

秦琰琰, 龚建东, 李泽椿. 集合卡尔曼滤波同化多普勒雷达资料的观测系统模拟试验[J]. 气象, 2012, 38(5): 513-525.

# 集合卡尔曼滤波同化多普勒雷达资料的观测系统模拟试验<sup>\* 1</sup>

秦琰琰<sup>1,2,3</sup> 龚建东<sup>4</sup> 李泽椿<sup>4</sup>

1 中国气象科学研究院, 北京 100081

2 中国科学院研究生院, 北京 100049

3 国家海洋环境预报中心, 北京 100081

4 国家气象中心, 北京 100081

**提 要:** 本文将集合卡尔曼滤波同化技术应用到对流尺度系统中, 实施了基于 WRF 模式的同化单部多普勒雷达径向风和反射率因子的观测系统模拟试验, 验证了其在对流尺度中应用的可行性和有效性, 并对同化系统的特性进行了探讨。试验表明: WRF-EnKF 雷达资料同化系统能较准确分析模式风暴的流场、热力场、微物理量场的细致特征; 几乎所有变量的预报和分析误差经过同化循环后都能显著下降, 同化分析基本上能使预报场在各层上都有所改进, 对预报场误差较大层次的更正更为显著; 约 8 个同化循环后, EnKF 能在雷达反射率、径向风观测与背景场间建立较可靠的相关关系, 使模式各变量场能被准确分析更新, 背景场误差协方差在水平方向和垂直方向都有着复杂的结构, 是高度非均匀、各项异性和流依赖的; 集合平均分析场做的确定性预报在短时间内能较好保持真值场风暴的细节结构, 但预报误差增长较快。

**关键词:** 集合卡尔曼滤波, 雷达资料同化, 背景场误差, 观测系统模拟试验(OSSE)

## Assimilation of Doppler Radar Observations with an Ensemble Kalman Filter: OSS Experiments

QIN Yanyan<sup>1,2,3</sup> GONG Jiandong<sup>4</sup> LI Zechun<sup>4</sup>

1 Chinese Academy of Meteorological Sciences, Beijing 100081

2 Graduate University of Chinese Academy of Sciences, Beijing 100049

3 National Marine Environment Forecast Centre, Beijing 100081

4 National Meteorological Centre, Beijing 100081

**Abstract:** The feasibility and availability of applying ensemble Kalman filter (EnKF) technique in convective scale systems were demonstrated by an observation system simulation experiment (OSSE) in this paper, which assimilates simulated radial velocity and reflectivity of one Doppler radar with an EnKF assimilation system based on WRF model. The experiment shows: the assimilation system has the ability to accurately analyze the detailed characters of flow fields, thermodynamic and microphysical fields of the storm, forecast and analysis errors of almost all variables decrease significantly after assimilation cycles, and forecast fields of all levels can be improved by assimilation, especially on levels of larger errors. Reliable correlativity between radar reflectivity, radial velocity observations and forecast fields can be established after 8 assimilation cycles, and the background error covariance has a complex structure and is highly inhomogeneous, anisotropic and flow-dependent. Determinate forecast of ensemble mean field can keep

\* 国家自然科学基金青年基金(41006016), 海洋公益性行业科研专项(201105018), 十二五科技支撑计划项目(2011BAC03B00), 973 计划项目(2010CB403500)和国家科技部 863 项目(2008AA09A404-2)联合资助

2011 年 5 月 26 日收稿; 2011 年 10 月 28 日收修定稿

第一作者: 秦琰琰, 主要从事大气模式资料同化的应用和研究工作. Email: qinyy@nmefc.gov.cn

the detail characters of truth storm in short period but errors grow fast.

**Key words:** ensemble Kalman filter, radar data assimilation, background error covariance, observation system simulation experiment (OSSE)

## 引 言

对流尺度天气往往造成严重的自然灾害,而目前利用数值模式对其进行准确预报还是个难点,主要原因是数值模式初始场中高分辨率探测资料的缺乏和在模式中利用该资料技术的不成熟。多普勒雷达资料具有高时空分辨率的特点,用它来改进数值模式初始场,对提高对流尺度灾害性天气预报的准确率有重要意义。

由于多普勒雷达观测数据不是常规气象要素,其在数值模式中的应用需要从探测的径向风和反射率来估计实际大气的状态,得到风、温、压、湿等模式变量场。目前同化雷达资料的方法主要有反演类算法<sup>[1-2]</sup>、热动力初始化方法<sup>[3-5]</sup>、松弛逼近方法<sup>[6]</sup>、变分类方法<sup>[7-12]</sup>,以及集合卡尔曼滤波(EnKF)方法。集合卡尔曼滤波同化方法利用一组预报集合来估计背景场误差协方差,自 Evensen<sup>[13]</sup> 1994 年首次引入到地球物理领域,EnKF 以其流依赖的背景场误差协方差、模式和观测算子可为非线性、不需要切线性伴随、易于并行等诸多优点,已经在多种大气和海洋模式中得到了广泛应用,成为目前最具业务发展前景的同化方法之一。近年来,在对流尺度系统中的应用 EnKF 方法同化雷达资料的研究取得了较大进展, Snyder 等<sup>[14]</sup> 首次将 EnKF 用于对流尺度系统实施同化雷达资料的观测系统模拟试验,他们采用 Sun 等<sup>[15]</sup> 发展的暖云模式同化了模拟的单多普勒雷达径向风观测,初步验证了 EnKF 在对流尺度的同化效果。Dowell 等<sup>[16]</sup> 在相同的暖云模式中尝试了真实的单多普勒雷达反射率和径向风资料的同化,并将结果与观测做了对比,结果表明风场的反演效果较好,温度场由于模式误差和观测的限制没有得到很好的反演。许小永等<sup>[17]</sup> 也采用此暖云模式同化了模拟的雷达观测和一次梅雨锋暴雨过程的双多普勒雷达资料,试验表明 EnKF 同化技术能够从雷达资料较准确地反演暴雨中尺度系统的动力场、热力场和微物理量场。Tong 等<sup>[18-19]</sup> 将此领域的研究推进到包含冰相微物理过程的同化,他们在 ARPS 模式中同化了模拟的雷达反射率和径向风,

得到了对风暴流场、热力场和各种微物理量场的准确分析。Zhang 等<sup>[20]</sup> 将 EnKF 同化系统应用于实际雷达资料同化,针对一次台风过程在 WRF 模式中同化了径向风观测,并进行了确定性预报和集合预报,试验表明,径向风同化能较好地分析出台风路径和强度,并有能力预报台风的快速形成和突然增强。

以上研究初步显示了 EnKF 同化方法在对流尺度的应用前景,但其针对雷达资料的同化技术不够成熟,目前国内这方面的研究也比较少,还需要开展进一步的研究工作。本文基于 WRF 模式,利用 EnKF 系统进行了同化雷达资料的观测系统模拟试验,并结合试验结果对 EnKF 雷达资料同化系统的性质进行了探讨和分析,以期获得对此同化系统更深入的认识,并为多普勒雷达资料在数值模式中的同化应用提供一些借鉴和科学依据。

## 1 同化系统及试验设计

### 1.1 预报模式

预报模式采用完全可压的欧拉非静力中尺度 WRF 模式,包含 6 个预报方程和一个静力诊断方程,预报变量为三维风场、扰动位温、扰动位势、干空气地面扰动气压,以及水汽、云水、雨水、冰晶、雪和霰的混合比等可选择的预报变量。

试验采用 WRF 模式中超级单体理想试验方案,积分区域格点数为  $42 \times 42 \times 41$ ,水平格距 2 km,垂直方向 0.5 km,雷达天线位于(0,0,0),即模式区域底部的西南角。积分时间步长为 12 秒,微物理过程选用 WSM6 方案,包含了 5 种水凝物的混合比,使用开放式侧边界条件。

### 1.2 同化分析系统

#### 1.2.1 集合卡尔曼滤波分析

WRF-EnKF 同化系统由奥克拉荷马大学风暴分析与预报中心提供,与 Tong 等<sup>[18]</sup> 相似,不同之处为本文的同化系统基于不扰动观测的集合方根卡尔曼滤波方法,即 EnSRF 方法,分析场为平均分析场  $\bar{x}^a$  与扰动分析场  $x'^a$  的和:

$$x_k^a = \bar{x}^a + x_k^{\prime a} \quad (1)$$

$$\bar{x}^a = \bar{x}^f + K(y^o - H\bar{x}^f) \quad (2)$$

$$x_k^{\prime a} = \gamma(I - \mu KH)x_k^{\prime f} \quad (3)$$

$$K = P^f H^T (HP_k^f H^T + R_k)^{-1} \quad (4)$$

式中,上标  $a$  和  $f$  分别表示分析场和预报场,下标  $k$  表示第  $k$  个同化时步,  $K$  为卡尔曼增益矩阵,  $y^o$  为同化的观测,  $P^f$  为背景场误差协方差,  $R$  为观测误差协方差,  $H$  为观测算子,  $\gamma$  为协方差扩大因子,一般略大于 1。

$$\mu = \left(1 + \sqrt{\frac{R}{HP^f H^T + R}}\right)^{-1} \quad (5)$$

在观测误差独立且与背景场误差不相关的假设下,同化观测的先后顺序不影响同化结果,逐个顺序同化观测,这时  $HP^f H^T$  和  $R$  都成为标量。

背景场误差协方差由预报集合估计得到,  $k$  时步的背景场误差协方差矩阵通过以下式子得到:

$$P^f H^T = \frac{1}{N-1} \sum_{k=1}^N (x_k^f - \bar{x}^f) [H(x_k^f) - \overline{H(x^f)}]^T \quad (6)$$

$$HP^f H^T = \frac{1}{N-1} \sum_{k=1}^N [H(x_k^f) - \overline{H(x^f)}] [H(x_k^f) - \overline{H(x^f)}]^T \quad (7)$$

