

DOI: 10.13718/j.cnki.xdzk.2017.08.006

# 基于马尔科夫链理论对 贵州锦屏白背飞虱发生程度的预测<sup>①</sup>

薛文鹏<sup>1,2</sup>, 杨 芮<sup>3</sup>, 杨 洪<sup>1</sup>, 杨 林<sup>4</sup>

1. 贵州大学 昆虫研究所, 贵阳 550025; 2. 贵阳市植保植检站, 贵阳 550081;  
3. 贵阳市林业有害生物防治检疫站, 贵阳 550002; 4. 贵州锦屏县植保植检站, 贵州 锦屏 556000

**摘要:** 白背飞虱是我国水稻上的重要害虫, 以贵州一季中稻区白背飞虱为研究对象, 运用马尔科夫链理论, 以代表站点锦屏 1990—2011 年的白背飞虱田间发生程度的时间序列资料, 构建了 1~5 阶的概率转移矩阵, 以预报年前 5 年的连续发生级别预测第 6 年的发生状态. 回报锦屏 17 年的结果, 得到了较高的历史符合率, 达到了 82%, 对锦屏 2012 年白背飞虱田间发生级别的预测结果也符合当年的实际发生情况.

**关键词:** 白背飞虱; 预测预报; 马尔科夫链; 发生程度

**中图分类号:** S431.2

**文献标志码:** A

**文章编号:** 1673-9868(2017)08-0043-06

作为一种迁飞昆虫, 白背飞虱 *Sogatella furcifera* (Horváth) 每年春季随西南气流由中南半岛进入我国, 而后逐代向北迁飞, 至秋季又从北往南回迁<sup>[1-2]</sup>, 同时我国的稻飞虱也可向日本、韩国迁飞<sup>[3-4]</sup>, 在迁飞过程中, 昆虫遇到降雨、下沉气流、风向等外力作用时会被集中迫降; 飞行能源物质耗尽或温度骤变低于其飞行温度阈值时也会选择主动降落<sup>[5]</sup>, 这种难以预测的大规模降虫, 给预测预报工作带来了很大困难, 加之气候条件和种植制度等因素的影响, 其在田间的发生情况具有较大的随机波动性<sup>[6]</sup>, 而马尔科夫链适用于预测分析具有多种状态的随机波动性较大的时间序列的预测<sup>[7]</sup>.

在贵州省各个稻区中, 黔南稻区和黔东稻区白背飞虱发生较为严重<sup>[1, 8]</sup>, 近年来对贵州主要稻区代表站点的系统调查结果同样如此<sup>[9]</sup>, 基于此, 本文以黔东南代表站点锦屏县 1990—2011 年田间发生量的历史资料建立了马尔科夫链模型, 对历史数据进行了回报检验, 并运用该模型对锦屏 2012 年的发生情况进行了预测.

## 1 材料与方法

### 1.1 数据来源及转换

锦屏县 1990—2011 年田间发生量数据来源于锦屏县植保植检站的田间系统调查, 调查方法采用盘拍法, 每年的 5 月 30 日—8 月 30 日每隔 5 d 调查 1 次, 用白瓷盘(33 cm×45 cm)进行平行跳跃式取样, 每块田调查 10 点, 每点查 2~5 丛(视虫量而定, 虫多少查, 虫少多查), 记录白背飞虱的数量, 然后折

① 收稿日期: 2015-08-25

基金项目: 贵州省农业科技攻关项目(黔科合 NY 字[2013]3006); 贵州省教育厅自然科学研究项目(黔教 2010011); 贵州省农业科技攻关项目(黔科合 NY 字[2010]3064); 973 计划前期研究专项(2009CB125908).

作者简介: 薛文鹏(1989-), 男, 河南孟州人, 硕士, 农艺师, 主要从事植保的研究.

通信作者: 杨 洪, 博士, 教授.

