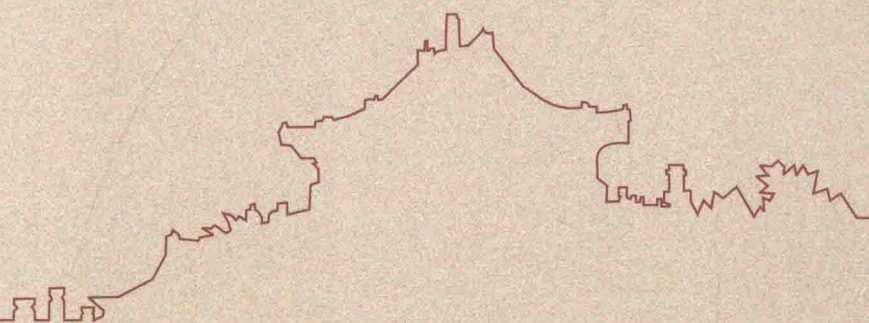


武汉大学优秀博士学位论文文库



Copula函数理论在 多变量水文分析计算中的 应用研究

陈璐 著



WUHAN UNIVERSITY PRESS

武汉大学出版社

武汉大学优秀博士学位论文文库

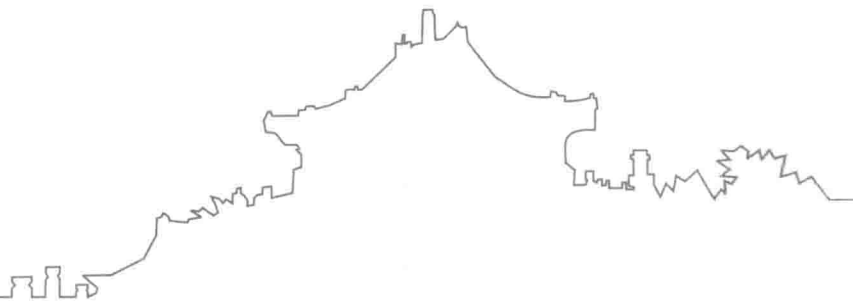


本书受国家自然科学基金项目

“Copula函数理论在多变量水文分析计算中的应用研究”（51079100）的资助

Copula函数理论在 多变量水文分析计算中的 应用研究

陈璐 著



WUHAN UNIVERSITY PRESS

武汉大学出版社

图书在版编目(CIP)数据

Copula 函数理论在多变量水文分析计算中的应用研究/陈璐著. —武汉: 武汉大学出版社, 2013. 11

武汉大学优秀博士学位论文文库

ISBN 978-7-307-11768-6

I. C… II. 陈… III. ①时间序列分析—应用—水文分析—研究 ②时间序列分析—应用—水文计算—研究 IV. P333

中国版本图书馆 CIP 数据核字(2013)第 224955 号

责任编辑:方慧娜

责任校对:鄢春梅

版式设计:马 佳

出版发行: 武汉大学出版社 (430072 武昌 珞珈山)

(电子邮件: cbs22@whu.edu.cn 网址: www.wdp.com.cn)

印刷:湖北恒泰印务有限公司

开本: 720 × 1000 1/16 印张: 11.25 字数: 158 千字 插页: 2

版次: 2013 年 11 月第 1 版 2013 年 11 月第 1 次印刷

ISBN 978-7-307-11768-6 定价: 26.00 元

版权所有, 不得翻印; 凡购我社的图书, 如有质量问题, 请与当地图书销售部门联系调换。

总 序

创新是一个民族进步的灵魂,也是中国未来发展的核心驱动力。研究生教育作为教育的最高层次,在培养创新人才中具有决定意义,是国家核心竞争力的重要支撑,是提升国家软实力的重要依托,也是国家综合国力和科学文化水平的重要标志。

武汉大学是一所崇尚学术、自由探索、追求卓越的大学。美丽的珞珈山水不仅可以诗意栖居,更可以陶冶性情、激发灵感。更为重要的是,这里名师荟萃、英才云集,一批又一批优秀学人在这这里砥砺学术、传播真理、探索新知。一流的教育资源,先进的教育制度,为优秀博士学位论文的产生提供了肥沃的土壤和适宜的气候条件。

