

doi:10.6041/j.issn.1000-1298.2022.03.036

土壤 Hg 和 NDVI 全局均值估计的样本点权重调整方法

张东悦¹ 朱庆伟¹ 董士伟^{2,3} 潘瑜春^{2,3} 吴亚楠^{2,3} 谭梦岩⁴(1. 西安科技大学测绘科学与技术学院, 西安 710054; 2. 北京市农林科学院信息技术研究中心, 北京 100097;
3. 国家农业信息化工程技术研究中心, 北京 100097; 4. 肇东市农业技术推广中心, 肇东 151100)

摘要: 样本点权重调整是统计推断样本点属性全局均值的重要手段。以北京市顺义区农业用地样本点为例, 提出了一种面向全局均值估计的样本点权重调整方法——泰森多边形权重调整法。首先构建样本点类型划分规则, 将样本点划分为聚集样本点、稀疏样本点和均匀样本点; 其次针对不同类型样本点, 分别确定样本点权重调整量和调整规则并进行样本点权重调整; 最后以农业用地土壤重金属 Hg 含量及 NDVI (归一化植被指数) 数据为例, 设置与原始样本点 (无权重调整)、常规面积比例权重调整法的对比实验, 通过 2 种评价参数的全局均值相对误差分别评价不同样本点权重调整方法的数据纠偏效果。结果表明: 2 个聚集样本点权重减小, 权重分别为 0.609 和 0.883; 1 个稀疏样本点权重增大, 权重为 1.068; 63 个均匀样本点权重不变, 权重为 1; 泰森多边形权重调整法统计推断农业用地土壤重金属 Hg 含量和 NDVI 的全局均值估计相对误差 (0.413%、1.617%) 小于原始样本点 (1.056%、2.500%) 和常规面积比例权重调整法 (2.933%、2.941%), 表明该权重调整方法的样本点全局均值估计更加准确、可靠。

关键词: 样本点; 全局均值估计; 泰森多边形; 数据纠偏; 统计推断; 权重调整

中图分类号: X53

文献标识码: A

文章编号: 1000-1298(2022)03-0338-08

OSID:



Weight Adjustment Method of Sampling Sites for Global Mean Estimation of Hg and NDVI

ZHANG Dongyue¹ ZHU Qingwei¹ DONG Shiwei^{2,3} PAN Yuchun^{2,3} WU Ya'nian^{2,3} TAN Mengyan⁴

(1. College of Geomatics, Xi'an University of Science and Technology, Xi'an 710054, China

2. Research Center of Information Technology, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China

3. National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China

4. Zhaodong Agricultural Technology Extension Center, Zhaodong 151100, China)

Abstract: Weight adjustment of the sampling sites is an important means of the statistical inference for the global mean of sampling sites attributes. Taking the sampling sites of agricultural land in Shunyi District, Beijing as an example, weight adjustment method (Thiessen polygon weight adjustment method) of sampling sites for global mean estimation was proposed. Firstly, the type division rules of sampling sites were constructed to divide the sampling sites into aggregated, sparse and even sampling sites. Secondly, weight adjustment amounts and adjustment rules for different types of sampling sites were determined, and the weights of the sampling sites were adjusted, respectively. Finally, taking soil heavy metal Hg content and NDVI data of agricultural land as an example, a comparative experiment between Thiessen polygon weight adjustment method and the original sampling sites (without weight adjustment) and the conventional area ratio weight adjustment method was set up to quantitatively evaluate the global mean estimation effect of different weight adjustment methods based on the relative errors of two evaluation parameters. The results suggested that the weights of the two aggregated sampling sites were reduced, and the weights were 0.609 and 0.883, respectively. The weight of one sparse sampling site was increased, and the weights was 1.068. The weights of sixty-three even sampling sites remained unchanged, and the weights of each sampling sites was 1. The relative errors (0.413%, 1.617%) of the global mean

收稿日期: 2021-12-03 修回日期: 2022-01-05

基金项目: 国家重点研发计划项目(2021YFD1500203)、北京市自然科学基金项目(8192015)和国家自然科学基金项目(41801276)

作者简介: 张东悦(1997—),女,研究实习员,主要从事样本点时空数据分析研究,E-mail: 485539936@qq.com

通信作者: 董士伟(1984—),男,高级工程师,主要从事时空数据分析研究,E-mail: dshiwei2006@163.com

estimated by the Thiessen polygon weight adjustment method statistical inference were smaller than those of estimated by the original sampling sites (1.056%, 2.500%) and the conventional area proportional weight adjustment method (2.933%, 2.941%), respectively. It was shown that the global mean estimation of the sampling sites for the Thiessen polygon weight adjustment method was much more accurate and reliable.

