

摘 要

洪水预报是防洪调度决策的主要依据,是防洪减灾工作中的一项重要的重要的非工程措施。大流域因其空间范围覆盖广、支流众多的特性,如何对其进行高精度的洪水预报一直是比较难的问题。对大流域预报结果进行实时校正是提高预报精度的有效手段之一。

本文在地理信息系统平台上完成了流域信息的提取、水文模型对象解析和流域概化,并在此基础上将大流域划分为不同级别的子系统,子系统之间由低级到高级逐级计算;研究了流域拓扑关系,各子系统内部根据该关系解决了对象最优演算次序的问题,构建了洪水预报方案。对现阶段主要的实时校正方法进行了比较,得出使用误差自回归和最小二乘在线估计耦合的自适应实时校正方法可以对洪水预报结果进行有效校正。最后将洪水预报方案与实时校正模型结合,提出了大流域的洪水预报实时校正方案。

将本文研究的成果应用于水文业务中,效果较好,洪水预报精度有了较大提高,是一种简便实用的洪水预报实时校正方法。

关键字: 大流域、拓扑关系、实时校正、预报方案

ABSTRACT

Flood forecasting is main dependence of flood controlling and decision, it is considered as an important non-structural measure on the flood control and mitigation . How to forecast its flood with high precision always is a difficult problem , for large basin's character of wide covering area and numerous branches. Real-time correct the large basin's forecasting result is one of the effective ways to improve the forecasting precision.

The paper accomplish basin information collection, hydrology object's parse and basin's generalization on GIS platform, firstly. and on the base, it compartmentalizes large basin to different level subsystem, and compute from base to advance between each subsystem; It study basin's topology relation, solve the problem of object's best figure order according to the relation between each subsystem, construct project of flood forecasting. It compares the main actual real-time correction methods and find that adopt the method of error autoregression combine with attenuation memory least-squares can effectively correct flood forecasting's result. Lastly, it combine flood forecasting project with real-time correction model to put forward the project of flood forecasting's real-time correction for large basin.

The study production of the paper is applied to hydrology can get good result, and flood forecasting's precision is much improved, so it is an applied method for flood forecasting 's real-time correction.

Key Words: large basin , topology relation , real-time correction , forecast project

学位论文独创性声明:

本人所提交的学位论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知,除了文中特别加以标注和致谢的地方外,论文中不包含其他人已经发表或撰写过的研究成果。与我一同工作的同事对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。如不实,本人负全部责任。

论文作者(签名): 段唯巍 2006 年 6 月 6 日

学位论文使用授权说明

河海大学、中国科学技术信息研究所、国家图书馆、中国学术期刊(光盘版)电子杂志社有权保留本人所送交学位论文的复印件或电子文档,可以采用影印、缩印或其他复制手段保存论文。本人电子文档的内容和纸质论文的内容相一致。除在保密期内的保密论文外,允许论文被查阅和借阅。论文全部或部分内容的公布(包括刊登)授权河海大学研究生院办理。

论文作者(签名): 段唯巍 2006 年 6 月 6 日

第一章 绪论

1.1 问题的提出

我国地域辽阔、河流众多、地形复杂。由于受季风和河流地理分布的影响，我国是一个水旱灾害频繁发生的国家。在与洪水斗争的长期实践中，人们逐渐认识到采用防洪措施可以减少洪水的灾害损失。防洪措施通常分为两大类^[1]：工程措施和非工程措施。非工程措施一般包括：加强防洪设施管理，提高防洪设施的防洪能力；建立健全通讯系统和预警系统；改进和发展洪水预报技术，提高防洪调度水平等。其中改进洪水预报方法、提高预报精度、实施洪水预报调度是非工程措施中最行之有效的办法之一^[2]。

水文预报是研究和运用水文变化规律，揭示和预测未来水文要素变化的一门应用学科。水文预报能够为防汛抗旱、水库的施工和调度管理、水资源的有效利用提供极为重要的依据。随着科技的发展，人民生活的需要，国家开始兴建很多大型水利工程，要保证这些工程高效的运转，精确的水文预报是必不可少的。这种预报一般涉及到上万甚至几十万平方公里的大流域，考虑到大流域测站多、资料情况参差不齐的特点，要能够得到比较好的预报结果，必须对资料稀缺大流域的预报方法进行研究。

本文正是针对这一问题展开讨论，在分析大流域特征的基础上，探求一种简单实用的预报方法，对预报结果进行实时校正，并将其实际应用在三峡入库站流量预报会商系统中。

1.2 国内外发展与研究现状

1.2.1 大流域预报模型的发展

现阶段在小流域上的水文预报研究相对较多，对于大流域的研究相对较少。

早期的大流域预报主要是采用降雨径流相关的方法来预报的。1978 年格雷克提出,在径流变化有明显季节性特征的流域,把前几年(率定期)的同一时期的实测流量的平均值,作为简单的季节性预报值,要比用性能较好的降雨径流模型得出的预报值好得多^[3]。这一理论的提出为后人提出大流域径流预报混合模型奠定了基础。1983 年, J·E·Nash 在格雷克等人研究的基础上,把降雨径流模拟的概念与纯粹的季节性预报方法结合起来,得到了更好的预报结果^[3]。他的研究成果表明,利用混合模型求得的结果要比单独用季节性预报模型或用较完善的降雨径流模型求得的结果好得多。因此,这类模型对具有明显季节性特征的大流域可能是很有用的^[3]。

1991 年梁庚辰等人提出多输入—单输出线性扰动模型及其识别方法在汇流方面的应用,并在中国长江上的两个不同的河段与一个支流流域上进行了验证。结果表明,这两类模型在河道汇流演算中都能给出很高的精度。对输入、输出关系复杂的大流域,线性扰动模型能够做出比简单线性模型好得多的预报^[4]。1995 年何少华在此基础上将多个水文系统模型在大流域汇流演进计算中进行了应用。系统模型较之概念性模型具有使用资料少、求解简便、可用严格的数字方法优化出多个输入的脉冲响应的优点^[5]。然而系统模型忽略了水文预报的物理意义,以数学方法进行计算,结果可能会导致很大的误差,甚至不符合预报结果的物理意义。

2004 年,李致家等人提出大流域的洪水预报和洪水调度管理需要采用水文学与水力学、确定性与随机性模型相结合的综合方法^[6]。对于降雨—径流模型通常采用新安江模型和经验预报方案进行流域流量过程预报;对于无资料的区间流域,采用水文模型进行流量预报。河道洪水演进采用一维或者二维非恒定流水力学计算方法。同时将这些确定性模型结合卡尔曼滤波或最小二乘递推模型等数学方法进行校正。这一理论虽然比较全面的概括出大流域预报的基本方法,但是由于大流域资料的不完整性,方法的复杂性以及预报人员经验的差异,真正将这些方法完全应用于生产项目中难度很大。

随着地理信息系统的蓬勃发展,近年来大流域预报研究又出现了一些新方法。2003 年陈利群等人提出运用 DEM 数据和 SWAT 模型对大尺度流域预报研究

[7]; 刘昌明等人运用 VIC-3L 模型对大尺度资料稀缺地区水文模拟进行研究^[8]。这些方法虽然能够解决一些问题,但是在使用中难度仍然很大。国内运用 SWAT 模型主要是靠软件来实现的,二次开发空间小,难以用于水情预报会商系统中; VIC-3L 模型对流域植被、土壤资料情况要求较高,这在现阶段在很多大流域也是难以实现的。

总之,国内外对于大流域洪水预报的研究较少,对于几十万平方公里流域洪水预报的研究更少,大多数研究仅仅停留在理论阶段,没有一种比较简单实用的方法能够真正用在大流域的预报中。

1.2.2 实时洪水校正模型

水文模型将复杂的水文过程加以概化成数学物理方程并在计算机上实现,这样整个流域的产汇流问题就归结为模型结构与模型参数问题。当流域的结构和参数确定后,系统就完全确定了,就能把任何输入过程确定地转化成输出过程。从这点来看,水文模型有它独特的优点^[9]。但预报模型在实际的应用中有时预报精度并不理想。这就要求在预报中加上校正模型,对预报结果进行校正以获得更高的精度。

实时洪水预报包括水文资料的实时传输、水文资料的实时检索、水文资料的实时插补、模型的预报以及实时校正^[10]。洪水预报实时校正方法大体上可以分为两类:预报与校正模型合一,一般的黑箱模型都是这样,如CLS模型;预报与校正模型耦合:如先用水文学模型或者水力学模型预报,而后采用误差自回归等方法进行实时校正。实时校正方法主要与所选用的预报模型有关。更确切地说,主要与预报模型的“数学表达形式”及“算法”有关。对线性系统模型已有许多卓有成效的实时校正模型,如卡尔曼滤波、递推最小二乘法、误差自回归模型等。

1.3 本文研究内容

针对目前大部分的方法尚处于理论研究阶段,对资料要求较高的情况,本文主要从生产需要出发,将理论与实际很好的结合起来,在前人研究的基础上对大流域洪水预报实时校正作进一步研究,以期能够找到一些简单实用的方法。

本文将研究解决以下两方面的问题:

(1)大流域预报建模的问题

大流域由于支流众多,计算有着严格的顺序,各水文模型对象的集成是建立洪水预报方案的关键所在。如何解决好模型之间数据的传递,提高系统运行的效率也是亟待解决的问题。目前国内外的研究还没有提出一个方便、有效的模型系统集成办法。研究一种方便的模型系统集成办法具有重要意义。

(2)大流域洪水预报实时校正的问题

要对洪水未来的变化趋势做出准确的预报,对预报结果实时校正是必不可少的。当前实时校正的方法很多,也取得了不错的效果,但是这些方法多数只是对于某一个小流域进行的。仅仅对大流域中一个小流域校正,其影响可能是微不足道的。大流域中还有很多支流,这些支流的校正结果对于最终出口断面可能是有影响的,必须探求一种方法能全面考虑大流域洪水预报实时校正的问题。本文这些做了详尽的研究比较工作,力求能找到一种便捷准确的能够对全流域进行实时校正以获得满意精度的方法。