致力于建设高水平的研究型大学,武汉大学素来重视研究生培养,是我国首批成立有研究生院的大学之一,不仅为国家培育了一大批高层次拔尖创新人才,而且产出了一大批高水平科研成果。近年来,学校明确将“质量是生命线”和“创新是主旋律”作为指导研究生教育工作的基本方针,在稳定研究生教育规模的同时,不断推进和深化研究生教育教学改革,使学校的研究生教育质量和知名度不断提升。

博士研究生教育位于研究生教育的最顶端,博士研究生也是学校科学研究的重要力量。一大批优秀博士研究生,在他们学术创作最激情的时期,来到珞珈山下、东湖之滨。珞珈山的浑厚,奠定了他们学术研究的坚实基础;东湖水的灵动,激发了他们学术创新的无限灵感。在每一篇优秀博士学位论文的背后,都有博士研究生们刻苦钻研的身影,更有他们的导师的辛勤汗水。年轻的学者们,犹如在海边拾贝,面对知识与真理的浩瀚海洋,他们在导师的循循善诱下,细心找寻着、收集着一片片靓丽的贝壳,最终把它们连成一串串闪闪夺

目的项链。阳光下的汗水，是他们砥砺创新的注脚；面向太阳的远方，是他们奔跑的方向；导师们的悉心指点，则是他们最值得依赖的臂膀！

博士学位论文是博士生学习活动和研究工作的主要成果，也是学校研究生教育质量的凝结，具有很强的学术性、创造性、规范性和专业性。博士学位论文是一个学者特别是年轻学者踏进学术之门的标志，很多博士学位论文开辟了学术领域的新思想、新观念、新视阈和新境界。

据统计，近几年我校博士研究生所发表的高质量论文占全校高水平论文的一半以上。至今，武汉大学已经培育出 18 篇“全国百篇优秀博士学位论文”，还有数十篇论文获“全国百篇优秀博士学位论文提名奖”，数百篇论文被评为“湖北省优秀博士学位论文”。优秀博士结出的累累硕果，无疑应该为我们好好珍藏，装入思想的宝库，供后学者慢慢汲取其养分，吸收其精华。编辑出版优秀博士学位论文文库，即是这一工作的具体表现。这项工作既是一种文化积累，又能助推这批青年学者更快地成长，更可以为后来者提供一种可资借鉴的范式亦或努力的方向，以鼓励他们勤于学习，善于思考，勇于创新，争取产生数量更多、创新性更强的博士学位论文。

武汉大学即将迎来双甲华诞，学校编辑出版该文库，不仅仅是为百廿武大增光添彩，更重要的是，当岁月无声地滑过 120 个春秋，当我们正大踏步地迈向前方时，我们有必要回首来时的路，我们有必要清晰地审视我们走过的每一个脚印。因为，铭记过去，才能开拓未来。武汉大学深厚的历史底蕴，不仅在于珞珈山的一草一木，也不仅仅在于屋檐上那一片片琉璃瓦，更在于珞珈山下的每一位学者和学生。而本文库收录的每一篇优秀博士学位论文，无疑又给珞珈山注入了新鲜的活力。不知不觉地，你看那珞珈山上的树木，仿佛又茂盛了许多！

李晓红

2013 年 10 月于武昌珞珈山

摘 要

水文分析计算是水利工程规划设计的一个重要环节。水文事件(过程)一般具有多个方面的特征属性,现行的单变量分析方法无法全面地反映事件的真实特征。本书结合国家自然科学基金等课题,系统地介绍了 Copula 函数理论方法,探讨了 Copula 函数在多变量水文分析中的应用,研究内容涉及分期设计洪水、洪水遭遇、干旱风险分析以及河流相关性分析等。主要研究成果和创新点如下:

(1) 综述了单变量水文频率分析方法,包括分期设计洪水、洪水遭遇以及干旱风险分析等的研究进展;概述了多变量水文分析计算的发展历程和 Copula 函数在水文分析计算中的应用。

