



中国矿业大学
CHINA UNIVERSITY OF MINING AND TECHNOLOGY

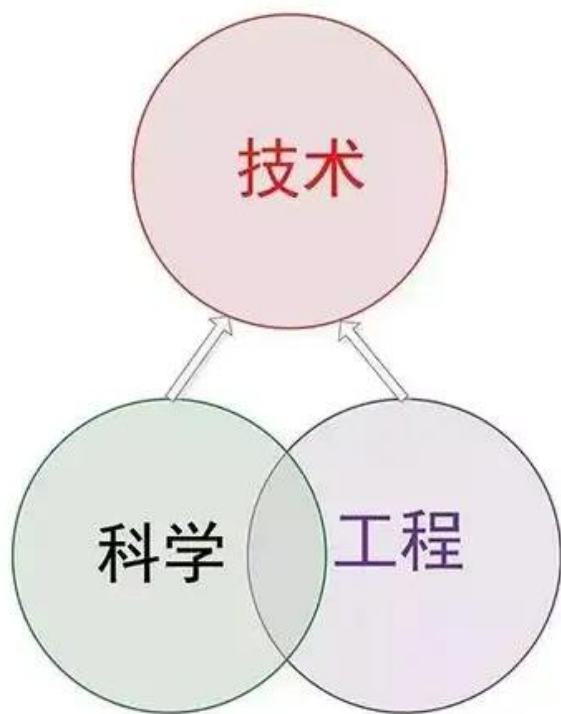
资源信息学

主讲：杨永均 副教授

中国矿业大学环境与测绘学院

第三章 资源空间信息统计分析

- 工学思维与理学科科研思维的差异。
- 大数据时代应当有数理统计思维。

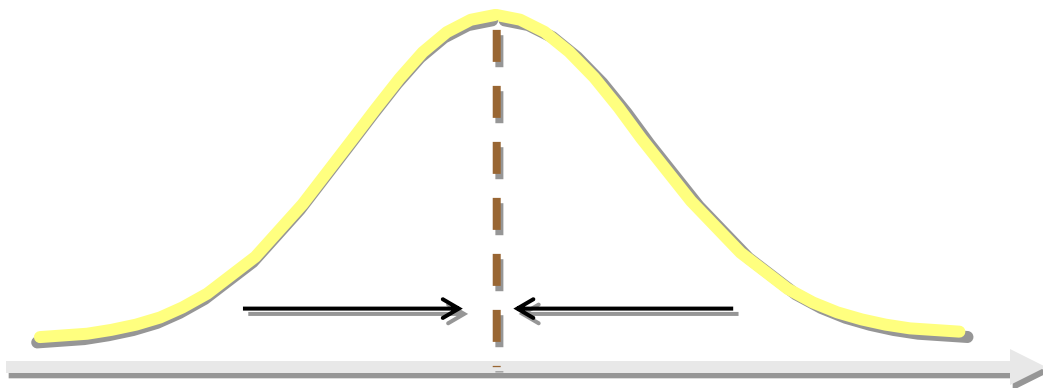


§ 3-1 资源空间信息的统计

一、集中趋势

1、定义

- 一组数据向其中心值靠拢的倾向和程度
- 测度集中趋势就是寻找数据水平的代表值或中心值
- 不同类型的数据用不同的集中趋势测度值
- 低层次数据的测度值适用于高层次的测量数据，但高层次数据的测度值并不适用于低层次的测量数据



§ 3-1 资源空间信息的统计

一、集中趋势

2、指标

(1) 分类数据：众数。一组数据中出现次数最多的变量值。

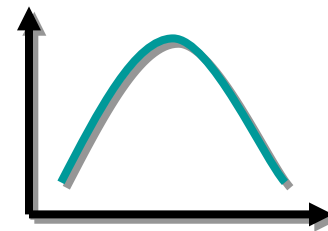
无众数

原始数据: 10 5 9 12 6 8



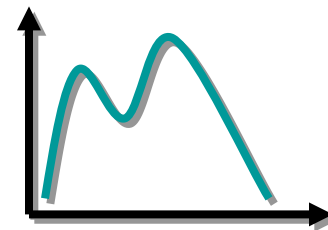
一个众数

原始数据: 6 5 9 8 5 5



多于一个众数

原始数据: 耕地、耕地、园地、林地、园地



§ 3-1 资源空间信息的统计

一、集中趋势

2、指标

(2) 顺序数据：中位数和分位数

排序后处于中间位置上的值



M_e

排序后处于25%和75%位置上的值



Q_L

Q_M

Q_U

资源质量		
质量类别	江苏省	
	个数	累计频数
低	24	24
较低	108	132
一般	93	225
较高	45	270
高	30	300
合计	300	—

§ 3-1 资源空间信息的统计

一、集中趋势

2、指标

(3) 数值型数据：简单、加权、几何平均数

设一组数据为： x_1, x_2, \dots, x_n (总体数据 x_N)

简单平均数

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

加权平均数

$$\bar{x} = \frac{M_1 f_1 + M_2 f_2 + \dots + M_k f_k}{f_1 + f_2 + \dots + f_k} = \frac{\sum_{i=1}^k M_i f_i}{n}$$

几何平均数

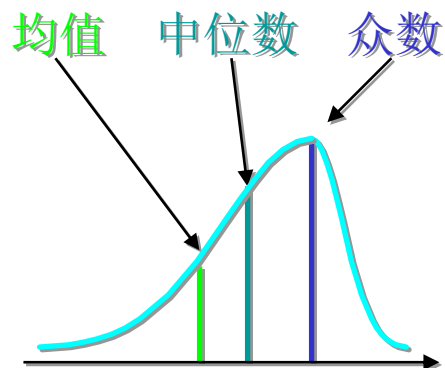
$$G_m = \sqrt[n]{x_1 \times x_2 \times \dots \times x_n} = \sqrt[n]{\prod_{i=1}^n x_i}$$

§ 3-1 资源空间信息的统计

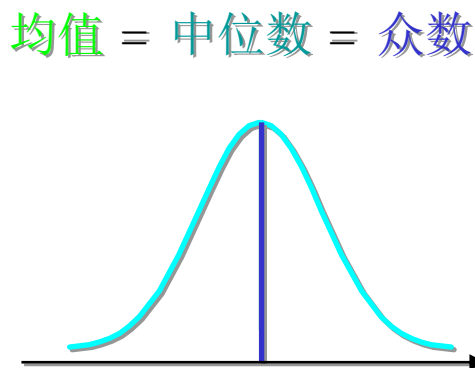
一、集中趋势

3、比较

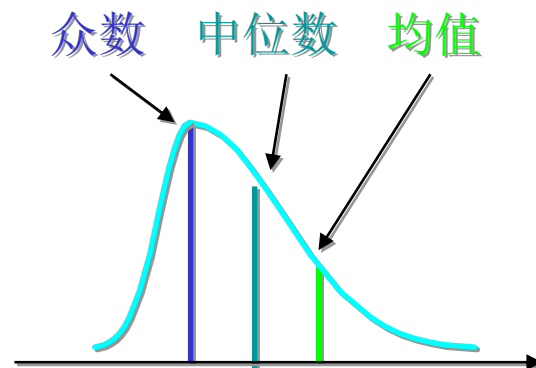
学习重点！



左偏分布



对称分布



右偏分布

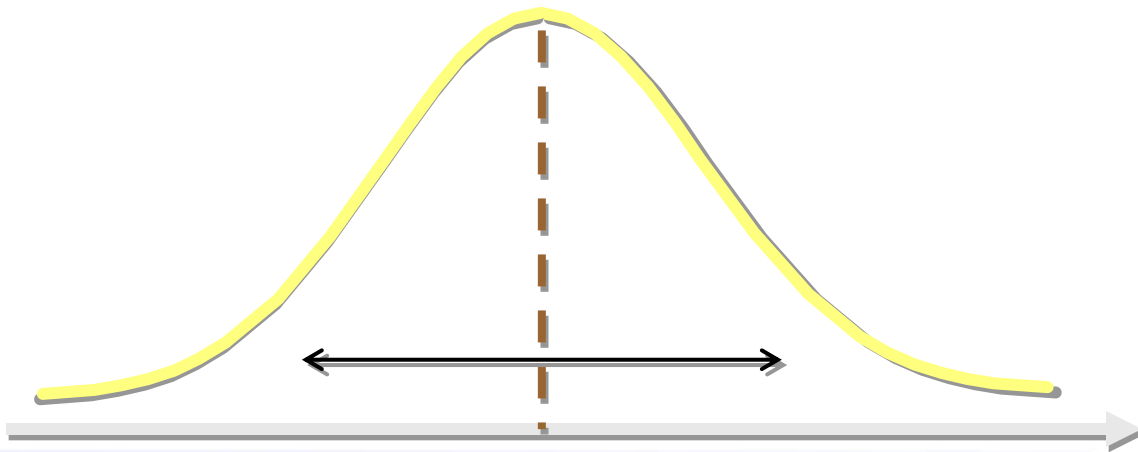
- 众数，不受极端值影响、具有不惟一性、数据分布偏斜程度较大且有明显峰值时应用
- 中位数、不受极端值影响、数据分布偏斜程度较大时应用，最稳健
- 平均数、易受极端值影响、数学性质优良、数据对称分布或接近对称分布时应用，最不稳健

§ 3-1 资源空间信息的统计

二、离散程度

1、定义

- 数据分布的另一个重要特征
- 反映各变量值远离其中心值的程度(离散程度)
- 从另一个侧面说明了集中趋势测度值的代表程度
- 不同类型的数据有不同的离散程度测度值



§ 3-1 资源空间信息的统计

二、离散程度

2、指标

(1) 方差和标准差

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

注意：

样本方差用自由度 $n-1$ 去除！

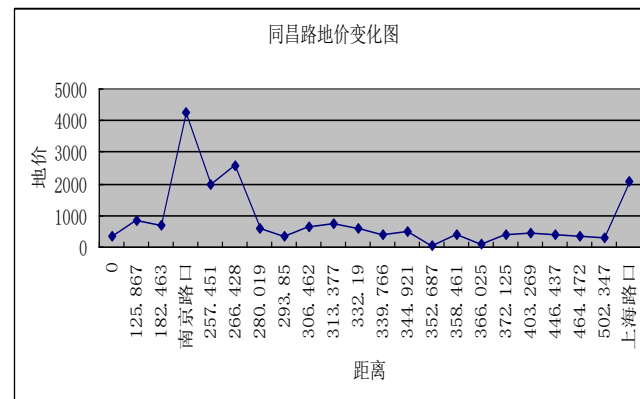
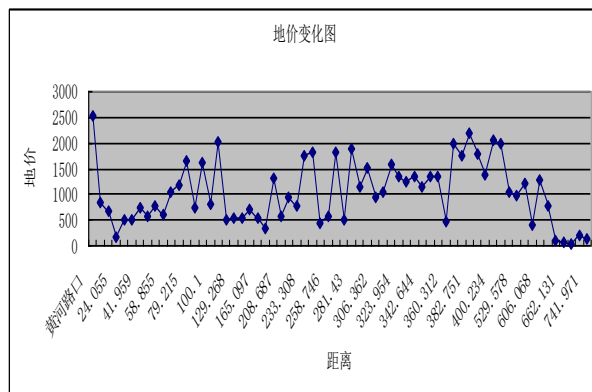
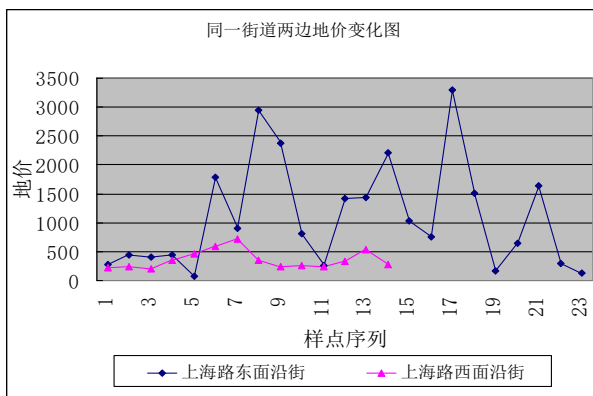
- 自由度是指数据个数与附加给独立的观测值的约束或限制的个数之差
- 从字面涵义来看，自由度是指一组数据中可以自由取值的个数
- 当样本数据的个数为 n 时，若样本平均数确定后，则附加给 n 个观测值的约束个数就是1个，因此只有 $n-1$ 个数据可以自由取值，其中必有一个数据不能自由取值
- 按着这一逻辑，如果对 n 个观测值附加的约束个数为 k 个，自由度则为 $n-k$

