

基于自由搜索人工神经网络的坡地入渗量预测

李新虎¹, 张展羽¹, 杨洁², 张国华³, 王斌⁴, 王超¹

(1. 河海大学水利水电学院, 南京 210098; 2. 江西水土保持科学研究所, 南昌 330000;
3. 中国灌溉排水发展中心, 北京 100054; 4. 东北农业大学水利与建筑学院, 哈尔滨 150030)

摘 要: 该文应用基于自由搜索算法的 BP (backpropagation) 网络模型对自然降雨条件下不同处理措施的红壤坡地入渗规律进行了预测, 选择降雨量、最大降雨强度、降雨历时、土壤初始含水率、土壤体积质量、通气孔度和下垫面状况 7 项指标作为网络输入, 土壤入渗量单项指标作为网络输出, 结果表明: 基于自由搜索算法的 BP 网络模型可以有效地预测自然降雨条件下不同处理措施坡地入渗规律, 预测的平均相对误差为 11.08%, 经 t 检验和回归分析表明预测值和实测值相差不大, 具有较好的一致性, 决定系数为 0.9715, 并和传统的 BP 网络进行了比较, 结果显示基于自由搜索算法的 BP 网络预测优于传统的 BP 网络, 模型具有较高的精度和稳定性。

关键词: 降雨, 入渗, 反向传播, 自由搜索算法, 坡地, 人工神经网络

doi: 10.3969/j.issn.1002-6819.2009.12.034

中图分类号: S210

文献标识码: A

文章编号: 1002-6819(2009)-12-0193-05

李新虎, 张展羽, 杨洁, 等. 基于自由搜索人工神经网络的坡地入渗量预测[J]. 农业工程学报, 2009, 25(12): 193—197.

Li Xinhu, Zhang Zhanyu, Yang Jie, et al. Prediction of slope infiltration based on artificial neural networks by free search[J]. Transactions of the CSAE, 2009, 25(12): 193—197. (in Chinese with English abstract)

0 引言

降雨入渗是坡地水文循环的中心环节, 水分入渗影响了降雨在地表径流、壤中流、地下水的再分配, 长期以来是土壤侵蚀、非点源污染、水资源管理等科学研究的重点内容^[1]。对于入渗模型许多学者进行了研究, 建立了一系列的经验性统计模型^[2]和物理性理论模型^[3-4], 由于影响入渗的因素较多且关系复杂, 所得模型的复杂性及适用性各异, 而且其经验参数的适用范围有限, 难以准确确定, 物理性理论模型具有一定的通用性, 但参数较多, 应用不便^[5]。人工神经网络 (artificial neural network, ANN) 特别是 BP (反向传播网络, backpropagation network) 网络模型具有自组织、自适应及自学习的功能, 非常适合用来模拟、处理影响因素多、关系复杂的系统, 但是 BP 网络模型存在容易陷入局部最优值等缺陷, 因此本文引入一种新的算法自由搜索算法 (free search, FS)^[6], 利用自由搜索算法来优化 BP 网络的初始权重来增加其速率和稳定性, 通过利用自由搜索算法和人工神经网络相结合对自然降雨条件下的坡地入渗规律进行预测, 为自然降雨条件下坡地入渗规律的进一步研究提供参考。

1 概述

1.1 FS 算法

收稿时间: 2009-03-02 修订时间: 2009-09-06

基金项目: 江苏省高校研究生创新计划 (CX09B_168Z); 水利部公益性行业科研专项经费项目 (200801048)

作者简介: 李新虎 (1981—) 男, 山东泰安人, 博士生, 主要从事农业水土工程的研究。南京 河海大学水利水电学院, 210098。

Email: lixinhu721@163.com

FS 是 Kalin Penev 和 Guy Littlefair 提出的一种新算法^[6], 该算法原理简单, 需要用户确定的参数不多, 操作也很简便, 是一种基于群体的优化方法。FS 中所仿照的是一些高等群居动物在进行生物行为时, 如寻找水源或食物, 从种群整体而言, 处于一种有序的进程之中, 其中的个体又存在着与整体进程相协调的个体随意行为^[7]。在 FS 算法中, 动物群体的搜索行为没有被限定, 而是通过概率描述, 这和其他算法不同^[8], 与粒子群算法 (particle swarm optimization, PSO)^[9-10]、遗传算法^[11] (genetic algorithm, GA) 和差分进化算法^[12] (differential evolution, DE) 相比, 其原理简单, 需要确定的参数不多, 操作也很简便, 具有较好的寻优能力。

