

应用马尔科夫链模型对草地螟发生程度的预测*

康爱国^{1**} 姜玉英² 王贺军³ 张玉慧¹ 李强¹
庞红岩¹ 沈成⁴ 李金有¹

(1. 河北省康保县植保站 康保 076650; 2. 全国农业技术推广中心 北京 100125;

3. 河北省植保植检站 石家庄 050011; 4. 张家口市植保站 张家口 075000)

摘要 草地螟 *Loxostege stictialis* L. 是我国北方农牧业生产上一种重要迁飞性、暴发性害虫,一旦暴发会给当地农牧生产造成严重危害。根据康保县 1977—2008 年 1 代草地螟幼虫发生程度的时间序列资料,应用马尔科夫链的转移概率预测法,构建了 1~3 阶转移概率矩阵,组建模型对该县 2009—2011 年 1 代草地螟发生程度进行了预测,结果与大田实际发生情况完全一致,准确率 100%。对 1980—2011 年的历史资料进行回检,历史符合率 89.9%,该方法可对草地螟进行长期预报,为草地螟长期预报提供了一种准确有效的方法,对草地螟发生程度的长期预报具有重要指导意义。

关键词 草地螟, 发生程度, 马尔科夫链, 预测

Predicting beet webworm occurrence with the Markov chain model

KANG Ai-Guo^{1**} JIANG Yu-Ying² WANG He-Jun³ ZHANG Yu-Hui¹ LI Qiang¹
PANG Hong-Yan¹ SHEN Cheng⁴ LI Jin-You¹

(1. Plant Protection Station of Kangbao County in Hebei Province, Kangbao 076650, China;

2. National Agrotech Extension and Service Center, Beijing 100125, China; 3. Plant Protection of Hebei

Province, Shijiazhuang 050011, China; 4. Plant Protection of Zhangjiakou City Hebei Province, Zhangjiakou 075000, China)

Abstract The beet webworm *Loxostege stictialis* (L.) is one of the main migratory pests in North China. Outbreak populations of this pest can seriously damage host crops. On the basis of data on the first generation of beet webworm larva collected from 1977 to 2008, we constructed a matrix of migration probability by applying the Markov chain model and then used this to predict the population of the first generation in 2009 – 2011. By using the transition probability method of Markov chain theory, a transition probability matrix of 1 to 3 steps was constructed based on time series data on the abundance of the pests' first generation from 1979 to 1999. This was used to predict the beet webworm abundance in Kangbao city from 2009 to 2011. All predictions were 100% accurate and the results show that the predictive accuracy of the data collected from 1980 to 2011 was 89.9%. The predicted population completely conformed to that actually observed in the field.

Key words *Loxostege stictialis*, degree of occurrence, Markov chain theory, prediction

草地螟 *Loxostege stictialis* L. 是我国北方农牧业生产上一种重要迁飞性、暴发性害虫,建国以来已 3 次周期性暴发。1995 年草地螟种群回升,间隔 12 年后 1997 年突然特大发生,进入了建国以来第 3 个暴发周期。暴发初期和暴发期,由于长期预报准确率低,药、械储备不足,贻误防治最佳时期,往往给当

地农牧业生产造成严重损失。1997 年河北省 1 代草地螟幼虫发生面积 100.4 万 hm^2 ,其中河北省张家口市草地螟幼虫发生 60 万 hm^2 ,经济损失 2.4 亿元。重发区康保县有 2.4 万 hm^2 农田因缺少药、械、水源等原因无法控制严重减产,绝收面积达 0.8 万 hm^2 (罗礼智等,1998)。因此,搞好草地螟的长期和

