

# 基于数理统计的 小型灌区样点灌区选取方法研究

吴松, 蔡守华

(扬州大学水利科学与工程学院, 江苏扬州 225009)

**摘要:**为合理确定样点灌区数量,减少灌溉水利用系数测算中的统计工作,提高灌溉水利用系数的测算精度,以扬州市仪征市小型灌区为例,对小型灌区样点灌区的选取数量、选取方法进行研究,运用统计学原理,采用概率抽样中的简单随机抽样和非概率抽样中的配额抽样的方法分别确定并选取样点灌区的数量并相互比较,为以后灌溉水利用系数测算过程中样点灌区的选取方法提供参考与借鉴。

**关键词:**数理统计; 概率抽样; 样点灌区; 方法研究

**中图分类号:**[TV93] **文献标识码:**B **文章编号:**1007-7839(2022)06-0056-0004

## Research on the selection method of sample irrigation areas in small irrigation areas

WU Song, CAI Shouhua

(College of Water Conservancy Science and Engineering, Yangzhou University, Yangzhou 225009, China)

**Abstract:** In order to reasonably determine the number of sample irrigation areas, reduce the statistical work in the calculation of irrigation efficiency and improve the calculation accuracy of irrigation efficiency, Taking the small-scale irrigation area in Yizheng City in 2018 as an example, this paper studies the selection quantity and selection method of sample irrigation areas in small-scale irrigation areas, and uses the statistical principle to determine and select the number of sample irrigation areas by using the methods of simple random sampling in probability sampling and quota sampling in non-probability sampling, It provides a reference for the selection method of sample irrigation area in the calculation of irrigation water utilization coefficient in the future.

**Key words:** mathematical statistics; probability sampling; sample irrigation area; method study

灌溉水利用系数是用来衡量农田灌溉水利用效率的指标,如何更加精准地测算灌溉水利用系数是目前需要攻克的难题,而样点灌区的选取则是灌溉水利用系数测算中一个重要的环节。样点灌区的选取数量多少不仅直接影响到灌溉水利用系数测算过程中的工作量,更直接反映在最终测算结果的精确程度上<sup>[1]</sup>。《国务院关于实行最严格水资源管

理制度的意见》(国发[2012]3号)明确要求,确立用水效率控制红线,到2030年,农田灌溉水有效利用系数提高到0.6以上。我国小型灌区工况及水源各异,分布范围广,数量众多,计量设施安装率低且大多没有专门的管理机构,所以有必要深入研究样点灌区选择方法,在准确估算样点灌区数量的基础上减少农田灌溉用水量统计的工作量,以及提高其估

收稿日期: 2022-01-15

作者简介: 吴松(1997—),男,硕士研究生,主要从事灌溉排水理论及技术研究。E-mail:982132396@qq.com

通信作者: 蔡守华(1964—),男,教授,主要从事灌溉排水及水生态工程技术研究。E-mail:caish@yzu.edu.com

算精度。

传统的样点灌区选取数量是依据《全国农田灌溉水有效利用系数测算分析技术指导细则》,综合考虑灌溉面积、种植作物类型等因素而确定。一般小型灌区样点灌区数量抽取比例为全省小型灌区数的0.5%,选取数量受个人主观意愿以及地区差异性等因素影响。张建华<sup>[2]</sup>采用概率抽样中重复性抽样的抽样公式对全国灌溉分区进行了样点灌区的选取。沈莹莹等<sup>[3]</sup>提出一种农田用水量推算工作中样点灌区的选取方法。李泽坤等<sup>[4]</sup>提出了使用抽样调查中样本容量计算公式代入相同降水代表年的测算数据进行典型样点灌区的选择。贾浩等<sup>[5]</sup>采用了首尾测算法测算了4个样点灌区的灌溉水利用系数,在采用主成分分析法对指标体系降维处理的基础上,用Copula函数构建PCA-Copula评价分析方法,对灌溉水利用系数各影响因素的影响程度和影响规律进行计算分析。本文对小型样点灌区的选取数量方法进行研究探讨,充分利用已有的灌溉水利用系数观测成果,分别采用数理统计中的简单随机抽样和配额抽样方法,验证或修改现有样点灌区数量。