1.1.1 算法模型

FS 适用于解决实数空间的优化问题, 在求解机制上采用的是一种具有个体意识的、自由的、不确定的个体行为方式。通过初始化、探查和终止 3 个步骤, 来确定目标函数 (优化函数) 的最优值 (最大值或最小值)。具体算法结构如下:

1) 初始化

$$x_{0ji} = X_{\min i} + (X_{\max i} - X_{\min i}) \text{random}_{ji}(0,1) \quad (1)$$

式中: $j=1, 2, \dots, m$; $i=1, 2, \dots, n$; x_{0ji} ——动物个体的初始位置分量; $X_{\min i}$, $X_{\max i}$ ——搜索空间边界; $\text{random}_{ji}(0,1)$ ——介于 [0, 1] 之间的随机数。

2) 探查

首先, 通过下式探查行走, 更新动物个体位置

$$x_{tji} = x_{0ji} - \Delta x_{tji} + 2\Delta x_{tji} \text{random}_{tji}(0,1) \quad (2)$$

式中: x_{tji} ——更新后的动物个体位置分量; t ——当前步伐, $t=1, 2, \dots, T$, T 为每次行走的限制步伐数;

$random_{ji}(0,1)$ ——介于 $[0, 1]$ 之间的随机数。

在自由搜索算法模型中, 个体移动一个搜索步, 每个搜索步包含 T 小步, 个体在多维空间作小步移动, 其目的是发现目标函数更好的解。

修改策略为 Δx_{ji}

$$\Delta x_{ji} = R_{ji}(X_{\max i} - X_{\min i})random_{ji}(0,1) \quad (3)$$

式中 R_{ji} ——邻域距离; $R_{ji} \in [R_{\min}, R_{\max}]$ 。

邻域距离是改变个体搜索范围的工具, 邻域距离反映了个体的灵活性, 它没有严格的定义, 仅受到整个搜索空间的约束。

在探查行走过程中, 动物个体行为可表示如下

$$f_{ij} = f(x_{ij}), f_j = \max(f_{ij}) \quad (4)$$

信息素 PH_j 按下式更新

$$PH_j = f_j / \max(f_j) \quad (5)$$

信息素大小和目标函数解的质量成正比, 完成 1 个搜索步后, 信息素将完全更新。

敏感性 SE_j 按下式更新

$$SE_j = SE_{\min} + \Delta SE_j \quad (6)$$

式中: $\Delta SE_j = (SE_{\max} - SE_{\min})random_j(0,1)$, $SE_{\min} = PH_{\min}$, $SE_{\max} = PH_{\max}$; $random_j(0,1)$ ——介于 $[0, 1]$ 之间的随机数, 与 (1), (2) 式 $random$ 意义相同, 下标表示各自的计算变量。

敏感性是自由搜索算法的 1 个重要参数, 个体可以搜索任何区域, 可以在其自身的当前最佳值邻域周围, 也可在其他个体或群体发现的当前最佳值邻域周围。增大敏感性, 个体将趋近整个群体的当前最佳值, 局部搜索。减小敏感性, 个体可以在其他邻域进行搜索, 全局搜索。敏感性是自由搜索算法中的一个创新, 在其他算法中没有类似的概念。

最后, 选择下一次探查行走的开始位置

$$x'_{0ji} = x_{ij}(PH_k \geq SE_j) \quad k = 1, 2, \dots, m \quad (7)$$

3) 结束

算法接受的结束准则如下

$$(1) \text{ 寻优准则} \quad f_{\max} \geq f_{\text{opt}} \quad (8)$$

$$(2) \text{ 迭代准则} \quad g \geq G \quad (9)$$

$$(3) \text{ 复合准则} \quad (f_{\max} \geq f_{\text{opt}}) \parallel (g \geq G) \quad (10)$$

式中: f_{\max} ——寻优结果; f_{opt} ——可接受的函数值;