Key words: sampling sites; global mean estimation; Thiessen polygon; data correction; statistical inference; weight adjustment

0 引言

我国工业化进程的快速发展,导致土壤重金属在污染水平、污染范围等方面面临日益严重的趋势。为了应对这种趋势,国家采取相关措施,如《土壤污染防治行动计划》(土十条)等开展土壤污染调查,农用地分类管理,加强农业污染源监管。不同时期土壤整体污染情况可以基于长期定位监测样本点进行预测与评估^[1-3],而样本点空间分布的均匀性和代表性是土壤整体污染情况监测准确性的重要指标。目前野外作业采集的样本点数据大多数存在样本点均匀性和代表性差的问题,导致样本点全局均值估计数据有偏,因此有针对性地提出样本点权重调整方法进行相应样本点数据纠偏非常关键。

空间采样是监测土壤整体污染情况及空间信息变化的重要手段,而统计推断是采样点布局优化的目的。样本点全局均值估计是统计推断的重要组成部分,其目的是准确估计研究区内某一数据的整体质量水平^[4]。通过土壤样本点重金属全局均值估计与土壤重金属背景值进行对比分析,可以判断土壤是否受到污染^[5]及污染程度^[6-7]。然而,样本点全局均值估计的准确性受样本点密度及分布合理性影响,样本点密度越大且分布越均匀其全局均值估计越准确^[8]。由于受研究区范围、抽样成本、地形等因素限制,难以获得高密度代表性样本点^[9-12],样本点空间分布易产生地理空间分布不均匀和特征空间代表性差的问题^[13-14],导致样本点全局均值估计值与全局均值真实值之间产生偏差。样本点数据纠偏主要包括样本点增删和样本点权重调整。样本点增删主要基于样本点均匀性检测结果,对密集分布样本点进行删除处理,对稀疏分布样本点进行加密处理^[15-16]。样本点增删可以有效提高样本点空间分布的均匀性和代表性,但由于样本点加密过程中需考虑加密样本点的土壤属性值,因此对加密样本点与原始样本点之间的时间间隔有很高的要求。样本点权重调整不需要加密或删除样本点,只需调整样本点属性值权重,实现对样本点数据纠偏。样本点权重调整方法主要包括层次分析法^[17-18]、专家打分法、格网法^[19]、自组织双分区法^[20-21]以及分层

权重调整法^[9]等。综合分析现有样本点权重调整方法后发现主要存在两大问题:①格网法、自组织双分区法直接对全部样本点进行权重调整。这种样本点权重调整模式无法确保大部分样本点属性值的真实值,降低样本点全局均值估计的可信度和准确性。②层次分析法、专家打分法所构建的样本点权重调整规则通常是静态的,即样本点权重不随评价因子状态变化而变化^[22],这将造成样本点在全局均值估计过程中产生偏差。综上所述,当前研究主要集中在对全部样本点通过统一的权重调整规则进行权重调整,而通过划分样本点类型确定待调整样本点数量并根据待调整样本点类型构建相应权重调整规则的方法亟待研究。

样本点权重调整的数据纠偏效果评价通常采用研究区大数据量的样本点某一属性值数据模拟估计的参考值作为真实值,或借助可获得真实值的属性值数据进行评价。研究区采样点数量越多,估计的参考值越准确^[2,23-24];由于受采样成本的限制,野外作业只能获取有限的样本点,难以估计获取准确的样本点土壤重金属全局均值,可以借助条件模拟方法进行模拟估计。归一化植被指数(NDVI)全局均值真实值通过研究区内遥感影像计算获取,也可以作为样本点权重调整的数据纠偏效果评价的参数。因此,本文以北京市顺义区农业用地为例,构建样本点类型划分规则以实现样本点类型划分,对非均匀样本点进行变权调整,基于土壤重金属Hg含量和NDVI 2个参数,分别设置与原始样本点、常规面积比例权重调整法的对比实验,并利用全局均值估计相对误差定量评价样本点权重调整的数据纠偏效果,旨在改善样本点空间分布的均匀性和代表性,进而提高样本点统计推断精度。

1 研究区概况与数据来源

1.1 研究区概况

顺义区位于北京市东北部,地理坐标为东经116°28'~116°58',北纬40°00'~40°18',境域东西长45 km,南北宽30 km,总面积为1 020 km²。其中农业用地面积为494.34 km²,是北京市重要的粮食和蔬菜生产基地。顺义区气候属暖温带半湿润大陆

性季风性气候,年平均气温为 11.5℃,年均降雨量约 625 mm。土壤为河流洪水携带沉积物质造成,表面堆积物主要是砂土、亚砂土,北部山地最高点海拔为 637 m,境内最低点海拔为 24 m。选取高丽营镇、赵全营镇、牛栏山镇和北石槽镇的农业用地为研究区域,其中农用地主要是菜地、水浇地、苗圃等,农业用地面积为 114.379 km²。由于该区域农业用地所占面积较大,对该区域农业用地土壤重金属污染状况进行监测有利于更好地了解顺义区土壤质量和农产品质量。

1.2 数据来源及预处理

采用 2017 年北京市顺义区样本点数据和 NDVI 数据。

(1) 样本点数据:样本点数据来源于北京市农林科学院农产品质量安全管理平台。该平台样本点数据采集工作于 2007 年开展,样本点的初始数量为 125 个,并且根据土地利用类型及面积的变化,定期对样本点进行增删处理。本研究样本点数据采集于 2017 年,样本点布局 and 数量根据田块的土地利用方式和面积确定,采样时利用 GPS 获得样本点的地理坐标。原始采样中,每个样本点在 10 m × 10 m 正方形网格 4 个顶点和中心点处选择 0 ~ 20 cm 耕层土壤混合,按四分法选取分析样品 1.0 kg,最终得到 66 个农业用地样本点。所有土样在室内自然风干,碾压磨碎后,过 100 目尼龙网筛,按照国家标准分析测定土壤重金属 Hg 含量。

