A comparative study in aquifer parameter estimation using MFree point collocation method with evolutionary algorithms

Alice Thomas, T. I. Eldho, A. K. Rastogi and Partha Majumder

ABSTRACT

In this study, we present a comparative assessment of simulation-optimization (S-O) models to estimate aquifer parameters such as transmissivity, longitudinal dispersivity, and transverse dispersivity. The groundwater flow and contaminant transport processes are simulated using the mesh-free radial basis point collocation method (RPCM). Four different S-O models are developed by combining the RPCM model separately with genetic algorithm (GA), differential evolution (DE), cat swarm optimization (CSO), and particle swarm optimization (PSO). The objective of the S-O model is to minimize a composite objective function with transmissivity, longitudinal dispersivity, and transverse dispersivity as decision variables. Hydraulic head and contaminant concentration at observation points are the state variables. The S-O models are used to estimate aquifer parameters of a confined aquifer with nine zones. It is found that RPCM-based DE, CSO, and PSO models are more accurate in estimating aquifer parameters than RPCM-GA. However, for noisy observed data, the RPCM-CSO model outperforms other models. The efficiency of the RPCM-CSO model over other models is further established by performing reliability analysis to the noisy observed data set. The comparative study reflects the efficacy of CSO over GA, DE, and PSO.

Key words | cat swarm optimization (CSO), differential evolutionary algorithm (DE), genetic algorithm (GA), particle swarm optimization (PSO), radial basis point collocation method (RPCM), simulation-optimization (S-O)

Alice Thomas
T. I. Eldho (corresponding author)
A. K. Rastogi
Department of Civil Engineering,
Indian Institute of Technology Bombay,
Mumbai,
India
E-mail: *eldho@civil.iitb.ac.in*

Partha Majumder
Department of Civil Engineering,
Utah State University,
Logan, UT,
USA

INTRODUCTION

Estimation of aquifer parameters such as storativity, transmissivity, porosity, and longitudinal dispersivity by the inverse modeling approach is very important for groundwater management (Peralta 2012). In groundwater inverse modeling, transmissivity, storativity, recharge, leakage, longitudinal dispersivity, transverse dispersivity, and molecular diffusion are the decision variables, whereas the hydraulic head and contaminant concentrations at predefined locations are the state variables (Medina & Carrera 1996). The quality of groundwater models cannot be simply improved by increasing the accuracy of forward solutions if inverse solutions are not properly solved (Sun 1994). In many previous studies, inverse modeling has been widely used to estimate aquifer parameters (Carrera & Neuman 1986a, 1986b; Zimmerman *et al.* 1998; Prasad & Rastogi 2001; Rajanayaka & Kulasiri 2001; Huggi & Rastogi 2003; Carrera *et al.* 2005; Franssen *et al.* 2009; Jin *et al.* 2009; Wang *et al.* 2017). The quantity and quality of the observed spatiotemporal groundwater data (hydraulic head and contaminant concentrations at observation location) greatly influence the results of inverse modeling and most of the time the data are generally inadequate for accurate prediction (Medina & Carrera 1996). Another major thing that governs the error in inverse modeling is

doi: 10.2166/hydro.2019.105

inaccurate model structure identification. By only changing performance criteria and optimization techniques, satisfactory results cannot be found for inverse modeling (Sun 2013). The inverse model determines aquifer parameters by using the spatiotemporal observed data.

Numerical modeling tools such as the finite difference method (FDM), finite element method (FEM), and analytical element method (AEM) are mostly used for groundwater flow and contaminant transport simulation (Fitts 2002). The FDM and FEM require preprocessing effort for mesh generation, hence they are computationally expensive for large aquifers. The AEM has limitations in simulating a highly heterogeneous medium and transient flow condition (Majumder & Eldho 2017). On the contrary, many studies claimed that mesh-free methods are computationally more efficient and accurate than FDM and FEM in simulating large-scale groundwater flow and contaminant transport problems (Guneshwor *et al.* 2018; Praveen Kumar & Dodagoudar 2008; Kovářík & Mužík 2013; Meenal & Eldho 2012; Patel & Rastogi 2017; Swathi & Eldho 2018; Thomas *et al.* 2014, 2018).

Heuristic optimizers such as GA, simulated annealing (SA), DE, particle swarm optimization (PSO), and ant colony optimization (ACO) were mostly used to estimate aquifer parameters by inverse modeling by earlier researchers (Abbaspour *et al.* 2001; Prasad & Rastogi 2001; Reed *et al.* 2001; Huggi & Rastogi 2003; Li *et al.* 2006; Fonna *et al.* 2013; Elçi & Ayvaz 2014; Gurarslan & Karahan 2015; Mategaonkar *et al.* 2018; Swathi & Eldho 2018; Thomas *et al.* 2018). Advantages of using heuristic optimizers over the traditional gradient-based optimization algorithms are that they do not require derivative computations and initial point to initiate the search operation (Majumder & Eldho 2016). A comparison carried out by Matott *et al.* (2006) among various optimization techniques for groundwater remediation studies shows the superiority of PSO over GA, random search algorithm (RSA), SSA, and conjugate gradient. However, PSO sometimes converges prematurely. In PSO, all swarms move towards the best solution (which may be the local best solution) which is found by a specific swarm and if the velocities of all the swarms are very low, then swarms cannot come out from the point. Thus, in such cases, premature convergence may take place and the associated error in the solution is known as stagnation point error (Saha *et al.* 2013; Guo *et al.* 2016; Majumder & Eldho 2016). Recently, cat swarm optimization (CSO) has been found to perform better when compared to PSO, GA, and SA for groundwater management problems (Majumder & Eldho 2016). Seeking mode and tracking mode are the search operations that take place in CSO. In seeking mode, each cat creates many copies around its surroundings, and there is randomness associated with each copy. Hence, the search space of CSO is vast. Due to this seeking mode, the premature convergence/stagnation point error is highly unlikely in CSO (Majumder & Eldho 2016). Successful application of CSO in estimating transmissivity values in a heterogeneous confined aquifer can be found in Thomas *et al.* (2018).

In this study, four S-O models are developed by coupling RPCM separately with GA, DE, CSO, and PSO. The models are applied to estimate transmissivity, longitudinal dispersivity, and transverse dispersivity of a hypothetical confined aquifer case study with nine zones. Further, a reliability analysis is performed to comprehend and intercompare the overall consistency of the output from the various models.

METHODOLOGY

A two-dimensional transient state groundwater flow equation for a heterogeneous confined aquifer can be expressed as (Bear 2012):

$$\frac{\partial}{\partial x}\left[T_x \frac{\partial h}{\partial x}\right] + \frac{\partial}{\partial y}\left[T_y \frac{\partial h}{\partial y}\right] = \mathrm{S}\frac{\partial h}{\partial t} + Q_w\,\delta(x - x_i)(y - y_i) - \mathrm{q} \tag{1}$$

where h is the hydraulic head, T_x and T_y are the transmissivities along x and y direction, Q_w is the vertically averaged source or sink term, q is inflow rate, and S is storativity.

The seepage velocity can be computed using Darcy's equation as (Bear 2012):

$$V_x = -\frac{K_x}{n_e}\frac{\partial h}{\partial x};\; V_y = -\frac{K_y}{n_e}\frac{\partial h}{\partial y} \tag{2}$$

where V_x and V_y are the seepage velocities along x and y direction, K_x and K_y are the hydraulic conductivities, and n_e is the porosity of the porous medium.

The vertically averaged advection-dispersion equation (ADE) describing contaminant transport processes can be expressed as (Freeze & Cherry 1979):

$$\frac{\partial C}{\partial t}=\frac{\partial}{\partial x}\left[D_{xx}\frac{\partial C}{\partial x}\right]+\frac{\partial}{\partial y}\left[D_{yy}\frac{\partial C}{\partial y}\right]-\frac{\partial(V_x C)}{\partial x}-\frac{\partial(V_y C)}{\partial y}+\frac{q_w C'}{ne} \tag{3}$$

where D_{xx} and D_{yy} are the dispersion coefficient, C is the concentration of dissolved species, C' is the concentration of the solute species injected at the sources which are governed by Dirac delta function, and q_w is the volumetric pumping rate.

Dispersion coefficient tensor ($\mathbf{D_{ij}}$) can be expressed as (Bear & Cheng 2010):

$$\mathbf{D_{ij}}=\alpha_T|v|\delta_{ij}+(\alpha_L-\alpha_T)\frac{v_i\times v_j}{|\mathbf{v}|}+\mathbf{D_{ij}^m};$$
$$\delta_{ij}=\begin{Bmatrix}1 & for\ i=j\\ 0 & for\ i\neq j\end{Bmatrix} \tag{4}$$

where δ_{ij} is the Kronecker delta; α_T is the transverse dispersivity (L); α_L is the longitudinal dispersivity (L); v_i and v_j is the seepage velocity in the ith and jth directions (LT^{-1}); $\mathbf{v}$ is the seepage velocity vector; $\mathbf{D_{ij}^m}$ is the effective molecular diffusion coefficient.

