

潘欣,马依依,毛程燕,等.2023.基于聚类分析的浙江省汛期OCF降水预报分区订正试验[J].暴雨灾害,42(6):716-723. PAN Xin, MA Yiyi, MAO Chengyan, et al. 2023. Regional correction calibration for OCF precipitation in flood season in Zhejiang Province based on cluster analysis [J]. Torrential Rain and Disasters, 42(6):716-723 (in Chinese). doi:10.12406/byzh.2023-042

## 基于聚类分析的浙江省汛期OCF降水预报分区订正试验

潘欣,马依依,毛程燕,郑倩

(浙江省衢州市气象局,衢州 324000)

**摘要:**汛期暴雨预报一直是气象预报业务工作中的重点及难点。首先,利用浙江省2227个气象站2016—2021年每年汛期(4—10月)逐日降水资料,以欧式距离作为相似度量标准,通过Kmeans聚类算法对浙江汛期降水进行分区;然后,对偏差订正法进行时空上的改进;最后,将改进后的偏差订正法与分区结合形成分区订正方法。以未与分区结合的全区订正作为对比,对浙江多模式客观集成预报(OCF)进行分区订正和检验。结果显示:(1)将浙江省划分为7个降水相似区,其结果呈现明显的区域特征,并与浙江省地形地貌密切相关。(2)经2021年汛期检验,相较于OCF预报,分区订正预报优于全区订正预报,其优势主要体现在能有效降低晴雨预报的降水空报率和大幅提升大雨以上量级降水的命中率,特别是暴雨以上量级预报命中率由0.25提升至0.41。(3)典型过程检验表明,对于系统性降水和对流性降水,分区订正预报均能改善暴雨以上量级降水强度和落区预报。特别是系统性降水,分区订正预报效果更明显,能预报出大暴雨以上强降水。

**关键词:**聚类分析;降水相似区;分区订正;汛期

中图法分类号:P456.5

文献标志码:A

DOI:10.12406/byzh.2023-042

## Regional correction calibration for OCF precipitation in flood season in Zhejiang Province based on cluster analysis

PAN Xin, MA Yiyi, MAO Chengyan, ZHENG Qian

(Quzhou Meteorological Administration of Zhejiang Province, Quzhou 324000)

**Abstract:** The forecast of rainstorm in flood season has always been the key and difficult point in the meteorological forecasting operation. First, we used the daily precipitation data from 2227 meteorological stations in Zhejiang Province from 2016 to 2021 during the flood season (April to October), and divided the precipitation region by applying the K-means clustering algorithm, which employed the Euclidean distance as the similarity measure. Then, the regional correction method is formed by combining the spatial-temporally improved bias correction method and divided regions. Finally, we applied this method to perform the regional correction and validation on the Zhejiang Multi-Model Objective Consensus Forecast (OCF), compared with the overall correction not combined with divided regions. The results are as follows. (1) The K-means clustering algorithm can divide Zhejiang Province into 7 precipitation-similar regions, which show distinct regional characteristics closely related to the topographic features of Zhejiang Province. (2) According to validation during the 2021 flood season, the regional correction performed better than the overall correction in the OCF forecasts. Its main advantages lie in effectively reducing the false alarm of precipitation in the clear-rain forecast and substantially increasing the hit rate (POD) for the heavy rain and above, especially for the rainstorm and above from 0.25 to 0.41. (3) The typical validation shows that, for both systematic and convective precipitation, the regional correction can significantly improve the intensity and falling area of precipitation for the rainstorm and above. Especially for the systematic precipitation, the regional correction demonstrated more remarkable effects, which can predict heavy rainstorms.

**Key words:** cluster analysis; precipitation-similar region; regional correction; flood season

## 引言

随着数值预报技术的迅速发展,数值预报产品

的精度、准确度不断提升,其已逐渐成为各级气象台站预报员制作天气预报的重要依据。然而,由于数值模式初值的不确定性和模式本身存在的系统

收稿日期:2023-03-07;定稿日期:2023-12-19

资助项目:浙江省气象局科技计划重点项目(2022ZD29);浙江省气象局青年科技基金项目(2020QN34)

第一作者:潘欣,主要从事天气预报及技术方法研究。E-mail: panxin1008@163.com

偏差,模式预报往往与实况存在偏差(穆穆等,2011;武英娇等,2019;赵宁坤等,2021)。因此,采用各种释用方法对模式预报产品进行后处理以改进预报效果,已成为目前解决模式预报偏差的常用手段。

降水量是气象预报中重要的预报要素之一,由于其本身的非正态性和不连续性,预报难度较大,因此研究模式降水预报的订正释用方法是气象预报关键技术研究中的重要课题。李俊等(2014)根据模式降水预报的偏差特征,采用基于频率(或面积)匹配的偏差订正方法显著改善了模式降水预报中雨量和雨区范围的系统性偏差。曹萍萍等(2018)基于地形地理分区,采用概率匹配方法有效订正了西南地区的逐12 h降水预报。赵瑞霞等(2020)将最优评分订正(Optimal Threat Score, OTS)、模式输出统计(Model Output Statistics, MOS)和纳入超前空间实况信息的模式输出统计(MOS with prior-spatial observation predictors, OMOS)方法及其优化组合应用于逐3 h降水预报中,结果表明MOS-OTS综合预报性能最好。曹萍萍等(2020)基于“配料法”有效订正了四川夏季的大雨、暴雨预报。从靖等(2021)将Gamma累积概率分布曲线应用于预报-实况概率匹配方法中,并以此改善了海河流域东北冷涡背景下ECMWF小雨、大雨和暴雨预报技巧。

