

小麦条锈病受灾率时间序列混沌特征研究^{*}

张 静^{1,2}, 洪新兰³

(1 襄樊学院 物理系, 湖北 襄樊 441000;

2 华中科技大学 控制科学与工程学院, 湖北 武汉 430074)

3 武汉中南民族大学 职业技术学院, 湖北 武汉 430074)

[摘 要] 利用混沌和分形理论, 对湖北省襄樊地区小麦条锈病受灾率进行了混沌识别研究, 对其功率谱、主分量、关联维度、最大Lyapunov指数进行了分析。结果表明, 小麦条锈病受灾率时间序列具有混沌特征, 属于混沌时间序列。因此建议对小麦条锈病进行建模预测时, 应主要采用非线性建模方法。

[关键词] 小麦条锈病; 受灾率时间序列; 混沌识别; 混沌特征

[中图分类号] S435.121.4⁺2; S11⁺9 **[文献标识码]** A **[文章编号]** 1671-9387(2005)09-0063-05

我国是农业大国, 小麦生产在农业中占有十分重要的战略地位, 但是作为小麦主要病虫害之一的小麦条锈病, 给小麦生长带来了严重危害, 导致小麦大幅减产、减收。小麦条锈病受灾率时间序列预报一直是相关科研工作者关注的重要课题, 对其研究具有重要的社会和经济意义, 但同时也是难度较大的研究领域。主要困难在于: 一方面生产实际中, 影响小麦条锈病受灾率的因素众多, 如与前一年的发病情况、大西北的外来菌、冬春季节西北风的强度和频度、春季气候温度和湿度等客观因素及农业部门的防治措施等, 可以认为小麦条锈病受灾率时间序列是许多变量组成的复杂的非线性系统; 另一方面, 时间序列数据实测每年均需大量的人力记录和汇总, 一些资料数据相对比较缺乏。

小麦条锈病受灾率时间序列预报的目的, 是根据其自身规律推断其未来。目前, 广泛采用数理统计回归分析方法进行定量分析预报, 只能在概率意义下作1种可能性预测, 其中大量模型未考虑非线性作用, 仅仅考虑线性因素。因此, 传统的基于统计学的方法, 难以获得满意的预测效果。

传统的时间序列预测方法都是基于一维空间进行的, 大量研究^[1~3]表明, 这种预测模型存在丢失实际演化过程信息的缺点。利用人工神经网络技术建立预测模型^[4], 是很好的方法, 但神经网络输入信号和输入节点数的确定, 只能基于经验或反复试验, 缺

乏有力的理论根据。

近年来, 混沌理论、分形理论的发展为研究这一问题提供了新的思路, 如果小麦条锈病受灾率时间序列具有混沌特征, 属于混沌时间序列^[2,3], 则可以通过相空间重构技术, 为确定系统性质及预测方法提供理论依据。本研究根据混沌理论、分形理论的方法, 尝试对湖北省襄樊地区小麦条锈病受灾率进行混沌识别研究, 以期小麦条锈病预测模型的建立提供理论依据。

1 原始数据

本研究以湖北省襄樊地区每年小麦条锈病发病面积为原始数据, 由于每年小麦播种面积不同, 因此受灾率定义为每年小麦条锈病发病面积占年小麦播种面积之百分比。取每年受灾率序列作为本研究讨论的时间序列, 具体见图1。

2 小麦条锈病混沌识别研究

2.1 小麦条锈病受灾率时间序列的功率谱

时间序列的功率谱是试验中可以直接观测的对象之一。时间序列的原始数据图像看上去并不规则, 但其功率谱却可能呈现规则性。一般, 周期性序列功率谱具有明显的周期特征, 频率为 f 的周期系统的功率谱在频率 f 及其高次谐波 $2f, 3f, \dots$ 处有 δ 函数形式尖峰; 基频 $f_1, f_2, \dots, f_k$ 准周期系统的功

^{*} [收稿日期] 2005-05-10

[基金项目] 湖北省教育厅科研项目(2001D69001)

[作者简介] 张 静(1967-), 女, 湖北襄樊人, 副教授, 主要从事建模预测研究, 现在华中科技大学控制与工程学院作访问学者。

E-mail: mindawang@163.com

率谱在 $f_1, f_2, \dots, f_k$ 及其线性组合处有 δ 函数形式尖峰, 非周期混沌系统的功率谱可能有尖峰, 但多少会增宽一些, 且功率谱上会出现宽带的噪声背景, 某些混沌系统其功率谱超过一定值以后, 谱线随频率指数衰减, 故可定性反映序列的混沌特征。

