

面向青藏高原地区耕地资源质量评估的空间抽样方法对比与应用

蒋嘉益^{1,2}, 宋长青^{1,2}, 叶思菁^{1,2*}, 高培超^{1,2}, 任书义²

(1. 北京师范大学地表过程与资源生态国家重点实验室, 北京, 100875; 2. 北京师范大学地理科学学部, 北京, 100875)

摘要: 充分认识青藏高原地区耕地资源质量的空间格局与时间变化特征, 是探索区域性耕地资源保护与可持续利用路径的重要基础, 对维护青藏高原地区的粮食安全和社会稳定具有重要意义。设计空间抽样方案, 利用较少的样点尽可能全面地表征区域耕地质量的总体特征, 对于开展耕地资源质量长期观测与机理研究具有重要意义。该研究以青海省互助土族自治县为研究区开展基于耕地质量指标数据驱动的空间抽样实证研究, 从信息熵、Kullback-Leibler 散度、相似度距离、样本对总体耕地质量空间异质性的表征能力、样本空间分布均匀性等视角, 对比分析随机抽样法 (random sampling method, RSM), 空间覆盖随机抽样法 (spatial coverage sampling and random sampling, SPCOSA), 条件拉丁超立方体法 (conditioned latin hypercube sampling, CLHS), 加入平面坐标的条件拉丁超立方体法 (CLHS with x and y coordinates as covariates, XY_CLHS), 空间覆盖随机抽样与条件拉丁超立方体抽样的混合抽样法 (spatial coverage sampling and random sampling-conditioned latin hypercube sampling, SPCOSA_CLHS) 5 种空间抽样方法在青藏高原县级区域耕地质量指标调查观测点位布局应用中的优劣特征与适用性, 并探索了研究区适宜的观测点位数量。结果表明: SPCOSA_CLHS 可以以较低空间约束的方式将 SPCOSA 指示的空间异质性特征集成到 CLHS 模型中, 在表达总体的耕地质量指标属性特征和空间异质性特征方面更具优势; 当抽样数量缩减到 40~50 时, 抽样结果对总体耕地质量指标属性信息量的表征能力与抽样数量为 100~200 时近似; 且 SPCOSA_CLHS 方法在表达耕地质量指标空间异质性、设计样点分布的空间均匀性、模拟总体耕地质量特征的准确性方面具有明显优势。该研究可以为青藏高原地区耕地资源质量调查监测工作提供方法支持, 进而对理解该地区耕地资源质量变化过程、探索耕地可持续利用路径具有支持作用。

关键词: 土地利用; 资源; 耕地质量; 空间抽样; 青藏高原; GIS

doi: 10.11975/j.issn.1002-6819.202306156

中图分类号: F323.21

文献标志码: A

文章编号: 1002-6819(2023)-21-0246-12

蒋嘉益, 宋长青, 叶思菁, 等. 面向青藏高原地区耕地资源质量评估的空间抽样方法对比与应用[J]. 农业工程学报, 2023, 39(21): 246-257. doi: 10.11975/j.issn.1002-6819.202306156 <http://www.tcsae.org>

JIANG Jiayi, SONG Changqing, YE Sijing, et al. Comparison and application of the spatial sampling methods for assessing the quality of arable land resources in Qinghai-Tibet Plateau[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(21): 246-257. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202306156 <http://www.tcsae.org>

0 引言

耕地是粮食生产的基本物质资源, 也是人类社会生存和发展的重要基础。中国的人口占世界总人口的 21%, 人均耕地面积却只有世界人均耕地面积的 1/4, 并面临快速工业化、城市化进程的诸多挑战^[1-2]。保护中国的耕地资源, 对实现联合国可持续发展目标, 保障国家粮食安全与社会稳定具有重要意义^[3]。青藏高原有“地球第三极”和“世界屋脊”之称, 是世界上最重要的高寒农牧区之一。该地区虽然耕地面积有限, 但因山盆结构高原地形导致粮食运输不便, 加之其特色食源青稞已成为藏区居民生活及藏族文化的有机组成部分^[4], 耕地资源对于保障区域粮食供给及社会文化至关重要。长期以来,

青藏高原地区因其独特的地理环境, 耕地利用面临低温冻害、干旱、水土流失等生态环境问题胁迫^[4-6], 同时, 农业机会成本的增加导致耕地撂荒现象频发; 不合理的耕地空间转换导致土壤肥力下降, 粮食安全风险增加^[5]。充分认识青藏高原地区耕地资源质量的空间格局与时间变化特征, 是探索区域性耕地资源保护与可持续利用路径的重要基础, 对维护青藏高原地区的粮食安全和社会稳定具有重要意义。已有青藏高原地区耕地利用相关研究, 主要集中在农田生态系统、农业集约化、耕地布局时空转换等方面^[4-8], 关于该地区耕地资源质量主题的研究较少。

作为表征耕地内在属性特征与外在利用状况的重要指标, 耕地质量内涵在学界尚未完全统一^[9-13]。通常认为耕地资源质量是对于耕地生产、生活、生态多元功能的量化表达^[14-16], 是由自然因素和人为因素综合作用而形成的满足作物生长的综合体^[17-18]。耕地质量因无法被直接观测而往往通过多维指标综合表征, 相关指标覆盖地形特征、土壤性状、耕作条件、环境状况、生物特性等维度^[19-21]。理解耕地资源质量空间格局并解析其主导驱动指标及临界条件, 可为区域耕地质量优化与利用模式调控提供重要科技支撑。点位尺度指标调查观测是耕

收稿日期: 2023-06-23 修订日期: 2023-11-03

基金项目: 第二次青藏高原综合考察研究 (2019QZKK0608); 国家自然科学基金项目 (42171250); 地表过程与资源生态国家重点实验室自主课题 (2022-ZD-04)

作者简介: 蒋嘉益, 研究方向为耕地资源质量与可持续利用。

Email: 202221051128@mail.bnu.edu.cn

*通信作者: 叶思菁, 博士, 副教授, 硕士生导师, 研究方向为耕地资源保护与可持续利用。Email: yesj@bnu.edu.cn

地资源质量研究的重要数据源。在调查实践中,存在指标类型多样、部分指标调查过程复杂并难以通过电化学传感器与遥感技术有效观测、人工调查成本高且效率低等难题^[22-26],导致耕地质量指标时空密集数据获取困难,成为解析耕地资源质量格局与变化过程的主要限制因素。设计空间抽样方案,利用较少的样点尽可能全面地表征区域耕地质量的总体特征,能够为耕地资源质量研究提供点位尺度的数据,并能有效缩减成本、提高效率,具有重要意义,特别是对地广人稀、环境恶劣的青藏高原地区。

空间抽样模型可以为耕地质量指标调查点位布局提供技术支持,即通过在研究区布设少量观测样点,有效反映全局耕地质量空间分异特征。由于总体的空间相关性和空间分异性,经典抽样不能简单地应用于空间对象的抽样。基于此,相关研究在传统抽样的基础上融合空间抽样对象的位置属性,逐渐形成多类型空间抽样方法。空间抽样方法主要有随机抽样、便利抽样和目的性抽样 3 类^[27]。随机抽样指给予抽样总体中每个样本一个固定的概率,并基于此抽取样本,包括简单随机抽样和分层随机抽样、系统抽样等加入其他限制条件的抽样方法。便利抽样优先考虑抽样成本,在易于获取样品的地方进行布样,如沿道路布设样点^[28]。目的性抽样指在目标指导下进行布样,目标函数法是其最主要的方法,通常综合考虑地理特征空间分布和属性特征空间分布。在缺乏先验知识和其他属性辅佐时,样点的空间分布均匀程度是常见优化目标,如平均最短距离、加权平均最短距离等方法^[29-30];而在抽样总体的空间变异函数可推导前提下,一般可行方案是将泛 Kriging 估计误差作为优化目标函数,以约束抽样点的选择与优化^[31-34]。部分空间抽样方法提出采用属性特征空间格局作为优化目标,如拉丁超立方体法 (latin hypercube sampling, LHS) 和条件拉丁超立方体法 (conditioned latin hypercube sampling, CLHS) 等^[35-36],其原理是使用总体的多维属性变量组成拉丁超立方体,并将每一维度的概率密度函数分割成非重叠层,然后在各层内随机或根据指定的规则选择并组合得到符合抽样值的样本点。GAO 等^[37]提出同时优化地理特征空间分布和属性特征空间分布的空间条件拉丁超立方体 (spatial conditioned latin hypercube sampling, SCLHS) 抽样方法。SCLHS 在基于 CLHS 方法构建样本属性特征空间优化目标基础上,通过将研究区域划分为若干规则空间格网,集成以每个格网尽可能提取相等数量样本点的地理特征空间优化目标。然而, SCLHS 对离散分布的空间抽样目标不适用,因为规则空间格网划分过程往往忽视空值区域的影响,而耕地地块在研究区往往是呈空间离散分布的。对于离散目标 (如耕地) 的空间抽样, WANG 等^[38]提出“三明治”空间抽样模型,模型由样本层、知识层 (分区) 及报告单元 (如行政单元) 相互独立的 3 层构成,在对行政单元内耕地面积的调查中得到广泛应用。WAN 等^[39]提出空间覆盖随机抽样与条件拉丁超立方体抽样的混合抽样法 (spatial coverage sampling and random sampling-conditioned latin hypercube sampling, SPCOSA_CLHS),并在省级尺度耕地质量监测样点空间布局的实证案例中验证 SPCOSA_CLHS 在空间均

