

# 局部多项式插值方法在多源海底沉积厚度数据融合中的应用

汪 俊, 高金耀, 吴招才, 张 涛

(国家海洋局 第二海洋研究所, 浙江 杭州 310012)

**摘要:**提出了在多源空间数据融合过程中出现不一致的问题;介绍了局部多项式插值方法的基本原理;通过该方法在海底沉积厚度模型构建中的成功应用,为基础数据的搜集和融合提供了新的思路。

**关键词:**局部多项式插值方法;空间数据融合;多源数据;沉积层厚度模型

**中图分类号:**P736; N37

**文献标识码:**A

**文章编号:**1000-3096(2009)04-0025-04

基础数据是科学研究的根本,基础数据种类的多样性、区域的广泛性,使非同源的数据的拼接、融合成为必然,这就要求我们用一些数学手段对这种情况进行处理,使数据可靠性得到保障,达到用于科学研究的目的。

## 1 数据融合方法的选择及原理介绍

空间数据融合的概念产生于 20 世纪 70 年代<sup>[1]</sup>。目前,被普遍接受的“数据融合”的概念是 1998 年 1 月,以 Buchroithner 和 Wald 的工作为基础给出的定义<sup>[1~3]</sup>：“数据融合是一个形式上的框架,在此框架下表达了融合的方式和工具,通过这些方式和工具将来自不同的源数据进行联合,其目的在于获取质量更好的信息,而质量的改善取决于应用。”本文中,所谓的数据融合实质上就是利用数学手段将来自非同源的空间数据的重叠部分进行处理,使多源数据重叠部分更好地过渡、拼接,以得到更大范围的可靠数据。

### 1.1 数据融合方法的选择

作者在上面提到的数学手段是指曲面处理方法<sup>[4]</sup>,包括曲面拟合和插值。实际应用中要根据需求来选择方法,主要有两点:一是尽可能不改变原始数据点的值,以保证数据的真实性;二是要较好地反映多源空间数据的重叠区域的曲面变化趋势,使数据尽可能圆滑地拼接。

鉴于以上两点原则,作者选择了插值方法作为空间数据融合的手段。当前常用的插值方法有多种<sup>[4]</sup>,下面先对其中一些方法及其特点作一些简要的介绍。

#### 1.1.1 反距离加权法(Inverse Distance to a Power)

计算速度快,但它只考虑离网格点最近的几个数据点,对曲面的整体趋势体现不够。

#### 1.1.2 克里格插值法(Kriging)

是当前最灵活,应用最广泛的一种插值方法,能

够兼顾曲面的整体趋势及局部细节,它对于同源数据的网格化无疑是最好的选择。

#### 1.1.3 多元回归分析法(Polynomial Regression)

是一种用于趋势面分析的插值方法,实质上是用简单的幂级数多项式来拟合复杂的实际曲面,结果只能反映曲面整体的变化趋势,为了准确地反映实际曲面,不能单独采用该方法。在实际应用中,一般结合其残差分析局部异常。

#### 1.1.4 局部多项式方法(Local Polynomial)

是一种局部加权最小二乘拟合方法,它引入了“距离权”的概念,比如求网格点 $(a, b)$ 的值,要考虑在局部范围内全部数据点对 $(a, b)$ 的贡献,距 $(a, b)$ 近的点权大,远的点权小,换一个网格点则形成另外一些权值。这样,每一个网格点值都对应一个多项式,求一个网格点值就要解一次联立方程。所以该方法一方面有趋势面法考虑全部数据点反映趋势性变化的优点,另一方面又有距离法反映局部特征的优点,对曲面作局部平滑的效果不错。

综合以上分析,本文选择了局部多项式方法作为数据融合的工具。

## 1.2 局部多项式插值方法的原理

如上所述,局部多项式方法实质上是一种局部加权最小二乘方法,它的算法原理可归结为以下三个主要步骤。

第一,选择插值函数。

最简单的插值函数是多项式<sup>[5]</sup>,一般常用的多项式

收稿日期:2008-09-03;修回日期:2009-02-17

基金项目:国家海洋局 908 专项项目(908-ZC-I-06,908-01-DW07/08)

作者简介:汪俊(1983-),男,浙江江山人,硕士研究生,主要从事海洋地质地球物理研究,电话:0571-81963157, E-mail: wangjun19830726@163.com

有三种,分别为一次、二次和三次多项式。在实践中,作者个人认为一般情况下二次多项式已能满足需求(公式1)。

$$F(X,Y)=a+bX+cY+dXY+eX^2+fY^2 \quad (1)$$

第二,确定“权”。

“权”的值由搜索范围、权系数和实际散点数据的几何分布(即“距离”)等因素决定,但在实际计算中往往只考虑其中几种<sup>[6]</sup>。所以在确定“权”的过程中要考虑这些因素的体现。

