Quasi Monte Carlo methods applied to equations in transient regime on the Theis equation

Juarez S. Azevedo,
CETEC-UFRB, Centro, 44380-000, Cruz das Almas-BA
E-mail: juarezsa@gmail.com

Resumo: In this study, we present the basic-concepts of ground-water hydraulic on stochastic media, whose the Theis equation is used in the transient movement of groundwater as result of pumping in a confined aquifer in saturated porous media under random parameters. A special importance is given on circumstances which requires a high-dimension stochastic to obtain a certain precision in probability space. As an alternative, we introduce quasi-Monte Carlo methods considering Sobol and Halton sequences for uncertainty assessment. Accuracy and efficiency are studied here with allusion to model problem using as reference, the solution obtained by Monte Carlo method and numerical experiments on two-dimensional random field identified some restrictions with the increase of realizations number.

Palavras-chave: Exponential covariance, Karhunen-Loève expansion, Monte Carlo, quasi-Monte Carlo, Theis equation

1 Model Problem

The problem is to determine the drawdown distribution around a well that penetrates a heterogeneous confined aquifer. It is assumed that the permeability, compressibility, and the thickness of each layer does not vary with the time but with the distance in an confined aquifer in saturated porous media whose discharge is given by a constant rate $Q$. In this case, the equation of continuity and Darcy’s law, which governs the movement of water respectively toward the well are

$$S(x)\frac{\partial p(x; \omega; t)}{\partial t} + \nabla \cdot q(x; \omega; t) = f(x)$$  (1)

and

$$q(x; \omega; t) = -T(x; \omega)\nabla p(x; \omega; t),$$  (2)

subjected to initial and boundary conditions

$$\begin{cases}
p(x; \omega; 0) = H_0, & x \in D \subset \mathbb{R}^2 \\
p(x; \omega; t) = h_0(r) & x \in \Gamma_D, 
\end{cases}$$  (3)

where $p(x; \omega; t)$ is hydraulic head, $H_0$ is the constant hydraulic head until the start of pumping of the well, and $h_0(r)$ is the prescribed head in the boundary $\Gamma_D$ formed by points at a radial distance $r$ from well large enough.

In addition, $q(x; \omega; t)$ is the discharge vector, $S(x)$ the elastic storage coefficient, $T(x; \omega)$ the hydraulic transmissivity field, and $f(x) = Q \cdot \delta(x)$ where $\delta(x)$ is the Dirac delta function.

Already in the spatial discretization by finite elements, the weak form of the equation (1) consist in find $p \in H_1^0(D)$ such that

$$\int_D \left( \frac{\partial p(x; \omega; t)}{\partial t} \phi(x) + T(x; \omega)\nabla p(x; \omega; t)\nabla \phi(x) \right) dx$$  (4)

$$= \int_D f(x)\phi(x) dx, \quad \forall \phi \in H_1^0(D),$$

983
where $H^0_0(D)$ represents the space of functions square integrable, whose derivative is square integrable and has compact support in $D$.

For this variational problem we use the functions $H_0^0(r)$ calculated from the deterministic solutions $h = h(x, t)$ given by Theis through drawdown $s$:

$$s = H^2_0 - h^2 = \frac{Q}{2\pi T} W \left( \frac{r^2}{4\nu t} \right),$$

with $W(u)$ defined as well function to nonleaky aquifers expressed by

$$W(u) = \int_{u}^{\infty} \frac{1}{x} \exp(-x) \, dx.$$  

The formula (5) has been widely used to analyze pumping-test data and to determine the parameters $S$ and $T$ of aquifers [3]. Note that in $t = 0$, the drawdown $s$ at the well is zero throughout the aquifer. Subsequently, after that the well is opened, the water level at the well drops to a lower level $s^w$ which is constant during the pumping period.

In the next section we will make the stochastic representation of the parameter $T(x; \omega)$.

