

ArcGIS 地统计分析实习指导书

实习内容：

- 练习 1：使用缺省参数创建一个表面
- 练习 2：数据检查
- 练习 3：制作臭氧浓度图
- 练习 4：模型比较
- 练习 5：制作超出某一临界值的臭氧概率图
- 练习 6：生成最终成果图

预备知识：

利用地统计分析模块，你可以根据一个点要素层中已测定 采样点、栅格层或者利用多边形质心，轻而易举地生成一个连续表面。这些采样点的值可以是海拔高度、地下水位的深度或者污染值浓度等。当与 ArcMap 一起使用时，地统计分析模块提供了一整套创建表面的工具，这些表面能够用来可视化、分析及理解各种空间现象。

美国环保局负责对加利福尼亚州的大气臭氧浓度进行监测。臭氧浓度值是通过遍布全州的监测站来测定的。已经知道所有监测站的臭氧值浓度，但是我们还想知道加利福尼亚州其他任一地方的臭氧浓度值，但是考虑到费用以及实用性问题，我们不可能在任何地方都建立监测站。地统计分析模块提供了许多工具，通过检测所有采样点之间的关系，生成一个关于臭氧浓度值、预测标差（不确定性）以及超出临界值的概率的连续表面，从而使对其他点的浓度值进行最佳预测成为可能。

在地统计分析模块的安装盘上，附有本次联系所需要的数据。这些数据包括：

数据集	描述
Ca_outline	加州轮廓图
Ca_ozone_pts	臭氧采样点数据（单位：ppm）
Ca_cities	加州主要城市位置图
Ca_hillshade	加州山体阴影图

臭氧数据集（Ca_ozone_pts）表示的是 1996 年中每八个小时时段内的臭氧平均浓度的最大值，以 ppm 为单位。根据这些臭氧采样点的测量值，你可以生成两个连续表面（或地图），以现有的采样点数据为基础来预测加州任意地方的臭氧浓度值。创建第一张图时你只需要简单地使用缺省选项即可，你会看到根据采样点数据生成表面是多么容易的事情。生成第二张图时，你可以较多的考虑采样点之间存在的空间关系，这时，你将要用到 ESDA（空间数据探索分析）工具来检查你的数据。你还会学习到一些地统计选项，利用这些选项，你可以创建诸如剔除趋势并且模拟空间自相关的表面。利用 ESDA 工具，通过对各种地统计参数的操作，你可以创建出更加精彩的表面。

很多时候，我们关心的并不是那些有损于健康的风险的实际值，而是它是否超出某一有毒水平，如果确实超出，必须采取行动。你创建的第三个表面将用来评估臭氧浓度超过临界值的概率。

对于本次练习，临界值规定如下：如果最大值超过 0.12ppm，那么这个地方就应该被严密监测。利用这个标准，你可以利用地统计分析模块来预测臭氧浓度超出临界值的概率大小。本练习被分为各个单独的任务，你可以根据自己的学习进度来逐步熟悉地统计分析模块各项功能。

练习 1 带你进入地统计分析模块，通过创建臭氧浓度表面的过程，你会发现使用缺省参数创建表面是一件很容易的事情。

练习 2 教你创建表面之前如何对数据进行检查。数据检查的目的是为了找出数据中那些

离群值并且发现数据中存在的趋势。

练习 3 你将创建第二个表面,这个表面更多地考虑了练习 2 中数据分析发现的空间关系,并且对练习 1 中生成的表面进行了改进。该练习还将向你介绍地统计学中的一些基本概念。

练习 4 教你如何对练习 1 和练习 3 中创建的表面进行比较,并判断哪个表面对未知值的预测更好。

练习 5 指导你创建臭氧浓度超出临界值的概率图,从而生成第三个表面。

练习 6 告诉你怎样利用 ArcMap 的功能将你在练习 3 和练习 5 中创建的表面放在一起做最终的显示。

你可能需要几个小时的时间集中精力来完成本次练习。当然如果你愿意,你也可以每次完成一个练习,在每次完成一个练习之后,把你的练习数据保存起来。

练习 1：利用缺省参数创建一个表面

在开始练习之前,首先启动 ArcMap 并激活地统计分析模块。

1.1 启动 ArcMap 并激活地统计分析模块

单击窗口任务栏的 Start 按钮,光标指向 Programs,再指向 ArcGIS,然后单击 ArcMap。在 ArcMap 中,单击 Tools,在单击 Extensions,选中 Geostatistical Analyst 复选框,单击 Close 按钮。

1.2 添加 Geostatistical Analyst 工具条到 ArcMap 中。

单击 View 菜单,光标指向 Toolbars,然后单击 Geostatistical Analyst。

1.3 在 ArcMap 中添加 数据层

一旦数据加入后,你就能利用 ArcMap 来显示数据,而且如果需要,你还可以改变没一层的属性设置(如符号等等)

1. 单击 Standard 工具条上的 Add Data 按钮。

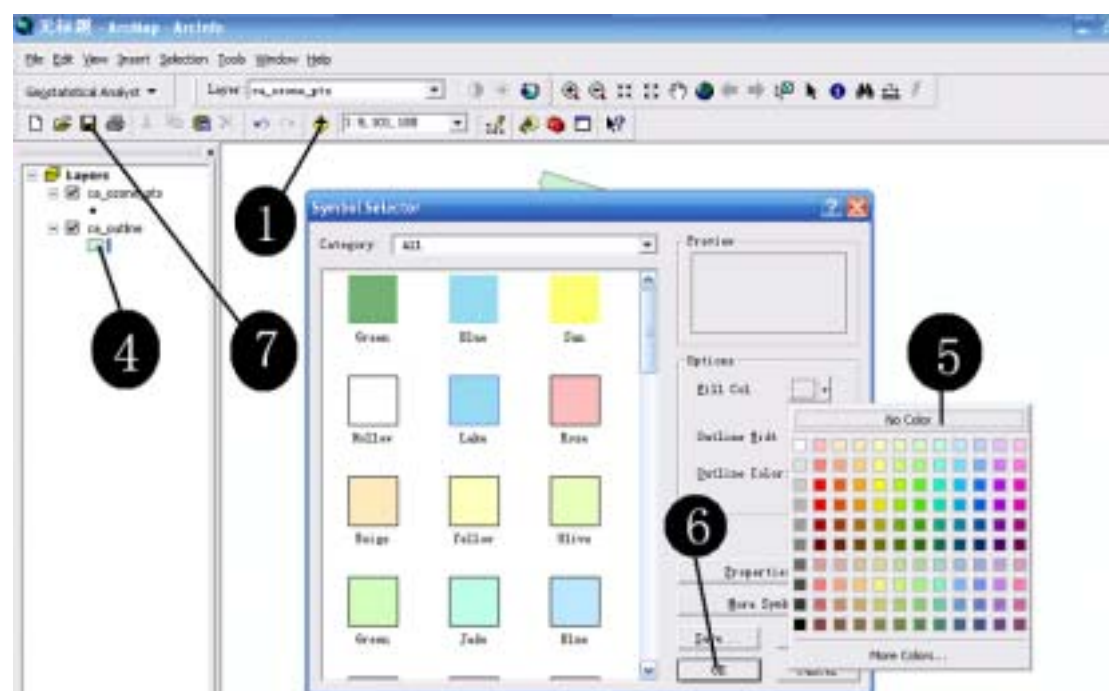
2. 找到你安装练习数据的文件夹(缺省安装路径是 C:\ArcGIS\ArcTutor\Geostatistics),按住 Ctrl 键,然后点击并高亮显示 Ca_ozone_pts 和 ca_outline 数据集。

3. 单击 Add 按钮。

4. 单击目录表中的 ca_outline 图层的图例,打开 Symbol Selector 对话框。

5. 单击 Fill Color 下拉箭头,然后单击 No Color。

6. 在 Symbol Selector 对话框中单击 OK 按钮。



现在 ca_outline 图层酒杯透明现实，只有轮廓可见，以便让你看见在在后面所要创建的其他图层。

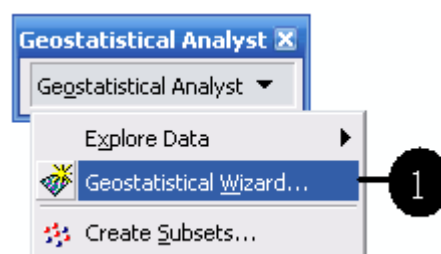
建议你在每个练习之后都要讲将地图保存。

7. 点击 Standard 工具条上的 Save 按钮。新建一个本地工作目录（如 C:\geostatistical），定位到本地工作目录。因为这是你第一次保存地图，所以你需要为它取个名字。（建议你将地图命名为 Ozone Prediction Map.mxd）。将来再点保存时，只需要点击 Save 按钮即可。

1.4 利用缺省值创建表面

下一步你将利用地统计分析模块的缺省设置来创建一个臭氧浓度表面。选择臭氧点数据集（ca_ozone_pts）作为输入数据集，利用普通克里格方法，通过插值即可得到那些未知点的臭氧浓度值。在许多对话框中你可以直接单击下一步按钮，接受缺省参数设置。你不必考虑本练习中的对话框细节，在后续练习中这些对话框你还会接触到。本练习的目的正是利用缺省选项来创建一个表面。

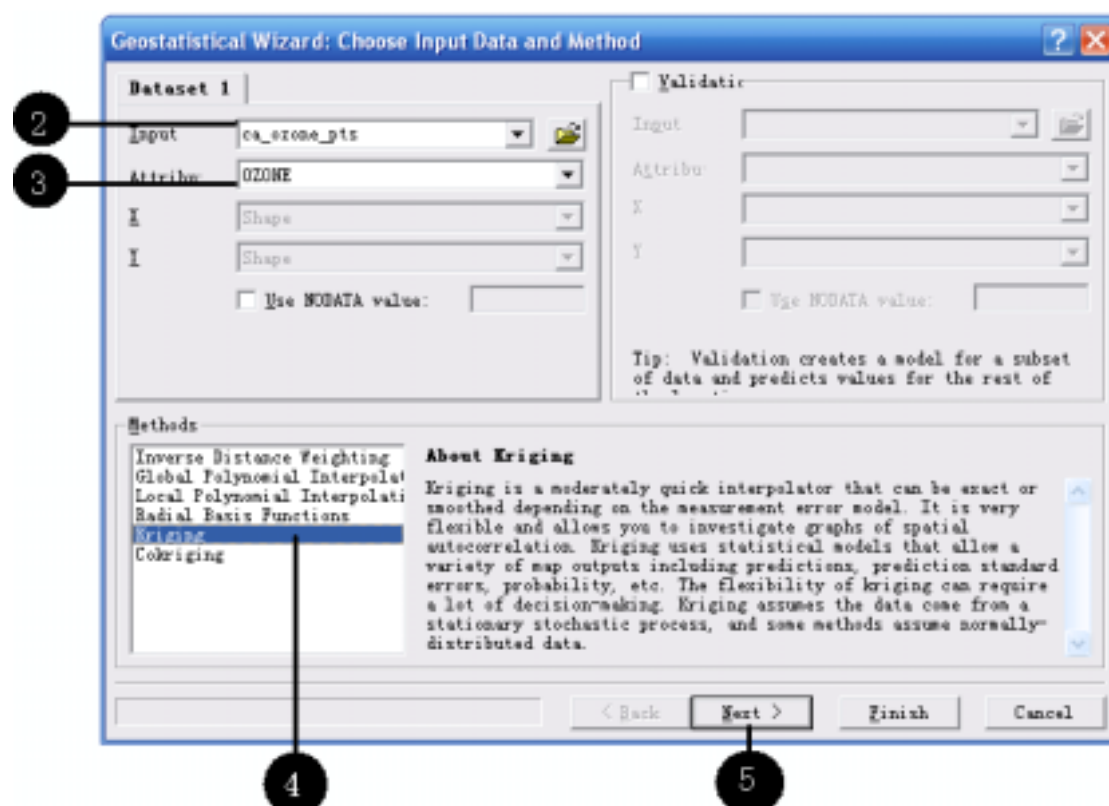
1. 单击 Geostatistical Analyst，然后单击 Geostatistical Wizard。



2. 点击 Input Data 下拉箭头，单击并选中 ca_ozone_pts。
3. 单击 Attribute 下拉框箭头，单击并选中属性 OZONE。
4. 在 Method 对话框中单击 Kriging。
5. 单击 Next 按钮。

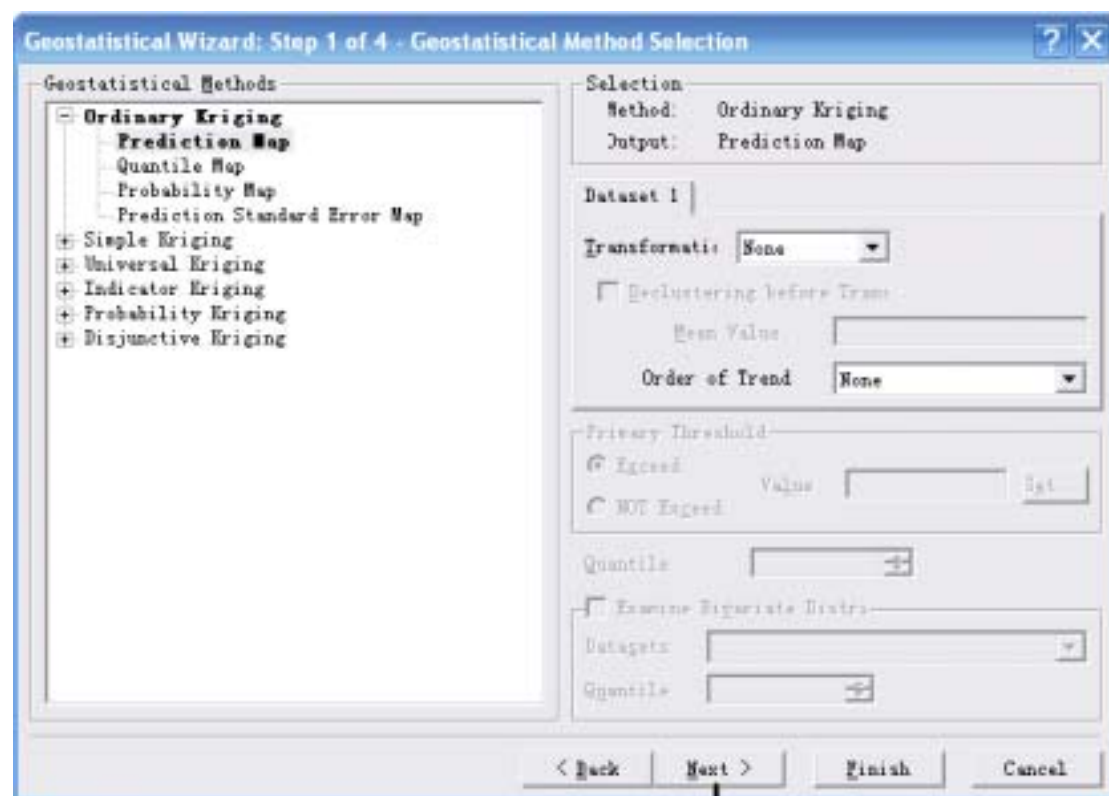
缺省情况下,在 Geostatistical Method Selection 对话框中,Ordinary Kriging 和 Prediction Map 被选中。

注意,在选定了臭氧浓度表面图的生成方法之后,你可以单击 Finish 按钮利用缺省参数来创建一个表面。不过,从第六步到第十步你可以看到许多不同的对话框。

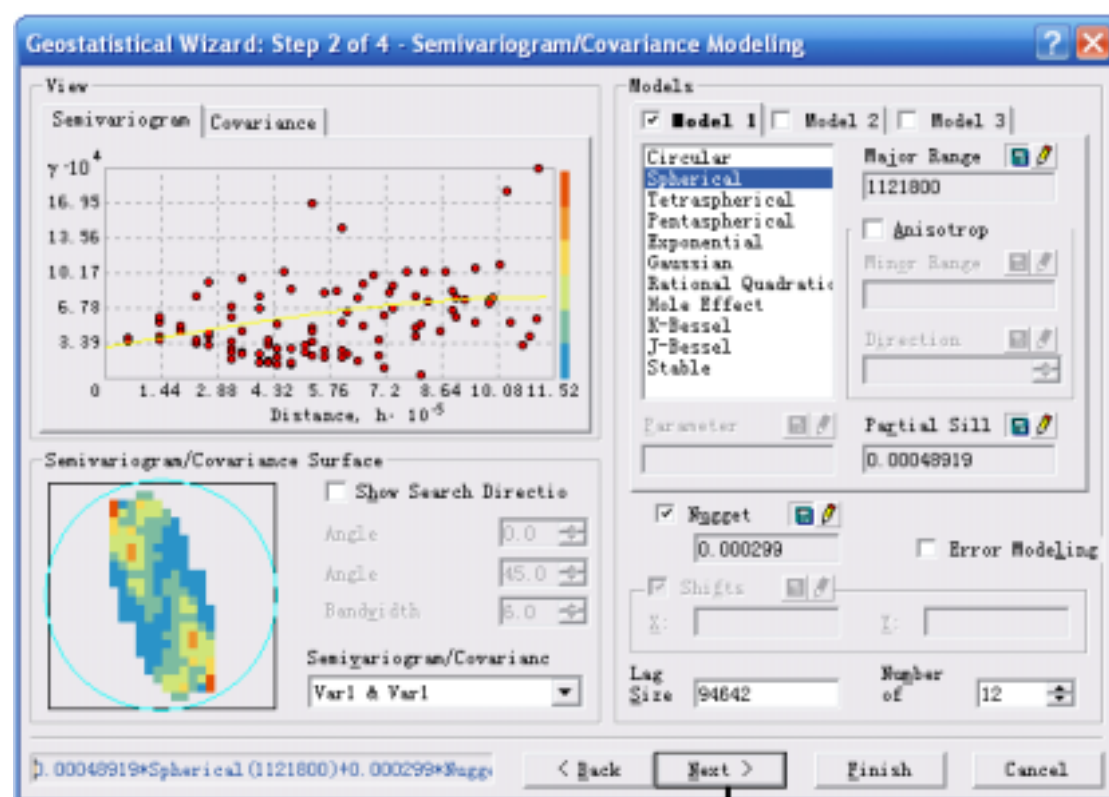


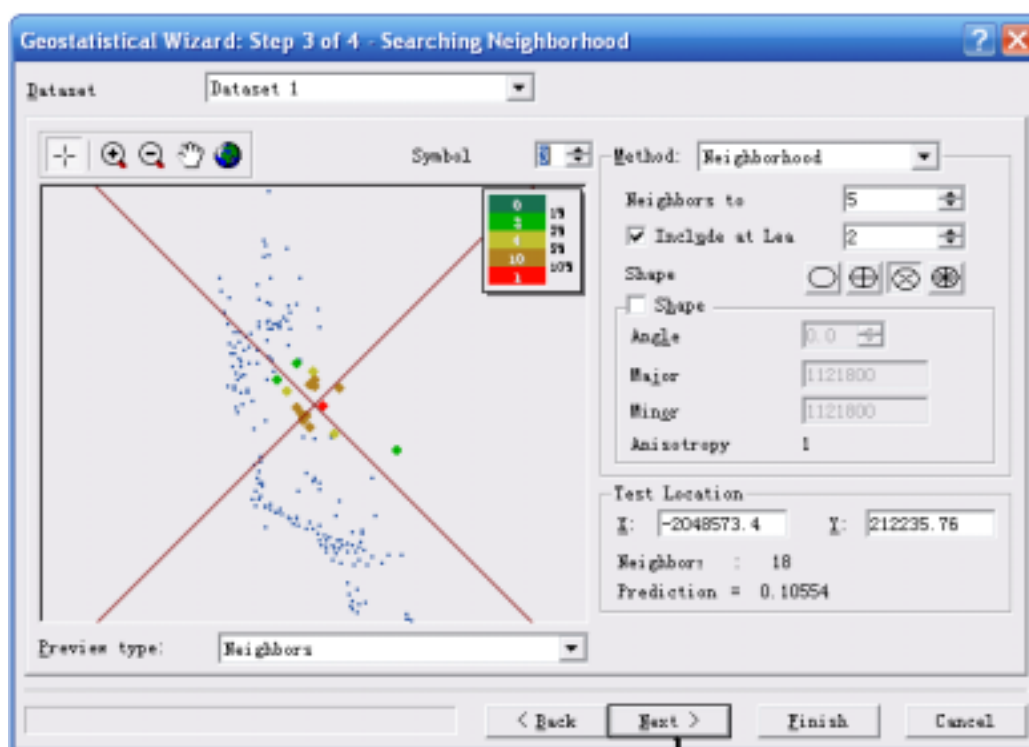
6. 在 Geostatistical Method Selection 对话框中单击 next 按钮。

