



数据同化组成部分对初始条件和误差增长的影响

王仲睿^{1#}, 孙浩昊^{1#}, 雷荔傑^{1,2*}, 谈哲敏¹, 张熠¹

1. 南京大学大气科学学院, 中尺度灾害性天气教育部重点实验室, 南京 210023;

2. 南京大学关键地球物质循环前沿科学中心, 南京 210023

* 通讯作者, E-mail: lililei@nju.edu.cn

共同第一作者

收稿日期: 2023-04-05; 收修改稿日期: 2023-11-12; 接受日期: 2023-11-21; 网络版发表日期: 2023-12-15

国家自然科学基金项目(批准号: 42192553、41922036、41775057)、关键地球物质前沿科学中心基金项目(编号: JBGS2102)和中央高校基础研究基金项目(编号: 0209-14380097)资助

摘要 若不考虑特定的数据同化方法, 数据同化通常可被分解为先验信息、观测算子、观测误差协方差和背景误差协方差等组成部分. 本文基于经典的Lorenz模式, 研究了数据同化各组成部分对初始条件误差和预报误差的影响, 以期设计不同尺度天气系统的数据同化方法提供理论基础. 研究表明, 预报误差经历三个典型阶段: 0~5天为预报误差的缓慢增长期; 5~15天为预报误差的快速增长期, 其中确定性预报和集合预报的误差增长速率具有显著差异; 15天后为预报误差的饱和期. 数据同化可通过提供更加准确的初始条件, 进而提升可预报性. 相比于静态背景误差协方差(B), 流依赖的背景误差协方差(P^f)可提供更精确的初始条件, 因此当瞬时观测或频繁的时间平均观测被同化时, 循环同化效果优于离线同化; 但当时间平均观测频率低时则结果相反, 这是因为循环同化在模式缺乏预报技巧时无法构造具有信息的先验估计, 且流依赖的 P^f 相比于静态的 B 不能有效地从含信息量低的观测中提取出观测信息. 瞬时观测相比于时间平均观测包含更多的信息, 因此在时间频率低的观测系统中, 瞬时观测应优先被考虑. 此外, 集合预报优于确定性预报, 且集合预报的优势在观测信息较少和模式预报技巧较低时更为显著.

关键词 资料同化, 大气可预报性, 背景误差协方差, 集合预报

1 引言

先进的数据同化(DA)为数值天气预报(NWP; 例如, Courtier等, 1994; Whitaker等, 2008; Buehner等, 2010; Bauer等, 2015; Bannister, 2017)和气候预测(例如, Meehl等, 2009; Balsamo等, 2012; Karspeck等, 2013; Massonnet等, 2013; Mochizuki等, 2016)的进步

做出了重要贡献. 数据同化基于观测、先验信息及其不确定性寻求动力系统状态的最优估计. 最优估计的系统状态可作为动力系统预报的初始条件(IC)(例如, Kalnay, 2002). 天气预报和季节性-年际(SI)气候预测可以看作是“初值问题”, 例如厄尔尼诺-南方涛动(ENSO)的预测, 但初值的重要性随着预报时效的增加而减小(Meehl等, 2009).

中文引用格式: 王仲睿, 孙浩昊, 雷荔傑, 谈哲敏, 张熠. 2024. 数据同化组成部分对初始条件和误差增长的影响. 中国科学: 地球科学, 54(1): 110–121, doi: 10.1360/N072023-0080

英文引用格式: Wang Z, Sun H, Lei L, Tan Z, Zhang Y. 2024. The importance of data assimilation components for initial conditions and subsequent error growth. Science China Earth Sciences, 67(1): 105–116, <https://doi.org/10.1007/s11430-023-1229-7>

地球系统的动力和物理过程涵盖了广阔的时间和空间尺度, 具有复杂相互作用和非线性的特征(例如, Lucarini等, 2014). 为了捕捉地球系统的特征, 需要进行跨越天气和气候时间尺度的无缝隙预测(Palmer等, 2008; Shapiro等, 2010; Ruti等, 2020), 而由天到年的预测的初始条件由数据同化提供给耦合地球系统模式(Brunet等, 2010). 为了预测涵盖不同时间尺度的非线性动力和物理过程, 特别是弥补地球系统各组成部分的有限观测的不足, 耦合数据同化被广泛应用(例如, Zhang, 2011; Yang等, 2013; Jung等, 2016; Laloyaux等, 2016; Penny和Hamill, 2017). Tardif等(2014, 2015)发现同化时间平均的大气观测可以改善大西洋经向翻转环流(AMOC)的预测, 尤其是在海洋观测不足的情况下. Lu等(2015a, 2015b)和Sun等(2020)提出了一种主平均耦合协方差(LACC)方法, 可有效利用大气快过程变量的观测更新海洋慢过程变量. 但是目前对慢时间尺度的观测如何影响快时间尺度大气变量的研究仍非常有限.

除了从年代际到季节-年际、再到次季节-季节(S2S)的无缝隙预测难题外, 数值天气预报还面临将天气预报技巧保持并延伸到更长时间尺度的挑战. Lorenz(1969a, 1969b)证明了小尺度的初始条件误差会引起大尺度误差的增长, 并且可预报性取决于平衡能谱. 通过湿过程引发的初始条件误差的快速增长会限制极端天气的可预报性(Tan等, 2004; Zhang, 2005), 且误差的向上增长经历三个阶段, 各个阶段具有从小尺度到大尺度的不同特征(Zhang等, 2007). 此外, 有证据表明一些极端天气事件的预报可以提前约7~10天(Hamill等, 2006), 并且误差增长特征随纬度变化, 这与热带、中高纬度和极地的动力学一致(Judt, 2020). Zhang等(2019)指出, 通过将初始条件误差减小一个量级, 可以将中纬度天气确定性预报的时效提升至15天.

为了给数值天气预报提供更加准确的初始条件, 数据同化持续发展, 其进展包括成功同化多种类型的观测资料, 如卫星辐射观测(例如, Geer等, 2018; Li等, 2021)、GPS无线电掩星观测(例如, Cucurull等, 2007; Poli等, 2010)、飞机数据(例如, Cardinali等, 2003; Tong等, 2018)和雷达观测(例如, Zhang等, 2009; Zeng等, 2021)等. 这些先进的观测数据可以是局地的或非局地的, 并且在观测数量、空间覆盖范围和观测频率等方面都各不相同. 同时, 数据同化方法本身也飞速发展,

从变分方法(例如, Courtier等, 1998; Rawlins等, 2007)、集合卡尔曼滤波(例如, Whitaker等, 2008; Houtekamer和Zhang, 2016)到混合集合-变分同化方法等(例如, Hamill和Snyder, 2000; Bannister, 2017). 为了捕捉从小尺度到大尺度误差的不同特征, 多尺度数据同化方法也应运而生, 这些方法通常采用迭代的同化策略或者对背景误差协方差进行分尺度的处理(例如, Xie等, 2011; Li等, 2015; Huang等, 2021). 但是, 数据同化方法对于不同观测类型的适用性、以及对误差动力学的认识仍非常有限.

与前人围绕数据同化中背景误差协方差的作用开展的研究不同, 本研究将从更广泛的视角将数据同化分解为不同的组成部分, 着重研究数据同化各组成部分对初始条件误差和误差增长的影响. 无论是使用变分同化还是集合卡尔曼滤波, 同化所得分析场主要由不同的数据同化组成部分来决定, 这些组成部分包括先验信息、由观测类型给定的前向观测算子、观测误差协方差和背景误差协方差等. 本文基于经典的Lorenz(2005)模式, 通过设计不同类型的数据同化组成部分, 在一个统一的集合框架内系统地分析了数据同化各组成部分对同化分析场和预报误差增长的影响. 理解数据同化各组成部分在不同尺度上重要性, 可以为设计相关尺度适用的数据同化策略提供理论基础.