本文第一章为总论,第二、三章为并列关系,在第二章使用 DEM 进行了产汇流分区和进行流域拓扑关系的确定,在第三章选择了实时校正模型并确定了模型的结构和参数,第四章运用这两章的结果进行了流域预报校正方案构建。

第二章 流域拓扑关系的确定

在流域水文模拟中,为了考虑降雨空间分布不均的影响及流域下垫面水文特性的空间差异,常常采用分散式水文模型。即将预报流域适当地分成多个单元流域,在各单元流域上进行产汇流计算,再将各单元流域的出流过程演算到流域出口断面(预报流域控制站),叠加后作为其流量预报过程。但是在具体实施过程中,大流域中各单元面积的计算是有严格顺序的,要先进行坡面产汇流计算,再进行河道汇流计算,只有确定了流域的拓扑关系,才能构建流域预报方案,高效的进行大流域预报计算。

为了确定流域拓扑关系,首先要提取流域下垫面信息,合理划分单元面积,对流域中的众多研究对象进行水文解析。这样才能将大流域中的众多对象划分为点对象、线对象和面对象,利用线对象的链接方式确定流域拓扑关系。

2.1 流域下垫面信息提取

在过去,流域下垫面信息的提取一般是采用人工方式进行流域信息的提取,操作费时费力而且精度得不到保证。随着数字地球概念的提出,各种以数字形式表达高程、土壤、土地利用、地质等信息的数据应运而生。这些数据是用一个任意坐标场中大量选择的已知 X , Y , Z 的坐标点对连续地面做一个简单的统计来表示地理信息的。当描述的地面特性是高程 Z 时,这种由 X , Y 水平坐标系统来描述的高程空间分布的数字地形模型称作数字高程模型 (Digital Elevation Model, 缩写为 DEM)。作为地球空间框架数据的基本内容和其它各种地理信息的载体,DEM 是各种地学分析的最基础的数据,也是 GIS 的基本内容。DEM 具有很大的实用价值,可对地形进行详细的描述,推求各种地形特征值。

俗话说水往低处走,不管是流域坡面上的产汇流过程,还是河道洪水波的演进过程,都遵循着这一规律。因此,在重力作用下,地形对径流的产生与运动,

起着决定性的作用。所以说, DEM 数据的普及为水文学的发展提供了新的数据源, 地理信息系统的发展与普及为水文学的发展提供了新的平台^[11]。

利用地理信息系统软件空间分析的功能, 以数字高程模型为基础, 可以使流域信息的提取更为方便、快捷。

2.1.1 DEM预处理

由于受高程数据采集误差(或 DEM 生成过程中的插值误差)及 DEM 格网大小的影响, 会使 DEM 中出现坑(或称洼地), 即某一网格或一组网格四周所有网格都高于这一网格或这一组网格。在这种情况下, 坑周围的网格方向将都指向坑, 在流水网中造成断点, 这意味着有些径流不能流出流域出口地面, 从而造成计算结果的不合理。为此, 需对 DEM 进行预处理, 对原始的 DEM 进行填洼处理。

2.1.2 流域流水网的推求

地形对流域汇流过程起着决定性作用。根据 DEM 数据, 对汇流过程进行描述的最好方法是流域流水网^[12]。每个网格上的径流, 将流向相邻网格中高度相对较低的网格, 据此可确定各网格的方向。将各网格方向首尾相连, 即为流域流水网。Beven 和 Kirkby^[13]和 Wharton^[14]都对排水网做了完整的定义。流水网可详细描述流域的径流汇流过程。

目前, 利用数字高程模型模拟流域地形、水系、集水面积、坡度等信息的方法很多, 很多软件也提供了这一功能。本文采用 ArcGis 软件作为推求流水网的主要工具。图 2-1 是利用 DEM 数据, 采用 ArcGis 软件生成的本文研究区域水系图。任何一种方法导出的流水网, 由于种种原因, 可能存在错误或误差。特别在平原地区, 由于地形差别不大, 这些误差和错误将导致生成的河网不能很好的代表真实河网。因此, 在流水网生成后, 需对流水网进行合理性检验。只有合理可靠的流水网, 才能以此为依据, 进行空间分析。模拟河网与蓝线河网(真实河网)拟合程度的好坏, 是检验流水网合理可靠的有效手段之一。图 2-2 是长江中上游

流域 1:25 万蓝线水系图分布图, 与图 2-1 相对比可看出, 模拟河网达到了较好的拟合精度, 能基本反映蓝线水系。为了更详细的对比分析, 图 2-3 中显示了模拟河网与 1:25 万蓝线河网的对比图。

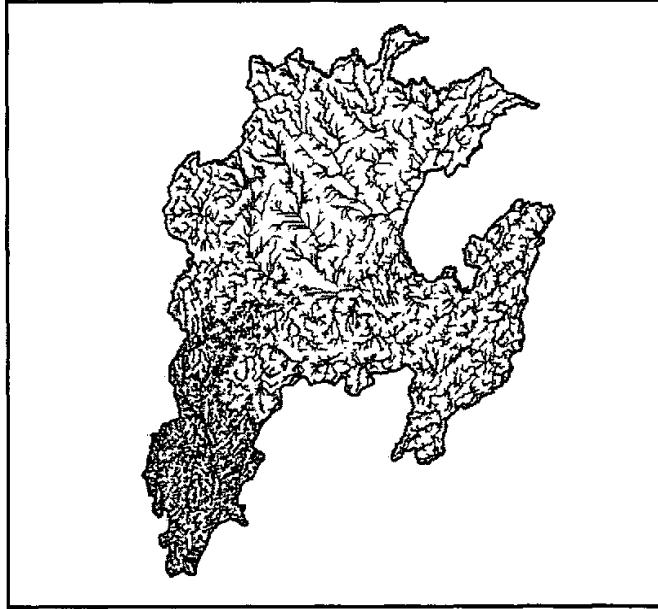


图 2-1 采用 ArcGis 软件生成的流域水系图

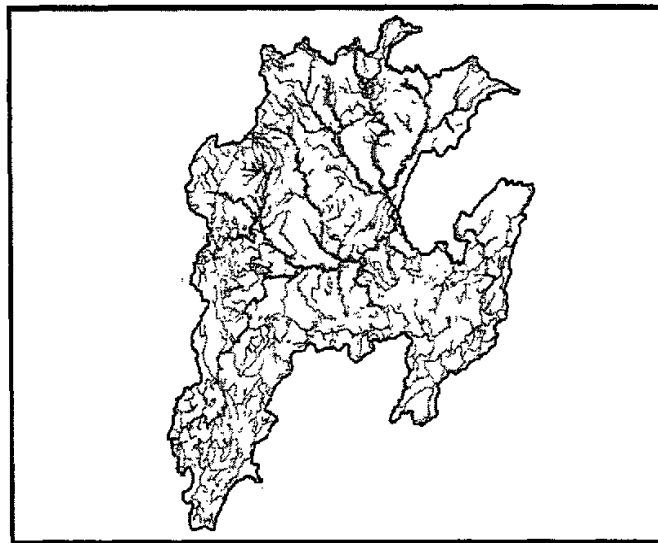


图 2-2 流域蓝线水系图

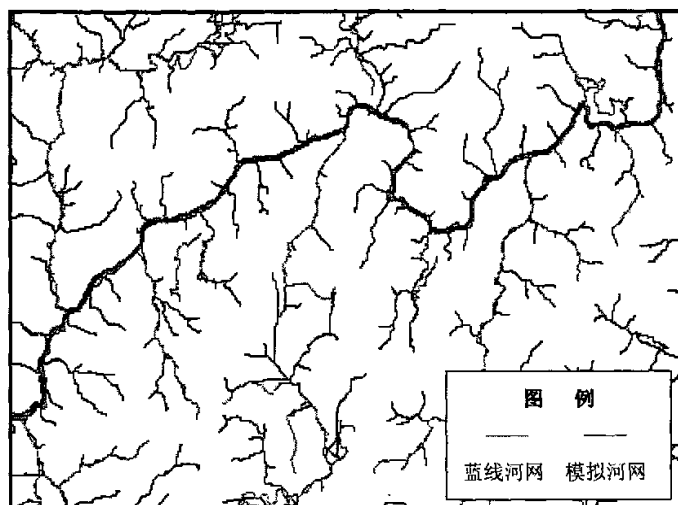


图 2-3 模拟河网与 1:25 万蓝线河网对比图

由于受 DEM 精度影响，一般情况下，高山丘陵地区地势起伏跌宕，利用 DEM 数据可以很好的模拟蓝线河网，达到了较高的精度，可以用于进一步的空间分析。而平原地区效果有可能差一些，能否用于进一步的分析还要进行详细的对照工作。

2.1.3 流域产汇流分区的建立

利用 ArcGis 软件能方便地按自然流域进行单元面积的划分^[15]，在流域流水分网生成后，只要给定集水面积的值，软件能自动按流域分水岭进行单元面积的划分。

但是软件在生成单元面积时，设置的集水面积值是对整个流域的，不能对流域的某一区域单独设置集水面积值，从而使整个流域上的单元面积大小都相差不大。但在实际的情况中，如果流域中的某一区域控制站较密，那么相应的单元面积数可以多点，就要将划分好的单元面积进行进一步的细分；但如果控制站较疏，那么相应的可以人工合并单元面积，增大单元面积面积。无论使用什么方法导出的流水分网，由于 DEM 数据本身的精度以及软件自身的问题，可能导致在流水分网基础之上生成的单元面积也存在同样的问题。在实际的应用过程中，要在地理信息的平台上，加以人工分析，根据蓝线河网校对划分的单元面积，以保证单元面

