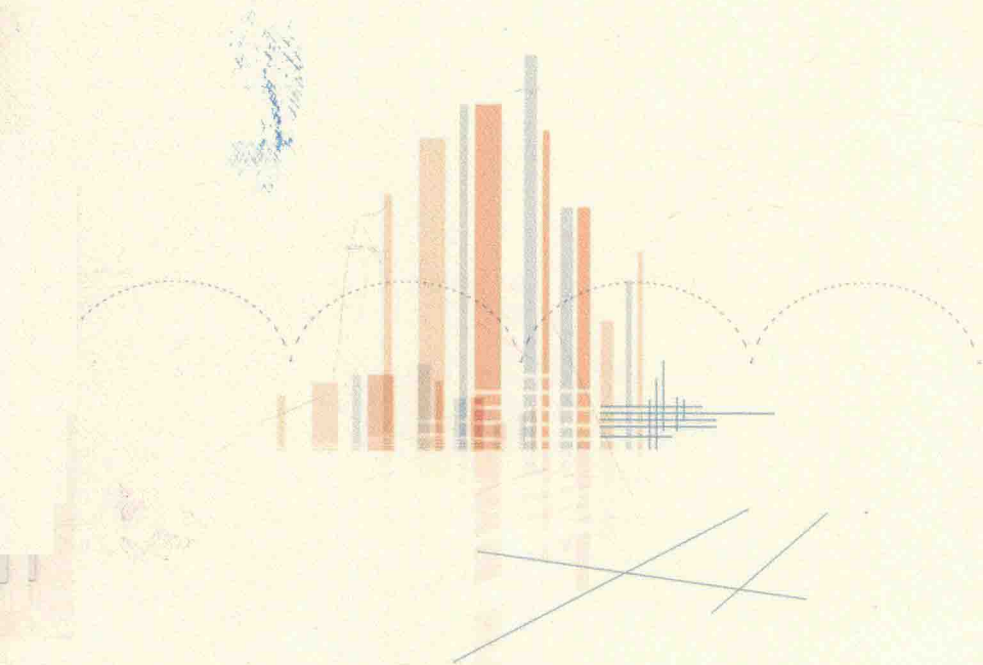


农作物面积空间抽样方法研究

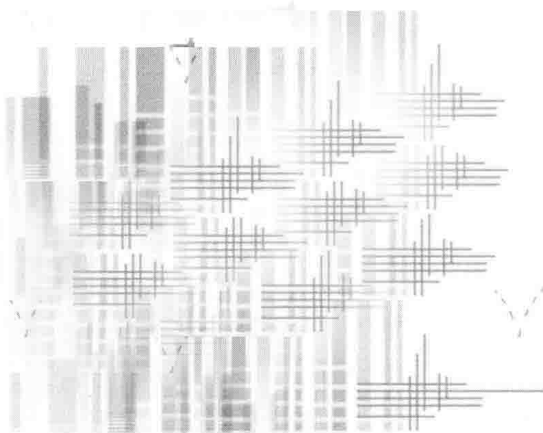
王迪 著



中国农业科学技术出版社

农作物面积空间抽样方法研究

王迪 著



中国农业科学技术出版社

图书在版编目 (CIP) 数据

农作物面积空间抽样方法研究 / 王迪著. —北京: 中国农业科学技术出版社, 2015. 12

ISBN 978 - 7 - 5116 - 2469 - 7

I. ①农… II. ①王… III. ①作物 - 种植面积 - 抽样调查统计 - 研究 IV. ①F307.1

中国版本图书馆 CIP 数据核字 (2015) 第 301559 号

责任编辑 闫庆健 张敏洁

责任校对 马广洋

出版者 中国农业科学技术出版社

北京市中关村南大街 12 号 邮编: 100081

电 话 (010)82106632 (编辑室) (010)82109702 (发行部)
(010)82109709 (读者服务部)

