Parallel Inverse Modeling and Uncertainty Quantification for Computationally Demanding Groundwater-Flow Models Using Covariance Matrix Adaptation

Ahmed S. Elshall[1]; Hai V. Pham[2]; Frank T.-C. Tsai, M.ASCE[3]; Le Yan[4]; and Ming Ye, A.M.ASCE[5]

Abstract: This study investigates the performance of the covariance matrix adaptation-evolution strategy (CMA-ES), a stochastic optimization method, in solving groundwater inverse problems. The objectives of the study are to evaluate the computational efficiency of the parallel CMA-ES and to investigate the use of the empirically estimated covariance matrix in quantifying model prediction uncertainty due to parameter estimation uncertainty. First, the parallel scaling with increasing number of processors up to a certain limit is discussed for synthetic and real-world groundwater inverse problems. Second, through the use of the empirically estimated covariance matrix of parameters from the CMA-ES, the study adopts the Monte Carlo simulation technique to quantify model prediction uncertainty. The study shows that the parallel CMA-ES is an efficient and powerful method for solving the groundwater inverse problem for computationally demanding groundwater flow models and for deriving covariances of estimated parameters for uncertainty analysis. **DOI: 10.1061/(ASCE)HE.1943-5584.0001126.** © *2014 American Society of Civil Engineers.*

Author keywords: Groundwater; Inverse modeling; Stochastic optimization; Covariance matrix; Evolution strategy; Uncertainty quantification; Parallel computing.

Introduction

The use of optimization algorithms for solving the inverse problem in subsurface hydrology is a common practice. The classes of optimization algorithms include local derivative algorithms, global heuristic algorithms, hybrid global-heuristic local-derivative algorithms, and global-local heuristic algorithms. While the local derivative algorithms are computationally efficient and can handle a larger number of unknown model parameters, this can be at the cost of finding local solutions instead of a near global solution. The second class of algorithms is the global heuristic algorithms, which are generally implemented when a gradient search is not successful. The heuristic algorithms are experience-based techniques that utilize simple to complex forms of learning to escape local optima and improve the solution. Many studies use global heuristic algorithms such as the genetic algorithm (ElHarrouni et al. 1996; Karpouzos et al. 2001; Solomatine et al. 1999; Bastani et al. 2010) and particle swarm optimization (Scheerlinck et al. 2009; Jiang et al. 2010; Krauße and Cullmann 2012) to avoid entrapment at local minima. This class of algorithms might experience poor local convergence properties. Thus, a third class of algorithms for solving the inverse problem in subsurface modeling is to use the hybrid global-heuristic local-derivative algorithms (Tsai et al. 2003a, b; Blasone et al. 2007; Matott and Rabideau 2008; Zhang et al. 2009), which run a global heuristic algorithm for exploring the search landscape followed by a local derivative algorithm for exploiting favorable search regions. The fourth class of algorithms is the global-local heuristic algorithms, which can perform both global search and local convergence without the need of combining two different algorithms. For solving the inverse problem in groundwater modeling and quantifying model parameter uncertainty, this study investigates the sequential and parallel performance of the covariance matrix adaptation-evolution strategy (CMA-ES) (Hansen and Ostermeier 2001; Hansen et al. 2003) as a global-local stochastic derivative-free algorithm for groundwater model calibration and parameter uncertainty quantification.

The CMA-ES is becoming popular within the environmental modeling community particularly for solving groundwater design-optimization problems (Bayer and Finkel 2004, 2007; Bürger et al. 2007; Bayer et al. 2008, 2010) and for solving parameter estimation problems (Skahill et al. 2009; Keating et al. 2010; Bledsoe et al. 2011; Elshall et al. 2013; Razavi and Tolson 2013; Tsai and Elshall 2013; Yu et al. 2013; Arsenault et al. 2014; Elshall and Tsai 2014). For parameter estimation, few studies (Skahill et al. 2009; Bledsoe et al. 2011; Arsenault et al. 2014) compared the performance of CMA-ES to other derivative and derivative-free algorithms. However, these studies used CMA-ES with the default population size. CMA-ES is quasi-parameter free with the population size being the only parameter to be tuned by the user. Increasing the population size will generally improve the solution precision since the local search of CMA-ES will become more global. However, increasing the population size encourages global exploration at the expense of

[1]Postdoctoral Fellow, Dept. of Scientific Computing, Florida State Univ., 400 Dirac Science Library, Tallahassee, FL 32306; formerly, Graduate Student, Dept. of Civil and Environmental Engineering, Louisiana State Univ., Baton Rouge, LA 70803. E-mail: aelshall@fsu.edu

[2]Graduate Student, Dept. of Civil and Environmental Engineering, Louisiana State Univ., 3418G Patrick F. Taylor Hall, Baton Rouge, LA 70803. E-mail: hpham28@lsu.edu

[3]Associate Professor, Dept. of Civil and Environmental Engineering, Louisiana State Univ., 3418G Patrick F. Taylor Hall, Baton Rouge, LA 70803 (corresponding author). E-mail: ftsai@lsu.edu

[4]Manager, HPC User Services, Louisiana State Univ., High Performance Computing, Frey Computing Services Center Building, Baton Rouge, LA 70803. E-mail: lyan1@tigers.lsu.edu

[5]Associate Professor, Dept. of Scientific Computing, Florida State Univ., 400 Dirac Science Library, Tallahassee, FL 32306. E-mail: mye@fsu.edu

Note. This manuscript was submitted on June 29, 2014; approved on October 14, 2014; published online on November 17, 2014. Discussion period open until April 17, 2015; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Hydrologic Engineering***, © ASCE, ISSN 1084-0699/04014087(11)/$25.00.

convergence speed (Skahill et al. 2009). While this is true for the sequential implementation, this study shows that increasing the population size will significantly improve the computational efficiency for the parallel CMA-ES implementation.

Parallel inverse modeling with CMA-ES can be implemented by an embarrassingly parallel technique. Algorithms that utilize independent solutions in each iteration allow for embarrassingly parallel computation (Vrugt et al. 2006, 2008; Tang et al. 2007, 2010). This is the most efficient parallel technique since the solutions in each iteration do not communicate, as explained later. The first objective of this study is to show that the parallel CMA-ES superiorly improves the calibration speed over the sequential CMA-ES. In addition, the speedup of parallel runs scales variably with increasing the number of processors, which is equal to the population size, up to a certain limit. This study makes an addition to the aforementioned literature on the CMA-ES since this is the first work to the authors' knowledge that examines the parallel performance of the CMA-ES in subsurface hydrology.

The second objective of this study is to investigate the use of the CMA-ES to quantify the model prediction uncertainty due to the parameter estimation error. The solution of the CMA-ES, which consists of a maximum likelihood estimate and a full covariance matrix of estimated model parameters, can be used for uncertainty analysis. Several studies have proposed the utilization of the covariance matrix for sampling target distributions (Haario et al. 1999, 2001; Kavetski et al. 2006a, b; Gallagher and Doherty 2007; Smith and Marshall 2008; Cui et al. 2011; Lu et al. 2012). However, all studies used a covariance matrix within the Markov chain Monte Carlo (MCMC) scheme such that the covariance matrix is used as a proposal distribution to initialize Metropolis-MCMC algorithms or for checking the convergence of MCMC. As pointed out by Müller (2010), the CMA-ES shares many common concepts and features with the derivative-free MCMC simulation algorithms (e.g., Haario et al. 1999, 2001, 2006; Andrieu and Thoms 2008; Müller and Sbalzarini 2010) because both MCMC algorithms and the CMA-ES algorithm are distribution estimation algorithms. However, the MCMC algorithms are more general since they can be used to estimate and sample any probability distribution of the parameters, while the CMA-ES algorithm assumes a multi-Gaussian distribution and estimates a full covariance matrix of the estimated parameters. A formal theoretical and computational comparison between the proposed sampling method and MCMC is beyond the scope of this study, yet this is warranted in a future study.

The sampling method proposed in this study has the following three practical advantages. First, the estimated covariance matrix, which is a by-product of the calibration process, is useful for modeling approaches that attempt to separate various sources of model prediction uncertainty (e.g., Elshall and Tsai 2014) in comparison to the approaches that provide lumped prediction uncertainty. Second, the estimated covariance matrix can be used to generate samples with efficient analytical techniques such as the Cholesky decomposition. Unlike the MCMC sampling methods, Cholesky decomposition sampling can be done independently without the need of model evaluation. Third, this study shows that after reaching a target distribution representing the maximum likelihood estimate of the model parameters and their covariances, the parameter estimation error can be quantified with a relatively small number of Monte Carlo realizations. To the authors' knowledge, this is the first work that investigates the use of the CMA-ES to quantify model parameter uncertainty and model prediction uncertainty.

