

DOI:10.3876/j.issn.1000-1980.2021.02.004

广义相加模型在乌江夏季径流预报中的应用

荣艳淑^{1,2}, 胡玉恒^{1,3}, 冯瑞瑞¹, 殷雨婷¹, 李崇浩⁴

(1. 河海大学水文水资源学院, 江苏 南京 210098; 2. 河海大学水利学科专业实验教学中心, 江苏 南京 210098;
3. 南京恩瑞特实业有限公司, 江苏 南京 211106; 4. 中国南方电网电力调度控制中心, 广东 广州 510623)

摘要: 基于前期冬季海温指数, 构建具有 4 个非线性指数和 9 个线性指数的广义相加模型(GAM), 对乌江流域洪家渡夏季径流进行了模拟与预测, 利用 5 种评估指标, 包括最小信息准则(AIC)、均方根误差(RMSE)、平均绝对误差(MAE)、概率空间线性误差(LEPS)和线性相关(r), 评估 GAM 和广义线性模型(GLM)模拟效果。结果表明: 在 AIC、RMSE 和 LEPS 评估指标上, GAM 的模拟效果与实测值相比均小于 GLM, 在线性相关系数上显著大于 GLM。因此, GAM 的模拟效果明显优于 GLM。利用留一法交叉验证对洪家渡夏季径流分别进行了 GAM 和 GLM 预测, 结果表明, GAM 与实测数据的相关系数提高到 0.41, 相对误差小于 30% 的预测值达到 60% 以上, 特别是在洪家渡典型洪水中, GAM 的预测误差仅在 10% 左右, 在典型干旱年中, GAM 的预测误差小于 10%, GAM 的预测结果比 GLM 明显改善。因此, 考虑了径流量与预报因子的非线性关系, 在径流预报中使用 GAM 进行建模, 能够有效改善线性回归模型的模拟效果和预测精度。

关键词: 广义相加模型(GAM); 广义线性模型(GLM); 留一法交叉验证; 夏季径流; 海温因子
中图分类号: P338 **文献标志码:** A **文章编号:** 1000-1980(2021)02-0121-06

Application of generalized additive model in summer runoff forecasting of Wujiang Basin

RONG Yanshu^{1, 2}, HU Yuheng^{1, 3}, FENG Ruirui¹, YIN Yuting¹, LI Chonghao⁴

(1. College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China;
2. Nanjing NRIET Industrial Co., Ltd., Nanjing 211106, China;
3. The Experimental Teaching Center of Water Resources of Hohai University, Nanjing 210098, China;
4. Power Dispatching Control Center, China Southern Power Grid, Guangzhou 510623, China)

Abstract: Based on the sea surface temperature (SST) indices in previous winters, a Generalized additive model (GAM) model was constructed with four non-linear and nine linear indices, to simulate and forecast the summer runoff of Hongjiadu hydropower station on the Wujiang Basin. Five evaluation indices were used to evaluate the simulation results of GAM and Generalized linear model (GLM), including the Akaike information criterion (AIC), the root mean square error (RMSE), the mean absolute error (MAE), the linear error in probability space (LEPS) and the linear correlation r . The results showed that the ratios of simulated data to real data by GAM were smaller than those by GLM in AIC, RMSE and LEPS, while the linear correlation coefficient of GAM was obviously larger than that of GLM, thus GAM performed better than GLM. Leave-one-out Cross validation was used to forecast the summer runoff of Hongjiadu hydropower station with GAM and GLM, respectively. The forecast results indicated that the correlation coefficient between the GAM result and observed data was improved to 0.41 and predicated data with relative error less than 30% was more than 60%. The prediction error of GAM was around 10% in the typical flood year of Hongjiadu and less than 10% in the typical dry year. It can be concluded that the prediction accuracy was improved in GAM compared with that in GLM. Consequently, considering the the non-linear relationship between runoff and prediction factors, the use of GAM in runoff forecasting can effectively improve the simulation result and prediction accuracy than the linear regression model.

