

DOI:10.13409/j.cnki.jdpme.20241017002

基于卫星遥感数据的任意地区双变量降雨事件 概率模型构建与应用*

吴晨光¹, 左秦朝², 张洁¹, 严周欣荣²

(1. 同济大学岩土及地下工程教育部重点实验室, 上海 200092;
2. 陕西工程勘察研究院有限公司, 陕西 西安 710068)

摘要: 滑坡是常见的地质灾害, 降雨是主要诱发因素。构建精确的降雨事件概率模型对于准确识别滑坡灾害的高风险区以及制定科学合理的防灾减灾策略至关重要。然而, 由于缺少降雨数据, 当前仅少数地区建立了双变量降雨事件概率模型。本文提出了一种基于卫星遥感降雨数据的任意地区双变量降雨事件概率模型构建方法。首先, 通过网络爬虫技术批量自动获取全球任意地区的卫星遥感降雨时间序列。以临海市邵家渡街道为例, 依据降雨事件间隔时间, 将连续降雨时间序列划分为独立降雨事件, 提取降雨量和降雨历时等变量。在此基础上, 基于Copula函数建立双变量降雨事件概率模型, 并利用K-S检验评估模型拟合优度。进一步地, 阐述了双变量降雨事件概率模型中重现期的定义, 并与单变量降雨事件概率模型分析结果进行了对比。最后, 构建了浙江省各地区的双变量降雨事件概率模型, 揭示了省内降雨变量的空间分布。建立的双变量降雨事件概率模型为降雨诱发滑坡风险评估提供了精确的降雨数据支撑。

关键词: 卫星遥感数据; 降雨变量; Copula函数; 双变量降雨事件概率模型

中图分类号: P 344 **文献标识码:** A **文章编号:** 1672-2132(2025)01-0059-12

Establishment and Application of Bivariate Probability Model for Rainfall Events in Arbitrary Regions Based on Satellite Remote Sensing Data

WU Chenguang¹, ZUO Qincao², ZHANG Jie¹, YAN Zhouxinrong²

(1. Key Laboratory of Geotechnical and Underground Engineering of the Ministry of Education, Tongji University, Shanghai 200092, China; 2. Shaanxi Institute of Engineering Investigation Co., Ltd., Xi'an 710068, China)

Abstract: Landslide is a common geological hazard, with rainfall being a major inducing factor. Establishing an accurate probability model for rainfall events is crucial for accurately identifying high-risk areas of landslides and developing scientific and reasonable disaster prevention and mitigation strategies. However, due to the limited availability of rainfall data, only a few regions have established bivariate probability models for rainfall events. This paper proposes a method for establishing bivariate probability model for rainfall events in any region based on satellite remote sensing data. First, satellite remote sensing rainfall time series from arbitrary regions in the world were automatically obtained using web crawler technology. Taking Shaojiadu Subdistrict in Linhai City as an example, the continuous rainfall

* 收稿日期: 2024-10-17; 修回日期: 2025-01-03

基金项目: 国家重点研发计划(2021YFB2600500)、国家自然科学基金项目(42072302, 52025094)资助

作者简介: 吴晨光(1997—), 男, 博士研究生。主要从事岩土工程可靠度分析。E-mail: wuchenguang@tongji.edu.cn

通信作者: 张洁(1980—), 男, 教授。主要从事岩土及地下工程风险分析与控制。E-mail: cezhangjie@tongji.edu.cn

time series were divided into independent rainfall events based on inter-event time definition (IETD), and variables such as rainfall amount and duration were extracted. On this basis, a bivariate probability model for rainfall events was established using Copula function, and K-S test was used to evaluate the goodness of fit of the model. Furthermore, the definition of the return period in the bivariate probability model for rainfall events was discussed and compared with the analysis results of the univariate probability model. Finally, the bivariate probability model for rainfall events was established for each region in Zhejiang Province, revealing the spatial distribution of rainfall variables across the province. The established bivariate probability model for rainfall events provides accurate rainfall data support for the risk assessment of rainfall-induced landslides.

Keywords: satellite remote sensing data; rainfall variable; Copula function; bivariate probability model for rainfall events

0 引言

我国滑坡灾害频发,其中降雨是主要诱发因素^[1-2]。受多种复杂气象因素影响,降雨事件具有显著的不确定性。建立精确的降雨事件概率模型,对于降雨诱发滑坡定量风险评估具有重要的意义^[3]。一般地,一场降雨事件可以用降雨量 V 、平均降雨强度 I 、降雨历时 D 和降雨事件发生频率等变量来表征^[4]。目前,基于降雨量或平均降雨强度的单变量降雨事件概率模型已在水文等领域获得了广泛的应用^[5]。单变量分析主要研究降雨事件中单一变量(如降雨量或平均降雨强度)的统计特性。基于单变量分析,所建立的单变量降雨事件概率模型仅针对单一变量,未考虑其他变量的影响^[6]。降雨诱发滑坡同时受到降雨量和降雨历时的影响,使得单变量降雨事件概率模型无法用于降雨诱发滑坡的风险分析^[7]。

与单变量分析相比,双变量分析主要研究降雨事件中两个相关变量(如降雨量与降雨历时、平均降雨强度与降雨历时)的联合分布特性。基于双变量分析,所建立的双变量降雨事件概率模型能同时考虑两个降雨变量及其相互关系,从而更加全面地描述降雨事件^[8]。在水文分析领域,Copula函数被广泛用于构建双变量联合分布^[9]。该方法通过考虑变量间的相关性,将变量的边缘分布函数连接成多维联合分布。Copula函数法不受边缘分布类型的限制,具有高度的灵活性,因此被广泛用于构建双变量降雨事件概率模型。例如,L. Zhang等^[4]采用Archimedean Copula构建了降雨强度与降雨历时、降雨强度与降雨量以及降雨历时和降雨量的联合

分布函数。许月萍等^[10]利用不同类型的Copula函数建立了降雨历时与降雨量的联合分布函数。C. Jun等^[11]采用Frank Copula建立了降雨历时和降雨强度的联合分布函数。H. Li等^[12]根据中国东南沿海8组雨量计50多年的日降雨数据,结合时变广义极值分布和Copula函数,构建了各站点年最大降雨量和降雨强度的边缘分布及其联合分布。

建立双变量降雨事件概率模型需要长期的降雨监测数据,然而,由于缺少降雨资料,许多地区难以建立此类模型。为此,本文提出了一种基于卫星遥感数据的任意地区双变量降雨事件概率模型构建方法。本文结构如下。本文首先提出了一种基于爬虫技术的全球任意地区降雨监测数据获取方法。随后,介绍了双变量降雨事件划分方法,并阐述了双变量降雨事件概率模型及标定方法。以临海市邵家渡街道为例,分别构建并比较了双变量降雨事件概率模型与单变量降雨事件概率模型。最后,以浙江省为例,构建了该省各地区的双变量降雨事件概率模型,并分析了不同地区的降雨特征差异。所建立的双变量降雨事件概率模型为降雨诱发灾害风险评估提供了准确的降雨输入。