The organization of the study is as follows. The parallel CMA-ES algorithm is presented in the next section. A challenging synthetic groundwater inverse problem is designed to test the CMA-ES performance against five other commonly used derivative and derivative-free algorithms. Then, the solution of the CMA-ES is analyzed to assess the plausibility of using the empirically estimated covariance matrix to quantify head prediction uncertainty. Impact of the population size is evaluated for the sequential and parallel performance of the CMA-ES. Finally, the parallel CMA-ES is applied to calibrate two computationally demanding groundwater models of the Baton Rouge aquifer system, Louisiana, and to analyze head prediction uncertainty.

Method

For complex groundwater models that generally take hours to run, using the sequential CMA-ES for solving the groundwater inverse problem is impractical due to the prohibitive computational cost. This study resolves this computational issue by implementing the CMA-ES in a high-performance computing (HPC) cluster using an embarrassingly parallel master/slave technique. The embarrassingly parallel master/slave technique treats the individual solutions as explicit tasks that do not communicate with each other, and assigns each task to a processor. Thus, embarrassingly parallel problems are the easiest to parallelize and have negligible parallelization overhead. This section explains the parallelization scheme and the role of population size in increasing the parallelization efficiency.

Given an optimization problem with a search space dimension n, λ candidate solutions, and μ_w best solutions, the iterative solution of the CMA-ES using a master/slave technique is as follows. First, at the master node at any search step $g \geq 1$, the $(\mu_w, \lambda)-$ CMA-ES samples $\mathbf{v}_i \in \mathbb{R}^n$, $\forall\ i = 1 \ldots \lambda$ candidate solutions from a multivariate normal distribution $\pi[\mathbf{v}^{(g+1)}|\mathbf{m}^{(g)}, \sigma^{(g)}\mathbf{C}^{(g)}]$ with mean vector $\mathbf{m}^{(g)}$, step size $\sigma^{(g)}$, and covariance matrix $\mathbf{C}^{(g)}$. The step size $\sigma^{(g)}$ is a scalar that controls the global step length. The covariance matrix $\mathbf{C}^{(g)}$ determines the shape of the search distribution ellipsoid. These candidate solutions are distributed to the slave nodes to run the simulation models and calculate the fitting errors $f(v_i) \in \mathbb{R}$, $\forall i = 1 \ldots \lambda$ accordingly. Since the internal computational time of the algorithm is negligible in comparison to a single groundwater forward model run, the parallel implementation is λ times faster than the sequential implementation, given a number of processors equal to the number of solutions. Third, the computing nodes pass the fitting errors to the CMA-ES at the master node to compute the weighted recombination of the best solutions μ_w out of all the solutions λ. Then, the search parameters are updated, and new solutions are sampled and passed to the slave nodes for the next iteration. The search parameters are the distribution mean $\mathbf{m}^{(g)} \in \mathbb{R}^n$, the overall step size $\sigma^{(g)} > 0$, the symmetric and positive definite covariance matrix $\boldsymbol{C}^{(g)} \in \mathbb{R}^{nxn}$, and the two self-adaptive search paths $\mathbf{p}_\sigma \in \mathbb{R}^n$ and $\mathbf{p}_c \in \mathbb{R}^n$ that act as conjugates for the step-size update and covariance matrix update, respectively. For the update procedures of these five parameters, the reader is referred to Hansen et al. (2003). A flowchart (Fig. S1) of the parallel implementation of the CMA-ES algorithm is provided in the supplementary materials. Here the authors only briefly discuss covariance matrix update since it has an important implication on the parallel implementation of the CMA-ES.

The adaptation of the covariance matrix $\mathbf{C}^g$, which learns all pairwise dependencies between the decision variables, follows a natural gradient approximation of the expected fitting error since there is a close relation between the covariance matrix and the Hessian matrix. The search path $\mathbf{p}_c^{(g)}$, which captures the relation between consecutive steps, is calculated to update the covariance matrix $\mathbf{C}^g$

$$\mathbf{C}^{(g+1)} = (1 - c_1 - c_\mu)\mathbf{C}^{(g)} + c_1 \underbrace{\mathbf{p}_c^{(g+1)} \mathbf{p}_c^{(g+1)T}}_{\text{rank-one update}} + c_\mu \underbrace{\sum_{i=1}^{\mu} w_i \frac{\mathbf{v}_{i:\lambda}^{(g+1)} - \mathbf{m}^{(g)}}{\sigma^{(g)}} \left[\frac{\mathbf{v}_{i:\lambda}^{(g+1)} - \mathbf{m}^{(g)}}{\sigma^{(g)}}\right]^T}_{\text{rank-}\mu\text{ update}} \quad (1)$$

where $c_1 \leq 1 - c_\mu$ are learning rates, $w_i \in \mathbb{R} \forall i = 1 \ldots \mu$ are the recombination weights following a log scale, and $\mathbf{v}_{i:\lambda}^{(g+1)}$ are the best solutions out of $\mathbf{v}_i^{(g+1)}$, $\forall i = 1 \ldots \lambda$. Eq. (1) shows that the covariance matrix adaptation is based on two principles, which are the rank-one update and the rank-μ update. The rank-one update depends on the search paths $\mathbf{p}_\sigma$ and $\mathbf{p}_c$ containing information about the correlation between successive successful steps, and the rank-μ update can be interpreted as the maximum likelihood estimation as the adaptation increases the weighted log-likelihood of previous samples with higher fitness values (Akimoto et al. 2012). The rank-one update of the covariance matrix reduces the number of function evaluations to adapt to a straight ridge from $O(n^2)$ to $O(n)$ (Hansen et al. 2003). Thus, important parts of the model can be learned in an iteration order of n. The rank-one update is important to maintain efficient and premature convergence for a small population size λ. The rank-μ update extends the update rule for large population sizes by using μ vectors to update $\mathbf{C}$ at each step, and thus increases the learning rate for large population sizes and consequently reduces the number of necessary iterations. Yet, increasing the population size λ will increase the number of function evaluations. However, for the parallel implementation of the CMA-ES, increasing the population size will not only improve the solution, but also accelerate the convergence, which is shown in both the synthetic groundwater flow model and the two Baton Rouge groundwater flow models.

Synthetic Groundwater Inverse Problem

A synthetic steady-state groundwater inverse problem is designed to analyze the CMA-ES performance for parameter estimation and uncertainty quantification. The numerical model consists of an unconfined aquifer with a thickness of 400 m and a confined aquifer with a thickness of 100 m, separated by an aquitard with a thickness of 100 m. The top elevation of the model is 200 m. The horizontal domain is 4,500 by 4,500 m and is discretized into 81 computational cells as shown in Fig. 1(a). Each cell is 500 by 500 m. The unconfined aquifer has a fixed head 1 m at the western boundary and is impervious for other three boundaries. The boundaries of the confined aquifer and the aquitard are impervious. Hydraulic conductivity (m/s) for the unconfined aquifer is of two zones, shown in Fig. 1(b):

$$K(x, y) = \begin{cases} 1 \times 10^{-2} & \text{for } x > 2{,}000 \text{ m} \\ 7 \times 10^{-2} & \text{for } x \leq 2{,}000 \text{ m} \end{cases} \quad (2)$$

The confined aquifer has a heterogeneous transmissivity field (m^2/s)

$$T(x, y) = -20\pi \cos(\pi x) \sin(\pi y) - 20\pi \sin(\pi x) \cos(\pi y) + 40\pi^2 (1 + x + y) \cos(\pi x) \cos(\pi y) \quad (3)$$

The vertical hydraulic conductance of the aquitard is 5×10^{-8} m^2/s. Two injection wells are located in the relatively low conductivity zone of the unconfined aquifer with the injection rate of 10 m^3/d for each well. One pumping well is located in the relatively high conductivity zone of the unconfined aquifer with pumping rate of 20 m^3/d. The well locations are shown in Fig. 1(a). The uniform surficial recharge rate of 5×10^{-5} m^2/s is applied to the unconfined aquifer. MODFLOW-2005 (Harbaugh 2005) is used to solve the steady-state flow problem.

Model Parameter Estimation

The 81 cell values of hydraulic conductivity of the unconfined aquifer are assumed not known and are estimated by minimizing the sum of squared errors as follows:

$$\min_{\mathbf{K} \in R^n} \sum_{i=1}^{L} [h_i^{\text{obs}} - h_i(\mathbf{K})]^2 \quad (4)$$

where h_i^{obs} is the ith observed groundwater head; $h_i(\mathbf{K})$ is ith simulated groundwater head; and $\mathbf{K} \in R^n$ is a vector of $n = 81$ unknown hydraulic conductivity values. A complete set of error-free head data (81 head values from the unconfined aquifer and 81 head values from the confined aquifer) is used to eliminate the issues of data error and data sufficiency while evaluating the algorithm performance. Therefore, the data set size is $L = 162$. The initial values of the CMA-ES parameters are $\boldsymbol{p}_\sigma^{(0)} = \boldsymbol{p}_c^{(0)} = 0$, $C^{(0)} = I$, $\boldsymbol{v} = \text{rand}(n)$, and $\sigma^{(0)} = 0.5$ with the default strategy parameters (Hansen et al. 2003).