Semivariogram/covariance modeling 对话框可以让你分析已测数据点之间空间关系,我们知道空间上距离越近的事物越相似,这个假设可以利用半变异函数来验证。在获取空间关系的同时对一个半变异函数模型来进行拟合,此过程称为变异估计。



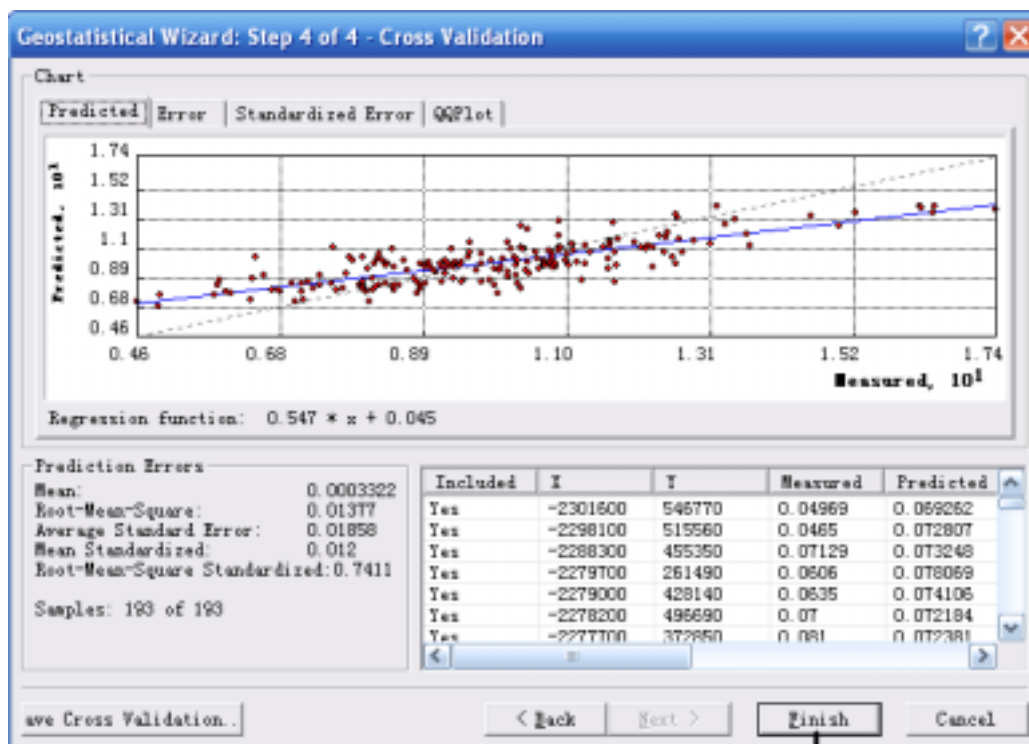
7. 点击 next 按钮。





十字丝处是一个未被测量的点。你可以利用已测点的值来预测十字丝处的值的情况。已测量点距离你需要预测的未知点越近，它们的值也就越相近。在上图中，红色点对未知点的值的影响势必要比绿色点大。根据周围的点，利用 Semi variogram Modeling 对话框中拟合的模型，你就能够为未知点预测出一个更精确的值。

8. 点击 Next 按钮。

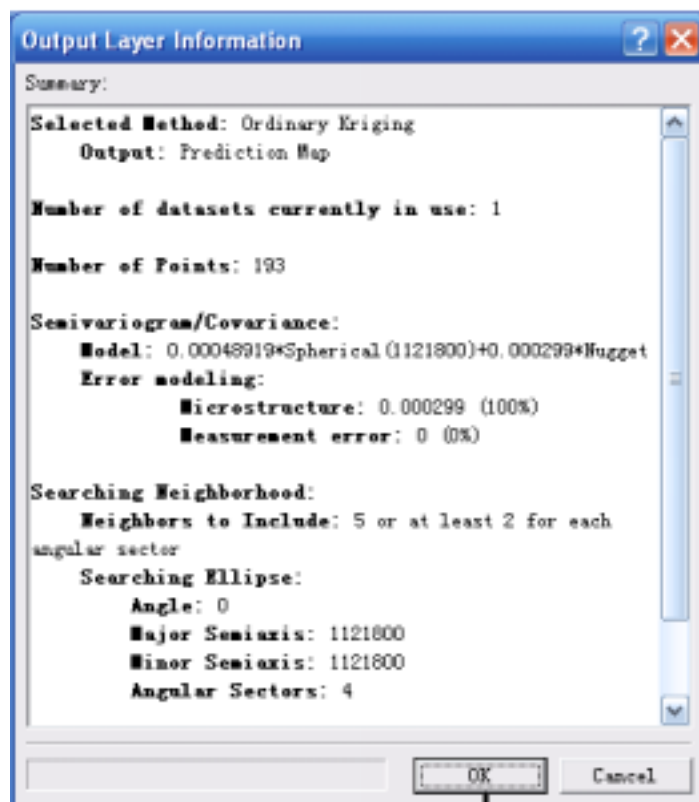


Cross Validation 对话框可以让你知道利用模型预测未知点的值的效果如何。在练习 4 中你将学到如何使用统计图，并理解统计表的意义。

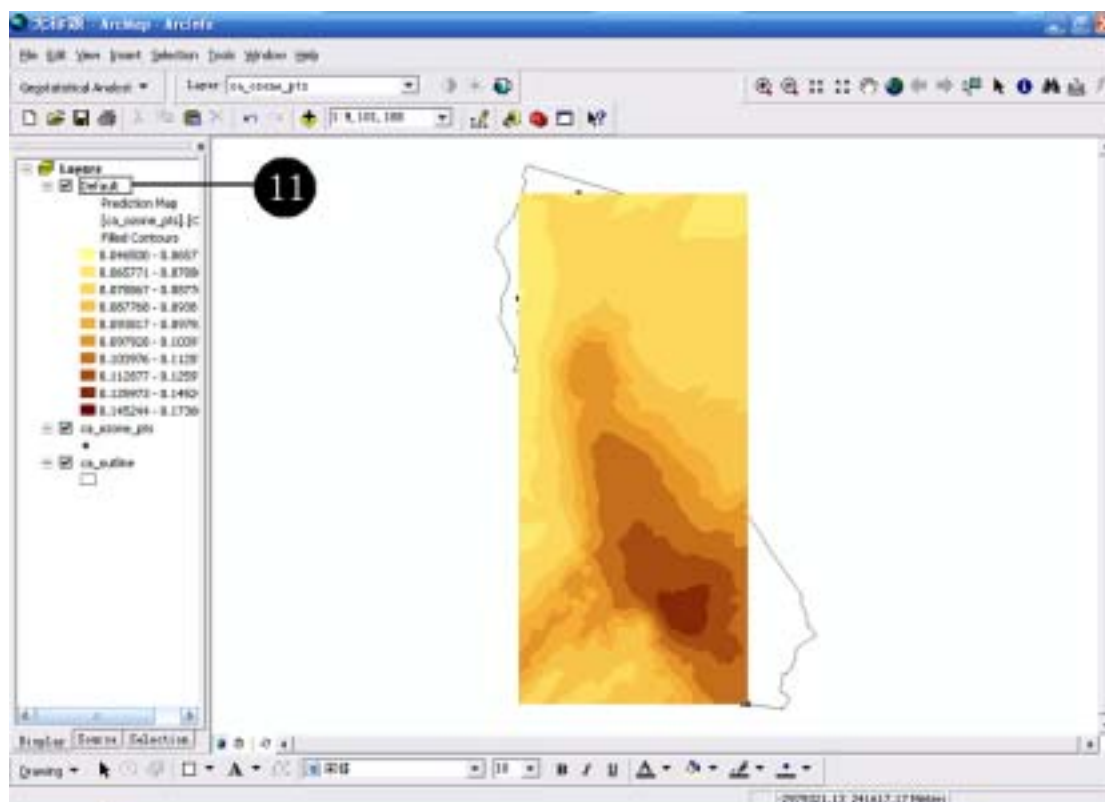
9. 点击 Finish 按钮。

Output Layer information(输出图层信息)对话框总结了用于创建输出表面的方法(及其相关参数) 信息。

10. 点击 OK 按钮。预测得到的臭氧图在目录表的顶层显示。



11. 在目录表中单击该图层以高亮显示，然后再单击一次，将图层名改为 Default。这种改动可以让你区别于练习 4 中创建的图层。



12. 注意，插值的过程一直延伸到海洋中。在练习 6 中你会学会如何将预测表面限制在加州范围内。

13. 保存地图。

练习 2：数据检查

在本练习中，你可以用三种方式对数据进行检验：

- 检测数据分布

- 发现数据可能存在的趋势

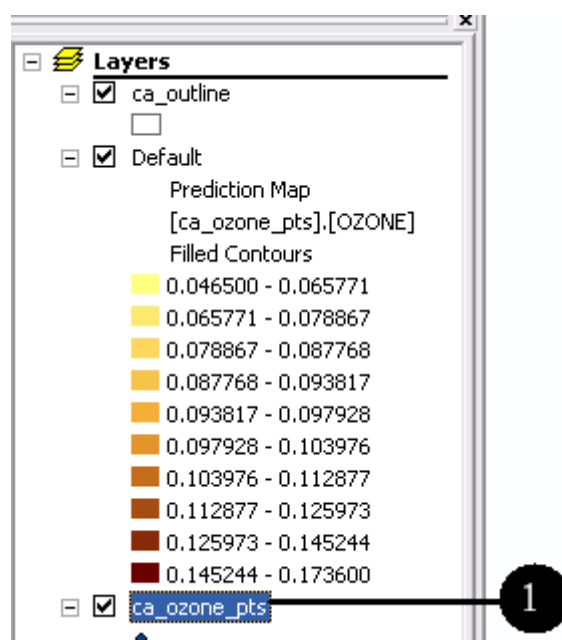
- 找出数据间的空间自相关以及方向效应

如果你在完成练习 1 后关闭了 ArcMap，请单击 File 菜单，再点击 Open。在弹出的对话框中单击 Look in 下拉箭头，找到你保存的地图文件（Ozone Prediction Map.mxd）

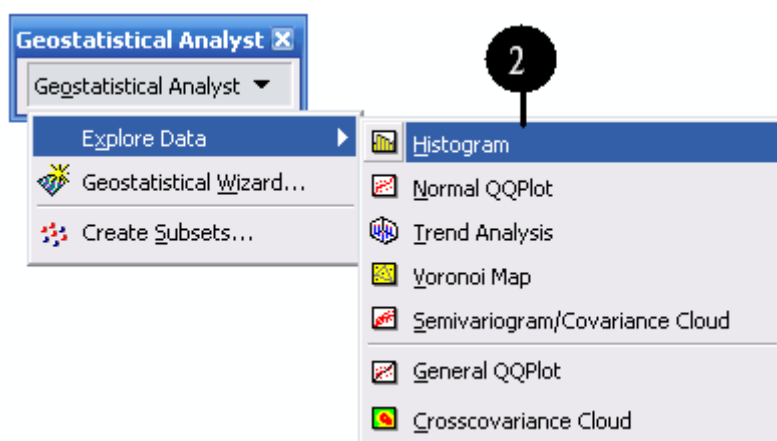
2.1 检查数据的分布

当数据服从正态分布时，**里用**插值方法生成表面的效果最佳。如果你的数据是偏态分布的，即向一边倾斜，则你可以选择数据变换使之服从正态分布。因此在创建表面之前了解你的数据分布非常重要。Histogram 工具描绘了数据属性的频率直方图，使你能够针对数据集的每一种属性检测其单变量分布。接下来，你就是要检查图层 ca_ozone_pts 的臭氧分布情况。

1. 单击 ca_ozone_pts，并将它移到目录表的顶层，然后将 ca_outline 置于 ca_ozone_pts 图层的下面。



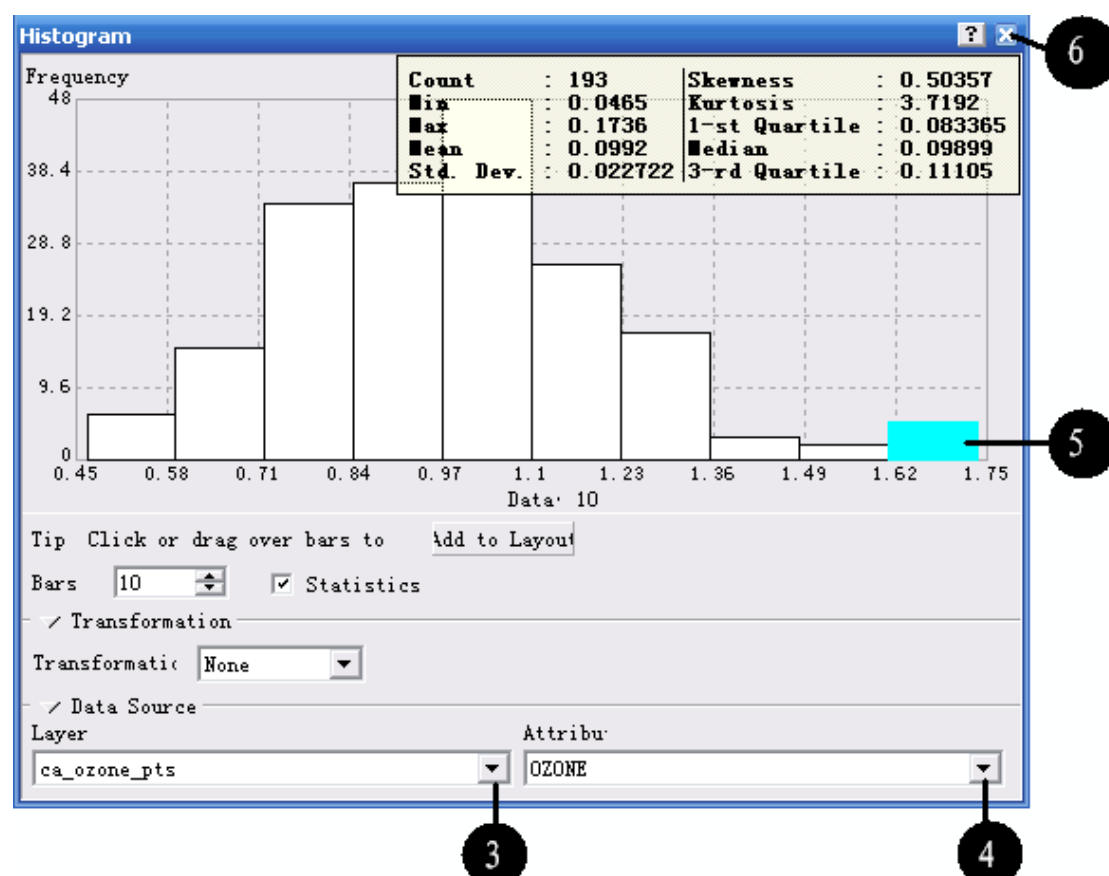
2. 单击 Geostatistical Analyst 工具条，指向 Explore Data，然后单击 Histogram。



你也可以改变 Histogram 对话框的大小以便能够看见地图，正如下图所示：

3. 单击 Layer 下拉箭头，点击并选择 ca_ozone_pts。

4. 单击 Attribute 下拉箭头，点击并选择 OZONE。



臭氧属性的分布情况是用一个直方图来描述的，该直方图将浓度值分为 10 级，每一级别中数量的相对比例（密度）通过每一个直方条柱子的高度来表示。

通常，描述数据分布的重要特征包括中值，他的展布以及对称性。对于正态分布，有一个快速检验的方法：如果平均值与中值大致相等，你就可以把它当作数据服从正台分布的证据之一。

上面显示的直方图表面数据是单峰分布的，而且具有较好的对称性，接近于正态分布。直方图的右册尾部表明，存在相对少量的具有较高臭氧浓度值的采样点。

5．单击直方图臭氧值在 0.162 至 0.175ppm 之间的直方条。

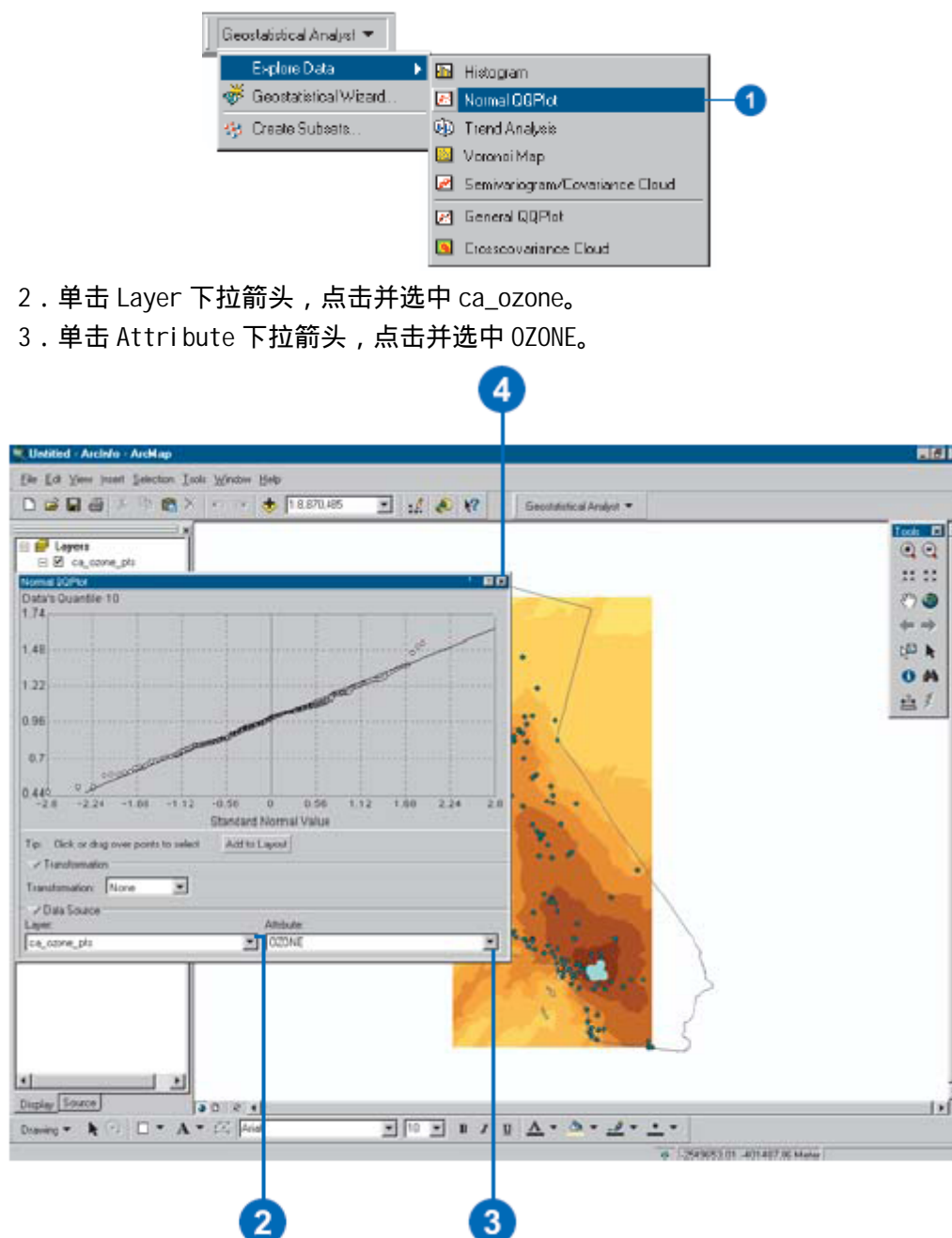
此范围内的采样点在地图中被高亮显示。

6．单击关闭对话框。

2.2 正态 QQ 图

QQ 图提供了另外一种度量数据正台分布的方法，利用 QQ 图你可以将现有数据的分布与标准正态分布对比，如果数据点接近一条直线，则它们越接近于服从正态分布。

1．单击 Geostatistical Analyst 工具条，指向 ExploreData，然后点击 Nomal QQplot。



在一个普通的 QQ 图上，两种分布的对应点一一对应。对于两种相同类型的分布，QQ 图应该是一条直线。因此通过绘制相对应的臭氧数据的分布点与标准正态分布的分布点，能够检查臭氧数据的正态分布情况。从上述正态 QQ 图可以看出，该图形非常接近于一条直线。而偏离直线的情况主要发生在臭氧浓度值较高时（因为这些偏离值在直方图中是高亮显示的，所以在这里它们也是被高亮显示的）。

