

GRAPES 集合卡尔曼滤波资料同化系统 I : 系统设计及初步试验^{* 1}

庄照荣^{1,2,3} 薛纪善³ 李兴良^{1,2}
ZHUANG Zhaorong^{1,2,3} XUE Jishan³ LI Xingliang^{1,2}

1. 国家气象中心,北京,100081
 2. 中国气象局数值预报中心,北京,100081
 3. 中国气象科学研究院灾害天气国家重点实验室,北京,100081
1. *National Meteorological Center, Beijing 100081, China*
 2. *Center for Numerical Prediction, CMA, Beijing 100081, China*
 3. *State Key Laboratory of Severe Weather, CAMS, Beijing 100081, China*
- 2009-05-04 收稿,2010-06-24 改回.

Zhuang Zhaorong, Xue Jishan, Li Xingliang. 2011. The GRAPES ensemble Kalman filter data assimilation system. Part I: Design and its tentative experiment. *Acta Meteorologica Sinica*. 69(4):620-630

Abstract The ensemble Kalman filter (EnKF) is able to obtain the flow-dependent background error covariance based on the statistics of the ensemble samples so that it is becoming a research focus in the current data assimilation field. In this paper, a practical GRAPES ensemble Kalman filter data assimilation system is established and its tentative experiment is carried out. In view of the difficulty of assimilating real observations, the observations are organized into batches that assimilated sequentially in this paper. The Schur product is employed in the horizontal and vertical direction for filtering the error correlation noise and alleviating the matrix singularity problem. The ensemble Kalman filter system consistent with the GRAPES model vertical coordinate and forecast variables has been also established. The generation of ensemble samples considers the spatial correlation of model variables and the correlation among the model variables. The initial perturbation field can be obtained by perturbing the control variable in the 3D-Var system. Through the ideal and actual observation experiments, the EnKF system has been verified. Furthermore, the experiment results of assimilating regional radiosonds show that the reasonable analysis of the GRAPES EnKF is obtained and the GRAPES EnKF system can be practically applied to the operation forecast. The 12 h forecast experiment using the GRAPES EnKF analysis indicates that the harmony of the GRAPES EnKF system is not as good as the GRAPES 3D-Var assimilation system.

Key words Data assimilation, Ensemble Kalman filter, Background error covariance

摘要 集合卡尔曼滤波资料同化方法,可以用集合样本统计出随天气形势变化的误差协方差,是当前资料同化领域的研究热点。主要介绍了 GRAPES 集合卡尔曼滤波资料同化系统的设计以及初步的试验结果。针对集合卡尔曼滤波同化实际观测资料难以实施的问题,采用成批观测同化的顺序同化方法进行多变量的集合卡尔曼滤波同化;为了滤除有限集合数造成的误差相关噪声和缓解求逆矩阵不满秩的问题,在水平和垂直方向都采用了 Schur 滤波;建立了与 GRAPES 预报模式的垂直坐标和预报变量一致的模式面集合卡尔曼滤波系统;集合样本的生成考虑了模式变量的空间相关和模式变量之间的相关,通过利用三维变分分析中的控制变量变换得到模式变量扰动场。通过比较 GRAPES 集合卡尔曼滤波资料同化系统和 GRAPES 区

* 资助课题:国家“十一五”科技攻关项目(2006BAC02B02)、中国气象局 GRAPES 专项、国家自然科学基金资助项目(40805045)、中国气象局成都高原气象研究所基本科研业务费专项(BROP200707)。

作者简介:庄照荣,主要从事资料同化研究。E-mail:zhuangzr@cma.gov.cn

域三维变分资料同化系统的单点观测资料同化分析结果,对比背景误差相关系数的分布,验证了 GRAPES 集合卡尔曼滤波系统的正确性。此外,同化区域探空观测资料试验结果表明,GRAPES 集合卡尔曼滤波资料同化系统能够得到合理的分析,并且具有实际运行能力。对分析结果进行 12 h 预报表明,GRAPES 集合卡尔曼滤波资料同化系统的分析协调性不如三维变分资料同化系统。

关键词 资料同化,集合卡尔曼滤波,背景误差协方差

中图法分类号 P456.7

1 引言

数值预报是数学物理中的一个典型的初值问题,模式积分的初值是根据初始时刻的气象观测资料通过特定的资料同化方案而形成的。因此,资料同化在数值预报中占有重要的地位。在过去的 40 年,由于观测资料的增多,计算机能力的提高,资料同化技术的进步(经过了逐步订正法、最优统计插值方法和变分同化方法等的发展),预报模式分辨率的提高及物理过程描述的改进(Eugenia, 2003),数值预报效果已经有较大的改善。当前,数值预报对天气尺度系统的可信预报时效在一周以上。但是,这一时效还短于理论上的上限,数值预报技术尚可进一步提高。而改进数值模式的初始场是数值预报技术提高的一个重要方面,即资料同化技术的改进对数值预报能力的提高起着重要作用。

20 世纪 80 年代初以来,变分同化技术得到了快速发展(Lorenc, 1986, 2000; Baker, 2004),20 世纪 90 年代以后,三维变分同化(3D-Var)技术在世界各国的气象中心逐渐成熟和业务化。然而,3D-Var 整个分析过程只对预报误差协方差矩阵做一次先验的估计,并将其高度模型化,例如采用空间均匀与各向同性的假定(Ingleby, 2001),使其与真实的随着环流而变化的背景误差协方差有很大差别,从而大大影响了同化的结果。为了克服这一问题,科学家致力于发展更高级的同化技术。在欧洲中期天气预报中心等单位,四维变分同化(4D-Var)分析技术已比较成熟,并对数值预报能力的提高有很大作用。在 4D-Var 中,由于预报模式的引入,隐式地考虑了模式预报误差协方差在分析时间窗中的演变,使得观测信息的传播反映了随着环流变化的特点。但 4D-Var 需要复杂的切线性和伴随模式,而对于具有复杂物理过程的数值预报模式,切线性和伴随模式的求取是困难的,甚至根本不存在,并且初始的背景误差还必须先验地指定(Fisher, 2001)。因

此,科学家还在发展非变分类的其他四维同化技术,其中卡尔曼滤波方法是在资料同化领域研究比较早的一个方法。在卡尔曼滤波中,预报或背景误差协方差由模式本身的演变来确定,它随天气流型和时间而变化。但由于卡尔曼滤波需要同时预报误差协方差矩阵,计算花费很大,对于像真实大气模式的超大规模问题,至少在当前的计算机条件下不可实现,一般只应用于较简单的模式。因此,很长时间以来卡尔曼滤波并没有成为一种实际的资料同化方案。在 20 世纪 90 年代中期,科学家提出了集合卡尔曼滤波(the Ensemble Kalman Filter, EnKF)同化方法,将卡尔曼滤波与集合预报结合在一起,形成一种有应用前途的卡尔曼滤波的简化方法。由于集合卡尔曼滤波方法不要求发展切线性和伴随模式,而且可以显式地提供集合预报的初始扰动,从而成为国际上资料同化研究的热点。

