

基于 BP 神经网络的东港灌区作物需水量预测研究

张 丽^{1,2}, 吴金亮², 杨国范²

(1. 吉林师范大学 旅游与地理科学学院, 吉林 四平 136000; 2 沈阳农业大学 水利学院, 沈阳 110866)

摘 要:为了指导东港地区的农业生产,对影响作物参考蒸散量(ET_0)的气象因子进行了分析,利用东港市 1999 年和 2000 年水稻生育期的常规气象资料,采用 Matlab 软件的神经网络工具箱,建立 ET_0 三层 BP 神经网络预测模型,网络的输入因子为日净辐射量、日平均相对湿度和日平均风速,输出因子为利用 Penman—Monteith 公式计算得到的同期 ET_0 值。结果表明:(1) 当模型的隐层节点数为 11,传递函数采用 tansig 函数,训练函数采用 trainlm 函数时,预测值与目标值的平均相对误差为 9%,预测效果较其他情况都好,由此确定模型的最终结构为 3—11—1 的 BP— ET_0 模型。(2) 用 BP— ET_0 模型预测参考作物蒸散量,预测值与目标值的变化趋势基本一致。

关键词:参考作物蒸散量; Matlab; BP 神经网络

中图分类号:S274.4

文献标识码:A

文章编号:1005-3409(2012)06-0207-04

Prediction Research for Crop Water Requirement Based on BP Neural Network in Donggang Irrigation District

ZHANG Li^{1,2}, WU Jin-liang², YANG Guo-fan²

(1. College of Tourism and Geographical Science, Jilin Normal University, Siping, Jilin 136000, China;

2. College of Hydraulic Engineering, Shenyang Agricultural University, Shenyang 110866, China)

Abstract: In order to guide agricultural production in Donggang region, the impact of different meteorological factors on reference crop evapotranspiration (ET_0) were analyzed. Using the neural network toolbox of Matlab software, a three-layer artificial neural network model was established to predict ET_0 according to the conventional meteorological data during maize growth and development period of year 1999 and 2000 in Donggang city. In this network, the input variables were daily net radiation, daily average relative humidity and daily average wind speed, and the output variable was ET_0 which was calculated by the Penman-Monteith formula. The results showed that: (1) the average relative error between prediction ET_0 values and target ET_0 values was 9%, when 11 neural nodes in the hidden layer was adopted, tansig function was chosen as transfer function and trainlm function was used to train the network, therefore, a 3—11—1 optimal network was determined; (2) the change tendency of the predicted values with BP— ET_0 model prediction and the target was the basic consistency.

Key words: reference crop evapotranspiration; Matlab; BP neural networks

参考作物蒸散量 (Reference Evapotranspiration, ET_0) 是一种假想的参考作物冠层的蒸腾速率,它非常类似于表面开阔、高度一致、生长旺盛、完全遮盖地面而不缺水的绿色草地的蒸腾量^[1-2]。 ET_0 的研究与估算,一直是灌溉和排水领域的重要研究课题之一,它的准确计算直接影响作物需水预报的精度,进而影响到区域水资源的优化配置。作物需水量的研究是节水农业的重要内容,对水资源规划和陆地水文

学的研究具有重要意义^[3]。传统方法计算 ET_0 时需要对风速等气象资料进行修正或需要一些不易测得的气象资料,这给一些条件相对较差地区的农业用水计算预报和调度带来很大困难,在一些气象资料缺测时具有明显的局限和不足^[4]。随着我国水资源的日趋紧张,有必要对参考作物蒸散量的计算预报模型进行简化,使其具有一定精度,方便农业生产的需要。正确估算作物需水量不仅仅是节约用水、计划用水、

工程建设与规划等需要解决的重要问题,也是农业经济用水、作物生长模拟、水资源平衡分析与评价研究中的一个重要因素^[5]。

20 世纪 80 年代迅速发展起来的人工神经网络(Artificial Neural Networks),是模拟生物神经网络进行信息处理的一种数学模型。它以对大脑的生理研究成果为基础,模拟大脑的某些机理和机制,实现一些特定的功能^[6]。由于其独特的分布并行处理、非线性映射、自适应学习能力等已经在诸多领域得到广泛应用,并取得了较好的成果^[7]。目前常用的神经网络模型包括前馈神经网络、反馈神经网络和随机神经网络等,其中应用较多的模型包括多层反向传播网络(BP 网络)、自组织特征映像网络(SOFM 网络)、径向基函数网络(RBFN 网络)和自适应共振网络(ARTNN 网络)等^[4,8]。其中 BP 神经网络是较为成熟也是应用最多的一种非线性函数逼近方法^[9],在预测预报和分类评价中较为适用^[10]。而 ET_0 与各影响因素之间可以看作是一种复杂的非线性关系。在对 ET_0 的原有成果基础上加以改进并将 BP 神经网络应用于 ET_0 的预测,具有较广泛的应用价值。