如果在直方图中或在正态 QQ 图中，数据都没有显示出正态分布，那么就有必要在应用某种克里格插值之前对数据进行转换，使之服从正态分布。

4. 退出对话框。

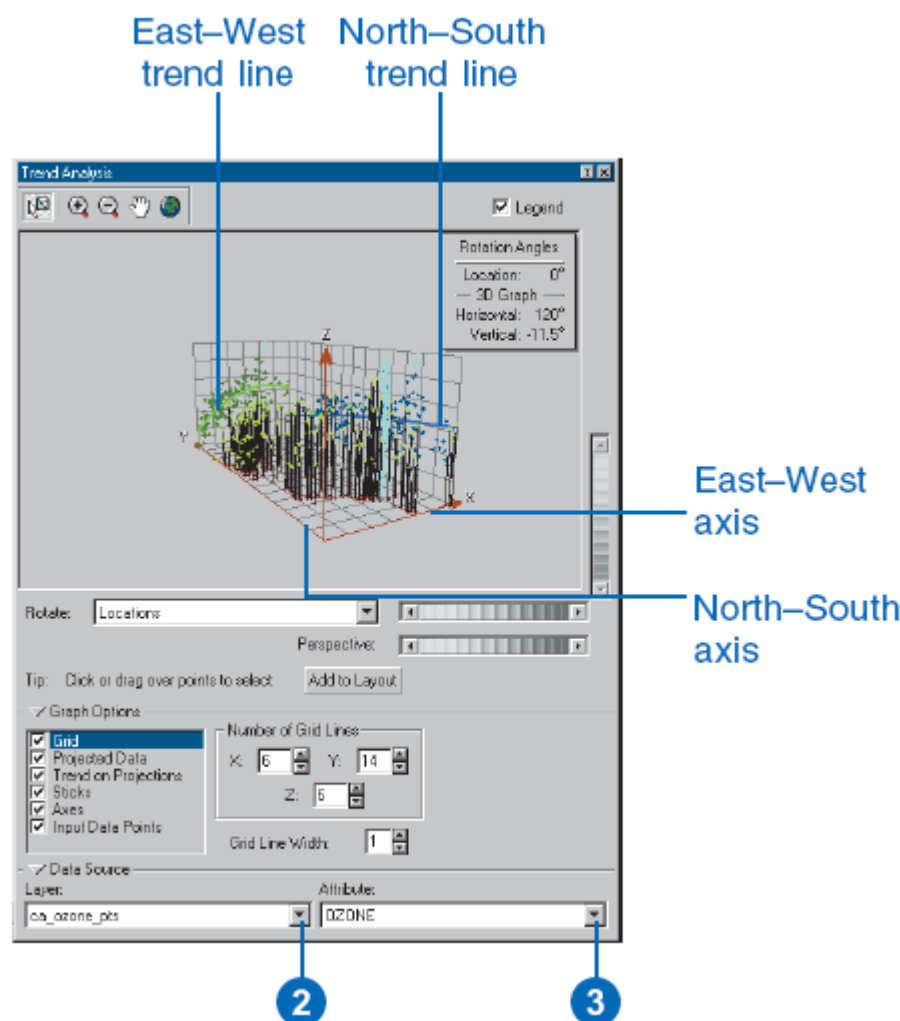
2.3 识别数据中的全局趋势

只有在你的数据中存在某种趋势时，你才可能利用某些数学公式对表面的非随机（确定

性的)成分进行表达。例如,一个缓倾斜的山坡可以用一个平面来表达,而山谷则可以利用一个能够生成“u”字形的更加复杂的公式(一个二次多项式)来表示。数学公式有时或许能够生成你想要的表面,但大多数时候,数学公式因为太过于平滑而不能精确地描述表面,因为没有哪个山坡是完完全全的平面,同样,也没有哪个山谷会是一个完美的“u”形。如果趋势面不能精确地描绘你实际需要的表面,你可能想到将其移去,通过建立趋势剔除后的残差的模型来继续你的分析。在建立残差模型时,你需要分析表面中的短程变异。这是理想平面或理想“u”型面所无法实现的内容。

Trend Analysis(趋势分析)工具使你能够找出在输入数据集中是否存在趋势。

1. 单击 Geostatistical Analysis 工具条,鼠标指向 Explore Data ,点击 Trend Analysis.
- 2 单击 Layer 下拉箭头,点击选中 ca_ozone_pts。



3. 单击 Attribute 下拉箭头,选中 OZONE ,

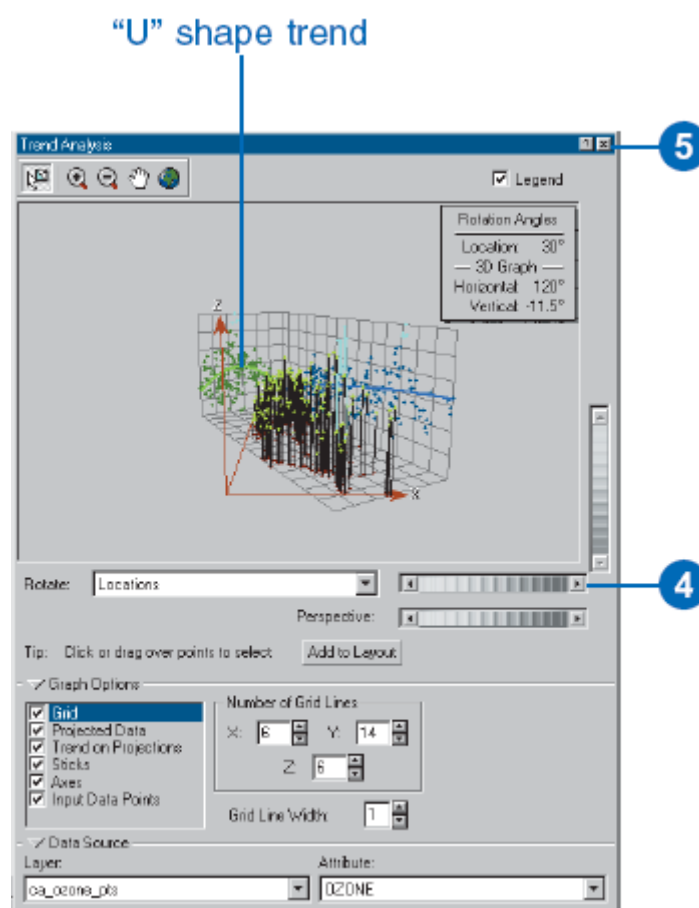
趋势分析图中的每一根竖棒代表了一个数据点的值(高度)和位置。这些点被投影到一个东西向的和一个南北向的正交平面上。通过投影点可以作出一条最佳拟合线(一个多项式),并用它来模拟特定方向上存在的趋势。如果该线是平直的,则表明没有趋势存在。不过,如果你注意看上图中的亮绿线,你会发现这条线从较低的值开始,向东移动时逐渐增加直到变平稳。这表明该数据在东西向上显示出一个很强的趋势,而在南北向的趋势则较弱。

4. 单击 Rotate Projection 滚动条并向左拖动,使旋转角为 30 度。

通过旋转,东西向趋势的形状可以看得更清楚。你能看到投影后确实显示了一个倒置的“u”型。既然该趋势呈“U”型,所以可以选择一个二阶多项式对其全局趋势进行较好的模

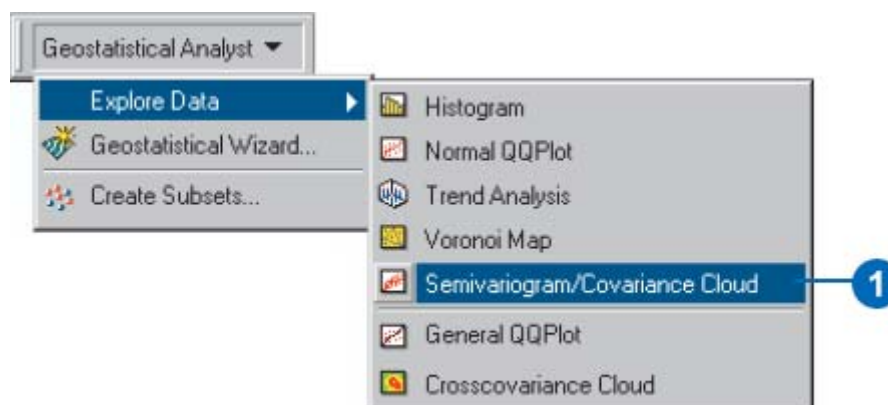
拟。尽管我们把这个趋势显示在了东西向的投影平面上，但因为我们把数据点旋转了 30 度，所以实际的趋势是北东-南西向。造成该趋势的一个可能的事实是，在沿海地区污染较轻，而在向内陆推进时，人口增多，污染增大。到了山区则人口又减少，污染也随之降低。在练习 4 中，这些趋势将要被剔除。

5. 单击退出对话框。



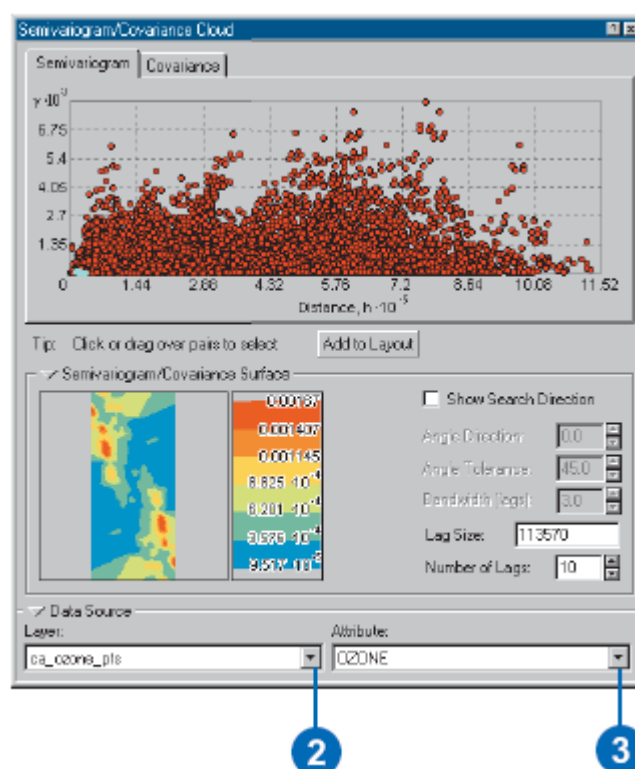
2.4 理解数据的空间自相关和方向效应

1. 单击 Geostatistical Analyst 工具条，指向 Explore Data，点击 Semivariogram / Covariance Cloud。



2. 单击 Layer 框下拉箭头，点击选中 ca_ozone_pts。

3. 单击 Attribute 框下拉箭头，单击选中 OZONE。



半变异函数 / 协方差函数云图使你能够检测已测样点间的空间自相关。空间自相关理论认为彼此之间距离越近的事物越相象。半变异函数 / 协方差函数云图使你能够对这种关系进行检测。为此，可以用 Y 轴表示半变异函数值，即每一样点对间测量值之差的平方，而相应地用 x 轴表示每对样点之间的距离。

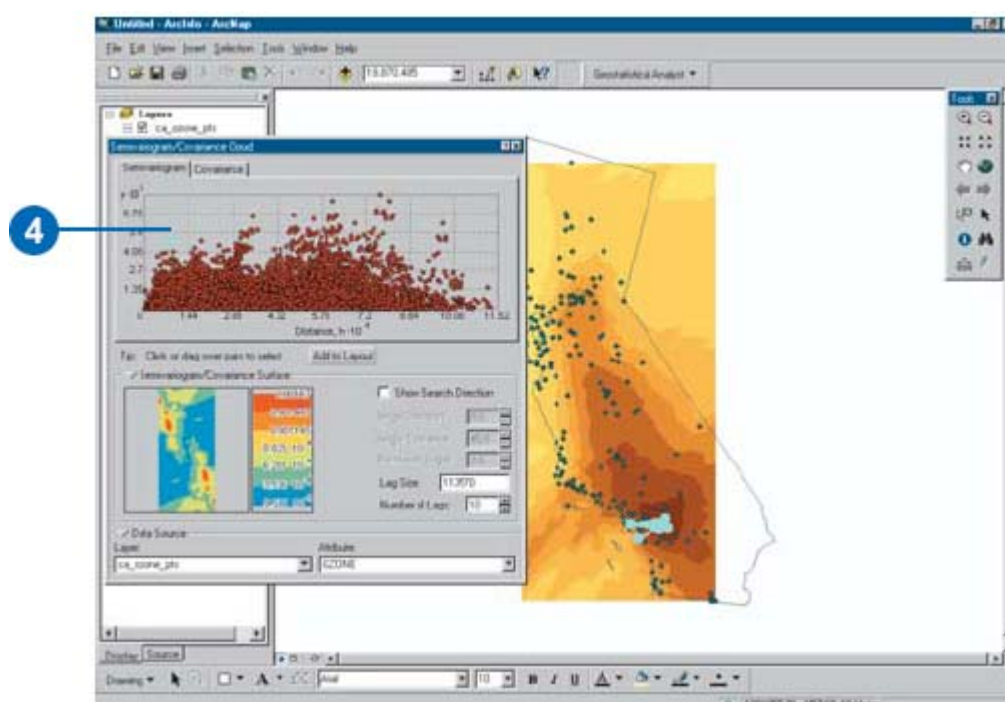
在半变异函数 / 协方差函数云图中，每个红点表示一对采样点。既然越近的点越相似，那么在半变异函数云图中邻近的点(在 x 轴的左边)应该有较小的半变异函数值(在 Y 轴的下部)。随着样点对间距离的增加(在 x 轴上向右移动)，中变异函数值也要相应增加(在 Y 轴上向上移动)。然而，当到达一定的距离后，云图变平，这表明超出这个距离时，样点对之间不再具有相关关系了。

观察半变异函数图，如果某些靠得很近的数据点(在 x 轴上接近于零)具有——一个异常的较高的半变异函数值(在 Y 轴的上部)时，你就应该仔细检查这些样点对，看看是不是这些数据不准确。

4. 在这些点上单击并拖动 Selector(选择)光标使之高亮显示。(可以使用下图作为指导，是否严格选择图中所示的那些点以高亮显示并不重要。)

在半变异函数图中选中的采样点将高亮显示在地图中，样点间通过直线相连，用以表示是一对采样点。有多种原因可以解释为什么洛杉矶地区的采样点数据值和其他地区相比差别很大。一种可能是洛杉矶地区的汽车比其他地区要多，这些汽车无疑将会造成更多的污染，从而在洛杉矶地区形成一个较高的臭氧累积。

除了前面章节中讨论的全局趋势外，影响数据的还有方向效应。这些方向效应的原因可能并不明了，但它们可以在统计上给予量化。这些方向效应将影响你在下一个练习中创建表面的精度。不过，一旦你知道某种方向效应确实存在，则地统计分析模块提供的工具可以在你创建表面的过程中对其给予解释。你可以利用方向查找工具，在半变异函数云图中检测某个方向效应。



5. 选中 ShowSearchDirection 复选框。

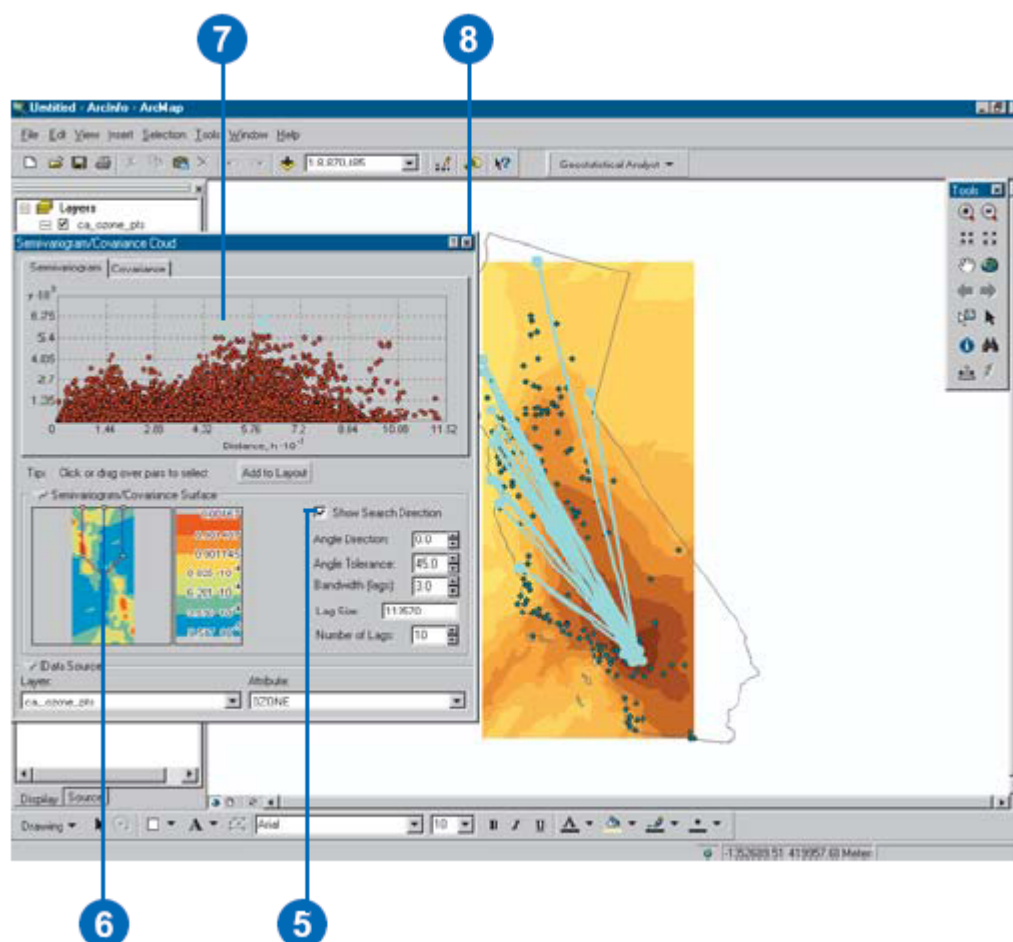
6. 点击并将方向指针移动到任意角度。

指针指向的方向决定了哪些样点对将会出现在半变异函数图中。例如，如果指针是东西方向，那么只有那些处于彼此东西方向的点对才会在半变异函数图中显示，这就使你能够去除你不感兴趣的那些点对从而来检查你数据中的方向效应