浙江汛期多暴雨过程、强度较大,模式对其预报偏差大,其精准预报一直是气象预报中的难点和重点。研究表明,频率匹配法能够有效利用历史降水资料对模式降水预报进行有效订正(智协飞和吕游,2019),此法是一种简便且实用的降水订正方法。然而,由于浙江地形复杂多样,呈“七山一水二分田”特征,同时降水天气气候背景复杂,浙西为梅汛期暴雨区,浙东为台汛期暴雨区。如果将全省作为整体进行频率匹配,显然缺乏合理性,而以降水分区为基础进行频率匹配降水订正更为合适。目前常用的降水分区方式是根据地形等地理信息进行人为主观分区(曹萍萍等,2018;智协飞和吕游,2019)。而为了保证降水分区的客观性和合理性,本文聚类分区方式采用Kmeans聚类算法(王伟等,2011;苟浩锋等,2020),其被广泛应用于数据分析、数据挖掘、模式识别等领域,是在科研和工业中应用最广泛的聚类算法之一。本文旨在通过聚类分析,对复杂地形地区的降水进行合理分区,在此基础上,对李俊等(2014)基于频率(或面积)匹配的偏差订正方法进行改进,独创基于聚类分析的偏差订正方法。利用该方法,对数值模式降水预报进行分区订正,以期提升降水(特别是暴雨以上)预报能力,为汛期防灾减灾提供参考。

## 1 资料与方法

### 1.1 资料说明

#### 1.1.1 降水实况资料

降水实况资料为2016—2021年每年汛期(4月15日—10月15日)浙江省2 227个气象站(包括国家气象站和区域自动气象站,其站点分布见图1)逐日(以20:00为日界,北京时,下同)降水资料,该资料来源于浙江省气象信息中心,对其中的缺测样本已剔除。

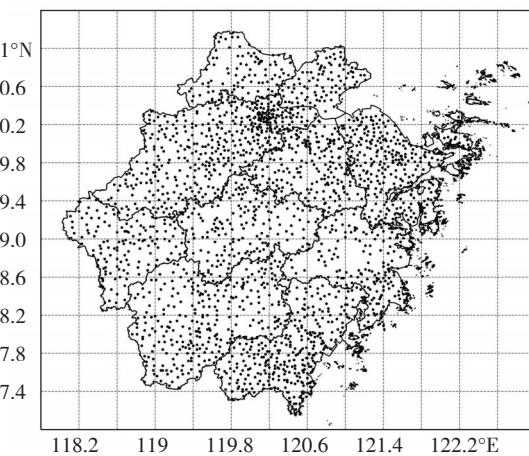


图1 浙江省2 227个气象站(黑色点)空间分布

Fig.1 Spatial distribution of 2227 automatic meteorological stations (black dots) in Zhejiang Province

#### 1.1.2 降水预报资料

降水预报资料为2017—2021年每年4—10月浙江省气象台下发的多模式客观集成预报(Objective Consensus Forecasting, OCF)降水资料(空间分辨率 $0.05^{\circ} \times 0.05^{\circ}$ )。考虑到日常预报业务中的实时性与实用性,选取每日08:00起报的未来12—36 h的降水预报值作为后一日的降水预报值(该预报值与以20:00为日界的日降水量实况值匹配)。采用最近邻法根据降水预报格点资料得到各个气象站点的降水预报资料。

### 1.2 研究方法

首先,选取2016—2020年汛期浙江省2 227个气象站逐日降水实况数据作为训练数据,采用Kmeans聚类算法,根据站点日降水序列间的相似性对所有站点进行分区。然后,对偏差订正进行时空上的改进。将改进后偏差订正法与分区结合形成分区订正方法。最后,以2021年汛期降水预报数据及降水实况数据为例,对OCF降水预报进行分区订正(如下文1.2.2所述,分区计算降水频率,分区订正降水),对分区订正和全区订正(即不对站点进行分区,而是全区统一订正)的预报结果进行对比检验。

#### 1.2.1 Kmeans聚类算法

Kmeans聚类算法作为一种无监督聚类算法,广泛

应用于科研与工业,特别是数据分析和数据挖掘方面(王伟等,2011;苟浩锋等,2020)。Kmeans聚类算法分类原理是将 $n$ 个样本划分为 $k$ 个簇,各个簇内的样本具有高度相似性而各个簇的样本间具有较大差异性。相似度的度量可以根据不同的需求采用不同度量方式。Kmeans聚类算法优点是不需要通过任何先验知识或标签,仅仅需要数据本身就可以对数据进行合理分类。虽然算法开始要求指定簇数,但可以通过尝试不同簇数,根据肘部法则和平均轮廓系数来确定聚类簇数,提升聚类合理性。

### 1.2.2 基于频率匹配的偏差订正法

采用基于频率匹配的偏差订正法,在原有的预报偏差订正法(李俊等,2014)基础上做时间和空间上的改进。Kmeans聚类算法中频率匹配方面和原预报偏差订正法一致,通过预报降水累积频率和实况降水累积频率的匹配来订正预报降水量。考虑到数据年限较短,为提升降水频率计算的合理性,根据Kmeans聚类降水分区结果,将原有的点对点或点对全区频率匹配改为点对区域匹配,将原本通过全区或单点预报/实况降水数据来计算得到的全区或单点预报/实况降水频率修改为各个区域所有站点集合计数得到区域预报/实况降水频率(李俊等,2014;曹萍萍等,2018)。为了进一步扩充数据,将过去20 d滑动平均法计算平均频率修改为过去、未来 $n$ 天滑动平均(准对称滑动训练期, $n$ 根据训练数据试验选取),并包含过去几年同时段数据计算平均频率,以便充分利用历史数据,提高对于极值降水的预测能力。如图2所示,准对称滑动训练期采用的数据包括当年当日及之前 $n$ 天的预报、观测资料和过去1年到 $i$ 年的当日及前后 $n$ 天的预报、观测资料( $i$ 为历史预报和观测资料年限)。为了提升降水订正效果,根据训练数据试验以降水检验评分指标选定 $n$ ( $n$ 的取值范围在5~15之间)。

