

1 基于并行计算的正交条件非线性最优扰动在正压准地转模式集合预报中的应用¹

3 汪叶¹ 段晚锁² 麻俊杰³

4 1 河南大学数学与统计学院, 河南开封 475004

5 2 中国科学院大气物理研究所地球系统数值模拟与应用重点实验室, 北京 100029

6 3 中北大学数学学院, 山西太原 030051

8 摘 要 正交条件非线性最优扰动 (Orthogonal conditional nonlinear optimal
9 perturbations, O-CNOPs) 是一种重要的集合预报初始扰动方法, 然而该扰动传
10 统的顺序优化算法 (S-CNOPs) 计算代价高昂。随后, 高效并行计算 O-CNOPs
11 的算法 (P-CNOPs) 被提出, 但该算法的初始实现基于简单的 LORENZ-96 理论
12 模型, 而本研究将采用更复杂的二维正压准地转模式, 研究 P-CNOPs 的可靠性
13 和高效性, 并从动力学和代数学两方面证明 P-CNOPs 在理论上的合理性, 数值
14 试验表明, P-CNOPs 提供了与 S-CNOPs 集合预报相当的预报技巧, 但前者计算
15 效率显著高于后者, 大大节省了计算资源。因此, P-CNOPs 是计算复杂模式
16 O-CNOPs 的潜在高效算法, 可望未来广泛应用于实际的数值天气预报模式。

17 关键词 集合预报; 初始扰动; 正交条件非线性最优扰动

18 文章编号 2024133B

19 doi: 10.3878/j.issn.1006-9895.2502.24133

20 The application of orthogonal conditional nonlinear optimal perturbations based
21 on parallel computing for ensemble forecasting in barotropic quasi-geostrophic
22 model

23 Wang Ye¹ Duan Wansuo² Ma Junjie³

24 1 School of Mathematics and Statistics, Henan University, Henan, Kaifeng 475004

25 2 State Key Laboratory of Numerical Modeling for Atmospheric Sciences and
26 Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of

收稿日期 2024-11-30; 网络预出版日期

作者简介 汪叶, 女, 1982 年 03 月, 博士, 讲师, 主要从事可预报性研究, E-mail: wangye@henu.edu.cn

通讯作者 段晚锁, E-mail: duanws@lasg.iap.ac.cn

资助项目 国家自然科学基金项目 42105057, 42375152

Funded by National Natural Science Foundation of China (Grants 42105057, 42375152)

Abstract Orthogonal conditional nonlinear optimal perturbations (O-CNOPS) is an important method for generating initial perturbations in ensemble forecasting. However, the traditional sequential optimization algorithm (S-CNOPS) for calculating O-CNOPS is computationally expensive. Subsequently, an efficient parallel algorithm for computing O-CNOPS (P-CNOPS) was proposed, but the initial implementation of this algorithm was based on the simple LORENZ-96 theoretical model. In this study, a more complex two-dimensional barotropic quasi-geostrophic model will be adopted to investigate the reliability and efficiency of P-CNOPS. The theoretical rationality of P-CNOPS will be demonstrated from both dynamical and algebraic perspectives. Numerical experiments show that P-CNOPS provide equivalent forecast skill to S-CNOPS ensemble forecasting, but the former has significantly higher computational efficiency than the latter, resulting in significant savings in computational resources. Therefore, P-CNOPS is a potentially efficient algorithm for computing O-CNOPS in complex models, and it is expected to be widely used in practical numerical weather prediction models in the future.

Keywords ensemble forecasting; initial perturbation; orthogonal conditional nonlinear optimal perturbations

1 引言

数值天气预报 (numerical weather prediction, 简称 NWP) 的本质是求解复杂偏微分方程组的初边值问题。然而, 观测误差和资料同化方法的局限性使得初始分析场存在初始误差。数值模式不可避免地存在模式误差。天气系统是一个混沌系统, 其对初值极其敏感, 具有高度的非线性不稳定性。基于以上三个因素, 对天气和气候事件的单一确定性预报存在较大的不确定性 (Lorenz, 1963; Palmer et al., 2014; Du et al., 2018)。集合预报可以定量估计预报结果的不确定性, 给出天气和气候事件未来状态的概率分布。天气和气候预报从单一确定性预报向概率预报转变更符合气象科学的实际, 而集合预报正是这一转变的关键基础 (杜钧和

陈静, 2010)。

经典的集合预报是在初始分析场上叠加一组不相关的初始扰动, 产生一组初值, 积分初值得到一组集合预报成员, 从而得到集合平均预报和集合离散度。其中, 集合平均预报能够改善单一确定性预报的预报技巧; 集合离散度用于定量估计预报结果的不确定性; 集合成员则提供了概率预报, 满足不同用户对天气和气候预报的多样化需求, 为人们应对天气和气候变化所作决策提供有力的支撑 (Yamaguchi et al., 2012; Palmer, 2019)。

最早的集合预报初始扰动生成方法是蒙特卡罗预报 (Monte Carlo Forecasting, MCF, Leith, 1974), 其在初始分析场上叠加多个随机扰动产生集合成员, 以描述未来时刻大气状态的概率分布。之后, Toth and Kalnay(1993)指出增长型的初始扰动更有利于集合预报技巧的提高, 提出了繁殖向量法(Breed vectors, BVs), 采用该方法构造的扰动是快速增长的初始扰动。同一时期, 欧洲中期天气预报中心 (European Centre for Medium Range Weather Forecasts, ECMWF) 采用奇异向量法 (Singular vectors, SVs; Lorenz, 1965) 进行业务集合预报并取得了巨大成功(Buizza et al., 1993)。奇异向量是从切线性模式推导出的一组相互独立的最快增长扰动。但是有研究指出大气和海洋系统是非线性动力系统, 而 SVs 是线性模式中的最快增长扰动, 因此存在线性局限性(Anderson, 1997; Hamill et al., 2000; Duan and Huo, 2016)。

基于 SVs 的线性局限性, Mu et al. (2003) 将第一 SV 推广到非线性领域, 提出了条件非线性最优扰动 (Conditional nonlinear optimal perturbation, CNOP) 方法。CNOP 代表了非线性模式中增长最快的初始扰动。CNOP 方法已经被广泛应用于大气、海洋、高影响天气和气候事件的可预报性研究中 (Duan et al., 2004; Mu et al., 2009; Duan et al., 2018; Jiang et al., 2024; Chen et al., 2021; Feng et al., 2022; Qin et al., 2023)。

随后, Jiang and Mu (2009) 用 CNOP 替换第一 SV 而保持其余 SVs 不变, 生成集合预报的初始扰动场, 提高了 SVs 方法对应的集合预报技巧 (Jiang et al., 2009)。进一步地, 考虑到相互独立的初始扰动在集合预报中的有用性, Duan and Huo (2016) 将所有的 SVs 推广到非线性领域, 提出了正交 CNOPs (Orthogonal conditional nonlinear optimal perturbations, O-CNOPs) 方法。与 SVs 和 BVs 相比,

O-CNOPs 方法在提高集合预报技巧方面具有更大的优势(Huo and Duan, 2019; Huo et al., 2019; Zhang et al., 2023)。上述研究展示了 O-CNOPs 方法在提高集合预报技巧方面的有效性和优越性。

由 Duan and Huo (2016) 可知求解 O-CNOPs 扰动的传统方法为顺序算法 (S-CNOPs), 其求解过程如下: 首先在整个扰动相空间中, 利用优化算法求解满足约束条件的最优扰动, 得到第一 CNOP (即全局 CNOP, Mu et al., 2003); 其次在与第一 CNOP 正交的子空间中, 求解满足约束条件的最优扰动, 得到第二 CNOP; 然后在与第一 CNOP 和第二 CNOP 正交的子空间中, 求解满足约束条件的最优扰动, 得到第三 CNOP; 这样依次顺序求解, 最终得到一组相互正交的最优扰动。可以看出这种顺序求解算法的计算效率有限, 计算代价高昂。于是, Ma et al. (2024)提出了一种基于并行思想求解 O-CNOPs 扰动的方法(P-CNOPs), 并将该方法用于 LORENZ-96 模式的集合预报中, 结果表明 P-CNOPs 大大减少了 O-CNOPs 扰动的计算时间, 并与 S-CNOPs 有相当的集合预报水平, 提高了 O-CNOPs 方法在集合预报中的适用性。

然而 LORENZ-96 模式是一个简单的理论动力学模型, 本文将采用更复杂的二维正压准地转模式, 研究 P-CNOPs 在理论上的合理性及其在集合预报中的有效性和时效性, 期望为实际数值天气预报提供一种高效且具有更高预报技巧的集合预报方法。另外, 研究指出初始扰动场的结构与初始扰动振幅的大小及优化时间的长短有关(Toth and Kalnay, 1993; Buizza et al., 2005; Descamps and Talagrand, 2007), 集合样本数也是影响集合预报技巧的关键因素(Daron and Stainforth, 2013; Scaife et al., 2014; 汪叶和段晚锁, 2019)。因此, 本文针对二维正压准地转模式, 首先采用传统的 S-CNOPs 产生初始扰动, 探讨 S-CNOPs 取得最高集合预报技巧时的参数配置(初始扰动振幅、优化时间和集合样本数的取值)并进行原因分析, 然后采用 P-CNOPs 产生新的初始扰动, 并进行集合预报, 考察 P-CNOPs 的合理性、有效性和时效性。

本文结构安排如下: 第二小节阐述 S-CNOPs 和 P-CNOPs 的基本原理, 第三小节介绍二维正压准地转模式, 第四小节进行 S-CNOPs 集合预报试验并做原因分析, 第五小节进行 P-CNOPs 集合预报试验, 并与 S-CNOPs 集合预报结果进行对比分析, 第六小节是总结和讨论。

2 方法介绍

正交条件非线性最优扰动 (O-CNOPs, Duan and Huo, 2016) 是一种重要的集合预报初始扰动方法, 本节将分别介绍 O-CNOPs 扰动的传统顺序算法 S-CNOPs 及其并行算法 P-CNOPs。

2.1 S-CNOPs

假定状态变量 $\mathbf{U}$ 的演变可以用下述非线性偏微分方程组表达:

$$\begin{cases} \frac{\partial \mathbf{U}}{\partial t} = \Psi(\mathbf{U}(\mathbf{x}, t)), \\ \mathbf{U}|_{t=0} = \mathbf{U}_0, \end{cases} \quad (1)$$

其中 $\mathbf{U}(\mathbf{x}, t) = (U_1(\mathbf{x}, t), U_2(\mathbf{x}, t), \dots, U_n(\mathbf{x}, t))$, $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \Omega \subset \mathbf{R}^n$, $t \in [0, T]$, $T < +\infty$, Ψ 是非线性微分算子, $\mathbf{U}_0$ 是初始状态。假设 M_t 是方程组(1)的非线性传播算子, 则在时刻 T 方程组(1)的数值解可由下式给出:

$$\mathbf{U}(T) = M_T(\mathbf{U}_0). \quad (2)$$

假设 $\mathbf{u}_0$ 是初始扰动, 那么 $\mathbf{u}_0$ 在时刻 T 的非线性发展可以表达为

$$J(\mathbf{u}_0) = \|M_T(\mathbf{U}_0 + \mathbf{u}_0) - M_T(\mathbf{U}_0)\|_f, \quad (3)$$

(3)式也表示由初始扰动 $\mathbf{u}_0$ 引起的预报轨迹对参考态 $M_T(\mathbf{U}_0)$ 的偏离大小。

基于上述表示, Duan and Huo (2016) 定义 O-CNOPs 如下: 如果初始扰动 $\mathbf{u}_{0j}^*$ 满足:

$$J(\mathbf{u}_{0j}^*) = \max_{\mathbf{u}_{0j} \in \Omega_j} \|M_T(\mathbf{U}_0 + \mathbf{u}_{0j}) - M_T(\mathbf{U}_0)\|_f, \quad (4)$$

其中

$$\Omega_j = \begin{cases} \{\mathbf{u}_{0j} \in \mathbf{R}^n \mid \|\mathbf{u}_{0j}\|_a \leq \delta\}, j=1, \\ \{\mathbf{u}_{0j} \in \mathbf{R}^n \mid \|\mathbf{u}_{0j}\|_a \leq \delta, \mathbf{u}_{0j} \perp \Omega_k, k=1, 2, \dots, j-1\}, j>1, \end{cases} \quad (5)$$