* 资助项目:国家科技攻关项目(2005BAD529A03,2005BAD529A04)、公益性行业(农业)科研专项(201103002)。

**E-mail: Lqzyh6@sina.com

收稿日期:2011-06-27,接受日期:2011-09-21

超长期准确预报,对科学决策,指导草地螟防治,保障农牧业生产安全,防灾减灾显得更为重要。本文应用马尔科夫链法对康保县 1 代草地螟幼虫发生程度进行了超长期预报。

马尔科夫链预测法又叫概率转移法,是一种超长期预测决策的方法。是根据转移概率由 i 时刻的状态预测 $i+n$ 时刻的状态,这种方法具有长期或超长期预测特点。农作物病虫害的发生,在时间序列上的变化,可视为一种时间离散、状态离散的过程,且具有“无后效应”和“遍历性”,这与马尔科夫链过程很相似,因此可参照马尔科夫链的求解方法,通过对历史资料进行统计分析,探究其“过去”、“现在”与“将来”的内在联系,并根据当前状态预测其未来趋势变化的概率保障,做出长期或超长期预报(胡伯海和姜瑞中,1997;张孝羲和张跃进,2006),目前这种方法已广泛应用于农林害虫、气象、金融等相关领域(谢成君,1997;陈观浩,2003;贾春生,2006;台文志,2008;田红霞和郭晨,2008;闫香慧等,2009)。鉴于此,作者应用马尔科夫链方法,依据康保县 1977—2008 年 32 年草地螟 1 代发生程度的历史资料,建立了草地螟的马尔科夫链预测模型,并对 2009、2010 和 2011 年 1 代草地螟发生程度进行了预测和实际发生的检测,还对历史上每年发生实况进行了回检。

现将研究结果报道如下。

1 材料与方法

1.1 数据来源与历史资料整理

使用的草地螟历史资料来自河北省康保县植保站 1977—2008 年 32 年草地螟系统监测历史资料,包括幼虫密度、发生面积和防治所占面积。调查方法:幼虫发生量、发生面积和化防所占面积采取田间系统监测和大面积普查相结合。系统调查:取不同环境,每种植被调查 5 块样地,每块地调查 10 个样点,每个样点取样面积 $33\text{ cm} \times 33\text{ cm}$ 或 $50\text{ cm} \times 50\text{ cm}$, 2~3 d 调查 1 次;大面积普查:调查全部乡镇,采取每个乡镇查 5 个村,每村、每种植被和样点调查同系统调查,于幼虫发生危害盛期和末期调查 2 次。根据每年草地螟的发生实况,结合全国草地螟测报规范发生级别的划分标准将其划分 5 级。由于使用的 1977—2008 年 32 年资料,是从第 2 个暴发周期开始至第 3 个暴发周期结束不完全的 2 个暴发周期资料,为了提高预报准确率,根据草地螟发生特点,在对指导防治无明显影响前提下,将历史资料发生级别由 5 级划分 3 级,草地螟 3 级发生程度划分指标见表 1,按新划分指标得到一新转换数据数列(表 2)。数据运算和建模在 EXCEL 上完成。

表 1 草地螟发生程度划分指标

Table 1 The degree of occurrence of *Loxostege stictialis*

发生程度 The degree of occurrence	虫量 (头/m ²) Number	化防面积所占比率(%) The proportion of chemical control (%)	备注 Remarks
1	≤10	≤5	不需化防,作物无明显受害损失 No chemical control, no significant loss of crop
2	11-80	6-30	需重点防治,不防治会造成减产 Need to fractional control, if not, cause dropped production
3	81-200 头及以上	31-41 及以上	需大面积防治,不防治会造成严重危害或绝收 Need to control a large area, if not, cause serious harm or total destruction

1.2 构建马尔科夫链 1 至 k 步的转移概率矩阵

在草地螟年度间发生状态序列中,把它的发生程度分为若干有限等级,根据草地螟历年级别(状态),经过 k 年(本研究取 $k=1,2,3$ 阶)从状态

E_i 转移到状态 E_j (将状态区分为 1,2,3 级)的转移概率用 $P_{ij}(k)$ 表示。

$$P_{ij}(k)_{N_{ij}(k) / B_i}$$

表 2 康保县 1977—2008 年 1 代草地螟发生程度划分级别

Table 2 Occurrence grade of the first generation of *Loxostege stictialis* from 1977 to 2008 in Kangbao city

