GRAPeS 三维变分同化系统的理想试验*

张 华 薛纪善 庄世宇 朱国富 朱宗申

(中国气象科学研究院数值预报研究中心, 北京, 100081)

摘 要

文中介绍了一种新的、适合格点模式的三维变分同化方案 GRAPeS-3DVAR。该方案采用相互独立的流函数、非平衡速度势函数、非平衡位势和水汽作为分析变量。通过变量变换对目标函数进行预调节, 不仅避免了直接计算背景误差协方差逆矩阵的困难, 而且改善了 Hessian 矩阵的性状, 提高了收敛速度。采用 EOF 分解方法, 将三维分析变量投影到垂直模态上, 分解成为二维场; 水平方向采用数字(递归)滤波器代替矩阵运算, 实现和简化了方案的求解。此外, 还考虑了质量场和风场之间的平衡约束关系。理想试验结果表明, GRAPeS-3DVAR 能够正确地反映多变量之间相互作用关系, 收敛迅速, 分析结果合理。

关键词: 三维变分同化, 变量变换, 预调节。

1 引 言

目前可获得的有效的大气信息来源主要有 3 种: (1) 观测资料; (2) 气象数值预报模式(包含大气运动控制规律); (3) 已知的大气运动观测误差统计特性。一个理想的资料同化系统应该能够充分有效地利用人类可获取的各种大气信息, 从理论上来讲, 三维变分同化方法可以满足上述要求, 它可以使用其分析变量能够描述的所有观测资料。与传统资料分析方法如最优插值 OI(Optimum Interpolation) 方法相比, 特别是在应用非常规观测资料如卫星、雷达等方面, 三维变分同化方法具有更加明显的优势[1,2]。此外, 在具体计算过程中, OI 方法是在人为分割的小区域内进行, 存在"资料选择"问题; 而三维变分方法则在全场进行分析, 其结果更加合理, 更加协调。

三维变分同化方法原理清晰, 但对其直接求解却非常困难, 主要原因是其目标函数所包含的背景误差协方差逆矩阵的计算非常困难, 此矩阵不仅规模巨大(阶数超过 $10^7 \times 10^7$), 而且还常常是病态的。在谱模式变分同化方案中, 可以利用其谱系数的正交性来解决这一问题[3,4]; 但对于格点模式变分同化方案, 问题则复杂得多。本项工作综合国际上对三维变分同化的最新研究进展, 根据 Lorenc[5] 和 Barker[6] 的工作, 结合中国实际, 设计了一种适合格点模式的三维变分同化方案; 该方案不仅能够避免直接计算背景误差协方差逆矩阵的困难, 而且在实际计算过程中收敛迅速。本文仅给出理想试验结果, 关于实际观测资料分析效果, 将另文详细介绍。

2 方案介绍——目标函数

2.1 目标函数及其梯度

三维变分同化方法的基本思想是将资料同化归结为一个二次泛函极小化问题。该泛函一般定义为:

$$\boldsymbol{J} = \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_\mathrm{b})^\mathrm{T}\boldsymbol{B}^{-1}(\boldsymbol{x} - \boldsymbol{x}_\mathrm{b}) + \frac{1}{2}(\boldsymbol{H}\boldsymbol{x} - \boldsymbol{y}_\mathrm{o})^\mathrm{T}\boldsymbol{O}^{-1}(\boldsymbol{H}\boldsymbol{x} - \boldsymbol{y}_\mathrm{o}) = \boldsymbol{J}_\mathrm{b} + \boldsymbol{J}_\mathrm{o} \tag{1}$$

其梯度为

$$\therefore \boldsymbol{J} = \boldsymbol{B}^{-1}(\boldsymbol{x} - \boldsymbol{x}_\mathrm{b}) + \boldsymbol{H}'^\mathrm{T}\boldsymbol{O}^{-1}(\boldsymbol{H}\boldsymbol{x} - \boldsymbol{y}_\mathrm{o}) \tag{2}$$

其中, $\boldsymbol{x}$ 是长度为 N 的分析变量向量, $\boldsymbol{x}_\mathrm{b}$ 是长度为 N 的背景向量, $\boldsymbol{y}_\mathrm{o}$ 是长度为 L 的观测向量, $\boldsymbol{B}^{-1}$ 是阶数为 $N \times N$ 背景误差协方差矩阵的逆, N 是分析

* 初稿时间: 2003 年 7 月 22 日; 修改稿时间: 2003 年 9 月 12 日。
资助课题: 十五国家重点科技攻关计划项目(2001BA607B)"中国气象数值预报系统技术创新研究"。

场的自由度数目, $\boldsymbol{O}^{-1}$ 是阶数为 $L \times L$ 观测误差协方差矩阵的逆, L 是观测数目。$\boldsymbol{H}$ 为观测算子, 它代表模式空间向观测空间的一种映射, 当 $\boldsymbol{y}_o$ 与 $\boldsymbol{x}$ 为相同类型的物理变量时, 观测算子简化为一个简单的插值算子; 如果 $\boldsymbol{y}_o$ 和 $\boldsymbol{x}$ 具有不同的物理属性, 这时观测量为分析变量的一些泛函信息, 算子 $\boldsymbol{H}$ 为模式空间向观测空间的具有某种复杂结构的映射算子, 如对卫星辐射率资料来说, 算子 $\boldsymbol{H}$ 即为辐射传输模式; $\boldsymbol{H}' = \therefore \boldsymbol{H}$ 为切线性观测算子。$\boldsymbol{J}_b$ 为背景项, $\boldsymbol{J}_o$ 为观测项, 上标 T 和 - 1 分别代表矩阵的转置和逆。

根据 Courtier[7], 本方案采用增量分析法, 目的是减少每一个迭代步的计算量, 令

$$\delta \boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}_b$$

对 $\boldsymbol{H}(\boldsymbol{x})$ 在 $\boldsymbol{x}_b$ 作泰勒展开, 截取前两项

$$\boldsymbol{H}(\boldsymbol{x}) = \boldsymbol{H}(\boldsymbol{x}_b + \delta \boldsymbol{x}) \approx \boldsymbol{H}(\boldsymbol{x}_b) + \boldsymbol{H}' \delta \boldsymbol{x}$$

令

$$\boldsymbol{d} = \boldsymbol{H}(\boldsymbol{x}_b) - \boldsymbol{y}_o$$

$$\boldsymbol{J} = \frac{1}{2} \delta \boldsymbol{x}^{\mathrm{T}} \boldsymbol{B}^{-1} \delta \boldsymbol{x} + \frac{1}{2} (\boldsymbol{H}' \delta \boldsymbol{x} + \boldsymbol{d})^{\mathrm{T}} \cdot \boldsymbol{O}^{-1} (\boldsymbol{H}' \delta \boldsymbol{x} + \boldsymbol{d}) \quad (3)$$

其梯度为

$$\therefore_{\delta \boldsymbol{x}} \boldsymbol{J} = \boldsymbol{B}^{-1} \delta \boldsymbol{x} + \boldsymbol{H}'^{\mathrm{T}} \boldsymbol{O}^{-1} (\boldsymbol{H}' \delta \boldsymbol{x} + \boldsymbol{d}) \quad (4)$$

其二阶导数(Hessian 矩阵) 为

$$\therefore^2_{\delta \boldsymbol{x}} \boldsymbol{J} = \boldsymbol{B}^{-1} + \boldsymbol{H}'^{\mathrm{T}} \boldsymbol{O}^{-1} \boldsymbol{H}' \quad (5)$$

