

# 数值天气预报检验方法邻域法的研究进展

庄园<sup>1,2</sup> 张悦含<sup>3</sup> 孙世玮<sup>1,2</sup> 王元<sup>4</sup>

(1 南京气象科技创新研究院, 南京 210041; 2 中国气象局交通气象重点开放实验室, 南京 210041;  
3 中国气象局气象干部培训学院, 北京 100081; 4 南京大学大气科学学院, 南京 210023)

**摘要:** 邻域检验方法被广泛应用于数值天气预报的检验, 随着高分辨率数值模式的不断发展, 新的邻域检验方法不断涌现, 邻域法的应用也逐渐从确定性预报扩展到集合概率预报。首先对不同邻域法进行了回顾, 并按照对预报和观测场处理方式的不同, 将邻域法分为升尺度法、邻域最大值法、格点概率法和其他方法。而后基于一套观测和预报资料, 对不同邻域法进行比较。不同邻域法适用情形不同, 在对预报进行检验时应根据实际需要选择合适的方法。在全面回顾的基础上, 指出邻域法的优势和不足, 并对未来发展方向进行展望。

**关键词:** 预报检验, 空间检验方法, 邻域法

**DOI:** 10.3969/j.issn.2095-1973.2023.06.004

## Research Progress of Neighborhood Methods for Numerical Weather Forecast Verification

Zhuang Yuan<sup>1,2</sup>, Zhang Yuehan<sup>3</sup>, Sun Shiwei<sup>1,2</sup>, Wang Yuan<sup>4</sup>

(1 Nanjing Joint Institute for Atmospheric Science, Nanjing 210041; 2 Key Laboratory of Transportation Meteorology, China Meteorological Administration, Nanjing 210041; 3 China Meteorological Administration Training Center, Beijing 100081; 4 School of Atmospheric Sciences, Nanjing University, Nanjing 210023)

**Abstract:** Neighborhood method has been widely used in numerical weather forecast verification. With the development of high-resolution numerical model, there have emerged new neighborhood methods, and the application of neighborhood method has been extended from deterministic forecasts to ensemble forecasts. Firstly, different neighborhood methods are reviewed. According to different processing methods for forecast and observation field, neighborhood methods are categorized into four types, upscaling method, neighborhood maximum method, grid probability method and other methods. Then different neighborhood methods are compared based on a set of observation and forecast data. The application of different neighborhood methods should be chosen according to the actual requirements. The advantages and disadvantages of neighborhood methods are pointed out and the future development direction is prospected based on a comprehensive review.

**Keywords:** forecast verification, spatial verification method, neighborhood method

### 0 引言

在过去几十年中, 数值天气预报模式的时空分辨率持续提高, 主观评估认为这些模式能够提供比粗分辨率模式更加真实的预报, 然而, 由于空间(或时间)位移误差和快速增长的小尺度误差会导致双重惩罚<sup>[1-2]</sup>, 使用传统逐网格检验方法对高分辨率模式预报进行检验并未显示出预报水平的提升<sup>[3]</sup>。为了协调

客观指标和主观评价之间的差异, 研究人员发展了大量空间检验方法, Dorminger等<sup>[1]</sup>将这些方法划分为邻域法、尺度分离法、基于特征的方法、变形场法和距离度量法。其中, 邻域法尤其适合高分辨率模式的检验, 被广泛应用于高分辨率模式的降水<sup>[4-8]</sup>、冰雹<sup>[9-10]</sup>、反射率<sup>[11-15]</sup>、上升气流螺旋度<sup>[16-19]</sup>等标量场以及矢量风场<sup>[20]</sup>的预报检验。

小尺度可预报性有限, 期望高分辨率模式在格点上精确是不现实的<sup>[21]</sup>, 邻域法放松了传统检验中预报和观测在格点上精确匹配的要求, 采用一定的不确定度, 如在某地点附近、在某时刻左右<sup>[19, 22-23]</sup>、在某大小左右<sup>[5]</sup>等来解释高分辨率模式的预报结果。最常用的邻域是空间邻域, 不同形状(方形或圆形)的邻域对检验结果没有明显影响<sup>[24]</sup>, 随着邻域大小的增加, 观测和预报的匹配要求越来越宽松, 通过一系列

收稿日期: 2022年9月28日; 修回日期: 2023年1月15日

第一作者: 庄园(1992—), Email: zhuangy@cma.gov.cn

通信作者: 张悦含(1994—), Email: zhangyuehan@cma.gov.cn

资助信息: 国家自然科学基金青年科学基金项目(42105151); 中国气象科学研究院基本科研业务费专项资金(2021Y008); 中国气象局气象干部培训学院青年科研项目(2022CMATCQN05)

不同大小邻域的检验结果可以解释预报和观测的内在不确定性<sup>[25]</sup>、衡量预报结果接近实际结果的程度以及判断预报在何种尺度上是有用的。不同的邻域方法对观测及预报的处理方式不同，或对于匹配要求的宽松程度不同，从不同角度来评估预报的质量。邻域法在对观测和/或预报进行邻域处理后，可使用常用的传统检验指标来评估预报的质量。对于连续变量的确定性预报，常使用偏差(bias)、均方根误差(RMSE)等连续变量检验指标；对于二分类事件的确定性预报，常建立列联表并计算公平兆得分(ETS)、频率偏差(BIAS)、命中率(POD)、空报率(F)等二元检验指标；对于概率预报，常使用Brier技巧评分(BSS)、相对作用特征(ROC)图、ROC曲线下面积、可靠性图等检验方式。另外，确定性预报和概率预报都可以使用相对价值<sup>[26]</sup>(RV)来评估预报的经济价值。

Ebert<sup>[24]</sup>回顾了2008年之前的邻域方法并系统阐述了使用邻域法检验高分辨率确定性预报的基本框架，但近十几年来又有许多新的邻域方法出现，邻域法的应用也从确定性预报扩展到了集合概率预报，有必要对目前常用的邻域方法进行介绍、比较与总结。随着邻域法的发展，Ebert<sup>[24]</sup>的基本框架出现了一定的局限性，比如他认为每种邻域法都采取以下两种匹配策略中的一种：一是预报邻域与观测邻域比较，二是预报邻域与观测格(站)点比较。目前有些邻域法的匹配策略并不属于这两类，因此有必要对邻域法的分类方

式进行调整。另外，正如Schwartz等<sup>[21]</sup>所指出的，许多研究在邻域法的术语、定义、实施和解释等方面存在一些混淆，有必要进行统一。综上所述，有必要对邻域法进行系统的梳理和总结。

本文首先回顾了较有代表性的邻域检验方法及指标，并按照对预报和观测处理方式的不同将邻域法进行分类，共分为升尺度法、邻域最大值法、格点概率法和其他方法4类；而后，基于一套观测和预报资料以及文献调研，对不同方法进行比较，为实际运用邻域法提供参考；在回顾与比较不同邻域方法的基础上，探讨了使用邻域法时需注意的问题以及邻域法的优势与不足，最后对未来邻域法的发展与应用进行了展望。

## 1 邻域检验的主要方法

表1和表2分别概括了确定性预报和集合预报的邻域检验方法。Ebert<sup>[24]</sup>根据匹配策略的不同将邻域法分为两类，一类为预报邻域与观测邻域比较，另一类为预报邻域与观测格点比较，而如表1及表2所示，目前还有预报格点与观测邻域比较以及混合匹配两种匹配策略，且一些方法可以根据观测和模式分辨率选择合适的匹配策略而非限定某种匹配策略。为了使分类更加合理，本文不沿用Ebert<sup>[24]</sup>的分类框架，而是根据对预报和观测场的处理方式进行分类，将确定性预报的邻域检验方法分为升尺度法、邻域最大值法、格点概率法和其他方法4类，其中升尺度法、邻域最大值法、

表1 确定性预报的邻域检验方法  
Table 1 Neighborhood verification methods for deterministic forecast

方法类别	方法名称	匹配策略	检验指标
升尺度法	平均值升尺度(AU)法/ 最大值升尺度(MU)法等	观测格点, 预报邻域 观测邻域, 预报格点	根据变量类型选择连续变量检验指标, 二元检验指标, RV等
		仅对预报处理 仅对观测处理 对预报和观测处理	
邻域最大值法	A01空间(多事件)列联表法	观测格点, 预报邻域	二元检验指标, RV, ROC图(用于多事件列联表法)等
	C10空间列联表法	混合匹配	
	M15空间列联表法	观测邻域, 预报格点	
	S16空间列联表法	观测邻域, 预报邻域	
格点概率法	邻域概率平均(NPA)法/ 高斯平滑法	观测格点, 预报邻域 观测邻域, 预报邻域	BS, BSS, 可靠性图, RV, ROC图, ROC面积等 FSS, 可靠性图, RV, ROC图, ROC面积, 二元检验指标(用于最低概率法)等
		仅对预报处理	
		对预报和观测处理 仅对观测处理(PP预报法)	
其他方法	区域相关RMSE	观测邻域, 预报邻域	区域相关RMSE
	SS列联表法	观测邻域, 预报邻域	二元检验指标

格点概率法已运用到集合预报检验中。表1和表2还给出不同方法可用的检验指标，特殊的邻域检验指标将在本节中介绍。另外，Ebert<sup>[24]</sup>文中的一些方法，如边缘分布不连贯的模糊逻辑列联表法、多被认为是尺度分离法的强度尺度法<sup>[2]</sup>等，不在本文讨论之列。

### 1.1 升尺度法

升尺度(upscaling, UC)法将观测和预报的检验区域均划分为若干一定大小的空间窗口，使得每个窗口中包含若干站点或网格点，选取一个值来代表每个空间窗口，最常采用空间的平均值<sup>[6-8, 27-29]</sup>或最大