集合卡尔曼滤波是基于随机动力预测理论发展而来(Epstein, 1969)。Evensen(1994)将集合卡尔曼滤波应用到海洋同化领域中。Evensen 设计的集合卡尔曼滤波是一个用蒙特卡罗的短期集合预报方法来估计预报误差协方差的四维同化方法。Houtekamer 等(1998)首先把集合卡尔曼滤波方法引入到大气资料同化系统中,并进行了一系列研究,提出用两组集合来做集合卡尔曼滤波,两组集合的使用可减少分析过程中的近亲繁殖问题。Houtekamer 等(2001)研究使用 Schur 乘积的方法将距观测较远的相关噪音逐渐滤除,来代替对相关噪音的直接截断;同时,为了解决当使用真实的观测资料和较大的集合数时,集合卡尔曼滤波的观测直接同化难以实施的问题,提出将观测分为多组来顺序同化,对于每组观测用 Cholesky 分解算法来求解系统的线性方程组;2005 年在集合卡尔曼滤波分析中同化了实际的观测资料,并用有限的集合获得了合理的分析结果。

中国近几年才开始关注集合卡尔曼滤波的研

究,刘成思(2005)对集合卡尔曼滤波系统初步建立了一个针对浅水方程模式的试验平台,针对多变量分析中变量不平衡问题,提出了半约束集合卡尔曼滤波方案,并进行了理想试验。集合卡尔曼滤波方法在中小尺度资料同化方面的工作也有所展开,许小永(2006)利用集合卡尔曼滤波在云数值模式中同化模拟多普勒雷达资料,并考察了不同条件下集合卡尔曼滤波同化方法的性能。

中国对集合卡尔曼滤波的研究尚处于初步阶段,对观测资料的同化主要以逐个观测资料同化的方式进行,计算资源耗费很大,很难进行真实的观测资料同化和业务运行。目前,GRAPES 三维变分资料同化系统已经发展得比较成熟(张华等,2004;庄世宇等,2005),考虑到集合卡尔曼滤波系统可为集合预报提供初始场,还可应用于预报对观测的敏感性试验研究,并可作为专门针对强非线性天气过程的资料同化方法,因而集合卡尔曼滤波资料同化方法是 GRAPES 同化系统的一个发展方向。因此,本文以构造可同化真实观测资料、可实施的集合卡尔曼滤波系统为目的,对集合卡尔曼滤波系统的实施方案进行详细描述,并且做了实际观测资料同化试验。

2 GRAPES 集合卡尔曼滤波同化系统的设计

2.1 集合卡尔曼滤波理论方法的简单介绍

集合卡尔曼滤波方法经过多年的发展已经比较成熟, Evensen(2003)对集合卡尔曼滤波的理论和实施方法进行了详细的介绍。在集合卡尔曼滤波中,背景场由一组集合组成,若背景集合成员为 $\mathbf{X}^b = (\mathbf{x}_1^b, \mathbf{x}_2^b, \dots, \mathbf{x}_N^b) \in \mathbf{R}^{s \times N}$, 其中 s 为模式状态变量的数目, N 为集合样本数,背景集合平均为 $\bar{\mathbf{x}}^b$, 则集合扰动为

$$\mathbf{X}' = \mathbf{X}^b - \bar{\mathbf{x}}^b \quad (1)$$

用集合样本估计的背景误差协方差矩阵为

$$\hat{\mathbf{P}}^b = \langle (\mathbf{x}_i^b - \bar{\mathbf{x}}^b)(\mathbf{x}_i^b - \bar{\mathbf{x}}^b)^T \rangle = \frac{\mathbf{X}'(\mathbf{X}')^T}{N-1} \quad (2)$$

假设观测集合为 $\mathbf{Y}^o = (\mathbf{y}_1^o, \mathbf{y}_2^o, \dots, \mathbf{y}_N^o) \in \mathbf{R}^{m \times N}$, 其中, m 为观测的数目,观测扰动为 $\mathbf{r} = (\epsilon_1, \epsilon_2, \dots, \epsilon_N) \in \mathbf{R}^{m \times N}$, 观测扰动服从正态分布 $\mathbf{r} \sim N(0, \mathbf{R})$, 同理,观测误差协方差为

$$\mathbf{R}_e = \langle (\mathbf{y}_i^o - \bar{\mathbf{y}}^o)(\mathbf{y}_i^o - \bar{\mathbf{y}}^o)^T \rangle = \frac{\mathbf{r}\mathbf{r}^T}{N-1} \quad (3)$$

若用 \mathbf{D}' 代表新息向量,则

$$\mathbf{D}' = \mathbf{Y}^o - \mathbf{H}\mathbf{X}^b \in \mathbf{R}^{m \times N} \quad (4)$$

式中, \mathbf{H} 为观测算子,则集合卡尔曼滤波的分析可以写为

$$\mathbf{X}^a = \mathbf{X}^b + \mathbf{X}'\mathbf{X}'^T\mathbf{H}^T[\mathbf{H}\mathbf{X}'\mathbf{X}'^T\mathbf{H}^T + \mathbf{r}\mathbf{r}^T]^{-1}\mathbf{D}' \quad (5)$$

从以上公式可以看出,集合卡尔曼滤波的背景误差协方差可由背景集合样本统计得到。集合卡尔曼滤波资料同化的最初方案则是利用蒙特卡罗方法设计预报状态的一组集合,该集合的平均为对大气的预报,集合成员间的离差统计即作为背景误差协方差的近似,该集合通过任何模式向前积分,每个样本分别更新分析变量,而对变量的最佳估计即为各更新分析变量的样本平均。

在当前的气象资料同化问题中,通常观测数的量级为 10^5 , 模式状态变量级的大小为 10^7 , 这样在式(5)中,对于大矩阵 $\mathbf{H}\mathbf{X}'\mathbf{X}'^T\mathbf{H}^T + \mathbf{r}\mathbf{r}^T$ 的求逆与背景误差相关 $\mathbf{X}'\mathbf{X}'^T\mathbf{H}^T$ 的计算是非常困难的,对实际观测资料的同化难以实施。另外,由于集合样本数有限,集合成员估计的误差相关有噪音,可造成分析不平滑。而且一般的集合成员数都远远少于观测维数,从而使矩阵 $\mathbf{H}\mathbf{X}'\mathbf{X}'^T\mathbf{H}^T + \mathbf{r}\mathbf{r}^T$ 的秩不大于样本数,矩阵是一个不满秩矩阵。对于一个不满秩矩阵,数学上是不存在逆矩阵的。针对以上问题,提出了解决集合卡尔曼滤波实施的以下关键技术问题。