算成百丛虫量。

根据锦屏白背飞虱的实际发生情况,结合贵州省白背飞虱发生级别的划分办法,设定发生级别如下:白背飞虱年发生量总量小于 1.5 万头为 1 级,1.5~2.5 万头为 2 级,2.5~3.5 万头为 3 级,3.5~4.5 万头为 4 级,大于 4.5 万头为 5 级(取值包括前值)。

## 1.2 建模方法及过程

### 1.2.1 马尔科夫链预测原理与转移概率矩阵

马尔科夫链又叫概率转移法,是时间序列分析中的一种方法,其研究对象是某种事物状态的转移概率。世间的各种事物每时每刻都处于不同的状态,随着时间的变换,事物的状态也在不断的改变,而每次的转移都有一定概率<sup>[7]</sup>,后一时刻的状态往往是前一时刻的状态按一定概率转移过来的,这种从一种状态转入另一种状态的现象,称为状态转移。若每次的状态转移都只与互相接引的前一次有关,而与过去的状态无关即状态转移过程是无后效性的,这种状态转移过程就称为马尔科夫链<sup>[7]</sup>。马尔科夫链有两个基本特征,一是“无后效性”,即如前文所说;二是“遍历性”,指的是不管事物的状态如何,从各状态现在的实际转移概率开始,经过多少连锁转移后,各状态的最终概率趋向一个稳定的数值<sup>[7]</sup>。对害虫发生的时间序列,可将年度的发生程度区分为若干有限的等级或状态,序列中的每一个时间只能出现其中的一个状态  $E_i$  ( $i=1, 2, \dots, n$ ),根据“无后效性”特征,系统由状态  $E_i(k)$  经过  $k$  步 ( $k=1, 2, \dots, m$  阶,即间隔 1, 2,  $\dots, m$  年)转移到状态  $E_j(k)$  的概率只与前  $k$  步的状态  $E_i(k)$  有关;根据“遍历性”特征,经过  $m$  步转移,其最终的概率将趋于稳定。本研究将白背飞虱年发生程度的时间序列视为马尔科夫链,通过前 5 年的连续发生数据计算状态转移概率,预测第 6 年的发生程度<sup>[10-12]</sup>。

按照本文 1.1 中的划定规则将白背飞虱历年的发生数量转化为 5 个发生级别,对应马尔科夫链的 5 个状态,经过  $k$  年 ( $k=1, 2, 3, 4, 5$ ) 从状态  $E_i$  (状态即白背飞虱田间发生的 1, 2, 3, 4, 5 级) 转移到状态  $E_j$  的概率记为  $P_{ij}(k)$ ,具体说来经过 1 个转移过程从状态 1 到状态 1 的概率记为  $P_{11}(1)$ ,依此类推,则有  $P_{12}(1), P_{13}(1), P_{14}(1), P_{15}(1)$ , 经过 1 个转移过程从状态 2 到状态 1 的概率记为  $P_{21}(1)$ , 同样则有  $P_{22}(1), P_{23}(1), P_{24}(1), P_{25}(1)$ ; 再经由 1 个转移过程从状态 3, 状态 4, 状态 5 分别转移到状态 1, 状态 2, 状态 3, 状态 4, 状态 5, 得出相应概率, 这些概率组合起来构成的矩阵  $P(1)$  就称为 1 阶转移概率矩阵, 那么当经过  $k$  个转移过程时所得概率组成的矩阵  $P(k)$  就称为  $k$  阶转移概率矩阵, 每一行的概率值包含了从该起始状态转移到其他状态的全部概率, 其和为 1。

$$P(1) = \begin{pmatrix} P_{11}(1) & P_{12}(1) & \cdots & P_{15}(1) \\ P_{21}(1) & P_{22}(1) & \cdots & P_{25}(1) \\ \cdots & \cdots & \cdots & \cdots \\ P_{51}(1) & P_{52}(1) & \cdots & P_{55}(1) \end{pmatrix}$$

$$P(k) = \begin{pmatrix} P_{11}(k) & P_{12}(k) & \cdots & P_{15}(k) \\ P_{21}(k) & P_{22}(k) & \cdots & P_{25}(k) \\ \cdots & \cdots & \cdots & \cdots \\ P_{51}(k) & P_{52}(k) & \cdots & P_{55}(k) \end{pmatrix}$$

