

基于马尔可夫链的油茶病虫害预测研究*

Camellia Pests Forecast by Markov Chain

庞正轰

PANG Zheng-hong

(广西生态工程职业技术学院, 广西柳州 545004)

(Guangxi Eco-engineering Vocational College, Liuzhou, Guangxi, 545004, China)

摘要:为了及时和有效地指导油茶病虫害防治工作,分别根据2000~2009年,2001~2010年,2002~2011年广西油茶病虫害发生危害的数据,预测2010年,2011年和2012年广西油茶病虫害发生危害的等级和发生面积。结果显示,2010年广西油茶重大病虫害发生危害面积为500 hm²,2011年病虫害发生量等级为2级(333.4~666.6 hm²),2012年病虫害发生量等级也为2级,而2010年广西油茶病虫害实际发生危害面积为496.7 hm²,2011年为435.7 hm²,2012年为656.7 hm²。可见3次预报的准确率都在90%以上,表明利用马尔可夫链方法进行油茶病虫害预测预报是可行的,但需要积累较长时间的历史数据。

关键词:马尔可夫链 油茶 病虫害 预测预报

中图分类号:S763.1 文献标识码:A 文章编号:1002-7378(2013)04-0278-03

Abstract: The pests and diseases of Camellia are very serious in Guangxi. In order to control the pests and diseases, the damage degree and area of the Camellia pests and diseases in Guangxi in 2010, 2011 and 2012, were predicted by Markov Chain, according to the history data of pests and diseases of Camellia in Guangxi from 2000 to 2011. The result showed that the accuracy of forecast was over 90%. Therefore, pests forecast of Camellia by Markov Chain is feasible, but it needs a long history data and the environment is in relative unchanged condition.

Key words: Markov Chain, camellia, pests and diseases, forecast

油茶(*Camellia oleifera* Abel.)是一种重要的经济林树种,广西的自然条件尤其适合油茶生长。至2010年,广西油茶面积约40万hm²,占全国油茶总面积的12%。病虫害是制约油茶产业健康发展的重大因子之一。据调查,2000~2009年,广西油茶重大病虫害年均严重发生危害面积580hm²,病虫害种类达60多种,主要有油茶毒蛾(*Euproctis pseudoconspersa* Strand)、油茶枯叶蛾(*Lebeda nobilis* Walker)、油茶尺蠖(*Biston marginata* Shiraki)和油茶炭疽病(*Colletotrichum camelliae* Masee)等,发生病虫害的地域很广,主要发生在融

水、融安、三江、巴马、凤山、田阳、右江、田林、隆林、岑溪等油茶生产重点县(市、区),已造成直接经济损失约1.5亿元。油茶产业在广西方兴未艾,根据广西省政府发展规划,到2020年,广西油茶林面积将达80万hm²,年产值将超过200亿元。如何科学有效地防治油茶病虫害已经成为发展油茶产业的一项重要任务。广西油茶病虫害严重的主要原因是,没有掌握油茶病虫害的发生规律和经济有效的防治方法,未能建立油茶病虫害预测预报机制^[1~5]。本文分别根据2000~2009年,2001~2010年,2002~2011年广西油茶病虫害发生的危害数据,利用马尔可夫链方法预测2010年,2011年和2012年广西油茶病虫害发生的危害等级和发生面积,旨在提高广西油茶病虫害预测预报水平,及时有效地指导油茶病虫害防治工作,为建立长期有效的油茶病虫害预测预报机制提供参考。

收稿日期:2013-06-20

修回日期:2013-09-10

作者简介:庞正轰(1957-),博士,教授,主要从事林业有害生物监测防治的科研和教学工作。

*广西重大科技项目(桂科攻10100012-1B)资助。

1 预测原理分析

如果对一个过程给定它的“现在”,那么,其“将来”就独立于它的“过去”,这样的随机过程就称为马尔可夫过程。马尔可夫过程具有无后效性,即过程对于最近之前是可“转化”的。马尔可夫链是时间离散、状态离散的马尔可夫过程。在马尔可夫链中,系统状态的转移可以用一个转移概率矩阵来表示,即 $P = [P_{ij}]$ 。转移矩阵中的元素,称为转移概率。转移矩阵有两个重要性质,即所有的元素是非负的,而且各项元素之和等于1。即 $P_{ij} \geq 0, \sum P_i = 1, 2, 3, \dots, n$ 。因为对于从任何一个固定的 i 状态出发,经过一次转移到 j 状态,就会生成状态 $E_1, E_2, E_3, \dots, E_n$ 。所以 P_{ij} 就形成一个概率矩阵。这是一阶转移概率矩阵。相应地,有2阶,3阶, n 阶(高阶)转移概率矩阵。高阶转移概率可用矩阵乘法求得: $P^{(n)} = P^{(1)} \cdot P^{(n-1)}$ 。

