

微
电
脑
在
农
业
科
学
中
的
应
用

胡秉民 编著

科学出版社

微电脑在农业科学中的应用

胡秉民 编著

科学出版社

1987

内 容 简 介

本书介绍微电脑在农业科学中的应用及其BASIC程序设计。内容包括应用微电脑程序设计实现对试验数据进行整理及其打印常规格式,小样本相关系数统计准确度的度量,两变量间最佳常用经验曲线拟合,作物病虫害防治预测预报中的逐步回归和逐步判别法,应用聚类和模糊聚类分析对作物种植区和地区小气候进行生态区划,遗传育种中的遗传力与配合力分析,品种区域试验中的品种稳定性分析,杂交亲本选配中亲本间遗传距离的测定和遗传亲疏组合的研究,农业经营中的优化设计及其农业经济预测和农业生产函数分析等,同时还介绍数据库应用技术。

本书可作为农业科技人员、高等院校农、林、生物各系科师生的工具书,也可供从事医学以及微电脑应用的科技人员参考。

微电脑在农业科学中的应用

胡秉民 编著

责任编辑 王爱琳

科 学 出 版 社 出 版

北京朝阳门内大街137号

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1987年10月第 一 版 开本: 787×1092 1/16

1987年10月第一次印刷 印张: 19 1/2

印数: 0001—2,000 字数: 451,000

统一书号: 13031·3927

本社书号: 5374·13—12

定价: 4.60元

序

随着微电脑技术的发展,广大农业科技工作者迫切要求在研究工作中推广应用微电脑,以摆脱试验数据处理中的庞杂运算过程。本书的目的就在于将农业科学各领域中的数字模型及其电算程序介绍给从事农、林、生物诸方面的科技工作者,以期把农业和生物科学研究手段提高到一个新的水平。

本书是在笔者为农业院校的研究生、进修教师和科技工作者培训班教学实践和科学研究成果的基础上,参照近几年国内外出版的有关专著和文章撰写而成。全书分十三章,除第四章矩阵与线性方程组是多变量分析程序设计的基础外,其他各章均是农业各领域中的实际应用问题。内容包括应用微电脑程序设计实现对试验数据进行整理、计算特征数及其打印常规格式,小样本相关系数统计准确度的度量,两变量间最佳常用经验曲线拟合,作物病虫害防治预测预报中的逐步回归和逐步判别方法,应用聚类和模糊聚类分析对作物种植区和地区小气候进行生态区划,遗传育种中的遗传力与配合力分析,品种区域试验中测定品种稳定性的方法,杂交亲本选配中亲本间遗传距离的测定和遗传亲疏组合的研究,农业经营中涉及生产计划、作物合理布局、畜禽饲料配方、运输合理调度诸方面的优化设计及其农业经济预测和农业生产函数分析等,同时还介绍数据库应用技术。典型示例来自试验实际,试验处理结果全部通过程序设计电算完成,并对程序注意事项加以说明,使之运行结果与实际计算步骤相符,程序通俗易懂、简便实用。

本书微电脑程序设计,选用BASIC语言。它是一种人机会话式语言,具有程序结构简单、语法浅显易懂、程序调试灵活等特点,是目前广泛应用又易于普及推广的算法语言。它不仅适用于一般微电脑,而且适用于多数可编程序型计算器(这些计算器只适用该种语言),应用面广。

应用本书程序时,可对照应用示例。全部程序在Apple II型CP/M操作系统和TRS-80型、国产机Z-3008上实算通过。对程序设计考虑到通用性,同样适用于IBM型和国产“长城”机等多种型号,由于机种型号繁多,不在此一一罗列。需要指出的是,在Apple II型CP/M操作系统和TRS-80型机中实现启动打印机功能时,需将PRINT改LPRINT,即PRINT结果是在显示屏上实现,LPRINT结果是在打印纸上实现。为了减少全书的程序篇幅,在此加以说明,但也有很多机型只要在原程序上加以命令即可实现打印纸打印结果。

全书蒙高明尉教授主审。由于微电脑在农业科学中的应用涉及面广,分章又征求各方面专家意见加以修改,在撰写过程中,得到张全德、陈闽桂、蒋学辉、章强华、徐加等同志的帮助并提供资料,同时得到浙江农业大学陈子元校长的热情支持和诸多研究生的关心与帮助,对上述同志,在此谨致衷心感谢。

