

基于 SVR 的无实测资料小流域山洪灾害 临界雨量预估模型及应用 ——以河南新县为例

王燕云^{1,2}, 原文林², 龙爱华¹, 向立云¹, 邓晓雅¹, 付磊², 宋汉振²

(1. 中国水利水电科学研究院流域水循环模拟与调控国家重点实验室, 北京 100038;

2. 郑州大学水利与环境学院, 河南 郑州 450001)

摘要:山洪灾害是我国防灾减灾工作中亟待解决的突出问题, 临界雨量是山洪灾害预警预报的重要指标, 传统的临界雨量计算方法各有所长, 对资料的要求各不相同, 但对于无实测资料小流域防灾对象临界雨量计算的研究却寥寥无几(这里无实测资料指的是既没有雨量、洪峰资料也无野外断面调查资料)。鉴于此, 基于支持向量机回归拟合算法(SVR), 应用比拟法建立模型, 对无资料小流域山洪灾害临界雨量进行推求。结果表明: (1)SVR 模型具有很好的泛化性, 预测精度高、误差小, 结果较优; (2)利用 SVR 模型预估的临界雨量值与真实值偏差小, 可以作为山洪灾害预警预报指标, 为当地决策部门提供参考。

关键词:无实测资料小流域; 山洪灾害; 临界雨量; 孕灾因子; 支持向量机回归算法

中图分类号:TV788

文献标识码:A

文章编号:1000-0852(2020)02-0042-06

形成山洪灾害的影响因素众多, 机理复杂^[1], 且突发性强, 破坏力大, 防御非常困难^[2]。受气候和地形地貌条件的影响, 我国山洪灾害频发, 是世界上山洪灾害最严重的国家之一^[3]。近年来, 随着我国经济、社会和生态环境的发展, 山洪灾害造成的危害和损失越来越大, 已成为我国防灾减灾工作中亟待解决的突出问题, 山洪灾害预警预报也成为该领域研究的热点与难点问题^[4-5]。

临界雨量和临界水位是山洪灾害预警预报的重要指标, 前者较后者更具普适性^[6-7]。临界雨量的计算分为两大类: 一类是基于数据驱动的统计学法, 一类是基于水文学及水力学物理机制的水文水力学法^[8]。统计学法主要包括单站/区域临界雨量法、灾害与降雨频率分析法、内插法和比拟法等, 而水文水力学法主要包括水位流量反推法、降雨径流关系曲线插值法、降雨驱动指标法、动态临界雨量法以及分布式模型法等^[9]。其

中, 基于数据驱动的统计法精度高, 对资料要求也高; 水文水力学法依据各地的暴雨图集和水文手册等基础资料进行产汇流计算, 现广泛用于无资料和缺资料地区, 但水文水力学法要求有野外沟道调查数据和成灾水位信息, 数据获取需进行野外调查, 耗费大量的人力物力财力, 同时一些河道地形险峻, 给野外测量工作带来了一定的难度。我国 2000 多个县级行政区具有山洪灾害防治任务, 工作任务繁重, 在雨量站布设稀疏且无野外调查资料的情况下, 若能寻找简便可行的方法来对小流域防灾对象临界雨量进行预估, 将为无资料或缺资料小流域山洪灾害预警预报工作, 提供参考和便利, 尽可能地减少山洪灾害造成的人身财产损失。

樊建勇等^[10]对江西省小流域临界雨量分析计算, 找到了流域面积、主沟道长度和主沟道比降与临界雨量之间的相关关系, 但其统计模型缺乏水文机制的考

收稿日期: 2019-01-11

基金项目: 国家重点研发计划项目“跨境流域水生态服务功能经济价值评估与生态补偿机制研究”(2016YFA0601602); “西北内陆区水资源安全状况与风险评估”(2017YFC0404301)

作者简介: 王燕云(1992-), 女, 河南焦作人, 博士研究生, 主要从事社会水文学、水资源规划与配置研究工作。E-mail: happywangyanyun@163.com
通讯作者: 龙爱华(1976-), 男, 湖南安仁人, 副教授, 主要从事社会水文学、跨界河流、水资源规划与配置研究工作。E-mail: aihuadragon@163.com