首先是确定搜索范围,它是局部多项式方法在“局部”特点上的体现。

定义:

$$T_{XX} = \frac{\cos\Phi}{R_1}, T_{XY} = \frac{\sin\Phi}{R_1}, T_{YX} = \frac{-\sin\Phi}{R_2},$$

$$T_{YY} = \frac{\cos\Phi}{R_2} \quad (2)$$

其中  $\Phi$  为搜索椭圆的搜索角度,  $R_1, R_2$  为搜索椭圆的长、短半径,两者决定了搜索范围。

又定义:

$$A_{XX} = T_{XX}^2 + T_{YX}^2, A_{XY} = 2(T_{XX}T_{XY} + T_{YX}T_{YY}),$$

$$A_{YY} = T_{YY}^2 + T_{XY}^2 \quad (3)$$

以上定义的  $A_{XX}, A_{XY}$  和  $A_{YY}$  仅仅是搜索椭圆的参数,只要搜索椭圆确定了,它们对于每个数据和网格节点来说都是定值。

其次是每个数据的几何分布,它是局部多项式方法在“距离权”特点上的体现。假定某一个散点数据位置为  $(X_i, Y_i)$ , 一个待求网格节点的位置为  $(X_0, Y_0)$ , 得出:

$$dX = X_i - X_0, dY = Y_i - Y_0 \quad (4)$$

$$\text{Minimize } \sum_{i=1}^N W_i [F(X_i, Y_i) - Z_i]^2 \quad (5)$$

最后就是选择“权系数”来确定“权”了:

$$W_i = (1 - R_i)^p \quad (6)$$

式中的  $W_i$  就是数据  $(X_i, Y_i)$  的“权”了,  $p$  是“权系数”。

第三,把以上分析的一个散点的情况,推广到搜索范围内的散点集合  $\{(X_i, Y_i, Z_i) \text{ for } i = 1, 2, \dots, N\}$ , 然后根据最小二乘原理(公式7),解出多项式(公式2)的系数  $a, b, c, d, e, f$ 。确定多项式,从而确定相应节点上的值。

$$\text{Minimize } \sum_{i=1}^N W_i [F(X_i, Y_i) - Z_i]^2 \quad (7)$$

## 2 局部多项式插值方法在多源海底沉积厚度模型构建中的应用

沉积层厚度模型作为研究该区域地质地球物理

背景重要的素材,需通过多种渠道进行搜集,主要包括:国际上公开的数据库、国内外已发表的文献资料的一部分可以直接获得的实测数据。本文的数据来源属于“国内外已发表的文献资料”这一类。

### 2.1 数据来源

本文实例中采用的数据来源于已发表的文献<sup>[7,8]</sup>, 主要包括两个区块(图1): A 区块和 B 区块。两者都处于台湾岛西南部,区块内涵盖了著名的澎湖盆地和台南盆地的大部分。

(1) A 区块, 相对处于北部, 区块范围  $(118^\circ \sim 120^\circ \text{E}; 22.5^\circ \sim 23.5^\circ \text{N})$ 。区内新生代以来的沉积厚度等值线如图2所示。

(2) B 区块, 相对处于南部, 区块范围  $(118^\circ \sim 120^\circ \text{E}; 22^\circ \sim 23^\circ \text{N})$ 。区内新生代以来的沉积厚度等值线如图3所示。

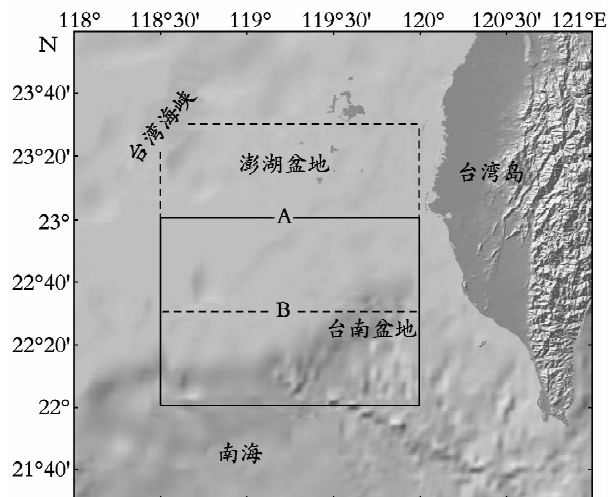


图1 研究区域的地形

Fig. 1 The topography map of research region

虚线框和实线框分别为 A 区块和 B 区块

The dashed frame and the real line frame represent A block and B block respectively

