

文章编号: 1001-0920(2005)06-0682-04

基于遗传算法和RBF网络的番茄生长模型辨识

张娟, 陈杰, 王珊珊
(北京理工大学自动控制系, 北京 100081)

摘要: 采用神经网络和遗传算法, 对温室栽培番茄生长过程中主要器官——茎的生长过程进行了建模。温室番茄的生长过程具有控制变量多、生长过程复杂等特点, 采用基于径向基函数(RBF)神经网络的辨识方法建立了温室栽培番茄生长的模型, 以温室中番茄的实测数据为训练和预测样本, 采用遗传算法进行训练。仿真结果表明, 该方法较其他方法更适用于温室番茄生长过程的建模。

关键词: 径向基函数网络; 遗传算法; 建模

中图分类号: TP273 **文献标识码:** A

Model building of growing process of tomato based on RBF neural network and genetic algorithm

ZHANG Juan, CHEN Jie, WANG Shan-shan

(Department of Automatic Control, Beijing Institute of Technology, Beijing 100081, China Correspondent: ZHANG Juna, E-mail: zhjuan@bit.edu.cn)

Abstract: The modeling of the growing status of the cauline of tomato in growing process is discussed, by using NN and genetic algorithm. The model of the growing process of tomato is constructed based on RBF neural network. The real data of the tomato planted in greenhouse are used as the sample data and the genetic algorithm is used for getting the parameters. Simulation results show that the model has better performance than those based on other method.

Key words: RBF neural network; GA; modeling

1 引言

现代化的温室大都具备先进完备的各种硬件系统, 能够对温度、湿度、光照等环境条件进行控制。在此基础上加强对作物自身生长发育的监控, 可使环境-作物-技术体系有机结合, 从而更好地实现对作物生长的控制。完善的生长模型可根据作物自身状态信息以及农田或设施小气候、土壤等环境信息给出作物长势诊断, 为决策管理提供依据。目前我国已在多种大田作物生产管理专家系统的研究方面取得了一定的成果, 并有几种常见蔬菜的生产管理专家系统正在研制。然而关于蔬菜作物生长模型的系统研究成果较少^[1]。

本文主要是为温室栽培的番茄茎的生长状况进行建模, 在建模过程中采用了遗传算法和径向基函数(RBF)神经网络算法相结合的方法。

2 问题分析

植物的生长发育过程十分复杂, 主要表现为: 1) 影响因素众多; 2) 各变量与生长结果之间可能具有较强的动态耦合变化过程; 3) 生长过程呈现随机性、非线性、多变性及突变性。

建立一个完整的生长过程模型, 并全面反映出以上特征是十分困难的。一般来说, 植物生长模型总是建立在一定的假设条件之下, 并根据某些特定研究内容, 对植物生长过程进行的简化数学处理^[2]。

收稿日期: 2004-09-13; 修回日期: 2004-11-15

基金项目: “十五”国家重点科技攻关项目(2001BA503B01)。

作者简介: 张娟(1976—), 女, 河北秦皇岛人, 博士, 从事约束系统稳定控制、智能控制等研究; 陈杰(1965—), 男, 福建福州人, 教授, 博士生导师, 从事非线性控制、智能控制等研究。

番茄栽培于温室中,影响番茄生长发育过程的因素很多,其中包括温室内的各项环境因子(温度、湿度、光照强度、二氧化碳浓度、栽培土壤的条件)以及温室管理控制因素(水肥灌溉条件、农药的喷洒等)。另外,番茄在生长过程中各个器官的生长状况各不相同

番茄生长过程建模的重点在于对产量形成直接相关的生长发育进程和生理生态指标的模拟,因此在建立模型的过程中,根据有关专家的经验,选择了温室室内的温度、湿度、光照强度和二氧化碳浓度作为番茄生长过程的主要影响因素,并选取番茄生长过程中的一个主要器官——茎的生长状况作为建模对象

3 番茄生长过程的 RBF 神经网络模型

近年发展起来的 RBF 理论具有拓扑结构简单和学习速度快等优点^[3]。遗传算法是模拟生物进化现象,采用进化的机制,快速有效解决问题的一种随机性迭代式概率搜索方法。遗传算法和神经网络相结合时,可只考虑体现问题的信息模式,并通过对信息模式的处理求得问题解,特别适用于非线性极强而难以建立模型的系统。基于以上原因,本文针对番茄生长过程的建模问题,提出一种 RBF 神经网络和遗传算法相结合的方法来建立温室栽培番茄的生长过程模型