## 1 样点灌区选取原则与方法

### 1.1 样点灌区选取原则

样点灌区的选取应当具有代表性,所选样点灌区应当将灌区现状、灌区管理结构、地形地貌、种植作物种类结构等因素考虑在内,能够代表区域内同类型的其他灌区。同时在选择的过程中应当综合考虑所选区域的灌溉方式、有无节水改造、灌溉面积等因素,尽量使所选的样点灌区能够反映区域灌溉灌区的现状。

### 1.2 样点灌区选取方法

采用抽样调查的方法选取样点灌区。抽样是与调查活动联系在一起,抽样调查是依据随机原则从总体中抽取部分单元作为样本,根据样本信息从而推断总体特征的一种现今最重要的调查方式。按照抽取样本方法的不同可以把抽样调查分为概率抽样和非概率抽样。本文分别选取概率抽样中的简单随机抽样方法和非概率抽样中的配额抽样方法来进行计算,并根据“样点灌区数量初步确定—样点灌区最低数量要求—样点灌区数量最终确定”3个阶段最终确定样点灌区数量。

#### 1.2.1 概率抽样

概率抽样又称随机抽样,是以概率论与数理统

计为基础,依据大数定律通过随机化的机械操作取得样本从而推算总体的性质和特征,并且能计算和控制抽样误差。概率抽样总体可分为:简单随机抽样、分层随机抽样、系统抽样、整群抽样。简单随机抽样顾名思义就是从总体中随机地抽取样本,在简单随机抽样中,每一个对象都有相等的机会被抽中。

#### 1.2.2 非概率抽样

非概率抽样又称非随机抽样,与概率抽样不同,非概率抽样是以抽样者的主观意愿为基础进行抽样,并不是严格按随机抽样原则来抽取样本,所以也失去了大数定律的存在基础从而无法控制误差。非概率抽样可分为方便抽样、雪球抽样、配额抽样等。

## 2 样点灌区数量初步确定

### 2.1 简单随机抽样

简单随机抽样是概率抽样的一种,其又可以分为重复抽样和非重复抽样。对于样点灌区的选取抽样应当采用非重复抽样方法。以目标区域内所有小型地表水源灌区为样本总体,通过确定总体规模、置信度、绝对误差限度和方差,并采用代表年度农田灌溉水有效利用系数测算小型灌区样点灌区为调查样本,估算确定样本方差,初步确定样点灌区数量。

在样本方差的计算过程中,以往的研究主要选取耕地灌溉单位面积用水量这一指标,本方案利用已选出的样点灌区数据,选取样点灌区的灌溉水利用系数,计算样本方差,从而得出样本容量来验证或修改已有的样点灌区数量。灌溉水利用系数,是评价灌区的工程状况和管理水平的一个重要指标。

根据抽样技术统计学原理<sup>[6]</sup>,以灌溉水利用系数估计的非重复性简单随机抽样样本容量公式为:

$$n = \frac{Nu_{a/2}^2 S^2}{Nd^2 + u_{a/2}^2 S^2} \quad (1)$$

$$S^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N-1} \quad (2)$$

式中: $N$ 为小型灌区样本数; $S^2$ 为方差; $x_i$ 为调查样本中第*i*个小型灌区灌溉水利用系数; $\mu$ 为调查样本中的平均小型灌区灌溉水利用系数; $d$ 为可接受的样本误差,在其他3个条件不变时,可接受的样本误差越小,所需的样本容量*n*越大; $u_{a/2}$ 为标准误

差的置信水平。置信度  $1-\alpha$  与置信水平  $u_{\alpha/2}$  的关系见表1。

表1 置信度与置信水平关系

$1-\alpha$	0.90	0.95	0.99
$\mu_{\alpha/2}$	1.645	1.96	2.58

## 2.2 配额抽样

配额抽样是非概率抽样的一种,在配额抽样中,总体会被分为不同的亚群,然后估计出总体中各亚群的比例,最后按照样本容量大小根据各亚群的比例从中抽样。依据配额抽样的样本作为参考的前提是要研究的总体特征量在总体中呈现均匀分布,也就是说所选样本与总体的结构一致是推断的前提。配额抽样样本容量的确定分为4个步骤:

(1)控制特征的选取。选取控制特征通常要考虑以下3个要素:研究的背景与目的、总体内部差异的程度、特征和要研究的目标变量之间的关联程度。

(2)依据总体中各控制特征所占比例,列出独立控制分配比例表。假定已选取的  $S$  个控制特征分别有  $m_1, m_2, \dots, m_s$  个控制水平,则可以得到  $m_1 + m_2 + \dots + m_s$  个控制特征比例,记第  $b$  个控制特征的第  $a$  个水平所占的比例为  $P_{ab}$ ,那么  $\sum_{a=1}^{m_b} P_{ab} = 1, b=1, 2, 3, \dots, S$ 。列出独立控制分配见表2。

表2 配额抽样独立控制比例分配

控制特征水平	1	2	...	$S$
1	$P_{11}$	$P_{12}$	...	$P_{1s}$
2	$P_{21}$	$P_{22}$	...	$P_{2s}$
3	$P_{31}$	$P_{32}$	...	$P_{3s}$
⋮	⋮	⋮	⋮	⋮
合计	1	1	1	1

(3)设计交叉控制的配额计算模型。

当配额抽样总的样本量  $n$  已知,样本量  $n$  在第  $1, 2, \dots, S$  个控制特征第一个水平上的配额就是  $n \times P_{11} \times P_{12} \times \dots \times P_{1s}$ ,理论上,任意一个这样的配额量都不应该为零,由此可得样本量计算模型如下:

$$\min_{1 \leq a_i \leq m_i} \left( n \times \prod_{b=1}^S P_{a_i b} \right) = 1 \quad (3)$$

很显然,当每个控制特征都取控制比例最小的那个控制水平进行组合得到的值就是最优解(满足条件的最小的样本量),即:

$$n = \frac{1}{\prod_{b=1}^S \min_{1 \leq a_i \leq m_i} \{P_{a_i b}\}} \quad (4)$$

(4)修正调整样本量。

依据上式计算出来的样本量  $n$  通常不是整数,但是抽取样本通常只能以完整的个体为单位,所以在利用上式计算出  $n$  后,还需要对样本量的非整数值进行调整和处理,同时需要兼顾最小样点灌区的要求。

## 3 案例分析

### 3.1 仪征市农业灌溉概况

仪征市位于长江三角洲西端,地跨长江、淮河两大流域,总面积  $859 \text{ km}^2$ 。全市属于宁镇杨丘陵区,地形地貌比较复杂,总体地势西北高,东南低。仪征市属亚热带沿江季风气候区,雨量充沛,光照资源丰富,四季分明,年平均气温  $16.0 \text{ }^\circ\text{C}$ 。根据全国第一次水利普查成果显示,仪征共有灌区 58 个,其中中型灌区 19 个,小型灌区 39 个,均属于提水灌溉灌区。各个灌区中,以塘坝为水源的灌区 13 个,以河道为水源的灌区 12 个,综合性水源灌区 33 个。2018 年农田总有效灌溉面积达到  $3.78 \text{ 万 hm}^2$ ,其中中型灌区农田灌溉面积  $3.04 \text{ 万 hm}^2$ ,小型灌区农田灌溉面积  $0.74 \text{ 万 hm}^2$ ,多采用提水灌溉,局部可自流灌溉。

### 3.2 仪征市小型样点灌区数量初步确定

#### 3.2.1 简单随机抽样选取样点灌区

以 2018 年仪征市农田灌溉水有效利用系数测算的小型地表水源样点灌区为调查样本数据,以小型灌区灌溉水利用系数为指标,各小型样点灌区灌溉水利用系数分别为 0.665、0.554、0.684、0.581、0.695、0.512,估算样本方差为 0.006,在方差  $S^2$  与总体数  $N$  确定后,所需样本容量  $n$  与绝对误差限度  $d$ 、标准误差的置信水平  $u_{\alpha/2}$  之间的关系如表 3 所示。

