

OCF 与集成成员气温预报的大误差过程对比检验

唐延婧¹ 慕建利² 袁彬² 廖波¹ 裴兴云¹ 杜正静¹

¹ 贵州省气象服务中心

² 中国气象局公共气象服务中心

提要: 为评价用于公众服务的多模式集成气温预报效果, 分析其误差成因, 以中国区域多模式集成 (OCF) 的日最高气温和日最低气温预报检验为切入点, 对服务影响较大的大误差日及其典型过程—降温天气开展检验分析, 并与集成成员 ECMWF 和 NCEP 气温高度订正预报进行对比。结果表明: OCF 气温预报总体上预报性能优于 ECMWF 和 NCEP 等集成成员, 准确率夏季高冬季低, 拉开了气温变化范围, 也有效减小了误差。OCF 气温预报大误差日较少, 但 2~3d 时效及冬半年的大误差日较 ECMWF 多, 与集成成员、降温天气相关。针对降温天气的检验分析发现: OCF、ECMWF 和 NCEP 在降温天气的预报性能有所下降, 尤其是 OCF 日最高气温预报误差增长快; OCF 对降温天气里的日最低气温、非降温区域的日最高气温进行了有效订正, 但其日最高气温误差在降温日的降温区域里预报偏高特征明显。基于过程检验结果, 进行了误差原因分析, 提出了改进方向, 说明针对过程的检验更利于发现客观预报及集成订正的问题。

关键词: 集成气温预报, 检验, 大误差, 降温过程

Temperature Forecast Verification on Consensus Model and Consensus members Focus on Big Error Processes

TANG Yanjing¹ Mu Jianli² YUAN Bing² LIAO Bo¹ PEI Xingyun¹ DU Zhengjing¹

¹ Meteorological Service Center of Guizhou Province, Guiyang 550002

² Public Meteorological Service Center, China Meteorological Administration, Beijing 100081

Abstract: In order to evaluate consensus and correction model base on objective numerical forecast, which was applied on public service, this paper focused on big error days and temperature-drop days, to make objective verification on OCF consensus model and main collection members' daily maximum and minimum temperature forecast. Result showed: OCF consensus model performed better than collection members overall, the accuracy was higher in summer and lower in winter. And OCF enlarged daily temperature range and effectively reduced error. OCF has less big error days than collection members, but showed larger errors in 2-3d forecasting periods, and winter half year. The big error days were related to collection members and obvious temperature-drop. It was found that the forecast performance of OCF, ECMWF and NCEP declined in temperature-drop days, especially the error of OCF daily maximum temperature forecast increases rapidly. In temperature-drop days, OCF has effectively corrected daily minimum temperature and daily maximum temperature in non-drop areas, but the daily maximum temperature error was obvious in temperature-drop area. Finally, process verification was conducive to discover the defects of objective forecasts and consensus, correction methods, therefore comes to the improve directions.

Key words: Consensus forecast of temperature, verification, big error, temperature-drop days

*国家重点研发计划 (2018YFC1507802) 资助

第一作者: 唐延婧, 主要从事专业气象服务研究, E-mail: wangwu@mail.com

通讯作者: 慕建利, 主要从事数值预报释用和旅游气象服务等研究, E-mail: mujl668@sina.com

引言

气温对人们的日常生活影响无时不在，因而备受关注，是公众气象预报服务中最为重要的预报要素之一。近二十年来，数值模式产品在时空精细化程度、预报准确性等方面均取得了巨大进展，但迄今为止没有一个数值模式能够完美地呈现近地面的气象要素，包括气温的变化特征（Hoffman et al, 2017）。因此，数值模式预报产品往往需要进行再加工处理。由此衍生出了多种数值预报释用技术，在气温的客观预报研究中应用最为广泛。这些方法包括：模式输出统计方法（Glahn and Lowry, 1972; 刘还珠等, 2004）、直接模式输出方法（Stensrud and Skindlov, 1996）、神经网络方法（赵声蓉, 2006; 张亚刚等, 2021）、多模式误差订正和集成（Woodcock and Engel, 2005; 李佰平和智协飞, 2012; 郝翠等, 2019; 盛春岩等, 2020），以及机器学习法（郭虎等, 2008; 任萍等, 2020）等，使得气温预报效果得到显著提升。

在上述国内外精细化预报和客观模式订正方法研究中，对气象要素预报效果的检验是必不可少的环节。客观而全面的检验能反映预报产品的性能优劣，也能暴露释用技术的问题，进而推动问题的解决和技术的进步。2005年中国气象局下发了中短期天气预报质量检验办法，制定了一套气象预报产品质量检验评分标准，国家气象中心建立了“国家级天气预报检验分析系统”（韦青等, 2019），对智能网格预报产品及其他相关模式预报产品和客观预报产品开展常规业务检验评估（韦青等, 2020）。平均绝对误差（MAE）、均方根误差（RMSE）等传统的检验指标被广泛用于气温等连续变量的检验评估（Murphy and Winkler, 1987; Murphy, 1996）。但对实际预报业务和集成订正预报而言，常规检验标准虽然能够定量描述气温预报整体的误差和准确率，但针对性的误差分析、过程分析等仍然非常必要。

中国气象局公共服务中心“精细化多模式集成预报服务产品”（以下简称 OCF），一直广泛用于公众气象服务。多年来，OCF 的多模式集成气温预报产品的平均效果已有常规检验佐证，但从集成订正方法改进的角度，以及从公众气象服务的关注点出发，都有对 OCF 气温预报产品进行针对性检验的需求。本文重点针对服务影响较大的大误差日，以及 OCF 气温预报的典型大误差过程——降温过程开展检验分析，以期发现气温模式预报和集成与订正的问题与规律，为 OCF 集成订正技术及其他客观数值预报释用方法评价提供参考。

1 多模式集成预报简介

国内外气象部门基于数值预报释用技术建立的客观预报业务系统，如美国的国家数字预报数据库（Glahn and Ruth, 2003）、德国基于优化集成预报方法提供的订正网格预报、澳大利亚的业务集成预报（Woodcock and Engel, 2005; Engel and Ebert, 2012）、奥地利气象局的无缝隙概率预报系统（Kann et al, 2018），中国气象局的智能网格预报（金荣花等, 2019），等等。中国气象局公共气象服务中心于 2012 年建立了多模式精细化预报服务产品（OCF）制作系统，并投入了实时业务化运行。OCF 的业务运用不仅提高了预报准确率，而且为当今公众气象服务主流的基于位置的精细化天气预报服务提供了基础的技术支撑。下面简单介绍 OCF 的 2m 气温集成预报模型技术。

