

应用马尔可夫链模型预测长爪沙鼠 和黑线仓鼠种群数量

武文华^① 付和平^① 武晓东^{①*} 杨玉平^② 董维惠^② 徐胜利^③

(^① 内蒙古农业大学生态环境学院 呼和浩特 010019; ^② 中国农业科学院草原研究所 呼和浩特 010021;

^③ 内蒙古水利科学研究院 呼和浩特 010020)

摘要: 依据1984~2004年21年在呼和浩特郊区对长爪沙鼠(*Meriones unguiculatus*)和黑线仓鼠(*Cricetus barabansis*)种群数量动态的调查数据,结合优选法(0.618法),应用加权马尔可夫链预测模型,建立了长爪沙鼠和黑线仓鼠种群数量预测模型。应用此模型分别预测了长爪沙鼠和黑线仓鼠2004年的种群数量,预测值与实测值相符,预测准确。同时,预测了未来三年(2005~2007年)长爪沙鼠和黑线仓鼠的种群数量。结果表明,此方法计算简便、准确、可靠,为鼠类种群数量的预测提供了新的方法。

关键词: 长爪沙鼠;黑线仓鼠;马尔可夫链模型;预测

中图分类号:Q958 文献标识码:A 文章编号:0250-3263(2007)06-69-10

Forecasting the Population Dynamics of *Meriones unguiculatus* and *Cricetus barabansis* by Applying Markov Model

WU Wen-Hua^① FU He-Ping^① WU Xiao-Dong^{①*} YANG Yu-Ping^②
DONG Wei-Hui^② XU Sheng-Li^③

(^① College of Ecology and Environmental Science, Inner Mongolia Agricultural University, Hohhot 010019;

^② Grassland Institute of the Chinese Academy of Agricultural Science, Hohhot 010021;

^③ Institute of Water Conservancy Science of Inner Mongolia, Hohhot 010020, China)

Abstract: The forecast model for population dynamics of *Meriones unguiculatus* and *Cricetus barabansis* was built by using Markov chain, combining with optimization method (0.618) based on of population dynamics data of these two species that collected in Hohhot suburbs of Inner Mongolia from 1984 to 2004. And a forecast on the population dynamics of *M. unguiculatus* and *C. barabansis* in 2004 was made in this paper. The result showed that the forecast was corresponded to the field data accurately. The Markov model was applied to forecast the population dynamics of *M. unguiculatus* and *C. barabansis* for three years, from 2005 to 2007. The result showed that the model provided a handy, accurate and reliable method for the prediction on rodent population dynamics.

Key words: *Meriones unguiculatus*; *Cricetus barabansis*; Markov model; Forecast

长爪沙鼠(*Meriones unguiculatus*)是北方农牧业的主要害鼠之一。在农区长爪沙鼠盗食并贮存大量粮食,严重时可减产20%~25%,在牧区消耗数量可观的牧草,破坏土层结构,同时长爪沙鼠是鼠疫病原体的主要贮存宿主^[1]。长爪沙鼠常见于荒漠草原,也分布于典型草原和农业地区。在疏松的沙质土壤、背风向阳、坡度不大并长有茂密的白刺(*Nitraria* spp.)、滨藜

(*Atriplex patens*)及小画眉草(*Eragrostis minor*)等

基金项目 国家自然科学基金(No. 30160019, 30560028),教育部“春晖计划”与内蒙古自然科学基金(No. 200508010511)项目资助;

* 通讯作者, E-mail: wuxiaodong_hgh@163.com;

第一作者介绍 武文华,女,硕士研究生;主要从事草原野生动物生态和草地保护的研究; E-mail: wuwenhua19800229@163.com。

收稿日期:2006-12-21, 修回日期:2007-07-14

植物的环境条件常常成为它们栖息的最适生境。长爪沙鼠种群数量的年间变化较大,季节变化较小。黑线仓鼠(*Cricetus barabansis*)对农业生产有较大的危害,一方面消耗部分粮食,另一方面在贮粮的过程中还要糟蹋更多的种子,在牧区则影响牧草的更新,同时也是鼠疫和钩端螺旋体病的贮存宿主^[1]。黑线仓鼠的栖息环境极为广泛,草原、半荒漠、农田、山坡及河谷的林缘、灌丛中都可栖息,在半荒漠地区,通常栖息于蒿草较高的地方或水塘附近,在农区,多集中于田埂、土坡或农田中的坟堆上,以及人工次生林等地。黑线仓鼠在内蒙古地区种群数量的季节动态表现有两个数量高峰,分别在5月和8月^[2]。

马尔可夫链是时间和状态都离散的马尔可夫过程,应用马尔可夫链模型预测各种问题的实质是可根据 n 时刻的状态预测 $n+1$ 时刻的状态。马尔可夫链在我国最早应用大约在20世纪70年代^[3-7],目前,已经被广泛应用于各个领域,包括经济学、医学、遗传学领域,以及自然灾害的预测、教育学等。如崔振洋等应用马尔可夫链理论和方法,对某地区1991~1995年水稻稻瘟病发生程度进行了预测^[8];巴剑波等应用马尔可夫链预测了1993~2003年海军疟疾疫情趋势^[9];葛键利用马尔可夫链进行分析、计算,给出了市场预测、利润预测及风险决策的数学模型^[10]。目前,有关鼠类种群数量预测使用的方法主要为线性回归模型或逐步回归模型,此方法需要测算的因子较多,同时有些因子很难通过直观观察或简单的计算而获得,有必要寻找一种简单方便而实用的预测方法,马尔可夫链预测方法只需考虑统计对象本身历史状态的演变即可,而现在较少应用马尔可夫链对鼠类种群数量进行预测。本文应用马尔可夫链模型预测了长爪沙鼠和黑线仓鼠2005~2006及2007年的种群捕获率,对应用马尔可夫链模型预测鼠类种群数量进行了研究。