1 基于爬虫技术的任意地区降雨数据获取方法

1.1 GPM IMERG 卫星遥感开源降雨数据简介

降水观测数据一般有地面观测、雷达估测及卫星观测反演三大来源^[13]。地面观测降水的精度高但空间分辨率低;雷达估测降水的时空分辨率高但测量空间范围有限且成本较高;相比之下,卫星遥感是当前唯一能够提供全球连续降水数据的观测

手段,凭借其广泛的空间覆盖和长期时序优势,近年来已得到广泛应用和关注。

目前,全球已有多个卫星降雨观测计划,其中影响力较大的有热带降雨测量任务卫星(Tropical Rainfall Measuring Mission, TRMM)、全球降水观测计划(Global Precipitation Measurement, GPM)、欧洲的气象卫星(Meteosat)以及中国的风云系列卫星(Fengyun)。这些卫星监测数据,通过各种反演算法,生成了一系列卫星遥感降雨产品。全球降水观测计划(GPM)是由美国航空航天局(NASA)和日本宇宙航空研究开发机构(JAXA)主导的国际项目,其卫星系统包括一颗主卫星和多颗副卫星。GPM IMERG 作为其核心数据产品,凭借高时空分辨率及开源优势,在极端天气事件分析、洪水预警等领域得到了广泛应用^[14]。

GPM IMERG 产品提供全球范围内空间分辨率为 0.1° (约 10 km)、时间分辨率为 0.5 h 的降水数据。该产品基于 GPM 卫星的降雨时间序列,融合了微波观测与地面站点校正等多源数据^[14]。根据发布时效,GPM IMERG 产品分为“Early-run”、“Late-run”和“Final-run”三个版本,其中 Final-run 版本经过地面站数据校正,是三者中精度最高的降水产品。本文使用 GPM IMERG 中“Final-run”产品作为降雨时间序列来源。

1.2 基于爬虫技术的降雨数据获取方法

用户在注明 GPM IMERG 数据来源的条件下,可以免费获取并使用该降雨数据。GPM IMERG 产品可根据用户对某一区域在某一起止时间内的降雨时间序列数据下载需求,生成统一资源定位符(URL)列表文件;用户点击 URL,即可下载指定的降雨数据。建立降雨事件概率模型需要大量的降雨数据。例如,本文案例分析中涉及 2001 年至 2020 年时间分辨率为 0.5 h 的降雨数据,共涉及 350 640 个 URL,手动逐一下载并不可行。为了 GPM IMERG 降雨数据的下载,NASA 提供了命令行工具 wget(World Wide Web + get),支持用户在注册 NASA Earthdata 账号后,通过爬虫技术实现数据的自动化下载与集成。不过,NASA 并未提供相应的爬虫代码。

爬虫技术是一种通过自动化程序模拟人类浏览行为,系统地遍历互联网并从中提取、收集和存储数据的软件技术。为方便自动获得任意区域的

降雨数据,因此本文基于 Python 开发了一套针对 GPM IMERG 降雨数据的自动化下载脚本,其核心代码见附表 A1。本段代码实现了一个高效的多线程爬虫程序,基于统一资源定位符列表文件,批量下载降雨数据文件。程序通过为每个 URL 分配独立的执行线程,充分利用计算资源,显著优化了数据获取的吞吐量与响应效率。在任务执行过程中,系统实时监控并输出请求的响应时间、状态码及任务进度等关键指标;若状态码返回 200,表明服务器已成功处理请求,降雨数据文件被完整下载并持久化存储至指定路径;若状态码异常,则判定为下载失败,系统自动将对应的 URL 归档至日志文件,支持后续的错误诊断与任务重试机制。通过循环算法,该代码实现了对目标文件的逐一下载任务调度,确保所有降雨数据文件被系统化、有序地获取,直至完整覆盖整个数据集。

基于该爬虫技术,能够高效获取全球任意地区的降雨数据,为区域降雨事件概率模型的构建与研究提供基础降雨数据支持。

2 双变量降雨事件的划分方法

由于降雨是一种复杂的时间序列现象。要对降雨进行定量分析,首先要对降雨事件进行划分。在单变量降雨事件分析中,常通过基于固定时间窗口(如 24、48 h)的滑动窗口法来划分降雨事件。滑动窗口法需预先设定降雨时长,无法考虑降雨历时的不确定性。在双变量降雨分析中,常采用基于间隔时间(IETD)的降雨事件划分方法^[5],如图 1 所示。与滑动窗口法不同,该方法不需要预设降雨历时,而是通过 IETD 将连续的降雨时间序列划分为独立的降雨事件。当两场降雨间的无降雨时间间隔 T 大于 IETD 时,视为两场独立的降雨事件;否则,判定为同一场降雨事件。由于雨量过小的降雨

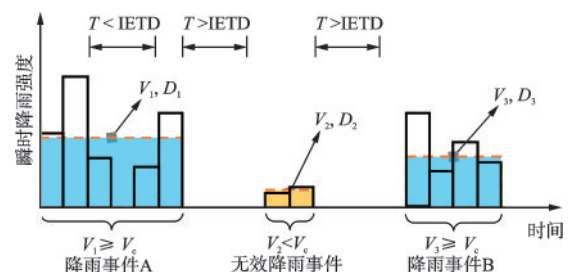


图 1 基于降雨事件间隔时间的划分方法

Fig.1 Classification method for rainfall events based on inter-event time definition

事件对工程设计的影响有限,根据具体应用场景,将降雨量小于临界降雨量 V_c 的事件定义为无效降雨事件。本文不考虑无效降雨事件。为简化表述,本文后续将有效降雨事件直接简称为降雨事件。

3 双变量降雨事件概率模型及标定方法

基于降雨事件间隔时间将连续降雨时间序列划分为独立降雨事件后,可提取降雨变量并分析其频率分布。在双变量降雨事件概率模型中,降雨可用降雨量 V 、降雨历时 D 、降雨平均强度 I 三个变量中任意两个变量进行描述。本文选取降雨量 V 和降雨历时 D 作为降雨变量。

Copula 函数是一种用于描述变量间相依关系的函数,其特点是不受变量边际分布类型的约束。通过 Copula 函数,可以将多个边际分布连接为一个多变量联合概率分布模型,因此被广泛应用于构建双变量降雨事件概率模型^[15]。在基于 Copula 函数建立双变量降雨事件概率模型时,需首先确定降雨变量的边际概率分布,然后通过变量相关性分析,构建双变量联合概率分布模型。

3.1 降雨变量边缘概率分布标定方法

对于上述降雨变量,水文分析中一般采用广义极值分布、广义帕托分布、伽玛分布等分布对其边缘分布进行拟合。作为示例,式(1)、(2)分别给出了广义帕累托分布 gp 和广义极值分布 gev 的累计分布函数:^[16]

$$F_{gp}(x) = 1 - \left(1 + k \frac{x - \mu}{\sigma}\right)^{-\frac{1}{k}} \quad (1)$$

$$F_{gev}(x) = \exp \left[- \left(1 + k \frac{x - \mu}{\sigma}\right)^{-\frac{1}{k}} \right] \quad (2)$$