那么称初始扰动 $\mathbf{u}_{0j}^*$ 为第 j 个 CNOP。这里 Ω_j 是扰动相空间的子空间, 符号“ $\perp$ ”代表正交, $\mathbf{u}_{0j}$ 是子空间 Ω_j 中的一个初始扰动, $\|\cdot\|_a$ 代表初始时刻范数, $\|\cdot\|_f$ 代表终止时刻范数。 $\|\mathbf{u}_{0j}\|_a \leq \delta$ 是初始扰动需要满足的约束条件, δ 是一个正常数,

代表初始扰动约束半径。由上述定义可知，第一（全局）CNOP 是在整个相空间中具有最大非线性发展的初始扰动(Mu et al., 2003)，第二 CNOP 是在与第一 CNOP 正交的子空间中有最大非线性发展的初始扰动，类似地，第 j 个 CNOP 是在与前 $j-1$ 个 CNOPs 正交的子空间中具有最大非线性发展的初始扰动。这些 CNOPs 是相互正交的，且这些扰动的非线性发展满足如下规律： $J(\text{CNOP}_1) > J(\text{CNOP}_2) > \dots > J(\text{CNOP}_n)$ 。

本文使用谱投影梯度优化算法 (Spectral projected gradient, SPG2; Birgin et al., 2000) 计算 O-CNOPs 扰动。具体地，首先将整个初始扰动相空间作为优化问题(4)的约束条件，用 SPG2 优化算法计算第一 CNOP；其次，将与第一 CNOP 正交的子空间内的所有初始扰动作为优化问题(4)的约束条件，采用 SPG2 优化算法计算第二 CNOP；再次，将与第一 CNOP 和第二 CNOP 正交的子空间内的所有初始扰动作为优化问题(4)的约束条件，计算第三 CNOP；以此类推，逐个求解，最终得到一组相互正交的、在各自子空间中具有最大非线性发展的初始扰动，即得到 O-CNOPs 扰动。

2.2 P-CNOPs

由 2.1 小节可知，O-CNOPs 扰动必须按顺序逐个求解，所需的计算时间成本较大，计算量也很大。为了提高计算效率，Ma et al. (2024) 基于并行思想提出了一种求解 O-CNOPs 扰动的高效算法 (P-CNOPs)。该方法能够同时输出一组正交的条件非线性最优扰动，简化了求解过程，降低了求解最优扰动的计算时间成本。P-CNOPs 方法的基本思想和主要步骤如下。

(1)初始化：设置初始扰动个数 M ，扰动约束半径 δ ，最大迭代步数 K ，在初始扰动相空间中给定一组随机的正交初始扰动 $\mathbf{u}_{0j}^{(0)} (j=1, 2, \dots, M)$ 作为初猜值，其中 $\mathbf{u}_{0j}^{(0)}$ 需要满足 $\mathbf{u}_{0i}^{(0)} \perp \mathbf{u}_{0j}^{(0)} (i \neq j, i, j=1, 2, \dots, M)$ 且 $\|\mathbf{u}_{0j}^{(0)}\|_a \leq \delta (j=1, 2, \dots, M)$ 。

(2)更新初始扰动组：每个初始扰动沿着其相应的目标函数的梯度方向同时进行优化，即

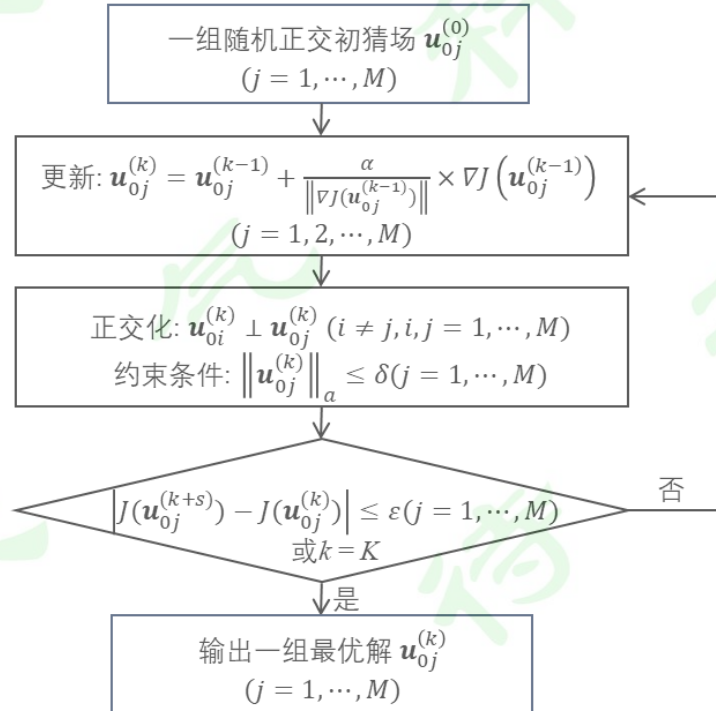
$$\mathbf{u}_{0j}^{(k)} = \mathbf{u}_{0j}^{(k-1)} + \frac{\alpha}{\|\nabla J(\mathbf{u}_{0j}^{(k-1)})\|} \times \nabla J(\mathbf{u}_{0j}^{(k-1)}), (j=1, 2, \dots, M), \quad (6)$$

其中 k 代表迭代步数, α 是超参数, 用于调整迭代步长, ∇J 指目标函数的梯度, $\|\nabla J\|$ 是梯度的 L_2 范数。

(3) 正交化及约束: 对步骤 (2) 中梯度更新后的初始扰动组 $\mathbf{u}_{0j}^{(k)} (j=1, 2, \dots, M)$ 做正交化处理及扰动半径约束, 即要求 $\mathbf{u}_{0i}^{(k)} \perp \mathbf{u}_{0j}^{(k)} (i \neq j, i, j=1, 2, \dots, M)$ 且 $\|\mathbf{u}_{0j}^{(k)}\|_a \leq \delta (j=1, 2, \dots, M)$ 。

(4) 终止条件: 判断初始扰动 $\mathbf{u}_{0j}^{(k)} (j=1, 2, \dots, M)$ 是否满足终止条件: $|J(\mathbf{u}_{0j}^{(k+s)}) - J(\mathbf{u}_{0j}^{(k)})| \leq \varepsilon$ 或者 $k = K$ 。其中, s 和 ε 是超参数, 分别代表连续迭代步数和一个很小的阈值。如果不满足终止条件, 重复 (2) 和 (3) 两个步骤。如果满足终止条件, 则一次性输出扰动组 $\mathbf{u}_{0j}^{(k)} (j=1, 2, \dots, M)$ 作为最优解, 结束程序。

从上述计算过程可以看出, P-CNOPS 是基于并行的思想, 在同一时刻对一组初始扰动并列执行梯度优化更新, 然后对于更新的扰动组进行正交化和约束处理, 多次循环这个过程, 在优化结束时同时输出一组正交的最优扰动, 从而减少了计算量, 提高了计算效率。P-CNOPS 的流程图如图 1 所示。



176 图 1 P-CNOPS 流程图。

177 Fig. 1 Flow chart of the P-CNOPS.

178 3. 模式介绍

179 本文采用的二维无量纲正压准地转模式如下：

$$180 \quad \begin{cases} \frac{\partial P}{\partial t} + \partial(\Phi, P) = 0, \\ P = \nabla^2 \Phi - F\Phi + f_0 + f_0 h_s, \quad \text{in } \Omega \times [0, t_0], \\ \Phi|_{t=0} = \Phi_0, \end{cases} \quad (7)$$

181 其中 P 是位势涡度， Φ 是流函数， h_s 是地形。行星尺度 Froude 数 $F = 0.102$ ，科
182 氏参数 $f_0 = 10.0$ 。 $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ 是拉普拉斯算子， x 和 y 是纬向和经向坐
183 标， t 是时间，水平雅可比算子 $\partial(\Phi, P) = \Phi_x P_y - \Phi_y P_x$ 。区域 $\Omega = [0, X] \times [0, Y]$ 具有
184 双周期边界条件。对于任意给定的时间 $t_0 > 0$ 和初始条件 $\Phi|_{t=0} = \Phi_0$ ，求解初值问
185 题 (7) 可以得到流函数 $\Phi(x, y, t_0)$ 。设 M 代表非线性传播算子，则
186 $\Phi(x, y, t_0) = M_{t_0}(\Phi_0)$ 代表初值问题(7)在时刻 t_0 的解（参见 Mu and Zeng, 1991）。
187 正压准地转模式是一个理论模式，已被广泛应用于可预报性和集合预报相关问题
188 的研究中（Cheung and Chan, 1999; Mu and Jiang, 2008; Durran and Gingrich, 2014;
189 Feng et al., 2016）。

190 本文数值试验中，对方程(7)中的拉普拉斯项采用五点离散差分格式，对雅
191 可比项采用 Arakawa 有限差分格式，时间离散采用 Adams-Bashforth 格式。区域
192 $\Omega = [0, 6.4] \times [0, 3.2]$ 对应于 $[0, 6400\text{km}] \times [0, 3200\text{km}]$ ，空间格距 $d = 0.2$ 相当于
193 200km ，对应格点数为 32×16 ，时间步长为 $dt = 0.006$ 相当于 10 分钟。参考 Mu and
194 Zhang（2006）中关于正压模式参考态的确定方式，本文选取
195 $\Phi_0 = 0.5 \sin(2\pi x/6.4) + 2 \sin(2\pi y/3.2) + 0.5$ 为初始条件，
196 $h_s = 0.112[\sin(4\pi y/3.2) + 1.0]$ 为地形，积分方程(7)得到非线性不稳定参考态
197 $\Phi(x, y, t_0)$ ，此时模式是一个具有混沌特性的非线性动力系统，可用于非线性误

差增长研究。本文将选择具有上述参数设置的正压模式进行集合预报试验。

4. S-CNOPS 集合预报试验

4.1. 试验设计

本文采用“完美模式”假定，即假定预报误差来源于初始场的不确定性。首先将模式自由积分 50 天，然后向前积分 1000 天生成状态变量 $\Phi(i, j)(i=1, 2, \dots, 32, j=1, 2, \dots, 16)$ 的时间序列，从该时间序列中每隔 20 天取一次状态变量的值作为初始状态，共得到 $\Phi(i, j)$ 的 50 个初始状态。将这些初始状态向前积分 10 天，得到 50 个用于检验预报的“真场”（或“个例”）。在每个格点的真值上，每隔 6 小时叠加一个具有正态分布 $N(0, 0.07^2)$ 的随机数构造观测场，这里观测误差的标准差约为气候态标准差的 10%。采用 Mu and Jiang (2008) 文中的四维变分资料同化方法（Four Dimensional Variational Data Assimilation, 4DVAR）同化观测场得到初始分析场。以初始分析场为初值，积分模式 10 天得到控制预报。对于 50 个个例，得到 50 个控制预报。初始分析场与真场的差为初始分析误差，初始分析误差的大小用 L_2 范数度量，记为 δ_a 。

观测误差的存在导致初始分析场存在误差，进而导致控制预报存在预报不确定性，因此需要采用集合预报的方法估计这种由初值不确定性导致的预报不确定性。本节使用 S-CNOPS 产生集合预报的初始扰动场。具体地，对于每个个例，以控制预报为基态，用 SPG2 优化算法求解优化问题(4)得到 S-CNOPS。在试验中分别采用 L_2 范数和能量范数度量初始扰动和扰动发展的范数，具体定义公式如下：

$$\|\varphi\|_{L_2}^2 = \int_{\Omega} \varphi^2 dx dy, \quad (8)$$

$$\|\varphi\|_E^2 = \int_{\Omega} (|\nabla \varphi|^2 + F|\varphi|^2) dx dy, \quad (9)$$

其中， $\|\cdot\|_E$ 代表能量范数， $\|\cdot\|_{L_2}$ 代表 L_2 范数， φ 是流函数。

前人的研究指出 O-CNOPS 扰动的空间结构和优化时间 T 、初始扰动振幅 δ 密切相关（Duan and Huo, 2016），因此在本文中，优化时间分别取为 $T=2$ 天、3

