

# Drift Analysis of Mutation for Biogeography-Based Optimization

Weian Guo<sup>1,2,3</sup>, Lei Wang<sup>1</sup>, Qidi Wu<sup>1</sup>, Shuzhi Sam Ge<sup>2</sup>, Hongliang Ren<sup>3</sup>

1. School of Electronics and Information, Tongji University, Shanghai 201804, China
2. Social Robotics Laboratory, Interactive Digital Media Institute, Electrical and Computer Engineering, National University of Singapore, 119077, Singapore
3. Department of Bioengineering, National University of Singapore, 117575, Singapore.

**Abstract:** As an essential factor of Evolutionary Algorithms (EAs), mutation operator plays an important role to explore the search space, retain the diversity of individuals and break away local optimums. In most standard classic evolutionary algorithms, the mutation operator is independent from the recombination operator. Nevertheless, in Biogeography Based Optimization (BBO), the mutation operator is not only predefined by users, but also relies on the design of the migration models which is a recombination operator. However to date, the relationship between the mutation and migration has never been investigated. To reveal the relationship and evaluate the mutation models, we utilize the drift analysis to investigate the first hitting time of BBO with different migration models. The analysis compares three different kinds of mutation models in mathematics and the conclusion is helpful to design migration models of BBO. Several standard benchmarks are tested to verify the analysis.

**Keywords:** Evolutionary Algorithm, Mutation operator, Migration operator, Biogeography Based Optimization, Drift analysis, First hitting time

## 1. Introduction

Biogeography-Based Optimization (BBO), which is a novel kind of Evolutionary Algorithm (EA) inspired from the science of biogeography, uses migration models in biogeography to solve optimization problems. Despite of just being proposed since several years ago [11], BBO has been researched by a lot of academicians. Simon compared this algorithm with Genetic Algorithm by Markov analysis [1] to illustrate

the characteristics of BBO. Ma compared the performance of BBO with several classical algorithms by numerical simulation to show BBO's dominant ability [12]. Guo investigated migration models of BBO mathematically and the conclusions are helpful to design the models [13]. There are also other studies hybridizing Fuzzy Set [2], Differential Evolution [3] with BBO and achieve a good performance. Besides the theoretical researches, BBO has also been well implemented in several practical applications, including exploring the groundwater possibility [4], solving the different economic load dispatch problems [5] and optimizing a home smart grid scheduling [6], all of which show the huge potential of BBO in optimization.

Like most other Evolutionary Algorithms, BBO has the common fundamental operators, including recombination and mutation. The migration in BBO is the recombination operator, which is analogous to the crossover in Genetic Algorithm, and the mutation in BBO is very similar with the mutation in Genetic Algorithm. In paper [12], Ma introduced six migration models and evaluated the performance by numerical simulation. In paper [13], Guo conducted researches on the migration models in mathematics to demonstrate the effect of migration on BBO's performance. Researches with respect to mutation operator were carried on as well, but only limited to numerical simulations and comparisons, such as studies about Gaussian mutation, Cauchy mutation, Lévy mutation [7] and differential mutation [8]. They are lacking theoretical analysis on the mutation operators which plays an important role in exploring the search space, maintaining the population diversity and breaking away local optimums [9,10]. Given the fact that the mutation operator in BBO is not only predefined by users but also affected by the migration operator, which make the design of mutation operator be more complex, it is explicit that the theoretical research on mutation operator will be helpful to improve algorithm's performance. However to date, the

relationship between the two operators has never been revealed. According to [15] and [19], first hitting time is a crucial indicator to evaluate the mutation operator. Hence we imported the indicator in this paper and use drift analysis to investigate the impact of migration operator on mutation operator.

The remainder of this paper is organized as follows. In Section 2, we briefly introduced the mechanism of BBO and illustrate the migration models and mutation models. In Section 3, the method of drift analysis and our previous work are introduced as preliminary. In Section 4, we analysed the mutation operator with different migration models in linear function, pseudo-modular function and almost positive function respectively. In Section 5, the three functions as above are used in the numerical simulation and analysis of results are enclosed. In addition, without loss of generality, we test the different mutation operator on other fourteen functions which were used in previous work. We ended this paper in Section 6 with conclusions and presented the suggestions for future work.

## **2. Migration Models and Mutation Models of BBO**

Biogeography-Based Optimization mimics the species distribution in nature biogeography. In the science of biogeography, a criterion to judge geographical areas whether are well suited as residences for biological species is habitat suitability index (HSI), which is correlated with the factors including climate, temperature, humidity and topographic features. All the factors that characterize habitability are called suitability index variables (SIVs). Fertile areas with a high HSI tend to have a large number of species, while the barren ones with low HSI have a small number species. Habitats with a high HSI has a low immigration rate and a high emigration rate since they are already nearly saturated with species, while habitats with low HSI has a high

immigration and a low emigration. In BBO, a good solution presents an area with a high HSI and the poor solution is analogous an area with a low HSI. All the features in solutions are considered as SIVs. Good solutions tend to share their features with poor solutions, while the poor solutions accept features from good solutions in a high probability. Meanwhile, poor solutions share their features with good solutions in a low probability. This is very similar to representatives of a species migrating between fertile areas and barren areas. Noted in [11] that term “island” presents an area which is geographically isolated from other habitats, we follow the expression in this paper. A linear model of species distribution in a single habitat is shown in Figure 1, where  $I$  is the maximum possible immigration rate and  $E$  represents the maximum possible emigration rate. The immigration rate and the emigration rate are functions of the number of species in the habitat. For an island, the immigration rate increases and the emigration rate decreases as the number of species increases.

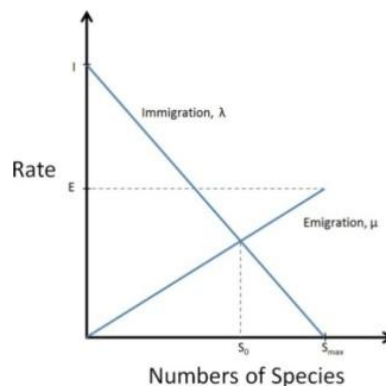


Figure 1 Species model of a single habitat

The migration model has been established in [11], which is stated that there are  $k$  species in the habitat throughout the evolutionary process with immigrants entering the habitat at an immigration rate  $\lambda_k$  and emigrants leaving the habitat at an emigration rate  $\mu_k$ . The largest possible species count that the habitat can support is  $n$ . Simon considered the probability  $P_k$  that the habitat contains exactly  $k$  species. The term

$P_k$  changes from time  $t$  to time  $t+\Delta t$  as (1) [11, 12].

$$P_k(t+\Delta t) = P_k(t)(1 - \lambda_k \Delta t - \mu_k \Delta t) + P_{k-1}(t)\lambda_{k-1} \Delta t + P_{k+1}(t)\mu_{k+1} \Delta t \quad (1)$$

