

郑嘉雯,高丽,任宏利,等,2019. 基于 T639 集合预报的我国 2016 年极端温度预报检验[J]. 气象,45(4):469-482. Zheng J W, Gao L, Ren H L, et al, 2019. Verification of China extreme temperature forecasts in 2016 based on T639 ensemble forecast[J]. Meteor Mon, 45(4):469-482(in Chinese).

基于 T639 集合预报的我国 2016 年 极端温度预报检验^{*}

郑嘉雯^{1,2,3} 高 丽² 任宏利^{1,3} 陈权亮¹ 蔡宏珂¹

¹ 成都信息工程大学, 成都 610225

² 国家气象中心, 北京 100081

³ 国家气候中心, 北京 100081

提 要: 基于我国 T639 集合预报系统的 2 m 温度模式实时预报和分析场资料, 对历史气候百分位分布进行了估算, 并对 2016 年我国极端高温和极端低温集合预报进行了检验评估。分析表明, 对于某一区域代表站点, 不同超前时间对应的气候百分位分布变化趋势均与分析场的基本一致, 但不同区域代表站点之间以及同一站点不同季节之间的气候百分位分布则存在差异。Talagrand 分布检验结果显示, T639 集合系统对哈尔滨和长沙站的预报存在暖偏差, 而对北京和拉萨站的预报则存在冷偏差, 其余典型站表现出预报离散度偏小的特点。基于历史气候百分位定义, 对 2016 年极端温度进行识别、预报和检验。TS 评分结果表明, T639 集合系统对于我国极端温度总体上具有一定预报性能, 预报技巧在各个区域有较大差别, 极端高温在江南和东北区域技巧相对较高、极端低温在华北和南方技巧较高。对于不同预报方法技巧比较显示, 集合平均法对极端信号有平滑效果, 总体技巧偏低, 集合众数法对低温技巧增加明显, 而采用集合最大值和最小值法对极端信号则有放大效果, 对高温技巧增加明显, 但对低温技巧则不明显。由此可见, 采用合适的方法对集合预报中极端信息的正确提取至关重要。

关键词: T639 集合预报, 极端高温, 极端低温, 模式, 检验

中图分类号: P456

文献标志码: A

DOI: 10.7519/j.issn.1000-0526.2019.04.003

Verification of China Extreme Temperature Forecasts in 2016 Based on T639 Ensemble Forecast

ZHENG Jiawen^{1,2,3} GAO Li² REN Hongli^{1,3} CHEN Quanliang¹ CAI Hongke¹

¹ Chengdu University of Information Technology, Chengdu 610225

² National Meteorological Centre, Beijing 100081

³ National Climate Centre, Beijing 100081

Abstract: Based on the real-time forecast and analysis data of 2 m temperature from the T639 ensemble forecast system in China, the climatic percentile distribution is estimated for both real-time forecast and analysis data, and the forecasts of extreme high temperature events and extreme low temperature at typical stations over China in 2016 are verified and evaluated. The analysis results show that the forecasted climate percentile distribution at different lead time is basically the same as the percentile distribution of the analysis for a given station site, but they are different to some degree between different regions and

^{*} 国家自然科学基金项目(41875138)、国家科技支撑计划项目(2015BAC03B01)、气象预报业务关键技术发展专项(YBGJXM2017-04)和中央引导地方科技发展专项(ZY18C12)共同资助

2018 年 1 月 17 日收稿; 2019 年 1 月 2 日收修定稿

第一作者: 郑嘉雯, 主要从事集合预报研究. Email: 546575108@qq.com

通信作者: 高丽, 主要从事大气动力学与集合预报研究. Email: gaol@cma.gov.cn

different seasons. The results of the Talagrand distribution show that there is a warm bias in the forecasts of Harbin and Changsha Stations by the T639 ensemble system and a cold bias in the predictions of Beijing and Lhasa Stations. The other typical stations show a low ensemble spread. Based on the definition of historical climate percentile, the 2016 extreme temperature identified, forecasted and verified. The results of TS score show that the T639 ensemble system has a certain forecasting performance for extreme temperatures in China, but the prediction skill has apparent difference in different regions. The extreme high temperature has higher prediction skills in the south to the Yangtze River and the Northeast China, and the extreme low temperature has higher prediction skills in the northern and southern parts of China. The comparisons of different forecasting methods show that the ensemble mean method has a smoothing effect on the extreme signals and the overall prediction skill is low. The ensemble model method improves obviously the extreme low temperature skill, while the method of using maximum or minimum in all ensemble members can amplify the extreme signals and increase significantly the prediction skill of extreme high temperature, but not clear for low temperature. Thus, we can see that proper extraction of extreme information in ensemble forecasting is of crucial importance.

Key words: T639 ensemble forecast, extreme high temperature, extreme low temperature, model, verification

引 言

极端天气事件是一定地区在一定时间内出现的历史上罕见的气象事件。其定义方法很多(Beniston et al, 2007; 罗亚丽, 2012), 21 世纪初, 政府间气候变化专门委员会(Intergovernmental Panel on Climate Change, IPCC)第三次评估报告与第四次评估报告(Houghton et al, 2001; Solomon et al, 2007)对极端天气事件给出明确定义: 极端天气事件是指发生概率小于观测记录概率密度函数第 10 个百分点数或超过第 90 个百分点数的天气事件。按照定义, 在绝对意义上, 极端天气特征因地区不同而异。

极端天气事件会对人类生命和财产造成指数型增长的损失, 近些年对极端天气事件的研究受到气象学者的广泛关注(Karl and Knight, 1998; 关吉平等, 2006; 张立凤和罗雨, 2010; 黄刚等, 2012; 段明铨等, 2012)。我国由于地域辽阔、地形复杂, 导致极端天气事件种类多、频次高、阶段性和季节性明显、区域差异大、影响范围广。前人研究表明, 近年来随着全球气温上升, 我国极端温度事件发生了显著变化, 高温日数增加, 极端低温频次明显下降(任国玉等, 2005)。然而, 极端天气事件研究仍是一个薄弱环节, 如何描述和监测极端天气或表征天气极值的各种统计特征及其变化规律, 已经成为重要课题之一(Meehl et al, 2000; Easterling et al, 2000)。

由于大气是一个混沌系统, 小概率的极端天气

事件的发展存在着很大的不确定性, 用单一初值的数值模式对其进行预报是远远不够的, 因此 20 世纪 70 年代前后国际上引入了集合预报的概念(Epstein, 1969; Leith, 1974)。经过近半个世纪的发展, 集合预报已经成为数值预报领域的一个重要方向。相对于单一初值的确定性预报, 集合预报能够考虑多个初值形成的不确定性, 其结果能够反映出未来天气的多种状况, 从而为预报员提供确定性预报之外的概率性预报信息(如预报结果的可信度等)(彭勇等, 2002; 杜钧, 2002; 皇甫雪官, 2002; Richardson, 2000; Palmer et al, 2000)。由于极端天气事件发生的历史样本很有限, 很难发掘统计上的规律性, 为此, 基于数值模式的集合概率预报逐渐成为极端天气事件预报的主要工具。李俊等(2015)对北京 2012 年 7 月 21 日的极端降水事件组织了一系列中尺度(水平格点间隔为 9 km)的集合预报试验, 在所有的集合预报之中, 利用了初始扰动、多物理模型、多初始值和边界条件的预报是最优预报, 并且对降水的强度和降水落区预报的准确度远远高于单一的确定性预报。

在对各预报要素(降水、风、温度和云量)的数值预报和集合预报系统的发展下, 预报员希望能利用集合预报系统对极端事件发出早期的预警信号。然而直接比较观测的气象要素和模式输出的预报之间的差别是很难的(Ghelli and Lalaurette, 2000), 因此提出了一种从集合预报系统中提取异常预报分布的方法。Lalaurette(2002)基于 ECMWF 集合预报

发展了一种极端天气预报方法——极端预报指数(extreme forecast index, EFI),其原理是计算累积的模式气候概率分布和集合预报概率分布之差。近年来许多气象学者对 EFI 进行了改进, Lalaurette (2003) 基于 Anderson-Darling 检验修正了原始的 EFI 公式, Zsótér (2006) 使用修正过的 EFI 公式对 2005 年 10 月西欧持续极端高温进行研究, 结果表明与原始 EFI 相比, 修正后的 EFI 能预测出更多发生极端高温的区域。

