

集合预报误差在 GRAPES 全球四维变分同化中的应用 研究 II：参数选取与数值试验*

龚建东 张林 王瑞春
GONG Jiandong ZHANG Lin WANG Ruichun

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
 1. *National Meteorological Centre, Beijing 100081, China*
 2. *Numerical Weather Prediction Center of CMA, Beijing 100081, China*
- 2020-04-24 收稿, 2020-10-13 改回.

龚建东, 张林, 王瑞春. 2021. 集合预报误差在 GRAPES 全球四维变分同化中的应用研究 II: 参数选取与数值试验. 气象学报, 79(1): 48-64

Gong Jiandong, Zhang Lin, Wang Ruichun. 2021. The study of introducing ensemble forecast errors in the GRAPES global 4DVar Part II: Parameters determination and numerical experiments. *Acta Meteorologica Sinica*, 79(1):48-64

Abstract In the first part of the study, the scientific scheme of how to effectively apply ensemble forecast errors is discussed, and the framework for the application of ensemble forecast errors in the global four dimensional variational data assimilation (4DVar) of the Global Regional Assimilation and PrEdiction System (GRAPES) is identified. Based on this work, the present paper further studies the application of ensemble forecast errors in the GRAPES global 4DVar. The study is focused on solving computational efficiency issues for efficient generation of ensemble samples close to or exceeding 100, as well as key parameters determination issues for matching with the GRAPES global 4DVar. The 4DVar based Ensemble of Data Assimilations (EDA) method is chosen to generate the ensemble samples. By using the preconditioning information generated during the minimization iteration of the first sample to precondition the minimization of other samples, the computational efficiency is increased by a factor of two. By using the Valid-Time-Shifting Perturbation (VTSP) method, the ensemble members are enlarged by a factor of three. The ensemble forecast errors variance inflation method is used, and the horizontal localization scale is selected to be 1.4 times of the horizontal length scale of the stream-function background error. The weight coefficients of background error and ensemble forecast errors are determined by numerical experiments, and the best weight for ensemble forecast error is 0.7 for 60 ensemble members. The results of two 52 d numerical experiments over the winter and summer seasons show that the improvement of ensemble four dimensional variational data assimilation (En4DVar) over 4DVar in the northern and southern hemisphere is mainly concentrated at 700—30 hPa in winter season and 400—150 hPa in the summer season. The tropical region is less affected by seasonal changes, and the improvement of geopotential height, wind field and temperature are obvious with En4DVar, and the improvement of meridional wind field is the most significant. The method developed in this study for the application of ensemble forecast errors in GRAPES Global 4DVar is reasonable and feasible.

Key words Ensemble data assimilation, Ensemble forecast errors, En4DVar, Parameter determination, Numerical experiments

摘 要 研究的第一部分讨论了如何有效应用集合预报误差的科学方案, 确定了集合预报误差在 GRAPES (Global Regional Assimilation and PrEdiction System) 全球 4DVar (four dimensional variational data assimilation) 中应用的分析框架。在此基础上研究了针对集合预

* 资助课题: 国家重点研发计划项目(2018YFC1506700、2018YFC1506705)。

作者简介: 龚建东, 主要从事数值预报资料同化、集合预报研究。E-mail: gongjd@cma.gov.cn

报误差实际应用于 GRAPES 全球 4DVar, 解决接近或超过 100 个集合样本数时高效生成的计算效率问题, 以及与 GRAPES 全球 4DVar 匹配的同化关键参数确定问题。选择基于 4DVar 的集合资料同化方法生成集合样本, 通过将第 1 个样本极小化迭代过程中产生的预调节信息用于其他样本极小化做预调节, 将计算效率提高了 2 倍。通过时间错位扰动方法增加集合样本数, 实现集合样本增加到 3 倍。对集合方差进行膨胀, 并选择水平局地化相关尺度为流函数背景误差水平相关的 1.4 倍。通过批量数值试验方法确定背景误差与集合预报误差的权重系数, 对 60 个集合样本当集合预报误差权重为 0.7 时预报效果最好。对北半球夏、冬两季各 52 d 的批量试验表明, 对于南、北半球 En4DVar (ensemble 4DVar) 较 4DVar 的改进在冬季主要集中在 700—30 hPa, 而在夏季主要集中在 400—150 hPa。赤道地区受季节影响较小, En4DVar 对位势高度、风场与温度的改进都较为明显, 且经向风场的改进最为显著。文中研发的集合预报误差在 GRAPES 全球 4DVar 中应用的方法合理可行。

关键词 集合资料同化, 集合预报误差, En4DVar, 参数选择, 数值试验

中图法分类号 P435 P456.7

1 引言

GRAPES 全球四维变分资料同化(4DVar) 系统已实现业务运行(Zhang, et al, 2019)。4DVar 使用气候背景误差(简称为背景误差), 该方法的基本原理表明, 虽然在其同化时间窗内能通过切线性与伴随模式积分来隐式计算流依赖背景误差, 但在下一次同化分析时 4DVar 对流依赖背景误差信息失去“记忆”能力。要克服这一缺陷需要在 4DVar 中引入集合样本估计的集合预报误差, 发展集合四维变分资料同化(En4DVar)方法, 从而实现在同化时间窗起始时刻就引入流依赖背景误差。为了构建和运行 En4DVar, 首要问题是生成适当的集合样本并在同化过程中合理应用。如何生成适合 GRAPES 同化的集合样本并在 GRAPES 全球 4DVar 中合理使用集合预报误差, 在 GRAPES 全球 En4DVar 的研发中还没有成熟经验可循。

利用集合卡尔曼滤波(EnKF, ensemble Kalman filtering)类方法或基于 4DVar 的集合资料同化(EDA)方法生成集合样本, 并将其用于变分资料同化, 动态估计随天气流型变化的集合预报误差, 是国际资料同化研究的重点和业务发展主流(Clayton, et al, 2013; Kleist, et al, 2015a, 2015b; Buehner, et al, 2013; Bonavita, et al, 2012; Raynaud, et al, 2011)。集合卡尔曼滤波(Evensen, 1994; Houtekamer, et al, 2005)、集合方根滤波(EnSRF, Whitaker, et al, 2002, 2008; Tippett, et al, 2003)、集合卡尔曼变换(ETKF, Bishop, et al, 2001)等分别被加拿大气象局、美国环境预报中心、英国气象局采用。以上方法中, EnKF 方法是基础性方法, 其他方法是 EnKF 方法的变形或简化(Tippett, et al, 2003)。其中, EnKF 与集合方根滤波都属于资料同化方法, 通过

集合样本估计卡尔曼增益矩阵, 并通过使用观测资料和观测误差计算分析场。这两种方法的差异在于如何在不同样本中引入观测扰动的影响, EnKF 通过一定的误差模型直接扰动观测资料(Burgers, et al, 1998), 但这种显式扰动会造成抽样误差增大; 集合方根滤波方法通过重新定义卡尔曼增益矩阵的表达式来隐式表述观测扰动的影响, 但其只能采用顺序同化方式每次使用一个观测资料, 无法像 EnKF 那样同时使用多个观测资料, 需要特殊的并行计算处理。集合卡尔曼变换方法不是一种资料同化方法, 而是一种集合预报样本产生方法, 只使用观测误差而不使用观测值, 并不求解卡尔曼增益矩阵。由于集合卡尔曼变换是在观测空间计算变换矩阵, 其优点是计算量小、生成集合样本快速; 其不足是集合预报误差在观测空间的变化与模式空间地理位置的分布特征不对应。

基于 EnKF 类方法生成集合样本, 可以平行更新每个集合样本的分析场(Evensen, 1994), 具有平行计算程度高可扩展性好、可以同时生成大量集合样本的优点, 但该方法要求在业务变分资料同化系统之外独立发展一套集合卡尔曼滤波系统, 开发和维持成本较高。基于 4DVar 的 EDA 方法利用扰动观测资料和表述模式不确定性的模式扰动技术来生成集合样本(Bonavita, et al, 2012), 为欧洲中期天气预报中心(ECMWF)、法国气象局所采用。与基于 EnKF 类方法的方案相比, EDA 直接使用业务已有 4DVar 系统, 只需对观测资料、边界条件、模式物理过程倾向叠加随机扰动并调整运行流程即可以计算分析样本, 更加方便业务数值预报系统的维护。而 EDA 方法的不足在于 4DVar 的计算量非常大, 限制了集合样本数的增加, 并且为了减少计算消耗系统一般在较低空间分辨率下运行。目前

在世界各主要业务中心的混合同化系统发展中采用接近或超过 100 个集合样本数并对集合预报误差给予较大的权重是主要趋势,因而提高 EDA 方法的计算效率非常关键。

Houtekamer 等 (2005) 研究指出,估计集合预报误差时会存在由于集合样本有限造成的抽样误差、观测误差估计不当、正演观测算子存在误差、预报与观测误差出现非高斯分布特征、预报模式中未能考虑的误差等问题,这些因素的综合作用会造成集合预报误差低估。为减缓低估产生的不利影响研究者们发展了不同的方差膨胀方法,如分析误差方差膨胀法 (Anderson, et al, 1999)、误差额外膨胀法 (Houtekamer, et al, 2005)、分析扰动幅度张弛到预报扰动幅度法 (Zhang, et al, 2004),以及分析误差幅度张弛到先验估计的集合预报误差幅度方法 (Whitaker, et al, 2012) 等。此外, Wang 等 (2003) 用新息向量来对分析扰动进行膨胀处理。ECMWF 在减缓集合预报误差存在低估现象时,一方面在观测扰动基础上,进一步引入针对多种物理过程的随机扰动来表征已有参数化方案的不确定性,另一方面通过比较集合离散度与集合预报误差来定量诊断集合预报误差低估幅度,并将全球划分为南、北半球、赤道 3 个区域,对每个区域采用不同的膨胀系数 (Bonavita, et al, 2012)。

中国在集合卡尔曼滤波方法上进行了多项研究探索,其中基于 GRAPES 区域模式的研究工作有 EnKF 方法 (庄照荣等, 2011),以及将 ETKF 估计的集合预报误差用于区域三维变分资料同化的混合资料同化方法 (Chen, et al, 2015)。在 GRAPES 全球模式中也开展了 ETKF 初值扰动方法研究 (马旭林等, 2008),利用模拟探空观测资料进行试验,还没有考虑卫星等非局地观测资料的局地化技术。刘柏年等 (2015) 开展了集合预报误差的方差滤波方法研究,采用谱滤波方法根据信号和噪声尺度的统计特征构造低通滤波器来滤除集合预报方差估计值中的大部分随机取样噪声。