Key words: Generalized additive model (GAM); Generalized linear model (GLM); Leave-one-out cross validation; summer runoff; SST factors

基金项目: “十三五”国家重点研发计划(2016YFA0601504); 国家自然科学基金重点国际(地区)合作研究项目(51420105014)
作者简介: 荣艳淑(1961—), 女, 教授, 博士, 主要从事水文气象和气候变化研究。E-mail: ysrn@hhu.edu.cn
引用本文: 荣艳淑, 胡玉恒, 冯瑞瑞, 等. 广义相加模型在乌江夏季径流预报中的应用[J]. 河海大学学报(自然科学版), 2021, 49(2): 121-126.
RONG Yanshu, HU Yuheng, FENG Ruirui, et al. Application of generalized additive model in summer runoff forecasting of Wujiang Basin[J]. Journal of Hohai University(Natural Sciences), 2021, 49(2): 121-126.

中长期径流预测一直是水文预报的一大难题。传统的统计水文预报模型,如时间序列分析和各种回归模型等,由于计算简便,模型参数有一定的物理意义而被广泛地应用于水文预报工作^[1]。如,基于多种气象因子建立的逐步回归模型、多元线性回归模型和时间序列分解模型对月、旬径流都有较好的预报效果^[2-4]。但是传统线性统计模型没有考虑地球气候系统中各要素与径流之间复杂的非线性关系,预测精度很难有进一步提高^[5]。随着计算机技术的发展,人工神经网络、改进人工神经网络和小波分解等,不仅能够处理变量之间的非线性映射关系^[6],而且预测结果明显优于线性回归模型,预报精度也有明显提高^[7-9]。但是这些方法存在模型物理意义不明确,模型通用性较差等缺陷^[10]。近年来涌现了一些考虑了非线性关系,或者使用多种模型融合的技术方法,明显提高了预测水平^[11-15]。广义相加模型(generalized additive model,GAM)就是综合了线性关系和非线性关系的一种复合模型,它承认预测变量与某些物理因子存在线性关系,与另外的物理因子可能存在非线性关系,并基于线性和非线性因子建立综合模型。这一模型在基于海温因子的降水预测、基于气温和降水的大气污染影响分析和基于NAO的阿尔卑斯山冬季极端气温的预测中都得到较好的应用^[16-18]。

乌江流域位于长江上游,流域内水电站网密布,为当地的农业灌溉、生活用水以及防洪发电带来了显著效益。但是,流域内径流预测准确性一直困扰着水电站网管理部门,特别是汛期季风盛行期间,降水变化显著,径流变化更难以预测。在这一背景下,基于广义相加模型,笔者对乌江流域的洪家渡夏季径流进行预测研究,并对GAM的建立、模拟和预测进行评估和探讨,以期能为乌江流域的径流预测提供一个新途径。

1 研究数据与方法

1.1 研究数据

以乌江流域的洪家渡夏季径流为研究对象,以多种海温指数作为预测因子,包括太平洋、大西洋和印度洋的26个海温指数等(表1)。径流资料来自洪家渡水电站,海温指数来自国家气候中心。由于海温与我国夏季降水之间存在较长的滞后相关性^[19],因此选取前期冬季海温指数进行研究。冬季指数为12月、翌年1—2月3个月的平均值,夏季径流为6—8月3个月的平均值。资料时间范围均为1952—2014年。

表1 海温指数与因子代码
Table 1 SST index and codes of factors

因子代码	海温指数名称	因子代码	海温指数名称
x1	NINO1+2 区海表温度距平指数	x14	印度洋暖池强度指数
x2	NINO3 区海表温度距平指数	x15	西太平洋暖池面积指数
x3	NINO4 区海表温度距平指数	x16	西太平洋暖池强度指数
x4	NINO3.4 区海表温度距平指数	x17	大西洋年代际振荡指数
x5	NINO W 区海表温度距平指数	x18	亲潮区海温指数
x6	NINO C 区海表温度距平指数	x19	西风漂流区海温指数
x7	NINO A 区海表温度距平指数	x20	黑潮区海温指数
x8	NINO B 区海表温度距平指数	x21	类 ENSO 指数
x9	NINO Z 区海表温度距平指数	x22	暖池型 ENSO 指数
x10	热带北大西洋海温指数	x23	冷舌型 ENSO 指数
x11	热带南大西洋海温指数	x24	热带印度洋全区一致海温模态指数
x12	西半球暖池指数	x25	热带印度洋海温偶极子指数
x13	印度洋暖池面积指数	x26	副热带南印度洋偶极子指数