2 Stochastic representation of parameter

We begin by presenting $(\Omega, \mathcal{F}, \mu)$ a complete probability space, with $\Omega$ representing the set of outcomes, $\mathcal{F} \subset 2^\Omega$ consisting of the $\sigma$-algebra of events, and $\mu : \mathcal{F} \rightarrow [0, 1]$ a probability measure.

We introduce the model uncertainty, quantified by the independent variable $\omega \in \Omega$ and by transmissivity $T(x, \omega)$ as function of a random lognormal representation, i.e:

$$T(x; \omega) = \exp(Y(x; \omega)),$$

where $Y(x; \omega)$ a stationary Gaussian process. We assume also that the covariance function $C(x, y)$ of the process $Y(x; \omega)$ satisfies $C(x, y) = C(y, x)$ e $C(x, x) > 0$ for all $x \neq 0$. In this form, $Y(x; \omega)$ can be represented in terms of the Karhunen-Loève expansion [1]

$$Y(x; \omega) = \langle Y(x; \omega) \rangle + \sum_{n=1}^{\infty} \sqrt{\lambda_n} \varphi_n(x) \xi_n,$$

in that $\langle Y(x; \omega) \rangle$ represent the mean of $Y$ and $\{\xi_1, \xi_2, \ldots\}$ is assumed to be a set of real mutually orthonormal Gaussian random variables with mean zero. For numerical purposes, it is necessary to truncate the expansion above as

$$Y_M(x; \omega) = \langle Y(x; \omega) \rangle + \sum_{n=1}^{M} \sqrt{\lambda_n} \varphi_n(x) \xi_n,$$

reducing the transmissivity

$$T(x; \omega) \approx T(x; \xi_1, \ldots, \xi_M).$$

The term $M$ represent the number of coordinates of the effective stochastic characterization and is referred as the stochastic dimension.

The terms $\lambda_n$ e $\varphi_n(x)$ from equation (8) represent the eigenvalue and eigenfunction respectively, associated with $C(x, y)$ and can be obtained from Fredholm equation

$$\int_{D} C(x, y) \varphi_n(x) \, dx = \lambda \varphi_n(y).$$

In this experiment we used the exponential covariance. Similar results were obtained by power-law covariance [2] and not will be shown here. The exponential covariance function is defined as

$$C(x, y) = \sigma_1^2 \exp(-|x - y|/\eta),$$
where \( \sigma^2 \) and \( \eta \) are the variance and correlation length of the stationary random process. According to [6] the choice of exponential covariance suggests that the decay of the eigenvalue depends primarily on correlation length so that for larger correlation length, the eigenvalue decay is fast and a small number of terms is necessary in the Karhunen-Loève expansion. Otherwise, for small correlation length, the eigenvalue decay slower and more terms are needed to achieve a better convergence.

### 3 Monte Carlo and quasi-Monte Carlo methods

In this section we shall briefly describe the classical stochastic methods for the calculation of statistical moments of head \( p \). We review the Monte Carlo method (MC) and the quasi-Monte Carlo method (QMC).

The MC iteratively evaluating a deterministic model using sets of random numbers as inputs where involves uncertain parameters. This method involves \( N_r \) realizations that converge asymptotically with \( 0(\sqrt{N_r}) \), independently of the size of the stochastic dimension \( M \). To constructs an approximation of the expected value of the solution by MC on finite element method, observe that the probability space \((\Omega, \mathcal{F}, \mu)\) can be replaced by \((\Gamma, B(\Gamma), \rho(\xi)d\xi)\), where \( \Gamma = \xi(\Omega) \subset \mathbb{R}^M \), \( B(\Gamma) \) denotes the Borel \( \sigma \)-algebra on \( \Gamma \), and \( \rho(\xi)d\xi \) is the probability measure of the vector \( \xi \) [4]. Therefore, we can proceed as follows:

1. Generate \( N_r \) realizations of random field \( Y \) based on the given distribution attributed to \( Y \). In this case, we consider sets of normally distributed, independent samples \( \xi^{(1)}, \ldots, \xi^{(N_r)} \).

