**西南大学**

**计算机与信息科学学院**

学年设计报告

**题 目：**  空间插值算法之普通Kriging

**年级、专业：** 2014 **级**  网络工程 **专业**  1 **班**

**学生姓名：**  吴梦露

**提交日期：**  2017 **年** 6 **月** 20 **日**

空间插值算法之普通Kriging

摘要

所谓空间数据插值，即通过探寻收集到的样点/样方数据的规律，外推/内插到整个研究区域为面数据的方法。空间插值方法分为两类：一类是确定性方法，另一类是地质统计学方法。确定性插值方法是基于信息点之间的相似程度或者整个曲面的光滑性来创建一个拟合曲面，地质统计学插值方法是利用样本点的统计规律，使样本点之间的空间自相关性定量化，从而在待预测的点周围构建样本点的空间结构模型。本文通过对比分析各类插值方法的原理及优缺点，利用ArcGIS10软件特定模块进行建模，并结合相关论文，探讨空间插值法在地理信息系统方面的应用。

**关键词：**空间插值；地质统计学；ArcGIS10软件；地理信息系统

Abstract

The so-called spatial data interpolation, which is the method of extrapolating/interpolating the entire area of the study by exploring the collected points/sample data. Space interpolation method is divided into two categories: one is deterministic method and the other is geological statistic method. The former is based on the similar degree of information points or smoothness of the entire surface to create a fitting surface; while geostatistics interpolation is to use the statistical law of sample points, quantifying the spatial autocorrelation among the sample points, to construct a spatial structure model of the sample points around the point to be predicted. There has a discuss in this article about analyzing the advantages and disadvantages of various interpolation methods, using ArcGIS10 software to model the specific modules, and exploring the application of spatial interpolation in Geographic Information System in terms of related papers.

**Key Words:** spatial interpolation; geostatistics; ArcGIS10; Geographic Information System

# 1 绪论

## 1.1 研究背景

随着GIS和计算机技术的不断发展及人们在研究工作中对空间高质量数据的要求，空间数据插值应用越来越广，受到人们的高度重视。空间插值既是数据维护方法，也是空间分析方法。ArcGIS软件作为一套完整的GIS平台产品，具有强大的地图制作、空间数据管理、空间分析、空间信息整合、发布与共享的能力。其地统计分析模块提供了一系列已知样点进行内插，生成研究对象表面图的内插技术。

## 1.2 主要工作和安排

本设计根据题目要求，需要概括、总结各类空间插值方法的原理、优缺点及应用场景，并选取一到两种插值方法具体分析，利用ArcGIS10软件建立模型，并阅读代码以及相关论文更进一步理解算法思想。时间计划安排如下：

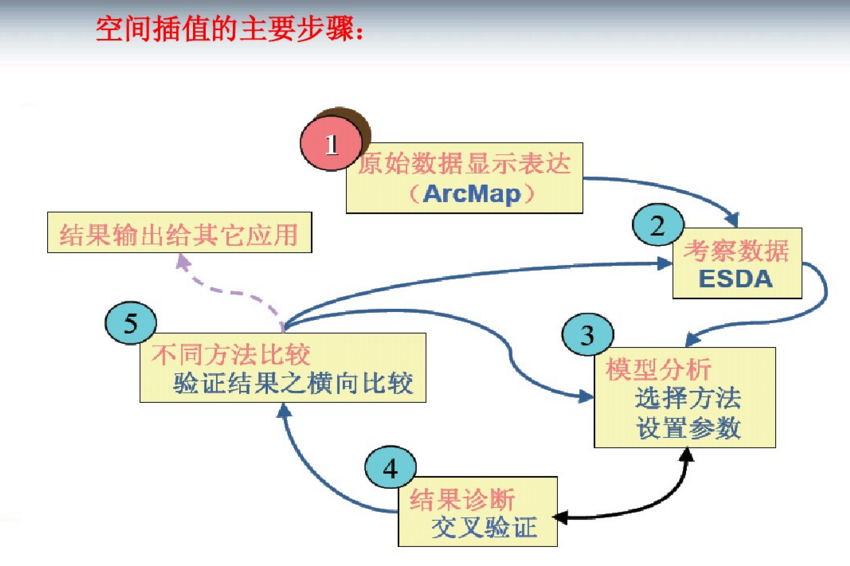
|  |  |
| --- | --- |
| 时间 | 任务 |
| 3.22-4.1 | 对空间插值算法有个大概的理解；并汇总比较各类空间插值算法 |
| 4.1-5.1 | 多方面熟悉Kriging基础知识；分析推导普通Kriging公式；结合实例加深理解；结合学习视频使用ArcGIS10软件建模 |
| 5.1-6.1 | 阅读相关论文 |
| 6.1-6.15 | 尝试去读IDW和Kriging的代码；整理设计报告 |

# 2 空间数据插值方法

## 2.1 基本原理

任何一种空间数据插值法都是基于空间相关性的基础上进行的。即空间位置上越靠近，则事物或现象就越相似；空间位置越远，则越相异或者越不相关。体现了事物/现象对空间位置的依赖关系。不同的插值方法对于如何获得最佳估计值都给予了一定的假设，但无论你选哪种插值方法，样本点越多，样本点分布越均匀，插值结果越接近实际值。

## 2.2 一般插值过程



①空间插值数据源获取；

②对数据进行分析，找出源数据的分布特性、统计特性，以利于选择最恰当的插值方法；

③插值方法的选择并进行插值计算；

④对插值结果的评价；

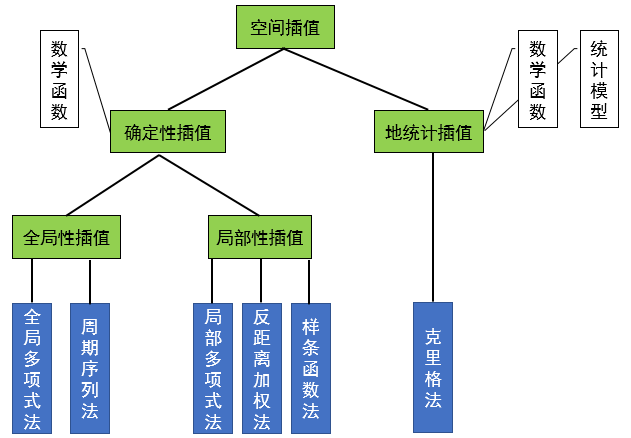
⑤运用多种插值方法进行计算，对各种方法的插值结果进行比较、分析并选择最佳的插值方法。

## 2.3 算法分类

1. 根据数学原理：

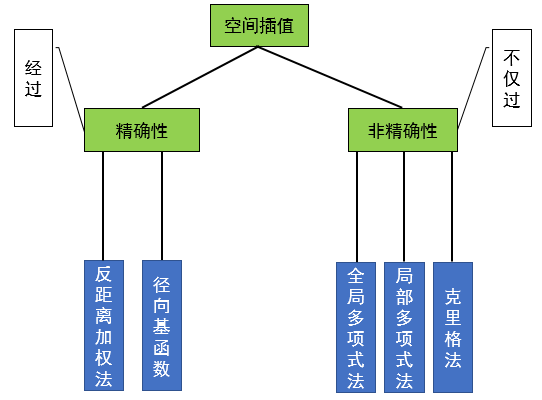
确定性插值法是使用数学函数进行插值，以研究区域内部的相似性，或者以平滑度为基础，由已知样点来创建预测表面的插值方法。地统计插值依赖于数学模型和统计模型，基于自相关性（测量点的统计关系），根据测量数据的统计特征产生曲面。地统计与确定性插值的最大区别在于，地统计插值引入了概率模型，即地统计插值认为从一个统计模型不可能完全精确地得出预测值，所以在进行预测时，应该给出预测值的误差，即预测值在一定概率内合理。

* 全局性插值（整体插值）：用研究区所有采样点数据进行全区特征拟合。整个区域的数据都会影响单个插值点，单个数据点变量值的增加、减少或者删除，都会对整个区域有影响。
* 局部性插值：只使用邻近的数据点来估计未知点的值。单个数据点的改变只影响其周围有限的数据点。



（2）根据是否能保证创建的表面经过搜有的采样点：

* 精确性插值：产生通过所有观测点的曲面。插值点落在观测点上，内插值等于估计值。
* 非精确（近似）插值：插值产生的曲面不通过所有观测点。当数据存在不确定性时，应该使用近似性插值，由于估计值替代了已知变量值，近似差值可以平滑采样误差。



## 2.4 综合比较

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **常用插值方法** | **基本思想** | **优点** | **缺点** | **应用** |
| **反距离加权插值法（IDW）** | 某未知采样点的数据值与其周围一定范围内的已知采样点的数据值有关，是这些临近已知点的数据值综合贡献的结果，其贡献程度与距离成反比。 | 算法简便易行，是一种精确插值算法。 | 对权重函数中幂次选择十分敏感，易受样本集群影响，造成“牛眼”效应。 | 样本点的分布均匀、密集且样本点含有区域特征信息。 |
| **克里格插值法（Kriging）** | 该方法认为在空间连续变化的属性是非常不规则的，用简单的平滑函数进行模拟将出现误差，用随机表面函数给予描述会比较恰当。关键在于权重系数的确定，该方法在插值过程中根据某种优化准则函数来动态地决定变量的数值，从而使内插函数处于最佳状态。  该算法不仅考虑待估点位置与已知数据位置的相互关系，而且还考虑变量的空间相关性。 | 给出统计误差，充分考虑空间变量相关性，有效弥补数据集存在的聚类影响，插值精度高。 | 作为一种统计学方法，其计算步骤繁琐且插值速度慢。 | 适用于采样区域的样本数据存在随机性和结构性两特征。 |
| **径向基函数插值法（RBF）** | 即基函数是由单个变量的函数构成的，是一系列精确插值法的统称。所有径向基函数插值法都是准确的插值器，它们都能尽量适应的数据；若要生成一个更圆滑的曲面，对所有这些方法都可以引入一个圆滑系数。 | 引入平滑因子，插值结果表面比较光滑。 | 当研究区域内采样点局部变异性大，采样数据具有较大的不确定时。难以保证数据的准确性。 | 采样数据局部变异不大时，像进行气象特征估计时可采用此法。 |
| **趋势面分析插值法** | 是一种多项式回归分析技术，即用多项式表示线或面，按最小二乘法原理对数据点进行拟合。通常用于分析趋势和异常，而不追求高的拟合精度，是一个非精确的插值方法。 | 它是一种极易理解的技术，至少在计算方法上易于理解。另外，大多数数据特征可以用低次多项式来模拟。 | 当研究区域范围较大，地形很复杂时，需要用高阶多项式拟合以提高精度，但高阶将增加其计算成本，因而需要进行改进。 | 感兴趣区域的表面在各位置间出现渐变时，可以将该表面与采样点拟合，例如，工业区的污染情况；检查或排除长期趋势或全局趋势的影响。 |
| **最近邻点插值法** | 又叫泰森多边形(Thiessen多边形)方法。在每个样点数据周边生成一个邻近区域，即Thiessen多边形，使得每个多边形内的任意一点离其内部的样点最近，在多边形内插值时只有其中心样点参与运算。 | 方法简单，计算率高，不受前提条件的约束。 | 插值结果受样本点的影响大，缺乏对其他空间因素和变量自身所具有的固有规律的考虑。 | 样本点分布均匀、完整，空间变异性不很明显，只存在少量数据的缺失。 |

