

吴启树, 韩美, 刘铭, 等. 基于评分最优化的模式降水预报订正算法对比. 应用气象学报, 2017, 28(3): 306-317.

DOI: 10.11898/1001-7313.20170305

基于评分最优化的模式降水预报订正算法对比

吴启树^{1)*} 韩美¹⁾ 刘铭¹⁾ 陈法敬²⁾

¹⁾(福建省气象台, 福州 350001) ²⁾(国家气象中心, 北京 100081)

摘 要

使用 2013 年 1 月 1 日—2016 年 1 月 7 日全国气象站观测资料, 应用准对称混合滑动训练期, 不改变雨带预报位置和形态, 基于模式降水预报订正结果的 TS 评分最优化及 ETS 评分最优化, 分别设计最优 TS 评分订正法 (OTS) 和最优 ETS 评分订正法 (OETS) 确定预报日各级降水订正系数, 对 2014—2015 年降水数值预报进行分级订正, 并与频率匹配法 (FM) 对比。结果表明: 在 24 h 累积降水的多个预报时效订正中, 无论是对欧洲中期天气预报中心、日本气象厅、美国国家环境预报中心和中国气象局的全球模式降水预报, 还是对 4 个模式的简单多模式平均, OTS 和 OETS 较 FM 在 TS 评分和 ETS 评分等传统降水检验指标上均更优秀, 其中 OTS 在所有时效均能提高模式降水预报质量, 为三者最优。在概率空间的稳定公平误差评分方面, OTS 在各时效、各单模式及多模式平均等方面优势明显。在预报员对应参考时效上, OTS 在 24~168 h 的 24 h 累积降水预报中的 TS 评分也优于主观预报。

关键词: 最优 TS 评分法; 最优 ETS 评分法; 频率匹配法; 降水分级订正; 训练期

引 言

数值预报模式输出的气象要素预报存在不同程度的系统性误差, 通过统计分析采用适当算法进行订正, 可以减少误差, 提高预报准确率^[1-11]。不同模式在降水预报的系统性误差方面表现各异, 如欧洲中期天气预报中心 (ECMWF)、日本气象厅 (JMA) 等机构的全球模式具有弱降水预报偏空报而强降水预报偏漏报的情况^[12-13], 区域模式如中国气象局武汉暴雨研究的中尺度模式 (AREM)^[14] 总体较实况偏强, 特别是中低量级的降水, 这些反映了模式对不同降水强度预报的误差非一致性和降水量在概率上的偏态分布。未分级的消除偏差集合平均法^[15-16] 预报质量优于单模式确定性预报, 但该方法对各量级的降水订正值相同, 存在小量级降水订正幅度过大而大量级降水订正过小的不足。孙靖等^[17] 在训练期和分级降水订正方面进行改进, 使训练期样本更加接近预报日前后的季节背景, 并对不同量级降水

进行相应的偏差订正, 但对中雨以上量级的模式降水预报阈值设定较为固定, 还有优化空间。周迪等^[18] 应用观测概率匹配订正法, 对四川盆地降水量为 50 mm 的集合预报平均值进行订正, 改善了 T213 模式的暴雨预报。陈博宇等^[12] 根据概率匹配和融合产品各自特点, 设计出概率匹配-融合法和融合-概率匹配法, 对大量级的降水预报较融合产品有一定提高, 但因融合产品百分位过于固定, 使大量级降水预报产生明显空报。李俊等^[14, 19] 利用实况降水频率结合概率匹配平均法, 能订正系统偏差和集合平均光滑的负作用。因此, 无论单模式还是多模式集成, 要提高模式降水订正后预报质量, 确定训练期实况与预报两者之间的关系至关重要。

频率匹配法 (FM) 在降水分级订正中以降水预报偏差^[19] 达到最优化为算法核心, 当训练期样本量足够大且比较接近于预报日前后的季节背景时, 该算法利用观测降水频率作为参考频率订正模式预报效果较好, 理论上能使预报与实况平均误差的数学期望值为零, 且空报率与漏报率大致相同。但该方

2016-10-13 收到, 2017-02-20 收到再改稿。

资助项目: 气象预报业务关键技术发展专项 (YBGJXM201703-06), 福建省自然科学基金社会发展引导性 (重点) 项目 (2017Y-008)

* email: 172475076@qq.com

法应用于业务时,因降水存在明显的季节变化,且当实况或预报在大量级降水上的样本量偏少时,订正系数易出现异常。即使订正后预报降水与实况面积相当,因降水落区与实况总是存在一定程度的差异,预报效果未必能达到最优。对于某一量级降水,在预报偏差为1的情况下,若将降水预报值适当提高(降低),该量级增加(减少)的预报站点带来TS评分^[20]增幅(降幅)大于(小于)原有评分时,相对于频率匹配法,预报质量可以得到改进,而订正效果是否达到最佳可以通过计算该量级的TS评分确定。

因此,本文基于训练期模式降水量预报调整后的TS评分最优化及ETS评分最优化进行建模,以确定预报日降水分级订正系数。分别设计最优TS评分订正法和最优ETS评分订正法,并与频率匹配法进行对比,优选订正质量既能在TS评分项目上达到最优,又能在ETS评分等降水预报检验项目上达到较优的算法。由于3个订正算法本质上是分别基于训练期模式降水订正后的TS评分、ETS评分和Bais评分^[19]的最优化设计,因此,本文还采用概率空间的稳定公平误差(stable equitable error in probability space, SEEPS)评分^[21-24]作为3种订正算法的第3方检验方法。SEEPS方法于2010年由Haiden等^[22]研究与应用,将站点按降水气候概率为基础划分成干、小雨、大雨,克服因地域降水概率差异带来检验的不公平性,并获得一个可以代表降水预报整体性能的单一评分,2011年被世界气象组织的确定性数值预报检验领导中心(LC-DNV)应用于业务,目前该方法也应用于中国气象局数值预报中心的模式检验,具有很高的公平性。

1 资料

1.1 实况资料

为确保计算效率及实况资料的可靠性,本文降水实况为2013年1月1日—2016年1月7日全国2414个气象站资料。本文个别降水检验方法需要应用各站点气候概率,则使用1984—2013年共30年的历史资料。

1.2 数值预报产品

产品包括中国气象局下发并在气象系统日常预报业务使用率较高的欧洲中期天气预报中心高分辨率模式(ECMWF)、日本气象厅高分辨率模式(JMA)、美国国家环境预报中心的全球模式