国外已经有不少学者利用极端预报指数对极端天气(如:极端降水、极端气温和极端洪水事件)进行研究。Guan and Zhu (2017) 的工作表明, EFI 适用于 2013/2014 年冬季的极端低温和极端降水预报。Lavers et al (2017) 通过对 2015/2016 年和 2016/2017 年冬季的极端降水的预测分析表明, 提前两周的极端降水预报中 IVT EFI 指数比极端降水预报指数更有技巧, 而在较短的预报时效中, 极端降水预报指数预报技巧比较高。国内基于数值模式集合的极端天气事件预报研究也已经得到开展, 夏凡和陈静(2012)和吴剑坤等(2015)基于我国 T213 集合预报系统资料, 利用极端天气预报指数对极端温度事件进行了预报试验和检验评估。结果表明, 极端天气预报指数对极端温度事件有较好的识别能力, 可以提前 3~7 d 发出极端温度预警信号, 随着预报时效的延长, 预报技巧逐渐降低。董全等(2017)通过个例分析和大样本的总结表明, ECMWF 基于 EPS 的 EFI 产品对我国极端高温、低温和极端强降水都具有较好的预报能力。

中国气象局 T639 全球集合预报业务系统, 较之 T213 集合系统性能上有了显著提升, 在气温概率预报业务上发挥了重要作用。但目前 T639 集合系统在极端天气预报性能方面还没有相关检验评估研究。本文拟利用 T639 实时集合预报的历史资料开展对极端温度预报的关键环节——模式气候百分位分布的估算和诊断分析, 并对我国 2016 年极端温度(高温和低温)进行 T639 集合系统预报的试验和检验, 为后续开展极端预报指数研究提供参考信息。

1 资料及方法

1.1 模式和资料介绍

本文所用模式预报资料来自于中国气象局数值预报中心发展的 T639 集合预报系统。该系统是由

中国气象局数值预报中心开发的全球集合数值预报业务系统, 于 2014 年 8 月 5 日正式投入业务运行。其初始扰动采用增长模繁殖法产生, 物理过程扰动采用随机参数化扰动方案, 格点空间的水平分辨率为 $0.28125^{\circ} \times 0.28125^{\circ}$ 、垂直分辨率为 60 层, 采用三维变分同化分析系统, 并采用了 Rayleigh 摩擦增加平流层的稳定性。与 T213 模式相比, T639 模式对物理过程的改进主要增加了次网格的对流活动, 包括云方案改进, 采用了与 T639 模式匹配的下垫面资料处理系统及合理的初始化方案等(管成功等, 2008); T639 模式动力框架的改进主要包括采用线性高斯格点, 同时改进 T639 模式的半拉格朗日时间积分方案, 使之在线性高斯格点下计算噪音问题得到有效克服, 并延长积分步长。

采用了 T639 集合预报系统 12 时(世界时, 下同)起报, 全国范围内的 2 m 温度历史实时预报资料, 其中 2009 年 1 月至 2015 年 7 月为控制预报(2014 年 8 月至 2015 年 7 月的集合成员预报数据由于系统技术原因只有不连续的部分保留, 本文没有采用); 2015 年 8 月至 2016 年 12 月为 15 个集合成员的集合预报。预报时效分别为 24、48、72、96、120、144、168、192、216 和 240 h 共 10 d。为了保持与模式预报资料空间和时间分辨率的一致性, 本文实况场采用的是同一时间段由 T639 集合系统同化得到的 2 m 气温分析场。另外, 文中使用的典型台站的数值由站点附近格点数值代替。

另外, 本文还采用了中国国家级地面气象站均一化气温日值 $0.5^{\circ} \times 0.5^{\circ}$ 格点数据集资料作为实况资料进行模式气候与观测气候之间差异的对比分析, 本数据集的数据来源包括 2 个部分: 由国家气象信息中心研制的“中国国家级地面气象站均一化气温日值数据集(V1.0)”; 由 GTOPO30 数据(分辨率为 $0.05^{\circ} \times 0.05^{\circ}$)经过重采样生产的中国陆地 $0.5^{\circ} \times 0.5^{\circ}$ 的数字高程模型 DEM。并且采用交叉验证质量检验法进行质量控制, 质量状况良好。

1.2 计算方法

在计算观测和模式分析场气候百分位时, 选取计算日期及前后 7 d 为滑动时间窗(15 d)、计算格点及附近 9 个格点为空间窗的分析场作为气候序列进行升序排序(序列长度为 $8 \text{ a} \times 15 \text{ d} \times 9$ 个格点), 每个百分位点(0.01, 0.02, ..., 0.99, 1)对应一个阈值, 即模式分析场的气候百分位分布是格点位置和

起报时间的函数。模式预报场气候百分位计算在分析场百分位计算的基础上考虑了使用多个集成员来增加样本数(序列长度为 $8\text{ a} \times 15\text{ d} \times 9\text{ 个格点} \times 15\text{ 个集成员}$),即模式气候百分位分布是格点位置、起报时间和预报时效的函数。输出全国范围分析场和模式气候每个格点,逐日各个预报时效 1~100 个百分位点对应的值。定义气候序列的第 5 个百分位或者以下的 2 m 温度值为极端低温,定义气候序列的第 95 个百分位或者以上的 2 m 温度值为极端高温。为了便于评估数值集合对于极端温度的预报性能,我们对于极端温度是以日数进行定义的。同时考虑到极端温度事件样本稀少、不利于进行每日滚动的集合预报检验,为此本文采用对极端温度日数为检验评估对象,以计算得到相对稳定的检验评分。

预报结果的检验是至关重要的。由于本文的检验评估考虑的因素为逐格点上极端温度的实况和预报两种状态,故利用表 1 所示的双态分类列联表来计算检验评分,包括命中率(R_H)、假警报率(R_{FA})和 TS 风险评分(TS),具体计算方法见公式(1)~公式(3)。

表 1 双态分类列联表

Table 1 Contingency table of dimorphic distribution

类别	预报出现	预报不出现
实况发生	A	C
实况不发生	B	D

$$R_H = \frac{A}{A + C} \quad (1)$$

$$R_{FA} = \frac{B}{B + D} \quad (2)$$

$$TS = \frac{A}{A + B + C} \quad (3)$$

TS 评分和命中率的取值范围为 0~1,越接近 1 代表预报效果越好;假警报率的范围也为 0~1,越接近于 0 代表预报效果越好。下面首先对分别基于观测和模式分析场的气候百分位进行对比分析,这里针对的是 8 个典型站。

图 1 可见,基于 T639 模式分析场历史资料计算的和实况观测场气候百分位概率分布存在一定差异。对于乌鲁木齐和成都,分析场气候百分位分布位于观测气候百分位分布右边,即 T639 模式 2 m 温度的预报存在正偏差,其余代表站点的分析场气

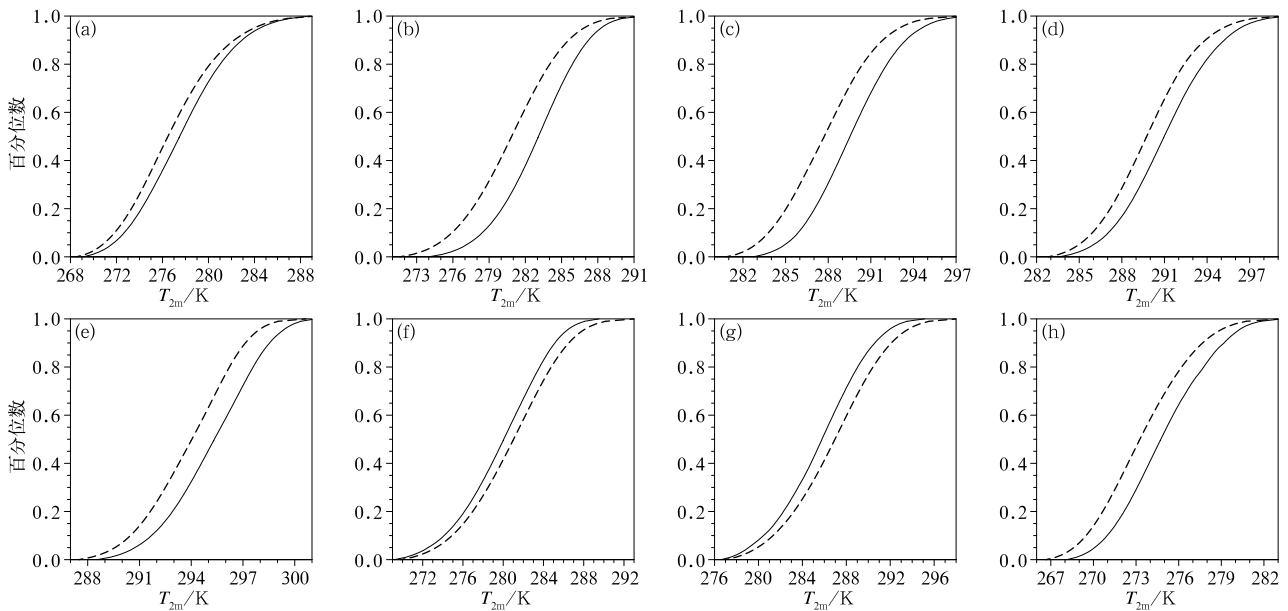


图 1 2016 年 6 月 15 日各区代表站点年平均分析场(虚线)和观测场(实线)气候百分位分布

(a) 哈尔滨, (b) 北京, (c) 南京, (d) 长沙, (e) 广州, (f) 乌鲁木齐, (g) 成都, (h) 拉萨

Fig. 1 Analysis (dashed line) and observed (solid line) climate percentile distributions

for the whole year average in all regions on 15 June 2016

(a) Harbin, (b) Beijing, (c) Nanjing, (d) Changsha, (e) Guangzhou,

(f) Urumqi, (g) Chengdu, (h) Lhasa

候百分位分布位于观测气候百分位分布左边,即 T639 模式 2 m 温度的预报存在负偏差。为了剔除系统偏差、并便于在统一的空间格点上进行预报检验,本文直接采用 T639 集合系统同化得到的 2 m 气温分析场作为实况场来计算检验评分。

