

确定性水文模型的 贝叶斯概率预报： 理论与方法

邢贞相 芮孝芳 著
付强 孙颖娜



科学出版社

确定性水文模型的贝叶斯概率预报： 理论与方法

邢贞相 芮孝芳 著
付 强 孙颖娜

科学出版社

北 京

内 容 简 介

本书对解决水文模型或水文水资源领域的的数据及模型参数的不确定性分析提供通用的研究方法,所构建的贝叶斯概率洪水预报系统能够充分考虑水文模型参数的不确定性和确定性水文模型的后验信息,同时给出洪水预报量数值和指定概率的置信区间,可为防洪决策提供预报值的不确定度。

本书采用 BP 神经网络和改进的蒙特卡罗数值算法研究贝叶斯预报系统的先验密度、似然函数的表达方法及后验密度获取技术,构建流域概率降雨预测模型和概率洪水预报系统。对各种算法的基本原理进行详尽介绍和实例应用研究。同时,建立基于贝叶斯理论的马尔可夫链蒙特卡罗模型算法,用于水文模型参数的不确定性分析。利用该算法分析获取 Nash 模型参数在不同研究流域的不确定性特征。

本书介绍的理论方法具有广泛的适用性,可供从事水文学及水资源、水利工程、农业工程、环境科学及其他相关专业的教学、研究和管理工作的读者借鉴和参考。

图书在版编目(CIP)数据

确定性水文模型的贝叶斯概率预报:理论与方法/邢贞相等著. —北京:科学出版社, 2015.6

ISBN 978-7-03-045076-0

I. ①确 II. ①邢 III. ①贝叶斯理论-应用-水文预报-概率预报-水文模型 IV. ①P338②P334

中国版本图书馆 CIP 数据核字 (2015) 第 131529 号

责任编辑:王丽平/责任校对:邹慧卿

责任印制:徐晓晨/封面设计:陈敬

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

北京京华虎彩印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2015 年 6 月第 一 版 开本: 720 × 1000 B5

2015 年 6 月第一次印刷 印张: 12 1/4

字数: 231 000

定价: 88.00 元

(如有印装质量问题,我社负责调换)

前 言

水文学大体上形成于 17 世纪中叶, 至今已经历了近 4 个世纪的发展. 最初, 水文学主要为修建水利工程防汛和抗旱、获取水利提供科学依据, “工程水文学”由此而产生. 此后, 由于经济发展和人口增长, 水资源供需矛盾逐渐突出, 为研究水资源合理开发、优化配置和有效管理的“水资源水文学”由此而发展起来. 与此同时, 一段时间的粗放式经济发展模式和环境意识淡薄使水环境污染问题凸现, 不仅进一步加深了水资源紧缺程度, 而且危及到经济社会的可持续发展, 于是, 旨在为水环境保护和修复寻找良策的“环境水文学”也应运而生. 20 世纪末, 近乎掠夺性开发使水资源紧缺和水污染程度不断加剧, 明显影响到一些地区的生态系统安全, 保护生态系统安全和生物多样性已成为人类面临的重要任务, 从而产生了“生态水文学”.

水文预报是工程水文学研究的重要主题. 随着人们对水文模型输入、输出、结构和参数等不确定性认识的加深, 不仅加强对水文预报实时校正的研究, 而且产生了“概率水文预报”这一分支学科. 1993 年美国学者 Krzysztofowicz 创建了世界上第一个概率水文预报理论框架, 即贝叶斯预报系统 (Bayesian Forecasting System, BFS). 这一理论框架将贝叶斯统计理论与确定性水文模型耦合, 能提供既包含确定性信息, 又包含随机性信息的水文预报结果, 有利于提高防洪减灾风险决策水平.

本人自 2004 年攻读博士学位以来, 致力于概率水文预报和水文模型参数不确定性的研究, 在研究过程中取得了些许成果, 同时也有许多疑惑想与同行探讨, 因此, 萌发撰写本书的想法, 旨在通过本书与大家交流研究概率水文预报的心得, 为推进概率水文预报研究尽绵薄之力.

本书绪论介绍概率水文预报的必要性、水文不确定性的来源, 以及国内外主要研究动态; 第 1 章介绍贝叶斯概率水文预报系统的基本理论框架及确定先验密度、似然函数等的基本方法; 第 2—4 章分别介绍遗传算法、蒙特卡罗算法和 BP 神经网络模型的基本原理及作者的一些改进, 并进行相应的应用测试; 第 5 章介绍与贝叶斯预报系统协同构建概率水文预报模型涉及的确定性水文模型; 第 6 章构建基于 BP 神经网络的贝叶斯概率洪水预报系统, 介绍其在乌苏里江挠力河流域的应用; 第 7 章构建基于 Nash 模型的贝叶斯概率洪水预报系统, 介绍其分别在长江沿渡河流域和乌苏里江挠力河流域的应用.

