

## 马尾松毛虫幼虫发生严重程度的预测研究

余 燕<sup>1</sup>, 李 尚<sup>2</sup>, 王振兴<sup>1</sup>, 毕守东<sup>1\*</sup>, 张国庆<sup>3</sup>,  
周夏芝<sup>2</sup>, 方国飞<sup>4</sup>, 邹运鼎<sup>2</sup>, 张 桢<sup>3</sup>, 宋玉双<sup>4</sup>

(1. 安徽农业大学理学院, 合肥 230036; 2. 安徽农业大学林学与园林学院, 合肥 230036;  
3. 安徽省潜山县林业局, 潜山 246300; 4. 国家林业局森林病虫害防治总站, 沈阳 110034)

**摘 要:** 为了提高马尾松毛虫幼虫发生严重程度的预测精度, 寻求简便准确的预测方法, 采用时间平稳序列法、回归预测法、马尔科夫链法、BP神经网络法和列联表多因子多级相关分析法对安徽省潜山县 1983—2014 年的马尾松毛虫越冬代、一代和二代幼虫发生的严重程度进行预测, 研究历史符合率, 并用 2015 年和 2016 年的实际发生情况验证。结果表明, 平稳时间序列法, 列联表多因子多级相关分析法计算简便, 预测结果准确; BP神经网络法和马尔科夫链法预测结果非常准确。回归模型中以当代卵盛期卵量预测当代幼虫发生严重程度的一元回归模型的预测结果准确性高, 其余一元回归模型预测结果稍差, 多元回归模型和逐步回归模型优于一元回归模型。BP神经网络模型是一种理想的预测模型。

**关键词:** 马尾松毛虫; 预测; 平稳时间序列; BP神经网络; 马尔科夫链

中图分类号: S763.305

文献标识码: A

文章编号: 1672-352X (2017)05-0882-12

## Study on the forecasting occurrence severity degree of the *Dendrolimus punctatus* larvae

YU Yan<sup>1</sup>, LI Shang<sup>2</sup>, WANG Zhenxing<sup>1</sup>, BI Shoudong<sup>1</sup>, ZHANG Guoqing<sup>3</sup>, ZHOU Xiazhi<sup>2</sup>,  
FANG Guofei<sup>4</sup>, ZOU Yunding<sup>2</sup>, ZHANG Zhen<sup>3</sup>, SONG Yushuang<sup>4</sup>

(1. School of Science, Anhui Agricultural University, Hefei 230036;

2. School of Forestry and Landscape Architecture, Anhui Agricultural University, Hefei 230036;

3. The Bureau of Qianshan Forestry, Anhui Province, Qianshan 246300;

4. General Station of Forest Pest Management, State Forestry Administration, Shenyang 110034)

**Abstract:** In order to improve the prediction accuracy of the occurrence severity degree of the *Dendrolimus punctatus* larvae and find simple and accurate forecasting method. The method of stationary time series, regression forecast, Markov chains, BP neural network and contingency table analysis were applied to establish the prediction model of the occurrence severity degree of overwintering generation, the first generation and the second generation *Dendrolimus punctatus* larvae from 1983 to 2014 in Qianshan county of Anhui Province and it was used to study the historical coincidence rate, and then the predicted result was proved with the actual happening situation in 2015 and 2016. Result shows that: The predicted results which used the more simple calculation methods of stationary time series and contingency table were accurate; BP neural network and Markov chain method to predict the result were very accurate. The predicted result of single regression model that the contemporary amount of eggs from the egg stage predicted the occurrence severity degree of the contemporary larvae was very accurate in the regression model and other predicted result with the single regression mode was a bit poor, so multiple regression model and the stepwise regression model were better than the single regression model. BP neural network model is a kind of ideal prediction model.

**Key words:** *Dendrolimus punctatus*; forecasting; stationary time series; BP neural network; markov chain

收稿日期: 2017-05-03

基金项目: 国家林业公益性行业科研专项(201404410)资助。

作者简介: 余 燕, 硕士研究生。E-mail: 544896437@qq.com

\* 通信作者: 毕守东, 教授。E-mail: bishoudong@163.com

马尾松毛虫 (*Dendrolimus punctatus* Walker), 在中国主要分布于安徽、河南、海南、广西和台湾等 16 个省区, 主要危害马尾松 (*Pinus massoniana* Lamb)、黑松 (*Pinus thunbergii* Parl)、火炬松 (*Pinus taeda* L.)、湿地松 (*Pinus elliotii*)、晚松 (*Pinus rigida* var. *serotina*)、海南松 (*Pinus fenzeliana*) 等松属植物。20 世纪中叶在中国森林害虫中马尾松毛虫是发生最广、危害面积最大且经常猖獗成灾的害虫。在广大丘陵地区, 虫害此起彼伏, 被害时枝叶常被吃光, 如同火烧, 造成了巨大的经济效益损失和生态效益损失。人们在林事活动中接触马尾松毛虫毒毛, 容易引发皮炎和关节肿痛<sup>[1-4]</sup>。进入 21 世纪, 由于各地采用封山育林、混交和间作等措施优化了森林生态环境, 加之科学地运用综合治理措施, 使马尾松毛虫的危害得到有效控制, 但该虫具有强大的繁殖潜力, 遇到有利的生态环境极易暴发成灾, 对其监测不能放松警惕。

马尾松毛虫发生的预测预报是对其进行综合防治的基本工作。陈绘画等<sup>[5]</sup>、王庆等<sup>[6]</sup>、贾春生<sup>[7]</sup>、张爱兵等<sup>[8]</sup>、田万银等<sup>[9]</sup>、许章华等<sup>[10]</sup>、费海泽等<sup>[11]</sup>、Zhang 等<sup>[12]</sup> 和 Park 等<sup>[13]</sup> 分别采用不同的预测方法预测马尾松毛虫的发生量、虫害等级、发生类别及发生空间格局, 为马尾松毛虫的综合防治工作提供了有力支持。由于各地气象条件、植被条件和地形地貌等不同, 马尾松毛虫的发生特点也不完全相同。马尾松毛虫幼虫是危害生态, 其发生数量的多少对林业生产的影响造成不同的经济损失, 本研究将一个世代的累计幼虫数量称作发生严重程度。为了有效地防治马尾松毛虫, 采用时间平稳序列法、回归预测法、BP 神经网络模型、马尔科夫链法和列联表多因子多级分析法研究安徽省潜山县马尾松毛虫幼虫的越冬代, 一代和二代的严重程度, 并对几种方法的预测结果进行比较, 以虫期为马尾松毛虫的综合治理提供科学依据。

## 1 材料与方法

### 1.1 材料来源

马尾松毛虫资料来自安徽省潜山县森林病虫害防治站, 气象资料来自国家气象局, 资料的时间跨度为 1983—2016 年, 其中缺 1998 年数据。

### 1.2 建模方法

**1.2.1 平稳时间序列预测法** 害虫发生量的多少, 在一定时间内彼此间存在一定相关性, 在未来的生态环境变化不大的情况下, 这种关系仍然保持不变, 对此可以用随机函数来处理<sup>[14]</sup>。

对害虫的发生量  $X(t)$  的随机函数, 若相关函数  $R_x(t_1t_2)$  在一定时间的值, 与日期  $t_1$  和  $t_2$  无关, 仅与相隔日期长度  $k(t_1-t_2)$  有关系, 即  $X(t)$  为平稳随机函数, 而其序列  $X(t_1), X(t_2), \dots, X(t_m)$  为平稳时间序列:

$$R_x(t_1t_2) = R_x(k) \quad k=t_2-t_1 \quad (1)$$

相关函数  $R_x(k)$  的绝对值随着而趋于零, 相关函数可表示成:

$$R_x(k) = \lim_{k \rightarrow \infty} \frac{1}{N} \int_0^{t_0} x(t+k)x(t)dt \quad (2)$$

其中  $t_0$  代表资料时间总长度。

采用三序列预报方程组, 其预报系数为  $a_1, a_2, a_3$  由下列方程组求解

$$\left. \begin{aligned} R_0a_1 + R_1a_2 + R_2a_3 &= R_3 \\ R_1a_1 + R_0a_2 + R_1a_3 &= R_2 \\ R_2a_1 + R_1a_2 + R_0a_3 &= R_1 \end{aligned} \right\} \quad (3)$$

