

doi:10.6041/j.issn.1000-1298.2022.08.029

# 基于回归克里格法的土壤盐分采样点布局优化

徐英 谢若禹 沈丽佳 冯绍元

(扬州大学水利科学与工程学院, 扬州 225009)

**摘要:** 为提高土壤盐分分布图绘制精度或降低野外采样和室内分析成本,优化土壤采样点布局,本文以回归克里格法估计方差(RKV)最小和普通克里格估计方差(OKV)最小为目标函数,运用空间模拟退火(SSA)算法对土壤盐分采样点布局进行优化,其中回归克里格法(RK)中的协变量选择与土壤盐分分布关系密切且容易获取的环境变量;通过最邻近距离分布函数( $G$ 函数)、间隔距离分布函数( $F$ 函数)和K-S检验对优化结果进行评价。结果表明:荒地分布在大尺度上( $400\text{ m} \times 400\text{ m}$ )与盐分相关性强,而在小尺度上( $50\text{ m} \times 50\text{ m}$ )与盐分相关性弱;以荒地分布(与荒地最短距离)作为协变量,利用RK法优化采样点布局(采样数为55)后得到的研究区平均RKV为0.0807,相较平均OKV(0.1143)缩小了29.39%;优化后采样点布局的 $G$ 函数曲线和 $F$ 函数曲线全部落在95%置信区间内,即 $F$ 函数和 $G$ 函数均与其理论值无显著差异,采样点布局良好覆盖了地理空间;K-S检验表明,优化后采样点上与荒地最短距离的分布与其总体分布服从同一概率分布,即采样点布局良好覆盖了属性空间。本研究可为河套灌区土壤盐渍化监测点布局提供参考。

**关键词:** 土壤采样优化; 土壤盐分; 回归克里格估计方差; 空间模拟退火算法

中图分类号: S152.1; S126

文献标识码: A

文章编号: 1000-1298(2022)08-0275-08

OSID:



## Layout Optimization of Soil Salt Sampling Points Based on Regression Kriging

XU Ying XIE Ruoyu SHEN Lijia FENG Shaoyuan

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

**Abstract:** In order to improve the drawing accuracy of soil salt distribution map or reduce the cost of field sampling and indoor analysis, it is very necessary to optimize the layout of soil sampling points. Taking the minimum regression Kriging estimated variance (RKV) and the minimum ordinary Kriging estimated variance (OKV) as the objective function, the layout of soil salt sampling points was optimized by spatial simulated annealing algorithm (SSA), the selection of covariates in regression Kriging (RK) method was closely related to soil salt distribution and easy to obtain environmental variables. The optimization results were evaluated by the nearest distance distribution function ( $G$ -function), interval distance distribution function ( $F$ -function) and K-S test. The results showed that wasteland distribution on a large scale ( $400\text{ m} \times 400\text{ m}$ ) was significantly correlated with salt, but weak correlated on a small scale ( $50\text{ m} \times 50\text{ m}$ ). Taking the distribution of wasteland (the shortest distance from wasteland) as the collaborative variable and optimizing the sampling layout by RK method (sampling number was 55), the mean RKV of the study area was 0.0807, which was 29.39% smaller than the mean OKV(0.1143). The  $G$ -function curve and  $F$ -function curve of the optimized sampling set all fall within the 95% confidence interval, there was no significant difference between the  $F$ -function and  $G$ -function and their theoretical values, the layout of the sampling points was well arranged and covered the geographical space. K-S test showed that the distribution of the shortest distance from wasteland on the optimized sampling point and its population obeyed the same probability distribution, the layout of the sampling points was well arranged and covered the attributes space. The research results can provide reference for the layout of soil salinization monitoring points in Hetao Irrigation Area.

**Key words:** soil sampling optimization; soil salinity; regression Kriging estimated variance; spatial simulation annealing algorithm

收稿日期: 2021-09-12 修回日期: 2021-10-17

基金项目: 国家重点研发计划项目(2017YFC0403301)

作者简介: 徐英(1972—),女,副教授,主要从事农业水土环境与灌排理论研究, E-mail: xuying@yzu.edu.cn

## 0 引言

土壤盐渍化是干旱、半干旱区农业和生态环境可持续发展的主要制约因素之一。掌握土壤盐渍化时空分布规律是防治土壤盐渍化的前提条件,而绘制土壤特性分布图是表达土壤特性空间分布最有效的途径之一<sup>[1-2]</sup>。目前,获取土壤盐分空间分布图的主要途径有3种:基于大量野外采样数据,借助空间插值方法绘制<sup>[3]</sup>;通过遥感技术反演盐分分布图<sup>[4]</sup>;利用遥感数据以及过往资料与影像等作为辅助变量(亦称为协变量),利用空间插值方法绘制<sup>[5-6]</sup>。无论哪种途径,都需要一定量的土壤盐分实测数据用以建模或验证。有研究表明,土壤特性分布图的全部误差由采样误差和数据分析误差组成,其中采样误差占全部误差的90%<sup>[7]</sup>,而80%~85%的采样误差源于样本的代表性不足<sup>[8]</sup>。可见,优化采样设计对提高土壤盐分分布图绘制精度或降低野外采样和室内分析成本具有重要意义。

