



# 农田土壤采样与空间变异性研究

张 玉 铭

栾城农业生态系统试验站

# 报告提纲

---

- 一、研究背景及意义
- 二、地统计学原理简介
- 三、农田土壤空间变异性研究案例
  - 1. 农田土壤样品采集
  - 2. 经典统计学分析
  - 3. 土壤养分的空间变异结构
  - 4. 讨论与小结

# 一、研究背景及意义

---

## 1. 农田土壤合理采样的重要性

土壤是复杂的历史自然实体，其属性在空间上的不均匀分布（空间变异性）有其独立性和不确定性（随机性），即必然性和偶然性。

农田土壤性质（如养分、水分）空间变异性是进行农业科学试验和实施精确耕作管理的主要障碍之一。土壤养分含量的非均一性分布会导致农作物产量空间变化，因而可靠的土壤采样方式是精确的土壤特征化诊断辨识的先决条件。进行土壤空间变异性研究对推动土壤科学定量研究与精准农业的实施具有重要意义。

# 一、研究背景及意义

---

## 2. 基于地统计学方法土壤特性空间变异性研究进展（1）

自60年代起尤其80年代以后，土壤属性空间变异性的研究日益受到重视。逐渐成熟的地统计学，能定量认识土壤空间变异性，使土壤变异性评价成为制定土壤采样方案和预示未采样点的一个新途径或基本步骤。基于地统计学方法研究土壤特性空间变异性，可为确定田间合理取样尺度和取样数目，为土壤过程的预测、模拟更接近农田土壤变化的实际情况提供有效途径，使人们能够更好地理解空间作用对土壤作物关系的重要性，而这些信息的定量化正是精准农业开展、实施不可或缺的基础资料和理论依据。

# 一、研究背景及意义

---

## 2. 基于地统计学方法土壤特性空间变异性研究进展（2）

不同水平和空间尺度的土壤空间变异性研究国内外已有大量报道。自1963年Andrew等研究土壤性质变异性开始，国外学者先后应用地统计学方法研究土壤水性质的空间变异特征、作物产量空间变异性及其与土壤特性空间变异特征、环境因素的结构特征的相关关系、土壤物理化学性质和景观制图单元的分类学构成上的变异性等等。

国内土壤性质变异性的研究八十年代起已有开展，雷志栋(1985, 1988)、陈志雄(1988)、吕军(1990)、李鸿杰(1993)、周慧珍(1996)、李保国(1998)等大部分研究针对土壤物理学方面，引导并推动了学科发展，尤其集中在土壤水分运动变异性上。进入21世纪，区域性土壤空间变异性相关研究报道大量涌现。

## 二、地统计学原理简介

---

### 1. 基本概念

➤ 区域化变量 (Regionalized Variable)

是在空间上与其位置有关的变量。区域化变量的变化遵循一定的规律并与其所处的空间位置相对应。

➤ 区域化变量同时具有空间分布结构性与随机性，区域化变量理论是地统计学的基础

## 二、地统计学原理简介

---

### 2. 具体内容

地统计学分析包括两个方面：以半方差函数为基本工具，描述区域化变量结构性与随机性；以克里格法为基本方法，给出有限区域内区域化变量的最佳无偏估计量。

## 二、地统计学原理简介

---

### 3. 分析方法

- 自相关 (autocorrelation) 分析

### 3. 分析方法

---

#### ➤ 自相关(autocorrelation)分析

土壤属性具有空间自相关性，自相关系数  $\rho(h)$  是描述随机函数  $Z(x)$  在空间位置上自相关程度的统计量，定义为：

$$\rho(h) = E\{[Z(x) - \mu][Z(x+h) - \mu]\} / \sigma^2,$$

其计算式为：

$$\rho(h) = \frac{\sum [Z(x_i) - \bar{Z}(x)][Z(x_i + h) - \bar{Z}(x)]}{\sqrt{\sum [Z(x_i) - \bar{Z}(x)]^2 \cdot \sum [Z(x_i + h) - \bar{Z}(x)]^2}}$$

$Z(x)$ 、 $Z(x+h)$  为位置  $x_i$ 、 $x_{i+h}$  处某性质测定值， $h$  为样本间距， $\mu$  为总体均值， $\sigma^2$  为总体方差。

$\rho(h)$  的范围在  $[-1, 1]$  之间。 $\rho(h)$  数值大小反映随机变量  $Z(x)$  相关程度强弱， $\rho(h)$  愈大， $Z(x)$  自相关性愈好，反之愈差

### 3. 分析方法

---

#### ➤ 半方差(semivariogram)分析

半方差函数(或称半变异函数)分析是另一种解释土壤属性空间结构(关系)的方法,半方差函数计算和模拟被认为是地统计学研究的核心。

半方差函数定义为区域化变量 $Z(x)$  的增量  $[Z(x+h)-Z(x)]$  方差的一半,即:

$$\gamma(h) = (1/2)\text{Var}[Z(x+h) - Z(x)]$$

其计算式:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2$$

式中 $h$ 为样本间距, 又称位差(lag);  $N(h)$ 是间距为 $h$ 的样本“成对(pair)” 数目。

## ➤ 半方差函数理论模型

半方差图由一系列离散点构成，可根据其形状用直线或曲线方程拟合，这种方程即为半方差函数理论模型。在土壤科学中常用到多种理论模型，如球状 (spherical)、指数 (exponential)、双曲线 (hyperbola)、高斯 (Gaussian)、线性 (linear)、乘幂 (power) 等模型。

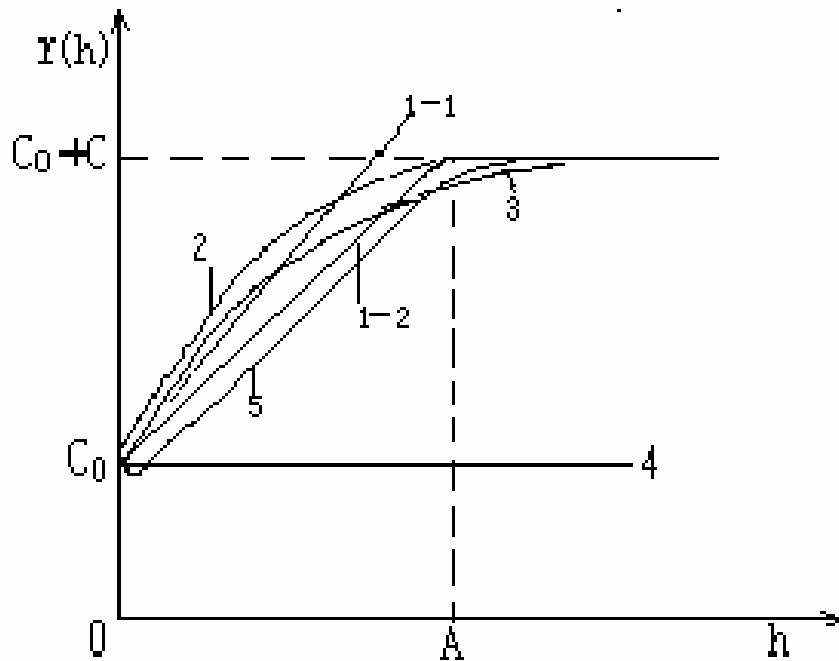


图 半方差函数理论模型示意图

**h:** 样本间距

**$C_0$ :** 块金系数: 间距  $h=0$  时的半方差

**C:** 结构方差

**$C_0 + C$ :** 基台值(又称平顶值或顶坎), 半方差随着间距递增到一定程度后, 出现的平稳值

**A:** 变程: 又称极限距离, 使半方差达到基台值时的样本间距

图中:

1-1 无基台线性模型

1-2 有基台线性模型

2. 球状模型

3. 指数模型

4. 纯块金效应模型

5. 高斯模型

## ➤ 半方差函数及半方差图参数的含义

**C<sub>0</sub>: 块金系数:** 块金效应是非空间性质观测值的变异, 因此, 它是在细微尺度模拟测量误差、操作偏差和空间变异的方差。

**A: 变程:** 它的意义因所遵循的模型而有不同。如对于球状和线性模型, A表明土壤性质存在空间变异结构的最大相关距离; 对于高斯模型和指数模型最大相关距离则分别为1.73A和3A。

**C<sub>0</sub>+C: 基台值:** 当 $h$ 趋于合适的间距 (**A**) 时半方差的极限值, 等于变量的方差, 超过此变程可以认为属性变量空间独立。

一般, 球状、指数或高斯模型采用块金值、基台和变程这3个基本属性参数来表征。

半方差图除用于分析土壤特性空间分布的方向性和相关距离外, 还可用于对未测点的参数进行最优内插估值和绘制等值线图。半方差函数的精确估计对进行属性值空间内插和生成栅格图很必要, 其可靠性取决于已知位差的测点数据“对”的数目, 反过来说, 它取决于可能承受的调查采样密度。

## ➤ 半方差与自相关系数的关系

半方差与自相关系数在理论上有如下关系：

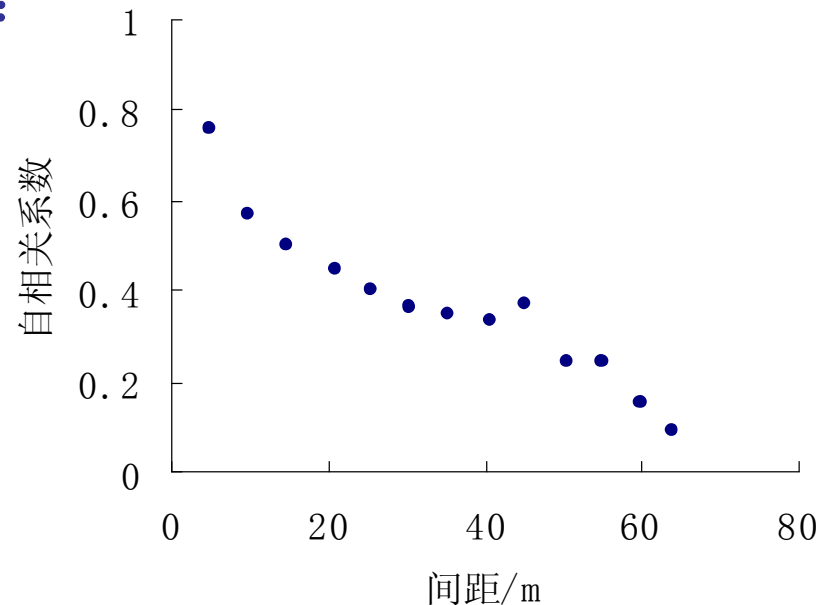
$$\gamma(h) = \sigma^2 [1 - \rho(h)]$$

随 $h$ 增大, 自相关系数减小, 半方差增大。

当 $h = a$  时,  $\gamma(h) = \sigma^2$ , 即半方差值等于随机变量的方差。

例如, 表土层**CI**的自相关图：

随着样点间距的增多, 自相关性减弱直至消失。



### 三、农田土壤空间变异性分析案例

---

采用地统计学理论的半方差函数，分析不同尺度下土壤养分的空间变异结构，以期制定合理土壤采样方案、降低采样误差、预示未采样点以及进行农田土壤肥力精确管理提供依据。

半方差函数计算和半方差图拟合应用Variowin软件。

# 1. 农田土壤样品采集

网格式采样。在栾城站示范区内以**20 m**间隔布点，测绳结合**GPS**定位，点位**230**个，用**5cm**直径土钻采集土壤样品，采样深度**0~15 cm**，样点布局见图1。