本书主要以邢贞相的博士学位论文为基础, 融汇另外三位作者近年来的工作中取得的一些研究成果. 本人博士学位论文是在河海大学芮孝芳教授指导下完成的. 近年来工作中取得的点滴成果主要得到东北农业大学付强教授、黑龙江大学谢永刚教授和合肥工业大学金菊良教授的指导或帮助. 在本书撰写过程中, 东北农业大

学水利与建筑学院硕士研究生纪毅、姚巍、郭皓、李晶、刘美鑫、杨兆睿等参与了部分章节的编写和校对。谨在此一并表示衷心的感谢!

由于作者水平所限,书中不足之处在所难免,恳请广大读者批评指正。

邢贞相

2014 年 9 月于哈尔滨

作者简介

邢贞相 男, 1976 年 8 月生, 汉族, 中共党员, 博士, 副教授, 硕士生导师。2007 年 6 月毕业于河海大学水文学及水资源专业, 自 2007 年 7 月起在东北农业大学任教。2009~2010 年在加拿大 Alberta 大学访问留学。现为中国自然资源学会水资源专业委员会委员、中国系统工程学会农业系统工程专业委员会委员。主要研究方向有水文不确定性分析、水资源系统优化配置与评价、水文预报、农业系统工程建模等; 在 *Chinese Geographical Science*、《农业工程学报》《水利学报》等国内外期刊上发表论文 40 余篇, 其中 EI 收录 14 篇、SCI 收录 1 篇、ISTP 收录 3 篇; 出版专著 1 部; 主持国家级科研项目 1 项、省级科研项目 2 项、厅级科研项目 4 项, 参加国家级、省部级科研项目 10 余项, 参加国际学术会议 10 余次。获省级科研奖励 3 项。

芮孝芳 男, 1939 年 12 月生, 汉族, 江苏省溧阳市人。1963 年毕业于华东水利学院 (1985 年改名为河海大学) 陆地水文专业, 同年留校任教。1981—1982 年公派赴荷兰 Delft 大学留学, 获硕士学位。河海大学资深教授、博士生导师, 长期从事水文学及水资源学科的教学和科学研究工作。兼任《水利水电科技进展》主编、《水科学进展》《河海大学学报》和《水文》等杂志编委。主要研究方向: 产汇流理论、流域水文模型、数字水文学、地貌水文学、防洪规划和洪水预报。已在国内外发表学术论文 120 余篇, 出版专著《水文学原理》《产汇流理论》《河流水文学》《径流形成原理》和《水文学研究进展》等。

付强 男, 1973 年 6 月生, 汉族, 中共党员, 博士, 教授, 博士生导师。2000 年 7 月在东北农业大学分别获得工学博士学位。2002 年 5 月在四川大学水利工程博士后流动站完成博士后研究工作, 2007 年 3 月在北大荒农垦集团公司博士后科研工作站完成博士后研究工作。多年来主要从事节水灌溉理论与技术、农业水土资源优化利用与管理等方面的研究。主持国家及省部级科研项目 10 余项, 发表学术论文 200 余篇, 其中被 SCI, EI, ISTP 收录 60 余篇。出版学术专著 5 部, 主编及副主编教材 4 部。获科研奖励 10 余项。

孙颖娜 女, 1976 年 8 月生, 汉族, 中共党员, 博士, 副教授, 硕士生导师。1999 年 7 月毕业于东北农业大学农田水利专业, 获得工学学士学位; 2002 年 6 月毕业于东北农业大学农业水土工程专业, 获得工学硕士学位; 2006 年 6 月毕业于河海大学水文学及水资源专业, 获得工学博士学位。主要从事产汇流不确定性理论、水文预报、流域水文模型等方面的研究, 在《水科学进展》《河海大学学报》《水电能源科学》《水文》等期刊上发表论文 20 余篇, 先后主持和参加国家及省部级科研项目 6 项。

目 录

绪论	1
0.1 人类面临的水问题	1
0.2 概率水文预报的必要性	3
0.3 概率水文预报的特点	3
0.4 水文不确定性的来源	4
0.5 水文不确定性研究概况	5
0.6 概率水文预报的研究进展	10
参考文献	14
第 1 章 贝叶斯概率预报的基本理论框架	19
1.1 概率水文预报的特点	19
1.2 贝叶斯概率水文预报的原理	19
1.3 输入不确定性处理	22
1.4 水文不确定性处理	23
1.4.1 水文不确定性处理简介	23
1.4.2 贝叶斯理论在水文不确定性分析中应用	24
1.5 先验分布的确定方法	26
1.5.1 利用先验信息确定先验分布	26
1.5.2 利用边缘分布 $m(x)$ 确定先验分布	27
1.5.3 无信息先验分布	27
1.5.4 共轭先验分布	29
1.5.5 多层先验	30
1.5.6 主观概率	30
1.6 似然函数的确定方法	30
参考文献	33
第 2 章 遗传算法	35
2.1 遗传算法简介	36
2.2 遗传算法可行的改进措施	41
2.2.1 控制参数的设置	41
2.2.2 编码方式的改进	41
2.2.3 选择算子的改进	42
2.2.4 杂交算子的改进	42
2.2.5 算法终止条件的改进	43