2 数据同化组成部分

给定状态向量 $\mathbf{x}^f$ 和观测 $\mathbf{y}$ 的先验信息, 三维变分(3DVAR; Courtier等, 1998)和集合卡尔曼滤波(EnKF; Whitaker等, 2008)所得的状态向量的最佳估计 $\mathbf{x}^a$ 具有等效的形式 $\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}[\mathbf{y} - H(\mathbf{x}^f)]$. 其中, H 是前向观测算子, 将状态变量转换为观测变量. 增益矩阵 $\mathbf{K}$ 的表达式为 $\mathbf{K} = \mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R})^{-1}$, 其中 $\mathbf{P}^f$ 是背景误差协方差矩阵, $\mathbf{R}$ 是观测误差协方差矩阵, $\mathbf{H}$ 是前向观测算子的雅可比矩阵, 即前向算子相对于状态变量的偏导.

Lei等(2021)将集合成员分解为集合平均和集合扰动, 用采样的气候态集合扰动近似估计静态的背景误差协方差, 从而实现了在统一的集合框架下进行集合卡尔曼滤波(EnKF; Whitaker等, 2008)、三维变分(3DVAR; Courtier等, 1998)和混合集合-变分同化(Hamill和Snyder, 2000; Penny, 2014). 令 $\mathbf{x}^f$ 表示控制预报

的先验信息, 其既可以来自确定性预报也可以是集合预报的平均值. 而先验集合扰动 $[\mathbf{x}_1^f, \mathbf{x}_2^f, \dots, \mathbf{x}_N^f]$ (N 为集合大小) 可以通过采样气候态扰动或由短期集合预报减去其集合平均而得到. 通过将先验集合扰动加到控制预报的先验信息上, 可构造 N 个先验集合成员 $\mathbf{x}_i^f = \bar{\mathbf{x}}^f + \mathbf{x}_i^f$ ($i=1, \dots, N$). 在集合框架中, EnKF将观测 $\mathbf{y}$ 与先验集合成员相结合, 得到后验的集合平均和集合扰动:

$$\begin{aligned} \bar{\mathbf{x}}^a &= \bar{\mathbf{x}}^f + \left\{ \mathbf{X}(\mathbf{H}\mathbf{X})^T [(\mathbf{H}\mathbf{X})(\mathbf{H}\mathbf{X})^T + \mathbf{R}]^{-1} \right\} \\ &\quad \left[\mathbf{y} - \overline{H(\mathbf{x}^f)} \right] \\ &= \bar{\mathbf{x}}^f + \left[\mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R})^{-1} \right] \left[\mathbf{y} - \overline{H(\mathbf{x}^f)} \right], \end{aligned} \quad (1)$$

$$\begin{aligned} \mathbf{x}_i^a &= \mathbf{x}_i^f - \left\{ \mathbf{X}(\mathbf{H}\mathbf{X})^T \left[(\sqrt{(\mathbf{H}\mathbf{X})(\mathbf{H}\mathbf{X})^T + \mathbf{R}})^{-1} \right]^T \right. \\ &\quad \left. \times \left[\sqrt{(\mathbf{H}\mathbf{X})(\mathbf{H}\mathbf{X})^T + \mathbf{R}} + \sqrt{\mathbf{R}} \right]^{-1} \right\} \mathbf{H} \mathbf{x}_i^f \\ &= \mathbf{x}_i^f - \left\{ \mathbf{P}^f \mathbf{H}^T \left[(\sqrt{\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R}})^{-1} \right]^T \right. \\ &\quad \left. \times \left[\sqrt{(\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R})} + \sqrt{\mathbf{R}} \right]^{-1} \right\} \mathbf{H} \mathbf{x}_i^f. \end{aligned} \quad (2)$$

由背景误差协方差矩阵 $\mathbf{P}^f = \mathbf{X}\mathbf{X}^T$ 可得, 式(1)和式(2)中归一化的先验集合扰动 $\mathbf{X}[\mathbf{x}_1^f, \mathbf{x}_2^f, \dots, \mathbf{x}_N^f] / \sqrt{N-1}$ 为 $\mathbf{P}^f$ 的平方根. 给定雅可比矩阵 $\mathbf{H}$, $\mathbf{H}\mathbf{X}$ 是观测空间中归一化的集合扰动.

因此, 由式(1)和式(2)给出的数据同化可以分为四个组成部分: 先验信息($\bar{\mathbf{x}}^f$), 前向观测算子(H 和 $\mathbf{H}$), 观测误差协方差($\mathbf{R}$)和背景误差协方差($\mathbf{P}^f$). 不同于以往研究多单独应用某一种数据同化方法, 数据同化组成部分可以像搭积木一样拆分混合. 因此, 可以独立而系统地分析每个数据同化组成部分对后验误差和预报误差的影响.

成功的数值天气预报依赖于对当前大气状态的准确估计(Kalnay, 2002), 因此数值天气预报是一个初值问题. 控制预报的先验信息 $\bar{\mathbf{x}}^f$ 通常由循环数据同化和预报所得, 即从前一次同化所得的分析场出发进行下一同化窗口的确定性/集合预报. 因此, 控制预报的先验信息包含了所有过去观测、模式动力和物理过程的

信息(例如, Kalnay, 2002; Bauer等, 2015), 这也意味着模式具有较高的预报技巧. 但是对于预报技巧较低的气候模拟, 尤其是当模式预报时间尺度长于观测的时间分辨率时, 来自循环同化的先验信息可能还不如从气候态样本中随机采样的先验信息(例如, Matsikaris等, 2015; Sun等, 2022). 因此, 离线数据同化使用从状态变量的气候态分布中随机采样得到的先验信息, 不需要循环积分模式, 从而降低了计算成本. 表1中循环(Cyc)和离线(Off)的先验选择对应天气和气候的数据同化.

以集合后验为初始条件进行短期集合预报得到的集合扰动可用来估计流依赖的背景误差协方差矩阵 $\mathbf{P}^f$. 而从气候态中随机采样也可以得到集合扰动, 该集合扰动可用来近似估计变分同化常用的静态背景误差协方差矩阵 $\mathbf{B}$. 因此, 若控制预报和集合扰动 $\mathbf{X}$ 来自相同的集合预报, 由式(1)和式(2)可得到传统的EnKF的解. 而若将控制预报 $\bar{\mathbf{x}}^f$ 和基于气候态采样的集合扰动 $\mathbf{X}$ 带入式(1)和式(2)进行集合数据同化, 则可得到等效于3DVAR的解. 混合数据同化方法使用由 $\mathbf{B}$ 和 $\mathbf{P}^f$ 混合而成的背景误差协方差矩阵. 为简单起见, 表1中所示的背景误差协方差矩阵或是静态的 $\mathbf{B}(\mathbf{B})$, 或是流依赖的 $\mathbf{P}^f(\mathbf{Pb})$. 如式(1)和式(2)所示, 同化的增量受背景误差协方差矩阵与观测误差协方差矩阵两者之间的相对权重大小影响, 因此本文中未讨论观测误差协方差矩阵对数据同化后验误差和预报误差的影响.

不论是循环同化还是离线同化, 控制预报的先验信息都可来自确定性预报或集合预报. 当模式状态变量更新后, 以同化所得的后验作为初始条件进行预报. 对于确定性预报, 从后验开始进行确定性预报得到下一个时刻的控制预报的先验; 对于集合预报, 从集合后验场进行集合预报得到下一个时刻的控制预报的集合先验. 表1以确定性(Det)和集合(Ens)标注这两种类型. 由于使用集合预报时很少完全使用静态 $\mathbf{B}$, 且在离线数据同化时也通常不会进行集合预报, 因此没有设计使用静态 $\mathbf{B}$ 的循环集合同化和使用流依赖 $\mathbf{P}^f$ 的离线集合同化这两组试验.