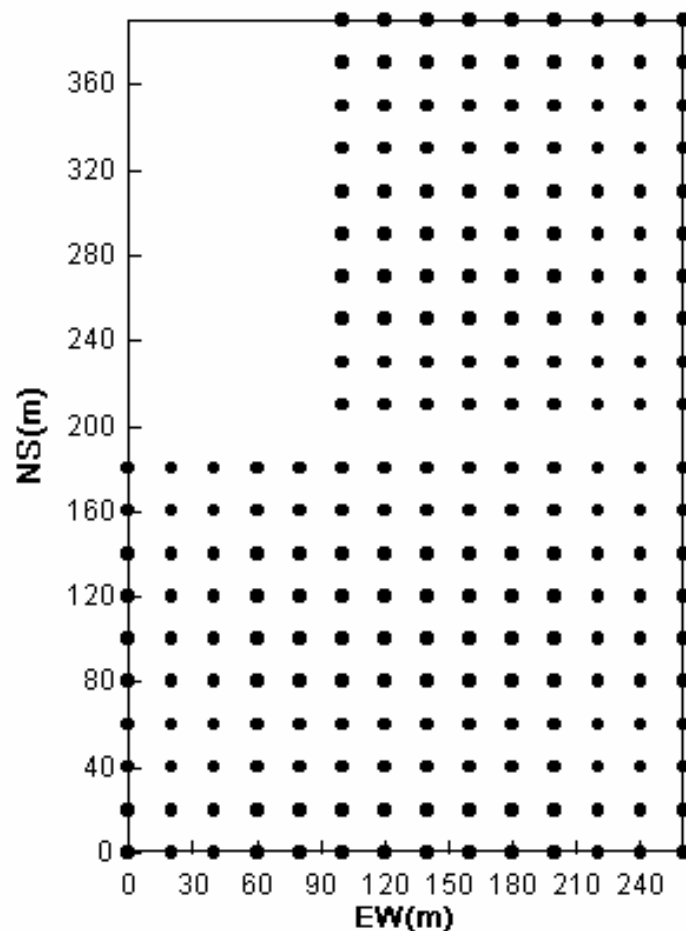
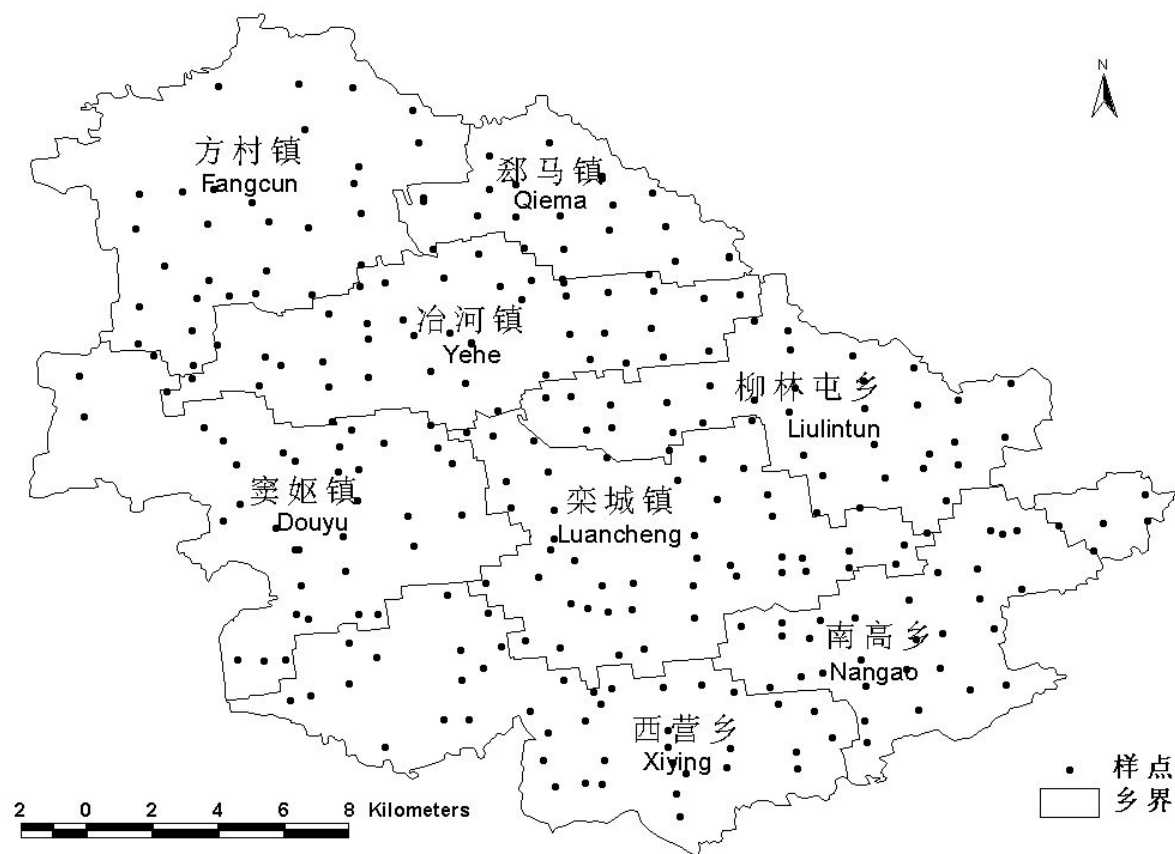


图3-1 栾城试验示范区样点布局示意图

# 1. 农田土壤样品采集



县域范围内基本上以1 km为间隔，GPS定位布点，采样深度0~20 cm，共采集土样289个，样点布局如图2所示。

图3-2 栾城县土壤采样点分布示意图

## 2. 经典统计学分析

经典统计所获取的均值、标准差、方差、变异系数等仅在某种程度上可以估测和描述总体，其中变异系数描述属性的作用最大。但是该方法的应用前提在于样本完全独立，服从正态分布，并且有足够多的样品数。实际上样品之间的相互独立性很难得到保证；如果设置过多的非独立性样点势必造成检测费用增加，也导致样本总方差的偏低。因此经典统计方法除了可以判别样本分布类型外，统计均值、标准差、变异系数等仅在一定程度上反映所研究区域的总体养分水平及变异状况。

表3-1 栾城县耕层土壤养分统计参数及分布类型检验

| 养分项目                                   | 均值    | 标准差   | 变异系数 | 峰度   | 偏度  | 极值          | 分布类型 |
|----------------------------------------|-------|-------|------|------|-----|-------------|------|
| 有机质(%)                                 | 1.734 | 0.238 | 13.7 | 4.7  | 1.2 | 1.177~3.068 | 正态   |
| 全氮(%)                                  | 0.111 | 0.014 | 12.4 | 0.3  | 0.4 | 0.079~0.157 | 正态   |
| 全磷(%)                                  | 0.084 | 0.012 | 14.2 | 2.9  | 1.1 | 0.057~0.142 | 正态   |
| 碱解氮 ( $\text{mg}\cdot\text{kg}^{-1}$ ) | 79.8  | 9.9   | 12.4 | 1.4  | 0.6 | 54.3~120.0  | 正态   |
| 速效磷 ( $\text{mg}\cdot\text{kg}^{-1}$ ) | 20.9  | 13.2  | 63.0 | 6.7  | 2.1 | 5.6~104.2   | 对数正态 |
| 速效钾 ( $\text{mg}\cdot\text{kg}^{-1}$ ) | 113.2 | 32.4  | 28.6 | 6.0  | 2.1 | 68.0~296.0  | 对数正态 |
| 铵态氮 ( $\text{mg}\cdot\text{kg}^{-1}$ ) | 2.7   | 2.0   | 74.0 | 10.6 | 2.2 | 0~15.5      | 对数正态 |
| 硝态氮 ( $\text{mg}\cdot\text{kg}^{-1}$ ) | 27.8  | 24.6  | 88.5 | 8.0  | 2.5 | 0.1~178.9   | 对数正态 |

## 2. 经典统计学分析

按照公认的变异程度分级标准来区分农田土壤样本总体变异的强弱：弱变异 $CV < 10\%$ ，中等变异 $CV = 10.0\% \sim 100\%$ ，强变异 $CV > 100\%$ 。

