2010 Esri 中国区域用户大会

ArcGIS 地统计分析介绍

张文

GIS–让人类认知世界

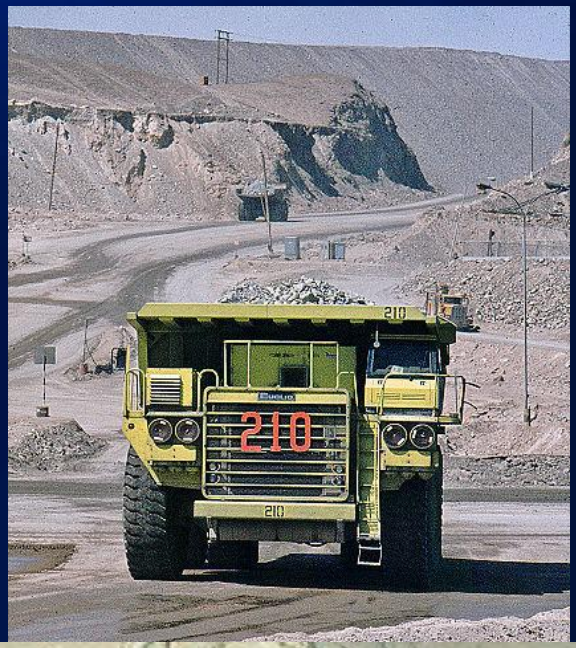
esri China BEIJING

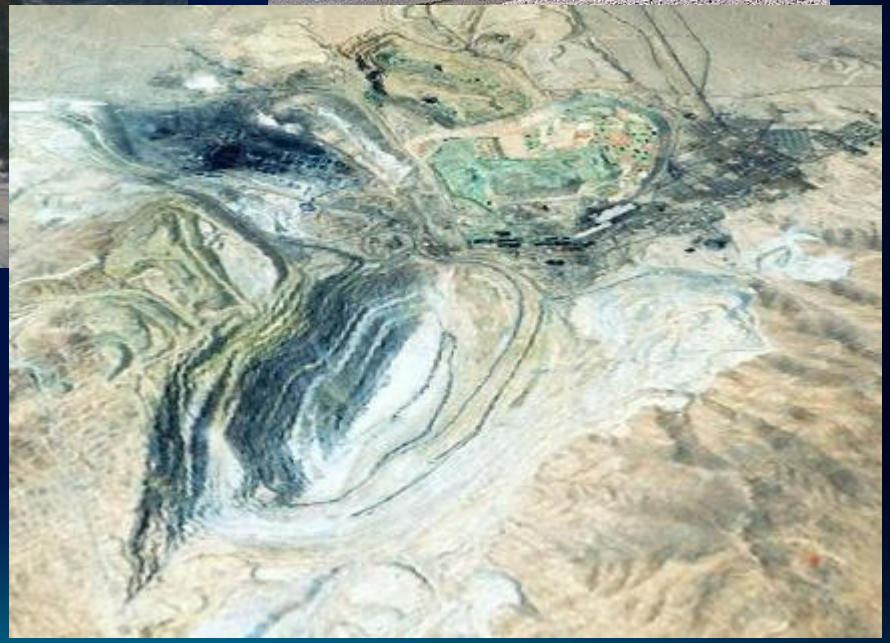
概要

- 什么是地统计？
- 如何使用地统计分析模块？
  - ✓数据分析（Explore Data）
  - ✓地统计分析向导（Geostatistical Wizard）
  - ✓创建数据子集（Create Subset）
  - ✓地统计分析工具（Tools）
- ArcGIS 10 地统计分析新特性

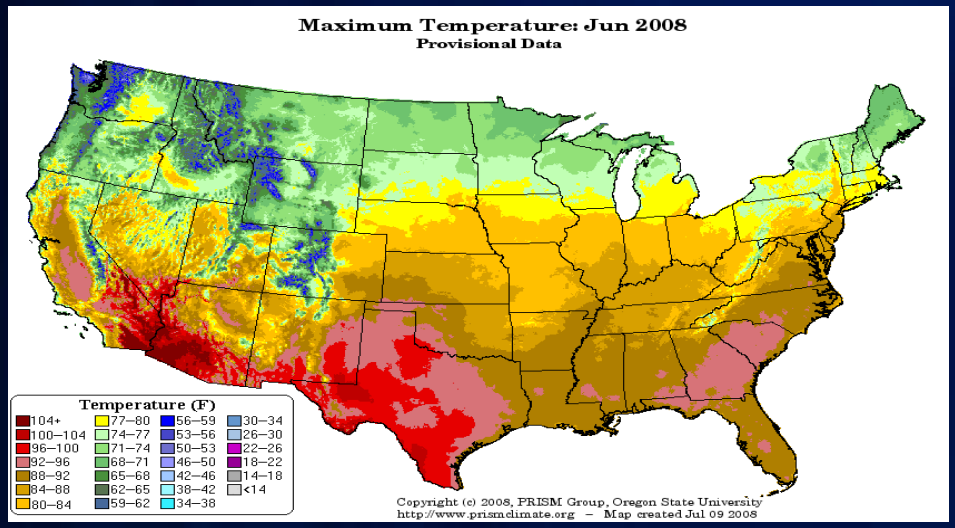
采矿业



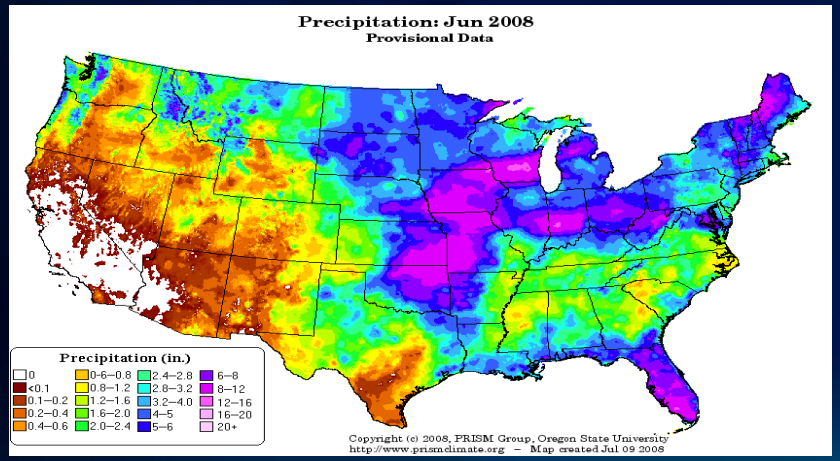




气象学

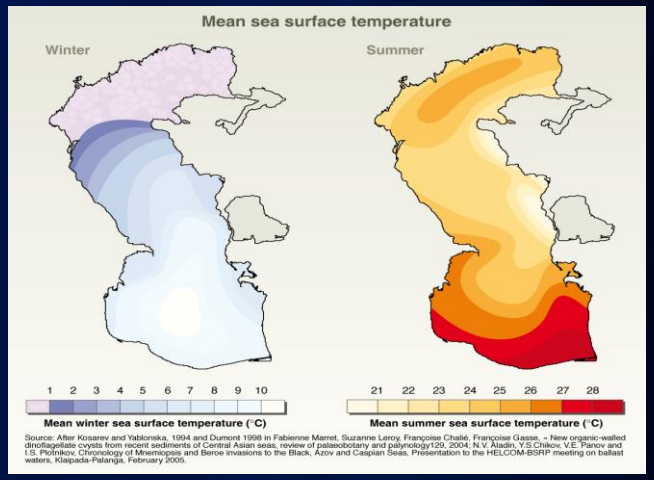


最高温度图



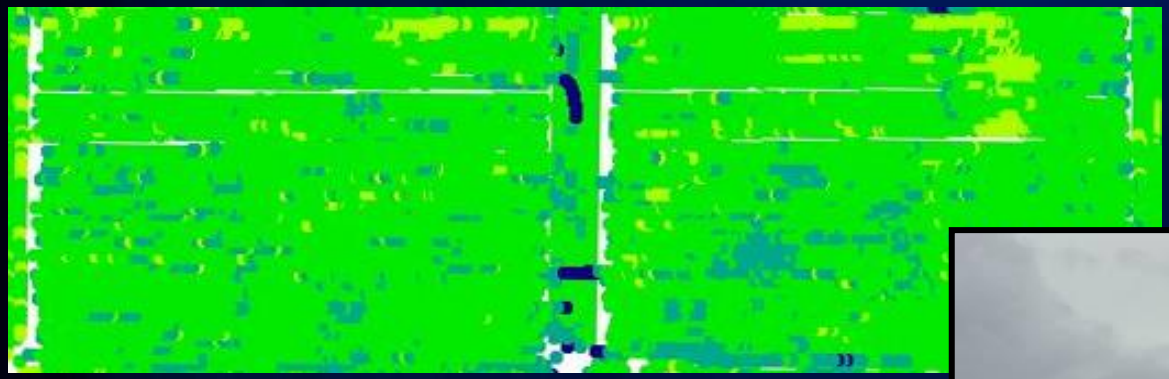
6月降水量

海洋学

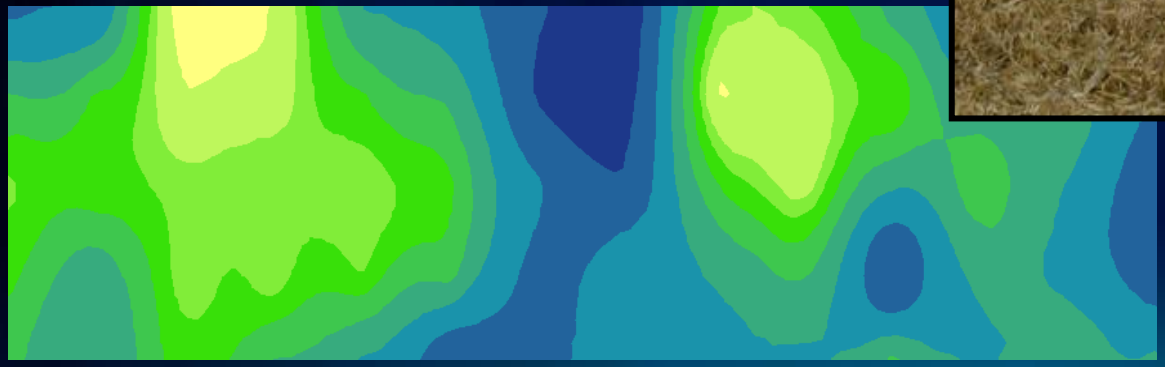


里海平均海面温度图

精确农业(Precision Agriculture)







收割产量

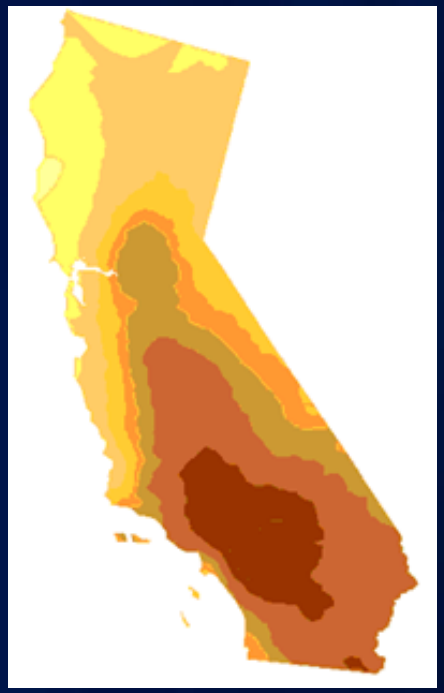


土壤肥力

地统计能够做什么？



样本点

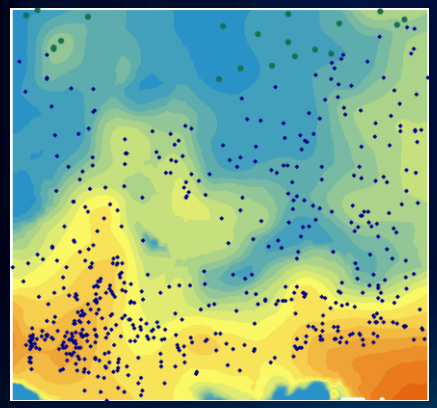


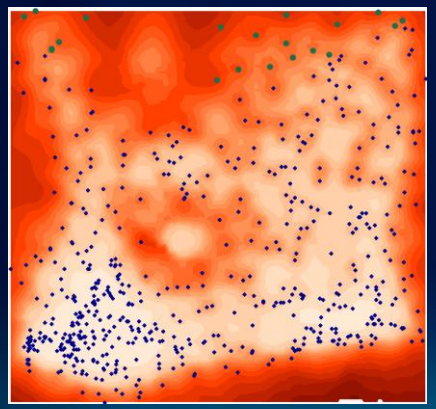
趋势面

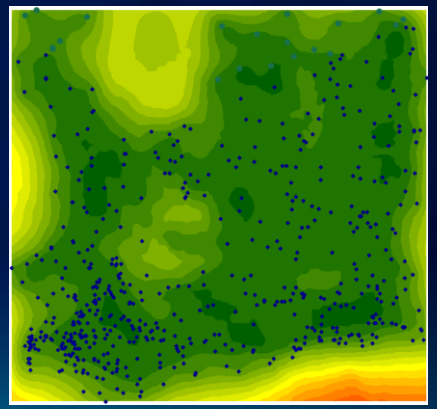
什么是地统计?