传 真 (010)82106650

网 址 <http://www.castp.cn>

经销者 各地新华书店

印刷者 北京华正印刷有限公司

开 本 850mm × 1 168mm 1/32

印 张 5.25 彩 插 12 面

字 数 124 千字

版 次 2015 年 12 月第 1 版 2015 年 12 月第 1 次印刷

定 价 20.00 元

— 版权所有 · 翻印必究 —

内 容 提 要

本书系依据作者承担的国家“863”计划统计遥感重点项目和国家自然科学基金项目的研究成果编著而成。全书针对农业抽样统计调查过程中存在的科学和技术问题，重点阐述了大区域尺度条件下的主要农作物种植面积空间抽样调查方案及其关键要素（样本容量、样本空间布局、抽样单元尺寸）的优化设计方法。共分6章，主要内容包括：①耕地面积空间抽样方案设计；②冬小麦种植面积空间抽样方案设计；③农作物面积空间抽样单元尺寸设计；④农作物面积空间抽样样本布局设计；⑤农作物面积空间抽样方案优化设计。

全书具有较强的系统性和创新性，可供农业统计、地学、生态、环境、空间信息及社会科学等领域的学者在抽样调查、统计推断和监测网优化设计中参考使用。

序

我国是农业生产大国。农作物种植面积、产量信息是国家制定粮食政策和经济计划的重要依据。及时了解、掌握农作物种植面积，对于准确估计和预测农作物产量，加强农业生产管理，确保我国粮食安全具有重要意义。

长期以来，我国统计部门主要采用省抽县、县抽乡、乡抽村、村抽农户的传统多级目录抽样方式进行农作物面积统计调查。受自身运行机制所限，目录抽样集中表现的问题为抽样框更新速度慢、调查对象稳定性差、调查过程中对空间信息利用不够。随着“3S”技术（遥感、地理信息系统和全球定位技术）的日益发展，将传统抽样与“3S”技术相结合的空间抽样方法已被广泛应用到大区域农作物面积监测与估计中。由于空间抽样将传统的以农户为调查对象的主观调查变为以耕地地块为调查对象的客观调查，具有准确、高效等诸多优点，可有效解决传统抽样调查存在的各种问题，因此，进一步深入研究和大力推广应用现有研究成果意义重大。

2006年，国家高技术研究发展计划（863计划）地球观测与导航技术领域专门设立了“国家统计遥感业务系统关键技术研究与应用”重点项目。其中，农作物面积空间抽样调

查方法是这一重点项目中的关键组成部分。中国农业科学院农业资源与农业区划研究所遥感研究室的科研团队多年来广泛开展了基于“3S”技术的农作物面积空间抽样调查方法研究工作。这些研究取得了重要成果。虽然基于“3S”技术的农作物面积空间抽样调查研究还存在一系列亟待解决的问题，但对已有成果进行一次系统地梳理十分必要，目的在于将成功的经验、方法和理论总结出来，使有待进一步研究和讨论的问题系统化，以便在今后的研究中解决和应用。

《农作物面积空间抽样方法研究》是作者数年努力的成果结晶。该书围绕国家农业统计调查业务对空间抽样方法的应用需求这一主题，详细介绍了农作物种植面积空间抽样调查的最新研究进展、数据获取与处理流程、样本抽选与统计推断原理、空间抽样方案及其基础要素优化设计方法，并有预见性地提出了基于“3S”技术的空间抽样方法在国家农业统计领域应用的前景。书中采用了大量第一手实验数据，写作风格质朴、内容翔实、案例具体，突出了新颖性和实用性。值此书出版之际，我愿意将其推荐给广大读者，希望该书对促进遥感技术在国家农业统计领域的进一步应用，推动我国农业向精准化、数字化方向发展起到积极作用。

周清波

2015年11月

目 录

第一章 绪论	(1)
第一节 农作物面积空间抽样目的意义	(1)
第二节 农作物面积空间抽样方法研究进展	(2)
第三节 空间抽样单元尺寸优化设计研究进展	(12)
第四节 空间抽样样本布局优化设计研究进展	(14)
第五节 农作物面积空间抽样研究现状分析	(17)
第二章 耕地面积空间抽样方案设计	(19)
第一节 概述	(19)
第二节 抽样总体分布正态性检验	(20)
一、试验区基本情况	(20)
二、样方尺寸优化设计方法	(21)
三、离散抽样区, 构建抽样框	(25)
四、抽样总体正态性分布检验	(25)
五、选取常用抽样技术, 计算样本容量、抽取样本、 外推总体及估计抽样误差	(26)
六、抽样基础单元空间变异性分析	(33)
第三节 多种抽样技术的效率比较	(37)
第四节 层数大小对分层抽样效率的影响	(39)
第五节 样方尺寸的优化设计	(42)
第六节 抽样基础单元空间变异性分析	(44)
第七节 样本空间布局的合理化调整	(47)

