

粒子群优化算法在确定河流水质参数中的应用

郭建青¹,李彦²,王洪胜¹,周宏飞²

(1. 长安大学环境科学与工程学院,陕西 西安 710054; 2. 中国科学院新疆生态与地理研究所,新疆 乌鲁木齐 830011)

摘要 将粒子群优化(PSO)算法应用于求解分析瞬时投放示踪剂情况下的一维河流水团示踪试验数据,以及确定河流水质参数的函数优化问题。分别就粒子数目和待估水质参数的初始取值范围对算法运算过程的影响进行了数值实验。结果表明:①PSO 算法能够有效地应用于求解分析河流水质试验数据,确定水质参数的函数优化问题;②粒子数目的多少对迭代次数、运算时间和算法是否收敛有一定的影响,在粒子数目较大的情况下,可以保证运算过程收敛;③待估参数初始猜测值的选取范围对迭代次数也有一定的影响,选取范围越大,需要的迭代次数越多。最后,指出了需要进一步研究的问题。

关键词 河流水质 参数估计 粒子群算法 算法收敛性 粒子数目

中图分类号 :X824 ;TP274 **文献标识码** :A **文章编号** :1006-7647(2007)06-0001-05

Application of particle swarm optimization algorithms to determination of water quality parameters of river streams//
GUO Jian-qing¹, LI Yan², WANG Hong-sheng¹, ZHOU Hong-fei² (1. *Institute of Environmental Science & Engineering, Chang'an University, Xi'an 710054, China*; 2. *Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China*)

Abstract : The particle swarm optimization (PSO) algorithms were applied to analysis of 1D tracing test data of river streams with tracers instantaneously injected, and further to optimization of functions to estimate the water quality parameters of river streams. The influences of the number of particles and the range of initial values adopted for water quality parameters to be evaluated on the convergence of the PSO algorithms were studied by numerical experiments. The result shows that the PSO algorithms are effective for function optimization to estimate water quality parameters of river streams with tracing test data. Some conclusions are also drawn: the convergence of PSO algorithms and the computation time are influenced by the number of particles and the large number of particles can ensure the convergence of the calculation; the range of initial values of the parameters to be evaluated is of certain influences on the number of iteration, and the larger the range, the more the iteration number and the computation time. Finally, some problems for further research were put forward.

Key words : water quality of river stream; parameter evaluation; particle swarm algorithm; convergence of algorithm; number of particle

水质参数是建立河流水质数学模型、预报河流水质状态变化的基本数据。目前获得水质参数的主要途径之一,就是利用以相应的解析解为基础而建立的数据分析方法,分析河流水团示踪试验数据。例如,分析瞬时投放示踪剂条件下的一维河流水团示踪试验数据,确定河流纵向离散系数的方法主要有以相应解析解为基础的单站法、双站法、演算法^[1]、直线图解法^[2]、相关系数极值法^[3]和非线性最小二乘法^[4]等。然而,这些方法在实际应用中均具有一定的局限性。近年来,智能优化算法的研究与应用越来越受到人们的重视^[5],特别是神经网络优化算法、遗传算法及相应的混合算法已在水环境模型参数识别方面得到了一定的应用^[6-8]。笔者也曾经应用改进模拟退火算法^[9]、混沌序列优化算法^[10]、模拟退火-单纯形混合算法^[11]等智能优化方法对求解分析瞬时投放示踪剂情况下的一维河流水团示踪试验数据及确定河流水质参数的函数优化问题进行了初步尝试,并取得了一些初步成果。从 20 世纪 90 年代初以来,一些学者模拟自然生物群体的行为方式,提出了群体智能优化方法。Dorigo 等从生物进化的机理中受到启发,通过模拟蚂蚁的群体寻径行为,提出了蚁群优化方法^[12];Kennedy 等通过

基金项目 国家自然科学基金(40671037)
作者简介 郭建青(1958—),男,陕西宜君人,教授,从事地下水动力学与环境水力学研究。E-mail: guojianqing@yahoo.com.cn

