

# 马尾松毛虫幼虫高峰期发生量的 预测模型研究\*

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

(1. 安徽农业大学林学与园林学院, 合肥 230036; 2. 安徽农业大学理学院, 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** 【目的】 为了提高马尾松毛虫 *Dendrolimus punctatus* Walker 发生量预测预报结果的准确性, 为选用合适的预测模型提供依据。【方法】 本文用平稳时间序列法、回归预测法、BP 神经网络法、马尔科夫链法和列联表多因子多级分析预测法研究建立安徽省潜山县 1983—2016 年 33 年的马尾松毛虫 1 代和 2 代幼虫高峰期发生量的预测模型, 并对 5 种模型进行比较。【结果】 以卵高峰期卵量为自变量的回归模型、多元回归模型和逐步回归模型预测结果与实际值相差 0.21~0.31 头/株, 其它 8 个一元回归预测结果与实际值相差 1.06~1.58 头/株。平稳时间序列预测 2015 和 2016 年的结果与实际值完全相符。BP 神经网络预测结果若以误差标准为 1 头/株, 1983—2014 年预报准确率 1 代为 90.32%, 2 代为 100%。马尔科夫链预测 2015 和 2016 年, 预测结果与实际值完全相符, 均为 1 级。列联表多因子多级综合相关分析法预测 2015 和 2016 年的结果与实际值完全相符, 1983—2014 年 1 代幼虫高峰期发生量预测的历史符合率均为 90.32%, 2 代为 83.47%。为了研究不同分级标准对预测值的影响, 将 2 代幼虫高峰期发生量的 1 级标准改为小于 3.5 头/株, 则历史符合率为 74.19%。【结论】 上述方法中, 回归预测法自变量的选择是预报准确的关键; 时间平稳序列法适用于害虫发生过程符合平稳时间序列的标准; 马尔科夫链法和列联表分析法分级标准科学与否直接影响预测结果的准确性; BP 神经网络法可用于自变量与预报量非线性关系的研究, 是一种比较理想的预报方法。

**关键词** 马尾松毛虫幼虫高峰期发生量, 平稳时间序列预测模型, 回归预测模型, BP 神经网络预测模型, 马尔科夫链预测模型, 列联表多因子多级分析预测模型

## Comparison of models for forecasting peak abundance of *Dendrolimus punctatus* larvae

ZHOU Xia-Zhi<sup>1\*\*</sup> WANG Zhen-Xing<sup>2</sup> YU Yan<sup>2</sup> LI Shang<sup>1</sup> BI Shou-Dong<sup>2\*\*\*</sup>  
ZHANG Guo-Qing<sup>3</sup> FANG Guo-Fei<sup>4</sup> ZOU Yun-Ding<sup>1</sup> ZHANG Zhen<sup>3</sup> SONG Yu-Shuang<sup>4</sup>

(1. School of Forestry and Landscape Architecture, Anhui Agricultural University, Hefei 230036, China; 2 School of Science, Anhui Agricultural University, Hefei 230036, China; 3. The Forest of Qianshan County, Anhui Province, Qianshan 246300, China;  
4. The Forest Disease and Pests Prevention and Control Station, Shenyang 110034, China)

**Abstract** [Objectives] To improve the accuracy of forecasting the peak occurrence of *Dendrolimus punctatus*, and provide a foundation for choosing suitable predictive models. [Methods] Models based on five different methods; stationary time series, regression forecasting, BP neural networks, Markov chains, and contingency table analysis, were used to predict peaks of abundance of first and second generation *Dendrolimus punctatus* larvae over a 33 year period from 1983 to 2016 in

\*资助项目 Supported projects: 国家林业公益性行业科研专项(201404410)

\*\*第一作者 First author, E-mail: zhouxz@ahau.edu.cn

\*\*\*通讯作者 Corresponding author, E-mail: bishoudong@163.com

收稿日期 Received: 2017-04-13, 接受日期 Accepted: 2017-07-19

Qianshan County Anhui Province. **[Result]** The single regression model used peak egg abundance as an independent variable, whereas the multiple regression and stepwise regression model used the actual 0.21 to 0.31 number of larvae per strain. The predicted and actual results of the single regression model differed by 1.06 to 1.58 larvae per strain. The actual results for 2015 and 2016 were identical to the predictions of the stationary time model for those years. If the standard error of the model based on the BP neural network was 1 larvae per strain, the accuracy of the forecast first generation from 1983 to 2014 was 90.32%, and that of the second generation was 100%. Forecasts based on Markov chains in 2015 and 2016 were highly consistent with the actual results. Forecasts based on contingency table analysis in 2015 and 2016 were all completely consistent with the actual results. The accuracy of forecasted peaks of larval abundance from 1983 to 2014 was 90.32% for the first generation, and 83.87% for the second generation. However, when the peak abundance of second generation larvae was less than 3.5 larvae per strain, then the forecast accuracy was 74.19%. **[Conclusion]** Choice of independent variable was the key to the accuracy of regression forecasts. Models based on stationary time series were suitable as temporal patterns of larval abundance conformed to stationary time series. Classification standards directly affect the accuracy of predictions based on contingency table analysis and Markov chains. BP neural networks can be used to investigate nonlinear relationships among independent variables and were a relatively ideal forecasting method.

**Key words** the peak occurrence quantity of *Dendrolimus punctatus* larvae, the prediction model of stationary time series, the prediction model of regression, the prediction model of BP neural network, the prediction model of Markov chains, the prediction model of contingency table analysis

马尾松毛虫 *Dendrolimus punctatus* Walker, 分布于安徽、河南、四川、贵州、陕西、云南、江西、江苏、湖南、浙江、福建、广东、台湾、海南、广西等省, 主要危害马尾松 *Pinus massoniana* Lamb, 还危害黑松 *Pinus thunbergii* Parl、火炬松 *Pinus taeda* L.、湿地松 *Pinus elliottii* Engelm、晚松 *Pinus rigida* var. *serotina* (Michx.) Loud ex Hoopes、海南松 *Pinus fenzeliana* Hand-Mazz 等松属植物。20 世纪中叶在我国森林害虫中马尾松毛虫是发生最广、危害面积最大、经常猖獗成灾的害虫。在广大丘陵地区虫害此起彼伏, 针叶常被吃光, 被害时如同火烧, 造成了巨大的经济效益损失和生态效益损失。再者, 人们在从事林业活动中接触马尾松毛虫毒毛, 容易引发皮炎和关节肿痛, 该虫不但影响林业生产, 还危害人身健康(侯陶谦, 1987; 邹运鼎等, 1990; 萧刚柔, 1992; 张真和李典谟, 2008)。进入 21 世纪, 由于各地采用封山育林、混交、间作等措施优化了森林生态环境, 科学地运用综合治理措施, 使马尾松毛虫的危害得到有效的控制, 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, 对其的监测不能放松警惕。

