

文章编号: 1001-1498(2003)02-0159-07

基于人工神经网络的马尾松毛虫 发生量预测模型的研究

陈绘画¹, 朱寿燕², 崔相富¹

(1. 浙江省仙居县林业局, 浙江 仙居 317300; 2. 浙江省仙居县气象局, 浙江 仙居 317300)

摘要: 运用人工神经网络的原理和方法, 根据相关系数法和逐步回归法分别选取与马尾松毛虫有虫面积、虫口密度、虫株率相关关系密切的气象因子作为样本的输入特征, 分别建立马尾松毛虫有虫面积、虫口密度、虫株率与气象因子的 BP 网络模型。结果表明: 所建立的各 BP 模型, 具有令人满意的拟合精度和预测精度。当隐含层神经元个数为 15 个, 预报因子数为 8 个时, 2 组预留有虫面积的 2 a 平均预测误差为 3.15%; 虫口密度 BP 模型的隐层神经元个数为 8 个, 预报因子数为 6 个时, 预留样本的平均预测误差为 5.91%; 虫株率 BP 模型的隐层神经元个数为 4 个, 预报因子数为 5 个时, 预留样本的平均预测误差为 10.65%。

关键词: 马尾松毛虫; 人工神经网络; 发生量; 预测预报; BP 模型

中图分类号: S763 文献标识码: A

森林害虫预测预报是森林害虫综合管理中极其重要的组成部分。加强对森林害虫发生趋势和危害程度的研究, 在其未大发生前及时加以控制, 可使森林害虫对森林资源危害造成的损失减少到最低限度, 同时还能使森林害虫的综合管理工作收到事半功倍之效。

马尾松毛虫 (*Dendrolimus punctatus* Walker) 是我国南方马尾松 (*Pinus massoniana* Lamb.) 林的最大害虫。马尾松毛虫的发生及其种群数量变化, 除受到其自身的遗传特性影响外, 还受到气候、林相、植被、食料和天敌 5 类因子的影响^[1]。但林相、植被、食料和天敌 4 类因子是定性变量, 很难进行数值表现, 更难以预测其变化, 而气象因子与马尾松毛虫的发生、生长及种群数量变化的关系也十分密切, 故前人根据气象因子与马尾松毛虫发生量的相互关系, 利用逐步回归分析^[2]、判别分析^[3]、模糊聚类分析^[4]、马尔可夫链分析^[5]等多种方法对其发生趋势或发生量进行预测预报。但气象因子与马尾松毛虫的关系并非线性关系, 所以利用这些方法预测马尾松毛虫的发生量, 由于还没有充分利用气象因子与马尾松毛虫有虫面积之间的相互关系, 无论是预测模型的拟合精度还是发生量的预测精度均有待进一步的提高, 而对马尾松毛虫的虫口密度、虫株率的预测预报则未见报道。

人工神经网络 (简称神经网络) 是一种基于生理学上的神经系统的理论抽象模型, 它是由大量被称为神经元的简要信息处理单元通过高度并联、互联而成的非线性动力学网络系统, 具有很强的自学习、自组织、自适应及容错性等特点, 善于联想、综合和推广^[6,7], 且特别适用于

非线性问题的处理, 并已在农业病虫害发生趋势预测方面取得了良好的效果^[8,9]。因此, 本文利用本县 1983 年至 2002 年的历年马尾松毛虫调查资料, 根据神经网络的原理和方法来研究马尾松毛虫的有虫面积以及有虫面积的虫口密度、虫株率与气象因子之间的关系, 分别建立马尾松毛虫有虫面积、虫口密度、虫株率的预测模型, 为人工神经网络理论在森林害虫预测预报中的应用进行积极的探索。

1 材料与方法

1.1 材料来源

马尾松毛虫资料来自浙江省仙居县森林病虫害防治站, 气象资料来自浙江省仙居县气象局。该县于 1982 年发现有马尾松毛虫危害, 自 1983 年起, 每年均在相对固定的时间内调查各代虫情的危害情况。

1.2 预报因子的选择

预报因子的选择是发生量预测结果正确与否的关键。马尾松毛虫在本县 1 年发生 2 代, 由于越冬(后)代马尾松毛虫对马尾松林当年生长威胁性大, 是生产中防治的重点, 故将越冬(后)代马尾松毛虫的有虫面积、虫口密度、虫株率作为因变量。从上年 4 月开始至当年 3 月的旬平均气温、旬降雨量、旬平均相对湿度等 8 个气象因子作为预报因子, 即自变量。因自变量个数太多(共计 288 个), 因此分 2 步选取预报因子。第 1 步, 先用相关系数法选取, 凡自变量与因变量的相关系数大于 $R_{0.9}$ 的即选入。第 2 步, 将相关系数法选出的自变量与因变量进行逐步回归, 选出与因变量相关最密切的气象因子作为预报因子。

