

基于免疫进化算法的逻辑斯谛曲线 水环境质量综合评价模型

苏鹏程¹, 倪长健²

(1. 中国科学院水利部成都山地灾害与环境研究所, 四川 成都 610041;

2. 成都信息工程学院地球环境科学系, 四川 成都 610041)

摘 要:在对现有水环境质量评价方法进行分析的基础上, 提出用一种新型的优化算法——免疫进化算法应用于非线性形的水环境质量综合评价模型, 即逻辑斯谛曲线的参数优化问题, 其结果优于目前广泛应用的遗传算法, 从而建立了一种新的数学模型。实例研究表明, 该算法易于操作、快速收敛于全局最优解, 可用于解决复杂的优化问题, 具有较为广阔的应用前景。

关键词:免疫进化算法; 水环境质量; 评价; 优化; 最小二乘法; 加速遗传算法

中图分类号: X824

文献标识码: A

水环境质量综合评价, 是根据已经获取的某些水环境质量指标值, 通过建立数学模型, 对水体的质量等级进行综合评价。它为水体的污染防治和开发利用提供科学依据, 是国民经济建设中进行决策的前提条件之一。

在进行水质评价时, 仅仅依靠水环境质量评价标准表进行水质评判往往有一定的局限性。目前用于水环境质量评价的方法比较多, 如神经网络法、模糊综合评价法、模糊灰色评价法、灰色聚类法、平分统计法和物元分析法等^[1], 但这些评价方法的计算结果大多是一些离散的水质等级, 是半定量化的, 等级的分辨率较粗, 对指导具体的水环境管理工作有所不便。相继人们提出了用逻辑斯谛曲线(Logistic Curve, 简称 LOG 模型) 作为水环境质量综合评价的数学模型, 并给出了基于加速遗传算法(Accelerating Genetic Algorithm, 简称 AGA 算法) 的建模方案^[2], 但通过对遗传算法原理和操作机制的分析, 改进后的加速遗传算法优化 LOG 模型主要依靠个人技巧^[3]。在实际应用中, 使用者须对 AGA 算法复杂的实现技术十分熟悉, 如果要加以推广的话, 比较困难。

本文尝试将免疫进化算法(Immune Evolutionary Algorithm, 简称 IEA 算法) 应用于水环境质量综合评价模型的参数优化, 以期建立一种能够应用于实际问题的统一数学模型, 并将计算结果与传统的遗传算法进行比较, 最后进行了实例研究。

1 免疫进化算法原理

受生物免疫系统自然防卫机制的启迪, IEA 把算法理解为免疫系统, 而把外来侵犯的抗原和免疫系统产生的抗体分别与实际求解问题的目标函数以及问题的解相对应。借鉴生物免疫机制, 免疫进化算法中子代个体的生成方式为

$$\begin{cases} x^{t+1} = x_{\text{best}}^t + \sigma^t N(0, 1) \\ \sigma^t = \sigma^0 \exp\left(-\frac{At}{T}\right) \end{cases} \quad (1)$$

式中 x^{t+1} 为子代个体的可行解; x_{best}^t 为父代最优个体; σ^t 为父代群体的标准差; A 为标准差动态调整系数; T 为总的进化代数; $N(0, 1)$ 为产生的服从标准正态分布的随机数; t 为进化的代数; σ^0 为对应

收稿日期(Received date): 2004- 01- 11; 改回日期(Accepted): 2004- 04- 20。

基金项目(Foundation item): 成都信息工程学院院管基金(CRF200310)资助。[Supported by the management foundation of Chengdu University of information technology.]

作者简介(Biography): 苏鹏程(1981-), 男, 安徽桐城人, 研究实习员, 学士, 主要从事灾害与环境的研究。E-mail: spcan@etang.com。[SU Pengcheng(1981-), male, A research assistant, work on research of hazards and environment.]