空间相关数据的统计分析

- 地统计学利用随机函数对不确定的现象进行探索分析，并结合采样点提供的信息对未知点进行估计和模拟。
  - 一个具有空间相关性的自然现象，样本点的选择
  - 样本值的大小，以及样本空间位置及样本间的距离，

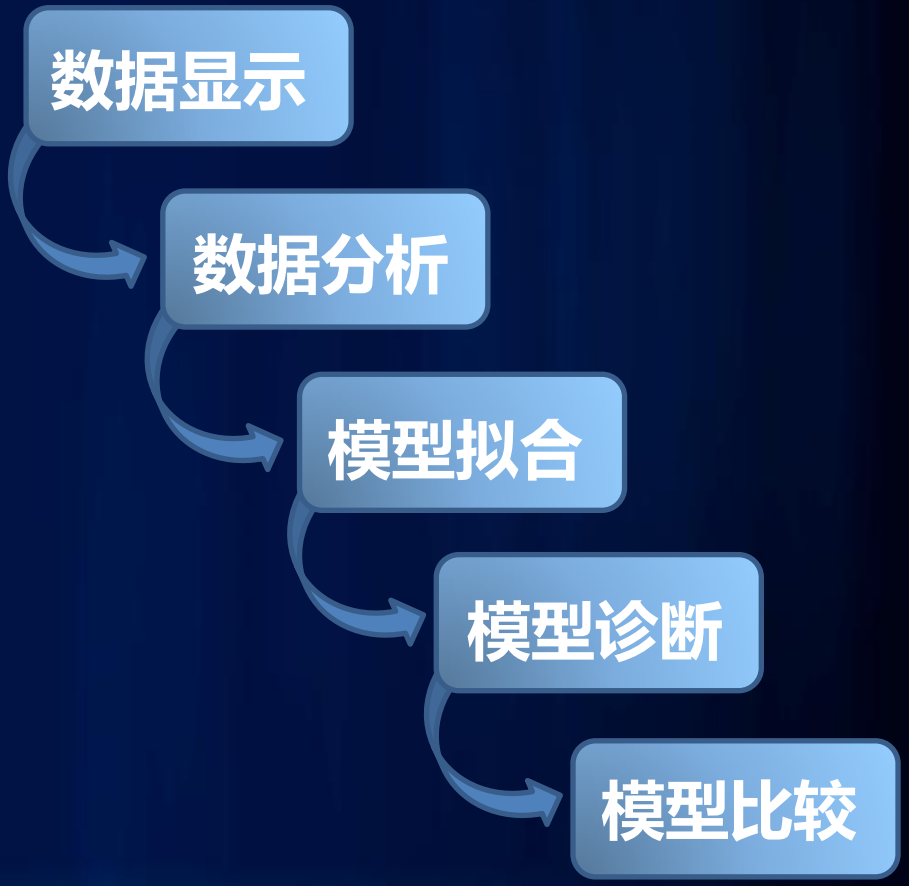






如何进行地统计分析?

- 地统计分析理论
  - 前提假设
  - 区域化变量
  - 变异分析
  - 空间估值



ArcGIS 中的地统计分析



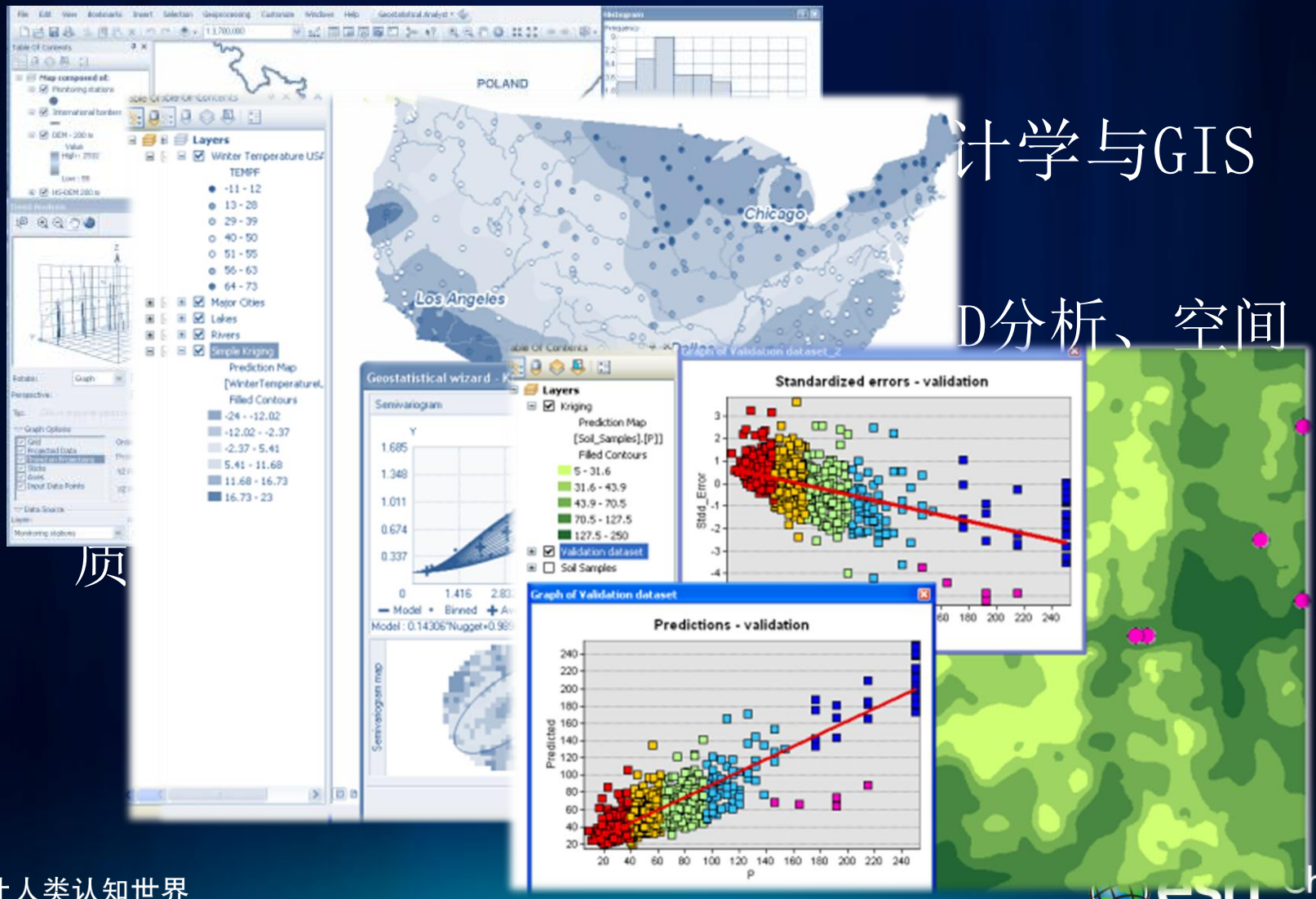
GIS-让人类认知世界

ArcGIS中的地统计分析

- 计学与GIS

D分析、空间

质



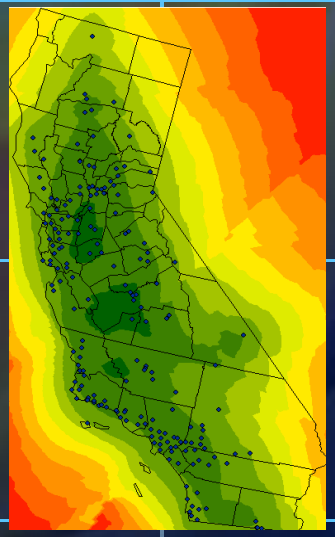
esri China BEIJING

地统计分析模块提供了…

- 空间数据分析（ESDA）
  - 直方图, QQplot, 趋势分析 …
- 确定性插值
  - IDW, GPI, RBF, LPI
- 地统计插值
  - 克里金/ 协同克里金
    - Ordinary, Simple, Universal, Indicator, Probability, Disjunctive
- 地统计模拟
- **含障碍的插值方法**
  - 核平滑, 扩散插值
- **采样网络设计**
- **…**


演示：地统计分析之初体验
地统计分析扩展
向导（wizard）
sand,stone
rock
sand,shells
Q(3)W 10s
rock
sand
"S. Lookinghaven"
Q(6)+LFl(1)W 15s6m3M
Old Channel
rock,shells
"Duke Rock"
rock,sand
No 2
sand,mu
Fl(1)Y 2s
MICKLEFIRTH
CHANNEL

插值方法（Interpolation）





地统计中的插值方法

确定性插值

地统计插值



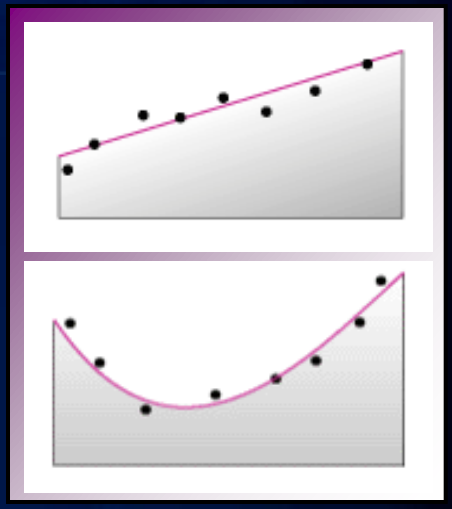
地统计中的插值方法

- 确定性插值
  - 全局多项式插值（GPI）
  - 径向插值（RBF）
  - 局部多项式插值（LPI）
  - 反距离权重（IDW）

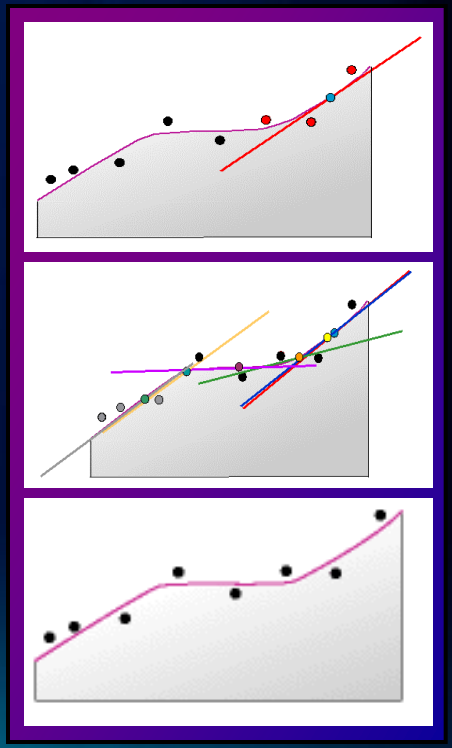
精确性插值

全局和局部多项式

- 全局多项式
  - 用一个平面或曲面（多项式）对整个研究区的样点数据集进行全局拟合
  - 常用于缓慢变化的现象



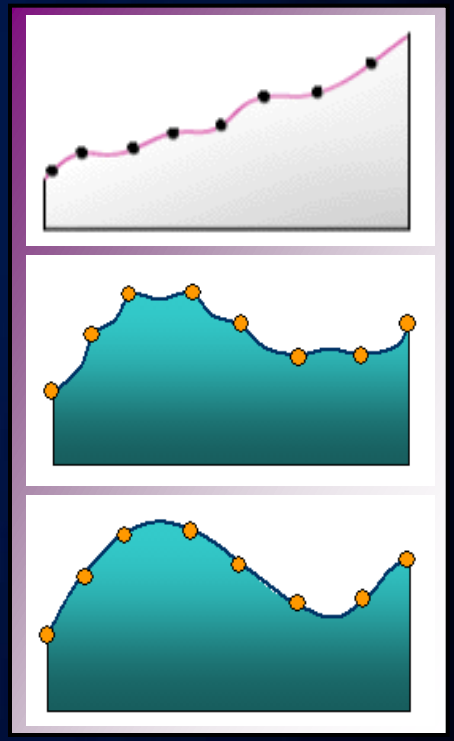
- 局部多项式
  - 利用多个小平面或曲面（多项式）来对研究区的数据集进行拟合
  - 领域设置：样本点的数量和领域搜索参数.
  - 模型考虑方向连续性
  - 常用于具有短程变异的现象



在使用过高阶的多项式时，过度平滑会生成无效结果

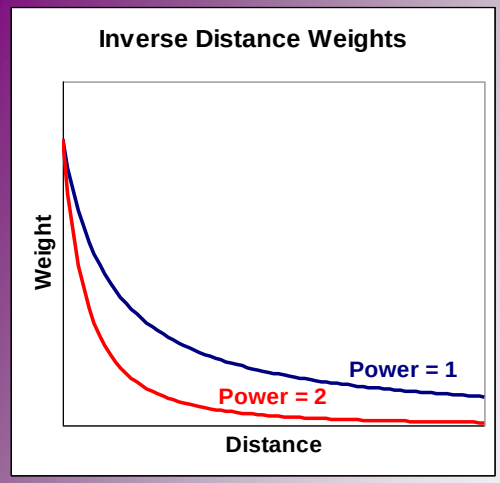
径向插值

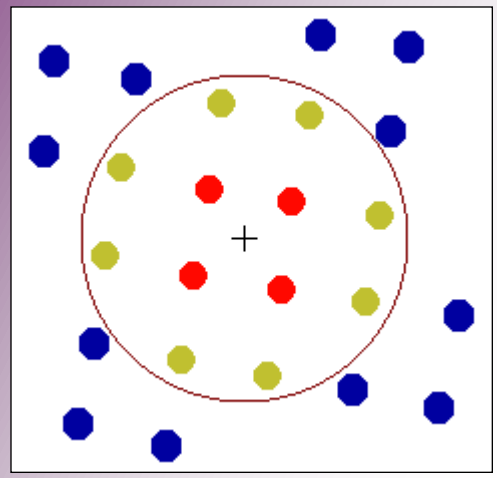
- 揉捏表面，表面经过每一个已知样本点
- 最小化表面的曲率.
- 通过选择函数来决定径向表面穿过样本点的方式
- 模型参数: 样本点个数和领域搜索参数.
- 模型考虑方向连续性
- 适用于大数据量的样本和缓慢变化的现象.