考虑到 GRAPES 全球模式还没有形成业务可用的 EnKF 系统,本研究借助现有 GRAPES 全球 4DVar 系统,通过扰动观测资料和扰动海温分析场,利用 EDA 方法来生成集合样本。研究的第一部分 (龚建东, 2021) 讨论了如何有效应用集合预报误差的科学方案,确定了集合预报误差在 GRAPES

全球 4DVar 中应用的分析框架。然而要将集合预报误差实际应用于 GRAPES 全球 4DVar 中,还需要解决如何高效生成集合样本的问题,以及确定与系统相匹配的同化关键参数。本研究重点是将实际估计的集合预报误差用于 GRAPES_GFS 全球 4DVar 中,通过批量数值试验合理给定多项关键参数,并通过数值试验来验证 En4DVar 相对于 4DVar 的优势。

2 EDA 集合样本生成方案

2.1 集合样本生成方法

EDA 方法利用不同扰动观测资料产生多组 4DVar 分析,并在模式积分时扰动海温分析场来考虑下垫面的不确定性,以及通过随机物理过程表示模式不确定性,最后获得多组集合样本。扰动观测和海温分析场可以表示为分析和预报步骤

$$J(\delta\tilde{\mathbf{X}}) = \frac{1}{2} \delta\tilde{\mathbf{X}}^T \mathbf{B}^{-1} \delta\tilde{\mathbf{X}} + \frac{1}{2} \sum_{i=0}^L [(\tilde{\mathbf{d}}_i - \mathbf{H}_i \mathbf{M}_{0,i} \delta\tilde{\mathbf{X}})^T \cdot \mathbf{R}^{-1} (\tilde{\mathbf{d}}_i - \mathbf{H}_i \mathbf{M}_{0,i} \delta\tilde{\mathbf{X}})] + J_c \quad (1)$$

$$\tilde{\mathbf{X}}^b = \mathbf{M}_{0,L} (\tilde{\mathbf{X}}^a + \delta\tilde{\mathbf{X}}) + \tilde{\mathbf{Q}} \quad (2)$$

式(1)为分析步骤,右端第 1 项为背景项,第 2 项为观测项, J_c 是基于数字滤波的弱约束项,用来抑制切线性模式时间积分中因分析场动力不平衡产生的噪音。 \mathbf{B} 是背景误差协方差, \mathbf{R} 是观测误差协方差, \mathbf{H} 为非线性观测算子。 (\cdot) 表示对一组集合样本中的任意一个样本进行分析或预报,形成分析或预报场。 $\delta\tilde{\mathbf{X}}$ 为集合样本中的任意一个样本的分析增量。 $\mathbf{M}_{0,i}$ 是非线性预报模式 \mathbf{M} 的切线性模式,下标 i 表示时间 t_i 。 $\tilde{\mathbf{d}}_i = \mathbf{y}^o + \boldsymbol{\eta}_i - \mathbf{H}_i \mathbf{M}_{0,i} (\tilde{\mathbf{X}}^b)$ 为在观测场 \mathbf{y}^o 上叠加观测误差扰动 $\boldsymbol{\eta}_i$ 后的新息向量, $\boldsymbol{\eta}_i$ 服从期望值为 0、方差为 \mathbf{R} 的观测误差分布。式(2)为预报步骤,其中 $\tilde{\mathbf{Q}}$ 为在模式预报中叠加的海温分析场随机扰动,它服从期望值为 0、方差为海温分析误差的随机扰动。EDA 方法在 GRAPES 模式中具体实现时,考虑了海温分析场误差存在水平相关的影响 (龚建东等, 2020)。式(1)–(2)所表示的集合样本生成的具体内容如表 1 所示。

上述 EDA 方法基于 GRAPES 全球 4DVar 系统,由于单个 4DVar 计算消耗较大限制了集合样本数的增加。Courtier 等 (1994) 研究表明,在 4DVar 中采用增量分析方案并适当降低内循环水平分辨

表 1 基于 GRAPES 全球 4DVar 的集合资料同化系统的扰动内容

扰动目的	扰动内容	是否应用
描述观测误差导致的分析场的不确定性	对4DVar同化的所有观测叠加随机扰动,扰动值满足期望为0、方差为系统中的观测误差方差的正态分布	是
描述海温分析误差导致的预报不确定性	对模式预报的边界条件海温分析场叠加随机扰动,扰动值满足期望为0、方差为海温分析场误差的正态分布	是
描述模式物理过程参数化不确定性导致的预报不确定性	在预报模式中叠加随机物理过程,如随机物理参数化倾向扰动SPPT方案,随机动能补偿方案SKEB方案等	否

率,可以在基本不影响分析结果的情形下显著减小计算消耗。GRAPES 全球 4DVar 就采用了这样的分析方案,模式轨迹及新息向量计算在高分辨率上进行,而切线性与伴随模式时间积分、目标函数极小化迭代计算在低分辨率上进行。本研究中 GRAPES 全球 4DVar 的高、低分辨率配置分别为 0.25°与 1.0°,由于 EDA 方法只需提供与全球 4DVar 内循环分辨率匹配的集合样本,因而其高、低分辨率可以设置为 1.0°与 1.5°,这显著减少了 EDA 方法所需的计算量。在此基础上顺序进行 EDA 分析样本和预报样本的生成计算,所有样本累计所耗时间仍然明显大于单次高分辨率 GRAPES 全球 4DVar 的计算时间,通过采用平行作业方法生成集合样本的计算时间还将继续缩短。

2.2 极小化预调节信息

分析 EDA 方法的计算消耗,其计算量主要集中在目标函数极小化迭代求解过程中切线性与伴

随模式的时间积分。张林等(2017)研究表明,间隔相近的两次 4DVar 分析对应的海森矩阵(目标泛函相对于控制变量的二阶偏导数构成的矩阵)变化不大,前一次极小化迭代过程中产生的信息可提供给下一次极小化做预调节,经预调节后的极小化算法的目标函数收敛效率能得到明显提高。对于同一时刻不同集合样本中的 4DVar 分析而言,由于它们的背景场和观测场均较为接近,对应的海森矩阵变化也不大。因而当第一个样本的 4DVar 完成计算后其在极小化迭代过程中产生的信息也可以用于后续其他样本。对 2018 年 12 月 1—30 日逐 6 h 同化循环的总计 120 个样本的统计表明,在引入极小化预调节的情形下,目标函数经过 20 次极小化迭代后的数值与未做预调节情形下 40 次迭代的结果相当,也即计算效率提高了 1 倍(图 1a)。进一步比较极小化迭代后目标函数梯度的下降比例,未做预调节情形下经 40 次迭代梯度下降为初始值的 7.5%,

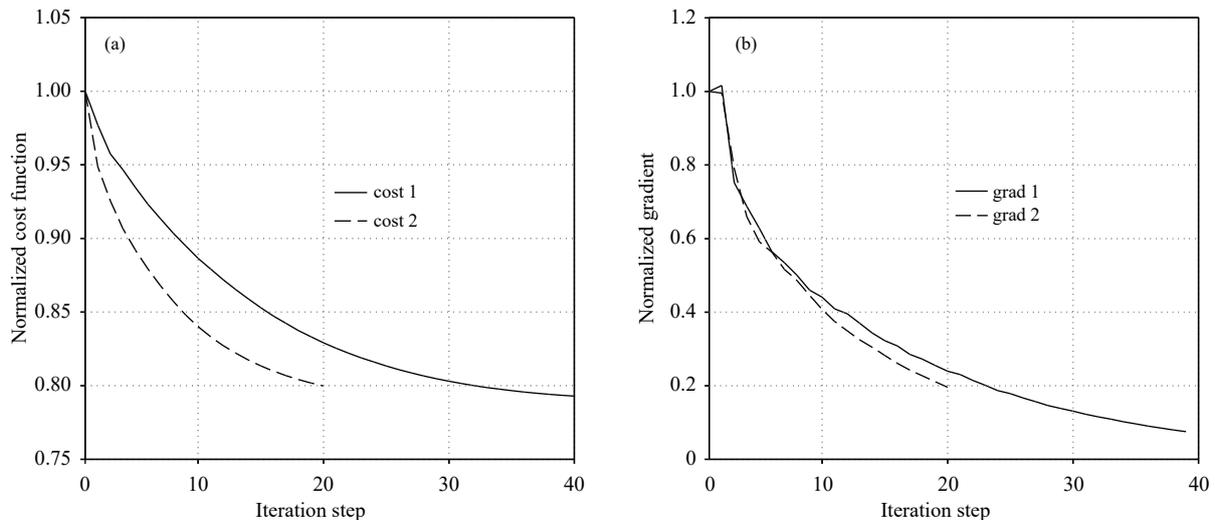


图 1 预调节信息对目标函数收敛的影响 (a. 归一化后的目标函数, b. 归一化后的目标函数梯度; 虚线与实线分别为使用或未使用预调节信息, 2018 年 12 月 1—30 日 6 h 同化循环的 120 个样本统计)

Fig. 1 Impacts of preconditioning information on cost function convergence (a. normalized cost function, b. normalized gradient of cost function; dashed and solid lines are for results with and without preconditioning information, respectively; 120 samples from 6 h data assimilation cycle in 1—30 December 2018 are used)

而使用预调节情形下经 20 次迭代梯度下降为初始值的 19.5%，梯度下降幅度相差近 2 倍 (图 1b)。由于 EDA 方法中的 4DVar 主要用于集合样本生成，并不涉及单次分析的精度，梯度下降程度上的差异可以接受。近 1 个月 EDA 方法的试验表明，使用极小化预调节后目标函数的减小程度和目标函数梯度下降程度逐日变化不大，表现稳定。以上是在 EDA 方法中顺序进行集合样本生成时使用前次预调节信息的方式。当集合样本采用并发作业生成时，借鉴 ECMWF 的经验可以使用高分辨率控制预报的预调节信息用于 EDA 方法中集合样本的计算。

3 集合预报误差应用

3.1 高频噪声滤除与异常扰动抑制

使用集合样本来估计集合预报误差时存在抽样误差，抽样误差主要与集合样本数有关，对 N 个集合样本其对应的蒙特卡洛抽样误差是 $\frac{1}{\sqrt{N}}$ 。抽样误差在不同空间尺度上表现不同 (Bonavita, et al, 2011)，长波对应的抽样误差小而短波对应的抽样误差大。对较短的波动而言其误差构成主要是无用的噪声，需要滤除。Raynaud 等 (2009)、Bonavita 等 (2012) 采用基于三角函数的低通滤波器，此处采用同样的滤波函数

$$\rho(n) = \cos^2\left(\frac{\pi}{2} \min(n, N_{\text{trunc}}) / N_{\text{trunc}}\right) \quad (3)$$