§ 3-1 资源空间信息的统计

二、离散程度

2、指标

(1) 方差和标准差



- 上海路地价标准差为471元/平方米，变异系数达到35.1%
- 不同街道地价差异大，这相邻三条街道地价平均值的标准差为403元/平米，变异系数为45%
- 上海路两侧差异很显著，平均值差异达到830元/平米；南京路、同昌路两侧地价差异很小；

§ 3-1 资源空间信息的统计

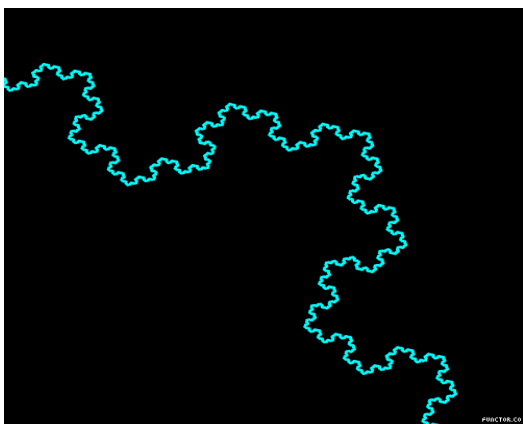
二、离散程度

2、指标

(2) 分形维数

分形：一个粗糙或零碎的几何形状，可以分成数个部分，且每一部分都（至少近似地）是整体缩小后的形状

分维：反映了复杂形体占有空间的有效性，它是复杂形体不规则性的量度。



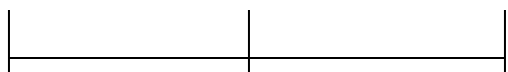
一条海岸线的长度不是绝对的，它取决于你用什么样的尺子来测量。考虑到海岸线被冲刷成的不规则图形，你采用的测量仪器越精细，这条海岸线就越长。

§ 3-1 资源空间信息的统计

二、离散程度

2、指标

(2) 分形维数



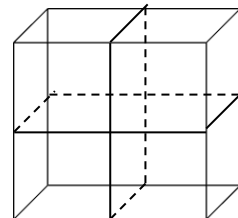
将长度为1的线段分为 n 等分,每段长为 r ,则

$$n \cdot r = 1$$



将面积为1的正方形 n 等分,每一个小正方形的边长为 r ,则

$$n \cdot r^2 = 1$$



将体积为1的正方体 n 等分,每一个小正方体的边长为 r ,则

$$n \cdot r^3 = 1$$

上面的等式中可以看到, r 的幂次实际就是该几何体的空间维数, 可以表示为: $n \cdot r^D = 1$ 。对上式两边取对数得: $D = -\ln n / \ln r$

显然, D 具有维数的概念。

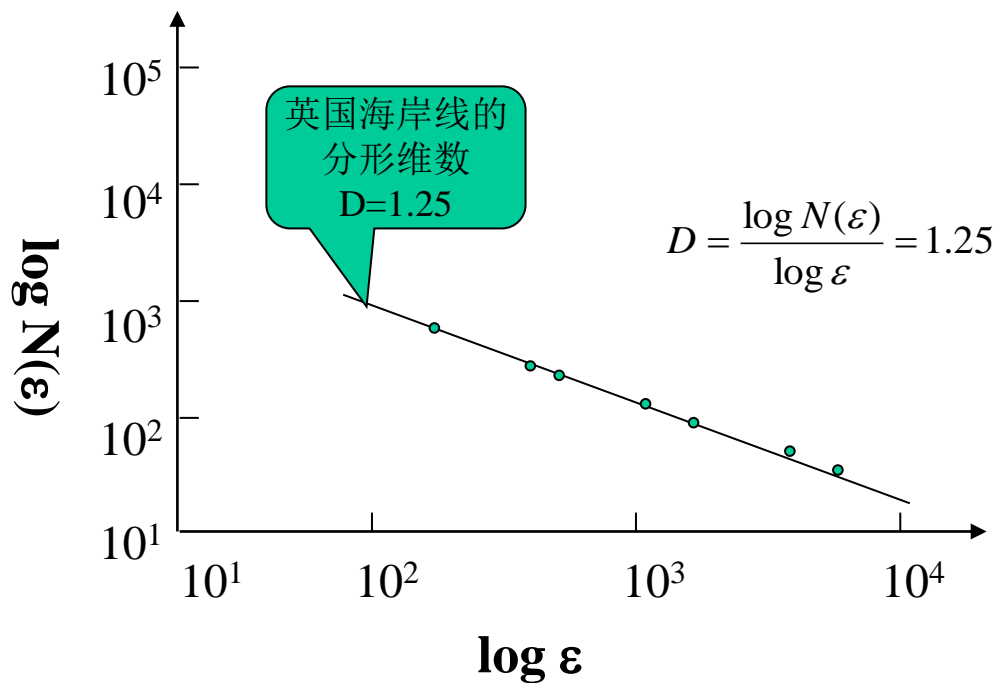
§ 3-1 资源空间信息的统计

二、离散程度

2、指标

(2) 分形维数

挪威东南部海岸线分维数1.52；英国海岸线分维数为1.25，说明挪威海岸线更曲折一些。



§ 3-1 资源空间信息的统计

二、离散程度

2、指标

(2) 分形维数

分维数可用来
较全面准确地
评价资源条件
因素的变化程
度

(2) 分维量测 按相同比例尺, 将 1: 2000 的矿井开拓平面图上的正断层转绘在透明纸上, 绘制成可供量测分维使用的断层迹线网络图(图 1)。采用网格覆盖法, 用边长为 r 的正方形网格覆盖研究区(图 2(a)), 记录含有断层迹线的网格数 $N(r)$ 。然后不断缩小网格尺寸, 若为 r_1, r_2, \dots , 则可得 $N(r_1), N(r_2), \dots$ 。在 $\lg \epsilon - \lg N(r)$ 坐标系中($\epsilon = 1/r$), $\lg \epsilon$ 与 $\lg N(r)$ 的最佳回归直线的斜率即是相似维(图 2(b))。若记录每一含有断层迹线的网格中的迹线长度 $l_i(r)$, 并求得 $P_i(r) = l_i(r)/L$ (L 为研究区断层迹线的总长度), 则在 $\lg \epsilon - \lg N(r)$ 坐标系中, $\lg \epsilon$ 与 $\lg N(r)$ 的最佳回归直线的斜率即是信息维(图 2(c))。

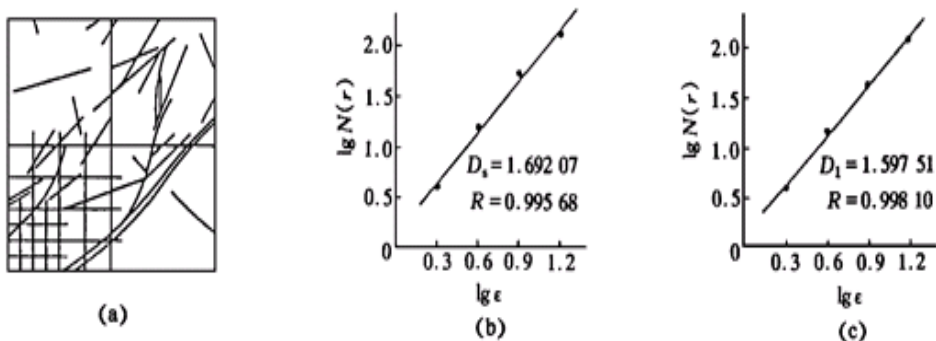


图 2 徐庄煤矿第 8 块段断层迹线及其分维

Fig. 2 The fault trace of the 8 th block and fractal dimension of Xuzhuang Coal Mine

在图 1(a)和图 1(b)上分别划分了 33 和 26 个正方形块段; 取其边长为 1, 则 $r = 1/2, 1/4, 1/8, 1/16$ 。最后采用计算机回归技术求得各块段的相似维(D_1)和信息维(D_1)。

§ 3-1 资源空间信息的统计

三、统计分布

1、定义

连续型随机变量可以取某一区间或整个实数轴上的任意一个值。

$$P(a < X \leq b) = \int_a^b f(x) dx$$

分布函数

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt \quad (-\infty < x < +\infty)$$

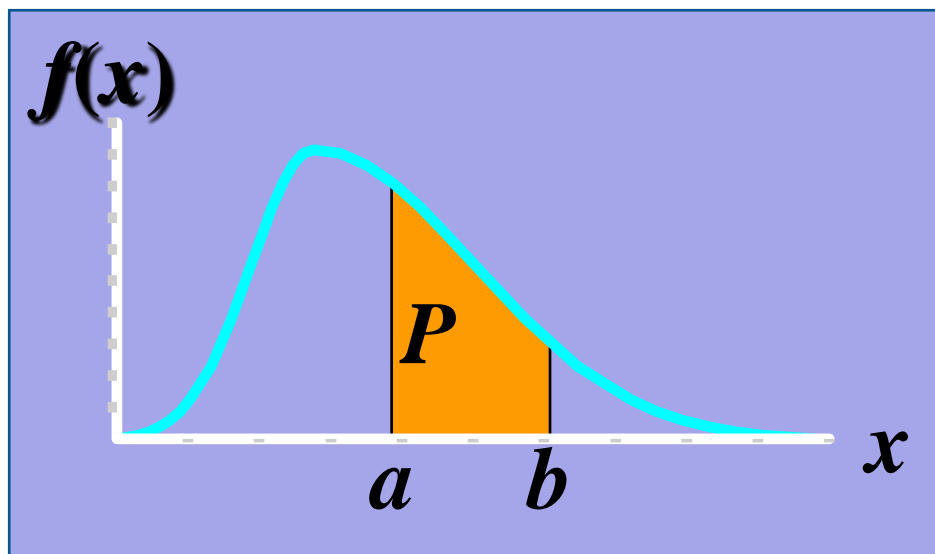
数学期望

$$E(X) = \int_{-\infty}^{+\infty} xf(x) dx = \mu$$

方差

$$D(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x) dx = \sigma^2$$

概率密度函数，表示X的所有取值x及其频数f(x)



