

沈军, 聂作先, 吴贤云, 等. 2016. 基于隐马尔可夫模型的中小河流致灾雨量阈值研究. 气象, 42(7): 865-874.

# 基于隐马尔可夫模型的中小河流 致灾雨量阈值研究<sup>\* 1</sup>

沈 军<sup>1,2,5</sup> 聂作先<sup>3</sup> 吴贤云<sup>1</sup> 郭海峰<sup>4</sup>

1 湖南省气候中心, 长沙 410118

2 湖南省气象技术装备中心, 长沙 410007

3 福建工程学院福建省数字化装备重点实验室, 福建 350118

4 湖南省气象科技服务中心, 长沙 410118

5 湖南省气象灾害省重点实验室, 长沙 410118

**提 要:** 基于数字高程模型(digital elevation model, DEM)高程数据和 D8 算法实现了湖南中小河流流域子流域划分。利用洪水预报网络传播模型, 构建中小河流各子流域径流拓扑结构并提出了空间子流域等效面雨量的概念。基于等效面雨量序列和径流观测序列, 构建了基于时空概念的隐马尔可夫降水-径流模型, 并利用该模型计算了中小河流各子流域不同时间尺度条件下暴雨致灾临界雨量阈值, 最后利用 2010—2015 年实况资料对阈值进行检验, 结果证明相比传统统计法, 新方法的计算结果与传统方法结果一致且具有很好的准确性和稳定性。基于 2015 年 6 月的一次暴雨灾害预报, 证明了该方法适应于业务化运行。

**关键词:** 中小河流, 暴雨致灾, 隐马尔可夫模型, 降水致灾阈值

**中图分类号:** P641, P49

**文献标志码:** A

**doi:** 10.7519/j.issn.1000-0526.2016.07.010

## Research of Critical Precipitation Threshold in Small Rivers Based on Hidden Markov Model

SHEN Jun<sup>1,2,5</sup> NIE Zuoxian<sup>3</sup> WU Xianyun<sup>1</sup> GUO Haifeng<sup>4</sup>

1 Hunan Climate Centre, Changsha 410118

2 Meteorological Technology Equip Center of Hunan, Changsha 410007

3 Fujian Provincial Key Laboratory of Digital Equipment, Fujian University of Technology, Fujian 350118

4 Hunan Meteorological Public Service Centre, Changsha 410118

5 Hunan Key Laboratory of Meteorological Disaster, Changsha 410118

**Abstract:** In this paper, sub-catchments of small rivers in Hunan are determined based on DEM data and D8 algorithm. According to the network propagation model for flood forecast of main rivers in Hunan Province, a directed graph related to the sub-catchments has been built, which represents the topological relationship among these areas both in temporal and spatial dimension. To build the relation between precipitation and runoff in a reasonable space, the concept of valid rain-area related to the flood graph is proposed. By using the valid rain-area data series and runoff series, the hidden Markov model is built to calculate critical rain-area threshold of flood disasters at different time scales in the sub-catchments in Hunan Province. Finally, the observed flood warning data (collected in 2010—2015) are used to test the proposed method. The results show that the new method has similar performance to traditional statistical ways but

\* 国家自然科学基金项目(41305056)、国家重点基础研究发展计划(973 计划)(2012CB955901)和福建省教育厅科技项目(JB11115)共同资助  
2015 年 1 月 26 日收稿; 2016 年 5 月 20 日收修定稿  
第一作者: 沈军, 主要从事数据处理、灾害风险区划和气候影响评价方面研究. Email: vafj5@163.com

is more accurate and stable. In addition, the successful application of the new method to the operation of forecasting flood disaster on 13 June 2015 further proved that the new method is suitable for flood forecasting.

**Key words:** small river, rainfall disaster, hidden Markov module, threshold of precipitation

## 引 言

中小河流是指流域面积在 200~2000 km<sup>2</sup> 的河流,该类河流网络覆盖了我国 85% 的城镇和农村地区。区域复杂性和时空变化多样性使得对中小河流子流域的识别存在较大难度。此外,中小河流洪水破坏强度大、致灾时间短和爆发性强等特点使得中小河流暴雨致灾预警预报成为当前各国政府和防灾服务领域关注的重点,其中,由短时强对流暴雨产生的中小河流灾害是近来气象灾害防御研究的方向,也是最近国内外研究的热点和难点。在国内,气象、水文等部门均开展了中小河流治理与洪涝灾害防御研究工作。研究中小河流暴雨致灾阈值是灾害管理的重要内容之一。暴雨致灾阈值研究的首要任务是对降水与径流的雨洪关系进行相关性分析,最终建立基于降水量的气象灾害预警机制。在水文模型的灾害预警研究领域,意大利 proGEA 公司基于 TOPKAPI 分布式水文模型和 GIS 技术开发了中小河流预警预报系统并得到推广和应用(刘志雨, 2005);希腊 Grillakis 等(2010)利用 HBV 模型和降水雷达短时降水预报对斯洛文尼亚热莱兹尼基(Zelezniki)区域的中小河流进行洪水预报,得到了很好的预报结果,美国国家气象服务水文管理中心采用 Sacramento Soil Moisture Accounting (SAC-SMA) 模型对全美 4000 个水文观测点进行逐日的水文预报,通过复杂的参数订正过程后取得较好效果(Smith et al, 2013);美国军方工程集团水文工程中心基于地理信息完成了分布式水文模型-HEC-HMS (Hydrologic Engineering Center\_hydrologic modeling system) 的构建受到国内水文领域学者的关注,如林峰等(2009)利用该模型建立了福建省晋江西流域降雨径流模型并研究了不同时间尺度条件下模拟步长对模型运算的影响。基于地理信息技术的中小河流暴雨致灾研究在我国起步较晚,水利部水文局 2010 年开展了中小河流山洪监测与预报技术研究。该课题系统性地研究了湖南湘江支流沅水流域和江西赣江支流遂川江流域暴雨灾害防御,