表2 集合预报的邻域检验方法  
Table 2 Neighborhood verification methods for ensemble forecast

方法名称	匹配策略	对集合成员的处理方式	检验指标
升尺度法	观测格点, 预报邻域	对各集合成员升尺度, 再对给定阈值求集合概率	BS, BSS, ROC图, ROC面积, 可靠性图, RV等
	观测邻域, 预报格点		
	观测邻域, 预报邻域		
邻域最大值法	观测邻域, 预报邻域	对各集合成员处理并转化为二元概率, 再对所有成员平均得到集合概率	BS, BSS, ROC图, ROC面积, 可靠性图, RV等
	观测格点, 预报邻域 (空间多事件列联法)	取不同发生概率(至少有m个成员为事件发生)阈值生成多个列联表	ROC图, RV
格点概率法	观测格点, 预报邻域	先求各集合成员的邻域概率再对所有成员平均得到集合邻域概率	BS, BSS, ROC图, ROC面积, 可靠性图, RV等
	观测邻域, 预报邻域		

值<sup>[6, 21, 27, 30]</sup>来代表空间窗口的值, 也可使用中位数、百分位数等分布变量<sup>[6]</sup>, 代表窗口的值与窗口内观测或预报值的确切位置无关。过去的研究对UC法的命名比较混乱<sup>[21]</sup>, 本文根据代表窗口的值, 称使用平均值代表窗口的方法为平均值升尺度法(averaging upscaling, AU), 称使用最大值代表窗口的方法为最大值升尺度法(maximum upscaling, MU), 以此类推。

UC法可以对观测和预报采取不同的邻域匹配策略, 可根据观测和预报的分辨率选择同时对观测和预报进行升尺度处理, 或仅对观测或预报进行处理。多数研究将观测和预报同时升尺度到更低分辨率进行检验<sup>[6, 8, 21, 27-30]</sup>, 但当观测(预报)的分辨率高于预报(观测)时, 也可以将观测(预报)升尺度到预报(观测)的分辨率来进行检验<sup>[7]</sup>。UC法对观测或预报的处理结果在低于原始格(站)点分辨率的区域上而非在原始格(站)点上, 不改变变量类型。

运用到集合预报时<sup>[6, 27]</sup>, 先对各集合成员进行升尺度, 再用生成一般集合概率预报的方法, 即给定阈值事件发生的集合成员数除以总集合成员数, 得到该空间窗口事件发生的概率, 对观测场先升尺度再用给定阈值转化为二元概率场(0或1)。

### 1.2 邻域最大值法

邻域最大值(neighborhood maximum, NM)法<sup>[21]</sup>在判定一个格(站)点上事件是否发生时, 将判定的条件放宽到邻域中, 若邻域中最大值大于给定阈值, 则判定该格(站)点为事件发生。等价的, 若搜索到邻域任意一处的值大于给定阈值, 则判定该格(站)点为事件发生, 因此也可称为搜索法。NM法对观测或预报场的处理结果是在原始格(站)点上的二元场

(即事件发生或不发生)。

在利用列联表对预报进行检验时, 不同方法对是否使用NM方法对观测或预报进行宽松的判定做出不同规定。Atger<sup>[5]</sup>定义了一种对观测严格而对预报宽松的列联表规则(A01), Clark等<sup>[31]</sup>定义的列联表规则(C10)对预报和观测采取混合宽松判定, McMillen等<sup>[32]</sup>定义了一种对于预报严格而对观测宽松的列联表规则(M15), Schwartz<sup>[33]</sup>给出对预报和观测都采取宽松判定的列联表规则(S16)。4种方法的列联表规则如表3所示, 其中 $o_i=1$  ( $f_i=1$ )代表格点*i*上观测(预报)为事件发生,  $O_i=1$  ( $F_i=1$ )代表格点*i*的观测(预报)邻域内有事件发生, 值为0代表事件未发生。一般地, 上述列联表采用空间邻域, 预报和观测采用相同的强度阈值来判定事件是否发生, 称为空间列联表法。Atger<sup>[5]</sup>对A01规则拓展了列联表的维度, 对于给定强度的观测阈值, 取不同强度的预报阈值(例如对于5 mm的观测阈值, 分别取1, 2, ..., 20 mm的预报阈值), 生成多个列联表, 对每个列联表求出命中率和空报率并通过ROC图和RV来评估预报质量, 该方法称为空间-多事件列联表法(space multi-event contingency table, SMCT)。

表3 邻域最大值方法中4种将格点*i*划分为列联表不同类别的规则

Table 3 Criteria for filling the contingency tables for the *i*th grid point using the four neighborhood maximum methods

规则	命中	空报	漏报	正确否定
A01	$o_i=1, F_i=1$	$o_i=0, F_i=1$	$o_i=1, F_i=0$	$o_i=0, F_i=0$
C10	$o_i=1, F_i=1$ 或 $f_i=1, O_i=1$	$O_i=0, f_i=1$	$o_i=1, F_i=0$	$o_i=0, f_i=0$
M15	$O_i=1, f_i=1$	$O_i=0, f_i=1$	$O_i=1, f_i=0$	$O_i=0, f_i=0$
S16	$O_i=1, F_i=1$	$O_i=0, F_i=1$	$O_i=1, F_i=0$	$O_i=0, F_i=0$

NM方法应用于集合预报时, Schwartz等<sup>[21]</sup>首先对各集合成员采用NM法, 得到给定阈值的二元预报, 再对所有成员平均得到概率预报, 对观测也使用NM方法进行处理, 而后采用概率预报检验方法进行检查。而Atger<sup>[5]</sup>采用SMCT法, 在不同强度阈值的基础上, 再加入不同发生概率的阈值(至少有1, 2, ..., *m*个集合成员预报了事件的发生), 对不同阈值组合分别生成列联表并通过RV来评估预报表现。

### 1.3 格点概率法

格点概率(grid probability, GP)法根据邻域中事件发生的分布, 采用一种数学方法来估计格点上发生事件的概率(或称分数), 通过比较观测和预报概率的差异来评估预报的质量。GP法对观测或预报场的处理结果是在原始格点上的概率(分数)场。

Davis等<sup>[34]</sup>最早提出, 假设一个格点周围邻域内所

有格点预报事件发生的可能性相同, 则该点预报事件发生的概率为邻域中发生事件的格点数除以邻域内总格点数。本文将这种计算格点概率的方法称为邻域概率平均法 (neighborhood probability averaging, NPA)。不同于NPA法, 高斯平滑法<sup>[16-17, 35]</sup>给不同格点分配不同权重来计算格点概率。两种方法本质上都是对区域内的概率进行平滑, 通过取不同的邻域大小或平滑参数来反映预报质量随尺度的变化, 但NPA法因其简单易用而更为常用<sup>[4, 22-23, 36-37]</sup>。

对预报进行邻域处理生成格点概率后, 若直接用给定阈值将观测转化为二元概率 (0或1), 则可用BSS等概率检验指标进行检验。若对观测也采用邻域处理得到分数型的概率, 则还可使用最低概率 (minimum coverage, MC) 法和分数技巧评分 (fractions skill scores, FSS) 进行检验。MC法<sup>[24]</sup>用给定阈值 (如50%) 将观测和预报概率转化为二元场, 使用二元检验指标进行检验。当最低概率为最宽松的条件 (至少有一个格点有事件发生) 时, 该方法与NM方法中的S16空间列联表法是等效的, 但NM方法不用计算格点上的概率。Robert等<sup>[4]</sup>提出的FSS, 能直接反应预报在哪个尺度上是有技巧的, 是目前最常用的邻域检验指标之一<sup>[23, 37-42]</sup>。在求得格点概率后, 计算邻域大小为 $n \times n$ 时观测和预报格点概率 (分数) 场的均方误差:

$$FBS_{(n)} = \frac{1}{N} \sum_{i=1}^N [f_{(n)}(i) - o_{(n)}(i)]^2, \quad (1)$$

式中,  $f(i)$ 和 $o(i)$ 分别为预报和观测在格点 $i$ 上的概率;  $N$ 为检验区域总格点数;  $n$ 为邻域大小。FBS依赖于观测降水频率, 因此用一个低技巧的参照来计算FSS:

$$FSS_{(n)} = \frac{FBS_{(n)} - FBS_{(n)reference}}{FBS_{(n)perfect} - FBS_{(n)reference}} = 1 - \frac{FBS_{(n)}}{FBS_{(n)reference}}, \quad (2)$$

式中,  $FBS_{reference}$ 为预报与观测概率场的最大可能均方误差:

$$FBS_{(n)reference} = \frac{1}{N} \left[ \sum_{i=1}^N o_{(n)}^2(i) + \sum_{i=1}^N f_{(n)}^2(i) \right]. \quad (3)$$

FSS值在0~1, 值越大预报技巧越高。当检验区域足够大时, FSS通常随着邻域的增大而增大,  $FSS_{(n)}$ 首次大于 $FSS_{uniform}$ 时, 认为尺度 $n$ 是预报有相当技巧的最小尺度, 其中 $FSS_{uniform} = 0.5 + v_o/2$ ,  $v_o$ 为检验区域中观测的事件发生频率。FSS的结果可使用折线图或分数图<sup>[43]</sup>等展示。

此外, 可以仅对观测使用高斯平滑法<sup>[35]</sup>或NPA法<sup>[24]</sup>得到观测的分数型概率, 称为“近乎完美的” (practically perfect, PP) 预报<sup>[35]</sup>, 可将PP预报与观测计算得到的检验指标作为预报与观测所得检验指标的

