

基于最新全球集合预报系统的可预报性评估



作者简介:朱跃建,气象学家,美国国家海洋与大气管理局(NOAA),环境预报中心(NCEP),环境模式中心(EMC)集合预报系统研发室主任,北美集合预报系统(NAEFS)美方首席代表和 NUOPC UEO 共同负责人。南京信息工程大学兼职教授,中国气象局国家气象中心咨询专家,中国气象局数值预报中心科学指导专家,现任和曾任多项世界气象组织(WMO)专家组成员,其中包括 WMO/WWRP/JWGFVR 专家组成员、WMO/WWRP/TIGGE-WG 专家组成员、WMO/WWRP/PDEF-WG 专家组成员、WMO/CBS/ET-EPS 美方首席代表、WMO/CBS/DPFS-ET-OFPS 美方首席代表、WMO/CBS/ET-OWFPS 美方首席代表、WMO/WWRP/TLFDP-ISSC 和 WMO/SCMREX-ISSC 科学指导专家,曾任美中大气科学联合工作组(JWG/US-China)美方数值预报首席代表。

朱跃建*

美国国家海洋与大气管理局(NOAA)/国家气象局(NWC)/美国国家环境预报中心(NCEP)环境模式中心(EMC),马里兰州 20740

摘要 为了提供有价值且可靠的概率(或者不确定性)预报,最新的全球集合预报系统已在美国国家环境预报中心日常业务运行,以满足社会需求。通过对各个关键要素的概率预报统计检验,可为广大用户提供这些概率预报的信心指数。但是预报(或集合预报)能力不仅取决于我们使用的预测要素,而且与时间和空间分辨率,极端事件或者高影响天气,以及预报时效有关。以大尺度天气系统预报为例,通常选择北半球 500 hPa 位势高度距平相关指数或概率指数表征模式的预报能力。如参照北半球 500 hPa 位势高度的距平相关指数(60% AC)或概率预报技巧指数(25% CRPSS),美国全球集合预报系统能够提供大约 10 d 的技巧预报。从全球集合预报系统输出的各预报要素,满足不同时空尺度需求的角度进行讨论,其可预报性(或预报极限)能够为模式研发人员、一线预报员和用户提供参考。尤其是对大气可预报性的深入研究,对于从科学与技术角度全面提升数值预报系统水平非常重要。当能够确定可预报性(或是预报误差)的真实来源时,科学家(包括模式研发人员)就能够有针对性地修改与完善。将传统的可预报性研究与改进的能够更客观地表述预报不确定性的集合预报相结合,所得可预报性将提供另一种有价值的参考。可预报性研究总体表明,全球集合预报系统对行星波、大尺度和天气尺度的系统(或者过程)可能分别具备约 15、12、10 d 的预报能力。对于热带天气过程的预报,如果进一步改善模式偏差和物理参数化过程,其 MJO (Madden-Julian Oscillation) 预报技巧可以延长至 32.5 d。

关键词 集合预报;预报的不确定性;可预报性

从 2005 年开始,世界气象组织(WMO)发起了为期 10 a(2005—2014 年)的世界天气研究计划(WWRP)——观测系统研究与可预报性试验(THORPEX),通过研究和开发全球集合预报系统,加速提升 1 天至 2 周高影响天气或者极端天气预报的准确率,以及改善对社会、经济和环境的影响(Bougeault, 2010; Swinbank et al., 2016)。美国国家气象中心(NCEP)全球集合预报系统(GEFS)是 THORPEX TIGGE (THORPEX 交互式大集合)项目的主要成员,旨在为研究和应用人员提供实时全球集合预报数据以及相关服务支持。GEFS 自 1992 年 12 月在美国国家环境预报中心投入业务预报至今,已为美国和世界各地的广大用户提供了有效的、可靠的、实时的概率预报(Toth and Kalnay,

* 联系人, E-mail: Yuejian.Zhu@noaa.gov

2019-11-01 收稿, 2019-12-21 接受

引用格式:朱跃建, 2020. 基于最新全球集合预报系统的可预报性评估[J]. 大气科学学报, 43(1): 193-200.

Zhu Y J, 2020. An assessment of predictability through state-of-the-art global ensemble forecast system[J]. Trans Atmos Sci, 43(1): 193-200. doi:10.13878/j.cnki.dqkxxb.20191101013. (in Chinese).

