

区域四维变分资料同化的数值试验^{*}

龚建东 邱崇践

(兰州大学大气科学系, 兰州, 730000)

王 强 陈伟民

(兰州中心气象台, 兰州, 730000)

摘 要

针对中尺度数值预报模式预报误差的主要来源, 尝试利用四维变分资料同化的方法来改善预报效果。在已建立的中尺度模式(MM4)四维变分资料同化系统基础上, 进行了若干数值试验, 通过比较同化前后的预报来检验同化的效果。这些试验中初始场、模式误差和侧边界条件被分别或同时作为控制变量来进行调整, 主要探讨了模式误差和侧边界条件对同化及预报的影响, 以及同时结合两者或三者的途径和方法。对两组个例分别进行的试验结果表明, 区域中尺度模式预报误差除了来源于初始误差外, 模式误差、侧边界条件也有不可忽视的作用。同化时应同时考虑初始场、模式误差和侧边界条件这三方面的共同作用, 仅修正其中某一个或某两个会把由于其它方面造成的预报误差转嫁到它们之上, 从而出现尽管目标函数下降很快而预报结果并没有相应改善的现象。

关键词: 区域四维变分资料同化, 误差来源, 同时同化, 数值试验。

1 引 言

四维变分资料同化方法将同化问题提为一个以动力模式为约束的极小化问题, 这里目标函数定义为预报与观测的距离。因此四维变分资料同化过程也可看作是将模式输出与观测相拟合。

顾震潮早在 1958 就提出应在数值预报中考虑历史演变信息的观点^[1], 并由丑纪范将利用多时刻资料信息等价的提为求解微分方程的反问题, 同时引入广义解的概念^[2]。随后的工作^[3~5]进行了一系列数值试验, 取得一些有意义的结果。不过, 这些试验采用的模式大都比较简单, 观测资料多由模式积分获得, 并主要讨论了初始误差对同化、预报效果的影响, 而对区域中尺度模式预报效果影响非常大的模式误差、边界条件却没有仔细研究过。实际区域中尺度模式预报误差同时来源于初始误差、模式误差和边界条件等多个方面, 单独考虑其中的某一方面并不完善。因此, 采用复杂的接近业务预报水平的预报模式和实际观测资料, 同时考虑初值、模式误差和边界进行四维变分资料同化时可能产生的问

* 初稿时间: 1997 年 7 月 31 日; 修改稿时间: 1998 年 3 月 10 日。

资助课题: 国家自然科学基金和甘肃自然科学基金资助项目。

题我们并不很清楚。本文以 Errico 等给出的中尺度模式(MM4)的伴随模式为基础,建立了一个四维变分资料同化系统,进行同化和预报试验。在试验中分别考虑了初值、模式误差和侧边界对同化和预报的影响,以及同时考虑这三方面误差来改善预报效果的可能性和途径。

2 理论概述

2.1 四维变分资料同化方法

一般的区域数值预报模式的控制方程经空间离散化后可以写为下面的常微分方程组:

$$\frac{d\mathbf{X}(t_{n+1})}{dt} = \mathbf{F}(\mathbf{X}(t_n), \mathbf{P}) \quad (1)$$

在 t_0 时刻有: $\mathbf{X}(t_0) = \mathbf{X}_0$

在侧边界上满足: $\mathbf{X}(t_n) = \mathbf{B}(t_n)$

四维变分资料同化方法的基本原理是通过修改模式的输入量(以下称为控制变量)使在同化时间区间内的模式输出量(或它们的某些函数)与相应的观测之间距离最小。为此,定义表征该距离的目标函数为:

$$J = \frac{1}{2} \sum_{r=1}^R [\mathbf{X}(t_r) - \mathbf{X}^{\text{obs}}(t_r)]^T \mathbf{W}(t_r) [\mathbf{X}(t_r) - \mathbf{X}^{\text{obs}}(t_r)] \quad (2)$$

问题成为在约束式(1)下求 J 的极小值。这里 \mathbf{P} 是表征模式误差的参数向量, $\mathbf{X}(t_r)$ 是 t_r 时刻的 N 维状态向量, R 是在同化时间区间内有观测的时间次数, $\mathbf{X}^{\text{obs}}(t_r)$ 是对应的观测值,上标 T 表示矩阵转置。

有三种选择控制变量的方案。第一种方案是以初值 \mathbf{X}_0 作为控制变量,这相当于认为模式精确。第二种方案以参数向量 \mathbf{P} (或侧边界条件 \mathbf{B}) 作为控制变量,这相当于认为初值精确,而模式有误差。第三种方案是将初值 \mathbf{X}_0 和 \mathbf{P} , \mathbf{B} 同时作为控制变量,亦即认为初值和模式都有误差。

实际上模式误差的显式形式无法获得。一般,可在模式中加入一个误差修正项以抵消模式误差的影响^[6],表示成 \mathbf{P} 与时间函数乘积的形式。相应的,预报模式可以写为:

$$\frac{d\mathbf{X}(t_{n+1})}{dt} = \mathbf{F}(\mathbf{X}(t_n)) + \lambda(t_n)\mathbf{P} \quad (3)$$