2.2 分析变量

在实际问题中, 分析变量由一组具体的气象变量场构成, 假定它们就是模式所用的变量场。那么是否能够直接根据式(3) 来求得它们的最优解呢? 众所周知, 各个模式变量之间存在一些约束关系, 它们是彼此相关的。为了简化同化分析问题, 最好采用彼此之间相互独立的量作为分析变量(这可以使背景误差协方差矩阵成为事实上的块对角矩阵)。通常的作法是取模式变量的非平衡部分作为分析变量。很显然, 它们是互不相关的。这样作还容易控制非平衡模态的增长。

根据上述原则, 分析变量取为

第一个变量为流函数 Ψ, 非平衡速度势函数 χ_u

$$\chi_u = \chi - \chi_g \quad (6)$$

第二个变量为非平衡位势 ϕ_u

$$\phi_u = \phi - \phi_g \quad (7)$$

χ_g, ϕ_g 分别为速度势和位势高度场平衡的部分。本文暂不考虑 χ_g 的存在, 令 $\chi_g = 0$; ϕ_g 是通过某种平衡关系得到的, 目前仅考虑地转平衡, ϕ_u 和风场在统计上是不相关的。

第三个变量为比湿 q 或相对湿度 R_h。

3 预调节——变量变换

预调节(Preconditioning) 的原因主要有两个: 其一 $\boldsymbol{B}^{-1}$ 往往是病态的, 需对其进行调节; 其二是背景误差协方差逆矩阵 $\boldsymbol{B}^{-1}$ 太大($10^7 \times 10^7$)。以 1997 年 ECMWF 为例, 其所需的存档能力是 ECMWF 的 1000 倍, 内存是当时的 10^7 倍[8], 现有的计算机都无法处理。

数学上, 矩阵的条件数描述了数值问题的性状(病态状况), 它是相对误差的放大倍数。在三维变分同化方法中, 计算过程收敛的性质与极小点处 Hessian 矩阵 $\boldsymbol{J}''$ 的特征值有关。有以下定理:

设 $\boldsymbol{J}(\boldsymbol{x})$ 存在连续二阶偏导数, $\boldsymbol{x}^*$ 是局部极小点, Hessian 矩阵 $\boldsymbol{J}''(\boldsymbol{x}^*)$ 最小特征值 $\lambda_{min} > 0$, 最大特征值 λ_{max}, 算法产生的序列 $\{\boldsymbol{x}^k\}$ 收敛于 $\boldsymbol{x}^*$, 则目标函数系列 $\{\boldsymbol{J}(\boldsymbol{x}^k)\}$ 以不大于 $(\lambda_{max} - \lambda_{min})^2/(\lambda_{max} + \lambda_{min})^2$ 的收敛比收敛于 $\boldsymbol{J}(\boldsymbol{x}^*)$。

若令 $r = \lambda_{max}/\lambda_{min}$ 为对称正定矩阵 $\boldsymbol{J}''(\boldsymbol{x}^*)$ 的条件数, 该定理表明条件数越小(接近 1) 收敛越快。条件数还反映了等值面的椭圆率(偏心率), 由于梯度并不正好指向最小值, 所以条件数描述了极小化问题的病态状况。条件数越大, 病态问题就越严重, 极小化过程就越难收敛。

根据式(5), 目标函数 $\boldsymbol{J}$ 的 Hessian 矩阵由两项组成, 一项是 $\boldsymbol{B}^{-1}$, 另一项是 $\boldsymbol{H}'^{\mathrm{T}} \boldsymbol{O}^{-1} \boldsymbol{H}'$, 其中第二项的性状难以确定, 且当观测误差互不相关时, 它是一个对角矩阵或阶数很少的块对角矩阵, 因此这里只考虑 $\boldsymbol{B}^{-1}$ 的性状对最优化的影响。由于背景误差协方差矩阵的逆 $\boldsymbol{B}^{-1}$ 是对称矩阵, 根据有关定理, 用于描述 $\boldsymbol{B}^{-1}$ 性状的条件数则取决于 $\boldsymbol{B}^{-1}$ 的最大特征值 λ_{max} 和 λ_{min} 之比, 即

$$\mathrm{cond}(\boldsymbol{J}'') = \lambda_{max}/\lambda_{min}$$

一般来说, 背景场误差主要来源于平衡部分, 它们对应于 $\boldsymbol{B}^{-1}$ 的小特征值; 而同时也存在一些非平衡的误差, 它们对应于 $\boldsymbol{B}^{-1}$ 的大特征值。这使得条件数是一个大值, 故 $\boldsymbol{J}''$ 或 $\boldsymbol{B}^{-1}$ 接近于病态。这给式(3) 极小化的求解带来困难。

对目标函数求极值需要采用下降算法 $\boldsymbol{B}^{-1}$ 的大条件数会减缓或停止收敛; 反之, 条件数接近 1 时会加速收敛。为了改善收敛性, 应设法先降低 $\boldsymbol{B}^{-1}$ 的条件数, 这就是预调节方法。

为了解决上述问题, 本方案通过变量变换对目标函数进行预调节。

对 $\boldsymbol{B}$ 分解如下

$$\boldsymbol{B} = \boldsymbol{U}\boldsymbol{U}^{\mathrm{T}} \tag{8}$$

定义新分析变量 $\boldsymbol{w}$

$$\delta\boldsymbol{x} = \boldsymbol{U}\boldsymbol{w} \tag{9}$$

将其代入目标函数 $\boldsymbol{J}$ 并对其求二阶导数, 可得其 Hessian 矩阵

$$\nabla^2_{\boldsymbol{w}}\boldsymbol{J} = \boldsymbol{I} + \boldsymbol{U}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{O}^{-1}\boldsymbol{H}\boldsymbol{U} \approx \boldsymbol{I} \tag{10}$$

与原目标函数 $\boldsymbol{J}$ 的 Hessian 矩阵

$$\nabla^2_{\delta\boldsymbol{x}}\boldsymbol{J} = \boldsymbol{B}^{-1} + \boldsymbol{H}^{\mathrm{T}}\boldsymbol{O}^{-1}\boldsymbol{H} \approx \boldsymbol{B}^{-1}$$

相比, 其条件数有了明显地改善。

3.1　物理变换

从分析变量空间到模式变量空间的变换 $\boldsymbol{U}$ 包括物理变换 $\boldsymbol{U}_{\mathrm{p}}$ 和空间变换 U_{s}, 即

$$\delta\boldsymbol{x} = \boldsymbol{U}\boldsymbol{w} = \boldsymbol{U}_{\mathrm{p}}\boldsymbol{U}_{\mathrm{s}}\boldsymbol{w} \tag{11}$$

在实际预报中, 模式变量(误差)在物理上是彼此相关的, 为了解决 $\boldsymbol{B}^{-1}$ 太大无法计算的困难, 首先需对模式变量进行变换, 目标是将模式大气变量分解成不相关但物理真实的误差模态。

$$\delta\boldsymbol{x} = \boldsymbol{U}_{\mathrm{p}}\boldsymbol{w}_{\mathrm{v}} \tag{12}$$

物理变换是从分析变量到模式变量的变换, 它利用平衡关系和无辐散关系设计 $\boldsymbol{U}_{\mathrm{p}}$。

3.1.1　风场变换

从分析变量到模式变量需经过 3 个阶段:

(1) 计算涡度 ζ 和散度 D

$$\boldsymbol{v}_\eta = (\boldsymbol{u}, \boldsymbol{v})_\eta = \boldsymbol{k} \times \nabla_\eta \Psi + \nabla_\eta \chi \tag{13}$$