病虫害在某个时间内的发生量处于一定的状态中。在病虫害种群变化的时间序列中,往往存在某种程度的相互关联性或者连续性,即前一年的种群数量是其下一年种群数量消长的基础,未来种群数量是其以前种群数量的延续,保持着整个种群的关联性和连续性。应用马尔可夫链方法进行病虫害种群数量动态预报,就是根据病虫害现在所处的状态向前相继转移的概率进行预报。如果某种害虫种群数量的初始状态的分布是已知的,那么,可计算出其各阶的转移概率,再根据转移概率的大小来预测将来害虫发生状态。

如果把病虫害所有可能发生的状态(发生量等级)分别记为 E_1, E_2, E_3, \dots ,把可能的发生转移时间(1a, 2a, 3a, ...)记为 t_1, t_2, t_3, \dots ,那么在马尔可夫链中,系统状态的转移可以用一个转移概率矩阵(又称随机矩阵) $P = [P_{ij}]$ 表示,转移矩阵中的元素,称为转移概率。可表示为

$$P = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix},$$

$$P_{ij} \geq 0, \sum P_{ij} = 1, i = 1, 2, 3, j = 1, 2, 3.$$

转移概率 P_{ij} 表示已知在时刻 t_{n-1} 系统处于状态 i 的条件下,系统在时刻 $t_n (t_n > t_{n-1})$ 处于状态 j 的概率。 P_{ij} 中的下标次序对应转移方向和顺序,即从状态 i 向状态 j 转移。对于从任何一个已知的 i 状态出发,经过一次转移到 j 状态,出现状态 E_1, E_2, E_3, \dots 中之一,所以 P_{ij} 就能形成一个概率矩

阵。对于有 n 个阶的转移概率矩阵,要做 n 次状态转移。如果有3个阶的转移矩阵就要做3次状态转移,根据这3次状态转移概率出现的最大可能值合计数进行比较。最大可能值合计数对应的等级即为预报年度的发生量等级。

2 预测方法

以2000~2009年广西油茶病虫害严重发生危害数据为例,简述基于马尔可夫链预测2010年广西油茶病虫害的方法。

步骤1 划分发生量等级。2000~2009年广西年均油茶病虫害严重发生危害面积 580hm^2 。根据马尔可夫链性质的要求,发生量等级可分为三级。1级: $< 333.3 \text{hm}^2$; 2级: $333.4 \sim 666.6 \text{hm}^2$; 3级: $> 666.6 \text{hm}^2$ 。2000~2009年广西油茶重大病虫害发生面积及发生量等级划分结果见表1。

表1 2000~2009年广西油茶重大病虫害发生危害面积及发生量等级划分表

时间(年度)	发生面积(hm^2)	发生等级
2000	100.0	1
2001	158.0	1
2002	173.3	1
2003	190.7	1
2004	210.7	1
2005	596.7	2
2006	590.0	2
2007	2850.0	3
2008	484.9	2
2009	442.7	2

步骤2 计算发生转移概率。首先计算1阶转移概率,然后计算2阶,3阶转移概率。

根据马尔可夫链性质,2000~2009年广西油茶病虫害发生等级频率及发生转移概率 $P^{(1)}$ 为:

$$P^{(1)} = \begin{bmatrix} 4 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0 & 0.67 & 0.33 \\ 0 & 1 & 0 \end{bmatrix},$$

$$P^{(2)} = P^{(1)} \cdot P^{(1)} = \begin{bmatrix} 0.64 & 0.28 & 0.08 \\ 0 & 0.78 & 0.22 \\ 0 & 0.67 & 0.33 \end{bmatrix},$$

$$P^{(3)} = P^{(1)} \cdot P^{(2)} = \begin{bmatrix} 0.51 & 0.38 & 0.11 \\ 0 & 0.75 & 0.25 \\ 0 & 0.78 & 0.22 \end{bmatrix},$$

步骤3 根据状态转移概率进行预测。2009年发生等级为2级,转移1次到2010年的概率从 $P^{(1)}$ 查出:1级0,2级0.67,3级0.33。2008年发生