1993;Toth et al.,2001;Zhu et al.,2002;Zhu,2005)。

近年来,随着全球经济的快速发展,国家环境预报中心及政府决策部门和普通用户对概率预报和无缝隙预报的需求已大大增加,该技术对提升公共安全、改善生活质量,制定发展计划以及推动国家经济增长提供了宝贵的指导性信息(Zhu et al.,2002;Zhu,2005;Zhu et al.,2018)。如能准确地了解大气可预报性(或潜在预报能力)和实际预报能力(目前的预报技巧),将会极大增强对地面降水、地表温度、地面气压、风、云量和对流有效位能(CAPE)等各种要素的预报减灾能力,并给预报员提供有价值的指导。时间尺度包括短临天气、两周,次季节到季节预报(Zhu et al.,2019a,2019b),对预测极端事件和高强度天气过程也一定有所帮助。

大气可预报性最初是作为一个科学问题从理论上进行探讨的(Lorenz,1969),然后根据天气预报的实际应用和一些假设对它进行了数值模拟和经验模型研究,得到了很多有应用意义的结论(Lorenz,1982,1996)。其后,许多气象学家还对季节和(或)气候的预测和可预报性(Shukla,1998)等进行了深入研究。直到最近,最新数值预报模型的内在预报能力已被用来进一步研究和探讨大气不同时空尺度运动的动能、温度、湿度和降水方面(Ying and Zhang,2017)的潜在可预报性。同时,基于全球大气系统(尤其是集合预报系统)的各种研究也已在各个区域展开,如北半球(NH)的阻塞过程(He et al.,2019)以及大尺度天气形势和热带环流(Zhu et al.,2019a,2019b)。通过考虑初始和模式不确定性,最新全球集合预报系统可以作为一个最优的数值预报系统,定量描述跨时空尺度天气过程存在的可预报性。如果模式动力学过程和物理参数化方案是完美的,其集合预报通过初始扰动形成的每一个成员的预报能力在统计意义上应该是等价的,其形成的集合预报应无偏,并且能够真正描述预报不确定性。如果初始分析及其扰动都能表示实际的观测(假设),并且假定模式完备,则预报结果都应该与观测“真值”相当。同时,集合平均预报相对于控制预报和扰动预报应统计最优。当集合预报系统中的扰动预报与控制预报相互独立时,那么集合平均预报和任一扰动预报(以及控制预报)的距平相关(或者均方根误差)技巧,即可预报能力(或者技巧)应当是最佳的。

1 集合预报系统

NCEP 全球集合预报系统(GEFS),最初使用增长繁殖矢量方法(经过重新尺度化,2015年12月之前)和EnKF(集合卡尔曼滤波)分析方法(2015年12月之后)生成的一组初始扰动,自从1992年12月投入NCEP业务运行以来,一直持续至今(Toth and Kalnay,1993;Wei et al.,2008;Zhou et al.2016;Zhou et al.,2017)。集合预报的演变过程可以如图1所示。一个可靠的集合预报系统,从一组初始扰动(分析的不确定性)开始,随着积分时间增长,预报误差(不确定性)不断增长,在实际应用中,称之为集合预报系统的“离散度”(也可以表述为无偏状态下的样本标准差)。理论上,集合离散度描述预报不确定性,最佳状态下的集合预报,其系统离散度应和集合平均预报的均方根误差一致,当其均方根误差达到饱和状态(稳定不增长),也失去了预报能力。另一方面,当所有包含扰动和未扰动的单一预报相互独立时,总体平均值(算术平均)预报应该是最优且被集合预报成员(集合云)所包围。图1中阴影部分从左至右分别表示初始时刻的扰动分布(分析不确定性), t_1 时刻预报的不确定范围和 t_2 时刻的预报概率空间。其预报不确定性随着时间演变而变大。理想的集合预报概率空间(t_1 或 t_2 时刻)应该包含了未来预报的所有可能,其实际观测值(真值)是不应该落在阴影区之外的。事实上,过去20多年来,NCEP全球集合预报系统(GEFS)一直在提供有价值的天气和2周概率预报指导。图2显示了1999年至今的NCEP全球集合预报系统其北半球(NH)500 hPa位势高度概率预报技巧评估的时间演变。基于不同的CRPSS技巧分级(CRPSS分别为0.25、0.35、0.45、0.55、0.65、0.75),图2纵坐标表示的是在某个技巧分级能力下的可参考预报天数。虚线表示其季节平均值形成的时间序列,而实线表示13个月的时间滑动平均。CRPSS为0.25(25%)相当于确定性预报的60%异常距平相关指数(AC)。从长时间滑动平均来看(平滑了季节变化),总体预报能力已从此前(1999年)的6 d提高到了最近的10 d(针对0.25 CRPSS),这表明在过去20 a,通过WMO的THORPEX研究计划,数值预报能力20 a来延长了4 d,比通常每10 a延长1 d预报能力整整提高了一倍,达到了WMO THORPEX研究计划的预期目标。从这近20 a的时间序列中可看出2009—2010年和2017—2018年的