式中, μ 、 σ 和 k 分别是位置参数、尺度参数和形状参数。在确定降雨变量的边缘分布后,可以通过 Copula 函数来确定变量的联合分布,相关内容将在下一节详细介绍。

3.2 降雨变量联合概率分布标定方法

基于阿基米德族的 Frank Copula 函数适用于描述具有正相关或负相关的两个变量,因此常用于构建双变量的联合分布。其联合分布函数可表

示为:^[17]

$$C_{\theta}[F(x_1), F(x_2)] = -\frac{1}{\theta} \ln \left\{ 1 + \frac{(e^{-\theta F(x_1)} - 1)(e^{-\theta F(x_2)} - 1)}{e^{-\theta} - 1} \right\} \quad (3)$$

式中, $F(x_1)$ 和 $F(x_2)$ 分别是变量 x_1 和 x_2 的边缘分布函数; θ 为 Frank copula 参数。Frank Copula 参数 θ 与肯德尔秩相关系数 τ 存在如下关系:^[18]

$$\tau = 1 + \frac{4}{\theta} \left(\frac{1}{\theta} \int_0^{\theta} \frac{y}{e^y - 1} dy - 1 \right) \quad (4)$$

依据式(4),就可由降雨变量间的肯德尔秩相关系数 τ 求解 Frank copula 参数 θ 。将 θ 和边缘分布函数代入公式(3),即可得到降雨量与降雨历时的联合分布函数。

3.3 拟合优度检验

在构建了降雨量与降雨历时联合分布后,需要评估其拟合优度。其中, K-S 检验被广泛用于评估双变量降雨事件概率模型的拟合优度^[18-19]。因此,本文也采用 K-S 检验。K-S 检验通过比较样本的经验分布函数 $F_1(x)$ 与拟合的分布函数 $F_2(x)$ 的差异来评估拟合优度,两者差值的最大值定义为 K-S 检验统计量 d , 即 $d = \max|F_1(x) - F_2(x)|$ 。若 d 小于拒绝临界值 $d(n, \alpha)$, 则认为样本服从或近似服从拟合的分布。拒绝临界值 $d(n, \alpha)$ 可以通过查表得到,其中 n 是样本容量, α 是显著性水平。

4 降雨事件回归周期确定方法

重现期被定义为大于等于某强度的降雨事件发生的平均时间间隔^[20], 在水文分析中被广泛用于表征降雨事件的罕见程度。在单变量降雨事件概率模型中, 令 $F_1(V)$ 代表给定降雨历时下年最大降雨量的累计概率分布函数, 其重现期 T 可由下式计算:^[21]

$$T = \frac{1}{1 - F_1(V)} \quad (5)$$

而在双变量降雨事件概率模型中, 降雨事件由降雨量和降雨历时两个变量共同定义, 其重现期包括“或”重现期和“且”重现期^[22]。其中, “或”重现期是指当 2 个变量中至少有 1 个变量超过特定频率时的重现期; “且”重现期则表示 2 个变量同时超过特

定频率时的重现期^[23-24]。在降雨条件下,边坡的入渗过程由降雨量和降雨历时共同决定。研究表明,滑坡多发生于降雨量较大且降雨历时较长的极端降雨事件,例如梅汛期和台汛期。因此,为了准确评估此类极端降雨事件对边坡稳定性的影响,分析时应优先采用“且”重现期^[25]。“且”重现期 T 可由下式计算^[24,26]:

$$T = \frac{1}{n_i [1 - F(x_1) - F(x_2) + C_\theta [F(x_1), F(x_2)]]} \quad (6)$$

依据公式(6),可计算双变量降雨事件概率模型中任意降雨历时与降雨量组合的重现期。通过对比公式(5)、(6),可以发现单变量降雨事件概率模型与双变量降雨事件概率模型在重现期定义上的差异:在单变量降雨事件概率模型中,重现期与降雨量存在一一对应关系;而在双变量降雨事件概率模型中,重现期由降雨量和降雨历时这两个变量的联合分布共同决定。

5 案例分析

5.1 研究区域概况

浙江省位于中国东南沿海,陆地面积约为10.5万平方公里,属于典型的亚热带季风气候区。浙江省山地和丘陵众多,降雨量充沛,使得滑坡和泥石流等地质灾害频发。此类地质灾害主要由暴雨或持续性降水诱发,大多发生在梅汛期、台汛期。因此,建立该地区精确的降雨事件概率模型对准确评估降雨引发的灾害风险具有重要意义。

本文选用GPM IMERG Final-run产品作为降雨数据源,数据时间范围为2001年至2020年,空间分辨率为 0.1° ,时间分辨率为0.5 h,共计下载350 640个降雨数据文件。在案例分析章节中,首先以浙江省临海市邵家渡街道(121.25°E , 28.85°N)为例,构建该地区的双变量降雨事件概率模型与单变量降雨事件概率模型,并分析比较两者的结果差异。随后,建立浙江省各地区的双变量降雨事件概率模型,并分析了不同地区的降雨变量分布差异。

5.2 邵家渡街道双变量降雨事件概率模型的构建

图2展示了邵家渡街道2001年1月至2020年12月的降雨时间序列,共计175 320 h。

当降雨事件间隔时间IETD为6 h且临界降雨

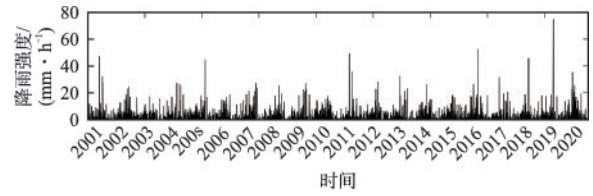


图2 邵家渡街道2001~2020年降雨时间序列

Fig.2 Rainfall time series in Shaojiadu Subdistrict from 2001 to 2020

量 V_c 为5 mm时,邵家渡街道降雨时间序列被划分为1 161个独立降雨事件,年均降雨58场。

在对降雨事件进行划分后,提取了每个降雨事件的降雨量和降雨历时,结果如图3所示。从图3可以看出,降雨量与降雨历时大多集中在左下方区域,这表明邵家渡街道的降雨事件具有降雨量较小、历时较短的特征。本文分别采用广义极值分布和广义帕累托分布对降雨历时和降雨量进行拟合,并标定了边缘分布参数,结果见表1。

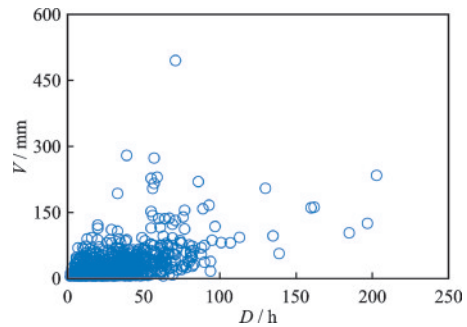


图3 邵家渡街道降雨历时与降雨量散点

Fig.3 Scatter plot of rainfall duration and rainfall amount in Shaojiadu Subdistrict