1.2 BP 算法^[6]

BP(Back Propagation)算法是一种采用误差反向传播算法的人工神经网络模型。它由一个输入层, 一个输出层和一个或多个隐层组成。每层有若干个互不连接的神经元节点, 各相邻层的节点之间单方向互联。这个算法的学习过程, 由正向传播和反向传播组成。在正向传播过程中, 一个输入信息从输入层经隐单元层逐层处理, 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, 在输出层得到网络输出。如果网络输出与期望输出有偏差, 则转入反向传播, 将误差信号沿原来的连接通路返回, 通过修改各层神经元的权值, 使网络输出与期望输出尽量一致。通过反复的修改即学习训练后, 使得误差信号最小。

BP 算法的步骤可概述如下^[10~13]:

(1) 权值和阈值初始化: 随机地给输入层到隐含层之间的连接权值 $W_{l,m}^1$ 、隐含层到输出层之间的连接权值 $W_{m,n}^2$ 和隐含层神经元的阈值 θ_m 、输出层神经元的阈值 Ψ_n 赋以初始值, l 为输入层第 l 个因子, m 为隐含层第 m 个神经元, n 为输出层第 n 个神经元。

(2) 从学习样本集 x 中取出一个样本 x_i , 将其信号输入网络学习训练, 并指明其期望输出 \hat{a}_k 。

(3) 信号向前传播, 经输入层到隐含层之间 S 型(sigmoid)作用函数

$$F(x) = 1 / [1 + \exp(-x)] \quad (1)$$

运算后, 得到隐节点的输出 a_j :

$$a_j = f(\sum W_{l,m}^1 x_j - \theta_m) \quad (2)$$

然后经隐层到输出层节点的线性激励函数作用后, 得到网络输出节点的输出 a_k :

$$a_k = f(\sum W_{m,n}^2 a_j - \Psi_n) \quad (3)$$

(4) 计算输出节点的输出误差:

$$\delta_k = a_k(1 - a_k)(\hat{a}_k - a_k) \quad (4)$$

(5) 计算隐节点的误差:

$$E_j = a_j(1 - a_j)(\sum \delta_k W_{kj}^2) \quad (5)$$

(6) 按下述递推公式向减少误差方向修正权值和阈值:

$$W_{m+1,n}^2 = W_{m,n}^2 + \alpha \delta_k a_j \quad (6)$$

$$W_{l+1,m}^1 = W_{l,m}^1 + \beta E_j x_i \quad (7)$$

$$\theta_{m+1} = \theta_m + \alpha \delta_k \quad (8)$$

$$\Psi_{n+1} = \Psi_n + \beta E_j \quad (9)$$

以上诸式中, α 、 β 为在 (0, 1) 之间取值的学习率。

(7) 依次输入学习样本, 重复上述步骤, 直到全部 N 个样本学完一遍后, 计算均方误差:

$$R = \sum (a_j - \hat{a}_j)^2 / N \quad (10)$$

若 $R \leq$ 指定精度 ε 或达到循环次数要求, 则输出结果。否则, 更新学习次数, 返回到步骤 (2), 将样本重新输入往复学习, 直到式 (10) 满足收敛条件为止。

2 结果与分析

2.1 预报因子的确定

2.1.1 有虫面积预报因子的确定 利用相关系数法共筛选出 29 个气象因子, 然后进行逐步回归, 最终入选 8 个气象因子: 当年 2 月下旬平均气温 (X_1); 当年 3 月上旬平均气温 (X_2); 当年 2 月中旬相对湿度 (X_3); 上年 10 月中旬降水量 (X_4); 上年 9 月中旬日照时数 (X_5); 上年 6 月上旬最低气温 (X_6); 上年 8 月中旬最低气温 (X_7), 上年 12 月中旬最低气温 (X_8)。历年的有虫面积及预报因子详见表 1。