According to [11], (1) holds because in order to have  $k$  species at time  $(t+\Delta t)$ , one of the following conditions must hold: (i) There were  $k$  species at time  $t$ , and no immigration or emigration occurred between  $t$  and  $(t+\Delta t)$ ; or, (ii) There were  $k-1$  species at time  $t$ , and one species immigrated; or, (iii) There were  $k+1$  species at time  $t$ , and one species emigrated. Assuming that  $\Delta t$  is small enough so that the probability of more than one immigration or emigration can be ignored. After taking the limit of (1) as  $\Delta t \rightarrow 0$ , we have

$$\dot{P}_k = \begin{cases} -\lambda_0 P_0 + \mu_1 P_1, & k=0 \\ -(\lambda_k + \mu_k)P_k + \lambda_{k-1}P_{k-1} + \mu_{k+1}P_{k+1}, & 1 \leq k \leq n-1 \\ -\mu_n P_n + \lambda_{n-1}P_{n-1}, & k=n \end{cases} \quad (2)$$

It is noted that (2) is valid for  $\mu_0=0$ ,  $\lambda_n=0$  and  $k=0, \dots, n$ . By defining

$P = [P_0, \dots, P_n]^T$  for notational simplicity, we obtain

$$\dot{P} = AP \quad (3)$$

where

$$A = \begin{bmatrix} -\lambda_0 & \mu_1 & 0 & \dots & \dots & \dots & 0 \\ \lambda_0 & -(\lambda_1 + \mu_1) & \mu_2 & \ddots & \ddots & \ddots & \vdots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \ddots & \ddots & \ddots & \ddots & \lambda_{n-2} & -(\lambda_{n-1} + \mu_{n-1}) & \mu_n \\ 0 & \dots & \dots & \dots & 0 & \lambda_{n-1} & -\mu_n \end{bmatrix} \quad (4)$$

**Theorem 1**<sup>[13]</sup> The steady-state probability of the number of each species is given by

$$P_0 = \left[ 1 + \sum_{j=1}^n \prod_{i=1}^j \left( \frac{\lambda_{i-1}}{\mu_i} \right) \right]^{-1} \quad (5)$$

$$P_k = P_0 \prod_{i=1}^k \left( \frac{\lambda_{i-1}}{\mu_i} \right) = \frac{1}{1 + \sum_{j=1}^k \prod_{i=j}^k \frac{1}{\alpha_i} + \sum_{l=k+1}^n \prod_{i=k+1}^l \alpha_i}, k \in \{1, 2, \dots, n\} \quad (6)$$

where  $\alpha_i = \frac{\lambda_{i-1}}{\mu_i}$ ,  $\lambda_{i-1} \neq 0$  and  $\mu_i \neq 0$  for all  $i$ .

According to the explanation in [11], if a given solution  $K$  has a low probability  $P_k$ , then it is surprising that it exists as a solution. Hence it is likely to mutate to some other solution. Conversely, a solution with high probability is less likely to mutate to a different solution. This can be implemented as a mutation rate that is inversely proportional to the solution probability as shown in (7)<sup>1</sup>.

$$P_{mutation} = m_{max} \left( 1 - \frac{P_s}{P_{max}} \right) \quad (7)$$

where  $m_{max} \in [0,1]$  is a user defined parameter,  $P_s$  is a transition probability with species number  $s$  and  $P_{max} = \max(P_1, P_2, \dots, P_n)$ . According to (7), mutation is relative to the probability transition matrix. Meanwhile according to the design of BBO [11], the probability transition matrix relies on migration model. Hence the mutation rate is a variable affected by migration models rather than a constant. The pseudo-codes of migration and mutation are shown in Figure 2 and Figure 3 respectively.

---

#### Algorithm for BBO Migration

---

```

Select  $H_i$  according to immigration rate  $\lambda_i$ 
If  $H_i$  is selected
    For  $j=1$  to  $n$ 
        Select  $H_j$  according to emigration rate  $\mu_i$ 
        If  $H_j$  is selected
            Replace SIV  $\sigma_i$  in  $H_i$  with  $\sigma_j$  in  $H_j$ 
        End
    End
End
End

```

---

Figure 2 Pseudocodes of Biogeography-Based Optimization Migration

<sup>1</sup> There is a typo in (14) of [11]. Simon shows the correct formula in <http://academic.csuohio.edu/simond/bbo/>.

---

**Algorithm for BBO Mutation**

---

```
For  $j=1$  to  $m$ 
  Use lambda and  $\mu_i$  to compute the probability  $P_i$ 
  Select SIV  $H_i(j)$  with probability  $P_i$ 
  If  $H_i(j)$  is selected
    Replace  $H_i(j)$  with a randomly generated SIV
  End
End
End
```

---

Figure 3 Pseudocodes of Biogeography-Based Optimization Mutation

### 3. Impact Analysis of Migration Rates on Mutation

In most other Evolutionary Algorithms, mutation operator is a predefined by users and it is fixed during the evolutionary progress. However, the mutation operator in Biogeography-Based Optimization is affected by probability transition matrices as shown in (7). Next, we will investigate the relationship between migration operator and mutation operator.

Paper [12] proposed six migration models that are considered as benchmark models in subsequent researches [13]. Based on the simulation results, we investigate three well-performed migration models which are shown as follows and analyse their impact on mutation.

#### **Linear Migration Model:**

$$\lambda_k = I \left( 1 - \frac{k}{n} \right) \quad (8)$$

$$\mu_k = \frac{k}{n} E \quad (9)$$

This model is presented in the original BBO paper [11], and the model is linear functions of species amount in the habitat.

#### **Quadratic Migration Model:**

$$\lambda_k = I \left( \frac{k}{n} - 1 \right)^2 \quad (10)$$

$$\mu_k = E \left( \frac{k}{n} \right)^2 \quad (11)$$

The migration model are convex quadratic functions of the number of species.

**Sinusoidal Migration Model:**

$$\lambda_k = \frac{I}{2} \left( \cos \left( \frac{k\pi}{n} \right) + 1 \right) \quad (12)$$

$$\mu_k = \frac{E}{2} \left( -\cos \left( \frac{k\pi}{n} \right) + 1 \right) \quad (13)$$

The migration model is sinusoidal functions of the number of species. In the next sections, we assume that  $I = E = 1$  for the consistency of previous work [12, 13].

First, we investigate the probability transition matrix.

According to (5), for a problem with a fixed population size, namely  $n$  is fixed,  $P_0$  can be considered as a constant matrix. From Figure 1, we know that the curve of immigration rate  $\lambda$  is decreasing and the curve of emigration rate  $\mu$  is increasing.

