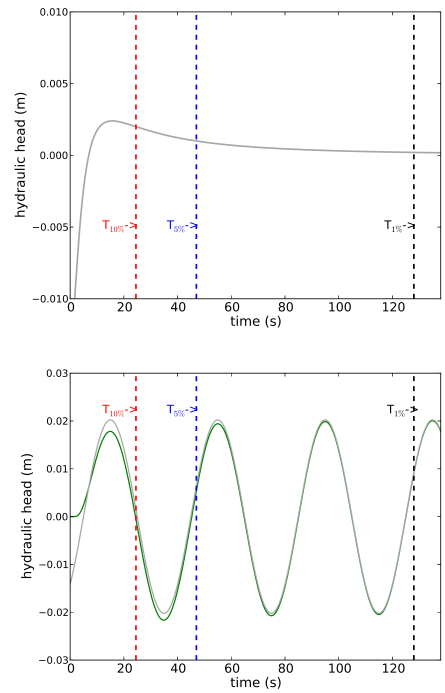

Figure 11: (top) The transient solution with marked lines denoting the time at which the magnitude of the transient drops to 1 (black), 5 (blue) and 10 percent of the amplitude of the signal. (bottom) Comparison of the signal (transient plus steady-state) and steady-state only. The parameters used in this example are $Y = 10^{-4}$ (m^2/s), $S = 10^{-5}$ (-) , $\omega = 2\pi/40$ (1/s) and $Q = 1.6$ (L/half cycle).

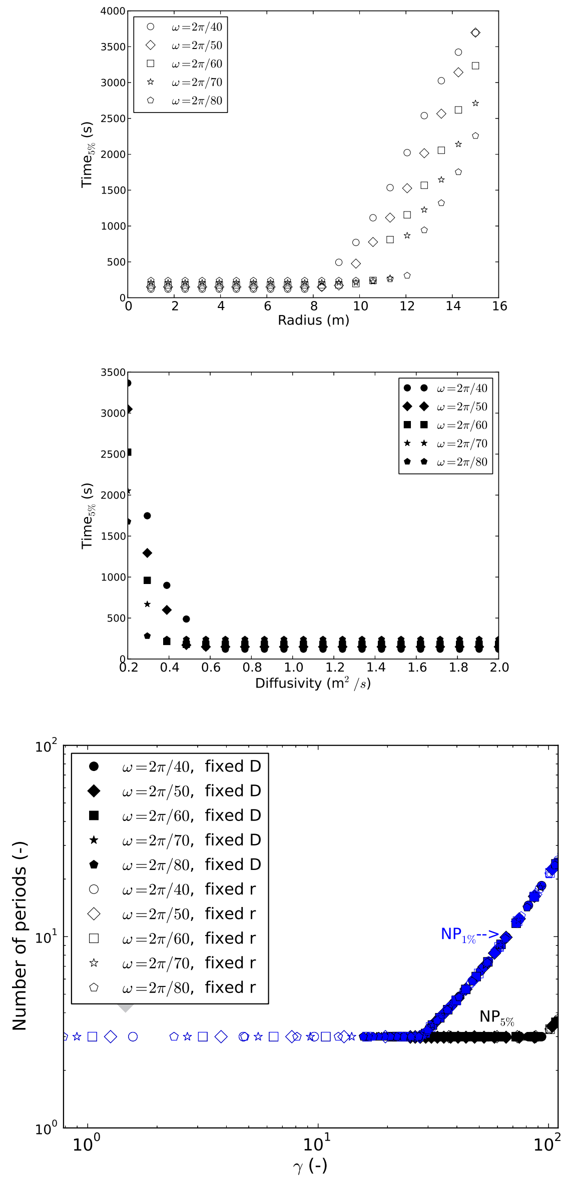

Figure 12: $q = Q_0 \cos(\omega t)(1 - \exp(-(t/T)^2))$, T being the period of oscillations. Behavior of T$_{5\%}$ as (top) diffusivity is fixed - $D = 0.1$ (m^2/s), radius varies and (middle) radius is fixed $r = 20$ (m), diffusivity varies. (bottom) loglog plot demonstrating data collapse using the scaling parameter γ. Note that the blue and black lines correspond to 1 and 5 percent respectively. The hollow symbols correspond to the case where radius is fixed, i.e. the top plot, and the shaded symbols to the case where the diffusivity is fixed, i.e the middle plot. Using the non-dimensional scaling, they collapse onto a single line. The minimum number of periods we considered was 3 periods.