1.3 极端温度的概率预报方法

集合预报已经成为目前进行天气预报的主流方法,但集合预报输出的信息量非常庞大,利用有效的方法对集合成员中的极端信息进行提取至关重要。集合均值(ensemble mean, EM)法,就是将集合成员进行等权重平均,这种方法可以过滤掉不可预报的随机信息,但是也会过滤掉少数成员预报出来的极端天气。吴岩等(2014)的研究表明,针对降水和温度预报,总体而言,集合成员中集合平均值预报表现相对优秀,可在业务中作为重要参考。段明铿等(2009)使用 NCEP 集合预报资料,对环流预报效果进行检验,结果表明预报时效大于 5 d 时,集合平均预报明显优于单一预报。最大值、最小值(maximum 或 minimum, Max 或 Min)法,是利用集合成员的最大值或最小值进行预报的一种方法。吴岩等(2014)的研究表明,集合预报成员对于暴雨预报漏报明显,而集合成员中的最大值,对暴雨预报有一定的启示作用,业务工作中应得到重视。包慧濛等(2015)利用 ECWMF 的集合预报资料对江西省降水的研究表明,集合最大值对于暴雨量级的预报有一定的指示意义。为此,本文将选取集合成员的最大值和最小值进行极端温度的预报。众数(Mode)法是采取集合成员众数进行天气预报的方法,胡海川等(2017)的研究表明,集合预报 10 m 风速众数在各个预报时效的正确率要高于集合中值和平均值。本文将分别利用这三种方法开展极端温度的预报及其检验。

2 气候百分位分布的计算和验证

2.1 同一时间、不同地点的气候百分位分布验证

我国地域辽阔,地理空间差异大,本文参照王岱等(2016)的方法将中国分成东北、华北、华东、华中、华南、西北、西南和高原 8 个区域,这 8 个区域同时

考虑了极端高温和极端低温的适用性,并分别选取哈尔滨、北京、南京、长沙、广州、乌鲁木齐、成都和拉萨站点作为各区域代表站点对 T639 模式(分析场)气候百分位进行验证。图 2 是以 2016 年 6 月 15 日为例给出了 8 个区域代表站点的模式(分析场)2 m 温度气候百分位分布曲线图,其中模式气候百分位分布的超前时间为 1~10 d。

从图 2 中不难看出,对于某一区域代表站点、不同超前时间对应的气候百分位分布变化趋势均与分析场的基本一致。总体而言,超前时间越短,模式预报的气候百分位分布越接近分析场的气候百分位分布。而且,不同超前时间的气候百分位分布在各代表站点之间略有差异,其中哈尔滨、北京、长沙和广州站点的不同超前时间对应的气候百分位分布存在微小差异,南京、乌鲁木齐、成都和拉萨站点则差异较明显。其中乌鲁木齐和成都站点在 0.1~0.8 气候百分位数区间内,不同超前时间的模式(分析场)气候百分位分布差异最大,并且超前时间越长, T639 模式的 2 m 温度气候百分位数偏低得越明显。这反映出计算气候百分位分布时,时效和空间格点的位置都需要分别考虑,也从侧面验证了气候百分位分布是这两者的函数。

图 2 中垂直竖线表征的是各站点 2016 年 6 月 15 日 2 m 温度的观测值,它处于气候百分位对应阈值范围的中间是最理想的状况。但图中可以看出,除了广州站点观测值基本处在气候百分位数的中间外,其余各区代表站点都不位于中间,其中乌鲁木齐和拉萨站点的预报存在冷偏差的情况,其他区域的代表站点的预报都存在不同程度的暖偏差。

对于一个好的集合预报系统,应该拥有一个适宜的离散度,即每个集合成员预报的发生概率均等,实况可能接近集合成员预报的任何一个。为了验证 T639 集合预报系统对不同区域代表站点的 2 m 温度预报是否存在不同的冷暖偏差,本文采用 Talagrand 分布的检验方法,将 T639 模式集合预报系统的 15 个集合成员按升序排序后构成 16 个区间,然后用实况值落在这 16 个区间的时间累积频率来描述集合预报系统的可靠性。若是一个理想的集合系统,实况落在每一个区间的概率是相同的,即 Talagrand 分布是平直的;若 Talagrand 分布呈“L”型,表明实况大多落在集合预报小值区,即该模式集合

预报系统有一定程度的暖偏差,反之亦然;若 Talagrand 分布呈“U”型,说明该模式集合预报系统的离散度偏小。

图 3 为 2015 年 8 月以来 8 个区域代表站点 2 m

温度 72 h 预报的 Talagrand 分布。可看出每个区域代表站点的 Talagrand 分布都有差异,并且各区域代表站点的实况值落在大多数区间的概率都低于理想概率,说明集合预报系统的离散度不够;其中哈

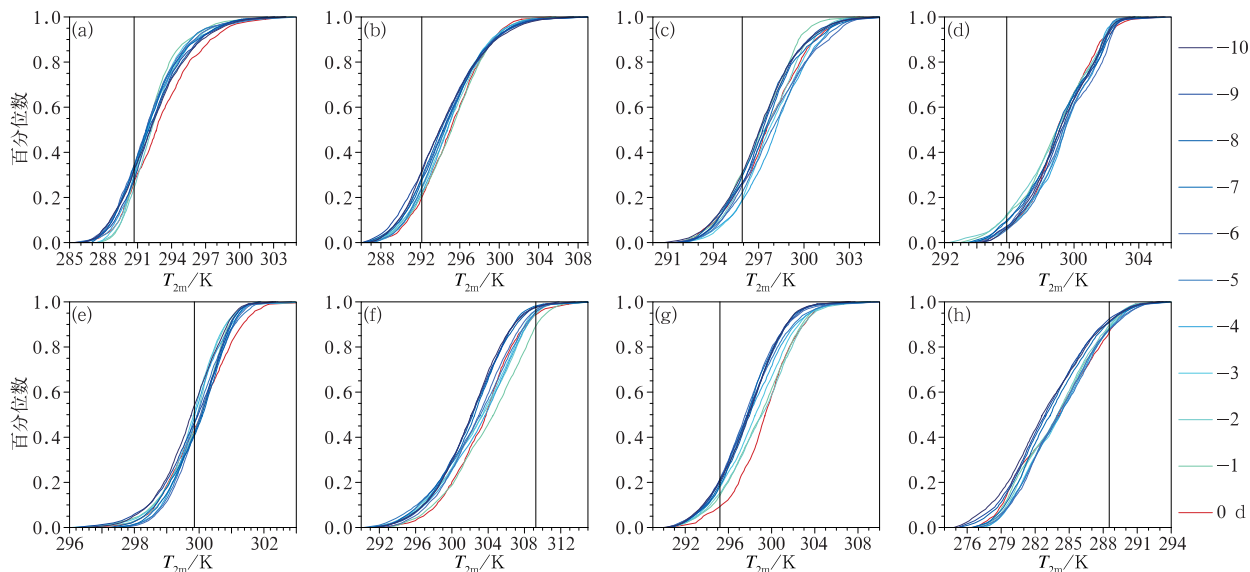


图 2 各区代表站点 2016 年 6 月 15 日起报的模式(分析场)气候百分位分布

(a)哈尔滨, (b)北京, (c)南京, (d)长沙, (e)广州, (f)乌鲁木齐, (g)成都, (h)拉萨

Fig. 2 Forecasted (analysis) climate percentile distribution in all regions on 15 June 2016