式中， n 为球谐谱函数对应的二维总波数， N_{trunc} 为最大谱截断波数。在试验中按照 Bonavita 等 (2012) 经验将最大谱截断波数取为 106，滤除随机噪声同时保留天气到中尺度主要扰动波数的扰动振幅。垂直方向的噪声直接来源于扰动观测资料对分析的影响。GRAPES 全球 4DVar 中由于位温扰动是通过气压扰动的垂直梯度计算获得，而垂直梯度计算对气压扰动非常敏感，对气压扰动进行噪音滤除能显著减小位温扰动存在“之”字形虚假振荡现象。采用简单的五点平滑方法进行气压扰动的噪声滤除就能收到较好效果。

通过 EDA 方法获得集合样本来估计集合预报误差时，在扣除集合平均后所生成的集合扰动样本之间的数值差异较大，当样本数较小时直接用这些扰动样本来估计集合预报误差时容易受到数值异常偏大样本的干扰而出现误差高估现象，且这种高

估不能通过高频噪声滤除的方式消除，需要对数值异常偏大值进行抑制。GRAPES 全球 4DVar 中，已经有纬向平均的气候背景误差，以此作为参考，当扰动样本数值明显大于气候背景误差时，采用衰减扰动幅度的方式控制数值。表示为

$$\begin{cases} v = v & v \leq a\sigma_c \\ v = \frac{v}{|v|} (a\sigma_c + (b-a)\sigma_c\gamma) & v > a\sigma_c \\ \gamma = a \tan\left(\frac{\pi}{2} \frac{\min(|v|/\sigma_c - a, b)}{b}\right) \end{cases} \quad (4)$$

式中， v 为任一预报变量， σ_c 为背景误差，系数 a 为扰动开始衰减的阈值系数， b 为衰减后的最大值阈值系数。式 (4) 限制了预报扰动的幅度，当扰动幅度大于 $a\sigma_c$ 时，开始采用三角函数进行衰减，最大扰动被限制在 $b\sigma_c$ 以下。试验中风与气压扰动取 $a = 3$ ， $b = 4.6$ ，对比湿 $a = 1.8$ ， $b = 3.0$ ，能获得较好的衰减效果。

3.2 集合预报误差与背景误差

利用 2018 年 6 月数据，比较纬向平均集合预报误差与背景误差。在南、北半球对流层 1000—100 hPa (对应模式 1—48 层) 风场集合预报误差存在低估现象，普遍低 50%—60%，赤道地区集合预报误差低估情况要小于南、北半球，低估 30%—40% (图 2a)。同样气压集合预报误差也存在低估，其低估程度比风场要大，南、北半球低估 60%—70%，赤道地区低估 50%—60% (图 2b)。值得注意的是，背景误差自身存在一定的偏差，并且这种偏差特征在赤道以及南、北半球表现并不一致。赤道地区的大气混沌特征明显，模式预报误差不易增长，用同一检验时刻不同预报时效预报的差异估计背景误差时容易在赤道地区造成背景误差低估 (龚建东等, 2006)，而在中高纬度地区容易出现背景误差高估。纬向平均的分布特征表明，相对于背景误差，集合预报误差明显小，需要进行膨胀处理。考虑到背景误差在赤道地区低估、南北半球存在高估的问题，不能仅参照背景误差来膨胀集合预报误差。

4 En4DVar 参数确定

4.1 权重系数与相关尺度

集合预报误差主要用于对流层大气。在对流层顶附近 (约 100 hPa, 模式 48 层)，随模式层次增加使用衰减函数将集合预报误差权重系数 (β_c) 的数值衰减为 0，背景误差权重系数 (β_c) 相应增大，两者

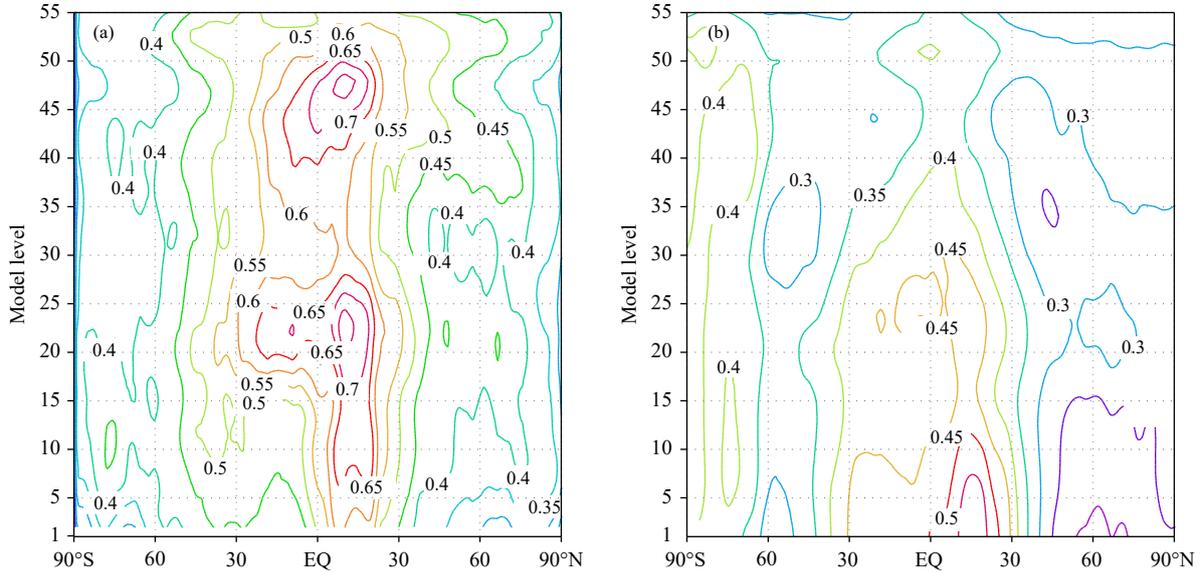


图 2 纬向平均集合预报误差相对于背景误差的比率 (a. 矢量风速, b. 无量纲气压; 使用 2018 年 6 月数据的统计结果; 模式垂直层次 12、18、25、36、48、55 分别大致对应气压层次 850、700、500、250、100、30 hPa)

Fig. 2 The ratios of zonal mean ensemble forecast errors to background error (a. vector wind, b. non-dimensional pressure; statistic results of data for June 2018 are used; model levels of 12, 18, 25, 36, 48, 55 correspond to pressure levels of 850, 700, 500, 250, 100, 30 hPa)

之和为 1。图 3a 给出 $\beta_e = 0.7$ 时权重系数廓线的垂直分布情况。对湿度分析批量试验结果表明, 引入集合预报湿度误差对改进 4DVar 没有明显效果, 其原因和现有 4DVar 中湿度观测数据量偏少 (Zhang, et al, 2018)、集合预报误差偏小有关系。在后续试验中暂不考虑集合预报湿度误差。

水平局地化的相关尺度需要考虑 4DVar 中背景误差水平相关尺度随高度的变化。GRAPES 全球 4DVar 中背景误差水平相关尺度随高度的变化如图 3b 所示, 对流层不同分析变量之间差异显著, 其中比湿的相关尺度最小 (约 175 km), 流函数约 600 km, 非平衡势函数最大 (700—1200 km), 非平衡无量纲气压在 500 km 左右。由于 GRAPES 全球 4DVar 的背景误差中流函数是主导控制变量, 其他变量的平衡部分由其计算得到 (Zhang, et al, 2018)。因此, 在选择水平局地化相关尺度 (L_{loc}) 时, 参照流函数背景误差水平相关尺度 (L_ψ) 的垂直分布特征来给出。批量试验的结果表明, 当 $L_{loc} = 1.4L_\psi$ 时 (图 3b), 也即在对流层 L_{loc} 取值在 850 km 左右时, 能获得较好的改进结果。

4.2 集合样本数影响

当集合样本足够大时能减少抽样误差影响, 可以合理地估计集合预报误差协方差。在增加集合

样本数的方法中, Xu 等 (2008a, 2008b) 提出在分析时刻前后利用集合样本在时间序列上扩展来增加集合预报数, 可以明显改进区域 EnKF 的分析与预报效果。Huang 等 (2018) 提出利用时间错位扰动 (VTSP, valid time shifting perturbation method) 方法, 通过增加样本在时间上的输出频次, 由时间错位来增加集合样本数。具体做法是, 对需要集合样本的时间 T , 使用由同一分析场起报、分别预报到 $T-\tau$ 、 T 、 $T+\tau$ 这 3 个时次的集合样本, 其中 τ 为时间错位的间隔。为了避免不同时间的大尺度场存在的相位误差, 采用扣除对应自身时间集合平均的扰动样本, 并采用 $T-\tau$ 与 $T+\tau$ 对称采样方式。研究表明, 这种方法对台风等系统扰动样本基本不存在相位误差。在 GRAPES 全球 EDA 方法生成集合样本时, 没有使用同一起报时间不同预报时长的样本, 而是采用当前起报时间的 T 、 $T+\tau$ 集合样本, 和 6 h 之前起报、时间对应 $T-\tau$ 时刻的集合样本。由式 (1) 可知, 同化时间窗起始和终止时间的 6 h, 是有观测约束的时段, 代表了分析场。本研究所使用集合样本, 所有样本的预报时长都大于 6 h, 代表了预报场。采用以上方法对每个集合预报通过增加 3 h 预报时效的较小计算代价, 实现集合样本增加到 3 倍, 最终产生 60 个集合样本。