## 2.5 插值方法选择的原则

* + 精确性。
  + 参数的敏感性：许多的插值方法都涉及到一个或多个参数，如距离反比法中距离的阶数等。有些方法对参数的选择相当敏感，而有些方法对变量值敏感。后者对不同的数据集会有截然不同的插值结果。希望找到对参数的波动相对稳定，其值不过多地依赖变量值的插值方法。
  + 耗时：一般情况下，计算时间不是很重要，除非特别费时。
  + 存储要求：同耗时一样，存储要求不是决定性的。特别是在计算机的主频日益提高，内存和硬盘越来越大的情况下，二者都不需特别看重。
  + 可视化、可操作性（插值软件选择）：三维的透视图等。

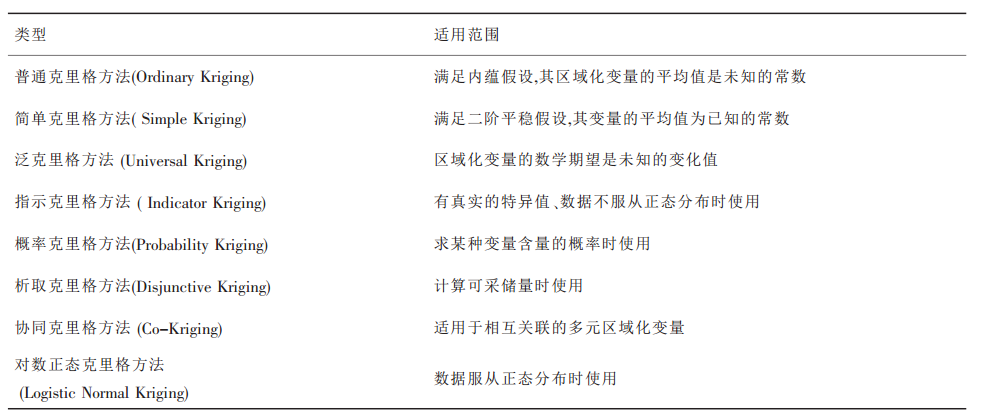
# 3 Kriging插值公式

克里金方法（kriging），是地质统计学的重要组成部分，也是地址统计学的核心。地统计方法不仅具有产生预测表面的功能，而且能够对预测的确定性或准确性提供某种度量。克里金方法最早是由法国地理学家Matheron和南非矿山工程师Krige提出的，用于矿山勘探。这种方法认为在空间连续变化的属性是非常不规则的，用简单的平滑函数进行模拟将出现误差，用随机表面函数给予描述会比较恰当。

克里金法工具可将数学函数与指定数量的点或指定半径内的所有点进行拟合以确定每个位置的输出值。克里金法是一个多步过程，它包括数据的探索性统计分析、变异函数建模和创建表面，还包括研究方差表面。克里金方法的关键在于权重系数的确定，该方法在插值过程中根据某种优化准则函数来动态地决定变量的数值，从而使内插函数处于最佳状态。克里金方法考虑了观测的点和被估计点的位置关系，并且也考虑各观测点之间的相对位置关系，在点稀少时插值效果比反距离权重等方法要好。所以利用克里金方法进行空间数据插值往往取得理想的效果。

在地质统计学中，根据应用目标的区别，发展了多种克里格方法如：

* 简单克里金(Simple-Kriging)、
* 普通克里金(Ordinary-Kriging)、
* 泛克里金(Universal-Kriging)、
* 对数正态克里金(Log-Normal Kriging)、
* 协同克里金(Cokriging)、
* 拟协克里金(Pseudo-Kriging)、
* 指示克里金(Indicator-Kriging)、
* 离析克里金(Disjunctive-Kriging）



在三维地质建模过程中，克里格被作为插值方法，能过最大的程度的保证地质界面与原始数据的吻合，且不依赖于网络。

## 3.1 假设条件

不同的kriging插值方法的主要差异就是假设条件不同，普通kriging方法是最普通和应用最广的kriging方法，这里仅介绍普通kriging插值的假设条件与应用。

普通kriging插值的假设条件为：空间属性Z是均一的。对于空间任意一点，都有同样的期望c与方差。即对任意点都有：

换一种说法：任意一点处的值，都由区域平均值c和该点的随机偏差组成，即：

其中，表示点处的偏差，其方差均为常数：

## 3.2 区域化变量

### 3.2.1 随机变量

为一个实值变量，可根据概率分布取不同的值。每次取值（观测）结果z为一个确定的数值，称为随机变量Z的一个实现。

* 连续型地质变量：构造深度、砂体厚度、有效厚度、孔隙度、渗透度、含油饱和度；
* 离散型地质变量：砂体、相、流动单元、夹隔层、断层；

### 3.2.1 随机变量的特征值

（1）数学期望：是随机变量的整体代表性特征数。相当于随机变量以其取值概率为权的加权平均数。从矩阵的角度说，数学期望是的一阶原点矩。

（2）方差：是随机变量的离散型特征数。若数学期望存在，则称它为的方差，记为，或，或。从矩阵的角度说，方差是的二阶中心矩。

方差的平方根为标准差，记为

### 3.2.3 随机函数

研究范围内的一组随机变量。

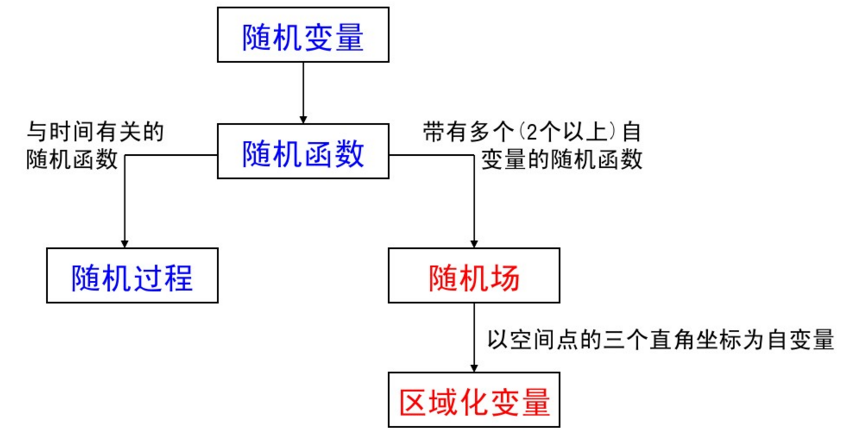
### 3.2.4 随机函数的特征值

协方差（Variance）：两个随机变量的协方差为二维随机变量的二阶混合中心距，记为 ，或。

### 3.2.5 区域化变量

地统计学是以区域化变量理论为基础，以变异函数为主要工具，研究那些在空间分布上既有随机性又有结构性，或空间相关和依赖性现象的学科。

当一个变量的取值与其空间位置有关时，就称为区域化变量。区域化变量常常反映某种空间现象的特征，用它来描述的现象称之为区域化现象。区域化变量，亦称区域化随机变量，它的取值（Z）与空间位置（x）有关，故将其即为Z(x)。区域化变量Z(x)，具有两个显著的特征：随机性和结构性特征。许多地理变量是区域化变量，有矿产的品位（铜矿、铁矿、煤矿、石油）、气温（中国气温由南向北逐渐降低，在同一区域里气温的高低有差异）、降水等。



区域化变量是一种随机函数，因而能同时反映空间变量的结构性和随机性。一方面，当空间点固定后，就是一个随机变量，这体现了其随机性。另一方面，在空间两个不同点与处的区域化变量值具有某种程度的相关性，这体现了其结构性。

## 3.3平稳要求

### 3.3.1 严格平稳

对于单变量而言：

可以从研究区内所有数据的累积直方图推断而得（将临近点当成重复取样点）

### 3.3.2 二阶平稳

当区域化变量满足下列两个条件时，则称其为二阶平稳或弱平稳：

①在整个研究区内有的数学期望存在，且等于常数，即：

随机函数在空间上的变化没有明显趋势，围绕值上下波动。

②在整个研究区内，的协方差函数存在且平稳（即只依赖于滞后，而与无关）：

特殊地，当时，上式变为，即方差存在且为常数。

协方差不依赖于空间绝对位置，而依赖于相对位置，即具有空间的平稳不变性。

### 3.3.3 本征假设

比二阶平稳更弱的平稳假设。

当区域化变量的增量满足下列两个条件时，称其为满足本征假设或内蕴假设。

①在整个研究区内有。（）

可出现不存在，但存在并为零的情况。、

②增量的方差函数（变差函数，）存在且平稳（即不依赖于），即：

相当于要求：的变差函数存在且平稳。

可出现协方差不存在时，但变差函数存在的情况。

准二阶平稳假设及准本征假设：若区域化变量在整个区域内不满足二阶平稳（或本征假设），但在有限大小的邻域内是二阶平稳（或本征）的，则称是准二阶平稳的（或准北征平稳的）。

## 3.3 变差函数

变差函数（或叫变程方差函数，或变异函数）是地质统计学所特有的基本工具。它既能描述区域化变量的空间结构性变化，又能描述其随机性变化。

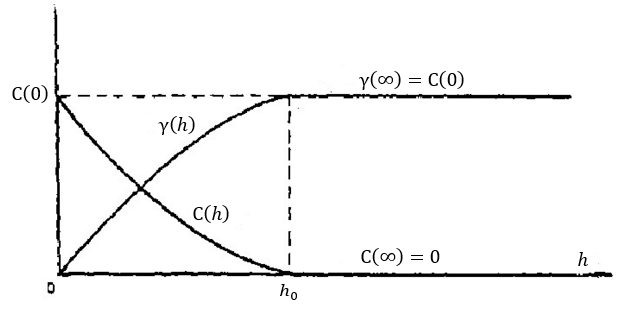
一维情况下的定义：

假设空间点只在一维的轴上变化，则将区域化变量在两点处的值之差的方差之半定义为在轴方向上的变差函数，记为。半方差函数（或半变异函数）。

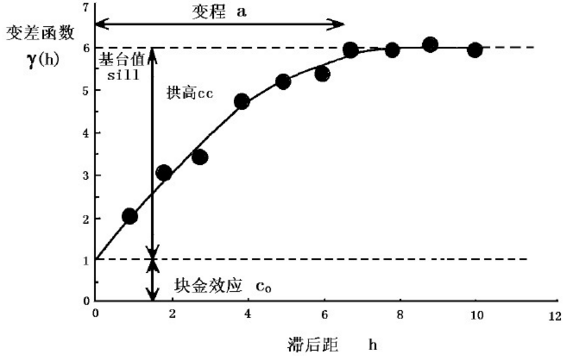
在二阶平稳假设，或作本征假设，此时：

所以：

地质统计学中最常用的基本公式之一



（二阶平稳假设条件下边查函数与写防查的关系）



（1）变程(Range)：指区域化变量在空间上具有相关性的范围。在变程范围之内，数据具有相关性；而在变程之外，数据之间互不相关，即在变程以外的观测值不对估计结果产生影响。

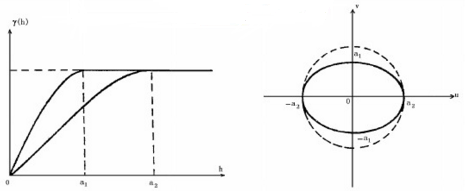
（2）块金值(Nugget)：变差函数如果在原点间断，在地质统计学中称为“块金效应”，表现为在很短的距离内有较大的空间变异性，无论多小，两个随机变量都不相关。它可以由测量误差引起，也可以来自矿化现象的微观变异性。在数学上，块金值相当于变量纯随机性的部分。