第八节 样本布局合理化调整前后抽样误差结果比较	(50)
第九节 结论与讨论	(51)
一、结论	(51)
二、讨论	(52)
第三章 冬小麦种植面积空间抽样方案设计	(53)
第一节 绪论	(53)
第二节 多种空间抽样方案效率比较	(55)
一、试验区基本情况	(55)
二、构建抽样框	(55)
三、遥感影像对地抽样	(58)
第三节 抽样基础单元尺寸优选	(81)
第四节 样本容量确定	(82)
第五节 结论与讨论	(84)
一、结论	(84)
二、讨论	(85)
第四章 农作物面积空间抽样单元尺寸设计	(86)
第一节 绪论	(86)
第二节 样方尺寸初选结果	(88)
一、研究思路	(88)
二、试验区基本情况	(88)
三、基础数据	(89)
四、样方尺寸初选	(89)
五、空间抽样方案设计	(91)
六、样本观测值获取	(94)
第三节 样方尺寸优选结果	(95)
第四节 结论与讨论	(98)
一、结论	(98)

二、讨论	(99)
第五章 农作物面积空间抽样样本布局设计	(100)
第一节 绪论	(100)
第二节 抽样单元内冬小麦种植面积比例描述性 统计学特征	(102)
一、研究思路	(102)
二、研究区概况	(102)
三、基础数据	(103)
四、样本初选方案设计	(104)
五、抽样单元空间关联阈值确定方法	(106)
六、样本布局优化设计方法	(107)
第三节 正态分布检验	(109)
第四节 抽样单元内冬小麦种植面积比例空间变异 特征分析	(111)
第五节 样本布局优化设计结果分析	(113)
第六节 结论与讨论	(115)
一、结论	(115)
二、讨论	(116)
第六章 农作物面积空间抽样方案优化设计	(119)
第一节 绪论	(119)
第二节 抽样单元尺度优选	(121)
一、研究思路	(121)
二、研究区概况	(122)
三、基础数据	(122)
四、空间抽样方案设计	(123)
第三节 分层标志选择	(129)
第四节 样本空间布设方式和抽样比优选	(131)
一、样本空间布设方式	(131)

二、抽样比优选	(131)
第五节 结论与讨论	(133)
一、结论	(133)
二、讨论	(134)
参考文献	(136)
符号表	(151)

表目录

表 2-1	试验区不同像元尺度下的全局空间自相关指数结果	(22)
表 2-2	样方大小设计结果	(24)
表 2-3	不同样方尺寸下北京耕地面积抽样框	(25)
表 2-4	按县区行政边界分层结果 (10 层和 7 层)	(30)
表 2-5	按县区行政边界分层结果 (5 层和 3 层)	(31)
表 2-6	按耕土比分层抽样结果	(33)
表 2-7	北京市耕地面积抽样总体分布正态性检验结果 ...	(37)
表 2-8	多种抽样技术下样本容量计算结果	(38)
表 2-9	多种抽样技术下抽样误差计算结果	(39)
表 2-10	不同层数下分层抽样效率结果 (行政边界为 分层标志)	(40)
表 2-11	不同层数下分层抽样效率结果 (耕土比为分 层标志)	(41)
表 2-12	按耕土比分层条件下各层样本容量计算结果 (7 层)	(41)
表 2-13	按耕土比分层条件下各层样本容量计算结果 (10 层)	(42)
表 2-14	各抽样技术在不同样方尺寸下的抽样误差结 果	(43)
表 2-15	8 种抽样框下以耕土比为分层标志的分层抽 样样本容量计算结果	(44)