本研究利用 BP 神经网络强大的非线性函数逼近功能,建立基于 BP 神经网络方法的参考作物蒸散量的计算模型(简称 BP— ET_0 模型),用以指导当地的农业用水宏观调控。

1 灌区数据收集与预处理

东港灌区位于辽宁省东港市管辖范围,地处鸭绿江和大洋河下游,东经 $123^{\circ}38'30''$ — $124^{\circ}21'25''$,北纬 $39^{\circ}47'24''$ — $40^{\circ}9'49''$,东西长 65 km,南北宽约 40 km,总控制面积 1 352 km²。灌区由铁甲分区和友谊分区组成,包括东港市的 21 个乡镇、农场,202 个行政村,总耕地面积 6.996 万 hm²。灌区设计灌溉面积 4.992 万 hm²,实际灌溉水田面积 3.707 万 hm²,是辽宁省大型灌区之一,也是丹东市的主要水稻产区。

对数据的规范化处理(数据单位转化及归一化处理)。其中,归一化处理采用的公式如下:

$$x_{\text{norm}} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}$$

式中: x_{norm} ——样本归一化的值; x_i ——样本初始值; x_{\max} ——样本最大值; x_{\min} ——样本最小值。

2 BP 神经网络建模注意事项

利用 Matlab 软件提供的工具箱编制 BP 网络程序解决非线性问题是一种便捷、有效的途径,在使用时要注意以下几个关键环节:

2.1 样本组织及输入、输出因子的选择

训练样本必须满足两个条件:一是样本要足够多,能反映事物的复杂程度;二是样本要具有代表性,尽可能包含能代表各类特征的模式,不能没有其中的任何一种^[11]。网络的输入因子的确定目前无定量指标,一般可依据经验和需要来确定。在 ET_0 的预测中,要把握其影响因子的轻重程度,作出合理的输入因子的选择。以关键因子作为输入项^[12]。同样,网络的输出因子也需根据网络的需要来确定。

2.2 网络层数的选择

BP 神经网络由输入层、隐含层和输出层及各层神经元之间连接组成^[13]。在模型样本相对较少的情况下,较少的隐层数可以实现模型样本空间的超平面划分。此时,三层 BP 即可行;当模式样本数很多时,增加一个隐层是必要的,但 BP 网络隐层数一般不超过两层。

2.3 隐层节点数的选择

隐层节点数的选择是一个十分复杂的问题,往往需要根据设计者的经验和多次实验来确定。如果隐含层神经元数目过少,网络很难识别样本,难以完成训练,并且网络的容错性也会降低;如果数目过多,则会增加网络的迭代次数,从而延长网络的训练时间,同时也会降低网络的泛化能力,导致预测能力下降。隐层节点数的选择至今没有统一的规范,可以用错试法来确定,即:首先给定较小的隐层节点数,构成一个较小的 BP 网络,进行训练,如果训练很多次或在规定的训练次数内没有满足收敛条件,停止训练。当训练次数骤然减少,以后再增加隐层节点数时,对训练次数影响不大,并且误差最小,此时的隐层节点数即是要求的隐层节点数。

3 模型的建立及应用检验

3.1 BP 神经网络建模

结合研究内容建立如下 BP— ET_0 模型,解决的问题。

3.1.1 训练、检验样本及输入、输出因子的确定 应用 BP 神经网络建模进行 ET_0 预测时,需先确定训练和检验样本集。为了能够指导东港地区的农业生产,选择 1999 年 5—9 月水稻生育期的逐日气象资料作为训练样本,2000 年同期的逐日气象资料作为检验样本。根据研究的需要,用 FAO 推荐的 Penman—Montieth 方法计算 ET_0 ^[14],作为输出因子。为了确定输入因子,将气温、净辐射、相对湿度、风速、气压等主要气象因子进行 ET_0 回归分析,选择与 ET_0 相关系数高的气象因子作为输入因子,如图 1 所示。

