技术与方法

马尾松毛虫幼虫发生高峰期的三种预测模型*

余 燕[1**] 王振兴[1] 李 尚[2] 周夏芝[2***] 毕守东[1***] 方国飞[3]
张国庆[4] 邹运鼎[2] 张 桢[4] 宋玉双[3]

（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. 国家林业局森林病虫害防治总站，沈阳 110034；4. 安徽省潜山县林业局，潜山 246300）

摘 要 **【目的】**为了科学确定马尾松毛虫 *Dendrolimus punctatus* 防治适期，提高防治效果。**【方法】** 采用平稳时间序列法、马尔科夫链法和 BP 神经网络法研究安徽省潜山县 1983-2014 年马尾松毛虫幼虫发生的高峰期预测模型，并用 2015 年和 2016 年发生资料进行验证。**【结果】** 显示：平稳时间序列法、马尔科夫链法、BP 神经网络法模型预测 2015 年和 2016 年 1 代幼虫高峰期均为 6 月 5 日，2 代高峰期均为 9 月 6 日，两年的 1、2 代实际发生期也是 6 月 5 日和 9 月 6 日，预测值与实际值完全一致。平稳时间序列法预测 1 代幼虫高峰期结果，若以大于 2 d 为误差标准，则 1983-2014 年的历史符合率为 96.77%；若以小于和等于 1 d 为误差标准，历史符合率为 74.19%，2 代幼虫高峰的预测结果的历史符合率与 1 代相同。BP 神经网络法预测结果若以 1 d 为误差标准，1983-2014 年则 1、2 代预测结果的历史符合率均为 100%。**【结论】**在 3 种方法中，平稳时间序列法适用于害虫发生过程符合平稳时间序列的情况；马尔科夫链法预报量的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，平稳时间序列法，BP 神经网络法，马尔科夫链法

The three models for forecasting the peak larval period of *Dendrolimus punctatus*

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

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

Abstract **[Objectives]** To scientifically determine the optimum period for controlling *Dendrolimus punctatus*, and thereby improve the effectiveness of current control methods. **[Methods]** Stationary time series, BP neural network and Markov Chain analysis were used to develop models predicting the timing of peak larval abundance from 1983 to 2016 in Qianshan County Anhui Province. The predictions of these models were verified in 2015 and 2016. **[Results]** The peak of the first larval generation was predicted to occur on June 5 and that of the second generation on September 6. The predicted dates were exactly the same as the actual dates in 2015 and 2016. The stationary time sequence method predicted the timing of peak of larval abundance with 96.77% accuracy from 1983 to 2014 if the error criterion was > 2 days. If the error criterion is less ≤1, the accuracy was 74.19%. The accuracy of the prediction of the second generation larval peak was the same as that for the first generation. For the BP neural network, if the error criterion was one day, the accuracy of the predicted outcome was 100%

*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：544896437@qq.com
***共同通讯作者 Co-corresponding authors，E-mail：zhouxz@ahau.edu.cn；bishoudong@163.com
收稿日期 Received：2017-07-14，接受日期 Accepted：2018-04-17

from 1983 to 2014. **[Conclusion]** The stationary time sequence method was applicable to stable time series for predicting peak larval abundance whereas the Markov Chain was directly influenced by the accuracy of the prediction results. The BP Neural Network method can be used to determine nonlinear relationships between independent variables and predicted larval abundance, which is an ideal prediction method.

Key words the larvae peak period of *Dendrolimus punctatus*, periodical and regression method, stationary time series method and BP neural network, Markov chains and contingency table analysis method

马尾松毛虫 *Dendrolimus punctatus* Walker，广泛分布于安徽、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*、晚松 *Pinus rigidavar* Serotina、海南松 *Pinus fenzeliana* 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，经济效益和生态效益损失巨大。再者，人们的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，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，加之科学地运用综合治理措施，使马尾松毛虫危害得到有效控制，但该虫具有极强的繁殖潜力，遇到有利的生态环境极易暴发成灾，对其监测仍不能放松警惕。