块金效应的尺度效应：如果品位完全是典型的随机变量，则不论观测尺度大小，所得到的实验变差函数曲线总是接近于纯块金效应模型。当采样网格过大时，将掩盖小尺度的结构，而将采样尺度内的变化均视为块金常数，这种现象即为块金效应的尺度效应。

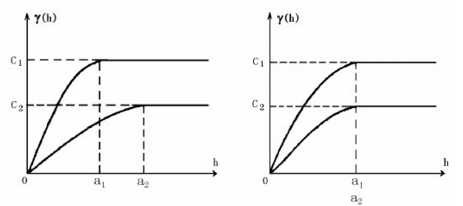
（3）基台值(Sill)：代表变量在空间上的总变异性大小。即为变差函数在大于变程时的值，为块金值和拱高值之和。拱高为在取得有效数据的尺度上，可观测得到的变异性幅度大小。当块金值等于0时，基台值即为拱高。

地址变量相关性的各向异性：

* 几何各向异性：变差函数在空间各个方向上的变程不同，但基台值不变（即变化程度相等）。这种情况能用一个简单的几何坐标变换将各向异性结构变换为各向同性结构。



* 带状各向异性：不同方向的变差函数具有不同的基台值，其中变程可以不同，也可以相同。这种情况不能通过坐标的线性变换转化为各向同性，因而结构套合是比较复杂的。



## 3.4 克里金法公式

由于克里金法可对周围的测量值进行加权以得出未测量位置的预测，因此它与反距离权重法类似。这两种插值器的常用公式均由数据的加权总和组成：

其中：

：测量值数

：点处的估计值，即

：第个位置处的测量值

：第个位置处的测量值的未知权重

这里是权重系数，它同样是用空间上所有已知点的数据加权求和来估计未知点的值。在反距离权重法中，权重仅取决于预测位置的距离。但是，使用kriging方法时，权重不仅取决于测量点之间的距离、预测位置，还取决于基于测量点的整体空间排列。要在权重中使用空间排列，必须量化空间自相关。因此，在普通克里金法中，权重取决于测量点、预测位置的距离和预测位置周围的测量值之间空间关系的拟合模型。

因此这里权重系数并非距离的倒数，而是能够满足点()处的估计值与真实值的差最小的一套最优系数。估计方差最小和无偏性被作为的选取标准，即：

同时满足无偏估计的条件：

Kriging插值法不仅考虑待估点位置已知数据位置的相互关系，而且还考虑变量的空间相关性。（应用随机函数理论）

## 3.5 无偏约束条件

先分析无偏估计条件，将带入则有：

又因为对任意的 Z都有，则：

可得到关系式：

这是的约束条件之一。

## 3.6 优化目标/代价函数J

分析估计误差。为方便公式推理，用符号表示，即：

则有：

## 3.7 代价函数的最优解

定义半方差函数，带入中，有

考虑到，

我们的目标是寻找使最小的一组，且是的函数，因此直接将对求偏导数令其为0即可，即：

但是要注意的是，我们要保证求解出来的最优满足公式，这是一个带约束条件的最优化问题。使用拉格朗日乘数法求解，求解方法为构造一个新的目标函数：

其中是拉格朗日乘数。求解使这个代价函数最小的参数集，则能满足其在约束下最小化。即：

由于，因此同样地，，所以：

在以上计算中我们得到了对于求解权重系数的方程组，写成线性方程组的形式就是：

写成矩阵形式即为：

对矩阵求逆即可求解，唯一未知的就是上文中定义的半方差函数。

## 3.8 半方差函数

上文中对半方差函数的定义为：

其等价形式为：

这也是半方差函数名称的由来，接下来证明这二者是等价的：

根据上文定义，有，则

又因为

于是有：

至此，便得证了，那么现在的问题是如何计算

这时需要用到地理学第一定律，空间上相似的属性相近。

表达了属性的相似度，空间的相似度就用距离来表达，定义与之间的集合距离：

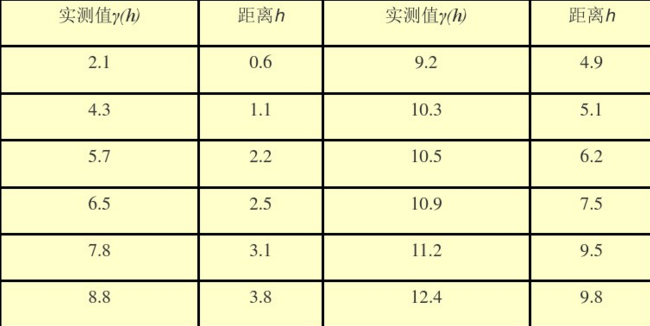
Kriging插值假设与存在着函数关系，这种函数关系可以是线性、二次函数、指数、对数关系。为了确认这种关系，我们需要首先对观测数据集计算任意两个点的距离和半方差，这时会得到个的数据对。

将所有的d和r绘制成散点图，寻找一个最优的拟合曲线拟合d与r的关系，的导函数关系式：

那么对于任意两点，先计算其距离，然后根据得到的函数关系就可以得到这两点的半方差。

# 4 Kriging具体步骤及优缺点

例1：某地区降水量是一个区域化变量，其变异函数的实测值及距离的关系见下表，下面我们试用回归分析方法建立其球状变异函数模型。



解：球状变异函数的一般形式为：

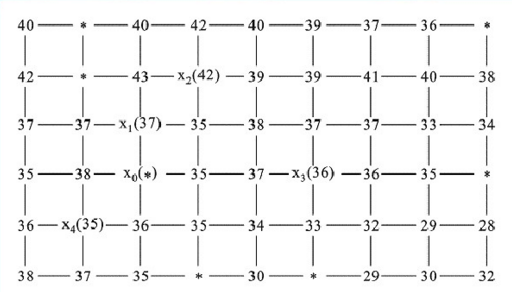
备注：这里的球状变异函数模型形式和上面的不太一样，这里的表示块金值，表示拱高值，基台值=块金值+拱高值。

当时，有

如果记，则可以得到线性模型：

根据表中的数据，对上式进行最小二乘拟合。它是一种数学上的近似和优化，利用已知的数据得出一条直线或者曲线，使之在坐标系上与已知数据之间的距离的平方和最小。得到：

比较上述两个等式，并做简单计算可知：，所以球状变异函数模型为：



4个观测点的观测值分别为，如果假设降水量的变异函数是向同性（即变异函数在各个方向的变化都相同）的二维球状模型，其具体形式为上式。现在，用普通克里格法父级观测点的降水量。

当时，。

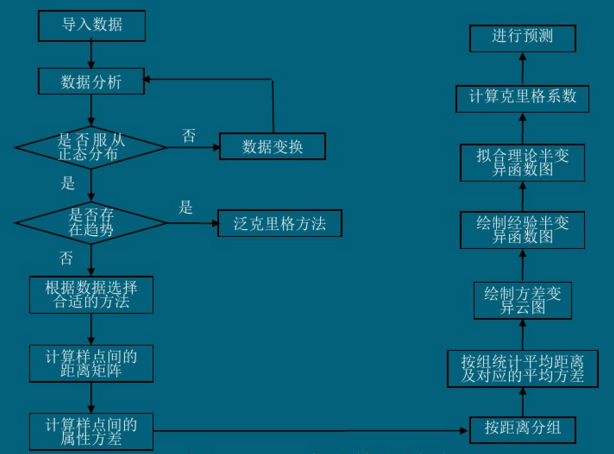
根据克里格矩阵的对称性，当时，

由此计算可得：

将以上计算结果代入克里格方程组中：

即克里格权重系数分别为，所以观测点的降水量的克里格估计值为：根据普通克里格法的基本原理，我们知道，估计的基本公式应该是：

## 4.1 克里格方法的具体步骤



克里格方法又称空间局部插值法，是以变异函数理论和结构分析为基础，在有限区域内对区域化变量进行无偏最优估计的一种方法，是地统计学长的主要内容之一。其实质是利用区域化变量的原始数据和变异函数的结构特点，对未知样点进行线性无偏、最优估计。无偏是值偏差的数学期望为0，最优是指估计值与实际值之差的平方和最小。

## 4.2 克里格方法的优缺点

### 4.2.1 优点

* 估计的无偏性
* 反映了变量的空间结构性
* 能得到估计精度

### 4.2.2 局限性

（1）克里格插值为局部估计方法，对估计值的整体空间相关性考虑不够，他保证了数据的估计局部最优，却不能保证数据的总体最优，因为克里格估计值的方差比原始数据的方差要小。因此，当井点较少且分布不均时可能会出现较大的估计误差，特别是井点之外的无井区误差可能更大。

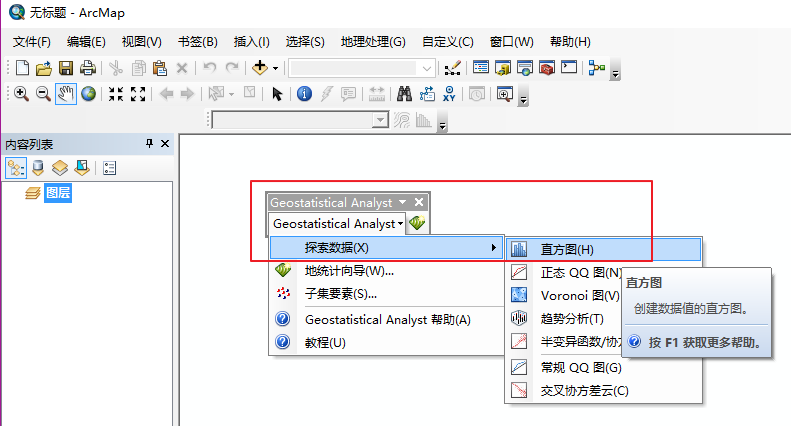
（2）克里格插值法为光滑内插方法，为减小估计方差而对真是观测数据的离散性进行了平滑处理，虽然可以得到由于光滑而更美观的等值线图或三维图，但一些有意义的异常带也可能被光滑作用而“光滑”掉了。所以有时候，克里格方法被称为一种“移动光滑窗口”。

# 5 ArcGIS10之地统计扩展模块

## 5.1 地统计模块简介

ArcGIS的地统计分析扩展模块的应用实现了地统计学与GIS学之间的结合，使得复杂的地统计方法可以在地理信息系统软件平台中实现。地统计扩展模块的三个主要功能模块：探索性数据分析、地统计分析向导、生成数据子集。使用这些基本功能模块，可以方便地完成多种地统计分析，创建完善的专题地图。

