

褐飞虱发生的相空间线性回归预测模型

许晓风¹, 马飞², 丁宗泽³, 程遐年²

(1. 南京师范大学生命科学学院, 南京 210097;

2. 南京农业大学, 农业部病虫监测与治理重点开放实验室, 南京 210095;

3. 江苏省太湖地区农科所, 苏州 215155)

摘要: 基于混沌理论, 通过对江苏省太湖地区农科所褐飞虱发生时间序列资料的分析, 组建了褐飞虱 *Nilaparvata lugens* Stål 发生时间序列的相空间线性回归预测模型。结果表明, 7 次预报的平均相对误差为 22.91%, 转化为发生等级, 预测准确率为 100%, 这为害虫的长期可预测性提供了一种有效的新方法。

关键词: 褐飞虱; 混沌; 相空间; 预测模型

中图分类号: Q968.1 文献标识码: A 文章编号: 0454-6296 (2002) 04-0548-04

A regression model of phase space for predicting brown planthopper (*Nilaparvata lugens* Stål) occurrence

XU Xiao-Feng¹, MA Fei², DING Zong-Ze³, CHENG Xia-Nian² (1. Life Science School of Nanjing Normal University, Nanjing 210097, China; 2. Key Laboratory of Monitoring and Management of Plant Diseases and Insects, Ministry of Agriculture, Nanjing Agricultural University, Nanjing 210095, China; 3. Taihu District Institute of Agricultural Science, Suzhou 215155, China)

Abstract: A regression model of phase space for predicting brown planthopper (BPH) occurrence in the rice growing area of Taihu District, Jiangsu Province, was established on the basis of chaos theory. Its accuracy was very high with a relative error of only 22.91%. Therefore, the model provides a new and more efficient method for the prediction of the long-term occurrence of this insect pest.

Key words: brown planthopper (BPH); chaos; phase space; regression prediction model

褐飞虱 *Nilaparvata lugens* Stål 是东南亚各国、日本、朝鲜、我国长江流域及以南稻区主要迁飞性害虫, 近二十多年来, 发生频次显著增加, 危害程度日趋严重。特别是 1987 年褐飞虱在全国大范围内暴发成灾, 发生面积占全国水稻面积的 45%, 损失稻谷 200 万吨; 1991 年, 褐飞虱不仅在全国南方稻区大发生, 天津、河北及西南部分省份的稻田也突遭危害, 直接经济损失 50 亿元人民币; 1997 年又遭受全国性的褐飞虱大发生, 给国民经济和农民带来了巨大损失 (胡伯海和姜瑞中, 1997)。显然, 如能对褐飞虱的灾害发生作出准确及超前预报, 势必减少损失, 推动农业可持续发展, 造福人类。

统计分析法是目前褐飞虱长期预测工作中使用较为广泛的预报方法之一。该方法计算简单、快

速, 主要缺点是, 当预报模型不合理或预报因子选取不当时, 预报效果不理想。近期一些研究论文 (Ma *et al.*, 2000; 马飞等, 2001) 不仅报道了褐飞虱发生系统具有分形性质和混沌特性, 而且还较为系统地阐述了褐飞虱发生系统的可预报尺度, 分析了长期预测准确性不稳定的内在原因。随着混沌理论和应用技术研究的不断深入, 混沌时间序列的建模和预测已成为混沌信息处理领域中的研究热点 (王明进和程乾生, 1997; 程乾生等, 2000; 张家树和肖先赐, 2000)。作者基于相空间重构理论, 提出了用于时间序列预测的相空间线性回归预测模型, 并用江苏省太湖地区农科所调查记载的褐飞虱发生量数据资料对该模型的预测效果进行了检验。

基金项目: “973” 项目 (G20000016210) 和 “948” 项目 (201065)

第一作者简介: 许晓风, 男, 1959 年生, 博士, 教授, 从事昆虫学与分子生物学研究, E-mail: xuxiaofeng@mail.njnu.edu.cn