式中,  $N$  为集合成员数。

### 1.2.2 观测算子

同化的观测包括雷达径向风和反射率因子,试验中假设观测误差间不相关,观测误差满足高斯分布。EnKF 中允许非线性观测算子的直接使用而无需线性化,本文雷达反射率的观测算子是非线性的。

雷达径向风的观测算子为:

$$V_r = u \cos \alpha \sin \beta + v \cos \alpha \cos \beta + w \sin \alpha \quad (8)$$

式中,  $V_r$  为雷达径向风观测,  $\alpha$  为雷达波束的仰角,  $\beta$  为雷达波束的方位角,  $u, v, w$  为模式输出的风分量。

雷达反射率因子的观测算子为:

$$Z = 10 \log_{10} \left( \frac{Z_e}{1 \text{ mm}^6 \cdot \text{m}^{-3}} \right) \quad (9)$$

式中,  $Z$  和  $Z_e$  分别是以 dBz 和  $\text{mm}^6 \cdot \text{m}^{-3}$  为单位的雷达反射率因子,

$$Z(\text{dBz}) = 10 \log_{10} \left( \frac{Z_e \text{ mm}^6 \cdot \text{m}^{-3}}{1 \text{ mm}^6 \cdot \text{m}^{-3}} \right) \quad (10)$$

总的雷达反射率因子由雨、雪、雹的反射率因子  $Z_{er}$ 、 $Z_{es}$ 、 $Z_{eh}$  加和得到。

$$Z_e = Z_{er} + Z_{es} + Z_{eh} \quad (11)$$

各水凝物反射率因子的计算基于 Smith 等<sup>[21]</sup>和 Lin 等<sup>[22]</sup>的算法求得,其中,雨滴的反射率因子为:

$$Z_{er} = \frac{10^{18} \times 720 (\rho q_r)^{1.75}}{\pi^{1.75} N_r^{0.75} \rho_r^{1.75}} \quad (12)$$

温度低于  $0^\circ\text{C}$  时,干雪的反射率因子为:

$$Z_{es} = \frac{10^{18} \times 720 K_i^2 \times \rho_s^{0.25} (\rho q_s)^{1.75}}{\pi^{1.75} K_r^2 N_r^{0.75} \rho_i^2} \quad (13)$$

温度高于  $0^\circ\text{C}$  时,湿雪的反射率因子为:

$$Z_{es} = \frac{10^{18} \times 720 (\rho q_s)^{1.75}}{\pi^{1.75} N_s^{0.75} \rho_s^{1.75}} \quad (14)$$

雹的反射率因子:

$$Z_{eh} = \left( \frac{10^{18} \times 720}{\pi^{1.75} N_h^{0.75} \rho_h^{1.75}} \right)^{0.95} (\rho q_h)^{1.6625} \quad (15)$$

式中,  $\rho$  为空气密度,  $\rho_r = 1000 \text{ kg} \cdot \text{m}^{-3}$ , 为雨水密度,  $\rho_s = 100 \text{ kg} \cdot \text{m}^{-3}$ , 为雪的密度,  $\rho_i = 917 \text{ kg} \cdot \text{m}^{-3}$ , 为冰晶的密度,  $\rho_h = 913 \text{ kg} \cdot \text{m}^{-3}$ , 为雹的密度;  $q_r, q_s, q_h$  分别为雨、雪、雹的比含水量;  $N_r = 8.0 \times 10^{-6} \text{ m}^{-4}$ ,  $N_s = 3.0 \times 10^{-6} \text{ m}^{-4}$ ,  $N_h = 4.0 \times 10^{-6} \cdot \text{m}^{-4}$  分别为雨、雪、雹的 Marshall-Palmer 指数型滴谱分布的截断参数,  $K_i^2 = 0.176$ 、 $K_r^2 = 0.93$  分别为冰晶和雨滴的介电系数。

### 1.2.3 协方差局地化

由有限集合样本数引入的样本误差会导致背景场误差协方差的估计误差,尤其会在观测较远处产生虚假的分析增量。Houtekamer 等<sup>[23]</sup>提出了 Schur 乘积方法解决这个问题,使用 Schur 乘积后卡尔曼增益变为:

$$K = [\rho \circ (P^f H^T)] [\rho \circ (HP^f H^T) + R]^{-1} \quad (16)$$

系统中采用了此算法,并使用 Gaspari 等<sup>[24]</sup>的 5 阶分段有理函数来定义局地相关函数。5 阶分段有理函数的形式为:

$$\rho(r, c) = \begin{cases} -\frac{1}{4} \left(\frac{r}{c}\right)^5 + \frac{1}{2} \left(\frac{r}{c}\right)^4 + \frac{5}{8} \left(\frac{r}{c}\right)^3 - \frac{5}{3} \left(\frac{r}{c}\right)^2 + 1 & 0 \leq r \leq c \\ \frac{1}{12} \left(\frac{r}{c}\right)^5 - \frac{1}{2} \left(\frac{r}{c}\right)^4 + \frac{5}{8} \left(\frac{r}{c}\right)^3 + \frac{5}{3} \left(\frac{r}{c}\right)^2 - 5 \left(\frac{r}{c}\right) + 4 - \frac{2}{3} \left(\frac{r}{c}\right)^{-1} & c \leq r \leq 2c \\ 0 & 2c \leq r \end{cases} \quad (17)$$

式中,  $r$  为观测与背景场上分析格点之间的距离,  $c$  为给定的长度尺度参数。

$$r = \sqrt{r_x^2 + r_y^2 + r_z^2} \quad (18)$$

#### 1.2.4 协方差调整

集合的离散度应该代表集合平均场的误差统计特征,然而有限集合数带来的样本误差会造成集合离散度与集合平均场误差的不一致,从而导致滤波发散。离散度过小会导致 EnKF 分析不吸收观测,而离散度过大则会使分析过度偏离背景场。集合离散度与集合平均场误差在统计上满足如下一致性关系<sup>[25]</sup>:

$$\frac{E\left[\frac{1}{N_e - 1} \sum_{n=1}^{N_e} |x_n^f - \hat{x}^f|^2\right]}{E(|x^f - \hat{x}^f|^2)} = \frac{N_e}{N_e + 1} \quad (19)$$

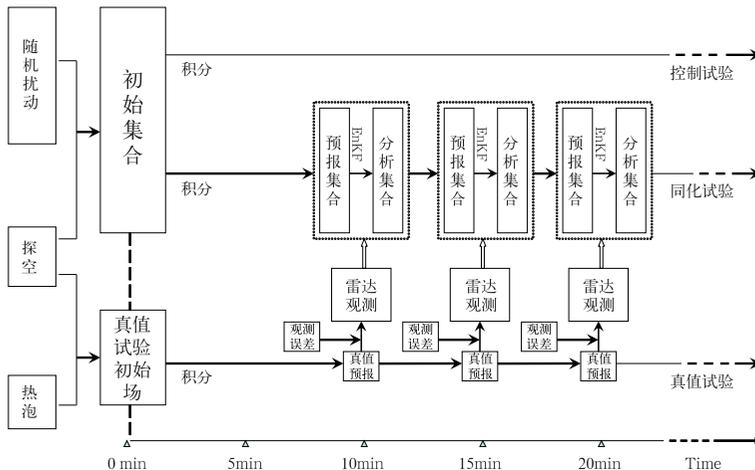


图 1 WRF-EnKF 雷达资料同化试验流程示意图

Fig. 1 Flow chart of radar data assimilation experiment based on WRF-EnKF

激发对流,热泡中心在  $x = 41 \text{ km}$ ,  $y = 41 \text{ km}$ ,  $z = 1.5 \text{ km}$  处,其水平影响半径为  $10 \text{ km}$ ,垂直影响半径为  $1.5 \text{ km}$ ,热泡以外  $\theta$  为水平均匀场。

真值实验探空如图 2 所示,取自于 WRF 模式超级单体理想试验,其 CAPE 值为  $2301 \text{ J} \cdot \text{kg}^{-1}$ , CIN 值为  $16 \text{ J}$ , LI 指数为  $-8$ ,为避免右移风暴离开模拟区域中心,将探空的纬向风减去了  $6 \text{ m} \cdot \text{s}^{-1}$ 。