式中下标 η 表示垂直坐标的普遍形式。

$$\begin{aligned} \zeta &= \nabla_\eta \times \boldsymbol{v}_\eta = \nabla^2_\eta \Psi \\ D &= \nabla_\eta \cdot \boldsymbol{v}_\eta = \nabla^2_\eta \chi \end{aligned} \tag{14}$$

(2) 解 Poisson 方程

$$\begin{aligned} \Psi &= \nabla^{-2}\zeta \\ \chi &= \nabla^{-2} D \end{aligned} \tag{15}$$

(3) 用式(6)计算非平衡速度势函数 χ_{u}。

3.1.2　质量场变换

根据散度方程

$$\nabla^2_\eta \phi = -\nabla_\eta \cdot \left[\left(\frac{\partial \boldsymbol{v}}{\partial t}\right)_\eta + \dot{\eta}\frac{\partial \boldsymbol{v}}{\partial \eta} + \boldsymbol{v}\cdot\nabla_\eta \boldsymbol{v} + f\boldsymbol{k}\times\boldsymbol{v} - \boldsymbol{F}\right] = T_1 + T_2 + T_3 + T_4 + T_5 \tag{16}$$

其中 T_1 是局地变化项, T_2 是垂直对流项, T_3 是非线性平衡方程中的一项, 它包括旋衡风场和质量场平衡效应, 且可以改善热带气旋的质量场分析, T_4 是最常见的地转平衡项, T_5 是摩擦项, 目前在本方案中仅考虑地转平衡项, 由此可导出地转平衡方程

$$\begin{aligned} \nabla^2 \phi_g &= \nabla\cdot f\nabla\Psi \\ &= \nabla f\cdot\nabla\Psi + f\nabla^2\Psi \end{aligned} \tag{17}$$

$$\phi_g = \nabla^{-2}\left[\nabla f\cdot\nabla\Psi + f\nabla^2\Psi\right] \tag{18}$$

用式(7)计算非平衡位势 ϕ_{u}。

3.2　空间变换

经过物理变换后, 各分析变量(误差)仍然存在水平和垂直相关, 必须进行进一步的变换, 空间变换包括垂直变换和水平变换, 即

$$\boldsymbol{U}_{\mathrm{s}} = \boldsymbol{U}_{\mathrm{v}}\boldsymbol{U}_{\mathrm{h}}$$

3.2.1　垂直变换

在三维场中, 空间点与点间仍存在着相关, 本方案采用 EOF 分解方法, 将分析变量投影到垂直模态上, 利用主分量的正交性, 使三维场分解成二维场。

$$\boldsymbol{B}_{\mathrm{v}} = \boldsymbol{E}\Lambda\boldsymbol{E}^{\mathrm{T}} \tag{19}$$

式中 $\boldsymbol{B}_{\mathrm{v}}$ 为垂直背景误差协方差矩阵, 是一个 $K\times K$ 阶对称正定矩阵, K 是垂直层数, Λ 是含有 K 个特征值 $\lambda(m)$ 的对角矩阵。

矩阵 $\boldsymbol{E}$ 的列是 $\boldsymbol{B}_{\mathrm{v}}$ 的 K 个特征值 $e(m)$, 它服从正交关系

$$\boldsymbol{E}\boldsymbol{E}^{\mathrm{T}} = \boldsymbol{I} \tag{20}$$

此外

$$\boldsymbol{B}_{\mathrm{v}} = \boldsymbol{U}_{\mathrm{v}}\boldsymbol{U}_{\mathrm{v}}^{\mathrm{T}} \tag{21}$$

式(21)与(19)相比可得

$$\boldsymbol{w}_{\mathrm{v}} = \boldsymbol{U}_{\mathrm{v}}\boldsymbol{w}_{\mathrm{h}} = \boldsymbol{E}\Lambda^{1/2}\boldsymbol{w}_{\mathrm{h}} \tag{22}$$

$\boldsymbol{U}_{\mathrm{v}}$ 变换具有以下特点: (1) 明显地节省机时; (2) 起预调节作用; (3) 特征向量是根据特征值大小排序, 即 $\lambda(m=1)$ 是主结构而 $\lambda(m=K)$ 包含低幅噪音, 因此在业务上可以通过略去对整个误差没有多少贡献的小尺度特征结构从系统中过滤次网格尺度噪音。

3.2.2　水平变换

经过垂直变换后, 垂直模态之间相互正交, 没有相关。因此, 可对各垂直模态二维场单独进行处理, 进行水平变换

$$\boldsymbol{w}_{\mathrm{h}} = \sqrt{\boldsymbol{B}_{\mathrm{h}}}\,\boldsymbol{w} \tag{23}$$

目标函数成为

$$\boldsymbol{J}(\boldsymbol{w}) = \frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} + \frac{1}{2}(\boldsymbol{H}'\boldsymbol{U}_{\mathrm{p}}\boldsymbol{U}_{\mathrm{v}}\sqrt{\boldsymbol{B}_{\mathrm{h}}}\boldsymbol{w} + \boldsymbol{d})^{\mathrm{T}} \cdot \boldsymbol{O}^{-1}(\boldsymbol{H}'\boldsymbol{U}_{\mathrm{p}}\boldsymbol{U}_{\mathrm{v}}\sqrt{\boldsymbol{B}_{\mathrm{h}}}\boldsymbol{w} + \boldsymbol{d}) \tag{24}$$

梯度

$$\nabla_{\boldsymbol{w}}\boldsymbol{J} = \boldsymbol{w} + \boldsymbol{H}'^{\mathrm{T}} \cdot \boldsymbol{O}^{-1} \cdot [\boldsymbol{H}'(\boldsymbol{U}_{\mathrm{p}}\boldsymbol{U}_{\mathrm{v}}\sqrt{\boldsymbol{B}_{\mathrm{h}}}\boldsymbol{w}) + \boldsymbol{d}] \tag{25}$$

式中 $\boldsymbol{B}_{\mathrm{h}}$ 为经过垂直变换后背景场误差协方差矩阵垂直模态二维场，按照下降算法迭代给出最后的 $\boldsymbol{w}$，由 $\boldsymbol{x} = \boldsymbol{x}_{\mathrm{b}} + \boldsymbol{U}_{\mathrm{p}}\boldsymbol{U}_{\mathrm{v}}\sqrt{\boldsymbol{B}_{\mathrm{h}}}\boldsymbol{w}$ 得到分析场。不过在实际问题中 $\sqrt{\boldsymbol{B}_{\mathrm{h}}}$ 仍是一个维数很高的矩阵（可达 $10^5 \times 10^5$，甚至更大），而且每次迭代都要作运算，占用空间很大，计算量也不小。

分析矩阵 $\sqrt{\boldsymbol{B}_{\mathrm{h}}}$ 和向量 W_{h} 相乘的作用

$$w_i = \sum_{j=1}^{N} b_{i,j} w_{\mathrm{h}j} \tag{26}$$

$b_{i,j}$ 是 $\sqrt{\boldsymbol{B}_{\mathrm{h}}}$ 的元素，在均匀格点下，上式相当于一个空间滤波运算，$b_{i,j}$ 是滤波系数，过滤范围是整个模式区间。按照这样的分析，可以用空间滤波运算 $\boldsymbol{U}_{\mathrm{h}}$ 代替 $\sqrt{\boldsymbol{B}_{\mathrm{h}}}$，这样的方法简称 VAF（variational analysis using a filter）。根据 Lorence[9]，采用递归滤波方法。关于空间滤波算子 $\boldsymbol{U}_{\mathrm{h}}$，可以根据 $\sqrt{\boldsymbol{B}_{\mathrm{h}}}$ 的特性进行设计。假定背景场误差各项均匀同性，具有高斯分布型联合概率密度，即可用高斯分布表示背景场误差相关函数。首先考虑一维问题，设计一个递归滤波器如下：