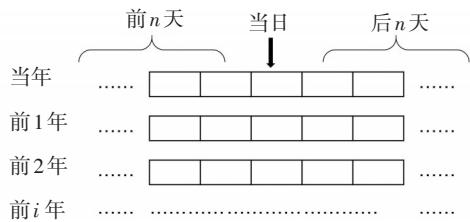


图2 准对称滑动训练期

Fig. 2 Quasi-symmetric sliding training period

### 1.2.3 降水检验方法

本文采用点对点的方法,对降水量的预报值与实况值之间的误差进行检验(赵滨和张博,2018)。检验指标包括TS评分、ETS评分、命中率POD、BIAS偏差评分。

TS(技巧评分, Threat Score, 记为 $T_S$ ), 表示在预报区域内满足某降水阈值的降水预报技巧。

$$T_S = \frac{T_p}{T_p + F_p + F_N} \quad (1)$$

ETS(公平技巧评分, Equitable Threat Score, 记为 $E_{TS}$ ), 表示在预报区域内满足某降水阈值的降水预报结果相对于满足同样降水阈值的随机预报的预报技巧。

$$E_{TS} = \frac{T_p - \frac{(T_p + F_p) \times (T_p + F_N)}{T_p + F_p + F_N + T_N}}{T_p + F_p + F_N - \frac{(T_p + F_p) \times (T_p + F_N)}{T_p + F_p + F_N + T_N}} \quad (2)$$

POD(命中率, Probability of Detection, 记为 $P_{OD}$ ), 预测的降水站点数占全部实际降水站点数的比重。

$$P_{OD} = \frac{T_p}{T_p + F_N} \quad (3)$$

BIAS(偏差评分, Bias Score, 记为 $B_S$ ), 主要用来衡量模式对某一量级降水的预报偏差, 该评分在数值上等于预报区域内满足某降水阈值的总站点数与对应实况降水总站点数的比值, 其值越接近1, 表明预报范围越接近实况。

$$B_S = \frac{T_p + F_p}{T_p + F_N} \quad (4)$$

式(1)~(4)中,  $T_p$ 为预报出现且实况出现站点数, 即命中数;  $F_p$ 为预报出现但实况未出现站点数, 即空报数;  $F_N$ 为预报未出现但实况出现站点数, 即漏报数;  $T_N$ 为预报未出现且实况未出现站点数, 即真负数。

## 2 气象站点分区

将2016—2020年汛期浙江省2 227个气象站逐日降水量数据作为训练数据,以欧氏距离作为相似度度量标准进行Kmeans聚类分区。为提升聚类合理性,根据肘部法则和平均轮廓系数来确定聚类簇数。综合肘部法则和平均轮廓系数(图略),7为Kmeans聚类分区合适的分区数。本文Kmeans聚类采用的各站训练数据中,仅包含逐日降水量数据,未加入任何站点空间地理位置信息。图3为浙江省气象站点分区结果,呈现明显的区域特征,存在南北、东西和海陆差异。分区结果与胡波等(2011)根据1961—2008年浙江省38个国家气象站汛期日降水资料,通过旋转经验正交函数(Rotated Empirical Orthogonal Function, REOF)方法得到的浙江省梅汛期降水分区类似,其中浙北北部区与4区对应,浙东南部区与1、3、6区对应,浙西南区与2、5区对应,浙中东区与7区对应。将以上分区结果与毛程燕等(2018)根据卫星资料反演得到的年平均降水空间分布对比,结果表明2 000 mm以上降水区对应5区,1 800~2 000 mm降水区对应2、3区,1 400 mm以下

降水区对应4、7区。分区区域与浙江省地貌地形存在明显相关性,用数字1~7及7种不同颜色代表浙江各区域(图3),其中1区为温州东部(瓯江流域),2区为衢州南部、金华南部和丽水西北部(仙霞岭区域),3区为温州南部、丽水东南部(洞宫山及南雁荡山区域),4区为嘉兴、湖州及杭州北部(杭嘉湖平原地区),5区为杭州中南部、衢州北部、金华北部和绍兴西部(浙西山区,包含白际山、千里岗山、龙门山),6区为台州和宁波、绍兴、金华、丽水、温州与台州交界小部分区域(括苍山及北雁荡山区域),7区为宁波、绍兴东部和舟山(会稽山、四明山、天台山以及浙北沿海区域)。