OCF 气温集成预报模型中，用到包括 ECMWF（时间分辨率为 3h，空间分辨率为 0.125° ）、NCEP（时间分辨率为 6h，空间分辨率为 0.5° ）、JMA（时间分辨率为 3h，空间分辨率为 0.125° ）、CMA_GFS（原 GRAPES 全球模式，时间分辨率为 3h，空间分辨率为 0.25° ）等数值预报模式 2m 气温预报产品，以及地面 2m 气温实时观测数据。OCF 在建模过程中，采取对最近 30 天的集成成员进行逐站点、逐时次滚动误差订正和动态变权重的方法建立预报模型，使集成预报结果随最近时段内各集成成员预报效果的优劣而变化。

OCF 集成预报模型具体步骤如图 1，以 20 时（全文均为北京时）起报的未来 7d 逐 3h 的集成预报产品为例，具体步骤如下。

①模式数据：读取各模式前 30 天 08 时起报气温预报场，解析未来 15~180h 时效内的逐 3h 的预报作为预报因子；

②模式格点至站点的插值:应用双线性插值方法将各模式气温预报插值至所需的站点上,在插值时考虑模式高度与站点实际地形高度差,对气温进行高度误差订正;

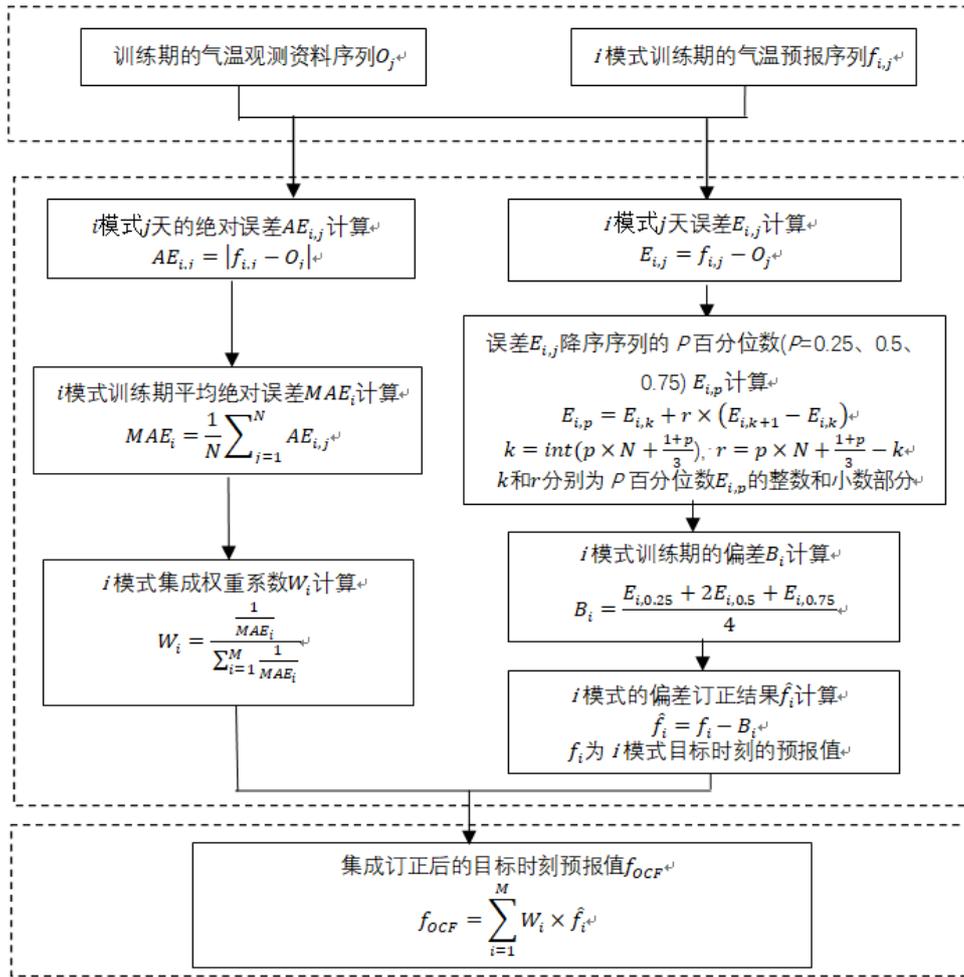
③模式误差计算:分别对各模式过去30天的预报时效的逐站点逐3h的预报与相对应时刻的实况观测数据相减,计算评估模式预报误差;

④模式预报偏差计算:对③计算的预报误差排序,并应用百分位四分法计算过去30天不同预报时次的预报偏差,算法具体参考 Engel and Ebert(2007)、Hyndman and Fan(1996);

⑤模式预报偏差订正后的预报:将通过②得到各模式不同预报时次的站点预报减去④对应时次的预报偏差,得出各模式不同时次的误差订正后的预报值;

⑥模式权重系数计算:通过过去30天各模式绝对预报误差倒数与集成成员绝对误差的倒数之比,确定多模式集成权重系数;

⑦集成预报计算:各模式逐站点逐时次的权重系数与对应模式预报偏差订正后的预报相乘,得到最后的集成预报值。



M 为参与集成模式数, i 为参与集成模式 ($i \in M$); N 为误差统计日数 (训练期) $N=30d$, j 为误差统计的某天 ($j \in N$)

图1. OCF 气温集成预报模型数据和计算流程图

Fig.1 Data and calculation flowchart of OCF consensus model's temperature forecast

2 检验方案

2.1 检验数据

本文检验 2020 年 6 月 1 日至 2021 年 5 月 31 日（其中去除缺少各模式预报数据的 12 天）OCF 实时业务 20 时起报的 24~168h 逐 24h（1~7d 时效）集成预报、OCF 集成成员相应时段的日最高气温和日最低气温的高度订正预报。需要说明的是，本文所用数据取自 OCF 实际业务，CMA_GFS 从 2021 年开始进入业务中，日本的 JMA 模式数据到 2021 年 3 月后停止提供，两个模式稳定进入 OCF 的为 1~3d 时效的预报数据，因此检验时段内 JMA 和 CMA_GFS 不完整，不开展非常规检验。检验站点选取中国区域 2401 个国家级观测站，考虑预报和实况的更替情况及数据完整率、一致性等情况，最终得到实际检验站点 1993 个，以下所有检验如不做特别说明均包含中国区域所有 1993 个站点。

