Krill herd: A new bio-inspired optimization algorithm

Amir Hossein Gandomi [a,*], Amir Hossein Alavi [b]

[a] Department of Civil Engineering, The University of Akron, Akron, OH, USA
[b] School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran

ABSTRACT

In this paper, a novel biologically-inspired algorithm, namely krill herd (KH) is proposed for solving optimization tasks. The KH algorithm is based on the simulation of the herding behavior of krill individuals. The minimum distances of each individual krill from food and from highest density of the herd are considered as the objective function for the krill movement. The time-dependent position of the krill individuals is formulated by three main factors: (i) movement induced by the presence of other individuals (ii) foraging activity, and (iii) random diffusion. For more precise modeling of the krill behavior, two adaptive genetic operators are added to the algorithm. The proposed method is verified using several benchmark problems commonly used in the area of optimization. Further, the KH algorithm is compared with eight well-known methods in the literature. The KH algorithm is capable of efficiently solving a wide range of benchmark optimization problems and outperforms the exciting algorithms.

Published by Elsevier B.V.

1. Introduction

Recently, metaheuristic optimization algorithms have been widely employed for solving complex optimization problems. These algorithms are found to be more powerful than the conventional methods that are based on formal logics or mathematical programming [1]. Intensification and diversification are two main features of the metaheuristic algorithms [2]. The intensification phase searches around the current best solutions and selects the best candidates or solutions. The diversification phase ensures that the algorithm explores the search space more efficiently. The specific objectives of developing modern metaheuristic algorithms are to solve problems faster, to solve large problems, and to obtain more robust methods [3].

The metaheuristic algorithms do not have limitations in using source of inspiration (e.g. music-inspired harmony search [4] or physic-inspired charged system search [5]). However, nature is a principal inspiration for proposing new metaheuristic approaches and the nature-inspired algorithms have been widely used in developing systems and solving problems [6]. Biologically-inspired algorithms are one of the main categories of the nature-inspired metaheuristic algorithms. The efficiency of the bio-inspired algorithms is due to their significant ability to imitate the best features in nature. More specifically, these algorithms are based on the selection of the fittest in biological systems which have evolved by natural selection over millions of years. Various bio-inspired optimization algorithms have been developed during the recent decades.

The bio-inspired algorithms may generally be divided into three main categories [7]:

(1) Evolutionary algorithms,
(2) Swarm intelligence, and
(3) Bacterial foraging algorithms.

---

* Corresponding author.
  E-mail addresses: a.h.gandomi@gmail.com, ag72@uakron.edu (A.H. Gandomi), ah_alavi@hotmail.com (A.H. Alavi).

The evolutionary algorithms are inspired from the genetic evolution process. Genetic algorithm (GA) [8], genetic programming (GP) [9], evolutionary strategy (ES) [10] and differential evolution (DE) [11] are the well-known paradigms of the evolutionary algorithms. These branches are population-based stochastic search algorithms working with best-to-survive criteria [7]. The evolutionary algorithms have been remarkably improved over the last decades. Stud genetic algorithm (SGA) [12] is a robust GA that only uses the best individual at each generation for doing the crossover. Gandomi and Alavi [13] proposed multi-stage genetic programming as an improvement of the GP for non-linear system modeling. It is based on incorporating the sole effect of predictor variable as well as the interactions between the variables to provide more accurate simulations. The mentioned algorithms have been widely used to solve different kind of optimization tasks (e.g. [14]). Simon [15] has proposed a new evolutionary algorithm, namely biogeography-based optimization (BBO). The BBO algorithm is used for global recombination and uniform crossover which are inspired from the GA literature.

The most well-known paradigms in the area of swarm intelligence are particle swarm optimization (PSO) [16] and ant colony optimization (ACO) [17]. These algorithms are based on the simulation of the collective behavior of animals. The PSO algorithm was first proposed by Eberhart and Kennedy [16]. PSO is a population-based method inspired from the social behavior of bird flocking or fish schooling. The ACO algorithm is inspired by the collective foraging behavior of ants [17]. These algorithms have been widely used in the literature to solve optimization problems. Several extensions to the major categories of the swarm algorithms have been presented in the literature [18].

The bacterial foraging behavior has been a source for developing a new bio-inspired optimization approach, called bacterial foraging algorithm [7,19]. The most well-known types of the bacterial foraging algorithms are computing systems of microbial interactions and communications (COSMIC) [20] and rule-based bacterial modeling (RUBAM) [21].

This paper presents a new bio-based swarm intelligence algorithm, called krill herd (KH). This method is based on the simulation of the herding of the krill swarms in response to specific biological and environmental processes. Nearly all of the necessary coefficients for implementing the proposed algorithm are obtained from real world empirical studies done in the literature. The fitness function of each krill individual is defined as its distances from food and highest density of the swarm. Three essential actions considered to determine the time-dependent position of an individual krill are (i) movement induced by other krill individuals, (ii) foraging activity, and (iii) random diffusion. This paper is organized as follows: Section 2 presents the basic aspects and the characteristics of the KH algorithm, including formulation of the algorithm by idealizing the herding behavior of the krill individuals. Two adaptive genetic mechanisms (crossover and mutation) are also introduced in this section. Numerical examples and comparison with well-known algorithms are presented in Section 3 to verify the efficiency of the KH algorithm. Finally, some concluding remarks and suggestions for future research are provided in Section 4.

2. Krill herd algorithm

2.1. Herding behavior of krill swarms

The formation of groupings of various species of marine animals are under-dispersed and non-random. Many studies have focused on capturing the underlying mechanisms governing the development of these formations [22,23]. The major mechanisms identified are related to the feeding ability, enhanced reproduction, protection from predators, and environmental conditions [24]. Some mathematical models have been developed to evaluate the relative contribution of these mechanisms based on experimental observations [22,23].

Antarctic krill is one of the best-studied species of marine animal. The krill herds are aggregations with no parallel orientation existing on time scales of hours to days and space scales of 10 s to 100 s of meters. One of the main characteristics of this specie is its ability to form large swarms [25,26]. Over the last three decades, several studies have been conducted to understand the ecology and distribution of krill. Although there are yet notable uncertainties about the forces determining the distribution of the krill herd [27], conceptual models have been proposed to explain the observed formation of the krill herds [28]. The results obtained by such conceptual frameworks revealed that the krill swarms form the basic unit of organization for this species. In order to better understand the formation of the krill swarms, the proximate causes and the factors that are adaptive advantages of aggregation formation (ultimate effects) should be distinguished [29].

When predators, such as seals, penguins or seabirds, attack krill, they remove individual krill. This results in reducing the krill density. The formation of the krill herd after predation depends on many parameters. The herding of the krill individuals is a multi-objective process including two main goals: (1) increasing krill density, and (2) reaching food. In the present study, this process is taken into account to propose a new metaheuristic algorithm for solving global optimization problems. Density-dependent attraction of krill (increasing density) and finding food (areas of high food concentration) are used as objectives which finally lead the krill to herd around the global minima. In this process, an individual krill moves toward the best solution when it searches for the highest density and food. That is, the closer the distance to the high density and food, the less the objective function. Generally, some coefficients should be determined for using multi-objective herding behavior for a single objective one. In this study, the coefficients are determined on the basis of a specialized literature review of the experimental observations of the krill behavior [24,30,31] and also after a trial study.