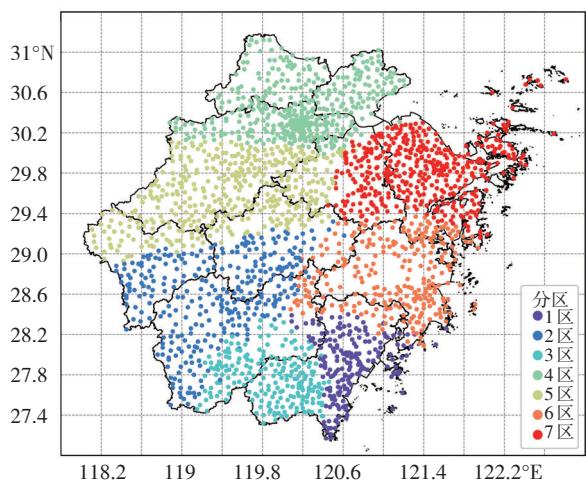


图3 基于Kmeans聚类算法的浙江省降水分区

Fig. 3 Precipitation division of Zhejiang province based on the K-means clustering algorithm

### 3 对OCF降水预报的分区订正效果

根据2021年汛期降水实况分别对OCF预报、全区订正及分区订正预报结果按不同降水阈值进行检验,检验评分结果如表1所示。从小雨以上量级降水( $\geq 0.1$  mm,即晴雨)评分指标来看,相较于OCF预报,全区订正和分区订正 $T_s$ 均由0.69提升至0.74左右,预报降水范围有所缩小,更接近实况( $B_s$ 更接近于1),但命中率由0.97下降至0.93。分析发现,全区订正和分区订正预报均通过合理减少降水预报数,以损失一定

命中率的代价,有效减少空报率(消空),明显提升预报技巧(包括 $T_s$ 和 $E_{Ts}$ ,下同),改善预报降水落区;其中全区订正预报方法消空效果略好于分区订正。从中雨以上量级降水( $\geq 10$  mm)评分指标来看,相较于OCF预报,全区订正和分区订正 $T_s$ 均由0.49下降至0.48, $E_{Ts}$ 均由0.48提升至0.49,中雨以上预报范围扩大,明显大于实况范围,命中率均由0.74提升至0.80左右;全区订正和分区订正预报相较OCF无明显优势。从大雨以上量级降水( $\geq 25$  mm)评分指标来看,相较于OCF预报,全区订正和分区订正预报技巧均有提升,特别是 $E_{Ts}$ 由0.27提升至0.36,大雨以上范围明显扩大(OCF预报范围大小更接近实况),命中率由0.46提升至0.64(全区订正和分区订正均值)左右;全区订正和分区订正将漏报率由0.54下降至0.35左右,空报率由0.51提升至0.59左右,以提升较小空报率为代价较明显提升命中率;全区订正和分区订正预报提升效果基本一致。从暴雨以上量级降水( $\geq 50$  mm)评分指标来看,相较于OCF预报,全区订正和分区订正预报技巧均有提升,分区订正提升更明显,暴雨以上范围大小更接近实况,OCF预报范围明显小于实况,全区订正和分区订正预报范围比实况略偏大),全区订正、分区订正预报命中率从OCF预报命中率的0.25分别提升至0.38、0.41。经分析发现,全区订正和分区订正预报均通过合理增加暴雨以上降水预报数(扩大预报范围),以增加一定空报率为代价(全区订正提升28%,分区订正提升24%),明显提高命中率(全区订正提高52%,分区订正提高64%),提升暴雨以上预报技巧,改善暴雨以上预报落区。综上,全区订正和分区订正预报均对OCF降水预报有较明显提升,对于小雨以上量级降水预报(晴雨预报),通过明显降低降水空报率,从而提升晴雨预报能力;对于大雨以上量级降水预报,通过合理扩大降水预报范围,提升预报能力,特别是对于暴雨以上量级降水分区订正方法提升更明显。

根据降水分区,进一步针对暴雨以上量级降水的

表1 OCF降水预报、全区订正以及分区订正预报方法在2021年汛期浙江省不同降水阈值中的检验评分

Table 1 Test scores of OCF precipitation forecast, overall correction, and regional correction method for different levels of precipitation during the flood season of 2021 in Zhejiang province

降水 量级	$T_s$			$E_{Ts}$			$B_s$			$P_{op}$		
	OCF	全区订正	分区订正	OCF	全区订正	分区订正	OCF	全区订正	分区订正	OCF	全区订正	分区订正
$\geq 0.1$ mm	0.692	0.742	0.74	0.692	0.745	0.725	1.377	1.186	1.217	0.972	0.93	0.931
$\geq 10$ mm	0.493	0.481	0.480	0.484	0.495	0.492	1.252	1.474	1.509	0.744	0.804	0.814
$\geq 25$ mm	0.313	0.333	0.335	0.265	0.363	0.362	0.931	1.535	1.596	0.460	0.633	0.651
$\geq 50$ mm	0.194	0.206	0.223	0.182	0.186	0.203	0.542	1.227	1.266	0.251	0.380	0.412

注: $T_s$ 为技巧评分, $E_{Ts}$ 为公平技巧评分, $B_s$ 偏差评分, $P_{op}$ 为命中率,下同

OCF、全区订正和分区订正预报结果进行检验。检验结果如表2所示。从中可见,不同区域暴雨以上降水OCF预报评分差异明显,7区最优,1区最差。全区订正和分区订正预报对1、2、3、4、5区暴雨以上预报均有明显提升效果,但6、7区提升效果不明显(仅有命中率一项有明显提升)。从两种订正方法评分指标对比来看,除5区分区订正预报评分略差于全区订正预报,其它区域分区订正的大多数评分优于全区订正。特别是4区,  $T_s$  全区订正提高15%,分区订正提高44%;  $E_{Ts}$

