

# 表征土壤有机碳区域分布的优化空间插值模型研究 ——以福建省龙海市为例

范胜龙, 黄炎和, 林金石

(福建农林大学 资源与环境学院, 福州 350002)

**摘要:**如何利用有限的样本数据来获得更为准确的土壤属性空间分布信息是土壤学研究的热点问题之一。利用福建省龙海市采集的 1 133 个耕地土壤样品, 设计了结合地貌类型、土壤类型和土地利用类型等信息的 5 种克里格插值模型, 研究县级尺度上土壤有机碳空间预测优化插值模型及其与样点密度的关系。结果表明: 设计的 5 种插值模型预测精度均高于普通克里格法, 但不同样点密度对插值结果影响较大。按  $0.5\text{ km} \times 0.5\text{ km}$  及以上的格网密度进行样点布设, 采用土地利用现状类型结合土壤类型信息的普通克里格法(KDLTR)插值结果误差较小; 按  $2\text{ km} \times 2\text{ km}$  的格网密度布设调查样点时, 采用土壤类型信息的普通克里格法(KTR)插值结果误差较小; 当格网大于  $4\text{ km} \times 4\text{ km}$  时, 由于样点数少, 各种模型的结果相差不大, 可直接采用普通克里格法(KYJZ)进行插值。

**关键词:**样点布设; 土壤有机碳; 空间插值模型; 克里格插值

中图分类号: S158

文献标识码: A

文章编号: 1005-3409(2011)06-0001-05

## The Optimized Interpolation Models and Its Relationship with Soil Sampling Density on Detecting Spatial Variability of Farmland Soil Organic Carbon

—A Case Study in Longhai City, Fujian Province

FAN Sheng-long, HUANG Yan-he, LIN Jin-shi

(College of Resource and Environment, Fujian Agriculture and Forestry University, Fuzhou 350002, China)

**Abstract:** How to use limited soil samplings to interpolation an accurate spatial distribution of soil properties is a hot issue in soil science. In order to study the optimized interpolation models and its relationship with soil sampling density, the research collected 1 133 soil samplings from Longhai County, Fujian Province, and designed 5 interpolate models combined with landform types, soil types, and land use types, etc. Results showed that the accuracy of 5 interpolate models are higher than Ordinary Kriging (OK) method. However, the size of soil sampling grid can affect interpolate results. Using interpolate model integrated with land use types and soil types can get more accuracy results when soil sampling collected from  $0.5\text{ km} \times 0.5\text{ km}$  grids. When taking sampling form  $2\text{ km} \times 2\text{ km}$  grids, the model integrated with soil types can get more accuracy results. But there are no variances between different models when sampling grid reaches  $4\text{ km} \times 4\text{ km}$ , so it can use OK method to interpolate directly.

**Key words:** sampling method; soil organic matter; interpolation models; Kriging method

土壤有机碳是土壤理化性质中最重要、最具代表性的因子, 其在空间分布上具有不均一性, 是发生变化的连续体<sup>[1-2]</sup>。对土壤有机碳储量、分布、转化、衰减机理进行研究, 并揭示其影响因素和生态效应是目前土壤有机碳研究的热点问题。而土壤样点的代表性是研究结果是否准确的前提, 大规模的土壤属性空

