

基于 PSO-BP 神经网络的参考作物蒸腾量预测

张志政, 周 威

(西安建筑科技大学环境与市政工程学院, 西安 710055)

摘要:针对以往 BP 神经网络收敛速度慢及易陷于局部极小值等问题, 引入粒子群算法优化 BP 网络的权值和阈值, 建立 PSO-BP 神经网络, 预测参考作物蒸腾量 ET_0 。以西安地区的相关资料为基础, 设计 9 种影响因子组合方案, 利用 PSO-BP 网络模型进行 ET_0 的预测, 结果表明, 该模型运算速度快, 预测精度较高; 对比分析 9 种方案的预测结果发现, 方案 7 为最优, 该方案只需选用平均温度、平均相对湿度、风速和日照时数四项影响因子, 即可获得较高精度的参考作物蒸腾量预测值。

关键词: 粒子群算法; PSO-BP 神经网络; 参考作物蒸腾量

中图分类号: TP273⁺.4 **文献标识码:** A

Reference Crop Evapotranspiration Prediction Based on PSO-BP Neural Network

ZHANG Zhi-zheng, ZHOU Wei

(School of Environment and Municipal Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China)

Abstract: Aiming at the shortages of the past BP neural network, including low convergence speed and easily falling into the local minimum, the particle swarm algorithm is adopted to optimize the weights and the thresholds of the BP neural network. Then PSO-BP neural network is set up to predict reference crop evapotranspiration, ET_0 . Based on the data of Xi'an city, this paper designs 9 kinds of impact factor combination schemes to predict ET_0 through the PSO-BP neural network. The results show the model has high operational speed and prediction accuracy. Comparatively analyzing the prediction result of the 9 kinds of impact factors, it is found that the 7th scheme is the best. It needs only 4 impact factors, including the average temperature, the average relative humidity, the wind speed and the sunlight hours, and it can obtain the prediction value of ET_0 with high degree precision.

Key words: particle swarm optimization; PSO-BP neural network; reference crop evapotranspiration

如何确定作物精量化需水量是现代化节水灌溉理论中的一大难点, 而作物蒸腾量的预测则是确定精量化需水量的关键。确定作物蒸腾量的方法一般有直接计算法和间接计算法。直接计算法是采用经验公式计算作物蒸腾量, 这种方法使用范围有限^[1]。目前, 在多种参考蒸腾量 ET_0 算法中, 联合国粮农组织 (FAO) 推荐的 Penman-Monteith 公式是公认的适用性较强的标准计算方法^[2], 但这种方法需要的数据量大, 且对数据质量有较严格的要求, 应用受到很大限制。间接计算法采用数学模型预测参考作物蒸腾量, 然后引入当地作物系数计算得到作物蒸腾量, 其关键在于如何准确的预测参考作物蒸腾量

ET_0 。随着 BP 神经网络理论的迅速发展, 一些学者将其应用于 ET_0 的预测, 并取得了较好的预测结果^[3,4], 但其隐含层节点数的确定具有很大主观性, 且 BP 学习算法收敛速度慢、易陷于局部极小值等问题难以克服, 本文针对相关问题进行网络优化研究。

1 BP 神经网络原理

BP 神经网络 (Back-Propagation Neural Networks) 即反向传播神经网络。是目前神经网络模型中发展最成熟、应用最为广泛的一种。BP 网络是一种具有三层或三层以上神经元的神

收稿日期: 2014-03-21

作者简介: 张志政 (1964-), 男, 副教授, 研究生导师, 从事水环境与水资源利用方面研究。E-mail: zhangzhizheng@xauat.edu.cn。

神经网络,包括输入层、中间层(隐含层)和输出层。上下层之间实现全连接,而每层神经元之间无连接^[5]。当一对学习样本提供给网络后,神经元的激活值从输入层经各中间层向输出层传播,在输出层的各神经元获得网络的输入响应。接下来,按照减少目标输出与实际输出之间误差的方向,从输出层反向经过各中间层回到输入层,从而逐层修正各连接权值,这种算法称为“误差反向传播算法”^[6]。随着这种误差逆向的传播修正不断进行,网络对输入模式响应的正确率也不断上升。直到误差达到预设最小值时网络训练结束。