G ——迭代限制条件数值; g ——当前迭代状态的数值。

在自由搜索优化算法中, 灵敏度、搜索步、信息素可以实现自适应调整, 它们通过概率的方法在随机搜索中完成。

1.2 BP 模型

人工神经网络是 19 世纪 80 年代中后期迅速兴起的一门非线性科学, 由于其大规模并行处理、分布式储存、高度非线性关系、学习能力强等优良特征, 在很多领域都有已经得到应用。BP 模型是 PDP (parallel distributed processing) 小组在 1985 年提出的, 是前向网

络的核心部分, 具有良好的非线性映射能力、结构简单, 性能良好, 它体现了神经网络中最精华最完美的内容。

2 模型构建

BP 网络的初始权值是随机赋的, 而最终权值的确定结果往往很大程度上依赖初始权值的选择, 因为在 BP 网络误差曲面存在多个局部最优点, 如果初始权值设置不理想, BP 网络容易陷入局部最优值^[13]; 在 FS 算法中, 动物群体的搜索行为没有被限定, 而是通过概率描述, 因此利用 FS 算法优化 BP 网络的权值, 通过优化后的权值输入网络, 从而使得传统 BP 神经网络对初始权值敏感和易陷入局部最优值等问题得到较好解。

按神经网络生成初始权重的常规办法来生成网络的权重, 任一组完整的神经网络权重 W_i ($i=1, 2, \dots, m$), 相当于 1 个动物个体, 这样的个体有 m 个, 即群体大小。

优化目标函数即为均方误差函数

$$E = \frac{1}{N} \sum_{i=1}^N (y_i^d - y_i)^2 \quad (11)$$

式中: E ——均方误差; N ——样本数; y_i^d ——第 i 个理想输出值; y_i ——第 i 个实际输出值;

通过优化使均方误差最小, 即通过 FS 算法搜索使网络误差平方和最小的网络权重, 即获得该函数的最小值 (最优值), 得到最优解即为对应的权重, 网络参数优化流程见图 1。

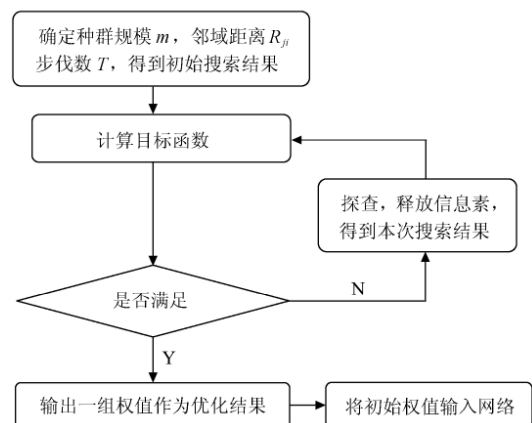


图 1 基于自由搜索算法的网络参数优化流程

Fig.1 Flow chart of parameter optimization base on free search

根据分析采用 3 层前馈型网络

$$Y = \log \text{sig}(w_2 \times \tan \text{sig}(w_1 \times p_n + B_1) + B_2) \quad (12)$$

式中: Y ——输出变量; $\log \text{sig}$ ——隐层到输出层的传递函数; w_1 ——输入层到隐含层的连接权重; w_2 ——隐含层到输出层的连接权重; p_n ——网络输入变量; B_1 ——输入层到隐含层的连接阈值; B_2 ——隐含层到输出层的阈值; $\tan \text{sig}$ ——输入层到隐含层的传递函数。按照 BP 网络的一般设计原则, 输入层与隐含层的传递函数为正切 S 型函数 $\tan \text{sig}$ 。隐含层与输出层的传递函数设定为对数 S 型函数 $\log \text{sig}$ 。

3 应用实例分析

3.1 研究区概况

试验地布设在江西水土保持生态科技园, 该园地处江西省北部鄱阳湖水系的德安县城郊燕沟小流域, 位于东经 $115^{\circ}42'38''\sim 115^{\circ}43'06''$, 北纬 $29^{\circ}16'37''\sim 29^{\circ}17'40''$ 之间。属亚热带季风气候区, 具有气候温和, 雨量充沛, 光照充足, 四季分明的特点。年平均降雨量 $1\,451.8\text{ mm}$, 因受季风气候影响而在季节分配上极不均匀, 形成明显的干季和湿季, 7—9月高温少雨。