2.2. Lagrangian model of the krill herding

Predation removes individuals, leads to reduction of the average krill density, and distances the krill swarm from the food location. This process is assumed to be the initialization phase in the KH algorithm. In the natural system, the fitness of each individual is supposed to be a combination of the distance from the food and from the highest density of the krill swarm. Therefore, the fitness (imaginary distances) is the value of the objective function. The time-dependent position of an individual krill in 2D surface is governed by the following three main actions [24]:

  i. Movement induced by other krill individuals;
  ii. Foraging activity; and
  iii. Random diffusion

It is known that an optimization algorithm should be capable of searching spaces of arbitrary dimensionality. Therefore, the following Lagrangian model is generalized to an n dimensional decision space:

$$\frac{dX_i}{dt} = N_i + F_i + D_i \tag{1}$$

where N_i is the motion induced by other krill individuals; F_i is the foraging motion, and D_i is the physical diffusion of the ith krill individuals.

2.2.1. Motion induced by other krill individuals

According to theoretical arguments, the krill individuals try to maintain a high density and move due to their mutual effects [24]. The direction of motion induced, α_i, is estimated from the local swarm density (local effect), a target swarm density (target effect), and a repulsive swarm density (repulsive effect) [24]. For a krill individual, this movement can be defined as:

$$N_i^{\text{new}} = N^{\text{max}}\alpha_i + \omega_n N_i^{\text{old}} \tag{2}$$

where,

$$\alpha_i = \alpha_i^{\text{local}} + \alpha_i^{\text{target}} \tag{3}$$

and N^{max} is the maximum induced speed, ω_n is the inertia weight of the motion induced in the range [0, 1], N_i^{old} is the last motion induced, α_i^{local} is the local effect provided by the neighbors and α_i^{target} is the target direction effect provided by the best krill individual. According to the measured values of the maximum induced speed [24], it is taken 0.01 (ms^{-1}).

The effect of the neighbors can be assumed as an attractive/repulsive tendency between the individuals for a local search. In this study, the effect of the neighbors in a krill movement individual is determined as follows:

$$\alpha_i^{\text{local}} = \sum_{j=1}^{NN} \widehat{K}_{i,j} \widehat{X}_{i,j} \tag{4}$$

$$\widehat{X}_{i,j} = \frac{X_j - X_i}{\|X_j - X_i\| + \varepsilon} \tag{5}$$

$$\widehat{K}_{i,j} = \frac{K_i - K_j}{K^{\text{worst}} - K^{\text{best}}} \tag{6}$$

where K^{best} and K^{worst} are the best and the worst fitness values of the krill individuals so far; K_i represents the fitness or the objective function value of the ith krill individual; K_j is the fitness of jth $(j = 1,2,\ldots,NN)$ neighbor; X represents the related positions; and NN is the number of the neighbors. For avoiding the singularities, a small positive number, ε, is added to the denominator.

The right sides of Eqs. (4)–(6) contain some unit vectors and some normalized fitness values. The vectors show the induced directions by different neighbors and each value presents the effect of a neighbor. The neighbors' vector can be attractive or repulsive since the normalized value can be negative or positive.

For choosing the neighbor, different strategies can be used. For instance, a neighborhood ratio can be simply defined to find the number of the closest krill individuals. Using the actual behavior of the krill individuals, a sensing distance (d_s) should be determined around a krill individual (as shown in Fig. 1) and the neighbors should be found.

The sensing distance for each krill individual can be determined using different heuristic methods. Here, it is determined using the following formula for each iteration:

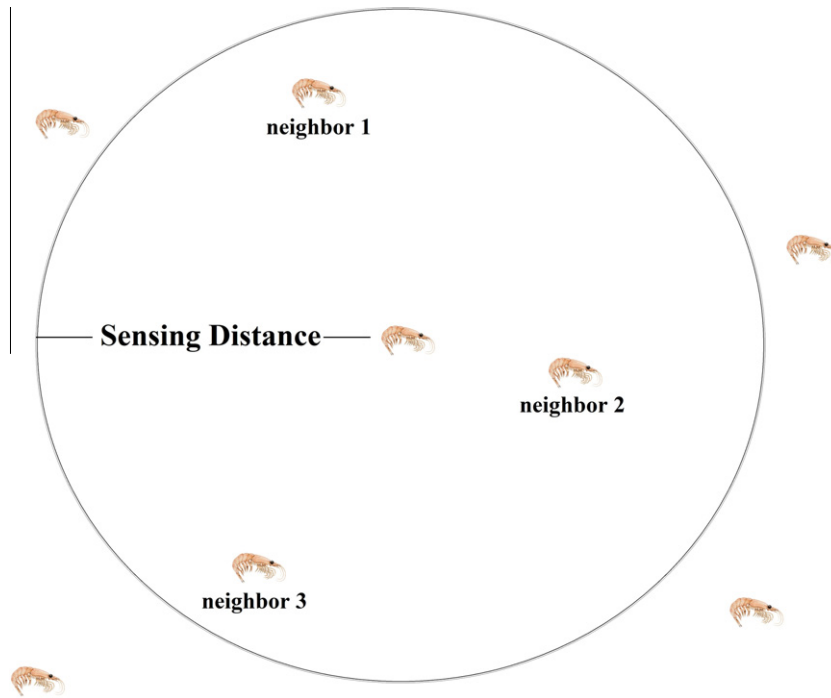$$d_{s,i} = \frac{1}{5N} \sum_{j=1}^{N} \|X_i - X_j\| \tag{7}$$

Fig. 1. A schematic representation of the sensing ambit around a krill individual.

where $d_{s,i}$ is the sensing distance for the ith krill individual and N is the number of the krill individuals. The factor 5 in the denominator is empirically obtained. Using Eq. (7), if the distance of two krill individuals is less than the defined sensing distance, they are neighbors.

The known target vector of each krill individual is the lowest fitness of an individual krill. The effect of the individual krill with the best fitness on the ith individual krill is taken into account using Eq. (8). This level leads it to the global optima and is formulated as:

$$\alpha_i^{\text{target}} = C^{\text{best}} \widehat{K}_{i,\text{best}} \widehat{X}_{i,\text{best}} \tag{8}$$

where, C^{best} is the effective coefficient of the krill individual with the best fitness to the ith krill individual. This coefficient is defined since α_i^{target} leads the solution to the global optima and it should be more effective than other krill individuals such as neighbors. Herein, the value of C^{best} is defined as:

$$C^{\text{best}} = 2\left(\text{rand} + \frac{I}{I_{\text{max}}}\right) \tag{9}$$

where $rand$ is a random values between 0 and 1 and it is for enhancing exploration, I is the actual iteration number and I_{max} is the maximum number of iterations.

2.2.2. Foraging motion

The foraging motion is formulated in terms of two main effective parameters. The first one is the food location and the second one is the previous experience about the food location. This motion can be expressed for the ith krill individual as follows:

$$F_i = V_f \beta_i + \omega_f F_i^{\text{old}} \tag{10}$$

where

$$\beta_i = \beta_i^{\text{food}} + \beta_i^{\text{best}} \tag{11}$$

and V_f is the foraging speed, ω_f is the inertia weight of the foraging motion in the range [0,1], is the last foraging motion, β_i^{food} is the food attractive and β_i^{best} is the effect of the best fitness of the ith krill so far. According to the measured values of the foraging speed [30], it is taken 0.02 (ms^{-1}).

The food effect is defined in terms of its location. The center of food should be found at first and then try to formulate food attraction. This cannot be determined but can be estimated. In this study, the virtual center of food concentration is estimated according to the fitness distribution of the krill individuals, which is inspired from "*center of mass*". The center of food for each iteration is formulated as:

$$X^{food} = \frac{\sum_{i=1}^{N} \frac{1}{K_i} X_i}{\sum_{i=1}^{N} \frac{1}{K_i}} \tag{12}$$

Therefore, the food attraction for the ith krill individual can be determined as follows:

$$\beta_i^{food} = C^{food} \widehat{K}_{i,food} \widehat{X}_{i,food} \tag{13}$$

where C^{food} is the food coefficient. Because the effect of food in the krill herding decreases during the time, the food coefficient is determined as:

$$C^{food} = 2\left(1 - \frac{I}{I_{max}}\right) \tag{14}$$

The food attraction is defined to possibly attract the krill swarm to the global optima. Based on this definition, the krill individuals normally herd around the global optima after some iteration. This can be considered as an efficient global optimization strategy which helps improving the globality of the KH algorithm.