其中:

$$R_0 = \frac{1}{N} \sum (x_i)^2 \quad (4)$$

$$R_1 = \frac{1}{N-1} \sum (x_1x_2) + (x_2x_3) + \dots + (x_{n-1}x_n) \quad (5)$$

$$R_2 = \frac{1}{N-2} \sum (x_1x_3) + (x_2x_4) + \dots + (x_{n-2}x_n) \quad (6)$$

$$R_3 = \frac{1}{N-3} \sum (x_1x_4) + (x_2x_5) + \dots + (x_{n-3}x_n) \quad (7)$$

一般采用此法的时间跨度要有足够长, 本研究是 31 年的资料, 建立的预测模型为

$$x_t = a_1x_{t-3} + a_2x_{t-2} + a_3x_{t-1} \quad (8)$$

**1.2.2 回归预测法** 利用与马尾松毛虫生长发育有关的生态因子与马尾松毛虫幼虫发生严重程度建立回归预测模型进行预测<sup>[14]</sup>。

**1.2.3 BP 神经网络法** 神经网络的模式很多<sup>[15-22]</sup>, 目前应用最为普遍的是 BP 型神经网络 (back propagation, 即误差后向传播神经网络), 它是利用最小方差的方式把结果反馈到中间层次的隐层, 通过改变它们的权重系数, 从而达到建网的目的。图 1 是一个 M-H-N 结构的 BP 网络, 其中 M 表示输入层 X 有 M 个输入节点, H 表示中间层即隐层 Y 有 H 个节点, N 表示输出层 Z 有 N 个输出节点。设输入层和隐层之间的权为  $w_{ij}^{(1)}$ ,

隐层与输出层间的权为  $w_{jk}^{(2)}$ ,

(其中  $i=1, 2, \dots, M$ ;  $j=1, 2, \dots, H$ ;  $k=1, 2, \dots, N$ ), BP 网的工作过程如下:

(1) 初始设置各权重  $w_{ij}^{(1)}$ ,  $w_{jk}^{(2)}$  和各阈值  $\theta$

为小随机数。

(2) 隐层单元的输入和输出:

输入:

$$net_j = \sum_{i=1}^m w_{ij}^{(1)} x_i$$

输出:  $O_j=f(net_j)$ , 其中函数  $f(x)=1/(1+e^{-x})$ 。

输出单元的输入和输出:

$$net_k = \sum_{j=1}^m w_{jk}^{(2)} x_j$$

(3) 输入:

输出:  $O_k=f(net_k)$ , 其中函数  $f(x)=1/(1+e^{-x})$ 。

(4)调整权值: 设学习样本为  $(x_i, z_i), i=1, 2, \dots, p$ , 设第  $R$  个学习样本的输出与期望输出值误差为  $E_R$ , 令

$$E_R = \frac{1}{2} \sum_{k=1}^n (O_k^R - Z_k^R)^2$$

则  $P$  个学习样本的总的误差为:

$$E = \sum_{i=1}^p E_R$$

BP 学习算法就是在权重空间沿梯度调整权重, 使总误差达到最小, 满足这一步要求权的调整量为:

$$\Delta w = -\eta \frac{\partial E}{\partial W}$$

公式中,  $\eta$  为学习率。考虑第  $R$  个学习样本, 若不满足精度条件, 则网络的权值和阈值的具体调整量值如下。

对输出层而言, 其网络权值改变量:

$$\Delta_R W_{jk}^{(2)} = \eta \delta_k^R \times O_j^R$$

$$\delta_k^R = -f'(net_k^R)(Z_k^R - O_k^R)$$

其中, 阈值改变量为:

$$\Delta_R \theta_k^R = \eta \delta_k^R$$

对隐层而言, 其网络权值改变量:

$$\Delta_R W_{ij}^{(1)} = \eta \delta_j^R \times O_i^R$$

其中:

$$\delta_j^R = -f'(net_j^R) \sum_{k=1}^n \delta_k^R W_{jk}^{(2)}$$

阈值改变量为:

$$\Delta_R \theta_j^R = \eta \delta_j^R$$

按照资料每计算 1 次, 权重系数就要修正 1 次, 直到满足要求为止。

**1.2.4 马尔科夫链方法**<sup>[23]</sup> 在马尾松毛虫年度间发生序列中, 将它的发生程度分为若干有限等级。序列中每一个结果只出现有限个或可数个两两相斥事件  $E_1, E_2, \dots, E_i$ , 等状态, 同时系统由  $E_i^{(k)}$  状

态经过 1 次转移到  $E_i^{(k+1)}$  状态的概率只与前一步的  $E_i^{(k)}$  状态有关, 因此马尾松毛虫年度间发生状态可视为马尔科夫链。马尾松毛虫发生状态经过  $m$  次转移, 由  $E_i$  状态转移到  $E_j$  状态的转移概率为:

$$P_{ij}(m) = N_j(m) / m_i$$

式中:  $m$  为状态转移次数,  $m_i$  为马尾松毛虫等级为  $E_i$  的总次数;  $N_j(m)$

$E_j$  为马尾松毛虫等级由  $E_i$  转移到的次数。  $m$  阶转移概率矩阵为:

$$P(m) = \begin{pmatrix} P_{11} & P_{12} & \dots & P_{1K} \\ P_{21} & P_{22} & \dots & P_{2K} \\ \vdots & \vdots & \vdots & \vdots \\ P_{n1} & P_{n2} & \dots & P_{nK} \end{pmatrix}$$

根据预报年的前 1 年, 前 2 年和前 3 年发生严重程度的级别, 分别查状态转移矩阵  $P(1)$ 、 $P(2)$  和  $P(3)$ , 计算出可能出现级别的平均概率, 以平均概率最大的那个状态的级别作为预报年可能发生级别的预报值。

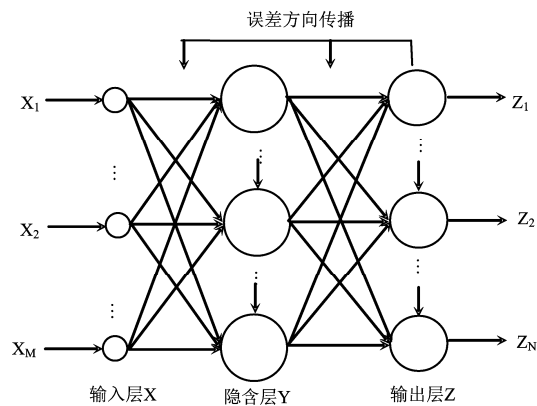


图 1 BP 网络结构原理  
Figure 1 Structure of BP neural network

**1.2.5 列联表多因子多级综合相关法分析预报** 用列联表进行害虫预报是一个比较简单的方法<sup>[23-24]</sup>。主要是如何划分等级, 同样的资料用不同的标准分级, 其预测结果相差很大, 预报量等级的划分即幼虫发生严重程度等级的划分常用等差分组法, 取历年的最高数值减去最低数值得到极差再除以分级数, 本研究分 4 级 ( $l=4$ )。

预报因子的分级即自变量等级的划分, 采用直线回归法, 首先确定自变量与预报量之间必须具有相关关系, 建立一元回归式, 将预报量分级标准的值代入一元回归式, 即可求出自变量对应的级别的值, 本研究预报量  $X_i$  也分为 4 级。

多因子多级综合相关法是一种概率相关分析法, 即对预报量  $y$  在各个预报因子  $X_i$  作用下出现的

概率做出估计, 计算一个因子在  $X_i$  处于  $k$  级相对应的  $y$  出现  $l$  级的概率有多大, 然后将多个预报因子  $X_i$  可能出现  $l$  级的概率进行综合, 估计预报对象  $y$  出现  $l$  级的概率, 进行预报。