其中 $\lambda(t)$ 是预先给定的时间演变函数。

可以给出目标函数 J 对初值、模式误差修正项的梯度分别为:

$$\dot{\mathbf{X}}_{x_0} J = \delta \hat{\mathbf{X}}(t_0) \quad (4)$$

$$\dot{\mathbf{X}}_{\mathbf{P}} J = \sum_{n=1}^N \lambda(t_n) \delta \hat{\mathbf{X}}(t_n) \quad (5)$$

其中 N 为预报模式的积分总步数, $\delta \hat{\mathbf{X}}(\)$ 表示伴随变量。对参数向量 \mathbf{B} 的梯度跟具体的侧边界方案有关。

2.2 权重系数

权重系数 $W(t_r)$ 主要用来度量各个预报变量的相对重要性, 是一个 $N \times N$ 维的对角矩阵, 对角元素表示各个预报变量的权重系数:

$$W = \text{diag}(W_U, W_V, W_T, W_{PS}, W_{TC}) \quad (6)$$

其中 $W_U, W_V, W_T, W_{PS}, W_{TC}$ 是子对角阵, 分别表示风场、温度场、地面气压场和地表温度。试验中权重系数阵的元素取成如下形式:

$$w_{\Omega} = \frac{1}{4 \sum_{r=1}^R \left(\Omega_{j,k} - \Omega_{j,k}^{\text{obs}} \right)^2} \quad (7)$$

式中 Ω 表示某一变量, i, j 表示水平方向格点数, k 表示垂直层次。

2.3 尺度化过程

在用下降算法求最优解时, 收敛速度与梯度的特点有关。当各个预报变量的梯度值相近时, 目标函数能迅速收敛。尺度化过程的目的在于通过归一化的办法使各个预报变量的梯度值相近。尺度化过程可以表述为:

$$\tilde{\Omega} = \Omega \tilde{\Omega}^{-1} \quad g_{\tilde{\Omega}}^s = S g_{\Omega} \quad (8)$$

对控制变量的梯度, g_{Ω} 为目标函数, 上标 s 表示经尺度化后的量, S 为尺度化对角矩阵。试验中 S 的子对角元素取成如下形式

$$s_{\Omega} = \max \left| \Omega_{j,k}^i - \Omega_{j,k}^{\text{obs}} \right| \quad (9)$$

3 四维变分资料同化系统和资料

区域中尺度模式(MM4)四维变分资料同化系统由非线性预报模式、伴随模式和下降算法等组成, 不考虑水汽作用, 下边界采用实际的地型高度和地表分类。试验中水平格距 40 km, 31 × 46 个网格点, 侧边界取 5 排“海绵”边界, 垂直分成等 σ 面 10 层, 顶部气压为 100 hPa。

试验选取了两组资料序列, 分别检验同化时间区间的长短对同化的影响。第一组资料序列从 1993 年 5 月 4 日 00 时到 5 日 12 时共 4 个时次客观分析场资料, 资料间隔 12 h。试验将 4 日 12 时作为“当前时刻”, 4 日 00 时资料作为前 12 h 预报初猜值。4 日 00 时与 4 日 12 时两个时次资料同时用于同化时间区间内的已知观测值, 试验将 5 日 00 时与 5 日 12 时作为实况来与预报作对比。另一组资料序列从 1997 年 10 月 11 日 00 时到 16 日 00 时共 6 个时次客观分析场资料, 资料间隔 24 h。试验以每 4 个连续时次资料做一次同化试验, 前两个时次资料用于同化试验中的已知观测值, 后两个时次资料作为未来实况, 同化时间区间 24 h, 预报时段 48 h。一共 3 组试验。

4 数值试验设计

针对不同的控制变量选择方案, 设计了以下几种同化试验方法:

(1) IC 试验: 仅对初始场进行调整。同通常四维变分资料同化作法。

(2) ME 试验: 仅对模式误差修正项进行调整。并对时间演变函数 $\lambda(t)$ 给出三种性质相对的方法:

ME-1: $\lambda(t)$ 在同化时间区间内为常数, 即振幅随时间无变化。

ME-2: $\lambda(t)$ 在同化时间区间初端为零, 随积分时间线性增加, 并在末端达到最大。

ME-3: 与 ME-2 相反。

在用 ME 方法同化后继续预报时, 采用了两种不同的方法, 一是在继续预报时不再考虑模式误差修正项。另一种是仍以适当的方式考虑模式误差修正项。前一种方法意味着在同化过程中反演出的误差修正项在随后的预报中没有意义。后一种方法则是认为误差修正项有一定的持续性。

(3) IC+ ME 试验: 对初始场、模式误差修正项二者进行调整。考虑三种方式: 先对初始场进行调整、先对模式误差修正项进行调整、同时对初始场和模式误差修正项进行调整。

(4) BDY 试验: 主要对侧边界进行调整。

(5) ALL 试验: 同时对初始场、模式误差修正项和侧边界进行调整。

5 目标函数与梯度的收敛特点

5.1 IC 同化试验

初始误差主要源于观测误差和资料的客观分析所产生的误差。当前的客观分析场大多由最优差值给出, 存在一定的缺陷。表现为由这种方法获得的初始场往往与动力预报模式不相协调, 从而对预报结果造成很坏的影响。IC 同化试验考虑到初始误差并对其进行调整。试验中同化过程迭代 10 步, 图 1 目标函数和梯度为随迭代步数的下降情况。