(2) NDVI 数据:基于欧洲空间局空间分辨率为 10 m 的 Sentinel-2 卫星数据计算获取。影像获取时间为 2017 年 9 月 5 日。首先利用欧洲空间局发布的针对哨兵卫星数据的 Sen2Cor 软件对其进行辐射定标及大气校正,然后利用红波段(Sentinel-2A 的第 4 波段)和近红外波段(Sentinel-2A 的第 8 波段)的反射率数据计算顺义区 NDVI^[25],并利用 ENVI 快速统计工具得到顺义区 NDVI 全局均值。

研究区内的样本点主要是作为样本点权重调整的实验数据,空间分布如图 1 所示。由于研究区内所布设的 66 个样本点无法准确估计土壤重金属全局均值的真实值,因此选择土壤重金属 Hg 的序贯高斯条件模拟全局均值与 NDVI 全局均值作为参数,分别用于定量评价样本点权重调整的数据纠偏效果。土壤重金属 Hg 含量侧重反映难以获取全局均值真实值的数据情景,NDVI 侧重反映可以获取全局均值真实值的数据情景。2 个参数的设置能够更好地反映样本点权重调整效果评价的客观性和科学性。

基于 ArcGIS 10.5 软件实现遥感影像裁剪、泰

森多边形计算和序贯高斯条件模拟;基于 ENVI 5.3 软件实现 NDVI 提取及相应计算。

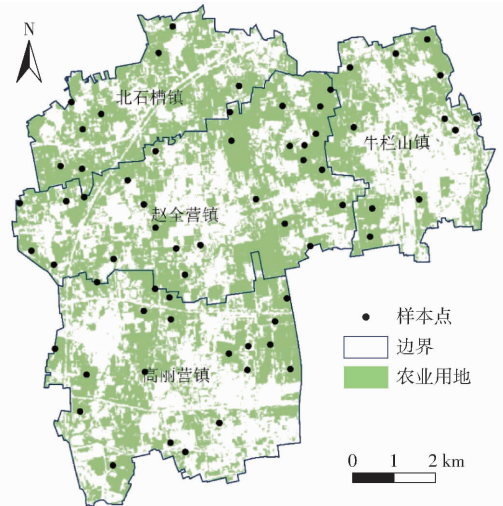


图 1 农业用地样本点空间分布

Fig. 1 Spatial distribution of sampling sites of agricultural land

2 研究方法

技术流程如图 2 所示,3 种权重调整方法包括泰森多边形权重调整法、原始样本点(无权重调整)、常规面积比例权重调整法。基于土壤重金属 Hg 含量和 NDVI 2 个参数,分别通过样本点全局均值估计相对误差对不同样本点权重调整的数据纠偏效果进行评价。其中常规面积比例权重调整法没有划分样本点类型,主要是将样本点所在泰森多边形面积与样本点泰森多边形平均面积的比值作为该样本点的调整权重,实现样本点权重调整。

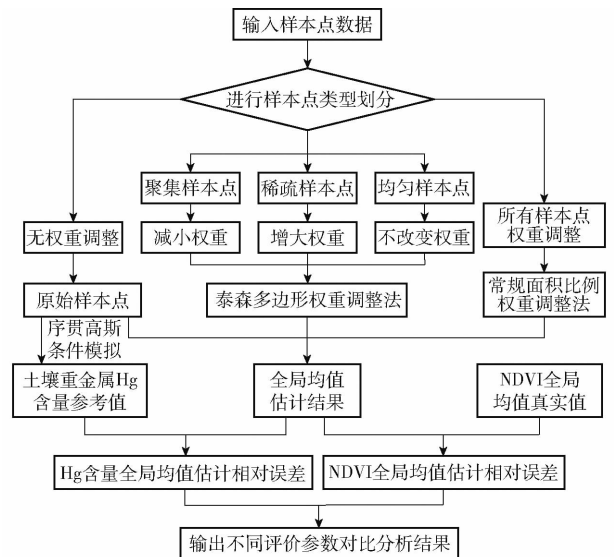


图 2 技术路线图

Fig. 2 Technical process

森多边形权重调整法主要步骤分为:①样本点类型划分:构建样本点泰森多边形,根据样本点所

在泰森多边形面积确定面积置信区间,实现样本点类型划分及不同类型样本点空间位置确定。②样本点权重调整:分别构建不同类型样本点对应的权重调整规则,实现研究区内非均匀样本点变权调整。③样本点权重调整的数据纠偏效果评价:通过全局均值估计相对误差对样本点权重调整的数据纠偏效果进行评价。

2.1 样本点类型划分规则

样本点类型划分是确定非均匀样本点位置及数量的关键,合理的样本点类型划分规则可以提高样本点全局均值估计的准确性^[25]。通过面积置信区间来划分样本点的类型,将泰森多边形面积小于面积置信区间下限的样本点划分为聚集样本点,将泰森多边形面积大于面积置信区间上限的样本点划分为稀疏样本点,将泰森多边形面积在面积置信区间之间的样本点划分为均匀样本点。