7. 单击并拖动选择工具，选中那些具有最大半变异函数值的点，使之在半变异函数图及地图中高亮显示。（可以使用厂图作为指导，是否严格选择图中所示的那些点以及是否使用相同的方向搜索并不重要。）

你会注意到，无论距离大小，大多数相连的样点对（代表了地图上的样点对）都会对应到洛杉矶地区的某一采样点上。对更多不同距离的点对的考虑结果表明，并非只有那些从洛杉矶地区延伸至海边的点对才具有很高的半变异函数值，许多从洛杉矶地区延伸到其他内陆地区的数据点对同样也有很高的半变异函数值。这是因为洛杉矶地区的臭氧值比加州其他任何地区都要高的多。

8. 点击退出对话框。



9. 单击 Selection 菜单，然后单击 Clear Selected Features 以释放地图中高亮显示的点。

在本练习中我们学到：

1. 臭氧数据接近于正态分布。正如我们在直方图中看见的那样呈单峰并且在平均值/中值两侧有较好的对称性。

2. 正态 QQ 图再次证实了数据是服从正态分布的，因为 QQ 图中的点构成了一条非常近似的直线。因此，无须进行数据转换。

3. 利用 Trend Analysis 工具可以看出数据显示出某种趋势，该趋势经过提取后，可以用一个南东-北西方向的二阶多项式对其进行最佳拟合。

4. 从半变异函数/协方差函数云图中我们发现洛杉矶地区距离很近的样点对却与其它地区那些相距较远的样点对一样具有较高的浓度值。

5. 半变异函数表面表明数据中存在空间自相关。

现在你已经知道数据集中没有离群值或错误的采样点，并且数据接近于正态分布，所以你可以放心地进行表面插值。而且你还知道数据中存在某个趋势，所以在插值的过程中通过某些调整，你可以创建一个更加精确的表面。

练习 3：制作臭氧浓度图

在练习 1 中，你已经利用缺省参数生成了臭氧浓度图。然而你并没有把采样数据的统计属性考虑进去。举例说，在练习 2 中进行数据检查时发现，数据显示出一种趋势。这种趋势可以在插值过程中表现出来。

本练习你将学习：

对练习 1 中生成的臭氧浓度图进行改进。

学习一些基本的统计思想。

你将再一次使用普通克里格插值法，并且在你的模型中加入趋势因素以便进行更好的预测。

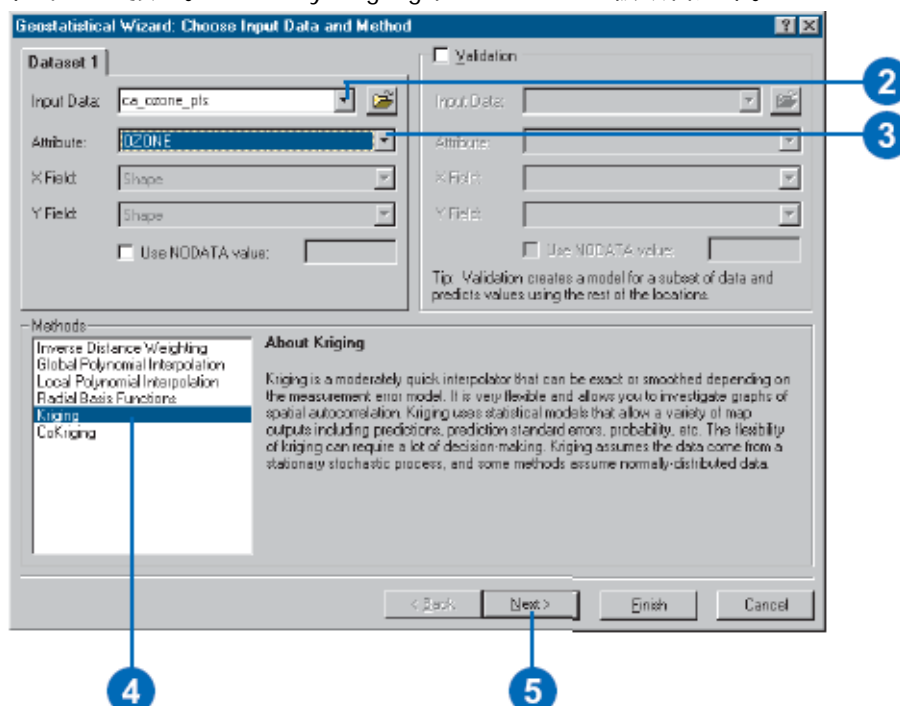
1. 单击 Geostatistical Analyst—工具条，然后单击 Geostatistical Wizard。

2. 单击 Input Data 下拉框箭头，点击选中 ca_ozone_pts。

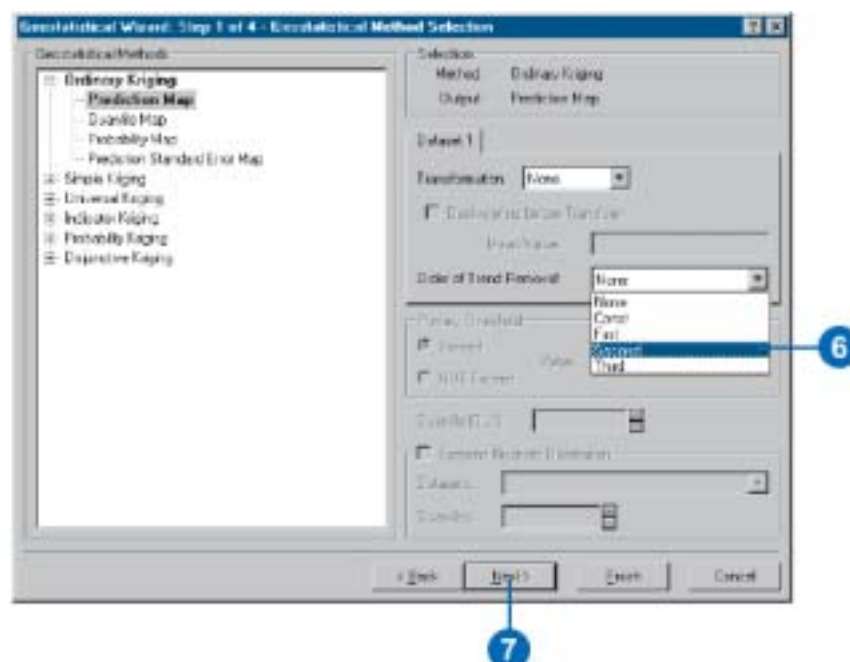
3. 单击 Attribute 下拉框箭头，点击属性 OZONE。

4. 在 Methods 框中选择 Kriging。

5. 单击 Next 按钮。Ordinary Kriging 和 Prediction 被缺省选中。



在练习 2 对你的数据进行检查时，你已经发现数据中存在一个全局趋势。利用 Trend Analysis 工具进行修正后，可知该趋势是南东-北西方向，并且可以用一个二阶多项式进行拟合。该趋势可以从数据中剔除，并可以用一个数学公式表达。一旦剔除全局趋势后，就可以对表面残差或表面的短程变异成分进行统计分析。在创建最终表面之前，该趋势还将自动添加回来以产生更有意义的结果。全局趋势剔除后所进行的分析将不再受其影响。而一旦将全局趋势再添加进来，就能够生成一个更加精确的表面。



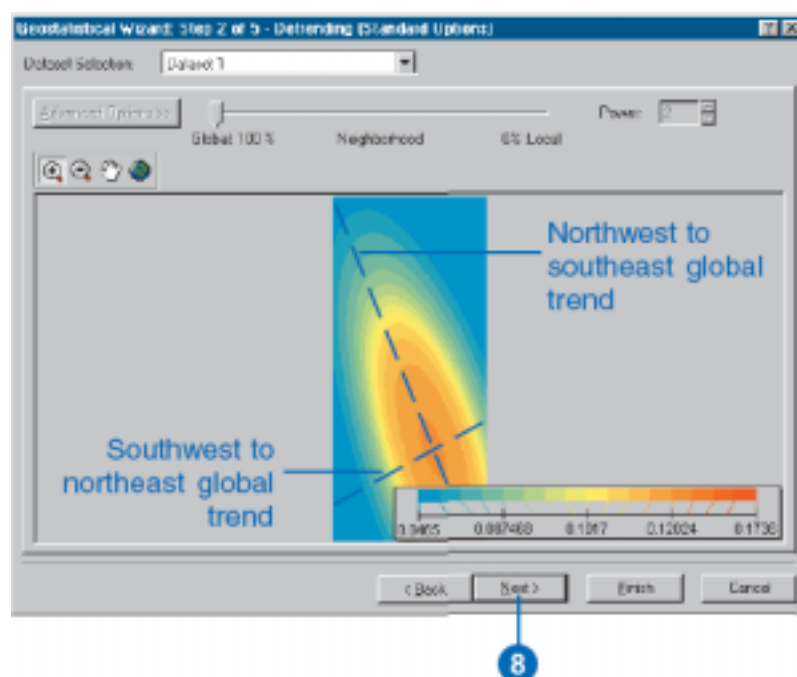
6. 在 Geostatistical MethodSelection 对话框中，单击 Order of Trend Removal 下拉箭头，选择 Second。因为在练习 2 的 TrendAnalysis 对话框中已经检测到一条南西-北东方向的“u”型曲线，所以选择二阶多项式拟合是合适的。

7. 在 Geostatistical Method Selection 对话框中点击 Next 按钮。

缺省情况下，地统计分析模块将绘制数据集中的全局趋势。从下图可以看出，南西-北东向的变化最快，而北西-南东向的变化则较平缓(从而形成椭圆形)。

趋势剔除必须合情合理。空气质量中的这种南两-北东向趋势可以归因于山区与沿海间的臭氧积累。山区的海拔高度以及在沿海地区占主导地位的风向导致山区与沿海地区的臭氧浓度值相对较低。人口密度过人导致在山区和沿海地区之间的污染程度较重。而北西-南东向的趋势变化较缓是由于洛杉矶附近的人口密度较高，而向圣弗朗叫斯科方向延伸时人口密度却降低。因此我们将这些趋势剔除是合理的。

8. 点击 Detrending 对话框中的 Next 按钮。



3.1 半变异函数 / 协方差函数模型

在练习 2 Semi variogram / Covsance Cloud 中, 你已经检测了已测点的所有的空间自相关。为此你需要检查半变异函数, 半变异函数显示了具有不同距离的样点对的方差。用半变异函数 / 协方差函数进行模拟的目的在于为其确定一个最佳拟合模型, 该模型将穿过半变异函数图中的那些点(如图中黄线所示)。

半变异函数是一个关于数据点的半变异值(或称变异性)与数据点间距离的函数。对它的图形表述可以得到一个数据点与其相邻数据点的空间相关关系图。

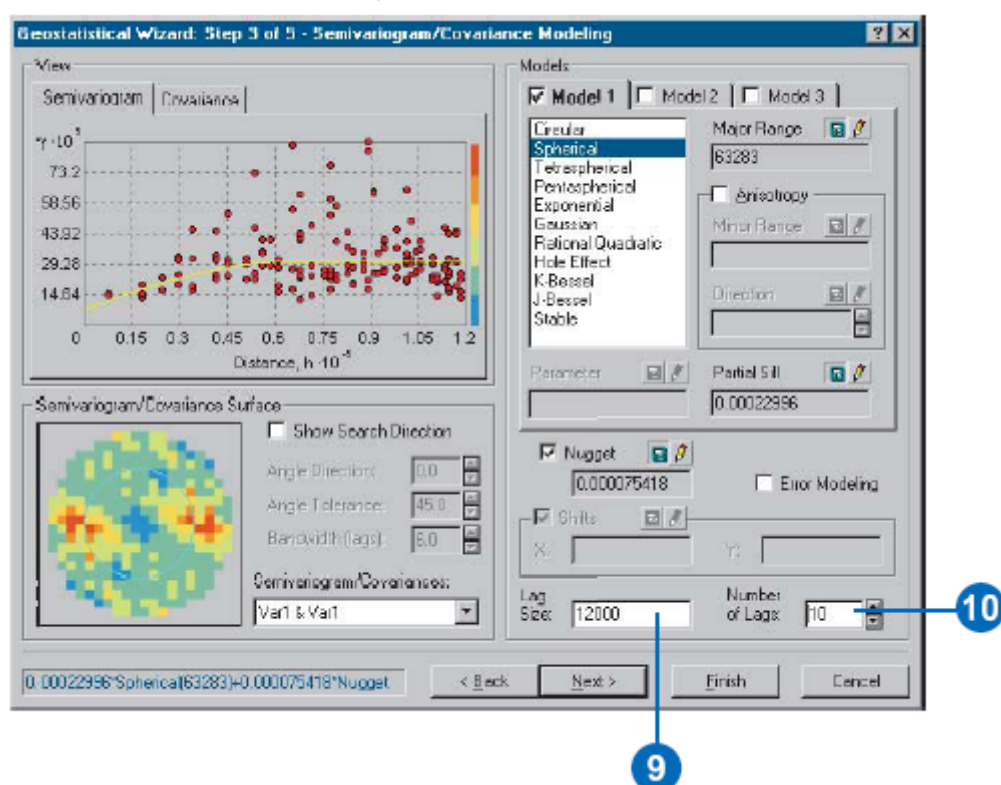
在 Semi variogra/Covariance Modeling 对话框中你可以模拟数据集的空间关系。缺省情况下, 将以球面半变异函数模型来计算其最佳参数值。首先, 地统计分析模块要为半变异函数值的划分确定一个合适的步长大小。为了减少大量可能的合并而将数据点对分成不同的距离级, 该距离级的大小就是步长。这种方法称为步氏分组。步长分组后, 半变异函数图中的点与练习 2 相比少了许多。一个好的步长大小也有助于揭示空间相关关系。该对话框用表面图和散点图的形式来显示半变异函数值与距离间的关系。然后拟合一个球面半变异函数模型(在各个方向都能拟合的很好)以及它们的相关参数值, 这些参数通常被称为块金效应、自白相关阈值及偏基台值(结构方差)。

可以尝试用小的步长(距离)来拟合半变异函数。可以采用不同的步长分组值, 通过改变步长的大小和步长组数目来重新拟合缺省球面模型。

9. 输入一个新步 K 值 12000。

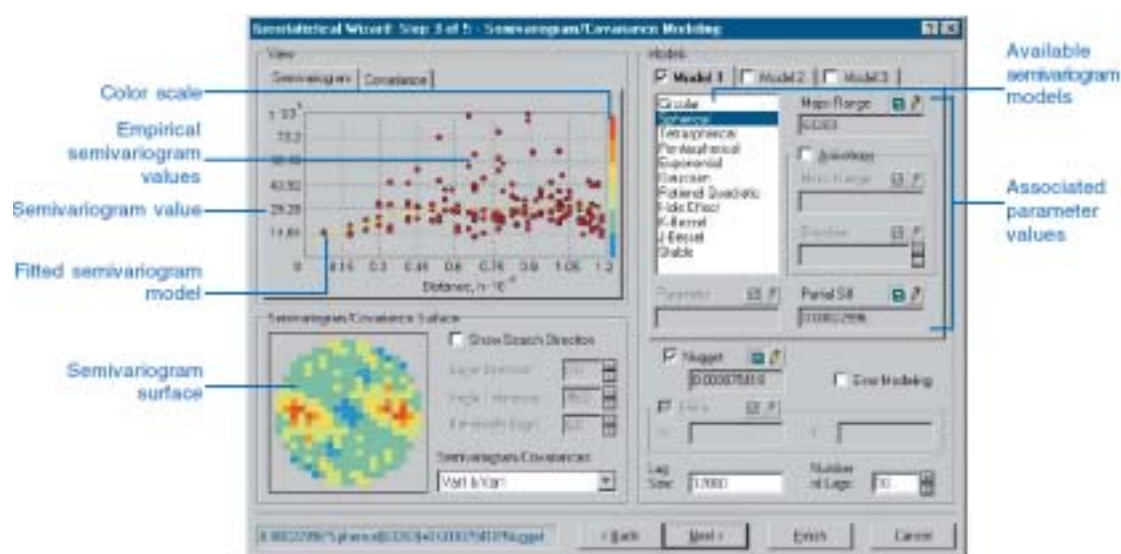
10. 单击输入框, 设定步长组的数目为 10。

减小步长的大小意味着你可以有效地放大并模拟相邻采样点间局部变异的细节。你会注意到当赋予一个较小的步长时, 拟合后的半变异函数图(图中黄线)急剧上升、接着变平。其自相关阈值即为曲线变平时经过的距离。半变异函数图的这种变平现象表明在超出其自相关阈值后自己相关关系就不再存在了。



提出趋势后, 半变异函数就可以模拟数据点间自相关而不用考虑数据中存在的趋势。该

趋势将在生成最终表面之前添加回来用于计算。



颜色比例尺代表了计算后的半变异函数值。它将散点图中的理论变异函数值与半变异函数表面土长的值直接联系起来。半变异函数表面图上的每个栅格单元用颜色进行编码，蓝色和绿色代表的值较低，而橙色和红色代表的值较高。可以用半变异函数散点图来表示半变异函数表面图上每个栅格单元的平均值，图中 x 轴表示从栅格单元中心到半变异函数表面图中心的距离，半变异函数值代表了变异性大小。在我们的例子中，当距离很小时，半变异函数值也很低（即越靠近的事物相似性越大）：随着距离的增加，其值也增大（即越远离的事物变异性越强）。从半变异函数表面图可以看山，在南西-北东向的变异性要比南东-北西方向增加得要快。在前面的操作中，你已剔除了一个大尺度的趋势，现在表现山的是细微尺度上的白相关的方向性成分，因此我们将在下一步对其进行模拟。

3.2 方向半变异函数

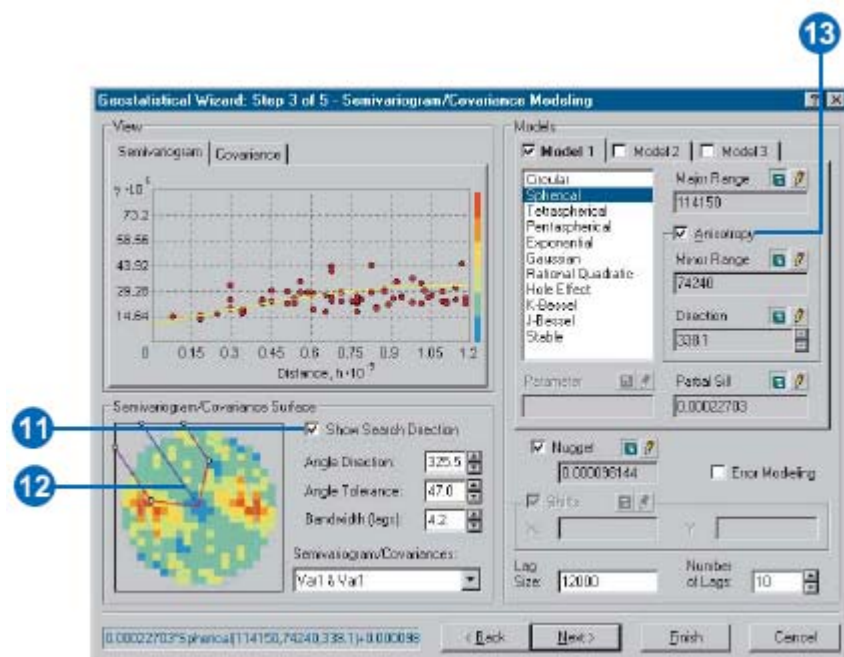
方向效应会对半变异函数中的点以及将要拟合的模型产生影响。邻近的事物在某些方向上的相似性比其他方向的相似性更强。方向效应又被称为各向异性，地统计分析模块能够对它们进行解释。引起各向异性的因素可能是风、侵蚀、地质构造或者许许多多其他过程。你可以对方向效应在统计上进行量化，并在制图时给予解释。