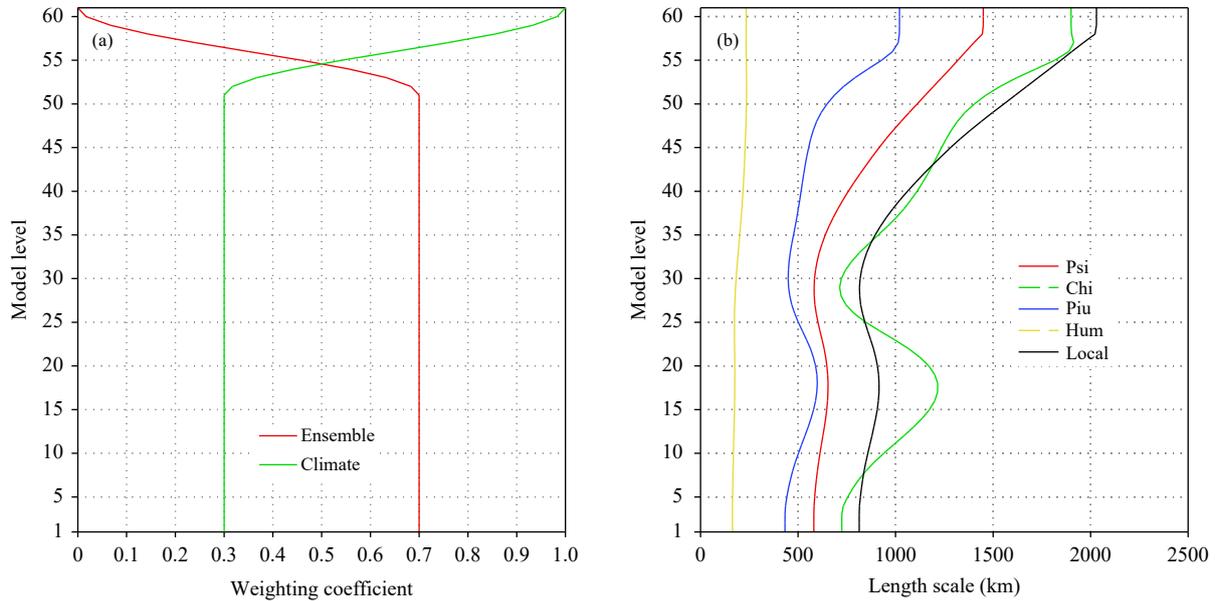


图3 误差权重与相关尺度廓线 (a. 集合预报误差与背景误差的权重系数, b. 背景误差及局地化的水平相关尺度)

Fig. 3 Profiles of error weighting and correlation scale (a. ensemble forecast error and background error weighing coefficients, b. horizontal correlation scale of background error and localization)

对EDA方法产生的20个集合样本,以及通过VTSP方法进行样本扩展后产生的60个集合样本,比较其引入4DVar后对分析与预报的影响。取背景误差权重系数 $\beta_c = 0.8$,集合预报误差权重系数 $\beta_e = 0.2$,集合预报误差未做膨胀处理,结果表明20个集合样本引入后,GRAPES全球En4DVar与4DVar的分析与预报效果相当,集合预报误差引入4DVar后的作用不明显。当集合样本增加到60个后,对南、北半球第1天的风、温度与位势高度预报效果均有改善,更长时效的改进不明显,对赤道地区的改善略大一些,特别是风场和位势高度场的改进较明显,且正影响可以达到5 d,En4DVar的效果略优于4DVar(图略)。其他多组对照试验表明,只引入20个集合样本,通过集合预报方差膨胀、集合预报误差权重系数加大,均不能产生明显改进效果,表明集合样本数对GRAPES全球En4DVar起到关键作用。

4.3 集合预报方差膨胀

集合预报误差与背景误差的比较(图2)表明,在利用集合样本估计集合预报误差时出现低估问题,且南、北半球比赤道地区的低估更加严重。Bonavita等(2012)在比较ECMWF EDA法生成集合样本的离散度和均方根误差时,发现集合样本的离散度明显低于均方根误差,且在南、北半球要高

于赤道地区。一种基于集合样本离散度和均方根误差的膨胀方法被发展出来,利用前5天样本序列的滑动平均来计算方差膨胀因子,北半球1和8月的平均膨胀幅度在1.5左右。由于EDA法生成的每个集合样本相当于单个4DVar分析预报循环,并不依赖于集合预报方差,集合预报方差主要用于En4DVar。在本研究中采用了静态膨胀处理法,具体做法是将EDA法生成的1个月集合样本的离散度与预先估计的气候误差进行比较来计算膨胀因子。预先估计的气候误差是将GRAPES全球4DVar背景误差均方根误差和全球集合预报计算的离散度进行加权平均得到,其中两者加权系数分别是0.6和0.4。计算得到的集合预报误差膨胀系数如图4所示。对风场南北半球的膨胀为1.3—1.4倍,而赤道地区基本没有膨胀。在垂直层次上风场膨胀主要在48层以下的对流层,48层以上系数逐步减小。结合图3a,在平流层权重主要放在背景误差上,集合预报误差的作用减少。对于无量纲气压膨胀发生在所有层次,其中48层以下膨胀幅度基本在1.6左右,在48层以上膨胀系数逐步减小,南北半球与赤道的差异不明显。可以发现风场与无量纲气压的膨胀特点在数值和分布上呈现出明显的不同。

取背景误差权重系数 $\beta_c = 0.2$,集合预报误差权

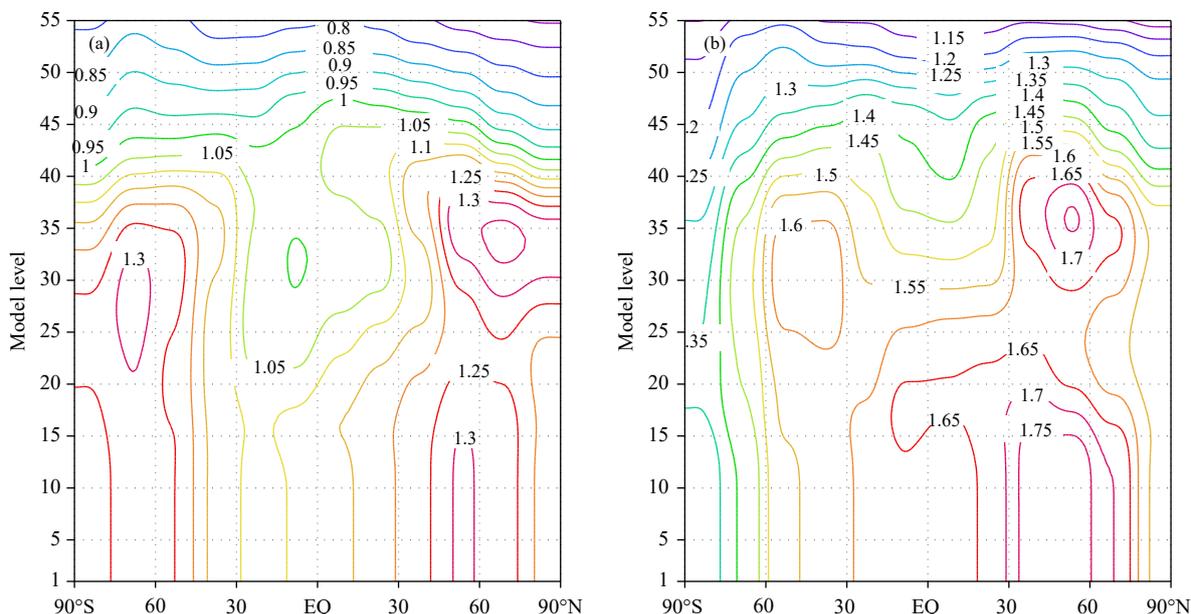


图 4 集合预报误差的膨胀系数 (a. 矢量风速, b. 无量纲气压)

Fig. 4 Inflation coefficients of ensemble forecast errors (a. vector wind, b. non-dimensional pressure)

重系数 $\beta_e = 0.8$, 采用 60 个集合样本, 比较集合预报误差未做方差膨胀和进行方差膨胀的影响。在进行方差膨胀时分别考虑风速与无量纲气压变量均膨胀 1.7 倍, 以及按照图 4 所示的系数进行膨胀, 并进行 1 个月的批量试验 (图略)。结果表明集合预报误差膨胀 1.7 倍对南、北半球有正影响, 而在赤道地区为负影响, En4DVar 不如 4DVar。而对风速与无量纲气压合理膨胀后北半球不同等压面层次的风场、位势高度场与温度场均有明显改进, 在南半球也有所改进但不及北半球显著。赤道地区的改善不明显。这表明对不同分析变量采用符合各自误差特点的膨胀系数要优于对分析变量采用均值膨胀系数。后续试验将采用图 4 所示的系数进行膨胀。

4.4 误差权重系数影响

合理给定背景误差与集合预报误差权重系数对改善 En4DVar 的分析与预报质量至关重要。在全球混合资料同化系统发展中, 世界上不同数值预报中心采用不同的权重系数。分析各业务数值预报中心的背景误差与集合预报误差的权重系数, 对 20 个左右的集合样本, 其抽样误差为 20%, 集合预报误差权重系数在 0.2 左右, 背景误差在 0.8 左右 (Clayton, et al, 2013), 而当集合样本数达到 80 个时, 抽样误差减小到 11%, 集合预报误差权重系数达到 0.75 (Kleist, et al, 2015a, 2015b)。文中使用

60 个集合样本, 其中 20 个通过 EDA 方法生成, 其余 40 个集合样本通过 VTSP 方法生成, 并不是完全独立的集合样本, 需要进行不同权重系数的试验来寻找最佳参数。对集合预报误差权重系数 β_e 分别取权重 0.5、0.7 和 0.9, 对 2018 年 6 月 10 日至 7 月 9 日, 以及 2018 年 11 月 21 日—2019 年 1 月 3 日, 进行夏、冬两个代表时段的数值试验。由图 5、6 所示北半球夏季结果可见, 随着 β_e 取值增大, 温度场在 100 hPa 以上层次改进增大, 但在对流层 250 hPa 附近, 当 $\beta_e = 0.9$ 时预报误差明显增大; 同时, 对流层风场第 1 天预报的误差明显减小, 但在 96 h 以后的预报时效, 当 $\beta_e = 0.7$ 时的效果最好。赤道和南半球也有类似结果 (图略), 表明权重系数取值 0.7 较为合理。对于北半球冬季试验结果类似。为此, 文中后续试验中设定 $\beta_e = 0.7$, 这与 Kleist 等 (2015a, 2015b) 选定的权重系数接近。

5 试验效果分析

通过数值试验确定集合预报误差与背景误差权重系数、水平局地化相关尺度、风速与无量纲气压的集合预报误差膨胀系数后, 还需要通过批量试验验证其对分析和预报的影响。分别进行北半球夏、冬两季各 52 d 的 En4DVar 相对于 4DVar 的批量试验, 重点关注对南、北半球及赤道地区的影响。

折射率资料, 卫星散射计测风资料, 以及极轨与静止气象卫星云导风资料 (Zhang, et al, 2019)。这些资料在 4DVar 中按对应时间分配在同化时间窗的 30 min 间隔的时间槽内。En4DVar 的系统配置与 4DVar 相同, 背景误差与集合预报误差的权重系数分别为 $\beta_c = 0.3$ 与 $\beta_e = 0.7$, 采用通过 VTSP 方法将 EDA 方法产生的 20 个集合样本扩展为 60 个集合样本。