衡性、可扩展性和减少抽样成本方面的性能与适用性。

面向青藏高原地区县级尺度耕地质量格局与过程研究需求,本文以青海省互助土族自治县为例,以耕地栅格单元总体的地形特征、土壤性状和耕作条件指标数据为驱动,在不同样点数量情景下应用随机抽样法 (random sampling method, RSM)、SPCOSA (按 SPCOSA 分区的样本大小进行等概率抽样)、CLHS、XY_CLHS (除属性变量外,在 CLHS 中使用 x 、 y 坐标作为协变量)、SPCOSA_CLHS 等方法设计空间抽样方案。在此基础上,对比不同抽样方案对总体耕地质量指标属性特征的信息量与空间异质性的表征能力,及样点分布的空间均匀度,以期对青藏高原地区耕地质量调查点位数量与空间布局设计提供技术支撑。

1 数据与方法

1.1 研究区概况

青海省互助土族自治县地处祁连山脉东段南麓,位于黄土高原与青藏高原交错嵌接地带,介于东经 $101^{\circ}46'$ ~ $103^{\circ}45'$ 和北纬 $36^{\circ}30'$ ~ $37^{\circ}90'$ 之间,海拔 $2\ 200$ ~ $4\ 484$ m 之间,县境内地形起伏多变,岭谷纵横交错,高差悬殊,整体地势呈北高南低。气候属半干旱大陆性季风气候,年降水量在 400 ~ 600 mm 之间,年均气温 3.4 $^{\circ}\text{C}$,热量资源水平时空分布不均衡,垂直地带差异十分明显^[40]。互助土族自治县是一个以农业为主的县,有着悠久的农业生产历史,全县总耕地面积 $69\ 548.11$ hm^2 ,以旱地为主,主要种植小麦、马铃薯和蚕豆等作物。根据自然资源部耕地质量评估试点工程,互助土族自治县中等地为 $9\ 794.46$ hm^2 ,占总面积 14.08% ;低等地为 $59\ 753.65$ hm^2 ,占总面积 85.92% (图 1)。

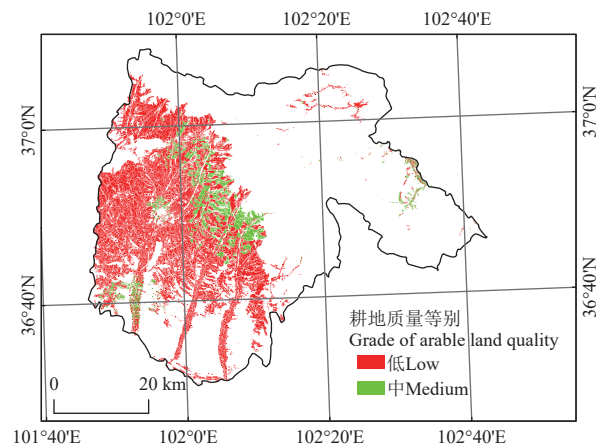


图 1 青海省互助县耕地质量等级空间格局

Fig.1 Spatial pattern of grade of arable land quality in Huzhu County in Qinghai Province

互助土族自治县耕地质量水平整体较低,中等地主要分布在县域中部、南部,包含部分水浇地;低等地以旱地为主,广泛分布在县域西北部、东部。

1.2 耕地质量指标数据

地形特征、土壤性状、耕作条件、环境状况、生物特性是表征耕地资源质量的重要维度^[10,41]。本文以互助土族自治县内 1 km 空间分辨率耕地栅格点的集合为抽

样总体,共 790 个格点,并从地形特征、土壤性状、耕作条件 3 个维度提取 21 个指标(表 1)构建总体耕地质量指标属性特征数据集,作为空间抽样的输入数据。环境

状况和生物特性指标在研究区内差异很小,因此本研究没有考虑。在数据处理方面,参考叶思菁等^[20]的转换方案,对不同量纲的耕地质量指标数据执行归一化处理。

表 1 作为空间抽样输入参数的耕地质量指标集合

Table 1 Indicator dataset of arable land quality as spatial sampling input

类别 Type	评价指标 Evaluation indicator	指标含义 Definition	数据源 Data source
地形特征 Topographic features	高程/m	海拔高度	ASTGTM 30 m 高程数据 ^[42]
	坡度/(°)	地表单元平缓的程度	
土壤性状 (0~60 cm 深) Soil properties (0~60 cm depth)	土体厚度/cm	—	数据来源于国家科技基础条件平台-国家地球系统科学数据中心-土壤分中心 ^[43] (http://soil.geodata.cn)
	土壤容重/(g·cm ⁻³)	—	
	黏粒含量/(g·kg ⁻¹)	—	
	粉粒含量/(g·kg ⁻¹)	—	
	砂粒含量/(g·kg ⁻¹)	—	
	阳离子交换量/(cmol·kg ⁻¹)	—	
	砾石含量/%	—	
	有机碳含量/(g·kg ⁻¹)	—	
	全钾含量/(g·kg ⁻¹)	—	
	全氮含量/(g·kg ⁻¹)	—	
全磷含量/(g·kg ⁻¹)	—	互助土族自治县国土局	
土壤 pH 值 (0~5 cm 深)	—		
耕作条件 Tillage technical conditions	田间道路通达度	与田间道路相交的耕地地块数(容差为 5 m)与耕地总地块数之比	互助土族自治县国土局
	农田防洪标准	以乡镇为尺度的农田防洪标准	互助土族自治县水利局
	灌溉保证度	灌溉保证度	互助土族自治县统计局
	农机化水平	以乡镇为尺度,由 2016 年的机耕面积、机播面积和机种面积计算得到	互助土族自治县国土局
	排水条件	空间连接得到每个村的排水条件,再根据评价单元的权属求得	互助土族自治县统计局
	田间输水方式	田间输水方式	互助土族自治县水利局
	耕作距离/m	农户居住地到所耕作地块的直线距离	互助土族自治县农牧局 LIU 等 ^[44]

此外,为了从更高维度对比分析不同空间抽样方案对总体耕地质量空间异质性特征表征能力的表征能力,本文选择研究区自然质量指数、技术水平指数、耕地产能指数 1 km 分辨率栅格数据集作为参考数据。自然质量指数、技术水平指数、耕地产能指数数据集均来源于自然资源部耕地质量评估试点工程^[20],其中自然质量指数是对耕地地形特征与土壤性状的综合表达,技术水平指数是对耕作条件指标的综合表达,耕地产能指数则反映气候、地形、土壤、耕作条件综合作用下耕地的生产潜力。图 2 为互助土族自治县耕地质量指标数据集及自然质量指数、技术水平指数、耕地产能指数数据集的空间格局。

