

分位数回归和 GAMLSS 模型在非一致性洪水频率分析中的比较

李婧¹, 闫磊¹, 屈春来¹, 刘章君², 鲁东阳¹

(1. 河北工程大学水利水电学院, 河北 邯郸 056000; 2. 江西省水利科学研究院, 江西 南昌 330029)

摘要:为了探索分位数回归模型在非一致性频率分析中的适用性,分别采用分位数回归模型和 GAMLSS 模型对渭河流域以及珠江流域 5 个站点的年最大洪水序列进行非一致性频率分析。模型概率覆盖率定性分析结果显示,分位数回归模型拟合效果优于 GAMLSS 模型;Filliben 相关系数定量分析结果显示,GAMLSS 模型拟合效果更好但优势不明显。综合分析可得,分位数回归模型的拟合优度整体优于 GAMLSS 模型。

关键词:洪水频率;非一致性频率分析;GAMLSS 模型;分位数回归;模型优度;Filliben 相关系数
中图分类号:P333.6 **文献标志码:**A **文章编号:**1006-7647(2020)05-0048-07

Comparison of quantile regression and GAMLSS model in non-stationary flood frequency analysis//LI Jing¹, YAN Lei¹, QU Chunlai¹, LIU Zhangjun², LU Dongyang¹(1. School of Water Conservancy and Hydroelectric Power, Hebei University of Engineering, Handan 056000, China; 2. Jiangxi Provincial Institute of Water Sciences, Nanchang 330029, China)

Abstract: In order to explore the applicability of the quantile regression model in non-stationary frequency analysis, the quantile regression model and the GAMLSS model were used to analyze the non-stationary frequency of the annual maximum flood series at five stations in the Weihe River Basin and the Pearl River Basin. Qualitative analysis of the probability coverage shows that the goodness of the fitting effect for the quantile regression model is better than that of the GAMLSS model. Quantitative analysis of the Filliben correlation coefficient shows that the GAMLSS model has a better fitting effect but the advantages are not obvious. Comprehensive analysis shows that the overall fitting goodness of the quantile regression model is better than the GAMLSS model.

Key words: flood frequency; non-stationary frequency analysis; GAMLSS model; quantile regression; model goodness; Filliben correlation coefficient

洪水频率分析对水利工程建设至关重要,一致性假设一直是洪水频率分析的基本假设条件,但受到气候变化以及人类活动的影响,降雨时空分配过程以及流域下垫面的产汇流条件发生了改变,导致洪水频率分析的一致性假设受到挑战^[1-4],传统频率计算方法构建模型的合理性受到质疑。因此,在变化环境下进行洪水序列的非一致性频率分析具有重要意义。非一致性频率分析成为变化环境下洪水分析计算的研究热点之一^[2,5]。当前最常用的非一致性频率分析方法是采用基于位置、尺度和形状参数的广义可加模型(generalized additive models for location scale and shape, GAMLSS)构建非一致性概率分布模型。Villarini 等^[6-7]应用 GAMLSS 模型分

别对罗马降水序列以及美国 Little Sugar Creek 流域的年最大洪峰序列进行了趋势分析,取得了较好的拟合效果。江聪等^[8]认为运用传统方法对水文序列均值的趋势性进行分析相对片面,通过 GAMLSS 模型引入方差和偏态系数等参数对宜昌站多年年流量序列进行了趋势分析。张冬冬等^[9]采用 GAMLSS 模型研究了大渡河流域年最大日降水极值序列的非一致性,发现 GAMLSS 模型可以模拟水文序列的方差、偏态系数等统计参数的非线性变化趋势。

然而,GAMLSS 模型构建过程较复杂,首先需要选定概率分布模型,然后假定所选概率分布模型统计参数与协变量的关系,模型统计参数多且结构较复杂。为解决这一问题,董洁等^[10-12]尝试引入非参

基金项目:国家自然科学基金(51909053);河北省自然科学基金(E2019402076);河北省高等学校科学技术研究项目(QN2019132)

作者简介:李婧(1996—),女,硕士研究生,主要从事水文分析计算。E-mail:386429384@qq.com

通信作者:闫磊(1990—),男,副教授,博士,主要从事水文分析计算。E-mail:yanl@whu.edu.cn

数方法进行非一致性水文频率分析。Koenker 等^[13]首先提出了分位数回归方法研究样本点在不同的分位数下随时间变化情况,随着该方法的进一步研究开发,目前广泛应用于水文分析中。Mazvimavi^[11]利用分位数回归方法对津巴布韦的降雨序列变化进行分析。Barbosa^[12]等运用分位数回归对波罗的海海平面变化进行分析,发现极大值处的斜率趋势性更大。Wang 等^[14]利用分位数回归方法对美国南部月降雨变化情况进行分析。冯平等^[15]利用分位数回归方法分析滦河流域年降雨、径流序列的变化特性,发现流域年降雨序列总体保持一致性,而年径流序列受人类活动等影响表现为减小趋势,且高分位数的减小速率大于低分位数。本文选取位于我国不同气候区的渭河流域和珠江流域的 5 个代表站点的年最大洪水序列作为研究对象,综合定性分析和定量分析,对比分位数回归模型与 GAMLSS 模型在进行非一致性洪水频率分析时的表现。

