

DOI: 10.13476/j.cnki.nsbdqk.2022.0092

黄赛男, 李文韬, 段青云. GEFSv12 降水再预报数据在淮河流域的适用性评估[J]. 南水北调与水利科技(中英文), 2022, 20(5): 925-934. HUANG S N, LI W T, DUAN Q Y. Evaluation of the performance of GEFSv12 precipitation reforecast dataset in the Huaihe River basin[J]. South-to-North Water Transfers and Water Science & Technology, 2022, 20(5): 925-934. (in Chinese)

GEFSv12 降水再预报数据在淮河流域的适用性评估

黄赛男^{1,2}, 李文韬^{1,2}, 段青云^{1,2}

(1. 河海大学水文水资源与水利工程科学国家重点实验室, 南京 210098;
2. 中国气象局-河海大学水文气象研究联合实验室, 南京 210098)

摘要: 新一代全球集合预报(global ensemble forecast system version 12, GEFSv12)是美国国家环境预报中心(National Centers for Environmental Prediction, NCEP)发布的最新的再预报数据集, 对该数据集中的降水数据在淮河流域的适用性进行评估。由于原始预报存在系统偏差, 所以应用贝叶斯联合概率模型(Bayesian joint probability, BJP), 对原始 GEFSv12 再预报数据进行统计后处理。预报验证的评价指标包括均方根误差(E_{RMS})、Brier 技巧评分(E_{BSS})、连续等级概率技巧评分(E_{CRPSS})、 α 指数及可靠性框图, 注重从预报偏差、精度以及可靠性 3 方面评价原始预报及经过后处理的预报。结果表明: GEFSv12 降水再预报数据对淮河流域具有较好的适用性, 在预见期 1~7 d 具有预报技能; BJP 统计后处理方法能有效降低预报系统性偏差, 增加预报精度以及预报可靠性, 经过后处理的降水集合预报将有利于进一步的水文集合预报应用。

关键词: 全球集合预报产品 GEFSv12; 再预报; BJP 统计后处理方法; 集合预报验证; 淮河流域

中图分类号: TV214 文献标志码: A 开放科学(资源服务)标识码(OSID):



在气象领域, 由于大气物理过程的复杂性以及大气系统的时间演变对系统初始状态的依赖性, 数值天气预报往往存在不确定性^[1], 通常采用初值扰动等方法生成一组集合成员来体现这种不确定性^[2]。在水文领域, 近年来水文气象耦合预报方法逐渐得到应用^[3-4], 将气象集合预报输入水文模型得到水文集合预报^[5]。相较于确定性预报而言, 集合预报能提高预报技巧, 并提供概率预报支持风险决策^[5-6], 在水文气象领域有着广泛的应用^[6-12]。

预报验证是评估集合预报质量的过程, 主要对预报和观测对的联合概率分布进行分析, 准确地验证在评估预报系统预报性能、表征预报系统的优势与不足及支持用户决策等方面有着重要作用^[1, 13]。预报评价指标主要分为确定性评价指标和概率评价指标两种, 可从偏差、关联性、精度、技巧、可靠

性、分辨率、锐度、分辨力以及不确定性 9 个方面展开^[14]。其中: 预报偏差常用指标包括相对偏差、均方根误差 E_{RMS} 等; 预报精度技巧方面包括 Brier 技巧评分 E_{BSS} ^[15]、连续等级概率技巧评分 E_{CRPSS} ^[16] 等; 预报可靠性方面包括 α 指数^[17]、可靠性框图^[1] 等。

GEFSv12 再预报数据集来自于美国国家环境预报中心(NCEP)的全球集合预报系统(GEFS), 由美国国家海洋与大气管理局(National Oceanic and Atmospheric Administration, NOAA)生成, 发布于 2020 年 9 月, 是 NCEP 最新的再预报数据集, 相较于前一版本的再预报数据集 GEFSv10, 分辨率从 0.50° 提升至 0.25°, 相对应的集合成员数从 11 个减少为 5 个。再预报数据是利用当前已有的预报模型回顾性地对过去十几年的天气进行重新预报, 是改善天气预

收稿日期: 2022-05-04 修回日期: 2022-07-25 网络出版时间: 2022-09-28

网络出版地址: <https://kns.cnki.net/kcms/detail/13.1430.TV.20220926.1856.005.html>

基金项目: 国家自然科学基金项目(42105142; 51979004; 41830752)

作者简介: 黄赛男(1997—), 女, 江苏南通人, 主要从事水文学及水资源研究。E-mail: 201301010067@hhu.edu.cn

通信作者: 段青云(1960—), 男, 湖南衡阳人, 教授, 博士, 主要从事水文集合预报及不确定性分析研究。E-mail: qyduan@hhu.edu.cn

报效果的重要数据集^[18],可用于水文气象领域的预报验证^[19]。

在淮河流域,已有 GEFSv10 再预报数据集的性能研究,证实 GEFS 再预报数据集在提高水文气象预报应用方面表现较好^[20]。然而,当前尚未有关于 GEFSv12 在淮河流域的预报性能评估研究,因而选取淮河流域为研究区,对该产品在淮河流域的适用性进行评价。由于初始条件和动力模型等方面存在误差,原始气象集合预报往往存在系统偏差,需要用统计后处理方法来进行校正^[1]。BJP^[21-22]统计后处理模型由 Wang 等提出,已成功应用于水文气象预报的校正和不确定量化问题。采用 BJP 统计后处理方法对 GEFSv12 原始降水集合预报产品进行统计后处理,借助多个评价指标评价原始 GEFSv12 集合预报及经过后处理的 GEFSv12 集合预报,并展现 BJP 统计后处理方法在校正偏差、改善预报精度技巧以及可靠性方面的效果。

1 研究区和数据源

淮河流域(30°55′~36°20′N, 111°55′~121°20′E)位于我国东部,流域面积约 27 万 km²。流域内包括淮河和沂沭泗河两大水系,流域由山丘区和平原(含湖泊和洼地)组成,其中流域西部、南部和东北部为山丘区,面积约占流域总面积的 1/3,见图 1。流域地处我国南北气候过渡带,多年平均降水量约 878 mm,受季风气候影响,流域内天气系统复杂多变,降水量年际变化大,年内分布极不均匀,汛期 6—9 月降水量约占全年的 50%~75%^[23]。

