

文章编号: 1001- 1749(2002)02- 0157- 06

等值线绘图软件 SURFER7. 0 中九种插值法介绍

白世彪, 陈 晔, 王 建

(南京师范大学 地理科学学院海岸与第四纪研究所, 江苏 南京 210097)

摘 要: SURFER 软件是一个功能强大的绘制等值线图及三维立体图软件包,能迅速地将离散的测量数据通过插值转换为连续的数据曲面。SURFER7. 0 提供的内插方法多达九种,其中每一种内插方法都有其意义及相关的参数设置。作者在本文中主要介绍了该软件中的九种内插方法,并进行了应用示例。

关键词: SURFER 软件; 插值方法; 地学应用

中图分类号: TP316 **文献标识码:** A

AN INTRODUCTION TO NINE GRIDDING METHODS AND THEIR APPLICATION IN SURFER VERSION 7. 0

BAI Shi-biao, CHEN Ye, WANG Jian

(Coastal and Quaternary Institute of Geo- Science School, Nanjing 210097, PRC)

Abstract SURFER is a full- function 3D surface modeling package, of which sophisticated interpolation engine can transform the dispersed X, Y, Z data into published- quality maps. There are nine gridding methods to choose in SURFER version 7. 0 and with each gridding method one can completely control over the gridding parameters, thus generate the high accuracy contour maps by computer.

Key words SURFER; gridding method; geosciences

0 引言

在地学的研究工作中,不可避免地要应用某种手段获取能够反映研究对象各方面特性或特征的数据,例如磁力、重力、化学勘探数据、钻探获得的岩面标高、海底或陆地的地形标高和地理要素指标等等。为了方便深入研究,要将数据处理并以等值线图的形式呈现。目前,绘制等值线的软件不少,但是,它们往往是为满足各自的需要而以小程序的形式开发出来,采用单一的插值方法,功能比较简单,适用性不广。这里介绍的 SURFER 是美国 GOLDEN 软件公司的产品,用于绘制等值线图及相应的三维图形,是科技工作者特别是地学研究人员必备的软件,应用极广。SURFER 目前的最高版本为 7. 03,它提供的内插方法多达九种,用户可根据不同的情况选用不同的内插方法。一般可使用 SURFER 默认的设置进行内插生成网格文件,但要科学地选择内插方法和灵活地进行参数设置,必须要熟悉各种内插方法的基本理论知识,下面将逐一介绍这九种内插方法。

1 SURFER九种插值法介绍

1.1 反距离加权插值法

反距离加权插值法 (Inverse Distance to a Power) 首先是由气象学家和地质工作者提出的, 后来由于 D. Shepard 的工作被称为谢别德法 (Shepard) 方法^[1]。它的基本原理是设平面上分布一系列离散点, 已知其位置坐标 (x_i, y_i) 和属性值 $z_i (i=1, 2, \dots, n)$, $p(x, y)$ 为任一格网点, 根据周围离散点的属性值, 通过距离加权插值求 P 点属性值。距离加权插值法综合了泰森多边形的邻近点法和多元回归法的渐变方法的长处, 它假设 P 点的属性值是在局部邻域内中所有数据点的距离加权平均值, 可以进行确切的或者圆滑的方式插值。周围点与 P 点因分布位置的差异, 对 $P(z)$ 影响不同, 我们把这种影响称为权函数 $w_i(x, y)$, 方次参数控制着权系数如何随着离开一个格网结点距离的增加而下降。对于一个较大的方次, 较近的数据点被给定一个较高的权重份额; 对于一个较小的方次, 权重比较均匀地分配给各数据点。计算一个格网结点时, 给予一个特定数据点的权值, 与指定方次的结点到观测点的距离倒数成比例。当计算一个格网结点时, 配给的权重是一个分数, 所有权重的总和等于 1.0。当一个观测点与一个格网结点重合时, 该观测点被给予一个实际为 1.0 的权重, 所有其它观测点被给予一个几乎为 0.0 的权重^[2,3]。换言之, 该结点被赋给与观测点一致的值, 这就是一个准确插值。权函数主要与距离有关, 有时也与方向有关, 若在 P 点周围四个方向上均匀取点, 那么可不考虑方向因素, 这时:

$$P(Z) = \frac{\sum_{i=1}^n \frac{Z_i}{[d_i(x, y)]^u}}{\sum_{i=1}^n \frac{1}{[d_i(x, y)]^u}}$$