1 自然概况与研究方法

1.1 自然概况 研究地点位于呼和浩特郊区

“农业部沙尔沁乡牧草资源重点野外科学观测试验站”(原中国农业科学院草原研究所呼和浩特市试验场)和托克托县永圣域乡。两个监测点分别在呼和浩特东南30 km和西南40 km处,两地相距约50 km。属于大陆性干旱气候,年平均气温6.8℃,最高气温37.3℃,最低气温-32.8℃,无霜期约130 d左右,年均降水量423 mm,境内有农田、栽培牧草、放牧地和林地等。

1.2 野外调查方法 在以上两个监测点采用直线夹日法调查鼠类种类和数量,两个监测点的基本生境条件相同。由于选择的监测点要作为长期的测报点,长时间的夹捕对鼠类种群数量会产生影响,鼠类数量的测报会出现误差。为了使研究结果更准确、更符合实际,减少误差,选择了两个调查点。1984~2005年每年的4~10月每月中旬进行一次野外夹捕,夹距5 m,行距约50 m,在每个研究地内取4个样方,每个样方100~200个夹日。每次选择的样方与前一次选择的样方不能重叠,在研究区域内随机选取。以花生米做诱饵,在不同生境内进行调查,记录捕获啮齿动物的种类和数量。以每百夹日捕获率表示种群相对数量或相对密度,同时进行测量、称重,并解剖记录相关的繁殖情况。

2 数学模型应用

2.1 马尔可夫链定义和性质^[11] 定义:设马尔可夫过程 $\{X_n, n \in T, \}$ 的参数集 T 是离散的时间集合, $T = \{0, 1, 2, \dots\}$, X_n 的所有可能取值的全体 $\{X_n\}$ 是离散的状态空间,记为 $E = \{x_1, x_2, \dots\}$,若对任意的正整数 $n \in T$ 及任意的 $x_1, x_2, \dots, x_n, x_{n+1} \in E$,都有 $P(X_{n+1} = x_{n+1} | X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = P(X_{n+1} = x_{n+1} | X_n = x_n)$ 则称 $\{X_n\}$ 为马尔可夫链。

在马尔可夫链中, $P_{ij}^{(n)} = P\{X_{n+1} = j | X_n = i\}$ 称为马尔可夫链 $\{X_n, n \in T, \}$ 在时刻 n 的1步转移概率,其中 $i, j \in E$,1步转移概率可用矩阵 P 表示:

$$P = \begin{bmatrix} P_{11} & P_{12} & \Lambda & P_{1m} \\ P_{21} & P_{22} & \Lambda & P_{2m} \\ \Lambda & \Lambda & & \\ P_{m1} & P_{m2} & \Lambda & P_{mm} \end{bmatrix}$$

t 到 $t+1$ 时刻,状态从 S_i 转移为 S_j 的频数 n_{ij} 与总频数 n 之比则为状态 S_i 转移为 S_j 的 1 步转移概率。

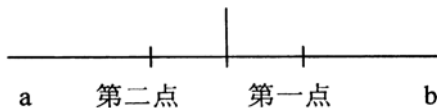
$$P_{ij} \geq 0 \text{ 且 } \sum_{j=1}^m P_{ij} = 1 \quad (i, j = 1, 2, \dots, m)^{[12]}。$$

一般地,转移概率 $P_{ij}^{(n)}$ 不仅与状态 i, j 有关,而且与时刻 n 有关。当 $P_{ij}^{(n)}$ 不依赖于时刻 n 时,表示马尔可夫链具有平稳转移概率。若对任意的 $i, j \in E$,马尔可夫链 $\{X_n, n \in T\}$ 的转移概率 $P_{ij}^{(n)}$ 与 n 无关,则称马尔可夫链是齐次的。马尔可夫链预测方法有三种:基于绝对分布的马尔可夫链预测、叠加马尔可夫链预测、加权马尔可夫链预测。本论文采用加权马尔可夫链预测。

2.2 加权马尔可夫链预测过程

2.2.1 对历史数据进行状态划分

本文采用优选法(0.618 法)对历史数据进行状态等级的划分,所谓优选法,是抓住具体问题的主要矛盾,运用数学的原理和方法,合理地安排试验点,以尽可能少的试验次数,迅速地找到最优点的科学试验方法^[13]。试验点的计算采用下列公式:



$$X_1 = a + 0.618(b - a)$$

$$X_2 = a + b - X_1$$

计算出的所有试验点根据马尔可夫链各阶转移概率矩阵的所有元素非负、所有各行之和约为 1 的特征及各种状态的概率大致相近的原则确定最后的分级标准。

2.2.2 马尔可夫性检验

随机序列是否具有马尔可夫性,是应用马尔可夫链模型分析和解决实际问题的必要前提,通常离散序列的马尔

可夫性可用 x^2 统计量来检验。

$$x^2 = 2 \sum_{i=1}^m \sum_{j=1}^m f_{ij} \left| \log \frac{P_{ij}}{P_{*j}} \right|$$

其中 P_{*j} 称为边际概率, $P_{*j} = \frac{\sum_{i=1}^m f_{ij}}{\sum_{j=1}^m f_{ij}}$ 表示

历史数据序列有 m 个可能的状态,用 f_{ij} 表示历史数据序列从状态 i 经过 1 步转移到状态 j 的频数, $i, j \in E$,其中 P_{ij} 为 1 步转移概率。

当 n 充分大时,统计量服从自由度为 $(m-1)^2$ 的 x^2 分布,给定显著性水平 α ,查表可得分位点 $x_{\alpha}(m-1)^2$ 值,若 $x^2 > x_{\alpha}(m-1)^2$,则可认为序列 $\{X_n\}$ 符合马尔可夫性,否则认为该序列不能作为马尔可夫链来处理。