1 流域概况和数据资料

渭河流域发源于甘肃省渭源县鸟鼠山,流经甘肃、陕西两省,至渭南市潼关县汇入黄河,河长 818 km,流域面积 134 766 km²,位于北纬 33°40′~37°26′,东经 103°57′~110°27′之间^[16-17]。渭河中下游径流的年际变化表现为南部小、北部大。渭河干流秋季径流量最大,约占年径流量的 38%~40%,夏季占 32.8%~34.2%,春季占 17.7%~19.1%,冬季为 8.3%~9.9%。渭河中下游降雨集中于 7—9 月,而且多大暴雨,洪涝灾害较多。

珠江发源于云南省曲靖市马雄山,流经云南、贵州、广西、广东、湖南、江西 6 省及越南北部,在下游从 8 个入海口注入南海,全长 2 320 km,流域面积 45 369 km²。珠江流域由西江、北江、东江和珠江三角洲诸河等 4 个水系组成^[18-19],流域多年平均降水量 1 200~2 200 mm,年径流量超 3 300 亿 m³,汛期 4—9 月径流量约占年径流量的 80%,夏季(6—8 月)则占年径流量的 50% 以上。

数据资料选取渭河的咸阳站、华县站、张家山站以及珠江流域的高道站和大湟江口站共 5 个重要站点的多年年最大洪水序列。5 个站点的基本概况详见表 1 及图 1。

2 研究方法

2.1 Mann-Kendall 趋势检验方法

采用 Mann-Kendall 非参数趋势检验法检测降

表 1 各站点基本概况					
流域	站名	控制流域面积/km ²	东经/(°)	北纬/(°)	资料年限
珠江流域	高道	7 007	113.17	24.16	1954—2014 年
	大湟江口	288 544	110.20	23.58	1954—2009 年
渭河流域	咸阳	46 827	108.70	34.32	1954—2011 年
	华县	106 498	109.76	34.58	1951—2012 年
	张家山	43 216	108.59	34.64	1954—2013 年

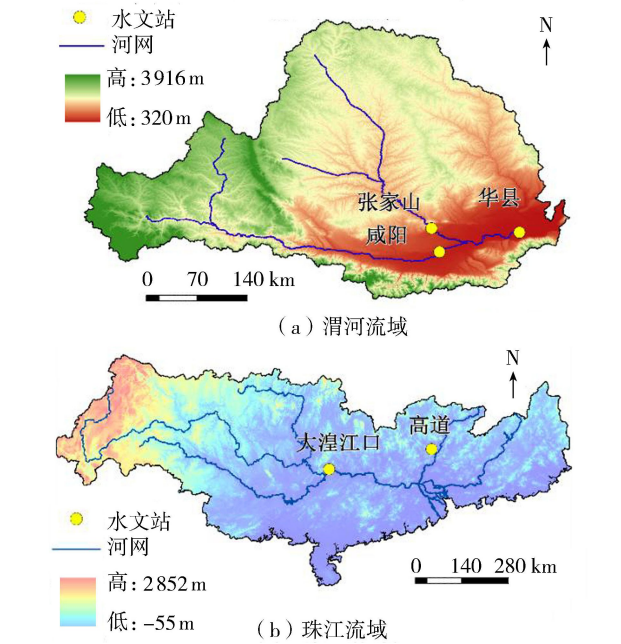


图 1 流域水文站点示意图

水和径流序列的长期变化趋势^[20]。Mann-Kendall 的检验方法^[15,21]如下:

设 $Y_i (i=1,2,\cdots,n)$ 为假设检验的随机变量, n 为样本的观测长度,标准化检验统计量 Z_c 为

$$Z_c = \frac{\varepsilon}{\psi_\varepsilon^2} \tag{1}$$

其中 $\varepsilon = \frac{4P'}{n(n-1)} - 1$ $\psi_\varepsilon^2 = \frac{2(2n+5)}{9n(n-1)}$ 式中: P' 为序列中所有对偶观测值 ($Y_i, Y_j, i < j$) 中 $Y_i < Y_j$ 出现的次数, 给定 α 显著性水平 (本文无特殊要求时设定 $\alpha=0.05$, 则统计量临界值 $Z_{\alpha/2}=1.96$), 若 $|Z_c| < Z_{\alpha/2}$ 则接受零假设, 认为样本序列不存在显著趋势; 当 $Z_c > 0$ 时则表示该样本序列呈现上升趋势, 反之则表现为下降趋势。

2.2 分位数回归法

分位数回归是依据因变量 Y 的条件分位数对自变量 X 进行回归, 得到所有分位数下的回归模型; 分位数回归对于模型中的随机干扰项不需要做概率分布假定, 且估计量基本不受异常值的影响而更加稳定, 可拟合出一簇曲线^[15]。