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

于初始群体的标准差; A 和 σ^0 的取值根据被研究的问题来确定, 通常 $A \in [1, 10]$, $\sigma^0 \in [1, 3]^{[4]}$ 。

2 免疫进化算法优化水环境质量综合评价的 LOG 模型

2.1 LOG 模型介绍

逻辑斯谛曲线是 1938 年由荷兰生物学家 P. F. Verhulst 提出的, 该模型的曲线分为三个部分, 先缓慢增加, 然后快速增加, 后又逐渐趋于稳定, 可反映事物的发生、发展和成熟的一般规律, 它又称为生长曲线。该曲线似“S”, 故又称 S 型曲线。由于 S 型曲线富有弹性, 已在地学、环境学、生态学、人口学和经济预测等领域得到了广泛应用。逻辑斯谛曲线的数学方程是

$$x(t)=\frac{K}{(1+e^{a-tr})}$$

(2)

式中 $x(t)$ 为时刻 t 的待预测变量; K 为预测变量的饱和值; r 为增长率; a 为积分常数^[5]。

2.2 IEA 用于对 LOG 模型参数的优化

为了与在文献[2] 中已获得比较深入研究的 AGA 算法进行对照, 这里同样以湖泊水质富营养化综合评价为例, 将两种方法应用于 LOG 模型的参数优化。在此摘录了湖泊富营养化的 4 项水质指标: 总磷、耗氧量、透明度、总氮以及他们的 5 级水质等级的具体标准^[1](表 1)。

表 1 湖泊水质评价标准

Table 1 The assessment standard of lake water quality

评价指标	极贫营养	贫营养	中营养	富营养	极富营养
	1	2	3	4	5
总 磷($\mu\text{g/l}$)	< 1	4	23	110	660
耗氧量(mg/l)	< 0. 09	0. 36	1. 80	7. 10	27. 10
透明度(m)	> 37	12	2. 40	0. 55	0. 170
总 氮(mg/l)	< 0. 02	0. 06	0. 31	1. 20	4. 600

注: 为便于比较, 总磷未取法定单位 mg/L ($1\mu\text{g/L}=10^{-3}\text{mg/L}$)。后同。

在表 1 的各级水质取值范围内随机产生 5 个指标值, 与对应的经验水质等级一起组成样本系列, 共 25 组。为了与 AGA 算法进行比较, 这里的各水质指标采用文献[2] 中同样的初始数据, 如表 2 所示。

表 2 湖泊水质等级经验值和用两种算法优化的 LOG 模型计算值的对比

Table 2 Comparison between empirical value and computations of two optimization LOG models

序 号	水质指标				水质等级		
	总磷($\mu\text{g/l}$)	耗氧量(mg/l)	透明度(m)	总氮(mg/l)	经验值	AGA	IEA
1	0. 658	0. 059	50. 857	0. 001	1. 0	0. 97	1. 06
2	0. 559	0. 051	42. 650	0. 015	1. 0	1. 22	1. 28
3	0. 106	0. 034	67. 902	0. 009	1. 0	0. 60	0. 70
4	0. 978	0. 061	42. 207	0. 010	1. 0	1. 23	1. 30
5	0. 646	0. 051	47. 659	0. 015	1. 0	1. 07	1. 15
6	1. 233	0. 259	36. 899	0. 040	2. 0	1. 44	1. 50
7	3. 225	0. 349	24. 079	0. 047	2. 0	1. 94	1. 94
8	2. 821	0. 133	23. 709	0. 042	2. 0	1. 94	1. 93
9	1. 426	0. 176	16. 260	0. 020	2. 0	2. 21	2. 18
10	1. 947	0. 098	19. 699	0. 035	2. 0	2. 09	2. 06
11	15. 338	0. 918	3. 359	0. 185	3. 0	3. 05	2. 93
12	22. 127	1. 799	4. 942	0. 240	3. 0	3. 13	3. 03
13	20. 269	1. 681	2. 606	0. 281	3. 0	3. 26	3. 13
14	9. 171	1. 479	3. 359	0. 185	3. 0	3. 01	2. 93
15	16. 739	1. 308	8. 979	0. 290	3. 0	2. 99	2. 87
16	24. 736	2. 621	1. 530	1. 182	4. 0	4. 22	3. 95
17	91. 674	3. 710	2. 106	0. 627	4. 0	3. 95	3. 82
18	105. 427	6. 403	1. 176	0. 661	4. 0	4. 13	4. 07
19	106. 776	5. 898	0. 993	1. 145	4. 0	4. 43	4. 30
20	93. 002	5. 341	1. 038	0. 606	4. 0	4. 02	3. 95
21	627. 669	20. 039	0. 409	2. 627	5. 0	4. 99	4. 98
22	297. 816	13. 171	0. 191	3. 050	5. 0	4. 97	4. 94
23	370. 796	21. 733	0. 460	3. 068	5. 0	4. 98	4. 98
24	279. 420	17. 600	0. 347	3. 850	5. 0	4. 99	4. 98
25	281. 502	22. 646	0. 317	1. 464	5. 0	4. 88	4. 90