式中 $d_i(x, y) = \sqrt{(x - x_i)^2 + (y - y_i)^2}$, 表示由离散点 (x_i, y_i) 至 $P(x, y)$ 点的距离。 $p(z)$ 为要求的待插点的值。权函数, $w_i(x, y) = 1/[d_i(x, y)]^u$, u 值一般取为 2。

反距离加权插值法是 GIS 软件根据点数生成规则格网数据文件的最常见的方法, 计算值易受数据点集群的影响, 计算结果常出现一种孤立点数据明显高于周围数据点的“鸭蛋”分布模式, 可在插值过程中通过动态修改搜索准则进行一定程度的改进。使用反距离加权插值法对数据进行格网化时, 可以指定一个圆滑参数。圆滑参数要求大于零。对于一个特定的结点, 没有哪个观测点被赋予全部的权值, 即使观测点与该结点重合也是如此。圆滑参数通过修匀已被插值的格网来降低“鸭蛋”影响。

1.2 克里金插值法

克里金 (Kriging) 插值法又称空间自协方差最佳插值法, 它是以法国 D. G. Krige 的名字命名的一种最优内插法。克里金法广泛地应用于地下水模拟、土壤制图等领域, 是一种很有用的地质统计格网化方法。它首先考虑的是空间属性在空间位置上的变异分布, 确定对一个待插点值有影响的距离范围, 然后用此范围内的采样点来估计待插点的属性值。该方法在数学上可对所研究的对象提供一种最佳线性无偏估计 (某点处的确定值) 的方法。它是考虑了信息样品的形状、大小及与待估计块段相互间的空间位置等几何特征以及品位的空间结构之后, 为达到线性、无偏和最小估计方差的估计, 而对每一个样品赋与一定的系数, 最后进行加权平均来估计块段品位的方法。但它仍是一种光滑的内插方法。在数据点多时, 其内插的结果可信度较高^[2,3]。

克里金法类型分常规克里金插值 (常规克里金模型、克里金点模型) 和块克里金插值。常规克里金插值其内插值与原始样本的容量有关, 当样本数量较少的情况下, 采用简单的常规克里金模型内插的结果图会出现明显的凹凸现象; 块克里金插值是通过修改克里金方程以估计子块 B 内的平均值来克服克里金点模型的缺点, 对估算给定面积实验小区的平均值或对给定格网大小的规则格网进行插值比较适用。块克里金插值估算的方差结果常小于常规克里金插值, 所以, 生成的平滑插值表面不会发生常规克里金模型的凹凸现象。按照空间场是否存在漂移 (drift) 可将克里金插值分为普通克里金和泛克里金, 其中普通克里金 (Ordinary Kriging 简称 OK 法) 常称作局部最优线性无偏估计, 所谓线性是指估计值是样本值的线性组合, 即加权线性平均, 无偏是指理论上估计值的平均值等于实际样本值的平均值, 即估计的平均误差为 0, 最优是指估计的误差方差最小^[4]。

1.3 最小曲率法

最小曲率法 (Minimum Curvature) 广泛应用于地球科学。用最小曲率法生成的插值面类似于一个通过各个数据值、具有最小弯曲量的长条形薄薄的弹性片。最小曲率法试图在尽可能严格地尊重数据的同时, 生成尽可能圆滑的曲面^[5]。最小曲率法不是一个精确的插值法, 也就是说在插值的过程中不可能总是完全尊重数据。

使用最小曲率法时要涉及到两个参数: 最大偏差参数 (Maximum Residuals) 和最大循环次数 (Maximum Iteration parameter) 参数来控制最小曲率的收敛标准, 而且最小曲率法要求至少有四个点。

1.4 改进谢别德法

改进谢别德法 (Modified Quadratic Shepard) 是由 Franke 及 Nielson 提出, 它仍是一个与距离成反比的加权方法。在使用反距离加权插值法时, 当增加、删除或改变一个点时, 需要重新计算权函数 $w_i(x, y)$, 为了克服反距离加权插值法的这一缺陷, 改进谢别德法同样使用距离倒数加权的最小二乘方的方法, 但有以下两个方面的改进^[6~8]。

