

基于GAMLSS模型的径流频率分析

郭路^{1,2}, 姚聪聪^{3*}, 喻晓琳^{1,2}, 陈凯^{1,2}, 张洪波^{3,4}

¹中国石油长庆油田公司勘探开发研究院, 陕西 西安

²低渗透油气田勘探开发国家工程实验室, 陕西 西安

³长安大学水利与环境学院, 陕西 西安

⁴长安大学旱区地下水文与生态效应教育部重点实验室, 陕西 西安

收稿日期: 2022年5月6日; 录用日期: 2022年6月17日; 发布日期: 2022年6月29日

摘要

河川径流数据非平稳性的日益凸显, 使流域或区域的水文设计与水资源评价面临极大挑战, 如何开展有效的水文分析, 指导工程设计与流域或区域水资源管理已成为水资源可持续开发与应急灾害管理领域的迫切需求。本文以GAMLSS模型为基础, 对马莲河流域不同站点的水文观测资料进行分析, 得到了全情景下的水文频率计算结果, 并与传统方法进行对比, 指出了应对不同工程或规划需要的水文数据筛选方法和处理建议。

关键词

河川径流, 水文频率, 非平稳性, GAMLSS, 马莲河流域

Frequency Analysis of River Runoff Based on GAMLSS Model

Lu Guo^{1,2}, Congcong Yao^{3*}, Xiaolin Yu^{1,2}, Kai Chen^{1,2}, Hongbo Zhang^{3,4}

¹Exploration and Development Research Institute of Petro China Changqing Oilfield Company, Xi'an Shaanxi

²National Engineering Laboratory for Exploration & Development of Low Permeability Oil/Gas Fields, Xi'an Shaanxi

³School of Water and Environment, Chang'an University, Xi'an Shaanxi

⁴Key Laboratory of Subsurface Hydrology and Ecological Effect in Arid Region of Ministry of Education, Chang'an University, Xi'an Shaanxi

Received: May 6th, 2022; accepted: Jun. 17th, 2022; published: Jun. 29th, 2022

Abstract

The increase of non-stationarity river runoff data poses great challenges to hydrological design and wa-

作者简介: 郭路(1978.08-), 男, 河北宣化人, 副高级, 硕士研究生, 从事水文学及水资源方面的研究, Email: guolu_cq@petrochina.com.cn
*通讯作者 Email: ycc19972563@163.com

ter resources assessment in a basin or a region. How to carry out effective hydrological analysis to guide basin or regional engineering design and water resource management has become an urgent demand for water resources sustainable development and emergency disaster management. The changes of hydrological observation data of various stations in the Malian River basin were analyzed systematically based on the GAMLSS model, and hydrological frequency under different scenarios was calculated. Compared with the traditional hydrological design methods, the issues and suggestions about hydrological data screening and processing to meet the needs of water project design and water resources planning are suggested.

Keywords

River Runoff, Hydrologic Frequency, Non-Stationarity, GAMLSS, Malian River Basin

Copyright © 2022 by author(s) and Wuhan University.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



1. 引言

气候变化和人类活动的耦合扰动下,全球很多地区的河川径流数据都表现出了显著的非平稳性(或称非一致性),已经引起了国际水文、气象、环境等领域专家学者的广泛关注[1]。随着近年来我国社会经济发展进程的持续推进,人类社会对水资源需求的日益增加,加之脆弱的生态环境和不合理的开发方式,我国正面临着严重的水资源短缺危机及生态环境安全问题,特别是非平稳水文变化的出现,使得流域或区域的水文情势变化更趋复杂,水安全问题凸显[2],亟待发展非平稳条件的水文分析方法,以提供变化环境下科学的水资源开发与生态保护。