你可以用 Search Direction 工具来分析某一方向上数据点的变异性，利用该工具你可以在半变异函数图上检测方向效应，这并不影响输出的表面。下面的步骤教你如何做到这一点：

11. 选中 Show Search Direction 复选框。注意个变异函数值数目的减少，因为只有那些位于你搜索方向上的点才会得到显示。

12. 在 Search Direction 的中心线上点击并按住鼠标，移动搜索工具的方向。当你改变搜索方向时，请注意半变异函数是如何变化的。只有在搜索方向上的半变异函数表面图上的值才会显示出来。

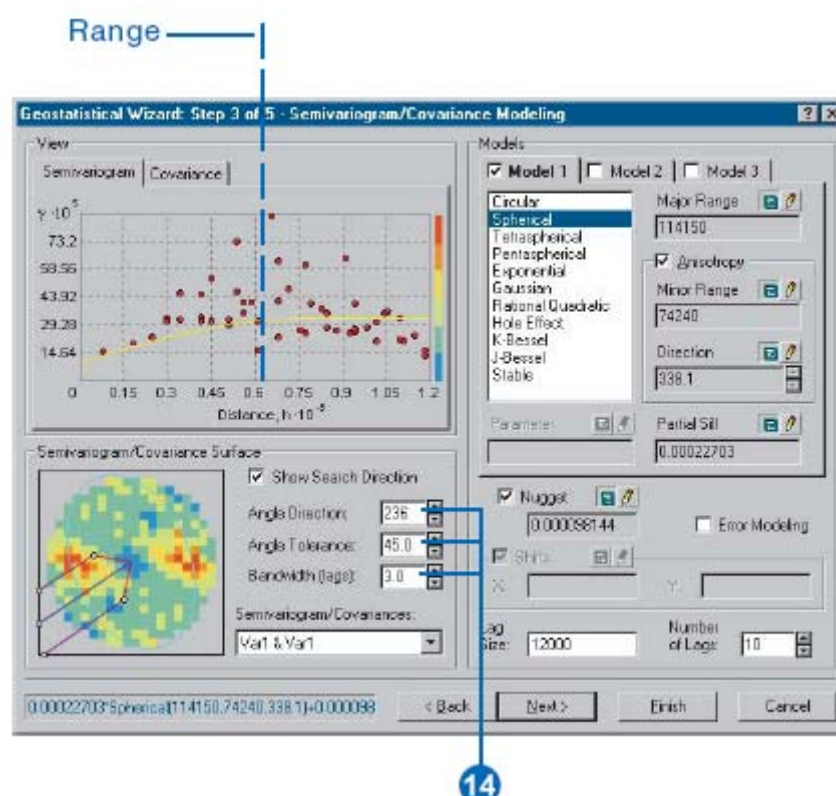
为实际说明用于表面计算的半变异函数模型中的方向效应，你必须计算各向异性半变异函数或者协调方差模型。



13. 选中 Anisotropy 复选框。

在变异函数表面图中的蓝色椭圆表明了不同方向上的半变异函数的自相关阈值。在这种情况下，椭圆主轴大致位于北北西-南南东方向。

现在你可以在模型中加入各向异性，用以调整输出表面中自相关的方向效应。



14. 为搜索方向输入下列参数，使方向指针与各向异性椭圆的短轴重合。

角度方向：236.0

角度容限值：45.0

带宽(步长) : 3.0

可以看到半变异函数曲线形状比它的基台值增长得要快。因为 x 轴和 Y 轴都是以米为单位, 所以此方向的白相关阈值人为 74 公里。

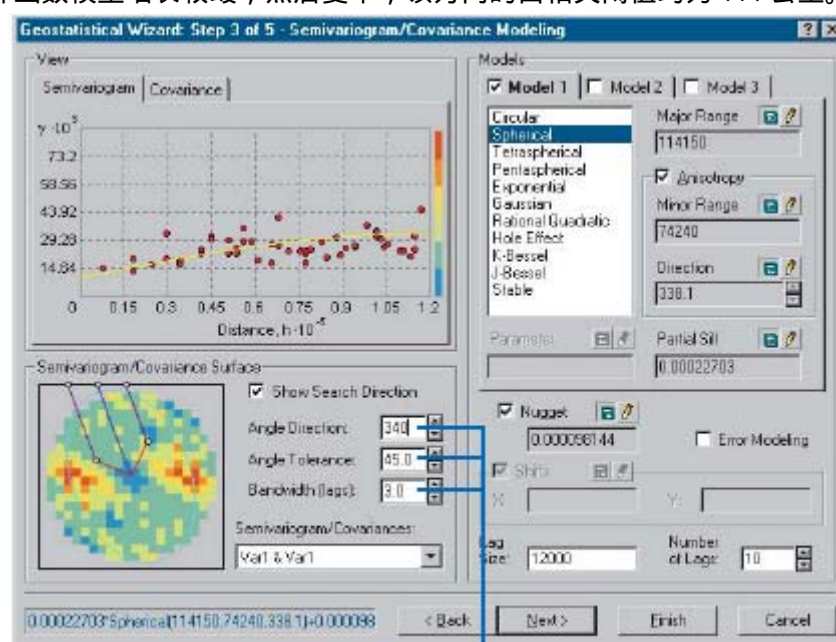
15. 为搜索方向输入下列参数, 使方向指针与各向异性椭圆的长轴重合。

角度方向 : 340.0

角度容限值 : 45.0

带宽(步长) : 3.0

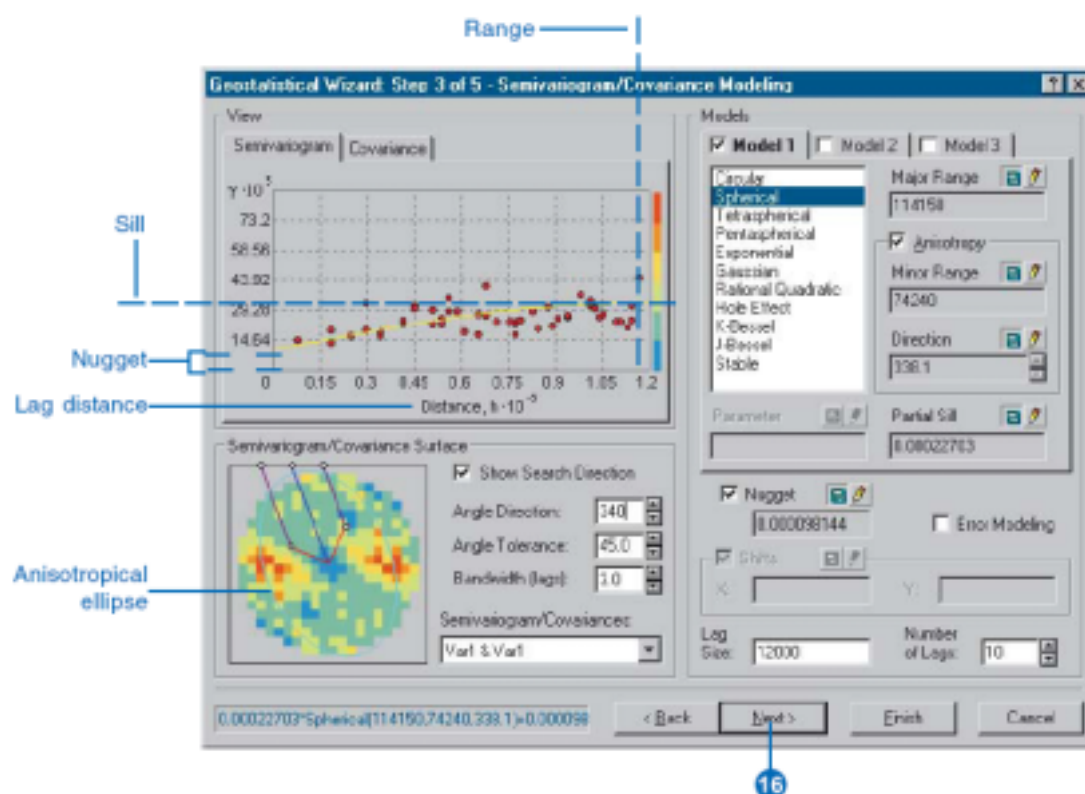
半变异函数模型增长较缓, 然后变平, 该方向的白相关阈值约为 114 公里。



在第 14 步和第 15 步中, 半变异函数模到达到的稳定水平时的值相同, 我们称之为 基台值。而半变异函数模达到其限值(基台值)时经过的距离, 即为模型的自相关阈值。超出这个自相关阈值后, 各点之间的变异性将随着步长距离的增加而变为常数。步长是通过点对间的距离来界定的。步长距离大于自相关阈值的样点对之间空间无关。块金效应代表了测量误差和(或)微观尺度的变异(该变异在空间尺度上太小以至于无法检测到)。如果你在每个点都有数个观测值, 或者你可以通过选中 Nugget Error 复选框将块金效应分解为测量误差和微观尺度变异, 这样你就有可能对测量误差进行估计。

16. 点击 Next 按钮。

现在你有了用来描述空间自相关的拟合模型, 并且该模型考虑了数据中的趋势剔除和方向效应。利用这个模型, 结合预测点周围的点的结构和测量值, 你就可以对它进行预测。但是怎样应用这些人工测定的点来进行计算呢?



3.3 领域搜索

定义一个圆(或者椭圆)，然后利用其内的点来预测那些未知点的值是一件简单的操作。

此外为避免某一特定方向上的偏差，可以把这个圆(或者椭圆)分为若干个小扇形，在各扇形内选取相同数目的点。利用 Searching Neighborhood 对话框，你可以指定点的数目(最大为 200)、半径(或者长 / 短轴)以及用来预测的圆(或者椭圆)中的扇形个数。

数据视图窗口中高亮显示的点表明了预测未知点时，各相关点的权重。本例中有四个点(红色)的权重大于 10%。某点的权重越大，其对未知值的预测影响也越大。

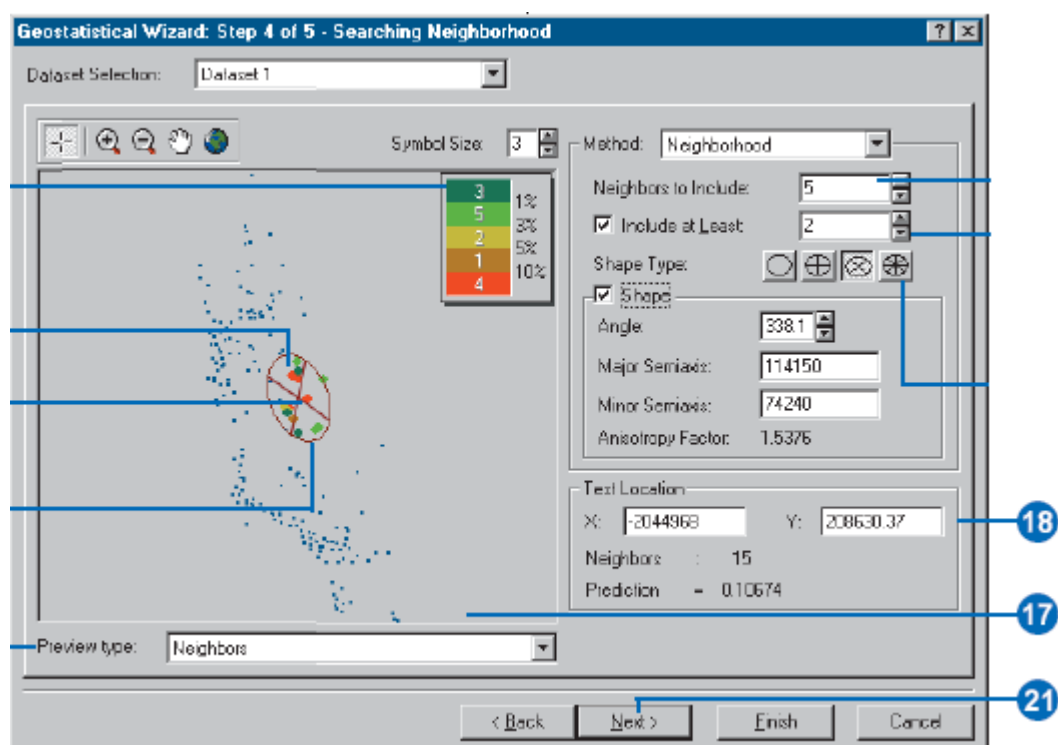
17. 在图形视图窗口内单击选择一个预测点(即十字丝处)。注意在选择那些用厂预测的点的过程中发生的变化(以及它们的权重变化)。

18. 为了教学目的，请在 Test Location 输入框中键入如 F 坐标：

X=2044968

Y=208630.37

19. 选中 Shape 复选框，并在 Angle 输入框中输入 90。注意形状是如何变化的。不过，为了说明方向效应，请把角度值改回 338.1。



20. 取消 Shape 复选框，地统计分析模块将使用缺省值(在前面 Semi variogram / Covariance 对话框中曾经计算过)。

21. 在 Searching Neighborhood 对话框中单击 Next 按钮。

在实际创建表面之前，你将在下一步中使用 Cross-validation 对话框对参数进行诊断以判断你的模型效果如何。

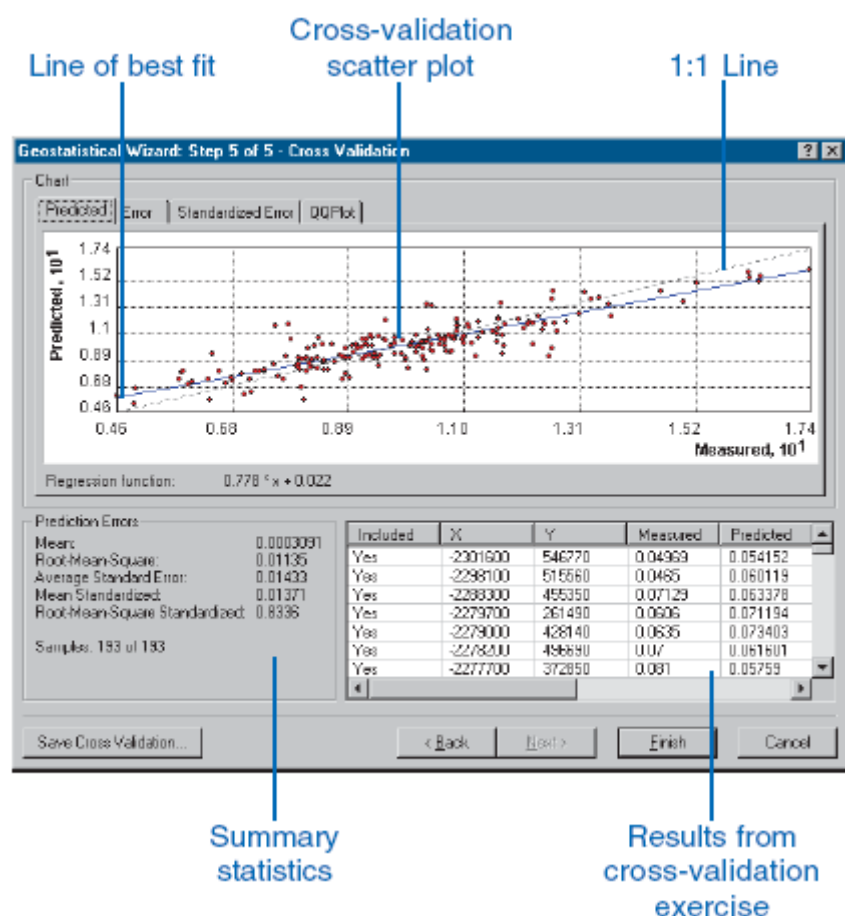
3.4 交叉验证

交叉验证可以让你知道你的模型对未知值的预测效果究竟怎样。

对于所有的点，交叉验证按顺序每次省略一个点，再利用剩余的数据来对该点的值进行预测。然后在实测值与预测值之间进行比较。可以用计算后的统计数据来判断该模型用于生成地图是否合适。

除了可以对角平分线附近的散点进行可视化表达外，还有许多统计方法可用来评估模型的性能。交叉验证的目的就是帮你做出周全的决定，让你知道哪个模型提供的预测最精确。对于一个预测精确的模型，其均差应接近于 0，其均方根误差和平均标准差应该尽可能地小(这在比较模型时很有用)，并且其均方根标准误差应该接近于 1。

术语“prediction error(预测误差)”是用来表示预测值与实际测量值之间的差异。对于一个能够精确预测的模型，如果是无偏估计，其平均预测误差应该接近于 0；如果标准差准确的话，其均方根标准预测误差应该接近于 1，如果预测值与实际测量值接近，则其均方根预测误差应该很小。

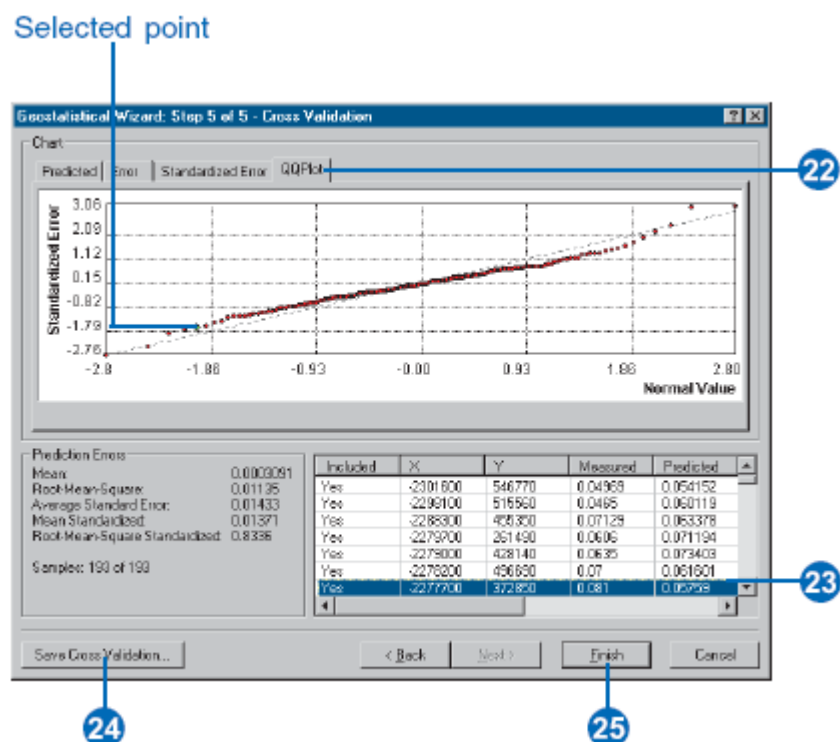


CrossValidation 对话框还可以以散点图的形式为每个数据点显示其误差、标准差及 QQ 图。

22. 点击 QQPlot 标签显示 QQ 图。从 QQ 图上你可以看出某些值稍微落在直线的上部，而某些值则稍稍落在直线的下部，但是大部分点都接近于一条平直的虚线，这表明预测误差近似于正态分布。

