

洪水概率预报评价指标研究*

蒋晓蕾, 梁忠民**, 胡义明, 王 军, 李彬权

(河海大学水文水资源学院, 南京 210098)

摘 要: 在分析现有洪水概率预报评价指标的基础上, 建立了洪水概率预报的“精度-可靠度”联合评价指标体系。“精度”指标用于评价倾向值预报的准确性, 包括确定性系数、相对误差等指标; “可靠度”指标用于评价区间预报的合理性, 包括覆盖率、区间离散度等指标。其中, 在“可靠度”评价指标中, 提出了一个新的评价指标, 即覆盖率判定系数, 用以评价多个区间预报结果的整体合理性。以淮河王家坝区间流域为例进行应用研究, 结果表明: “精度-可靠度”联合评价指标体系可以对洪水概率预报结果的合理性做出更全面的评价。

关键词: 洪水概率预报; 倾向值预报; 区间预报; 覆盖率判定系数; 精度-可靠度联合评价指标体系; 淮河流域

Research on assessment criteria in probabilistic flood forecasting*

JIANG Xiaolei, LIANG Zhongmin**, HU Yiming, WANG Jun & LI Binquan

(College of Hydrology and Water Resources, Hohai University, Nanjing 210098, P.R.China)

Abstract: Based on the analysis of the evaluating indicators used in the study of probabilistic flood forecasting, a system of assessment criteria in probabilistic flood forecasting is established named as the “accuracy-reliability” joint assessment criteria system. In this system, the “accuracy” estimation refers to assess the precision of preferred prediction using the indicators like Nash-Sutcliffe efficiency, relative errors and so on. While the “reliability” assessment means to evaluate the reasonability of predicted interval using the indicators like containing ratio, deviation index and so on. Meanwhile, a new “reliability” assessment indicator named the containing ratio coefficient is proposed to assess the comprehensive reasonability of multiple predicted intervals. The accuracy-reliability joint assessment criteria system is tested in Wangjiaba subbasin in Huaihe River. The results suggested that this system gives a synthetic evaluation of the probabilistic flood forecasting.

Keywords: Probabilistic flood forecasting; preferred value; interval prediction; containing ratio coefficient; accuracy-reliability joint assessment criteria system; Huaihe Basin

由于自然过程的复杂性和人类认识水平的局限性, 使得单一预报结果不可避免的存在不确定性, 因此发展起来了集合预报与概率预报^[1]。集合预报给出了预报量可能的结果集, 是采用集合形式量化预报的不确定性。而概率预报, 顾名思义, 估计了预报量(如流量)的概率分布, 通过概率形式量化预报的不确定性。一方面, 集合预报仅提供了预报量有限的可能结果, 而概率预报量化了预报量所有结果的可能性, 因此, 概率预报是集合预报的一种“升华”, 较集合预报提供更为全面的信息。另一方面, 由于很难直接估计预报量的理论分布函数, 因此, 往往采用预报量有限的可能结果集来近似替代其分布函数, 估计其预报不确定性, 即集合预报可以视为概率预报的一种表现形式^[2]。本文将集合预报视作概率预报的一种表现形式, 对洪水概率预报结果的评价体系开展研究。

就方法而言, 目前的洪水概率预报方法可以大致分为两类途径: 全要素耦合途径和总误差分析途径。

全要素耦合途径是指分别识别、量化洪水预报过程中各阶段的不确定性, 并将其与确定性预报模型耦合, 从而实现洪水概率预报。这些不确定性一般包括输入(如降雨)^[3-4]、模型结构^[5]、模型参数和模型状态^[6]

* 2019-05-16 收稿; 2019-08-09 收修改稿。

国家重点基础研发计划项目(2016YFC0402709)和国家自然科学基金项目(41730750)联合资助。

** 通信作者; E-mail: zmliang@hhu.edu.cn.

的不确定性. 代表性方法包括贝叶斯总误差分析方法 BATEA^[7] (BAYesian Total Error Analysis)、贝叶斯综合不确定性估计方法 IBUNE^[8] (Integrated Bayesian UNCertainty Estimator) 等. 这类方法虽然可以量化洪水预报过程中的各类不确定性要素, 识别不确定性来源, 但往往计算耗时, 不利于生产实际应用. 总误差分析途径是指不识别洪水预报具体过程的不确定性, 仅对最终预报结果进行分析, 量化其不确定性, 进而实现概率预报^[9-10]. 常用的方法包括水文不确定性处理器 HUP^[11] (Hydrologic Uncertainty Processor)、模型条件处理器 MCP^[12] (Model Conditional Processor)、三维误差矩阵^[13] 等. 这一途径避免了处理洪水预报子过程的不确定性, 往往计算省时, 结构灵活, 可以与各类确定性预报模型进行衔接, 但是无法溯源预报的不确定性.

随着洪水概率预报方法研究的进展, 其评价问题也得到重视, 但相较而言, 对洪水概率预报评价的研究相对滞后, 尚未形成较完整的评价指标体系. Murphy^[14] 构建了概率预报制定-应用过程的检验框架: 从一致性 (Consistency)、预报质量 (Quality) 和预报价值 (Value) 三方面进行检验. 其中“预报质量”指预报结果与观测数据的一致性, 即一般意义上预报结果的好坏. Murphy 就观测分布 $f(Q_{\text{obs}})$ 、预报分布 $f(Q_{\text{sim}})$ 、条件分布 $f(Q_{\text{obs}}|Q_{\text{sim}})$ 和 $f(Q_{\text{sim}}|Q_{\text{obs}})$ 的关系总结了偏差 (Bias, 预报分布与观测分布均值的差异)、技巧 (Skill, 相对的预报精度)、分辨率 (Resolution, 预报条件分布与观测分布的差异) 等一系列评价预报质量的指标, 为后续的概率预报评价研究提供了理论基础. 而在实际洪水概率预报研究中, 往往关注条件分布 $f(\hat{Q}|Q_{\text{sim}})$ (以确定性预报为条件的实际流量的分布函数) 对实际观测值 Q_{obs} 的“拟合情况”, 如任政等^[15] 采用区间覆盖率 (预报区间覆盖实测流量的比例)、区间平均宽度和区间对称性对概率预报区间进行评估; 徐兴亚等^[6] 采用 60% 和 90% 置信度预报区间的落入频率 (即覆盖率) 对河道洪水概率预报的区间预报进行评价; 王艳兰等^[16] 采用 90% 置信度的覆盖率和平均离散度评估概率预报结果, 并给出了 Q50 分位数预报的洪峰误差、洪量误差和纳什系数; 徐伟等^[17] 采用单位区间离散化系数对比评估了两个概率预报模型的 90% 置信度预报区间的合理性; 梁忠民等^[18] 采用 90% 置信度覆盖率、平均相对带宽 (即平均离散度) 和平均相对偏移度对洪水概率预报结果进行评价, 并计算了中位数预报的洪峰误差和纳什系数. 此外, 随着水文气象研究的深入, 气象集合预报的评价方式在水文领域得到了发展, 如 Alfieri 等^[19] 采用连续排序概率得分 CRPS 和连续排序概率技能得分 CRPSS 对集合流量预报结果进行了评价.

尽管洪水概率预报评价指标研究得到了一定程度的发展, 但仍存在一些问题: 如不同研究中采用的评价指标, 有些虽然名称不同, 但含义相同, 尚欠缺相关研究对现有的评价指标进行归纳总结; 这些指标大多仅对某一置信度条件下的区间预报结果进行合理性评估, 如 90% 置信度区间预报的覆盖率、离散度等, 尚没有一种评价指标可以对概率预报的整体合理性做出估计 (即同时考虑所有置信度区间预报的合理性); 不同研究对概率预报结果的侧重各有不同, 采用的评价指标亦有所差异, 尚没有形成一套通用的系统的概率预报评价指标体系. 为此, 本文根据实际洪水概率预报研究和应用的需要, 针对条件分布 $f(\hat{Q}|Q_{\text{sim}})$ 对实际观测值 Q_{obs} “拟合情况”的评价问题, 分析并总结了现有研究中常见的评价指标, 形成了洪水概率预报“精度-可靠度”联合评价指标体系, 并在可靠度评价指标中, 提出一个新的评价指标-覆盖率判定系数, 用以评估概率预报的整体合理性 (同时考虑各个置信度的区间预报). 在淮河王家坝区间流域, 对“精度-可靠度”联合评价指标体系进行应用研究.