马尾松毛虫发生 1-5 代/年，随不同地区而异，河南信阳地区发生 2 代/年，长江流域发生 2-3 代/年，广东、广西、福建南部发生 3-4 代/年，海南发生 4-5 代/年，安徽省潜山县发生 3 代/年。马尾松毛虫发生预测预报是对其进行综合防治的基本工作之一。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；许章华等，2014；费海泽等，2014；王庆等，2016）等分别采用不同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，为马尾松毛虫综合防治工作提供了有力支持。发生期是决定防治适期的基础，幼虫发生高峰期直接与防治有关，但有关幼虫高峰期报道极少，由于各地气象条件、植被条件和地形地貌等不同，马尾松毛虫发生特点也不尽相同。过去发生期预测通常采用累积发育进度预测法（历期预测法），加权平均累计发育进度历期预测法、虫量比例历期预测法、期距预测法、有效积温预测法和回归预测法，随着科学进步和电子计算机的应用，新方法不断涌现，大大提高了其预测结果的准确性（邹运鼎和王弘法，1989）。为有效地防治马尾松毛虫，本研究采用平稳时间序列法、人工 BP 神经网络法和马尔科夫链法预测安徽省潜山县马尾松毛虫幼虫的 1 代和 2 代发生高峰期，并对几种方法预测结果的优缺点进行比较，以期为马尾松毛虫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国家森林病虫测报中心之一的安徽省潜山县森林病虫防治站，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。调查采用国家林业局森林病虫防治总站制定的森林害虫测报方法，即踏查与详查相结合的办法，对不同的虫态采取不同的观察方法，发现虫情立即设标准地。采用平行线抽样法抽取 20 株标准株详查。

卵期调查，在雌蛾羽化高峰后 1-2 d，调查平均卵块数和卵块平均卵粒数；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 以下的树直接调查统计树冠上的幼虫数，大树用“虫粪粒推算法”调查。越冬期幼虫，调查树干基部树皮缝中幼虫数推算全株虫口。蛹期调查，在结茧盛期后 2-5 d 剖茧，调查雌雄性比，平均雌蛹重死亡率和寄生率。成虫

期调查，在成虫羽化前 2-3 d 及羽化期内用黑光灯诱杀，统计其诱杀数量和雌性性比。

1.2 建模方法

1.2.1 平稳时间序列预测法 在害虫预测预报中，把按时间顺序产生和排列的观察数据序列称为时间序列，它可能是历年某种害虫某一世代的发生数量，也可能是害虫对作物的为危害程度或发生时间，这些数据的特点是其观测值排列顺序的重要性，正是这些数据的顺序与大小反映了数据所包含的信息以及数据内部的相互联系，这些数据的重要性在于这些观察值之间的依赖关系或相关性。分析这种相依有序的数据集合的方法，即时间序列分析。害虫发生期的早迟，两个相邻的发生期间的时间长度在一定时间内彼此间存在一定相关性，在未来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（邹运鼎和王弘法，1989）。

对害虫发生期 $X(t)$ 的随机函数，若相关函数 $R_x(t_1t_2)$ 在一定时间的值（或转移概率），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，则称此随机过程关于时间为齐性，即时齐性，而其序列 $X(t_1)$，$X(t_2),\cdots,X(t_m)$ 为平稳时间序列，

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

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

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

其中 t_0 代表资料时间总长度。

在害虫发生的实际随机过程中有时存在非平稳的时间序列，对这类时间序列可以用第一次差分的 Δx_t 或第二次差分 Δ^2x_t 的方法使之变成平稳时间序列，$\Delta x_t=x_t-x_{t-1}$。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

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

其中

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

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

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

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

一般采用此法的时间跨度要有足够长，本文是 N=31 年的历史资料，N 为数据的个数，本文为年数，建立其预测模型为:

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

式中 x_t，x_{t-3}，x_{t-2}，x_{t-1} 为第 t 年，t－3 年，t－2 年，t－1 年某代幼虫发生高峰期的分级值。

1.2.2 BP 神经网络法 神经网络的模式很多（Rumelhart and McClellan，1986；Lippman *et al*.,1987:Homik *et al*.,1989:Mhaskar and Miccheii，1992；李祚泳和彭荔红，1999；汪四水和张孝羲，2000；陈绘画，2003；汪四水，2003）

