A grey-encoded hybrid accelerating genetic algorithm for parameter optimisation of environmental models

Xiaohua Yang*, Zhifeng Yang and Zhenyao Shen

School of Environment,
State Key Laboratory of Water Environment Simulation,
Beijing Normal University, Beijing 100875, China
Fax: 86 10 58800829   E-mail: y1x1h1@sohu.com
E-mail: zfyang@bnu.edu.cn   E-mail: z.y.shen@163.com
*Corresponding author

Jianqiang Li

Water Strategy Research Division,
Water Resources and Hydropower Planning and
Design General Institute, Beijing 100011, China
E-mail: ljqiang263@263.net

Abstract: A new method, grey-encoded hybrid accelerating genetic algorithm, is presented for the parameter optimisation of environmental models. With the shrinking of searching range, the method gradually directs to optimal result with the excellent individuals obtained by grey genetic algorithm embedding Nelder-Mead simplex searching operator. The convergence theorem is given for guaranteeing the global convergence of the new genetic algorithm. The global optimisation of the new genetic algorithm is analysed. Its efficiency is verified by application of ten test functions. The comparison of our algorithm with six other algorithms is presented. This algorithm overcomes any Hamming cliff phenomena in existing genetic algorithms, and it is good for the parameter optimisation for the practical O'Connor water quality model and rainfall-runoff model.

Keywords: environmental models; global optimisation; grey-encoded genetic algorithm; Nelder-Mead simplex algorithm; accelerating convergence.

Biographical notes: Xiaohua Yang is a Professor in School of Environment and State Key Laboratory of Water Environment Simulation at Beijing Normal University, China. She received her PhD in Hydrology and Water Resource System Engineering from Hohai University, China. In 2004, she finished her Postdoctoral work in Environment Science in School of Environment at Beijing Normal University, China. Her research interests include issues related to optimisation, prediction, assessment and modelling in hydrology, water resource and environment science. She has published papers in many water resource and environment system engineering journals.

Zhifeng Yang is a Professor in School of Environment and State Key Laboratory of Water Environment Simulation at Beijing Normal University, China. He received his PhD from Tsinghua University, China. He is a member of the Standing Board of Directors for the China Environment Association and the China Committee for International Water Resource Commission. He is the Vice Chair of the Department of Geosciences in the Scientific and Technology Commission of the Ministry of Education, China. His major research fields are water resources and water environment, eco-environmental planning and management and environment systems engineering. He has published papers in many water and environment science journals.

Zhenyao Shen is a Professor in School of Environment and State Key Laboratory of Water Environment Simulation at Beijing Normal University, China. He received his PhD from Tsinghua University, China. His research interests include issues related to water resource and environment assessment. He has published papers in many water resource and environment assessment journals.

Jianqiang Li is a Professor in Water Strategy Research Division, Water Resources and Hydropower Planning and Design General Institute, China. He received his Master's Degree from Hohai University, China. His research interests include issues related to water resource assessment and modelling. He has published papers in many hydrology and water resource journals.

1 Introduction

Many environmental models consist of a large number of parameters. The parameter optimisation of complicated non-linear environmental models is intractable mathematically. Once an objective function has many local extreme points, the traditional optimisation methods may not obtain the global optimisation efficiently. A genetic algorithm (GA) based on the genetic evolution of a species was proposed by Holland (1975). The detailed genetic algorithm and its implementation were given by Goldberg (1989). De Jong (1975) showed that the standard binary-encoded GA (SGA) could constitute an interesting alternative to perform the global optimisation of a function depending on several continuous variables (Andre et al., 2001). The genetic algorithm and its extensions are powerful in their search for the global optimum (Harrouni et al., 1997; Yuping et al., 2002; Leung and Yuping, 2001). GAs have been applied in many fields (Yang et al., 2004; Cheng et al., 2002; Chau, 2002; Chau and Albermani, 2002, 2003; Chau, 2004). However, the computational amount is very large and premature convergence phenomena exist in SGA. To reduce computations and improve the computing precision, the binary-encoded accelerating genetic algorithm (BAGA), real-encoded accelerating genetic algorithm (RAGA) and integer-encoded genetic algorithm were developed (Jin et al., 2001; Davis, 1991; Janikow and Michalewicz, 1991; Renders and Flasse, 1996; Linei et al., 2003). However, these genetic algorithms cannot be effectively applied for continuous variable global optimisations in non-linear models. The Hamming distance between two closest integers in binary code is very large. For instance, integers 511 and 512 are expressed by 0111111111 and 1000000000 in binary code, respectively. All of the codes must be changed if we turn 511 into 512 in binary code. This operation reduces the efficiency of genetic algorithms.