1.3 地理探测器

抽样对象具备显著空间异质性特征是空间抽样方法适用的重要理论前提。本文使用地理探测器来验证抽样总体的空间分层异质性,进而说明空间抽样方法的适用性。空间分层异质性指地层内比地层间相似度高的现象,此处的“地层”是一个统计概念,大致对应于地理上的子区域。地理探测器是王劲峰等^[45]提出的一种空间统计方法,用于检测空间分层异质性并揭示驱动因素。如式(1)所示, h 代表变量 X 的地层; N_h 和 N 分别是地层 h 和全局的单元数; σ_h^2 和 σ_t^2 分别是地层内平方和与全局总平方和。 q 的取值范围为[0,1], q 值越大表示空间分层异质性越强。

$$q = 1 - \frac{\sum_{h=1}^L N_h \sigma_h^2}{N \sigma_t^2} \quad (1)$$

1.4 SPCOSA_CLHS 方法

拉丁超立方体法(LHS)使用多维属性变量构成多维拉丁超立方体^[35],并将每一维度的概率密度函数分割成非重叠层,然后在各层内随机的或根据指定的规则选择并组合得到特征向量,满足特征向量的样本点,即为抽样点。然而,现实中的样本点数量有限,总体中不一定存在满足特征向量的样本点。MINASNY 等^[36]基于 LHS 提出了条件拉丁超立方体法(CLHS),通过设计目标函数,并使用模拟退火算法在总体中寻找可能满足 LHS 分布的样本点,从而将样点的选择问题转为优化求解问题,选择能够最小化目标函数的样本点。目标函数如式(2)所示。其中, O 是全局优化目标; O_1 、 O_2 、 O_3 分别代表连续变量、分类变量以及样本与总体的相关矩阵差的子目标函数(式(3)~式(5))。 w_1 、 w_2 、 w_3 是各个子函数对应的权重,一般设置为 1。 O_1 、 O_2 的设计是为了使每一层的样本数尽可能为 1,使得目标函数的值为 0。

$$O = w_1 O_1 + w_2 O_2 + w_3 O_3 \quad (2)$$

$$O_1 = \sum_k \sum_{l=1}^f |\eta(q_l^k \leq s_l \leq q_l^{k+1}) - 1| \quad (3)$$

$$O_2 = \sum_j \left| \frac{\eta(s_j)}{e} - s_j \right| \quad (4)$$

$$O_3 = \sum_k \sum_{l=1}^f |c_{kl} - t_{kl}| \quad (5)$$

式中 k 和 l 代表层和属性空间维度; e 和 f 代表选择的样

本点数和辅助变量； $\eta(q^k_i \leq s_i \leq q^{k+1}_i)$ 表示某个辅助变量 l 和某个层 k 的样本数； $\eta(s_i)$ 为属于 l 类的抽样数； s_i 为属

于 l 类的总体数量； c_{kl} 代表总体相关系数矩阵的元素； t_{kl} 代表样本相关系数矩阵的元素。

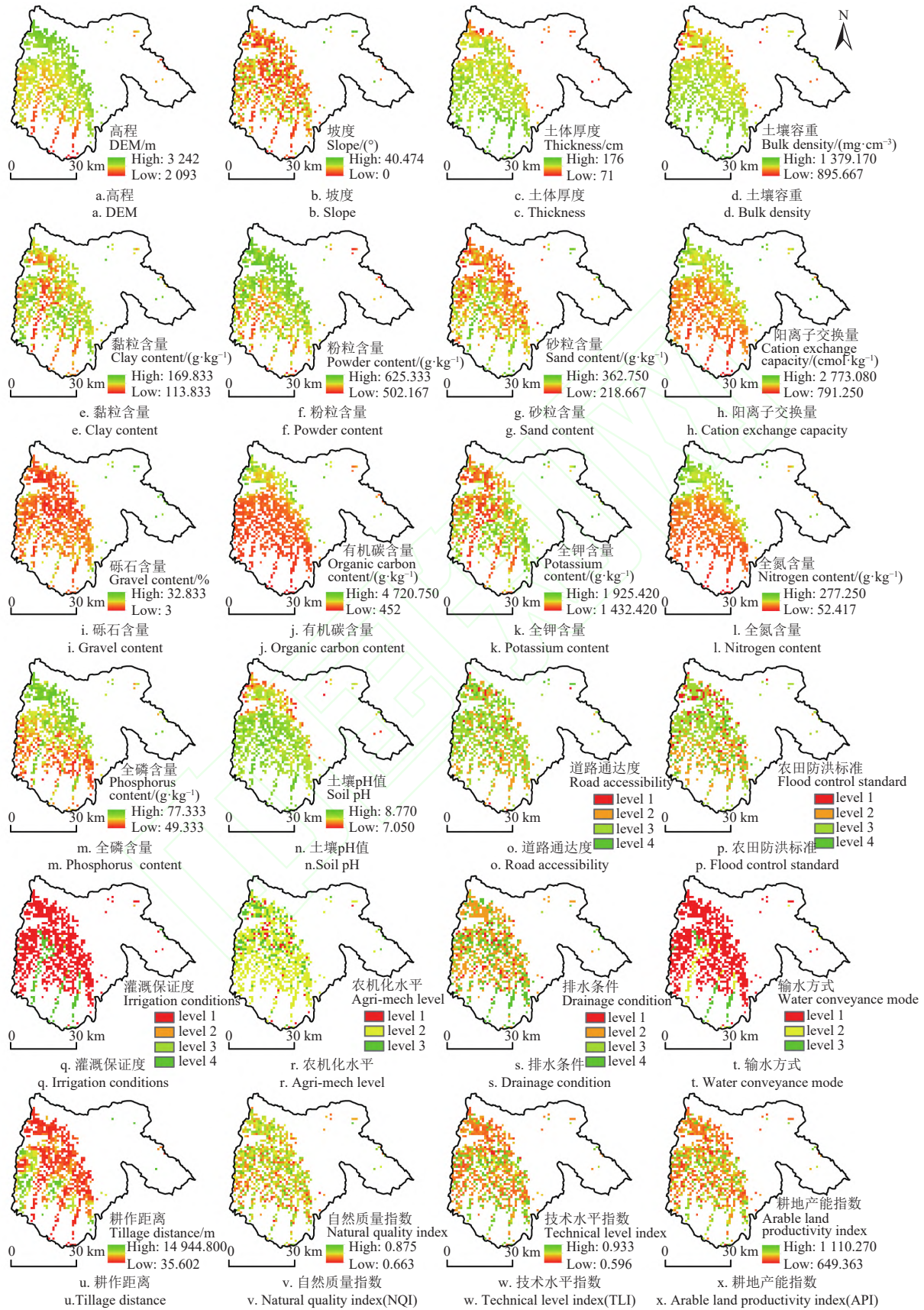


图 2 互助县耕地质量指标与综合指数空间格局 (分级越高表示指标分值越大, 耕作条件越充分)

Fig.2 Spatial pattern of arable land quality indicators and comprehensive indices in Huzhu County (The higher the grading, the greater the indicator score and the more sufficient the tillage technical conditions are)

在实践中, 耕地图斑呈空间离散分布, 并具有一定的空间聚集性与空间异质性, 如果仅应用耕地质量指标作为 CLHS 方法的输入参量, 将难以调控调查样点空间布局的整体性与匀质性, 需要融合空间参量。空间覆盖随机抽样法 (SPCOSA) 是 WALVOORT 等^[46] 基于 k-means 开发的一种典型的地理分层抽样方法, 选择均方最短距离作为目标函数, 将研究区域划分为地理紧凑的地层。它避免了样本位置的聚类, 可以更好地实现空间均匀性, 从而提高在地理空间中的表示。WAN 等^[39] 结合二者提出了同时考虑地理空间和属性空间分层的 SPCOSA_CLHS 方法。首先, 使用 SPCOSA 将研究区域划分为等面积的子区域并对其进行编号; 然后, 将这些编号作为空间参量添加输入到 CLHS 中, 用作地理空间约束, 并将 1 km 分辨率的耕地质量指标栅格数据集用作属性空间约束, 在兼顾地理空间和属性空间约束基础上求解抽样优化方案。本文在不同样点数量情景下应用 5 种空间抽样方法设计研究区耕地质量调查样点空间布局方案。

1.5 抽样结果评估方法

本文从信息量、样本对总体耕地质量空间异质性的表征能力、样本空间分布均匀性等视角对比分析不同抽样方法的优劣及适用性。

1.5.1 信息熵

信息熵 $H(X)$ 描述了信息源各可能事件发生的不确定性^[47], 通过式 (6) 计算获取, 其中, m 表示耕地质量评价指标的数量, 对于每个耕地质量评价指标 j , i 表示指标属性值类型, n 表示指标属性值类型 i 的数量, $P(X_i^j)$ 表示指标 j 的某一属性值类型 i 出现的频率。本文通过对比不同抽样方法抽样结果与总体的信息熵, 衡量抽样结果以少量样点表达总体信息量的能力。抽样结果信息熵越接近总体信息熵, 表示对总体信息量的表征能力越强。

$$H(X) = - \sum_{j=1}^m \sum_{i=1}^n P(X_i^j) \log_2 P(X_i^j) \quad (6)$$

1.5.2 Kullback-Leibler 散度 (KL 散度)

从熵相似性的角度来看, KL 散度被用来衡量抽样总体和样点之间的多样性, 以度量使用样点分布来近似总体分布时所损失的信息量。对于每个耕地质量评价指标 j , 采用式 (7) 计算 KL 散度, 其中, $Q(X_i)$ 是样点概率分布, $P(X_i)$ 是原始总体概率分布, 如果两个分布完全匹配, $KL(P||Q)=0$; 否则, $KL(P||Q)>0$ 。本文通过对比不同抽样方法抽样结果的 KL 散度, 衡量抽样结果样点频率分布与总体频率分布的相似程度。KL 散度越小, 样点频率分布和总体频率分布之间的差异越小, 样点频率分布对总体频率分布的表示性就越好。

$$KL(P||Q) = \sum_{i=1}^n P(X_i) \times \log_2 \left[\frac{P(X_i)}{Q(X_i)} \right] \quad (7)$$