间信息的采集需要花费大量的人力、物力和财力。因此, 研究如何利用有限的样本数据来获得更为详尽的土壤属性空间分布信息的方法具有重要意义。

空间插值模型是实现土壤有机碳含量从离散的点状信息向面状连续信息转换的有力工具, 是表征土壤有机碳空间分布特征的重要手段。目前, 在土壤属

性空间预测的研究方面,地统计学被证明为最有效的空间插值方法<sup>[3-6]</sup>。综观学者们的研究,克里格插值是使用最多的空间预测方法,但在地形复杂地区,普通克里格方法的应用有一定的局限性。辅助信息用以提高目标变量的预测精度,在土壤学的研究中已经达成共识。协同克里格法就是一种利用辅助的方法,并在土壤科学中得到了广泛的应用<sup>[7]</sup>。但相对于县级或更大研究区域而言,为了达到优化插值目的,而多调查另一数据,无疑增加了调查成本,且通过调查所得两组数据一般样点数目相同,因此,采用协同克里格法增加了样点调查成本。本研究以福建省龙海市为典型区,以耕地中耕层土壤有机碳为研究对象,设计结合地貌类型、土壤类型和土地利用类型等信息的克里格插值模型,研究县级尺度上土壤有机碳空间预测的优化空间插值模型及其与样点密度的关系。

## 1 材料与方法

### 1.1 研究区概况

龙海市位于福建省漳州市东部沿海,全市总面积 1 289.72 km<sup>2</sup>,属南方丘陵区。境内地势北部、西部、南部三面环山、中部为平原、东南部临海。土壤形成受成土母质的影响极大,由于母质分布的区域性不同,导致不同区域土壤变化过程具有明显的差异性,加上长期以来自然因素和人为因素的作用,形成土壤类型的多样性。根据第二次土壤普查资料,龙海市土壤分为 6 个土类,16 个亚类,52 个土属,70 个土种。

### 1.2 研究方法

1.2.1 土壤样品采集及测定方法 考虑到格网法是有关土壤属性调查样点布设的最常用方法且能够较为准确地表征土壤属性的空间变异<sup>[8-10]</sup>,本研究采用格网法布设样点。根据前人相关研究成果和研究区域的大小,本研究采用 0.5 km×0.5 km( $X_1$ )、1 km×1 km( $X_2$ )、2 km×2 km( $X_3$ )、4 km×4 km( $X_4$ ) 四种格网布设样点。

土壤样品采集的具体做法是在样点附近 20 m 的范围内取 5 个耕作层(0—20 cm)土样,然后混合成一个土壤样品,用四分法取 1.0 kg 带回实验室,采用常规的低温外加热重铬酸钾氧化—滴定法测定土壤有机碳含量。在采样的同时,用 GPS 记录每个样点的经纬度信息,并描述各样点的土壤、土地利用类型、地貌及相关环境信息。所有土壤样品采集在 2009 年 11—12 月农作物收割完成后进行。

1.2.2 土壤有机碳空间插值模型 由于地貌类型、土壤类型和土地利用类型对土壤有机碳空间分布均有重要影响,可以将其作为提高土壤属性预测精度的

辅助信息<sup>[11]</sup>。为研究上述辅助信息对优化空间插值结果的影响,本研究在采用格网法取样的基础上,分别设计了结合地貌类型信息的克里格方法(KDM)、结合土地利用类型信息的克里格方法(KDL)、结合土壤类型信息的克里格方法(KTR)、结合地貌—土壤类型信息的克里格方法(KDMTR)、结合土地利用类型—土壤类型信息的克里格方法(KDLTR)五种结合类型信息的方法,并与直接采用各样点土壤有机碳含量数据进行普通克里格插值(KYJZ)的结果进行比较分析。上述结合类型信息的方法将每一个样点的土壤有机碳含量值  $z(x_{kj})$  分为相同类型均值  $\mu(t_k)$  和残差  $r(x_{kj})$  之和。用公式表示为:

$$z(x_{kj}) = \mu(t_k) + r(x_{kj}) \quad (1)$$

式中: $z(x_{kj})$ ——样品的土壤有机碳含量; $\mu(t_k)$ ——相同类型样品的均值; $r(x_{kj})$ ——样品土壤有机碳含量与其相同类型样品的均值之差,称为“残差”。