5.1.1 对分析的影响

首先分析集合预报误差引入 4DVar 后对分析场的影响, 取 ECMWF 再分析场作为“实况”, 夏季 52 d 的统计结果如图 7。南、北半球, 集合预报误

差引入 4DVar 后, 从对流层 700 hPa 一直到平流层 50 hPa 的位势高度场均有改善, 且南半球更加显著, 主要表现在均方根误差减小、偏差略有减小。赤道地区, En4DVar 对风场, 尤其是经向风场改善明显。

En4DVar 相对于 4DVar 在赤道地区的改善也可以从均方根误差的时间序列上表现出来, 以赤道地区 300 hPa 为例, 位势高度场、温度与经向风场误差的演变情况如图 8 所示。可以看到在 7 月 20 日之前, En4DVar 的误差虽然略小于 4DVar, 但差异并不大, 而在 7 月 20 日之后的时段内 4DVar 误差出现了突然增大的现象, 而 En4DVar 则抑制了

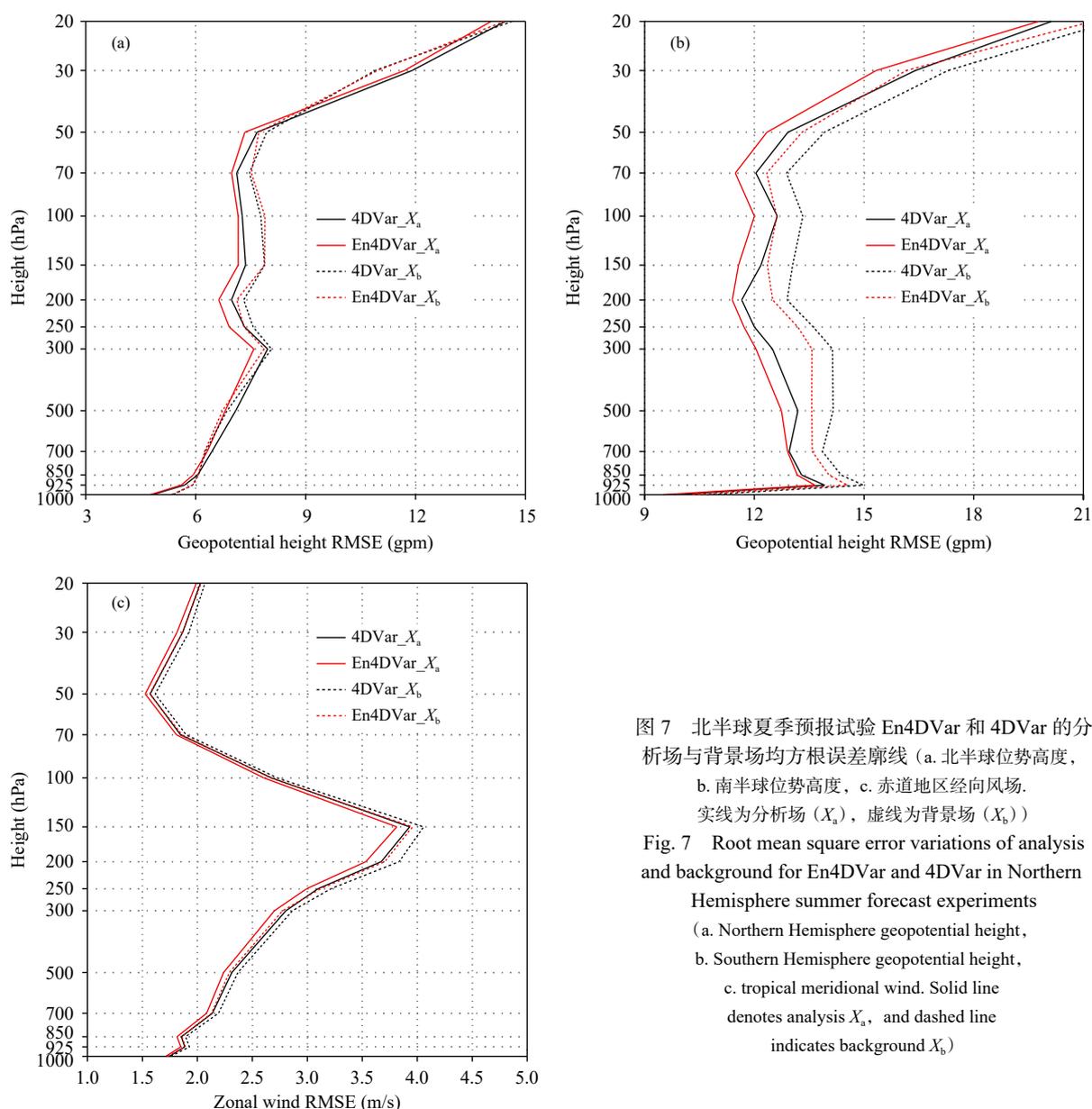


图 7 北半球夏季预报试验 En4DVar 和 4DVar 的分析场与背景场均方根误差廓线 (a. 北半球位势高度, b. 南半球位势高度, c. 赤道地区经向风场。实线为分析场 (X_a), 虚线为背景场 (X_b))

Fig. 7 Root mean square error variations of analysis and background for En4DVar and 4DVar in Northern Hemisphere summer forecast experiments (a. Northern Hemisphere geopotential height, b. Southern Hemisphere geopotential height, c. tropical meridional wind. Solid line denotes analysis X_a , and dashed line indicates background X_b)

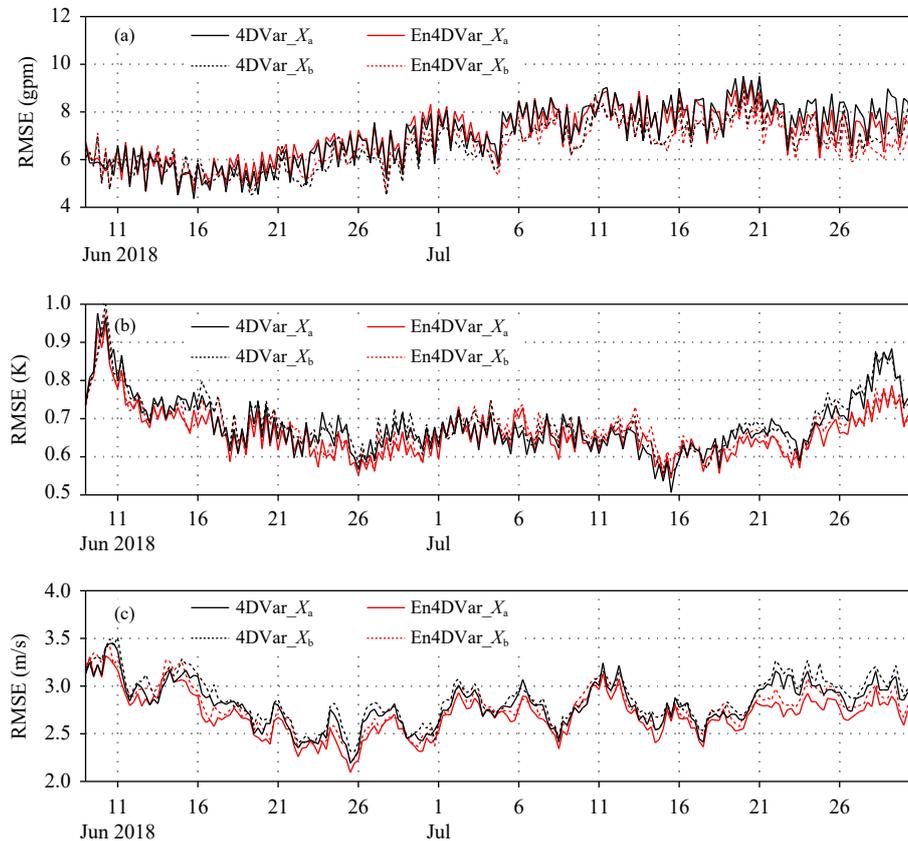


图 8 北半球夏季预报试验赤道地区 300 hPa 逐 6 h 分析与背景误差时间演变
(a. 位势高度, b. 温度, c. 经向风; 实线为分析场 (X_a), 虚线为背景场 (X_b))

Fig. 8 Root mean square error variations of 6 h analysis and background in the tropical region for En4DVar versus 4DVar of Northern Hemisphere summer forecast experiments (a. geopotential height, b. temperature, c. meridional wind; solid line is for analysis (X_a), and dashed line is for background (X_b))

这一现象。有必要对此做进一步分析,以理解En4DVar表现出正效果的内在原因。

选择7月22日00时,分析集合预报误差的特征及对分析造成的影响。60个集合样本估计的集合预报误差在大西洋赤道东部地区及非洲大陆中南部,出现风场集合预报误差大值区。此外,在太平洋中东部和南美洲北部地区也存在较大的风场集合预报误差(图9a)。当集合预报误差被引入4DVar后对应的分析增量出现明显的变化(图9b),在大西洋赤道东部地区及非洲大陆中南部出现分析值差异的大值区,在太平洋中东部地区也出现明显的分析差异。集合预报误差和分析值差异存在较好的对应关系,这是En4DVar分析误差和6h预报误差减小的内在原因。

5.1.2 对北半球夏季预报的影响

En4DVar相对于4DVar的预报,对北半球24h

预报有明显改进,48h有改善,更长时效的预报技巧与4DVar相当;对南半球48h预报有明显改进,72h有改善;对赤道地区72h预报有明显改善,120h有改善。北半球预报改进的时效短与北半球夏季的气候背景有关。夏季中纬度天气尺度到小尺度天气系统较为活跃,集合预报误差的数值偏小,造成集合预报误差的影响较小。在后文中进行了冬季预报试验,来比较季节变化对北半球的影响。从预报效果来看,改进最显著的是赤道地区。图10给出的是赤道地区250hPa温度与风场的预报误差,可以看到En4DVar对应的误差更小,通过显著性检验的有改进时效可以达到7d。