表3-2 示范区耕层土壤养分统计参数及分布类型检验

| 养分项目                                     | 均值    | 标准差   | 变异系数  | 峰度    | 偏度   | 极值          | 分布类型 |
|------------------------------------------|-------|-------|-------|-------|------|-------------|------|
| 有机质(%)                                   | 1.654 | 0.171 | 10.4  | 1.4   | 0.7  | 1.121~2.280 | 对数正态 |
| 全氮(%)                                    | 0.101 | 0.007 | 7.2   | 0.1   | 0.1  | 0.081~0.119 | 正态   |
| 全磷(%)                                    | 0.074 | 0.008 | 10.7  | 44.8  | 4.9  | 0.057~0.153 | 对数正态 |
| 碱解氮 ( $\text{mg} \cdot \text{kg}^{-1}$ ) | 71.6  | 9.4   | 13.2  | 1.0   | 0.4  | 42.9~106.1  | 正态   |
| 速效磷 ( $\text{mg} \cdot \text{kg}^{-1}$ ) | 8.3   | 6.7   | 79.8  | 146.0 | 1.2  | 3.5~98.2    | 对数正态 |
| 速效钾 ( $\text{mg} \cdot \text{kg}^{-1}$ ) | 114.4 | 17.2  | 15.0  | 2.3   | 10.9 | 87.0~193.0  | 对数正态 |
| 铵态氮 ( $\text{mg} \cdot \text{kg}^{-1}$ ) | 1.8   | 2.0   | 111.1 | 7.4   | 1.9  | 0~14.6      | 对数正态 |
| 硝态氮 ( $\text{mg} \cdot \text{kg}^{-1}$ ) | 18.2  | 22.8  | 125.3 | 8.4   | 2.6  | 0.5~140.8   | 对数正态 |

### 3. 土壤养分的空间变异结构

结构分析主要目的是建立一个最优变异函数的理论模型，定量地描述区域化变量的随机性和结构性，并对变量背景和变异函数理论模型进行专业分析和解释。如果变异函数和相关分析的结果表明区域化变量空间相关性不存在，则空间局部插值方法不适用。利用示范区农田土壤养分数据绘制全方位半方差值计算和半方差图如下：