The effect of the best fitness of the ith krill individual is also handled using the following equation:

$$\beta_i^{best} = \widehat{K}_{i,ibest} \widehat{X}_{i,ibest} \tag{15}$$

where K_{ibest} is the best previously visited position of the ith krill individual.

2.2.3. Physical diffusion

The physical diffusion of the krill individuals is considered to be a random process. This motion can be express in terms of a maximum diffusion speed and a random directional vector. It can be formulated as follows:

$$D_i = D^{max} \delta \tag{16}$$

where D^{max} is the maximum diffusion speed, and δ is the random directional vector and its arrays are random values between -1 and 1. Wolpert and Macready [31] proposed a range for the maximum diffusion speed of the krill individuals as $D^{max} \in [0.002, 0.010]\,(ms^{-1})$ and a random number in this range is also used in this study. The better the position of the krill is, the less random the motion is. Thus, another term is added to the physical diffusion formula to consider this effect. The effects of the motion induced by other krill individuals and foraging motion gradually decrease with increasing the time (iterations). Referring to Eq. (16), the physical diffusion is a random vector and does not steadily reduce with the increases of the iteration number. Thus, another term (Eq. (17)) is added to Eq. (16). This term linearly decreases the random speed with the time and works on the basis of a geometrical annealing schedule:

$$D_i = D^{max}\left(1 - \frac{I}{I_{max}}\right)\delta \tag{17}$$

2.2.4. Motion Process of the KH Algorithm

In general, the defined motions frequently change the position of a krill individual toward the best fitness. The foraging motion and the motion induced by other krill individuals contain two global and two local strategies. These are working in parallel which make KH a powerful algorithm. According to the formulations of these motions for the ith krill individual, if the related fitness value of each of the abovementioned effective factor (K_j, K^{best}, K^{food} or K_i^{best}) is better (less) than the fitness of the ith krill, it has an attractive effect; otherwise, it has a repulsive effect. It is also clear from the above formulations that a better fitness is more effective on the movement of ith krill individual. The physical diffusion performs a random search in the proposed method. Using different effective parameters of the motion during the time, the position vector of a krill individual during the interval t to $t + \Delta t$ is given by the following equation:

$$X_i(t + \Delta t) = X_i(t) + \Delta t \frac{dX_i}{dt} \tag{18}$$

It should be noted that Δt is one of the most important constants and should be carefully set according to the optimization problem. This is because this parameter works as a scale factor of the speed vector. Δt completely depends on the search space and it seems it can be simply obtained from the following formula:

$$\Delta t = C_t \sum_{j=1}^{NV} (UB_j - LB_j) \tag{19}$$

where NV is the total number of variables, and LB_j and UB_j are lower and upper bounds of the jth variables ($j = 1, 2, \ldots, NV$), respectively. Therefore, the absolute of their subtraction shows the search space. It is empirically found that C_t is a constant number between [0,2]. It is also obvious that low values of C_t let the krill individuals to search the space carefully.

2.3. Genetic operators

To improve the performance of the algorithm, genetic reproduction mechanisms are incorporated into the algorithm. The introduced adaptive genetic reproduction mechanisms are crossover and mutation which are inspired from the classical DE algorithm.

2.3.1. Crossover

The crossover operator is first used in GA as an effective strategy for global optimization. A vectorized version of the crossover is also used in DE which can be considered as a further development to GA. In this study, an adaptive vectorized crossover scheme is employed.

The crossover is controlled by a crossover probability, Cr, and actual crossover can be performed in two ways: (1) binomial and (2) exponential. The binomial scheme performs crossover on each of the d components or variables/parameters. By generating a uniformly distributed random number between 0 and 1, the mth component of X_i, $x_{i,m}$, is manipulated as:

$$x_{i,m} = \begin{cases} x_{r,m} & rand_{i,m} < Cr \\ x_{i,m} & else \end{cases} \tag{20}$$

$$Cr = 0.2\widehat{K}_{i,best} \tag{21}$$

where $r \in \{1,2,\ldots,i-1, i+1,\ldots,N\}$. Using this new crossover probability, the crossover probability for the global best is equal to zero and it increases with decreasing the fitness.

2.3.2. Mutation

The mutation plays an important role in evolutionary algorithms such as ES and DE. The mutation is controlled by a mutation probability (Mu). The adaptive mutation scheme used herein is formulated as:

$$x_{i,m} = \begin{cases} x_{gbes,m} + \mu(x_{p,m} - x_{q,m}) & rand_{i,m} < Mu \\ x_{i,m} & else \end{cases} \tag{22}$$

$$Mu = 0.05/\widehat{K}_{i,best} \tag{23}$$

where p, $q \in \{1,2,\ldots,i-1, i+1,\ldots,K\}$ and μ is a number between 0 and 1. It should be noted In $\widehat{K}_{i,best}$ the nominator is $K_i - K^{best}$. Using this new mutation probability, the mutation probability for the global best is equal to zero and it increases with decreasing the fitness.

2.4. Methodology of the KH algorithm

Various krill-inspired algorithms can be developed by idealizing the motion characteristics of the krill individuals. Generally, the KH algorithm can be introduced by the following steps:

   I. Data Structures: Define the simple bounds, determination of algorithm parameter(s) and etc.
  II. Initialization: Randomly create the initial population in the search space.
 III. Fitness evaluation: Evaluation of each krill individual according to its position.
 IV. Motion calculation:
- Motion induced by the presence of other individuals,
- Foraging motion
- Physical diffusion

  V. Implement the genetic operators
 VI. Updating: updating the krill individual position in the search space.
VII. Repeating: go to step III until the stop criteria is reached.
VIII. End

A basic representation of the KH algorithm is presented in Fig. 2.

2.5. Simulation Using the KH algorithm

As an example, the proposed algorithm is benchmarked using the Peak Function (see Fig. 3). The Peak Function is defined as:

$$Z = X_1 e^{-(X_1^2 + X_2^2)}$$

In this problem, 10 krill individuals are used to find the minima with the time interval equal to 2 second. Different positions of the krill individuals and the virtual food place are shown in Fig. 4. The positions in spread of the individuals after 1st, 5th,
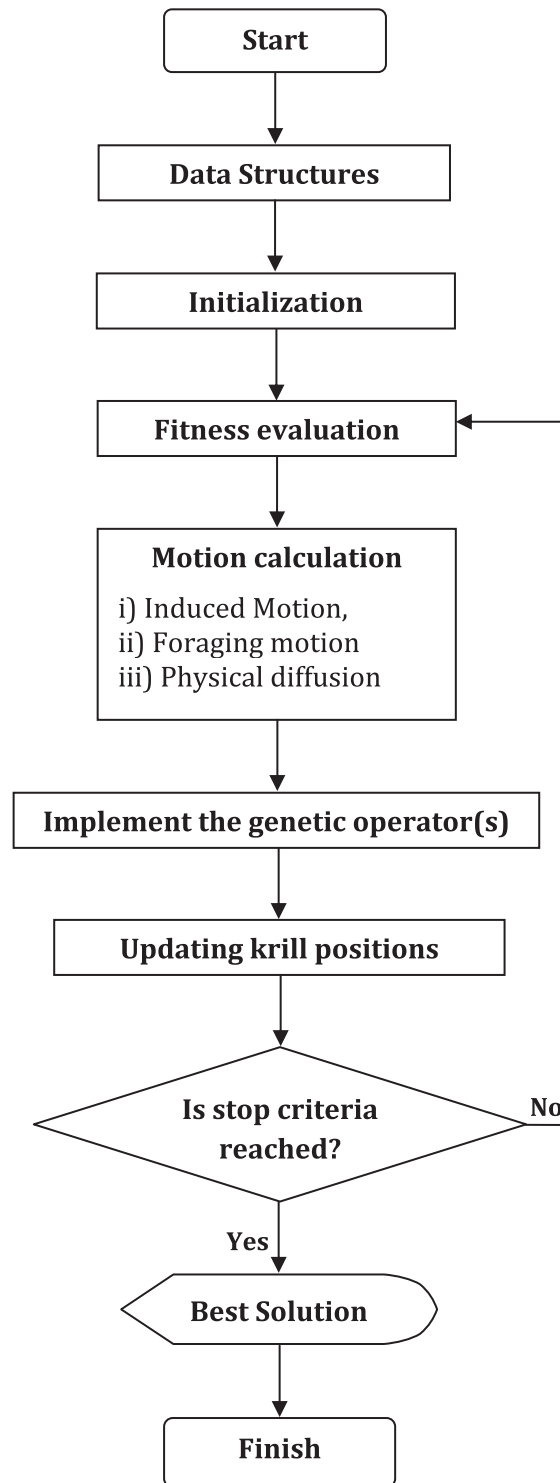
Fig. 2. Simplified flowchart of the krill herd algorithm.