表1 双变量频率分析参数表

Table 1 Parameters of bivariate frequency analysis			
分布类型	参数	值	
广义极值分布	降雨历时 D/h	μ	17.90
		σ	12.17
		k	0.27
广义帕累托分布	降雨量 V/mm	μ	5.00
		σ	16.28
		k	0.29
Frank Copula	τ	3.41	
年均降雨场次	$n_i/场$	58	

在建立降雨量和降雨历时的边缘分布后,本文采用Frank Copula函数构建两者的联合分布函数。标定的Frank Copula参数为3.41。至此,可依据公式(3)构建降雨量与降雨历时的联合分布,结果如图4所示。

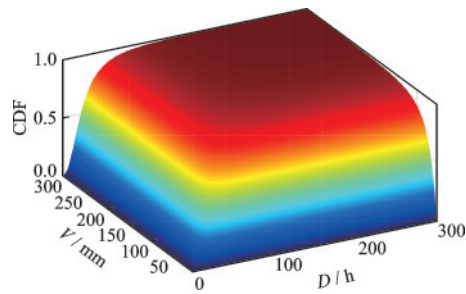


图4 降雨量与降雨历时的联合分布函数

Fig.4 Joint distribution function of rainfall amount and rainfall duration

5.3 降雨事件间隔时间和临界降雨量对降雨事件划分的影响

上一节的分析是基于降雨事件间隔时间 IETD 为 6 h 且临界降雨量 V_c 为 5 mm 这一条件的, 本节将进一步探究 IETD 和 V_c 的对降雨事件的划分的影响。

图 5 展示了在临界降雨量 V_c 为 5 mm 条件下, 降雨事件间隔时间对年均降雨场次的影响。从图中可以看出, 随着 IETD 的增加, 年均降雨场次呈现下降趋势。这一现象可归因于较大的 IETD 会将原本相互接近的降雨事件归并为同一降雨事件, 从而减少了统计意义上的降雨事件数量。

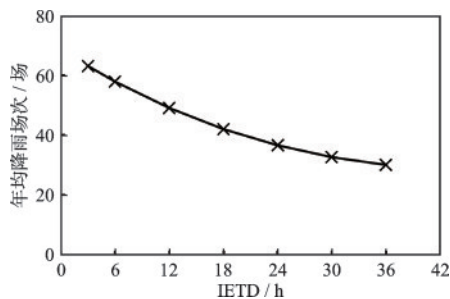


图5 降雨事件间隔时间对年均降雨场次的影响

Fig.5 Effect of IETD on annual average rainfall frequency

图 6 展示了在 IETD 为 6 h 条件下, 临界降雨量对年均降雨场次的影响。从图中可以看出, 随着临界降雨量的增大, 年均降雨场次显著减少。这一变化可以解释为, 较高的临界降雨量排除了更多降雨量较小的降雨事件, 从而导致年均降雨场次的减少。

IETD 和 V_c 的取值受降雨数据集分辨率、研究地区降雨特性及工程应用场景等多种因素的影响。目前尚无统一的标准来确定这两个参数^[27]。D. Dunkerley^[28]对 26 篇相关文献进行了归纳分析, 发现 IETD 的取值范围从 0.5 h 到 24 h 不等, 其中常见

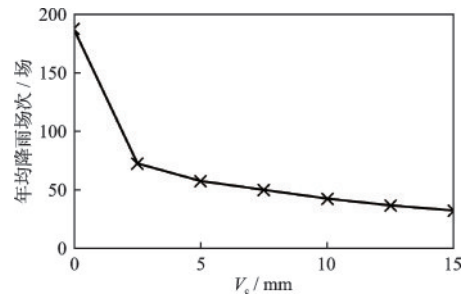


图6 临界降雨量 V_c 对年均降雨场次的影响

Fig.6 Effect of critical rainfall amount V_c on annual average rainfall frequency

的取值为 2 h 和 6 h, 这些取值大多基于经验确定^[11,27]; 而临界降雨量 V_c 的取值范围为 0.1~10 mm, 常用值为 5 mm^[29-31]。在本文的后续分析中, IETD 定为 6 h, V_c 定为 5 mm。

5.4 双变量降雨事件概率模型中降雨事件的重现期

在建立双变量降雨事件概率模型后, 可依据式 (6) 计算任意降雨量与降雨历时对应降雨事件的重现期, 其结果如图 7 所示。从图中可以看出, 在双变量降雨事件概率模型中, 降雨事件的重现期是关于降雨量与降雨历时的一条等值线。这表明, 在某一重现期下, 降雨量并非一个固定值, 其值还与降雨历时有关, 这一现象与 J. Shiau^[32] 和 C. H. Lee 等^[33] 的观察一致。

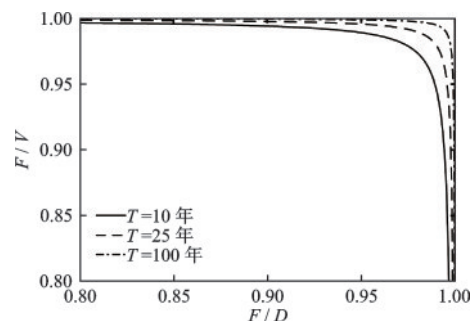


图7 双变量降雨事件概率模型中重现期的等值线

Fig.7 Contour map of return periods in bivariate probability model for rainfall events

5.5 邵家渡街道单变量降雨事件概率模型的构建

在构建邵家渡街道单变量降雨事件概率模型时, 选取了 24 h 和 48 h 两个降雨历时。根据《室外排水设计标准》^[34] (GB 50014—2021), 年最大值法适用于具有 20 年及以上连续降雨记录的地区。本文使用的降雨数据跨度达 20 年, 因此采用年最大值

法提取年最大降雨量样本序列。随后,利用广义极值分布对样本数据进行拟合,标定得到的模型参数如图8所示。本文采用K-S检验来评估模型的拟合优度,结果表明,在显著性水平 $\alpha = 0.05$ 的条件下,降雨量服从广义极值分布。

标定邵家渡街道单变量降雨事件概率模型后,可以通过公式(5)确立降雨量与重现期的关系,相关分析将在下一节中详细介绍。

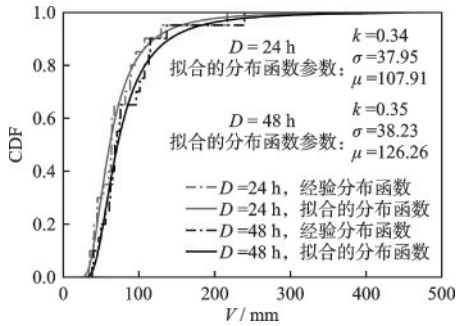


图8 不同降雨历时下降雨量的经验分布函数与拟合的分布函数

Fig.8 Empirical and fitted distribution functions of rainfall amount under different rainfall durations