### 1.2.2 回报与预测

根据回报年或预报年的前 1, 2,  $\dots, m$  年的状态  $i$  分别取转移矩阵概率  $P(1), P(2), \dots, P(m)$  中的第  $i$  行的相应概率值, 建立概率预测矩阵, 该矩阵各列之和即为可能发生状态的  $j$  ( $j=1, 2, 3, 4, 5$ ) 的概率, 以概率最大值或与之最接近的状态作为回报或预测结果<sup>[13]</sup>。

## 2 结果与分析

### 2.1 白背飞虱发生级别的转移概率矩阵

按照 1.1 中的发生级别(状态)划分方法, 得到 1990—2011 年锦屏白背飞虱逐年田间发生级别, 见表 1, 并

依据此发生状态的时间序列建立 5 阶转移概率矩阵  $P_{JP}(1), P_{JP}(2), P_{JP}(3), P_{JP}(4), P_{JP}(5)$ , 详情如下:

$$P_{JP}(1) = \begin{pmatrix} 0.500 & 0.250 & 0.000 & 0.000 & 0.250 \\ 0.333 & 0.167 & 0.333 & 0.000 & 0.167 \\ 1.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 1.000 & 0.000 & 0.000 & 0.000 \\ 0.250 & 0.000 & 0.000 & 0.500 & 0.250 \end{pmatrix}$$

$$P_{JP}(2) = \begin{pmatrix} 0.250 & 0.125 & 0.125 & 0.125 & 0.375 \\ 0.600 & 0.200 & 0.200 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 1.000 \\ 0.500 & 0.500 & 0.000 & 0.000 & 0.000 \\ 0.250 & 0.500 & 0.000 & 0.250 & 0.000 \end{pmatrix}$$

$$P_{JP}(3) = \begin{pmatrix} 0.375 & 0.125 & 0.000 & 0.125 & 0.375 \\ 0.250 & 0.250 & 0.250 & 0.000 & 0.250 \\ 0.000 & 0.000 & 0.000 & 1.000 & 0.000 \\ 0.500 & 0.000 & 0.500 & 0.000 & 0.000 \\ 0.500 & 0.500 & 0.000 & 0.000 & 0.000 \end{pmatrix}$$

$$P_{JP}(4) = \begin{pmatrix} 0.375 & 0.125 & 0.000 & 0.125 & 0.375 \\ 0.500 & 0.000 & 0.000 & 0.250 & 0.250 \\ 0.000 & 1.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 1.000 & 0.000 & 0.000 & 0.000 \\ 0.500 & 0.250 & 0.250 & 0.000 & 0.000 \end{pmatrix}$$

$$P_{JP}(5) = \begin{pmatrix} 0.375 & 0.250 & 0.000 & 0.250 & 0.125 \\ 0.500 & 0.250 & 0.000 & 0.000 & 0.250 \\ 1.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 1.000 \\ 0.000 & 0.333 & 0.333 & 0.000 & 0.333 \end{pmatrix}$$

## 2.2 历史符合率

本文利用连续 5 阶转移概率矩阵计算历史符合率, 对锦屏 1995—2011 年共 17 年的历史资料进行回报检验, 以 1995 年白背飞虱发生情况为例, 其前 5 年发生级别依次为 1(1994 年), 3(1993 年), 2(1992 年), 1(1991 年), 2(1990 年) 级(表 1), 取上文中锦屏白背飞虱田间发生程度时间序列的 1~5 阶转移概率矩阵  $P_{JP}(k)$  相应行的元素, 得如下预测矩阵:

$$\begin{pmatrix} 0.500 & 0.000 & 0.250 & 0.375 & 0.500 \\ 0.250 & 0.000 & 0.250 & 0.125 & 0.250 \\ 0.000 & 0.000 & 0.250 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.125 & 0.000 \\ 0.250 & 1.000 & 0.250 & 0.375 & 0.250 \end{pmatrix}$$

预测矩阵各列之和的算数平均数 0.325, 0.175, 0.050, 0.025, 0.425 分别为 1989 年发生 1~5 级的平均概率, 其中发生 5 级的概率最大( $P=0.425$ ), 因此回报 1995 年发生级别为 5 级, 该回报与锦屏 1995 年实际发生程度相符.