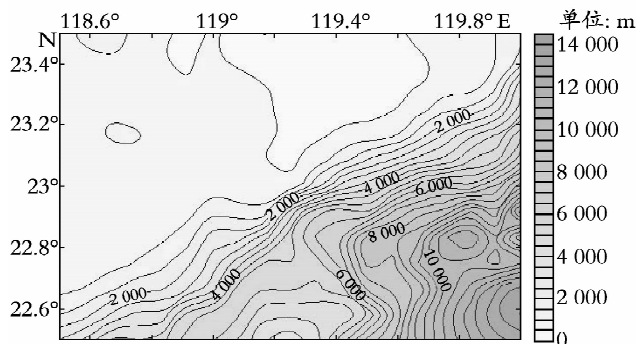


图2 A 区块内新生代以来的沉积厚度

Fig. 2 The Cenozoic sedimentary thickness map of A block

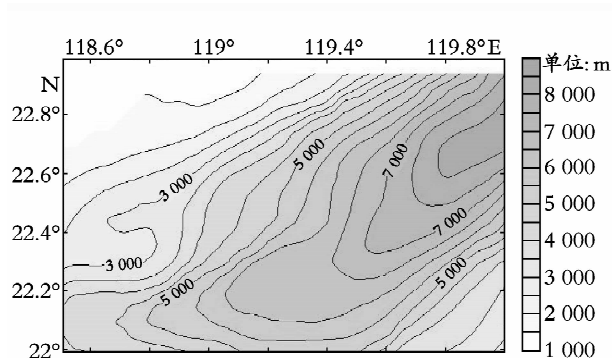


图3 B区块内新生代以来的沉积厚度

Fig. 3 The Cenozoic sedimentary thickness map of B block

## 2.2 数据拼接过程中出现的问题

从图2和图3可以看出,A区块和B区块在重合区域沉积层厚度的大体变化趋势相同,在局部细节上A区块明显更为丰富。若将两区块的数据拼接时,只是对两个区块数据的重合部分取平均值,势必会在重合区域出现等值线无法平滑过渡的情况(图4),但这种等值线的扭曲决不是实际异常,是两块非同源数据在融合上出现的不一致问题,在实际工作中十分常见。

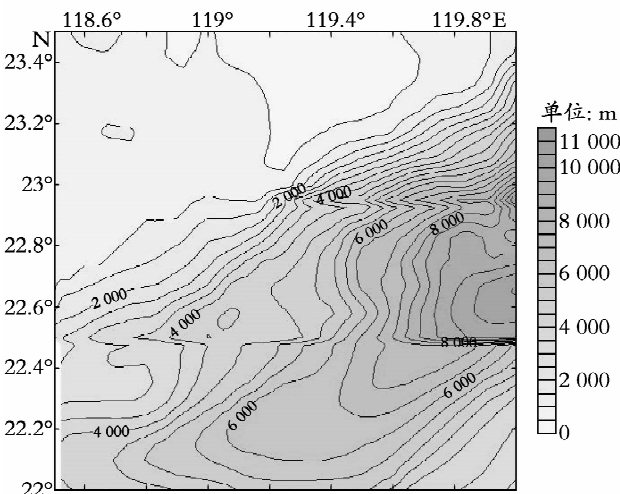


图4 只对A区块和B区块重合区域的数据求平均值的沉积厚度拼接

Fig. 4 The sedimentary thickness mosaic map of A block and B block which only calculated the average in reduplicate area

## 2.3 局部多项式插值方法的应用效果

面对以上出现的问题,本文选用局部多项式插值方法对A区块和B区块的沉积厚度数据做一些处理。本文在局部多项式拟合方法的应用过程中,选取的参数如下:

插值多项式:如公式2;

搜索椭圆长、短半径和搜索角度分别取: $R_1 = 1$ ,

$R_2 = 1, \Phi = 0$ ;