IC 试验中目标函数的下降较为平缓, 迭代 10 步后仅下降了 17%。而且主要是前几步

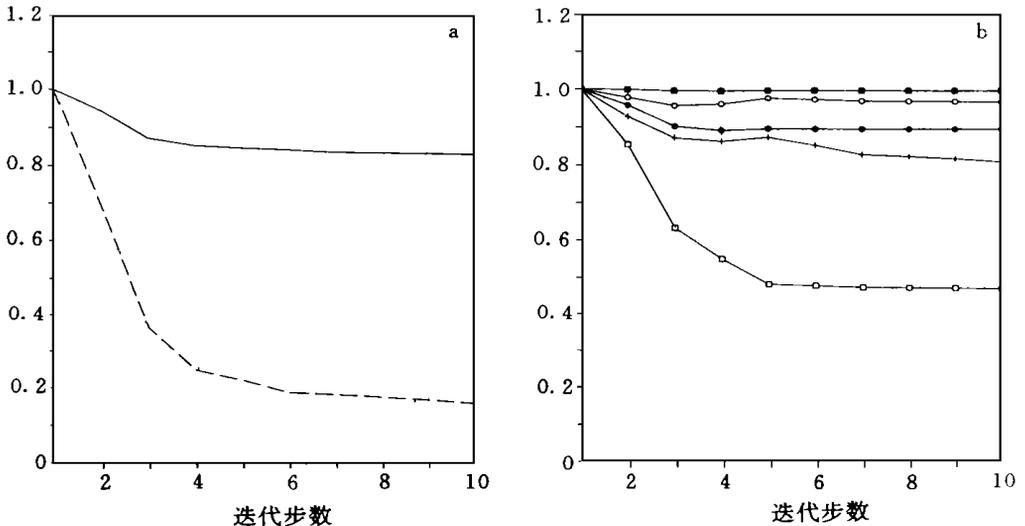


图 1 目标函数、梯度和各预报变量目标函数随迭代步数的下降
(实线为目标函数, 虚线为梯度, 实心方框线为地面温度, 空心圆线为 v 场,
实心圆线为温度场, 十字叉线为 u 场, 空心方框线为地面气压)

下降明显, 迭代 3, 4 步后目标函数基本上不下降。在各个预报变量中, 气压场下降最快, 有 53%, 地温场下降不明显。在同化区间末端时刻, 同化后初始场的预报与未经同化的预报对比, 可以发现这两个预报场之间没有明显差别, 表明 IC 试验中同化效果并不显著。

5.2 ME 同化试验

模式误差来自数值计算方法、次网格尺度过程参数化、物理过程参数化等多个方面。一般, 并不可能完全清楚模式误差的来源, 试验通过引入模式误差修正项的办法对模式误差进行抵偿, 以减弱模式误差的影响。其中的时间演变函数表示了模式误差随积分时间振幅的变化, 无法预知^[6]。试验分别给出它取 ME-1, ME-2, ME-3 这 3 种不同方式时相应的目标函数下降情况。

时间演变函数取 3 种不同方式时目标函数的下降有差别, 其中以 ME-2 下降最为显著, 下降了 90%, ME-3 最不好, 只有 58%。ME-1 居中, 为 80%。这说明模式误差随积分时间的增加而逐步的显著起来。ME-1, ME-2 两种方式没有显著差别, 为了便于在随后的试验中继续考虑模式误差修正项的作用, 时间演变函数取 ME-1 振幅随时间无变化形式。图 3 为 ME 试验迭代 10 步时目标函数和梯度随迭代步数的下降情况。

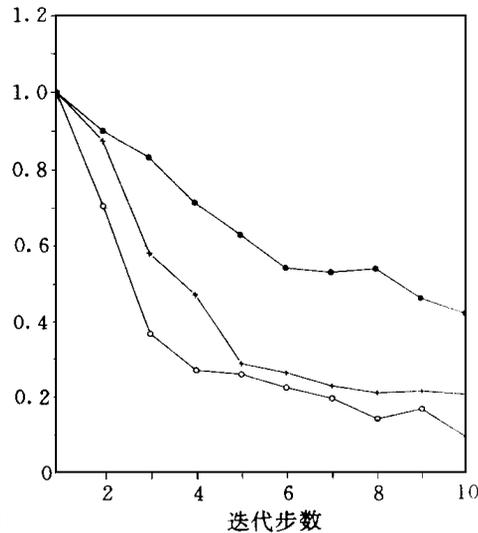


图 2 时间演变函数不同时目标函数随迭代步数的下降 (十字为 ME-1, 空心圆为 ME-2, 实心圆为 ME-3)