10th and 15th iterations are respectively shown in Fig. 4(a)–(d). The krill individuals are shown by • marks and the virtual food place is shown by ∗ mark. As it is seen, the positions of the krill individuals and the location of the food concentration go toward the global minimum with increasing the iterations.
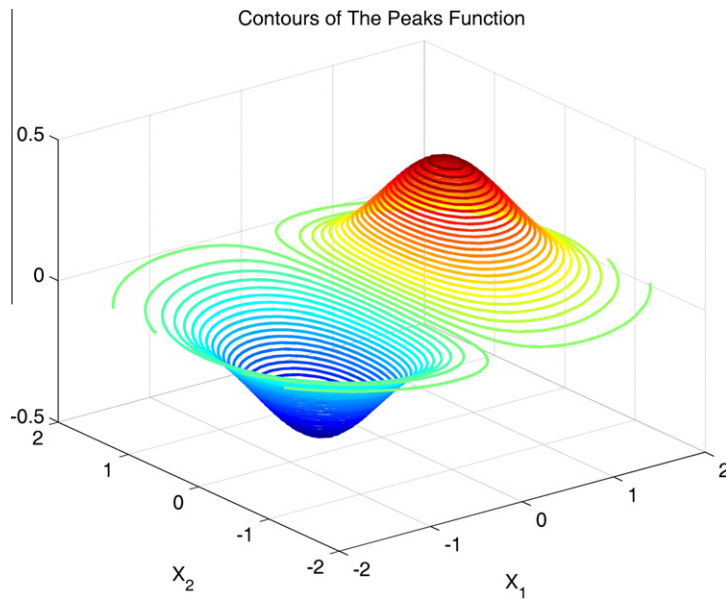
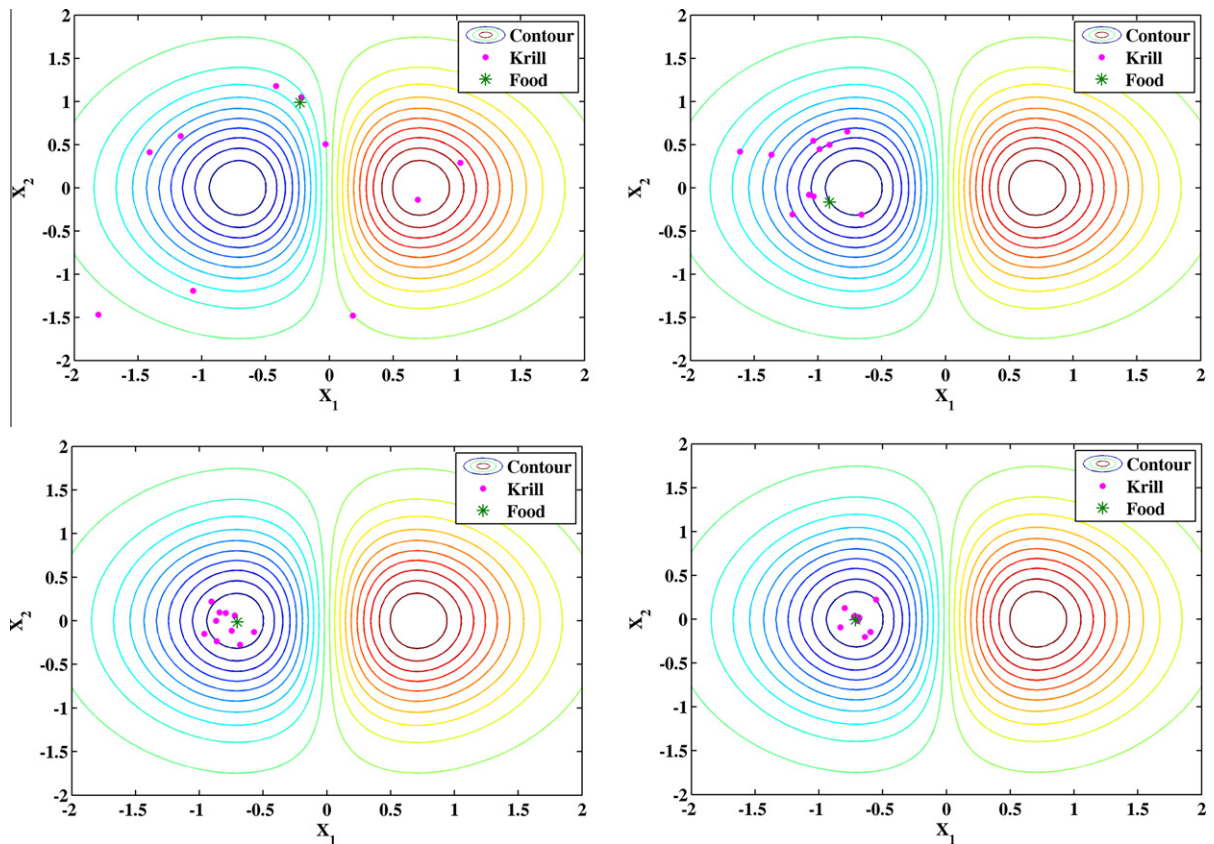**Fig. 3.** 3-D contours of the Peak Function.

Fig. 4. The positions of the krill individuals and food after (a) 1st iteration, (b) 5th iteration, (c) 10th iteration, and (d) 15th iteration.

