

2019-2021年北京春季温度预报的精细化检验评估

李妮娜 刘凑华 林建 代刊 韦青 赵声蓉
国家气象中心, 北京 100081

摘要

利用 2019-2021 年春季 (2-4 月) 北京地区国家级台站逐 3 h 观测资料, 基于统计分析方法对当前主流数值模式欧洲中心天气预报 (ECMWF)、中国气象局全球区域同化预报系统 (CMA-GFS) 以及智能网格预报国家级指导报 (SCMOC)、国省融合预报 (SMERGE) 的温度预报开展了精细化检验评估。结果表明, ECMWF 和 CMA-GFS 对北京春季的温度预报多出现负偏差, 且山区和平原地区无明显差异, 但在夜间更加突出。SCMOC 和 SMERGE 对温度预报有较好订正能力, 温度预报平均误差集中在 $-1\sim 1^{\circ}\text{C}$, 温度预报准确率较高且平均绝对误差也较小。四种产品对于 24 h 变温以及日内昼夜温差的预报存在问题: 所有产品对强变温的变化幅度预报不足, 且智能网格预报产品也未体现明显的订正能力; 此外, 所有产品昼夜温差幅度的预报存在 $1\sim 3^{\circ}\text{C}$ 的正偏差, SCMOC 有一定订正能力, 而 SMERGE 高估昼夜温差的情况更突出。模式中昼夜温差预报偏大与低温 (05 时温度) 预报偏低密切相关, 而智能网格预报中高温 (14 时) 预报偏高也不容忽视。通过精细化检验分析, 说明智能网格预报在关注整体准确率 (平均绝对误差) 提升 (减小) 的同时, 还应关注重要天气过程发展演变特征的预报改进。

关键词: 精细化检验, 偏差特征, 日变化, 过程演变
中图分类号: P456

Refined Evaluation of Spring Temperature Forecast in Beijing during 2019— 2021

LI Nina LIU Couhua LIN Jian DAI Kan WEI Qing Zhao Shengrong

National Meteorological Centre, Beijing 100081

Abstract: Using the 3-h national site observation data at Beijing area in spring (February to April) during 2019—2021, refined evaluation of temperature forecasts by current main-stream numerical models (European Centre for Medium-Range Weather Forecasting ECMWF and Global and Regional Assimilation and Prediction

国家重点研发计划 (2021YFC3000904) 资助

第一作者: 李妮娜, 主要从事精细化预报检验评估工作, E-mail: linn@cma.gov.cn

System CMA-GFS), national gridded guidance forecast product (SCMOC) and provincial and municipal revised feedback gridded forecast product (SMERGE) are conducted. The results show that, the temperature forecasts for spring in Beijing often show negative bias in the models of ECMWF and CMA-GFS. This bias has no significant difference between mountainous and plain areas, but is more prominent during nighttime periods. The gridded forecast products (SCMOC and SMERGE) have a good ability to correct the temperature forecasted by the models. The temperature forecast biases of the gridded products are concentrated between -1 to 1°C , and the forecast accuracy is higher and the mean absolute error is lower than that of model forecast. There are issues with the forecasts of 24-h temperature change and daily temperature range for the four products. The amplitude of strong 24-h temperature change forecasted by all products is relatively smaller than that of observation, and the gridded forecast products do not demonstrate significant correction ability. In addition, the daily temperature range forecasted by all products has a positive deviation of $1\sim 3^{\circ}\text{C}$ compared to the observation. SCMOC has better correction ability for the daily temperature range forecasted by the models, while SMERGE overestimates more prominently than that in models. The positive deviation of daily temperature range forecast is closely related to the underestimation of low temperature (at 05:00 BT) in the model, while the overestimation of high temperature (at 14:00 BT) can not be ignored in the grid forecast. It is suggested that gridded forecast products should not only focus on improving (reducing) overall accuracy (biases), but also on the improvement of the forecasts of development and evolution of important synoptic processes.

Key words: refined evaluation, bias characteristics, diurnal cycle, process evolution

华北地区位于我国第二、三级地形阶梯，西倚太行山，北靠燕山山脉，东临渤海。多变的地形及特殊的地理环境使得该地区的天气状况复杂。特别地，在全球变化背景下，极端降水、高温热浪、低温寒潮等灾害性天气频发，这类恶劣天气过程严重影响了社会生产和人民生活（Zhai et al, 2005; Zhang and Zhai, 2011; Liao et al, 2021; Chen et al, 2021）。对于发展程度高，人口密集是北京更是如此。气象灾害的防御亟需提高天气预报的准确性和及时性。为应对人们对气象服务更高的需求，在天气预报业务的现代化进程中，无缝隙精细化智能网格天气预报业务得以迅速发展（赵声蓉等, 2012; Hamill et al, 2015; 毕宝贵等, 2016; 代刊等, 2016; 金荣花等, 2019），即便如此，智能网格天气预报也很难满足极端性天气的预报需求。为了进一步提高极端灾害性天气的预报准确率和防灾减灾能力，通过精细化的检验评估来促进数值模式、智能网格预报技术的发展势在必行（毛冬艳等, 2014; 陈昊明等, 2021）。

早期的检验评估多聚焦于单模式或单变量（王雨, 2003; 管成功等, 2006, 2008），且多关注逐日或逐 12h 时间尺度的预报（韩洁等, 2015; 季晓东和漆梁波, 2018）。随着精细化预报产品的增加，为应对传统检验方法所带来的问题，出现了一系列空间检验方法（Davis et al, 2009; Gilleland et al, 2008; 刘凑华和牛若芸, 2013; 潘留杰等, 2014; 符娇兰和代刊, 2016; 赵滨和张博, 2018; 陈法敬等, 2019）。但是这些研究多注重空间上的对比，对于天气过程在发生、持续和结束等特征的关注较少（金荣花等, 2019）。因此，越来越多的研究开展了各产品间逐 3h 或逐 1h 的对比检验，以便了解数值模式或客观产品对日内尺度的精细化特征的预报能力（卢冰等, 2017; 许晨璐等, 2017; 陈龙等, 2024）。韦青等（2020）通过对国家级网格指导报、省市级订正反馈网格预报以及模式输出的降水和温度预报进行全面检验，发现智能网格预报可以改善午后降水峰值预报显著提前的偏差，且最高/最低温预报较数值模式改进明显，并指出智能网格预报需进一步改进强降水预报空报率较高、最低温预报模型构建等问题。Yuan et al (2020) 指出 ECMWF 预报的西南地区降水事件开始以及达到峰值的时间明显偏早且午后降水峰值被高估，这些偏差在西南地区西部更加突出。值得注意的是，这些精细化的检验多针对降水，针对温度的检验相对较少。

北京春季受西风槽频繁活动影响，地面冷锋活跃，天气复杂多变。气温变化快，昼夜温差大且大风天气频发，加之受北京周边复杂地形影响，给精细化预报带来了极大挑战。因此，本研究聚焦北京春季的温度预报，针对当前主流数值模式以及智能网格预报产品，开展精细化检验评估，分析不同产品的预报偏差特征、预报难点及业务参考价值。一方面为客观产品在业务预报中的合理应用提供参考，另一方面为智能网格预报产品后期研发与改进提供可靠依据。