真值模拟中,液相水凝物约于  $5 \sim 10 \text{ min}$  时在  $900 \sim 500 \text{ hPa}$  开始生成,冰相水凝物约于  $10 \sim 15 \text{ min}$  时在  $650 \sim 300 \text{ hPa}$  生成,垂直上升气流在  $70 \text{ min}$  达  $47 \text{ m} \cdot \text{s}^{-1}$ 。约  $40 \text{ min}$  单体开始分裂,分别向南北方向移动,北边的单体在  $90 \text{ min}$  左右分裂为两个,之后不断分裂减弱,南边的单体逐渐成为主导单体,于  $145 \text{ min}$  后逐渐移出模拟区域。

式中  $N_e$  为集合成员数。协方差调节是解决滤波发散问题的有效方法,即式(3)中给扰动分析场乘以一个扩大因子  $\beta$  来调节集合成员的发散度,从而缓解滤波发散问题。

### 1.3 试验设计

设计了真值模拟试验、控制试验和同化试验三组试验,试验流程见图 1。

#### 1.3.1 真值试验

假设模式完美,在不考虑模式误差的影响下可将模拟结果当成大气真实状态,与同化试验结果进行对比以检验同化效果。真值试验积分开始时刻在探空环境中植入一高于环境场  $3 \text{ K}$  的椭圆热泡以

#### 1.3.2 同化试验

同化试验的初始集合由环境场叠加初始扰动得到,环境场为由探空资料提供的水平均匀场,初始场各层  $w = 0$ 。初始扰动采用高斯分布的随机扰动,  $u$ 、 $v$ 、 $w$  的均方差都取为  $3 \text{ m} \cdot \text{s}^{-1}$ ,  $\theta$  的均方差取  $1 \text{ K}$ 。集合成员取 20 个,从  $t = 0 \text{ min}$  开始积分,  $10 \text{ min}$  后进行第一次集合卡尔曼滤波分析,之后每  $5 \text{ min}$  进行一次同化分析,共做 20 个同化循环。同化的观测为模拟的雷达径向风和雷达反射率因子,径向风和反射率因子观测由下式得到:

$$V_r = V_r' + \epsilon_{V_r} \quad (20)$$

$$Z = Z' + \epsilon_Z \quad (21)$$

其中,  $V_r'$  和  $Z'$  分别为雷达径向风和反射率观测的真值,从观测算子计算得到,  $\epsilon_{V_r}$  和  $\epsilon_Z$  分别为雷达径

向风和反射率的观测误差,假设观测误差分布为高斯型,其均方差分别为  $1 \text{ m} \cdot \text{s}^{-1}$  和  $5 \text{ dB}$ 。分析变量为三维风场、扰动位温、水汽混合比以及水汽、云水、雨水、云冰、雪、雹 5 种水凝物的混合比。

### 1.3.3 控制试验

为与同化试验结果做对比,设计了不进行同化的集合预报作为控制试验,初始集合与同化试验一样,集合成员也取 20 个,从  $t=0 \text{ min}$  开始做集合预报,预报 120 min,不同化雷达观测。

## 2 试验结果

### 2.1 对集合平均分析场的分析能力

图 3 将 5 km 高度上 20、60 和 100 min 的同化

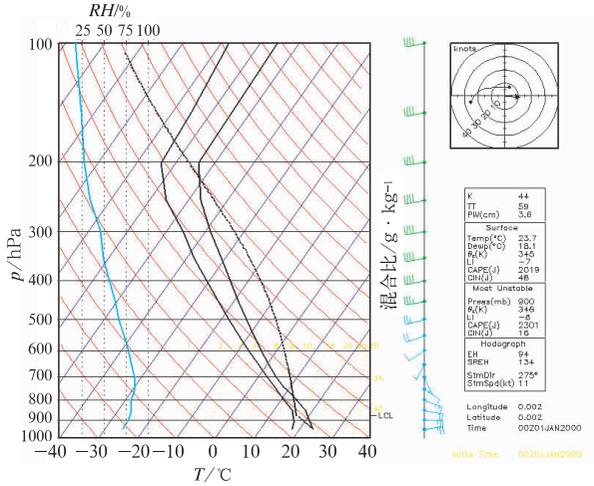


图 2 探空斜温图

Fig. 2 Skew T diagram for the environmental sounding

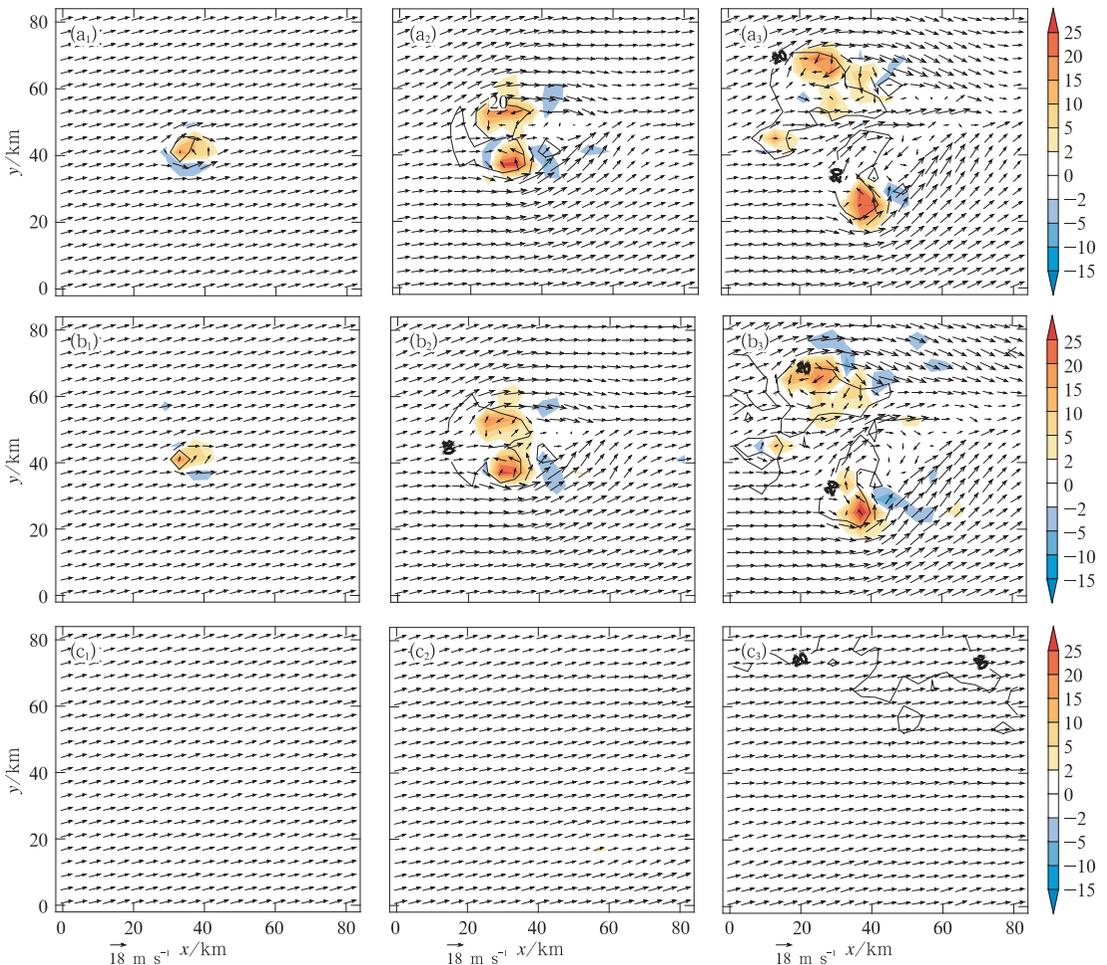


图 3 5 km 高度上真值场(第一行, a)、同化试验集合平均分析场(第二行, b)和控制试验集合平均场(第三行, c)在 20 ( $a_1, b_1, c_1$ ), 60 ( $a_2, b_2, c_2$ ), 100 ( $a_3, b_3, c_3$ ) min 时垂直速度场(阴影区)、水平速度场(矢量图)和扰动位温的分布 Fig. 3 Vertical velocity (shaded), horizontal wind (vector) and perturbed potential temperature (contours) at 5 km in the truth simulation (top row), the EnKF analysis (ensemble mean, middle row) and the control run (ensemble mean, bottom row) which are shown at  $t=20 \text{ min}$  (left panel), 60 min (middle panel), 100 min (right panel)

后集合平均分析场与真值场及控制试验集合平均场做了比较。由图可见,风暴附近的 中层水平风分析场在 3 个同化循环后就能较准确地反映真值场的流型,垂直上升、下沉运动的位置与真值场吻合,但强度上稍微偏弱,风暴附近扰动位温的高值区也与真值吻合较好。60 min,垂直上升运动的强度仍稍偏弱,但水平流场和扰动位温都与真值场非常一致。11 个同化循环后,流场和扰动位温场与真值场基本完全吻合。而控制试验中,风暴一直没发展起来,各变量场的大小和空间分布从开始到结束几乎没有变化,与真值场相差甚远。