假设随机变量 Y 的分布函数为 $F(y) = P(Y \leq y)$, 则 Y 的第 τ 分位数为

$$Q(\tau) = \inf\{y: F(y) \geq \tau\} \quad (0 < \tau < 1) \tag{2}$$

根据不同的分位数 τ , 可得到不同的分位数函数 $Q_Y(\tau) = \omega_i \beta(\tau)$ ($Q_Y(\tau)$ 为第 τ 分位数时 Y 的分位数函数, $\beta(\tau)$ 为参数值, $i=1, 2, \dots, n$)。分位数回归的目标函数是一个加权的绝对残差和, 目的是使样本值与拟合值之间的距离最短。给定观测数据 $(\omega_1, y_1), (\omega_2, y_2), \dots, (\omega_n, y_n)$, 由此求解 τ 分位数的回归估计:

$$\min_{\beta \in R} \sum_{i=1}^n \rho_{\tau}(y_i - \omega_i \beta) \quad (3)$$

其中 $\rho_{\tau}(\cdot)$ 是一个非对称的损失函数, 定义为

$$\rho_{\tau}(y_i - \omega_i \beta) = \begin{cases} (y_i - \omega_i \beta) \tau & y_i - \omega_i \beta \geq 0 \\ (y_i - \omega_i \beta) (1 - \tau) & y_i - \omega_i \beta < 0 \end{cases} \quad (4)$$

可通过下式得到参数估计值, 根据一个 τ 值确定一组 $\beta(\tau)$ 值:

$$\beta(\tau) = \arg \min_{\beta \in R} \sum_{i=1}^n \rho_{\tau}(y_i - \omega_i \beta) \quad (5)$$

2.3 GAMLSS 模型

GAMLSS 模型是基于位置、尺度、形状的广义可加模型, 可以描述统计参数与协变量之间的线性和非线性关系, 模拟随机变量分布函数类型广泛, 适用于高偏度、高峰度、超峰度、平顶峰度等的随机变量序列^[8, 22]。

2.3.1 模型的定义

在 GAMLSS 模型中, 假设某一时刻 t ($t=1, 2, \dots, n$) 的相对独立随机变量观测值 y_t 的分布函数可以表示为 $F(y_t | \theta_t)$, 其中 $\theta_t = (\theta_{t1}, \theta_{t2}, \dots, \theta_{tf})$ 为时刻 t 对应的分布统计参数向量, f 为分布参数的个数, n 为观测值的个数^[8]。记 $g_k(\theta_k)$ 为 θ_k 与相应的协变量 Y_k 的单调函数关系, 一般表示为

$$g_k(\theta_k) = \eta_k = Y_k \beta_k + \sum_{j=1}^{jk} Z_{jk} \gamma_{jk} \quad (6)$$

式中: η_k 为长度 n 的向量; $\beta_k = (\beta_{1k}, \beta_{2k}, \dots, \beta_{lk})^T$, 为长度为 l_k 的回归参数向量; Y_k 为 $n \times l_k$ 的协变量矩阵; Z_{jk} 为已知的 $n \times q_{jk}$ 矩阵; γ_{jk} 为 q_{jk} 维的正态分布随机变量向量; $Z_{jk} \gamma_{jk}$ 为第 j 项随机效应^[8]。

若不考虑随机效应的影响, 则 $g_k(\theta_k) = \eta_k = Y_k \beta_k$, 针对位置参数 μ 、尺度参数 σ 和形状参数 v , 以时间 t 为协变量的不考虑随机效应的全参数模型为

$$\begin{cases} g_1(\mu_t) = \beta_{11} + \beta_{21}t + \dots + \beta_{l_11}t^{l_1-1} \\ g_2(\sigma_t) = \beta_{12} + \beta_{22}t + \dots + \beta_{l_22}t^{l_2-1} \\ g_3(v_t) = \beta_{13} + \beta_{23}t + \dots + \beta_{l_33}t^{l_3-1} \end{cases} \quad (7)$$

式中: μ_t, σ_t, v_t 分别为时变的位置参数、尺度参数、形状参数, 可体现非平稳序列统计参数随时间的变化特征, 因为形状参数对模型影响较大, 所以一般不做考虑, 可通过 RS 法^[23] 对 β 进行参数估计。

2.3.2 模型的评价准则

选取 GAIC 准则 (generalized akaike information criterion, 简称 GAIC) 作为 GAMLSS 模型拟合评价指标, 其值 V_{GAIC} 的计算公式为

$$V_{GAIC} = V_{GD} + \#d_f \quad (8)$$

式中: $V_{GD} = -2\ln L(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)$ 为 GAMLSS 模型的全局拟合偏差, $\ln L(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)$ 为回归参数估计值对应的对数似然函数; d_f 为模型的整体自由度; $\#$ 为惩罚因子, 取 $\# = 2$ 则为 AIC 值 (akaike information criterion), AIC 值越小, 模型表现越优。

2.4 拟合优度检验

2.4.1 正态 QQ 图

通过正态 QQ 图来分析最优模型的残差分布状况。正态 QQ 图是在平面坐标系中以经验残差为纵坐标、理论残差为横坐标绘制而成的, 数据点与 1:1 的线偏差越小, 说明模型的拟合优度越好。

2.4.2 模型概率覆盖率检验

通过模型概率覆盖率偏差数值的大小定性分析分位数回归模型与 GAMLSS 模型的拟合优度。模型概率覆盖率是各条分位数曲线覆盖范围内的实测样本点数量与总样本点数量的比值。计算模型概率覆盖率与对应分位数的差值, 差值越小说明模型拟合优度越好。以 50% 分位数曲线为例, 理论上讲, 如果 50% 分位数曲线覆盖范围内的实测样本点数量与总样本点数量的比值越接近 50%, 则模型概率覆盖率与 50% 差值越小, 模型拟合优度越好。