2 BP 确定网络结构

(1)网络层数。理论上已经证明,对于在任何闭区间内的一个连续函数,都可以用一个单隐含层的 BP 网络来逼近,因而一个三层 BP 网络就可以完成任意的 n 维到 m 维的映射^[7]。本文采用一个单隐含层的三层 BP 网络。

(2)各层节点数。影响作物参考蒸腾量 ET_0 的因素有很多,一般认为最高温度、平均温度、最低温度、平均相对湿度、日序数、日照时数和风速等七项是其主要因素。我们可根据需要对上述七项影响因子进行组合,作为 BP 网络的输入项,进而确定网络的输入层节点数。本文所建立的神经网络预测输出层变量为作物参考蒸腾量 ET_0 ,输出层的节点数为 1。

隐含层节点数的确定一直是确定网络结构的难点和关键,对此目前尚无严格的理论依据。一个合理的方法是先通过经验公式确定隐含层节点数的一个适宜的范围,然后在此范围内逐一实验,对比其网络误差大小,选择误差最小值对应的隐含层节点数。隐含层节点数与输入层节点数及输出层节点数有如下关系,输入层、隐含层和输出层的节点数分别为 n_i 、 n_h 和 n_o ,取:

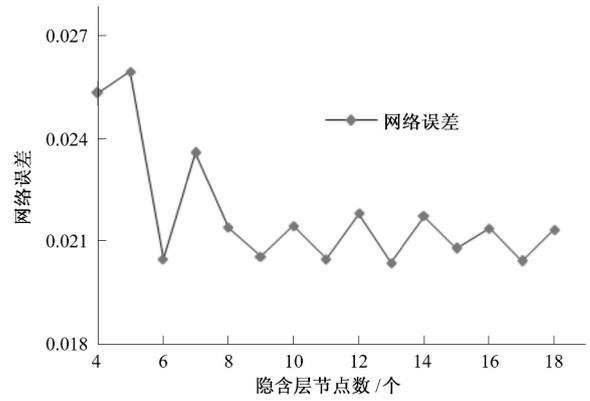
$$a = \frac{n_i + n_o}{2} \leq n_h \leq (n_i + n_o) + 10 = b \quad (1)$$

则隐含层节点数应在区间 $[a, b]$ 之内取得。据此,可得出选取不同输入输出节点数目时对应的隐含层节点数的适宜范围,如表 1 所示。当不同输入量个数或相同输入量个数但组合方式不同的时候,在上述区间内依次实验,取网络误差值达到最小值时对应的隐含层节点数。当以最高温度、平均温度、最低温度、平均相对湿度、日序数、风速和日照时数等七项作为输入量,参考作物蒸腾量为输出量构建 BP 网络时,隐含层节点数范围为 $[4, 18]$,取该区间内每一个整数值依次进行实验,实验误差结果如图 1 所示。

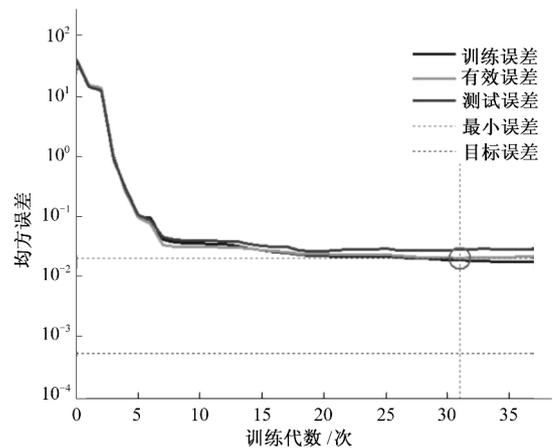
表 1 不同输入输出节点数对应隐含层节点区间

输入层节点数	输出层节点数	隐含层节点数区间
7	1	[4,18]
6	1	[3,17]
5	1	[3,16]
4	1	[2,15]
3	1	[2,14]

由图 1 知,当隐含层节点数取 13 时, BP 网络取得最小误差为 0.020 368,故隐含层节点数取 13。最优网络结构为 7-13-1 的三层 BP 神经网络预测模型。