Performance Evaluation and Comparison

Using the synthetic model for which the true solution is known, five commonly used global heuristic and local derivative algorithms were used to demonstrate the relative performance of CMA-ES in handling several practical challenges, such as high

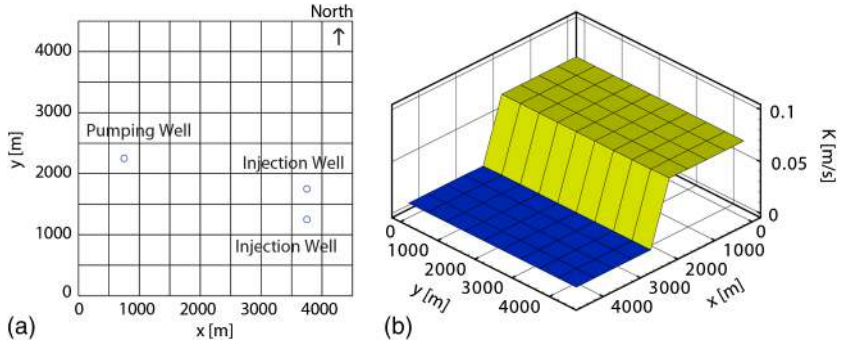

Fig. 1. Synthetic problem: (a) plan view of the pumping well and injection wells; (b) two-zone hydraulic conductivity field for the unconfined aquifer

dimensionality, nonseparablity, and noise. Four global population-based algorithms are considered, which are the ant colony optimization for real domain (Socha and Dorigo 2008), the particle swarm optimization (Iwasaki et al. 2006), the modified deferential evolution (Babu and Angira 2006), and the genetic algorithm (Haupt and Haupt 2004). The ant colony optimization for real domain (ACOR) is selected since it shares the feature of probability distribution estimation with CMA-ES. The particle swarm optimization (PSO) is selected since it is known for its computational efficiency and it is the second most published heuristic algorithm after the genetic algorithm (GA). The modified deferential evolution (mDE) is selected since it belongs to the same class of evolutionary computation of the CMA-ES. In addition, the local derivative Levenberg-Marquardt (L-M) algorithm (Marquardt 1963) is considered. Parameters for each algorithm were tuned to achieve its most effective and efficient performance. In addition, five runs were conducted for each algorithm and the best solution was selected.

Fig. 2 shows the best solutions of the six algorithms and only the CMA-ES succeeded in reaching the true solution. The results are consistent with the findings of Arsenault et al. (2014) that mDE, GA, and PSO performances are generally less robust in comparison to the CMA-ES. The mDE, GA, and ACOR are unable to handle the search difficulties. However, the poor performance of ACOR was unexpected since theoretically the ACOR can handle non-separable functions by invoking correlations between decision parameters and can adapt to a rotating search space. This poor performance may be attributed to the fixed step size of ACOR. The L-M and PSO succeed in recognizing the two hydraulic conductivity zones, but are unable to overcome the noise at the eastern boundary of the low conductivity zone, resulting in imprecise hydraulic conductivity estimation. Since the L-M and PSO cannot effectively exploit the correlation of distinct conductivity zones, the high-conductivity zone does not smooth out. The CMA-ES overcomes these two pitfalls by utilizing the second-order learning through the adaptation of the covariance matrix along with the careful adaptation of the step size to allow for systematic convergence.

The aim of the algorithm comparison is to demonstrate the relative performance of the CMA-ES in handling several practical challenges, such as high dimensionality, non-separability and noise. This is done to introduce the results in Fig. 3 to show that precise estimation of the covariance matrix is not a trivial task since it requires precise estimation of the model parameters.

The CMA-ES performance is further analyzed by showing the solution progress up to 5,000 iterations in Fig. 3. The CMA-ES first detects the hydraulic conductivity structure at iteration 400 and then overcomes the noise through careful adaptation of the step size. After iteration 5,000, the sum of squared errors is negligible and the estimated hydraulic conductivity field is close to the true field [Fig. 1(b)]. The variances of estimated conductivities decrease as the best solution improves and the candidate solutions converge. As the fitting error is significantly reduced at iteration 5,000, the two-zone variance field reflects the two-zone K field and estimation error.

One of the advantages of using the CMA-ES is to quickly obtain the variances of estimated heads once the optimal solution is reached. Fig. 4(a) shows the convergence of the mean head variance [m^2] of the 81 cells of the unconfined aquifer derived from 10,000 realizations of the hydraulic conductivity field for every 100 iterations. The realizations are obtained from the Monte Carlo simulation given the best solution and the covariance matrix of the estimated K field obtained by the CMA-ES. At early iterations, the mean head variance is not necessarily monotonically decreasing because different local minima are sampled along the iterations. The mean head variance decreases and gradually converges after 3,000 iterations. As shown in Fig. 4(b), at early iterations the convergence of the mean head variance requires a large number of realizations and the resulted variance is large due to using incorrect mean and incorrect covariance matrix. After reaching the optimal solution, the mean head variance quickly converges within 100 realizations and the magnitude of the variance is small. Achieving quick convergence of head variances by only a small number of realizations is mainly because the CMA-ES is able to derive smaller parameter estimation variances and more accurate correlations between the estimated parameters.

CMA-ES Population Size

The population size has an important influence on the efficiency of the CMA-ES in terms of the number of iterations and the number of function evaluations to reach stopping criteria. Fig. 5(a) shows the impact of the population size on the number of function evaluations in the sequential run of the CMA-ES. In general, more function

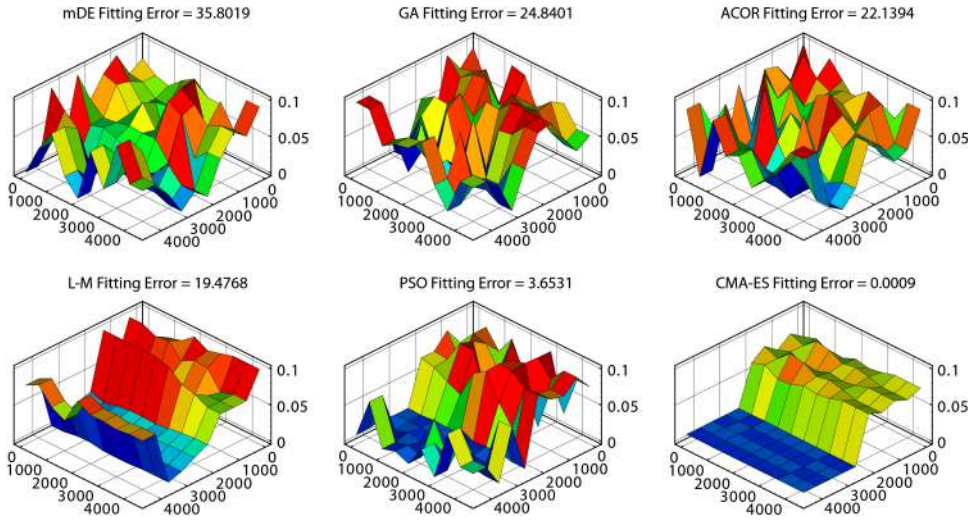

Fig. 2. Hydraulic conductivity solutions for the unconfined aquifer: modified deferential evolution (mDE), genetic algorithm (GA), ant colony optimization for real domain (ACOR), Levenberg-Marquardt (L-M), particle swarm optimization (PSO), and CMA-ES