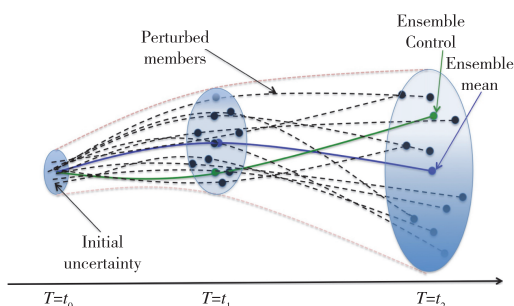


图1 从最初的分析/扰动开始的集合预报系统演变示意图,在 $t=0$ 时具有完全的分析不确定性(其中包括观测的不确定性,图中最左侧的阴影部分),随着时间增长到 t_1 和 t_2 时刻(集合预报的离散度或不确定性增加)。绿色实线表示未扰动(控制预报)预报;蓝色实线表示集合预报平均值;黑色虚线表示每一个集合预报成员的预报。其阴影区域分别表示 t_1 和 t_2 时刻预报的不确定性范围。理想的集合预报其预报集(实线)应该包括所有可能发生的事件

Fig.1 Schematic diagram of an ensemble forecast starting from initial analyses/perturbations with full analysis uncertainty at $t=0$, growing (increasing ensemble spread or uncertainty) with time evolution to t_1 and t_2 . Solid green is for unperturbed (ensemble control) forecast; solid blue represents ensemble mean; and dash blacks are perturbed ensemble forecasts. The shaded areas represent uncertainties of initial and forecasts at t_1 and t_2 .

年变化很大。这种现象可能与 ENSO 事件有关,而与模型系统本身无关。事实已经证明,数值模式的预报能力在不断改善。尤其在过去的几年中,随着数值模式物理参数化和数据同化,观测系统和先进的高性能计算机(HPC)的许多改善,目前的 NCEP 全球集合预报系统(GEFS),尤其是 SubX(次季节预报多模型实验)版本,从天气到次季节,在所有尺度上都达到了历史上最好的表现(Zhu et al., 2017, 2018; Pegion et al., 2019)。在此项研究中使用了 NCEP GEFS 的两个最新版本,其中一个为 GEFS SubX 版本(Zhu et al., 2018; Guan et al., 2019; Li et al., 2019; Zhu et al., 2019a, 2019b)。另一个是 FV3 GEFS 版本(Zhou et al., 2019),使用新的动力框架(FV3),更高(且均匀)模式水平分辨率,GFDL 的微物理过程(代替了 Zhao-Carr 微物理过程)和许多其他修正调整的物理参数。与 GEFS SubX 版本相比, FV3-GEFS 版本提供了更高且均匀的水平分辨率

(约 25 km)(GEFS SubX 版本的水平解析度为 33 km(0~8 d),55 km(8~35 d))。试验选取 2017 年 4 月 1 日至 2018 年 4 月 6 日期间,所有集合预报都包含 20 个扰动预报和一个控制(无扰动)预报。两种模式试验都使用 64 个相同的混合垂直坐标,预报至 35 天用以支持美国气候预测中心(CPC)第 3 和第 4 周的指导预报。

2 检验方法介绍

通过图 2 的时间演变结果以及许多早期的研究,GEFS 模式的天气学性能(0~14 d)已经得到了总结(Zhou et al., 2017; Zhu et al., 2017, 2018),但是延伸预报至 35 天的 GEFS 已经覆盖了次季节(3~4 周)的时间尺度,其预报能力和效果直到最近才进行了系统性的评估(Zhu et al., 2018; Guan et al., 2019; Li et al., 2019; Pegion et al., 2019),作为 NOAA SubX 项目的一部分,GEFS 提供了 18 a(1999—2016 年)的预报回算(reforecast or hindcast)。本文提出的研究包括将 GEFS SubX 结果与新近研发的基于 FV3 的 GEFS(下一代业务模式)结果进行比较,其中包括动力模式、水平分辨率、微物理方案和其他物理参数化方案的调整和设计。各种评估指标可用于评估预报系统的预报技能和可预报性,当然这样的评估高度依赖于不同的时空尺度,不同的预报区域和不同的预报要素等。在这项研究中,使用北半球 500 hPa 位势高度的距平相关(AC)评分来评估天气预报技能和天气的潜在预报能力(或者称可预报性)。同时,传统的实时多变量(RMM)MJO(Madden-Julian 振荡)指数(WH index; Wheeler and Hendon, 2004; Lin et al., 2008; Gottschalack et al., 2010)也用于评估热带大气环境预报能力和潜在预报技能(可预报性)。