1.5.3 相似度距离

利用相似度距离 D 计算个体间信息熵的差异^[39], 如式 (8) 所示, 其中, $h(P_j)$ 、 $H(P_j)$ 分别为不同抽样方法

和总体的辅助变量 j 的信息熵。本文使用相似度距离来表征不同抽样方法抽样结果基于每个指标的信息熵差异, 以衡量抽样结果对总体的代表性。相似度距离 D 越小, 样点对总体属性特征的代表性越强。

$$D = \sqrt{\frac{\sum_{j=1}^m [h(P_j) - H(P_j)]^2}{m}} \quad (8)$$

1.5.4 焦点统计与点统计

采用焦点统计方法量化多种抽样方法在表达耕地质量局部空间异质性方面的效果, 如式 (9) 所示^[48], 对于每个抽样点 z , 生成一个半径为 3 km 的缓冲区 C ; 然后, 分别提取与相邻区域 C 相交的所有耕地格点的自然质量指数、技术水平指数、耕地产能指数 y , 计算标准差 σ_z , 每个 z 对应一个 σ_z ; 最后, 计算 σ_z 的平均值, 其中, t 表示与 C 相交的耕地网格总数, \bar{y} 表示所有 y 的平均值, s 表示抽样点的总数。 $\bar{\sigma}$ 越高, 表示抽样结果对耕地质量局部空间异质性的表现越好。

$$\sigma_z = \sqrt{\frac{\sum (y - \bar{y})^2}{t+1}}, \bar{\sigma} = \frac{\sum \sigma_z}{s} \quad (9)$$

此外, 本文采用点统计法评价抽样结果的空间分布均匀性。其算法原理如式 (10) 所示, 对于每个耕地网格 G , 生成半径为 5 km 的缓冲区, 覆盖的样本点数计为 C_G 。计算所有 C_G 的平均值 \bar{C} 和标准差 σ , 其中 n_{um} 表示耕地网格的总数。具有较高 \bar{C} 和较低 σ 的抽样结果具有更好的空间分布均匀性。

$$\bar{C} = \frac{\sum C_G}{n_{um}}, \sigma = \sqrt{\frac{\sum (C_G - \bar{C})^2}{n_{um} + 1}} \quad (10)$$

1.5.5 双变量空间自相关

空间自相关是检验某一要素属性值是否显著与其相邻空间点上的属性值相关联的重要指标^[49]。双变量空间自相关可反映两个不同属性变量之间的空间聚集关系^[50]。如式 (11) 所示, a 、 b 是用于双变量空间自相关计算的两个属性, X_a^p 是空间单元 p 属性 a 的值, X_b^q 是空间单元 q 属性 b 的值, δ_a 、 δ_b 是属性 a 、 b 的方差, W_{pq} 是空间权重值。本文分别计算基于不同抽样方案的反距离加权插值结果与总体自然质量指数、技术水平指数、耕地产能指数数据集的双变量空间自相关特征, 以评估不同空间抽样方法对模拟总体耕地资源质量特征的准确性。莫兰指数 (I_{ab}^p) 越高, 正相关的耕地格点越多, 说明抽样方法模拟总体耕地资源质量特征的准确性越高。

$$I_{ab}^p = \frac{X_a^p - \bar{X}_a}{\delta_a} \sum_q W_{pq} \frac{X_b^q - \bar{X}_b}{\delta_b} \quad (11)$$

2 结果与分析

2.1 耕地质量指标空间异质性分析

表 2 为总体的空间自相关和空间异质性计算结果。

本文选择 5 000 m 作为欧氏距离来计算空间权重矩阵，并计算 21 个耕地质量指标及自然质量指数、技术水平指数、耕地产能指数的莫兰指数。结果表明所有指标均表现出显著的全局空间自相关性，其中高程、土体厚度、土壤容重、阳离子交换量、土壤有机碳含量、全氮含量、全磷含量、耕作距离的莫兰指数在 0.5 以上。

本文在 SPCOSA 将研究区域划分为 20 个子区域的基础上，采用地理探测器计算总体耕地质量指标的分层空间异质性，结果显示除田间道路通达度、农田防洪标

准、农机化水平、排水条件、自然质量指数这 5 个指标外，其他所有指标均表现出显著的分层空间异质性，且 SPCOSA 分区方案可以超过 50% 解释高程、土体厚度、阳离子交换量、土壤有机碳含量、全氮含量的分层空间异质性。

综上，说明研究区耕地质量指标数据具有空间相关性和空间异质性特征，符合空间抽样假设；且 SPCOSA 分区结果可以较好地表达这种空间异质性，适合作为地理空间约束参量输入空间抽样方法。

表 2 各指标空间自相关与空间异质性计算结果

Table 2 Results of spatial autocorrelation and spatial heterogeneity of each indicator

编号 No.	指标 Indicator	莫兰指数 Moran's I	q 值 q value	编号 No.	指标 Indicator	莫兰指数 Moran's I	q 值 q value
1	高程	0.753**	0.611**	13	全磷含量	0.567**	0.482**
2	坡度	0.289**	0.190**	14	土壤 pH 值	0.452**	0.333**
3	土体厚度	0.569**	0.500**	15	田间道路通达度	0.042**	0.056
4	土壤容重	0.586**	0.492**	16	农田防洪标准	0.048**	0.051
5	黏粒含量	0.281**	0.076*	17	灌溉保证度	0.399**	0.242**
6	粉粒含量	0.483**	0.433**	18	农机化水平	0.024**	0.048
7	砂粒含量	0.378**	0.281**	19	排水条件	0.068**	0.075
8	阳离子交换量	0.671**	0.565**	20	田间输水方式	0.408**	0.268**
9	砾石含量	0.416**	0.386**	21	耕作距离	0.524**	0.426**
10	有机碳含量	0.717**	0.592**	22	自然质量指数	0.023**	0.044
11	全钾含量	0.321**	0.153**	23	技术水平指数	0.118**	0.090**
12	全氮含量	0.780**	0.666**	24	耕地产能指数	0.089**	0.079*

注：显著*，极显著**。
Note: Significant*, Extreme significant**.

2.2 基于信息量的多空间抽样方法适用性对比

本文分别采用 5 种空间抽样方法，从总体中抽取样本点集，对于每种方法，本文在设定多组不同的样本数量情景下分别输出样本点集。样本数量情景包括 10、20、30、40、50、75、100、150、200。在此基础上，面向样本点集分别计算信息熵、KL 散度、相似度距离。样本点集的信息熵越接近总体信息熵，KL 散度或相似度距离越小，表明抽样结果对总体耕地质量属性特征的代表性越强。

不同样本数量下，基于各空间抽样方法分别输出的样本集合的信息熵见图 3a。总体信息熵 E_t 为 6.535。样本数量在 75~200 的区间变化时，SPCOSA_CLHS、

XY_CLHS 和 CLHS 样本点集的信息熵较为稳定，在区间 ($E_t-0.020, E_t+0.040$) 内轻微波动；而 SPCOSA 和 RSM 样本点集的信息熵波动较大，在区间 ($E_t-0.200, E_t+0.040$) 内波动。

在样本量从 75 减少到 10 的过程中，样本信息熵相对 E_t 的偏离程度加剧。与其他方法相比，SPCOSA_CLHS 表现出更稳定的波动性，表明其抽样结果对总体属性特征的代表性受样本量变化的影响较小。随着样本量减少，基于其他方法的抽样结果均呈现出信息熵大幅波动的趋势，表明样本数量不足会导致空间抽样方法的适用性发生突变。