参考, 从而间接地反映预报的质量<sup>[35, 44]</sup>。用不同概率阈值, 将PP预报转化为二元预报, 生成不同概率阈值下PP预报与观测的列联表并计算二元检验指标。概率阈值趋近0% (即预报员预报事件可在任何位置发生) 时的评分即为合理预报能得到的最低评分, 评分随概率阈值先增大后减小, 最大值即为预报预期能达到的最高评分。原始预报的评分在PP预报最大和最小评分间的相对大小为预报技巧。

GP法应用于集合预报时<sup>[21, 37, 45]</sup>, 先对各集合成员计算格点的邻域概率, 再对格点上所有集合成员的概率值进行平均。对于NPA法, 等价的方法是对所有格点求集合概率, 再对集合概率求邻域概率。对观测可选择是否进行邻域处理。

## 1.4 其他方法

Stein等<sup>[46]</sup>定义了一种基于邻域、包含误差补偿的列联表 (SS)。对于一个格点, 考虑其周围边长为 $n$ 的邻域范围内的预报和观测情况, 当邻域中有一对空报和漏报时, 移动预报事件的位置从而将空报和漏报转化成命中和正确否定, 得到该格点考虑误差补偿的局地列联表:

$$t_{(n)} = \begin{bmatrix} a + \min(b, c) & b - \min(b, c) \\ c - \min(b, c) & d + \min(b, c) \end{bmatrix}, \quad (4)$$

式中,  $a, b, c, d$ 分别表示命中、空报、漏报和正确否定。将各格点的局地列联表除以 $n^2$ 并累加得到整个检验区域的列联表:

$$T_{(n)} = \sum_{i=1}^N \frac{t_{(n)}(i)}{n^2}, \quad (5)$$

式中,  $N$ 为检验区域总格点数, 这样单个格点对区域列联表的贡献仍为1, 但却在四个类别中重新分配了。

Rezacova等<sup>[47]</sup>定义了一个区域相关RMSE。对于检验区域的每一个格点, 将格点邻域内的观测和预报值分别按从小到大的顺序排列, 然后计算这两列有序数列的RMSE, 即为该格点在该邻域尺度的区域相关RMSE, 整个检验区域的平均区域相关RMSE代表预报的整体误差大小。

## 2 不同邻域法的比较

为了对上述邻域法进行比较, 本节使用不同邻域法对2007年6月20日15—18时 (世界时, 下同) 3 h累计降水的预报进行检验。检验的预报和观测资料均来自空间检验方法比较计划MesoVICT<sup>[1]</sup>资料集, 确定性预报选取COSMO-2模式的18 h预报, 观测为VERA分析资料, 集合预报选取COSMO-LEPS模式的6 h预报, 共16个集合成员, 观测为集合VERA中的参考分析。确定性和集合预报分别被插值到对应的观测网格, 分

分辨率均为8 km, 检验区域为观测和模式范围重合的部分, 分别包含 $137 \times 88$ 和 $154 \times 98$ 个格点。使用边长为1 (即传统的点对点检验)、3、5、9、17和33个格点的方形邻域进行检验, 对应空间尺度分别为8、24、40、72、136和264 km。本节仅选择表1、表2中给出的部分检验指标进行检验。

首先对确定性预报进行检验。如图1a所示, 检验区域中部的阿尔卑斯山脉及其以北地区有强对流, 图1b显示模式较好地预报了沿地形的降水, 但预报的雨带较观测偏东且强度偏大。

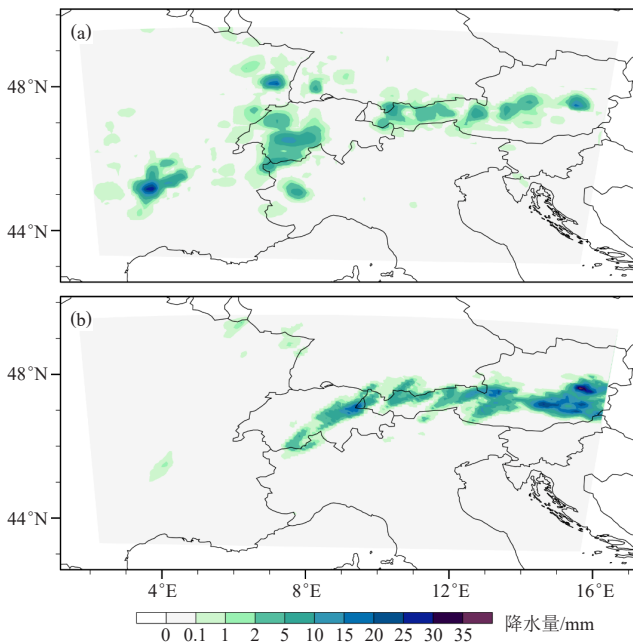


图1 2007年6月20日15—18时观测降水量 (a) 和COSMO-2模式的18 h预报降水量 (b)

Fig. 1 Observed (a) and COSMO-2 model 18-h forecast (b) rainfall during 15:00—18:00 UTC on 20 June 2007

升尺度方法保持变量类型不变, 可以给出图片以便直观了解升尺度后降水的分布情况。如图2, 在使用平均/最大值升尺度法将预报和观测升尺度到更大邻域尺度后, 弱/较强降水的分布直观看来更加相似。可用RMSE对降水预报误差大小进行衡量。如表4所示, 在平均到更大尺度后, 预报和观测平均降水量的差异越来越小, 而MU法的RMSE随邻域尺度的增大而增大, 这与图1中东北和西南两处明显的最大值差异相对应。表4中还给出区域相关RMSE随邻域大小的变化情况, 可知预报误差随邻域尺度增加略有增大。

阈值-邻域尺度图是运用较广的检验结果呈现形式, 其横纵坐标分别为变化的阈值和邻域空间尺度, 填色为使用的检验指标的值, 传统二元检验指标、FSS、BSS等指标都可以使用阈值-邻域尺度图来

呈现。多数邻域方法可建立列联表使用二元检验指标进行检验, 图3给出不同邻域法得到的ETS的降水阈值-邻域尺度图, 当观测和预报都没有超过阈值的格点时, ETS无法计算,  $ETS < 0$ 代表预报相对于随机预报没有技巧, ETS越大预报越好。图3表明, 20 mm以上降水的预报仅在使用SS法 (图3h) 检验时在264 km的邻域上略优于随机预报, 除此之外与随机预报相比均没有技巧, 这与图1中降水极大值预报位置的明显偏差相对应, 下面对图3的讨论不包括20 mm阈值。AU法得到的ETS (图3a) 在阈值为0.1 mm时随邻域尺度增大而增大, 对于1 mm和2 mm的降水阈值, ETS随尺度先增大再减小, 空间的平均减弱了局地的强降水并扩大了弱降水的范围, 使得在一定范围的更大尺度上较弱降水分布与观测更为相似, 而对于更高阈值的降水, 原始格点的ETS最高, 空间平均平滑了强降水信息。MU法得到的ETS (图3b) 多在邻域尺度较大时取值较大, 因为邻域越大越容易满足该区域有事件发生, 而左上角ETS为0是因为观测全部为事件发生。图3c和3e显示, A01和M15方法得到的ETS整体较低, 且在多数阈值下随邻域尺度不是单调变化的, 这是因为A01/M15方法将预报/观测判定放宽到邻域, 命中随尺度增大而增加, 空报不断增加/减少, 漏报不断减少/增加, 因此ETS的变化是不确定的。C10方法的ETS (图3d) 整体远高于其他方法, ETS随邻域尺度增加而增大, 因为当邻域尺度增大时, 命中增加, 空报、漏报减少而正确否定保持不变, 邻域尺度为264 km时预报对1 mm降水的落区全部报对。S16方法得到的ETS (图3f) 在阈值为0.1 mm时随邻域尺度的增大先增大后减小, 这主要是因为尺度较大时空报过多, 其他降水阈值的ETS随邻域尺度的变化与C10类似, 根据S16的判定条件, 随尺度的增大, 命中增加而正确否定减小, 因此ETS值比较适中。使用格点概率法对观测和预报进行处理得到概率场, 可用MC法进行检验, 图3g给出使用NPA法计算概率, 最低概率阈值为50%时的结果, 阈值为0.1 mm时ETS随邻域尺度先增大后减小, 其他阈值时ETS随着尺度的增大而减小, 这是因为邻域中一半以上格点为事件发生的条件越来越难以达到导致命中率减小。SS方法得到的ETS (图3h) 随邻域尺度增大而增大, 阈值较大时增大幅度较小, 这是因为预报在较大阈值有明显高估 (更多空报, 更少漏报), 较难通过增大邻域尺度来获得更多局地补偿。图3i中PP预报是由NPA方法得到的, 随着邻域尺度增大, PP预报与原始观测的差异加大, 预报的相对质量提高。