虑。基于此,针对无资料小流域,本文提出基于支持向量机回归算法(support vector machine for regression, 简称 SVR) 的无实测资料小流域临界雨量预估模型,即:对于无实测资料的小流域防灾对象,找邻近地形地貌条件和植被覆盖条件相似、产汇流机制相同的有资料小流域,找出该小流域的孕灾因子与各个时段临界雨量值的相关关系;采用地理信息技术与比拟法,将所求小流域的地形地貌特征参数和水文特性参数代入前述相关关系,即可推求无资料小流域各防灾对象的临界雨量估值。

1 基于 SVR 的临界雨量预估模型

基于支持向量机回归的临界雨量预估模型,主要用来模拟山洪灾害孕灾因子与临界雨量之间的相关关系,模型构建主要包括三部分:一是山洪灾害孕灾因子识别与分析,也即模型指标的选取;二是模型的构建与检验;三是无实测资料小流域临界雨量的预估。

1.1 SVR 原理简介

SVR 是由 Vapnik 等人在 SVM (support vector machine, 简称 SVM) 分类的基础上,引入 ε 不敏感损失函数得到的,该算法没有数据维数的限制,具有很好的性能和效果,可应用于仿真模拟^[11],目前在很多领域有较为成功的应用^[12-14]。

SVM 的基本思想是寻找一个最优分类曲面使得两个样本分开,而支持向量机回归算法(简称 SVR)旨在多维空间寻找一个最优分类曲面,使得所有训练样本离该分类曲面的总误差最小。具体求解过程参照史峰等^[15]对 SVR 案例分析的详细介绍。

1.2 模型构建

模型构建大体可以分为产生训练集/测试集、创建/训练 SVR 回归模型、仿真测试和性能评价 4 个步骤,本研究应用 MATLAB 进行建模。

(1)产生训练集/测试集。产生的训练集和测试集要可以进行样本互换,即可将样本进行随机排序,并抽取前半部分作为训练集,后半部分作为测试集。

(2)创建/训练 SVR 回归模型。在进行 SVR 回归模型的创建时需安装 libsvm 软件包,软件包中的函数 svmpredict 可以实现模型的创建和训练。在设计时需综合考虑归一化、核函数的类型、参数的取值对回归模型的影响。核函数的选取是建立 SVR 模型至关重要的一步,目前新的核函数层出不穷,但被广泛认同的 4 个基本核函数为线性核函数、d 阶多项式核函数、径向基

核函数和 Sigmoid 核函数。由于径向基核函数(radial basis kernel function, 简称 RBF)涉及参数少、计算简便,故本文选用径向基核函数。模型中参数的选取参照文献[15],采用交叉验证方法寻找最佳的参数,利用最佳的参数训练模型。

(3)仿真测试。建立模型后,输入相关参数,返回的第一个参数为多维空间映射出的预测值,第二个参数为测试集的均方误差(mean square error, MSE)和决定系数 R^2 。MSE 的值越小,预测模型描述实验数据精确度更好。 R^2 越接近 1 时,相关的方程式参考价值越高。计算公式见式(1)~(2):

$$MSE = \frac{1}{t} \sum_{i=1}^t (\hat{y}_i - y_i)^2 \quad (1)$$

$$R^2 = \frac{(1 \sum_{i=1}^t \hat{y}_i y_i - \sum_{i=1}^t \hat{y}_i \sum_{i=1}^t y_i)^2}{(1 \sum_{i=1}^t \hat{y}_i^2 - (\sum_{i=1}^t \hat{y}_i)^2)(1 \sum_{i=1}^t y_i^2 - (\sum_{i=1}^t y_i)^2)} \quad (2)$$

式中: y_i 为第 i 个样本的真实值; $\hat{y}_i$ 为第 i 个样本的预测值。

(4)性能评价。利用模型返回的均方误差 MSE 和决定系数 R^2 , 可对所建立的 SVR 回归模型的性能进行评价。若均方误差 MSE 过大或者决定系数 R^2 过小,则可对模型参数进行修正或重新选择核函数类型,直到满足要求为止。

2 案例应用——以河南省新县为例

2.1 研究区概况

河南省新县位于河南省东南部大别山区,东经 $114^{\circ}33' \sim 115^{\circ}12'$,北纬 $31^{\circ}28' \sim 31^{\circ}46'$,总面积 $1\,612\text{ km}^2$,其中:淮河流域 $1\,307.5\text{ km}^2$,占总面积的 80.5%;长江流域为 304.5 km^2 ,占总面积的 19.5%。新县地处北亚热带与北暖温带的交界线,兼具夏季季风和副热带高压双重气候,年平均降水量 $1\,313.8\text{ mm}$ 。新县山洪灾害具有分布广、来势猛、流速快、破坏力大、突发性强的特点(见图 1)。据新县民政部门统计,2016~2018 三年来新县均遭受不同程度的山洪灾害,2016 年全县 14 个乡镇 89 个村告警,其中 7 个乡镇 43 个村达危险级别,7 个乡镇 46 个村达警戒级别^[16]。