(NCEP)和中国气象局数值预报中心T639模式(T639)的确定性预报降水产品。2013—2015年每日两次,起报时间分别为00:00和12:00(世界时,下同)。ECMWF水平分辨率为 $0.25^\circ \times 0.25^\circ$;JMA和NCEP水平分辨率为 $0.5^\circ \times 0.5^\circ$;T639水平分辨率为 $1.125^\circ \times 1.125^\circ$ 。ECMWF预报时效最长,为0~240h,其他3个模式为84~216h不等。

模式的格点资料采用Cressman客观插值法进行站点插值^[25]。

2 方法

2.1 检验方法

本文设10个降水量级(k)应用于分级订正,阈值为 O_k ,10个降水量级分别为0,1,1,5,10,25,35,50,75,100,150mm。

结合天气预报业务的站点检验方法,本文采用TS评分、ETS评分(E)^[17]、空报率、漏报率、预报偏差(B)^[19-20]、SEEPS技巧评分^[21-24]和HSS技巧评分(H)^[24]等检验方法。对于第 k 量级的累积降水,部分检验方法计算如下:

$$E = \frac{N_A - R_a}{N_A + N_B + N_C - R_a}, \quad (1)$$

$$B = \frac{N_A + N_B}{N_A + N_C}, \quad (2)$$

$$H = \frac{N_A + N_D - R_a - R_d}{N_A + N_B + N_C + N_D - R_a - R_d}, \quad (3)$$

$$R_a = \frac{(N_A + N_B) \cdot (N_A + N_C)}{N_A + N_B + N_C + N_D}, \quad (4)$$

$$R_d = \frac{(N_B + N_D) \cdot (N_C + N_D)}{N_A + N_B + N_C + N_D}. \quad (5)$$

其中, N_A 为预报正确站(次)数, N_B 为空报站(次)数, N_C 为漏报站(次)数, N_D 为预报和实况均未达到阈值的正确站(次)数。 R_a 和 R_d 分别为空报站数与漏报站数相当时的随机预报 N_A 和 N_D 的数学期望。

TS评分、ETS评分、空报率、漏报率、预报偏差为目前我国气象系统降水预报业务检验考核项目。其中,ETS评分是对TS评分的改进,能对空报或漏报进行惩罚,使评分相对后者更加公平^[24]。SEEPS技巧评分计算方法相对复杂,计算方法参见文献^[21-23],并应用各站点30年降水气候概率;该技巧评分对预报误差的惩罚与降水气候概率相关,可自动适应不同气候区域或降水季节的站点降水概率,

使得多个站点的评分融合更加合理。HSS技巧评分也能对空报或漏报进行惩罚,且给随机预报和常量预报的期望评分为0,其数学属性为线性的且具有渐进公平性,也是公平的降水检验之一^[24]。

2.2 算法设计原理及实现

本文对比3种降水分级订正算法,其中前两个算法为本文设计。对于某一模式、某一时效,模式降水订正算法分别为算法1、算法2、算法3。

算法1:最优TS评分订正法(optimal threat score method, OTS)。

$$y = \begin{cases} 0, & x < F_1; \\ O_k + (O_{k+1} - O_k) \times \frac{x - E_k}{F_{k+1} - F_k}, & F_k \leq x < F_{k+1} (k \text{ 为 } 1 \sim 9); \\ x \times \left(\frac{O_{10}}{F_{10}}\right), & x \geq F_{10}。 \end{cases} \quad (6)$$

式(6)中, x 和 y 分别为模式降水的预报值和订正值, O_k 为第 k 量级降水阈值(k 为 $1, 2, \dots, 10$), F_k 为预报降水量订正到 O_k 时,该量级对应模式降水阈值。 $x < F_1$ 时, y 为0,称消空订正; y 与 x 比值为订正系数^[19], F_k 相对 O_k 越小(大)订正系数就越大

(小)。 F_k 从训练期调整订正系数使第 k 量级累积降水TS评分达到最高时求得。

算法2:最优ETS评分订正法(optimal equitable threat score method, OETS)。该订正算法与OTS相似,但模式订正阈值 F_k 由第 k 量级的累积降水ETS评分达到最高时的订正系数确定。

算法3:频率匹配法(frequency matching method, FM)^[14,19]。其中,参考频率为训练期实况降水,预报频率为训练期模式预报降水。具体步骤与文献[12]相似。

3种算法均只改变预报降水强弱,不改变预报的雨带位置和形态。假设经过FM订正后,某一量级(如暴雨)预报站点数与实况降水站点数相同,但只部分重合(图1a),那么OTS对降水TS评分提高的可能主要有两种情况。一种是订正时适当将预报降水量调大(图1b),这有可能造成空报率和漏报率的减小,从而提高TS评分等检验结果;另一种是适当调小(图1c),可能带来漏报率增大及空报率的减小,当空报率的减小幅度大于前者时在一定程度上也能提高TS评分等。

实际计算中,以暴雨量级为例, F_k 求解意见

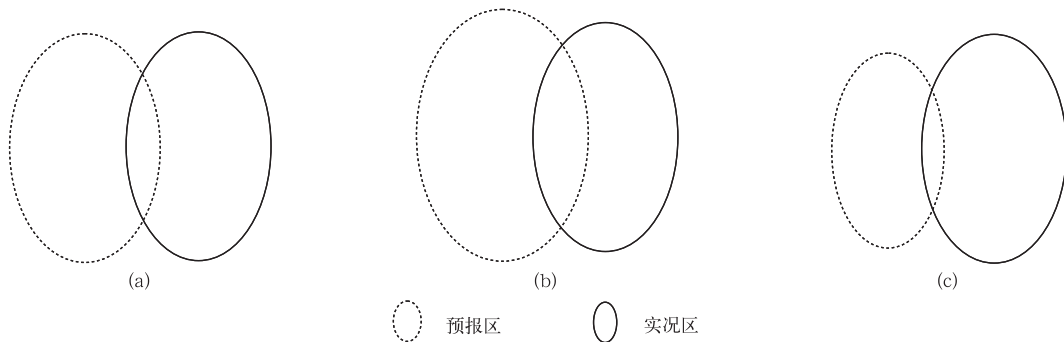


图1 TS评分或ETS评分提高的不同情况示意图

(a) 预报站点数与实况相同, (b) 预报站点数多于实况, (c) 预报站点数少于实况

Fig. 1 Illustration of different conditions of TS or ETS performance improvement

(a) the number of forecast stations is in consistent with the observation,

(b) the number of forecast stations is larger than the observation,

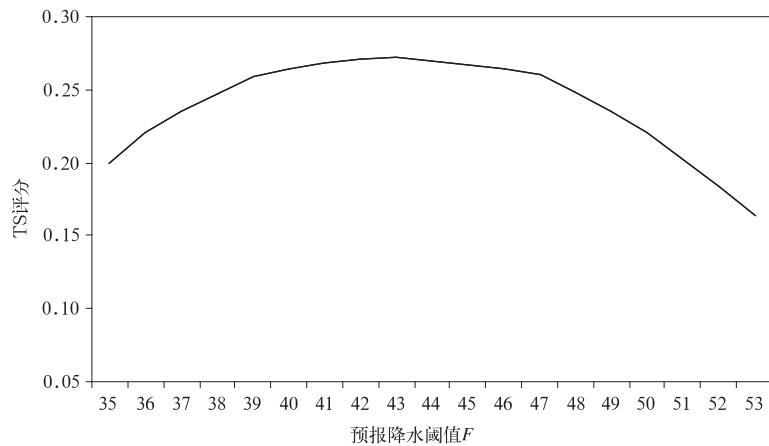
(c) the number of forecast stations is less than the observation