3. Implementation and numerical experiments

The computational procedures described above have been implemented in MATLAB™ computer program. In order to evaluate the performance of the KH algorithm, it is validated using some well-know benchmark problems obtained from
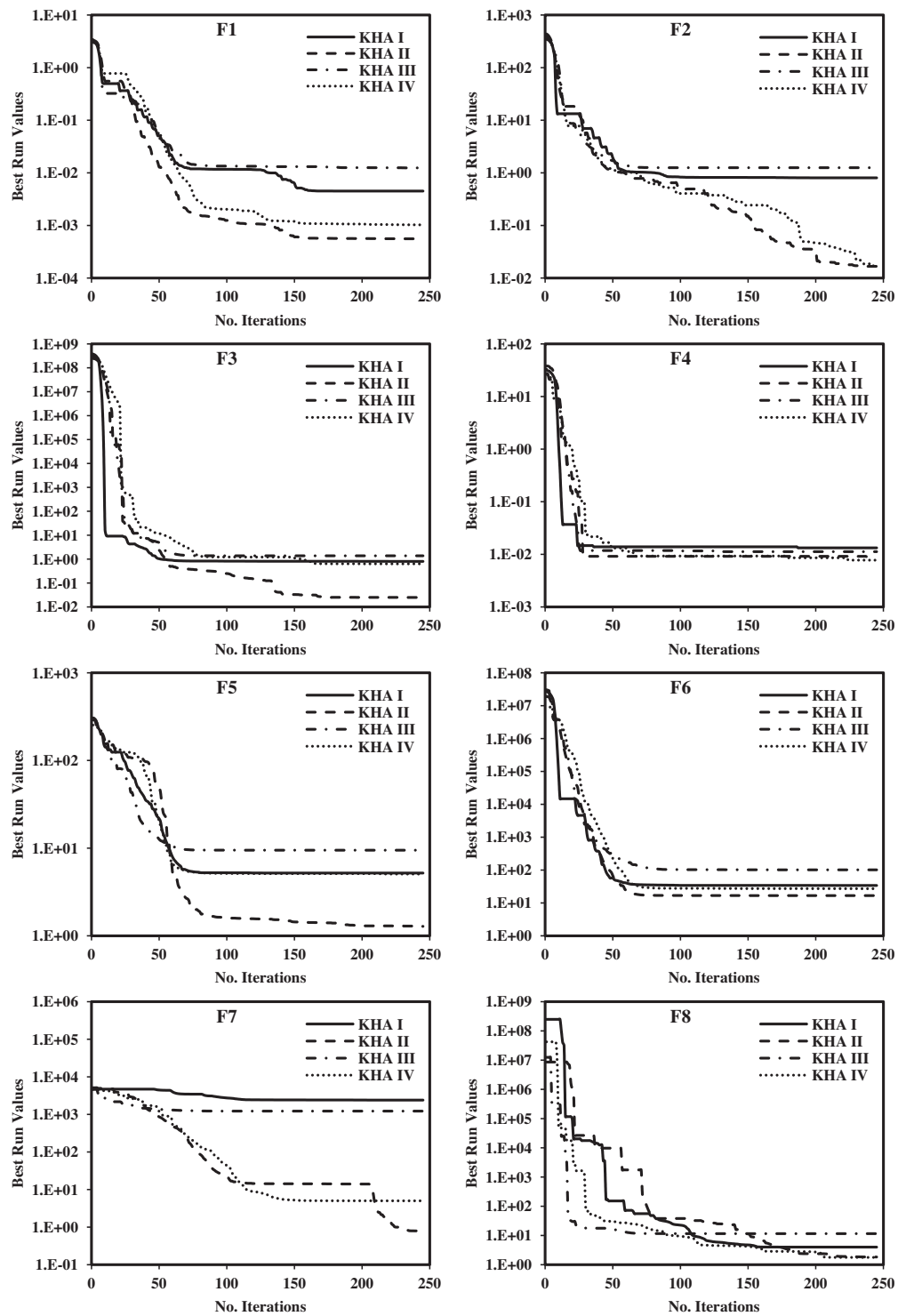
Fig. 5. Comparison of the convergence history of different KH algorithms for the benchmark problems.

the literature [32]. The descriptions of the functions are presented in Appendix A. The benchmarks include both low and high dimension problems, and each of them has ten well-known functions.
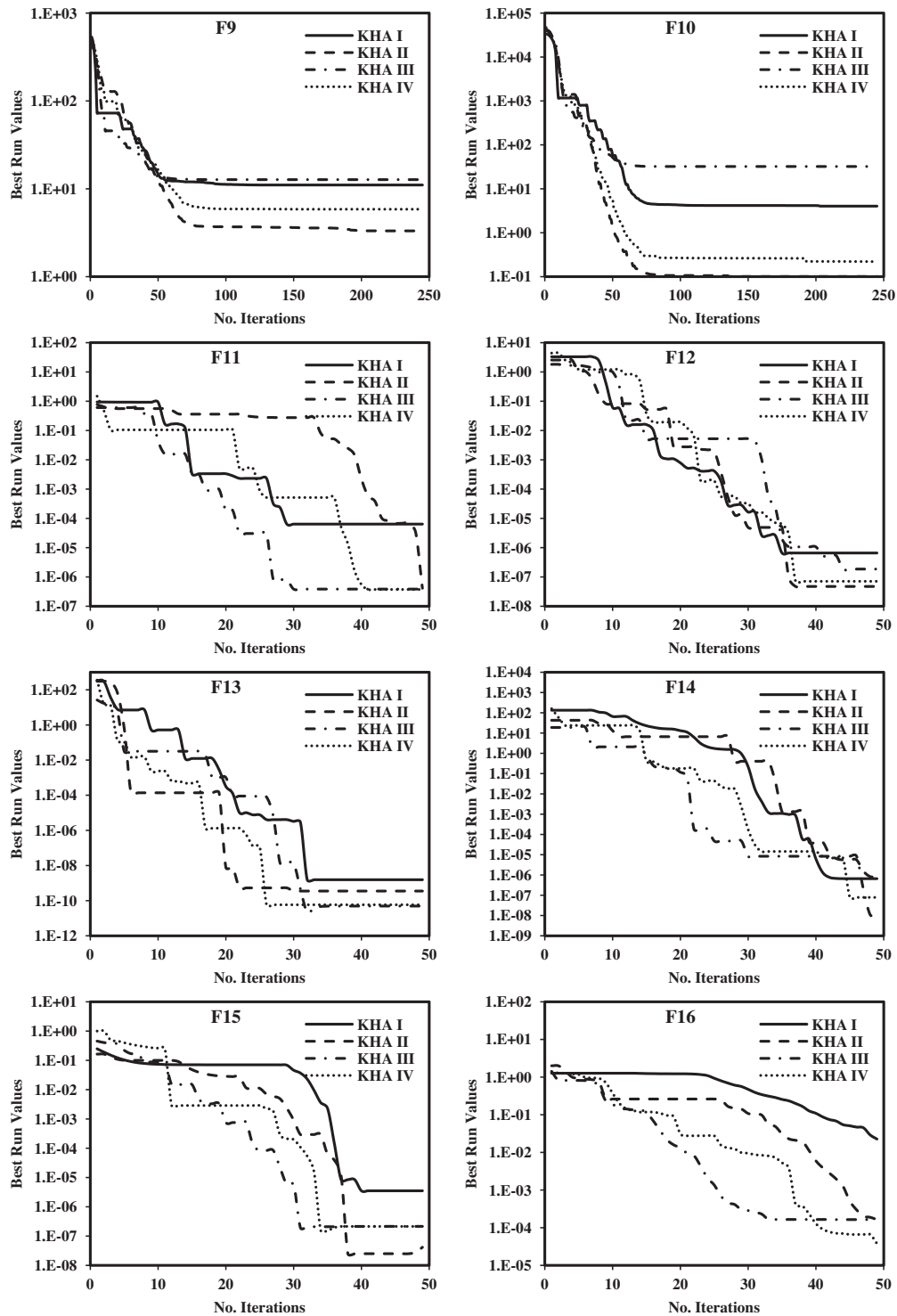
Fig. 5 (*continued*)

For solving the benchmark problems, four different KH algorithms are tested, namely:

  I. KH without any genetic operators (KH I);
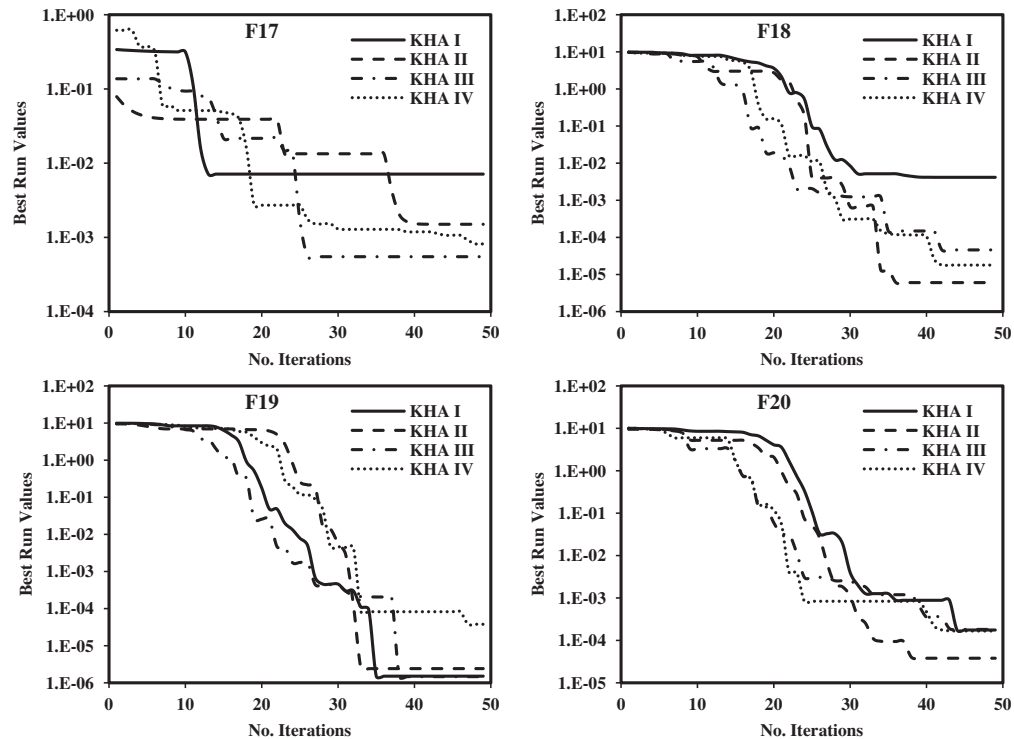 II. KH with crossover operator (KH II);

Fig. 5 (*continued*)

Table 1
Normalized statistical results of KH algorithms and GA, ES, BBO, ACO, DE, HDE, PSO, APSO for the benchmark problems.

ID	KH I	KH II	KH III	KH IV	GA	ES	BBO	ACO	DE	HDE	PSO	APSO
F1	0.995	1.000	0.987	0.999	0.964	0.000	0.989	0.395	0.838	0.998	0.390	1.000
F2	0.989	1.000	0.983	1.000	0.000	0.859	0.906	0.556	0.484	1.000	0.973	1.000
F3	1.000	1.000	1.000	1.000	0.000	0.306	0.967	0.325	0.780	0.997	0.891	1.000
F4	0.968	1.000	0.933	0.969	0.734	0.858	0.940	0.808	0.909	0.301	0.000	0.798
F5	0.865	0.934	0.596	0.892	0.905	0.000	1.000	0.697	0.925	0.958	0.998	0.943
F6	0.320	1.000	0.654	0.999	0.973	0.993	0.990	0.986	0.984	0.000	1.000	0.452
F7	0.989	0.995	0.969	0.995	0.683	0.000	1.000	0.831	0.902	1.000	0.933	1.000
F8	0.690	0.852	0.499	0.783	0.176	0.000	1.000	0.660	0.498	0.999	0.198	0.481
F9	0.721	0.918	0.679	0.853	0.976	0.000	1.000	0.479	0.823	1.000	0.999	1.000
F10	0.999	1.000	0.990	1.000	1.000	0.000	1.000	0.989	0.911	1.000	1.000	1.000
F11	1.000	1.000	1.000	1.000	0.666	0.000	0.666	0.666	0.666	1.000	0.666	1.000
F12	1.000	1.000	1.000	1.000	0.993	0.000	1.000	0.997	1.000	1.000	1.000	1.000
F13	1.000	1.000	1.000	1.000	0.797	0.249	0.969	0.722	0.965	1.000	0.000	1.000
F14	1.000	1.000	1.000	1.000	0.968	0.000	0.995	0.744	0.960	1.000	0.974	1.000
F15	1.000	1.000	1.000	1.000	0.000	0.180	0.740	0.520	0.616	1.000	0.667	1.000
F16	0.990	1.000	1.000	1.000	0.209	0.000	0.534	0.299	0.124	1.000	0.282	1.000
F17	0.966	1.000	1.000	1.000	0.315	0.421	0.974	0.000	0.257	1.000	0.885	0.999
F18	1.000	1.000	1.000	1.000	0.984	0.000	1.000	0.822	0.333	0.999	1.000	1.000
F19	0.999	1.000	0.999	0.998	0.993	0.000	0.999	0.877	0.997	0.983	0.999	1.000
F20	1.000	1.000	1.000	1.000	0.991	0.000	1.000	0.998	0.998	0.993	0.999	1.000
$\sum$	18.491	19.697	18.288	19.487	13.325	3.866	18.669	13.373	14.971	18.229	14.854	18.673
Rank	5	1	6	2	11	12	4	10	8	7	9	3

III. KH with mutation operator (KH III); and
IV. KH with crossover and mutation operators (KH IV).

Owing to the random nature of the KH method and other metaheuristic algorithms, their performance cannot be judged by the result of a single run. Many trials with independent population initializations should be made to obtain a useful conclusion of the performance of the approach. Therefore, in this study the results are obtained in 50 trials.

Table 2
Benchmark problems.