马尾松毛虫一年发生 2~4 代, 发生世代的多

少, 随不同地方而异, 在河南省信阳地区一年发生 2 代为主, 在长江流域诸省一年发生 2~3 代, 而在广东、广西、福建南部一年发生 3~4 代, 海南一年发生 4~5 代(萧刚柔, 1992)。安徽潜山县一年发生 3 代, 即 4—6 月上旬为越冬代, 6 月上旬—8 月中下旬为一代, 8 月中下旬—12 月为第 2 代。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。科研工作者(张爱兵等, 2001; Park *et al.*, 2003; 贾青生, 2006; Zhang *et al.*, 2008; 陈绘画等, 2011; 田万银和徐华潮, 2012; 费海泽等, 2014; 许章华等, 2014; 王庆等, 2016) 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、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。为了有效地防治马尾松毛虫, 本研究采用时间平稳序列法、回归预测法、人工 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马尾松毛虫害虫态——幼虫的 1 代和 2 代高峰期的发生量, 并对几种方法预测结果的优缺点进行比较, 以期对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 1998 年缺如。根据国家林业局森林病虫害防治总站 (2013) 制定的调查方法进行调查, 即采用踏查和详查相结合的办法, 对不同虫态采取不同的观测方法。沿林斑线、林道、公路、铁路等线路调查, 目测发生范围, 危害状况, 发现虫情或灾情立即设临时标准地, 采取平行线抽样法抽取 20 株标准株详查。卵期调查是在雌蛾羽化高峰后 1~2 d 调查平均卵块数, 卵块平均粒数; 幼虫期调查, 1~2 龄幼虫调查枯黄卷曲的枝数, 推算幼虫数, 3 龄以上幼虫 3 m 以下小树直接调查合计树冠上的幼虫数, 大树用“虫粪粒推算法”调查, 幼虫越冬期调查树干基部树皮缝中的幼虫数推算全部虫口。蛹期调查, 在结茧盛期后 2~5 d 剖茧, 调查雌雄性比, 平均雌蛹重、死亡率、寄生率; 成虫期调查, 在成虫羽化前 2~3 d 及羽化期用黑光灯诱杀, 统计其诱杀数量和雌雄性比。

## 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。

对害虫的发生量  $X(t)$  的随机函数, 若相关函数  $R_x(t_1 t_2)$  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  $k = t_2 - t_1$  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,

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

相关函数  $R_x(k)$  的绝对值随着  $|k| \rightarrow \infty$  而趋于零, 相关函数可表示成

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

其中  $t_0$  代表资料时间总长度。

本文采用 3 次序列预报方程组, 其预报系数

为  $a_1$ 、 $a_2$ 、 $a_3$  由下列方程组求解

$$\begin{cases} R_0 a_1 + R_1 a_2 + R_2 a_3 = R_3 \\ R_1 a_1 + R_0 a_2 + R_1 a_3 = R_2 \\ R_2 a_1 + R_1 a_2 + R_0 a_3 = R_1 \end{cases} \quad (3)$$

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

$$R_1 = \frac{1}{N-1} \sum (x_1 x_2) + (x_2 x_3) + \dots + (x_{n-1} x_n) \quad (5)$$

$$R_1 = \frac{1}{N-2} \sum (x_1 x_3) + (x_2 x_4) + \dots + (x_{n-2} x_n) \quad (6)$$

$$R_1 = \frac{1}{N-3} \sum (x_1 x_4) + (x_2 x_5) + \dots + (x_{n-3} x_n) \quad (7)$$

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

$$x_t = a_1 x_{t-3} + a_2 x_{t-2} + a_3 x_{t-1} \quad (8)$$

式中  $x_t$ 、 $x_{t-3}$ 、 $x_{t-2}$  和  $x_{t-1}$  为第  $t$  年、第  $t-3$  年、 $t-2$  年和  $t-1$  年的发生量等级值。 $a_1$ 、 $a_2$  和  $a_3$  是预测式的相应系数。

**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高峰期发生量之间建立回归预测模型进行预测 (汪四水和张孝羲, 2000), 预测式为  $y = a + bx$  或  $y = a + b_1 x_1 + b_2 x_2 + b_3 x_3$ ,  $y$  为某代幼虫高峰期发生量,  $a$  为截距,  $b$  或  $b_i$  为斜率,  $x_i$  为自变量, 即为前一虫态发生量或某时段的气象因子。

**1.2.3 BP 神经网络法** 神经网络的模式很多 (汪四水和张孝羲, 2000; 张孝羲, 2006; 许章华等, 2014), 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)。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:

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

隐层单元的输入和输出:

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

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

输出单元的输入和输出：

$$\text{输入：} \text{net}_k = \sum_{i=1}^m w_{jk}^{(2)} x_j$$

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

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：

$\Delta w = -\eta \frac{\partial E}{\partial W}$ ，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： $\Delta_R W_{jk}^{(2)} = \eta \delta_k^R \times O_j^R$ ，其中  $\delta_k^R = -f'(\text{net}_k^R)(Z_k^R - O_k^R)$ ，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， $\delta_j^R = -f'(\text{net}_j^R) \sum_{k=1}^n \delta_k^R W_{jk}^{(2)}$ ，阈值改变量为  $\Delta_R \theta_j^R = \eta \delta_j^R$ 。

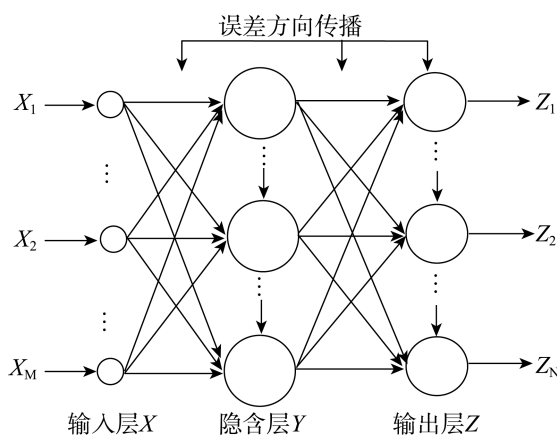


图 1 BP 网络结构原理图

Fig. 1 Structure of BP neural network

$X_i$  表示输入层第  $i$  个节点的输入，即  $X_i$  为与某代高峰期发生量有关的气象因子和前一虫态的发生量。 $W_{ij}$  为隐含层第  $j$  个节点到输入层第  $i$  个节点之间的权值， $Z_j$  表示隐含层第  $j$  个节点的阈值， $(j=1, \dots, M)$ ； $O_j$  表示隐含层的激励函数， $W_{ij}$  表示输出层第  $k$  个节点到隐含层第  $j$  个节点之间的权值， $O_k$  表示输出层的激励函数， $Z_k$  表示输出层第  $k$  个节点的输出，本文为一个节点，即高峰期发生量。输出层的输入变量  $X_i$  为与幼虫高峰期发生量有关的因子，输出层的输出变量为预报量，即本文的某代幼虫高峰期发生量预测值。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** 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（邹运鼎和王弘法，1989）。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，同时系统由状态  $E_i^{(k)}$  经过 1 次转移到状态  $E_i^{(k+1)}$  的概率只与前一步的状态  $E_i^{(k)}$  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，由状态  $E_i$  转移到状态  $E_j$  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； $N_j(m)$  为马尾松毛虫等级由  $E_i$  转移到  $E_j$  的次数。 $m$  阶转移概率矩阵为

$$P(m) = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1k} \\ P_{21} & P_{22} & \cdots & P_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ P_{n1} & P_{n2} & \cdots & P_{nk} \end{bmatrix}$$

式中状态  $E_i$  即第  $i$  年某世代幼虫高峰期发生量，状态  $E_j$  即第  $j$  年某世代幼虫高峰期发生量， $P_{(m)}$  为以预报年相距  $m$  年各级值的转移概率，相距 1 年为  $P_{(1)}$ ，相距 2 年为  $P_{(2)}$ ， $P_{1k}$  为由 1 级转移到  $K$  级的概率， $P_{2k}$  为由 2 级转移到  $k$  级的概率，计算出转移概率之和，以概率总和最大的那个状态的级别作为预报年可能发生级别的预测值，对历年高峰期发生量进行分级方可进行计算。

### 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(邹运鼎和王弘法, 1989; 李祚泳和彭荔红, 1999)。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高峰期发生量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, 本文分 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, 将预报量分级的值代入一元回归式, 即可求出自变量对应的级别的值, 本文预报量  $X_i$  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  $X_i$  作用下出现的概率作出估计, 计算一个因子在  $X_i$  处于  $k$  级相对应的  $y$  出现 1 级的概率有多大, 然后将多个预报因子  $X_i$  可能出现 1 级的概率进行综合, 估计预报对象  $y$  出现 1 级的概率, 以进行预报。