图2,在调整预报降水订正阈值时,若43.0 mm时的TS评分达0.272,为最高,则43.0 mm就为暴雨量级的 F 值。OETS对预报质量的提高原理与OTS相似,3种订正算法的优劣由预报检验进行评估。

为了控制因大量级降水样本偏少而引起订正系数异常偏大或偏小的情况,并考虑到多数全球模式

对大量级降水预报较实况偏弱的特点,因此,对于任一算法,设35 mm及以上量级降水订正系数的上(下)限分别为1.6(0.8)。

以最优TS评分订正法对ECMWF降水预报的订正为例,简称为ECMWF_OTs,另两个算法类似,分别简称为ECMWF_OETS和ECMWF_FM。

图2 阈值 F 求解示意图Fig. 2 Illustration of threshold F solution

为了能更好地自适应预报日前后相似的季节背景,建模时应用1年期的准对称滑动训练期^[25],即滑动各取当年预报日之前和上一年预报日之后20 d、每日两次共80次全国所有气象站样本,每次建模样本量为193120站(次),滑动计算出各量级降水订正阈值 F_k ,统一应用于预报日所有站点降水的订正。

3 算法对比

对全国气象站2014—2015年每日两次预报进行检验。准对称混合滑动训练期需要应用2013年资料。

3.1 对ECMWF降水预报的订正

3.1.1 24 h 预报时效检验

24 h 累积降水的24 h 预报TS评分(图3a),3种算法对ECMWF降水预报有明显的订正能力,尤其在弱降水和强降水提高更为显著。3种算法中,OTS仅在150 mm量级上低于OETS,其他量级均为最优。OETS除了在0.1 mm和5 mm两个量级上略低于FM,其他量级均优于后者。因此,在TS评分上,OTS最优,OETS次之。

ETS评分因ECMWF在中量级(5 mm和10 mm)降水预报较好(图3b),故3种算法在该量级评分提升幅度较小,但在其他所有量级均较模式有明显提升,特别是在弱降水和强降水方面提高更明显。3种算法中,OTS仅在150 mm量级上略低于OETS,其他量级均为最优。OETS在0.1 mm,1 mm和5 mm 3个量级上低于FM,其他7个量级

优于后者。因此,在ETS评分上,OTS最优,OETS次之。

ECMWF降水预报存在低量级偏空报、大量级偏漏报的情况(图3c~图3e),因此,3种算法通过对最低量级的消空以及对其他低量级应用较低的订正系数减少空报和降低湿偏差^[14],同时在大量级上应用较高的订正系数减少漏报。在样本量较多的中低量级(35 mm及以下),FM的空报率与漏报率相当(该算法的特点),且空报率为3种算法中最小,但在样本量较少的大量级(50 mm及以上),FM过度湿订正^[14],Bias评分超过1.3,且随量级增大而增加,最多超过2.2(150 mm量级),空报率为3种算法中最高。与FM相比,OTS和OETS算法虽然在中低量级上(25 mm及以下)湿偏差和空报略多,但湿偏差幅度较小(Bias评分为1.06~1.40)且漏报较少,略增的湿偏差使漏报率降幅较空报率增幅更加显著,TS评分更优。在较大量级上(35 mm及以上),OTS总体优于OETS,两者不仅在漏报率方面与FM相当,且在空报率方面显著优于FM。

HSS技巧评分(图3f)表明,3种算法对ECMWF降水预报有明显的订正能力,在弱降水和强降水提高更为显著。3种算法中,OTS仅在75 mm和150 mm两个量级上略低于OETS,其他量级为最优。OETS在0.1 mm,1 mm和5 mm 3个量级上低于FM,其他7个量级优于后者。因此,HSS技巧评分上OTS最优,OETS次之。