目前应用最为普遍的是 BP 型神经网络（Back propation，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 $I=1,2,\cdots,M$；$j=1,2,\cdots H$；$k=1,2,\cdots,N$）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 θ 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 $f(x)=1/(1+e^{-x})$。

③ 输出单元的输入和输出：

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

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,⋯, 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，满足这一步要求权重的调整量为：$\Delta w = -\eta \frac{\partial E}{\partial W}$，式中 η 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：$\Delta_R W w_{jk}^{(2)} = \eta \delta_k^R \times O_j^R$，其中 $\delta_k^R = -f'(net_k^R)(Z_k^R - O_k^R)$，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 $\delta_j^R = -f'(net_j^R)\sum_{k=1}^{n}\delta_k^R W_{jk}^{(2)}$，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。

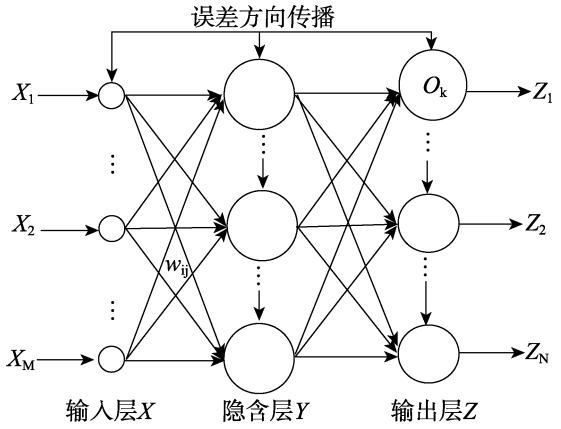

图 1　BP 网络结构原理图

Fig. 1　Structure of BP neural network

网络输入层神经元节点数就是系统的特征因子（自变量）个数，输出层神经元节点数就是系统目标个数，即预报量的个数，隐含层节点数根据经验选取，一般设为输入层节点数的 75%。X_i 表示输入层第 i 个节点的输入，即 x_i 为与某代幼虫高峰期有关的气象因子或前一虫态的发生期，w_{ij} 为隐含层第 j 个节点到输入层第 i 个节点的权值，Z_j 表示隐含层第 j 个节点的阈值，（j=1,2,⋯,m）；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，即本文的马尾松毛虫幼虫某世代的发生高峰期。

1.2.3　马尔科夫链方法　马尔可夫链是一种最简单的马尔科夫过程，它涉及到马尔科夫过程和状态转移概率。昆虫的一生有不同虫态即不同状态，害虫发生严重程度也分为轻、中、重等状态。害虫发生时期可分为早、中、迟等状态。各阶段的状态往往是从前一状态转变过来的，这即为状态转移。随着时间的变化所作的状态转移即为状态转移过程，若每次状态转移却只与相互接引的前一次有关，与过去的状态无关，即无后效性的随机过程，该状态转移过程即马尔科夫过程，从某种状态出发，下一步转移到其它状态的可能性即称为状态转移概率。如果此概率与所处的开始时间无关，而只与状态和时间间隔有关，则称此随机过程关于时间为齐性，即时齐性，马尔可夫链研究的范围就是局限在其有时齐性的马尔科夫过程。在马尾松毛虫年度间发生序列中，将其幼虫高峰期分为若干有限等级。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，$E_2, \cdots, E_i$ 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，马尾松毛虫年度间某世代幼虫高峰期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 为马尾松毛虫幼虫高峰期等级为 E_i 的总次数；$N_j(m)$ 为马尾松毛虫幼虫高峰期等级由 E_i 转移到 E_j 的次数。m 阶转移概率矩阵为

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

预报时，以预报年相距 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 $p(3)$，$p11$ 为在若干个从“ 1 ”级出发转移出去的状态，转移中有一定数量是从“ 1 ”级转入下一年的“ 1”级的概率，$p12$ 同理。计算出转移概率总和，以转移概率总和最大的那个状态级别作为预报年某代幼虫高峰期可能发生级别的预测值。