由于微电脑在农业上应用这一先进技术手段在我国还是近些年来刚开始应用,很多方面理论模式还不够完善,而且作者的水平也有限,不成熟与谬误之处敬请读者指正。

浙江农业大学

胡秉民

一九八五年十二月

目 录

序	(v)
第一章 导论	(1)
§ 1. 农业试验设计结果分析的数值计算	(1)
§ 2. 育种工作中的微电脑应用	(8)
§ 3. 作物病虫害防治的预测预报	(5)
§ 4. “土壤-植物-大气”系统的研究	(7)
§ 5. 农业经营管理	(8)
第二章 试验资料的整理和总体均值估计	(10)
§ 1. 编号为N个个体所组成的有限总体中, 抽取M个个体组成随机样本 的电算程序	(10)
〔有限总体中抽取样本个体有重复、无重复的抽样程序〕	
§ 2. 数据资料的顺序排序	(12)
〔数据间逐一比较和堆法分类进行排序程序〕	
§ 3. 样本资料的特征数	(14)
〔计算样本资料的最大值、最小值、全距、平均数、标准差、标准 误、变异系数的电算程序〕	
§ 4. 总体的均值估计	(17)
〔一定概率保证下, 估算总体均值区间电算程序〕	
§ 5. 样本的频数分布	(19)
〔打印样本资料频数分布的电算程序〕	
第三章 试验资料的基本打印格式	(22)
§ 1. 单向分组资料的整理与打印格式	(22)
〔组内观察数目相等、不等、以及单向分组组内又分亚组资料整理打 印格式电算程序〕	
§ 2. 两向分组资料的整理与打印格式	(28)
〔组合内无重复、有重复观测值资料的整理打印格式电算程序〕	
第四章 矩阵与线性方程组	(34)
§ 1. 矩阵及其运算	(34)
〔矩阵排序、加(减)法、数乘、矩阵间相乘电算程序〕	
§ 2. 几种特殊矩阵	(40)
〔单位矩阵排序、转置矩阵、逆矩阵电算程序〕	
§ 3. 特征根与特征向量	(48)
〔特征根与特征向量电算程序〕	
§ 4. 线性方程组的数值解法	(51)
〔解线性方程组电算程序〕	
第五章 小样本相关系数统计准确度的度量	(55)
§ 1. 问题的提出	(55)

§ 2. 引导程序法的原理和步骤.....	(57)
§ 3. 引导程序法的应用示例.....	(57)
§ 4. 引导程序法的电算程序.....	(59)
第六章 两变量间最佳常用经验曲线拟合	(64)
§ 1. 试验资料拟合的常用经验曲线类型.....	(64)
§ 2. 建立常用经验曲线的步骤以及选择最佳拟合曲线.....	(67)
§ 3. 应用实例和最佳常用经验曲线选择的电算程序.....	(70)
§ 4. 在一定误差精度保证下高次多项式拟合的电算程序.....	(79)
第七章 作物病虫害防治的预测预报	(84)
§ 1. 应用逐步回归分析建立作物病虫害的预测方程.....	(84)
〔逐步回归电算程序〕	
§ 2. 应用逐步判别法建立作物病虫害的预测方程.....	(98)
〔逐步判别分析电算程序〕	
第八章 生态区域的划分	(119)
§ 1. 以聚类分析划分生态区的数值分类法.....	(119)
〔联合区域化试验方差分析和地区间类平均法聚类分析电算程序〕	
§ 2. 模糊聚类法进行生态区划分的数值分类法.....	(129)
〔由相似关系阵 R 出发, 建立模糊等价矩阵和模糊聚类电算程序〕	
〔夹角余弦法计算相似系数阵 R 进行模糊聚类电算程序〕	
〔数量积法计算相似系数阵 R 进行模糊聚类的电算程序〕	
第九章 遗传育种中的遗传力与配合力分析	(144)
§ 1. 广义遗传力和狭义遗传力.....	(144)
〔由原始观测数据估算广义遗传力的电算程序〕	
〔由原始观测数据出发, 从 F_2 和回交世代估算狭义遗传力的电算程序〕	
§ 2. 杂种优势与配合力分析.....	(148)
〔多个性状显著性检验的电算程序〕	
〔完全双列杂交第一种试验方法(固定模型)和第四种试验方法(固定模型)各性状配合力的方差分析、估算一般和特殊配合力效应值以及亲本特殊配合力方差的电算程序〕	
第十章 品种区域试验中的品种稳定性分析	(174)
§ 1. 品种稳定性的回归分析法.....	(174)
〔Eberhart和Russell (1966) 测定品种稳定性模型电算程序〕	
§ 2. 品种稳定性的基因型分组法.....	(193)
〔品种稳定性的基因型分组法计算品种性状平均数和变异系数的电算程序〕	
第十一章 亲本间遗传距离的测定和遗传亲疏组合	(197)
§ 1. 实施步骤的统计方法和示例分析.....	(197)
§ 2. 电算程序及其说明.....	(206)
〔多个数量性状、方差分析应用随机模型进行性状筛选及其计算遗传	

相关矩阵的电算程序]	
〔遗传相关矩阵的特征根与特征向量电算程序〕	
〔主成分值计算电算程序〕	
〔遗传距离测定及其品系聚类电算程序〕	
〔遗传通径分析电算程序〕	
第十二章 农业经营中的优化设计与农业经济预测	(221)
§ 1. 农业生产计划安排的最优决策.....	(221)
〔线性规划电算程序〕	
§ 2. 农作物合理种植结构 (作物布局)	(227)
§ 3. 畜禽饲料配方的优化设计.....	(232)
§ 4. 运输的合理调度.....	(235)
§ 5. 经济预测.....	(238)
〔一元和多元线性回归预测式及其预测精度的电算程序〕	
§ 6 农业生产函数分析.....	(246)
〔抛物线拟合农业生产函数电算程序〕	
第十三章 数据库应用技术	(251)
§ 1. 数据库系统及其组成.....	(251)
§ 2. 微电脑数据库管理系统dBASE II	(253)
§ 3. 数据库的建立和简单使用.....	(255)
附录 APPLE II型微电脑使用说明	(265)
§ 1. BASIC, MBASIC和开机操作.....	(265)
§ 2. 实用操作命令.....	(266)
§ 3. BASIC 语言的基本概念和语句.....	(270)
附表	(282)
附表 1 t 值表 (两尾)	(282)
附表 2 5% (上) 和 1% (下) F 值表 (一尾)	(283)
附表 3 25%, 10%, 2.5%和0.5% F 值表 (一尾)	(287)
附表 4 Duncan's新复极差测验5%和1%SSR值表.....	(291)
附表 5 5% q 值表 (两尾)	(293)
附表 5 (续) 1% q 值表 (两尾)	(294)
附表 6 χ^2 值表 (一尾)	(295)
附表 7 r 和 R 显著性检验表.....	(296)
附表 8 百分数反正弦 ($\sin^{-1}\sqrt{x}$) 转换表.....	(296)
附表 9 自然对数表.....	(298)
附表 10 指数函数 e^x 和 e^{-x} 表.....	(300)
附表 11 弧度和度的换算表.....	(301)
参考文献	(303)