3.2 材料与方法

研究建立 3 个处理小区 (表 1)。在修筑前选择一块地形、土壤等条件基本一致的坡地 (测量坡度为 14°), 将原土块取出, 分层堆放。待全部取完土后, 并作防渗处理, 然后将土块按原样回填至 1.1 m 土深。每个处理小区面积为 $5\text{ m}\times 15\text{ m}=75\text{ m}^2$, 坡度恢复至原地面的 14° , 自上至下总共设置 4 个出水口, 地表径流观测是在小区下部设置 1 个承水槽, 承接地表径流与泥沙, 将出水口用塑胶管连接并引入地表径流池, 同时地表径流地下径流池池配置自记水位计, 能全天候的记录径流和入渗动态过程。在每个处理的坡面上, 离上坡边缘的 3.5 、 7 及 10.5 m 处还埋设从美国引进的土壤水分传感器 (SM200), 埋设深度为 30 、 60 及 90 cm , 分别量计各层土壤剖面的土壤含水率, 以观测土壤中水分分布及变化规律, 试验区配有自记雨量计能够记录每场降雨过程, 试验观测时间为 2001—2005 年。

表 1 试验处理

Table 1 Experiment treatment

处理	措施	特征
A	百喜草覆盖	全园种植百喜草, 覆盖度 100%
B	百喜草敷盖	百喜草刈割后敷盖于地表, 敷盖度 100%, 厚度约 15 cm
C	裸露对照	地表完全裸露

3.3 结果分析

2001—2005 年 5 a 观测期内, 试验区总降雨量 $7\,259.1\text{ mm}$, 年平均降雨量为 $1\,451.8\text{ mm}$ 。2001—2005 年降雨量依次为 $1\,164.3$ 、 $1\,808.5$ 、 $1\,433.0$ 、 $1\,302.4$ 、 $1\,550.9\text{ mm}$ 。观测期内的月平均及年平均降雨量及入渗量见图 2 和图 3, 从图中可以看出入渗量为: 百喜草覆盖 > 百喜草敷盖 > 裸露对照, 反映了不同下垫面导致降雨在土壤中的再分配差异, 覆盖和敷盖对土壤入渗均起到促进作用, 入渗量年际分布特征明显, 与降雨量的分布规律基本一致, 主要集中在 4—9 月 (汛期), 占总入渗量的 $63.7\%\sim 98.8\%$ (图 2、3), 3 种处理下的土壤水分入渗量 4—9 月差异显著, 而 1—3 月、10—12 月差异不明显 (图 2), 而大雨和暴雨基本集中在 4—9 月, 说明在大雨量、高雨强的情况下, 覆盖与敷盖对土壤入渗的促进作用更为显著 (图 4)。

对于影响入渗的因素许多学者进行了研究, 大量的试验结果^[14-15]表明影响入渗的主要因素有土壤初始含水率、土壤体积质量、坡度、降雨量、降雨强度、降雨时间、通气孔度和下垫面状况等。对于本研究由于采用的

土壤水分渗漏装置是仿照试验自然坡地坡度设计 (施工后回填至原地面的 14°), 坡度是固定不变的, 因此坡度不作为影响入渗的因素; 3 个处理的下垫分别为百喜草覆盖、敷盖和裸露, 不同下垫面由于植被管理措施不同对降雨入渗的影响效果也不同; 雨型、降雨量、降雨强度和降雨历时均能影响土壤的入渗过程, 也是影响土壤侵蚀和洪峰流量的主要因素, 除此之外, 土壤体积质量和通气孔度对入渗都有较大的影响, 因此这里取降雨历时、最大降雨强度、降雨量、土壤初始含水率、土壤体积质量、通气孔度和下垫面状况 7 项指标作为网络输入, 土壤入渗量单项指标作为网络输出, 即网络输入层节点数为 $n=7$, 输出层节点数 $m=1$ 。选用 2001—2005 年有代表性的 90 场降雨入渗的观测数据, 90 场降雨包括了暴雨 (11 场)、大雨 (13 场)、中雨 (61 场) 和小雨 (5 场), 因此非常具有代表性, 选取前 80 场降雨为训练样本, 后 10 场降雨为检验样本, 预测样本为 240 个, 检验样本为 30 个。结果见图 5、6。

