

# BP神经网络在烟草蚜传病毒病预测中的应用\*

任广伟 王凤龙 高汉杰 王永

## 摘要

利用1980~1998年在山东青州调查的数据资料,结合相关气象因子及烟草蚜传病毒病的病情指数,采用BP神经网络方法建立了烟草蚜传病毒病的预测模型。利用该模型对1999~2002年的发病情况进行预测,预测准确度较高。BP神经网络为烟草蚜传病毒病的预测预报提供了一种新方法。

关键词:烟草 蚜传病毒病 BP神经网络 预测模型

中图分类号:S431

文献标识码:A

文章编号:1004-5708(2004)04-0023-04

烟草病毒病是烟草整个生育期的重要病害,在我国绝大多数烟区发生和危害都相当严重。在山东省,自20世纪60年代起至今,以蚜虫为主要传播媒介的病毒病种类如黄瓜花叶病毒病(CMV)、马铃薯Y病毒病(PVY)等已上升为山东烟区的主要病毒病害<sup>[1~3]</sup>。

对于烟草蚜传病毒病的防治,目前仍然遵循“预防为主,综合防治”的策略。因此,提前对其发生情况作出预测,对于制定合理的综合治理方案尤其重要。国内有关研究人员曾采用逐步回归的方法建立了烟草蚜传病毒病的预测模型<sup>[3~5]</sup>,但大都存在或总体样本数量少,或只有一个待测样本,或没有待测样本等问题,因此无法对模型进行足够的验证。近年来,人工神经网络理论得到了迅速发展,其中的BP(Back Propagation)神经网络已在多个领域有所应用。在农业病虫害的预测预报中,有关学者采用BP神经网络的方法建立了预测预报模型,取得了较好的预测结果<sup>[6~9]</sup>。在烟草病虫害的预测预报中,未见采用BP神经网络方法进行预测的报道,本研究将利用BP神经网络建立烟草蚜传病毒病的预测预报模型,为烟草蚜传病毒病的综合治理提供理论依据。

## 1 BP神经网络的基本原理

人工神经网络是一种模拟人的神经系统而建立起来的非线性动力学模型,由大量的被称为神经元的简要信息处理单元通过高度并联、互联而组成,每个神经元从它邻近的神经元接受信息,同时也向邻近的其他神经元发出信息,整个网络系统的信息处理是通过神经元之间的相互作用来完成的。

目前,已经发展了几十种神经网络,误差反向传递学习算法(BP算法,Back Propagation),实现了多层网络设想,是近年来使用最多的神经网络之一。BP神经网络模型采用的是并行网络结构,包括输入层、隐含层和输出层(图1)。对于输入信号,先向前传播到隐含层节点,经作用函数后,再把隐节点的输出信号传播到输出节点,最后给出输出结果。该算法的学习过程由正向传播和反向传播组成。在正向传播的过程中,输入信息从输入层经隐含层逐层处理,并传向输出层。每一层神经元的状态只影响下一层神经元的状态。如果输出层得不到期望的输出结果,则转入反向传播,将误差信号沿原来的连接通道返回,通过修改各层神经元的权值,使得误差信号最小<sup>[10]</sup>。

## 2 BP神经网络在烟草蚜传病毒病预测中的应用

### 2.1 田间调查

田间烟草蚜传病毒病及烟蚜的调查均在中国农科

\* 任广伟,男,30岁,硕士,青州烟草研究所,山东青州,262500

王凤龙,通讯地址同第一作者

高汉杰,王永,青州卷烟厂,山东青州,262500

收稿日期:2004-01-07

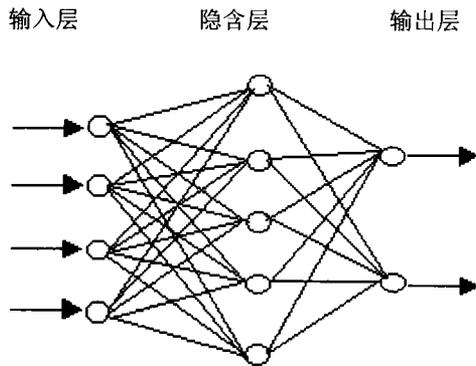


图1 BP神经网络模型

院烟草研究所试验农场内进行。蚜传病毒病的病情指数系在感病品种 NC89 和 G28 上调查所得,有翅蚜量系在感蚜品种 G28 上调查所得。

烟草蚜传病毒病的田间调查依据烟草行业标准《烟草病害分级及调查方法》(YC/T 39-1996)进行,并根据田间调查结果计算病情指数。分级标准如下(以株为单位):