收稿日期 Received: 2001-06-08; 接受日期 Accepted: 2002-03-12

1 原理与方法

1.1 相空间重构

设时间序列 $\{x(t_n)\}$ ，其时间间隔是 Δt （单位时间）。显然，吸引子的结构特性就包含在这时间序列中。为了从实际的单变量时间序列中估计出系

$$\begin{array}{cccc}
 x(t_1) & x(t_2) & \cdots x(t_i) & \cdots x(t_n - (d - 1)\tau) \\
 x(t_1 + \tau) & x(t_2 + \tau) & \cdots x(t_i + \tau) & \cdots x(t_n - (d - 2)\tau) \\
 x(t_1 + 2\tau) & x(t_2 + 2\tau) & \cdots x(t_i + 2\tau) & \cdots x(t_n - (d - 3)\tau) \\
 \cdots & \cdots & \cdots & \cdots \\
 x(t_1 + (d - 1)\tau) & x(t_2 + (d - 1)\tau) & \cdots x(t_i + (d - 1)\tau) & \cdots x(t_n) \\
 Y(t_1) & Y(t_2) & \cdots Y(t_i) & \cdots Y(t_n)
 \end{array} \tag{1}$$

(1) 式的各列表示了 d 维空间的各相点 $Y(t_i)$ (见 (2) 式) 的 d 个分量值，分别称为第 1, 2, \cdots 及第 d 个分量。式中的 τ 为延滞时间，它是 $\{x(t_n)\}$ 序列的间隔 Δt 的整数倍， d 为整数。通常 $d = 3, 4, 5, 6, \cdots$ 。 τ, d 的取值可在模式用于不同参数预报试算中确定。相点数 $m = n - (d - 1)\tau$ 。

做预报，即是想知道下一时刻的 $x(t_n + 1)$ 值，由 (1), (2) 可知， $Y(t_m + 1)$ 的 d 个分量中只有 $x(t_n + 1)$ 是未知的，它正是我们待求的量。

1.2 相空间线性回归模式的建立

某一参考态 $Y(t_m)$ 附近有许多相轨道，找出 $Y(t_m)$ 的几个最近邻态 $Y_{nbi}(t_i)$ 并考查其所处相轨道的统计行为 (关系)，利用相似原理将其移植于参考态 $Y(t_m)$ 所处相轨道，以求得下一时刻的参量

$$b = \frac{k \sum_k \| Y_{nbi}(t_i)_k \| \| Y_{nbi}(t_i + T)_k \| - \sum_k \| Y_{nbi}(t_i) \| \sum_k Y_{nbi}(t_i + T)_k}{k \sum_k \| Y_{nbi}(t_i)_k \|^2 - (\sum_k \| Y_{nbi}(t_i)_k \|^2)} \tag{6}$$

从 k 个 $Y_{nbi}(t_i)$ 及 k 个 $Y_{nbi}(t_i + T)$ 的有关资料里 [它们完全由 (1)、(2) 确定]，求得系数 a, b 。将近邻相轨道的相似行为 (4) 移植于参考态 $Y(t_m)$ 所处的相轨道，则预报方程为

$$\| Y(t_m + T) \| = a + b \| Y(t_m) \| \tag{7}$$

在相空间里，我们可以逐一来考查不同相点在不同时刻的演化情景，使预报模式的函数形式或系数可因时而变动。(7) 式给出的线性形式在各相点虽不变，但系数 a, b 却随相点移动而变化，这就避免了传统统计模式的缺陷。

统的信息，Packard (1980) 提出了用时间序列重构吸引子的相空间图象。

Takens (1981) 提出嵌入定理，表明只要相空间的维数 d 足够大 (通常为 $2D_0 + 1$)，它就可以刻划 D_0 维混沌吸引子。

把时间序列 $\{x(t_n)\}$ 延拓成 d 维相空间的一个相型分布：

$x(t_n + 1)$ 。具体过程如下：

首先找出参考 $Y(t_m)$ 的最近邻态 $Y_{nbi}(t_i)$

$$Y_{nbi}(t_i) = \min \| Y(t_m) - Y(t_j) \| \tag{3}$$

$$j = 1, 2, \Lambda(m - 1)$$

式中 \min 表示最小， $\| \|$ 表欧氏模。由 (3) 可以找到最小、次小 $\cdots k$ 个近邻点 $Y_{nbi}(t_i)$ 。然后考查这 k 个相轨道的统计行为。 $Y_{nbi}(t_i)$ 经时间 T ($T = \Delta t, 2\Delta t, \Lambda$) 后，在其轨道上演变成 $Y_{nbi}(t_i + T)$ 。在 T 不太大的情况下 (即相轨道演变极短)，假设 $Y_{nbi}(t_i + T)$ 和 $Y_{nbi}(t_i)$ 服从以下的线性关系：