2.4.3 Filliben 检验

可通过 Filliben 相关系数来确定洪水序列最优分布形式, 通过 Filliben 相关系数的大小来判定序列最优拟合分布, Filliben 值越大则分布拟合优度越好^[24]。

假设标准化残差 $r_1, r_2, \dots, r_n$ 的实际分布服从正态分布, 升序统计量为 $r_{(1)}, r_{(2)}, \dots, r_{(n)}$; 理论残差计算公式为 $M_i = \varphi^{-1}[(i-0.375)/(n+0.25)]$, 升序统计量与理论残差存在线性关系, 用 $\bar{r}$ 表示 r_i 的均值, $\bar{M}$ 表示 M_i 的均值, Filliben 相关系数大于 0.975 时通过显著性水平为 5% 的检验^[25-26], 定义 Filliben 相关系数为

$$r_{\text{Filliben}} = \frac{\sum_{i=1}^n (r_i - \bar{r})(M_i - \bar{M})}{\sqrt{\sum_{i=1}^n (r_i - \bar{r})^2 \sum_{i=1}^n (M_i - \bar{M})^2}} \quad (9)$$

3 结果和讨论

3.1 Mann-Kendall 趋势分析

使用 R 语言中 trend 程序包对各站点历史资料

进行 Mann-Kendall 检验,显著性水平设定为 5%,得出华县站、高道站、大湟江口站、咸阳站、张家山站 5 个站点 P 、 Z_c 和 S 值见表 2。

表 2 各站点趋势性分析结果

站点	P	Z_c	S
华县站	2.117×10^{-5}	-4.252 2	-701
高道站	3.596×10^{-2}	2.097 4	338
大湟江口站	5.727×10^{-2}	1.901 3	270
咸阳站	3.236×10^{-4}	-3.595 6	-537
张家山站	1.225×10^{-2}	-2.504 8	-384

华县站、高道站、咸阳站、张家山站 4 个站点通过显著性水平 5% 的检定,而大湟江口站通过显著性水平为 10% 的检定,同时根据 Z_c 值正负关系可得出高道和大湟江口 2 个站点呈上升趋势,其他 3 个站点则呈下降趋势。故华县站、咸阳站、张家山站 3 个站点显著下降,高道站显著上升,而大湟江口站呈现不显著上升趋势。

3.2 GAMLSS 最优模型拟合效果分析

3.2.1 GAMLSS 最优模型判定

本算例为选定最优概率分布类型,选取我国不同气候区的渭河、珠江流域的 5 个站点最大洪水序列为研究对象,时间 t 为协变量,AIC 为评价准则,分别对比了 P-Ⅲ 分布、Gamma 分布(简称 GA 分布)、Lognormal 分布(简称 LN 分布)和 GEV 分布 4 种分布形式,通过正态 QQ 图判别拟合优度。每个站点每种分布类型根据位置参数和尺度参数分别随时间变化可形成 4 种不同的模型,总计 16 种,相应 AIC 取值详见表 3。在表 3 模型代码中,以“L”表示

位置参数,“S”表示尺度参数,“0”表示参数不变,“t”表示参数随时间协变量变化,此外粗体 AIC 值表示该站点的最佳模型。

表 3 各站点模型 AIC 值对比

模型代码	华县	高道	大湟江口	咸阳	张家山
P-Ⅲ_I0_S0	1070.78	1064.88	1159.02	970.66	914.98
P-Ⅲ_Lt_S0	1055.08	1062.09	1155.29	963.63	909.36
P-Ⅲ_I0_St	1072.28	1064.59	1160.49	969.59	916.73
P-Ⅲ_Lt_St	1055.84	1062.43	1157.23	958.05	910.60
GA_I0_S0	1067.70	1061.64	1159.68	969.63	906.85
GA_Lt_S0	1053.73	1058.00	1155.26	963.40	901.13
GA_I0_St	1066.76	1060.30	1161.63	966.93	908.84
GA_Lt_St	1051.84	1057.64	1158.02	955.81	903.07
LN_I0_S0	1069.93	1063.52	1161.631	974.87	899.82
LN_Lt_S0	1052.91	1062.08	1158.438	963.46	894.26
LN_I0_St	1070.71	1062.15	1163.60	973.41	901.82
LN_Lt_St	1052.33	1061.34	1160.01	956.03	896.15
GEV_I0_S0	1073.06	1063.08	1159.87	974.68	901.51
GEV_Lt_S0	1057.44	1065.08	1155.40	961.69	899.33
GEV_I0_St	1073.42	1059.02	1158.60	976.82	903.92
GEV_Lt_St	1058.11	1061.02	1160.37	978.52	905.43

由此可得出华县站、高道站、咸阳站的最优模型为位置参数和尺度参数均随时间变化的 Gamma 分布,AIC 值分别为 1 051.84、1 057.64、955.81;大湟江口站最优模型则为位置参数随时间变化而尺度参数保持不变的 Gamma 分布,AIC 值为 1 155.26;张家山站最优模型则为位置参数随时间变化而尺度参数保持不变的 Lognormal 分布,AIC 值为 894.26。