3.1 模型建立

为预测番茄生长过程中茎高的变化,取温室环境因子:温度、湿度、光照强度和 CO₂ 浓度为输入变量,以茎在生长过程中每天的增长量为输出变量,建立如图 1 所示的模型辨识结构。选择番茄生长过程中的第 10 天到第 174 天中的 165 个株高样本对 RBF 神经网络进行训练与测试

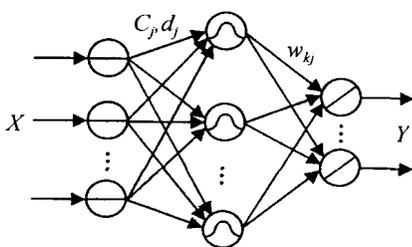


图1 RBF 网络拓扑结构图

图1中, X 为 N 维输入向量, $X \in R^N$ ($N = 4$); Y 为 P 维输出向量, $Y \in R^P$ ($P = 1$); C_j 和 d_j 分别为第 j 个隐层神经元的中心向量和宽度向量, C_j 和 $d_j \in R^N$; w_{kj} 为第 k 个输出神经元与第 j 个隐层神经元之间的调节权重

通常 RBF 神经网络隐层中的基函数使用径向

基函数,它对输入激励产生一个局部化的响应,即仅当输入落在输入空间中一个很小的指定区域时,隐单元才作出有意义的非零响应。输出层的输出为各隐单元的输出加权和。输入到隐层单元的权值固定为 1,隐层单元到输出单元之间的权值可调^[4]。本文采用的基函数是高斯函数,即

$$G(X, C_j) = G\left(\frac{X - C_j}{d_j}\right) = \exp\left[-\frac{\|X - C_j\|^2}{2d_j^2}\right], \quad j = 1, 2, \dots, M. \quad (1)$$

其中: $G(X, C_j)$ 为隐层第 j 个单元的输出, M 为隐层数。在 RBF 神经网络的拓扑结构中,需要确定 3 种权重参数和隐层神经元数。本文中,隐层神经元数 M 采用试凑方法来确定。选定后采用遗传学习算法对网络进行训练以求解上述 3 种权重参数

3.2 训练算法

遗传算法^[5]建立在自然选择和种群遗传基础上,模拟自然界“物竞天择,适者生存”的进化过程,在问题空间进行全局并行的、随机的搜索优化,使得种群全局最优地收敛。遗传算法的计算过程是将实际的优化问题编码成符号串,也称染色体,实际问题的对象函数则用染色体的适应函数表示。在最初随机产生一群染色体的基础上,根据各染色体的适应函数进行复制、交叉和变异等遗传操作,产生下一代染色体

3.2.1 权重参数的求解步骤

上述权重参数的求解过程如下:

- 1) 对所有 3 种权重参数进行编码,形成一个染色体串
- 2) 随机产生初始种群,由多个染色体串组成,不同数量的染色体串构成不同种群规模
- 3) 对种群中的每个串进行评价,确定串的评价值(或适应度值)。
- 4) 将遗传算子(选择、交叉、变异)作用于种群,产生新一代种群
- 5) 评价新一代种群,如果某一个串满足要求,则计算终止,对应于该串的参数即为所求参数;否则转至 4),继续遗传操作

如果进化(迭代)次数达到规定次数时仍不能满足要求,则需要调整 M 值,直到有一个 M 值满足要求为止

3.2.2 实现遗传操作的方法

1) 编 码

常用的编码方式有二进制和实数两种。二进制编码在计算运行速度和精度方面不如实数编码^[5],因此本文采用实数编码。一个实数对应于一个具体

网络参数,串的长度 n 为网络参数个数的总和 ($n = 2NM + MP + P$, 输出层 P 个神经元的阈值按调节权重处理)。对于 RBF 神经网络,中心是重要参数,对网络性能有很大影响。不同的中心对应着不同的宽度,中心变化,宽度也随之变化,所以采用中心和宽度交替排列的编排顺序。这样编排的好处是,在交叉这一主要算子作用下,中心和与之对应的宽度在一个串上同时变化的概率较高,容易实现中心变化宽度也应随之变化的要求。对于一个染色体串,具体的编码方式是:不同的中心和对应的宽度顺次排列完后,再顺序编排各个调节权重。