2.2 指标体系及参数获取方法

临界雨量与地形因子及水文要素紧密相关,新县山洪灾害主要受到自然地理和暴雨特性影响;降雨是山洪灾害最直接的驱动力;地质情况通过影响土壤下渗能力来影响产流;地貌情况的变化直接影响洪水的

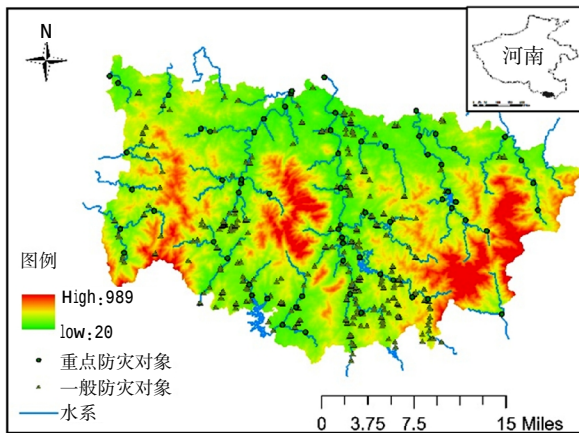


图1 新县地理位置及防灾对象分布

Fig.1 The geographical location and distribution of disaster prevention objects in Xinbian County

汇流路径及汇流时间；而人类活动通过改变下垫面条件间接影响山洪灾害。本文以河南省新县 161 个防灾对象为例^[17]，通过对山丘区小流域山洪灾害孕灾因子分析，进行指标选取和指标体系构建。

根据孕灾因子类型结合实际山洪灾害分析评价流程，考虑以下流域特征参数和产汇流参数。

(1)流域特征参数。流域特征参数主要包括防灾对象所在流域面积 F 、控制断面沟道长度 L 和沿河道平均比降 J 。流域面积 F 通常指流域分水岭所包围区域的平面投影面积，可利用 GIS 工具在流域 DEM 工作底图上提取流域边界后计算得到。沟道长度 L 指河源沿主河道至防灾对象所采用控制断面的距离，通过量算流域工作底图得到。沿河道平均比降 J 可根据流域的 DEM 进行提取。

(2)产汇流计算所涉及到的参数。各时段的雨量均值 $\bar{H}$ 。雨量均值反映了一个地区的降雨水平，可通过《河南省中小流域设计暴雨洪水图集（1984 年）》^[18] 的雨量均值等值线图查找得到。

变差系数 C_v 又称离差系数或离势系数，是衡量系列相对离差程度的一个参数。对水文现象来说，变差系数的大小反映了降雨在多年中的变化情况，可通过《河南省中小流域设计暴雨洪水图集（1984 年）》^[18] 的变差系数等值线图查得。

暴雨递减指数 n 。参考《山洪灾害分析评价方法指南》^[19]，暴雨参数计算方式见式(3)：

$$n_x = 1 + a \ln \left(\frac{H_{x1}}{H_{x2}} \right) \quad (3)$$

式中： a 为系数； n_x 为暴雨递减指数，只与时段雨量均

值 H 有关。由式(3)可得，暴雨递减指数 n 的影响可用雨量均值的影响来代替。

点面折减系数。求得设计点暴雨量后，利用点面折减系数可折算出设计面暴雨量。点面折减系数与流域面积和形状相关，流域面积小于一定数值时可取 1，即忽略点面折减系数的影响。

降雨损失参数 μ 为产流历时内流域的平均下渗率，可根据水文分区由相应的水文手册查得降雨损失参数 μ 的范围，再结合当地下垫面情况和流域特征参数综合确定。本研究通过降雨径流关系曲线图推求净雨，故可忽略降雨损失参数 μ 的影响。

综上，本文小流域山洪灾害临界雨量预估模型选取的指标主要包括流域面积 F 、河道长度 L 、比降 J 、雨量均值 $\bar{H}$ 、 C_v 值五个。本研究构建的模型及临界雨量预估流程如图 2 所示。

