

直接搜索—模拟退火算法在水质模型参数识别中的应用^{*}

申 玮¹, 郭宗楼¹, 周新超²

(1 浙江大学 生物系统工程与食品科学学院, 浙江 杭州 310027;

2 杭州市自来水总公司, 浙江 杭州 310016)

[摘 要] 在研究新近发展起来的模拟退火算法及其各种改进算法的基础上, 提出并构造了一种能识别水质模型参数的、以记忆为基础的直接搜索—模拟退火算法, 并在 O'Connor 水质模型参数识别的实例应用中, 得到了更优的全局最优解, 为水质模型参数识别提供了一条新途径。

[关键词] 直接搜索—模拟退火算法(DSA); 水质模型; 参数识别

[中图分类号] X824

[文献标识码] A

[文章编号] 1671-9387(2004)05-0101-04

参数识别是水质模型建立过程中的核心技术。从运用经验公式/半经验公式的单一参数估算法, 到从模型整体出发的多维参数同时最优估值方法, 如一阶梯度法可以求得不同初始区间的目标局部最优解, 复合形法可以求得非线性多变量整体目标最优解, 阻尼最小二乘法有效地解决了非线性高维目标参数识别, 这些都大大提高了水质模型的可靠性^[1~3]。逐步迭代参数识别优化算法不受初始值的影响, 而且单调收敛, 能迅速求得全局最优解^[4], HSY (Hornberger-Spear-Young) 算法还有效地解决了稀疏数据条件下识别参数并建立水质模型的问题^[5]。但这些参数识别方法都存在受初始条件影响明显, 算法过于复杂, 运算效率低的缺点。

近年来, 遗传算法、模拟退火算法等新型随机非数值优化算法为水质模型的参数识别提供了新途径。遗传算法是基于达尔文生物进化规律的一种群体优化算法, 它同时从多个状态出发, 通过选择、交叉、变异等手段, 不断地向最优解逼近; 而模拟退火是基于热力学原理建立起来的随机优化算法, 它在优化结果解的精度和优化速度上优于遗传算法^[6]。但遗传算法存在种群退化, 模拟退火存在收敛速度慢等不足。一些学者在深入研究的基础上, 对上述2种算法进行了改进, 特别是在模拟退火算法方面取得了喜人的成果。如将 POWELL 算法与模拟退火算法相结合构造的混合优化算法, 通过缩小搜索区

域, 改善了优化解的精度^[7]。还有一种改进后的自适应模拟退火算法, 它通过依据算法搜索过程中的进展信息, 将临域搜索与温度控制有机结合, 从而动态地确定温度参数的变化和对不同的临域施加的搜索次数, 该算法能独立于初始温度的选择, 具有更强的鲁棒性^[8]。但以上这些算法均未从根本上改进模拟退火算法的运算效率和稳定性。

作者基于对模拟退火算法的分析以及对各种模拟退火的改进算法的综合, 提出并构造了一种水质模型参数识别的直接搜索—模拟退火算法(Direct search-Simulated Annealing(DSA)), 经过水质模型参数识别的大量实验测试, 能够显著地改进模拟退火算法的运算效率及其稳定性。现将该方法介绍如下。

1 模拟退火算法在水质模型参数识别中的运用

1.1 水质模型概述

O'Connor 模型是描述河流 DO-BOD (溶解氧-生化需氧量) 水质变化的一种较精确的模型^[9], 其一般形式是:

$$\frac{dL}{dt} = -(k_1 + k_3)L \quad (1)$$

$$\frac{dC}{dt} = -k_1L - \alpha k_4L_N + k_2(C_s - C) \quad (2)$$

^{*} [收稿日期] 2003-04-28

[基金项目] 国家 863 计划资助项目(2001AA 242121); 教育部博士点基金资助项目(1999049306)