1.2 GAM 模型

GAM 是由 Hastie 等^[20]在广义线性模型(GLM)^[21]的基础上改进而来,该模型包括两部分,分别描述了自变量与因变量的线性和非线性关系,因此 GAM 存在参数和非参数两部分,属于半参数模型。这一改进打破了线性模型中关于自变量与因变量线性相关的假定,使得模型应用有了进一步拓展。GAM 可表示为

$$g(y_i) = \beta_0 + \sum_{k=1}^n s(x_{ki}) + \sum_{j=1}^m \beta_j x_{ji} + \varepsilon_i \quad (i = 1, 2, \cdots, l)$$
 (1)

式中: $g(y_i)$ ——因变量 y_i 的连接函数; l ——序列长度; β_0 ——截距; $s(x_{ki})$ ——平滑函数,本文采用薄板样条函数作为平滑函数; n ——非线性因子的个数; x_{ki} ——非线性因子; β_j ——线性部分的回归系数; m ——线性因子的个数; x_{ji} ——线性因子; ε_i ——随机误差,服从高斯正态分布。

取 $l = 63$,为避免平滑过程中出现过度拟合,将薄板样条函数节点的最大个数固定为5。GAM 参数部分

检验使用 t 检验,非参数部分使用 F 检验。

1.3 模型模拟和预测的评估方法

1.3.1 模型模拟的评估方法

基于式(1)对洪家渡夏季径流序列进行模拟,模拟效果选用最小信息准则(Akaike information criterion, AIC)、均方根误差(root mean square error, RMSE)、平均绝对误差(mean absolute error, MAE)、概率空间线性误差(linear error in probability space, LEPS)^[22]和线性相关系数(correlation, r^2)5 种评价指标进行评估。根据最小信息准则,AIC 值越小表明模型模拟效果越好;RMSE 用来衡量模拟值与观测值之间的相对偏差,其值越小表明偏差越小;MAE 能够反映模拟值与实测值的绝对偏离程度,其值越小模拟效果越好;LEPS 使用了 Potts 等^[23]改进的 LEPS,其值越接近于 1,模拟效果越好。 r^2 刻画 2 个序列之间的线性相关,其值越接近 1 表明整个序列模拟值与实测值的相关性越高。

1.3.2 模型预测的评估方法

一般的模型都是取序列长度的一部分数据建立模型,剩余数据作为预测样本,用以评估模型的模拟能力^[24],但是这种方法存在一定的主观性以及测试样本长度的局限性,特别是当预测样本过长时,模型易出现衰退,导致预测效果不好。

交叉验证是一个可以克服上述问题的模型预测检验方法,其主要思想是将数据样本分成两部分,一部分作为训练集进行模型训练,另一部分作为测试集进行模型预测,但是该方法是将数据循环分组,对每个原数据都进行模拟,这样通过多次重复试验能得到较好的模型估计结果^[23]。为了充分利用已有的样本数据得到更可靠的预测结果,本文使用了广义交叉验证^[25]确定薄板样条函数平滑自由度的取值,使用留一法交叉验证^[26]进行模型预测结果的评估。

留一法交叉验证的主要思想是从 l 个样本数据中取一个作为测试集,剩下的 $l-1$ 个作为训练集,例如,对于某时间序列的第一个样本,用 $i = 2,3,\cdots,l$ 个样本建立模型,获得模型参数,再对第一个样本进行模拟;对第二个样本,则是用 $i = 1$ 和 $i = 3,4,\cdots,l$ 个样本建立模型,再对第二个样本进行模拟,如此往复,共进行 l 次训练,得到 l 个预测结果。最后,通过模拟序列和实测序列的相关性分析和预报相对误差频数统计,对 GAM 预报效果进行评价。