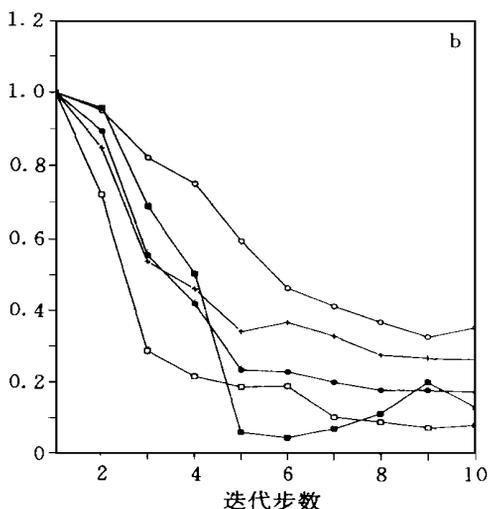
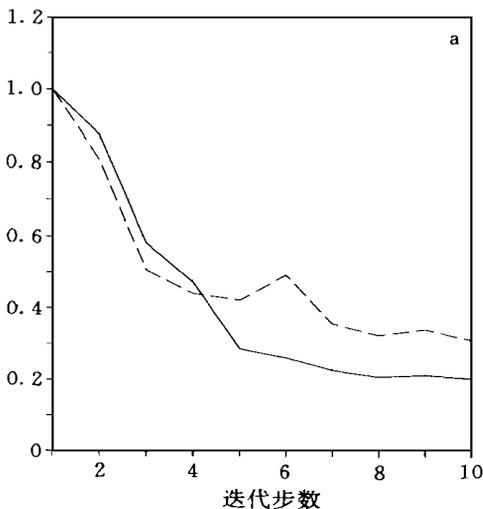


图 3 目标函数、梯度和各预报变量目标函数随迭代步数的下降

(图中关于各种曲线的说明同图 1)

相对于同化初始场而言,同化模式误差目标函数收敛迅速,迭代 10 步目标函数下降达 78%,而且各个场都有不同程度的下降,其中下降最快的是气压场和地温场,达到 90%,最慢的经向风场也下降了 65%。

5.3 IC+ ME 同化试验

IC+ ME 试验同时考虑了初值、模式误差对预报的影响。试验分成 3 种情况,即先迭代 5 次同化初猜值,紧接着迭代 5 次同化模式误差的 ICME 试验;先迭代 5 次同化模式误差,然后迭代 5 次同化初猜值的 MEIC 试验,以及迭代 10 次同时同化初猜值和模式误差的 IMEC 试验。这 3 种试验的目的在于对比不同的结合方式以找到一个最佳的方式。图 4 给出了 3 种试验的目标函数的下降图,作为对比,同时给出了 IC 和 ME 试验的结果。

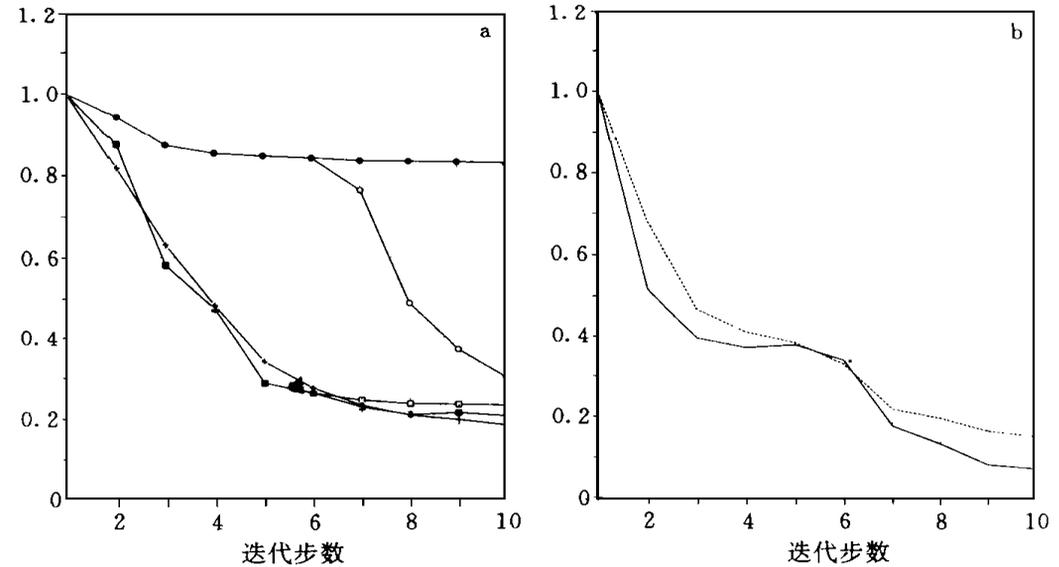


图 4 目标函数和 IMEC 试验中初值、模式误差的梯度随迭代步数的下降

(实心圆线为 IC 实验,空心圆线为 ICME 实验,

实心方框线为 ME 实验,空心方框线为 MEIC 实验,

十字叉线为 IMEC 实验;实线为初值的梯度,虚线为模式误差的梯度)

3 种方式稍有差别,以 IMEC 试验的下降效果最好,ICME 试验的下降效果最差。IMEC 试验的结果和 ME 试验的结果接近,IMEC 试验中各个预报变量目标函数的下降情况与 ME 试验中各个预报变量的收敛也相似,但经、纬向风场收敛要好于 ME 试验。而且 IMEC 试验对初值和模式误差均有较好的收敛性。

5.4 侧边界影响的诊断

试验以常用的“海绵”侧边界为例讨论侧边界条件对区域模式预报的影响。“海绵”侧边界条件应在预报前给出,一般多用连续两个时次的侧边界资料线性内插出两时次之间的边界值。首先,对模式的初、边值做敏感性试验。图 5 给出目标函数对初、边值的梯度的合成图。