表中 $x(1, i)$, $x(2, i)$, $x(3, i)$ 和 $x(4, i)$ 分别表示水体总磷、耗氧量、透明度、总氮的质量指标, 其中 $i = 1, 2, \dots, 25$ 为样本数据的个数; $Z(i)$ 表示环境质量的计算等级; $c(j)$ 表示 LOG 模型的参数, 其中 $j = 0, 1, \dots, 4$ 。在这里设最低水环境质量等级为 1 级, 最高等级设为 5 级, 当指标值超过 5 级限值时就判定为最高等级, 当指标低于 1 级限值时就判定为最低等级, 当指标值介于这两个限值之间时则判定为中等等级。

2.2.1 LOG 模型合理性验证

首先这里对各个指标 $x(1, i)$, $x(2, i)$, $x(3, i)$ 和 $x(4, i)$ 的 25 个数据依照从小到大的顺序进行排

列, 然后用 MATLAB 软件将它们分别与相应的水质等级 $y(i)$ 进行拟合, 得出图 1。从图中可以看出, 这是一种上下段有限, 中间段变化迅速的复杂函数关系, 各指标值(横坐标)与水质等级(纵坐标)之间呈现单调递增关系, 与水体由澄清、被污染到水质恶化过程中水质等级的变化一致。他们都符合逻辑斯谛曲线描述的特征, 表明采用 LOG 模型是合理的。对应于 4 项水质指标, 根据式(2)本例中 LOG 模型的具体形式表述为

$$Z(i) = \frac{N}{[1 + e^{c(0) + C(1)x(1,i) + c(2)x(2,i) + c(3)x(3,i) + c(4)x(4,i)}]}$$

(3)

