

在四维变分同化中运用集合协方差的试验^{* 1}

张 蕾 邱崇践 张述文
ZHANG Lei QIU Chongjian ZHANG Shuwen

半干旱气候变化教育部重点实验室,兰州大学大气科学学院,兰州,730000

Key Laboratory for Semi-Arid Climate Change of the Ministry of Education, College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China

2008-08-28 收稿,2008-11-13 改回.

Zhang Lei, Qiu Chongjian, Zhang Shuwen. 2009. Experiments on the 4D-variation with ensemble covariances. *Acta Meteorologica Sinica*, 67(6):1124–1132

Abstract Using the two-dimensional shallow water equation model and model simulated data, a set of numerical experiments were conducted to evaluate the impacts of three different specification schemes of the background error covariance matrix on the four-dimensional variational (4DVAR) data assimilation in the case of different observation densities and observation errors. The three schemes are as follows: (1) for a single control variable, the background error covariance is assumed to be a diagonal matrix; (2) the background error covariance is simplified to a Gaussian form with the homogeneous and isotropic assumptions; (3) the background error covariance is restructured through using the ensemble forecasts and the solving of the inverse of the background error covariance matrix is carried out by using the singular value decomposition (SVD) technique. The results show that the background error covariance plays an important role in 4DVAR data assimilation. When the observational spatial density is not high enough, there is no satisfied analysis available if the background error covariance matrix is simply reduced to a diagonal matrix. The Gaussian filter scheme has the ability to improve the analysis accuracy, but this it is sensitive to the length scale of background error correlations. The third method shows a stable performance. In this method, the background error covariance matrix is calculated implicitly so the computation of the inverse of background error covariance matrix is avoided. When observations are sparse or large errors exist in the observations, the third method will behave better compared to the other two methods.

Key words Data assimilation, Background error covariance, 4D variation, Ensemble forecasts, Singular value decomposition

摘 要 利用浅水方程模式和模式模拟资料进行数值试验比较 3 种不同的背景误差协方差矩阵处理方法对四维变分(4DVAR)资料同化的影响。3 种背景误差协方差矩阵分别是:(1)对单一变量将背景误差协方差矩阵简化为对角矩阵;(2)将背景误差协方差矩阵的作用简化为高斯过滤;(3)由预报集生成背景误差协方差矩阵并利用奇异值分解技术解决矩阵的求逆。通过一系列数值试验,比较不同观测密度、不同观测误差下 3 种背景误差协方差处理方法对 4DVAR 同化效果的影响。结果表明,背景误差协方差的结构对 4DVAR 有重大影响。当观测资料的空间密度不够高时,采用对角矩阵得不到满意的结果。高斯过滤方案可以明显改善同化结果,但是对背景误差特征长度比较敏感。第 3 种方法采用的背景误差协方差矩阵是流型依赖的,而且并不以显式的方式出现在目标函数中,避免了对它求逆的复杂运算。由于做了降维处理,在观测点的密度较低和观测误差较大时可望取得较好的同化结果,同化效果较为稳定。

关键词 资料同化,背景误差协方差,四维变分,预报集合,奇异值分解

中图法分类号 P435

^{*} 资助课题:国家高技术研究计划(863 计划)(2007AA12Z140),国家自然科学基金项目(40875063、40775065)和科技部公益性行业(气象)科研专项(GYHY200806029)。

作者简介:张蕾,主要从事卫星资料的反演与同化研究。E-mail: lzhangsea@163.com

通讯作者:邱崇践,主要从事数值模拟、大气资料同化、Doppler 雷达和卫星遥感资料的反演方法。E-mail: qiucj@lzu.edu.cn

1 引言

四维变分方法(4DVAR)(Lewis, et al, 1985; Talagrand, et al, 1987; Le Dimet, et al, 1986)将同化问题表述为微分方程的反问题,在同化过程中利用了完整的模式方程作为动力约束,可以同时同化多个时刻的观测资料,是目前在天气预报业务中成功运用的一种资料同化方法(Zou, et al, 1997; Park, et al, 2003)。4DVAR 综合处理来自背景场和观测两方面的信息(Daley, 1991),按照最优估计理论,这两种信息在分析场中的贡献依赖于各自的误差统计特征(Stephen, 1997)。其中背景误差协方差矩阵可以起到将观测信息传递给周围地区的作用,这在观测资料稀疏或者缺失区域尤为重要(Zhang, et al, 2003; Lorenc, et al, 2003; 龚建东等, 2006),因此正确地给出背景误差协方差对同化结果有不小的影响。现在业务运行的 4DVAR 系统是通过统计的方法来估计背景误差协方差(Gandin, 1965; Rutherford, 1972; Hollingsworth, et al, 1986; Lonnberg, et al, 1986; Dee, et al, 1999; Fisher, 1998; Parrish, et al, 1992),其中运用最多的是 NMC 方法(Parrish, et al, 1992)。这种由过去预报样本统计得到的背景误差协方差是不依流型而变的。在变分同化中,给出背景误差协方差后,为了计算目标函数及其相对控制变量的梯度,还需要计算背景误差协方差矩阵的逆矩阵。由于分析变量的自由度很大,对于大型的业务预报模式,协方差矩阵的维数高达 10^{14} 量级,并且是病态的,不可能直接求逆。为了简化计算,实际业务中通常假设背景误差是均匀、各向同性的,并且通过适当的变换让各分析变量之间不存在相关关系,这被认为是 4DVAR (以及 3DVAR) 的一个重大缺点。集合卡尔曼滤波(EnKF)方法(Evensen, 1994; Houtekamer, et al, 2001; Hamill, et al, 2001; Tippett, et al, 2003)通过一组随时间演变的预报样本来估计预报误差协方差,所得到的预报误差协方差是依流型而变的,这被认为是它相对于 4DVAR 的一大优点。将 4DVAR 和 EnKF 结合自然成为大家研究的一个目标。Lorenc (2003) 将 4DVAR 和 EnKF 作了比较后指出将变分方法和