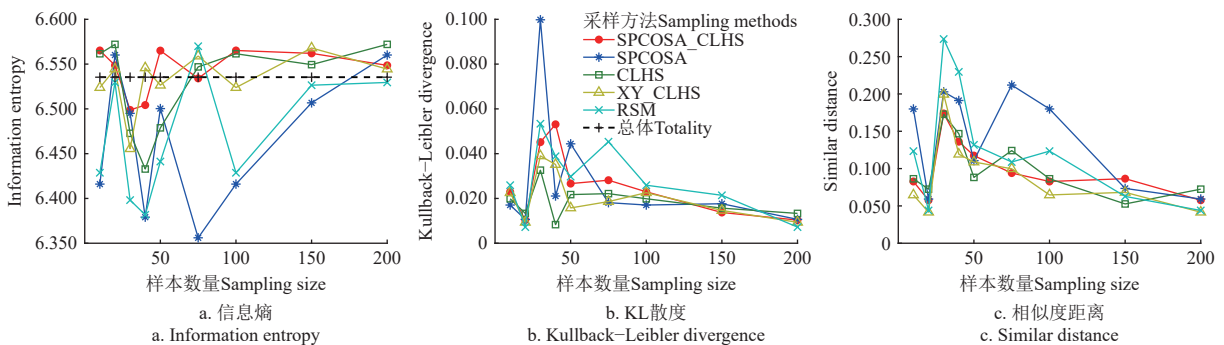


图 3 总体和多种抽样方法不同样本数量的抽样结果的信息熵、KL 散度和相似度距离结果

Fig.3 Result of the information entropy, the Kullback–Leibler divergence and the similar distance of the totality and samples generated by several sampling methods with multigroup sample sizes

图 3b 表现了不同空间抽样结果的 KL 散度差异。当样本量较高 (100~200) 时，基于各种方法抽样结

果的 KL 散度相似且均在 0.030 以下，从小到大依次为 SPCOSA_CLHS、XY_CLHS、CLHS、SPCOSA、RSM。随

着样本量减少, SPCOSA 和 RSM 抽样结果的 KL 散度增加最为剧烈。样本量为 75 时, RSM 抽样结果的 KL 散度大于 0.040, 其他抽样结果的 KL 散度均在 0.030 以下; 样本量为 30 时, SPCOSA 抽样结果的 KL 散度突变至 0.100, RSM 抽样结果的 KL 散度大于 0.050, 抽样结果的 KL 散度均在 0.050 以下。SPCOSA_CLHS、XY_CLHS 抽样结果的 KL 散度相对稳定, CLHS 抽样结果的 KL 散度波动最小。

图 3c 表现了不同空间抽样结果的相似度距离差异。不同空间抽样结果的信息熵与总体信息熵的相似度距离显示出与 KL 散度较为一致的变化特征。样本量在 150~200 的变化区间中, 多种方法产生的抽样结果的相似度距离相似 (均小于 0.100)。随着样本量的减少, 所有方法都显示出相似度距离的强烈增加与波动, SPCOSA、RSM、XY_CLHS 样本点集的相似度距离最高在 0.200 以上, SPCOSA_CLHS 和 CLHS 样本点集的相似度距离较小 (最高为 0.174 和 0.173), 说明 SPCOSA_CLHS 和 CLHS 抽样方法的稳定性较好。

当抽样数量在 100~200 区间时, 各空间抽样方法都很好地表征总体的耕地质量指标属性特征: 抽样结果的信息熵近似于总体信息熵, KL 散度和相似度距离较小, 受样本量变化的影响相对较小。当样本量低于 100 时, 不同方法的抽样结果之间差异增加, 其中 SPCOSA 和 RSM 抽样结果波动最大, SPCOSA_CLHS 抽样结果的信息熵、KL 散度和相似距离在收敛性和稳定性方面比其他方法更好。因此, SPCOSA_CLHS 比其他方法更适用于互助土族自治县耕地质量指标调查观测的样点设置。在实践中, 考虑到传感器设备的运行和维护成本, 应权衡样本数量和样本对总体属性特征的表征能力, 即用尽可能少的样本有效反映总体特征。当抽样数量设定在 40~50 时, 抽样结果对总体属性特征的表征能力与抽样数量为 100~200 时近似。

2.3 不同空间抽样结果的空间分布特征对比

根据前文结果, 设置抽样数量为 40, 对比分析不同抽样结果对总体自然质量指数、技术水平指数、耕地产能指数空间异质性特征的表征能力 (表 3) 及样点分布的均匀性 (表 4)。

表 3 焦点统计方法下各指数标准差

Table 3 Result of expressing local spatial heterogeneity through focal statistics method

指数 Index	SPCOSA_CLHS	SPCOSA	CLHS	XY_CLHS	RSM
NPI	16.223	15.864	16.180	14.880	16.204
TLI	23.551	22.461	21.651	20.991	20.328
API	20.978	20.402	20.082	19.767	18.720

表 4 点统计方法下样本点均值和标准差

Table 4 Result of multiple sampling methods on expressing spatial through point statistics method

指标 Indicator	SPCOSA_CLHS	SPCOSA	CLHS	XY_CLHS	RSM
均值 Average	1.989	2.041	1.885	1.941	1.800
标准差 SD	1.301	1.174	1.336	1.271	1.410

从表 3 可知。对于自然质量指数、技术水平指数和耕地产能指数 3 个指标, SPCOSA_CLHS 抽样结果焦点统计的平均标准差均高于其他抽样方法。结果表明, SPCOSA_CLHS 抽样结果对总体耕地质量局部空间异质性的表征能力优于其他方法, SPCOSA_CLHS 与仅考虑地理空间或属性空间的抽样方法相比, 可以更好地权衡地理空间均匀性约束和属性均匀性约束。

表 4 为点统计的均值和标准差结果, 每个耕地格点半径 5 km 的缓冲区平均覆盖 1.800~2.041 个样本点, SPCOSA_CLHS 方法的平均样本数高于 CLHS、XY_CLHS 和 RSM, 略低于 SPCOSA。从不同抽样方法点统计的标准差结果上看, SPCOSA 最小, 其次是 XY_CLHS 和 SPCOSA_CLHS。结果表明, 仅考虑地理空间的 SPCOSA 方法在样点空间分布均匀性上表现最好, 其次是 SPCOSA_CLHS 方法。

2.4 基于双变量局部空间自相关的抽样方法性能分析

图 4 为不同抽样结果的反距离权重插值结果与总体自然质量指数、技术水平指数、耕地产能指数数据双变量局部空间自相关分析结果, 以此来量化不同抽样结果对模拟总体耕地质量特征的准确性。

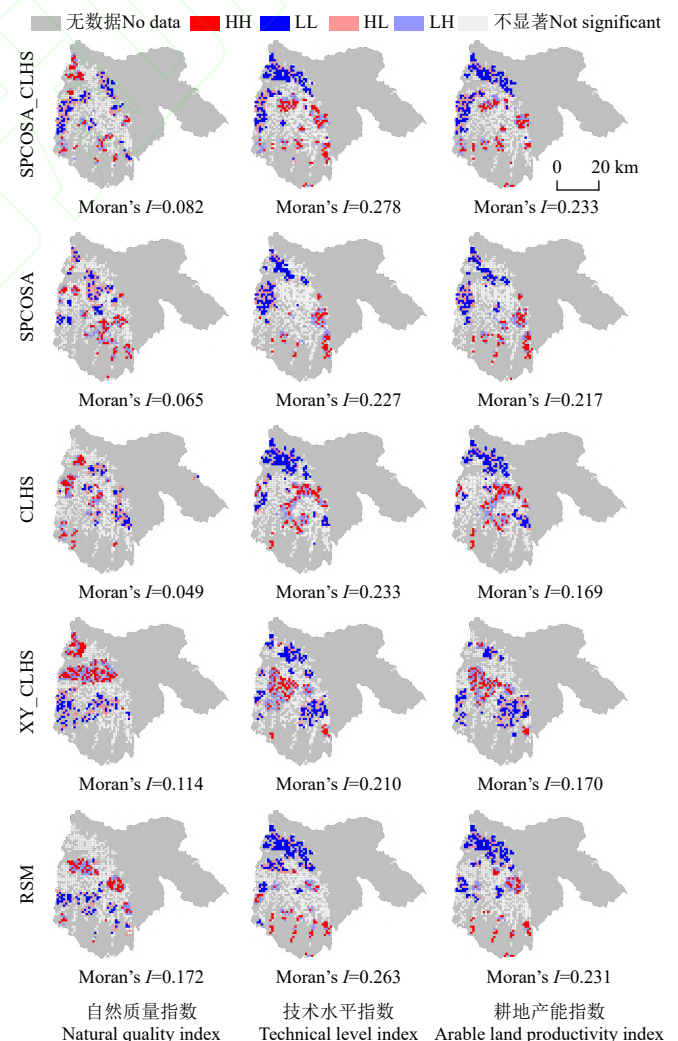


图 4 基于不同抽样方法的双变量局部空间自相关分析
Fig.4 Result of multivariate local spatial autocorrelation analysis based on multiple sampling methods

