

估值类数据同化方法中背景场的确定——三维可分离的情况

马寨璞¹, 井爱芹^{1,2}

(1. 河北大学 生命科学学院, 河北 保定 071002; 2. 河北省基因工程技术应用中试基地, 河北 保定 071002)

摘要:以相关性为基础, 详细研究了估值类数据同化方法中背景场的确定问题, 并以渤海 3 月份到 7 月份的温度场为例, 具体说明了如何确定背景场: 当要素场在水平方向和垂直方向上可以进行分离变量时, 背景场的确定可以通过单独确立水平相关性与垂直相关性完成。研究表明: (1) 当没有温跃层时, 温度在垂向上的相关是沿着整个深度的; (2) 当有温跃层时, 温度在垂向上的相关应该限定在上准等温层的下界; (3) 在温度的垂向分布方面, 其相关性与温跃层有着对应关系, 对比相关性与温跃层可以看出, 垂向相关性与温跃层位置的变化趋势一致。

关键词: 数据同化; 背景场; 跃层; 渤海

中图分类号: P732.6; O141.4 文献标识码: A 文章编号: 1000-3096(2006)10-0025-05

估值类^[1]的数据同化方法是在估值理论基础上发展起来的一种同化方法, 该类数据同化方法的基本特点是以方差作为确定权重的依据, 以相关性作为研究要素间相互影响的手段, 所有属于该类的同化方法都会涉及到方差和相关性的问题。在这一类的同化方法中, 常常涉及到 3 种场^[2], 它们分别是: (1) 背景场; (2) 观测场; (3) 分析场。要进行数据同化, 首先要准备这 3 种场。

到目前为止, 所研究的 3 种场中总是存在着一些不确定性。对于不确定性的描述, 通常使用的描述方法为概率密度函数^[3]。设 X 为某种场, 定义 X_i 为该场的真值, 那么场 X 与其真值 X_i 之间的差异定义为 ε , 则称 ε 为误差变量。那么它的统计特性可以用概率密度函数来描述。在估值理论类的同化方法中, 最需要确定的量就是背景协方差矩阵。在观测站点, 背景值与观测值通常有些差异 (即观测增量), 同化的目的之一就是将该差异插值到或传播到各分析点, 通常多用背景误差协方差的空间结构来完成上述的插值或传播。

作者围绕着背景场, 详细研究了方差与相关性方面的准备问题。

1 三维可分离的情况

在背景场的确定过程中, 如果水平场与垂直场能够进行分离, 那么背景误差协方差的确定就会相

对简单一些。参照文献[4], 定义三维坐标: $r = r(x, y, z)$, 其中的 x, y 指定了水平坐标, z 指定了垂向坐标, 则定义在 r_k, r_l 处的背景误差协方差为:

$$C_B(r_k, r_l) = C_B^H(x_k, y_k, x_l, y_l) C_B^V(z_k, z_l) \quad (1)$$

式中, C_B^H, C_B^V 分别是水平与垂向的各自的方差。其中的下角标 B 表示背景场, k 和 l 表示不同的两个测点。上角标 H, V 分别表示水平方向和垂直方向。将 C_B^H, C_B^V 分别讨论, 得到水平与垂向的模型。因此, 将相关模型分为水平与垂向两个方向的。

水平误差协方差一般是假设各向同性的^[5]。在这一假设下, 引入自相关函数 $\rho_{\psi\psi}(r, \theta)$, 存在:

$$C_{\psi\psi}(r, \theta) = C_{\psi\psi}(0) \cdot \rho_{\psi\psi}(r, \theta) = \langle \psi^2 \rangle \cdot \rho_{\psi\psi}(r, \theta) \quad (2)$$

式中, 下标 ψ 表示不同方向。在各向同性条件下, (2) 式表明背景误差协方差就等于单一常数 (方差) 乘以背景误差相关矩阵。

对于海洋数据同化, 其同化效果的优劣在很大程度上依赖于背景误差协方差矩阵的表述形式。在实际应用中, 水平背景误差协方差 $C_B^H(r, \theta)$ 或相关函数 $\rho_B(r, \theta)$ 是预先定义好的, 改进它们的确定是至关重要的。观测数据可用来估计背景误差方差及其相关