(a) Harbin, (b) Beijing, (c) Nanjing, (d) Changsha, (e) Guangzhou, (f) Urumqi, (g) Chengdu, (h) Lhasa

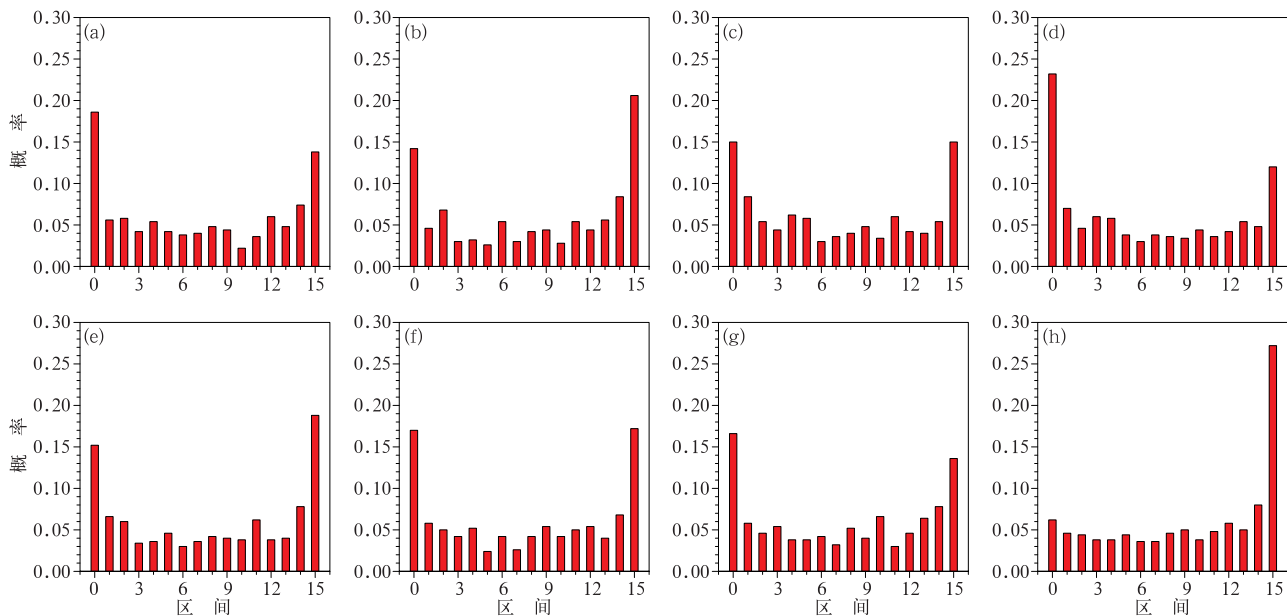


图 3 我国各区域代表站点 2 m 气温 72 h 预报的 Talagrand 分布

(a)哈尔滨, (b)北京, (c)南京, (d)长沙, (e)广州, (f) 乌鲁木齐, (g)成都, (h)拉萨

Fig. 3 Talagrand distribution of 72 h forecast of 2 m temperature in all regions of China

(a) Harbin, (b) Beijing, (c) Nanjing, (d) Changsha, (e) Guangzhou, (f) Urumqi, (g) Chengdu, (h) Lhasa

尔滨和长沙站点的 Talagrand 分布呈“L”型,即实况值落在集合预报最小值之外的概率较大,可看出 T639 集合预报系统对这两个区域的 2 m 温度预报,尤其是华中区域存在暖偏差;而在华北和高原区域的代表站点的 Talagrand 分布则呈“反 L”型,即 T639 集合预报系统对这两个区域的 2 m 温度预报,尤其是高原区域存在冷偏差;其余大部分区域代表站点的 Talagrand 分布呈“U”型,说明 T639 集合预报系统的离散度偏小,这也是目前国内外数值天气集合预报系统中普遍存在的问题,会对极端温度的预报产生一定影响。

2.2 同一地点、不同时间的气候百分位分布验证

对于不同时间模式(分析场)气候百分位分布的验证,本文选取华中的代表站点,即长沙站点进行验证,并对 1、4、7 和 10 月的气候百分位分布进行月平均作为各季节模式(分析场)气候百分位分布的代表进行验证(图 4)。

由图可见,整体而言,超前时间越短,模式气候百分位分布曲线越接近历史气候百分位分布曲线。不同季节的情况有所差异,其中冷暖季(1 和 7 月)各超前时间的气候百分位分布差异较小,即长沙站点冷暖季可预报时效较长。但夏季(7 月),各超前时间在模式气候百分位分布的高端尾部部分有明显偏冷的表现,即 T639 集合系统对长沙站点夏季极端高温的预报强度偏弱。转换季(4 和 10 月)各超前时间的气候百分位分布差异较大,秋季的差异最大,尤其是在高端尾部部分,即 T639 集合系统对长沙站点秋季极端高温的预报强度偏强,且超前时间越短,这种偏强的表现越不明显。反映出在进行历史(模式)气候百分位分布的计算中,需要考虑季节

差异,在转换季则要着重考虑不同预报时效之间的差异。

3 2016 年极端温度的预报检验

本文对我国 2016 年极端温度进行识别和预报的检验,并针对极端高温、极端低温分别选取浙江省的龙泉站、北京站(华北代表站点)作为对象进行研究。首先利用分析场结合历史气候百分位分布对各代表站点的极端高/低温进行识别。其次,利用集合成员的均值、最大/小值和众数结合模式气候百分位对极端温度进行预报。最后,对分析场识别的极端温度与模式预报的进行对比分析。

3.1 2016 年极端温度的识别

利用历史气候百分位和分析场可对极端温度进行识别,图 5 展示的是 2016 年我国极端高温和极端低温发生的日数分布。由图可见,2016 年我国极端低温的发生频次较低,主要在大兴安岭北麓、内蒙古中西部以及云南的发生频次较高;与极端低温相比,2016 年极端高温发生的频次较高,全国而言,东北三省、河套地区以及江南地区发生极端高温的频次较高。

3.2 2016 年极端高温的预报检验

图 6 为 T639 集合系统对 2016 年极端高温 72 h 预报的 TS 评分全国分布图,分析显示:T639 集合系统对于我国极端高温总体上具有一定的预报性能。预报技巧在各个区域有较大的差别,在江南和东北区域技巧相对较高。从全国 TS 评分的平均值看出,采取集合成员最大值法对极端高温进行预报

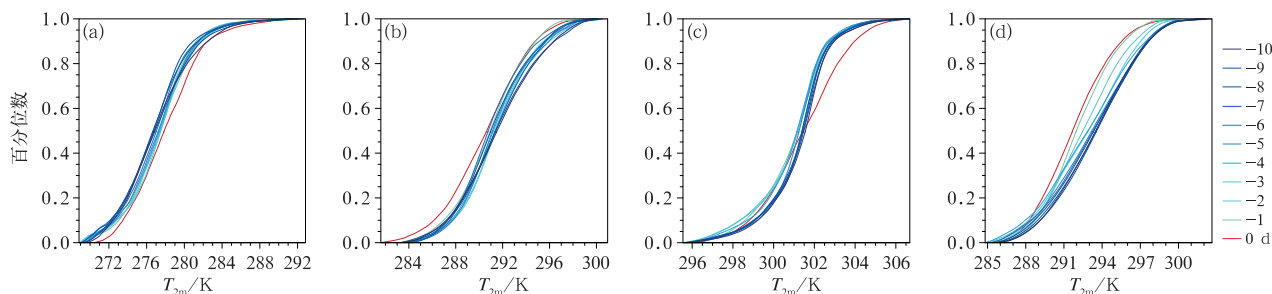


图 4 长沙站点(a)1月、(b)4月、(c)7月、(d)10月模式(分析场)平均气候百分位分布

Fig. 4 Averaged forecasted (analysis) climate percentile distribution in (a) January, (b) April, (c) July and (d) October at Changsha Station

