

基于 WSN 的温室番茄光合速率预测*

王伟珍 张漫 蒋毅琼 沙莎 李民赞

(中国农业大学现代精细农业系统集成研究教育部重点实验室, 北京 100083)

摘要: 为了提高 CO₂ 气肥的利用率,对日光温室番茄开花期光合速率变化进行了研究。采用无线传感器网络系统对温室环境信息进行实时监测;采用 LI-6400XT 型光合仪测定番茄植株叶片净光合作用速率,并对叶片的环境状况按照一定的规律进行调控。将经过主成分分析后的环境信息作为输入参数,将光合作用速率作为输出参数,利用 BP 神经网络建立了番茄开花期单叶净光合作用速率的预测模型,并对预测模型进行了性能评估。结果表明,所建立的光合作用速率模型预测值和实测值相关系数为 0.99,均方根误差为 0.288,具有较好的预测效果。在一定环境条件下改变 CO₂ 浓度的输入值,得到的光合作用速率预测曲线与实际曲线变化趋势一致,该模型可以作为温室番茄开花期 CO₂ 施肥量化调控的依据。

关键词: 番茄 温室 光合作用速率 无线传感器网络 BP 神经网络

中图分类号: S625.5⁺1; Q945.11 **文献标识码:** A **文章编号:** 1000-1298(2013)S2-0192-06

Photosynthetic Rate Prediction of Tomato Plants Based on Wireless Sensor Network in Greenhouse

Wang Weizhen Zhang Man Jiang Yiqiong Sha Sha Li Minzan

(Key Laboratory of Modern Precision Agriculture System Integration Research, Ministry of Education, China Agricultural University, Beijing 100083, China)

Abstract: In order to improve the utilization of CO₂ fertilizer, the photosynthetic rate of tomato plants in the flowering stage was studied. A wireless sensor network system was used to real time monitor greenhouse environmental parameters. A LI-6400XT portable photosynthesis analyzer was used to measure the photosynthetic rate of tomato plants, and the environmental parameters of leaves were controlled according to the presetting rule. The photosynthetic rate prediction models of single leaves were established based on the back-propagation (BP) neural network. The environmental parameters were used as input neurons after processed by principal component analysis (PCA), and the photosynthetic rate was taken as the output neuron. The performance of the prediction model was evaluated. The prediction results of the models showed that the correlation coefficient between the simulated and observed data sets was 0.99, RMSE was 0.288. Furthermore, when different CO₂ concentrations were selected as the input to predict the photosynthetic rate, the simulated and observed data showed the same trend. According to the above analysis, it was concluded that the model could be used as the basis of the quantitative regulation of CO₂ fertilization to tomato plants in greenhouse.

Key words: Tomato Greenhouse Photosynthetic rate Wireless sensor network BP neural network

引言

无线传感器网络 (WSN) 应用于温室有利于温

室现场环境信息数据的自动获取,进而有效发挥温室环境可调控的优势,提高劳动生产率和资源产出率,促进高产优质和增收增效^[1]。作物生长模型是

收稿日期: 2013-06-24 修回日期: 2013-07-05

* 国家自然科学基金资助项目(31271619)和高等学校博士学科点专项科研基金资助项目(20110008130006,20100008110030)

作者简介: 王伟珍,硕士生,主要从事无线传感器网络农业应用研究,E-mail: wwzbetter@aliyun.com

通讯作者: 张漫,副教授,博士生导师,主要从事精细农业及其支持技术研究,E-mail: cauzm@cau.edu.cn

温室作物生产环境优化调控的有力工具^[2-3]。对于番茄净光合作用速率模型的建立,已有大量的研究^[4-7],大多采用机理建模及相关参数的修正。绿色植物进行光合作用的主要来源是 CO_2 , CO_2 浓度对增加作物产量和提升品质都具有重要意义^[8]。为了有效利用由无线传感器网络采集的大量环境数据,以便更科学地指导温室生产,本文利用 WSN 技术和作物生长建模方法,将温室的环境信息与作物的生长信息进行结合,对温室番茄光合作用速率模型进行研究探讨。以 CO_2 浓度的调控作为切入点,利用 WSN 技术获取温室环境因子,通过 BP 神经网络建立单叶净光合作用速率与主要环境因子的关系模型,为一定环境条件下 CO_2 精细施肥调控提供管理建议。

1 材料和方法

1.1 系统总体结构

实验于 2012 年 12 月在中国农业大学水利与土木工程学院日光温室中进行,实验面积约 6 m^2 ,实验对象为处于开花期的番茄。图 1 为所用无线传感器网络采集系统的总体结构框图。