研究将残差作为一个新的区域变量  $r(x_{kj})$  进行普通克里格插值,空间插值利用 ArcGIS 软件中的地统计分析模块完成。各待估点的土壤有机碳含量预测值  $z^*(x_{kj})$  为类型均值  $\mu(t_k)$  与残差估计值  $r^*(x_{kj})$  之和。

$$z^*(x_{kj}) = \mu(t_k) + r^*(x_{kj}) \quad (2)$$

本研究中,TR 法、DMTR 法和 DLTR 法中的土壤类型均划分到土属级别。土壤类型资料来自于第二次土壤普查成果中的 1:5 万土壤图和《龙海土壤》;土地利用现状资料来源于第二次全国土地调查标准时点统一更新数据库;地貌划分来源于福建省农用地分等成果更新项目成果中的指标区图。

1.2.3 预测结果检验 为了检验各种空间插值方法的效果,本研究在全市范围内随机选择 259 个验证样点,验证样点包含了预测样点中所有地类(土壤类型和地貌类型),且每种类型的验证样点数量不少于 3 个。同时,在选择验证样点时也兼顾空间分布的均匀性。通过验证点的实测值与该验证点通过各种插值方法所得的预测值的相关系数( $r$ )及其均方根误差(RMSE)来评价其预测精度的高低。 $r$  值越大、RMSE 越小则预测精度越高,反之精度越低。

1.2.4 样点密度对土壤有机碳空间插值模型精度的影响研究 在选定空间插值方法的基础上,为探讨样点密度与土壤有机碳空间预测精度之间的量化关系,本研究利用布设的 4 种不同密度格网所得样点,采用设计的 6 种结合类型信息的克里格法分别进行空间预测,并通过各验证点实测值与预测值的 RMSE,对其预测精度进行验证,研究满足精度要求的空间插值模型及其所需样点密度。

## 2 结果与分析

### 2.1 龙海市土壤有机碳的空间分布特征

本研究采用4种不同密度格网叠套于龙海市1:1万土地利用现状图中进行样点布设,全市共获得1133个耕地图斑上的预测样点,在所有样点中,土壤有机碳含量最小值为3.50 g/kg,最大值为86.70 g/kg,两者相差了近23倍,土壤有机碳含量均值为

24.45 g/kg,变异系数43.35%,说明南方丘陵区县级尺度上农田土壤有机碳含量变异较大。

### 2.2 结合不同类型信息的土壤有机碳含量空间插值

利用ArcGIS软件对4种格网密度下耕层有机碳含量数据去除各种类型均值后残差采用普通克里格法进行空间插值,结果如图1所示。受篇幅的限制,本研究仅以0.5 km×0.5 km格网为例,列出结果图并进行分析。