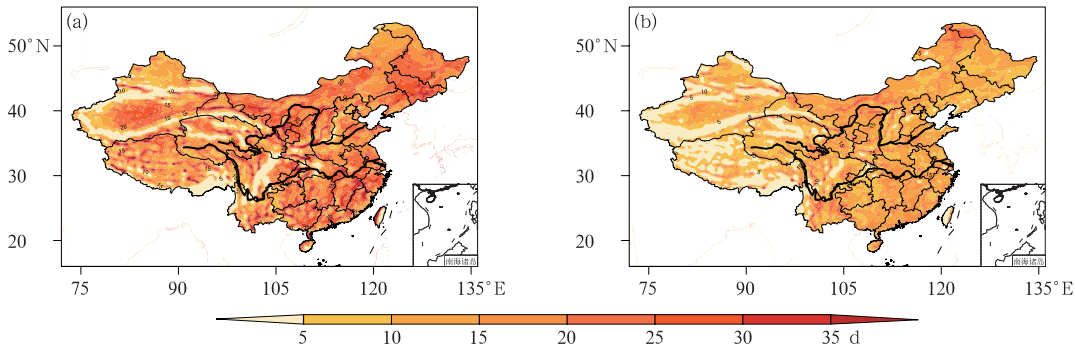


图 5 2016 年极端高温(a)和极端低温(b)日数分布图

Fig. 5 Days of extreme high temperature (a) and extreme low temperature (b) in 2016

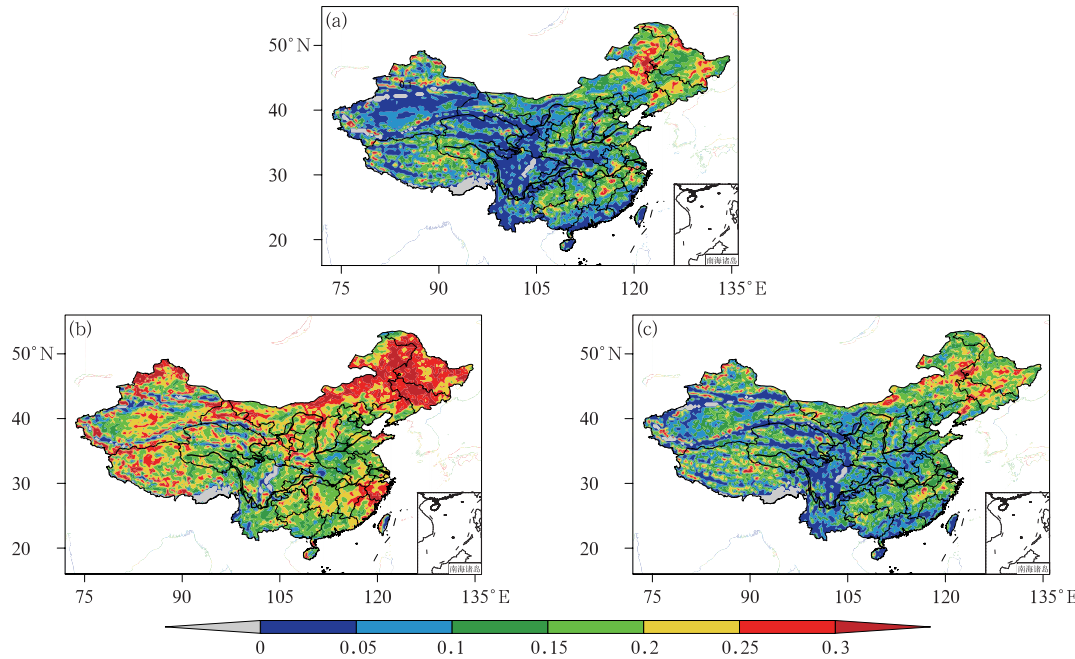


图 6 T639 集合系统对我国 2016 年极端高温 72 h 预报的 TS 评分

(a) 集合均值法(EM), (b) 最大值法(Max), (c) 众数法(Mode)

Fig. 6 TS scores of 72 h forecasted extreme high temperature in 2016 by T639 ensemble system

(a) ensemble mean, (b) maximum, (c) mode

能有效提高预报技巧,尤其在东北、西北和江南地区。集合成员众数与集合成员平均值相比,TS 评分在西北地区略有提升,但全国平均而言提高程度不强。

结合 T639 集合成员均值、最大值和众数对极端高温预报检验结果全国平均统计(表 2)可看出,对于中国极端高温的预报,集合成员最大值对极端信号有放大效果,命中率最高,其次为集合众数,集合均值对极端信号有平滑效果,其命中率低于集合众数。就误报率而言,集合成员最大值为最高,其次

表 2 2016 年 T639 集合成员均值、最大值和众数三种方法对极端高温 72 h 预报检验结果全国平均统计表

Table 2 National average scores of 72 h forecasted extreme high temperature in 2016 using the ensemble mean, maximum and mode of T639 ensemble system

	R_H	R_{FA}	TS
EM	0.1295	0.0142	0.0968
Max	0.6037	0.1211	0.2037
Mode	0.2258	0.0399	0.1271

为集合众数,集合均值的误报率最低。综合而言,从 TS 评分来看,T639 集合预报系统对我国极端高温的预报有一定的预报性能,与集合均值法相比,集合众数法的技巧略有提升,但提升不明显,采用集合成员相对的极值的最大值法预报技巧增高明显。

图 7 为 T639 集合预报系统对我国 2016 年极端高温预报的年平均 TS 评分分布,红色、蓝色和紫色的线条分别代表利用集合众数法、集合最大值法和集合均值法的 TS 评分。对于我国 2016 年极端高温的预报而言,在不同的预报时效下采取集合最大值法进行预报的 TS 评分都最大,其次是集合众数法,集合均值对我国 2016 年极端高温预报的 TS 评分最小,预报时效为 10 d 时,集合均值对我国 2016 年极端高温预报 TS 评分几乎接近 0,预报技巧较低。

图 8 为龙泉站 2016 年分析场识别的和模式 72 h 预报的极端高温。图中虚线/点线为逐日历史(模式)气候态的第 95/5 个百分位数,黑色实线为逐日分析场(72 h 预报)对应的 2 m 温度值。我们定义气候序列的第 95 个百分位或者以上的 2 m 温度值为极端高温,即实线超出虚线则为极端高温。由分析场识别的 2016 年极端高温发生日数为 46, T639 集合系统 72 h 预报的集合均值法预报的日数为 6,集合成员最大值法为 82,集合成员众数法为 15,从对极端高温预报的日数可看出集合均值对极

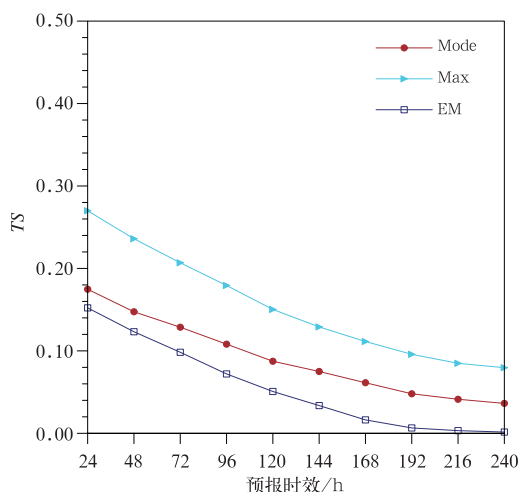


图 7 T639 集合系统对我国 2016 年极端高温预报的年平均 TS 评分

Fig. 7 Annual-mean TS scores of forecasted extreme high temperature in 2016 by T639 ensemble system

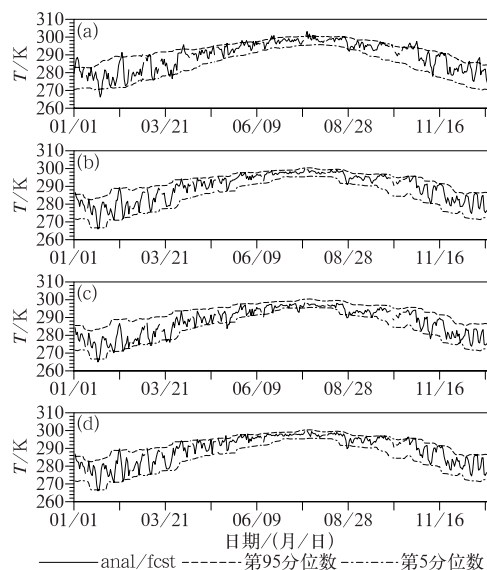


图 8 龙泉站点 2016 年分析场识别的和模式 72 h 预报的极端高温

(a) 分析场实况 (Anal), (b) 集合均值法 (EM), (c) 最大值法 (Max), (d) 众数法 (Mode)

Fig. 8 Extreme high temperatures at Longquan Station identified by analysis and 72 h forecast in 2016

(a) analysis, (b) ensemble Mean, (c) maximum, (d) mode

端信号有平滑效果,对最大值有放大效果,而集合众数对极端高温预报的日数与分析场识别的日数差别最小。另外,从图中可看出龙泉站点暖季 2 m 温度变化幅度较小,冷季 2 m 温度的变化幅度较大,冷暖季温度变化幅度的这种差异会使模式对极端预报的准确度在冷暖季造成一定的差异。