[作者简介] 申 玮(1976-), 女, 新疆阿克苏人, 讲师, 硕士, 主要从事水环境规划与控制等研究。

式中, L 为 BOD 的浓度 (mg/L); L_N 为氨氮的浓度 (mg/L); C_s 为饱和 DO 的浓度 (mg/L); C 为 DO 的浓度 (mg/L); α 为硝化系数 (mg/L); k_1 为 BOD 的耗氧系数, d^{-1} ; k_2 为大气复氧系数, d^{-1} ; k_3 为 BOD 沉积系数, d^{-1} ; k_4 为氨氮耗氧系数, d^{-1} 。(1)和(2)式的解析解为:

$$L = L_0 e^{-(k_1 + k_2)t} \quad (3)$$

$$C = C_0 e^{-(k_1 + k_3)t} + \frac{k_1 L_0}{k_1 + k_3 - k_2} (e^{-(k_1 + k_3)t} - e^{-k_2 t}) + \frac{\alpha k_4 L_{N0}}{k_4 + k_3 - k_2} (e^{-(k_4 + k_3)t} - e^{-k_2 t}) + C_s (1 - e^{-k_2 t}) \quad (4)$$

式中, L_0 , L_{N0} 和 C_0 分别为 BOD、氨氮和 DO 的初始浓度。用(3)和(4)式能很容易地计算出起始断面以下河水中 BOD 和 DO 的浓度,从而可估算河流的水环境容量,为总量控制提供依据。但是在采用这 2 个公式之前,首先需要确定或识别公式中 k_1 , k_2 , k_3 和 k_4 这 4 个参数值。通常根据水团追踪实验进行识

别,由于(4)式是高度非线性模型,其误差函数是一个多谷多峰函数,过去在采用一般的非线性优化方法识别参数时,获得的参数往往精度很低,致使计算结果误差较大,难以满足工程技术要求。近年来,模拟退火算法在优化领域里相当活跃,提供了一种崭新的优化方法。作者以某河流的水团追踪实验为例,探讨改进型模拟退火算法在 O'Connor 模型参数中的应用及其优点。

1.2 模拟退火算法在水质模型参数识别中的应用

1.2.1 水质模型参数识别问题的提出

以某河流的水质模型参数识别为例。为计算此河流的水环境容量,在该河流进行了水团跟踪实验。实验前已测定河流中 DO 的初始浓度 C_0 为 4.71 mg/L , BOD 的初始浓度 L_0 为 2.69 mg/L , 氨氮的初始浓度 L_{N0} 为 2.81 mg/L , $\alpha = 2.70 \text{ mg/L}$, 平均气温为 29°C 。DO 的浓度随距离和时间变化的实测数据结果^[9]见表

表 1 某河流水团跟踪实验溶解氧实测数据

Table 1 Data of DO of water pursuing experimentation

断面编号 Number of transversals	时间/d Time	DO 浓度/ ($\text{mg} \cdot \text{L}^{-1}$) Concentration of DO	饱和 DO 的 浓度/ ($\text{mg} \cdot \text{L}^{-1}$) Equilibrium concentration of DO	断面编号 Number of transversals	时间/d Time	DO 浓度/ ($\text{mg} \cdot \text{L}^{-1}$) Concentration of DO	饱和 DO 的 浓度/ ($\text{mg} \cdot \text{L}^{-1}$) Equilibrium concentration of DO
1	0.23	4.03	7.63	4	0.82	4.04	7.58
2	0.35	4.81	7.74	5	0.99	4.91	7.54
3	0.57	3.96	7.62	6	1.16	4.68	7.61

将表 1 中的时间 t_i 和饱和溶解氧浓度 C_{si} 代入(4)式,如果存在一组参数(k_1, k_2, k_3, k_4),使得计算浓度 CC_i 与实测浓度 C_i 的偏差最小,那么这组参数即为所求的最佳参数,也就是完成了该水质模型参