ID	Name	Dimension	Function	Bounds	Global minimum
High-dimensional functions					
F1	Ackley	20	$f(X) = -a \exp\left(-0.02\sqrt{n^{-1}\sum_{i=1}^{n}x_i^2}\right) - \exp\left(n^{-1}\sum_{i=1}^{n}\cos(2\pi x_i)\right) + a + e, a = 20$	$X \in [-32,32]^{20}$	0
F2	Griewank	20	$f(X) = 1 + \frac{1}{4000}\sum_{i=1}^{n}x_i^2 - \prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right)$	$X \in [-600,600]^{20}$	0
F4	Quartic	20	$f(X) = \sum_{i=1}^{n}ix_i^4 + rand$	$X \in [-1.28,1.28]^{20}$	0
F5	Rastrigin	20	$f(X) = 10n + \sum_{i=1}^{n}(x_i^2 - 10\cos(2\pi x_i))$	$X \in [-5.12,5.12]^{20}$	0
F6	Rosenbrock	20	$f(X) = \sum_{i=1}^{n-1}(100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2)$	$X \in [-30,30]^{20}$	0
F7	Schwefel 2.26	20	$f(X) = \sum_{i=1}^{n} -x_i \sin(\sqrt{|x_i|})$	$X \in [-500,500]^{20}$	0
F8	Schwefel 2.22	20	$f(X) = \sum_{i=1}^{n}|x_i| + \prod_{i=1}^{n}|x_i|$	$X \in [-100,100]^{20}$	0
F9	Schwefel 2.21	20	$f(X) = \max\{|x_i|, 1 \leqslant i \leqslant n\}$	$X \in [-100,100]^{20}$	0
	Schwefel 1.2	20	$f(X) = \sum_{i=1}^{30}(\sum_{j=1}^{i}x_j)^2$	$X \in [-100,100]^{20}$	0
F10	Sphere	20	$f(X) = \|X\|, \quad \|X\| = \sqrt{\sum_{i=1}^{n}x_i^2}$	$X \in [-100,100]^{20}$	0
Low-dimensional functions					
F11	Branin	2	$f(X) = a(x_2 - bx_1^2 + cx_1 - d)^2 g(1 - h)\cos(x_1) + g$ $a = 1, \ b = 1.25\pi^{-2}, \ c = 5\pi^{-1}, \ d = 6, \ g = 10, \ h = 0.125\pi^{-1}$	$x_1 \in [-5, 10] x_2 \in [0, 15]$	$5/4\pi$
F12	Camel Back – 6 Hump	2	$f(X) = 4x_1^2 - 2.1x_1^4 + \frac{x_1^6}{3} + x_1 x_2 - 4x_2^2 + 4x_2^4$	$X \in [-5,5]^2$	-1.0316285
F13	De Joung (Shekel's Foxholes)	2	$f(X) = \left(\frac{1}{500} + \sum_{j=1}^{25}\frac{1}{j+\sum_{i=1}^{n}(x_i - a_{ji})^2}\right)$, where $a_{1j} = -32, -16, 0, 16, 32$ for $j = 1, 2, \ldots, 5,$ $a_{1k} = a_{1j}$ for $k = j+5, j+10, j+15, j+20,$ and $j = 1, 2, \ldots, 5,$ $a_{2j} = -32, -16, 0, 16, 32$ for $j = 1, 6, 11, 16, 21,$ $a_{2k} = a_{2j}$ for $k = j+1, j+2, j+3, j+4,$ and $j = 1, 6, 11, 16, 21.$	$X \in [-65.536, 65.536]^2$	1
F14	Goldstein and Price	2	$f(X) = (1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 - 6x_1 x_2 + 3x_2^2)) \times (30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1 x_2 + 27x_2^2))$	$X \in [-2,2]^2$	3
F15	Hartman 1	4	$f(X) = -\sum_{i=1}^{4}c_i \exp(-\sum_{j=1}^{4}a_{ij}(x_j - p_{ij})^2)$ The constants (c, a and p) are presented in Table 3.	$X \in [0,1]^4$	-3.86
F16	Hartman 2	6	$f(X) = -\sum_{i=1}^{4}c_i \exp(-\sum_{j=1}^{6}a_{ij}(x_j - p_{ij})^2)$ The constants (c, a and p) are presented in Table 4.	$X \in [0,1]^6$	-3.32
F17	Kowalik	4	$f(X) = \sum_{i=1}^{11}\left(a_i - \frac{x_1(1+x_2 b_i)}{(1+x_3 b_i + x_4 b_i^2)}\right)^2$ The constants (a and b) are presented in Table 5.	$X \in [-5,5]^4$	3.0748×10^{-4}
F18	Shekel 1	4	$f(X) = -\sum_{i=1}^{5}((X - a_i)(X - a_i)^T + c_i)^{-1}$	$X \in [0,10]^4$	-10.1532
F19	Shekel 2	4	$f(X) = -\sum_{i=1}^{7}((X - a_i)(X - a_i)^T + c_i)^{-1}$	$X \in [0,10]^4$	-10.4029
F20	Shekel 3	4	$f(X) = -\sum_{i=1}^{10}((X - a_i)(X - a_i)^T + c_i)^{-1}$ Constants (c and a) for Shekel's family problems are presented in Table 6	$X \in [0,10]^4$	-10.5364

The maximum number of function evaluations is set to 10,000 for high dimensional functions (F1 − F10), and 1000 for low dimensional functions (F11 − F20).

Here, C_t is set to 0.5 and the inertia weights (ω_n, ω_f) are equal to 0.9 at the beginning of the search to emphasize exploration. These parameters are linearly decreased to 0.1 at the end to encourage exploitation.

3.1. Simulation Results

The convergence plots of the benchmark problems for four different KH algorithms are presented in Fig. 5. It is notable that the global optima of the low-dimensional benchmarks (F11 − F20) are shifted to the origin to draw the semi-logarithmic convergence plots. From these figures, the proposed KH algorithms are able to find the global minima in all cases with high degree of accuracy. As it can be seen in these figures, the performances of KH-I and KH-IV are better than the other ones. This shows the crossover operator is very effective and can clearly improve the performance.

3.2. Comparison

Here, the proposed KH algorithm is compared with some well-known algorithms including ES, GA, DE, BBO, PSO and ACO. The results of these well-know algorithms are presented in [33]. The KH algorithm is classified as a swarm intelligence algorithm and its genetic operators have some similarities with DE. Therefore, the proposed algorithm is further compared

Table 3
Constants for Hartman 1 Problem.

i	c_i	a_{ij}			p_{ij}		
		$j = 1$	2	3	$j = 1$	2	3
1	1	3	10	30	0.3689	0.117	0.2673
2	1.2	0.1	10	35	0.4699	0.4387	0.747
3	3	3	10	30	0.1091	0.8732	0.5547
4	3.2	0.1	10	35	0.03815	0.5743	0.8828

Table 4
Constants for Hartman 2 Problem.

i	c_i	a_{ij}						p_{ij}					
		$j = 1$	2	3	4	5	6	$j = 1$	2	3	4	5	6
1	1	10	3	17	3.5	1.7	8	0.1312	0.1696	0.5569	0.0124	0.8283	0.5886
2	1.2	0.05	10	17	0.1	8	14	0.2329	0.4135	0.8307	0.3736	0.1004	0.9991
3	3	3	3.5	1.7	10	17	8	0.2348	0.1451	0.3522	0.2883	0.3047	0.665
4	3.2	17	8	0.05	10	0.1	14	0.4047	0.8828	0.8732	0.5743	0.1091	0.0381

Table 5
Constants for Kowalik Problem.

i	1	2	3	4	5	6	7	8	9	10	11
a_i	0.1957	0.1947	0.1735	0.16	0.0844	0.0627	0.0456	0.0342	0.0323	0.0235	0.0246
b_i	0.25	0.5	1	2	4	6	8	10	12	14	16

Table 6
Constants for Shekel's family problems.

I	c_i	d_{ij}			
		$j = 1$	2	3	4
1	0.1	4	4	4	4
2	0.2	1	1	1	1
3	0.2	8	8	8	8
4	0.4	6	6	6	6
5	0.4	3	7	3	7
6	0.6	2	9	2	9
7	0.3	5	5	3	3
8	0.7	8	1	8	1
9	0.5	6	2	6	2
10	0.5	7	3.6	7	3.6

with new and robust variants of PSO and DE, namely accelerated PSO (APSO) [1] and hybrid DE with eagle strategy (HDE) [34].

It is worth mentioning that the results have been also obtained after 50 runs with equal number of function evaluations, 10,000 for high-dimensional functions and 1,000 for low-dimensional functions, to the number used for KH algorithms.

One of the effective strategies to perform a comparative study between metaheuristic algorithms is to use the oracle-based view of computation [35]. According to this strategy, the best solution should be found within a certain number of function evaluations. Herein, the best values can be used for comparison because of the equal number of function evaluation for all methods in all cases.

In order to compare the algorithms, the results of best run are normalized between 0 and 1 and therefore the worst and best values of each best solution are changed to 0 and 1, respectively. The normalized results for all cases are presented in Table 1. To achieve a general conclusion based on the oracle-based view, the sum of scores and rank of each algorithm is presented in this table. At first glance, it is clear that the KH algorithms works very well because the KH I to IV have ranks of 5, 1, 6, 2, respectively, among 12 algorithms. Although the new variants of DE and PSO (HDE and APSO) clearly outperform their basic algorithms, they cannot outperform the KH II and IV. The KH algorithm without any genetic operators (KH I) outperforms the other well-known algorithms except for BBO, HDE and APSO. This shows the power of Lagrangian model used here. Comparing all the algorithms, the KH II has the best performance with the score of 19.697 to 20 which confirms the robustness of the proposed algorithm.