对于自然质量指数，基于 SPCOSA_CLHS 方法抽样结果的莫兰指数高于 SPCOSA 和 CLHS 方法，对于技术水平指数和耕地产能指数，基于 SPCOSA_CLHS 方法抽样结果的莫兰指数高于其他抽样方法。根据 LISA 聚类图，SPCOSA_CLHS 方法插值结果与总体数据局部空间自相关显著的耕地格点数量较多，并且正相关（HH、LL）的格点数多于负相关的格点数，说明 SPCOSA_CLHS 抽样结果的插值结果和总体数据在空间上表现出更高的正相关特征，可以更准确地模拟总体耕地质量特征。

3 讨论与结论

3.1 讨论

耕地质量抽样过程中，如果考虑耕地质量的多重属性特征，结果将获得更大的信息增益，如 CLHS 的抽样结果在减少样本量的情况下对总体属性特征具有更好的代表性和稳定性；如果考虑总体的空间分布特征，结果则可以反映总体的空间结构，如 SPCOSA 的抽样结果空间分布更加均匀；如果完全随机抽样，结果将具有较大的不确定性，RSM 的抽样结果对总体属性特征的代表性在样本量较多时表现良好，但随样本量的减小而剧烈波动。然而，样本的属性特征和空间分布特征在空间抽样中可能会发生冲突，因此权衡这两类特征很重要。SPCOSA_CLHS 以低空间约束的方式将空间信息整合到 CLHS 中，在表达总体属性特征、空间异质性和样本空间分布均匀性方面具有优势，更加适合作为县域耕地质量调查观测的空间抽样方法。

受限于数据的分辨率，本文使用的 1 km 格网不能很好地表达地形、土壤等细节信息，环境状况和生物特性指标也由于缺乏精细的数据在县级尺度的抽样过程中没有考虑。因此，本文与指导实地抽样之间仍有一定距离，更加侧重耕地质量空间抽样方法的实证对比，进一步设计抽样方案还需要更加高分辨率的准确数据支撑。本文应用的方法也同样适用于更高分辨率的耕地质量指标空间数据。同时，本文选取青藏高原地区青海省互助土族自治县作为研究区，进行耕地质量空间抽样方法的实证对比，研究结果可以为青藏高原及以外地区耕地质量抽样调查提供参考。未来，需要在更多地区以更高分辨率的数据开展关于多空间抽样方法的研究，在降低偶然性的同时，探索适用于不同地区的耕地资源质量空间抽样方法。

近几十年来，中国自然资源部数次开展全国耕地质量等级调查与评定工程，为认识中国耕地数量与质量格局提供重要数据成果，进而为耕地保护红线划定、耕地占补平衡政策实施提供信息支持。然而，受限于人工调查的数据收集模式成本高、耗时长缺点，该项工程在实现多维耕地质量指标长期调查观测方面尚存挑战，难以满足解析耕地质量变化过程、拐点及驱动因素，探索区域适宜的耕地质量提升策略的需求。在已有耕地质量指标调查成果数据驱动下，构建耕地质量调查样点布设方案，权衡样点调查成本与样点调查成果对总体特征的

代表性，进而发展基于参与式移动调查技术与便携式耕地资源质量指标观测装备，是提升耕地资源质量调查实施精度与效率的重要举措。另一项亟待发展的技术是基于长期耕地质量指标样点调查数据发展融合星空地的土壤性质及耕作条件指标遥感反演模型，进而推动区域尺度栅格单元耕地质量指标模拟。

3.2 结论

本文以青海省互助土族自治县为研究区开展基于耕地质量指标数据驱动的空间抽样实证研究，从信息熵、KL 散度、相似度距离、对总体耕地质量空间异质性的表征能力及样本空间分布均匀性等视角，对比分析 5 种不同空间抽样方法在研究区耕地质量指标调查观测点位布局中的优劣特征与适用性，并探索了研究区耕地资源质量监测适宜的点位数量。研究结论如下：

1) 与其他空间抽样方法相比，空间覆盖随机抽样与条件拉丁超立方体抽样的混合抽样法（spatial coverage sampling and random sampling-conditioned latin hypercube sampling, SPCOSA_CLHS）可以较低空间约束的方式将 SPCOSA 指示的空间异质性特征集成到 CLHS 模型中，在表达总体的耕地质量指标属性特征和空间异质性特征方面更具优势。

2) 当样点数量设置为 100~200 时，5 种抽样方法在表达总体耕地质量指标属性特征方面具有相似的适用性。随着样本量减少到 100 以下，SPCOSA 和随机抽样法（random sampling method, RSM）波动程度增加，基于 SPCOSA_CLHS 方法空间抽样方案的信息熵、KL 散度和相似度距离在收敛性和稳定性方面均表现出比其他方法更好的性能。

3) 当样点数量控制在 40~50 时，空间抽样结果能表现出与样本量在 100~200 时相似的总体耕地质量指标属性特征，并且此时 SPCOSA_CLHS 方法在表达耕地质量指标空间异质性、设计样点分布的空间均匀性、模拟总体耕地质量特征的准确性方面具有明显优势。综上，本研究表明在青海省互助土族自治县农田部署近 50 个样点，能以相对低的成本达到近 100~200 观测样点对研究区耕地质量指标总体特征的表征能力；SPCOSA_CLHS 方法相比其他空间抽样方法，在信息熵、KL 散度、相似度距离、对总体耕地质量空间异质性的表征能力、样本空间分布均匀性、模拟总体耕地质量特征的准确性方面具有优势，在研究区耕地质量观测点位设计方面具有更高的适用性。

本研究可以为青藏高原地区耕地资源质量调查监测工作提供方法支持，进而对理解该地区耕地资源质量变化过程、探索耕地可持续利用路径具有支持作用。同时，对青藏高原以外的其他地区耕地资源质量调查抽样具有参考意义。

[参 考 文 献]

- [1] YE S, SONG C, SHEN S, et al. Spatial pattern of arable land-use intensity in China[J]. Land Use Policy, 2020, 99: 104845.