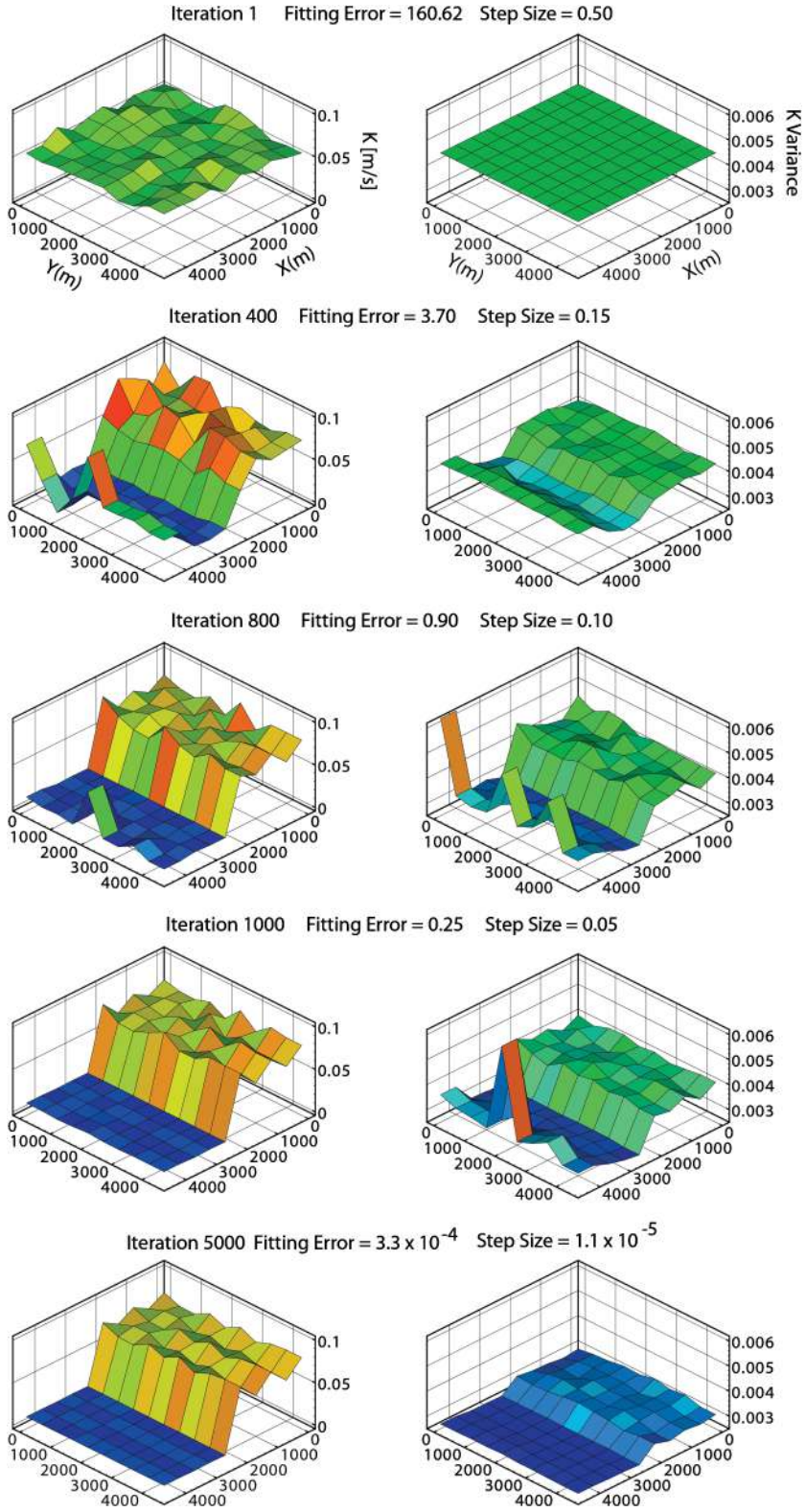

Fig. 3. Solution progress of fitting error, step size, estimated hydraulic conductivity (K), and variance after 1, 400, 800, 1,000, and 5,000 iterations

evaluations are needed to reach solutions that produce smaller fitting error. The default population size $\lambda = 4 + \lfloor 3\ln(n) \rfloor = 17$ and the population size $\lambda = 50$ cannot meet the fitting error criterion 10^{-3}. Since the CMA-ES can detect the global topology by increasing the population size (Hansen and Kern 2004), increasing the population size improves the solution precision. However, increasing the population size increases the number of function evaluations to reach the stopping criteria. According to Fig. 5(a), the optimum population size is 100 for the sequential run.

Increasing the population size is advantageous for the parallel runs of the CMA-ES. Since rank-μ update can effectively exploit the information contained in large population sizes, required iterations can be significantly reduced to reach a certain fitting error. As shown in Fig. 5(b), more iterations are needed to reach solutions

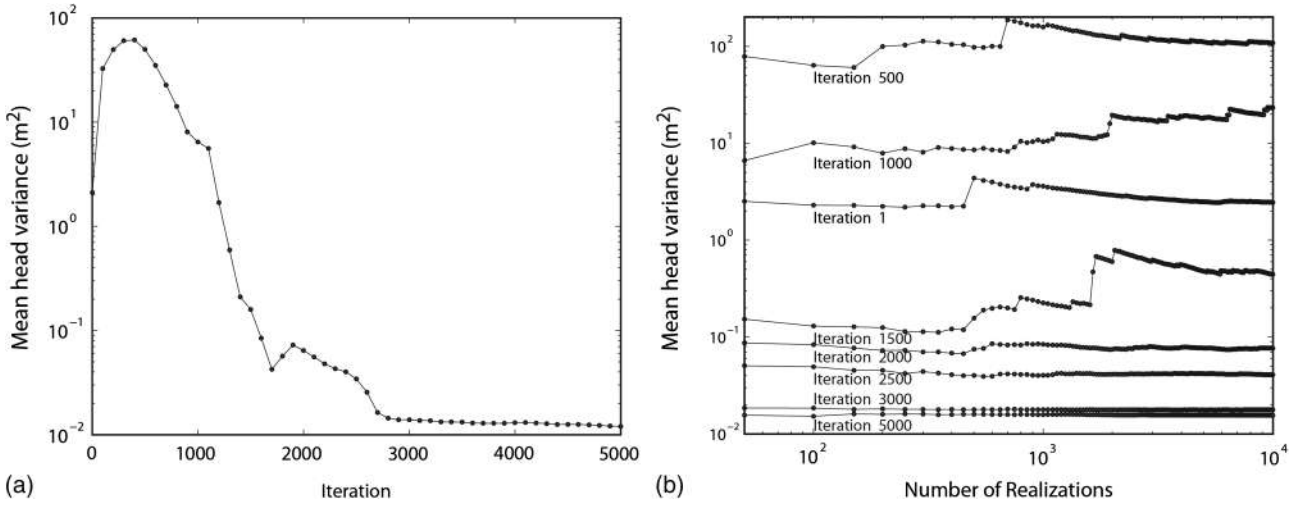

Fig. 4. (a) Convergence profile of mean head variances with respect to the number of iterations; (b) convergence profiles of mean head variances with respect to 10,000 realizations after different iterations for the unconfined aquifer

that produce smaller fitting error. By distributing the candidate solutions of size λ to a number of processors λ, the paralell CMA-ES scales favorably with increasing the number of processors. For example, distributing 100 candidate solutions to 100 processors, the required number of iterations to reach a fitting error 1×10^{-3} is 2,499. The number of iterations is reduced to 964 for using 600 candidate solutions on 600 processors. However, the favorable scaling with increasing the number of processors has a limit. For example, using the population size 700 needs more iterations than using the population size 600 in this case. This finding is consistent with the results of Hansen and Kern (2004) on eight test functions, which show that the scaling could have a convex shape. The optimum population size for the parallel run is about $\lambda \approx 7.4n$ for the synthetic problem.

Case Study: Baton Rouge Groundwater Models

The study area shown in Fig. 6 includes a major part of the Baton Rouge metropolitan area. The Baton Rouge aquifer system consists of complexly interbedded series of fluvial sand and clay units with variable thickness from 6 to 91 m that extends to a depth of 914.4 m (Tomaszewski 1996; Chamberlain 2012). The east-west trending Baton Rouge fault and Denham Springs-Scotlandville fault cut across these sand and clay units. This study develops groundwater models for the 1,200-ft sand, the 1,500-ft sand, the 1,700-ft sand, and the 2,000-ft sand. These sand units were named by their approximate depth below ground level in the Baton Rouge industrial district (Meyer and Turcan 1955). The hydrostratigraphic architecture study (Elshall et al. 2013) shows that the 1,200-ft sand, the 1,500-ft sand, and the 1,700-ft sand between the two faults are interconnected and should be modeled together while the 2,000-ft sand is a separate aquifer. Fig. 7 shows the active computational cells of the 1,200-1,500-1,700-ft sands model and the 2,000-ft sand model, which have 45 and 29 layers, respectively. The layer thickness varies from 1 to 6 m. Each layer consists of 93 rows and 137 columns with a cell size 200×200 m.

The time-varied constant-head boundary condition is assigned to all boundary active cells through extrapolation of the nearby head observation data. Inactive cells represent the clay unit. Detailed pumpage data are available from the Capital Area Ground Water Conservation Commission of Louisiana. In December 2010,

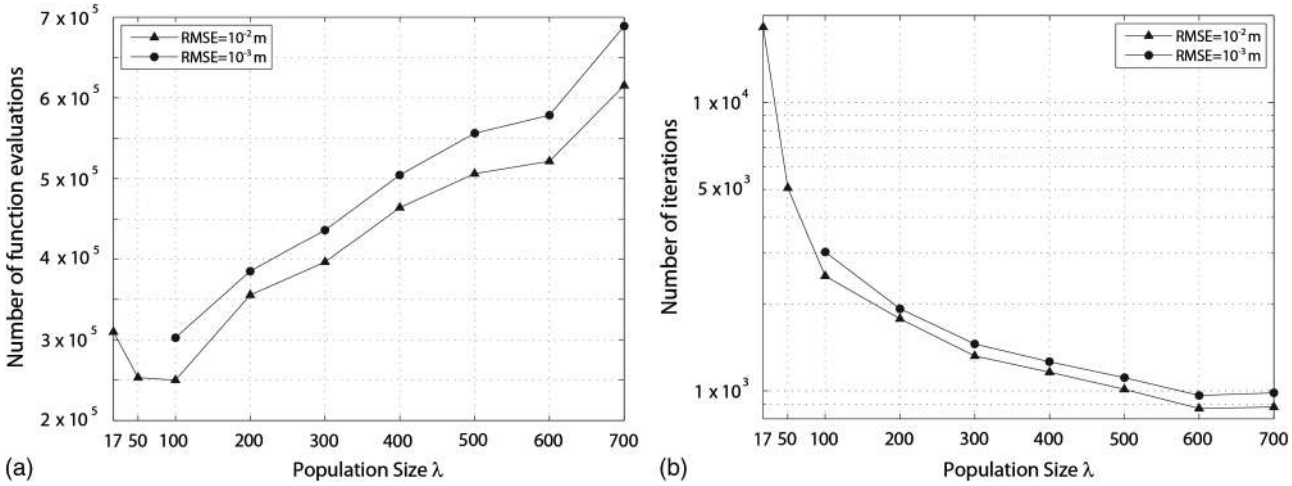

Fig. 5. Impact of the population size on (a) the number of function evaluations to reach the stopping criteria in the sequential run; (b) the number of iterations to reach the stopping criteria for parallel run