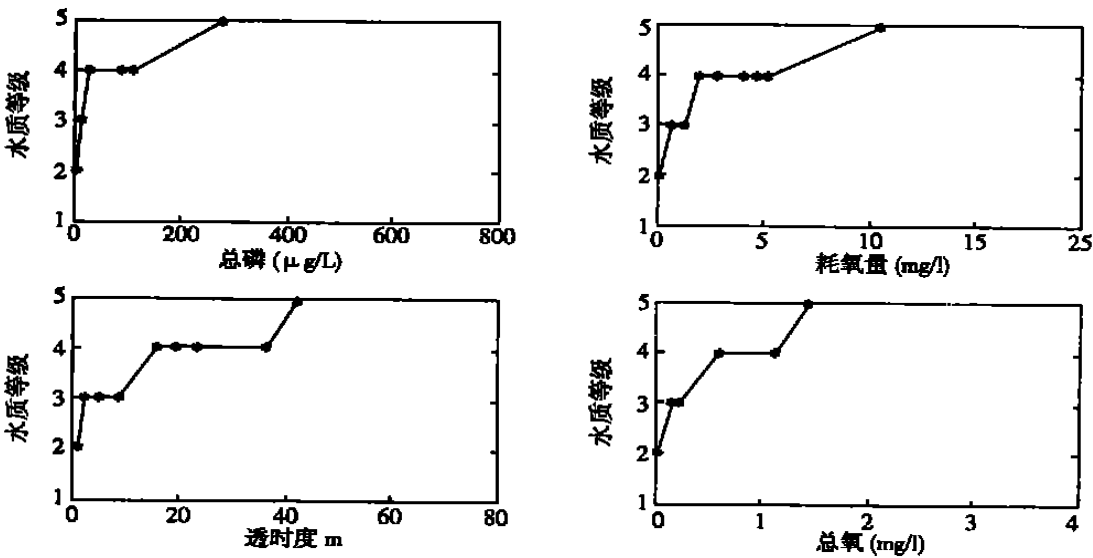


图 1 各水质指标初始数据系列与水质等级的关系

Fig. 1 Relation between original date of the water quality and water quality class

2.2.2 确定目标函数

为了对其参数 $c(j)$ 进行合理估计, 这里采用最小二乘法并构造免疫进化算法所要求的如下目标函数

$$\min(f(x)) = (Z(1) - y(1))^2 + (Z(2) - y(2))^2 + \dots + (Z(25) - y(25))^2$$

(4)

2.2.3 LOG 模型参数的 IEA 优化过程

第一步, 初始参数赋值。标准差动态调整系数 $A = 5$, 初始标准差 $\sigma^0 = 1.0$ 。编程时给定的总的进化代数数为 $T = 300$, 初试群体的规模 $S = 200$ 。

第二步, 初始数据处理。将 LOG 模型的 5 个参数 $c(0)$, $c(1)$, $c(2)$, $c(3)$ 和 $c(4)$ 的取值区间统一定为 $[-10.000, 10.000]$ 。相对于文献[2], 从表 3 中可以看出它们的区间范围增大了, 这也意味着搜

索难度的加大。为便于计算, 把表 2 中各数据 $x(1, i)$, $x(2, i)$, $x(3, i)$ 和 $x(4, i)$ 分别除以 660, 27.10, 74 和 4.60 代入式(3)得到 $Z(i)$, 并与水质等级经验值 $y(i)$ 一起代入式(4), 即得此例的目标函数 $\min(f(x))$ 。

第三步, 对目标函数 $\min(f(x))$ 进行优化^[4]。(1) 解空间内随机生成初始群体, 并计算其适应度, 确定最优个体 x_{best}^0 ; (2) 根据式(1)进行进化操作, 在解空间内生成子代群体, 规模为 S ; (3) 计算子代群体的适应度, 确定最优个体 x_{best}^{t+1} , 若 $f(x_{best}^{t+1}) < f(x_{best}^t)$, 则选定最优个体为 x_{best}^{t+1} ; 否则, 最优个体取为 x_{best}^t ; (4) 反复执行步骤(2)和(3), 直至达到终止条件, 选择最后一代的最优个体作为寻优结果, 结果见表 3。

表 3 用两种算法优化的 LOG 模型参数及计算结果的对比

Table 3 Comparison between results from two optimizing LOG model and their parameters

		$c(0)$	$c(1)$	$c(2)$	$c(3)$	$c(4)$	$min(f(x))$
AGA	搜索区间	$[-5.0, 5.0]$	$[-10.0, 0]$	$[-10.0, 0]$	$[0, 10.0]$	$[-10.0, 0]$	
算法	优化结果	-0.276 5	-2.024 0	-1.198 5	2.476 6	-4.928 8	0.988 0
IEA	搜索区间	$[-10, 10]$	$[-10, 10]$	$[-10, 10]$	$[-10, 10]$	$[-10, 10]$	
算法	优化结果	-0.192 5	-2.139 0	-2.008 0	2.206 7	-3.498 4	0.779 9