EnKF 结合的混合方法可能是一种具有吸引力的同化方法。在此之前 Hamill 等(2002)也曾提出过一个 EnKF-3DVar 混合方案,在他们的混合方案中背景误差协方差一部分由常规的统计方法产生,是均匀、各向同性的,另一部分由 EnKF 提供,是依流型而变的。另一方面,一些研究还提出将低阶近似用于卡尔曼滤波来克服高维病态矩阵求逆的困难(Cohn, et al, 1996; Verlaan, et al, 1997; Farrell, et al, 2001),这些研究利用奇异值分解(SVD)技术将背景误差协方差矩阵的维数大大降低。Qiu 等(2006)也指出应该在大气吸引子上求解同化问题,据此提出了一种集合降维同化方法,其基本思想是将 SVD 用于预报集合得到一组基向量,同化是在这组基向量形成的子空间进行。Cao 等(2007)也将合适正交分解方法(Proper Orthogonal Decomposition, 简称 POD)应用于四维变分同化,提出一种降维四维变分同化方法。这类降维方法使得模式变量的自由度大大减少,为在 4DVAR 中采用依流型而变的背景误差协方差矩阵带来了方便。本文将降维同化方法和 4DVAR 相结合,给出一种在 4DVAR 中利用依流型而变的背景误差协方差矩阵的实施方案,利用浅水方程模式和模式模拟资料进行的数值试验对该方法的效果加以检验。

2 方法介绍

4DVAR 被表述为极小化的目标函数

$$J(\mathbf{x}_0) = \mathbf{J}_b + \mathbf{J}_0 \quad (1)$$

$$\mathbf{J}_b = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^T \mathbf{P}^{-1}(\mathbf{x}_0 - \mathbf{x}_b) \quad (2)$$

$$\mathbf{J}_0 = \frac{1}{2} \sum_{k=0}^K (\mathbf{H}(\mathbf{x}_k) - \mathbf{y}_k)^T \mathbf{R}_k^{-1} (\mathbf{H}(\mathbf{x}_k) - \mathbf{y}_k) \quad (3)$$

其中,向量 $\mathbf{x}_0$ 和 $\mathbf{x}_b$ 分别是初始时刻的模式变量和背景场变量, $\mathbf{P}$ 和 $\mathbf{R}$ 分别是背景和观测的误差协方差矩阵, $\mathbf{H}$ 是观测算子, $\mathbf{y}_k$ 是在 k 时刻的观测, $\mathbf{x}_k$ 是在相应时刻的预报,它是从初始场 $\mathbf{x}_0$ 出发积分下面的预报方程得到:

$$\mathbf{x}_k = \mathbf{M}_k(\mathbf{x}_0) \quad (4)$$

令 $\Delta \mathbf{x}_0 = \mathbf{x}_0 - \mathbf{x}_b$, 代入式(2)得到

$$\mathbf{J}_b = \frac{1}{2}(\Delta \mathbf{x}_0)^T \mathbf{P}^{-1} \Delta \mathbf{x}_0 \quad (5)$$

通常背景误差协方差矩阵 $\mathbf{P}$ 是一个维数很高

的矩阵,不可能直接求逆,实际运用时必须采取某种近似,比较3种方案。

方案1:假设矩阵 $\mathbf{P}$ 是对角矩阵,这样它只起到权重系数的作用,是一种最简单的方案,在一些4DVAR系统中也经常使用。

方案2:假设两个点之间背景误差相关在水平方向是均匀、各向同性的,矩阵 $\mathbf{P}$ 中的元素用下面高斯指数函数表示

$$p_{i,j} = \sigma_p^2 \exp[-(r_{i,j}/L)^2] \quad (6)$$

其中, σ_p^2 是该平面上所有点平均的背景误差方差, $r_{i,j}$ 是水平方向上 i,j 两点之间的距离, L 是一个水平特征尺度因子,实际运用时会选择一个影响半径,两点间距离 $r_{i,j}$ 超过影响半径时可近似认为 $p_{i,j} = 0$ 。为了避免矩阵 $\mathbf{P}$ 的求逆,Lorenc(1988)和Derber(1989)在3DVAR中提出变换。

$$\text{令 } \boldsymbol{\gamma} = \mathbf{P}^{-1}(\mathbf{x}_0 - \mathbf{x}_b), \text{ 即 } \mathbf{x}_0 = \mathbf{x}_b + \mathbf{P}\boldsymbol{\gamma} \quad (7)$$

在4DVAR中仍然可以作类似的变换。将式(7)分别代入式(2)和(4)得到

$$\mathbf{J}_b = \frac{1}{2} \boldsymbol{\gamma}^T \mathbf{P} \boldsymbol{\gamma} \quad (8)$$

$$\mathbf{x}_k = \mathbf{M}_k(\mathbf{x}_b + \mathbf{P}\boldsymbol{\gamma}) \quad (9)$$

方案三:利用预报集合给出背景误差协方差矩阵 $\mathbf{P}$ 的估计,并利用SVD实现降维,同时避免背景误差协方差矩阵求逆。

通过扰动初始场得到同化时刻的 M 个预报样本组成的集合,记为 $\mathbf{A} = (\mathbf{x}_1, \mathbf{x}_2 \cdots \mathbf{x}_M)$,将平均场 $\bar{\mathbf{x}}$ 近似为真实场,记 $\Delta\mathbf{A} = (\mathbf{x}_1 - \bar{\mathbf{x}}, \mathbf{x}_2 - \bar{\mathbf{x}} \cdots \mathbf{x}_M - \bar{\mathbf{x}})$,则有 $\mathbf{P} = \Delta\mathbf{A}(\Delta\mathbf{A})^T / (M-1)$ 。对 $\Delta\mathbf{A}$ 作奇异值分解得到