对真值模拟、同化试验和控制试验沿风暴中心剖面的水汽扰动和各水凝物混合比做了比较,可见在 3 个同化循环后,水汽和微物理量场的位置与真值场吻合得很好,量上略微偏弱。5 个同化循环后 EnKF 对水汽和微物理量场的细致特征也能准确刻画。

下面通过误差分析来检验同化分析对热力场、流场、水凝物场的分析和反演能力,采用均方根误差作为指标进行分析。集合平均预报误差  $\text{rms}_f$  和集合平均分析误差  $\text{rms}_a$  由式(22)和(23)计算:

$$\text{rms}_f = \sqrt{\frac{1}{N} \sum_{i=1}^N (\bar{X}_i^f - X^t)^2} \quad (22)$$

$$\text{rms}_a = \sqrt{\frac{1}{N} \sum_{i=1}^N (\bar{X}_i^a - X^t)^2} \quad (23)$$

$N$  为全场或雨区的总格点数,  $\bar{X}_i^f$  为集合平均预报场,  $\bar{X}_i^a$  为集合平均分析场,  $X^t$  为真值场。

图 6 给出了控制试验预报误差和同化试验的集合平均场各主要变量的均方根误差分布。由图可见,控制试验预报误差增长很快,误差值也很大,而经过同化循环后,所有变量的预报误差和分析误差都比控制试验显著减小,并且基本上在每次同化分析之后,都比预报场误差有所下降,说明同化分析对预报有显著改进。

风场预报误差在头两个循环迅速增大,20~55 min 预报和分析误差增长缓慢,之后误差略有下降并在震荡中逐渐收敛, $u$ 、 $v$ 、 $w$  的分析误差分别收敛至约 2.5、2 和 1.5  $\text{m} \cdot \text{s}^{-1}$  左右。流场在第一个循环预报集合误差很小,主要是由于同化前初始集合经过了一段时间积分,使得预报场与真值场差别不大。温度扰动场( $p_t$ )、云水( $q_r$ )、云冰( $q_i$ )等水凝物也有误差在开始两三个循环增加,之后趋于平稳,最后逐渐收敛的特点。

同化初期各变量场预报误差迅速增加,主要是由于这段时间风暴发展迅速,流场、热力场及水凝物

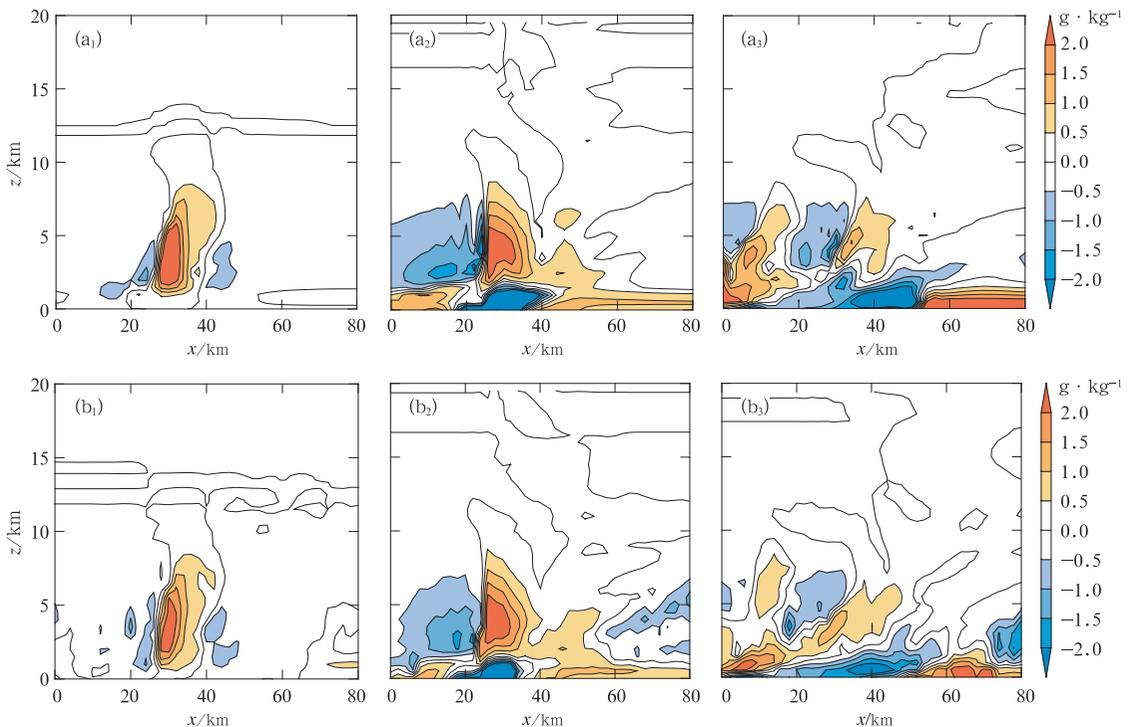


图 4 真值场(a)和同化试验集合平均分析场(b)的扰动水汽混合比  $q'_v$  在 20(a<sub>1</sub>, b<sub>1</sub>)、60(a<sub>2</sub>, b<sub>2</sub>) 和 100(a<sub>3</sub>, b<sub>3</sub>) min 沿  $y=41$  km 的  $x-z$  剖面图

Fig. 4 Perturbation  $q'_v$  of truth simulation (top row) and EnKF analysis (bottom row) in the  $x-z$  plane at  $y=41$  km at 20 min (left panel), 60 min (middle panel), 100 min (right panel)

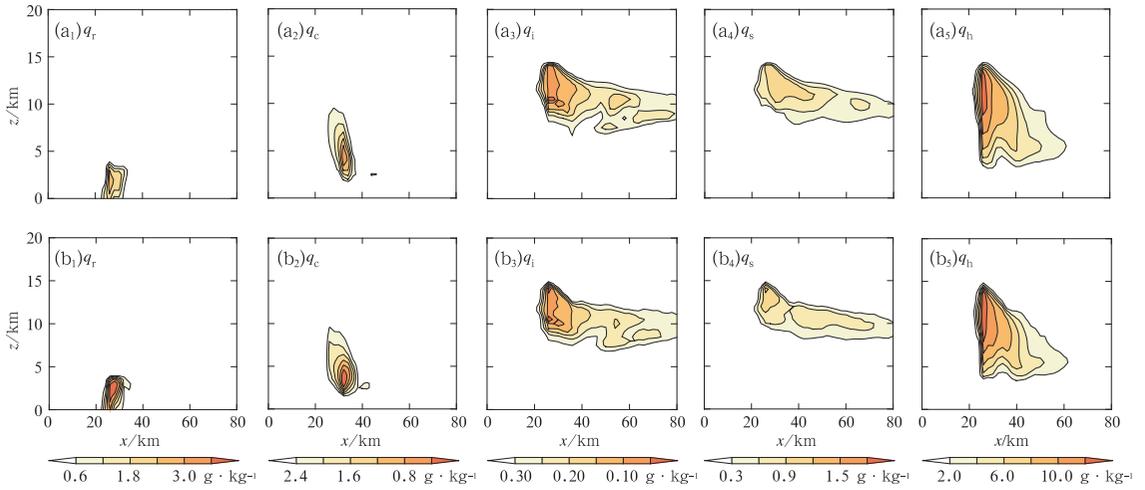


图 5 真值场(a)和同化试验(b)的集合平均分析场中各水凝物(从左至右依次为雨水  $q_r$ 、云水  $q_c$ 、冰晶  $q_i$ 、雪  $q_s$  及雹  $q_h$  的混合比)在 60 min 沿  $y=41$  km 的  $x-z$  剖面图

Fig. 5  $q_r, q_c, q_i, q_s$  and  $q_h$  of truth simulation (a) and EnKF analysis (b, ensemble mean) in the  $x-z$  plane at  $y=41$  km at 60 min