2.2 检验量和检验方法

2.2.1 常规检验

公众接收到的气象预报信息以及评价，都是围绕所在地点的小范围开展的，OCF 集成订正技术方法及输出产品也都是对应观测站点的，因此采用点对点的常规检验进行日气温预报性能的评价。本文选取平均误差（ E ，公式见图 1）、均方根误差 RMSE（公式详见郝翠等，2019）、气温预报准确率（绝对误差 $AE \leq 2^\circ\text{C}$ 的站次占总站次的比率， AE 公式见图 1）作为常规检验指标。

2.2.2 大误差检验

本文从公众服务的效果和公众敏感度考虑，重点分析气温预报中出现大范围大误差的过程。大误差及检验相关过程的标准定义如下：

(1) 大误差：根据公众气象服务经验，当绝对误差 $AE > 4^\circ\text{C}$ 时，公众会产生气温预报不准确的感觉，因此将 $AE > 4^\circ\text{C}$ 称为大误差；将区域内 $AE > 4^\circ\text{C}$ 的站点占比率称为大误差率 ER ；将中国区域单日的大误差率 $ER \geq 20\%$ 定义为大误差日。

$$ER = N_{AE>4}/N$$

式中： $N_{AE>4}$ 为 $AE > 4^\circ\text{C}$ 的站点数， N 为中国区域内的检验站点总数（本文为 1993 站）。

(2) 变温日：为探讨大误差日与气温变化的关系，将实况的日最高气温和日最低温的平均值 \bar{T} 求取日变温 $\Delta\bar{T}$ （以下简称高低温平均日变化）作为气温变化的判断指标， $\Delta\bar{T} \geq 4^\circ\text{C}$ 为明显升温， $\Delta\bar{T} \leq -4^\circ\text{C}$ 为明显降温，两者统称明显变温；单日 $\Delta\bar{T} \geq 4^\circ\text{C}$ 明显升温站点比率 $\geq 20\%$ 作为升温日，单日 $\Delta\bar{T} \leq -4^\circ\text{C}$ 明显降温站点比率 $\geq 20\%$ 升温日，以上两者统称变温日。

$$\bar{T} = (T_{max} + T_{min}) / 2, \Delta\bar{T} = \bar{T}_t - \bar{T}_{t-1}$$

式中： \bar{T} 为日最高气温和日最低温的平均值， T_{max} 为日最高气温， T_{min} 为日最低气温， j 代表某日， \bar{T}_{t-1} 为前 1 日的日最高气温和日最低温的平均值。

3 检验结果分析

3.1 OCF 与集成成员模式预报性能对比分析

首先针对中国区域整个检验时段，统计 OCF 集成预报与 2 个集成成员 ECMWF、NCEP 各预报时效的日最高、日最低气温的均方根误差（RMSE）、预报准确率，对比分析各预报在中国区域的整体预报性能。

从 1~7d 各预报时效的日最高气温、日最低气温均方根误差和预报准确率（图 2）来看，OCF 预报效果最好、ECMWF 次之、NCEP 最差。其中，OCF 日最高气温均方根误差为 $2.0\sim 3.2^\circ\text{C}$ ，较 ECMWF 小 $0.2\sim 0.8^\circ\text{C}$ 、较 NCEP 小 $1.1\sim 1.5^\circ\text{C}$ ；OCF 日最低气温均方根误差为 $2.0\sim 2.6^\circ\text{C}$ ，较 ECMWF 小 $0.2\sim 0.3^\circ\text{C}$ 、较 NCEP 小 $1.1\sim 1.3^\circ\text{C}$ 。OCF 日最高气温预报准确率分

别为 49%~74%，较 ECMWF 高 1%~10%，较 NCEP 高 4%~15%；OCF 日最低气温预报准确率为 60%~75%，较 ECMWF 高 2%~5%，较 NCEP 高 10%~13%。

OCF、ECMWF 和 NCEP 的日最高气温和日最低气温均方根误差均随着预报时效的增长而增加。日最高气温预报的均方根误差明显大于日最低气温的，且日最高气温预报的均方根误差随时效增长的幅度较日最低气温明显。

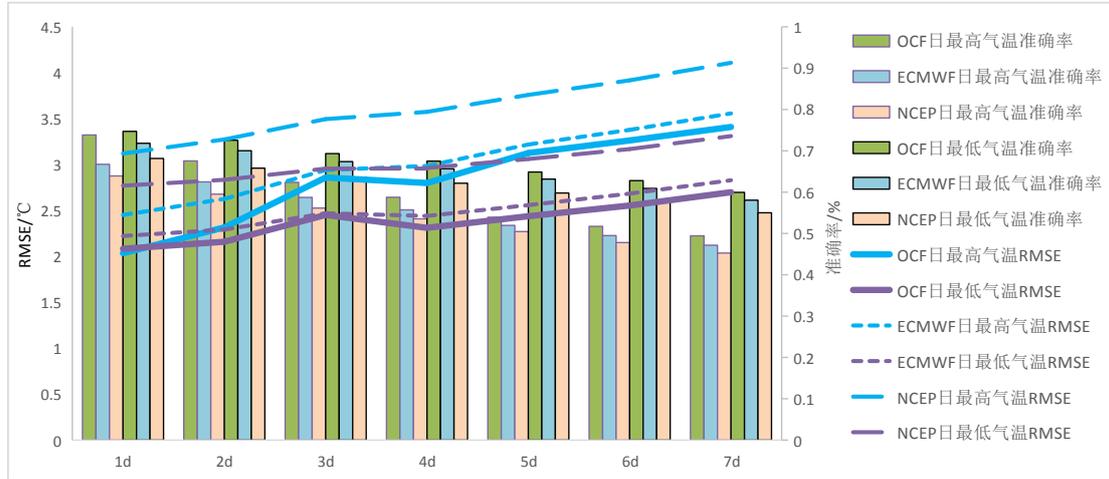


图2. 2020年6月至2021年5月OCF、ECMWF和NCEP在不同预报时效日最高气温、最低气温预报的RMSE和准确率

Fig.2 Daily maximum and minimum temperature's RMSE and accuracy of OCF, ECMWF, NCEP from June 2020 to May 2021