按照上述首先对预报量  $y$  (幼虫高峰期发生量) 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 1 级的条件概率公式  $P_{ik}^i = \frac{n_{kl}}{n_k}$  求出其条件概率,  $P_{ik}^i$  为第  $i$  个因子处于  $k$  级的条件下,  $y$  出现 1 级的条件概率  $P$ ,  $n_{kl}$  为预报量  $y$  处于 1 级的频率,  $n_k$  是历史上该预报因子出现  $k$  级的总频率。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  $\hat{y}_t$ ,

$$\hat{y}_t = \frac{1}{m} \sum_{i=1}^m P_{kl}^i = \frac{1}{m} \sum_{i=1}^m \frac{n_{kl}^i}{n_k^i}, \quad n_{kl}^i \text{ 为第 } i \text{ 个因子处于 } k$$

级预报量  $y$  处于 1 级的频率,  $n_k^i$  为第  $i$  个因子处于  $k$  级, 预报量  $y$  处于  $k$  级的总频率, 取  $\hat{y}_t$  最大值的 1 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

根据安徽省潜山县马尾松毛虫 1983—2014 年的资料统计, 一年发生 3 代, 即越冬代, 1 代

和 2 代。1—6 月为越冬代, 1—4 月下旬为幼虫期, 4 月下旬—5 月中下旬为蛹期, 5 月下旬—6 月上旬为成虫期, 5 月中下旬—8 月中旬为第 1 代, 5 月中下旬—6 月上旬为卵期, 6 月上旬—7 月中下旬为幼虫期, 7 月中下旬—8 月上旬为蛹期, 8 月上旬—8 月中旬为成虫期, 8 月中下旬—12 月为第 2 代, 8 月中下旬—9 月上旬为卵期, 9 月上旬—11 月上旬为幼虫期, 温度适宜时少量幼虫可继续发育。

### 2.1 用时间平稳序列法预测幼虫高峰期发生量模型

**2.1.1 1 代幼虫高峰期发生量预测模型** 将 1983—2014 年 1 代幼虫高峰期的发生量进行分级并列于表 1, 1998 年因缺如, 用内插法进行估值, 2015 年和 2016 年作为验证样本。

运用 (4) (5) (6) (7) 式求得  $R_0=5.0313$ ,  $R_1=4.2258$ ,  $R_2=4.3667$ ,  $R_3=3.7931$ , 代入 (3), 则预测模型为  $X_t = -0.1769X_{t-3} + 0.6182X_{t-2} + 0.4742X_{t-1}$ 。

将表中 2012、2013 和 2014 年的  $X_t$  值代入, 求出 2015 年为 0.9155, 约等于 1 级, 实际发生值为 1 级, 将中 2013、2014 和 2015 年的  $X_t$  值代入, 求出 2016 年为 0.9155, 约等于 1 级, 实际发生值为 1 级, 预测的 2015 年、2016 年 1 代幼虫高峰期发生量与实际相符。

**2.1.2 2 代幼虫高峰期发生量预测模型** 将 1983—2014 年 1 代幼虫高峰期的发生量分级后列于表 1, 1998 年因缺如, 用内插法进行估值, 运用 (4) (5) (6) (7) 式求得  $R_0=5.1563$ ,  $R_1=4.0323$ ,  $R_2=4.0333$ ,  $R_3=4.1034$ , 代入 (3) 式, 则预测模型为  $X_t = 0.3486X_{t-3} + 0.2865X_{t-2} + 0.2853X_{t-1}$ 。

将表 1 中 2012、2013 和 2014 年的  $X_t$  值代入上式, 求出  $X_{2015}=0.9204 \approx 1$ , 即 2015 年的发生量预测值为 1 级, 与实际发生值相符, 将中 2013、2014 和 2015 年的  $X_t$  值代入上式, 求出  $X_{2016}=0.9204 \approx 1$ , 即 2016 年的发生量预测值为 1 级, 预测的 2015 年、2016 年 1 代幼虫高峰期发生量预测值与实际值完全吻合。

表 1 幼虫发生量的  $X_t$  值 (分级值)  
Table 1 The value of  $X_t$  for the occurrence quantity of larvae (grading value)

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

a: 1 代幼虫高峰期发生量 (头/株); b: 2 代幼虫高峰期发生量 (头/株)。下表同。分级标准均为: 小于 4 头/株为 1 级, 4~6.49 头/株为 2 级, 6.5~9 头/株为 3 级, 大于 9 头/株为 4 级。2015 年和 2016 年括号内数字为预测的数值。

Letter "a" represents the peak occurrence quantity of the first generation larvae (head/plant). Letter "b" represents the peak occurrence quantity of the second generation larvae (head/plant). The same below. The level standard of occurrence quantity is as follows. The level one was less than 4 heads per plant. The level two was 4-6.49 heads per plant. The level three was 6.5-9 heads per plant and the level four was more than 9 heads per plant. The numbers in brackets of 2015 and 2016 were the predictive values.

## 2.2 用回归预测法建立幼虫高峰期发生量预测模型

2.2.1 1 代幼虫高峰期发生量的回归预测模型  
用 1983—2014 年资料建立模型, 2015 年和 2016 年资料作为验证资料。

1) 1 代幼虫高峰期发生量  $y_1$  与越冬代蛹高峰期发生量 ( $x_1$ ) 的关系

$$y_1 = 2.9024 + 5.4697x_1, r = 0.5476^{**}$$

( $df = 29, r_{0.01} = 0.4567$ );

2) 2 代幼虫高峰期发生量  $y_1$  与越冬代成虫高峰期发生量 ( $x_2$ ) 的关系

$$y_1 = 2.4199 + 6.2894x_2, r = 0.5760^{**};$$

3) 1 代幼虫高峰期发生量  $y_1$  与 1 代卵盛期卵量 ( $x_3$ ) 的关系

$$y_1 = 0.0594 + 0.8841x_3, r = 0.9820^{**};$$

4) 1 代幼虫高峰期发生量  $y_1$  与当代马尾松毛虫卵的赤眼蜂寄生率 ( $x_4$ ) 的关系

$$y_1 = 13.4496 - 0.5836x_4, r = -0.5831^{**};$$

5) 1 代幼虫高峰期发生量  $y_1$  与越冬代幼虫累计虫口 ( $x_5$ ) 的关系

$$y_1 = 1.8219 + 0.2170x_5, r = 0.7416^{**};$$

6) 1 代幼虫高峰期发生量  $y_1$  与 4 月上旬降

雨量 ( $x_6$ ) 的关系

$$y_1 = 2.0554 + 0.0880x_6, r = 0.5616^{**};$$

7) 1 代幼虫高峰期发生量  $y_1$  的多元回归预测模型

选择与 1 代幼虫高峰期发生量  $y_1$  相关系数绝对值大的  $x_3$ 、 $x_4$  和  $x_5$  组建多元回归预测模型:  
 $y_1 = 0.5167 + 0.8616x_3 - 0.0303x_4 + 0.0051x_5$ , 其复相关系数  $r$  为 0.9826。

8) 1 代幼虫高峰期发生量  $y_1$  的逐步回归预测模型

对与 1 代幼虫高峰期发生量  $y_1$  达到相关水平的 6 个因子进行逐步回归运算, 建立逐步回归预测模型:  $y_1 = -0.1575 - 1.7398x_1 + 0.8780x_3 + 0.0662x_5$ , 其复相关系数  $r$  为 0.9831。