2.1.2 虫口密度预报因子的确定 利用相关系数法共筛选出 17 个气象因子, 然后进行逐步回归, 最终入选 6 个气象因子: 当年 3 月下旬降水量 (X_{11}); 上年 4 月下旬日照时数 (X_{12}); 上年 5 月上旬相对湿度 (X_{13}); 上年 9 月中旬相对湿度 (X_{14}); 上年 9 月中旬极端最高气温 (X_{15}); 上年 10 月上旬极端最高气温 (X_{16})。历年的预报因子及虫口密度详见表 2。

2.1.3 虫株率预报因子的确定 利用相关系数法共筛选出 19 个气象因子, 然后进行逐步回归, 最终入选 5 个气象因子: 当年 3 月上旬极端最低气温 (X_{21}); 上年 8 月中旬降水量 (X_{22}); 上年 4 月下旬日照时数 (X_{23}); 上年 5 月上旬相对湿度 (X_{24}); 上年 9 月中旬极端最高气温 (X_{25})。历年的虫株率及预报因子详见表 3。

2.2 网络隐层神经元个数的确定

隐层神经元个数的选择是一个十分复杂的问题, 不同的研究者有不同的观点^[6, 10-14]。结合网络的泛化功能, 经多次拟合和筛选, 本文有虫面积 BP 网络的隐层神经元数最终确定为 8 个, 虫口密度 BP 网络的隐层神经元数最终确定为 8 个, 虫株率 BP 网络的隐层神经元数最终确定为 4 个。

表 1 各年份的预报因子与有虫面积的实际值、理论值和预测检验结果

年份	气象预报因子							有虫面积实际值/ ($\times 10^3 \text{ hm}^2$)	有虫面积理论值/ ($\times 10^3 \text{ hm}^2$)	误差	
	$X_1/^\circ\text{C}$	$X_2/^\circ\text{C}$	$X_3/\%$	X_4/mm	X_5/h	$X_6/^\circ\text{C}$	$X_7/^\circ\text{C}$				$X_8/^\circ\text{C}$
1983	7.0	9.3	73	0.1	27.6	12.8	23.7	-3.5	4.333	4.333	0
1984	5.0	7.6	83	44.7	41.5	17.2	20.8	-3.3	4.826	4.826	0
1985	7.0	7.6	83	61.3	24.1	19.2	21.1	0.4	4.210	4.210	0
1986	5.4	7.7	78	26.5	71.1	15.5	23.3	-4.7	3.106	3.106	0
1987	6.2	11.1	84	46.3	35.4	15.4	22.5	-4.1	3.205	3.205	0
1988	7.8	5.5	85	9.6	74.8	11.9	20.4	-2.9	4.018	4.018	0
1989	8.4	8.1	81	0.1	52.7	15.3	22.9	-5.9	4.310	4.310	0
1990	7.2	8.9	88	14.3	24.1	16.0	19.6	-2.0	4.178	4.178	0
1991	8.4	11.6	75	2.6	42.9	17.9	23.1	1.5	0.579	0.579	0
1992	10.7	9.9	80	3.5	61.2	21.4	20.4	0.0	0.913	0.835	0.085 4
1993	7.2	9.2	79	0.0	37.7	15.1	22.3	-1.9	2.137	2.137	0
1994	7.4	9.5	84	11.9	61.8	15.5	23.9	-1.9	1.088	1.089	0.000 9
1995	7.9	11.0	85	7.6	39.1	16.6	22.1	0.7	0.748	0.748	0
1996	3.7	10.3	68	4.0	52.7	16.1	23.0	0.5	0.756	0.834	0.103 2
1997	13.1	13.5	62	0.0	71.8	19.3	23.3	-2.5	0.453	0.433	0.044 2
1998	10.0	10.8	86	1.4	70.5	17.7	21.8	-0.8	0.171	0.171	0
1999	10.2	12.5	64	9.3	61.2	18.1	23.8	0.3	0.413	0.433	0.048 4
2000	6.1	9.9	78	0.4	60.8	19.1	23.2	-2.3	0.747	0.747	0
(平均值)											0.015 7
2001 ^①	11.7	10.9	75	16.1	54.8	16.7	23.2	2.2	1.052	0.834	0.207 2
2002 ^①	13.1	12.9	71	0.0	34.1	16.1	21.9	-2.7	1.175	1.164	0.009 4