**Corollary 1** For linear migration model, quadratic migration model and sinusoidal migration model, we define the maximum steady-state probability as  $P_{\max}$ , then

$P_{\max} = P_j$ , where  $j = \text{ceil} \left( \frac{n+1}{2} \right)$ ,  $\text{ceil}(X)$  rounds  $X$  to the nearest integer towards infinity.

**Proof:** In (6), it is obvious that  $P_k \geq P_{k-1}, k \in \{1, 2, \dots, n\}$  if  $\frac{\lambda_{j-1}}{\mu_j} \geq 1$  and

$P_k < P_{k-1}, k \in \{1, 2, \dots, n\}$ , if  $\frac{\lambda_{j-1}}{\mu_j} < 1$ .

For linear migration model, according to (8) and (9),



$$\frac{\lambda_{j-1}}{\mu_j} = \frac{n+1}{j} - 1 \quad (14)$$

Hence if  $j \leq \frac{n+1}{2}$ ,  $\frac{\lambda_{j-1}}{\mu_j} = \frac{n+1}{j} - 1 \geq 1$  holds. Since  $j$  is an integer and

$j = \max[1, 2, \dots, k]$ , we obtain that  $j = \text{ceil}\left(\frac{n+1}{2}\right)$  to guarantee  $P_{\max} = P_j$ , where

$\text{ceil}(X)$  rounds  $X$  to the nearest integer towards infinity.

For quadratic migration model, according to (10) and (11), we can obtain that

$$\frac{\lambda_{j-1}}{\mu_j} = \left(\frac{n+1}{j} - 1\right)^2 \quad (15)$$

If  $\frac{\lambda_{j-1}}{\mu_j} = \left(\frac{n+1}{j} - 1\right)^2 \geq 1$ , then  $j \leq \frac{n+1}{2}$ . Since  $j$  is an integer and

$j = \max[1, 2, \dots, k]$ , we can obtain that  $j = \text{ceil}\left(\frac{n+1}{2}\right)$ .

For sinusoidal migration model, according to (12) and (13), we can obtain that

$$\frac{\lambda_{j-1}}{\mu_j} = \frac{\cos\left(\frac{(j-1)\pi}{n}\right) + 1}{-\cos\left(\frac{j\pi}{n}\right) + 1} \quad (16)$$

If  $\frac{\lambda_{j-1}}{\mu_j} \geq 1$ , then

$$\begin{aligned} \cos\left(\frac{(j-1)\pi}{n}\right) + 1 &\leq -\cos\left(\frac{j\pi}{n}\right) + 1 \\ \Rightarrow \cos\left(\frac{(j-1)\pi}{n}\right) + \cos\left(\frac{j\pi}{n}\right) &\leq 0 \\ \Rightarrow 2\cos\left(\frac{(2j-1)\pi}{2n}\right)\cos\left(\frac{-\pi}{2n}\right) &\leq 0 \end{aligned} \quad (17)$$

Since  $2\cos\left(\frac{-\pi}{2n}\right) \geq 0$ , we know that  $\cos\left(\frac{(2j-1)\pi}{2n}\right) \leq 0$ . Since  $j \in [1, 2, \dots, n]$ , we

know  $j \leq \frac{n+1}{2}$ . Since  $j$  is an integer and  $j = \max[1, 2, \dots, k]$ , we can obtain that

$$j = \text{ceil}\left(\frac{n+1}{2}\right).$$

For the sake of comparison visually, we set the population size as 50, maximum mutation rate as 0.05, and draw the mutation probability with the three migration models in Figure 4. In Figure 4, it is obvious that even users predefine the same maximum mutation rate, the mutation models are different due to the differing migration models. Up to now, the analysis of mutation operator in Biogeography-based Optimization has never been discussed, which will be the focus of this paper.

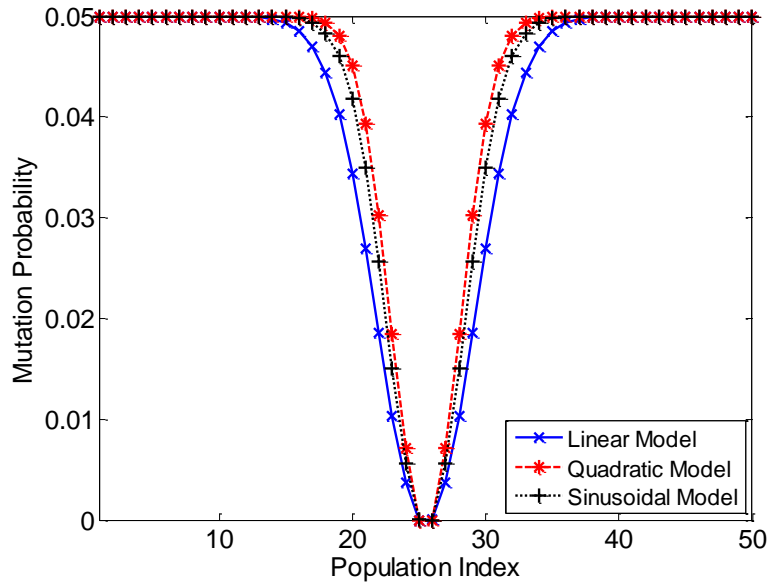


Figure 4 Mutation Probabilities in Different Migration Models

**Lemma 1** (Jordan Inequality) If  $0 \leq |x| < \frac{\pi}{2}$ , the following equality holds,

$$\frac{2}{\pi} \leq \frac{\sin x}{x} < 1. \quad (18)$$

**Theorem 2** By defining  $p_{lm}$  as the mutation probability in linear migration model,  $p_{qm}$  as the mutation probability in quadratic migration model and  $p_{sm}$  as the mutation probability in sinusoidal migration model, the following equality holds,

$$p_{qm}(k) \geq p_{sm}(k) \geq p_{lm}(k). \quad (19)$$

where  $k$  denotes the  $k$ th individual in population

**Proof:** According to Corollary 1 and the definition in (7), we know that

$$p_{mutation} = m_{\max} \left( 1 - \frac{P_k}{P_{\max}} \right) = m_{\max} \left( 1 - \frac{P_0 \prod_{i=1}^k \left( \frac{\lambda_{i-1}}{\mu_i} \right)}{P_0 \prod_{i=1}^j \left( \frac{\lambda_{i-1}}{\mu_i} \right)} \right) = \begin{cases} m_{\max} \left( 1 - \prod_{i=j}^k \left( \frac{\lambda_{i-1}}{\mu_i} \right) \right), & \text{if } k \geq j \\ m_{\max} \left( 1 - \frac{1}{\prod_{i=k}^j \left( \frac{\lambda_{i-1}}{\mu_i} \right)} \right), & \text{if } k < j \end{cases} \quad (20)$$