对鸟群、鱼群寻食行为的模拟,于 1995 年提出了粒子群优化 (particle swarm optimization, PSO) 算法^[13]。通常单个自然生物的行为并不是智能的,但是整个生物群体的行为却表现出处理复杂问题的能力,群体智能就是这些团体行为在人工智能优化问题中的应用。由于 PSO 算法也是基于群体的优化方法,与其他进化方法相比,它的主要特点表现在^[14] ①每一个体(称为粒子)都被赋予了一个随机速度,并在问题的状态空间中运动;②个体具有记忆功能;③个体的进化主要通过个体之间的合作与竞争来实现;④PSO 算法的运算过程实现简单,需要调整的算法控制参数少。作为一种高效并行运算的优化方法,PSO 算法能够有效地应用于求解非线性、不可微和多峰值的复杂函数优化问题。因此,其应用得到了迅速的发展。

笔者尝试应用 PSO 算法求解分析瞬时投放示踪剂情况下的一维河流水团示踪试验数据、估计河流水质参数的函数优化问题,并对诸如粒子数目和惯性权重系数等算法控制参数及待估水质参数的初始取值范围对算法收敛速度的影响进行数值实验,根据实验结果分析和讨论这些因素对算法收敛性的影响,并指出需要进一步研究的问题。

1 粒子群算法原理与实现

1.1 算法原理

PSO 算法不像遗传算法那样对个体进行选择、交叉和变异操作,而是将群体中的每一个个体视为多维搜索空间中的没有质量和体积的粒子(点),这些粒子在搜索空间中以一定的速度飞行,并根据粒子本身的飞行经验以及同伴的飞行经验对自己的飞行速度进行动态调整,即每一个粒子通过统计迭代过程中自身的最优值和群体的最优值来不断地修正自己的运动方向和速度大小,从而形成群体寻优的正反馈机制。PSO 算法就是这样依据每个粒子对环境的适应度将个体逐步移到较优的区域,并最终搜索、寻找到问题的最优解。

1.2 算法的实现

在 PSO 算法中,用粒子的位置表示待优化问题的解,每个粒子性能的优劣程度取决于待优化问题的目标函数确定的适应值,每个粒子由一个速度矢量决定其飞行方向和速率大小。设在一个 D 维的目标搜索空间中,有 m 个粒子组成一个群体,其中,在第 t 次迭代时粒子 i 的位置表示为 $x_i(t) = (x_{i1}(t), x_{i2}(t), \dots, x_{iD}(t))$,相应的飞行速度表示为 $v_i(t) = (v_{i1}(t), v_{i2}(t), \dots, v_{iD}(t))$ 。在开始运用 PSO 算法进行计算时,首先随机初始化 m 个粒子的

位置和速度,然后通过迭代寻找最优解,在每一次迭代中,粒子通过跟踪两个极值来更新自己的速度和位置:一个极值是粒子本身迄今搜索到的最优解,称为个体极值,表示为 $p_i(t) = (p_{i1}(t), p_{i2}(t), \dots, p_{iD}(t))$;另一个极值是整个粒子群到目前为止找到的最优解,称为全局极值,表示为 $p_g(t) = (p_{g1}(t), p_{g2}(t), \dots, p_{gD}(t))$ 。在第 $t+1$ 次迭代计算时,粒子 i 根据下列规则来更新自己的速度和位置:

$$v_{ik}(t+1) = \omega v_{ik}(t) + c_1 \text{rand}_1(p_{ik}(t) - x_{ik}(t)) + c_2 \text{rand}_2(p_{gk}(t) - x_{ik}(t)) \quad (1)$$

$$x_{ik}(t+1) = x_{ik}(t) + v_{ik}(t+1) \quad (2)$$

式中: ω 为惯性权重系数; c_1 和 c_2 为加速因子,一般取 $c_1 = c_2 = 2.0$; $k = 1, 2, \dots, D$; rand_1 和 rand_2 为在 $[0, 1]$ 区间内变化的随机函数。

1.3 算法步骤

标准 PSO 算法的步骤为^[14] ①随机初始化粒子群体的位置和速度。通常是在允许的范围内随机产生,每个粒子的 p_{best} 坐标设置为其当前位置,且计算出相应的个体极值(即个体的适应度值),而全局极值(即全局的适应度值)就是个体极值中最好的,记录该最好值的粒子序号,并将 g_{best} 设置为该最好粒子的当前位置。②计算每个粒子的适应值,粒子初始速度值一般取为零。③对每个粒子,将其适应值与个体极值进行比较,如果较优,则更新当前的个体极值。④对每个粒子,将其适应值与全局极值进行比较,如果较优,则更新当前的全局极值。⑤根据式(1)、式(2),更新每个粒子的飞行速度和位置。⑥如果达到预先设定的停止准则(一般设置为目标函数小于某个数值或者允许最大迭代次数 N),则停止计算;若未达到,则返回步骤②继续进行迭代计算。