统计中国区域 OCF 和各集成成员 1~3d 短期时段（图 3）和 4~7d 中期时段（图略）各月的准确率及平均误差，其中 JMA 和 CMA_GFS 仅参与 1~3d 短期时段检验。OCF、ECMWF、NCEP 短期时段和中期时段的准确率、平均误差月分布类似，只是短期时段准确率较中期时段高 5%~15% 左右，误差也是短期时段较中期时段低。

分月来看，各集成成员短期时段预报准确率（图 3a）各有优劣变化，ECMWF、NCEP 均表现为夏季高冬季低的变化趋势，OCF 也有同样的季节变化趋势，但始终优于各集成成员。OCF 日最低气温、日最高气温短期时段各月平均预报准确率分别为 72%、68%，中期时段准确率分别为 64%、53%，仍是各预报中最高的；ECMWF 次优，日最低气温、日最高气温短期时段各月平均预报准确率分别为 68%、58%，中期时段准确率分别为 61%、49%。

平均误差可以说明模式系统预报偏高（平均误差为正，简称正误差）或偏低（平均误差为负，简称负误差）的平均状况。从图 3b 可以看出，ECMWF 各月日最高、最低气温都呈负误差，这种一致性的误差特征说明其气温预报偏低更多是系统性误差，存在进一步订正的空间；NCEP 冬季时段预报的日最高、最低气温都呈负误差，1 月日最高气温的平均误差甚至达到 -4℃ 以上，春、秋季检验时段的最低气温多为正误差、最高气温负误差相对较大，夏季最低气温和最高气温误差很小，说明其气温预报误差既有模式系统误差的影响，也有季节变化的影响；JMA 和 CMA_GFS 在其各自检验时段内也有最低气温正误差、最高气温负误差的特征。而 OCF 大部分月份的平均误差较集成成员明显减小，且呈现最低气温负误差、最高气温正误差，说明加入实况后的多模式集成订正 OCF 气温预报效果较好，拉开了气温变化的范围，也有效减小了误差。

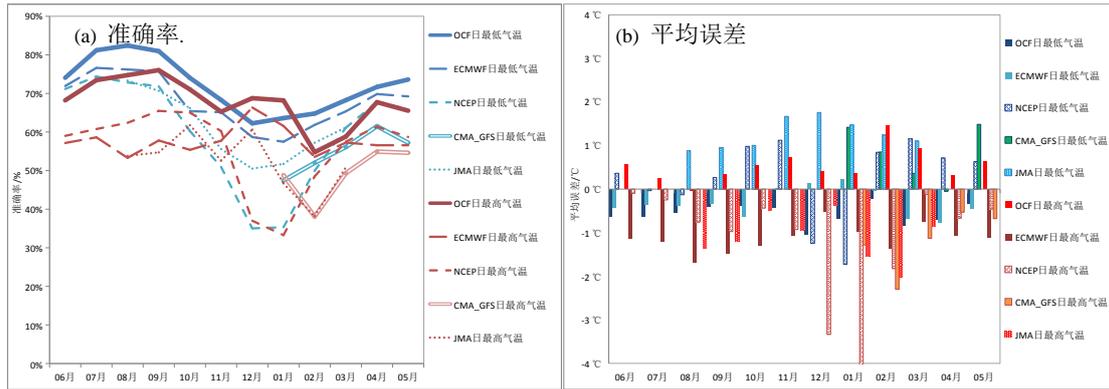


图3. 2020年6月至2021年5月各预报场1~3d短期时段日最高、最低气温的逐月 (a) 准确率, (b) 平均误差

Fig.3 Daily maximum and minimum temperature's (a) accuracy and (b) average error in each month of each short term forecasting feilds from June 2020 to May 2021

3.2 集成预报大误差分析

3.2.1 大误差日分布特征

由于 JMA 和 CMA_GFS 数据不完整, 以下检验及分析仅针对 OCF、ECMWF 和 NCEP 展开。根据前文的大误差日定义, 统计检验时段内 1~7d 时效各预报场的大误差日总数, OCF 为 608d, ECMWF 为 753d, NCEP 为 1017d, OCF 大误差日数总体少于 2 个集成成员。OCF、ECMWF 和 NCEP 的大误差日都随预报时效增长而增多(图 4a), NCEP 的大误差日数明显高于其他两者。OCF 和 ECMWF 各有优劣: 1d 时效的大误差日, 两者都小于 10d, OCF 略少; 2~3d 时效, OCF 大误差日数较 ECMWF 多 1/3 左右; 4~7d 时效 OCF 的大误差日较 ECMWF 少 1/3 左右。这说明, 随着预报时效的延长, OCF 订正集成有正效应, 尤其在 6~7d 时效订正效果更加明显; 2~3d 时效 OCF 大误差日多于 ECMWF, 可能与 OCF 集成算法有关; 另外由于 1~3d 时效, JMA 和 CMA_GFS 参与了 OCF 集成, 而图 3 可见 JMA 和 CMA_GFS 的预报性能不如 ECMWF, 这可能也是导致 OCF 在 2~3d 时效大误差日较 ECMWF 多的原因之一。

进一步分析大误差日的各月分布(图 4b)发现, 9—10 月 OCF、ECMWF 和 NCEP 的大误差日都比较少, 均在 40d 以下, 2—3 月、5—6 月则均超过了 50d, NCEP 冬季 12—1 月的预报缺陷非常明显。值得注意的是, 5—10 月期间, OCF 的大误差日数, 大多少于 ECMWF 和 NCEP, 尤其 7—10 月期间, 较两个集成成员少了 70% 以上, 这体现了其优越的集成订正效果; 但 11—3 月, OCF 大误差日多于 ECMWF, 11—12 月则较 ECMWF 偏多 30 d 左右。