天, 初始扰动振幅分别取为 $\delta = 0.2\delta_a$ 、 $0.4\delta_a$ 、 $0.6\delta_a$ 、 $0.8\delta_a$ 和 δ_a , 共构成 10 种 T - δ 组合。我们的前期工作 (汪叶和段晚锁, 2019) 指出, 排序在前的 O-CNOPS 扰动是快速增长型扰动, 排序在后的 O-CNOPS 扰动是慢速增长或不增长型扰动。其中, 快速增长型初始扰动在预报的初始阶段即被放大并影响到整个物理场, 对集合预报更重要, 对提高集合预报技巧起着积极作用 (Toth and Kalnay, 1993; 汪叶和段晚锁, 2019)。而正压模式的空间维数为 512 维, 因此本文对于每种 T - δ 组合, 对于每个个例, 求解了前 82 个 S-CNOPS。然后, 从这 82 个 S-CNOPS 中取前 n 个 S-CNOPS, 以正负扰动对的形式, 将其叠加在控制预报的初始场上构造 $2n$ 个对称的初始场, 从而得到 $2n$ 个集合成员, 再加上未受到扰动的控制预报, 一共构成 N ($N=2n+1$) 个集合成员。其中 $N=15$ 、31、45、61、75、91、105、121、135、151、165。为了减少统计波动性, 对集合预报技巧评分的计算均基于 $50 \times N$ 个样本的统计结果。

集合预报系统的检验分为确定性预报和概率预报两个方面。本文采用均方根误差 (RMSE) 和距平相关系数 (ACC; Murphy and Epstein, 1989) 度量确定性预报的预报技巧。其中, RMSE 越小, ACC 越大, 代表预报技巧越高。一般认为 ACC 大于 0.6 时预报有效。采用 BS 评分 (Brier score; Descamps and Talagrand, 2007) 和 ROCA 评分 (the area under the relative operating characteristic curve, ROCA; Mason and Graham, 1999) 度量概率预报的预报技巧。其中, BS 评分越小, ROCA 评分越大, 则概率预报技巧越高。一般地, 当 ROCA 大于 0.5 时表明概率预报有效。为了度量概率预报, 定义了两个二分类事件, 分别为高频事件 event1 : $\Phi(i, j) > \Phi_c(i, j) (i=1, 2, \dots, 32, j=1, 2, \dots, 16)$ 和低频事件 event2 : $\Phi(i, j) > \Phi_c(i, j) + \sigma(i, j) (i=1, 2, \dots, 32, j=1, 2, \dots, 16)$, 其中 Φ_c 和 σ 分别是每个格点上变量 Φ 的气候态平均值和标准差, 高频事件和低频事件发生的频率分别约为 0.498 和 0.166。另外, 有研究指出, 在同等度量方式下, 一个可靠的集合预报系统, 其集合平均预报误差与集合离散度无论在时间上还是在空间上都应该相当的, 这样, 集合预报系统可以通过集合离散度来估计集合平均预报误差 (Buckingham et al., 2010; Fortin et al., 2014; Zhu et al., 2023)。因此, 本文还考察了集合平均预报误差和集合离散度在时间上和空间上的关系, 用以检验集合预报系

统的可靠性。

4.2 试验结果

首先从 50 个个例中任意选取一个个例，考察真实初始场 Φ_0^T 和对应的初始分析场 Φ_0^a 的空间结构的差异。以 L_2 范数为度量，计算得到真实初始场的大小为 $\|\Phi_0^T\|_{L_2} = 12.1416$ ，初始分析误差（初始分析场与真实初始场之差值）的大小为 0.3268，初始分析误差的 L_2 范数占真实场 L_2 范数的 2.69%，表明初始时刻控制预报对真实场的模拟效果较好。其余个例有类似结果，此处不再一一展示。

下面探讨初始扰动振幅和优化时间及集合样本数对 S-CNOPS 集合预报技巧的影响。首先针对每个个例，针对 10 种 $T-\delta$ 组合中的每种组合，采用 11 种集合样本数 ($N=15, 31, \dots, 165$) 进行 10 天的集合预报试验。然后针对每种 $T-\delta-N$ 组合，综合每个个例及其在每一预报时刻 (6h、12h、18h、...、10d) 的预报结果，计算了集合平均预报的 RMSE 和 ACC。同时，如果将每个个例中对每个变量 $\Phi(i, j)$ 的预报看做一次预报过程，那么共有 $512 \times 50 = 25600$ 次预报过程。针对每种 $T-\delta-N$ 组合，基于 25600 次预报过程和每一个预报时刻，也计算了两个二分类事件的 BS 评分和 ROCA 评分。

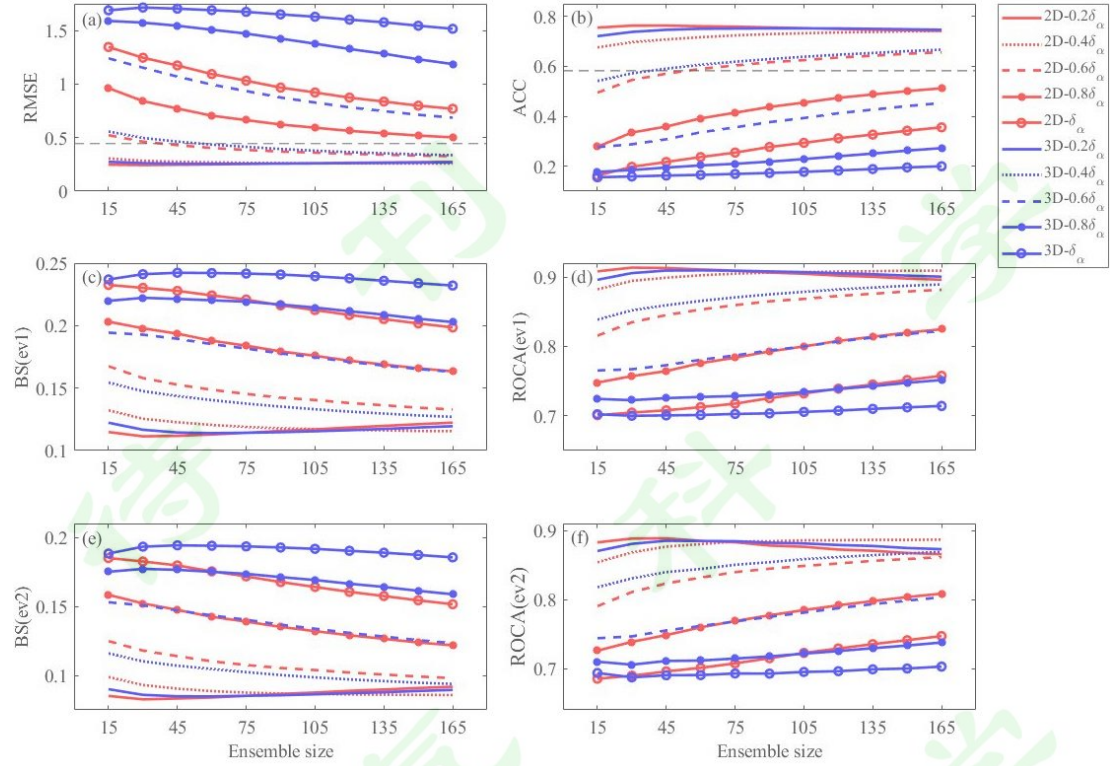


图 2 优化时间为 2 天（红）和 3 天（蓝）时 S-CNOPs 集合预报技巧随集合样本数的演变。
 (a) RMSE; (b) ACC; (c) 事件 ev1 的 BS 评分; (d) 事件 ev1 的 ROCA 评分; (e)
 事件 ev2 的 BS 评分; (f) 事件 ev2 的 ROCA 评分。所有技巧评分是针对 50 个个例和 10
 天内所有预报时刻的平均值。实线、点线、虚线、圆点实线和圆圈实线分别对应于初始扰动
 振幅 $\delta = 0.2\delta_a$ 、 $0.4\delta_a$ 、 $0.6\delta_a$ 、 $0.8\delta_a$ 和 δ_a 。(a) 和 (b) 中黑色虚线分别代表控制预报
 的 RMSE 和 ACC。

Fig. 2 Skills of ensemble forecasts based on the S-CNOPs with optimization time intervals of $T=2$
 days (red) and 3 days (blue) as a function of ensemble size. (a)RMSE, (b)ACC, (c)BS for the
 event ev1, (d)ROCA for the event ev1, (e)BS for the event ev2, (f)ROCA for the event ev2. The
 skills are estimated by averaging over 50 cases with all the lead times within 10 days. The solid
 lines, dotted lines, dashed lines, solid lines with dots, solid lines with circles correspond to the
 initial perturbation magnitude δ of $0.2\delta_a$, $0.4\delta_a$, $0.6\delta_a$, $0.8\delta_a$ and δ_a , respectively.
 The black dashed lines in figures (a) and (b) represent RMSE and ACC of control forecast,
 respectively.

图 2 给出了当优化时间分别为 2 天和 3 天时，不同扰动振幅下，S-CNOPs

集合预报的预报技巧随集合样本数的演变。结果表明，在相同初始扰动振幅下，S-CNOPS 集合预报在优化时间为 2 天时其预报技巧基本上均高于优化时间为 3 天时的预报技巧。而当固定优化时间为 2 天时，较小的初始扰动振幅（ $\delta = 0.2\delta_a$ 、 $0.4\delta_a$ 、 $0.6\delta_a$ ）对应的集合平均预报能够改善控制预报，且此时的概率预报也是相对更高的。这表明初始扰动振幅较小且优化时间较短，有利于 S-CNOPS 集合预报获得更高的预报技巧。事实上，CNOPS 是非线性系统中的最快增长扰动，虽然其初始振幅较小，但是其快速增长的行为使得它们生成的集合成员有着较大且合适的离散度，能够抓住真值，从而获得更高的集合预报技巧。但是，当初始振幅较大时，扰动成员将会过高估计预报误差的非线性增长，离散度过大，导致集合预报不能改善控制预报。

下面我们探讨在何种参数配置下，S-CNOPS 集合预报系统最可靠。图 3 给出了优化时间为 2 天，初始扰动振幅 $\delta = 0.2\delta_a$ 、 $0.4\delta_a$ 、 $0.6\delta_a$ 时，S-CNOPS 集合平均预报的预报误差和集合离散度的关系。结果表明，随着扰动振幅的增大，集合离散度快速增大；当优化时间为 2 天，扰动振幅 $\delta = 0.2\delta_a$ ，集合样本数为 135 时，集合平均预报误差和集合离散度的比值最接近于 1，集合预报系统最可靠。

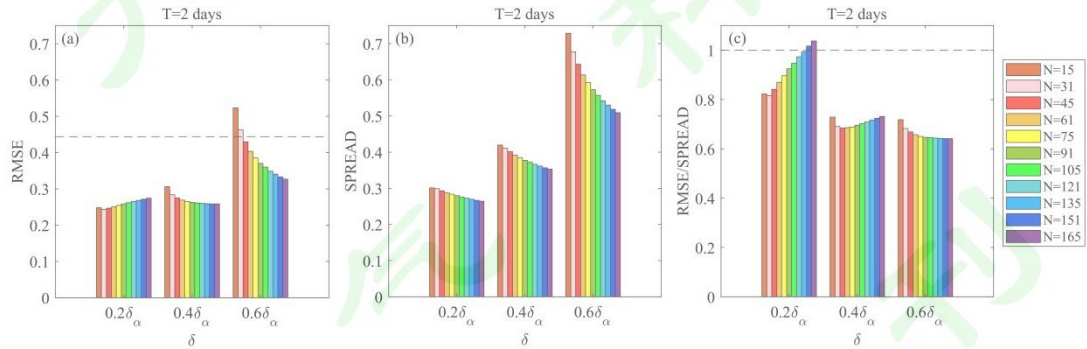


图 3 优化时间为 2 天时，不同初始扰动振幅下，（a）集合平均预报误差，（b）集合离散度，（c）RMSE 与 SPREAD 的比值随集合样本数的演变。其中，RMSE 和 SPREAD 是针对 50 个个例和 10 天内所有预报时刻计算的平均值。每幅图中有三簇柱状图分别对应于初始扰动振幅 $\delta = 0.2\delta_a$ 、 $0.4\delta_a$ 、 $0.6\delta_a$ 。每簇柱状图中从左到右不同颜色的柱对应于集合样本数由小（ $N=15$ ）变到大（ $N=165$ ）。（a）中黑色虚线代表控制预报的预报误差。

Fig. 3 Evolution of (a) RMSE of ensemble mean forecast, (b)SPREAD, (c)the ratio of RMSE to

SPREAD with respect to ensemble sizes N under different initial perturbation magnitudes when the optimization time interval is 2 days. RMSE and SPREAD are estimated by averaging over 50 cases with all the lead times within 10 days. There are three clusters of bars in each figure that correspond to the initial perturbation magnitudes δ of $0.2\delta_a$, $0.4\delta_a$, $0.6\delta_a$, respectively. The bars of different colors from left to right in each cluster of bars correspond to different ensemble sizes from small ($N=15$) to large ($N=165$). The black dashed line in (a) represents RMSE of control forecast.