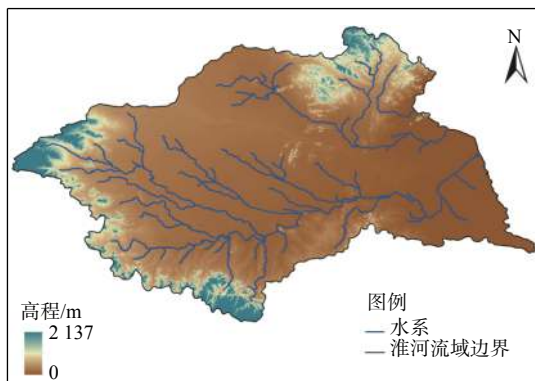


图 1 淮河流域区域概况

Fig. 1 The topography of the Huaihe River basin

使用的观测数据是中国 3 万多个自动气象观测站小时观测降水与 CMORPH 卫星反演降水产品融合的逐小时降水量网格数据集(1.0 版)^[24],可以通过国家气象科学数据中心下载(<http://data.cma.cn/site/>

showSubject/id/101.html)。数据融合的核心算法是概率密度匹配和最优插值融合概念模型,数据空间分辨率为 0.1°,时间序列长度为 2008 年 1 月 1 日至最新。研究使用的原始降水集合预报数据是 GEFSv12 再预报数据,下载自美国国家海洋和大气管理局(<https://noaa-gefs-retrospective.s3.amazonaws.com/index.html#GEFSv12/forecast/>)。GEFSv12 降水集合预报在每日的 00UTC 发布 1 次,以逐 3 小时与逐 6 小时累积降水交替的形式发布,预报时效 1~10 d 内对应的空间分辨率为 0.25°。该原始预报时间序列为 2000 年 1 月 1 日至 2019 年 12 月 31 日,集合成员为 5 个,其中包括 1 个控制预报和 4 个扰动成员。为获得更精细化的降水预报结果,使用双线性插值方法将 0.25°预报数据通过双线性插值到 0.10°,使降水预报数据与观测数据空间分辨率保持一致。有研究^[25]发现淮河流域雨季起止日期为第 24 候至 53 候,根据多年月平均降水量计算结果,得出 5—9 月是包含雨季在内且降水量最大的 5 个月。由于 2017 年 8 月以后的观测资料存在缺失值,研究选取 2008 年 1 月 1 日至 2017 年 7 月 31 日时段内每年的 5—9 月作为样本的时间范围,共计 10 a。

2 研究方法

统计后处理方法采用“留一交叉验证”方法,即将 10 a 数据中的 9 a 视为训练数据,1 a 作为验证数据,共遍历 10 次。对于后处理而言,以月为尺度进行后处理模型训练。以 5 月份为例,训练期为所选取的 9 a,训练期窗口为以 5 月 15 日为中心的 61 d,共计 9×61=549 d,输入后处理模型进行参数训练,验证期为剩余的 1 a,验证期对应的 5 月数据(1×31=31 d)输入训练好的后处理模型,获得该验证期后处理后的 5 月的预报值,遍历 10 次,获得 2008—2017 年经过后处理得到的 5 月份的预报数据。

对预报验证而言,通常情况下要选取气候态作为参考预报,具体方法为:基于观测样本采用留一法得到气候态,即 2008 年的集合预报为 2009—2017 年的观测值系列,集合成员数与后处理后的预报成员数保持一致,定为 25 个。例如 2008 年 5 月 1 日对应的集合预报为 2009 年 5 月 1 日、2010 年 5 月 1 日、…、2017 年 5 月 1 日当天及前后各 1 d,共计 3×9=27 d,取前 25 d 组成气候态集合。

2.1 BJP 统计后处理方法

采用 power 变换,对降水的原始集合预报均值

以及观测值分别进行正态化变换,公式为

$$x = z^p \quad (1)$$

式中: z 为原始集合预报均值或观测值; p 为未知参数通过极大似然法估计得到; x 为经正态化转换后的集合预报均值或观测值。对于含有大量零值的降水预报,将小于某阈值 z_c 的数据视为删失数据(此处定为 0.1 mm/d)。

进行联合概率分布建模。假设经过正态化变换的集合预报均值 x 以及观测值 y 服从二元正态分布:

$$\begin{bmatrix} x \\ y \end{bmatrix} \sim N \left(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix} \right) \quad (2)$$

式中: μ_x 、 σ_x 分别为经过正态化变换后的预报的边缘分布的均值和方差; μ_y 、 σ_y 为正态化变换后的观测的边缘分布的均值和方差;未知参数为变换空间内预报和观测的相关系数 ρ ,可通过极大似然法估计。

得到观测的条件概率分布及生成后处理后的集合预报。给定预报条件下观测的条件分布,即经过后处理的概率预报为

$$p(y|x) \sim N \left(\mu_y + \rho \frac{\sigma_y}{\sigma_x} (x - \mu_x), (1 - \rho^2) \sigma_y^2 \right) \quad (3)$$

抽取 1 000 个正态分布随机数,获得 1 000 个集合成员,取等间距的 25 个分位数,即 2%、6%、...、94%、98%分位数,获得 25 个集合成员组成的后处理后的集合预报。采用 power 变换的逆变换将集合预报变换至原始空间,按照 Schaake shuffle 方法^[26]对降水集合预报进行排序,使得降水集合预报具有合理的时空相关性。

2.2 评价指标

在所研究的时间范围内,通过每日的预报和观测构成样本对来计算指标。数值型预报验证指标,包括 E_{RMS} 、 E_{BSS} 、 E_{CRPS} 及 α 指数,在淮河流域的格点尺度上进行评估,评价成果用计算指标的中值以及[10%,90%]分位数区间表示。图形预报验证指标,包括可靠性框图,是在淮河流域的流域尺度上进行评估。

2.2.1 均方根误差

均方根误差(E_{RMS})衡量了集合预报均值与观测值之间的误差总体水平,取值范围为 $[0, +\infty]$,值越接近 0,结果越好。

$$E_{\text{RMS}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\bar{f}_i - o_i)^2} \quad (4)$$

式中: n 为预报和观测的总样本对数量; \bar{f}_i 为第 i 个集合预报均值; o_i 为第 i 个观测值。

2.2.2 Brier 评分以及 Brier 技巧评分

Brier 评分(E_{BS})考虑对观测值大于或等于某个阈值的降水事件的预报评价指标,观测对应的二值事件为 1,未发生则为 0,其本质上是概率预报与对应的观测发生与否的二值事件间的均方误差。 E_{BS} 取值范围为 $[0, 1]$,值越小,精度越高,完美预报 E_{BS} 为 0。

$$E_{\text{BS}} = \frac{1}{n} \sum_{i=1}^n [F_{f_i}(q) - F_{o_i}(q)]^2 \quad (5)$$