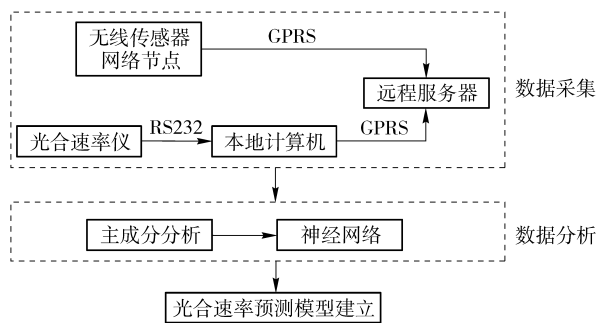


图 1 系统总体结构

Fig. 1 System architecture of WSN

通过无线传感器网络节点采集作物的环境因子信息,同时,为了研究不同的光辐射强度、 CO_2 浓度、空气湿度、空气温度、土壤温度和土壤湿度等环境因子对番茄光合作用速率的影响,进而量化 CO_2 浓度对作物净光合作用速率的影响,本研究采用 LI-6400XT 型便携式光合速率仪作为测量仪器,在人工控制环境条件下对叶片进行测量,并记录相关参数,主要包括净光合速率、大气 CO_2 浓度、光量子通量密度和空气相对湿度等。

1.2 数据采集

番茄植株处的环境信息由无线传感器网络节点负责测量,系统每 2 min 自动记录一次数据并通过 GPRS 方式发送到远程控制终端。

为了测量番茄植株在不同 CO_2 浓度、不同光强时光合作用速率的变化情况,以扩展数据范围,采用

人工方式控制 CO_2 浓度和光强。增施 CO_2 采用钢瓶释放法,光照采用人工光源。测量依照一定的规律进行,在光辐射强度分别为饱和光强、1 500、1 300、1 200、1 000、900、800、600、400 $\mu\text{mol}/(\text{m}^2 \cdot \text{s})$ 的条件下,分别测定 CO_2 摩尔比为 350、250、200、150、100、50、350、450、550、700、900、1 100、1 200、1 300、1 500、1 700、2 000 $\mu\text{mol}/\text{mol}$ 时的番茄叶片净光合作用速率。其他环境参数均为自然条件。

因为冬季外界光照强度最大值也远远低于叶片饱和光强,为了提高测量的准确性,需对叶片进行光诱导。选取自上而下第 3 花序下面的功能叶片进行测量。由于在自然条件下生长的 C3 植物叶片光合作用的饱和光强大约在 $1\,000 \mu\text{mol}/(\text{m}^2 \cdot \text{s})$ 左右^[9],首先将待测叶片置于叶室中,让待测定叶片在 $1\,000 \mu\text{mol}/(\text{m}^2 \cdot \text{s})$ 的光照下,在空气中诱导约 30 min,待叶片光合诱导期结束后,开始进行光合作用对光响应的观测。在光响应的观测中,先从光强为 $1\,000 \mu\text{mol}/(\text{m}^2 \cdot \text{s})$ 开始,以每次增加 $100 \mu\text{mol}/(\text{m}^2 \cdot \text{s})$ 的幅度依次提高光强,当光合作用速率不再随光强的增加而增加时,在每个光强下停留 2~3 min 后记录光合速率。在通过光合作用对光强变化响应的观测确知自然条件下生长的叶片光合作用饱和光强的前提下,先让待测叶片在饱和光强下和普通空气中诱导约 60 min。

光合作用达到稳态即诱导期结束后,在整个观测过程中,每次 CO_2 浓度变化后都需要进行 1 次光合速率仪的匹配操作,让同一气体流经样品室和参比室,使其重新平衡;然后再停留 3 min 左右后记录这一 CO_2 浓度下的净光合作用速率,使叶片适应此时环境的变化,以得到更精确的结果。每组数据记录 3 次,取平均值作为最终测量值。

1.3 数据分析

由于所测叶片的环境是经过人工调控的,为了保证一致性,空气温、湿度、 CO_2 浓度、光照强度采用 LI-6400XT 型光合速率仪中的环境数据,只采用传感器节点的土壤温、湿度。

