	水利	学报	
2005年6月	SHUILI	XUEBAO	第36卷 第6期

文章编号:0559-9350(2005)06-0674-06

遗传单纯形混合算法在复杂环境模型参数识别中的应用

王建平,程声通

(清华大学 环境科学与工程系,北京 100084)

摘要:参数识别是数学模型应用的一个重要环节。为提高复杂环境模型参数识别的性能和效率,引入了遗传单纯 形法(GASM),该方法融合了遗传算法和单纯形法两类算法的不同搜索机制,具有很强的广度搜索和深度搜索能 力。本研究以密云水库水质模拟为例,将 GASM 算法应用于模拟地表水水质的 WASP 模型中 10 个参数的优化识 别。计算结果表明,无论是没有扰动的情况还是有扰动的情况,GASM 算法均高效可靠地搜索到水质模型参数的全 局最优解,说明此方法应用于复杂环境模型参数搜索是可行的实用的。同时,通过不同算法的比较也说明了 GASM 算法在搜索性能和效率方面的优越性。

关键词:参数识别;环境模型;遗传算法;单纯形法;全局优化;混合算法

中图分类号:X524 文献标识码:A

环境模型参数识别属非线性约束优化问题,约束是模型系统,故参数识别优化又是一个仿真优化问题,即每一次参数调整都需进行仿真模拟。仿真优化是基于仿真的目标优化问题,一个显著特点是系统 仿真过程较费时,以致于整个优化过程很慢^[1]。传统的优化方法,如梯度法和单纯形法对模型结构、优 化准则要求严格,受初始条件影响较大,通常只能得到局部最优解。基于随机采样的统计方法,如 HSY 算法、GLUE 算法等由于其随机抽样机制,参数个数超过 3~5 个时,率定过程将非常耗时^[2]。因此,在环 境模型参数识别研究中,开发时间复杂性可以接受的优化算法至关重要。

1 遗传单纯形混合算法

遗传算法(genetic algorithms, GA)是 J. Holland 于 1975 年受生物进化论的启发而提出的。GA 是一种高度并行、随机和自适应的通用优化算法,其编码技术和遗传操作比较简单,优化不受限制性条件的约束,两个最显著特点是隐含并行性和全局解空间搜索。遗传算法在环境模型参数识别中以水文模型^[3]、地下水模型^[4]等居多,而且研究开始较早。遗传算法用于水质模型参数率定的国内外研究文献还不多,多集中在应用研究,并以河流模型系统为主,优化的参数较少^[5]。

尽管遗传算法在理论上具有概率1的收敛特性,但实际应用时往往出现"早熟"收敛或收敛缓慢等缺点^[6]。不少学者进行了 GA 的改进研究^[7,8]。算法设计方面,Grefenstette^[9]将 GA 的参数选取作为一个优化问题,提出用 GA 优化 GA 参数的二级数值方法。尽管此方法适用范围较广,但工作量较大,且二级算法本身的参数也有待优化,因此很少得到实际应用。Goldberg^[7]引入分享(Sharing)思想将解空间分成若干子空间,然后在子空间中产生子群体成员分别进行优化,以求得整个问题的解,避免算法只收敛到某个局部解。然而,这些方法很难从根本上避免"早熟"收敛和收敛缓慢的现象。算法混合的思想已发展成为提高算法优化性能的一个重要且有效的途径,其出发点就是使各种单一算法相互取长补短,产生更好的优化效率。

收稿日期:2004-07-19

资助项目:北京市重大科技项目(H012110010119)

作者简介:王建平(1977-),男,内蒙古清水河人,博士生,主要从事环境系统分析的理论与应用研究。E-mail:wangjp@tsinghua.org.cn.

674

遗传算法为指导性搜索算法,全局搜索能力较强,但其局部搜索能力较弱,导致优化解质量不高。 单纯形法(Simplex Method,SM)^[10],也称可变多面体搜索法,是确定性下降方法,局部搜索能力很强。将 搜索机制上存在如此差异的两种算法进行混合,有利于丰富搜索行为,增强全局和局部意义下的搜索能 力和效率。基于这样的思路,本文研究了遗传单纯形混合算法在环境模型参数识别中的适用性。遗传 算法和单纯形法的混合算法在 Duan 等 1993 年提出的 SCE-UA 算法中已有应用^[11]。2003 年,Vrugt 又对 SCE-UA 进行了改进,增加了 Metropolis 准则控制随机搜索、种群进化和复合形演化的操作(SCEM-UA), 并在水文模型中也得到成功应用^[12]。