但研究重点在于中小河流站网监测、水文预测和风险预警等规划工作[水利部水文局(水利信息中心), 2010],对降水与洪水致灾等相关课题的研究较少。气象部门也开展了相关研究,文明章等(2013)利用 Floodarea 水动力模型对给定流域进行淹没模拟,实现淹没过程再现,同时推算出给定点的临界雨量阈值,然而该方法涉及到高精度的高程数据,需要大量的计算资源,对中小河流流域阈值计算能力有限;当前,中小河流暴雨致灾研究由于资料匮乏、河道变化复杂及人工调蓄工程等因素的影响使得实用成果有限,业务应用与支撑系统更少。强降水是中小河流径流变化致灾的主要原因,然而,受土地类型、植被指数、河道糙率和水资源调蓄等多种因素的影响,降水与径流并非简单的线性关系,而是以一定非线性关系存在。此外,某一流域径流变化不仅受本流域面雨量变化的影响,而且受上游流域径流变化的影响。基于数字高程模型(digital elevation model, DEM)高程数据和 D8 算法,本文实现了与湖南水利观测站网一致的中小河流流域划分,形成了湖南中小河流流域集合,利用湖南“四水”流域洪水传播时序关系,构建了各中小河流子流域相互影响的有向拓扑图,描述了时空范围内各子流域间的相互关系并提出了中小河流流域等效面雨量的概念和计算算法;利用 40 年逐日降水观测资料和等效面雨量计算算法得到了各子流域的等效面雨量序列,结合子流域水文控制站观测序列,实现了用于暴雨致灾阈值分析的隐马尔可夫模型(许博等, 2012)的构建。利用水文控制站相关参数和隐马尔可夫模型计算了中小河流洪水灾情发生前 1~5 d 的累计面雨量序列概率,从而实现了不同时间尺度条件下的各子流域暴雨致灾的临界雨量阈值的确定过程。

## 1 数据与方法

本文研究空间范围为湖南(25.0°~30.0°N, 108.8°~114.3°E)主要降水汇流区域。研究对象为降水条件下湖南湘江、资江、沅江、澧水四大流域所属中小河流各子流域的暴雨致灾临界雨量确定问

题。文中采用的水文观测资料为湖南省水利局整理的 1965—1980 年 87 个水文控制站汛期(4—9 月)逐日的水文观测资料,降水数据为国家气象信息中心整理下发的 1970—2000 年逐日降水的历史整编资料。考虑到水文观测资料的完整性结合中国气象局“山洪暴雨风险普查工作实施方案”<sup>①</sup>的相关要求,选用 60 个与中小河流流域相配的水文站观测资料为研究数据。考虑中小河流降水汇流的空间影响,结合湖南洪水传播时间时序关系并基于本文提出的中小河流等效面雨量的概念和相应的计算算法,生成了与中小河流流域相匹配的从 1965—1980 年的逐日等效流域面雨量序列。

中小河流暴雨致灾临界雨量阈值确定方法是近些年来气象防灾减灾研究的新内容,基于模式和统计等多种方法不断地被得到完善和修正,在总结中国气象局山洪普查与阈值确定工作的基础上,本文将隐马尔可夫模型用于暴雨致灾雨量的确定工作,主要采用如下方法:(1)基于 D8 算法利用 GIS 工具实现基于 DEM 高程数据的子流域划分;(2)基于国家级观测站和国家基准站的逐日降水观测数据,使用 Kriging 插值方法构建流域面雨量数据集;(3)有向图递归算法:利用洪水传播时间表,结合等效流域面雨量,构建流域洪水传播有向图,使用有向图递归算法计算中小河流各子流域等效面雨量;(4)分类方法:为构建合理的隐马尔可夫模型状态集,本文针对不同特征的数据序列,采用不同的分类方法,对等效面雨量序列(离散性明显),采用自适应 K-均值分类方法,对水文观测序列(有一定连续性且变化区间较小)采用直方图分类法;(5)期望值最大化算法(expectation maximization algorithm, EM):利用降水资料和水文观测资料形成了隐马尔可夫模型的系统状态集合和观测状态集合以及状态转换概率矩阵,实现了隐马尔可夫模型的构建,基于水文观测状态序列和 EM 方法,求得达到灾害临界(或警戒水位)条件下,最大概率的降水状态序列最优参数估计值和相应的概率值。

## 2 湖南水文气候状况

湖南为大陆性热带季风湿润气候区域,年平均

降水量为 1403.1 mm,降水时空特征明显,每年 4—9 月为集中降水期,其中,从 4 月开始,湖南降水量经历两个时间剧增(4 月第 2 至第 4 候、5 月第 1 候至 6 月第 5 候),最大候平均降水量为 49 mm(吴贤云等,2015),段亚雯等(2014)利用 1961—2010 年逐日降水、气温观测资料,对中国典型区域进行了降水集中指数(PCI)研究,证明了长江中下游 PCI 指数在 12~17 变化,降水集中在 6—7 月。根据湖南极端气候事件监测系统显示,湖南 1 d 极端降水最大值为 455.5 mm,连续 2 d 最大降水量为 541.4 mm,连续 3 d 最大降水量为 619.7 mm,连续 5 d 最大降水量为 623.1 mm,强降水集中期、独特的地形地貌和复杂的中小河流特征导致湖南中小河流汛期暴雨灾害频发,灾情损失严重。仅 2014 年,因洪水灾害造成的社会损失占湖南自然灾害损失总量 94.2%,暴雨灾害共导致 53 人死亡,直接受灾人口 153.3 万人<sup>②</sup>。