(1) 通过修改反距离加权插值法的权函数 $w_i(x, y) = 1/[d_i(x, y)]^r$, 使其只能在局部范围内起作用, 以改变反距离加权插值法的全局插值性质, 即它利用了局部最小二乘方法来消除或减少所生成等值线的“鸭蛋”外观。

(2) 同时用节点函数 $Q(x, y)$ 来代替离散点 (x_i, y_i) 的属性值 z_i , $Q(x, y)$ 是一个插值于 (x_i, y_i) 点的二次多项式, 即有 $Q(x_i, y_i) = z_i (i = 1, 2, \dots, n)$ 。而且 $Q(x, y)$ 在点 (x_i, y_i) 附近与函数属性值 $z(x, y)$ 具有局部近似的性质。因此, 如果认为距离 (x_i, y_i) 较远的点对 $Q(x, y)$ 影响不大, 则可以认为在 (x_i, y_i) 点附近, $Q(x, y)$ 就可以近似地表示函数属性值 $z(x, y)$ 了。改进谢别德法可以是一个准确或圆滑插值器。

在用改进谢别德法作为格网化方法时要涉及到圆滑参数的设置。圆滑参数是使改进谢别德法能够像一个圆滑插值器那样工作, 增加圆滑参数的值可增强圆滑的效果。

1.5 自然邻点插值法

自然邻点插值法 (Natural Neighbor) 是 Surfer 7.0 才有的网格化新方法。自然邻点插值法广泛应用于一些研究领域。其基本原理是对于一组泰森 (Thiessen) 多边形, 当在数据集中加入一个新的数据点 (目标) 时, 就会修改这些泰森多边形, 而使用邻点的权重平均值将决定待插点的权重, 待插点的权重和目标泰森多边形成比例^[9]。实际上, 在这些多边形中, 有一些多边形的尺寸将缩小, 并且没有一个多边形的大小会增加。同时, 自然邻点插值法在数据点凸起的位置并不外推等值线 (如泰森多边形的轮廓线)。

1.6 最近邻点插值法

最近邻点插值法 (Nearest Neighbor) 又称泰森多边形方法, 泰森多边形 (Thiessen, 又叫 Dirichlet 或 Voronoi 多边形) 分析法是荷兰气象学家 A. H. Thiessen 提出的一种分析方法。最初用于从离散分布气象站的降雨量数据中计算平均降雨量, 现在 GIS 和地理分析中经常采用泰森多边形进行快速的赋值^[2]。实际上, 最近邻点插值的一个隐含的假设条件是任一网格点 $p(x, y)$ 的属性值都使用距它最近的位置点的属性值, 用每一个网格节点的最邻点值作为待的节点值^[3]。当数据已经是均匀间隔分布, 要先将数据转换为 SURFER 的网格文件, 可以应用最近邻点插值法; 或者在一个文件中, 数据紧密完整, 只有少数点没有取值, 可用最近邻点插值法来填充无值的数据点。

有时需要排除网格文件中的无值数据的区域, 在搜索椭圆 (Search Ellipse) 设置一个值, 对无数据区域赋予该网格文件里的空白值。设置的搜索半径的大小要小于该网格文件数据值之间的距离, 所有的无数据网格节点都被赋予空白值。在使用最近邻点插值网格化法, 将一个规则间隔的 XYZ 数据转换为一个网格文件时, 可设置网格间隔和 XYZ 数据的数据点之间的间距相等。

最近邻点插值网格化法没有选项, 它是均质且无变化的, 对均匀间隔的数据进行插值很有用, 同时, 它对填充无值数据的区域很有效。

1.7 多元回归法

多元回归 (Polynomial Regression) 是用来确定数据的大规模的趋势和图案。多元回归实际上不是插值器, 因为它并不预测未知的 z 值, 它只是根据空间的采样数据, 拟合一个数学曲面, 用该数学曲面来反映空间分布的变化情况, 它实际上是一个趋势面分析作图程序。