1 检验数据和方法

实况数据为 2019-2021 年 2-4 月北京地区 20 个国家级台站的温度及风场观测资料。预报数据为欧洲中心天气预报中心 European Centre for Medium-Range Weather Forecasting（简称 ECMWF）、中国气象局全球区域同化预报系统 Global and Regional Assimilation and Prediction System（简称 CMA-GFS）、国家级指导报（简称 SCMOC）以及国省融合预报（简称 SMERGE）的 2019-2021 年 2-4 月逐日

08 时和 20 时起报的逐 3 h（预报时效在 108 小时后为逐 6h）温度预报产品。其中，全球数值模式 ECMWF 预报产品的空间分辨率为 12.5km，CMA-GFS 于 2016 年 6 月正式进入业务运行并向全国下发相关产品，其空间分辨率为 25km（沈学顺等，2017）。对于两套智能网格预报，国家级指导报 SCMOC 是基于 CMA-GFS 采用 GMOS 方法的精细化预报产品，由国家气象中心制作下发；国省融合预报 SMERGE 为省级业务单位在 SCMOC 基础上修改订正后上传得到；两者的空间分辨率均为 5km（金荣花等，2019；韦青等，2020）。

采用双线性插值方法将逐 3 h 温度预报产品插值到北京地区 20 个国家级台站后开展检验分析。

检验内容包括平均误差、平均绝对误差、预报准确率、分类误差频率统计以及日内尺度的变化特征等。

其中，温度预报准确率参照中国气象局中短期天气预报质量检验办法（中国气象局，2005）定义为 $R_c = \frac{N_{r2}}{N}$ ； N 为总站次， N_{r2} 为温度绝对误差 $\leq 2^\circ\text{C}$ 的站次。其实际含义为观测与预报温度之差的绝对值 $\leq 2^\circ\text{C}$ 的占比。

24 h 变温：利用逐 3h 温度数据，计算每个时刻 t 及 $t-24$ 时刻的温度差，获得逐 3h 滑动的 24h 变温，即 $\Delta T_{24t} = T_t - T_{t-24}$ ，其中 t 为研究时段内逐 3h 时刻，定义 $\Delta T_{24t} \geq 6^\circ\text{C}$ 为强升温时刻， $\Delta T_{24t} \leq -6^\circ\text{C}$ 为强降温时刻。检验针对 ΔT_{24} 开展。

昼夜温差：利用 14 时（全文均为北京时）和 05 时的温差表征昼夜温差，以检验不同产品对温度日内变化幅度的预报能力。

2 检验结果分析

2.1 温度预报基本性能检验

从 2019-2021 年 2-4 月全球模式 ECMWF、CMA-GFS 以及国家级指导报 SCMOC、国省融合预报 SMERGE 对北京地区的温度预报来看（图 1），大部分产品的温度预报准确率（绝对误差 $\leq 2^\circ\text{C}$ ）随预报时效增加而逐渐降低。CMA-GFS 温度预报准确率随预报时效变化较小，集中在 0.3~0.45，较其他预报产品的准确率偏低。ECMWF 模式预报准确率明显高于 CMA-GFS，8d 以上的预报准确率高于 0.35。国家级指导报 SCMOC 与 ECMWF 模式预报准确率相当，在短期预报时效（ $< 72\text{h}$ ）优于 ECMWF。SMERGE 订正作用明显，在绝大部分预报时效下准确率最高。平均绝对误差随着预报时效增加而逐渐增大。CMA-GFS 在各时效下绝对误差均最大，大部分预报时效下误差均超过 3°C ，10d 平均绝对误差达 3.40°C 。除 CMA-GFS 外，其余产品 5d 内的平均绝对误差小于 3°C ，5d 以上平均绝对误差不超过 4°C ；SMERGE 误差最小，10d 平均绝对误差为 2.45°C 。对于平均误差，不同时效下 ECMWF 和 CMA-GFS 模式均呈现温度预报偏低的特征。ECMWF（CMA-GFS）温度预报偏低 1°C （ 2°C ）左右，对于 SCMOC 和 SMERGE 两种预报产品来说，大部分预报时效下温度预报平均误差略低于 0°C 。

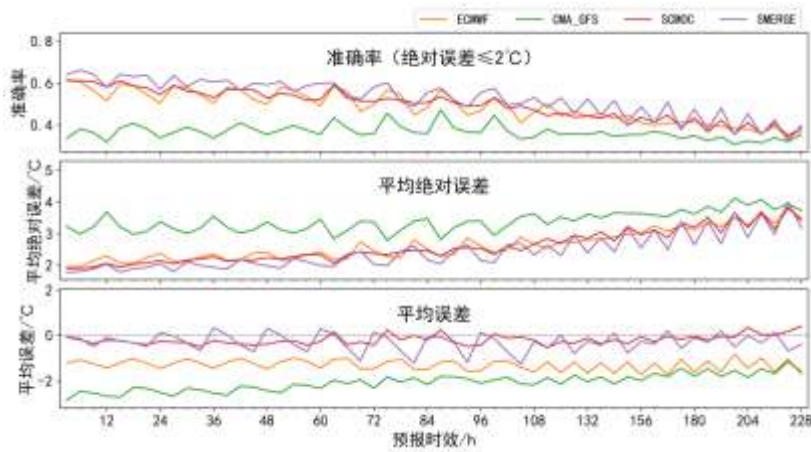
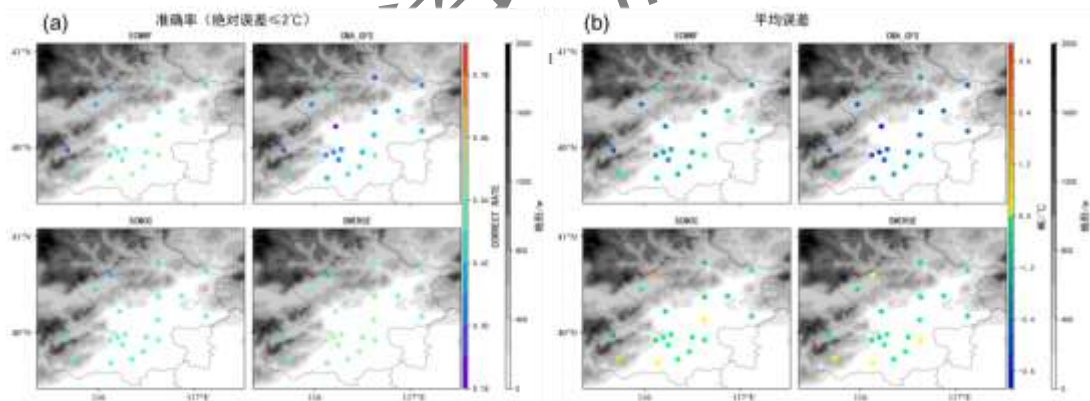


图 1 2019-2021 年 2-4 月北京地区不同预报产品温度预报 (a) 准确率 (绝对误差 $\leq 2^{\circ}\text{C}$), (b) 平均绝对误差和 (c) 平均误差随预报时效的变化

Fig.1 The variation of (a) accuracy (absolute error $\leq 2^{\circ}\text{C}$), (b) mean absolute error and (c) mean error of temperature forecast in Beijing with valid time from February to April during 2019—2021

图 2a 给出北京春季温度预报准确率的空间分布，整体来看，各预报产品均呈现出低海拔地区的预报准确率略高于高海拔地区的特征。从平均误差的空间分布来看（图 2b），四种预报产品对春季北京地区温度预报整体偏低。CMA-GFS 温度预报偏低尤为明显，中部地区/城区大多在 3°C 以上，ECMWF 温度预报平均误差集中在 $-2\sim-1^{\circ}\text{C}$ 。SCMOC 以及 SMERGE 对温度预报有明显订正能力，平均误差集中在 $-1\sim-1^{\circ}\text{C}$ 。



注：灰色阴影为地形高度。

图 2 2019-2021 年 2-4 月北京地区不同预报产品 (a) 温度预报准确率, (b) 温度预报平均误差空间分布

Fig.2 The spatial distribution of (a) accuracy, (b) mean error of temperature forecast in Beijing from February to April during 2019—2021.