1 概率预报“精度-可靠度”联合评价指标体系

传统的洪水确定性预报提供了每一时刻预报量的预测值, 可以通过比较预测值与观测值的差异来评价确定性预报结果的准确性, 即精度评价. 以流量预报为例, t 时刻真实流量 \hat{Q} 的观测值为 Q_{obs} , 传统的确定性预报可以给出流量的单一模拟/预报值 Q_{sim} , 可以通过比较 Q_{obs} 和 Q_{sim} 的数值来评价确定性预报效果. 而概率预报估计的是每一时刻预报量的条件分布, 即以模拟/预报值 Q_{sim} 为条件的流量 \hat{Q} 的条件分布 $f(\hat{Q}|Q_{\text{sim}})$. 通过 $f(\hat{Q}|Q_{\text{sim}})$ 不仅可以获得预报量的倾向值预报 (某一分位数, 如中位数 Q50 或均值等), 亦可获得具有一定置信度的区间预报等结果. 因此, 对概率预报进行评价时, 应对倾向值预报的准确性 (精度评价) 和区间预报的合理性 (可靠度评价) 分别进行评估.

本节首先对现有的精度评价指标和可靠度评价指标进行分析和总结, 在可靠度评价指标中, 提出覆盖率判定系数用以评价概率预报的整体可靠度. 在此基础上, 建立了洪水概率预报“精度-可靠度”联合评价指

标体系.

1.1 精度评价指标

对任一预报时刻,可以通过条件分布 $f(\hat{Q}|Q_{\text{sim}})$ 来估计预报量的某一分位数(中位数 Q_{50} 或均值等),获得倾向值预报. 所以,倾向值预报是一种定值预报,可以采用现有确定性预报的评价指标进行精度评价. 以流量预报为例,常用的评价指标包括:洪峰相对误差、洪量相对误差、洪峰滞时、纳什系数等:

(1) 洪峰相对误差 $REP(\%)$ 和洪量相对误差 $REV(\%)$.

《水文情报预报规范》^[20] 规定:降雨径流预报以实测洪峰流量的 20% 作为许可误差. 因此, REP 的许可误差为 $\pm 20\%$.

将洪量平铺于流域面积,可以计算出相应的径流深,因此,径流深误差与洪量误差相等. 《水文情报预报规范》^[20] 规定:径流深预报以实测值的 20% 作为许可误差. 因此,可以认为 REV 的许可误差为 $\pm 20\%$.

REP 和 REV 均评价了预报结果的偏差($Bias$, 预报是否偏大/偏小于观测).

(2) 洪峰滞时 ET .

洪峰滞时指预测洪峰滞后于实测洪峰的时段长. 《水文情报预报规范》^[20] 规定:滞时绝对值的许可误差为预见期的 30%, 许可误差小于 3 h 或一个计算时段长时,则以 3 h 或一个计算时段长作为许可误差.

(3) 确定性系数/纳什系数^[21] (Deterministic Coefficient/Nash-Sutcliffe Efficiency, NSE).

$$NSE = 1 - \frac{\sum_{i=1}^N [Q_{\text{obs}}(i) - Q_{\text{pre}}(i)]^2}{\sum_{i=1}^N [Q_{\text{obs}}(i) - \bar{Q}_{\text{obs}}]^2} \quad (1)$$

式中, $Q_{\text{obs}}(i)$ 为第 i 时刻的实测流量, m^3/s ; $Q_{\text{pre}}(i)$ 为第 i 时刻的倾向值预报流量, m^3/s ; \bar{Q}_{obs} 为实测流量的平均值, m^3/s ; N 为洪水过程时段数.

NSE 评价了洪水过程的拟合效果,其取值范围是 $(-\infty, 1]$, NSE 越接近于 1, 表明洪水过程的拟合效果越好. 除了 NSE 以外,均方误差 (Mean Squared Error, MSE)^[22] 也可以用来评估洪水过程的拟合情况:

$$MSE = \frac{1}{N} \sum_{i=1}^N [Q_{\text{obs}}(i) - Q_{\text{pre}}(i)]^2 \quad (2)$$

理论上, NSE 与 MSE 存在转化关系^[22]:

$$NSE = 1 - \frac{MSE}{\sigma_{\text{obs}}^2} \quad (3)$$

式中, σ_{obs}^2 为实测流量的方差.

由公式(2)和(3)可知, MSE 的取值受到了流量级别的影响,不利于对比不同量级洪水过程的拟合效果;而 NSE 在一定程度上消除了流量量级的影响,因此,在实际应用中 NSE 比 MSE 应用更为广泛,更具说服力. 对 NSE 进行分解,可以将 NSE 分解为 3 个评价因子^[22]:

$$NSE = 2 \cdot \alpha \cdot r - \alpha^2 - \beta^2 \quad (4)$$

$$\alpha = \frac{\sigma_{\text{pre}}}{\sigma_{\text{obs}}} \quad (5)$$

$$\beta = \frac{\mu_{\text{pre}} - \mu_{\text{obs}}}{\sigma_{\text{obs}}} \quad (6)$$

$$r = \frac{Cov_{\text{pre,obs}}}{\sigma_{\text{obs}} \cdot \sigma_{\text{pre}}} \quad (7)$$

式中, σ_{pre} 为倾向值预报的标准差; σ_{obs} 为实测流量的标准差; μ_{pre} 为倾向值预报的均值; μ_{obs} 为实测流量的均值; $Cov_{\text{pre,obs}}$ 为倾向值预报与实测流量的协方差; α 为方差因子; β 为均值因子; r 为预报与实测的线性相关系数.

由公式(4)~(7)可知, NSE 评估了倾向值预报与实测流量之间均值和标准差的差异,同时考虑了两者的线性相关程度. 当倾向值可以完美拟合实测流量过程时(完美预报),两者的均值相等、标准差相等、线性

相关程度最高, 即 $\alpha = 1; \beta = 0; r = 1$, 此时 $NSE = 1$. 根据 NSE 的评价因子 (β, α, r), Gupta 等^[22] 提出了 Kling-Gupta Efficiency (KEG) 评价指标:

$$KEG = 1 - \sqrt{G_1 + G_2 + G_3} \quad (8)$$

$$G_1 = \left(\frac{\sigma_{pre}}{\sigma_{obs}} - 1 \right)^2 \quad (9)$$

$$G_2 = \left(\frac{\mu_{pre}}{\mu_{obs}} - 1 \right)^2 \quad (10)$$

$$G_3 = (r - 1)^2 \quad (11)$$

式中, 变量含义与式(5)~(7)中变量相同. KEG 的取值范围为 $(-\infty, 1]$, KEG 越接近于 1, 表明洪水过程的拟合效果越好.

类似于 NSE , KEG 也考虑了均值因子 G_2 、方差因子 G_1 和线性相关性因子 G_3 共 3 个评价因子. 不同于 NSE , KEG 中因子对 KEG 的影响关系更直观: KEG 与 3 个因子均为单调递减关系. 因此, 通过计算 KEG 及其评价因子, 可以提供预报流量更细致的精度评价.

(4) 基准系数^[23] (Benchmark Efficiency, BE).

$$BE = 1 - \frac{\sum_{i=1}^N [Q_{obs}(i) - Q_{pre}(i)]^2}{\sum_{i=1}^N [Q_{obs}(i) - Q_b(i)]^2} \quad (12)$$

式中, $Q_b(i)$ 为基准预报流量, m^3/s ; 其他变量含义同前.