趋势面分析是对地质特征的空间分布进行研究和分析的一种方法,它是用某种形式的函数所代表的曲面来逼近该地质特征的空间分布。这个函数从总体上反映了采样数据的区域性变化趋势,称为趋势面部分;采样数据的实测值与这个函数对应值之差,称为偏差部分,它反映了局部性的变化。这就是说,把采样数据的实测值分解成两部分,趋势面部分和偏差部分,趋势面部分用一个函数表示,它反映采样数据的总体变化,可以认为是由大范围的系统性因素引起的;偏差部分反映了局部性的变化特点,可以认为由局部因素和随机因素引起的^[10]。

使用多元回归法进行趋势面分析要考虑两个方面的问题:一是趋势面函数(数学表达式)的确定;二是拟合精度的确定。通常用的趋势面函数主要是多项式趋势面,因为多项式能够逼近任意连续函数,因此,用多项式作趋势面能较好地反映连续变化的分布趋势,这在地质科学中常用到。一般说多项式次数越高,则趋势面与实测数据偏差越小,但是,并不能说它就与实际情况最符合,这还要在实践中检验,次数较高的趋势面只在采样点附近效果较好,在外推和内插的效果方面不好,因而在实际应用的效果并不理想。在实际应用中,对起伏变化比较缓和的简单采样数据配合次数较低的趋势面,就可以反映出区域背景;而变化复杂且起伏较多的采样数据要配合次数较高趋势面^[11]。

使用多元回归法时要涉及到曲面定义和指定 X 、 Y 的最高次数设置,在曲面定义中选择所需的多项式类型,可选用的曲面类型:简单平面(Simple planar surface)、双线性鞍(Bi-linear saddle)、二次曲面(Quadratic surface)、三次曲面(Cubic surface)和用户自定义多项式(User defined Polynomial)。参数设置是指定多项式方程中 X 、 Y 组元的最高次数。

1.8 径向基函数插值法

所谓径向基函数(Radial Basis Function),其基函数是由单个变量的函数构成的。一个点 (x, y) 的这种基函数的形式往往是 $h_i(x, y) = h(d_i)$,这里的 d_i 表示由点 (x, y) 到第 i 个数据点的距离。径向基函数插值法是多个数据插值方法的组合。根据生成一个圆滑曲面适应数据的能力,许多人认为其中的复二次函数是最好的方法^[7]。所有径向基函数插值法都是准确的插值器,它们都能尽量适应你的数据。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。

函数类型:最基本的函数类似于克里金中的方差图。当对于一个网格点插值时,这些函数为数据点规定了一套最佳权重。

基函数类型有:倒转复二次函数(Inverse Multiquadric):

$$B(h) = \frac{1}{h^3 + R^2}$$

复对数(Multilog): $B(h) = \log(h^2 + R^2)$

复二次函数(Multiquadratic): $B(h) = \frac{1}{h^2 + R^2}$

自然三次样条函数(Natural Cubic Spline): $B(h) = (h^2 + R^2)^{3/2}$

薄板样条法函数(Thin Plate Spline): $B(h) = (h^2 + R^2) \log(h^2 + R^2)$

式中 h 为表示由点 (x, y) 到第 i 个数据点的距离; R 参数是用户指定的平滑因子。

其中的复二次函数(Multiquadric)方法是由 R. L. Hardy 在 1971 年提出来的。它是最早提出并且应用得最为成功的一种径向基函数插值法。现在复二次函数(Multiquadric)方法在水文测量、大地测量、地质及采矿、地球物理等领域都得到了广泛应用,效果良好。在数据点数量不太大的情况下(例如,几百个数据点),计算也不太复杂。

薄板样条法(Thin Plate Spline)是由 R. L. Harder 及 R. N. Desmarais 在 1972 年提出来的,后来得到了 J. Duchon 及 J. Meinguet 等人进一步发展。从力学的观点看,这一方法的实质是使插值函数所代表的弹性薄板受限于插值点,并且具有最小的弯曲能量。

1.9 三角网线性插值法

三角网线性插值法(Triangulation with Linear Interpolation)使用最佳的 Delaunay 三角形,连接数据点间的连线形成三角形。原始数据点的连结方法是这样:所有三角形的边都不能与另外的三角形相交,其结果构成了一张由三角形拼接起来的覆盖网格范围的网。每一个三角形定义了一个覆盖该三角形内网格节点的面^[12]。三角形的倾斜和标高由定义这个三角形的三个原始数据点确定。给定三角形内的全部节点