最后把表 3 所得到的模型参数优化值代回式 (3), 即得此例水环境质量综合评价的 LOG 模型

$$Z(i)=\frac{5}{[1+e^{-0.1925-2.139x(1,i)-2.008x(2,i)+2.2067x(3,i)-3.4984x(4,i)}]} \tag{5}$$

把 25 个初试样本数据代入该模型, 对应的水质等级计算值参见表 2。

此外在程序设计过程中, 本例有意记录了 IEA 算法在进化过程中群体的平均适应度和最优个体的适应度随进化代数变化的情况(图 2)。

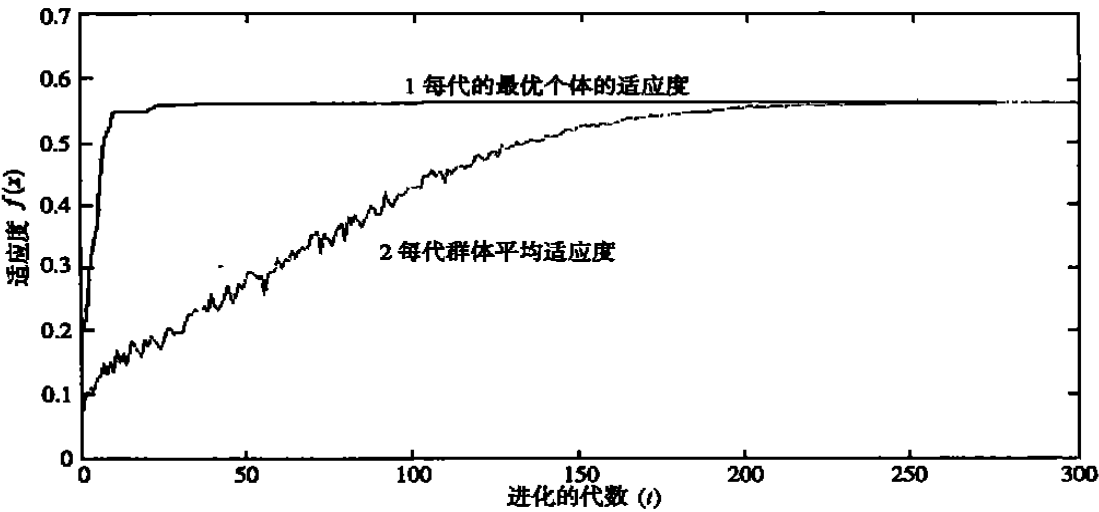


图 2 IEA 算法优化 LOG 模型参数的过程中最优个体和群体平均适应度随进化代数 t 的变化曲线

Fig.2 The curve of the optimal unit and the whole average fitness change with evolution era t during the process of IEA for LOG

2.2.3 对 IEA 算法优化结果的比较和分析

常规的逻辑斯谛曲线估计方法有目测法、三点法、平均值法、四点式平均值法^[6]、麦夸尔特法、枚举法优选法^[7]、逐次加密搜索法^[8]、正交设计法^[9]等。金菊良等人的研究表明了 AGA 算法的效果明显优于上述方法^[10]。从表 3 可以看出 IEA 在 AGA 的基础上又将目标函数的误差减少了 21%, 可见这里 IEA 算法的优化效果比 AGA 算法要好。

同时, 对比表 2 中水质等级的经验值和两种优化方法的计算值以及表 1 的水质等级标准, 可以看出 LOG 模型的计算值是合理的, 它进一步刻画了各水质指标值的数量差异对水质等级判定的影响。从图 2 中可以看出, IEA 算法在对参数系列的优化过程中, 随着进化代数的增加, 它能很快逼近全局最优

解。同时它以其自身的特点, 在进化的过程中, 群体的平均适应度虽然呈现一定幅度的震荡, 但在总体上又呈稳定的增加, 这显示它能很好的平衡全局和局部最优的搜索。

2.3 应用实例

根据式 (5) 的 LOG 模型和被文献[2] 采用的来自文献[11] 的 5 个湖泊水质指标实测数据, 对各湖泊水质富营养化等级进行评判, 其结果见表 4。表 4 中同时也列出了文献[1] 模糊神经网络方法和文献[2] AGA 算法的评价结果。