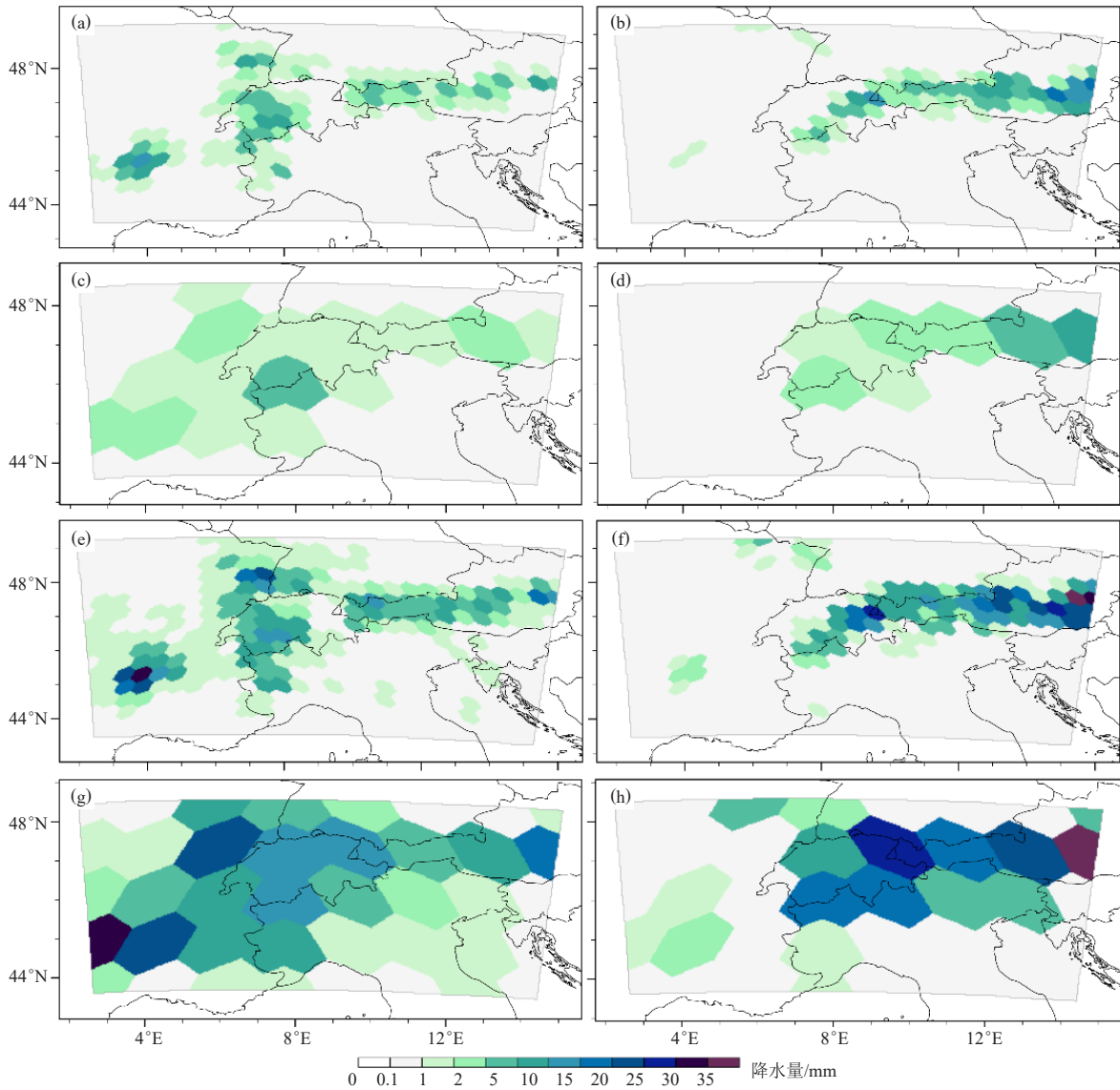


图2 观测降水 (a) 和COSMO-2模式的18 h预报降水 (b) 使用平均值升尺度法得到的邻域尺度为40 km的降水分布; (c~d) 同 (a~b), 但邻域尺度为136 km; (e~f) 同 (a~b), 但为使用最大值升尺度法得到的结果; (g~h) 同 (e~f), 但邻域尺度为136 km

Fig. 2 Rainfall distribution at a scale of 40 km using averaging upscaling method for the observed rainfall (a) and COSMO-2 model 18-h forecast rainfall (b); (c-d) the same as (a-b), but at a scale of 136 km; (e-f) the same as (a-b), but using maximum upscaling; (g-h) the same as (e-f), but at a scale of 136 km

表4 区域相关RMSE法、平均值升尺度法和最大值升尺度法检验COSMO-2模式预报得到的不同邻域空间尺度下的RMSE  
Table 4 RMSEs for forecast from COSMO-2 model using area-related RMSE, averaging upscaling and maximum upscaling under different spatial scales

邻域空间尺度/km	区域相关RMSE法	平均值升尺度法	最大值升尺度法
8	--	2.45	2.45
24	0.70	2.15	3.82
40	0.76	1.90	5.06
72	0.88	1.64	6.34
136	1.05	1.33	8.92
264	1.34	0.89	15.10

对观测和预报使用NPA法时可用FSS进行检验, 如图4a, FSS基本随邻域尺度增大而增大, 但在阈值为20 mm时, FSS全为0, 表明不论考虑多大尺度的邻域, 20 mm以上降水的预报都没有技巧。对于0.1、1、2、5和10 mm的降水阈值,  $FSS_{uniform}$  的值分别为0.67、0.57、0.54、0.51和0.50, 由图4a可知, 对于1 mm的降水阈值, 预报在136 km的尺度上具有相当技巧, 对于0.1和2 mm的降水阈值, 预报在264 km的尺度上才具有相当技巧, 而对于5和10 mm的降水阈值, 预报未达到相当技巧。

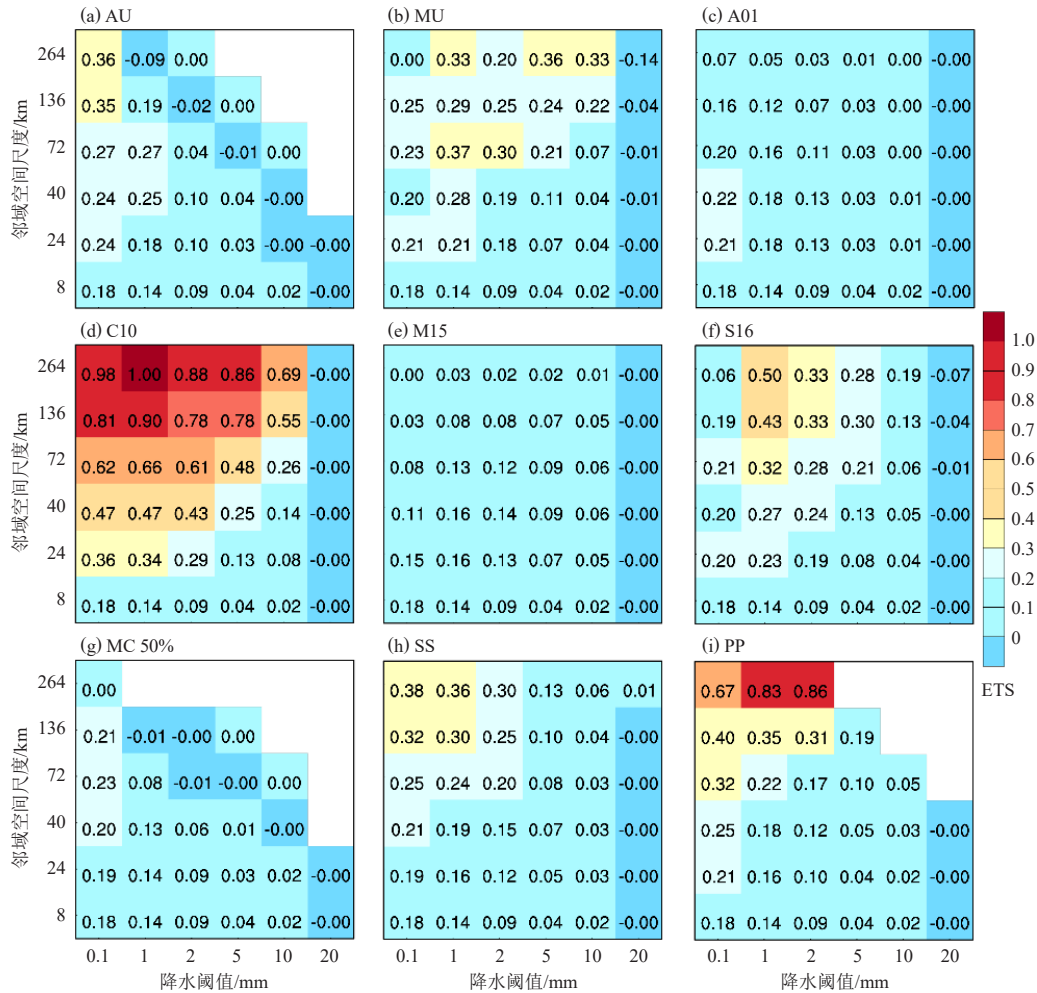


图3 使用平均值升尺度法 (a)、最大值升尺度法 (b)、A01空间列联表法 (c)、C10空间列联表法 (d)、M15空间列联表法 (e)、S16空间列联表法 (f)、最低概率(50%)法 (g)、SS列联表法 (h) 和近乎完美预报法 (i) 检验COSMO-2模式预报得到的ETS评分

Fig. 3 ETS for COSMO-2 model forecast verification using averaging upscaling method (a), maximum upscaling method (b), A01 spatial contingency table method (c), C10 spatial contingency table method (d), M15 spatial contingency table method (e), S16 spatial contingency table method (f), minimum coverage (50%) method (g), SS contingency table method (h) and practically perfect forecast method (i)

仅对预报使用NPA法可用BSS进行检验, 这里将原始预报的BS评分作为参考计算BSS。由图4b可知, BSS的值基本随邻域尺度增大而增大, 且降水阈值越高BSS越大, 使用NPA法得到的概率预报相对于原始预报更有技巧, 表明对于较高的降水阈值, 对邻域中概率进行平均能明显改善预报质量。