2) 产生初始种群

用随机数发生器产生 p 个染色体串,整数 p 的取值范围一般在 $[50, 100]$ 之间选取,本文选择 $p = 10$ 。

3) 确定评价函数

评价函数定义为

$$F_i = 1/E, \quad i = 1, 2, \dots, p. \quad (2)$$

其中 F_i 为第 i 个染色体串对应的评价值, E 为神经网络训练集中所有训练样本的网络输出值与期望值的误差平方和。

4) 选择操作

遗传算法最常用的选择方法是比例选择法。然而计算研究发现,这种方法在进化的初中时期是一种简单有效的方法,但在进化后期,由于种群内的各个染色体串的适应度值比较接近,比例选择几乎处于纯粹的随机选择过程中,导致进化计算陷入迟钝状态。对此,本文采用先比例选择后排序选择的混合选择机制,克服了计算后期的迟钝现象。比例、排序选择的转换条件定义如下:设种群中最大的适应度值为 P_{\max} , 最小的为 P_{\min} , 如果 $P_{\max}/P_{\min} > r$ 则进行比例选择, 否则进行排序选择。经计算试验认为, 比值 = 1.5 时进行选择方法的转换是合适的。

5) 交叉操作

本文采用两点交叉操作方法。设种群中第 i 个串的表达形式为 $S^i = (S^i_1, S^i_2, \dots, S^i_n)$; S^1_k 和 S^2_k 是串 S^1 和 S^2 在两个交叉点 i 与 j 之间的第 k 个交叉实数(两交叉点是随机选取的), 那么经交叉后产生的两个子代串在 k 位置处的实数 S^{1h}_k 和 S^{2h}_k 可描述为

$$\begin{aligned} S^{1h}_k &= \beta S^1_k + (1 - \beta) S^2_k, \\ S^{2h}_k &= \beta S^2_k + (1 - \beta) S^1_k, \end{aligned} \quad (3)$$

这里取交叉概率 $\beta = 0.75$ 。

6) 变异操作

本文使用的遗传算法采取按位变异方法。在二进制编码中,被选中的变异位如果是“0”则变为“1”,如果是“1”则变为“0”。但是,实数编码不同于

二进制编码,以下是一种实数编码的按位变异方法。

设权重参数取值范围是 $[a, b]$, S^i_k 是第 i 个串上被选中变异的第 k 个数(相当于一个变异位, $k = 1, 2, \dots, n$)。 S^{ih}_k 是变异后对应 S^i_k 的数, γ 是 $(0, 1)$ 之间的随机数, 则 S^{ih}_k 可表示为

$$S^{ih}_k = \gamma S^i_k + d(1 - \gamma), \quad (4)$$

其中 $d = a + (b - a)\gamma$ 。

这种变异方法可保证后代参数仍在规定的取值范围内,而且当参数 γ 的取值趋于 0 时, S^{ih}_k 与 S^i_k 关系不大; 参数 γ 的取值趋于 1 时, S^{ih}_k 与 S^i_k 的差别很小。这样,在演化过程中改变 γ 的取值范围,使其取值逐渐增加,从而使得参数的变化范围逐渐减小,有助于提高算法的精度。

本文中变异概率的取值为 0.02。

4 仿真结果

用基于遗传算法的 RBF 神经网络对番茄生长过程的株高进行学习和预测。选择番茄生长过程中的第 10 天到第 174 天中的 165 个株高数据的前 130 个作为学习样本,后 35 个作为预测样本。

番茄生长过程的每天株高增长量的学习和预测的拟合曲线如图 2~4 所示,图中实线代表实际数据,虚线代表网络的输出曲线。图 2~4 分别为基于遗传算法的 RBF 神经网络、基于 K-Means 算法的 RBF 网络和基于遗传算法的 3 层前馈网络预测后的结果。