Mesh free radial point collocation method

MFree methods use a set of nodes that carry values of the field variables called field nodes which are scattered in the problem domain as well as on the boundaries. They are not for discretizing but rather just to represent the domain. MFree methods do not require any prior information regarding the relationship of these field nodes for the approximation of the unknown functions of field variable as in the case of mesh-based methods (Liu 2002). The radial point collocation method (RPCM) described in Thomas *et al.* (2018) is used in the present study for developing the groundwater flow and transport models. Within the local support domain, the approximation of a function $h(x)$ is expressed as a linear combination of n radial basis function (RBF) $\phi^{\boldsymbol{T}}$ (Liu & Gu 2005) as follows:

$$h(x)=\sum_{i=1}^{n}R_i(x)a_i=R^T(x)a=\phi^{\boldsymbol{T}}(x)h_s \tag{5}$$

where n is the number of nodes in the support domain; a_i is the unknown coefficient to be determined. Here, multiquadric radial basis function (MQ-RBF) is used to develop shape functions, $Ri(x)=(r_i^2+(\alpha_c d_c)^2)^q$ is the radial basis function (RBF); q and α_c, are the shape parameters; $r_i=\sqrt{(x-x_i)^2+(y-y_i)^2}$, d_c is the average nodal spacing for all the nodes in local support domain; (x, y) are the coordinates of the point of interest and (x_i, y_i) are the coordinates of any node in the support domain of the point of interest. $\phi^{\boldsymbol{T}}(x)$ is known as shape function. By using the shape function that is derived as explained in Thomas *et al.* (2018), the hydraulic head can be approximated as:

$$h(x,t)=\phi^T(x)h_s(t)=\sum_{i=1}^{n}\phi_i(x)h_i(t) \tag{6}$$

where n is the number of nodes in the support domain and h is the hydraulic head. Discretization of the governing flow equation is done by simple collocation at all the internal nodes after the head is approximated (Liu & Gu 2005). Thus, by collocation at the point of interest X_r (x_r, y_r), Equation (1), for a homogeneous and isotropic aquifer, is discretized as:

$$T(\boldsymbol{x}_r)\left[\frac{\partial^2\phi^{\boldsymbol{T}}}{\partial x^2}+\frac{\partial^2\phi^{\boldsymbol{T}}}{\partial y^2}\right]h_s(t)=S(\boldsymbol{x}_r)\left(\frac{\partial h}{\partial t}\right)_{x_r}+Q_w\delta(x_r-x_i)(y_r-y_i)-q(\boldsymbol{x}_r) \tag{7}$$

The complete RPCM formulation for solving the groundwater flow equation can be referred from Thomas *et al.* (2018). On rearranging Equation (7) after application of forward difference fully implicit scheme (Rastogi 2007) for the time derivative, the solution of governing the flow equation can be written in matrix form as:

$$h_s^{t+\Delta t}\left([K_1]-\frac{T\Delta t}{S}([K_4]+[K_5])\right)=h_s^t[K_1]-\frac{\Delta t}{S}\delta(x_r-x_i)(y_r-y_i)+q\frac{\Delta t}{S} \tag{8}$$

where $[K_1]$ is global matrix (GM) for shape function; $[K_4]$ is GM for the seccond derivative of shape function with respect to x, and $[K_5]$ is GM for the second derivative of shape function with respect to y.

Seepage velocities V_x, V_y can be evaluated from the known hydraulic head value as:

$$V_x = -K_x \frac{\partial \phi^T}{\partial x} h_s; \quad V_y = -K_y \frac{\partial \phi^T}{\partial y} h_s \tag{9}$$

The Dirichlet boundary condition is handled by directly substituting in the approximation equations. If $h(x,\ y) = h_0$ is the constant head boundary condition, it can be easily discretized as $\phi^T h_s = h_0$. For implementing the derivative boundary condition, a direct collocation approach has been implemented. A direct collocation method for derivative boundary condition implementation is illustrated here by using Equation (10):

$$A(x)\frac{\partial h}{\partial x} + B(x)h + q_b(x) = 0 \tag{10}$$

Equation (10) can be discretized by direct collocation as:

$$\left[A(x)\frac{\partial \phi^T}{\partial x} + B(x)\phi^T\right] h_s = -q_b(x) \tag{11}$$

In a similar manner, RPCM is used to generate a system of algebraic equations in matrix form for ADE Equation (3):

$$\left[\frac{[K_1]}{\Delta t} - \left[D_{xxr}[K_4] + D_{yyr}[K_5] - V_{xr}[K_1] - V_{yr}[K_2]\right]\right] C_s^{t+\Delta t} = [K_1]\frac{C_r^t}{\Delta t} + \frac{q_w\, C}{n}\delta(x_r - x_i)(y_r - y_i) \tag{12}$$

where $[K_1]$ is GM for shape function; $[K_2]$ is GM for the first derivative of shape function with respect to x; $[K_3]$ is GM for first derivative of shape function with respect to y; $[K_4]$ is GM for second derivative of shape function with respect to x; $[K_5]$ is GM for second derivative of shape function with respect to y.

Genetic algorithm (GA)

The genetic algorithm is a metaheuristic optimization technique. Based on Darwin's survival of the fittest principle, GA finds the best and fittest solution. GA can easily find the global optimum solution of large complex nonlinear models unlike traditional optimization methods. The basic elements of GA are reproduction, crossover, and mutation. More information about GA can be found in Deb (2012). In this study, we have directly used the GA optimization toolbox which is available in MATLAB.

Differential evolution (DE)

Differential evolution is a meta-heuristic evolutionary optimization algorithm proposed by Storn & Price (1995). The basic elements of DE are mutation, crossover, and selection. The steps of DE are explained below (Storn & Price 1997).

Step 1: Randomly initialize the position of N populations in D dimensional space within specified upper bound and lower bound.

$$x_{n,i} = x^l_{n,i} + rand_1 \times (x^u_{n,i} - x^l_{n,i}) \quad i = 1, 2, 3, 4, \ldots..D;\ \ n = 1, 2, 3, 4, \ldots..N \tag{13}$$

x_i^l and x_i^u are the lower bound and upper bound of the variable x_i; $rand_1$ is a random number $\in (0,1)$.

Step 2: (Mutation) For each parameter vector, randomly select three other vectors and generate a mutant vector (v_n^{g+1}).

$$v_n^{g+1} = x^g_{r1n} + F(x^g_{r2n} - x^g_{r3n}) \tag{14}$$

$r_1,\ r_2,\ r_3 \in (1, 2, 3, \ldots N)$ and F is a user-defined mutation factor $\in (0, 2)$.

Step 3: (Crossover) Generate a trial vector $(u_{n,i}^{g+1})$ from the target vector $(x_{n,i}^g)$ and the mutant vector $(v_{n,i}^{g+1})$,

$$u^{g+1}_{n,i} = \begin{Bmatrix} v^{g+1}_{n,i} & if\, rand_2 \le CR\, or\, i = I_{rand} \ \ i = 1, 2, 3, \ldots..D\ \& \\ x^g_{n,i} & if\, rand_2 > CR\, or\, i \ne I_{rand} \ \ n = 1, 2, 3, \ldots\ldots N \end{Bmatrix} \tag{15}$$

$rand_2$ is a random number $\in (0,1)$ and CR is the user-defined crossover probability $\in (0,1)$;

I_{rand} is an integer random number $\in (1,2, 3, \ldots..D)$.

Step 4: (Selection) Compare target vector $(x_{n,i}^g)$ and trial vector $(u_{n,i}^{g+1})$. For the next generation select the vector which results in the lowest fitness function value.

$$x_n^{g+1} = \begin{Bmatrix} u^{g+1}_{n,i} & if\, f(u_n^{g+1}) < f(x_n^g) \\ x_n^g & else \end{Bmatrix} \tag{16}$$

Step 5: Continue step 2 to step 4 until the termination criterion is satisfied.

Among the various variants of DE, DE/rand/1/bin is used in the present study.

Cat swarm optimization (CSO)

Chu & Tsai (2007) proposed a metaheuristic optimization algorithm, namely, cat swarm optimization, by imitating the food searching behavior of the cat. In CSO, food searching takes place by seeking mode and tracking mode. Seeking mode is a local search process in which the cat looks for prey with very small changes in position. There are four parameters in seeking mode. These are seeking memory pool (SMP), seeking range of selected dimensions (SRD), count of dimension to change (CDC), and self-position consideration (SPC). Tracking mode is a global search process in which the cat chases the prey with a certain velocity. The velocity decreases with respect to generation number. It also signifies that the cat is moving closer to catch the prey with respect to generation number. The velocity is also multiplied by an inertia weight (w), which varies linearly from 0.9 to 0.4 with respect to generation number (Majumder & Eldho 2016). The seeking mode and tracking mode are connected by a ratio called mixture ratio (MR). The cat mostly spends time in seeking prey and little time in chasing prey. Thus, the MR is assigned in such a way that a higher number of cats belongs to seeking mode and a lower number of cats belongs in tracking mode.

A typical CSO randomly generates the initial positions and velocities of N number of cats in the D dimensional space within the specified position and velocity bound. Thereafter, evaluation of the fitness function value for each cat is done. The cat (g_{best}) with the best fitness value (F_g) is found and is stored in the memory. Then, all the cats are randomly distributed in the seeking and tracking modes depending upon the MR. The results of seeking and tracking mode processes are combined, to yield new position and fitness values of all the cats in the vector form. The cat with the best fitness value (F_t) is found and it replaces F_g, if F_t is better than F_g. Also g_{best} is replaced by the respective cat. Until the termination criterion is satisfied the steps from randomly distributing cats depending upon mixture ratio to seeking mode and tracking mode are repeated. The termination criterion is generally the specified maximum number of iterations/generations (Majumder & Eldho 2016). The more detailed description and equations of CSO can be found in Thomas *et al.* (2018).

Particle swarm optimization (PSO)

In particle swarm optimization, particles move across the search space according to global best position of all particles and local best position of the individual particle (Eberhart & Kennedy 1995). In PSO, there is a velocity update equation and a position update equation. A typical PSO has the following steps (Robinson & Rahmat 2004):

(i) Generate N number of particles in D dimensional space.

(ii) Randomly initialize position (x_i) and velocity (v_i) of each particle within the allowable limits.

(iii) Compute fitness function values of all the particles. Find global best position ($gbest[t]$), global best fitness function value ($Fgbest[t]$), individual best position ($pbest_i[t]$), individual best fitness function value ($Fpbest_i[t]$) before starting iteration and store them. Here, t is the iteration number. Before starting iteration $t = 0$. For the first iteration $t = 1$.

(iv) Update velocity and position of each particle for the next generation using the following equations (Robinson & Rahmat 2004):

$$v_i(t+1) = w \times v_i(t) + c_1 \times r_1[pbest_i(t) - x_i(t)] + c_2 \times r_2[gbest(t) - x_i(t)] \quad (17)$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (18)$$

Check whether velocity and position are within the allowable limits. If velocity and position are out of bounds, then set them to their allowable limits. In the above expressions, w is the inertia weight, which with respect to iteration number decreases linearly from 0.9 to 0.4; r_1 and r_2 is the random number between 0 and 1, $c_1 = c_2 = 1.49$ (Robinson & Rahmat 2004).

(v) With the updated position, compute the fitness value of all the particles.