（1）递归滤波（Recursive Filters）器的定义

考虑一格距为 $\delta\boldsymbol{x}$ 的一维格点场，A_i 是格点的原始值，B_i 是经过 $i = 1 \cdots n$ 滤波后的值，C_i 是经过左右两个方向滤波后的值，a 是滤波系数。则递归滤波器定义为：

$$\begin{aligned} B_i &= \alpha B_{i-1} + (1-\alpha)A_i \qquad i = 1, 2 \cdots n \\ C_i &= \alpha C_{i+1} + (1-\alpha)B_i \qquad i = n \cdots 2, 1 \end{aligned} \tag{27}$$

（2）递归滤波器的边界条件

如果所有观测都位于滤波区内，被滤波的初始场在边界以外为零。为确定边界条件，将滤波区域延伸到无穷远。根据 Hayden 和 Purser[10] 给定的边界条件，有

一次滤波：
$$B_1 = (1-\alpha)A_1 \tag{28}$$

二次滤波：
$$(C_n, B_1) = \frac{(B_n, A_1)}{1+\alpha} \tag{29}$$

二次以上滤波：
$$(C_n, B_1) = \frac{1-\alpha}{(1-\alpha^2)^2}[(B_n, A_1) - \alpha^3(B_{n-1}, A_2)] \tag{30}$$

（3）滤波器分析

将方程组（27）中两式合并，有：

$$A_i = C_i - \frac{\alpha}{(1-\alpha)^2}(C_{i-1} - 2C_i + C_{i+1}) \tag{31}$$

式（31）亦称反滤波器，从该式中可以求得其谱响应函数，并进而得到递归滤波器的谱响应函数（为反滤波器谱响应函数的倒数）。用波长为 $\frac{2\pi}{k}$ 的波动解表示 A_i 和 C_i，即令：

$$\begin{aligned} A_i &= \sum_k A_k \sin k(i-1)\delta\boldsymbol{x} \\ C_i &= \sum_k C_k \sin k(i-1)\delta\boldsymbol{x} \end{aligned} \tag{32}$$

以波动解代入式（31）右边，得

$$\begin{aligned} &C_i - \frac{\alpha}{(1-\alpha^2)}(C_{i-1} - 2C_i + C_{i+1}) \\ &= \sum_k C_k \left\{ \sin[k(i-1)\delta\boldsymbol{x}] - \frac{\alpha}{(1-\alpha)^2}(\sin[k(i-2)\cdot \delta\boldsymbol{x}] - 2\sin[k(i-1)\delta\boldsymbol{x}] + \sin[ki\delta\boldsymbol{x}]) \right\} \\ &= \sum_k C_k \left\{ \sin[k(i-1)\delta\boldsymbol{x}] + \frac{\alpha}{(1-\alpha)^2} \cdot 2\sin[k(i-1)\delta\boldsymbol{x}][1-\cos(k\delta\boldsymbol{x})] \right\} \\ &= \sum_k C_k \sin[k(i-1)\delta\boldsymbol{x}] \left\{ 1 + \frac{2\alpha}{(1-\alpha)^2}[1-\cos(k\delta\boldsymbol{x})] \right\} \end{aligned}$$

比较上式与式（32）的第一式，可得：

$$\begin{aligned} A_k &= C_k \left\{ 1 + \frac{2\alpha}{(1-\alpha)^2}[1-\cos(k\delta\boldsymbol{x})] \right\} \\ &= C_k \left\{ 1 + \frac{\alpha}{(1-\alpha)^2}[2\sin(\frac{k\delta\boldsymbol{x}}{2})]^2 \right\} \end{aligned}$$

即

$$S_{1\mathrm{Pass}}(k) = \frac{1}{1 + \frac{\alpha}{(1-\alpha)^2}\left[2\sin\left(\frac{k\delta\boldsymbol{x}}{2}\right)\right]^2} \tag{33}$$

这是一次向前、一次向后递归滤波的振幅响应因子。对于 N_r 次滤波的振幅响应因子：

$$S_{N\mathrm{Pass}}(k) = \left\{ 1 + \frac{\alpha}{(1-\alpha)^2}[2\sin(\frac{k\delta\boldsymbol{x}}{2})]^2 \right\}^{-n} \tag{34}$$

如果 $k\delta\boldsymbol{x}$ 足够小，即 $k\delta\boldsymbol{x} \to 0$，则其振幅响应因子为：

$$S_{N\mathrm{pass}}(k) \approx 1 - \frac{\alpha}{(1-\alpha)^2}N_r k^2 \delta\boldsymbol{x}^2 \tag{35}$$

（4）滤波系数的确定

滤波系数的确定需满足使滤波器的作用与用一

个协方差函数对一个场卷积的效果相似, 下面讨论两种相关函数。

(1) 高斯型相关函数(GAUSSIAN)

考虑一维情况, 有

$$b(x_i, x_j) = \varepsilon_b^2 \exp(-(x_i - x_j)^2/2R^2) \quad (36)$$

其中, ε_b^2 是方差, R 是影响半径。则其谱响应为

$$S_{\text{Gaussian}}(k) = \varepsilon_b^2 \sqrt{2\pi} R \exp(-R^2 k^2/2) \quad (37)$$

故递归滤波器式(27)的结果需乘一个因子 β

$$\beta = \varepsilon_b^2 \sqrt{2\pi} R \quad (38)$$

将式(37)中的 e 指数展开取前两项, 与式(35)对比

$$\frac{\alpha}{(1-\alpha)^2} N_r \delta \boldsymbol{x}^2 = \frac{R^2}{2} \quad (39)$$

令

$$E = N_r \delta \boldsymbol{x}^2 / R^2 \quad (40)$$

可得

$$\alpha = 1 + E - \sqrt{E(E+2)} \quad (41)$$

(2) 二阶自回归型相关函数(SOAR)

研究表明, Gaussian 型相关函数对小尺度作用不大, 这时可采用一个更有力的二阶自回归型相关函数(Second-order auto-regressive (SOAR))。SOAR 模型为

$$b(x_i, x_j) = \varepsilon_b^2 (1 + |x_i - x_j|/R) \cdot \exp(-|x_i - x_j|/R) \quad (42)$$

其谱响应为:

$$S_{\text{SOAR}}(k) = \varepsilon_b^2 \frac{4s}{(1 + k^2 R^2)^2} \quad (43)$$

$$E = N_r \delta \boldsymbol{x}^2 / 4R^2 \quad (44)$$

上式代入式(41)可得 α。

从前面递归滤波的叙述可以知道, 背景误差协方差矩阵 $\boldsymbol{B}_h$ 可以用空间滤波运算代替, 对于 Gaussian 型函数, 在二维情况下, 它相当于两个独立方向的滤波的乘积。

$$\boldsymbol{B}_h \approx \underbrace{\boldsymbol{R}_F \boldsymbol{R}_F \cdots \boldsymbol{R}_F}_{N_r} \quad (45)$$