上述关于 S-CNOPS 集合预报系统可靠性的分析是基于 50 个个例和 10 天预报时间内所有预报时刻的平均值做出的，那么 S-CNOPS 集合平均预报误差和离散度随着预报时间的演变是否也能够保持较好的一致性呢？图 4（a）给出了控制预报误差和集合平均预报误差及集合离散度随预报时间的演变，图 4（b）给出了控制预报和集合平均预报的距平相关系数 ACC 随预报时间的演变。结果表明，随着预报时间的增加，S-CNOPS 集合平均预报对控制预报的改善越来越大，改善程度达 40%。从 ACC 的角度来看，S-CNOPS 集合平均预报将控制预报的有效预报时效从 5 天提高到 10 天以上。考察离散度发现，在预报前期集合预报的离散度较小，这是因为初始扰动振幅较小，只占初始分析误差的 20%；随着预报时间的增加，离散度迅速增大，在预报后期（5-10 天）离散度赶上并略大于集合平均预报误差。总之，在整个预报时间段内集合离散度和集合平均预报误差都具有较好的一致性关系。

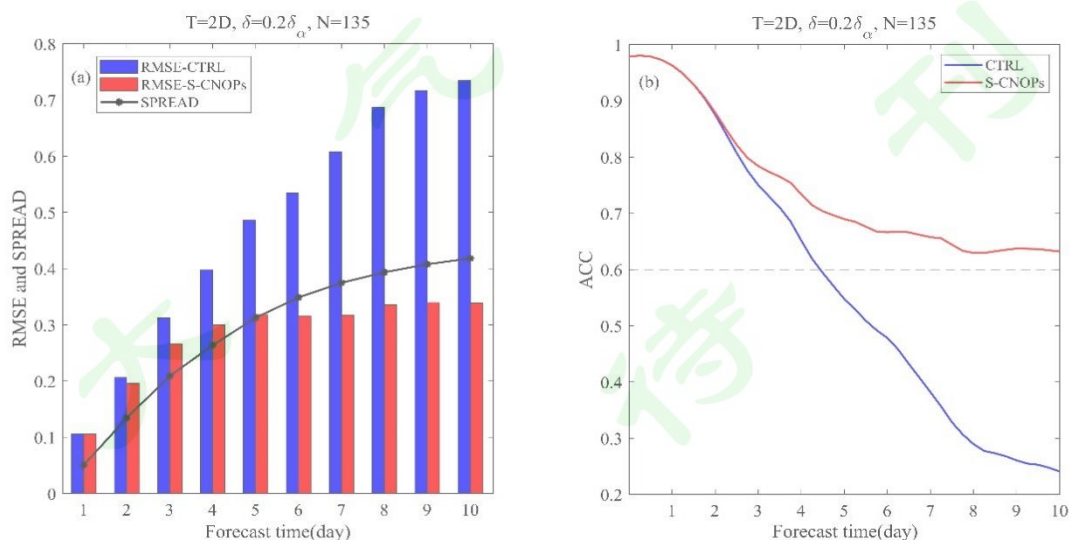


图 4 控制预报（蓝）和 S-CNOPS 集合平均预报（红）的（a）预报误差及集合离散度（黑色圆点实线）和（b）ACC 随预报时间的演变。

Fig. 4 Evolution of (a) the RMSE and SPREAD (black solid line with dots) and (b) ACC for the control forecast (blue) and S-CNOPS ensemble forecast (red) as a function of forecast time.

进一步地，我们考察 S-CNOPS 集合预报系统在空间上的可靠性。图 5 表明在整个研究区域内集合平均预报误差与集合离散度的比值均位于 $[0.9, 1.1]$ 之间，在大部分区域中比值都接近于 1，显示了一种几乎完美的 RMSE-SPREAD 比值关系。综上，S-CNOPS 集合预报能够显著提高控制预报技巧，极大延长有效预报时效，集合离散度无论在时间上还是在空间上均与集合平均预报误差保持较好的一致性关系，表明了 S-CNOPS 集合预报系统的有效性和可靠性。

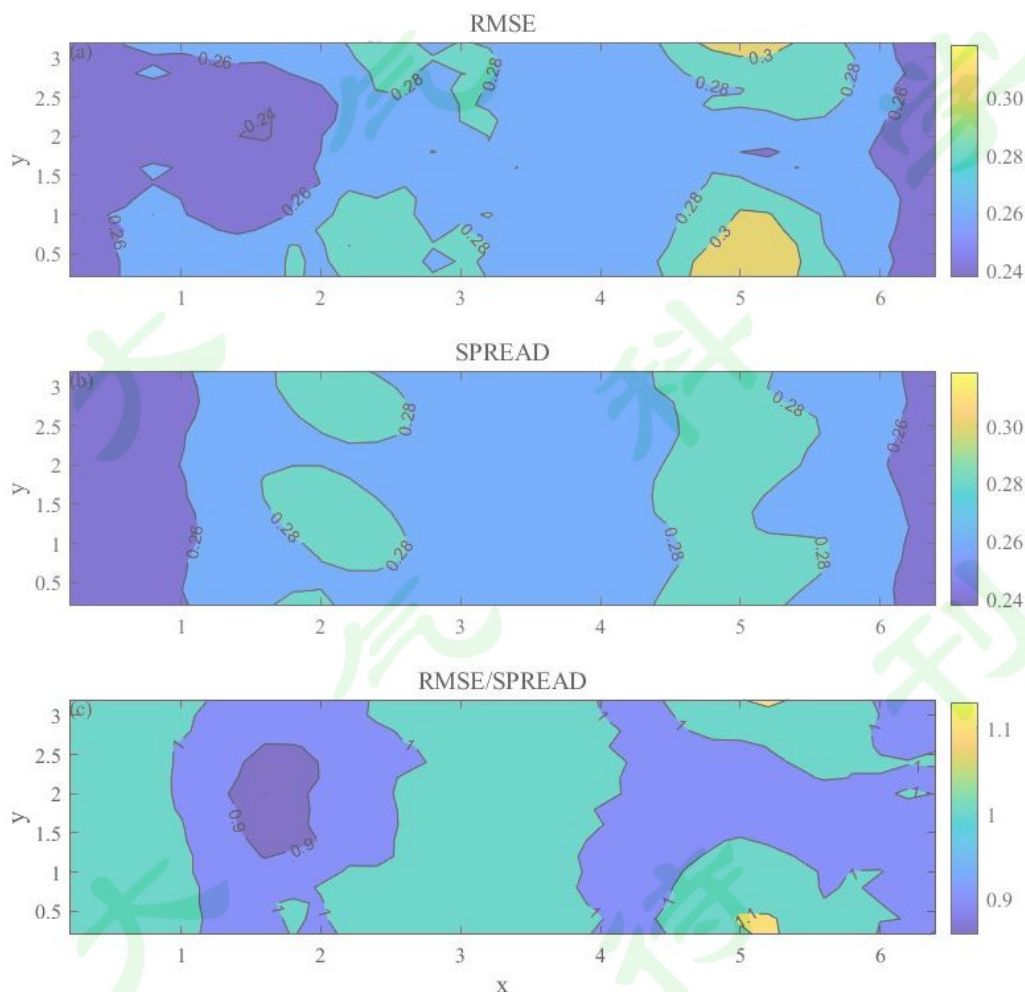


图 5 （a）S-CNOPS 集合平均预报误差 RMSE、（b）离散度及（c）两者比值的空间分布。

Fig. 5 The spatial distributions of (a) the RMSE of S-CNOPS ensemble mean, (b) ensemble spread

and (c) the ratio of RMSE to SPREAD.

4.3 原因分析

本节我们尝试分析 S-CNOPS 集合平均预报显著改善控制预报的原因。有研究指出，在完美模式条件下，集合预报的效果主要依赖于初始扰动能否较好地表征初始分析误差。为了探讨分析误差和集合扰动的关系，假定 Φ^T 为真实状态向量， Φ^M 为集合平均预报， Φ^C 为控制预报，令 $\mathbf{e}_1 = \Phi^C - \Phi^T$ ， $\mathbf{e}_2 = \Phi^M - \Phi^C$ ，其中 $\mathbf{e}_1$ 代表分析误差的非线性发展模态， $\mathbf{e}_2$ 代表集合平均扰动的非线性发展模态，那么有 $\Phi^M - \Phi^T = \mathbf{e}_1 + \mathbf{e}_2 = \mathbf{e}_3$ 。如图 6 所示，只有当分析误差 $\mathbf{e}_1$ 和集合平均扰动 $\mathbf{e}_2$ 呈现相反发展态势时，即 $\mathbf{e}_1$ 和 $\mathbf{e}_2$ 的相似系数为负值时，集合平均扰动才能抵消分析误差的非线性发展，从而提高控制预报的预报技巧。这里相似系数 S 定义为

$$S = \langle \mathbf{e}_1, \mathbf{e}_2 \rangle / (\|\mathbf{e}_1\| \cdot \|\mathbf{e}_2\|). \quad (10)$$

可以看出 $\mathbf{e}_1$ 和 $\mathbf{e}_2$ 的负相关越高，集合平均预报技巧越高。

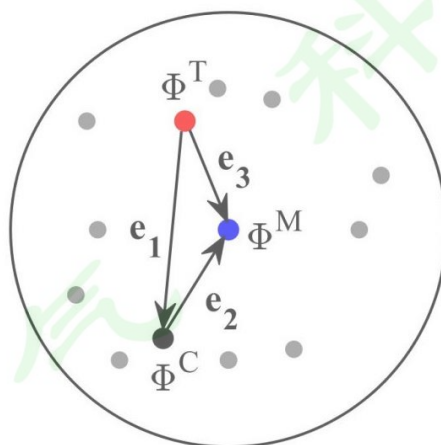


图 6 集合预报示意图，红点、蓝点、黑点和灰点分别代表真值、集合平均预报、控制预报和集合成员。

Fig. 6 Schematic diagram of ensemble forecast, where the red dot, blue dot, black dot and gray dots represent the truth run, ensemble mean forecast, control forecast and ensemble members, respectively.

下面对分析误差和集合平均扰动的 10 天非线性发展序列进行经验正交函数 (Empirical Orthogonal Function, EOF) 分析。图 7 分别给出了分析误差和集合平均扰动的非线性发展的前五个 EOF 模态,其解释方差分别占各自总方差的 81.3% 和 79.8% (表 1)。然后计算了分析误差的每个 EOF 模态与相对应的集合平均扰动 EOF 模态的相似系数 (表 2), 结果表明分析误差和集合平均扰动的对应模态之间均呈现负相关关系, 和上述理论分析一致。

进一步地, 我们给出分析误差和 S-CNOPs 集合平均扰动在第 0、2、4、6、8、10 天的非线性发展 (图 8), 给出分析误差和 S-CNOPs 集合平均扰动的非线性发展的相似系数随预报时间的演变图 (图 9)。可以看出分析误差的非线性发展和集合平均扰动的非线性发展随预报时间的增长呈现越来越强的负相关。两者负相关程度越高, 集合平均扰动的发展对分析误差的发展的抵消力度越大, 因此集合平均预报越准确, 对控制预报的改善程度越大。综上, 正是因为 S-CNOPs 扰动充分抓住了分析误差的非线性演变过程, 使得 S-CNOPs 集合平均预报的轨迹更加靠近真值, 使得 S-CNOPs 集合平均预报能够显著地改善控制预报。