2.2 GRAPES 集合卡尔曼滤波同化的实现

2.2.1 GRAPES 预报模式与模式变量的分析

集合卡尔曼滤波系统可以在等压面上构建,也可以在模式面上实现。在等压面集合卡尔曼滤波中,大气状态变量为位势高度 ϕ 、风速 u 和 v 、湿度 q/rh , 而垂直坐标则选为气压,这样设计的同化系统不针对具体的预报模式可以灵活地与不同模式联结。但由于实际预报模式并不采用等压面坐标,在形成预报-同化循环时,当模式预报作为同化的背景场和同化分析作为模式积分的初始场,则都需要进行模式面到等压面或者相反方向的插值,从而会带来额外的插值误差。此外,模式面和等压面之间还必须进行变量间的转换,从而增加了新的误差。为了消除这两项误差,在等压面同化方案基础上,需要发展完全针对模式的模式面同化系统。

GRAPES 模式的预报变量是无量纲气压(Exner 函数) PI 、位温 θ 、风的三个分量 u, v, w 以及水

成物变量。模式的垂直坐标是地形追随的标准化高度

$$\tilde{z} = Z_T \frac{z - Z_s(x, y)}{Z_T - Z_s(x, y)} \quad (6)$$

式中, Z_T 为模式顶高度, $Z_s(x, y)$ 为地形高度。针对 GRAPES 模式, 选取 PI 作为同化系统的状态变量, 而把 θ 作为 PI 的导出量。在模式面分析中, 模式变量与等压面上的观测要素不一致, 需要对观测算子进行重新构造, 即要利用上述模式预报变量将观测要素表达出来。而且模式面上的变量分布在垂直方向采用 Charney-Phillips 跳点, 即垂直速度 w 、位温 θ 和湿度在模式的整数层上, u 、 v 和 PI 在模式的中间层(半层); 水平方向采用 Arakawa-C 跳点, 因而分析中模式变量的水平网格结构和垂直分层要重新组织。

在 GRAPES 非静力预报模式的三维变分系统(马旭林, 2009)基础上, 发展了模式面集合卡尔曼滤波同化系统, 即在与模式完全一致的垂直坐标系下, 通过集合卡尔曼滤波同化系统直接得到相应的模式预报变量。

2.2.2 扰动集合样本的生成

集合卡尔曼滤波方法需要一组集合样本作为背景场, 首先需要构造合理的初值扰动场。集合成员是否协调合理对分析和预报结果影响很大, 因而初值扰动生成方法很重要。一般扰动场的特征要求与实际分析资料中可能的误差分布一致, 以保证每个初始场都可能代表大气的实际状态, 而且每个初始扰动在模式中的演变方向要尽可能大的发散, 以保证预报集合最大可能地包含实际大气所能出现的状况(杨学胜, 2001)。

初值扰动的生成方法有两类, 第一类是模拟分析误差的概率分布, 如蒙特卡罗随机扰动法(Houtekamer, 1995); 第二类是从数值预报误差分析中产生的, 考虑相空间误差增长方向形成初始扰动。从理论上说, 沿着预报系统相空间中最不稳定的方向扰动初始条件可以描述对预报的不确定性有最明显贡献的初始扰动, NCEP 和 ECMWF 相继应用增长模繁殖法和奇异向量法形成初始扰动。

本文采用的初值扰动方法是基于蒙特卡罗随机扰动, 考虑了模式变量的空间相关, 以及模式变量之间的相关, 通过利用三维变分中的控制变量变换得到模式变量扰动场。由于控制变量是归一化的变

量, 其误差均方差为 1, 对三维变分中的控制变量 W 加入均值为 0、方差为 1 的随机高斯扰动, 然后通过控制变量转换(递归滤波、EOF 分解、平衡变换、物理变换), 可以得到平衡的、尺度化、相关的背景场集合(Barker, 2005)。

以上方法得到的背景场集合的误差相关分布近似高斯型, 但不是随天气形势而变化的。一般可以把以上背景场集合预报 6 或 12 h, 用此时的预报场集合作为集合卡尔曼滤波的背景场, 它的误差协方差分布是随着天气形势变化的。

此外, 在集合卡尔曼滤波系统中, 对观测资料也进行了扰动。不考虑观测资料之间的相关, 对观测资料的扰动采用均值为 0、均方差为观测误差大小的随机高斯扰动。

2.2.3 顺序同化方法

针对集合卡尔曼滤波难以实施的问题, 文中采用了顺序同化方法。考虑到在观测站点之间互不相关的情况下, 观测资料的逐点顺序同化与同时同化具有等价性, 若采用单点(单个站点、单个变量、单层)的顺序同化方法, 矩阵求逆就变为标量的倒数。但是, 单点的顺序同化方法有大量的重复计算, 计算效率较低, 在业务运行中难以实施。考虑了顺序同化的优点和计算代价, 采取逐个观测廓线成批同化的方法或多个观测廓线分组的同化方法。这样, 矩阵 $HX'X'^T H^T + rr^T$ 不满秩的问题也会得到很大程度的缓解, 它的秩和观测维数就相差不大。例如, 一般常规无线电探空观测站最多的观测层为 200 层, 观测变量为质量场、 u 、 v 风场和湿度场, 若采用逐个观测廓线同化, 矩阵 $HX'X'^T H^T + rr^T$ 的维数最大为 800×800 , 此时, 对矩阵求逆是可行的。文中对矩阵求逆采用了奇异值分解, 即得到

$$HX'X'^T H^T + rr^T = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T \quad (7)$$

式中, \mathbf{U} 为奇异向量, $\mathbf{\Lambda}$ 为奇异值, 是对角阵。把式(7)代入式(5), 则

$$\mathbf{X}^a = \mathbf{X}^b + \mathbf{X}'(\mathbf{H}\mathbf{X}')^T \mathbf{U}\mathbf{\Lambda}^{-1} \mathbf{U}^T \mathbf{D}' \quad (8)$$

对于以上的分析式, 考虑计算的效率, 尽量在维数较低的集合空间和观测空间计算, 文中采用以下的计算步骤:

第 1 步, 把矩阵求逆转换为奇异值的倒数, 在观测空间计算奇异值的逆和奇异向量的积

$$\mathbf{A}_1 = \mathbf{\Lambda}^{-1} \mathbf{U}^T \in \mathbf{R}^{m \times m} \quad (9)$$

第 2 步, 在观测空间上, 用集合成员的线性组合

表示 \mathbf{A}_1 和新息向量的积

$$\mathbf{A}_2 = \mathbf{A}_1 \mathbf{D}' \in \mathbf{R}^{m \times N} \quad (10)$$

第3步,在观测空间上,计算奇异向量和 \mathbf{A}_2 的积

$$\mathbf{A}_3 = \mathbf{U} \mathbf{A}_2 \in \mathbf{R}^{m \times N} \quad (11)$$

第4步:在集合空间,计算

$$\mathbf{A}_4 = (\mathbf{H} \mathbf{X}')^T \mathbf{A}_3 \in \mathbf{R}^{N \times N} \quad (12)$$

第5步:在模式变量空间,计算分析增量 $\mathbf{X}' \mathbf{A}_4$ 及分析

$$\mathbf{X}^a = \mathbf{X}^b + \mathbf{X}' \mathbf{A}_4 \in \mathbf{R}^{s \times N} \quad (13)$$