(a)网络误差



(b)训练至 31 代取得最小有效误差 0.020 368

图 1 网络误差曲线图

3 粒子群优化 BP 神经网络

(1)粒子群算法原理。粒子群优化算法(Particle Swarm Optimization)的核心思想是模拟鸟类群体飞行过程中的觅食行为^[8]。每个粒子对应于鸟群中的一只鸟,寻找最优解的过程,对应于鸟群寻找食物的过程,在此过程中往往是有某一只鸟先发现食物,带动周围的鸟聚集至食物周围,最终整个群体都得到食物。算法中,每个粒子的速度决定了它的运动方向以及运动距离,通过目标函数求得个体的适应度值,让粒子追寻周围拥有最好位置的个体和整个群体的最好位置来不断更新自己的位置,以使自己达到最佳位置^[9]。

(2)粒子群算法优化 BP 神经网络。近年来, BP 神经网络在相关预测研究中被广泛应用。但是在 BP 神经网络算法中,梯度下降法的训练效果过于依赖初始权值的选择,且 BP 算法有着训练时间长、易陷入局部极小等问题,而 PSO 算法具有收敛速度快、记忆性强和全局搜索能力较强等特点。

BP 神经网络与粒子群算法都是通过模拟生物的特性而成,具有一定的共性,从而可以将两种算法结合起来,取长补短,利用粒子群算法优化神经网络的权值,可以弥补 BP 网络在学习能力和收敛速度上的不足,并且可以充分发挥神经网络的非线性映射能力^[10]。

本文使用 PSO 优化 BP 网络的权值与阈值。选用训练误差绝对值之和作为 PSO 算法的适应度函数。优化过程如下:

①初始化。根据 BP 神经网络的输入样本,建立 BP 神经网络的拓扑结构,输入层节点数 n_i , 隐层节点数 n_h , 输出层节点数 n_o ; 初始化粒子的位置和速度, 以及粒子数、最大迭代次数、惯性权重、学习因子等参数。②通过网络训练, 计算粒子的适应度值, 得到粒子的个体最优值与全局最优值。将粒子适应度值与个体最优值和全局最优值相比较, 记录当前粒子所经历的最好位置。③考察每一个粒子的适应度值。若该值优于个体最优, 则将当前值置为个体最优, 并更新该粒子的个体最优; 若粒子中的个体最优优于当前的全局最优, 则将个体最优置为全局最优, 并更新全局最优值。④将经过 PSO 优化的权值和阈值作为 BP 神经网络的初始权值和阈值代入 BP 网络, 训练至满足网络的性能指标均方误差小于预先设定的误差要求或达到最大迭代次数时, 停止迭代, 输出结果, 否则转到 ②, 继续迭代直至算法收敛。⑤由训练样本和测试样本完成神经网络的训练和测试, 输出预测值。

4 建立 PSO-BP 网络模型预测参考作物蒸腾量

4.1 原始数据处理

本文采用西安地区 2001—2004 年的气象资料(包括最高温度 T_{max} 、平均温度 T 、最低温度 T_{min} 、平均相对湿度 RH_{mean} 、日序数 J 、日照时数 n 和风速 u) 和 Penman-Monteith 方程计算所得的 ET_0 作为训练数据, 用 2005 年的气象资料和 Penman-Monteith 方程计算的 ET_0 作为检验数据。训练数据占总样本数的 80%, 符合神经网络预测模型的数据划分原则。

输入向量物理意义和单位的不统一, 会影响网络的训练过程。因此, 将输入向量和期望输出向量值按公式(2)进行归一化处理。

$$X'_i = \frac{X_i - 0.5(X_{max} + X_{min})}{0.5(X_{max} - X_{min})} \quad (2)$$

反归一化公式见式(3):

$$X_i = 0.5(X_{max} - X_{min}) X'_i + 0.5(X_{max} + X_{min}) \quad (3)$$

式中: X_{max} 、 X_{min} 为原始样本数据中的最大值和最小值; X'_i 、 X_i 为原始样本数据和对应处理后的数据。

4.2 模型运算流程

本文所建立 PSO-BP 预测模型采用 MATLAB 软件编程实现, 实验运行环境为: Windows 7 操作系统、2 G 内存、Intel I5 2.7 GHz 处理器。建立 PSO-BP 预测模型包括 BP 神经网络的程序设计及利用粒子群算法优化 BP 网络的权值和阈值。首先建立 BP 神经网络并对原始样本数据进行初始化, 初始化的程序传给主程序, 主程序在编码解码的同时随机生成初始权值阈值, 再调用粒子群算法优化权值阈值, 然后将优化后的权值阈值赋给建立的 BP 神经网络进行网络训练, 最后利用训练好的 PSO-BP 网络模型预测参考作物蒸腾量。PSO-BP 神经网络预测模型流程图见图 2。