上述分析表明,OTS在TS评分、ETS评分、Bias评分、HSS技巧评分等方面均为3种算法中最优,OETS次之。

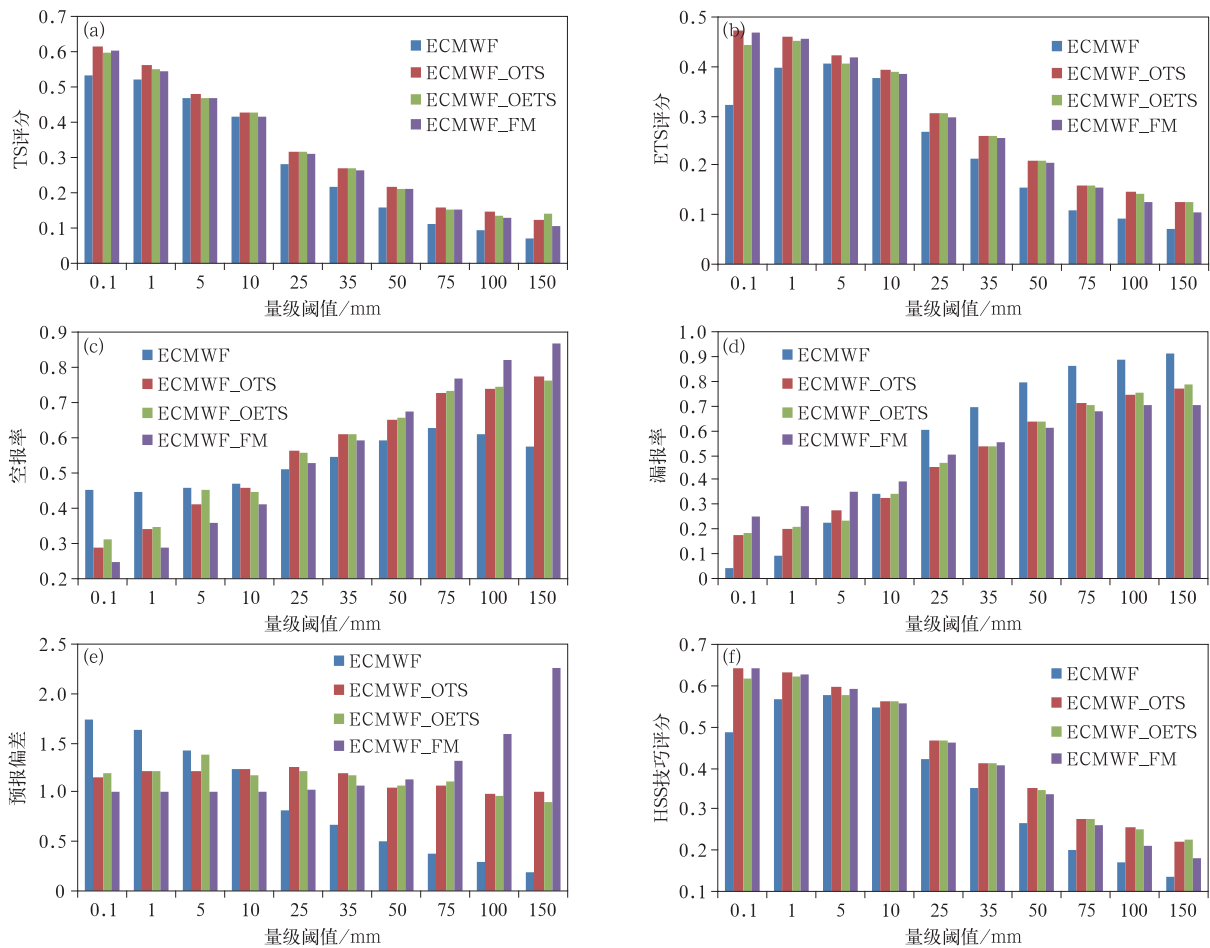


图3 2014—2015年ECMWF降水预报和FM,OTS,OETS 3种算法订正的24 h累积降水24 h预报的TS评分(a)、ETS评分(b)、空报率(c)、漏报率(d)、预报偏差(e)和HSS技巧评分(f)

Fig. 3 TS(a), ETS(b), false alarm rate(c), miss rate(d), Bias(e) and HSS(f) of 24 h accumulative precipitation by FM, OTS and OETS based on ECMWF with lead time of 24 h during 2014—2015

3.1.2 多预报时效 TS 评分

图4给出4个代表量级(0.1 mm, 10 mm, 25 mm, 50 mm)24~240 h预报时效TS评分。在所有量级各预报时效上,OTS均有一定的订正能力且为3种算法中最优(图4a~图4d)。在以10 mm为代表的中量级上,各算法提升幅度均最小,其中FM订正效果最差,甚至在120 h及以上预报时效出现反订正^[17]。在0.1 mm量级上,FM订正效果也最差,其在192 h及以上是预报时效为反订正,OETS虽在216 h及以上预报时效也出现反订正,但总体上优于前者。在25 mm量级和50 mm量级上,3种算法在各时效均优于模式自身,且TS评分均为OTS最优,OETS次之。

目前,天气预报业务上,降水预报以TS评分为主要考核项目,又因24~240 h预报时效的其他传统检验结果与3.1.1相近,不再详述。

3.1.3 SEEPS 技巧评分

除HSS技巧评分外,因TS评分、ETS评分、空报率、漏报率、Bias评分等传统的检验方法不是完全意义上的公平评分方法^[24],预报员可以根据单一检验方法指标的不同特性,规避不利的预报;且OTS,OETS和FM 3种算法是分别基于TS评分、ETS评分和Bias评分的最优化设计;3.1.1节检验的几个方法不是对所有量级的综合检验。因此,本文应用SEEPS检验方法作为3种算法的第3方检验。

SEEPS技巧评分显示(图5),各算法的技巧评分随预报时效增加而单调递减。FM仅在24 h时效对ECMWF降水预报略有正订正^[17],其他时效均为反订正,OTS在所有时效、OETS在216 h以下时效对ECMWF降水预报均有正订正,OTS更为明显。因此,SEEPS技巧评分上,OTS为最优,OETS次之。

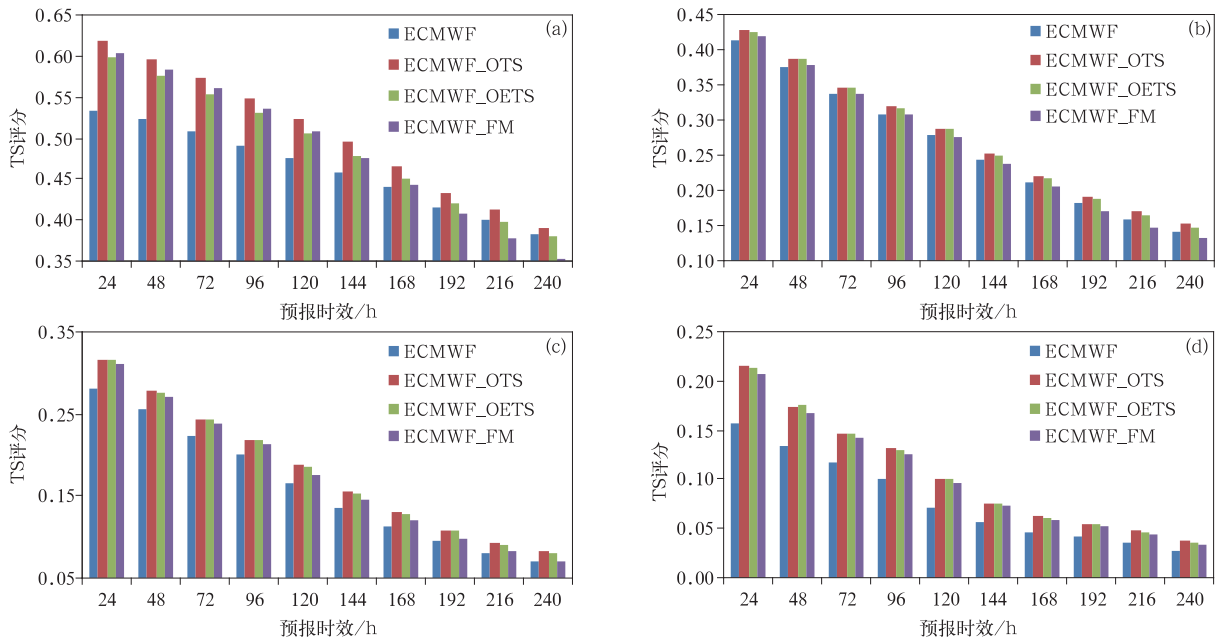


图 4 2014—2015 年 ECMWF 降水预报和 FM,OTS,OETS 3 种算法订正的 0.1 mm(a),10 mm(b), 25 mm(c)和 50 mm(d)24 h 累积降水 24~240 h 预报的 TS 评分