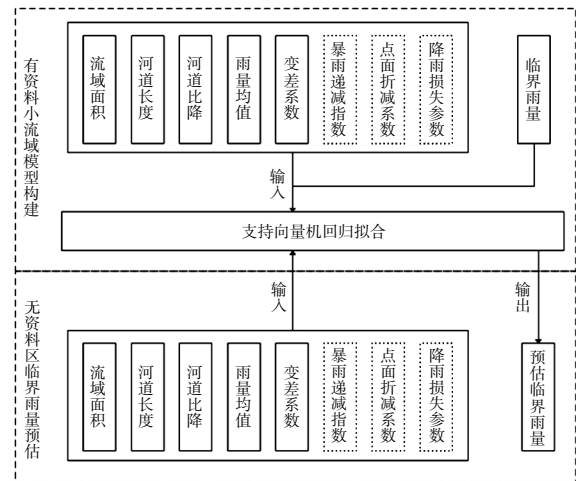


图2 模型构建及临界雨量预估流程

Fig.2 The flow of model construction and rainfall threshold estimation

2.3 模型训练与检验

参照《山洪灾害分析评价方法指南》^[19]，选取新县预警时段为 1、3、6h 的数据，分别进行建模。为简化前期影响雨量对临界雨量的影响，参照《河南省新县山洪灾害分析评价报告》^[17] 选取 0.2W M、0.5W M、0.8W M 三种土壤含水量进行建模，其中 W M=65mm 为土壤最大含水量，也即田间持水量。为排除偶然性，每个时段随机选取 151 个样本数据为训练集，剩下 10 个样本数据为测试集对模型性能进行检验。为避免过拟合现象对模型精度造成的影响，模拟中舍弃了训练集拟合度高而测试集拟合度不高的运行结果。

结果表明,三种前期土壤含水量情况下各时段运行出来的预估值与计算值均接近。图 3 和表 1 展示了 0.5W M 前期土壤含水量的某次运行结果。其中,真实值是通过对流域出口沟道断面进行实地测量,应用水文水力学法计算得到。同时,通过对比当地记载的实发洪水的数据和对受灾地区调查走访,对真实值进行检验复核,以确保真实值的可靠性。

由图 3 和表 1 可以看出,0.5W M 的前期影响雨量下 1h、3h、6h 训练集和测试集的拟合精度都较高,其中均方误差均小于 0.02,决定系数均大于 0.90。新县 161 个防灾对象模拟结果表明,所建立的 SVR 模型具有非常好的泛化能力,性能较好,仿真模拟精度高,可以应用到无资料或缺资料的相似小流域进行山洪灾害临界雨量的确定。

2.4 模型预测与结果分析

通过训练与检验,在模型的可靠度以及精度达到要求的情况下,将下垫面条件相似、产汇流机制相同的无资料或缺资料小流域防灾对象的相关指标代入模型,即可预估无实测资料小流域临界雨量值。新县山洪灾害分布范围广,许多山丘区无雨量站或水文站且沟道断面调查困难,因此本文应用模型进行临界雨量预估时仍以新县小流域为研究对象,具体为新县箭厂河乡、酒店乡、田铺乡共 10 个防灾对象。

通过新县 DEM 数据提取以及《河南省中小流域设计暴雨洪水图集》^[19]暴雨参数查询,得到 10 个防灾对象的基本信息(见表 2)。

将 10 个防灾对象的基本信息代入模型,即可得临界雨量的预估值,这里以 0.5W M 前期影响雨量为例进行防灾对象临界雨量预估。为了进一步验证模型的适用性及精度,针对 10 个防灾对象进行沟道断面的实地测量,通过水文水力学法进行产汇流计算

以及水位流量关系推求得到临界雨量真实值,对真实值进行检验复核后与预估值进行比较,结果如表 3 所示。

由表 3 可知,10 个防灾对象中 1h、3h、6h 临界雨量预估值与真实值的最大偏差绝对值,分别为 1.93%、2.60%和 1.46%,通过 SVR 模型预估的临界雨量值与真实值偏差较小,预估结果可以作为无资料区山洪灾害雨量预警指标,为当地防洪防灾预警及决策提供参考。

表1 训练集和测试集的均方误差和决定系数
Table1 The mean square errors and determination coefficients of training set and test set

时段	训练集		测试集	
	均方误差	决定系数	均方误差	决定系数
1h	0.0015	0.9922	0.0066	0.9503
3h	0.0009	0.9953	0.0158	0.9432
6h	0.0011	0.9926	0.0040	0.9576