反距离权重插值

- 表面经过所有的已知样本点
- 使用先入为主的空间相关性
- 根据周围样本值的加权平均来进行预测
- 权重随着距离增大而递减，越高级数递减越快
- 模型参数: 样本点个数和领域搜索参数.
- 模型考虑方向连续性
- 适用于通过定义搜索领域能够捕捉到局部变化的现象。

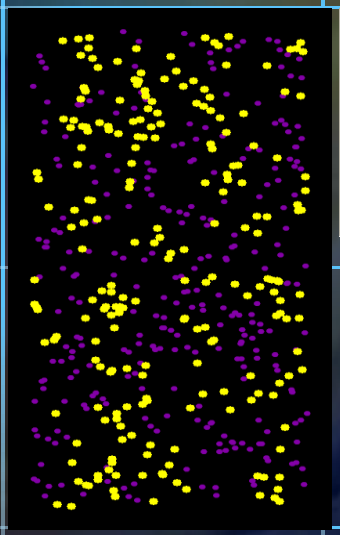


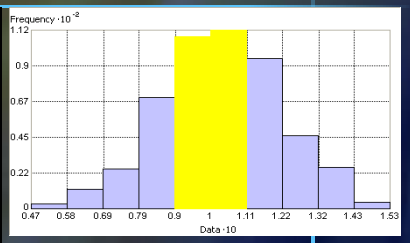


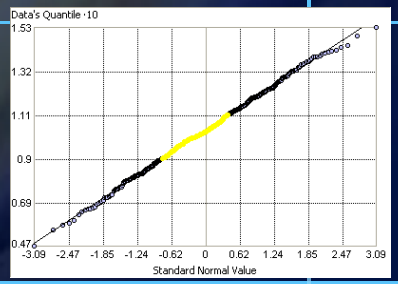
生成预测表面的方法



空间数据分析（ESDA）




Frequency ·10 -2
1.12
0.9
0.67
0.45
0.22
0
0.47 0.58 0.69 0.79 0.9 1 1.11 1.22 1.32 1.43 1.53
Data ·10

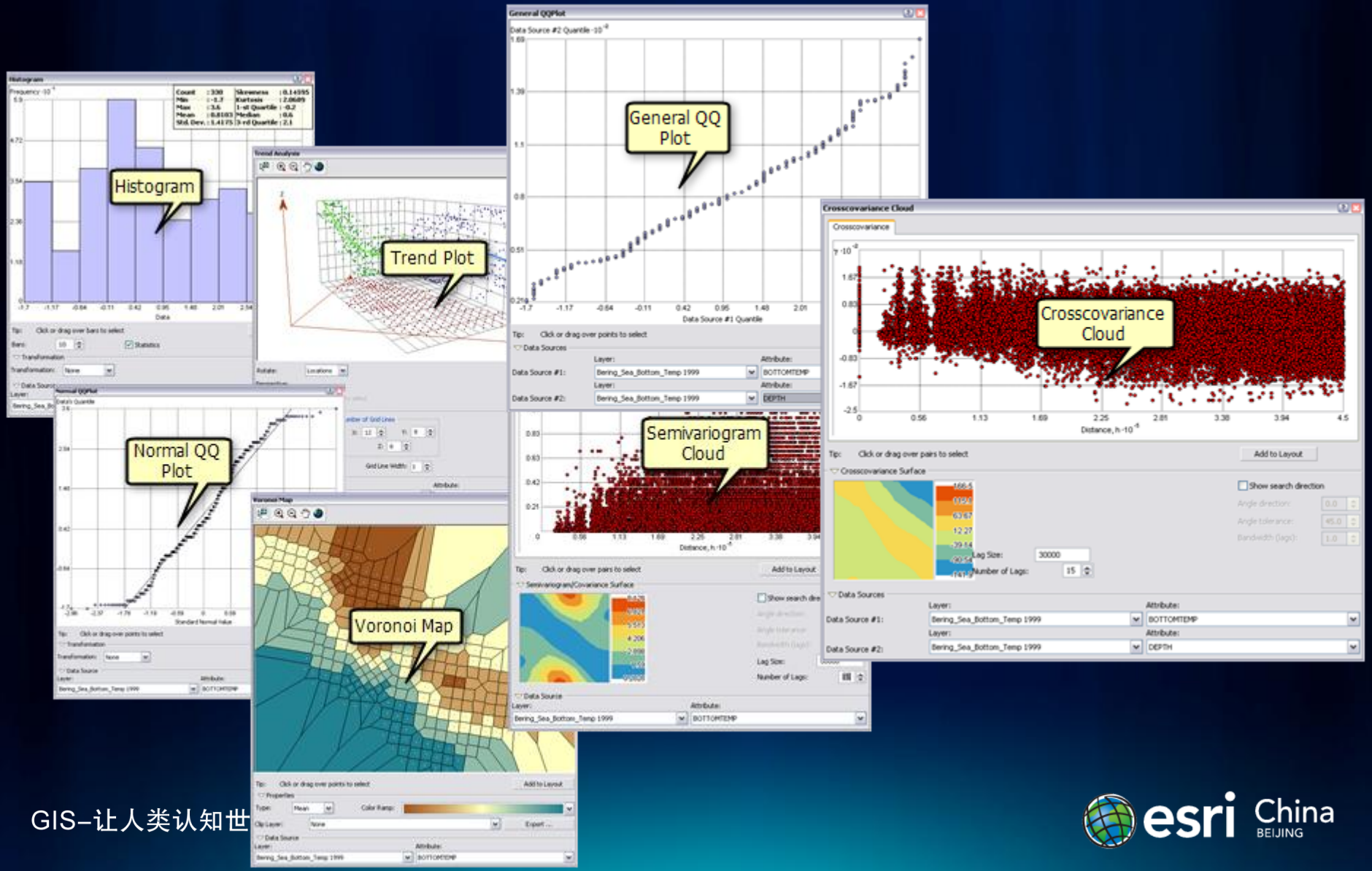

Data's Quantile ·10
1.53
1.32
1.11
0.9
0.69
0.47
-3.09 -2.47 -1.85 -1.24 -0.62 0 0.62 1.24 1.85 2.47 3.09
Standard Normal Value


GIS
geography
esri
opening
the world
to everyone

空间数据分析(ESDA)

- 空间数据分析可以增强对数据的了解，从而在进行插值的时候选择最合适的模型和参数
  - 异常值
  - 空间依赖性
  - 空间分布趋势
  - 平稳性

空间数据分析(ESDA)


Histogram
Trend Plot
General QQ Plot
Crosscovariance Cloud
Normal QQ Plot
Semivariogram Cloud
Voronoi Map

地统计插值

- 克里金
  - 通过对已知样本点赋权重来求得未知样点的值
  - 权重值:
    - 基于空间相关性模型（变异函数）.
    - 线性无偏最优估计
    - 考虑了已知样本点的空间分布及与未知样点的空间方位关系

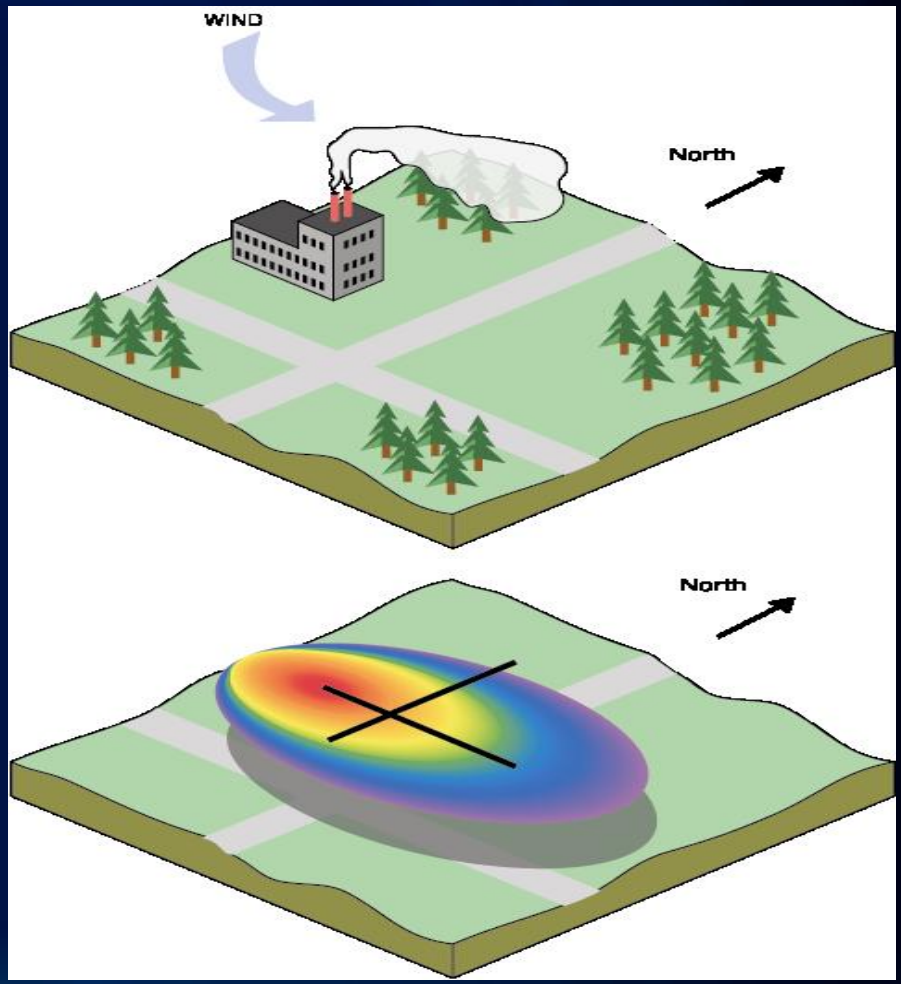
基于区域化变量存在空间相关性的假设.

克里金插值方法

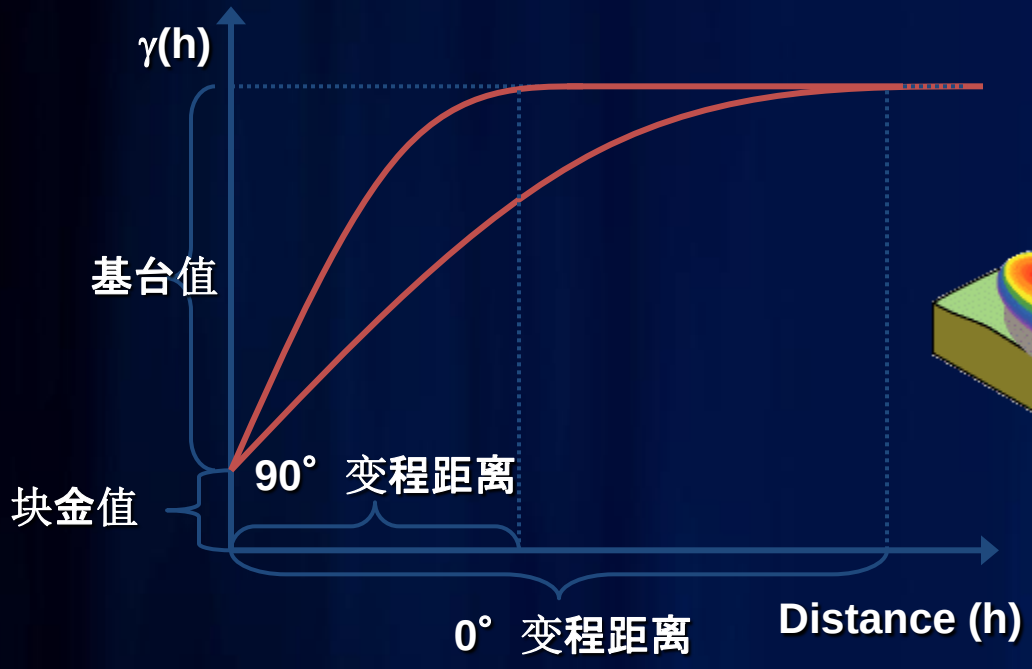
- 普通克里金
  - 广泛使用.
  - 属性的期望值是未知
  - 克里金方差与样本点位置相关
- 简单克里金
  - 属性的期望值为某一已知常数
  - 克里金方差与样本点的位置和样本值相关
- 指示克里金
  - 通常用于二进制数组 (0和 1).
  - 用于分析属性值是否超过某一阈值的问题
- ...

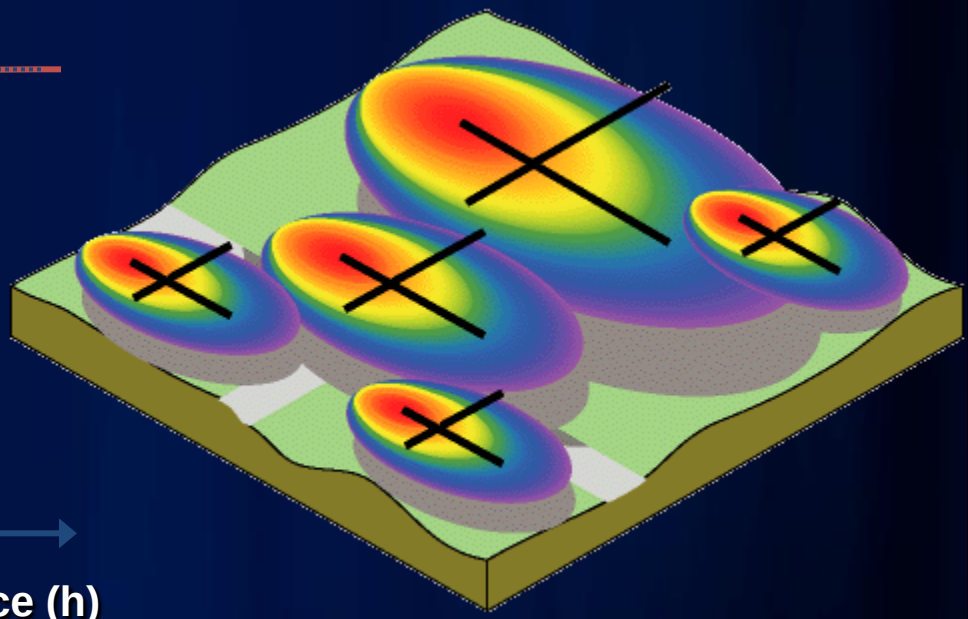
趋势特征

- 排除趋势
  - 单独分析剩余部分
  - 生成最后结果之前再自动结合进来



搜索方法



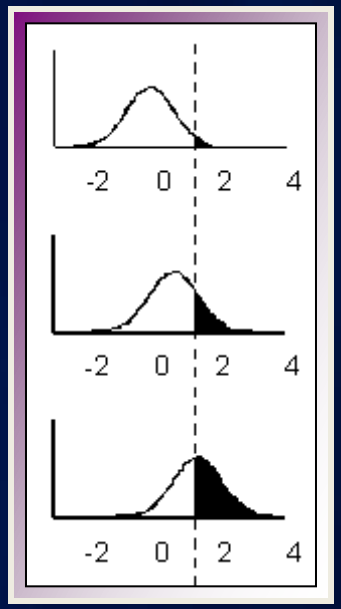


各向异性: 各个方向表现出的自相关性有所区别

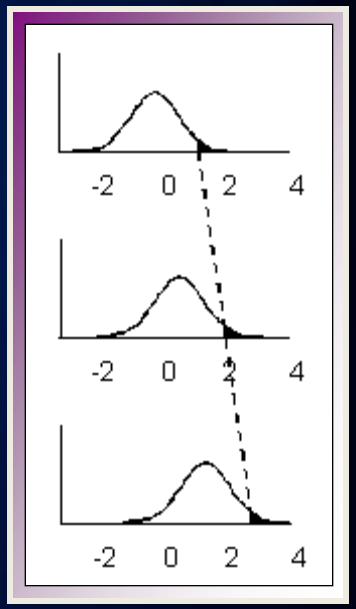
克里金插值结果

需求	输出类型
非样本区域的预测值	预测图 Prediction map
预测值的不确定性分析	标准差图 Standard Error map
预测将会超出阈值的概率	概率图 Probability map
获得与某特定概率相关的值	分位数图 Quantile map