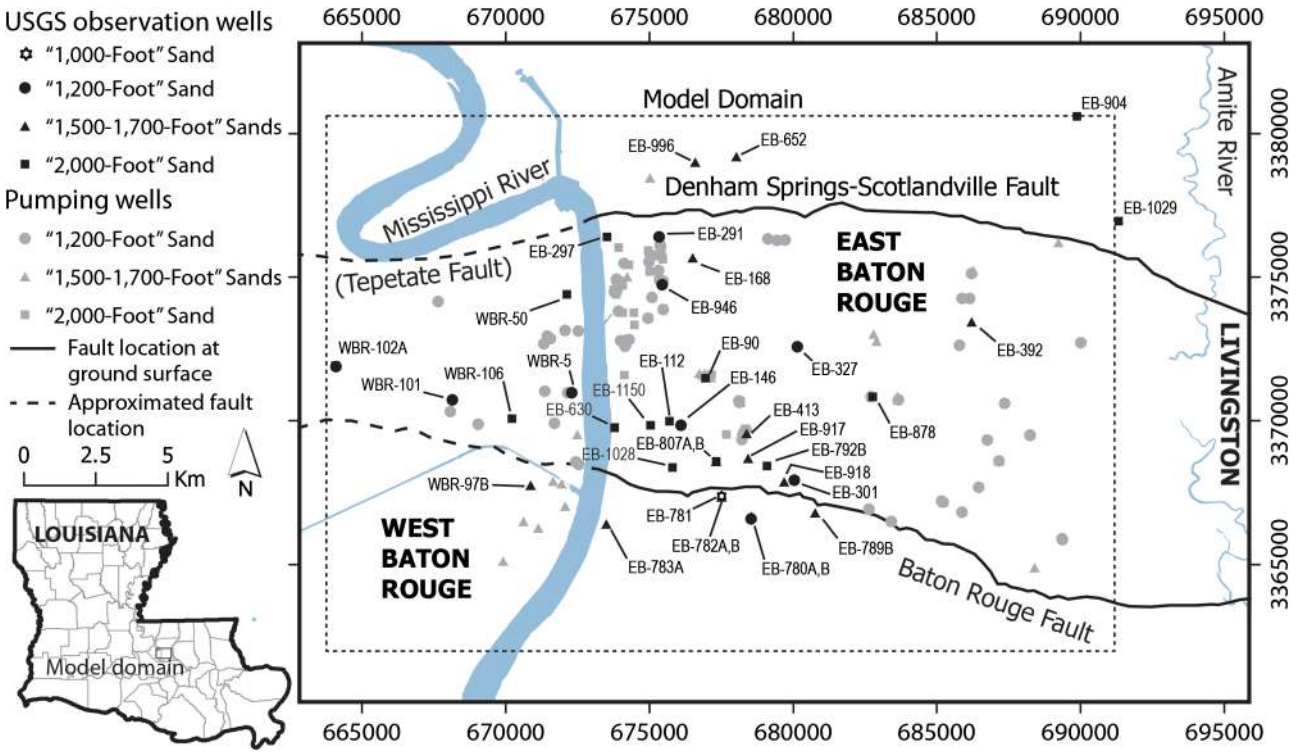

Fig. 6. Map of the study area; the coordinate system is meters, UTM (NAD83)

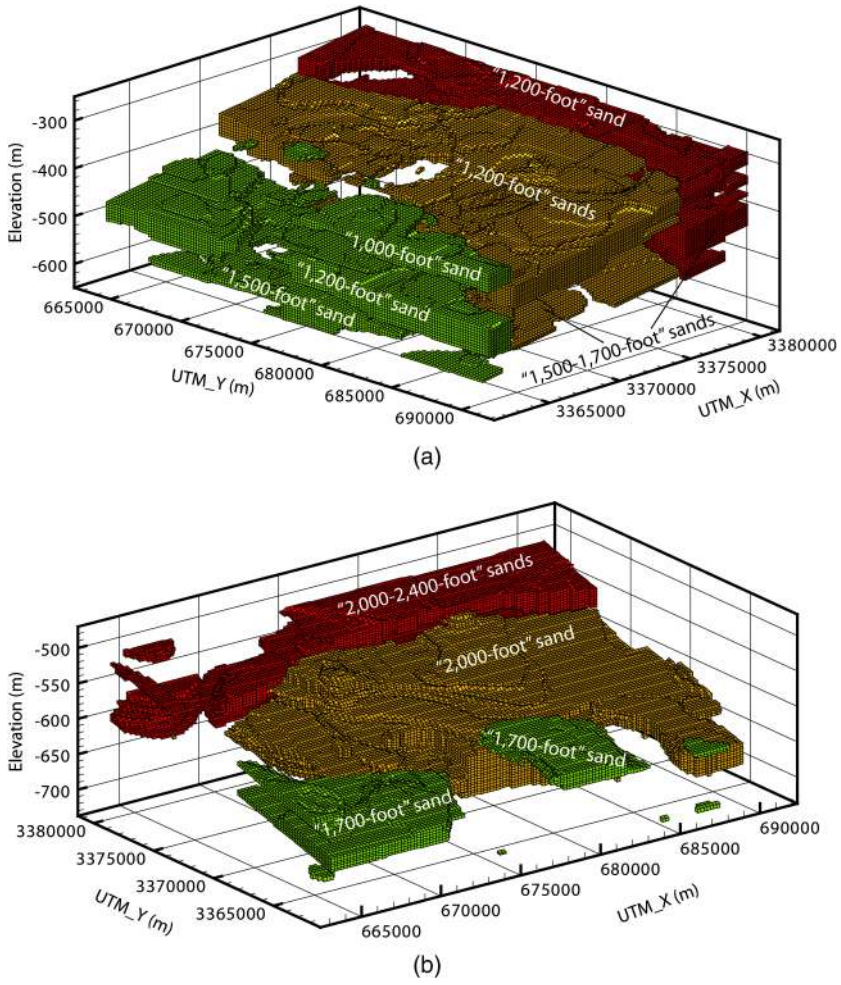

Fig. 7. Active computational cells for (a) the 1,200-1,500-1,700-ft sands model; (b) the 2,000-ft sand model

J. Hydrol. Eng.

Table 1. Parameter Ranges and Estimated Values of the Unknown Model Parameters for Different Sand Units

	1,200-ft sand		1,500-ft sand and 1,700-ft sand		2,000-ft sand	
Parameter	Range	Estimated	Range	Range	Range	Estimated
Hydraulic conductivity [m/d]	$15.00 \sim 35.00$	23.13	$20.00 \sim 32.00$	25.64	$70.00 \sim 170.00$	144.86
Specific storage [1/m]	$2.00 \times 10^{-6} \sim 1.00 \times 10^{-4}$	5.27×10^{-6}	$2.00 \times 10^{-6} \sim 1.00 \times 10^{-4}$	2.82×10^{-6}	$1.00 \times 10^{-5} \sim 3.00 \times 10^{-5}$	1.86×10^{-5}
Vertical anisotropy ratio	—	—	—	—	1.00–5.00	1.01
Hydraulic characteristic of the BR[a] fault [1/d]	$1.00 \times 10^{-4} \sim 1.00 \times 10^{-2}$	2.64×10^{-3}	$1.00 \times 10^{-4} \sim 1.00 \times 10^{-2}$	2.48×10^{-4}	$1.00 \times 10^{-4} \sim 1.00 \times 10^{-2}$	4.20×10^{-3}
Hydraulic characteristic of the DSS[b] fault [1/d]	$1.00 \times 10^{-3} \sim 1.00 \times 10^{-1}$	6.08×10^{-3}	$1.00 \times 10^{-3} \sim 1.00 \times 10^{-1}$	5.00×10^{-2}	$1.00 \times 10^{-6} \sim 1.00 \times 10^{-3}$	1.34×10^{-6}
Boundary condition adjustment factor [m]	—	—	—	—	$-5.00 \sim 5.00$	1.36

[a]Baton Rouge.
[b]Denham Springs-Scotlandville.

the 1,200-1,500-1,700-ft sands model has 87 pumping wells extracting about 112,556 m^3/day and the connector well (EB-1293) injecting about 2,600 m^3/day of groundwater from the 800-ft sand to the 1,500-ft sand. The 2,000-ft sand model has 29 pumping wells extracting about 78,457 m^3/day in December 2010. The 1,200-1,500-1,700-ft sands model is calibrated using 2,805 groundwater head records from 20 USGS observation wells from January 1975 to December 2010. The 2,000-ft sand model uses 1,285 head records from 18 USGS observation wells for the same period.