表 2-16	各抽样框下的抽样基础单元空间变异性模拟结果	(45)
表 2-17	6 000m × 6 000m 抽样框下分层抽样中各层抽样基础单元空间变异模拟结果	(45)
表 2-18	5 000m × 5 000m 抽样框下分层抽样中各层抽样基础单元空间变异模拟结果	(46)
表 2-19	2 000m × 2 000m 抽样框下分层抽样中各层抽样基础单元空间变异模拟结果	(46)
表 2-20	1 000m × 1 000m 抽样框下分层抽样中各层抽样基础单元空间变异模拟结果	(46)
表 2-21	700m × 700m 抽样框下分层抽样中各层抽样基础单元空间变异模拟结果	(47)
表 2-22	500m × 500m 抽样框下分层抽样中各层抽样基础单元空间变异模拟结果	(47)
表 2-23	样本空间布局调整前后抽样误差计算结果	(51)
表 3-1	试验区冬小麦生长物候历	(55)
表 3-2	国内外遥感抽样调查样方尺寸总结	(57)
表 3-3	样方大小设计结果	(57)
表 3-4	各抽样框总体单元数及总体总值统计结果	(58)
表 3-5	快速近似法计算分层界限	(64)
表 3-6	各抽样框初级单元划分结果	(69)
表 3-7	各抽样框初级和二级单元抽样数计算结果	(71)
表 3-8	代码法抽取初级样本单元过程	(74)
表 3-9	分层泊松抽样实施过程	(76)
表 3-10	3 000m × 3 000m 抽样框条件下各抽样方法样本抽取与总体外推结果	(78)
表 3-11	2 000m × 2 000m 抽样框条件下各抽样方法样本抽取与总体外推结果	(78)

表 3 - 12	1 000m × 1 000m 抽样框条件下各抽样方法样本抽取与总体外推结果	(79)
表 3 - 13	500m × 500m 抽样框条件下各抽样方法样本抽取与总体外推结果	(79)
表 3 - 14	300m × 300m 抽样框条件下各抽样方法样本抽取与总体外推结果	(80)
表 3 - 15	各抽样方法抽样外推总体误差结果	(80)
表 3 - 16	各抽样方法所需样本容量计算结果	(81)
表 3 - 17	不同样方尺寸水平下样本外推总体误差结果 ..	(81)
表 3 - 18	300m × 300m 样方水平下样本容量计算及外推总体误差结果	(82)
表 3 - 19	500m × 500m 样方水平下样本容量计算及外推总体误差结果	(83)
表 3 - 20	1 000m × 1 000m 样方水平下样本容量计算及外推总体误差结果	(83)
表 3 - 21	2 000m × 2 000m 样方水平下样本容量计算及外推总体误差结果	(83)
表 3 - 22	3 000m × 3 000m 样方水平下样本容量计算及外推总体误差结果	(84)
表 4 - 1	8 种样方尺寸下利用分层抽样外推研究区冬小麦种植面积总体与误差估计结果	(97)
表 5 - 1	8 种尺寸下抽样单元内冬小麦种植面积比例基本统计学特征 (基于研究区 2009—2010 年冬小麦空间分布数据)	(108)
表 5 - 2	8 种尺寸下抽样单元内冬小麦种植面积比例统计分布检验 (基于研究区 2009—2010 年冬小麦空间分布数据)	(110)

表 5-3	8 种尺寸下抽样单元内冬小麦种植面积比例的 变异函数模型及其参数（基于研究区 2009— 2010 年冬小麦空间分布数据）	(112)
表 5-4	样本布局优化设计前后的研究区 2009 和 2010 年冬小麦种植面积空间抽样外推总体与误差估 计结果	(113)

第一章 绪 论

第一节 农作物面积空间抽样目的意义

我国是农业生产大国，粮食作物产量的丰欠历来受到社会、政府部门的重视。粮食作物的播种面积、产量等信息更是国家制定粮食政策和经济计划的重要依据（陈水森等，2005）。及时了解、掌握主要农作物的种植面积，对于准确估计和预测农作物产量（张养贞，1995；Quarmby 等，1993；Fulu 等，2005；Xavier 等，2005；Loveland 等，1991；Reynolds 等，2000），加强作物生产管理，确保我国粮食安全具有重要的意义（Reynolds 等，2000；Quarmby 等，1993；Allen 等，2002）。

长期以来，我国耕地和农作物播种面积等重要农业统计数据，主要采用村级起报、逐级汇总上报的全面统计方法。在全国范围内开展全面调查方式统计作物播种面积，存在调查工作量庞大、财力和物力耗费高及调查周期长等诸多缺陷。随着我国社会经济的快速发展，政府决策部门、社会公众对农作物统计数据的需求越来越高，全面调查方法已无法适应当前市场经济条件下农村新形势发展和反映农业结构调整变化的需要（国家统计局农村调查总队，2002；国家统计局农村社会经济调查司，2006）。