预测值的概率> 1



百分之95的值



概率图和分位数图需要数据满足正态分布

esri China BEIJING

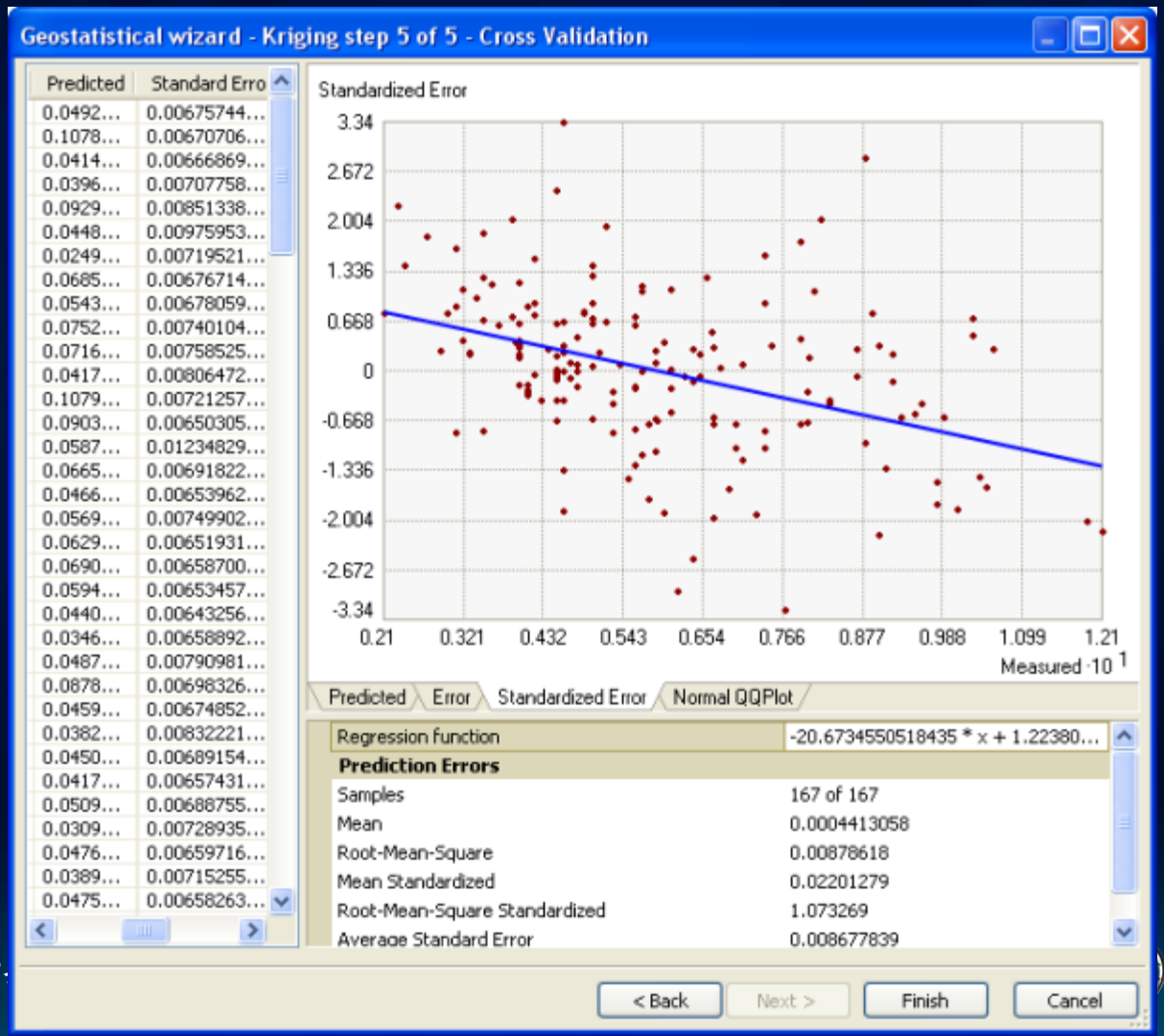
交叉验证

预测表面的准确性?



交叉验证

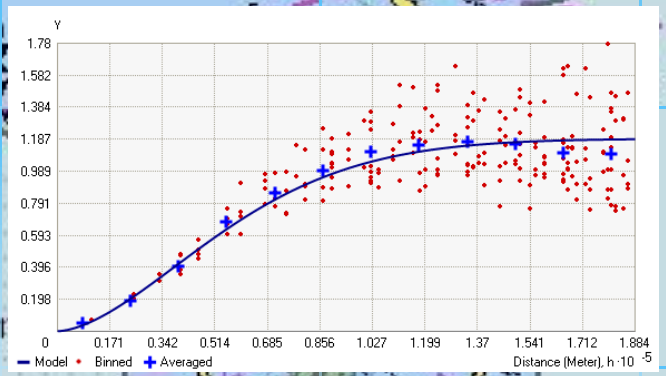
- 地统计提供了测量值与预测值的散点图和统计信息。



esri China BEIJING

演示：克里金插值

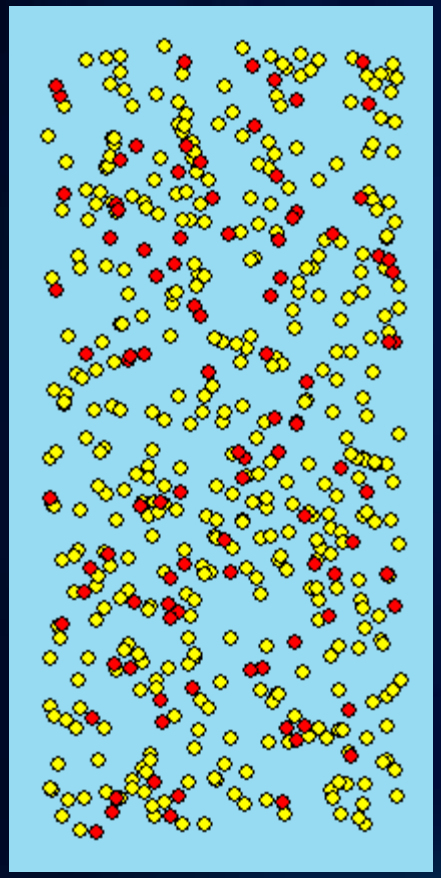




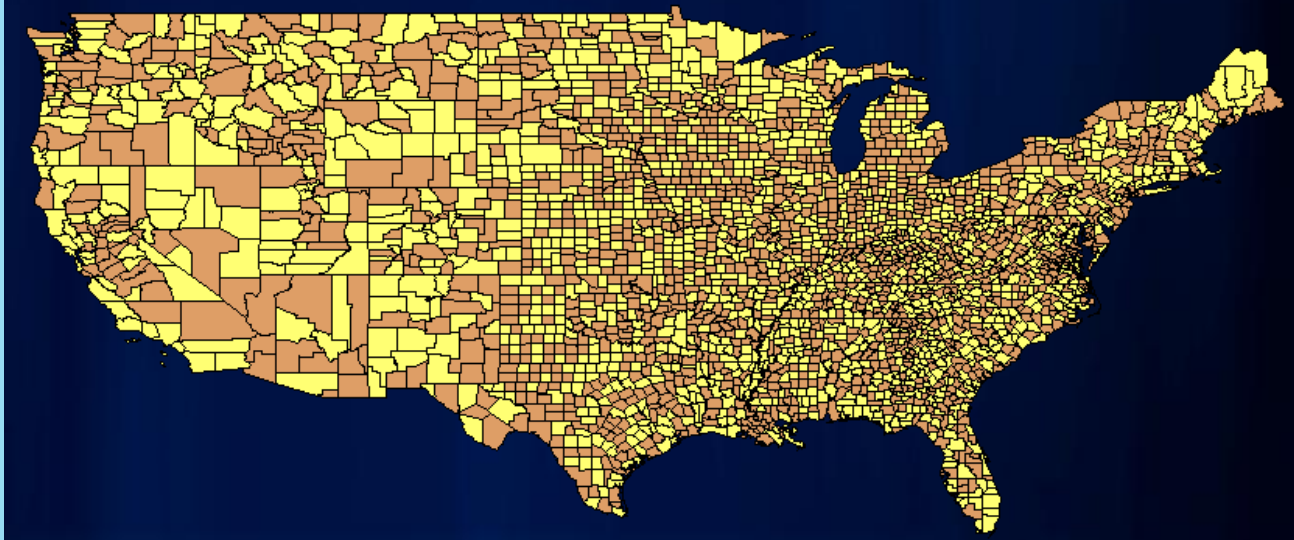
Optimize

Help

子类要素



…通过从一个在[0,1]区间均匀分布的随机值来分割数据



地统计分析GP工具





高斯地统计模拟（GGS）

9.3版本后推出的新特性

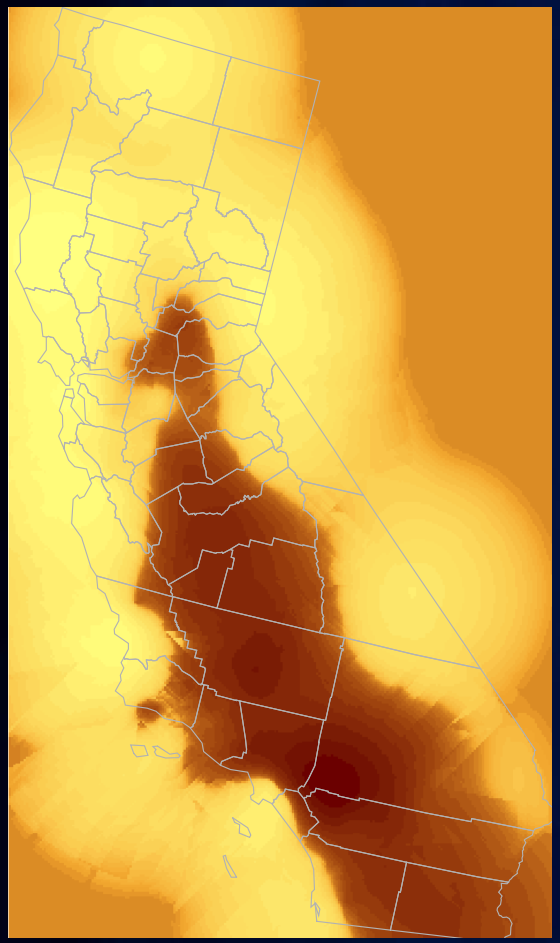
- 一般插值方法将在一个位置上生成一个预测值。
  - 不考虑每个位置上的变化
  - 插值结果比真实现象要平滑