第一章 导 论

农业和生物学研究受复杂的自然生态系统影响, 试验中随机因素比之其他领域复杂得多, 因而定性地或用简单对比法描述农业试验结果和生物现象成为一段时期内的主要手段. 随着农业现代化的进展, 计算工具得到日益改善, 微电脑在农业和生物学领域中逐步推广运用. 不论是对大农业作系统分析, 还是农业生产和经营管理中的优化设计、试验模型的统计分析、遗传育种中的杂种优势利用、品种稳定性与适应性的区域化试验、生态区划、植物病虫害防治、生物模拟以及农产品加工的最优综合评估等领域, 应用微电脑都收到很好的实际效果, 前景十分广阔. 下面就微电脑在农业上应用的几个主要方面加以概述.

§ 1. 农业试验设计结果分析的数值计算

农业试验设计统计模型的理论基础一般来自于概率论与数理统计, 但数理统计多数基本理论又源于农业科学试验. 例如, 早在1922年Hayes研究亲代与子代的蛋白质含量相关不密切时, 提出了相关分析的理论. 1923年Fisher和Maekenzie研究肥料对马铃薯产量的影响, 首先应用统计中的方差分析方法, 提出了遗传型与环境相互作用并非简单相加, 而建立起交互作用的概念. 1925年Engedow研究指出品种产量高低顺序, 在不同地区不同年份表现不一致, 说明遗传型与表现型之间之所以存在着差异, 主要因为相应地存在着交互作用. 上述交互作用效应概念的建立, 为以后农业试验模型的统计分析提供了十分重要的依据.

在我国, 农业科学工作者在进行科学实验中积累了大量的观测数据, 迫切要求把定性地描述生物现象的方法改为定量的方法, 以揭示生物现象的本质. 尤其在近一时期, 努力寻求试验设计模型及其统计分析方法, 并使其与电子计算机计算(以下简称电算)程序设计有效地结合起来, 从而把农业的研究手段提高到一个新的水平. 微电脑具有速度快、精度高、计算准确等优点, 应用微电脑可以使农业科学实验中工作量大、难以完成的多变量统计分析问题迎刃而解.

农业试验统计相应的软件发展也异常迅速. 它不仅应用于传统的常用统计方法的程序设计, 如数据整理、常用函数的数值计算、参数估计、显著性检验、方差分析以及回归分析等, 同时也对常用的试验设计, 如随机区组试验设计、拉丁方试验设计、裂区试验设计、条区试验设计、正交试验设计, 也都有了专用软件. 在多元统计分析方面, 目前已有诸如协方差分析、多元回归、逐步回归、通径分析、主成分分析、因子分析、典范分析、聚类分析和判别分析等计算机软件. 这些试验设计和统计方法与微电脑结合, 大大推动了这一领域的科学实验活动.

值得一提的是, 微电脑不仅使多个因子分析工作变得容易, 而且使一些对一般试验统计分析并非十分精通的农业科技工作者, 也可借助已有的电算程序得到所要求的结

论, 这方面工作已引起广大农业科技人员的注意。

就试验统计工作的过程来看, 农业试验工作者首先要针对试验全群, 将收集到的样本资料加以整理, 包括诸如在微电脑中贮存和打印原始数据表格, 计算样本平均数、标准差、标准误和变异系数等特征数。必要时, 可进一步打印出频率分布图和累计频数分布, 在此基础上, 进一步对统计资料进行分析。

在农业试验中, 同一品种的几个小区或单株间的产量常存在差异, 这种差异是由于试验中随机因素所致。如果不同品种或处理在不同小区或单株间试验产生了差异, 那就自然会问, 这样的差异主要是由于品种或处理(如施肥方案)不同而引起的呢? 还是由于偶然因素所引起的? 这种判断新品种与原地方品种或两种不同的处理是否在产量上存在明显的差异是属显著性检验问题, 有时称统计假设检验。其方法一般采用 t 检验法, 在设计方法上多采用成组比较试验设计或成对比较试验设计。

假若实际问题中需比较多个样本(样本数 $k \geq 3$), 则所代表各自总体间差异的显著性常采用方差分析法, 并经多重比较后可阐明品种间(或处理间)的差异显著性。

农业试验设计尤以田间试验设计为主。田间试验常用的试验设计有随机区组试验设计、拉丁方试验设计、裂区试验设计、正交试验设计和条区试验设计等。这些试验设计的任务, 就是要确立一套行之有效的统计分析方法。它由田间试验提供大量数据, 在千差万异的数据中, 通过统计分析作出符合实际的科学结论。

随机区组试验设计可适用于检验单因素试验处理效应的显著性, 也可用于多因素试验处理效应的显著性。该试验设计的好处是能有效地降低单向土壤肥力的差异, 提供较为正确的误差估计, 所以常为广大农业科技人员所采用。

拉丁方试验设计是从两个方向控制土壤差异, 因而精确度比随机区组设计高, 适用于要求精度较高的试验。但拉丁方试验设计要遵循处理数与重复数、直行区组数、横行区组数全部相等, 同时又要使方差分析中自由度不能过小, 拉丁方本身的排序又缺乏灵活性, 因此使它有很大的局限性。