在式(8)中,采用逐个观测廓线同化,因而 $\mathbf{X}' \mathbf{X}'^T \mathbf{H}^T$ 表示为分析格点和某观测站上所有层的背景误差相关。若背景误差协方差水平和垂直方向可以分解,则某变量集合 $\mathbf{x}_{i,k}$ (水平位置为 i , 垂直位置为模式层第 k 层) 与某观测点处的某变量集合 $\mathbf{x}_{j,m}$ (水平位置为 j , 垂直位置为 m) 的背景误差协方差为

$$p_{i,j,k,m} = \frac{\mathbf{x}'_{i,k} \mathbf{x}'_{j,m}{}^T}{N-1} \approx \frac{\mathbf{x}'_{i,k} \mathbf{x}'_{j,k}{}^T}{N-1} \left[\frac{\mathbf{x}'_{j,k} \mathbf{x}'_{j,k}{}^T}{N-1} \right]^{-1} \frac{\mathbf{x}'_{j,k} \mathbf{x}'_{j,m}{}^T}{N-1} \quad (14)$$

式中, $\frac{\mathbf{x}'_{i,k} \mathbf{x}'_{j,k}{}^T}{N-1}$ 为模式层第 k 层 i 与 j 点的水平误差协方差, $\frac{\mathbf{x}'_{j,k} \mathbf{x}'_{j,m}{}^T}{N-1}$ 为在观测点 j 处的第 k 层与第 m 层的垂直误差协方差, $\frac{\mathbf{x}'_{j,k} \mathbf{x}'_{j,k}{}^T}{N-1}$ 为观测站点 j 处所在模式层第 k 层的方差。因此,在第4步和第5步把背景和观测相当量的误差协方差水平和垂直分离,则

$$\hat{\mathbf{P}}^b \mathbf{H}^T = \frac{\mathbf{X}' (\mathbf{H} \mathbf{X}')^T}{N-1} \approx \mathbf{B}_h \mathbf{B}_m^{-1} \mathbf{B}_v \quad (15)$$

上式表明分析格点和观测的背景误差相关 ($\hat{\mathbf{P}}^b \mathbf{H}^T$) 可以用分析格点与观测所在位置的一层背景误差水平协方差 (\mathbf{B}_h) 和观测所在位置的模式层与观测点

层的误差垂直协方差 (\mathbf{B}_v) 的积表示,式中, \mathbf{B}_m 为观测点所在的模式层间的背景误差方差。

采用以上的计算步骤,在逐个观测廓线的同化过程中,由于分析过程大部分在观测空间和集合空间进行,大大提高了计算效率。

2.2.4 Schur 滤波

在集合卡尔曼滤波方法中,有限集合数造成矩阵不满秩和估计的相关有虚假噪音,为了滤除距观测较远的相关噪音,减轻不满秩问题,通常在集合卡尔曼滤波中采用截断半径或者 Schur 乘积。截断半径的使用会使分析变量不连续,造成分析不协调。本文采用 Schur 乘积算法,把从集合样本估计的背景误差协方差和局地相关函数相乘,得到的也是协方差函数。若误差协方差矩阵为 $\hat{\mathbf{P}}^b$, 局地相关矩阵为 \mathbf{A} , 则 Schur 乘积为 \mathbf{C} , 有 $C_{ij} = \mathbf{A}_{ij} \hat{\mathbf{P}}^b_{ij}$ 。对背景误差协方差采用 Schur 乘积,对于增益矩阵

$$\hat{\mathbf{K}} = [(\rho \circ \hat{\mathbf{P}}^b) \mathbf{H}^T] [\mathbf{H}(\rho \circ \hat{\mathbf{P}}^b) \mathbf{H}^T + \mathbf{R}_e]^{-1} \quad (16)$$

其中, $\rho \circ \hat{\mathbf{P}}^b$ 为相关矩阵 \mathbf{A} 和协方差矩阵 $\hat{\mathbf{P}}^b$ 的 Schur 乘积, ρ 为局地相关函数。

在以上设计方案中,由于采用了观测廓线的顺序同化方法,对于 $\mathbf{H} \hat{\mathbf{P}}^b \mathbf{H}^T$ 不用加 Schur 算子。对 $\mathbf{X}' \mathbf{X}'^T \mathbf{H}^T$ 采用水平和垂直方向分解后,对水平和垂直误差相关都采用 Schur 乘积算子,将式(15)代入第4步和第5步,即有

$$\mathbf{X}^a = \mathbf{X}^b + (N-1) \rho_h \circ \mathbf{B}_h \mathbf{B}_m^{-1} \mathbf{B}_v \circ \rho_v \mathbf{A}_3 \quad (17)$$

式中, ρ_h 和 ρ_v 分别为在水平、垂直方向采用的局地相关函数。

文中定义 ρ 为 5 阶分段有理函数 (Gaspari, 1999), 这个函数是三角函数的自卷。它是各向同性的, 依赖于一个长度尺度参数 (Schur 半径), 并随着距离单调下降的函数。5 阶分段有理函数为

$$\rho(r, c) = \begin{cases} -\frac{1}{4}(|r|/c)^5 + \frac{1}{2}(|r|/c)^4 + \frac{5}{8}(|r|/c)^3 - \frac{5}{3}(|r|/c)^2 + 1 & 0 \leq |r| \leq c \\ \frac{1}{12}(|r|/c)^5 - \frac{1}{2}(|r|/c)^4 + \frac{5}{8}(|r|/c)^3 + & \\ \frac{5}{3}(|r|/c)^2 - 5(|r|/c) + 4 - \frac{2}{3}c/|r| & c < |r| \leq 2c \\ 0 & 2c < |r| \end{cases} \quad (18)$$

式中, c 为 Schur 半径, r 为任意两点的距离。 ρ 在两倍 Schur 半径距离内是非零的, 之外为零。

Schur 乘积可以逐步滤除距观测较远的相关噪声,这种局地化策略和截断半径一样,提高了矩阵 $P^b H^T$ 和 $HP^b H^T$ 的条件数,缓解求逆矩阵不满秩的问题。而且,由于 5 阶分段有理函数是平滑和单调下降的,Schur 乘积可以减少和平滑观测点之间的影响效果,所以得到分析增量是平滑的。另外,Schur 乘积的使用可以使集合卡尔曼滤波的计算量大大减少。

3 集合卡尔曼滤波理想和实例分析试验

为了与三维变分分析系统比较,用空间相关近似高斯分布的初始集合场作为背景场,进行了模式面集合卡尔曼滤波同化单点观测及实际探空观测资料的试验,验证集合卡尔曼滤波系统的正确性,并且比较集合卡尔曼滤波和三维变分分析系统的分析增量和分析变量间的协调性。

3.1 试验设置

试验对 2005 年 7 月 10 日 00 时(世界时,下同)的无线电探空观测资料进行模式面的集合卡尔曼滤波和三维变分同化分析以及 12 h 预报试验。试验的区域为(4.5°—65.8°N,54.0°—146.8°E),分辨率为 $0.5625^\circ \times 0.5625^\circ$,水平格点数为 166×110 ,垂