数的识别。

求一组参数(k_1, k_2, k_3, k_4),使得下面目标函数取得最小值(目标函数的理论最小值应为 0):

$$f(k_1, k_2, k_3, k_4) = \sum_{i=1}^6 (CC_i - C_i)^2 =$$

$$\sum_{i=1}^6 \left\{ \left[C_0 e^{-(k_1 + k_3)t_i} + \frac{k_1 L_0}{k_1 + k_3 - k_2} (e^{-(k_1 + k_3)t_i} - e^{-k_2 t_i}) + \frac{\alpha k_4 L_{N0}}{k_4 + k_3 - k_2} (e^{-(k_4 + k_3)t_i} - e^{-k_2 t_i}) + C_s (1 - e^{-k_2 t_i}) \right] - C_i \right\}^2$$

1.2.2 模拟退火算法(SA)在水质模型参数识别中的应用

模拟退火算法是基于 Monte Carlo 方法,按照 Metropolis 准则进行迭代求解的一种启发式随机搜索算法^[10]。它从一个定义域内随机产生的初始点出发进行迭代,在每一步迭代中,算法从当前点 x 产生下一个点 y ,如果 $f(x) \geq f(y)$,则 y 为被接受当前点,否则 y 仅以概率 $\exp[-f(y) - f(x)]/t_k$ 被接受。重复上述过程并减小温度控制参数 t_k ,直至求得最佳解停止。

其算法用伪代码描述如下:

Procedure Simulated-A nnealing;

begin

initialize (i_0, t_0, L_0);

$k := 0$;

$i := i_0$;

repeat

for $l := 1$ to L_k do

begin

generate (j from s_i);

if $f(j) \leq f(i)$ then $i := j$

else

if $\exp\{(f(i) - f(j))/t_k\} > \text{random}[0, 1]$

then $i := j$

end;

```

    k := k + 1;
    Calculate L length ( $L_k$ );
    Calculate control( $t_k$ );
until stopcriterion
end;

```

经编程并在计算机上实现,得出的结果不理想并且还存在不稳定,甚至出现恶化解等现象,这是模拟退火算法本身的缺陷所致。模拟退火算法是一种“随机漫步”优化搜索算法,没有利用已搜索试探过的区域所得的信息引导下一步算法,且很难判断搜索空间中哪些区域更有机会得到最优解。为克服模拟退火算法的上述缺陷,特提出并构造了一种水质模型参数识别的DSA新算法,其在水质模型参数识别实验中得到了更为理想的结果。

2 DSA 算法在水质模型参数识别中的应用

2.1 DSA 算法的思想与描述

以记忆为基础的DSA算法不同于其他改进模拟退火算法,其最主要的特点在于它采用了一组工作点而不是仅仅一个工作点,而且新算法在执行过程中不需要对原有各部分功能进行优化。

DSA算法的基本思想是:考虑优化求解 $\min f(x)$ 时,首先从一个包含 N 个点及其目标值的集合 A 和初始点 x 开始(x 作为初始Markov链的起点),集合 A 中的 N 个点与初始点 x 均在定义域内随机生成,集合 A 中目标函数最优值为 f_l , 目标函数最差值为 f_h ; 在控制参数 t_0 下,由当前点 x 的邻域产生下一个点 y , 如果 $f(y) \leq f_h$, 则 y 被接受为当前点, 否则 y 仅以 $\exp[-(f(y) - f_h)/t_0]$ 被接受; 如果 y 点被接受, 则 y 和 $f(y)$ 取代集合 A 中的最差点与 f_h , 并在集合 A 中重新找出新的 f_h 和 f_l , 重复上述过程直至Markov链终止, 然后减小控制参数 t_0 的值, 重复上述过程, 如此反复, 直至满足某个停止准则时算法终止。其中, 在控制参数为 t_0 时, 其Markov链的长度为 L_t^d , 如果 $f(y) < f_l$, 当前Markov链即终止。

DSA算法用伪代码描述如下:

```

Procedure Direct search-Simulated-Annealing;
begin
    initialize ( $T_0, x$ );
    initialize an array  $A$  of  $N$  points with function values;
    stopcriterion := false;  $t := 0$ ;