(vi) If $Fgbest[t+1]$ is better than $Fgbest[t]$, then $Fgbest[t] = Fgbest[t+1]$; also replace $gbest(t) = gbest(t+1)$. If $Fgbest[t+1]$ is not better than $Fgbest[t]$, then keep $Fgbest[t]$ and $gbest(t)$ unchanged.

(vii) Also, for each particle, if $Fpbest_i[t+1]$ is better than $Fpbest_i[t]$, then set $Fpbest_i[t] = Fpbest_i[t+1]$ and $pbest_i[t] = pbest_i[t+1]$; otherwise, keep $Fpbest[t]$ and $pbest(t)$ unchanged.

(viii) Repeat step (iv) to step (vii), until the termination criterion is achieved. Although there are many variants of PSO, we have used standard original PSO in this study (Robinson & Rahmat 2004).

In all the optimization algorithms mentioned above, the termination criterion is a user-defined number of generations.

Verification of optimization models

The four optimization methods discussed above are used to optimize the Michalewicz function. The Michalewicz function has $d!$ local minima and it is a multimodal test function. The Michalewicz function can be expressed as (Molga & Smutnicki 2005):

$$f(x) = -\sum_{i=1}^{d} \sin(x_i)\left[\sin\left(\frac{ix_i^2}{\pi}\right)\right]^{2m} \tag{19}$$

The search space of the function is $x_i = [0, \pi]$. The Michalewicz function is optimized by GA, DE, PSO, and CSO by assuming dimension$(d) = 10$. The parameters' settings for various optimization methods are given in Table 1. The total number of generations (N) is kept as 1,000. The true global minimum of the Michalewicz function for the dimension $(d) = 10$ is -9.66015 (Molga & Smutnicki 2005). The optimum values obtained by DE and CSO are very close to the true optimum (Table 2). However, CSO takes about 600 generations to reach the global optimum value whereas DE takes almost 950 generations to reach the global optimum value (Figure 1). There are no improvements in the results of PSO and GA after 300 generations. This is because the solutions of both PSO and GA are trapped in local minima. From Table 2, it can be seen that even GA and PSO have failed to reach very close to the

Table 1 | Parameter settings of GA, DE, PSO, and CSO

Optimization model	Parameter	Range of value
GA	Population size (N)	40
	Mutation ratio (MR)	0.01
	Crossover (CR)	0.8
DE	Population size (N)	40
	Mutation factor (F)	0.8
	Crossover (CR)	0.9
PSO	Population size (N)	40
	C1	1.49
	C2	1.49
	w	0.9–0.4; linearly decreasing with respect to the iteration number
CSO	Population size (N)	40
	MR	0.8
	SMP	15
	SRD	0.15
	CDC	80%
	w	0.9–0.4, linearly decreasing with respect to the iteration number

Table 2 | Minimum values of Michalewicz's function obtained by GA, DE, CSO, and PSO

	GA	DE	PSO	CSO	True value
Fitness value of Michalewicz function	−8.02	−9.621	−8.11	−9.632	−9.66015
Computational time (seconds)	21.52	5.12	6.31	17.78	

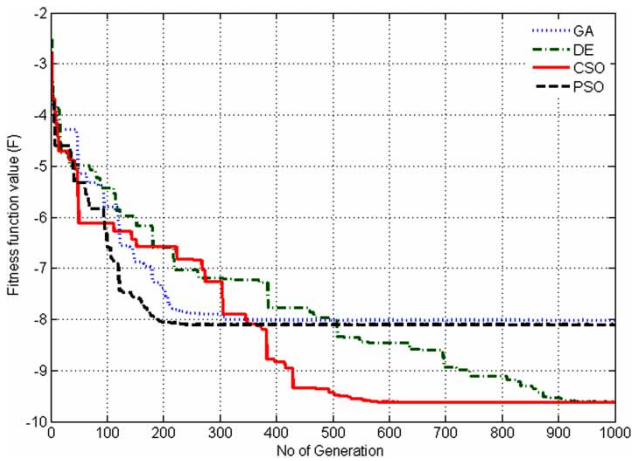

Figure 1 | Best fitness value versus number of generations using GA, DE, CSO, and PSO for Michalewicz's function.

true optimum value even after 1,000 generations. To comprehend the computational performance of CSO, DE, PSO, and GA for the Michalewicz function, a comparative study is carried out. The computational times of the CSO, PSO, DE, and GA models are given in Table 2. As can be seen from the table, the DE is computationally most efficient followed by PSO, CSO, and GA (Table 2). In this analysis, we used a computer with a Core i7 processor, a speed of 3.1 GHz and 8 GB RAM.

GROUNDWATER INVERSE MODEL DEVELOPMENT

Mathematically, a groundwater inverse model for aquifer parameter estimation can be expressed as (Sun 1994):

$$Min\ (J) = \sum_{L=1}^{L}\sum_{t=t_o}^{t_f} w\ \ (h_{sim}(P) - h_{obs})^2 + \sum_{L=1}^{L}\sum_{t=t_o}^{t_f} W\ \ (C_{sim}(P) - C_{obs})^2 \quad (20)$$

subject to:

$$T_{i,min} \le T_i \le T_{i,max} \quad (20a)$$

$$\alpha_{L,min} \le \alpha_{L_i} \le \alpha_{L,max} \quad (20b)$$

$$\alpha_{T,min} \le \alpha_{T_i} \le \alpha_{T,max} \quad (20c)$$

where J is the objective function, h_{obs} is observed hydraulic head; $h_{sim}(P)$ is the simulated hydraulic head; C_{obs} is observed contaminant concentration; $C_{sim}(P)$ is simulated contaminant concentration; L is the total number of monitoring wells; t_o and t_f are beginning and ending time of observations. T_i is transmissivity at node i; w and W are the weighting coefficients. The first term of the composite objective function is the summation of weighted squared differences of observed and simulated hydraulic head, and the second term is the weighted squared differences of observed and simulated contaminant concentrations.

Initially, the MFree-based RPCM model is developed to simulate groundwater flow and contaminant transport processes. By coupling the RPCM model separately with GE, DA, CSO, and PSO, four inverse models are developed for aquifer parameter estimation. The flowchart for the RPCM-DE model is shown in Figure 2. The flowchart of RPCM-CSO and RPCM-PSO models can be found in Thomas *et al.* (2018). The coupled RPCM-GA, RPCM-DE, RPCM-CSO, and RPCM-PSO models are applied to estimate the aquifer parameters of a hypothetical confined aquifer by minimizing Equation (20).

MODEL APPLICATION AND ANALYSIS

The RPCM-GA, RPCM-CSO, RPCM-DE, and RPCM-PSO models are applied to estimate the aquifer parameters of a hypothetical confined aquifer with an area of 36 sq. km (Rastogi & Huggi 2009). The aquifer domain is shown in Figure 3. The southern boundary is a constant head boundary of 100 m. The western boundary is a constant flux boundary with an inflow rate of 0.25 m^2/day. Both the northern and eastern boundaries are no-flow boundaries. Two aquitard recharge zones exist at the north with recharge at a rate of 0.00015 m/day and 0.00025 m/day, respectively. Also total dissolved solids (TDS) concentration in the two recharge zones are 80 ppm and 100 ppm, respectively. Table 3 gives the true values of transmissivity, longitudinal dispersivity, and transverse dispersivity values of the nine-zoned aquifer. The study area has a recharge well with a flow rate of 500 m^3/d. TDS concentration of injected water in the recharge well is 1,000 ppm. There is also one pumping well with a flow rate of 1,200 m^3/d. As shown in Figure 4, 18 observations nodes (wells) are chosen within the flow region. The storage coefficient of the aquifer is assumed as 0.001. Transmissivities vary from 5 to 150 m^2/day for the nine zones. The flow and transport processes of the aquifer are simulated using RPCM by scattering 49 field nodes in the aquifer domain (Figure 4).

A sensitivity analysis with three different q values mentioned in Liu & Gu (2005) is conducted to know the best parameter suited for the problem attempted with values of 0.5, 0.98, and 1.03. It has been observed that for the particular problem, $q = 0.98$ gives better results (Tables 4 and 5). It is also observed that all three values do not significantly affect the accuracy of the solution. Mean absolute percentage error (MAPE) of hydraulic head and contaminant concentration estimated with different q values are given