小麦条锈病受灾率时间序列功率谱如图 2 所

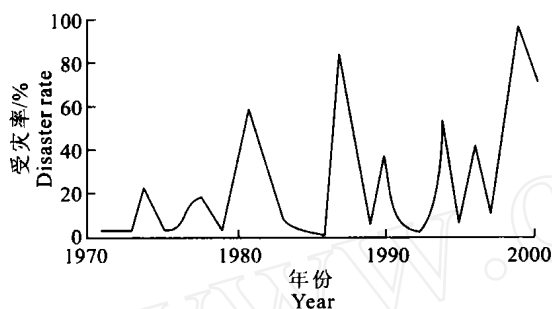


图 1 小麦条锈病受灾率时间序列原始数据

Fig. 1 Disaster rate of wheat stripe rust

示。从图 2 可以看出, 所研究的受灾率时间序列功率谱无明显的周期性, 从形状上看接近混沌序列, 初步分析受灾率时间序列可能具有一定的混沌特征。当然, 由于实测数据有限, 无法确定其是否是混沌序列, 或长周期序列, 因此需用其他方法定量分析计算, 以确定其是否具有混沌特征。

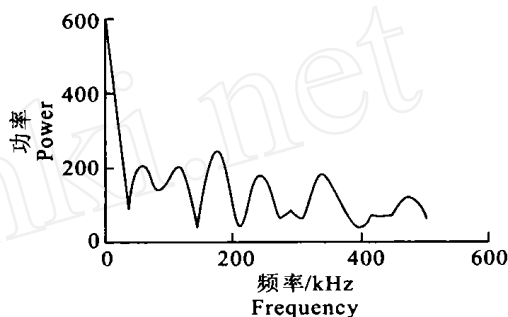


图 2 小麦条锈病受灾率时间序列功率谱

Fig. 2 Power spectrum for disaster rate of wheat stripe rust

2.2 小麦条锈病受灾率时间序列的关联维数

2.2.1 相空间重构 时间序列是许多物理因子相互作用的综合反映, 蕴藏着参与运动的全部变量的痕迹。可能包含混沌运动的信息, 而混沌运动至少要在三维自治动力系统中才能出现。因此, 要将时间序列扩展到三维或更高的相空间中, 才能将时间序列的混沌信息充分显现出来, 这就是时间序列的重建相空间。

Packard 等^[5]和 Takens^[6]提出, 证明了可以找到 1 个合适的嵌入空间维数 $m \geq 2D + 1$ (D 为系统的关联维数), 在这个嵌入空间里可把有规律的轨迹(吸引子)恢复出来。即从 1 个时间间隔一定的一维时间序列 $\{x_1, x_2, \dots, x_N\}$ (设总点数为 N) 出发, 构造一批 m 维 ($m \geq 2D + 1$) 的矢量 $Y(t_i)$, $Y(t_i)$ 的表示式为:

$$Y(t_i) = [x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(m-1)\tau}] \quad (1)$$

式中, m 为嵌入维数; τ 为时间延迟; $Y(t_i)$ 为 m 维相空间的相点; $i = 1, 2, \dots, M$, M 为相点数, 且 $M = N - (m - 1)\tau$

$\{Y(t_i) | i = 1, 2, \dots, M\}$ 支起一个嵌入空间, 在拓朴等价的意义下恢复原动力学性态。

相空间重构是分维数和 Lyapunov 指数计算的关键。重构相空间的关键在于嵌入空间维数 m 和时间延迟 τ 的选择。

2.2.2 小麦条锈病受灾率时间延迟的确定 为了从一维时间序列构造出相空间, 必须给出合适的采

样间隔, 即确定时间延迟 τ 。确定时间延迟 τ 的方法很多, 自相关函数法是非常成熟的方法, 其主要思想是提取序列间的线性相关性, 其具体步骤是求 $R_{xx}(j\tau)$ 。对一维时间序列 $\{x_1, x_2, \dots, x_N\}$ (设总点数为 N), 时间跨度为 $j\tau$ 的序列 $\{x_i\}$ 的自相关函数为 $R_{xx}(j\tau)$ 。 $R_{xx}(j\tau)$ 计算公式为:

$$R_{xx}(j\tau) = \frac{1}{N} \sum_{i=1}^{N-j\tau} x_i x_{i+j\tau} \quad (2)$$

式中, τ 为时间延迟; $j\tau$ 为时间跨度; N 为时间序列总点数; x_i 为 i 时刻的时间序列值; $x_{i+j\tau}$ 为 $i + j\tau$ 时刻的时间序列值。固定 (2) 式中 j , 取 $\tau = 1, 2, \dots$, 求自相关函数 $R_{xx}(j\tau)$, 当 (2) 式中自相关函数下降到初始值时的 $1 - \frac{1}{e}$, 所得的时间即是重构相空间的最佳时间延迟 τ 。本研究使用该法求得 $\tau = 2$ 。