```

```

while not stopcriterion do
begin
    check := false;
     $i := 0$ ;
    repeat
    begin
        generate  $y$  from  $x$ ;
        if  $f_y - f_h \leq 0$  then  $x := y$ ;
        else if  $\exp(-(f_y - f_h)/T_t) > \text{random}[0, 1)$ 
then  $x := y$ ;
        if ( $f_y < f_l$ ) then check := true;
        replace the worst point in  $A$  and find the new worst, best points;
         $i := i + 1$ ;
    end;
    until check or  $i = L_t^d$ ;
     $t := t + 1$ ;
    compute  $T_t$ ;
end;
end;

```

为了保证算法的收敛特性,对于冷却进度表,DSA采取如下技术措施:

(1) Markov链的长度(L_t^d): 取决于问题的规模,不能随意决定。这里取 $L_t^d = L_t + L_t \times F$, $L_t = 10 \times n$ (其中 n 是问题的维数); $F = 1 - \exp[-(f_h - f_l)]$, 很明显 $L_t^d \in (L_t, 2L_t)$ 。这样取值的目的,在于确保在给定温度下,出现 f_h 与 f_l 差值增大的情况时,目标函数值具有得以提高的可能。然而,正如前面提到的,如果生成点的目标函数值优于 f_l , 则不需要Markov链的长度加大。

(2) 控制参数衰减因子(α): $\alpha \in [0.8, 0.9]$ 。

(3) 停止准则: 在当前温度比较小,并且集合 A 中的点很密集时,算法可以取得优解。即选用如下停止准则:

$$T_t \leq \epsilon \quad (6)$$

与

$$f_h - f_l \leq \epsilon \quad (7)$$

其中,当目标值较高或较低时,式(7)可以标准化为 $|1 - f_l/f_h| \leq \epsilon$, 也可以得到同样的精确度。

2.2 DSA 算法在水质模型参数识别中的应用

利用上述河流水团跟踪实验已取得的实验数据,运用DSA算法重新识别上述水质模型参数的 k_1, k_2, k_3, k_4 。

经编程并在计算机上再次实现,改进的DSA算法得出了较为理想的结果。将之与其他几种算法对

上述水质模型识别参数 k_1, k_2, k_3, k_4 的计算结果进行比较, 结果见表 2。

由表 2 的对比结果可以得出, 单纯形算法最优解受初始值的影响较大; 遗传算法与改进的 DSA 算法均求得了相对最优解, 但 DSA 算法具有编程简单

易实现, 运算效率高、运用方便的优点; 与模拟退火算法相比, DSA 算法在优化解、运算效率、收敛速度以及稳定性等方面有较大改善, 特别是在结果的稳定性方面得到了显著改进。

表 2 几种算法的结果比较

计算结果 Result	单纯形算法 Simplex algorithm					模拟退火算法 SA	
	算例 1 Example 1		算例 2 Example 2		k_i 范围 Range of value	算例 1 Example 1	
	初始解 Initial solution	最优解 Optimal solution	初始解 Initial solution	最优解 Optimal solution		初始解 Initial solution	最优解 Optimal solution
k_1	0 1	0 1	0 3	0 525	[0 1, 1 0]	0 211 542	0 398 068
k_2	0 1	0 15	0 4	1 050	[0 1, 1 0]	0 337 779	0 998 737
k_3	0 1	0 1	0 5	0 875	[0 1, 1 0]	0 938 804	0 632 249
k_4	0 1	0 1	0 3	0 525	[0 1, 1 5]	0 704 743	0 204 08
目标函数 f Value of aim function		1 740 6		0 867 1			1 405 51
计算结果 Result	遗传算法 Genetic algorithms			直接搜索-模拟退火算法 DSA			
	k_i 范围 Range of value	算例 1 Example 1		算例 1 Exapm le 1		算例 2 Example 2	
		最优解 Optimal solution	k_i 范围 Range of value	初始解 Initial solution	最优解 Optimal solution	初始解 Initial solution	最优解 Optimal solution
k_1	[0 1, 1 0]	0 461 2	[0 1, 1 0]	0 339 120	0 222 406	0 211 542	0 258 068
k_2	[0 1, 1 0]	0 943 8	[0 1, 1 0]	0 472 777	0 999 944	0 337 779	0 999 833
k_3	[0 1, 1 0]	0 775 0	[0 1, 1 0]	0 470 388	0 569 409	0 938 804	0 577 117
k_4	[0 1, 1 5]	0 381 6	[0 1, 1 5]	1 353 960	0 516 488	0 704 743	0 479 087
目标函数 f Value of aim function		0 725 8			0 723 462		0 723 481