基准系数 BE 评价了模型的预报技巧 (Forecast Skill, 预报相对于某一标准是否具有较高的精度), 一般用于对比倾向值预报 Q_{pre} 与基准预报 Q_b 在洪水过程拟合中的相对好坏: $BE = 0$ 说明 Q_{pre} 与 Q_b 在洪水过程拟合中表现相当; $BE > 0$ 说明 Q_{pre} 比 Q_b 更优; $BE < 0$ 说明 Q_{pre} 比 Q_b 拟合效果差. BE 在形式上与 NSE 相似, 可以认为 NSE 是以实测平均值作为基准预报的 BE 值.

1.2 可靠度评价指标

Murphy^[14] 将预报量条件分布函数 $f(\hat{Q}|Q_{sim})$ 与真实流量的分布函数 $f(\hat{Q})$ 的一致性定义为可靠性 (Reliability). 由于 $f(\hat{Q})$ 往往无法获得, 因此往往通过评估 $f(\hat{Q}|Q_{sim})$ 对观测值 Q_{obs} 的“模拟预报能力”来评价概率预报的可靠度. 一般的, 以置信度为 90% 的区间预报为代表, 常采用离散度、覆盖率等指标对区间预报的合理性进行评价, 进而反映 $f(\hat{Q}|Q_{sim})$ 对观测值 Q_{obs} 的“模拟预报能力”.

(1) 区间离散度^[24-26]:

$$DI = \frac{\sum_{i=1}^N D(i)}{N} \quad (13)$$

式中, DI 指某一置信度条件下预报区间的平均离散度; N 为预报时刻数; $D(i)$ 指第 i 时刻预报区间的离散度 (预报区间宽度与实测的比值), 具体计算如下:

$$D(i) = \frac{q_u(i) - q_d(i)}{Q_{obs}(i)} \quad (14)$$

式中, $q_u(i)$ 、 $q_d(i)$ 指预报区间的上限和下限, m^3/s ; 其他变量含义同前.

$D(i)$ 的计算过程消除了流量级别对区间宽度 ($b(i) = q_u(i) - q_d(i)$) 的影响, 因此 DI 的取值不受流量级别影响, 可用以对比不同量级洪水的区间预报结果; 离散度越小区间预报效果越好, 但是有关 DI 阈值的研究仍较为匮乏. 我国《水文情报预报规范》^[20] 中规定洪峰预报允许误差为 20%, 借鉴这一概念, 本文将洪峰处的区间离散度定义为洪峰离散度 D_{peak} , 并确定其阈值如下:

将洪峰预报区间边界与实测洪峰相对距离的最大允许值定为 0.2, 即:

$$\frac{|q_{u, peak} - Q_{obs, peak}|}{Q_{obs, peak}} \leq 0.2 \quad (15)$$

$$\frac{|Q_{\text{obs,peak}} - q_{\text{d,peak}}|}{Q_{\text{obs,peak}}} \leq 0.2 \quad (16)$$

式中, $q_{\text{u,peak}}$ 和 $q_{\text{d,peak}}$ 分别为洪峰预报区间的上下边界, m^3/s ; $Q_{\text{obs,peak}}$ 为洪峰观测流量, m^3/s .

当 $q_{\text{d,peak}} \leq Q_{\text{obs,peak}} \leq q_{\text{u,peak}}$ 时, 即实测洪峰位于区间预报范围内, 将式(15)和(16)相加可知, 洪峰离散度 $D_{\text{peak}} = \frac{q_{\text{u,peak}} - q_{\text{d,peak}}}{Q_{\text{obs,peak}}} \leq 0.4$. 当 $q_{\text{d,peak}} \leq q_{\text{u,peak}} \leq Q_{\text{obs,peak}}$ 时, $D_{\text{peak}} < \frac{Q_{\text{obs,peak}} - q_{\text{d,peak}}}{Q_{\text{obs,peak}}} \leq 0.2 < 0.4$; 同理可知, 当 $Q_{\text{obs,peak}} \leq q_{\text{d,peak}} \leq q_{\text{u,peak}}$ 时, $D_{\text{peak}} < 0.4$. 因此, 本研究将洪峰离散度的最大允许值确定为 0.4.

(2) 区间覆盖率^[24-25] (Containing Ratio, CR):

$$CR = \frac{\sum_{i=1}^N k(i)}{N} \quad (17)$$

$$k(i) = \begin{cases} 1, & (q_{\text{d}}(i) \leq Q_{\text{obs}}(i) \leq q_{\text{u}}(i)) \\ 0, & (Q_{\text{obs}}(i) < q_{\text{d}}(i) \text{ 或 } Q_{\text{obs}}(i) > q_{\text{u}}(i)) \end{cases} \quad (18)$$

理论上来说, CR 值越接近区间置信度, 概率预报结果越合理. 如置信度为 90% 的区间预报结果, 其 CR 值越接近 90%, 该预报区间越合理.

(3) 覆盖率判定系数 (Containing Ratio Coefficient, CRC).

区间覆盖率 CR 只能考虑某一个置信度 (如 90%) 预报区间包含实测的情况, 并不能代表整个概率预报结果的合理性, 因此, 本文提出覆盖率判定系数 CRC 评价指标, 对所有置信度区间预报结果做出整体评估.

根据区间覆盖率的含义, 当概率预报为完美预报时, 对所有的置信度 $X = 0, \dots, 100\%$, 其对应的覆盖率 CR, 应满足 $CR = X$. 点绘 (X, CR) 时, 所有的点据应位于 45° 线上. 然而, 实际中不可能做到完美预报, (X, CR) 点往往散落在 45° 线附近, 此时, 可以计算这些点据与 45° 线的确定性系数 R (即 CRC), 来判断概率预报的合理性, 覆盖率判定系数计算公式如下:

$$CRC = 1 - \frac{\sum_{j=1}^F (CR_j - X_j)^2}{\sum_{j=1}^F (X_j - \bar{X})^2} \quad (19)$$

式中, X_j 为置信度, $X_j \in (0, 100\%]$; \bar{X} 为置信度的均值; CR_j 是对应于置信度 X_j 的区间覆盖率; F 为置信度个数, 理论上 F 应取无穷大, 即 X_j 为无穷多个, 实际应用时, 建议置信度 X_j 在 $[10\%, 90\%]$ 区间内间隔 5% 选取, 即 $X_j = 90\%、85\%、\dots、15\%、10\%$, 此时 $F = 17$.

CRC 的取值范围是 $(-\infty, 1]$, $CRC = 1$ 为完美概率预报. 当 CRC 大于某一阈值时, 线性关系 $CR = X$ 显著, 概率预报结果合理, 而这一阈值随实际研究问题不同而变化, 应通过统计检验来确定. 在水文实际应用中, 一般认为相关系数 $r \geq 0.8$ 时变量存在较显著线性相关性, 根据 r 与确定性系数 R 的关系 ($r^2 = R$), 本文认为在不便于采用统计检验方式确定 CRC 阈值时, 可以将 0.64 作为 CRC 的阈值, 认为 $CRC \geq 0.64$ 时 45° 线可以较好的拟合 (X, CR) 点据, 即概率预报结果合理.

由 CRC 计算公式可知, CRC 同时考虑了所有置信度条件下的区间预报结果, 可以衡量概率预报的整体合理性.

就思路而言, CRC 与 Q-Q (Quantiles-Quantiles) 图方法^[27]、概率图 (Probability Plot) 法^[28] 类似, 三者的 45° 线均代表了完美预报. 然而, Q-Q 图方法评价的是各个分位数的合理性, 概率图法评价了累计概率的合理性, CRC 考虑的是区间覆盖率的合理性, 三者的关注点不同. 考虑到实际洪水概率预报中多关注区间预报的覆盖率, 因此本文认为 CRC 方法较 Q-Q 图方法、概率图方法更适用于实际生产应用.