裂区试验设计主要用于当试验中某一因素的主效比另一因素的主效要求更高的精确度时, 或因素间的互作比其主效更为重要时, 宜作本设计。此时需将要求精度较高的因素作为副处理安排在副区, 另一因素作为主处理, 安排在主区。

条区试验适用于两个因素都需要较大的面积, 尤其是需要长条区的较大面积时采用。

上述试验设计都属全面实施的试验设计。但当一个试验中因素增多、水平数增加时, 要做全面试验就相当困难。例如, 一个 4 因素 4 水平的试验, 就有 $4^4 = 256$ 个处理组合搭配。一个区组内包含的处理组合数太多, 区组太大, 土壤差异就难以控制。运用正交试验设计则可针对上述情况解决下面几个问题:

(1) 在全部处理组合中, 应该挑选几个处理组合来做试验?

(2) 在全部处理组合中, 具体挑选哪几个处理组合来做试验?

就其结论而言, 可解决:

(1) 对试验指标(如产量)的影响, 在诸因素中, 哪个因素重要, 哪个因素不重要?

(2) 哪个处理组合能使试验指标获得最佳效应?

试验统计分析中的回归分析, 可以通过电算程序设计, 对实验数据进行经验曲线的

配置与拟合度分析。这方面工作主要包括一元回归和多元线性回归，在一元回归中又包含直线和曲线两种类型。

近年来，多变量统计分析中诸如通径分析、主成分分析和聚类分析已广泛应用于农业科学实验，也受到农业科学工作者的重视。下面对这几种分析方法作一概述。

在农业科学研究中，不论对作物或动物进行试验研究，科学工作者总要收集多个变数资料。例如，研究作物产量与其产量构成因素穗数、粒数、粒重之间关系。研究这些多变数的一个重要目的是想知道这些变数间关系的密切程度和性质。以往研究采用简单回归和相关，或多元回归分析。但是简单回归和相关分析不能完全考察变数间的相互关系，致使结果带来一定的片面性；多元回归分析，虽在一定程度上能够消除变数之间的混淆，能够真实地表现出各个自变数和依变数的真实关系。但多元回归在分析偏回归系数时，使原因对结果的效应不能直接进行比较。同时在研究变量 1 对 y 的关系时，要把变量 2、变量 3、……等固定在一个水平上，没有考虑因素之间的相互作用。

然而，通径系数能够有效地表示相关变量间原因对结果的直接影响效应，能够估计出原因因素对效应因素的间接效应，从而能直接比较各原因因素的相对重要性。所以建立在通径系数概念基础上的通径分析，使多变数资料统计分析农业上含义的解释较之多元回归分析为优。

与通径分析类似，但由于性状数增多，而各性状之间具有一定的相关性，如单株籽粒重量与单株粒数之间有明显的相关性，所以自然会提出这样的问题：能否将多个彼此相关的性状指标变换为较少几个彼此独立的综合指标，而又不失原来多性状指标的信息，主成分分析就是针对上述提出的问题进行处理的一种科学方法。所以主成分分析的实质是将多个彼此相关的指标变换为少数彼此独立的综合指标，并要求这些较少的综合指标（又称主成分）能反映原来性状的几乎全部信息。近年来，遗传育种工作者将主成分分析应用于动、植物数量性状遗传距离预测和杂种优势的利用方面，做了大量的工作，收到很好的实际效果。

聚类分析的实质是，根据各样品的相似关系将样品分成若干组，从而实现对各群落的比较客观的分类，所以是“物以类聚”的一种科学分析方法。它在预测杂种优势、区域规划和群落分类中均有广泛的应用。其方法是利用相似系数（如相关系数）或距离（如常用的欧氏距离）作为考察指标进行聚类。相似或距离较近的点（样品）归为一类，距离较远或不相似的点应属其他的类。划类过程中，使同类内的成员（样品）尽量相似，不同类的成员（样品）则尽量相异，其实质是一种群分析。如在亲本杂交选配方面，就要选择类间距离大的亲本进行杂交，其子代表现出的优势一般来说就比较强。同时可画出相应的谱系图，更形象地阐明聚类的整个过程。有时还可进一步建立目标函数以进行判别分析。

上述所论及的农业试验设计模型结果分析的数值计算，可参见有关文献通过电算程序实现，本书不再一一罗列。

§ 2. 育种工作中的微电脑应用

作物育种在提高作物产量、改进品质、抵抗病虫害和各种不良环境条件以及适应特

殊需要等方面都有很大成就。对于保证高产稳产、增加收益、满足市场消费要求都具有十分重要的意义。选育良种不仅对于大田作物十分重要，而且对于畜牧、家蚕、蔬菜、园艺等领域都具有同样的意义，它是微电脑在农业上应用的一个重要课题。

育种工作中的微电脑应用可以从下述诸方面加以考虑：首先种质资源数据库的建立是微电脑在农业上应用的一个行之有效的方面。它是利用微电脑的“记忆功能”和庞大的“记忆容量”，建立作物的生长季节、温度范围、需要降雨量、抗病性以及其它环境因素和性状指标的数据。这些庞大的数据管理系统已成为育种家们不可缺少的工具。利用数据库技术可存贮和检索成千上万的农情信息，这对加速培育综合性状优良的新品种和开展遗传试验研究均具有十分重要的意义。