按照上述步骤首先对预报量  $y$  分为 4 级, 预报因子  $X_i$  也分为 4 级, 用分级数值排成列联表, 然后按预报量  $y$  出现在  $l$  级的条件概率公式

$$P_{lk}^i = \frac{n_{kl}}{n_k}$$

求出其条件概率, 再将多因子的条件概率列成新的概率列联表, 由概率列联表综合得到预报年  $m$  个因子的条件概率之和  $\hat{y}_t$ ,

$$\hat{y}_t = \frac{1}{m} \sum_{i=1}^m P_{kl}^i = \frac{1}{m} \sum_{i=1}^m n_{kl}^i$$

取  $\hat{y}_t$  最大值的  $l$  作为预报  $y$  的发生量级别。

上述计算工作用 DPS 软件进行。

## 2 结果与分析

### 2.1 用平稳时间序列法预测各代幼虫发生严重程度

**2.1.1 越冬代幼虫发生严重程度预测模型** 将各代幼虫发生严重程度分级值列于表 1。利用式 (4)、(5)、(6) 和 (7) 求得越冬代的  $R_0=5.5313$ ,  $R_1=5.1613$ ,  $R_2=4.4667$ ,  $R_3=4.1379$ , 代入式 (4)、(5)、(6) 和 (7), 列方程组求得  $a_1=0.8358$ ,  $a_2=-1.6492$ ,  $a_3=1.7971$ 。

将 2012、2013 和 2014 的越冬代的分级值代入上式,  $X_{2015}=0.9837$ , 约等于 1, 则预测 2015 年越冬代幼虫发生严重程度为 1 级。

将 2013、2014 和 2015 年的越冬代的分级值代入上式, 预测 2016 年越冬代的发生级别, 则  $X_{2016}=0.9837$ , 约等于 1, 即预测为 1 级。

表 1 各代幼虫虫口分级值

Table 1 The value of grading for the population of each generation of larvae

世代 Generation	1983	1984	1985	1986	1987	1988	1989	1990	1991
a	3	4	4	3	2	4	4	4	2
b	3	3	4	4	4	2	2	4	1
c	4	2	3	4	4	2	4	2	1
世代 Generation	1992	1993	1994	1995	1996	1997	1998	1999	2000
a	1	1	2	2	4	4	2	1	1
b	1	1	4	3	4	1	2	1	1
c	1	4	4	3	4	1	2	1	1
世代 Generation	2001	2002	2003	2004	2005	2006	2007	2008	2009
a	1	1	2	1	1	1	1	1	1
b	1	2	2	1	1	1	1	1	1
c	1	2	2	3	1	1	1	1	1
世代 Generation	2010	2011	2012	2013	2014	2015	2016		
a	2	2	1	1	1	1 (1)	1 (1)		
b	4	2	1	1	1	1 (1)	1 (1)		
c	4	2	1	1	1	1 (1)	1 (1)		

注: 一代和二代累计虫口的分级标准均为  $y < 14$  为 1 级,  $14 \leq y < 19$  为 2 级,  $19 \leq y < 24$  为 3 级,  $y > 24$  为 4 级, 括号内为预测值。a: 越冬代幼虫发生严重程度 (头·株<sup>-1</sup>), b: 一代幼虫发生严重程度 (头·株<sup>-1</sup>) c: 二代幼虫发生严重程度 (头·株<sup>-1</sup>) 则越冬代幼虫发生严重程度的预测模型为:  $X_t = 0.8358X_{t-3} - 1.6492X_{t-2} + 1.7971X_{t-1}$

Note: The grading standard of the generation and the second generation: level 1:  $y < 14$  level 2:  $14 \leq y < 19$ , level 3:  $19 \leq y < 24$ , level 4:  $y > 24$ , Parentheses is predictive value. a: The severity of the larvae of the overwintering generation (head/strain); b: The severity of a generation of larvae(head/strain); c: The severity of the second generation of larvae(head/strain); A prediction model of the severity of overwintering generation:  $X_t = 0.8358X_{t-3} - 1.6492X_{t-2} + 1.7971X_{t-1}$

2015 年和 2016 年越冬代幼虫实际发生量为 6.2 头·株<sup>-1</sup> 和 7.1 头·株<sup>-1</sup>, 均为 1 级, 预测值和实际值完全一致。

**2.1.2 一代幼虫发生严重程度预测模型** 利用式 (4)、(5)、(6) 和 (7) 求得以及累计幼虫虫口的  $R_0=5.2188$ ,  $R_1=4.4516$ ,  $R_2=4.333$ ,  $R_3=3.8966$ , 代

入(3)式,求得  $a_1=-0.0693$ ,  $a_2=0.4139$ ,  $a_3=0.5575$ , 则预测一代幼虫发生严重程度的模型为  $X_t = -0.0693X_{t-3} + 0.4139X_{t-2} + 0.5575X_{t-1}$ 。

将2012、2013年和2014年的一代累计幼虫虫口的分级值代入上式,则  $X_{2015}=0.9021 \approx 1$ 。

将2013、2014年和2015年的一代累计幼虫出口的分级值代入上式,求2016年一代累计虫口,则  $X_{2016}=0.9031 \approx 1$ 。2015年和2016年一代累计幼虫虫口实际为6.3头·株<sup>-1</sup>和6.6头·株<sup>-1</sup>,即均为一级,预测结果与实际值符合。

**2.1.3 二代幼虫发生严重程度的预测模型** 利用式(4)、(5)、(6)和(7)求得二代累计幼虫虫口的  $R_0=6.1563$ ,  $R_1=2.1935$ ,  $R_2=4.9313$  和  $R_3=4.7241$ , 代入(3)式,求得  $a_1=0.1566$ ,  $a_2=0.2235$ ,  $a_3=0.5344$ , 则二代幼虫发生严重程度的预测模型为  $X_t = 0.1566X_{t-3} + 0.2235X_{t-2} + 0.5344X_{t-1}$ 。

将表1中2012、2013和2014年二代累计幼虫虫口分级值代入上式,2015年二代累计幼虫虫口  $X_{2015}=0.9145 \approx 1$ 。

将表1中2013、2014和2015年二代累计幼虫虫口分级值代入上式,预测2016年二代累计幼虫虫口,  $X_{2016}=0.9145 \approx 1$ 。

预测的2015、2016年2代累计幼虫虫口均为一级,实际发生量为7.4头·株<sup>-1</sup>和8.1头·株<sup>-1</sup>,即均为一级,两者完全吻合。

**2.2 预测各代累计幼虫虫口的回归预测模型**

将1983—2014年的马尾松毛虫各代幼虫累计虫口与各种因子进行回归分析,2015年和2016年作为验证材料。

**2.2.1 预测越冬代累计幼虫虫口  $y_1$  的预测模型** (1) 越冬代累计幼虫虫口  $y_1$  与4月份均温 ( $x_1$ ) 的关系  $y_1 = 91.0929 - 4.2973x_1$   $r = -0.4569^{**}$  ( $df=29$ ,  $r_{0.05}=0.355$ ,  $r_{0.01}=0.456$ )

(2) 越冬代累计幼虫虫口 ( $y_1$ ) 与去年二代马尾松毛虫卵的赤眼蜂寄生率 ( $x_2$ ) 的关系  $y_1 = 36.9981 - 1.1721x_2$   $r = -0.6265^{**}$  ( $df=28$ ,  $r_{0.05}=0.361$ ,  $r_{0.01}=0.463$ )

(3) 越冬代累计幼虫虫口  $y_1$  与头年底二代累计幼虫虫口 ( $x_3$ ) 的关系  $y_1 = 3.4928 + 0.6967x_3$   $r = 0.6426^{**}$

(4) 越冬代累计幼虫虫口  $y_1$  与二月下旬均温 ( $x_4$ ) 的关系  $y_1 = 32.2077 - 1.9394x_4$   $r = -0.4569^{**}$

(5) 多元回归预测模型,选择与  $y_1$  相关系数大的  $x_2$  和  $x_3$  与  $y_1$  之间的多元回归模型  $y_1 = 20.4621 - 0.7853x_2 + 0.4857x_3$   $R = 0.7424^{**}$  ( $df=27$ ,  $r_{0.01}=0.367$ ,  $r_{0.05}=0.470$ )

(6) 逐步回归预测模型  $y_1 = 73.9075 - 3.3608x_1 + 0.5252x_3 - 1.3627x_4$   $R = 0.7831$  ( $df=27$ )