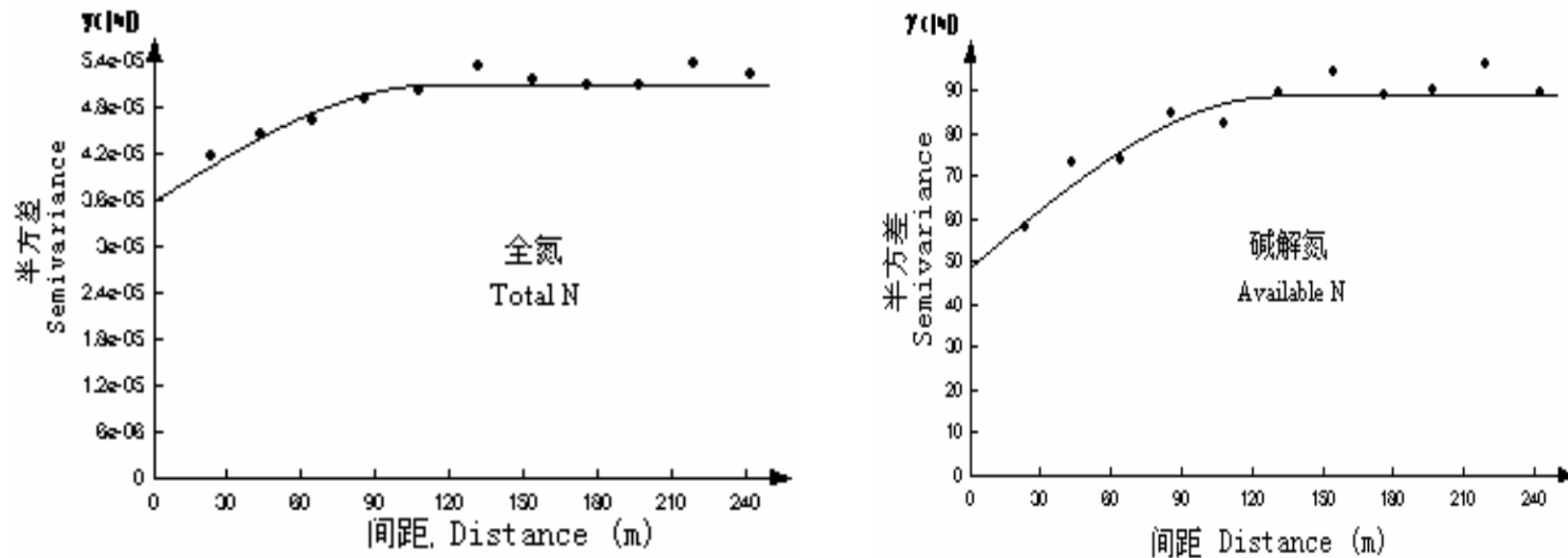


图3-3 试验示范区土壤全氮、碱解氮半方差图

图中小实点是土壤属性的半方差图，实线为其理论模型曲线

示范区土壤全氮、碱解氮半方差实验模型遵从球状模型。

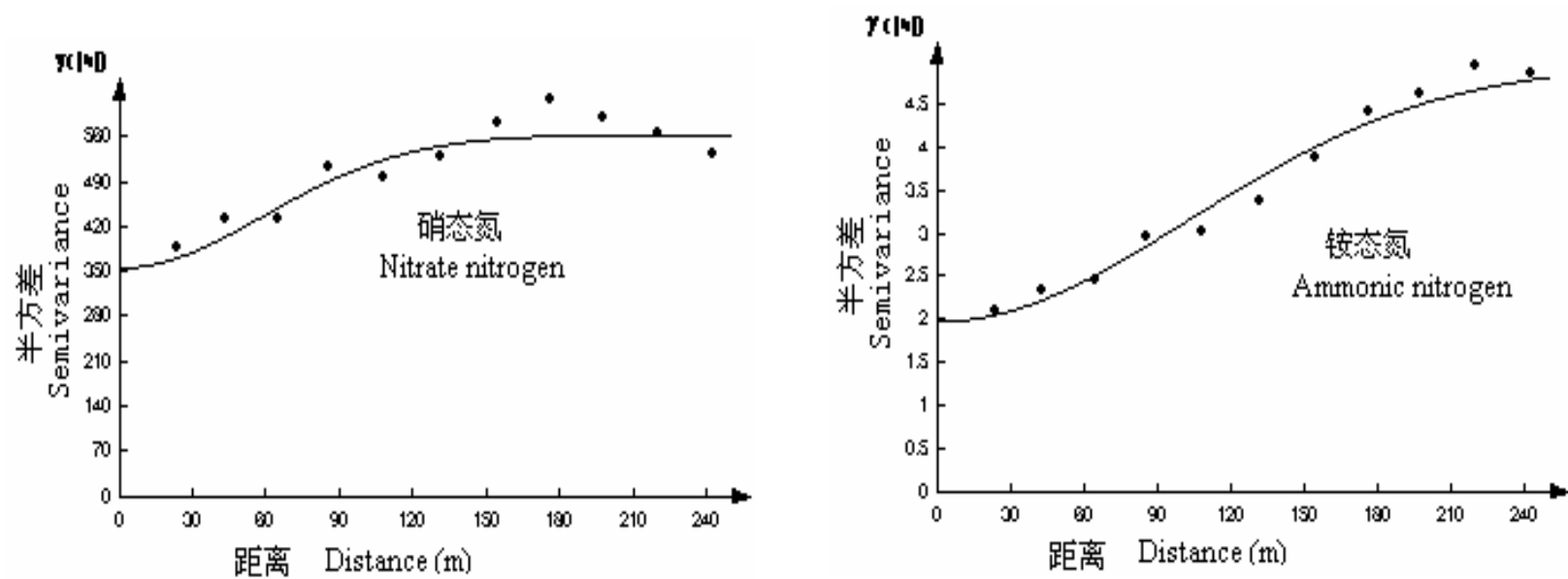


图3-4 试验示范区土壤硝态氮、铵态氮半方差图

铵态氮、硝态氮半方差实验模型遵从高斯模型。

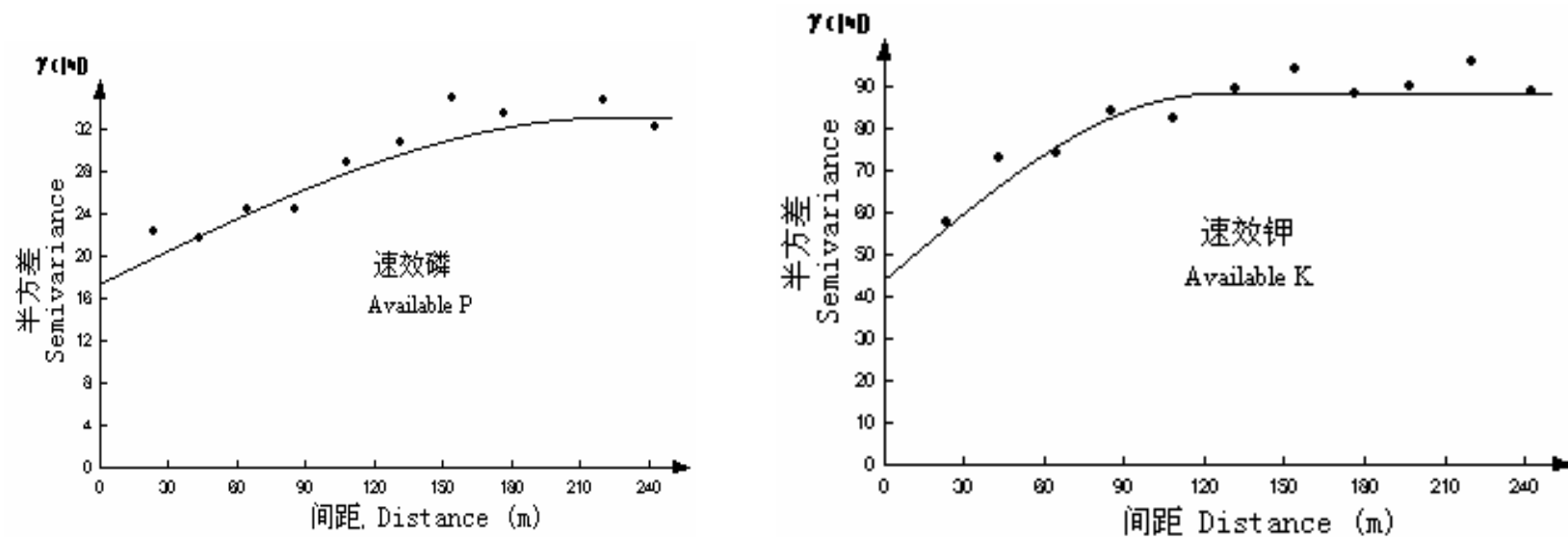


图3-5 试验示范区土壤速效磷、速效钾半方差图

速效磷、速效钾半方差实验模型遵从球状模型。

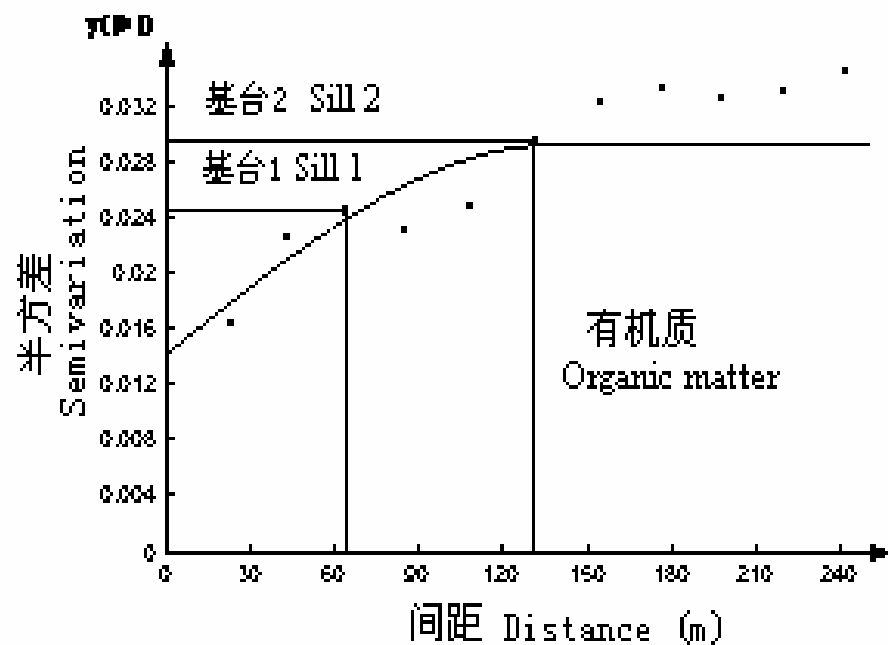


图3-6 试验示范区土壤有机质半方差图

有机质含量半方差图存在两个基台，它们对应于不同的变程，即小尺度空间变异嵌套于大的空间变异之中。若用**1**条曲线模拟半方差图会造成较大偏差，以基台**1**对应于相关间距**70 m**、基台**2**对应于**137 m**的参数配置曲线，其结果则更符合实际。