以上算法步骤可以用如图 1 所示的流程表示。

1.4 算法控制参数

根据上述介绍,可以知道标准 PSO 算法的主要控制参数有粒子的数目 n 、惯性权重系数 ω 和 2 个加速因子 c_1 与 c_2 。在有关研究 PSO 算法的文献中,对惯性权重系数 ω 的介绍较多,研究结果表明,较大的 ω 值有利于跳出局部极小值,较小的 ω 值有利于算法收敛,在实际运算中一般采用动态的 ω ,即在搜索初始阶段采用较大的数值,然后按一定的规则逐渐减小。本文采用文献^[15]介绍的公式进行线性衰减:

$$\omega = \omega_{\max} - k \frac{(\omega_{\max} - \omega_{\min})}{N} \quad (3)$$

式中: $\omega_{\max}$ 为最大惯性权重系数,一般取 0.9; $\omega_{\min}$ 为最小惯性权重系数,一般取 0.1; k 为实际循环次数; N 为最大迭代次数。

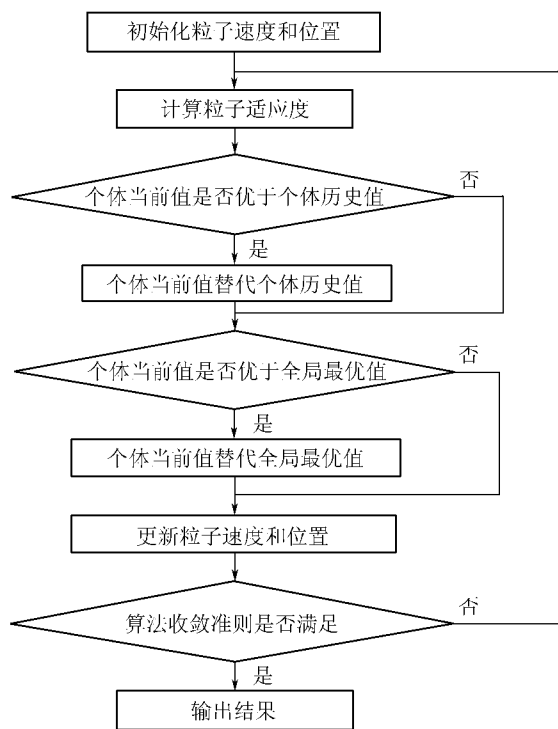


图1 粒子群优化算法流程

另一方面,目前有关粒子数目的多少和待估参数初始取值范围对算法收敛性影响的研究报道较少。为了便于将粒子群优化算法应用于求解分析河流水质试验数据,确定河流水质参数,本文将对其不同的取值进行数值实验,分析它们对算法收敛性的影响。

2 河流水质参数优化问题的构造

在分析河流水团示踪试验数据,确定纵向离散系数 D_L 与断面平均流速 v 时,多采用以下描述瞬时投放示踪剂条件下的一维河流水团示踪试验的解析解^[1] :

$$c = \frac{M_0}{A \sqrt{4\pi D_L t}} \exp\left[-\frac{(x - vt)^2}{4D_L t}\right] \quad (4)$$

式中: c 为示踪剂质量浓度, $[\text{M} \cdot \text{L}^{-3}]$; M_0 为瞬时投放示踪剂的质量, $[\text{M}]$; A 为河流横断面面积, $[\text{L}^2]$; D_L 为河流纵向离散系数, $[\text{L}^2 \cdot \text{T}^{-1}]$; x 为采样点与投放点间的距离, $[\text{L}]$; t 为时间, $[\text{T}]$; v 为河流断面平均流速, $[\text{L} \cdot \text{T}^{-1}]$ 。需要说明的是,在应用式(4)时,应注意满足其使用条件^[1]。