2. Solve the deterministic problem

\[
(S\partial_t p_h(:\xi^{(j)}), \phi)_{L^2(D)} + (T(:\xi^{(j)}) \nabla p_h(:\xi^{(j)}), \nabla \phi)_{L^2(D)} = (f, \phi)_{L^2(D)},
\]

such that \( \phi \in W_h(D) \) for each transmissivity \( T(x; \xi^{(j)}) \), \( j = 1, \ldots, N_r \), based on random variables \( \{\xi^{(j)}\}_{n=1}^{N_r} \), and find a corresponding approximation \( p_h \) in the piecewise linear finite element \( W_h(D) \subset H^1_0(D) \);

3. Finally, calculate the statistical moments;

The first moment of \( p_h(x; \xi) \), defined by the \( M \)-dimensional integral

\[
\mu_{p_h}(x) = \int_{\Omega} p_h(x; \xi) \rho(\xi) \, d\xi
\]

is approximated by the equally-weighted average

\[
\mu_{p_h}^{MC}(x) \approx \frac{1}{N_r} \sum_{k=1}^{N_r} p_h(x; \xi^{(k)}).
\]

In the QMC, deterministic points \( X^{(1)}, X^{(2)}, \ldots, X^{(N_r)} \) are chosen in hypercube \([0, 1]^M\). This points are uniformly distributed so that their degree of uniformity is established by difference between the discrete uniform distribution and the continuous uniform distribution on \([0, 1]^M\). The points that satisfy this property are elements of a low-discrepancy sequence [5]. In the experiments we deal Sobol and scramble Halton sequences [8, 7].

As we assume a Gaussian process in the experiments, a change of variables is necessary in the low-discrepancy sequence \( \{X^{(j)}\}_{j=1}^{N_r} \). Thus, given an univariate standard normal cumulative distribution function \( \Phi : \Gamma \to [0, 1]^M \), the equation (12) can be rewritten as:

\[
(S\partial_t p_h(:\Phi^{-1}(X^{(j)}), \phi)_{L^2(D)} + (T(:\Phi^{-1}(X^{(j)}) \nabla p_h(:\Phi^{-1}(X^{(j)})), \nabla \phi)_{L^2(D)} = (f, \phi)_{L^2(D)}, \quad \phi \in W_h(D).
\]
Consequently the moments of (15) are approximated by

$$\mu_{ph}^{QMC}(x) \approx \frac{1}{N_r} \sum_{k=1}^{N_r} p_h(x; \Phi^{-1}(X^{(k)})),$$

$$\sigma^2_{ph}^{QMC}(x) \approx \frac{1}{N_r} \sum_{k=1}^{N_r} (p_h(x; \Phi^{-1}(X^{(k)})) - \mu_{ph}^{QMC}(x))^2.$$  

(16)

4 Results

In this section we examine the performance of the methods in computing the expected values of some quantities of interest for our transient model problem in 2D. In particular, we consider (1) on $D = [0, 1000]^2$ with

$$f(x) = Q \delta(x - x_1) - Q \delta(x - x_2)$$

and subject to a pumping rate was set at $Q = 0.125 \text{ m}^3/\text{s}$. The wells are located in positions $x_1 = (270, 500)$ and $x_2 = (730, 500)$. The average transmissivity is given by $T_G = \exp(\langle Y \rangle) = 0.0038 \text{ m}^2/\text{s}$ and the total time of observation was fixed to 2 and 100 days.

No flow is prescribed under the bottom or top of the aquifer and flow in the lateral boundaries extending to infinity with hydraulic load $H_0 = 380 \text{ m}$. In the following examples, we establish the boundary conditions through the Theis formula (5). For the purpose of comparison, we conducted Monte Carlo simulations. In each case, we use 10000 two-dimensional unconditional realizations generated on the grid of $40 \times 40$ nodes with the separable covariance function given by (11), based on (9) with 1600 terms. The unsteady state, saturated flow equation is solved for each realization of the log hydraulic conductivity, using code Matlab®.