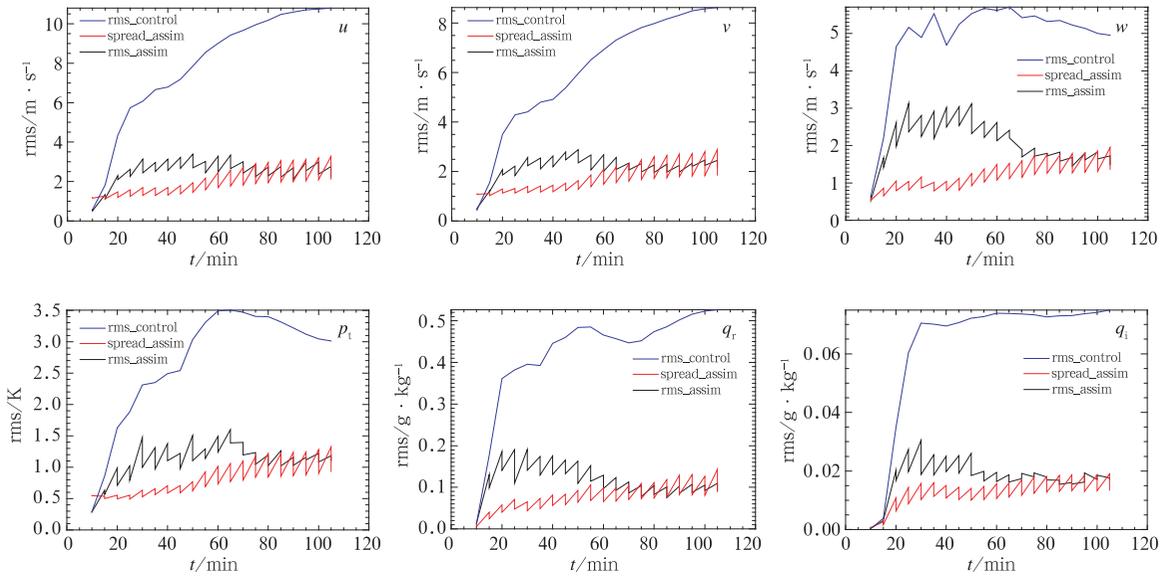


图 6 控制试验预报误差和同化试验的集合平均场各变量的均方根误差(包括预报误差和分析误差),以及同化试验中的集合成员发散度

(蓝线表示控制试验误差,黑线表示同化试验误差,红线表示同化试验中成员的集合发散度)

Fig. 6 The rms errors of variables ( $u, v, w, p_1, q_r, q_i$ ) of ensemble mean forecast and analysis from EnKF (black) and of control run (blue), ensemble spread of EnKF (red)

场的迅速变化使得同化后分析场的强度跟不上风暴的发展而造成;同化初期分析场的误差下降不明显主要是因为同化开始阶段观测和模式变量间合理的协相关关系还没有建立起来,此时分析还不能准确更正模式变量的预报误差,但几个循环之后,分析对预报场的更正明显增强,说明 3~4 个同化循环之后,观测与模式变量场之间的协相关关系开始趋于合理,观测信息逐渐能对预报场进行准确的调整,60

min 之后,可靠和协调的背景场误差协方差基本建立起来,误差逐渐收敛。如图 7 所示,60~70 min 之后集合离散度/集合平均场误差比率在  $N_e/(N_e + 1)$  附近振荡[参考本文式(19)],这时集合的发散度对集合平均场的误差统计特征有很好的代表性,EnKF 分析在这时准确可靠。

从图 8 各变量误差的垂直分布可见,误差值一般在该变量发展较强的高度较大,在同化初始时刻

误差大值区的高度较低,值也较小,随着对流风暴逐渐向高层发展,误差大值区逐渐向上移动,值也迅速变大,各变量误差大值高度基本代表了该变量发展最旺盛的高度。60 min 时流场误差最大值高度发

展到约 15 km;水汽在低层较大,其误差大值区始终在 10 km 以下,45 min 后误差的最大值一直在 1 km 以下;水凝物的误差只在含有这种水凝物的高度上有值,其他层次上误差值为零。

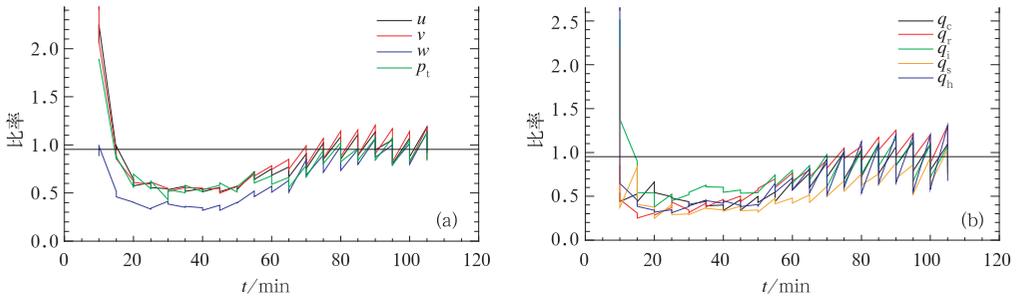


图 7 集合离散度与集合平均场误差的比率 [黑色直线为  $N_e/(N_e+1)$ ]

Fig. 7 Ratio of the ensemble variance to the squared error of the ensemble mean, the horizontal line indicates  $N_e/(N_e+1)$

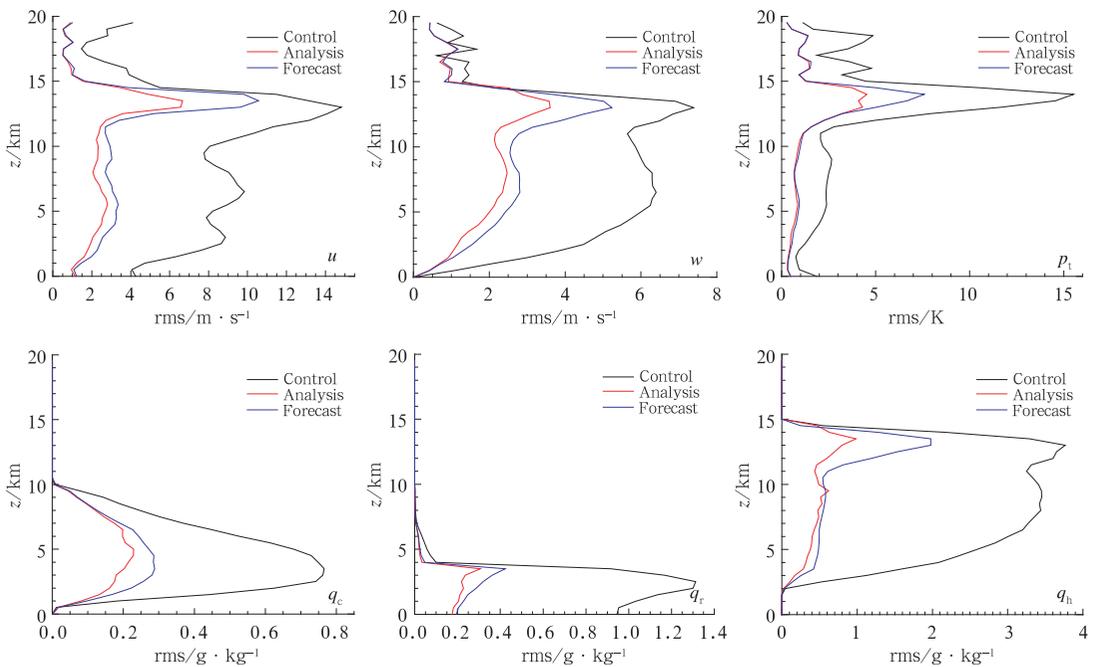


图 8 控制试验预报误差和同化试验集合平均场各变量在 60 min 的水平均方根误差的垂直廓线 (蓝线、红线、黑线分别表示同化试验预报误差、分析误差和控制试验预报误差)

Fig. 8 Vertical profiles of rms errors of EnKF forecast (blue), EnKF analysis (red) and control run (black) at 60 min

综合各变量来看,在误差较大的层次,分析对预报场的更正非常明显,在误差较小的高层和低层分析对预报的更正很小或几乎为零。由图也可见,同化试验误差始终小于控制试验误差,分析对预报场的改进在同化的前期和中期比较显著,在误差较大的层次比较显著。而控制试验的误差从高层到低层都大于同化试验的预报误差,在误差较大的层次,控

制试验的误差更大。

## 2.2 背景场误差协方差分析

由预报集合估计的流依赖的背景场误差协方差在 EnKF 的变量分析中扮演着非常重要的角色。背景场误差协方差包含两部分内容:相关和方差,方差大值区一般对应于系统发展剧烈的区域,相关性

数则表示了某点上某变量的增量对其他点某变量变化的影响程度。图 9 和 10 为 60 min 时点(16, 19, 19)(为上升运动强中心, 图中粗黑点标出)与  $z=19$  水平面及  $y=19$  剖面上各点上流场, 温度场, 水汽场、水凝物场同变量间的背景场误差的水平(图 9)

和垂直(图 10)相关系数。

60 min 时, 预报场各变量的背景场误相关系数不论在水平方向还是垂直方向都与三维变分方法假定的背景场误差协方差高斯型分布有些相似, 基本上有与自身相关系数为 1, 离目标点越远相关越小