全区订正提升13%,分区订正提升44%;全区订正命中率提升至0.43,分区订正命中率提升至0.52。综合来看,对于暴雨以上量级降水,分区订正和全区订正除对浙中北沿海地区(6、7区)订正效果不佳外,其它地区效果较好,且分区订正明显优于全区订正。其中6、7区订正效果不佳主要原因是2021年汛期浙中北沿海地区(即6区)暴雨以上量级降水多为对流性降水,局地性强、突发性强,订正效果较差;OCF预报对7区的预报效果较好,较难进行有效订正。

表2 OCF降水预报、全区订正以及分区订正预报方法在2021年汛期浙江省暴雨以上量级降水中的检验评分

Table 2 Test scores of OCF precipitation forecast, overall correction, and regional correction method for the magnitude of rainfall above in Zhejiang province during the flood season of 2021

降水分区	$T_s$			$E_{Ts}$			$B_s$			$P_{ob}$		
	OCF	全区订正	分区订正	OCF	全区订正	分区订正	OCF	全区订正	分区订正	OCF	全区订正	分区订正
1区	0.115	0.158	0.183	0.104	0.137	0.161	0.351	1.028	1.121	0.139	0.276	0.328
2区	0.214	0.203	0.221	0.201	0.185	0.203	0.700	1.466	1.442	0.300	0.416	0.443
3区	0.124	0.245	0.246	0.113	0.230	0.224	0.306	0.840	0.997	0.144	0.368	0.395
4区	0.181	0.209	0.261	0.170	0.192	0.245	0.702	1.511	1.500	0.261	0.434	0.518
5区	0.162	0.181	0.176	0.151	0.163	0.159	0.496	1.225	1.185	0.209	0.341	0.327
6区	0.179	0.165	0.171	0.164	0.161	0.146	0.547	1.234	1.428	0.235	0.316	0.354
7区	0.309	0.259	0.285	0.295	0.239	0.265	0.599	1.204	1.172	0.377	0.454	0.482

#### 4 分区订正预报方法在2021年汛期降水预报中的应用效果

以两次降水过程为例,对分区订正预报效果进行分析、检验。第一次过程为2021年第6号台风“烟花”登陆前其外围影响造成的一次系统性暴雨到大暴雨过程(7月23日20:00—24日20:00)。给出此次过程OCF、全区订正、分区订正降水预报和实况分布图4。从中可知,OCF预报降水主要集中在浙中北地区,降水量级为大雨,暴雨预报区域主要分布于浙中沿海小部分区域,范围小且较分散。经全区订正后,大雨区和暴雨区范围有所扩大,浙北和浙中沿海预报了较明显的暴雨区。经分区订正,大雨区范围扩大,暴雨区主要集中在浙北和浙东沿海地区,范围明显扩大,且在浙东沿海(宁波)、浙北(湖州、杭州交界处)预报有大暴雨区。与降水实况对比发现,OCF预报暴雨区范围明显偏小,落区与实况偏差大,降水中心强度明显小于实况;全区订正对暴雨落区和强度有所改善;经分区订正后,暴雨区范围明显扩大,与实况暴雨区范围最为接近,且基本覆盖了大部分浙北和浙东沿海的实况暴雨区,同时预报出了大暴雨区(其中四明山区、天目山区大暴雨落区与实况基本重合)。综上,从暴雨以上降水的落区和强度来看,经分区订正后降水预报效

果有明显提升。

第二次过程为2021年8月12日20:00—13日20:00的一次对流性暴雨、局部大暴雨过程。给出此次过程OCF、全区订正、分区订正预报和实况降水分布情况图5。从中可见,OCF预报此次过程大雨以上降水位于浙北,暴雨区集中在杭嘉湖交界处。经全区订正,大雨范围扩大至浙中北,暴雨扩大至浙北大部分地区。经分区订正,暴雨区向浙东沿海扩大,浙北出现大暴雨区。由于此次降水过程为对流性降水,暴雨区小而分散,整体预报效果较差。OCF预报暴雨落区整体偏北,与实况有较明显偏差,且存在大暴雨漏报。全区订正后预报的暴雨落区囊括了杭州北部、杭州绍兴嘉兴交界处的暴雨区,但漏报衢州西北部、浙东沿海暴雨区以及大暴雨区。相对而言,分区订正后降水预报效果更好,在全区订正的基础上进一步预报出了浙东沿海暴雨区,且预报出大暴雨区(与实况相比落区偏北)。综上,对于分散性对流降水,在原始预报暴雨落区偏差较大的情况下,分区订正方法能够在一定程度上改进暴雨落区和强度预报,但出现较大范围空报。

综合来看,相较于OCF预报,分区订正预报方法对于受台风“烟花”外围影响造成的系统性暴雨到大暴雨过程有明显的订正效果,可以有效提升暴雨以上降水的预报效果;对于对流性暴雨到大暴雨过程分区