4.1 Numerical study of statistical convergence

As described in the previous section a reference solution for the purpose of comparing the accuracy of the QMC methods is computed by MC. To numerically illustrate the statistical convergence of the MC, we define the relative error of the moments involved $m_p$ as

$$E_{rel}(m_p(N_r)) = \frac{\|m_p(N_r + 1) - m_p(N_r)\|}{\|m_p(N_r + 1)\|},$$

(17)

where $m_p(N_r)$ denotes the mean or variance of $N_r$ samples generated by the Monte Carlo method based on the expansion (9), which was computed by a piecewise-constant finite element approximation of the Fredholm integral equation (10).

Figures 1-2 illustrate the statistical convergence of mean and variance for two times, with exponential covariance both evolved in time $t_1 = 2$ and $t_2 = 100$ days. The reference solution relative to Monte Carlo method with 10000 realizations was chosen because the fluctuation of the relative error in this number of realizations is below $10^{-6}$ in mean and $10^{-3}$ in variance for both covariance.

4.2 Error and computational efficiency - Exponential Covariance.

Figure 3 describe the profiles of mean and variance for the QMC method in the time $t_1 = 2$ and $t_2 = 100$ days, considering the exponential covariance with $N_r = 1000$ realizations and parameters $\eta = 1$ and $\sigma = 1$. In Figures 3(a) and 3(b) we have the profiles of the mean for the times of 2 and 100 days, respectively. We can observe that during the initial heterogeneity there are large oscillations in the head to certain regions of the domain, but with the advance of time, there is a dissipation of these oscillations. In addition, when the stochastic dimension $M$ increases, the proposed solution approaches the target solution. Figures 3(c) and 3(d) depict
Figura 1: Statistical convergence of the head for the expected value of exponential covariance.

Figura 2: Statistical convergence of the head for the variance value of exponential covariance.

the profiles of the variance of head with increase in large periods of time. This effect may be related to the influence of initial data which decrease for large periods of time and consequently the variance increases. We did not include the Halton sequence results in these figures since a similar behavior was observed for this sequence.

Figure 4 presents the evolution of the relative errors of the mean and variance in terms of number of realizations \( N_r \) in time \( t = 2 \) and \( t = 100 \) using QMC methods on Halton (solid lines) and Sobol (dashed lines) sequences. In this case, we consider a stochastic dimension \( M = 10, 100 \) and 400. Figures 4(c)-4(d) clearly illustrates the influence of the origin on small number of realizations, however this difference loses effect when \( N_r > 1000 \). This shows that the advantage of these low discrepancy sequences does seem to decrease with increasing \( N_r \).
Figura 3: Cross sections of pressure head mean and variance along the line $(x_1, x_2) = (x, x)$ for QMC with Sobol sequence, using the exponential covariance with $N_r = 1000$.

Figura 4: Error (on $L_2$ norm) of the mean and variance with respect to the number of realizations on QMC with Halton (solid lines) and Sobol (dashed lines) sequences using exponential covariance.
5 Conclusion

Stochastic models for transient flow in saturated regions are developed in this study and an approximate solution to this equation is found in the two-dimensional sense considering a spatial discretization by finite element method. In this experiment, the aim was to assess the performance of QMC method on Theis formula, showing the nonsteady hydraulic response as a result of pumping in a confined aquifer on heterogeneous media with the inclusion of the stochastic transmissivity.

This analysis has shown that the relative error of variance has a large fluctuation at the source for exponential covariance. Furthermore, the increase in the number of realizations ($N_r > 1000$) can not lead to important gains in simulation performance over the use of low discrepancy sequences. The results also complement other results concerning some problems with low discrepancy sequences as high-dimensional clustering, possible high correlation of neighboring dimensions and possible difficulties in the transformation methods for the creation of non-uniformly distributed low discrepancy sequences.

Referências