这里 $\boldsymbol{R}_F$ 是一次向前一次向后的递归滤波。由于递归滤波是对称的, 故式(45)可写为:

$$\boldsymbol{B}_h \approx \underbrace{\boldsymbol{R}_F \boldsymbol{R}_F \cdots \boldsymbol{R}_F}_{N_r/2} \underbrace{\boldsymbol{R}_F^T \boldsymbol{R}_F^T \cdots \boldsymbol{R}_F^T}_{N_r/2} \quad (46)$$

所以

$$\sqrt{\boldsymbol{B}_h} \approx \underbrace{\boldsymbol{R}_F \boldsymbol{R}_F \cdots \boldsymbol{R}_F}_{N_r/2} \quad (47)$$

即 $\sqrt{\boldsymbol{B}_h}$ 可以用 $N_r/2$ 次向前—向后递归滤波来逼近。

4　背景误差协方差矩阵

4.1　背景误差协方差矩阵的结构

背景误差协方差矩阵定义为

$$\boldsymbol{B} = E\left\{(\boldsymbol{x}_b - \boldsymbol{x}_t)(\boldsymbol{x}_b - \boldsymbol{x}_t)^T\right\} \quad (48)$$

为了进一步考察 $\boldsymbol{B}$ 的结构, 先研究二维问题, 假设随机变量 x, y 具有高斯分布型联合概率密度

$$p(x, y) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left[-\frac{1}{2}\left(\frac{x - \overline{x}}{\sigma_x}\right)^2\right] \times \frac{1}{\sqrt{2\pi}\sigma_y(1-\rho_{xy}^2)} \times \exp\left[-\frac{1}{2}\left(\frac{y - \overline{y} - \rho_{xy}\sigma_x/\sigma_y)(x - \overline{x})}{\sqrt{1-\rho_{xy}^2}\sigma_y}\right)^2\right] \quad (49)$$

式中 $\overline{x}$, $\overline{y}$ 为随机变量 x, y 的均值, ρ_{xy} 为变量 x, y 相关系数, 相应式(48), 二维随机变量 $(x, y)^T$ 的误差协方差矩阵为

$$\boldsymbol{P}_{xy} = E\left\{\left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} \overline{x} \\ \overline{y} \end{bmatrix}\right)\left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} \overline{x} \\ \overline{y} \end{bmatrix}\right)^T\right\} = \begin{pmatrix} E\left\{(x-\overline{x})(x-\overline{x})^T\right\} & E\left\{(y-\overline{y})(x-\overline{x})^T\right\} \\ E\left\{(x-\overline{x})(y-\overline{y})^T\right\} & E\left\{(y-\overline{y})(y-\overline{y})^T\right\} \end{pmatrix} \quad (50)$$

根据数学上期望的定义

$$E\{f\} = \int f(x, y) p(x, y) \, dx \, dy \quad (51)$$

可得

$$\boldsymbol{P}_{xy} = \begin{pmatrix} \sigma_x^2 & \sigma_x \sigma_y \rho_{xy} \\ \sigma_x \sigma_y \rho_{xy} & \sigma_y^2 \end{pmatrix} \quad (52)$$

推广到 n 维情形

$$\boldsymbol{B} = \begin{pmatrix} (\sigma_x^2)_{i,i} & (\sigma_x \sigma_y \rho_{xy})_{i,j} \\ (\sigma_x \sigma_y \rho_{xy})_{i,j} & (\sigma_y^2)_{i,i} \end{pmatrix} \quad (1 \leqslant i \leqslant n, \ 1 \leqslant j \leqslant n) \quad (53)$$

4.2　NMC 方法

NMC 方法是由 Parrish 和 Derber[11] 提出, 它用同一时间不同预报时效的预报偏差计算背景误差协方差矩阵, 即假设这种偏差可以代表预报误差。

$$\boldsymbol{X}_{12}^{\text{NMCf}}(t_a) = \boldsymbol{X}_{24}^f(t_a) - \boldsymbol{X}_{12}^f(t_a) = \boldsymbol{M}_{t_a, t_a-24}[\boldsymbol{X}^a(t_a - 24)] - \boldsymbol{M}_{t_a, t_a-12}[\boldsymbol{X}^a(t_a - 12)] \quad (54)$$

式中 $\boldsymbol{X}_{24}^f(t_a)$, $\boldsymbol{X}_{12}^f(t_a)$ 分别代表同一分析时刻 t_a 的 24 和 12 h 预报, $\boldsymbol{M}$ 为预报模式, NMC 方法假设预

报误差和预报偏差具有近似的结构,对其求月平均

$$\overline{\boldsymbol{X}}_{12}^{\mathrm{NMCf}}(t_a)=\frac{1}{I}\sum_{i=1}^{I}\boldsymbol{X}_{12}^{\mathrm{NMCf}}(t_a) \tag{55}$$

$$(\sigma^{\mathrm{f}})_l=(\overline{\boldsymbol{X}}_{12}^{\mathrm{NMCf}}(t_a))_l^2 \tag{56}$$

5 极小化算法

3D-Var 问题的求解可以归结为一个目标函数的极小化问题。根据最优化理论,可以通过调用某种迭代下降算法来达到解的最优,其基本思想是,给定一个初始点 $\boldsymbol{x}_0\in R^n$,按某种迭代规则产生一个点列 $\boldsymbol{x}^{(k)}$,使得当 $\boldsymbol{x}^{(k)}$ 是有穷点列时,其最后一个点是最优化模型问题的最优解。一个好的算法应具备的典型特征为:迭代点 $\boldsymbol{x}^{(k)}$ 能稳定地接近局部极小点 $\boldsymbol{x}^*$ 的邻域,然后迅速收敛于 $\boldsymbol{x}^*$,当给定的某种收敛准则满足时,迭代即终止。在最优化理论[11]中,常用的下降算法有共轭梯度法、牛顿法、拟牛顿法等。

牛顿法的基本思想是利用目标函数的二次 Taylor 展开,并将其极小化,牛顿法的突出优点是收敛很快,但该方法除了要计算目标函数的梯度,还要计算其二阶偏导矩阵(Hessian 矩阵)的逆,计算量大,Hessian 矩阵还可能出现奇异的情况。为了克服这一缺点,人们提出了拟牛顿法,基本思想是用目标函数值及其梯度的信息所得到的 $\boldsymbol{H}_k$ 代替 Hessian 矩阵的逆,再沿方向 $\boldsymbol{d}_k=-\boldsymbol{H}_k\boldsymbol{g}_k$ 作一维搜索($\boldsymbol{g}_k=\nabla\boldsymbol{J}(\boldsymbol{x}^{(k)})$)。由于构造 $\boldsymbol{H}_k$ 的方法不同,因而有不同的拟牛顿法。BFGS 算法被认为是迄今最好的拟牛顿法,它是由 Broyden, Fletcher, Goldfarb 和 Shanno 等在 1970 年给出的,LBFGS 算法是解决变量个数很大的无约束最优化问题的有效方法,二者仅有的区别是矩阵的更新。

由于变分同化问题中变量个数 N 通常是 10^7 或以上,本方案采用有限记忆的变尺度方法 LBFGS 进行极小化运算。

6 三维变分同化方案测试及理想试验

利用单个观测并假定背景场均匀分布,对三维变分同化系统进行测试。

6.1 水平递归滤波分析

为检验用递归滤波处理背景场误差协方差矩阵计算的效果,本文进行了一维和二维水平递归滤波试验。考察矩阵 $\sqrt{\boldsymbol{B}_{\mathrm{h}}}$ 和向量 $\boldsymbol{w}_{\mathrm{h}}$ 的乘法,即式(26),$\boldsymbol{w}_{\mathrm{h}}$ 的第 j 个分量是

$$w_{\mathrm{h}j}=\begin{cases}1.0 & i=11\\ 0.0 & \text{其他}\end{cases} \tag{57}$$