为了解北京山区（站点海拔 $\geq 200\text{ m}$ ）和平原地区（站点海拔 $< 200\text{ m}$ ）昼夜温度预报特征差异，图 3a,3b 分别给出两类地形区不同时段温度预报的平均误差。ECMWF 和 CMA-GFS 模式预报的温度存在系统性偏差，无论是山区还是平原地区，

两个模式的温度预报平均误差都呈现为负偏差，这一特征在白天（08-17时）和夜间（20-05时）均存在。值得注意的是，夜间时段温度预报的平均误差幅度大于白天时段。智能网格预报 SCMOC 和 SMERGE 平均误差幅度明显低于模式预报。夜间时段平均误差由模式的-2~-3℃减小至-0.5℃左右；白天时段的预报平均误差表现为较小的正偏差。分析不同大小的误差出现频率（图 3c），发现 ECMWF、CMA-GFS 温度预报误差为负的出现频率明显偏高，分别为 67.35%和 73.29%。SCMOC 和 SMERGE 的温度预报误差较小，-1~1℃频率最高，但负偏差出现频率仍略高于正偏差。考虑到温度预报以负偏差为主，图 3d 进一步分析了各预报产品负偏差出现频率的日变化特征。对于模式预报 ECMWF 与 CMA-GFS，它们在夜间 23 时的负偏差频率最高。SCMOC 以及 SMERGE 负偏差出现频率的峰值时刻较模式预报偏晚，分别出现在凌晨 02 时和清晨 05 时。

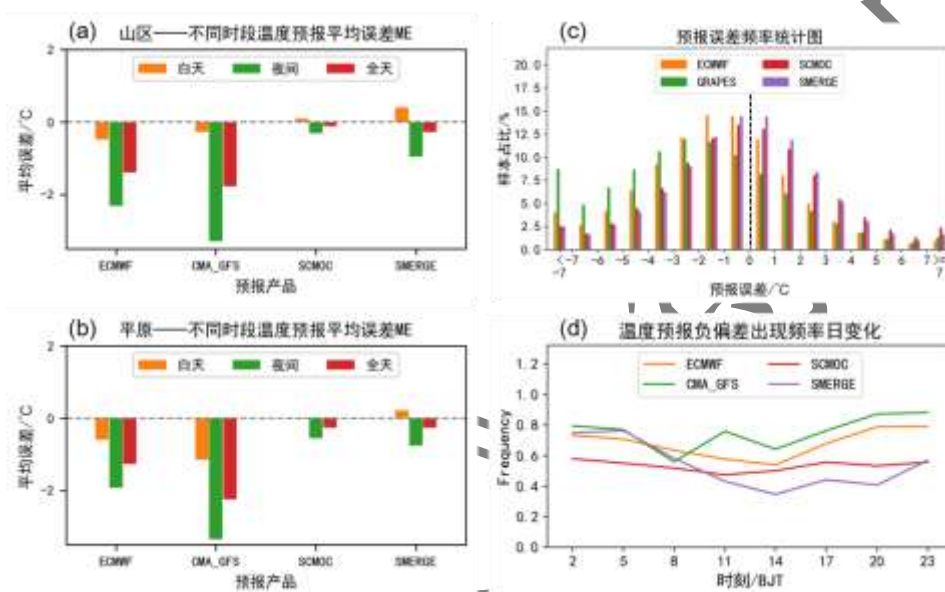


图 3 2019-2021 年 2-4 月北京 (a) 山区和 (b) 平原地区各预报产品不同时段温度预报平均误差, (c) 温度预报误差频率统计和 (d) 温度预报负偏差出现频率日变化

Fig.3 The mean error of temperature forecast at different time periods in (a) mountainous and (b) plain areas in Beijing, and (c) the frequency statistic at different error bins (units: %), and (d) the diurnal cycle of occurrence frequency of negative deviation of temperature forecast from February to April during 2019—2021.

总体来说，CMA-GFS 模式对北京温度预报效果较差，但 SCMOC 产品预报效果明显提升，与 ECMWF 模式效果相当，甚至在短期时效优于 ECMWF，SMERGE 产品对北京春季温度预报效果最优。这四种预报产品均倾向于低估北京地区的春季温度，以 CMA-GFS 模式预报的温度偏低最为明显。北京的山区和平原地区均存在温度预报的负偏差，这种偏差在夜间时段更大，且夜间时段出现负偏差的频率更高。

2.2 24 h变温

考虑到北京春季经常受西风槽频繁活动影响，转折性天气过程较多，气温升

降幅大，24 h 变温是温度预报关注的重点问题之一，这部分将聚焦 24 h 变温开展检验评估。

基于逐 3h 滑动的 24 h 变温进行统计，图 4 给出观测与预报中不同变温幅度的出现频率以及北京不同地区强升温及强降温的频率。从图 4a 可以看到，观测中北京春季出现升温的频率 (58.21%) 高于降温频率 (41.79%)，其中以升温 0~2℃ 的频率最高，可达 22.49%。但关注强变温过程可以发现，北京地区强降温频率 (7.88%) 略高于强升温频率 (6.21%)，且山区发生强变温的频率高于平原地区 (图 4b,4c)。四种产品对变温幅度 <2℃ 的过程预报频率偏高，而对强变温过程的预报频率明显不足。这种特征在国省融合产品 SMERGE 中表现尤为明显，强变温预报频率甚至不足观测的 1/2。尽管从预报准确率、整体平均误差来看 SMERGE 温度预报较好，但 SMERGE 并没有很好地把握升降温过程。

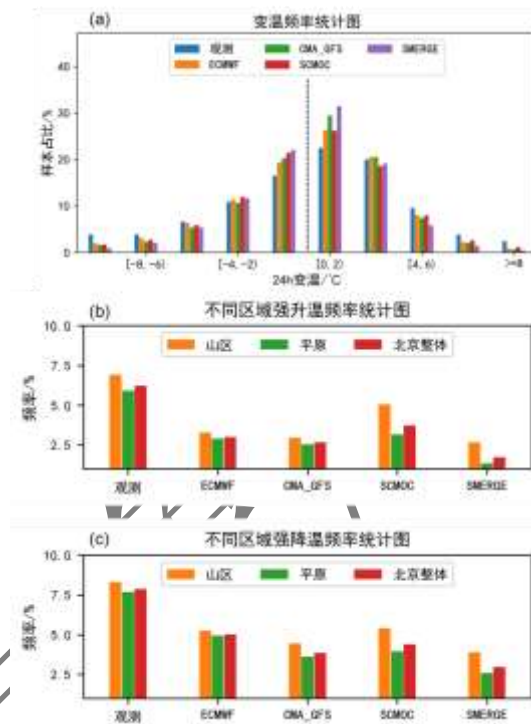


图 4 2019-2021 年 2-4 月北京地区观测和各预报产品 (a) 不同变温幅度的频率，及不同区域 (b, c) 强变温的发生频率

(b) 强升温 ($\Delta T_{24h} \geq 6^{\circ}\text{C}$), (c) 强降温 ($\Delta T_{24h} \leq -6^{\circ}\text{C}$)

Fig.4 The (a) frequency statistic at different 24-h temperature change range, and the occurrence frequency of the strong temperature (b) rising ($\Delta T_{24h} \geq 6^{\circ}\text{C}$) and (c) cooling ($\Delta T_{24h} \leq -6^{\circ}\text{C}$) from February to April during 2019 to 2021.