- 高斯模拟生成多个的现实，以不同的表现方式表现相同的空间结构
  - 提供了无样本值区域的不确定性.
  - 每一个现实都是模拟的，而不是预测的

基于数据平稳性的假设

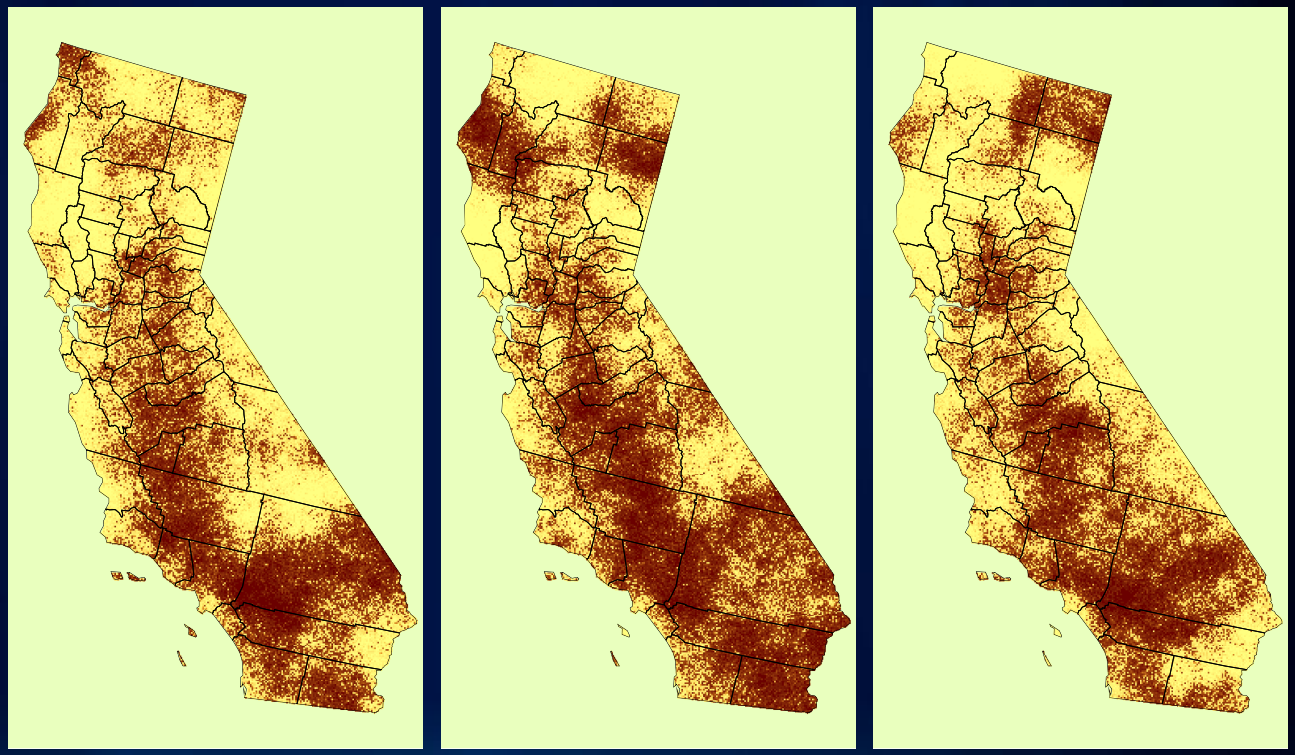
– **均值，方差和空间结构不随着空间位置的变化而变化**

高斯地统计模拟（GGS）



每一个表面都可能是真实的现象

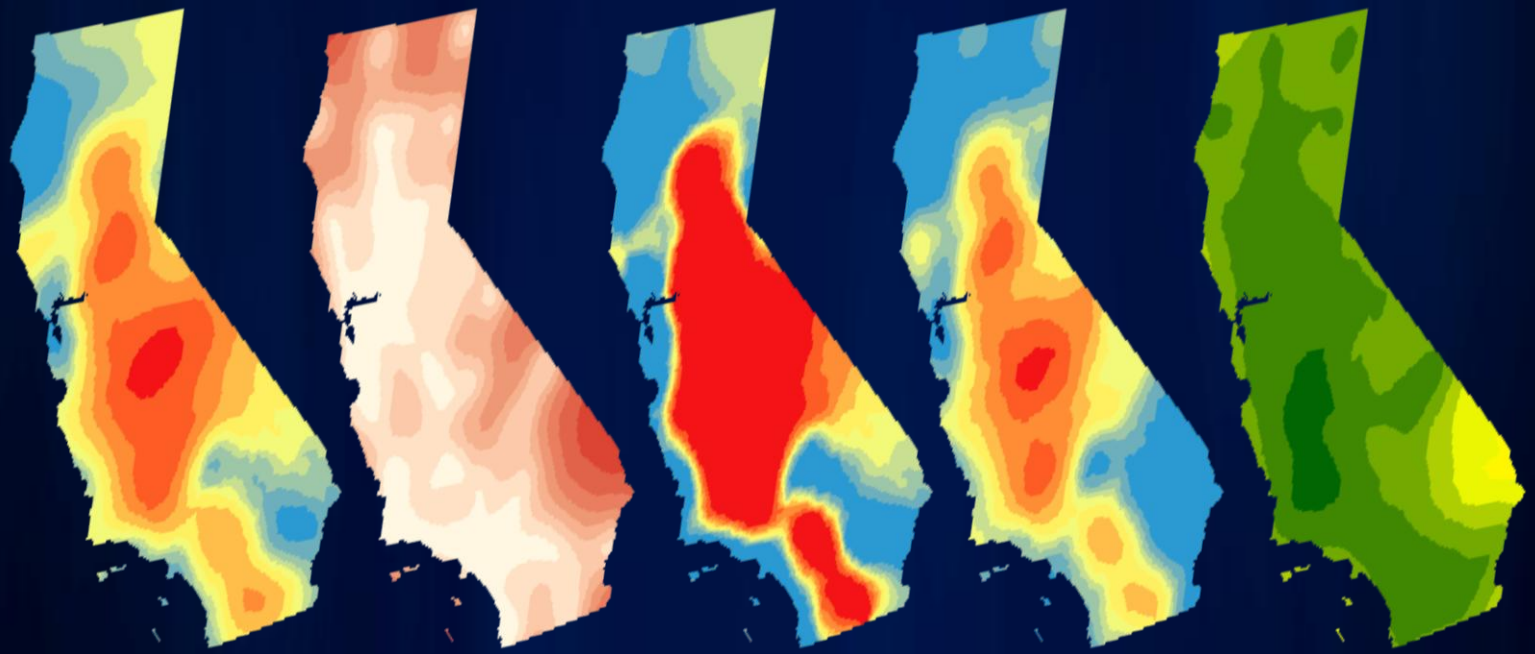
S_0 S_1 S_2



esri China BEIJING

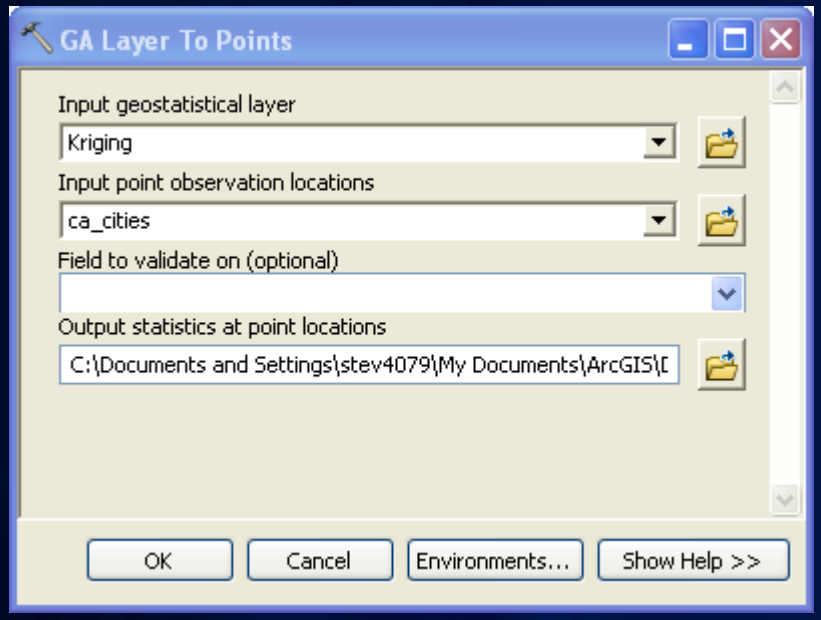
插值结果

- 预测图，预测值标准差，概率图，分位数图，系数图



输出预测结果

- 导出对应位置上的预测值

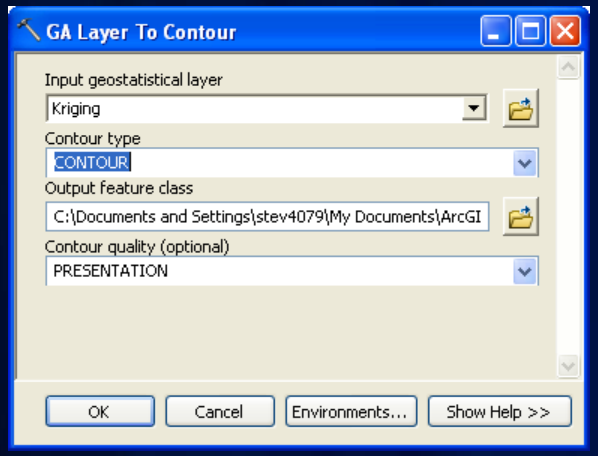


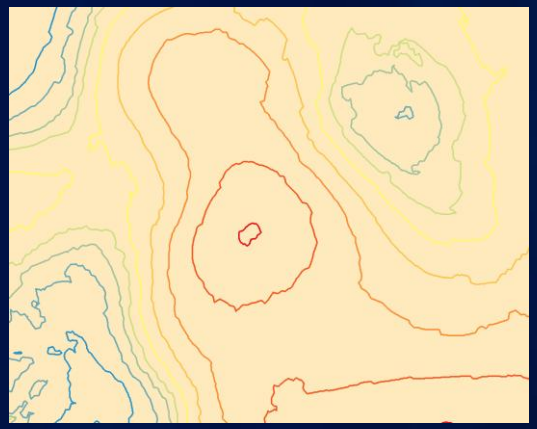
Shape *	AREANAME	Predicted	Standard Error
Point	Acton	0.059379	0.003906
Point	Adelanto	0.05981	0.003798
Point	Agoura Hills	0.045074	0.00336
Point	Alameda	0.030991	0.002774
Point	Alamo	0.043789	0.002526
Point	Albany	0.032968	0.002622
Point	Alhambra	0.043275	0.002134
Point	Aliso Viejo	0.04102	0.003714
Point	Alondra Park	0.035341	0.003856
Point	Alpine	0.052307	0.004391
Point	Altadena	0.048678	0.002699
Point	Alta Sierra	0.085087	0.002836
Point	Alturas	0.050335	0.011539
Point	Amador City	0.086782	0.003456