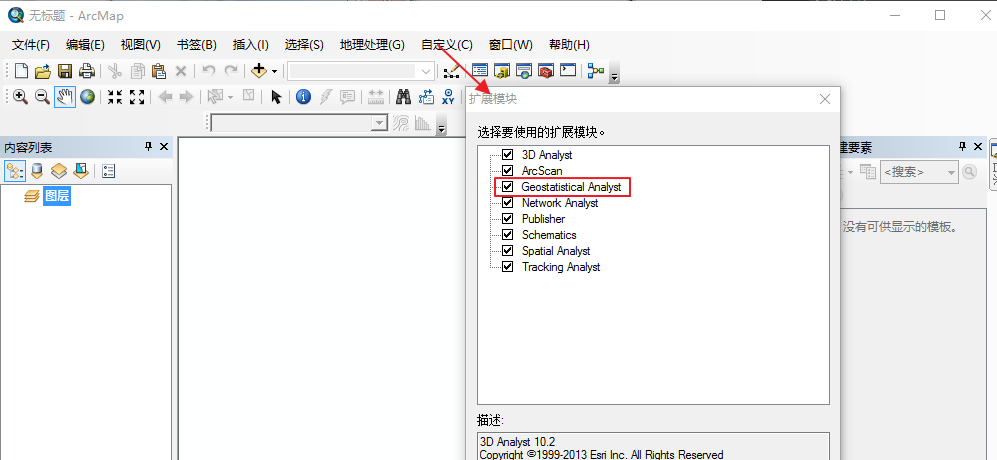
ArcGIS地统计扩展模块的菜单非常简单，但由此却可以完成完整的空间数据分析过程。一个完整的空间数据分析过程，或者说表面预测模型，一般为：拿到数据，首先要检查数据，发现正态分布、有没有趋势效应、各向异性等等（此功能主要由【探索数据】菜单及其下级菜单完成）行表面预测，这其中包括半变异模型的选择和预测模型的选择；最后检验模型是否合理或几种模型主要由【地统计向导】菜单完成）。【子集要素】菜单的作用是为把采样点数据分成两本，一部分作为检验样本。



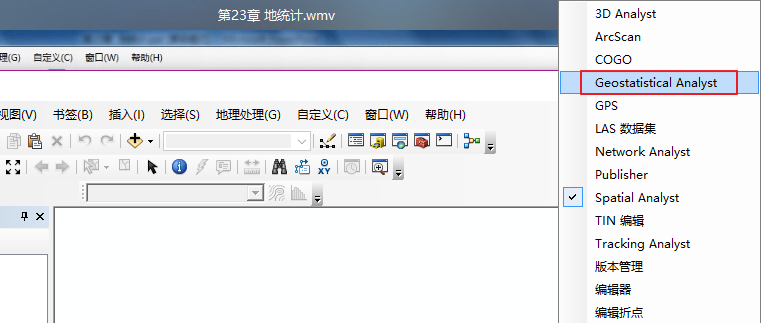
## 5.1 探索数据

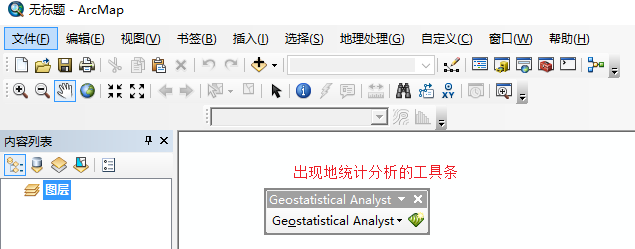
数据探索模块可以让用户全面了解所使用的数据，以更好确定合适的参数及方法，如数据是否服从正态分布、是否存在某种趋势等。数据检查，即空间数据探索分析（ESDA）。

自定义——》扩展模块——》选中Geostatistical Analyst：



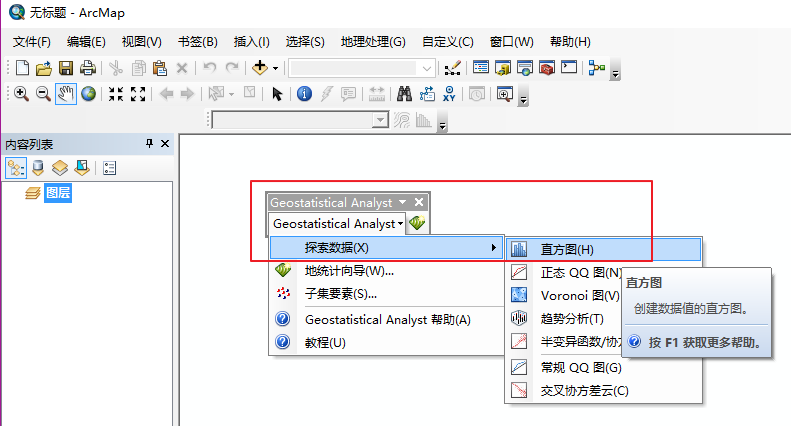
在主界面菜单中右击，弹出快捷菜单，选择Geostatistical Analyst，则将出现其工具条：



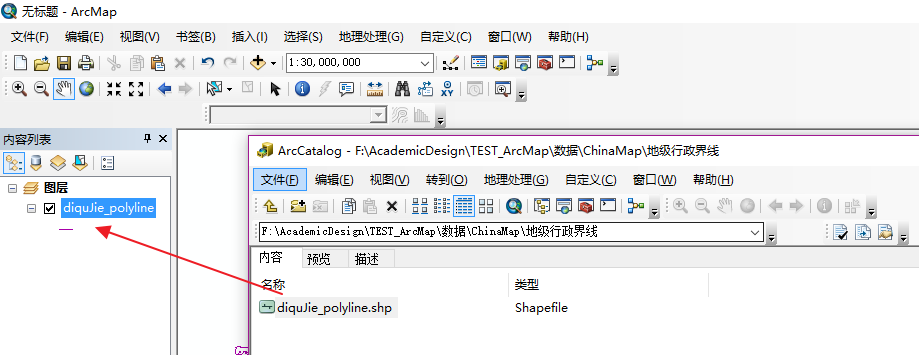


### 5.1 直方图

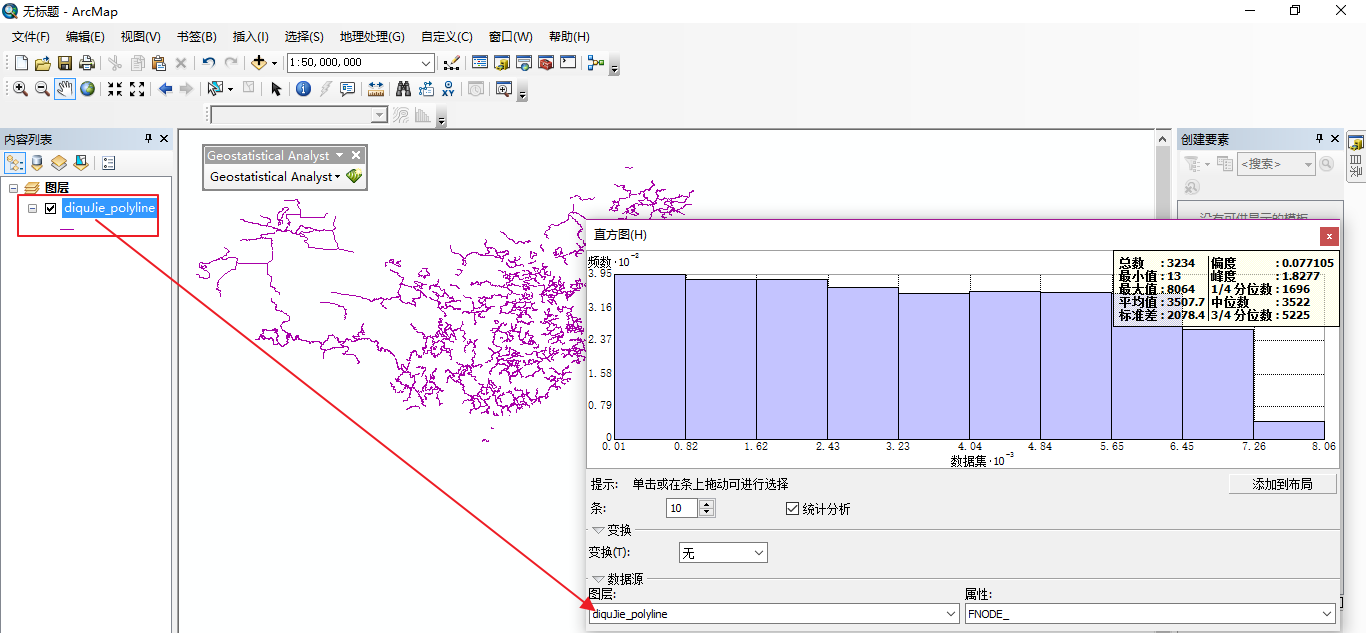
克里格方法对正态数据的预测精度最高，而且有些空间分析方法特别要求数据为正态分布。