2 结果与讨论

2.1 模型因子挑选

本文选择前期冬季作为海温影响的关键时期,将前期冬季 26 个海温指数作为洪家渡夏季径流预测的潜在预报因子。同时,为了验证 GAM 模拟的效果,使用 GLM 进行对比分析。根据逐步回归挑选因子的最小信息准则^[27](AIC)分别筛选出 GAM 和 GLM 模型的预报因子。GAM 挑选出 13 个预报因子,分别为 x_1 、 x_2 、 x_3 、 x_4 、 x_6 、 x_7 、 x_{10} 、 x_{14} 、 x_{17} 、 x_{20} 、 x_{22} 、 x_{24} 、 x_{26} ;GLM 挑选出 14 个预报因子,分别为 x_4 、 x_5 、 x_7 、 x_8 、 x_9 、 x_{10} 、 x_{11} 、 x_{12} 、 x_{13} 、 x_{14} 、 x_{17} 、 x_{20} 、 x_{22} 、 x_{24} 。2 个模型共有的预报因子是 8 个,这是线性模型和非线性模型对自变量影响因变量的物理机制存在差异的原因。

用薄板样条函数分别对 GAM 中的 13 个因子进行平滑处理,平滑自由度(用 edf 表示)的值为 1 时,表明自变量与因变量为线性关系,超过 1 时,表明自变量与因变量为非线性关系,其值越大,表明非线性关系越明显。经过平滑处理,有 9 个因子与洪家渡径流是线性关系,有 4 个因子与洪家渡径流是非线性关系(表 2)。

由表 2 可见,NINO1+2 区海表温度距平指数(x_1)、黑潮海温指数(x_{20})、热带印度洋全区一致海温指数(x_{24})和副热带南印度洋偶极子指数(x_{26})均对洪家渡径流有非线性影响,而其他指数,包括 NINO3 区(x_2)、NINO4 区(x_3)、NINO3.4 区(x_4)、NINO C 区(x_6)和 NINO A 区(x_7)海表温度距平指数、热带北大西洋海温指数(x_{10})、印度洋暖池强度指数(x_{14})、大西

表 2 GAM 中各预报因子的线性和非线性属性

Table 2 Linear and Non-Linear relationships of prediction factors in GAM

因子代码	edf	线性或非线性	因子代码	edf	线性或非线性
x_1	1.780	非线性	x_{14}	1.000	线性
x_2	1.000	线性	x_{17}	1.000	线性
x_3	1.000	线性	x_{20}	2.531	非线性
x_4	1.000	线性	x_{22}	1.000	线性
x_6	1.000	线性	x_{24}	3.469	非线性
x_7	1.000	线性	x_{26}	2.118	非线性
x_{10}	1.000	线性			

洋年代振荡指数(x17)和暖池型 ENSO 指数(x22)与洪家渡径流的关系可用线性关系描述。

以 x20 和 x24 为例,图 1 给出它们与模拟径流量偏差的相关关系,图中实线为平滑函数值,点线之间的区域为信度 95% 水平的范围。可以看出,x20 和 x24 与模拟值偏差的散点基本分布在信度 95% 水平之间,表明对 x20 和 x24 作为非线性因子的判定是合理的。

2.2 模型模拟效果评估

由图 2 可以看出 GAM 和 GLM 对径流量模拟的趋势相似,且与实测序列基本相一致。从一些异常年份看,比如 1954 年、1957 年、1967 年、1998 年、2001 年、2008 年、2012 年等几个流量异常偏多的年份,以及 2005 年之后几个流量异常偏少的年份,均得到了较为准确的模拟。通过 GAM 与 GLM 的对比分析,在 1954 年、1957 年、1998 年、2001 年、2008 年等几个流量异常偏多的年份,GAM 的模拟效果更接近实测值。在 2 个模型模拟值明显不一致的年份,如 1967 年、1985 年、1994 年,GAM 的模拟值与实测值更为接近,模拟效果更好。

由表 3 可知,GAM 的 AIC 值为 733.04,而 GLM 的 AIC 值为 752.57,表明 GAM 的模拟拟合效果比 GLM 好。GAM 的 RMSE 与 MAE 值均低于 GLM,说明 GAM 在整体序列上的模拟精度高于 GLM。GAM 模拟值与实测值序列的相关性(r^2)与 GLM 相比也有很大的提升。综合分析,GAM 模拟效果比 GLM 模拟效果有了显著提升。