**2.2.2 一代累计幼虫  $y_2$  虫口的预测模型** (1) 一代累计幼虫虫口  $y_2$  与一代马尾松毛虫卵的赤眼蜂寄生率 ( $x_5$ ) 的关系模型  $y_2 = 42.1857 - 1.8571x_5$   $r = -0.6089^{**}$  ( $df=29$ ,  $r_{0.05}=0.355$ ,  $r_{0.01}=0.456$ )

(2) 一代幼虫累计虫口  $y_2$  与一代卵盛期卵量 ( $x_6$ ) 的关系模型  $y_2 = 0.0009 + 0.8269x_6$   $r = 0.9728^{**}$

(3) 一代幼虫累计虫口  $y_2$  与一代幼虫高峰期发生量 ( $x_7$ ) 的关系  $y_2 = 0.2834 + 3.0316x_7$   $r = 0.9938^{**}$

(4) 一代幼虫累计虫口  $y_2$  的多元回归预测模型:选择与  $y_2$  相关系数大的  $x_6$  和  $x_7$  建立的多元回归模型是:

$$y_2 = 0.0008 + 0.8271x_6 - 0.0008x_7 \quad R = 1$$

(5) 一代幼虫累计虫口  $y_2$  的逐步回归预测模型  $y_2 = 0.0008 + 0.8271x_6 - 0.0008x_7$   $R = 1$

表2 越冬代幼虫累计虫口实际值与预测值

Table 2 The prediction and realization results of the population of the overwintering generation larvae

年份 Year	项目 Item	预报因子 Forecast factor						备注 Remark
		$X_1$	$X_2$	$X_3$	$X_4$	多元回归式	逐步回归式	
2015	预测值 Predicted value	19.5429	11.0947	8.4394	18.3167	6.5554	11.9188	2015年实际值为6.2头·株 <sup>-1</sup>
	误差 Error	13.3429	4.8947	2.2394	12.1164	0.3554	5.7188	
	准确率排序 Accuracy ranking	6	3	2	5	1	4	
2016	预测值 Predicted value	15.3745	9.4538	8.6484	14.7531	5.6017	6.3124	2016年实际值为7.1头·株 <sup>-1</sup>
	误差 Error	8.2745	2.3538	1.5484	7.6531	-1.4983	-0.79	
	准确率排序 Accuracy ranking	6	4	3	5	2	1	

**2.2.3 二代累计幼虫虫口  $y_3$  的预测模型** (1) 二代幼虫累计虫口  $y_3$  与二代马尾松毛虫卵的赤眼蜂

寄生率 ( $x_8$ ) 的关系  $y_3 = 34.4073 - 0.8354x_8$   $r = -0.4782^{**}$  ( $df=29$ ,  $r_{0.05}=0.355$ ,

$r_{0.01}=0.456$ )

(2) 二代幼虫累计虫口  $y_3$  与一代幼虫累计虫口 ( $x_9$ ) 的关系

$y_3=8.1925+0.6802x_8$   $r=0.7404^{**}$  ( $df=29$ ,  $r_{0.05}=0.355$ ,  $r_{0.01}=0.456$ )

(3) 二代幼虫累计虫口  $y_3$  与二代卵盛期卵量 ( $x_{10}$ ) 的关系

$y_3=1.4240+2.2324x_{10}$   $r=0.9673^{**}$  ( $df=29$ ,  $r_{0.05}=0.355$ ,  $r_{0.01}=0.456$ )

(4) 二代幼虫累计虫口  $y_3$  与二代幼虫高峰期

发生量 ( $x_{11}$ ) 的关系

$y_3=3.1057+3.0236x_{11}$   $r=0.9632^{**}$  ( $df=29$ ,  $r_{0.05}=0.355$ ,  $r_{0.01}=0.456$ )

(5) 多元回归预测模型: 选择与  $y_3$  相关系数大的  $x_9$ ,  $x_{10}$  和  $x_{11}$  建立的多元回归模型是:

$y_3=2.6560+0.0113x_9-0.1193x_{10}+3.2647x_{11}$   $R=0.9929$

**2.2.4 逐步回归预测模型**  $y_3=4.1958-0.0954x_8-0.0165x_9+3.0823x_{11}$   $R=0.9939$

为了比较各预测模型的预测结果与实际值的误差, 将其列于表 2~表 4。

表 3 一代幼虫累计虫口实际值与预测值

Table 3 The prediction and realization results of the population of the first generation larvae

年份 Year	项目 Item	预报因子 Forecast factor					备注 Remark
		$X_5$	$X_6$	$X_7$	多元回归式 Multiple regression	逐步回归式 Stepwise regression	
2015	预测值 Predicted value	10.0579	6.30	6.81	6.3016	6.3016	2015 年实际值为 6.3 头·株 <sup>-1</sup>
	误差 Error	3.7579	0	0.51	0.0016	0.0016	
	准确率排序 Accuracy ranking	1					
2016	预测值 Predicted value	10.807	6.6	7.104	6.5993	6.5993	2016 年实际值为 6.6 头·株 <sup>-1</sup>
	误差 Error	4.2007	0	0.5045	0.0007	0.0007	
	准确率排序 Accuracy ranking	1					

表 4 二代幼虫累计虫口实际值与预测值

Table 4 The prediction and realization results of the population of the second generation larvae

年份 Year	项目 Item	预报因子 Forecast factor						备注 Remark
		$X_8$	$X_9$	$X_{10}$	$X_{11}$	多元回归式 Multiple regression	逐步回归式 Stepwise regression	
2015	预测值 Predicted value	14.7754	12.4778	8.2328	8.8808	8.6001	7.7371	2015 年实际值为 7.4 头·株 <sup>-1</sup>
	误差 Error	7.3754	5.0778	0.8328	1.4808	1.2001	0.3371	
	准确率排序 Accuracy ranking	6	5	2	4	3	1	
2016	预测值 Predicted value	10.6819	12.6818	9.4160	9.8186	9.6164	8.6253	2016 年实际值为 8.1 头·株 <sup>-1</sup>
	误差 Error	2.5819	4.5818	1.3160	1.7786	1.5164	0.5253	
	准确率排序 Accuracy ranking	5	6	2	4	3	1	

可看出越冬代的多元回归模型以头年二代幼虫累计虫口为自变量的一元回归模型预测结果误差较小。一代的多元回归模型、逐步回归模型和一代卵盛期卵量为自变量的一元回归模型预测结果误差较小。二代的逐步回归模型、多元回归模型和二代卵盛期卵量为自变量的一元回归模型预测结果误差较小。从一、二代预测结果分析, 以当代卵盛期卵量为自变量预测当代幼虫累计虫口准确性最高。

### 2.3 BP 神经网络预测模型

**2.3.1 越冬代幼虫发生严重程度的 BP 神经网络预测模型** 对 1984—2014 年的资料作为学习样本进行分析, 2015 年和 2016 年资料作为预测样本检验预测准确性。利用相关系数法从 85 个自变量中最终筛选

出 2 月下旬均温, 4 月均温, 头年二代马尾松毛虫卵的赤眼蜂寄生率和头年二代幼虫虫口 4 个因子作为输入样本, 采用 4-4-1 的 BP 神经网络模型, 通过本系统运算出越冬代幼虫发生严重程度的数值, 预测结果列于表 5, 2015 年和 2016 年实际发生值为 6.20 头·株<sup>-1</sup> 和 7.10 头·株<sup>-1</sup>, 预测值为 7.66 头·株<sup>-1</sup> 和 7.36 头·株<sup>-1</sup>, 误差为 1.46 头·株<sup>-1</sup> 和 0.26 头·株<sup>-1</sup>。若以 4 头·株<sup>-1</sup> 为误差标准, 则 1983 年—2014 年预测结果的历史符合率为 86.67%。若按表中  $y_1$  的分级标准的级差为 5 头·株<sup>-1</sup>, 则历史符合率为 100%。