为了验证模型预测结果的准确性, 将预测结果列于表 2, 可看出对 2015 年和 2016 年两年预测结果准确性高的依次是多元回归式、逐步回归式和 1 代卵盛期卵量作为自变量的一元回归式。

## 2.2.2 2 代幼虫高峰期发生量 $y_2$ 的回归预测模型

1)  $y_2$  与 1 代蛹高峰期发生量 ( $x_7$ ) 的关系

$$y_2 = 2.6519 + 2.9262x_7, r = 0.6464^{**};$$

2)  $y_2$  与 1 代成虫高峰期发生量 ( $x_8$ ) 的关系

$$y_2 = 2.2696 + 3.2522x_8, r = 0.7037^{**};$$

表 2 马尾松毛虫 1 代幼虫高峰期发生量的预测值和实际值  
Table 2 The prediction and realization results of the peak occurrence quantity of the first generation *Dendrolimus punctatus* larvae

| 年份<br>Year | 预报因子 Forecast factor                       |        |        |        |        |        | 多元回归<br>Multiple<br>regression | 逐步回归<br>Stepwise<br>regression | 备注<br>Remark |
|------------|--------------------------------------------|--------|--------|--------|--------|--------|--------------------------------|--------------------------------|--------------|
|            | $X_1$                                      | $X_2$  | $X_3$  | $X_4$  | $X_5$  | $X_6$  |                                |                                |              |
| 2015       | 预测值<br>Predictive value                    | 3.6681 | 3.6149 | 2.4288 | 3.3533 | 3.1890 | 15.8538                        | 2.3332                         | 2.3624       |
|            | 误差 Error                                   | 1.5450 | 1.4918 | 0.3057 | 1.2302 | 1.0659 | 13.7307                        | 0.2101                         | 0.2393       |
|            | 准确率排序<br>Sort according to<br>the accuracy | 7      | 6      | 3      | 5      | 4      | 8                              | 1                              | 2            |
| 2016       | 预测值<br>Predictive value                    | 3.7229 | 3.7406 | 2.5702 | 3.5868 | 3.2541 | 18.5026                        | 2.4887                         | 2.5451       |
|            | 误差 Error                                   | 1.4723 | 1.4900 | 0.3196 | 1.3362 | 1.0035 | 16.2520                        | 0.2381                         | 0.2945       |
|            | 准确率排序<br>Sort according to<br>the accuracy | 6      | 7      | 3      | 5      | 4      | 8                              | 1                              | 2            |

3)  $y_2$  与 2 代卵盛期卵量量 ( $x_9$ ) 的关系  
 $y_2 = -0.3855 + 0.7179x_9, r = 0.9764^{**}$  ;

4)  $y_2$  与 1 代幼虫累计虫口 ( $x_{10}$ ) 的关系  
 $y_2 = 1.8208 + 0.2171x_{10}, r = 0.7416^{**}$  ;

5)  $y_2$  与 2 代马尾松毛虫卵的赤眼蜂寄生率 ( $x_{11}$ ) 的关系  
 $y_2 = 9.8288 - 0.2457x_{11}, r = -0.4413^*$  ;

6)  $y_2$  的多元回归预测模型  
选择与 2 代幼虫高峰期发生量  $y_2$  相关系数绝对值大的  $x_8$ 、 $x_9$  和  $x_{10}$  组建多元回归预测模型： $y_1 = -0.5808 - 3.3343x_8 + 0.7054x_9 + 0.2134x_{10}$  , 其复相关系数  $r$  为 0.9853。

7)  $y_2$  的逐步回归预测模型  
通过逐步回归运算，组建回归预测模型：  
 $y_2 = -0.3768 - 0.1654x_8 + 0.7373x_9$  , 其复相关系数  $r$  为 0.9767。

将上述模型预测结果列于表 3，可看出对 2015 年和 2016 年两年预测结果准确性高的前 3 位预测模型依次为逐步回归式、2 代卵盛期卵量 ( $X_9$ ) 作为自变量的一元回归模型和多元回归模型。

2.3 用人工 BP 神经网络法预测马尾松毛虫幼虫高峰期发生量

2.3.1 1 代幼虫高峰期发生量的 BP 神经网络法预测模型 对 1983—2014 年作为学习样本进行分析，2015 年和 2016 年资料作为预测样本进行预报结果验证。对旬均温、旬降雨量等 85 个预报因子利用相关系数筛选出越冬代蛹高峰期发生量 ( $x_1$ ) 越冬代成虫高峰期发生量 ( $x_2$ ) 1 代卵盛期卵量 ( $x_3$ ) 1 代马尾松毛虫卵的赤眼蜂寄生率 ( $x_4$ ) 越冬代幼虫累计虫口 ( $x_5$ ) 和 4 月上旬降雨量 ( $x_6$ ) 6 个因子作为输入样本，采用 6-6-1 的人工 BP 神经网络模型，通过本系统运算可输出 1 代幼虫高峰期发生量，结果列于表 4，学习样本中 1990 年、2003 年和 2010 年预测值与实际值误差大于 1 头/株，若以 1 头/株为误差标准，预测正确率为 90.32%。2015 年和 2016 年作为预测样本，若以 1 头/株作为标准误差，预测准确率均为 100%。

2.3.2 2 代幼虫高峰期发生量的 BP 神经网络预测模型 对 1983—2014 年作为学习样本进行分析，2015 年和 2016 年资料作为预测样本进行预报结果验证。对 6 月—9 月上旬的旬均温、旬降

表 3 马尾松毛虫 2 代幼虫高峰期发生量的预测值和实际值  
Table 3 The prediction and realization results of the peak occurrence quantity of the second generation *Dendrolimus punctatus* larvae

| 年份<br>Year                              | 预报因子 Forecast factor |        |          |          |          |                             |                             | 备注<br>Remark                    |
|-----------------------------------------|----------------------|--------|----------|----------|----------|-----------------------------|-----------------------------|---------------------------------|
|                                         | $X_7$                | $X_8$  | $X_9$    | $X_{10}$ | $X_{11}$ | 多元回归<br>Multiple regression | 逐步回归<br>Stepwise regression |                                 |
| 2015 预测值<br>Predictive value            | 3.9102               | 3.4079 | 1.7969   | 3.1885   | 4.0619   | 1.7410                      | 1.7885                      | 2015 年<br>实际值为<br>1.9092<br>头/株 |
| 误差 Error                                | 2.0010               | 1.4987 | - 0.1123 | 1.2793   | 2.1527   | - 0.1682                    | - 0.1207                    |                                 |
| 准确率排序<br>Sort according to the accuracy | 6                    | 5      | 2        | 4        | 7        | 3                           | 1                           |                                 |
| 2016 预测值<br>Predictive value            | 3.8809               | 3.5054 | 2.1845   | 3.2537   | 3.7354   | 2.0859                      | 2.1998                      | 2016 年<br>实际值为<br>2.3560<br>头/株 |
| 误差 Error                                | 1.5249               | 1.1494 | - 0.1715 | 0.8977   | 1.3794   | - 0.2701                    | - 0.1562                    |                                 |
| 准确率排序<br>Sort according to the accuracy | 7                    | 5      | 2        | 4        | 6        | 3                           | 1                           |                                 |

雨量等 95 个预报因子利用相关系数筛选出 1 代蛹高峰期发生量 ( $x_7$ )、1 代成虫高峰期发生量 ( $x_8$ )、2 代卵盛期卵量 ( $x_9$ )、1 代幼虫累计虫口 ( $x_{10}$ ) 和 1 代马尾松毛虫卵的赤眼蜂寄生率 ( $x_{11}$ ) 等 5 个因子作为输入样本, 采用 5-5-1 的人工 BP 神经网络模型, 通过本系统运算可输出 2 代幼虫高峰期发生量, 结果也列于表 4, 学习样本中若以大于 1 头/株为误差标准, 预测正确率为 100%。2015 年和 2016 年作为预测样本, 若以 1 头/株作为标准误差, 预测准确率均为 100%。