For the 1,200-1,500-1,700-ft sands model, the 1,500-ft and the 1,700-ft sand are considered to have the same hydrogeological parameter values, which are different from the 1,200-ft sand. The two faults are considered as horizontal flow barriers and their permeability is characterized by the hydraulic characteristic (Hsieh and Freckleton 1993). Hydraulic conductivity, specific storage, and two hydraulic characteristics for the two faults are estimated. Thus, the model has eight unknown model parameters. The 2,000-ft sand model has the following six unknown parameters: hydraulic conductivity, vertical anisotropy ratio, specific storage, two hydraulic characteristics for the two faults, and a boundary head adjustment factor for the eastern boundary between the two faults. The prior parameter ranges and the estimated parameter value are shown in Table 1. The model parameter values are estimated using parallel CMA-ES by minimizing the root-mean squared error (RMSE) between the simulated and observed groundwater heads.

The calibration results in Table 1 show that the 2,000-ft sand has higher hydraulic conductivity and specific storage than the other three sands. The table shows isotropic hydraulic conductivity for the 2,000-ft sand. The Baton Rouge fault and the Denham Springs-Scotlandville fault are low-permeability faults. The Baton Rouge fault permeability for the 1,500-ft sand and the 1,700-ft sand is relatively lower than that for the 1,200-ft sand and 2,000-ft sand. The very low hydraulic characteristic of the Denham Springs-Scotlandville fault for the 2,000-ft sand indicates that groundwater in this sand is mainly from the east.

Parallel Performance

The parallel computation was carried using SuperMike-II, a high performance computer cluster at Louisiana State University with 440 compute nodes. Each compute node is equipped with two 8-Core processors operating at a core frequency of 2.6 GHz. The execution time for a single model simulation is around 1.89 ± 0.10 h for the 1,200-1,500-1,700-ft sands model and 1.28 ± 0.1 h for the 2,000-ft sand model. It is impractical to conduct model calibration using a single processor. For example, given that the algorithm parallelization time is less than 1 s per iteration, the parallel calibration time using 80 processors for the 1,200-1,500-1,700-ft sands model is 75.6 h given 40 iterations with the forward model run being 1.89 h. The equivalent sequential calibration would be $75.6 \times 80 = 6,048$ h, or about 252 days.

For parallel computation, the population size λ is equal to the number of processors. The optimal population size for the groundwater models is determined by performing model calibration with different population sizes $\lambda = 16$, 32, 48, 64, and 80. For the best performance, Hansen and Ostermeier (2001) and Hansen et al. (2003) recommended $4 + \lfloor 3 \ln(n) \rfloor \leq \lambda \leq 10n$. Thus, $\lambda = 80$ was selected as the maximum population size for the 1,200-1,500-1,700-ft sands model given $n = 8$, and $\lambda = 64$ as the maximum population size for the 2,000-ft sand model given $n = 6$. Note that the population size $\lambda = 64$ for the 2,000-ft sand is slightly over the recommended range of $10n$.

Fig. 8 shows the number of iterations required to meet different stopping criteria of the RMSE versus the number of processors. The required number of iterations to reach convergence for the 1,200-1,500-1,700-ft sands model is about double that of the 2,000-ft sand model. This can be attributed to the larger number of unknown parameters, the complexity of the geological structure, and the plausibility of the prior parameter ranges (Table 1). For both groundwater models, more iterations are needed for small RMSE criteria. Small RMSE criteria may not be met using smaller population sizes as the search becomes less global. For example, as shown in Fig. 8(a) the population size $\lambda = 16$ does not reach the RMSE 1.50-m criterion. Increasing the population size will always reduce the number of iterations. Thus, the optimal population size for the 1,200-1,500-1,700-ft sands model is $\lambda = 80 = 10n$ and for the 2,000-ft sand model is $\lambda = 64 \approx 10.67n$. For both models, the optimal population size is at the upper limit $\lambda = 10n$.

Similar to the synthetic problem, Fig. 8 shows that increasing the population size requires less iterations to reach the fitting error criteria. As shown in Fig. 8(a), given RMSE 1.57 m for the 1,200-1,500-1,700-ft sands model, a speedup $84/41 = 2.05$ is achieved by the optimal population size ($\lambda = 80$) with respect to the default population size ($\lambda = 16$). Given RMSE 2.95 m for the 2,000-ft sand model, a speedup $55/24 = 2.29$ is achieved by the optimal population size ($\lambda = 64$) with respect to the default population size ($\lambda = 16$) as shown in Fig. 8(b).

Groundwater Head Uncertainty

The mean and the covariance matrix of the estimated model parameters are used for Monte Carlo simulation to quantify head

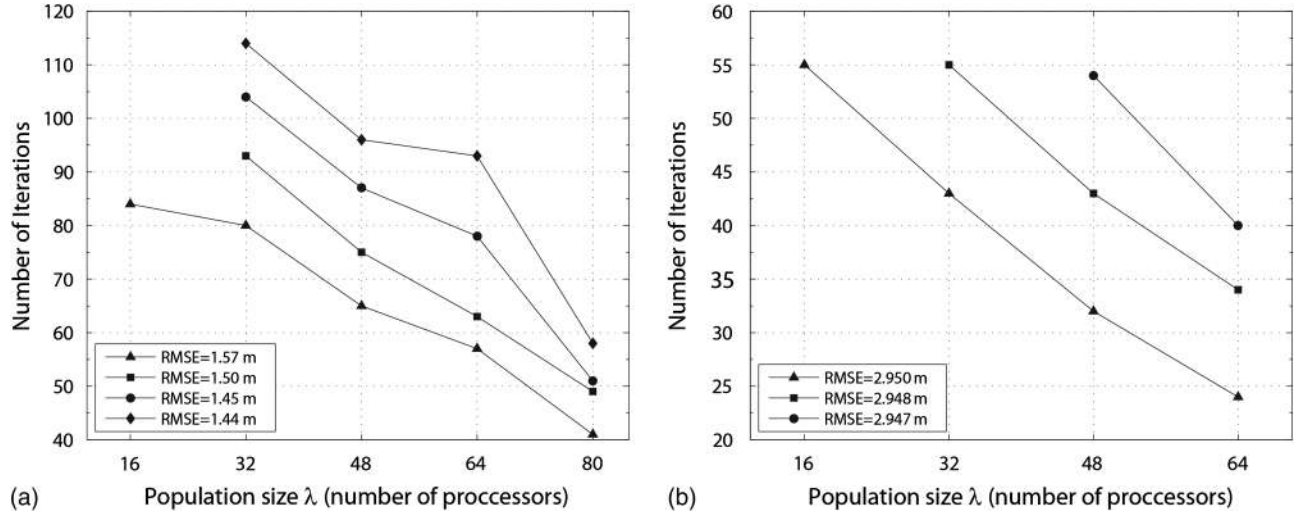

Fig. 8. Number of iterations for different population sizes required to reach several target fitting errors for (a) the 1,200-1,500-1,700-ft sands model; (b) the 2,000-ft sand model

uncertainty due to parameter estimation uncertainty. Head variances at USGS observation well EB-291 (screened at the 1,200-ft sand) in August 1978 and observation well WBR-106 (screened at the 2,000-ft sand) in September 1976 are calculated since these two data points show the highest head standard deviation among all observation data. The location of EB-291 and WBR-106 is shown in Fig. 6.