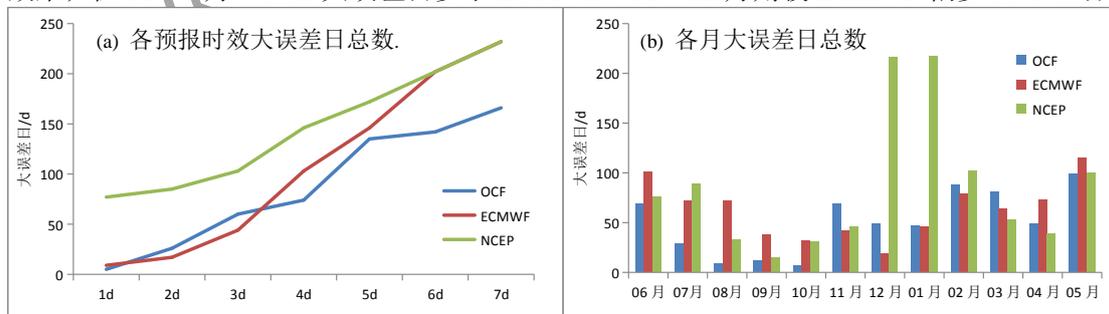


图4. 2020年6月至2021年5月 OCF、ECMWF 和 NCEP 气温预报

(a) 各预报时效大误差日总数, (b) 各月大误差日总数

Fig. 4 Big error days of OCF, ECMWF, NCEP's temperature forecasting feilds from June 2020 to May 2021 (a) in each forecasting periods, (b) monthly distribution

总之，OCF 大误差日总数少于 ECMWF 和 NCEP，尤其在 4d 以后预报时效、5—10 月期间，大误差日较 ECMWF 和 NCEP 少，但 2~3d 时效、11—3 月则多于 ECMWF。由此可见，OCF 集成订正在气温相对稳定的仲暮春、夏季和初仲秋具有很好的效果，而在降温多发、气温变化起伏较大的深秋、初春和冬季，则效果不佳，集成订正效果与气温变化过程的关系值得进一步探讨。

3.2.2 集成预报大误差日的气温变化特征分析

为分析大误差日的典型过程，按前文气温变化相关定义，统计 OCF、ECMWF、NCEP 在 1~3d 时效各大误差日，和检验时段内逐日实况的明显变温站点比率，进行对比(图 5a,5b)。由于 4d 时效以上各预报的误差日超过了 1/3，气温变化特征与实况对比不明显，因此不开展讨论。

分析发现 OCF 集成预报 1~3d 时效大误差日中，降温站点比率比两个集成成员以及实况明显偏高，集成成员与实况相当；OCF 在 1d 时效、ECMWF 在 1~2d 时效大误差日中明显升温站点比率较实况明显偏高；相比之下 OCF 在 1~3d 大误差日的降温特征较升温特征明显。

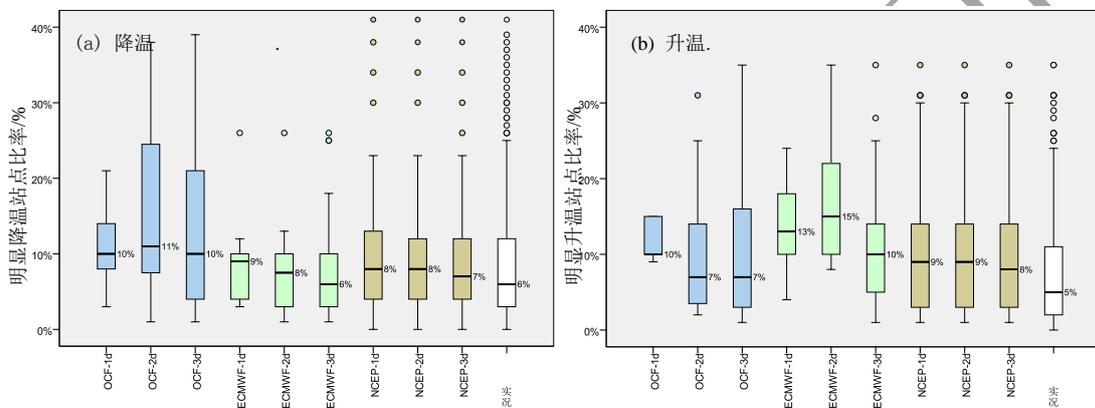


图 5 2020 年 6 月至 2021 年 5 月 OCF、ECMWF 和 NCEP 1~3d 时效大误差日及实况观测的 (a) 明显降温、(b) 明显升温站点比率

(箱体中横线及标注为中位数，“。”为离群值)

Fig. 5 (a) Obvious Temperature-drop rate, (b) obvious Temperature-rise rate of big error days in 1~3d forecasting times of OCF, ECMWF, NCEP's temperature forecasting fields and all observed time from June 2020 to May 2021

(Horizontal line and marked as median, “。” indicates outlier value)

进一步统计各预报大误差日中的变温日(表 1)，分析大误差日和变温日关系。根据 2.2 中变温日的定义，统计得到中国区域检验时段实况有降温日 39d，升温日 32d。OCF 大误差日中的降温日较升温日多，降温日占 1~3d 时效大误差日的 20% 以上，较实况的降温日发生概率高了 1 倍左右，4~7d 中期时段大误差日涵盖了大部分的降温日；ECMWF 1~4d 大误差日中，升温日较降温日多，其大误差日中的降温日是各模式中最少的；NCEP 大误差日中的降温日较升温日多，与 OCF 类似，其 4~7d 中期时段大误差日涵盖了大部分的降温日。

表 1. 2020 年 6 月至 2021 年 5 月 OCF、ECMWF 和 NCEP 大误差日中的变温日统计 (单位: d)

Table 1. Statistics of temperature-change days in big error days of OCF, ECMWF, NCEP from June 2020 to May 2021(unit: d)

预报 时效	OCF			ECMWF			NCEP		
	大误差日数	降温日	升温日	大误差日数	降温日	升温日	大误差日数	降温日	升温日
1	5	1	0	9	1	2	77	11	10
2	26	10	4	17	1	6	85	11	13
3	60	17	10	44	3	7	103	13	14
4	74	19	4	103	7	12	146	19	16
5	135	28	7	146	12	13	172	22	18
6	142	28	7	202	20	15	202	28	19
7	166	26	12	232	21	22	232	29	21