(4) 单位区间离散化系数.

一般而言, 区间宽度越大, 离散度越大, 区间覆盖率越大, 反之亦然, 区间离散度和覆盖率在某种程度上是一组矛盾的评价指标. 为了解决这一矛盾, 提出了单位区间离散化系数^[17, 29] (Percentage of observations bracketed by the Unit Confidence Interval, PUCI), 用以描述“平均相对区间宽度”所包含的实测点据比例.

$$PUCI_j = \frac{1 - |CR_j - X_j|}{DI_j} \quad (20)$$

式中变量含义同前. 理论上, $PUCI$ 的取值范围是 $[0, \infty)$. $PUCI$ 值越大, 表明该置信度的区间预报结果合理性越差.

对不同置信度预报区间的 $PUCI$ 值进行平均, 可以得到平均单位区间离散化系数 (ACI):

$$ACI = \frac{1}{F} \sum_{j=1}^F PUCI_j \quad (21)$$

式中变量含义同前. 通过计算 ACI 值, 可以估计区间预报的整体合理性: ACI 越小, 概率预报结果越合理. 然而, 由于 ACI 的允许误差很难界定, 因此不建议采用 ACI 对概率预报结果合理性进行评估.

(5) 连续排位评分 (Continuous Ranked Probability Score, $CRPS$) 和连续排位技巧评分 (Continuous Ranked Probability Skill Score, $CRPSS$)

$CRPS$ ^[30-31] 评估了概率预报的预报能力. 对任一时刻, $CRPS$ 计算方式如下:

$$CRPS = \int_{-\infty}^{\infty} [\tau(y) - \tau_0(y)]^2 dy \quad (22)$$

式中, y 为预报量, 即流量; $\tau(y)$ 为该时刻的预报累计概率分布; $\tau_0(y) = \begin{cases} 0, y < Q_{obs} \\ 1, y \geq Q_{obs} \end{cases}$.

$CRPS$ 值^[31-32] 越小, 说明预测分布在观测值附近越呈现“尖瘦”形态, 预测分布的“准确性”越高. 但目前研究尚未给出 $CRPS$ 的合理阈值, 往往采用 $CRPS$ 比较多个预报的相对预报能力^[33], 进而提出了 $CRPSS$ ^[19], 其计算公式如下:

$$CRPSS(\text{fore}, \text{ref}) = \frac{CRPS_{\text{ref}} - CRPS_{\text{fore}}}{CRPS_{\text{ref}}} \quad (23)$$

式中, $CRPS_{\text{fore}}$ 为待评价概率预报的 $CRPS$ 值; $CRPS_{\text{ref}}$ 为参考预报的 $CRPS$ 值, 对于确定性预报, 其 $CRPS$ 值即为平均绝对误差^[30].

连续排位技巧评分 $CRPSS$ 是一种技巧 (Skill) 评分指标, 取值范围是 $(-\infty, 1]$. 当 $CRPSS > 0$ 时, 概率预报较参考预报具有更高的预报能力; 当 $CRPSS < 0$ 时, 待评价概率预报的“准确性”不如参考预报; 当 $CRPSS = 0$ 时, 待评价概率预报与参考预报预报能力相当.

(6) 区间对称度和对称比

对称度^[24] A_1 评估了某一置信度预报区间关于实测流量的对称性, 其计算公式如下:

$$A_1 = \frac{1}{N} \sum_{i=1}^N \lambda_1(i) \quad (24)$$

$$\lambda_1(i) = |h(i) - 0.5| \quad (25)$$

$$h(i) = \frac{q_u(i) - Q_{obs}(i)}{b(i)} \quad (26)$$

式中变量含义同前. 由 A_1 计算公式可知, 当 i 时刻的预报区间关于实测流量对称时, $h(i) = 0.5$, $\lambda_1(i) = 0$. A_1 值越小, 预报区间关于实测对称性越强. 一般的, 要求 $A_1 < 0.5$. 此外, 还可以采用平均对称度 A_2 对区间对称性进行评价:

$$A_2 = \frac{1}{N} \sum_{i=1}^N \lambda_2(i) \quad (27)$$

$$\lambda_2(i) = \left(\frac{|(q_u(i) - Q_{obs}(i))^3 - (q_d(i) - Q_{obs}(i))^3|}{b(i)} \right)^{\frac{1}{3}} \quad (28)$$

式中变量含义同前. A_2 值越小, 预报区间关于实测对称性越强, 当 $A_2 = 0$ 时, 预报区间关于实测流量完全对称. 一般的, 要求 $A_2 < 1$.

区间对称比^[15] A_3 采用实测值大于区间上界的个数与小于区间下界个数的比值来衡量预报区间关于实测流量的对称性:

$$\Lambda_3 = \frac{\sum_{i=1}^N J_u(i)}{\sum_{i=1}^N J_d(i)} \quad (29)$$

$$J_u = \begin{cases} 1(Q_{\text{obs}}(i) > q_u(i)) \\ 0(Q_{\text{obs}}(i) \leq q_u(i)) \end{cases} \quad (30)$$

$$J_d = \begin{cases} 1(Q_{\text{obs}}(i) < q_d(i)) \\ 0(Q_{\text{obs}}(i) \geq q_d(i)) \end{cases} \quad (31)$$

由于区间预报结果往往并不要求对称,同时,区间覆盖率与离散度已经在某种程度上反映了预报区间相对实测流量的对称性.因此,对称性评价指标(对称度和对称比)在实际生产中的应用较少.

1.3 建议的“精度-可靠度”联合评价指标

由于洪水概率预报结果一般可以分为倾向值预报和区间预报两个部分,因此,应采用“精度-可靠度”联合评价方式对洪水概率预报结果进行评估,即分别评价倾向值预报的准确性(精度评价)和区间预报的合理性(可靠度评价),具体评价指标如下:

在精度评价时,根据前文的介绍与分析,参考《水文情报预报规范》^[20]的相关规定,建议采用:纳什系数 NSE 评价洪水过程,洪峰误差 REF 、洪峰滞时 ET 评价洪峰拟合情况,洪量误差 REV 评估水量平衡情况,基准系数 BE 比较不同定值预报(倾向值/确定性)对洪水过程的拟合情况.当需要对定值预报过程进行更为细致的评价时,可以采用 KGE 及其 3 个评价因子 G_1 、 G_2 、 G_3 对洪水过程的拟合效果进行评价.

在“可靠度”评价时,常关注置信度为 90% 的区间预报结果,因此,建议采用:洪峰处区间离散度 D_{peak} 、单位区间离散化系数 $PUCI$ 和覆盖率 CR 对该区间预报结果进行评价.由于 90% 预报区间仅是众多预报区间中的一个,并不能代表其他置信度的预报区间结果,因此,应采用覆盖率判定系数 CRC 给出概率预报结果合理性的整体评估.此外,可以采用连续排位技巧评分 $CRPSS$ 对比多个概率预报结果的相对预报能力.

相关评价指标的具体计算公式见前文,此处不再赘述.

2 洪水概率预报评价指标应用

以淮王家坝区间流域为例,在新安江模型预报基础上,采用水文不确定性处理器 (Hydrologic Uncertainty Processor, HUP)^[11] 进行洪水概率预报,并采用“精度-可靠度”联合评价指标体系对洪水概率预报结果进行评估.

2.1 研究区概况

王家坝区间流域集水面积为 7110 km². 区域内年平均降水量 800~1200 mm,降水年际变化大,时空分布不均匀.洪水主要来源为上断面(息县、潢川、班台)河道来水和区间降雨产生的径流,区间降雨的汇流时间约 72 h(雨峰至洪峰出现时间).研究区域示意图见图 1.