(1)通过样本点平均采样面积划分样本点所在泰森多边形类型。首先构建样本点泰森多边形,将样本点泰森多边形总面积与总采样点数的比值确定为样本点平均采样面积 S_0 ;将样本点所在泰森多边形面积按从小到大顺序排列,当样本点所在泰森多边形面积小于 S_0 时,则划分为I类泰森多边形(V_1),反之则划分为II类泰森多边形(V_2)。

(2)根据划分的两种类型泰森多边形面积确定面积置信区间。计算 V_1 的面积均值 μ_1 及面积标准差 σ_1 来确定面积置信区间下限,计算 V_2 的面积均值 μ_2 及面积标准差 σ_2 来确定面积置信区间上限,面积置信区间计算式为

$$\begin{cases} \hat{\theta}_1 = \mu_1 - 2\sigma_1 \\ \hat{\theta}_2 = \mu_2 + 2\sigma_2 \end{cases} \quad (1)$$

式中 $\hat{\theta}_1$ ——面积置信区间下限,km²

$\hat{\theta}_2$ ——面积置信区间上限,km²

(3)样本点类型划分。当样本点所在泰森多边形面积小于 $\hat{\theta}_1$ 定义为聚集样本点;当样本点所在泰森多边形面积大于 $\hat{\theta}_2$ 定义为稀疏样本点;当样本点所在泰森多边形面积在 $[\hat{\theta}_1, \hat{\theta}_2]$ 之间定义为均匀样本点。

2.2 样本点权重调整规则

样本点权重调整规则基于划分样本点类型确定,聚集样本点减小权重,稀疏样本点增大权重,均匀样本点无需进行权重调整。

(1)聚集样本点权重调整

在特定区域范围内,由于聚集样本点比理想均匀分布的样本点数量多,应该减小单个样本点的权重。聚集样本点权重调整公式为

$$W_i = 1 - \Delta W_i \quad (2)$$

$$\text{其中} \quad \Delta W_i = \frac{|S_i - \hat{\theta}_1| |S_i - S_0|}{\hat{\theta}_1 S_0} \quad (3)$$

式中 ΔW_i ——第*i*个聚集样本点权重调整量

S_i ——第*i*个聚集样本点所在泰森多边形面积,km²

(2)稀疏样本点权重调整

在特定区域范围内,由于稀疏样本点比理想均匀分布的样本点数量少,应该增大单个样本点的权重。稀疏样本点权重调整公式为

$$W_j = 1 + \Delta W_j \quad (4)$$

$$\text{其中} \quad \Delta W_j = \frac{|S_j - \hat{\theta}_2| |S_j - S_0|}{\hat{\theta}_2 S_0} \quad (5)$$

式中 ΔW_j ——第*j*个稀疏样本点权重调整量

S_j ——第*j*个稀疏样本点所在泰森多边形面积,km²

(3)均匀样本点权重调整

均匀样本点不影响全局均值估计的准确性,则无需对这类样本点进行权重调整,对应的调整权重为1。

2.3 样本点权重调整的数据纠偏效果评价

利用全局均值估计相对误差来对样本点权重调整的数据纠偏效果进行评价。全局均值估计相对误差表示样本点全局均值估计值与全局均值真实值之间的偏离程度,计算公式为

$$E_g = \frac{|M_r - M_e|}{M_r} \times 100\% \quad (6)$$

式中 E_g ——全局均值估计相对误差,%

M_r ——全局均值真实值

M_e ——全局均值估计值

当计算得到的全局均值估计相对误差越大时,全局均值估计值与全局均值真实值越偏离;当计算得到的全局均值估计相对误差越小时,全局均值估计值与全局均值真实值越接近。

2.4 序贯高斯条件模拟

序贯高斯条件模拟(Sequential Gaussian conditional simulation, SGCS)是一种对连续变量进行随机模拟的地统计学方法,该模拟方法结果在现有条件样本点处的模拟值为该点实测的真实值^[26]。SGCS将现有数据作为条件数据计算待模拟点值的条件概率分布,从该分布中随机抽取一个值作为模拟值,连同原始数据、模拟数据一起作为条件数据代入下一点的模拟。该模型能产生多个可能的空间分布格局,有效避免普通克里格插值结果的平滑效应^[27]。SGCS将研究数据作为一个整体,最大程度模拟研究区土壤重金属真实空间分布,模拟的结果

与实测数据的直方图相同,并且模拟结果与实测结果具有相同的空间自相关关系,因此 SGCS 可以实现研究区土壤重金属含量模拟。

3 结果与分析

3.1 样本点类型划分结果

以北京市顺义区农业用地样本点为例,构建农业用地样本点泰森多边形并计算泰森多边形面积,进行农业用地样本点类型划分,结果如图 3 所示。图中样本点所在泰森多边形面积按照从小到大顺序排列。通过样本点划分规则判断存在 2 个聚集样本点,所在泰森多边形面积分别为 0.265、0.423 km²; 1 个稀疏样本点,所在泰森多边形面积为 3.439 km²; 63 个均匀样本点,所在泰森多边形面积在 0.653 ~ 3.109 km²之间。不同类型样本点的空间分布如图 4 所示,图中 2 个聚集样本点分别是聚集样本点 A 和聚集样本点 B。