聚焦强变温过程，表 1 给出发生强变温的个例情况，正如前文所提到的，强降温的发生频次高于强升温频次。基于挑选的 2132/2642 个强升温/强降温站次，分析强变温发生前及发生时的预报特征。图 5 给出强变温过程发生前一天及发生时不同温度预报误差的频率分布。对于强升温过程 (图 5a, 5b)，升温前 ECMWF 与 CMA-GFS 温度预报的正、负偏差基本相当，SCMOC 和 SMERGE 温度预报偏高的频率略高于温度预报偏低的频率 (图 5a)。而升温日，温度预报出现了明显的

负偏差。四种产品温度预报出现负偏差的频率均在 80%以上，CMA-GFS 中预报误差超过-7℃的频率高达 35%，其余产品温度预报误差在该区间的频率也超过了 15%（图 5b）。对于强降温过程（图 5c, 5d），降温发生前温度预报多呈现负偏差（图 5c），四种产品预报出现负偏差的频率在 78.56~92.89%。降温发生时，ECMWF 和 CMA-GFS 温度预报正、负偏差出现的频率较为接近，SCMOC 及 SMERGE 倾向于高估温度；四种产品中温度预报偏高超过 7℃的频率均在 5%以上。

综上所述，当出现强变温过程时，四种产品的变温幅度预报不足；对于强升温过程来说，主要是升温日温度预报偏低造成升温幅度预报不足，对于强降温过程来说，降温前温度预报偏低、降温日温度预报偏高共同影响了降温幅度的预报。

表 1 强变温主要发生日期、发生日数、时次数以及总站次数

Table 1 The main occurrence date, number of days and the frequency of strong 24-h temperature change, as well as the total number of times at every stations occurring strong 24-h temperature change

	主要发生时段	主要发生日数 (单位: d)	主要发生时次数 (单位: h)	总发生站次数 (单位: 站 *h)
强升温 ($\Delta T_{24h} \geq 6^\circ\text{C}$)	2019 年 3 月 9 日、3 月 15 日、3 月 19 日、3 月 24-25 日、3 月 31 日、4 月 4 日、4 月 11 日、4 月 18 日、4 月 21 日、4 月 23 日、4 月 30 日	33	84	2132
	2020 年 2 月 8 日、2 月 13 日、3 月 1 日、3 月 9 日、3 月 12 日、3 月 14-15 日、3 月 17 日、3 月 20 日、4 月 4 日、4 月 11-12 日、4 月 24 日、4 月 29 日			
	2021 年 2 月 6 日、2 月 19-21 日、2 月 28 日、3 月 13 日、3 月 22 日、3 月 30 日			
强降温 ($\Delta T_{24h} \leq -6^\circ\text{C}$)	2019 年 3 月 6-7 日、3 月 20-21 日、3 月 28 日、4 月 5-6 日、4 月 9-10 日、4 月 19 日、4 月 24-25 日、4 月 27 日	41	119	2642
	2020 年 2 月 14-15 日、3 月 8 日、3 月 13 日、3 月 16 日、3 月 19 日、3 月 21-22 日、3 月 24 日、3 月 26-27 日、4 月 2 日、4 月 5 日、4 月 13 日、4 月 25-26 日			
	2021 年 2 月 7-8 日、2 月 15-17 日、2 月 22-23 日、3 月 6 日、3 月 12 日、3 月 26 日、4 月 1 日、4 月 4 日			

*表中主要发生时段、日数和时次的统计是基于某一时次发生强变温站数超过 50%（10 站）得到的

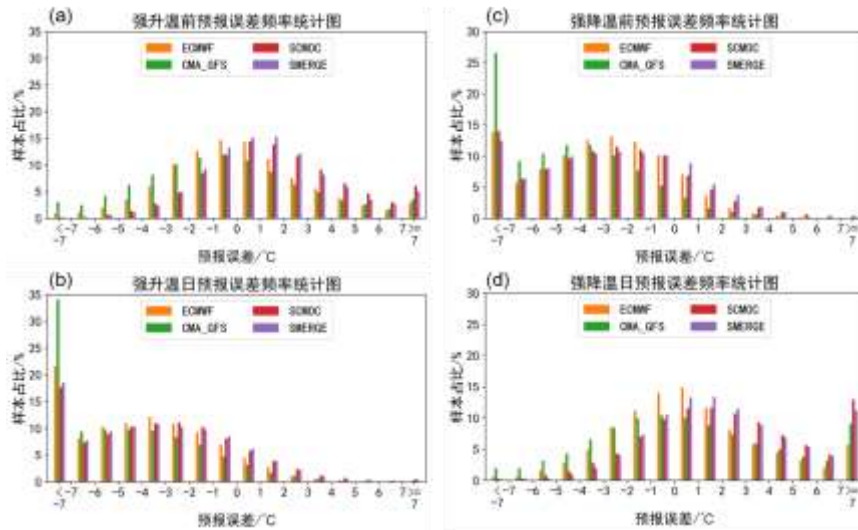
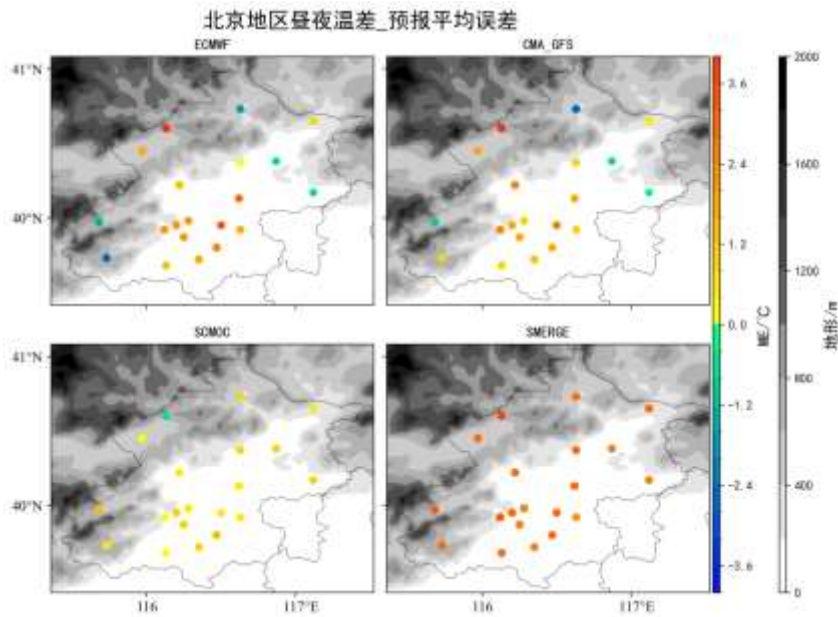


图 5 2019-2021 年 2-4 月北京地区不同预报产品在强升温 (a, c) 发生前及 (b, d) 发生日的温度预报误差频率分布 (a, b) 强升温 ($\Delta T_{24t} \geq 6^{\circ}\text{C}$), (c, d) 强降温 ($\Delta T_{24t} \leq -6^{\circ}\text{C}$)