$$\Delta\mathbf{A} = \mathbf{B}\boldsymbol{\Lambda}\mathbf{V}^T \quad (10)$$

这里 $\boldsymbol{\Lambda}$ 是矩阵 $\Delta\mathbf{A}$ 的奇异值组成的对角矩阵,按照由大到小的顺序排列, $\mathbf{B}$ 和 $\mathbf{V}$ 分别是 $\Delta\mathbf{A}$ 的左奇异向量和右奇异向量组成的正交矩阵。由式(10)得到 $\mathbf{Q} = \Delta\mathbf{A}^T(\Delta\mathbf{A}) / (M-1) = \mathbf{V}\boldsymbol{\Lambda}^2\mathbf{V}^T / (M-1)$ 及 $\mathbf{P} = \Delta\mathbf{A}(\Delta\mathbf{A})^T / (M-1) = \mathbf{B}\boldsymbol{\Lambda}^2\mathbf{B}^T / (M-1)$ 。 $\mathbf{V}$ 的第 m 个列向量 $\mathbf{v}_m$ 是矩阵 $\mathbf{Q}$ 的第 m 个特征向量, $\mathbf{B}$ 的第 m 个列向量 $\mathbf{b}_m$ 是矩阵 $\mathbf{P}$ 的第 m 个特征向量。 $\Delta\mathbf{A}$ 可以根据 $\mathbf{v}_m$ 和 $\mathbf{b}_m$ 近似重构

$$\Delta\mathbf{A}_m = \mathbf{x}_m - \bar{\mathbf{x}} = \sum_{k=1}^r \lambda_k \mathbf{v}_{m,k} \mathbf{b}_k \quad (11)$$

$m = 1, 2 \cdots M$

这里 r 是截断阶数。类似展开式(11)将向量 $\mathbf{b}_m$ 作为基向量,分析增量 $\Delta\mathbf{x}_0$ 依 $\mathbf{b}_m$ 展开得到

$$\Delta\mathbf{x}_0 = \mathbf{x}_0 - \mathbf{x}_b \approx \sum_{k=1}^r \alpha_k \mathbf{b}_k = \mathbf{B}\boldsymbol{\alpha} \quad (12)$$

预报模式写为

$$\mathbf{x}_k = \mathbf{M}_k(\mathbf{x}_b + \mathbf{B}\boldsymbol{\alpha}) \quad (13)$$

将式(12)代入式(5)得到

$$\mathbf{J}_b = \frac{1}{2} (\mathbf{B}\boldsymbol{\alpha})^T \mathbf{P}^{-1} (\mathbf{B}\boldsymbol{\alpha}) = \frac{1}{2} (M-1) \boldsymbol{\alpha}^T \boldsymbol{\Lambda}^{-2} \boldsymbol{\alpha} \quad (14)$$

这里利用了近似 $\mathbf{P} = \mathbf{B}\boldsymbol{\Lambda}^2\mathbf{B}^T / (M-1)$

3 数值试验

利用浅水方程模式和模式资料比较上面提到的3种处理背景误差协方差矩阵的方案对4DVAR同化效果的影响。

3.1 模式和资料

考虑了地形的浅水方程模式是

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} - fv + g \frac{\partial h}{\partial x} = 0 \quad (15a)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + fu + g \frac{\partial h}{\partial y} = 0 \quad (15b)$$

$$\frac{\partial h}{\partial t} + \frac{\partial[u(h+H)]}{\partial x} + \frac{\partial[v(h+H)]}{\partial y} - \frac{\partial(uh_s)}{\partial x} - \frac{\partial(vh_s)}{\partial y} = 0 \quad (15c)$$

其中, $f = 10^{-4} \text{ s}^{-1}$, $H = 3000 \text{ m}$, 地形高度 $h_s = h_0 \sin(4\pi x/L_x) \sin(\pi y/L_y)$, $h_0 = 200 \text{ m}$, L_x 和 L_y 分别是矩形计算区域的两个边长,计算网格为 45×45 ,网格距 300 km ,积分时间步长 360 s ,采用双周期边界条件。从下面的初始场出发

$$h = 240 \sin(\pi y/L_y) + 120 \cos(2\pi x/L_x) \sin(2\pi y/L_y) \quad (16a)$$

$$u = -f^{-1} g \partial h / \partial y \quad (16b)$$

$$v = f^{-1} g \partial h / \partial x \quad (16c)$$

模式积分48 h得到第一个同化循环开始时刻的“真实场”(图1a)。将模式积分240 h,每3 h一次输出的结果作时间平均后作为该时刻的背景场(图1b)。从初始“真实场”继续积分得到随后需要的所有“真实场”,这些“真实场”用于检验同化效果和产生观测场。在所有试验中假定模式没有误差,只有高度观测,在真实场上叠加一个方差为 100 m^2 的随机扰动场得到观测值。

3.2 试验设计

假定同化时间窗长度为12 h,每隔3 h有一次

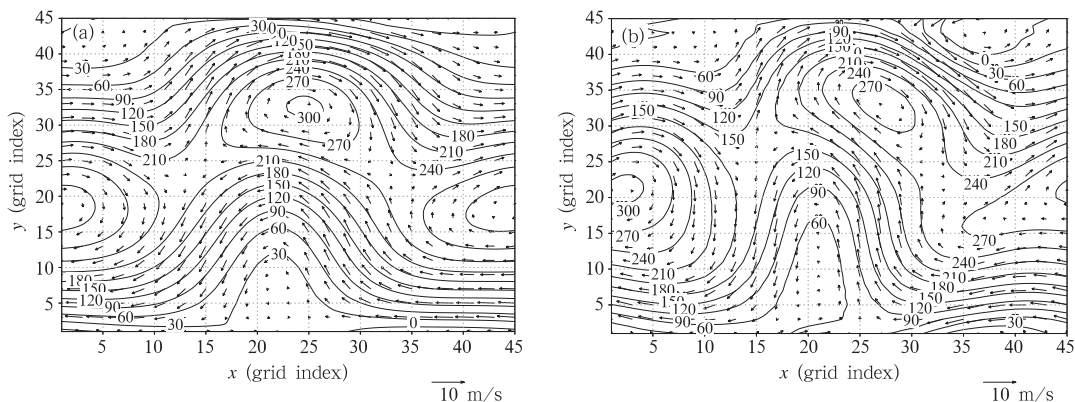


图1 第一个同化循环开始时刻真实场(a)与初猜场(b)