为了研究不同尺度的观测数据, 设计了两种类型的观测, 分别为瞬时观测和时间平均观测. 瞬时观测为给定时间的模式格点的值, 以模拟传统的天气观测. 时间平均观测为给定时间窗口内模式格点的时间平均, 以模拟时间尺度更长的观测, 如海表温度等. 瞬时

表 1 数据同化试验的设置^{a)}

试验名称	同化循环	背景误差协方差矩阵	观测类型	预报类型
Cyc_Pb_Ens_Inst_x	循环	流依赖 $\mathbf{P}^f$	瞬时	集合
Cyc_Pb_Det_Inst_x	循环	流依赖 $\mathbf{P}^f$	瞬时	确定性
Cyc_B_Det_Inst_x	循环	静态 $\mathbf{B}$	瞬时	确定性
Off_Pb_Det_Inst_x	离线	流依赖 $\mathbf{P}^f$	瞬时	确定性
Off_B_Ens_Inst_x	离线	静态 $\mathbf{B}$	瞬时	集合
Off_B_Det_Inst_x	离线	静态 $\mathbf{B}$	瞬时	确定性
Cyc_Pb_Ens_Tavg_x	循环	流依赖 $\mathbf{P}^f$	时间平均	集合
Cyc_Pb_Det_Tavg_x	循环	流依赖 $\mathbf{P}^f$	时间平均	确定性
Cyc_B_Det_Tavg_x	循环	静态 $\mathbf{B}$	时间平均	确定性
Off_Pb_Det_Tavg_x	离线	流依赖 $\mathbf{P}^f$	时间平均	确定性
Off_B_Ens_Tavg_x	离线	静态 $\mathbf{B}$	时间平均	集合
Off_B_Det_Tavg_x	离线	静态 $\mathbf{B}$	时间平均	确定性

a) 试验名称中的x表示观测频率(即同化窗长度), 包括6小时、1天、5天和10天

观测的观测算子及其雅可比矩阵为空间插值, 而时间平均观测的观测算子及其雅可比矩阵则是空间和时间的插值. 表1中的瞬时(Inst)和时间平均(Tavg)观测可分别对应大气和海洋的观测资料. 当观测频率增加时, 时间平均观测趋近于瞬时观测.

3 试验设计

本文试验使用单尺度的Lorenz (2005) 模式II来检验数据同化各组成部分对后验误差和预报误差的影响. Lorenz (2005) 模式可以模拟在某一纬度和高度的标量大气变量, 且能够很好地表征大气特征波长和误差增长速率, 已被广泛应用于数据同化的研究(例如, Brajard 等, 2020; Lei 等, 2021). 单尺度Lorenz (2005) 模式II的控制方程为

$$\frac{dX_n}{dt} = [X, X]_{K,n} - X_n + F, \quad (3)$$

其中下标 n 为模式格点的序号(共有960个模式格点), K 和 F 是常数, 分别设置为32和16. 平流项 $[X, X]_{K,n}$ 为

$$[X, X]_{K,n} = -W_{n-2K}W_{n-K} + \sum_{j=-J}^J W_{n-K+j}X_{n+K+j}/K, \quad (4)$$

其中, K 为偶数时 $J=K/2$, $W_n = \sum_{i=-J}^J X_{n-i}/K$; K 为奇数时

$J=(K-1)/2$, $W_n = \sum_{i=-J}^J X_{n-i}/K$. $\sum'$ 是一个特殊的求和算

子, 除首项与末项需除以2外与通常的求和算子相同.

观测模拟系统试验采用完美模式假设, 真值和模拟采用同样的试验参数. 真值、确定性和集合预报的初始条件均从气候态样本中抽取. 模拟观测通过在真值上添加正态分布 $N(0, R)$ 的随机扰动而得. 同化试验使用两种类型的观测, 观测网络的密度为每8个模式格点一个观测, 观测误差的方差 R 为2.0. 瞬时观测为同化窗中间时刻的格点值, 时间平均观测为同化窗内格点值的时间平均. 观测密度和误差方差的选取范围介于天气和气候观测密度和误差之间. 观测频率包括6小时、1天、5天和10天. 对于Lorenz (2005)模式, 0.05无量纲时间单位大约等于6小时. 同化窗口的设置与观测频率相同, 使得观测频率增加时可以同化更多的观测数据, 减小分析误差.

集合数据同化和预报的集合数目为40. 为减少因有限的集合数目导致的样本误差, 对于使用流依赖的背景误差协方差矩阵 $\mathbf{P}^f$ 的同化试验, 应用协方差膨胀和协方差局地化. 协方差膨胀采用常系数乘法膨胀(Anderson和Anderson, 1999), 以增加集合离散度、防止滤波发散. Gaspari和Cohn(GC; Gaspari和Cohn, 1999)函数被用于限制观测的影响, 减小观测与远距离状态变量间的虚假相关. 对于每个同化试验, 协方差膨胀和局地化参数均使用最优参数. 使用静态背景误差协方差矩阵 $\mathbf{B}$ 的同化试验未应用协方差膨胀和局地化, 因为气候态扰动的集合数目足以准确估计静态 $\mathbf{B}$. 对

于各频率的瞬时观测及6小时和1天频率的时间平均观测, 气候态扰动的样本数量级为 10^4 . 对于5天和10天频率的时间平均观测, 由于时间平均, 气候态扰动的样本数量级为 10^3 .

同化试验的时长为500天, 前50天的同化试验被舍弃以避免误差未饱和和对结果的影响, 将剩余450天同化试验的确定性后验或集合后验的平均与真值比较, 以评估同化各组成部分的影响. 为比较各同化组分对预报误差增长的影响, 以同化分析场为初始条件启动自由预报. 为了使不同同化窗口的预报样本数相同, 从剩余450天同化窗中间每隔10天启动一个自由预报, 每个自由预报的时长为半个同化窗加20天. 同化试验的具体设置如图1所示. 每6小时计算确定性预报或集合预报时空平均的均方根(RMS)误差. 为了检查误差差异的显著性, 基于45个样本(450天中每10天一个)进行成对样本 t 检验, 误差差异在99%置信水平上显著认为具有统计显著性. 因为不同数据同化组成部分提供的初始条件不同, 因而不同同化试验的预报误差增长方式不同. 为了公平地比较不同数据同化组成部分的预报误差增长, 每个预报误差曲线的起始点被移动到最接近参考试验预报误差的时间点. 同化试验组和相关的参考试验参见下文. 误差增长分为三个阶段: 第一个阶段是从0小时到5天, 第二个阶段是从5天到大约15天, 最后为误差饱和阶段, 大约为15天后. 第二个误差增长阶段的结束时间点由线性回归的最小残差来确定.

4 数据同化组成部分对后验误差的影响

首先讨论同化瞬时观测的试验结果. 使用流依赖背景误差协方差矩阵 $\mathbf{P}^f$ 进行确定或集合循环同化(Cyc_Pb_Det_Inst 和 Cyc_Pb_Ens_Inst), 后验误差会随着观测频率的降低而增大, 这是因为当观测变得稀疏时观测信息也随之变少(图2a和表2). 同样地, 使用静态 $\mathbf{B}$ 进行确定性循环同化(Cyc_B_Det_Inst)的后验误差随着观测频率的降低而增大. 其次, 对于确定性循环同化, 使用 $\mathbf{P}^f$ 的后验误差比使用静态 $\mathbf{B}$ 的确定性循环同化的后验误差更小, 尤其是当观测频率较高时. 无论是使用 $\mathbf{P}^f$ 的循环同化或静态 $\mathbf{B}$ 的离线同化, 集合同化的后验误差均显著小于确定性同化的后验误差. 此外, 在不同观测频率下, 使用静态 $\mathbf{B}$ 的确定性循环同化比