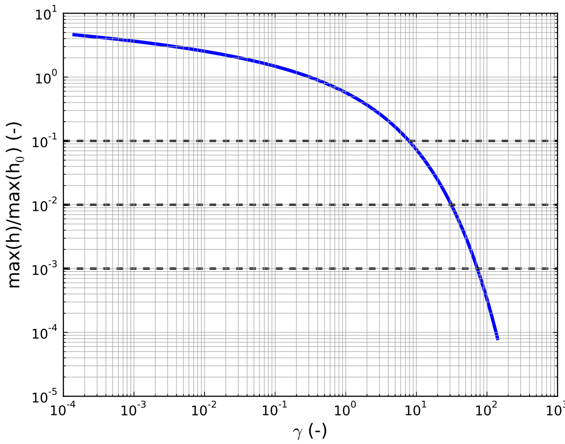

Figure 13: The attenuation of the signal with γ. $h_0 = \frac{Q}{2\pi T}$. Note that at $\gamma \approx 100$, $h \approx \frac{Q}{2\pi T} * 10^{-3}$. With typical values of $Q = 0.1$ (L/s) and $T = 10^{-4}$ (m^2/s), the signal for $\gamma = 100$ would be $h \approx 2 * 10^{-4}$ (m).

source and the hydraulic head respectively. Define $M^{1/2} = U\Lambda^{1/2}U^T$ where the columns of U are the eigenvectors of M. Then by multiplication of (A.1) with $M^{-1/2}$,

$$A\tilde{h}(x,t) + \frac{\partial \tilde{h}(x,t)}{\partial t} = \tilde{b}e^{i\omega t} \tag{A.2}$$

where $A = M^{-1/2}KM^{-1/2}$ is a symmetric positive definite matrix, $\tilde{h} = M^{1/2}h$ and $\tilde{b} = M^{-1/2}b$. The solution to (A.2) is given by the variation-of-constants formula (Hochbruck and Ostermann, 2010).

$$\tilde{h}(x,t) = \int_0^t e^{-(t-s)A}\tilde{b}e^{i\omega s}ds \tag{A.3}$$

Assume a diagonalization of A, $A = VDV^T = \sum_{j=1}^n \lambda_j v_j v_j^T$, where V is the matrix whose columns are the eigenvectors of A, v_j, and D is a diagonal matrix whose diagonal is comprised of the eigenvalues of A, λ_j. Evaluating (A.3),

$$\tilde{h}(x,t) = \left(\sum_{j=1}^n \frac{e^{i\omega t} - e^{-\lambda_j t}}{\lambda_j + i\omega} v_j v_j^T\right)\tilde{b} \tag{A.4}$$

As $t \to \infty$, $\tilde{h}(x,t)$ reaches a quasi-steady state. Using the property that $\|V\|_2 = \|V^T\|_2 = 1$,

$$\|\tilde{h}(x,t) - \sum_{j=1}^{n} \frac{e^{i\omega t}}{\lambda_j + i\omega} v_j v_j^T \tilde{b}\|_2 \leq \frac{e^{-\lambda_{min} t}}{\sqrt{\lambda_{min}^2 + \omega^2}} \|\tilde{b}\|_2 \tag{A.5}$$

For a given tolerance *tol*, the time needed to wait until the hydraulic head reaches quasi-steady state globally is,

$$T = \frac{1}{\lambda_{min}} \log \left(\frac{\|\tilde{b}\|_2}{tol * (\sqrt{\lambda_{min}^2 + \omega^2})} \right) \tag{A.6}$$

1. We analyze the behavior of the transient for homogeneous aquifers and show it scales with $\omega r^2/D$.
2. We derive bounds for the duration of the initial transient for heterogeneous aquifers.
3. We showcase the denoising properties of linear regression on signals subjected to various types of noise.
4. We perform a joint inversion for storativity and transmissivity on synthetic data.