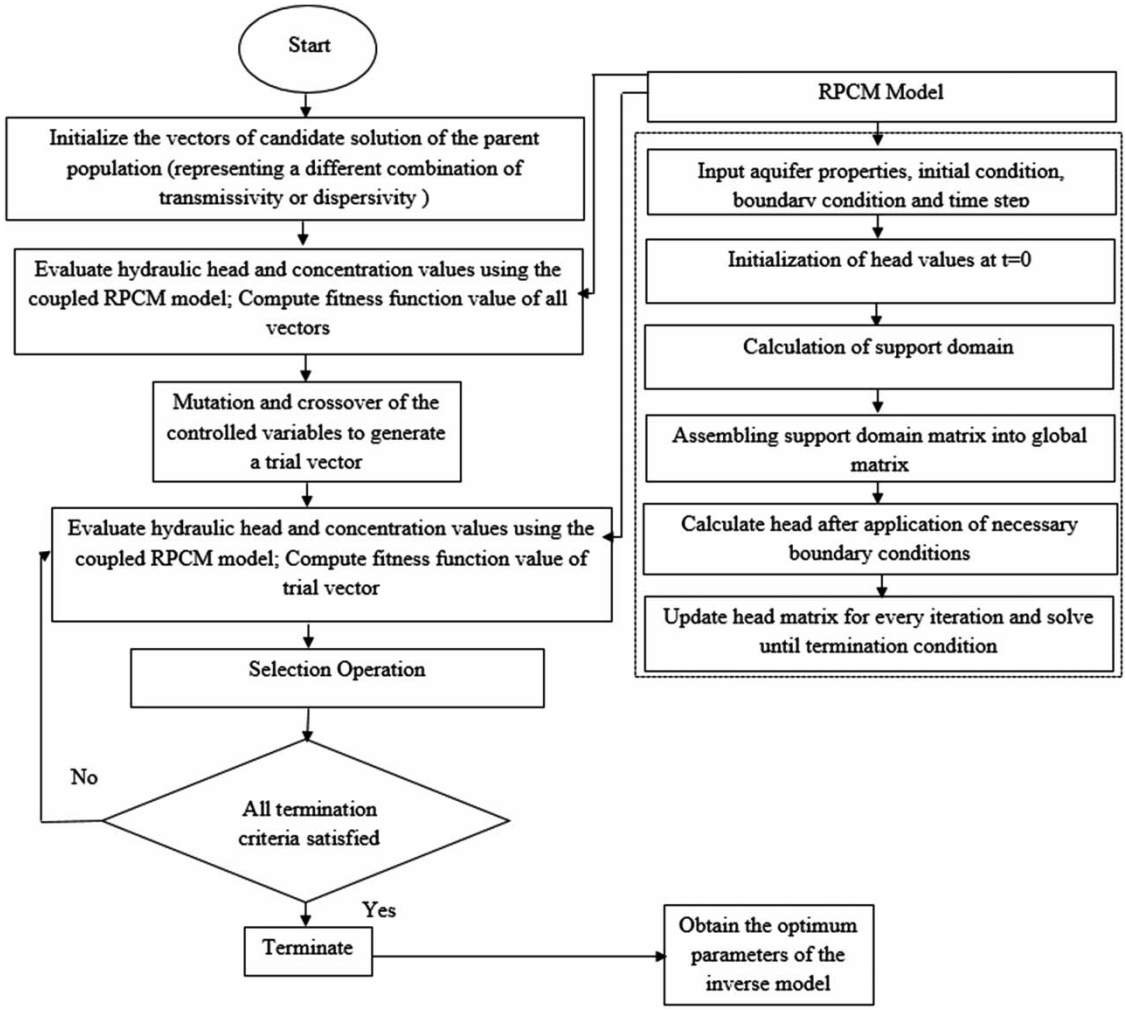

Figure 2 | Simulation optimization flowchart using RPCM-DE.

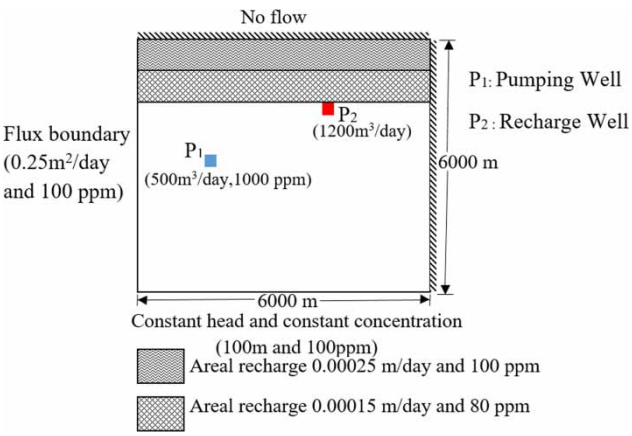

Figure 3 | Aquifer domain and boundary conditions.

Table 3 | True aquifer parameters for various zones

Zone	Transmissivity value (m^2/d)	Longitudinal dispersivity (m)	Transverse dispersivity (m)
1	150	60	6
2	150	60	6
3	50	40	4
4	150	60	6
5	50	40	4
6	15	15	1.5
7	50	40	4
8	15	15	1.5
9	5	10	1

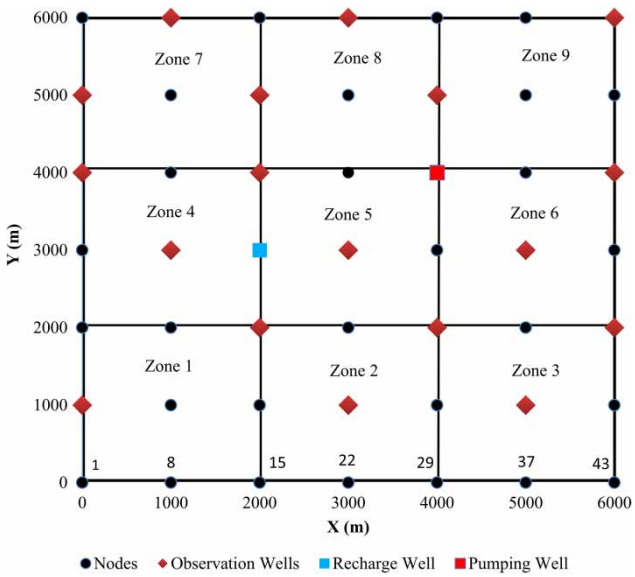

Figure 4 | Nodal distribution and zonation pattern of the aquifer domain.

in Table 6. Since MAPE of both estimated hydraulic head and concentration was lower with the shape parameter $q = 0.98$, the present study is carried out with q as 0.98.

The other shape parameter of the RBF is α_c. The solution accuracy is found to depend on α_c as it has a profound impact on the quality of the interpolation within a support domain. Although the choice of the optimum value of this parameter is crucial, there is no established method for finding the optimum value of this parameter. In this study, it is observed that for $\alpha_c = 3$, the result obtained by RPCM is in close agreement with FEM (Rastogi & Huggi 2009).

The time step considered for the flow and transport simulation is considered to be 1 day. The groundwater head and concentration values obtained using the RPCM model for the known transmissivity and dispersivity values for the flow and transport problem are taken as observed head and concentration in inverse modeling. The flow and transport simulations are carried out by RPCM for a total duration of 1,000 days. The head and concentration plots after 1,000 days are plotted in Figures 5 and 6. Comparison of hydraulic head and concentration distribution obtained after 1,000 days using RPCM and FEM are shown in Table 5. As can be seen from Table 5, the hydraulic head

Table 4 | Hydraulic head (in meter) obtained by FEM and RPCM for different q values

Node number	Head (FEM))	Head RPCM ($q = 0.5$)	Absolute % difference	Head RPCM ($q = 0.98$)	Absolute % difference	Head RPCM ($q = 1.03$)	Absolute % difference
2	103.16	101.73	1.386	102.17	0.960	102.13	0.998
23	100.30	99.84	0.459	100.56	0.259	100.27	0.030
37	99.30	98.69	0.614	99.89	0.594	99.76	0.463
17	102.47	101.31	1.132	103.47	0.976	103.29	0.800
31	98.92	94.35	4.620	95.40	3.558	95.27	3.690
45	98.15	99.17	1.039	99.26	1.131	99.68	1.559
11	105.35	100.31	4.784	100.01	5.069	100.75	4.366
25	98.15	97.16	1.009	98.75	0.611	98.86	0.723
39	93.62	99.21	5.971	98.03	4.711	97.91	4.582
5	108.34	106.85	1.375	107.22	1.034	107.08	1.163
19	104.04	108.67	4.450	107.56	3.383	108.28	4.075
47	91.26	91.42	0.175	91.65	0.427	92.49	1.348
6	111.65	110.21	1.290	110.84	0.725	111.75	0.090
20	105.16	107.55	2.273	107.92	2.625	108.12	2.815
34	91.18	93.47	2.512	93.26	2.281	93.89	2.972
14	108.49	103.17	4.904	103.42	4.673	103.71	4.406
28	98.69	97.71	0.993	98.50	0.193	99.16	0.476
49	91.74	91.58	0.174	91.83	0.098	92.04	0.327

Table 5 | Contaminant concentration (in ppm) obtained by FEM and RPCM for different q values

Node number	Concentration (FEM)	Concentration RPCM ($q = 0.5$)	Absolute % difference	Concentration RPCM ($q = 0.98$)	Absolute % difference	Concentration RPCM ($q = 1.03$)	Absolute % difference
2	101.55	105.94	4.323	105.42	3.811	105.73	4.116
23	171.99	177.67	3.303	177.44	3.169	177.49	3.198
37	101.11	104.64	3.491	104.42	3.274	104.48	3.333
17	456.00	448.97	1.542	448.94	1.548	448.92	1.553
31	197.00	193.91	1.569	193.74	1.655	193.83	1.609
45	100.07	104.97	4.897	104.56	4.487	104.84	4.767
11	129.76	132.84	2.374	132.47	2.088	132.61	2.196
25	833.44	838.24	0.576	837.65	0.505	837.98	0.545
39	88.13	90.97	3.223	90.43	2.610	90.53	2.723
5	114.58	118.15	3.116	117.47	2.522	117.88	2.880
19	343.92	351.90	2.320	351.54	2.216	351.54	2.216
47	101.39	102.94	1.529	102.47	1.065	102.62	1.213
6	113.55	117.76	3.708	117.24	3.250	117.53	3.505
20	66.43	68.94	3.778	68.67	3.372	68.81	3.583
34	75.17	79.27	5.454	78.16	3.978	78.11	3.911
14	104.26	108.42	3.990	107.55	3.156	108.16	3.741
28	102.06	106.34	4.194	105.49	3.361	106.04	3.900
49	100.00	100.00	0.000	100.00	0.000	100.00	0.000

Table 6 | Mean absolute percentage error of head and concentration estimation with different q values using RPCM

Shape parameter (q) value	Mean absolute % error for head	Mean absolute % error for concentration
0.5	2.176	2.966
0.98	1.850	2.559
1.03	1.938	2.722

and contaminant concentration obtained by both the methods are in good agreement.

Further, the inverse models, namely, RPCM-GA, RPCM-DE, RPCM-CSO, and RPCM-PSO are applied to the hypothetical confined aquifer. All four models can predict the aquifer parameters quite accurately (Tables 7–9). However, if we look closely, it can be found that DE, CSO, and PSO are more efficient in predicting aquifer parameters than GA.

Additionally, to investigate the model's performance with poor quality of observed data, noises are added to hydraulic head and contaminant concentration. Data set 2 is generated by assuming normally distributed error with zero mean and unit variance. The mean percentage error in the estimation of transmissivity, longitudinal dispersivity, and transverse dispersivity values obtained by CSO, PSO, DE, and GA are tabulated in Table 10. It can be seen that for the noisy data set (data set 2), mean percentage error is lowest for CSO. Hence, the finding is that with the noisy data set, estimated aquifer parameters by CSO are closer to the actual values followed by DE, PSO, and GA.