2.2 确定性预报

新安江模型是河海大学赵人俊^[34]教授设计的国内第一个完整的流域水文模型.三水源新安江模型蒸发计算采用三层模型;产流计算采用蓄满产流模型;采用自由水蓄水库结构将总径流划分为地表径流、壤中流和地下径流 3 种;流域汇流计算采用线性水库;河网汇流采用滞洪演算法;河道汇流采用马斯京根分段连续演算法.三水源新安江模型共有 17 个参数,包括 7 个产流参数、10 个汇流参数.

以王家坝区间流域洪水为研究对象,上断面河道来水采用马斯京根法进行汇流计算.选用 1990—2013 年共 24 年的资料进行日模型率定,对其间的 12 场洪水资料进行次洪模型计算(其中 8 场洪水用于率定,4 场洪水用于验证),本文以计算时段为 3 h 为例开展相关研究,新安江模型参数见表 1,确定性预报精度统计见表 2.

由表 2 可知:王家坝区间流域新安江模型预报的洪峰误差在 4%~31%,最小洪峰误差为 4.07%,有两场洪水洪峰误差超过 20%,合格率为 83%.12 场洪水的洪量误差均在 15% 以内;洪峰滞时在 3 个时段(9 h)以内;确定性系数均大于 0.75.综合而言,新安江模型预报洪水精度尚可.

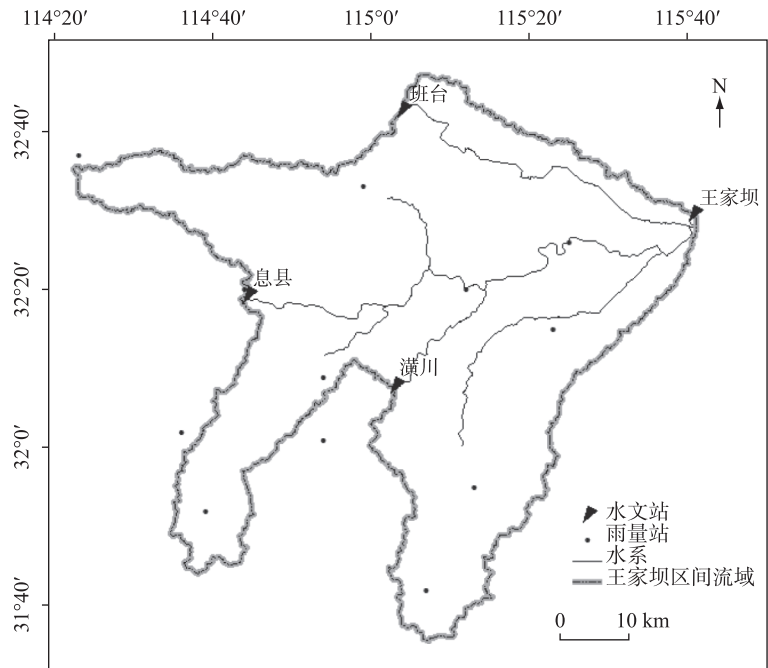


图 1 研究区域示意图
Fig.1 The sketch of research area

表 1 王家坝区间流域新安江模型参数
Tab.1 The parameters of Xinanjiang model in Wangjiaba subbasin

层次	符号	参数意义	参数取值
蒸散发计算	KC	流域蒸散发折算系数	1.4
	WUM	上层张力水容量/mm	20
	WLM	下层张力水容量/mm	90
	C	深层蒸散发折算系数	0.3
产流计算	WM	流域平均张力水容量/mm	150
	B	张力水蓄水容量曲线方次	0.4
	IM	不透水面积占全流域面积的比例	0.03
水源划分	SM	表层自由水蓄水容量/mm	5.63
	EX	表层自由水蓄水容量曲线方次	1.5
	KG	表层自由水蓄水库对地下水的日出流系数	0.5
	KI	表层自由水蓄水库对壤中流的日出流系数	0.2
汇流计算	CI	壤中流消退系数	0.88
	CG	地下水消退系数	0.98
	CS	河网蓄水消退系数	0.89
	L	河网汇流滞时/h	5
	KE	马斯京根法演算参数/h	3
	XE	马斯京根法演算参数	-2.35
	N	河段数	11

表2 新安江模型精度统计
Tab.2 The performance of Xinanjiang model

时期	洪号	洪峰误差/%	洪量误差/%	洪峰滞时 <i>ET</i>	<i>NSE</i>
率定期	19910524	-4.71	-10.83	2	0.94
	19910629	30.24	-2.03	-1	0.88
	19950707	13.65	-2.39	1	0.96
	19980509	-13.95	-13.58	3	0.93
	19990622	10.19	9.85	2	0.76
	20020506	-16.98	-13.32	0	0.90
	20030719	4.12	-6.30	3	0.93
	20120907	8.62	5.19	-1	0.96
验证期	19910612	25.66	4.59	-2	0.88
	19980630	6.27	5.26	1	0.97
	20031005	4.07	-1.10	-1	0.95
	20050513	-18.74	-14.86	0	0.91

2.3 概率预报模型

HUP 模型^[11]由 Krzysztofowicz 和 Herr(2001 年)提出,已被广泛用于洪水概率预报研究中^[35-38],本文仅对该模型进行简单介绍,其他细节可参考文献[11,39-40].

HUP 模型中,首先采用正态分位数转换技术^[41](Normal Quantile Transform)将确定性预报(新安江模型预报)结果与实测结果转换至正态空间.然后在正态空间进行线性似然函数假定,并根据贝叶斯理论推求预报量的后验分布函数.最后通过正态分位数逆转换过程,将正态空间的预测分布函数转换至原始空间,进而实现概率预报.预报量 h_n 的后验概率密度函数可以表示为:

$$\Phi(h_n | s_n, h_0) = \Omega^{-1}\left(\frac{\Omega(\Gamma_n(h_n)) - A_n \Omega(\Gamma_s(s_n)) - D_n \Omega(\Gamma_0(h_0)) - B_n}{T_n}\right) \quad (32)$$

式中, s_n 为 n 时刻的预报量; n 代表预见期,本文预见期为 3 h(与确定性预报一致); h_0 为起报时刻实测流量值; A_n 、 B_n 、 D_n 、 T_n 为模型参数,结果见表 3; $\Omega(\cdot)$ 表示正态分位数转换, $\Omega^{-1}(\cdot)$ 表示正态分位数的逆转换过程; $\Gamma(\cdot)$ 表示新安江模型预报或实测流量的边际分布,一般采用三参数对数威布尔分布,对数威布尔分布函数形式如下:

$$F(x) = 1 - e^{-\left(\frac{\ln(x)}{a}\right)^b} \quad (33)$$

式中, x 为新安江模型预报流量或实测流量; a 、 b 、 c 为对数威布尔分布的 3 个参数.采用上述 8 场洪水率定 HUP 模型相关参数,参数率定结果见表 3,边际分布拟合情况如图 2 所示.

由图 2 可知,对数威布尔分布理论概率曲线可以很好地拟合流量的经验累积概率.计算对数威布尔分布与经验概率的平均绝对值误差: h_0 、 h_1 和 s_1 的平均绝对值误差均不超过 0.02,因此,概率预报模型的边际分布拟合效果较好.

2.4 概率预报评价

在精度评价中,采用洪峰误差 *REF*、洪量误差 *REV*、洪峰滞时 *ET* 对 HUP 模型的倾向值预报(Q50 预报值和均值预报)进行评估,并采用基准系数 *BE* 对比倾向值预报与新安江模型预报结果,相关数据见表 4.