This phenomenon is termed the 'Hamming cliff'. To overcome these difficulties relating to binary encoding for continuous variable optimisations, the SGA was improved with grey encoding for parameters (Andre et al., 2001; Ming and Shudong, 2001; Yang, 2002). For the grey-encoded genetic algorithm (GGA), the integers 511 and 512 are expressed by 0100000000 and 1100000000. It was found that this algorithm still needs a large amount of computation (Yang et al., 2005). In this paper, a Grey-encoded Hybrid Accelerating Genetic Algorithm (GHAGA) will be presented to reduce computations and to improve the computational efficiency. The convergence theorem will be given. This approach will apply to ten non-linear functions and two practical environmental models for verification.

2 The steps of GHAGA

Consider the global optimisation of a non-linear environmental model as

$$\min f(x_1, x_2, \ldots, x_p) \tag{1}$$

$$\text{st. } a_j \leq x_j \leq b_j, \quad \text{for } j = 1, 2, \ldots, p.$$

where $x = \{x_j, j = 1, 2, \ldots, p\}$, x_j is a variable to be optimised, f an objective function and $f \geq 0$ (simple modifications can always turn the problem into such a function).

The procedure for GHAGA is shown as follows.

Step 1: Grey encoding. Suppose the Grey-encoding length is e in every variable, the jth variable range is the interval $[a_j, b_j]$ and then each interval is divided into $2^e - 1$ sub-intervals

$$x_j = a_j + I_j \times c_j, \tag{2}$$

where the length of the sub-interval of the jth variable $c_j = (b_j - a_j)/(2^e - 1)$ is constant. The searching location I_j is an integer, and $0 \leq I_j < 2^e$ for $j = 1, 2, \ldots, p$. The Grey-code array of the jth variable is denoted by the grid points of $\{d(j, k) \mid k = 1, 2, \ldots, e\}$ for every individual

$$I_j = \sum_{m=1}^{e} (\overset{e}{\underset{k=m}{\oplus}} d(j,k)) \times 2^{m-1}, \tag{3}$$

where $\oplus$ denotes the operator of addition modulo 2 on $\{0, 1\}$ (Yang et al., 2005). GHAGA's process operates on a population of individuals (also called Grey-code array strings or chromosomes). Each individual represents a potential solution to the problem.

Step 2: Random creation of the initial population. To cover homogeneously the whole solution space and to avoid individuals entering into the same region, large uniformly random populations are selected in this algorithm. The initial population of n chromosomes is generated at random. The n-corresponding chromosomes are $d(j, k, i)$ (the ith individual) for $j = 1, 2, \ldots, p$; $k = 1, 2, \ldots, e$; $i = 1, 2, \ldots, n$.

Once the initial father population is finished, the decoding and fitness evaluation should be carried out.

Step 3:   Evaluation of the fitness value of each individual (chromosome). Decoding of $d(j, k, i)$ ($j = 1, 2, …, p$; $k = 1, 2, …, e$; $i = 1, 2, …, n$) works through equations (3) and (2) and then the corresponding variable $x_j(i)$ is obtained. Substitution of $x_j(i)$ into equation (1) produces the objective function $f(i)$. The smaller the value $f(i)$ is, the higher the fitness of its corresponding ith chromosome is. So, the fitness function of the ith chromosome is defined:

$$F(i) = \frac{1}{[f(i)]^2 + 0.1} \, .$$

Step 4:   Selection. Select chromosome pairs randomly depending on their fitness value (using roulette wheel method) from the initial population. Two groups of n-chromosomes $d_1(j, k, i)$, $d_2(j, k, i)$, for $j \in \{1, 2, …, p\}$, $k \in \{1, 2, …, e\}$, $i \in \{1, 2, …, n\}$, are obtained.

Step 5:   Two-point crossover. Perform crossover on each chromosome pair according to probability p_c to generate two offspring. For two-point crossover, two crossing points $I_1 = \text{int}(U_1 \times (e + 1))$ $I_2 = \text{int}(U_2 \times (e + 1))$ are randomly chosen, where U_1, U_2 are uniformity random numbers, where $\{U_1, U_2\} \in [0, 1]$. The two-point crossover between two individuals is performed through the crossing probability p_c:

$$d'_1(j,k,i) = \begin{cases} d_2(j,k,i) & k \in [I_1, I_2], \\ d_1(j,k,i) & k \notin [I_1, I_2]; \end{cases} \tag{4}$$

$$d'_2(j,k,i) = \begin{cases} d_1(j,k,i) & k \in [I_1, I_2], \\ d_2(j,k,i) & k \notin [I_1, I_2]. \end{cases} \tag{5}$$

In order to enhance the diversity of population, the crossing probability is set as $P_c = 1$.