式中: n 为总样本对数; $F_{f_i}(q)$ 为集合预报 f_i 中集合成员超过确定性阈值 q 的概率; $F_{o_i}(q)$ 为观测值 o_i 是否超过确定性阈值 q 的二值事件。

$$F_{o_i}(q) = \begin{cases} 1, & o_i \geq q \\ 0, & o_i < q \end{cases} \quad (6)$$

Brier 技巧评分(E_{BSS})描述待评估预报相对于参考预报的预报精度技巧:当 $E_{\text{BSS}} > 0$ 时,表明待评估预报优于参考预报; E_{BSS} 越接近 1,相对于参考预报的预报精度技巧越高。

$$E_{\text{BSS}} = 1 - \frac{E_{\text{BS}}}{E_{\text{BS,ref}}} \quad (7)$$

式中: $E_{\text{BS,ref}}$ 为参考预报对应的 E_{BS} 值。

计算原始 GEFSv12 降水集合预报以及经过后处理后的降水集合预报相对于气候态参考预报的 E_{BSS} 值的过程中,由于原始 GEFSv12 降水集合预报集合成员为 5 个,气候态集合成员数为 25 个,存在成员数不对等的情况,需要利用公式获得原始集合预报 5 个集合成员的 E_{BS} 值转换为 25 个集合成员对应 E_{BS} 值的无偏估计,计算公式^[27]为

$$E_{\text{BS}_M} = E_{\text{BS}_m} - \frac{M-m}{M(m-1)n} \sum_{i=1}^n F_{f_i}(q)[1 - F_{f_i}(q)] \quad (8)$$

式中: n 为总样本对数; $F_{f_i}(q)$ 为原始集合预报 f_i 中集合成员超过确定性阈值 q 的概率; m 为原始集合预报集合成员数($m=5$); M 为气候态的集合成员数($M=25$)。

2.2.3 连续等级概率评分以及连续等级概率技巧评分

连续等级概率评分(E_{CRPS})是对概率预报总体精度的评价指标,描述预报和观测累积分布函数间的

方差在所有可能阈值上的积分,也可视为 E_{BS} 在所有可能阈值上的积分。对于多组预报和观测样本对,通常对 E_{CRPS} 计算平均值 \bar{E}_{CRPS} , \bar{E}_{CRPS} 取值范围为 $[0, +\infty]$, 完美预报为 0。

$$\bar{E}_{CRPS} = \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{+\infty} [F_{f_i}(y) - F_{o_i}(y)]^2 dy \quad (9)$$

式中: n 为总样本对数量; F_{f_i} 为集合预报 f_i 对应集合成员算得的累积概率分布函数; $F_{f_i}(y)$ 为集合预报 f_i 中成员 y 对应的累积概率值; F_{o_i} 为阶跃函数, 定义为

$$F_{o_i}(y) = \begin{cases} 0, & o_i \geq y \\ 1, & o_i < y \end{cases} \quad (10)$$

E_{CRPSS} 描述待评估预报相对于参考预报在预报精度方面的预报技巧情况: 当 $E_{CRPSS} > 0$ 时, 表示待评估预报优于参考预报; \bar{E}_{CRPSS} 越接近 1, 预报技巧越高。

$$E_{CRPSS} = 1 - \frac{\bar{E}_{CRPS}}{\bar{E}_{CRPS_{ref}}} \quad (11)$$

式中: $\bar{E}_{CRPS_{ref}}$ 为参考预报对应的 \bar{E}_{CRPS} 值。

计算原始 GEFsV12 降水集合预报相对于气候态参考预报的 E_{CRPSS} 值时, 同样需要利用公式获得原始集合预报 5 个集合成员的 \bar{E}_{CRPS} 值转换为 25 个集合成员对应的 \bar{E}_{CRPS} 值的无偏估计, 计算公式^[27] 为

$$\bar{E}_{CRPS_m} = \bar{E}_{CRPS_m} - \frac{M-m}{2Mmn} \sum_{i=1}^n A_i \quad (12)$$

$$A_i = \frac{1}{m(m-1)} \sum_{j \neq k} |f_{i,j} - f_{i,k}| \quad (13)$$

式中: n 为总样本对数; A_i 为原始集合预报 f_i 对应的各集合成员 $\{f_{i,1}, \dots, f_{i,m}\}$ 算得的基尼平均差; m 为原始集合预报集合成员数 ($m=5$); M 为气候态的集合成员数 ($M=25$)。

2.2.4 α 指数

E_{PIT} 是集合预报 f_i 对应的集合成员算得的累积概率分布函数 F_{f_i} 在观测值 o_i 处的取值, 定义为

$$E_{PIT_i} = F_{f_i}(o_i) \quad (14)$$

在对降水计算 E_{PIT} 值时, 由于降水量存在大量零值, 对观测值小于等于降水阈值 o_c 的情况下(此处定为 0.1 mm/d), 采用伪 E_{PIT} (pseudo- E_{PIT}) 值, 生成 $[0, F_{f_i}(o_c)]$ 区间内均匀分布的随机数作为 E_{PIT} 值^[28]。

如果预报具有完美的可靠性, E_{PIT} 值是服从均匀分布的。

集合预报的可靠性可用 α 指数来表征, α 指数定义为

$$\alpha = 1.0 - \frac{2}{n} \sum_{i=1}^n \left| E_{PIT_i} - \frac{i}{n+1} \right| \quad (15)$$

式中: n 为总样本对数量; E_{PIT} 从小到大排序的 E_{PIT} 值中排在第 i 位的 E_{PIT} 值; α 指数取值范围为 $[0, 1]$, 值越高, 表示 E_{PIT} 值越接近均匀分布, 即可靠性越好。

2.2.5 可靠性框图

可靠性框图是针对给定阈值条件下降水发生与否事件, 表征预报概率与观测频率接近程度的指标。根据不同阈值大小, 计算每个集合预报样本在对应阈值下的预报概率值; 再根据算得的预报概率值以及设置的概率区间, 获得对应概率区间内的预报概率以及观测值频率。由于原始集合预报集合成员有 5 个, 成员数较少, 设置概率区间为 $[0, 0.2)$ 、 $[0.2, 0.4)$ 、 \dots 、 $[0.8, 1.0]$, 后处理后的预报集合成员为 25 个, 成员数相对较多, 设置概率区间为 $[0, 0.1)$ 、 $[0.1, 0.2)$ 、 \dots 、 $[0.9, 1.0]$ 。