图 5 示出, 西侧边界处目标函数对边值的梯度远大于对初值的梯度。可见侧边界条件, 特别是西侧边界条件, 对预报有不可忽视的影响, 在同化试验中应当考虑侧边界条件的作用。

5.5 BDY 同化试验

试验仅考虑侧边界对同化的影响, 不考虑初值、模式误差对预报的影响。这时有两个初猜值, 分别对应同化时间区间初、末时次的侧边界值, 需同时调用两次下降算法进行计算。这与 IC 试验、ME 试验不同。迭代 10 次后目标函数下降了 58% (图 6), 好于同化初值情况 (26%), 但比同化模式误差要差一些 (78%)。

6 同化后的预报

6.1 初值和模式误差同化后的预报效果

尽管同化后在同化时间区间末端的预报误差减小了, 一个更切实际的问题是, 在随后的预报中同化后的预报场是否优于同化前的预报场, 预报误差能不能相应的减小, 如果没有, 同化过程就没有实际应用的意义。这可通过比较同化前后的预报均方根误差来说明。图 7 示出试验结果。IC 试验的预报效果与未做同化试验时 4 日 00 时的预报结果接近。ME 试验的结果表明, 在预报模式中加入模式误差修正项的确可以减小同化时间区间内的预报误差, 期望这一修正项可以继续持续到以后的预报中。为此做了一个对比试验, 一是在同化后的预报中不再考虑模式误差修正项的作用, 另一是继续考虑模式误差修正项, 它的时间项仍取振幅随时间无变化这一种形式。

试验结果表明, 继续考虑模式误差修正项能减小经、纬向风场和温度场的预报误差, 并有一定的持续作用。但对地面气压和地温来说, 没有起到改善作用。在以后的试验中将继续考虑对风场和温度场的模式误差修正项的作用。试验给出同化初值、同化模式误差、和同时同化初值模式误差这 3 种情况下随后预报的均方根误差。

IC+ ME 试验明显改善了预报效果, 其中经向风场和温度场的改善最明显, IMEC 试验效果最好。ICME 试验的预报效果与同化前的预报效果相当, 而 MEIC 试验预报效果同

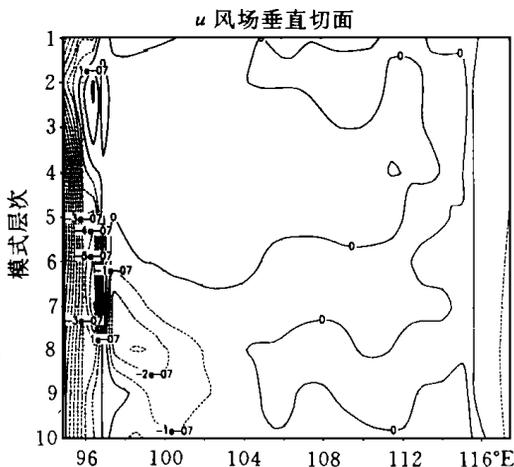


图 5 经过模式区域中心东西方向梯度的垂直剖面

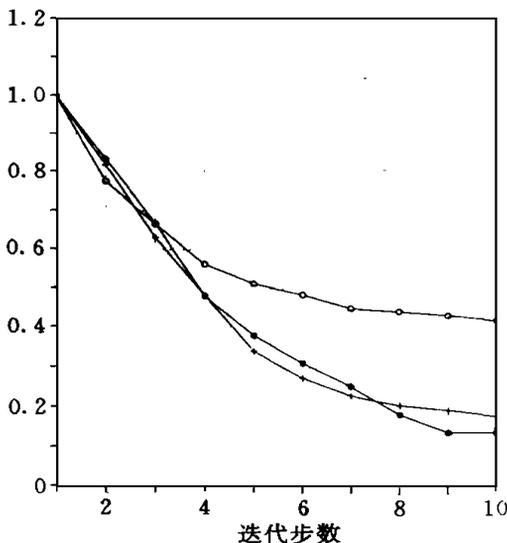


图 6 目标函数随迭代步数下降
(空心圆线为 BDY, 十字叉线为 IMEC, 实心圆线为 ALL)