所谓的河川径流非一致性,即水文序列的概率分布或期望、方差等统计参数时变,导致很难用一种模型实现对观测数据的拟合[3]。为此,国内外学者对此进行了系统的研究,并提出了许多不同层面的研究成果,如极端降水的非一致性研究[4]、降水与径流的非一致性研究、全球变暖背景下水文非一致性研究[5]、洪水非一致性研究[6]等。而在方法上,也有很多具有说服力的成果,如 Kharin 和 Zwiers [4]使 GEV 分布的位置和尺度参数随时间发生变化,对非一致性条件下未来全球气温和降水极值的变化进行了研究。刘丙军等[7]运用时变矩法对西北江三角洲地区非一致性洪水频率进行了分析,当水文序列分布的统计参数随时间变化时,能较好反映变化环境下水文要素的非一致性变化特征。叶长青等[8]采用时变矩法对年最大日流量序列进行了非一致性分析,得到了较优的拟合效果。熊立华和郭生练[9]应用 Gumbel-Hougaard Copula 函数建立了长江流域水文站点的洪峰和洪量的联合分布,取得了更好的洪水频率分析效果。

基于位置、尺度和形状参数的广义加法模型(GAMLSS)是由 Rigby 和 Stasinopoulos [10]提出,该模型具有广泛的随机变量分布类型,包括了高偏度和高峰度的分布,并且可以对序列任何分布参数随协变量的线性变化或非线性变化进行描述。Villarini 等[6] [11]于 2015 年最先将 GAMLSS 模型用于分析水文序列的非一致性,之后 GAMLSS 模型被广泛应用于水文领域非一致性研究中序列的趋势性变化。李婧等[12]采用 GAMLSS 模型对渭河流域和珠江流域 5 个站点的年最大洪水序列进行非一致性频率分析,有较好的拟合效果。鲁帆等[13]采用 GAMLSS 模型对黄河干流 6 个水文站的年径流系列进行拟合,得到各站年天然径流对应的设计径流量有随时间减少的趋势。高洁[14]以美国 Little Sugar Creek 为研究案例,利用 GAMLSS 模型对其 83 年的洪峰流量进行了分析,得到该流域 2006 年 10 年一遇的洪峰流量相当于 1999 年 20 年重现期的洪峰,并和 1989 年 100 年一遇洪峰量级接近。综上,大量的研究结果已表明,GAMLSS 模型可以灵活地描述随机变量分布的任何参数与解释变量

之间的线性或非线性关系，在非平稳水文频率计算方面具有有效性，且已被众多学者认可和接受。

因此，本文拟选取 GAMLSS 模型进行非一致性水文频率模型构建，并对马莲河流域不同水文站点的观测数据进行水文频率计算，得到不同频率下的径流设计值。通过与传统方法的对比，指出在变化环境下水文数据筛选和方法应用的建议。

2. 研究区概况与数据来源

马莲河流域地处泾河流域东北部，界于东经 106°37'~108°34'，北纬 35°18'~37°23'，流域主要包括了甘肃省庆阳市除镇远县外的其他七个县区，涉及了黄土高原沟壑区、黄土丘陵沟壑区和黄土丘陵区三种地貌类型，总面积 19,086 km²。马莲河上游主源西川，自发源地麻黄山南流到洪德镇，与东北方向来的发源于白玉山西麓的东川汇合。西川全长 82 km，流域面积 4700 km²，多年平均年径流量 0.378 亿 m³。东川全长 52 km，流域面积 2093 km²，多年平均年径流量 0.079 亿 m³。两川汇合后称为环江。环江过洪德镇东南流，右岸依次纳入玄城沟、马坊川、城西川等主要支流后到达环县县城。环江继续向东南流，左岸纳入安山川，右岸纳入合道川，继而到达曲子镇，即出环县县境进入庆阳县境。干流继续东南向流动，分别纳入马岭东沟、柳树河沟、蔡家沟等小支流到达庆阳县城(庆城镇)，左岸纳入柔远河，自此以下环江改称马莲河。马莲河过庆阳县继续东南流，左纳合水川后呈正南流向，后左岸纳入固城川和九龙河，最后于甘陕边界处的政平村汇入泾河，整个马莲河水系图如图 1 所示。马莲河全长 374.8 km，多年平均流量 13.9 m³/s，年输沙量 1.34 亿吨。