2 结果与分析

根据安徽省潜山县 1983-2014 年的资料进行建模分析，用 2015 年和 2016 年资料验证模型的准确性。

安徽省潜山县马尾松毛虫一年发生 3 代，以幼虫越冬，1-6 月上旬为越冬代，6 月上旬-8 月中旬为 1 代，其中 5 月下旬-6 月上旬为卵期，6 月上旬-7 月中旬为幼虫期，7 月中旬-8 月上旬为蛹期，8 月上旬-8 月中旬为成虫期，8 月中旬-12 月为 2 代，其中 8 月中旬-9 月上旬为卵期，9 月上旬-10 月份为幼虫期。1983-2014 年 1 代幼虫高峰期的范围为 6 月 5 日-12 日，2 代幼虫高峰期范围为 9 月 6 日-9 月 11 日，现将 1983-2016 年安徽省潜山县马尾松毛虫幼虫高峰期资料列于表 1。

表 1 马尾松毛虫幼虫高峰期（潜山县）

Table 1 The larvae peak period of *Dendrolimus punctatus* (Qianshan county)

年份 Year	1 代幼虫高峰期（月.日）The first generation of larvae (month.day)	2 代幼虫高峰期（月.日）The second generation of larvae (month.day)	年份 Year	1 代幼虫高峰期（月.日）The first generation of larvae (month.day)	2 代幼虫高峰期（月.日）The second generation of larvae (month.day)
1983	6.12	9.11	2000	6.6	9.6
1984	6.12	9.11	2001	6.6	9.6
1985	6.12	9.11	2002	6.6	9.6
1986	6.12	9.11	2003	6.6	9.6
1987	6.12	9.11	2004	6.6	9.6
1988	6.12	9.11	2005	6.6	9.6
1989	6.10	9.10	2006	6.5	9.6
1990	6.10	9.10	2007	6.5	9.6
1991	6.10	9.10	2008	6.5	9.6
1992	6.9	9.10	2009	6.5	9.6
1993	6.9	9.8	2010	6.5	9.6
1994	6.8	9.7	2011	6.5	9.6
1995	6.8	9.7	2012	6.5	9.6
1996	6.8	9.7	2013	6.5	9.6
1997	6.8	9.7	2014	6.5	9.6
1998	6.8	9.7	2015	6.5	9.6
1999	6.8	9.7	2016	6.5	9.6

6 月 4 日为-4，5 日为-3，6 日为-2，7 日为-1，8 日为 0，9 日为 1，10 日为 2，11 日为 3，12 日为 4。
June 4th for -4, 5th for -3, 6th for -2, 7th for -1, 8th for 0, 9th for 1, 10th for 2, 11th for 3, 12th for 4.

2.1 用时间平稳序列法预测幼虫高峰期模型

2.1.1 1 代幼虫高峰期预测模型 将 1983-2014 年 1 代幼虫高峰期的发生期进行分级并将其列于表 2，1998 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=6.774 2，

表 2　1 代幼虫高峰期的 X_t 值（分级值）和预测值（X_t'）

Table 2　The value of X_t (grading value) and predictive value of X_t' for the first generation larvae peak period

年份 Year	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
X_t	4	4	4	4	4	4	2	2	2	1	1
X_t'	4	4	4	4	4	4	3.72	2.59	2.05	1.86	1.29
年份 Year	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
X_t	0	0	0	0	0	0	-2	-2	-2	-2	-2
X_t'	1.02	0.37	0	0	0	0	0	-1.13	-1.67	-1.86	-1.86
年份 Year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	
X_t	-2	-2	-3	-3	-3	-3	-3	-3	-3	-3	
X_t'	-1.86	-1.86	-1.86	-2.42	-2.67	-2.79	-2.79	-2.79	-2.79	-2.79	

R_1=6，R_2=5.793 1，R_3=5.535 7，代入（3），则预测模型为 X_t=0.092 9X_{t-3}+0.272 7X_{t-2}+0.564 7X_{t-1}

将表中 2012、2013 和 2014 年 1 代幼虫高峰期的 X_t 值代入预测模型，求出 2015 年为 -2.790 9≈-3，即 1 代幼虫高峰期日为 6 月 5 日，与实际相符。将 2013、2014 和 2015 年 1 代幼虫高峰期的 X_t 值代入预测模型，求出 2016 年为 -2.790 9≈-3，即 2016 年 1 代幼虫高峰期日为 6 月 5 日，预测结果与实际完全相符。1983-2014 年的历史符合率若以大于和等于 2 d 为误差标准时为 96.77%，若以小于 1 d 为误差标准则为 74.19%。