可靠性框图的 x 轴为落在各个概率区间内的预报概率的平均值, y 轴为对应预报概率区间内的观测值频率, 计算公式为

$$f_i = \frac{1}{N_k} \sum_{i \in I_k} F_{y_i}(q) \quad (16)$$

$$P_{f_i}(o = 1 | f_i) = \frac{1}{N_k} \sum_{i \in I_k} p(o_i \geq q) \quad (17)$$

式中: q 为确定性阈值; $F_{y_i}(q)$ 为集合预报 y_i 中集合成员超过确定性阈值 q 的概率; N_k 为落于第 k 个概率区间内的样本数; I_k 为落于第 k 个概率区间内的样本索引。

为使组成可靠性框图的散点具有统计意义, 绘制可靠性框图时须舍去落在概率区间内预报概率样本数小于 10 个的散点。

3 评价结果分析

3.1 降水空间分布特征

预见期为第 1 天时, 淮河流域观测降水、原始 GEFsV12 降水集合预报及后处理后的降水集合预报多年日平均降水量空间分布见图 2。以观测降水为基准, 原始 GEFsV12 降水预报在全流域的格点上存在明显高估的现象, 而后处理后的降水预报除少

量格点存在低估现象外, 预报降水量与观测降水基本一致, 降水集合预报的偏差情况有所改善。后处理后的降水预报在空间分布上与观测保持一致, 后处理后的预报相较于原始 GEFSv12 预报保持了合理的时空相关性。

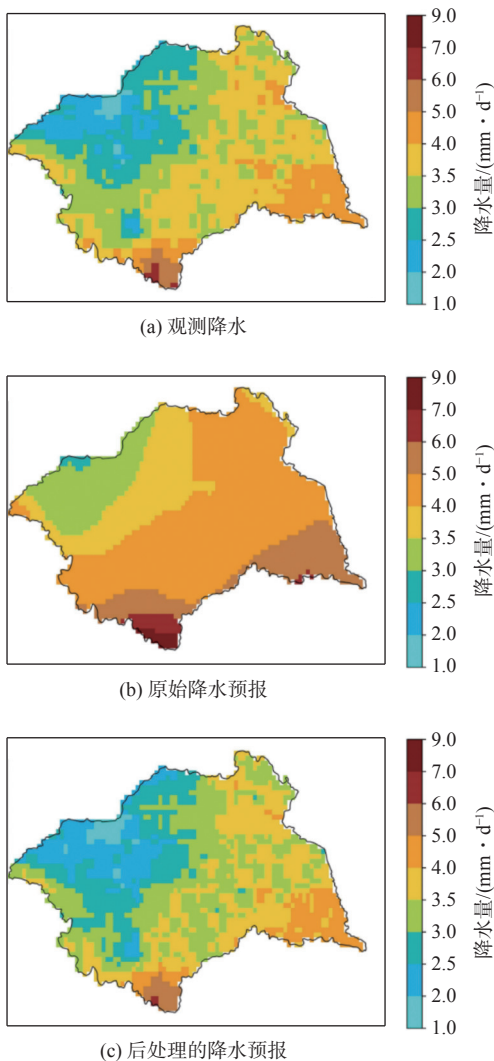
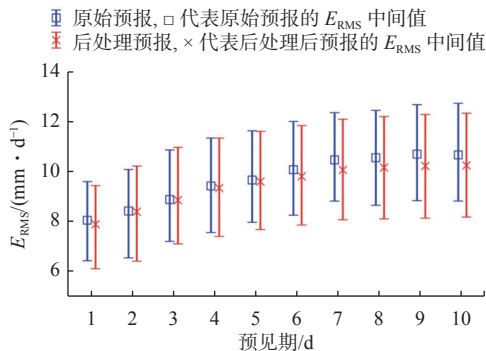


图2 预见期为第1天时, 2008—2017年5—9月淮河流域日平均降水量空间分布

Fig. 2 Spatial variation of precipitation in the Huaihe River basin at 1 day lead time. The precipitation is averaged over the periods 1 May to 30 September from 2008 to 2017

3.2 均方根误差

淮河流域原始 GEFSv12 降水集合预报以及后处理后的降水集合预报, 均方根误差随预见期的变化情况见图3。原始预报均方根误差的中值随预见期的增加而增大, 取值在 8.10~10.73 mm/d。后处理后的预报随预见期的变化趋势与原始预报相一致, 相较于原始预报, 其均方根误差中值有所降低, 降幅在 0.3%~4.4%, 偏差情况有所改善。



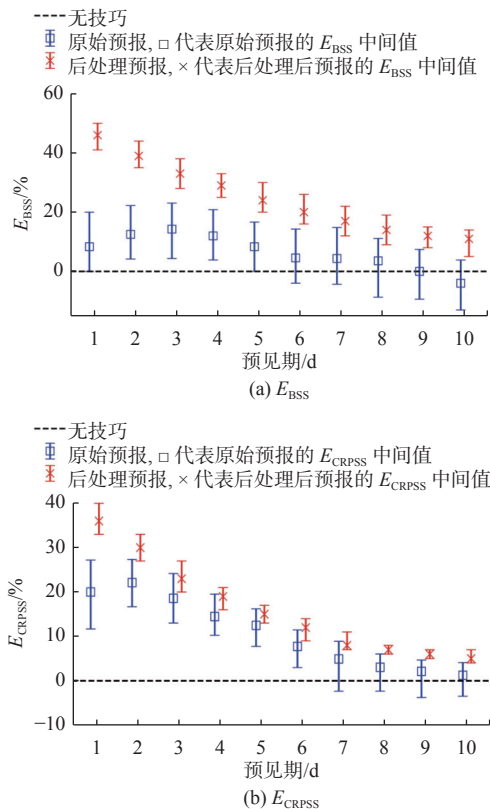
注: 上下两根短横线中间距离表示 E_{RMS} 的 [10%, 90%] 分数区间。

图3 淮河流域原始 GEFSv12 预报以及后处理后的预报 E_{RMS} 随预见期的变化区间

Fig. 3 The E_{RMS} of raw GEFSv12 forecasts (blue) and BJP post-processed forecasts (red) at different lead days in the Huaihe River basin