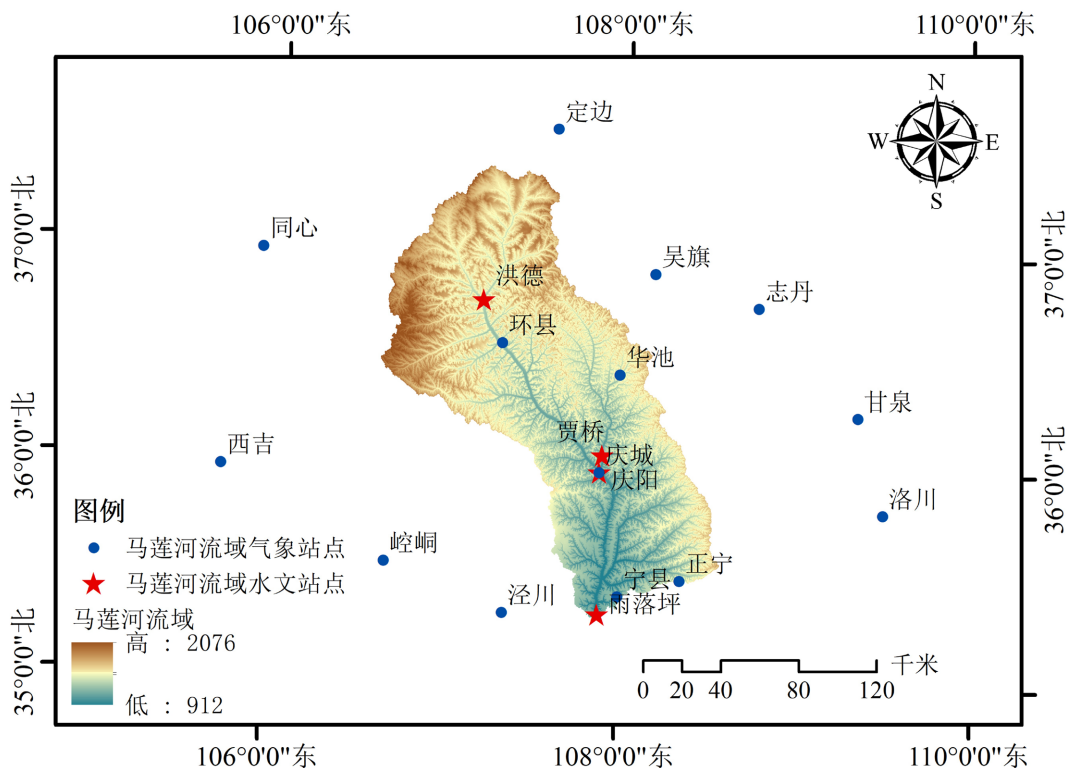


Figure 1. Malian River drainage and hydrological site distribution map
图 1. 马莲河水系及水文站点分布图

本文选取马莲河干流洪德站、庆阳站、雨落坪站的实测径流数据(序列长度为 1964~1997、2006~2012 和 2015~2019 年)和支流柔远河贾桥站的实测径流数据(序列长度为 1979~1997、2006~2012 和 2015~2019 年)，所有径流数据均来源于黄河水利委员会印制的水文数据年鉴。图 2 显示了不同时期四个水文站实测径流的箱图，由