积和分水岭地划分正确。

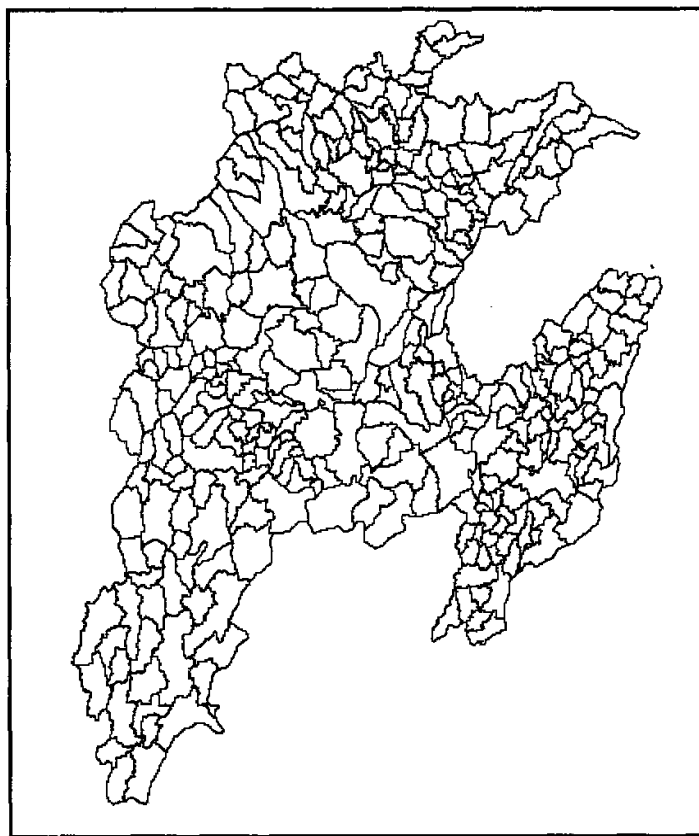


图 2-4 流域单元面积分布图

图 2-4 是利用流水网生成单元面积,再进行人工校对后得到的流域单元面积分布图,其空间范围是从长江干流金沙江龙街站以下至干流寸滩站,包括乌江流域、嘉陵江流域、岷江流域、沱江流域、金沙江下游、长江上游干流区。

在建立产汇流分区的时候,仍需同时考虑蓝线河网和生成河网,有时还需参考纸质的流域图,进行分析,考虑单元面积的划归。产汇流分区分为两种类型:小流域和区间流域。小流域是指没有上游来水,主要用降雨径流模型计算坡面产流和汇流的区域;区间流域指以上游来水为主,主要用河道演算模型计算河网汇流的区域。对于区间流域,划分时还需要考虑到河道测站的影响,尽量将上下游两个水文(水位)站之间的区域划分在一个产汇流分区内。这样做主要是为了更多的利用流域内的水位流量信息,达到更好的预报校正效果。图 2-5 是流域产汇流分区的分布图,在这一区域,一共划分了 53 个产汇流分区,包括 20 个小流域

和 33 个区间流域。

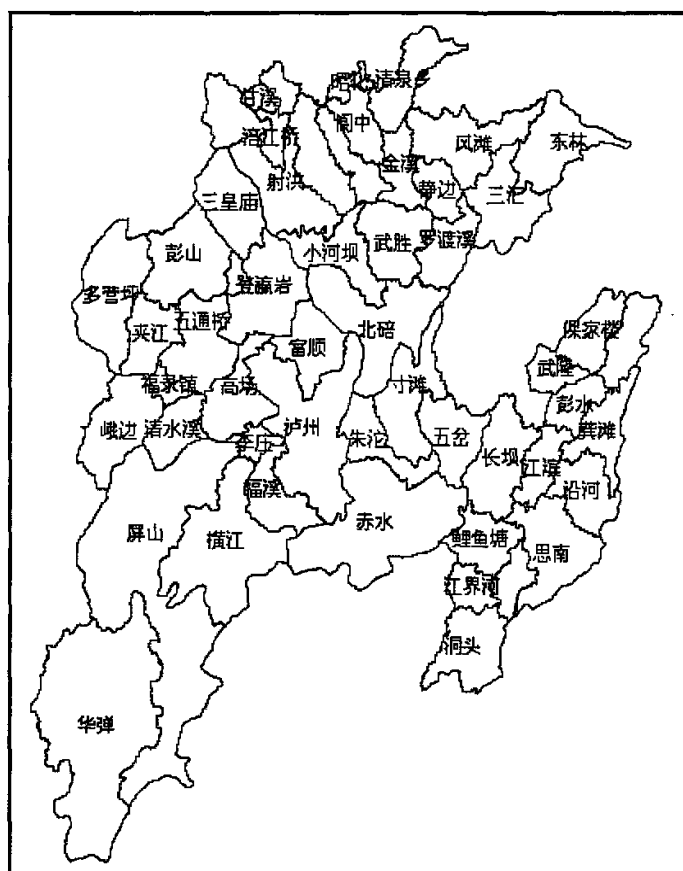


图 2-5 流域产汇流分区分布图

2.2 流域可视化构模

流域产汇流分区中有些分区需要采用降雨径流模型,有些分区中包含大的河流,有上下断面的流量资料可以利用,还需要采用河道演算模型。流域可视化构模就是要在地理信息的平台上根据采用模型的不同将流域中的研究对象抽象为一个个点、线、面,对这些点、线、面赋予不同的属性。如雨量站可以抽象为点,河道可以抽象为线,产汇流分区可以抽象为面。这样整个流域就可以抽象成一张概化图,对不同的对象可以采用不同的模型,达到高效率的预报。对这些对象赋予属性就可以将概化图以表格的形势储存在数据库中,方便使用。如何将这些分

区根据其采用模型类型不同划分不同对象,对象应该包括哪些属性,模型参数怎么确定,正是本节需要研究的问题。

2.2.1 相关水文模型介绍

水文预报模型一般分为确定性模型和随机模型,确定性模型又分为概念性模型和黑箱模型。概念性模型是由描述流域降雨径流形成的数学函数构成的一种数学物理模型,物理意义清晰,严格满足流域水量平衡原理^[16],但需要率定的参数较多;黑箱模型只关心模拟结果的精度,不考虑输入输出之间的因果关系。在过去的二三十年里概念性流域水文模型得到了长足的发展,在生产实践中,也得到了广泛的应用。我国赵人俊教授提出的新安江模型就是流域水文模型的一种,新安江模型是分散型模型,它把全流域分成很多单元面积,对每个单元面积的出流过程相加,得出单元面积的出口流量过程,再进行出口断面以下的河道洪水演算,求得流域出口的流量过程^[17]。

新安江模型总体结构可分成三个部分,产流计算、分水源和汇流。三者基本上是相互独立的。由降雨、蒸发和土壤含水量,计算出径流量;采用自由水蓄水量模拟的方法,将径流量分成地表水、壤中流和地下水三种水源;由于三种水源汇流过程中所通过的介质不同,具有不同的汇流特性,汇流过程可为三个阶段,单元面积内坡面汇流、单元面积内河网汇流、单元面积出口至流域出口断面的河道汇流。不同的汇流过程,采用不同的模拟方法。新安江模型的产流部分为蓄满产流模型,蒸发采用的是三层蒸发模型,河道洪水演算采用的是马斯京根河道演算法的线性解。

在集水面积较大的河段,其上游断面的来水量常比区间入流量大,所以在流域的洪水预报中,河道流量演算是比较重要的环节。河道流量演算模型的成功与否,关系到各单元流域洪水波演进规律(传播历时及变形)的模拟,组合结果决定着出口断面流量过程的所有位相,对整体预报方案的性能起到至关重要的作用^[18]。多数大流域区间来水量比较小,所以本文主要对河道流量演算方法进行了深入研究。

流量演算法是在圣维南方程组进行简化的基础之上,利用河段的水量平衡原理和蓄泄关系把河段上游断面的入流过程演化成下游断面的出流量过程的方法。换言之,该法是用以计算河槽调蓄作用对洪水波形状与运动的影响^[19]。这种方河段短期洪水预报和河道洪水分析计算中都被广泛采用。我国水文预报中常用的流量演算法有特征河长法、马斯京根法等,根据长江水利委员会多年长江水文预报中心的作业预报实践经验,常规洪水预报方案采用马斯京根河道演算法比较合理。

马斯京干法是美国一个叫麦卡锡(G.T.MaCarthy)的人于1938年提出的一种河道流量的演算方法,由于最先使用于美国的马斯京干河而得名。

马斯京干法认为槽蓄量方程的表达式为:

$$W = K[xI + (1-x)O] \quad (2-2-1)$$

$$\text{令} \quad Q' = xI + (1-x)O \quad (2-2-2)$$

$$\text{则} \quad W = KQ' \quad (2-2-3)$$

式中 Q' ——为示储流量 (m^3/s)

K ——为槽蓄流量关系曲线的坡度 (h)

x ——为流量比重因素 (无因次)

再结合水量平衡方程:

$$\frac{1}{2}(I_1 + I_2)\Delta t - \frac{1}{2}(O_1 + O_2)\Delta t = W_2 - W_1 \quad (2-2-4)$$

式中 I_1 、 I_2 ——计算时段始、末的河段入流量 (m^3/s)

O_1 、 O_2 ——计算时段始、末的河段出流量 (m^3/s)

可以推导出马斯京干法的计算公式:

$$O_2 = C_0 I_2 + C_1 I_1 + C_2 O_1 \quad (2-2-5)$$

其中参数 C_0 、 C_1 、 C_2 的计算公式如下，下式中的 K, x 的意义同上。

$$\begin{aligned}
 C_0 &= \frac{0.5\Delta t - Kx}{\nabla} \\
 C_1 &= \frac{0.5\Delta t + Kx}{\nabla} \\
 C_2 &= \frac{K - Kx - 0.5\Delta t}{\nabla} \\
 \nabla &= K - Kx + 0.5\Delta t \\
 C_0 + C_1 + C_2 &= 1
 \end{aligned} \tag{2-2-6}$$