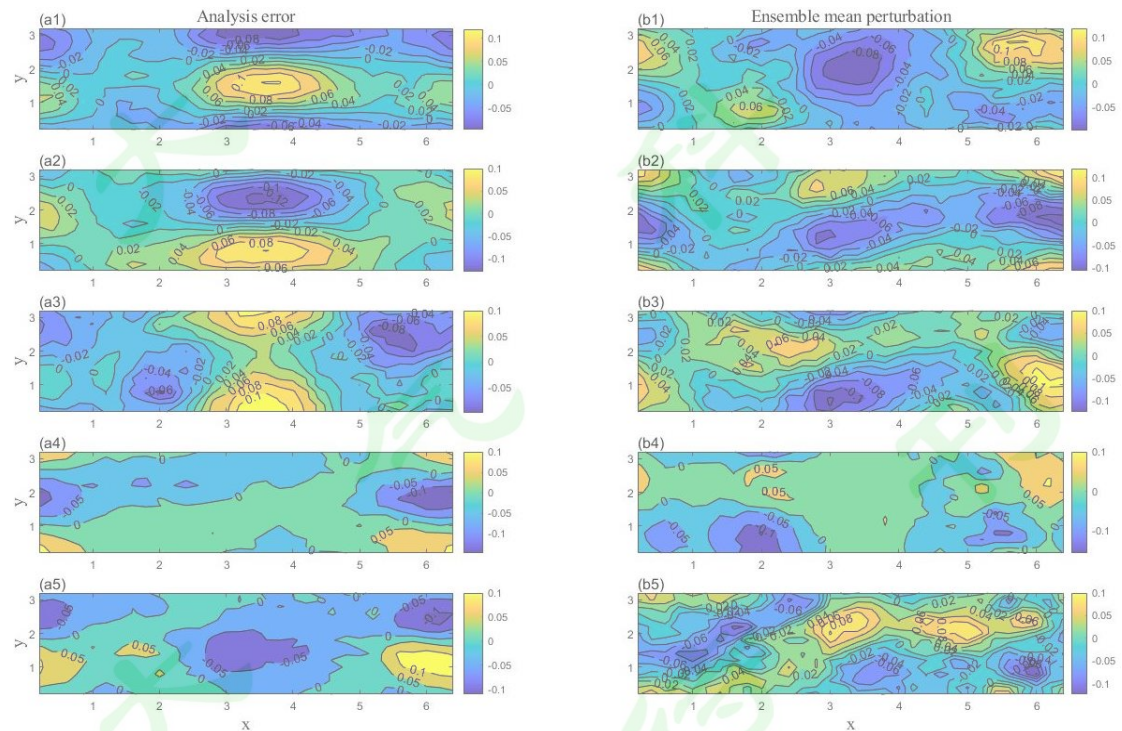


图 7 (a1) 至 (a5) 是分析误差的前五个 EOF 模态, (b1) 至 (b5) 是集合平均扰动的前五个 EOF 模态。

Fig. 7 (a1) to (a5) are the first five EOF modes of analysis error, and (b1) to (b5) are the first five

EOF modes of ensemble mean perturbation.

表 1 分析误差和集合平均扰动的前五个 EOF 模态所占方差和总方差

Table 1 The variances and total variance of the first five EOF modes of analysis error and ensemble mean perturbation

	EOF1	EOF2	EOF3	EOF4	EOF5	Total variance
Analysis error	0.281	0.252	0.109	0.092	0.079	0.813
Ensemble mean	0.284	0.217	0.155	0.093	0.049	0.798

表 2. 分析误差的前五个 EOF 模态与集合平均扰动的对应模态的相似系数

Table 2 The similar coefficients between the EOF modes of analysis error and the corresponding ones of ensemble mean perturbation

Mode	EOF1	EOF2	EOF3	EOF4	EOF5
Similar coefficient	-0.367	-0.233	-0.394	-0.458	-0.411

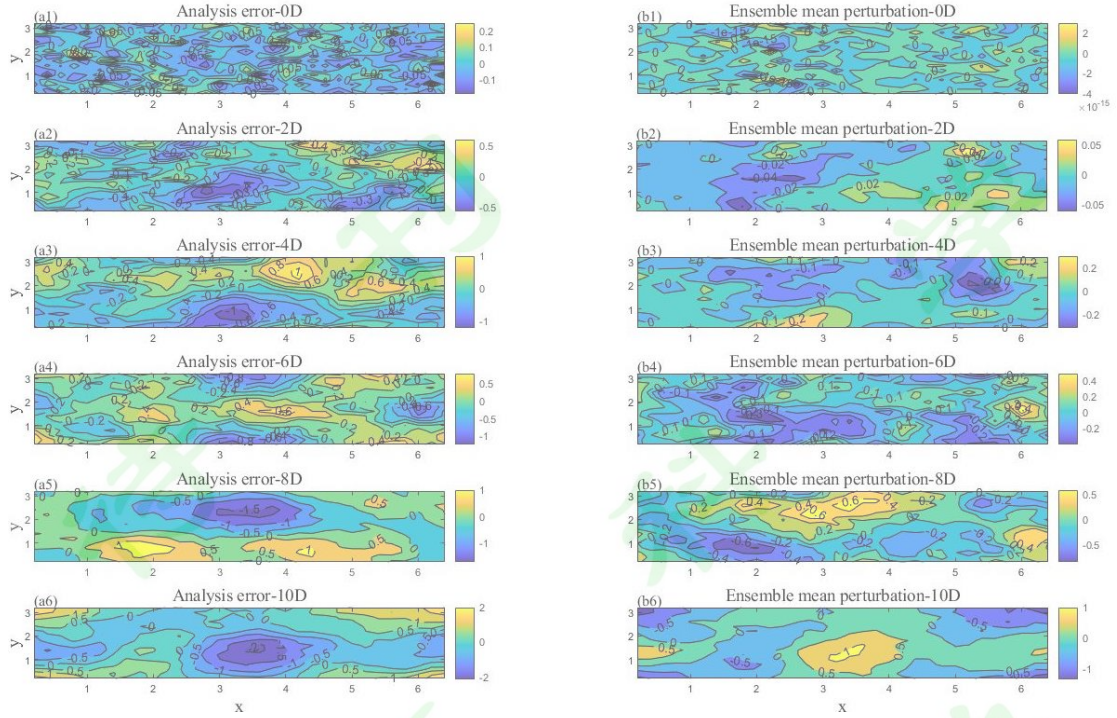


图 8 (a1) 至 (a6) 是分析误差在第 0、2、4、6、8、10 天的非线性发展，(b1) 至 (b6) 是集合平均扰动的非线性发展。

Fig. 8 (a1) to (a6) are the nonlinear evolution of analysis error on day 0, 2, 4, 6, 8 and 10, and (b1) to (b6) are the nonlinear evolution of ensemble mean perturbation at the corresponding day.

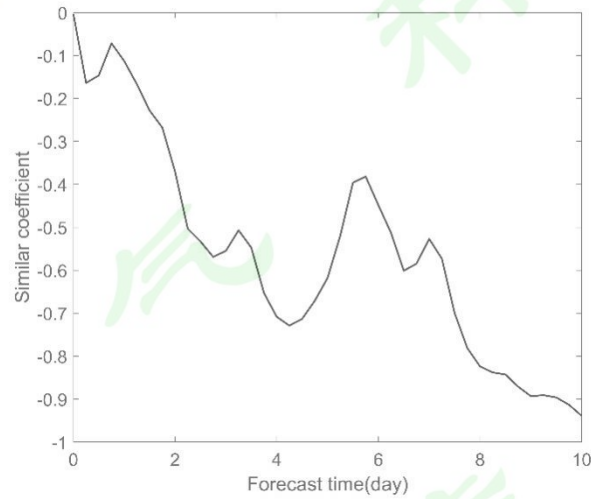


图 9 分析误差和集合平均扰动的非线性发展的相似系数随预报时间的演变图。

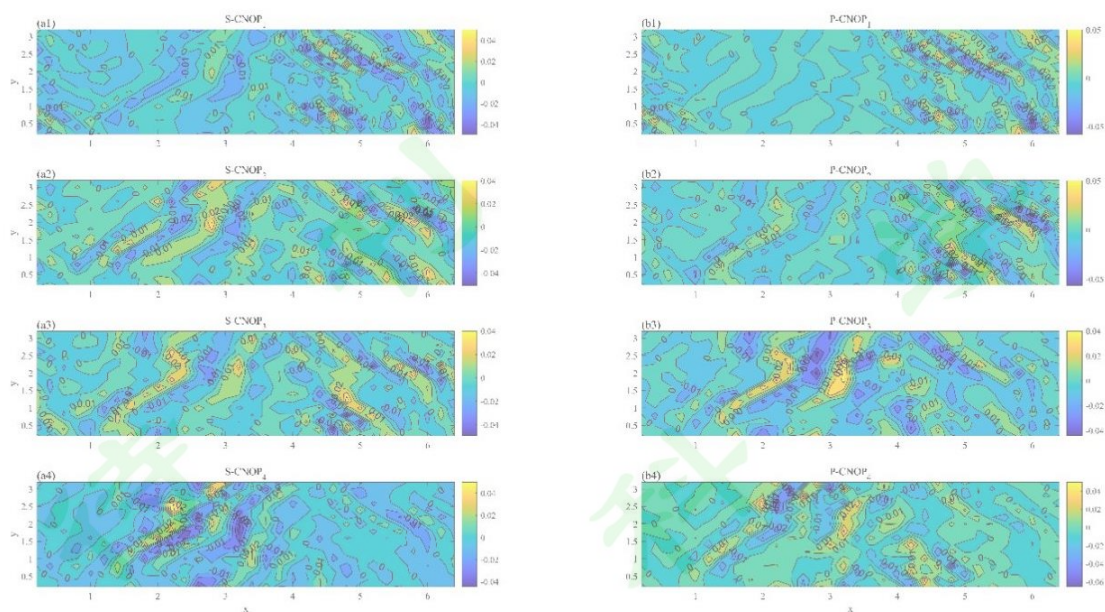
Fig. 9 The temporal evolution of similar coefficient between the nonlinear evolution of analysis error and that of ensemble mean perturbation.

5. P-CNOPs 集合预报试验

由第 4 节可知, 在 $T=2D$, $\delta = 0.2\delta_a$, $N=135$ 的参数配置下, S-CNOPs 集合预报对控制预报的改善程度达到 40%, 有效预报时效从 5 天提高到 10 天以上, 集合预报系统最可靠。因此, 本节试验将基于优化时间 $T=2D$ 且初始扰动振幅 $\delta = 0.2\delta_a$ 这一 T - δ 组合, 对于每个个例, 采用第 2.2 节介绍的 P-CNOPs 求解出 82 个 P-CNOPs 扰动, 研究 P-CNOPs 扰动的物理和动力特征, 证明 P-CNOPs 的合理性。然后采用第 4.1 节中 S-CNOPs 集合预报试验相同的试验方案构造集合成员, 进行集合预报试验。最后对比 P-CNOPs 集合预报技巧和 S-CNOPs 集合预报技巧的差异, 检验 P-CNOPs 集合预报方法的有效性和时效性。

5.1 P-CNOPs 扰动的性质

P-CNOPs 采用梯度优化算法求解正交的条件非线性最优扰动, 那么 P-CNOPs 扰动能否作为 O-CNOPs 扰动的一种近似呢? 我们首先考察 P-CNOPs 扰动的空间结构是否与传统算法 S-CNOPs 方法求出的扰动相同或者接近? 图 10 给出了随机选取的一个个例的前 4 个 S-CNOPs 和前 4 个 P-CNOPs 的空间结构, 表 3 给出了它们的相似系数。可以看出, 第一 P-CNOP 和第一 S-CNOP 空间结构很相似, 相似系数达 0.894, 而其余的 P-CNOP 和相对应的 S-CNOP 的空间结构展示了不同, 相似系数也越来越低。然后我们计算 50 个个例中每个个例的第一 P-CNOP 和第一 S-CNOP 的相似系数 (图 11), 结果表明相似系数的平均值为 0.811, 其中半数以上的个例其相似系数高于 0.9, 表明大多数个例的第一 P-CNOP 和第一 S-CNOP 的空间结构都很相似。



411

412 图 10. (a1) 至 (a4) 是前 4 个 S-CNOPS, (b1) 至 (b4) 是前 4 个 P-CNOPS。

413 Fig. 10 (a1) to (a4) are the first four S-CNOPS, and (b1) to (b4) are the first four P-CNOPS.

414 表 3 前 4 个 S-CNOPS 和对应的前 4 个 P-CNOPS 的相似系数

415 Table 3 The similar coefficients between the first four S-CNOPS and their corresponding

416 P-CNOPS

Perturbation	S-CNOP ₁	S-CNOP ₂	S-CNOP ₃	S-CNOP ₄
	-P-CNOP ₁	-P-CNOP ₂	-P-CNOP ₃	-P-CNOP ₄
Similar coefficient	0.894	0.384	0.206	-0.252

417

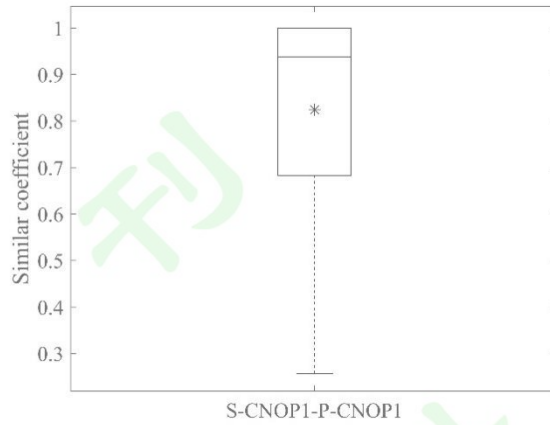


图 11 50 个个例的第一 S-CNOP 和第一 P-CNOP 的相似系数的箱线图,其中星星代表平均值。

Fig. 11 Box-whisker plot of the similar coefficients between the first S-CNOP and the first P-CNOP of 50 cases, where the asterisk represents the mean.