图可知, 马莲河径流量在过去的 50 余年中, 已表现出了较为明显的减少趋势。

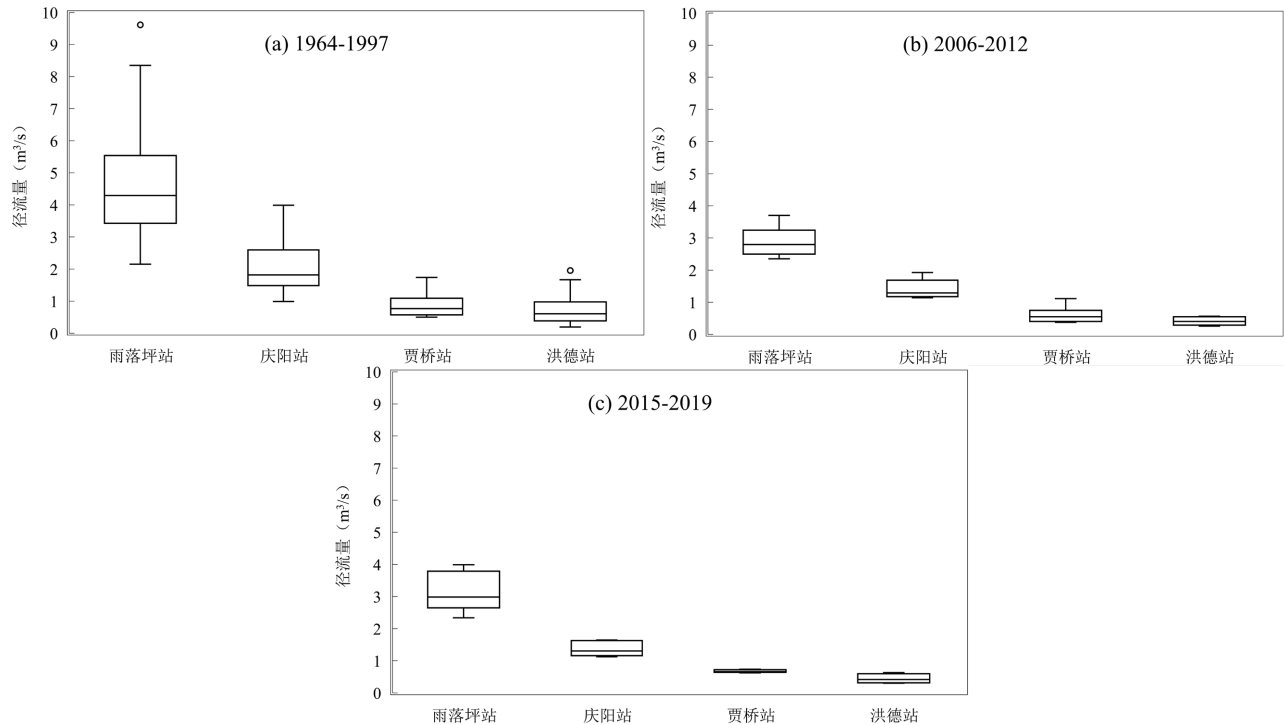


Figure 2. Map of observed runoff box at different stations of Malian River

图 2. 马莲河不同站点实测径流箱图

3. 研究方法

GAMLSS 模型是一种用于非一致性分析的(半)参数回归模型, 可考量广泛的随机变量频率分布类型[3]。GAMLSS 模型是对 GLM (Generalized Linear Models)广义线性模型和 GAM (Generalized Additive Models)广义加法模型的进一步优化与修正。模型的建立不仅只针对于均值的分布, 也考虑了标准差、峰度、偏度等统计量的分布情况, 对非一致性序列具有良好的适用性。

本文将不同水文站点的月径流量作为研究对象, 构建 GAMLSS 模型, 使用 poly 多项式模型对各月径流量进行拟合, 并藉由评价准则, 确定最佳多项式次数和拟合分布类型, 最后依据最优拟合模型, 绘制分位数图, 得到不同频率下的月径流量。

评价准则采用 AIC 和 SBC 准则。AIC 和 SBC 值越小, 表明模型拟合的越好, 因此以最小 AIC 和 SBC 值对应的模型为最优模型[15]。

4. 结果与讨论

应用 GAMLSS 模型对马莲河流域洪德、贾桥、庆阳、雨落坪站的实测水文数据进行概率分布拟合, 并根据拟合分布类型的评价结果, 均选择正态分布 LOGNO, 多项式次数选择 3 次, 可得到不同站点的不同月份不同频率下的径流量设计值, 如图 3 所示。

将其与 P-III 曲线所计算得到的结果进行对比, 如图 4 所示。由图可知, 不考虑径流序列本身的非一致性, 直接应用 P-III 曲线进行拟合适线, 所得到的不同频率的设计径流量与面向水文非一致性的 GAMLSS 方法差异较大。总体上表现为, 特枯水年型(95%和 90%), GAMLSS 模型的计算结果小于 P-III 曲线, 而在其他年型, 则表现为 GAMLSS 模型的计算结果大于 P-III 曲线。