2.2.3 转移概率

较多概率根据历史数据序列的状态,写出频率矩阵和各阶转移概率矩阵。

各阶转移概率矩阵 $P_{ij} = \frac{f_{ij}}{\sum_{j=1}^m f_{ij}}$ 表示历史数据序

列包含了 m 个状态,用 f_{ij} 表示历史数据序列中,从状态 i 经过 1 步 2 步 3 步或 m 步转移到状态 j 的频数, $i, j \in E$,将转移频数矩阵的第 i 行第 j 列元素 f_{ij} 除以各行的总和和所得的值称为转移概率。

2.2.4 计算马尔可夫链的权重

首先计算各阶自相关系数 $r_k, k \in E$,对各阶自相关系数规范化,即 $w_k = |r_k| / \sum_{k=1}^m |r_k|$,将 w_k 作为各阶时滞(步长)的马尔可夫链的权重, m 为按时间需要计算到的最大阶数。

2.2.5 预测 $n+1$ 时段的状态

分别以前面若干年份的历史数据为初始状态,结合其相应的各阶转移概率矩阵预测该段的状态概率 $P_i^{(k)}, i \in E, k$ 为时滞(步长), $k = 1, 2, \dots, m$ 。将同一状态各预测概率加权和作为指标值处于该状态的预测概率,即 $P_{(j)} = \sum_{k \in E} w_k P_i^{(k)}$,时刻 $n+1$ 时所处状态 j 满足 $P_{(j)} = \max\{P_{(i)}\} i \in E$ 。

待 $n+1$ 时段的状态确定后,将其加入到原始序列中,再重复 1~5,预测 $n+2$ 时的状

态,以此类推。

3 种群数量的预测过程及结果分析

3.1 长爪沙鼠和黑线仓鼠马尔可夫链模型的建立

连续不间断21年的野外调查获得1984~2004年长爪沙鼠和黑线仓鼠种群数量(捕获率)实测数据(表1和表2),以2004年长爪沙鼠和黑线仓鼠种群捕获率作为模型检验预测的实例。

表1 1984~2004年长爪沙鼠种群捕获率实测值

Table 1 The measured values on the captured rate of *Meriones unguiculatus* population (1984-2004)

年度	捕获率(%)	年度	捕获率(%)	年度	捕获率(%)
Year	Capture rate	Year	Capture rate	Year	Capture rate
1984	0	1985	0	1986	0
1987	0	1988	0	1989	0.05
1990	0.06	1991	0.05	1992	2.04
1993	2.92	1994	6.26	1995	4.82
1996	1.73	1997	0.86	1998	0.55
1999	0.32	2000	0.27	2001	0.43
2002	1.44	2003	0.05	2004	0.02

表2 1984~2004黑线仓鼠种群捕获率实测值

Table 2 The measured values on the captured rate of *Citellus barabansis* population (1984-2004)

年度	捕获率(%)	年度	捕获率(%)	年度	捕获率(%)
Year	Capture rate	Year	Capture rate	Year	Capture rate
1984	7.74	1985	4.22	1986	2.17
1987	1.30	1988	1.35	1989	1.10
1990	0.75	1991	1.18	1992	1.19
1993	1.23	1994	0.72	1995	0.25
1996	0.85	1997	2.01	1998	1.65
1999	1.74	2000	2.20	2001	2.30
2002	0.87	2003	1.68	2004	1.71

3.1.1 状态等级划分 根据1984~2004年长爪沙鼠和黑线仓鼠种群数量的调查数据表1、表2,从表1数据中得出1984~2004年长爪沙鼠的最大捕获率为6.26,最小的为0。根据优选法首先做第一点。

第一点: $0 + (6.26 - 0)0.618 = 3.87$

第二点: $0 + 6.26 - 3.87 = 2.39$

第三点: $0 + 3.87 - 2.39 = 1.48$

第四点: $0 + 2.39 - 1.48 = 0.91$

第五点: $0 + 1.48 - 0.91 = 0.57$

第六点: $0 + 0.91 - 0.57 = 0.34$

第七点: $0 + 0.57 - 0.34 = 0.23$

第八点: $0 + 0.34 - 0.23 = 0.11$

再将第一次余下的线段左折:

第一点: $3.87 + (6.26 - 3.87)0.618 = 5.35$

第二点: $3.87 + (6.26 - 5.35) = 4.79$

由上述数据,根据马尔可夫链各阶转移概率矩阵的所有元素非负,所有各行之和约为1的特征及各种状态的概率大致相近的原则,将长爪沙鼠种群捕获率划分为五个等级,分别为:0.11以下为一级,0.11~0.34为二级,0.34~2.39为三级,2.39~3.87为四级,3.87以上为五级。同理,由表2实测数据将黑线仓鼠种群捕获率状态进行划分,划分的等级分别为:0.93以下为一级,0.93~1.34为二级,1.34~3.11为三级,3.11以上为四级。长爪沙鼠和黑线仓鼠种群捕获率状态等级划分,分别见表3和表4。

表3 1984~2004年长爪沙鼠捕获率状态等级划分表

Table 3 Statistics of state hierarchical division on the captured rate of *Meriones unguiculatus* (1984-2004)

年度	捕获率(%)	等级	年度	捕获率(%)	等级
Year	Capture rate	Hierarchy	Year	Capture rate	Hierarchy
1984	0	1	1995	4.82	5
1985	0	1	1996	1.73	3
1986	0	1	1997	0.86	3
1987	0	1	1998	0.55	3
1988	0	1	1999	0.32	2
1989	0.05	1	2000	0.27	2
1990	0.06	1	2001	0.43	3
1991	0.05	1	2002	1.44	3
1992	2.04	3	2003	0.05	1
1993	2.92	4	2004	0.02	1
1994	6.26	5			