结合上面极端高温的预报检验结果可看出,集合均值的假警报率较低,但由于命中率也较低,所以 TS 评分与集合最大值和集合众数相比最低;集合最大值的命中率最高,假警报率也随之有明显的增高。总体而言,TS 风险评分与集合均值和集合众数法相比为最高;集合成员众数的命中率与集合均值相比,误报率较集合最大值低,其 TS 风险评分与集合均值相比为有所提高,但提高不明显。

3.3 2016 年极端低温的预报检验

图 9 为 T639 集合系统对 2016 年极端低温 72 h 预报的 TS 评分全国分布,分析显示:与极端高温相比,T639 集合系统对我国极端低温的预报性能较高。整体而言,TS 评分在我国华北和南方地区较

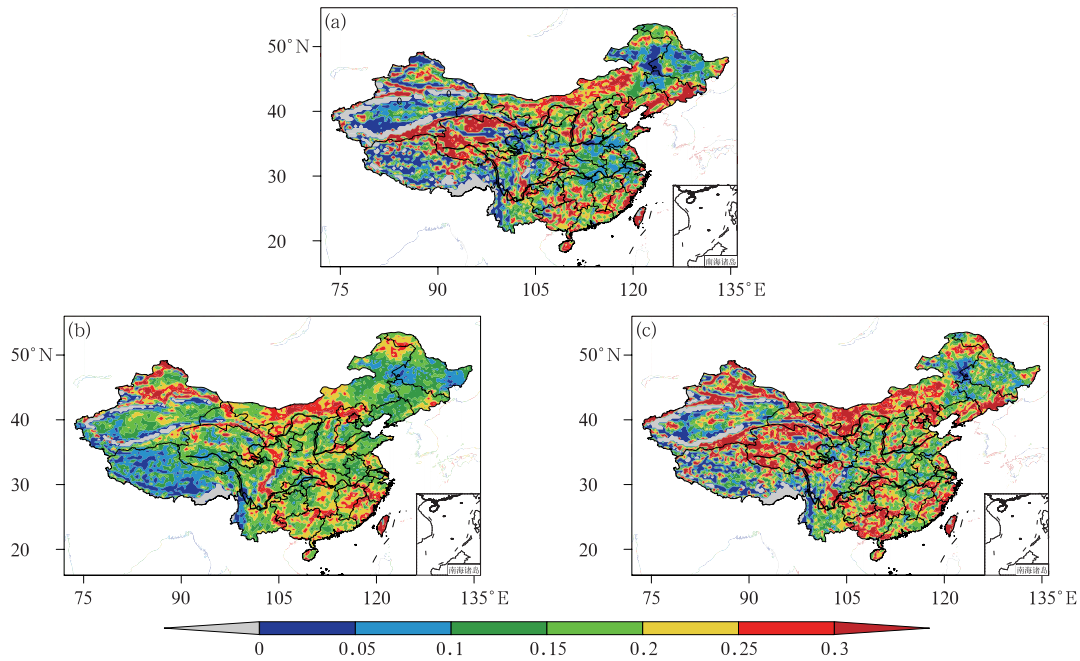


图 9 同图 6, 但为极端低温
Fig. 9 Same as Fig. 6, but for extreme low temperature

高。集合成员最小值的预报技巧与集合成员均值和众数相比较差, 仅在江淮地区略有提高, 集合成员众数与集合成员平均值相比, 72 h 预报的 TS 评分整体而言有所提高, 主要在新疆、内蒙古和广西等地提高明显。

结合 T639 集合成员均值、最小值和众数对极端低温 72 h 预报检验结果全国平均统计(表 3)可看出, 对于中国的极端低温的预报, 集合成员最小值对极端信号有放大效果, 命中率最高, 其次为集合众数, 集合均值对极端信号有平滑效果, 其命中率低于集合众数。就误报率而言, 集合成员最小值为最高, 其次为集合众数, 集合均值的误报率较低。综合来看, 就 TS 评分而言, T639 集合预报系统对我国极端低温 72 h 的预报具有一定的预报性能, 采用集合成员相对的极值的最小值法预报技巧最低, 与集合均值法相比, 集合众数法的技巧提升较明显。

表 3 同表 2, 但为极端低温 72 h 预报
检验结果全国平均统计表

Table 3 Same as Table 2, but for national average scores of 72 h forecasted extreme low temperature

	R_H	R_{FA}	TS
EM	0.2043	0.0041	0.1748
Min	0.6625	0.1025	0.1545
Mode	0.2741	0.0123	0.1907

图 10 为 T639 集合预报系统对我国 2016 年极端低温预报的年平均 TS 评分分布。由图可见, 在 1~8 d 预报时效下采取集合众数法进行预报的 TS 评分都最高, 在 9~10 d 预报时效下采取集合最小值法 TS 评分最高。集合均值对我国极端低温预报随预报时效的增加下降速度最快, 在预报时效为 10 d 时集合均值预报的 TS 评分接近 0, 预报效果较差。

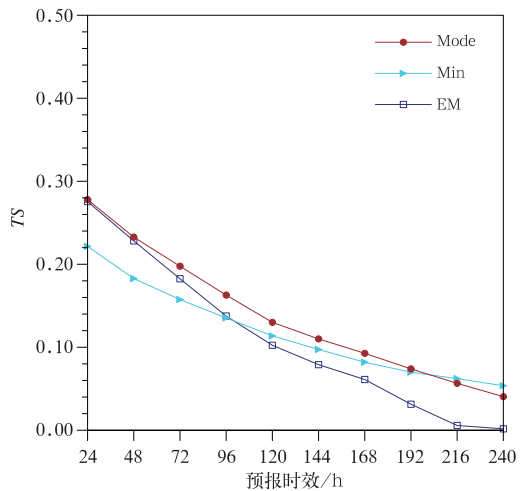


图 10 同图 7, 但为极端低温预报
Fig. 10 Same as Fig. 7, but for extreme low temperature

图 11 为北京站 2016 年分析场识别的和模式 72 h 预报的极端低温。图中虚线/点线为北京站点 2016 年逐日历史(模式)气候态的第 95/5 个百分位数,实线为逐日分析场(72 h 预报)对应的 2 m 温度值。我们定义气候序列的第 5 个百分位或者以下的 2 m 温度值为极端低温,即实线低于点线则为极端低温。由分析场识别的 2016 年极端低温的发生日数为 10, T639 集合系统 72 h 预报的集合均值法预报的日数为 2, 集合成员最小值预报的为 38, 集合成员众数的为 5, 从对北京站点极端低温识别的日数可看出集合均值对极端信号有平滑效果, 最小值有放大效果, 而集合众数对极端低温预报的日数与分析场识别的差别最小。

结合 2016 年 T639 数值集合预报模式中, 集合成员均值、最小值和众数对北京站点极端低温的 72 h 预报, 及其命中率、假警报率和 TS 风险评分的检验结果统计(表略)可看出, 集合均值的假警报率最低, 但命中率也为最低, 所以 TS 评分比集合成员最小值和众数低; 而集合成员最小值的命中率高, 但假警报率也随之升高明显, 所以, TS 评分提高不明显; 集合成员众数的命中率与均值相比高, 误报率也有所升高, 但升高不明显, 所以 TS 风险评分与集合均值和集合最小值相比为最高。

综上所述, T639 集合系统对于我国极端温度具有一定的预报性能, 预报技巧在各个区域有较大差

别, 极端高温在江南和东北区域技巧相对较高、极端低温在华北和南方技巧较高。对于不同预报方法技巧对比显示, 集合均值法对极端温度的极端信号有平滑效果, 总体技巧偏低, 集合众数法技巧略有提升, 而采用集合成员相对的极值的最大值或最小值法对极端信号则有放大效果, 对高温技巧增加明显, 对低温技巧增加则不明显, 但就极端温度发生的日数而言, 集合众数的预报效果是最好的。由此可见, 采用合适的方法对集合预报中极端信息的正确提取至关重要。

选取 2016 年 1 月 21 日影响我国的极端低温作为个例(图 12), 图中阴影部分为分析场识别的低温范围, 不同的颜色代表 T639 集合预报各成员预报此次极端低温的提前天数, 可以看出, 不同成员对极端低温预报的提前天数和预报范围各不相同, 大部分都能提前 3 d 以上预报出此次极端低温, 部分成员对部分区域的预报甚至能达到 8 d, 各成员预报的极端低温发生范围也较大。然而, 采用集合平均进行预报时, 预报出来的提前天数下降明显, 并且预报的极端低温范围和各集合成员相比缩小明显。可看出对于极端温度的预报, 各集合成员预报的提前天数都较长, 预报的范围也较大, 但是采用集合平均对极端温度进行预报则会较大地缩短提前预报天数和缩小极端温度发生范围, 因此, 采用适合的方法提取集合预报中的极端信息尤为重要。