Fig.5 The frequency statistic at different temperature forecast error bins in Beijing from February to April during 2019 to 2021 for (a) one-day before and (b) on the day of occurrence of strong temperature rising ($\Delta T_{24t} \geq 6^{\circ}\text{C}$), and that for (c) one-day before and (d) on the occurrence of strong temperature cooling ($\Delta T_{24t} \leq -6^{\circ}\text{C}$)

2.3 昼夜温差

北京春季温度另一突出特征是昼夜温差较大。本文利用 14 时（高温）和 05 时（低温）的温度差考察不同产品对逐 3h 定时温度日内变化幅度的预报情况。从预报平均误差图中可以看出四种预报产品均高估了北京春季的昼夜温差（图 6）。ECMWF 和 CMA-GFS 中，大部分站点的昼夜温差预报偏高 2°C 左右，CMA-GFS 的预报平均误差略小于 ECMWF。SCMOC 对于昼夜温差的预报有明显订正能力，其预报平均误差减小至 1°C 以内，仅为 0.70°C 。但对于 SMERGE 来说，北京地区昼夜温差预报平均误差甚至高于模式预报结果，最大昼夜温差的平均误差可达 3.85°C ，最小也超过 2°C 。区分 08 和 20 时起报的预报发现，在两种不同起报时间下昼夜温差的预报平均误差基本一致（图略），表明昼夜温差幅度预报偏大属于系统性误差。



注：灰色阴影为地形高度。

图 6 2019-2021 年 2-4 月北京地区不同预报产品昼夜温差预报的平均误差空间分布

Fig.6 The spatial distribution of mean error of daily temperature range forecast in Beijing from February to April during 2019—2021.

不同昼夜温差出现频率的统计结果显示北京地区春季昼夜温差集中在 $6\sim 14^{\circ}\text{C}$ (出现频率 48.83%), $8\sim 10^{\circ}\text{C}$ 出现频率最高 (达 12.76%) (图 7)。对于昼夜温差小于 8°C 的情况, 四种预报均明显低估其出现频次, SMERGE 对昼夜温差在 6°C 以内的频率预报甚至不足观测的 1/2。对于昼夜温差较大的情况, 四种预报产品均会高估其频率。另外, 预报中最频繁出现的昼夜温差区间也较观测偏大: ECMWF (CMA-GFS) 预报的昼夜温差更多集中在 $12\sim 14$ ($10\sim 12$) $^{\circ}\text{C}$, SMERGE 预报的昼夜温差主要集中在 $12\sim 16^{\circ}\text{C}$, 仅 SCMOC 能够做出较为合理的预报, 其预报的昼夜温差主要集中在 $8\sim 12^{\circ}\text{C}$, 最接近观测结果。

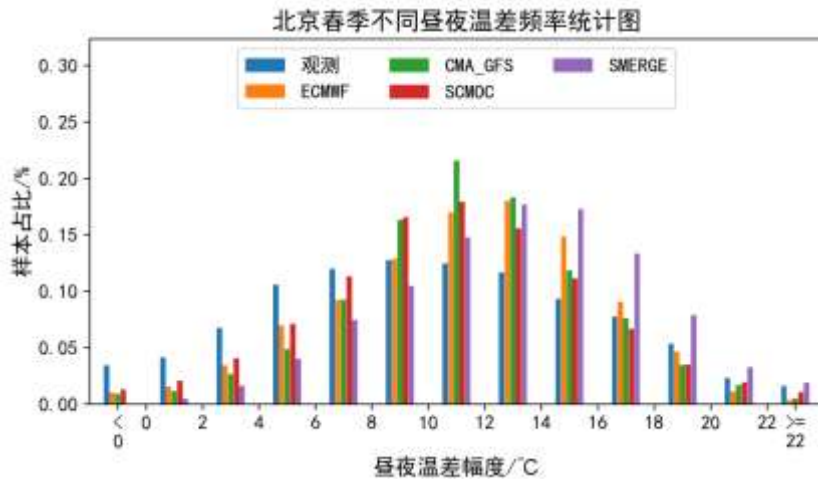


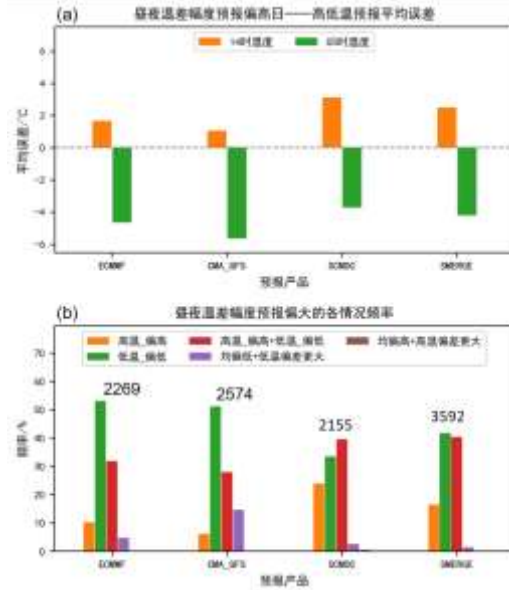
图 7 2019-2021 年 2-4 月北京地区观测和各预报产品不同昼夜温差的出现频率

Fig.7 The frequency statistic at different daily temperature ranges in Beijing from February to April during 2019—2021.

进一步分析昼夜温差预报偏大的情况，挑选各产品中预报误差超过 4°C 的日期考察该日 14 时和 05 时温度预报平均误差。如图 8 所示，昼夜温差幅度预报误差较大时，模式产品的 05 时温度预报误差幅度明显大于 14 时温度预报误差，低温预报误差是高温预报误差幅度的 3 倍甚至更高；对于智能网格预报产品 SCMOC 和 SMERGE 来说，高低温预报误差幅度基本相当。综合高低温的预报特征，将昼夜温差预报呈现正偏差的情况分为了 5 种类型，分别是：（1）低温预报较好（误差在 $\pm 2^{\circ}\text{C}$ 之间），但高温预报偏高（误差 $> 2^{\circ}\text{C}$ ）；（2）高温预报较好，低温预报偏低（误差 $< -2^{\circ}\text{C}$ ）；（3）高温预报偏高且低温预报偏低；（4）高、低温预报均偏低（误差均 $< -2^{\circ}\text{C}$ ）但低温预报偏低幅度更大以及（5）高、低温预报均偏高（误差均 $> 2^{\circ}\text{C}$ ）但高温预报偏高幅度更大。

图 8b 给出各类型预报误差的出现频率以及不同产品昼夜温差预报过大的站次。从图中可以看出四种产品昼夜温差预报偏大的站次存在一定差异。其中，CMA-GFS 昼夜温差预报偏大的站次数高于 ECMWF 预报。SCMOC 对昼夜温差的预报略有改善，正偏差较大的站次数模式预报结果均有所减少。对应较高的昼夜温差预报平均误差（图 6），SMERGE 产品中昼夜温差偏大的站次数也明显高于其余三种产品。关注不同类型偏差出现情况可以发现，模式与智能网格预报昼夜温差偏大的情况有所区别。对于模式产品 ECMWF 和 CMA-GFS 而言，昼夜温差幅度预报偏大很大程度上是由于低温预报偏低（绿色柱）造成的，这种情况出现频率分别为 53.02% 和 51.09%。第二类主要的偏差为高温预报偏高与低温预报偏低，分别为 31.86% 和 27.89%。值得注意的是，相较于智能网格预报产品，高低温预报均偏低但低温偏低幅度更大（紫色柱）这种情况的出现频率也较高。整体来说，低温预报偏低是模式产品昼夜温差幅度预报偏大的最主要原因。