1.3.1 数据预处理

首先,将传感器节点和 LI-6400XT 型光合速率仪采集的信息进行匹配,之后,还需要对错误数据进行剔除。实验时,当 CO_2 浓度或光照强度设置太低时,光合作用速率会出现在零左右波动的数值,这种数据在数据预处理时删除。其次,在实验过程中,若有人进入温室或者通风速率发生较大变化时,测量精度会受到影响,这种影响在人工控制环境下相对较小,但是这些活动仍需要特别记录,以利于数据分析。最后,将同一条件下的多组数据的平均值作为

测量值。

(1) 归一化处理

因为所测量的参数处于不同数量级,为了避免小数据被大数据淹没,发挥不出其应有的作用,同时为了使得各参数指标具有相同的量纲,需要对数据进行变换处理。

$$x' = \frac{2(x_{ij} - \min_{1 \leq k \leq n} x_{kj})}{\max_{1 \leq k \leq n} x_{kj} - \min_{1 \leq k \leq n} x_{kj}} - 1 \tag{1}$$

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{\sqrt{s_{jj}}} \tag{2}$$

 $(i = 1, 2, \cdots, n; j = 1, 2, \cdots, p)$

式中 x_{ij} ——第 j 个变量 X_j 的第 i 个观测值
 x'_{ij} —— x_{ij} 变换后的值
 $\max_{1 \leq k \leq n} x_{kj}, \min_{1 \leq k \leq n} x_{kj}$ ——变量 X_j 的最大、最小观测值
 $\bar{x}_j$ ——变量 X_j 的观测值的平均值
 $\sqrt{s_{jj}}$ ——变量 X_j 的观测值的标准差

式(1)、(2)为对 p 维向量 $X = (X_1, X_2, \cdots, X_p)$ 的两种变换方式。

(2) 主成分分析

考虑到所采集变量之间存在某种冗余,采用综合变量的方法,可达到减少变量间相关性且降维的目的。

主成分分析 (Principal components analysis, PCA) 能在损失很少信息的前提下把多个指标转换为几个综合指标,这些新的指标彼此既互不相关,又能综合反映原来多个指标的信息,是原来多个指标的线性组合。综合后的新指标称为原来指标的主成分或主分量,第 k 个主成分与原变量的系数称作因子载荷量。因子载荷量是主成分分析中重要的解释依据,由因子载荷量在主成分中的绝对值大小来刻画该主成分的主要意义及其组成因素^[10-11]。

在 SAS 系统主成分分析中,给出的因子载荷量与式(3)中的 b_{ij} 对应。而变量 x_i 对应的为经过式(2)标准化后的变量。

$$\begin{cases} Z_1 = b_{11}x_1 + b_{12}x_2 + \cdots + b_{1m}x_m \\ Z_2 = b_{21}x_1 + b_{22}x_2 + \cdots + b_{2m}x_m \\ \vdots \\ Z_m = b_{m1}x_1 + b_{m2}x_2 + \cdots + b_{mm}x_m \end{cases} \tag{3}$$

利用 $c_{ij} = \sqrt{\lambda_i}b_{ij}$ 对公式进行变换, λ_i 为与第 i 个主成分对应的特征值,得到的 c_{ij} 就是主成分 Z_i 与原始变量 x_i 的相关系数。此时,因子载荷量的大小和方向反映了主成分与相应变量的密切程度和变化方向。

1.3.2 光合速率预测模型建立

神经网络类似于一个黑盒子,不需要知道内部

机理,即可根据数据建立模型,找出输入属性和可预测属性之间的平滑而连续的非线性关系,因此适合于像光合作用预测这类复杂的模型。前馈型 BP (Back-propagation) 网络,即误差逆传播神经网络是最常用的一种神经网络^[12]。

神经网络的一个缺点是结果解释比较困难,这就需要较好的结果评估^[13-14]。一方面是对模型训练质量的度量,使训练误差达到最小;另一方面是对实际效果的评估。本文采用的训练结果评估参数分别是相关系数 (R)、平均相对误差 (ARE)、平均绝对误差 (MAE) 和均方根误差 (RMSE)。

由于数据量有限,为了充分利用数据,增加模型的稳定性,采用了交叉验证方法。交叉验证 (Cross-validation) 是为了有效的估测泛化误差所设计的实验方法,有 Hold out 验证、 K 折交叉验证及留一验证等多种形式。

本研究采用 K 折交叉验证。其过程为:将训练样本集随机分为 K 个集合,通常分为 K 等份,对其中的 $K-1$ 个集合进行训练,得到一个模型,并用该模型对剩下的一个集合进行样本测试。使用此方法可以从有限的学习数据中获取尽可能多的有效信息,可有效的避免陷入局部最小值,也可在一定程度上避免过拟合问题^[15]。同时,该过程重复 K 次,取 K 次过程中评估结果的平均值作为推广误差。