注: ①代表预测样本,下同。

表 2 各年份的预报因子与虫口密度的实际值、理论值和预测检验结果

年份	气象预报因子					虫口密度实际 值/(条·株 ⁻¹)	虫口密度理论 值/(条·株 ⁻¹)	误差	
	X_{11}/mm	X_{12}/h	$X_{13}/\%$	$X_{14}/\%$	$X_{15}/^\circ\text{C}$				$X_{16}/^\circ\text{C}$
1983	37.2	54.8	82	88	31.6	33.8	3.61	3.614 2	0.001 2
1984	31.0	52.2	84	88	33.3	35.6	3.11	3.110 0	0
1985	29.6	62.6	76	88	27.9	34.0	4.29	4.288 1	0.000 4
1986	73.6	81.7	84	85	35.0	31.9	2.93	2.930 2	0.000 1
1987	57.0	56.0	80	84	34.8	28.8	1.16	1.160 0	0
1988	95.7	63.0	82	82	29.8	32.8	2.37	2.370 4	0.000 2
1989	23.1	64.5	82	83	30.2	28.3	3.42	3.417 5	0.000 7
1990	50.5	33.3	82	90	31.3	27.9	1.38	1.379 9	0.000 1
1991	108.6	34.7	76	86	33.1	26.8	0.13	0.245 2	0.886 2
1992	149.7	31.7	80	82	34.9	31.3	0.073	0.073 2	0.002 7
1993	55.4	41.8	79	85	35.5	33.7	0.33	0.298 7	0.094 8
1994	30.2	60.2	81	86	37.0	27.9	0.22	0.267 1	0.214 1
1995	60.8	34.9	66	83	29.2	28.2	0.33	0.245 7	0.255 5
1996	67.8	7.6	75	80	31.6	30.7	0.29	0.245 8	0.152 4
1997	44.7	67.9	73	81	34.8	30.8	0.31	0.306 6	0.011 0
1998	139.9	56.0	76	76	32.0	30.3	0.14	0.139 8	0.001 4
1999	67.2	46.1	82	79	33.0	32.0	1.34	1.339 7	0.000 2
2000	15.6	43.1	78	81	36.7	34.6	0.34	0.340 6	0.001 8
(平均值)									0.090 2
2001 ^①	45.2	34.0	70	71	31.7	34.7	0.27	0.270 6	0.002 2
2002 ^①	121.2	29.9	82	80	32.0	28.8	1.49	1.317 1	0.116 0

表 3 各年份的预报因子与虫株率的实际值、理论值和预测检验结果

年份	气象预报因子					虫株率 实际值	虫株率 理论值	误差
	$X_{21}/^{\circ}\text{C}$	X_{22}/mm	X_{23}/h	$X_{24}/\%$	$X_{25}/^{\circ}\text{C}$			
1983	- 2 1	11.0	54 8	82	31. 6	0 421 5	0. 440 5	0. 045 1
1984	- 2 3	27. 3	52 2	84	33. 3	0 385 4	0. 336 1	0. 127 9
1985	0	19.0	62 6	76	27. 9	0 485 3	0. 490 3	0. 010 3
1986	- 5 1	14. 2	81 7	84	35. 0	0 562 7	0. 555 9	0. 012 1
1987	- 0 2	4. 5	56 0	80	34. 8	0 339 0	0. 325 7	0. 039 2
1988	- 1 1	10. 2	63 0	82	29. 8	0 508 9	0. 506 9	0. 003 9
1989	- 2 0	38. 8	64 5	82	30. 2	0 437 4	0. 438 1	0. 001 6
1990	0	99.0	33 3	82	31. 3	0 312 0	0. 288 3	0. 076 0
1991	4 5	153. 2	34 7	76	33. 1	0 087 6	0. 087 1	0. 005 7
1992	2 9	14. 3	31 7	80	34. 9	0 042 6	0. 038 7	0. 091 5
1993	- 1 2	89. 4	41. 8	79	35. 5	0 235 0	0. 229 2	0. 024 7
1994	- 1 9	64. 6	60 2	81	37. 0	0 157 6	0. 150 4	0. 045 7
1995	- 1 3	26. 7	34 9	66	29. 2	0 236 0	0. 289 2	0. 225 4
1996	- 0 9	43. 2	7. 6	75	31. 6	0 166 4	0. 167 8	0. 008 4
1997	1 3	79. 0	67. 9	73	34. 8	0 245 2	0. 268 0	0. 093 0
1998	0 6	222. 8	56 0	76	32. 0	0 142 9	0. 142 8	0. 000 7
1999	3 2	40. 2	46 1	82	33. 0	0 234 2	0. 242 1	0. 033 7
2000	4 0	78. 2	43 1	78	36. 7	0 103 3	0. 101 1	0. 021 3
(平均值)								0. 048 1
2001 ^①	0 1	87. 8	34 0	70	31. 7	0 231 3	0. 279 4	0. 208 0
2002 ^①	1 4	37. 3	29 9	82	32. 0	0 179 6	0. 178 7	0. 005 0