#### 2.4 利用马尔科夫链方法预测马尾松毛虫幼虫高峰期发生量

1983—2014 年的各个分级值以相距 1 年各级值的转移概率为  $p_{(1)}$ , 相距 2 年为  $p_{(2)}$ , 相距 3 年为  $p_{(3)}$ , 逐个计算转移概率矩阵, 1 代幼虫高峰期发生量的转移概率矩阵如下:

$$p_{(1)} = \begin{vmatrix} 0.58 & 0.25 & 0.08 & 0.08 \\ 0.25 & 0.50 & 0 & 0.25 \\ 1 & 0 & 0 & 0 \\ 0.33 & 0.33 & 0 & 0.33 \end{vmatrix}$$

$$p_{(2)} = \begin{vmatrix} 0.64 & 0.18 & 0.09 & 0.09 \\ 0.42 & 0.33 & 0 & 0.25 \\ 0 & 1 & 0 & 0 \\ 0.17 & 0.50 & 0 & 0.33 \end{vmatrix}$$

$$p_{(3)} = \begin{vmatrix} 0.30 & 0.40 & 0.10 & 0.20 \\ 0.50 & 0.33 & 0 & 0.17 \\ 1 & 0 & 0 & 0 \\ 0.50 & 0.33 & 0 & 0.17 \end{vmatrix}$$

2 代幼虫高峰期发生量的转移概率矩阵如下:

$$p_{(1)} = \begin{vmatrix} 0.64 & 0.14 & 0.14 & 0.07 \\ 0.44 & 0.33 & 0 & 0.22 \\ 0 & 0.50 & 0 & 0.50 \\ 0.33 & 0.50 & 0 & 0.17 \end{vmatrix}$$

$$p_{(2)} = \begin{vmatrix} 0.46 & 0.31 & 0.15 & 0.08 \\ 0.67 & 0.11 & 0 & 0.22 \\ 0.50 & 0.50 & 0 & 0 \\ 0.33 & 0.33 & 0 & 0.33 \end{vmatrix}$$

$$p_{(3)} = \begin{vmatrix} 0.42 & 0.42 & 0.08 & 0.08 \\ 0.67 & 0.11 & 0.11 & 0.11 \\ 0.50 & 0 & 0 & 0.50 \\ 0.50 & 0.17 & 0 & 0.33 \end{vmatrix}$$

列出与 2015 年相距 1、2、3 年马尾松毛虫

表 4 马尾松毛虫幼虫高峰期发生量的人工 BP 神经网络预测结果  
Table 4 The prediction results by using BP neural network of the peak occurrence quantity of *Dendrolimus punctatus* larvae

| 年份<br>Year | a                   |                         |             | b                   |                         |             | 年份<br>Year | a                   |                         |             | b                   |                         |             |
|------------|---------------------|-------------------------|-------------|---------------------|-------------------------|-------------|------------|---------------------|-------------------------|-------------|---------------------|-------------------------|-------------|
|            | 实际值<br>Actual value | 预测值<br>Predictive value | 误差<br>Error | 实际值<br>Actual value | 预测值<br>Predictive value | 误差<br>Error |            | 实际值<br>Actual value | 预测值<br>Predictive value | 误差<br>Error | 实际值<br>Actual value | 预测值<br>Predictive value | 误差<br>Error |
| 1983       | 6.29                | 6.77                    | - 0.48      | 10.23               | 10.87                   | - 0.63      | 2001       | 2.11                | 2.26                    | - 0.15      | 3.62                | 2.99                    | 0.62        |
| 1984       | 6.09                | 6.18                    | - 0.09      | 5.13                | 5.08                    | 0.06        | 2002       | 3.11                | 2.64                    | 0.48        | 4.99                | 4.06                    | 0.93        |
| 1985       | 11.94               | 12.71                   | - 0.77      | 5.67                | 5.81                    | - 0.13      | 2003       | 6.53                | 5.35                    | 1.17        | 5.87                | 5.65                    | 0.21        |
| 1986       | 13.92               | 13.64                   | 0.28        | 12.20               | 12.98                   | - 0.78      | 2004       | 2.63                | 2.41                    | 0.22        | 5.67                | 5.81                    | - 0.14      |
| 1987       | 14.30               | 14.38                   | - 0.08      | 13.38               | 12.83                   | 0.55        | 2005       | 5.25                | 4.54                    | 0.71        | 3.11                | 3.88                    | - 0.77      |
| 1988       | 5.66                | 5.43                    | 0.22        | 3.63                | 4.51                    | - 0.88      | 2006       | 3.47                | 3.54                    | - 0.07      | 3.14                | 3.44                    | - 0.30      |
| 1989       | 5.93                | 5.94                    | - 0.01      | 9.31                | 9.51                    | - 0.20      | 2007       | 4.13                | 3.66                    | 0.47        | 2.93                | 3.28                    | - 0.35      |
| 1990       | 15.26               | 13.87                   | 1.39        | 4.18                | 4.18                    | - 0.01      | 2008       | 4.19                | 4.17                    | 0.02        | 3.47                | 3.46                    | 0.02        |
| 1991       | 1.92                | 2.44                    | - 0.52      | 3.18                | 3.79                    | - 0.61      | 2009       | 4.56                | 4.62                    | - 0.07      | 3.34                | 3.66                    | - 0.32      |
| 1992       | 2.35                | 2.44                    | - 0.09      | 2.64                | 3.02                    | - 0.38      | 2010       | 5.52                | 7.08                    | - 1.56      | 8.04                | 8.24                    | - 0.20      |
| 1993       | 2.10                | 2.34                    | - 0.24      | 6.73                | 6.62                    | 0.11        | 2011       | 5.15                | 5.66                    | - 0.51      | 4.53                | 4.30                    | 0.22        |
| 1994       | 11.77               | 11.31                   | 0.47        | 12.32               | 12.70                   | - 0.38      | 2012       | 2.35                | 2.62                    | - 0.28      | 2.31                | 2.74                    | - 0.44      |
| 1995       | 5.66                | 5.19                    | 0.48        | 5.57                | 5.47                    | 0.10        | 2013       | 2.30                | 2.48                    | - 0.18      | 1.92                | 2.42                    | - 0.51      |
| 1996       | 13.07               | 13.54                   | - 0.46      | 15.89               | 15.21                   | 0.67        | 2014       | 2.16                | 2.49                    | - 0.33      | 1.85                | 2.64                    | - 0.79      |
| 1997       | 2.17                | 2.48                    | - 0.31      | 3.08                | 3.05                    | 0.03        | 2015       | 2.12                | 2.74                    | - 0.62      | 1.91                | 2.58                    | - 0.67      |
| 1999       | 2.02                | 2.39                    | - 0.37      | 3.07                | 2.70                    | 0.38        | 2016       | 2.25                | 2.81                    | - 0.56      | 2.24                | 2.61                    | - 0.37      |
| 2000       | 2.43                | 2.29                    | 0.15        | 3.59                | 2.74                    | 0.85        |            |                     |                         |             |                     |                         |             |