4.3 模型评价指标

本文选用以下指标作为评价模型的标准。

(1) 标准偏差。

$$\delta = \sqrt{\frac{\sum_{i=1}^n (Y_{PM,i} - Y_{BP,i})^2}{n-1}} \quad (4)$$

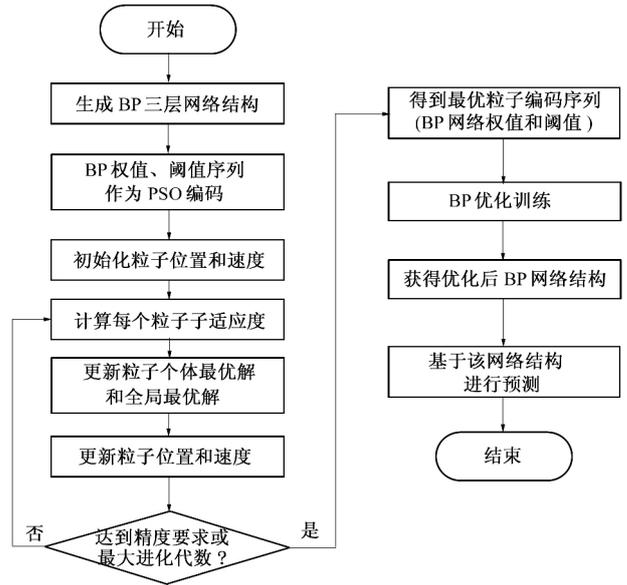


图 2 PSO-BP 神经网络预测模型流程图

(2) 相对误差绝对值的平均值。

$$\bar{\epsilon} = \frac{\sum_{i=1}^n |Y_{PM,i} - Y_{BP,i}|}{Y_{BP,i} \cdot (n-1)} \quad (5)$$

(3) 预测合格率。预测合格率设为相对误差绝对值 $\epsilon < 15\%$ 的预测样本占总样本数的百分数。

(4) a, R^2 。 a 为预测值对实际值的线性回归方程的系数, R^2 为相关系数的平方。两者越接近于 1, 说明预测值与实际值吻合得越好。

(5) 为了验证 PSO-BP 模型的稳定性, 我们通过 10 次独立重复实验, 记录每次实验结果的预测相对误差、预测标准偏差、预测合格率, 绘制其统计盒状图。

4.4 预测结果分析

本文采用西安地区 2001—2004 年的数据为训练数据, 2005 年的数据作为预测检验数据。选择若干典型气象因子组合作为输入量, 用 PSO-BP 网络预测 ET_0 , 研究不同的影响因子对 ET_0 预测值的影响, 对各种评价指标做分析对比, 结果如表 2 所示。

(1) 对比方案 1、2、3 的预测结果可知, 各项评价指标数值上较为接近, 说明最高温度、平均温度和最低温度三项具有极强的相关性, 只需选择其中的平均温度, 与其他五项影响因子组合作为输入量即可取得精度较高的预测结果。

(2) 对比方案 4、5、6、7、8 的预测结果可知, 将输入因子减少至四项时, 方案 7 的预测结果明显优于方案 4、5、6、8。说明平均温度、平均相对湿度、日序数、风速、日照时数五项影响因子中日序数对预测结果的影响较小, 其余四项均为影响预测的结果的重要因素。

(3) 对比方案 9 与方案 4、5、6、7、8 的预测结果, 进一步证实, 平均温度、平均相对湿度、风速和日照时数四项影响因子是预测作物参考蒸腾量的必不可少的影响因子。