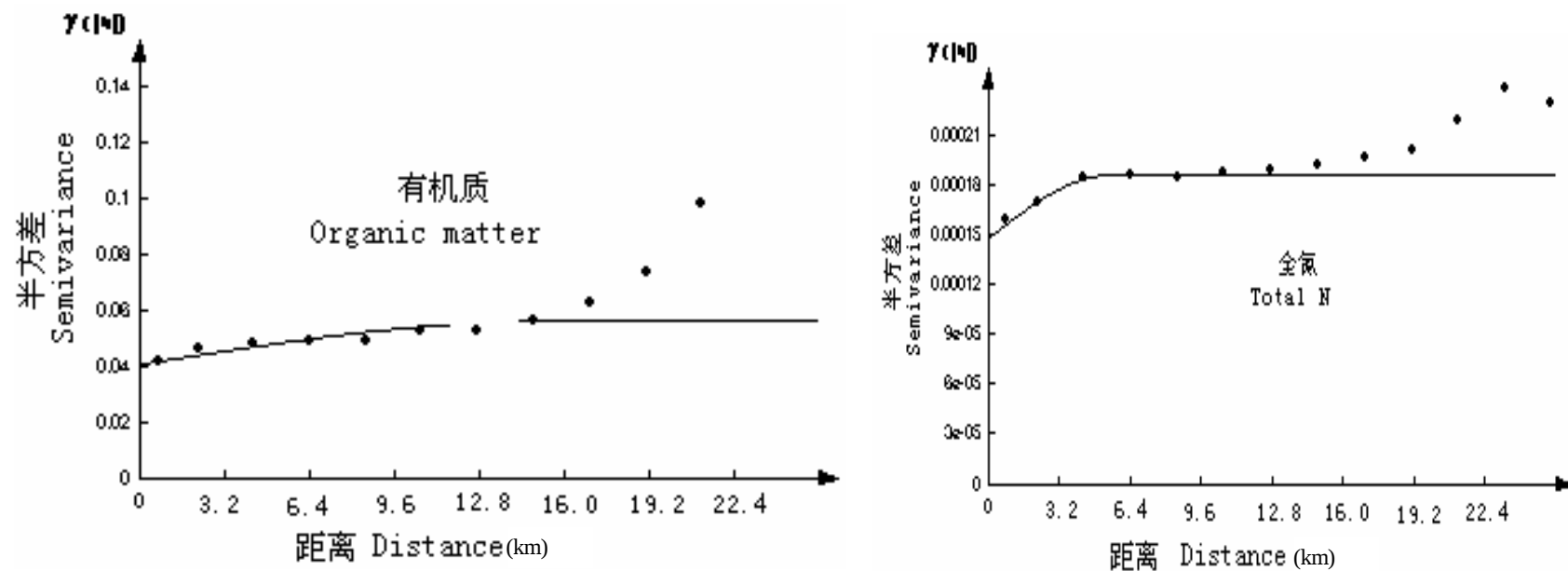


图3-7 栾城县土壤有机质、全氮半方差图  
图中小实点是土壤属性的半方差图，实线为其理论模型曲线

县域有机质、全氮半方差实验模型遵从球状模型。

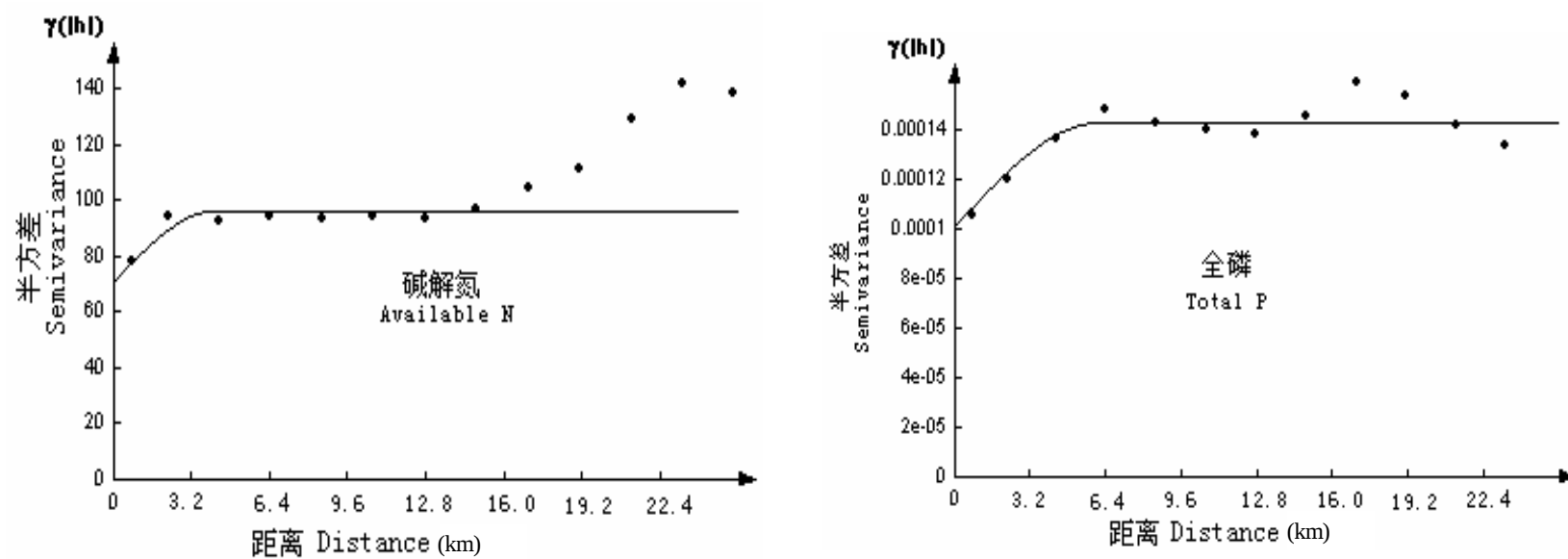


图3-8 栾城县土壤碱解氮、全磷半方差图

县域碱解氮、全磷半方差实验模型遵从球状模型。

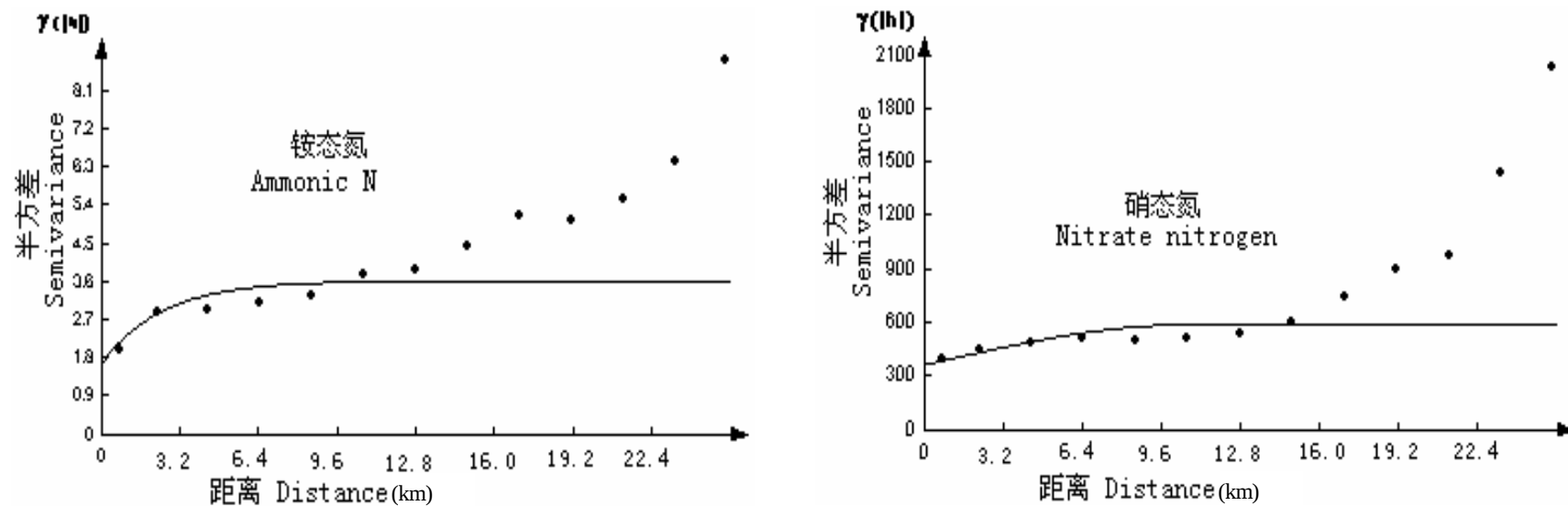


图3-9 栾城县土壤铵态氮、硝态氮半方差图

县域铵态氮、硝态氮半方差实验模型遵从球状模型。

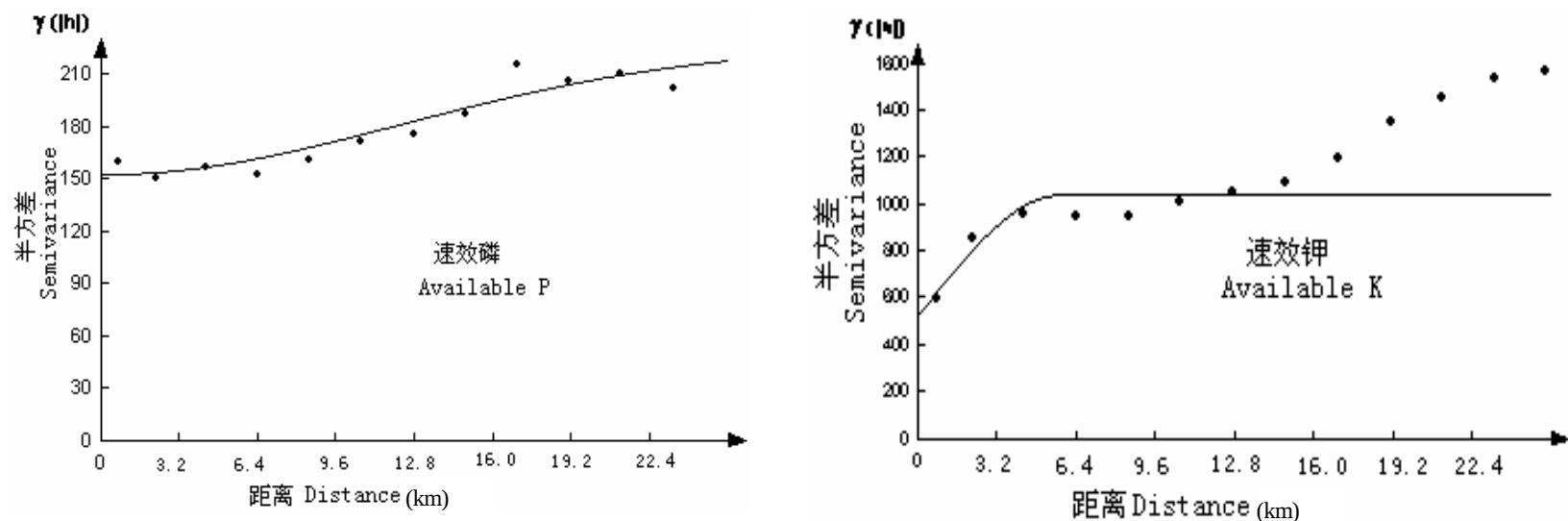


图3-10 栾城县土壤速效磷、速效钾半方差图

栾城县土壤速效磷半方差实验模型遵从高斯模型，而速效钾遵从球状模型。

顺便提一下，可以按一定宽容度和不同方向，作半方差函数分析，计算估测值来说明属性空间变异是否有方向性。从地统计学原理来理解，全方位性半方差最能代表指定区域的空间结构变异，因而不必追求每个方向上的精确结果。

在县域范围内，各土壤养分参数的总空间变异性中随机因素为**50%～85%**，随机因素均大于结构性因素，说明存在不可忽视的小尺度人为（非区域）因素如施肥、耕作等农业管理措施的影响而引起的土壤养分的变异过程。

表3-3 栾城县耕层土壤养分属性半方差函数模型结构及其参数值

| 养分项目 | 块金方差     | 基台值      | 块金方差/基台值 | 最大相关距离(km) | 模型 |
|------|----------|----------|----------|------------|----|
| 有机质  | 0.0414   | 0.057    | 0.726    | 15.6       | 球状 |
| 全氮   | 0.000152 | 0.00019  | 0.800    | 5.5        | 球状 |
| 全磷   | 0.000102 | 0.000144 | 0.708    | 6.2        | 球状 |
| 碱解氮  | 71.5     | 97.0     | 0.737    | 4.2        | 球状 |
| 速效磷  | 153.0    | 178.2    | 0.859    | 13.6       | 高斯 |
| 速效钾  | 528.0    | 1045.0   | 0.505    | 5.8        | 球状 |