另一方面,受能源需求的影响,湖南拥有水库 14121 座,总库容量为 530.7 亿  $\text{m}^3$ ,水电站 4467 座(装机容量为 1534.8 万 kW),总水利工程规模位居全国首位(湖南省水利厅等,2013)。系列的调蓄工程破坏了自然条件下的降水-径流关系,增加了基于水文模型的开展暴雨阈值研究的复杂性和暴雨致灾阈值确定的不确定性。

## 3 基本原理与理论

对某一研究流域,其日流域面雨量( $\mathbf{R}$ )可分布在  $k$  个不同的区间内,每个区间出现的概率为  $\boldsymbol{\pi}_i$ [即  $P(\mathbf{R}_{i=1,2,\dots,k}) = \boldsymbol{\pi}_i$ ],在观测时间  $t$  内( $t$  为出现暴雨灾害前连续降水日数),该流域水文控制站观测观测序列为  $l_t, l_{t-1}, \dots, l_1$ 、同步的降水序列为  $o_t, o_{t-1}, \dots, o_1$ ,日降水量  $o_k$  条件下产生径流  $l_t$  的概率为  $\mathbf{b}_{k,t}$  ( $\mathbf{b}_{k,t} \in \mathbf{B}$ ),降水量  $o_t$  转变成  $o_k$  的概率为  $\alpha_{t,k}$  ( $\alpha_{t,k} \in \mathbf{A}$ )。则,对一给定的水文观测序列  $l_t, l_{t-1}, \dots, l_1$ ,基于降水转换概率矩阵  $\mathbf{A}$ ,径流状态转换概率矩阵  $\mathbf{B}$  和初始降水概率向量  $\boldsymbol{\pi}$  的四元组条件  $\boldsymbol{\lambda}$  下[令  $\boldsymbol{\lambda} = (\mathbf{L}, \mathbf{O}, \mathbf{A}, \mathbf{B}, \boldsymbol{\pi})$ ],利用式(1)得到有效面雨量序列  $o_t, o_{t-1}, \dots, o_1$  的概率。在给定的径流序列条件下,若存在多个  $o_t, o_{t-1}, \dots, o_1$  观测序列与其对

<sup>①</sup> 国家气候中心. 2013. 暴雨洪涝灾害致灾临界(面)雨量确定技术指南.

<sup>②</sup> 湖南省民政厅. 2014. 2014 年湖南省自然灾害年度报告.

应,那么取概率最大的序列为灾害发生条件下对应的雨量序列,该序列对应的降水累计值为灾害发生条件下的暴雨致灾面雨量阈值。

$$P(\mathbf{O}/\lambda) = \sum P(\mathbf{O}, \mathbf{L}/\lambda) = \sum \pi_{i_1} b_{i_1}(o_1) a_{i_1 l_2} b_{i_2}(o_2) a_{i_2 l_3} \cdots a_{i_{t-1}} b_{i_t}(o_t) \quad (1)$$

满足:

$$a_t(i) = p(o_1, o_2, \dots, o_t = i | \lambda) \text{ 且 } a_1(i) = P(o_1 = i, l_1 | \lambda)$$

式中,  $a_t(i)$  表示在概率模型  $\lambda$  条件下, 观测值  $o_i$  在  $t$  时刻等于  $i$  的概率。

## 4 基于 HMM 模型临界雨量分析

### 4.1 基于 DEM 高程数据的中小河流流域划分

中小河流临界雨量阈值研究的本质是研究中小河流域内降水形成的径流与本流域内面雨量的关系。因此,中小河流流域边界的确定和划分是研究降水阈值的基础。中小河流洪水来源于相关的山洪沟流域,流域划分应与降水汇流的物理过程一致同时也需要防止流域内因高程差异带来的边界不连续性(孤岛现象)情况。此外,受区域降水空间差异的影响、河道断面特征、河道长度、河沟坡降和流域地表持水率的不同,相同的降水量也会产生不同的径流量。此外,中小河流流域划分也需要兼顾水文观测站网的分布,尽可能保证每个流域内至少有一个水文观测站,以确保每个研究流域有可靠的水文观测数据支持。本文以 1:50000 的 DEM 高程数据为数据源,根据湖南水文观测站分布情况,利用汇流模拟 D8 算法(赵健等,2006)和 ArcGIS 的 hydrological tools 工具对目标区域流域进行划分,得到全



图 1 基于 DEM 高程数据、D8 算法和水文观测站网生成的湖南省中小河流流域  
Fig. 1 Distribution of small rivers in Hunan based on D8 algorithm, DEM and hydrological observation

省中小河流流域子流域划分(图 1)所示。图 1 中,红色多边形为基于 DEM 进行流域分割的结果,绿色圆点为基于研究流域内主要河流主干流域监测的水文控制站,可以看出本文流域划分结果与水文观测站网基本一致。

### 4.2 子流域面雨量的计算

为了得到合理的流域面雨量,采用地质统计分析法(刘爱利等,2012)对日降水量进行空间插值,形成流域面雨量序列。为得到适合湖南省降水条件下 Kriging 插值的最优参数,本文通过计算 30 年平均降水量的空间分布,得到各雨量站间的变异数;利用线性变差模型、指数变差模型、球面变差模型和高斯变差模型对平均雨量进行分析,并调节相应块金方差获得稳定性的最优值,得到的分析结果如表 1 所示。

表 1 为在不同插值模型条件下获得的相关参

表 1 不同参数条件下运用 kriging 方法对降水数据的插值结果

Table 1 Performance of interpolation of rainfall data based on Kriging with different parameters

模型类型	拟合误差	块金	偏基台值	变程	偏基台值/基台值
线性模型	447	21.35	25.41	2.7559	0.208
球面模型	465	7.338	23.56	0.15	0.533
指数模型	465	3.9	23.5	0.009	0.834
高斯模型	465	4.868	23.5	0.0416	0.83