为了展示潜在的预报能力,将以下主要假设应用于此评估:1)模式的初始扰动代表了观测系统和分析系统的真实误差(不确定性);2)数值模式是完备的且“无偏差”;3)集合预报系统是完备的;4)集合预报离散度确实代表真实大气的未来不确定性;5)所有扰动(未扰动-控制)预报成员等价的,并且每一个成员都可能是真值(在 1, 2, 3 假设的前提下);6)集合预报平均值是最佳预报值。为了显示最新全球集合预报系统的预报效果,图 3 显示了北半球 500 hPa 位势高度平均的均方根(RMS)误差与集合预报离散度的随预报时效的统计分布。与 GEFS SubX 预报相比, FV3-GEFS 的 RMS 误差更

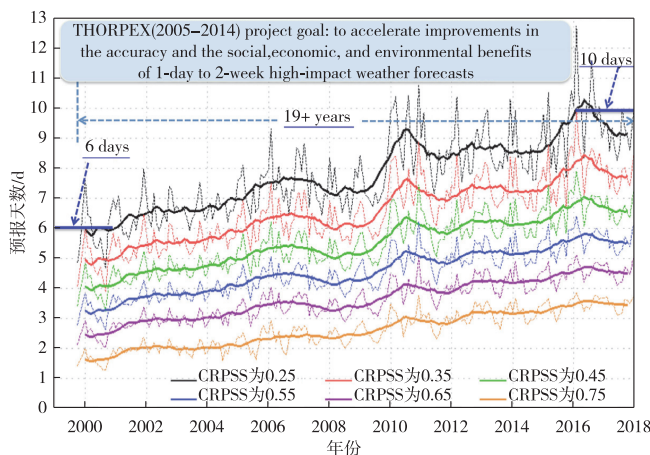


图2 过去近 20 a (1999—2018 年)NCEP 业务运行的全球集合预报系统 (GEFS) 的北半球 (NH) 500 hPa 位势高度的 CRPSS 技巧评分。其时间序列分别表示 CRPSS 为 0.25 (黑色), 0.35 (红色), 0.45 (绿色), 0.55 (蓝色), 0.65 (紫色), 0.75 (橙色) 的月平均值 (细线), 和 13 个月的滑动平均 (粗线)

Fig.2 Northern Hemisphere (NH) 500 hPa geopotential height CRPSS for NCEP operational Global Ensemble Forecast System (GEFS) of past 19+ years (1999—2018). The time series displays the monthly average scores (thin lines) for CRPSS = 0.25 (black), 0.35 (red), 0.45 (green), 0.55 (blue), 0.65 (purple), 0.75 (orange) respectively, and 13 months running mean scores (thick lines). The horizontal axis is for years from 1999 to 2018; and vertical axis is for forecast days of presented CRPSSs

小, RMS 误差与离散度的比例更好。大约第 17 天前, 两个实验的 RMS 误差都超过了气候误差 (水平), 这意味着日常预报在第 17 天之后, 就超出了其预报能力而无预报技巧。同时, 集合离散度已经达到饱和水平 (不再增长), 其集合离散度的饱和值和气候值 (分析) 之间的 RMS 差相近 (绿色粗虚线), 这足以表明此集合系统已经达到或接近最佳状态。此外, 通过 RMS 误差和集合离散度比例的整体空间分布, 可以证明系统的完备性 (图略)。集合平均误差和集合离散度的散点图 (scatter plot, 图略) 也从另一方面通过其平均统计数据进一步验证了集合预报系统的完备性 (图 3)。

3 结果分析比较

通过为期一年 (2017 年 4 月至 2018 年 4 月) 的试验, 其中包含了 20 个扰动和 1 个未扰动 (控制) 预报, 预报时效为 35 d。使用 NCEP 的实时全球数据同化系统 (GDAS) 分析和 NCEP/NCAR 40 a 再分析的气候场作为参考, 进行了集合预报系统的天气尺度预报和 MJO 评估。