3.1.2 马尔可夫链的检验 由表1、表2长爪沙鼠和黑线仓鼠种群捕获率的实测数据计算频度转移矩阵 f_{ij} 和1步转移概率矩阵 $P_{(1)}$,由表1计算长爪沙鼠频度转移矩阵为:

$$f_{ij} = \begin{bmatrix} 7 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

表4 1984~2004年黑线仓鼠捕获率状态等级划分表
Table 4 Statistics of state hierarchical division on the captured rate of *Cricetus barabansis* (1984-2004)

年度	捕获率(%)	等级	年度	捕获率(%)	等级
Year	Capture rate	Hierarchy	Year	Capture rate	Hierarchy
1984	7.74	4	1995	0.25	1
1985	4.22	4	1996	0.85	1
1986	2.17	3	1997	2.01	3
1987	1.30	2	1998	1.65	3
1988	1.35	3	1999	1.74	3
1989	1.10	2	2000	2.20	3
1990	0.75	1	2001	2.30	3
1991	1.18	2	2002	0.87	1
1992	1.19	2	2003	1.68	3
1993	1.23	2	2004	1.71	3
1994	0.72	1			

结合1步转移概率矩阵 $P_{(1)}$,

$$P_{(1)} = \begin{bmatrix} 0.875 & 0 & 0.125 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 & 0 \\ 0.1667 & 0.1667 & 0.5 & 0.1667 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0.50 & 0.5 & 0 \end{bmatrix}$$

可计算长爪沙鼠的统计量 X^2 的值34.445,同理计算黑线仓鼠的 X^2 为16.987,在给定的显著性水平 $\alpha = 0.05$ 查表可得 $x_{\alpha}^2(16) = 26.296$, $x_{\alpha}^2(9) = 16.919$,由于 $X^2 > x_{\alpha}^2[(m-1)^2]$,因此长爪沙鼠和黑线仓鼠的历史数据满足马尔可夫性,可作为马尔可夫链来处理。

3.1.3 各阶马尔可夫链的权重 由表1、表2实测数据计算长爪沙鼠和黑线仓鼠1984~2003年捕获率的平均值分别为 $\bar{x}_{长} = 1.0925$, $\bar{x}_{黑} = 1.825$,根据1984~2003年实测数据和平均值计算长爪沙鼠和黑线仓鼠的自相关系数,由自相关系数计算权重,结果见表5、表6。

表5 长爪沙鼠各阶自相关系数和各个步长的马尔可夫链的权重

Table 5 Each orders auto correlation coefficient and the weigh of Markov chain of each steps of *Meriones unguiculatus*

项目	时滞Time-delay(年 Years)				
	1	2	3	4	5
自相关系数 r_k	0.7082	0.3358	-0.0060	-0.2221	-0.3073
权重 w_k	0.4484	0.2126	0.0038	0.1406	0.1945

表6 黑线仓鼠各阶自相关系数和各个步长的马尔可夫链的权重

Table 6 Each orders auto correlation coefficient and the weigh of Markov chain of each steps of *Cricetus barabansis*

项目	时滞Time-delay(年 Years)			
	1	2	3	4
自相关系数 r_k	0.4228	0.1014	-0.0012	-0.0180
权重 w_k	0.7780	0.1866	0.0023	0.0331

3.1.4 模型的计算与检验 根据呼和浩特郊区1984~2003年长爪沙鼠和黑线仓鼠20年的捕获率数据(2004年数据除外,留做对模型的检验),计算长爪沙鼠种群捕获率各阶转移概率矩阵为:

$$P_{(1)} = \begin{bmatrix} 0.875 & 0 & 0.125 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 & 0 \\ 0.1667 & 0.1667 & 0.1667 & 0.1667 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0.5 & 0 & 0.5 \end{bmatrix}$$

$$P_{(2)} = \begin{bmatrix} 0.7865 & 0.0208 & 0.1719 & 0.0208 & 0 \\ 0.0833 & 0.3333 & 0.5 & 0.833 & 0 \\ 0.2292 & 0.1667 & 0.3542 & 0.0833 & 0.1667 \\ 0 & 0 & 0.5 & 0 & 0.5 \\ 0.0833 & 0.0833 & 0.5 & 0.0833 & 0.625 \end{bmatrix}$$

$$P_{(3)} = \begin{bmatrix} 0.7168 & 0.0391 & 0.1947 & 0.0286 & 0.0208 \\ 0.1563 & 0.2500 & 0.4271 & 0.0833 & 0.0833 \\ 0.2595 & 0.1424 & 0.3724 & 0.0590 & 0.1667 \\ 0.0833 & 0.0833 & 0.5 & 0.0833 & 0.25 \\ 0.1563 & 0.1250 & 0.4271 & 0.0833 & 0.2083 \end{bmatrix}$$

$$P_{(4)} = \begin{bmatrix} 0.6596 & 0.0520 & 0.2169 & 0.0324 & 0.0391 \\ 0.2079 & 0.1962 & 0.3997 & 0.0711 & 0.1250 \\ 0.2892 & 0.1332 & 0.3732 & 0.0621 & 0.1424 \\ 0.1563 & 0.1250 & 0.4271 & 0.0833 & 0.2083 \\ 0.2079 & 0.1337 & 0.3998 & 0.0712 & 0.1875 \end{bmatrix}$$

$$P_{(5)} = \begin{bmatrix} 0.6133 & 0.0621 & 0.2364 & 0.0361 & 0.0520 \\ 0.2485 & 0.1647 & 0.3864 & 0.0666 & 0.1337 \\ 0.3152 & 0.1288 & 0.3605 & 0.0621 & 0.1332 \\ 0.2079 & 0.1337 & 0.3997 & 0.0711 & 0.1875 \\ 0.2485 & 0.1335 & 0.3864 & 0.0666 & 0.1649 \end{bmatrix}$$