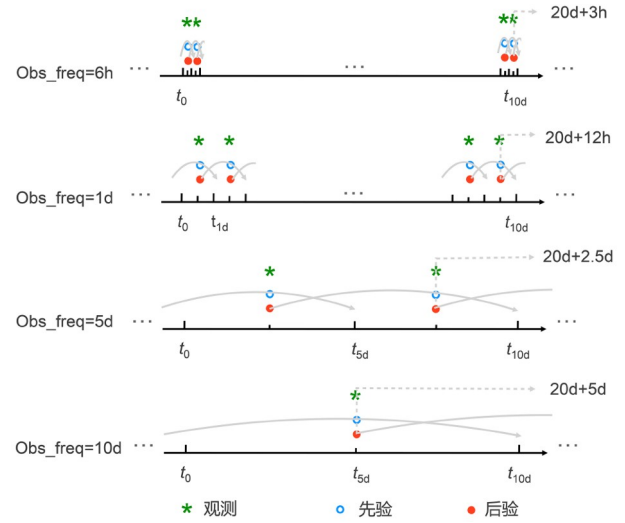


图1 观测频率分别为6小时、1天、5天和10天的同化试验示意图

灰色实线表示循环同化和预报, 虚线表示每隔10天启动一次的自由预报, 预报时长为20天加上半个同化窗口

确定性或集合的离线同化的后验误差小(Off_B_Det_Inst/Off_B_Ens_Inst), 并且对于较高的观测频率(6小时和1天)时误差差异是显著的. 但是, 使用 $\mathbf{P}^f$ 的确定性离线同化(Off_Pb_Det_Inst)的后验误差显著大于其他同化方法, 特别是当观测频率较高时. 这是因为 $\mathbf{P}^f$ 会低估离线采样的先验误差, 而后者在量级上约等于气候态误差, 因此较小的流依赖的集合离散度不能够代表离线先验的较大不确定性, 从而导致过小的同化增量, 不能有效地修正先验误差. 因此, 对于不同观测频率的瞬时观测, 集合同化优于确定性同化, 循环同化优于离线同化. 并且流依赖的 $\mathbf{P}^f$ 较静态 $\mathbf{B}$ 对后验误差有更大的影响, 特别是对于频繁的观测.

同化时间平均观测的结果与同化瞬时观测的结果相似(图2b和表2). 但前者的后验误差比后者大, 尤其是观测频率较低时. 说明时间平均观测包含的信息比瞬时观测少, 而且平均时间长度越长, 观测信息损失的越多. 与瞬时观测相比, 当观测频率较低时(5天和10天), 集合同化相对于确定性同化的优势更加明显. 这与Sun等(2022)的结果一致, 因为当预报时长较长且观测较少时, 模式的预报技巧较低, 进行集合平均对于构造先验信息非常重要. 与先前同化瞬时且高频观测不同, 当同化低频的时间平均观测时, 使用静态 $\mathbf{B}$ 会优于流依赖的 $\mathbf{P}^f$, 同时使用离线同化可优于循环同化. 因为

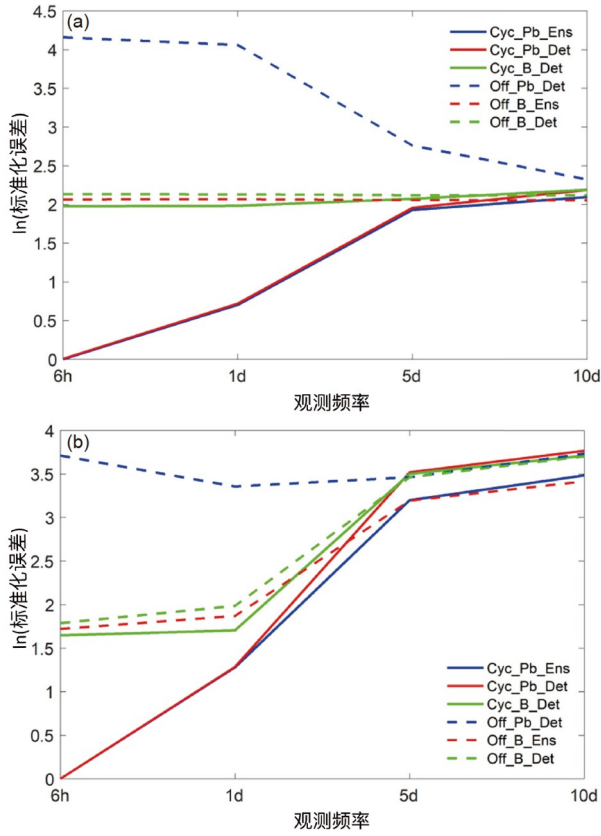


图2 不同观测频率时数据同化试验所得的后验误差

(a) 同化瞬时观测; (b) 同化时间平均观测. 误差均以参考试验Cyc_Pb_Ens_Inst_6h的后验误差为标准进行归一化处理. 具体数值见表2

当模式缺乏预报技巧且观测包含较长时间平均的信息时, 循环同化构造的先验信息不如从气候态分布中随机采样得到的先验估计, 同时, 静态 $\mathbf{B}$ 相比流依赖的 $\mathbf{P}^f$ 能更好地提取观测信息. 因此, 对于低频的观测系统, 瞬时观测优于时间平均观测. 为了更好地从低频的时间平均观测中提取信息, 需要使用集合先验, 而此时使用静态 $\mathbf{B}$ 进行离线同化比使用流依赖 $\mathbf{P}^f$ 进行循环同化更有优势.

5 数据同化组成部分对预报误差增长的影响

5.1 使用流依赖 $\mathbf{P}^f$ 的循环同化

图3a为使用流依赖 $\mathbf{P}^f$ 循环同化瞬时观测的试验(Cyc_Pb_Ens_Inst 和 Cyc_Pb_Det_Inst)所得的平均预报误差的时间序列. 由于各同化试验的后验误差各不相同, 各同化试验的平均预报误差的起始点均移至其

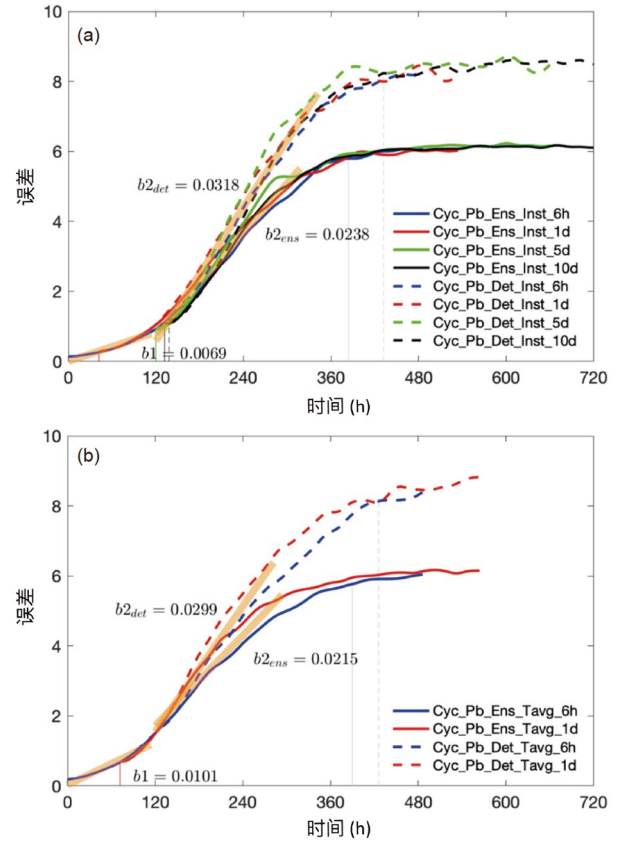


图3 使用流依赖的 $\mathbf{P}^f$ 进行循环同化和预报的误差

(a) 同化瞬时观测; (b) 同化时间平均观测. 实线和虚线分别表示集合和确定性预报. 蓝色、红色、绿色和黑色分别表示同化的观测频率为6小时、1天、5天和10天. 橙色粗线为线性回归线, 其斜率标注在一侧. 竖直灰色实线和虚线分别标注集合和确定性预报的误差达到其气候态误差的95%的时间. 其他竖直线标注相应的同化试验的后验误差等于参考试验(a)Cyc_Pb_Ens_Inst_6h 和(b)Cyc_Pb_Ens_Tavg_6h预报误差的时间

后验误差最接近参考试验 Cyc_Pb_Ens_Inst_6h 的预报误差的时刻. 单个模式格点的预报误差与模式格点空间平均的预报误差的增长模态相似, 并且集合预报的集合平均误差和集合离散度的量级也一致(图略). 由图3可知, 预报误差的增长经历了三个典型的阶段. 对于6小时和1天的观测频率, 初始条件(即后验)的误差量级为 10^{-1} , 误差以 0.0069h^{-1} 的速率缓慢增长, 直至第5天. 第5天时, 同化6小时和1天观测频率试验的预报误差以及同化5天和10天观测频率试验的初始条件误差量级均为 10^0 . 第5天到约第15天, 确定性和集合预报的预报误差增长速率不同. 集合预报的误差增长速率