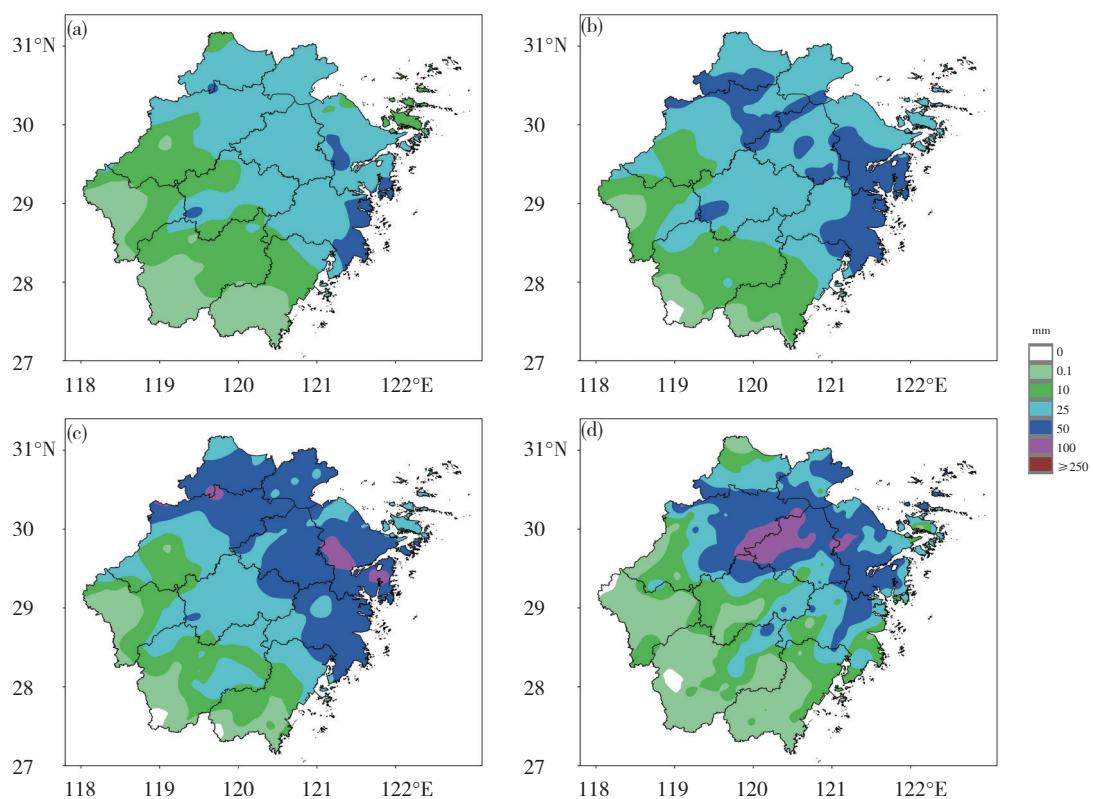


图4 2021年7月23日08:00起报的OCF(a)、全区订正(b)、分区订正(c)的浙江省12—36 h降水量与24 h降水量实况(d)

Fig.4 Diagram of future 12–36 hour precipitation forecasted by (a) OCF, (b) overall correction, (c) regional correction forecast and (d) observed precipitation in Zhejiang province starting from 08:00 BT on 23 July 2021

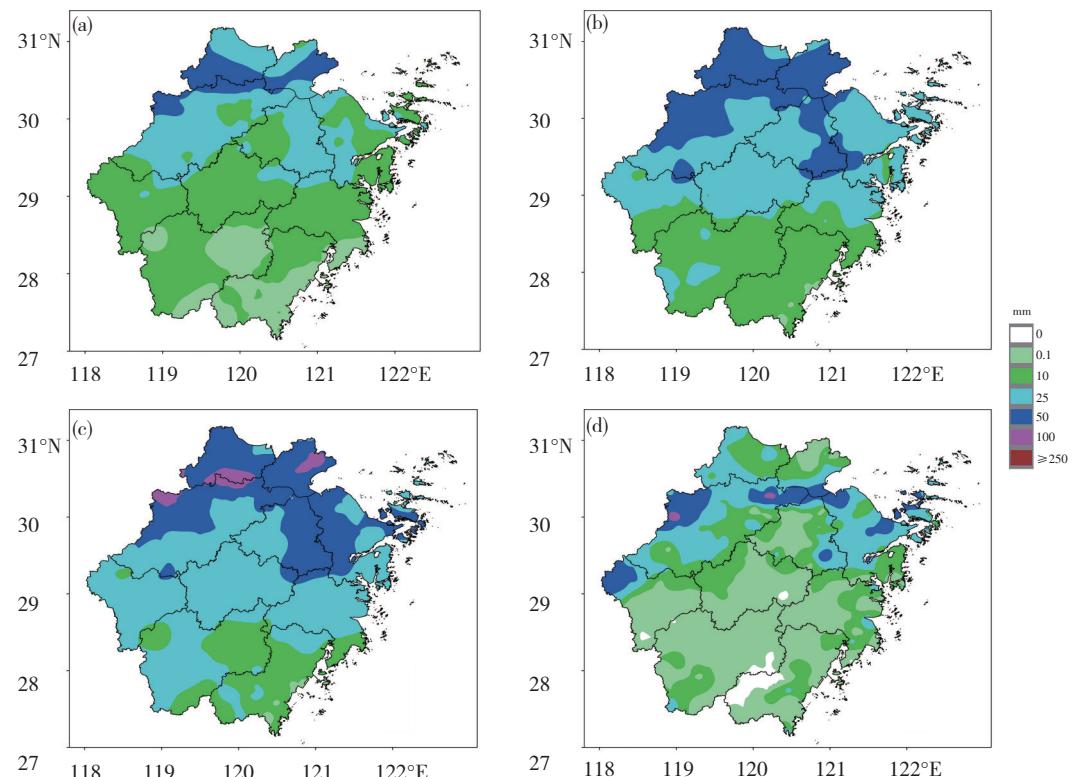


图5 2021年8月12日08:00起报的OCF(a)、全区订正(b)、分区订正(c)的浙江省12—36 h降水量与24 h降水量实况(d)