5.6 双变量与单变量降雨事件概率模型中降雨量差异分析

图9和图10分别展示了在降雨历时为24 h和48 h的条件下,单变量与双变量降雨事件概率模型中降雨量与重现期的关系。如5.4节所述,在双变量降雨事件概率模型中,同一重现期对应的降雨量并非固定值。为便于模型对比分析,双变量降雨事件概率模型中某一重现期对应的降雨量是在指定降雨历时条件下计算得出的。

从图9和图10可以看出,两个模型中,降雨量

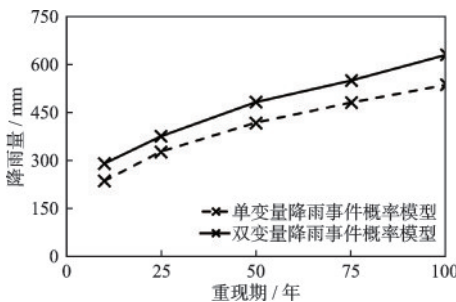


图9 24 h降雨历时下单变量与双变量降雨事件概率模型中降雨量-重现期关系

Fig.9 Rainfall amount-return period relationships in univariate and bivariate probability models for rainfall events for 24-hour rainfall duration

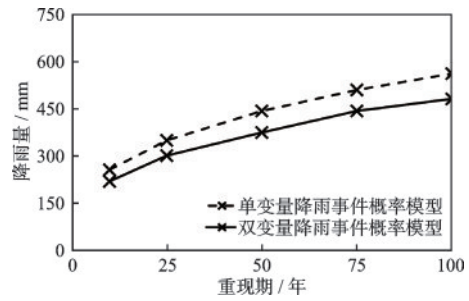


图10 48 h小时降雨历时下单变量与双变量降雨事件概率模型中降雨量-重现期关系

Fig.10 Rainfall amount-return period relationships in univariate and bivariate probability models for rainfall events for 48-hour rainfall duration

随重现期的增加而增大。当降雨历时为24 h且重现期为100年时,基于双变量降雨事件概率模型计算的降雨量为629 mm,较单变量降雨事件概率模型的535 mm高出约18%。在其他重现期下,双变量降雨事件概率模型得到的降雨量普遍比单变量降雨事件概率模型高出15%~21%。如图10所示,当降雨历时为48 h时,双变量降雨事件概率模型计算的降雨量比单变量降雨事件概率模型低14%~18%。

造成这种差异的主要原因在于降雨事件划分方式的不同以及重现期定义差异^[24]。单变量降雨事件概率模型中采用滑动窗口法划分降雨事件,仅考虑了降雨量这一变量;而双变量降雨事件概率模型则基于降雨事件间隔时间进行划分,并且能综合考虑降雨历时与降雨量的联合分布特性^[24,35]。

5.7 年均降雨量与年均降雨场次空间分布

在构建临海市邵家渡街道双变量降雨事件概率模型的基础上,本文进一步将该方法应用于浙江省其他地区,构建了全省范围内的双变量降雨事件概率模型。图11(a)展示了浙江省各地区的年均降雨量分布,结果表明年均降雨量呈现南多北少、西多东少的空间分布特征,范围为1 258~1 997 mm。其中,杭州、宁波、温州、嘉兴、湖州、绍兴、金华、衢州、舟山,台州和丽水的年均降雨量分别是1 650 (1 567, 5%)、1 620 (1 526, 6%)、1 788 (1 848, -3%)、1 395 (1 222, 14%)、1 453 (1 389, 5%)、1 521 (1 469, 4%)、1 643 (1 528, 8%)、1 868 (1 838, 2%)、1 304 (1 297, 1%)、1 699 (1 664, 2%)、1 727 (1 778, -3%)。其中,括号的数值为本文分析得到的年均降雨量,单位是mm;括号内的第一项为浙江省水利厅发布的《浙江省水资源公报》^[36]中给出的年均降雨量,单位是mm;括号内的第二项为本文分

析结果与《浙江省水资源公报》^[36]结果的相对误差。分析表明,本文计算的年均降雨量与《浙江省水资源公报》^[36]的结果差异较小,验证了所用降雨数据的准确性。

图 11(b)展示了浙江省各地区的年均降雨场次,其值介于 38~63 场,整体呈现南多北少、西多东少的空间分布格局。进一步分析表明,年均降雨量与年均降雨场次之间具有显著的正相关性,如图 11(c)所示,即年均降雨量较大的地区往往伴随着更频繁的降雨事件。

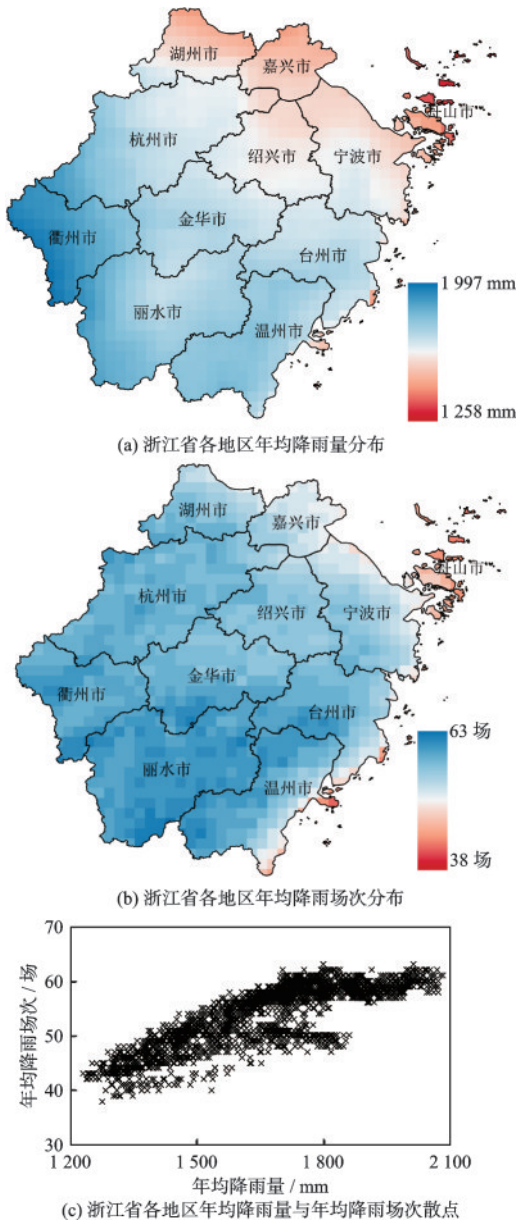


图 11 浙江省各地区年均降雨量与年均降雨场次空间分布
Fig.11 Spatial distribution of annual average rainfall amount and annual average rainfall frequency across regions in Zhejiang Province

5.8 各地区降雨量与降雨历时的边缘分布参数

在将各地区的降雨时间序列划分为独立的降雨事件后,可以提取各地区的降雨变量并进行频率分析。各地区的降雨量和降雨历时分别采用广义帕累托分布和广义极值分布进行拟合,并采用极大似然法标定边缘分布参数。标定结果如图 12 和图 13 所示。