收稿日期: 2004-04-12; 修回日期: 2004-08-18

作者简介: 马寨璞 (1970), 男, 河北保定人, 副教授, 博士, 从事物理海洋数据同化研究, 电话: 0312-5079364, E-mail: mazhai pu@m ail. hbu. edu. cn

函数。设有一个稠密的各向同性的数据观测网,假设该观测网某要素场的时间序列已经得到,观测数据可按照垂向分开,则观测站处背景场可以得到,而偏差也可从该观测数据场中剔除。设观测网共 k 个测站,在每一个测站有许多观测数据,则相关性表达为:

$$R_{lk} = \frac{\overline{(O_k - B_k)(O_l - B_l)}}{\sqrt{(\overline{O_k - B_k})^2 (\overline{O_l - B_l})^2}} \quad (3)$$

式中: O 表示观测场, B 表示背景场。上划线表示观测站 l 与 k 的长期观测序列的平均。 R_{lk} 表示观测站 l 与 k 处观测增量的相关性。计算各个观测站间的相关性系数,可得到相关性系数的大致分布。

若相关函数可通过分析确定的话,那么背景误差协方差矩阵的计算会更简便一些。多年来,人们提出了几种较实用的相关函数模型。它们多数以指数的表达形式出现,文献[6, 7]给出了形式相近的几种表达。

$$\rho_B(r) = \left[\cos(cr) + \frac{\sin(cr)}{Lc} \right] \cdot e^{-r} \quad (4)$$

式中, r 是测站之间的距离, L 是最大相关距离常数, c 是特定的常数,当 $c \rightarrow 0$ 时,有:

$$\rho_B(r) = \exp\left[-\frac{r^2}{L^2}\right] \quad (5)$$

2 渤海温度场的水平相关性研究

渤海是中国的内海,对它的温度场的研究已经有许多文献可供参考,如黄大吉^[8]等。为了研究渤海温度场的水平相关情况,作者利用文献[8, 9]提供的模型,并采用当年实际测量的天气、风力等要素作为外力条件,计算了渤海温度分布的水平相关系数,选取该区域内的5个点,它们的点位分布如图7所示,各点的坐标分别为: A(40.166 6° N, 121.166 6° E); B(39.333 3° N, 120.416 6° E); C(38.666 6° N, 118.333 3° E); D(38.166 6° N, 119.916 6° E); E(38.333 3° N, 122.083 3° E)。选取这5个点,是考虑在渤海区域内均匀分布,A点位于辽东湾中部,代表着湾内的情况。B, C, D三点居于渤海中部区域外围, E点则位于和黄海的交界处,代表着水深较大的情况。

在具体讨论使用(3)式时,要涉及到观测场的数据,对于特定观测点的数据,考虑到实测数据的欠缺,文中采用了变通的处理方法,即采用伪数据作为观

测场,令计算值加上一个很小的合理的误差量作为观测场,这样的处理方法可以说明上述公式的应用和图像处理中滤波算法的研究相类似。

经过计算,给出相关系数的分布如图2~图6所示,图2~图6给出了5个点处各自的相关系数(平均0.9以上)。从图2~图6描述的5点的水平相关系数分布可知,在夏季,渤海表层水温的强相关距离约100 km左右,且相关区域多为一个近似各向同性的区域。