在应用 PSO 算法时,要求欲估计的河流水质参数值能够使式(5)达到极小:

$$\varphi(\theta) = \frac{1}{n} \sum_{i=1}^n (c_i^o - c_i^c)^2 \quad (5)$$

式中: c_i^o 为在第 i 时刻观测到的实际示踪剂质量浓度值, $[\text{M} \cdot \text{L}^{-3}]$; c_i^c 为利用式(1)计算的第 i 时刻的示踪剂质量浓度值, $[\text{M} \cdot \text{L}^{-3}]$; θ 为待估河流水质参

数向量; i 为示踪剂浓度观测时间序列数 ($i = 1, 2, \dots, n$)。选取适当的参数 θ 值,使得浓度计算值与浓度观测值的离差平方和的均值达到极小,此时对应的河流水质参数值即为问题所求。对于由式(5)构成的河流水质参数的确定问题,设河流的纵向离散系数 D_L 为 θ_1 , 河流断面平均流速 v 为 θ_2 , 排放参数 $\frac{M_0}{A}$ 为 θ_3 。

在此,需要指出的是在构成目标函数时,除了要求待优化函数能够计算其函数值外,再无其他附加的要求。因此,粒子群优化算法也能够用于估计其他水质解析模型情况下的水质参数,也就是说,粒子群优化算法在水质参数估计方面应具有更宽的应用范围。

3 算例与讨论

3.1 实验条件与结果

3.1.1 实验数据

原始数据引自文献[2]。表1给出了参数真值 $\theta_1 = D_L = 3000 \text{ m}^2/\text{min}$, $\theta_2 = v = 30 \text{ m/min}$ 与 $A = 20 \text{ m}^2$ 时,不同时间所对应的原始数据 c_i , 且已知 $M_0 = 10 \text{ kg}$, $x = 500 \text{ m}$, $\theta_3 = \frac{M_0}{A} = 0.5$ 。

表1 $c_i \sim t_i$ 原始数据

t_i / min	6	10	12	14	16	20	24	36
$c_i / (\text{mg} \cdot \text{L}^{-1})$	0.254	0.583	0.649	0.663	0.642	0.552	0.444	0.197

3.1.2 实验条件

数值实验是采用 VB 语言,根据图1所示框图编制的程序进行的。实验中,粒子数目分别取5, 10, 20, 50, 100, 200, 500 和 1000 等8种情况,待估水质参数初始取值范围为:最小值均为0,最大值分别为参数真值的2, 5, 10, 20, 50 和 100 倍。

本算例根据目标函数绝对值的大小判断算法是否收敛,其标准为 $\epsilon_1 = 1.0 \times 10^{-6}$ 。如果算法的迭代次数超过最大迭代次数 $N = 1000$ 次,且参数计算结果与其真值的相对误差大于2%,则认为计算失败。在利用实际试验资料进行计算时,需要根据实际情况调整收敛标准。

3.1.3 实验结果

不同粒子数目、待估水质参数初始取值范围条件下的迭代次数和运算时间的数值实验结果见表2。在此有必要说明的是表2中不同情况下的迭代次数能够在重复实验中重现,而运算时间会因计算机在不同时刻可能运行着不同的后台程序,在重复运算中可能会出现不同的结果,但其结果能够反

表 2 算法迭代次数和粒子数目及运算时间之间的关系

粒子数目	待估参数的 倍数	迭代次数	运算时间 / 10^{-5} s	粒子数目	待估参数的 倍数	迭代次数	运算时间 / 10^{-5} s	粒子数目	待估参数的 倍数	迭代次数	运算时间 / 10^{-5} s
5	2	655	< 1.0	20	50	*	*	200	10	550	4.63
5	5	655	< 1.0	20	100	*	*	200	20	623	5.79
5	10	788	< 1.0	50	2	428	1.16	200	50	748	6.94
5	20	*	*	50	5	495	1.16	200	100	937	9.26
5	50	*	*	50	10	631	2.32	500	2	322	6.94
5	100	*	*	50	20	681	1.16	500	5	496	11.57
10	2	528	< 1.0	50	50	902	2.32	500	10	527	11.57
10	5	564	1.16	50	100	*	*	500	20	564	12.73
10	10	687	1.16	100	2	405	1.16	500	50	715	16.20
10	20	*	*	100	5	422	2.31	500	100	815	18.52
10	50	*	*	100	10	622	3.47	1000	2	305	13.89
10	100	*	*	100	20	642	2.31	1000	5	441	19.68
20	2	453	1.16	100	50	793	3.47	1000	10	439	19.68
20	5	657	< 1.0	100	100	*	*	1000	20	543	24.31
20	10	718	< 1.0	200	2	365	3.47	1000	50	718	21.25
20	20	761	1.16	200	5	464	4.63	1000	100	801	35.88