**2.3.2 一代幼虫发生严重程度的 BP 神经网络的模型** 将 1983—2014 年的资料作为学习样本进行分析, 2015 年和 2016 年资料作为预测样本检验预测结

果的准确性。利用相关系数法从 89 个自变量因子筛选出一代马尾松毛虫卵的赤眼蜂寄生率。一代卵盛期卵量和越冬代幼虫累计虫口 3 个因子作为输入样本, 采用 3-3-1 的 BP 神经网络模型。通过本系统运算可输出一代幼虫发生严重程度值。预测结果见表 5, 2015 年和 2016 年的实际发生值为 7.4 头·株<sup>-1</sup>和

8.1 头·株<sup>-1</sup>, 预测值为 10.15 头·株<sup>-1</sup>和 10.61 头·株<sup>-1</sup>, 误差为 2.75 头·株<sup>-1</sup>和 2.51 头·株<sup>-1</sup>。以 4 头·株<sup>-1</sup>为误差标准, 1983—2014 年预测结果的历史符合率为 96.77%。若按表中  $y_2$  的分级标准, 级差为 5 头·株<sup>-1</sup>, 历史符合率为 100%。

表 5 马尾松毛虫幼虫发生严重程度的实际值和预测值

Table 5 The prediction and realization results of the serious occurrence degree of *Dendrolimus* larve

项目 Item		1983	1984	1985	1986	1987	1988	1989	1990	1991
越冬代 Overwintering generation	实际值		32.3	34.4	19.2	18.7	36.3	42.5	37.3	18.5
	预测值		31.18	42.49	19.36	21.08	37.04	35.18	36.18	20.24
	误差		1.13	-8.09	-0.16	-2.38	-0.74	7.32	1.12	-1.74
一代 First generation	实际值	34.0	18.2	19.3	40.4	42.6	16.2	34.1	18.4	13.2
	预测值	36.61	18.26	20.81	40.32	42.66	13.70	34.61	16.21	12.28
	误差	-2.61	-0.06	-1.51	0.08	-0.06	2.50	-0.51	2.19	0.92
二代 Second generation	实际值	34.0	18.2	19.3	40.4	42.6	16.2	34.1	18.4	13.2
	预测值	36.61	18.26	20.81	40.32	42.66	13.70	34.61	16.21	12.28
	误差	-2.61	-0.06	-1.51	0.08	-0.06	2.50	-0.51	2.19	0.92
项目 Item		1992	1993	1994	1995	1996	1997	1999	2000	2001
越冬代 Overwintering generation	实际值	12.2	7.8	18.6	18.2	38.6	42.1	6.2	7.6	6.2
	预测值	6.21	11.56	18.57	17.34	40.98	41.15	7.29	7.05	7.82
	误差	5.99	-3.76	0.03	0.86	-2.38	0.95	-1.09	0.55	-1.62
一代 First generation	实际值	11.8	26.5	40.4	19.6	52.6	12.4	11.6	12.3	12.6
	预测值	11.28	23.75	43.10	20.05	47.33	12.07	12.04	13.19	13.21
	误差	0.52	2.75	-2.70	-0.45	5.27	0.33	-0.44	-0.89	-0.61
二代 Second generation	实际值	11.8	26.5	40.4	19.6	52.6	12.4	11.6	12.3	12.6
	预测值	11.28	23.75	43.10	20.05	47.33	12.07	12.04	13.19	13.21
	误差	0.52	2.75	-2.70	-0.45	5.27	0.33	-0.44	-0.89	-0.61
项目 Item		2002	2003	2004	2005	2006	2007	2008	2009	2010
越冬代 Overwintering generation	实际值	7.1	15.7	7.6	6.4	6.6	6.4	6.6	8.4	16.2
	预测值	7.36	9.51	6.46	7.32	7.00	6.78	6.93	7.19	15.58
	误差	-0.26	6.19	1.14	-0.92	-0.40	-0.38	-0.33	1.21	0.62
一代 First generation	实际值	16.4	18.8	22.6	12.6	11.8	12.1	13.2	13.8	26.8
	预测值	17.22	20.89	19.18	12.45	12.36	11.96	13.07	12.81	28.31
	误差	-0.82	-2.09	3.42	0.15	-0.56	0.14	0.13	0.99	-1.51
二代 Second generation	实际值	16.4	18.8	22.6	12.6	11.8	12.1	13.2	13.8	26.8
	预测值	17.22	20.89	19.18	12.45	12.36	11.96	13.07	12.81	28.31
	误差	-0.82	-2.09	3.42	0.15	-0.56	0.14	0.13	0.99	-1.51
项目 Item		2011	2012	2013	2014	2015	2016			
越冬代 Overwintering generation	实际值	16.6	7.2	7.2	6.4	6.20	7.10			
	预测值	16.70	6.95	7.95	7.93	7.66	7.36			
	误差	-0.10	0.25	-0.75	-1.53	-1.46	-0.26			
一代 First generation	实际值	16.7	8.9	7.2	7.1	7.4	8.1			
	预测值	16.04	10.73	10.17	10.08	10.15	10.61			
	误差	0.66	-1.83	-2.97	-2.98	-2.75	-2.51			
二代 Second generation	实际值	16.7	8.9	7.2	7.1	7.4	8.1			
	预测值	16.04	10.73	10.17	10.08	10.15	10.61			
	误差	0.66	-1.83	-2.97	-2.98	-2.75	-2.51			

实际值:actual value;预测值:predicted value;误差:error

**2.3.3 二代幼虫发生严重程度的 BP 神经网络预测模型** 将 1983—2014 年的资料作为学习样本进行分析, 2015 年和 2016 年资料作为预测样本验证预测结果的准确性, 利用相关系数法从 85 个自变量最终筛选出二代马尾松毛虫卵的赤眼蜂寄生率、一代幼虫发生量、二代卵盛期卵量, 二代幼虫高峰期发生量 4 个因子作为输入样本, 采用 4-4-1 的 BP 神经网络模型。通过本系统的运算以输出二代幼虫发生严重程度的数值, 预测结果列于表 5。2015 年和 2016 年二代幼虫累计虫口实际值为 7.4 头·株<sup>-1</sup> 和 8.1 头·株<sup>-1</sup>, 预测值为 10.15 头·株<sup>-1</sup> 和 10.61 头·株<sup>-1</sup>, 误差 2.75 头·株<sup>-1</sup> 和 2.51 头·株<sup>-1</sup>。若按表 1 中  $y_3$  的分级标准级差为 5 头·株<sup>-1</sup>, 1983—2014 年误差小于 5 头·株<sup>-1</sup>, 历史符合率为 100%。

#### 2.4 幼虫发生严重程度的马尔科夫链预测模型

将 1983—2016 年马尾松毛虫越冬代, 一代和二代幼虫发生严重程度列于表 1, 分别计算出一代转移概率矩阵  $P(1)$  和高阶转移矩阵  $P(2)$  和  $P(3)$ 。

越冬代幼虫发生严重程度转移概率矩阵  $P(1)$ 、 $P(2)$  和  $P(3)$ :

$$P(1) = \begin{vmatrix} 0.786 & 0.214 & 0 & 0 \\ 0.5 & 0.25 & 0 & 0.25 \\ 0 & 0.5 & 0 & 0.5 \\ 0 & 0.286 & 0.143 & 0.571 \end{vmatrix}$$

$$P(2) = \begin{vmatrix} 0.615 & 0.385 & 0 & 0 \\ 0.625 & 0 & 0 & 0.375 \\ 0 & 0 & 0 & 1 \\ 0.286 & 0.429 & 0.143 & 0.143 \end{vmatrix}$$

$$P(3) = \begin{vmatrix} 0.583 & 0.333 & 0 & 0.083 \\ 0.5 & 0.25 & 0 & 0.25 \\ 0 & 0 & 0.5 & 0.5 \\ 0.571 & 0.286 & 0 & 0.143 \end{vmatrix}$$