对照表 1 的评价标准, 可以看出优化后的 LOG 模型计算的结果是合理的、精确的。例如西湖和东湖的水质指标十分接近, 东湖的总磷指标值已接近 4 级(富营养)水质标准的上限值, 其余指标值都在 5

级(极富营养)水质范围内。IEA 算法优化后的 LOG 模型评价的结果分别为 4.85 级和 4.72 级, 而用神经网络评价方法则认为这两湖泊分别属于 5 级和 4 级, 模糊综合评价方法和模糊灰色评价方法则认为这两湖泊均为 4 级。这里体现出了采用 LOG 模型的优势: 它是一简明的代数函数式, 根据模型参

数值的大小, 可进一步分析各指标值的等级标准是否合理, 如不合适则可适当调整, 如此反复, 可增强评价标准的客观性。另一方面, 从表 3 中可知虽然加大了搜索空间, 但 IEA 优化后的模型比 AGA 算法优化的结果更接近实际值, 且其优化的过程比 AGA 更为简洁。

表 4 用不同模型评价的五大湖泊水质评价结果
Table 4 Assessments of water quality of five lakes using different models

湖 泊	水质指标			水质等级模型计算值			
	总 磷 (mg/L)	耗氧量 (mg/L)	透明度 (m)	总 氮 (mg/L)	神经网络方法	AGA 算法	IEA 算法
杭州西湖	130	10.30	0.35	2.76	5	4.92	4.85
武汉东湖	105	10.70	0.40	2.00	4	4.80	4.72
青海湖	20	1.40	4.50	0.22	3	3.09	2.99
巢湖	30	6.26	0.25	1.67	4	4.59	4.41
滇池	20	10.13	0.50	0.23	4	3.67	3.81

3 结论

通过运用免疫进化算法(IEA)优化 LOG 模型的参数, 可以得出如下结论:

- 1. IEA 算法是一种容易操作的算法, 相对于 AGA 算法, 它的参数只有动态调整系数(A)和初始标准差(σ^0), 且优化效果更好。
- 2. IEA 算法不同于 AGA 算法的随机的交叉, 它通过在父代最优个体上叠加一个服从正态分布的随机数, 加大了对父代最优个体附近空间的投影密度, 同时也以较小概率寻找其他优秀的个体, 确保结果逼近全局最优解。因此它体现出的是一种遗传与变异的综合特征。
- 3. 在进化后期, 随着进化代数的增加, 它会以更快的速度收敛, 这比 AGA 在中后期的随机搜索更为有效。

实例研究表明, IEA 算法能有效地解决复杂的非线性优化问题, 通过这一方法可以建立统一的、指导性数学模型, 可以广泛地应用于水环境质量评价。同时 IEA 作为一种新型的优化方法, 也可以广泛应用于其他领域。

致谢: 刘希林研究员 对本文的修改提出了宝贵意见, 特此深表感谢!

参考文献(References):

[1] Hu Mingxing, Guo Dazhi. Evaluation of lake eutrophication by means of fuzzy neural network [J]. *Research of Environmental Sciences*, 1998, **11**(4): 40~ 42. [胡明星, 郭达志. 湖泊水质富营养化评价的模型神经网络方法[J], 环境科学研究, 1998, **11**(4): 40~ 42.]

[2] Jin Juliang, Ding Jing. Genetic algorithm and its application to water science[M]. Chengdu: Sichuan University Press, 2000. 118~ 122. [金菊良, 丁晶. 遗传算法及其在水科学中的应用[M]. 成都: 四川大学出版社, 2000. 118~ 122.]

[3] Yan Pingfan, Zhang Changshui. Manual work nerve network and simulating evolution arithmetic[M]. Beijing: Tsinghua University Press, 2000. [阎平凡, 张长水. 人工神经网络与模拟进化算法[M]. 北京: 清华大学出版社, 2000.]