表 2 不同同化试验设置、不同观测频率同化所得的后验误差

观测频率	6小时	1天	5天	10天
Cyc_Pb_Ens_Inst	0.1347	0.2717	0.9271	1.0942
Cyc_Pb_Det_Inst	0.1356	0.2768	0.9539	1.1995
Cyc_B_Det_Inst	0.9722	0.9782	1.0699	1.2029
Off_Pb_Det_Inst	8.6262	7.8057	2.1304	1.3742
Off_B_Ens_Inst	1.0615	1.0639	1.0519	1.0494
Off_B_Det_Inst	1.1371	1.1342	1.1219	1.1205
Cyc_Pb_Ens_Tavg	0.1910	0.6863	4.6809	6.2307
Cyc_Pb_Det_Tavg	0.1912	0.6904	6.4494	8.2697
Cyc_B_Det_Tavg	0.9923	1.0501	6.3407	7.8168
Off_Pb_Det_Tavg	7.8179	5.4851	6.0992	7.9959
Off_B_Ens_Tavg	1.0672	1.2390	4.6599	5.8193
Off_B_Det_Tavg	1.1420	1.3916	6.0949	7.7589

为 0.0238h^{-1} ,而确定性预报的误差增长速率为 0.0318h^{-1} ,比集合预报快33.6%. 相较于确定性预报,集合预报的优势可通过分析预报误差得到更好体现. 大约15天后,确定性和集合预报的预报误差逐渐饱和至气候态误差,其数值分别为8.55和6.16. 集合预报误差小于确定性预报误差,这与预期的结果一致,因为当误差饱和时,集合平均的方差比任何一个确定性预报的方差小 N 倍. 集合(确定性)预报的误差在384小时(432小时)时增长到气候态误差的95%.

图3b为使用流依赖 $\mathbf{P}^f$ 循环同化时间平均观测(Cyc_Pb_Ens_Tavg 和 Cyc_Pb_Det_Tavg)所得的平均预报误差的时间序列,其所得预报误差特征与同化瞬时观测相似,均为三个典型的误差增长阶段. 同化6小时观测频率的初始误差量级为 10^{-1} ,并以 0.0101h^{-1} 的速率缓慢增长,直至第5天. 该误差增长率大于同化瞬时观测的试验,因为其初始场误差大于同化瞬时观测的初始场误差. 第5天时,同化6小时和1天观测频率的预报误差量级达到 10^0 ;第5天到约第15天,集合和确定性预报的误差分别以 0.0215h^{-1} 和 0.0299h^{-1} 的速率增长. 同化时间平均观测得到的第二阶段误差增长率比同化瞬时观测小,因为前者第5天的预报误差大于后者. 确定性预报的误差增长速率比集合预报快39.1%. 因此在同化时间平均观测时,集合预报相对于确定性预报的优势仍然存在. 第15天后,集合预报和确定性预报的预报误差逐渐达到气候态误差. 在390小时(426小时),集合(确定性)预报的预报误差达到气候态误差的95%. 不

论是同化瞬时观测还是时间平均观测,对于不同的观测频率,预报误差在相似的预报时长后达到饱和.

不同于同化瞬时观测,采用流依赖 $\mathbf{P}^f$ 循环同化(Cyc_Pb_Ens_Tavg 和Cyc_Pb_Det_Tavg)观测频率较低的时间平均观测(5天和10天),得到更大的后验误差(图4). 并且,对于使用静态 $\mathbf{B}$ 的循环同化和使用静态 $\mathbf{B}$ 或流依赖 $\mathbf{P}^f$ 的离线同化,也得到相似的结果(图4). 由于同化观测频率为5天和10天的时间平均观测的初始条件误差较大,预报误差迅速达到气候态误差水平. 因此无论采用哪种同化策略,同化5天和10天观测频率的时间平均观测,预报误差不会经历前两个误差增长阶段.

5.2 使用静态 $\mathbf{B}$ 的离线同化

图3中使用流依赖的 $\mathbf{P}^f$ 循环同化频率越低的观测,所得后验误差越大. 但与使用静态 $\mathbf{B}$ 离线同化不同,同化观测频率为6小时到10天的瞬时观测或者观测频率为6小时和1天的时间平均观测(Off_B_Ens_Inst 和 Off_B_Det_Inst)时,所得的后验误差相似(图5). 通过将同化试验的预报误差时间序列的起始点移动到最接近参考试验 Cyc_Pb_Ens_Inst_6h(图5a)和 Cyc_Pb_Ens_Tavg_6h(图5b)预报误差的时间,可以看出使用静态 $\mathbf{B}$ 离线同化的预报误差经历第二和第三阶段,但不经历图3中的第一阶段. 因为使用静态 $\mathbf{B}$ 同化瞬时和时间平均观测的初始场误差均为 10^0 ,比使用流依赖的 $\mathbf{P}^f$ 循环同化的初始场误差大一个数量级,因此使用静态 $\mathbf{B}$ 离线同化试验并不会像图3一样经历第一阶段的误差

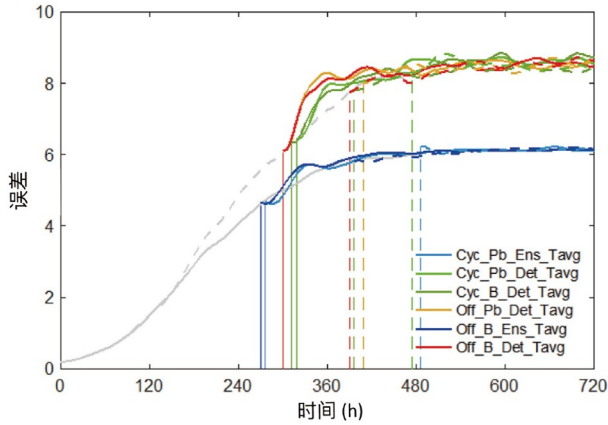


图 4 同化观测频率为5天和10天的时间平均观测的预报误差
实线表示5天观测频率, 虚线表示10天观测频率. 灰色的实线和虚线分别表示使用流依赖 $\mathbf{P}^f$ 循环同化时间平均观测(Cyc_Pb_Ens_Tavg和Cyc_Pb_Det_Tavg)的参考试验. 竖直线标记相应试验的后验误差约等于参考试验预报误差的时间

增长. 第5天到约第15天, 确定性预报和集合预报的预报误差增长速率不同. 同化瞬时观测, 集合和确定性预报的误差增长率分别为 0.0268h^{-1} 和 0.0346h^{-1} . 同化时间平均观测, 集合和确定性预报的误差增长率分别为 0.0296h^{-1} 和 0.0386h^{-1} . 确定性预报的误差增长率比集合预报在同化瞬时观测和时间平均观测试验时分别快29.1%和30.4%. 故在使用静态 $\mathbf{B}$ 离线同化时, 集合预报优于确定性预报. 使用静态 $\mathbf{B}$ 离线同化在第二阶段的误差增长率高于使用流依赖的 $\mathbf{P}^f$ 循环同化, 因此证明流依赖的 $\mathbf{P}^f$ 循环同化能够更好地限制预报误差增长. 与使用流依赖的 $\mathbf{P}^f$ 循环同化类似, 使用静态 $\mathbf{B}$ 离线同化的确定性和集合预报误差在约15天后逐渐达到气候态误差.