综上所述, 方案 1、2、3 和 7 的效果较好, 从尽量减少输入因子、选择合理网络结构以减少网络运算量的角度考虑, 本文

表 2 不同输入组合的预测结果评价指标对比

方案	输入因子	网络结构	PSO-BP 网络误差	δ	$\bar{\epsilon}/\%$	$\epsilon < 15\% / \%$	a	R^2
1	$T_{max}, T, T_{min}, RH_{mean}, J, u, n$	7-13-1	0.002 3	0.076	5.70	95.34	0.987	0.981
2	$T_{max}, T, RH_{mean}, J, u, n$	6-11-1	0.002 2	0.076	5.72	94.52	0.994	0.980
3	T, RH_{mean}, J, u, n	5-9-1	0.002 2	0.073	5.65	96.71	1.003	0.981
4	T, J, u, n	4-7-1	0.004 4	0.124	9.21	81.92	0.991	0.953
5	RH_{mean}, J, u, n	4-6-1	0.003 7	0.123	9.09	81.37	1.040	0.953
6	T, RH_{mean}, J, u	4-7-1	0.009 7	0.145	11.66	71.23	1.059	0.934
7	T, RH_{mean}, u, n	4-6-1	0.002 5	0.079	6.02	94.52	1.006	0.977
8	T, RH_{mean}, J, n	4-6-1	0.015 2	0.138	10.94	75.07	1.138	0.906
9	T, H_{mean}, n	3-5-1	0.014 6	0.129	10.45	74.52	1.186	0.918

选择方案 7 作为预测模型,只需选用平均温度、平均相对湿度、风速和日照时数四项输入因子,就可快速获得较高精度的预测参考蒸腾量(标准偏差为 0.079,误差绝对值的均值为 6.02%,预报合格率 $\epsilon < 15\%$ 为 94.52%)。图 3 为方案 7 模型的预测合格率盒状图及平均相对误差与标准偏差盒状图;图 4 和图 5 分别为方案 7 预测值 ET_0 与 PM 公式计算值的对比图及线性拟合图。