由图 3(a)可以看出,模型残差的分布大致呈正态分布的规律,而当值大于 100 以及小于-50 时,部分散点并不完全处于 1:1 直线上;图 3(b)也显示出残差分布出现了略呈偏态的特征,这表明残差分布并不完全服从正态分布。采用 Shapiro 等^[28]提出的 W' 正态检验方法,计算了统计量 $W'=0.991$,达到了统计显著性 95% 信度水平,认为残差通过了正态性检验,服从正态分布。Montgomery^[29]指出使用 GAM 时,样本低于 100 时有可能出现类似图 3 的现象。

为进一步验证模型的适用性,探讨了模型残差散点分布以及实测流量与模拟流量散点的拟合情况。由图 4(a)可知,残差大部分处于-50~50 之间,分布基本合理。从图 4(b)可知,其值大致呈现 1:1 线性分布的趋势,流量存在一个异常大值(>500)也得到了较为准确的模拟,表明模型的拟合效果较好。

综上所述,可以认定 GAM 用于本文的径流模拟是合适的。

2.3 模型预测效果评估

采用留一法交叉验证对 GAM 和 GLM 的预测结果进行对比分析,其中样本数据为 1952—2014 年。根据留一法交叉验证的原理,每次均选取其中的 62 a 作为训练集,剩下的 1 a 作为测试集验证模型的优劣,循环往复,共进行 63 次模拟与预测。

图 5 为实测值与 2 种模型预

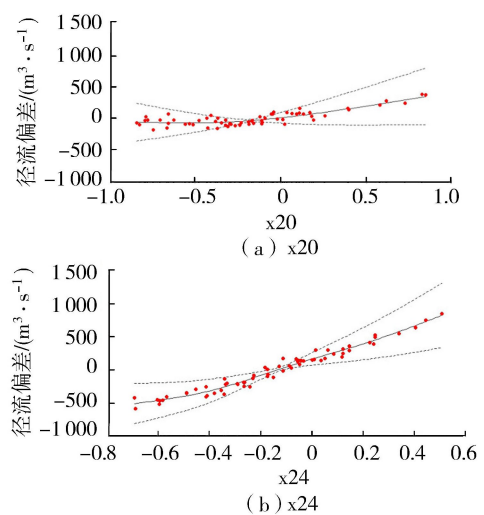


图 1 x20 和 x24 与洪家渡夏季径流偏差的相关关系

Fig.1 Relations between summer runoff deviation at Hongjiadu and factors, x20 and for x24