其中分析场格点数取 21 个格点,分两种情况:

(1) 矩阵 $\sqrt{\boldsymbol{B}_{\mathrm{h}}}$ 为高斯分布函数型

$$b_{i,j}=\varepsilon_{\mathrm{b}}^2\exp[-\frac{1}{2}(r_{i,j}/R)^2] \tag{58}$$

(2) 矩阵 $\sqrt{\boldsymbol{B}_{\mathrm{h}}}$ 为二阶自回归型相关函数型(SOAR)

$$b_{i,j}=\varepsilon_{\mathrm{b}}^2(1+|r_{i,j}|/R)\cdot\exp(-|r_{i,j}|/R) \tag{59}$$

$r_{i,j}$ 是格点 i,j 之间的距离,$\varepsilon_{\mathrm{b}}^2$ 是背景场误差标准差取为 1.0,水平特征尺度为 4 倍格距。图 1 中的黑实线为式(57)~(59)的计算结果,其余曲线为一维递归滤波计算结果,其中图 1a 的背景场误差协方差

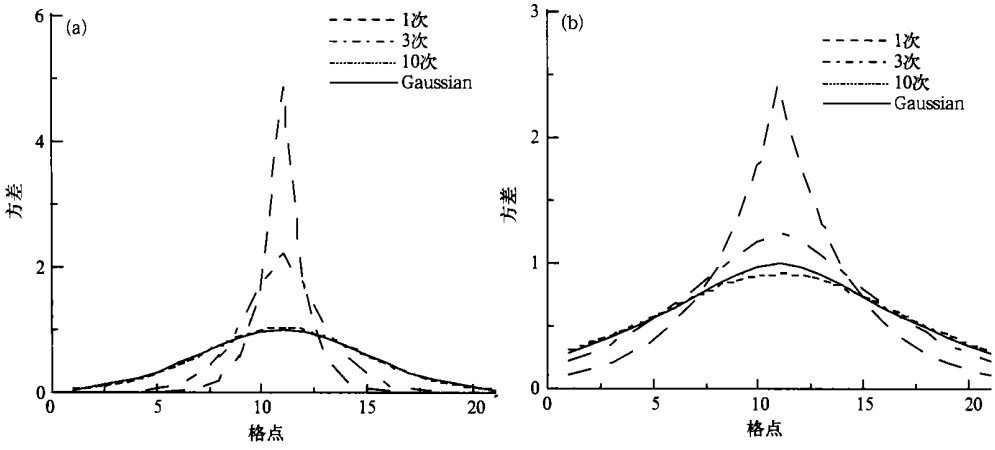


图 1 单点观测一维递归滤波

(a. Gaussian 分布; b. SOAR 分布)

Fig. 1 Single observation for one-dimensional recursive filter

(a. Gaussian; b. SOAR)

为高斯分布型(GAUSSIAN), 图 1b 的背景场误差协方差为二阶自回归型(SOAR)。由图 1a, b 可知, 无论是高斯分布型还是二阶自回归型, 随着滤波次数的增加, 用递归滤波方法计算的结果都可以很好地逼近原函数解析解, 高斯分布型需要 10 次滤波, 二阶自回归型只需要 3 次滤波即可。

图 2 为二维递归滤波计算结果, 其背景场误差协方差为高斯分布型, 其余与一维相同, 经过 10 次滤波, 其中心点值为 1, 自内向外逐步减小, 呈环状分布。

上述结果表明, 可以用空间滤波运算代替 $\sqrt{\boldsymbol{B}}_{\mathrm{h}}$, 这样矩阵 $\sqrt{\boldsymbol{B}}_{\mathrm{h}}$和 $\boldsymbol{w}_{\mathrm{h}}$ 的乘积相当于用一个高斯滤波器作用到 $\boldsymbol{w}_{\mathrm{h}}$ 上, 实际计算时应用的是一个截断的高斯滤波器或 SOAR 滤波器, 这样可以大大减少内存和计算量, 试验结果表明递归滤波方法的确可以很好地逼近原函数解析解。

6.2　垂直变换分析

在测试方案中, 垂直暂取为 15 层, 背景场误差

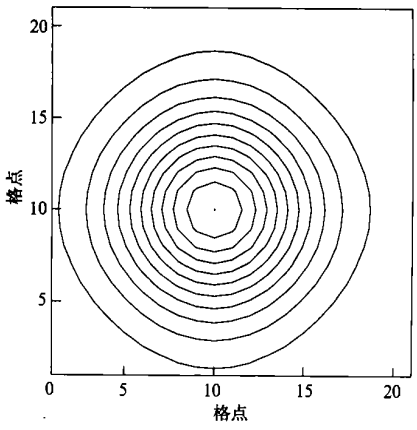


图 2　单点观测二维水平递归滤波, Gaussian 分布

Fig. 2　Single observation for two-dimensional recursive filter

方差暂取自国家气象中心的 HLAFS 模式系统的最优内插 OI 方案(表 1), 背景场误差相关系数假定为

$$r = \frac{1.0}{1.0 + k_p(\lg p_{l_i} - \lg p_{l_j})^2} \quad (i, j = 1, 2 \cdots 15) \tag{60}$$

表 1　各层位势高度均方差 σ_φ

Table 1　RMSE of geopotential height at each level

p_l(hPa)	1000	925	850	700	600	500	400	300
σ_φ(m)	10.0	11.0	12.0	12.0	12.0	12.0	15.0	19.0
p_l(hPa)	250	200	150	100	70	50	30	
σ_φ(m)	25.0	28.0	30.0	40.0	45.0	50.0	55.0	

其中, k_p 为分析变量数, p_{l_i}, p_{l_j} 分别为第 i 和 j 层大气压。流函数 Ψ 的方差通过下式计算

$$\sigma_\Psi^2 = \frac{\sigma_\varphi^2}{f_0^2} \tag{61}$$

若取前 $m(m < 15)$ 个主分量, 则其累积解释方差为

$$G(m) = \frac{\sum_{i=1}^{m} \lambda_i}{\sum_{i=1}^{15} \lambda_i} \tag{62}$$

其中, λ_i 为第 i 个特征值。

从表 2 可以看出, 当 $m \geqslant 5$ 时, 其累积方差贡献百分率已超过 90%。关于 m 的确定, 从理论上来说应考虑两方面的因素, 其一是累积解释方差; 其二是必须考虑各个垂直模态对选入的前 m 个主分量的贡献。可以看出, 当 $m = 11$ 时应有很好的代表性。

6.3　单点观测的三维变分试验分析

6.3.1　多变量分析

水平分析区域取 21 × 21 个格点, 背景场值为 0, 在分析区域中心设一单个观测, 进行如下 3 个试验(表 3)。

图 3 是单点位势高度(试验 1) 分析结果, 图 3a 是位势高度场, 呈圆环状分布, 自内向外逐步减小。图 3b 是纬向风速 u 场, 通过地转平衡关系, 形成区

表 2　累积解释方差

Table 2　Cumulative relative RMSE

m	4	5	6	7	8	9	10	11	12	13	14	15
$G(m)$	86.1	91.1	94.7	97.1	98.5	99.3	99.7	99.9	99.9	99.9	99.9	100