First, we discuss the case that  $k \geq j$ . We define  $\alpha_k = \prod_{i=j}^k \left( \frac{\lambda_{i-1}}{\mu_i} \right)$ , then

$$\alpha_{k,quadratic} = \prod_{i=j}^k \left( \frac{\left( \frac{i-1}{n} - 1 \right)^2}{\left( \frac{i}{n} \right)^2} \right), \quad \alpha_{k,sinusoidal} = \prod_{i=j}^k \left( \frac{\cos \left( \frac{(i-1)\pi}{n} \right) + 1}{-\cos \left( \frac{i\pi}{n} \right) + 1} \right).$$

$$\frac{\alpha_{k,quadratic}}{\alpha_{k,sinusoidal}} = \prod_{i=j}^k \left( \frac{\left( \frac{\left( \frac{i-1}{n} - 1 \right)^2}{\left( \frac{i}{n} \right)^2} \right)}{\left( \frac{\cos \left( \frac{(i-1)\pi}{n} \right) + 1}{-\cos \left( \frac{i\pi}{n} \right) + 1} \right)} \right) = \prod_{i=j}^k \left( \left( \frac{n+1-i}{i} \right)^2 \left( \frac{1 - \cos \left( \frac{i\pi}{n} \right)}{1 + \cos \left( \frac{(i-1)\pi}{n} \right)} \right) \right) \quad (21)$$

$$= \prod_{i=j}^k \left( \left( \frac{n+1}{i} - 1 \right) \frac{\sin \left( \frac{i\pi}{2n} \right)}{\cos \left( \frac{(i-1)\pi}{2n} \right)} \right)^2 = \prod_{i=j}^k \left( \left( \frac{n+1}{i} - 1 \right) \frac{\sin \left( \frac{i\pi}{2n} \right)}{\sin \left( \frac{(n+i-1)\pi}{2n} \right)} \right)^2$$

According to Lemma 1, we know that

$$\begin{aligned}
\frac{\alpha_{k,quadratic}}{\alpha_{k,sinusoidal}} &\leq \prod_{i=j}^k \left( \left( \frac{n+1}{i} - 1 \right) \frac{\frac{i\pi}{2n}}{\frac{2(n+i-1)\pi}{\pi \cdot 2n}} \right)^2 \\
&= \prod_{i=j}^k \left( \left( \frac{n+1}{i} - 1 \right) \frac{\frac{i\pi}{2n}}{\frac{2(n+i-1)\pi}{\pi \cdot 2n}} \right)^2 = \prod_{i=j}^k \left( \frac{\pi}{2} \left( \frac{n+1-i}{n+i-1} \right) \right)^2
\end{aligned} \tag{22}$$

Since  $j \leq i \leq n$ , the following inequalities hold,

$$\begin{aligned}
\frac{\alpha_{k,quadratic}}{\alpha_{k,sinusoidal}} &\leq \prod_{i=j}^k \left( \frac{\pi}{2} \left( \frac{n+1-i}{n+i-1} \right) \right)^2 \\
&\leq \prod_{i=j}^k \left( \frac{\pi}{2} \left( \frac{n+1-\frac{n+1}{2}}{n+\frac{n+1}{2}-1} \right) \right)^2 = \prod_{i=j}^k \left( \frac{\pi}{2} \left( \frac{n+1}{3n-1} \right) \right)^2 \\
&\leq \prod_{i=j}^k \left( \frac{\pi}{6} \right)^2 \leq 1
\end{aligned} \tag{23}$$

Therefore,  $\alpha_{k,quadratic} \leq \alpha_{k,sinusoidal}$ . Considered (20), we know  $p_{qm}(k) \geq p_{sm}(k)$  when  $k \geq j$ .

Second, we discuss the case that  $k < j$ . We define  $\beta_k = \prod_{i=k}^j \left( \frac{\lambda_{i-1}}{\mu_i} \right)$ , then

$$\beta_{k,quadratic} = \prod_{i=k}^j \left( \frac{\left( \frac{i-1}{n} - 1 \right)^2}{\left( \frac{i}{n} \right)^2} \right), \quad \beta_{k,sinusoidal} = \prod_{i=j}^k \left( \frac{\cos\left(\frac{(i-1)\pi}{n}\right) + 1}{-\cos\left(\frac{i\pi}{n}\right) + 1} \right). \text{ Then we obtain}$$

$$\begin{aligned}
\frac{\beta_{k,quadratic}}{\beta_{k,sinusoidal}} &= \prod_{i=k}^j \left( \frac{\left( \frac{\left( \frac{i-1}{n} - 1 \right)^2}{\left( \frac{i}{n} \right)^2} \right)}{\left( \frac{\cos\left( \frac{(i-1)\pi}{n} \right) + 1}{-\cos\left( \frac{i\pi}{n} \right) + 1} \right)} \right) = \prod_{i=k}^j \left( \frac{n+1}{i} - 1 \right) \frac{\cos\left( \frac{(n-i)\pi}{2n} \right)}{\cos\left( \frac{(i-1)\pi}{2n} \right)} \right)^2 \\
&= \prod_{i=k}^j \left( \frac{n+1}{i} - 1 \right) \frac{\sin\left( \frac{(2n-i)\pi}{2n} \right)}{\sin\left( \frac{(i-1+n)\pi}{2n} \right)} \right)^2 \geq \prod_{i=k}^j \left( \frac{n+1}{i} - 1 \right) \frac{2\pi \left( \frac{(2n-i)\pi}{2n} \right)}{(i-1+n)\pi} \right)^2 \quad (24) \\
&= \prod_{i=k}^j \left( \frac{4n^2 + 2i^2 - 6ni - 2i + 4n}{(i^2 - i + ni)\pi} \right)^2
\end{aligned}$$

When  $1 \leq i < j$ , the following inequalities hold,

$$\frac{\beta_{k,quadratic}}{\beta_{k,sinusoidal}} \geq \prod_{i=k}^j \left( \frac{4n^2 + 2i^2 - 6ni - 2i + 4n}{(i^2 - i + ni)\pi} \right)^2 \geq \prod_{i=k}^j \left( \frac{4n-2}{\pi} \right)^2 \quad (25)$$

When  $n > 1$ ,  $\frac{\beta_{k,quadratic}}{\beta_{k,sinusoidal}} \geq 1$ , which means  $\beta_{k,quadratic} \geq \beta_{k,sinusoidal}$ . Considered (20), we

know  $p_{qm}(k) \geq p_{sm}(k)$  when  $k < j$ .