3.1 北半球 (NH) 中高纬度天气预报

在中高纬度的天气预报检验中, 有许多预报量和检验方法可供选择, 但是 500 hPa 位势高度的距平相关系数 (AC) 是最具代表性和常用的一种。通过北半球 (NH) 中高纬度 (20° ~ 80°N) 集合预报平均的距平相关系数可以表达当前实际预报能力和未来可预报能力 (图 4)。显然, 一年的统计平均, FV3 GEFS (深黑色-粗线) 的实际预报技能比 SubX GEFS (深红色-粗线) 稍好。FV3 GEFS 和 SubX GEFS 都显示出相似的潜在预报技能 (浅红色/浅黑色-细线), 同时也表明通过使用 50% AC 作为其参考的预报能力, 可以将当前的预报技能再延长 2 d (第 12 至 14 天)。SubX GEFS 的潜在预报技能 (或可预报性, 深红色-粗线) 比 FV3 GEFS (深黑色-粗线) 略好, 其潜在预报能力时效较长 (约为 14 d 以上)。可能的原因是: 1) SubX GEFS 的预报离散度比同期的 FV3 GEFS 要小 (under-dispersion, 图 3 黑色虚线), 这可能导致有更高的 AC 评分; 2) 目前的 FV3 GEFS 系统的设计可能并不是最佳的选择, 可能需要进一步的改进。

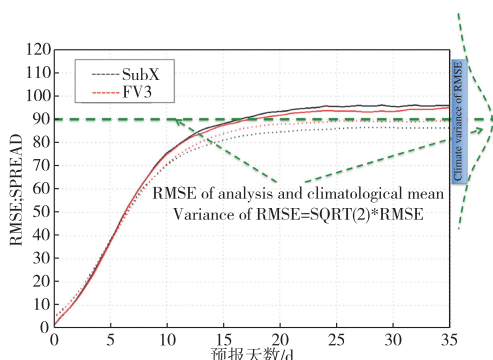


图3 基于两个集合预报系统(SubX版本(黑线)和FV3版本(红线))的一年平均(2017年4月至2018年4月)的北半球(NH)500 hPa位势高度均方根(RMS)误差(实线)和集合预报离散度(虚线)的35 d预报。观测/分析和气候平均值之间的均方根(RMS)误差以绿色虚线(约90 m)表示。右侧纵坐标所绘是均值附近的气候变化空间

Fig.3 Northern Hemisphere (NH) 500 hPa geopotential height RMS error (solid) and ensemble spread (dash) of 35 days forecasts from one year experiments based on two ensemble systems (SubX version (black) and FV3 version (red)). The RMS difference between analysis and climatological mean shows in green dash line (90 m). The climatological variance around mean has been drawn on the right

3.2 北半球(NH)中高纬度不同时空尺度的预报

500 hPa的位势高度场已被分解为行星尺度(1~3个纬向波),长波(4~9个纬向波)和天气波(10~20个纬向波)。图5给出这三组的北半球(NH)中高纬度500 hPa位势高度AC得分。由于FV3 GEFS的总体表现类似于SubX GEFS的预报系统(长达14 d,图4),因此图5仅显示SubX GEFS模式的AC分数,以便清楚地表达三组不同尺度的天气预报在当前的实际预报能力和潜在的预报能力。结果表明:1)不同尺度之间的预报能力差异很大;2)与大气长波(10.5 d)和天气尺度预报(7 d)相比,大气行星波具有更强的预报能力(12.5 d);3)与当前的实际预报技能相比,这三种时间空间尺度的预报均具有大约1~2.5 d的附加预报能力。行星波(1~3波)可能具有15 d以上(>2周)的潜在预报能力;长波(4~9波)将具有12 d的潜在预报能,而当下的实际预报能力仅为10.5 d;天气尺度系统

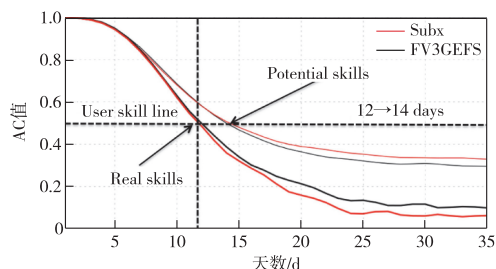


图4 一年试验时间(2017年4月至2018年4月),21个集合预报成员,预报35 d,集合预报平均值的北半球(NH)500 hPa距平相关(AC)。比较两个试验,分别是SubX版本(红线)和FV3版本(黑线),给出了目前的预报能力(粗线)和潜在的预报技巧(细线),并将其标记为12 d(真实)和14 d(潜在)