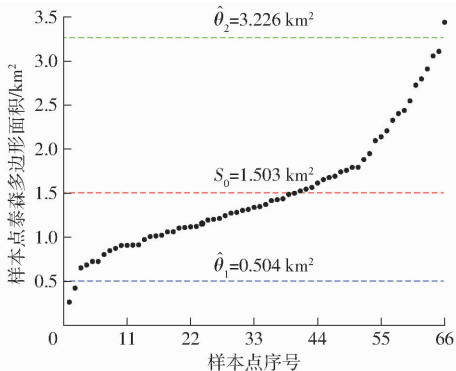


图 3 农业用地样本点类型划分结果

Fig.3 Classification of sampling sites of agricultural land

表 1 不同权重调整方法的样本点权重调整结果

Tab.1 Weight adjustment results of sampling sites for different weight adjustment methods

权重调整方法	样本点	样本点数量	泰森多边形面积/km ²	原始权重	权重调整量	调整权重
泰森多边形权重调整法	聚集样本点 A	1	0.264	1	0.391	0.609
	聚集样本点 B	1	0.423	1	0.117	0.883
	稀疏样本点	1	3.439	1	0.068	1.068
	均匀样本点	63	0.653 ~ 3.109	1	0	1
常规面积比例权重调整法	所有样本点	66	0.264 ~ 3.439	1	-0.824 ~ 1.289	0.176 ~ 2.289

别设置与原始样本点及常规面积比例权重调整法的对比实验,计算不同方法对于不同评价参数的样本点全局均值估计相对误差,定量评估泰森多边形权重调整法数据纠偏效果。

3.3.1 土壤重金属 Hg 全局均值相对误差

基于泰森多边形权重调整法、原始样本点和常规面积比例权重调整法计算样本点的调整权重,顺义区农业用地土壤重金属 Hg 含量的全局均值估计结果分别为 0.069 9、0.070 9、0.068 1 mg/kg。基于原始土壤重金属 Hg 含量数据序贯高斯条件模拟得到顺义区农

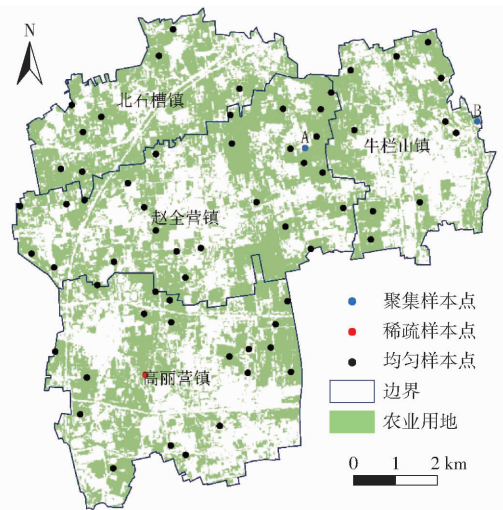


图 4 不同类型样本点的空间分布

Fig.4 Spatial distribution of different types of sampling sites

3.2 样本点权重调整结果

利用泰森多边形权重调整法实现非均匀样本点权重调整,通过式(3)计算得到聚集样本点 A 和聚集样本点 B 权重调整量分别为 0.391 和 0.117,通过式(2)计算得到对应的调整权重分别为 0.609 和 0.883;通过式(4)、(5)分别计算得到稀疏样本点权重调整量为 0.068,调整权重为 1.068;对于所有均匀样本点不需进行权重调整,调整权重均为 1。利用常规面积比例权重调整法计算得到全部样本点的调整权重,其对应范围为 0.176 ~ 2.289。不同权重调整方法的样本点权重调整结果如表 1 所示。

3.3 数据纠偏效果评价与对比分析

基于土壤重金属 Hg 含量和 NDVI 2 个参数,分

用地土壤重金属 Hg 含量全局均值参考值。首先,基于 ArcGIS 10.5 软件的探索数据工具探索最优变换方式及最优参数设置,通过地统计向导工具得到简单克里金插值结果;其次,基于插值结果进行序贯高斯条件模拟,为了避免由于模拟随机性带来的模拟误差,实验分别进行了 10、15、20、25、30、35、40、45、50、55、60 等不同次数模拟,直至模拟均值结果趋于平稳,获取最适的条件模拟次数为 50;最后,对样本点进行 50 次序贯高斯条件模拟得到土壤重金属 Hg 含量全局均值为 0.070 2 mg/kg。通过式(6)计算得到泰森多边形

权重调整法、原始样本点以及常规面积比例权重法的土壤重金属 Hg 含量的全局均值相对误差分别为 0.413%、1.056% 和 2.933%, 如表 2 所示。结果表

明, 泰森多边形权重调整法具有较好的数据纠偏效果, 可以有效地提高土壤重金属 Hg 含量全局均值估计结果的准确性。

表 2 不同方法样本点重金属 Hg 含量全局均值估计相对误差

Tab. 2 Relative errors of Hg mean estimation of sampling sites for different methods