数。可以得知当 Kriging 模型选择指数模型时,块金值最低且模拟稳定性最好(偏基台值/基台值=0.834)。因此,利用与指数模型对应的最优块金值和偏基台值参数,本文对空间域站点日降水数据序

列进行 Kriging 插值,得到子流域逐日面雨量序列。

### 4.3 基于流域面雨量构建等效流域网络

中小河流流域水情变化不仅受本流域内降水影

响,同时受到上游流域来水的影响。因此,需要从时间-空间的二维空间上考虑降水对给定流域径流的影响。邓柯(2004)和叶泽纲等(2002)研究“湖南四水”内主要子流域在汛期期间的洪水传播过程

(表 2),该模型考虑了河道长度、河道比降、河道植被和河道糙率等地理条件下中小河流河道洪水传播的速度和子流域间的传播关系,是真实条件下降水汇流方向和汇流时间的经验模型。

表 2 主要流域洪水传播时间表

Table 2 Flood propagation time of main catchments in Hunan

湘江流域			沅江流域			资江流域			澧水流域		
河段	河长 /km	传播时间 /h	河段	河长 /km	传播时间 /h	河段	河长 /km	传播时间 /h	河段	河长 /km	传播时间 /h
全州—老埠头	100	13	黔城—安江	35	7	罗家庙—邵阳	53	7	南岔—张家界	73	9
老埠头—归阳	141	20	安江—普市	141	17	罗阳—冷水江	78	8	张家界—石门	143	12
归阳—衡阳	140	20	普市—沅陵	80	5	冷水江—新化	29	8	石门—津市	64	11
衡阳—衡山	67	10	沅陵—桃源	30	9	新化—柘溪	121	19	长潭河—石门	53	10
衡山—株洲	135	16	桃源—常德	49	7	柘溪—桃江	140	17	皂市—石门	14	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
长沙—湘阴	72	24	陶伊—普市	33	3	益阳—城陵矶	166	58			

根据洪水传播表和中小河流流域划分的结果,构建各子流域之间的传播关系(洪水传播有向网络图)(图 2),其中节点号代表流域号,权值为洪水自上游到下游的平均经历时间。此外,为研究空间域各子流域降水对径流影响,本文提出了等效流域和等效流域面雨量的概念。

定义 1:等效流域( $C_i$ ):一个流域集合,该集合内任意流域的面雨量产生的径流在相同时间段  $U_i$  内可以通过固定参考节点;

令  $t_{i,j}$  为流域  $i$  的径流通过固定河道流经流域  $j$  所需时间,则:  $C_i = \{t_k \mid t_{k,j} < U_j, \forall k, i \in D\}$

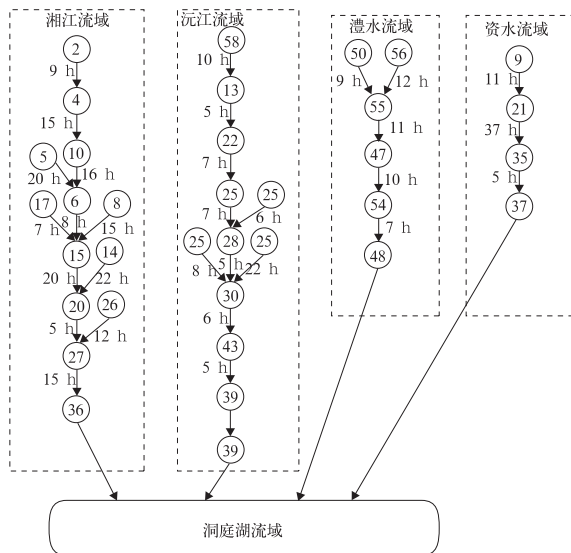


图 2 湖南“四水”洪水传播有向图  
Fig. 2 Directed flood propagation graph of the four rivers in Hunan

定义 2:等效流域面雨量( $R_i$ ):单位时间内等效流域  $C_i$  内的单位面积的降水量。

对一给定的中小河流流域,流域等效面雨量是本流域内径流产生的主要来源,为便于动态计算所有子流域等效面雨量,基于有向拓扑图,提出了求解中小河流各子流域的等效流域面雨量的递归算法(图 3)。

```

Algorithm: Catch Rainfall(i, L, R, t)
{
// 算法说明:利用递归算法计算流域 时间 t 的等效流域面雨量
// Input: 有向图的邻接矩阵  $(G_{i,j} \mid i, j \in N)$ , 其中:  $G_{i,j}$  表示自然条件下河道洪水由流域  $i$  至流域  $j$  所需要的时间;  $N$  为子流域个数; 子流域面雨量集合  $\{R_{i,j} \mid i \in N\}$ ; 单位时间内子流域  $i$  的在时间为  $t$  时间的面雨量。  $Tmp24, tmp48$  为临时数组变量,记录与待处理流域相关的固定时间内  $\zeta_r$  (为研究方便本例取  $\zeta_r = 24$  小时)内上游临界流域和  $2 * \zeta_r$  (48 小时)内上游河流关联流域;  $A$  为流域面积数组;
// Output: 每个节点时间  $t$  的等效流域面雨量  $L_{r,t} (r \in N)$ 。

tmp24 = null; tmp48 = null
Step1: 如果存在对任意整数  $k(k \in N)$ , 均有  $G_{k,i} = \infty$ , 则  $L_{i,t} = R_{i,t}$ , 返回  $L_{i,t}$ ; 否则, 进入 Step2;
Step2: 若  $G_{k,i} < 24$  且  $L_k$  为无穷大, 则
{
 $L_{k,t} = CatchRainfall(k, G, L, R, t)$ ; // 递归遍历计算上游临界流域值

tmp24 = tmp24 + {k};
}
否则若  $\zeta_r \leq G_{k,i} < 2 * \zeta_r$ 
{
tmp24 = tmp24 + {k};
tmp48 = tmp48 + {k};
}
Step3: 对集合 tmp48 的每一个元素, 计算与该元素对应的  $t-1$  天的子流域面雨量; 对集合 tmp24 的每一个元素, 计算与该元素对应的  $t$  天的等效流域面雨量, 并将结果置于集合 tmp24 中;
Step4: 基于数据集 tmp24 和流域面积数组  $A$ , 利用泰森多边形法则计算  $L_{r,t}$ , 返回  $L_{r,t}$ 
}