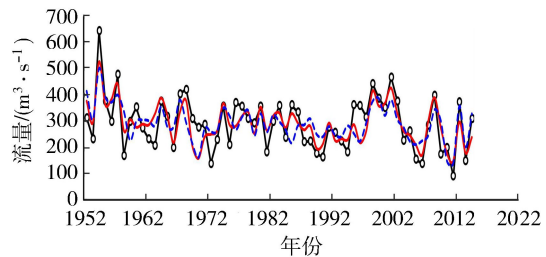


图 2 GAM 和 GLM 模拟值与实测值比较

Fig.2 Comparison between simulated values of GAM and GLM and observed values

表 3 GAM 和 GLM 评价指标对比

Table 3 Comparison of evaluation index values for GAM and GLM

模型	AIC	RMSE	LEPS	MAE	r^2
GAM	733.04	58.39	0.47	48.35	0.81
GLM	752.57	73.69	0.31	62.54	0.66

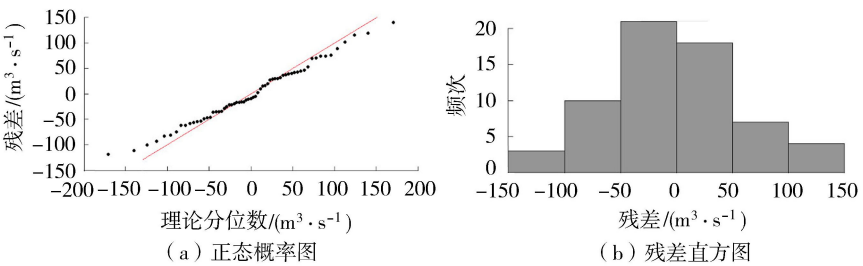


图 3 模型残差正态概率图和残差直方图

Fig.3 Normal probability plot and histogram of model residuals

测值的对比,其中实测值与 GLM 预测序列的相关性为 0.267,通过了信度 95% 的显著性水平;实测值与 GAM 预测序列的相关系数为 0.407,通过了信度 99.99% 的显著性检验。因此,相较于 GLM, GAM 的预测能力也有了显著提升。在一些流量异常年份的预测上,如 1957 年、1968 年、2001 年、2008 年几个流量偏多年份, GAM 的预测效果相较于 GLM 均有明显提升,而在 1985 年和 1998 年 2 个 GLM 预测值误差显著偏大的年份, GAM 的预测结果也有了很大的改善。虽然也有部分年份 GLM 预测更接近实测值,但从整体趋势及一些流量异常的年份看, GAM 的预测能力均高于 GLM。

从图 6 可以看到在 10% 的相对误差绝对值范围内, GAM 预报正确年份为 13 a,与 GLM 相比少 1 a,而在 10% ~ 20% 以及 20% ~ 30% 的误差范围内 GAM 预报准确的年份比 GLM 分别多 3 a 和 5 a。而当相对误差绝对值大于 40% 时, GLM 预报的年份要明显多于 GAM。因此,在频数分布上也可以看出, GAM 模型的预报效果明显优于 GLM 的预报效果。

如果以预测相对误差绝对值小于等于 30% 为合格的预测结果, GAM 和 GLM 的预测合格率分别为 85.5% 和 79.0%。表明 GAM 的预测合格率高于 GLM, 预测效果更好。

文献[15]使用了 5 种模型(时间序列方法、周期均值叠加法、逐步回归、岭回归和主成分回归)对 2004 年和 2005 年分别进行了预测分析,从 6—8 月平均相对误差看,这 2 a 平均相对误差分别是 18.0% 和 142.0%,而 GAM 对这两年预测的相对误差分别是 9.0% 和 68.8%,预测精度均明显提高。这也印证了考虑前期海温因子的非线性作用后,模型模拟及预测能力均有明显提高。因此, GAM 能够对传统线性回归模型的预报效果加以改进。

3 结 语

在前期海温因子中,挑选了 4 个非线性因子和 9 个线性因子,构建的 GAM 对洪家渡的夏季径流预测非常有效,通过 AIC、RMSE、MAE、LEPS、相关性 5 种评价指标评估, GAM 都优于 GLM。

对洪家渡夏季径流的预测评估结果表明, GAM 的预测效果比 GLM 明显提高,从相对误差绝对值的频数、极值预报误差和预测合格率等 3 个方面的分析表明, GAM 的误差范围更小,极值预测的精度和预测合格率均比较高。

径流的中长期预测对水库调度和水资源管理有重要作用,研究高精度的径流预测模型意义重大。本文的尝试尽管提高了洪家渡夏季径流的预测能力,但是,今后通过因子影响机理分析和消除因子之间的共线性等问题,选取更好的预测指标,或许可以进一步提高 GAM 的模拟精度。

参考文献:

[1] 范钟秀. 中长期水文预报[M]. 南京: 河海大学出版社, 1999.

[2] 李红霞, 何清燕, 彭辉, 等. 基于耦合相似指标的最近邻法在年径流预测中的应用[J]. 水科学进展, 2015, 26(2): 161-168. (LI Hongxia, HE Qingyan, PENG Hui, et al. Nearest neighbor method based on a coupled similarity indicator and its application in annual runoff prediction[J]. Advances in Water Science, 2015, 26(2): 161-168. (in Chinese))

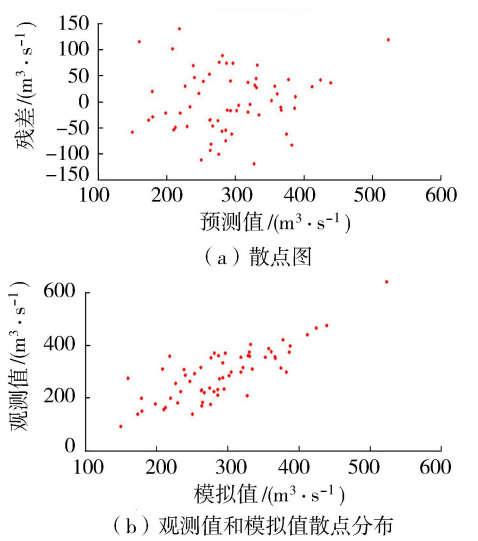


图 4 残差散点图及流量的实测值和模拟值散点分布

Fig.4 Residual scatter plot and scatter plot of observed runoff versus predicted runoff