2.2.6 改进父代替换方式	43
2.3 基于实数编码的加速遗传算法的改进	43
2.3.1 算法的计算原理	43
2.3.2 算法的测试	46
2.3.3 RAGA 在水资源工程中的应用	47
参考文献	64
第 3 章 马尔可夫链蒙特卡罗算法	65
3.1 马尔可夫链蒙特卡罗算法	65
3.1.1 马尔可夫链蒙特卡罗算法的基本原理	66
3.1.2 马尔可夫链蒙特卡罗期望值目标函数	67
3.1.3 马尔可夫链蒙特卡罗处理机会约束	68
3.2 Adaptive Metropolis 算法	68
3.3 MCMC 的抽样方法	70
3.3.1 AM-MCMC	70
3.3.2 Metropolis 算法	71
3.3.3 Metropolis-Hastings 算法	71
3.3.4 Gibbs 采样	72
3.4 AM 算法的改进	72
3.4.1 基于贝叶斯理论的接受概率	73
3.4.2 初始样本及协方差选择的优化	73
3.5 收敛准则	74
3.6 RAGA-AM-MCMC 算法性能测试	74
3.7 基于 BAM-MCMC 的 BFS 基本框架	77
参考文献	79
第 4 章 BP 神经网络模型	80
4.1 人工神经网络设计	81
4.1.1 ANN 的基本知识	81
4.1.2 ANN 的拓扑结构	83
4.2 BP 学习算法	84
4.2.1 BP 算法的网络误差的确定和权值的调整	84
4.2.2 BP 算法的反向传播流程	85
4.2.3 BP 算法的限制与不足	86
4.2.4 BP 算法的改进与应用	86
4.2.5 SABP 算法的应用	89
参考文献	94

第 5 章 确定性水文模型	95
5.1 新安江模型	97
5.1.1 新安江模型简介	97
5.1.2 模型的基本思想	98
5.1.3 三水源新安江模型的结构	98
5.1.4 基于 XAJ 模型的贝叶斯概率洪水预报	100
5.1.5 实例应用	102
5.2 Nash 模型	107
参考文献	109
第 6 章 基于 BP 神经网络的贝叶斯概率洪水预报	111
6.1 模型的构建	111
6.1.1 先验密度的确定	112
6.1.2 后验信息的获取及似然函数的确定	112
6.1.3 后验密度的获取	112
6.1.4 降雨量概率预报	113
6.2 BP 模型数据预处理	113
6.3 网络特征值矩阵的建立	113
6.3.1 BP 参数的选定及网络结构的确定	114
6.3.2 BP 网络仿真模拟预测	114
6.4 实例应用	114
6.4.1 基础数据	114
6.4.2 先验密度的确定	115
6.4.3 似然函数的构建	119
6.4.4 基于 BAM-MCMC 的贝叶斯概率降雨预报	122
参考文献	127
第 7 章 基于 Nash 模型的贝叶斯概率洪水预报	128
7.1 研究区域概况	128
7.1.1 沿渡河流域自然情况简介	128
7.1.2 挠力河流域研究区域概况	130
7.2 基于 RAGA-AM-MCMC 的 BFS 基本框架	133
7.3 基于 Nash 模型贝叶斯概率洪水预报	134
7.3.1 沿渡河流域概率洪水预报	134
7.3.2 襄阳—皇庄段的概率洪水预报	156
7.3.3 挠力河流域概率洪水预报的研究	166
参考文献	183

绪 论

水文学的发展与人类对水的认识和需求密切相关. 随着人类对水的认识的不断加深, 以及对水需求的不断拓展, 使得水文学的研究领域不断延伸与丰富^[1]. 在人类社会发展的早期, 人们试图通过建造水利工程来防治水害, 获取水利, 因而工程水文方面的内容得到了发展, 并于 20 世纪 30~60 年代得到快速发展, 形成了分支学科——工程水文学. 自 20 世纪 70 年代以来, 由于经济迅猛发展和人口快速增长, 各种自然资源出现了不同程度的紧张局面, 水资源的供需矛盾显得更加突出, 于是一个专事水资源供需预测、合理开发、优化配置和有效管理的水资源水文研究方向由此而迅速发展起来. 与此同时, 由于经济发展模式上的不合理, 出现了环境严重污染问题, 水环境污染首当其冲. 这不仅进一步加深了水资源的紧缺程度, 而且危及经济社会的可持续发展. 在此背景下, 旨在为水环境保护和修复寻找良策的环境水文研究方向应运而生^[2].