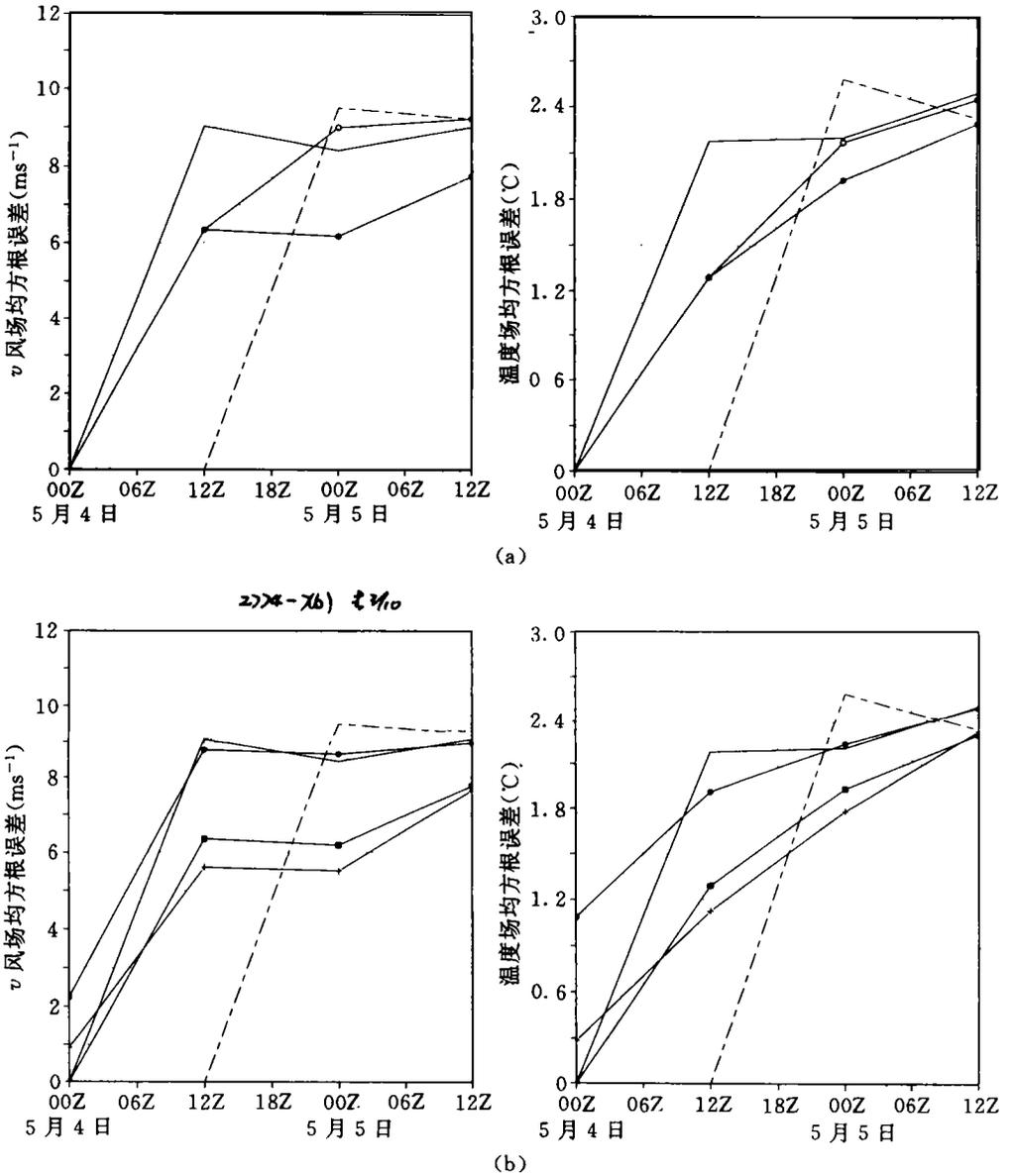


图7 不考虑(空心圆)和继续考虑误差修正项(实心圆)时预报的均方根误差和IC(图b中实心圆)、ME(实心方框)、IMEC(十字叉)试验的均方根误差,实线和虚线为变分同化试验前的原有预报结果

IMEC 试验接近。

分析ME 试验在同化时段末端的预报结果(图 8a),在西、南侧边界处出现涡度、散度的极值中心,这和观测相差甚远。其原因是目标函数对模式误差梯度中的一部分源于模式误差本身,另一部分则来源于侧边界和初值。在仅选择模式误差作为控制变量时,模式并不能区分,而是全部当作模式误差来处理,结果导致西侧边界处出现扭曲。因而有必要同