(等值线代表高度场,间隔 30 m;矢量线代表水平风场)

Fig. 1 Model-produced “true” initial fields (a) and imperfect initial state (b) at the starting time ($t=0$) of the first assimilation cycle (Contours are every 30 m for the height fields. Vectors are for the wind fields, and the vector scale (10 m/s) is labeled at the bottom of each panel)

高度观测。一个同化循环结束后,将末时刻的分析场作为下一个同化循环的背景场,一共进行 10 个同化循环。针对不同的观测密度设计了 3 个试验来比较 3 种背景场方案的同化效果。试验 1,在所有 2025 个网格点上都有观测;试验 2,有 202 个观测,它们随机分布在网格点上;试验 3,有 101 个观测,随机分布在网格点上。

所有试验中假设不同点的观测误差独立,按照所给的高度观测误差方差的倒数给出目标函数中观测的权重。背景误差协方差采用对角方案时,背景权重取为空间平均的背景场误差方差的倒数。背景误差协方差采用高斯方案时,假定观测的影响半径为 4 个网格距,高斯函数式(6)中的水平尺度 L 根据背景误差的统计结构确定。背景误差协方差采用 SVD 方案时,用于产生背景误差协方差的预报样本是在 12 h 前的背景场上叠加不同的随机扰动场后积分 12 h 得到的,随机扰动场采用 Evensen (2003) 生成随机扰动场的方法产生。在按照式(12)展开时截断阶数 $r=75$,每个同化循环结束后都按同样的方法重新产生预报样本。

4 试验结果

利用定义的相对误差来比较不同方案的同化效果,对于高度场

$$E_n(h) = \frac{\sum_{i=1}^S (h_{n,i}^a - h_{n,i}^t)^2}{\sum_{i=1}^S (h_{1,i}^f - h_{1,i}^t)^2}$$

这里 $E_n(h)$ 表示第 n 次同化循环起始时刻的高度分

析误差, h_n^a 和 h_n^t 分别是该时刻的分析值和真实值, h_1^f 和 h_1^t 分别是第一次同化开始时刻的预报值和真实值,求和是对所有网格点进行。对于风场(u, v),有类似的定义。将二者平均, $E_n(V) = [E_n(u) + E_n(v)]/2$, 作为水平风的分析误差。

采用方案 2 时需要确定高斯滤波函数的水平特征尺度。我们根据同化起始时刻所给的背景场误差分别统计出 h, u 和 v 3 个变量背景误差的相关系数与距离的关系,利用式(7)对统计结果进行拟合,由此确定对 3 个变量高斯分布函数的水平特征尺度。图 2 是根据同化开始时刻的预报误差统计的两

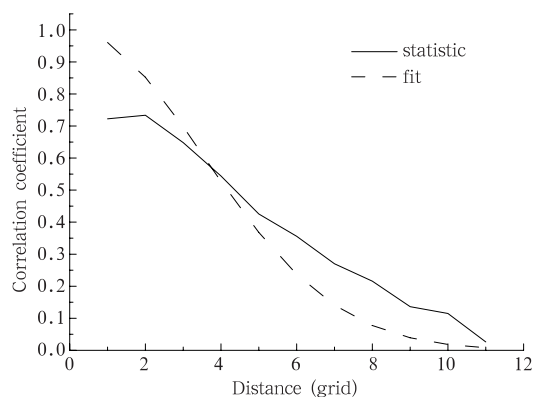


图2 同化区域中两点间水平风 u 分量背景误差相关系数随距离(单位:网格距)的变化 (实线是由预报样本统计的结果,虚线是高斯函数拟合结果)

Fig. 2 The correlation coefficient of the background error of u component plotted as functions of distance (unit: grid distance) (Solid line: the statistical results from the forecasting samples, dash line: the fitness results with Gauss function)

点间水平风 u 分量预报误差相关系数随距离的变化及用高斯函数拟合的结果, 可以看到, 高斯滤波函数的拟合结果与实际统计结果是比较相近的。

从所有格点都有观测时(试验 1)3 种方案得到的高度和风场的分析误差随同化循环次数的变化(图 3)可以看到, 无论是高度场还是风场, 方案 1 和 2 的同化效果都很好, 一次同化后分析误差就很小, 以后一直保持较小值。而方案 3 的分析效果较差, 虽然每次同化分析误差都有所减少, 但是减少的幅度不够大, 在 10 个同化循环后分析误差仍然大于前两种方案。方案 3 的分析误差大的一个重要原因

是将分析变量按照 SVD 基向量展开会产生较大的截断误差。为了说明这一点我们计算了第一次同化开始时真实的增量场(即背景场误差)按照式(12)展开(截断阶数 $r=75$)的截断误差(相对误差), 对于高度和水平风它们分别是 0.38 和 0.56, 这时候得到的分析误差是 0.55 和 0.71, 这说明截断误差分别占到了分析误差的 69% 和 79%。这一试验说明在所有网格点都有多个时刻的观测时, 背景场误差协方差在同化中的作用很小, 观测资料已经能够提供足够的信息来由高度场反演风场。不过在实际运用中很难满足这种条件。