参数识别可转化为如下的最优化问题

$$f = \min \left\{ \sum Y(t) - Y_0(t) \right)^{\mathrm{T}} (Y(t) - Y_0(t)) \right\}$$
(1)

式中: $Y_0(t)$ 为模型系统的真实输出;Y(t)为搜索参数下的模型系统输出。

遗传单纯形混合算法是遗传算法和单纯形法结合的 混合算法。遗传算法为启发式全局搜索法,单纯形法是确 定性下降方法,选择机制上存在如此差异的两种算法进行 混合,有利于丰富搜索行为,增强全局和局部意义下的搜 索能力和效率。具体算法流程如图1所示,其混合思想是 以遗传算法的基本框架为基础,由单纯形法替代标准遗传 算法的交叉操作和变异操作,本研究中应用的单纯形算法 是单纯形算法在约束问题上的推广,也称复合形算法^[11]。 遗传单纯形混合算法不仅具有遗传算法通用、简单、全局 随机搜索的优点,而且融入了局部搜索法的快速寻优并收 敛的优点, Duan 指出由启发式搜索算法和局部优化法组 成的混合算法是求解非线性复杂模型采用随机搜索方法 最成功的途径之一¹¹¹。有关遗传单纯形法详见文献 [11],其中标准遗传算法的参数选取和设计详见相关书 籍^[6],此外种群 S 的划分方法是按目标值排序由高到低 依次划分为p个子群。



算法控制参数包括:算法总进化代数 gen,种群大小 N,子种群个数 p,单纯形顶点数 C,单纯形内循 环次数 T,即每一代调用参数形算法次数。

2 WASP 模型和数据

2.1 模型及待识别参数简介 研究采用的水质模型为 WASP 模型系统。WASP^[13]是由美国国家环保局开发的用于地表水水 质模拟的模型,它提供了一个灵活的动态模拟系统。如图 2 所示,WASP 可以模拟 8 个指标,分别为:氨氮(NH₃)、硝酸盐氮(NO₃)、溶解性磷酸盐(OPO₄)、叶绿素 a(Chl-a)、碳生化需氧量(CBOD)、溶解氧(DO)、有机氮(ON)和有机磷(OP)。WASP 模型系统中水质模块 EUTRO5 的参数有 42 个之多,通过灵敏度分析提取灵敏度较高的参数如表 1 所示。
2.2 数据序列产生 本文采用合成的"观测"数据进行研究,即在已知参数值的情况下(参数值见表 1)应用 WASP 模型,产生"真实值"时间序列,得到"观测值"时间序列,然后利用这些数据进行模型参数识别。这样做的目的是为了使系统的真实



图 2 水质模拟反应动力学关系

情况在掌握之中,同时模型没有结构上的误差,参数估计的所有误差仅来源于参数初值和参数识别本 身。这使得有可能在排除结构误差的情况下,单独研究遗传单纯形混合算法的参数识别效果和收敛效 率。合成数据方法在许多模型分析中有广泛的使用。为简洁起见,在不引起混淆的情况下,不再对"观 测值"和"真实值"加上引号。

用于参数识别的合成数据序列为:(1)没有扰动的数据序列。模型最优解已知,目标函数值为零,后 文中简称"没有扰动情况";(2)两组带有扰动的数据序列。为了模拟观测误差的影响,给观测值(没有扰 动情况的数据序列)加上了一定的随机挠动。随机扰动为: $\zeta = N(0,\sigma)$,即均值为0,标准差为 σ 的正态 分布, σ 分别为各水质状态变量浓度均值的5%和10%。观测值为: $C'(t) = C(t) + \zeta$,显然,观测误差 满足独立、正态和同方差的性质。叙述中分别简称为"5%扰动情况"和"10%扰动情况"。

参数名称	物理意义	参数取值	参数范围
K12C	20℃条件下的硝化速度系数/d ⁻¹	0.200	0.05 ~ 0.35
K20 C	20℃条件下的反硝化速度系数/d ⁻¹	0.050	0.01 ~ 0.20
K1 C	浮游植物的饱和生长率/d ⁻¹	2.500	$1.50 \sim 4.00$
K1 RC	20℃条件下浮游植物的内源呼吸速率/d ⁻¹	0.100	0.05 ~ 0.20
NCRB	浮游植物内的氮碳比,mg/mg,缺省值为0.25	0.250	0.20~0.30
PCRB	浮游植物内的磷碳比,mg/mg,缺省值为0.025	0.025	0.02 ~ 0.03
KDC	20℃条件下的 CBOD 降解速率/d ⁻¹	0.030	0.01 ~ 0.10
K2	20℃条件下,水体的复氧速度常数/d ⁻¹	0.150	0.10~0.20
K71 C	溶解有机氮的矿化速度/d ⁻¹	0.030	0.01 ~ 0.10
K83 C	溶解有机磷的矿化速度/d ⁻¹	0.030	0.01 ~ 0.10