To sum up, we know that  $p_{qm}(k) \geq p_{sm}(k)$ . Similarly, we can prove that

$$p_{sm}(k) \geq p_{lm}(k).$$

## 4. Drift Analysis of Mutation Rates in BBO

### 4.1 Drift Analysis

Assume that  $x^*$  is an optimal point, and let  $d(x, x^*)$  be the distance between a point  $x$  and  $x^*$ . If there exists more than one optimal point, we define the optimal

solutions as a set  $S^*$  and  $d(x, S^*) = \min\{d(x, x^*): x^* \in S^*\}$  as the distance between  $x$  and the set  $S^*$ . For the sake of convenience, we use  $d(x)$  to denote the distance. It is obvious that  $d(x^*) = 0$  and  $d(x) > 0$  for any  $x \notin S^*$ . Assume that  $\xi_k$  is the population at generation  $k$ . The sequence  $\{d(\xi_k); k = 0, 1, 2, \dots\}$  is a random sequence and can be modelled by a homogeneous Markov chain if adaptive strategy is used [14].

**Definition 1** Given a population  $X = \{x, \dots, x_N\}$ , let  $d(X) = \min\{d(x): x \in X\}$  be the distance of the population to the optimal solution.

**Definition 2** The drift of the random sequence  $\{d(\xi_k); k = 0, 1, 2, \dots\}$  at generation  $k$  is defined by  $\Delta(d(\xi_k)) = d(\xi_{k+1}) - d(\xi_k)$ .

**Definition 3** Define the stopping time as  $\tau = \min\{k: d(\xi_k) = 0\}$ , which is the first hitting time on the optimal solution.

**Condition 1** Let  $d_{\max} = \{d(x): x \in S\}$ , and the interval  $[0, d_{\max}]$  be divided into  $L+1$  sub-intervals:  $d_0 \equiv 0 < d_1 < \dots < d_L < d_{L+1} \equiv d_{\max}$ , where  $L > 0$  is an integer.

(a) For any  $l (0 \leq l \leq L)$ , if at time  $k$ , the population  $\xi_k$  enters the interval  $[0, d_l]$ , i.e.  $d(\xi_k) \leq d_l$ , then after that time

(b) At any time  $k$ , if the population  $\xi_k$  is in the interval  $(d, d_l]$ , then the drift satisfies the following inequality,

$$E[d(\xi_k) - d(\xi_{k+1}) | d_l < d(\xi_k) \leq d_{l+1}] \geq \frac{1}{h_l(n)} \quad (26)$$

where  $h_l(n) > 0$ .

**Theorem 3**<sup>[15]</sup> If series  $\{d(\xi_k); k = 0, 1, 2, \dots\}$  meet Condition 1, then starting from any

initial population  $\xi_0$  with  $d(\xi_0) > 0$ ,

$$E[\tau | d(\xi_0) > 0] \leq \sum_{l=0}^L h_l(n)(d_{l+1} - d_l) \quad (27)$$

**Condition 2** Let  $d_{\max} = \{d(x) : x \in S\}$ , and the interval  $[0, d_{\max}]$  be divided into  $L+1$  subintervals:  $d_0 \equiv 0 < d_1 < \dots < d_L < d_{L+1} \equiv d_{\max}$ , where  $L > 0$  is an integer. At any time  $k \geq 0$ , if the population  $\xi_k$  is in the interval  $(d_l, d_{\max}]$ , the drift satisfies:

$$E[d(\xi_k) - d(\xi_{k+1}) | d(\xi_k) \geq d_l] \geq \frac{1}{h_l(n)} \quad (28)$$

where  $h_l(n) > 0$ .

**Theorem 4** <sup>[15]</sup> If series  $\{d(\xi_k); k = 0, 1, \dots\}$  meet Condition 2, then the starting from any initial population  $\xi_0$  with  $d(\xi_0) > 0$ ,

$$E[\tau | d(\xi_0) > 0] \leq \sum_{l=0}^L h_l(n)(d_{l+1} - d_l) \quad (29)$$

**Condition 3** For population  $X$ ,  $d(X) \geq h_0(n)$ , where  $h_0(n) > 0$  is a function of problem size  $n$ .

**Condition 4** There exists a polynomial function,  $h_1(n) > 0$ , of problem size  $n$  such that

$$E[d(\xi_k) - d(\xi_{k+1}) | \xi_k = X] \leq \frac{1}{h_1(n)} \quad (30)$$

For any generation  $k \geq 0$  and population  $X$  with  $d(X) > 0$ .

**Theorem 5** <sup>[15]</sup> If series  $\{d(\xi_k); k = 0, 1, \dots\}$  meet Condition 3 and Condition 4, then starting from the initial population with  $d(X) \geq h_0(n)$ , the expected value of the first hitting time is shown as follows,

$$E[\tau | \xi_0 = X] \geq h(n) \quad (31)$$

where  $h(n) = h_0(n)h_1(n)$  is a function of problem size  $n$ .

## 4.2 Drift Analysis for Linear Functions

A function  $F$  is linear if  $F(X) = c_0 + \sum_{i=1}^n c_i x_i$  where coefficients  $c_i \in R$  and

[16]. Here we suppose  $X$  is in a binary space. For maximum problem, if  $c_i \geq 0$ , then it is clear that  $x^* = (1, 1, 1, \dots, 1)$  is the only optimal solution. For the sake of convenience, we set  $c_0 = 0$  and  $c_1 = c_2 = \dots = c_n = 1$ , then

$$F(X) = \sum_{i=1}^n x_i. \quad (32)$$

**Theorem 6** The expected first hitting time of BBO for function (32) is

$$E[\tau] \leq \sum_{l=0}^L \frac{(1 - p_{\text{mutation}})^{n-1}}{(1 + p_{\text{mutation}})^n - 1} \text{ when the mutation rate is } p_{\text{mutation}} = \frac{1}{n}, \text{ where } 0 \leq l \leq n.$$

**Proof:** Obviously,  $F(X)$  only takes  $n+1$  different values  $a_0, a_1, a_2, \dots, a_n$ , where  $a_i = i$ . For (32), the maximum solution is  $x^* = (1, 1, \dots, 1)$ . For any solution  $x \neq x^*$ , let  $l$ , where  $0 < l \leq n$ , be the number of zero valued bits in  $x$ , then we obtain

$$\begin{aligned} E[d(\xi_k) - d(\xi_{k+1}) | \xi_k = X] &= \sum_{k=1}^l (a_l - a_{l-k}) P[d(\xi_{k+1}) = a_{l-k} | d(\xi_k) = a_l] \\ &\quad + \sum_{k=1}^{n-l} (a_l - a_{l+k}) P[d(\xi_{k+1}) = a_{l+k} | d(\xi_k) = a_l] \end{aligned} \quad (33)$$

Let  $m$  be the number of zero valued bits flip to one and at the meanwhile time the other  $n-m$  bits keep unchanged. Then we can obtain

$$\begin{aligned} E[d(\xi_k) - d(\xi_{k+1}) | \xi_k = X] &\geq \sum_{m=1}^l m \binom{l}{m} (p_{\text{mutation}})^m (1 - p_{\text{mutation}})^{n-m} \\ &\quad + \sum_{m=1}^{n-l} (-m) \binom{n-l}{m} (p_{\text{mutation}})^m (1 - p_{\text{mutation}})^{n-m} \end{aligned} \quad (34)$$