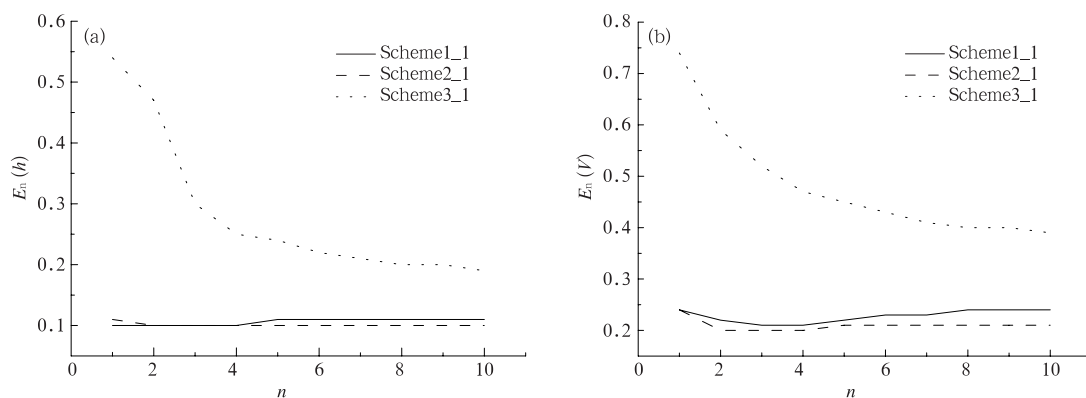


图 3 试验 1(所有格点均有高度观测)中 3 种方案得到的高度(a)和风(b)的分析误差随同化循环次数(n)的变化(粗实线: 方案 1; 间断线: 方案 2; 点线: 方案 3)

Fig. 3 Height rms errors index (a) and wind rms errors index (b) plotted as functions of cycle number (n) in experiment 1 (all height observations at each grid points are available; Scheme 1: solid line; Scheme 2: dash line; Scheme 3: dot line)

第 2 个试验是假定分析区域中只有 202 个点的高度观测。我们注意到这时候式(6)中的水平尺度 L 对方案 2 的同化结果影响明显, 因此在运用方案 2 时采用了不断变化 L (方案 2a)和固定 L (方案 2b)两种做法。我们统计了运用同化方案 2 时高度和风的背景误差特征尺度随同化循环的变化(图 4), 可以看出无论是对高度还是风, L 在头两个同化循环迅速减小, 以后就变化不大(对高度)或者缓慢下降(对风)。固定 L 的试验(方案 2b)是在初始时刻给 L 一个固定值并在同化循环中不加改变。不断变化 L 的试验(方案 2a)则是每个同化循环结束后重新计算 L 供下次同化用。从 3 种方案的分析误差随同化循环次数的变化(图 5)可以看到, 无论高度场(图 5a)还是风场(图 5b), 方案 1 的分析误差始终比其他两种方案大, 说明观测点比较稀疏时,

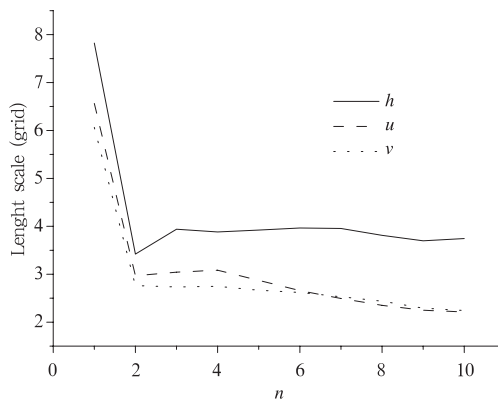


图 4 试验 2 中方案 2a 背景误差特征长度(单位: 网格距)随同化循环次数的变化

Fig. 4 The background error characteristic scale (unit: grid distance) of Scheme 2a plotted as functions of cycle number (n) in Experiment 2a (202 observations are available; solid line: height; dash line: u component; dot line: v component)

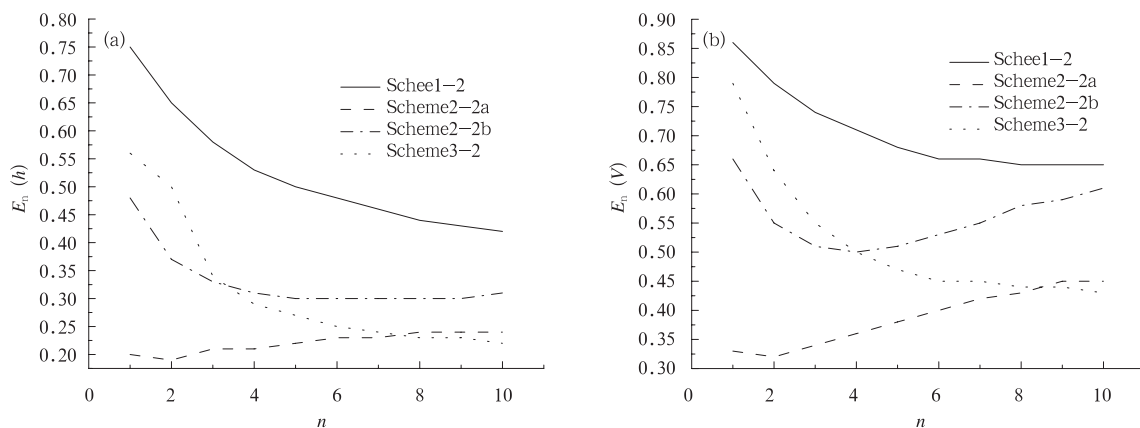


图 5 试验 2(202 个点有高度观测)中 3 种方案的高度(a)和风(b)的分析误差随同化循环次数(n)的变化

(粗实线:方案 1;间断线:方案 2;点线:方案 3)

Fig. 5 Height rms errors index (a) and wind rms errors index (b) plotted as functions of cycle number (n)

in Experiment 2 (202 height observations are available; Scheme 1: solid line;

Scheme 2a: dash line; Scheme 2b: dash dot line; Scheme 3: dot line)