3.3 预报精度技巧

淮河流域原始 GEFSv12 降水集合预报以及后处理后的降水集合预报的 E_{BSS} 、 E_{CRPSS} 随预见期的变化情况见图4。 E_{BSS} 是概率预报评价指标, 能够评价在不同概率阈值条件下, 如图4(a)阈值为 0.1 mm/d (即发生小雨以上量级的降水), 原始预报以及 BJP 后处理后的预报相对于气候态的精度评价技巧。



注: 上下两根短横线中间距离表示 E_{BSS} 、 E_{CRPSS} 的 [10%, 90%] 分数区间。

图4 淮河流域原始 GEFSv12 预报以及后处理后预报的 E_{BSS} 、 E_{CRPSS} 随预见期的变化区间

Fig. 4 The E_{BSS} and E_{CRPSS} of raw GEFSv12 forecasts (blue) and BJP post-processed forecasts (red) for different lead days in the Huaihe River basin

当预见期为第 1 天时,原始预报 E_{BSS} 取值范围为 0~20.0%,中值为 8.3%;后处理后的预报在 41.0%~50.0%,中值为 46.0%;使用统计后处理方法后,后处理后的预报中值从 8.3% 提升至 46.0%,预报精度技巧明显得到改善。在其余预见期,原始预报 E_{BSS} 中值取值范围为-4.0%~14.3%,后处理后的预报中值取值范围为 11.0%~39.0%,预报精度技巧均有所改善。当预见期 ≥ 6 d 时,原始预报部分格点的 E_{BSS} 在 0 值以下,对该部分格点而言,气候态的预报精度要优于原始预报。随着预见期增加,后处理后的预报 E_{BSS} 总体趋势为逐渐减小,但始终大于 0,即预报技巧有所下降,但后处理后的预报始终优于气候态,对比原始预报,可以发现统计后处理方法可以改善预报在较长预见期(预见期 ≥ 6 d)上的预报技巧。

与 E_{BSS} 不同, E_{CRPSS} 描述的是所有阈值下的概率预报结果的总体精度评价技巧。后处理后的预报 E_{CRPSS} 中值取值范围为 5.0%~36.0%,原始预报 E_{CRPSS} 中值在 1.2%~22.1%。与 E_{BSS} 的计算结果相类似,在各个预见期,后处理后的预报的 E_{CRPSS} 始终大于原始预报的 E_{CRPSS} ,即统计后处理方法改善了原始预报的精度技巧。随着预见期增加,原始预报以及后处理后的预报的 E_{CRPSS} 总体趋势为逐渐下降。当预见期 ≥ 7 d 时,原始预报部分格点 E_{CRPSS} 小于 0,对于这些格点而言,气候态参考预报要优于原始预报。后处理后的预报 E_{CRPSS} 始终大于 0,说明统计后处理方法可以改善预报在较长预见期(预见期 ≥ 7 d)上的预报技巧。

3.4 预报可靠性

3.4.1 α 指数

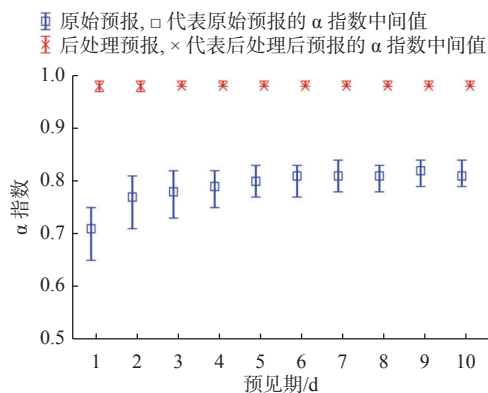
α 指数是表征 E_{PIT} 分布均匀程度的指标,反映了预报的整体可靠性。原始 GEFSv12 降水集合预报以及后处理后的降水集合预报的 α 指数随预见期的变化情况见图 5。原始预报的 α 指数中值取值范围为 0.71~0.82,后处理后的预报中值取值始终在 0.98 左右,后处理后的预报相比于原始预报提升了约 20%~38%,取值更接近于 1,可靠性更好,即统计后处理方法提高了预报的可靠性。

3.4.2 可靠性框图

可靠性框图是表现超过某一阈值的降水预报概率与实际发生的观测频率的接近程度。在两种阈值 0.1 mm/d、10 mm/d 下,即发生小雨以上以及中雨以上量级降水事件,概率预报的可靠性情况见图 6。

在理想情况下,预报概率与观测频率相一致,即散点落在 1:1 线附近。由图 6 可以看出:预见期为

第 1 天、第 3 天时,原始预报预报概率均明显大于实际发生的观测频率,存在明显高估现象;而后处理后的预报较原始预报接近理想情况,存在略微低估现象,即统计后处理方法使得预报概率与观测频率大体一致,有效提高了预报的可靠性。通过对比两种降水阈值,可以发现阈值为 10 mm/d 时,后处理后的预报更加接近理想情况,说明后处理后的预报对中雨以上降水事件的预报可靠性要高于小雨以上降水事件的可靠性。



注:上下两根短横线中间距离表示 α 指数的 [10%, 90%] 分数区间。

图 5 淮河流域原始 GEFSv12 预报以及后处理后的预报 α 指数随预见期的变化区间

Fig. 5 The α index of raw GEFSv12 forecasts (blue) and BJP post-processed forecasts (red) for different lead days in the Huaihe River basin

4 结论

降水是水文模型的输入变量,获取准确的降水预报至关重要。对原始 GEFSv12 降水集合预报采用 BJP 进行统计后处理,并对原始 GEFSv12 降水预报以及经过后处理后的降水预报进行了系统评估。主要结论如下:

在预报偏差方面,根据降水的空间分布以及 E_{RMS} 结果可知,原始 GEFSv12 预报数据偏差较大,难以重现观测值的空间分布,而统计后处理方法能够有效减少预报的系统偏差。

在预报精度方面,原始 GEFSv12 预报在预见期 1~7 d 具有优于气候态的预报技能;随着预见期的增加,预报精度技巧的总体趋势为下降趋势。根据 E_{BSS} 可知,当预见期 ≥ 6 d 时,GEFSv12 对淮河流域部分格点的小雨量级以上降水事件的预报精度要低于气候态参考预报,而使用统计后处理方法能够改善原始预报在长预见期(≥ 6 d)上的预报技能。根据 E_{CRPSS} 可知,当预见期 ≥ 7 d 时,GEFSv12 对淮河流域部分格点的预报精度要低于气候态参考预报,而使用统计后处理方法能够改善在长预见期(≥ 7 d)上的预报技能。