表2 10个防灾对象基本信息
Table2 The basic information of 10 disaster prevention objects

防灾对象	流域面积 F/km ²	河道长度 L/m	河道比降 J	雨量均值	变差系数 C _v
杨畈组	22.204	7289	0.02	46	0.409
北河组	56.615	14804	0.02	46	0.406
范西组	56.615	14804	0.02	46	0.406
五房河组	16.225	8144	0.04	46	0.405
余河组	15.092	7220	0.04	46	0.405
边店组	10.318	6025	0.04	46	0.408
草河组	10.633	6286	0.04	46	0.408
泗湾组	59.606	16337	0.01	46	0.406
蔡边组	59.606	16337	0.01	46	0.406
夏湾组	21.568	10499	0.06	45	0.403

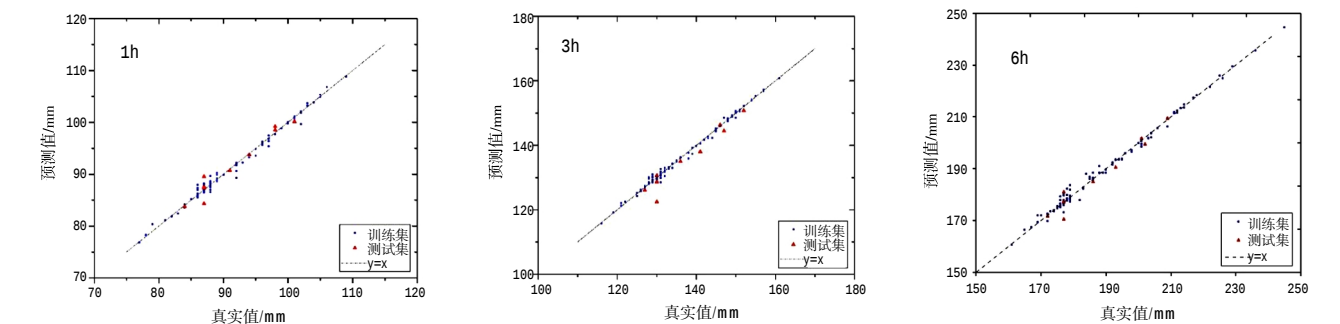


图 3 0.5W M 训练集和测试集的预测结果对比

Fig.3 The comparison of prediction results between 0.5W M training set and test set

表3 10个防灾对象0.5W M预测值与真实值对比分析

Table3 The comparison between the predicted value and real value in 0.5W M for 10 disaster prevention objects

防灾对象	1h			3h			6h		
	预测值 /mm	真实值 /mm	偏差绝对值	预测值 /mm	真实值 /mm	偏差绝对值	预测值 /mm	真实值 /mm	偏差绝对值
杨岷组	94.00	93.85	0.16%	141.00	137.34	2.60%	193.00	190.19	1.46%
北河组	85.00	83.43	1.84%	128.00	126.86	0.89%	175.00	173.85	0.66%
范西组	85.00	83.43	1.84%	128.00	126.86	0.89%	175.00	173.85	0.66%
五房河组	97.00	97.76	0.78%	147.00	147.84	0.57%	204.00	204.57	0.28%
余河组	85.00	85.64	0.75%	129.00	129.61	0.48%	177.00	176.57	0.24%
边店组	86.00	87.22	1.42%	130.00	129.56	0.34%	177.00	176.94	0.03%
草河组	86.00	87.24	1.44%	130.00	129.62	0.29%	177.00	177.03	0.01%
泗湾组	85.00	83.36	1.93%	128.00	125.45	1.99%	175.00	173.69	0.75%
蔡边组	85.00	83.36	1.93%	128.00	125.45	1.99%	175.00	173.69	0.75%
夏湾组	84.00	84.23	0.27%	128.00	127.43	0.45%	177.00	175.59	0.80%