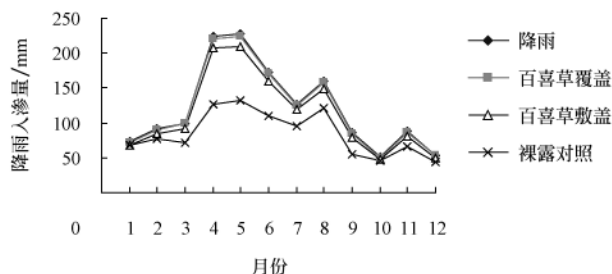


图 2 试验期间月平均降雨入渗量

Fig.2 Monthly mean infiltration and precipitation during experiment period

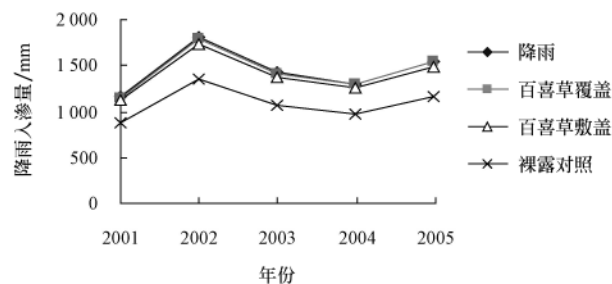


图 3 试验期间年平均降雨入渗量

Fig.3 Average annual infiltration and precipitation during experiment period

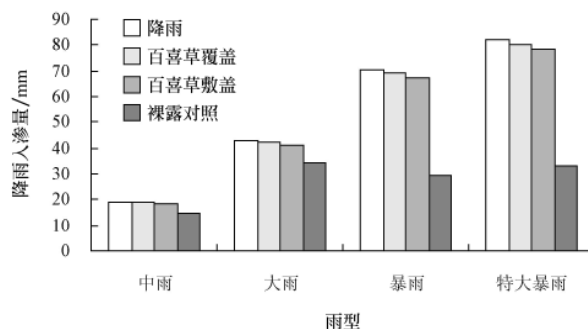


图 4 不同雨型下降雨入渗量特征

Fig.4 Characteristics of infiltration and precipitation under different types of rainfall

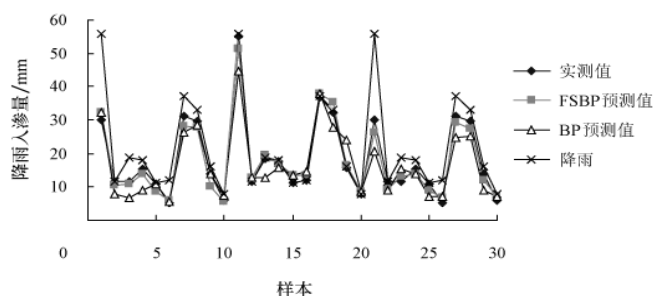


图5 入渗实测值与预测值比较

Fig.5 Comparison between observed and predicted data of infiltration

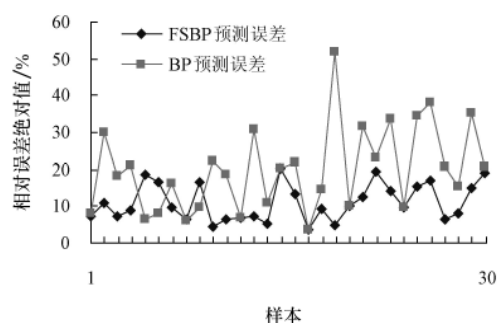


图6 BP模型与FSBP模型预测误差比较

Fig.6 Comparison of prediction errors between BP model and FSBP model

从图5可以看出实测值和FSBP预测值相差较小,平均误差只有11.08% (最小相对误差为3.62%),通过分析预测值和实测值相关关系可以看出预测值和实测值具有较好的相关性,决定系数为0.9715 (图7),经t检验其显著水平Sig.取值分别为0.168大于0.05, (表2),预测值和实测值没有显著差异,说明模型具有较高的精度。