2.2.3 小麦条锈病受灾率时间序列关联维数 D 的确定 混沌体系由奇怪吸引子的不规则轨线来描述, 奇怪吸引子为分形结构, 而分形结构的特点是分数维。关联维数 D 可对吸引子的几何特征及基于吸引子的轨道随时间的演化情况进行数量上的描述, 是序列混沌特征的重要不变量。Grassberger 和 Procaccia^[7]提供了 1 种易于从一维时间序列试验数据中提取关联维数 D 的算法, 即 G-P 算法。其思想是: 对重构相空间中的矢量 $Y(t_i)$, 定义矢量间距离为 $\|Y(t_i) - Y(t_j)\|$, 规定凡是距离小于给定正数 r 的矢量均称为有关联矢量。重构相空间中有关联矢量对的数量与一切可能的 M^2 种配对之比为关联积

分 $C(r)$ 。 $C(r)$ 的计算公式为:

$$C(r) = \frac{1}{M^2} \sum_{i,j=1}^M \theta(r - \|Y(t_i) - Y(t_j)\|)$$

$$\text{其中 } \theta(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0 \end{cases} \quad (3)$$

式中, $Y(t_i)$ 和 $Y(t_j)$ 为 m 维相空间的相点; $i, j = 1, 2, \dots, M$; M 为相点个数; r 为给定的正数。则有 $\lim_{r \rightarrow 0} C(r) \propto r^D$, 其中 D 即为关联维数。适当地选择给定的正数 r , 近似的有:

$$D = \frac{\ln(C(r))}{\ln r} \quad (4)$$

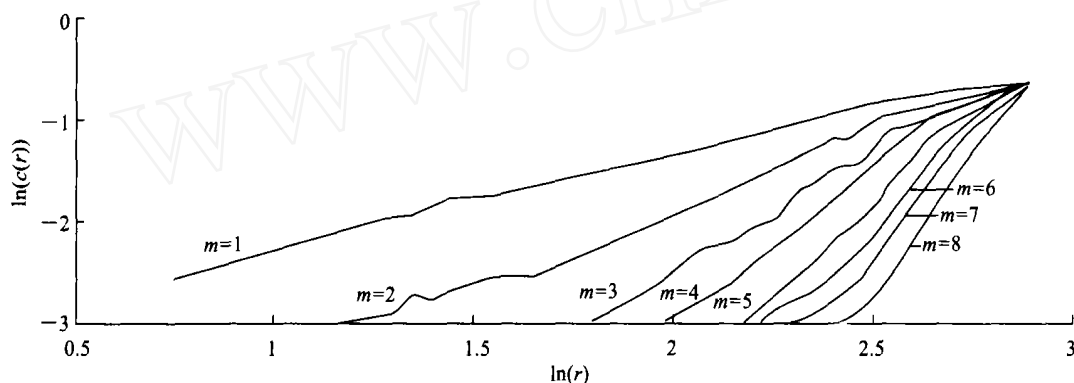


图3 小麦条锈病受灾率时间序列 $\ln(C(r)) - \ln r$ 曲线

Fig 3 Curve of $\ln(C(r))$ vs $\ln r$ for disaster rate of wheat stripe rust

图3是小麦条锈病受灾率时间序列 $\ln(C(r)) - \ln r$ 曲线。由图3的 $\ln(C(r)) - \ln r$ 曲线束可以看出, 当 $m = 1 \sim 5$ 时, 随着 m 的增大, $\ln(C(r)) - \ln r$ 曲线的斜率随之增大; 当 $m = 6, 7$ 和 8 时, $\ln(C(r)) - \ln r$ 曲线的斜率逐渐趋于稳定, 3条直线趋于平行。因此 $m = 6, 7$ 和 8 时, $\ln(C(r)) - \ln r$ 直线的斜率就是关联维数 D 。