总之，OCF 大误差日的降温特征较升温明显，降温日占 1~3d 时效大误差日 20% 以上，中期时段几乎所有的降温日都出现 20% 以上的大误差；ECMWF 在 1~4d 时效预报更倾向于在升温日出现大误差；NCEP 因大误差日多，大误差的气温变化特征不明显。可以说降温日是 OCF 集成预报的一种典型大误差过程。

3.3 集成预报在降温天气中的预报性能分析

3.3.1 降温日与非降温日的预报性能差异

对比中国区域内 39d 降温日与其余的非降温日 OCF、ECMWF、NCEP 在 1~7d 各时效预报的 RMSE 平均值（表 2），来分析降温过程中的预报效果。发现 OCF 与 NCEP 日最高气温、日最低气温，以及 ECMWF 的日最低气温预报误差在降温日较非降温日都有所增长，ECMWF 的日最高气温预报误差在非降温日和降温日差异不大；OCF 在降温日的日最低气温误差仍小于集成成员，集成预报的优势仍然明显，但 OCF 在降温日的日最高气温误差略大于 ECMWF 模式 0.03℃。

表3-表2. 2020年6月至2021年5月OCF、ECMWF和NCEP各预报时效平均的降温日与非降温日的RMSE(单位: °C)

Table 2.The average RMSE of OCF, ECMWF, NCEP's temperature forecasting feilds in each forecasting periods on temperature-drop days and non-temperature-drop days from June 2020 to May 2021 (unit: °C)

分类	OCF	ECMWF	NCEP
降温日, 日最高气温	3.49	3.21	5.03
非降温日, 日最高气温	2.61	3.16	4.60
降温日, 日最低气温	2.55	2.92	4.24
非降温日, 日最低气温	2.43	2.59	3.86

从 1~7d 各时效降温日与非降温日的 RMSE 增长幅度来看（图 6 中以色块区域表示），日最低气温的误差：NCEP 最高（平均增长 0.38℃），ECMWF 次之（平均增长 0.33℃），OCF 最低（平均增长 0.12℃）；日最高气温的误差：OCF 最高（平均增长 0.88℃），NCEP 次之（平均增长 0.43℃），ECMWF 最低（平均增长 0.24℃）。逐预报时效来看，OCF 1~4d 时效的日最低气温预报误差增长不明显，5~7d 时效误差增长幅度缓慢变大，日最高气温误差增长明显，尤其 2~3d 时效误差增长最快；ECMWF 日最高气温的误差增长较为稳定，降温日和非降温日几乎没有变化，日最低气温的误差增长平稳；NCEP 日最高、日最低气温的误差增长幅度也都较为平稳。将降温日中 OCF 和表现较优的集成成员 ECMWF 的各预报时效 RMSE 差值绘制于图 6a，发现 OCF 各预报时效的日最低气温预报，以及 1~2d 时效的最高气温误差都小于 ECMWF，但 OCF 日最高气温的 RMSE 从 3d 时效开始大于 ECMWF。可见，OCF 降温日中的日最低气温和 2d 时效内日最高气温预报有较好集成订正效果，但 3d 时效以上的日最高气温预报误差增长明显，且误差大于 ECMWF，可以作为 OCF 集成订正模型改进的方向。

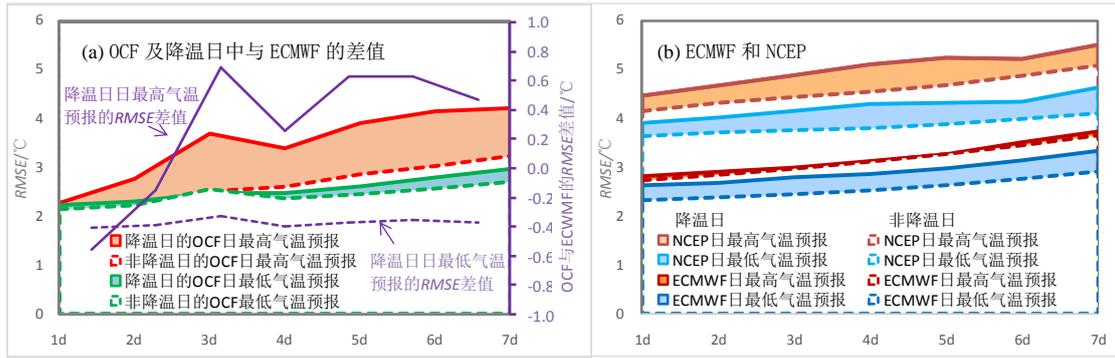


图 6 2020 年 6 月至 2021 年 5 月降温日和非降温日各预报时效 OCF、ECMWF 和 NCEP 预报场的 $RMSE$

(a) OCF 及降温日中 OCF 与 ECMWF 的差值 (b) ECMWF 和 NCEP

(色块部分表示各预报在降温日与非降温日的 $RMSE$ 增长)

Fig. 6 $RMSE$ of OCF, ECMWF, NCEP each time's forecast on temperature-drop days and non-temperature-drop days

from June 2020 to May 2021

(a) daily maximum temperature (b) daily minimum temperature

(Color block part indicates the $RMSE$ increase on temperature-drop days and non-temperature-drop days of each forecast)

3.3.2 降温日误差分析

为分析降温日中的预报误差与气温变化的关系，将各降温日按照高低温平均日变温 $\Delta\bar{T}$ 对将整个检验区域划分为三个变温区域：降温区域 ($\Delta\bar{T} \leq -4^\circ\text{C}$)，升温区域 ($\Delta\bar{T} \geq 4^\circ\text{C}$)，和变温幅度不大的弱变温区域 ($-4^\circ\text{C} < \Delta\bar{T} < 4^\circ\text{C}$)。各降温日都按变温区域对 OCF、ECMWF、NCEP 分 1~3d 短期时段 (图 7a,7b)、4~7d 中期时段 (图略) 统计预报误差频率分布，主要从准确率 ($AE \leq 2^\circ\text{C}$ ，或 $-2^\circ\text{C} \leq E \leq 2^\circ\text{C}$ ，图 7a,7b 中黄色区域)、负误差 ($E < -4^\circ\text{C}$ ，图 7a,7b 中蓝色区域)、正误差 ($E > 4^\circ\text{C}$ ，图 7a,7b 中红色区域) 等进行分析。