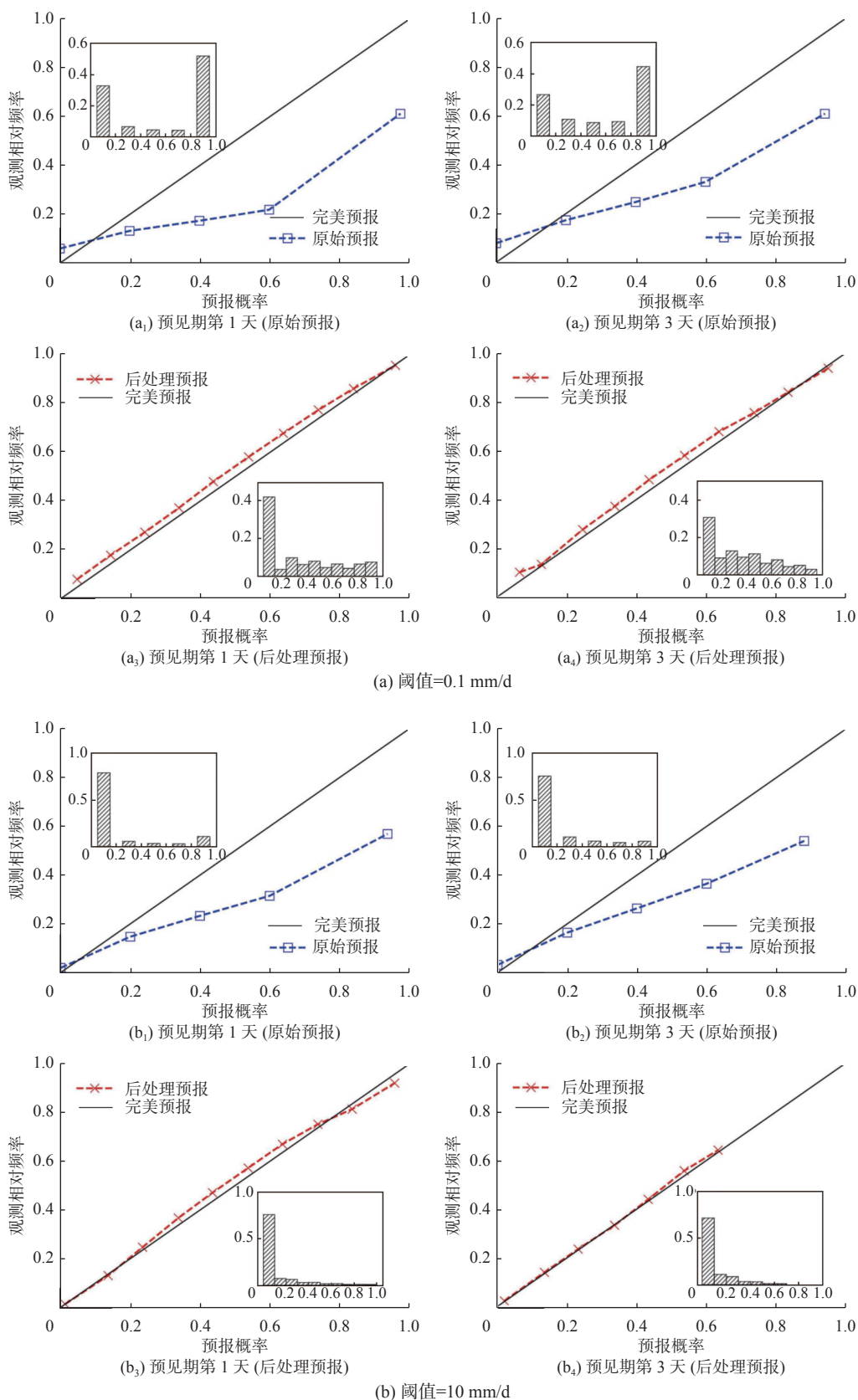


图 6 不同阈值条件下, 预见期分别为第 1、3 天, 原始 GEFSv12 预报以及后处理后的预报对应的可靠性框图

Fig. 6 The reliability diagrams of raw GEFSv12 forecasts (blue) and BJP post-processed forecasts (red) for different thresholds at lead days of 1 and 3 in the Huaihe River basin

在预报可靠性方面, 根据 α 指数以及可靠性框图可知, 对小雨和中雨以上量级降水事件的预报,

GEFSv12 存在明显高估, 使用统计后处理后预报可靠性得到明显提高。

对比后处理前后的 GEFSv12产品以证实降水预报产品在淮河流域的适用性,后处理后的降水集合预报可进一步应用于水文集合预报等方面。在后续的研究中,将对 GEFSv12 与前一代产品 GEFSv10 进行比较,以及两套产品在我国其他流域的适用性进行进一步分析。

参考文献(References):