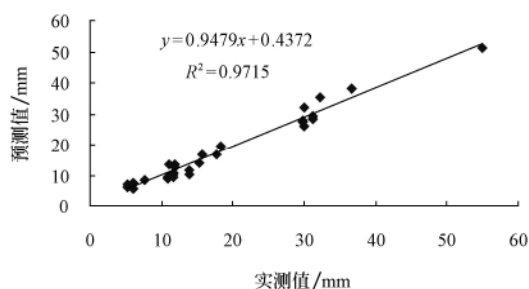


图7 入渗实测值与预测值相关关系

Fig.7 Correlation between observed and predicted data of infiltration

表2 预测值与实测值t检验结果

Table 2 Results of t examination between observed and predicted data of infiltration

项目	均值	标准差	标准 误差均值	检验 统计量	自由度	显著水平 (2-尾)
结果	0.51767	2.00368	0.36582	1.425	29	0.168

与传统BP模型相比预测平均相对误差低8.9%,从误差范围来看基于自由搜索算法的BP模型(FSBP)误差范围分别比传统BP模型低31.45% (图6),由此可以

看出,无论是从模型预测精度上还是稳定性上基于自由搜索算法的BP模型(FSBP)都要优于传统的BP模型。

虽然自然降雨条件下坡地入渗的过程机理及相关参数非常复杂,但是基于自由搜索算法的BP网络模型预测精度是比较理想的,与传统BP模型相比精度高,说明基于自由搜索算法的BP模型在(FSBP)能够对复杂的坡地入渗进行预测。

4 讨论与结论

本文应用基于自由搜索算法的BP网络对自然降雨条件下不同处理措施红壤坡地的入渗规律进行了预测,结果表明基于自由搜索算法的BP网络模型对入渗规律的预测效果较好,平均误差只有11.08%,决定系数达0.9715,与传统BP模型相比有较高的精度和稳定性,说明基于自由搜索算法的BP网络能对复杂的坡地入渗规律进行预测。基于自由搜索算法的BP网络相对物理模型操作简单,计算量小,在物理参数较少的情况下也可以取得较高的精度,具有一定的实用性。

针对自由搜索算法的研究刚刚开始,公开发表的文献不多,和其他算法相比还没有形成成熟的系统分析方法,许多问题仍待进一步研究。由于试验是在固定坡度下进行的,限于试验条件没有同时进行不同坡度入渗试验的观测,而且对于坡地入渗的其他影响因素(土壤温度、水稳性团粒等)均可作为模型的输入因子,限于试验条件没有全部进行观测,对模型的通用性产生了一定的影响,因此对于该模型方法仍需做进一步的深入研究。

[参 考 文 献]

- [1] 傅斌,王玉宽,朱波,等. 紫色土坡耕地降雨入渗试验研究[J]. 农业工程学报, 2008, 24(7): 39—43.
Fu Bin, Wang Yukuan, Zhu Bo, et al. Experimental study on rainfall infiltration in sloping farmland of purple soil[J]. Transactions of the CSAE, 2008, 24(7): 39—43. (in Chinese with English abstract)
- [2] 蒋定生,黄国俊. 黄土高原土壤入渗速率的研究[J]. 土壤学报, 1986, 23(4): 299—304.
Jiang Dingsheng, Huang Guojun. Study on the filtration rate of soils on the loess plateau of china[J]. Acta Pedologica Sinica, 1986, 23(4): 299—304. (in Chinese with English abstract)
- [3] 雷志栋,杨诗秀,谢森传. 土壤水动力学[M]. 北京: 清华大学出版社, 1988: 223—257.
- [4] 芮孝芳. 水文学原理[M]. 北京: 中国水利水电出版社, 2004: 76—100.
- [5] 赵西宁,王万忠,吴普特,等. 坡面入渗的人工神经网络模型研究[J]. 农业工程学报, 2004, 20(3): 48—50.
Zhao Xining, Wang Wanzhong, Wu Pute, et al. Artificial neural network model for soil infiltration in slope farmland [J]. Transactions of the CSAE, 2004, 20(3): 48—50. (in Chinese with English abstract)
- [6] Kalin P, Guy L. Free search-a comparative analysis[J]. Information Science, 2005, 172(1/2): 173—193.