注：* 表示在迭代 1000 次后计算结果仍没有满足精度要求。

映出一定的运算耗时的趋势。

3.2 实验结果分析

3.2.1 粒子数目的影响

从表 2 可以看出,粒子数目的多少对算法的运算过程具有明显的影响。当粒子数目大于或等于 200 时,在所有的待估水质参数取值范围情况下运算过程均收敛。随着粒子数目的增大,迭代次数虽然减少,但运算时间却在增加。在实验条件下,当粒子数目小于或等于 100 和待估水质参数取值范围较大的情况下,会出现运算失败(不收敛)的情况。这种现象表明,在应用 PSO 算法时宜采用较多的粒子数目进行运算,尽管在这种情况下需要的运算时间较长,但可以保证运算过程收敛。

3.2.2 参数初值取值范围的影响

从表 2 可以明显地看出,待估水质参数初始取值范围对运算过程也有明显的影响。就总体趋势而言,参数初始值范围越大,需要的迭代次数越多,相应的运算时间越长。另一方面,在输入较小的参数初始值范围的情况下,如待估水质参数初始猜测值的可能最大值等于或小于参数真值的 10 倍时,在实验条件下所有的运算过程均收敛,而且,取值范围越小收敛速度越快;在待估水质参数初始猜测值的可能最大值等于或大于参数真值的 20 倍的情况下,当粒子数目较小时会出现运算失败的现象,而在粒子数目较大时,可以保证不同待估水质参数初始取值范围搜索过程均收敛。上述现象表明,分析者在应用 PSO 算法时,应根据自己的专业知识,初步估计较为适当的水质参数的可能取值范围,以保证较快得到运算结果;另一方面,如果不能选取比较接近实际参数的初值,则需要设置较大的粒子数目,这样尽

管迭代次数增加,但能保证运算过程收敛。

4 需要进一步研究的问题

根据文中数值实验的初步结果可以知道,PSO 算法能够有效地应用于求解分析瞬时投放示踪剂情况下的一维河流水团示踪试验数据、估计河流水质参数的函数优化问题,较其他现有方法具有原理简单、编程容易、需要人们预先确定和调整的算法控制参数少及运算过程较为稳定等优点。然而,由于这一算法出现的时间较短,目前的研究多集中在算法本身上,就算法的验证而言,多采用数学上的标准测试函数的运算结果评价算法的优劣,而将其应用于解决实际水环境数学模型参数确定问题的研究工作开展得较少。就笔者的初步研究实践而言,认为需要进一步研究以下问题:

a. PSO 算法需要调整的算法控制参数较少,其中最主要的是惯性权重系数。根据相关算法研究的文献报道,随着迭代次数的增加,其数值应由大逐步减小,以利于算法在迭代初始阶段进行全局搜索,在运算后期加强局部搜索能力。文中采用的是文献 [15] 推荐的如式(3)表示的线性变化方式,这种变化方式存在的主要问题是其变化过程受到设定最大迭代次数 N 的影响,因此,需要进一步研究的问题是提出新的惯性权重系数的变化方式,即保证算法既具有良好的收敛性,又使其不受设定最大迭代次数 N 的影响。

b. 文中的研究建议在以式(4)表示的河流水质解析模型的基础上,这个解析模型中含有 3 个待定的河流水质参数,运算过程为在三维空间上的状态搜索,可以推断 PSO 算法对于具有少于 3 个待定参

数的河流水质解析模型也具有好的应用性。理论上讲,也应能够应用于含有多于 3 个待估参数的更高维的河流水质解析模型,但在具体运算时,适当设置粒子数目和最大迭代次数,则属于需要进一步研究的问题。

c. 根据文中的数值实验结果可以知道,PSO 算法的收敛过程明显受到待估水质参数初始值选取范围的影响。因此,需要进一步研究 PSO 算法与其他智能优化算法的混合,提出新的混合算法,以削弱待估水质参数初始值选取范围对运算过程的影响,使其具有更好的适用性。

参考文献：

[1] 余常昭. 环境水力学[M]. 北京:清华大学出版社,1992: 180-181.
[2] 郭建青,温季. 示踪试验确定河流纵向弥散系数的直线图解法[J]. 环境科学,1990,11(2): 24-27.
[3] 郭建青,王洪胜,李云峰. 确定河流离散系数的相关系数极值法[J]. 水科学进展,2000,11(4): 387-391.
[4] 郭建青. 非线性最小二乘法在分析河流水文示踪试验数据中的应用[J]. 中国给水排水,1991,7(5): 13-17.
[5] 王凌. 智能优化算法及其应用[M]. 北京:清华大学出版社,2001: 17-33.