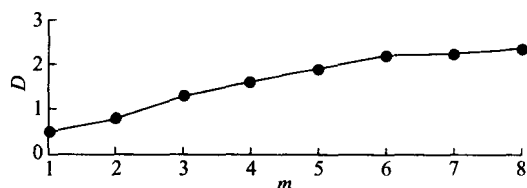


图4 小麦条锈病受灾率时间序列 $D - m$ 曲线

Fig 4 Correlation of D vs m for disaster rate of wheat stripe rust

从图4可以看出, $m = 6, 7$ 和 8 时, 关联维数 D 趋于一个稳定值, 即 $D = 2.2$ 。该关联维数 D 的存在, 表明小麦条锈病受灾率时间序列具有混沌特征。嵌入空间维数按 $m \geq 2D + 1$ 的原则选取, 则 $m = 6$ 。

但Wolf等^[8]认为, 大多数情况下 m 的取值, 可

在(1), (3)和(4)式计算中, 令嵌入空间维数 m 从小到大, 求出 $\ln(C(r)) - \ln r$ 曲线。若 $\ln(C(r)) - \ln r$ 曲线的斜率随 m 增大而增加, 当 m 足够大时, $\ln(C(r)) - \ln r$ 曲线的斜率将趋于一个稳定值, 相应的 $\ln(C(r)) - \ln r$ 曲线束趋于平行, 则说明存在关联维数。同时, 求出每个 m 对应的 $D = \ln(C(r)) / \ln r$, $\ln(C(r)) / \ln r$ 的稳定值即是关联维数 D 。

使用G-P算法计算小麦条锈病受灾率时间序列, 其 $\ln(C(r)) - \ln r$ 曲线 $D - m$ 曲线如图3, 4所示。

较 $m \geq 2D + 1$ 确定的值小一些, 在计算Lyapunov指数时, 只计及沿轨道的平均性质, 这时取 $D < m < 2D$ 即可, 因此下面论述中, 在计算Lyapunov指数时取 $m = 3$ 。

2.3 小麦条锈病受灾率时间序列的主分量分析

主分量(Principle component analysis, PCA)分析法是一种能有效识别混沌和噪声的方法^[9]。其具体方法如下: 对时间序列 $\{x_1, x_2, \dots, x_N\}$ (设总点数为 N) 重构相空间, 取嵌入维数 m , 则由该时间序列可形成轨线矩阵 $X_{M \times m}$ 。 $X_{M \times m}$ 的计算公式如下:

$$X_{M \times m} = \frac{1}{M^2} \begin{bmatrix} x_1 & x_2 & \cdots & x_m \\ x_2 & x_3 & \cdots & x_{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_M & x_{M+1} & \cdots & x_N \end{bmatrix} \quad (5)$$

式中, N 为时间序列的总点数; m 为嵌入维数; $M = N - (m - 1)$ 。 $X_{M \times m}$ 的协方差矩阵为:

$$A_{m \times m} = \frac{1}{M} X_{M \times m}' X_{M \times m} \quad (6)$$

式中, $X_{M \times m}'$ 为 $X_{M \times m}$ 的转置矩阵。

$A_{m \times m}$ 的特征值和相应的特征向量为 λ_i 和 U_i ($i = 1, 2, \dots, m$)。将特征值按从大到小排序即 $\lambda_1 \geq \lambda_2 \geq \dots$

$\dots \geq \lambda_n$, 以特征值和相应的特征向量为主分量。

$$r = \sum_{i=1}^m \lambda_i \quad (7)$$

式中, λ_i 为 $A_m \times m$ 的第 i 个特征值; $i = 1, 2, \dots, m$ 。以指标 i 为横坐标, $\ln(\lambda_i/r)$ 为纵坐标作图, 称 $\ln(\lambda_i/r) - i$ 图为主分量谱, 该谱图反映了信号能量

在对应主轴上的大小分布情况。混沌信号的主分量谱图应是 1 条过定点且斜率为负的直线。Duffing 方程的主分量谱图见图 5-a。根据 $m \geq 2D + 1$, 取 $m = 5, 6$ 和 7 , 分别由公式 (5), (6) 和 (7) 求出 $\ln(\lambda_i/r)$ ($i = 1, 2, \dots, m$), 得到小麦条锈病受灾率时间序列主分量谱图, 如图 5-b, c, d 所示。