采样设计包括确定采样数量和采样点分布。采样数量通常由预算经费、允许采样时间以及可用劳力决定,为保证一定的绘图精度,优化采样点空间布局显得尤为重要<sup>[7]</sup>。根据目的不同,采样设计还要考虑不可采样区域(如本研究区的居民区)、如何利用以往采样数据以及其他辅助信息等约束条件,在绘图精度和采样成本之间取得平衡<sup>[9]</sup>。在具有土壤实测数据或先验方差的情况下,研究者们利用克里格估计方差只依赖于采样点空间布局、与实际取值无关的特点,以平均普通克里格估计方差(Ordinary Kriging estimated variance, OKV)最小(采样数固定)为准则优化土壤采样布局<sup>[10-12]</sup>。普通克里格法(Ordinary Kriging, OK)采样设计能很好地表达出研究变量的地理空间变异<sup>[11]</sup>,变量均质性越强,采样也越趋于规则性空间分布,有利于表达变程范围内主变量的变异,但因不能很好地覆盖受环境因素(包括人类活动相关的因素)约束或影响的属性空间,故对特异值极其不敏感<sup>[12]</sup>。与此相反,考虑土壤特性与各种协变量(包括环境变量、地理坐标、遥感数据等)之间的线性、非线性关系,以平均估计方差最小为准则优化采样点布局,可得到更趋于聚集性空间分布的样本<sup>[12-13]</sup>。这样的采样点布局虽然对地理空间覆盖性较差,但适合表征变量在属性空间(由各协变量所代表的属性空间)的变异,有助于包含更多极值点,进而降低外延误差<sup>[14-15]</sup>,对于应用遥感数据优化采样过程非常有效<sup>[16]</sup>。因土壤特性同时受成土过程和环境因素的影响,以绘制土壤特性分布图为目的的采样布局,不仅需要良

好的覆盖地理空间,同时需要在属性空间有良好的分布。为此,研究者们引入各种环境因子作为协变量,用协变量与土壤特性之间的回归关系表达属性空间分布,用残差表达地理空间分布,试图在二者之间寻找平衡。BRUS等<sup>[13]</sup>将环境因子作为协变量,以平均回归克里格估计方差(Regression Kriging estimated variance, RKV)最小为准则,通过空间模拟退火(Spatial simulated annealing, SSA)算法优化了荷兰 Noord - Brabant 省某区域地下水位动态监测网,结果表明回归克里格法(Regression Kriging, RK)得到的监测点布局能在属性空间和地理空间之间取得很好的平衡,估值精度也比 OK 法和多元线性回归法更高。SZATMARI等<sup>[9]</sup>以地面高程和土地利用类型为辅助变量,在匈牙利中部 Mezőföld 地区优化了土壤有机质采样点布局,结果也表明 RK 法优化得到的采样点布局能很好地覆盖属性空间和地理空间,同时 SSA 允许优化中考虑不可采样区域等约束条件。尽管 RK 法被广泛地用于优化各种土壤特性的采样点布局,但针对土壤盐分采样点布局优化的研究很少。

受降雨、蒸发、地下水位、高程、土质等自然因素与灌溉、排水、荒地分布、作物布局等人类活动的共同影响,内蒙古河套灌区土壤盐分时空分布格局在不断发生变化。为了更好地进行盐渍化防治和盐渍化土壤管理,建立相对稳定的土壤盐分监测网络非常必要。为满足各种插值方法的需求,土壤盐分采样在保证插值精度、降低成本的前提下,其样点分布应在属性空间和地理空间均有较好的代表性。为此,本文拟引入协变量,用 RK 法(即平均 RKV 最小化为优化准则,以 SSA 为寻优方法)对永济灌域一研究区进行土壤盐分采样布局优化,以期河套灌区土壤盐渍化监测提供参考。

## 1 材料与方 法

### 1.1 研究区概况与数据采集

#### (1) 研究区概况

内蒙古河套灌区永济灌域属于温带大陆性气候,干旱少雨,蒸发强烈,因大量引黄灌溉造成地下水位埋深较浅,土地盐碱化严重。

研究区(东经 108°49' ~ 108°52', 北纬 40°48' ~ 40°52')位于内蒙古河套灌区永济灌域合济渠控制范围内,如图 1 所示,研究区南北约 7 km,东西约 4 km,面积约为 1 574 hm<sup>2</sup>。研究区平均地下水埋深为 1.7 m;高程在 1 040 ~ 1 041 m 之间且地势平缓(自然坡降平均为 1/5 000),区内土壤以粉砂质壤土为主,部分区域为砂质壤土;研究区多年平均降水

量 125.2 mm, 多年蒸发量在 1 999 ~ 2 346 mm 之间; 平均气温为 8.3℃, 年日照时数 3 100 h 左右; 主要作物有小麦、玉米、葵花、瓜菜等。