年份 Years	年序 Numbers	级别 Grades	年份 Years	年序 Numbers	级别 Grades
1977	1	2(1-2)	1995	19	2(1-2)
1978	2	1	1996	20	2(1-2)
1979	3	3	1997	21	3
1980	4	2	1998	22	3(2-3)
1981	5	3	1999	23	3(3-2)
1982	6	3	2000	24	3
1983	7	2(2-3)	2001	25	3
1984	8	3(3-2)	2002	26	2(2-3)
1985	9	2	2003	27	2(2-3)
1986	10	1	2004	28	3(3-2)
1987	11	1	2005	29	2
1988	12	1	2006	30	2(2-3)
1989	13	1	2007	31	1
1990	14	1	2008	32	1
1991	15	1			
1992	16	1			
1993	17	1			
1994	18	1			

式中, $P_{ij}(k)$, 为 k 阶草地螟发生程度 E_i 转移到 E_j 的次数, B_i 为草地螟发生程度为状态 E_i 的总次数。以 $P_{ij}(k)$ 为元素构成的矩阵 $P(k)$ 成为 k 阶转移概率矩阵。该矩阵各行之和等于 1。

$$P(k) = \begin{bmatrix} p_{11}(k) & p_{12}(k) & p_{13}(k) \\ p_{21}(k) & p_{22}(k) & p_{23}(k) \\ p_{31}(k) & p_{32}(k) & p_{33}(k) \end{bmatrix}$$

1.3 数据分析方法

草地螟逐年发生级别的转移概率矩阵和 2009 年预测结果采用唐启义和冯明光《实用统计分析及其计算机处理平台》数理统计软件 DPS v3.01 专业版分析所得。回检采用回报年的前 1, 2, 3 年的状态变量 i , 分别取转移概率矩阵 $P(1)$, $P(2)$, $P(3)$ 中第 i 行元素各阶值, 该矩阵各列之和即为可能发生状态 j 的概率, 以概率值最大或与之最接近的状态为回检值或称预测值。

2 结果与分析

2.1 草地螟发生级别的转移概率矩阵

1977—2008 年 1 代草地螟逐年发生级别见表 1。本研究取 1, 2, 3 阶, 从表 1 中的 32 个年序的各

个分级, 分别计算草地螟各阶的转移概率矩阵 $P(k)$ 如下:

$$P(1) = \begin{bmatrix} 0.818 & 0.091 & 0.091 \\ 0.300 & 0.300 & 0.400 \\ 0.000 & 0.500 & 0.500 \end{bmatrix}$$

$$P(2) = \begin{bmatrix} 0.700 & 0.300 & 0.000 \\ 0.300 & 0.200 & 0.500 \\ 0.100 & 0.400 & 0.500 \end{bmatrix}$$

$$P(3) = \begin{bmatrix} 0.600 & 0.200 & 0.200 \\ 0.333 & 0.444 & 0.222 \\ 0.200 & 0.300 & 0.500 \end{bmatrix}$$

2.2 预报 2009 年草地螟发生级别(程度)

为了便于数学表达, 在此以 E_i 代表预测项(即 2009 年), j 为 E_i 将处状态(级别), $j=1, 2, 3$ 。 $P(E_i=j)$ 代表预测项 E_i 所处状态(j)的概率。列出与第 32 年序(2009 年)相距 1, 2, 3 年(2006, 2007, 2008 年)草地螟发生程度级值, 计算转移概率平均值, 依据转移概率平均值的最大值为预报值见表 3, 表 3 数据表明, 在转移概率平均值栏中, 以状态“1”的概率最大为 0.617, 故预报 2009 年 1 代草地螟 1 级(轻发生), 实际发生发生级别为

1 级(轻发生), 预报值与实测值相吻合。同样运用模型对 2010 年和 2011 年 1 代草地螟幼虫发生程度进行预测, 预测值均为 1 级(轻发生), 预测值和实测值相吻合, 连续 3 年预报准确率 100%。

草地螟 32 年的发生级别, 预测值和实测值只有 1985, 1986, 1997 相差 1 级, 2007 相差 2 级, 其余 28 年预测结果与实际发生情况相吻合, 历史回检符合率 89.9% (表 3)。误差年份主要出现在每个暴发周期开始和结束的 1~2 年。

2.3 历史符合率回检

用上述计算方法逐一回报 1980—2011 年 1 代

表 3 康保县 1977—2011 年 1 代草地螟发生级别回报检验表
Table 3 Test of the predictive accuracy of the model on *Loxostege stictialis* occurrence from 1977 to 2008 in Kangbao city