下面对集合概率预报进行检验。如图5所示, 集合概率预报较好地反映了检验区域中部的弱降水位置, 但区域北侧和西侧降水分布与观测有明显差异, 对于10 mm以上的降水, 集合概率未能提供较好的预报。常用于集合概率预报检验的4种邻域法为AU、MU、NM和NPA法, 用概率预报检验指标进行检验, 生成ROC图时, 概率阈值取0.05,

0.15, ..., 0.95, 生成可靠性图时, 将概率预报分为[0, 0.05), [0.05, 0.15), ..., [0.95, 1]等11个区间<sup>[48]</sup>。

图6a~6d给出不同邻域法得到的BSS, 将预报概率为样本观测频率的预报作为参考预报计算BSS。AU法在阈值为0.1、5、10、20 mm时BSS基本为正, 在大阈值和大邻域尺度上有很高的BSS, 是因为对区域取平均值后, 强降水的预报概率和观测频率均接近0。MU法和NM法在阈值为0.1 mm时以及在格点尺度上BSS基本不为负值, 其他阈值及邻域尺度集合概率预报技巧都不如概率为观测频率的预报。NPA法仅在阈值为0.1 mm、邻域尺度为264 km时BSS为正值。

对观测也使用NPA法可使用FSS来进行检验, 如图6e所示, FSS随邻域尺度增大而增大, 根据

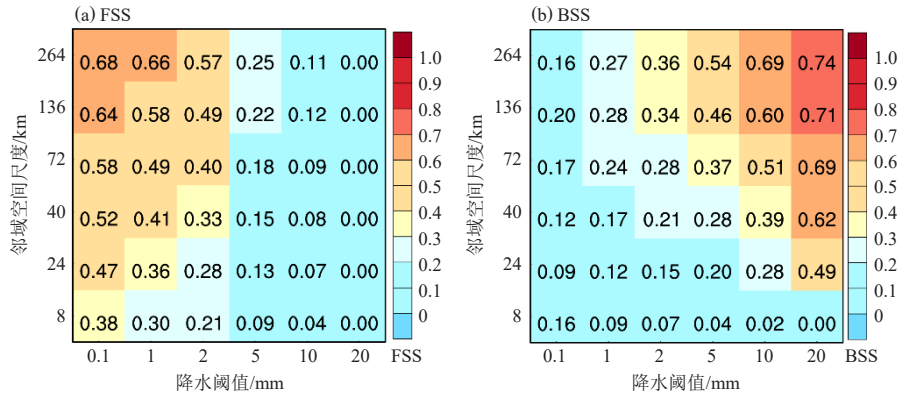


图4 使用邻域概率平均法对COSMO-2模式预报检验得到的FSS评分 (a) 和BSS评分 (b) 的降水阈值-邻域尺度图  
Fig. 4 Intensity-scale diagrams of FSS (a) and BSS (b) for COSMO-2 model forecast verification using neighborhood probability averaging methods

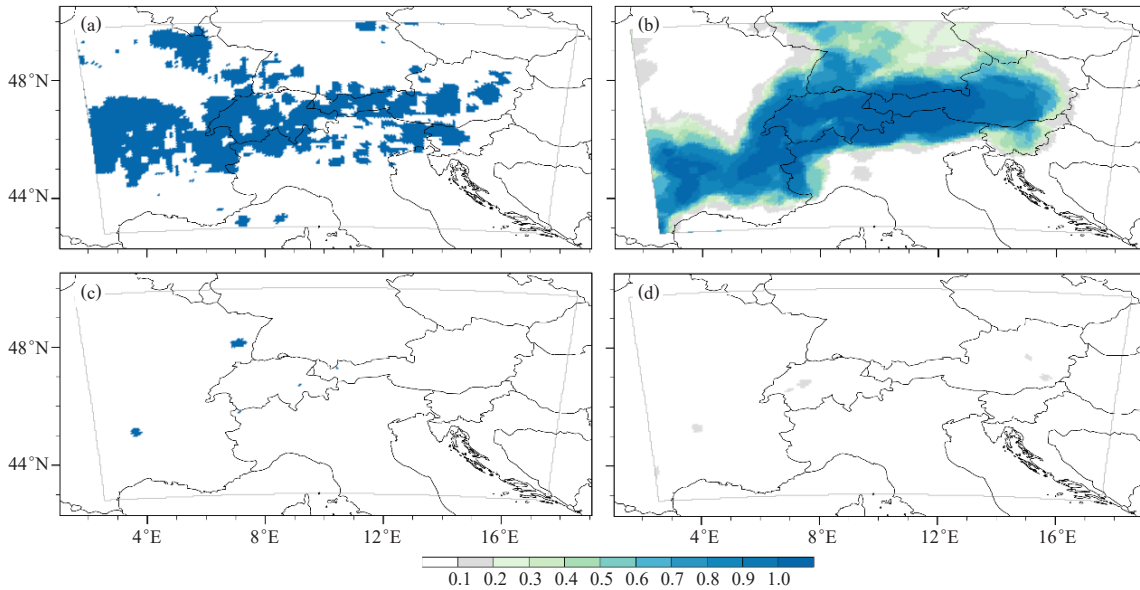


图5 观测在阈值为0.1 mm (a)、10 mm (c) 时的频率分布, COSMO-LEPS模式在阈值为0.1 mm (b)、10 mm (d) 时的集合概率预报  
Fig. 5 Observed frequency distribution for threshold 0.1 mm (a) and 10 mm (c); Probability distribution for probability forecast derived from COSMO-LEPS ensemble forecasts for threshold 0.1 mm (b) and 10 mm (d)

FSS<sub>uniform</sub>, 对0.1、1和2 mm的阈值, 预报分别在40、136和264 km尺度上是有相当技巧的, 其他阈值的预报则未达到相当技巧。

ROC曲线下面积可定量反映预报分辨事件和非事件的能力, ROC面积超过0.7时认为预报是有用的<sup>[21, 49]</sup>。由图7可知, 对于AU法和NPA法 (MU和NM法), 2 (5) mm以下降水的概率预报在不同邻域尺度下都是有用的。对于较弱降水, NM和MU法所得ROC面积在不同邻域尺度差异不大, 对于5和10 mm降水, NM和MU在大邻域尺度得到的概率预报分辨力远高于原始集合概率预报。

可靠性图可以反映预报的可靠性, 曲线越接近对

角线表明预报越可靠。由于升尺度方法在大邻域尺度时格点太少, 故不作比较。由图8可知, NM法得到的概率预报的可靠性在阈值为1和5 mm时均高于NPA法和原始集合概率预报, NPA法得到的概率预报在5 mm时可靠性基本高于原始集合概率预报。

由上述对特定预报的检验可知, 不同邻域用法不同, 所得检验结果有较大差异, 检验时应当根据需求进行选择。结合文献调研, 此处对各方法的特点和用途进行总结和比较。AU法衡量预报与观测在一定大小区域内平均值的一致程度, 比较适用于流域平均或总降水量的预报检验<sup>[28]</sup>, 而MU法衡量一定范围内观测与预报最大值的一致程度, 适用于评估预报

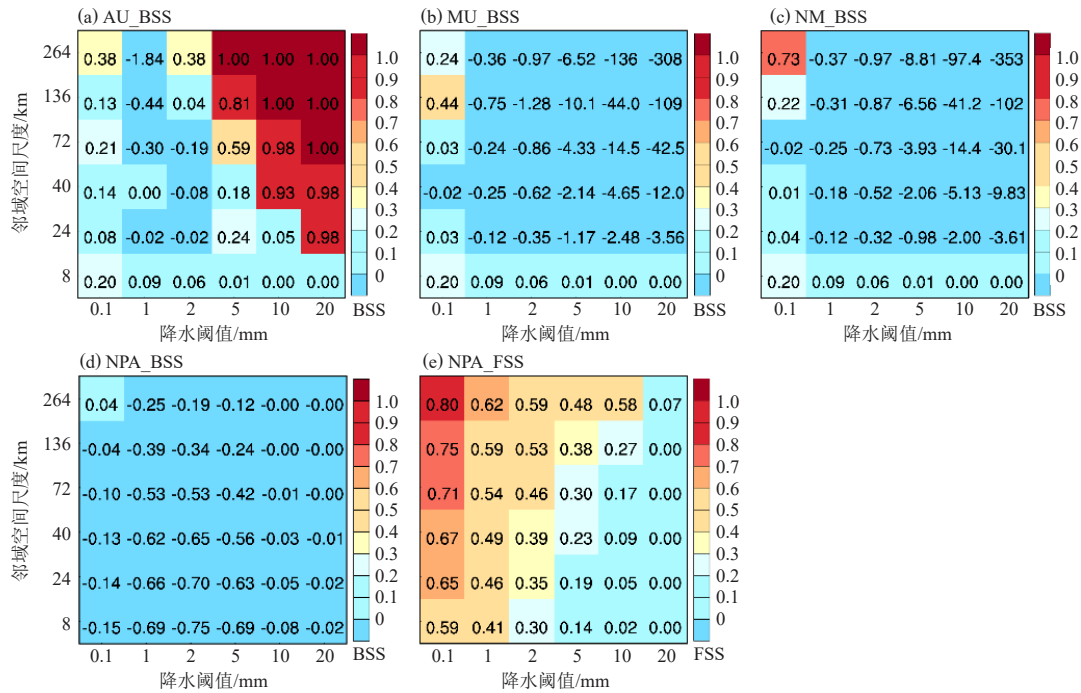


图6 使用平均值升尺度法 (a)、最大值升尺度法 (b)、邻域最大值法 (c)、邻域概率平均法 (d) 对COSMO-2模式预报检验得到的BSS评分以及使用邻域概率平均法得到的FSS评分 (e) 的降水阈值-邻域尺度图