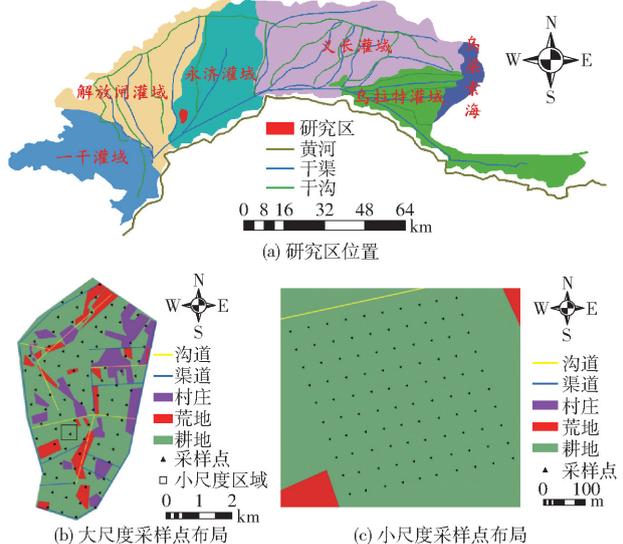


图 1 研究区位置及样点分布

Fig. 1 Distributions of study area and sampling point

(2) 土样采集

采用嵌套式采样, 即大尺度嵌套小尺度, 大尺度采样范围(下称大尺度区)即为整个研究区, 小尺度采样区(下称小尺度区)位于研究区的中部偏南, 面积约为 20.25 hm<sup>2</sup>。

在谷歌地图上分别勾勒出大尺度区和小尺度区, 以 400 m × 400 m 规则网格在大尺度区布置 89 个采样点(图 1b), 以 50 m × 50 m 规则网格在小尺度区布置 100 个点位(图 1c), 注意点位避开村庄和道路。所有采样点采样深度均为 0 ~ 10 cm、10 ~ 20 cm、20 ~ 40 cm(每个采样点重复 2 次), 以 0 ~ 40 cm 土层平均盐分进行研究。采样时间为 2020 年 9 月 1 日(秋浇前)。

(3) 数据测定

盐分测定: 将土样自然风干、研磨、过 2 mm 筛, 按土水质量比 1:5 进行土壤浸提、振荡、过滤, 土壤溶液的电导率测定使用 DDSJ-308A 型电导率仪, 并以经验公式换算成土壤全盐量, 公式为

$$M = 2.882EC_{1:5} + 0.183$$

式中  $M$ ——全盐量, g/kg

$EC_{1:5}$ ——土水质量比 1:5 电导率, mS/cm

与荒地距离测定: 在谷歌地图中勾勒出研究区荒地的边界, 将其导入 ArcGIS 10.0 软件中计算各网格结点与荒地边界的最短距离。

与沟道距离测定: 在谷歌地图中画出研究区沟道, 将其导入 ArcGIS 10.0 软件中计算各网格结点与沟道的最短距离。

质地测定: 使用激光粒度分析仪(马尔文 Mastersizer 2000 型)测定土壤颗粒组成, 土粒分级采用美国制分级标准(粒径小于 0.002 mm 为粘粒; 粒径 [0.002 mm, 0.05 mm] 为粉粒; 粒径 (0.05 mm, 2 mm) 为砂粒)。

1.2 研究方法

本文基于 RK 法计算研究区内平均估计方差, 并使用 SSA 寻找平均估计方差全局最小的采样点布局, 最后对寻找到的采样点布局进行评价。

1.2.1 回归克里格法(RK)

RK 是一种混合插值方法, 由线性回归和 OK 法结合而成, 即通过建立协变量(通常为容易获得的环境变量)和主变量(所研究变量)之间的回归方程, 分离趋势项, 并对残差进行 OK 插值, 最后将回归预测的趋势项和残差的普通克里格估计值进行空间叠加, 从而得到未采样点  $x_0$  处主变量的估值。因此, RK 法中趋势项表达了协变量对主变量空间变异的影响, 残差估值反映了去除协变量影响后的主变量空间变异。此时,  $x_0$  处的估计方差计算式为<sup>[17]</sup>

$$\sigma_{RK}^2 = (c(0) - c_0^T \cdot C^{-1} \cdot c_0) + (q_0 - q^T \cdot C^{-1} \cdot c_0)^T (q^T \cdot C^{-1} \cdot q)^{-1} \cdot (q_0 - q^T \cdot C^{-1} \cdot c_0) \quad (1)$$

式中  $\sigma_{RK}^2$ ——回归克里格估计方差(RKV)

$c(0)$ ——残差的方差

$c_0$ ——采样点和未采样点残差之间的协方差向量

$C$ ——采样点残差的方差-协方差矩阵

$q$ ——采样点的协变量矩阵

$q_0$ ——未采样点的协变量向量

式(1)整合了回归方差和残差的普通克里格估计方差, 因此, RKV 既反映了协变量代表的属性空间插值精度, 又反映了地理空间的插值精度。采样点布局优化问题可以转换为在采样空间内寻求平均 RKV 最小的采样集。

为了评价 RK 法在优化土壤盐分采样布局方面的表现, 以 OK 法(即寻求 OK 估计方差平均值最小)作为对照方法。普通克里格估计方差计算见文献[18]。

1.2.2 空间模拟退火算法(SSA)

SSA 是目前最常用的空间寻优算法之一, 其运行原理是模拟固体退火过程: 固体受热熔解为液体后冷却时, 粒子运动随温度下降而逐渐有序, 最后围绕晶体格点做微小振动。退火过程中, 系统能量随温度降低趋于最小值。

本文将研究区按照 50 m × 50 m 离散成 6 352 个网格节点, 去除村庄内的节点剩余 5 187 个节点, 以

这些节点为潜在采样点进行优化。

使用 SSA 将从一个初始解(采样方案)出发(初始解在已知采样点中随机抽取生成),经过大量扰动和概率接受后,最终求得满足目标函数的全局最优解。

### (1) 温度参数

温度参数决定了算法的循环次数,本研究中初始温度  $T_0 = 100^\circ\text{C}$ , 结束温度  $T_e = 0.1^\circ\text{C}$ , 衰减系数  $\alpha$  为 0.995。