块金方差/基台值 ( $c_0/(c+c_0)$ ) 表示空间异质程度，如果该值比较高，说明随机部分引起的空间变异性程度起主要作用，相反，则由空间自相关部分引起的空间变异性程度较大。与县域各土壤养分参数空间变异性相比，示范区中，受随机因素影响减小，结构性因素加大。

表3-4 示范区土壤养分属性半方差函数模型结构及其参数值

| 养分项目 | 块金方差     | 基台值      | 块金方差/基台值 | 最大相关距离 (m) | 模型 |
|------|----------|----------|----------|------------|----|
| 有机质  | 0.014    | 0.029    | 0.517    | 137.5      | 球状 |
| 全氮   | 0.000036 | 0.000052 | 0.698    | 112.5      | 球状 |
| 全磷   | 0.000031 | 0.000063 | 0.494    | 115        | 球状 |
| 碱解氮  | 49.09    | 89.14    | 0.551    | 130        | 球状 |
| 速效磷  | 17.48    | 33.23    | 0.526    | 223        | 球状 |
| 速效钾  | 126.00   | 305.54   | 0.412    | 172        | 球状 |

## 取样数与变异性的关系

农田土壤属性不均匀性或空间变异性，限制着对其管理、研究及制图综合表达。在田间试验中常用小区大小和形状、区组大小和形状、重复次数等来克服可能存在的趋势性误差；也可以根据变异性，在限定到一定精度要求下，采取适量样品来解决。后者假设取样相互独立和样点足够多，中心极限定理成立，则取样数N可按固定的关系式计算。**Wilding**和**Drees（1983）**描述的变差系数(CV)和样品数间的关系为：

$$N = T^2 CV^2 / A^2$$

式中，T为设想的概率水平，A为估计均值的偏差(以%表示)。从式可见，若要达到相同置信水平和估值精度，样点数与变差系数的平方成正比，对变异大的田块，为满足同样置信水平显然所需样点数远较变异性小的田块要多，也就是说总体均值的精确估计依赖取样数目。为了在给定空间范围以显著精度和概率水平或可信水平估计土壤属性的均值，必须确定合适的采样数目。取样点过少，不能真实反映所定总体特征，如此制定技术措施会造成失误；相反若取样点过多，虽然提高精度，但势必费时、费工和资金，故应从精度和经济两方面综合考虑确定适宜的取样数量，即必须平衡考虑高可信度、精度与可行的采样方案。