从降温日 1~3d 短期时段各日最低气温预报在各变温区域的误差频率分布可见 (图 7a)，OCF、ECMWF、NCEP 在各变温区域的误差分布特征类似：降温区域的预报准确率高于升温 and 弱变温区域，在升温 and 弱变温的负误差较多，即预报气温较实况偏低；OCF、ECMWF 在降温区域正误差的比率增加，各模式在升温区域的负误差较降温和弱变温区域增加；各变温区域预报准确率最高的均是 OCF。

OCF、ECMWF、NCEP 日最高气温预报误差分布在各变温区域差异较大：各预报在升温区域的负误差都较降温和弱变温区域增加，降温区域的正误差较升温 and 弱变温区域增加，尤其 OCF 降温区域正误差比率明显增加， 4°C 以上误差比率平均达到了 20%，而 -4°C 以下误差比率平均仅为 3%，正误差倾向明显，即 OCF 日最高气温预报在降温区域预报偏高；OCF 日最高气温的预报准确率在降温区域达 68%，明显高于升温区域 (44%)，而 ECMWF、NCEP 的预报准确率在降温区域较高。各预报比较发现：升温区域预报准确率 OCF (68%) 高于 ECMWF (52%)、NCEP (33%)，降温区域预报准确率 ECMWF (70%)、NCEP (55%) 高于 OCF (44%)；ECMWF、NCEP 对最高气温预报偏低 4°C 以上 (负误差) 的比率较 OCF 明显偏高约 10% 以上，OCF 有效订正了两个集成成员在降温日中日最高气温预报偏低的误差倾向；但 OCF 在升温 and 弱变温区域正误差比率较 ECMWF、NCEP 明显偏高，说明在降温日的非降温区域，OCF 集成订正效果不佳，反而产生了预报偏高的误差倾向。

降温日 4~7d 中期时段的各预报，较 1~3d 短期时段的准确率降低，日最低气温和最高气温的正误差都有增长，但负误差增长不明显。可见在中期时段的气温预报中，更体现出现

报偏高的误差倾向，进而导致对降温强度预报偏弱；各预报各变温区域的误差特征分布与1~3d 短期时段类似，不再赘述。

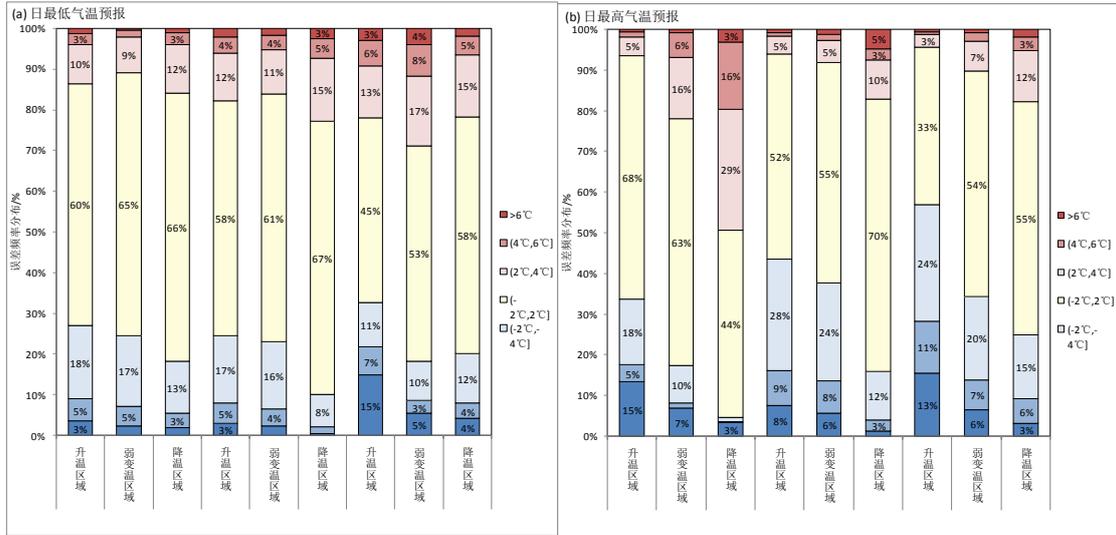


图7 检验时段内降温日中各变温区域1~3d 短期时段 OCF、ECMWF 和 NCEP 气温预报误差频率分布 (a) 日最低气温, (b) 日最高气温

Fig. 7 OCF, ECMWF, NCEP 1~3d forecasting times' error frequency distribution of each temperature-change region in verified temperature-drop days (a) daily minimum temperature (b) daily maximum temperature

总之，在降温日，OCF 对各变温区域的日最低气温、非降温区域的日最高气温都具有集成订正效果；但其在降温区域日最高气温预报偏高，准确率较低，正误差明显增多，与 RMSE 较大对应。OCF 这一突出的误差分布特征与集成成员不同，说明误差并非成员的影响，而是集成订正的方法问题；OCF 有效订正了降温日中各变温区域的最高气温偏低的误差，但同时也造成了降温区域的最高气温正误差增多。对各降温过程单独考察，发现 OCF 这种突出的误差特征并非少数过程造成，且大部分过程的降温强度预报偏弱，可关联到 OCF 在降温区域日最高气温预报偏高的误差倾向。

4 结论与讨论

(1) 从常规检验来看，OCF 集成订正及集成成员在中国区域各预报时效的气温预报性能，体现为：预报误差随预报时效增长而增大，日最低气温预报效果明显优于日最高气温的预报；在全部检验时段内 OCF 较 ECMWF 和 NCEP 总体上预报性能较优，尤其在 1~2d 时效的日最高气温预报，其优势明显；分月来看 OCF 及各集成成员准确率多夏季高冬季低，OCF 集成预报拉开了气温变化范围，也有效减小了误差。

(2) 针对大误差日的分析发现：OCF 大误差日总数少于 ECMWF 和 NCEP，但 2~3d 时效时段的大误差日多于 ECMWF，在气温相对稳定的暮春到仲秋（5—10 月）大误差日较少，而在降温多发、气温变化起伏较大的深秋到冬季（11—3 月）大误差日甚至多于 ECMWF。OCF 大误差日的降温特征较升温明显，中期时段大部分降温日都出现 20% 以上的大误差，降温日是 OCF 的一种典型大误差过程。