### (2) 目标函数

目标函数代表退火过程中的系统能量,SSA 的目标就是找到目标函数全局最小的解。本研究目标函数设为研究区内平均 RKV,其公式为

$$\phi(S) = \sum_{j=1}^{n_e} \frac{\sigma^2(x_{e,j}|S)}{n_e} \quad (2)$$

式中  $\sigma^2$ ——研究区内潜在在采样点处的 RKV

$x_{e,j}$ ——研究区域内采样点的点位

$S$ ——当前采样点组合

$n_e$ ——潜在采样点数

以平均 RKV 最小为目标函数可以同时在地空间特征空间内对采样点布局进行优化,同时 RKV 的计算独立于采样点的观测值,可以在实际采样前计算<sup>[17]</sup>。

### (3) Metropoils 准则

Metropoils 准则定义了扰动后新解被接受的概率  $P_c$ , 计算见文献[10],本文中  $c$  初始值为 0.025, 衰减系数为 0.995。Metropoils 准则使得 SSA 从一开始能以一定概率接受劣化解,随着算法运行,接受劣化解的概率逐渐较小,这样避免优化结果陷入局部最优。

### (4) 扰动

扰动的作用是在当前解的基础上产生一个新解,扰动的原理是在当前解中任取一采样点,以该点为圆心,扰动范围  $h$  为半径画圆,将原采样点替换为圆内随机一个潜在采样点。 $h$  初始值取研究区最大长度,本文  $h$  初始值为 6 000 m,  $h$  随温度衰减而衰减,衰减系数为 0.995。

## 1.2.3 合理采样数计算

根据经典统计学,使用 Cochran 公式确定土壤盐分的合理采样数,计算公式为

$$n' = \frac{t_a^2 \sigma^2}{(\bar{X}K)^2} \quad (3)$$

式中  $n'$ ——合理采样数

$t_a$ ——置信度为  $a$  时的概率度

$\sigma$ ——样本标准差

$\bar{X}$ ——样本平均值

$K$ ——采样允许误差, %

## 1.2.4 采样布局评价方法

为评价 RK 法在土壤盐分采样布局优化中的表现,本文除了通过与 OK 法比较,还通过  $G$  函数、 $F$  函数和  $K-S$  检验来评价优化结果的代表性。

$G$  函数即为最邻近距离分布函数(Nearest neighbour distribution function),  $F$  函数为间隔距离分布函数(Empty space distribution function)。 $G$  函数与  $F$  函数常一起分析样点的空间分布情况,检验它们对地理空间的覆盖程度或代表性。 $G$  函数描述了从任意采样点到其最近采样点的距离分布。将优化采样集(本文指优化后的采样数和采样位置)的  $G$  函数与其理论分布(本文在潜在采样点中随机抽取 99 组样本近似计算  $G$  函数的 95% 置信限,样本容量与优化后采样点数相同)比较,优化采样集的  $G$  函数值相对理论值越大,表明样点越趋于聚集性空间分布,反之则样点趋于规则(均匀)空间分布。 $F$  函数描述空间任意点(潜在采样点)到其最近采样点的距离分布。仍然将优化采样集的  $F$  函数与相应的理论分布(获取方法与  $G$  函数相似)比较,优化采样集  $F$  函数值相对理论值越小,表明研究区域内的间隔空间越少,样点越趋于聚集性空间分布,反之则表明样点趋于规则(均匀)空间分布。 $G$  函数和  $F$  函数计算式为

$$F' = \frac{\sum_{i=1}^n I(r_i \leq r)}{n} \times 100\% \quad (4)$$

其中

$$I(r_i \leq r) = 1$$

式中  $F'$ —— $G$  或  $F$  函数, %

$r$ ——预先设定的距离

$I(\cdot)$ ——指示函数

$n$ ——采样数( $G$  函数)或潜在采样点数( $F$  函数)

$r_i$ ——第  $i$  个采样点到其他采样点的最小距离( $G$  函数)或第  $i$  个潜在采样点到采样点的最小距离( $F$  函数)

科尔莫戈罗夫-斯米尔诺夫检验(Kolmogorov-Smirnov test,  $K-S$ )常用于检验两组数据是否来自同一个分布。本文用于检验协变量样本(由优化后采样点上的协变量取值组成)对协变量总体分布(潜在采样点上协变量取值估计)的代表性,即采用  $K-S$  法检验优化采样集对属性空间的覆盖程度或代表性。

本文变异函数模型通过  $GS^+$  软件确定,其他计算均用 Java 语言编程完成;用 Excel 绘制变异函数图,其他图形在 ArcGIS 10.0 平台上完成。

## 2 结果与分析

### 2.1 协变量选择

#### 2.1.1 环境变量与土壤含盐量的相关关系

协变量通常选择与主变量(本研究为土壤含盐量)关系密切且较容易获取的环境变量。本文初步确定的备选变量为土质(本研究区以粉砂质壤土为主,土质用砂粒含量表示;无采样数据时,可用当地的数字化土壤地图)、沟道分布(用与沟道的最短距离表征)和荒地分布(用与荒地的最短距离表征)。因研究区面积不大,且地形平坦,以往研究也表明研究区高程对土壤盐分分布影响不明显<sup>[19]</sup>,因此,备选变量未包括高程。不同采样尺度上,3个备选变量与土壤含盐量的皮尔逊相关系数见表 1。由表 1 可以看出,大尺度(采样分辨率为 400 m × 400 m)上土壤含盐量与荒地分布、土质均呈极显著相关,且荒地分布与盐分分布关系更密切,相关系数为 -0.509;小尺度(采样分辨率为 50 m × 50 m)上盐分分布受荒地分布影响较小(相关系数仅为 -0.057),受土质影响最大,其次为与沟道最短距离,且相关系数均达到极显著水平。可见,不同尺度上影响土壤盐分分布的主要环境变量不同,这表明协变量选择受采样尺度的影响。