```

图 3 基于有向图的有效流域面雨量的递归算法  
Fig. 3 Recursive algorithm to calculate valid area rainfall of catchment based on directed graph



#### 4.4 基于隐马尔可夫模型的等效流域临界雨量阈值确定

##### 4.4.1 状态的分类

为构建合理的隐马尔可夫模型,需要对观测序列进行分类以构建系统状态集和观测状态集。受天气过程的影响,某一站点日降水序列在时间维上表现“聚集性”特征。Sivakumar 等(2009)证明了日降水数据为混沌序列。然而相对于日降水序列,某一

站点的径流数据序列和水位序列均有一定的连续性;因此,对降水序列和水位序列,本文采用了不同的分类方法。对降水序列采用 K-均值分析法,而对水位序列采用直方图统计分析法结合曲线拟合求拐点的分类方法(简称直方图分类法)(图 4)。图 4a 为运用 K-均值方法对某一雨量数据序列进行 5 个区间的分类,图 4b 为基于直方图分析法对水位数据序列进行分析,并利用曲线拟合求拐点方法将该数据进行分类。

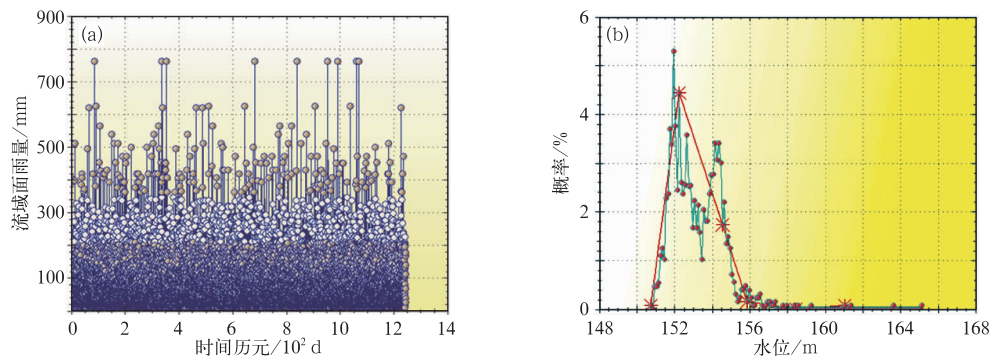


图 4 基于 K-均值分析法的降水序列分类(a)和基于直方图的水位数据序列分类(b)

Fig. 4 Classification of rainfall series based on K-means (a) and classification of water-level data series based on histogram (b)

##### 4.4.2 基于隐马尔可夫模型的临界雨量阈值确定

对每个中小河流子流域,利用历史观测资料和水文同步观测资料,构建隐马尔可夫模型的降水状态转换矩阵  $O$ ,径流状态转换概率矩阵  $A$  和降水-径流转换矩阵  $B$  和径流初始状态向量  $\pi$ 。对各水文控制站出现灾害前  $N$  天降水序列构建径流状态序列,选择概率最大的观测序列为预警条件下的径流序列,基于隐马尔可夫模型并利用 EM 算法求得在给定径流序列条件下对应的最大流域面雨量序列和相应的累计面雨量值。

## 5 检验与应用

在实际应用过程中由于灾害预警直接与水位关联,且每个水文控制站均有降水-径流关系曲线图,为方便中小河流洪水预警业务运行同时降低阈值确定的中间变换过程,用水位替代径流,直接建立降水和水位的隐马尔可夫模型。本文采用的研究流域为图 1 中第 10 号子区域,从图可知,流域 2、4 子流域为该流域的上游流域,基于有效流域划分方法,形成了以 10 号流域为下游的包含流域 2 和流域 4 的等

效流域,流域面积为 1.9876 万  $\text{km}^2$ ,其中归阳水文观测站为等效流域的水文控制站。图 5 为归阳水文站 2004—2010 年汛期(4—9 月)的逐日流域面雨量数据序列和水位-降水关系图,其中图 5b~5f 表示了观测时间为 24、48、72、96 和 120 h 条件下,利用统计模型(章国材,2014)计算湘江(归阳站)由起涨水位至警戒水位条件下所需要的累计面雨量(圆圈所示),相应的降水临界值如表 3 所示。

利用 K-均值分类方法(Kanungo et al, 2002),将该等效流域面雨量数据序列划分成 5 个部分。利用直方图分析方法(Glasbey, 1993),将水文观测水位序列划分成 5 个子区间,数据区间如表 4 所示。

基于系统状态集和观测状态集形成 HMM 模型并利用 EM 方法得到 1~5 d 警戒水位条件下的最大概率降水序列和概率值,将各区间的有效值进行求和,得到该尺度条件下的临界面雨量阈值,计算结果如表 5 所示。

数据结果显示两种方法对归阳站在不同观测时间尺度条件下计算的临界面雨量基本一致。另外,对图 1 所示的子流域集合中选择 5 个子流域进行对比分析,并将计算结果列于表 6。