Fig. 6 Intensity-scale diagrams of BSS for COSMO-2 model forecast verification using averaging upscaling method (a), maximum upscaling method (b), neighborhood maximum method (c), neighborhood probability averaging method (d) and FSS (e) using neighborhood probability averaging

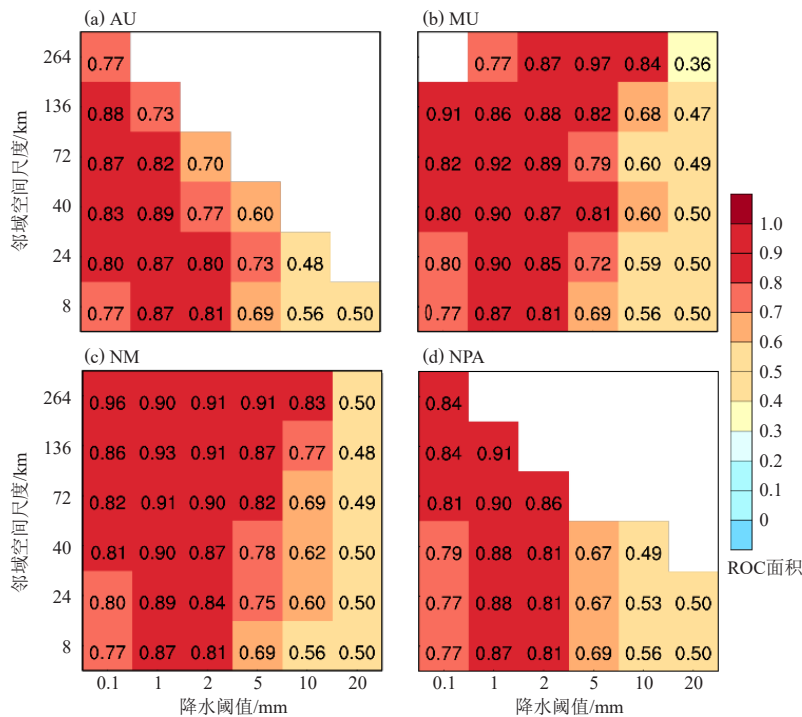


图7 使用平均值升尺度法 (a)、最大值升尺度法 (b)、邻域最大值法 (c)、邻域概率平均法 (d) 对COSMO-2模式预报检验得到的ROC面积的降水阈值-邻域尺度图

Fig. 7 Intensity-scale diagrams of ROC area for COSMO-2 model forecast verification using averaging upscaling method (a), maximum upscaling method (b), neighborhood maximum method (c), and neighborhood probability averaging method (d)

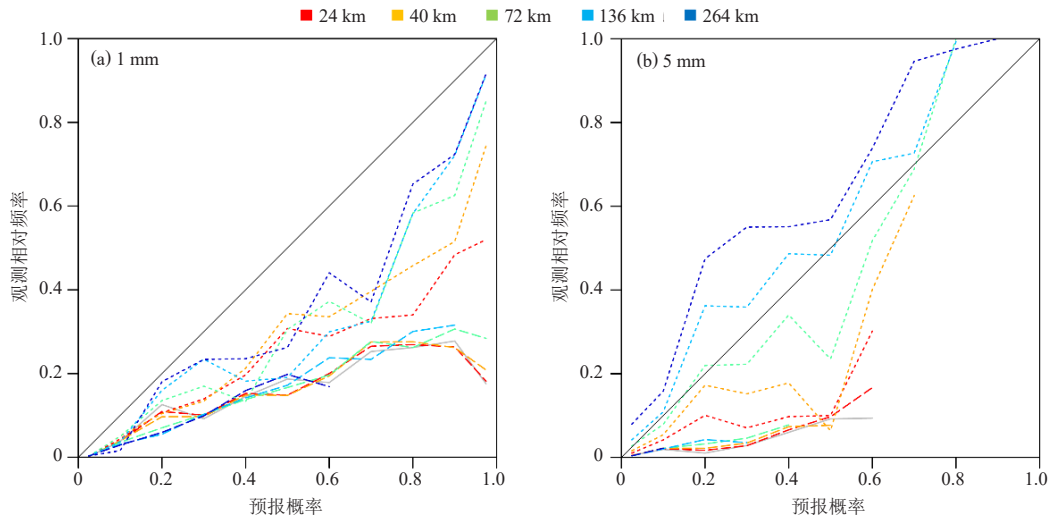


图8 降水阈值为1 mm (a)和5 mm (b)时,原始集合概率预报(灰色实线)以及使用邻域概率平均法(虚线)和邻域最大值法(点线)处理得到的概率预报的可靠性图解(不同颜色代表不同邻域尺度)

Fig. 8 Reliability diagrams for thresholds 1 mm (a) and 5 mm (b). The grey solid line represents the original COSMO-LEPS ensemble forecast reliability curves. The dashed lines and dotted lines correspond to probabilistic forecasts derived by neighborhood probability averaging method and neighborhood maximum method respectively. Different colors represent different neighborhood scales

能否为一片区域提供有效预警信息<sup>[6, 30]</sup>。NM方法中4种用于确定性预报检验的不同空间列联表规则A01、C10、M15和S16差异较大, C10方法过于宽松,而A01、M15方法又过于严苛, S16方法则较为适中<sup>[33]</sup>。A01/M15对观测严格/宽松而对预报宽松/严格,这种不对称的判定导致更容易产生空报/漏报而不易产生漏报/空报,使得指标随邻域尺度的变化不总是单调的。A01方法拓展得到的SMTC方法灵活性较大,通过对观测和预报选取不同阈值,可以更充分地获取原始预报中的信息,为不同需求的用户提供参考,适用于单个位置点(城市、机场等)预报的检验。M15方法可用于检验站点主观确定性预报<sup>[50-52]</sup>, C10方法较常应用于强降水预报检验<sup>[53-56]</sup>。Schwartz<sup>[33]</sup>指出S16方法对观测和预报的对称处理要比其他空间列联表方法更为公正。SS法的优势在于其对预报偏差的处理, SS法的BIAS评分不随邻域尺度的变化而变化,且局地补偿总是在预报无偏时最大,因此不会产生偏强的预报评分优于无偏预报的现象<sup>[57]</sup>。PP法以近乎完美的预报作为参考,尤其适用于小概率事件预报技巧的评估。格点概率方法是最常用的邻域法之一,一方面可评估预报质量<sup>[4, 12, 37, 39-41]</sup>,另一方面其产生的概率预报可作为概率指导产品并受到检验<sup>[16-17, 21-22, 36, 45]</sup>,不过,格点概率方法的本质是对概率进行平滑,会不可避免地降低预报的锐度<sup>[30]</sup>,且可能会降低小概率事件预报的可靠性或分辨力,因此较少被用于检验小概率事件的预

报。NM法和MU法都是利用一定大小区域中最大值是否超过给定阈值来判断事件是否发生,不同在于NM法判定格点上事件是否发生,而MU法则是判定大于格点的区域上事件是否发生。使用最大值的邻域方法比使用平均值的方法更为宽松,更适用于小概率事件预报能力的评估<sup>[30, 58]</sup>。需要注意的是,各邻域法通常在更大邻域尺度上检验评分更高,但这也忽略了高分辨率的信息,与使用高分辨率模式的初衷相悖,因而在实际使用时应当根据需求选取适当的邻域尺度检验范围。

### 3 总结与展望

邻域法放松了传统检验中观测与预报在格点上精确匹配的要求,认为在空间、时间、强度等方面与观测接近的预报是有效的。邻域法简单易行,能显著改善双重惩罚问题,常被用来检验高分辨率模式在一段时间内的预报能力、比较不同预报(来自不同模式、参数化方案等)间的差异、考察模式分辨率的改变对预报效果的影响以及评估预报在多大尺度是有用的等等。另外,邻域法也被广泛应用于生成概率指导预报<sup>[17, 19, 21, 30, 37, 41, 59]</sup>。

本文对邻域法进行了回顾与梳理,按照对观测及预报处理方式的不同,将邻域法分为了升尺度法、邻域最大值法、格点概率法和其他方法共4大类,各种方法的匹配策略、常用的检验指标以及对集合预报的处理方式等在表1和表2中简要列出,不同邻域法适用

情形不同, 应根据需要选择合适的方法对预报进行检验。在使用邻域法时, 有以下问题需要注意: 1) 邻域方法繁多, 且不同作者对同种方法的命名存在不同, 因此在检验时必须详细说明邻域法是如何实施的, 而不是简单地声明“使用了某种邻域法”。表1和表2给出常用的邻域检验指标, 在实际检验时只需根据需求选取几个指标即可。2) 本文为了说明和比较邻域方法仅检验了一个时次的预报, 实际应用中更多的是检验一组预报。3) 通常来说, 预报在更大邻域尺度上检验评分更高, 但这也忽略了高分辨率的信息, 因此应当根据需求选择适当的邻域尺度。合适的邻域尺度取决于网格大小、天气系统类型、天气现象强度等, 因此单个邻域尺度不可能适用于所有预报, 通常的做法是采用一系列变化的邻域尺度, 变化方式可以是不等间距<sup>[46, 60]</sup>或等间距的<sup>[43, 61]</sup>, 邻域尺度的数量从几个<sup>[60]</sup>到几十个<sup>[61]</sup>均可, 通过检验结果随邻域尺度的变化可以判断预报在何种尺度表现更优<sup>[61-65]</sup>或在哪些尺度上是有用的<sup>[17, 31, 66-67]</sup>。4) 邻域法的采样方式在预报质量较差、邻域面积相对检验区域较大时可能会对检验结果有不可忽略的影响<sup>[43]</sup>, 对于升尺度方法, 区域的划分对结果有一定影响, 对于其他方法, 应当确保靠近检验区域边缘格点的邻域始终在检验区域内。5) 需要估测检验评分的置信区间时可采用重新抽样法<sup>[4, 21, 29, 46, 68]</sup>。6) 用百分数阈值来代替固定值阈值可以减小模式偏差带来的影响<sup>[4, 61]</sup>, 但在解释结果时不应忽略偏差的存在。