水文预报是水文科学的重要组成部分, 属于应用水文学范畴. 水文预报就是根据已知的信息 (测验或分析的信息) 对未来一定时期内水文要素的状态做出定量或定性的预测^[3]. 它是一项重要的水利基本工作和防洪非工程措施. 水文预报按预见期的长短可分为短期、中期和长期预报; 按预报对象可分为洪水预报、枯季径流预报、墒情预报、冰情预报、融雪径流预报、台风暴潮预报和水质预报等.

在实际水文预报工作中, 特别是防洪减灾中, 洪水预报是调度决策的重要依据, 关系到人民生命财产的安全、工农业生产和社会经济的稳定发展, 而发布的洪水预报将在很短的时间内得到检验和验证. 因此, 水文预报是一项理论性强、应用性要求高的工作. 水文学的分支学科“水资源水文学”和“环境水文学”几乎在 20 世纪 70 年代同时形成. 长期以来, 由于呈加剧趋势的水污染和水资源紧缺, 已越来越明显地影响到一些地方的生态系统安全, 保护生态系统安全和生物多样性已成为人类面临的重要任务, 与之相应, 近 10 年来生态水文学得到较快发展, 并成为水文学新的分支学科^[2].

0.1 人类面临的水问题

(1) 水旱灾害是人类面临的主要自然灾害. 人类为了生存与发展, 与水旱灾害做斗争已有几千年历史, 通过不断地总结经验和教训, 认识和掌握自然界中水文现象及其运动规律, 从而形成水文科学理论, 对水文形势进行分析, 预测未来可能出现的水文要素的变化. 但至今, 水旱灾害造成的损失仍居诸自然灾害之首. 据统计, 在世界内因水旱灾害造成的损失占各种自然灾害总损失的 55%, 其中水灾导致的损失占 40%, 旱灾导致的损失占 15%^[4].

新中国成立前, 由于防洪减灾手段少、技术落后, 我国洪旱灾害频繁, 危害范围广、生命财产损失惨重. 新中国成立后, 党和政府高度重视水旱灾害的防治和水资

源的开发利用,进行了大规模的水利基本建设,其中对防洪起到骨干性调控作用的大型水库就建立了 442 座;此外,从中央到流域机构和各地方省市都成立了水文监测、预报和管理机构,全方位地研究、监控、预报和管理洪旱灾害,从非工程措施角度防治和减轻灾害损失。我国每年因水灾、旱灾造成的损失占各种自然灾害总损失的比例要大于 55%。据统计,1901~2000 年,在全国发生的最严重的 30 次大灾害中,16 次是由洪水引起的,7 次由干旱引起,其余为 4 次地震、1 次风暴潮、1 次鼠疫和 1 次天灾人祸^[5,6]。其中,新中国成立前的 49 年发生的 22 次大灾害中,洪灾有 13 次,占 59.1%,而新中国成立后的 51 年共发生 8 次大灾害,其中洪灾 3 次,只占 37.5%。洪灾频率大大降低,洪灾损失大大减少,这与洪水预报的积极作用密不可分。据中国水利年鉴(2004 年)统计,仅 2003 年全国水文情报预报减灾效益就达 180 亿^[5]。

(2) 全球气候变暖增加了解决水问题的难度。人口的增长、工业的发展,导致二氧化碳等温室气体大量的排放,“温室效应”加剧,全球气候变暖,海平面上升,水文循环发生重要变化。全球气候变暖对我国产生了较大的影响:①使我国一些地区降水量减少,如山西汾河流域多年平均降水量已由过去的 558mm 减少到现在的 449mm;②海平面明显上升,据分析,近百年来我国海平面平均每年上升了 14mm,其中天津、江苏、上海和广东沿海上升超过了 200mm。降水量的减少加重了一些地区的干旱缺水,海平面上升加重了沿海地区和感潮河段的水灾^[1]。

(3) 水污染加剧的势头还未得到有效控制。改革开放初期,因一味追求经济发展,而水环境保护意识淡薄,留下了许多环境方面的后遗症。全国河流因大量排入污水而受到污染,使得 23.3%的河段不宜灌溉。符合饮用水、渔业用水标准的只占 14.1%,特别是北方的一些河流,流域内人口集中,工业发达,排入污水量大,河流水量又较南方小,年径流量变幅大,在枯水期,主干流水质甚至已降到不能利用的程度。南方河流水量虽大,但由于工业排污量大,且 80%以上不经处理直接排入河流湖库等水体,致使一些大城市附近的河段,已出现局部的水污染,有些河段污染已经相当严重。水污染的加剧,不仅带来了严重的生态与环境问题,同时也增加了一些缺水地区和缺水城市的缺水程度,甚至出现了缺乏饮用水的危机^[1]。