We have generated bar diagrams to check errors in estimated transmissivity, longitudinal dispersivity, and transverse dispersivity values for data set 2. It can be seen from Figures 7–9 that CSO shows the least error in estimating transmissivity, longitudinal dispersivity, and transverse dispersivity values compared to DE, PSO, and GA.

RELIABILITY ANALYSIS

The estimated transmissivity, longitudinal dispersivity, and transverse dispersivity values are tested for reliability analysis. Reliability analysis is ascertained by evaluating the

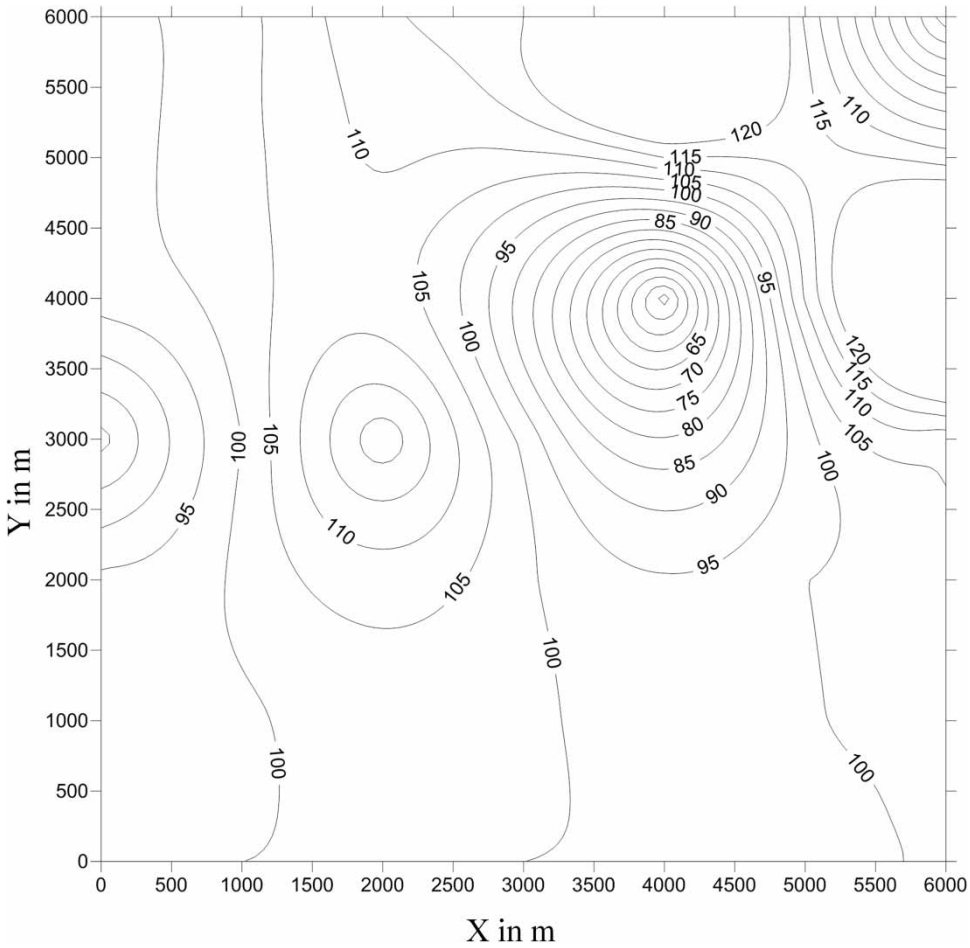

Figure 5 | The contour of hydraulic head (m) after 1,000 days.

coefficient of variation (CV) and variance-covariance matrix (VCM).

Coefficient of variation

The CV is a dimensionless parameter which is equal to the ratio of the standard deviation of the estimated parameter. Lower value of CV indicates a better estimate of the zonal parameter, i.e., better reliability. The CV can be estimated as (Kovar & Van der Heijde 1996):

$$\mathrm{CV} = \left\{ \frac{(\text{Variance of estimated parameters T})^{0.5}}{\text{Estimated parameter T}} \right\} \tag{21}$$

Table 11 shows the CV values, which are estimated for all four models with noisy data (data set 2). Lower CV values of RPCM-CSO indicate better reliability of RPCM-CSO in comparison to RPCM-DE, RPCM-PSO, and RPCM-GA.

Variance-covariance matrix

In the VCM, the diagonal elements represent the individual parameter variances, whereas the off-diagonal elements represent the correlation between the parameters. It is used as a measure of the reliability of the estimated parameters. If a problem has only three zones the VCM can be expressed as (Bard 1974):

$$M = \begin{bmatrix} (T_1 - T_{1e})^2 & (T_1 - T_{1e})(T_2 - T_{2e}) & (T_1 - T_{1e})(T_3 - T_{3e}) \\ (T_2 - T_{2e})(T_1 - T_{1e}) & (T_2 - T_{2e})^2 & (T_2 - T_{2e})(T_3 - T_{3e}) \\ (T_3 - T_{3e})(T_1 - T_{1e}) & (T_3 - T_{3e})(T_2 - T_{2e}) & (T_3 - T_{3e})^2 \end{bmatrix} \tag{22}$$

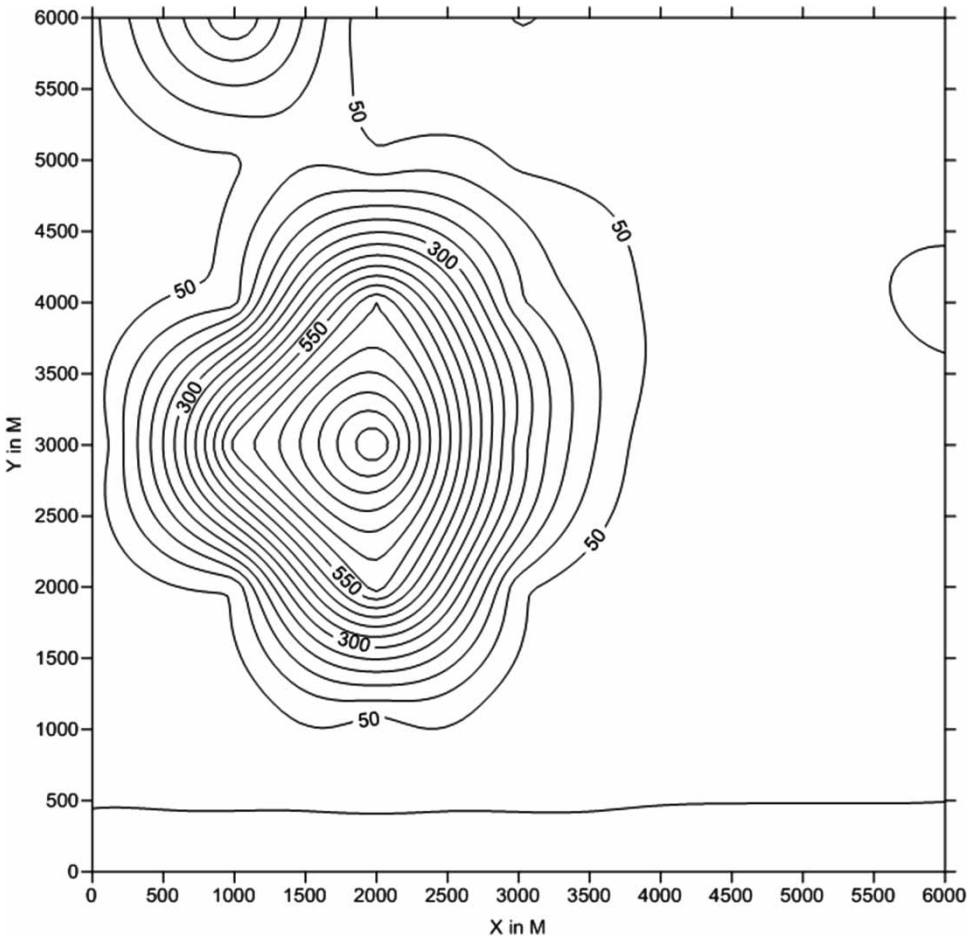

Figure 6 | The contour of contaminant concentration (ppm) after 1,000 days.

where T_1, T_2, and T_3 are the true parameters and T_{1e}, T_{2e}, and T_{3e} are the estimated parameters. The confined aquifer problem of this study has nine zones. Hence, the VCM will be 9×9 matrices. Here, the VCM for transverse dispersivity is computed for data set 2 and the results are shown in Tables 12–15. There are larger magnitude terms in the VCM of RPCM-PSO, RPCM-DE, and RPCM-GA in comparison with RPCM-CSO. The findings suggest that the RPCM-CSO model has better reliability than RPCM-PSO, RPCM-DE, and RPCM-GA.