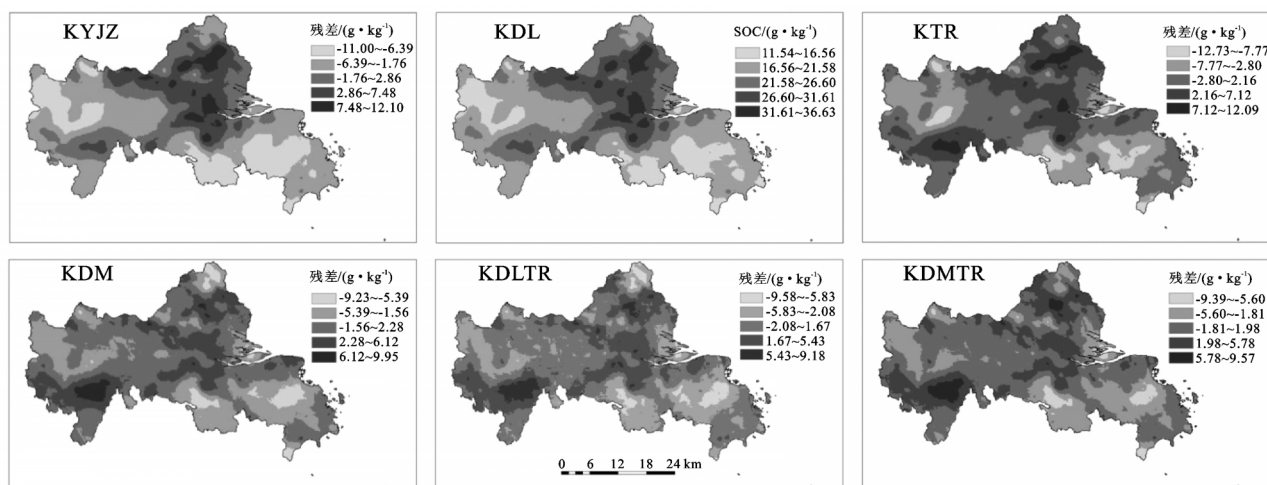


图1 不同方法所得插值结果

从图1中可以看出,各种方法所得插值图反映的土壤有机碳的空间分布趋势基本一致,即北部和中部九龙江下游冲积平原的土壤有机碳含量较高,而东部港尾和隆教两乡镇沿海的太武山区土壤有机碳含量最低。各种方法所得插值图反映的土壤有机碳的这种分布趋势与龙海市土壤有机碳实际的空间分布情况基本一致,表明所用各种空间插值方法能基本准确反映出研究区的土壤有机碳分布。

尽管各种方法所得空间插值图有一定的相似之处,但各种方法所得成果图的细部具有较大差别。KYJZ、KDL所得的插值图最为相似,图斑较大,平滑效应明显,表明这两种方法所得结果基本一致。而KDLTR和KDMTR法所得插值图明显较上述两种方法所得图斑来得破碎,表明对于研究区而言,土壤类型信息较土地利用类型信息和地貌类型信息对反映土壤有机碳含量的分布影响较大。KDLTR法所得图斑最为破碎,分辨率最高,表明其更好地表达了研究区土壤有机碳含量的空间变异特点,采用该方法具有较好的空间插值效果。

### 2.3 结合不同类型信息的空间插值模型预测精度分析

ArcGIS软件所进行的交叉验证是利用参加空间插值的数据组进行的自我验证,仅能反映基于该数据组的最佳插值模型及其参数,从理论上说具有一定的“欺骗性”。在ArcGIS软件所进行的交叉验证的基

础上,为检验不同样点密度下,各种不同空间插值方法的预测精度,本研究对验证样点的预测值与相同位置的实测值进行相关性分析,并作散点图(图2)进行检验,拟合出两者间的回归方程。同时,采用均方根误差(RMSE)(图3)来衡量各种方法预测结果的准确性和预测方法的系统误差。

从图2中可以看出,各方法的实测值与预测值均达到了极显著相关水平( $p < 0.01$ ),且其拟合线的斜率不同。同时各种格网密度下采用的各种空间插值方法间的均方根误差最大值与最小值相差均达到0.5 g/kg以上,表明在同一格网密度内采用的各种预测方法所得的预测结果的精度具有较明显的差距。

在0.5 km×0.5 km的样点密度中,KDLTR法的预测值和实测值的相关系数最大( $r = 0.786^{**}$ )其拟合线的斜率也较接近于参考线,KDL法中的相关系数最小( $r = 0.744^{**}$ ),其拟合线的斜率与参考线的斜率相差也较大,而其它方法的相关系数介于两者之间,由大到小分别为KTR、KDMTR、KYJZ和KDM。同时,KDLTR法所得预测结果的均方根误差最小,为3.6419 g/kg,KDL法的均方根误差最大,达4.4226 g/kg,两者相差0.7807 g/kg;其余方法中均方根误差由小到大依次为KTR、KDMTR、KDM和KYJZ。KDLTR法相对于KYJZ法,其均方根误差下降了17.38%,即其精度提高了17.38%。从各种方法