表3 样本容量与绝对误差限度、置信水平关系

$n$	$d$				
	0.10	0.15	0.20	0.25	0.30
$u_{\alpha/2}$					
1.645	1.56	0.71	0.40	0.26	0.18
1.960	2.18	1.00	0.57	0.37	0.25
2.580	3.60	1.70	0.97	0.63	0.44

对于样点灌区的选取,其样本容量的确定不用过于精确,因此置信度可以取低些,因此选取0.95的置信度 $1-\alpha$ ,0.1的绝对误差限度 $d$ ,得到的样本容量 $n$ 为2.18,修正后取整,综合考虑选取2个样本作为典型样点灌区。

3.2.2 配额抽样选取样点灌区

依据现有的统计资料,选取有效灌溉面积规模、灌溉水源类型、节水改造工程实施情况3个控制特征,计算各控制特征独立分配比例见表4所示。

根据现有的统计资料,仪征市小型灌区共有39个,总体样本数量较少,因此在抽选样本时在样本代表性方面做一些牺牲,放弃当每个控制特征都取控制比例最小的那个控制水平进行组合的考虑,找出3个特征分层中比例乘积最大的,计算样点灌区数量。控制特征比例最大的是有效灌溉面积2 000亩~5 000亩、灌溉水源类型为河道且无节水改造的灌区,依据表4以及配额计算模型可得:

$$n = \frac{1}{51.32\% \times 58.3\% \times 92.7\%} \approx 3.61$$

经过修正调整后,样本总数确定为4个。这样,在样本总数确定之后,由3个特征联合控制的样本配额也都可以计算出来了,从而完成交叉控制的样本配额的计算,如表5所示。

3.3 仪征市小型样点灌区数量最终确定

样点灌区个数应根据本省(区、市)小型灌区(或小型水利工程控制的灌溉区域)的实际情况确定;同时,样点灌区应包括提水和自流引水两种水源类型,不同水源类型的样点灌区个数应与该类型灌区数量所占的比例相协调。仪征市小型灌区均为提水灌溉灌区,综合考虑简单随机抽样方法与配额抽样的计算结果,通过上述方法确定,仪征市2018年共需布设4个小型地表水源样点灌区,而2018年度仪征市农田灌溉水有效利用系数测算布设的小型地表水源样点灌区为6个,通过对计算确定的样点灌区数量与现有2018年度农田灌溉水有效利用系数测算布设的样点灌区进行比较分析,可以发现:

(1)总体来说,计算确定的小型灌区样点灌区数量与现有的2018年度仪征市农田灌溉水有效利用系数测算布设的小型灌区样点灌区数量相差不大,现有系数测算样点灌区只需适当调整即可满足农田灌溉用水量的统计要求。

(2)现有的仪征市小型灌区样点灌区数量为6个,与现有样点灌区相比,需要减少的仪征市小型灌区样点灌区数为2个。

表4 小型灌区控制特征分类比例

有效灌溉面积规模/hm <sup>2</sup>	比例/%	灌溉水源类型	比例/%	有无节水改造工程	比例/%
< 133.3	35.86	塘坝	10.9	有节水改造	92.7
133.3~333.3	51.32	河道	58.3	无节水改造	7.3
> 333.3	12.82	综合性水源	30.8		

表5 样本配额实际分配

单位:hm<sup>2</sup>

	有节水改造			无节水改造		
	< 133.3	133.3~333.3	> 333.3	< 133.3	133.3~333.3	> 333.3
塘坝	0	0	0	0	0	0
河道	1	2	0	0	0	0
综合性水源	0	1	0	0	0	0

4 结 语

本研究遵循样点灌区选取原则,提出一种农田用水量统计工作中样点灌区的选取数量确定方法,方法采用简单随机抽样与配额抽样相结合。为以后农田灌溉水有效利用系数测算工作过程中样点

灌区的选取数量问题的解决提供了参考。主要结论有:(1)采用简单随机抽样中非重复性抽样的样本容量计算公式计算的方法,以及选取控制特征建立配额抽样样本量计算模型的方法,两种方法综合确定小型灌区样点灌区选取数量;(2)根据上述方

(下转第63页)

挥,并影响大坝安全与下游安全,需及早采取必要工程措施处理,建议采取如下处置对策:

(1)探测云图虽表明坝基局部存在渗漏薄弱部位,但运行过程中并未暴露坝基渗漏问题,当前渗漏处置可以左岸接触渗流与坝体渗漏为主。

(2)建议采用垂直防渗措施对全坝段坝体进行防渗处理,防渗体轴线布置在坝顶,其中左侧坝段(桩号0+000~0+040)防渗体深入坝基长度不宜小于1.0 m,其他坝段防渗体底高程控制在29.00 m。贴坡反滤措施可视垂直防渗处理后的效果选用。

(3)增设必要的雨水情测报和大坝安全监测设施如水尺、测压管,监控大坝渗流安全状况,指导工程安全运行。

(4)及时清理坝坡杂草,疏通坝脚导渗沟,改善工程形象。

(5)在消除隐患前,控制水位在30.90 m以下运行,并加强巡视检查。

## 4 结 语

本文采用高密度电法对桃园凹塘坝渗漏进行探测,通过对探测成果进行模型反演和后处理分析。

(1)坝体左岸和中部存在“低阻异常区”,经分析判断,左岸山体以及与坝体接触部位、坝体中部为大坝渗流较为薄弱部位,左坝肩与坝体结合部位存在接触渗漏,探测分析结果与现场检查结论一致。

(2)高密度电法为大坝渗漏探测提供了较为丰富的地电断面信息,与现场检查信息相互补充、互为印证,可以对除险加固效果进行有效评价,并为渗漏处置对策的确立提供依据。

(3)根据现场检查和探测分析结果,提出了采用垂直防渗措施对全坝段坝体进行防渗处理,增设必要的雨水情测报和大坝安全监测设施,疏通坝脚导渗沟,隐患消除前控制低水位运行并加强巡视检查等相应的渗漏处置对策,为充分发挥工程的社会及经济效益提供了决策支持。

### 参考文献:

- [1] 严加永,孟贵祥,吕庆田,等. 高密度电法的进展与展望[J]. 物探与化探,2012,36(4):576-584.
- [2] 中华人民共和国水利部. SL 326—2005 水利水电工程物探规程[S]. 北京:中国水利水电出版社,2005.
- [3] 皮雷,谭磊,李波. 综合物探方法在水库绕坝渗漏隐患探测中的应用[J]. 中国农村水利水电,2021(1):6.
- [4] 严飞. 高密度电法在深切峡谷大跨径桥址勘察中的应用研究[J]. 中国农村水利水电,2013(10):75-77.
- [5] JIN L, QI Z A, LIANG L C, et al. Monitoring and characterizing the deformation of an earth dam in Guangxi Province, China[J]. Engineering Geology, 2019, 248: 50-60.
- [6] 王志鹏,刘江平,王鸣谦,等. 高密度电法对带破碎状溶壳溶洞探测数值模拟[J]. 工程地球物理学报,2019,16(2):193-202.

(上接第59页)

法,2018年仪征市小型灌区共需布设4个样点灌区,与现有的样点灌区数量相比,需要减少的样点灌区数量为2个;(3)考虑灌区动态变化情况,对样点灌区及时进行调整。

### 参考文献:

- [1] 沈莹莹,张绍强,吉晔. 我国农业灌溉用水量统计方法的确定及工作开展情况[J]. 中国农村水利水电,2016(11):133-134,138.
- [2] 张建华. 基于样点灌区选取方法的农田灌溉用水量统计分析[J]. 黑龙江水利科技,2021,49(10):113-115.

- [3] 沈莹莹,陈梦婷,崔静,等. 农田灌溉用水量统计工作中的样点灌区选取方法研究[J]. 中国农村水利水电,2020(4):112-115.
- [4] 李泽坤,关迦文,杨波. 农田灌溉水利用系数测算中小型样点灌区选择方法研究[J]. 内蒙古水利,2021(8):25-27.
- [5] 贾浩,王振华. 基于主成分分析和Copula函数的灌溉水利用系数影响因素研究——以新疆建设兵团第十二师中型灌区为例[J]. 干旱地区农业研究,2020,38(6):167-175,233.
- [6] 邵志强. 抽样调查中样本容量的确定方法[J]. 统计与决策,2012(22):12-14.