上述 3 种方法预测的平均相对误差分别为 0.097, 0.115 281 和 0.107。由实验结果可以看出,

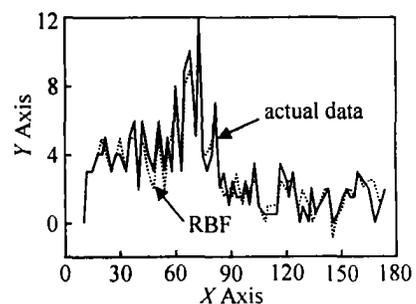


图2 基于遗传算法的 RBF 网络

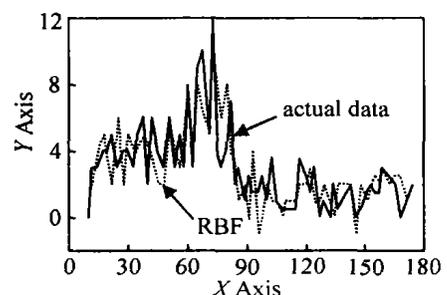


图3 基于 K-Means 的 RBF 网络

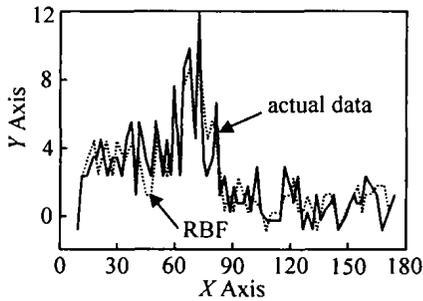


图 4 基于 GA 的 3 层前馈网络

基于遗传算法的 RBF 网络比基于 K-Means 算法的 RBF 网络和基于遗传算法的 3 层前馈网络有更好的预测效果

5 结 论

本文采用基于遗传算法的 RBF 神经网络实现了对温室栽培番茄茎的生长过程的建模。该模型结构简单, 易于实现, 且具有较高的精度, 可用于温室栽培番茄生长过程中主要器官——茎的生长状况的预估。

参考文献(References)

[1] 宋有洪, 贾文涛, 郭焱, 等. 虚拟作物研究进展[J]. 计算

机与农业, 2000, (6): 6-8

(Song Y H, Jia W T, Guo Y, et al. A review virtual crop research[J]. *Computer and Agriculture*, 2000, (6): 6-8)

[2] 金之庆. 作物模拟的发展趋势与应用前景[J]. *世界农业*, 1999, (6): 26-28

(Jin Z Q. The developmental trends and applied prospects of crop simulation [J]. *World Agriculture*, 1999, (6): 26-28)

[3] Kario W iedl, John D Heketh. *Plant growth modeling for resource management* [M]. Florida: CRC Press, 1994

[4] Vittorio Maniezzo. Genetic evolution of the topology and weight distribution of neural networks[J]. *IEEE Trans on Neural Networks*, 1994, 5(1): 39-53

[5] 胡宏银, 朱绍文, 张大斌, 等. 用实数编码的遗传算法构造斜决策树[J]. *计算机科学*, 2001, 28(2): 108-110

(Hu H Y, Zhu S W, Zhang D B, et al. Oblique decision tree construction with decimal-coded genetic algorithm [J]. *Computer Science*, 2001, 28(2): 108-110)

(上接第 681 页)

[2] Hubber H E, Redmond J A, McDonald J R. Multi-attribute trade-off analysis applied to long-term generation planning [J]. *IEEE Conf Publication*, 1994, 388 (2): 906-909

[3] 王欣荣, 樊治平. 一种基于自然语言评价信息的多指标群决策方法[J]. *系统工程学报*, 2003, 18(2): 173-176
(Wang X R, Fan Z P. The multi-index group decision making based the natural linguistic information[J]. *The System Engineering Transaction*, 2003, 18 (2): 173-176)

[4] Herrera F, Herrera-Viedma E, Verdegay J L. A linguistic decision processing in group decision process in group decision making [J]. *Group Decision Negotiation*, 1996, 31(5): 165-176

[5] Herrera F, Herrera-Viedma E. Linguistic decision analysis: Steps for solving decision problems under linguistic information [J]. *Fuzzy Set and Systems*, 2000, 115(1): 67-82

[6] Cheng C H, Lin Y. Evaluating the best main battle tank using fuzzy decision theory with linguistic criteria evaluation [J]. *European J of Operational Research*, 2002, 142(1): 174-1869

[7] 李洪燕, 樊治平. 一种基于二元语义的多指标群决策方法[J]. *东北大学学报(自然科学版)*, 2003, 24(5): 495-

498

(Li H Y, Fan Z P. Multi criteria group decision making method based on two-tuple linguistic information processing [J]. *J of Northeastern University (Natural Science)*, 2003, 23(5): 495-498)

[8] 陈岩, 樊治平. 群决策中基于语言评价信息的 TOPSIS 方法[J]. *南京工业大学学报*, 2004, 26(3): 27-30

(Chen Y, Fan Z P. TOPSIS method for group decision making based on linguistic assessment information [J]. *J of Nanjing University of Technology*, 2004, 26(3): 27-30

[9] 邱凯昌, 李德仁, 李德毅. 云理论及其在空间数据挖掘和知识发现中的应用[J]. *中国图像图形学报*, 1999, 4(11): 929-935

(Di K C, Li D Y, Li D Y. The cloud theory and application in the space data mining and knowledge discovery [J]. *The Chinese Graph and Image Transaction*, 1999, 4(11): 929-935)

[10] 邱凯昌. *空间数据挖掘与知识发现* [M]. 武汉: 武汉大学出版社, 2001

[11] 邱凯昌, 李德仁, 李德毅. 语言云模型的扩展及其在空间数据挖掘中的应用[A]. *第九届全国图像图形学术会议* [C]. 西安, 1998, 5: 521-526