的均方根误差还可看出(图 3),KDLTR、KDMTR 和 KTR 这三种结合了土壤类型信息的方法均较其它方

法的预测精度有较为明显的提高,表明在县级尺度下,土壤类型信息对于土壤有机碳含量的影响较大。

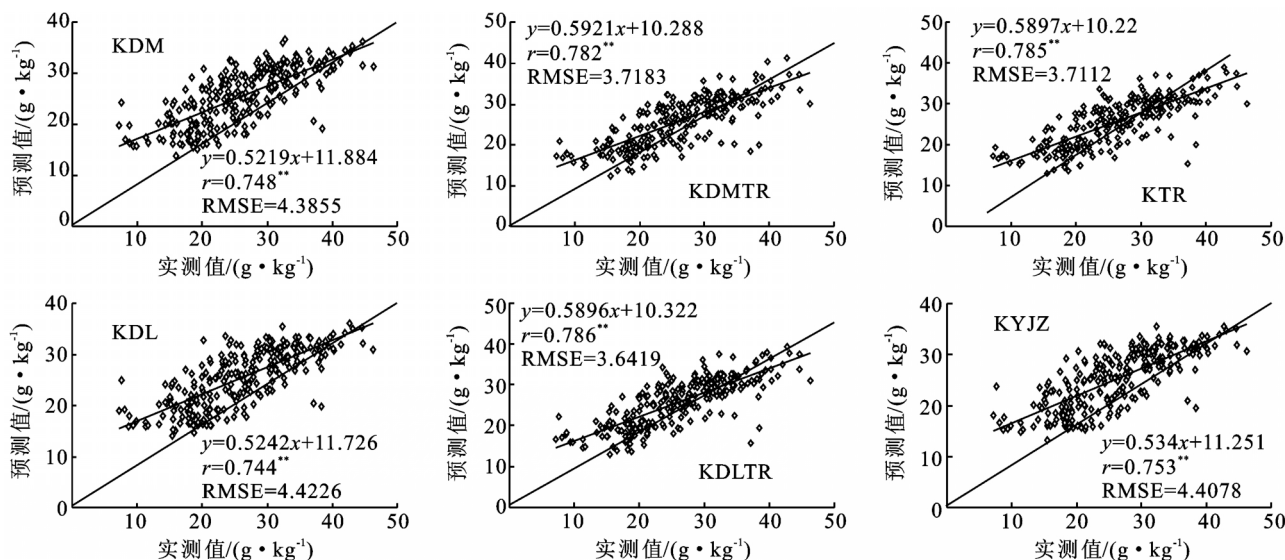


图 2 不同方法的实测值与预测值线性回归分析

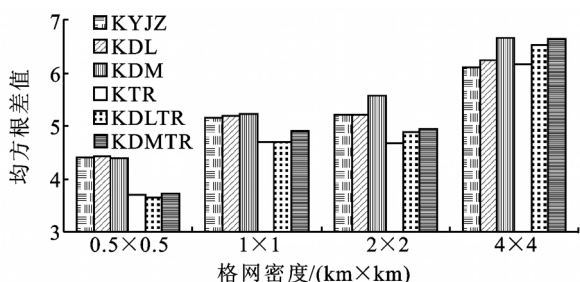


图 3 不同格网密度内各种空间插值方法的均方根误差

在  $1 \text{ km} \times 1 \text{ km}$  的样点密度中 KDMTR 法的预测值和实测值的相关系数最大( $r = 0.792^{**}$ ), KDM 法两者的相关系数最小( $r = 0.689^{**}$ )。其它方法的相关系数介于两者之间,由大到小分别为 KTR、KDLTR、KDL 和 KYJZ。同时, KDLTR 法所得预测结果的均方根误差最小,为  $4.6505 \text{ g/kg}$ , KDM 法的均方根误差最大,为  $5.2320 \text{ g/kg}$ ,两者相差  $0.5815 \text{ g/kg}$ ;其余方法中均方根误差由小到大依次为 KTR、KDMTR、KYJZ 和 KDL。KDLTR 法相对于 KYJZ 法,其均方根误差下降了  $9.64\%$ 。