4. Summary and conclusion

In the present study, the new KH algorithm, inspired from the krill herding motions, is introduced for solving optimization problems. The three essential actions determining the time-dependent position of an individual are: movement induced by the presence of other individuals, foraging activity, and random diffusion. Each of the first two actions contains a global and a local optimizer. Two global and two local strategies work in parallel resulting in notable efficiency of the KH algorithm. For random searching, the third action is also added to the proposed methodology. To complete this bio-inspired algorithm, the adaptive evolutionary operators (crossover and mutation) are also added to it. The derivative information is not necessary in the KH algorithm because it uses a stochastic random search instead of a gradient search.

For each metaheuristic algorithm, it is important to tune its related parameters. One of interesting parts of the proposed algorithm is that it carefully simulates the krill behavior and it uses the real world empirical studies to obtain the coefficients. Because of this fact, only time interval (C_t) should be fine-tuned in the KH algorithm. This can be considered as a remarkable advantage of the proposed algorithm in comparison with other nature-inspired algorithms. However, it is suggested to employ an efficient meta-optimization technique for fine-tuning of the parameters of the KH algorithm for each problem.

Other features and advantages of the proposed method are as listed below:

- Each agent can contribute to the moving process according to its fitness.
- Each neighbor has an attractive/repulsive effect on the movement of the krill individual. Therefore, these effects can act as a local search for each krill individual.
- Center of food determined according to the fitnesses of all of the krill individuals is regarded as an estimation for the global best.

Despite different advantages of the KH algorithm, its basic aspects, such as induced and foraging motions, are not as simple as the classical swarm intelligence algorithm, PSO.

To validate the proposed method and to evaluate the effect of genetic operators, four KH algorithms are verified using several benchmark problems. The initial testing of the KH algorithms indicates that this method is very encouraging for its future application to optimization tasks. Moreover, the conducted empirical study shows that applying a crossover operator is very effective. Further improvements in the design of the KH algorithm are possible. Future research in this context may focus on the following issues:

- An optimal strategy for choosing the initial krill distribution;
- Changing and improving the basic motions in the KH algorithm. As an instance, the physical diffusion can be defined using other random processes (e.g. Levy walk).
- An efficient strategy for determining the KH parameters, in particular Δt.

The performance of the proposed KH algorithms is compared with that of the GA, ES, BBO, ACO, DE, HDE, PSO, and APSO algorithms. The results show that the KH algorithms can outperform these well-known algorithms.

Appendix A

Tables 2–6.

Reference

[1] Yang X-S. Nature-inspired metaheuristic algorithms. Luniver Press; 2008.
[2] Gandomi AH. Yang X-S, Alavi AH. Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems. Engineering with Computers, in press. DOI 10.1007/s00366-011-0241-y.
[3] Talbi EG. Metaheuristics: from design to implementation. Hoboken, New Jersey, USA: John Wiley & Sons; 2009.
[4] Geem ZW, Kim J-H, Loganathan GV. A new heuristic optimization algorithm: harmony search. Simulation 2001;76(2):60–8.
[5] Kaveh A, Talatahari S. A novel heuristic optimization method: charged system search. Acta Mech 2010;213:267–89.
[6] Zang H, Zhang S, Hapeshi K. A review of nature-inspired algorithms. J Bionic Eng 2010;7:S232–7.
[7] Tang WJ, Wu QH. Biologically inspired optimization: a review. Trans Inst Meas Control 2009;31:495–515.
[8] Goldberg DE. Genetic algorithms in search, optimization and machine learning. Reading, MA: Addison-Wesley; 1989.
[9] Koza JR. Genetic programming: on the programming of computers by natural selection. Cambridge, MA: MIT Press; 1992.
[10] Fogel LJ, Owens AJ, Walsh MJ. Artificial intelligence thorough simulated evolution. Chichester, UK: John Wiley & Sons, Ltd.; 1966.
[11] Storn R, Price KV. Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces. J Global Optim 1997;11(4):341–59.
[12] Khatib W, Fleming P. The stud GA: A mini revolution? In: Eiben A, Back T, Schoenauer M, Schwefel H, editors. Parallel Problem Solving from Nature. New York: Springer; 1998.
[13] Gandomi AH, Alavi AH. Multi-stage genetic programming: a new strategy to nonlinear system modeling. Inf Sci 2011;181(23):5227–39.
[14] Gandomi AH, Alavi AH. A new multi-gene genetic programming approach to nonlinear system modeling. Part I: Materials and structural engineering problems. Neural Comput Appl 2012;21(1):171–87.
[15] Simon D. Biogeography-based optimization. IEEE Trans Evolut Comput 2008;12:702–13.
[16] Eberhart RC, Kennedy J. A new optimizer using particle swarm theory. In: Proceedings of the sixth international symposium on micro machine and human science. Japan: Nagoya; 1995.
[17] Dorigo M, Maniezzo V, Colorni A. The ant system: optimization by a colony of cooperating agents. IEEE Trans Syst Man Cybern B 1996;26(1):29–41.
[18] Yang XS, Gandomi AH. Bat algorithm: a novel approach for global engineering optimization. Eng Comput 2012;29(5).
[19] Passino KM. Biomimicry of bacterial foraging for distributed optimization and control. IEEE Control Syst Mag 2002;22:53–67.
[20] Gregory R, Paton RC, Saunders JR, Wu QH. Parallelising a model of bacterial interaction and evolution. BioSystems 2004;76:121–31.
[21] Vlachos C, Paton RC, Saunders JR, Wu QH. A rule-based approach to the modelling of bacterial ecosystems. BioSystems 2005;84:49–72.
[22] Flierl G, Grunbaum D, Levin S, Olson D. From individuals to aggregations: the interplay between behavior and physics. J Theor Biol 1999;196:397–454.
[23] Okubo A. Dynamical aspects of animal grouping: swarms, schools, flocks, and herds. Adv Biophys 1986;22:1–94.
[24] Hofmann EE, Haskell AGE, Klinck JM, Lascara CM. Lagrangian modelling studies of Antarctic krill (Euphasia superba) swarm formation. ICES J Mar Sci 2004;61:617–31.
[25] Hardy AC, Gunther ER. The plankton of the South Georgia whaling grounds and adjacent waters, 1926–1927. Disc Rep 1935;11:1–456.
[26] Marr JWS. The natural history and geography of the Antarctic krill (Euphausia superba Dana). Disc Rep 1962;32:33–464.
[27] Nicol S. Living krill, zooplankton and experimental investigations. Proceedings of the international workshop on understanding living krill for improved management and stock assessment marine and freshwater behaviour and physiology 2003;36(4):191–205.
[28] Murphy EJ, Morris DJ, Watkins JL, Priddle J. Scales of interaction between Antarctic krill and the environment. In: Sahrhage D, editor. Antarctic Ocean and resources variability. Berlin: Springer-Verlag; 1988. p. 120–30.
[29] Miller DGM, Hampton I. Krill aggregation characteristics: spatial distribution patterns from hydroacoustic observations. Polar Biol 1989;10:125–34.
[30] Price HJ. Swimming behavior of krill in response to algal patches: a mesocosm study. Limnol Oceanogr 1989;34:649–59.
[31] Morin A, Okubo A, Kawasaki K. Acoustic data analysis and models of krill spatial distribution. Scientific Committee for the Conservation of Antarctic Marine Living Resources, Selected Scientific Papers, Part I; 1988. p.311–29.
[32] Yao X, Liu Y, Lin G. Evolutionary programming made faster. IEEE Trans Evolut Comput 1999;3:82–102.
[33] Ma H. An analysis of the equilibrium of migration models for biogeography-based optimization. Inf Sci 2010;180:3444–64.
[34] Gandomi AH, Yang XS, Talatahari S, Deb S. Coupled eagle strategy and differential evolution for unconstrained and constrained global optimization. Comput Math Appl. 2012;63(1):191–200.
[35] Wolpert DH, Macready WG. No free lunch theorems for optimization. IEEE Trans Evol Comput 1997;1(1):67–82.