图 2 给出了最优模型的正态 QQ 图的拟合优度检验结果。结果显示,根据 AIC 值选择的最优模型

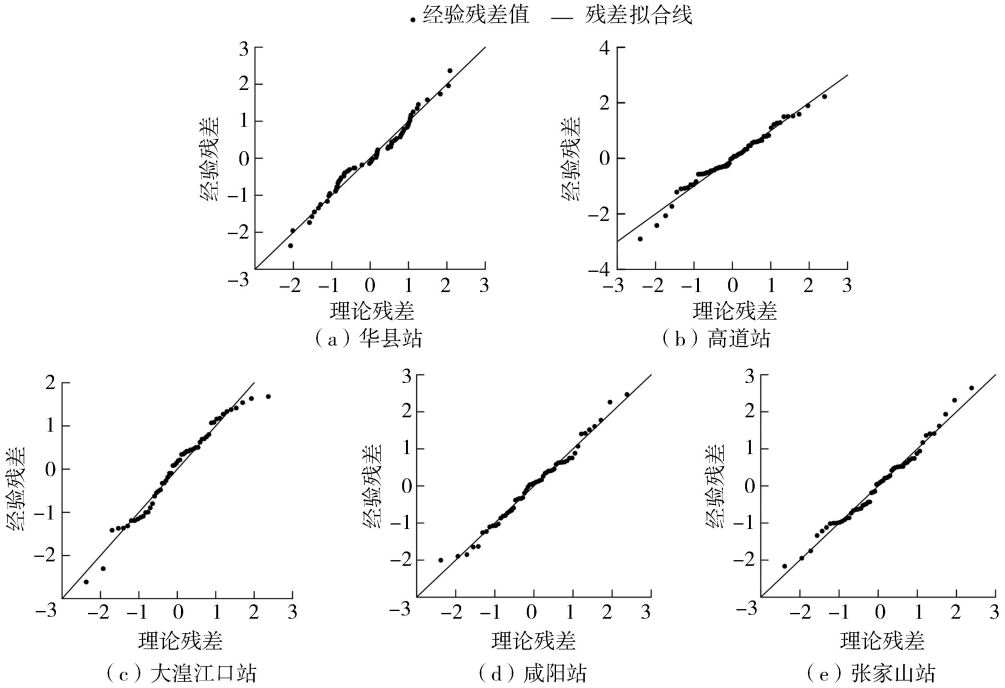


图 2 各站点 QQ 图

点均匀分布在 1 : 1 直线附近,经验残差和理论残差基本一致,表明模型具有良好的拟合效果。

3.2.2 GAMLSS 最优模型的定性分析

对于该算例的实际样本点来说,通过对比各站点 GAMLSS 模型的分位数曲线图,定性分析模型概率覆盖率,可得出华县站 GAMLSS 模型的概率覆盖率偏差最小为 0.16%,最大为 7.26%,高道站、大湟江口站、咸阳站、张家山站 GAMLSS 模型的偏差范围分别为 1.56% ~ 6.97%、1.43% ~ 5.36%、0.17% ~ 5.17% 和 0.08% ~ 2.97%,具体见图 3 与表 4。

3.2.3 GAMLSS 最优模型的定量分析

通过对比 Filliben 相关系数进行定量对比分析,基于式(9)进行相应计算可得 GAMLSS 模型的 Filliben 相关系数:华县站为 0.989 96;高道站为 0.987 843;大湟江口站为 0.986 808;咸阳站为 0.994 089;张家山站为 0.991 817。

3.3 分位数回归模型拟合效果分析

3.3.1 分位数回归模型的定性分析

通过对比各站点分位数回归模型的分位数曲线图,可定性分析模型概率覆盖率,可得出华县站分位数回归模型下模型概率覆盖率偏差最小为 0.81%,

最大为 4.84%;高道站、大湟江口站、咸阳站、张家山站分位数回归模型下偏差范围分别为 0.08% ~ 5.74%、0 ~ 0.36%、0 ~ 2.59% 和 0.08% ~ 1.61%,具体见图 4 与表 5。

3.3.2 分位数回归模型的定量分析

可通过对比 Filliben 相关系数进行定量对比分析,基于式(9)进行相应计算可得分位数回归模型的 Filliben 相关系数:华县站,0.983 076;高道站,0.983 149;大湟江口站,0.983 534;咸阳站,0.983 022;张家山站,0.984 437。

3.4 分位数回归模型与 GAMLSS 模型拟合效果比较

根据模型概率覆盖率进行定性分析,可发现各站点分位数回归模型均优于 GAMLSS 模型,华县站显示优度为 0.3% ~ 6.4%,高道站为 1.6% ~ 3.3%,大湟江口站为 1.1% ~ 5.4%,咸阳站为 1.4% ~ 5.2%,张家山站为 0 ~ 1.8%;根据 Filliben 相关系数进行定量分析,两种模型 Filliben 相关系数较高,均通过显著性水平 5% 时,Filliben 相关系数临界值为 0.975 的检验^[25-28],两种模型的拟合效果 GAMLSS 模型优于分位数回归模型的范围