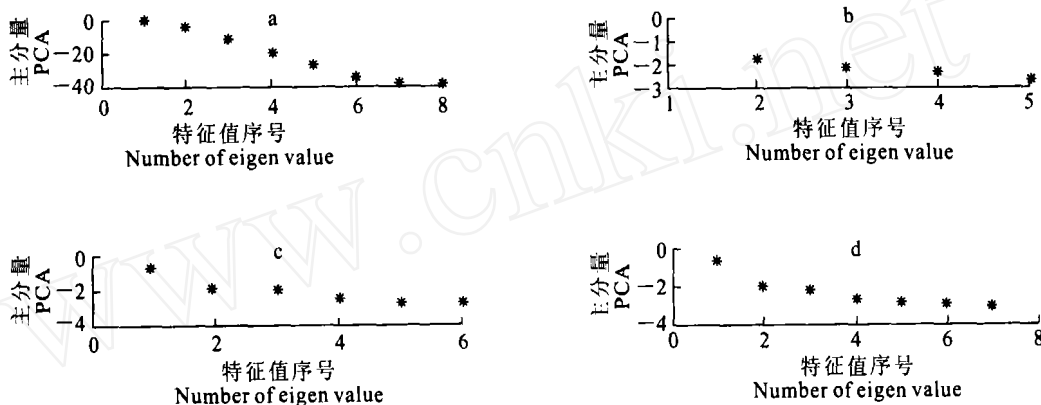


图 5 Duffing 方程和小麦条锈病受灾率时间序列的主分量谱图

a Duffing 方程; b $m = 5$; c $m = 6$; d $m = 7$

Fig. 5 PCA chart for Duffing and disaster rate of wheat stripe rust in different embedding dimension

a Duffing; b $m = 5$; c $m = 6$; d $m = 7$

从图 5 可以看出, 当 $m = 5, 6$ 和 7 时, 考虑误差因素, 则小麦条锈病受灾率时间序列主分量谱图中各点构成的直线, 基本符合混沌信号的主分量谱图的条件, 即为 1 条过定点且斜率为负的直线。再将其与 Duffing 方程的主分量谱图比较, 可以认定小麦条锈病受灾率时间序列具有混沌特征。

2.4 Lyapunov 指数

目前, 在表征混沌运动方面, Lyapunov 指数是具有重大意义的统计量, 是相空间中相近轨道的平均收敛性或平均发散性的一种度量。在 Lyapunov 指数 $\lambda < 0$ 的方向, 相体积收缩, 运动稳定, 对初始条件不敏感; 在 $\lambda > 0$ 方向, 轨道迅速分离, 长时间行为对初始条件敏感, 运动为混沌状态; 在 $\lambda = 0$ 方向, 边界稳定, 属临界状态。若系统最大的 Lyapunov 指数 $\lambda_1 > 0$, 且系统有界, 则系统一定是混沌状态。

Wolf 方法^[8]是一种能从试验数据中计算非负最大的 Lyapunov 指数的算法。其计算方法如下: 从一维时间序列 $\{x_1, x_2, \dots, x_N\}$ (设总点数为 N) 出发, 重构相空间矢量 $Y(t_i)$, 取初始点 $Y(t_0)$, 其与最近邻点 $Y_0(t_0)$ 的距离为 L_0 , 追踪这两点时间的演化, 直到 t_i 时刻, $Y(t_0)$ 演化为 $Y(t_i)$, $Y_0(t_0)$ 演化为 $Y_0(t_i)$, 其间距超过某规定值 $\epsilon > 0$, $L_0 = |Y(t_0) - Y_0(t_0)| > \epsilon$,

保留 $Y(t_i)$, 并在 $Y(t_i)$ 邻近找 $Y_1(t_i)$, 使得 $L_1 = |Y(t_i) - Y_1(t_i)| < \epsilon$ 并且与之夹角尽可能的小。继续该过程, 直至 $Y(t)$ 到达时间序列的终点 N , 这时演化过程的总迭代次数为 M , 则最大的 Lyapunov 指数为 λ_1 , 其计算公式为:

$$\lambda_1 = \frac{1}{t_M - t_0} \sum_{i=1}^M \ln \frac{L_i}{L_0} \quad (8)$$

式中, t_0 为初始时刻; t_M 为演化过程的总迭代次数对应的时刻; M 为演化过程的总迭代次数; L_i 为 t_i 时刻点 $Y(t_i)$ 与其最近邻点 $Y_i(t_i)$ 的距离; L_0 为 t_0 时刻到 t_{i+1} 时刻, $Y(t_0)$ 的演化点 $Y(t_{i+1})$ 与 $Y_i(t_0)$ 的演化点 $Y_i(t_{i+1})$ 之间的距离。利用 Wolf 方法计算受灾率时间序列最大的 Lyapunov 指数 λ_1 , 可知 $\lambda_1 = 1.66$ 。系统为实际受灾率时间序列, 则系统有界; 且 $\lambda_1 > 0$, 因此小麦条锈病受灾率时间序列是混沌状态。