然后我们计算了 P-CNOPs 扰动彼此之间的相似系数,并与 S-CNOPs 扰动进行了对比(图 12),结果表明,和 S-CNOPs 一样,P-CNOPs 扰动彼此之间相互正交,表明采用 P-CNOPs 扰动构造集合成员能够保证集合成员的多样性。进一步考察 P-CNOPs 扰动的非线性发展(计算公式见公式(3)),结果表明 P-CNOPs 扰动和 S-CNOPs 扰动的非线性发展基本相等(图 13),且 P-CNOPs 扰动的非线性发展也基本满足如下增长规律: $J(P-CNOP_1) > J(P-CNOP_2) > \dots > J(P-CNOP_{82})$ 。

以上表明, P-CNOPs 扰动具备非线性最优扰动的动力学性质:正交性和增长性。

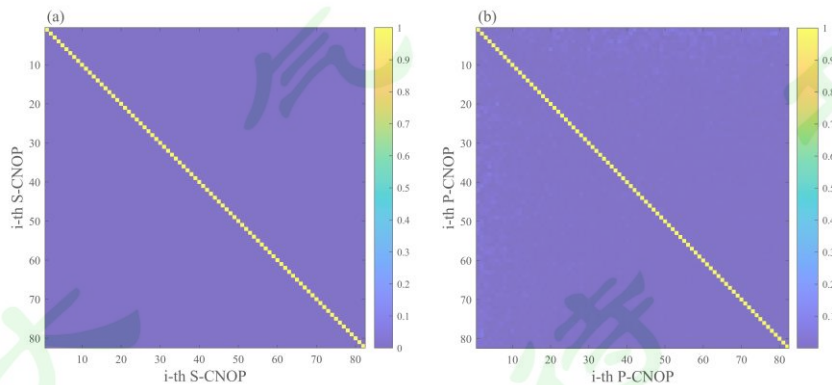


图 12 不同扰动之间的相似系数, (a) S-CNOPs, (b) P-CNOPs。

Fig. 12 The similar coefficients between different perturbations, (a)S-CNOPs, (b)P-CNOPs.

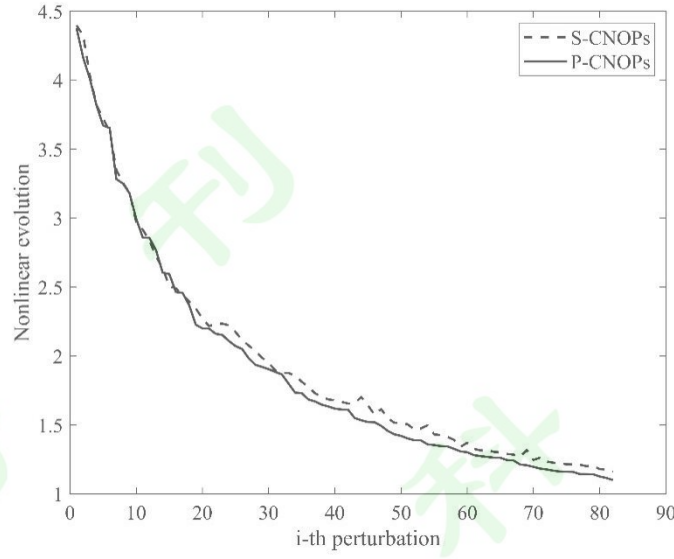


图 13 82 个 S-CNOPS 扰动（虚线）和 82 个 P-CNOPS 扰动（实线）的非线性发展。横坐标代表 82 个扰动，纵坐标代表扰动相应的非线性发展。

Fig. 13 The nonlinear evolution for 82 S-CNOPS perturbations (dashed line) and 82 P-CNOPS perturbations (solid line). The horizontal axis denotes 82 perturbations and the vertical axis represents the nonlinear evolutions.

下面从代数学的角度进一步考察 P-CNOPS 扰动和 S-CNOPS 扰动的关系。为了方便起见，将初始扰动相空间记为 V ，则 V 是 512 维线性空间。将 S-CNOPS 扰动记为 $\alpha_1, \alpha_2, \dots, \alpha_{82}$ ，P-CNOPS 扰动记为 $\beta_1, \beta_2, \dots, \beta_{82}$ ，那么 $\alpha_1, \alpha_2, \dots, \alpha_{82}$ 和 $\beta_1, \beta_2, \dots, \beta_{82}$ 是 V 中的两个向量组，且 $\alpha_1, \alpha_2, \dots, \alpha_{82}$ 两两正交， $\beta_1, \beta_2, \dots, \beta_{82}$ 两两正交， $\|\alpha_i\|_{L_2} = \|\beta_i\|_{L_2} = \delta (i=1, 2, \dots, 82)$ ，因此有

$$(\alpha_i, \alpha_j) = (\beta_i, \beta_j), (i, j = 1, 2, \dots, 82). \quad (11)$$

以 $\alpha_1, \alpha_2, \dots, \alpha_{82}$ 为基底可以构成线性空间 V 的一个 82 维线性子空间 $V_1 = \{k_1\alpha_1 + \dots + k_{82}\alpha_{82} | k_i \in \mathbb{R}, i=1, 2, \dots, 82\}$ ，类似地，以 $\beta_1, \beta_2, \dots, \beta_{82}$ 为基底也可以构成线性空间 V 的一个 82 维线性子空间 $V_2 = \{l_1\beta_1 + \dots + l_{82}\beta_{82} | l_i \in \mathbb{R}, i=1, 2, \dots, 82\}$ 。由代数学知识可知，两个维数相等的有限维线性空间一定同构，因此线性子空间 V_1 与 V_2 同构。

取与 V_1 正交的子空间的一组标准正交基 $\alpha_{83}, \alpha_{84}, \dots, \alpha_{512}$, 取与 V_2 正交的子空间的一组标准正交基 $\beta_{83}, \beta_{84}, \dots, \beta_{512}$, 那么得到 V 的两组基 $\alpha_1, \alpha_2, \dots, \alpha_{82}, \dots, \alpha_{512}$ 和 $\beta_1, \beta_2, \dots, \beta_{82}, \dots, \beta_{512}$, 这两组基的度量矩阵相等, 都等于

$$\begin{pmatrix} (\alpha_1, \alpha_1) & (\alpha_1, \alpha_2) & \cdots & (\alpha_1, \alpha_{82}) \\ (\alpha_2, \alpha_1) & (\alpha_2, \alpha_2) & \cdots & (\alpha_2, \alpha_{82}) \\ \vdots & \vdots & & \vdots \\ (\alpha_{82}, \alpha_1) & (\alpha_{82}, \alpha_2) & \cdots & (\alpha_{82}, \alpha_{82}) \\ & O & & E \end{pmatrix} = \begin{pmatrix} \delta^2 & 0 & \cdots & 0 \\ 0 & \delta^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \delta^2 \\ & O & & E \end{pmatrix} = C, \quad (12)$$

其中 E 是 430×430 的单位矩阵, O 代表零矩阵。由于 $\alpha_1, \alpha_2, \dots, \alpha_{512}$ 是 V 的一组基,

所以 V 中的任意向量 α 可以表示为 $\alpha = x_1 \alpha_1 + \cdots + x_{512} \alpha_{512}$ 。定义 V 的线性变换 $\mathcal{A}$ 如

下: $\mathcal{A}\alpha = x_1 \beta_1 + \cdots + x_{512} \beta_{512}$, 则 $\mathcal{A}$ 是 V 的一个线性变换, 且满足

$\mathcal{A}\alpha_i = \beta_i (i=1, 2, \dots, 512)$ 。对于 V 中的任意向量 β , 有 $\beta = y_1 \alpha_1 + \cdots + y_{512} \alpha_{512}$, 所

以 $\mathcal{A}\beta = y_1 \beta_1 + \cdots + y_{512} \beta_{512}$, 且有

$$(\mathcal{A}\alpha, \mathcal{A}\beta) = (x_1, x_2, \dots, x_{512}) \cdot C \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{512} \end{pmatrix} = (\alpha, \beta), \quad (13)$$

所以 $\mathcal{A}$ 是一个正交变换, 且满足 $\mathcal{A}\alpha_i = \beta_i (i=1, 2, \dots, 82)$, 即有

$$\mathcal{A}(\alpha_1, \alpha_2, \dots, \alpha_{82}) = (\beta_1, \beta_2, \dots, \beta_{82}). \quad (14)$$

这意味着正交变换 $\mathcal{A}$ 将基底 $\alpha_1, \alpha_2, \dots, \alpha_{82}$ 变换成基底 $\beta_1, \beta_2, \dots, \beta_{82}$ 。

公式(14)也可以理解为正交变换 $\mathcal{A}$ 将扰动 $S\text{-CNOP}_1, S\text{-CNOP}_2, \dots, S\text{-CNOP}_{82}$ 旋转一定的角度变成扰动 $P\text{-CNOP}_1, P\text{-CNOP}_2, \dots, P\text{-CNOP}_{82}$, 因此 $P\text{-CNOPs}$ 扰动一定具备 $S\text{-CNOPs}$ 扰动具有的性质--正交性。由于 $S\text{-CNOPs}$ 扰动是预报初始时段内非线性模式中的最快增长扰动, 所以以 $S\text{-CNOPs}$ 为基底构造的线性空

间 V_1 代表了动力学意义下最不稳定的相空间。以 P-CNOPS 扰动为基底构造的线性子空间 V_2 和 V_1 同构，这就意味着在同构的意义下，两组扰动可以相互线性表出。因此，在同构的意义下，P-CNOPS 扰动也是快速增长扰动。上述分析表明，从代数学角度来看，P-CNOPS 扰动也具备正交性和增长性两大特征。综上，我们从动力学和代数学两个角度考察，证明了 P-CNOPS 扰动的合理性。

5.2 P-CNOPS 集合预报与 S-CNOPS 集合预报结果对比

探讨了 P-CNOPS 扰动的合理性之后，我们要研究的问题是：使用 P-CNOPS 扰动构造集合成员进行集合预报，是否能够得到与 S-CNOPS 集合预报相同或者相当的预报技巧呢？图 14 给出了当优化时间为 2 天且扰动振幅 $\delta = 0.2\delta_a$ 时，P-CNOPS 集合预报技巧随集合样本数的演变情况，这里的确定性预报技巧评分（RMSE 和 ACC）以及概率预报技巧评分（BS 评分和 ROCA 评分）都是基于 50 个个例和 10 天预报时间内的所有预报时刻计算的平均值。结果表明，无论集合样本数为多少，无论是确定性预报还是概率预报，P-CNOPS 集合预报与 S-CNOPS 集合预报的预报技巧都是相当的。进一步地，在 $T=2D$, $\delta = 0.2\delta_a$, $N=135$ 的参数配置下，考察 P-CNOPS 集合预报技巧随预报时间的演变（图 15），结果表明 P-CNOPS 方法和 S-CNOPS 方法在每个预报时刻也都具有几乎相同的集合预报技巧。从图 15 可以看到 P-CNOPS 方法的集合预报技巧比 S-CNOPS 方法的集合预报技巧略微高一点，可能是因为 S-CNOPS 扰动比 P-CNOPS 扰动增长更快（见图 13），在发展过程中可能高估了初始误差的不确定性，而 P-CNOPS 扰动发展相对较慢，可能提供了更好的误差估计。

最后，我们统计了两种方法的计算耗时。对于 50 个个例，将每个个例计算 82 个 S-CNOPS 和 82 个 P-CNOPS 的时间统计出，并给出箱线图（图 16）。结果表明 50 个个例计算 S-CNOPS 的时间中位数约为 9293s，平均值为 9283s，计算 P-CNOPS 的时间中位数是 2710s，平均值为 2771s，按平均值计算，计算时间节省了约 70.15%，表明 P-CNOPS 计算效率更高。综上，采用 P-CNOPS 构造集合成员并进行集合预报，其预报水平与 S-CNOPS 相当，而且计算耗时显著减少，表明了 P-CNOPS 在集合预报领域中的有效性和时效性。