由表 4 可知:倾向值预报的 *REF* 在 6% 以内,小于新安江模型洪峰误差;倾向值预报的 *REV* 在 7% 以内;洪峰滞时 *ET* 在 1 个时段(3 h)以内.通过倾向值预报的 *REF*、*REV* 和 *ET* 比较结果可知,概率预报的 Q50 和

表3 HUP 模型相关参数

Tab.3 The parameters of HUP model

分类	参数符号	参数取值
h_0 边际分布	a	4.95538
	b	5.20253
	c	2.14912
h_n 边际分布	a	4.84989
	b	5.11157
	c	2.25552
s_n 边际分布	a	4.84766
	b	4.79721
	c	2.16981
后验分布	A_n	0.11166
	B_n	0.00001
	D_n	0.88879
	T_n	0.00274

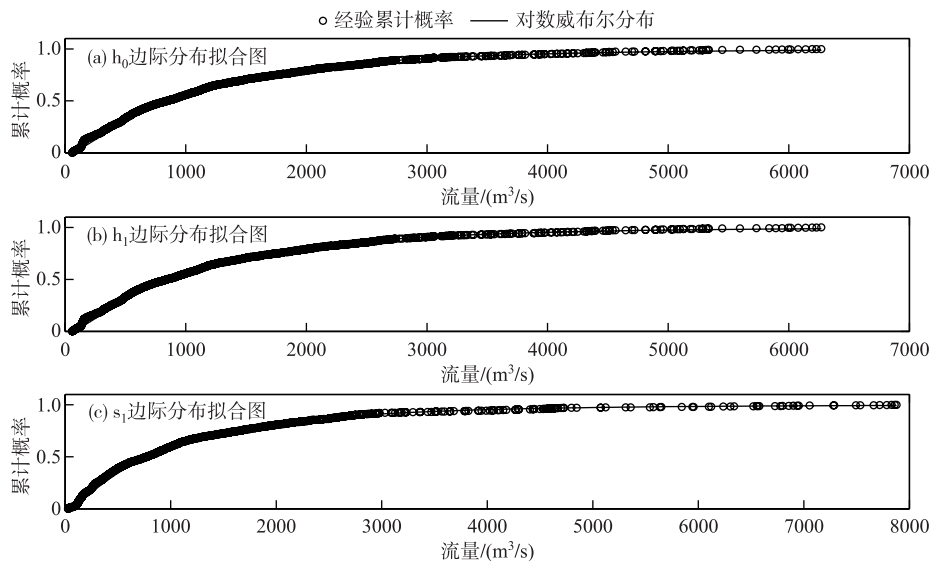


图 2 边际分布拟合图
Fig.2 Fitting graph of marginal distribution

表 4 洪水概率预报精度评价
Tab.4 The accuracy assessment of probabilistic flood forecast

时期	洪号	洪峰误差/%		洪量误差/%		洪峰滞时 ET		基准系数	
		Q50	均值	Q50	均值	Q50	均值	Q50	均值
率定期	19910524	-0.34	2.09	-0.56	2.07	0	0	0.85	0.81
	19910629	2.85	5.08	-0.02	2.40	-1	-1	0.93	0.91
	19950707	1.50	3.98	0.30	2.99	0	0	0.85	0.78
	19980509	-1.77	0.63	-0.80	1.90	-1	-1	0.90	0.89
	19990622	2.53	5.64	3.72	6.98	-1	-1	0.97	0.92
	20020506	-0.99	1.77	-0.68	2.23	-1	-1	0.83	0.81
	20030719	-0.17	2.06	-0.50	1.92	-1	-1	0.86	0.83
	20120907	1.44	4.03	1.55	4.41	-1	-1	0.80	0.69
验证期	19910612	2.33	4.48	0.58	2.96	-1	-1	0.94	0.91
	19980630	0.56	2.82	0.83	3.37	-1	-1	0.83	0.74
	20031005	0.79	3.25	0.54	3.26	-1	-1	0.85	0.74
	20050513	-0.34	2.09	-0.56	2.07	0	0	0.85	0.81

均值预报均比新安江模型预报精度更高. 此外,倾向值预报的 BF 均大于 0,说明 Q50 预报和均值预报的洪水过程拟合效果比新安江模型更好. 为了更细致的对比倾向值预报的精度,分别计算了新安江模型和倾向值预报(Q50 和均值)的 KGE 及其 3 个评价因子取值,见表 5.

由表 5 可知:概率预报倾向值预报(Q50 和均值)的 KGE 均比新安江模型大,说明 Q50 和均值对实际流量过程的拟合精度较新安江模型预报更高. 此外,新安江模型的 3 个评价因子取值较小,说明新安江模型预报在方差、均值和线性相关性方面的模拟精度较高,而倾向值预报(Q50 和均值)的 3 个评价因子几乎都小于新安江预报,说明倾向值预报在均值、方差和线性相关性的模拟精度都大于新安江模型.

在可靠度评价中,采用洪峰处区间离散度 D_{peak} 、覆盖率 CR 和单位区间离散化系数 $PUCI$ 对置信度为 90% 的区间预报结果进行评价,并采用覆盖率判定系数 CRC 对概率预报结果进行整体评估. 概率预报可靠

度评价结果见表 6. 图 3 点绘了验证期 4 场洪水置信度分别为 90%、85%、……、10%时,对应的区间覆盖率.

表 5 倾向值预报 KGE 计算结果
Tab.5 The KGE of preferred prediction

时期	洪号	KGE			G_1			G_2			G_3		
		新安江	Q50	均值	新安江	Q50	均值	新安江	Q50	均值	新安江	Q50	均值
率定期	19910524	0.89	0.99	0.96	4E-04	3E-05	9E-04	1E-02	3E-05	4E-04	2E-05	2E-05	2E-05
	19910629	0.78	0.98	0.95	5E-02	3E-04	1E-03	4E-04	2E-08	6E-04	9E-04	1E-05	1E-05
	19950707	0.86	0.98	0.95	2E-02	2E-04	2E-03	6E-04	9E-06	9E-04	5E-05	6E-06	6E-06
	19980509	0.83	0.99	0.98	1E-02	1E-04	2E-04	2E-02	6E-05	4E-04	1E-04	1E-05	1E-05
	19990622	0.84	0.96	0.92	7E-03	1E-04	2E-03	1E-02	1E-03	5E-03	9E-03	5E-06	5E-06
	20020506	0.79	0.99	0.96	3E-02	8E-07	7E-04	2E-02	5E-05	5E-04	1E-04	7E-05	7E-05
	20030719	0.87	0.99	0.96	1E-02	1E-04	1E-03	4E-03	2E-05	4E-04	3E-04	2E-05	2E-05
	20120907	0.83	0.97	0.94	3E-02	5E-04	2E-03	3E-03	2E-04	2E-03	3E-05	1E-05	1E-05
验证期	19910612	0.72	0.98	0.95	8E-02	6E-04	2E-03	2E-03	3E-05	9E-04	3E-04	1E-05	1E-05
	19980630	0.92	0.99	0.96	4E-03	6E-05	9E-04	3E-03	7E-05	1E-03	9E-05	6E-06	6E-06
	20031005	0.83	0.98	0.94	3E-02	5E-04	2E-03	1E-04	3E-05	1E-03	6E-05	1E-05	1E-05
	20050513	0.80	0.99	0.97	2E-02	5E-05	4E-04	2E-02	4E-05	5E-04	8E-05	2E-05	2E-05

表 6 洪水概率预报可靠度评价
Tab.6 The reliability assessment of probabilistic flood forecast

时期	洪号	90% 洪峰离散度	90% 覆盖率/%	90% $PUCI$	CRC
率定期	19910524	0.15	85.29	0.24	0.96
	19910629	0.14	88.10	0.24	0.71
	19950707	0.15	83.65	0.30	0.71
	19980509	0.15	91.07	0.21	0.68
	19990622	0.19	88.50	0.27	0.90
	20020506	0.17	80.36	0.33	0.82
	20030719	0.14	87.50	0.27	0.95
	20120907	0.16	83.04	0.32	0.92
验证期	19910612	0.13	86.67	0.31	0.92
	19980630	0.14	92.50	0.21	0.84
	20031005	0.15	95.31	0.20	0.91
	20050513	0.16	95.45	0.23	0.83

由表 6 可知,置信度为 90% 的区间预报:洪峰离散度均在 20% 以内;覆盖率在 80.36%~95.45%,接近 90%;单位区间离散化系数 $PUCI$ 在 0.35 以内,数值较小. 因此认为概率预报置信度 90% 区间预报结果较为合理. 同时, CRC 均在 0.64 以上,因此,概率预报结果整体上是合理的.