各气象要素与 ET_0 的相关系数由高到低依次是净辐射、相对湿度、风速、气温、水气压和气压。因此，选择日净辐射量、日平均相对湿度和日平均风速 3 个气象因子作为模型的输入因子。

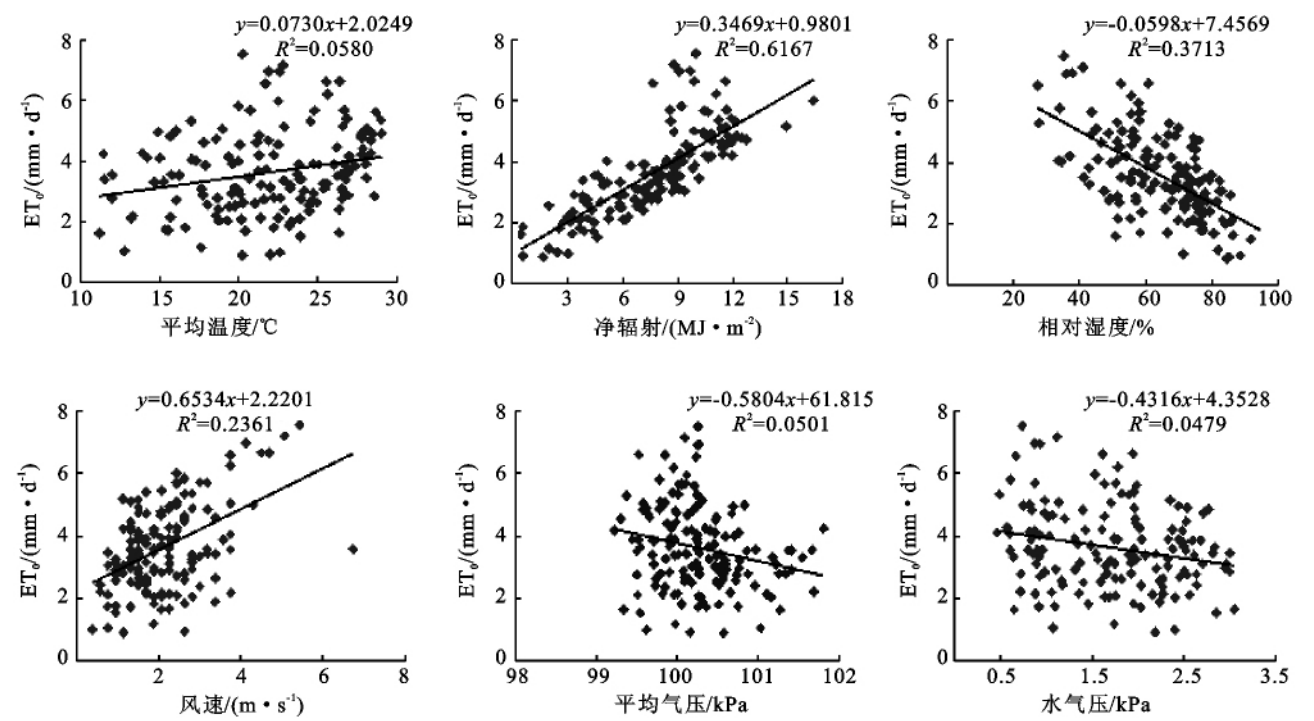


图 1 1999 年水稻生育期 ET_0 与各气象因子的关系图

3.1.2 网络层数和隐层节点数的确定 数学上已经证明多层前馈网络具有很强的函数映射功能,一个三层前馈人工神经网络足以解决一般函数的拟合逼近问题。 ET_0 的预测问题实际上是函数映射或拟合的问题,因此选择三层的 BP 网络。

对于给定的训练样本数,存在一个最佳 BP 网络结构,即最少隐层节点数,使网络训练次数最少且具有最广泛的预测能力。确定模型的隐层节点数分别从模型的训练次数与隐层节点数两方面考虑,如图 2、表 1 所示。

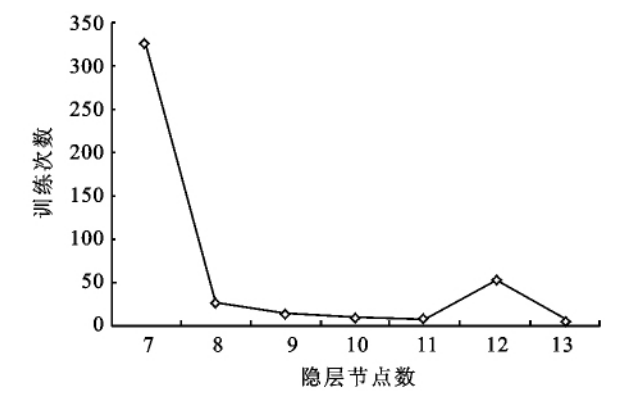


图 2 训练次数与隐层节点数关系图 (goal=0.0001)