Since the fact that  $\sum_{m=0}^n \binom{n}{m} a^m = (1+a)^n$  exists,

$$\begin{aligned}
E\left[d(\xi_k) - d(\xi_{k+1}) \mid \xi_k = X\right] &\geq \left((1 + p_{\text{mutation}})^n - np_{\text{mutation}}\right) \sum_{m=1}^l m(1 - p_{\text{mutation}})^{n-m} \\
&\quad - \left((1 + p_{\text{mutation}})^n - np_{\text{mutation}}\right) \sum_{m=1}^{n-l} m(1 - p_{\text{mutation}})^{n-m} \\
&= \left((1 + p_{\text{mutation}})^n - np_{\text{mutation}}\right) \sum_{m=1}^{\lfloor 2l-n \rfloor} m(1 - p_{\text{mutation}})^{n-m} \\
&\geq \frac{(1 + p_{\text{mutation}})^n - np_{\text{mutation}}}{(1 - p_{\text{mutation}})^{n-1}}
\end{aligned} \tag{35}$$

Since  $p_{\text{mutation}} = \frac{1}{n}$ ,

$$E\left[d(\xi_k) - d(\xi_{k+1}) \mid \xi_k = X\right] \geq \frac{(1 + p_{\text{mutation}})^n - 1}{(1 - p_{\text{mutation}})^{n-1}} \tag{36}$$

According to Theorem 4, we know that  $E[\tau \mid d(\xi_0) > 0] \leq \sum_{l=0}^L \frac{(1 - p_{\text{mutation}})^{n-1}}{(1 + p_{\text{mutation}})^n - 1}$ .

**Corollary 2** By defining  $m_{\max} = \frac{1}{n}$  in (7), a large  $p_{\text{mutation}}$  can help BBO reduce the upper bound of expected value of first hitting time for linear functions.

**Corollary 3** To solve the Linear Functions, the design of mutation in Quadratic Model is superior to that in Sinusoidal Model, and the design of mutation in Sinusoidal Model is superior to that in Linear Model.

**Proof:** According to Theorem 2, Theorem 6 and Corollary 2, Corollary 3 holds.

### 4.3 Drift Analysis for Pseudo-modular Functions

A function  $F : X \rightarrow R$  is pseudo-modular if the following inequalities (37) and (38) are met,

$$\min\{F(Y), F(Z)\} \leq \max\{F(Y \wedge Z), F(Y \vee Z)\} \tag{37}$$

$$\max \{F(Y), F(Z)\} \geq \min \{F(Y \wedge Z), F(Y \vee Z)\} \quad (38)$$

for all  $Y, Z \in X$  [16]. An example of this kind function is shown in (39).

$$F(X) = \sum_{i=1}^n \prod_{j=1}^i x_j \quad (39)$$

Without loss of generality, we consider pseudo-modular function as a fitness function as follows,

$$\text{Maximize } F(X) = \sum_{i=1}^n \prod_{j=1}^i x_j \quad (40)$$

where  $X$  is in a binary space.

**Theorem 7** The expected first hitting time of BBO for function (40) is  $E[\tau] \leq p_{mutation}^{-1} (1 - p_{mutation})^{-n+l+1}$ , where  $0 \leq l \leq n$ , and  $p_{mutation}$  is the mutation rate.

**Proof :** For (40), the optimal solution is  $(1, 1, \dots, 1)$ . Define the distance function  $d(x)$  as follows:

$$d(x) = n - \sum_{i=1}^n \prod_{j=1}^i x_j \quad (41)$$

We divide the whole interval  $[0, n]$  into  $n$  subintervals  $d_0 < d_1 < \dots < d_n$  where  $d_l = l$  for  $0 \leq l \leq n$ . Since we adopt the elitism strategy, Condition 1(a) holds automatically. At any generation  $k \geq 0$ , if the fitness of population  $\xi_k$  is in the interval  $(d_l, d_{l+1}]$ , where  $0 \leq l \leq n-1$ , then there exists at least one individual  $x$  in  $\xi_k$  such that  $d(x) = l+1$ . The probability of  $x$  becoming better is no less than  $p_{mutation} (1 - p_{mutation})^{n-l-1}$ . Hence,

$$E[d(\xi_k) - d(\xi_{k+1})] \geq p_{mutation} (1 - p_{mutation})^{n-l-1} \quad (42)$$

According to Theorem 3, we have

$$E[\tau | \xi_0] \leq p_{mutation}^{-1} (1 - p_{mutation})^{-n+l+1} \quad (43)$$

**Proposition 1** By defining  $m_{\max} = \frac{1}{n}$  in (7), a large  $p_{mutation}$  can help BBO reduce the upper bound of expected value of first hitting time.

**Proof:** Let  $G = p_{mutation} (1 - p_{mutation})^{n-l-1}$ , we can obtain

$$\frac{\partial G}{\partial p_{mutation}} = (1 - p_{mutation})^{n-l-2} [1 - (n-l)p_{mutation}] \quad (44)$$

Since  $p_{mutation} \leq m_{\max} = \frac{1}{n} < 1$ , we know that  $(1 - p_{mutation})^{n-l-2} \geq 0$  and

$1 - (n-l)p_{mutation} > 0$ . Hence,

$$\frac{\partial G}{\partial p_{mutation}} \geq 0 \quad (45)$$

Therefore,  $G$  increases as  $p_{mutation}$  increases. According to (43), we know

$$E[\tau | \xi_0] \leq \frac{1}{G} \quad (46)$$

Hence a large  $p_{mutation}$  can help BBO reduce the upper bound of expected value of first hitting time.

**Corollary 4** To solve the Pseudo-modular Functions, the design of mutation in Quadratic Model is superior to that in Sinusoidal Model, and the design of mutation in Sinusoidal Model is superior to that in Linear Model.

**Proof:** According to Theorem 2 and Proposition 1, Corollary 4 holds.

#### 4.4 Drift Analysis for Almost Positive Functions

A function  $F : X \rightarrow R$  is almost positive if the coefficients of all nonlinear terms are non-negative [16]. An example is shown as in (47).

$$F(X) = n - \sum_{i=1}^n x_i + (n+1) \prod_{i=1}^n x_i \quad (47)$$

For this problem, we define  $d(x) = \sum_{i=1}^n |x_i - 1|$ .

We consider this problem as a maximum optimization problem as follows,

$$\text{Maximize } F(X) = n - \sum_{i=1}^n x_i + (n+1) \prod_{i=1}^n x_i \quad (48)$$

where  $X$  is in a binary space.

**Theorem 8** The expected first hitting time of Biogeography-Based Optimization for almost positive function (48) is  $E[\tau] = p_m^{-n}$  with mutation rate  $p_m$  and  $d(\xi_0) = n$ .