由图 3 可知:4 场洪水的点据总体上位于 45° 线(图中实线)附近,仅在置信度小于 50% 左右时, (X, CR) 点据与 45° 线的距离有所偏大. 总体而言,验证期 4 场洪水的概率预报结果比较合理. 图 4 以 90% 置信度的区间预报为例,给出了验证期 4 场洪水的概率预报过程图.

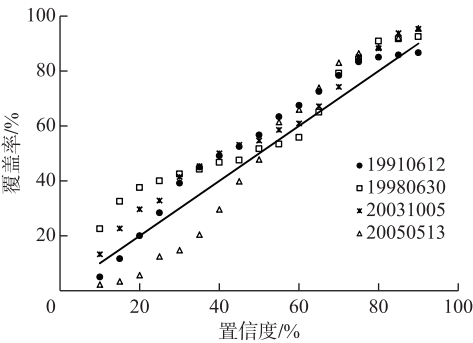


图 3 验证期洪水置信度与覆盖率的关系
Fig.3 X vs. CR for verification floods

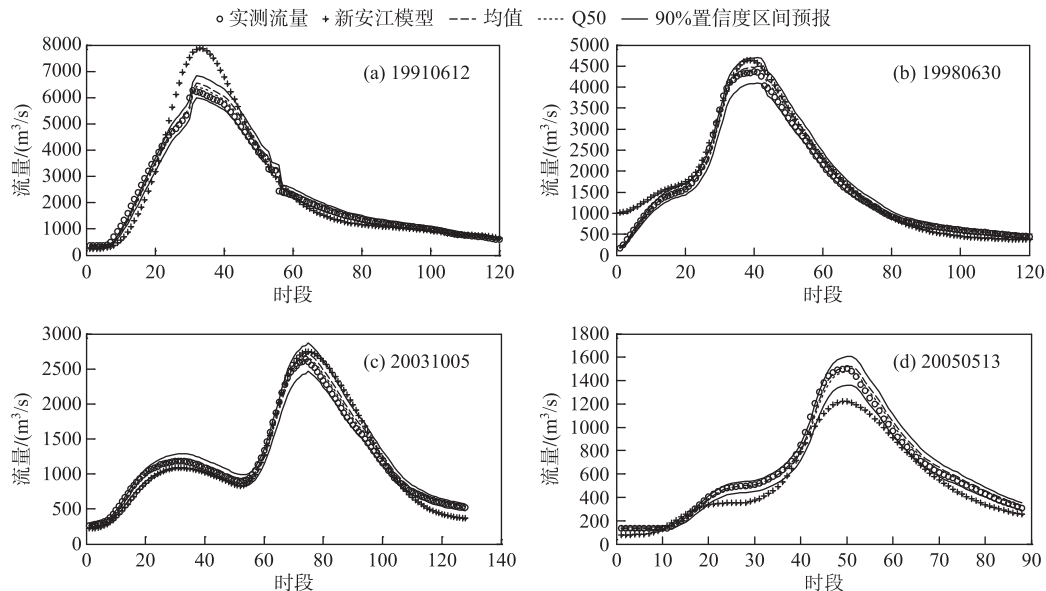


图4 验证期洪水概率预报过程(计算时段为3 h)
Fig.4 The hydrograph of probabilistic forecast for verification floods

3 结论

本文分析并总结了目前洪水概率预报研究中常用的评价指标,在此基础上,建立了“精度-可靠度”联合评价指标体系。在淮河王家坝区间流域,采用新安江模型作为确定性预报模型,采用 HUP 模型进行洪水概率预报,对“精度-可靠度”联合评价指标体系进行应用,主要结论如下:

在“精度”评价时,采用洪峰误差、洪量误差、洪峰滞时和基准系数指标对倾向值预报(Q50和均值)进行评价,结果表明:Q50和均值在洪峰流量、径流量、洪峰出现时间和洪水过程预报中的表现均优于初始确定性预报。

在“可靠度”评价时,采用洪峰离散度、覆盖率和单位区间离散化系数指标可以评价某一置信度(以90%为例)的区间预报结果的合理性。覆盖率判定系数指标同时考虑了多个置信度的区间预报结果,可以用以估计概率预报结果的整体合理性。

“精度-可靠度”联合评价指标体系分别量化了概率预报倾向值的准确性和区间预报的合理性,实现了洪水概率预报结果的综合评价。

4 参考文献

- [1] Ye SZ, Xia J. Century's retrospect and looking into the future of hydrological science. *Advances in Water Science*, 2002, **13**(1): 93-104. [叶守泽, 夏军. 水文科学研究的世纪回眸与展望. 水科学进展, 2002, **13**(1): 93-104.]
- [2] Cloke HL, Pappenberger F. Ensemble flood forecasting: A review. *Journal of Hydrology*, 2009, **375**(3/4): 613-626. DOI: 10.1016/j.jhydrol.2009.06.005.
- [3] Wen YH, Li ZJ, Sun MK *et al.* Influence of rainfall input on real-time flood forecasting accuracy and forecast period. *J Lake Sci*, 2019, **31**(1): 39-51. DOI: 10.18307/2019.0104. [温娅惠, 李致家, 孙明坤等. 降雨输入对实时洪水预报精度与预见期的影响. 湖泊科学, 2019, **31**(1): 39-51.]
- [4] Zhao LN, Liu Y, Dang HF *et al.* The progress on application of ensemble prediction to flood forecasting. *Journal of Applied Meteorological Science*, 2014, **25**(6): 641-653. DOI: 10.11898/1001-7313.20140601. [赵琳娜, 刘莹, 党皓飞等. 集合数值预报在洪水预报中的应用进展. 应用气象学报, 2014, **25**(6): 641-653.]