3 结 论

本研究根据混沌理论、分形理论的方法, 对湖北省襄樊地区小麦条锈病受灾率进行了混沌识别研究, 分析其不同的特征指数, 不同的方法结果均显示其具有混沌特征, 属于混沌时间序列。因此, 在对小麦条锈病建模预测时, 应采用非线性建模方法。

根据混沌学理论, 最大 Lyapunov 指数 λ_1 的倒

数 T_m 表示混沌系统确定性预测的时间上界, 即最长预报时间。但在本研究中, 由于客观因素, 如最大 Lyapunov 指数计算过程中客观实测数据不够长, 混沌特征值提取过程中某些参数的确定方法(如时间延迟 τ 的选择)不够成熟, 以及人为因素(如 Wolf 法距离和夹角的选取)等, 无法确认小麦条锈病的最长预报时间 T_m ($T_m < 1$ 年), 这实际意味着小麦条锈病无法预测。

虽然目前国内尚无与本研究的类似报道, 且本研究所得的结论对小麦条锈病防灾减灾工作也应该

有一定意义, 但本研究还存在以下问题:

(1) 每年实测时间序列数据需要大量人力记录和汇总, 实测数据不够长;

(2) 小麦条锈病发病地区涉及多省, 由于外来菌源会随冬春季西北风向东南方向传播, 因此所有发病地区属于同一个大系统, 而且从菌源看, 从西北到东南有“上下游”关系, 如果在大系统内讨论特征和建模预测, 则信息会更充分全面, 结论也会更加准确。

[参考文献]

- [1] 安鸿志, 陈 敏 非线性时间序列分析[M]. 上海: 上海科学技术出版社, 1998
- [2] 吕金虎, 陆君安, 陈士华 混沌时间序列及其应用[M]. 湖北武汉: 武汉大学出版社, 2002
- [3] 王心元 复杂非线性系统中的混沌[M]. 北京: 电子工业出版社, 2003
- [4] 张 静 基于神经网络的受灾面积预测精度研究[J]. 襄樊学院学报, 2004, 25(2): 24- 26
- [5] Packard N H, Crutchfield J P, Farmer J D, et al Geometry from a time series[J]. Phys Rev Lett, 1980, 45: 712- 716
- [6] Takens F. Detecting strange attractor in turbulence[J]. Lect Notes in Math, 1981, 898: 366- 381.
- [7] Grassberger P, Procaccia I Measuring the strangeness of strange attractors[J]. Physica D, 1983, 9(1- 2): 189- 208
- [8] Wolf A, Swift J B, Swinney H L, et al Determining lyapunov exponents from a time series[J]. Physica D, 1985, 16(3): 285- 317.
- [9] 龚云帆, 徐建学 混沌信号与噪声[J]. 信号处理, 1997, 13(2): 112- 118

Chaotic analysis of chaotic features of wheat stripe rust disaster rate of crops time sequence

ZHANG Jing^{1, 2}, HONG Xin-lan³

(1 Department of Physics, Xiangfan University, Xiangfan, Hubei 441000, China;

2 Department of Control Science and Engineering, Huazhong University of Science & Technology, Wuhan, Hubei 430074, China;

3 College of Vocational Technology of South-Central University of Nationalities, Wuhan, Hubei 430074, China)

Abstract: Forecast the disaster rate of crops is important, unfortunately, it is also difficult. Proving whether disaster rates are chaotic is useful for modeling. Based on the chaotic theory, there are several methods developed for the chaos identification at present. In this paper, power spectrum and PCA of the disaster rates are analyzed; correlation dimension and Lyapunov exponent are computed by the reconstruction of phase space, G-P algorithm and method of determining Lyapunov exponent from time series for identifying chaotic character of the time sequence, and the time sequence is proved to be chaotic indeed. So methods of modeling nonlinear systems can be used in modeling and forecasting the disaster rate of crops time sequence.

Key words: wheat stripe rust; the disaster rate of crops time sequence; chaos identification; chaotic character