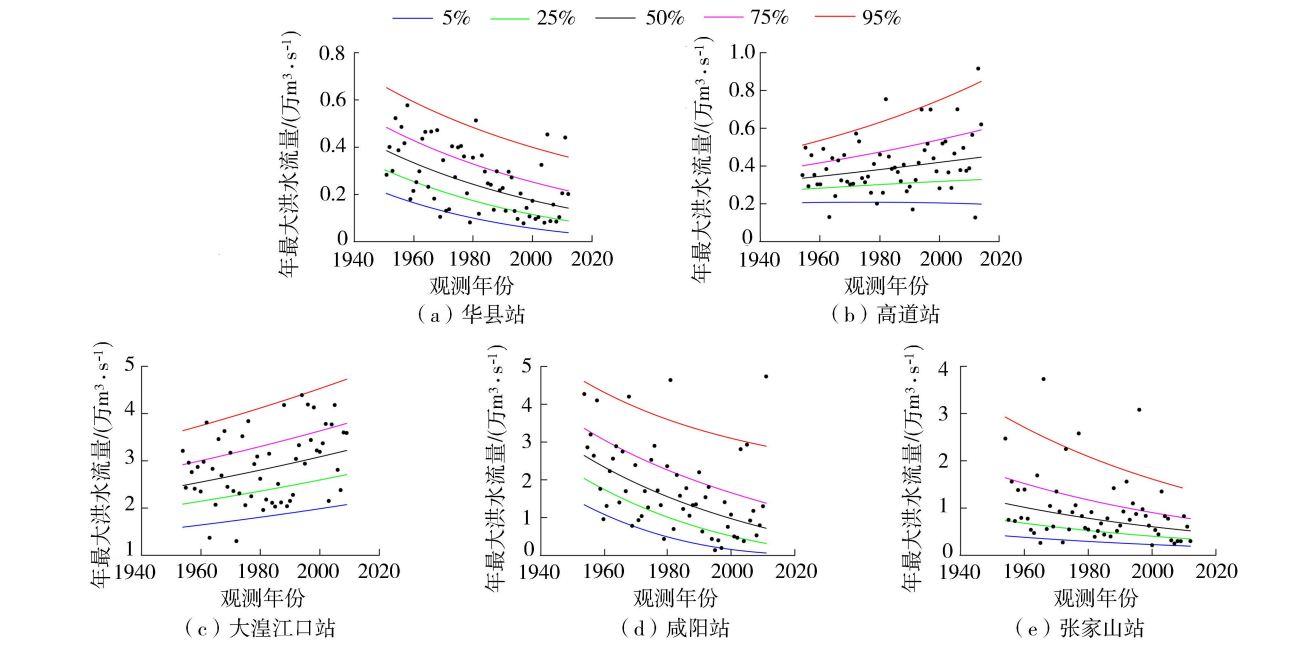


图 3 各站点基于 GAMLSS 模型的分位数曲线

表 4 各站点不同分位数 GAMLSS 模型拟合优度对比

站点	模型概率覆盖率/%					模型概率覆盖率偏差/%				
	5%	25%	50%	75%	95%	5%	25%	50%	75%	95%
华县站	3.23	32.26	45.16	69.35	95.16	-1.77	7.26	-4.84	-5.65	0.16
高道站	6.56	18.03	47.54	78.69	96.72	1.56	-6.97	-2.46	3.69	1.72
大湟江口站	3.57	26.79	44.64	73.21	98.21	-1.43	1.79	-5.36	-1.79	3.21
咸阳站	6.90	27.59	44.83	79.31	94.83	1.90	2.59	-5.17	4.31	-0.17
张家山站	5.08	25.42	47.46	77.97	94.92	0.08	0.42	-2.54	2.97	-0.08