发生级值,依据转移概率值总和的最大值为预报值(表 5)。

由表 5 可知,在转移状态概率合计栏中,以状态 1 的概率最大,1 代、2 代幼虫高峰期发生量转移概率合计均为 1.52,故预报 2015 年马尾松毛虫 1 代、2 代幼虫高峰期发生量均为 1 级。为了预测 2016 年马尾松毛虫的幼虫高峰期发生量,将 2015 的发生程度级列入基础数据,重新组建转移矩阵,根据以上方法预测 2016 年马尾松毛虫的 1 代、2 代幼虫高峰期发生量均为 1 级,根据级数划分 2015 年和 2016 年 1、2 代幼虫高峰期发生量实际值均为 1 级,与预测值完全一致。

2.5 利用列联表分析法进行马尾松毛虫幼虫高峰期发生量的多因子多级预报

2.5.1 1 代幼虫高峰期发生量预测 对 1 代幼虫

高峰期发生量  $y_1$  按极差法进行分级见表 1, $y_1$  的级数代入与其相关显著的自变量( $x_i$ )的直线回归式进行分级,分级结果列于表 6,将因子对  $y_1$  的条件概率列联表列于表 7,将 2015 年  $x_1$ 、 $x_3$ 、 $x_4$  和  $x_6$  相应的条件概率进行计算得 2015 年  $\hat{y}_1=0.6275$ ,  $\hat{y}_2=0.2475$ ,  $\hat{y}_3=0$ ,  $\hat{y}_4=0.125$ ,  $\hat{y}_1$  最大,即 2015 年 1 代幼虫高峰期发生量为 1 级,实际值也是 1 级。2016 年  $\hat{y}_1=0.6275$ ,  $\hat{y}_2=0.2475$ ,  $\hat{y}_3=0$ ,  $\hat{y}_4=0.125$ ,  $\hat{y}_1$  最大,即 2016 年 1 代幼虫高峰期发生量为 1 级,实际值也是 1 级。计算出 1983—2014 年 31 年的预测值列于表 6,预测值与实际值的历史符合率为 90.32%,误报相差 1 级,误报的预测值与实际值相差 0~2.5 头/株。

2.5.2 2 代幼虫高峰期发生量预测 对 2 代幼虫

表 5 2015 年马尾松毛虫幼虫高峰期发生量预报

Table 5 The prediction of the peak occurrence quantity of *Dendrolimus punctatus* larvae in 2015

| 年份<br>Year | 起始状态<br>Initial state | 转移步数<br>Transfer steps | 转移概率 (a) Transition probability |      |      |      | 转移概率 (b) Transition probability |      |      |      |
|------------|-----------------------|------------------------|---------------------------------|------|------|------|---------------------------------|------|------|------|
|            |                       |                        | 1                               | 2    | 3    | 4    | 1                               | 2    | 3    | 4    |
| 2014       | 1                     | 1                      | 0.58                            | 0.25 | 0.08 | 0.08 | 0.64                            | 0.14 | 0.14 | 0.07 |
| 2013       | 1                     | 2                      | 0.64                            | 0.18 | 0.09 | 0.09 | 0.46                            | 0.31 | 0.15 | 0.08 |
| 2012       | 1                     | 3                      | 0.30                            | 0.40 | 0.10 | 0.20 | 0.42                            | 0.42 | 0.08 | 0.08 |
| 合计 Total   |                       |                        | 1.52                            | 0.83 | 0.27 | 0.37 | 1.52                            | 0.87 | 0.38 | 0.23 |

表 6 预报要素分级值及预测值

Table 6 The forecast factors of grading and prediction value

| 年份<br>Year | $X_1$ | $X_3$ | $X_4$ | $X_6$ | $Y_1$ | 预测值<br>Predictive value | $X_7$ | $X_8$ | $X_9$ | $Y_2$ | 预测值<br>Predictive value |
|------------|-------|-------|-------|-------|-------|-------------------------|-------|-------|-------|-------|-------------------------|
| 1983       | 2     | 2     | 4     | 3     | 2     | 2                       | 2     | 3     | 4     | 4     | 4                       |
| 1984       | 4     | 2     | 3     | 2     | 2     | 2                       | 3     | 2     | 2     | 2     | 2                       |
| 1985       | 4     | 4     | 4     | 4     | 4     | 4                       | 4     | 4     | 2     | 2     | 2                       |
| 1986       | 3     | 4     | 4     | 3     | 4     | 4                       | 4     | 4     | 4     | 4     | 4                       |
| 1987       | 3     | 4     | 3     | 3     | 4     | 4                       | 4     | 4     | 4     | 4     | 4                       |
| 1988       | 3     | 2     | 3     | 1     | 2     | 2                       | 2     | 2     | 1     | 1     | 1                       |
| 1989       | 4     | 2     | 4     | 4     | 2     | 2                       | 2     | 2     | 3     | 3     | 3                       |
| 1990       | 4     | 4     | 3     | 3     | 4     | 4                       | 4     | 4     | 2     | 2     | 2                       |
| 1991       | 2     | 1     | 3     | 3     | 1     | 1                       | 1     | 1     | 1     | 1     | 1                       |
| 1992       | 2     | 1     | 2     | 2     | 1     | 1                       | 1     | 1     | 1     | 1     | 1                       |
| 1993       | 1     | 1     | 3     | 2     | 1     | 1                       | 1     | 1     | 2     | 2     | 1                       |
| 1994       | 2     | 4     | 3     | 2     | 4     | 4                       | 3     | 3     | 4     | 4     | 4                       |
| 1995       | 2     | 2     | 2     | 1     | 2     | 2                       | 2     | 2     | 2     | 2     | 2                       |
| 1996       | 3     | 4     | 2     | 2     | 4     | 4                       | 3     | 3     | 4     | 4     | 4                       |
| 1997       | 4     | 1     | 2     | 2     | 1     | 1                       | 1     | 1     | 1     | 1     | 1                       |
| 1999       | 1     | 1     | 1     | 3     | 1     | 1                       | 1     | 1     | 1     | 1     | 1                       |
| 2000       | 1     | 1     | 1     | 1     | 1     | 1                       | 1     | 1     | 1     | 1     | 1                       |
| 2001       | 1     | 1     | 1     | 2     | 1     | 1                       | 1     | 1     | 1     | 1     | 1                       |
| 2002       | 1     | 1     | 1     | 2     | 1     | 1                       | 1     | 1     | 2     | 2     | 1                       |
| 2003       | 3     | 2     | 2     | 3     | 3     | 2                       | 2     | 2     | 2     | 2     | 2                       |
| 2004       | 1     | 1     | 1     | 2     | 1     | 1                       | 1     | 1     | 3     | 3     | 1                       |
| 2005       | 2     | 2     | 2     | 2     | 2     | 2                       | 2     | 2     | 1     | 1     | 1                       |
| 2006       | 1     | 2     | 2     | 2     | 1     | 2                       | 2     | 2     | 1     | 1     | 1                       |
| 2007       | 2     | 2     | 2     | 1     | 2     | 2                       | 2     | 2     | 1     | 1     | 1                       |

| 续表 6 (Table 6 continued) |       |       |       |       |       |                            |       |       |       |       |                            |
|--------------------------|-------|-------|-------|-------|-------|----------------------------|-------|-------|-------|-------|----------------------------|
| 年份<br>Year               | $X_1$ | $X_3$ | $X_4$ | $X_6$ | $Y_1$ | 预测值<br>Predictive<br>value | $X_7$ | $X_8$ | $X_9$ | $Y_2$ | 预测值<br>Predictive<br>value |
| 2008                     | 2     | 2     | 1     | 3     | 2     | 1                          | 2     | 2     | 1     | 1     | 3                          |
| 2009                     | 2     | 2     | 2     | 2     | 2     | 2                          | 2     | 2     | 1     | 1     | 2                          |
| 2010                     | 2     | 3     | 2     | 2     | 2     | 2                          | 2     | 2     | 4     | 4     | 4                          |
| 2011                     | 2     | 2     | 1     | 1     | 2     | 2                          | 2     | 2     | 2     | 2     | 1                          |
| 2012                     | 1     | 1     | 1     | 1     | 1     | 1                          | 1     | 1     | 1     | 1     | 1                          |
| 2013                     | 1     | 1     | 1     | 1     | 1     | 1                          | 1     | 1     | 1     | 1     | 1                          |
| 2014                     | 1     | 1     | 1     | 1     | 1     | 1                          | 1     | 1     | 1     | 1     | 1                          |
| 2015                     | 1     | 1     | 1     | 4     | 1     | 1                          | 1     | 1     | 1     | 1     | 1                          |
| 2016                     | 1     | 1     | 1     | 4     | 1     | 1                          | 1     | 1     | 1     | 1     | 1                          |