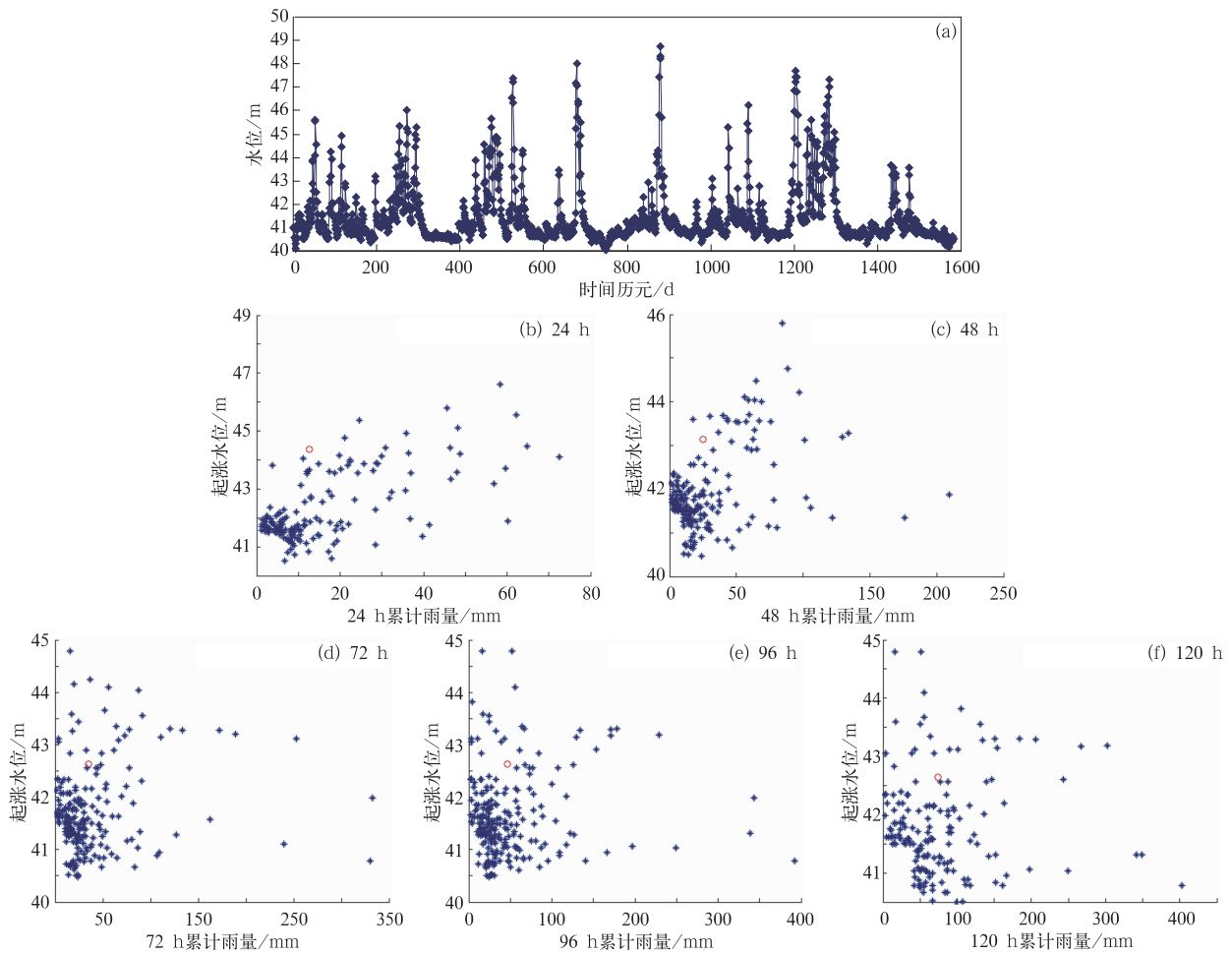


图 5 2004—2010 年湖南归阳水文观测站水位观测序列(a)以及该站在不同时间尺度条件下的起涨水位与临界水位分布图(b~f)

Fig. 5 Hydrological observations from 2004 to 2010 at Guiyang observation station (a) and distribution of start-up point and critical level at different time scales by Guiyang Observation Station (b-f)

表 3 基于统计模型计算相同条件下归阳站面雨量阈值

Table 3 Rainfall threshold at Guiyang Hydrological Station calculated by statistical method

观测时间尺度/h	24	48	72	96	120
阈值/mm	11.0	30.1	46.09	73.9	97.6

表 6 的结果进一步证明本方法与统计方法基本一致。为验证本方法的有效性,以 2010—2016 年实际观测雨量数据为依据,对表 6 涉及的流域临界雨量验证,验证所用雨量为相关流域出现警戒水文条件前 24 h 累计面雨量,并利用式(2)计算两种方法的阈值与观测样本的几何距离,结果如表 7 所示。

表 4 水位观测序列与面雨量观测序列划分

Table 4 Partition of level series and rainfall series

水文划分区间 (直方图法)	区间起始值 /m	区间终止值 /m	面雨量划分区间 (K-均值)	区间起始值 /0.1 mm	区间终止值 /0.1 mm
1	40.07	40.57	1	10.068	52.718
2	40.58	40.67	2	52.719	117.98
3	40.68	41.27	3	117.989	210.1
4	41.28	42.97	4	210.11	391.2
5	42.98	43.95	5	391.3	760.87

表 5 基于 HMM 模型计算的归阳站临界雨量结果

Table 5 Rainfall threshold at Guiyang Hydrological

Station calculated by HMM method

观测时间尺度/h	最大概率状态序列	临界雨量阈值/mm
24	3	11.7
48	14	22.0
72	144	43.02
96	1445	82.1
120	14445	103.2

$$R = \frac{1}{K} \sqrt{\sum_{i=1}^K (x_i - \theta)^2} \quad (2)$$

式中,  $R$  为阈值与警戒水位对应的实际面雨量的平均误差, 单位为 mm;  $K$  为出现警戒水位的过程数;  $x_i$  为第  $i$  次出现警戒水位过程前的等效流域面雨量值;  $\theta$  为对应的阈值。