(2) 系统地介绍了 Copula 函数的理论和方法,着重介绍 Archimedean Copula 和椭圆 Copula 函数、多维 Copula 函数的构建方法、参数估计以及拟合检验的方法;分析了 Copula 函数在多变量水文分析计算应用中的优势及局限性。

(3) 采用年最大和超定量取样方法,从洪水发生时间和量级两个方面考虑,建立了基于 Copula 函数的洪水发生时间和量级的两变量联合分布,推求了分期设计洪水。结果表明:估计的分期设计洪水既能符合防洪标准的要求,又能提供更多分期洪水发生概率的信息,提高了洪水资源利用率和综合利用效率,为分期设计洪水的分析计算提供了一条新的途径。

(4) 提出和建立了考虑洪水发生时间和量级的洪水遭遇风险分析计算模型。分别采用混合 von Mises 分布描述年最大洪水发生时间的多峰特征,采用 P-III 型分布描述年最大洪水量级;基于多维 Copula 函数分别建立了长江上游干支流洪水发生时间和量级的联合分布;估计了洪水发生时间和量级的遭遇可能性及条件概率,并与

实测资料的统计结果进行比较。该方法不仅具有理论基础,而且结果合理可信,为洪水遭遇分析计算提供了新思路。

(5) 以标准降水指数 SPI 作为干旱指标,对于旱事件进行识别,选择干旱历时、干旱程度及其最小 SPI 值和干旱间隔时间作为研究变量。采用椭圆 Copula 和 Archimedean Copula 函数构造四维联合概率分布,分析研究汉江上游干旱事件的单变量重现期、联合重现期和条件重现期。结果表明,四维联合分布能够更加全面地反映干旱的真实特征。

(6) 采用信息熵和 Copula 函数相结合的非线性技术,即 Copula 函数的熵,探索了河流的上游和下游、干流和支流以及支流和支流之间的相关关系。公式推导得出变量总相关值等于 Copula 函数负熵的结论。基于多维的 Copula 函数建立了长江上游多站年最大洪水的联合分布,采用多重积分法和 Monte Carlo 模拟的方法计算 Copula 函数的熵值。结果表明:基于不同的 Copula 函数计算得到的总相关值相差较大,多重积分法和 Monte Carlo 模拟法计算值基本相同。该方法为涉及多个变量,且变量间存在非线性相关的水文事件提供了新的研究途径。

关键词:多变量;水文分析计算;Copula 函数;分期设计洪水;洪水遭遇;干旱分析;相关性分析

ABSTRACT

Hydrological frequency analysis is of great importance in hydraulic design, flood and risk controls. Since hydrological events (or processes) usually have multivariate characteristics, the traditional univariate analysis methods are unable to describe these events. Supported by the Chinese National Natural Science Fund, the theory and methodology of Co-pula function, as well as its application in hydrological field were studied. The research mainly focuses on seasonal design flood, flood coincidence, drought risk and correlation analysis. The main conclusions and innovations were summarized as follows:

(1) The application of traditional univariate analysis methods in seasonal design flood, flood coincidence and drought risk analysis were briefly reviewed and discussed. The developments of multivariate analysis and applications of copulas in hydrological field were reviewed.

(2) Copula theories, mainly including the Archimedean copulas and meta-elliptical copulas, constructions of multivariate copulas, estimation of parameters and methods of goodness of fit were introduced. The advantages and limitations of copulas in hydrological applications were also discussed.

(3) A new seasonal design flood method which considers the flood occurrence dates and magnitudes of the peaks was proposed and established based on Copula functions. Case study results show that the proposed method could satisfy the flood prevention stand-

ards and provide more information about the flood occurrence probability in each sub-season, and enhance floodwater utilization rate and comprehensive benefits. The proposed method may provide a new approach for seasonal design flood estimation.