(4) 不合理的工程措施和管理也产生负面影响。盲目砍伐森林,不合理的筑坝拦水、围垦、跨流域调水、引水灌溉和开采地下水等,都有可能带来负面影响。对森林的乱砍滥伐,致使水土流失严重,恶化了当地生态与环境,造成了河道淤积,加之不合理的围垦,减少了水体的调蓄能力和输水能力,从而降低了江河防洪标准,污染地下水。不合理的引水灌溉,可能造成灌区次生盐碱化,也可能引起河流盐化。流域内大量修建蓄水工程,或不合理使用河川径流,或不合理跨流域调水,可能使河川径流不合理地减少,甚至断流,导致下游河道流淤积萎缩,防洪能力降低,湿地减小,河口水环境恶化,生物多样性减少^[1]。

0.2 概率水文预报的必要性

产汇流的研究是水文预报的主要研究内容,而揭示并模拟流域出口断面或河道上任一断面流量过程的形成规律是产汇流研究的根本问题^[7]。大量的实测资源表明,在产汇流过程中,存在着许多难以预测和控制的随机性因素。它是一个十分复杂的过程,人们至今还不能确切掌握影响产汇流过程的全部因素。水文模型描述的一些物理过程,如降雨径流转换、下渗、河道演算等都不得不建立在物理简化的基础上,加之水文模型的输入和初始状态本身都具有随机性。即使将来,随着人们认识世界水平的提高及手段的改进,也不可能完全做到这一点。具体地说,由于降雨和下垫面特性的时空分布都表现出不同程度的随机性,人们对产汇流机理尚不能确切掌握。产汇流规律显然既有确定性的一面,又有随机性的一面。因此,产汇流系统的输入及其结构和参数均具有一定的随机性,但目前广泛应用的产汇流计算模型通常认为降雨输入和模型结构及参数都是确定性的,即对一个产汇流系统来说,认为其系统响应是唯一的,而将其与实际发生的系统响应的差值视为误差作实时校正来考虑。这种产汇流计算方法没有从本质上考虑产汇流过程的随机性,是不完全符合实际情况的。综上,水文预报具有不确定性。因数据或模型的不确定性均可用概率分布来描述,即可将系统的输入、模型的参数视为符合一定分布的随机变量,故系统的输出也可用概率分布来描述。水文预报从确定性的向概率性的转变是一种发展趋势。可与任一确定性的流域水文模型协同工作的贝叶斯预报系统是一个进行概率预报的通用理论框架,在这个框架下,发展了许多适用于不同目标的概率预报系统^[8],对水文预报则可给出水文预报量的分布或其统计特征(均值、方差和置信区间等),其中均值反映确定性成分影响的结果,方差反映在不确定性成分影响下可能的变动范围。因此,这种水文预报称为概率水文预报。这种预报方法综合了各种随机因素对水文预报结果的影响,统一处理了包括在一种物理过程内的确定性规律部分和随机性规律部分,具有明显的合理性。因此,借助于概率水文预报这一新途径可以同时提供水文模型的计算结果和模型计算结果的不确定度。这样建立起来的产汇流模型将更具科学性。

0.3 概率水文预报的特点

概率水文预报能尽可能充分地利用预报过程中出现的各种信息,以概率分布的形式描述水文预报的不确定性。它不仅给出水文预报的均值,还能给出其方差和指定概率的置信区间,这就可以描述水文变量发生的不确定性程度。水文预报的不确定性随着预见期的不同而变化,且需要根据不同预报时刻获取新的信息实时修正其概率分布。与传统的确定性水文预报相比,概率水文预报有其自身的优越性,主要体现在以下三方面。

(1) 由于水文过程本身具有非线性和随机性,因此,概率水文预报比确定性预报更具科学性与合理性。

(2) 概率水文预报给出的不是唯一的预报值, 而是给出不同预报值与相应的超过频率, 这就能对预报的不确定度进行定量。

(3) 概率水文预报可使决策者将风险考虑到决策中去, 实现预报与决策有机结合, 更好地体现预报的价值。

0.4 水文不确定性的来源

作为地学分支的水文学, 其水文现象深受气候、气象、地形、地貌和下垫面等条件的影响, 因此, 人们既不能准确地获取水文要素 (降雨、流量等), 也不能获得流域内水循环要素 (降雨、蒸发、下渗、地表及地下径流等) 可靠的时空分布。故水文模型的概化给水文模型带来诸多不确定因素。水文现象本身的随机性与模糊性则是导致水文预报不确定性的根本原因。水文现象中的不确定性与确定性是互补的。这就是说无论采用哪种水文模型进行水文预报, 都会不可避免产生水文预报的不确定性问题。国内外许多学者, 如 Kuczera^[9]、赵人俊^[10]、Beven 和 Binley^[11,12]、郭生练等^[13]、张洪刚等^[14-16]、李向阳等^[17,18] 和芮孝芳^[2,19] 等都分析和讨论水文模型的不确定性问题, 主要归纳为以下三个方面。