Using the optimal parallel population size, Figs. 9(a and b) show that the improvement of RMSE is negligible after about 50 iterations and 30 iterations for the 1,200-1,500-1,700-ft sands model

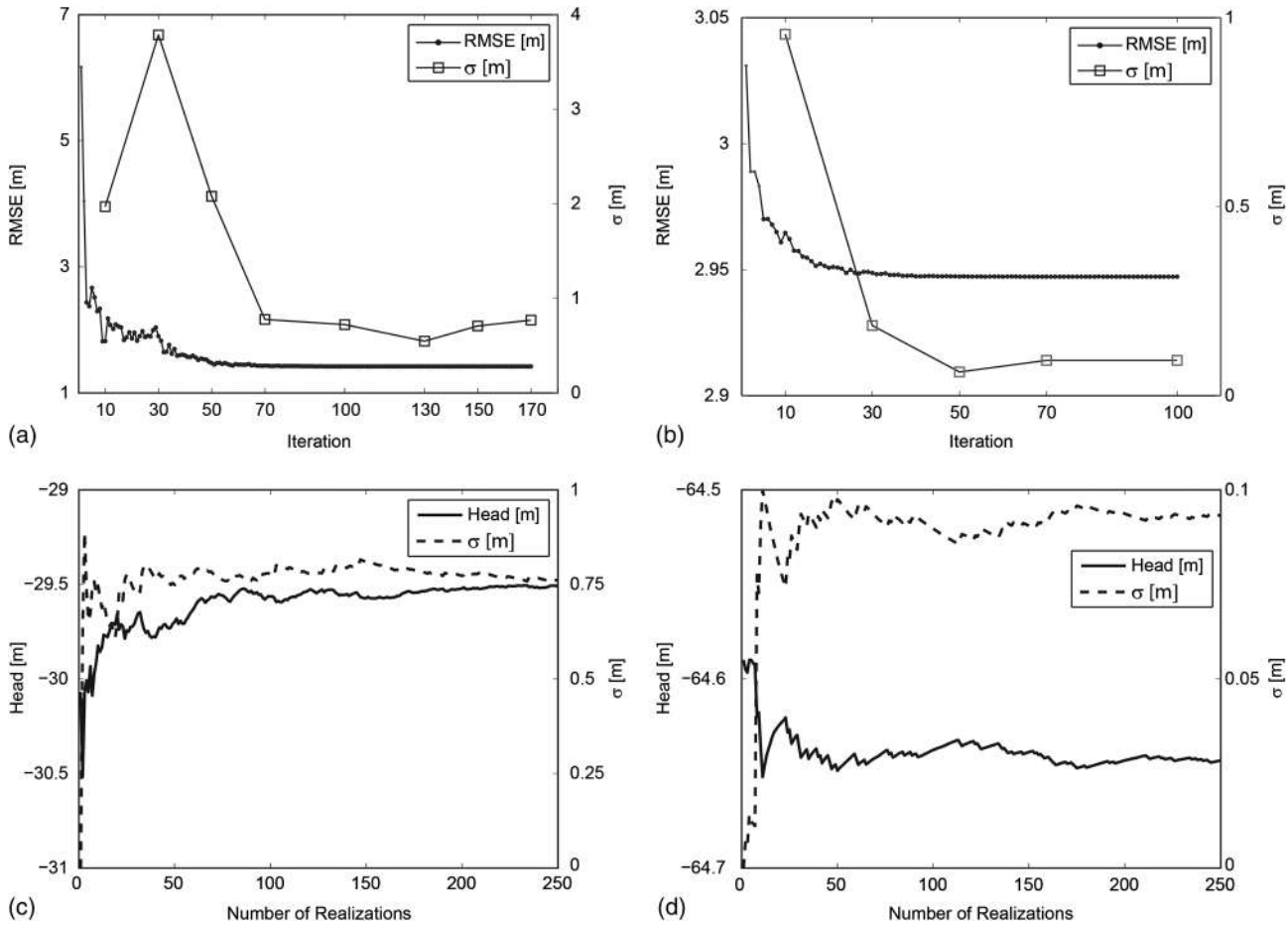

Fig. 9. Convergence profiles of the RMSE [m] and head standard deviation [m] of the selected observation points based on 250 realizations after each iteration for (a) the 1,200-1,500-1,700-ft sands model; (b) the 2,000-ft sand model. Convergence profiles of head [m] and head standard deviation [m] for (c) EB-291 in August 1978 after 170 iterations; (d) WBR-106 in September 1976 after 100 iterations

and the 2,000-ft sand model, respectively, yet the calibration was not terminated to ensure the convergence of the covariance matrix estimation. The head standard deviation [m] is calculated based on 250 realizations after each iteration for the selected observation points as shown in Figs. 9(a and b). The results indicate that the CMA-ES reaches convergence with respect to covariance matrix estimation within 170 and 100 iterations for the 1,200-1,500-1,700-ft sands model and 2,000-ft sand model, respectively. Similar to the synthetic case it is observed that the estimation of covariance matrix requires about three times more iterations than the estimation of the mean solution.

Figs. 9(c and d) show the head prediction and the head prediction standard deviation for the selected observation points at the last iteration intervals for both models. The results are consistent with the synthetic case study that after reaching the optimal solution, the mean head variance quickly converges in less than 100 realizations.

Conclusions

The results of the synthetic groundwater inverse problem show that the CMA-ES is able to detect and overcome the noise at the fixed head boundary through careful adaptation of the step size along with the covariance matrix. However, this is computationally demanding and requires more iterations than detecting the correct shape of the hydraulic conductivity field.

The authors clarify the difference between an adequate model and a precise inverse solution by the case studies. The results show that the retrieved head standard deviation for the two groundwater models in the real-world case study is very small in comparison to the RMSE. The small head standard deviation is due to small estimated parameter variance, which is a measure of the precision of the inverse solution regardless of the adequacy of the model. This is not the case with the synthetic case study since it has no data and model structure error. On the contrary, the synthetic model has negligible fitting error, yet relatively high variance due to nonuniqueness.

The study provides several practical hints about the computational cost of the CMA-ES for groundwater model calibration and uncertainty quantification. First, the population size tuning follows a specific pattern such that the search becomes more global by increasing the population size as shown in the synthetic and the real-world problems. Second, the parallel CMA-ES encourages the development of realistic groundwater models due to the significant alleviation of the computational cost. Increasing the population size reduces the number of iterations to meet the stopping criteria.

The study shows that the empirically estimated covariance matrix can be used for Monte Carlo sampling to quantify parameter-related uncertainty. After the covariance matrix is estimated, only a small number of realizations are required for the convergence of the mean head prediction and head prediction variance.

Acknowledgments

This study was supported in part by the National Science Foundation under Grant No. EAR-1045064, the U.S. Geological Survey under Grant/Cooperative Agreement No. G10AP00136, and LSU Economic Development Assistantship (EDA). The authors acknowledge computing resources from the LSU High Performance Computing. The authors thank two anonymous reviewers for their helpful comments.

Supplemental Data

A flowchart (Fig. S1) of the parallel implementation of CMA-ES algorithm is available online in the ASCE Library (www.ascelibrary.org).

References

Akimoto, Y., Nagata, Y., Ono, I., and Kobayashi, S. (2012). "Theoretical foundation for CMA-ES from information geometry perspective." *Algorithmica*, 64(4), 698–716.

Andrieu, C., and Thoms, J. (2008). "A tutorial on adaptive MCMC." *Stat. Comput.*, 18(4), 343–373.

Arsenault, R., Poulin, A., Côté, P., and Brissette, F. (2014). "Comparison of stochastic optimization algorithms in hydrological model calibration." *J. Hydrol. Eng.*, 10.1061/(ASCE)HE.1943-5584.0000938, 1374–1384.

Babu, B. V., and Angira, R. (2006). "Modified differential evolution (mDE) for optimization of non-linear chemical processes." *Comput. Chem. Eng.*, 30(6–7), 989–1002.

Bastani, M., Kholghi, M., and Rakhshandehroo, G. R. (2010). "Inverse modeling of variable-density groundwater flow in a semi-arid area in Iran using a genetic algorithm." *Hydrogeol. J.*, 18(5), 1191–1203.

Bayer, P., Bürger, C. M., and Finkel, M. (2008). "Computationally efficient stochastic optimization using multiple realizations." *Adv. Water Resour.*, 31(2), 399–417.

Bayer, P., de Paly, M., and Bürger, C. M. (2010). "Optimization of high-reliability-based hydrological design problems by robust automatic sampling of critical model realizations." *Water Resour. Res.*, 46(5), W05504.

Bayer, P., and Finkel, M. (2004). "Evolutionary algorithms for the optimization of advective control of contaminated aquifer zones." *Water Resour. Res.*, 40(6), W06506.

Bayer, P., and Finkel, M. (2007). "Optimization of concentration control by evolution strategies: Formulation, application, and assessment of remedial solutions." *Water Resour. Res.*, 43(2), W02410.

Blasone, R. S., Madsen, H., and Rosbjerg, D. (2007). "Parameter estimation in distributed hydrological modelling: Comparison of global and local optimisation techniques." *Nordic Hydrol.*, 38(4–5), 451–476.

Bledsoe, K. C., Favorite, J. A., and Aldemir, T. (2011). "A comparison of the covariance matrix adaptation evolution strategy and the Levenberg-Marquardt method for solving multidimensional inverse transport problems." *Ann. Nucl. Energy*, 38(4), 897–904.

Bürger, C. M., Bayer, P., and Finkel, M. (2007). "Algorithmic funnel-and-gate system design optimization." *Water Resour. Res.*, 43(8), W08426.

Chamberlain, E. L. (2012). "Depositional environments of upper Miocene through Pleistocene siliciclastic sediments, Baton Rouge aquifer system, southeastern Louisiana." M.Sc. thesis, Louisiana State Univ., Baton Rouge, LA.

Cui, T., Fox, C., and O'Sullivan, M. J. (2011). "Bayesian calibration of a large-scale geothermal reservoir model by a new adaptive delayed acceptance Metropolis Hastings algorithm." *Water Resour. Res.*, 47(10), W10521.

ElHarrouni, K., Ouazar, D., Walters, G. A., and Cheng, A. H. D. (1996). "Groundwater optimization and parameter estimation by genetic algorithm and dual reciprocity boundary element method." *Eng. Anal. Boundary Ele.*, 18(4), 287–296.