(3) 降温日和非降温日的对比检验揭示了 OCF 和 ECMWF 和 NCEP 在降温日的误差都有所增长，尤其是 OCF 在 3d 时效以上的日最高气温预报的误差增长明显。对降温日各预报划分变温区域分析误差分布发现，OCF 对各变温区域的日最低气温、非降温区域的日

最高气温都具有集成订正效果；但其在降温区域日最高气温预报偏高，准确率低于非降温区域，也低于集成成员 ECMWF 和 NCEP，正误差明显增多。

(4) 在降温过程中最高气温的变幅往往很大，易出现预报误差增大；降温区域预报偏高，升温区域预报偏低的误差特征是符合常理的。但 OCF 日最高气温预报在降温日的降温区域明显偏高，可理解为其有效订正了集成成员预报偏低的误差倾向，但同时也造成了在降温区域的正误差增多。这种误差主要是集成订正的方法问题，可作为集成预报改进的方向之一。另外，OCF 集成订正误差或也与其他集成成员预报效果差有关。采取针对降温日的最高气温、日变温等来构建不同于常规的调整训练周期、预报偏差指标、集成权重等方式来优化集成策略，或是进一步提高集成订正预报性能的道路之一。

(5) 在 OCF 整体预报性能较优的认识下，通过对大误差的典型过程研究，揭示了 OCF 集成订正方法的改进方向，说明针对过程的检验更利于发现客观预报及集成订正方法的问题。

参考文献

- 郭虎, 王建捷, 杨波, 等, 2008. 北京奥运演练精细化预报方法及其检验评估[J]. 气象, 34(6): 17-25. Guo H, Wang J J, Yang B, et al, 2008. Refined forecast techniques and related verification during 2007 Olympic drilling[J]. Meteor Mon, 34(6): 17-25 (in Chinese).
- 郝翠, 张迎新, 王在文, 等, 2019. 最优集合预报订正方法在客观温度预报中的应用[J]. 气象, 45(8): 1085-1092. Hao C, Zhang Y X, Wang Z W, et al, 2019. Application of analog ensemble rectifying method in objective temperature prediction[J]. Meteor Mon, 45(8): 1085-1092(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 R 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).
- 李佰平, 智协飞, 2012. ECMWF 模式地面气温预报的四种误差订正方法的比较研究[J]. 气象, 38(8): 897-902. Li B P, Zhi X F, 2012. Comparative study of four correction schemes of the ECMWF surface temperature forecasts[J]. Meteor Mon, 38(8): 897-902 (in Chinese).
- 刘还珠, 赵声蓉, 陆志善, 等, 2004. 国家气象中心气象要素的客观预报—MOS 系统[J]. 应用气象学报, 15(2): 181-191. Liu H Z, Zhao S R, Chen Z S, et al, 2004. Objective element forecasts at NMC- a MOS system[J]. J Appl Meteor Sci, 15(2): 181-191 (in Chinese).
- 任萍, 陈明轩, 曹伟华, 等, 2020. 基于机器学习的复杂地形下短期数值天气预报误差分析与订正[J]. 气象学报, 78(6): 1002-1020. Ren P, Chen M X, Cao W H, et al, 2020. Error analysis and correction of short-term numerical weather prediction under complex terrain based on machine learning[J]. Acta Meteor Sin, 78(6): 1002-1020 (in Chinese).
- 盛春岩, 范苏丹, 荣艳敏, 等, 2020. 几种气温客观预报方法对比及最优集成预报研究[J]. 气象, 46(10): 1351-1361. Sheng C Y, Fan S D, Rong Y M, et al, 2020. Comparison of several objective methods and optimal consensus forecast study of temperature[J]. Meteor Mon, 46(10): 1351-1361 (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).
- 韦青, 李伟, 彭颂, 等, 2019. 国家级天气预报检验分析系统建设与应用[J]. 应用气象学报, 30(2): 245-256. Wei Q, Li W, Peng S, et al, 2019. Development and application of national verification system in CMA[J]. J Appl Meteor Sci, 30(2): 245-256 (in Chinese).
- 张亚刚, 杨银, 张成军, 等, 2021. 基于径向基函数神经网络温度预报订正方法及评估[J]. 热带气象学报, 37(1): 136-144. Zhang Y G, Yang Y, Zhang C J, et al, 2021. Temperature forecast correction method and evaluation based on radial basis function neural network[J]. Journal of Tropical Meteorology, 37(1): 136-144 (in Chinese).
- 赵声蓉, 2006. 多模式温度集成预报[J]. 应用气象学报, 17(1): 52-58. Zhao S R, 2006. Multi-model consensus forecast for temperature[J]. J Appl Meteor Sci, 17(1): 52-58 (in Chinese).
- Engel C, Ebert E, 2007. Performance of hourly operational consensus forecasts (OCFs) in the Australian region[J]. Wea Forecasting, 22(6): 1345-1359.
- Engel C, Ebert E, 2012. Gridded operational consensus forecasts of 2-m temperature over Australia[J]. Wea Forecasting, 27(2): 301-322.
- Glahn H R, Lowry D A, 1972. The use of model output statistics (MOS) in objective weather forecasting[J]. J Appl Meteor Climatol, 11(8): 1203-1211.
- Glahn H R, Ruth D P, 2003. The new digital forecast database of the national weather service[J]. Bull Amer Meteor Soc, 84(2): 195-202.
- Hyndman R J, Fan Y, 1996. Sample Quantiles in statistical packages[J]. Amer Statist, 50(4): 361-365.
- Hoffman R R, LaDue D S, Mogil H M, et al, 2017. Minding the Weather: How Expert Forecasters Think[M]. Cambridge: The MIT Press.
- Kann A, Wang Y, Atencia A, et al, 2018. Seamless probabilistic analysis and forecasting: from minutes to days ahead[C]//Proceeding of the 20th EGU General Assembly. Vienna: EGU: 7962.
- Murphy A H, 1996. The Finley affair: a signal event in the history of forecast verification[J]. Wea Forecasting, 11(1): 3-20.
- Murphy A H, Winkler R L, 1987. A general framework for forecast verification[J]. Mon Wea Rev, 115(7): 1330-1338.
- Stensrud D J, Skindlov J A, 1996. Gridpoint predictions of high temperature from a mesoscale model[J]. Wea Forecasting, 11(1): 103-110.
- Woodcock F, Engel C, 2005. Operational consensus forecasts[J]. Wea Forecasting, 20(1): 101-111.