图 2 显示,当隐层节点数为 8 时,训练次数骤然减少且趋于稳定;以后再增加隐层单元数,对训练次数影响不大;当隐层节点数小于 8 时,训练次数增加且不能满足收敛条件,当隐层节点数为 8~11 时,训

练次数出现一些波动,但是相对稳定。由此,初步确定隐层节点数为 8,9,10,11。

表 1 不同隐层节点数时网络预测的相对误差

节点数	8	9	10	11
最大相对误差	0.27	0.65	0.58	0.31
最小相对误差	0.00	0.00	0.00	0.00
平均相对误差	0.11	0.12	0.11	0.09

注:此表格数据均为相对误差的绝对值。

由表 1 可知,以平均相对误差为主要参考对象,以最大、最小相对误差为次要参考对象,当隐层节点数为 11 时,预测效果较好,确定 11 为隐层节点数。

3.1.3 传递函数的确定 网络中间层的神经元传递函数采用 S 型正切函数 tansig,输出层神经元传递函数采用 S 型对数函数 logsig。函数的输出位于区间 (0—1),满足网络输出的要求。

3.1.4 训练函数的确定 不同的训练函数对网络的收敛速度、性能、网络训练及预测结果等均有重要的影响,经比较确定网络的训练函数如表 2 所示。

由表 2 可知,当训练函数为 trainlm 函数时,网络的训练次数最少,预测效果最好。当训练函数为其他函数时,网络的训练次数较多,预测效果也相对较差。

3.1.5 网络参数的最终确定 根据以上研究,确定优化的 BP— ET_0 网络模型的具体参数为:网络结构:3—11—1;输入因子:日净辐射、日平均相对湿度和日平均风速;输出因子:FAO56 Penman—Montieth 计

算出的 ET_0 值;传递函数:输入、输出隐层均为 tansig 函数;训练函数: trainlm 函数。

表 2 不同训练函数的结果比较

训练函数	训练次数	训练 MSE 值(10^{-5})	预测相对误差
traindx	>10000	—	—
traincgf	435	9.89	0.11
traincgp	773	10.0	0.13
traincgb	633	9.96	0.12
trainscg	908	10.0	0.11
trainbfg	380	9.99	0.12
trainoss	1313	8.61	0.12
trainlm	9	6.70	0.09

3.2 模型的应用检验

3.2.1 ET_0 的预测结果与趋势分析 通过对比 BP— ET_0 模型的预测值及由 FAO56 Penman—Monteith 计算出的 ET_0 目标值的变化趋势,分析模型的趋势预测效果,结果如图 3 所示。

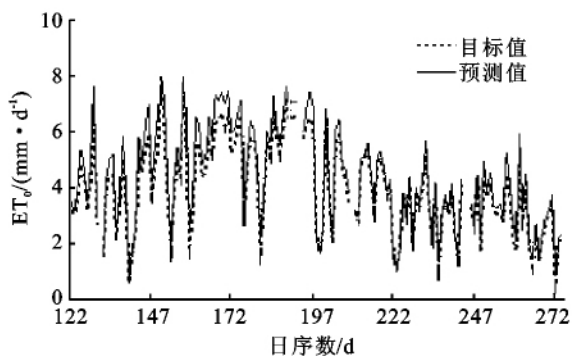


图 3 ET_0 目标值与预测值变化趋势图

由图 3 可知,BP— ET_0 模型的预测值与目标值变化趋势基本一致,可作 ET_0 的变化趋势预测。

3.2.2 绝对误差分析 通过分析 ET_0 的预测值与目标值的绝对误差,分析模型的模拟预测效果,如图 4 所示。

由图 4 可知,BP 模型的最小绝对误差为 0,最大绝对误差为 1.30 mm,绝对误差绝对值的平均值为 0.39 mm,并由绝对误差大多为正值可以看出用 BP— ET_0 模型预测,其预测值较目标值偏大。