表1 待识别水质模型参数

注:参数范围确定参考文献包括 Ditoro and Mtystik^[14], Thomann and Fitzpatrick^[15], Bowie et al^[16].

3 参数识别结果及其分析

研究中为避免随机数产生器种子(seed)的影响,每一组试验均随机模拟 5 次,参数搜索范围见表 1 最后一列。优化目标为式(1),即模拟值和观测值的最佳拟合,优化方法为遗传单纯形算法。算法控制 参数: gen = 200, N = 50, p = 2, C = 11, T = 25。模型运行结果及分析如下。

3.1 没有扰动的情况 在没有扰动的情况下,最优参数值已知,最优目标函数值为0。从表2可知,遗 传单纯形算法能有效地识别水质模型参数,目标值达到最优目标值,水质参数的最优解全部准确识别。 研究同时考察了 GA 和 SM 分别单独运行时的参数优化效果,如表2所示,GA 和 SM 均收敛到局部较优 解。理论上 GA 可实现全局最优,但在有限计算量要求下,往往很难获得最优解。综上可知,将 GA 和 SM 两种算法混合,可大大提高算法的优化性能,这一点从后面的结果中也可得到证实。

去 # な #	44 金米目山田	遗传单纯形法		GA		SM	
参数名称 参数最优值	优化参数	相对误差(%)	优化参数	相对误差(%)	优化参数	相对误差(%)	
K12C	0.200	0.20000	0.0	0.21055	5.3	0.16908	15.5

表 2 没有扰动情况下的参数识别结果

目标值		0		2.08921		0.39597	
K83 C	0.030	0.03000	0.0	0.02991	0.3	0.02977	0.8
K71 C	0.030	0.03000	0.0	0.03121	4.0	0.02690	10.3
K2	0.150	0.14999	0.0	0.14982	0.1	0.14766	1.6
KDC	0.030	0.02999	0.0	0.03484	16.1	0.03212	7.1
PCRB	0.025	0.02501	0.0	0.02320	7.2	0.02433	2.7
NCRB	0.250	0.24990	0.0	0.23945	4.2	0.22220	11.1
K1 RC	0.100	0.10000	0.0	0.10772	7.7	0.10130	1.3
<i>K</i> 1 <i>C</i>	2.500	2.50000	0.0	2.67722	7.1	2.53098	1.2
K20 C	0.050	0.05000	0.0	0.04949	1.0	0.04985	0.3

注:目标值按式(1)计算。相对误差计算公式为:|优化参数-参数最优值|/参数最优值。

676

图 3 给出了遗传单纯形、GA、SM 算法的优化进程 对比,其中模拟次数为调用水质模型的次数,可以看 出,尽管 GA 在理论上具有概率 1 的收敛特性,但收敛 速度缓慢,计算冗长,说明其局部搜索能力较弱。SM 算 法属确定性搜索算法,局部搜索能力很强,如图 3 SM 优 化进程曲线,目标值在一定模拟次数内迅速下降并达 到局部最优解。遗传单纯形混合算法融合了 GA 的广 度搜索能力和 SM 的局部搜索能力,目标值稳定下降,





并最终收敛到全局最优解。试验结果表明,遗传单纯形混合算法通过对两种算法的不同搜索机制的融合,丰富了搜索行为,增强了全局和局部意义下的搜索能力和效率。

3.2 带有扰动的情况 带有扰动情况下的参数识别结果如表 3 所示,结果分析如下。

(1)研究为避免随机数产生器种子(seed)的影响,每一组试验随机模拟了5次。考察5次平行试验 目标函数值的"标准差/平均值","标准差/平均值"反映目标函数值的离散程度,可指示算法是否收敛到 全局最优解。如表3所示,不同扰动条件下目标函数值的"标准差/平均值"均为0.0%,说明5次平行试 验均收敛到或近似收敛到一个目标函数值上,算法全局收敛。