- [7] 高波, 田东, 胡亮. free search—一种新的进化算法[J]. 仪器仪表学报, 2006, 27(6): 819—821.
Gao Bo, Tian Dong, Hu Liang. Free search—a novel evolutionary algorithm[J]. Chinese Journal of Scientific Instrument, 2006, 27(6): 819—821. (in Chinese with English abstract)
- [8] 王斌, 张展羽, 张国华, 等. 一种新的优化灌溉制度算法—自由搜索[J]. 水科学进展, 2008, 19(5): 738—741.
Wang Bin, Zhang Zhanyu, Zhang Guohua. Free search: A novel algorithm for optimizing irrigation schedule[J]. Advance In Water Science, 2008, 19(5): 738—741. (in Chinese with English abstract)
- [9] 张国华, 张展羽, 邵光成, 等. 基于粒子群优化算法的灌溉渠道配水优化模型研究[J]. 水利学报, 2006, 37(8): 1004—1008.
Zhang Guohua, Zhang Zhanyu, Shao Guangcheng, et al. Optimization model for discharge distribution of irrigation channels based on particle swarm optimizer[J]. Journal of Hydraulic Engineering, 2006, 37(8): 1004—1008. (in Chinese with English abstract)
- [10] Kennedy J, Eberhart R C. Particle swarm optimization[C]// IEEE International Conference on neural networks. Perth australia, 1995: 1942—1948.
- [11] 付强. 数据处理方法及其农业应用[M]. 北京: 科学出版社, 2006: 283—323.
- [12] 刘波, 王凌, 金以慧. 差分进化算法研究进展[J]. 控制与决策, 2007, 22(7): 721—729. 283—323.
Liu Bo, Wang Ling, Jin Yihui. Advances in differential evolution[J]. Control and Decisio, 2007, 22(7): 721—729. 283—323. (in Chinese with English abstract)
- [13] Yi D, Ge X R. An improved PSO based ANN with simulated annealing technique[J]. Neurocomputing, 2005, 63(11): 527—533.
- [14] 吴发启, 赵西宁, 余雕. 耕坡地土壤水分入渗影响因素分析[J]. 水土保持通报, 2003, 23(1): 17—18.
Wu Faqi, Zhao Xining, Yu Diao. Analysis on affecting factors of soil infiltration in slope farmland[J]. Bulletin of Soil and Water Conservation, 2003, 23(1): 17—18. (in Chinese with English abstract)
- [15] 章俊霞, 李小军, 左长清. 南方红壤入渗影响因素研究[J]. 中国水土保持, 2008, (6): 27—28.
Zhang Junxia, Li Xiaojun, Zuo Changqing. Affecting factors of red loam infiltration of the south[J]. Soil and Water Conservation in China, 2008, (6): 27—28. (in Chinese with English abstract)

Prediction of slope infiltration based on artificial neural networks by free search

Li Xinhui¹, Zhang Zhanyu¹, Yang Jie², Zhang Guohua³, Wang Bin⁴, Wang Chao¹

(1. College of Water Conservancy and Hydropower Engineering, Hehai University, Nanjing 210098, China;

2. Soil and Water Conservation Research Institute of Jiangxi Province, Nanchang 330000, China;

3. China Irrigation and Drainage Development Center, Beijing 100054, China;

4. College of Water Conservancy and Building Engineering, Northeast Agricultural University, Harbin 150030, China)

Abstract: Attempts of using FSBP were made to predict infiltration of natural rainfall on the slop surface of red soil under different land use patterns. Seven indexes such as precipitation, maximum rainfall intensity, rainfall duration, initial soil water content, soil bulk density, soil porosity and underlaying surface were selected as input variable, and the amount of infiltration as output variable. Results show that the mean relative error of the prediction is 11.08%, and t test and regression analysis indicates that the predicted value differs just slightly from the observed value and their correlation coefficient was 0.9715. The model is quite high in accuracy and stability, and serves as useful tool in further research on prediction of infiltration of nature rainfall on slopes.

Key words: rain, infiltration, backpropagation, free search, slope, artificial neural networks