§ 3-1 资源空间信息的统计

三、统计分布

2、类型

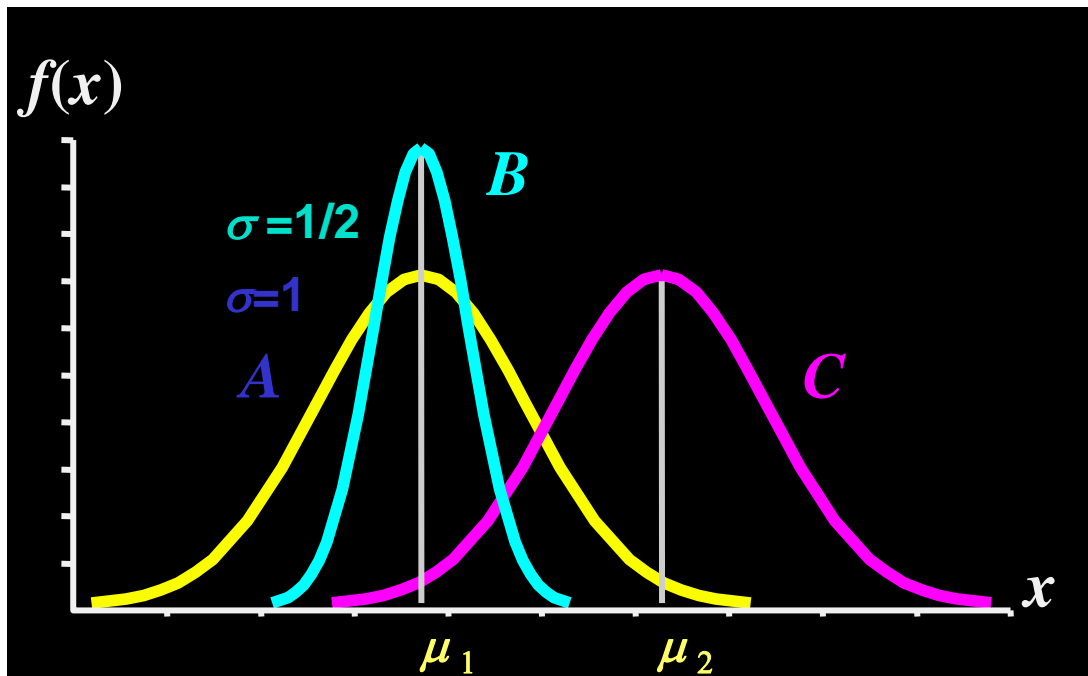
(1) 正态分布

- $f(x)$ = 随机变量 X 的频数
- μ = 正态随机变量 X 的均值
- σ^2 = 正态随机变量 X 的方差
- $\pi = 3.1415926$; $e = 2.71828$
- x = 随机变量的取值 ($-\infty < x < \infty$)

$$X \sim N(\mu, \sigma^2)$$

$$Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad -\infty < x < \infty$$



§ 3-1 资源空间信息的统计

三、统计分布

2、类型

(2) 对数正态分布

记作：

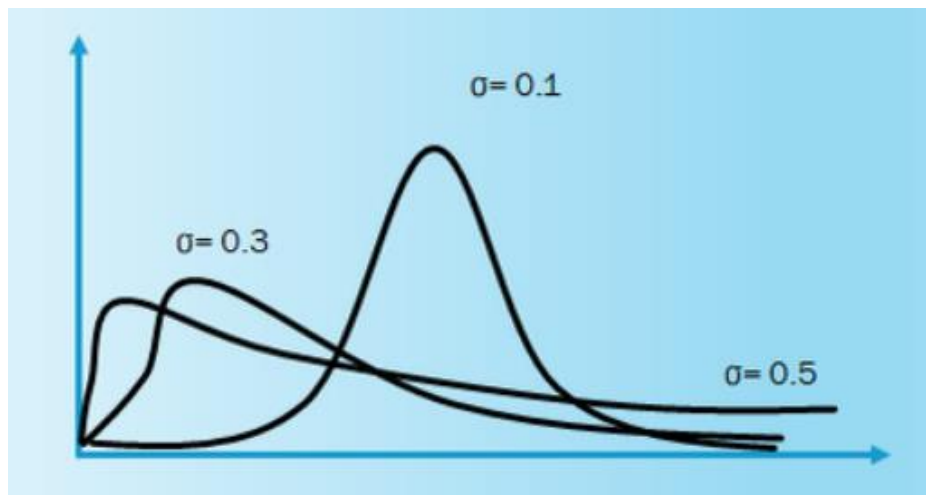
$$f(x) = \begin{cases} \frac{1}{x\sigma\sqrt{2\pi}} e^{-(\log x - \mu)^2 / (2\sigma^2)} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

$$X \sim \text{LogN}(\mu, \sigma)$$

数学期望和方差

$$E(X) = e^{\mu + \sigma^2/2}$$

$$D(X) = (e^{\sigma^2} - 1) e^{2\mu + \sigma^2}$$



§ 3-1 资源空间信息的统计

三、统计分布

2、类型

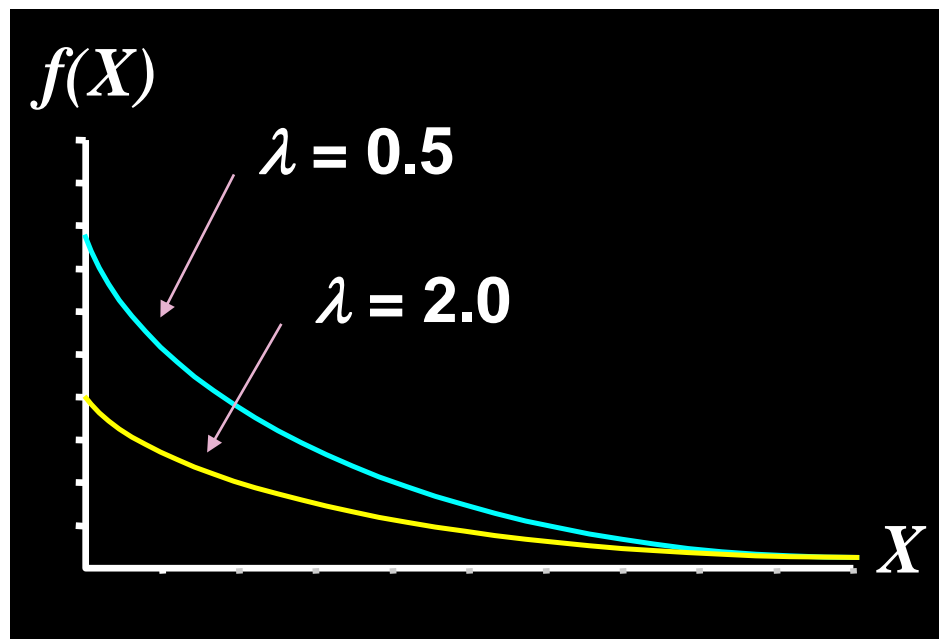
(3) 指数分布

称 X 服从参数为 λ 的指数分布，记为 $X \sim E(\lambda)$

数学期望和方差

$$E(X) = \frac{1}{\lambda} ; \quad D(X) = \frac{1}{\lambda^2}$$

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 (\lambda > 0) \\ 0 & \text{其他} \end{cases}$$



3、应用

功能区土壤划分 Zoning of soil by function	文献量 Volumn of literature	样点数 Number of sampling sites	分布类型 Distribution type	范围 Range (mg kg ⁻¹)	几何均值 Geometric mean (mg kg ⁻¹)	变异系数 CV (%)	P_{k-s}
清洁土壤Clean soil	6	269	正态分布	18.80 ~ 39.00	26.17 (1.372)	31.27	0.914
城区土壤Soil in urban area	4	126	正态分布	20.74 ~ 65.56	32.88 (1.758)	57.45	0.915
工矿土壤Soil in industrial and mining area	25	299	对数正态分布	45.90 ~ 6 350	244.5 (4.581)	185.9	0.152
非矿区农用土壤 Farmland soil in agricultural area	19	722	正态分布	26.18 ~ 80.95	42.93 (1.466)	39.96	0.541
矿区农用土壤 Farmland soil in industrial and mining area	32	1 019	对数正态分布	42.50 ~ 4 382	347.6 (2.530)	139.0	0.971
合计Total	86	2 435	对数正态分布	18.80 ~ 6 350	147.9 (4.150)	220.8	0.068

注 D = 0.05 市三口丁太八女。坪口市数据由日后二津芳 M₁ D = 0.05 “ / \ ”

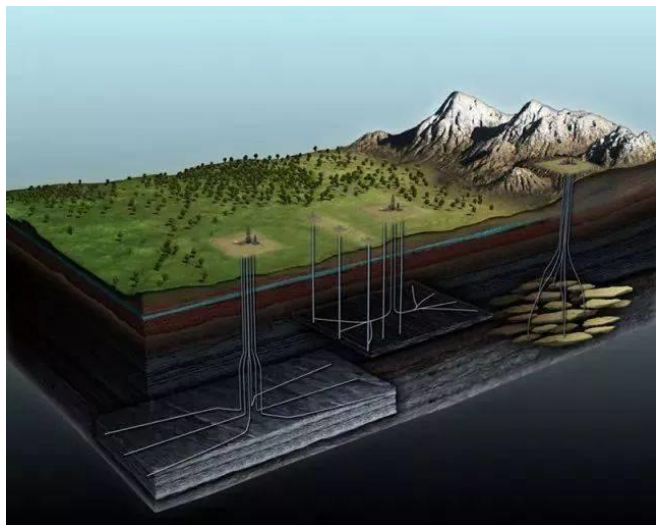
§ 3-2 资源空间信息的估计

一、研究程度

学习重点！

1、定义

- 衡量勘探工作质量和勘探充分程度的指标
- 当采样点间的距离一定时，变化程度愈低，则研究程度愈高。反之，则研究程度愈低。
- 研究程度与变化程度成反比。当变化程度一定时，采样数目愈多，研究程度愈高。反之，研究程度愈低。
- 研究程度与采样的数目成正比。



§ 3-2 资源空间信息的估计

一、研究程度

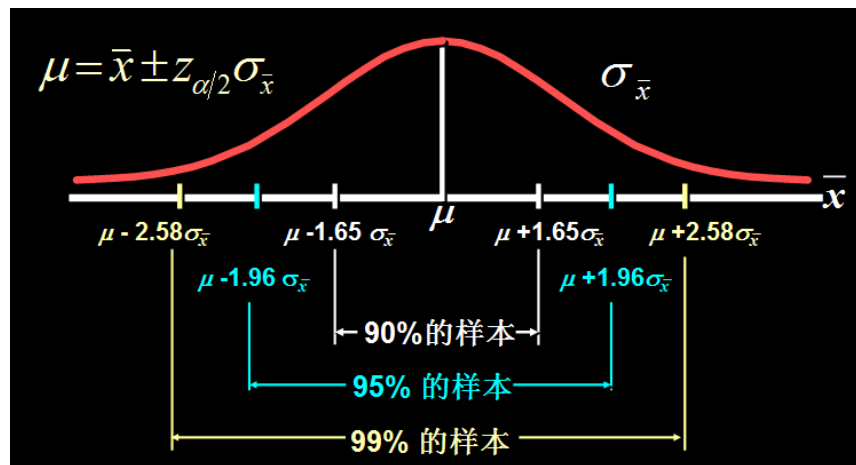
2、方法

某一铜矿进行勘探时，共打了107个钻孔，求得矿体的平均厚度为11.6m，标准差8.66m。今要求以95%的概率、期望参数平均值的误差不超过1m，试问目前勘探工程数目是否足够？

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \quad \text{或} \quad \bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}} (\sigma \text{ 未知})$$

$$m = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 1.96 * \frac{8.66}{\sqrt{107}} = 1.64 \text{米}$$

$$n = \left(Z_{\alpha/2} \frac{\sigma}{m} \right)^2 = \left(1.96 * \frac{8.66}{1} \right)^2 = 288 \text{个}$$