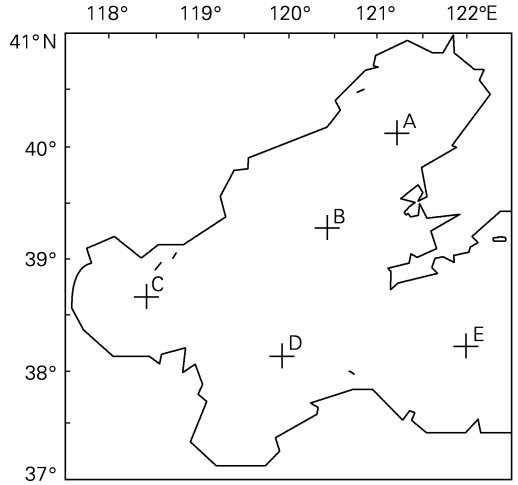


图1 数据采样点分布

Fig. 1 Sampling point distribution

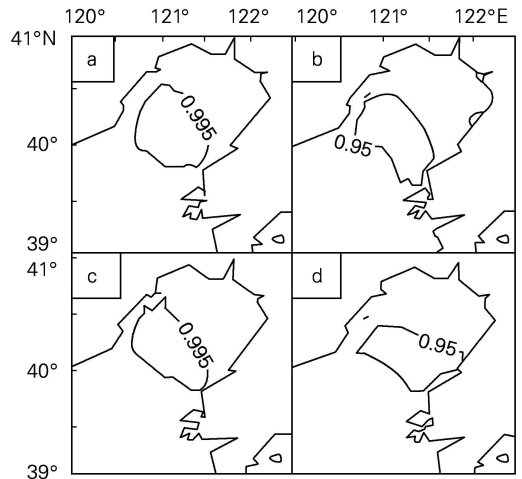


图2 A 采样点的相关系数

Fig. 2 Relative coefficient at point A

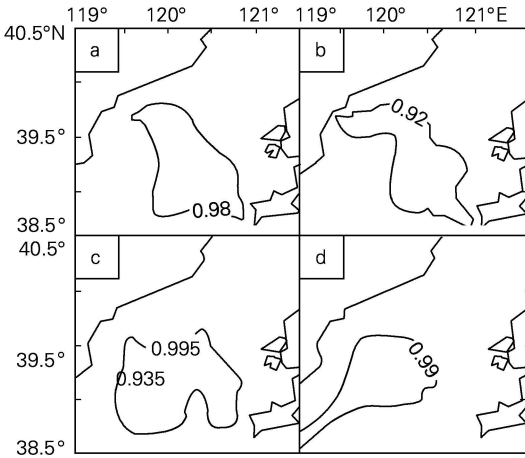


图3 B采样点的相关系数

Fig. 3 Relative coefficient at point B

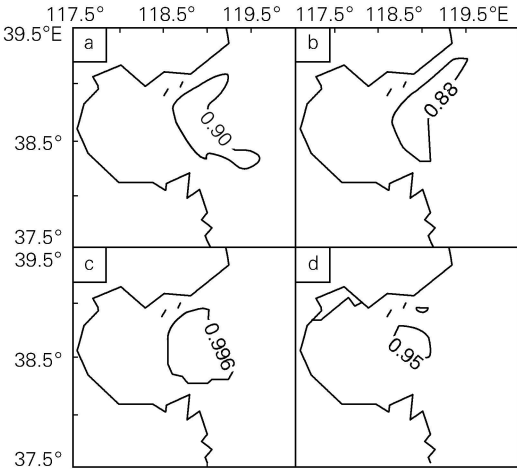


图4 C采样点的相关系数

Fig. 4 Relative coefficient at point C

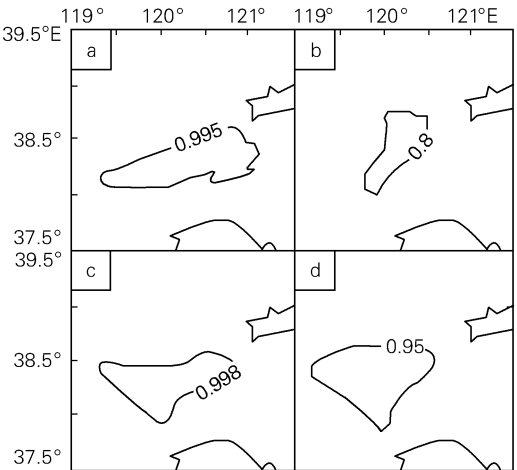


图5 D采样点的相关系数

Fig. 5 Relative coefficient at point D

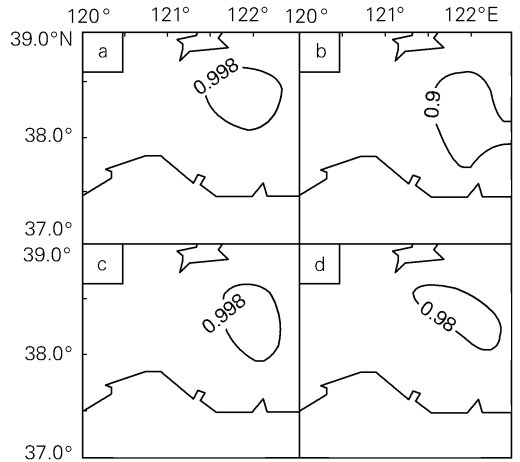


图6 E采样点的相关系数

Fig. 6 Relative coefficient at point E

3 渤海温度场的垂向相关性研究