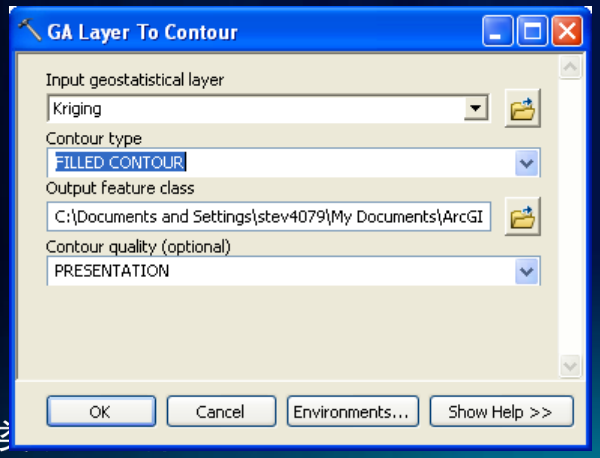
esri China BEIJING

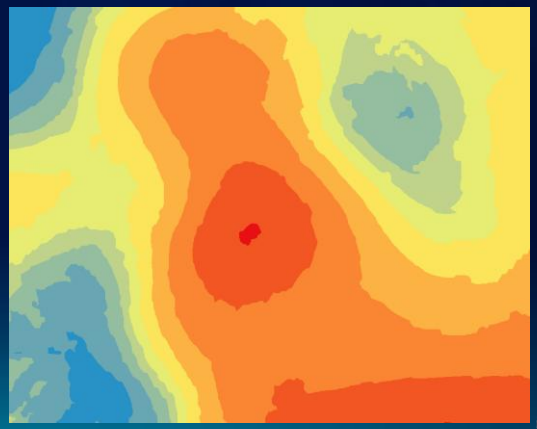
输出预测结果

- 将插值结果导出成等高线



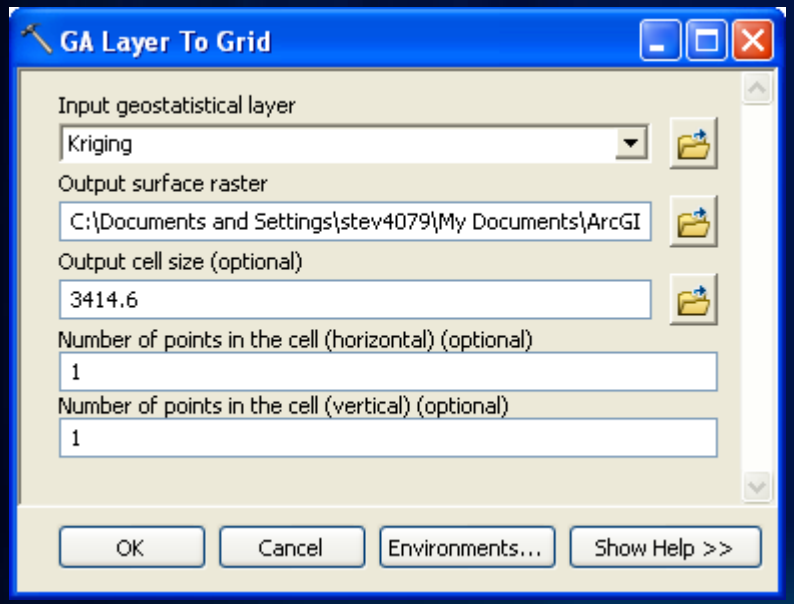


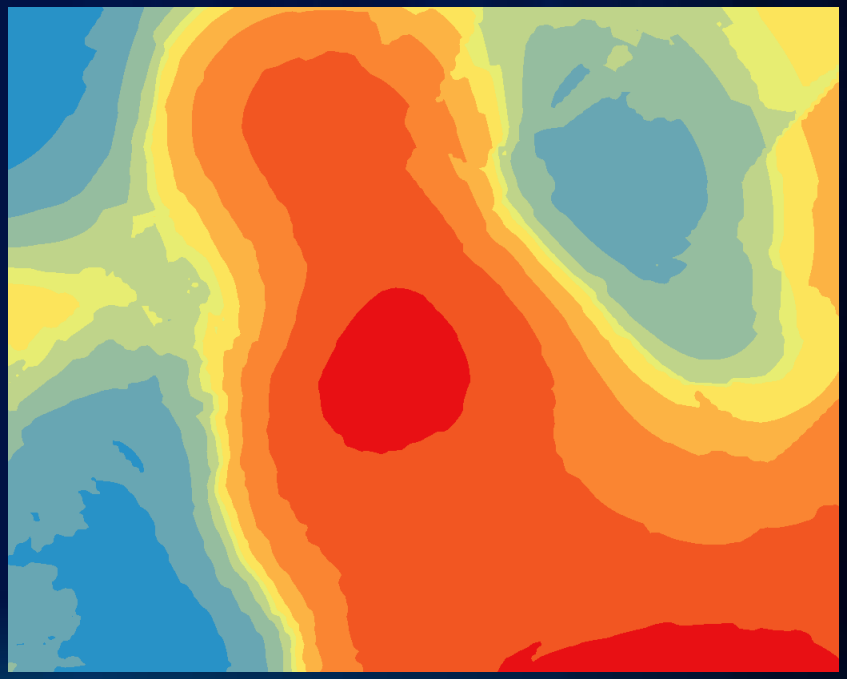




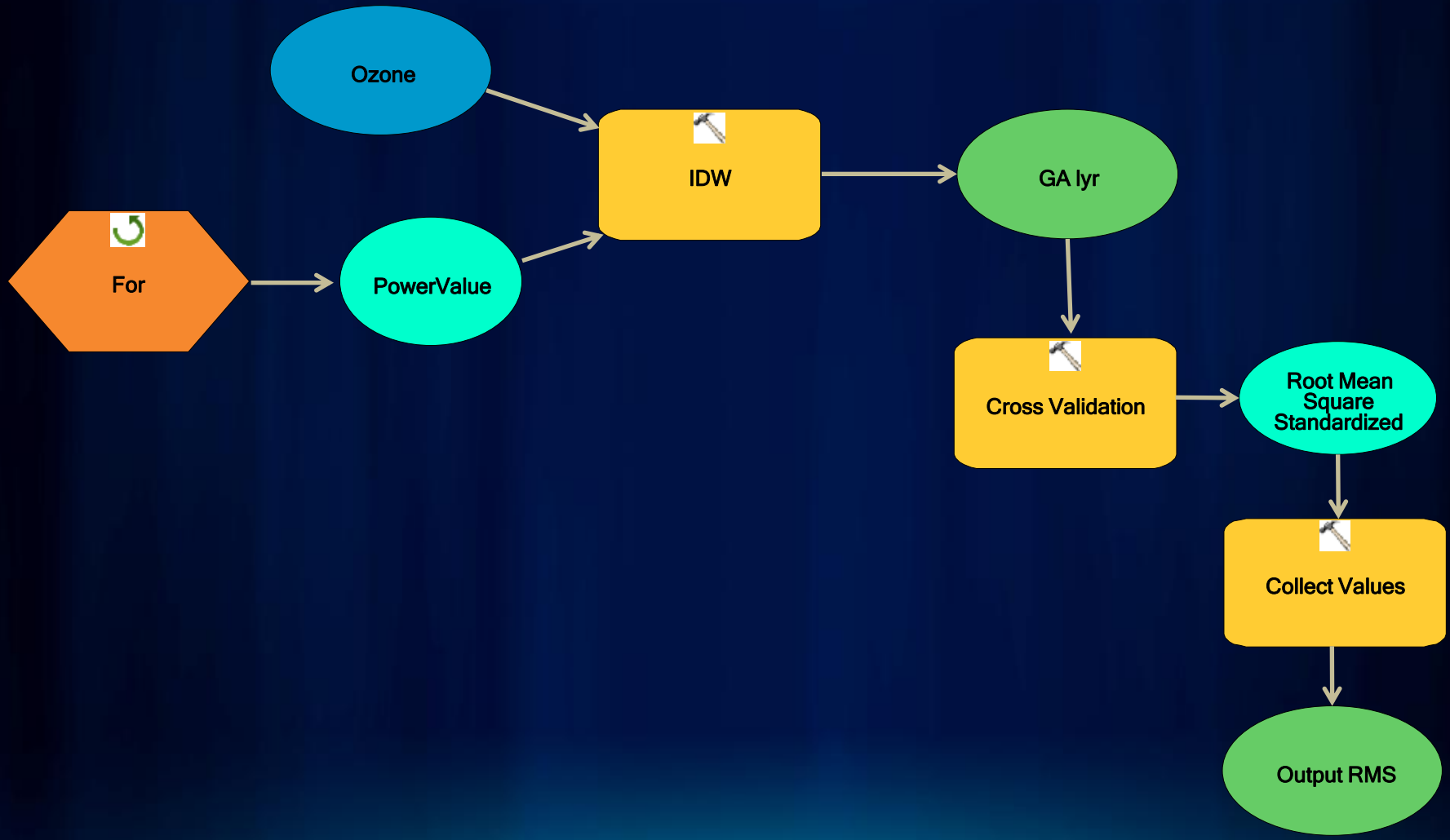
输出预测结果

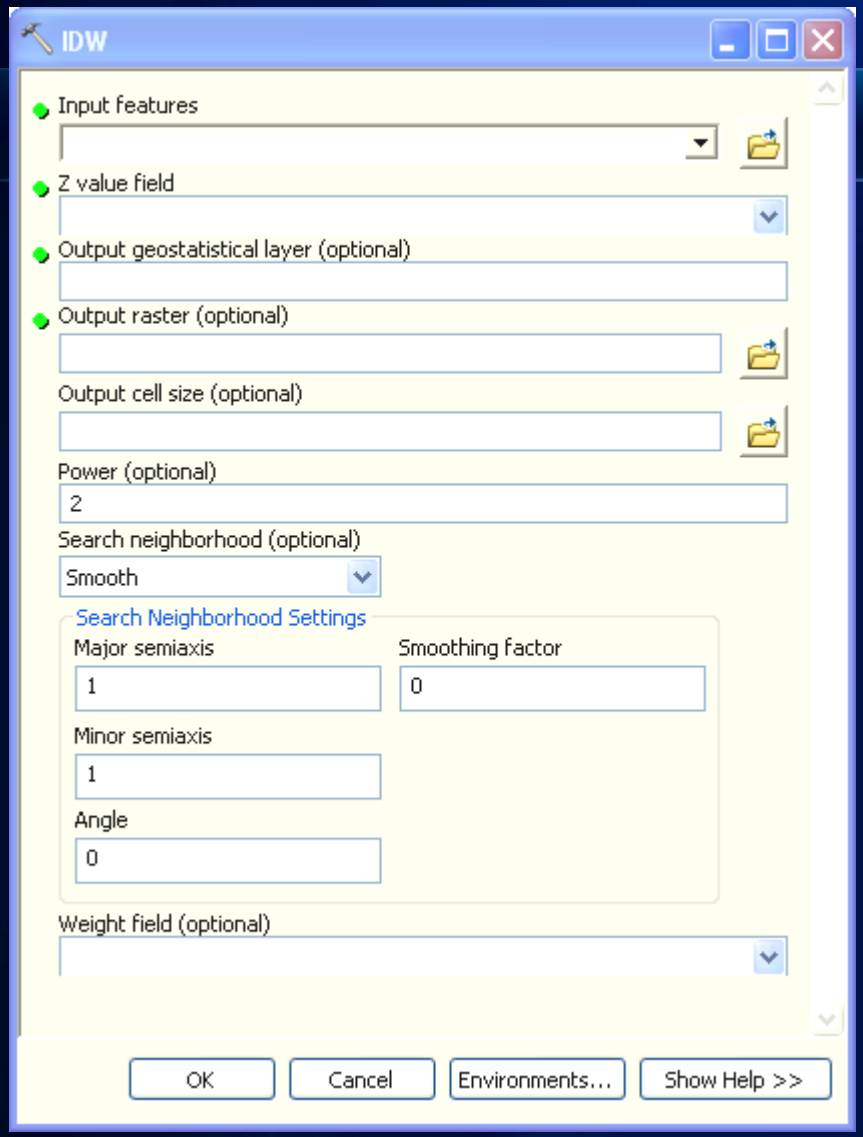
- 将插值图层导出成栅格
  - DEM
  - 趋势面

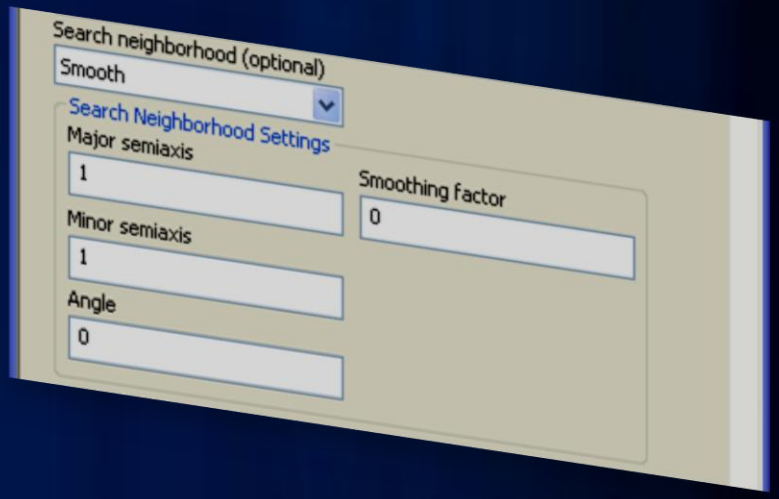




Model Builder



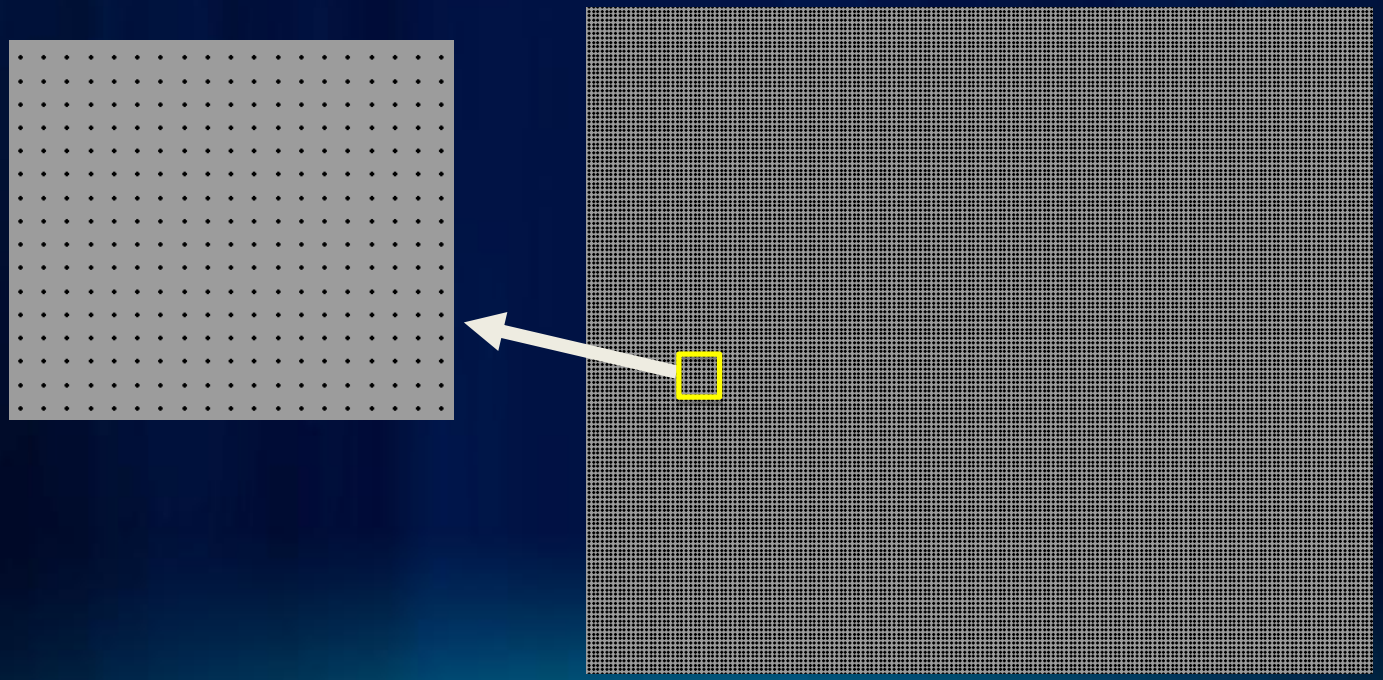




"NBRTYPE=Smooth S_MAJOR=1 S_MINOR=1 ANGLE=0 SMOOTH_FACTOR=0"