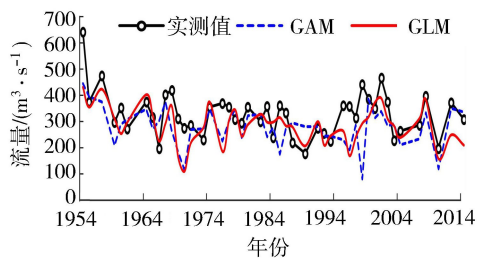


图 5 2 种模型的预测值与实测值对比

Fig.5 Comparison between prediction values of GAM and GLM and observed values

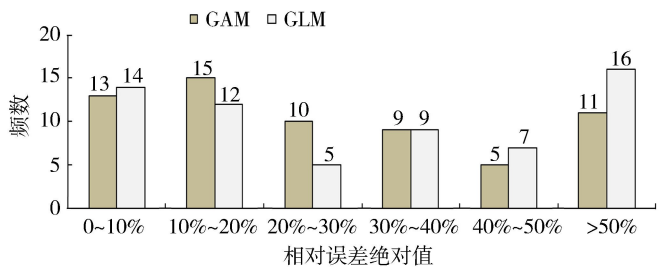


图 6 GAM 和 GLM 预测值相对误差频数分布

Fig.6 Frequency distribution of prediction relative error of GAM and GLM

- [3] 王力磊, 袁晶瑄, 徐炜, 等. 桓仁流域汛期旬径流预报方法研究[J]. 水文, 2012, 32(6): 52-55. (WANG Lilei, YUAN Jingxuan, XU Wei, et al. Research on 10-day runoff forecasting method for Huanren Basin in flood season[J]. Journal of China Hydrology, 2012, 32(6): 52-55. (in Chinese))
- [4] 张闻胜, 王银堂. 时间序列分解模型在水文要素中长期预报中的应用[J]. 水文, 2001, 21(1): 21-24. (ZHANG Wensheng, WANG Yintang. Application of the time series decomposable model in medium and long-term hydrologic forecasting [J]. Journal of China Hydrology, 2001, 21(1): 21-24. (in Chinese))
- [5] 朱灵子. 基于统计模型的汛期降水预测研究[D]. 金华: 浙江师范大学, 2015.
- [6] 朱大奇, 史慧. 人工神经网络原理及应用[M]. 北京: 科学出版社, 2006.
- [7] 董全, 黄小玉, 宗志平. 人工神经网络法和线性回归法对降水相态的预报效果对比[J]. 气象, 2013, 39(3): 324-332. (DONG Quan, HUANG Xiaoyu, ZONG Zhiping. Comparison of artificial neural network and linear regression methods in forecasting precipitation types[J]. Meteorological Monthly, 2013, 39(3): 324-332. (in Chinese))
- [8] 冯亚文, 任国玉, 张丽, 等. 长江上游月降水人工神经网络预测模型[J]. 长江流域资源与环境, 2011, 20(1): 40-47. (FENG Yawen, REN Guoyu, ZHANG Li, et al. Artificial neural network models for forecasting monthly precipitation in the upper Yangtze River[J]. Resources and Environment in the Yangtze Basin, 2011, 20(1): 40-47. (in Chinese))
- [9] 曹启辉, 王文圣, 汤成友. 一种新的小波网络组合预测模型[J]. 人民长江, 2006, 37(11): 65-67. (CAO Qihui, WANG Wensheng, TANG Chengyou. A new forecasting model combined wavelet and network[J]. Yangtze River, 2006, 37(11): 65-67. (in Chinese))
- [10] 杨晓帆, 陈廷槐. 人工神经网络固有的优点和缺点[J]. 计算机科学, 1994, 21(2): 23-26. (YANG Xiaofan, CHEN Tinghuai. Inherent advantages and disadvantages of artificial neural networks[J]. Computer Science, 1994, 21(2): 23-26. (in Chinese))
- [11] 霍文博, 朱跃龙, 李致家, 等. 新安江模型和支持向量机模型实时洪水预报应用比较[J]. 河海大学学报(自然科学版), 2018, 46(4): 283-289. (HUO Wenbo, ZHU Yuelong, LI Zhijia, et al. Comparison of Xin'anjiang model and support vector machine model in the application of real-time flood forecasting[J]. Journal of Hohai University(Natural Sciences), 2018, 46(4): 283-289. (in Chinese))
- [12] 李培都, 司建华, 冯起, 等. 疏勒河年径流量变化特征分析及模拟[J]. 水资源保护, 2018, 34(2): 52-60. (LI Peidu, SI Jianhua, FENG Qi, et al. Analysis and simulation on annual runoff variation characteristics of Shule River[J]. Water Resources Protection, 2018, 34(2): 52-60. (in Chinese))