为了研究渤海区域内的垂向温度的相关性,作者检测了渤海5个点的垂向水温变化及垂向关系数的变化。从直观上讲,温跃层可看作一个界面,将温度垂向分布分成三部分,跃层以上的上准等温层,温度正梯度层(这里不含温度负梯度层),跃层以下的下准等温层。并且跃层以上和以下的分布都自成系统,两者的关系通过跃层将这两部分连接起来。因此,从相关性上讲,相关系数在经过温跃层时,应该有较大的变化,而在跃层以上和以下的准等温层中的变化应该平缓。为此,作者计算了5点的相关系数分布(和表层的相关,基准层为表层),下面给出的是A, B, C, D, E这5个点的3月到7月的根据每月15日一天中的8次采样计算得到的相关系数。

从图中可以看出,当没有温跃层时,垂向相关系数随深度距离增加而衰减,在沿着整个深度上都正相关,图7中A点很明显地说明了这一点。当有温跃层时,虽然每月的相关系数有所变化,但各点处的相关趋势是一致的。在温跃层以上,相关值都较大,但在温跃层内部,相关性系数迅速衰减到0或负相关。因此,在考虑垂向相关时,作者认为,要以该位置的温跃层为参考,将相关距离限定在上准等温层的下界处。总结为:(1)当没有温跃层时,温度在垂向上的相关是沿着整个深度的;(2)当没有温跃层时,温度在垂向上的相关应该限定在上准等温层的下界;(3)在温度的垂向分布方面,其相关性与温跃层有着某种

对应关系,从图上的对比可看出:垂向相关性与温跃层位置的变化趋势一致(为了节省篇幅,只给出了E点的垂向温度分布,其余略)。

为了检验垂向相关与温跃层之间的显著性,以E点为代表,对7月份的温度数据和相关系数进行显著性检验,经过回归,得到在温跃层(含跃层本身)以上,二者的回归方程为:

$$y = -0.052037 + 0.043463x$$

式中, x 表示温度, y 表示相关系数。对回归方程进行显著性检验,结果表明:检验值小于临界值,两者存在着显著性相关。

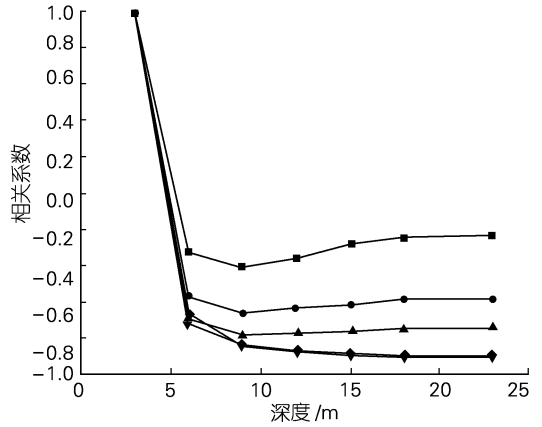


图9 C点每月15日温度垂向相关系数

Fig. 9 Vertical relative coefficient at point C

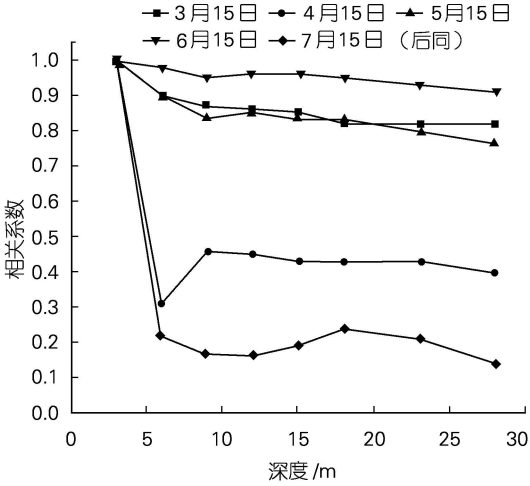


图7 A点每月15日温度垂向相关系数

Fig. 7 Vertical relative coefficient at point A

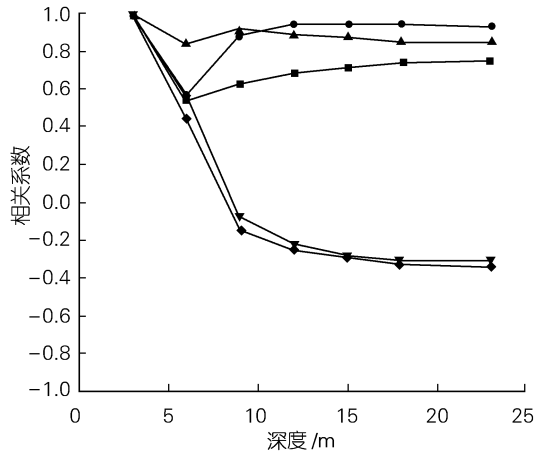


图10 D点每月15日温度垂向相关系数

Fig. 10 Vertical relative coefficient at point D

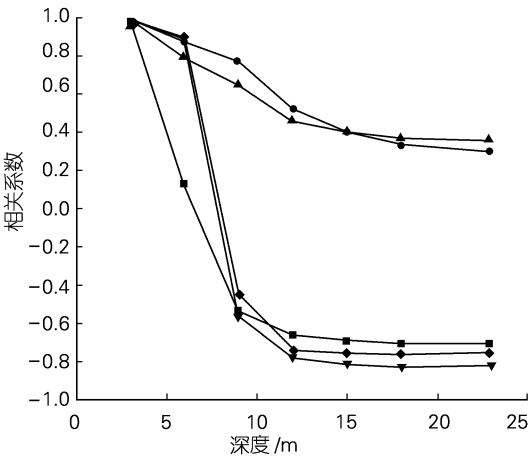


图8 B点每月15日温度垂向相关系数

Fig. 8 Vertical relative coefficient at point B

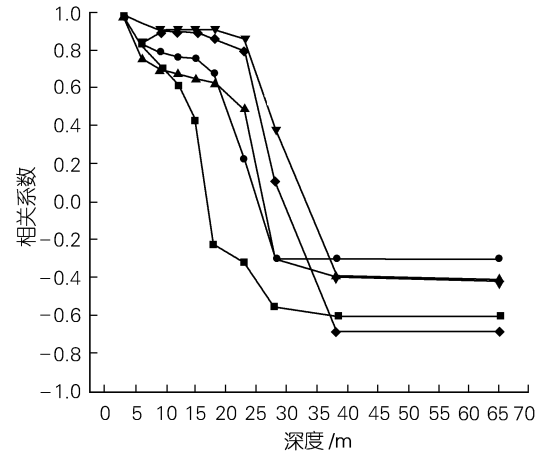


图11 E点每月15日温度垂向相关系数

Fig. 11 Vertical relative coefficient at point E

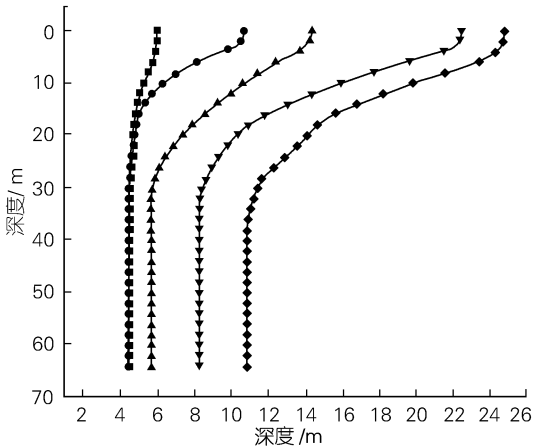


图 12 E 点每月 15 日温度垂向分布

Fig. 12 Temperature vertical distribution at E

4 三维时间相关的理论研究

前边研究了同一时段内不同观测地点的观测数据的相关情况,并以渤海温度场为实例,具体研究了不同区域内在水平和垂向上的相关情况。这种情况适用于观测时刻相同或者近似相同的不同地点的同一个要素的相关,可用来处理和距离有关的背景场的准备情况。文献[10,11]就是基于这种相关情况进行数据同化的具体例子。