根据各阶转移概率矩阵、各状态的权重值和初始年的状态,预测2004年长爪沙鼠的捕获率状态(表7)。

表7 2004 年长爪沙鼠种群捕获率情况预测表

Table 7 Forecast of capture rate of *Meriones unguiculatus* population of 2004

初始年 Initial year	状态 State	时滞 Time-delay (年 Years)	状态权重 The stated weight	1	2	3	4	5	概率来源 The source of probability
2003	1	1	0.448 4	0.875	0	0.125	0	0	$P_{(1)}$
2002	3	2	0.212 6	0.229 2	0.166 7	0.354 2	0.083 3	0.166 7	$P_{(2)}$
2001	3	3	0.003 8	0.259 5	0.142 4	0.372 4	0.059 0	0.166 7	$P_{(3)}$
2000	2	4	0.140 6	0.207 9	0.196 2	0.399 7	0.071 2	0.125 0	$P_{(4)}$
1999	2	5	0.194 5	0.248 5	0.164 7	0.386 4	0.066 6	0.133 7	$P_{(5)}$
加权值 $\sum p_i$ The weighted value				0.519 6	0.095 6	0.264 2	0.040 9	0.079 7	

同理,计算1984~2003 年黑线仓鼠种群捕获率各阶转移概率矩阵如下:

$$P_{(1)} = \begin{bmatrix} 0.4 & 0.2 & 0.4 & 0 \\ 0.4 & 0.4 & 0.2 & 0 \\ 0.142 9 & 0.285 7 & 0.571 4 & 0 \\ 0 & 0 & 0.5 & 0.5 \end{bmatrix}$$

$$P_{(2)} = \begin{bmatrix} 0.297 1 & 0.274 3 & 0.428 6 & 0 \\ 0.348 6 & 0.297 1 & 0.354 3 & 0 \\ 0.253 1 & 0.306 1 & 0.440 8 & 0 \\ 0.071 4 & 0.142 9 & 0.535 7 & 0.25 \end{bmatrix}$$

$$P_{(3)} = \begin{bmatrix} 0.289 8 & 0.291 6 & 0.418 6 & 0 \\ 0.308 9 & 0.289 8 & 0.401 3 & 0 \\ 0.286 6 & 0.299 0 & 0.414 3 & 0 \\ 0.162 2 & 0.224 5 & 0.488 3 & 0.125 \end{bmatrix}$$

$$P_{(4)} = \begin{bmatrix} 0.292 4 & 0.294 2 & 0.413 4 & 0 \\ 0.296 8 & 0.292 4 & 0.410 8 & 0 \\ 0.293 5 & 0.295 3 & 0.411 2 & 0 \\ 0.224 4 & 0.261 7 & 0.451 3 & 0.062 5 \end{bmatrix}$$

从表7 可以看出,在长爪沙鼠状态栏中,将同一状态的各预测概率加权求和之后,状态“1”的概率最大为0.519 6,故2004 年长爪沙鼠的捕获率状态等级为“1”级,在0.11 以下,与实测数据0.02 相符。同理应用马尔可夫链模型预测黑线仓鼠2004 年的捕获率状态等级,计算结果见表8。由表8 可以看出,在黑线仓鼠状态栏中,将同一状态的各预测概率加权求和之后,状态“3”的概率最大为0.539 1,可见2004 年黑线仓鼠的预测状态是“3”级,故2004 年黑线仓鼠的捕获率状态等级为“3”级,在1.34~3.11 之间,与实测数据1.71 相符,预测比较准确。

3.2 应用马尔可夫链对长爪沙鼠和黑线仓鼠种群数量进行预测

把2004 年长爪沙鼠和黑线仓鼠捕获率数据放入调查数据中,计算出1984~2004 年长爪沙鼠和黑线仓鼠的各阶转移概率矩阵。根据马尔可夫链预测模型和各阶转

表8 2004 年黑线仓鼠种群捕获率情况的预测表

Table 8 Forecast of capture rate of *Citellus barabansis* population of 2004

初始年 Initial year	状态 State	时滞 Time-delay (年 Years)	状态权重 The stated weight	1	2	3	4	概率来源 The source of probability
2003	3	1	0.778 0	0.142 9	0.285 7	0.571 4	0	$P_{(1)}$
2002	1	2	0.186 6	0.297 1	0.274 3	0.428 6	0	$P_{(2)}$
2001	3	3	0.002 3	0.286 6	0.299 0	0.414 3	0	$P_{(3)}$
2000	3	4	0.033 1	0.293 5	0.295 3	0.411 2	0	$P_{(4)}$
加权值 $\sum p_i$ The weighted value				0.177 0	0.283 9	0.539 1	0	

移概率矩阵预测长爪沙鼠和黑线仓鼠未来三年(2005 2006 2007 年)的捕获率情况(表9、10)。从表9 可以看出长爪沙鼠三年的状态栏中都以

“1”的概率最大,所以未来三年的状态级别都为“1”级,所以在未来三年内长爪沙鼠的捕获率都在0.11 以下,不会造成危害,应注意其他优势

表9 2005~2007 年长爪沙鼠种群数量预测表

Table 9 Predicted statistics of population density of *Meriones unguiculatus* (2005 - 2007)

年度 Year	状态分布概率 Probability of state distribution					预测状态 Predicted state	预测范围 Predicted range
	状态1 State 1	状态2 State 2	状态3 State 3	状态4 State 4	状态5 State 5		
	2005	0.674 4	0.052 1	0.205 0	0.024 4		
2006	0.690 5	0.044 7	0.198 6	0.023 3	0.042 8	1	0.11 以下
2007	0.741 6	0.033 7	0.176 6	0.019 2	0.029 0	1	0.11 以下

