

全国高等农业院校教学参考书

全国高等农业院校
教材指导委员会审定
植物保护专业用



作物病虫害 数理统计预报

程极益 编著

农业出版社

全国高等农业院校教学参考书

作物病虫害数理统计预报

程极益 编著

植物保护专业用

农 业 出 版 社

422769

S435

128A

(京)新登字060号

全国高等农业院校教学参考书
作物病虫害数理统计预报

程极益 编著

* * *

责任编辑 张洪光

农业出版社出版 (北京市朝阳区农展馆北路2号)
新华书店北京发行所发行 农业出版社印刷厂印刷

787×1092mm32开本 12.25 印张 245 千字

1992年10月第1版 1992年10月北京第1次印刷

印数 1—4,050 册 定价 4.50 元

ISBN 7-109-02121-1/S·1395

编 审 者 名 单

编者 程极益 (南京农业大学 副教授)
审稿 汪世泽 (西北农业大学 教授)

内 容 简 介

本书介绍了国内外最新的作物病虫害数理统计预报方法。内容包括聚类分析、判别分析、相关分析、回归分析、时间序列分析等。既有理论，又有很强的实用性。每种方法均有手工计算实例与习题，并为读者提供了 BASIC 电子计算机程序。可以做为植物保护、畜牧兽医、卫生等专业本科生的教材，也可供农林部门广大植物保护科技工作者学习使用。

前 言

《作物病虫害数理统计预报》原为自编讲义，自1983年开始，做为南京农业大学植物保护、农业昆虫、植物病理专业本科生教材使用。编者并曾多次在农业部的“全国病虫害测报进修班”、农垦系统的“病虫害测报培训班”与林业部的“森林害虫预测预报训练班”，使用该讲义讲授作物病虫害数理统计预测预报方法。这次编写在原有讲义的基础上，充实了理论部分，吸收了国内外的最新科研成果，给各种方法配置了BASIC电子计算机程序与应用计算实例。

全书共分八章。一至五章为多元分析方法，包括预报因子的选择，简易预报方法，聚类分析，判别分析，逐步判别分析，逐步回归分析等。六、七两章为随机时间序列分析方法。第八章为预报质量的评定。考虑到目前市级以下植保部门推广使用的需要，书中所提供的BASIC电子计算机程序，多数是按“PC-1500”型电子计算机的要求编写的，但只要做一点小的改动，也可适用于其它型号的电子计算机。书中所有的练习题，可使用小型计算器进行手工计算。

本书可以做植物保护专业本科生的教材，也可供农林部门广大植物保护科技工作者，畜牧兽医、卫生等专业的本科生学习使用。读者应具备“高等数学”、“生物统计”、

“概率论”等方面的基础知识。

本书在张孝羲教授指导下完成。书稿由西北农业大学汪世泽教授审阅并提出了许多宝贵意见。在编辑出版过程中，曾得到南京农业大学校、系、教研室各级领导，许多学者、同行的支持与帮助，收到过许多听过编者讲课的同志们的来信，表达了他们对本书的热切希望，在此一并表示衷心的感谢。

由于编者学识、经验有限，书中难免出现错误和缺点，恳请读者提出批评指正。

编 者

1989年12月于南京农业大学

目 录

第一章 选择预报因子的方法	1
第一节 普查法	3
一、方法原理和步骤	3
二、几点说明	10
三、程序	16
练习一 用普查法选择预报因子	20
第二节 相关系数法	21
一、相关系数的意义	21
二、相关系数的计算	23
三、相关系数的统计测验	28
四、程序	32
练习二 线性相关系数	35
第二章 简易预报方法	36
第一节 点聚图法	38
练习三 点聚图	41
第二节 复相关表法	41
练习四 复相关表	44
第三节 多因子综合相关法	45
一、计算海明距离	45
二、相似基数法	47
三、线性综合法	49
第三章 聚类分析	53

第一节 聚类分析概述	53
第二节 一次聚合法	58
一、原始资料标准化	58
二、计算各样本之间的欧氏距离	61
三、绘制谱系图	63
四、预报	65
五、程序	67
第三节 模糊聚类分析	71
练习五 一次聚合法	75
第四章 回归分析	76
第一节 相关的基本概念	76
一、相关关系	76
二、直线回归方程	81
练习六 直线回归方程	90
三、可变换成直线的曲线回归方程	93
第二节 多元回归分析	102
一、计算准备工作	103
二、建立多元回归方程	108
三、求解回归系数	109
四、总体回归效果的显著性测验	127
五、各回归系数 b_i 的显著性测验	129
六、多元回归分析讨论	142
七、附录	150
练习七 多元回归分析	155
第三节 逐步回归分析	156
一、逐步回归的基本思想	156
二、逐步回归的计算步骤	157
三、讨论	179
四、程序	181
练习八 逐步回归分析	191