表 1 各环境变量与土壤含盐量的皮尔逊相关系数

Tab.1 Pearson correlation coefficient between environment variable and soil salt

尺度	砂粒含量	与沟道最短距离	与荒地最短距离
大尺度	-0.456 **	( $P < 0.001$ )	-0.509 **
小尺度	-0.539 **	-0.314 **	-0.057

注:\*\*表示在 0.01 水平(双侧)上极显著相关。

#### 2.1.2 研究区协变量确定以及 RK 法中实测点处残差计算

本文主要研究大尺度上采样点的优化布局,小尺度采样仅是为了提高变异函数的计算精度,因此,选取在大尺度上与土壤盐分相关系数最大的与荒地最短距离为协变量,不仅因为它在大尺度上与盐分关系最密切,而且因为该变量获取非常容易。此时,因为小尺度上盐分与荒地分布关系不密切,故而,在 RK 法中,仅采用大尺度网格结点上的土壤含盐量和与荒地最短距离确定回归方程( $y = -0.4036x$ ,  $x$  和  $y$  分别表示与荒地最短距离和土壤含盐量),分离趋势项,计算所有点(包括大尺度和小尺度采样点)的残差。

图 2 为残差的变异函数。图 2 表明,所有点(189 个点)残差的变异函数可较好地用指数模型拟合,小尺度上点的参与能更细致地刻画较小滞后距

上的变异特征,且图形比较连续、无明显突变;与仅用大尺度数据计算的残差变异函数相比,小尺度数据加入后,块金值缩小 42.9%,块金值/基台值缩小 32.2%(表 2)。可见,用大尺度上的回归方程分离趋势项、计算残差是合理的。

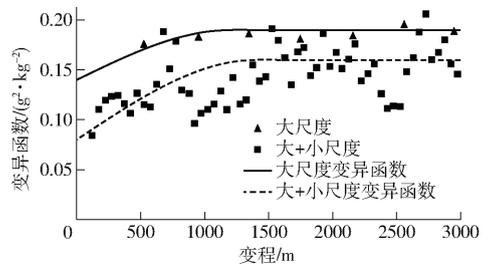


图 2 土壤盐分残差变异函数变化曲线

Fig.2 Variogram of soil salt residuals

表 2 变异函数模型参数

Tab.2 Variogram model parameters

采样尺度	理论模型	块金值	基台值	块金值/基台值/%	变程/m
大尺度	球状模型	0.14	0.19	73.7	1200
大+小尺度	球状模型	0.08	0.16	50.0	1400

需说明的是:因为土壤盐分符合对数正态分布(文中未列出有关计算),因此,所有计算均是基于实测土壤含盐量对数变换后的值进行的。

### 2.2 优化结果与评价

#### 2.2.1 合理采样数确定

由式(3)可知,精度越高,对应的采样数越多。计算结果表明在 95% 置信度下相对误差分别为 5%、10% 和 15% 时,合理采样数分别为 77、55 和 37。

初步选择合理采样数为 55,但该采样数仅是从统计学角度出发,为了保证该采样数具有更好的合理性,绘制了不同数量样点(从 89 个已知点中随机抽取)用于回归克里格插值时所有潜在采样点平均估计方差的变化曲线(图 3)。由图 3 可以看出,随着样点数的增加,平均估计方差逐渐减小并趋于平缓,样点数为 55 时的平均估计方差相对于大尺度 89 个采样点时的平均估计方差增大 6.7%。因此认

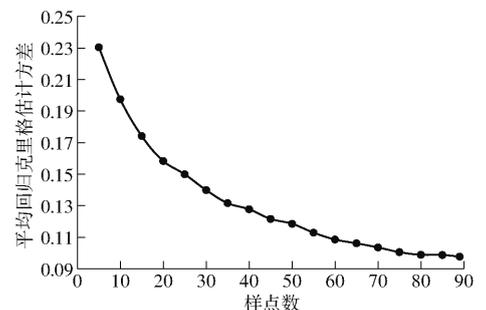


图 3 不同采样数的平均回归克里格估计方差变化曲线

Fig.3 Mean RKV based on different sampling numbers

为图3选取55为优化方案采样数合理。

### 2.2.2 SSA优化结果

最终优化后采样点布局如图4所示,分别以采样方案的平均RKV最小和平均OKV最小为目标函数,对采样方案进行优化。潜在采样点的平均RKV和平均OKV分别为0.0807和0.1143,RK法的估计方差估值精度相较OK法提高了29.39%。这表明以荒地分布作为协变量,采用RK法优化采样布局,可以有效提高土壤盐分估值精度。