年份 Years	实测级别 Actual grades	各级转移状态及概率 Transfer condition and probability of each grade			预测级别 Forecasting grades	误差 SE	检验 Test
		1	2	3			
1977	2 (1-2)						
1978	1						
1979	3						
1980	2	0.344	0.415	0.247	2	0	√
1981	3	0.333	0.300	0.367	3	0	√
1982	3	0.167	0.333	0.500	3	0	√
1983	2 (2-3)	0.145	0.448	0.407	2	0	√
1984	3 (3-2)	0.100	0.333	0.467	3	0	√
1985	2	0.167	0.333	0.500	3	-1	×
1986	1	0.244	0.381	0.374	2	-1	×
1987	1	0.439	0.197	0.364	1	0	√
1988	1	0.617	0.278	0.104	1	0	√
1989	1	0.706	0.197	0.097	1	0	√
1990	1	0.706	0.197	0.097	1	0	√
1991	1	0.706	0.197	0.097	1	0	√
1992	1	0.706	0.197	0.097	1	0	√
1993	1	0.706	0.197	0.097	1	0	√
1994	1	0.706	0.197	0.097	1	0	√
1995	2 (1-2)	0.706	0.197	0.097	1	-0.5	√
1996	2 (1-2)	0.533	0.267	0.200	1	-0.5	√
1997	3	0.400	0.233	0.367	1 或 3	-1	×
1998	3 (2-3)	0.211	0.381	0.407	2-3	0	√
1999	3 (3-2)	0.144	0.448	0.407	2-3	0	√
2000	3	0.100	0.400	0.500	3	0	√
2001	3	0.100	0.400	0.500	3	0	√
2002	2 (2-3)	0.100	0.400	0.500	2-3	0	√
2003	2 (2-3)	0.200	0.333	0.467	3-2	0	√
2004	3 (3-2)	0.267	0.267	0.467	3	0	√
2005	2	0.211	0.381	0.407	2-3	0	√
2006	2 (2-3)	0.244	0.381	0.374	2-3	0	√
2007	1	0.267	0.267	0.467	3	-2	×
2008	1	0.484	0.125	0.271	1	0	√
2009	1	0.617	0.278	0.104	1	0	√
2010	1	0.781	0.199	0.100	1	0	√
2011	1	0.807	0.183	0.091	1	0	√

3 结论与讨论

目前国内对草地螟的发生程度长期预测主要采用上一代残虫量和气象部门长期气候预测情况综合评价后来预报下一代或来年的发生量(张跃进,2006),这种方法准确率低,特别是省、市和县级基层测报站预报准确率更低,原因是草地螟是一种迁飞性害虫,受气候环境、植被等因素影响,上一代残虫量与下一代幼虫发生密度和面积关系不显著,相关系数仅 0.1889 和 0.3469(康爱国等,2004)。1996 年康保县越冬虫量平均 0.05 头/ m^2 ,1997 年越冬代成虫 1 盏黑光灯累计诱蛾 298 333 头,日最高有蛾量 201 748 头,为当地历史极值,1 代幼虫密度最高达 3 000 多头/ m^2 ,1 代幼虫大发生;2008 年 1 代草地螟在国内发生偏轻,发生面积 93.33 万 hm^2 ,残虫面积小,密度低,预报 2 代呈偏轻发生态势。进入 7 月下旬至 8 月中旬大批境外虫源迁入我国,造成 2 代草地螟在我国境内大发生,全国 2 代草地螟幼虫发生面积 1 160 万 hm^2 。2008 年全国草地螟越冬虫量是历史上最高年份 2003 年的 4.0 倍,各地越冬虫量密度平均虫量 17.2 头/ m^2 ,为国内有记载最高值(姜玉英等,2009)。根据越冬虫量经验预报 2009 年 1 代草地螟幼虫在我国北方大面积中偏重或大发生,实测只在东北、内蒙西部局部偏重发生。2008 年康保县越冬面积 5.2 万 hm^2 ,越冬基数平均 21.1 头/ m^2 ,最高 274.5 头/ m^2 ,是 1979 年有记载以来最高的 1 年。根据当地和全国草地螟越冬虫量,结合当地 2009 年全年气候展望预报,经验预测 2009 年 1 代草地螟幼虫在康保县中偏重至大发生,实测轻发生。而采用马尔科夫链预测法预报 2009 年 1 代草地螟幼虫轻发生,实测轻发生,预测值和实测值相吻合。