与其他空间检验方法相比较, 邻域法最显著的优势在于其操作简单且可以直接衡量预报在何种尺度有技巧<sup>[2]</sup>, 但邻域法与尺度分离法类似, 不能衡量预报的结构误差, 对距离误差敏感但不能直接衡量, 而基于特征的方法、变形场法和距离度量法能够直接衡量距离误差, 其中基于特征的方法还能够衡量结构误差<sup>[69]</sup>。值得关注和期待的是, 各种空间检验方法并不是完全割裂的, 一种方法中的技巧时常被用于改善另一种方法, 比如邻域法中的平滑步骤被用于基于特征的方法来更好地识别特征, 再比如将邻域法与尺度分离法相结合以改善尺度分离法中强度-尺度评分过低的问题<sup>[63]</sup>。邻域法及其他空间检验方法必然会有新的发展, 而本文整理的分类方式和常用检验操作仍具备一定的适用性和参考性。

尽管一些邻域法已被用于业务预报的常规检验, 比如我国采用预报点对观测邻域来检验站点主观降水预报<sup>[50-51]</sup>, 英国气象局采用FSS来检验模式高分辨率降水预报<sup>[57, 69]</sup>, 但其他邻域方法以及对集合预报的邻

域检验还不普遍, 多限于对某个模式某个区域某段时间的个别检验, 如何在实际业务预报中更好地运用各种邻域法提供的信息从而提高预报质量, 还需要更多思考与实践。

随着高分辨率模式尤其是对流可分辨模式的发展, 邻域法的应用将变得更加普遍。本文提到的诸多邻域方法多用于降水预报的检验, 未来可将更多种邻域方法运用到其他预报场的检验中去。本文侧重介绍了空间邻域的处理方法, 目前已有研究将M15方法<sup>[65]</sup>和NPA方法<sup>[22-23, 70-71]</sup>拓展至时间邻域, 其他邻域法也可用于处理时间邻域, 未来可进行更多探索。此外, 如何将邻域法与其它空间检验方法相融合, 充分利用各种方法的优势, 从而更加全面地检验预报将是重要且很有挑战性的工作。

#### 参考文献

- [1] Dorninger M, Gilleland E, Casati B, et al. The setup of the MesoVICT project. *Bulletin of the American Meteorological Society*, 2018, 99(9): 1887-1906.
- [2] Gilleland E, Ahijevych D, Brown B, et al. Intercomparison of spatial forecast verification methods. *Weather and Forecasting*, 2009, 24(5): 1416-1430.
- [3] Mass C, Owens D, Westrick K, et al. Does increasing horizontal resolution produce more skillful forecasts? *Bulletin of the American Meteorological Society*, 2002, 83(3): 407-430.
- [4] Roberts N, Lean H. Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Monthly Weather Review*, 2008, 136(1): 78-97.
- [5] Atger F. Verification of intense precipitation forecasts from single models and ensemble prediction systems. *Nonlinear Processes in Geophysics*, 2001, 8(6): 401-417.
- [6] Marsigli C, Montani A, Paccagnella T. A spatial verification method applied to the evaluation of high-resolution ensemble forecasts. *Meteorological Applications*, 2008, 15(1): 125-143.
- [7] Anna G, François L. Verifying precipitation forecasts using upscaled observations. *ECMWF Newsletter*, 2000, 87: 9-17.
- [8] Zepeda-Arce J, Foufoula-Georgiou E, Droegemeier K. Space-time rainfall organization and its role in validating quantitative precipitation forecasts. *Journal of Geophysical Research: Atmospheres*, 2000, 105(D8): 10129-10146.
- [9] Snook N, Jung Y, Brotzge J, et al. Prediction and ensemble forecast verification of hail in the supercell storms of 20 May 2013. *Weather and Forecasting*, 2016, 31(3): 811-825.
- [10] Gagne D, McGovern A, Brotzge J, et al. Day-ahead hail prediction integrating machine learning with storm-scale numerical weather models//*Innovative Applications of Artificial Intelligence Conference*. AAAI Press, 2015.
- [11] Stratman D, Coniglio M, Koch S, et al. Use of multiple verification methods to evaluate forecasts of convection from hot- and cold-start convection-allowing models. *Weather and Forecasting*, 2012, 28(1): 119-138.
- [12] Hitchcock S, Coniglio M, Knopfmeier K. Impact of MPEX upsonde observations on ensemble analyses and forecasts of the 31 May 2013 convective event over Oklahoma. *Monthly Weather Review*, 2016, 144(8): 2889-2913.
- [13] Snook N, Xue M, Jung Y. Multiscale EnKF assimilation of radar and conventional observations and ensemble forecasting for a tornadic mesoscale convective system. *Monthly Weather Review*, 2015, 143(4): 1035-1057.
- [14] Kober K, Craig G, Keil C, et al. Blending a probabilistic nowcasting method with a high-resolution numerical weather prediction

- ensemble for convective precipitation forecasts. *Quarterly Journal of the Royal Meteorological Society*, 2012, 138(664): 755-768.
- [15] Scheufele K, Kober K, Craig G, et al. Combining probabilistic precipitation forecasts from a nowcasting technique with a time-lagged ensemble. *Meteorological Applications*, 2014, 21(2): 230-240.
- [16] Sobash R, Kain J, Bright D, et al. Probabilistic forecast guidance for severe thunderstorms based on the identification of extreme phenomena in convection-allowing model forecasts. *Weather and Forecasting*, 2011, 26(5): 714-728.
- [17] Sobash R, Schwartz C, Romine G, et al. Severe weather prediction using storm surrogates from an ensemble forecasting system. *Weather and Forecasting*, 2015, 31(1): 255-271.
- [18] Yussouf N, Gao J, Stensrud D, et al. The Impact of mesoscale environmental uncertainty on the prediction of a tornadic supercell storm using ensemble data assimilation approach. *Advances in Meteorology*, 2013(2): 731647.
- [19] Clark A J, Gao J, Marsh P, et al. Tornado pathlength forecasts from 2010 to 2011 using ensemble updraft helicity. *Weather and Forecasting*, 2013, 28(2): 387-407.
- [20] Skok G, Hladnik V, Skok G, et al. Verification of gridded wind forecasts in complex Alpine terrain: A new wind verification methodology based on the neighborhood approach. *Monthly Weather Review*, 2017, 146(1): 63-75.
- [21] Schwartz C, Sobash R. Generating probabilistic forecasts from convection-allowing ensembles using neighborhood approaches: A review and recommendations. *Monthly Weather Review*, 2017, 145(9): 3397-3418.
- [22] Theis S, Hense A, Damrath U. Probabilistic precipitation forecasts from a deterministic model: A pragmatic approach. *Meteorological Applications*, 2005, 12(3): 257-268.
- [23] Duc L, Saito K, Seko H. Spatial-temporal fractions verification for high-resolution ensemble forecasts. *Tellus A: Dynamic Meteorology and Oceanography*, 2013, 65(1): 18171.
- [24] Ebert E. Fuzzy verification of high-resolution gridded forecasts: A review and proposed framework. *Meteorological Applications*, 2008, 15(1): 51-64.
- [25] Casati B, Wilson L, Stephenson D, et al. Forecast verification: Current status and future directions. *Meteorological Applications*, 2008, 15(1): 3-18.
- [26] Richardson D. Skill and relative economic value of the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 2000, 126(563): 649-667.
- [27] Marsigli C, Boccanera F, Montani A, et al. The COSMO-LEPS mesoscale ensemble system: Validation of the methodology and verification. *Nonlinear Processes in Geophysics*, 2005, 12(4): 527-536.
- [28] Yates E, Anquetin S, Ducrocq V, et al. Point and areal validation of forecast precipitation fields. *Meteorological Applications*, 2006, 13(1): 1-20.
- [29] Weusthoff T, Ament F, Arpagaus M, et al. Assessing the benefits of convection-permitting models by neighborhood verification: Examples from MAP D-PHASE. *Monthly Weather Review*, 2010, 138(9): 3418-3433.
- [30] Ben Bouallègue Z, Theis S. Spatial techniques applied to precipitation ensemble forecasts: From verification results to probabilistic products. *Meteorological Applications*, 2014, 21(4): 922-929.
- [31] Clark A, Gallus W, Weisman M. Neighborhood-based verification of precipitation forecasts from convection-allowing NCAR WRF model simulations and the operational NAM. *Weather and Forecasting*, 2010, 25(5): 1495-1509.
- [32] McMillen J, Steenburgh W. Capabilities and limitations of convection-permitting WRF simulations of lake-effect systems over the Great Salt Lake. *Weather and Forecasting*, 2015, 30(6): 1711-1731.
- [33] Schwartz C. A comparison of methods used to populate neighborhood-based contingency tables for high-resolution forecast verification. *Weather and Forecasting*, 2016, 32(2): 733-741.
- [34] Davis C, Carr F. Summary of the 1998 workshop on mesoscale model verification. *Bulletin of the American Meteorological Society*, 2000, 8(4): 809-819.
- [35] Hitchens N, Brooks H, Kay M. Objective limits on forecasting skill of rare events. *Weather and Forecasting*, 2013, 28(2): 525-534.
- [36] Johnson A, Wang X. Verification and calibration of neighborhood and object-based probabilistic precipitation forecasts from a multimodel convection-allowing ensemble. *Monthly Weather Review*, 2012, 140(9): 3054-3077.
- [37] Schwartz C, Romine G, Weisman M, et al. A real-time convection-allowing ensemble prediction system initialized by mesoscale ensemble Kalman filter analyses. *Weather and Forecasting*, 2015, 30(5): 1158-1181.
- [38] Zacharov P, Rezacova D. Using the fractions skill score to assess the relationship between an ensemble QPF spread and skill. *Atmospheric Research*, 2009, 94(4): 684-693.
- [39] Romine G, Schwartz C, Berner J, et al. Representing forecast error in a convection-permitting ensemble system. *Monthly Weather Review*, 2014, 142(12): 4519-4541.
- [40] Schwartz C, Romine G, Smith K, et al. Characterizing and optimizing precipitation forecasts from a convection-permitting ensemble initialized by a mesoscale ensemble Kalman filter. *Weather and Forecasting*, 2014, 29(6): 1295-1318.
- [41] Yussouf N, Kain J, Clark A. Short-term probabilistic forecasts of the 31 May 2013 Oklahoma tornado and flash flood event using a continuous-update-cycle storm-scale ensemble system. *Weather and Forecasting*, 2016, 31(3): 957-983.
- [42] Mittermaier M, Roberts N, Thompson S. A long-term assessment of precipitation forecast skill using the fractions skill score. *Meteorological Applications*, 2013, 20(2): 176-186.
- [43] Nachamkin J, Schmidt J. Applying a neighborhood fractions sampling approach as a diagnostic tool. *Monthly Weather Review*, 2015, 143(11): 4736-4749.
- [44] Germann U, Zawadzki I. Scale dependence of the predictability of precipitation from continental radar images. Part II: Probability forecasts. *Journal of Applied Meteorology*, 2004, 43(1): 74-89.
- [45] Schwartz C, Kain J, Weiss S, et al. Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Weather and Forecasting*, 2010, 25(1): 263-280.
- [46] Stein J, Stoop F. Neighborhood-based contingency tables including errors compensation. *Monthly Weather Review*, 2018, 147(1): 329-344.
- [47] Rezacova D, Sokol Z, Pesice P. A radar-based verification of precipitation forecast for local convective storms. *Atmospheric Research*, 2007, 83(2): 211-224.
- [48] Buizza R, Hollingsworth A, Lalaurette F, et al. Probabilistic predictions of precipitation using the ECMWF ensemble prediction system. *Weather and Forecasting*, 1999, 14(2): 168-189.
- [49] Mullen S, Buizza R. The impact of horizontal resolution and ensemble size on probabilistic forecasts of precipitation by the ECMWF ensemble prediction system. *Weather and Forecasting*, 2002, 17(2): 173-191.
- [50] 郑永光, 周康辉, 盛杰, 等. 强对流天气监测预报预警技术进展. *应用气象学报*, 2015, 26(6): 641-657.
- [51] 韦青, 李伟, 彭颂, 等. 国家级天气预报检验分析系统建设与应用. *应用气象学报*, 2019, 30(2): 245-256.
- [52] 陈法敬, 陈静, 韦青, 等. 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. *气象学报*, 2019, 77(1): 28-42.
- [53] Schumacher R, Clark A, Xue M, et al. Factors influencing the development and maintenance of nocturnal heavy-rain-producing convective systems in a storm-scale ensemble. *Monthly Weather Review*, 2013, 141(8): 2778-2801.
- [54] Dahl N, Xue M. Prediction of the 14 June 2010 Oklahoma City extreme precipitation and flooding event in a multiphysics multi-initial-conditions storm-scale ensemble forecasting system. (下转72页)