(4) The flood coincidences by considering flood magnitudes and time (dates) of occurrence analysis method were proposed and established. The mixed Von-Mises and Pearson Type III distributions were selected as the marginal distributions of flood occurrence dates and flood magnitude for annual maximum flood series, respectively. Four-dimensional copula functions were developed for the joint distribution of flood magnitudes and occurrence dates in the upper Yangtze River basin. The coincidence probabilities and conditional probabilities were calculated and compared with the observed ones. It is shown that the proposed method not only has theoretic background, but also has reliable and rational results, which may provide a new way for flood coincidence analysis.

(5) The monthly Standardized Precipitation Index (SPI) time series was used as drought index to identify drought event. Drought characteristics, namely, drought duration, severity, interval time and minimum SPI values, were determined. Several copulas from the Archimedean and meta-elliptical families were applied to construct four-dimensional joint distributions. The dependence structure in each drought state was investigated and drought probabilities and return period were calculated and analyzed in the upper Han River basin. It is shown that joint distribution can completely describe drought characteristics.

(6) The total correlation and copula-entropy method was applied to measure the dependence. This method only needs to calculate the copula entropy instead of the marginal or joint entropy, which estimates the total correlation more directly and avoids the accumulation of systematic bias. The bivariate and multivariate Ar-

chimedean and meta-elliptical copulas were employed and multiple integration and Monte Carlo methods were used to calculate the copula entropy. The methodology was applied to the upper Yangtze River basin which has five major tributaries. Results showed that there is a significant difference in total correlation values when different copula functions are used, the copula entropy calculated using the multiple integration and Monte Carlo methods leads to similar results. The proposed method provides a new way for correlation analysis.

Keywords: multi-variables, hydrological frequency analysis, Co-pula function, seasonal design flood, flood coincidence, drought analysis, correlation analysis

目 录

第 1 章 绪论	1
1.1 研究背景及意义	1
1.2 现行水文分析计算方法及评述	3
1.2.1 单变量分布线型	3
1.2.2 分期设计洪水计算	4
1.2.3 洪水遭遇分析	10
1.2.4 干旱分析研究进展	11
1.3 多变量水文分析方法综述	13
1.3.1 多元联合分布函数	14
1.3.2 非参数方法	19
1.4 Copula 函数研究进展	21
1.4.1 Copula 函数理论的研究	22
1.4.2 二维 Copula 函数的应用	23
1.4.3 多维 Copula 函数的应用	25
1.5 本书研究内容	26
第 2 章 Copula 函数理论与方法	28
2.1 Copula 函数的基本定义和属性	28
2.2 Copula 函数的分类	30
2.2.1 Archimedean Copula 函数	30
2.2.2 椭圆 Copula 函数	31
2.2.3 Plackett Copula 函数	32
2.2.4 经验 Copula 函数	34

2.3 Archimedean Copula 函数	34
2.3.1 Gumbel-Hougaard Archimedean Copula 函数	34
2.3.2 Frank Archimedean Copula 函数	35
2.3.3 Cook-Johnson (Clayton) Archimedean Copula 函数	35
2.3.4 Ali-Mikhail-Haq Archimedean Copula 函数	35
2.3.5 Joe Archimedean Copula 函数	36
2.4 椭圆 Copula 函数	37
2.4.1 正态 Copula 函数	37
2.4.2 t -Copula 函数	37
2.5 高维联合分布的构建	38
2.5.1 多维椭圆 Copula 函数	39
2.5.2 多维 Archimedean Copula 函数	40
2.5.3 条件混合法	44
2.6 Copula 函数的参数估计	45
2.6.1 Kendall 相关系数法	45
2.6.2 极大似然法	45
2.6.3 椭圆 Copula 的参数估计	47
2.7 Copula 函数的拟合检验	47
2.7.1 图形分析法	47
2.7.2 Genest 和 Rivest 方法	48
第3章 基于 Copula 函数的两变量分期设计洪水研究	49
3.1 汛期分期原理和方法	50
3.1.1 气候特征与天气系统成因分析	50
3.1.2 矢量统计法	50
3.1.3 相对频率法	52
3.2 基于 Copula 函数的两变量分期设计洪水	52
3.2.1 联合分布的建立	53