$$\| Y_{nbi}(t_i + T)_k \| = a + b \| Y_{nbi}(t_i)_k \| \tag{4}$$

用最小二乘法可以得到

$$a = \frac{\sum_k \| Y_{nbi}(t_i + T)_k \| - b \sum_k \| Y_{nbi}(t_i)_k \|}{k} \tag{5}$$

(7) 式右端之 a, b 由 (5)、(6) 给出， $Y(t_m)$ 由 (1)、(2) 给定，(7) 式左端的 $Y(t_m + T)$ 中只有 $x(t_m + T)$ 分量是未知的，由 (7) 式可求出 $x(t_m + T)$ 。 T 是序列 $\{x(t_n)\}$ 的时间间隔 Δt 的整数倍，它可随预报要求而定。

2 方法的应用

研究中使用的褐飞虱发生资料，取自种植单季晚稻为主的江苏省太湖地区农科所 (120.30° E,

31.50° N) 1986~1998 年 13 年间 7~9 月田间观察圃系统调查资料, 序列长度为 381。在进行混沌时间序列相空间线性回归预测模型的建立时, 我们采取了以年为间隔抽取、以旬为单位等间隔数据点的

方法。以积累的 7 月下旬的数据可预测下一年 7 月下旬的褐飞虱发生量, 同样以积累的 8 月上旬的数据可预测下一年 8 月上旬的褐飞虱发生量, 依次类推可以得到全年 7~9 月各旬的褐飞虱发生量预测。

表 1 1998 年 7~9 月各旬褐飞虱发生量的预测结果检验

Table 1 Test of the predictive accuracy of the model on *N. lugens* (BPH) occurrence between July and September, 1998

月.日 Month. Day	实际值 Observed value	预测值 Forecasted value	相对误差绝对值 (%) Absolute value of relative error	模拟发生程度* Simulated occurrence degree	实际发生程度 Observed occurrence degree	检验 Test
7.30	36	39.6735	10.20	轻发生 light occurrence	轻发生 light occurrence	✓
8.10	140	109.1692	22.02	轻发生 light occurrence	轻发生 light occurrence	✓
8.20	266	162.5923	38.88	轻发生 light occurrence	轻发生 light occurrence	✓
8.30	250	173.460	30.62	轻发生 light occurrence	轻发生 light occurrence	✓
9.10	1 630	1 860	14.11	中发生 medium occurrence	中发生 medium occurrence	✓
9.20	4 460	5 512	23.59	大发生 severe occurrence	大发生 severe occurrence	✓
9.30	5 340	6 458	20.94	大发生 severe occurrence	大发生 severe occurrence	✓
平均误差 average error			22.91			

* 根据当地褐飞虱发生程度等级划分标准: 百穴虫量 < 500 头, 轻发生; 百穴虫量在 500~1 000 头, 偏轻发生; 百穴虫量在 1 000~2 000 头, 中等发生; 百穴虫量在 2 000~3 000 头, 偏重发生; 百穴虫量 > 3 000 头, 大发生

* Occurrence degree of BPH was based on local classification standards: 1) BPH number/100 rice plants (B/100R) < 500, light occurrence; 2) B/100R = 500 - 1 000, a little more than light occurrence; 3) B/100R = 1 000 - 2 000, medium occurrence; 4) B/100R = 2 000 - 3 000, a little more than medium occurrence; 5) B/100R > 3 000, severe occurrence