5.3 使用交叉组分的同化

与使用静态 $\mathbf{B}$ 的离线同化以及使用流依赖的 $\mathbf{P}^f$ 的循环同化相比, 循环同化可以使用静态 $\mathbf{B}$, 而离线同化也可以利用流依赖的 $\mathbf{P}^f$. 因此类似于前述讨论, 将使用静态 $\mathbf{B}$ 的确定性循环同化(Cyc_B_Det_Inst/Tavg)和使用流依赖 $\mathbf{P}^f$ 的确定性离线同化(Off_Pb_Det_Inst/Tavg)的预报误差时间序列的起始点移到最接近参考试验Cyc_Pb_Det_Inst_6h(图6a)和Cyc_Pb_Det_Tavg_6h(图6b)预报误差的时间. 可以看出使用交叉组分同化与使用静态 $\mathbf{B}$ 的确定性离线同化的预报误差相似, 但

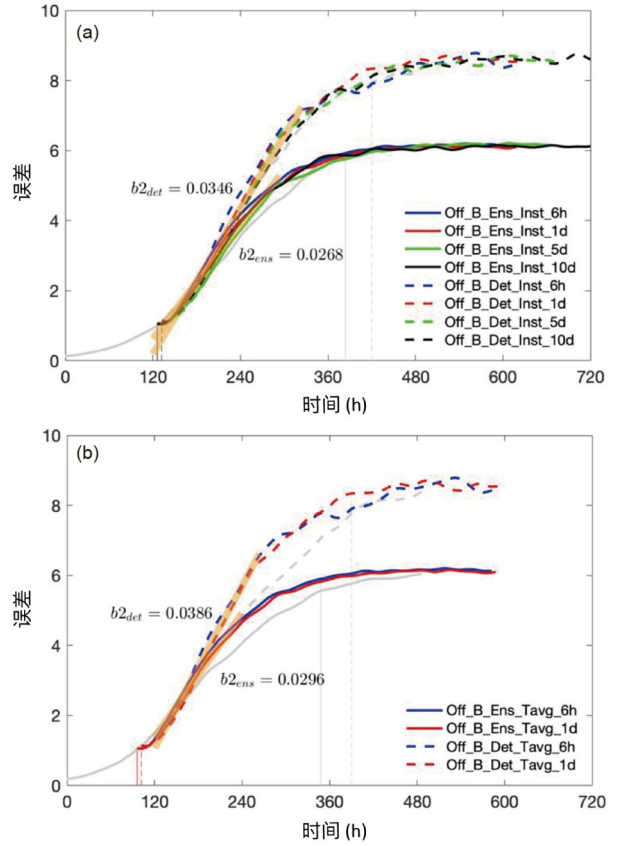


图 5 使用静态 $\mathbf{B}$ 进行离线数据同化和预报的误差

(a) 同化瞬时观测; (b) 同化时间平均观测. 灰色实线和虚线分别表示使用流依赖的 $\mathbf{P}^f$ 进行集合和确定性预报的6小时循环同化的参考试验

不包含使用流依赖的 $\mathbf{P}^f$ 离线同化较频繁观测的情况, 因为此时有较大的初始误差. 使用交叉组分同化的预报误差也经历第二和第三增长阶段. 由于初始误差量级为 10^0 , 因此交叉组分同化没有经历类似于使用流依赖的 $\mathbf{P}^f$ 循环同化的第一阶段的误差增长. 第5天到约第15天, 确定性同化瞬时观测和时间平均观测的预报误差分别以 0.0334h^{-1} 和 0.0371h^{-1} 的速率增长, 这与使用静态 $\mathbf{B}$ 确定性离线同化的误差增长率相似. 约15天后, 交叉组分同化的确定性预报误差逐渐达到气候态误差.

使用静态 $\mathbf{B}$ 离线同化和使用交叉组分同化的初始条件误差至少比使用流依赖的 $\mathbf{P}^f$ 循环同化的误差大一个数量级. 因此, 上述的同化配置不同于使用流依赖的 $\mathbf{P}^f$ 循环同化经历预报误差缓慢增长的第一阶段, 只经历第二和第三阶段, 其中对于第二阶段误差增长速率, 使用静态 $\mathbf{B}$ 离线同化比使用交叉组分同化更快. 因此,

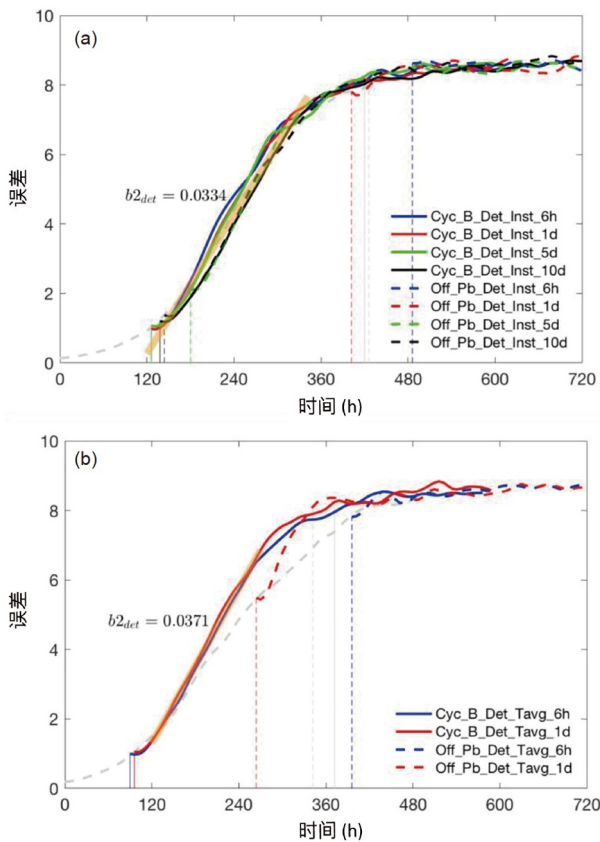


图6 使用交叉组分同化的确定性预报误差

(a) 同化瞬时观测; (b) 同化时间平均观测. 橙色粗线是线性回归线, 其斜率标注在一侧. 灰色竖直线标记使用静态 B 确定性循环同化(实线)和使用流依赖的 P^f 离线同化(虚线)的误差分别接近其95%气候误差的时间. 其他竖直线标记相应的同化试验的后验误差等于参考试验(使用流依赖的 P^f 确定性循环同化观测频率为6小时的观测)预报误差的时间

在第二阶段使用流依赖的 P^f 循环同化, 相比其他同化配置可以更好地约束误差增长.

6 总结与讨论

为了更好地理解同化不同组成部分对初始条件和预报误差增长的影响, 可将数据同化从更广的视角分解成四个组成部分, 即先验信息, 观测前向算子, 观测误差协方差以及背景误差协方差. 背景误差协方差与观测误差协方差两者之间相对权重大小决定同化增量的量级, 因此本文中未考虑观测误差协方差, 只研究了背景误差协方差、先验信息与观测前向算子对初始条

件和预报误差增长的影响.

在Lorenz(2005)模式II中, 预报误差从量级为 10^{-1} 开始经历三个典型阶段: 第0天到第5天, 预报误差以小于 $0.01h^{-1}$ 的速率缓慢增长; 第5天到第15天, 集合预报和确定性预报的误差分别以 $0.02h^{-1}$ 和 $0.03h^{-1}$ 的速率快速增长; 15天以后, 预报误差达到饱和的气候态误差. 数据同化可通过提供更加准确的初始条件, 进而提升可预报性. 当同化更频繁的观测时, 使用流依赖 P^f 循环同化比使用静态 B 离线同化或者交叉组分同化(使用静态 B 的循环同化、使用流依赖的 P^f 的循环同化)的初始误差更小, 因此使用流依赖 P^f 的循环同化可较其他同化配置获得5天的可预报性. 但是当同化信息较少的观测时(比如5天或10天的时间平均观测), 初始条件因为不能被任何同化配置很好地限制, 从而无法延长可预报性. 并且由于第一阶段的误差增长很慢, 通过对初始条件误差的限制可以对短时期预报技巧带来巨大提升.

在循环或离线同化中, 可以从确定性预报或集合预报中构造先验信息. 对于后验误差, 集合先验信息相比于确定性先验信息更具优势, 并且集合先验信息的优势在同化信息量低的观测、以及模式预报技巧低的先验信息时表现得更大. 同时, 集合预报也比确定性预报更有优势, 这种优势相比后验误差更显著, 当预报误差饱和时, 集合预报相比确定性预报的优势达到最大. 对于频繁的观测, 循环同化优于离线同化, 但对于低频的观测则相反. 这是因为当模式预报技巧较低、同时同化几乎无信息的观测时, 从气候态分布中随机取样构造的先验相比循环同化预报所得的先验包含的信息更多. 因此, 循环同化的重要性随着动力系统的时间尺度增大而降低.