模式方程的动力约束不足以从观测比较好地反演出初始场,这时背景误差协方差在同化中起到重要作用。在同化的头几个循环,方案 2 的分析误差比方案 3 小,但是随后误差有逐渐增大的趋势,特别是当 L 固定不变时(方案 2b),风的分析误差较大而且上升较快。如果能够随同化进程最优地选择 L ,方案 2a 会有很好的结果,只有从第 8 次同化循环开始方案 3 才略好于方案 2a。上述试验表明背景误差特征尺度对于 4DVAR 同样是很重要的。由于我们的试验中准确场是已知的,可以较好由预报误差的统计结构决定特征尺度,在实际运用中要比较难准确给出背景误差特征长度会有一些困难。

图 6 给出 3 种方案第一次同化得到的高度的分析增量和真实增量,3 种方案得到的分析增量场上主要的正负值区域和真实场都大体符合,但是方案 1 得到的分析增量场表现出许多小的扰动,说明方案 1 在拟合观测点的观测时不能将周围格点的模式变量很好地反演出来,同时观测中的随机误差对结果影响也比较大。方案 2b 得到的分析增量场和真实增量很接近。方案 3 得到的分析增量场结构与真实增量场结构大体是一致的,但强度较弱,这和截断误差较大有关。

在方案 3 中考虑了不同变量间的预报误差协相关,为了考察这种协相关对同化结果的影响,我们在方案 3 中让不同变量间的协相关为零,将同化结果

与原来的结果进行对比。结果表明,对于有观测的变量(高度场)而言,是否考虑变量间的协相关对分析结果的影响很小(图略),但是对于没有观测的变量(风场)而言(图 7),考虑不同变量间的协相关对同化有明显的改进,尤其对于观测资料比较稀疏的情形。

为了比较不同方案对观测误差的敏感性,在试验 2 中将观测误差增大 1 倍作同样的比较,结果见图 8。可以看到,观测误差增加后对方案 3 的影响很小,在第 3 个同化循环后它的分析误差已经小于方案 2,之后分析误差呈逐渐减小的趋势(高度)或略有增加(风场)。方案 1 和 2 对观测误差则比较敏感,随着同化循环的进行,风场分析误差增长会很快,特别是采用固定的误差特征长度(方案 2b)尤其明显,最后风的分析误差甚至会大于方案 1 的分析误差,这进一步看到合理地选择背景误差特征长度对方案 2 的重要性。

试验 3 是保留原来的观测误差不变而将观测点减少到 101 个,这时候方案 3 有较明显的优势(图 9),在第 5 次同化后其误差已经和方案 2 大体相当,以后分析误差一直呈现下降趋势,而其他两种方案的分析误差或下降很慢(高度场)或有所上升(风场),到 10 次同化结束时,方案 3 的分析误差已经明显小于前两个方案。如果特征长度不能准确给出,方案 2 的分析误差会增长很快。

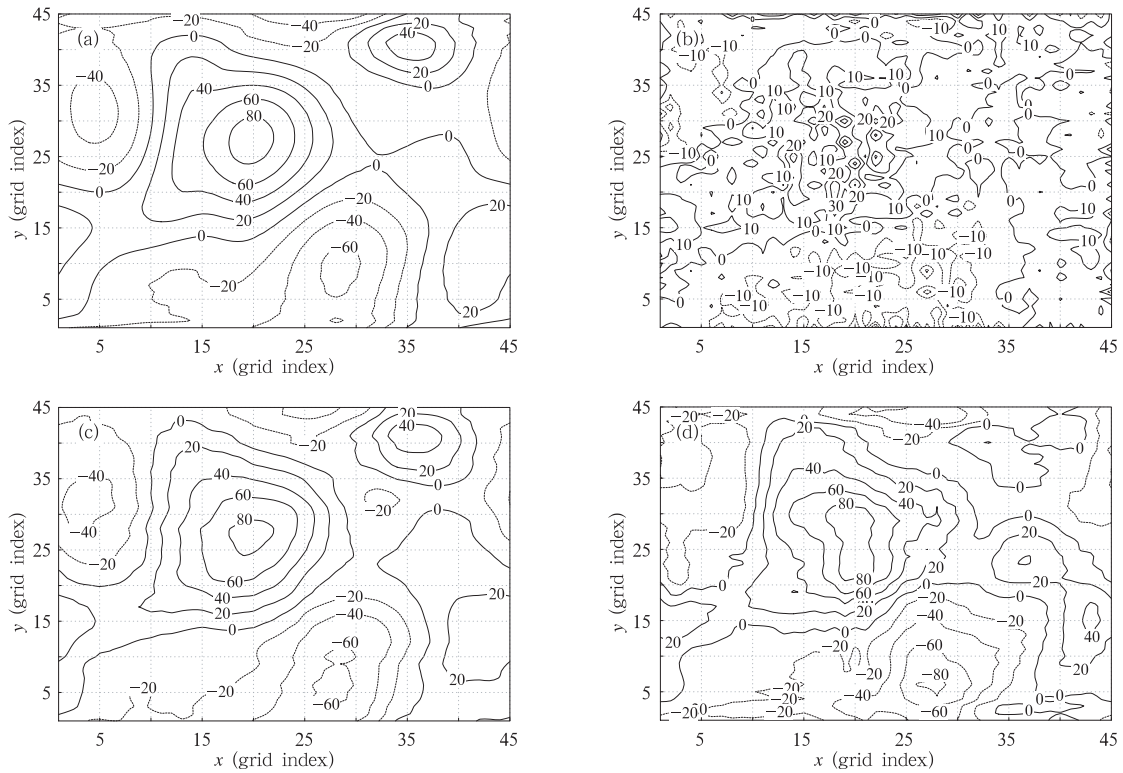


图6 试验2中第一个同化循环得到的高度场增量(单位:m)

(a. 真实增量, b. 方案1分析增量, c. 方案2a分析增量, d. 方案3分析增量)