2.1.2　2 代幼虫高峰期预测模型　将 1983-2014 年 2 代幼虫高峰期分级后列于表 3，1998 年因缺如，用内插法进行估值，运用（4）、（5）、（6）、（7）式求得 R_0=4.354 8，R_1=4.166 7，R_2=3.896 6，R_3=3.607 1，代入方程组（3）式，则预测模型为 X_t=0.934 4X_{t-3}+0.401 7X_{t-2}－0.332 3X_{t-1}。

将表中 2012、2013 和 2014 年的 2 代幼虫高峰期 X_t 值代入上式，求出 2015 年的 X_t 值为 -2.007 6≈-2，即 9 月 6 日，实际发生期为 9 月 6 日，将 2013、2014 和 2015 年 2 代幼虫高峰期的 X_t 值代入上式，求出 X_{2016}= -2.005 1≈-2，即 9 月 6 日，预测值与实际值完全吻合。若误差标准为大于和等于 2 d，则 1983-2014 年的历史符合率为 96.77%，若以小于和等于 1 d 为误差标准则历史符合率为 74.19%。

表 3　2 代幼虫高峰期的 X_t 值（分级值）和预测值（X_t'）

Table 3　The value of X_t (grading value) and predictive value of X_t' for the second generation larvae peak period

年份 Year	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
X_t	3	3	3	3	3	3	2	2	2	2	0
X_t'	3	3	3	3.01	3.01	3.01	3.01	3.34	2.94	2.01	2.01
年份 Year	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
X_t	-1	-1	-1	-1	-1	-1	-2	-2	-2	-2	-2
X_t'	2.67	-2.20	-0.07	-1.00	-1.00	-1.00	-1.00	-0.67	-1.07	-2.01	-2.01
年份 Year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	
X_t	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	
X_t'	-2.01	-2.01	-2.01	-2.01	-2.01	-2.01	-2.01	-2.01	-2.01	-2.01	

9 月 5 日为-3，6 日为-2，7 日为-1，8 日为 0，9 日为 1，10 日为 2，11 日为 3。

September 5th for -3, 6th for -2, 7th for -1, 8th for 0, 9th for 1, 10th for 2, 11th for 3.

2.2 用 BP 神经网络法预测马尾松毛虫幼虫高峰期

2.2.1 1 代幼虫高峰期的 BP 神经网络预测模型

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进行预报结果验证。对 3-5 月的旬均温、旬降雨量等 85 个预报因子利用相关系数筛选出 4 月上旬均温（x_1）、4 月中旬均温（x_2）和 5 月中旬均温（x_3）和 1 代幼虫始盛期（x_4）等 4 个因子作为输入样本，采用 4-4-1 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，结果列于表 4，学习样本中若以 1 d 为误差标准，预测正确率为 100%。2015 年和 2016 年预测值分别为 66.20 和 66.70，实际值都是 66，若以 1 d 作为标准误差，预测准确率均为 100%。

2.2.2 2 代幼虫高峰期的 BP 神经网络预测模型

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进行预报结果验证。对 6-9 月上旬的旬均温、旬降雨量等 95 个预报因子利用相关系数筛选出 6 月中旬均温（x_5）、6 月下旬均温（x_6）和 7 月上旬均温

表 4 马尾松毛虫幼虫高峰期的人工 BP 神经网络预测结果

Table 4 The prediction results by using BP neural network of the *Dendrolimus punctatus* larvae peak period