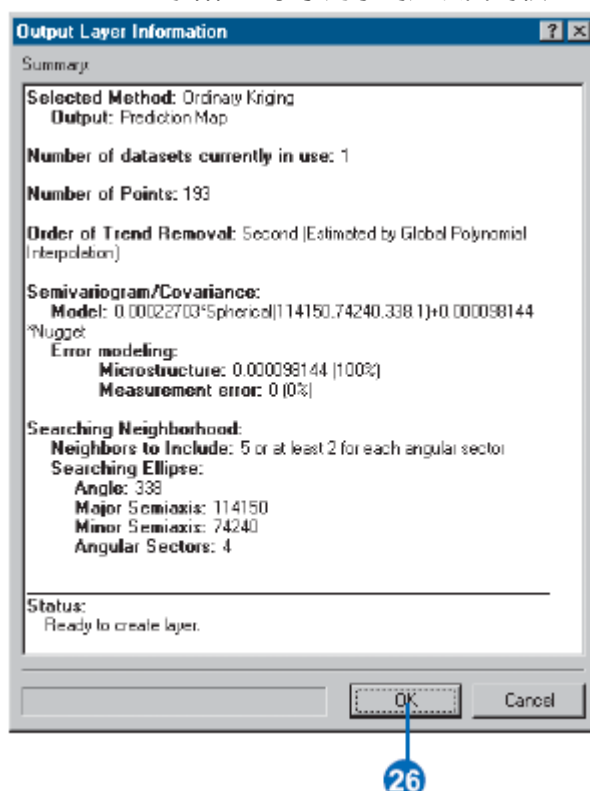
23. 要想高亮显示某个特定的点，可以在表中单击与之相关的行。所选的点在散点图中以绿色高亮显示。

24. 作为可选项，你可以单击 SaveCrossValidation 按钮来保存该表以便对结果做进一步的分析。



25. 单击 Finish 按钮。

Output Layer information 对话框显示了用于创建表面的模型的信息摘要。



26. 点击 OK 按钮。预测得到的臭氧图就会在 ArcMap 中作为顶层显示出来。

图层的默认名称是你用来创建表面所使用的克里格方法的名字（例如，Ordinary Kriging）。

27. 单击图层名以高亮显示图层，再次单击，将图层名称改为“Trend removed”。

同样你也可以创建一个预测标准误差表面来检查你的预测质量。

28. 在你创建的图层 “Trendremoved” 上右击鼠标，点击 Create Prediction Standard Error Map 菜单项。

29. 在 Standard 工具条上点击 Save 按钮。

预测标准误差可以对你创建的表面上的每个点的不确定性进行量化。根据经验，如果数据服从正态分布，在 95% 的情况下，表面的真实值应该在预测值 ± 2 倍预测标准误差的范围内浮动。可以看出在预测标准误差表面上，靠近采样点附近的位置通常其误差值也较小。

在练习 1 中创建表面时，使用的是地统计分析模块中的缺省参数而没有考虑表面中存在的某些趋势，也没有使用较小的步长值或者各向异性半变异函数模型。本练习中你创建的预测表面考虑了数据的全局趋势，并对半变异函数中的步长大小及局部方向效应(各向异性)进行了调整。

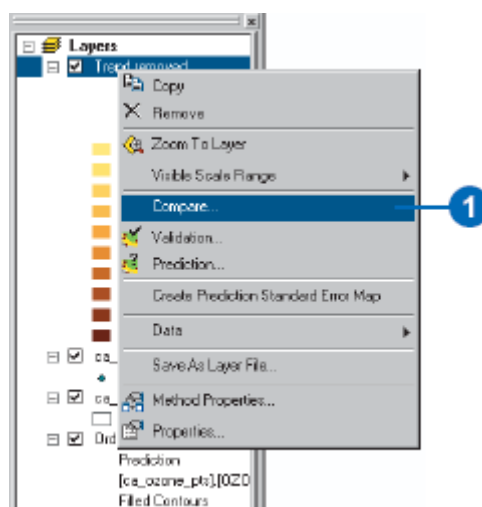
在练习 4 中，通过对两种模型的比较，你可以看出哪种模型对未知值的预测更好。

注意：你会再次看到插值表面一直延伸到海洋中。在练习 6 你将学习如何将预测表面限制在加州范围内。

练习 4：模型对比

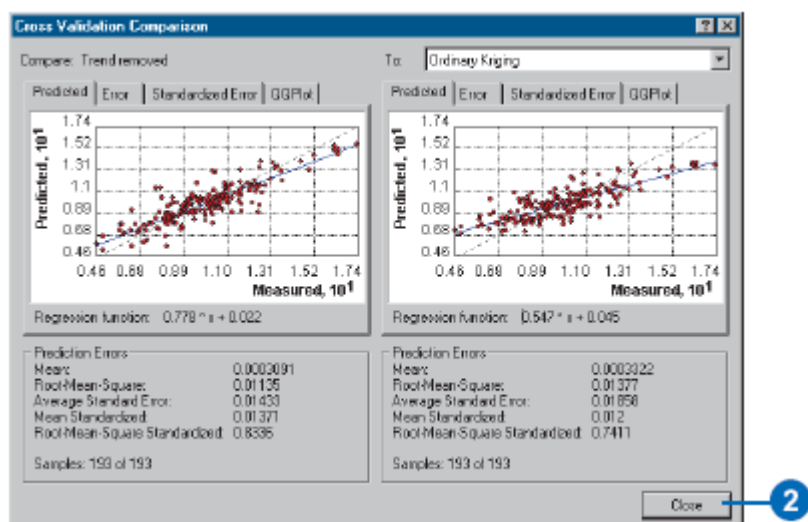
利用地统计分析模块，你可以对两种成图结果进行对比。结合交叉验证统计表，你可以判断哪个结果的预测更精确。

1. 在 “Trendremoved” 层上右击，在快捷菜单中点击 “Compare...”，你就能够将 “Trendremoved” 层与你在练习 2 中创建的 “Default” 层进行比较。

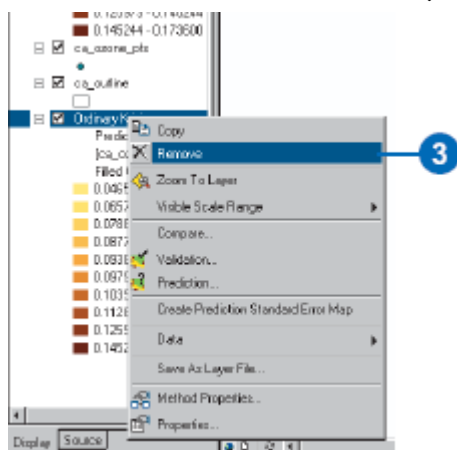


因为 “Trendremoved” 层的均方根预测误差较小，其均方根标准预测误差接近于 1，而平均预测误差则接近于 0，所以你有理由相信 “Trend removed” 模型更优更有效。所以你现在不再需要 “default” 层了，可以把它移去。

2. 在 CrossValidationComparison(交叉验证对话框)中点击 Close 按钮。



3. 在“Default”层上右击，在快捷菜单中选择 Remove(移去)。



4. 点击“Trendremoved”层，把它移到内容列表的底部，这样你就能够看见采样点及加州轮廓线。

5. 在 Standard 工具条上单击 Save 按钮。

现在你已经找出最佳预测表面，但你还想生成几种其他类型的表面。

练习 5：创建臭氧超出某一临界值的概率图

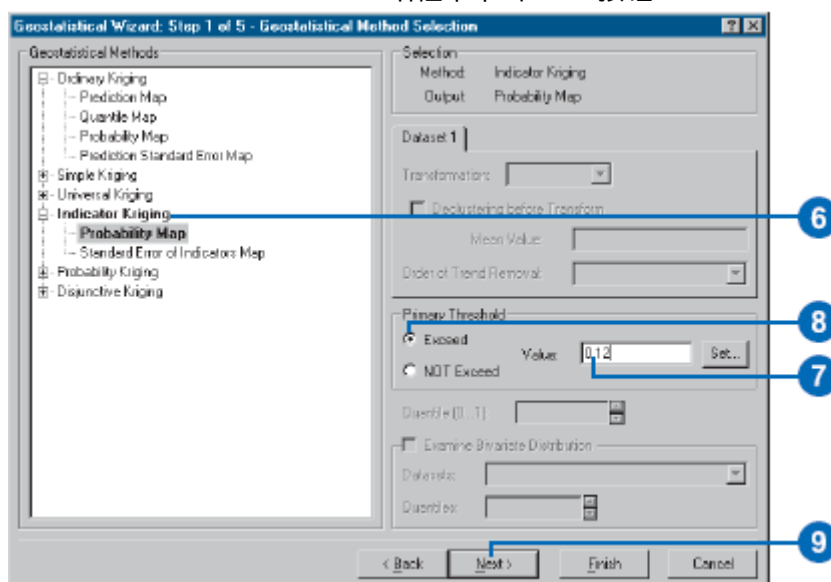
在练习 1 和练习 3 中，你使用不同参数的普通克里格法创建了臭氧浓度图。在决策阶段，利用预测得到的臭氧图来识别危险区域时，务必要谨慎，因为你需要了解预测中的不确定性。例如，对于一个 8 小时的时段，假设臭氧的临界值是 0.12ppm，你可能想判断山那些超出该临界值的地区。你可以使用地统计分析模块来生成臭氧浓度超出临界值的概率图，这可以对你的决策过程给予帮助。

尽管地统计分析模块提供了大量的可以完成此任务的工具，但在本练习你使用的是指示克里格法。这种方法不要求数据集一定要服从某种特定分布。根据数据值是高于或者低于一个临界值来将数据值转换为一系列的 0 和 1。如果利用 0.12ppm 作为临界值的话，任何低于它的数据值都将被赋予 0，而高于它的值则被赋予 1。然后指示克里格法使用一个根据转换后的 0—1 数据集计算得到的半变异函数模型进行计算。

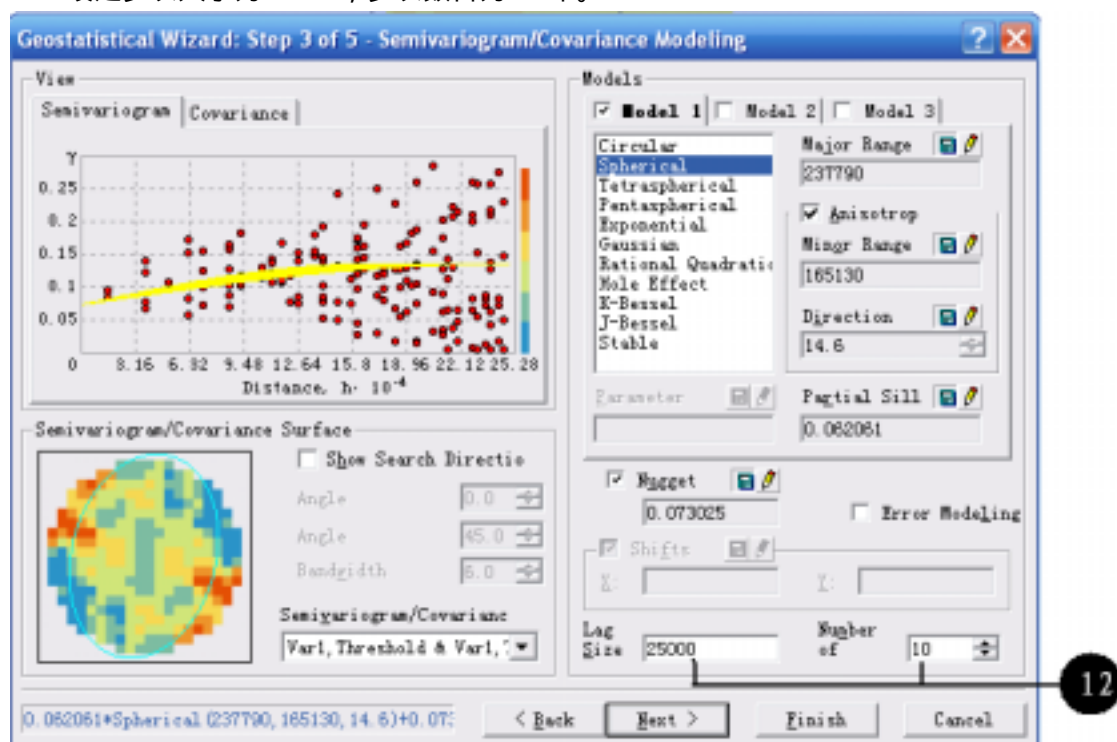
1. 点击 Geostatistical Analyst 工具条，然后单击 Geostatistical Wizard。

2. 单击 LayerF 拉箭头，选择 ca_ozone_pts。

3. 单击 Attribute 下拉箭头，选择属性 OZONE。
4. 在 Method 框中单击 Kriging。
5. 在 Choose Input Data and Method 对话框中单击 Next 按钮。
6. 单击 Indicator Kriging，注意在其下一级菜单中，概率图被选中。
7. 设置 Primary Threshold Value(主临界值)为 0.12ppm。
8. 单击选中 Exceed radial 单选按钮。
9. 在 Geostatistical Method Selection 对话框中单击 Next 按钮



10. 在 Additional Cutoffs Selection 对话框中单击 Next 按钮
11. 单击 Anisotropy，说明数据的方向性。
12. 设定步长大小为 25000，步长数为 10 个。



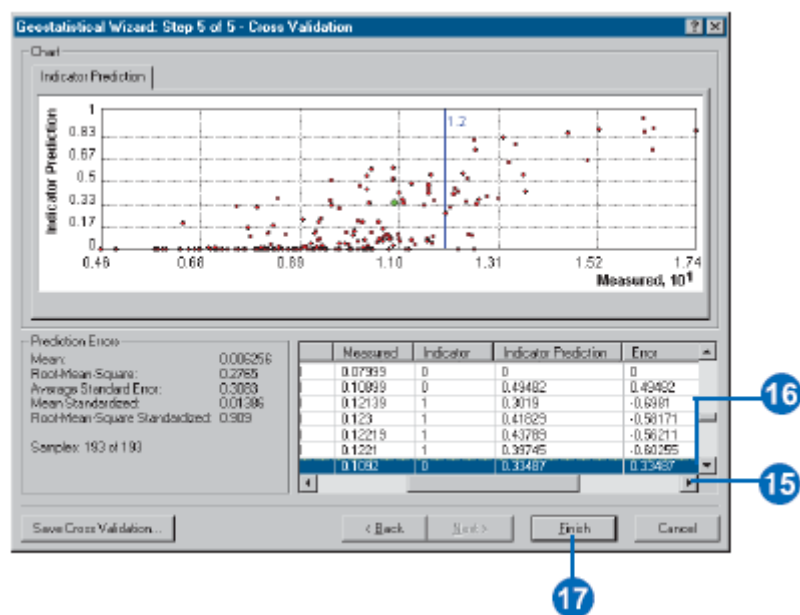
13. 在 Semivariogram / Covariance Modeling 对话框中单击 Next 按钮。
14. 对 Search Neighborhood 对话框中单击 Next 按钮。

蓝线代表临界值(0.12ppm)，在其左边的点的指示变换值为 0，而其右边的点的指示变换值则为 1。

15. 单击并向右滚动窗口，显示 Measured, Indicator, 及 Indicator Prediction 等栏

16. 点击并高亮显示表中指示值为 0 的那一行，则该点将在散点图中以绿色高亮显示在蓝色临界值的左边。

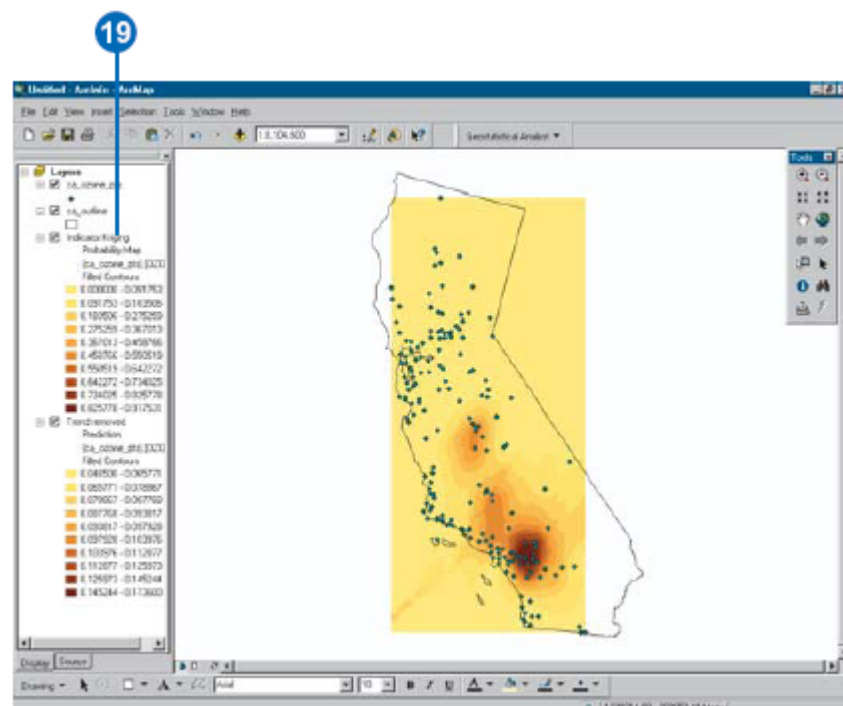
测定和指示栏显示的是每个样点的实际的和转换过的值。指示预测值表示超出临界值的概率，该值是利用二值(0,1)数据的半变异函数模拟计算得到的，并作为原始数据的指示转换。交叉验证函数按顺序每次省略一个点，然后计算该省略点的指示预测值。举个例子，假设最大实测值为 0.1736，如果该点实际上没有被测量，则根据指示克里格模拟算出的预测值超出临界值的可能性约为 85%。