Fig. 6 The true and analysis increment (unit: m) of height field for three schemes after one assimilation cycle in Experiment 2 (a. true increment, b. analysis increment of Scheme 1, c. analysis increment of scheme 2a, d. analysis increment of Scheme 3)

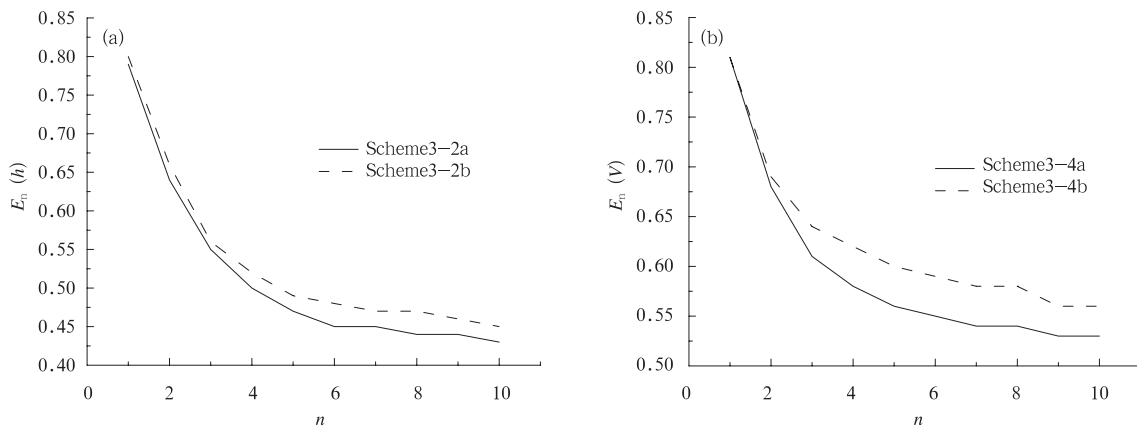


图7 在202个(a)和101个(b)点有高度观测时,不同变量间的协相关为零(间断线)

和不为零(实线)时方案3的风场分析误差随同化循环次数的变化

Fig. 7 Wind field rms errors index plotted as functions of cycle number (n) in Scheme 3 with (solid line) and without (dash line) correlations between different variables when 202 (a) and 101 (b) height observations are available

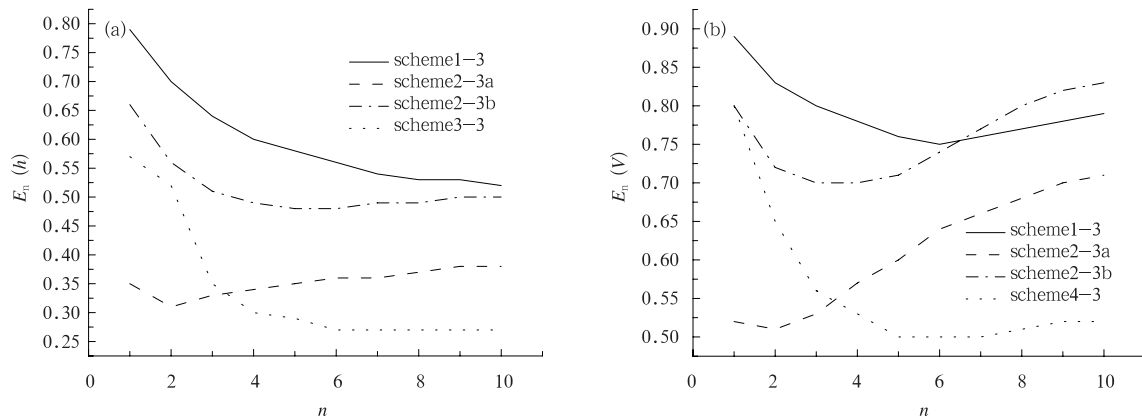


图 8 观测误差增大一倍,其余同图 5

Fig. 8 Same as Figure 5 except that the observation error is doubled

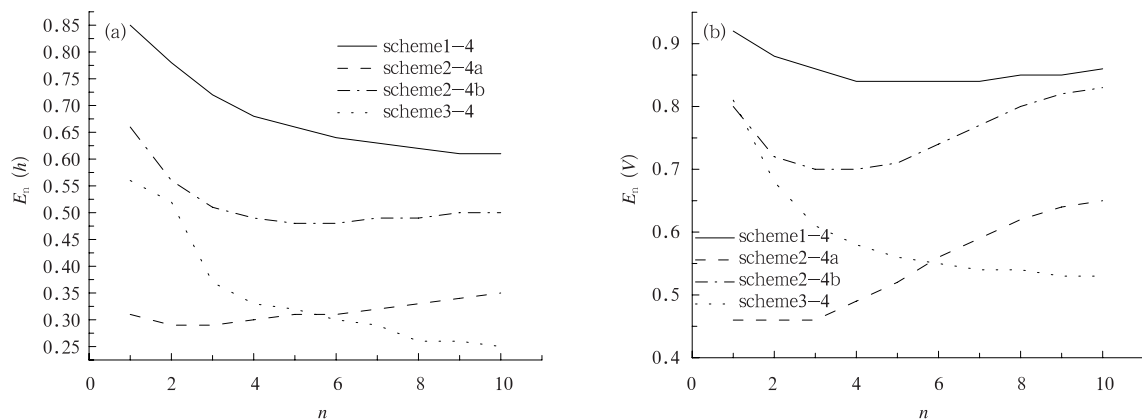


图 9 101 个点有高度观测,其余同图 5

Fig. 9 Same as Figure 5 except that only 101 height observations are available

5 结 论

一般认为背景误差协方差矩阵的结构在三维变分资料同化中起到决定性的作用,本文通过数值试验说明它对四维变分资料同化也有不可忽视的影响。当观测量的空间密度较低时,简单地将背景误差协方差矩阵当作对角矩阵得到的同化效果不好,将背景误差协方差矩阵的作用简化为高斯过滤的方法增加的计算量不大,但是可以明显改善同化结果。不过这一方案对背景误差特征长度比较敏感,如果能够比较好地给出背景误差特征长度,可望得到好的同化结果。本文将降维同化方法和 4DVAR 相结合,给出一种在 4DVAR 中利用依流型而变的背景误差协方差矩阵的实施方案,在计算中背景误差协方差矩阵并不以显式的方式出现在目标函数中,避免了背景误差协方差矩阵求逆的复杂运算。由于做