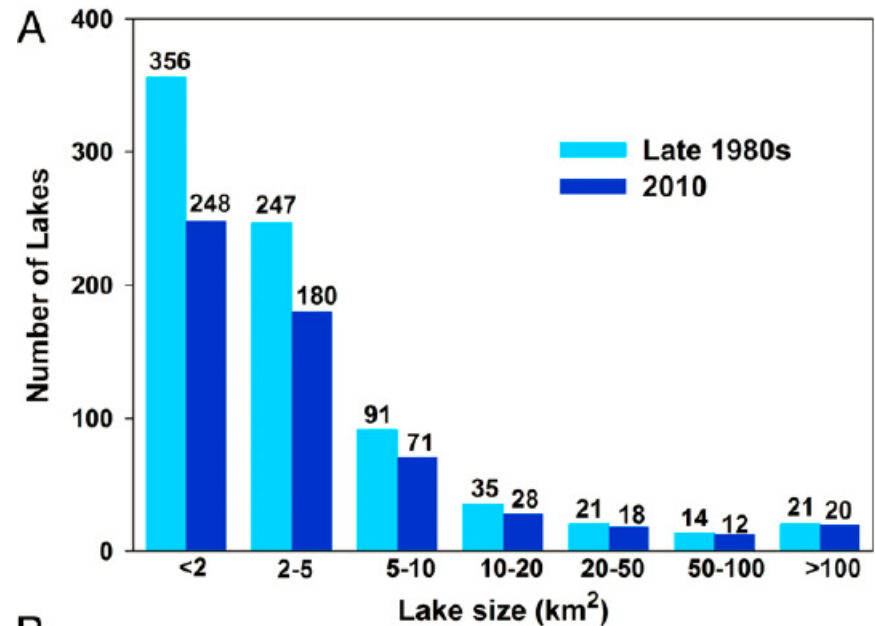
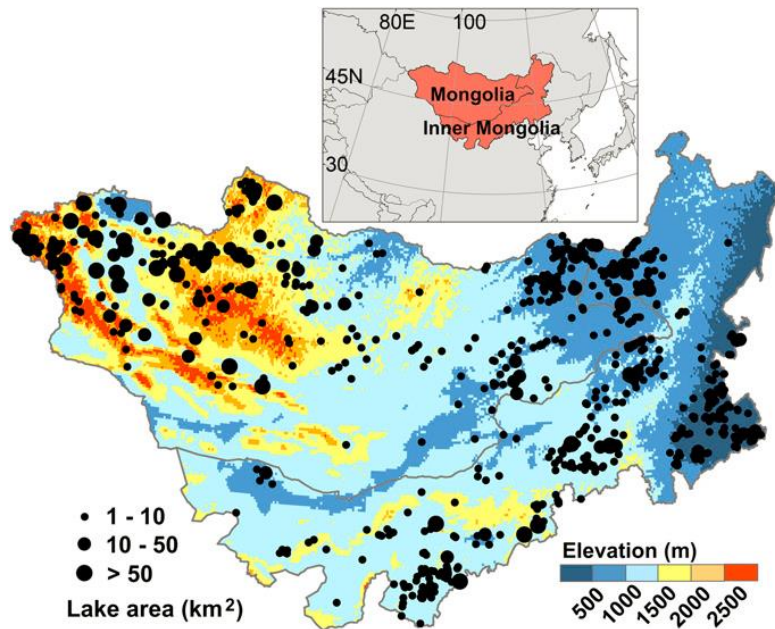


变化程度大，目前研究程度不够，
勘探工程数目还需要增加288-
107=181个。

§ 3-3 资源空间信息的建模

一、回归分析

1、定义



北京大学方精云院士课题组发现蒙古高原湖泊减少，是什么原因造成的？煤炭开采、GDP、人口、农业，哪个因素起主要作用？

§ 3-3 资源空间信息的建模

一、回归分析

1、定义

一个因变量与两个及两个以上自变量的回归

描述因变量 y 如何依赖于自变量 x_1, x_2, \dots, x_k 和误差项 ε 的方程, 称为多元回归模型

涉及 k 个自变量的多元回归模型可表示为

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \Lambda + \beta_k x_k + \varepsilon$$

- $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ 是参数
- ε 是被称为误差项的随机变量
- y 是 x_1, x_2, \dots, x_k 的线性函数加上误差项 ε
- ε 包含在 y 里面但不能被 k 个自变量的线性关系所解释的变异性

§ 3-3 资源空间信息的建模

一、回归分析

2、应用

SPSS Statistics 数据编辑器

文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 图形(S) 实用程序(U) 附加内容(O) 窗口(W) 帮助

13: 可见: 6 变量的 6

	台站	经度X度	海拔m	年降水量mm	年蒸发量mm	变量	变量	变量	变量	变量	变量	变量
1	安西	98	1170.80	48.25	2835.57							
2	白银	104	1707.20	193.72	1947.97							
3	定西	104	1908.80	413.94	1538.10							
4	古浪	102	2072.40	358.60	1756.79							
5	和政	103	2136.40	615.04	1317.64							
6	徽县	106		752.42	1167.44							
7	会宁	108		435.43	1632.93							
8	靖远	104		238.55	1594.28							
9	酒泉	98		87.85	2005.45							
10	兰州	103		316.00	1410.15							
11	礼县	108		503.73	1318.59							
12	临洮	103		554.04	1229.31							
13	临夏	103		502.07	1282.17							
14	玛曲	102		611.78	1279.50							
15	岷县	104		603.66	1159.48							
16	秦安	105		501.67	1414.59							
17	天水	108		540.16	1277.33							
18	天祝松山	103	2726.70	264.15	1705.98							
19	通渭	108	1765.00	427.11	1295.52							
20	通渭华家岭	104	2450.00	513.09	1303.09							
21	武山	104.88	34.7	1495.00	478.21							
22	榆中	104.08	35.9	1873.70	395.25							
23	成县	105.72	33.8	970.00	650.14							

报告 描述统计 表(T) BFM 分析 比较均值(M) 一般线性模型(G) 广义线性模型 混合模型(X) 相关(C) 回归(R) 对数线性模型(O) 神经网络 分类(F) 降维 度量(S) 非参数检验(N) 预测(I) 生存函数(S) 多重响应(U) 缺失值分析(Y) 多重回归(I) 复杂抽样(L) 权重控制(Q) ROC 曲线图(V) 线性(L) 曲线估计(C) 部分最小平方... 二元 Logistic... 多项 Logistic... 有序... Probit... 非线性(N) 权重估计(W) 两阶最小二乘法(2) 最佳尺度(CATREG)...

数据视图 变量视图

线性(L)...

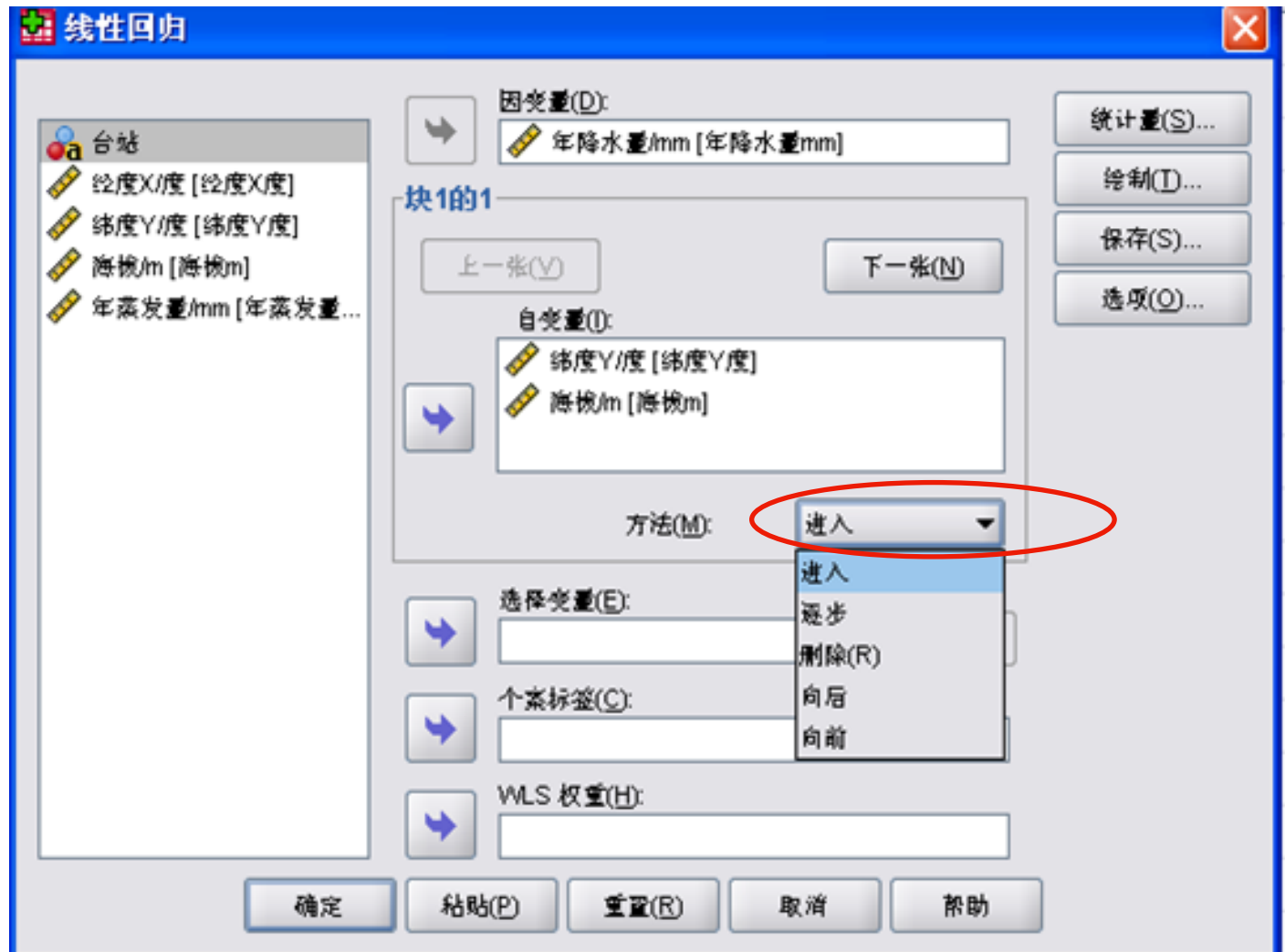
SPSS Statistics 处理器已就绪 0KB/S 0KB/S

开始 任务栏 1 ppt F:\专业书\ 一元线性回... 3.pdf - Fo... *输出1 [文... *未标题2 [自由的我 15:07

§ 3-3 资源空间信息的建模

一、回归分析

2、应用



§ 3-3 资源空间信息的建模

一、回归分析

2、应用

输入 / 移去的变量

模型	输入的变量	移去的变量	方法
1	海拔/m, 纬度Y/度 ^a	.	输入

a. 已输入所有请求的变量。

模型汇总

模型	R	R 方	调整 R 方	标准 估计的误差
1	.909 ^a	.827	.820	91.20807

表示变量在一起,
可以解释因变量
82.7%的变异性

a. 预测变量: (常量), 海拔/m, 纬度Y/度。

§ 3-3 资源空间信息的建模

一、回归分析

2、应用

Anova^b

模型		平方和	df	均方	F	Sig.
1	回归	1985197.596	2	992598.798	119.318	.000 ^a
	残差	415945.593	50	8318.912		
	总计	2401143.189	52			