0级:全株无病。

1级:心叶脉明或轻微花叶,或上部 1/3 叶片花叶但不变形,植株无明显矮化。

2级:1/3 至 1/2 叶片花叶,或少数叶片变形,或主脉变黑,植株矮化为正常株高的 2/3。

3级:1/2 至 2/3 叶片花叶,或变形或主侧脉坏死,或植株矮化为正常株高的 1/2 至 2/3。

4级:全株叶片花叶,严重变形或坏死,植株矮化为正常株高的 1/3 至 1/2。

## 2.2 预测因子及预报对象

选取上年 10 月份到当年 5 月份的月降雨量、月均温及移栽期、有翅蚜迁飞高峰期蚜量以及有翅蚜迁飞高峰期(移栽期和蚜虫迁飞高峰期以 5 月 1 日为 0)为预测因子,以 1980~2002 年(缺少 1984、1985、1989 年的数据)6 月下旬烟草蚜传病毒病发生高峰期的病情指数为预报对象。

## 2.3 预测准确度

采用肖悦岩的“最大误差参照法”计算公式<sup>[11]</sup>:

$$S_i(\%) = [1 - |X_i' - X_i| / X_i V(X_{\max} - X_i)] \times 100$$

其中:  $S_i$  表示预测准确度,  $X_i$  表示实测值,  $X_i'$  表示预测值,  $X_i V(X_{\max} - X_i)$  表示在  $X_i$  和  $X_{\max} - X_i$  中取较大值,即最大误差。

## 2.4 数据处理

为了消除各变量量纲的影响,将预测因子进行标准化。方法如下:

$$X' = (X - a) / S$$

其中  $a$  为平均值,  $S$  为标准差。

以 1980~1998 年数据资料为基础建立预测模型, 1999~2002 年为待测样本,采用 DPS 数据处理系统<sup>[10]</sup>进行数据分析。

## 2.5 BP神经网络模型的建立与分析

2.5.1 长期预测模型 根据逐步回归分析的结果,选取与预测对象相关性较强的 7 个变量( $X_1 \sim X_7$ )为预测因子,详见表 1。

表1 烟草蚜传病毒病的预测因子

符号	预测因子	符号	预测因子
X1	上年 10 月份降雨量	X5	上年 11 月份月均温
X2	上年 12 月份降雨量	X6	上年 12 月份月均温
X3	当年 3 月份降雨量	X7	当年 1 月份月均温
X4	上年 10 月份月均温	X8	当年 5 月份月均温

在 DPS 系统中,经反复调整,选取如下参数建立了 BP 神经网络模型。

输入层节点:7; 隐含层节点:6; 学习速率:0.2; 动态参数:0.7; 误差:0.00001; Sigmoid 参数:0.9; 最大学习次数:1000。

将预测因子数据进行标准化转换,经过学习后,收敛误差达到标准,学习结束。预测结果见表 2。

由表 2 可知, BP 神经网络长期预测模型对 16 个样本的预测结果与实测值非常接近,除 1982、1988 年的预测准确度分别为 98.24%、97.49%外,其余样本的预测准确度都在 99% 以上。

为了验证所建立的长期预测模型,将未参与训练的 1999~2002 年 4 个样本作为待测样本进行预测。由表 3 可知, 1999 和 2002 年的预测准确度分别为 99.95%、98.61%; 2000 和 2001 年的预测结果与实测值基本相符,准确度分别为 70.15%、83.47%。

该预测模型涉及到冬春的一些气象因子,可提前对下一个生长季节的烟草蚜传病毒病发生情况进行预测,根据肖悦岩等<sup>[12]</sup>的划分方法,可定为长期预测模型。

表2 BP神经网络模型的回测结果(长期预测模型)

年份	预测值	实测值	准确度%
1980	41.95	41.80	99.64
1981	6.88	6.59	99.59
1982	6.69	5.42	98.24
1983	24.35	24.29	99.90
1986	20.03	20.00	99.96
1987	15.53	15.60	99.89
1988	75.55	77.50	97.49
1990	20.37	20.54	99.70
1991	22.05	22.00	99.91
1992	42.94	42.86	99.81
1993	49.72	49.70	99.96
1994	45.33	45.34	99.97
1995	62.85	62.91	99.90
1996	48.65	48.54	99.77
1997	42.81	42.80	99.98
1998	46.54	46.60	99.87

表3 BP神经网络模型的预测结果(长期预测模型)

年份	预测值	实测值	准确度%
1999	70.97	71.00	99.95
2000	73.69	56.75	70.15
2001	75.24	64.57	83.47
2002	9.57	10.50	98.61

2.5.2 中期预测模型 根据逐步回归分析的结果,选取与预测对象相关性较强的8个变量(X1~X8)为预测因子(表1)。

在DPS系统中,选取如下参数建立了BP神经网络模型。

输入层节点:8,其他参数同长期预测模型。预测结果见表4。

由表4可知,BP神经网络中期预测模型对16个样本的预测结果与实测值非常接近,除1988年的预测准确度为97.81%外,其余样本的预测准确度都在99%以上。

将未参与训练的1999~2002年4个样本作为待测样本进行预测,结果见表5。其中1999、2001、2002年的预测准确度在94%以上;2000年的预测结果与实测值基本相符,预测准确度为74.13%。

在山东烟区,该预测模型可提前20天左右对烟草蚜传病毒病的发生情况进行预测。根据肖悦岩等<sup>[12]</sup>的

表4 BP神经网络模型的回测结果(中期预测模型)