对于两套智能网格预报产品 SCMOC 和 SMERGE 来说，低温预报偏低、高低温预报均偏低的情况明显减少。昼夜温差幅度预报偏大，在很大程度上是由于高温预报偏高和低温预报偏低（红色柱）共同导致的。此外，仅高温预报偏高（橙色柱）的频率明显增加；对应于图 8a，智能网格预报产品中低温预报偏低幅度略小于模式产品，但高温预报正偏差明显增大。结合图 2、图 3 分析可以推断，智能网格预报产品在订正温度时倾向于系统性增加温度来减小温度预报误差，但聚焦昼夜温差的预报时，就会过多地出现高温预报明显偏高的情况。后续需要结合模式预报偏差的主要问题，针对性做出温度预报的订正，使其在整体温度预报准确率提升的同时，也能对温度的日变化演变特征做出更精准预报。



注：图 b 柱状图上方数值表示昼夜温差幅度预报偏高 4°C 以上的站次数

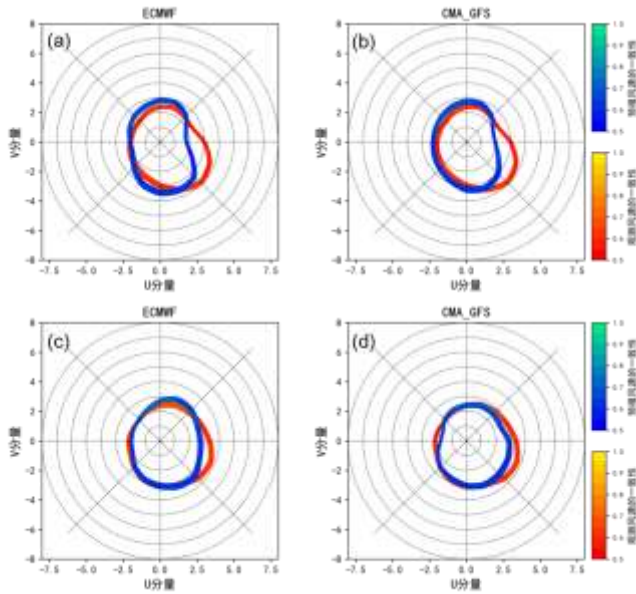
图 8 2019-2021 年 2-4 月北京地区不同预报产品 (a) 昼夜温差幅度预报偏大日的高低温的预报平均误差, (b) 昼夜温差幅度预报偏大的 5 种情况概率分布)

Fig.8 (a) The bias of high and low temperature on days with large positive biases of forecast daily temperature range in Beijing from February to April during 2019 to 2021, (b) The frequency statistics of five cases with large positive biases of forecast daily temperature range for the period in Beijing from February to April during 2019 to 2021

2.4 预报误差的可能原因

整体来看, 北京春季温度预报难点为: (1) 变温过程中预报变温幅度偏小; (2) 温度的日内变化 (昼夜温差) 幅度预报偏大。针对上述两方面的预报难点, 进一步探讨其可能原因。

通过分析研究时段内北京地区所有出现强变温日的风场预报情况 (图 9) 可以发现, 对于强降温过程来说, 无论是 ECMWF 还是 CMA-GFS, 其预报的西北-偏西方向的风速仅为 2~3 m/s, 明显小于观测风速 (3~4 m/s)。北京地区通常是在以西北气流为主的冷空气影响下发生降温, 而模式预报的西北-偏西风风速偏弱, 说明北京地区的降温幅度预报不足可能与冷空气的平流输送不足有关。对于强升温过程, 模式预报的风场在各方向的风速、出现频率基本相当, 与观测的风场特征较为相近, 说明北京地区的升温幅度预报不足可能更多与辐射过程相关。



注：红色、蓝色分别表示观测、预报风速的一致性。

图 9 2019-2021 年 2-4 月观测和不同预报产品北京地区 (a, b) 发生强降温日 ($\Delta T_{24t} \leq -6^{\circ}\text{C}$) 以及 (c, d) 发生强升温日 ($\Delta T_{24t} \geq 6^{\circ}\text{C}$) 风矢量分布

(a, c) 为 ECMWF, (b, d) 为 CMA-GFS,

Fig.10 The observation and forecast wind vector distribution on the days of occurrence of strong temperature (a-b) cooling ($\Delta T_{24t} \leq -6^{\circ}\text{C}$) and (c-d) rising ($\Delta T_{24t} \geq 6^{\circ}\text{C}$) in Beijing from February to April during 2019—2021

(a, c) ECMWF, (b, d) CMA-GFS

考虑到云辐射强迫对气温有重要影响, 针对昼夜温差幅度预报偏大这一问题, 进一步分析研究时段内所有误差偏大日, 14 时和 05 时的云量预报平均误差。从图 10 可以看出, 除 ECMWF 预报的白天云量较观测略偏高外, 模式预报的云量通常较观测偏少, 平均误差约 -10%, 甚至超过 -20%。白天的少云会带来更强的辐射增温, 而夜间云量的偏少会使向外长波辐射增强, 日内的降温幅度更加明显。同时, 图 10 也显示了夜间的云量偏少较白天更加突出, 这与前文提到的模式产品低温预报偏低是最主要误差原因相对应。也就是说, 昼夜温差幅度预报偏高与当日云量的预报误差具有一定关联。以 2021 年 3 月 1-5 日过程为例, 通过检验北京站当时总云量的预报可以发现, 观测总云量可以达到 90% 左右, 而对应的模式预报结果显示基本无云 (总云量约 0) (图略), 造成这段时间白天 (夜间) 温度预报偏高 (偏低)。正如秦庆昌等 (2022) 指出, 此次春季温度预报失误的原因主要与观测中持续性低云天气过程有关, 模式无法再现这类低云特征。

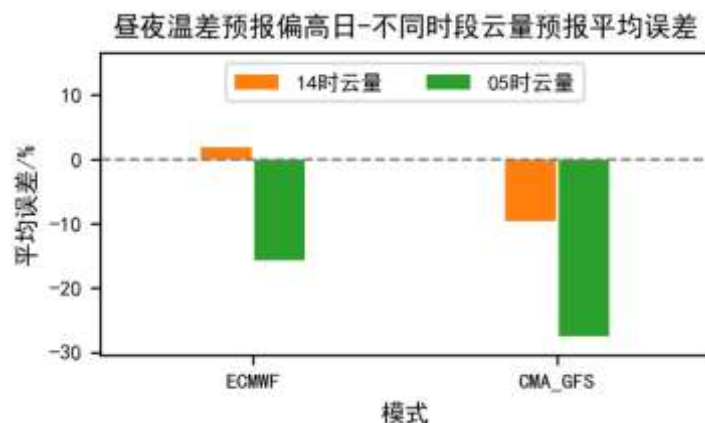


图 10 2019-2021 年北京地区昼夜温差幅度预报偏高日 14 时和 05 时云量预报平均误差

Fig.10 The mean error of cloud cover forecast at 14 and 05 BJT on days with large positive biases of forecast daily temperature range in Beijing from February to April during 2019—2021