DISCUSSION

This study is an attempt to check the effectiveness of four important optimization algorithms (GA, DE, CSO, PSO) in estimating aquifer parameters (transmissivity, longitudinal dispersivity, and transverse dispersivity) of a hypothetical confined aquifer. It has been observed that, for non-noisy observed data (hydraulic head and contaminant concentration), DE, CSO, and PSO are more accurate in estimating aquifer parameters than GA. However, for noisy observed data, CSO outperforms the DE, PSO, and GA. This is an important finding, in the sense that in real field scenarios, the observed data are usually erroneous, and in such cases use of CSO for inverse modeling will be a better choice than GA, DE, and PSO. However, the computational time of CSO-based optimization models is greater compared to DE and PSO. In the seeking mode of CSO, each cat creates its multiple copies in the search space. Hence, in CSO for every iteration, the number of fitness function evolutions are greater in comparison with

Table 7 | Comparison of estimated transmissivity (m^2/day) values in various zones using DE, GA, PSO, and CSO

Data set	Zone	True values (m^2/d)	Estimated values using CSO (m^2/d)	Absolute % error CSO	Estimated values using PSO (m^2/d)	Absolute % error PSO	Estimated values using DE (m^2/d)	Absolute % error DE	Estimated values using GA (m^2/d)	Absolute % error GA
Data set 1	1	150	149.99	0.01	149.99	0.01	149.97	0.02	149.23	0.51
	2	150	149.98	0.01	149.98	0.01	149.98	0.01	149.37	0.42
	3	50	49.98	0.03	49.90	0.20	49.96	0.08	49.34	1.32
	4	150	149.97	0.02	149.89	0.07	149.95	0.03	149.85	0.10
	5	50	49.97	0.06	50.10	0.20	49.96	0.08	49.95	0.10
	6	15	14.99	0.09	14.97	0.19	14.96	0.27	14.92	0.53
	7	50	49.98	0.04	49.98	0.04	49.97	0.06	49.48	1.04
	8	15	14.98	0.16	14.97	0.18	14.97	0.20	14.64	2.40
	9	5	4.99	0.20	4.99	0.28	4.94	1.20	4.73	5.40
Data set 2: $\mu = 0$, $\sigma = 1$	1	150	148.93	0.71	148.79	0.81	149.14	0.57	148.712	0.86
	2	150	149.48	0.35	148.94	0.71	148.92	0.72	149.03	0.65
	3	50	49.16	1.68	49.22	1.56	49.24	1.52	49.31	1.38
	4	150	149.07	0.62	148.91	0.73	148.74	0.84	148.36	1.09
	5	50	49.13	1.74	49.09	1.82	49.26	1.48	49.18	1.64
	6	15	14.43	3.80	14.20	5.33	14.61	2.60	13.975	6.83
	7	50	48.98	2.04	49.18	1.64	48.52	2.96	49.27	1.46
	8	15	14.66	2.27	14.37	4.20	14.67	2.20	14.16	5.60
	9	5	4.72	5.60	4.49	10.20	4.48	10.40	4.37	12.60

Table 8 | Comparison of estimated longitudinal dispersivity (m) values in various zones using DE, GA, PSO, and CSO

Data set	Zone	True values (m)	Estimated values using CSO (m)	Absolute % error CSO	Estimated values using PSO (m)	Absolute % error PSO	Estimated values using DE (m)	Absolute % error DE	Estimated values using GA (m)	Absolute % error GA
Data set 1	1	60	59.98	0.03	59.96	0.06	59.98	0.03	59.67	0.55
	2	60	59.96	0.07	59.93	0.12	59.96	0.07	59.18	1.37
	3	40	39.98	0.05	39.98	0.05	39.99	0.02	39.43	1.43
	4	60	59.99	0.02	59.96	0.07	59.98	0.03	59.50	0.83
	5	40	39.99	0.02	39.96	0.10	39.95	0.12	39.64	0.90
	6	15	14.98	0.13	14.84	1.07	14.99	0.07	14.63	2.47
	7	40	39.99	0.03	40	0.00	39.96	0.10	39.73	0.68
	8	15	15	0.00	14.97	0.20	14.94	0.40	14.87	0.87
	9	10	9.97	0.30	10	0.00	9.96	0.40	9.72	2.80
Data set 2: $\mu = 0$, $\sigma = 1$	1	60	59.19	1.35	59.24	1.27	60.26	0.43	59.15	1.42
	2	60	59.64	0.60	59.43	0.95	59.53	0.78	59.53	0.78
	3	40	39.04	2.40	39.58	1.05	39.37	1.58	38.46	3.85
	4	60	59.72	0.47	59.29	1.18	59.79	0.35	59.06	1.57
	5	40	39.64	0.90	39.54	1.15	40.37	0.92	39.42	1.45
	6	15	14.72	1.87	14.62	2.53	14.66	2.27	14.37	4.20
	7	40	39.68	0.80	39.57	1.08	39.61	0.98	40.36	0.90
	8	15	14.45	3.67	14.39	4.07	14.5	3.33	14.22	5.20
	9	10	9.54	4.60	9.47	5.30	9.42	5.80	9.28	7.20

other optimization algorithms. However, the computational speed of CSO can be drastically reduced by using high-performance computing systems, a graphics processing unit, vectorizing the codes, parallel computing, and using surrogate simulators. The RPCM model can be used to train surrogate models such as artificial neural network, deep learning, radial basis function, support vector machine, kriging, etc. The SO model developed by coupling a

Table 9 | Comparison of estimated transverse dispersivity (m) values in various zones using DE, GA, PSO, and CSO

Data set	Zone	True values (m)	Estimated values using CSO (m)	Absolute % error CSO	Estimated values using PSO (m)	Absolute % error PSO	Estimated values using DE (m)	Absolute % error DE	Estimated values using GA (m))	Absolute % error GA
Data set 1	1	6	5.98	0.25	5.99	0.02	5.98	0.33	5.81	3.17
	2	6	5.91	1.50	5.99	0.03	5.96	0.67	5.72	4.67
	3	4	3.93	1.75	4	0.00	3.94	1.50	3.82	4.50
	4	6	5.98	0.33	6	0.00	5.99	0.17	5.92	1.33
	5	4	3.99	0.25	3.99	0.02	3.98	0.50	3.94	1.50
	6	1.5	1.48	1.33	1.5	0.00	1.49	0.67	1.46	2.67
	7	4	4	0.00	4.01	0.25	3.98	0.50	3.91	2.25
	8	1.5	1.5	0.00	1.479	1.40	1.5	0.00	1.48	1.33
	9	1	1	0.00	1	0.00	0.98	2.00	0.98	2.00
Data set 2: $\mu = 0$, $\sigma = 1$	1	6	5.73	4.50	5.67	5.50	5.66	5.67	5.58	7.00
	2	6	5.61	6.50	5.58	7.00	5.7	5.00	5.59	6.83
	3	4	3.70	7.50	3.63	9.25	3.65	8.75	3.53	11.75
	4	6	5.61	6.50	5.48	8.67	5.63	6.17	5.48	8.67
	5	4	3.77	5.75	3.57	10.75	3.69	7.75	3.5	12.50
	6	1.5	1.42	5.33	1.40	6.67	1.38	8.00	1.29	14.00
	7	4	3.84	4.00	3.68	8.00	3.77	5.75	3.62	9.50
	8	1.5	1.47	2.00	1.39	7.33	1.43	4.67	1.36	9.33
	9	1	0.95	5.00	0.93	7.00	0.96	4.00	0.93	7.00

Table 10 | Comparison of mean absolute percentage error for CSO, PSO, DE, and GA

	Without noise			With noise		
Data	Mean absolute % error for transmissivity	Mean absolute % error for longitudinal dispersivity	Mean absolute % error for transverse dispersivity	Mean absolute % error for transmissivity	Mean absolute % error for longitudinal dispersivity	Mean absolute % error for transverse dispersivity
CSO	0.07	0.07	0.60	2.09	1.85	5.23
PSO	0.13	0.19	0.19	3.00	2.06	7.80
DE	0.22	0.14	0.70	2.59	1.85	6.19
GA	1.31	1.32	2.60	3.57	2.95	9.62

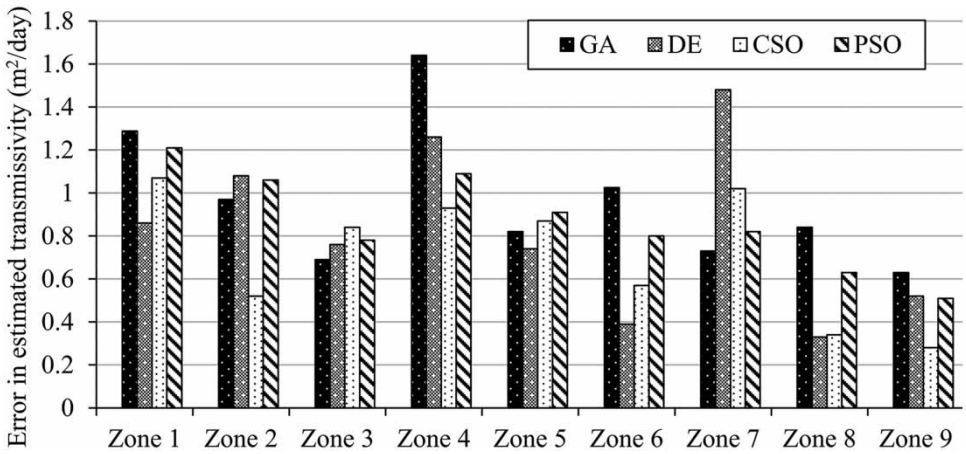

Figure 7 | Comparison of error in estimated transmissivity (m^2/day) for noisy data using inverse models.

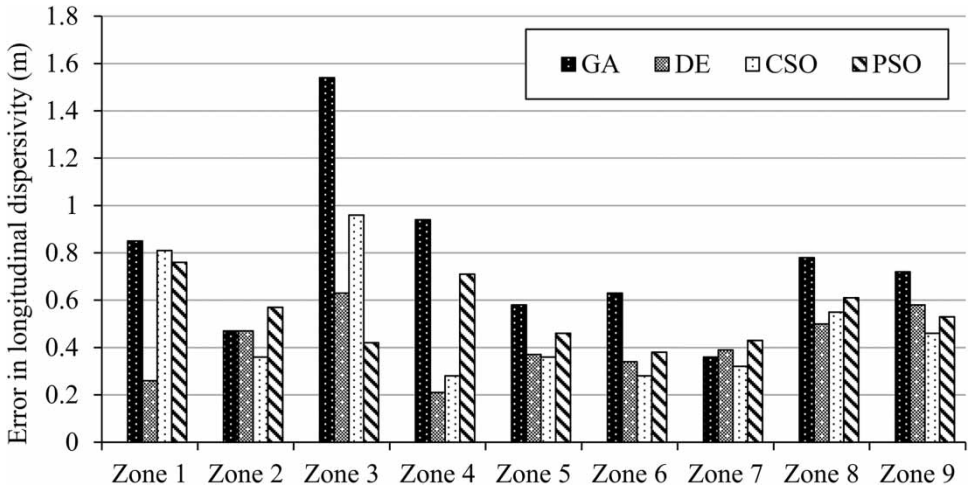

Figure 8 | Comparison of error in estimated longitudinal dispersivity (m) values for noisy data using inverse models.