- [13] 薛联青, 魏卿, 魏光辉. 塔里木河干流地表水与地下水耦合模拟[J]. 河海大学学报(自然科学版), 2019, 47(3): 195-201. (XUE Lianqing, WEI Qing, WEI Guanghui. Coupled simulation of surface water and groundwater in the main stream of Tarim River[J]. Journal of Hohai University (Natural Sciences), 2019, 47(3): 195-201. (in Chinese))
- [14] 张华, 陈会飞. 基于历史数据的系统隐患预测研究[J]. 电信工程技术与标准化, 2015(11): 68-72. (ZHANG Hua, CHEN Huifei. Research on prediction of system hidden trouble based on historical data[J]. Telecom Engineering Technics and Standardization, 2015(11): 68-72. (in Chinese))
- [15] 顾月红. 洪家渡水电站月径流量预报方案的研究[D]. 南京: 河海大学, 2007.
- [16] UNDERWOOD F M. Describing long-term trends in precipitation using generalized additive models[J]. Journal of Hydrology, 2009, 364(3/4): 285-297.
- [17] CHAVEZ-DEMOULIN V, DAVISON A C. Generalized additive modelling of sample extremes[J]. Applied Statistics, 2005, 54(1): 207-222.
- [18] PEARCE J L, BERINGER J, NICHOLLS N, et al. Quantifying the influence of local meteorology on air quality using generalized additive models[J]. Ultrasonics, 2011, 45(6): 1328-1336.
- [19] 郭玲. 中国夏季降水影响因子评估和统计预测研究[D]. 南京: 南京信息工程大学, 2011.
- [20] HASTIE T, TIBSHIRANI R. Generalized additive models[J]. Statistical Science, 1986, 1(3): 297-310.
- [21] NELDER J A, BAKER R J. Generalized linear models[J]. Journal of the Royal Statistical Society, 1972, 135: 370-384.
- [22] WARD M N, FOLLAND C K. Prediction of seasonal rainfall in the north nordeste of Brazil using eigenvectors of sea-surface temperature[J]. International Journal of Climatology, 1991, 11(7): 711-743.
- [23] POTTS J M, FOLLAND C K, JOLLIFFE I T, et al. Revised 'LEPS' scores for assessing climate model simulations and long-range forecasts[J]. Journal of Climate, 1996, 9(1): 34-53.
- [24] 王增学, 周志华, 邹晓天, 等. 中长期水文预报方法应用研究[J]. 东北水利水电, 2012, 30(4): 43-44. (WANG Zengxue, ZHOU Zhihua, ZOU Xiaotian, et al. Application study on mid long term hydrological forecasting method[J]. Water Resouces & Hydropower of Northeast China, 2012, 30(4): 43-44. (in Chinese))
- [25] 范永东. 模型选择中的交叉验证方法综述[D]. 太原: 山西大学, 2013.
- [26] GEISSER S. A predictive approach to the random effect model[J]. Biometrika, 1974, 61(1): 101-107.
- [27] SAKAMOTO Y, ISHIGURO M, KITAGAWA G. Akaike information criterion statistics[M]. Berlin: Springer Netherlands, 1986.
- [28] SHAPIRO S S, FRANCIA R S. An approximate analysis of variance test for normality[J]. Journal of the American Statistical Association, 1972, 67(337): 215-216.
- [29] MONTGOMERY D. Design and analysis of experiments[M]. Hoboken: John Wiley & Sons Inc., 2001.