种群的防治工作。

从表10 可以看出:黑线仓鼠未来三年的状态栏中都以“3” 的概率最大,所以三年的状态级别都为“3” 级,捕获率在1.34 ~ 3.11 之间,应加强黑线仓鼠的防治工作。

由2005 年的实际调查资料知,2005 年长爪沙鼠的实测捕获率为0.01,与预测值相符;黑线仓鼠的实测捕获率为1.56,捕获率在1.34 ~ 3.11 之间,预测值和实测值相符。这说明马尔可夫链对2005 年长爪沙鼠和黑线仓鼠种群数量的预测比较准确。同时利用Senseth 等^[14] 的变动指数 S 计算长爪沙鼠和黑线仓鼠种群数量变动的周期性,设 D_i 为第 i 年的鼠类平均相对密度,则各年以10 为底的对数 $\ln(D_i)$ 的标准差 S 为变动指数。当 $S < 0.5$ 时,群落或种群为非周期性波动;当 $S > 0.5$ 时,为周期性波动。计算长爪沙鼠和黑线仓鼠的变动指数,分别为

0.825 1 0.306 5,其中长爪沙鼠的 $S > 0.5$,具有周期性,黑线仓鼠 $S < 0.5$,为非周期性波动。长爪沙鼠种群的周期性变动见图1,可以看出长爪沙鼠2004 年以后正处于周期性中的低谷期。这与应用马尔可夫链模型预测的长爪沙鼠未来三年(2005 2006 2007 年)种群动态情况正好相符,也说明了应用马尔可夫链预测模型对长爪沙鼠的预测是比较准确的。

表10 2005~2007 年黑线仓鼠种群数量预测表

Table 10 Predicted statistics of population density of *Cricetus barabansis* (2005 - 2007)

年度 Year	状态分布概率 Probability of state distribution				预测状态 Predicted state	预测范围 Predicted range
	状态1 State 1	状态2 State 2	状态3 State 3	状态4 State 4		
	2005	0.150 6	0.257 0	0.592 4		
2006	0.150 6	0.257 0	0.592 4	0	3	1.34 ~ 3.11
2007	0.150 5	0.257 1	0.592 4	0	3	1.34 ~ 3.11

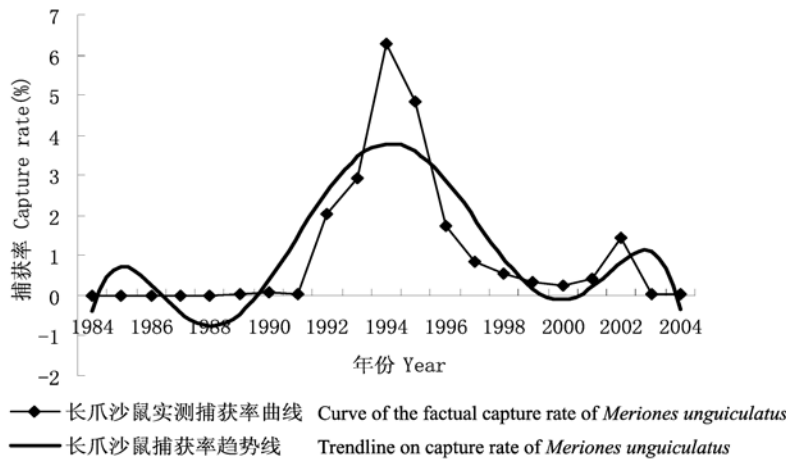


图1 1984~2004 年长爪沙鼠平均相对密度变化图

Fig.1 Variance diagram of average relative density of *Meriones unguiculatus* (1984 - 2004)

通过对长爪沙鼠和黑线仓鼠种群数量的预测,在未来三年内长爪沙鼠处于种群数量的低谷期,而黑线仓鼠处于数量的相对高峰期。因为这两种是呼和浩特郊区鼠类群落的优势种群,所以未来三年对呼和浩特郊区鼠类进行防治时,应着重加强黑线仓鼠种群数量的防治,制定相应的防治措施,减少对农田的危害。

本文利用马尔可夫链模型同时预测了长爪沙鼠和黑线仓鼠捕获率的等级状态,预测的结果比较准确。这说明对于预测长爪沙鼠和黑线仓鼠种群的动态数量,马尔可夫链模型是比较好的模型。马尔可夫链模型对数据的要求不高,只要有多年连续的种群动态数量数据,种群数量的变化符合马尔可夫链预测模型的基本思想,即事物变动过程符合马尔可夫链形成过程中事物变动的随机性和转移状态的无后效应两个主要特点,便可应用其方法进行预测,不考虑其他影响种群数量的因素。长爪沙鼠和黑线仓鼠种群数量的数据符合上面的要求,所以应用马尔可夫链模型进行种群数量预测是合理的,今后对当地这两个种群数量进行预测时,只需补充当年的种群捕获率,求出各阶的转移概率矩阵,并进行相应的概率转移,即可预测下一年的发生数量,这种方法简单易行,准确度较高。本方法在种群数量动态方面的应用是广阔的,值得人们对其进一步的开发。

4 讨论

种群数量动态是生态学研究的重要内容之一。自20世纪50年代,我国开始对小型兽类种群数量动态进行研究,但多数研究时间较短。较长时间的定点、定时研究较少,对其进行长期逐月定点定时的研究就更少。丁平等^[15]于1985~1992年在浙江省萧山围垦农区,对小型兽类种群数量动态进行了长期逐月定点定时的研究。80年代以来,我国学者对农田重要害鼠种群数量的预测预报开展了卓有成效的研究工作,对鼠类种群数量调节因子的研究也比较深入。洪朝长等^[16]利用雌性成体比、成体性比、怀孕率、胎仔数等种群特征建立了中期预测模