Fig.4 Northern Hemisphere (NH) 500 hPa geopotential height Pattern Anomaly Correlation (PAC) of the ensemble mean for 35 days forecast from one year experiments (April 2017—April 2018). There are two experiments which are SubX version (red) and FV3 version (black). The real forecast skills (thick) and potential prediction skills (thin) are presented, and marked as 12 days (real) and 14 days (potential)

(10~20波)也将由7 d的实际预报技能提升为8 d。因此,系统的预报能力和可预报性是高度依赖于预报对象,以及空间和时间尺度。

3.3 热带MJO评估

在热带地区,MJO是影响大气环流和能量转换的重要活动之一。特别是,中纬度的高/低空急流可能与MJO活动高度相关。对RMM(或MJO技能:RMM1+RMM2)的双变量距平相关性评估仅针对FV3 GEFS版本的一年试验时间(2017年4月至2018年4月),其MJO的预报技巧得分略高于SubX GEFS版本(未显示)。和图4、图5一样,图6展示了当前的MJO预报能力(黑色,约20 d,50% AC)和潜在预测能力(红色,约32.5 d,50% AC)。这样的结果可能表示:1)20 d的MJO预报能力客观地描述了当前集合预报系统和与其相关的模式系统,其预报能力符合客观实际,同中高纬度的其他评估是基本一致的;2)MJO的潜在预报能力或可预报性(32.5 d)远高于目前的实际预报水平(20 d);3)模式系统具有很大的改善空间,包括动力学过程的描述和设计、物理过程参数化方案、初始扰动和模式扰动以及许多其他因素(例如大气-海洋耦合)等以达到可预期的目标。

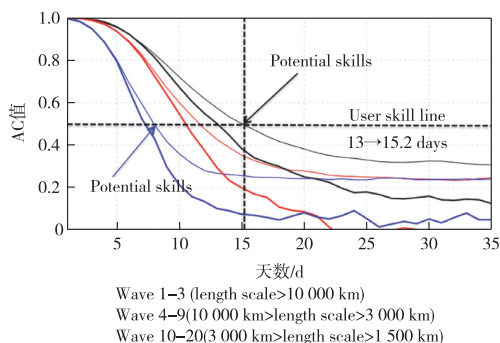


图5 一年试验时间(2017年4月—2018年4月),21个集合预报成员,预报35 d。预报场被分解为1~3个纬向波(黑色),4~9个纬向波(红色)和10~20个纬向波(蓝色),图中曲线显示了集合预报平均值的北半球(NH)500 hPa距平相关(AC)。两个实验分别是SubX版本(红线)和FV3版本(黑线)。显示了实际的预测能力(粗线)和潜在的预测技能(细线)

Fig.5 Northern Hemisphere (NH) 500 hPa geopotential height Pattern Anomaly Correlation (PAC) of re-grouped ensemble mean (zonal wave 1—3 (black), 4—9 (red) and 10—20 (blue)) for 35 days forecast from one year experiments (April 2017—April 2018). There are two experiments which are SubX version (red) and FV3 version (black). The real forecast skills (thick) and potential prediction skills (thin) are presented for each group

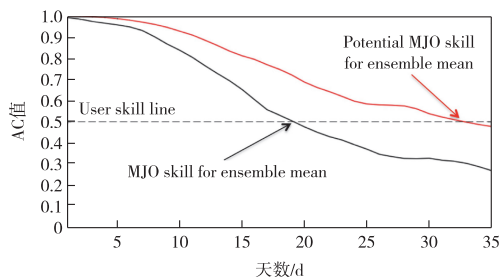


图6 与图3—5相同的一年试验,但仅仅考虑了FV3 GEFS版本。其真实的MJO预报能力(黑色,用分析场作为参考场)和潜在的MJO预报能力(红色,用扰动和不扰动的预报场作为参考)

Fig.6 The same 1-year experiment as figure 3—5, but for FV3 GEFS version only. A real MJO skill (black, i.e. analysis as reference) and potential MJO skill (red, i.e. control member as reference) of ensemble mean are presented