- [2] 蔡运龙, 傅泽强, 戴尔阜. 区域最小人均耕地面积与耕地资源调控[J]. 地理学报, 2002, 57(2): 127-134.
CAI Yunlong, FU Zeqiang, DAI Erfu. The minimum area per capita of cultivated land and its implication for the optimization of land resource allocation[J]. *Acta Geographica Sinica*, 2002, 57(2): 127-134. (in Chinese with English abstract)
- [3] 叶思菁, 宋长青, 高培超, 等. 地理空间视角下耕地资源新认知体系构建[J]. 农业工程学报, 2023, 39(9): 225-240.
YE Sijing, SONG Changqing, GAO Peichao, et al. Cogstruction of the new cognitive system for arable land resources from geospatial perspective[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(9): 225-240. (in Chinese with English abstract)
- [4] 潘开文, 何燕, 唐亚伟, 等. 青藏高原农田生态系统与粮食安全科学考察研究专题进展[J]. 应用与环境生物学报, 2022, 28(4): 813-821.
PAN Kaiwen, HE Yan, TANG Yawei, et al. Progress of the scientific expedition and research on farmland ecosystem and food security in Qinghai-Tibet Plateau[J]. *Chinese Journal of Applied and Environmental Biology*, 2022, 28(4): 813-821. (in Chinese with English abstract)
- [5] 张丽萍, 张镜铨, 阎建忠, 等. 青藏高原东部山地农牧区生计与耕地利用模式[J]. 地理学报, 2008, 63(4): 377-385.
ZHANG Liping, ZHANG Yili, YAN Jiangzhong, et al. Livelihood diversification and agro-pastoral mountainous region cropland use pattern in of the eastern Tibetan Plateau[J]. *Acta Geographica Sinica*, 2008, 63(4): 377-385. (in Chinese with English abstract)
- [6] 崔鹏, 苏凤环, 邹强, 等. 青藏高原山地灾害和气象灾害风险评估与减灾对策[J]. 科学通报, 2015, 60(32): 3067-3077.
CUI Peng, SU Fenghuan, ZOU Qiang, et al. Risk assessment and disaster reduction strategies for mountainous and meteorological hazards in Tibetan Plateau[J]. *Chinese Science Bulletin*, 2015, 60(32): 3067-3077. (in Chinese with English abstract)
- [7] 刘玉洁, 吕硕, 陈洁, 等. 青藏高原农业现代化时空分异及其驱动机制[J]. 地理学报, 2022, 77(1): 214-227.
LIU Yujie, LV Shuo, CHEN Jie, et al. Spatio-temporal differentiation of agricultural modernization and its driving mechanism on the Qinghai-Tibet Plateau[J]. *Acta Geographica Sinica*, 2022, 77(1): 214-227. (in Chinese with English abstract)
- [8] 王梦溪, 潘开文, 伍小刚, 等. 青藏高原耕地面积时空变化特征及其驱动因子[J]. 应用与环境生物学报, 2022, 28(4): 859-868.
WANG Mengxi, PAN Kaiwen, WU Xiaogang, et al. Temporal and spatial patterns and driving factors of cultivated land change in Qinghai-Tibet Plateau[J]. *Chinese Journal of Applied and Environmental Biology*, 2022, 28(4): 859-868. (in Chinese with English abstract)
- [9] 姜芸, 王军, 滕浩, 等. 基于 TOPSIS 模型的典型黑土区耕地质量评价及土壤侵蚀耦合协调分析[J]. 农业工程学报, 2023, 39(12): 82-94.
JIANG Yun, WANG Jun, TENG Hao, et al. Coupling coordination analysis of the quality evaluation of cultivated land and soil erosion in typical black soil areas using TOPSIS method[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(12): 82-94. (in Chinese with English abstract)
- [10] 赵海乐, 徐艳, 张国梁, 等. 基于限制因子改良与耕地质量潜力耦合的耕地整治分区[J]. 农业工程学报, 2020, 36(21): 272-282.
ZHAO Haile, XU Yan, ZHANG Guoliang, et al. Farmland consolidation zoning based on coupling of improved limiting factors and farmland quality potential[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2020, 36(21): 272-282. (in Chinese with English abstract)
- [11] 孙晓兵, 孔祥斌, 张青璞, 等. 基于指标综合特征的耕地遗传质量和动态质量评价[J]. 农业工程学报, 2020, 36(22): 254-265.
SUN Xiaobing, KONG Xiangbin, ZHANG Qingpu, et al. Evaluation of inherent quality and dynamic quality of cultivated land based on comprehensive characteristics of indexes[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2020, 36(22): 254-265. (in Chinese with English abstract)
- [12] 卫新东, 林良国, 罗平平, 等. 耕地多功能耦合协调发展时空格局与驱动力分析[J]. 农业工程学报, 2022, 38(4): 260-269.
WEI Xindong, LIN Liangguo, LUO Pingping, et al. Spatiotemporal pattern and driving force analysis of multi-functional coupling coordinated development of cultivated land[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2022, 38(4): 260-269. (in Chinese with English abstract)
- [13] 陈红, 杨润佳, 叶艳妹. 中国土地资源的可耕性评价及其保护策略[J]. 农业工程学报, 2023, 39(5): 192-200.
CHEN Hong, YANG Runjia, YE Yanmei, et al. Cultivability evaluation and conservation strategies of land resources in China[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(5): 192-200. (in Chinese with English abstract)
- [14] 孙瑞, 金晓斌, 赵庆利, 等. 集成“质量-格局-功能”的中国耕地整治潜力综合分区[J]. 农业工程学报, 2020, 36(7):

- 264-275.
- SUN Rui, JIN Xiaobin, ZHAO Qingli, et al. Comprehensive zoning of cultivated land consolidation potential integrating quality-pattern-function in China[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(7): 264-275. (in Chinese with English abstract)
- [15] 杜国明, 闫佳秋, 张娜, 等. 面向多元主体需求的耕地质量体系新认知[J]. 农业工程学报, 2023, 39(1): 212-222.
- DU Guoming, YAN Jiaqiu, ZHANG Na, et al. New cognition of the cultivated land quality system for the needs of multiple subjects[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(1): 212-222. (in Chinese with English abstract)
- [16] 郭贯成, 韩小二. 考虑粮食安全和耕地质量的县域基本农田空间布局优化[J]. 农业工程学报, 2021, 37(7): 252-260.
- GUO Guancheng, HAN Xiao'er. Spatial layout optimization of basic farmland considering food security and cultivated land quality at county scale[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(7): 252-260. (in Chinese with English abstract)
- [17] 温良友, 孔祥斌, 张蚌蚌, 等. 基于可持续发展需求的耕地质量评价体系构建与应用[J]. 农业工程学报, 2019, 35(10): 234-242.
- WEN Liangyou, KONG Xiangbin, ZHANG Bangbang, et al. Construction and application of arable land quality evaluation system based on sustainable development demand[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(10): 234-242. (in Chinese with English abstract)
- [18] 孔祥斌, 张蚌蚌, 温良友, 等. 基于要素-过程-功能的耕地质量理论认识及其研究趋势[J]. 中国土地科学, 2018, 32(9): 14-20.
- KONG Xiangbin, ZHANG Bangbang, WEN Liangyou, et al. Theoretical framework and research trends of cultivated land quality based on elements-process-function[J]. China Land Science, 2018, 32(9): 14-20. (in Chinese with English abstract)
- [19] 张蚌蚌, 孔祥斌, 郟文聚, 等. 我国耕地质量与监控研究综述[J]. 中国农业大学学报, 2015, 20(2): 216-222.
- ZHANG Bengbeng, KONG Xiangbin, YUN Wenju, et al. A review on quality and monitoring of arable land in China[J]. Journal of China Agricultural University, 2015, 20(2): 216-222. (in Chinese with English abstract)
- [20] 叶思菁, 宋长青, 程锋, 等. 中国耕地健康产能综合评价与试点评估研究[J]. 农业工程学报, 2019, 35(22): 66-78.
- YE Sijing, SONG Changqing, CHENG Feng, et al. Cultivated land health-productivity comprehensive evaluation and its pilot evaluation in China[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(22): 66-78. (in Chinese with English abstract)
- [21] 宋长青, 程昌秀, 杨晓帆, 等. 理解地理“耦合”实现地理“集成”[J]. 地理学报, 2020, 75(1): 3-13.
- SONG Changqing, CHENG Changxiu, YANG Xiaofan, et al. Understanding geographic coupling and achieving geographic integration[J]. Acta Geographica Sinica, 2020, 75(1): 3-13. (in Chinese with English abstract)
- [22] YE S, ZHU D, YAO X, et al. Development of a highly flexible mobile GIS-based system for collecting arable land quality data[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 7(11): 4432-4441.
- [23] 叶思菁, 朱德海, 姚晓闯, 等. 基于移动 GIS 的作物种植环境数据采集技术[J]. 农业机械学报, 2015, 46(9): 325-334.
- YE Sijing, ZHU Dehai, YAO Xiaochuang, et al. Mobile GIS based approach for collection of crop planting environment data[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(9): 325-334. (in Chinese with English abstract)
- [24] 叶思菁, 张超, 王媛, 等. GF-1 遥感大数据自动化正射校正系统设计与实现[J]. 农业工程学报, 2017, 33(S1): 266-273.
- YE Sijing, ZHANG Chao, WANG Yuan, et al. Design and implementation of automatic orthorectification system based on GF-1 big data[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2017, 33(S1): 266-273. (in Chinese with English abstract)
- [25] 王媛, 叶思菁, 岳彦利, 等. 面向高分一号遥感影像的自动几何配准算法对比[J]. 农业机械学报, 2015, 46(S1): 260-266.
- WANG Yuan, YE Sijing, YUE Yanli, et al. Contrast of automatic geometric registration algorithms for GF-1 remote sensing image[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(S1): 260-266. (in Chinese with English abstract)
- [26] 叶思菁. 大数据环境下遥感图谱应用方法研究——以作物干旱监测为例[J]. 测绘学报, 2018, 47(6): 892.
- YE Sijing. Research on application of remote sensing Tupu: Take monitoring of meteorological disaster for example[J]. Acta Geodaetica et Cartographica Sinica, 2018, 47(6): 892. (in Chinese with English abstract)
- [27] 高秉博, 王劲峰, 胡茂桂, 等. 中国陆表自然资源综合观测台站布点优化[J]. 资源科学, 2020, 42(10): 1911-1920.
- GAO Bingbo, WANG Jinfeng, HU Maogui, et al. Optimization of integrated observation station layout for terrestrial surface natural resources[J]. Resources Science, 2020, 42(10): 1911-1920. (in Chinese with English abstract)
- [28] ZARTMAN R. Sampling for natural resource monitoring[J].