表 3 试验方案

Table 3 Experimental scheme

	φ	u	v
试验 1	100.0	0	0
试验 2	0	1.0	0
试验 3	0	0	1.0

域上半部为西风, 下半部为东风。图 3c 是经向风速 v 场, 同样通过地转平衡关系, 形成区域左半部为南风, 右半部为北风。

图 4 是单点东西风(试验 2) 分析结果, 图 4a 是位势高度场, 通过地转平衡关系, 形成区域上半部为负, 下半部为正的两个扁圆。图 4b 是纬向风速 u 场, 其中部为西风, 上、下两边为东风。图 4c 是经向风速 v 场, 分析区域 1, 3 象限为南风, 2, 4 象限为北风。

图 5 是单点南北风(试验 3) 分析结果, 图 5a 是位势高度场, 通过地转平衡关系, 形成区域左半部为负, 右半部为正的两个扁圆。图 5b 是纬向风速 u 场, 分析区域 1, 3 象限为西风, 2, 4 象限为风。图 5c 是经向风速 v 场, 其中部为南风, 左、右两侧为北风。

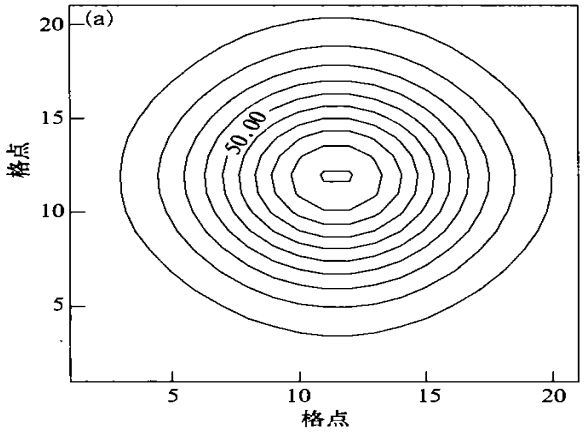
由图 3, 4, 5 可以看出, 三维变分分析结果与经典的位势高度场与风场的背景场误差水平相关模型(即地转平衡关系) 的形状完全一致, 这反映了该三维变分分析系统能正确地反映多变量之间的相互作用关系。

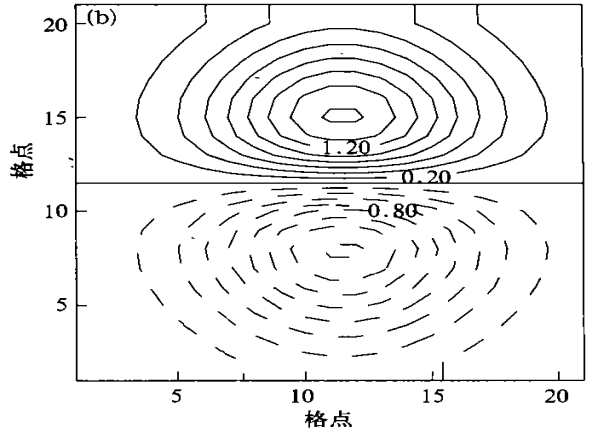
6.3.2 极小化过程目标函数的收敛分析

试验中迭代 10 次, 背景场取为 0, 单站观测值在垂直方向为正弦曲线分布, 即

$$y_{\mathrm{Obs}}(k) = 10\sin\left[\frac{2(k-1)\pi}{k_{\max}-1}\right] \qquad (63)$$

分析图 6 结果, 经过 10 次迭代后, 分析场从背景场迅速向观测值收敛, 二者非常接近。图 7, 8 分别为目标函数和目标函数梯度随迭代步数的下降曲线, 可以看到目标函数经过 10 次迭代后, 其值迅速下降为原来的 6.2%; 与目标函数相似, 其梯度值也下降为原来的 7.3%。





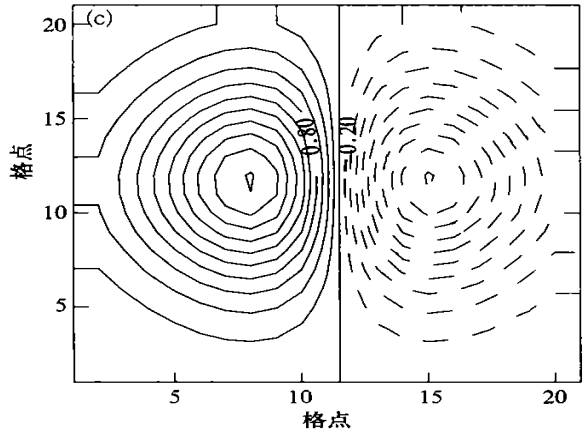


图 3 单点位势高度 φ 观测多变量分析结果

(a. 位势高度场 φ, b. 东西风 u, c. 南北风 v)

Fig. 3 A multivariate correlation model on the singleobservation of geopotential height

(a. geopotential height φ, b. wind component u, c. wind component v)

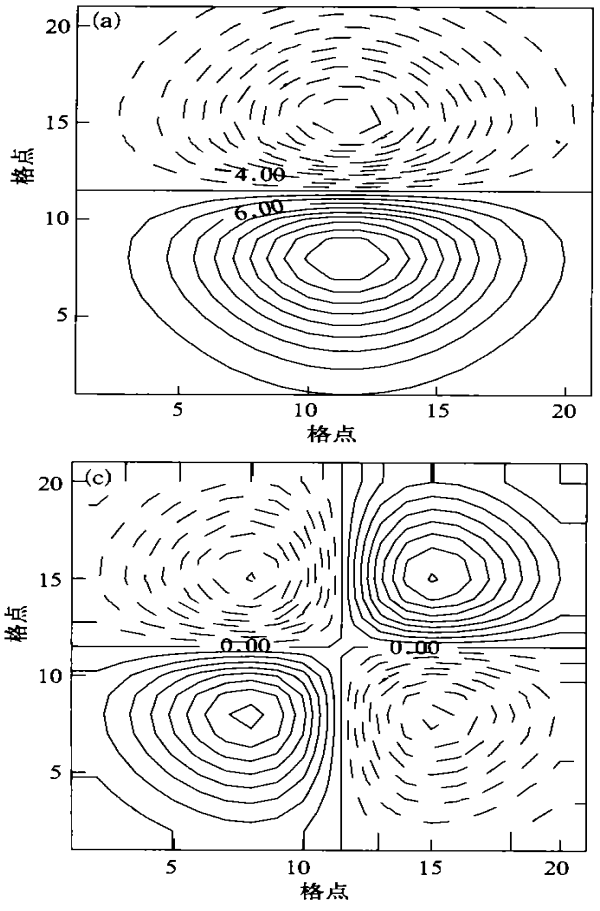


图 4　单点东西风 u 观测多变量分析结果
(a. 位势高度场 φ, b. 东西风 u, c. 南北风 v)

Fig. 4　A multivariate correlation model on the singleobservation of wind component u
(a. geopotential height, b. wind component u, c. wind component v)