4 讨论与总结

使用NCEP GEFS评估北半球(NH)中高纬度和热带天气预报的潜在预报能力或可预报性,目前仍然只是一个初步尝试。北半球(NH)中高纬度500 hPa位势高度的AC评分表明,FV3 GEFS和SubX GEFS都具有相似的实际预报能力,但FV3 GEFS略有优势。与实际预报能力相比,潜在的预报能力或可预报性对于短期天气预报来说,两个GEFS模式是大体相当的,但SubX GEFS在延伸期(图4)的情况下具有更高的潜在预报能力(或更高的可预报性),这可能是SubX GEFS的离散度较小(离散度小于集合预报平均的均方根误差;或者说是一个不完美的系统)所造成的可预报能力偏高,或者解释为FV3 GEFS可能需要做更多的工作来改进和增强其潜在预报能力。对于FV3 GEFS的热带预报,MJO的实际预报能力和潜在预报能力(或可预报性)之间的巨大差异表明,NCEP FV3 GEFS系统具有很大的潜在能力(从20天到32.5天)。潜在的预报能力(或可预报性)的源头可能来自于对未来集合预报的不确定性的改进、更客观地描述大气环流异常、完善MJO相关的热带对流参数化方案和物理过程,以及考虑大气与海洋的耦合和相互作用等。

模式的实际预报能力和可预报性是和预报对象紧密相关的。总的来说,时空尺度较大的过程其预报能力也较强,相反,时空尺度较小的系统其预报能力也较弱。在未来的工作中,可以对可预报性做进一步探索并且扩展到其他关键的预报要素,例如降水、地表温度、地表风等;还有其他关键性预报事件,例如热带风暴,温带气旋,阻塞高压,冬季风暴以及其他一些高影响天气等。

致谢:感谢EMC集合预报团队成员提供的所有帮助和建设性讨论。特别是Dr. Yan Luo提供了图2, Mr. Eric Sinsky提供了图3、图4和图5, Dr. Wei Li提供了图6。这项研究得到了NWS OSTI和NOAA气候研究项目办公室(CPO)在模拟、分析、预报以及预测(MAP)计划层面的部分支持资助。其主要成果已在2018年10月召开的年度气候诊断和预测研讨会上提交。为了庆祝母校——南京气象学院(南京信息工程大学)建校(校)60周年,特将此项研究工作从英文转译成中文发表,并将在今后对此项研究工作进一步完善。

参考文献 (References)