**Proof:** For the problem (48), the optimal solution is  $(1,1,\dots,1)$ . For solution  $(0,0,\dots,0)$ , it is the second maximum point, but it is the farthest from the optimal solution with  $d(x) = n$ . First we define the initial solution as  $(0,0,\dots,0)$  since its maximum drift distance. Meanwhile we define  $h_0(n) = n$ , then Condition 3 holds. Second, at any generation  $k \geq 0$ , except that we obtain the solution  $(1,1,\dots,1)$ , solution  $(0,0,\dots,0)$  will be kept as the best solution because of the elitism strategy. Hence only two events may happen:  $I\{d(\xi_{k+1}) = n\}$  and  $I\{d(\xi_{k+1}) = 0\}$ . Therefore, we can obtain

$$E[d(\xi_k) - d(\xi_{k+1}) | d(\xi_k) = n] \leq (n-0)(p_m)^n = n(p_m)^n \quad (49)$$

Let  $h_1(n) = n^{-1} p_m^{-n}$ , the Condition 4 holds. According to Theorem 5, we have

$$E[\tau] \geq n \times n^{-1} p_m^{-n} = p_m^{-n} \quad (50)$$

**Corollary 5** To solve the Almost Positive Functions, large  $p_m$  can help Biogeography-Based Optimization reduces the expected value of first hitting time.

**Proof:** From (50), it is obvious that a large  $p_m$  is helpful to reduce the lower bound of expected value of first hitting time.

**Corollary 6** To solve the Almost Positive Functions, the design of mutation in Quadratic Model is superior to that in Sinusoidal Model, and the design of mutation in Sinusoidal Model is superior to that in Linear Model.

**Proof:** According to Theorem 2 and Corollary 5, Corollary 6 holds.

## 5. Numerical Simulation and Comparisons

In this section, we use the numerical simulation to verify the analysis. Based on our analysis, higher mutation rate will help BBO obtain better solutions. The following functions are considered,

$$F(X) = \sum_{i=1}^n x_i \quad (51)$$

$$F(X) = \sum_{i=1}^n \prod_{j=1}^i x_j \quad (52)$$

$$F(X) = n - \sum_{i=1}^n x_i + (n+1) \prod_{i=1}^n x_i \quad (53)$$

They are linear function, pseudo-modular function and almost positive function respectively. To search the maximum value of the three functions, we define the dimension is 1000, we use linear migration model, quadratic migration model and sinusoidal migration models to test the performance. We ran the simulation 50 times and for each time we ran 1000 generations. There is no any recombination in evolution but only mutation. The numerical simulation results are shown in Table 1 and

Table 2. From Table 1 and

Table 2, it is obvious that Quadratic Model performs better than Sinusoidal Model and Sinusoidal Model is prior to Linear Model.

Table 1 Comparison of numerical simulation results of best value of Sinusoidal Model, Quadratic Model and Linear Model for linear function, pseudo-modular function and almost positive function

Function	Sinusoidal Model	Quadratic Model	Linear Model
Function (51)	995.6	996.2	992.8
Function (52)	996.2	997.6	995.2
Function (53)	29	39.4	22.2

Table 2 Comparison of numerical simulation results of mean value of Sinusoidal Model, Quadratic Model and Linear Model for linear function, pseudo-modular function and almost positive function

Function	Sinusoidal Model	Quadratic Model	Linear Model
Function (51)	999	999	995
Function (52)	999	999	996
Function (53)	34	48	28

In addition, we will conduct several numerical simulations not only the functions analysed above but also several other classical benchmarks. In paper [1], Simon used 14 functions [17,18,19] as benchmarks to compare BBO with other EAs. In this paper, we use the them as well, which can help us know more about the mutation. The details about the benchmarks are given in Table 3. In addition, the property of these benchmarks is shown in Table 4.

Table 3 Benchmark Functions

Function	Name	Dimension	Domain
F1	Ackley's Function	20	$[-30, 30]^D$
F2	Fletcher-Powell	20	$[-\pi, \pi]^D$
F3	Generalized Griewank's function	20	$[-600, 600]^D$
F4	Generalized Penalized function 1	20	$[-50, 50]^D$
F5	Generalized Penalized function 2	20	$[-50, 50]^D$
F6	Quartic function	20	$[-1.28, 1.28]^D$
F7	Generalised Rastrigin's function	20	$[-5.12, 5.12]^D$
F8	Generalized Rosenbrock's function	20	$[-2.048, 2.048]^D$
F9	Schwefel' Problem 1.2	20	$[-65.535, 65.535]^D$
F10	Schwefel' Problem 2.21	20	$[-100, 100]^D$
F11	Schwefel' Problem 2.22	20	$[-10, 10]^D$
F12	Schwefel' Problem 2.26	20	$[-512, 512]^D$
F13	Sphere Model	20	$[-5.12, 5.12]^D$
F14	Step Function	20	$[-200, 200]^D$

From Table 5, Quadratic Model performs the best for 9 times, Sinusoidal Model 5 times and Linear Model only 1 time. This means that Quadratic Models has a higher expected value to obtain optimal solutions, which verify the analysis in this paper. In

Table 6, for benchmarks F1, F6, F7, F11, F13 and F14, the three models are competitive. Except the benchmarks, Quadratic Model wins for F4, F5, F10, and Sinusoidal Model wins for F9, F12, and Linear Model wins for F2, F8. Since Table 6 compares the best value, the results show that the models are competitive to get the best values.

Table 4 Property of benchmarks

Function	Multimodal	Separable	Regular
Ackley's Function	√		√
Fletcher-Powell	√		
Generalized Griewank's function	√		√
Generalized Penalized function 1	√		√
Generalized Penalized function 2	√		√
Quartic function		√	√
Generalised Rastrigin's function	√	√	√
Generalized Rosenbrock's function			√
Schwefel' Problem 1.2			√
Schwefel' Problem 2.21			
Schwefel' Problem 2.22	√		
Schwefel' Problem 2.26	√	√	
Sphere Model		√	√
Step Function		√	

Table 5 Comparison of numerical simulation results of Best value of Sinusoidal Model, Quadratic Model and Linear Model for 14 standard benchmarks

	Sinusoidal Model	Quadratic Model	Linear Model
F1	3.1822E-01	<b>3.0894E-01</b>	3.2927E-01
F2	2.7962E+03	3.2420E+03	4.4758E+03
F3	1.0019E+00	<b>1.0018E+00</b>	1.0019E+00
F4	5.8215E-03	<b>5.5064E-03</b>	6.8182E-03
F5	<b>1.6071E-02</b>	1.9529E-02	2.1795E-02
F6	1.6000E-08	<b>1.4400E-08</b>	1.6000E-08
F7	<b>9.9342E-02</b>	0.0000E+00	0.0000E+00
F8	1.0130E+01	<b>6.8275E+00</b>	7.6850E+00
F9	8.0016E-01	<b>7.6113E-01</b>	8.4164E-01
F10	<b>2.1739E+02</b>	2.5714E+02	2.3228E+02
F11	1.0000E-02	<b>6.0000E-03</b>	7.0000E-03
F12	<b>1.2480E+00</b>	1.3420E+00	1.2940E+00
F13	<b>0.0000E+00</b>	<b>0.0000E+00</b>	<b>0.0000E+00</b>
F14	1.0000E-01	<b>6.0000E-02</b>	8.0000E-02