用以上方法依次回报锦屏 1995—2011 年白背飞虱田间发生级别(表 1), 结果表明, 回报 17 年中仅有 3 年发生级别与实际发生级别不符: 2008 年实际发生 4 级, 回报级别 1 级, 误差 3 级; 2010 年实际发生 2 级, 回报级别 1 级, 误差 1 级; 2011 年实际发生 3 级, 回报级别 1 级, 误差 2 级. 其余年份的回报级别与实际发

生级别均相符, 17 年锦屏白背飞虱田间发生级别的历史回报符合率为 82%。

表 1 1990—2011 年锦屏白背飞虱田间发生级别的回报检验

年份	各 级 概 率					回报级别	实际级别	误差
	1	2	3	4	5			
1990							2	
1991							1	
1992							2	
1993							3	
1994							1	
1995	0.325	0.175	0.050	0.025	0.425	5	5	0
1996	0.271	0.075	0.025	0.425	0.200	4	4	0
1997	0.225	0.400	0.000	0.075	0.125	2	2	0
1998	0.467	0.258	0.092	0.025	0.108	1	1	0
1999	0.495	0.019	0.019	0.050	0.075	1	1	0
2000	0.200	0.392	0.142	0.025	0.242	2	2	0
2001	0.292	0.058	0.142	0.100	0.433	5	5	0
2002	0.420	0.140	0.004	0.150	0.250	1	1	0
2003	0.425	0.225	0.025	0.125	0.200	1	1	0
2004	0.425	0.225	0.025	0.125	0.200	1	1	0
2005	0.425	0.200	0.075	0.050	0.025	1	1	0
2006	0.300	0.192	0.092	0.075	0.342	5	5	0
2007	0.325	0.125	0.025	0.225	0.300	1 或 5	5	0
2008	0.325	0.200	0.000	0.250	0.225	1	4	3
2009	0.300	0.475	0.000	0.125	0.100	2	2	0
2010	0.442	0.333	0.117	0.050	0.058	1	2	1
2011	0.487	0.190	0.323	0.000	0.100	1	3	2

### 2.3 预测应用

用与回报检验相同的方法预测锦屏 2012 年白背飞虱田间发生程度(表 2), 预测结果显示, 2012 年锦屏田间白背飞虱的预测级别与实际发生级别相符, 预测发生级别 1 级或 2 级, 实际发生级别为 2 级, 预测结果准确。

表 2 2012 年锦屏白背飞虱田间发生程度的预测结果

站点	各 级 概 率					回报级别	实际级别	误差
	1	2	3	4	5			
锦屏	0.370	0.357	0.157	0.000	0.150	1 或 2	2	0

## 3 结论与讨论

稻飞虱是贵州省重要的水稻害虫, 属于典型的 r-对策昆虫, 迁飞能力强、内禀增长率高, 其发生受气候条件(降水、温度、湿度等)影响较大, 在一定的迁入基数和适宜的条件下常能爆发成灾<sup>[14-15]</sup>, 20 世纪六、七十年代, 在全国稻飞虱科研协作组和各地推广部门的通力配合下, 基本建成了全国各稻区的稻飞虱综合防治体系<sup>[16-17]</sup>, 其中也形成了一系列的中短期预测方法, 但对其发生程度的长期预测研究方法却较为缺乏<sup>[18]</sup>, 所以本着预防为主植保方针, 应加强对稻飞虱发生程度的长期预测研究, 以增强面对稻飞虱爆发式发生时的预警和应对能力。

基于马尔科夫链理论的预测方法就具有长期预测的特点, 且方法较为成熟, 在病虫害预测预报中已有较多的应用<sup>[19-20]</sup>. 本文以贵州省白背飞虱发生较为严重的锦屏为例, 采用 1990—2011 年锦屏白背飞虱田间发生程度的时间序列资料, 基于马尔科夫链理论, 构建了 1~5 阶的概率转移矩阵, 以预报年前 5 年的连