Fig. 4 TS of 0.1 mm(a),10 mm(b),25 mm(c)和 50 mm(d) of 24 h accumulative precipitation by FM, OTS and OETS based on ECMWF with lead time from 24 h to 240 h during 2014—2015

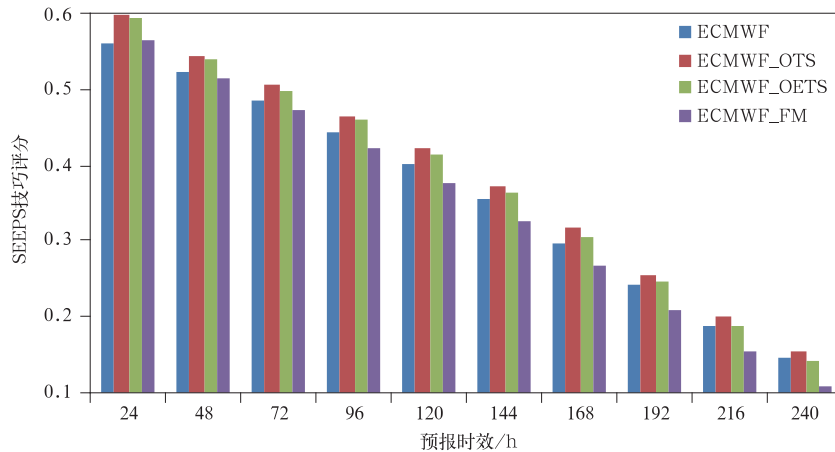


图 5 2014—2015 年 ECMWF 降水预报和 FM,OTS,OETS 3 种算法的 24 h 累积降水 24~240 h 预报的 SEEPS 技巧评分

Fig. 5 SEEPS skill scores of 24 h accumulative precipitation by FM, OTS and OETS based on ECMWF with lead time from 24 h to 240 h during 2014—2015

3.2 对 JMA,NCEP 和 T639 降水预报的订正

为进一步说明 3 种算法对比结果对模式的普适性,再对其他 3 个模式的降水预报进行订正并检验。JMA,NCEP 和 T639 3 种模式预报时效不同,因不影响结果,本文仅计算三者均有的 24 h 累积降水的 24~72 h 降水预报。

3.2.1 TS 评分检验

对于 JMA,NCEP 和 T639 降水预报的订正,因其他预报时效结论相近,本文仅列出 24 h 累积降水的 24 h 预报 TS 评分对比(图 6)。可以看出,对于 JMA 和 NCEP,3 种算法在所有量级均有明显的订正能力,OTS 仅在 NCEP 降水预报的 150 mm 量级

上低于 OETS,其他均为最优,FM 表现最差。T639 降水预报,3 种算法总体对模式有一定订正能力,但幅度小且集中在不大于 25 mm 的量级上;OTS 占优的量级最多,OETS 总体略好于 FM。总体上,3 种算法中 OTS 为最优,OETS 次之。需要指出的

是,对于 T639 的强降水预报,3 种算法对模式提高幅度有限。这是由于 3 种方法虽能订正模式降水的相对强弱,但并不能改变降水落区,当预报落区与实况不太一致时(T639 在 50 mm 及以上量级 TS 评分低于 0.2),仅通过增大或减小降水强度,预报质量不