4 结 论

本文基于我国 T639 集合预报系统的模式实时预报资料, 对历史气候百分位和模式气候百分位进行估算, 并对 2016 年我国极端高温和极端低温进行了检验评估。利用模式气候态第 95 个百分位数(第 5 个百分位数)作为阈值对极端高温(低温)进行预报, 并且利用历史气候百分位识别的极端高温(低温)对预报的极端温度进行识别、预报和检验, 得到以下结论:

(1) 对于不同格点, 气候百分位分布存在差异, 不同超前时间对应的气候百分位分布变化趋势均与分析场的基本一致, 超前时间越短, 模式气候百分位分布曲线越靠近分析场气候百分位分布曲线。不同代表站点, 不同季节气候百分位分布存在差异, 就长沙站点而言, 冷暖季差异较小, 转换季各预报时效的气候百分位分布差异较大。

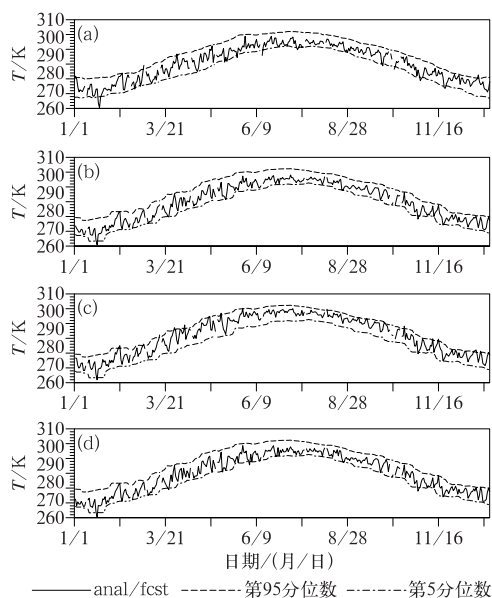


图 11 同图 8, 但为北京站点的极端低温
Fig. 11 Same as Fig. 8, but for extreme low temperature at Beijing Station

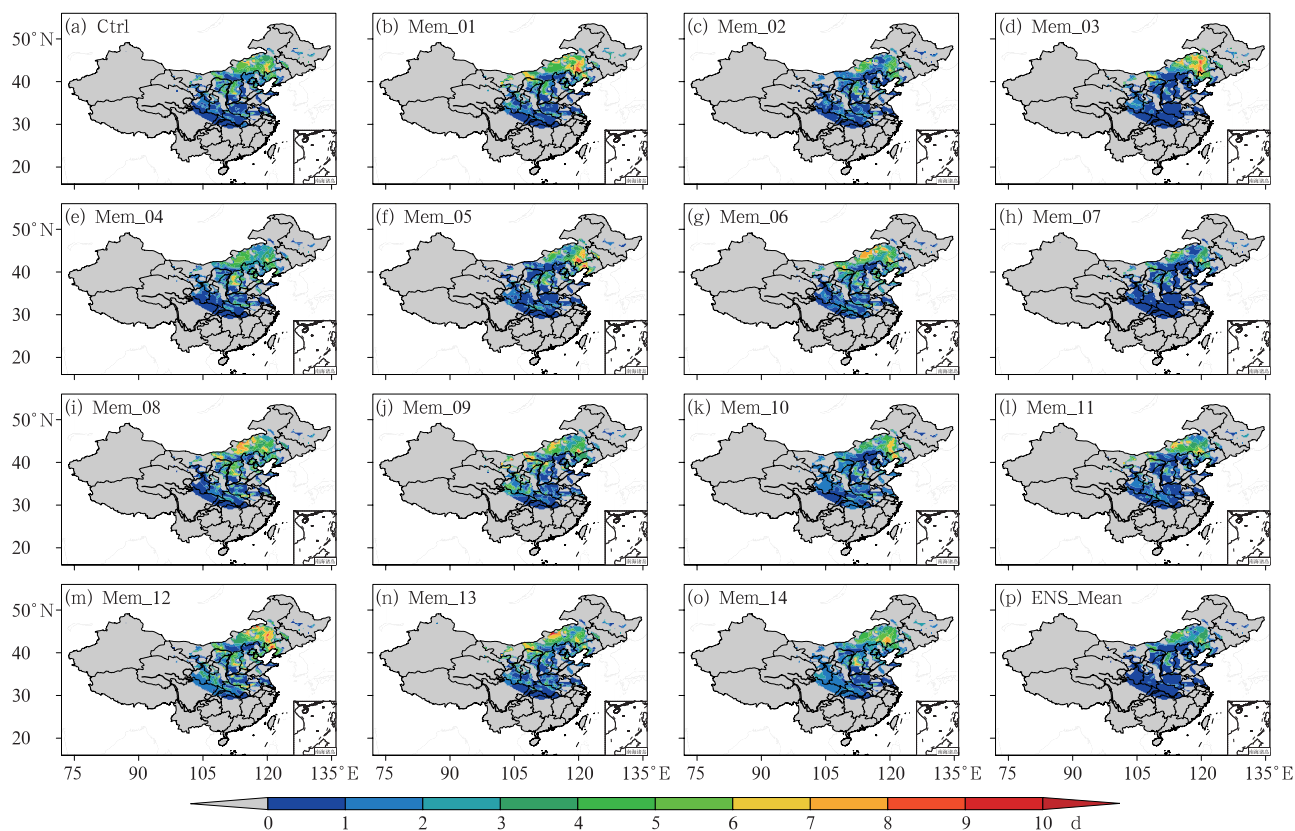


图 12 2016 年 1 月 21 日极端低温 T639 集合系统不同成员 1~15(a~o)和集合平均(p)预报的提前日数图

Fig. 12 Lead days of forecasted extreme cold on 21 January 2016 by T639 ensemble member 1–15 (a–o) and ensemble mean (p) respectively

(2) T639 集合预报系统 2 m 温度的 Talagrand 分布检验结果显示,哈尔滨和长沙站点存在不同程度的暖偏差;而北京和拉萨站点则存在冷偏差;其余典型站表现出离散度偏小的特点。

(3) T639 集合预报系统对我国极端高温具有一定的预报技巧,在江南、东北区域和高原区域技巧相对较高。从全国平均 TS 评分来看,利用集合成员最大值法预报的技巧最高。针对所选取的龙泉市代表站点的极端高温分析可看出,集合平均对极端信号有平滑效果,集合成员极值对极端信号则有放大效果,就极端发生的日数的预报而言,集合众数的预报效果是最好的,但命中率不高,所以 TS 评分提高不明显,从 TS 评分来看,利用集合成员最大值法的预报技巧最高。

(4) T639 集合预报系统对极端低温的预报技巧在我国华北和南方区域相对较高。利用集合成员众数法预报的全国平均 TS 评分最高。针对典型代表站点的极端低温分析可看出,集合均值对极端信号有平滑效果,集合成员极值对极端信号则有放大

效果。对北京站点的极端低温预报日数而言,集合众数的预报效果是最好的,集合众数法预报的 TS 评分也为最高。

本文研究表明,T639 集合系统对于我国极端温度总体上具有一定预报性能,但预报技巧在各个区域有较大差别。从全国平均 TS 评分而言,集合最大值法对极端高温技巧增加明显,集合众数法对低温技巧增加明显。但是,由于目前该系统集合预报的扰动方法难以形成充分离散化的集合成员扰动、而且集合成员数太少也可能对极端温度的预报产生影响。另外,对其他气象要素(降水、风)不同集合方法的预报效果,还需要在今后的工作中进行单独研究。目前只利用集合均值法、集合最大值或最小值法和集合众数法对极端温度的预报进行检验,采用何种合适的方法对集合预报中极端信息的正确提取还有待研究。