权系数: $p = 2$ 。

将处理的结果(图5)和图4比较,处理效果很明显,图4中数据重合区域的等值线扭曲得到了很好的解决,而且保持了源数据中的沉积层厚度的大体变化趋势。

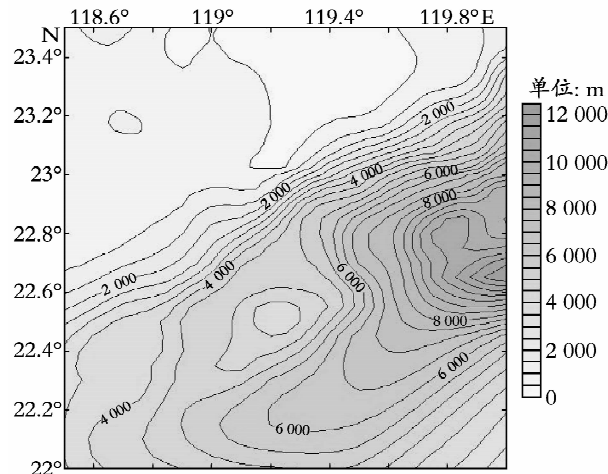


图5 经局部多项式插值方法处理的A、B区块沉积厚度  
Fig. 5 The sedimentary thickness mosaic map of A block and B block which calculated with local polynomials interpolation method in reduplicate area

## 3 结束语

本文从利用局部多项式插值方法解决多源海底沉积厚度数据的融合问题的过程中,取得了以下认识:(1)在局部多项式插值方法应用过程中,要根据实际需要调整搜索椭圆的参数( $\Phi, R_1$ 和 $R_2$ );(2)局部多项式插值方法作为一种曲面处理方法,具有既考虑全部数据点反映趋势性变化,又能较好地反映局部特征的优点。在本文多源海底沉积厚度数据融合的实例中,取得了良好的效果。

### 参考文献:

- [1] 郭黎,崔铁军,吴正升,等. 多源空间数据融合技术探讨[J]. 地理信息世界, 2007, 1: 62-66.
- [2] 杨华,林卉. 数据融合的研究综述[J]. 矿山测量, 2005, 3: 24-28.
- [3] 王春娟,杜德文,朱志伟,等. 空间数据融合模型及方法[J]. 海洋地质动态, 2008, 24(1): 37-4.
- [4] 安琪. 几种离散数据网格化方法的对比分析[J]. 内蒙古科技与经济, 2001, 1: 87-89.
- [5] 袁慰平,孙志忠,吴宏伟,等. 计算方法与实习[M]. 南京:东南大学出版社, 2000. 101-106.
- [6] Polzehl J, Spokoiny V. Varying coefficient regression modeling[J]. *WIAS*, 2003, 818: 15-21.
- [7] Lin A T, Watts A B, Hesselbo S P. Cenozoic stratigraphy and subsidence history of the South China Sea margin in the Taiwan region[J]. *Basin Research*.

2003, 15: 453-478.

- [8] McDonnella S L, Maxb M D, Cherkis N Z, *et al.* Tectono-sedimentary controls on the likelihood of gas hydrate occurrence near Taiwan [J]. **Marine and Petroleum Geology**, 2000, 17: 929-936.

- [9] Marine Geoscience Data system at Lamont-Doherty Earth Observatory of Columbia University. GeoMap App [EB/OL]. <http://www.geomapp.org>, 2008-08-08.

## The local polynomials fitting methods on geospatial data merging: An application to the multi-resources marine sediment thickness data

WANG Jun, GAO Jin-yao, WU Zhao-cai, ZHANG Tao

(The Second Institute of Oceanography, State Oceanic Administration, Hangzhou 310012, China)

**Received:** Sep. , 3, 2008

**Key words:** local polynomials interpolation method; geospatial data merging; multi-resources data; the model of marine sedimentary thickness

**Abstract:** This paper presents an inconsistent problem on multi-resources data merging. The theory of the local polynomials interpolation method was introduced. The successful application on constructing the model of marine sedimentary thickness was described. Experiences for data collecting and merging were provided.

(本文编辑:刘珊珊)

(上接第 8 页)

## Determination of taurine in *Nevertia didyma* (Roding) and *Asterias rollestoni* Bell by HPLC

ZHANG Liang, GENG Yue, TU Wen-li, ZHANG Jing-jing

(Key Laboratory of Animal Resistant Biology, College of Life Science, Shandong Normal University, Jinan 250014, China)

**Received:** Feb. , 10, 2007

**Key words:** taurine; *Nevertia didyma* (Roding); *Asterias rollestoni* Bell; HPLC

**Abstract:** The contents of taurine in liver and abdominal foot of *Nevertia didyma* (Roding) and in viscera and body wall of *Asterias rollestoni* Bell were analyzed by high-performance liquid chromatography (HPLC). The analysis was based on the HPLC separation of the ophthaldialdehyde derivative of taurine. The results showed that the taurine contents of liver and abdominal foot of *N. didyma* (Roding), viscera and body wall of *A. rollestoni* Bell were determined to be 535.2, 1298.7, 60.4 and 0.317 mg/g wet weight sample, respectively.

(本文编辑:张培新)