Step 6:   Two-point mutation. The operator of two-point mutation is for four random numbers $\{V_1, V_2, V_3, V_4\} \in [0, 1]$. If $V_1 \le 0.5$, the offspring is computed by equation (4). Otherwise, the offspring $d'(j, k, i)$ is computed by equation (5). Let the mutating probability be $p_\text{m} \in [0, 1]$ and the two points be $J_1 = \text{int}(V_2 \times (e + 1))$, $J_2 = \text{int}(V_3 \times (e + 1))$, then the above offspring is mutated from

$$\{d_3(j,k,i)\}_{(j,k,i)} = \begin{cases} 0 & k = J_1, V_4 \le p_m, \{d'(j,k,i)\}_{(j,k,i)} = 1 \\ 1 & k = J_1, V_4 \le p_m, \{d'(j,k,i)\}_{(j,k,i)} = 0 \\ 0 & k = J_2, V_4 \le p_m, \{d'(j,k,i)\}_{(j,k,i)} = 1 \\ 1 & k = J_2, V_4 \le p_m, \{d'(j,k,i)\}_{(j,k,i)} = 0 \\ \{d'(j,k,i)\}_{(j,k,i)} & \text{otherwise} \end{cases} \tag{6}$$

and a new offspring $d_3(j, k, i)$ can be computed by a mutating probability p_m.

Step 7:   Nelder-Mead evolution. The Nelder-Mead algorithm is a useful, local descent algorithm, which does not make use of the objective function derivatives (Nelder and Mead, 1965). The best point in the previous phase becomes a new initial solution in the Nelder-Mead simplex algorithm, and then a new best point is obtained by this Nelder-Mead simplex algorithm. The new best point inside the offspring will be inserted to replace the worst one in the previous phase. Repeat Step 3 to Step 7 until the evolution times Q is met.

Step 8: Accelerating cycle. The variable ranges of n_e-excellent individuals (the n_e individuals nearest to the optimum solution in the evolution) obtained by Q-times of the Nelder-Mead simplex evolution alternating are regarded as the new ranges of the variables, and then the whole process goes back to the Grey encoding (Yang et al., 2005). The GHAGA computation is over until the algorithm running times gets to the design T times or there exists a chromosome c_{fit} whose fitness satisfies a given criterion. In the former case, the c_{fit} is the fittest chromosome or the most excellent chromosome in the population. The chromosome c_{fit} represents the solution.

3 GHAGA's theory

Convergence Theorem: Suppose the initial range of the variables in equation (1) is a closed p-dimensional space B_0:

$$B_0 = \{(x_1, x_2, ..., x_p) \mid a_j^0 \le x_j \le b_j^0, \ j = 1, 2, ..., p\}, \tag{7}$$

and for a moment t, the new ranges of the variables in the accelerating evolution generations are

$$B_t = \left\{(x_1, x_2, ..., x_p) \mid a_j^t \le x_j \le b_j^t, \ j \in \{1, 2, ..., p\}\right\} \tag{8}$$

and $0 \le b_j^t - a_j^t \le b_j^{t-1} - a_j^{t-1}$. For a given constant $a < 1$, the GHAGA is convergent if the shrinking interval ratio k_j^t has the following relation (Yang et al., 2005)

$$k_j^t = \frac{b_j^t - a_j^t}{b_j^{t-1} - a_j^{t-1}} \le a < 1. \tag{9}$$

Proof: For the GHAGA, the shrinking interval ratio is defined as

$$k_j^t = \frac{b_j^t - a_j^t}{b_j^{t-1} - a_j^{t-1}}.$$

As $T \to \infty$ for $j \in \{1, 2, ..., p\}$, if equation(9) holds, we have

$$0 \le \prod_{t=1}^{T} k_j^t \le \prod_{t=1}^{T} a \to 0,$$

which implies

$$B_0 \supset B_1 \supset \cdots \supset B_t \supset B_{t+1} \supset \cdots$$
$$\|x_i - y_i\| \to 0 \text{ as } t \to \infty \text{ for } \{x_t, y_t\} \in B_t,$$

where $\|\cdot\|$ denotes a norm. According to the *Banach* fixed point theorem, we can conclude that the GHAGA is convergent (Yang et al., 2005).

Considering the Nelder-Mead simplex evolution generations Q, excellent individuals n_e, optimised variable p and accelerating evolution generations T, the probability p_e of excellent individuals surrounding the optimum point is $p_e = (1 - 2^{Qn_e})^{pT}$

(Yang et al., 2005). The probability p_e is given in Table 1. From the above theorem and Table 1, we can conclude that the global optimisation of the GHAGA is not only accurate but also stable. And the GHAGA is global convergence with probability $p_e = 1.000000$ and $n_e = 10$, $Q \geq 3$.