权重调整方法	样本点数量	调整样本点数	估计值/(mg·kg ⁻¹)	真实值/(mg·kg ⁻¹)	相对误差/%
泰森多边形权重调整法	66	3	0.069 9	0.070 2	0.413
原始样本点	66	0	0.070 9	0.070 2	1.056
常规面积比例权重调整法	66	66	0.068 1	0.070 2	2.933

3.3.2 农业用地 NDVI 全局均值相对误差

基于泰森多边形权重调整法、原始样本点和常规面积比例权重调整法计算样本点的调整权重, 顺义区农业用地 NDVI 的样本点全局均值估计结果分别为 0.691、0.697 和 0.700。基于研究区农用地全部 996 148 个像元计算获得 NDVI 全局均值为 0.680, 以此作为全局均值真实值。通过式(6)计算得到泰森多边形法、原始样本点和常规面积比例权重法的 NDVI 全局均值相对误差分别为 1.617%、2.500% 及 2.941%, 如表 3 所示。结果表明, 与原始样本点和常规面积比例权重法相比, 泰森多边形权重调整法 NDVI 全局均值相对误差最小, 说明该方法对样本点的数据纠偏效果最好。

表 3 不同方法样本点 NDVI 全局均值估计相对误差

Tab. 3 Relative errors of NDVI mean estimation of sampling sites for different methods

权重调整方法	样本点数量	调整样本点数	估计值	真实值	相对误差/%
泰森多边形权重调整法	66	3	0.691	0.680	1.617
原始样本点	66	0	0.697	0.680	2.500
常规面积比例权重调整法	66	66	0.700	0.680	2.941

由表 2、3 可以看出, 泰森多边形权重调整法对于不同类型的评价参数均能有效地进行数据纠偏, 可以提高样本点在全局均值估计的代表性, 使其统计推断结果更加准确。分析其原因可能包括: ①泰森多边形权重调整法仅对非均匀样本点进行权重调整, 能确保大部分样本点属性值的真实性, 使样本点全局均值估计结果更加可靠。②泰森多边形权重调整法构建不同权重调整规则可实现对非均匀样本点变权调整, 针对性和目标性更强, 调整后的样本点更具代表性, 进而提高样本点全局均值估计的准确性。

4 讨论

泰森多边形权重调整法是一种面向全局均值统计推断的样本点数据纠偏方法, 通过对样本点类型划分来确定不同类型样本点数量及其空间位置, 并

基于样本点所在泰森多边形面积构建不同类型样本点权重调整规则及其权重调整量, 进而实现对聚集样本点和稀疏样本点的变权调整。该方法能够确保大部分样本点属性值的真实性, 并且对非均匀样本点变权调整, 提高样本点的代表性, 最终获得更加准确、可靠的全局均值估计结果。与常规样本点权重调整方法相比, 设置的土壤重金属 Hg 含量及 NDVI 数据纠偏效果评价对比实验, 均可验证泰森多边形权重调整法对全局均值估计的样本点数据纠偏效果更好。面向全局均值估计的样本点权重调整方法包括两个具体的应用场景: ①用于国家及各级政府监测管理所属管辖行政单元的土壤重金属整体污染评价, 辅助宏观政策决策与管理。②应用于某个行政单元内不同空间分区的土壤重金属污染评价, 辅助区域土壤重金属污染治理与防治。然而, 该方法在研究过程中存在一定的不确定性。

(1) 不同类型样本点数量划分存在不确定性。不同类型样本点数量通过面积置信区间确定, 在不同置信水平下, 面积置信区间的范围会随之发生变化, 导致不同类型样本点的数量也会随之改变, 权重调整样本点数量较多时会对样本点数据纠偏效果产生影响。

(2) 尽管序贯高斯条件模拟方法能够获取比较准确的土壤重金属含量模拟参考值^[22-23], 但是参考值不一定等于真实值, 未来可以通过高密度代表性采样来进一步模拟逼近真实值的土壤重金属元素全局均值。

(3) 泰森多边形权重调整法调整样本点并未改变样本点空间位置, 而是通过改变样本点某个属性值的权重实现对样本点的数据纠偏, 这仅仅是通过对样本点特征空间来进行调整, 但并未对样本点的地理空间进行调整, 未来将进一步研究耦合地理空间和特征空间的样本点权重调整方法, 以兼顾样本点在地理空间的均匀性和特征空间的代表性。

5 结论

(1) 泰森多边形权重调整法划分得到 2 个聚集

样本点、1 个稀疏样本点和 63 个均匀样本点。2 个聚集样本点调整权重分别为 0.609 和 0.883, 稀疏样本点调整权重为 1.068, 均匀样本点调整权重均为 1。

(2) 泰森多边形权重调整法、原始样本点、常规面积比例权重调整法进行顺义区农业用地土壤重金属 Hg 含量和 NDVI 的全局均值估计相对误差分别为 0.413%、1.056%、2.933% 和 1.617%、2.500%、2.941%。泰森多边形权重调整法对于不同评价参