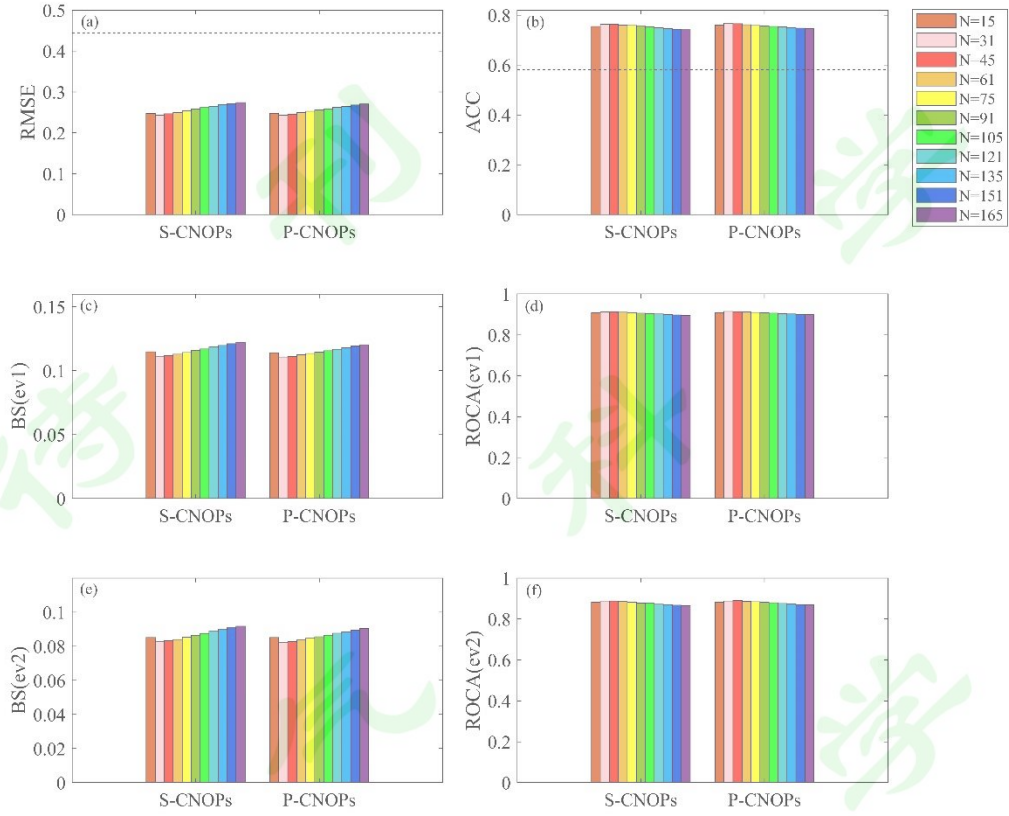
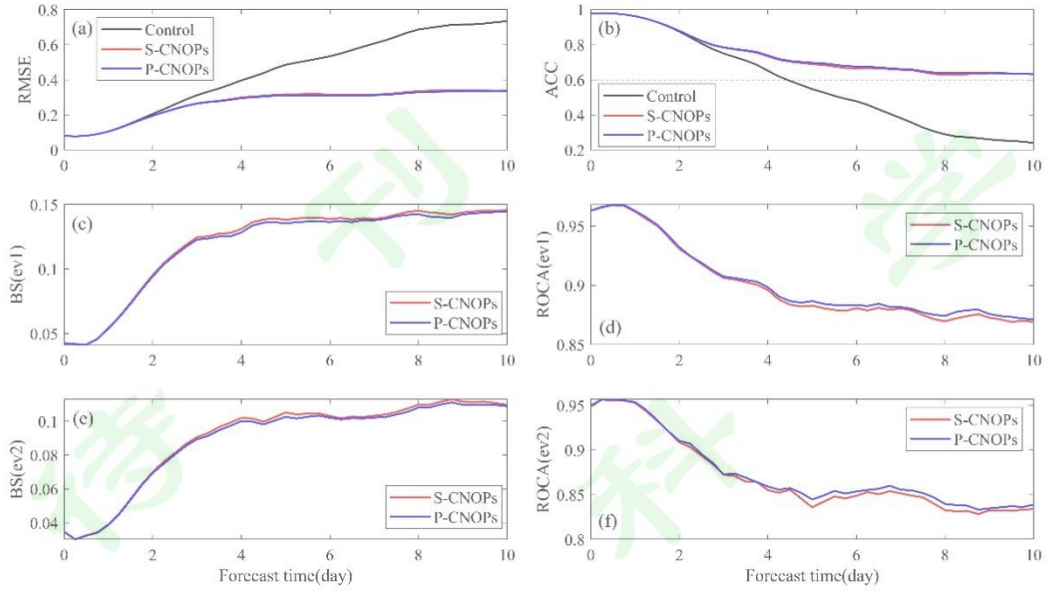


图 14 优化时间为 2 天且扰动振幅 $\delta = 0.2\delta_a$ 时 S-CNOPs 和 P-CNOPs 集合预报技巧随集合样本数的演变。(a) RMSE; (b) ACC; (c) 事件 ev1 的 BS 评分; (d) 事件 ev1 的 ROCA 评分; (e) 事件 ev2 的 BS 评分; (f) 事件 ev2 的 ROCA 评分。所有技巧评分是针对 50 个个例和 10 天内所有预报时刻的平均值。每幅图中有两簇柱状图分别对应于 S-CNOPs 和 P-CNOPs。每簇柱状图中从左到右不同颜色的柱对应于集合样本数由小 ($N=15$) 到大 ($N=165$)。 (a) 和 (b) 中黑色虚线分别代表控制预报的 RMSE 和 ACC。

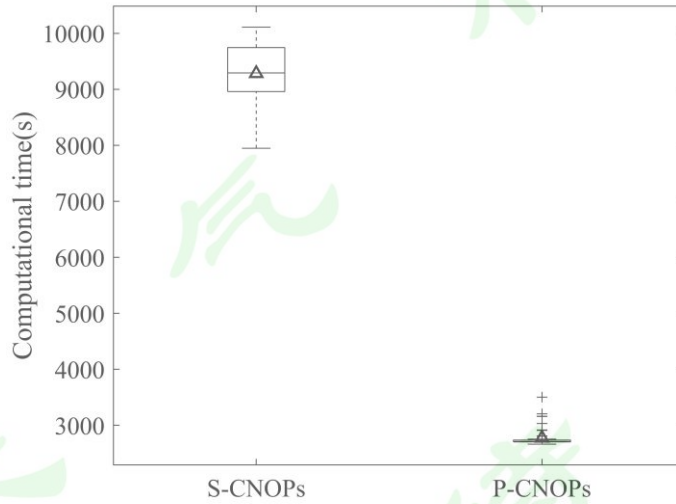
Fig. 14 Evolution of ensemble forecast skills of S-CNOPs and P-CNOPs with respect to ensemble size N with the optimization time interval of 2 days and perturbation magnitude of $\delta = 0.2\delta_a$. (a)RMSE, (b)ACC, (c)BS for the event ev1, (d)ROCA for the event ev1, (e)BS for the event ev2, (f)ROCA for the event ev2. The skills are computed by averaging over 50 cases with all the lead times within 10 days. There are two clusters of bars in each figure that correspond to S-CNOPs and P-CNOPs, respectively. The bars of different colors from left to right in each cluster of bars correspond to different ensemble sizes from small ($N=15$) to large ($N=165$). The black dashed lines in (a) and (b) represent RMSE and ACC of control forecast, respectively.



508

509 图 15 优化时间为 2 天，扰动振幅 $\delta = 0.2\delta_a$ 且集合样本数 $N=135$ 时 S-CNOPS（红）和
510 P-CNOPS（蓝）集合预报技巧随预报时间的演变，其中技巧评分同图 14.

511 Fig. 15 The temporal evolution of ensemble forecast skills of S-CNOPS (red) and P-CNOPS (blue)
512 with the optimization time interval of 2 days, perturbation magnitude of $\delta = 0.2\delta_a$ and
513 ensemble size $N=135$, where the skills are as Fig. 14.



514

515 图 16 S-CNOPS 和 P-CNOPS 的计算时间的箱线图，其中三角代表均值。

516 Fig. 16 Box-whisker plot of the computational times for S-CNOPS and P-CNOPS, where the
517 triangle represents the mean.

6. 总结与讨论

正交条件非线性最优扰动 (O-CNOPs) 是一种重要的集合预报初始扰动方法, 近年来已被广泛应用于集合预报的相关研究中并取得了成功。然而该扰动传统的顺序算法 (S-CNOPs) 计算效率有限, 计算代价高昂, 限制了 O-CNOPs 方法在集合预报中的进一步应用。基于并行思想计算 O-CNOPs 的并行算法 (P-CNOPs) 随后被提出, 但该算法的初始实现是基于非常简单的 LORENZ-96 理论模型, 本文则采用更复杂的二维正压准地转模式, 研究 P-CNOPs 的合理性、可靠性和高效性。具体结论如下:

首先, 在完美模式假定下, 针对二维正压准地转模式, 采用 S-CNOPs 产生集合预报成员, 探讨了优化时间、初始扰动振幅和集合样本数对 S-CNOPs 集合预报技巧的影响。结果表明, 在较短的优化时间 ($T=2D$)、小于初始分析误差大小的初始扰动振幅 ($\delta = 0.2\delta_a$) 及合适的集合样本数 ($N=135$) 下, S-CNOPs 集合预报能够极大改善控制预报技巧, 改善程度达 40%, 同时将控制预报的有效预报时效由 5 天延长到 10 天以上。在时间上和空间上均协调一致的离散度--集合平均预报误差关系, 表明了 S-CNOPs 集合预报系统的可靠性。然后采用相似系数和 EOF 分析方法, 探讨了 S-CNOPs 集合预报改善控制预报的原因。发现分析误差的前五个 EOF 模态与 S-CNOPs 集合平均扰动的前五个 EOF 模态呈现负相关关系, 且随着预报时间的增加 S-CNOPs 集合平均扰动的非线性发展与分析误差的非线性发展的负相关越来越强, 表明 S-CNOPs 扰动能够很好地刻画分析误差的非线性发展的演变过程, 所以能够更好地捕捉到真值, 进而改善控制预报。

进一步地, 采用 P-CNOPs 求解正交条件非线性最优扰动。一方面, 从动力学角度, 验证了 P-CNOPs 扰动彼此两两正交, P-CNOPs 扰动和 S-CNOPs 扰动的非线性发展基本相等, 表明 P-CNOPs 扰动具备正交性和增长性。另一方面, 从代数学的角度, 证明了 P-CNOPs 扰动和 S-CNOPs 扰动可以通过第一类正交变换, 即旋转变换, 相互转化, 表明 P-CNOPs 扰动具备正交性。又由于以 P-CNOPs 扰动为基底构成的线性子空间和以 S-CNOPs 扰动为基底构成的线性子空间是同构的, 表明两组基底可以相互线性表出。因此, 在同构的意义下, P-CNOPs 扰动也是快速增长扰动, 即 P-CNOPs 扰动具备增长性。上述研究从动力学和代数学

两方面证实了 P-CNOPs 的合理性。然后采用 P-CNOPs 扰动构造集合成员进行集合预报试验, 结果表明无论是确定性预报还是概率预报, P-CNOPs 集合预报与 S-CNOPs 集合预报的预报技巧都是相当的, 同时, P-CNOPs 极大地减少了求解条件非线性最优扰动的计算时间, 节省了计算资源, 表明了 P-CNOPs 的有效性和时效性。

综上所述, 本文基于二维正压准地转模式, 首先探讨了 S-CNOPs 集合预报方法在提高确定性预报技巧方面的有效性, 通过时间上和空间上均协调一致的离散度--集合平均预报误差关系证明了 S-CNOPs 集合预报系统的可靠性, 并进行了原因分析; 然后采用 P-CNOPs 计算正交条件非线性最优扰动, 并进行集合预报试验, 从动力学和代数学两方面证明了 P-CNOPs 在理论上的合理性, 揭示了 P-CNOPs 提供了与 S-CNOPs 集合预报相当的预报技巧, 同时前者计算效率显著高于后者, 从而拓展了 O-CNOPs 方法在数值天气预报和气候预测中的应用。