近年来，国际植物遗传资源机构对于种质资源档案资料的管理十分重视。如美国、日本、法国、英国、联邦德国以及国际水稻研究所、墨西哥小麦和玉米改良中心等都建立了电脑信息贮存和检索系统。目前世界上种质库建立的发展速度异常迅速，育种工作者利用电脑检索，可以在很短的时间内获得所需特征的种子。我国作物的种质资源十分丰富，也是世界作物体系的一支主流。国际上有些作物和果蔬品种就是利用我国的品种资源育成的。同时，我国搜集到的农家品种和野生近缘植物为主的品种已达30万份之多，今后随着国外引种、新品种选育，种质资源数量将会大幅度增加。如此众多的品质资源是国家一项极其宝贵财富，因此，建立我国的种质资源数据库就显得十分紧迫和重要。这一工作已引起了我国农业科技人员的关注，各地在建立数据库方面也做了不少探索工作并取得一定成绩。

数据库系统的主要任务，是使种质资源收集者、育种工作者以及其他科学工作者更好地利用种质资源，节省更多的时间去进行种质贮存、评价和作物改良等方面的研究。同时，由于科研情报资料检索工作的自动化，也有利于与别的机构合作，相互提供种质科学研究方面的情报资料，以便加强种质交换工作，使之形成情报网络系统，促进种质资源工作的发展。

数据库系统的建立是一项较为复杂的工作，最为简便的方法一般是在已有的DBMS基础上建立的。它是通过调查分析资源现状，汇总一份完整的种质档案资料。管理系统能做到对各种性状自动地进行分类、检索和编制各种表格。编制包括输入、修改、分类、检索、输出等方面的程序块，然后运行程序进行系统调试，并作出对管理系统的功能进行评价，编制数据库字典这一套工作流程来实现的。目前在植物遗传资源利用与贮存工作方面，可以处理种质资源目录和性状、育种世代设计、试验统计分析等。在建立数据库系统的过程中，各个环节还会出现反复，需逐步地达到相对完善的地步。另外，数据库系统的使用方法，对于不同的系统其差别很大，但每一数据库系统都为用户规定了一套使用本系统的方法介绍。其共同点在于做好原始数据的整理、核实、编辑与分析、最后实现数据库的咨询服务工作。

微电脑应用的第二方面，可以在试验数据结果的统计分析、物种间遗传距离的测定及其在杂种优势的利用诸方面。基本手段在第一节已作介绍，这儿主要说明数量性状遗传规律方面的研究问题。数量性状的遗传参数是动、植物品种的重要种性之一，也是育种工作的重要依据。掌握性状的遗传规律，可以减少育种工作的盲目性，提高育种的工作效率；同时还可以排除试验点和年度环境差异对资料的影响，所以，为了加速动、

植物选育良种工作的顺利开展,对物种数量性状遗传规律的研究是十分必要的。诸如数量遗传中的遗传力、遗传相关、遗传进度、选择指数和配合力的估算,多个性状指标的通径分析,各性状指标对考核指标贡献的主成分分析,建立在相似系数和遗传距离基础上的聚类分析,并由系谱追踪对评价以往育种成果、评价育种群体的近交程度都可以作为杂种优势利用时选配亲本的重要遗传参数。现代育种需要应用群体遗传学和统计遗传学的原理和方法,只有电脑才能对大量数据资料进行快速、准确运算和分析,以帮助预测后代的性能,指导选种和选配。

对于已选育成的作物品种,为了全面评价其优劣,进行品种的稳定性评估,了解品种推广的适宜种植范围以及品种受不同年份的不可预测环境因素影响的适应性,必须做好品种的多点试验、多年试验以及多点多年试验工作,以利于作出科学的鉴定和搞好推广工作。电算程序对于上述三种试验的综合分析,无疑具有很大的现实意义。

对于试验模型的统计分析,同时必须注意正确选取因素的效应模型(固定模型、随机模型及混合模型)和期望均方。对于效应模型的选取参见文献[1],此处不再赘述。

§ 3. 作物病虫害防治的预测预报

正确地预测预报作物病虫害的发生期、发生量,做到及时防治是确保农作物稳产、高产的重要环节。以往植病预测预报主要依据于田间调查和气象因子,凭籍经验推断或借助于简单的预报公式进行预测,这样往往准确性较差。事实上,由于近代农业的发展,使植保方面面临许多新的情况,如耕作制度的改变、栽培措施的提高、新品种的引入、品种栽培的单一化以及农药的广泛使用等,每一因素的改变,都通过影响农业生态体系而导致某些病虫害的变化。

利用微电脑进行作物病虫害防治预测预报的研究,大致从下述两个方面进行。一个是建立各种作物主要病虫害的一系列预测方程式,用以预测预报病虫害的发生期与发生量。为了有效地筛选因子并建立预测模式,我们以微电脑为工具,通常采用多元统计分析中多元回归分析、逐步回归分析、时序分析、判别分析和模糊数学(特别是模糊聚类分析)等途径,用以建立各种作物主要病虫害一系列长、中、短期预测式。

多元回归分析尤以建立多元线性回归预测式为多见,它是研究三个或三个以上变量之间定量关系的一种方法。在病虫害预测中,通常把害虫发生量或病害发生率等作为依变量,用 y 表示,而把对依变量有影响的因子,如气温、雨量或其他因子作为自变量,用 $x_1, x_2, \dots, x_m$ 表示。其主要任务是通过各因子的观测数据建立预测式。