$F > 5.06$ (查表结果)

a. 预测变量: (常量), 海拔/m, 纬度Y/度。

b. 因变量: 年降水量/mm

系数^a

模型		非标准化系数		标准系数	t	Sig.
		B	标准 误差	试用版		
1	(常量)	3295.128	205.455		16.038	.000
	纬度Y/度	-81.174	5.386	-.892	-15.071	.000
	海拔/m	.036	.021	.103	1.733	.089

a. 因变量: 年降水量/mm

$$P = 3295.128 - 81.174y + 0.036a$$

§ 3-3 资源空间信息的建模

二、时间序列分析

1、定义

- 同一现象在不同时间上的相继观察值排列而成的数列
- 形式上由现象所属的时间和现象在不同时间上的观察值两部分组成
- 排列的时间可以是年份、季度、月份或其他任何时间形式

遥感指数（NDVI、NPP、NDWI、GPP.....）

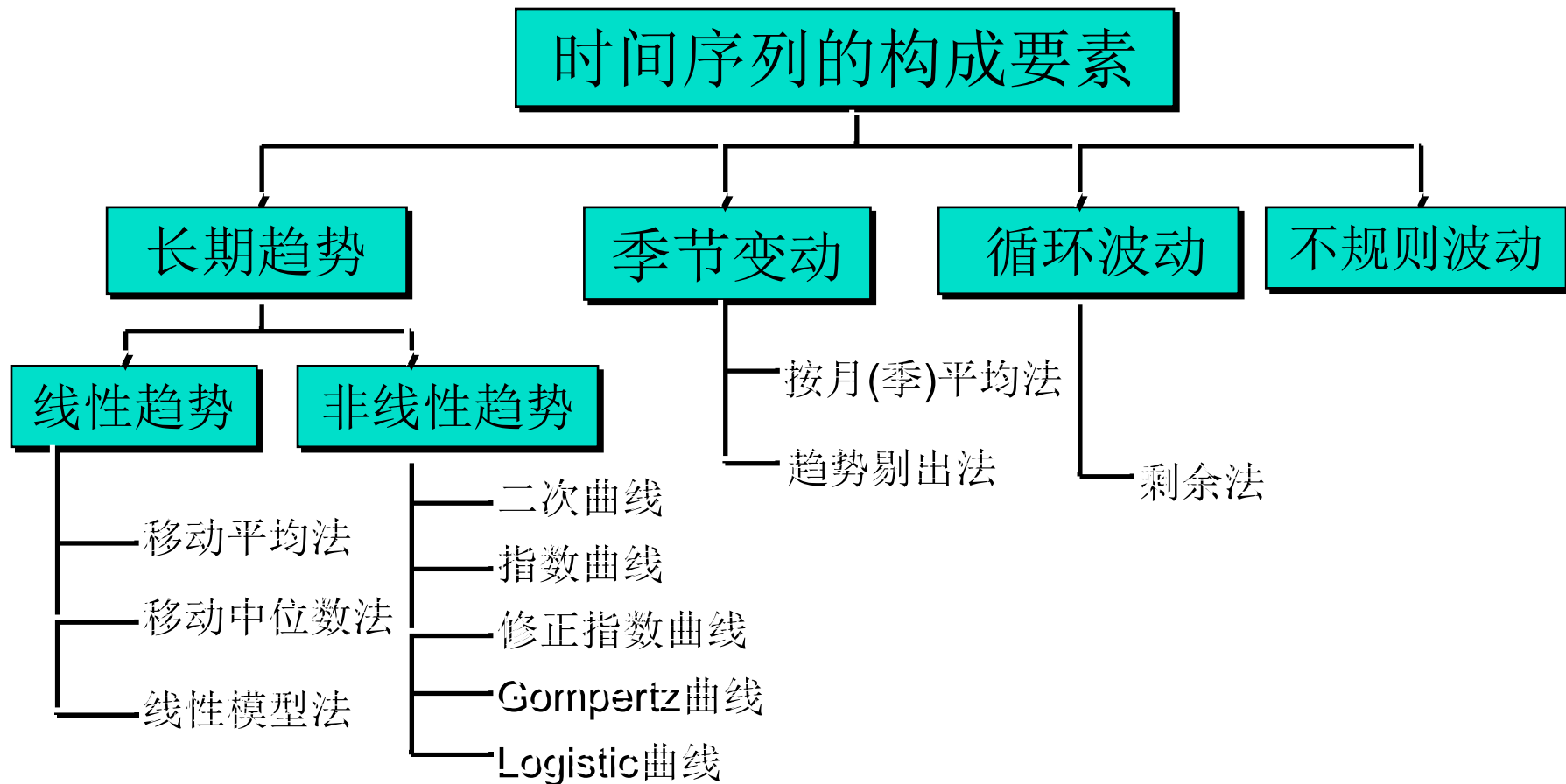
统计资料（GDP、GEP、人口、土地面积、矿产量.....）

定位观测（气象、土壤、植被、大气.....）

§ 3-3 资源空间信息的建模

二、时间序列分析

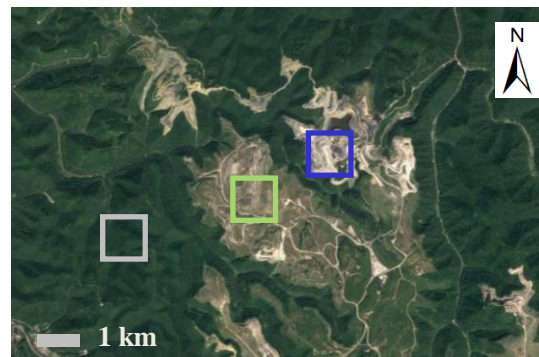
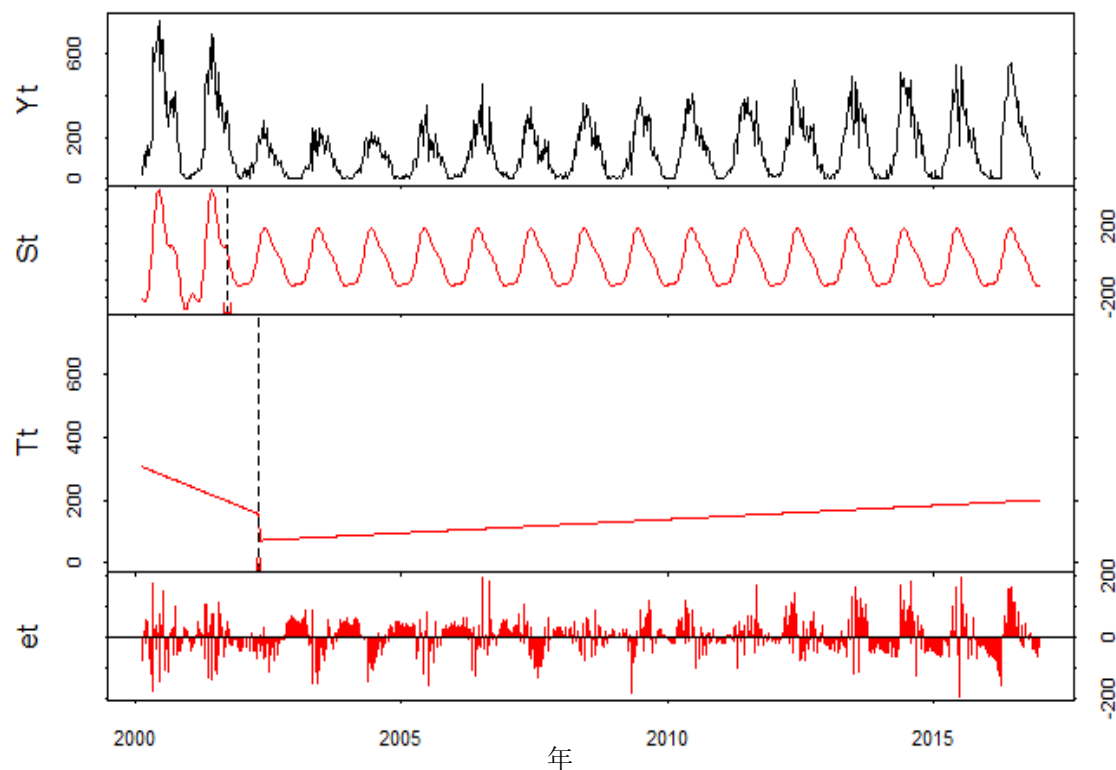
1、定义



§ 3-3 资源空间信息的建模

二、时间序列分析

2、应用



拟合方程

$$Y_t = T_t + S_t + e_t$$

长期趋势

$$T_t = \alpha_i + \beta_i t$$

季节趋势

$$S_t = \sum_{k=1}^K \alpha_{j,k} \sin\left(\frac{2\pi kt}{f} + \delta_{j,k}\right)$$

§ 3-3 资源空间信息的建模

三、空间插值分析

学习重点！

1、区域化变量

(1) 区域化变量：

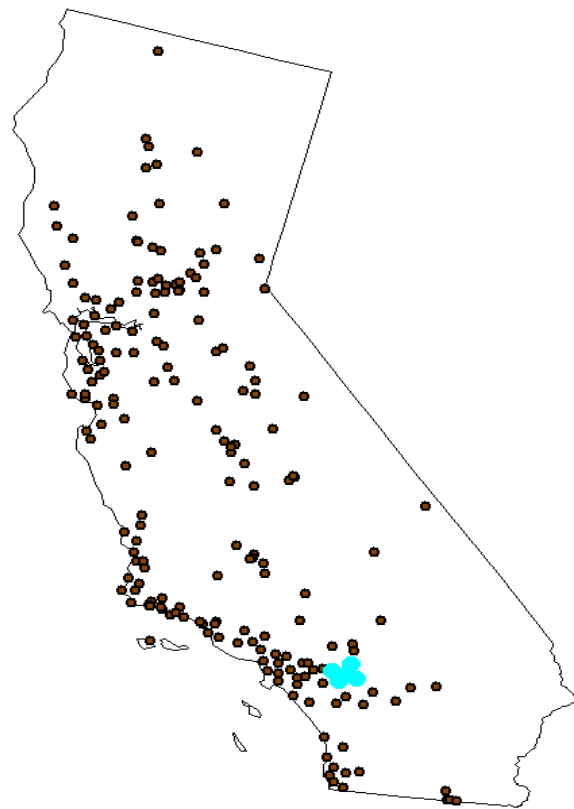
以空间点 x 的三个直角坐标 (x_1, x_2, x_3) 为自变量的随机场称为区域化变量, 或者区域化随机变量。

(2) 区域化变量的性质：

首先区域化变量是一个随机函数，具有随机性质。其次，区域化变量具有一般的或者平均的结构性，即变量在点 x 和 $x+h$ 处的数值 $Z(x)$ 和 $Z(x+h)$ 具有某种程度的自相关，这种自相关依赖于两点间的距离及变量特征，这就是其结构性。

(3) 区域化变量的特点：

空间的局限性；不同程度的连续性；不同类型的各向异性

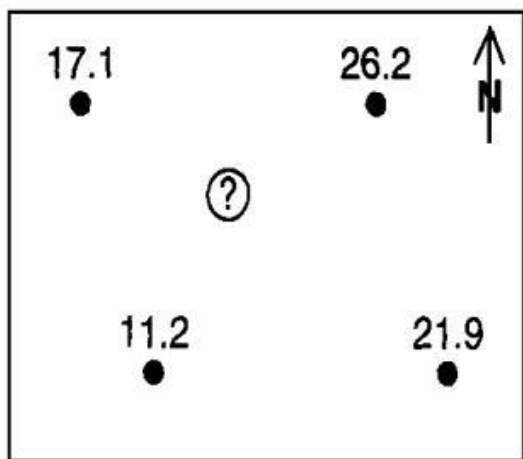


§ 3-3 资源空间信息的建模

三、空间插值分析

2、空间插值

空间插值常用于将离散点的测量数据转换为连续的数据曲面，空间插值的理论假设是：**空间位置上越靠近的点，越可能具有相似的特征值，而距离越远的点，其特征值相似的可能性越小。**