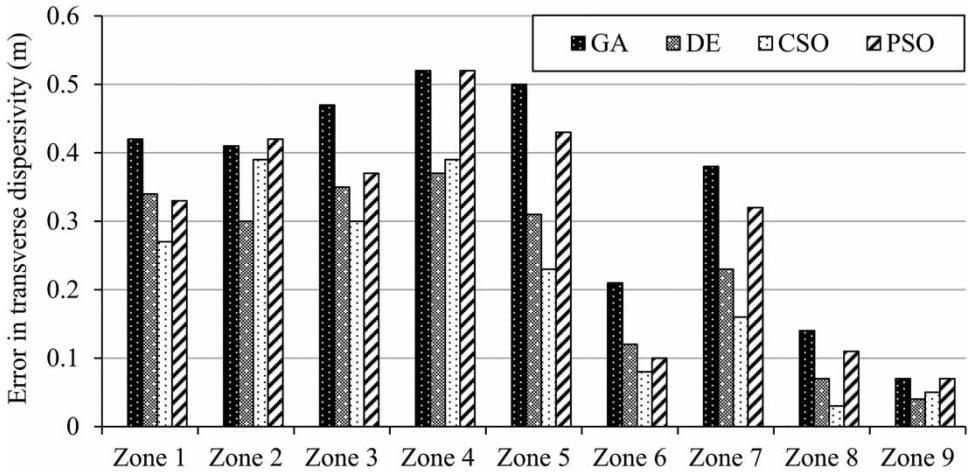

Figure 9 | Comparison of error in estimated transverse dispersivity (m) values for noisy data using inverse models.

surrogate model with CSO will be computationally more efficient than RPCM-CSO.

CONCLUSIONS

In this study, a mesh-free RPCM model is developed to simulate the groundwater flow and contaminant transport processes of a confined aquifer. The hydraulic head and contaminant concentration obtained by the RPCM model are found to be in good agreement with the solution of FEM. Further, four optimization models are developed based on GA, DE, PSO, and CSO. By coupling the RPCM model separately with GA, DE, PSO, and CSO, the S-O models are developed. The S-O models are used to compute transmissivity, longitudinal dispersivity, and transverse dispersivity by minimizing a composite objective function. The composite objective is a summation of sum of the weighted squared differences of observed and simulated hydraulic head and sum of the weighted squared differences of observed and simulated contaminant concentrations at the monitoring wells. The S-O model is applied to compute aquifer parameters of a hypothetical confined aquifer with nine zones. It is observed that RPCM-CSO, RPCM-DE, and RPCM-PSO models are more accurate in estimating aquifer parameters than RPCM-GA. The performance of

Table 11 | Comparison of CV for the estimated transmissivity, longitudinal dispersivity, and transverse dispersivity in various zones for data set 4 using RPCM-GA, RPCM-DE, and RPCM-CSO with the RPCM-PSO model

	Transmissivity				Longitudinal dispersivity				Transverse dispersivity			
Zone no.	CV-RPCM-CSO	CV-RPCM-PSO	CV-RPCM-DE	CV-RPCM GA	CV-RPCM-CSO	CV-RPCM-PSO	CV-RPCM-DE	CV-RPCM-GA	CV-RPCM-CSO	CV-RPCM-PSO	CV-RPCM-DE	CV-RPCM-GA
Zone 1	0.0072	0.0081	0.0058	0.0087	0.0137	0.0128	0.0043	0.0144	0.0471	0.0582	0.0601	0.0753
Zone 2	0.0035	0.0071	0.0073	0.0065	0.0060	0.0096	0.0079	0.0079	0.0695	0.0753	0.0526	0.0733
Zone 3	0.0171	0.0158	0.0154	0.0140	0.0246	0.0106	0.0160	0.0400	0.0811	0.1019	0.0959	0.1331
Zone 4	0.0062	0.0073	0.0085	0.0111	0.0047	0.0120	0.0035	0.0159	0.0695	0.0949	0.0657	0.0949
Zone 5	0.0177	0.0185	0.0150	0.0167	0.0091	0.0116	0.0092	0.0147	0.0610	0.1204	0.0840	0.1429
Zone 6	0.0395	0.0563	0.0267	0.0733	0.0190	0.0260	0.0232	0.0438	0.0563	0.0714	0.0870	0.1628
Zone 7	0.0208	0.0167	0.0305	0.0148	0.0081	0.0109	0.0098	0.0089	0.0417	0.0870	0.0610	0.1050
Zone 8	0.0232	0.0438	0.0225	0.0593	0.0381	0.0424	0.0345	0.0549	0.0204	0.0791	0.0490	0.1029
Zone 9	0.0593	0.1136	0.1161	0.1442	0.0482	0.0559	0.0616	0.0776	0.0526	0.0753	0.0417	0.0753

Table 12 | VCM for noisy data using the RPCM-GA model for the estimated transverse dispersivity in various zones

	1	2	3	4	5	6	7	8	9
1	0.1764	0.1722	0.1974	0.2184	0.21	0.0882	0.1596	0.0588	0.0294
2	0.1722	0.1681	0.1927	0.2132	0.205	0.0861	0.1558	0.0574	0.0287
3	0.1974	0.1927	0.2209	0.2444	0.235	0.0987	0.1786	0.0658	0.0329
4	0.2184	0.2132	0.2444	0.2704	0.26	0.1092	0.1976	0.0728	0.0364
5	0.21	0.205	0.235	0.26	0.25	0.105	0.19	0.07	0.035
6	0.0882	0.0861	0.0987	0.1092	0.105	0.0441	0.0798	0.0294	0.0147
7	0.1596	0.1558	0.1786	0.1976	0.19	0.0798	0.1444	0.0532	0.0266
8	0.0588	0.0574	0.0658	0.0728	0.07	0.0294	0.0532	0.0196	0.0098
9	0.0294	0.0287	0.0329	0.0364	0.035	0.0147	0.0266	0.0098	0.0049

Table 13 | VCM for noisy data using the RPCM-DE model for the estimated transverse dispersivity in various zones

	1	2	3	4	5	6	7	8	9
1	0.1156	0.102	0.119	0.1258	0.1054	0.0408	0.0782	0.0238	0.0136
2	0.102	0.09	0.105	0.111	0.093	0.036	0.069	0.021	0.012
3	0.119	0.105	0.1225	0.1295	0.1085	0.042	0.0805	0.0245	0.014
4	0.1258	0.111	0.1295	0.1369	0.1147	0.0444	0.0851	0.0259	0.0148
5	0.1054	0.093	0.1085	0.1147	0.0961	0.0372	0.0713	0.0217	0.0124
6	0.0408	0.036	0.042	0.0444	0.0372	0.0144	0.0276	0.0084	0.0048
7	0.0782	0.069	0.0805	0.0851	0.0713	0.0276	0.0529	0.0161	0.0092
8	0.0238	0.021	0.0245	0.0259	0.0217	0.0084	0.0161	0.0049	0.0028
9	0.0136	0.012	0.014	0.0148	0.0124	0.0048	0.0092	0.0028	0.0016

Table 14 | VCM for noisy data using the RPCM-CSO model for the estimated transverse dispersivity in various zones

	1	2	3	4	5	6	7	8	9
1	0.09	0.105	0.096	0.135	0.093	0.039	0.084	0.018	0.021
2	0.105	0.1225	0.112	0.1575	0.1085	0.0455	0.098	0.021	0.0245
3	0.096	0.112	0.1024	0.144	0.0992	0.0416	0.0896	0.0192	0.0224
4	0.135	0.1575	0.144	0.2025	0.1395	0.0585	0.126	0.027	0.0315
5	0.093	0.1085	0.0992	0.1395	0.0961	0.0403	0.0868	0.0186	0.0217
6	0.039	0.0455	0.0416	0.0585	0.0403	0.0169	0.0364	0.0078	0.0091
7	0.084	0.098	0.0896	0.126	0.0868	0.0364	0.0784	0.0168	0.0196
8	0.018	0.021	0.0192	0.027	0.0186	0.0078	0.0168	0.0036	0.0042
9	0.021	0.0245	0.0224	0.0315	0.0217	0.0091	0.0196	0.0042	0.0049

Table 15 | VCM for noisy data using the RPCM-PSO model for the estimated transverse dispersivity in various zones

	1	2	3	4	5	6	7	8	9
1	0.1563	0.18235	0.16672	0.23445	0.16151	0.06773	0.14588	0.03126	0.03647
2	0.1461	0.17045	0.15584	0.21915	0.15097	0.06331	0.13636	0.02922	0.03409
3	0.1116	0.1302	0.11904	0.1674	0.11532	0.04836	0.10416	0.02232	0.02604
4	0.1743	0.20335	0.18592	0.26145	0.18011	0.07553	0.16268	0.03486	0.04067
5	0.1296	0.1512	0.13824	0.1944	0.13392	0.05616	0.12096	0.02592	0.03024
6	0.0543	0.06335	0.05792	0.08145	0.05611	0.02353	0.05068	0.01086	0.01267
7	0.1206	0.1407	0.12864	0.1809	0.12462	0.05226	0.11256	0.02412	0.02814
8	0.0321	0.03745	0.03424	0.04815	0.03317	0.01391	0.02996	0.00642	0.00749
9	0.027	0.0315	0.0288	0.0405	0.0279	0.0117	0.0252	0.0054	0.0063

RPCM-CSO and RPCM-DE are comparable and it is difficult to ascertain the superiority of one over the other.

Further, high measurement errors (noises) are introduced to the observed hydraulic head and contaminant concentrations' values to check the performance of the S-O model in estimating aquifer parameters when observed data sets are noisy. Here, it is observed that the average percentage errors in estimating aquifer parameters by the RPCM-CSO model are less than those obtained by RPCM-GA, RPCM-PSO, and RPCM-DE. For the noisy observed data set, the reliability analysis also suggests the superiority of the RPCM-CSO model over RPCM-GA, RPCM-DE, and RPCM-PSO. It is evident from the results that CSO is more accurate in estimating aquifer parameters than GA, DE, and PSO with the noisy observed data set. The RPCM-CSO model can be effectively used for estimating aquifer parameters or contaminant source identification using inverse modeling for aquifers with more complex hydrogeological features.

REFERENCES

Abbaspour, K. C., Schulin, R. & Van Genuchten, M. T. 2001 Estimating unsaturated soil hydraulic parameters using ant colony optimization. *Advances in Water Resources* **24** (8), 827–841.