年份 Year			1983	1984	1985	196	1987	1988	1989	1900
a		日期（月/日）Date (month/day)	6/12	6/12	6/12	6/12	6/12	6/12	6/10	6/10
	实际值 Actual value	数值 Value	73	73	73	73	73	73	71	71
	预测值 Predictive value	数值 Value	72.75	72.85	72.89	72.61	72.33	72.52	71.13	71.83
		误差 Error	- 0.25	- 0.15	- 0.11	- 0.39	- 0.67	- 0.48	0.13	0.83
b		日期（月/日）Date (month/day)	9/11	9/11	9/11	9/11	9/11	9/11	9/10/	9/10
	实际值 Actual value	数值 Value	164	164	164	164	164	164	163	163
	预测值 Predictive value	数值 Value	163.54	163.78	163.77	163.71	163.69	163.71	163.1	163.03
		误差 Error	- 0.46	- 0.22	- 0.23	- 0.29	- 0.31	- 0.29	0.1	0.03
年份 Year			1991	1992	1993	1994	1995	1996	1997	1999
a		日期（月/日）Date (month/day)	6/10	6/9	6/9	6/8	6/8	6/8	6/8	6/8
	实际值 Actual value	数值 Value	71	70	70	69	69	69	69	69
	预测值 Predictive value	数值 Value	70.97	69.97	70.29	69.04	68.94	69.18	69.08	68.75
		误差 Error	0.03	0.03	- 0.29	- 0.04	0.06	- 0.18	- 0.08	0.25
b		日期（月/日）Date (month/day)	9/10	9/10	9/8	9/7	9/7	9/7	9/7	9/7
	实际值 Actual value	数值 Value	163	163	161	160	160	160	160	160
	预测值 Predictive value	数值 Value	162.72	163	160.85	159.91	160.68	160.08	159.65	159.56
		误差 Error	0.28	0	0.15	0.09	- 0.68	- 0.08	0.35	0.44

续表 4（**Table 4 continued**）

年份 Year			2000	2001	2002	2003	2004	2005	2006	2007
a		日期（月/日）Date (month/day)	6/6	6/6	6/6	6/6	6/6	6/6	6/6	6/5
	实际值 actual value	数值 Value	67	67	67	67	67	67	67	66
	预测值 Predictive value	数值 Value	66.77	66.88	67.26	66.39	66.62	66.84	66.56	66.25
		误差 Error	0.23	0.12	-0.26	0.61	0.38	0.16	0.44	-0.25
b		日期（月/日）Date (month/day)	9/6	9/6	9/6	9/6	9/6	9/6	9/6	9/6
	实际值 actual value	数值 Value	159	159	159	159	159	159	159	159
	预测值 Predictive value	数值 Value	159.09	159.10	159.06	159.09	159.03	159.10	159.07	159.03
		误差 Error	－0.09	－0.10	－0.06	－0.09	－0.03	－0.10	－0.07	－0.03

年份 Year			2008	2009	2010	2011	2012	2013	2014
a		日期（月/日）Date (month/day)	6/5	6/5	6/5	6/5	6/5	6/5	6/5
	实际值 actual value	数值 Value	66	66	66	66	66	66	66
	预测值 Predictive value	数值 Value	66.36	66.28	66.13	66.21	66.43	66.4	66.49
		误差 Error	－0.36	－0.28	－0.13	－0.21	－0.43	－0.40	－0.49
b		日期（月/日）Date (month/day)	9/6	9/6	9/6	9/6	9/6	9/6	9/6
	实际值 actual value	数值 Value	159	159	159	159	159	159	159
	预测值 Predictive value	数值 Value	159.01	159.09	159.08	159.00	159.02	159.02	159.02
		误差 Error	－0.01	－0.09	－0.08	0	－0.02	－0.02	－0.02

a, b：发生日期数值化时，4 月 1 日为 1，即 5 月 1 日为 31，依次类推。

a, b: Occurrence date numeric，April 1st for 1, May 1st for 31, so on.

（x_7）和 2 代幼虫始盛期（x_8）等 4 个因子作为输入样本，采用 4-4-1 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，结果也列于表 4，学习样本中若以大于 1 d 为误差标准，预测正确率为 100%。2015 年和 2016 年作为预测样本，预测值分别为 159.00 和 159.01（即 9 月 6 日），实际发生值也是 9 月 6 日，若以 1 d 作为标准误差，预测准确率均为 100%。