Table 1   The probability p_e of excellent individuals for the GHAGA

Q	p	T	p_e
2	10	5	0.999952
3	10	5	1.000000
5	10	2	1.000000
5	10	5	1.000000
10	20	10	1.000000

4   Numerical simulations

Three main criteria appeared to be very important when trying to determine the performances of an algorithm: convergence, speed and robustness (Andre et al., 2001). The following is performed to demonstrate the efficiency of GHAGA.

The parameters of GHAGA are selected as follows: The length $e = 10$, population size $n = 300$, the number of excellent individuals $n_e = 10$, the times of Nelder-Mead simplex evolution alternating $Q = 3$, the crossover probability $p_c = 1.0$, the mutation probability $p_m = 0.5$ and the times of Nelder-Mead simplex searching $m \leq 600$.

The aim of this study is the search for global optimum of non-linear models. So, it is of importance that our GA effectively converges to this optimum. The global optimisation of the ten test functions (Andre et al., 2001) is accomplished by the use of the following methods: Standard Binary-encoded GA (SGA, Andre et al., 2001), Improved Grey-encoded GA (IGGA, Andre et al., 2001) and Grey-encoded Hybrid Accelerating Genetic Algorithm (GHAGA). To compare with the global optimisation ability of the three algorithms objectively, less than or equal to 18,000 computations of the objective functions are performed, and one of the three termination criteria is used for ensuring the optimisation precision and avoiding algorithm invalidation.

Criteria one: Relative error criteria. The relative error E_{rel} of calculating optimisation values is used each time if it is possible:

$$E_{rel} = \frac{\left| f_{algo} - f_{exact} \right|}{\left| f_{exact} \right|}.$$

Criteria two: Absolute error criteria. When the optimum is 0, it is no longer possible to use the above expression, so we calculate the absolute error E_{abs} of calculating optimisation values:

$$E_{abs} = \left| f_{algo} - f_{exact} \right|.$$

We let the absolute error or relative error of calculating optimisation values in neighbour generations be less than or equal to 0.01 and the distance of computed optimisation points in the closest generations be less than or equal to 0.01.

Criteria three: Total computation amount criteria. The total computation amount of function is less than or equal to 18,000, or the interval $b_j^t - a_j^t$ is less than or equal to 0.00001.

For each of these criteria and each evolution of our algorithm, the verifications are carried out in a statistic method and 100 tests for each of ten functions (Andre et al., 2001) are completed, which implies that the result may be reliable.

The computational results of the ten non-linear test functions on global optimisation, convergent speed and calculating accuracy in the SGA (Andre et al., 2001), IGGA (Andre et al., 2001) and GHAGA are given in Table 2. It is clearly observed that the GHAGA is the best one both in efficiency and in accuracy (see minimum found and number of evaluation of the functions in Table 2). Moreover, the 'Hamming cliff' phenomena are avoided in GHAGA.

Table 2   Results with the SGA, IGGA and GHAGA

Name of the functions	Theoretical minimum	Minimum found			Number of evaluation of the functions		
		SGA	IGGA	GHAGA	SGA	IGGA	GHAGA
F1	−1.12323	−1.12323	−1.12323	−1.12323	5,566	784	300
Branin	0.39789	0.39789	0.39791	0.39789	8,125	2,040	300
Camelback	−1.03163	−1.03163	−1.03163	−1.03163	1,316	1,316	300
Goldprice	3.00000	3.00000	3.00028	3.00000	8,185	4,632	303
Pshubert	−186.73091	−186.73100	−186.70469	−186.73090	7,303	4,116	402
Shubert	−186.73091	−186.73100	−186.72802	−186.73090	6,976	2,364	321
Shekel	−10.53641	−10.40340	−10.51404	−10.52122	8,521	36,772	696
Hartman1	−3.86278	−3.86249	−3.86114	−3.86278	1,993	1,680	300
Hartman2	−3.32237	−3.30652	−3.31383	−3.31953	19,452	53,792	768
Hosc45	1.00000	1.99506	1.00943	1.00000	11,140	126,139	3,522

Lastly, we have performed a comparison of our GA with six other methods of iterative improvement listed in Table 3: pure random search (PRS, Anderssen et al., 1972), multistart (MS, Rinnoy and Timmer, 1987), simulated diffusion (SD, Aluffipentini et al., 1985), simulated annealing (SA, Dekkers and Aarts, 1991), tabu search (TS, Cvijovic and Klinowski, 1995) and improved grey-encoded genetic algorithm (IGGA, Andre et al., 2001). Efficiency was qualified in terms of the number of function evaluations necessary to find the global optimum. Each program was stopped as soon as the relative error between the best point found and the known global optimum was less than 1%. The numbers of function evaluations used by the various algorithms to optimise five test functions are listed in Table 4. It should be pointed out that we did not program ourselves the competitive algorithms used for the comparison, but we reported the results published by Cvijovic and Klinowski (1995) and Andre et al. (2001). We can see that results achieved with our GA are satisfactory in global optimum and convergent speed (see the numbers of function evaluations in Table 4). In addition, our results were the average outcome of 100 independent runs. For some published methods, the number of runs was equal to four or unspecified (Andre et al., 2001).