Fig.5 Diagram of future 12–36 hour precipitation forecasted by (a) OCF, (b) overall correction, (c) regional correction forecast and (d) observed precipitation in Zhejiang province starting from 08:00 BT on 12 August 2021

订正预报方法能够在一定程度上改善暴雨以上落区和强度,但出现较大范围空报。同时也可以从图4中OCF预报、分区订正与实况降水分布的对比中发现,虽然分区订正预报暴雨落区与实况接近,但大暴雨落区中心(绍兴与杭州交界处)漏报。此外,分区订正预报最大降水量为123.4 mm,虽然高于OCF预报最大值(78.9 mm),但仍然与实况最大值(256.6 mm)存在较大偏差。图5中同样存在大暴雨落区和极值偏差问题。

## 5 结论与讨论

首先,利用2016—2020年每年汛期浙江省2 227个气象站逐日降水资料作为训练数据(不包含任何地理特征信息),采用Kmeans聚类算法,以欧氏距离为相似度度量标准,度量各站逐日降水序列相似性,对2 227个站点进行分区。然后,对偏差订正进行时空上的改进,将改进后偏差订正法与分区结合形成分区订正方法。最后,以2021年汛期OCF降水预报数据及浙江省同期降水实况为例,对该订正方法进行检验。主要结论如下:

(1) 采用Kmeans聚类算法,将浙江省分为7个降水相似区,分区结果具有明显的区域特征,与浙江省主要山区、平原、流域分布密切相关。

(2) 2021年汛期降水订正检验结果表明,相较于OCF预报,全区订正和分区订正预报均能明显提升汛期降水预报效果。在晴雨预报中,全区订正预报消空能力略好,预报结果略优于分区订正。针对中雨以上量级预报,全区订正和分区订正预报均无明显订正效果。针对大雨以上量级预报,全区订正和分区订正预报结果有所提升。针对暴雨以上量级预报,全区订正和分区订正后预报效果有较明显提升,特别是分区订正提升幅度更大,暴雨以上降水预报落区接近实况,命中率全区订正由0.25提升至0.38,提升51%;分区订正提升至0.41,提升64%。分区检验表明,针对暴雨及以上量级降水,分区订正预报对于各个分区订正效果不一,对浙中北沿海地区订正效果不佳,其它地区分区订正效果明显提升。主要原因是浙中北沿海地区(即6区)暴雨以对流性降水为主,局地性强、突发性强,订正效果较差;OCF预报对7区的预报效果较好,较难进行有效订正。

(3) 选取2021年台风“烟花”登陆前造成的系统性暴雨、局部大暴雨过程和一次对流性暴雨、局部大暴雨过程对分区订正预报方法进行检验。对于系统性暴雨、局部大暴雨过程,分区订正后暴雨落区更接近实况,且预报出大暴雨区(其中四明山区、天目山区大暴雨落区与实况基本重合)。针对对流性暴雨、局部大

暴雨过程,暴雨以上落区和强度预报效果有所改善,但出现较大范围空报。

值得注意的是,不论是系统性降水过程还是对流性降水过程,虽然订正后暴雨预报范围接近实况,但对于大暴雨中心位置和强度的预报仍存在明显偏差。其主要原因在于:偏差订正法是通过对原有的预测值进行放大和缩小来调整降水量,它能有效订正降水量的大小,通过订正降水量可在一定程度上对降水落区做出订正,但很难直接对整体降水落区进行修正;同时它订正的基础来源于历史观测和预报数据的内插,很难预报极端或大量级降水(特别是超历史极值的降水)。这是分区降水订正方法的缺陷,也是需要继续提升的方向。分区订正预报方法是由偏差订正方法改进而来,经过2021年汛期检验,结果表明该方法能依据各个降水相似区的降水特征对汛期降水预报进行有效订正(特别是暴雨以上量级降水预报)。但它仍存在偏差订正法固有的难以改进强降水落区偏差、极端降水预报订正能力欠缺等问题。可以考虑采用机器学习方法如随机森林、梯度提升树模型等根据“配料法”来建立降水分类模型或总结分析不同天气形势下预报大量级降水偏差特征,根据该特征对大量级降水落区进行修正。极端降水预报订正还可以结合集合预报(李俊等,2020)来做进一步探索。

## 参考文献(References):

- 从靖,吴振玲,田笑,等.2021.海河流域东北冷涡背景下的降水预报订正研究[J].气候与环境研究,26(5):1-13. Cong J, Wu Z L, Tian X, et al. 2021. Correction of precipitation forecast under the Northeast Cold Vortex in the Haihe River Basin [J]. Climatic and Environmental Research,26(5):1-13 (in Chinese). doi:10.3878/j.issn.1006-9585.2021.20144
- 曹萍萍,肖递祥,徐栋夫,等.2018.基于概率匹配的西南区域模式定量降水订正试验[J].气象科技,46(1):102-111. Cao P P, Xiao D X, Xu D F, et al. 2018. Calibration of SWC-WARMS quantitative precipitation based on probability matching method [J]. Meteorological Science and Technology,46(1):102-111 (in Chinese). doi:10.19517/j.1671-6345.20170063
- 曹萍萍,康嵒,王佳津,等.2020.基于ECMWF模式的四川夏季强降水订正试验[J].暴雨灾害,39(1):63-70. Cao P P, Kang L, Wang J J, et al. 2020. Calibration of ECMWF model forecast of summer heavy rainfalls in Sichuan [J]. Torrential Rain and Disasters,39(1):63-70 (in Chinese). doi:10.3969/j.issn.1004-9045.2020.01.007
- 苟浩锋,王遂缠,杨银,等.2020.基于聚类分析的兰州地区自动站降水特征分析[J].沙漠与绿洲气象,14(1):108-114. Gou H F, Wang S C, Yang Y, et al. 2020. Analysis of AWS precipitation characteristics based on K-Means clustering method in Lanzhou Area [J]. Desert and Oasis Meteorology,14(1):108-114 (in Chinese). doi:10.12057/j.issn.1002-0799.2020.01.014