续发生级别预测第 6 年的发生状态, 回报锦屏的历史数据, 得到了 82% 的历史符合率, 对锦屏 2012 年的预测结果也符合当年的实际发生情况, 总体而言取得了令人满意的结果。

与投影寻踪理论<sup>[21]</sup>、人工神经网络<sup>[22]</sup>、像空间重构<sup>[23]</sup>、大气环流特征<sup>[14]</sup>、海温遥相关<sup>[24]</sup>等其他长期预测方法相比, 马尔科夫链无需考虑繁复的外界因子, 只考虑样本本身历史状态的演变规律, 计算状态转移概率即可预测未来可能发生的状态<sup>[25]</sup>, 相关资料的获取难度和技术操作难度都大大降低, 且有较好的预测准确性. 这种方法减少了测报工作的难度, 利于基层测报工作者运用, 在测报经费花销巨大和测报人员紧张的今天不失为一个优良的预测预报方法; 但是马尔科夫链同样具有一定的缺点, 首先, 需要积累多年的田间虫情资料, 且不能间断; 其次, 此方法没有考虑多种因素对害虫种群发生的影响, 回避了种群内个体行为和外界条件对种群消长的作用. 在以后的测报工作中, 要把中短期预测和长期预测有机结合起来<sup>[26]</sup>, 以长见短, 以短防长, 更有效地开展病虫害的综合防治工作。

## 参考文献:

- [1] 贵州省稻飞虱联合测报组. 贵州稻飞虱研究(1981—1983 年) [J]. 贵州农业科学, 1984(5): 24—28.
- [2] KISIMOTO R, ROSENBERG L J. Long-Distance Migration in Delphacid Planthopper [M] // Denno R F, Perfect T J. Planthopper: Their Ecology and Management. New York: Chapman and Hall, 1994.
- [3] OTUKA A, MATSUMURA M, SANADA-MORIMURA S, et al. The 2008 Overseas Mass Migration of the Small Brown Planthopper, *Laodelphax striatellus* (Fallén), and Subsequent Outbreak of Rice Stripe Disease in Western Japan [J]. Applied Entomology and Zoology, 2010, 45(2): 259—266.
- [4] SYOBU S, OTUKA A, MATSUMURA M. Trap Catches of the Small Brown Planthopper, *Laodelphax striatellus* (Fallen)(Hemiptera: Delphacidae), in Northern Kyushu District, Japan in Relation to Weather Conditions [J]. Applied Entomology and Zoology, 2011, 46(1): 41—50.
- [5] 翟保平, 张孝羲. 迁飞过程中昆虫的行为: 对风温场的适应与选择 [J]. 生态学报, 1993, 13(4): 356—363.
- [6] 程遐年, 吴进才, 马 飞. 褐飞虱研究与防治 [M]. 北京: 中国农业出版社, 2003.
- [7] 张孝羲, 张跃进. 农作物有害生物预测学 [M]. 北京: 中国农业出版社, 2006.
- [8] 金 星, 王德其, 金道超, 等. 贵州省稻飞虱发生危害规律与防治对策 [J]. 山地农业生物学报, 1998, 17(4): 208—214.
- [9] 薛文鹏, 金道超, 杨 洪. 贵州白背飞虱灯下种群发生动态区域性比较研究 [J]. 环境昆虫学报, 2014(4): 487—493.
- [10] 陈观浩. 应用马尔可夫链法预测晚稻飞虱发生程度 [J]. 应用昆虫学报, 2003, 40(2): 176—178.
- [11] 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): 16—22.
- [12] 闫香慧, 赵志模, 刘 怀, 等. 应用马尔可夫链理论对褐飞虱和白背飞虱发生程度的预测 [J]. 生态学报, 2009, 29(11): 5799—5806.
- [13] 夏乐天, 彭支行, 沈永梅. 加权马尔科夫链在农作物年景预测中的应用 [J]. 数学的实践和认识, 2005, 35(12): 30—35.
- [14] 高 苹, 武金岗, 陈 宁, 等. 大气环流特征量的水稻白背飞虱发生程度预报模型的研究 [J]. 生态学杂志, 2005, 24(2): 146—152.
- [15] 祝增荣, 程家安.  $r$ -对策型害虫暴发成灾 [M] // “10000 个科学难题”农业科学编委会. 10000 个科学难题: 农业科学卷. 北京: 科学出版社, 2011.
- [16] 杜正文, 朱绍先, 蒋文烈, 等. 我国五大稻区水稻病虫害综合防治的系统对策 [J]. 中国农业科学, 1986, 19(5): 65—70.
- [17] 李汝铎, 丁锦华, 胡国文, 等. 褐飞虱及其种群管理 [M]. 上海: 复旦大学出版社, 1996.
- [18] 巫国瑞, 俞晓平, 陶林勇. 褐飞虱和白背飞虱灾害的长期预测 [J]. 中国农业科学, 1997, 30(4): 25—29.
- [19] 高汉杰, 任广伟, 王 永, 等. 马尔可夫链在烟蚜发生程度预测中的应用 [J]. 中国烟草科学, 2004, 25(2): 46—48.
- [20] 康爱国, 姜玉英, 王贺军, 等. 应用马尔可夫链模型对草地螟发生程度的预测 [J]. 应用昆虫学报, 2012, 49(5): 1243—1248.
- [21] 娄伟平, 陈先清, 吴利红, 等. 基于投影寻踪理论的稻飞虱发生程度预测模型 [J]. 生态学杂志, 2008, 27(8): 1438—1443.