Elshall, A. S., and Tsai, F. T.-C. (2014). "Constructive epistemic modeling of groundwater flow with geological structure and boundary condition uncertainty under the Bayesian paradigm." *J. Hydrol.*, 517, 105–119.

Elshall, A. S., Tsai, F. T.-C., and Hanor, J. S. (2013). "Indicator geostatistics for reconstructing Baton Rouge aquifer-fault hydrostratigraphy." *Hydrogeol. J.*, 21(8), 1731–1747.

Gallagher, M., and Doherty, J. (2007). "Parameter estimation and uncertainty analysis for a watershed model." *Environ. Model. Software*, 22(7), 1000–1020.

Haario, H., Laine, M., Mira, A., and Saksman, E. (2006). "DRAM: Efficient adaptive MCMC." *Stat. Comput.*, 16(4), 339–354.

Haario, H., Saksman, E., and Tamminen, J. (1999). "Adaptive proposal distribution for random walk Metropolis algorithm." *Comput. Stat.*, 14(3), 375–395.
Haario, H., Saksman, E., and Tamminen, J. (2001). "An adaptive Metropolis algorithm." *Bernoulli*, 7(2), 223–242.
Hansen, N., and Kern, S. (2004). "Evaluating the CMA evolution strategy on multimodal test functions." *Parallel problem solving from nature-PPSN VIII*, X. Yao, et al., eds., Springer, Berlin, Germany, 282–291.
Hansen, N., Müller, S. D., and Koumoutsakos, P. (2003). "Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES)." *Evol. Comput.*, 11(1), 1–18.
Hansen, N., and Ostermeier, A. (2001). "Completely derandomized self-adaptation in evolution strategies." *Evol. Comput.*, 9(2), 159–195.
Harbaugh, A. W. (2005). "MODFLOW-2005, The U.S. Geological Survey modular groundwater model- the ground-water flow process." *U.S. Geological Survey Techniques and Methods 6-A16*, U.S. Geological Survey, Reston, VA.
Haupt, R. L., and Haupt, S. E. (2004). *Practical genetic algorithms*, 2nd Ed., Wiley, New York.
Hsieh, P. A., and Freckleton, J. R. (1993). "Documentation of a computer program to simulate horizontal-flow barriers using the U.S. Geological Survey modular three- dimensional finite-difference ground-water flow model." *U.S. Geological Survey Open-File Rep. 92-477*, U.S. Geological Survey, Sacramento, CA.
Iwasaki, N., Yasuda, K., and Ueno, G. (2006). "Dynamic parameter tuning of particle swarm optimization." *IEEJ Trans. Electr. Electron. Eng.*, 1(4), 353–363.
Jiang, Y., Liu, C., Huang, C., and Wu, X. (2010). "Improved particle swarm algorithm for hydrological parameter optimization." *Appl. Math. Comput.*, 217(7), 3207–3215.
Karpouzos, D. K., Delay, F., Katsifarakis, K. L., and de Marsily, G. (2001). "A multipopulation genetic algorithm to solve the inverse problem in hydrogeology." *Water Resour. Res.*, 37(9), 2291–2302.
Kavetski, D., Kuczera, G., and Franks, S. W. (2006a). "Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory." *Water Resour. Res.*, 42(3), W03407.
Kavetski, D., Kuczera, G., and Franks, S. W. (2006b). "Calibration of conceptual hydrological models revisited: 2. Improving optimisation and analysis." *J. Hydrol.*, 320(1–2), 187–201.
Keating, E. H., Doherty, J., Vrugt, J. A., and Kang, Q. (2010). "Optimization and uncertainty assessment of strongly nonlinear groundwater models with high parameter dimensionality." *Water Resour. Res.*, 46(10), W10517.
Krauße, T., and Cullmann, J. (2012). "Towards a more representative parametrisation of hydrologic models via synthesizing the strengths of particle swarm optimisation and robust parameter estimation." *Hydrol. Earth Syst. Sci.*, 16(2), 603–629.
Lu, D., Ye, M., and Hill, M. C. (2012). "Analysis of regression confidence intervals and bayesian credible intervals for uncertainty quantification." *Water Resour. Res.*, 48(9), W09521.
Marquardt, D. W. (1963). "An algorithm for least-squares estimation of nonlinear parameters." *J. Soc. Ind. Appl. Math.*, 11(2), 431–441.
Matott, L. S., and Rabideau, A. J. (2008). "Calibration of complex subsurface reaction models using a surrogate-model approach." *Adv. Water Resour.*, 31(12), 1697–1707.
Meyer, R. R., and Turcan, A. N., Jr. (1955). "Geology and ground-water resources of the Baton Rouge area, Louisiana." *U.S. Geological Survey Water Supply Paper 1296*, U.S. Government Printing Office, Washington, DC.
Müller, C. L. (2010). "Exploring the common concepts of adaptive MCMC and covariance matrix adaptation schemes." *Theory of evolutionary algorithms*, A. Auger, J. L. Shapiro, L. D. Whitley, and C. Witt, eds., Dagstuhl, Germany, 1–10.
Müller, C. L., and Sbalzarini, I. F. (2010). "Gaussian adaptation as a unifying framework for continuous black-box optimization and adaptive Monte Carlo sampling." *2010 IEEE Congress on Evolutionary Computation (CEC)*, IEEE, NJ, 1–17.
Razavi, S., and Tolson, B. A. (2013). "An efficient framework for hydrologic model calibration on long data periods." *Water Resour. Res.*, 49(12), 8418–8431.
Scheerlinck, K., Pauwels, V. R. N., Vernieuwe, H., and De Baets, B. (2009). "Calibration of a water and energy balance model: Recursive parameter estimation versus particle swarm optimization." *Water Resour. Res.*, 45(10), W10422.
Skahill, B. E., Baggett, J. S., Frankenstein, S., and Downer, C. W. (2009). "More efficient PEST compatible model independent model calibration." *Environ. Model. Software*, 24(4), 517–529.
Smith, T. J., and Marshall, L. A. (2008). "Bayesian methods in hydrologic modeling: A study of recent advancements in Markov chain Monte Carlo techniques." *Water Resour. Res.*, 44(12), W00B05.
Socha, K., and Dorigo, M. (2008). "Ant colony optimization for continuous domains." *Eur. J. Operat. Res.*, 185(3), 1155–1173.
Solomatine, D. P., Dibike, Y. B., and Kukuric, N. (1999). "Automatic calibration of groundwater models using global optimization techniques." *Hydrol. Sci. J.*, 44(6), 879–894.
Tang, G., D'Azevedo, E. F., Zhang, F., Parker, J. C., Watson, D. B., and Jardine, P. M. (2010). "Application of a hybrid MPI/OpenMP approach for parallel groundwater model calibration using multi-core computers." *Comput. Geosci.*, 36(11), 1451–1460.
Tang, Y., Reed, P. M., and Kollat, J. B. (2007). "Parallelization strategies for rapid and robust evolutionary multiobjective optimization in water resources applications." *Adv. Water Resour.*, 30(3), 335–353.
Tomaszewski, D. J. (1996). "Distribution and movement of saltwater in aquifers in the Baton Rouge area, Louisiana, 1990–92." *Water Resources Technical Rep. No. 59*, Louisiana Dept. of Transportation and Development, Baton Rouge, LA.
Tsai, F. T.-C., and Elshall, A. S. (2013). "Hierarchical Bayesian model averaging for hydrostratigraphic modeling: Uncertainty segregation and comparative evaluation." *Water Resour. Res.*, 49(9), 5520–5536.
Tsai, F. T.-C., Sun, N.-Z., and Yeh, W. W.-G. (2003a). "A combinatorial optimization scheme for parameter structure identification in ground water modeling." *Ground Water*, 41(2), 156–169.
Tsai, F. T.-C., Sun, N.-Z., and Yeh, W. W.-G. (2003b). "Global-local optimization for parameter structure identification in three-dimensional groundwater modeling." *Water Resour. Res.*, 39(2), 1043.
Vrugt, J. A., Nuallain, B. O., Robinson, B. A., Bouten, W., Dekker, S. C., and Sloot, P. M. A. (2006). "Application of parallel computing to stochastic parameter estimation in environmental models." *Comput. Geosci.*, 32(8), 1139–1155.
Vrugt, J. A., Stauffer, P. H., Wöhling, T. B., Robinson, A., and Vesselinov, V. V. (2008). "Inverse modeling of subsurface flow and transport properties: A review with new developments." *Vadose Zone J.*, 7(2), 843–864.
Yu, X., Bhatt, G., Duffy, C., and Shi, Y. (2013). "Parameterization for distributed watershed modeling using national data and evolutionary algorithm." *Comput. Geosci.*, 58(8), 80–90.
Zhang, Y., Gallipoli, D., and Augarde, C. E. (2009). "Simulation-based calibration of geotechnical parameters using parallel hybrid moving boundary particle swarm optimization." *Comput. Geotech.*, 36(4), 604–615.