2.3 网络模拟结果

在带有神经网络工具箱的 MATLAB 语言 5.3 版环境下, 建立一个含一个隐层的 BP 网络, 将表 1 中 8 个气象因子作为输入样本, 各年度的有虫面积作为目标矢量。输入层到隐层的传递函数为 tansig , 隐层到输出层的传递函数为 purelin , 学习函数为 learnqdm , 训练函数为 trainlm (各训练参数均采用默认值), 利用 $\text{network}/\text{sim}$ 函数计算网络输出。当训练了 853 次时, 平均均方误差 $MSE = 7.19136 \times 10^{-4}$ 达到最小, 于是中断网络学习, 输出 18 组学习信息的网络输出值, 结果列于表 1 中。

同理建立虫口密度的 BP 网络模型, 其训练次数为 202 次, $MSE = 1.42092 \times 10^{-3}$, 网络的模拟结果列于表 2; 虫株率模型的训练次数为 30 次, $MSE = 3.96074 \times 10^{-4}$, 网络的模拟结果列于表 3。

从表 1~3 的结果可以看出, 用人工神经网络所建立的有虫面积 BP 模型, 平均模拟误差为 0; 虫口密度模型的平均模拟误差为 9.02%; 虫株率模型的平均模拟误差为 4.81%。说明用人工神经网络方法模拟马尾松毛虫的有虫面积、虫口密度、虫株率, 其效果是非常理想的。

2.4 预报与验证

经过训练后的 BP 网络, 还具有联想学习功能(网络的泛化学习功能), 即可以预测出没有参与建模年份的有虫面积、虫口密度、虫株率。

利用 BP 网络的联想学习功能, 计算出表 1 中预留的 2 个年份有虫面积的理论值: 2001 年, $0.834 \times 10^3 \text{ hm}^2$, 预测误差为 20.72%; 2002 年, $1.164 \times 10^3 \text{ hm}^2$, 预测误差为 0.94%。两年平均

预测误差为 10.83%，预测准确率为 100%。

表 2 中预留的 2 个年份虫口密度的理论值: 2001 年, $0.2706 \text{ 条} \cdot \text{株}^{-1}$, 预测误差为 0.22%; 2002 年, $1.3171 \text{ 条} \cdot \text{株}^{-1}$, 预测误差为 11.60%。两年平均预测误差为 5.91%, 预测准确率为 100%。

表 3 中预留的 2 个年份虫株率的理论值: 2001 年, 0.2794, 预测误差为 20.80%; 2002 年, 0.1787, 预测误差为 0.50%。两年平均预测误差为 10.65%, 预测准确率为 100%。

3 小结与讨论

(1) 经计算用逐步回归预测方法进行预测, 2001 年有虫面积的预测值为 $0.6145 \times 10^3 \text{ hm}^2$, 预测误差为 41.59%; 2002 年的预测值为 $2.8977 \times 10^3 \text{ hm}^2$, 预测误差为 146.61%, 两年平均预测误差为 94.1%。虫口密度: 2001 年, 预测值为 $-0.0249 \text{ 条} \cdot \text{株}^{-1}$, 预测误差为 109.22%; 2002 年, 预测值为 $0.5953 \text{ 条} \cdot \text{株}^{-1}$, 预测误差为 60.05%, 两年平均预测误差为 84.64%。虫株率: 2001 年, 预测值为 0.1592, 预测误差为 31.17%; 2002 年, 预测值为 0.2436, 预测误差为 35.63%, 两年平均预测误差为 33.40%。由此可见虽然所选择的预报因子一样, 但由于气象因子与森林害虫的关系并非单纯的线性关系, 人工神经网络模型比逐步回归方程模型更充分地利用了每一年马尾松毛虫的有虫面积、虫口密度、虫株率与各预报因子间相互关系的信息, 所以无论是所建立的 BP 模型的拟合精度还是模型的预测精度, BP 模型均能比逐步回归模型取得令人更满意的效果。