Table 3   Global optimisation methods used for performance analysis

Method	Name	Reference
PRS	Pure random search	Anderssen et al. (1972)
MS	Multistart	Rinnoy and Timmer (1987)
SD	Simulated diffusion	Aluffipentini et al. (1985)
SA	Simulated annealing	Dekkers and Aarts, (1991)
TS	Tabu search	Cvijovic and Klinowski (1995)
IGGA	Improved gray-encoded genetic algorithm	Andre et al. (2001)
GHAGA	Grey-encoded hybrid accelerating genetic algorithm	This work

Table 4   Number of function evaluations in global optimisation of five functions by the seven different methods defined in Table 3

Method	Function				
	Goldprice	Branin	Hartman1	Hartman2	Shubert
PRS	5,125	4,850	5,280	18,090	6,700
MS	4,400	1,600	2,500	6,000	–
SD	5,439	2,700	3,416	3,975	2,41,215
SA	563	505	1,459	4,648	780
TS	486	492	508	2,845	727
IGGA	4,632	2,040	1,680	53,792	2,364
GHAGA	303	300	300	768	321

5   Practical example

Example 1: O'Connor water quality model

$$c = \left(10 - \frac{15k_1}{k_1 - k_2} - \frac{0.5k_3}{k_3 - k_2} - 10\right)e^{-\frac{k_2 x}{4}} + \frac{15k_1}{k_1 - k_2}e^{-\frac{k_1 x}{4}} + \frac{0.5k_3}{k_3 - k_2}e^{-\frac{k_3 x}{4}} + 10$$

where k_1 is the oxygen consumption coefficient (h^{-1}), k_2 is the reaeration coefficient (h^{-1}), k_3 is the NBOD reduction coefficient (h^{-1}), x is flow distance (km) and c is the dissolved oxygen concentration at point x. The three parameters k_1, k_2 and k_3 are required in this model. In this work, the parameters are estimated with respect to one criterion, namely the sum of least squares. The form of the objective function is described as

$$\text{Minimise } f = \sum_{i=1}^{5}(d_i - c_i)^2$$

where d_i and c_i are the measured and estimated values of dissolved oxygen concentration at point x_i, respectively.

The least residual square sum *f* is 0.5046 with GHAGA. GHAGA runs one second for optimisation of this model only. The computational results of the above model are given in Table 5. For GHAGA, the evaluation of acceleration is only one time and the evaluation number of the objective function is 300. For BAGA (Jin and Ding, 2000), the evolution of acceleration is eight times, the evaluation number of the objective function is 2400 and the least residual square sum *f* is 1.2516. The simulation results of BAGA and GHAGA for the practical example are shown in Figure 1(a), (b), respectively. From Table 5, Figure 1(a), (b), we can see that results achieved with our GA are satisfactory in global optimum and convergent speed.

Table 5 The calculating result of several methods for the practical example

Method	Evaluation number for f	Parameters			Least residual square sum f
		k_1	k_2	k_3	
GHAGA	600	0.0642	0.1560	0.0001	0.5046
BAGA	2,400	0.0780	0.1660	0.0630	1.2516

Figure 1 (a) Simulating result of BAGA for the practical example and (b) Simulating result of GHAGA for the practical example 1

(a)

(b)
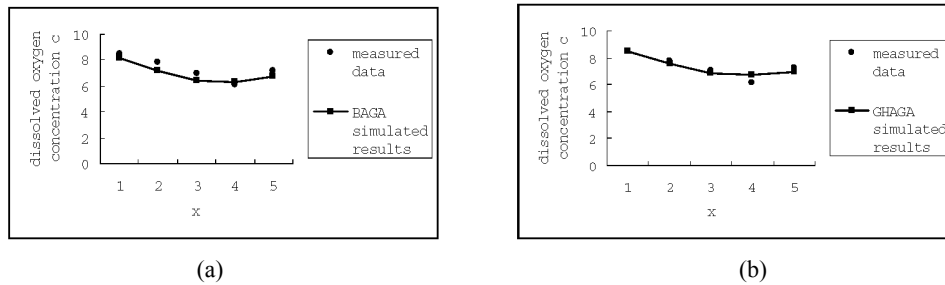
Example 2: An automatic calibration for the rainfall-runoff model (Xinanjiang model).