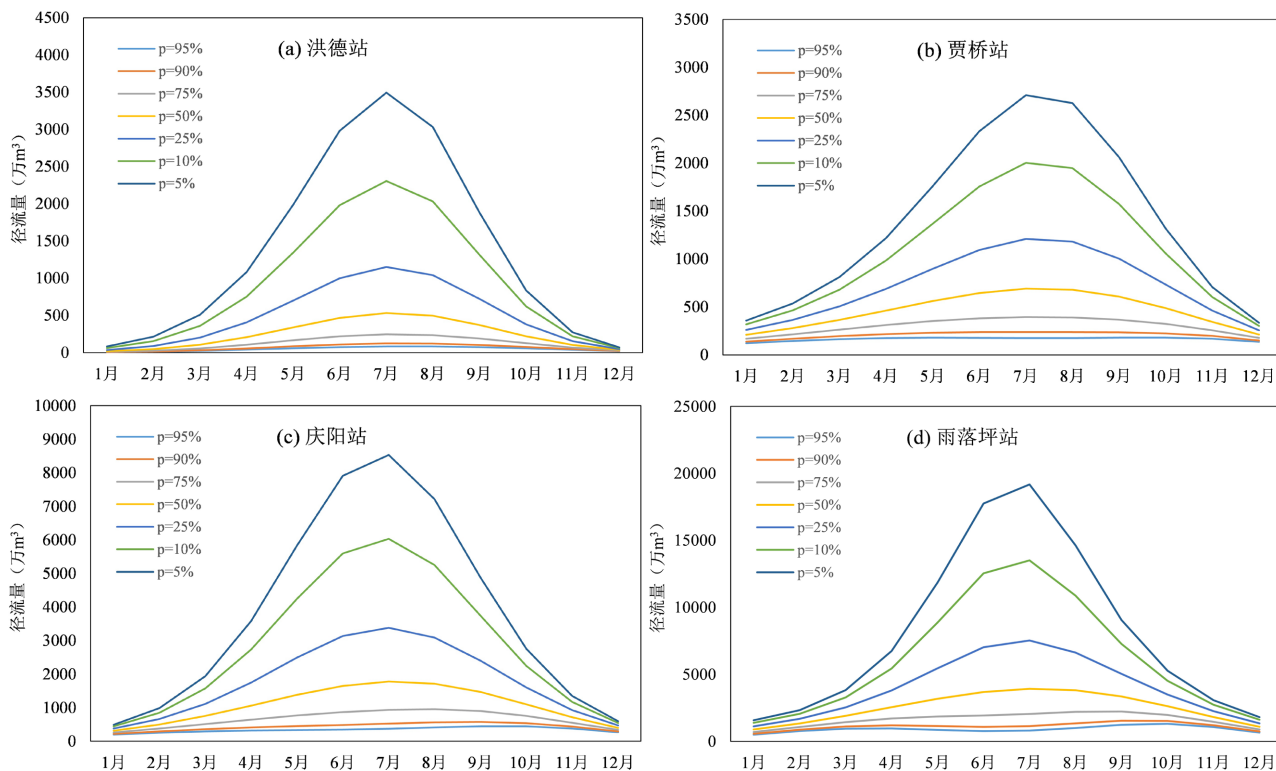


Figure 3. Design values of annual runoff distribution measured at different hydrological stations