截止到目前, 我们的研究表明 P-CNOPs 在 LORENZ-96 模式和二维正压准地转模式的集合预报中均表现良好, 能够显著提高 S-CNOPs 集合预报的时效性, 节省计算资源。因此, P-CNOPs 是计算复杂模式 O-CNOPs 的潜在高效算法, 可望未来能广泛应用于实际的数值天气预报模式。在未来的研究中, 我们会将 P-CNOPs 推广到更复杂的业务数值预报模式的集合预报中, 例如 Weather Research Forecast (WRF) 模式和 Global/Regional Assimilation and Prediction System (GRAPES) 模式等。期望通过努力, 能够为天气预报和气候预测提供一个具有较好理论基础、更高预报技巧且具有更高时效性的集合预报方法。

参考文献 (References)

- Anderson J L. 1997. The impact of dynamical constraints on the selection of initial conditions for ensemble predictions: Low-order perfect model results[J]. Mon. Wea. Rev., 125(11): 2969-2983. doi:10.1175/1520-0493(1997)125<2969:TIODCO>2.0.CO;2
- Birgin E G, Martínez J M, Raydan M. 2000. Nonmonotone spectral projected gradient methods on convex sets[J]. SIAM Journal on Optimization, 10(4): 1196-1211. doi:10.1137/S1052623497330963
- Buckingham C, Marchok T, Ginis I, et al. 2010. Short- and medium-range prediction of tropical

575 and transitioning cyclone tracks within the NCEP global ensemble forecasting system[J].
 576 Wea. Forecasting, 25(6): 1736-1754. doi:10.1175/2010WAF2222398.1

577 Buizza R, Tribbia J, Molteni F, et al. 1993. Computation of optimal unstable structures for a
 578 numerical weather prediction model[J]. Tellus A, 45(5): 388-407.
 579 doi:10.3402/tellusa.v45i5.14901

580 Buizza R, Houtekamer P L, Pellerin G, et al. 2005. A comparison of the ECMWF, MSC, and
 581 NCEP global ensemble prediction systems[J]. Mon. Wea. Rev., 133(5): 1076-1097.
 582 doi:10.1175/MWR2905.1

583 Chen G, Wang B, Liu J J. 2021. Study on the sensitivity of initial perturbations to the development
 584 of a vortex observed in Southwest China [J]. J. Geophys. Res.: Atmospheres, 126.
 585 doi:10.1029/2021JD034715

586 Cheung K K W, Johnny C L Chan. 1999. Ensemble forecasting of tropical cyclone motion using a
 587 barotropic model. Part I: perturbations of the environment [J]. Mon. Wea. Rev., 127(6):
 588 1229–1243. doi:10.1175/1520-0493(1999)127<1229:EFOTCM>2.0.CO;2

589 Daron J D, Stainforth D A. 2013. On predicting climate under climate change[J]. Environ. Res.
 590 Lett., 8(3): 034021. doi: 10.1088/1748-9326/8/3/034021

591 Descamps L, Talagrand O. 2007. On some aspects of the definition of initial conditions for
 592 ensemble prediction[J]. Mon. Wea. Rev., 135(9): 3260-3272. doi:10.1175/MWR3452.1

593 杜钧, 陈静. 2010. 单一值预报向概率预报转变的基础: 谈谈集合预报及其带来的变革[J]. 气
 594 象, 36(11):1-11. Du Jun, Chen Jing. 2010. The Corner Stone in Facilitating the Transition
 595 from Deterministic to Probabilistic Forecasts-Ensemble Forecasting and Its Impact on
 596 Numerical Weather Prediction[J]. Meteor. Mon. (in Chinese), 36(11): 1-11.
 597 doi:10.7519/j.issn.1000-0526.2010.11.001

598 Du J, Berner J, Buizza R, et al. 2018. Ensemble Methods for Meteorological Predictions. NCEP
 599 Office Notes 493, 66pp. doi: 10.7289/V5/ON-NCEP-493

600 Duan W S, Mu M, Wang B. 2004. Conditional nonlinear optimal perturbations as the optimal
 601 precursors for El Nino-southern oscillation events[J]. J. Geophys. Res.: Atmospheres,
 602 109(D23). doi:10.1029/2004JD004756

603 Duan W, Huo Z. 2016. An approach to generating mutually independent initial perturbations for

ensemble forecasts: orthogonal conditional nonlinear optimal perturbations[J]. J. Atmos. Sci., 73(3): 997-1014. doi:10.1175/Jas-D-15-0138.1

Duan W S, Li X Q, Tian B. 2018. Towards optimal observational array for dealing with challenges of El Niño-Southern Oscillation predictions due to diversities of El Niño[J]. Climate Dyn., 51, 3351-3368. doi:10.1007/s00382-018-4082-x

Durran D R, M Gingrich. 2014. Atmospheric predictability: Why butterflies are not of practical importance[J]. J. Atmos. Sci., 71(7):2476–2488. doi:10.1175/JAS-D-14-0007.1

Feng J, Ding R, Li J, et al. 2016. Comparison of nonlinear local Lyapunov vectors with bred vectors, random perturbations and ensemble transform Kalman filter strategies in a barotropic model[J]. Adv. Atmos. Sci., 33: 1036-1046. doi:10.1007/s00376-016-6003-4

Feng J, Qin X H, Wu C Q, et al. 2022. Improving typhoon predictions by assimilating the retrieval of atmospheric temperature profiles from the FengYun-4A's Geostationary Interferometric Infrared Sounder (GIIRS)[J]. Atmospheric Research, 280, 106391. doi:10.1016/j.atmosres.2022.106391

Fortin V, Abaza M, Ancil F, et al. 2014. Why Should Ensemble Spread Match the RMSE of the Ensemble Mean? [J]. Journal of Hydrometeorology, 15(4): 1708-1713. doi:10.1175/Jhm-D-14-0008.1

Hamill T M, Snyder C, Morss R E. 2000. A comparison of probabilistic forecasts from bred, singular-vector, and perturbed observation ensembles[J]. Mon. Wea. Rev., 128(6): 1835-1851. doi: 10.1175/1520-0493(2000)128<1835:ACOPFF>2.0.CO;2

Huo Z, Duan W. 2019. The application of orthogonal conditional nonlinear optimal perturbations method to typhoon track ensemble forecasts[J]. Science China Earth Sciences, 62: 376-388. doi:10.1007/s11430-018-9248-9

Huo Z H, Duan W S, Zhou F F. 2019. Ensemble Forecasts of Tropical Cyclone Track with Orthogonal Conditional Nonlinear Optimal Perturbations[J]. Adv. Atmos. Sci., 36, 231-247. doi:10.1007/s00376-018-8001-1

Leith C E. 1974. Theoretical skill of Monte Carlo forecasts[J]. Mon. Wea. Rev., 102(6), 409–418. doi: 10.1175/1520-0493(1974)102<0409:TSOMCF>2.0.CO;2

Lorenz E N. 1963. Deterministic Nonperiodic Flow [J]. J. Atmos. Sci., 20(2): 130-141. doi: 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2

634 Lorenz E N. 1965. A study of the predictability of a 28-variable atmospheric model[J]. *Tellus*,
635 17(3): 321-333. doi:10.3402/tellusa.v17i3.9076

636 Jiang L, Duan W S, Wang H. 2024. The Sensitive Area for Targeting Observations of Paired
637 Mesoscale Eddies associated with Sea Surface Height Anomaly Forecasts[J]. *J. Geophys.*
638 *Res.: Oceans*, 129. doi:10.1029/2023JC020572

639 Jiang Z, Mu M. 2009. A comparison study of the methods of conditional nonlinear optimal
640 perturbations and singular vectors in ensemble prediction[J]. *Advances in Atmospheric*
641 *Sciences*, 26(3): 465-470. doi:10.1007/s00376-009-0465-6

642 Jiang Z, Mu M, Wang D. 2009. Ensemble prediction experiments using conditional nonlinear
643 optimal perturbation[J]. *Sci. China Ser. D-Earth Sci.* 52(4): 511-518. doi:
644 10.1007/s11430-009-0042-y

645 Ma J, Duan W, Liu Z, et al. 2025. A new method to calculate nonlinear optimal perturbations for
646 ensemble forecasting[J]. *Adv. Atmos. Sci.*, doi:10.1007/s00376-024-4069-y

647 Mason S J, Graham N E. 1999. Conditional probabilities, relative operating characteristics, and
648 relative operating levels [J]. *Wea. Forecasting*, 14(5): 713-725. doi:
649 10.1175/1520-0434(1999)014<0713:CPROCA>2.0.CO;2

650 Mu M, Zeng Q. 1991. New development on existence and uniqueness of solutions to some models
651 in atmospheric dynamics[J], *Adv. Atmos. Sci.*, 8, 383-398. doi:10.1007/BF02919262

652 Mu M, Duan W S, Wang B. 2003. Conditional nonlinear optimal perturbation and its applications
653 [J]. *Nonlinear Processes in Geophysics*, 10(6): 493-501. doi:10.5194/npg-10-493-2003

654 Mu M, Zhang Z Y. 2006. Conditional nonlinear optimal perturbations of a Two-Dimensional
655 Quasi-geostrophic model[J]. *J. Atmos. Sci.*, 63, 1587–1604. doi:10.1175/JAS3703.1

656 Mu M, Jiang Z N. 2008. A new approach to the generation of initial perturbations for ensemble
657 prediction: Conditional nonlinear optimal perturbation[J]. *Chinese Science Bulletin*, 53(13):
658 2062-2068. doi:10.1007/s11434-008-0272-y

659 Mu M, Zhou F F, Wang H L. 2009. A method for identifying the sensitive areas in targeted
660 Observations for tropical cyclone prediction: Conditional Nonlinear optimal perturbation[J].
661 *Mon. Wea. Rev.*, 137, 1623-1639. doi:10.1175/2008MWR2640.1

662 Murphy A H, Epstein E S. 1989. Skill scores and correlation coefficients in model verification [J].
 663 Mon. Wea. Rev., 117(3): 572-582.
 664 doi:10.1175/1520-0493(1989)117<0572:SSACCI>2.0.CO;2
 665 Palmer T N, Döring A, Seregin G. 2014. The real butterfly effect [J]. Nonlinearity, 27(9): 123-141.
 666 doi:10.1088/0951-7715/27/9/R123
 667 Palmer T. 2019. The ECMWF ensemble prediction system: Looking back (more than) 25 years
 668 and projecting forward 25 years [J]. Quart. J. Roy. Meteor. Soc., 145: 12-24.
 669 doi:10.1002/qj.3383
 670 Qin X H, Duan W S, Chan P W, et al. 2023. Effects of Dropsonde Data in Field Campaigns on
 671 Forecasts of Tropical Cyclones over the Western North Pacific in 2020 and the Role of CNOP
 672 Sensitivity[J]. Adv. Atmos. Sci., 40, 791–803. doi:10.1007/s00376-022-2136-9
 673 Scaife A A, Arribas A, Blockley E, et al. 2014. Skillful long-range prediction of European and
 674 North American winters[J]. Geophys. Res. Lett., 41(7): 2514-2519.
 675 doi:10.1002/2014GL059637
 676 Toth Z, Kalnay E. 1993. Ensemble forecasting at NMC: The generation of perturbations[J]. Bull.
 677 Amer. Meteor. Soc., 74(12): 2317-2330.
 678 doi:10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2
 679 汪叶, 段晚锁. 2019. 初始扰动振幅和集合样本数对 CNOPs 集合预报的影响[J]. 大气科学,
 680 43(4): 915-929. Wang Ye, Duan Wansuo. 2019. Influences of Initial Perturbation Amplitudes
 681 and Ensemble Sizes on the Ensemble Forecasts Made by CNOPs Method[J]. Chinese Journal
 682 of Atmospheric Sciences (in Chinese), 43(4): 915-929. doi:
 683 10.3878/j.issn.1006-9895.1810.18219
 684 Yamaguchi M, Nakazawa T, Hoshino S. 2012. On the relative benefits of a multi-centre grand
 685 ensemble for tropical cyclone track prediction in the western North Pacific[J]. Quart. J. Roy.
 686 Meteor. Soc., 138(669): 2019-2029. doi:10.1002/qj.1937
 687 Zhang H, Duan W S, Zhang Y C. 2023. Using the Orthogonal Conditional Nonlinear Optimal
 688 Perturbations Approach to Address the Uncertainties of Tropical Cyclone Track Forecasts
 689 Generated by the WRF Model[J]. Wea. Forecasting, 38, 1907-1933.

690 doi:10.1175/WAF-D-22-0175.1

691 Zhu Y, Fu B, Yang B, et al. 2023. Quantify the coupled GEFS forecast uncertainty for the weather
692 and sub-seasonal prediction[J]. J. Geophys. Res.: Atmospheres, 128, e2022JD037757.
693 doi:10.1029/2022JD037757