经过实况资料和等效流域面雨量计算结果显示统计法和 HMM 方法的阈值几何距离平均值分别为:

表 6 基于两种方法对多个子流域阈值计算的结果

Table 6 Rainfall thresholds of multiple catchments calculated by statistical and HMM method respectively

计算方法	流域编号	流域面雨量阈值/mm				
		24 h	48 h	72 h	96 h	120 h
统计法	2	21	42.3	50.3	74.2	80.5
	4	23.3	49.8	59.7	117.2	136.5
	6	27.7	56	83.7	96.4	105.8
	8	11	17.7	29.3	38.8	48.6
	15	17.4	34.5	55.1	65.5	78.2
HMM 法	2	24.7	38.1	51.5	98.0	111.4
	4	17.0	58.4	99.8	218.8	260.65
	6	30	40	79.1	113.8	145.1
	8	26.4	62.7	99.1	191.8	228.2
	15	15	35.1	55.1	102.1	122.2

表 7 两种方法得出的阈值对观测样本的几何距离

Table 7 Geometric distance between the observation samples and the threshold calculated by statistical method and HMM respectively

流域号	控制站	降水量/mm					平均误差( $R$ )/mm		
		2010 年	2011 年	2012 年	2013 年	2014 年	2015 年	统计法	HMM 法
2	道县	38.2, 21.3	无	无	无	无	13.4, 27.5	4.976	4.536
4	双牌	18.8, 23.3	28.6	15.5, 23.3	11.0, 13.6, 19.1, 41.1	12.2, 19.8	27.2	4.13	2.618
6	衡阳	25.8	无	15.9, 21.2	无	15.2, 17.6, 23.3	21.4	3.2	4.0
8	耒阳	无	18.8	无	无	38.5, 10, 11.5	无	7.15	5.7
15	衡山	24.1, 13.2	无	15.5, 23.5	无	18.3	20.2, 12.2	1.6	2.0

注: 为降低观测数据误差干扰, 将警戒水位设定成区间, 范围为  $[A \times 99.75\%, A \times 100.25\%]$  ( $A$  为警戒水位值)。

Note: In order to reduce the influence of observed noise, a warning water level is regarded as a interval, such as  $[A \times 99.75\%, A \times 100.25\%]$  where  $A$  is the warning level value in theory.

4.21 和 3.78 mm, 对应方差为 4.26 和 2.2。由此可见, 相对于统计法, 基于本文提出的方法有较好的准确率和更好的稳定性。

为验证阈值在气象防灾减灾中的业务应用情况, 利用本文提出的方法, 开发了相应的阈值分析平台和洪水预警分析平台, 利用 2015 年 6 月 12—13 日的降水预报对湖南中小河流暴雨洪水开展预估。图 6a 为湖南省 6 月 12 日 08 时至 13 日 08 时的降

水预报图, 图 6b 为基于该降水预报结果的中小河流洪水预报结果, 图 6c 为湖南省水文局实况预警图。可以看出, 基于本方法的阈值能成功预报出双牌、老埠头、归阳和衡山站的警戒水位。但也存在错报情况, 如罗家庙、石门和南嘴站, 经过检查发现, 罗家庙所在流域(9 号流域)错误主要是受降水预报误差的影响, 其他则主要是该流域阈值计算偏低造成的, 故需要根据业务运行效果对阈值进行检验和修订。



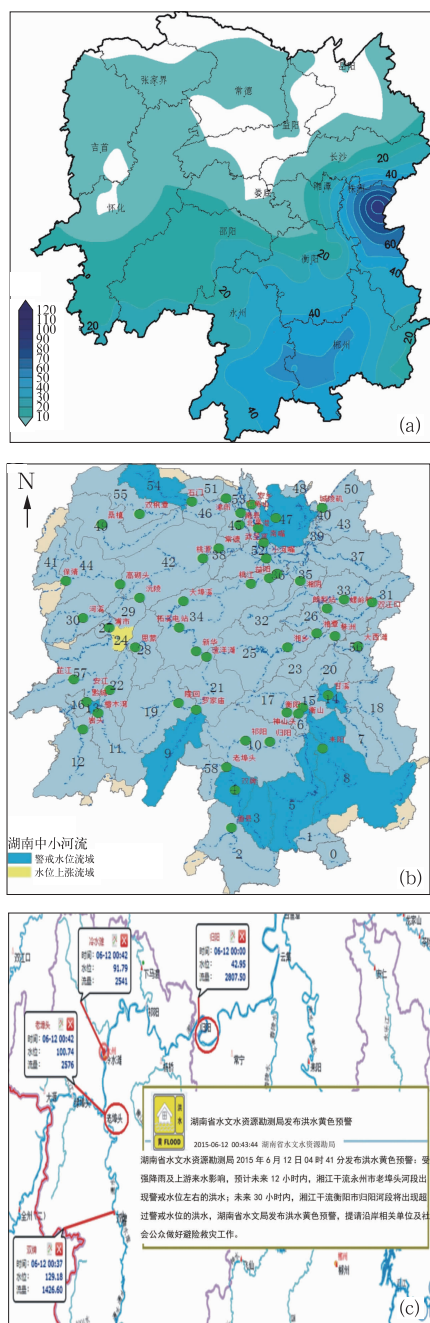


图 6 面雨量阈值在业务中的应用  
(a)24 h 预报, (b)24 h 预警分析,  
(c)实际达到警戒水位河流的分布

Fig. 6 Application of the critical threshold in early warning system (a) 24 h forecast, (b) 24 h early warning analysis, (c) actual warning stage of river