| 表 7 各因子对 $Y_1$ 的条件概率列联表<br>Table 7 The contingency table for each factor to $Y_1$ of conditional probability |   |       |      |      |      |       |   |       |      |      |      |
|--------------------------------------------------------------------------------------------------------------|---|-------|------|------|------|-------|---|-------|------|------|------|
|                                                                                                              |   | $Y_1$ |      |      |      |       |   | $Y_1$ |      |      |      |
|                                                                                                              |   | 1     | 2    | 3    | 4    |       |   | 1     | 2    | 3    | 4    |
| $X_1$                                                                                                        | 1 | 0.71  | 0.29 | 0.00 | 0.00 | $X_3$ | 1 | 1.00  | 0.00 | 0.00 | 0.00 |
|                                                                                                              | 2 | 0.25  | 0.50 | 0.13 | 0.13 |       | 2 | 0.08  | 0.83 | 0.08 | 0.00 |
|                                                                                                              | 3 | 0.00  | 0.25 | 0.00 | 0.75 |       | 3 | 0.00  | 1.00 | 0.00 | 0.00 |
|                                                                                                              | 4 | 0.20  | 0.40 | 0.00 | 0.40 |       | 4 | 0.00  | 0.00 | 0.00 | 1.00 |
| $X_4$                                                                                                        | 1 | 0.80  | 0.20 | 0.00 | 0.00 | $X_6$ | 1 | 0.50  | 0.50 | 0.00 | 0.00 |
|                                                                                                              | 2 | 0.30  | 0.50 | 0.10 | 0.10 |       | 2 | 0.54  | 0.31 | 0.00 | 0.15 |
|                                                                                                              | 3 | 0.29  | 0.29 | 0.00 | 0.43 |       | 3 | 0.25  | 0.25 | 0.13 | 0.38 |
|                                                                                                              | 4 | 0.00  | 0.50 | 0.00 | 0.50 |       | 4 | 0.00  | 0.50 | 0.00 | 0.50 |

| 表 8 各因子对 $Y_2$ 的条件概率列联表<br>Table 8 The contingency table for each factor to $Y_2$ of conditional probability |   |       |      |      |      |       |   |       |      |      |      |
|--------------------------------------------------------------------------------------------------------------|---|-------|------|------|------|-------|---|-------|------|------|------|
|                                                                                                              |   | $Y_2$ |      |      |      |       |   | $Y_2$ |      |      |      |
|                                                                                                              |   | 1     | 2    | 3    | 4    |       |   | 1     | 2    | 3    | 4    |
| $X_7$                                                                                                        | 1 | 0.75  | 0.17 | 0.08 | 0.00 | $X_8$ | 1 | 0.75  | 0.17 | 0.08 | 0.00 |
|                                                                                                              | 2 | 0.50  | 0.25 | 0.08 | 0.17 |       | 2 | 0.50  | 0.33 | 0.08 | 0.17 |
|                                                                                                              | 3 | 0.00  | 0.33 | 0.00 | 0.67 |       | 3 | 0.00  | 0.00 | 0.00 | 1.00 |
|                                                                                                              | 4 | 0.00  | 0.50 | 0.00 | 0.50 |       | 4 | 0.00  | 0.50 | 0.00 | 0.50 |
| $X_9$                                                                                                        | 1 | 1.00  | 0.00 | 0.00 | 0.00 |       |   |       |      |      |      |
|                                                                                                              | 2 | 0.00  | 1.00 | 0.00 | 0.00 |       |   |       |      |      |      |
|                                                                                                              | 3 | 0.00  | 0.00 | 1.00 | 0.00 |       |   |       |      |      |      |
|                                                                                                              | 4 | 0.00  | 0.00 | 0.00 | 1.00 |       |   |       |      |      |      |

高峰期发生量  $y_2$  按极差法进行分级见表 1, 将  $y_2$  的分级数代入与其相关显著的自变量用直线回归式进行分级, 分级结果列于表 6, 将因子对  $y_2$  的条件概率列联表列于表 8, 2015 年  $x_7$ 、 $x_8$  和  $x_9$  都为“1”级, 将相应的条件概率进行计算得,  $\hat{y}_1 = 0.8333$ ,  $\hat{y}_2 = 0.1133$ ,  $\hat{y}_3 = 0.0533$ ,  $\hat{y}_4 = 0$ ,  $\hat{y}_1$  最大, 即 2015 年 2 代幼虫高峰期发生量为 1 级, 实际值也是 1 级。2016 年  $x_7$ 、 $x_8$  和  $x_9$  都为 1 级, 将相应的条件概率进行计算得,  $\hat{y}_1 = 0.8333$ ,  $\hat{y}_2 = 0.1133$ ,  $\hat{y}_3 = 0.0533$ ,  $\hat{y}_4 = 0$ ,  $\hat{y}_1$  最大, 即 2016 年 2 代幼虫高峰期发生量为 1 级, 实际值也是 1 级。将 1983—2014 年的预测值列于表 6, 31 年的历史符合率为 83.87%, 为了研究不同分级标准对预测值的影响, 将表 1 中  $y_2$  的 1 级标准 4 头/株改为小于 3.5 头/株, 其他分级值不动, 结果是历史符合率为 74.19%。

### 3 小结与讨论

本文采用回归预测法、平稳时间序列预测法、人工神经网络法、马尔科夫链预测法、列联表多因子多级综合相关法对安徽省潜山县 1983—2014 年的马尾松毛虫 1、2 代幼虫高峰期发生量进行分析, 预测验证 2015 年和 2016 年马尾松毛虫 1 代、2 代幼虫高峰期发生量, 本研究的 5 种方法相比结果是: 以卵高峰期卵量为自变量的回归模型、多元回归模型和逐步回归模型预测结果与实际值相差 0.21~0.31 头/株, 其它 8 个一元回归预测结果与实际值相差 1.06~1.58 头/株。平稳时间序列预测 2015 和 2016 年的结果与实际值完全相符。BP 神经网络预测结果若以误差标准为 1 头/株, 1983—2014 年预报准确率 1 代为 90.32%, 2 代为 100%。马尔科夫链预测 2015 和 2016 年, 预测结果与实际值完全相符, 均为 1 级。列联表多因子多级综合相关分析法预测 2015 和 2016 年的结果与实际值完全相符, 1983—2014 年 1 代幼虫高峰期发生量预测的历史符合率均为 90.32%, 2 代为 83.47%。为了研究不同分级标准对预测值的影响, 将 2 代幼虫高峰期发生量的 1 级标准改为小于 3.5 头/株, 则历史符合