我们构造了 10 维相空间, 通过计算各点与 $Y(t_m)$ 之间的欧氏距离找出 $Y(t_m)$ 的最近邻域点 k 为 7 个。根据公式 (4)、(5)、(6)、(7) 得到 1998 年 7~9 月各旬褐飞虱的发生量, 结果见表 1。从表 1 给出的 1998 年 7~9 月各旬褐飞虱发生量的实际情况与预报结果的对比中不难看出, 预报结果是令人满意的, 7 次的预报平均相对误差为 22.91%, 转化成的发生程度与实际发生程度相符, 预报准确率达 100%。

3 结论与讨论

(1) 本文作者基于混沌动力系统的相空间延迟坐标重构, 利用混沌序列固有的确定性和非线性, 提出了用于混沌时间序列预测的相空间线性回归预测模型, 通过对 1998 年 7~9 月各旬的褐飞虱发生量实测数据资料进行预报, 得到了 7 次预报的平均相对误差仅为 22.91%, 转化为发生程度与实际发生程度相符, 预报准确率达 100% 的预报结果。这就为害虫时间序列的可预测性及新型预测模型的构建提供了可靠的依据。

(2) 相空间线性回归预测方法要求数据愈多愈

好, 本试验仅用了 1986~1998 年的历史资料, 数据总数不足 400 个, 相点数 $m = n - (d - 1)\tau$ 也在 400 以内。故相轨道不是很稠密, 影响近邻点的选择以及 a 、 b 的计算, 最终影响预报的精度。如果能够获得更多的资料, 用于实时预报, 精度将会进一步提高, 误差大于 22.91% 的份额也将会降低。

(3) 混沌时间序列预测方法是一新的很有发展前途的预测方法。本文作者提出的相空间线性回归预测模型, 具有其独特的优点, 主要表现在能够有效地减小测报误差。在实际应用过程中, 随着对害虫发生系统内部运行机制的进一步揭示, 长期预报——这一长期困扰害虫测报的难题将会得到较好的解决。

参 考 文 献 (References)

- Cheng Q S, Wu L W, Wang S Z, 2000. Combination prediction based on attributive cluster net and radial basic function. *Chinese Science Bulletin*, 45 (11): 1 211 - 1 216. [程乾生, 武连文, 王守章, 2000. 基于属性聚类网络和径向基函数的融合预测. 科学通报, 45 (11): 1 211 - 1 216]
- Hu B H, Jiang R Z, 1997. Long-term Occurrence Rules and Forecasting of Crops' Diseases and Pest Insects. Beijing: China Agricultural Press. 7 - 12. [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. 7 - 12]

- Ma F, Ding Z Z, Cheng X N, 2000. Chaos and predictable time-scale of brown planthopper *Nilaparvata lugens* (Stål) occurrence system. Inter-country Forecasting System and Management for Brown Planthopper in East Asia, Zhejiang University, Hangzhou, China. 3 - 14.
- Ma F, Li H B, Cheng X N, 2001. Study on fractal property of *Nilaparvata lugens* (BPH) occurrence. *Acta Ecol. Sin.*, 21 (2): 179 - 285. [马飞, 李红兵, 程遐年, 2001. 褐飞虱 (*Nilaparvata lugens*) 发生的分形性质研究. 生态学报, 21 (2): 179 - 285]
- Packard N H, Cratchfield J P, Farmer J D *et al.*, 1980. Geometry from a time series. *Phys. Rev. Letts.*, 45 (9): 712 - 716.
- Tekens F, 1999. Detecting strange attractors in turbulence. In: Rand D A, Young L S eds. *Dynamical Systems and Turbulence*. London: Oxford Press. 134 - 156.
- Wang M J, Cheng Q S, 1997. Application of Kohonen self-organization net in chaotic time series prediction. *Theo. Appl. System Engineering*, 2: 12 - 17. [王明进, 程乾生, 1997. Kohonen 自组织网络在混沌时间序列预测中的应用. 系统工程理论与实践, 2: 12 - 17]
- Zhang J S, Xiao X C, 2000. Predicting low dimensional chaotic time series using Volterra adaptive filters. *Acta Physica Sinica*, 49 (3): 403 - 407. [张家树, 肖先赐, 2000. 混沌时间序列的 Volterra 自适应预测. 物理学报, 49 (3): 403 - 407]