3.2.3 相对误差分析 通过对比 ET_0 的预测值与目标值的相对误差,分析模型的模拟预测效果,如图 5 所示。

由预测的相对误差可知,BP 模型的最小相对误差为 0,最大相对误差为 30.5%,相对误差绝对值的平均值为 9%。

3.2.4 均方误差及根均方误差分析 通过分析 ET_0 的预测值与目标值的均方误差(MSE)和根均方误差(RMSE),分析模型的模拟预测效果,计算公式如下:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - z_i)^2$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - z_i)^2}$$

式中: n ——样本数; y_i ——预测值; z_i ——目标值。由计算可得,预测值与目标值的均方误差为 0.23,根均方差为 0.48。

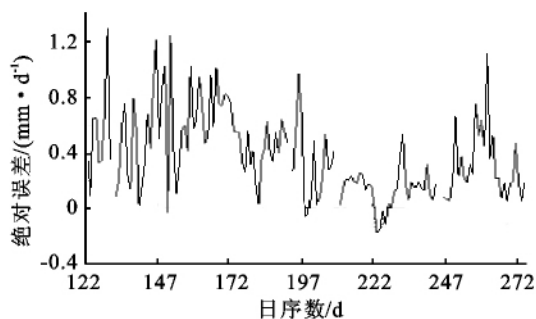


图 4 ET_0 预测值与目标值绝对误差

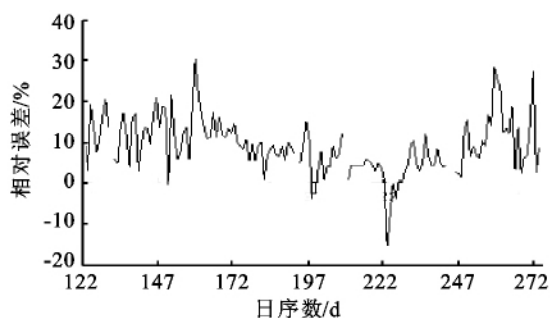


图 5 ET_0 预测值与目标值相对误差

4 结论

从多个气象因子的 ET_0 回归分析出发,确定模型的输入因子。通过对选择不同隐层节点数和选用不同的传递函数及训练函数的预测值与目标值的相对误差分析,得到了预测 ET_0 的最优 BP— ET_0 神经网络模型。通过检验结果分析,得到如下结论:

(1) 通过训练建立了 BP— ET_0 的预测模型,网络训练时具体参数是:网络结构为 3—11—1;输入因子分别为日净辐射量、日平均相对湿度和日平均风速;输出因子为用 Penman—Monteith 公式计算得到的同期 ET_0 值;传递函数为 tansig 函数;训练函数为 trainlm 函数。

(2) 用 BP— ET_0 模型预测出的 ET_0 与用 Penman—Monteith 公式计算出的 ET_0 随时间变化的趋势基本一致,即,可以用此模型对 ET_0 的变化趋势进行预测。

(3) 用 BP— ET_0 模型预测出的 ET_0 与用 Penman—Monteith 公式计算出的 ET_0 平均绝对误差为 0.39 mm;平均相对误差为 9%;均方差为 0.23,根均方差为 0.48 mm。预测值和目标值的绝对误差和相对误差大多为正值,BP— ET_0 模型预测值较目标值偏大。

(下转第 216 页)

参考文献:

- [1] 李秀彬. 全球环境变化研究的核心领域: 土地利用/土地覆被变化的国际研究动向[J]. 地理学报, 1996, 51(6): 553-558.
- [2] Turner B, Skole D, Sanderson D. Land-use and Land-Cover Changes: Science/research plan, International Geosphere-Biosphere Program [R]. Stockholm: IGBP Secretariat, 1995.
- [3] 张文忠, 王传胜, 吕昕, 等. 珠江三角洲土地利用变化与工业化和城市化的耦合关系[J]. 地理学报, 2003, 58(5): 677-685.
- [4] 何丹, 刁承泰. 重庆江津市土地利用变化及社会驱动力分析[J]. 水土保持研究, 2006, 13(2): 24-26.
- [5] 蒙吉军, 吴秀芹, 李正国. 河西走廊土地利用/覆被变化的景观生态效应: 以肃州绿洲为例[J]. 生态学报, 2004, 24(11): 2535-2541.
- [6] 杨卓, 李全, 魏斌, 等. 典型东北农牧交错区土地利用/覆被变化分析[J]. 水土保持研究, 2010, 17(4): 212-216.
- [7] 李小雁, 许何也, 马育军, 等. 青海湖流域土地利用/覆被变化研究[J]. 自然资源学报, 2008, 23(2): 285-296.
- [8] 刘建飞, 杨勤科, 梁伟, 等. 近 30 年来陕北黄土高原土地利用动态变化分析[J]. 水土保持研究, 2009, 16(2): 112-116.
- [9] 张怡芳, 邢元志, 濮励杰, 等. 苏州市土地利用变化的生态环境效应研究[J]. 水土保持研究, 2009, 16(5): 98-103.
- [10] 韩超峰, 陈仲新. LUCC 驱动力模型研究综述[J]. 中国农学通报, 2008, 24(4): 365-368.
- [11] 田彦军, 郝晋珉, 韩亮, 等. 县域土地利用程度评估模型构建及应用研究[J]. 农业工程学报, 2003, 19(6): 293-297.
- [12] 谢书明. 土地利用现状图及其数据资源的修测更新[J]. 测绘科技通讯, 1996, 19(1): 21-25.
- [13] 杨立民, 朱智良. 全球区域尺度土地覆盖土地利用遥感研究的现状和展望[J]. 自然资源学报, 1999, 14(4): 340-344.
- [14] 张杨, 刘艳芳, 丁庆, 等. 1996—2006 年武汉市土地利用/覆被变化研究[J]. 生态环境学报, 2010, 19(11): 2534-2539.
- [15] 张杨, 刘艳芳, 顾渐萍, 等. 武汉市土地利用覆被变化与生态环境效应研究[J]. 地理科学, 2011, 31(10): 1280-1285.
- [16] 杨述河, 闫海利, 郭丽英. 北方农牧交错带土地利用变化及其生态环境效应[J]. 地理科学进展, 2004, 23(6): 49-55.
- [17] Liu Y S, Gao J, Yang Y F. A holistic approach towards assessment of severity of land degradation along the great wall in Northern Shaanxi Province, China [J]. Environmental Monitoring and Assessment, 2003, 82(2): 187-202.
- [18] 李晓文, 方创琳, 黄金川, 等. 西北干旱区城市土地利用变化及其区域生态环境效应: 以甘肃省河西地区为例[J]. 第四纪研究, 2003, 23(3): 280-290.

(上接第 210 页)

参考文献:

- [1] 刘钰, Pereira L S. 气象数据缺测条件下参照腾发量的计算方法[J]. 水利学报, 2001(3): 11-17.
- [2] 李荣超. 水稻覆膜旱作节水高产灌溉模式研究[D]. 南京: 河海大学, 2000.
- [3] 王志良, 黄军学. 基于 MATLAB 神经网络工具箱的作物需水量预测[J]. 华北水利水电学院学报, 2009, 30(1): 5-6.
- [4] 徐俊增, 彭世彰. BP 神经网络在农田水利学科的应用[J]. 灌溉排水学报, 2003(专): 41-33.
- [5] 尚虎君, 马孝义, 高建恩, 等. 赵西宁作物需水量计算模型组件研究与应用[J]. 节水灌溉, 2011(8): 66-72.
- [6] 尹京川, 马孝义, 孙永胜, 等. 基于 BP 神经网络与 GIS 可视化的作物需水量预测[J]. 中国农村水利水电, 2012(2): 13-15, 18.
- [7] 苑希民, 李鸿雁. 神经网络和遗传算法在水科学领域的应用[M]. 北京: 水利水电出版社, 2002.
- [8] 王旭, 王宏, 王文辉. 人工神经网络原理与应用[M]. 沈阳: 东北大学出版社, 2002.
- [9] 武开福. 基于灰色关联度与 BP 神经网络模型的日参考作物腾发量预测[J]. 水土保持研究, 2011, 18(2): 237-240.
- [10] 王丽霞, 任志远, 孔金玲. 基于 BP 模型的延河流域社会经济需水预测[J]. 干旱区资源与环境, 2011, 25(4): 106-110.
- [11] 余世明. 作物需水量的灰色拓扑预测[J]. 四川水利, 1996(2): 18-21.
- [12] 马黎华, 康绍忠, 粟晓玲, 等. 农作区净灌溉需水量模拟及不确定性分析[J]. 农业工程学报, 2012, 28(8): 11-18.
- [13] 张兵, 袁寿其, 成立, 等. 基于 L—M 优化算法的 BP 神经网络的作物需水量预测模型[J]. 农业工程学报, 2004, 20(6): 73-76.
- [14] 贾德彬, 刘艳伟, 张永平, 等. 内蒙古河套灌区春小麦高效用水灌溉制度研究[J]. 干旱区资源与环境, 2008, 22(5): 174-177.