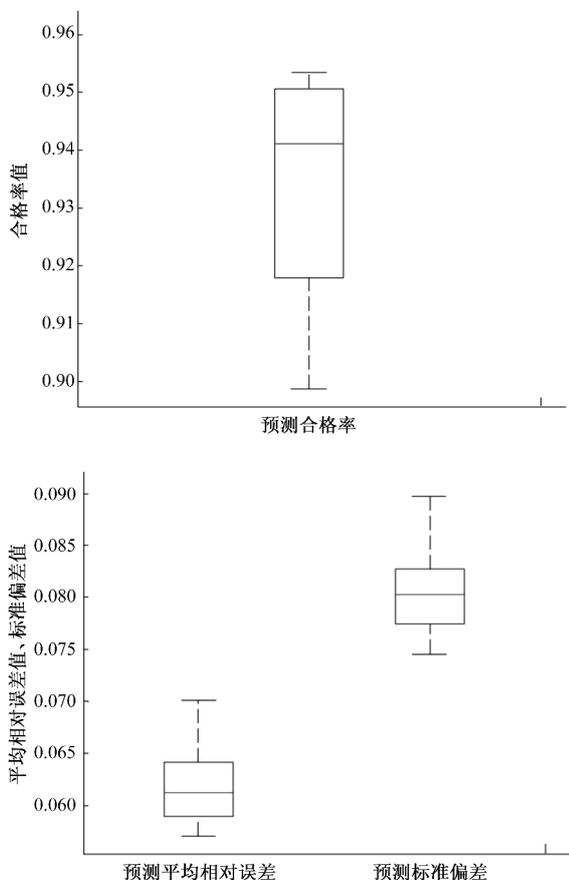


图 3 预测合格率盒状图及平均相对误差与标准偏差盒状图

5 结 语

本文首次将粒子群优化 BP 神经网络(PSO-BP)模型应用于参考作物蒸腾量的预测中,克服了 BP 神经网络训练时间长、

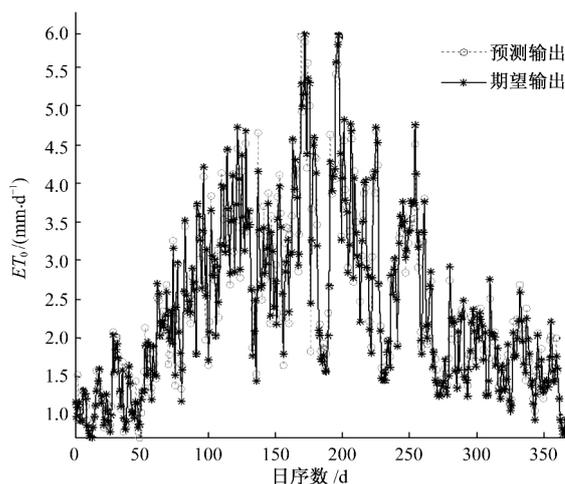


图 4 PSO-BP ET_0 与 PM ET_0 的对比

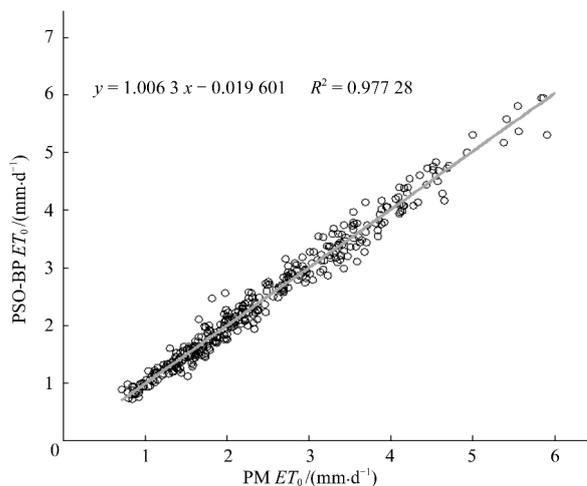


图 5 线性拟合

易陷入局部极小等缺陷,较之以往的遗传算法优化 BP 神经网络,也避免了过早收敛,需要的控制参数过多等问题。结果表明,该模型不仅预测精度高,而且需要的输入量较少,对气象资料缺失情况下准确预测参考作物蒸腾量提供了一个可行的方法,对快速准确地预测作物需水量及节水灌溉的广泛实施具有一定的指导意义。(下转第 95 页)

mm。地势平坦,土壤类型以渗育水稻土和潜育水稻土为主,自然肥力较高,水系纵横交错,水资源较丰富,水利条件较好。地面灌溉是主要的灌溉方式,但由于经济的发展,工业废水、生活污水、农药、化肥等污染加剧,水质型缺水日趋严重。

主要节水农业发展模式:水资源保护+地面灌溉改进技术。

模式内容:①保护水资源,加大水环境综合治理力度,建立污染源与河网水质监测体系,加强灌溉水源水质监测。在控制工业点源污染排放的基础上,控制灌溉水质量。②加快地面灌溉改进技术的应用,如喷灌、滴灌、微灌节水技术,有条件的地区可采用自动化控制灌溉技术,提高灌溉用水管理水平。③结合高标准农田建设,扩大灌溉暗渠,明排衬砌,水稻采用浅湿、湿润灌溉等技术。④调整农业种植结构,减少对土地资源要求较高、低经济价值的粮食作物种植比例。⑤合理施用化肥和农药,降低化学氮肥和化学农药的使用量,防止农业面源污染,改善水体环境。

5 结 语

本文从江苏省气候特征、地形地貌、土壤类型、灌排方式、农业生产和农村经济因素等几个方面提出 11 个节水农业相关指标,构建了江苏省节水农业分区的评价指标体系。并通过因子分析方法遴选节水农业关键指标,结合层次聚类分析方法,对江苏省 13 市 65 县进行了聚类分析,将该区域划分为徐淮平原水资源紧缺区、里下河平原季节性缺水、沿海平原盐碱化区、沿江平原高沙土区、宁镇扬及宜溧丘陵工程性缺水、太湖平原水质性缺水 6 类区域,并根据各区域特点提出了各区域节水农业发展中存在的主要问题和未来发展方向。

参考文献:

- [1] 王 燕. 安徽省节水农业分区初步研究[J]. 中国农村水利水电, 2001, (7): 13—15.
- [2] Trung T C, Gupta A D, Babel M S, et al. Assessment of different irrigation management models in Vietnam[J]. International Journal of Water Resources Development, 2005, 21(3): 525—535.