主要的内插方法有：

- 反距离加权 (Inverse Distance Weighted)
- 全局多项式 (Global Polynomial Interpolation)
- 局部多项式 (Local Polynomial Interpolation)
- 径向基函数 (Radial Basis Functions)
- 克里格内插 (Kriging)

§ 3-3 资源空间信息的建模

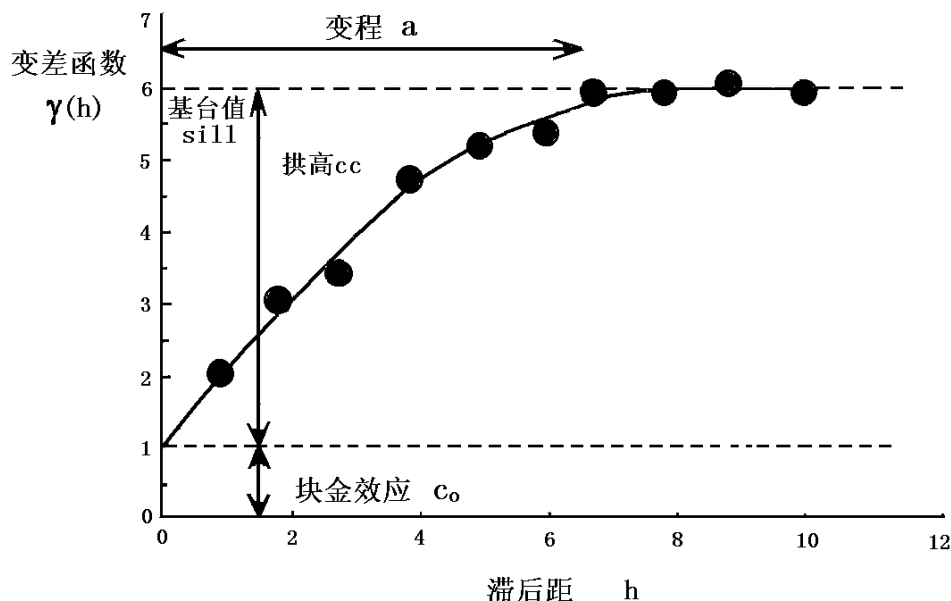
三、空间插值分析

学习重点！

3、半变异函数

(1) 半变异函数

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



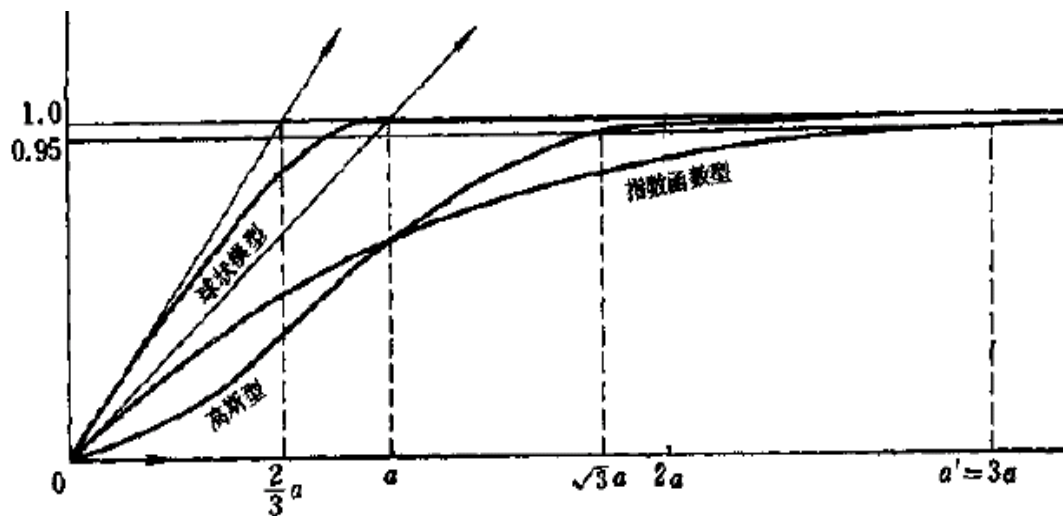
- 半变异值的变化随着距离的加大而增加，这主要是由于半变异函数是事物空间相关系数的表现，当两事物彼此距离较小时，它们是相似的，而半变异值较小；反之，半变异值较大。
- 当变异函数随着间隔距离 h 的增大，从非零值达到一个相对稳定的常数时，该常数称为基台值 C_0+C 。
- 当间隔距离 $h=0$ 时， $\gamma(0)=C_0$ ，该值称为块金值或块金方差（nugget variance）。
- 基台值是系统或系统属性中最大的变异，变异函数达到基台值时的间隔距离 a 称为变程。
- 变程表示在 $h \geq a$ 以后，区域化变量 $Z(x)$ 空间相关性消失。

§ 3-3 资源空间信息的建模

三、空间插值分析

学习重点!

3、半变异函数



球状模型

$$\gamma(h) = \begin{cases} 0 & h = 0 \\ C_0 + C(\frac{3h}{2a} - \frac{h^3}{2a^3}) & 0 < h \leq a \\ C_0 + C & h > a \end{cases}$$

指数模型

$$\gamma(h) = \begin{cases} 0 & h = 0 \\ C_0 + C(1 - e^{-\frac{h}{a}}) & h > 0 \end{cases}$$

高斯模型

$$\gamma(h) = \begin{cases} 0 & h = 0 \\ C_0 + C(1 - e^{-\frac{h^2}{a^2}}) & h > 0 \end{cases}$$

模型	通过原点切线与基台值线交点的横坐标	变程	原点处的性状
球状	$2a/3$	a	直线
指数	a	$3a$	直线
高斯	无交点	$\sqrt{3} a$	抛物线

§ 3-3 资源空间信息的建模

三、空间插值分析

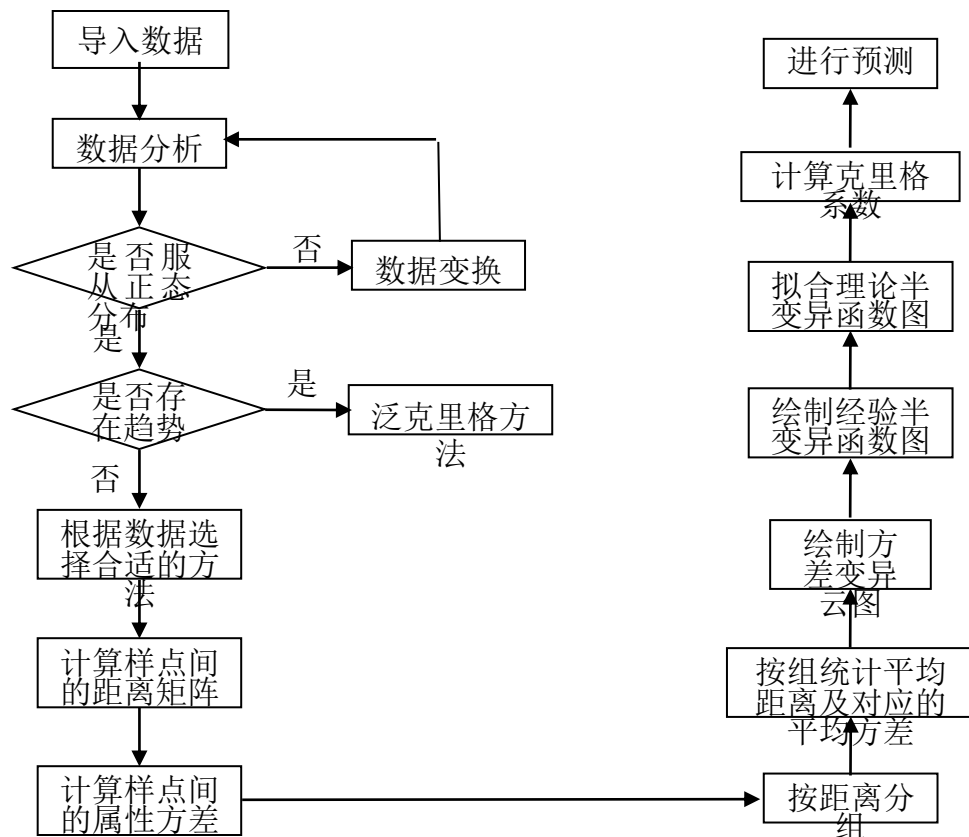
4、插值方法

上机实验！

(2) 克里金插值

$$z^*(x_0) = \sum_{i=1}^n \lambda_i z(x_i)$$

$$\begin{cases} \sum_{j=1}^n \lambda_j \gamma(x_i, x_j) + \mu = \gamma(x_i, x) \\ \sum_{i=1}^n \lambda_i = 1 \end{cases}$$



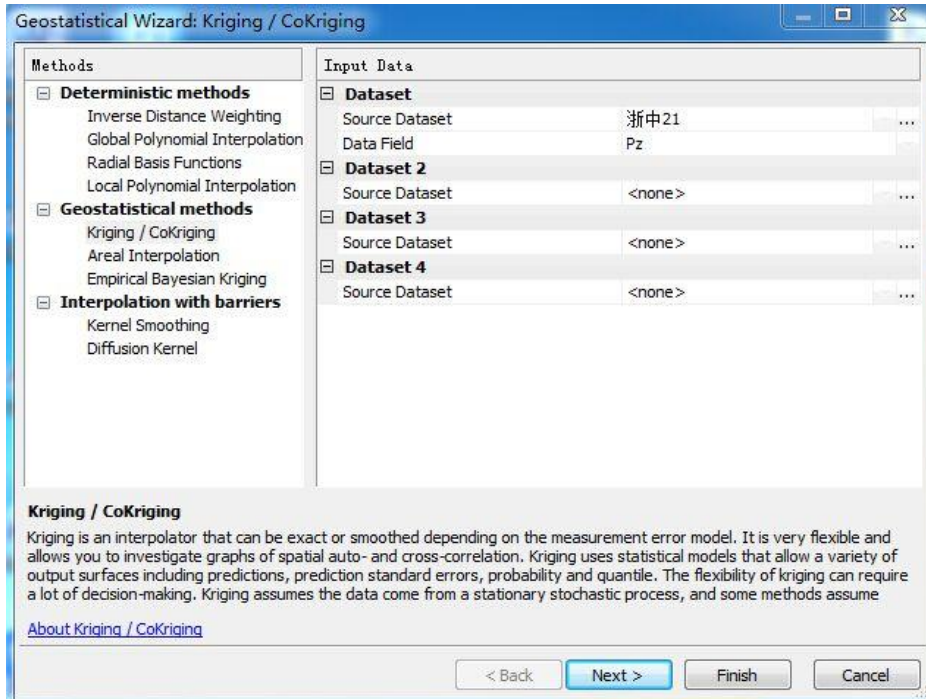
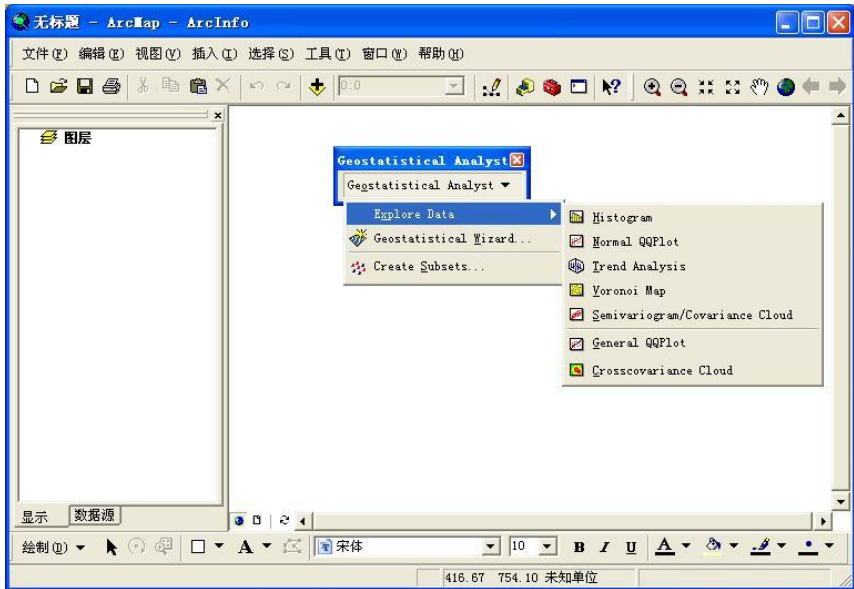
克里格方法流程图