(2) BP 网络具有自适应、抗干扰和容错性强等能力, 个别学习样本的分量偏差对网络的学习效果影响较小, 因此模型的稳定性较好。

(3) BP 网络建模精度高低与预报因子的选取有关。因此, 在选择预报因子时宜采用逐步回归法选取。

(4) 由于马尾松毛虫多发生在人工纯林区中, 人类活动对其种群变化影响较大, 如 20 世纪 80 年代过度的砍伐灌木引起松林内生态平衡的破坏; 长期采取化学、生物等防治, 改变了马尾松毛虫自然种群变动规律; 近年来的封山育林使松林的生态环境有了很大改善。所有这些使得气候对马尾松毛虫发生的影响作用受到干扰, 同时由于缺少系统的有关食料、天敌、林相、植被等因素对马尾松毛虫发生影响的资料, 因此马尾松毛虫发生量的预报精度有待今后进一步的提高。

参考文献:

- [1] 陈昌洁. 松毛虫综合管理[M]. 北京: 中国林业出版社, 1990. 345
- [2] 薛贤清. 在马尾松毛虫测报中应用逐步回归电算方法的研究[J]. 林业科学, 1984, 20(1): 42~ 49
- [3] 薛贤清, 冯晋臣, 张石新, 等. 马尾松毛虫定量测报的判别分析模型[J]. 南京林产工业学院学报, 1982(1): 134~ 153
- [4] 薛贤清, 杨么明, 史顺平, 等. 模糊聚类分析在林业害虫测报中应用的研究[J]. 广东林业科技, 1986(5): 1~ 5
- [5] 江士玲, 金崇华. 微电脑在马尾松毛虫预测预报中的应用[J]. 森林病虫害通讯, 1986(3): 23~ 24
- [6] 焦李成. 神经网络系统理论[M]. 西安: 西安电子科技大学出版社, 1996. 1~ 36
- [7] 楼顺天, 施阳. 基于 MATLAB 的系统分析与设计——神经网络[M]. 西安: 西安电子科技大学出版社, 1999. 1~ 126
- [8] 孙凡, 陈汉容. 稻瘿蚊发生趋势的 BP 网络预测模型[J]. 计算机仿真, 2000, 17(3): 61~ 63
- [9] 李祚泳, 彭荔红. 基于人工神经网络的农业病虫害预测模型及其效果检验[J]. 生态学报, 1999, 19(5): 759~ 762

- [10] 袁曾任. 人工神经网络及其应用[M]. 北京: 清华大学出版社, 1999. 118~ 130
- [11] 闻新, 周露, 王丹力, 等. MATLAB 神经网络应用设计[M]. 北京: 科学出版社, 2000
- [12] 吴简彤, 王建华. 神经网络技术及其应用[M]. 哈尔滨: 哈尔滨工业大学出版社, 1998
- [13] 王科俊, 王克成. 神经网络建模、预报与控制[M]. 哈尔滨: 哈尔滨工业大学出版社, 1996
- [14] 丛爽. 面向 MATLAB 工具箱的神经网络理论与应用[M]. 合肥: 中国科学技术大学出版社, 1998. 59
- [15] 郎奎健, 唐守正. IBM-PC 系列程序集[M]. 北京: 中国林业出版社, 1989

A Study on the Forecast Model of *Dendrolimus punctatus* Occurrence Based on Artificial Neural Network

CHEN Hui-hua¹, ZHU Shou-yan², CUI Xiangfu¹

(1. Xianju Forestry Bureau of Zhejiang Province, Xianju 317300, Zhejiang, China;

2. Xianju Meteorological Bureau of Zhejiang Province, Xianju 317300, Zhejiang, China)

Abstract: The principle and methodology of artificial neural network were used to select some meteorological factors closely correlated to the occurrence area, population density and damage rate by the methods of correlation coefficients and step regression. The BP network models of meteorological factors and occurrence area, population density and damage rate of *Dendrolimus punctatus* were established. The results showed that these BP models established have satisfied fitting and forecast precision. When the amount of implicit layer neuron is 15 and the amount of forecast factor is 8, the mean error of forecast of 2 groups of reserved occurrence zone was 3.15% in two years. When the amount of implicit layer neuron is 8 and the amount of forecast factor is 6, the mean error of forecast of reserved occurrence zone was 5.91%, while when the amount of implicit layer neuron is 4 and the amount of forecast factor is 5, the mean error of forecast of reserved occurrence zone was 10.65%.

Key words: *Dendrolimus punctatus*; artificial neural network; occurrence; forecast; BP model