The rainfall-runoff model used in this study is based on the latest version of the Xinanjiang model (Zhao, 1992; Cheng et al., 2002). In this paper, the model is applied to the Tankou basin in China. The basin, with a catchment area of 1755.4 km^2, includes Tankou area of 992.3 km^2, Qimen area of 519.3 km^2 and Likou area of 243.8 km^2. The parameters WM, X, Y, KE, B, SM, EX, KI, KG, IMP, C, CI, CG, N_1, NK_1, N_2, NK_2, N_3,NK_3 are required in this model. The significance of these parameters is shown in Table 6. In this work, the 19 parameters are estimated with respect to one criterion, namely the sum of least squares. The form of the objective function is described as

$$\text{Minimise } f = \sum_{i=1}^{N}[Q_0(i) - Q_c(i)]^2$$

where Q_0 is the observed discharges, Q_c is the calculated discharges and N is the number of time intervals. Parameter ranges in search are also shown in Table 6. Most of the ranges have been selected according to the recommendation in Zhao (1992).
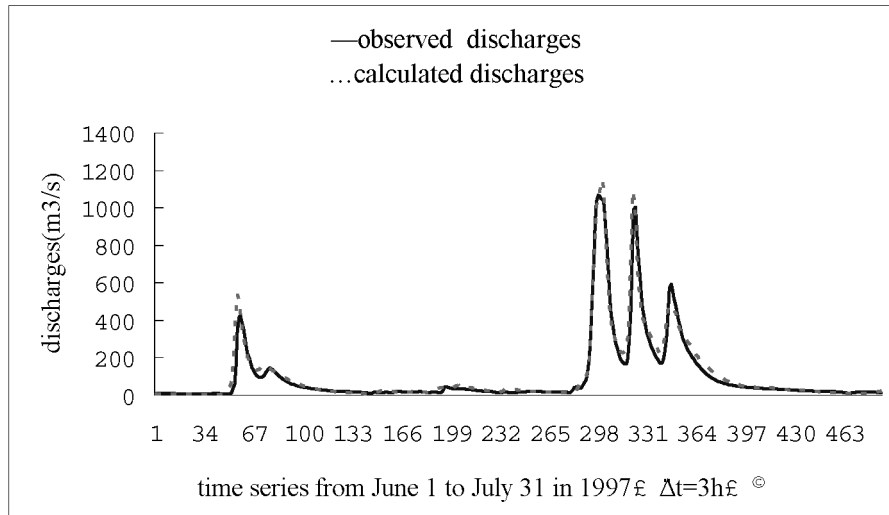
Table 6   Parameters of the Xinanjiang model

| Parameters	Physical meaning	Range
WM	Area mean tension water capacity	100 ~ 200 mm
X	Ratio of the WUM to WM, where WUM is average basin storage capacity of the upper layer	0.05 ~ 0.3
Y	Ratio of the WLM to $(1 - X)$ WM, where WLM is average basin storage capacity of the lower layer (WM = WUM + WLM + WDM)	0.1 ~ 0.9
KE	Ratio of potential evapotranspiration to pan evaporation	0.5 ~ 2.0
B	Exponent in the spatial distribution curve of soil moisture storage capacity	0.15 ~ 0.4
SM	Area mean free water storage capacity	10 ~ 50 mm
EX	A parameter in the distribution of free water storage capacity	0.05 ~ 2.0
KI	Outflow coefficients of the free water storage to interflow relationships	0.2 ~ 0.48
KG	Outflow coefficients of the free water storage to groundwater relationships	0.2 ~ 0.48
IMP	Proportion of impermeable area to the total area	0.01 ~ 0.5
C	Evapotranspiration coefficient from deep layer	0.01 ~ 0.2
CI	The interflow recession coefficient	0.7 ~ 1.0
CG	The groundwater recession coefficient	0.7 ~ 1.0
N_1	Number of cascade linear reservoir for runoff routing for Tankou of 992.3 km^2 area	1 ~ 10
NK_1	Scale parameter of cascade linear reservoir for Tankou of 992.3 km^2 area	1 ~ 20
N_2	Number of cascade linear reservoir for runoff routing for Qimen of 519.3 km^2 area	1 ~ 10
NK_2	Scale parameter of cascade linear reservoir for Qimen of 519.3 km^2 area	1 ~ 20
N_3	Number of cascade linear reservoir for runoff routing for Likou of 243.8 km^2 area	1 ~ 10
NK_3	Scale parameter of cascade linear reservoir for Likou of 243.8 km^2 area	1 ~ 20

The least residual square sum f is 0.08544 with GHAGA. For GHAGA, the evaluation number of the objective function is 1,800. The computational results of the above model are given in Table 7. For RAGA (real-encoded accelerating genetic algorithm), the evaluation number of the objective function is 3,000 and the least residual square sum f is 0.10540. Eight years' data, starting from 1st January 1989, of three-hour values of rainfall and potential evapotranspiration were used for the calibration period (data series from 1 to 23,296). Two years' data, starting from 1st January 1997, of three-hour values of rainfall and potential evapotranspiration were used for the verification period (data series from 23,297 to 29,120). Figure 2(a) gives the observed and calculated discharges with GHAGA in Tankou basin from 1st June to 31st July in 1997, and Figure 2(b) gives the observed and calculated discharges with GHAGA in Tankou basin

from 1st June to 31st July in 1998. From Table 7, Figure 2(a), (b) we can see that the results achieved with our GA are satisfactory in convergent speed. In terms of minimising the objective function, GHAGA has been shown to be suitable for the rainfall-runoff model.