目前国内和前苏联利用太阳黑子的活动规律对草地螟周期性暴发进行了预测研究已取得成效(罗礼智等,1999;黄绍哲等,2008);国内利用历期法,有效蛾峰日 3~4 级,雌蛾超过 50% 向后推 10~14 d(日均温 20~26℃ 条件下,产卵前期 4~6 d,卵期 3~4 d,1 龄期 3~4 d)即为 2 龄幼虫期,即为防治适期,对草地螟开展中短期预报,准确率在 90% 以上(张跃进,2006)。查阅文献,目前国内对草地螟长期或超长期预报多为经验预测,预报准确率低,需要考虑因素多,繁琐,技术要求高,需要

有丰富实践经验,一般测报人员很难掌握。而马尔科夫链与其他方法不同,不需考虑许多复杂的外界因子,只需有多年的病虫害发生程度历史资料,通过计算机辅助软件很快计算出结果,基层测报人员很容易掌握和应用,一些因子通过计算和观察很容易获得。因此,应用马尔科夫链方法对 1 代草地螟开展长期或超长期预报具有简便易行,使用准确,可以今后在预测预报中使用。

通过对 1 代草地螟马尔科夫链长期预测模型回检,回检历史符合率达到 89.9%,对于草地螟这样一种大尺度,迁飞性,周期性暴发害虫来说,预报准确率是令人满意的,况且参与分析数据还是 2 个不完整发生周期。2008 年后,进入轻发生期,预测年份准确率提高,之后预测模型回检率可达 90% 以上。从回检结果分析来看,出现预报值与实测值不相吻合的年份主要出现在每个暴发周期起始、结束的 1~2 年,进入每个暴发周期和结束年份后,预报准确率高。对出现预测偏差的年份,还可通过中短期预报来弥补,指导当地草地螟的防治。

参考文献(References)

- 陈观浩,2003.应用马尔科夫链法预测晚稻稻飞虱发生程度.昆虫知识,40(2):176—178.
- 胡伯海,姜瑞中,1997.农作物病虫害长期运动规律与预测.北京:中国农业出版社.43—46.
- 黄绍哲,姜幸福,雷朝亮,2008.草地螟周期性大发生与太阳黑子活动相关性.生态学报,28(10):4823—4829.
- 贾春生,2006.利用马尔科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.
- 姜玉英,张跃进,杨宝胜,2009.草地螟 2008 年越冬虫源分布特点和 2009 年发生趋势分析.中国植保导刊,29(1):39—41.
- 康爱国,杨海珍,李强,2004.草地螟越冬代虫量与第一代草地螟发生关系的研究.昆虫知识,41(1):70—72.
- 罗礼智,李光博,曹雅忠,1999.草地螟第三个猖獗危害周期已经来临.植物保护,36(1):11—14.
- 罗礼智,张红杰,康爱国,1998.1997 年张家口市一代草地螟大发生原因分析.自然灾害学报,7(3):119—123.
- 台文志,2008.利用马尔科夫链模型预测股票市场的近期走势.西南民族大学学报(自然科学版),3(3):447—481.
- 唐启义,冯明光,1997.实用统计分析及其计算机处理平台.北京:中国农业出版社.1—407.

田红霞, 郭晨, 2008. 马尔科夫链在预测太原市降水量中的应用. 太原科技大学学报, 29(6):444—446.

谢成君, 1997. 用马尔科夫链法预报黑绒金龟发生程度. 昆虫知识, 34(2):88—90.

闫香慧, 赵志模, 刘怀, 2009. 应用马尔科夫链理论对褐飞虱和白背飞虱发生程度的预测. 生态学报, 29(11):

5799—5805.

张孝义, 张跃进, 2006. 农作物有害生物预测学. 北京: 中国农业出版社. 456—457.

张跃进, 2006. 农作物有害生物测报技术手册. 北京: 中国农业出版社. 59—69.