3 结论和讨论

聚焦北京春季的温度，通过开展精细化检验评估，分析了当前主流预报模式 ECMWF、CMA-GFS 以及智能网格预报产品 SCMOC 和 SMERGE 的预报能力以及预报偏差特征，主要结论如下：

(1) CMA-GFS 模式对北京地区春季温度预报准确率低，平均绝对误差大；但基于 CMA-GFS 采用 GMOS 方法的 SCMOC 产品预报效果明显提升，与 ECMWF 模式效果相当；SMERGE 产品对北京春季温度预报准确率最高，平均绝对误差最小。所有预报产品均倾向于低估北京地区的春季温度。温度预报偏低这一问题在山区和平原地区无明显差异，但在夜间时段更加突出。智能网格预报 SCMOC 和 SMERGE 对温度预报的负偏差具有较好的订正能力，温度预报平均误差集中在 $-1\sim 1^{\circ}\text{C}$ 。

(2) 北京春季转折性天气较多，24 h 变温剧烈。当前四种产品对转折性天气预报能力较差，强变温过程中 24 h 变温幅度预报偏小； 6°C 及以上的强升温过程中幅度预报不足主要是受升温日温度预报偏低影响， -6°C 及以下强降温过程幅度预报不足是降温前温度预报偏低和降温日温度预报偏高两者共同影响；强升温（强降温）幅度预报不足可能与辐射过程（平流输送）预报偏差相关。

(3) 北京春季日内昼夜温差较大，以 $8\sim 10^{\circ}\text{C}$ 的温差最常出现。模式预报产品各均过高估计了温度的日内变化幅度，昼夜温差预报偏高 $1\sim 3^{\circ}\text{C}$ 。SCMOC 对昼夜温差幅度高估的问题有所改善，但 SMERGE 对昼夜温差的预报甚至不如模式预报产品。对模式产品来说，低温预报偏低是昼夜温差幅度预报偏大最主要的原因，这可能与云量的预报偏差相关；智能网格预报产品中，高温预报偏高且低温预报偏低的出现频率有所增加，且仅高温预报偏高的情况也不容忽视。

本文通过细致分类，从统计检验和个例分析的角度全面给出了温度预报的偏差特征，并通过建立 24 h 变温、昼夜温差幅度这两个具有物理意义的评估指标

进行检验,指出所有的预报产品对北京春季温度具有较好预报能力,但面对重要天气过程时,依旧存在预报误差大、预报难度高的情况。针对这些问题,所有预报产品应该在关注整体预报准确率提升的同时,重点聚焦天气过程演变特征,做出精准预报。另外,本文对偏差的成因进行了初步的探讨,发现强降温幅度预报偏低可能与模式预报的西北气流强度偏弱有关。但影响强降温的环流型众多,除了主导的西北气流型外,还可能存在偏北气流型、偏西气流型以及西南气流型等环流形势的影响(杜晓辉等,2016;杨晓玲等,2016)。未来将拓展研究时段,对影响强变温的环流型进行分类,探究不同类型下的预报误差特征,明晰偏差出现时的主导环流系统以及关键影响因子。在此基础上开展针对性客观预报技术研究,完善当前技术体系中的预报因子和建模样本方案、偏差订正方案、预报融合集成方案等,减小针对转折性、高影响天气预报误差偏大的问题,为最终提升预报服务能力提供支撑。

参考文献

- 毕宝贵,代刊,王毅,等,2016. 定量降水预报技术进展[J]. 应用气象学报, 27(5): 534-549. Bi B G, Dai K, Wang Y, et al, 2016. Advances in techniques of quantitative precipitation forecast[J]. J Appl Meteor Sci, 27(5): 534-549 (in Chinese).
- 陈法敬,陈静,韦青,等,2019. 一种基于可预报性的暴雨预报评分新方法 II: 暴雨检验评分模型及评估试验[J]. 气象学报, 77(1): 28-42. Chen F J, Chen J, Wei Q, et al. 2019. A new verification method for heavy rainfall forecast based on predictability II: verification method and test[J]. Acta Meteor Sin, 77(1): 28-42 (in Chinese).
- 陈昊明,李普曦,赵妍,2021. 千米尺度模式降水的检验评估进展及展望[J]. 气象科技进展, 11(3): 155-164. Chen H M, Li P X, Zhao Y, 2021. A review and outlook of verification and evaluation of precipitation forecast at convection-permitting resolution[J]. Adv Meteor Sci Technol, 11(3): 155-164 (in Chinese).
- 陈龙,陈静静,胡媚,等,2024. 华南快速循环同化模式在湖南不同环流型下的小时降水预报性能检验[J]. 气象, 50(4): 434-448. Chen L, Chen J J, Hu M, et al, 2024. Evaluation of hourly quantitative precipitation forecast of the rapid updating cycle assimilation and forecasting system in south china based on circulation classification in Hunan Province[J]. Meteor Mon, 50(4): 434-448 (in Chinese).
- 代刊,曹勇,钱奇峰,等,2016. 中短期数字化天气预报技术现状及趋势[J]. 气象, 42(12): 1445-1455. Dai K, Cao Y, Qian Q F, et al, 2016. Situation and tendency of operational technologies in short- and medium-range weather forecast[J]. Meteor Mon, 42(12): 1445-1455 (in Chinese).
- 杜晓辉,代华光,德庆,2016. 西藏强降温时空分布特征与天气概念模型[J]. 西藏科技, (8): 62-64. Du X H, Dai H G, De Q, 2016. Spatial and temporal distribution characteristics and synoptic conceptual type of strong cooling in Tibet[J]. Tibet Sci Technol, (8): 62-64 (in Chinese).
- 符娇兰,代刊,2016. 基于 CRA 空间检验技术的西南地区东部强降水 EC 模式预报误差分析