对于有区间入流的情况，可将水量平衡方程改为：

$$\frac{1}{2}(I_1 + I_2)\Delta t - \frac{1}{2}(O_1 + O_2)\Delta t + \frac{1}{2}(q_2 - q_1) = W_2 - W_1 \tag{2-2-7}$$

槽蓄方程中的示储流量公式可改为：

$$Q' = xI + (1-x)O + q \tag{2-2-8}$$

其中 q ——区间入流。

式 (2-2-7) 就是马斯京干槽蓄曲线方程式。

根据式 (2-2-7)、(2-2-8) 可以推导出有区间入流的马斯京干法计算公式如下：

$$O_2 = C_0 I_2 + C_1 I_1 + C_2 O_1 + C_3 q_1 + C_4 q_2 \tag{2-2-9}$$

$$\begin{aligned}
 C_0 &= \frac{0.5\Delta t - Kx}{\nabla} \\
 C_1 &= \frac{0.5\Delta t + Kx}{\nabla} \\
 C_2 &= \frac{K - Kx - 0.5\Delta t}{\nabla} \\
 C_3 &= \frac{K + 0.5\Delta t}{\nabla} \\
 C_4 &= \frac{0.5\Delta t - K}{\nabla} \\
 \nabla &= K - Kx + 0.5\Delta t
 \end{aligned} \tag{2-2-10}$$

马斯京干法的基本出发点就是引用“示储流量” Q' 与蓄量 W 成单一线性关系, 而 Q' 则是入流 I 和出流 Q 的函数。

经过几十年的应用和发展, 马斯京干法已形成许多不同的应用形式, 主要由分段马法、变参数马法、非线性马法和分层马法等^[20]。

(1) 分段马法。为了避免出现负出流等不合理现象, 保证上、下断面的流量在计算时段内呈线性变化和在任何时刻流量在河段内沿程呈线性变化, 一般要求 $\Delta t \approx K$ 。1962 年赵人俊教授提出了马斯京干法分段连续演算法。

该方法将演算河段分成 N 个子河段, 每个子河段的参数 K_L, x_L 与未分河段时的参数 K, x 的关系为

$$K_L = \frac{K}{N}$$

$$x_L = \frac{1}{2} - \frac{N}{2}(1-2x)$$

分段连续演算的每段推流公式仍是 (2-2-9) 和 (2-2-10), 但式 (2-2-9) 中的 K 和 x 必须用分段后的 K_L, x_L 代替。

(2) 变参数马法。由于河道水力特性的非线性, 不同数量级的洪水传播时间和洪峰流量的衰减程度也不一样。为解决这一问题, 可以通过计算大量的历史洪水, 求出 $x = f(I_m)$ 、 $K = f(I_m)$, 并取 $N = \text{INT}(K/\Delta t)$ (式中, I_m 为入流站洪峰流量, INT 为取整函数) 计算时是根据上游站的洪峰流量选用本次洪水的参数 K, x , 然后再用分段马法进行推流。

变参数马法是利用 K, x 和 N 与入流洪峰 I_m 的关系来处理河道水力特性非线性问题。实质上, 除了当 $x=0$ (水库) 时, K 是入流的函数之外, 大多数情况下, K 和 x 都不仅是入流量的函数, 而且也是出流量的函数。换句话说。出流量是前期入流的函数。我们经常遇到, 在同一河段上, 一些同量级的洪水 (洪峰流量相当), 到达下游断面的传播时间和洪峰衰减程度大不相同, 主要原因就在于此。但在一般情况下, 点出的 $K \sim I_m$ 和 $x \sim I_m$ 关系比较散乱, 而且在使用时

由于洪水量级参数的不同,会导致洪水上下波动,不符合洪水预报的物理过程。

(3) 非线性马法。该法要求 K 、 x 与示储流量 Q 关系能够呈非线性的关系来表示,一般情况下要找到符合这种情况的河道才能使用该方法。

(4) 分层马法。该法是根据河道断面特性,确定出分层数目,然后把入流过程进行分层。演算是否需要分层,分层多少和各层参数的确定,一般是根据河道特性和实测流量过程线分析确定的。

这四种方法中,分段马法和变参数马法使用最为普遍。大流域在计算过程中节点众多,节点与节点之间距离较长,节点之间的河段采用分段马斯京根法进行演算比较合理。考虑到变参数马法会引起流量波动,而且可能引起水量不平衡的问题,本文在计算时主要采用分段马斯京根河道演算法。

在预报时还要考虑到计算时段长的选择,即系统进行预报的时间间隔。对于大流域计算时段长,宜短不宜长,在我国湿润地区一般取 1 小时,北方取 0.5 小时或者更短一些^[21]。

2.2.2 水文模型对象分析

水文模型是对流域上发生的水文过程进行模拟所建立的数学模型,对于复杂流域,如何将各水文模型应用对象集成才是建立洪水预报方案的关键所在。各水文模型对象的集成首先要从地理空间角度和水文角度对水文模型对象进行解析,分析研究各个模型对象的属性和方法,进而完成水文模型类的抽象。

对象的地理解析是指在地理信息系统平台上,将流域在地理空间上抽象概化成为点、线、面的集合体^[22],并赋与相应的属性。通过对水文模型对象的分析后,可以将其抽象成为点对象(测站、水量交汇点)、线对象(河段)和面对象(产汇流分区),这些对象都有其相应的属性。如点对象一般指测站,属性为站名、站码、经度、纬度等;线对象一般指河道,属性有河段名称、河段长度、上下游名称等;面对象一般指产汇流分区,属性有分区名、分区代码、分区面积(真实流域面积)等。

对象的水文解析是指结合水文模型应用实际,进一步扩充水文模型对象属性,其中较为关键的是分析流域水文特性,大致确定水文模型对象的模型参数属性。点对象除了其空间地理属性外,还有其水文属性,包括站码的确定、站类甄别等;线对象也有其水文属性,如对于河道汇流模型中的马斯京根法模型,其对应的模型对象是河段,那么此河段就要增加其河道演算的参数属性:子河段数、河段平均汇流时间、河段流量比重系数等;面对象的属性相对复杂,如新安江模型是一个分散性模型,考虑分散性,那么就应该在产汇流分区中继续进行单元面积的划分。现在对单元面积的划分比较合理的一种方式是按自然流域进行划分,这部分在前一节已经有过详细介绍。流域产汇流模型就是对每个单元面积进行计算,得出单元面积出口的流量过程,再进行出口以下的河道汇流过程,得出至流域出口的流量过程。将所有单元面积至流域出口的流量过程叠加,就求得了流域出口的总出流过程。每个单元面积进行产汇流计算,所需的水文属性包括:单元面积的“外界”输入(面雨量与蒸发量),单元面积的面积、单元面积出口至流域出口的汇流路径长度,单元面积产汇流模型参数。

综合上述对象解析过程,并结合流域洪水预报系统实际,进一步运用抽象、聚类、分解等信息建模方法,可先将水文模型对象主要分为三大基本类:测站、水量交汇点类、河段类、产汇流分区类。这些类的属性确定后就完成了水文模型的类抽象,这样洪水预报方案的构建将变得非常容易,以水文模型对象概化某一流域的过程,仅仅是结合该流域的水文模型类的实例化过程。

2.2.3 资料稀缺地区参数确定方法研究

在模型对象水文分析过程中,最重要的是分析流域水文特性,大致确定水文模型对象的模型参数属性。无论是线对象还是面对象,都有一些模型参数需要确定,由于采用的是概念性水文模型,参数具有明确的物理意义,一般并不难求出。但是在资料比较少的地区,其参数的确定方法还是需要进行一些研究工作的。

新安江模型参数简单,调试起来比较方便,对于资料稀缺地区新安江模型参数确定方法,井立阳^[23]已经经过详细研究,这里不再赘述。

由于大流域上游断面的来水量常比区间入流量大,所以在洪水预报中河道演算的精度如何直接关系到预报结果的精度。马斯京根法方程中仅有两个参数, x 表示上下游断面流量在槽蓄量中的相对权重,反映楔蓄对流量演算的作用,槽蓄作用大, x 小;反之, x 就大。 K 为蓄量常数。

对于 K 的性质和范围,大家的看法一般是一致的,在实际分段马法计算中, $\Delta t \approx K$ 。

而对于 x 的本质和数值变化范围,水文工作者的看法却不一致。按照 Yevjevich 和 Barnes(1970)的说法,参数 x 与模拟河段出口断面处流量的差别有关,它是通过测验河段的波形获得,原则上随着不同类型的波面不同。Cunge (1969) 提出应 x 从 0~0.5 的范围选择。Dooge (1973) 提出他同意 x 取负值的观点。在实际应用中, x 取负值并没有与任何条件下传递函数或模型脉冲相应相矛盾,可允许 x 的变动范围为 $x \leq 0.5$ [24]。

我们再从水力学的公式来看,由水力学求 x 的公式

$$x = x_1 - \frac{l}{2L} \quad (2-2-11)$$

式中 x_1 ——河道水面线的形状系数

l ——特征河长 (km)

L ——河长 (km)

当水面线为直线时 $x_1 = \frac{1}{2}$, 一般在计算时 x_1 就取 $\frac{1}{2}$; 特征河长 l 和实际的河长 L 均为正值,所以 x 的取值范围在 $(-\infty, 0.5)$ 。

求 K 和 x 值的方法很多,早期的方法比较麻烦,一般采用试算法,即由实测入出流过程逐时段摘录 I 和 O 的值,然后计算时段内蓄水变量 ΔW 和各时刻的槽蓄量 W ,假定不同的 x 值,按公式 (2-2-8) 计算 Q' ,并与同时刻的点 W 值点绘关系图,当涨洪段与落洪段基本合拢成一单一线,即求得需要的 x , $K = \Delta W / \Delta Q' * \Delta t$ 。实际工作中可以取多场洪水资料,分别按以上步骤求出 K 和 x ,然后进行分析和综合归纳,求出所需要的 K 和 x 。但是有些河段,不论 x 怎么取