直分为 31 层。

在模式面的集合卡尔曼滤波和三维变分分析系统中,模式变量都选用无量纲气压 PI 、 u 风场、 v 风场和比湿 q 。其中在集合卡尔曼滤波系统中,采用集合数为 50;水平和垂直相关方案采用集合样本统计和 Schur 滤波,水平方向的 Schur 半径为 1000 km,垂直方向的 Schur 半径选用 2,矩阵求逆采用了奇异值分解,采用观测成批同化的顺序同化方法。

三维变分分析用 T213 的 2005 年 7 月 9 日 18 时的 6 h 预报场作为背景场,集合卡尔曼滤波分析采用三维变分控制变量 W 扰动的方法得到的扰动场集合叠加到 T213 的 2005 年 7 月 9 日 18 时的 6 h 预报场上作为背景场集合,此时的背景集合平均为 T213 预报场,而背景集合样本统计的误差结构与三维变分采用的误差模型近似。

为了得到合理的分析,文中对背景误差和观测误差进行了调整(图 1)。从图 1 可以看出,位势高度的背景误差大于观测误差,为 12—24 gpm,风场的背景误差也大于观测误差,为 3—6 m/s。在模式面集合卡尔曼滤波和三维变分同化系统中都使用以上的背景误差和观测误差,或者以上误差的导出量。

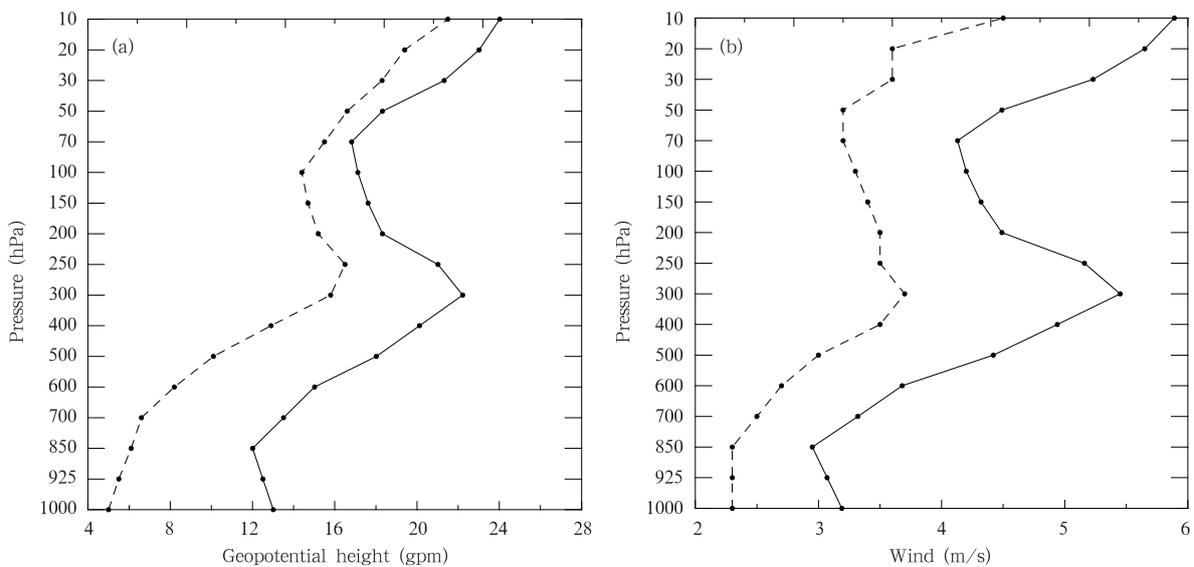


图 1 背景误差和观测误差随高度的分布

(实线为背景误差,虚线为观测误差;a. 位势高度(单位:gpm), b. 风场(单位:m/s))

Fig. 1 Standard deviations of the background (solid line) and observation (dashed line) values of (a) the geopotential height (gpm) and (b) the wind (m/s) with the height

3.2 单点观测资料同化试验

为了验证集合卡尔曼滤波分析系统的正确性,首先进行单点气压观测资料同化试验,在单点试验中,集合卡尔曼滤波分析没有采用 Schur 滤波。在 2005 年 7 月 10 日 00 时,用三维变分和集合卡尔曼

滤波同化单点气压观测资料,图 2 为同化单点气压观测资料得到的三维变分和集合卡尔曼滤波的水平分析增量场,以及观测点 A(处于 16 层,31.5°N,99°E)与其他格点的背景误差水平相关系数的分布,其中集合卡尔曼滤波分析增量为集合平均的结果。