抽样技术作为一种非全面统计调查技术，因具有节省调查费用、缩短调查周期及改善调查数据质量等优点而得以在农村社会经济调查中广泛应用。我国现有农作物面积、产量抽样调查采用

省抽县、县抽乡、乡抽村、村抽农户的传统多级目录抽样方式(国家统计局农调总队, 2002)。受自身技术体系所限, 目录抽样调查集中表现的问题为: 抽样框更新速度慢、误差大, 难以满足各级部门业务需求; 调查过程中对空间信息利用不够以及调查结果空间化表达较差等。针对目前国家农村抽样统计调查中存在的问题, 本研究拟通过遥感(RS)、地理信息系统技术(GIS)和应用抽样技术及地统计学的联合应用, 进行作物面积空间化样本抽选与推断方法研究, 具体研究内容包括作物面积空间抽样框架(空间抽样技术)及抽样基础要素(空间抽样过程中涉及的基础关键因子, 包括样本容量、样方尺寸及样本空间布局)优化设计, 旨在为改进现有作物面积抽样调查技术体系提供参考依据。本研究的实施将有利于提升国家农村抽样统计调查的业务水平, 有利于拓展遥感技术的应用领域。

第二节 农作物面积空间抽样方法研究进展

抽样是将总体集按某种规律划分为样本集且不损失总体主要信息的方法; 空间抽样则是针对在地理空间上分布且相互间有关联性的研究对象而言, 它是具有空间关联性的抽样。

早在1974—1986年, 美国开展的“大面积农作物估产试验”(LACIE)和“农业和资源的空间遥感调查”(AGRISTARS)就采用面积抽样框(Area Sampling Frame)方法进行作物面积抽样估算。两计划针对不同级别区域采用不同的抽样布点方式。在有国家级详细历史数据的地区(如州)采用二阶随机分层抽样布点; 在低于州一级的国家地理区域采用标准分层抽样布点(Fecso等, 1986; Cotter和Nealon, 1987)。FAO在其刊印的统计发展丛书中对3种常用农业调查方法(目录框、面积框和多样框抽样调查)的抽样设计、精度评价及适用范围进行了详尽

介绍 (FAO, 1996; FAO, 1998), 相关报道中指出, 目录框调查是最早应用于农业方面的概率抽样调查方法, 虽然该法具有实施方便等优点, 但由于受到目录框 (例如覆盖全国范围的目录框) 不完整、不准确以及更新速度慢等因素影响, 往往表现为抽样结果误差较大。面积框方法由于克服了目录框方法中存在的上述缺点, 因而在农业面积抽样调查中逐渐得到广泛应用 (Houseman, 1975; Vogel, 1986; Kish, 1989)。多样框法实际上是将面积框法和目录框法结合起来的一种概率抽样方法 (Nealon, 1984)。该法主要为具有高度偏态分布的农业变量提供合理估计 (对数目较少但占总体比重很大的调查单位进行目录框抽样, 对数目较多但占总体比重较小的调查单位采用面积框抽样)。在作物面积、产量及长势等农业调查方面, 综合比较 3 种抽样调查方法, 面积框法迄今应用最为广泛 (多样框法实际是在面积框法的基础上融入目录框法混合而成)。在面积框抽样方法中, FAO 主要针对 3 种类型地物边界即切块 (segments) 边界进行了样本设计: ①具有可识别永久性自然边界样本设计; ②方形切块样本设计; ③含有与农业经营单位土地吻合的切块样本设计。随后, 针对第一种类型切块边界的面积抽样设计、相关估计量构建及方差评价进行了具体阐述。其中在面积抽样设计部分, 具体的实施步骤如下: 1) 基于耕地比例、作物类型、特殊农业耕作方式、都市农业区或其他土地利用特征在拟选面积样框内划分土地利用层。2) 在每一土地利用层内划分互不重叠的初级抽样单位 (PSUs), 每个 PSUs 的规模等于它所包含的切块数。然后对每层的 PSUs 按相似性进行排队, 最后采用 PPS 抽样 (与规模成比例的概率抽样) 抽选 PSUs。3) 在每一抽中的 PSUs 中, 按等概率抽选切块。与目录框法相比, 面积框抽样法虽然具有不受样本更新影响, 抽样误差范围小等优点, 但同时也存在对包含在样框内所有土地的制图资料 (地图、卫星成像及航空照片)