5.2 北半球冬季预报试验

5.2.1 对分析影响

北半球冬季试验的配置与夏季一致,具体时段为2018年11月20日至2019年1月22日,其中预

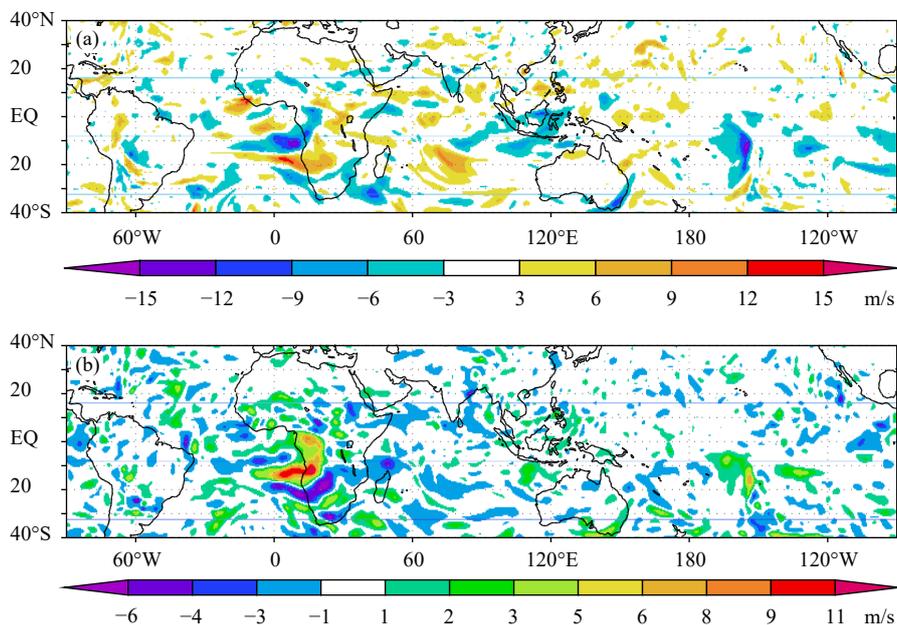


图 9 (a) 赤道地区 300 hPa 集合预报风场误差, (b) En4DVar 相对于 4DVar 的风场分析差异

Fig. 9 (a) Distribution of ensemble forecast wind error at 300 hPa on tropical region, (b) wind analysis difference for En4DVar against 4DVar

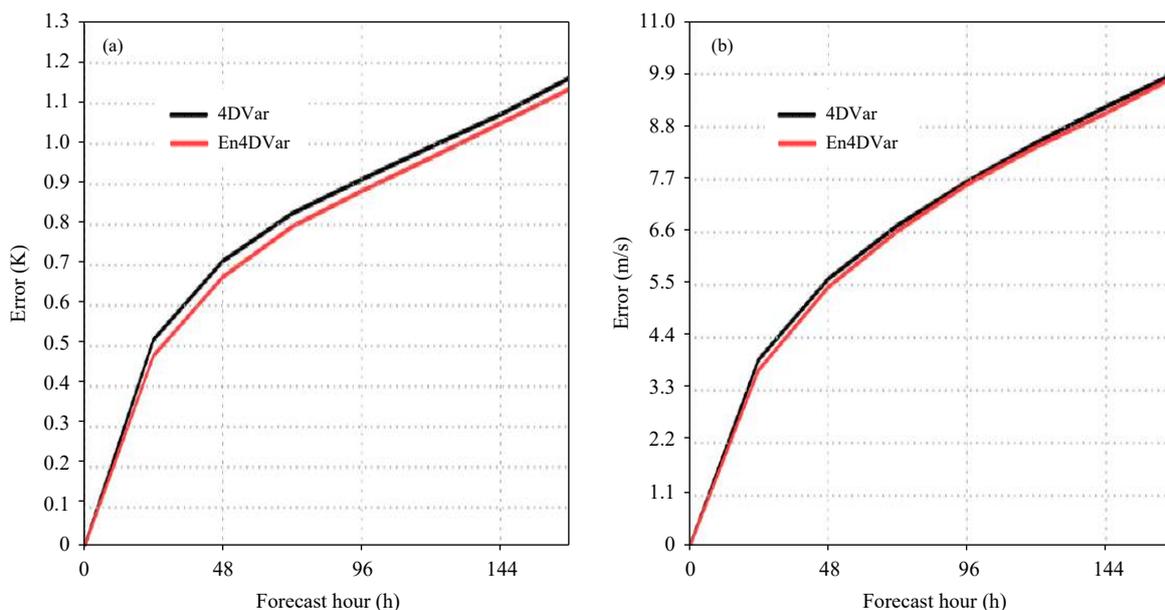


图 10 北半球夏季预报试验赤道地区预报误差 (a. 500 hPa 温度场, b. 250 hPa 矢量风场)

Fig. 10 Forecast root mean square errors in the tropical regional for summer forecast experiments (a. 500 hPa temperature, b. 250 hPa vector wind)

报时段是 2018 年 12 月 1 日至 2019 年 1 月 22 日。冬季试验主要侧重讨论北半球在季节变换时集合预报误差的影响。首先分析集合预报误差在冷、暖季节的变化, 结果如图 11 所示。对于风场误差而言, 南、北半球冷暖季的数值相当, 冷季误差略小一

些。赤道地区低层风场误差大值区出现随季节南北移动的现象, 这与赤道对流辐合带的季节变动相对应。此外, 北半球冷季风场的集合预报误差在边界层增大明显。对于无量纲气压冷、暖两季的变化特征明显, 北半球冷季的集合预报误差明显大于暖

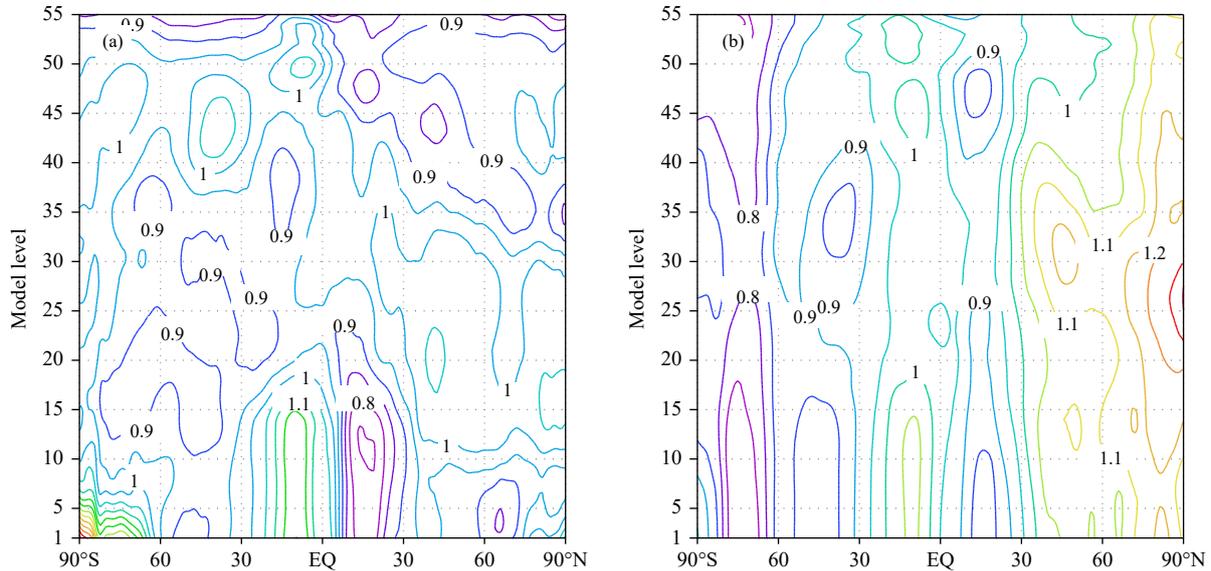


图 11 冬季试验与夏季试验的集合预报误差比值 (a. 风速误差, b. 无量纲气压)

Fig. 11 Ratios of ensemble forecast errors of winter against summer forecast experiments
(a. wind error, b. non-dimensional pressure)

季, 而南半球冷季的集合预报误差要明显小于暖季, 这与风场存在明显的不同。而在赤道地区无量纲气压的误差特征同样体现出对流辐合带随季节南北移动的影响。

北半球冬季 52 d 的统计结果和夏季预报试验接近, 主要差异是对北半球位势高度的均方根误差, En4DVar 较 4DVar 的改进主要集中在 700—30 hPa, 较夏季试验结果更加深厚, 且在 200—50 hPa 的改进幅度较大 (图 12a), 而夏季试验主要改进在 400—150 hPa。北半球位势高度的均方根误差变化与集合预报无量纲气压误差有明显的对应关系。北半球冬季试验时南半球处于暖季, En4DVar 主要改进集中在 700—150 hPa, 而对应的南半球冷季如图 7b 所示, 改进主要位于 700—30 hPa。这可能与冷、暖季节主导的天气系统不同有关, 冷季以大尺度锋面天气系统为主, 天气系统比较深厚, 而暖季以中小尺度天气系统为主, 其影响主要是强降雨、强对流天气等过程的高度。对于北半球风场的改进, 同样是冷季 (图 12b) 大于暖季 (图 12c), 且改进的垂直高度更加深厚。而对于南半球风场冷、暖季的改进相近。

以 300 hPa 位势高度的均方根误差为例来讨论集合预报误差对分析的影响。2018 年 12 月 25—30 日东亚发生了一次寒潮天气过程, 影响范围广,

中国中东部地区普遍降温 6—10℃。从形势场来看北半球被 5 波大尺度系统控制 (图 14b)。从分析误差来看位势高度的分析误差与背景误差在寒潮天气过程期间, En4DVar 要明显小于 4DVar (图 13), 对风场和温度场表现不明显。检查 25—30 日所使用的观测资料的数量并不能发现明显的差异。改进的原因与对照试验 4DVar 采用背景误差, 而 En4DVar 可以使用集合预报误差, 在资料相同的情况下可以更好地使用观测资料有关。由图 14a 可见, 在北极冷涡附近、亚洲等与形势场槽线对应的位置, 气压的集合预报误差明显增大, 表明背景场的不确定性较大, 这一预报误差在 En4DVar 中使用后, 与集合预报误差大值区对应的位置, 如在北极冷涡附近 En4DVar 相对于 4DVar 的分析增量差异明显增大, 其他区域的分析增量也有明显变化 (图 14b), 相应的分析误差也有所减小 (图略)。这表明集合预报误差能反映天气形势变化的特点, 可以改善分析质量。

5.2.2 对冬季预报的影响

北半球冬季预报试验效果要明显好于夏季预报试验效果, 特别是对北半球。从距平相关系数和均方根误差上看, En4DVar 相对于 4DVar 在北半球 48 h 预报有明显改进, 风场改进可以达到 72 h, 南半球 72 h 也有明显改进。改进最显著的是赤道