参考文献

包慧濛,郭达烽,肖潇,2015. ECMWF 集合预报统计量在江西降水预

- 报中的检验分析[J]. 气象与减灾研究, 38(3): 60-67. Bao H M, Guo D F, Xiao X, 2015. Validation of ECMWF ensemble forecast statistics in Jiangxi precipitation forecast[J]. Meteor Disaster Reduct Res, 38(3): 60-67(in Chinese).
- 董全, 代刊, 陶亦为, 等, 2017. 基于 ECMWF 集合预报的极端天气预报产品应用和检验[J]. 气象, 43(9): 1095-1109. Dong Q, Dai K, Tao Y W, et al, 2017. Application and verification of extreme weather forecast products of ECMWF ensemble prediction system[J]. Meteor Mon, 43(9): 1095-1109(in Chinese).
- 杜钧, 2002. 集合预报的现状和前景[J]. 应用气象学报, 13(1): 16-28. Du J, 2002. Present situation and prospects of ensemble numerical prediction[J]. J Appl Meteor Sci, 13(1): 16-28(in Chinese).
- 段明铨, 王盘兴, 李驰钦, 2012. NCEP 集合预报系统在亚欧和北美区域的预报效果对比[J]. 大气科学学报, 35(6): 641-651. Duan M K, Wang P X, Li C Q, 2012. Comparisons on forecast performance of NCEP ensemble prediction system between Eurasia and North America regions[J]. Trans Atmos Sci, 35(6): 641-651(in Chinese).
- 段明铨, 王盘兴, 吴洪宝, 等, 2009. 夏季亚欧中高纬度环流的集合预报效果检验[J]. 应用气象学报, 20(1): 56-61. Duan M K, Wang P X, Wu H B, et al, 2009. The ensemble forecasting verification on the summer Eurasian middle-high latitude circulation[J]. J Appl Meteor Sci, 20(1): 56-61(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).
- 关吉平, 张立凤, 张铭, 2006. 集合预报研究现状与展望[J]. 气象科学, 26(2): 228-235. Guan J P, Zhang L F, Zhang M, 2006. Review of the ensemble forecast[J]. Sci Meteor Sin, 26(2): 228-235(in Chinese).
- 胡海川, 黄彬, 魏晓琳, 2017. 我国近海洋面 10 m 风速集合预报客观订正方法[J]. 气象, 43(7): 856-862. Hu H C, Huang B, Wei X L, 2017. Objective correction method of ensemble forecast of 10 m winds on Chinese offshore[J]. Meteor Mon, 43(7): 856-862(in Chinese).
- 皇甫雪官, 2002. 国家气象中心集合数值预报检验评价[J]. 应用气象学报, 13(1): 29-36. Huangfu X G, 2002. The verification for ensemble prediction system of National Meteorological Center[J]. J Appl Meteor Sci, 13(1): 29-36(in Chinese).
- 黄刚, 张立凤, 王秋良, 等, 2012. 基于集合预报产品的降尺度降水预报试验[J]. 气象科学, 32(5): 508-514. Huang G, Zhang L F, Wang Q L, et al, 2012. Precipitation forecasting experiments based on the products of ensemble forecast by using downscaling method[J]. J Meteor Sci, 32(5): 508-514(in Chinese).
- 李俊, 杜钧, 刘羽, 2015. 北京“7. 21”特大暴雨不同集合预报方案的对比试验[J]. 气象学报, 73(1): 50-71. Li J, Du J, Liu Y, 2015. A comparison of initial condition-, multi-physics-and stochastic physics-based ensembles in predicting Beijing “7. 21” excessive storm rain event[J]. Acta Meteor Sin, 73(1): 50-71(in Chinese).
- 罗亚丽, 2012. 极端天气和气候事件的变化[J]. 气候变化研究进展, 8(2): 90-98. Luo Y L, 2012. Changes in weather and climate extremes[J]. Adv Climate Change Res, 8(2): 90-98(in Chinese).
- 彭勇, 王萍, 徐炜, 等, 2002. 气象集合预报的研究进展[J]. 南水北调与水利科技, 10(4): 90-97. Peng Y, Wang P, Xu W, et al, 2002. Review of weather ensemble prediction[J]. South-to-North Water Transf Water Sci Technol, 10(4): 90-97(in Chinese).
- 任国玉, 初子莹, 周雅清, 等, 2005. 中国气温变化研究最新进展[J]. 气候与环境研究, 10(4): 701-716. Ren G Y, Chu Z Y, Zhou Y Q, et al, 2005. Recent progresses in studies of regional temperature changes in China[J]. Climatic Environ Res, 10(4): 701-716(in Chinese).
- 王岱, 游庆龙, 江志红, 等, 2016. 基于均一化资料的中国极端地面气温变化分析[J]. 高原气象, 35(5): 1352-1363. Wang D, You Q L, Jiang Z H, et al, 2016. Analysis of extreme temperature changes in China based on the homogeneity-adjusted data[J]. Plat Meteor, 35(5): 1352-1363(in Chinese).
- 吴剑坤, 高丽, 乔林, 等, 2015. 基于 T213 集合预报的中国极端温度预报方法研究[J]. 气象科学, 35(4): 438-444. Wu J K, Gao L, Qiao L, et al, 2015. Research on Chinese extreme temperature forecasting method based on T213 ensemble forecast[J]. J Meteor Sci, 35(4): 438-444(in Chinese).
- 吴岩, 张惠君, 关铭, 等, 2014. 两次暴雨过程的主观及集合预报检验[J]. 黑龙江气象, 31(4): 1-3, 6. Wu Y, Zhang H J, Guan M, et al, 2014. Subjective and ensemble prediction test of twice rainstorm[J]. Heilongjiang Meteor, 31(4): 1-3, 6(in Chinese).
- 夏凡, 陈静, 2012. 基于 T213 集合预报的极端天气预报指数及温度预报应用试验[J]. 气象, 38(12): 1492-1501. Xia F, Chen J, 2012. The research of extreme forecast index based on the T213 ensemble forecast and the experiment in predicting temperature[J]. Meteor Mon, 38(12): 1492-1501(in Chinese).
- 张立凤, 罗雨, 2010. 初始场对暴雨数值预报的影响及集合预报试验[J]. 气象科学, 30(5): 650-656. Zhang L F, Luo Y, 2010. The effect of initial condition on numerical precipitation prediction and ensemble forecast[J]. Sci Meteor Sin, 30(5): 650-656(in Chinese).
- Beniston M, Stephenson D B, Christensen O B, et al, 2007. Future extreme events in European climate: an exploration of regional climate model projections[J]. Climatic Change, 81(S1): 71-95.
- Easterling D R, Evans J L, Groisman P Y, 2000. Observed variability and trends in extreme climate events: a brief review[J]. Bull Amer Meteor Soc, 81(3): 417-426.
- Epstein E S, 1969. Stochastic dynamic prediction[J]. Tellus, 21(6): 739-759.
- Ghelli A, Lalaurette F, 2000. Verifying precipitation forecasts using up-scaled observations[J]. ECMWF Newsl, 87: 9-17.
- Guan H, Zhu Y J, 2017. Development of verification methodology for extreme weather forecasts[J]. Wea Forecasting, 32(2): 479-491.
- Houghton J T, Ding Y H, Griggs D J, et al, 2001. Climate change

- 2001; the scientific basis. Contribution of working group 1 to the third assessment report of the intergovernmental panel on climate change[R]. Cambridge: Cambridge University Press.
- Karl T R, Knight R W, 1998. Secular trends of precipitation amount, frequency, and intensity in the United States[J]. *Bull Amer Meteor Soc*, 79(2): 231-241.
- Lalaurette F, 2002. Early detection of abnormal weather conditions using a probabilistic extreme forecast index[J]. *Quart J Roy Meteor Soc*, 129(594): 3037-3057.
- Lalaurette F, 2003. Two proposals to enhance the EFI response near the tails of the climate distribution[EB/OL]. <http://www.ecmwf.int/products/forecasts/efi-guide.pdf>.
- Lavers D A, Zsoter E, Richardson D S, et al, 2017. An assessment of the ECMWF extreme forecast index for water vapor transport during boreal winter[J]. *Wea Forecasting*, 32(4): 1667-1674.
- Leith C E, 1974. Theoretical skill of Monte Carlo forecasts[J]. *Mon Wea Rev*, 102(6): 409-418.
- Meehl G A, Karl T, Easterling D R, et al, 2000. An introduction to trends in extreme weather and climate events: observations, socioeconomic impacts, terrestrial ecological impacts, and model projections[J]. *Bull Amer Meteor Soc*, 81(3): 413-416.
- Palmer T N, Branković Č, Richardson D S, 2000. A probability and decision-model analysis of PROVOST seasonal multi-model ensemble integrations[J]. *Quart J Roy Meteor Soc*, 126(567): 2013-2033.
- Richardson D S, 2000. Skill and relative economic value of the ECMWF ensemble prediction system[J]. *Quart J Roy Meteor Soc*, 126(563): 649-667.
- Solomon M S, Qin D H, Manning M, et al, 2007. Climate change 2007: The physical science basis. Contribution of working group 1 to the fourth assessment report of the intergovernmental panel on climate change[R]. Cambridge, United Kingdom and New York, USA: Cambridge University Press.
- Zsótér E, 2006. Recent developments in extreme weather forecasting [J]. *ECMWF Newsl*, 107: 8-17.