Figure 2　The observed and calculated discharges with GHAGA in Tankou basin from 1st June to 31st July in (a) 1997 and (b) 1998

(a)
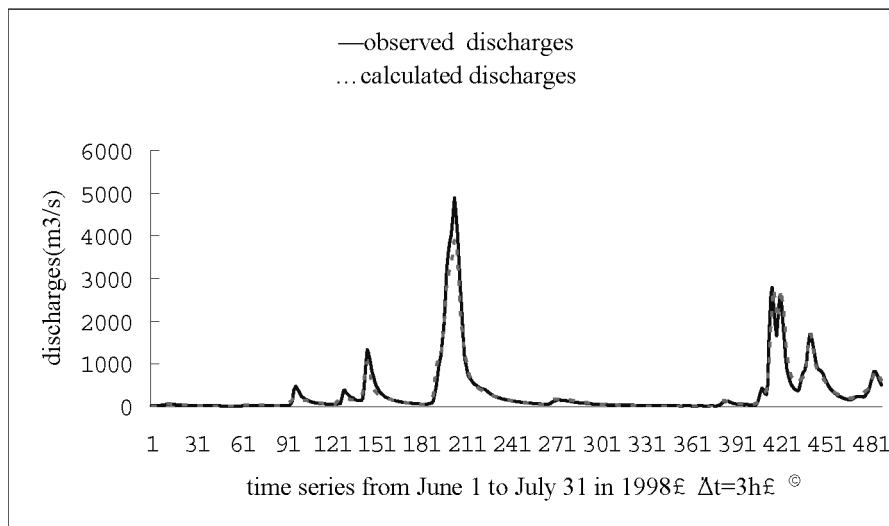
(b)

Table 7   Results for optimising the parameter of the Xinanjiang model in Tankou

Parameter	Optimising value
WM	140.36
X	0.10
Y	0.40
KE	1.51
B	0.38
SM	16.71
EX	0.05
KI	0.28
KG	0.20
IMP	0.01
C	0.20
CI	0.90
CG	0.97
N_1	8.74
NK_1	4.48
N_2	4.32
NK_2	6.62
N_3	9.99
NK_3	11.28

6   Conclusion

In this paper, the grey-encoded hybrid accelerating genetic algorithm (GHAGA) with the Nelder-Mead simplex searching operator is developed for the parameter optimisation of non-linear environmental models. In order to address the trade-off between precision, reliability and computing time in the grey-encoded genetic algorithm for global optimisation on continuous variables, we emphasise two deciding alterations made to the algorithm: the definition interval of each variable is shrunk with the excellent individuals of each generation and the hybrid integration of the grey-encoded genetic algorithm and Nelder-Mead simplex algorithm. The circulating mechanism of GHAGA has been studied. The corresponding convergence theorem and global optimisation ability are analysed for guaranteeing the global convergence of the new genetic algorithm. Because the steps of grey encoding, Nelder-Mead simplex hybrid searching operator and accelerating cycle are adopted, the efficiency and accuracy of the new algorithm are very high compared to existing algorithms. This algorithm overcomes any Hamming-cliff phenomena in existing genetic methods, it has been applied to ten non-linear test functions and the practical O'Connor water quality model and rainfall-runoff model and the results are encouraging.

Acknowledgements

References

Aluffipentini, F., Parisi, V. and Zililli, F. (1985) 'Global optimization and stochastic differential equations', *J. Opt. Theor. Appl.,* Vol. 47, pp.1–16.

Anderssen, R., Jennings, L. and Ryan, D. (1972) *Optimization*, University of Queensland Press, St. Lucia, Brisbane, Australia, pp.27–34.

Andre, J., Siarry, P. and Dognon, T. (2001) 'An improvement of the standard genetic algorithm fighting premature convergence in continuous optimization', *Advances in Engineering Software*, Vol. 32, pp.49–60.

Chau, K.W. (2002) 'Calibration of flow and water quality modeling using genetic algorithm', *Lecture Notes in Artificial Intelligence*, Vol. 2557, pp. 720.

Chau, K.W. and Albermani, F. (2002) 'Genetic algorithms for design of liquid retaining structure', *Lecture Notes in Artificial Intelligence*, Vol. 2358, pp.119–128.