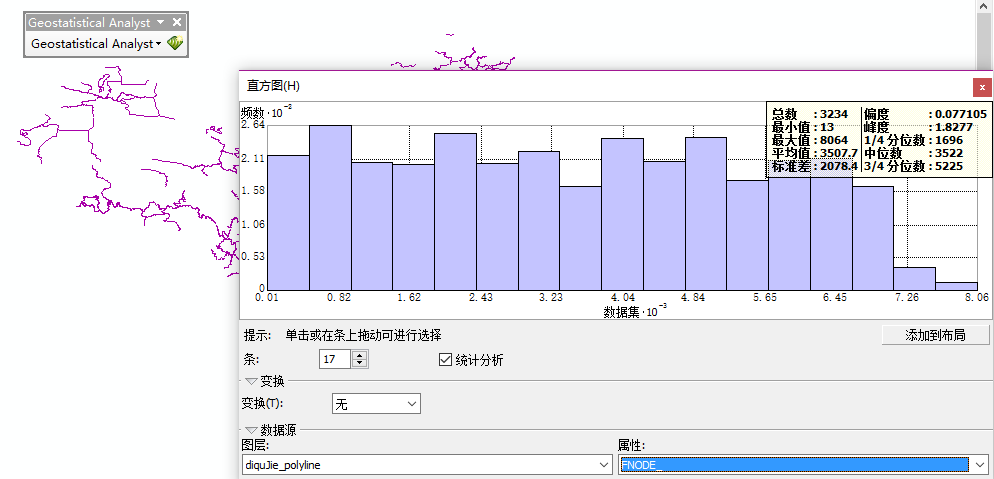


在ArcCatalog上打开一个shapefile，并复制到ArcMAP中作为直方图统计的原图层：



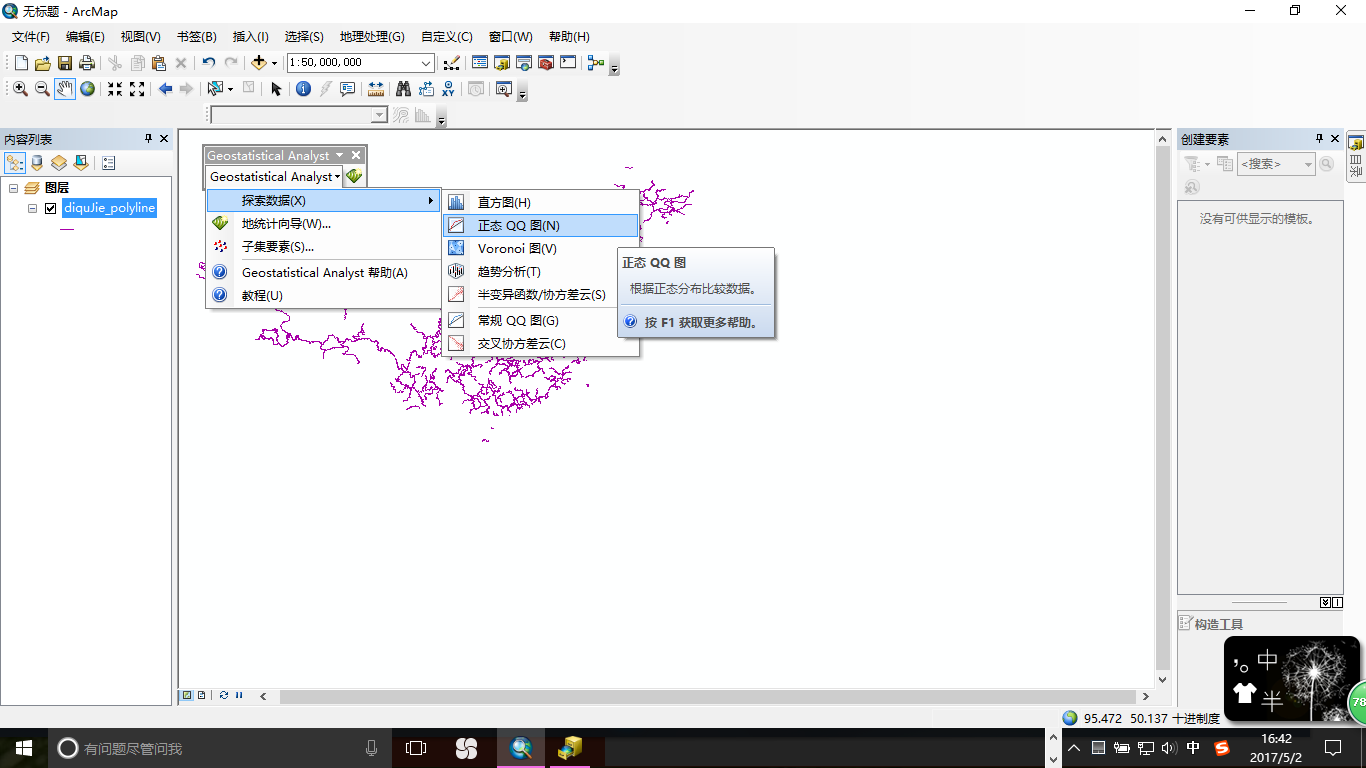
直方图用来显示数据的概率分布特征以及概括性的统计指标。若预测直方图基本呈现正态分布，则说明通过普通克里金插值得到预测值误差较小，能够较真实地反应测量数据，较完善弥补地质工作人员采集数据不足的缺点，从而能够较高地描述靶区的地质特征。

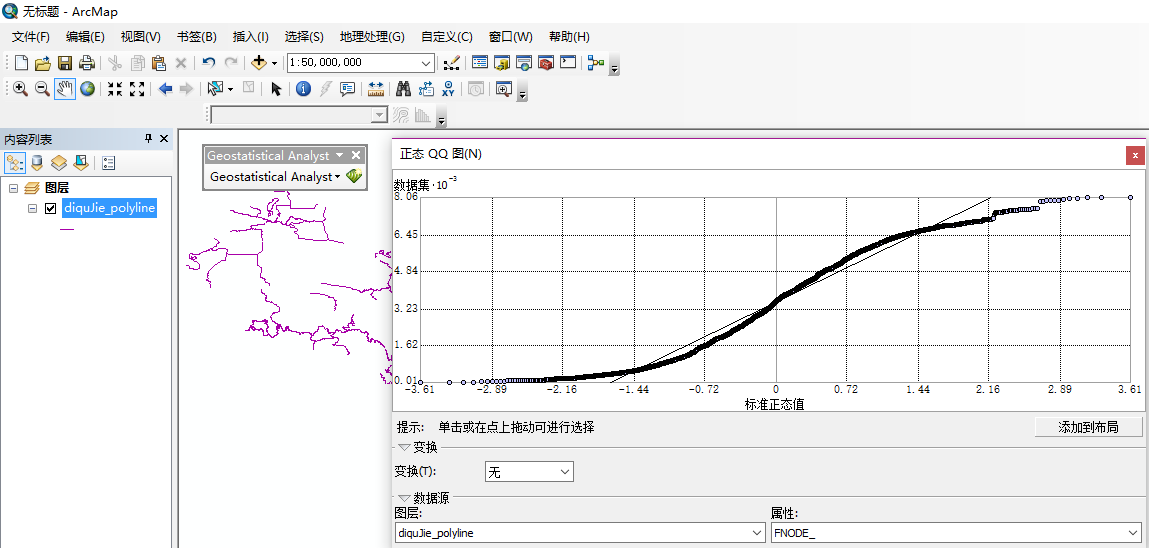




### 5.2 正态QQ图

正态QQ图可以检查数据的正态分布情况。作图原理是用分位图思想。直线表示正态分布，从图中可以看出数据很接近正态分布（左上角几个偏离的点被选中）。

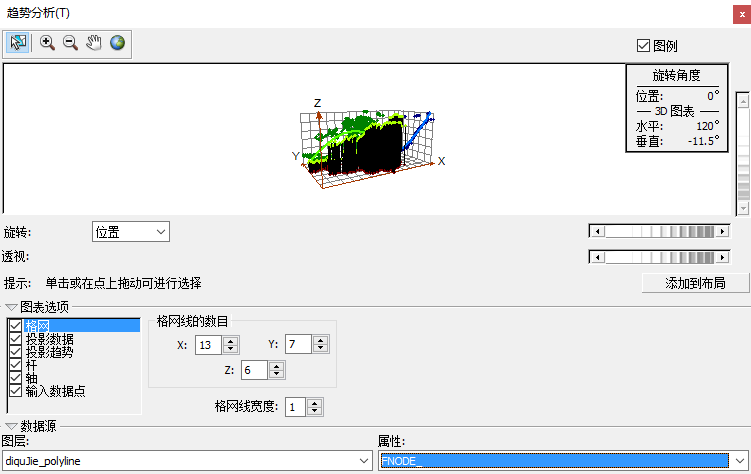




### 5.3 趋势分析

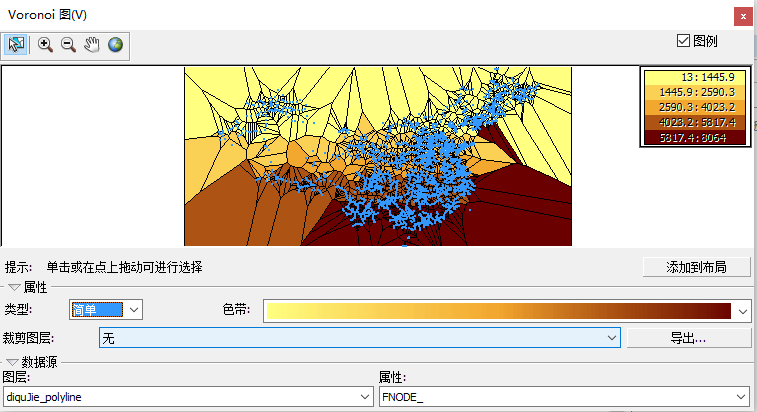
趋势分析图可以反映不同方向的趋势：蓝线表示南北方向；绿线表示东西方向，呈倒"U"形，可用二阶曲线拟合，在后面进行表面预测时将会去除。

点击Rotete右边的方向旋转箭头（横向箭头），可旋转趋势图，更明显地显示某一个方向的趋势。



### 5.4 Voronoi图

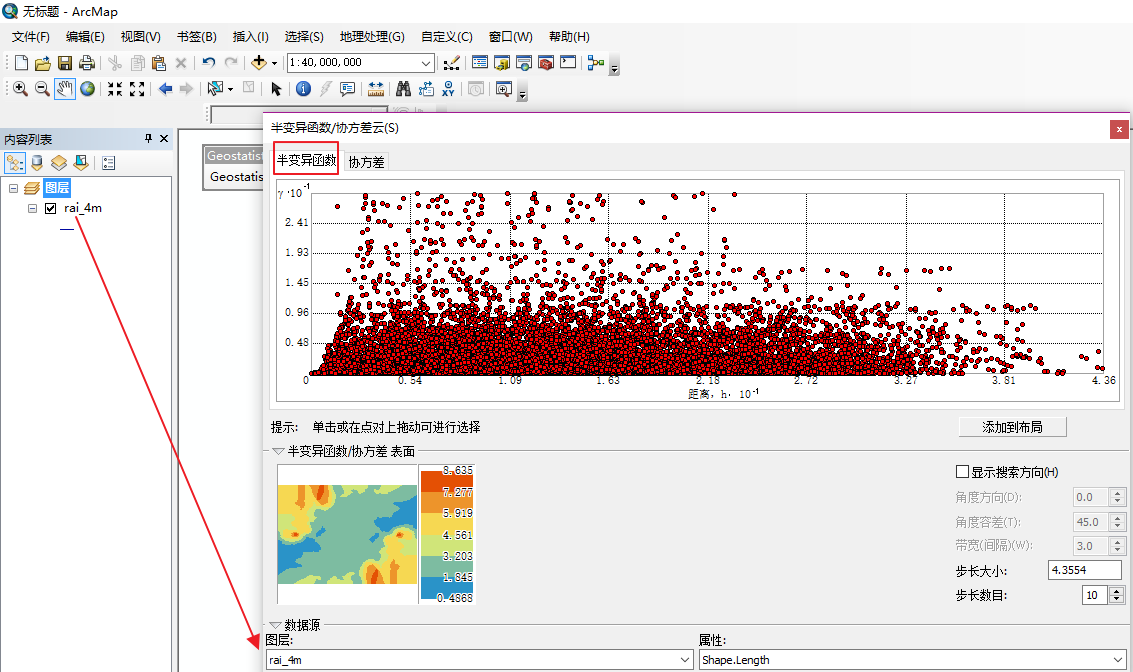
Voronoi图在实际上用途上主要用来发现离群值。Voronoi图的生成方法：每个多边形内有一个样点，多变形内任一点到该点的距离都小于其他多边形到该点的距离，生成多边形后。某个样点的相邻样点便会与该样点的多边形有相邻边。至于多边形值的计算有多种方法，可以用生成多边形的样点。值作为多边形的值（Simple方法），也可以以相邻样点的平均值为多边形的值（Mean方法），具体计算方法可以在Type下拉菜单中选择。