数的全局均值估计相对误差均为最小, 与全局均值真实值更为接近, 表明泰森多边形权重调整法数据纠偏效果更好。

(3) 泰森多边形权重调整法准确定位到非均匀样本点位置并实现非均匀样本点的变权调整, 可以提高样本点的代表性, 更加准确地进行全局均值估计。但该方法仅在特征空间对样本点进行权重调整, 且在面积置信区间确定、序贯高斯条件模拟土壤重金属含量等方面存在一定的不确定性。

参 考 文 献

- [1] MCEWING R K, FISHER P J, ZONA D. Environmental and vegetation controls on the spatial variability of CH₄ emission from wet-sedge and tussock tundra ecosystems in the Arctic[J]. *Plant and Soil*, 2015, 388(1/2): 37–52.
- [2] 叶回春, 黄珊瑜, 张世文, 等. 土壤有机碳空间变异性对采样密度的响应研究[J]. *农业机械学报*, 2014, 45(12): 215–223.
YE Huichun, HUANG Shanyu, ZHANG Shiwen, et al. Spatial variability response of soil organic carbon to sampling density change[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2014, 45(12): 215–223. (in Chinese)
- [3] NELSON P N, BANABAS M, GOODRICK I, et al. Soil sampling in oil palm plantations: a practical design that accounts for lateral variability at the tree scale[J]. *Plant and Soil*, 2015, 394(1/2): 421–429.
- [4] 高秉博, 郝朝展, 李发东, 等. 面向土壤环境质量等级划分的统计推断与加密采样优化方法研究综述[J]. *农业环境科学学报*, 2021, 40(4): 712–722.
GAO Bingbo, HAO Chaozhan, LI Fadong, et al. Statistical inference and the additional sampling optimization method for soil environmental quality grade classification: a review[J]. *Journal of Agro-Environment Science*, 2021, 40(4): 712–722. (in Chinese)
- [5] 徐仲雨, 周春财, 孙浩, 等. 新庄孜矿区土壤重金属分布与来源分析[J]. *中国煤炭地质*, 2017, 29(9): 41–44.
XU Zhongyu, ZHOU Chuncai, SUN Hao, et al. Soil heavy metal distribution and source analysis in Xinzhuangzi coalmine area[J]. *Coal Geology of China*, 2017, 29(9): 41–44. (in Chinese)
- [6] 齐杏杏, 高秉博, 潘瑜春, 等. 基于地理探测器的土壤重金属污染影响因素分析[J]. *农业环境科学学报*, 2019, 38(11): 2476–2486.
QI Xingxing, GAO Bingbo, PAN Yuchun, et al. Influence factor analysis of heavy metal pollution in large-scale soil based on the geographical detector[J]. *Journal of Agro-Environment Science*, 2019, 38(11): 2476–2486. (in Chinese)
- [7] 魏洪斌, 罗明, 吴克宁, 等. 长江三角洲典型县域耕地土壤重金属污染生态风险评价[J]. *农业机械学报*, 2021, 52(11): 200–209, 332.
WEI Hongbin, LUO Ming, WU Kening, et al. Ecological risk assessment of heavy metal pollution in cultivated soil at typical county level in Yangtze River Delta[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021, 52(11): 200–209, 332. (in Chinese)
- [8] CROW W T, BERG A A, COSH M H, et al. Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products[J]. *Reviews of Geophysics*, 2012, 50(2): 372–392.
- [9] 刘振民. 区域土壤环境质量现状一致性估计及实现[D]. 廊坊: 北华航天工业学院, 2021.
LIU Zhenmin. Consistency estimation and realization of regional soil environmental quality status[D]. Langfang: North China Institute of Aerospace Industry, 2021. (in Chinese)
- [10] 杨琳, 朱阿兴, 秦承志, 等. 一种基于样点代表性等级的土壤采样设计方法[J]. *土壤学报*, 2011, 48(5): 938–946.
YANG Lin, ZHU Axing, QIN Chengzhi, et al. A soil sampling method based on representativeness grade of sampling points[J]. *Acta Pedologica Sinica*, 2011, 48(5): 938–946. (in Chinese)
- [11] 李洪建, 高玉凤, 严俊霞, 等. 不同取样尺度下亚高山草甸土壤呼吸的空间变异特征[J]. *环境科学*, 2014, 35(11): 4313–4320.
LI Hongjian, GAO Yufeng, YAN Junxia, et al. Spatial heterogeneity of soil respiration in a subalpine meadow at different sampling scales[J]. *Environmental Science*, 2014, 35(11): 4313–4320. (in Chinese)
- [12] 刘京, 朱阿兴, 张淑杰, 等. 基于样点个体代表性的大尺度土壤属性制图方法[J]. *土壤学报*, 2013, 50(1): 12–20.
LIU Jing, ZHU Axing, ZHANG Shujie, et al. Large-scale soil attribute mapping method based on individual representativeness of sample points[J]. *Acta Pedologica Sinica*, 2013, 50(1): 12–20. (in Chinese)
- [13] RICARDO A O. Declustering of clustered preferential sampling for histogram and semivariogram inference[J]. *Mathematical Geology*, 2007, 39(5): 453–467.
- [14] 赵业婷, 常庆瑞, 李志鹏, 等. 基于 Cokriging 的耕层土壤全氮空间特征及采样数量优化研究[J]. *土壤学报*, 2014,