17. 在 Cross Validation 对话框中点击 Finish 按钮。

18. 在 Output Layer Information 对话框中点击 OK 按钮。

所得的概率图将显示在 ArcMap 数据视图窗口的顶层。该图显示的指示预测值是指超出临界值 0.12ppm 的概率，时间是 1996 年的某一天或某几天。



从图中可以清楚看出，在洛杉矶附近，臭氧浓度很有可能超出安全值(一般说来，一年中的任意一个 8 小时时段的浓度值小于 0.12ppm 为安全值)。

19. 点击并按住 IndicatorKriging 层，将其拖到 ca_outline 层和 trendremoved 层之间。

练习 6 教你如何在 ArcMap 中，利用你在练习 3 创建的预测表面及在本练习中创建的概

率表面图来生成一幅符合制图规范的令人满意的地图。

练习 6：生成最终成果图

现在你可以生成一幅显示用的最终成果图。你要用 ArcMap 来生成一幅包含预测表面及概

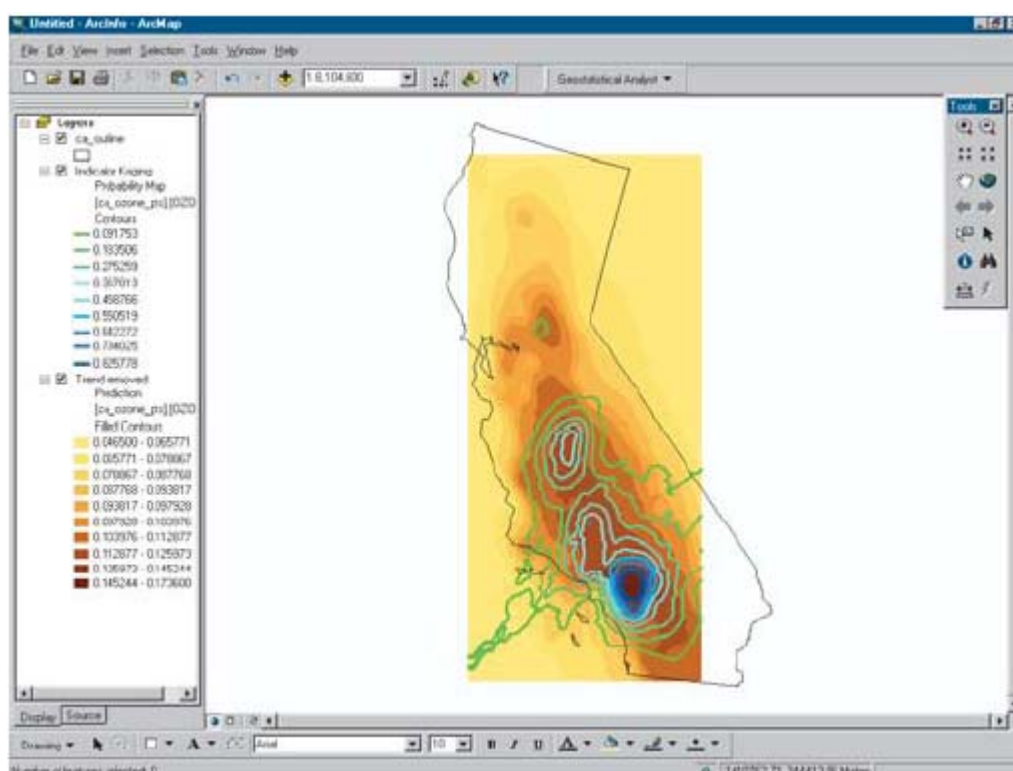
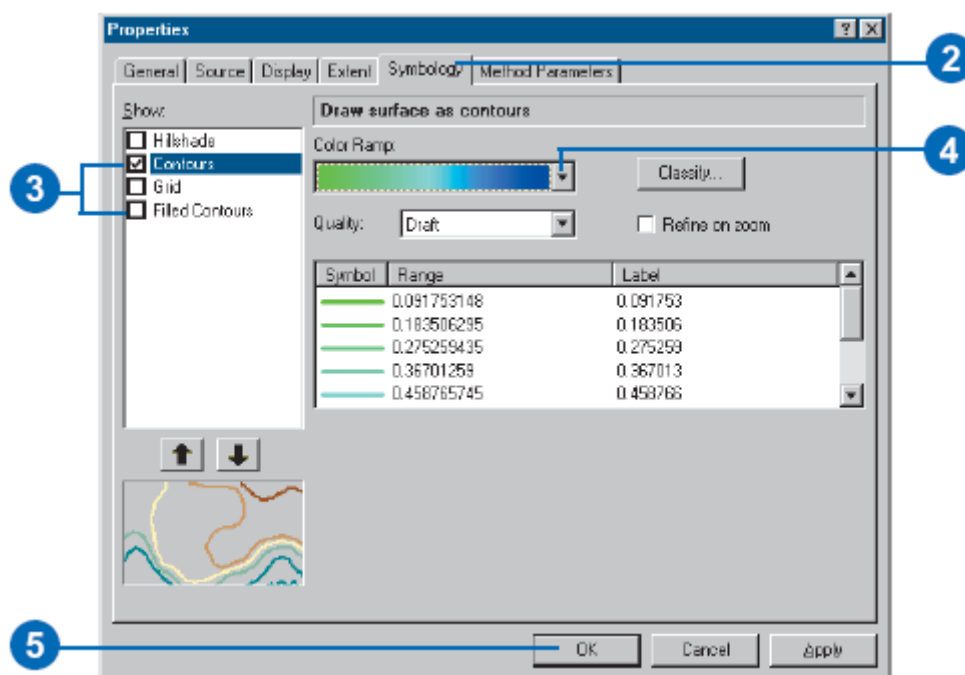
率表面的最终成果图。

6.1 显示两个表面

调整概率图的显示使你可以同时看见概率图和预测图。概率值将以等高线的形式显示。

1. 在 IndicatorKriging 层上右击鼠标，单击 Properties。
2. 点击 Symbology 标签。
3. 取消 FilledContours 复选框，然后选中 Contours 复选框。
4. 点击 ColorRamp 下拉箭头，选择一种别的颜色梯度。
5. 点击 OK 按钮。

现在你可以同时看见概率图(等高线)和预测图，如下图所示。



6.2 外推臭氧值

缺省时，地统计分析模块对落于采样数据的南北和东西范围内的任一位置上的可选变量的值进行内插。但是，预测得到的臭氧图并不能覆盖加州全部地理范围(图层 ca_outline)。要解决这个问题，你必须对两个表面进行外推求值(即预测缺省边界之外的点的值)。

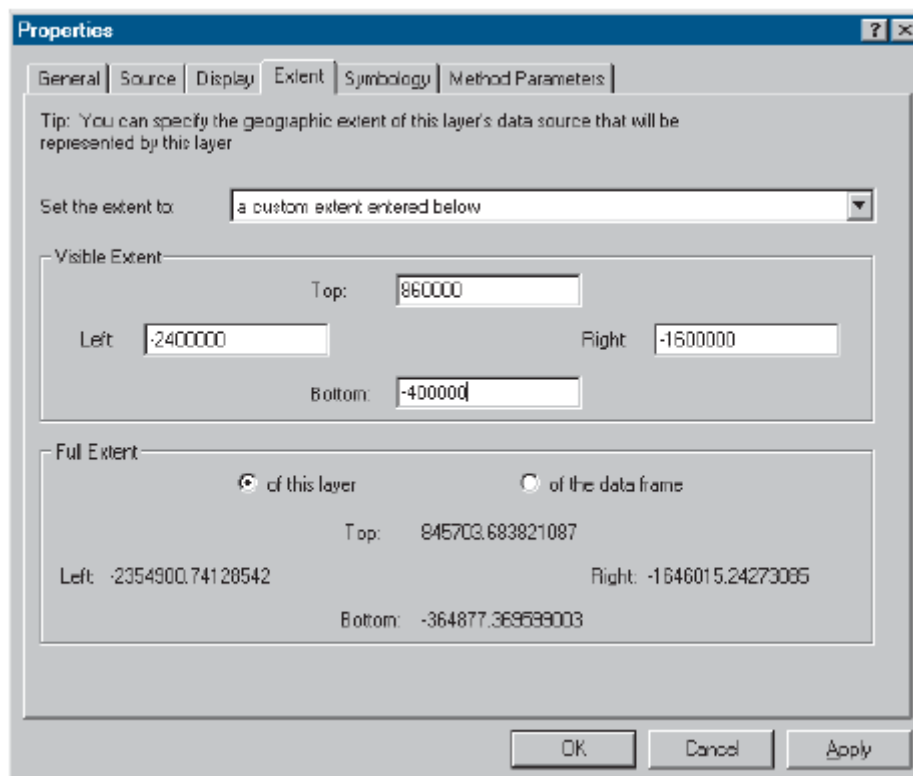
1. 在目录表中右击 IndicatorKriging 层，再单击 Properties。在 Properties 对话框

中点击 Extent 标签。在 Set the extent to 框中选择 a custom extent entered below，在 Visible Extent 区域中输入下面的值，然后点击 OK 按钮。

左：-2400000 右： -1600000

上： 860000 下： -400000

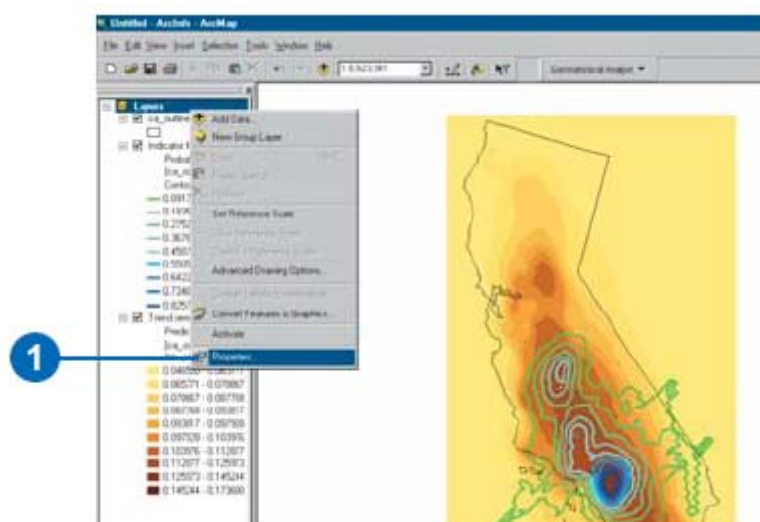
对图层 Trendremoved 重复上述步骤。



6.3 以加州轮廓图为参考对图层进行剪裁

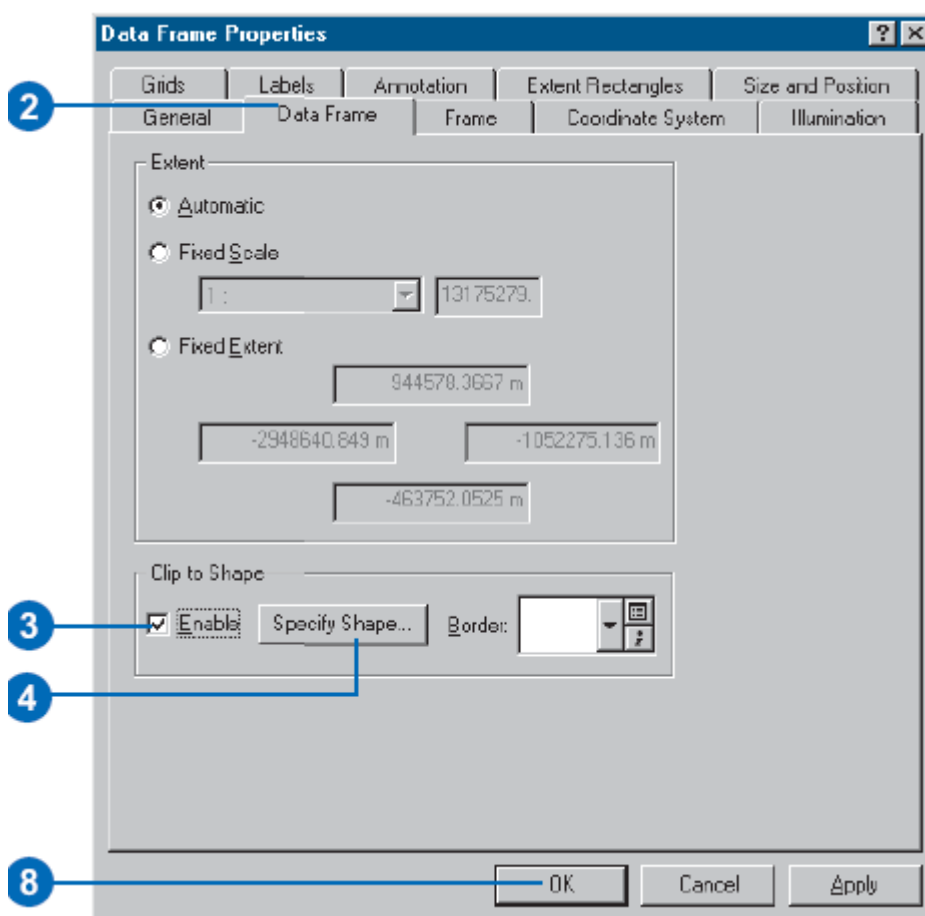
因为你只想制作加州范围内的臭氧浓度图，所以你可以以图层 ca_outline 为标准对图层进行剪裁，以生成一幅更吸引人的地图。

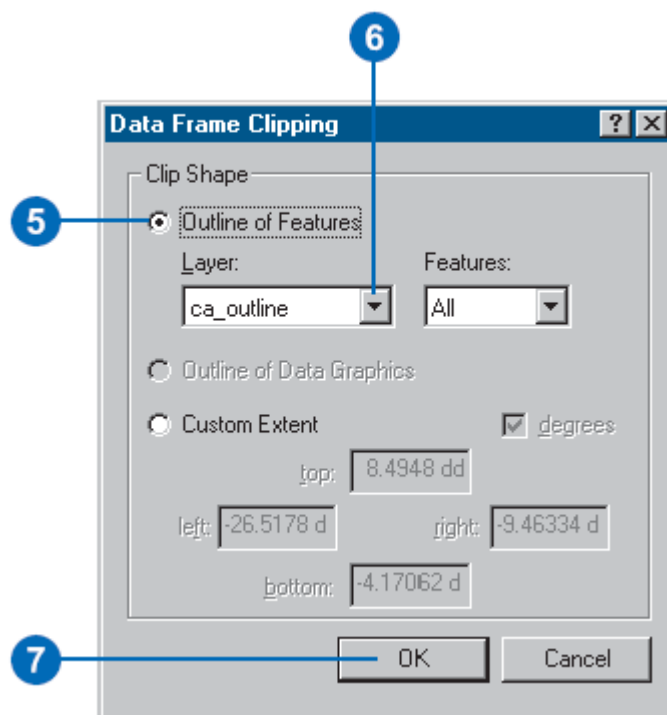
1. 在目录中右击 Layer，然后单击 Properties。



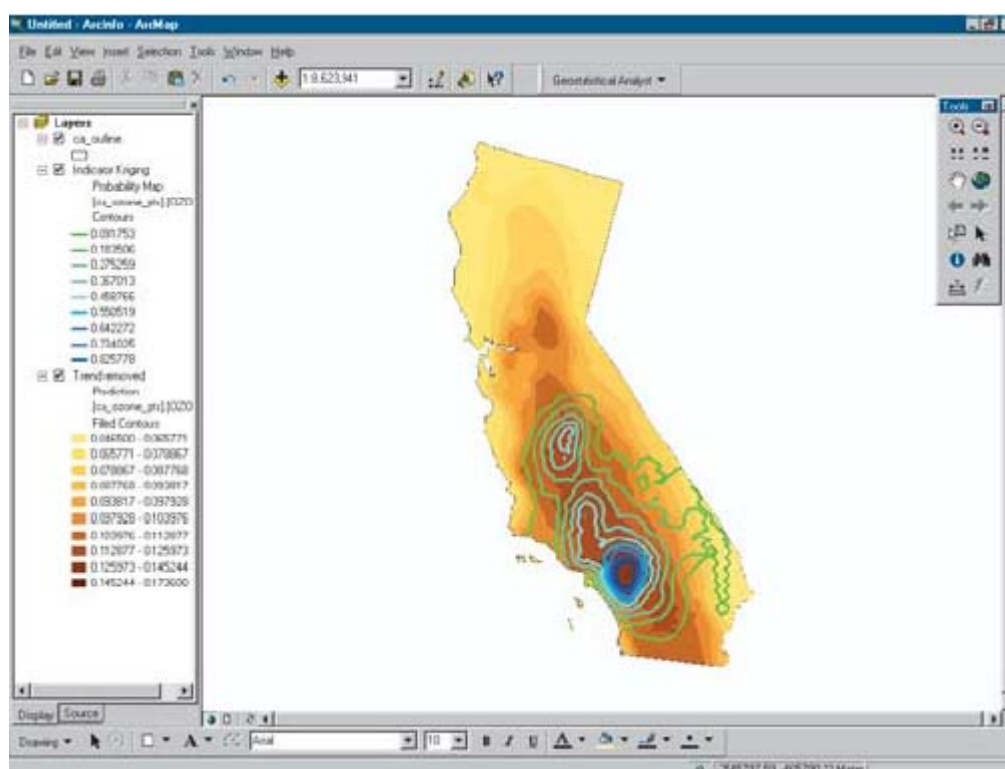
2. 在 Properties 对话框中单击 DataFrame 标签。
3. 选中 EnableClipToShape 复选框。
4. 点击 Specify Shape 按钮。

5. 在弹出的对话框中单击 Outline of Features 按钮。
6. 在 Layer 下拉框中单击 ca_outline。
7. 点击 OK 按钮。
8. 单击 OK 按钮以关闭 Data Frame Properties 对话框。





裁剪后的地图如下图所示。

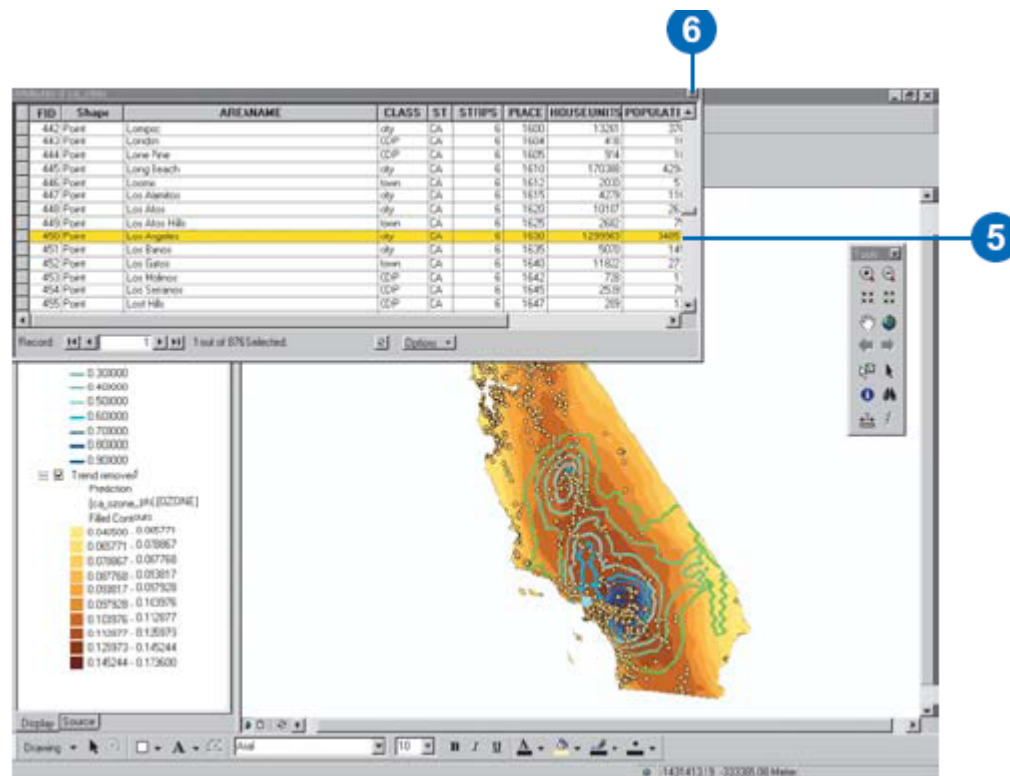


6.4 定位洛杉矶城

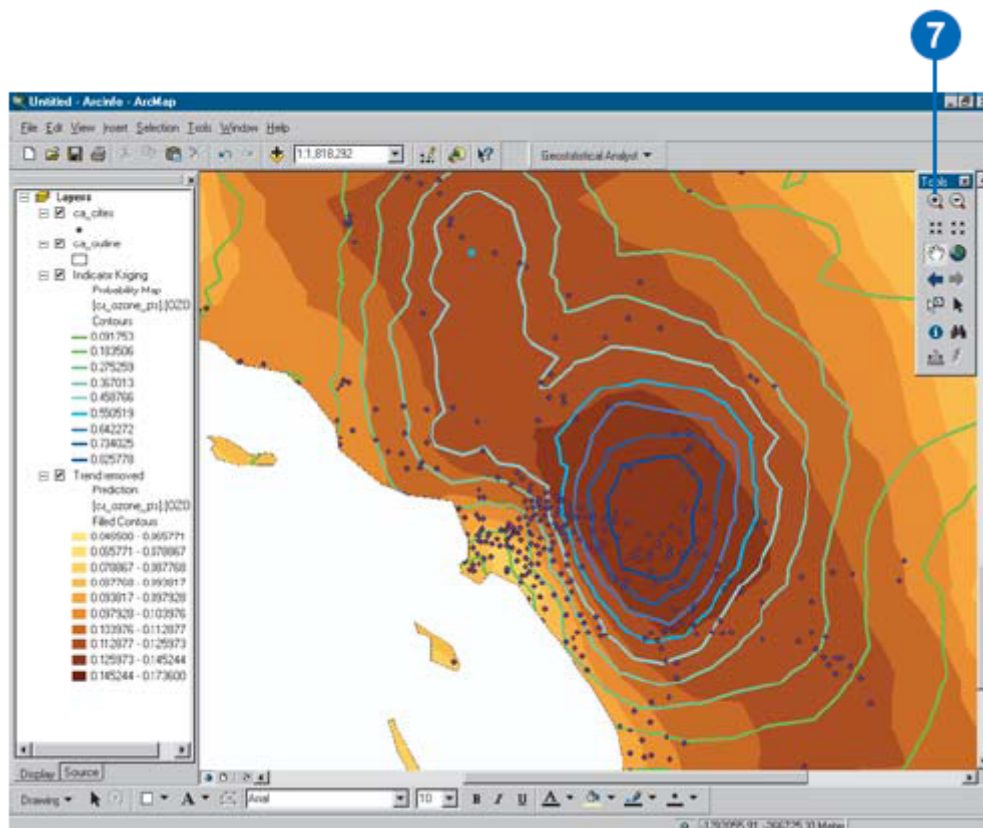
1. 在 Standard 工具条上单击 AddData 按钮。
2. 找到你安装快速入门练习数据的文件夹(缺省安装路径是 C:\ArcGIS\ArcTutor\Geostatistics), 然后选中 ca_cities。
3. 点击 Add 按钮。一幅加州城市位置图即显示出来。
4. 在图层 ca_cities 上右击, 在快捷菜单中单击 OpenAttributeTable。