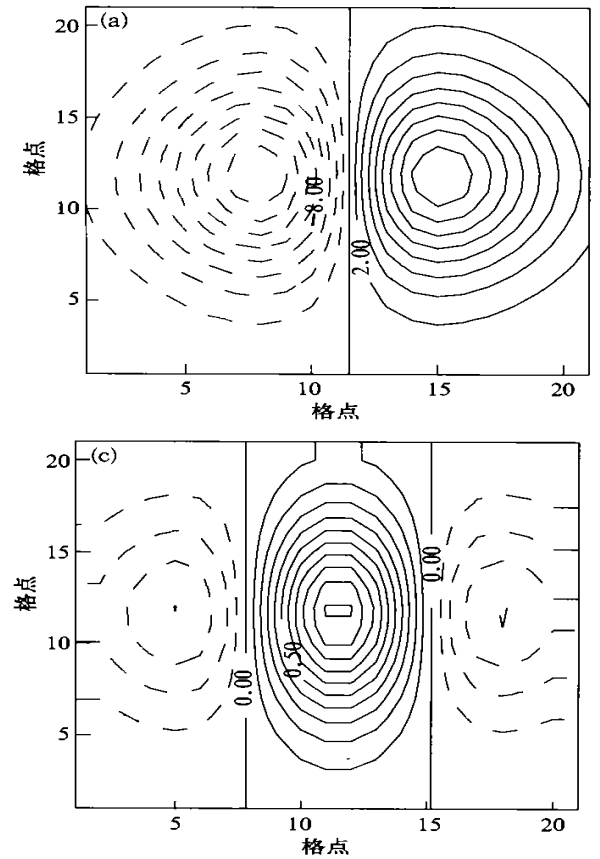


图 5　单点南北风 v 观测多变量分析结果
(a. 位势高度场 φ, b. 东西风 u, c. 南北风 v)

Fig. 5　A multivariate correlation model on the singleobservation of wind component v
(a. geopotential height, b. wind component u, c. wind component v)

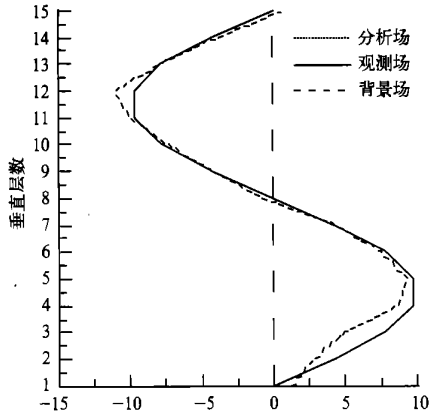


图 6 单点观测变分分析结果
Fig. 6 Variational assimilation on single observation

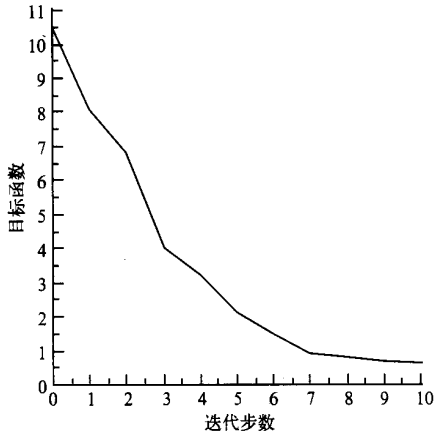


图 7 目标函数随迭代步数的下降曲线
Fig. 7 The cost-function in minimizing

7 小 结

本文给出一种适用于格点模式的三维变分分析方案,其主要特点:

(1) 目标函数采用增量形式;

(2) 分析变量取为流函数 Ψ,非平衡速度势函数 χ_u,非平衡位势 Φ_u 和比湿 q(或相对湿度),考虑了质量场和风场之间的平衡约束关系;

(3) 通过变量变换进行预调节(Preconditioning),明显地降低了目标函数 Hessian 矩阵的条件数,改善了收敛性,避免了直接计算背景误差协方差矩阵的逆 $\boldsymbol{B}^{-1}$ 的困难;

(4) 下降算法调用有限内存的 BFGS 方法(LBFGS);

(5) 理想试验结果表明,该三维变分同化系统能够正确地反映多变量之间相互作用关系,收敛迅速,分析结果合理。

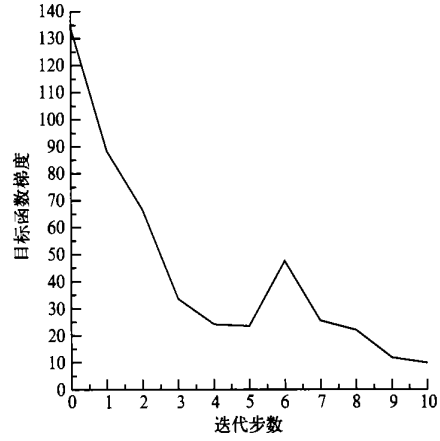


图 8 目标函数梯度随迭代步数的下降曲线
Fig. 8 The gradient of the cost-function in minimizing

上述结果是非常初步的,该三维变分同化系统还需要进一步的改进和完善。

参考文献

1 潘宁,董超华,张文建. ATOVS 辐射率资料的直接变分同化试验研究. 气象学报, 2003, 61(2): 226~ 235

2 翁永辉,徐祥德,柏晶瑜. 应用 TOVS 资料变分分析技术增加青藏高原地区模式初始场信息. 气象学报, 2000, 58(6): 679~ 691

3 Parrish F D, Derber C J. The national meteorological center's spectral-interpolation analysis system. Mon Wea Rev, 1992, 120: 1747~ 1763.

4 Courtier P, etal. The ECMWF implementation of three-dimensional variational assimilation (3D- Var). I: Formulation. Quart J Roy Meteor Soc, 1996, 124: 1783~ 1807.

5 Lorenc A C. Development of an operational variational assimilation scheme. J Meteor Soc Japan, 1997, 75(1B): 339~ 346

6 Barker D M. A General Formulation for 3DVAR Control Variables. 1999

7 Courtier P, Thepaut J- N, Hollingsworth A. A strategy for operational implementation of 4D- Var using an incremental approach. Quart J Roy Meteor Soc, 1994, 120: 1367~ 1387.

8 Courtier P. Variational Methods. J Meteor Soc, 1997, 75(1B): 211~ 218.

9 Lorenc A C. Iterative analysis using covariance functions and filters. Quart J Roy Meteor Soc, 1992, 118: 569~ 591.

10　Hayden C M, Purser R J. Recursive filter objective analysis of meteorological fields: Application to NESDIS operational processing. J Appl Meteor, 1995, 34, 3~ 15

11　陈宝林. 最优化理论与算法. 北京: 清华大学出版社, 1989. 535pp

IDEA EXPERIMENTS OF GRAPES THREE- DIMENSIONAL VARIATIONAL DATA ASSIMILATION SYSTEM

Hua Zhang　Jisan Xue　Shiyu Zhuang　Guofu Zhu　Zhu Zhongchen

(Chinese Academy of Meteorological Sciences, Beijing 100081*)*

Abstract

This paper introduces a new scheme of three- dimensional variational assimilation (GRAPeS- 3DVAR), which is suitable for a grid-point model. The domain includes China area on horizontal dimension and 15 mandatory levels on vertical dimension. The scheme adopts stream function, unbalanced potential velocity, unbalanced geopotential height and humidity as analysis variables which are independent of each other. The cost function takes the way of incremental method. To solve the problems that $\boldsymbol{B}^{-1}$ are ill-conditioned and too large to be computed, the background term is preconditioned using a sequence of variable transform, which improves the convergence in the minimization process and avoids to calculate $\boldsymbol{B}^{-1}$ directly. EOF decomposition is used as vertical component of analysis variable transform to project onto vertical modes and separate each three-dimensional field into two-dimensional field. Recursive filters is used as horizontal component of analysis variable transform. LBFGS method is used on solution of optimization. Simple geostrophic balanced relationship is used. The system has passed verifications strictly and proved to be credibility. In order to test the functions of the system, A series of idea experiments were made. The results show a reasonable relationship among the multi-variables and a fast convergence in the minimized process.

Key words: Three-dimensional Variational assimilation, Variable transform, Preconditioning.