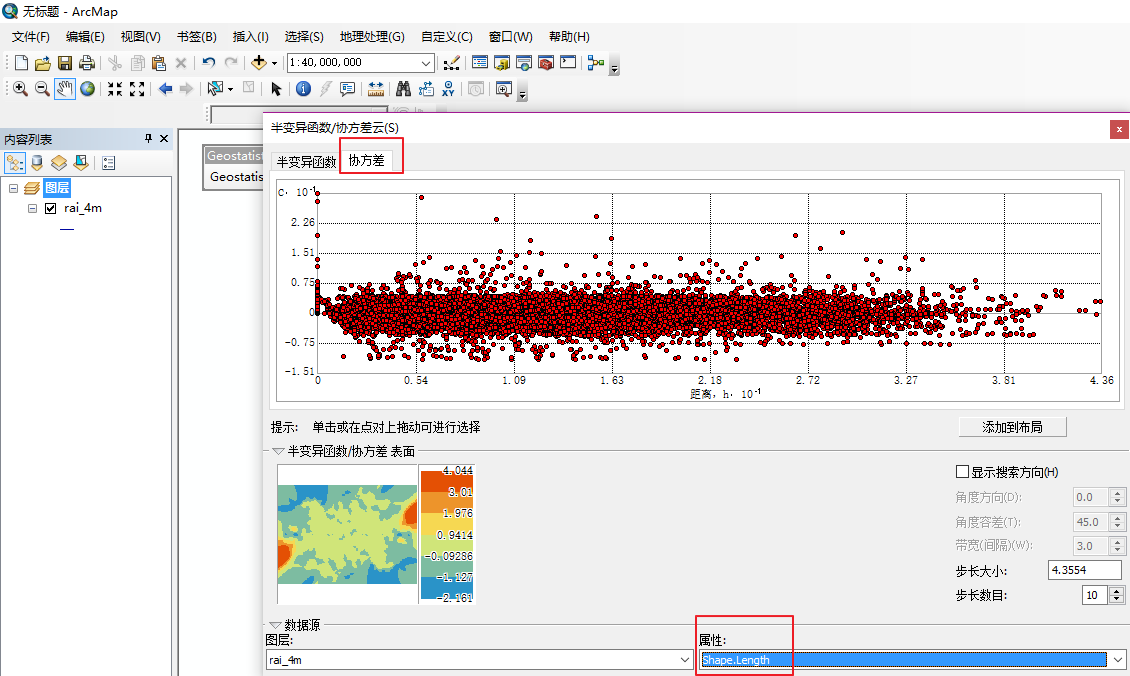


### 5.5 半变异函数/协方差云图

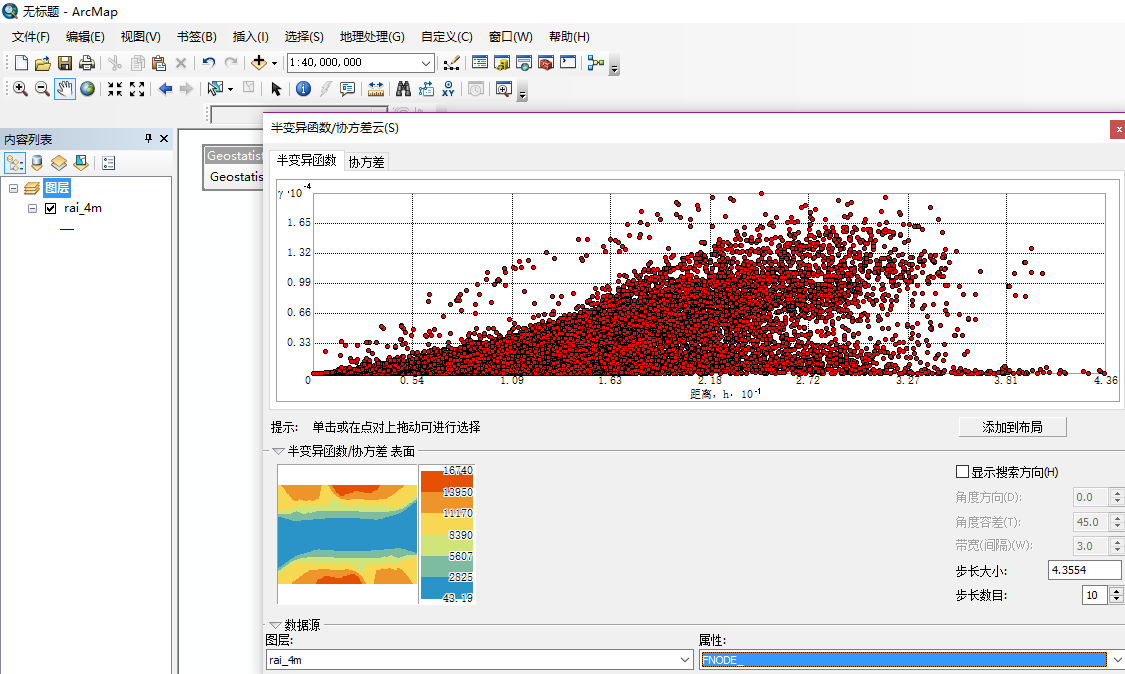
主要反映数据的空间相关程度。当数据存在空间相关关系时，才有必要进行空间插值法。图表的横坐标表示任两点的空间距离，纵标表示该两点的半变异函数值。根据距离越近越相似的原理，因而x值越小，y值应该越小。

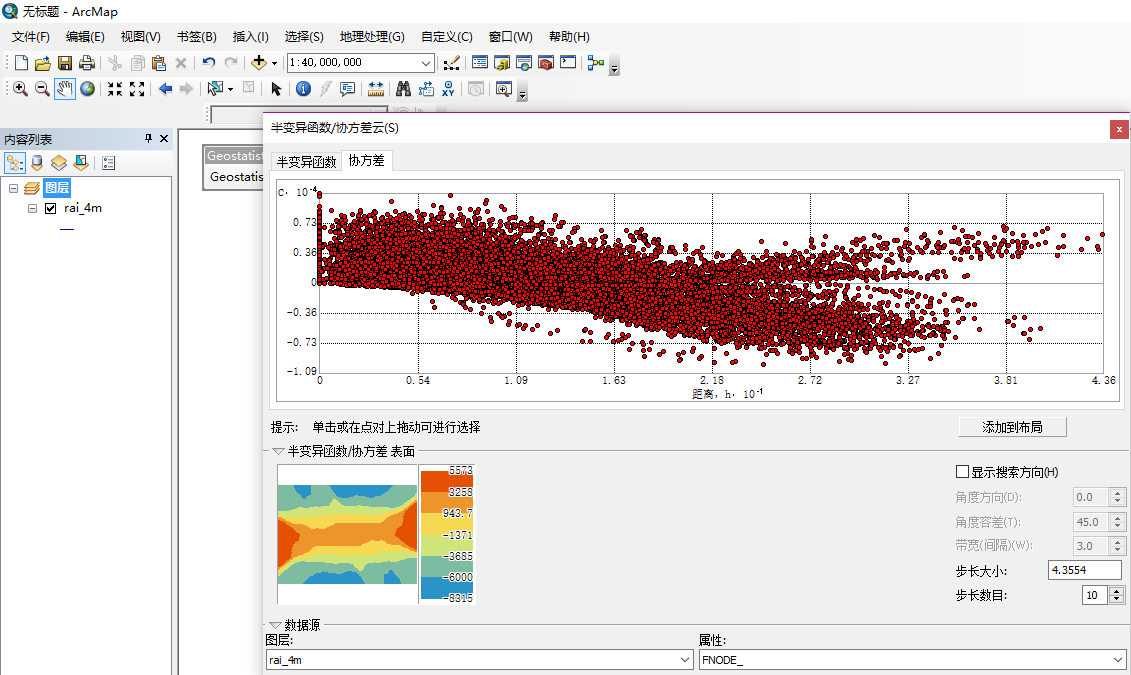
如果任意两点的值都要计算，当采样点很多时，数据量便很大，因而根据距离和方向对样点距离进行了分组。下列参数便是为此要求而设置：步长值；步长数量。步长值和步长组数之乘积应小于采样点区域的坐标范围的一半。