大数据量的样本情况

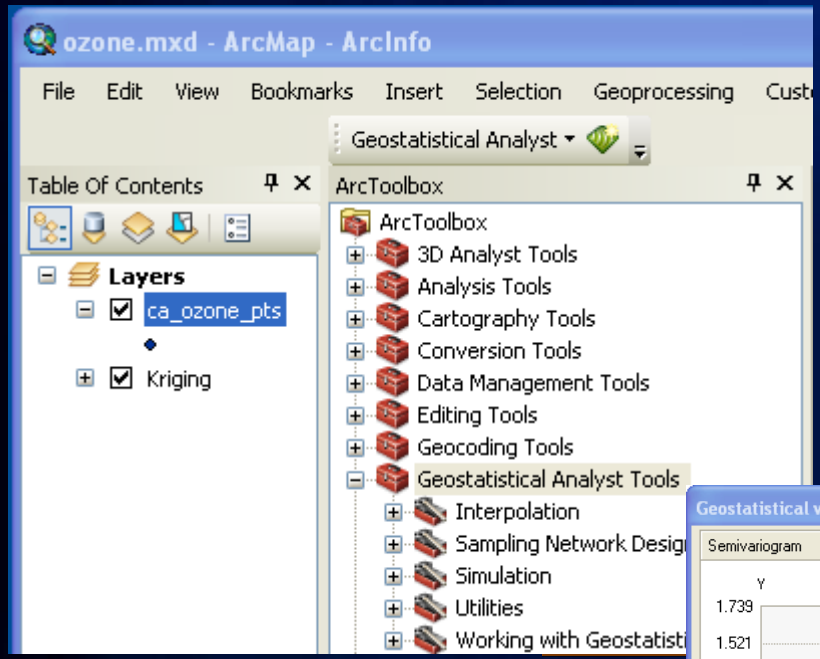
- 激光遥感数据插值成DEM
  - 海量样本点
  - 高精度DEM

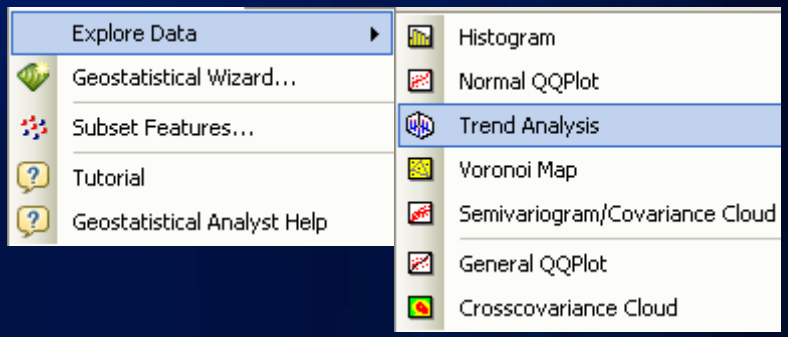


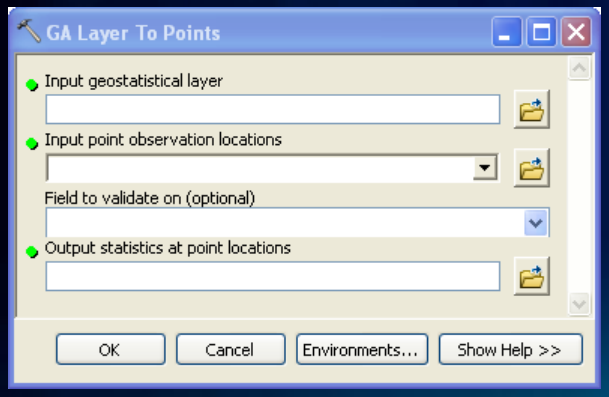
ArcGIS 10 地统计新特性

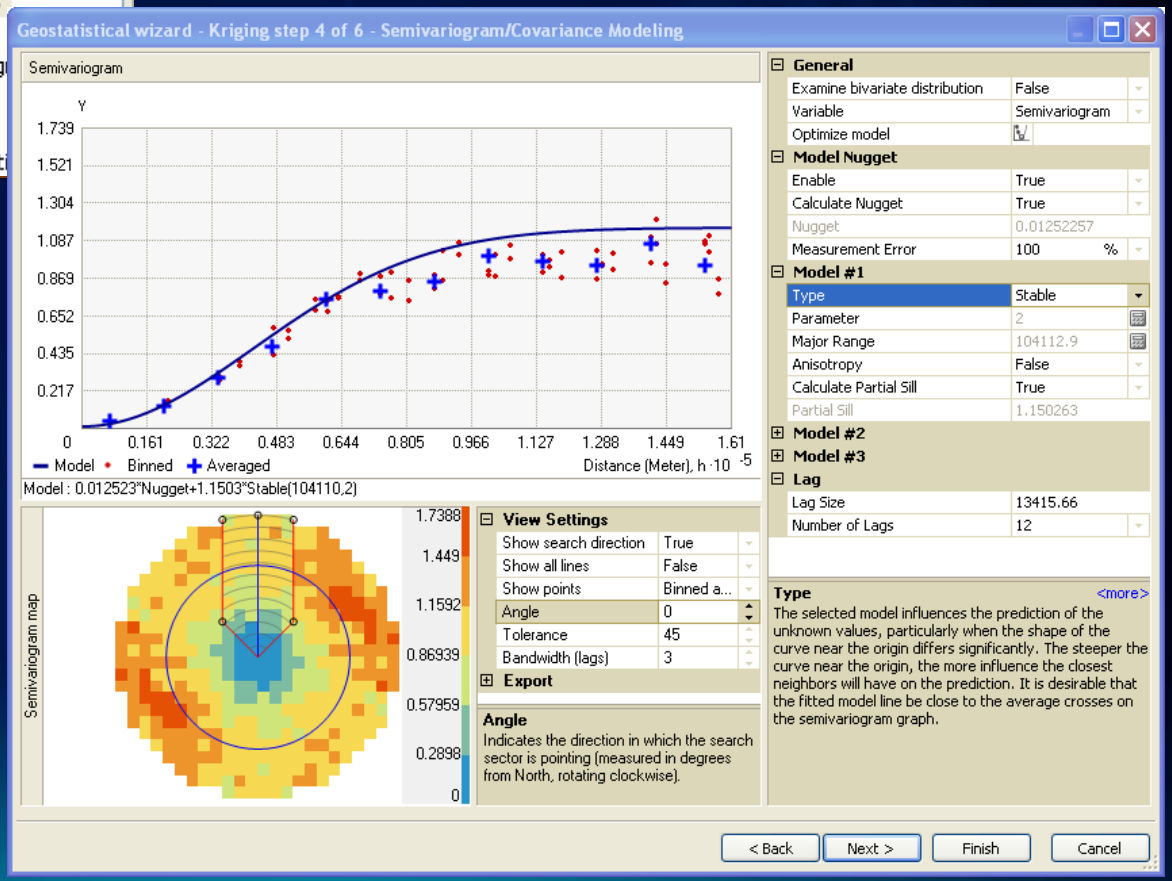


GIS-让人类认知世界


ozone.mxd - ArcMap - ArcInfo
File Edit View Bookmarks Insert Selection Geoprocessing
Geostatistical Analyst
Table Of Contents
ArcToolbox
Layers
ca_ozone_pts
Kriging
ArcToolbox
3D Analyst Tools
Analysis Tools
Cartography Tools
Conversion Tools
Data Management Tools
Editing Tools
Geocoding Tools
Geostatistical Analyst Tools
Interpolation
Sampling Network Desig
Simulation
Utilities
Working with Geostatisti


Explore Data
Geostatistical Wizard...
Subset Features...
Tutorial
Geostatistical Analyst Help
Histogram
Normal QQPlot
Trend Analysis
Voronoi Map
Semivariogram/Covariance Cloud
General QQPlot
Crosscovariance Cloud


GA Layer To Points
Input geostatistical layer
Input point observation locations
Field to validate on (optional)
Output statistics at point locations
OK
Cancel
Environments...
Show Help >>

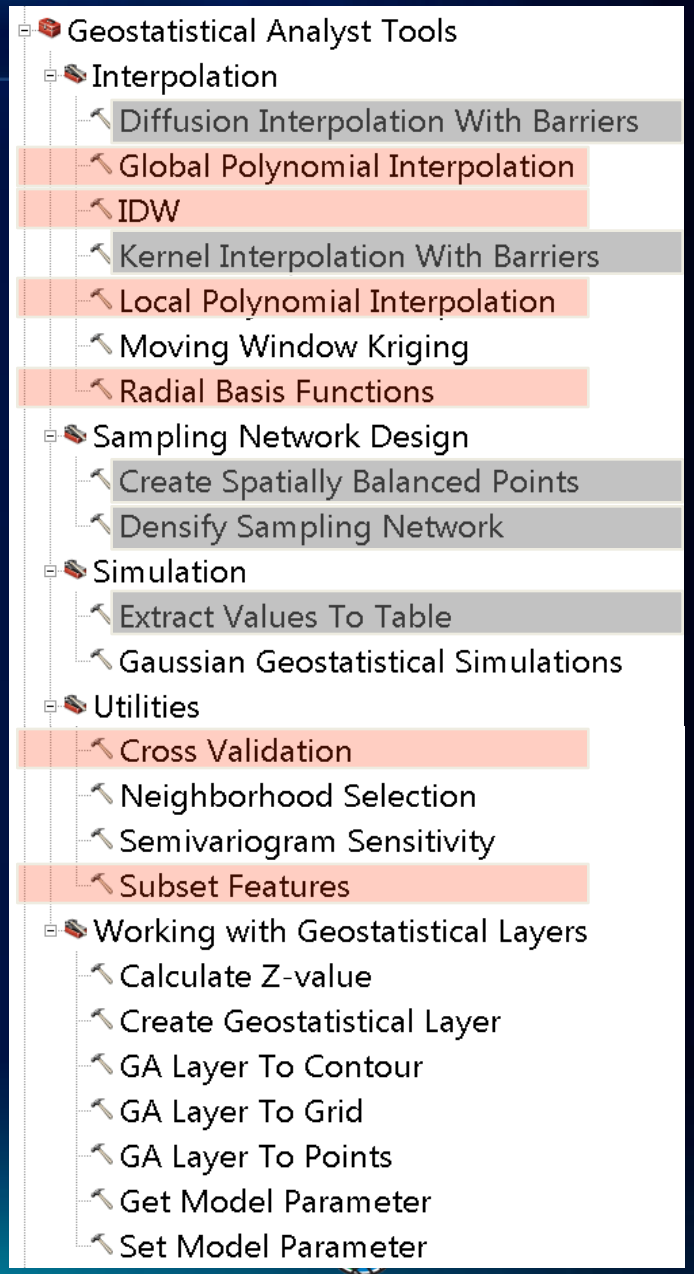

Geostatistical wizard - Kriging step 4 of 6 - Semivariogram/Covariance Modeling
Semivariogram
Y
1.739
1.521
1.304
1.087
0.869
0.652
0.435
0.217
0
0.161
0.322
0.483
0.644
0.805
0.966
1.127
1.288
1.449
1.61
Model
Binned
Averaged
Distance (Meter), h
Model : 0.012523*Nugget+1.1503*Stable(104110,2)
Semivariogram map
1.7388
1.449
1.1592
0.86939
0.57959
0.2898
0
View Settings
Show search direction
Show all lines
Show points
Angle
Tolerance
Bandwidth (lags)
Export
Angle
Indicates the direction in which the search sector is pointing (measured in degrees from North, rotating clockwise).
General
Examine bivariate distribution
Variable
Optimize model
Model Nugget
Enable
Calculate Nugget
Nugget
Measurement Error
Model #1
Type
Parameter
Major Range
Anisotropy
Calculate Partial Sill
Partial Sill
Model #2
Model #3
Lag
Lag Size
Number of Lags
Type
<more>
The selected model influences the prediction of the unknown values, particularly when the shape of the curve near the origin differs significantly. The steeper the curve near the origin, the more influence the closest neighbors will have on the prediction. It is desirable that the fitted model line be close to the average crosses on the semivariogram graph.
< Back
Next >
Finish
Cancel

ArcGIS 10的新特性

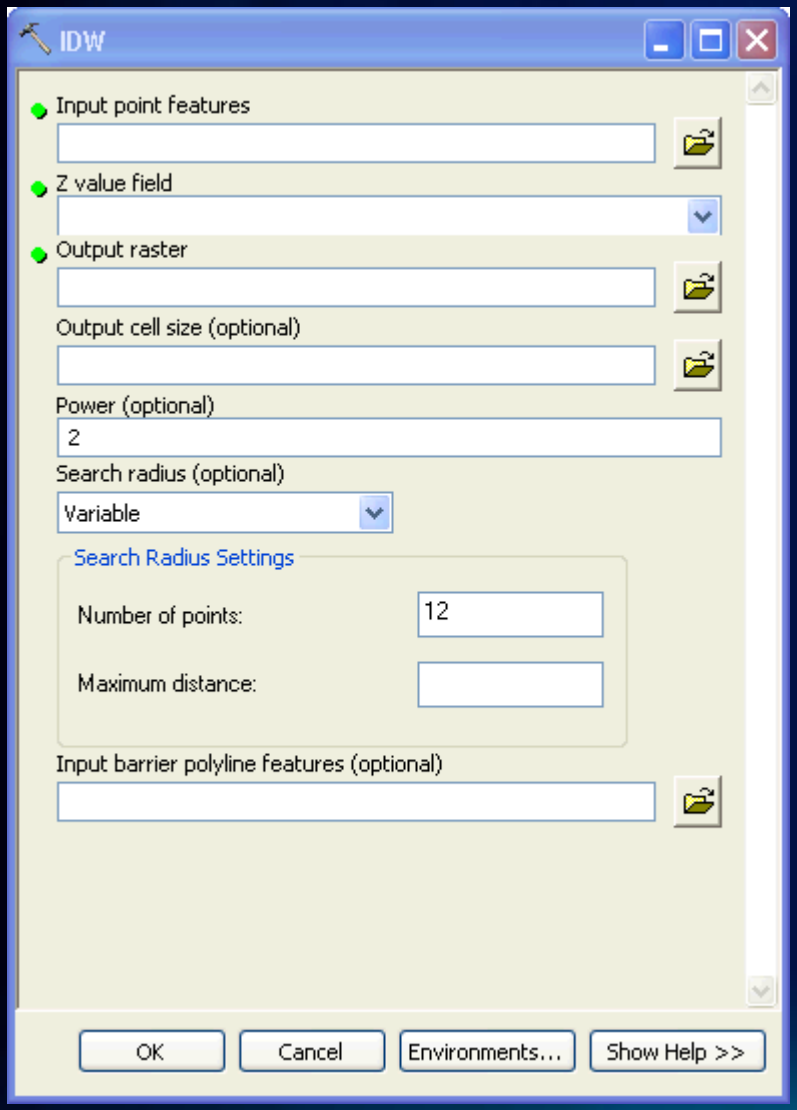
- 地统计分析向导的增强
  - 面板及窗口的自由缩放
  - 新的对话框布局和功能
  - 模型参数的优化
  - 变异函数图中包含平均 值
  - 提供参数的帮助
    - 对话框上查看
    - 弹出帮助