第五章 判别分析	192
第一节 Fisher准则下二类判别分析的原理	192
一、基本思想	192
二、Fisher准则	199
第二节 二类判别分析实例与程序	202
一、预报因子资料的分类	202
二、计算 a_i 、 L_{ij}	202
三、求解判别系数 c_i	205
四、建立判别函数	206
五、预报决策与历史样本的检验	206
六、两类判别效果的统计测验	209
七、讨论	215
八、程序	226
练习九 Fisher准则下的二类判别分析	235
第三节 逐步判别分析与程序	236
一、Bayes准则下的多组判别分析	236
二、判别效果的显著性测验	245
三、各因子判别能力的显著性测验	248
四、逐步判别分析	251
五、逐步判别程序	261
第六章 时间序列分析	277
第一节 历史曲线演变法	278
第二节 方差分析周期外推法	282
一、方差分析周期外推法概述	282
二、方差分析周期外推法实例	286
三、讨论	293
练习十 方差分析周期外推法	294
四、程序	295
第三节 平稳随机时间序列预报方法	299
一、AR(P)自回归模型	299

二、选点法与程序	302
练习十一 选点法	314
第七章 马尔科夫链预报方法	315
一、马尔科夫链的概念	315
二、一阶转移概率	316
三、高阶转移概率	317
四、以转移概率为基础的预报方法	322
练习十二 马尔科夫链预报方法	325
五、程序	326
第八章 预报质量的评定	329
第一节 分类预报的预报质量评定方法	330
一、预报符合率	330
二、Hedike的预报技术评分	330
三、两分类预报的定性相关系数	335
第二节 定量预报的预报质量评定方法	337
一、平均绝对误差与标准差	337
二、离均差值和离均差符号的联合评分	339
附录 练习答案	342
练习一 用“普查法”选择预报因子	342
练习二 线性相关系数	342
练习三 点聚图	342
练习四 复相关表	343
练习五 一次聚合法	344
练习六 直线回归方程	345
练习七 多元回归分析	345
练习八 逐步回归分析	350
练习九 Fisher 准则下的两类判别分析	354
练习十 方差分析周期外推法	356
练习十一 选点法	360

练习十二 马尔科夫链预报方法	361
附表一 正态分布双侧置信限表.....	363
附表二 χ^2 分布的置信限表	364
附表三 t 分布的双侧置信限表	366
附表四 F 分布的置信限表	368
参考文献	378

第一章 选择预报因子的方法

数理统计是数学的一个分支，借助统计学和数学中的方法研究随机事件的客观规律，可以作为分析病虫害资料 and 使病虫害预测预报客观化的工具。分析病虫害变化的统计规律性以及预报因子和预报对象的数量关系，建立数学模型与图表，对未来的病虫害作出预测预报，这就是病虫害数理统计预测预报，简称统计测报。

病虫害的预测预报工作是进行病虫害综合防治的必要前提。只有对病虫害发生为害的预测预报做得及时、准确，才能正确地拟定综合防治计划，及时采取必要的措施，经济有效地压低病虫害的发生，保证作物的高产、稳产。

病虫害的预测预报内容，通常包括：①发生期预测预报，预测病虫害的发生时期或为害时期；具有迁飞习性的害虫，预测其迁出或迁入本地田间的时期。②发生量预测预报。③分布区预测预报，包括迁移、扩散的预测预报。④为害程度和损失的预测预报。

预报的种类有短期预报：预报期限10天以内或害虫一个世代以内；中期预报：预报期限10天以上或害虫一个世代以上的季节性预报；长期预报：预报期限为30天以上或1年以上的预报；多年预测：预报期限为5年以上的预报。

本书所介绍的各种数理统计预报方法适用于上述各项病虫害的预测预报,适用于长期、中期、短期与多年的预测预报。

通常我们把对未来的病虫害有指示性的现象,称为预报因子,简称因子。预报因子用英文字母 x 来表示。把需要预报的某种病虫害称为预报对象。预报对象用英文字母 y 来表示。

选择良好的预报因子是病虫害数理统计预报成败的关键。经验证明,同一种预报方法,采用不同的预报因子,其预报效果往往存在很大的差别。采用良好的预报因子其预报效果是良好的,采用低劣的预报因子其预报效果则肯定是低劣的。

必须注意要从与病虫害的发生、变化有关的因素中选择预报因子,而不能随意将那些风马牛不相及的因素取来,盲目地进行统计计算。从已有的研究来看,病虫害大体上与下列因素有关:

1. 病原物和虫源 病害的发生和流行与病原物(菌源)的数量、飞散和传播有关。虫害的发生期与发生量与害虫的越冬、繁殖数量以及发育速度、迁飞等有关。这是病虫害发生、蔓延的物质基础。