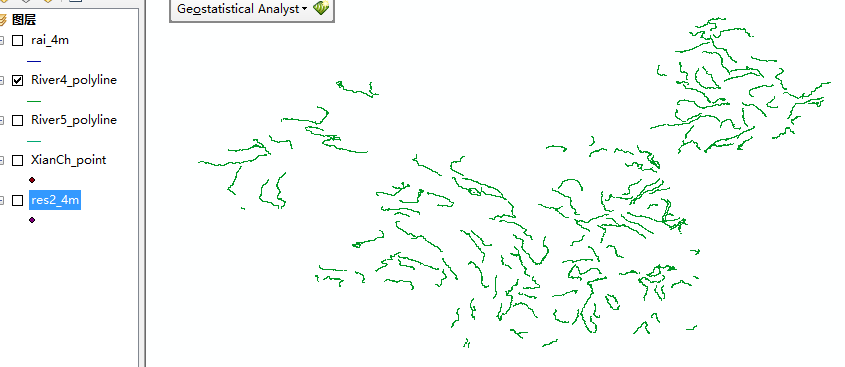
改变属性：

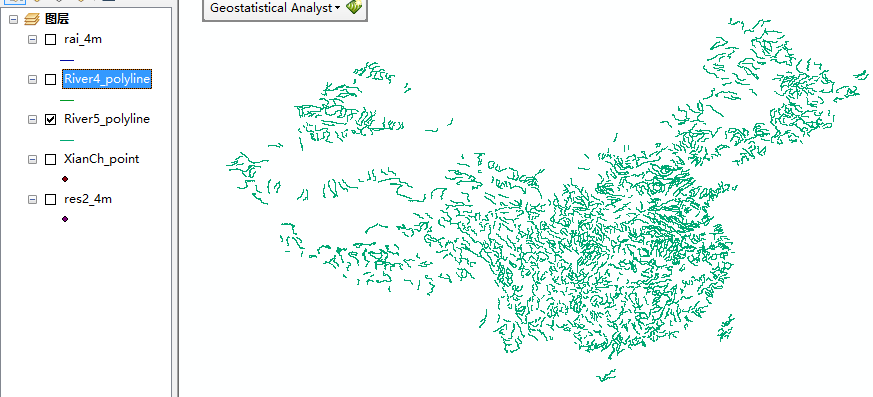


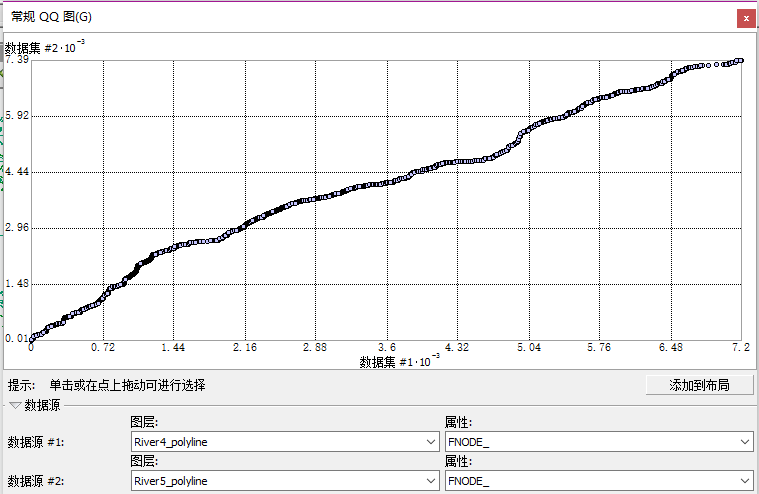


### 5.6 常规QQ图

主要用来评估数据集分布的相似程度，使用两个数据集中具有相同累积分布式的数据值来做图：

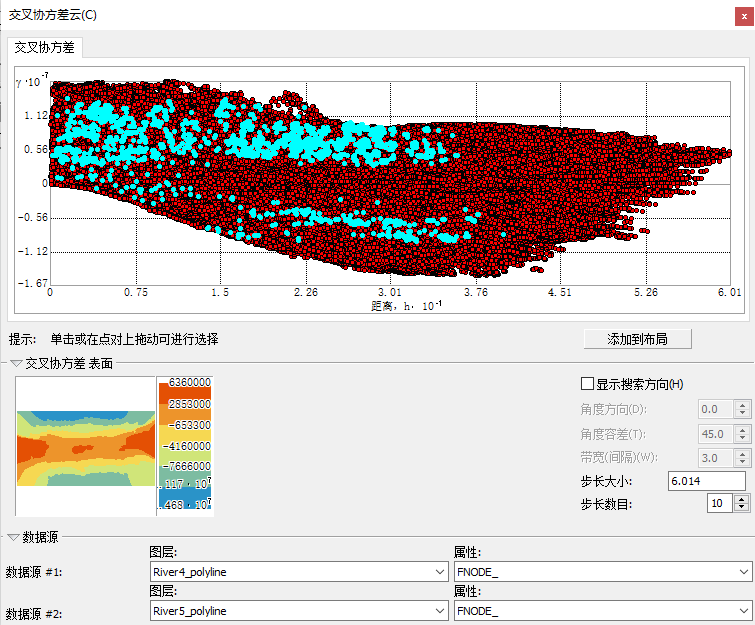






### 5.7 交叉协方差云图

横坐标表示两点间的距离，纵坐标表示两点间的距离所对应的样点对的理论正交协方差：

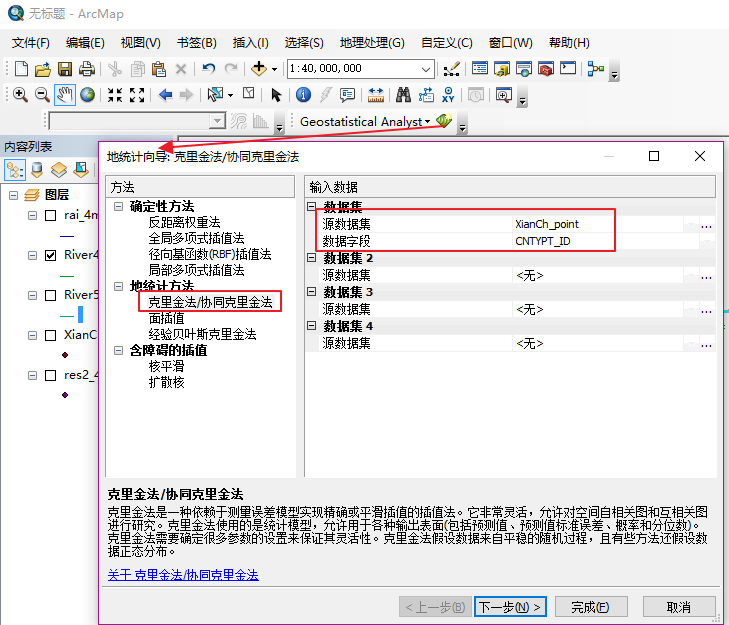


这些图彼此相关联，并与ArcMap中的图层相关联。即，在某个分析图中选择某些点，在ArcMap图层及其他分析图中同样会选中这些点。

## 5.2 地统计向导

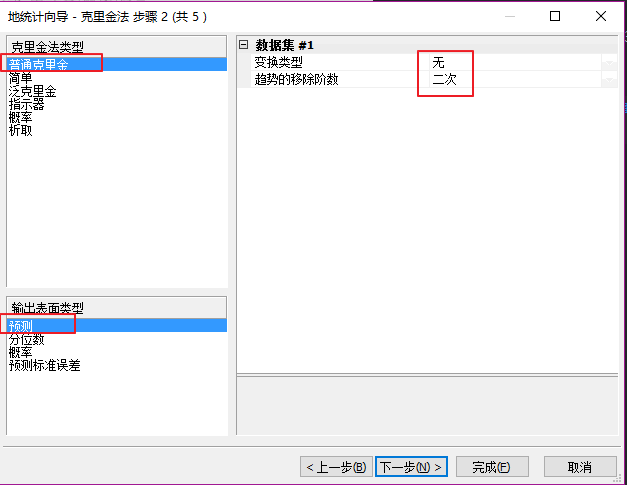
地统计分析模块提供了一系列利用已知样点进行内插生成研究对象表面图的内插技术。在ARCGIS的地统计分析扩展模块中，地统计分析向导命令从了解数据、选择内插模型、评估内插精度到最后完成表面预测和误差建模。

### 5.2.1 选择输入数据和方法面板



预测方法的选择要根据数据分析的结果而定。现在假如选择普通Kriging方法（其实所谓地统计方法，最主要并且用的最多的就是Kriging方法的几种变化形式），即使用普通克里金方法生成一个表面预测图。普通克里金方法是最常用的地统计分析方法，其他几种依次为简单克里格、泛克里格、指示克里格、概率克里格、析取克里格。这几种克里金的区别是由于克里格的形式及其数据特点的不同。

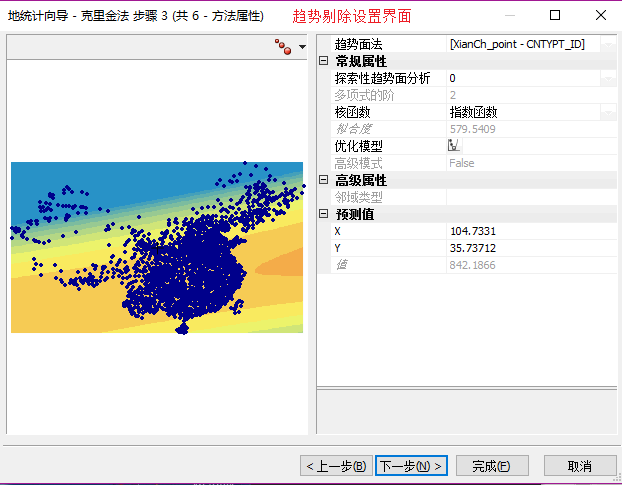
### 5.2.2 地统计方法选择面板



变换类型：对数据集进行转换，由于某些方法要求数据正态分布，因此如果数据与正态分布差距很大，可以在此选择一种方法对数据进行转换。 趋势的移除阶数：如果数据在某方向上存在趋势，则为了提高预测的准确性，一般要剔除趋势。在此处选择趋势方程的阶数：线性、一阶、或无趋势等。数据的趋势有无以及阶数在数据检查时得到，即用Explore Data菜单下的Trend analysis来分析得到。

### 5.2.3 趋势剔除面板

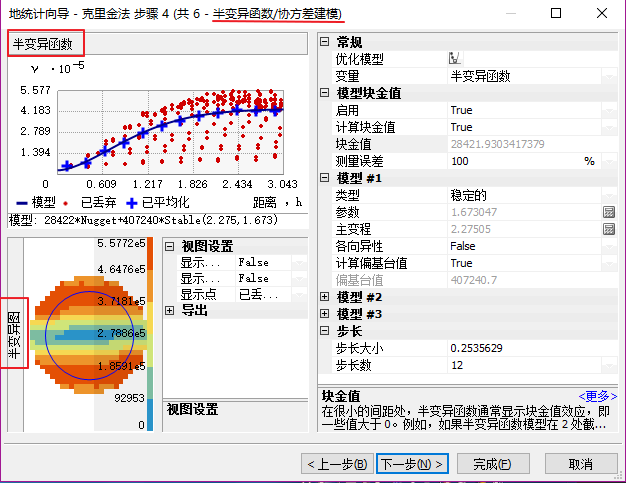
此面板只有在第二步中选择了【趋势的移除阶数】选项时才会出现，一般为缺省即可。



### 5.2.4 半变异函数/协方差模型面板

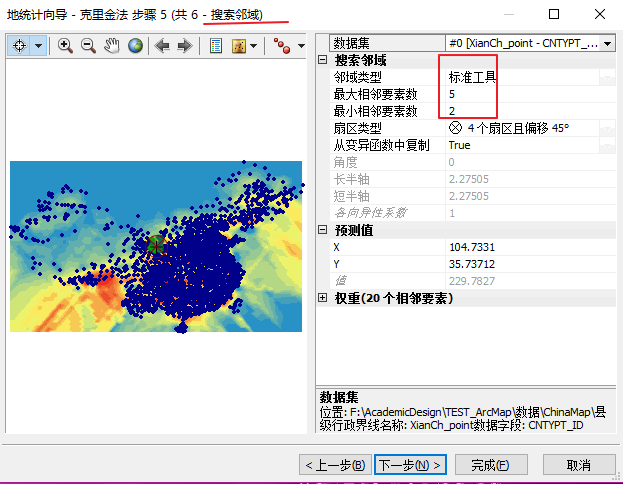
此步的主要功能为半变异函数建模，是预测过程中的实质性阶段。在此面板中需要设定许多与拟合半变异函数相关的选项以及半变异函数的参数，是克里格预测中十分关键的部分。

半变异函数：显示的是拟和的模型，蓝线即半变异函数曲线。模型：model1,model2,model3表示可以用多个通用函数来拟和半变异函数模型。如果数据为各向异性，则需要选中Anisotropy（其实大多数空间数据是各向异性的，各向同性只是相对的），当选中此选项时，蓝线变为多条，表示多个方向的拟合函数。 Show Search Direction选项选中后，表示只搜索某个方向的半变异函数。 Nugget：块金值，函数参数之一，即函数与y轴相交的y值。 Error Modeling：如果数据中有测量误差（比如一起原因等）的话，则选中此项，预测表面将光滑许多。



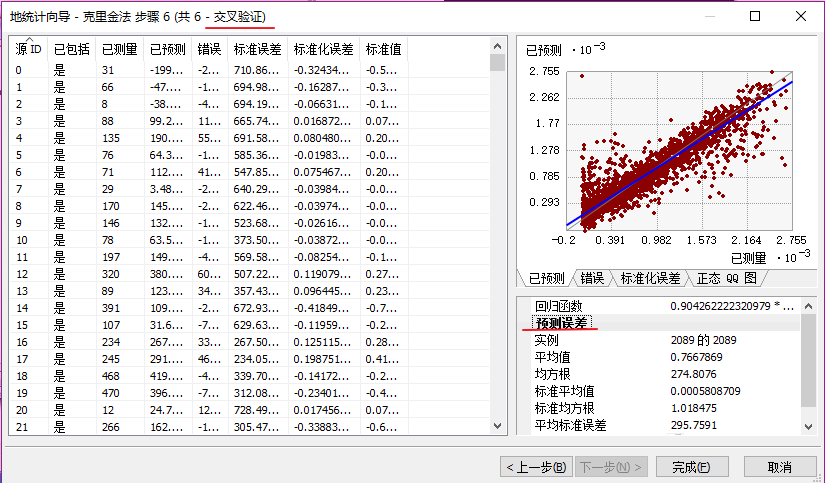
### 5.2.5 搜索区域面板

此面板的主要功能是设定预测某点数值时如何搜索邻近的已测量点。 主要有样点数（neighbours to）和搜索形状（shape）两个选项。离预测点太远的样点对预测无意义，所以最大相邻要素数不应设置过大。



### 5.2.6 交叉验证面板

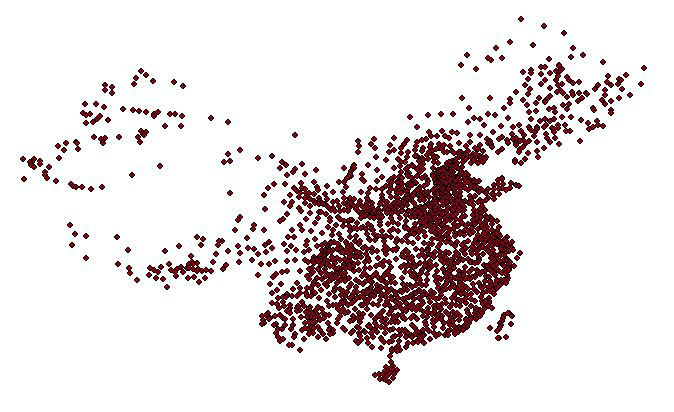
在此面板中查看预测的精度，右上方的图表中，横坐标为测量制值，纵坐标为预测值，最理想的情况是数据呈1:1线，即图中的破折线。 左边的预测误差（precited error）项是预测误差的一些统计值，可很好的体现预测的好坏。