Chau, K.W. and Albermani, F. (2003) 'Knowledge-based system on optimum design of liquid retaining structures with genetic algorithm', *Journal of Structural Engineering*, ASCE, Vol. 129, pp.1312–1321.

Chau, K.W. (2004) 'A two-stage dynamic model on allocation of construction facilities with genetic algorithm', *Automation in Construction*, Vol. 13, pp.481–490.

Cheng, C.T., Ou, C.P. and Chau, K.W. (2002) 'Combining a fuzzy optimal model with a genetic algorithm to solve multiobjective rainfall-runoff model calibration', *Journal of hydrology*, Vol. 268, pp.72–86.

Cvijovic, D. and Klinowski, J. (1995) 'Taboo search: an approach to the multiple minima problem', *Science*, Vol. 267, pp.664–666.

Davis, L. (1991) *Handbook of Genetic Algorithms*, Van Nostrand Reinhold, New York.

De Jong, K.A. (1975) 'An analysis of the behavior of a class of genetic adaptive systems', *Doctoral dissertation*, University of Michigan, Ann Arbor, MI. Dissertation abstracts international, Vol. 36, pp.5140, 5141.

Dekkers, A. and Aarts, E. (1991) *Mathematical Programming*, Vol. 50, pp.367, 376.

Goldberg, D.E. (1989) *Genetic Algorithms: Search, Optimization and Machine Learning*, Addison-Wesley, New York.

Harrouni, K.E.I., Ouazar, D., Walters, G.A. and Cheng, A.H-D. (1997) 'Groundwater optimization and parameter estimation by genetic algorithm and dual reciprocity boundary element method', *Engineering Analysis with Boundary Elements*, Vol. 18, pp.287–296.

Holland, J.H. (1975) *Adaptation in Natural and Artificial Systems*, University of Michigan press, Ann Arbor.

Janikow, C.Z. and Michalewicz, Z. (1991) 'An experimental comparison of binary and floating point representation in genetic algorithms', *Proceedings of the Fourth International Conference on Genetic Algorithms*, San Francisco, Morgan Kaufman, pp.31–36.

Jin, J.L. and Ding, J. (2000) *Genetic Algorithm and its Applications to Water Science*, Sichuan University, Sichuan.

Jin, J.L., Yang, X.H. and Ding, J. (2001) 'An improved simple genetic algorithm – accelerating genetic algorithm', *Xitong Gongcheng Lilun Yu Shijian*, Vol. 21, pp.8–13.

Leung, Y.W. and Yuping, W. (2001) 'An orthogonal genetic algorithm with quantization for global numerical optimization', *IEEE Trans. On Evolutionary Computation*, Vol. 5, pp.41–53.

Linei, Y., Guiming, W., Ke, G. and Junsong, L. (2003) 'An application genetic algorithms with initial group floating for mixed integer nonlinear programming', *Computing Techniques for Geophysical and Geochemical Exploration*, Vol. 25, pp.253–258.

Ming, Z. and Shudong, S. (2001) *Genetic Algorithms: Theory and Applications*, Defence industry press, Beijing.

Nelder, J. and Mead, R. (1965) 'A simplex method for function minimization', *The Computer Journal*, Vol. 7, pp.308–313.

Renders, J.M. and Flasse, S.P. (1996) 'Hybrid methods using genetic algorithms for global optimization', *IEEE Trans Systems, Man Cybernetics–part B*, Cybernetics, Vol. 26, pp.243–258.

Rinnoy, K. and Timmer, G. (1987) 'Stochastic global optimization methods. Part II: multi-level methods', *Mathematical Programming*, Vol. 39, pp.57–78.

Yang, X.H. (2002) 'Study on parameter optimization algorithm and its application in hydrological model', *Doctoral dissertation*, Hohai University (in Chinese), Nanjing.

Yang, X.H., Yang, Z.F., Shen, Z.Y. and Li, J.Q. (2004) 'An ideal interval method of multi-objective decision-making for comprehensive assessment of water resources renewability', *Science in China Series E-engineering and Materials Science*. Vol. 47, Suppl. S, pp.42–50.

Yang, X.H., Yang, Z.F., Lu, G.H. and Li, J.Q. (2005) 'A gray-encoded, hybrid-accelerated, genetic algorithm for global optimizations in dynamical systems', *Communications in Nonlinear Science and Numerical Simulation*, Vol. 10, No. 4, pp.355–363.

Yuping, W., Yongchang, J. and Fushun, Z. (2002) 'A new evolution algorithm for unconstrained nonlinear global optimization problems and its convergence', *Electronica Sinca*, pp.1867–1869 (in Chinese).

Zhao, R.J. (1992) 'The Xinanjiang model applied in China', *Journal of Hydrology*, Vol. 135, pp.371–381.