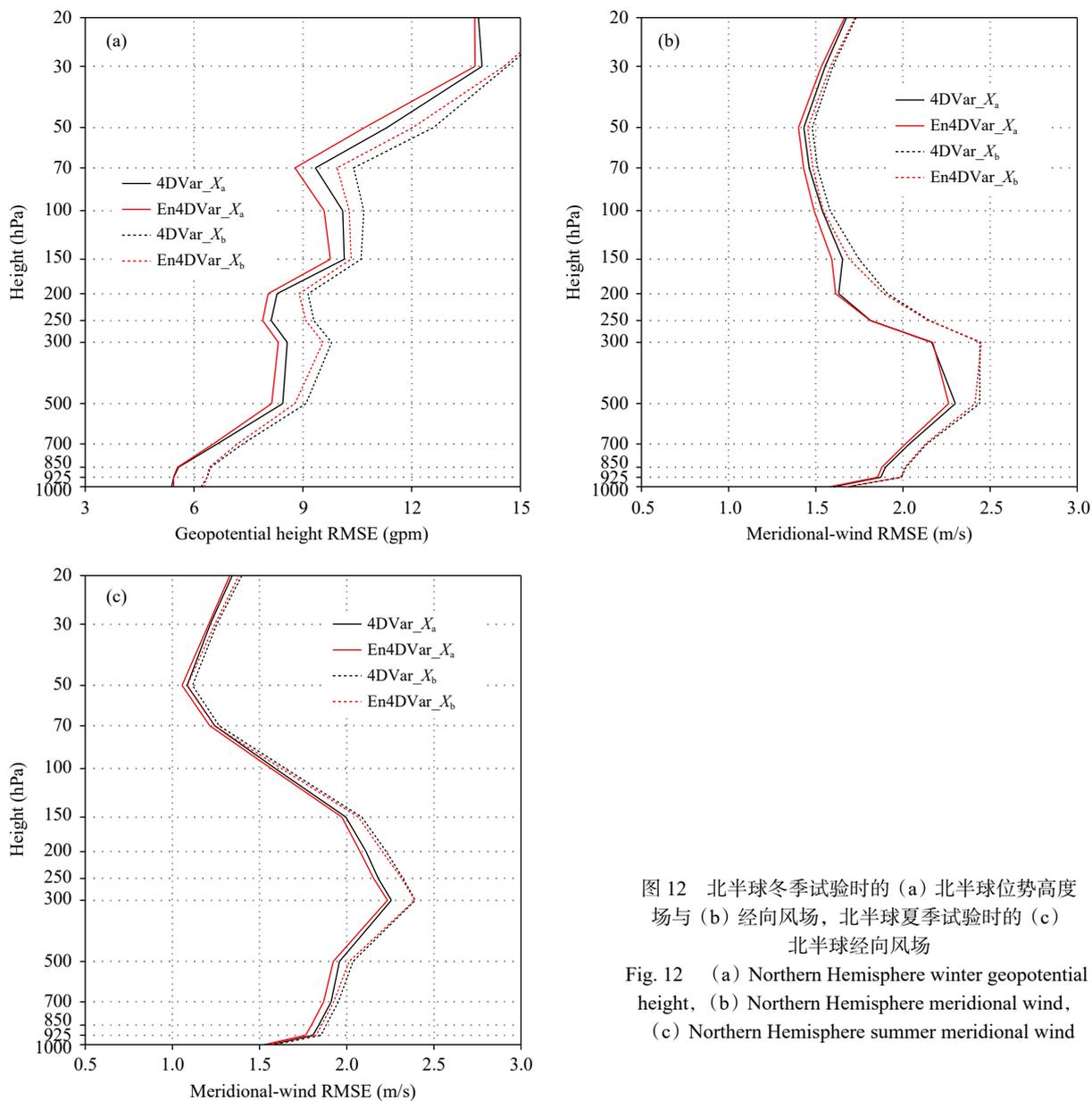


图 12 北半球冬季试验时的 (a) 北半球位势高度场与 (b) 经向风场, 北半球夏季试验时的 (c) 北半球经向风场

Fig. 12 (a) Northern Hemisphere winter geopotential height, (b) Northern Hemisphere meridional wind, (c) Northern Hemisphere summer meridional wind

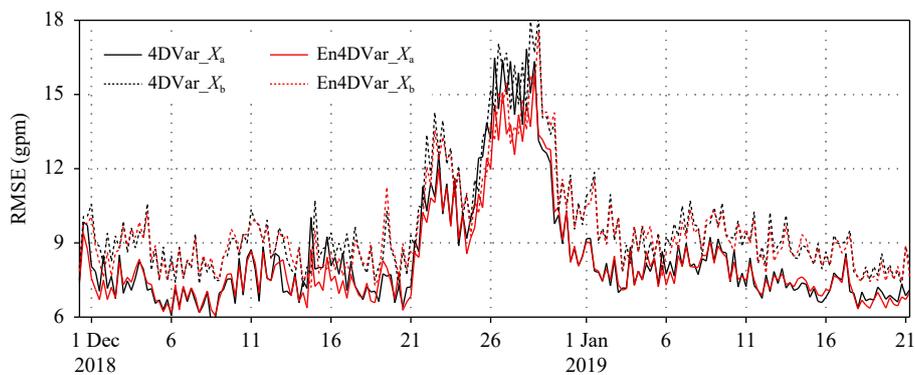


图 13 北半球 300 hPa 位势高度背景误差与分析误差的时间演变

Fig. 13 Temporal evolutions of 300 hPa geopotential height background and analysis errors in the Northern Hemisphere

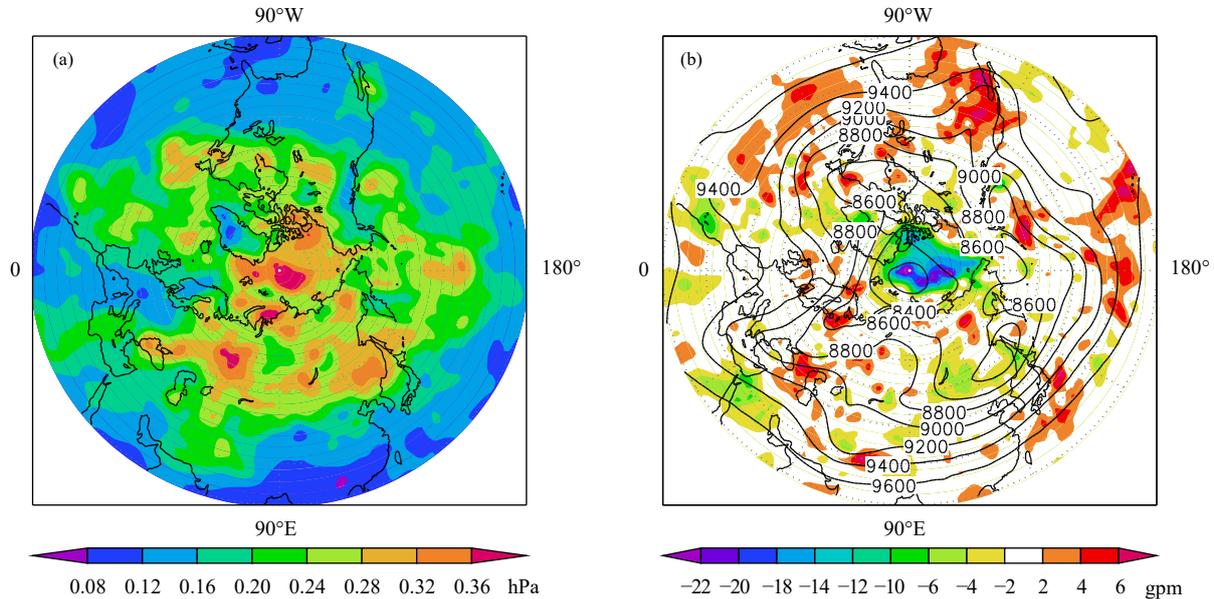


图 14 (a) 北半球 300 hPa 集合预报气压误差, (b) En4DVar 相对于 4DVar 的 300 hPa 位势高度分析增量变化

Fig. 14 (a) Northern hemisphere 300 hPa ensemble forecast pressure error, (b) 300 hPa geopotential height analysis increment difference of En4DVar against 4DVar

地区, 前 120 h 预报评分明显提高(图略), 这些改进在 700—250 hPa 整个层次表现非常一致。

6 总结与讨论

在 GRAPES 全球业务 4DVar 系统中引入集合样本估计的集合预报误差, 发展 En4DVar 同化方法, 进一步改进系统的分析和预报质量。第一部分(龚建东, 2021)重点讨论了集合预报误差引入方式、水平与垂直局地化对分析平衡性的影响, 以及 En4DVar 的计算效率。第二部分重点讨论集合预报误差与背景误差在 GRAPES 全球 4DVar 中合理使用问题, 并通过批量试验验证改进效果。

考虑到 En4DVar 系统包括 4DVar 和集合样本生成两个系统, 研发面向业务应用, 集合样本生成的计算效率是非常重要的一个方面。集合样本生成方法上选择基于 GRAPES 全球 4DVar 的 EDA 方法。EDA 方法生成的每个集合样本都需要通过 4DVar 的分析和预报来实现, 计算量较大。除了采用低分辨率轨迹计算和更低分辨率的内循环计算外, 本研究重复使用第一个样本极小化迭代过程中产生的预调节信息, 并用于其他样本极小化做预调节, 使得 20 次极小化迭代后目标函数的减小程度和 40 次相当, 将计算效率提高了 1 倍。此外, 借鉴

Xu 等(2008a, 2008b)的连续时间采样方法和 Huang 等(2018)时间错位扰动方法, 通过增加样本在时间上的输出频次, 由时间错位来增加集合样本数, 在对每个集合预报通过增加 3 h 预报的少许计算代价, 实现集合样本额外增加 2 倍, 最终产生 60 个集合样本。比较 EDA 方法产生的 20 个集合样本, 以及通过 VTSP 方法进行样本扩展后产生的 60 个集合样本, 多组对照试验表明, 20 个集合样本引入后, 通过集合预报方差膨胀、集合预报误差权重系数加大, 均不能产生明显改进效果, 而当集合样本增加到 60 个后, 对分析与预报有改进效果。