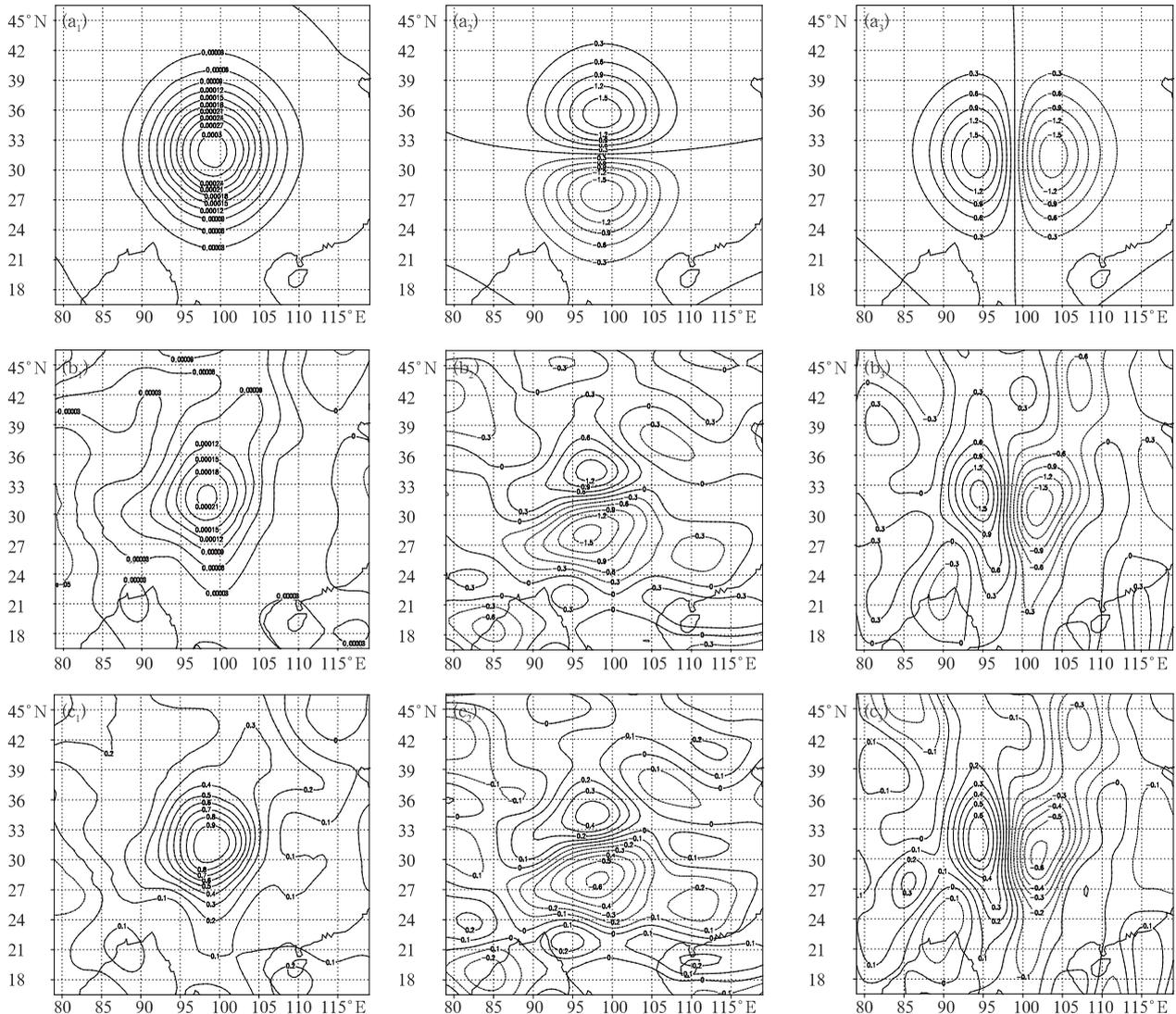


图 2 3D-Var 的水平分析增量 (a_1 . PI , a_2 . u (m/s), a_3 . v (m/s)), EnKF 的水平分析增量 (b_1 . PI , b_2 . u (m/s), b_3 . v (m/s)), 观测点 A 无量纲气压与其他格点的背景误差相关系数的分布 (c_1 . $PI-PI$, c_2 . $u-PI$, c_3 . $v-PI$)

Fig. 2 Horizontal analysis increments of the 3D-Var (a_1) PI , (a_2) u (m/s), (a_3) v (m/s); the horizontal analysis increments of the EnKF (b_1) PI , (b_2) u (m/s), (b_3) v (m/s); and the correlation fields with respect to the observation point A (c_1) $PI-PI$, (c_2) $u-PI$, (c_3) $v-PI$

集合卡尔曼滤波同化得到的水平分析增量(图 2 b_1 — b_3)的大小和分布情况与三维变分得到的分析增量(图 2 a_1 — a_3)相似。在集合卡尔曼滤波系统中,背景误差协方差用 50 个集合样本统计得到,由于有

限集合数造成的样本误差,使观测信息的传播不是均匀的高斯分布,表现在分析增量的分布不是很规则。此外,没有使用 Schur 算子,因而背景误差较小相关的不准确描述造成距观测较远处存在分析噪

音。

从图 2c₁ 可以看出,观测点处的无量纲气压 PI 与其他格点上 PI 的背景误差水平相关从观测点到四周近似同心圆,也就是近似高斯型分布。观测点处的无量纲气压 PI 与其他格点上 PI 、 u 、 v 风场的误差水平相关决定了观测信息向四周传播的情况,也就是相关结构决定了分析增量的分布情况。比较

图 2b₁—b₃ 和 c₁—c₃ 可以看出集合卡尔曼滤波分析增量的分布和背景误差相关系数的分布一致。

另外,在集合卡尔曼滤波中,垂直误差相关也是用集合样本统计得到,在垂直方向没有使用 Schur 滤波,图 3 给出同化单点气压观测的分析增量的垂直分布和观测点 A(处于 16 层, 31.5°N, 99°E) 与其他格点的背景误差垂直相关情况。