光合作用速率模型建立流程如图 2 所示。

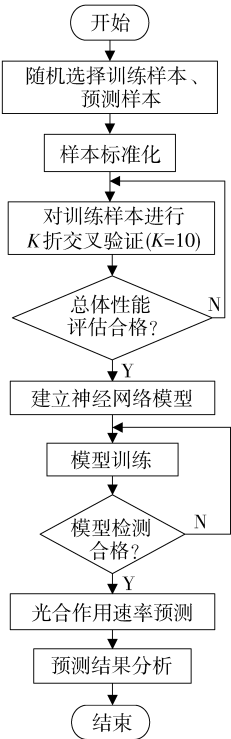


图 2 光合作用速率模型建模流程图

Fig. 2 Flowchart of photosynthetic rate modelling

2 结果与分析

2.1 主成分分析

为了减少输入变量的个数并使得网络训练效果更佳,采用 SAS 8.0(SAS Institute, Cary NC, USA) 的 PRINCOMP 过程进行主成分分析,分析结果如表 1 所示。

表 1 主成分特征值及其贡献率

Tab.1 Eigenvalue and contribution rate of each principal component

主成分	特征值	贡献量/%	累积贡献量/%
1	2.232 2871 1	37.20	37.20
2	1.486 256 90	24.77	61.97
3	1.301 878 51	21.70	83.67
4	0.500 385 50	8.34	92.01
5	0.460 364 84	7.67	99.68
6	0.018 827 14	0.32	100

从表 1 可以看出,前 4 个主成分已经包含了原始变量 92.01%(一般不小于 85%)的信息,但是考虑到主成分 5 与主成分 4 的贡献量很接近,所以保留 5 个主成分。

采用 CORR 过程求出各个变量与主成分之间的相关系数,如表 2 所示,第 1 行为相关系数,第 2 行为在相关系数为零的假设条件下的检验概率。从相关系数表可以看出,第 1 主成分中除空气温度外,其他变量在 0.05 水平上与主成分的相关性都是显著的;决定第 2 主成分的主要参数为空气温度、空气湿度、CO₂ 浓度和土壤温度;决定第 3 主成分的主要参数为空气温度、CO₂ 浓度和土壤湿度;决定第 4 主成分的主要参数为光强和土壤温度;决定第 5 主成分的主要参数为空气湿度和 CO₂ 浓度。

表 2 主成分与各变量间相关系数

Tab.2 Correlation coefficients between principal component and variables

	主成分 1	主成分 2	主成分 3	主成分 4	主成分 5
空气温度	-0.064 31 0.578 4	-0.284 43 0.012 2	0.943 89 <0.000 1	0.068 71 0.552 7	-0.123 43 0.284 9
空气湿度	0.467 41 <0.000 1	-0.727 89 <0.000 1	-0.007 19 0.950 5	-0.165 90 0.149 3	0.473 35 <0.000 1
光照强度	0.795 43 <0.000 1	-0.024 39 0.833 2	-0.143 10 0.214 4	0.588 19 <0.000 1	0.009 59 0.934 1
CO ₂ 浓度	-0.245 99 0.031 0	0.771 23 <0.000 1	0.352 20 0.001 7	0.111 53 0.334 2	0.456 31 <0.000 1
土壤湿度	0.835 65 <0.000 1	0.161 12 0.161 5	0.484 33 <0.000 1	-0.152 06 0.186 8	-0.100 44 0.384 8
土壤温度	0.786 22 <0.000 1	0.504 18 <0.000 1	-0.178 34 0.120 7	-0.294 31 0.009 4	-0.051 68 0.655 3

由表 2 可得式(4),进而得出一组新的数据集,用于后续模型的建立。

$$\left\{\begin{aligned} X'_1 &= -0.064x_1 + 0.467x_2 + 0.795x_3 - \\ &\quad 0.246x_4 + 0.836x_5 + 0.786x_6 \\ X'_2 &= -0.284x_1 - 0.728x_2 - 0.024x_3 + \\ &\quad 0.771x_4 + 0.161x_5 + 0.504x_6 \\ X'_3 &= 0.944x_1 - 0.007x_2 - 0.143x_3 + \\ &\quad 0.352x_4 + 0.484x_5 - 0.178x_6 \\ X'_4 &= 0.069x_1 - 0.166x_2 + 0.588x_3 + \\ &\quad 0.112x_4 - 0.152x_5 - 0.294x_6 \\ X'_5 &= -0.123x_1 + 0.473x_2 + 0.010x_3 + \\ &\quad 0.456x_4 - 0.100x_5 - 0.052x_6 \end{aligned}\right. \quad (4)$$