背景误差协方差可以是流依赖 P^f 或者是静态 B . 使用循环同化, 即更多信息的先验, 流依赖 P^f 的后验误差一般比静态 B 小, 说明采样估计的流依赖 P^f 相比于静态 B 能更好地提取观测信息. 但是在离线同化中, 即更少信息的先验, 静态 B 相比于取样估计的流依赖 P^f 能更好地提取观测信息, 其后验误差一般小于使用流依赖 P^f . 使用静态 B 离线同化与交叉组分同化在第二阶段的预报误差增长速率是类似的, 但都大于使用流依赖 P^f 循环同化. 因此使用流依赖 P^f 循环同化相比于使用静态 B 的离线同化和交叉组分同化, 可以更好地限制预报误差的增长.

关于观测前向算子(即观测类型), 6小时到10天的时间平均观测比相应的瞬时观测含有更少的信息. 当增大观测的平均窗口, 即观测频率降低, 观测信息以及模式的可预报性会损失更多. 第二阶段的预报误差增长率由阶段开始时的初始误差以及阶段持续时间决定. 对于各种不同配置的同化组成成分, 同化时间平均观测相比于瞬时观测的初始误差更大(图3), 第二阶段预报误差增长持续时间更短(图5和图6). 因此, 相比于时间平均观测, 同化瞬时观测更有效, 特别是对于低频的观测系统. 这些结论对分钟到秒级的超频繁观测情景可能不适用, 因为高频观测的扰动会导致较大的观测噪声从而会干扰观测信息的提取.

除了以上讨论的四种同化组成部分, 模式误差也是同化的重要组成部分之一, 需要进一步研究其对同化及后续预报的影响. 同化组成部分对后验误差和预报误差的影响与模式的预报技巧相关, 需要在耦合模式中探索各同化组成部分的重要性. 此外, 本文使用的是一个经典但简单的Lorenz模式, 同化组成部分在真实模式中的影响仍需检验.

致谢 感谢审稿人提出的宝贵修改意见. 文中用于产生真值、观测以及同化试验的代码和输出数据, 均可参见网页 <https://meso.nju.edu.cn/xwdt/20220718/i226029.html>.

参考文献

- 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: 2741–2758
- Balsamo G, Albergel A, Beljaars A, Boussetta S, Brun E, Cloke H L, Dee D. 2012. ERA-Interim/Land: A global land-surface reanalysis based on ERA-Interim meteorological forcing. ERA-Report series: 13
- Bannister R N. 2017. A review of operational methods of variational and ensemble-variational data assimilation. *Quart J R Meteorol Soc*, 143: 607–633
- Bauer P, Thorpe A, Brunet G. 2015. The quiet revolution of numerical weather prediction. *Nature*, 525: 47–55
- Brajard J, Carrassi A, Bocquet M, Bertino L. 2020. Combining data assimilation and machine learning to emulate a dynamical model from sparse and noisy observations: A case study with the Lorenz 96 model. *J Comput Sci*, 44: 101171
- Brunet G, Shapiro M, Hoskins B, Moncrieff M, Dole R, Kiladis G N, Kirtman B, Lorenc A, Mills B, Morss R, Polavarapu S, Rogers D, Schaake J, Shukla J. 2010. Collaboration of the weather and climate communities to advance subseasonal-to-seasonal prediction. *Bull Amer Meteor Soc*, 91: 1397–1406
- Buehner M, Houtekamer P L, Charette C, Mitchell H L, He B. 2010. Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: Description and single-observation experiments. *Mon Weather Rev*, 138: 1550–1566
- Cardinali C, Isaksen L, Andersson E. 2003. Use and impact of automated aircraft data in a global 4DVAR data assimilation system. *Mon Weather Rev*, 131: 1865–1877
- Courtier P, Thepaut J N, Hollingsworth A. 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart J R Meteorol Soc*, 120: 1367–1387
- Courtier P, Andersson E, Heckley W, Vasiljevic D, Hamrud M, Hollingsworth A, Rabier F, Fisher M, Pailleux J. 1998. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Quart J R Meteorol Soc*, 124: 1783–1807
- Cucurull L, Derber J C, Treadon R, Purser R J. 2007. Assimilation of Global Positioning System radio occultation observations into NCEP's global data assimilation system. *Mon Weather Rev*, 135: 3174–3193
- Gaspari G, Cohn S E. 1999. Construction of correlation functions in two and three dimensions. *Quart J R Meteorol Soc*, 125: 723–757
- Geer A J, Lonitz K, Weston P, Kazumori M, Okamoto K, Zhu Y, Liu E H, Collard A, Bell W, Migliorini S, Chambon P, Fourrié N, Kim M, Köpken-Watts C, Schraff C. 2018. All-sky satellite data assimilation at operational weather forecasting centres. *Quart J R Meteorol Soc*, 144: 1191–1217
- Hamill T M, Snyder C. 2000. A hybrid ensemble Kalman filter–3D variational analysis scheme. *Mon Wea Rev*, 128: 2905–2919
- Hamill T M, Whitaker J S, Mullen S L. 2006. Reforecasts: An important dataset for improving weather predictions. *Bull Amer Meteor Soc*, 87: 33–46
- Houtekamer P L, Zhang F. 2016. Review of the ensemble kalman filter for atmospheric data assimilation. *Mon Weather Rev*, 144: 4489–4532
- Huang B, Wang X, Kleist D T, Lei T. 2021. A simultaneous multiscale data assimilation using scale-dependent localization in GSI-based hybrid 4DVar for NCEP FV3-based GFS. *Mon Weather Rev*, 149: 479–501
- Judt F. 2020. Atmospheric predictability of the tropics, middle latitudes, and polar regions explored through global storm-resolving simulations. *J Atmos Sci*, 77: 257–276
- Jung T, Gordon N D, Bauer P, Bromwich D H, Chevallier M, Day J J, Dawson J, Doblas-Reyes F, Fairall C, Goessling H F, Holland M,