了降维处理,它要求的观测量可以比较少,当观测点稀疏时,这一方法会得到较好的结果。该方法对观测误差的敏感性也很小,应该更适用于同化观测误差比较大的各种遥感资料。但是除非用于产生预报集合的样本足够多,否则将分析变量按照截断的奇异向量展开时会有较大的截断误差,这限制了分析精度的提高,是该方法存在的一个问题。

References

- Cao Yanhua, Zhu Jiang, Navon I M, et al. 2007. A reduced order approach to four-dimensional variational data assimilation using proper orthogonal decomposition. *Int J Numer Meth Fluid*, 53: 1571-1583
- Cohn S, Todling R. 1996. Appropriate data assimilation schemes for stable and unstable dynamics. *J Meteor Soc Japan*, 74: 63-75
- Daley R. 1991. *Atmospheric Data Analysis*. Cambridge University Press, 457pp

- Dee D P, Da Silva A M. 1999. Maximum-Likelihood estimation of forecast and observation error covariance parameters. Part I: Methodology. *Mon Wea Rev*, 127: 1822-1834
- Derber J, Rosatti A. 1989. A global oceanic data assimilation system. *J Phys Oceanogr*, 19: 1333-1347
- Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. *J Geophys Res*, 99: 10143-10162
- Evensen G. 2003. The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn*, 53: 343-367
- Farrell B F, Ioannu P J. 2001. State estimation using a reduced order Kalman filter. *J Atmos Sci*, 58: 3666-3680
- Fisher M. 1998. Development of a simplified Kalman filter. ECMWF Technical memorandum 260. Available from ECMWF, Shinfield Park, Reading RG2 9AX, United Kingdom
- Gandin L S. 1965. Objective Analysis of Meteorology Fields (in Russian). Israel Program for Scientific Translation. 242pp
- Gong Jiandong, Wei Li, Tao Shiwei et al. 2006. Accurate estimation and application of 3-D error covariance structures in global data assimilation. Part I: Accurate estimation of error covariance in observation space. *Acta Meteor Sinica* (in Chinese), 64(6): 669-683
- Gong Jiandong, Zhao Gang. 2006. Accurate estimation and application of 3-D error covariance structures in global data assimilation. Part II: Background error covariance structure adjustments and numerical experiments. *Acta Meteor Sinica* (in Chinese), 64(6): 684-698
- Hamill T M, Snyder C. 2002. Using improved background-error covariances from an ensemble kalman filter for adaptive observations. *Mon Wea Rev*, 130: 1552-1572
- Holliingsworth A, Lonnberg P. 1986. The statistical structure of short-range forecast errors as determined from radiosond data. Part I: the wind field. *Tellus*, 38A: 111-136
- Houtekamer P L, Mitchell H L. 2001. A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon Wea Rev*, 129: 123-137
- Le Dimet F X, Talagrand O. 1986. Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, 38A: 97-111
- Lewis L, Deber J. 1985. The use of adjoint equations to solve a variational adjustment problem with advective constraint. *Tellus*, 37A: 309-322
- Lonnberg P, Holliingsworth A. 1986. The statistical structure of short-range forecast errors as determined from radiosond data. Part II: the covariance of height and wind error. *Tellus*, 38A: 137-161
- Lorenc A C. 1988. Optimal nonlinear objective analysis. *Q J R Meteor Soc*, 114: 205-240
- Lorenc A C. 2003. Modeling of error covariances by 4D-VAR data assimilation. *Q J R Meteor Soc*, 129: 3167-3182
- Park S K, Zupanski D. 2003. Four-dimensional variational assimilation for mesoscale and storm-scale applications. *Meteor Atmos Phys*, 82: 173-208
- Parrish D F, Derber J C. 1992. The national meteorological center's spectral statistical-interpolation analysis system. *Mon Wea Rev*, 120: 1747-1763
- Qiu C, Chou J. 2006. Four-dimensional data assimilation method based on SVD: Theoretical aspect. *Theor Appl Climat*, 83: 51-57
- Rutherford I D. 1972. Data assimilation by statistical interpolation of forecast error fields. *J Atmos Sci*, 29: 809-815
- Stephen E C. 1997. An introduction to estimation theory. *J Meteor Soc Japan*, 75: 257-288
- Talagrand O, Courtier P. 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation I: Theory. *Q J R Meteor Soc*, 113: 1311-1328
- Tippett M K, Chang P. 2003. Some theoretical considerations on predictability of linear stochastic dynamics. *Tellus*, 55A: 148-157
- Verlaan M, Heemink A W. 1997. Tidal flow forecasting using reduced rank square root filters. *Stoch Hydrol Hydraul*, 11(5): 349-368
- Zhang S, Anderson J L. 2003. Impact of spatially and temporally varying estimates of error covariance on assimilation in a simple atmospheric model. *Tellus*, 55A: 126-147
- Zou X, Vandenberghe F, Pondeva M, et al. 1997. Introduction to adjoint techniques and the MM5 adjoint modeling system. NCAR Tech. Note. NCAR/TN-435-STR, 107pp

附中文参考文献

- 龚建东, 魏丽, 陶士伟等. 2006. 全球资料同化中误差协方差三维结构的准确估计与应用 I: 观测空间协方差的准确估计. *气象学报*, 64(6): 669-683
- 龚建东, 赵刚. 2006. 全球资料同化中误差协方差三维结构的准确估计与应用 II: 背景误差协方差调整与数值试验分析. *气象学报*, 64(6): 684-698