值, $W \sim Q'$ 关系呈“8”字形, 而不是单一线; 有些河段的 $W \sim Q'$ 关系虽然呈单一线, 但是为曲线。试算法计算比较复杂, 效率不高, 且不容易编制参数自动计算的计算机程序, 本文不予采用。

通过水力学方法, 也可以计算 x 。某一特征河长的河段, 其下断面流量 O 与该河段槽蓄水量 W 呈单值关系, 即具有水库型性质。

特征河长的计算公式是^[25]:

$$l = \frac{O_0}{S_0} \left(\frac{\partial Z}{\partial O} \right)_0 \quad (2-2-12)$$

实际运用是采用上式的有限差形式:

$$l = \frac{O_0}{S_0} \left(\frac{\Delta Z}{\Delta O} \right)_0 \quad (2-2-13)$$

式中 O_0 ——下断面流量

S_0 ——比降

$\left(\frac{\Delta Z}{\Delta O} \right)_0$ ——水位流量关系曲线坡度

该方法对于 x 的推求理论性是比较强的, 利用式 (2-2-13) 计算时, 需要河段上下断面的水位流量关系, 但是在流域中, 河道节点众多, 资料往往难以查找齐全, 所以在最终确定 x 值时, 很多河段利用特征河长计算的 x 值只能给后面参数值的确定提供一个参考。

翟家瑞在《常用水文预报算法和计算程序》一书中给出的求一次洪水参数 K 和 x 的计算公式为^[20]:

$$x = - \frac{\sum_{i=1}^n O_i (I_i - O_i)}{\sum_{i=1}^n (I_i - O_i)^2} \quad (2-2-14)$$

$$K = \frac{\Delta t \sum_{i=1}^n O_i [\sum_{j=1}^{i-1} (I_j - O_j) + \frac{1}{2}(I_i - O_i)]}{\sum_{i=1}^n [x(I_i - O_i) + O_i]^2} \quad (2-2-15)$$

这个公式比较容易编制程序,使用起来也比较方便,但是在本文研究使用时,用它求得的参数值并不理想,很多断面的计算结果不符合其物理意义。

近年来由于计算机的普及,越来越多的数学方法被引用进水文专业,最小二乘估计法就是其中的一种,其原理简单介绍如下:

洪水预报中使用的线性系统数学模型有差分方程、响应函数、状态空间等形式,但归结到参数估计实施的时候,一般可将它们转化为通用的线性方程组的形式,即:

$$Z_N = \Phi_N \theta + e_N \quad (2-2-16)$$

式中: Φ_N 为系统输入矩阵, $N \times n$ 维; Z_N 为系统输出变量, N 维; θ 为系统参数向量, n 维; e_N 维模型噪声向量, N 维。

确定估计参数结果“最优”的最小二乘准则为:用估计的参数还原计算系统的输出估计值与已知值的误差平方和达到最小。相应的估计参数 $\hat{\theta}$ 就称为最小二乘估计值。

最小二乘估计准则的目标函数记为:

$$E\{\theta\} = \sum_{i=1}^N e_i^2 = e_N^T e_N = \min \quad (2-2-17)$$

根据最小二乘准则,式(2-2-17),可以推导出模型式(2-2-16)的参数计算公式:

$$\Phi_N^T \Phi_N \theta = \Phi_N^T Z_N \quad (2-2-18)$$

式(2-2-18)称为(2-2-16)的正规方程组或者法方程组,容易求出它的解向量 $\hat{\theta}$ 计算公式为:

$$\hat{\theta} = (\Phi_N^T \Phi_N)^{-1} \Phi_N^T Z_N \quad (2-2-19)$$

这就是 (2-2-16) 参数的最小二乘估计计算公式。

根据实测资料, 就可以利用以上方法求得最小二乘的系统输入矩阵。最小二乘法具有收敛性, 当使用有限的样本数的观测数据对系统参数进行估计时, 参数的估计值 $\hat{\theta}$ 随样本数的增大而逐渐接近 (收敛) 于系统的真实参数。

本文应用实测的出入流流量列出矛盾方程, 将最小二乘估计的原理用于 K 和 x 的求解。

根据公式 (2-2-11) 和 (2-2-13) 可以推导出以下公式:

$$\left(\frac{I_1 + I_2}{2} - \frac{O_1 + O_2}{2}\right)\Delta t = k[x(I_2 - I_1) + (1-x)(O_2 - O_1)] \quad (2-2-20)$$

将 (2-2-20) 展开, 得:

$$\left(\frac{I_1 + I_2}{2} - \frac{O_1 + O_2}{2}\right)\Delta t = kx(I_2 - I_1) + k(1-x)(O_2 - O_1) \quad (2-2-21)$$

令 (2-2-21) 中

$$Y = \left(\frac{I_1 + I_2}{2} - \frac{O_1 + O_2}{2}\right)\Delta t, \quad X_1 = I_2 - I_1, \quad X_2 = O_2 - O_1, \quad B_1 = kx, \quad B_2 = k(1-x)$$

$$\text{得: } Y = B_1 X_1 + B_2 X_2 \quad (2-2-22)$$

根据一场洪水过程资料可以列出很多这样的方程, 组成一个矛盾方程组, 利用最小二乘法估计参数 B_1 和 B_2 , 然后再求解二元一次方程组就可以求得 K 和 x 。然后对多场洪水分别求得 K 和 x , 进行综合分析最终确定其值。对于大部分的流域, 这一方法取得了较好的效果。

2.3 流域拓扑关系

防洪决策支持系统中的一个关键技术就是如何将相对独立的各个模型集成在一个大的系统内进行整体计算, 如何解决好模型之间数据的传递, 提高系统运行的效率。目前国内外的研究还没有提出一个方便、有效的模型系统集成的办法。

研究一种方便的模型系统集成办法具有重要意义^[26]。

各模型对象的集成,也就是各模型对象间的链接问题。只有搞清楚各水文模型对象间是如何链接的、对象间是如何传递数据的、对象的计算顺序又如何等问题才能将一个分散的、独立的水文模型对象组成有机的整体,才能构建流域洪水预报方案。水文模型对象的链接就涉及到了对象之间的拓扑关系。因此只要建立好水文模型对象间的拓扑关系,那么就可以解决复杂流域洪水预报方案中的水文模型对象集成问题^[11]。

2.3.1 流域拓扑关系中的对象

为了计算更为方便,在采用模型计算时,首先要对流域进行概化,使得复杂流域被抽象成为点、线、面对象的集合。对象的类型主要有产汇流分区(面)、河段(线)、测站(点)、水量交汇点(点)。不同类型的对象,使用的计算模型是不同的,如产汇流分区对象采用产汇流模型,可以是新安江模型、水箱模型等;河段对象采用水文学或水力学河道演算模型,如马斯京根法、一维非恒定流法等。流域的概化图中可以采用对象符号的形式表现,这里用P代表产汇流分区对象;R代表河段对象,采用水文学或水力学方法进行河道演算;S代表测站对象,N代表水量交汇点对象,图例如图2-9。

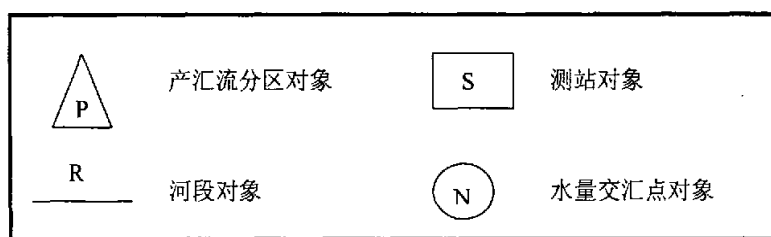


图 2-9 对象类型图例

2.3.2 对象链接方式

流域内各个对象之间的链接模式大致可分为三种:串联、并联及混联(既有串连又有并联)。对于某些简单流域,对象链接可能仅是串联,而对于复杂流域,一般各种链接模式都存在。图2-10表示了流域单元对象(P)、河段对象(R)、

测站对象(S)、水量交汇点对象(N)之间的链接关系,各对象符号的意义同前,箭头方向表示水流方向。

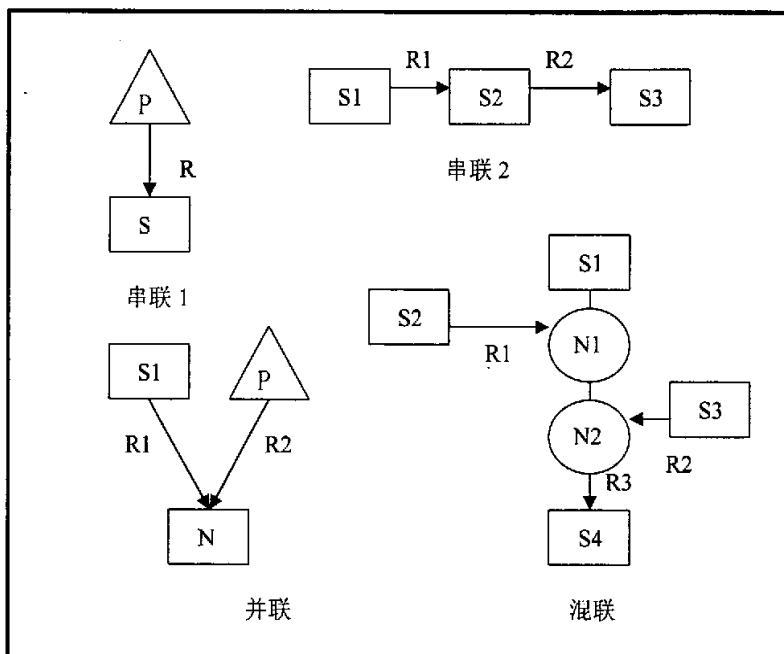


图 2-10 各对象类型的链接模式

图中各子图表示意义如下:

串联1所表示的流域是产汇流分区(P),是洪水产生的源节点,没有上游来水,产流后通过河道汇流流向出口控制测站;串联2所表示的是干流河道的洪水演进过程的概化。S1测站断面的输入为上游来水,经过R1河段的河道演算,到达S2测站,再经过R2河段的河道演算,到达S3测站断面,这种河系连续洪水演算过程同样也是串联模式;并联所表示的是有支流的河道洪水演进过程。P对象进行完产汇流计算后,在N处作为集中入流加入河道演算,这种有多个对象作为输入的称之为并联模式;混联所表示的是有多条支流的河道洪水演进过程,是一个复杂的演进过程。在干流河道上,S1测站断面的输入为上游来水,经R3河段的河道演算,到达S3测站断面,在演算过程中,有支流汇入(汇入的方式有集中入流与旁侧入流,还有更复杂的既有汇流又有分流),S1测站断面的输入是上游来水,经R1河段的河道演算,在N1交汇点入干流河道,那么对于N1交汇点其有两个对象的输入,为并联模式,同样N2交汇点也为并联模式,但干流河道的链接模式又为

串联，所以这种链接方式成为混联模式。

2.3.3 对象链接的拓扑表示方法

在建立洪水预报方案的时候，首先要将流域进行概化形成流域概化图。这张概化图能够反映洪水在流域内的演进方向，同时也能直观反映洪水演算的顺序，在上游断面没有计算之前，下游断面是不能进行计算的。最常见的情况就是在有支流汇入的情况下，如果支流的出流是未知的，即使知道上游断面的入流，下断面的计算结果也是不准确的。

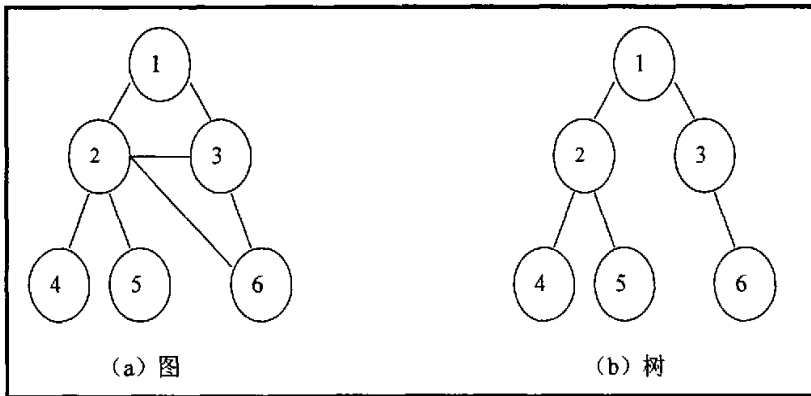


图 2-11 数据结构中的图与树示例

流域的概化图在数据结构上可以表示为数据节点图，也可以表示为树，图和树的区别在于：图中的任意两个节点之间都有可能相关，节点之间的邻接关系是任意的，见图2-11(a)；而在树形结构中，节点具有层次关系，每一层的节点只能与上一层中最多一个节点相关，见图2-11(b)。树可以认为是图的子集^[27]。

对于复杂流域的概化图，综合两种方法的优点后，可以采用“图”、“树”耦合的结构来表示概化图：用“图”来表示概化图中线对象的拓扑关系，用“树”来表示概化图中点对象的关系^[11]。

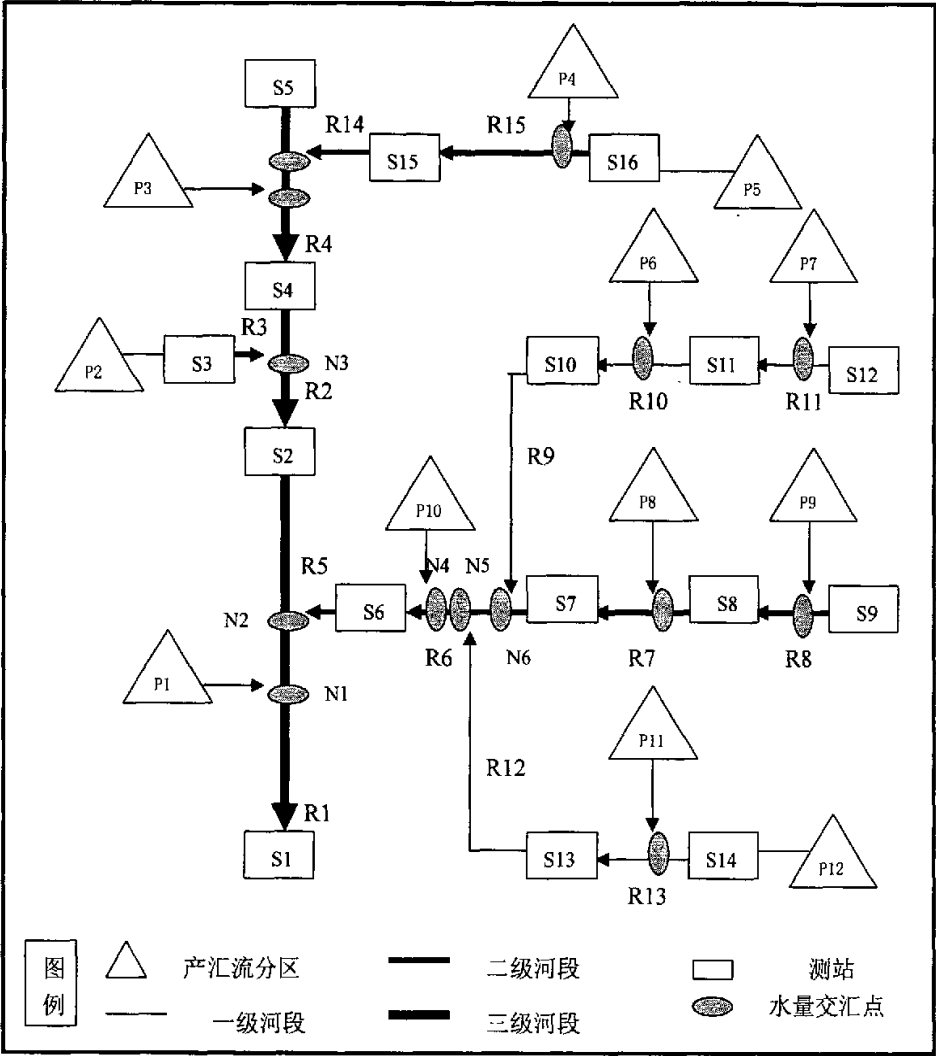


图2-12是基于水文模型对象的复杂流域概化图，要确定流域的拓扑关系，必须先对预报对象编码。因为流域概化图是由各个不同的对象组成的，包括产汇流分区、河段、测站、水量交汇点等，对象的编码方式可自行决定，但是不要遗漏对象，确保全流域对象都已编码；然后建立拓扑关系：用图邻接链表表示概化图中线对象拓扑关系，如表2-1；用树链接表表示概化图中的点对象拓扑关系，如表2-2。经过这些工作，河道概化图已经以表格形式存入数据库。

流域的拓扑关系确定后，就可以根据流域概化图编制预报方案，由于在预报方案构建中涉及到实时校正问题，预报方案的编制工作将放在第四章详细论述。

表2-1 线对象图邻接链表

链接对象	起始对象	终止对象
R1	S2	S1
R2	S4	S2
R3	S3	N3
R5	S6	N2
R6	S7	S6
R9	S10	N6
R12	S13	N5
R18	S19	N2
...	...	...
...	...	...

表2-2 点对象树链接表

根结点	左支	右支
S1	N1	
N1	P1	N2
N2	S2	S19
S2	N3	
N3	S3	S4
S3	P2	
S6	N4	
N4	P10	N5
N5	N6	S13
N6	S10	S7

2.4 本章小结

本章首先使用 DEM 数据进行了下垫面信息的提取工作，划分了单元面积，建立了产汇流分区。然后在地理信息平台上对水文模型应用对象进行了分析，将流域抽象概化成为点、线、面的集合体，并赋与相应的属性。对资料稀缺地区参数的确定方法也进行了一定的研究。最后对流域进行了模型类抽象并确定了流域的拓扑关系，使用“图”“树”结合的方式来进行流域概化图的编制工作。这一切工作已经为下一步的预报方案构建打下了基础。

第三章 实时校正模型和基本原理

无论什么样的预报模型,什么样的预报方案,预报值与实测值(流量或水位)之间总是存在一个差值。对于一个大流域,不仅出口断面有误差,每个水文站如果有实测流量(水位)信息,都会发现有误差,如何进行全面合理的修正,以获得比较高的预报精度,就是需要解决的问题。

在进行预报结果的校正时,要考虑到最新的实测资料对预报的影响,将这些最新的数据利用起来,以获得更好的校正结果。实时校正就是根据预报时刻的实时信息,对预报模型的参数、状态变量、输入向量或预报值进行某种修正,使其更符合客观实际,以提高预报精度。实时校正是实时洪水预报的核心之一,没有实时校正能力的模型,随着时间的推移,精度可能会越来越低。在大流域的预报中,每一个断面都有预报误差,如果不对上断面预报结果进行校正,最终出口断面的误差可能是相当大的。由此可见,对结果预报值进行误差实时校正是十分必要的。

实时校正有多种方法,但是大致上可以分为两类:预报与校正模型合一,一般的黑箱模型都是这样,如CLS模型;预报与校正模型耦合;如先用水文学模型或者水力学模型预报,而后采用误差自回归等方法进行实时校正。