Bard, Y. 1974 *Nonlinear Parameter Estimation*. Academic Press, Orlando, FL, USA, p. 341.

Bear, J. & Cheng, A. H. D. 2010 *Modeling groundwater flow and contaminant transport*, Vol. 23. Springer Science & Business Media, Dordrecht, The Netherlands.

Bear, J. 2012 *Hydraulics of Groundwater*. Courier Corporation, Chelmsford, MA, USA.

Carrera, J. & Neuman, S. P. 1986a Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood method incorporating prior

information. *Water Resources Research* **22** (2), 199–210. 10.1029/WR022i002p00199.

Carrera, J. & Neuman, S. P. 1986b Estimation of aquifer parameters under transient and steady state conditions: 3. Application to synthetic and field data. *Water Resources Research* **22** (2), 228–242.

Carrera, J., Alcolea, A., Medina, A., Hidalgo, J. & Slooten, L. J. 2005 Inverse problem in hydrogeology. *Hydrogeology Journal* **13** (1), 206–222.

Chu, S. C. & Tsai, P. W. 2007 Computational intelligence based on the behaviour of cats. *International Journal of Innovative Computing, Information and Control* **3** (1), 163–173.

Chu, S. C., Tsai, P. W. & Pan, J. S. 2006 Cat swarm optimization. In: *PRICAI 2006: Trends in Artificial Intelligence*. Springer, Berlin, Heidelberg, Germany, pp. 854–858.

Deb, K. 2012 *Optimization for Engineering Design: Algorithms and Examples*. PHI Learning Pvt. Ltd, New Delhi, India.

Eberhart, R. & Kennedy, J. 1995 A new optimizer using particle swarm theory. In: *MHS'95, Proceedings of the Sixth International Symposium Micro Machine and Human Science*, 4–6 October, Nagoya, Japan, pp. 39–43.

Elçi, A. & Ayvaz, M. T. 2014 Differential-evolution algorithm based optimization for the site selection of groundwater production wells with the consideration of the vulnerability concept. *Journal of Hydrology* **511**, 736–749.

Fitts, C. R. 2002 *Groundwater Science*. Elsevier, Amsterdam, The Netherlands.

Fonna, S., Huzni, S., Ridha, M. & Ariffin, A. K. 2013 Inverse analysis using particle swarm optimization for detecting corrosion profile of rebar in concrete structure. *Engineering Analysis with Boundary Elements* **37** (3), 585–593.

Franssen, H. H., Alcolea, A., Riva, M., Bakr, M., Van der Wiel, N., Stauffer, F. & Guadagnini, A. 2009 A comparison of seven methods for the inverse modelling of groundwater flow. Application to the characterisation of well catchments. *Advances in Water Resources* **32** (6), 851–872.

Freeze, R. A. & Cherry, J. A. 1979 *Groundwater*. Prentice-Hall, Englewood Cliffs, NJ, USA, p. 604.

Guneshwor, L., Eldho, T. I. & Kumar, A. V. 2018 Identification of groundwater contamination sources using meshfree RPCM simulation and particle swarm optimization. *Water Resources Management* **32** (4), 1517–1538.

Guo, L., Meng, Z., Sun, Y. & Wang, L. 2016 Parameter identification and sensitivity analysis of solar cell models with cat swarm optimization algorithm. *Energy Conversion and Management* **108**, 520–528.

Gurarslan, G. & Karahan, H. 2015 Solving inverse problems of groundwater-pollution-source identification using a differential evolution algorithm. *Hydrogeology Journal* **23** (6), 1109–1119.

Huggi, V. P. & Rastogi, A. K. 2003 Estimation of solute transport parameters of groundwater systems using genetic algorithm. *Water and Energy International* **60** (4), 38–45.

Jin, X., Mahinthakumar, G. K., Zechman, E. M. & Ranjithan, R. S. 2009 A genetic algorithm-based procedure for 3D source identification at the Borden emplacement site. *Journal of Hydroinformatics* **11** (1), 51–64.

Kovar, K. & Van der Heijde, P. 1996 Calibration and reliability in groundwater modelling. In: *Proceedings of the ModelCARE'96 Conference*, 24–26 September, Golden, CO, USA.

Kovářík, K. & Mužík, J. 2013 A meshless solution for two dimensional density-driven groundwater flow. *Engineering Analysis with Boundary Elements* **37** (2), 187–196.

Li, S., Liu, Y. & Yu, H. 2006 Parameter estimation approach in groundwater hydrology using hybrid ant colony system. In: *Computational Intelligence and Bioinformatics: International Conference on Intelligent Computing, ICIC 2006*, 16–19 August, 2006, Kunming, China. Proceedings, Part III, pp. 182–191.

Liu, G. R. 2002 *Mesh Free Methods: Moving Beyond the Finite Element Method*. CRC Press, Boca Raton, FL, USA.

Liu, G. R. & Gu, Y. T. 2005 *An Introduction to Meshfree Methods and Their Programming*. Springer, Dordrecht, The Netherlands.

Majumder, P. & Eldho, T. I. 2016 A new groundwater management model by coupling analytic element method and reverse particle tracking with cat swarm optimization. *Water Resources Management* **30** (6), 1953–1972.

Majumder, P. & Eldho, T. I. 2017 Vectorized simulation of groundwater flow and contaminant transport using analytic element method and random walk particle tracking. *Hydrological Processes* **31** (5), 1144–1160.

Mategaonkar, M., Eldho, T. I. & Kamat, S. 2018 In-situ bioremediation of groundwater using a meshfree model and particle swarm optimization. *Journal of Hydroinformatics* **20** (4), 886–897.

Matott, L. S., Rabideau, A. J. & Craig, J. R. 2006 Pump-and-treat optimization using analytic element method flow models. *Advances in Water Resources* **29** (5), 760–775.

Medina, A. & Carrera, J. 1996 Coupled estimation of flow and solute transport parameters. *Water Resources Research* **32** (10), 3063–3076.

Meenal, M. & Eldho, T. I. 2012 Two-dimensional contaminant transport modeling using meshfree point collocation method (PCM). *Engineering Analysis with Boundary Elements* **36** (4), 551–561.

Molga, M. & Smutnicki, C. 2005 *Test Functions for Optimization Needs*. http://new.zsd.iiar.pwr.wroc.pl/files/docs/functions.pdf.

Patel, S. & Rastogi, A. K. 2017 Meshfree multiquadric solution for real field large heterogeneous aquifer system. *Water Resources Management* **31** (9), 2869–2884.

Peralta, R. C. 2012 *Groundwater Optimization Handbook: Flow, Contaminant Transport, and Conjunctive Management*. CRC Press, Roca Baton, FL, USA.

Prasad, K. L. & Rastogi, A. K. 2001 Estimating net aquifer recharge and zonal hydraulic conductivity values for Mahi Right Bank Canal project area, India by genetic algorithm. *Journal of Hydrology* **243** (3), 149–161.

Praveen Kumar, R. & Dodagoudar, G. R. 2008 Two-dimensional modelling of contaminant transport through saturated porous media using the radial point interpolation method (RPIM). *Hydrogeology Journal* **16** (8), 1497–1505.

Rajanayaka, C. & Kulasiri, D. 2001 Investigation of a parameter estimation method for contaminant transport in aquifers. *Journal of Hydroinformatics* **3** (4), 203–213.

Rastogi, A. K. 2007 *Numerical Groundwater Hydrology*. Ulhas Phatak for Penra International (I) Pvt. Ltd, Mumbai, India.

Rastogi, A. K. & Huggi, V. P. 2009 Parameter assessment in flow through porous media. *ISH Journal of Hydraulic Engineering* **15** (Suppl. 1), 272–296.

Reed, P., Minsker, B. S. & Goldberg, D. E. 2001 A multiobjective approach to cost effective long-term groundwater monitoring using an elitist nondominated sorted genetic algorithm with historical data. *Journal of Hydroinformatics* **3** (2), 71–89.

Robinson, J. & Rahmat-Samii, Y. 2004 Particle swarm optimization in electromagnetics. *IEEE Transactions on Antennas and Propagation* **52** (2), 397–407.

Saha, S. K., Ghoshal, S. P., Kar, R. & Mandal, D. 2013 Cat swarm optimization algorithm for optimal linear phase FIR filter design. *ISA Transactions* **52** (6), 781–794.

Storn, R. & Price, K. 1995 Differential evolution – a simple and efficient adaptive scheme for global optimization over continuous spaces. ICSI Technical Report TR-95-012.

Storn, R. & Price, K. 1997 Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization* **11** (4), 341–359.

Sun, N. Z. 1994 *Inverse Problems in Groundwater Modeling. Theory and Applications of Transport in Porous Media*, Vol. 6. Springer, Dordrecht, The Netherlands, p. 338.

Sun, N. Z. 2013 *Inverse Problems in Groundwater Modeling*, Vol. 6. Springer Science & Business Media, Dordrecht, The Netherlands.

Swathi, B. & Eldho, T. I. 2018 Aquifer parameter and zonation structure estimation using meshless local Petrov–Galerkin method and particle swarm optimization. *Journal of Hydroinformatics* **20** (2), 457–467.

Thomas, A., Eldho, T. I. & Rastogi, A. K. 2014 A comparative study of point collocation-based MeshFree and finite element methods for groundwater flow simulation. *ISH Journal of Hydraulic Engineering* **20** (1), 65–74.

Thomas, A., Majumdar, P., Eldho, T. I. & Rastogi, A. K. 2018 Simulation optimization model for aquifer parameter estimation using coupled meshfree point collocation method and cat swarm optimization. *Engineering Analysis with Boundary Elements* **91**, 60–72.

Wang, X., Jardani, A. & Jourde, H. 2017 A hybrid inverse method for hydraulic tomography in fractured and karstic media. *Journal of Hydrology* **551**, 29–46.

Zimmerman, D. A., Marsily, G. D., Gotway, C. A., Marietta, M. G., Axness, C. L., Beauheim, R. L., Bras, R. L., Carrera, J., Dagan, G., Davies, P. B. & Gallegos, D. P. 1998 A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. *Water Resources Research* **34** (6), 1373–1413.

First received 25 September 2018; accepted in revised form 25 January 2019. Available online 22 February 2019