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \dots + b_mx_m,$$

其中 b_i ($i = 1, 2, \dots, m$) 为偏回归系数,如 b_1 ,其含义是在 $x_2, x_3, \dots, x_m$ 均保持一定时 x_1 每增加一个单位时对 y 的效应。在此基础上,还需对所建立的预测式和偏回归系数进行显著性检验。若偏回归系数不显著则需逐一进行自变量剔除而重新建立预测式的工作。自变量的增多,会引起计算工作量的迅速增大。但由于有了微电脑及其软件设计,大大改善了计算技术,所以预报的准确性迅速提高。

逐步回归分析与多元回归分析不同之处是,对因子的选择是按其对预报量作用的显著程度而逐个引进的。在引进过程中,若已引进的因子由于后来因子的引入而变得不显

著时，则随时予以剔除。

时序分析方法是基于作物害虫种群变动的时间序列，反应了一切内在和外在因素对害虫种群若干代或更长时间的持续影响，并在时间序列上表现出了一定的规律性。根据病虫发生变化的以往过程，运用数字时间序列分析的方法去揭示其规律，这种方法用于长期测报有一定的参考意义。目前时序分析中较常用Box-Jenking方法，同时已编制了求Box-Jenking模型用的简便易行的电算程序。

判别分析是解决分类预报的问题。事实上，目前对病虫害要实现精确的定量预报有一定的困难。对病虫害的发生期、发生程度能作出分类别的预报，已能满足目前生产上的需要。例如，预报发生期分早、中、迟三个类别，发生量分重、中、轻三个类别。为实现这一目的，我们根据历史调查观测的数据，可以对某一病虫的发生情况进行分类或分级，并用来预测即将发生的类别。

已往在进行分类预报中，常采用点聚图。但点聚图的不足之处在于区域的划分有一定的人为性，分类标准的不统一造成预报结果的不一致性。判别分析则可从定量角度判定一个样本究竟属于何种已知母体。判别分析的内容包括两组判别和多组判别，又称两级判别和多级判别。两组判别在于判别一个样品到底属于已知两组中A与B的哪一组，但它必须属于其中的一组。有时我们认为这样的判别还太粗，需要好几组区别开来，这就是多组判别问题。解决多组判别问题的途径，一种方法是可采用前面所述的两组判别的方法，先一分为二，分成两组，再各自一分为二，依次分下去，实现逐步两组判别的方法；另一种方法是采用直接把病虫资料分成 g 级，再根据资料多个变量的观测数据，建立多级判别函数，用以判别它应归属哪一类。

在实际测报中，对于病虫问题确有一定的模糊性。比如预报量等级的隶属关系就有模糊不清的问题。例如，我们规定一个标准：二化螟全年灯下蛾量451只至600只为4级，600只以上为5级，但若灯下蛾量为598只时，就很难说它不属于5级。同时，因子观测数据由于一些随机因素的影响，使得现有数据往往只是近似地反映真实情况，所以也带有一定的模糊性。因此，在测报方法上有必要引进模糊数学的一些处理方法。

目前病虫测报方面研究利用模糊数学的基本理论主要有模糊综合评判、模糊关系方程以及模糊聚类分析等。例如，模糊聚类分析就是利用样本因子数据之间的模糊等价关系来实现分类，以达到预报的目的。而应用上述模糊数学的好处是可以把人们的经验结合进去以提高测报效果。

上述主要介绍了多元统计分析中的一些方法，这些方法在计算技术上均依赖于电脑。同时需要注意的是，这些多种方法虽然各有其独立的预测模式，但在实际应用时应全面考虑，相互补充，并结合可行性和生态学上的意义加以讨论。

微电脑在测报工作的第二方面应用是模拟各种病虫害流行规律，建立作物病虫害发生的模拟模型。模拟模型是把与病虫害发生有关的基本因素及其作用模型化，以电算程序形式，用微电脑模拟地把自然的病虫害发生过程再现出来的一种方法。在这种模型中，能提供全面分析害虫的消长和病原菌增值的可靠信息。

构造模型需涉及多学科的知识，如生物学、数学和电脑知识等，它是多学科交叉的产物。构造模型着重需要解决两个问题：一是综合生物学方面的信息使其转变成数学模式；二是根据数学模式编制电算程序，使其实现用微电脑模拟的再现过程。程序调试后

运行的结果通常称第一代模型，经过与实际结果对比，并逐步取得改进可形成第二代、第三代、…模型。

1969年美国的Waggoner发表了世界上第一个番茄早疫病发生的模拟模型EPIDEM。他为了模拟番茄早疫病的发生过程，首先求出了病原菌与植株在各个生育阶段所处环境之间的定量关系，然后利用这些数学式制作了模型，利用这个模型模拟了在各种环境条件下早疫病发生的过程，其相对发生量与田间调查情况酷似。而后美国、联邦德国、加拿大、日本、法国、荷兰、新西兰、苏联等国建立了诸如菊花花腐病、玉米小叶斑病和大叶斑病、马铃薯晚疫病、柑桔黑点病、大小麦条锈病、芹菜斑点病、小麦颖枯病、苹果黑星病、稻瘟病、柑桔红蜘蛛、棉铃虫、果树红叶螨等模拟模型。七十年代美国在棉铃虫防治研究中，通过设计一种描述棉铃虫种群动态的模拟模型，全面模拟了棉铃虫的发生、成长动态和恶果，从可能采取的各种模拟防治策略中找出了最适时、合理的防治策略和措施。