图 3. 不同水文站实测径流年内分布设计值

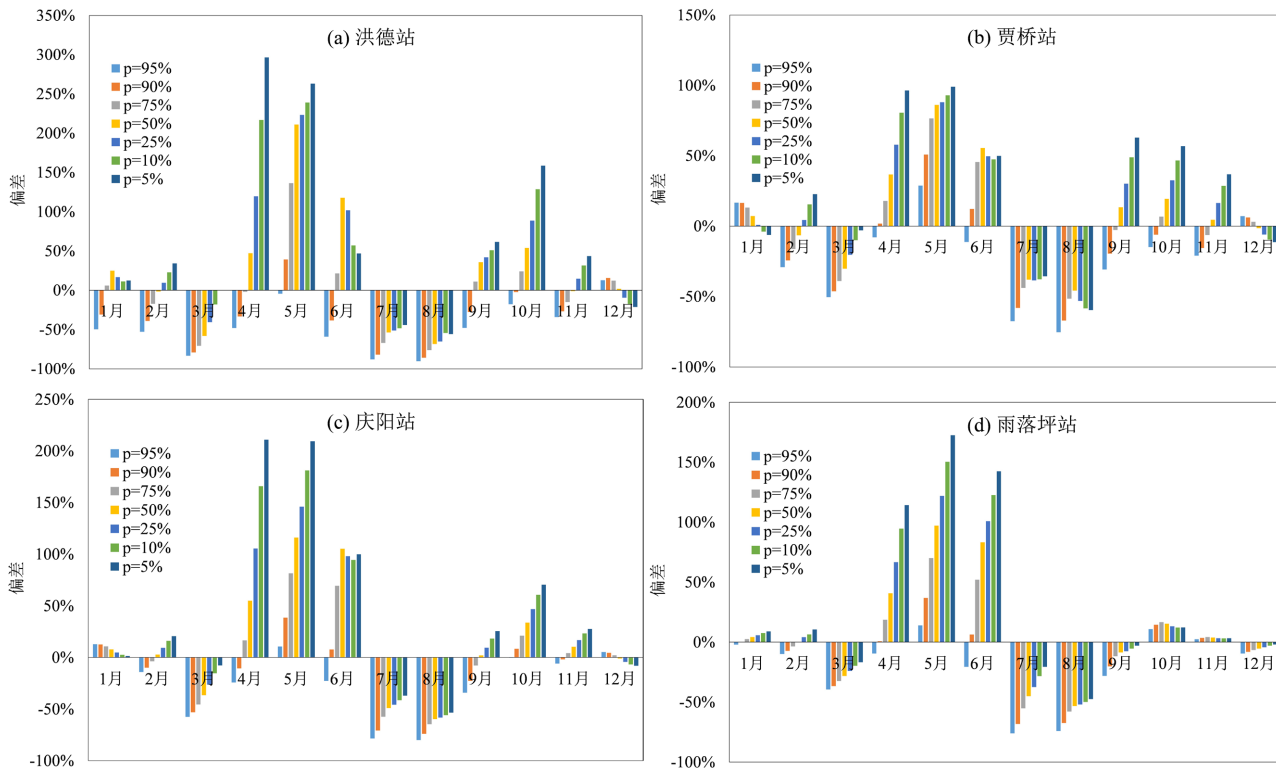


Figure 4. Deviation of the calculated results between GAMLSS model and P-III curve

图 4. GAMLSS 模型与 P-III 曲线计算结果的偏差

将不同频率下两种方法的计算结果的最大偏差进行统计分析(图 5), 可知特枯水期(95%, 90%)计算的最大偏差主要集中在-67%~-90%, 即 GAMLSS 模型的计算结果小于 P-III 曲线, 枯水期(75%)计算的最大偏差主要集中在 70%~136%, GAMLSS 模型的计算结果大于 P-III 曲线, 平水期(50%)计算的最大偏差主要集中在 86%~210%, 丰水期(25%)计算的最大偏差主要集中在 88%~223%, 特丰水期(10%, 5%)计算的最大偏差主要集中在 92%~296%。

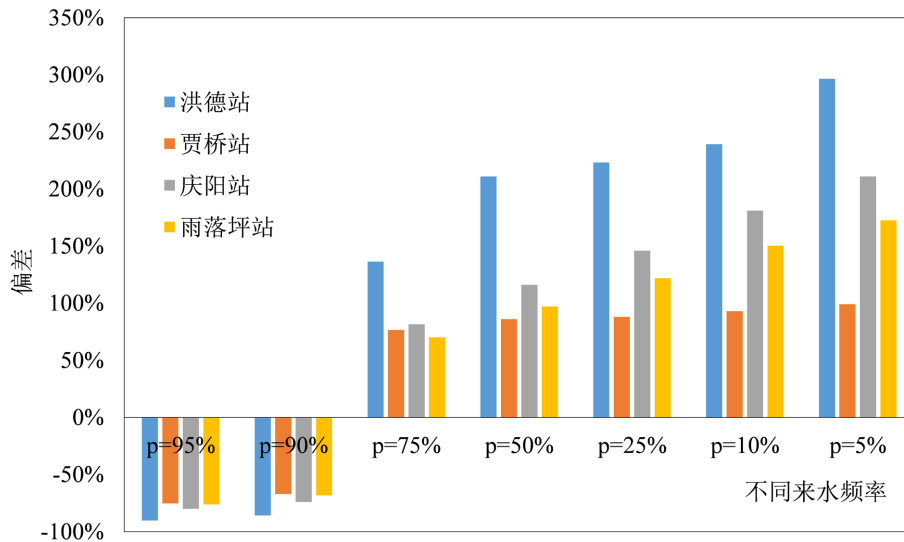


Figure 5. Maximum deviation of the calculated results between GAMLSS model and P-III curve