表 3 带有扰动情况下的参数识别结果

参数名称	参数最优值		5%扰动情况			10%扰动情况		
		平均值	标准差	相对误差(%)	平均值	标准差	相对误差(%)	
K12C	0.200	0.219	0.001	9.52	0.196	0.000	1.99	
K20 C	0.050	0.053	0.000	5.56	0.063	0.000	25.84	
K1 C	2.500	2.419	0.000	3.25	2.764	0.000	10.57	
K1 RC	0.100	0.097	0.000	3.27	0.112	0.000	11.98	
NCRB	0.250	0.262	0.000	4.99	0.220	0.000	12.03	
PCRB	0.025	0.027	0.000	9.90	0.025	0.000	0.06	
JDC	0.030	0.029	0.000	1.77	0.032	0.000	7.71	
K2	0.150	0.185	0.000	23.30	0.158	0.000	5.00	
K71 C	0.030	0.030	0.000	0.45	0.029	0.000	3.73	
K83 C	0.030	0.032	0.000	6.73	0.032	0.000	6.98	
目,	标值	79.681	0.002		350.57	0	_	

注:"参数最优值"为数据序列合成时的参数取值;"平均值"为5次平行试验参数识别结果的平均值;"标准差"为5次平行试验参数 识别结果的标准偏差;相对误差为+优化参数平均值 - 参数最优值+/参数最优值。"目标值"为算法优化结果的目标函数值(按式(1)计 算)。相对误差反映了参数优化结果偏离"最优目标值"的距离。

(2)在存在观测误差的情况下,遗传单纯形算法同样能较好识别模型参数,大多数参数的相对误差 能控制在15%以下,基本趋势是扰动越大,参数识别误差越大。个别参数的相对误差较大的主要原因 是参数灵敏度和相关性的影响,如 K2 属低灵敏度参数,较小扰动就可导致较大的不确定性。

(3)由于观测误差的存在,最优解一般未知,大多数研究都把合成数据序列时的参数值看作是"最优参数值",表 3 的相对误差计算也是这种思路。研究中将"最优参数值"带入模型,重新计算目标函数值, 5%和 10%扰动情况下的目标函数值分别为:83.6194 和 372.3457,对比表 3,实际优化结果均小于这两

个值,说明表3中的最优参数值并不是真正的最优参数。



图 4 5% 扰动下的参数识别结果



图 5 10% 扰动下的参数识别结果

677

(4)图 4 和图 5 给出了以叶绿素(chl-a)为例的模型模拟值与观测值对比图,从图中可以看出,观测值基本上在水质模型输出曲线上下分布比较均匀,且误差均比较小,参数识别效果较好。

4 结论

(1)GA 是一种高度并行、随机和自适应的通用优化算法。尽管遗传算法在理论上具有概率1的收敛特性,但实际应用时往往出现"早熟"收敛或收敛缓慢等缺点。为克服 GA"早熟"收敛和收敛缓慢的缺点,相关研究人员开展了大量的研究工作。当前的热点方向是开发混合优化算法,融合算法不同的搜索机制,可丰富搜索行为,增强全局和局部意义下的搜索能力和效率。遗传单纯形法就是这样一类算法。

(2)试验结果表明:无论是没有扰动的情况还是有扰动的情况,遗传单纯形混合算法均高效可靠地 搜索到最优解,说明遗传单纯形法应用于环境模型参数优化是可行的。(3)文献研究中一般将合成数据 序列时的参数值看作是"最优参数值",研究表明,在有限的时间序列长度下该参数值并非真正意义上的 最优参数值,表3中的参数识别误差也有这方面因素的影响。

参考 文献:

- Boesel J, Bowden J R, Glover F, et al. Future of simulation optimization [A]. Proc Winter Simulation Conference [C].
 2001.1466 1469.
- [2] Jorgensen S E. An improved parameter estimation procedure in lake modeling [J]. Lake & Reservoirs: Research and management, 1998,3:139 142.
- [3] Wang Q J. The genetic algorithm and its application to calibrating conceptual rainfall-runoff models [J]. Water Resources Research. 1991,27(9):2467 2471.
- [4] McKinney D C, Min D L. Genetic algorithm solution of groundwater management models [J]. Water Resources Research.
 1994,30(6):1897 1906.
- [5] Mulligan A E, Brown L C. Genetic algorithms for calibrating water quality models [J]. Journal of Environmental Engineering, 1998, 124(3): 202 - 211.
- [6] 王凌.智能优化算法及其应用[M].北京:清华大学出版社,2001.
- [7] Goldberg D. Genetic Algorithms in Search, Optimization, and Machine Learning [M]. Addison-Wesley, Reading, MA, 1989.
- [8] 孟庆春,周长久,纪洪波.带有对称编码的遗传算法在动力学系统中的应用[J].电子学报,1999,27(2):59-63.
- [9] Grefenstette J J. Optimization of control parameters for genetic algorithms [J]. IEEE Trans. On Systems, Man, and Cybernetics, SMC-16(1):122 128.
- [10] Nelder J A, Mead R. A simplex method for function minimization [J]. Computer J, 1965,7(2):308 313.
- [11] Duan Q Y, Gupta V K, and Sorooshian S. Shuffled evolution approach for effective and efficient global minimization [J]. Journal of Optimization Theory, 1993,76(3):501 – 521.
- [12] Vrugt J A, Gupta H V, Bastidas L A, et al. Effective and efficient algorithm for multiobjective optimization of hydrologic
 - models [J]. Water Resources Research, 2003, 39(8): No. 1214.
- [13] Ambrose R B, Wool T A, Martin J L, et al. WASP5.x, A Hydrodynamic and Water Quality Model Theory, User's Manual, and Programmer's Guide [M]. Draft: Environmental Research Laboratory, US Environmental Protection Agency, 1993.
- [14] DiToro D M and Matystik J W F. Mathematical Models of Water Quality in Large Lakes Part 1: Lake Huron and Saginaw
 Bay. U. S[R]. EPA/600/3-80-056. Duluth, Minnesota: Environmental Protection Agency, 1980.
- [15] Thomann R V, Fitzpatrick J J. Calibration and Verification of a Mathematical Model of the Eutrophication of the Potomac Estuary [R]. Prepared for Department of Environmental Services. Washington, D C: Government of the District of Columbia, 1982.
- [16] Bowie G L, Mills W B, Porcella D B, et al. Rates, Constants, and Kinetic Formulations in Surface Water Quality 678

Modeling, Second Edition [R]. EPA/600/3 - 85/040. Athens: U.S. Environmental Protection Agency, 1985.

Application of genetic algorithm and simplex method in parameter identification of complicated environmental model

WANG Jian-ping, CHENG Sheng-tong

(Tsinghua University, Beijing 100084, China)

Abstract: Parameter identification plays an important role in application of mathematical model. In order to improve the performance and efficiency of parameter identification in complicated environmental models, the genetic algorithm and simplex method(GASM) are introduced. This hybrid method integrates the search mechanisms of these two methods and greatly elevates the ability of exploration and exploitation. The simulation of water quality in the Miyun Reservoir is presented as an example to demonstrate the application of this method. The optimal identification of 10 parameters in the model simulating the surface water quality(WASP) is carried out. The calculation result indicates that the global

optimized solutions of these parameters can be reliably and effectively acquired either in the condition of non-disturbed data or in the condition of disturbed data. The comparisons among different optimization methods also show that GASM possesses particular advantages in performance and efficiency. Key words: parameter identification; environmental model; genetic algorithm; simplex method; global optimization; hy brid Algoribhms

(责任编辑:吕斌秀)

(上接第 673 页)

•

3-D viscoelastic dynamic analysis of transversely isotropic saturated poroelastic soil in time domain

ZHU Yan-zhi, ZHONG Zheng

(Tongji University, Shanghai 200092, China)

Abstract: Based on the 3-D dynamic equations for transversely isotropic saturated poroelastic media the Fourier expansion approach and Laplace-Hankel transformation are used to solve the viscoelastic dynamic response of the transversely isotropic saturated soil. The general solutions are derived in terms of solid matrix displacements and fluid displacement relative to the soil matrix. The process for calculating dynamic response in poro-viscoelastic half-space is given. The solutions in time domain can be obtained by using numerical inverse Laplace-Henkel transforms, and the computational program is developed. The boundary-value problems of transverse isotropic saturated poro-viscoelastic half-space under the action of vertical and horizontal dynamic loads are solved when initial condition and boundary condition are given. The numerical results for solid matrix displacement and pore pressure are presented. The comparison shows that the proposed solution possesses enough accuracy. Key words: Biot's wave equations; transversely isotropic; saturated poroelastic soil; viscoelastic; dynamic response; Laplace-Hankel transform; numerical integration

(责任编辑:韩昆)