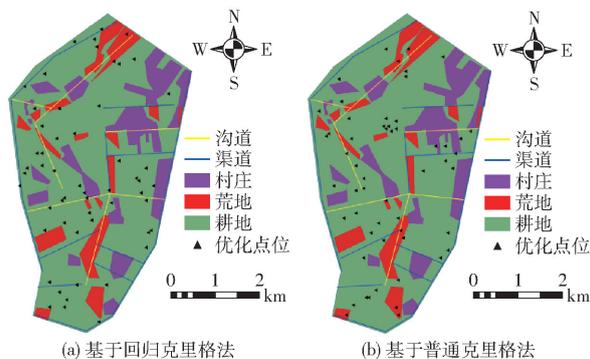


图4 SSA优化采样点布局图

Fig.4 Optimized sampling point layout by SSA

### 2.2.3 采样布局评价

从估计方差的全局最优性、采样点对地理空间和属性空间的代表性3方面,对基于RK法的优化结果进行评估。

(1)评估优化结果的全局最优性。本文跟踪了优化过程中目标函数值的变化。由图5可以看出,随着温度衰减次数增加,平均估计方差的总体趋势下降,在下降过程中一直上下波动,最终趋于平稳。估计方差的上下波动体现了SSA接受劣化解的能力,避免了优化结果陷入局部最优的状况,随着温度衰减,SSA接受劣化解的概率逐渐下降,平均估计方差仅在小范围内波动,最后不变。这表明优化结果具有全局最优性。

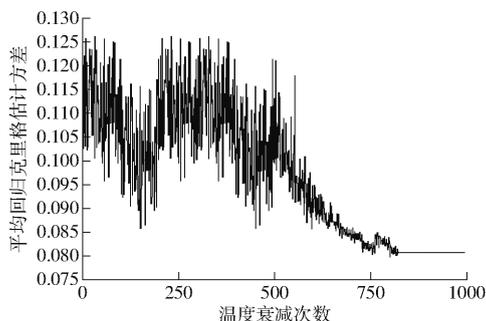


图5 优化过程中平均回归克里格估计方差变化曲线

Fig.5 Change of mean RKV during optimization

(2)评价优化方案对地理空间的代表性。图6为优化后采样点布局的G函数与F函数。由图6可知,优化后采样点布局的G函数与F函数均位于

随机分布95%置信区间内,即G函数与F函数均与其理论值无显著差异,这表明优化后的采样点布局在地理空间内既不是规则分布,也不是聚集型分布,而是随机分布并较好覆盖整个研究区。实际上,由图4a也可以看出,虽然个别点相距较近,但并未出现大量点聚集在同一区域的现象。此外,由F函数可以看出任意潜在采样点在600m(盐分变程)邻域内找到距其最近的采样点的概率为96.5%,这为空间插值提供邻域保证。因此,可以认为优化后的采样点布局较好地覆盖了地理空间,具有较好的地理空间代表性。

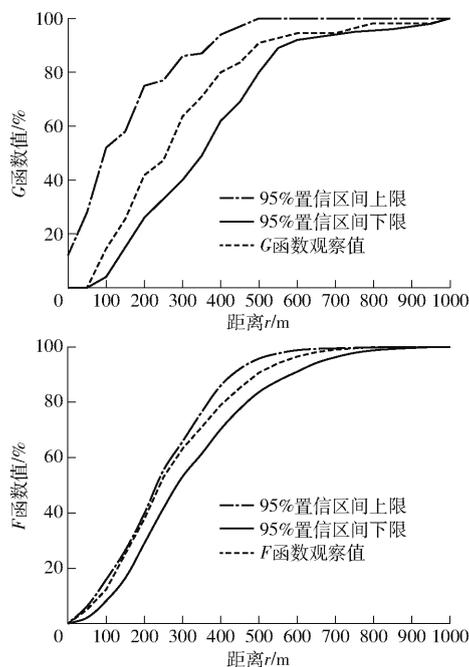


图6 G函数和F函数变化曲线

Fig.6 Changing curves of G-function and F-function

(3)评价优化方案对属性空间的代表性。如前所述,本文采用K-S法检验优化采样集对协变量所表达的属性空间代表性。由优化后采样点与荒地最短距离组成的样本和由50m×50m离散结点与荒地最短距离组成的总体属于同一分布。通过SPSS软件分析,得到相伴概率P为0.275,远大于0.05,零假设成立,这表明优化后的协变量样本来自于协变量总体,优化采样集在荒地分布所表达的属性空间上具有较好的代表性。

## 3 讨论

本研究仍是基于地统计学理论对土壤采样布局进行优化,这就要求研究变量的变异函数已知或具有估计的条件。当研究变量的变异函数确定后,则可以计算出研究区平均估计方差,因此本文的研究方法可以适用于大多数土壤属性的采样点布局优化,例如金属离子、土壤有机质等。对土壤盐分而

言,河套灌区的盐分研究已经积累了大量不同尺度、不同时期的盐分采样数据<sup>[20-24]</sup>,这可为盐分变异函数模型的建立和采样方案的优化提供有力支撑;研究还表明,河套灌区某一区域周年盐分空间分布格局相对稳定<sup>[23]</sup>,这也有利于盐分采样布局的优化,综上本文的研究方法可以很好地适用于河套灌区的盐分采样点布局优化。而当无过往数据或先验方差可利用时,本文的方法则难以使用。为获取精度较高的变异函数模型,可使用嵌套采样,此时,如何在综合考虑不同尺度环境协变量基础上,设计嵌套采样方案还需要进一步探讨。