1981年我国曾士迈等制作了小麦条锈病春季流行模型。之后，随着电脑的普及，我国越来越多的科学工作者从事这方面的研究。目前模拟方法多采用单一状态变量的简单模型、多个状态变量的系统模型和概率随机模型等，并使建立的模型对已知的参量、模型的状态变量值能与实际情况相吻合。

§ 4. “土壤-植物-大气”系统的研究

利用微电脑开展“土壤-植物-大气”系统的研究，内容主要包括两个方面。其一是揭示“土壤-植物-大气”系统中的基本理论，探索建立数学模型，用它来更深刻地揭示出植物生长与大气和土壤的关系。在这个系统中，由于构成多因素多环节的复杂系统，各因素在各环节都可能表现出他们的单独以及联合的作用。测定这个系统产生影响的主要因素及其相互作用关系的表现，归纳整理分析，运用系统工程、电脑模拟和多变量回归分析等方法找出其反映作物特性的规律性本质，用以建立作物产量、经济收益等预测模型和提出各种因素对农业的综合影响，优选农艺措施的组合方案。农业系统着眼于综合最优决策，即综合考虑生态（植物、土壤、气候）与经济多种条件，运用系统分析方法确定对当地生态条件较适应、经济收益较高的最优决策。同时要注意到农业系统中的一切综合最优决策都是相对的，这是由于农业系统构成因素的复杂性及其差异所致，同时最优决策又随着时间的不断变化而改变，这就要求系统分析各项技术和经济措施，做到因地制宜、因时制宜。

其二是研究这个系统中各个环节对作物生长所形成的影响。诸如确定各生产环节中最佳施肥方案的选择；各阶段作物对土壤水分的需求以实现科学灌溉；探求作物不同生育阶段的最佳模式，进行控制条件下的作物生长电脑模拟；园艺温室电脑控制装置及小气候环境模拟等。利用微电脑解决农业系统中存在的多种局部最优决策，这里既包括生态最优，同时又包含经济最优方面的问题。

在确定最佳施肥量的研究中，微电脑将记载主要耕作条件并对进行土样分析的地块及时地提出具体的施肥意见，还可以根据普查结果中土壤有效养分、条件和布局，在较大范围内提出对肥料的需求，从而也为化肥的生产和分配提供科学的依据。最佳施肥量

的计算往往要通过大量田间肥料试验,运用多元回归分析以模拟施肥量与产量之间的施肥生产函数,而施肥生产函数的获得又为最佳施肥量的计算提供了可能性。应用微电脑编制施肥建议,是针对不同地区、不同作物而提出的应施用的化肥品种、施肥量、最佳施肥方法以及最适宜的施肥时期。考虑经济效益方面,还可列出不同肥料搭配下对作物产量的影响和收益情况,用以确定在获得较高经济效益下的经济施肥量。

微电脑在作物产量预报方面工作,目前一般采用统计数值预报法。这种预报方式有的采用把实际产量划分成两个部分:一部分是在正常气候条件下,当时、当地生产水平所能达到的产量;另一部分则是由于当地、当年的气象异常所引起的产量波动,称之为修正产量。在做作物产量预报时,就是根据历史产量资料和气象资料对当时、当地正常水平所能达到的产量和修正产量作出估算。当时、当地正常水平所能达到的产量可采用正交多项式逐次拟合实际产量中估得,并可建立必要的气象数据文件。而有的对作物产量估算是以品种基因型及环境生态因子的综合制约为依据,根据环境生态因子的组合特征,定量分析一定生态区内环境生态因子组合对产量构成因素的效应,同时阐明主导因子的作用。统计方法可采用如逐步回归、通径分析和主分量分析等,用以探讨在环境因子处于最理想状态下的品种生产潜力。

作物产量始终受复杂的气象因子的综合作用,充分探明这种复杂关系对于农业科学的发展以及农业生产实践都具有十分重要的意义。就作物本身的生命过程来看,一方面受环境因子的强烈制约,另一方面又反作用于环境。对于如此复杂的作用规律,往往需要在模拟设施中加以研究。这种人工环境模拟设施,通常指与自然大气相隔离的人工气候箱或人工气候室。随着电脑的出现,才能为模拟自然提供有力的手段。目前微电脑在塑料大棚、大型温室控制技术的管理系统中已得到应用与推广。

在农业气象方面,由于气象资料复杂且预报又要求及时,而气象又是影响农业生产主要因素之一,为了更有效地为农业提供气象情报,利用电脑建立农业情报观测网络和资料传播系统则是一项很有意义的工作,目前,气象部门在引进电脑技术中作了不少工作。

作物栽培技术中,科学灌溉是一项重要的农技措施。微电脑可用以根据作物生长、土壤情况适时提供给水、排水方案,根据作物对土壤水分平衡的特定性用以控制作物的产量。在制订灌溉计划时,采用微电脑技术与土壤水分传感器和灌溉自动化控制系统相结合的方式,可做到省水节能。

§ 5. 农业 经 营 管 理

农业经济中的管理及其优化设计,是微电脑应用的一个重要方面。发展农业生产和搞好经营管理是增加物质财富的两个方面。农业科学现代化除了实现电气化、机械化方面工作之外,要抓好现代经营的管理工作。抓好投入、产出的综合平衡模型的研究、资源最优利用与分配的研究、统筹方法与经营管理的研究、经济预测的研究、生产和运输中的合理调配研究等等方面。这些方面的大量工作都可以同微电脑结合起来,从而发挥其最大的经济效益。