3.2.2 分期设计洪水的推求	54
3.3 应用研究	56
3.3.1 隔河岩水库流域汛期分期	56
3.3.2 经验频率的计算	60
3.3.3 分布函数拟合检验	60
3.3.4 隔河岩水库汛期分期设计洪水	64
3.3.5 比较分析	68
3.4 本章小结	74
第4章 洪水遭遇风险分析	75
4.1 多维 Copula 函数	76
4.2 洪水发生时间和量级的边缘分布函数	76
4.3 洪水遭遇风险分析	77
4.4 长江上游洪水干支流遭遇分析	78
4.4.1 研究区域概况	78
4.4.2 边缘分布的建立	80
4.4.3 联合分布的建立	83
4.4.4 洪水遭遇风险分析	85
4.4.5 条件概率分析	91
4.4.6 比较分析	93
4.5 本章小结	94
第5章 基于 Copula 函数的干旱分析	95
5.1 SPI 降水标准指标	96
5.2 干旱事件的定义	96
5.3 联合分布的建立	97
5.4 重现期分析	97
5.4.1 单变量重现期	98
5.4.2 多变量重现期	99
5.4.3 条件重现期	103
5.5 应用研究	103

5.5.1 研究区域概况	103
5.5.2 降雨模拟	104
5.5.3 相关性分析	107
5.5.4 边缘分布的建立	107
5.5.5 联合分布的建立	108
5.5.6 干旱概率分析	110
5.5.7 重现期分析	110
5.5.8 分级干旱事件联合分布的建立	113
5.6 本章小结	117
第 6 章 基于 Copula 函数熵的河流相关性分析	118
6.1 熵理论	119
6.1.1 随机变量的熵	119
6.1.2 最大熵原理	121
6.2 Copula 熵	123
6.2.1 Copula 熵的定义	123
6.2.2 Copula 熵的计算方法	125
6.3 总相关	126
6.4 应用研究	127
6.4.1 区域概况	127
6.4.2 河流相关性分析	131
6.5 本章小结	143
第 7 章 结论与展望	145
7.1 主要研究成果与结论	145
7.2 展望	147
参考文献	149
致 谢	164

第1章 绪 论

1.1 研究背景及意义

水是生命之源,最早的人类文明都出现于河流的沿岸,如中国的黄河、埃及的尼罗河、印度的印度河和美索不达米亚的幼发拉底河。然而,水又具有利害两重性,既能兴利,又能为害。纵观历史,人类的进步史就是一部与水的抗争史。从古代到现代,人类通过各种措施对自然水资源进行整治,使其变害为利,美索不达米亚平原上的灌溉系统,罗马的水道,中国的灌溉工程、运河和黄河的治理等,都是显著的例子^[1]。

在各类自然灾害当中,水灾害是与人类关系最为密切的。自然界总会同时出现两种情形:一面是洪水泛滥,另一面却是干旱成灾。随着人类社会的发展,人类活动导致易损性增加和生态自然平衡的改变,对我们赖以生存的环境破坏也比以往更大。不仅如此,由于受气候变迁、环境退化以及厄尔尼诺等现象的影响,与水有关的灾害其频度和强度都将持续增强^[2],水灾害所造成的损失也在逐渐增加^[3]。据统计,世界范围内水灾害所造成的经济损失占60%以上,如洪涝灾害占40%,旱灾占15%,雪崩、海啸等占5%以上^[3]。2002年约有11 000人死于自然灾害,其中约有半数死于洪水;1992—2001年报道有25万人因干旱致死^[2]。

我国由于受季风气候及地形、地貌等自然条件的影响,降水时空分布不均,进而导致水资源时空分布严重不均,自古以来就是一个水灾害频繁发生且影响范围大、危害严重的国家,暴雨、洪水和干旱等极值事件对社会经济的发展造成了重大的危害。据史书记载,从公元前