预测 2015 年越冬代幼虫发生严重程度, 根据表 2012 年, 2013 年和 2014 年马尾松毛虫

越冬代发生的级别, 均为一级, 求得:  $\bar{P}_1 = 0.6613$ ,  $\bar{P}_2 = 0.3107$ ,  $\bar{P}_3 = 0$ ,  $\bar{P}_4 = 0.0277$ ,  $\bar{P}_1$  最大, 即 2015 年预测为 1 级。同样根据 2013 年, 2014 年和 2015 年越冬代马尾松毛虫发生严重程度与级值均为 1 级, 求得:  $\bar{P}_1 = 0.6613$ ,  $\bar{P}_2 = 0.3107$ ,  $\bar{P}_3 = 0$ ,  $\bar{P}_4 = 0.0277$ ,  $\bar{P}_1$  最大。预测 2016 年越冬代幼虫发生严重程度为一级, 2015 年和 2016 年的实际发生严重程度均为 1 级, 预测结果和实际值完全一致。

一代幼虫高峰期发生量的转移概率矩阵如下:

$$P(1) = \begin{vmatrix} 0.733 & 0.2 & 0 & 0.067 \\ 0.429 & 0.429 & 0 & 0.143 \\ 0 & 0 & 0.333 & 0.667 \\ 0.333 & 0.167 & 0.167 & 0.333 \end{vmatrix}$$

$$P(2) = \begin{vmatrix} 0.571 & 0.286 & 0.071 & 0.071 \\ 0.857 & 0 & 0 & 0.143 \\ 0.333 & 0 & 0 & 0.667 \\ 0.167 & 0.5 & 0 & 0.333 \end{vmatrix}$$

$$P(3) = \begin{vmatrix} 0.462 & 0.308 & 0.077 & 0.154 \\ 1 & 0 & 0 & 0 \\ 0 & 0.333 & 0 & 0.667 \\ 0.5 & 0.333 & 0 & 0.167 \end{vmatrix}$$

预测 2015 年一代幼虫发生严重程度, 根据表 1 中 2012, 2013 和 2014 年马尾松毛虫一代幼虫发生级别, 即均为 1 级, 求得  $\bar{P}_1 = 0.5887$ ,  $\bar{P}_2 = 0.2647$ ,  $\bar{P}_3 = 0.0493$ ,  $\bar{P}_4 = 0.0973$ ,  $\bar{P}_1$  最大, 即 2015 年马尾松毛虫一代幼虫发生严重程度为 1 级。同样根据表中 2013 年, 2014 年和 2015 年的马尾松毛虫幼虫发生级别预测 2016 年一代幼虫发生级别为 1 级, 2015 年和 2016 年一代幼虫实际发生值均为 1 级, 预测值和实际值完全吻合。

二代幼虫一阶和高阶转移概率矩阵  $P(1)$ ,  $P(2)$  和  $P(3)$ :

$$P(1) = \begin{vmatrix} 0.692 & 0.154 & 0 & 0.154 \\ 0.429 & 0.143 & 0.286 & 0.143 \\ 0.333 & 0 & 0 & 0.667 \\ 0.125 & 0.5 & 0.125 & 0.25 \end{vmatrix}$$

$$P(2) = \begin{vmatrix} 0.5 & 0.25 & 0 & 0.25 \\ 0.571 & 0.143 & 0.143 & 0.143 \\ 0.667 & 0 & 0 & 0.333 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{vmatrix}$$

$$P(3) = \begin{vmatrix} 0.364 & 0.273 & 0.182 & 0.182 \\ 0.714 & 0 & 0 & 0.286 \\ 0.333 & 0.667 & 0 & 0 \\ 0.5 & 0.125 & 0 & 0.375 \end{vmatrix}$$

预测 2015 年二代幼虫发生严重程度, 根据表 1 中 2012、2013 和 2014 年马尾松毛虫二代幼虫发生级别均为 1 级, 求得  $\bar{P}_1 = 0.41$ ,  $\bar{P}_2 = 0$ ,  $\bar{P}_3 = 0.045$ ,  $\bar{P}_4 = 0.045$ ,  $\bar{P}_1$  值最大, 即预测 2015 年二代幼虫发生严重程度为 1 级。同样根据表 1 中 2013, 2014 和 2015 年马尾松毛虫二代幼虫发生级别均为 1 级, 求得  $\bar{P}_1 = 0.83$ ,  $\bar{P}_2 = 0.0625$ ,  $\bar{P}_3 = 0.0225$ ,  $\bar{P}_4 = 0.085$ ,  $\bar{P}_1$  值最大, 即预测 2016 年马尾松毛虫二代幼虫发生严重程度为 1 级, 2015 年和 2016 年二代幼虫发生级别实际均为 1 级, 预测值与实际值完全一致。



表 6 预报量和预报要素分级

Table 6 The value of prediction and forecast factors of grading

年份 Year	$X_1$	$X_2$	$X_3$	$Y_1$	预测值 Predicted value	$X_5$	$X_6$	$X_7$	$Y_2$	预测值 Predicted value	$X_8$	$X_9$	$X_{10}$	$X_{11}$	$Y_3$	预测值 Predicted value
1983	4	4	4	3		4	4	3	3	4	4	3	4	4	4	4
1984	3	4	2	4	4	4	4	2	3	2	4	3	3	2	2	2
1985	1	4	2	4	4	4	4	4	4	4	4	4	3	3	3	4
1986	2	4	4	3	4	4	4	4	4	4	4	4	4	4	4	4
1987	2	4	4	2	4	3	4	4	4	4	4	4	4	4	4	4
1988	2	4	2	4	4	4	4	2	2	2	4	3	2	2	2	2
1989	1	4	4	4	4	4	4	2	2	2	4	3	4	4	4	4
1990	1	4	3	4	4	4	4	4	4	4	4	4	2	2	2	2
1991	2	3	1	2	2	3	1	1	1	1	3	1	2	1	1	1
1992	2	3	1	1	1	2	1	1	1	1	4	1	1	1	1	1
1993	1	3	4	1	1	3	1	1	1	1	4	1	3	3	4	3
1994	2	3	4	2	1	3	4	4	4	4	4	4	4	4	4	4
1995	4	3	3	2	2	2	4	2	3	2	3	3	2	3	3	2
1996	4	1	4	4	4	2	4	4	4	4	2	4	4	4	4	4
1997	2	2	1	4	4	2	1	1	1	1	3	1	1	1	1	1
1999	2	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1
2000	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2001	3	1	1	1	1	1	1	1	1	1	2	1	1	2	1	1
2002	2	2	2	1	1	1	1	1	2	1	3	2	2	2	2	1
2003	2	2	3	2	1	2	4	3	2	2	3	3	3	3	2	2
2004	2	1	4	1	1	1	1	1	1	1	2	1	3	3	3	1
2005	2	1	1	1	1	2	3	2	1	2	2	3	2	1	1	1
2006	1	1	1	1	1	2	2	1	1	1	2	2	1	1	1	1
2007	2	1	1	1	1	2	2	1	1	1	2	2	1	1	1	1
2008	2	1	1	1	1	1	2	1	1	1	2	2	2	1	1	1
2009	1	1	2	1	1	2	2	2	1	1	2	2	2	1	1	1
2010	2	1	4	2	2	2	3	2	2	2	1	3	4	4	4	4
2011	2	1	2	2	2	1	3	2	2	2	1	3	2	2	2	2
2012	2	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1
2013	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2014	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1
2015	4	1	1	1	1	4	1	1	1	1	2	1	1	1	1	1
2016				1	1	4	1	1	1	1	1	1	1	1	1	1

注:  $x_1$ .10 月上旬降雨量,  $x_2$ .头年二代卵的赤眼蜂寄生率,  $x_3$ .头年底二代幼虫累计虫口,  $y_1$ .越冬代累计虫口,  $x_5$ .一代马尾松毛虫卵的赤眼蜂寄生率,  $x_6$ .一代卵量,  $x_7$ .一代幼虫高峰期发生率,  $y_2$ .一代幼虫累计虫口,  $x_8$ .二代马尾松毛虫卵的赤眼蜂发生率,  $x_9$ .二代幼虫累计虫口,  $x_{10}$ .二代卵盛期卵差,  $x_{11}$ .二代幼虫高峰发生量,  $y_3$ .二代幼虫累计虫口。

Note:  $x_1$ .The rainfall in early October,  $x_2$ .The parasitic rate of trichogrammatid in the second generation of the egg,  $x_3$ .The second generation of larvae accumulates in the end of the first year,  $y_1$ .The accumulation of *Dendrolimus punctatus* Walker of overwintering generation,  $x_5$ .The parasitic rate of trichogrammatid of a generation of *Dendrolimus punctatus* Walker eggs,  $x_6$ .The first generation amount of *Dendrolimus punctatus* Walker eggs,  $x_7$ .The occurrence rate of a generation of larvae,  $y_2$ .A generation of larvae accumulates,  $x_8$ .The incidence rate of trichogrammatid in the second generation of *Dendrolimus punctatus* Walker,  $x_9$ .The second generation of larvae accumulates,  $x_{10}$ .The egg stage of egg quantity difference of the second generation;  $x_{11}$ .The peak amount of the second generation larvae,  $y_3$ .The second generation of larvae accumulates.