型;朱盛侃等^[17]、陈安国等^[18]分析了种群基数、繁殖指数、年龄结构、气候因子等同新疆小家鼠(*Mus musculus*)种群动态的关系,并建立双方程多元回归预测模型;郑智民等^[19]分析了气温、相对湿度和降雨对黄毛鼠(*Rattus rattoides*)数量的影响;汪笃栋等^[20]以怀孕率、平均胎仔数、性比和雌性成亚比等作为数量调节因子,分析了江西农田黑线姬鼠(*Apodemus agrarius*)的发生数量,建立了预测模型;何森等^[21]得出了板齿鼠(*Bandicota indica*)种群数量中长期预测的时间序列模型;王勇等^[22]对洞庭湖稻区黑线姬鼠种群数量进行了预测。有关鼠类种群数量预测使用的方法主要为线性回归模型或逐步回归模型,如姜运良等^[23]以逐步回归分析法,预报了山东阳谷县黑线仓鼠的种群数量,并得到了回归方程。董维惠等^[24]应用线性回归模型用当月捕获率分别预测下月和隔月长爪沙鼠及黑线仓鼠的捕获率,用当年10月捕获率预测黑线仓鼠翌年4月的捕获率,用当年9月捕获率分别预测长爪沙鼠翌年4月和5月的捕获率,以及用当年10月繁殖指数预测翌年4月的捕获率。而应用马尔可夫链模型对鼠类种群进行预测预报的较少。对主要害鼠数量动态进行长期监测,深入研究其数量变动规律,阐明其内在变动机制的意义,已得到生态学工作者的广泛认可。

虽然我国学者已对不同地区林业、农业、草原等常见害鼠种群数量动态及预测预报进行了研究,较成功地建立了多种数学模型,使啮齿动物种群动态预测的准确度达到了一定的水平,促进了啮齿动物种群生态学的发展,但这些模型主要是线性和非线性回归模型。而建立线性回归预测模型或逐步回归模型,需要对害鼠的生物学、生态学特征有较全面的了解,对不同的鼠种需要选择不同的生物、生态学特征及关键环境因子作为模型参量^[21],涉及的因子较多。而有些关于种群动态机制的信息在自然条件下难以直观观察或由简单计算而获得,再加上那些难以量化统计的因子以及随机因子的干扰,如人为干扰、天敌捕食及动物在栖息地间的迁移等使预测仍有偏差。马尔可夫链模型是一种

简单方便而实用的方法,是应用广泛、理论完备的随机模型,在近代统计学中有极其重要的地位。马尔可夫链与其他方法不同,不需从繁杂的测报因子中寻求各个因子间的相关规律,只需考虑统计对象本身历史状态的演变特点即可,通过计算状态转移概率预测未来状态的变化^[25],但需要积累连续多年的种群动态调查资料。

本文中1984~2004年的数据是引述的,2005年的数据是未发表的原始数据,利用马尔可夫链法进行预测时,状态等级划分的不同得到的结果有很大的差异,通常状态划分的方法有数据序列约定俗成的分组法,即样本均值-均方差分组法、有序聚类分组法,都是凭借经验、查相关的手册或经过多次试验来获得。为了使预测结果更加准确,本文采用了优选法(0.618法)进行等级划分。邵崇斌等^[25]在利用马尔可夫链预测杨树天牛发生量时即应用了优选法,得到的结果证明利用优选法进行分级,比常规的分级法进行预测明显优越,且能够进行质量分析,是有推广价值的方法之一。

马尔可夫链模型预测的结果是区间预测,虽然使预测的精度降低,却提高了预测的准确度。由于预测结果取决于各阶转移概率矩阵,并假设转移概率矩阵必须逐期保持不变,而这个矩阵是不会一成不变的,所以马尔可夫链模型近期预测结果较好,要使预测结果有更高的准确度,就要把实际调查的数据不断加入到历史数据中,不断对初始状态和状态转移概率矩阵作一定的调整,使其更符合事物客观发展的变化趋势。马尔可夫链模型预测的是否准确,关键在于各阶转移概率矩阵的可靠性^[26],因此,该预测模型要求有足够多的和准确的统计数据,历史数据越多、越精确,预测也越可靠,而这样的统计数据要求有长时期变化的详细记录,而在实际中要得到这些资料是比较困难的,需要大量的人力物力,这也是本方法的缺点之一。加权马尔可夫链模型中计算权重时对于时滞的选取直接关系到预测精度的高低,对于原数列的波动不大,时滞宜于选择年数长一些,这

样可以加重远期观测值对预测值的权数,如果数据波动较大,权重的时滞宜于选择的年数短一些,可以加重近期观测值的权数,在实际工作中时滞的多少并不易判断,由于预测的目的不同,预测对象对预测的模型精度的要求具有多样性。本文时滞选取4~5年,是由于长爪沙鼠和黑线仓鼠的寿命较短,一般1~2年,近期的观测值对预测值的影响较大,所以选择的时滞较短。本文分析了21年连续不间断的数据,得到了较好的预测结果。种群数量变动是多种因素综合作用的结果^[23],马尔可夫链预测模型对未来突发性事件难以做出准确的预测,这也是所有预测模型的弱点。