2.2 光合作用速率模型建立与性能评估

2.2.1 神经网络结构

采用 3 层神经网络结构,即输入层、单隐含层和输出层。将主成分分析后所得的 5 个综合指标作为神经网络的输入变量,光合作用速率作为输出变量。由于所得的综合变量的数值范围并不在[-1,1]之间,所以先采用式(1)对数据进行归一化处理。

隐含层节点数根据式(5)选取,通过采用由小到大逐渐递增的方法逐次训练。本研究采用的数值范围为[3,13],结合预测效果最终选取 10 个隐含层节点。

$$n_h = \sqrt{n_i + n_o} + l \quad (5)$$

式中

n_h ——隐含层节点数

n_i ——输入层节点数

n_o ——输出层节点数

l ——常数,取值范围(1,10)

2.2.2 传输函数

传输函数包括多种,其中较为常用的有对数 S 型(logsig)、双曲正切 S 型(tansig)和线性(purelin)传输函数^[16]。在本模型中,输入层和隐含层之间采用 logsig 传输函数,隐含层和输出层之间采用线性传输函数。

2.2.3 光合速率模型性能评估

对所有数据随机分为训练集和测试集,其中训练集为 70 组,占 91%,测试集为 7 组,占 9%。进行 10 折交叉验证后,对总体的验证结果进行平均,得到的平均性能为:相关系数为 0.790;平均相对误差为 0.328;平均绝对误差为 0.678;均方根误差为 1.02。

图 3 为对测试数据的预测结果。模型的性能评估结果分别是:相关系数为 0.99;平均相对误差为 0.177;平均绝对误差为 0.243;均方根误差为 0.288。可以看出,预测结果优于交叉验证的总体平均性能,说明该模型具有较好的泛化性能。

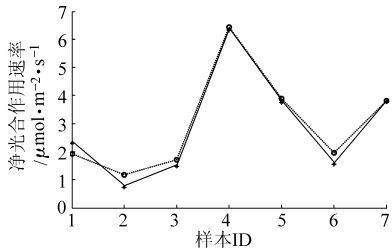


图 3 BPNN 训练效果

Fig. 3 Testing results of BPNN

2.3 光合速率模型应用

在一定的环境条件下,通过在一定范围内改变CO₂浓度,使用该模型对番茄的光合作用速率进行预测,从而找到使光合作用速率达到最大时的CO₂浓度,即找到最优的CO₂调控浓度。如图4所示,其中空气温度为21.14℃,空气湿度为44.41%,土壤温度为14.9℃,土壤湿度为48.8%,光照强度为701 μmol/(m²·s)。从图4可以看出光合作用速率最大时CO₂摩尔比的最优值在650 μmol/mol左右。实际测量结果如图5所示,可以看出预测值与实测值趋势大体相似,但是由于预测值计算时是假定其他环境是固定的,而实际测量值的环境条件都在变

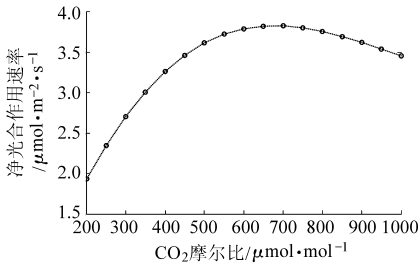


图 4 不同 CO₂ 浓度对应的 光 合 作 用 速 率 (预 测)

Fig. 4 Simulated relationship between CO₂ and photosynthetic rate

化,所以存在一定差异。采用神经网络进行建模需要的时间虽然较长,但当模型建立好进行预测时速度较快,因此可以满足实际需求。

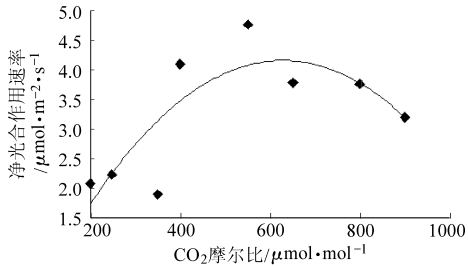


图 5 不同 CO₂ 浓度对应的 光 合 作 用 速 率 (实 测)