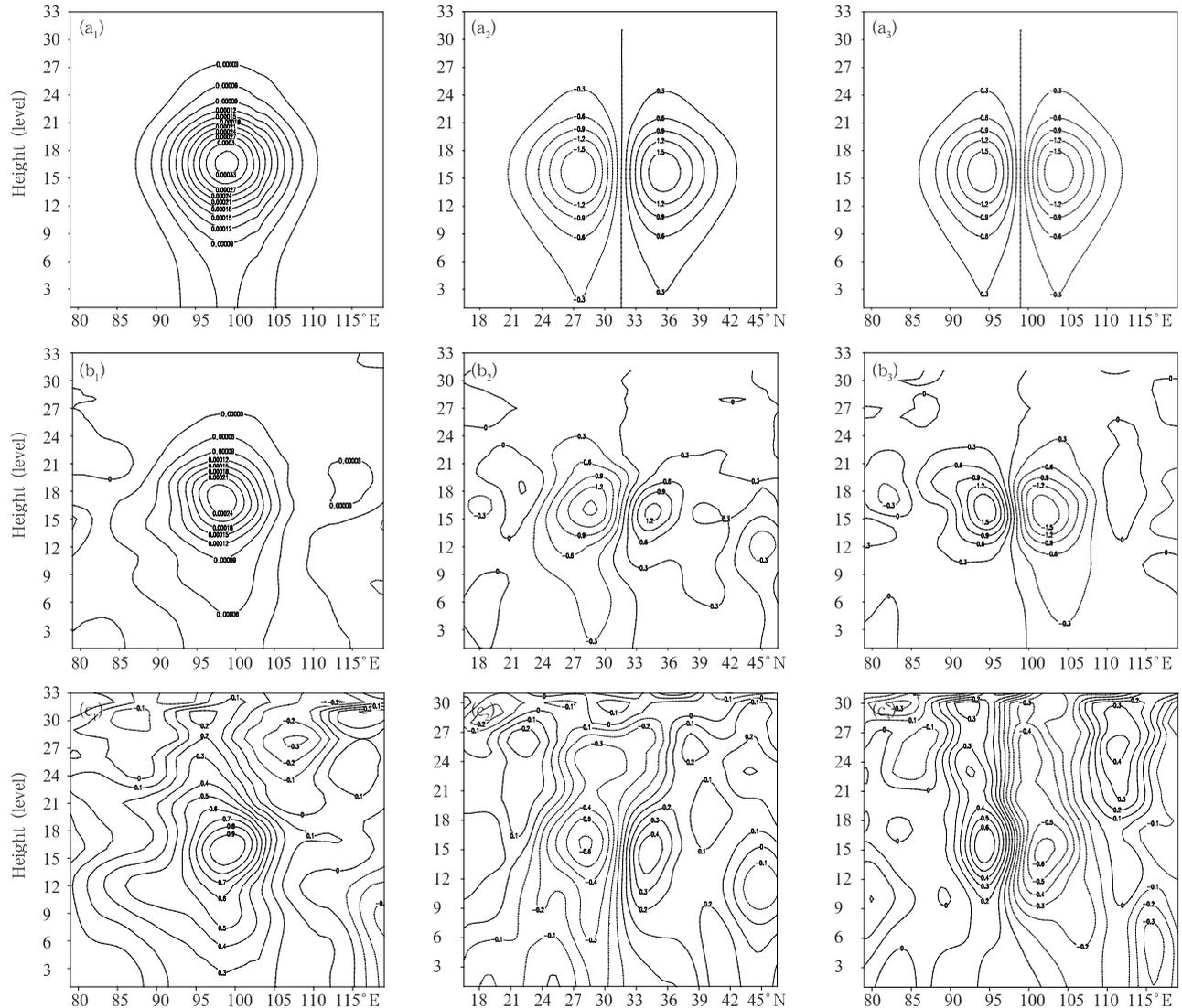


图 3 3D-Var 的垂直分析增量(a₁. PI , a₂. u (m/s), a₃. v (m/s)), EnKF 的垂直分析增量 (b₁. PI , b₂. u (m/s), b₃. v (m/s)), 观测点 A 与其他格点的背景误差相关系数的分布 (c₁. $PI-PI$, c₂. $u-PI$, c₃. $v-PI$)
 Fig. 3 Vertical analysis increments of the 3D-Var (a₁ PI , (a₂ u , (a₃ v); the vertical analysis increments of the EnKF (b₁ PI , (b₂ u , (b₃ v); and the correlation fields with respect to the observation point A (c₁ $PI-PI$, (c₂ $u-PI$, (c₃ $v-PI$)

三维变分的垂直分析增量场很平滑和规则(图 3a),而经过集合卡尔曼滤波同化后的垂直分析增量

场(图 3b)略不规则,但分析增量场的分布和大小与三维变分的近似。比较图 3b₁—b₃ 和 c₁—c₃ 可以看

出,集合卡尔曼滤波的垂直分析增量的结构由背景误差垂直相关系数的结构所决定,二者的分布基本一致。

从单点气压观测的分析可以看出,在基本相同的误差设置下,集合卡尔曼滤波和三维变分的分析结果相近,说明集合卡尔曼滤波在同化气压观测时是正确的。

同理,用单点 u 、 v 风场观测或者湿度观测的分析也可以验证集合卡尔曼滤波系统同化其他变量观

测的正确性(图略)。

由上可知,在基本相同的误差设置条件下,通过比较集合卡尔曼滤波和三维变分的单点观测资料同化分析增量场,对比背景误差相关系数的分布,验证了集合卡尔曼滤波系统的正确性。

3.3 实例分析

2005年7月10日00时,同化区域所有的探空观测资料,三维变分分析的分析增量及集合卡尔曼滤波分析的集合平均分析增量的分布情况如图4。

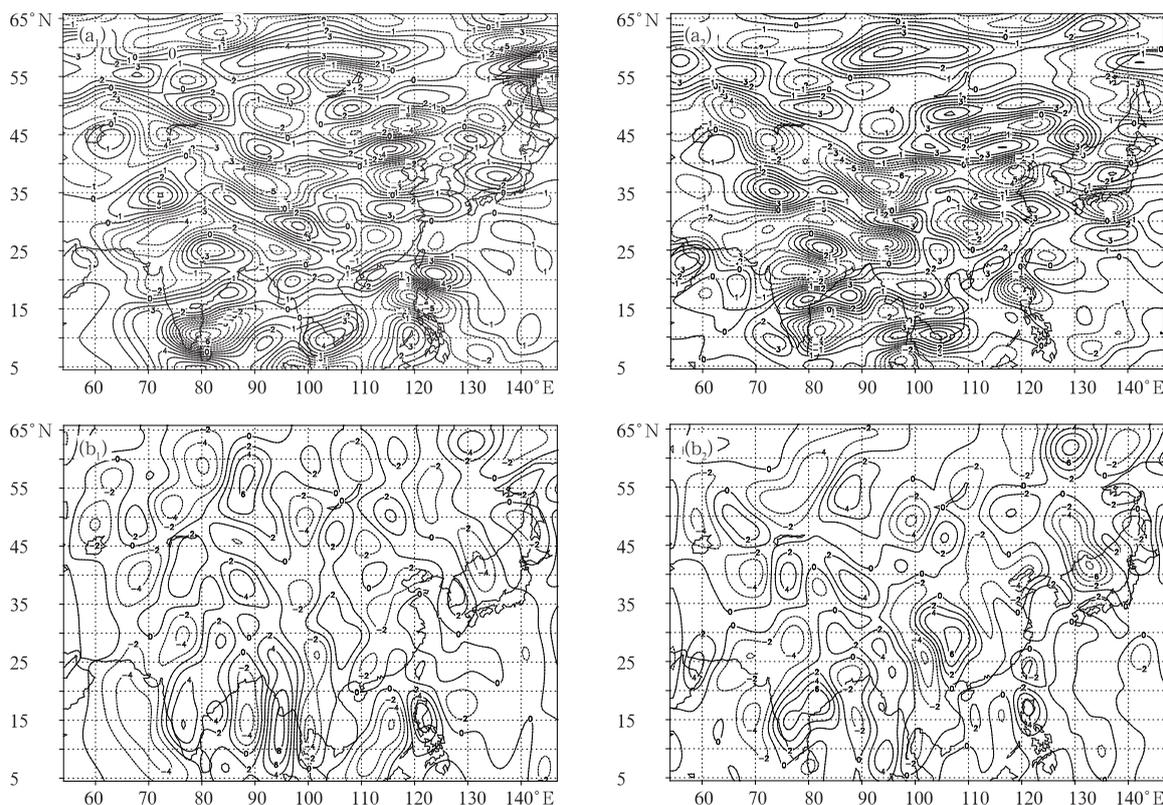


图4 2005年7月10日00时 $u(a_1, a_2)$ 、 $v(b_1, b_2)$ 分析增量场分布
(a_1, b_1 . 三维变分; a_2, b_2 . 集合卡尔曼滤波; 单位: m/s)

Fig. 4 Analysis increments of the 3D-Var (a_1) u and (b_1) v ; and the analysis increments of the EnKF (a_2) u and (b_2) v (unit: m/s) at 00:00 UTC 10 July 2005

从 u 风场的分析增量分布可以看出(图 4 a_1 、 a_2),集合卡尔曼滤波和三维变分的分析增量的整体结构相似,增量场极值中心的大小略有差别。而从 v 风场也可以看出集合卡尔曼滤波和三维变分的分析增量分布相似。从无量纲气压 PI 和比湿的分析增量也可以看出(图略),两种分析方法得到的分析增量中心的分布和大小基本相同。

由以上分析可知,由于采用控制变量扰动的初

始集合场作为背景场,使集合卡尔曼滤波的背景误差相关分布近似高斯分布,而且背景和观测的误差设置和三维变分同化分析系统中的一致,所以同化探空观测资料后,集合卡尔曼滤波和三维变分分析增量的结构和大小基本相同,从而验证了集合卡尔曼滤波系统的正确性。