元数据可用性, 优化完善社会化保障业务支撑能力, 以试点带动省级本地化需求与平台的功能融合, 通过智能告警算法建立覆盖全面、投放精准的监控告警与业务引导策略等。在倡导国家级业务系统规范化、集

约化建设的大背景下, 平台持续升级优化功能、改善体验, 为我国综合气象观测业务的信息化建设提供可靠支撑。

### 深入阅读

杜建苹, 任杰, 2014. 气象装备全寿命自动跟踪编码设计. 气象科技, 42(5): 759-763.  
方海涛, 钱毅, 窦炜明, 2015. 省级气象装备动态管理系统设计与实现. 中国管理信息化, 18(7): 73-76.  
李峰, 秦世广, 周薇, 等, 2014. 综合气象观测运行监控业务及系统升级设计. 气象科技, 42(4): 539-544.  
梁海河, 孟昭林, 张春晖, 等, 2011. 综合气象观测运行监控系统. 气象, 37(10): 1292-1300.  
刘媛媛, 何文春, 王妍, 等, 2021. 气象大数据云平台归档系统设计及实现. 气象科技, 49(5): 697-706.  
裴翀, 宋连春, 吴可军, 等, 2011. 我国综合气象观测运行监控系统的设计与实践. 气象, 37(2): 213-218.  
气象观测装备编码规则 (GB/T 40215-2021). 北京: 国家市场监

督管理总局, 中国国家标准化管理委员会.  
宋智, 徐晓莉, 张常亮, 等, 2019. 应用分布式存储技术优化省级CIMISS数据服务能力. 气象科技, 47(3): 433-438.  
辛园园, 钮俊, 谢志军, 等, 2018. 微服务体系结构实现框架综述. 计算机工程与应用, 54(19): 10-17.  
杨紫超, 孙文博, 袁文辉, 2014. 3MS省级气象计量检定业务系统简介. 气象灾害防御, 21(3): 43-44.  
张文建, 2010. 世界气象组织综合观测系统 (WIGOS). 气象, 36(3): 1-8.  
质量管理体系GB/T 19001-2016应用指南 (GB/T 19002-2018). 北京: 国家市场监督管理总局, 中国国家标准化管理委员会.  
WMO, 2019. WIGOS metadata standard (WMO-No.1192).

(作者单位: 秦世广、刘健、李巍、邵楠、周薇、刘洁、石锐, 中国气象局气象探测中心; 刘健, 内蒙古自治区赤峰市气象局; 庞晶, 中国气象局)

(上接40页)

- Weather and Forecasting, 2016, 31(4): 1215-1246.
- [55] Squitieri B, Gallus W. WRF Forecasts of Great Plains nocturnal low-level jet-driven MCSs. Part I: Correlation between low-level jet forecast accuracy and MCS precipitation forecast skill. Weather and Forecasting, 2016, 31(4): 1301-1323.
- [56] Pytharoulis I, Kotsopoulos S, Tegoulis I, et al. Numerical modeling of an intense precipitation event and its associated lightning activity over northern Greece. Atmospheric Research, 2016, 169: 523-538.
- [57] Ebert E. Neighborhood verification: A strategy for rewarding close forecasts. Weather and Forecasting, 2009, 24(6): 1498-1510.
- [58] Hardy J, Gourley J, Kirstetter P, et al. A method for probabilistic flash flood forecasting. Journal of Hydrology, 2016, 541: 480-494.
- [59] Sobash R, Kain J, Bright D, et al. Probabilistic forecast guidance for severe thunderstorms based on the identification of extreme phenomena in convection-allowing model forecasts. Weather and Forecasting, 2011, 26(5): 714-728.
- [60] Ebert E. Fuzzy verification of high-resolution gridded forecasts: A review and proposed framework. Meteorological Applications, 2008, 15(1): 51-64.
- [61] Azorin-Molina C, Tijm S, Ebert E, et al. Sea breeze thunderstorms in the eastern Iberian Peninsula. Neighborhood verification of HIRLAM and HARMONIE precipitation forecasts. Atmospheric Research, 2014, 139: 101-115.
- [62] 罗聪, 时洋, 吴乃庚, 等. 基于邻域最优概率的GRAPES快速更新循环同化预报系统逐时降水订正方法. 热带气象学报, 2021, 37(4): 569-578.
- [63] 马申佳, 陈超辉, 智协飞, 等. 基于时空不确定性的对流尺度集合预报效果评估检验. 气象学报, 2018, 76(4): 578-589.
- [64] 刘雪晴, 陈静, 陈法敬, 等. 降水邻域集合概率方法尺度敏感性试验. 大气科学, 2020, 44(2): 282-296.
- [65] 张小雯, 唐文苑, 郑永光, 等. GRAPES\_3km数值模式对流风暴预报能力的多方法综合评估. 气象, 2020, 46(3): 367-380.
- [66] Skok G, Roberts N. Analysis of fractions skill score properties for random precipitation fields and ECMWF forecasts. Quarterly Journal of the Royal Meteorological Society, 2016, 142(700): 2599-2610.
- [67] Mittermaier M, Roberts N. Intercomparison of spatial forecast verification methods: Identifying skillful spatial scales using the fractions skill score. Weather and Forecasting, 2010, 25(1): 343-354.
- [68] Wolff J, Harrold M, Fowler T, et al. Beyond the basics: Evaluating model-based precipitation forecasts using traditional, spatial, and object-based methods. Weather and Forecasting, 2014, 29(6): 1451-1472.
- [69] Gilleland E, Ahijevych D, Brown B, et al. Verifying forecasts spatially. Bulletin of the American Meteorological Society, 2010, 91(10): 1365-1373.
- [70] 吴瑞姣, 邱学兴, 周昆, 等. 安徽省WRF模式短时强降水的预报检验. 气象科技, 2020, 48(2): 254-262.
- [71] Lockhoff M, Zolina O, Simmer C, et al. Evaluation of satellite-retrieved extreme precipitation over Europe using gauge observations. Journal of Climate, 2013, 27(2): 607-623.