在  $2 \text{ km} \times 2 \text{ km}$  的样点密度中 KTR 法的预测值和实测值的相关系数最大( $r = 0.752^{**}$ ), KDM 法两者的相关系数最小( $r = 0.622^{**}$ )。其它方法的相关系数介于两者之间,由大到小分别为 KDMTR、KDLTR、KYJZ 和 KDL。同时, KTR 法所得预测结果的均方根误差最小,为  $4.6805 \text{ g/kg}$ , KDM 法的均方根误差最大,为  $5.5784 \text{ g/kg}$ ,两者相差  $0.8979 \text{ g/kg}$ ;其余方法中均方根误差由小到大依次为 KDLTR、KDMTR、KYJZ 和 KDL。KTR 法相对于 KYJZ 法,其均方根误差下降了  $10.04\%$ 。

在  $4 \text{ km} \times 4 \text{ km}$  的格网密度中,相关系数最大的为 KTR 法, ( $r = 0.617^{**}$ ), KDM 法的相关系数最小,为  $0.461^{**}$ 。同时, KYJZ 法所预测结果的均方根误差最小,为  $6.1010 \text{ g/kg}$ ,各种方法所得均方根误差均大于  $6.0 \text{ g/kg}$ ,平均达  $6.39 \text{ g/kg}$ ,比  $2 \text{ km} \times 2 \text{ km}$  样点密度的平均均方根误差增加了  $25.72\%$ ,其主要原因仍应归于样点数量的减少,不能充分表征土壤有机碳的空间变异。研究结果表明  $4 \text{ km} \times 4 \text{ km}$  的样点密度已不适用于南方丘陵区县级尺度的土壤有机碳含量的空间预测。

## 2.4 样点密度变化对各种预测方法的精度影响

将各种预测方法在不同样点密度下所得预测结果与实测值的相关系数进行比较分析可知,除 KTR 和 KDMTR 法外,其余方法的相关系数均随格网密度减小而减小。

比较各种预测方法在不同样点密度下所得预测结果与实测值的均方根误差(图 4),可知各种预测方法的均方根误差随着格网密度的减少而增大。 $0.5 \text{ km} \times 0.5 \text{ km}$  的样点密度下,其各种方法的空间插值精度均较其它密度等级有明显提高;而  $1 \text{ km} \times 1 \text{ km}$  和  $2 \text{ km} \times 2 \text{ km}$  两种样点密度的空间插值精度差距很小,表明  $1 \text{ km} \times 1 \text{ km}$  的样点密度对其耕层有机碳含量空间预测精度的影响不明显; $4 \text{ km} \times 4 \text{ km}$  的样点密度下,各种空间插值方法所得结果的误差均有较大幅度的提高。就各种方法而言, KTR 法在各个样点密度等级下均有较良好的表现,是一种较好的空间插值方法,也表明在县级尺度上,土壤类型是影响土壤有机碳的重要因素,而仅结合土地利用现状或地貌

类型信息的克里格插值方法,对于土壤有机碳含量的空间预测效果则不理想。KDLTR法在样点密度大于 $1\text{ km}\times 1\text{ km}$ 时,其预测结果的精度最高;当样点密度小于等于 $1\text{ km}\times 1\text{ km}$ 时,则KTR法的预测精度最高。