本文研究方法适用于所有存在盐渍化问题的区域,但由于自然因素和人类活动的不同,不同地区影响土壤盐分分布的因素也不同。如张芳等<sup>[25]</sup>研究表明,新疆维吾尔自治区奇台绿洲地形因素对表层和中层土壤盐分分布格局影响较大,其中高程与盐分呈极显著相关关系;刘洋等<sup>[26]</sup>研究认为黄河三角洲地区土壤盐分空间变异程度与距海洋远近以及人类活动等因素有关;李亮等<sup>[27]</sup>研究表明,河套灌区内荒地,在作物生育期有排水积盐作用,荒地对土壤含盐量存在影响。本文选择荒地分布作为协变量,主要因为它是众多与土壤盐分关系密切的环境变量中更容易获取的。此外,本研究中大、小尺度上各环境变量和土壤盐分的相关分析以及其他相关研究<sup>[28]</sup>表明,同一个环境变量在不同研究尺度上与主

变量关系的密切程度可能不同,不同尺度上土壤盐分分布的主控因子不同。因此,在实际使用本文研究方法中,应根据地区特点和研究尺度选择合适的协变量。

综上所述,本研究以研究区内平均估计方差最小为目标,运用 SSA 对土壤盐分采样点布局进行优化,优化后的采样点布局在地理空间和属性空间均能较好地覆盖。研究结果对提高土壤盐分分布图绘制精度或降低野外采样和室内分析成本具有重要意义。

## 4 结论

(1) 在采样数为 55 时,以荒地分布作为协变量、运用 RK 法优化土壤盐分采样点布局,可使 50 m × 50 m 潜在采样点上的平均 RKV (0.080 7) 相较 OK 法优化结果的平均 OKV (0.114 3) 缩小 29.39%; *F* 函数图和 *G* 函数图表明,优化后的采样点布局在地理空间具有较好的代表性; *K-S* 检验则表明该采样点布局在荒地分布所代表的属性空间亦具有良好的代表性。

(2) 因不同研究尺度上,影响土壤盐分分布的主要环境因素不同,所以协变量具有尺度依赖性。本研究区在大尺度(采样网格 400 m × 400 m, 范围 1 574 hm<sup>2</sup>)上以荒地分布为协变量是合理的,但小尺度上或多尺度嵌套采样优化时,协变量需要另外择优选用。

## 参 考 文 献

- [1] FRANZEN D W, PECK T R. Field soil sampling density for variable rate fertilization[J]. *Journal of Production Agriculture*, 1995, 8(4): 568 - 574.
- [2] 徐剑波, 宋立生, 彭磊, 等. 土壤养分空间估测方法研究综述[J]. *生态环境学报*, 2011, 20(8-9): 1379 - 1386.  
XU Jianbo, SONG Lisheng, PENG Lei, et al. Research review on methods of spatial prediction of soil nutrients[J]. *Ecology and Environmental Sciences*, 2011, 20(8-9): 1379 - 1386. (in Chinese)
- [3] 陈天恩, 陈立平, 王彦集, 等. 基于地统计的土壤养分采样布局优化[J]. *农业工程学报*, 2009, 25(2): 49 - 55.  
CHEN Tianen, CHEN Liping, WANG Yanji, et al. Optimal arrangement of soil nutrient sampling based on geo-statistics[J]. *Transactions of the CASE*, 2009, 25(2): 49 - 55. (in Chinese)
- [4] 陈俊英, 姚志华, 张智韬, 等. 大田葵花土壤含盐量无人机遥感反演研究[J]. *农业机械学报*, 2020, 51(7): 178 - 191.  
CHEN Junying, YAO Zhihua, ZHANG Zhitao, et al. UAV remote sensing inversion of soil salinity in field of sunflower[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2020, 51(7): 178 - 191. (in Chinese)
- [5] SANAZ Z, SEYED R F S, SEYED A A. Weakly-coupled geo-statistical mapping of soil salinity to stepwise multiple linear regression of MODIS spectral image products[J]. *Journal of African Earth Sciences*, 2019, 152: 101 - 114.
- [6] AHMED E, LUIS G. Detecting soil salinity in alfalfa fields using spatial modeling and remote sensing[J]. *Soil Science Society of America Journal*, 2008, 72(1): 201 - 211.
- [7] LAME F P J, DEFIZE P R. Sampling of contaminated soil: sampling error in relation to sample size and segregation[J]. *Environ. Sci. Technol.*, 1993, 27(10): 2035 - 2044.
- [8] SIQUEIRA D S, MARQUES J, PEREIRA G T. The use of landforms to predict the variability of soil and orange attributes[J]. *Geoderma*, 2010, 155(1): 55 - 66.
- [9] SZATMARI G, BARTA K, PASZTOR L. An application of a spatial simulated annealing sampling optimization algorithm to support digital soil mapping[J]. *Hungarian Geographical Bulletin*, 2015, 64(1): 35 - 48.
- [10] 林清火, 郭澎涛, 罗微, 等. 基于空间模拟退火算法的橡胶园土壤取样布局优化[J]. *热带作物学报*, 2018, 39(10): 1920 - 1927.  
LIN Qinghuo, GUO Pengtao, LUO Wei, et al. Optimization of spatial soil sampling using continuous simulated annealing in