图 5. GAMLSS 模型与 P-III 曲线计算结果的最大偏差

从年内时程分布上看, 两种模型的计算偏差具有较强的聚集特征, 主要集中于 4~8 月份, 95%和 90%年型, 主要偏差集中于 7~8 月, 而在 75%、50%、25%、10%和 5%年型下, 这主要集中于 4~6 月份。

透过以上分析, 不难发现对于非一致性径流序列, 若不使用正确的频率计算方法, 所得的径流设计值可能产生极大的偏差, 继而影响工程规模的确定。从影响的程度看, 丰水期 > 平水期 > 枯水期。

5. 结论

本文使用 GAMLSS 模型, 对马莲河流域不同水文站点的观测数据进行水文频率计算, 得到如下结论:

1) 通过对马莲河流域水文站点观测数据进行箱图分析, 得知该流域 50 余年径流量呈现明显下降趋势。
2) 应用 GAMLSS 模型对马莲河流域洪德、贾桥、庆阳、雨落坪站的实测水文数据进行概率分布拟合, 得到不同站点的不同月份不同频率下的径流量。

3) 通过对比 GAMLSS 模型设计值与 P-III 曲线设计值可知: 不考虑径流序列本身的非一致性, 直接应用 P-III 曲线进行拟合适线, 所得到的不同频率的设计径流量与面向水文非一致性的 GAMLSS 方法差异较大。总体上表现为, 特枯水年型(95%和 90%), GAMLSS 模型的计算结果小于 P-III 曲线, 而在其他年型, 则表现为 GAMLSS 模型的计算结果大于 P-III 曲线。

4) 根据两种方法计算结果最大偏差的统计分析结果可知, 特枯水期(95%, 90%)计算的最大偏差主要集中在 -67%~-90%, 即 GAMLSS 模型的计算结果小于 P-III 曲线, 枯水期(75%)计算的最大偏差主要集中在 70%~136%, GAMLSS 模型的计算结果大于 P-III 曲线, 平水期(50%)计算的最大偏差主要集中在 86%~210%, 丰水期(25%)计算的最大偏差主要集中在 88%~223%, 特丰水期(10%, 5%)计算的最大偏差主要集中在 92%~296%。

5) 从年内时程分布上看, 两种模型的计算偏差具有较强的聚集特征, 主要集中于 4~8 月份。其中, 95%和

90%年型, 主要偏差集中于 7~8 月, 而在 75%、50%、25%、10%和 5%年型下, 主要偏差集中于 4~6 月份。

综上所述, 对于非一致性径流序列, 如何选择正确的频率计算方法是至关重要的。频率分析计算所得设计值直接影响后续水利工程规模的确定。有效的水文分析是工程设计和流域区域水资源管理的根本。