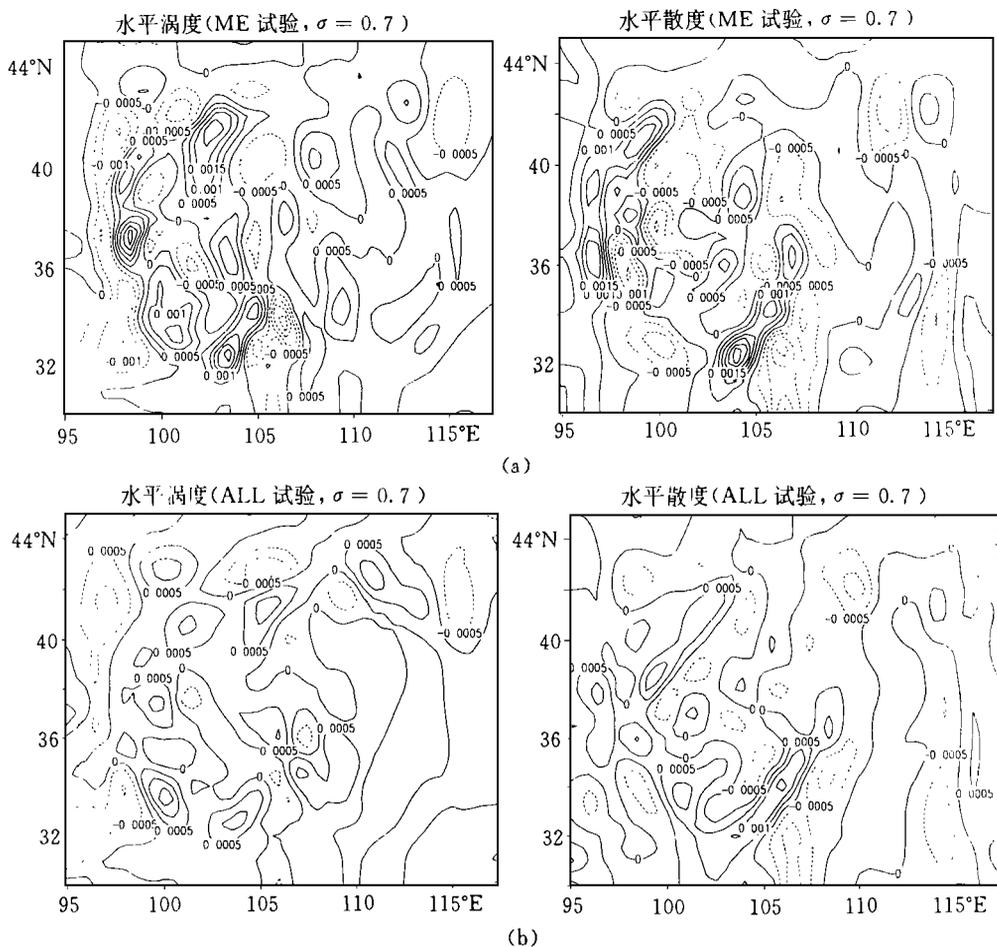


图 8 ME 和 ALL 试验预报水平涡度和散度

(a. ME 试验; b. ALL 试验。涡度和散度单位: s^{-1})

时考虑初值、模式误差和侧边界条件的共同作用。

6.2 侧边界同化后的预报效果

对侧边界同化后进行预报时发现, 尽管目标函数下降了 58%, 但预报不仅没有相应的改善, 反而比同化前的预报差(图 9)。这和两个方面的原因有关, 一方面侧边界条件是通过前后两个时次的侧边界线性内插得出, 同化时段内的结果不能对随后的侧边界有所改善。另一方面跟 ME 试验类似, 同化时把本是由初值或模式误差造成的预报误差转移到侧边界上, 出现目标函数收敛快而预报差的现象。

7 ALL 同化、预报试验

同时考虑初值、边值和模式误差这三个方面的误差时需要同时调用四次下降算法。图 6 给出了试验结果。同时考虑了初值、边值和模式误差的同化效果在所有试验中效果最

好,目标函数下降了 87%。在随后的预报中,预报效果也有一定程度上的改善(图 9)尤其是纬向风改善明显,但经向风不如 IMEC 试验好。在同化时段末端对比 ALL 试验与 ME 试验的预报结果,ALL 试验西、南侧边界处的涡度、散度明显减小,与观测非常接近,而且正、负涡度配置也与观测一致(图 8b)。可见同时同化在一定程度上克服了误差的转移。

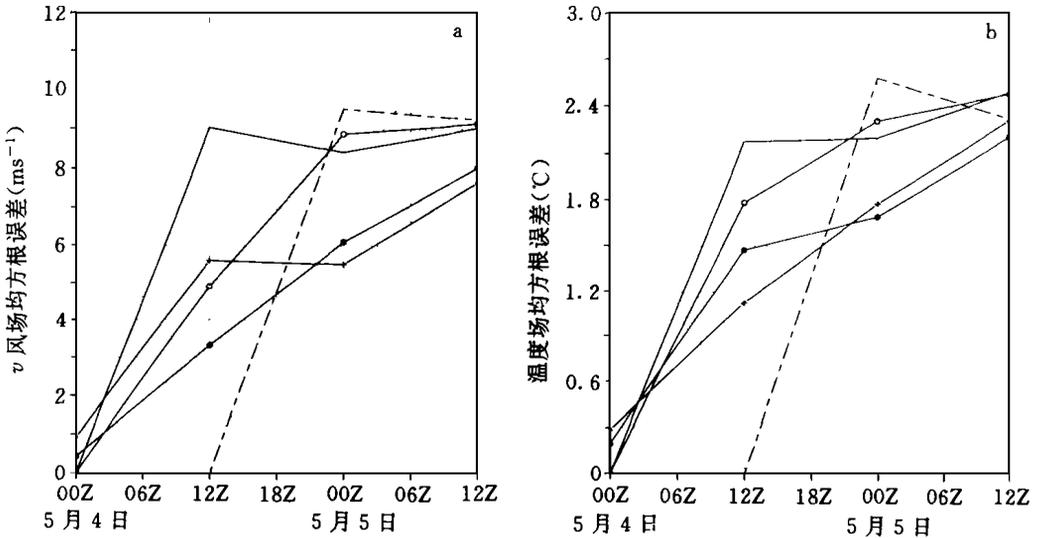


图 9 同化前后的均方差

(空心圆线为 BDY, 十字叉线为 IMEC, 实心圆线为 ALL)

当同化时间区间延长至 24 h, 表 1 给出用资料二进行 IC, ME, IMEC, ALL 试验时同化结果。相对同化时间区间 12 h 而言,目标函数收敛明显减慢。

表 1 IC, ME, IMEC, ALL 试验中目标函数下降百分比

试验序号	试验方法和下降百分比			
	IC 试验	ME 试验	IMEC 试验	ALL 试验
试验 1	14%	56%	59%	61%
试验 2	13%	47%	50%	60%
试验 3	14%	50%	55%	62%