- [1] WILKS D S. Statistical methods in the atmospheric sciences[M]. 4th ed. Amsterdam: Elsevier, 2020. DOI: [10.1016/C2017-0-03921-6](https://doi.org/10.1016/C2017-0-03921-6).
- [2] 彭勇, 王萍, 徐炜, 等. 气象集合预报的研究进展[J]. *南水北调与水利科技*, 2012, 10(4): 90-96. (PENG Y, WANG P, XU W, et al. Review of weather ensemble prediction[J]. *South-to-North Water Transfers and Water Science & Technology*, 2012, 10(4): 90-96. (in Chinese)) DOI: [10.3724/SP.J.1201.2012.04090](https://doi.org/10.3724/SP.J.1201.2012.04090).
- [3] 高冰, 杨大文, 谷湘潜, 等. 基于数值天气模式和分布式水文模型的三峡入库洪水预报研究[J]. *水力发电学报*, 2012, 31(1): 20-26. (GAO B, YANG D W, GU X Q, et al. Flood forecast of Three Gorges reservoir based on numerical weather forecast model and distributed hydrologic model[J]. *Journal of Hydroelectric Engineering*, 2012, 31(1): 20-26. (in Chinese))
- [4] 郭生练, 张俊, 郭靖, 等. 基于气象模式的汉江流域洪水预报系统[J]. *水利水电科技进展*, 2009, 29(3): 1-5. (GUO S L, ZHANG J, GUO J, et al. Flood forecasting system of Hanjiang basin based on meteorological model[J]. *Advances in Sciences and Technology of Water Resources*, 2009, 29(3): 1-5. (in Chinese)) DOI: [10.3880/j.issn.1006-7647.2009.03.001](https://doi.org/10.3880/j.issn.1006-7647.2009.03.001).
- [5] DUAN Q Y, PAPPENBERGER F, WOOD A, et al. Handbook of hydrometeorological ensemble forecasting[M]. Berlin: Springer, 2019. DOI: [10.1007/978-3-642-39925-1](https://doi.org/10.1007/978-3-642-39925-1)
- [6] 叶爱中, 段青云, 徐静, 等. 基于GFS的飞来峡流域水文集合预报[J]. *气象科技进展*, 2015, 5(3): 57-61. (YE A Z, DUAN Q Y, XU J, et al. The hydrological ensemble prediction based on GFS in Feilaixia basin[J]. *Advances in Meteorological Science and Technology*, 2015, 5(3): 57-61. (in Chinese)) DOI: [10.3969/j.issn.2095-1973.2015.03.005](https://doi.org/10.3969/j.issn.2095-1973.2015.03.005).
- [7] 包红军, 赵琳娜. 基于集合预报的淮河流域洪水预报研究[J]. *水利学报*, 2012, 43(2): 216-224. (BAO H J, ZHAO L N. Flood forecast of Huaihe River based on TIGGE ensemble predictions[J]. *Journal of Hydraulic Engineering*, 2012, 43(2): 216-224. (in Chinese)) DOI: [10.13243/j.cnki.slxb.2012.02.001](https://doi.org/10.13243/j.cnki.slxb.2012.02.001).
- [8] PAPPENBERGER F, BARTHOLMES J, THIELEN J, et al. New dimensions in early flood warning across the globe using grand-ensemble weather predictions[J]. *Geophysical Research Letters*, 2008, 35(3): 1956-1964. DOI: [10.1029/2008GL033837](https://doi.org/10.1029/2008GL033837).
- [9] 李俊, 廖移山, 张兵, 等. 集合数值预报方法在山洪预报中的初步应用[J]. *高原气象*, 2007, 26(4): 854-861. (LI J, LIAO Y S, ZHANG B, et al. The preliminary application of ensemble prediction in flash flood forecasting[J]. *Plateau Meteorology*, 2007, 26(4): 854-861. (in Chinese))
- [10] 刘硕, 王国利, 张琳. TIGGE降雨信息在柴河流域洪水预报中可利用性评估[J]. *水文*, 2018, 38(5): 17-22. (LIU S, WANG G L, ZHANG L. Availability assessment of TIGGE precipitation data in flood forecasting for Chaihe River basin[J]. *Journal of China Hydrology*, 2018, 38(5): 17-22. (in Chinese)) DOI: [10.3969/j.issn.1000-0852.2018.05.004](https://doi.org/10.3969/j.issn.1000-0852.2018.05.004).
- [11] 巴欢欢, 郭生练, 钟逸轩, 等. 考虑降水预报的三峡入库洪水集合概率预报方法比较[J]. *水科学进展*, 2019, 30(2): 182-197. (BA H H, GUO S L, ZHONG Y X, et al. Comparative study on probabilistic ensemble flood forecasting considering precipitation forecasts for the Three Gorges Reservoir[J]. *Advances in Water Science*, 2019, 30(2): 182-197. (in Chinese)) DOI: [10.14042/j.cnki.32.1309.2019.02.004](https://doi.org/10.14042/j.cnki.32.1309.2019.02.004).
- [12] 刘莉. 基于TIGGE数据的分布式集合洪水预报研究[D]. 杭州: 浙江大学, 2020. (LIU L. Research on distributed ensemble flood forecasting based on TIGGE dataset[D]. Hangzhou: Zhejiang University, 2020. (in Chinese)) DOI: [10.27461/d.cnki.gzjdx.2020.002486](https://doi.org/10.27461/d.cnki.gzjdx.2020.002486).
- [13] 雍燕兰, 杨云川, 赵铜铁. 水文集合预报检验指标调研及其在西江流域季节降水预报中的应用[J]. *水文*, 2020, 40(6): 31-39. (YONG Y L, YANG Y C, ZHAO T T G. Verification metrics of hydrological ensemble forecasts and their applications to seasonal precipitation forecasts of Xijiang River basin[J]. *Journal of China Hydrology*, 2020, 40(6): 31-39. (in Chinese)) DOI: [10.19797/j.cnki.1000-0852.20190348](https://doi.org/10.19797/j.cnki.1000-0852.20190348).
- [14] MURPHY A H. What is a good forecast? An essay on the nature of goodness in weather forecasting[J]. *Weather and Forecasting*, 1993, 8(2): 281-293. DOI: [10.1175/1520-0434\(1993\)008<0281:WIAGFA>2.0.CO;2](https://doi.org/10.1175/1520-0434(1993)008<0281:WIAGFA>2.0.CO;2).
- [15] HAMILL T M, JURAS J. Measuring forecast skill: Is it real skill or is it the varying climatology?[J]. *Quarterly Journal of the Royal Meteorological Society*, 2006, 132(621C): 2905-2923. DOI: [10.1256/qj.06.25](https://doi.org/10.1256/qj.06.25).