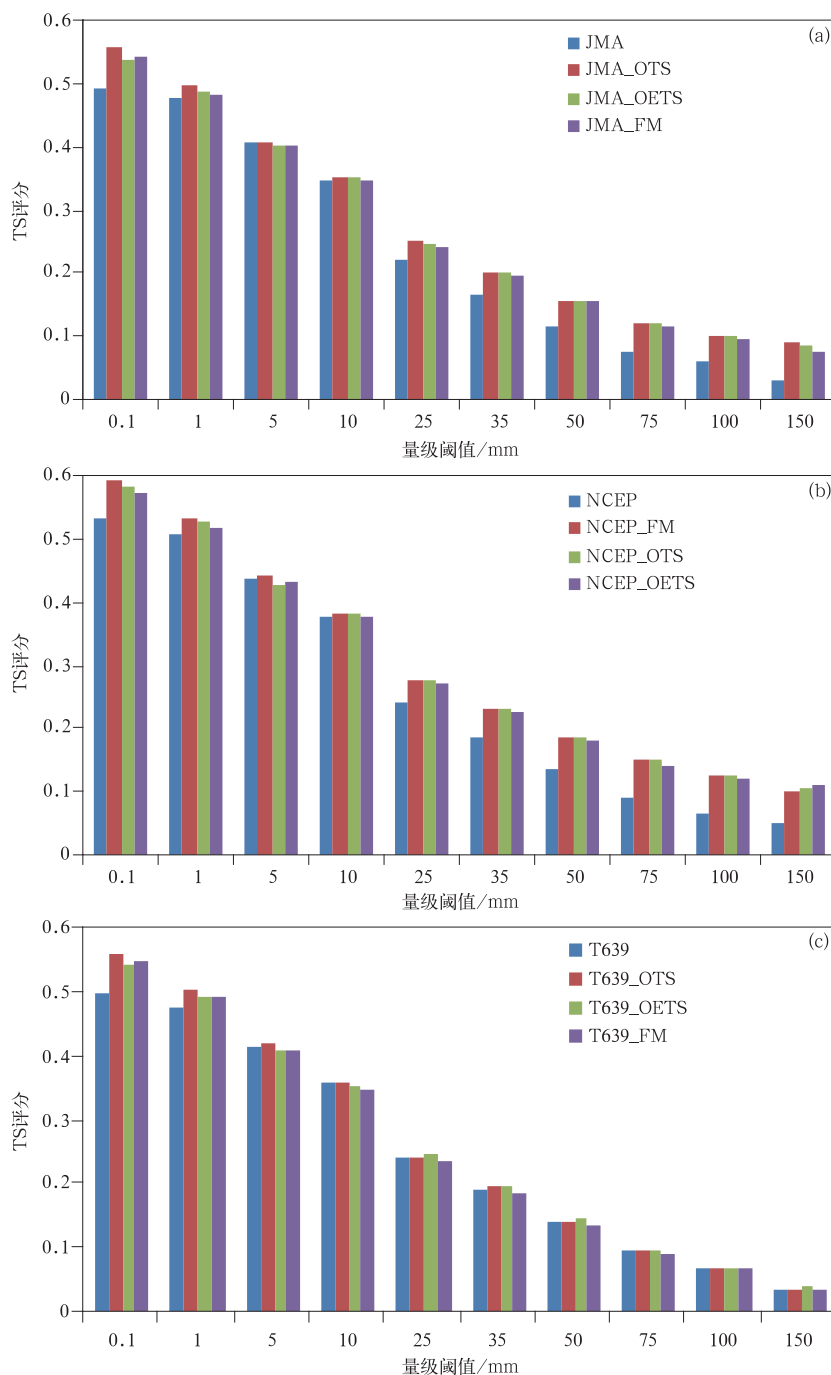


图6 2014—2015年JMA(a)、NCEP(b)和T639(c)降水预报及FM,OTS,OETS 3种算法订正的24h累积降水24h预报TS评分

Fig. 6 TS of 24 h accumulative precipitation by FM, OTS and OETS based on JMA(a), NCEP(b) and T639(c) with lead time of 24 h during 2014—2015

一定能得到提升。但一般情况下,3 种算法仍可对模式降水预报起到一定的正订正,且 OTS 最优。

3.2.2 SEEPS 技巧评分检验

从 SEEPS 技巧评分检验(表 1)可以看到,OTS 和 OETS 两种算法对 JMA,NCEP 和 T639 降水预报的订正均为正技巧,以 OTS 算法为最优;而 FM

对 JMA 和 NCEP 降水预报在 48 h 及以下预报时效上有正技巧,72 h 预报时效上为负技巧;FM 对 T639 降水预报所有 3 个时效订正均为负技巧。因此,在 3 个模式降水预报的多个时效,OTS 仍为最优,OETS 次之。

表 1 2014—2015 年 FM,OTS 和 OETS 3 种算法对 JMA,NCEP 及 T639 降水预报订正的 24 h 累积降水 24~72 h 预报的 SEEPS 技巧评分

Table 1 SEEPS skill scores of 24 h accumulative precipitation by FM,OTS and OETS based on JMA,NCEP and T639 with lead time from 24 h to 72 h during 2014—2015

模式	算法	24 h 预报	48 h 预报	72 h 预报
JMA	未订正	0.562	0.506	0.466
	FM	0.626	0.562	0.504
	OTS	0.640	0.572	0.513
	OETS	0.632	0.567	0.511
NCEP	未订正	0.525	0.493	0.474
	FM	0.572	0.533	0.495
	OTS	0.594	0.556	0.526
	OETS	0.588	0.550	0.519
T639	未订正	0.512	0.472	0.405
	FM	0.567	0.518	0.425
	OTS	0.602	0.546	0.453
	OETS	0.599	0.541	0.451

3.3 简单多模式降水预报平均

3.3.1 方法

对 ECMWF,JMA,NCEP 和 T639 4 个模式降水进行简单多模式平均^[16](ensemble mean,EMN)形成降水预报场,再应用算法订正后检验对比。本文多模式平均计算公式为

$$X_{EMN} = \frac{1}{4} \sum_{i=1}^4 x_i \quad (7)$$

式(7)中, x_i 为第 i 个模式的预报值。

3.3.2 TS 评分检验

因结论相近,本节仅对 24 h 预报的 TS 评分进行对比,其他时效略。3 种算法对简单多模式降水平均有明显的订正能力(图 7),特别是在降水量级两端(小量级和大量级)订正能力较强,这主要是由于简单多模式降水平均会产生小量级降水面积过

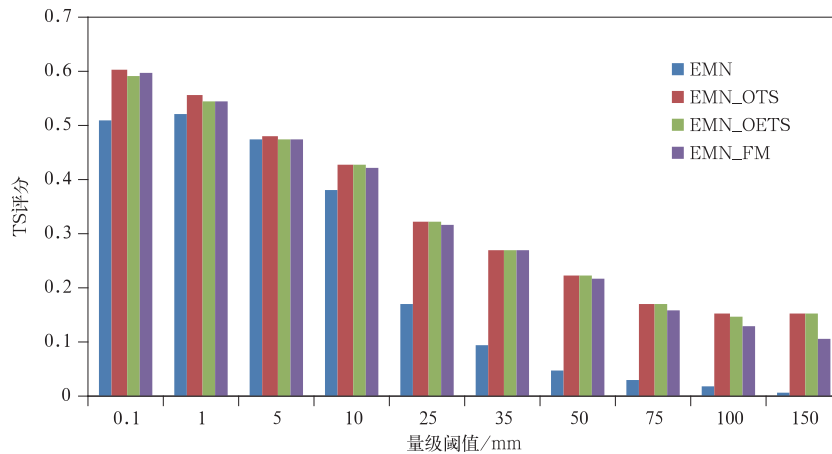


图 7 2014—2015 年简单多模式降水预报平均和 FM,OTS,OETS 3 种算法订正的 24 h 累积降水 24 h 预报 TS 评分

Fig. 7 TS of 24 h accumulative precipitation by FM, OTS and OETS based on ensemble mean forecasts with lead time of 24 h during 2014—2015

多、大量级降水面积偏小和强度偏弱的缺点。OTS 仅在 25 mm 和 75 mm 量级上略低于 OETS,其他量级均为 3 种算法中最优;而 FM 在所有量级上均为最差。

3.3.3 SEEPS 技巧评分检验

3 种算法对简单多模式降水预报平均的 SEEPS

技巧评分提高显著(图 8),以 OTS 为最优,OETS 次之,FM 较差。简单多模式降水预报平均未订正的 SEEPS 技巧评分 24 h,48 h,72 h 预报分别为 0.492,0.467,0.431,OTS 算法则分别达 0.601,0.556,0.516,提高值分别为 0.109,0.089 和 0.085,幅度较大;其他两种算法也有较好的表现。