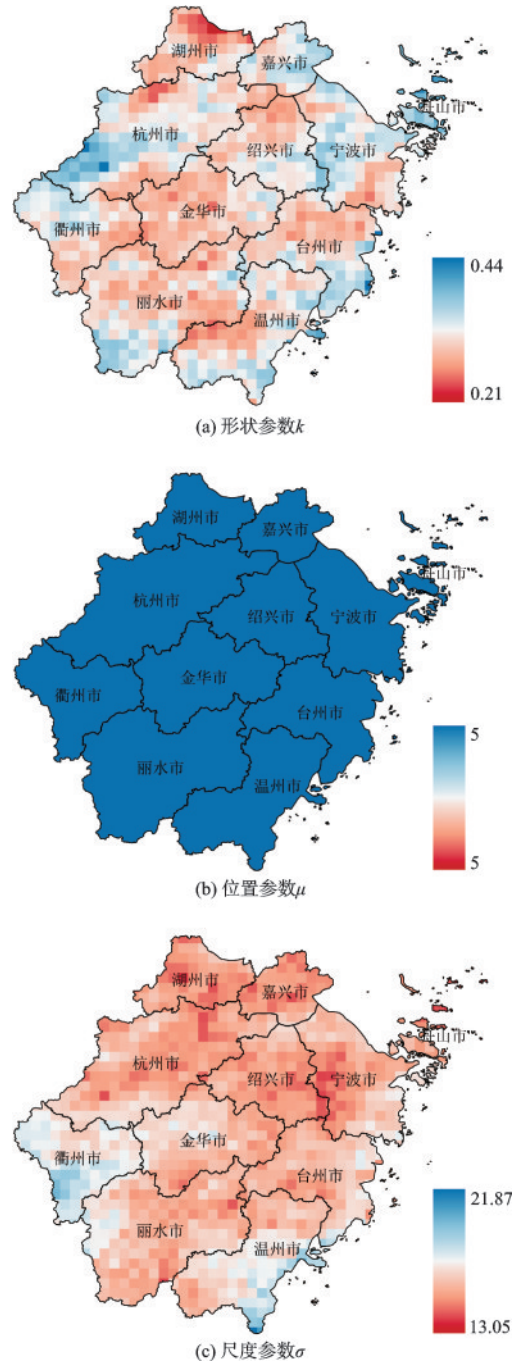
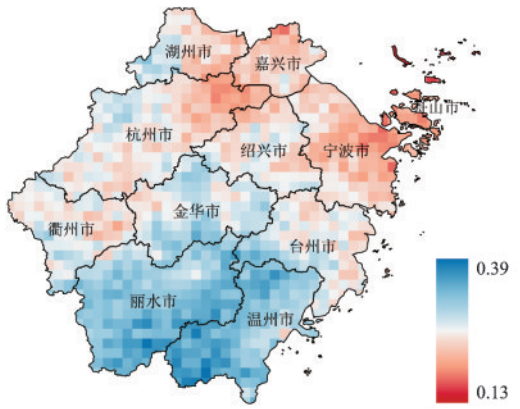
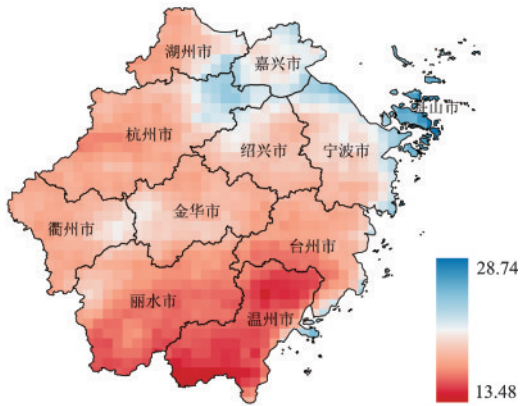


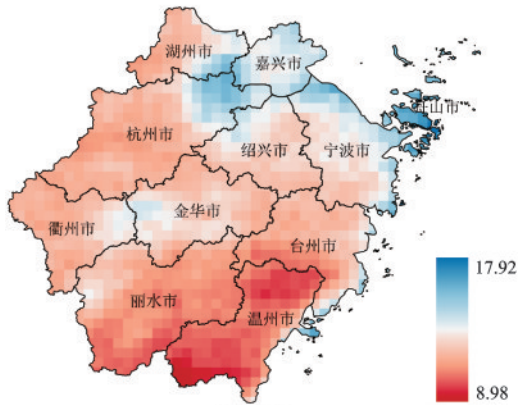
图 12 降雨量的边缘分布参数
Fig.12 Marginal distribution parameters of rainfall amount



(a) 形状参数 k



(b) 位置参数 u



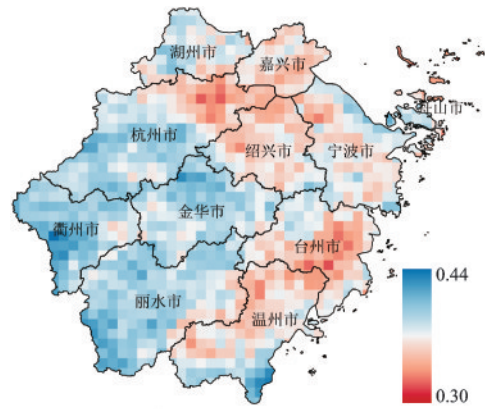
(c) 尺度参数 σ

图 13 降雨历时的边缘分布参数

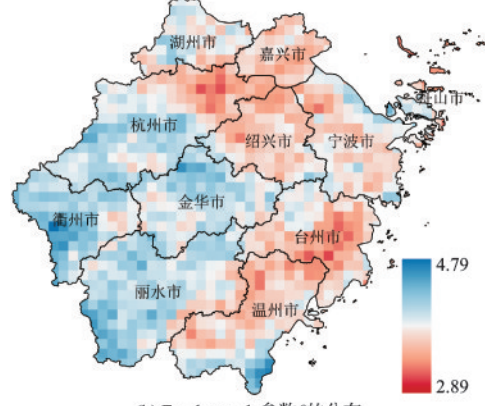
Fig.13 Marginal distribution parameters of rainfall duration

5.9 各地区降雨量与降雨历时的联合分布参数

图 14(a)展示了浙江省各地区肯德尔秩相关系数的分布。从图中可以看出,各地区的肯德尔秩相关系数为 0.30~0.44,表明降雨量与降雨历时呈现正相关关系。基于肯德尔秩相关系数,利用公式(4)可以计算 Frank Copula 参数,结果如图 14(b)所示。从图中可以看出, Frank copula 参数值为 2.89~4.79,其分布趋势与肯德尔秩相关系数高度一致。