Table 6 Comparison of numerical simulation results of mean value of Sinusoidal Model, Quadratic Model and Linear Model for 14 standard benchmarks

	Sinusoidal Model	Quadratic Model	Linear Model
F1	<b>0.0000E+00</b>	<b>0.0000E+00</b>	<b>0.0000E+00</b>
F2	2.7228E+02	2.2552E+02	<b>1.4647E+02</b>
F3	<b>1.0002E+00</b>	1.0003E+00	1.0004E+00
F4	6.5450E-05	<b>1.5705E-32</b>	1.5705E-03
F5	2.0000E-03	<b>1.3498E-32</b>	1.0000E-03
F6	<b>0.0000E+00</b>	<b>0.0000E+00</b>	<b>0.0000E+00</b>
F7	<b>0.0000E+00</b>	<b>0.0000E+00</b>	<b>0.0000E+00</b>
F8	1.2088E+00	1.2681E+00	<b>5.4877E-01</b>
F9	<b>1.6120E-01</b>	1.8224E-01	1.8701E-01
F10	5.9846E+01	<b>3.2906E+01</b>	4.5170E+01
F11	<b>0.0000E+00</b>	<b>0.0000E+00</b>	<b>0.0000E+00</b>
F12	<b>4.0000E-01</b>	7.0000E-01	8.0000E-01
F13	<b>0.0000E+00</b>	<b>0.0000E+00</b>	<b>0.0000E+00</b>
F14	<b>0.0000E+00</b>	<b>0.0000E+00</b>	<b>0.0000E+00</b>

## 6. Conclusions

In BBO, the mutation rate is not a constant but a variable affected by migration models, which is different from other standard evolutionary optimizations. The relationship between migration and mutation has been investigated in this paper. Three objective functions including linear function, pseudo-modular function and almost positive function are considered to investigate the mutation operator of BBO by drift analysis, which shows that the migration model which can generate a high mutation rate can enhance the performance of BBO. This conclusion is helpful to design the migration models. In the future, the indicator that how to improve migration model to obtain a better mutation operator should be investigated and will be a useful tool to measure the convergency of BBO.

## Acknowledge:

This work is sponsored by the National Natural Science Foundation of China under Grant No.70871091, No.61075064, No.61034004 and No.61005090, Program



for New Century Excellent Talents in University of Ministry of Education of China, Ph.D. Programs Foundation of Ministry of Education of China (20100072110038). The author Guo Weian express the best regards to Professor Shuzhi Sam Ge for the help and guidance and appreciates China Scholarship Council (CSC) to support the expense in Singapore.

## **Reference:**

- 
- 1 D. Simon, M. Ergezer, D. Du, and R. Rarick, Markov models for biogeography-based optimization, *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*, 41 (1) (2011) 299-306
  - 2 S. Kumar, P. Bhalla, and A. Singh, Fuzzy rule base generation from numerical data using biogeography-based optimization, *Institution of Engineers Journal of Electronics and Telecomm Engineering*, 90 (2009) 8-13
  - 3 A. Bhattacharya and P. Chattopadhyay, Closure to Discussion of "Hybrid Differential Evolution With Biogeography-Based Optimization for Solution of Economic Load Dispatch," *IEEE Transactions on Power Systems*, 27 (1) (2012) 575
  - 4 V. Panchal, H. Kundra, and A. Kaur, An integrated approach to biogeography based optimization with case-based reasoning for exploring groundwater possibility, *The Delving: Journal of Technology and Engineering Sciences*, 1 (1) (2009) 32-38
  - 5 A. Bhattacharya and P. Chattopadhyay, Biogeography-based optimization for different economic load dispatch problems, *IEEE Transactions on Power Systems*, 25 (2) (2010) 1064-1077
  - 6 W. Guo, D. Yang, J. Hu, C. Huang, L. Wang, Q. Wu, Optimal management of a home smart grid scheduling, *Journal of Computational Information*

---

Systems, 8(5)(2012) 1921-1928

- 7 W Gong, Z Cai, C. X. Ling, H. Li, A real-coded biogeography-based optimization with mutation, *Applied Mathematics and Computation* 216 (2010) 2749–2758
- 8 Y. Wang, Z. Cai, A Novel Hybrid Biogeography-based Optimization with differential mutation, *Electronic and Mechanical Engineering and Information Technology (EMEIT)*, 2011 International Conference on, Volume: 5 2710 - 2714, Harbin, Heilongjiang, China
- 9 H. Dong, J. He, H. Huang, W. Hou, Evolutionary programming using a mixed mutation strategy, *Information Sciences*, 177 (1) (2007) 312–327
- 10 L. Hu, F. Sun, H. Xu, H. Liu, X. Zhang, Mutation Hopfield neural network and its applications, *Information Sciences*, 181 (1) (2011) 92–105
- 11 D. Simon, Biogeography-Based Optimization, *IEEE Transactions on Evolutionary Computation* 12 (6) (2008) 702-713
- 12 H. Ma, An analysis of the equilibrium of migration models for biogeography-based optimization, *Information Sciences* 176 (8) (2010) 3444-3464.
- 13 W. Guo, L. Wang, Q. Wu, An Analysis of the Migration Rates for Biogeography Based Optimization, *Information sciences*, *Accept.*
- 14 J. He, L. Kang, On the convergence rate of genetic algorithms, *Theoretical Computation Science*, 229 (1999) 23–39.
- 15 J. He, X. Yao, Drift analysis and average time of evolutionary algorithms, *Artificial Intelligence*, 127(2001) 57-85.
- 16 G. Rudolph, Finite Markov chain results in evolutionary computation: A tour d’horizon, *Fundamenta Informaticae* 35 (1–4) (1998) 67–89.
- 17 T. Back, *Evolutionary Algorithms in Theory and Practice*. Oxford, U.K.: Oxford Univ. Press, 1996.

- 
- 18 Z. Cai and Y. Wang, A multi-objective optimization-based evolutionary algorithm for constrained optimization, *IEEE on Evolutionary Computation* 10 (6) (2006) 658-675.
- 19 X. Yao, Y. Liu, and G. Lin, Evolutionary Programming made faster, *IEEE Transactions on Evolutionary Computation*. 3 (2) (1999) 82-102.