3 结论

我国小流域数量众多,全面测量掌握各小流域预警山洪灾害的必要资料,将耗费大量的人力物力,可操作性不强。临界雨量是山洪灾害预警预报的基础性指标,在雨量站布设稀疏且小流域防灾对象无野外调查资料的情况下,探索出简便可行的模型方法来对小流域防灾对象临界雨量进行预估,将为无资料或缺资料小流域山洪灾害预警预报工作提供较大便利。本文结合推求临界雨量的实际问题,对孕灾因子指标选取和量化,根据有资料小流域山洪灾害孕灾因子与临界雨量之间的相关关系,探索构建了基于 SVR 的无实测资料小流域山洪灾害临界雨量预估模型及方法,利用均方误差和决定系数来判断模型方法的可行性,并以经检验的模型推求无实测资料小流域山洪灾害临界雨量。结果表明,基于 SVR 的无实测资料小流域山洪灾害临界雨量预估模型合理可行、简单有效,通过 SVR 模型计算得出的临界雨量值可以作为山洪灾害预警预报指标,且基于 SVR 的无实测资料小流域防灾对象山洪灾害临界雨量预估方法,既基于一定的物理基础,又有一定的统计学意义。采用该模型方法既可省去小流域产流和汇流的复杂计算,简便高效,又可大大减少繁重的野外调查测量,为无实测资料区小流域临界雨量的确定提供了新的思路和方法。但需要指出的是,此方法的应用需要借助邻近有资料小流域的临界雨量成果。同时在模型模拟过程中可能出现过度拟合现象,如何避免过拟合现象的产生仍需进一步的探索研究。

参考文献:

- [1] 赵刚, 庞博, 徐宗学, 等. 中国山洪灾害危险性评价 [J]. 水利学报, 2016,47(9):1133-1142. (ZHAO Gang, PANG Bo, XU Zongxue, et al. Assessment on the hazard of flash flood disasters in China [J]. Journal of Hydraulic Engineering, 2016,47(9):1133-1142. (in Chinese))
- [2] 郭良, 丁留谦, 孙东亚, 等. 中国山洪灾害防御关键技术[J]. 水利学报, 2018,49(9):1123-1136. (GUO Liang, DING Liuqian, SUN Dongya, et al. Key techniques of flash flood disaster prevention in China [J]. Journal of Hydraulic Engineering, 2018,49(9):1123-1136. (in Chinese))
- [3] 王文川, 和吉, 邱林. 我国山洪灾害防治技术研究综述[J]. 中国水利, 2011,(13):35-37. (WANG Wenchuan, HE Ji, QIU Lin. Review on control techniques of mountain torrent disaster in China [J]. China Water Resources, 2011,(13):35-37. (in Chinese))
- [4] Miao Q, Yang D, Yang H. Establishing a rainfall threshold for flash flood warnings in China's mountainous areas based on a distributed hydrological model [J]. Journal of Hydrology, 2016,541: 371-386.
- [5] 原文林, 高倩雨, 张晓蕾. 基于成灾信息的山洪灾害临界雨量综合检验复核方法[J]. 水电能源科学, 2018,36(7):47-51. (YUAN Wenlin, GAO Qianyu, ZHANG Xiaolei. Comprehensive assessment and recheck method of rainfall threshold for flash flood based on disaster information [J]. Water Resources and Power, 2018,36(7):47-51. (in Chinese))
- [6] Hapuarachchi H A P, Wang Q J, Pagano T C. A review of advances in flash flood forecasting [J]. Hydrological Processes, 2011,25(18):2771-2784.
- [7] 李昌志, 郭良. 山洪临界雨量确定方法述评 [J]. 中国防汛抗旱, 2013,23(6):23-28. (LI Changzhi, GUO Liang. Methods of rainfall indicator for flash flood [J]. China Flood & Drought Management, 2013,23(6):23-28. (in Chinese))
- [8] 程卫帅. 山洪灾害临界雨量研究综述 [J]. 水科学进展, 2013,24(6): 901-908. (CHENG Weishuai. A review of rainfall thresholds for triggering flash floods [J]. Advances in Water Science, 2013,24(6): 901-908. (in Chinese))
- [9] 吴承君. 小流域山洪灾害临界雨量检验复核及区域预警指标估算