[6] 刘毅,陈吉宁,杜鹏飞,等. 环境参数优化方法的比较[J]. 环境科学,2002,23(2): 1-6.
[7] 王建平,程声通. 遗传单纯形混合算法在复杂环境模型参数识别中的应用[J]. 水利学报,2005,36(6): 40-45.
[8] 杨晓华,杨志峰,酆建强. 水环境模型参数识别的一种新方法[J]. 水科学进展,2003,14(5): 21-24.
[9] 郭建青,李彦,王洪胜,等. 利用改进 SA 算法估计河流水质参数的仿真实验[J]. 系统仿真学报,2003,15(12): 94-96,106.
[10] 郭建青,李彦,王洪胜,等. 利用混沌优化算法确定河流水质模型参数[J]. 水力发电学报,2004,15(4): 95-99.
[11] 郭建青,李彦,王洪胜,等. 应用单纯形-模拟退火混合算法估计河流水质参数[J]. 水科学进展,2004,15(6): 78-82.
[12] DORIGO M, MANIEZZO V, COLORI A. Ant system: optimization by a colony of cooperating agents[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1996, 26(1): 28-41.
[13] KENNEDY J, EBERHART R C. Particle swarm optimization [C]//IEEE International Conference on Neural Networks. Perth, Australia: IEEE Service Center, 1995: 1942-1948.
[14] 范娜,云庆夏. 粒子群优化算法及其应用[J]. 信息技术, 2006(1): 53-56.
[15] 孟红记,郑鹏,梅国晖,等. 基于混沌序列的粒子群优化算法[J]. 控制与决策,2006(3): 25-28.

(收稿日期: 2007-02-07 编辑: 骆超)

· 简讯 ·

2007 重大水利水电科技前沿院士论坛暨
首届中国水利博士论坛在河海大学举办

2007 重大水利水电科技前沿院士论坛暨首届中国水利博士论坛于 11 月 10~11 日在河海大学举办。全国政协副主席、中国工程院院士钱正英出席开幕式。江苏省委书记、省长梁保华会见前来出席论坛的 20 位院士和水利部领导。省委副书记张连珍、水利部副部长胡四一出席并致辞,水利部原部长汪恕诚,长江三峡工程开发总公司原总经理、中国工程院院士陆佑楣以及江苏省人民政府副省长黄莉新出席开幕式。

本次论坛由中国工程院土木、水利、建筑工程学部和中国水利学会主办,河海大学、南京水利科学研究所、水文水资源与水利工程科学国家重点实验室、水资源高效利用与工程安全国家工程研究中心承办。论坛开幕式由河海大学党委书记林萍华主持。中国科学院院士张楚汉、林皋、汪集,中国工程院院士陆佑楣、崔俊芝、吴中如、郑守仁、陈厚群、马洪琪、曹楚生、王浩、周君亮、梁应辰作大会主题报告,中国科学院院士薛禹群、刘昌明,中国工程院院士王思敬、张耀明、茆智、卢耀如以及来自全国水利水电系统的 400 余名专家、学者和博士研究生出席了论坛。除大会主题报告外,与会代表还分别在“土石坝、地下工程和高边坡的设计理论与方法”、“混凝土坝和大型水工结构的设计、施工与安全”、“水利水电工程与水力学问题”、“水资源利用与可持续发展”、“大型水利水电工程对生态、环境与社会的影响和对策”5 个分会场进行了特邀报告与论文交流。

本次论坛既是院士论坛,也是博士论坛,其宗旨是为水利学科的学者们提供学术交流与科研创新的高层次平台,促进水利科技的发展以及水利工程科学与人文社会科学的交叉和融合,促进水利学科拔尖创新人才的培养。

(本刊编辑部供稿)