[4] Ni Changjian, Ding Jing, Li Zuoyong. Immune evolutionary algorithm[J]. *Journal of Southwest Jiaotong University*, 2003, **38**(1): 87~ 91. [倪长健, 丁晶, 李祚泳. 免疫进化算法[J]. 西南交通大学学报, 2003, **38**(1): 87~ 91.]

[5] Jin Juliang, Ding Jing. Water resource system engineering[M]. Chengdu: Sichuan Science & Technology Press, 2002, 178~ 180. [金菊良, 丁晶. 水资源系统工程[M]. 成都: 四川科学技术出版社, 2002, 178~ 180.]

[6] Wang Zhenzhong, Ling Kongxun. Four-point average value estimation of logisitic curve K value[J]. *Acta Ecologica Sinica*, 1987, **7**(3): 193~ 197. [王振中, 林孔勋. 逻辑斯谛曲线 K 值的四点式平均值估计法[J]. 生态学报, 1987, **7**(3): 193~ 197.]

[7] Wan Changxiu, Liang Zhongyu. A imitating method of logistic curve[J]. *Acta Ecologica Sinica*, 1983, **3**(3): 288~ 296. [万昌秀, 梁中宇. 逻辑斯谛曲线的一种拟合方法[J]. 生态学报, 1983, **3**(3): 288~ 296.]

- [8] Zhou Saihua. Estimating the parameters of logistic equation[J]. *application of Statistics And Management*, 1992, **11**(5): 32~35. [周赛花. 逻辑斯谛方程中参数的估计[J]. 数理统计与管理, 1992, **11**(5): 32~ 35.]
- [9] SiMa Xisheng. Estimator of Pearl Curv Parameters[J]. *Forecasting*, 1994, (1): 51~ 52. [司马锡生. 皮尔曲线的参数估计[J]. 预测, 1994, (1): 51~ 52.]
- [10] Jin Juliang, Yang Xiaohua, Liu Yuming, *et al.* Application of genetic algorithm for estimating logistic curve parameters[J]. *System Sciences and Comprehensive Studies in Agriculture*, 1997, **13**(3): 186~ 190. [金菊良, 杨晓华, 刘宇敏, 等. 基因算法在 logistic 曲线估计中的应用[J]. 农业系统科学与综合研究, 1997, **13**(3): 186~ 190.]
- [11] Shu Jinhua. Eutrophication degree elementary assessment and prevention countermeasure of our country primary lakes[J]. *Techniques and Equipment for Environmental Pollution Control*, 1986, **7**(2): 1~ 9. [舒金华. 我国主要湖泊的富营养化程度的初步评价与防治对策[J]. 环境污染治理技术与设备, 1986, **7**(2): 1~ 9.]

Immune Evolutionary Algorithm and Its Application to Quality Assessment of Water Environment

SU Pengcheng¹, NI Changjian²

(1. *Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610041, China;*

2. *Department of Earth Environment Science, Chengdu University of Information Technology, Chengdu 610041, China*)

Abstract: Based on analysing existing water environment quality assessment model and optimization technology, a new optimization technology called Immune Evolutionary Algorithm(IEA) is applied to optimize the parameters for nonlinear water environment quality integrated assessment model. According to element of least square method, a new optimization model is set up. IEA is simply operated because it only has two dynamic adjustable parameters. An example is shown that IEA has a much better result than Accelerating Genetic Algorithm (AGA) which is ameliorated by Genetic Algorithm and applied extensively. Comparing IEA with existing optimization method, there are two conclusions, one is that the solution is not jumpiness during the process of evolution, it incarnates a integrated characteristic of descendibility and aberrance, an other is that it is more effective at the latter period. Therefore, this technology is good and new algorithm, which also can be used in many other fields.

Keywords: Immune Evolutionary Algorithm (IEA); water environment; assessment; optimization; least square method; Accelerating Genetic Algorithm (AGA)