- [16] HERBACH H. Decomposition of the continuous ranked probability score for ensemble prediction systems[J]. *Weather and Forecasting*, 2000, 15(5): 559-570. DOI: 10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2.
- [17] RENARD B, KAVETSKI D, KUCZERA G, et al. Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors[J]. *Water Resources Research*, 2010, 46: W05521. DOI: 10.1029/2009WR008328.
- [18] HAMILL T M, WHITAKER J S, MULLEN S L. Re-forecasts: An important dataset for improving weather predictions[J]. *Bulletin of the American Meteorological Society*, 2006, 87(1): 33-46. DOI: 10.1175/BAMS-87-1-33.
- [19] GUAN H, ZHU Y J, SINSKY E, et al. GEFSv12 re-forecast dataset for supporting subseasonal and hydrometeorological applications[J]. *Monthly Weather Review*, 2022, 150(3): 647-665. DOI: 10.1175/MWR-D-21-0245.1.
- [20] YANG C L, YUAN H L, SU X. Bias correction of ensemble precipitation forecasts in the improvement of summer streamflow prediction skill[J]. *Journal of Hydrology*, 2020, 588: 124955. DOI: 10.1016/j.jhydrol.2020.124955.
- [21] WANG Q J, ROBERTSON D E, CHIEW F H S. A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites[J]. *Water Resources Research*, 2009, 45: W05407. DOI: 10.1029/2008WR007355.
- [22] SHRESTHA D L, ROBERTSON D E, BENNETT J C, et al. Improving precipitation forecasts by generating ensembles through postprocessing[J]. *Monthly Weather Review*, 2015, 143(9): 3642-3663. DOI: 10.1175/MWR-D-14-00329.1.
- [23] 水利部淮河水利委员会.流域介绍[EB/OL]. <http://www.hrc.gov.cn/main/lyjs.jhtml>. (The Huaihe River Commission of the Ministry of Water Resources, P.R.C.Introduction to the Huaihe River basin[EB/OL].<http://www.hrc.gov.cn/main/lyjs.jhtml>.(in Chinese))
- [24] SHEN Y, ZHAO P, PAN Y, et al. A high spatiotemporal gauge-satellite merged precipitation analysis over China[J]. *Journal of Geophysical Research:Atmospheres*, 2014, 119(6): 3063-3075. DOI: 10.1002/2013JD020686.
- [25] 于群, 吴炜, 周发琇, 等. 中国东部降水的气候模态及雨季划分[J]. *大气科学学报*, 2014, 37(3): 378-384. (YU Q, WU W, ZHOU F X, et al. Climate modes of precipitation in eastern China and division of rainy season[J]. *Transactions of Atmospheric Sciences*, 2014, 37(3): 378-384. (in Chinese)) DOI: 10.13878/j.cnki.dqkxxb.2014.03.012.
- [26] CLARK M, GANGOPADHYAY S, HAY L, et al. The Schaake Shuffle: A method for reconstructing space-time variability in forecasted precipitation and temperature fields[J]. *Journal of Hydrometeorology*, 2004, 5(1): 243-262. DOI: 10.1175/1525-7541(2004)005<0243:TSSAMF>2.0.CO;2.
- [27] FERRO C A T, RICHARDSON D S, WEIGEL A P. On the effect of ensemble size on the discrete and continuous ranked probability scores[J]. *Meteorological Applications*, 2008, 15(1): 19-24. DOI: 10.1002/met.45.
- [28] WANG Q J, ROBERTSON D E. Multisite probabilistic forecasting of seasonal flows for streams with zero value occurrences[J]. *Water Resources Research*, 2011, 47: W02546. DOI: 10.1029/2010WR009333.

Evaluation of the performance of GEFSv12 precipitation reforecast dataset in the Huaihe River basin

HUANG Sainan^{1,2}, LI Wentao^{1,2}, DUAN Qingyun^{1,2}

(1. State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering and College of Hydrology & Water Resources, Hohai University, Nanjing 210098, China; 2. CMA-HHU Joint Laboratory for Hydrometeorological Studies, Nanjing 210098, China)

Abstract: Due to the complexity of atmospheric physical processes, numerical weather prediction often suffers from uncertainty. Ensemble forecasts can be used to represent the uncertainty and provide uncertainty information for risk-based decision-making. In recent years, hydrometeorological ensemble forecasting has become widely applied, which uses the ensemble outputs from the numerical weather prediction models to drive the hydrological models. Bias-free, accurate and reliable ensemble precipitation forecasts are important for producing accurate streamflow forecasts. Forecast verification is the process of assessing the quality of forecasts, which involves the investigation

of the properties of the joint distribution of forecasts and observations. Objective evaluations of forecast quality can help monitor operational forecasts, assess the specific strengths and weaknesses of forecast systems and support decision-making.

Global Ensemble Forecast System, version 12 (GEFSv12), is a new generation of reforecast dataset generated by the National Centers for Environmental Prediction (NCEP). Reforecast dataset is generated by retrospectively re-forecasting weather for previous years using the current dynamical model. It is an important dataset for improving weather predictions. The performance of GEFSv12 precipitation forecasts has not been discussed in the Huaihe River basin, so we evaluated the performance of GEFSv12 precipitation forecasts in the Huaihe River basin in this research. Due to the systematic bias of the raw forecasts, the Bayesian Joint Probability (BJP) model was used to perform statistical post-processing on the raw GEFSv12 reforecasts. The quality of the raw and post-processed GEFSv12 ensemble reforecasts was evaluated in three aspects, i. e. , bias, accuracy and reliability. Verification metrics included Root Mean Squared Error (E_{RMS}), Brier Skill Score (E_{BSS}), Continuous Ranked Probability Skill Score (E_{CRPSS}), α index and Reliability Diagram.

The main results are as follows:(1) In terms of forecast bias, according to the map of the spatial variation of precipitation and the results of E_{RMS} , raw GEFSv12 precipitation forecasts suffer from a large bias, and have difficulty in reproducing the spatial variability of precipitation within the Huaihe River basin. Statistical post-processing approach can effectively reduce the systematic bias of the raw forecasts. (2) In terms of forecast accuracy, the accuracy of raw GEFSv12 precipitation forecasts is better than that of climatology at lead times of 1-7 days. The accuracy of the raw and post-processed forecasts decreases with lead times. According to the E_{BSS} , when the lead times are up to 6 days, the accuracy of raw GEFSv12 precipitation forecasts for light rain events at some grid cells in the Huaihe River basin is lower than that of climatology. The statistical post-processing approach can improve the forecast accuracy of raw forecasts for light rain at long lead times(≥ 6 days). According to the E_{CRPSS} , when the lead times are up to 7 days, the accuracy of raw GEFSv12 precipitation forecasts at some grid cells in the Huaihe River basin is lower than that of climatology. The statistical post-processing approach can improve the forecast accuracy of raw forecasts at long lead times(≥ 7 days). (3) In terms of forecast reliability, according to the α index and reliability diagram, the raw GEFSv12 precipitation forecasts suffer from overestimation for light and moderate rain events. Statistical post-processing approach can significantly improve forecast reliability.

The results show that the raw GEFSv12 forecasts perform well in the Huaihe River basin at lead times of 1-7 days. BJP post-processing approach can effectively reduce the systematic bias of the raw forecasts and increase forecast accuracy and reliability. Post-processed precipitation ensemble forecasts can be applied in further applications such as hydrological ensemble forecasting.

Key words: Global Ensemble Forecast System GEFSv12; reforecast; BJP post-processing approach; ensemble forecasts verification; Huaihe River basin