表 7 越冬代幼虫发生严重程度有关因子对  $y_1$  的条件概率列联

Table 7 The contingency table about factor of the serious occurrence degree of overwintering generation of the conditional probability of the  $y_1$

		$y_1$						$y_1$			
		1	2	3	4			1	2	3	4
$x_1$	1	0.17	0.50	0.17	0.17	$x_2$	1	0.79	0.14	0.00	0.07
	2	0.67	0.22	0.00	0.11		2	0.67	0.33	0.00	0.00
	3	0.67	0.25	0.00	0.33		3	0.40	0.40	0.00	0.20
	4	0.00	0.00	0.00	1.00		4	0.00	0.25	0.13	0.63
$x_3$	1	1.00	0.00	0.00	0.00						
	2	0.17	0.33	0.17	0.33						
	3	0.33	0.33	0.00	0.33						
	4	0.11	0.44	0.00	0.44						

表 8 一代幼虫发生严重程度的有关因子对  $y_2$  的条件概率列联

Table 8 The contingency table about factor of the occurrence severity degree of the first generation of the conditional probability of the  $y_2$

		$y_2$						$y_2$			
		1	2	3	4			1	2	3	4
$x_5$	1	0.80	0.20	0.00	0.00	$x_6$	1	0.91	0.09	0.00	0.00
	2	0.60	0.20	0.10	0.10		2	1.00	0.00	0.00	0.00
	3	0.50	0.00	0.00	0.50		3	0.33	0.67	0.00	0.00
	4	0.00	0.29	0.29	0.43		4	0.00	0.25	0.25	0.50
$x_7$	1	0.93	0.07	0.09	0.00						
	2	0.25	0.50	0.25	0.00						
	3	0.00	0.50	0.50	0.13						
	4	0.00	0.00	0.00	1.00						

表 9 二代幼虫发生严重程度的有关因子对  $y_3$  的条件概率列联

Table 9 The contingency table about factor of the serious occurrence degree of the second generation of the conditional probability of the  $Y_3$

		$y_3$						$y_3$			
		1	2	3	4			1	2	3	4
$x_8$	1	0.50	0.25	0.00	0.25	$x_9$	1	0.82	0.00	0.09	0.09
	2	0.82	0.00	0.09	0.09		2	0.80	0.20	0.00	0.00
	3	0.40	0.40	0.20	0.00		3	0.11	0.44	0.11	0.33
	4	0.09	0.27	0.09	0.54		4	0.00	0.17	0.17	0.67
$x_{10}$	1	1.00	0.00	0.00	0.00	$x_{11}$	1	1.00	0.00	0.00	0.00
	2	0.44	0.44	0.11	0.10		2	0.17	0.73	0.00	0.00
	3	0.00	0.40	0.40	0.20		3	0.00	0.20	0.60	0.20
	4	0.00	0.00	0.00	1.00		4	0.00	0.00	0.00	1.00

2.5 利用列联表分析法的马尾松毛虫幼虫发生严重程度的多因子多级分析预报模型

2.5.1 越冬代幼虫发生严重程度的预测模型 对越冬代幼虫累计虫口  $y_1$  按级层法进行分级 (见表 1),  $y_1$  的数值代入与其相关的自变量 ( $x_i$ ) 的直线回归式中, 求出  $x_i$  的分级标准, 按标准进行分级, 分级结果列于表 6, 将因子对  $y_1$  的条件概率列联表  $\hat{y}_i$  列于表 7, 根据 2015 年的  $x_i$  相应的条件概率进行计算。

2015 年:  $\hat{y}_1 = 0.6533$ ,  $\hat{y}_2 = 0.2133$ ,  $\hat{y}_3 = 0.567$ ,

$\hat{y}_4 = 0.08$

2016 年:  $\hat{y}_1 = 0.5976$ ,  $\hat{y}_2 = 0$ ,  $\hat{y}_3 = 0$ ,  $\hat{y}_4 = 0.3567$

2015 年的最大, 即预测 2015 年越冬代幼虫发生严重程度为 1 级。2016 年的也最大, 即 2016 年越冬代幼虫发生严重程度为 1 级。2015 年和 2016 年越冬代幼虫的累计虫口为 6.2 头·株<sup>-1</sup> 和 7.1 头·株<sup>-1</sup>。按照表 1 中的分级, 均为 1 级。表明两年的预测值和实际值完全吻合。统计 1983—2014 年历史符合率为 80.65%, 误报的 1986 年、1987 年、1995 年、2003

年、2010年和2011年的预测值和实际值只差1个等级。

**2.5.2 一代幼虫发生严重程度的预测模型** 一代幼虫累计虫口  $y_2$  分级结果见表1,  $y_2$  的值代入与其相关的自变量 ( $x_i$ ) 的直线回归式, 求出  $x_i$  的分级标准, 按此标准对  $x_i$  进行分级, 求得结果列于表6, 将因子对  $y_2$  的条件概率列联表于表8。

预测2015年一代幼虫发生严重程度, 表6中的  $x_5=4$ ,  $x_6=1$ ,  $x_7=1$ , 由表8计算3个因子的条件概率之和  $\hat{y}_2$

计算结果是:  $\hat{y}_1=0.6133$ ,  $\hat{y}_2=0.15$ ,  $\hat{y}_3=0.1433$ ,  $\hat{y}_4=0.1433$ 。

最大, 即预测2015年一代幼虫发生严重程度为1级, 2016年表6中的  $x_5=4$ ,  $x_6=1$ ,  $x_7=1$ , 计算3个因子的条件概率之  $\hat{y}_i$  和  $\hat{y}_1$  最大, 也是1级。2015年和2016年一代幼虫发生量为  $6.3 \text{ 头} \cdot \text{株}^{-1}$  和  $6.6 \text{ 头} \cdot \text{株}^{-1}$ 。按照表1  $\hat{y}_2$  的分级, 表明预测值和实际值完全吻合。统计1983—2014年的历史符合率为87.10%, 误报的1983年, 1984年, 2002年和2005年的预测值和实际值相差1个等级。

**2.5.3 二代幼虫发生严重程度预测模型** 二代幼虫累计虫口  $y_3$  的分级结果见表1,  $y_2$  的值代入与其相关的  $x_i$  的直线回归式, 求出  $x_i$  的分级标准, 按照此标准  $x_i$  的分级结果列于表6, 将因子对  $y_3$  的条件概率列联表于表9。

预测2015年二代幼虫发生严重程度, 表6中的  $x_8$  是2级,  $x_9$  是1级,  $x_{10}$  是1级,  $x_{11}$  是1级, 由表9计算4个因子的条件概率之和  $\hat{y}_i$ , 计算结果是  $\hat{y}_1=0.91$ ,  $\hat{y}_2=0$ ,  $\hat{y}_3=0.045$ ,  $\hat{y}_4=0.045$ 。  $\hat{y}_1$  最大, 即2015年二代幼虫发生严重程度为1级。预测2016年二代幼虫发生严重程度,  $x_8$ 、 $x_9$ 、 $x_{10}$  和  $x_{11}$  均为1级, 由表9计算4个因子的条件概率之和  $\hat{y}_i$