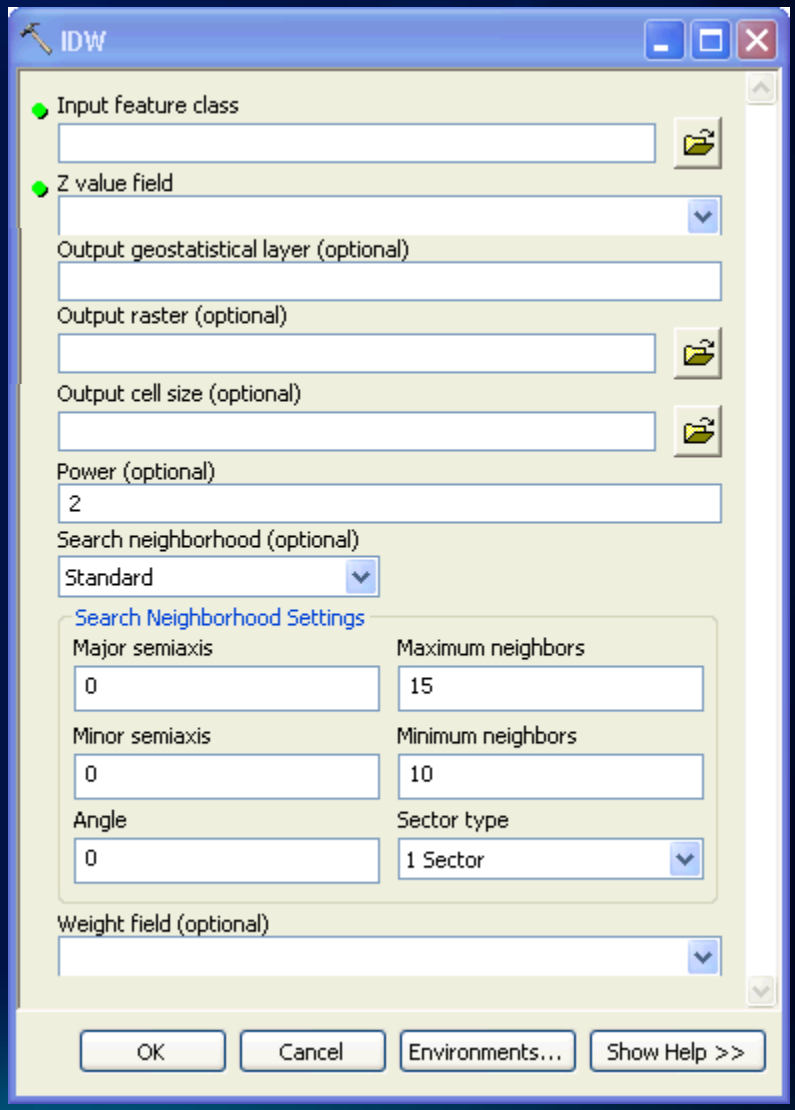
ArcGIS 10的新特性

- 在地统计分析工具箱中，新增了11个GP工具
  - 新增功能
  - 之前版本中仅限地统计分析向导或者地统计分析工具条



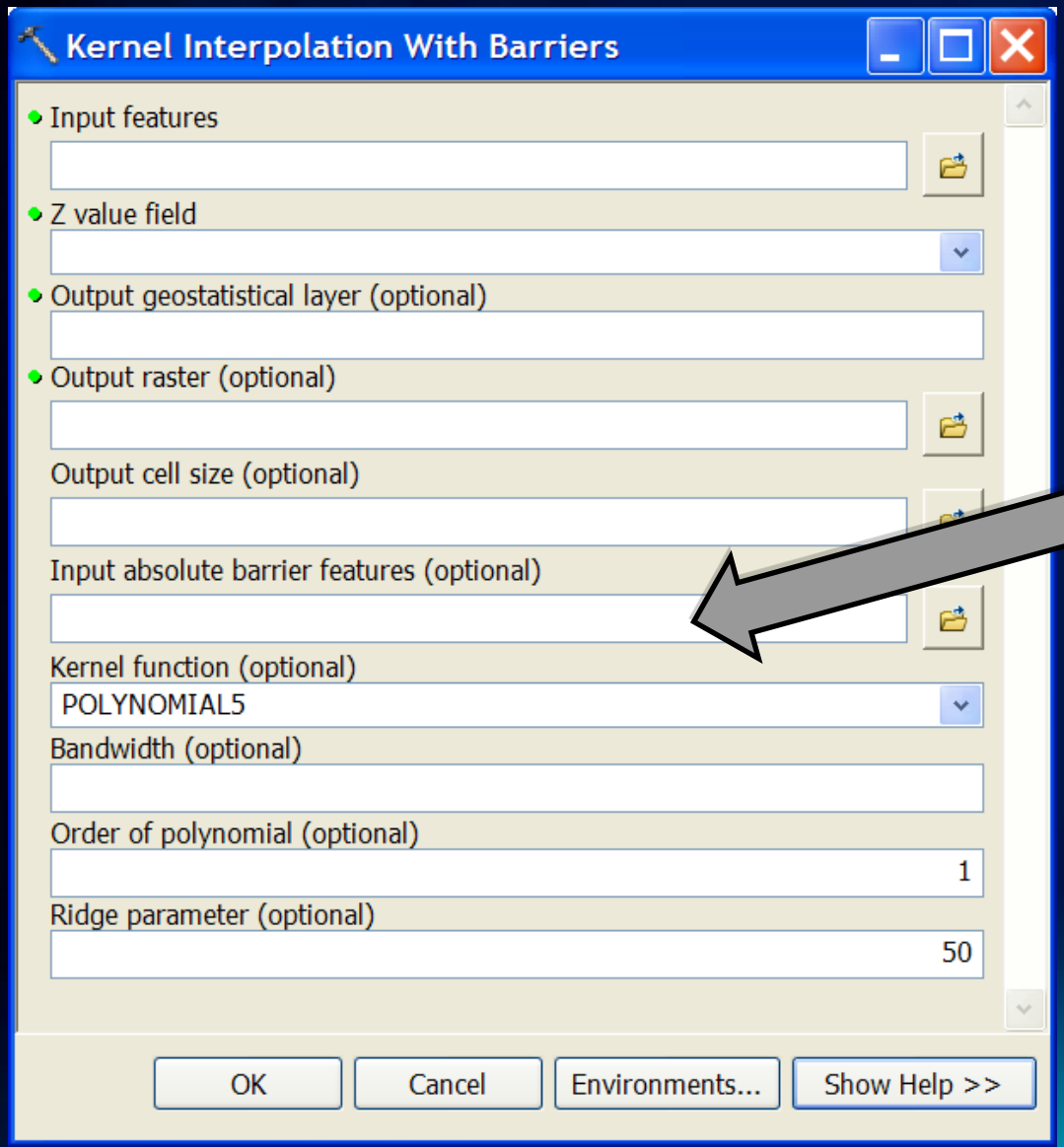
地统计分析中的IDW工具

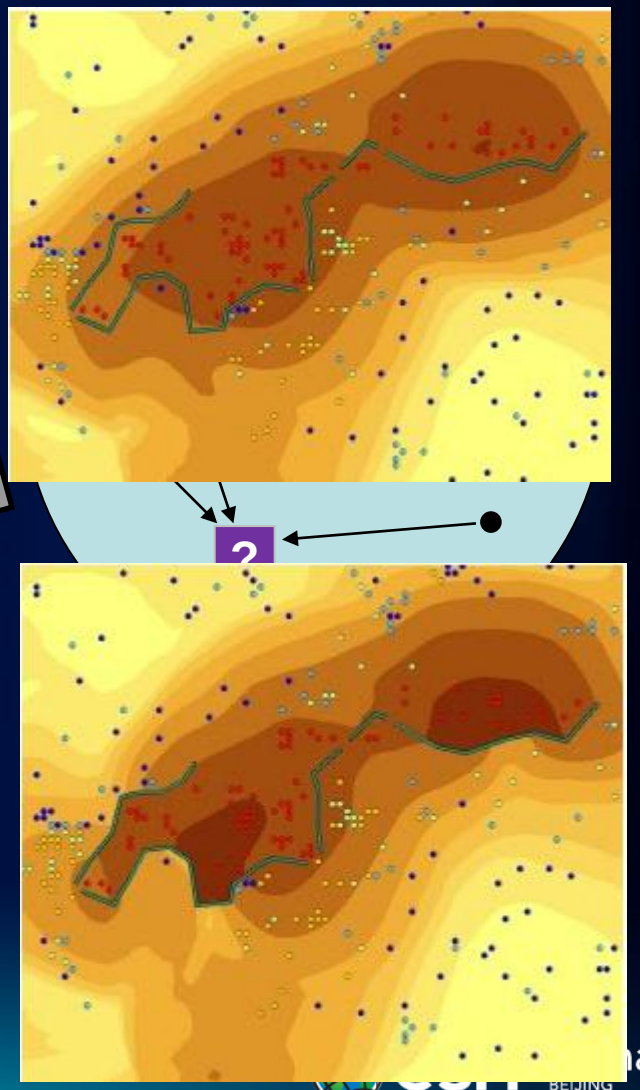




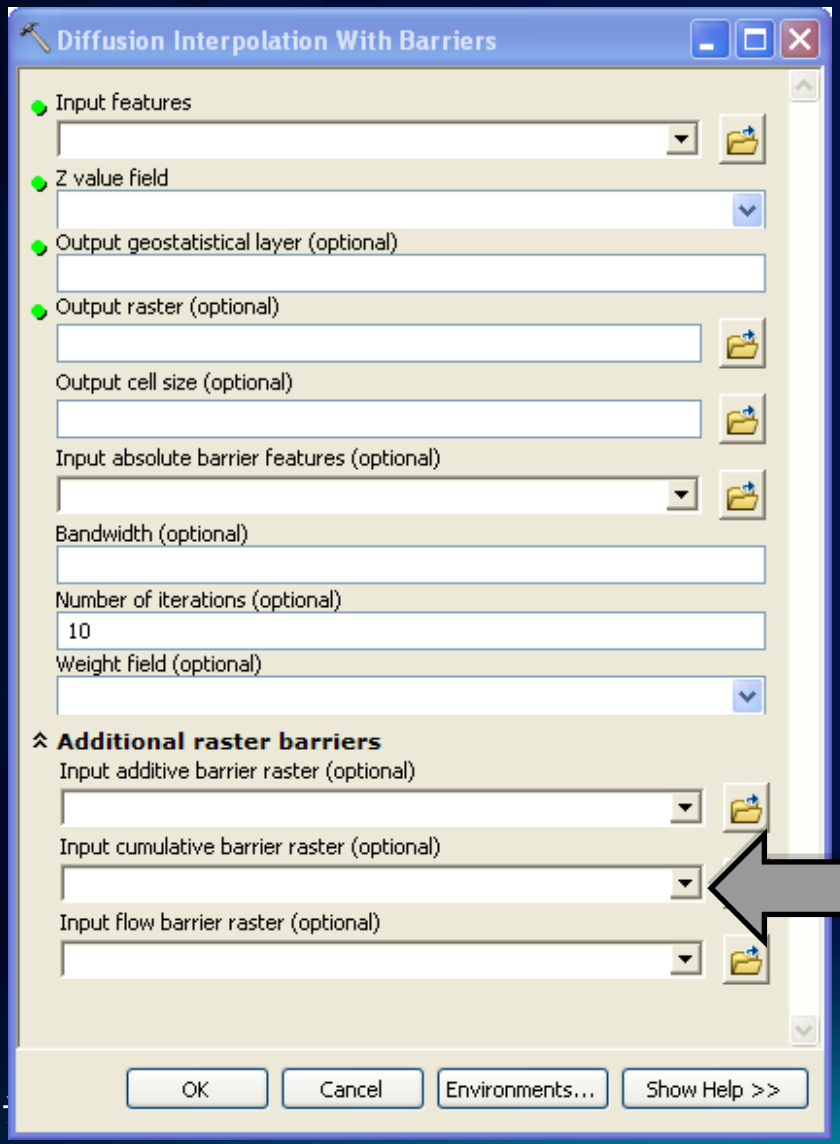
esri China BEIJING

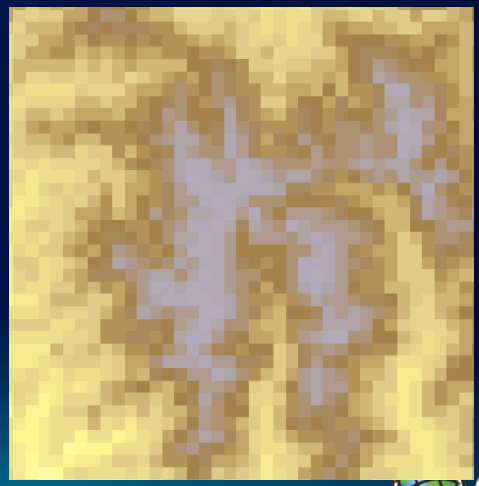
含障碍的核插值


Kernel Interpolation With Barriers
Input features
Z value field
Output geostatistical layer (optional)
Output raster (optional)
Output cell size (optional)
Input absolute barrier features (optional)
Kernel function (optional)
POLYNOMIAL5
Bandwidth (optional)
Order of polynomial (optional)
1
Ridge parameter (optional)
50
OK
Cancel
Environments...
Show Help >>



含障碍的扩散插值





esri China BEIJING

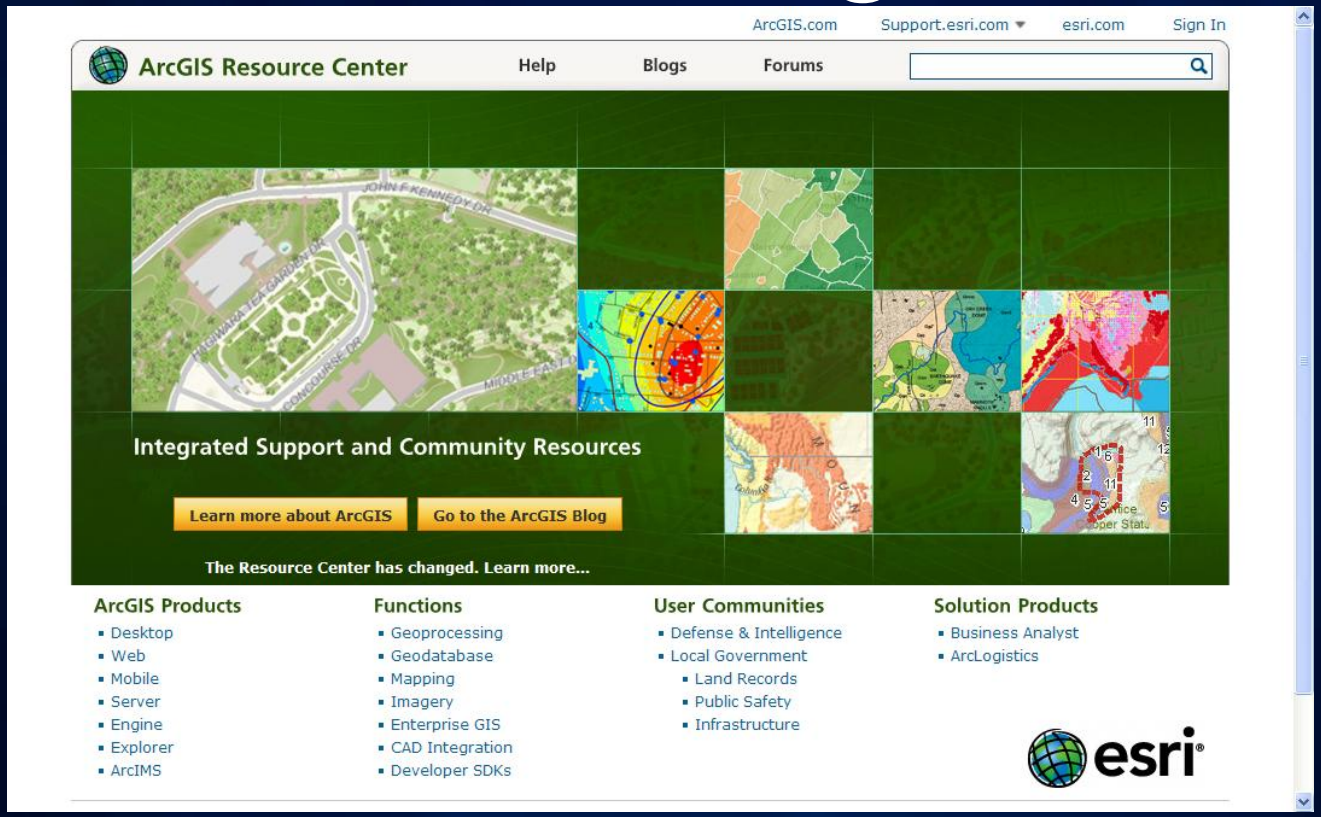
采样网络设计



- 创建空间平衡点
  - 基于预先得到的概率结果
  - 输出的结果样本点是空间平衡的
    - 样本点的空间依赖性被最小化
- 增密采样网络
  - 基于预定义的地统计克里金图层
  - 使用预测表明的标准差来决定哪个位置上的样本需要保留，哪一个需要被移除。

相关资料

resources.arcgis.com



精彩预告

- 什么是Geometric Network？
- 如何使用Geometric Network对目标网络进行分析？

石羽 ----- “ArcGIS 设施网络分析”

谢谢！