- Publications of the American Statistical Association, 2006, 103(2): 889-890.
- [29] VAN G, STEIN A. Constrained optimization of spatial sampling using continuous simulated annealing[J]. Journal of Environmental Quality, 1998, 27(5): 1078-1086.
- [30] VANGROENIGEN J. Spatial simulated annealing for optimizing sampling - Different optimization criteria compared[Z]. GeoENV I - Geostatistics for Environmental Applications. 1997: 351-361.
- [31] TANG S, YANG J, CHEN Y, et al. Spatial sampling design for monitoring the quality of cultivated land by simulated annealing[J]. Sensor Letters, 2014, 12(3): 1-6.
- [32] BRUS D, HEUVELINK G. Optimization of sample patterns for universal kriging of environmental variables[J]. Geoderma, 2007, 138(1): 86-95
- [33] 杨建宇, 岳彦利, 宋海荣, 等. 基于空间模拟退火算法的耕地质量布样及优化方法[J]. 农业工程学报, 2015, 31(20): 253-261.
YANG Jianyu, YUE Yanli, SONG Hairong, et al. Sampling and optimizing methods of cultivated land quality based on spatial simulated annealing algorithm[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2015, 31(20): 253-261. (in Chinese with English abstract)
- [34] 杨建宇, 岳彦利, 宋海荣, 等. 基于空间平衡法的县域耕地质量监测布样方法[J]. 农业工程学报, 2015, 31(24): 274-280.
YANG Jianyu, YUE Yanli, SONG Hairong, et al. Sampling distribution method for monitoring quality of arable land in county area based on spatial balanced[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2015, 31(24): 274-280. (in Chinese with English abstract)
- [35] MCKAY M, BECKMAN R, CONOVER W. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code[J]. Technometrics, 1979, 21(2): 239-245.
- [36] MINASNY B, MCBRATNEY A. A conditioned Latin hypercube method for sampling in the presence of ancillary information[J]. Computers & Geosciences, 2006, 32(9): 1378-1388.
- [37] GAO B, PAN Y, CHEN Z, et al. A spatial conditioned Latin hypercube sampling method for mapping using ancillary data[J]. Transactions in GIS, 2016, 20(5): 735-754.
- [38] WANG J, LIU J, ZHUAN D, et al. Spatial sampling design for monitoring the area of cultivated land[J]. Land Degradation & Development, 2002, 23(2): 263-284.
- [39] WAN C, KUZYAKOV Y, CHENG C, et al. Soil sampling design for arable land quality observation by using SPCOSA-CLHS hybrid approach[J]. Land Degradation & Development, 2021, 32(17): 4889-4906.
- [40] 张海峰, 刘林山, 王有宁, 等. 青藏高原东北部互助县域土地利用变化对粮食安全的影响[J]. 地域研究与开发, 2012, 31(2): 95-100.
ZHANG Haifeng, LIU Linshan, WANG Youning, et al. Impact of land-use changes on food safety in Huzhu County, NE Tibetan Plateau[J]. Areal Research and Development, 2012, 31(2): 95-100. (in Chinese with English abstract)
- [41] 孔祥斌, 张青璞. 中国西部区耕地等别空间分布特征[J]. 农业工程学报, 2012, 28(22): 1-7.
KONG Xiangbin, ZHANG Qingpu. Spatial distribution characteristics of arable land grade in Western China[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2012, 28(22): 1-7. (in Chinese with English abstract)
- [42] ABRAMS M, YAMAGUCHI Y, CRIPPEN R. ASTER global DEM (GDEM) version 3 [Z]. XXIV ISPRS Congress Imaging Today, Foreseeing Tomorrow, Commission IV. 2022, 43-B4, 593-598.
- [43] LIU F, WU H, ZHAO Y, et al. Mapping high resolution national soil information grids of China[J]. Science Bulletin, 2022, 67(3): 328-340.
- [44] LIU C, SONG C, YE S, et al. Estimate provincial-level effectiveness of the arable land requisition-compensation balance policy in China's mainland in the last 20 years[J]. Land Use Policy, 2023, 131: 106733.
- [45] 王劲峰, 徐成东. 地理探测器: 原理与展望[J]. 地理学报, 2017, 72(1): 116-134.
WANG Jinfeng, XU Chengdong. Geodetector: Principle and prospective[J]. Acta Geographica Sinica, 2017, 72(1): 116-134. (in Chinese with English abstract)
- [46] WALVOORT D, BRUS D, GRUIJTER J. An R package for spatial coverage sampling and random sampling from compact geographical strata by k-means[J]. Computers & Geosciences, 2010, 36(10): 1261-1267.
- [47] SHANNON C. A mathematical theory of communication[J]. Bell Systems Technical Journal, 1948, 27(4): 623-656.
- [48] HYNDMAN R, FAN Y. Sample quantiles in statistical packages[J]. The American Statistician, 1996, 50(4): 361-365.
- [49] ANSELIN L. Local indicators of spatial association—LISA[J]. Geographical Analysis, 1995, 27(2): 93-115.
- [50] ANSELIN L. A local indicator of multivariate spatial association: Extending Geary's c[J]. Geographical Analysis, 2018, 51(2): 133-150.

Comparison and application of the spatial sampling methods for assessing the quality of arable land resources in Qinghai-Tibet Plateau

JIANG Jiayi^{1,2}, SONG Changqing^{1,2}, YE Sijing^{1,2*}, GAO Peichao^{1,2}, REN Shuyi²

(1. State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China;

2. Faculty of Geographical Science, Beijing Normal University, Beijing, 100875, China)

Abstract: Spatial patterns and temporal variation of arable land resources can greatly contribute to regional protection and sustainable utilization, particularly for food security and social stability in the Qinghai-Tibet Plateau. A spatial sampling with fewer points is required to comprehensively characterize the overall features of regional arable land. Long-term observation is also of great significance in evaluating the quality of arable land resources. In this study, a data-driven spatial sampling was presented to determine the indicators of arable land quality in the Tu Autonomous Country of Huzhu, Qinghai Province, China. The grid units of the sample population were set as 1 km of arable land, with a total of 790 points. 21 indicators were extracted from three dimensions of topographic features, soil properties, and tillage technical conditions. An indicator system was then constructed for the sampling of arable land at the point scale, such as the slope, soil bulk density, organ carbon content, and agriculture mechanization level. The performance of spatial sampling was quantified by the natural quality, technical level, and arable land productivity index, which were collected from the pilot project of the Ministry of Natural Resources of China. Local spatial heterogeneity was represented to simulate the accuracy of the overall quality of arable land. Multiple indicators were also calculated from five dimensions of topographic features, soil properties, tillage technical conditions, environmental conditions and biological characteristics, in order to evaluate the quality and productivity of arable land. RSM (random sampling), SPCOSA (spatial coverage sampling and random sampling), CLHS (conditioned Latin hypercube sampling), XY_CLHS (CLHS with x and y coordinates as covariates), and SPCOSA_CLHS (spatial coverage sampling and random sampling-conditioned Latin hypercube sampling) were compared from the perspectives of information entropy, Kullback–Leibler divergence, similarity distance, the spatial heterogeneity and distribution uniformity of samples in the overall quality of arable land. A survey of observation points was carried out on the indicators of arable land quality in the county-level areas. Entropy-based tests were also performed on each sampling using the nine groups with the multigroup sample sizes: 10, 20, 30, 40, 50, 75, 100, 150, and 200. A suitable number of points were explored in the study area. The results show that the SPCOSA_CLHS model integrated the SPCOSA into the CLHS model, and then represented the spatial heterogeneity with the lower spatial constraints. A better performance was achieved to express the overall index attribute and spatial heterogeneity of arable land quality. When the number of sampling points was between 100 and 200, five sampling models shared similar applicability for the attribute in the overall quality indicators of arable land. Once the sample size dropped below 100, there were the greatest differences among sampling models, where the SPCOSA and RSM offset most. A better performance of SPCOSA_CLHS was also obtained in the information entropy, KL divergence, and similarity distance, in terms of convergence and stability. When the number of sample points was 40-50, a similar spatial sampling was observed as the sample size of 100-200. Therefore, SPCOSA_CLHS can be expected to describe the spatial heterogeneity of arable land quality indicators, the spatial uniformity of sample point distribution, and the simulated accuracy of the overall quality. This finding can provide strong support to the survey and monitoring of arable land quality in the Qinghai-Tibet Plateau. In turn, the evolution of arable land quality can also be used to explore the sustainable use pathways of arable land.

Keywords: land use; resources; arable land quality; spatial sampling; Qinghai-Tibet Plateau; GIS