- 方法研究 [D]. 郑州: 郑州大学, 2017. (WU Chengjun. Critical Rainfall Inspection of Flash Flood Disaster in Small Watershed and Study of Regional Early-warning Indicators Estimation Method [D]. Zhengzhou: Zhengzhou University, 2017. (in Chinese))
- [10] 樊建勇, 单九生, 管珉, 等. 江西省小流域山洪灾害临界雨量计算分析 [J]. 气象, 2012, 38 (9): 1110-1114. (FAN Jianyong, SHAN Jiusheng, GUAN Min, et al. Research on analysis and calculation method of critical precipitation of mountain torrents in Jiangxi province [J]. Meteorological Monthly, 2012, 38 (9): 1110-1114. (in Chinese))
- [11] Vladimir C, Ma Y. Practical selection of SVM parameters and noise estimation for SVM regression [J]. Neural Networks, 2004, 17 (1): 113-126
- [12] 刘靖旭. 支持向量回归的模型选择及应用研究 [D]. 长沙: 国防科学技术大学, 2006. (LIU Jingxu. Research on Model Selection for Support Vector Regression and Application of It [D]. Changsha: National University of Defense Technology, 2006. (in Chinese))
- [13] 周秀平, 王文圣, 黄伟军. 支持向量机回归模型在径流预测中的应用 [J]. 水电能源科学, 2006, (4): 4-7. (ZHOU Xiuping, WANG Wensheng, HUANG Weijun. Application of support vector machine regression in runoff forecast [J]. Water Resources and Power, 2006, (4): 4-7. (in Chinese))
- [14] 李波, 顾冲时, 李智录, 等. 基于偏最小二乘回归和最小二乘支持向量机的大坝渗流监控模型 [J]. 水利学报, 2008, 39 (12): 1390-1394. (LI Bo, GU Chongshi, LI Zhilu, et al. Monitoring model for dam seepage based on partial least-squares regression and partial least square support vector machine [J]. Journal of Hydraulic Engineering, 2008, 39 (12): 1390-1394. (in Chinese))
- [15] 史峰, 王辉, 郁磊, 等. MATLAB 智能算法 30 个案例分析 [M]. 北京: 北京航空航天大学出版社, 2011. (SHI Feng, WANG Hui, YU Lei, et al. Analysis of 30 Cases of Intelligent Algorithms in MATLAB [M]. Beijing: Beijing University of Aeronautics and Astronautics Press, 2011. (in Chinese))
- [16] 郑州大学综合设计研究院有限公司. 新县分析评价成果检验与复核报告 [R]. 2016. (Zhengzhou University Multi-functional Design and Research Academy Co. Ltd. Report on the inspection and review of the results of analysis and evaluation in Xinxian county [R]. 2016. (in Chinese))
- [17] 河南黄河水文勘测设计院. 河南省新县山洪灾害分析评价报告 [R]. 2015. (Henan Yellow River Hydrological Survey and Design Institute. Analysis and evaluation report of mountain flood disaster in Xinxian county, Henan province [R]. 2015. (in Chinese))
- [18] 河南省水利厅. 河南省中小流域设计暴雨洪水图集 [R]. 1984. (Water Resource Department of Henan Province. Atlas of design storm flood in medium and small river basins in Henan Province [R]. 1984. (in Chinese))
- [19] 全国山洪灾害防治项目组. 山洪灾害分析评价方法指南 [Z]. 2015. (National Project Group for Mountain Flood Disaster Prevention and Control. Guidelines for the analysis and evaluation of mountain flood disasters [Z]. 2015. (in Chinese))

Threshold Rainfall Prediction Model for Flash Flood Disaster in Ungauged Small Watersheds Based on SVR: A Study Case in Xinxian County of Henan Province

WANG Yanyun^{1,2}, YUAN Wenlin², LONG Aihua¹, XIANG liyun¹, DENG Xiaoya¹, FU Lei², SONG Hanzhen²

(1. State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing 100038, China; 2. School of Water Conservancy and Environment, Zhengzhou University, Zhengzhou 450001, China)

Abstract: Flash flood disaster is a prominent problem to be solved in disaster prevention and mitigation in China. The rainfall threshold is one of the important indicators of flash flood disaster forecasting. The traditional calculation methods have their own advantages and different requirements for data. For some ungauged area, however, data are hard to collect. There are few studies on the calculation of rainfall in disaster prevention objects in small watersheds without measured data. In view of the condition mentioned above, based on the support vector machine regression fitting method (SVR), this paper established a model to deduce the rainfall threshold of flash flood disaster in small watersheds without data by using the comparison method. The results show that: (1) the SVR model has good generalization, high prediction accuracy, small error and better forecasting; (2) the difference between modeled rainfall threshold and real value is small and this model could be used as an early warning and forecasting of flash flood disaster. Additionally, it can also provide reference for departments of local decision-making.

Key words: ungauged small watershed; flash flood disaster; rainfall threshold; risk factor; support vector machine regression algorithm