年份	预测值	实测值	准确度%
1980	41.82	41.80	99.95
1981	6.72	6.59	99.82
1982	5.79	5.42	99.49
1983	24.32	24.29	99.94
1986	20.01	20.00	99.98
1987	15.54	15.60	99.90
1988	75.80	77.50	97.81
1990	20.57	20.54	99.95
1991	21.96	22.00	99.93
1992	42.85	42.86	99.98
1993	49.74	49.70	99.92
1994	45.32	45.34	99.96
1995	62.89	62.91	99.97
1996	48.58	48.54	99.92
1997	42.76	42.80	99.91
1998	46.56	46.60	99.91

表5 BP神经网络模型的预测结果(中期预测模型)

年份	预测值	实测值	准确度%
1999	75.17	71.00	94.13
2000	71.43	56.75	74.13
2001	62.39	64.57	96.62
2002	11.36	10.50	98.72

划分方法,可定为中期预测模型。

### 3 结论

3.1 本文运用BP神经网络的方法建立了烟草蚜传病毒病的长期和中期预测模型,对待测样本的预测结果较为准确,为烟草病害的预测预报提供了一种新方法。

3.2 运用BP神经网络建立预测预报模型,具有如下优点:(1)容错能力强。在神经网络模型中,数据信息分布存储于各神经单元中,个别输入信号误差不会引起识别错误;(2)预测准确率高;(3)BP神经网络具有很强的自组织、自适应能力,通过对典型例子的学习,能够掌握事物的本质特征。

3.3 国内对于烟草蚜传病毒病的预测,大都采用逐步回归的方法,本研究首次尝试采用BP神经网络的方法建立了烟草蚜传病毒病的长期和中期预测预报模型,对于预测的准确性和可靠性,还需要在生产实践中

进一步验证。烟草蚜传病毒病的发生流行受到多种因素的影响,如烟株的营养水平、品种的抗病性、毒源、种植制度以及烟蚜的消长与迁飞等。在进行预测预报时,除了参考以上模型的预测结果外,还要综合分析相关的影响因素,特别是一些未选入模型中的不可量化的因素,从而得到准确度较高的预测结果。

致谢:文中烟草蚜传病毒病的有关数据由我所植保室病毒课题组提供。

#### 参考文献

- 1 陈瑞泰,韩晓东,史万华,等. 五省区烟草病毒种类的初步研究[J]. 中国烟草,1986(1):1~4.
- 2 韩晓东,李林森,周嘉平,等. 山东烟草病毒病鉴定与防治的初步研究[J]. 中国烟草,1980(3):14~18.
- 3 王彦廷,朱贤朝. 烟草病虫害预测预报及综合防治技术研究进展[M]. 北京:中国农业科技出版社,2002.
- 4 成巨龙,安德荣,孙渭. 陕西省烟草蚜传病毒病的发生规律及预测预报模型的研究[J]. 中国烟草学报,1998,4(2):43~48.
- 5 金思明,王培琳,程增林,等. 烟草黄瓜花叶病测报方法初探[J]. 中国烟草,1987(2):1~3.
- 6 马飞,许晓风,张夕林,等. 神经网络预警系统及其在害虫预测中的应用[J]. 昆虫知识,2002,39(2):115~119.
- 7 高苹,局为民,陈宁,等. 人工神经网络方法在赤霉病预报中的应用研究[J]. 中国农业气象,2001,22(2):21~24.
- 8 林高飞,林乃铨,尤民生. 人工神经网络在三化螟预测中的应用[J]. 华东昆虫学报,1999,8(2):74~77.
- 9 胡小平,杨之为,李振岐,等. 汉中地区小麦条锈病的BP神经网络预测[J]. 西北农业学报,2000,9(3):28~31.
- 10 唐启义. 实用统计分析及其DPS数据处理系统[M]. 北京:科学出版社,2002,268~278.
- 11 肖悦岩. 预测预报准确度评估方法的研究[J]. 植保技术与推广,1997,17(4):3~6.
- 12 肖悦岩,季伯衡,杨之为,等. 植物病害流行与预测[M]. 北京:中国农业大学出版社,1998,80~81.

### Applying BP neural network to predict tobacco virus diseases transmitted by aphids

Ren Guangwei<sup>1</sup> Wang Fenglong<sup>1</sup> Gao Hanjie<sup>2</sup> Wang Yong<sup>2</sup>

1 Qingzhou Tobacco Research Institute, Qingzhou, Shandong 262500

2 Qingzhou Cigarette Factory, Shandong 262500

#### Abstract

Based on data of meteorological factors and tobacco virus diseases transmitted by aphids from 1980 to 1998 in Qingzhou of Shandong province prediction models were established by means of BP neural network. Prediction of 1999-2002 was carried out respectively and results showed that the accurate ratio of prediction was high. BP neural network can be used as a new effective method to predict the trend of tobacco virus diseases transmitted by aphids.

**Key words:** Tobacco virus diseases Aphids BP neural network Prediction model