1. 水文资料或信息的不确定性

(1) 水文变量随机分布特性的均化问题。水文变量的时程变化是连续的, 而在模型计算时采样却是离散的, 进而导致时段内的均化给模型计算带来了误差。

(2) 水文变量空间分布特性与数学期望的代表性问题, 如雨量, 由于降雨的空间分布的变动性, 导致固定雨量站网接收输入信息误差的变动性。

(3) 直接测量得到的水文要素的误差来源是多方面的, 如流量大都用流速仪的断面流速乘以断面面积算得流量。其中, 断面流速与断面面积的测得均可能存在误差, 且将水流简化为垂直于断面的一维水流在很多情况下是粗略的。在数据读取和录入过程中也可能存在人为的错误。

(4) 仪器测量精度的误差及其产生的系统误差。

(5) 部分水文要素至今还缺乏可靠的信息来源, 如流域土壤含水量、壤中流、地下径流的划分及预见期内降雨预报误差等。

2. 模型结构的不确定性

(1) 由于对水文现象或水文过程认识还不够深刻, 使现实采用的模型结构不尽合理而无法真实地反映实际水文过程。模型的参数主要依据降雨和径流资料率定, 而这样求得的模型参数必然带有经验统计性, 只能反映有关影响因素对流域径流形成过程的平均作用。

(2) 模型输入的空间分散性和不均匀性。流域水文模型的输入是流域上各点的降雨过程, 输出是流域出口断面的流量或水位过程, 这种具有分散输入和集中输出的模型与现有的流域水文模型在结构上并不匹配, 在实际应用中考虑这一问题时, 多是采用将全流域按雨量站划分若干个单元面积的方法, 认为当面积小到一定程度

时,即可作集中输入和集中输出的流域水文模型来模拟该单元面积的径流,最后将各单元面积对全流域出口断面输出的贡献叠加起来作为出口断面的输出.显然,这种处理方法仍然不够完善,例如,如何考虑不同单元面积在径流形成机制和模型参数上的差别、单元面积划分不同对模型输出的影响、各单元面积对全流域出口断面的贡献是否满足叠加原理等都未能解决.

(3) 水文模型之间的确定性联系很复杂,而模型往往用大量简化的数学物理方程去近似模拟其联系.大多数集总式模型忽略了流域空间分布面上产汇流的随机性.

(4) 许多模型没有考虑环境变化(如全球变化,人类活动使地形、地貌发生的变化等)对流域产汇流机制的影响.

(5) 由于水文工作者的个人经验、对模型的理解、对模型的喜好等因素,同一流域不同的人选择模型不同而产生了模型适用性所导致的误差.

3. 模型参数优选的不确定性

(1) 对一个水文模型来讲我们希望用最少的参数以便率定并减少误差将模型的不确定性降到最低,而多数具有多个参数.理论上讲,模型参数可以从流域直接或间接获得,但由于水文模型参数既有其物理意义,又有其概化的成分,故大部分模型参数只能在对实测资料分析筛选的基础上通过参数率定来确定,其中便增加了资料的选取、优化方法的选择、目标函数的确定,以及参数组合等因素而产生的模型参数率定的不确定性.水文学者在水文模型参数优选技术上做了大量的研究,主要方向之一就是如何有效地分析模型结构和参数的不确定性及这种不确定性对模型输出的影响^[20-22].

(2) 不同的率定样本优选出的参数也不同.率定参数时所用水文资料的质量对模型参数率定的影响远大于所选取水文资料的数量对模型参数优选的影响^[23].水文资料的质量依赖于数据中所含有关水文过程信息的多少及数据本身存在的误差,数据所含信息取决于水文过程的变幅,如涵盖了丰水年、中水年、枯水年,则认为数据中包含了较多的水文信息.

0.5 水文不确定性研究概况

1. 水文不确定性研究的进展

水文过程本身的随机性特点必然使确定性水文预报走向概率水文预报,也就是说,认识水文过程中不确定性,进而实现随机水文预报是将来水文预报的大势所趋.当前,对于水文现象不确定性的研究主要基于随机不确定性、混沌不确定性、灰色不确定性与模糊不确定性而展开.

(1) 随机不确定性.确定性水文模型与随机性水文模型是两种不同描述水文现象的方法,当前水文模型的发展趋势是由起初的两者独立地工作而走向两者之间的协同工作,即在一个水文模型当中同时考虑水文的确定性与不确定性.通常的做法

是在确定性水文模型当中加入随机项。周文德^[24]认为,自然界中的水文过程极为复杂,解释它的现象也相当困难的,最恰当的方法是将随机数学的理论与实际问题的物理过程联系起来,这就是所谓的随机水文学。在水文水资源系统中引入随机理论最初起源于水库规划的需要^[25,26]。水文过程具有随机性,这是不争的事实。为此,Thomas 和 Fiering^[27]及 Yevjevich^[28]首先开始引入马尔可夫模型来描述水文过程。20 世纪 60 年代末,随机水文学发展日新月异,出现了自回归滑动平均求和模型、解集模型、散粒噪声模型、分数高斯噪声模型、快速分数高斯噪声模型、折线模型、门限自回归模型等^[29]。1993 年, Krzysztofowicz^[8]提出用 BFS 来研究水文过程的非线性,基于贝叶斯原理,根据已有确定性水文模型的输出信息来修正原有水文预报量的先验信息。后来,他提出降雨不确定性处理 (PUP)^[30]、概率定量降雨预报系统 (PQPF)^[31]、水文不确定性处理 (HUP) 及亚高斯转换^[32]和正态线性假设模型^[33,34]来研究水文过程的随机性,取得了大量的研究成果^[35,36]。芮孝芳教授^[37]就入流系统为白噪声过程时,推导出了 Nash 模型参数与汇流系统出流过程的相关函数。此后,孙颖娜等^[38]利用随机微分方程进一步研究了 Nash 模型输入与参数为随机的随机预报方法,并在沿渡河流域的汇流计算中得到了较好的预报结果。