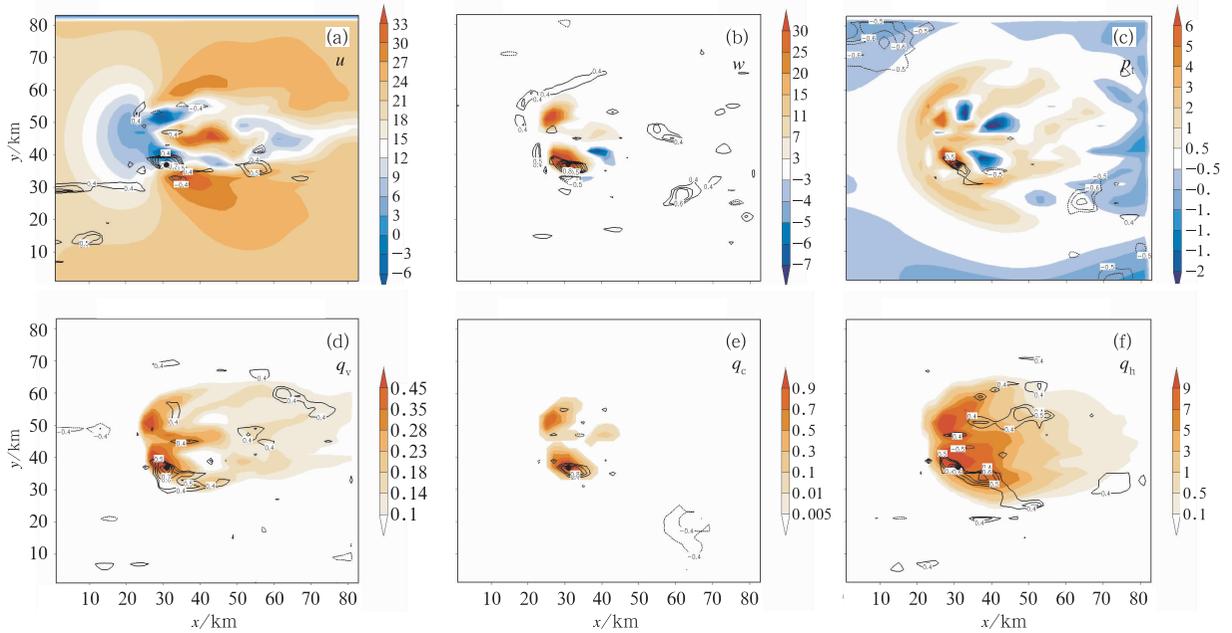


图 9 60 min 时点(16, 19, 19)与  $z=19$  (9 km) 水平面上各点同变量(依次为:  $u$ 、 $w$ 、 $p_t$ 、 $q_v$ 、 $q_c$ 、 $q_h$ ) 间的背景场误差水平相关系数, 填色图为此变量的真值场

Fig. 9 Background error correlations estimated from an ensemble at 60 min in the  $x$ - $y$  plane along  $z=9$  km

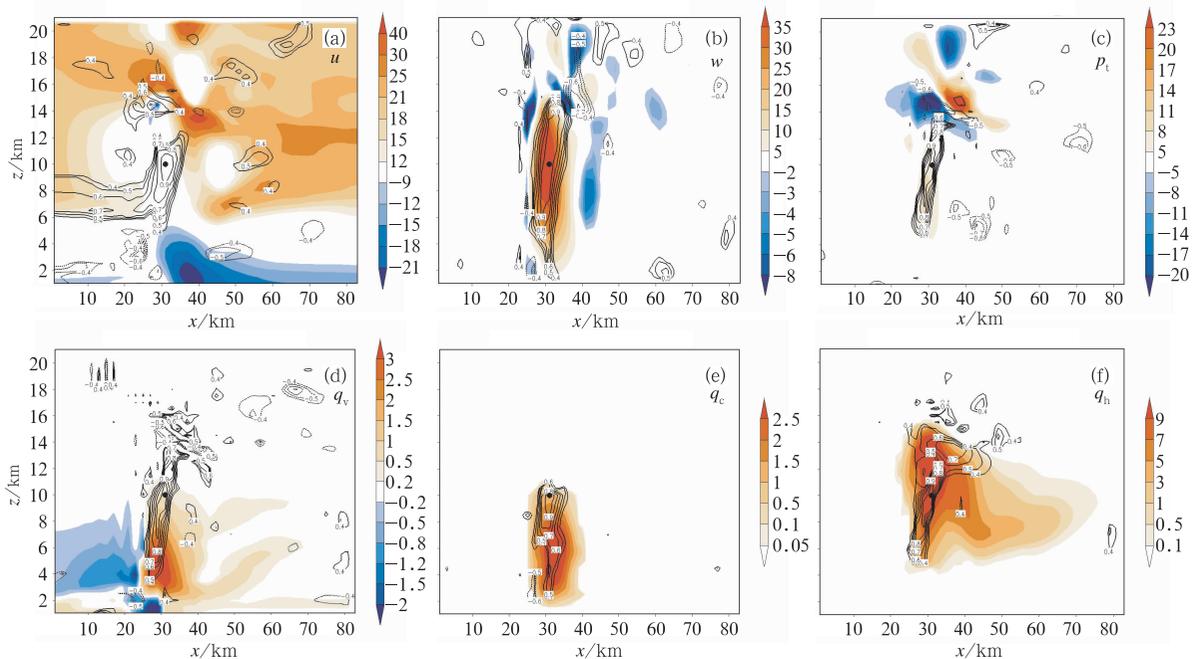


图 10 60 min 时点(16, 19, 19)与  $y=19$  (37 km) 剖面上各点同变量(依次为:  $u$ 、 $w$ 、 $p_t$ 、 $q_v$ 、 $q_c$ 、 $q_h$ ) 间的背景场误差垂直相关系数, 填色图为此变量的真值场

Fig. 10 Background error correlations estimated from an ensemble at 60 min in the  $x$ - $z$  plane along  $y=37$  km

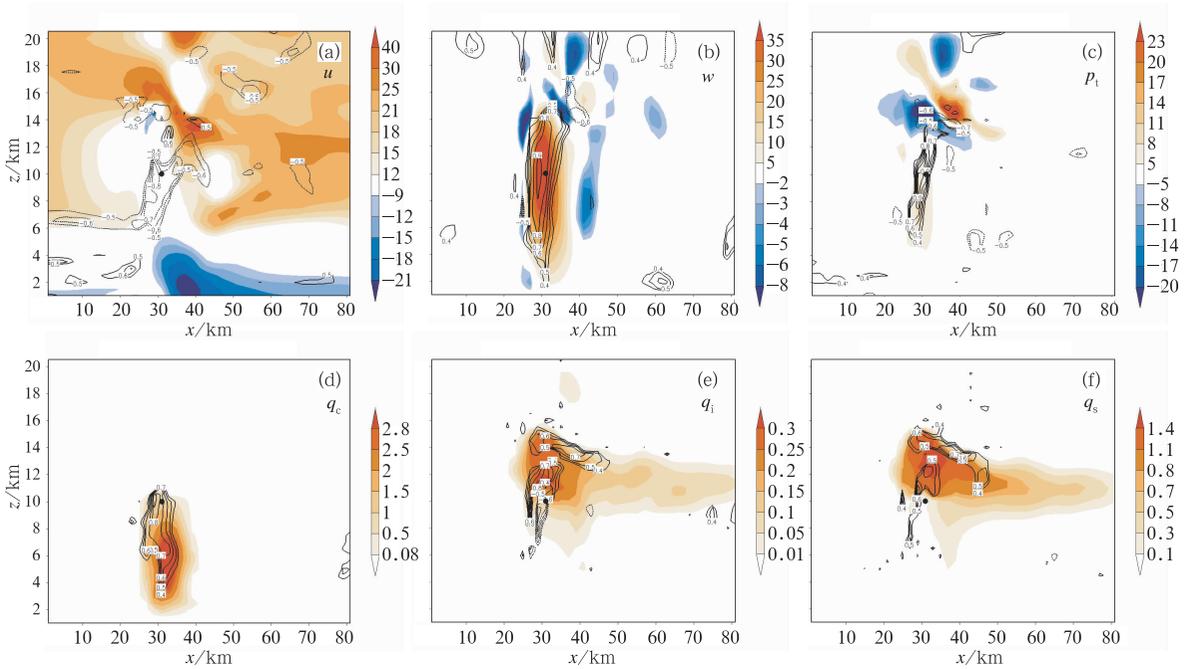


图 11 60 min 时预报集合的雷达反射率因子与沿  $y=19(37\text{ km})$  剖面的各分析变量的背景场误差相关图 (等值线, 实线为正相关, 虚线为负相关), 粗黑点为反射率观测所在位置, 填色图为此时的真值场各变量, 变量依次为:  $u$ 、 $w$ 、 $p_t$ 、 $q_c$ 、 $q_i$ 、 $q_s$

Fig. 11 Background error correlations estimated from an ensemble at 60 min in the  $x-z$  plane along  $y=37\text{ km}$  between reflectivity (indicated by a black dot) and  $u$ ,  $w$ ,  $p_t$ ,  $q_c$ ,  $q_i$ ,  $q_s$ , respectively, thick solid (dashed) contours represent positive (negative) correlations