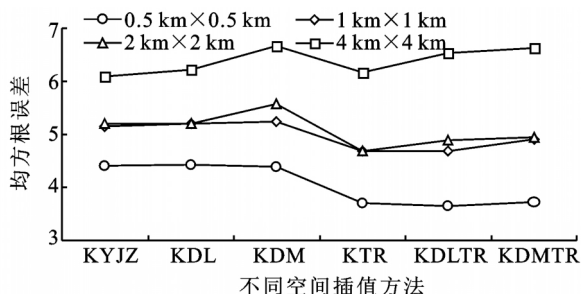


图4 不同空间插值方法的精度对格网密度的响应

以上分析表明,当样点密度小于等于 $4\text{ km}\times 4\text{ km}$ 的格网密度时,由于样点数过少,不适用于南方丘陵区县级尺度的土壤有机碳空间预测。如受条件制约,仅能布设到该密度等级时,采用结合类型信息的克里格插值法的预测精度不如采用原始有机碳含量数据直接进行普通克里格插值法所得结果的预测精度。因此,在县级尺度上,当研究所需较高的空间预测精度时,样点的布设须达到 $0.5\text{ km}\times 0.5\text{ km}$ 及以上的样点密度,此时应采用结合土地利用现状和土壤类型信息的普通克里格法进行空间预测;当研究仅需达到中等精度的空间预测结果时,则仅需以 $2\text{ km}\times 2\text{ km}$ 的格网密度布设样点,此时宜采用结合土壤类型信息的普通克里格法进行空间预测。

### 3 结论

(1)龙海市土壤有机碳含量在县级研究区域上存在较强的空间变异性。

(2)结合不同类型信息空间插值模型对于土壤有机碳区域分布的预测精度均有不同程度的提高,同时样点密度直接影响了空间插值方法的选择。根据研究目的的不同,当研究只要求获得中等精度的空间预测结果时,应按 $2\text{ km}\times 2\text{ km}$ 的格网密度布设调查样点,并采用结合划分到土属级别的土壤类型信息进行普通克里格法(KTR)插值;当研究要求获得高精度的空间预测结果时,则需按 $0.5\text{ km}\times 0.5\text{ km}$ 及以上的格网密度进行样点布设,并采用结合土地利用现状类型和土壤类型信息的普通克里格法(KDLTR)进行

空间插值;而格网密度小于 $4\text{ km}\times 4\text{ km}$ 时,由于样点数太少,不能得出准确反映土壤有机碳含量的空间分布,若受条件限制,仅能得到该格网密度时,应直接采用普通克里格法(KYJZ)进行空间预测。

(3)在南方丘陵区县级尺度上,土壤类型较土地利用类型和地貌类型对土壤有机碳含量的空间分布影响大。凡是涉及结合土壤类型信息的克里格方法(KTR、KDLTR、KDMTR)所得空间插值结果均明显好于未结合土壤类型信息的方法(KYJZ、KDL、KDM)。因此在县级尺度上研究土壤有机碳的空间分布时最好考虑土壤类型的影响。

#### 参考文献:

- [1] 华孟,王坚. 土壤物理学[M]. 北京:北京农业大学出版社,1992:214-243.
- [2] 杨玉玲,文启凯,田长彦. 土壤空间变异研究现状及展望[J]. 干旱地区农业研究,2001,18(2):50-55.
- [3] 胡克林,李保国,林启美,等. 农田土壤养分的空间变异性特征[J]. 农业工程学报,1999,15(3):33-38.
- [4] 王军,邱扬. 土地质量的空间变异与尺度效应研究进展[J]. 地理科学进展,2005,24(4):28-35.
- [5] 张淑娟,方慧,何勇. 精细农业田间信息采集策略[J]. 农业机械学报,2004,35(4):88-92.
- [6] Thomas J S, David W M. The Spatial variation of plant-available phosphorus in pastures with contrasting management[J]. Soil Science Society of America Journal, 2003,67:826-836.
- [7] Wu J, Norvell W A, Hopkins D G, et al. Improved prediction and mapping of soil copper by kriging with auxiliary data for cation-exchange capacity [J]. Soil Science Society of America journal,2003,67(3):919-927.
- [8] 姜城,杨俐苹. 土壤养分变异与合理取样数量[J]. 植物营养与肥料学报,2001,7(3):262-270.
- [9] 王宏庭,金继运,王斌,等. 土壤速效养分空间变异研究[J]. 植物营养与肥料学报,2004,10(4):249-354.
- [10] Wang X J, Qi R. The effect of sampling design on spatial structure analysis of contaminated soil[J]. The science of the total environment,1998,224:29-41.
- [11] Liu T L, Juang K W, Lee D Y. Interpolating soil properties using kriging combined with categorical information of soil maps[J]. Soil Science Society of America Journal,2006,70(4):1200-1209.