§ 3-3 资源空间信息的建模

三、空间插值分析

5、软件实现

上机实验！

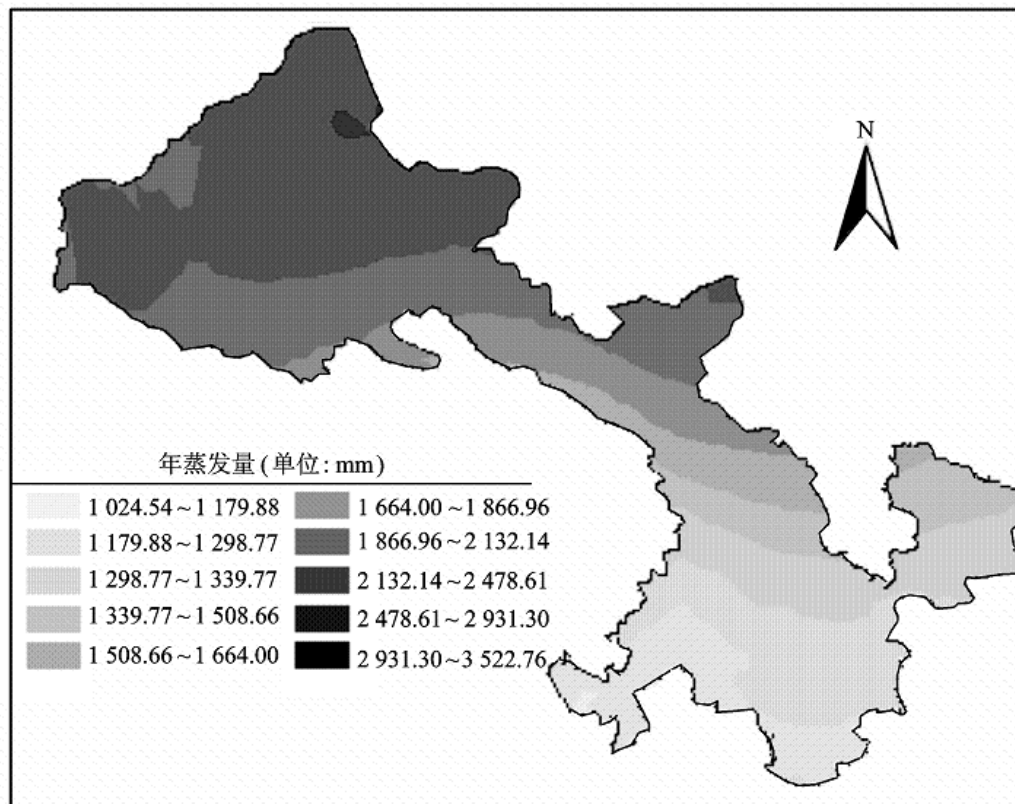


§ 3-3 资源空间信息的建模

三、空间插值分析

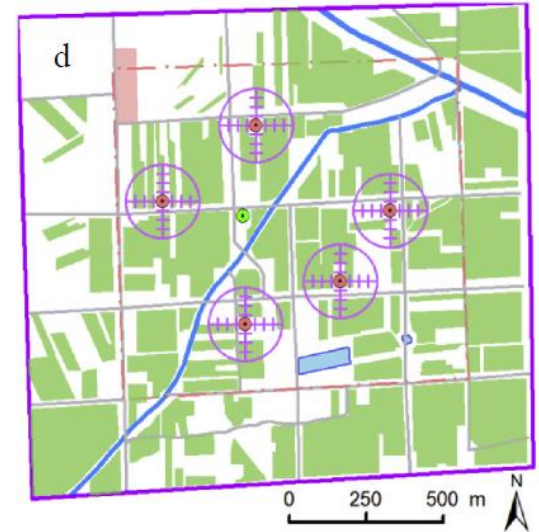
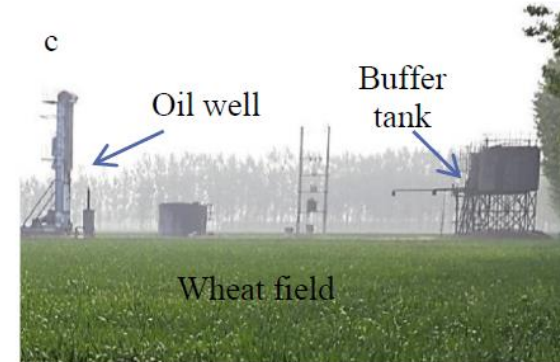
上机实验！

5、软件实现



§ 3-4 资源空间信息案例

一、基本概况



- | | | | |
|---------------------|-------------------|-------------|----------------|
| Core operating area | Injection station | Village | Road |
| Study area | Production well | Water Body | River |
| Boundary of Gaoqing | City or Town | Wheat Field | Sampling point |

§ 3-4 资源空间信息案例

一、基本概况

合同分类：技术服务

正本

2013230

技术服务合同

项目名称：国家科技支撑计划课题五“大规模燃煤电厂烟
气 CO₂ 捕集纯化与封存工程示范及评价”子课题
胜利油田 CO₂ 驱油示范区植被生态监测研究

委 托 方：中石化石油工程设计有限公司

受 托 方：中国矿业大学

签订时间：2013 年 9 月 29 日

签订地点：山东省东营市

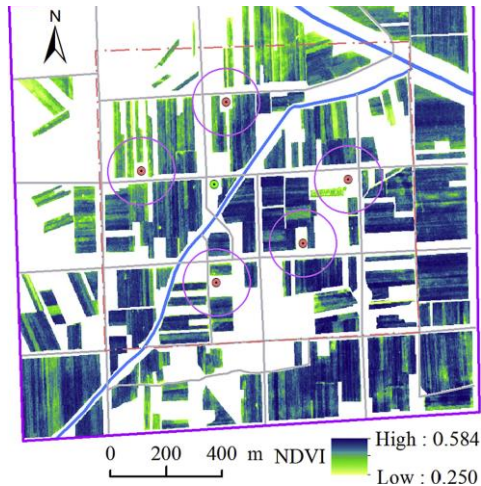
有效期限：2013 年 6 月至 2015 年 6 月

<https://lcei.cumt.edu.cn/info/1027/1113.htm>

§ 3-4 资源空间信息案例

二、NDVI及空间变异

$$NDVI = (IR - R) / (IR + R)$$



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

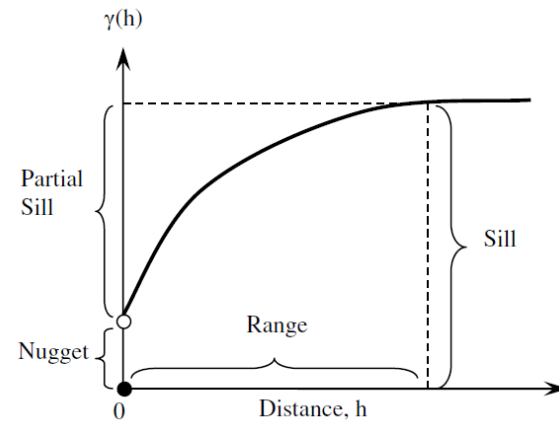


Table 1
Comparisons of the NDVI and semivariogram parameters.

Group	Wheat fields affected by CO ₂ emission	Reference wheat fields
Amount of wheat fields	15	18
Average NDVI	0.524 ± 0.013 ^a	0.529 ± 0.008 ^a
Standard deviation of NDVI	0.018 ± 0.004 ^a	0.018 ± 0.004 ^a
variation coefficients of NDVI	0.034 ± 0.009 ^a	0.034 ± 0.008 ^a
Range	33.362 ± 8.402 ^a	32.136 ± 6.534 ^a
Nugget	1.93E-05 ± 2.03E-05 ^a	2.06E-05 ± 2.60E-05 ^a
Partial Sill	3.00E-04 ± 1.00E-04 ^a	2.00E-04 ± 1.00E-04 ^a

Note: Different lower-case letters indicate significant differences between the two groups.

§ 3-4 资源空间信息案例

三、NDVI变异原因

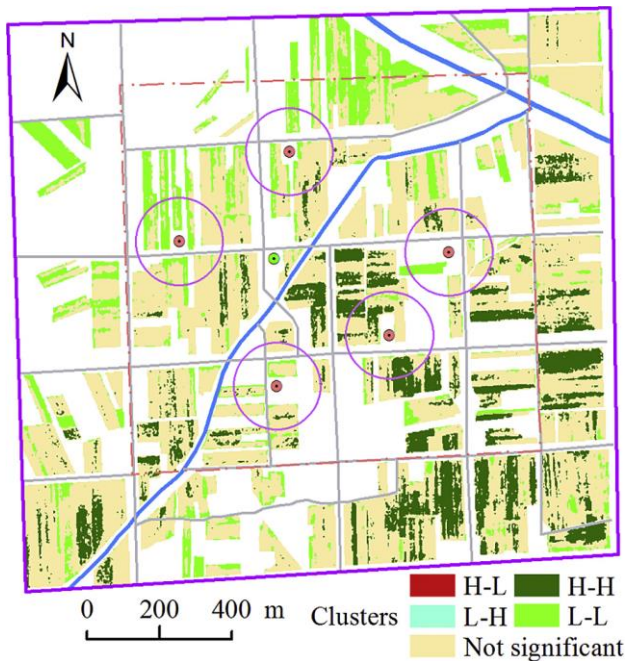


Table 2

Correlation and multiple GLM analysis of the relationships between NDVI and environmental factors.

Variable	Method			
	Correlation analysis		Multiple GLM regression	
	<i>r</i>	<i>p</i>	SS, %	<i>p</i>
Azimuth	0.194	0.156	5.680*	0.020
Distance	0.336*	0.012	0.459	0.827
Elevation	-0.437*	0.001	0.263	0.682
Average CO ₂	-0.283*	0.036	0.435	0.598
Soil moisture	0.790*	0.000	22.604*	0.000
Soil density	-0.107	0.437	0.164	0.746
Soil organic matter	0.656*	0.000	33.951*	0.000
Available N	0.626*	0.000	20.916*	0.001
Available P	-0.061	0.660	1.921	0.271
Available K	0.376*	0.005	10.757*	0.012
pH	0.246	0.070	1.313	0.361
Residuals	/	/	1.539	/

Note: * indicates $p < 0.05$; SS, proportion of variances explained by the variable.