(2) 混沌不确定性。在一定的参数范围内求解确定的非线性方程有时会无法获得确定解,而是类似随机的混沌解(状态空间呈现无穷无尽的扭曲和折叠的运动轨迹),人们将这种确定的非线性系统出现类似随机行为的起因称为系统的内随机性。混沌理论的引入使得许多复杂的水文水资源问题有了新的解释^[39,40],且对水文变量给出了新的描述指标,分维数便是一类最重要的特征^[41,42],当某系统的分维数大于 2 时,就表现出一种对初始条件非常敏感且有不可预见的混沌现象。丁晶等^[43]系统地综述了随机水文学领域的基本理论和方法及研究成果,利用混沌分析将确定性和内随机性联系在一起,提出了基于混沌分析的相空间预测方法。混沌可以理解为确定的随机性:“确定的”是指它由内在原因而非外界干扰所产生;“随机性”是指其不规则和混乱的行为,确定的随机性说明确定性与随机性之间存在由此及彼的桥梁^[15]。这与芮孝芳教授^[2]提出的水文模型中确定性与随机性互补的观点不谋而合。分形理论及耗散理论等非线性理论也被尝试用于洪水预报领域。例如,温泉等^[44]利用局部水文相空间模式并采用最小二乘法来拟合函数,在葛洲坝和隔河岩的入库日径流预测中取得了较好的试验结果。

(3) 灰色不确定性。灰色理论把自然界中介于已知与未知信息之间的系统叫灰色系统。它始于邓聚龙的一篇论文^[45],主要研究部分信息已知、部分信息未知或未确定的系统。在国内,夏军^[46]较早将其应用于水文非线性系统模型参数的识别与预测。蓝永超等^[47]利用灰色系统中的残差序列周期修正 GM(1, 1) 模型,对龙羊峡水库入库径流的近期及未来变化趋势进行了预测。在国外,学者 Bass^[48]曾将灰色系统理论应用到全球环流模型预测评定的研究中。

(4) 模糊不确定性。模糊数学诞生于 1965 年,经过四十多年的发展,它已经渗

透到许多科学领域,水文水资源领域也不例外.模糊数学的出现使人们对水文现象非确定性的理解有了拓广与深化.模糊性是水文流域系统中客观存在的另一类不确定性,主要源于数据观测(记录)过程与知识提取(概念形成)过程. Bogardi [49] 最早将模糊理论引入水文水资源中,对地区水管理模型进行了研究. 陈守煜 [50] 对水文水资源与环境模糊集分析进行了简单集成. 他认为水文水资源系统中许多概念的外延存在不确定性,对立概念之间的划分具有中间过渡阶段,这都是客观存在的模糊现象. 为此,提出了“模糊水文学”的概念 [51]. 蔡华等 [52] 运用模糊集理论,建立了一个新的专家系统,不仅具有传统系统的功能,还对预报结果的综合分析引入了模糊综合评判的概念. 周念来等 [53] 利用模糊因果聚类分析法克服了预测因子的随机性和模糊性带来的困难.

2. 水文不确定性研究的方法

现行流域水文模型,由于其对水文物理过程的描写与定量计算过于粗略和概化,实际上只具有“模拟”功能,属于“模拟模型”. 这种模型容易做到模拟或复演过去已经发生的水文现象,利用计算机的长处,只要不断地调整模型参数的组合可能有多个最优参数组使所获得的输出具有相同的拟合精度. 而预报者在预报时往往只是从中选择一组认为最优的参数来进行预报,这就必然存在不确定性. 这就是所谓的“异参同效 (Equifinality)”,产生这一现象的原因至少有: 目标函数是多极值的; 模型中包含的参数之间存在相互补偿作用; 模型参数具有随机性 [54,55]. 因此,水文模型参数与水文预报的不确定性在国际上引起了广泛关注. 目前提出的主要方法有 Beven 和 Binley 提出的 GLUE(Generalized Likelihood Uncertainty Estimation) 法 [11], 马尔可夫链蒙特卡罗 (Markov Chain Monte Carlo, MCMC) 法 [56] 和 BFS [8] 等.