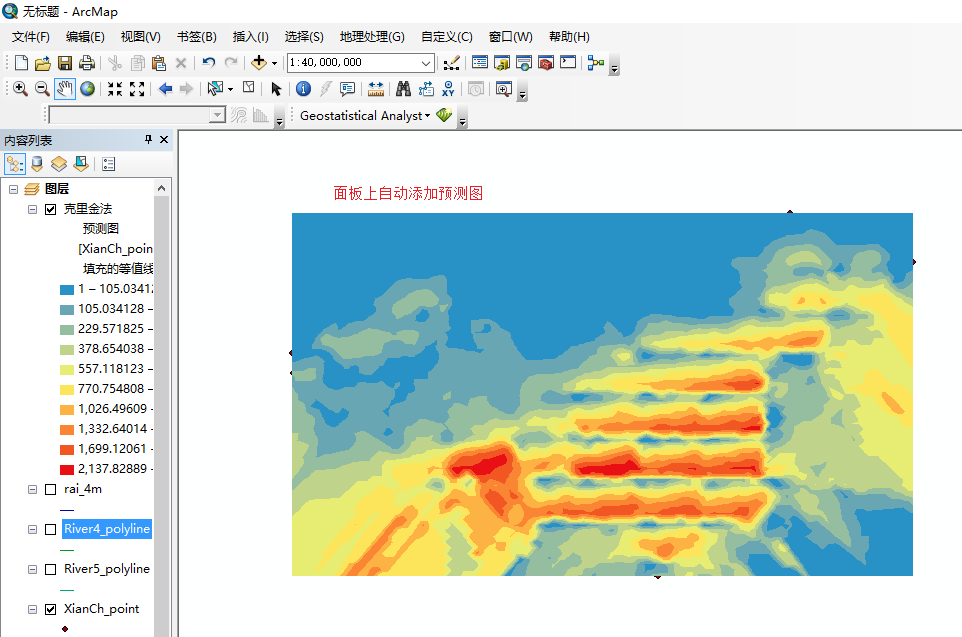


### 5.2.7 数据图层信息面板

该面板中显示了在数据预测过程中设置的参数，可以查看。点击OK，即可生成预测图。



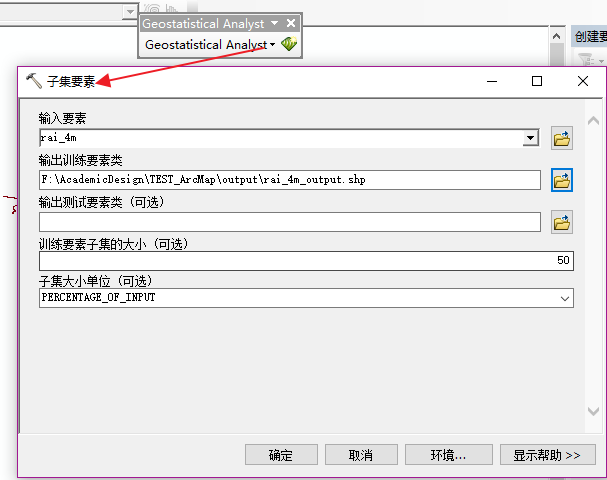
原图：

预测图：

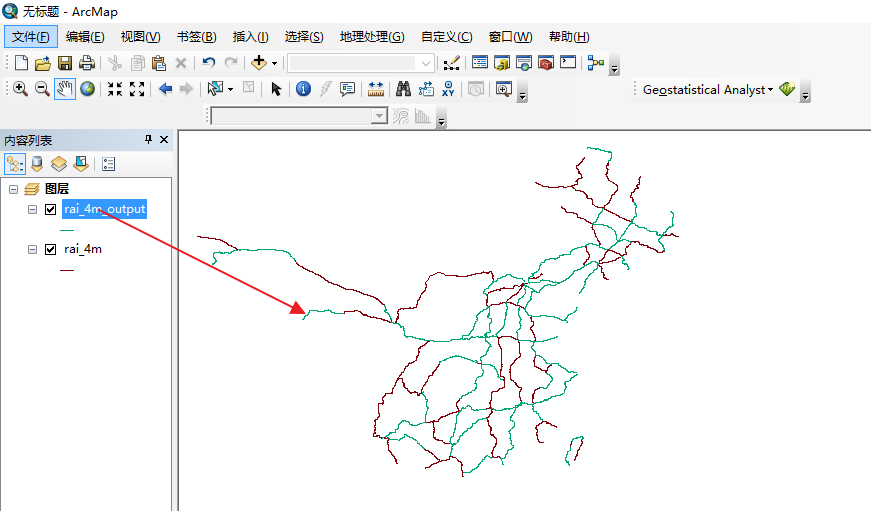
一般情况下，有时候某些参数难以判断，因而会生成几个预测表面，然后比较不同表面的精度，选择精度最高的作为结果。（普通Kriging表面是用上述过程中的方法生成的预测表面，default是用缺省的参数得到的预测表面）

## 5.3 子集要素

通常，对输出表面进行质量评价的最严格方法是将观测值与预测值进行比较，在实际应用中，这种方法的优点是不需要到研究区采集独立的验证数据集。

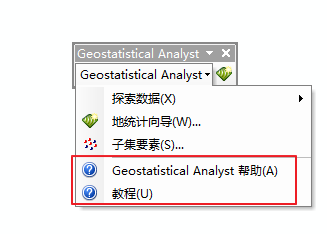






## 5.4 使用帮助

该帮助系统分为两大部分，一部分是地统计扩展模块的教程，一部分是帮助文档：



# 6 总结与探讨

通过阅读几篇论文，可以加深对空间插值算法的理解。在读论文的过程中，确实存在一些看不懂的，例如涉及到的偏向于地理方面的专业用语、表达式的推导过程、数据的验证建模转换过程，也有看不明白分析的结果的。

拿我最终整理的四篇论文《ArcGIS中的地统计克里格插值法及其应用》、《Kriging空间分析法及其在地价评估中的应用》、《基于GIS的气象要素空间插值方法研究》、《普通Kriging法在空间插值中的运用》来说，可以发现：①由于区域范围、气象要素种类及其时间尺度的不同，造成数据的时空变异特征不同，其空间插值方法也不尽相同，应当根据插值数据的特征、插值对精度的要求以及计算试验效果分析等因素综合选择合适的插值方法。 ②Kriging 以最佳线性无偏差估计而在地统计学中得到广泛应用，其中，普通Kriging法插值结果可以较真实地模拟地质特征。③半变异函数是Kriging插值的核心，它模拟了区域化变量随空间距离增大而产生的变异情况，它是空间未知点取值的依据。通过交叉验证5个检验指标(平均误差、均方根误差、平均标准误差、标准化平均误差和标准化均方根误差)来评价各个半变异函数模型的拟合效果。④实验中插值的过程大概都是：分析源数据特性——>选择2~3种比较适合的插值方法进行插值计算——>分别对插值结果进行验证比较——>选择最佳的插值方法。⑤在利用普通克里金插值法对而在进行Kriging插值时，选用什么样的模型和参数也很重要。所需参数较多，每个参数的设置都影响到下一步的结果。⑥针对数字地面模型的空间内插值，球形模型能够较为准确地套合实验变异函数值，得到较为精确的块段值、变程、基台值。

# 参考来源

中国大学imooc《计量地理学》视频教程第11讲 地统计分析

《ArcGIS中的地统计克里格插值法及其应用》

《Kriging空间分析法及其在地价评估中的应用》

《普通Kriging法在空间插值中的运用》

《地理信息系统中不同空间插值方法的比较研究》

《东北地区逐日气象要素的空间插值方法应用研究》

《基于GIS的空间插值方法研究》

《基于GIS的气象要素空间插值方法研究》

《基于Kriging法的森林土壤养分空间插值》

《基于地理信息系统的空间插值算法研究》

《空间插值算法研究及其在遥感数据模拟中的应用》

空间插值——插值方法的适用范围：

<http://www.cnblogs.com/zhuyuchen/p/3810811.html>

地理科学学院地理信息系统专业网络课程教程第7章——空间数据分析模型：

<http://kc.njnu.edu.cn/dky/nb/page/2000-3-3/2000332117262480.htm>

空间插值方法对比整理版：

<https://wenku.baidu.com/view/d34f957827284b73f2425069.html>

空间插值方法简介：

<https://wenku.baidu.com/view/82c52c21192e45361066f5e3.html>

空间插值算法汇总：

<https://wenku.baidu.com/view/7941bbc0aa00b52acfc7ca90.html>

空间数据插值方法的分析与比较：

<https://wenku.baidu.com/view/c2476b75a417866fb84a8e6f.html>

10种插值方法在物探数据处理中的对比：

<https://sanwen8.cn/p/400ebZM.html>

第九章地理信息系统空间插值：

<https://wenku.baidu.com/view/9a7d3921b9f3f90f77c61b74.html>

空间插值与地统计：

<https://wenku.baidu.com/view/d44e307c8e9951e79b892742.html>

地理信息系统课程GIS空间插值：

<https://wenku.baidu.com/view/fccc39f9a98271fe900ef95d.html>

克里格法（Kriging）：

<http://blog.sina.com.cn/s/blog_49babfed010007wd.html>

Kriging插值法：

<http://blog.sina.com.cn/s/blog_4aa4593d01014gwc.html>

克里金插值(kriging)：

<https://wenku.baidu.com/view/b497d80979563c1ec5da712a.html>

插值算法（二）：反距离加权法IDW：

<http://blog.sina.com.cn/s/blog_816800900101f2lk.html>

西南大学图书馆：

<http://202.202.96.94/>

地图上的各种等高线是怎么画出来的：

<http://daily.zhihu.com/story/9228981>

ArcGIS 帮助（10.2、10.2.1 和 10.2.2）：

<http://resources.arcgis.com/zh-cn/help/main/10.2/index.html#//009t00000002000000>

arcgis10.2操作教程 第二章 ArcGIS 10基础操作：

<http://www.92to.com/xuexi/2016/10-08/11319886.html>

ArcGIS 10用户使用指南

<https://wenku.baidu.com/view/7b41b729aaea998fcc220ead.html>

Arcgis地统计图绘制：

<https://wenku.baidu.com/view/bab8b3057cd184254b3535be.html>

克里金(Kriging)插值的原理与公式推导：

<https://xg1990.com/blog/archives/222>

kriging基础知识：

<https://wenku.baidu.com/view/3d142be827284b73f24250d7.html>

Kriging方法的公式推导：

<https://wenku.baidu.com/view/2827d58db7360b4c2f3f6461.html>

普通Kriging法在空间插值中的运用：

<https://wenku.baidu.com/view/df8155d3195f312b3169a5a6.html>

**成绩：**

**指导教师（签字）：**

**年 月 日**