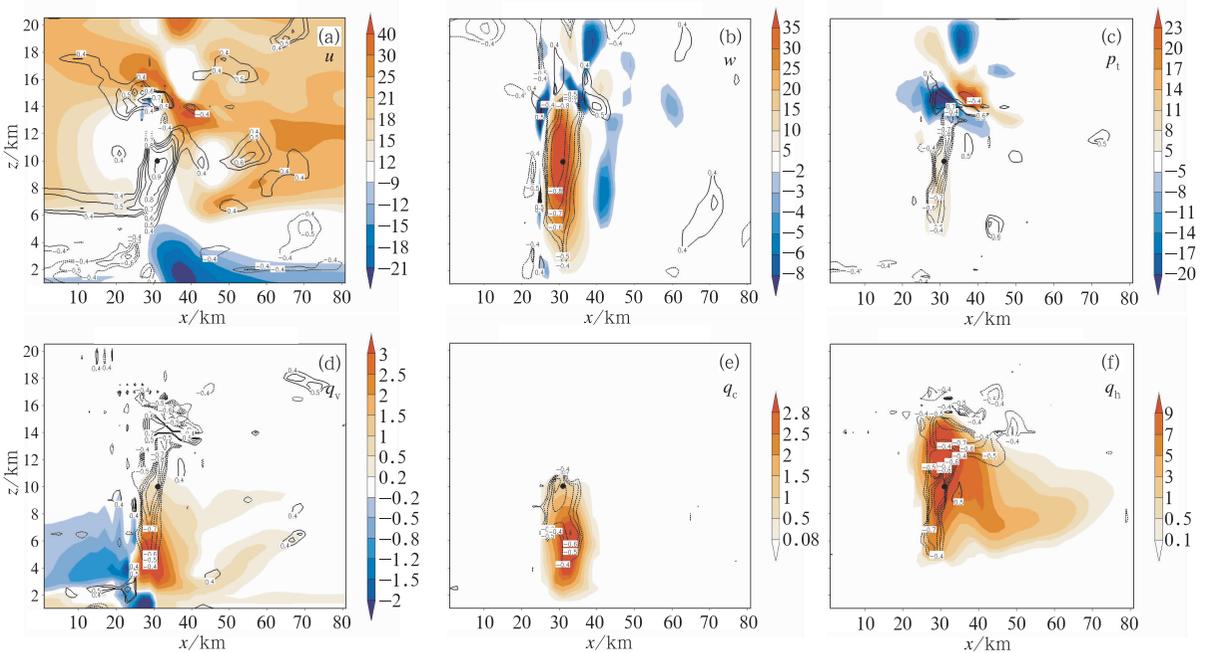


图 12 60 min 时预报集合的雷达径向风与沿  $y=19(37\text{ km})$  剖面的各分析变量的背景场误差相关图 (等值线, 实线为正相关, 虚线为负相关), 粗黑点为反射率观测所在位置 (16, 19, 19), 填色图为此时的真值场各变量 (各变量依次为:  $u$ 、 $w$ 、 $p_t$ 、 $q_v$ 、 $q_c$ 、 $q_h$ )

Fig. 12 Background error correlations estimated from an ensemble at 60 min in the  $x-z$  plane along  $y=37\text{ km}$  between radial velocity (indicated by a black dot) and  $u$ ,  $w$ ,  $p_t$ ,  $q_v$ ,  $q_c$ ,  $q_h$ , respectively, thick solid (dashed) contours represent positive (negative) correlations

的趋势。同时,由图清楚可见背景场误差相关系数有着复杂的结构,与三维变分方法所假定的同心圆结构不同,它是高度非均匀、各项异性和流依赖的。

集合卡尔曼滤波雷达资料同化分析的准确性主要依赖于非模式变量的径向风、反射率观测与模式变量间正确协相关关系的建立。图 11 和 12 分别给出了南边单体上升气流中心所在位置,即点(16,19,19)上的背景场反射率观测和径向风观测与  $y=19$  剖面上各点各变量的背景场误差相关系数。

由图 11 可见,60 min 时背景场的反射率观测与各模式变量建立的相关关系较好地反映了真值场各变量的结构,结合此时的真值场可以得到,与 5 种水凝物场和水汽场的相关,正相关大值区很好地对应着此时真值场此变量的大值区域,这是合理的,反射率增量增加了水凝物和水汽,应该对应于对水凝物分析和水汽分析的增加;对于与垂直速度的相关,正相关的高值区较好地对应了垂直上升运动的大值区域,它反映了垂直上升运动带来水汽凝结而对应的反射率增大;对于扰动位势,正相关的高值区对应于真值场扰动位势的大值区,强的上升运动可以带来扰动位势的增大,而强上升运动也对应着水汽的凝结和大的反射率值。由图可以看到,  $q_c$ 、 $q_h$ 、 $q_v$ 、 $\theta$ 、 $\omega$  等变量与雷达反射率是高度相关的,其与反射率相关的最大值都超过了 0.9。

由图 12 可见,60 min 时点(16,19,19)上的径向风观测与各模式变量间也建立了可靠的相关关系。其中与水平风  $u$  和  $v$  有很强的正相关( $v$  分量图略),其最大相关值分别超过了 0.9 和 0.8,并且呈近似高斯型的分布,与  $u$  风的强相关区位于上升气流区所对应的弱水平气流区;径向风与  $\omega$  在上升气流中心有强的负相关,径向风与上升气流区所对

应的正扰动位温区域有较强的负相关,而与上升气流顶部的负扰动位温区有较强的正相关。径向风与扰动位势在扰动位势的高值区有较强的负相关。径向风与 5 种水凝物和扰动水汽场也都有较强的负相关,其负相关的高值区都对应着这些物理量的大值区。

反射率、径向风观测与各变量的相关图也很好反映了观测与各模式各变量间相关关系的复杂结构,它是高度非均匀、各项异性和流依赖的。

### 3 集合平均分析场做的确定性预报

为分析同化对其后确定性预报的影响,这一节使用 40、50 和 60 min 的集合平均分析场做确定性预报,都报到 160 min,检验其预报误差增长情况。

图 13 给出了用 50 min 的集合平均分析场做的确定性预报与真值场在不同时刻的流场和总水凝物场的变化图。用集合平均分析场做的前 20 多分钟预报较好地保持了真值场风暴的细节结构,包括水平流场的流型、大小,总水凝物的强度、位置,垂直上升运动的结构和位置。75 min 开始北边单体上升气流的强度稍微偏弱,总水凝物场的结构与真实风暴也发生了一些差异,尤其在中层,北边单体总水凝物场的强度偏大,南边单体总水凝物的位置较真实风暴偏北,量也偏大,两单体间的水平气流也较真值场偏弱。90 min 时,平均分析场的预报场北边的单体在中层(约 7 km)开始分裂,并产生其后方大片的虚假水凝物,随后此虚假单体发展迅速,至 110 min 时在高层(约 10 km)与原单体强度相当,而原单体不断减弱,至 120 min 基本将原单体取代。150 min 时,真实单体已逐渐开始移出模拟区域,平均分析场

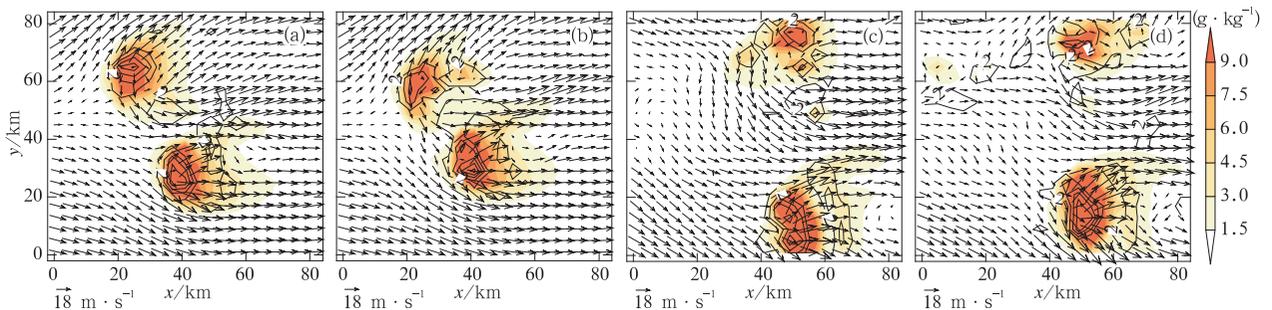


图 13 50 min 的集合平均分析场做的确定性预报(b,d)与真值场(a,c)在 10 km 高度上 100 min(a,b)和 150 min(c,d)的水平风场(矢量)、垂直速度(等值线)、总水凝物场(填色)

Fig. 13 Horizontal wind (vector), vertical velocity (contours) and total hydrometeor (shading) of 100 min (a, b) and 150 min (c, d) at 10 km level for the truth simulation (a, c) and the forecast (b, d) beginning from the ensemble mean analysis at 50 min