- [J]. 气象, 42(12): 1456-1464. Fu J L, Dai K, 2016. The ECMWF model precipitation systematic error in the east of Southwest China based on the contiguous rain area method for spatial forecast verification[J]. Meteor Mon, 42(12): 1456-1464 (in Chinese).
- 管成功, 王克敏, 陈晓红, 2006. 2002—2005 年 T213 数值降水预报产品分析检验[J]. 气象, 32(8): 70-76. Guan C G, Wang K M, Chen X H, 2006. Analyses and Verifications of T213 rainfall forecast from 2002 to 2005[J]. Meteor Mon, 32(8): 70-76 (in Chinese).
- 管成功, 陈起英, 佟华, 等, 2008. T639L60 全球中期预报系统预报试验和性能评估[J]. 气象, 34(6): 11-16. Guan C G, Chen Q Y, Tong H, et al, 2008. Experiments and evaluations of global medium range forecast system of T639L60[J]. Meteor Mon, 34(6): 11-16 (in Chinese).
- 韩洁, 庞翻, 陈卫东, 2015. WRF 模式对宝鸡地区 2013 年夏季温度和降水预报的评估检验[J]. 陕西气象, (1): 1-6. Han J, Pang F, Chen W D, 2015. Evaluation and verification of summer temperature and precipitation forecast in Baoji, Shaanxi Province in 2013 by WRF model[J]. J Shaanxi Meteor, (1): 1-6 (in Chinese).
- 胡萍, 冉仙果, 徐大红, 等, 2021. 铜仁市寒潮天气分型及成因分析[J]. 中低纬山地气象, 45(3): 40-47. Hu P, Ran X G, Xu D H, et al, 2021. Type and cause analysis of cold wave weather in Tongren City[J]. Mid-low Latitude Mountain Meteor, 45(3): 40-47 (in Chinese).
本条文献在正文中未被引用季晓东, 漆梁波, 2018. ECMWF 模式降水预报与极端天气预报指数在暴雨预报中的评估与应用[J]. 暴雨灾害, 37(6): 566-573. Ji X D, Qi L B, 2018. Evaluation and application of ECMWF model precipitation and extreme weather forecast index of precipitation on heavy rainfall forecast[J]. Torr Rain Dis, 37(6): 566-573 (in Chinese).
- 金荣花, 代刊, 赵瑞霞, 等, 2019. 我国无缝隙精细化网格天气预报技术进展与挑战[J]. 气象, 45(4): 445-457. Jin Q H, Dai K, Zhao R X, et al, 2019. Progress and challenge of seamless fine gridded weather forecasting technology in China[J]. Meteor Mon, 45(4): 445-457 (in Chinese).
- 刘凑华, 牛若芸, 2013. 基于目标的降水检验方法及应用[J]. 气象, 39(6): 681-690. Liu C H, Niu R Y, 2013. Object-based precipitation verification method and its application[J]. Meteor Mon, 2013, 39(6): 681-690 (in Chinese).
- 卢冰, 孙继松, 仲跻芹, 等, 2017. 区域数值预报系统在北京地区的降水日变化预报偏差特征及成因分析[J]. 气象学报, 75(2): 248-259. Lu B, Sun J S, Zhong J Q, et al, 2017. Analysis of characteristic bias in diurnal precipitation variation forecasts and possible reasons in a regional forecast system over Beijing area[J]. Acta Meteor Sin, 75(2): 248-259 (in Chinese).
- 毛冬艳, 朱文剑, 樊利强, 等, 2014. GRAPES_MESO V3.3 模式强天气预报性能的初步检验[J]. 气象, 40(12): 1429-1438. Mao D Y, Zhu W J, Fan L Q, et al, 2014. Preliminary synoptic verification on the performance of GRAPES_MESO V3.3 in severe weather forecasting[J]. Meteor Mon, 40(12): 1429-1438 (in Chinese).
- 潘留杰, 张宏芳, 王建鹏, 2014. 数值天气预报检验方法研究进展[J]. 地球科学进展, 29(3): 327-335. Pan L J, Zhang H F, Wang J P, 2014. Progress on verification methods of numerical weather prediction[J]. Adv Earth Sci, 29(3): 327-335 (in Chinese).

- 秦庆昌, 于波, 徐路扬, 等, 2022. 造成北京春季气温预报失误的持续性低云过程分析[J]. 气象, 48(6): 705-718. Qin Q C, Yu B, Xu L Y, et al, 2022. Analysis of a continuous low-cloud process causing spring temperature forecast errors in Beijing[J]. Meteor Mon, 48(6): 705-718 (in Chinese).
- 沈学顺, 苏勇, 胡江林, 等. 2017. GRAPES_GFS 全球中期预报系统的研发和业务化[J]. 应用气象学报, 28(1): 1-10. Shen X S, Su Y, Hu J L, et al, 2017. Development and operation transformation of GRAPES global middle-range forecast system[J]. J Appl Meteor Sci, 28(1): 1-10 (in Chinese).
- 王雨, 2003. 2002年主汛期国家气象中心主客观降水预报对比检验[J]. 气象, 29(5): 21-25, doi: 10.3969/j.issn.1000-0526.2003.05.005. Wang Y, 2003. Verification of NMC subjective and objective precipitation prediction during the main flood season in 2002[J]. Meteor Mon, 29(5): 21-25, doi: 10.3969/j.issn.1000-0526.2003.05.005 (in Chinese).
- 韦青, 代刊, 林建, 等, 2020. 2016—2018年全国智能网格降水及温度预报检验评估[J]. 气象, 46(10): 1272-1285. Wei Q, Dai K, Lin J, et al, 2020. Evaluation on the 2016-2018 Fine gridded precipitation and temperature forecasting[J]. Meteor Mon, 46(10): 1272-1285 (in Chinese).
- 许晨璐, 王建捷, 黄丽萍, 2017. 千米尺度分辨率下 GRAPES-Meso4.0 模式定量降水预报性能评估[J]. 气象学报, 75(6): 851-876. Xu C L, Wang J J, Huang L P, 2017. Evaluation on QPF of GRAPES-Meso4.0 model at convection-permitting resolution[J]. Acta Meteor Sin, 75(6): 851-876 (in Chinese).
- 杨晓玲, 丁文魁, 马中华, 等, 2016. 河西走廊东部强降温变化特征和典型环流型[J]. 气象, 42(6): 756-763. Yang X L, Ding W K, Ma Z H, et al, 2016. Change characteristic and typical circulation type of strong cooling in Eastern Hexi Corridor[J]. Meteor Mon, 42(6): 756-763 (in Chinese).
- 赵滨, 张博, 2018. 邻域空间检验方法在降水评估中的应用[J]. 暴雨灾害, 37(1): 1-7. Zhao B, Zhang B, 2018. Application of neighborhood spatial verification method on precipitation evaluation[J]. Torr Rain Dis, 37(1): 1-7 (in Chinese).
- 赵声蓉, 赵翠光, 赵瑞霞, 等, 2012. 我国精细化客观气象要素预报进展[J]. 气象科技进展, 2(5): 12-21. Zhao S R, Zhao C G, Zhao R X, et al, 2012. The Development of objective meteorological element forecast in China[J]. Adv Meteor Sci Technol, 2(5): 12-21 (in Chinese).
- 中国气象局, 2005. 中短期天气预报质量检验办法(试行)[R]. China Meteorological Administration, 2005. Verification methods for short and medium range weather forecast (trial)[R]. (in Chinese).
- Chen Y, Liao Z, Shi Y, et al, 2021. Detectable increases in sequential flood-heatwave events across China during 1961-2018[J]. Geophys Res Lett, 48(6): e2021GL092549.
- Davis C A, Brown B G, Bullock R, et al, 2009. The Method for Object-Based Diagnostic Evaluation (MODE) applied to numerical forecasts from the 2005 NSSL/SPC spring program[J]. Wea Forecasting, 24(5): 1252-1267.

-
- Gilleland E, Lee T C M, Gotway J H, et al, 2008. Computationally efficient spatial forecast verification using Baddeley's Delta image metric[J]. *Mon Wea Rev*, 136(5): 1747-1757.
- Hamill T M, Scheuerer M, Bates G T, 2015. Analog probabilistic precipitation forecasts using GEFS reforecasts and climatology-calibrated precipitation analyses[J]. *Mon Wea Rev*, 143(8): 3300-3309.
- Liao Z, Chen Y, Li W, et al, 2021. Growing threats from unprecedented sequential flood - hot extremes across China[J]. *Geophys Res Lett*, 48(18): e2021GL094505.
- Yuan W H, Hu X L, Li Y S, 2020. Evaluation of the hourly rainfall in the ECMWF forecasting over southwestern China[J]. *Meteor Appl*, 27(4): e1936.
- Zhai P M, Zhang X B, Wan H, et al, 2005. Trends in total precipitation and frequency of daily precipitation extremes over China[J]. *J Clim*, 18(7): 1096-1108, doi: 10.1175/JCLI-3318.1.
- Zhang H, Zhai P M, 2011. Temporal and spatial characteristics of extreme hourly precipitation over eastern China in the warm season[J]. *Adv Atmos Sci*, 28(5): 1177-1183, doi: 10.1007/s00376-011-0020-0.