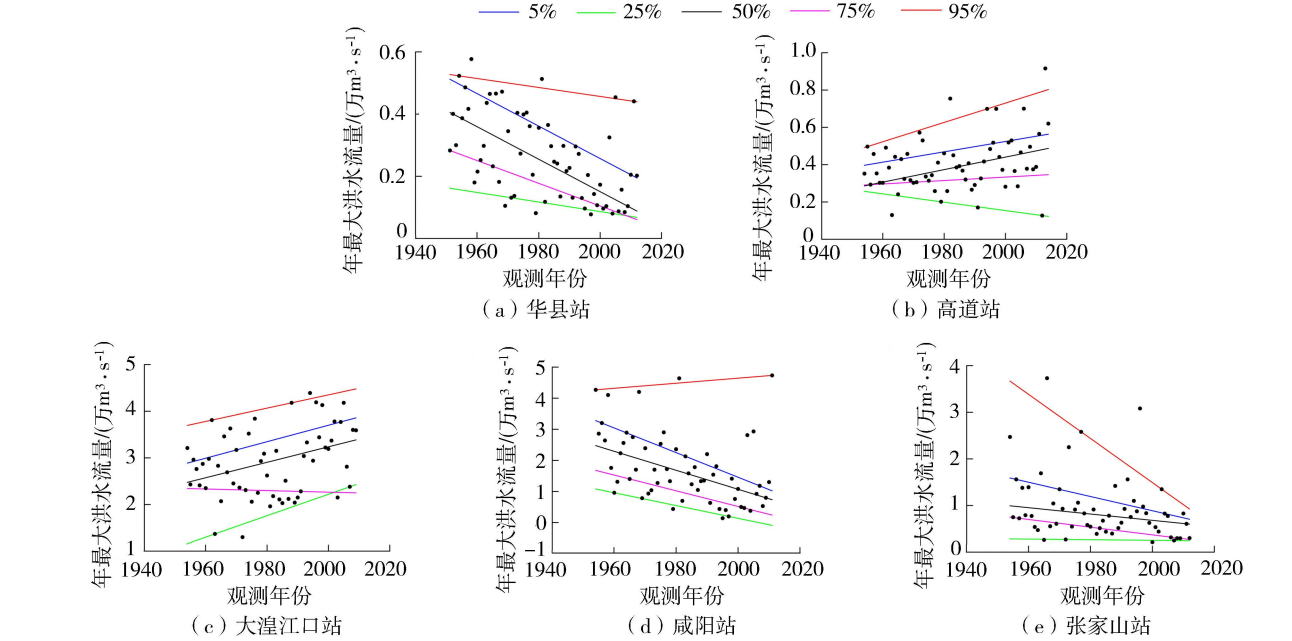


图4 各水文站基于分位数回归模型的分位数线

表5 各站点不同分位数回归模型拟合优度对比

站点	模型概率覆盖率/%					模型概率覆盖率偏差/%				
	5%	25%	50%	75%	95%	5%	25%	50%	75%	95%
华县站	6.45	24.19	45.16	75.81	93.55	1.45	-0.81	-4.84	0.81	-1.45
高道站	3.28	21.31	44.26	73.77	95.08	-1.72	-3.69	-5.74	-1.23	0.08
大湟江口站	5.36	25.00	50.00	75.00	96.43	0.36	0.00	0.00	0.00	1.43
咸阳站	5.17	24.14	50.00	72.41	96.55	0.17	-0.86	0.00	-2.59	1.55
张家山站	3.39	23.73	50.85	76.27	94.92	-1.61	-1.27	0.85	1.27	-0.08

基本位于0.3%~0.9%之间。基于Filliben相关系数的定量分析结果显示,GAMLSS模型拟合效果更好但优势不明显;基于模型概率覆盖率的定性分析结果表明,分位数回归模型的拟合效果更好。因此,综合上述定性分析和定量分析结果,总体来说,分位数回归模型的拟合效果更优。

4 结 论

a. 通过Mann-Kendall趋势检验可以得出5%显著性水平下华县站、咸阳站、张家山站3个站点显著下降,高道站显著上升,而大湟江口站呈不显著上升趋势。

b. 各站点在GAMLSS模型中拟合效果最好的分布,以Gamma分布为主、Lognormal分布为辅。华县站、高道站、咸阳站表现为Gamma分布下的位置参数和尺度参数均随时间变化的模型效果最佳,大湟江口站则为位置参数随时间变化而尺度参数保持不变的Gamma分布效果最佳,张家山站表现为Lognormal分布下的位置参数随时间变化而尺度参数保持不变的模型最佳。

c. 通过定性分析对比各站点两种模型概率覆盖率,发现分位数回归模型整体拟合效果优于

GAMLSS模型。定量分析Filliben相关系数,分位数回归模型与GAMLSS模型拟合效果基本一致。故综合定性分析和定量分析可得出分位数回归模型的拟合效果优于GAMLSS模型。

d. GAMLSS模型可以考虑多种因素的影响以及更多的分布类型,进行更加复杂的模拟运算,但是统计参数较多,且统计参数与协变量的关系可能不只是线性,实际变化情况更加复杂,使得模型构建过程较为烦琐,而基于非参数方法的分位数回归模型构建更加简单且易于操作,模型拟合更加灵活,无须确定各项参数以及参数与协变量之间的复杂关系。

参考文献:

[1] 吴孝情,陈晓宏,唐亦汉,等. 珠江流域非平稳性降雨极值时空变化特征及其成因[J]. 水利学报, 2015, 46(9): 1055-1063. (WU Xiaoqing, CHEN Xiaohong, TANG Yihan, et al. Spatiotemporal variations and the causes of non-stationary extreme precipitation in the Pearl River basin[J]. Journal of Hydraulic Engineering, 2015, 46(9):1055-1063. (in Chinese))

[2] 梁忠民,胡义明,王军. 非一致性水文频率分析的研究进展[J]. 水科学进展, 2011, 22(6): 864-871. (LIANG Zhongmin, HU Yiming, WANG Jun. Advances