(上接第 90 页)

参考文献:

- [1] 魏新光, 王密侠, 张 倩. 基于主成分分析的灌区参考作物蒸发蒸腾量预测模型研究[J]. 节水灌溉, 2011, (1): 29—36.
- [2] Allen R G, Pereira L S, Raes D, et al. Crop evapotranspiration guidelines for computing crop water requirements[M]. FAO Irrigation and Drainage Paper 56, 1998: 300.
- [3] 刘丙军, 邵东国, 沈新平. 基于小波变换的 BP 神经网络参考作物腾发量预测模型[J]. 武汉大学学报, 2007, 40(1): 69—73.
- [4] 李帅莹, 迟道才, 刘 丽, 等. 基于 Elman 和 BP 神经网络的逐月参考作物腾发量预测[J]. 中国农村水利水电, 2009, (12): 11—14.

- [3] Goncalves J M, Pereira L S, Fang S X, et al. Modelling and multicriteria analysis of water saving scenarios for an irrigation district in the upper Yellow River Basin[J]. Agricultural Water Management, 2007, 94(1/3): 93—108.
- [4] 尹 剑, 王会肖, 王艳阳, 等. 关中地区农业节水分区研究[J]. 中国生态农业学报, 2012, 20(9): 1 173—1 179.
- [5] 胡长忠. 江苏省水资源供需平衡现状分析[J]. 水利水电科技进展, 2000, 20(4): 35—49.
- [6] 吴玉柏, 常本春, 王亦斌. 江苏省节水农业问题研究[J]. 水资源保护, 2002, (2): 53—56.
- [7] 刘力兵, 邵孝侯, 徐 茂, 等. 江苏省节水农业技术体系研究[J]. 灌溉排水学报, 2007, (2): 48—51.
- [8] 俞双恩, 张展羽. 江苏省水稻高产节水灌溉技术体系研究[J]. 河海大学学报(自然科学版), 2002, (6): 30—34.
- [9] 江苏省水利厅. 江苏省“十一五”水利发展思路[C]//水利部发展研究中心. 2005 年中国水利发展报告, 2005.
- [10] 江苏省水利厅. 江苏省水利发展基本情况[DB/OL]. http://www.jswater.gov.cn/art/2009/12/8/art_55_27962.html.
- [11] 程 焯, 王 静. 土地用途分区管制研究[M]. 北京: 地质出版社, 2003.
- [12] 李英能, 黄修桥. 水土资源评价与节水灌溉规划[M]. 北京: 水利水电出版社, 1998.
- [13] 胡永宏, 贺思辉. 综合评价方法[M]. 北京: 经济科学出版社, 2000.
- [14] Sudheer R Satti, Jennifer M Jacobs, Suat Irmak. Agricultural water management in a humid region: sensitivity to climate, soil and crop parameters [J]. Agricultural Water Management, 2004, 10(70): 51—65.
- [15] 马智民, 乔 亮. 聚类分析在土地利用功能分区中的应用——以西安市雁塔区为例[J]. 国土资源科技与管理, 2007, 24(6): 90—93.
- [16] 赵荣钦, 黄贤金, 钟太洋, 等. 聚类分析在江苏沿海地区土地利用分区中的应用[J]. 农业工程学报, 2010, 26(6): 310—314.
- [17] 张爱军, 彭世彰, 郭相平. 江苏省徐淮地区农业节水问题及对策研究[J]. 中国农村水利水电, 2004, (1): 29—31.

- [5] 葛彩莲, 蔡焕杰, 王 健, 等. 基于 BP 神经网络的降雨量预测研究[J]. 节水灌溉, 2010, (1): 7—10.
- [6] 葛哲学, 孙志强. 神经网络理论与 MATLAB R2007 实现[M]. 北京: 电子工业出版社, 2007: 59—61.
- [7] 刘婧然, 马英杰, 王 喆, 等. 基于 RBF 神经网络与 BP 神经网络的核桃作物需水量预测[J]. 节水灌溉, 2013, (3): 16—19.
- [8] 魏秀业. 基于粒子群优化齿轮箱智能故障诊断研究[D]. 太原: 中北大学, 2009.
- [9] 刘贵明, 李晓英. 水库优化调度的粒子群算法[J]. 中国农村水利水电, 2013, (6): 156—158.
- [10] 亢丽君. 粒子群优化 BP 神经网络在应急物资需求预测中的应用研究[D]. 兰州: 兰州交通大学, 2013.