## 6 结论与讨论

基于 DEM 高程数据和湖南水文观测网络,构

建了阈值研究的有效空间范围(中小河流子流域),利用水文传播时序关系,构建了流域洪水传播有向图,结合子流域和加权有向图,提出了等效流域面雨量的概念,结合有效流域面雨量与同步水文观测数据,本文提出了基于隐马尔可夫模型的中小河流临界雨量阈值分析方法,并将该方法对实况灾情雨量比较,实验证明,相对传统统计法,新方法具有准确性高,误差空间小的特点;基于该阈值系统构建的阈值服务平台的实际业务运行也证明了新方法具有较好的业务支持能力,但在应对复杂天气过程的计算中也存在准确性不足的问题,需要在业务中进一步应用和验证。本阈值确定方法和预警系统具有如下特点:

(1) 不同于传统的基于站点的面雨量预报技术,本文提出了基于洪水流域时间响应和流域等效面雨量概念开展中小河流暴雨临界雨量阈值研究,在时间维和空间维对降水-径流进行研究,提高了临界雨量阈值确定的准确性和阈值空间的稳定性。因此,在阈值应用过程中,可以采用动态计算等效面雨量的方法和阈值进行比较开展中小河流实时预警服务,有较好的业务应用价值。

(2) 受水文观测站上游水坝等人工调蓄的影响,阈值统计法在阈值计算和预警服务过程中会受到干扰,需要大量的人工干预去识别有效起涨水位,相关统计参数选择的带主观性,难以满足业务的实时预警服务。基于 HMM 方法,采用了观测数据分类和相关状态转换概率的计算,一定程度上降低了随机干扰的影响,同时方便阈值分析自动化和预警业务化应用。

(3) 基于 HMM 方法进行的阈值确定效果受 HMM 模型结构的影响,尤其是观测数据的分类和状态集合的构建过程对隐马尔可夫模型的结果有一定的影响,状态分类对隐马尔可夫模型的影响也是当前智能计算领域研究的热点,也是未来研究的重要内容之一。

(4) 降水-径流的因果关系是本文研究的理论基础,然而受数据获取的完整性和洪水预警的现实需要,阈值确定用降水-水位的关系代替降水-径流的方法在一定程度上降低了阈值确定的准确性,因此阈值也需要在业务应用中不断得到修订和完善。

(5) 在阈值的洪水预警的应用过程中,洪水预报准确性受到降水量预报准确性的影响,同时也受水资源人工调蓄的影响。因此,阈值的修订与预报

应该与水文观测实况资料相结合,也是本工作未来努力和研究的方向。

### 参考文献

- 邓柯. 2004. 湘江干流洪水传播时间分析. 湖南水利水电, 5(1): 20-22.
- 段亚雯, 朱克云, 马柱国, 等. 2014. 中国区域 1961—2010 年降水集中度指数(PCD)的变化及月分配特征. 大气科学, 38(6): 1124-1136.
- 湖南省水利厅, 湖南统计局. 2013. 湖南省水利普查公报. [http://www.hnwr.gov.cn/xxgk/fzgh/201309/t20130924\\_75626.htm](http://www.hnwr.gov.cn/xxgk/fzgh/201309/t20130924_75626.htm).
- 林峰, 陈桂芳. 2009. HEC-HMS 分布式水文模型的时间尺度效应研究. 吉林师范大学学报(自然科学版), 30(3): 132-135.
- 刘爱利, 王陪法, 丁园园. 2012. 地统计学概论. 北京: 科学出版社.
- 刘志雨. 2005. ArcTOP: TOPKAPI 与 GIS 紧密连接的分布式水文模型系统. 水文, 25(4): 18-22.
- 水利部水文局(水利信息中心). 2010. 中小河流山洪监测与预警预测技术研究. 北京: 科学出版社.
- 文明章, 林昕, 游立军, 等. 2013. 山洪灾害风险雨量评估方法研究. 气象, 39(10): 1325-1330.
- 吴贤云, 丁一汇, 叶成志. 2015. 江南西部雨季降水区域特征及其受热带海洋海表温度异常的影响分析. 气象, 41(3): 286-295.
- 许博, 陈鸣, 魏祥麟. 2012. 基于隐马尔可夫模型的 P2P 流识别技术. 通信学报, 33(6): 35-47.
- 叶泽纲, 胡新明, 游兴, 等. 2002. 澧水干流张家界至慈利河段洪水水面线分析. 水文, 22(3): 34-39.
- 章国材. 2014. 自然灾害风险评估与区划原理和方法. 北京: 气象出版社.
- 赵健, 贾忠华, 罗纨. 2006. ArcGIS 环境下基于 DEM 的流域特征提取. 水资源与水工程学报, 17(1): 74-76.
- Glasbey C A. 1993. An analysis of histogram-based thresholding algorithms. CVGIP: Graphical Models and Image Processing, 15(6): 532-537.
- Grillakis M G, Tsanis I K, Koutroulis A G. 2010. Application of the HBV hydrological model in a flash flood case in Slovenia. Nat Hazards Earth Syst Sci, 10(12): 2713-2725.
- Kanungo T, Mount D M, Netanyahu N S, et al. 2002. An efficient k-means clustering algorithm: Analysis and implementation. IEEE Transactions on, Pattern Analysis and Machine Intelligence, 24(7): 881-892.
- Sivakumar Bellie, Berndtsson R, Olsson J, et al. 2009. Evidence of chaos in the rainfall runoff process. Hydrolog Sci J, 46(1): 131-145.
- Smith M B L, Donald P K, Victor I R, et al. 2013. Hydrologic Model Calibration in the National Weather Service in Calibration of Watershed Models. American Geophysical Union, 133-152.