(a) 肯德尔秩相关系数 r 的分布



(b) Frank copula 参数 θ 的分布

图 14 降雨量与降雨历时的联合分布参数

Fig.14 Joint distribution parameters of rainfall amount and rainfall duration

结合 Frank copula 参数、降雨量和降雨历时的边缘分布函数,依据公式(3)可构建浙江省各地区的双变量降雨事件概率模型。

5.10 各地区双变量降雨事件概率模型拟合优度检验

本文采用 K-S 检验评估双变量降雨事件概率模型的拟合优度。图 15 展示了各地区 K-S 检验统计

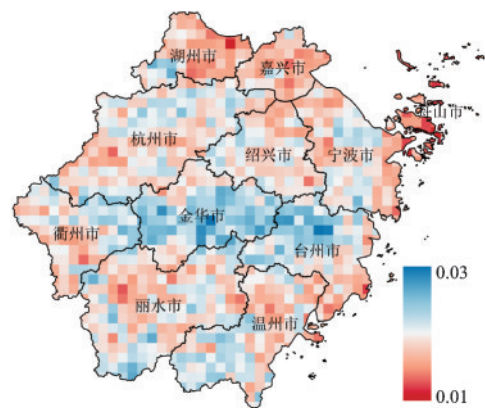


图 15 K-S 检验中 d 值的分布

Fig.15 Distribution of d values in K-S test

值 d 的分布。结果表明,各地区 d 值均小于显著性水平 α 为 0.05 条件下的临界值,表明双变量降雨事件概率模型的拟合效果良好,均通过 K-S 检验。

6 结 论

(1) 基于网络爬虫技术,实现了全球范围内任意地区卫星遥感降雨监测数据的自动批量下载。所获取数据覆盖 2001 年至 2020 年,具备 0.5 h 的时间分辨率和 0.1° 的空间分辨率。该数据的引入不仅填补了部分地区降雨监测数据的空白,同时为区域降雨事件概率模型的构建及精细化分析提供了可靠的数据支撑。

(2) 基于降雨事件间隔时间,将连续降雨时间序列划分为独立的降雨事件,并提取了降雨量、降雨历时等降雨量,进而构建了降雨量与降雨历时的边缘分布函数。在此基础上,采用 Copula 函数建立了降雨量与降雨历时的双变量降雨事件概率模型。同时,基于年最大值法,构建了单变量降雨事件概率模型。以邵家渡街道为例,进行了双变量降雨事件概率模型与单变量降雨事件概率模型的对比分析。结果表明,在降雨历时为 24 h 的情况下,双变量降雨事件概率模型计算得到的降雨量比单变量降雨事件概率模型高 15%~21%;而在降雨历时为 48 h 的情况下,双变量降雨事件概率模型计算得到的降雨量比单变量降雨事件概率模型低 14%~18%。

(3) 构建了浙江省各地区的双变量降雨事件概率模型。分析结果表明,2001 至 2020 年间,浙江省各地区年均降雨量介于 1 258~1 997 mm,呈现西多东少、南多北少的空间分布特征。此外,年均降雨量与年均降雨场次呈显著正相关关系。进一步分析发现,降雨事件的降雨量与降雨历时之间也存在显著的正相关性,其肯德尔秩相关系数为 0.30 至 0.44。

参考文献:

[1] 刘红军,武闻禹,耿林,等. 全风化花岗岩滑坡稳定性与降雨关系分析[J]. 防灾减灾工程学报,2024,44(2): 322-332.
Liu H J, Wu W Y, Geng L, et al. Analysis of the relationship between stability of fully weathered granite landslides and rainfall[J]. Journal of Disaster Prevention and Mitigation Engineering, 2024, 44(2): 322-332. (in Chinese)

[2] 闫家康,寇磊,曾斌,等. 基于有效降雨阈值的鄂西山区五峰集镇滑坡危险性评价[J]. 防灾减灾工程学报, 2023,43(6):1405-1414.
Yan J K, Kou L, Zeng B, et al. Landslide hazard assessment based on effective rainfall threshold in Wufeng Town, Western Hubei[J]. Journal of Disaster Prevention and Mitigation Engineering, 2023, 43(6): 1405-1414. (in Chinese)

[3] 文海家,钱龙,李卓航,等. 山地城镇洪涝滑坡灾害韧性评估研究[J]. 防灾减灾工程学报,2024,44(4): 751-761.
Wen H J, Qian L, Li Z H, et al. Study on resilience assessment of flood and landslide disasters in mountainous urban areas[J]. Journal of Disaster Prevention and Mitigation Engineering, 2024, 44(4): 751-761. (in Chinese)

[4] Zhang L, Singh V P. Bivariate rainfall frequency distributions using Archimedean copulas[J]. Journal of Hydrology,2007,332(1/2):93-109.

[5] 梅超,刘家宏,王浩,等. 城市设计暴雨研究综述[J]. 科学通报,2017,62(33):3873-3884.
Mei C, Liu J H, Wang H, et al. Review on urban design rainstorm[J]. Chinese Science Bulletin, 2017, 62(33): 3873-3884. (in Chinese)

[6] 刘成林,周玉文,隋军,等. 设计降雨特征的多变量分析方法[J]. 北京工业大学学报,2015,41(6):940-945.
Liu C L, Zhou Y W, Sui J, et al. Methodology of multivariate analysis of design storm [J]. Journal of Beijing University of Technology, 2015, 41(6): 940-945. (in Chinese)

[7] Lu M, Zhang J, Zheng J, et al. Assessing annual probability of rainfall-induced slope failure through a mechanics-based model[J]. Acta Geotechnica, 2022, 17(3): 949-964.

[8] Vandenberghe S, Verhoest N, De Baets B. Fitting bivariate copulas to the dependence structure between storm characteristics: A detailed analysis based on 105 year 10 min rainfall[J]. Water resources research, 2010, 46(1): 1-17.

[9] 许月萍,李佳,曹飞凤,等. Copula在水文极限事件分析中的应用[J]. 浙江大学学报(工学版),2008(7): 1119-1122.
Xu Y P, Li J, Cao F F, et al. Applications of Copula in hydrological extreme analysis[J]. Journal of Zhejiang University(Engineering Science), 2008(7): 1119-1122. (in Chinese)

[10] 许月萍,童杨斌,富强,等. 几种 Copulas 模拟不同历时降雨量的影响分析[J]. 浙江大学学报:工学版,2009, 43(6):1107-1111.
Xu Y P, Tong Y B, Fu Q, et al. Impact analysis for