3.4 分析协调性

为了比较集合卡尔曼滤波和三维变分同化系统

分析的协调性,下面给出两种同化方法同化探空观测资料后预报 12 h 的地面位势高度倾向变化,以及直接用背景场预报 12 h 的地面位势高度倾向变化。

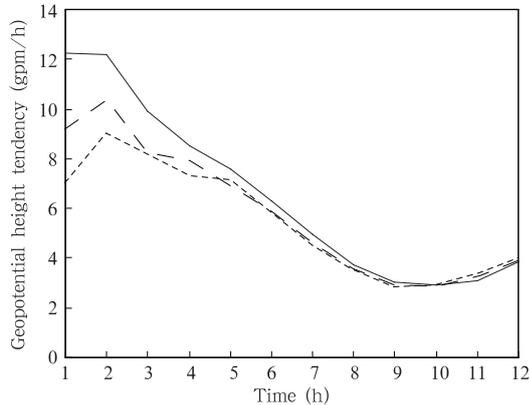


图5 背景场(点线)、三维变分(虚线)、集合卡尔曼滤波(实线)12 h 预报的地面位势高度倾向变化(单位:gpm/h)

Fig. 5 Ground geopotential height tendency of the background field (dotted line), the 3D-Var (dashed line) and the EnKF (solid line) (unit: gpm/h)

从图5可以看出,背景场变量之间最协调,开始的地面位势高度倾向为 7.06 gpm/h;其次是三维变分同化系统,开始的地面位势高度倾向为 9.22 gpm/h;集合卡尔曼滤波分析场的协调性不如三维变分,它开始的地面位势高度倾向为 12.26 gpm/h。经过模式 6 h 积分以后,变量之间基本协调,两种分析预报场的位势倾向变化基本与背景预报场的一致。

集合卡尔曼滤波的分析协调性不如三维变分分析,这主要是由于在集合卡尔曼滤波分析中,不仅背景场进行了扰动,观测也进行了扰动,多次扰动使变量之间的协调性降低;而且分析中也没有考虑变量之间的平衡关系;此外,采用的观测顺序同化方法对分析的协调性也有影响。

4 结论与讨论

本文概述了 GRAPES 集合卡尔曼滤波同化系统的技术方案及实施方法,为了解决集合卡尔曼滤波同化实际观测资料难以实施的问题,文中采用了把观测分批同化的策略,这样不仅解决了增益矩阵很难求逆问题,而且有限集合数造成的不满秩现象

也得到缓解。此外,由于有限的集合样本造成的远距离相关噪声严重影响分析的质量和协调性,文中采用了 Schur 乘积来滤除远距离的相关噪声。

文中通过理想试验表明在集合卡尔曼滤波分析中观测信息按照背景集合统计出的背景误差相关系数的分布传播,因而分析增量的分布与背景误差相关系数一致。当集合卡尔曼滤波的背景误差相关分布近似高斯分布,而且在同样的背景和观测误差设置下,集合卡尔曼滤波和三维变分同化探空观测资料得到分析增量的结构和大小基本相同,验证了集合卡尔曼滤波系统的正确性。经过观测资料同化后,集合卡尔曼滤波的分析协调性不如三维变分分析,这是集合卡尔曼滤波系统自身造成的。

文中用集合卡尔曼滤波只同化了常规探空观测,同化计算所消耗的时间远比三维变分同化运行时间长。如果采用大批量卫星、雷达等观测资料,需要对集合卡尔曼滤波采用并行计算。集合卡尔曼滤波各个成员在分析预报循环中,除了对误差协方差分析时,运算基本上是独立的,可以实现高效率的并行化。集合卡尔曼滤波可以为集合预报提供一组统计意义最优的初值,集合预报也可以为同化提供尽可能准确的误差统计,集合卡尔曼滤波和集合预报相结合是未来发展趋势。同化和集合预报在中国系统地研究起步晚,和其他国差距大,而对于基于集合的同化理论和技术的研究,不但能促进资料同化的理论技术发展,而且必然会推动集合预报理论和技术发展。

致谢:本文完成过程中,国家气象中心刘成思给予了很大的帮助和支持,国家气象中心朱国富和中国气象科学研究院庄世宇也给予宝贵的建议,在此一并表示感谢。

参考文献

- 刘成思. 2005. 集合卡尔曼滤波资料同化方案的设计和研发[D]. 北京:中国气象科学研究院,1-69
- 马旭林,庄照荣,薛纪善等. 2009. GRAPES 非静力数值预报模式的三维变分资料同化系统的发展. 气象学报, 67(1):50-60
- 许小永,刘黎平,郑国光. 2006. 集合卡尔曼滤波同化多普勒雷达资料的数值试验. 大气科学, 30(4):712-728
- 杨学胜. 2001. 业务集合预报系统的现状及展望. 气象,27(6):3-9
- 张华,薛纪善,庄世宇等. 2004. GRAPES 三维变分同化系统的理想试验. 气象学报,62(1):31-41
- 庄世宇,薛纪善,朱国富等. 2005. GRAPES 全球三维变分同化系统——基本设计方案与理想试验. 大气科学, 29(6): 872-884
- Barker D M, Huang W, Guo Y R, et al. 2004. A three-dimensional

- variational data assimilation system for MM5: implementation and initial results. *Mon Wea Rev*, 132(4):897-914
- Barker D M. 2005. Southern high-latitude ensemble data assimilation in the Antarctic mesoscale prediction system. *Mon Wea Rev-special section*, 133:3431-3449
- Epstein E S. 1969. Stochastic dynamic prediction. *Tellus Ser A*, 21:739-759
- Eugenia K. 2003. Atmospheric modeling, data assimilation and predictability. Cambridge University Press, 341pp
- Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J Geophys Res*, 99(C5):10143-10162
- Evensen G. 2003. The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dyn*, 53:343-367
- Fisher M, Andersson E. 2001. Developments in 4D-Var and Kalman Filtering. ECMWF Technical Memorandum No. 347:1-36
- Gaspari G, Cohn S E. 1999. Construction of correlation functions in two and three dimensions. *Quart J Roy Meteor Soc*, 125:723-757
- Houtekamer P L, Derome J. 1995. Methods for ensemble prediction. *Mon Wea Rev*, 123:2181-2195
- Houtekamer P L, Mitchell H L. 1998. Data assimilation using an ensemble Kalman filter technique. *Mon Wea Rev*, 126(3):796-811
- Houtekamer P L, Mitchell H L. 2001. A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon Wea Rev*, 129(1):123-137
- Houtekamer P L, Mitchell H L, Pellerin G, et al. 2005. Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon Wea Rev*, 133(3): 604-620
- Ingleby N B. 2001. The statistical structure of forecast errors and its representation in The Met. Office Global 3-D Variational Data Assimilation Scheme. *Quart J Roy Meteor Soc*, 127:209-231
- Lorenz A C. 1986. Analysis methods for numerical weather prediction. *Quart J Roy Meteor Soc*, 112:1177-1194
- Lorenz A C, Ballard S P, Bell R S, et al. 2000. The Met. Office global three-dimensional variational data assimilation scheme. *Quart J Roy Meteor Soc*, 126:2991-3012