- Inoue J, Iversen T, Klebe S, Lemke P, Losch M, Makshtas A, Mills B, Nurmi P, Perovich D, Reid P, Renfrew I A, Smith G, Svensson G, Tolstykh M, Yang Q. 2016. Advancing polar prediction capabilities on daily to seasonal time scales. *Bull Am Meteorol Soc*, 97: 1631–1647
- Kalnay E. 2002. Atmospheric Modeling, Data Assimilation, and Predictability. Cambridge: Cambridge University Press. 341
- Karspeck A R, Yeager S, Danabasoglu G, Hoar T, Collins N, Raeder K, Anderson J, Tribbia J. 2013. An ensemble adjustment kalman filter for the CCSM4 ocean component. *J Clim*, 26: 7392–7413
- Lalouaux P, Balmaseda M, Dee D, Mogensen K, Janssen P. 2016. A coupled data assimilation system for climate reanalysis. *Quart J R Meteorol Soc*, 142: 65–78
- Lei L, Wang Z, Tan Z M. 2021. Integrated hybrid data assimilation for an ensemble kalman filter. *Mon Weather Rev*, 149: 4091–4105
- Li J, Geer A J, Okamoto K, Otkin J A, Liu Z, Han W, Wang P. 2021. Satellite all-sky infrared radiance assimilation: Recent progress and future perspectives. *Adv Atmos Sci*, 39: 9–21
- Li Z, McWilliams J C, Ide K, Farrara J D. 2015. A multiscale variational data assimilation scheme: Formulation and illustration. *Mon Weather Rev*, 143: 3804–3822
- Lorenz E N. 1969a. Three approaches to atmospheric predictability. *Bull Amer Meteor Soc*, 50: 345–349
- Lorenz E N. 1969b. The predictability of a flow which possesses many scales of motion. *Tellus A-Dynamic Meteor Oceanography*, 21: 289–307
- Lorenz E N. 2005. Designing chaotic models. *J Atmos Sci*, 62: 1574–1587
- Lu F, Liu Z, Zhang S, Liu Y. 2015a. Strongly coupled data assimilation using leading averaged coupled covariance (LACC). Part I: Simple model study. *Mon Weather Rev*, 143: 3823–3837
- Lu F, Liu Z, Zhang S, Liu Y, Jacob R. 2015b. Strongly coupled data assimilation using leading averaged coupled covariance (LACC). Part II: CGCM experiments. *Mon Weather Rev*, 143: 4645–4659
- Lucarini V, Blender R, Herbert C, Ragone F, Pascale S, Wouters J. 2014. Mathematical and physical ideas for climate science. *Rev Geophys*, 52: 809–859
- Massonnet F, Mathiot P, Fichet T, Goosse H, König Beatty C, Vancoppenolle M, Lavergne T. 2013. A model reconstruction of the Antarctic sea ice thickness and volume changes over 1980–2008 using data assimilation. *Ocean Model*, 64: 67–75
- Matsikaris A, Widmann M, Jungclaus J. 2015. On-line and off-line data assimilation in palaeoclimatology: A case study. *Clim Past*, 11: 81–93
- Meehl G A, Goddard L, Murphy J, Stouffer R J, Boer G, Danabasoglu G, Dixon K, Giorgetta M A, Greene A M, Hawkins E, Hegerl G, Karoly D, Keenlyside N, Kimoto M, Kirtman B, Navarra A, Pulwarty R, Smith D, Stammer D, Stockdale T. 2009. Decadal prediction. *Bull Amer Meteor Soc*, 90: 1467–1486
- Mochizuki T, Masuda S, Ishikawa Y, Awaji T. 2016. Multiyear climate prediction with initialization based on 4D-Var data assimilation. *Geophys Res Lett*, 43: 3903–3910
- Palmer T N, Doblas-Reyes F J, Weisheimer A, Rodwell M J. 2008. Toward Seamless prediction: Calibration of climate change projections using seasonal forecasts. *Bull Amer Meteor Soc*, 89: 459–470
- Penny S G. 2014. The hybrid local ensemble transform Kalman filter. *Mon Weather Rev*, 142: 2139–2149
- Penny S G, Hamill T M. 2017. Coupled data assimilation for integrated Earth system analysis and prediction. *Bull Am Meteorol Soc*, 98: ES169–ES172
- Poli P, Healy S B, Dee D P. 2010. Assimilation of Global Positioning System radio occultation data in the ECMWF ERA-Interim reanalysis. *Quart J R Meteorol Soc*, 136: 1972–1990
- Rawlins F, Ballard S P, Bovis K J, Clayton A M, Li D, Inverarity G W, Lorenc A C, Payne T J. 2007. The Met Office global four-dimensional variational data assimilation scheme. *Quart J R Meteorol Soc*, 133: 347–362
- Ruti P M, Tarasova O, Keller J H, Carmichael G, Hov Ø, Jones S C, Terblanche D, Anderson-Lefale C, Barros A P, Bauer P, Bouchet V, Brasseur G, Brunet G, DeCola P, Dike V, Kane M D, Gan C, Gurney K R, Hamburg S, Hazeleger W, Jean M, Johnston D, Lewis A, Li P, Liang X, Lucarini V, Lynch A, Manaenkova E, Jae-Cheol N, Ohtake S, Pinardi N, Polcher J, Ritchie E, Sakya A E, Saulo C, Singhee A, Sopaheluwakan A, Steiner A, Thorpe A, Yamaji M. 2020. Advancing research for seamless Earth system prediction. *Bull Am Meteorol Soc*, 101: E23–E35
- Shapiro M, Shukla J, Brunet G, Nobre C, Bèland M, Dole R, Trenberth K, Anthes R, Asrar G, Barrie L, Bougeault P, Brasseur G, Burridge D, Busalacchi A, Caughey J, Chen D, Church J, Enomoto T, Hoskins B, Hov Ø, Laing A, Le Treut H, Marotzke J, McBean G, Meehl G, Miller M, Mills B, Mitchell J, Moncrieff M, Nakazawa T, Olafsson H, Palmer T, Parsons D, Rogers D, Simmons A, Troccoli A, Toth Z, Uccellini L, Velden C, Wallace J M. 2010. An earth-system prediction initiative for the twenty-first century. *Bull Amer Meteor Soc*, 91: 1377–1388
- Sun H, Lei L, Liu Z, Ning L, Tan Z. 2022. An analog offline EnKF for paleoclimate data assimilation. *J Adv Model Earth Syst*, 14: e2021MS002674
- Sun J, Liu Z, Lu F, Zhang W, Zhang S. 2020. Strongly coupled data assimilation using leading averaged coupled covariance (LACC). Part III: Assimilation of real world reanalysis. *Mon Weather Rev*, 148: 2351–2364

- Tan Z M, Zhang F, Rotunno R, Snyder C. 2004. Mesoscale predictability of moist baroclinic waves: Experiments with parameterized convection. *Mon Weather Rev*, 61: 1794–1804
- Tardif R, Hakim G J, Snyder C. 2014. Coupled atmosphere-ocean data assimilation experiments with a low-order climate model. *Clim Dyn*, 43: 1631–1643
- Tardif R, Hakim G J, Snyder C. 2015. Coupled atmosphere-ocean data assimilation experiments with a low-order model and CMIP5 model data. *Clim Dyn*, 45: 1415–1427
- Tong M, Sippel J A, Tallapragada V, Liu E, Kieu C, Kwon I H, Wang W, Liu Q, Ling Y, Zhang B. 2018. Impact of assimilating aircraft reconnaissance observations on tropical cyclone initialization and prediction using operational HWRF and GSI ensemble-variational hybrid data assimilation. *Mon Weather Rev*, 146: 4155–4177
- Whitaker J S, Hamill T M, Wei X, Song Y, Toth Z. 2008. Ensemble data assimilation with the NCEP global forecast system. *Mon Weather Rev*, 136: 463–482
- Xie Y, Koch S, McGinley J, Albers S, Bieringer P E, Wolfson M, Chan M. 2011. A space-time multiscale analysis system: A sequential variational analysis approach. *Mon Weather Rev*, 139: 1224–1240
- Yang X, Rosati A, Zhang S, Delworth T L, Gudgel R G, Zhang R, Vecchi G, Anderson W, Chang Y S, DelSole T, Dixon K, Msadek R, Stern W F, Wittenberg A, Zeng F. 2013. A predictable AMO-like pattern in the GFDL fully coupled ensemble initialization and decadal forecasting system. *J Clim*, 26: 650–661
- Zeng Y, Janjić T, de Lozar A, Welzbacher C A, Blahak U, Seifert A. 2021. Assimilating radar radial wind and reflectivity data in an idealized setup of the COSMO-KENDA system. *Atmos Res*, 249: 105282
- Zhang F. 2005. Dynamics and structure of mesoscale error covariance of a winter cyclone estimated through short-range ensemble forecasts. *Mon Weather Rev*, 133: 2876–2893
- Zhang F, Bei N, Rotunno R, Snyder C, Epifanio C C. 2007. Mesoscale predictability of moist baroclinic waves: Convection-permitting experiments and multistage error growth dynamics. *J Atmos Sci*, 64: 3579–3594
- Zhang F, Weng Y, Sippel J A, Meng Z, Bishop C H. 2009. Cloud-resolving hurricane initialization and prediction through assimilation of doppler radar observations with an ensemble Kalman filter. *Mon Wea Rev*, 137: 2105–2125
- Zhang F, Sun Y Q, Magnusson L, Buizza R, Lin S J, Chen J H, Emanuel K. 2019. What is the predictability limit of midlatitude weather? *J Atmos Sci*, 76: 1077–1091
- Zhang S. 2011. A study of impacts of coupled model initial shocks and state-parameter optimization on climate predictions using a simple pycnocline prediction model. *J Clim*, 24: 6210–6226

(责任编辑: 段晚锁)