- rainfall depthsimulation of different durations through several Copulas[J]. *Journal of Zhejiang University(Engineering Science)*, 2009, 43(6): 1107-1111. (in Chinese)
- [11] Jun C, Qin X, Gan T Y, et al. Bivariate frequency analysis of rainfall intensity and duration for urban stormwater infrastructure design [J]. *Journal of Hydrology*, 2017, 553:374-383.
- [12] Li H, Wang D, Singh V P, et al. Non-stationary frequency analysis of annual extreme rainfall volume and intensity using archimedean copulas: A case study in eastern China[J]. *Journal of Hydrology*, 2019, 571:114-131.
- [13] 雷晓辉,王浩,廖卫红,等. 变化环境下气象水文预报研究进展[J]. *水利学报*, 2018, 49(1):9-18.
Lei X H, Wang H, Liao W H, et al. Advances in hydro-meteorological forecast under changing environment [J]. *Journal of Hydraulic Engineering*, 2018, 49(1): 9-18. (in Chinese)
- [14] Huffman G, Stocker E, Bolvin D, et al. GPM IMERG final precipitation 13 half hourly 0.1 degree x 0.1 degree V07 (GPM_3IMERGHH) [EB/OL]. (1998-01-01) [2013-10-21]. https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGHH_07/summary?keywords=GPM%20IMERG%20Early%20Precipitation.
- [15] Balistocchi M, Bacchi B. Modelling the statistical dependence of rainfall event variables through copula functions[J]. *Hydrology and Earth System Sciences*, 2011, 15(6):1959-1977.
- [16] Pickands J. Statistical inference using extreme order statistics[J]. *The Annals of Statistics*, 1975:119-131.
- [17] Ariff N, Jemain A, Ibrahim K, et al. IDF relationships using bivariate copula for storm events in Peninsular Malaysia[J]. *Journal of Hydrology*, 2012, 470:158-171.
- [18] Singh V P, Zhang L. IDF curves using the Frank Archimedean copula [J]. *Journal of Hydrologic Engineering*, 2007, 12(6):651-662.
- [19] Salvadori G, De Michele C. Statistical characterization of temporal structure of storms [J]. *Advances in Water Resources*, 2006, 29(6):827-842.
- [20] 左斌斌,徐宗学,叶陈雷,等. 基于Copula函数的北京市设计降雨研究[J]. *水力发电学报*, 2021, 40(2): 77-88.
Zuo B B, Xu Z X, Ye C L, et al. Study on design rainfalls in Beijing based on Copula functions[J]. *Journal of Hydroelectric Engineering*, 2021, 40(2): 77-88. (in Chinese)
- [21] Koutsoyiannis D, Kozonis D, Manetas A. A mathematical framework for studying rainfall intensity-duration-frequency relationships [J]. *Journal of Hydrology*, 1998, 206(1/2):118-135.
- [22] 蔡奇霖,徐丹丽,王兆礼,等. 基于Copula函数的水库强风暴雨遭遇分析研究[J]. *人民珠江*, 2021, 42(12): 42-49.
Cai Q L, Xu D L, Wang Z L, et al. Encounter analysis between strong wind and rainstorm of reservoir based on Copula functions [J]. *Pearl River*, 2021, 42(12): 42-49. (in Chinese)
- [23] Salvadori G, De Michele C. On the use of copulas in hydrology: theory and practice [J]. *Journal of Hydrologic Engineering*, 2007, 12(4):369-380.
- [24] Yoo C, Park C, Jun C. Evaluation of the concept of critical rainfall duration by bivariate frequency analysis of annual maximum independent rainfall event series in Seoul, Korea [J]. *Journal of Hydrologic Engineering*, 2016, 21(1):05015016.
- [25] Poulin A, Huard D, Favre A-C, et al. Importance of tail dependence in bivariate frequency analysis [J]. *Journal of Hydrologic Engineering*, 2007, 12(4):394-403.
- [26] Kim T-W, Valdés J B, Yoo C. Nonparametric approach for estimating return periods of droughts in arid regions [J]. *Journal of Hydrologic Engineering*, 2003, 8(5):237-246.
- [27] Tu A, Zeng J, Liu Z, et al. Effect of minimum inter-event time for rainfall event separation on rainfall properties and rainfall erosivity in a humid area of southern China [J]. *Geoderma*, 2023, 431:116332.
- [28] Dunkerley D. Identifying individual rain events from pluviograph records: a review with analysis of data from an Australian dryland site [J]. *Hydrological Processes: An International Journal*, 2008, 22(26):5024-5036.
- [29] Ziegler A D, Negishi J N, Sidle R C, et al. Impacts of logging disturbance on hillslope saturated hydraulic conductivity in a tropical forest in Peninsular Malaysia [J]. *Catena*, 2006, 67(2):89-104.
- [30] Fornis R L, Vermeulen H R, Nieuwenhuis J D. Kinetic energy-rainfall intensity relationship for Central Cebu, Philippines for soil erosion studies [J]. *Journal of Hydrology*, 2005, 300(1-4):20-32.
- [31] Negishi J, Sidle R C, Noguchi S, et al. Ecological roles of roadside fern (*Dicranopteris curranii*) on logging road recovery in Peninsular Malaysia: Preliminary results [J]. *Forest Ecology and Management*, 2006, 224(1/2): 176-186.
- [32] Shiau J. Return period of bivariate distributed extreme hydrological events [J]. *Stochastic Environmental Research and Risk Assessment*, 2003, 17:42-57.
- [33] Lee C H, Kim T-W, Chung G, et al. Application of bivariate frequency analysis to the derivation of rainfall -

- frequency curves [J]. Stochastic Environmental Research and Risk Assessment, 2010, 24: 389-397.
- [34] 室外排水设计标准: GB 50014—2021[S]. 北京: 中国计划出版社, 2021.
- [35] Park M, Yoo C, Kim H, et al. Bivariate frequency analysis of annual maximum rainfall event series in Seoul,

- Korea[J]. Journal of Hydrologic Engineering, 2014, 19(6): 1080-1088.
- [36] 浙江省水利厅. 2020年浙江省水资源公报[EB/OL]. (2021-08-30) [2024-12-31]. http://slt.zj.gov.cn/art/2021/8/30/art_1229243017_4719972.html.

(本文编辑:张艺)

附录

表 A-1 爬取降雨数据的核心代码

Table A-1 The core code for scraping rainfall data

```
import re
import threading
import requests

def scraping_rainfall_data():
    input_urls = [] # 用于储存输入的 urls
    fail_urls = [] # 用于储存爬取失败的 urls
    with open('E:\\urls.txt', 'r', encoding="utf-8") as f:
        for line in f:
            line = line.strip()
    input_urls.append(line)
    total_urls = len(input_urls) # 总共的 urls 数目

    def get_save_page(url):
        # 输入单个链接, 下载储存到地址 filename
        filename = re.findall(r'(3B.*\.nc4)', url)[0]
        FILENAME = 'E:\\download\\1\\' + filename
        # 发送请求
        result = requests.get(url)
        print(result.elapsed) # 显示响应时间
        print(result.status_code) # 显示响应状态
        if result.status_code == 200:
            with open(FILENAME, 'wb') as f:
                f.write(result.content)
        else:
            fail_urls.append(url)

    i = 0
    for url in input_urls:
        i = i + 1
    print(i, '/', total_urls)
    t = threading.Thread(target=get_save_page, args=(url,))
    t.start()
    t.join()

    # 记录下载失败的 url 链接
    with open('E:\\fail_urls.txt', "w") as outfile:
        for item in fail_urls:
            outfile.write(item + "\n")
```