206 年至 1949 年中华人民共和国成立的 2 155 年间,大水灾就发生了 1 029 次,几乎每两年就有一次。作为中华民族母亲河的长江,唐代至清代的 1300 年间,共发生洪灾 223 次。唐代发生水灾 16 次,平均每 18 年发生一次;宋、元代 79 次,平均每 5.2 年一次;明清时期 128 次,平均每 4.2 年 1 次,近代洪灾变得更加频繁。黄河的水患灾害较长江有过之而无不及,在历史上曾决口 1 500 多次,大的改道 26 次,平均每 3 年有一次决口,每 100 年有一次大改道。据邓拓《中国救荒史》的统计结果,自公元前 1766 年至公元 1937 年,旱灾共 1 074 次,平均每 3.4 年便有一次。由此可见,洪水、干旱等水文极值事件严重地制约着我国社会经济的发展,并威胁着千百万人民的生命财产安全。了解水文事件的特性及其变化规律,为人类战胜洪水与干旱、充分合理开发和利用水资源、不断改善人类生存和发展的环境条件提供了科学依据。

水文事件一般具有多个方面的特征属性,是一个包含频域、时域和空间域的复杂过程。如洪水事件包括洪峰、洪量和洪水过程线;暴雨事件包括暴雨历时和强度;干旱事件包括干旱历时、干旱强度和干旱间隔时间等;洪水遭遇事件包括洪水发生时间和量级的遭遇。此外,水文事件各个变量之间又存在一定程度的相关性。如洪水过程包括洪峰、1 天洪量、3 天洪量、7 天洪量等,这些变量之间一般具有较强的相关性。水文分析计算就是研究频域、时域和空间域这三个自然属性的分布特征。传统的水文分析与计算方法,仅挑选某一特征变量进行单变量的分析与计算,这种分析无法全面反映事件的真实特性,忽略了变量间的相关性特征,会造成低估或高估水文特征值的后果。Chebana 和 Ouarda(2011)指出在水文事件是由几个相关随机变量来定义的情况下,单变量频率分析不能完整地评估水文事件的发生概率;通过建立联合分布,综合考虑所有相关变量,才能更好地理解水文现象^[4]。因此,近年来多变量水文分析计算成为一个研究热点,并已证实它比单变量的分析能更好地描述水文事件的内在规律和各个特征属性之间的相互关系^[5]。

本书依托国家自然科学基金项目“Copula 函数理论在多变量水文分析计算中的应用研究”(51079100),以长江流域水系作为研究区域,致力于 Copula 函数在多变量水文分析中的应用研究。

1.2 现行水文分析计算方法及评述

1.2.1 单变量分布线型

分布线型选择是水文频率计算中的基本问题^[6]。对于设计洪水而言,基于实践应用检验和统计检验的结果,不同的区域可能会根据实际情况选择最适合自己的分布形式。分布线型的选择一般要求理论依据充分、应用简单便捷、形式灵活稳健、易于接受^[7]。如 1999 年英国出版的《洪水估算手册》推荐采用广义逻辑分布(Generalized logistic, GL)作为频率曲线分布线型;20 世纪 60 年代,我国根据大量长期洪水系列分析结果和多年来设计工作的实际经验,规定 P-Ⅲ型分布作为我国水文分析计算中的线型^[8]。世界各国设计洪水频率分析所用的分布线型如表 1-1 所示^[7]。

表 1-2 总结了水文分析计算中常用的几种分布线型:正态分布常用来模拟无偏的水文变量,指数分布常用来模拟洪水发生的次数、干旱的历时和干旱的间隔时间,Gamma 分布常用来模拟降雨量或干旱程度等,P-Ⅲ型分布常用来模拟降水量、径流量或洪峰流量等,Generalized Pareto (GP)分布常用来模拟超定量序列,广义极值分布常用来模拟水文极值事件。

表 1-1 洪水(暴雨)频率分布线型

分布线型	国 家
皮尔逊Ⅲ型分布(P-Ⅲ)	中国、奥地利、保加利亚、匈牙利、波兰、罗马尼亚、瑞士、泰国
对数皮尔逊Ⅲ型分布(LP-Ⅲ)	美国、澳大利亚、加拿大、新西兰、墨西哥以及拉丁美洲和南美等国
广义极值分布(GEV)	法国、爱尔兰等欧洲国家
广义逻辑分布(GL)	英国