5. 在表中拖动滚动条, 找到字段 AreaName 的值为 Los Angeles 的那一行并单击, 洛杉矶即在地图中高亮显示。

6. 点击关闭属性表。

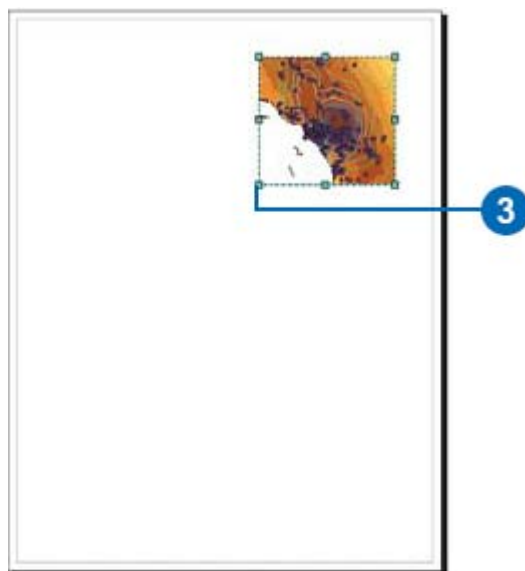


7. 在 T00LS 工具条上点击 Zoom in 工具, 将 City of Los Angeles 放大显示。可以看到山具有最大臭氧浓度值的地区实际上位于洛杉矶的东部。



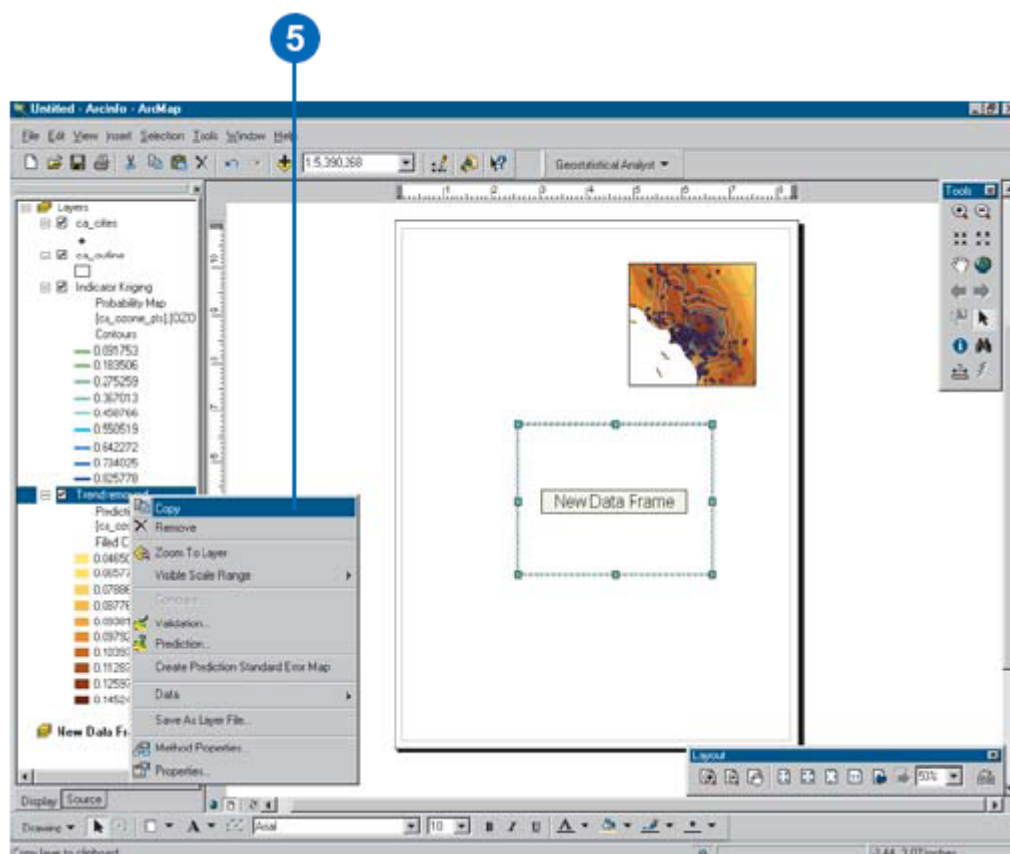
6.5 创建图版

1. 在主菜单上单击 View，然后单击 Layout View。
2. 单击地图以高亮显示。
3. 单击并拖动数据框的左下角以改变地图的大小。

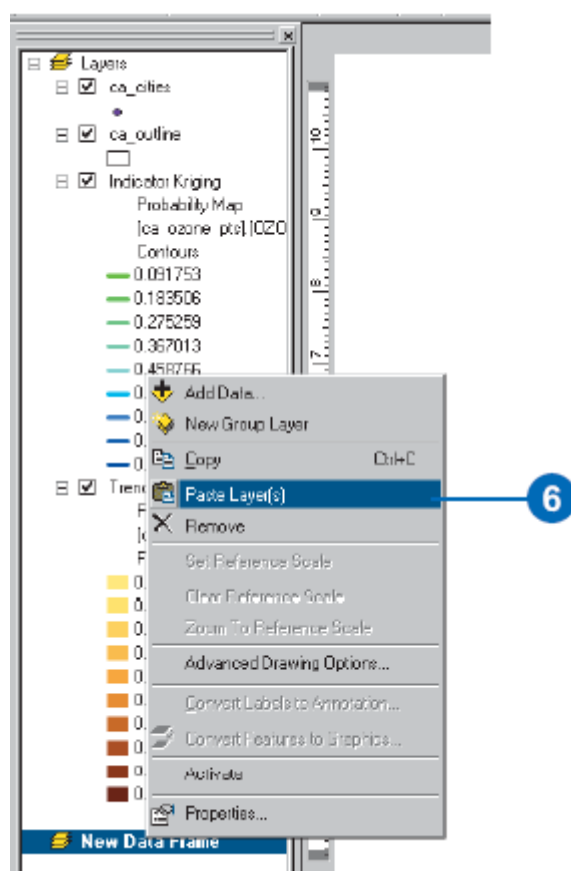


4. 在主菜单上单击 Insert，再点击 DataFrame。

一个新的数据框就插入到地图中。现在你可以将第一个数据框中的所有图层拷贝到这个新的数据框中以在臭氧图的旁边显示全加州的臭氧浓度值图，该图以洛杉矶为中心进行放大。

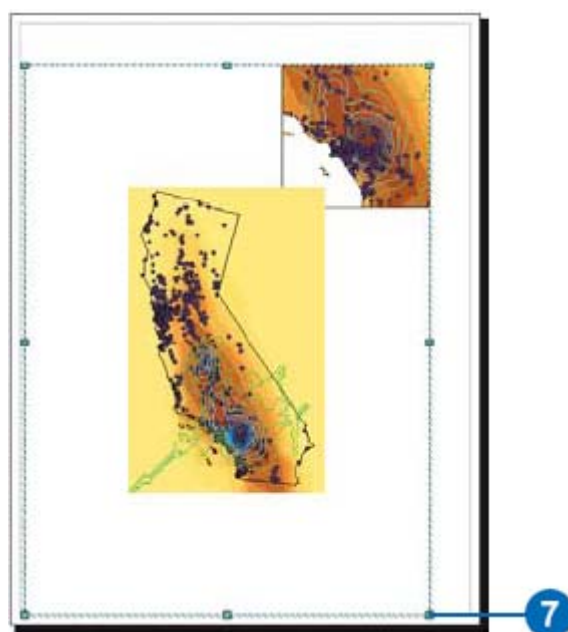


5. 在图层 Trendremoved 上右击，在快捷菜单中单击 Copy 选项。
6. 在目录表中右击 NewDataFrame，然后点击 PasteLayer(s)。



对所有其他图层重复步骤 5 和步骤 6 操作。

7. 点击并拖动新数据框以充满整个页面。

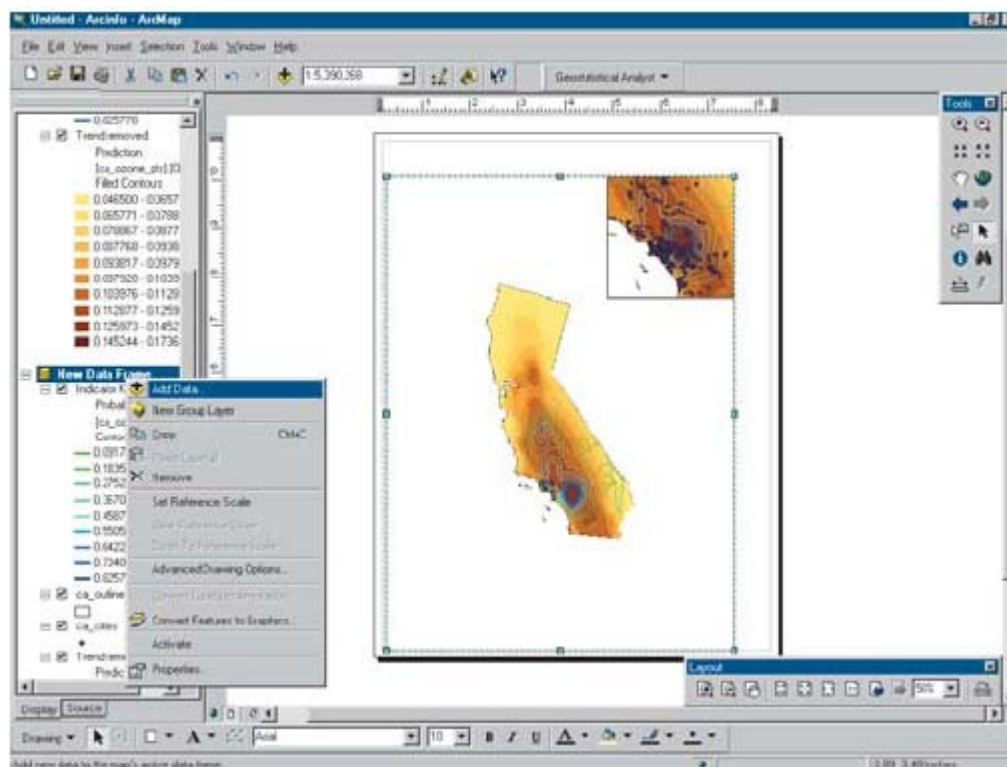


8. 在 Tools 工具条上单击 FullExtent 按钮来查看新数据框中地图的整幅范围。
9. 右击 NewDataFrame 并点击 Properties。

10. 与你在对第一个数据框进行的操作相同, 点击 DataFrame 标签, 选中 EnableClip to Shape 复选框并单击 SpecifyShape 按钮, 选择图层 ca_outline 用于剪裁, 然后点击 OK 按钮。

6.6 添加一个透明的山体阴影图

1. 在 NewDataFrame 上右击, 在快捷菜单中单击 AddData 选项。

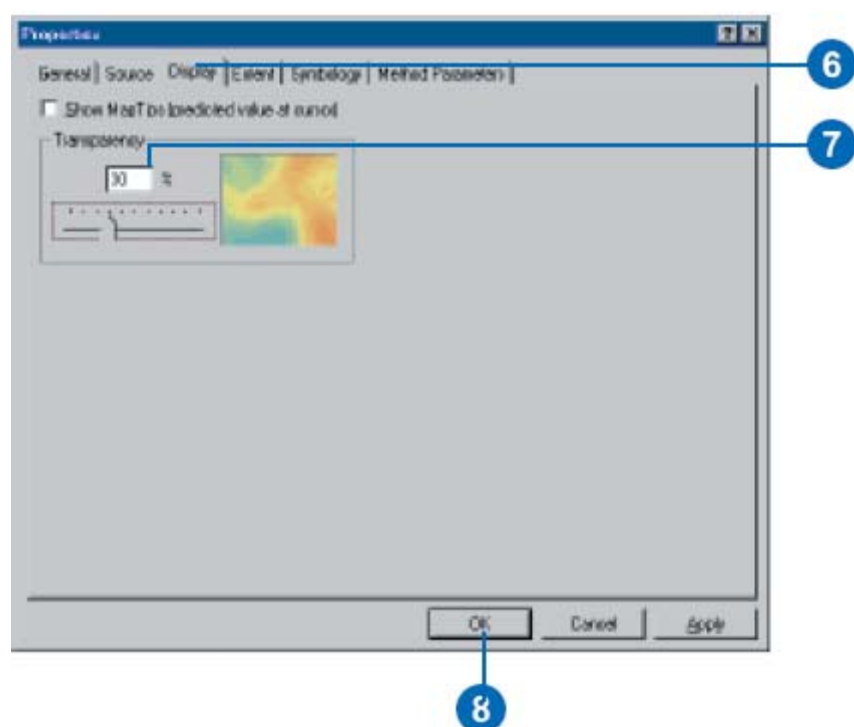


2. 找到你安装快速入门数据的文件夹(缺省安装路径是 C:\ArcGIS\ArcTutor\Geostatistics), 然后选中 ca_hillshade。

3. 点击 Add 按钮。一幅加州山体阴影图即显示出来。

4. 点击图层 ca_hillshade, 将其移动到内容列表的底部。

5. 在 NewDataFrame 目录表中右击 Trendremoved 层, 在快捷菜单中选择 Properties。



6. 单击 Display 标签。

7. 输入透明度为 30。

8. 单击 OK 按钮。

这样山体阴影图就被透明地显示在 Trendremoved 层的下面。

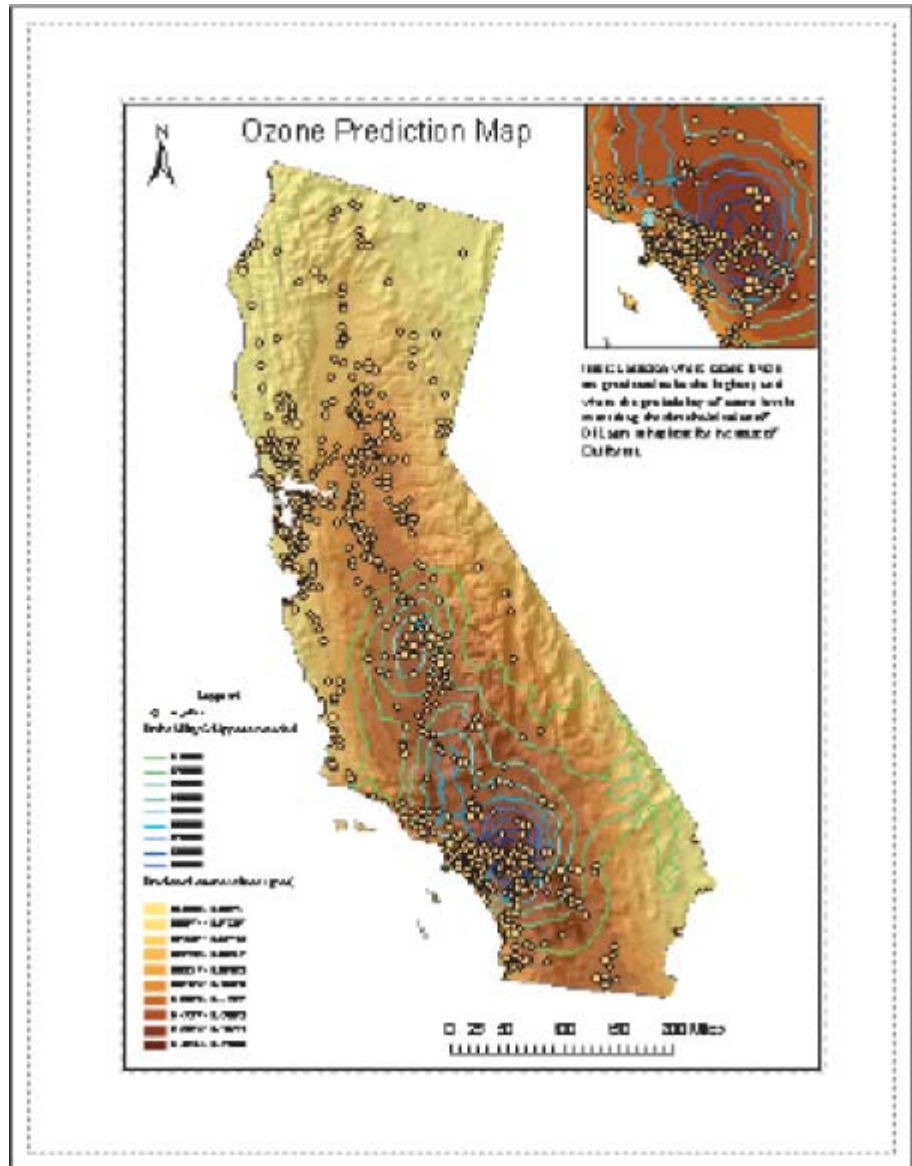
6.7 添加地图要素

1. 在主菜单上单击 Insert，单击 Legend。

2. 将图例移到布局的左下角。

3. 如果你愿意，你还可以单击 Insert 菜单以添加一个指北针，比例尺以及文本等。

下图显示的是你利用 ArcMap 创建的成果地图。如果你需要了解如何在布局中插入地图要素，请参阅“Using ArcMap”。



从图上可以看出，洛杉矶以东地区的臭氧预测值最高，并且在 1996 年中，至少有一个 8 小时时段该地区的臭氧值极有可能超出临界值(0.12ppm)。由于这是分析中的一个案例(但要记住原始数据被改动过)，你可以把精力放在这些地区，分析臭氧的时间序列值以精确确定出那些具有潜在风险的地区。