在此基础上, 通过高频噪声滤除与异常扰动抑制, 使得集合样本估计的集合预报误差分布特征更加合理。最后采用了静态膨胀处理方法, 将 EDA 方法生成的 1 个月集合样本的离散度与预先估计的气候误差进行比较, 计算膨胀因子, 由该膨胀因子对每个集合样本进行放大。处理后的集合样本进入到 4DVar 中, 用于估计集合预报误差。文中主要考虑在 4DVar 中应用对流层大气的集合预报误差。在对流层顶附近(约 100 hPa 的位置, 模式 48 层), 随模式层次增加使用衰减函数将集合预报误差权重系数衰减为 0。通过批量试验, 选择水平局地化相关尺度(L_{loc})为流函数背景误差水平相关

(L_y) 的 1.4 倍, 在对流层约为 850 km。此外, 通过批量数值试验方法来确定背景误差与集合预报误差的权重系数, 当 $\beta_e = 0.7$ 时的预报效果最好。

文中进一步通过北半球夏、冬季各 52 d 的批量试验, 验证 En4DVar 相对于 4DVar 的改进效果。结果表明, 北半球 En4DVar 对位势高度误差的改进冷季更加深厚, 暖季试验主要改进在 400—150 hPa, 而冷季的改进主要集中在 700—30 hPa, 且在 200—50 hPa 的改进幅度较大。北半球位势高度的均方根误差变化与集合预报无量纲气压误差有明显的对应关系。南半球 En4DVar 在暖季的改进主要集中在 700—150 hPa, 在冷季改进主要位于 700—30 hPa。赤道地区受季节影响较小, En4DVar 对位势高度、风场与温度的改进都较为明显, 其中最显著的是经向风场。预报试验表明, 引入 En4DVar 后南、北半球 2—3 d 的预报有明显改进, 赤道地区 5 d 有明显改进。

En4DVar 还有一些方面需要继续优化。首先, 集合样本估计的湿度预报误差还没有在 4DVar 中使用。GRAPES 全球模式在赤道边界层顶附近存在较大的湿度偏差, 围绕湿度偏差的改进工作正在开展, 后续将考虑使用湿度集合预报误差。再者, 目前集合预报误差采用静态膨胀处理方法, 区分了不同变量、不同高度和纬度的膨胀系数, 但还不能实现随季节变化的动态变化。ECMWF 发展了一种基于集合样本离散度和均方根误差的膨胀方法, 利用前 5 d 样本序列的滑动平均来计算方差膨胀因子 (Bonavita, et al, 2011), 后续工作将考虑动态的膨胀处理技术, 更好的利用集合预报误差。

此外, 将集合预报误差用于 4DVar 目前还没有针对台风进行专门的工作, 如台风涡旋位置和强度扰动、台风涡旋重定位等。在夏季试验期间的台风路径预报没有发现改进效果, 针对台风还需要进行研究工作。

最后, 使用基于 4DVar 的 EDA 方法生成集合样本, 其最大优势是实现比较简单, 但计算量较大, 样本需要逐一生成。从业务应用上看, 发展基于集合的同化方法, 如四维集合变分同化 (4DEnVar, four dimensional ensemble variational data assimilation) 方法、DRP-4DVar (dimension reduced projection 4DVar) 方法 (Wang, et al, 2010), 来并行生

成大量的集合样本是未来发展的一个趋势, 这方面的探索工作已经在开展中。

致 谢: 感谢中国气象局数值预报中心陆慧娟博士、刘艳博士、王金成博士在全球四维变分同化基础版本上所进行的基础性工作, 以及支持完成本研究提供的帮助。

参考文献

- 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. Gong J D, Zhao G. 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*, 64(6): 684-698 (in Chinese)
- 龚建东, 张林, 王金成. 2020. 背景误差水平相关结构对四维变分资料同化的影响研究. *气象学报*, 78(6): 988-1001. Gong J D, Zhang L, Wang J C. 2020. The impact study of background error horizontal correlation structure on 4D-Var. *Acta Meteor Sinica*, 78(6): 988-1001 (in Chinese)
- 龚建东, 王瑞春. 2021. 集合预报误差在 GRAPES 全球四维变分同化的应用研究 I: 局地化方案设计与动力平衡特征分析. *气象学报*, 79(1): 31-47. Gong J D, Wang R C. 2021. The study of introducing ensemble forecast errors in the GRAPES global 4DVar. Part I: Localization scheme and dynamic balance analysis. *Acta Meteor Sinica*, 79(1): 31-47 (in Chinese)
- 刘柏年, 张卫民, 曹小群等. 2015. 集合资料同化中方差滤波技术研究及试验. *地球物理学报*, 58(5): 1526-1534. Liu B N, Zhang W M, Cao X Q, et al. 2015. Investigations and experiments of variances filtering technology in the ensemble data assimilation. *Chinese J Geophys*, 58(5): 1526-1534 (in Chinese)
- 马旭林, 薛纪善, 陆维松. 2008. GRAPES 全球集合预报的集合卡尔曼变换初始扰动方案初步研究. *气象学报*, 66(4): 526-536. Ma X L, Xue J S, Lu W S. 2008. Preliminary study on ensemble transform Kalman filter-based initial perturbation scheme in GRAPES global ensemble prediction. *Acta Meteor Sinica*, 66(4): 526-536 (in Chinese)
- 张林, 刘永柱. 2017. GRAPES 全球四维变分同化系统极小化算法预调节. *应用气象学报*, 28(2): 168-176. Zhang L, Liu Y Z. 2017. The preconditioning of minimization algorithm in GRAPES global four-dimensional variational data assimilation system. *J Appl Meteor Sci*, 28(2): 168-176 (in Chinese)
- 庄照荣, 薛纪善, 李兴良. 2011. GRAPES 集合卡尔曼滤波资料同化系统 II: 区域分析及集合预报. *气象学报*, 69(5): 860-871. Zhuang Z R, Xue J S, Li X L. 2011. The GRAPES ensemble Kalman filter data assimilation system. Part II: Regional analysis and ensemble prediction. *Acta Meteor Sinica*, 69(5): 860-871 (in Chinese)
- Anderson J L, Anderson S L. 1999. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and

- forecasts. *Mon Wea Rev*, 127(12): 2741-2758
- Bishop C H, Etherton B J, Majumdar S J. 2001. Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon Wea Rev*, 129(3): 420-436
- Bonavita M, Raynaud L, Isaksen L. 2011. Estimating background-error variances with the ECMWF ensemble of data assimilations system; Some effects of ensemble size and day-to-day variability. *Quart J Roy Meteor Soc*, 137(655): 423-434
- Bonavita M, Isaksen L, Hólm E. 2012. On the use of EDA background error variances in the ECMWF 4D-Var. *Quart J Roy Meteor Soc*, 138(667): 1540-1559
- Buehner M, Morneau J, Charette C. 2013. Four-dimensional ensemble-variational data assimilation for global deterministic weather prediction. *Nonlin Processes Geophys*, 20(5): 669-682
- Burgers G, Van Leeuwen P J, Evensen G. 1998. Analysis scheme in the ensemble Kalman filter. *Mon Wea Rev*, 126(6): 1719-1724
- Chen L L, Chen J, Xue J S, et al. 2015. Development and testing of the GRAPES regional ensemble-3DVar hybrid data assimilation system. *J Meteor Res*, 29(6): 981-996
- Clayton A M, Lorenc A C, Barker D M. 2013. Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Quart J Roy Meteor Soc*, 139(675): 1445-1461
- Courtier P, Thépaut J N, Hollingsworth A. 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart J Roy Meteor Soc*, 120(519): 1367-1387
- Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J Geophys Res Oceans*, 99(C5): 10143-10162
- Houtekamer P L, Mitchell H L. 2005. Ensemble Kalman filtering. *Quart J Roy Meteor Soc*, 131(613): 3269-3289
- Huang B, Wang X G. 2018. On the use of cost-effective Valid-Time-Shifting (VTS) method to increase ensemble size in the GFS hybrid 4DVar system. *Mon Wea Rev*, 146(9): 2973-2998
- Kleist D T, Ide K. 2015a. An OSSE-based evaluation of hybrid variational-ensemble data assimilation for the NCEP GFS. Part I: System description and 3D-hybrid results. *Mon Wea Rev*, 143(2): 433-451
- Kleist D T, Ide K. 2015b. An OSSE-based evaluation of hybrid variational-ensemble data assimilation for the NCEP GFS. Part II: 4DVar and hybrid variants. *Mon Wea Rev*, 143(2): 452-470
- Raynaud L, Berre L, Desroziers G. 2009. Objective filtering of ensemble-based background-error variances. *Quart J Roy Meteor Soc*, 135(642): 1177-1199
- Raynaud L, Berre L, Desroziers G. 2011. An extended specification of flow-dependent background error variances in the Météo-France global 4D-Var system. *Quart J Roy Meteor Soc*, 137(656): 607-619
- Tippett M K, Anderson J L, Bishop C H, et al. 2003. Ensemble square root filters. *Mon Wea Rev*, 131(7): 1485-1490
- Wang B, Liu J J, Wang S D, et al. 2010. An economical approach to four-dimensional variational data assimilation. *Adv Atmos Sci*, 27(4): 715-727
- Wang X G, Bishop C H. 2003. A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *Mon Wea Rev*, 60(9): 1140-1158
- Whitaker J S, Hamill T M. 2002. Ensemble data assimilation without perturbed observations. *Mon Wea Rev*, 130(7): 1913-1924
- Whitaker J S, Hamill T M, Wei X, et al. 2008. Ensemble data assimilation with the NCEP global forecast system. *Mon Wea Rev*, 136(2): 463-482
- Whitaker J S, Hamill T M. 2012. Evaluating methods to account for system errors in ensemble data assimilation. *Mon Wea Rev*, 140(9): 3078-3089
- Xu Q, Lu H J, Gao S T, et al. 2008a. Time-expanded sampling for ensemble Kalman filter: Assimilation experiments with simulated radar observation. *Mon Wea Rev*, 136(7): 2651-2667
- Xu Q, Wei L, Lu H J, et al. 2008b. Time-expanded sampling for ensemble-based filters: Assimilation experiments with a shallow-water equation model. *J Geophys Res Atmos*, 113(D2): D02114
- Zhang F, Snyder C, Sun J Z. 2004. Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon Wea Rev*, 132(5): 1238-1253
- Zhang L, Liu Y Z, Liu Y, et al. 2019. The operational global four-dimensional variational data assimilation system at the China Meteorological Administration. *Quart J Roy Meteor Soc*, 145(722): 1882-1896
- Zhang L H, Gong J D, Wang R C. 2018. Diagnostic analysis of various observation impacts in the 3DVAR assimilation system of global GRAPES. *Mon Wea Rev*, 146(10): 3125-3142