2. 寄主和食料 病虫害的发生、蔓延与受害作物品种、生长状况以及所处的发育期有关。显然,抗病性弱的品种,生长差的植株以及易感病的发育期,病害会加重。虫害也有类似情况。

3. 环境条件 环境条件包括气象、土壤、天敌等。在病原物、虫源和寄主、食料一定的情况下,环境条件往往对病虫害的发生、蔓延起决定性作用。其中尤以气象条件的影响

更值得注意。

本章介绍两种从上述因素的历史资料中选择预报因子的方法。通常，不同的统计预报方法，对因子有不同的要求，挑选因子的方法也有所不同。其它选择预报因子的方法将在有关章节中另行讨论。

第一节 普查法

一、方法原理和步骤

假如我们要寻找预报无锡地区5月下旬三化螟发生量的预报因子。大家都知道水稻三化螟的越冬幼虫死亡率与气象条件关系密切。若冬季温暖且降水多则死亡少，若冬季寒冷干燥则死亡率高。因此，我们可以把无锡地区1月份上、中、下旬的平均相对湿度视为无锡地区水稻三化螟越冬幼虫的一个环境条件，从中查找能指示无锡地区5月下旬三化螟发生量多少的统计规律。采用1961—1968年无锡1月上、中、下旬的平均相对湿度资料做为查找预报因子的历史资料，对应上述各年无锡地区5月下旬三化螟发生量做为预报对象，记为 y 。上述有关项目的资料如表(1.1)所示。根据表(1.1)的历史资料，采用“普查法”选择预报无锡地区5月下旬三化螟发生量的预报因子，方法步骤如下：

(一) 分组

首先，把预报对象 y 划分为个数相等的两组。将 y 的历年资料依数值的大小排列，把前面数值较大的4年资料划为一组，用符号“+”表示，而把排在后面数值较小的4年资

料划为另一组,用符号“-”表示。由表(1.1)中y的资料很容易地得到y的分组符号表,如表(1.2)所示。y的分组符号表(1.2)我们称之为“游尺”。

表(1.1) 无锡地区历年5月下旬三化螟发生量与1月各旬平均相对湿度

年 份		1961	1962	1963	1964	1965	1966	1967	1968
y (头数)		26	1034	134	1	108	13	18	67
平湿 均度 相 对%	1 月上旬	81	71	61	86	71	81	75	69
	1 月中旬	69	71	56	82	72	82	68	67
	1 月下旬	73	55	44	80	85	79	81	86

表(1.2) 预报对象y的分组符号表(游尺)

年 份	1961	1962	1963	1964	1965	1966	1967	1968
符 号	-	+	+	-	+	-	-	+

其次,把1月上旬、中旬、下旬的平均相对湿度资料也分别依数值的大小进行排列,把排列在前面较大的4年资料分为一组,用符号“+”表示,把排列在后面较小的4年资料分为另一组,用符号“-”表示。由表(1.1)中1月上旬、中旬、下旬平均相对湿度资料可以很快地做出它们的分组符号表,如表(1.3)所示。

(二) 统计符号重合次数 如果要求用当年1月上旬、中旬、下旬平均相对湿度资料查找出能指示5月下旬三化螟

表 (1.3) 无锡1月上旬、中旬、下旬平均相对湿度分组符号表

年 份		1961	1962	1963	1964	1965	1966	1967	1968
符 号	1月上旬	+	-	-	+	-	+	+	-
	1月中旬	-	+	-	+	+	+	-	-
	1月下旬	-	-	-	+	+	-	+	+

发生量多少的预报因子，则把游尺放在表 (1.3) 上，将游尺的年份与表 (1.3) 中的年份对齐，然后上下移动游尺，分别统计出游尺符号与 1 月上旬、中旬、下旬平均相对湿度符号相重合的次数。若用英文字母 m 表示游尺与被查找的资料符号相重合的次数，则游尺与 1 月上旬平均相对湿度符号重合的次数记为 m_1 ，游尺与 1 月中旬平均相对湿度符号重合的次数记为 m_2 ，游尺与 1 月下旬平均相对湿度符号重合的次数记为 m_3 ，它们分别是：

$$m_1 = 0, m_2 = 4, m_3 = 4$$

(三) 计算符号重合率，初选预报因子 用英文字母 N 表示所采用的历史资料的总个数，在本例中 $N = 8$ ，我们可以计算符号重合次数占总次数的百分率， $\frac{m}{N}$ 。若将此百分率（又叫做频率）近似地看做为概率，则可令

$$P = \frac{m}{N}$$

称之为符号重合率。

根据上一步统计的符号重合次数： m_1 、 m_2 、 m_3 ，可以

422769