基金项目

本研究获得国家自然科学基金项目(51979005)及陕西省自然科学基金基础研究计划项目(2020JM-250)支持。

参考文献

- [1] KHALIQ, M. N., OUARDA, T. B. M. J., ONDO, J. C., et al. Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review. *Journal of Hydrology*, 2006, 329(3): 534-552. <https://doi.org/10.1016/j.jhydrol.2006.03.004>
- [2] 谢平, 陈广才, 雷红富, 等. 水文变异诊断系统[J]. *水力发电学报*, 2010, 29(1): 85-91.
XIE Ping, CHEN Guangcai, LEI Hongfu, et al. Hydrologic variation diagnostic system. *Journal of Hydroelectric Engineering*, 2010, 29(1): 85-91. (in Chinese)
- [3] 熊立华, 江聪, 杜涛, 等. 变化环境下非一致性水文频率分析研究综述[J]. *水资源研究*, 2015, 4(4): 310-319.
XIONG Lihua, JIANG Cong, DU Tao, et al. A review of inconsistent hydrological frequency analysis in changing environments. *Journal of Water Resources Research*, 2015, 4(4): 310-319. (in Chinese)
- [4] KHARIN, V. V., ZWIERS, F. W. Estimating extremes in transient climate change simulations. *Journal of Climate*, 2005, 18(8): 1156-1173. <https://doi.org/10.1175/JCLI3320.1>
- [5] 宋浩然. 气候变化影响下辽西地区水文非一致性概率分析研究[J]. *地下水*, 2019, 41(3): 176-178.
SONG Haoran. Probability analysis of hydrological inconsistency under the influence of climate change in western Liaoning Province. *Journal of Groundwater*, 2019, 41(3): 176-178. (in Chinese)
- [6] VILLARINI, G., SMITH, J. A., SERINALDI, F., et al. Flood frequency analysis for nonstationary annual peak records in an urban drainage basin. *Advance in Water Resource*, 2009, 32(8): 1255-1266. <https://doi.org/10.1016/j.advwatres.2009.05.003>
- [7] 刘丙军, 邱凯华, 廖叶颖. 基于 TVM 的西北江三角洲地区非一致性洪水频率分析[J]. *中山大学学报(自然科学版)*, 2016, 55(4): 130-135.
LIU Bingjun, QIU Kaihua and LIAO Yeying. Frequency analysis of inconsistent flood in northwest River Delta based on TVM. *Acta Scientiarum Naturalium Universitatis Sunyatseni*, 2016, 55(4): 130-135. (in Chinese)
- [8] 叶长青, 陈晓宏, 张家鸣, 等. 具有趋势变异的非一致性东江流域洪水序列频率计算研究[J]. *自然资源学报*, 2013, 28(12): 2105-2116.
YE Changqing, CHEN Xiaohong, ZHANG Jiaming, et al. Frequency calculation of flood Sequence with trend variation in the Dongjiang River basin. *Journal of Natural Resources*, 2013, 28(12): 2105-2116. (in Chinese)
- [9] 熊立华, 郭生练. 两变量极值分布在洪水频率分析中的应用研究[J]. *长江科学院院报*, 2004, 21(2): 35-37.
XIONG Lihua, GUO Shenglian. Application of extreme value distribution of two variables in flood frequency analysis. *Journal of Yangtze River Scientific Research Institute*, 2004, 21(2): 35-37. (in Chinese)
- [10] RIGBY, R. A., STASINOPOULOS, D. M. Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 2005, 54(3): 507-554. <https://doi.org/10.1111/j.1467-9876.2005.00510.x>
- [11] VILLARINI, G., SERINALDI, F., SMITH, J. A., et al. On the stationarity of annual flood peaks in the continental United States during the 20th century. *Water Resources Research*, 2009, 45(8): W08417. <https://doi.org/10.1029/2008WR007645>
- [12] 李婧, 闫磊, 屈春来, 等. 分位数回归和 GAMLSS 模型在非一致性洪水频率分析中的比较[J]. *水利水电科技进展*, 2020, 40(5): 48-54+75.
LI Jing, YAN Lei, QU Chunlai, et al. Comparison of quantile regression and GAMLSS model in inconsistent flood frequency analysis. *Advances in Science and Technology of Water Resources*, 2020, 40(5): 48-54+75. (in Chinese)
- [13] 鲁帆, 肖伟华, 戴雁宇, 等. 黄河干流年径流量非一致性频率计算[J]. *水力发电学报*, 2020, 39(12): 76-84.
LU Fan, XIAO Weihua, DAI Yanyu, et al. Calculation of inconsistent frequency of annual runoff of the Yellow River. *Journal of Hydroelectric Engineering*, 2020, 39(12): 76-84. (in Chinese)
- [14] 高洁. 基于 GAMLSS 模型的水文系列非平稳性研究[J]. *水力发电*, 2019, 45(7): 1-6.
GAO Jie. Study on hydrologic series nonstationarity based on GAMLSS model. *Water Power*, 2019, 45(7): 1-6. (in Chinese)
- [15] 关晓明. 基于 GAMLSS 模型的辽西地区气候变化对暴雨洪水极值影响评估[J]. *水利技术监督*, 2021(1): 52-55.
GUAN Xiaoming. Impact assessment of climate change on extreme value of rainstorm and flood in western Liaoning Province based on GAMLSS model. *Technical Supervision in Water Resources*, 2021(1): 52-55. (in Chinese)