然而,对某一要素场的观测,有时候在时间上是连续的,这就涉及到和时间有关的相关背景场处理。在上述(1),(2)式的基础上,将相关的意义延伸,可处理和时间相关的背景场的准备问题,具体的处理过程和上述空间相关类似,不再详细讨论。

5 结论

估值类数据同化方法的应用难点之一是背景场的确定,这其中主要是如何应用观测数据。相关性为处理此类问题提供了基础。作者详细研究了估值类的数据同化方法中背景场的确定问题,说明了如何确定背景场,并以渤海3月份到7月份的温度场为例,具体处理了渤海水温要素的相关:当要素场能够在水平方向和垂直方向进行分离变量时,背景场的

确定可以通过单独确立水平相关性与垂直相关性完成。研究表明:在垂向上,当没有温跃层时,温度的相关是沿着整个深度的;当有温跃层时,其相关应该限定在上准等温层的下界;在温度的垂向分布方面,其相关性与温跃层有着对应关系,对比相关性与温跃层可以看出:垂向相关性与温跃层位置的变化趋势一致。

参考文献:

- [1] Robinson A R, Lermusiaux P F J, Sloan N Q. The Sea [M], New York: Wiley, 2001. 541-594.
- [2] Bouttier F, Courtier P. Data assimilation concepts and methods[EB/OL]. http://www.ecmwf.int/newsevents/training/rcourse_notes/DATA_ASSIMILATION/ASSIM_CONCEPTS/Assim_concepts2.html, 2002-12-06/2004-02-18.
- [3] 郭尚来. 随机控制[M]. 北京:清华大学出版社, 1999. 7-21.
- [4] Daley R. Atmospheric Data Analysis [M]. Cambridge: Cambridge University Press, 1991. 457.
- [5] Daley R. Forecast error statistics for homogeneous and inhomogeneous observation networks[J]. *Monthly Weather Review*, 1992, **120** (B): 627-643.
- [6] Lermusiaux P F J, Robinson A R. Data assimilation via error subspace statistical estimation Part I: theory and schemes [J]. *Monthly Weather Review*, 1999, **127** (7): 1385-1407.
- [7] 朱江,徐迎春,王赐震. 海温数值预报资料同化试验 I. 客观分析的最优插值法试验[J]. *海洋学报*, 1995, **17** (6): 9-20.
- [8] 黄大吉,苏纪兰,陈宗镛. 三维陆架海模式在渤海中的应用 II. 温度的季节性变化[J]. *海洋学报*, 1996, **18** (6): 8-18.
- [9] 黄大吉,陈宗镛,苏纪兰. 三维陆架海模式在渤海中的应用 I. 潮流、风生环流及其相互作用[J]. *海洋学报*, 1996, **18**(5): 1-13.
- [10] 马寨璞,黄大吉,章本照. 应用 SVD 分解简化卡尔曼增益计算的理论研究[J]. *浙江大学学报(工学版)*, 2003, **37**(1): 87-93.
- [11] 马寨璞,黄大吉,章本照. 融合法及其在数据同化中的应用研究[J]. *海洋学报*, 2003, **25**(2): 33-41.

(下转第 60 页)

Preparation for the background of the estimating data assimilation method——the separated background branch

MA Zhai-pu¹, JIN G Ai-qin^{1,2}

(1. Department of Ocean Science, Hebei University, Baoding 071002, China; 2. Pilot Center of Gene Engineering and Technology Application of Hebei Province, Baoding 071002, China)

Received: Apr. , 12. 2004

Key words: data assimilation; background; thermocline; the Bohai Sea

Abstract: Based on the relativity, the preparation for the background of the estimating data assimilation method is studied. The Bohai temperature, as an example cited in this paper from March to July, is detailed to prepare the background when its horizontal distribution and vertical one can be separated from each other. The research reaches the conclusions that the covariances of horizontal and vertical ones can be constructed respectively when such separation can be done; that the relative coefficient along vertical direction strides the total depth where there no thermocline; that the relative coefficient along vertical direction is limited strictly above and below the thermocline where there one exists; and that the thermocline can be regarded as a barrier layer; that the trend of the relative coefficient is akin to that of the temperature along the vertical direction.

(本文编辑:刘珊珊)