- [5] Liu KL, Hu YB, Wang YJ *et al.* Performance and parameterization of the BMA model applied in the Huaihe River Basin. *J Lake Sci*, 2017, **29**(6): 1520-1527. DOI: 10.18307/2017.0624. [刘开磊, 胡友兵, 汪跃军等. BMA 集合预报在淮河流域应用及参数规律初探. 湖泊科学, 2017, **29**(6): 1520-1527.]
- [6] Xu XY, Fang HW, Zhang YF *et al.* A real-time probabilistic channel flood forecasting model and application based on particle filters. *Advances in Water Science*, 2015, **26**(3): 356-364. [徐兴亚, 方红卫, 张岳峰等. 河道洪水实时概率预报模型与应用. 水科学进展, 2015, **26**(3): 356-364.]
- [7] Kavetski D, Kuczera G, Franks SW. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. *Water Resources Research*, 2006, **42**(3). DOI: 10.1029/2005WR004368.
- [8] Ajami NK, Duan QY, Sorooshian S. An integrated hydrologic Bayesian multimodel combination framework; Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resources Research*, 2007, **43**(1): W1403. DOI: 10.1029/2005WR004745.
- [9] Han SS, Coulibaly P. Bayesian flood forecasting methods; A review. *Journal of Hydrology*, 2017, **551**: 340-351. DOI: 10.1016/j.jhydrol.2017.06.004.
- [10] Liu ZJ, Guo SL, He SK *et al.* Multivariate hydrologic uncertainty processor based on Copula function. *Shuili Xuebao*, 2018, **49**(3): 332-342. [刘章君, 郭生练, 何绍坤等. 基于 Copula 函数的多变量水文不确定性处理器. 水利学报, 2018, **49**(3): 332-342.]
- [11] Krzysztofowicz R, Herr HD. Hydrologic uncertainty processor for probabilistic river stage forecasting: precipitation-dependent model. *Journal of Hydrology*, 2001, **249**(1/2/3/4): 46-68. DOI: 10.1016/S0022-1694(01)00412-7.
- [12] Todini E. A model conditional processor to assess predictive uncertainty in flood forecasting. *International Journal of River Basin Management*, 2008, **6**(2): 123-137. DOI: 10.1080/15715124.2008.9635342.
- [13] Van Steenbergen N, Ronsyn J, Willems P. A non-parametric data-based approach for probabilistic flood forecasting in support of uncertainty communication. *Environmental Modelling & Software*, 2012, **33**: 92-105. DOI: 10.1016/j.envsoft.2012.01.013.
- [14] Murphy AH. What is a good forecast an essay on the nature of goodness in weather forecasting. *Weather and Forecasting*, 1993, (2): 281-293.
- [15] Ren Z, Sheng D. Study on Uncertainty of Xin'anjiang model parameters based on Multi-objective GLUE Method. *Water Resources and Power*, 2016, **34**(3): 15-18. [任政, 盛东. 基于多目标 GLUE 算法的新安江模型参数不确定性研究. 水电能源科学, 2016, **34**(3): 15-18.]
- [16] Wang YL, Liang ZM, Wang K *et al.* Probabilistic flood forecasting based on multi-model MCP. *South-to-North Water Transfers and Water Science & Technology*, 2018, **16**(6): 39-45. [王艳兰, 梁忠民, 王凯等. 基于多模型 MCP 方法的洪水概率预报. 南水北调与水利科技, 2018, **16**(6): 39-45.]
- [17] Xu W, Jiang HG, Yang X *et al.* Efficiency of Bayesian probabilistic hydrological forecast system based on Box-Cox transformation. *Journal of Hydroelectric Engineering*, 2018, **37**(11): 15-23. DOI: 10.11660/slfdbx.20181102. [徐炜, 姜宏广, 杨洵等. 基于 Box-Cox 变换的贝叶斯概率水文预报效率. 水力发电学报, 2018, **37**(11): 15-23.]
- [18] Liang ZM, Jiang XL, Cao YX *et al.* Probabilistic flood forecasting considering rainfall uncertainty. *Journal of Hohai University: Natural Sciences*, 2016, **44**(1): 8-12. [梁忠民, 蒋晓蕾, 曹炎煦等. 考虑降雨不确定性的洪水概率预报方法. 河海大学学报: 自然科学版, 2016, **44**(1): 8-12.]
- [19] Alfieri L, Pappenberger F, Wetterhall F *et al.* Evaluation of ensemble streamflow predictions in Europe. *Journal of Hydrology*, 2014, **517**: 913-922. DOI: 10.1016/j.jhydrol.2014.06.035.
- [20] GB/T 22482-2008, Standard for hydrological information and hydrological forecasting. [GB/T 22482-2008, 水文情报预报规范.]
- [21] Nash J, Sutcliffe J. River flow forecasting through conceptual models. Part I: A discussion of principles. *Journal of Hydrology*, 1970, (10): 282-290.
- [22] Gupta HV, Kling H, Yilmaz KK *et al.* Decomposition of the mean squared error and NSE performance criteria; Implications for improving hydrological modelling. *Journal of Hydrology*, 2009, **377**(1/2): 80-91. DOI: 10.1016/j.jhydrol.2009.08.003.
- [23] Schaefli B, Gupta HV. Do Nash values have value? *Hydrological Processes*, 2007, **21**(15): 2075-2080. DOI: 10.1002/hyp.6825.

- [24] Xiong LH, Wan M, Wei XJ *et al.* Indices for assessing the prediction bounds of hydrological models and application by generalised likelihood uncertainty estimation. *Hydrological Sciences Journal*, 2009, **54**(5): 852-871. DOI: 10.1623/hysj.54.5.852.
- [25] Li L, Xia J, Xu CY *et al.* eds. Analyse the sources of equifinality in hydrological model using GLUE methodology. In: *Hydroinformatics in Hydrology, Hydrogeology and Water Resources*. Italy: IAHS Press, 2009: 130-138.
- [26] Jin XL, Xu CY, Zhang Q *et al.* Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model. *Journal of Hydrology*, 2010, **383**(3/4): 147-155. DOI: 10.1016/j.jhydrol.2009.12.028.
- [27] Renard B, Kavetski D, Kuczera G *et al.* Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resources Research*, 2010, **46**(5). DOI: 10.1029/2009WR008328.
- [28] Laio F, Tamea S. Verification tools for probabilistic forecasts of continuous hydrological variables. *Hydrol Earth Syst Sci*, 2007, **11**: 1267-1277. DOI: 10.5194/hess-11-1267-2007.
- [29] Li L, Xu CY, Xia J *et al.* Uncertainty estimates by Bayesian method with likelihood of AR (1) plus Normal model and AR (1) plus Multi-Normal model in different time-scales hydrological models. *Journal of Hydrology*, 2011, **406**(1/2): 54-65. DOI: 10.1016/j.jhydrol.2011.05.052.
- [30] Matheson JE, Winkler RL. Scoring rules for continuous probability distributions. *Management Science*, 1976, **22**: 1087-1095.
- [31] Hersbach H. Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting*, 2000, **15**(5): 559-570. DOI: 10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2.
- [32] Biondi D, De Luca DL. Performance assessment of a Bayesian Forecasting System (BFS) for real-time flood forecasting. *Journal of Hydrology*, 2013, **479**: 51-63. DOI: 10.1016/j.jhydrol.2012.11.019.
- [33] Pappenberger F, Ramos MH, Cloke HL *et al.* How do I know if my forecasts are better? Using benchmarks in hydrological ensemble prediction. *Journal of Hydrology*, 2015, **522**: 697-713. DOI: 10.1016/j.jhydrol.2015.01.024.
- [34] Zhao RJ ed. Basin hydrological simulation—Xinjiang model and Shanbei model. Beijing: Water Resources and Electric Power Press, 1984. [赵人俊. 流域水文模拟—新安江模型和陕北模型. 北京: 水利电力出版社, 1984.]
- [35] Zhang Y, Liang ZM. Application research of BFS to flood forecasting. *Water Resources and Power*, 2009, (5): 44-47. [张宇, 梁忠民. BFS在洪水预报中的应用研究. 水电能源科学, 2009, (5): 44-47.]
- [36] Biondi D, Versace P, Sirangelo B. Uncertainty assessment through a precipitation dependent hydrologic uncertainty processor: An application to a small catchment in southern Italy. *Journal of Hydrology*, 2010, **386**(1/2/3/4): 38-54. DOI: 10.1016/j.jhydrol.2010.03.004.
- [37] Xing ZX, Rui XF, Fu Q *et al.* Nash model parameter uncertainty analysis by AM-MCMC based on BFS and probabilistic flood forecasting. *Chinese Geographical Science*, 2011, **21**(1): 74-83. DOI: 10.1007/s11769-010-0433-1.
- [38] Wang J, Liang ZM, Hu YM. Application and improvement of BFS in flood forecasting. *Journal of Hohai University: Natural Sciences*, 2012, (1): 52-58. [王军, 梁忠民, 胡义明. BFS在洪水预报中的应用与改进. 河海大学学报: 自然科学版, 2012, (1): 52-58.]
- [39] Zhang HG. Bayesian probability hydrological forecasting system and its application [Dissertation]. Wuhan: Wuhan University, 2005. [张洪刚. 贝叶斯概率水文预报系统及其应用研究[学位论文]. 武汉: 武汉大学, 2005.]
- [40] Jiang XL, Liang ZM, Wang CQ *et al.* Application of BFS-HUP model to flood probabilistic forecasting of Tongguan Station. *Yellow River*, 2015, **37**(7): 13-15. [蒋晓蕾, 梁忠民, 王春青等. BFS-HUP模型在潼关站洪水概率预报中的应用. 人民黄河, 2015, **37**(7): 13-15.]
- [41] Krzysztofowicz R. Transformation and normalization of variates with specified distributions. *Journal of Hydrology*, 1997, **197**(1/2/3/4): 286-292.