- 胡波,俞善贤,滕卫平,等.2011.浙江省汛期降水定量预测技术[C]/2011年第二十八届中国气象学会年会. Hu B, Yu S X, Teng W P, et al. 2011. Quantitative prediction technology of flood season precipitation in Zhejiang Province [C]/The 28th Annual Meeting of the Chinese Meteorological Society in 2011
- 李俊,杜钧,陈超君.2014.降水偏差订正的频率(或面积)匹配方法介绍和分析[J].气象,40(5):580-588. Li J, Du J, Chen C J. 2014. Introduction and analysis to frequency or area matching method applied to precipitation forecast bias correction [J]. Meteorological Monthly,40(5):580-588 (in Chinese). doi:10.7519/j.issn.1000-0526.2014.05.008
- 李俊,杜钧,许建玉,等.2020.一次特大暴雨过程高分辨率集合预报试验的检验和评估[J].暴雨灾害,39(2):176-184. Li J, Du J, Xu J Y, et al. 2020. The assessment and verification of high-resolution ensemble forecast for a heavy rainstorm [J]. Torrential Rain and Disasters,39(2):176-184 (in Chinese). doi:10.3969/j.issn.1004-9045.2020.02.008
- 毛程燕,龚理卿,廖君钰,等.2018.基于卫星探测资料分析浙江省降水的时空分布特征[J].中低纬山地气象,42(2):7-12. Mao C Y, Gong L Q, Liao J Y, et al. 2018. The spatial temporal features of precipitation in Zhejiang Province based on satellite data [J]. Mid-low Latitude Mountain Meteorology,42(2):7-12 (in Chinese). doi:10.3969/j.issn.1003-6598.2018.02.002
- 穆穆,陈博宇,周非凡,等.2011.气象预报的方法与不确定性[J].气象,37(1):1-13. Mu M, Chen B Y, Zhou F F, et al. 2011. Methods and uncertainties of meteorological forecast [J]. Meteorological Monthly, 37(1): 1-13 (in Chinese). doi:10.7519/j.issn.1000-0526.2011.1.001
- 毛程燕,龚理卿,廖君钰,等.2018.基于卫星探测资料分析浙江省降水的时空分布特征[J].中低纬山地气象,42(2):7-12. Mao C Y, Gong L Q, Liao J Y, et al. 2018. The spatial temporal features of precipitation in Zhejiang Province based on satellite data [J]. Mid-low Latitude Mountain Meteorology,42(2):7-12 (in Chinese). doi:10.7519/j.issn.1000-0526.2018.02.001
- 王伟,宋卫国,刘士兴,等.2011. Kmeans聚类与多光谱阈值相结合的MODIS云检测算法[J].光谱学与光谱分析,31(4):1061-1064. Wang W, Song W G, Liu S X, et al. 2011. A cloud detection algorithm for MODIS images combining Kmeans clustering and multi-spectral threshold method [J]. Spectroscopy and Spectral Analysis,31(4):1061-1064 (in Chinese). doi:10.3964/j.issn.1000-0593 (2011)04-1061-04
- 武英娇,杨浩,钱仙桃,等.2019. ECMWF集合预报在安徽大别山区降水预报中的检验[J].暴雨灾害,38(1):66-71. Wu Y J, Yang H, Qian X T, et al. 2019. Verification of ECMWF ensemble prediction in precipitation forecast of Ta-pieh Mountains area in Anhui Province [J]. Torrential Rain and Disasters,38(1):66-71 (in Chinese). doi:10.3969/j.issn.1004-9045.2019.01.008
- 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]. Torrential Rain and Disasters,37(1):1-7 (in Chinese). doi:10.3969/j.issn.1004-9045.2018.01.001
- 赵宁坤,张秀年,孙俊奎,等.2021.高分辨率区域模式降水预报在云南的检验[J].暴雨灾害,40(1):78-86. Zhao N K, Zhang X N, Sun J K, et al. 2021. Verifications of high-resolution regional numerical model precipitation forecast in Yunnan Province [J]. Torrential Rain and Disasters,40(1):78-86 (in Chinese). doi:10.3969/j.issn.1004-9045.2021.01.010
- 赵瑞霞,代刊,金荣花,等.2020.OTS,MOS和OMOS方法及其优化组合应用于72 h内逐3 h降水预报的试验分析研究[J].气象,46(3):420-428. Zhao R X, Dai K, Jin R H, et al. 2020. Comparison of OTS, MOS, OMOS methods and their combinations applied in 3h precipitation forecasting out to 72h [J]. Meteorological Monthly,46(3):420-428 (in Chinese). doi:10.7519/j.issn.1000-0526.2020.03.013
- 智协飞,吕游.2019.基于频率匹配法的中国降水多模式预报订正研究[J].大气科学学报,42(6):814-823. Zhi X F, Lü Y. 2019. Calibration of the multimodel precipitation forecasts in China using the frequency matching method [J]. Transactions of Atmospheric Sciences (in Chinese),42(6):814-823 (in Chinese). doi:10.13878/enu.dqkxxb.20190109001

(责任编辑 何明琼)