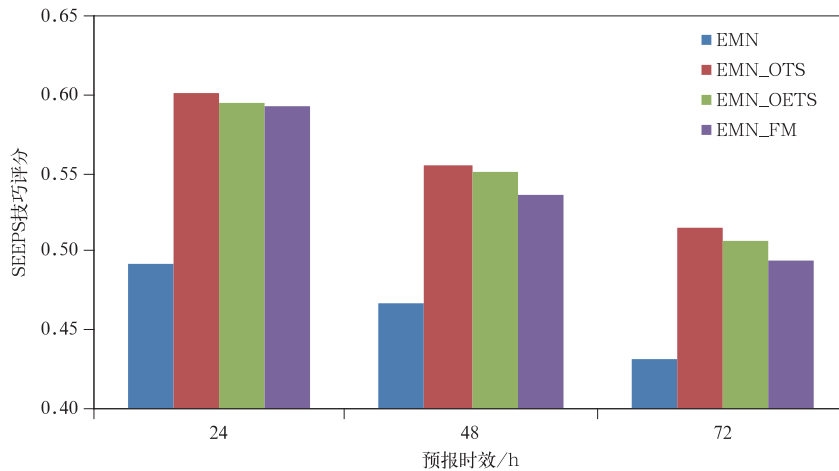


图 8 2014—2015 年简单多模式降水预报平均和 FM,OTS,OETS 3 种算法订正的 24 h 累积降水 24~72 h 预报 SEEPS 技巧评分

Fig. 8 SEEPS skill scores of 24 h accumulative precipitation by FM,OTS and OETS based on ensemble mean forecasts with lead time from 24 h to 72 h during 2014-2015

4 业务应用对比

为了更好地说明 3 种算法在预报业务中的应用效果,本章对比 2015 年主观预报与 3 种算法中的最优算法预报即 ECMWF_OTS。因目前业务上预报员降水预报主要以量级为主,不进行降水量预报,故检验以 TS 评分为主,未采用 SEEPS 技巧评分。

4.1 主观与客观降水预报

当前国家气象中心主观定量降水预报主要为每日 00:00 和 12:00 两次,核心任务及业务考核为 24 h 累积降水预报,00:00 起报为 24~168 h 预报时效,12:00 起报为 24~72 h 预报时效。基于预报员的可应用性,采用模式滞后 12 h 客观订正预报^[25],24 h 累积降水的 36 h 时效订正预报与预报员的 24 h 主观预报对应,其他时效类推。主观预报

的检验数据来自第 3 方检验机构即国家气象中心预报系统开放实验室。

4.2 检验对比

表 2 为 2015 年国家气象中心预报员主观预报与 ECMWF_OTS 的 24 h 累积降水 TS 评分。在小雨到大暴雨共 5 级累积降水预报 TS 评分中(大暴雨数据略),所有起报时间、预报时效,ECMWF_OTS 评分均优于预报员,其中 00:00 起报的 24 h 预报,预报员和 ECMWF_OTS 的暴雨 TS 评分分别为 0.175,0.189,后者较前者偏高 0.014;12:00 起报的 24 h 降水预报,预报员、ECMWF_OTS 的暴雨 TS 评分分别为 0.186,0.199,后者也较前者高 0.013。表明采用 OTS 算法对 ECMWF 订正后,降水预报准确率高于同期预报员主观预报,具有较高的业务应用价值。

表 2 2015 年中国国家气象中心预报员与 ECMWF_OTs 24 h 累积降水 TS 评分对比
Table 2 TS comparisons of 24 h accumulative precipitation between China NMC forecasters and ECMWF_OTs all over China stations in 2015

起报时间	预报时效/h	小雨		中雨		大雨		暴雨	
		预报员	ECMWF_OTs	预报员	ECMWF_OTs	预报员	ECMWF_OTs	预报员	ECMWF_OTs
00:00	24	0.594	0.608	0.395	0.414	0.281	0.302	0.175	0.189
	48	0.572	0.590	0.359	0.376	0.250	0.263	0.147	0.153
	72	0.552	0.565	0.328	0.335	0.225	0.233	0.118	0.136
	96	0.522	0.538	0.298	0.305	0.198	0.206	0.094	0.118
	120	0.500	0.514	0.271	0.276	0.168	0.178	0.073	0.093
	144	0.473	0.486	0.242	0.244	0.145	0.148	0.080	0.083
	168	0.442	0.448	0.208	0.214	0.124	0.133	0.050	0.062
12:00	24	0.590	0.609	0.392	0.414	0.285	0.304	0.186	0.199
	48	0.574	0.584	0.354	0.372	0.255	0.264	0.158	0.163
	72	0.550	0.559	0.323	0.337	0.222	0.236	0.125	0.150

5 结论和讨论

为了提高模式降水预报的订正效果,本文采用 ECMWF, JMA, NCEP 和 T639 的 4 个全球模式降水预报和全国气象站实况资料,设计了基于 TS 评分最优化订正的 OTS 算法及基于 ETS 评分最优化订正的 OETS 算法,并与基于 Bias 评分最优化订正的 FM 算法对比,同时结合业务需求进行多时效、多种检验方法对比,结果表明:

1) OTS, OETS 和 FM 3 种算法对 ECMWF, JMA, NCEP 和 T639 的 4 个模式降水预报及 4 个模式的简单多模式降水平均均有明显订正能力。从 TS 评分、ETS 评分等检验看,前两种算法较第 3 种更优,尤其是在大量级上两者表现更突出。以降水预报业务质量考核更为常用的 TS 评分来看,无论在时效、量级及模式上,OTS 占优项最多,为三者最优,OETS 次之。SEEPS 技巧评分对比中,OTS 在各时效、各模式中均明显更优。因此,OTS 的优异具有普适性。

2) OTS 和 OETS 对模式降水预报进行分级订正后,大部分量级预报与实况保持一定的湿偏差,但湿偏差幅度小,与 FM 相比,前两者在中低量级上有略多的湿偏差,略增的湿偏差带来的漏报率减小幅度大于空报率增加的幅度,使 TS 评分更高。对 ECMWF 降水预报订正中,虽然大量级上样本较少,但 OTS 和 OETS 的 Bias 评分仍保持在 1.1 左右,相对稳定,而 FM 不稳定,其 Bias 评分在最大量级上超过 2.2,过度湿订正。

3) 应用 2015 年全国气象站资料,在预报员可参考的对应时效,与国家气象中心定量降水预报对比,OTS 对 ECMWF 降水订正的各项检验均较前者更优,其 24 h 预报的全国暴雨 TS 评分平均达 0.194,表现出很好的可应用性。

本文分级订正算法采用的是全国所有站点作为一个区进行统一订正方案,并不一定是最优的分区订正方案,由于降水量在概率上的偏态分布,故采用单站独立订正方案会产生模式降水订正阈值不稳定的情况,订正质量低于前者,若对不同气候区进行更优的分区订正,质量可提升。本文订正算法对比的是站点预报,针对格点进行订正后、再插值到站点进行检验,结论与上述一致,目前算法也应用于业务的站点和网格预报。另外,因不影响结论,本文并未对多模式集成进行更深入研究。但若采用更优的权重对多模式集成后,通过 OTS 算法进行降水分级订正,预报质量可能更优。

参考文献

- [1] 周兵,赵翠光,赵声蓉.多模式集合预报技术及其分析与检验.应用气象学报,2006,17(增刊 I):104-109.
- [2] 赵声蓉.多模式温度集成预报.应用气象学报,2006,17(1):52-58.
- [3] 林春泽,智协飞,韩艳,等.基于 TIGGE 资料的地面气温多模式超级集合预报.应用气象学报,2009,20(6):706-712.
- [4] 范丽军,符宗斌,陈德亮.统计降尺度法对华北地区未来区域气温变化情景的预估.大气科学,2007,31(5):887-897.
- [5] 赵声蓉,裴海瑛.客观定量预报中降水的预处理.应用气象学报,2007,18(1):21-28.
- [6] 赵声蓉,赵翠光,赵瑞霞,等.我国精细化客观气象要素预报进展.气象科技进展,2012,2(5):12-21.