在实时校正理论与方法的研究中,黑箱模型仅仅使用出口断面流量来进行实时校正,而对于流域内部许多宝贵的实时流量信息弃而不用,仅靠一个出口误差,包含信息太少,不论采用什么方法,都难以取得好的效果^[21],所以应重视实时校正方法与水文规律的有机结合,即重视预报与校正模型耦合的方式。

概念性模型和校正模型耦合的过程中,由于概念性流域水文模型使得水文模型呈现为较复杂的“隐式”结构,模型参数识别存在着计算的复杂性,很难用实测数据实时修正模型参数。目前在利用定常参数的流域水文模型进行洪水实时预报时,采用的有效方法是用确定性流域水文模型加上实时校正处理算法进行洪水实时校正预报即预报与校正模型耦合的方式。

在进行误差处理时基本上有三种方法,一种是用相邻时段的实测差值与预报差值之比来建立自动反馈模拟处理的 $FACT$ 因子校正系数法;一种是用时间序列分析探讨误差规律,对预报值进行逐时修正的 n 阶自回归模型计算法;另一种是用跟踪系统参数来实现对动态参数进行处理的方法,如卡尔曼滤波法、最小二乘法动态识别法^[28]。

3.1 反馈模拟实时校正模型

反馈模拟实时校正系统,就是一个系统把信息输出后,又将其作用结果以信息的方式返送回来,并对输入信息再输出,起到控制的作用^[29]。反馈模拟实时校正正是最大限度地利用了各种信息对后期预报值进行校正以提高预报精度的一种有效办法^[20]。

反馈法的原理及计算公式为^[29]:

首先对实测流量与预报流量进行相关分析,计算相关系数 R_c 和确定性系数 D_r ,

$$R_c = \frac{\sum_{i=1}^n (Q_{f(i)} - \bar{Q}_f)(Q_{ob(i)} - \bar{Q}_{ob})}{[\sum_{i=1}^n (Q_{f(i)} - \bar{Q}_f)^2 \cdot \sum_{i=1}^n (Q_{ob(i)} - \bar{Q}_{ob})^2]^{1/2}} \quad (3-1-1)$$

$$D_r = R_c^2 \quad (3-1-2)$$

式中 $Q_{f(i)}$ ——与实测流量对应的预报流量

$\bar{Q}_f$ ——预报流量的平均值

Q_{ob} ——实测流量

$\bar{Q}_{ob}$ ——实测流量平均值

n ——实测流量个数

然后求相邻时刻实测流量间的差值:

$$DQ_{ob(i)} = Q_{ob(i)} - Q_{ob(i-1)} \quad (3-1-3)$$

式中 $i \in (2, n)$, $DQ_{ob(1)} = 0$

预报流量间的差值为

$$DQ_{f(i)} = Q_{f(i)} - Q_{f(i-1)} \quad (3-1-4)$$

式中 $i \in (2, LAP)$, $DQ_{f(1)} = 0$, $LAP > n$

最后, 计算相邻两个时段实测流量差值之和与预报流量差值之和的比值 $FACT$ 因子, 表达式为

$$FACT(i-1) = \frac{DQ_{ob(i)} + DQ_{ob(i-1)}}{DQ_{f(i)} + DQ_{f(i-1)}} \quad (3-1-5)$$

对于 $FACT$ 的取值范围, 宜用 $FACT \in (0.45, 2.21)$

求 $F_{(i,j)}$, 其表达式为: $F_{(i,j)} = FACT(i)^{0.75^j}$, ($j=1, 2, \dots, 6$)

在实际水文预报反馈模拟中, 实时校正分涨水段和退水段, 当 $DQ_{f(i)} \geq 0$ 时, 为涨水段, 当 $DQ_{f(i)} < 0$ 时, 为退水段。

(1) 涨水段反馈模拟实时校正

①如果 $i-(n+6) \geq 0$, 且 $i > 7$, 则实时校正流量为:

$$Q_{ob(i)} = Q_{ob(i-1)} + DQ_{f(i)} \quad (3-1-6)$$

②如果 $i-(n+6) < 0$, 则实时校正系数

$$FAC = \frac{F_{(i-6,6)} + F_{(i-5,5)} + \dots + F_{(n,i-n)}}{7+n-i} \quad (3-1-7)$$

$$Q_{ob(i)} = Q_{ob(i-1)} + DQ_{f(i)} \cdot FAC \quad (3-1-8)$$

(2) 退水段反馈模拟实时校正

退水段反馈模拟实时校正计算公式为：

$$Q_{ob(i)} = Q_{f(i)} \cdot \frac{Q_{ob(i-1)}}{Q_{f(i-1)}} \quad (3-1-9)$$

(3) 当 $n=1$ 时的涨水段反馈实时校正计算公式为：

$$Q_{ob(i)} = Q_{ob(i-1)} + DQ_{f(i)} \quad (3-1-10)$$

在整个模型中，程序的输入是 LAP 个时段的预报流量和 n 个时段的实测流量；输出为实时校正后 LAP 个时段的流量，其中前 n 个为实测值，后 $LAP-n$ 个是根据实测流量与预报流量的关系对未来出现的预报流量做了校正的流量值。

反馈法是比较简单的方法，其物理意义明确，没有需要率定的参数，容易编制程序，所有的参数在程序中都可以直接计算出来。但是正是由于其数学理论性不强，所以在使用过程中误差的累计比较厉害，预见期比较短的时候还能取得一定的效果，预报外推长的话，几乎没有什么校正能力，有时候校正后的结果精度反而还会降低。但是由于这种方法与其他方法原理相差比较大，本文依然将其作为一种校正方法来研究比较。

3.2 水文时间序列模型

水文上常用的时间序列模型主要有三种：自回归(AR)模型，滑动平均(MA)模型，自回归滑动平均(ARMA)模型。其中最常用的为AR模型和ARMA模型。

3.2.1 自回归模型

自回归模型(autoregressive model, 即AR模型)是描述时间序列相依特性的数学模型。本模型早在上世纪60年代就广泛地应用于水文序列分析中，这是因为其具有时间相依的特性(即现在时刻的状态依赖于过去时刻的状态和形式)

和简单方便的应用形式。

设时间序列 $\{y(t)\}$ 为平稳序列，且是均值为 u ，方差为 σ^2 的正态分布，具有自回归相关。一般地，阶数为 p 的自回归模型 $AR(p)$ 可以表达为：

$$y(t) = \sum_{i=1}^p y(t-i)\theta(i) + \varepsilon(t) \quad (3-2-1)$$

式中 $y(t)$ 是相依时间序列（变量）， $\varepsilon(t)$ 与 $y(t)$ 独立，且是均值为 0，方差为 σ^2 的正态分布序列。系数 $\theta(i)$ 为自回归系数，这些系数可由实测资料估计或率定。

3.2.2 自回归滑动平均模型

自回归滑动平均模型，简记为 ARMA 模型，是在自回归成分 AR 上再加一个滑动平均成分 MA 来表示。一般地，一个由 p 阶 AR 模型 ($AR(p)$) 和一个 q 阶 MA 模型 ($MA(q)$)，可以组成一个 (p, q) 阶的自回归滑动平均模型，简记为 $ARMA(p, q)$ 模型，其数学表达式为：

$$y(t) = \sum_{i=1}^p y(t-i)\theta(i) + \varepsilon(t) - \sum_{j=1}^q \varepsilon(t-j)\varphi(j) \quad (3-2-2)$$

式中系数 $\varphi(j)$ 为滑动平均系数， $ARMA(p, q)$ 模型共有 $p+q+2$ 个参数需要估计。

对于 $AR(p)$ 序列，只要知道 $x_t, x_{t-1}, \dots, x_{t-p+1}$ 这几个历史数据无需知道更早的数据就可以对该序列作任意步的平稳最小方差的预报。但是对 $MA(q)$ 序列，当数据有限时， MA 序列只能近似的平稳最小方差预报值。因此 MA 序列的预报需要有充分多的历史数据。这就是 MA 序列和 AR 序列的根本区别。计算 AR 序列的预报值时需要对大量的历史数据加权求和，且每次取得新数据后就必须这样算一次，使预报过程中数据存储量和计算工作量不断增加，由此可见，在水文实时预报中，单独采用滑动平均序列做预报是不太适宜的^[30]。

$ARMA(p,q)$ 序列与 $MA(q)$ 序列不同, 不像 $MA(q)$ 序列那样, 超过 q 步的预报值并非为零, 因此在实时预报中, 宜采用自回归或自回归滑动平均模型^[30]。

两种自回归模型中, AR 模型比较适宜于流量、水位或其他预报对象的变化过程较缓的预报, 如主要受地下水、流域和河网蓄水量支配的枯季预报, 洪水的退水段或涨落段比较平缓的河道预报。当区间来水较多、河道汇流并不占主导地位时, 宜在自回归成分上再加上滑动平均成分, 即采用 $ARMA$ 模型。

一般来讲, $ARMA$ 模型由于其参数相对较少而且模型有较大的弹性, 在使用过程中可以取得良好的效果, 但是如何确定自回归和滑动平均部分值得研究。李纪人^[30]认为, 自回归部分可有流域出口断面的流量序列组成, 而上游断面的流量序列或流域上各实测雨量站的雨量序列组成滑动平均部分。平原地区雨量随地形变化不大, 用实测雨量站的雨量组成滑动平均部分可以取得不错的效果, 而西南各省由于流域内地形变化大, 雨量站大都设在河谷附近, 虽然分布还算均匀, 但对面雨量控制比较差。要解决这个问题需要在无人居住的高山地区大大增加雨量站的密度才行, 这在实际上是很难办到的。根据目前的情况来看, 对于地形变化比较大的流域应主要把流量利用起来, 用实测流量组成滑动平均部分。

自回归模型是一种使用比较广泛的模型, 要按“预报”用途进行修改, 增加实时校正的能力, 最简单的方法是附加一个误差自回归模型^[21], 这一方法在大多数流域取得了良好的效果。