等级为2级,转移2次到2010年的概率从 $P^{(2)}$ 查出:1级0,2级0.78,3级0.22;2007年发生等级为3级,转移3次到2011年的概率从 $P^{(3)}$ 查出:1级0,2级0.78,3级0.22。

由此可见,2010年油茶病虫害发生面积等级的概率为:1级为0,2级为2.23,3级为0.77。所以,可以预报2010年广西油茶病虫害发生等级为2级,发生危害面积为333.4~666.6 hm^2 ,中值为500亩;可靠性在95%的情况下,其置信区间466.7~533.3 hm^2 。

3 预测结果

利用马尔可夫链预测2010年广西油茶重大病虫害发生危害面积为500 hm^2 ,而实际发生危害面积为496.7 hm^2 (数据来源于广西森林病虫害防治站,下同),可见预测准确率为99.3%。表明预报准确。

同理,根据2001~2010年广西油茶重大病虫害发生危害的数据预测2011年发生量等级为2级,发生危害面积为333.4~666.6 hm^2 ,实际发生危害面积为435.7 hm^2 ;根据2002~2011年广西油茶重大病虫害发生危害的数据预测2012年发生量等级为2级,发生危害面积为333.4~666.6 hm^2 ,实际发生危害面积为656.7 hm^2 。上述结果都表明预报是准确的。

再根据2003~2012年广西油茶重大病虫害发生危害的数据预测2013年发生量等级为3级(1级400 hm^2 以下,2级400~600 hm^2 ,3级600 hm^2 以上),发生危害面积为600 hm^2 以上,实际发生危害面积为多少,有待实践检验。

4 结论

薛贤靖等^[1]于20世纪80年代采用马尔可夫链方法对湖南省湘乡县马尾松毛虫(*Dendrolimus punitatus* Walker)发生量和发生期进行了预测,预报结果与实际发生情况基本吻合。庞正轰等^[6]于1990年同样应用马尔可夫链方法对广西马尾松毛虫灾害进行了长期预测,预报结果与实际发生情况一致。本研究,利用马尔可夫链方法进行油茶病虫害长期预报,同样取得比较满意的结果。这表明马

尔可夫链方法应用于森林病虫害预测预报具有一定的可行性。而且利用马尔可夫链方法进行病虫害预测预报具有如下特点:它不需要从许多繁杂因子(气象、林分、土壤、天敌、人为活动等)中去寻找相关规律,只需利用油茶病虫害本身的历史状态的演变特点,就能预测出它未来的发展态势。油茶病虫害每年发生数量都处于一定的变化状态之中,前一年的发生状态是下一年的基础,下一年的发生数量是上一年发生状态的延续,两者保持着密切的关系,维持着整个种群演变的连续性。马尔可夫链方法正好可以揭示油茶病虫害种群数量变动的内在规律,因此,利用马尔可夫链方法对油茶病虫害进行预测预报是可行的。但是利用马尔可夫链进行预测预报,需要连续多年的历史资料 and 进行等级划分,预报结果只是一个数值范围,因此,使用马尔可夫链进行病虫害预测预报需要考虑如下几个方面的影响因素:一是要积累较长时间的历史资料,一般要求10a以上;二是等级划分要比较合理,等级划分越多,预报结果越精细,但准确率可能下降,根据马尔可夫链性质的要求,等级划分一般为3~4级为宜;三是客观条件要相对稳定,如果通过人为措施对病虫害进行了大面积防治,或通过人工造林短期内大量地增加寄主(油茶林)面积,在这种情况下采用马尔可夫链方法预测下一年度的发生量就可能与实际不符。

参考文献:

- [1] 薛贤靖. 森林病虫害预测预报[M]. 北京:中国林业出版社,1992.
- [2] 黄敦元,王森. 油茶病虫害防治[M]. 北京:中国林业出版社,2010.
- [3] 束庆龙,张良富. 中国油茶栽培与病虫害防治[M]. 北京:中国林业出版社,2009.
- [4] 周国英,宋光桃,李河. 油茶病虫害防治现状及应对措施[J]. 中南林业科技大学学报,2007,27(6):179-182.
- [5] 杨坚. 油茶主要病虫害的发生及防治[J]. 广东农业科学,2007(7):66-68.
- [6] 庞正轰,杨秀好. 马尔可夫链方法在广西松毛虫长期预报中的应用[J]. 广西植保,1992(4):15-17.

(责任编辑:尹 闯)