- Bougeault P, 2010. The THORPEX interactive grand global ensemble [J]. *Bull Amer Meteor Soc*, 91: 1059-1072. doi: 10.1175/2010BAMS2853.1.
- Gottschalck J, Wheeler M, Weickmann K, et al., 2010. A framework for assessing operational madden-Julian oscillation forecasts [J]. *Bull Amer Meteor Soc*, 91(9): 1247-1258.
- Guan H, Zhu Y J, Sinsky E, et al., 2019. Systematic error analysis and calibration of 2-m temperature for the NCEP GEFS reforecast of the subseasonal experiment (SubX) project [J]. *Wea Forecasting*, 34(2): 361-376.
- He B, Liu P, Zhu Y J, et al., 2019. Prediction and predictability of Northern Hemisphere persistent maxima of 500 hPa geopotential height eddies in the GEFS [J]. *Clim Dyn*, 52(7/8): 3773-3789.
- Li W, Zhu Y J, Zhou X Q, et al., 2019. Evaluating the MJO prediction skill from different configurations of NCEP GEFS extended forecast [J]. *Clim Dyn*, 52(7/8): 4923-4936.
- Lin H, Brunet G, Derome J, 2008. Forecast skill of the madden-Julian oscillation in two Canadian atmospheric models [J]. *Mon Wea Rev*, 136(11): 4130-4149.
- Lorenz E N, 1969. Atmospheric predictability as revealed by naturally occurring analogues [J]. *J Atmos Sci*, 26(4): 636-646.
- Lorenz E N, 1982. Atmospheric predictability experiments with a large numerical model [J]. *Tellus*, 34(6): 505-513.
- Lorenz E N, 1996. Predictability: a problem partly solved [M] // Palmer T, Hagedorn R. *Predictability of Weather and Climate*. Cambridge: Cambridge University Press; 40-58.
- Pegion K, Kirtman B P, Becker E, et al., 2019. The subseasonal experiment (SubX): a multimodel subseasonal prediction experiment [J]. *Bull Amer Meteor Soc*, 100(10): 2043-2060.
- Shukla J, 1998. Predictability in the midst of chaos: a scientific basis for climate forecasting [J]. *Science*, 282(5389): 728-731.
- Swinbank R, Kyouda M, Buchanan P, et al., 2016. The TIGGE project and its achievements [J]. *Bull Amer Meteor Soc*, 97(1): 49-67.
- Toth Z, Kalnay E, 1993. Ensemble forecasting at NMC; the generation of perturbations [J]. *Bull Amer Meteor Soc*, 74(12): 2317-2330.
- Toth Z, Zhu Y J, Marchok T, 2001. The use of ensembles to identify forecasts with small and large uncertainty [J]. *Wea Forecasting*, 16(4): 463-477.
- Wei M Z, Toth Z, Wobus R, et al., 2008. Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system [J]. *Tellus* 59A: 62-79. doi: 10.3402/tellusa.v60i1.15267.
- Wheeler M C, Hendon H H, 2004. An all-season real-time multivariate MJO index; development of an index for monitoring and prediction [J]. *Mon Wea Rev*, 132(8): 1917-1932.
- Ying Y, Zhang F Q, 2017. Practical and intrinsic predictability of multiscale weather and convectively coupled equatorial waves during the active phase of an MJO [J]. *J Atmos Sci*, 74(11): 3771-3785.
- Zhou X Q, Zhu Y J, Hou D C, et al., 2016. A comparison of perturbations from an ensemble transform and an ensemble kalman filter for the NCEP global ensemble forecast system [J]. *Wea Forecasting*, 31(6): 2057-2074.
- Zhou X Q, Zhu Y J, Hou D C, et al., 2017. The NCEP global ensemble forecast system with the EnKF initialization [J]. *Wea Forecasting*, 32: 1989-2004.
- Zhou X Q, Zhu Y J, Fu B, et al., 2019. The development of next NCEP global ensemble forecast system [J]. *Climate Bulletin*: 159-163.
- Zhu Y J, 2005. Ensemble forecast: a new approach to uncertainty and predictability [J]. *Adv Atmos Sci*, 22(6): 781-788.
- Zhu Y J, Toth Z, Wobus R, et al., 2002. The economic value of ensemble-based weather forecasts [J]. *Bull Amer Meteor Soc*, 83(1): 73-83.
- Zhu Y J, Zhou X Q, Peña M, et al., 2017. Impact of sea surface temperature forcing on weeks 3 and 4 forecast skill in the NCEP global ensemble forecasting system [J]. *Wea Forecasting*, 32(6): 2159-2174.
- Zhu Y J, Zhou X Q, Li W, et al., 2018. An assessment of subseasonal forecast skill using an extended global ensemble forecast system (GEFS) [J]. *J Geophys. Res.*, 123: 6732-6745. doi: 10.1029/2018JD028506.
- Zhu Y J, Li W, Zhou X Q, et al., 2019a. Stochastic representation of NCEP GEFS to improve sub-seasonal forecast [M] // Springer Atmospheric Sciences. Singapore: Springer Singapore; 317-328.
- Zhu Y J, Li W, Sinsky E, et al., 2019b. An investigation of prediction and predictability of NCEP global ensemble forecast system (GEFS) [J]. *Climate Bulletin*: 154-158.

An assessment of predictability through state-of-the-art global ensemble forecast system

ZHU Yuejian

Environmental Modeling Center, National Oceanic and Atmospheric Administration (NOAA)/National Weather Service (NWS)/National Centers for Environmental Prediction (NCEP), College Park, Maryland, 20740, USA

In order to provide valuable and reliable numerical guidance include probabilistic forecast, state-of-the-art global ensemble forecast system has been implemented into National Centers for Environmental Prediction daily operation to service the general public. The forecast skills for various elements through statistical verification are offering the levels of confidence to general users when they apply these guidance. However, the skills are all different rely on the forecast elements, spatial and temporal resolutions, and special events or the forecast extremes. As an example of large scale pattern prediction, the global ensemble forecast system (global ensemble mean) could provide about 10 days skillful forecast on average based on 60% Northern Hemisphere 500 hPa geopotential height anomaly correlation. The predictability or limit of prediction we will discuss in this article could be a good reference for model developers, stakeholders and general users in terms of their requirements from various forecast elements to the different spatial-temporal scales. Especially, the investigation of sources predictability are very important to advance our numerical system in terms of science and technowledge. When we understand the source of predictability, the scientists will know where to work on, and how to improve the system. To combine the traditional predictability study and recently developed global ensemble forecast system with full representation of uncertainty, the predictability and limit of prediction through state-of-art global ensemble forecast system could be another valuable reference. As a summary of predictability investigation, there could be about 15-, 12-, 10-days forecast skills for planetary wave, large scale and synoptic scale patterns respectively. For tropical prediction, a skill of Madden-Julian Oscillation could be extended to 32.5 days if we could reduce model bias and improve others.

ensemble forecast; forecast uncertainty; predictability

doi:10.13878/j.cnki.dqkxxb.20191101013

(责任编辑:刘菲)