§ 3-4 资源空间信息案例

四、风速与CO₂扩散

$$C(x, y, z, H) = \frac{q \times 1000}{2\pi\bar{u}\sigma_y\sigma_z} \exp\left(\frac{-y^2}{2\sigma_y^2}\right) \left\{ \exp\left[\frac{-(z-H)^2}{2\sigma_z^2}\right] + \exp\left[\frac{-(z+H)^2}{2\sigma_z^2}\right] \right\}$$

$$F(C) = F(u, q) = \int_0^u \int_0^q f(u, q) du dq$$

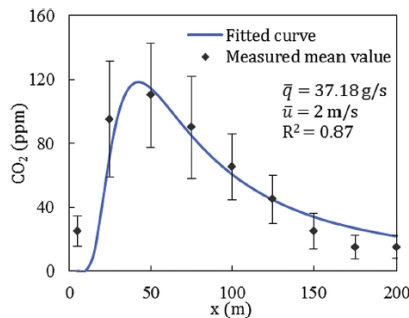


Fig. 3. Concentrations of emitted CO₂ in the downwind direction. The y axis is the mean of groups data of CO₂ measured from production wells, while the x axis represents the distance to the buffer tank.

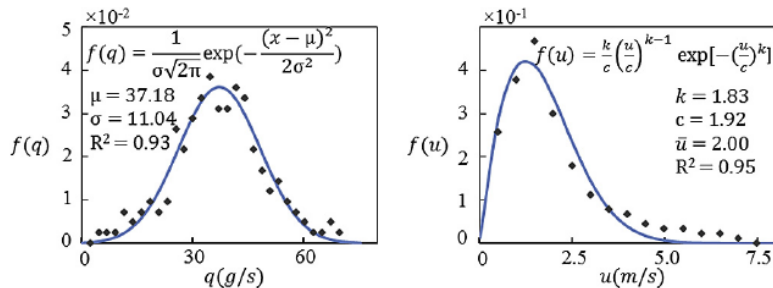


Fig. 6. The probability density of CO₂ emission intensity (q) and wind speed (u).

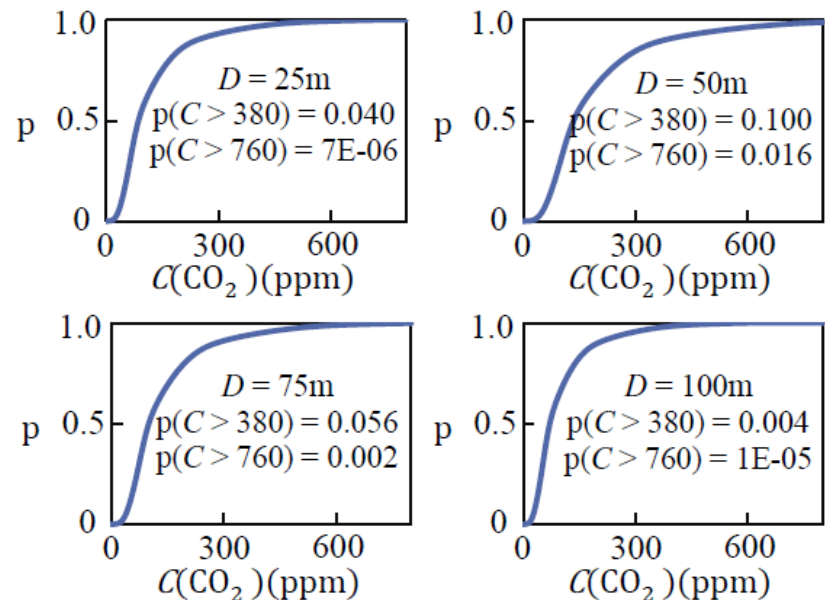


Fig. 7. The probability of the emitted CO₂ concentration at different site along the downwind direction, considering the background average CO₂ concentration in near-surface atmosphere is 380 ppm. D indicates the distance to the buffer tank.

§ 3-4 资源空间信息案例

五、CO₂逸散

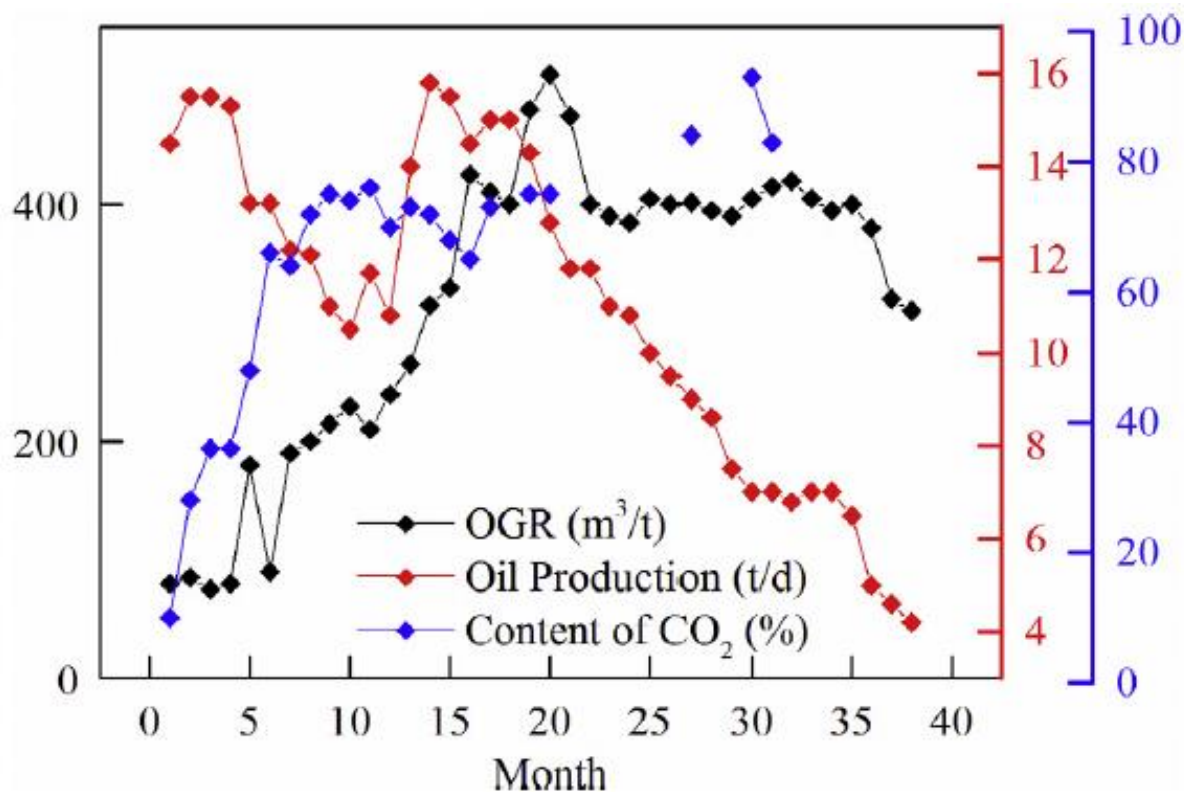


Fig. 8. Oil production parameters (oil-gas rate (OGR), oil production and CO₂ content) after CO₂ injection, summarized from Peng (Peng, 2013).

§ 3-4 资源空间信息案例



Monitoring the impact of fugitive CO₂ emissions on wheat growth in CCS-EOR areas using satellite and field data

Yongjun Yang^a, Yue Li^b, Shaoliang Zhang^{b, **}, Fu Chen^{a, c, *}, Huping Hou^b, Jing Ma^a

^a Low Carbon Energy Institute, China University of Mining and Technology, Xuzhou 221008, China

^b School of Environment Science and Spatial Informatics, China University of Mining and Technology, Xuzhou 221008, China

^c College of Science, Engineering and Health, Royal Melbourne Institute of Technology University, Melbourne, VIC 3001, Australia

ARTICLE INFO

Article history:

Received 15 August 2016

Received in revised form

8 March 2017

Accepted 8 March 2017

Available online 9 March 2017

Keywords:

CCS-EOR

Ecological monitoring

Fugitive CO₂ emissions

NDVI

Wheat

Geostatistical analysis

ABSTRACT

In-situ monitoring of the environmental impact is essential for the verification of clean operation of carbon capture and storage (CCS) projects. We conducted an empirical study based on remotely sensed data and field observations from an enhanced oil recovery (EOR) site in China. Geostatistical analysis and general linear model regression were performed to detect the impact of fugitive CO₂ emission from oil buffer tanks. Results show that the emitted CO₂ resulted in CO₂ enrichment about 25–100 m away from the buffer tanks. The spatial pattern and semivariogram parameters of normalized differential vegetation index (NDVI) in the CCS core operating area had not been altered significantly. The CO₂ concentration is not a statistically significant explanatory factor for the variation of wheat growth in the present CCS-EOR site that located in Gaoqing County in the east of China. The impact of fugitive CO₂ emission on wheat growth appears limited because of the instability and rapid diffusion of emitted CO₂. However, we emphasize that these results were extracted from the in-situ monitoring that characterized by macro-level and short term. Long-term biology-based study at a micro-level is imperative for further understanding and determining the environmental impact of fugitive CO₂ emission. Moreover, considering the incidental CO₂ breakthrough and the unknown impact of emitted gas, including CO₂ and hydrocarbons, on the quality of wheat plant and grain, long-time field monitoring, and improvement of production equipment and technique are essential to ensure the clean production of CCS-EOR.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

With rising levels of atmospheric carbon dioxide (CO₂), carbon capture and storage (CCS) has been recognized as an important means for reducing greenhouse gas emissions and mitigating global warming (Huisingsh et al., 2015; Zhu et al., 2015). The technologies of CO₂ storage with enhanced industrial production, such as enhance oil recovery (CO₂-EOR), natural gas recovery (CO₂-EGR), coal-bed methane (CO₂-ECBM), saline water/brine recovery (CO₂-EWR) and geothermal energy (CO₂-EGS) have recently attracted interest worldwide since they have great potential to reduce the overall cost and increase the economic output (Li et al., 2015; Wang and Du, 2016; Zhang and Huisingsh, 2017). Among of them, CO₂-EOR

has been globally expanded to over a hundred oil sites because this technology can produce more crude oil while storing the CO₂ below ground to provide environmental benefits (Cámara et al., 2013; Jenkins et al., 2015). However, the potential environmental impact is a challenge for achieving wide public acceptance and clean operation of CCS (Cámara et al., 2013; Li et al., 2013; Wennersten et al., 2015). To assure safety and environmental friendliness of CCS technology, more detailed environmental monitoring and quantitative verification are necessary (Bowden et al., 2013; Jones et al., 2015; Li et al., 2016a,b).

Since 2005, there have been attempts to reveal the potential impact of CO₂ leakage on plant growth in the research field of CCS (Jones et al., 2015). According to existing field studies, no real CO₂ leakage and consequent impacts have been detected (Jones et al., 2015). Therefore, research scholars are mainly relying on the simulation experiments and analogous sites such as ZERT, ASCARD, CO₂ Field Lab, PISCO₂, Ginninderra, Ressacada Farm and natural CO₂ seepage sites in Italy and Germany (Govindan et al., 2011; Noble et al., 2012). The results of these studies suggest that CO₂ emitted from below ground will impose stress on plant growth, thereby

* Corresponding author. Low Carbon Energy Institute, China University of Mining and Technology, Xuzhou 221008, China.

** Corresponding author.

E-mail addresses: shzhang@cumt.edu.cn (S. Zhang), chenfu@cumt.edu.cn (F. Chen).

http://dx.doi.org/10.1016/j.jclepro.2017.03.058
0959-6526/© 2017 Elsevier Ltd. All rights reserved.



国科奖社证第0265号

绿色矿山科学技术奖

获奖证书

获奖项目：CO₂驱替开采生态环境影响监测与评估技术研究

奖励类型：基础研究

获奖等级：一等

获奖者：杨永均

奖励年度：2020年



证书号：LKJ-P-2020-1-005-1-004