- [22] 马 飞, 许晓风, 张夕林, 等. 神经网络预警系统及其在害虫预测中的应用 [J]. 应用昆虫学报, 2002, 39(2): 115—119.
- [23] 马 飞, 许晓风, 张夕林, 等. 相空间重构与神经网络融合预测模型及其在害虫测报中的应用 [J]. 生态学报, 2002b, 22(8): 1297—1301.
- [24] 高 苹, 武金岗, 杨荣明, 等. 江苏省稻纵卷叶螟迁入期虫情指标与西太平洋海温的遥相关及长期预报模型 [J]. 应用生态学报, 2008, 19(09): 2056—2066.
- [25] 邵崇斌, 郎杏茹. 应用马尔科夫链预测杨树天牛发生量 [J]. 陕西林业科技, 1996(4): 28—32.
- [26] 卿雨文, 李 春, 张志东, 等. 泸州地区白背飞虱种群动态及预测预报 [J]. 西南农业大学学报, 2000, 22(1): 49—52.

## Forecasting the Occurrence Degree of *Sogatella furcifera* (Horváth) in Jinping of Guizhou Province by Applying Markov Chain Theory

XUE Wen-peng<sup>1,2</sup>, YANG Rui<sup>3</sup>, YANG Hong<sup>1</sup>, YANG Lin<sup>4</sup>

1. Institute of Entomology, Guizhou University, Guiyang 550025, China;

2. Guiyang Pant Protection and Quarantine Station, Guiyang 550081, China;

3. Guiyang Forestry Pest Control Station, Guiyang 550002, China

4. Jinping Plant Protection Station, Jinping, Guizhou 556000, China

**Abstract:** White-backed planthopper, [*Sogatella furcifera* (Horváth)], is an important rice pest insect in China. In a study reported in this paper, the time series data of *S. furcifera* field occurrence level from 1900 to 2011 in Jinping, a quaternary middle rice-growing area in Guizhou, were selected to establish a probability transfer matrix of 1 to 5 steps according to Markov chain theory, and to forecast, based on the continuous occurrence level of the previous five years, the occurrence of the pest in the sixth year. When used to reversely forecast the occurrence of *S. furcifera* in Jinping in the preceding 17 years, this matrix gave a satisfactory historical coincident rate (82%), and the results based on prediction in 2012 of Jinping were also consistent with the real situation of the year.

**Key words:** white-backed planthopper (*Sogatella furcifera*); forecast; Markov chain; occurrence degree

责任编辑 周仁惠

崔玉洁