都要受到该三角形的表面的限制。因为各个三角形都是用原始数据点来定义的,这样就把三角形和你的数据紧密联系起来

三角网/线性插值法将在网格范围内均匀分配数据,地图上稀疏的区域将会形成截然不同的三角面

1.10 九种插值方法应用示例

作者在本文中选用了在研究工作中得到的某组离散数据,要对数据所反映的地质特征的空间分布进行研究和分析,分别使用 SURFER中的九种不同的插值方法对离散数据进行网格化,得到规则的网格文件,然后分别绘制等值线图^[13],如图 1所示。图 1右边为原始的离散数据,左边为分别绘制的等值线图。

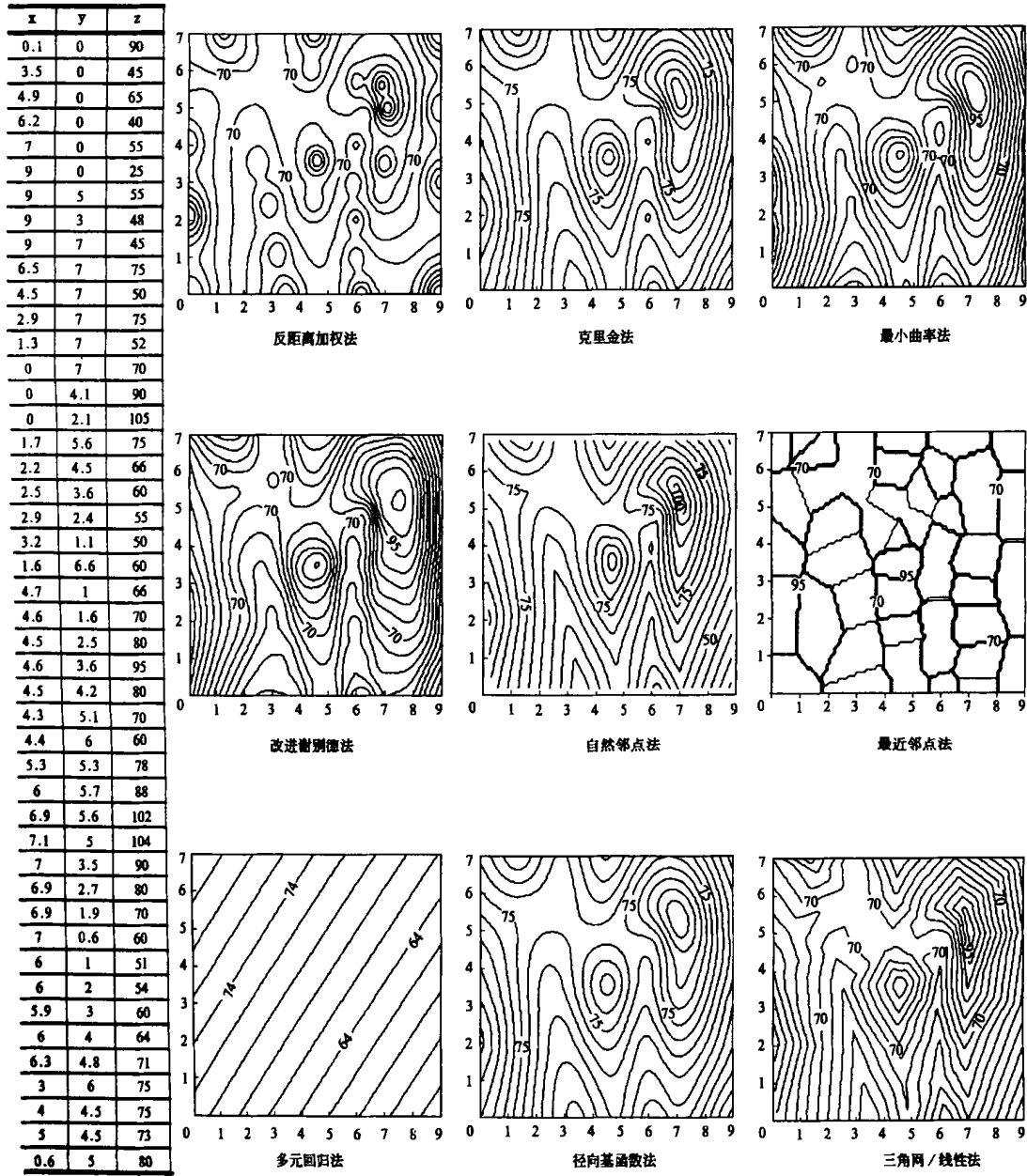


图 1 Surfer 九种不同插值方法生成的网格文件绘制的等值线图

Fig. 1 Contour maps drawn from the grid file by nine gridding methods in surfer

2 结论

SURFER7.0所采用的九种离散数据内插方法,几乎包括了目前所有的插值方法。该软件具有如此种类繁多的内插方法,这是其它同类软件所不能比拟的。一般可使用 SURFER默认的设置进行内插,生成网

格文件,然后再绘制所需的图件。熟悉了各种内插方法的基本理论知识,根据各种数据的不同特点,结合数据分析的目的,就能科学地选择内插方法,灵活地进行参数设置内插生成网格文件,在此基础上绘制正确的图件,供我们提取更多的地学信息。

致谢:感谢导师刘泽纯教授对本文的指导!

参考文献:

[1] Franke R. Scattered Data Interpolation: Test of Some Methods[J]. Mathematics of Computations, 1982, 33(157): 181.

[2] 邬伦,刘瑜,张晶,等.地理信息系统——原理、方法和应用[M].北京:科学出版社,2001.

[3] 陈述彭,鲁学军,周成虎.地理信息系统导论[M].科学出版社,2001.

[4] 王正权.地统计学及在生态学中的应用[M].北京:科学出版社,1999.

[5] Briggs I C. Machine Contouring Using Minimum Curvature[J], Geophysics, 1974, 39(1): 39.

[6] Franke R, Nielson G. Smooth Interpolation of Large Sets of Scattered Data[J]. International Journal for Numerical Methods in Engineering, 1980, 15(2): 1691.