- 51(2): 415–422.
- ZHAO Yeting, CHANG Qingrui, LI Zhipeng, et al. Study on spatial distribution of total nitrogen in arable soil layer and optimization of number of soil samples for the study based on Cokriging[J]. *Acta Pedologica Sinica*, 2014, 51(2): 415–422. (in Chinese)
- [15] 唐炬彪, 朱庆伟, 董士伟, 等. 农业用地土壤重金属样本点数据精化方法——以北京市顺义区为例[J]. *农业环境科学学报*, 2020, 39(10): 2288–2296.
- TANG Guibiao, ZHU Qingwei, DONG Shiwei, et al. Data refinement method for sampling sites of agricultural soil heavy metals: a case study in Shunyi District, Beijing, China[J]. *Journal of Agro-Environment Science*, 2020, 39(10): 2288–2296. (in Chinese)
- [16] 吴丹, 吉红霞, 刘玉, 等. 华北平原土地整治项目区农用机井的空间布局优化[J]. *经济地理*, 2015, 35(6): 154–160.
- WU Dan, JI Hongxia, LIU Yu, et al. Optimizing the distribution pumping wells for agricultural land consolidation in the North China Plain[J]. *Economic Geography*, 2015, 35(6): 154–160. (in Chinese)
- [17] 李懋峰, 谭煌, 王嘉航, 等. 滴灌水肥条件对樱桃产量、品质和土壤理化性质的影响[J]. *农业机械学报*, 2017, 48(7): 236–246.
- LI Pingfeng, TAN Huang, WANG Jiahang, et al. Effect of water and fertilizer conditions under drip irrigation on yield, quality of cherry and physicochemical properties of soil[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2017, 48(7): 236–246. (in Chinese)
- [18] 樊桂菊, 李钊, 毛文华, 等. 基于工作空间的果园作业平台结构参数优化与试验[J]. *农业机械学报*, 2021, 52(4): 34–42, 265.
- FAN Guiju, LI Zhao, MAO Wenhua, et al. Structure parameter optimization and experiment of orchard platform based on workspace[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021, 52(4): 34–42, 265. (in Chinese)
- [19] DEUTSCH C. DECLUS; a fortran 77 program for determining optimum spatial declustering weights[J]. *Computers and Geotechnics*, 1989, 15(3): 325–332.
- [20] PAN Y C, REN X H, GAO B B, et al. Global mean estimation using a self-organizing dual-zoning method for preferential sampling[J]. *Environmental Monitoring and Assessment*, 2015, 187(3): 121.
- [21] REN X H, WEI N, GAO B B, et al. Experimental study on estimation of global mean with preferential spatial samples[J]. *Intelligent Automation & Soft Computing*, 2014, 20(4): 625–635.
- [22] 郭永奇. 基于惩罚型变权的农地生态安全预警评价——以新疆生产建设兵团为例[J]. *地域研究与开发*, 2014, 33(5): 149–154.
- GUO Yongqi. Forewarning evaluation of farmland ecological security based on punished variable weight: a case of Xinjiang Production and Construction Corps[J]. *Areal Research and Development*, 2014, 33(5): 149–154. (in Chinese)
- [23] 范曼曼, 吴鹏豹, 张欢, 等. 采样密度对土壤有机质空间变异解析的影响[J]. *农业现代化研究*, 2016, 37(3): 594–600.
- FAN Manman, WU Pengbao, ZHANG Huan, et al. Effect of sampling density on spatial variability analysis of soil organic matter[J]. *Research of Agricultural Modernization*, 2016, 37(3): 594–600. (in Chinese)
- [24] 陈光, 贺立源, 詹向雯. 耕地养分空间插值技术与合理采样密度的比较研究[J]. *土壤通报*, 2008, 39(5): 1007–1011.
- CHEN Guang, HE Liyuan, ZHAN Xiangwen. Comparison of spatial interpolation technique of soil nutrient and reasonable sampling density[J]. *Chinese Journal of Soil Science*, 2008, 39(5): 1007–1011. (in Chinese)
- [25] 李若溪, 周翔, 吕婷婷, 等. 植被异质性样区真实性检验的优化采样策略[J]. *农业工程学报*, 2021, 37(8): 177–186.
- LI Ruoxi, ZHOU Xiang, LÜ Tingting, et al. Optimum sampling strategy for authenticity test in heterogeneous vegetated areas[J]. *Transactions of the CSAE*, 2021, 37(8): 177–186. (in Chinese)
- [26] 陈剑, 瞿明凯, 王燕, 等. 长三角平原区县域土壤磷素流失风险及其空间不确定性的快速评估[J]. *生态学报*, 2019, 39(24): 9131–9142.
- CHEN Jian, QU Mingkai, WANG Yan, et al. Rapid assessment of soil phosphorus loss risk and its spatial uncertainty in county areas of the Yangtze River Delta Plain[J]. *Acta Ecologica Sinica*, 2019, 39(24): 9131–9142. (in Chinese)
- [27] REMY N, BOUCHER A, WU J. Applied geostatistics with SGeMS: a user's guide[M]. New York: Cambridge University Press, 2009: 135–139.