在优化设计方面,微电脑还可用以模拟农产品加工过程是否合理,以建立最佳的加工体系及加工自动化。农产品加工中还可根据不同动物对饲料营养成分的不同要求,确

定最佳饲料配方。随着微电脑及配合饲料工业的发展,应用电算技术进行配合饲料生产已成现实。其原理是根据配合饲料原料的库存量、畜禽对各种营养成分的需求,以及市场价格变动等条件,并提出约束条件,微电脑即可用这些信息计算出成本最低的全价饲料配合方案。目前饲料配方的电算程序多采用线性规划方法,要求解一个由多种原料(多变量)配合、满足多种营养指标(约束条件),而使配合饲料成本最低(目标函数值最小)这样一个线性规划问题,其实质是求某一线性目标函数,在一定约束条件下的最小值问题。

微电脑还可存贮农业生产过程的各项数据以及市场信息,如作物、耕作措施、土地资源、劳力、机械设备以及能源利用与价格等,为提高管理效率和优化决策提供依据。目前国外还利用电脑开展咨询服务,根据所拥有的土地数量、环境条件和资金状况,为用户提供种植作物门类、预计产量和收益、购买农机具型号和数量等商业性服务。

实现科学的经营管理,其前提是搞好农业发展规划。规划包括农业发展总体规划、作物优化种植结构以及农业生态区划等方面的问题。系统工程方法被应用到农业发展总体规划领域,它是把组成农业系统的各个因素结合在一起用以研究多重目标。农业系统中既要考虑生产,又要考虑消费与流通;既要考虑产量,又要考虑经济效益;既要考虑经济目标,又要考虑自然生态环境的保护;既要考虑所在部门的局部利益,又要考虑国家的整体利益,还需注意把系统工程方法和多年计划工作实践经验的常规计划方法相结合,使规划更为完善。

农业规划需要考察多个因素的影响,它需要考虑天时、地理、作物布局、经济状况以及科学文化诸因素,没有大功能的电脑技术去处理这方面的信息是难以实现的,电脑在这方面的主要任务用以建立信息数据库和优化模型。在农业模型制作中,采用以电脑为工具的计划和情报系统,使其能对于资源分配、能源利用、农业生产结构以及政策手段(指标、价格、税收、补贴等)作出决策。在区划工作中,应用遥感新技术也是获得数据资料的一个方面。电脑对用遥感技术得到的大量信息数据自动地进行记录、识别和分类处理,并且能够准确地直接得到图象的判读结果。

生态区的划分可利用电脑进行多因素分析,即使对于某类作物的布局,也可作出合理的区划。过去作物栽培适应区域主要按行政区划分试验点进行试验,如每个县设置良种场进行试验。但更为科学的是应该在不同的生态区有代表性地进行试验,而对于同一生态区原则上减少重复设点,以节省大量人力、物力。目前已开始应用微电脑编制作物种植计划,并代替人们来安排农事。

上述几节仅罗列了微电脑在农业上应用的几个方面,但这些方面并非包括全部应用范围,很多方面还远未开发。例如在动植物的分类方面,过去该项工作多属定性分类,一定程度上凭借经验。但有了电脑后,可以根据数值性状指标,利用诸如判别分析和聚类分析方法对各数量性状作出综合评估,使其为分类提供更多的信息,以求分类更趋合理。又如农业文摘电脑检索方面,利用电脑检索文献来代替手工检索。读者只要把检索要求填入一张提问单内,再经检索人员把读者检索要求编制成机读形式,电脑系统即可从文摘磁带内的期刊中检出读者所需要的文献。

上述大量事实可以看出,电脑技术不仅是实现农业现代化所不可缺少的工具,而且也是当代衡量一个国家科学技术水平的标志之一,随着这一新技术的推广应用,必将在农业科学各个领域中发挥越来越大的作用。

第二章 试验资料的整理和总体均值估计

农业试验的目的在于探讨被研究总体的特征，而具体研究时往往又只能从总体中抽取一部分个体作为样本加以研究，因而针对农业试验的全群，如何合理地去收集资料，做好观察、测定工作，就成为农业试验资料整理和分析的前提。农业试验中常用的取样方法有简单随机取样、分层取样、整群取样、两级取样和顺序取样等。本章着重介绍简单随机抽样的电算技术，同时将收集到的资料通过编制算法程序加以整理，并相应计算某些重要的特征数和对总体作出均值估计。

§ 1. 编号为N个个体所组成的有限总体中，抽取M个个体组成随机样本的电算程序

为了考虑样本的代表性，要求样本的各个体必须从总体中随机抽取，这样从总体中随机抽取的样本称为随机样本。为了取得随机样本，就要求在抽样方法上不依主观因素，而使总体中的每一个个体被抽中是等可能的，即每一个个体具有相等的被抽取机会。

总体为有限时，可对总体内个体进行编号，设编号由N个个体组成。今假设要求从N个个体所组成的有限总体中，抽取M个个体组成的随机样本，以往常采用抽签法或查随机数字表获得其随机样本。本节介绍有限总体中所抽取的随机样本个体，可以重复和不可以重复两种抽样程序。

1.1 有限总体中抽取样本个体可以重复的抽样程序

〔程序〕

```
10 REM RANDOM SAMPLE FROM FINITE POPULATION WITH REPLAC  
EMENT  
20 PRINT "WHAT IS THE POPULATION SIZE ?"  
30 INPUT N  
40 PRINT "WHAT IS THE SAMPLE SIZE ?"  
50 INPUT M  
60 PRINT  
70 FOR I=1 TO M  
80 LET K=1+INT(RND(1)*N)  
90 PRINT USING "#####";K  
100 NEXT I  
110 END
```