[7] 唐泽圣.三维数据场可视化[M].北京:清华大学出版社,1999.

[8] Renka R. J. Multivariate Interpolation of Large Sets of Scattered Data[J]. ACM Transaction on Mathematical Software, 1988, 14(2): 139.

[9] Barnett V. Interpreting multivariate data[M]. New York, 1981: 21.

[10] 徐振邦,姜元仁.数学地质基础[M].北京:北京大学出版社,1994.

[11] 郭仁忠.空间分析[M].武汉:武汉测绘大学出版社,2000.

[12] Lee D T, Schachter B J. Two Algorithms for Constructing a Delaunay Triangulation. International Journal of Computer and Information Sciences[J]. 1980, 9(3): 219.

[13] 田黔宁.利用 WIN-surf 软件绘制物化探图件[J].物探化探计算技术,1998,20(4): 366.

作者简介:白世彪,(1974—),男,1999年毕业于兰州大学地理学基地班,现为南京师范大学地理科学学院硕士研究生,主要研究 GIS 技术及环境演变模拟。

(上接第 133 页)

5 结论

针对塔河油田六区石炭系碎屑岩储层特点,把地震资料和测井资料有机地结合起来,以地震、测井互为约束条件贯穿于整个反演过程,形成对塔河油田六区石炭系储层的横向预测方法技术。以地震地层学研究为基础,首先,利用砂岩的测井响应识别出砂岩,进行井约束下的地震反演;然后进行地震约束下的属性体反演。实际应用表明,利用反演的波阻抗和岩性反演剖面可有效地预测卡拉沙依组上、下部具有不同的储层分布规律,砂体的横向变化情况,为圈闭储层预测、圈闭条件分析时,提供有效的技术保证。

反演提高了地震资料的分辨能力,识别、分辨砂体的能力提高,具有较好的可视性,波形活跃,信噪比较高,为塔河油田在石炭系进一步开发展示了良好的前景。

参考文献:

[1] 顾家裕.沉积相与油气[M].北京:石油工业出版社,1994.

[2] 王捷.勘探阶段油藏描述技术[M].北京:石油工业出版社,1996.

[3] 刘震.储层地震地层学[M].北京:地质出版社,1997.

作者简介:张卫红,(1966—),女,工程师,从事 VSP 资料处理解释及地震资料反演工作。