- in hydrological frequency analysis of non-stationary time series[J]. *Advances in Water Science*, 2011, 22(6): 864-871. (in Chinese))
- [3] 胡义明, 梁忠民, 姚轶, 等. 变化环境下水文设计值计算方法研究综述[J]. *水利水电科技进展*, 2018, 38(4): 89-94. (HU Yiming, LIANG Zhongmin, YAO Yi, et al. Review on computational methods of hydrological design values in changing environment[J]. *Advances in Science and Technology of Water Resources*, 2018, 38(4):89-94. (in Chinese))
- [4] 熊立华, 闫磊, 李凌琪, 等. 变化环境对城市暴雨及排水系统影响研究进展[J]. *水科学进展*, 2017, 28(6): 930-942. (XIONG Lihua, YAN Lei, LI Lingqi, et al. Advances in analysis of impacts of changing environments on extreme urban rainfall and drainage infrastructure[J]. *Advances in Water Science*, 2017, 28(6):930-942. (in Chinese))
- [5] XIONG L H, JIANG C, DU T, et al. Review on nonstationary hydrological frequency analysis under changing environments[J]. *Journal of Water Resources Research*, 2015, 4(4):310-319.
- [6] VILLARINI G, SMITH A J, NAPOLITANO F. Nonstationary modeling of a long record of rainfall and temperature over Rome [J]. *Advances in Water Resources*, 2010, 33(10):1256-1267.
- [7] VILLARINI G, SMITH A J, SERINALDI F, et al. Flood frequency analysis for nonstationary annual peak records in an urban drainage basin [J]. *Advances in Water Resources*, 2009, 32(8):1255-1266.
- [8] 江聪, 熊立华. 基于GAMLSS模型的宜昌站年径流序列趋势分析[J]. *地理学报*, 2012, 67(11): 1505-1514. (JIANG Cong, XIONG Lihua. Trend analysis for the annual discharge series of the Yangtze River at the Yichang Hydrological Station based on GAMLSS[J]. *Acta Geographica Sinica*, 2012, 67(11): 1505-1514. (in Chinese))
- [9] 张冬冬, 鲁帆, 周翔南, 等. 基于GAMLSS模型的大渡河流域极值降水非一致性分析[J]. *水利水电技术*, 2016, 47(5): 12-15. (ZHANG Dongdong, LU Fan, ZHOU Xiangnan, et al. GAMLSS model-based analysis on nonstationarity of extreme precipitation in Daduhe River Basin[J]. *Water Resources and Hydropower Engineering*, 2016, 47(5):12-15. (in Chinese))
- [10] 董洁, 谢悦波. 洪水频率分析的非参数变换回归方法[J]. *水利水电科技进展*, 2004, 24(3): 19-20. (DONG Jie, XIE Yuebo. Nonparametric transformed kernel estimation of flood frequencies[J]. *Advances in Science and Technology of Water Resources*, 2004, 24(3):19-20. (in Chinese))
- [11] MAZVIMAVI D. Investigating changes over time of annual rainfall in Zimbabwe [J]. *Hydrology and Earth System Sciences*, 2010, 14(12):2671-2679.
- [12] BARBOSA M S. Quantile trends in Baltic sea level[J]. *Geophysical Research Letters*, 2008, 35(22):L22704.
- [13] KOENKER R, BASSETT G. Regression quantiles [J]. *The Econometric Society*, 1978, 46(1):33-50.
- [14] WANG H, KILLICK R, FU X. Distributional change of monthly precipitation due to climate change: comprehensive examination of dataset in southeastern United States [J]. *Hydrological Processes*, 2014, 28(20):5212-5219.
- [15] 冯平, 商颂, 李新. 基于分位数回归的滦河流域降水及径流变化特性[J]. *水力发电学报*, 2016, 35(2): 28-36. (FENG Ping, SHANG Song, LI Xin. Temporal variation characteristics of annual precipitation and runoff in Luan River Basin based on quantile regression[J]. *Journal of Hydroelectric Engineering*, 2016, 35(2):28-36. (in Chinese))
- [16] 左德鹏, 徐宗学, 李景玉, 等. 气候变化情景下渭河流域潜在蒸散量时空变化特征[J]. *水科学进展*, 2011, 22(4): 455-461. (ZUO Depeng, XU Zongxue, LI Jingyu, et al. Spatiotemporal characteristics of potential evapotranspiration in the Weihe River Basin under future climate change[J]. *Advances in Water Science*, 2011, 22(4):455-461. (in Chinese))
- [17] 栗晓玲, 梁箴. 关中地区气象水文综合干旱指数及干旱时空特征[J]. *水资源保护*, 2019, 35(4): 17-23. (SU Xiaoling, LIANG Zheng. Meteorology-hydrology drought index and spatial and temporal characteristics of drought in Guanzhong area[J]. *Water Resources Protection*, 2019, 35(4):17-23. (in Chinese))
- [18] 朱永楠, 林朝晖, 郝振纯. 珠江流域大尺度陆面水文耦合模式的构建及应用[J]. *水文*, 2015, 35(1): 14-19. (ZHU Yongnan, LIN Zhaohui, HAO Zhenchun. Development of large scale coupled land surface and hydrologic model system and its application in Pearl River Basin[J]. *Journal of China Hydrology*, 2015, 35(1):14-19. (in Chinese))
- [19] 陈晓丽, 雷勇, 黄国如. 北江飞来峡库区流域非点源污染现状评价[J]. *水资源保护*, 2019, 35(2): 44-48. (CHEN Xiaoli, LEI Yong, HUANG Guoru. Evaluation on non-point source pollution in Feilaixia Reservoir Area of Beijiang River[J]. *Water Resources Protection*, 2019, 35(2):44-48. (in Chinese))
- [20] 王东升, 袁树堂, 杨祺. 金沙江流域云南片水文极小值演变及生态基流保障分析[J]. *水资源保护*, 2019, 35(4):35-41. (WANG Dongsheng, YUAN Shutang, YANG Qi. Analysis of hydrological minimum evolution and ecological base flow guarantee in Yunnan area of Jinshajiang River Basin, China [J]. *Water Resources Protection*, 2019, 35(4):35-41. (in Chinese))

(下转第 75 页)