IC 试验对预报改善不明显, IMEC 试验在同化时段末端均方差最小, 在 24 h, 48 h 预报时与不做同化的情况接近。对经、纬向风场、温度场的预报效果要显著好于气压和地温的预报。ALL 试验在 24 h, 48 h 经向风场预报也好于其它试验。这与同化时间区间 12 h 的预报结果没有明显差别(图略)。

8 结果与讨论

本文利用四维变分资料同化方法对区域中尺度模式进行了一系列数值试验, 试验中初值、边值和模式误差被分别或同时作为控制变量来进行调整, 并检验了同化和预报的效

果。由于实际资料独立于动力模式, 目标函数收敛较为缓慢。试验中同化时间区间内的边值和模式误差对预报误差的贡献要比初值明显, 体现在目标函数的收敛速度和预报效果改善上。试验区域中侧边界所占格点与整个区域所占格点之比接近 $1/2$, 而一般中尺度侧边界条件问题解决的还不能令人满意, 因而, 试验中初值误差对预报误差的贡献相对于侧边界条件误差而言显得不那么显著。这一试验结果与区域中尺度模式的预报事实一致, 即通过放大模式区域的办法减小侧边界条件对区域中心的影响。但正如同化试验所揭示的那样, 侧边界条件的误差主要集中于区域的西侧边界处, 这一区域主要是西风平流作用将动量、能量、水汽等向区域内输送, 因而这一区域处理的好坏对预报的效果起着决定性作用。虽然对侧边界进行调整能使同化时间区间内的目标函数迅速收敛, 但由于侧边界通过两个时次的值插值给出, 而不具有由初值决定的演变特性, 从而使得同化时间区间内的同化作用不能继续影响到随后的预报中, 这在 BDY 试验中很明显的表现出来。

在选择控制变量时, 由于区域中尺度模式预报误差同时来源于初始误差、模式误差和边界条件等多个方面, 将初值、侧边界或模式误差单独做为控制变量时, 往往将某种误差来源归于另一种误差来源, 造成收敛快的假象, 而预报效果却并不好。试验的结果也表明, 同时考虑初值、侧边界和模式误差不但可能, 而且对于预报的改进也有益。

在将侧边界或模式误差单独做为控制变量时, 造成收敛快的假象跟目标函数的定义有关。试验中目标函数定义为距离形式, 不能反映出位相的差异, 单凭总体距离来判断同化好坏有不足之处。一个比较好的目标函数应既能反映距离的差异, 又能反映出位相的差异, 这个工作还应进一步开展。

致谢: 本文作者感谢 Dr. Errico 提供中尺度模式 MM4 的伴随模式软件。

参考文献

- 1 顾震潮. 作为初值问题的天气形势数值预报与由地面天气历史演变做预报的等值性. 气象学报, 1958, 29(2): 176 ~ 186
- 2 丑纪范. 天气数值预报中使用过去资料的问题. 中国科学, 1974, (6): 635 ~ 644
- 3 邱崇践. 模式误差对变分同化过程影响的数值研究. 高原气象, 1994, 13: 449 ~ 456
- 4 郜吉东, 丑纪范. 数值天气预报中的两类反问题及一种数值解法——理想试验. 气象学报, 1994, 52(2): 129 ~ 137
- 5 蒲朝霞, 丑纪范. 对中尺度遥感资料进行四维同化的共轭方法及其数值研究. 高原气象, 1994, 13: 419 ~ 429
- 6 Zupanski M. Regional four-dimensional variational data assimilation in a quasi-operational forecasting environment. Mon Wea Rev, 1993, 121: 2396~ 2408
- 7 Courtier P and Talagrand O. Variational assimilation of meteorological observations with the adjoint vorticity equation. Quart J Roy Meteor Soc, 1987, 113: 1311~ 1328
- 8 Errico R M, Raeder K and Vukicevic T. Mesoscale Adjoint Modeling System Version 1. NCAR Tech Note, 1994
- 9 Derber J C. A variational continuous assimilation technique. Mon Wea Rev, 1989, 117: 2437~ 2466

THE NUMERICAL EXPERIMENT IN AREA FOUR- DIMENSIONAL VARIATIONAL DATA ASSIMILATION

Gong Jiandong Qiu Chongjian

(*Department of Atmospheric Sciences, Lanzhou University, Lanzhou, 730000*)

Wang Qiang Chen Weimin

(*Lanzhou Central Meteorological Bureau, Lanzhou, 730000*)

Abstract

In this paper, a number of numerical experiments are carried out with area mesoscale four-dimensional variational data assimilation system. In those experiments, initial condition, model error and boundary condition are selected as control variable to adjust. The results of data assimilation and following prediction are tested. The main propose of those experiments is to find out the actions of model error and boundary condition in data assimilation, and to select the method for those three as control variable at same time. The results show that model error and boundary condition have large affection to forecast error except initial condition in area mesoscale model. The experiments use a method to correct model error, and the results show this correction makes cost function converge quickly in data assimilation and has persistence in the following prediction. The adjustment to boundary condition makes cost function converge quickly. The results also show that initial condition, model error, and boundary condition should select as control variable at same time in data assimilation. Otherwise, it will transpose one error to another error because the model has no ability to distinguish where forecast error really comes from, and makes cost function converge quickly but the following forecast has no improvement.

Key words: Area four-dimensional variational data assimilation, Error source, Assimilation simultaneity, Numerical experiment.