率为 74.19%。分级标准科学与否, 直接影响预测结果的准确性。

一元回归方法计算简便预测结果比较直观准确, 但要具有一定年份的资料积累才能分析出自变量与因变量间的实质关系。时间平稳序列法预测, 预报量必须具有一定的连续资料积累, 从表面上看, 时序分析撇开了环境因子等系统变量对害虫发生情况的影响, 但事实上由于时序中反映了曾经发生过的所有环境因子的影响, 时序分析则从总的方面进行考察, 当所需要的环境因子如气象资料等与错综复杂或有关的数据资料无法得到, 时间序列分析法就是理想的分析方法, 平稳随机序列预报法要求害虫发生的实际随机过程中必须是平稳的时间序列, 若不符合平稳时间序列特点, 则应通过非平稳时间序列的转化方可利用时间平稳序列方法分析。

列联表分析法要求必须具备系统的历年观察数据, 用列联表法最重要的是如何划分等级, 同样的资料用不同的标准分级, 其预测效果相差悬殊, 且每一个入选的自变量必须和因变量单相关关系达到相关水平方可, 由于影响害虫发生因子常常是多个, 利用多元回归计算时计算量大, 且计算过程极易出现误差, 与多元回归计算相比利用列联表分析进行多因子多级预报法大大减少工作量, 且预报比较准确。

马尔科夫链预报法要求害虫发生情况符合马尔科夫过程, 通过运算求出状态转移概率, 根据预报年的前一年、前两年和前三年的害虫发生级别分别查状态转移概率  $p_{(1)}$ 、 $p_{(2)}$  和  $p_{(3)}$ , 计算出可能出现级别的平均概率, 用平均概率最大的那个状态作为预报年可能发生级别的预报值, 预测结果比较准确。

人工神经网络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, 模型的稳定性好, BP 神经网络建模精度高低与预报因子的选取有关, 因子选取得当, 预测结果准确率高, 但计算内容复杂, 尚需迭代上千次, 但在电子计算机计算技术普及的情况, 该方法不失为

一种比较理想的方法。

## 参考文献 (References)

- Chen HH, Wang JY, Xu ZH, 2011. Application of response-surface methodology to Chaos detection and forecast of occurrence quantity of *Dendrolimus punctatus*. *Journal of Northeast Forestry University*, 39(9): 94–96. [陈绘画, 王坚娅, 徐志宏, 2011. 基于响应面方法的马尾松毛虫发生量混沌特性检测及其预测. 东北林业大学学报, 39(9): 94–96.]
- Fei HZ, Wang HB, Kong XB, Zhang Z, Zhang SS, Song XG, 2014. Selection and prediction of meteorological factors correlated with *Dendrolimus punctatus* outbreak. *Journal of Northeast Forestry University*, 41(1): 136–140. [费海泽, 王鸿斌, 孔祥波, 张真, 张苏芳, 宋雄刚, 2014. 马尾松毛虫发生相关气象因子筛选及预测. 东北林业大学学报, 41(1): 136–140.]
- Homik K, Stinchcombe M, White H, 1989. Multilayer feedforward networks are universal approximators. *Neural Networks*, (2): 359–366.
- Hou TQ, 1987. The Pine Caterpillars in China. Beijing: Science Press. 188–191. [侯陶谦, 1987. 中国松毛虫. 北京: 科学出版社. 188–191.]
- Jia CS, 2006. Forecast of occurrence level of *Dendrolimus punctatus* by Markov Chains. *Journal of Northeast Forestry University*, 34(5): 21–22. [贾春生, 2006. 利用马尔可夫链方法测报马尾松毛虫发生级别. 东北林业大学学报, 34(5): 21–22.]
- LI ZY, Peng LH, 1999. Prediction model of agricultural plant diseases and insect pests based on artificial neural network and its verification. *Acta Ecologica Sinica*, 19(5): 759–762. [李祚泳, 彭荔红, 1999. 基于人工神经网络的农业病虫害预测模型及其效果检验. 生态学报, 19(5): 759–762.]
- Park YS, Cereghino R, Compin A, 2003. Applications of artificial neural networks for patterning and predicting aquatic insect species richness in running waters. *Ecological Modelling*, 160(3): 265–280.
- The Forest Disease and Pests Prevention and Control Station, 2013. Forest Pest Monitoring and Forecasting Technology. Beijing: Chinese Forest Press. 117–118. [国家林业局森林病虫害防治总站, 2013. 林业有害生物监测预报技术. 北京: 中国林业出版社. 117–118.]
- Tian WY, Xu HC, 2012. The prediction model of *Dendrolimus punctatus* walker in coastal shelterbelt of Zhejiang. *Journal of Environmental Entomology*, 34 (4): 401–406. [田万银, 徐华潮, 2012. 浙江沿海防护林马尾松毛虫的预测预报模型. 环境昆虫学报, 34(4): 401–406.]
- Xiao GR, 1992. Forest Insects of China (2<sup>nd</sup> edition). Beijing: Chinese Academy of Forestry. 948–953. [萧刚柔, 1992. 中国森林害虫(第二版). 北京: 林业出版社. 948–953.]
- Wang SS, Zhang XX, 2000. A neutral network approach to long-term forecasting for rice leaf roller. *Acta Phytophylacica Sinica*, 27(4): 313–316. [汪四水, 张孝羲, 2000. 基于神经网络的稻纵卷叶螟长期预测. 植物保护学报, 27(4): 313–316.]
- Wang Q, Bi M, Du T, 2016. Spatial pattern of *Dendrolimus punctatus* incidence rate based on meteorological factors. *Forest Research*, 29(2): 256–260. [王庆, 毕猛, 杜婷, 2016. 基于气象因子的马尾松毛虫发生率空间格局研究. 林业科学研究, 29(2): 256–260.]
- Xu ZH, Li CH, Liu J, 2014. Fisher discriminant analysis of *Dendrolimus punctatus* Walker pest levels. *Transactions of the Chinese Society for Agricultural Machinery*, 45(6): 401–406. [许章华, 李聪慧, 刘健, 2014. 马尾松毛虫害等级的 Fisher 判别分析. 农业机械学报, 45(6): 401–406.]
- Zhang AB, Chen J, Wang ZJ, Li DM, Tian J, 2001. The application of BP model and logit model to prediction of forest insect pests. *Acta Ecologica Sinica*, 21(12): 2159–2156. [张爱兵, 陈建, 王正军, 李典谟, 田洁, 2001. BP 网络模型和 LOGIT 模型在森林害虫测报上的应用初报. 生态学报, 21(12): 2159–2156.]
- Zhang XX, 2006. Insect Ecology and Forecast. (3<sup>rd</sup> edition). Beijing: China Agriculture Press. 205–221. [张孝羲, 2006. 昆虫生态及预测预报(第三版). 北京: 中国农业出版社. 205–221.]
- Zhang Z, Li DM, 2008. Approach to outbreak mechanism of *Dendrolimus punctatus* (Lepidoptera : Lasiocampidae). *Scientia Silvae Sinicae*, 44(1): 140–150. [张真, 李典谟, 2008. 马尾松毛虫暴发机制分析. 林业科学, 44(1): 140–150.]
- Zhang WJ, Zhong XQ, Liu GH, 2008. Recognizing spatial distribution patterns of grassland insects: neural network approaches. *Stoch Environ. Res. Risk Assess*, 22(2): 207–216.
- Zou YD, Cheng FJ, Zha GJ, 1990. Effect of the chemical components of pine needle on the existence and growth of pine caterpillars *Dendrolimus punctatus*. *Scientia Silvae Sinicae*, 26(2): 142–148. [邹运鼎, 程扶玖, 查光济, 1990. 松针内含物与马尾松毛虫生存发育关系的研究. 林业科学, 26(2): 142–148.]
- Zou YD, Wang HF, 1989. Insect Ecology of Agriculture and Forestry. Heifei: Anhui Science and Technology Press. 462–469. [邹运鼎, 王弘法, 1989. 农林昆虫生态学. 合肥: 安徽科学技术出版社. 462–469.]