3 结束语

直接搜索-模拟退火(DSA)算法改进了水质模型参数识别的精度、运算效率、收敛速度和稳定性。

目前虽已研究并提出了各种水质模型参数识别的方法, 但在实际应用中还要考虑到各随机因素对水质模型参数识别过程中的影响, 需深入研究各随机因素影响下的水质模型参数识别方法。

[参考文献]

- [1] 张鸿志, 成均德 河流水质模型的多维参数同时估值法[J]. 工程勘察, 1990, (6): 29- 31
- [2] 何秉宇, 朱建新 河流水质模型参数的优化估算探讨[J]. 新疆环境保护, 1996, (3): 5- 8
- [3] 刘苏峡 梁溪河多组分水质模型参数的阻尼最小二乘估算法[J]. 水文, 1991, (4): 24- 29
- [4] 王 黎, 唐允吉 河流水质模型参数估计的逐步迭代算法[J]. 四川水力发电, 1991, (6): 11- 15
- [5] 邓义祥, 陈吉宁, 杜鹏飞 HSY 算法中水质模型参数识别的应用探讨[J]. 上海环境科学, 2002, (8): 497- 500
- [6] 郑红星, 李丽娟 水质模型参数的非数值随机优化[J]. 地理研究, 2001, (2): 97- 102
- [7] 王惠东, 张金泉 混合优化算法[J]. 山东矿业学院学报, 1998, (9): 306- 308
- [8] 庞哈利, 郑秉霖, 徐心和 一种自适应的模拟退火算法[J]. 控制与决策, 1999, (9): 478- 480
- [9] 刘国东, 姚 建, 张 翔 水质模型参数识别的遗传算法[J]. 环境污染与防治, 1999, (5): 605- 607
- [10] 康立三, 谢 云, 尤矢勇, 等 非数值并行算法(第一册). 模拟退火算法[M]. 北京: 科学出版社, 1994

(下转第 109 页)

An exploration into various proofs of a theorem about characteristic vectors and zero-input responses

WANG Zhao-qi, XIE Yuan-yuan

(Shaanxi Polytechnic Institute, Xi'an, Shaanxi 712000, China)

Abstract Based on discussing the relationship between characteristic vectors and zero-input-responses, a theorem is established which is applicable for linear, time-invariant and continuous systems. In addition, different proofs and several examples are given to illustrate the theorem.

Key words: linear and time-invariant system; characteristic vector; zero-input response

(上接第 104 页)

Using direct search-simulated annealing algorithm to identify parameters of a water quality model

SHEN Wei¹, GUO Zong-lou¹, ZHOU Xin-chao²

(1 College of Biosystem Engineering and Food Sciences, Zhejiang University, Hangzhou 310027, China;

2 Hangzhou Water Company, Hangzhou, Zhejiang 310016, China)

Abstract Parameter identification is a key technique in the establishment of water quality model. A direct search-simulated annealing algorithm (DSA) for parameter identification of a water quality model, which is memory-based, is proposed and has been applied in the parameter identification of O'Connor water quality model. Results of a case show an improved performance in finding the global minimum. DSA provides a new way to identify parameters of water quality model.

Key words: direct search-simulated annealing (DSA); water quality model; parameter identification