- [7] 车钦,赵声蓉,范广洲. 华北地区极端温度 MOS 预报的季节划分. 应用气象学报, 2011, 22(4): 429-436.
- [8] 刘还珠,赵声蓉,陆志善,等. 国家气象中心气象要素的客观预报——MOS 系统. 应用气象学报, 2004, 15(2): 181-191.
- [9] 马清,龚建东,李莉,等. 超级集合预报的误差订正与集成研究. 气象, 2008, 34(3): 42-48.
- [10] 王在文,郑祚芳,陈敏,等. 支持向量机非线性回归方法的气象要素预报. 应用气象学报, 2012, 23(5): 562-570.
- [11] 胡邦辉,刘善亮,席岩,等. 一种 Bayes 降水概率预报的最优子集算法. 应用气象学报, 2015, 26(2): 185-192.
- [12] 陈博宇,代刊,郭云谦. 2013 年汛期 ECMWF 集合统计量产品的降水预报检验与分析. 暴雨灾害, 2015, 34(1): 64-73.
- [13] 张宏芳,潘留杰,杨新. ECMWF、日本高分辨率模式降水预报能力的对比分析. 气象, 2014, 40(4): 424-432.
- [14] 李俊,杜钧,陈超君. 降水偏差订正的频率(或面积)匹配方法介绍和分析. 气象, 2014, 40(5): 580-588.
- [15] 王海霞,智协飞. 基于 TIGGE 多模式降水量预报的统计降尺度研究. 气象科学, 2015, 35(4): 430-437.
- [16] 智协飞,季晓东,张璟,等. 基于 TIGGE 资料的地面气温和降水的多模式集成预报. 大气科学学报, 2013, 36(3): 257-266.
- [17] 孙靖,程光光,张小玲. 一种改进的数值预报降水偏差订正方法及应用. 应用气象学报, 2015, 26(2): 173-184.
- [18] 周迪,陈静,陈朝平,等. 暴雨集合预报-观测概率匹配订正法在四川盆地的应用研究. 暴雨灾害, 2015, 34(2): 97-104.
- [19] 李俊,杜钧,陈超君. “频率匹配法”在集合降水预报中的应用研究. 气象, 2015, 41(6): 674-684.
- [20] 王雨,闫之辉. 降水检验方案变化对降水检验评估效果的影响分析. 气象, 2007, 33(12): 53-61.
- [21] Rodwell M J, Richardson D S, Hewson T D, et al. A new equitable score suitable for verifying precipitation in numerical weatherprediction. *Quart J Roy Meteor Soc*, 2010, 136: 1344-1363.
- [22] Haiden T M, Rodwell M J, Richardson D S. Intercomparison of global model precipitation forecast skill in 2010/11 using the SEEPS score. *Mon Wea Rev*, 2012, 140: 2720-2733.
- [23] 陈法敬,陈静. “SEEPS”降水预报检验评分方法在我国降水预报中的应用试验. 气象科技进展, 2015, 5(5): 6-14.
- [24] Jolliffe I T, Stephenson D B. 预报检验. 李应林,译. 北京:气象出版社, 2016: 58-68.
- [25] 吴启树,韩美,郭弘,等. MOS 温度预报中最优训练期方案. 应用气象学报, 2016, 27(4): 426-434.

A Comparison of Optimal-score-based Correction Algorithms of Model Precipitation Prediction

Wu Qishu¹⁾ Han Mei¹⁾ Liu Ming¹⁾ Chen Fajing²⁾

¹⁾ (*Meteorological Observatory of Fujian Province, Fuzhou 350001*)

²⁾ (*National Meteorological Center, Beijing 100081*)

Abstract

Based on data from national meteorological stations, one year quasi-symmetrical mixed running training period (QSRTP), and precipitation prediction from CMA (T639), ECMWF, NCEP, JMA, both optimal threat score (OTS) method and optimal equitable threat score (OETS) method are designed to conduct a comparison experiment on correction algorithms for model precipitation with frequency matching (FM) method. Through classification correction, three methods are used merely to calibrate model precipitation amount with the predicted rain-belt location and shape kept unchanged. The OTS method figures out correction coefficients of different precipitation classes by optimizing threat score (TS) of corrected precipitation within training period. OETS is similar to OTS but achieved by optimizing ETS. Correction experiments are conducted twice a day with forecast time at 0000 UTC and 1200 UTC, respectively. To consider seasonal background, 20 days before the forecast day and 20 days after the same day in the previous year are adopted to constitute training period. For each national meteorological station, there are 80 samples in total. The correction experiment shows that for either precipitation products of ECMWF, JMA, NCEP, CMA, or their ensemble mean, both OTS and OETS show much better performance than FM in 24 h accumulated precipitation classification calibration with different lead time according to traditional verification methods like TS and ETS. In particular, OTS is the best and can improve precipitation prediction in all lead times. After correction, both OTS and OETS incline to forecast larger precipitation area than observation for most classes but less precipitation amounts. Compared to FM, both methods tend to produce a little higher false alarm rates in middle and low classes, which is much less than the reduced missing rate, thereby leading to a higher threat score. In terms of ECMWF correction, OTS and OETS have a relatively stable Bias score of 1.1, although there are much fewer samples in high class. By contrast, FM produces an unstable Bias score, especially in maximum class with score over 2.2, indicating an excessively high missing rate. As for stable equitable error in probability space (SEEPS), OTS has superiorities over all lead times, all single models and multi-model mean. Furthermore, TS of corrected ECMWF precipitation using OTS method in 2015 are also better than subjective forecast from all aspects, with national averaged threat score of 1 d rainstorm forecast reaching 0.194.

Key words: optimal threat score method; optimal equitable threat score method; frequency matching method; precipitation classification calibration; training period