2.3　利用马尔科夫链方法预测马尾松毛虫幼虫高峰期

对马尾松毛虫幼虫高峰期进行分级，分级标准值列于表 5，1983-2014 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 代幼虫高峰期的转移概率矩阵如下：

$$p(1)=\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0.17 & 0.83 & 0 & 0 \\ 0 & 0.2 & 0.8 & 0 \\ 0 & 0 & 0.17 & 0.83 \end{vmatrix}$$

$$p(2)=\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0.33 & 0.67 & 0 & 0 \\ 0 & 0.4 & 0.6 & 0 \\ 0 & 0 & 0.33 & 0.67 \end{vmatrix}$$

$$p(3)=\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0.6 & 0.4 & 0 \\ 0 & 0 & 0.5 & 0.5 \end{vmatrix}$$

2 代幼虫高峰期的转移概率矩阵如下：

$$p(1)=\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0.17 & 0.83 & 0 & 0 \\ 0 & 0.25 & 0.75 & 0 \\ 0 & 0 & 0.17 & 0.83 \end{vmatrix}$$

$$p(2)=\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0.33 & 0.67 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0.33 & 0.67 \end{vmatrix}$$

$$p(3)=\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0.75 & 0.25 & 0 \\ 0 & 0 & 0.5 & 0.5 \end{vmatrix}$$

表 5　马尾松毛虫幼虫高峰期分级值

Table 5　The grading value of the *Dendrolimus punctatus* larvae peak period

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

a：1 代幼虫高峰期分级标准为：66-67（6 月 5 日-6 月 6 日）1 级；68-69（6 月 7 日-6 月 8 日）2 级；70-71（6 月 9 日-6 月 10 日）3 级；72-73（6 月 11 日-6 月 12 日）4 级。b：2 代幼虫高峰期分级标准为：159（9 月 6 日）1 级；160-161（9 月 7 日-9 月 8 日）2 级；162-163（9 月 9 日-9 月 10 日）3 级；164（9 月 11 日）4 级。2015 年和 2016 年括号内数字为预测的数值。发生期的数值化是令 4 月 1 日为 1，则 5 月 1 日为 31，6 月 1 日为 62，余类推。

a: The peak stage grading standard of the first generation of larvae: 66-67(June 5th - June 6th) level 1; 68-69(June 5th- June 8th) level 2; 70-71(June 9th- June 10th) level 3; 72-73(June 11th- June 12th) level 4.

b: The peak stage grading standard of the second generation of larvae: 159(September 6th) level 1; 160-161(September 7th- September 8th) level 2; 162-163(September 9th- September 10th) level 3; 164(September 11th) level 4. The figures in brackets for 2015 and 2016 are the predicted values. Numeralization of the occurrence period: April 1st for 1, May 1st for 31, June 1st for 62, and so on.

列出与 2015 年相距 1、2、3 年马尾松毛虫发生级值，依据转移概率值总和的最大值为预报值（表 6）。

由表 6 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 代幼虫高峰期转移概率合计均为 3，即对应的为“1”级，故预

表 6　2015 年马尾松毛虫幼虫高峰期预报
Table 6　The prediction of the *Dendrolimus punctatus* larvae peak period in 2015

年份 Year	起始状态 Initial state	转移步数 Transfer steps	1 代幼虫高峰期转移概率 A generation of caterpillars transfer probability				2 代幼虫高峰期转移概率 The second generation of caterpillars transfer probability			
			1	2	3	4	1	2	3	4
2014	1	1	1	0	0	0	1.00	0.00	0.00	0.00
2013	1	2	1	0	0	0	1.00	0.00	0.00	0.00
2012	1	3	1	0	0	0	1.00	0.00	0.00	0.00
合计 Total			3.00	0.00	0.00	0.00	3.00	0.00	0.00	0.00

报 2015 年马尾松毛虫 1 代、2 代幼虫高峰期也均为 1 级。为了预测 2016 年马尾松毛虫的幼虫高峰期，将 2015 的发生程度级列入基础数据，重新组建转移矩阵，计算结果见表 6，根据马尔科夫链法预测 2016 年马尾松毛虫的 1 代、2 代幼虫高峰期也均为 1 级，根据级数划分 2015 年和 2016 年 1、2 代幼虫高峰期预测值为 1 级，实际值也为 1 级，两者完全一致。