- rubber plantation[J]. *Chinese Journal of Tropical Crops*,2018,39(10):1920–1927. (in Chinese)
- [11] VASAT R, HEUVELINK G B M, BORUVKA L. Sampling design optimization for multivariate soil mapping[J]. *Geoderma*, 2010,155(3–4):147–153.
- [12] GLECIO M S, JORGE D D, ANTONIO P G, et al. Spatial soil sampling design using apparent soil electrical conductivity measurements[J]. *Bragantia*,2016,75(4):459–473.
- [13] BRUS D J, HEUVELINK G B M. Optimization of sample patterns for universal kriging of environmental variables[J]. *Geoderma*,2007,138(1–2):86–95.
- [14] FITZGERALD G J, LESCH S M, BAMES E M, et al. Directed sampling using remote sensing with a response surface sampling design for site-specific agriculture[J]. *Comput. Electron. Agr.*,2006,53(2):98–112.
- [15] SAMUEL-ROSAA A, HEUVELINK G B M, VASQUES G M, et al. Do more detailed environmental covariates deliver more accurate soil maps? [J]. *Geoderma*,2015,243:214–227.
- [16] LESCH S M. Sensor-directed response surface sampling designs for characterizing spatial variation in soil properties[J]. *Comput. Electron. Agr.*,2005,46(1–3):153–179.
- [17] SZATMARI G, LASZLO P, TAKACS K, et al. Optimization of second-phase sampling for multivariate soil mapping purposes: case study from a wine region, Hungary[J]. *Geoderma*,2019,352(3):373–384.
- [18] 张仁铎. 空间变异理论及应用[M]. 北京:科学出版社,2005.
- [19] 化骞寂. 永济灌域典型区土壤盐分时空变异分析与评价[D]. 扬州:扬州大学,2020.  
HUA Qianji. Analysis and evaluation of spatial and temporal variation of soil salinization in Yongji irrigation area [D]. Yangzhou:Yangzhou University,2020. (in Chinese)
- [20] 葛洲. 永济灌域典型区土壤盐碱化时空分布控制因素研究[D]. 扬州:扬州大学,2019.  
GE Zhou. Study on controlling factors of spatial-temporal distribution of soil salinization in typical area of Yongji irrigation area [D]. Yangzhou:Yangzhou University,2019. (in Chinese)
- [21] 徐旭,黄权中,黄冠华,等. 区域尺度农田水盐动态模拟模型——GSWAP[J]. *农业工程学报*,2011,27(7):58–63.  
XU Xu, HUANG Quanzhong, HUANG Guanhua, et al. Regional scale model for simulation soil water flow and solute transport processes—GSWAP[J]. *Transactions of the CASE*,2011,27(7):58–63. (in Chinese)
- [22] 寇微. 内蒙古河套灌区土壤水盐空间变异性研究[D]. 兰州:西北师范大学,2008.  
KOU Wei. Study on spatial variability of soil water and salinity in Hetao District in Inner Mongolia, China [D]. Lanzhou: Northwest Normal University,2008. (in Chinese)
- [23] 化骞寂,冯绍元,徐英,等. 河套灌区典型区周年内耕层土壤盐分时空变异研究[J]. *灌溉排水学报*,2020,39(8):26–34.  
HUA Qianji, FENG Shaoyuan, XU Ying, et al. Intra-annual spatiotemporal variation in salt content in the plough layer in Hetao Irrigation District[J]. *Journal of Irrigation and Drainage*,2020,39(8):26–34. (in Chinese)
- [24] 黄权中,徐旭,任东阳,等. 基于遥感反演河套灌区土壤盐分分布及对作物生长的影响[J]. *农业工程学报*,2018,34(1):102–109.  
HUANG Quanzhong, XU Xu, REN Dongyang, et al. Soil salinity distribution based on remote sensing and its effect on crop growth in Hetao Irrigation District[J]. *Transactions of the CASE*,2018,34(1):102–109. (in Chinese)
- [25] 张芳,熊黑钢,田源,等. 区域尺度地形因素对奇台绿洲土壤盐渍化空间分布的影响[J]. *环境科学研究*,2011,24(7):731–739.  
ZHANG Fang, XIONG Heigang, TIAN Yuan, et al. Impacts of regional topographic factors on spatial distribution of soil salinization in Qitai oasis[J]. *Research of Environmental Sciences*,2011,24(7):731–739. (in Chinese)
- [26] 刘洋,姜德锋,邹晓霞. 黄河三角洲地区棉田土壤盐分的时空分布特征研究[J]. *山东农业科学*,2017,49(5):68–73.  
LIU Yang, JIANG Defeng, ZOU Xiaoxia. Study on spatial and temporal distribution of soil salinity in cotton field in the Yellow River Delta[J]. *Shandong Agricultural Sciences*,2017,49(5):68–73. (in Chinese)
- [27] 李亮,史海滨,张义强,等. 内蒙古河套灌区耕地与荒地间水盐补排规律的研究[J]. *灌溉排水学报*,2010,29(5):73–77.  
LI Liang, SHI Haibin, ZHANG Yiqiang, et al. Transport of soil moisture and salt between irrigated land and saline land in the Inner Mongolia Hetao Irrigation District[J]. *Journal of Irrigation and Drainage*,2010,29(5):73–77. (in Chinese)
- [28] DONGYANG R, BOYU W, BEMARD E, et al. Analyzing spatiotemporal characteristics of soil salinity in arid irrigated agroecosystems using integrated approaches[J]. *Geoderma*,2019,356:113935.