Fig. 5 Real relationship between CO₂ and photosynthetic rate

由于本次实验时间为冬季番茄开花期,叶片长期处于光照较弱的情况下,植株叶片很薄,番茄的长势也不是很理想,所以叶片光合作用速率整体较低。

3 结 束 语

通过建立无线传感器网络系统实时采集温室环境信息,通过人工调控环境的方法对环境进行调控,并记录相应环境信息及光合作用速率,将两者整合后的环境信息经主成分分析后作为BP神经网络的输入变量,建立了温室番茄环境信息与光合作用速率的预测模型。数据分析结果表明,所建立模型的预测值与实测值的相关系数高达0.99,可较好地预测番茄开花期的光合作用速率。另外,本研究假定其他环境处于较为适宜的状态时进行CO₂浓度的调控研究。但是事实上由于冬季番茄长期生长在低光照下,虽然对叶片进行了光诱导,但可能由于诱导不充分所致,在改变光强及CO₂浓度时,光合作用速率变动幅度较小。

参 考 文 献

1 Gonda L, Cugnasca C E. A proposal of greenhouse control using wireless sensor networks [C] // Computers in Agriculture and Natural Resources, 4th World Congress Conference, ASABE Paper 701P0606, 2006.

2 李天来, 颜阿丹, 罗新兰, 等. 日光温室番茄单叶净光合作用速率模型的温度修正[J]. 农业工程学报, 2010, 26(9): 274 ~ 279.

Li Tianlai, Yan Adan, Luo Xinlan, et al. Temperature modified model for single-leaf net photosynthetic rate of greenhouse tomato [J]. Transactions of the CSAE, 2010, 26(9): 274 ~ 279. (in Chinese)

3 罗新兰, 李天来, 李国春, 等. 日光温室番茄叶片净光合速率与气象因子的关系[J]. 江苏农业科学, 2007(4): 89 ~ 92.

4 El-Sharkawy M A. Overview: early history of crop growth and photosynthesis modeling [J]. Biosystems, 2011, 103(2): 205 ~ 211.

5 Gary C, Jones J W, Tchamitchian M. Crop modeling in horticulture: state of the art [J]. Scientia Horticulture, 1998, 74: 3 ~ 20.

6 Jones J W, Dayan E, Allen L H, et al. A dynamic tomato growth and yield model (TOMGRO) [J]. Transactions of the ASAE, 1991, 34(2): 663 ~ 672.

7 Zhang J, Wang S X. Simulation of the canopy photosynthesis model of greenhouse tomato [J]. Procedia Engineering, 2011, 16: 632 ~ 639.

8 贾探民, 杜双田, 周雷. 日光温室沼气 CO₂ 施肥[J]. 农村实用工程技术, 2003(10): 19.

9 黄燕, 吴平. SAS 统计分析及应用[M]. 北京: 机械工业出版社, 2006.

10 陈根云, 俞冠路, 陈悦, 等. 光合作用对光和二氧化碳响应的观测方法探讨[J]. 植物生理与分子生物学学报, 2006,

32(6):691~696.

Chen Genyun, Yu Guanlu, Chen Yue, et al. Exploring the observation methods of photosynthetic responses to light and carbon dioxide [J]. Journal of Plant Physiology and Molecular Biology, 2006, 32(6):691~696. (in Chinese)

11 贺昌政,俞海. BP 人工神经网络主成分分析预测模型及应用[J]. 数量经济技术经济研究,2001,18(9):104~106.

12 韩磊,李锐,朱会利. 基于 BP 神经网络的土壤养分综合评价模型[J]. 农业机械学报,2011,42(7): 109~115.

Han Lei, Li Rui, Zhu Huili. Comprehensive evaluation model of soil nutrient based on BP neural network [J]. Transactions of the Chinese Society for Agricultural Machinery, 2011, 42(7):109~115. (in Chinese)

13 Ehret D L, Hill B D, Raworth D A, et al. Artificial neural network modeling to predict cuticle cracking in greenhouse peppers and tomatoes [J]. Computers and Electronics in Agriculture, 2008, 61(2): 108~116.

14 项美晶. 基于信息融合的温室环境因子调控优化研究[D]. 镇江:江苏大学,2009.

15 韩萌,丁剑. 基于交叉验证的 BP 算法的改进与实现[J]. 计算机工程与设计,2008,29(14):3 738~3 739.

Han Meng, Ding Jian. Improvement of BP algorithm based on cross-validation method and its implementation [J]. Computer Engineering and Design, 2008,29(14):3 738~3 739. (in Chinese)

16 周开利,康耀红. 神经网络模型及其 MATLAB 仿真程序设计[M]. 北京:清华大学出版社,2007.