预报的单体向南的移速偏慢,直到 160 min 左右才逐渐移出模拟区域。

从上面的分析可见,预报误差主要来自于虚假的分裂单体,单体的传播速度与真实单体不一致也是误差的来源之一。在开始的 10~20 min,预报单体基本能保持实际单体的结构、位置和强度,但预报误差增长很快,预报单体很快与真实单体出现较大差异。预报误差的迅速增长,一方面可能与高度非线性的对流风暴系统发展迅速、预报时效短有关,小误差经过模式积分后误差的非线性增长非常迅速,

另一方面也与作为预报初始场的集合平均分析场的误差有关系。

图 14 给出了用 40、50 和 60 min 的集合平均分析场做确定性预报,报到 160 min 时各变量的预报误差。由图可见,预报误差随时间的增加都逐渐增长,垂直速度,云、雪、雹比含水量在 110~140 min 基本达到最大值,之后有减小趋势;60 min 时,可信的背景场误差协方差基本建立,用此时的集合平均分析场做的确定性预报误差更小。

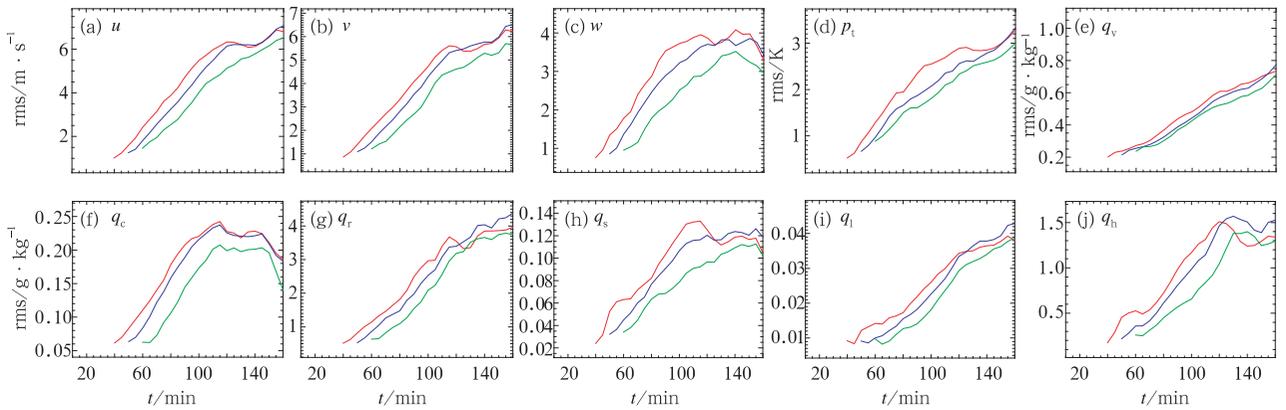


图 14 40 min(红线)、50 min(蓝线)和 60 min(绿线)的集合平均分析场做的确定性预报的各变量预报误差

Fig. 14 The rms errors of variables of the forecasts begin from ensemble mean analysis at 40 min (red), 50 min (blue) and 60 min (green)

## 4 主要结论

本文实施了 EnKF 同化模拟雷达资料的观测系统模拟试验,验证了 WRF-EnKF 雷达资料同化系统的同化效果,得到结论如下:

(1) WRF-EnKF 雷达资料同化系统有能力准确分析模式风暴的风场、热力场、微物理量场的细致特征。试验中 EnKF 在约 3 个同化循环后就能比较准确地重建模式风暴,更多的同化循环后,各变量分析场的位置、结构、量级都与真值场非常一致。

(2) 预报误差在每次同化分析后都有所下降,试验中所有变量的预报和分析误差在经过同化循环后都显著减小;各变量预报、分析误差最大的高度都对应该变量发展最旺盛的高度,同化分析基本上能使预报场在各层上都有所改进,对预报场误差较大的层次的更正更为显著。

(3) 背景场误差协方差在 EnKF 分析中扮演着

重要角色,试验表明,约 8 个同化循环后,EnKF 能在雷达反射率、径向风观测与背景场间建立较可靠的相关关系,使模式各变量场能被准确分析更新,背景场误差协方差在水平方向和垂直方向都有着复杂的结构,是高度非均匀、各项异性和流依赖的。

(4) 用集合平均分析场做的确定性预报显示,在短时间内(约 30 min)预报能较好地保持真值场风暴的细节结构,但预报误差增长很快,预报误差主要来自于虚假的分裂单体和单体传播速度与真实单体的不一致。误差的迅速增长,一方面可能与非线性的对流风暴系统发展迅速、预报时效短有关,另一方面也与作为预报初始场的集合平均分析场的误差有关系。

## 参考文献

- [1] Gal-Chen T. A method for the initialization of the anelastic equations; Implications for matching models with observations[J]. Mon Wea Rev, 1978, 106: 587-602.
- [2] Hane C E, Ray P S. Pressure and buoyancy fields derived

- from Doppler radar data in a tornadic thunderstorm[J]. *J Atmos Sci*, 1985, 42:18-35.
- [3] Lin Y, Ray P S, Johnson K W. Initialization of a modeled convective storm using Doppler radar-derived fields[J]. *Mon Wea Rev*, 1993, 121: 2757-2775.
- [4] Xue M, Wang D, Hou D, et al. Prediction of the 7 May 1995 squall line over the central U. S. with intermittent data assimilation[R]. 12th Conference on Numerical Weather Prediction, 1998, 191-194.
- [5] Zhang J, Carr F, Brewster K. ADAS Cloud Analysis[R]. 12th Conference on Numerical Weather Prediction, 1998, 185-188.
- [6] Jones C D, Macpherson B. A latent heat nudging scheme for the assimilation of precipitation data into an operational mesoscale model[J]. *Meteor Appl*, 1997, 4:269-277.
- [7] 杨毅, 邱崇践, 龚建东, 等. 利用3维变分方法同化多普勒天气雷达资料的试验研究[J]. *气象科学*, 2008, 28(2):124-132.
- [8] Sun J, Crook N A. Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint: Part I. Model development and simulated data experiments[J]. *J Atmos Sci*, 1997, 54:1642-1661.
- [9] 李媛, 刘健文, 董佩明, 等. GRAPES-3Dvar 雷达资料直接同化对江淮暴雨数值预报影响的分析研究[J]. *气象*, 2011, 37(4):403-411.
- [10] 许小永, 郑国光, 刘黎平. 多普勒雷达资料4DVAR同化反演的模拟研究[J]. *气象学报*, 2004, 62(4):410-422.
- [11] 杨艳蓉, 王振会, 杨洪平, 等. 多普勒雷达反射率与径向风资料在数值模式中的应用试验[J]. *气象*, 2008, 34(6):95-110.
- [12] 王叶红, 赵玉春, 崔春光. 雷达降水资料一维变分同化的敏感性研究[J]. *气象*, 2004, 30(4):6-10.
- [13] Evensen G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics[J]. *J Geo Res*, 1994, 99(C5):10143-10162.
- [14] Snyder C, Zhang Fuqing. Assimilation of simulated doppler radar observations with an ensemble Kalman filter[J]. *Mon Wea Rev*, 2003, 131:1663-1677.
- [15] Sun J, Crook N A. Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint: Part I. Model development and simulated data experiments[J]. *J Atmos Sci*, 1997, 54:1642-1661.
- [16] Dowell D, Zhang F, Wicker L J, et al. Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma supercell: Ensemble Kalman filter experiments[J]. *Mon Wea Rev*, 2004, 132(8):1982-2005.
- [17] 许小永, 刘黎平, 郑国光. 集合卡尔曼滤波同化多普勒雷达资料的数值试验[J]. *气象学报*, 2006, 30(4):712-728.
- [18] Tong Mingjing, Xue Ming. Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS Experiments[J]. *Mon Wea Rev*, 2005, 133(7): 1789-1807.
- [19] Xue Ming, Tong Mingjing. An OSSE framework based on the ensemble square root Kalman filter for evaluating the impact of data from radar networks on thunderstorm analysis and forecasting[J]. *J Atmos Oce Tech*, 2006, 23(1):46-66.
- [20] Zhang Fuqing, Weng Yonghui. Cloud-resolving hurricane initialization and prediction through assimilation of doppler radar observations with an ensemble Kalman filter[J]. *Mon Wea Rev*, 2009, 137(7):2105-2125.
- [21] Smith P L, Myers Jr C G, and Orville H D. Radar reflectivity factor calculations in numerical cloud models using bulk parameterization of precipitation processes[J]. *J Appl Meteor*, 1975, 14:1156-1165.
- [22] Lin Y-L, Farley R D, and Orville H D. Bulk parameterization of the snow field in a cloud model[J]. *J Climate Appl Meteor*, 22, 1065-1092.
- [23] Houtekamer P L, Mitchell H L. A sequential ensemble Kalman filter for atmospheric data assimilation[J]. *Mon Wea Rev*, 2001, 129:123-137.
- [24] Gaspari G, and Cohn S E. Construction of correlation functions in two and three dimensions[J]. *Quart J Roy Meteor Soc*, 1999, 125: 723-757.
- [25] Murphy J M. The impact of ensemble forecasts on predictability[J]. *Quart J Roy Meteor Soc*, 1988, 114, 463-493.