计算结果是:  $\hat{y}_1=0.83$ ,  $\hat{y}_2=0.0625$ ,  $\hat{y}_3=0.0225$ ,  $\hat{y}_4=0.085$ 。  $\hat{y}_1$  值最大, 即预测2016年二代幼虫发生严重程度为1级, 实际上2015年和2016年二代幼虫累计虫口为  $7.4 \text{ 头} \cdot \text{株}^{-1}$  和  $8.1 \text{ 头} \cdot \text{株}^{-1}$ 。按照表1分级标准均为1级, 所以预测结果和实际值完全吻合。统计1983—2014年的历史符合率为83.81%, 误报的1985年, 1993年, 1995年、2002年和2004年的预测值和实际值也只差1个等级。

### 3 讨论与结论

对安徽省潜山县1983—2014年马尾松毛虫幼

虫发生严重程度及其相关资料用平稳时间序列法、回归预测法、马尔科夫链法、BP神经网络法和列联表多因子多级相关分析预测法建立该县马尾松毛虫越冬代, 一代和二代幼虫发生严重程度的预测模型, 并用2015年和2016年的实际发生资料进行验证, 分析5种预测模型预测结果的准确率, 结果是:

时间序列法预测模型预测2015年和2016年越冬代, 一代和二代幼虫发生严重程度与实际值完全吻合。

回归预测法中的以当代卵盛期卵量为自变量的一元回归模型预测当代幼虫发生严重程度的准确性高, 其余一元回归模型预测结果的准确性次之, 多元回归和逐步回归模型预测结果的准确性高。

BP神经网络模型预测结果: 若按幼虫发生严重程度的分级标准其级差为  $5 \text{ 头} \cdot \text{株}^{-1}$ , 则预测结果的历史符合率为100%, 2015年和2016年的预测结果与实际完全吻合。

马尔科夫链法模型对2015年和2016年的预测结果越冬代、一代和二代幼虫发生严重程度均为1级, 实际发生也均是1级。

列联表多因子多级相关分析法预测模型对2016年和2016年的预测结果越冬代, 一代和二代幼虫发生严重程度均为1级, 与实际完全吻合。1983—2014年的历史符合率为80.65%~87.10%, 若按表1中  $y_1$  的分级标准, 级差为  $5 \text{ 头} \cdot \text{株}^{-1}$ , 误报的预测结果与实际只差一级, 历史符合率应该更高一些。

5种预测模型时间平稳序列法和列联表多因子多级相关分析法进行等级预报, 不仅历史符合率和预报准确性较高, 而且由于分级数字简单, 极易进行计算, 大大减少了运算工作量, 但是最重要的是如何划分等级, 因为同样的资料用不同的标准分级, 其预报结果相差悬殊, 列联表的预报因子必须与预报量 ( $y$ ) 相关。否则不能用直线回归法分级, 时间平稳序列法适用于符合时间平稳序列的资料。

BP神经网络模型是模拟人脑神经的活动过程而提出的一种计算方法, 是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, 有很强的自学习, 自组织, 自适应及容错性等特点, 特别适用于非线性问题的处理, 模型稳定性好。它不仅是有处理多因子的功能, 而且还有很强的非线性处理能力<sup>[25-26]</sup>, 建模过程比较简单。只要对样本数据进行学习, 模型就会自动找到最佳刻划样本数据的函数, 克服了由于刻划函数太复杂或难以想象而造成的工作量无法进行下去的缺点, 但在应用上仍有一些技术关键, 一是因子的选择,

找一些合适的因子对提交预测精度起着十分重要的作用, 再者是网络结构的选取, 三层的神经网络各层的节点数取多少为宜, 即如何筛选出最合适的网络模型, 尚需深入探讨, 总之该方法是值得推广的一种方法。回归预测法是常用的预测方法, 回归模型中自变量的选择是非常关键的工作, 对自变量的选择不光具有一定的数学知识, 更重要的是还需要依赖生物学的专业知识。

### 参考文献:

- [1] 萧刚柔. 中国森林害虫[M]. 2 版. 北京: 林业出版社, 1992: 948-953.
- [2] 侯陶谦. 中国松毛虫[M]. 北京: 科学出版社, 1987: 188-191.
- [3] 邹运鼎, 程扶玖, 查光济. 松针内含物与马尾松毛虫生存发育关系的研究[J]. 林业科学, 1990, 26(2): 142-14.
- [4] 张真, 李典谟. 马尾松毛虫暴发机制分析[J]. 林业科学, 2008, 44(1): 140-150.
- [5] 陈绘画, 王坚娅, 徐志宏. 基于响应面方法马尾松毛虫发生量混沌特性检测及其预测[J]. 东北林业大学学报, 2011, 39(9): 94-96.
- [6] 王庆, 毕猛, 杜婷, 等. 基于气象因子的马尾松毛虫发生率空间格局研究[J]. 林业科学研究, 2016, 29(2): 256-260.
- [7] 贾春生. 利用马尔可夫链方法测报马尾松毛虫发生级别[J]. 东北林业大学学报, 2006, 34(5): 21-22.
- [8] 张爱兵, 陈建, 王正军, 等. BP 网络模型和 LOGIT 模型在森林害虫测报上的应用初报[J]. 生态学报, 2001, 21(12): 2159-2156.
- [9] 田万银, 徐华潮. 浙江沿海防护林马尾松毛虫的预测预报模型[J]. 环境昆虫学报, 2012, 34(4): 401-406.
- [10] 许章华, 李聪慧, 刘健, 等. 马尾松毛虫害等级的 Fisher 判别分析[J]. 农业机械学报, 2014, 45(6): 275-283.
- [11] 费海泽, 王鸿斌, 孔祥波, 等. 马尾松毛虫发生相关气象因子筛选及预测[J]. 东北林业大学学报, 2014, 41(1): 136-140.
- [12] ZHANG W J, ZHONG X Q, LIU G H. Recognizing spatial distribution patterns of grassland insects: neural network approaches[J]. *Stoch Env Res Risk*, 2008, 22(2): 207-216.
- [13] PARK Y S, CÉRÉGHINO R, COMPIN A, et al. Applications of artificial neural networks for patterning and predicting aquatic insect species richness in running waters[J]. *Ecol Model*, 2003, 160(3): 265-280.
- [14] 邹运鼎, 王弘法. 农林昆虫生态学[M]. 合肥: 安徽科学技术出版社, 1989: 462-469.
- [15] 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.
- [16] 陈绘画, 朱寿燕, 崔相富. 基于人工神经网络马尾松毛虫发生量预测模型的研究[J]. *林业科学研究*, 2003, 16(2): 159-165.
- [17] 汪四水, 张孝羲, 张夕林. 稻纵卷叶螟发生程度的神经网络预警[J]. *生物数学学报*, 2003, 18(1): 93-97.
- [18] 汪四水, 张孝羲. 基于神经网络的稻纵卷叶螟长期预测[J]. *植物保护学报*, 2000, 27(4): 313-316.
- [19] HORNİK K, STINCHCOMBE M, WHITE H. Multilayer feedforward networks are universal approximators[J]. *Neur Net*, 1989, 2(5): 359-366.
- [20] MHASKAR H N, MICCHELLI C A. Approximation by superposition of sigmoidal and radial basis functions[J]. *Adv Appl Math*, 1992, 13(3): 350-373.
- [21] RUMELHART D E, MCCLELLAND J L, PDP Research Group. Parallel distributed processing: Explorations in the microstructures of cognition. Volume 1: Foundations[M]. Cambridge: The MIT Press, 1986.
- [22] LIPPMANN R. An introduction to computing with neural nets[J]. *IEEE Assp Magazine*, 1987, 4(2): 4-22.
- [23] 张孝羲. 昆虫生态及预测预报 [M]. 3 版. 北京: 中国农业出版社, 2006, 205-221.
- [24] 国家林业局森林病虫害防治总站. 林业有害生物监测预报技术[M]. 北京: 中国林业出版社, 2013: 117-118.
- [25] HORNİK K, STINCHCOMBE M, WHITE H. Multilayer feedforward networks are universal approximators[J]. *Neural Networks*, 1989, 2(5): 359-366.
- [26] MHASKAR H N, MICCHELLI C A. Approximation by superposition of sigmoidal and radial basis functions[J]. *Adv Appl Math*, 1992, 13(3): 350-373.