(1) GLUE 方法. 该方法由 Beven 和 Binley [11] 于 1992 年提出的,主要目的是为了解决水文模型的“异参同效”现象对水文预报的影响. GLUE 方法建立的基础是 Hornberger 和 Spear 提出的 RSA (Regionalized Sensitivity Analysis) 方法 [57]. 该方法认为: 导致模型模拟结果好坏的原因是所有模型参数的组合,而不是某一个模型参数所致. 其解决问题的基本思路是先根据已有的知识设定模型参数(起决定性作用的模型参数)的分布空间,即先验分布. 再按照先验分布抽取模型的各参数组合,以各参数组确定的模型模拟水文过程. 选取适当的似然函数,计算模型输出与实测之间的似然函数值,再将函数值归一化后作为各参数组的似然值. 设定一个似然函数的临界值,凡低于这个临界值的似然值均被赋为零,表示这些参数组无法表征模型的功能特征; 凡高于这个临界值的似然函数值所对应的参数组则表示它们能够表征模型的功能特征. 最后,按归一化权重对各组参数进行抽样,利用各抽得的参数组样本分别模拟某一水文过程,再由这些模拟结果求出该水文过程在指定置信度下模型输出的不确定性范围.

GLUE 在流域水文模型的不确定性预报中得到广泛应用 [58-61]. Freer 等 [58] 将该法应用在法国 Ringelbach 流域的 TOPMODEL 模型不确定性预报中,结果表

明, 预报流量的分布为非高斯分布, 且形状随时间的变化而变化. Franks 等^[59] 研究了无水文资料情况下 SVAT 模型的不确定性, 采用 GLUE 方法根据贝叶斯公式由新的资料对原似然值进行更新, 利用比较更新前后的不确定性估计就可以评价新增信息的价值. Camron 等^[60] 在威尔士瓦伊河流域采用 TOPMODEL 模型在 GLUE 框架下研究了 21 年降雨径流数据的洪水频率. 结果表明, TOPMODEL 模型能有效的完成洪水频率分析和连续模拟. Montanari^[61] 通过试验得出这样的结论: GLUE 方法应用于人造大样本数据时往往低估了水文模型模拟所产生的不确定性.

(2) 蒙特卡罗法. 蒙特卡罗模拟 (Monte Carlo Simulation)^[62] 也称随机模拟、统计试验, 是随机模拟的基本方法, 其理论基础是数理统计, 模拟手段是随机抽样. 蒙特卡罗法是威勒蒙 (Velleman) 和冯·诺依曼 (von Neumann) 在 20 世纪 40 年代为研制核武器而首次提出来的. 由于工作的保密要求就给这种方法起了一个能象征性表明该法特点的代号为蒙特卡罗, 即摩纳哥一个世界闻名的赌城名字. 在此之前, 作为该方法的基本思想, 实际上早已被统计学家所发现和利用. 例如, 在 17 世纪的时候, 人们就知道按频数来决定概率. 又如, 19 世纪末, 很多人曾利用随机投针试验来验证大数定律, 并根据针与地面上平行线 (距离均 2 倍针长) 相交的概率 P 等于 π 的倒数, 用频数 n/N 替代概率 P , 进而得到 $\pi \approx \hat{\pi} = N/n$. 现今, 蒙特卡罗已成为数值模拟试验的一个专用术语. 因其在粒子的随机行为和飞机轰炸任务的研究中非常有效, 故被广泛应用.

根据概率论中的强大数定律, 随机模拟蒙特卡罗法的估计值 $\hat{I}_N$ 依概率收敛于 I , 即满足

$$P\left(\lim_{N \rightarrow \infty} \hat{I}_N = I\right) = 1 \quad (1)$$

的充分必要条件是随机变量 $h(X)$ 满足条件

$$E(|h|) = \int |h(x)|f(x)dx < \infty \quad (2)$$

式中, N 为随机模拟产生的伪随机数的个数.

蒙特卡罗模拟的实质是从某种分布中随机抽取随机数来模拟现实系统中可能出现的随机现象. 每次模拟试验都可以模拟现实中可能出现的情况, 根据大数定理和中心极限定理就可得到我们所关注的统计特征. 由于近些年来计算机性能的飞速发展, 使蒙特卡罗法的应用更加广泛.

蒙特卡罗法工作开始从已知的分布中抽取随机样本, 即随机抽样. 因此, 均匀随机数的产生是其基础, 通过某种方法可以根据均匀随机数来产生其他任意分布的随机数. 产生随机数的方法分为随机数表法、物理随机数发生器法和数学方法等三类. 其中数学方法在计算机产生随机数中应用最多, 通常采用线性同余法, 利用递推的数学公式来产生随机数, 故也称伪随机数.

蒙特卡罗法的特点: ①简单. ②收敛速度与问题的维数无关. 在置信水平一定的情况下, 其误差除了与问题所确定的随机变量的方差有关, 只取决于样本容量, 而与样本观测值是多少维空间上的点无关. 这是其他计算方法所无法具有的, 因此,