表 3-5 取样数目估计（个）

|         | 有机质 | 全氮  | 全磷  | 碱解氮 | 速效磷   | 速效钾  |
|---------|-----|-----|-----|-----|-------|------|
| 栾城县耕层土壤 | 7.2 | 6.1 | 7.8 | 5.9 | 153.2 | 31.5 |
| 示范区耕层土壤 | 4.1 | 1.8 | 4.5 | 6.6 | 250.3 | 8.7  |

# 空间插值

空间插值分析是根据农田土壤采样点的已知数据，来估算未采样位置数据的一种方法。通过插值可以将离散的田间采样数据，生成连续的土壤信息分布图。**Kriging**是一种局部的加权平均内插方法，以栾城县农田土壤全氮含量插值图为例。

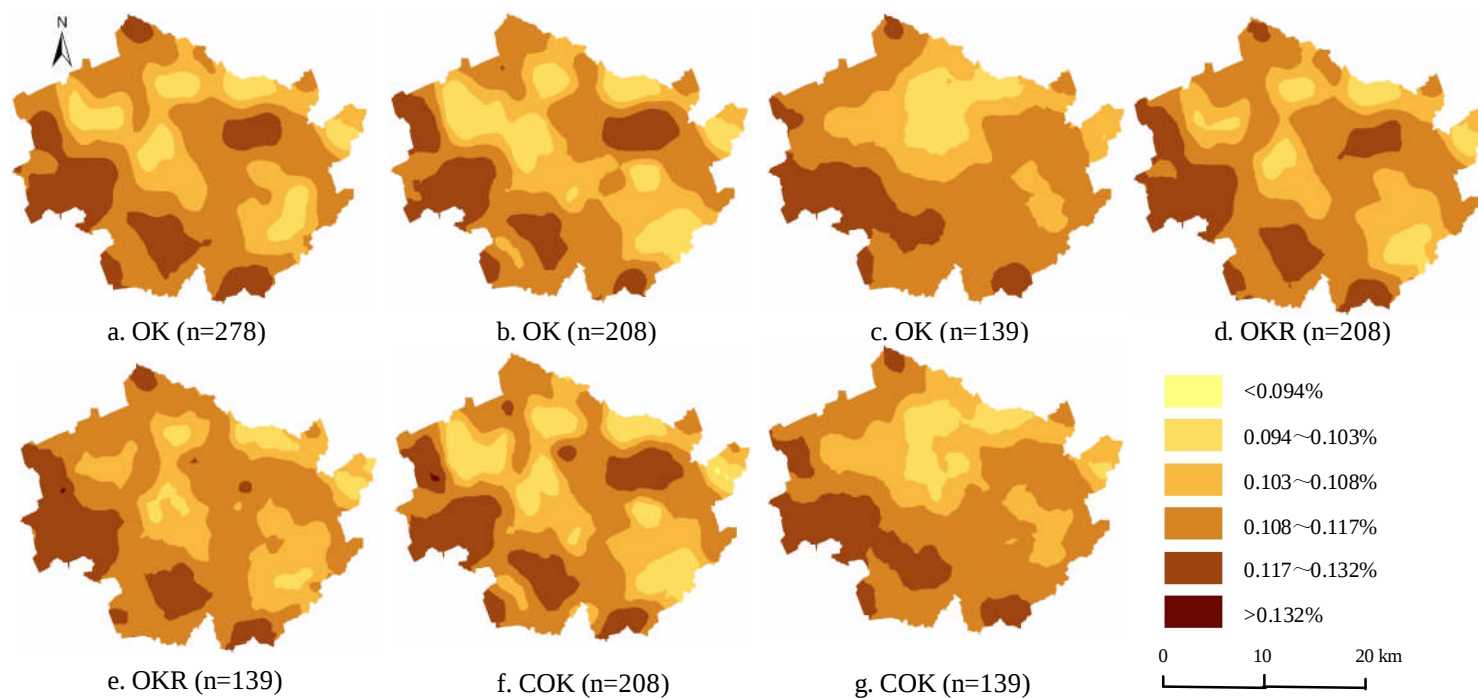


图3-11 在不同样点数下全氮kriging(OK、OKR和COK法)插值图

## 插值精度检验

采用交叉验证法(**cross-validation**)比较不同**kriging**法和不同取样间距下的插值精度。均方根误差(**RMSE**)越小, 预测值与实测值越接近。

从总体中按一定间隔抽取样本, 使样本数减少, 比较插值方法的精度如表**3-12**。

表3-12 Kriging法插值精度及实测值与预测值间相关系数

| 方法          | 普通克里格法<br>(OK) |         |         | 结合回归模型的普通<br>克里格法(OKR) |         | 协同克里格法<br>(COK) |         |
|-------------|----------------|---------|---------|------------------------|---------|-----------------|---------|
| 样点数         | 278            | 208     | 139     | 208                    | 139     | 208             | 139     |
| <i>RMSE</i> | 0.01309        | 0.01354 | 0.01325 | 0.01276                | 0.01281 | 0.01108         | 0.01193 |
| <i>r</i>    | 0.303**        | 0.288** | 0.275** | 0.294**                | 0.285** | 0.604**         | 0.484** |

农田土壤全氮含量的3种克里格插值, 均方根误差(**RMSE**)均较小, 相关性均达极显著水平, 说明预测值与实测值接近; 以**COK**法最佳。

## 4. 讨论与小结

---

土壤养分含量是典型的区域化变量，随着取样间距由小变大，样品变化的随机成分不断增加，小型结构特征逐渐被掩盖，在空间结构分析中体现为块金效应。

不同尺度不同采样间距下，土壤养分含量的半方差图主要遵从球函数模型，仅有速效磷服从高斯模型。在不同的采样间距下，县域范围内和示范区最大相关距离分别为4.2-15.6km和112-223m。

本次示范区采样密度远较一般土壤养分调查为高，在一定程度上反映了小尺度土壤养分含量的空间分异特征；而县域范围内1km采样间距的养分调查则在相当程度上反映了区域空间分布格局。

较小采样间距一般更能揭示土壤养分的变异性，在较大采样间距一般只能揭示土壤养分的空间分布格局，而难以反映其在较小空间尺度上的变异性。

由于土壤参数相关距离差别极大，有关农田尺度的合理采样间距的确定需要进一步研究。

敬请各位专家与同行批  
评指正！

欢迎光临栾城站指导工作！！