马尔可夫的预测方法有基于绝对分布的马尔可夫链预测法、叠加马尔可夫链预测法、加权马尔可夫链预测法。基于绝对分布的马尔可夫链预测方法是指对于一系列相依的随机变量,用步长为1的马尔可夫链模型和初始分布推算出未来时段的绝对分布来做预测分析^[27];叠加马尔可夫链预测法是指利用各阶(多步长)马尔可夫链求得的绝对分布叠加来做预测分析^[28];加权马尔可夫链预测法先考虑分别依其前面若干时段的历史数据值的状态进行预测,然后按前面各年与该年相依关系的强弱进行加权求和,充分合理地利用信息进行预测。对于前两种预测方法都存在一定的局限性,对于基于绝对分布的马尔可夫链预测方法,默认所有的马尔可夫链满足“齐次性”缺乏依据,另外该法没有考虑到对应各阶(各种步长)马尔可夫链的绝对分布在预测中所起的作用,因此没有能充分利用已知数据资料的信息,而对于叠加马尔可夫链预测方法没有考虑到各阶马尔可夫链对应的绝对概率在叠加中所起的作用,即认为各阶马尔可夫链的绝对概率所起的作用是相同的,事实上满足马尔可夫性的相依时间序列,其各阶自相关性是不一致的^[29]。一系列相依的随机变量,其各阶自相关系数刻画了各种时滞状态间的相关关系的强弱,因此应用加权马尔可夫链预测方法能充分合理地利用信息进行预测。张宗国也证明了加权马尔可夫链预测模型预测效果优

于基于绝对分布的马尔可夫链法预测*。因此本文选用了加权马尔可夫链预测方法。

在某些情况下需要计算数据的具体值,如何根据最后计算出的状态概率分布求出某一数据的具体值仍是一个有待解决的问题。

参 考 文 献

- [1] 甘肃农业大学编著. 草原保护学 第一分册 草原啮齿动物学(第二版). 北京:中国农业出版社,1984 229~241.
- [2] 侯希贤,董维惠,张鹏利等. 呼和浩特栽培牧草地黑线仓鼠生态学的调查. 中国草地,1989,11(5):53~57.
- [3] 王梓坤,朱成熹,李漳南等. 地震迁移的统计预报. 地质科学,1973,(4):294~306.
- [4] 王启明,孙惠文,徐道一等. 地震迁移的统计预报. 数学学报,1974,(1):5~19.
- [5] 王梓坤,吴荣,钱尚玮. 预测大地震的一种方法. 数学物理学报,1975,18(2):118~126.
- [6] 王梓坤,吴荣. 华北地区地震的统计预报(一). 南开大学学报(自然科学版),1977,(1):8~15.
- [7] 王梓坤,钱尚玮. 华北地区地震的统计预报(二). 南开大学学报(自然科学版),1977,(1):16~24.
- [8] 崔振洋,李晓亮,王伟. 马尔可夫链预测模型及其在农业病虫害预报中的应用. 山西农业大学学报,1994,14(1):96~98.
- [9] 巴剑波,方旭东,徐雄利. 马尔可夫链在海军疟疾疫情预测中的应用. 解放军预测医学杂志,2001,19(2):114~116.
- [10] 葛键. 马尔可夫链在经济预测上的应用. 陕西经贸学院学报,2000,13(4):97~99.
- [11] 刘嘉,王公恕. 应用随机过程(第二版). 北京:科学出版社,2004 90~168.
- [12] 吕雄编著. 概率论与数理统计. 呼和浩特:内蒙古大学出版社,2000,191~195.
- [13] 陕西省卫生系统编著. 优选法在医药上的应用. 西安:陕西省卫生局,1975,1~11.
- [14] Stenseth N C, Framstad E. Reproductive effort and optimal reproductive rates in small rodents. *OIKOS*, 1980, 34:23.
- [15] 丁平,鲍毅新,诸葛阳. 萧山围垦农区小型兽类种群动态的研究. 兽类学报,1994,14(1):35~42.
- [16] 洪朝长,袁高林,郑本栋. 黄毛鼠的种群动态研究及数量预测的意见. 兽类学报,1989,9(2):137~145.
- [17] 朱盛侃,陈安国. 小家鼠生态特征与预测. 北京:科学出版社,1993,197~250.
- [18] 陈安国,朱盛侃,李春秋等. 新疆北部农业区害鼠的研究(六) 小家鼠种群数量消长同气候的关系. 见:青海省生物研究所编著. 灭鼠和鼠类生物学研究报告. 北京:科学出版社,1981,69~93.
- [19] 郑智民,黄应修. 黄毛鼠种群数量季节变动及其影响因素的研究. 兽类学报,1988,8(3):199~207.
- [20] 汪笃栋,叶正襄,龙丘陵等. 用逐步回归方法建立鼠类预报模型的探讨. 兽类学报,1991,11(3):238~240.
- [21] 何森,林继球,翁文英. 板齿鼠种群数量中长期预测的时间序列模型. 兽类学报,1996,16(4):297~301.
- [22] 王勇,陈安国,郭聪等. 洞庭湖湖区黑线姬鼠种群数量预测. 兽类学报,1997,17(2):125~130.
- [23] 姜运良,卢浩泉,李玉春等. 山东阳谷县黑线仓鼠种群数量预测预报. 兽类学报,1994,14(3):195~202.
- [24] 董维惠,侯希贤,杨爱莲等. 鼠类种群数量动态及持续控制研究. 呼和浩特:内蒙古大学出版社,2005,21~28.
- [25] 邵崇斌,董丽芬,赵文英等. 应用马尔科夫链预测杨树天牛发生量. 陕西林业科技,1996,(4):28~37.
- [26] 崔振洋,李晓亮,王伟. 马尔可夫链预测模型及其在农业病虫害预报中的应用. 山西农业大学学报,1994,14(1):96~98.
- [27] 冯强,王昂生. 用马尔可夫模型预测长江中下游地区旱涝灾害的试验. 中国减灾,1997,7(4):33~39.
- [28] 孙才志,张戈,林学钰. 加权马尔可夫链在降水丰枯状况预测中的应用. 系统工程理论与实践,2003,(4):100~105.
- [29] 冉景江,赵燮京,梁川. 基于加权马尔可夫链的降水预测应用研究. 人民黄河,2006,4(28):32~34.

* 张宗国. 马尔可夫链预测方法及其应用研究. 南京:河海大学硕士学位论文,2005,3~50.