3　小结与讨论

为了科学确定马尾松毛虫防治适期，本文分别采用平稳时间序列法、马尔科夫链法和 BP 神经网络法对安徽省潜山县 1983-2014 年马尾松毛虫资料进行研究建立幼虫高峰期预测模型，并用 2015 年和 2016 年发生资料进行验证。

（1）时间平稳序列法预测 2015 年、2016 年 1 代、2 代幼虫高峰期均为 6 月 5 日和 9 月 6 日。1、2 代幼虫高峰期的预测结果，若以大于和等于 2 d 为误差标准，1983-2014 年的历史符合率为 96.77%。若以小于和等于 1 d 为误差标准，则历史符合率为 74.19%。

（2）马尔科夫链法预测 2015 年、2016 年 1 代、2 代幼虫高峰期均为 6 月 5 日和 9 月 6 日。

（3）BP 神经网络法预测 2015 年、2016 年 1 代、2 代幼虫高峰期均为 6 月 5 日和 9 月 6 日，若以 1 d 为误差标准，1983-2014 年以该法 1、2 代预测结果的历史符合率均为 100%，2015 年和 2016 年 3 种预测方法预测结果与实际完全符合。

上述 3 种方法中，马尔科夫链法预测结果准确与否取决于预报量选什么样的分级标准，相同的资料选择不同的分级标准其预测结果大相径庭。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，其优点是从表面上看，时间序列分析撇开了环境因子等系统变量对害虫发生期的影响，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错综复杂或有关数据资料无法得到时，时间序列分析法就是理想的分析方法。该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。BP 神经网络法不仅具有处理多因子的功能，而且还有很强的非线性处理能力，其建模的过程比较简单：只要对样本数据进行学习，模型就会自动找到最佳刻画样本数据的函数。这样克服了由于刻画函数太复杂或难以想象而造成的研究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，选择一些合适的因子对提高预测精度将起十分重要的作用，再者是网络结构的选取，三层的神经网络结构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，尚需要进一步的探讨，总之 BP 神经网络法是一种值得大力推广的预测方法。

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 punctatu. Journal of Northeast Forestry*

Uniwersity, 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.]

Chen HH, Zhu SY, Cui XF, 2003. A study on the forecast model of *Dendrolimus punctatus* occurrence based on artificial neural network. *Forest Research*, 16(2): 159–165. [陈绘画, 朱寿燕, 崔相富, 2003. 基于人工神经网络的马尾松毛虫发生量预测模型的研究. 林业科学研究, 16(2): 159–165.]

Fei HZ, Wang HB, Kong XB, 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, 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.]

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.]

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. 科学出版. 188–191.]

Jia CS, 2006. Forecast of occurrence level of *Dendrolimus punctatus* by Markov Chains. *Journal of Northeast ForestryUniwersity*, 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.]

Lippmann RP, 1987. An introduction to computing with neural nets. *Institute of Electrical and Electronics Engineers*, 4(2): 4–22.

Maskar HN, Micchelli CA, 1992. Approximation by superposition of sigmoid and radial basis functions. *Adv. Appl. Math.*, (3): 350–373.

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.

Rumelhart DE, McClellan JL, 1986. Parallel Distributed Processing-Exploration in the Microst ructure of Cognition, Vol. 1 and Vol.2. Cambridge: MIT Press. 163–174.

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 (2nd edition). Beijing: Chinese Academy of Forestry. 948–953. [萧刚柔, 1992. 中国森林害虫(第 2 版). 北京: 林业出版社. 948–953.]

Wang SS, Zhang XX, Zhang XL, 2003. Predictive warning system of neural for occurrence level of rice leaf roller. *Journal of Biomathematics*, 18(1): 93–97. [汪四水, 张孝羲, 张夕林, 2003. 稻纵卷叶螟发生程度的神经网络预警. 生物数学学报, 18(1): 93–97.]

Wang SS, Zhang XX, 2000. A neutral network appoach 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. (3rd edition). Beijing: China Agriculture Press. 205–221. [张孝羲, 2006. 昆虫生态及预测预报(第 3 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 Y D, Cheng FJ, Zha GJ, 1990. Effect of the chemical components of pine needle on the existence and growth of pine caterplilars *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. Hefei: 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.]