3.3 参数动态识别实时校正模型

参数动态识别实时校正模型是指实时校正模型能够根据最新的信息实时改变校正模型的参数, 达到最佳预报的效果。参数动态实时校正模型最常用的两种方法为卡尔曼滤波动态识别法和实现最小二乘在线递推的自适应实时校正法。

卡尔曼滤波动态识别法(简称卡尔曼滤波法)因对系统的状态变量进行最优估计, 既可以达到最小方差, 又不损失预见期, 是一种比较理想的实时校正方法。在实时洪水预报中可选择作为状态变量的有洪水预报模型的参数、预报对象和预报误差等。卡尔曼滤波法实质上是一种线性无偏最小方差估计, 可用于任何线性

随机系统,并可综合处理模型误差和量测误差^[33]。

卡尔曼滤波在工业自动控制中和国防尖端科学等领域中,显示出了明显的效果,但是在水文领域在很多条件下使用效果并不好。原因在于前者可利用的信息量大,简单校正技术不能充分利用这些信息,卡尔曼滤波的精细结构刚好满足了这些领域校正的需要;而洪水预报误差信息量,只够提供简单修正技术,甚至连简单修正方法所需的信息量也不够,将复杂的修正技术应用在该领域,没能增加可利用实时信息量,修正效果自然也就难以提高^[34]。根据包为民^[35]的研究结果:该方法虽然设计思路科学,设计结构精细,但用于水文预报中,由于能利用的实时信息量限制,往往效果与简单的自回归方法相近。目前国内虽然研究卡尔曼滤波的人很多,但是真正能将其在生产项目中使用并能够取得良好效果的并不多。

根据以上分析,由于卡尔曼滤波模型与自回归模型校正效果相近,可以采用自回归模型(或自回归滑动平均模型)与在线参数估计方法相结合的自适应实时校正模型进行校正。

自回归模型参数校正分为在线校正和离线校正,校正参数的估计一般采用最小二乘法或极大似然法。离线校正的自回归参数是根据历史洪水事先估计好的,校正的结果比较稳定,反映了洪水资料的平均情况。在线校正能够根据最新输入与输出信息,给现时预报误差一定的权重以校正模型参数来进行实时预报的,反映了预报时的最新情况。自回归模型在线校正属于参数在线识别(也称动态识别),能反映预报时刻的校正模型参数状态^[33]。该法既拥有动态识别参数的优点,又有自回归模型简单易用的优点,是一种比较好的校正方法,可以广泛的应用于洪水的校正。它可以根据各场实际洪水的变化自适应地调整参数,提高洪水预报精度,故可称为自适应实时校正法。

最小二乘估计是一种应用十分广泛的参数估计方法,对于不同的系统模型,需要选择不同的递推算法,以便更好地估计模型参数;极大似然估计法比最小二乘估计方法理论上更为严密,而且能够比较高地估计精度。但是,极大似然估计算法要复杂得多,而且要求样本容量大。由于似然函数最大是从概率意义上获得的,因此在大量的数据基础上增加一组新数据对似然函数的影响是很难估计的。所以,最大似然法没有严格意义上的在线算法^[32]。所以在线估计方法采用最小

二乘法比较合理。

最小二乘法的在线估计方法主要有以下几种：

(1) 无限记忆在线最小二乘估计法

对于一个运行中的系统，如果只是使用离线最小二乘估计，就不能够运用最新的观测数据获得最新的系统输入矩阵，根据历史数据离线求出的输入矩阵在使用一段时间以后，可能已经无法适合现在的情况，也就不能对现在的情况进行准确的预报。最小二乘在线递推算法正是为了解决这个问题而研制的。

式 (2-3-16) 的矩阵表达形式为：

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix} = \begin{bmatrix} \varphi_{11}\varphi_{12}\cdots\varphi_{1n} \\ \varphi_{21}\varphi_{22}\cdots\varphi_{2n} \\ \vdots \\ \varphi_{N1}\varphi_{N2}\cdots\varphi_{Nn} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix} + e_N \quad (3-4-1)$$

$$\text{令 } \Psi_i = [\Psi_{i1}, \Psi_{i2}, \dots, \Psi_{in}]^T = \Phi_N, P_N = (\Phi_N^T \Phi_N)^{-1}$$

有 N 组观测数据进入程序，已经得到了它的离线最小二乘估计值

$$\hat{\theta} = (\Phi_N^T \Phi_N)^{-1} \Phi_N^T z_N \quad (3-4-2)$$

当第 $N+1$ 组新观测数据进入程序之后， Φ_{N+1} 在 Φ_N 的基础上增加了新的行向量，输出向量 z_{N+1} 比 z_N 也增加了一个新元素 z_{N+1} 。这时，可以推导出根据 N 组观测求出的最小二乘估计值 $\hat{\theta}_N$ 和根据 $N+1$ 组观测求出新的估计值 $\hat{\theta}_{N+1}$ 的递推关系^[32]。

$$P_{N+1} = P_N - P_N \Psi_{N+1} r_N \Psi_{N+1}^T P_N \quad (3-4-3)$$

$$\text{式中 } r_N = (1 + \Psi_{N+1}^T P_N \Psi_{N+1})^{-1} \quad (3-4-4)$$

式 (3-4-3) 得到了从 P_N 阵递推到计算 P_{N+1} 阵的公式。因为 P_N 阵实际上包含了在 $1 \sim N$ 组输入观测数据 Φ_N 的全部信息，故称 P_N 为记忆矩阵。

同样，也可以导出 $\hat{\theta}_{N+1}$ 与 $\hat{\theta}_N$ 的递推关系。

$$\hat{\theta}_{N+1} = \hat{\theta}_N + P_N \Psi_{N+1} r_N (z_{N+1} - \Psi_{N+1}^T \hat{\theta}_N) \quad (3-4-5)$$

式 (3-4-3)、(3-4-4) 和 (3-4-5) 就构成了从 $\hat{\theta}_N$ 计算 $\hat{\theta}_{N+1}$ 的最小二乘算法的递推公式。

(2) 有限记忆最小二乘估计法

从在线最小二乘估计公式推导可见，它与离线估计完全等价，仅在实现方式上运用了递推的技巧。每增加一组新观测数据后，如果要计算系统参数的变化，就要使用离线估计方法把全部的数据重新进行一次计算，工作量大，速度低。在一组新的观测数据 Φ_{N+1} 、 z_{N+1} 进入到记忆阵 P_{N+1} 和 $\hat{\theta}_{N+1}$ 的计算结果中去后，新的数据和旧的数据在后续递推中永远同时存在，不会消失，故这种方式被称为具有“无限记忆”的递推方法。当 N 不大时，每一组新数据都对 P_{N+1} 和 $\hat{\theta}_{N+1}$ 具有一定的影响和修正作用。但是，当 N 到达比较大的数量级，如上万个时，新的数据（新息）的修正作用也就越来越小了， $\hat{\theta}_{N+1}$ 与 $\hat{\theta}_N$ 的差异可以说是微乎其微的，在线最小二乘估计也就失去了意义。

从“无限记忆”这一特征出发，对在线最小二乘估计方法的一种改进就是实现“有限记忆”的递推。即：设定一个系统记忆长度 M （即记忆观测数据的组数，也可以代表一个时间长度）。当有一组新的观测数据进入识别程序时，把原先记忆的 M 组数据中最早进入的一组数据同时舍弃，保持用 M 组数据进行最小二乘估计。早于 M 组的观测数据对于新的参数估计就不再起作用，这种方式估计所得的参数永远是反映当前最新的 M 组观测数据的结果。

有限记忆最小二乘法公式为^[32]：

$$P_{N+1}^* = P_{N+1} + P_{N+1} \Psi_1 (1 - \Psi_1^T P_{N+1} \Psi_1)^{-1} \Psi_1^T P_{N+1} \quad (3-4-6)$$

$$\Delta z_{N+1} = z_{N+1} - \hat{z}_{N+1} = z_{N+1} - \Psi_{N+1}^T \hat{\theta}_N \quad (3-4-7)$$

$$\hat{\theta}_{N+1} = \hat{\theta}_N + P_{N+1}^* (\Psi_{N+1} \Delta z_{N+1} - \Psi_1 \Delta z_1) \quad (3-4-8)$$

式 (3-4-3), (3-4-6)、(3-4-7) 和 (3-4-8) 就是有限记忆最小二乘估计的递推公式。在递推公式里, 记忆长度 $M=N$

有限记忆与无限记忆差别在于:

参加递推计算既要记入信息 Δz_{N+1} , 要去掉最初一组数据信息 Δz_1 的影响。所以, 有限记忆方法需要将 M 组数据始终保留在计算机中, 用来完成加入、剔除的计算。本法对参数的修正取决于 Δz_{N+1} 与 Δz_1 这两组信息的加权差值, 当这个差值特别大或者观测噪声较大时, 识别的参数可能出现跳动。而且, 记忆长度 M 的选择也有很大关系: M 过大, $\hat{\theta}_{N+1}$ 的变化对新息反映不够灵敏; 选择过小, 参数不够稳定, 易出现跳动。

(3) 衰减记忆最小二乘估计法和时变遗忘因子最小二乘法

鉴于有限记忆的最小二乘法的的一些不足, 人们又研制了衰减记忆最小二乘法, 该法是将最小二乘目标函数的等权结构改为加权结构, 加大新息在推导参数的作用。根据以上思想, 将目标函数 (2-3-17) 改为:

$$E\{\theta\} = \sum_{i=1}^N \rho^{N-i} e_i^2 = \min \quad (3-4-9)$$

式中: ρ 为加权因子, $0 < \rho \leq 1$, 若 $\rho = 1$, 就是等权的情况, 相当于无限记忆最小二乘法。所以加权最小二乘法在更高一层次上概括了普通等权方法。当从 N 组观测数据向 $(N+1)$ 组递推时, 按照 (3-4-13) 的原理, 最新观测数据 Ψ_{N+1} 和 z_{N+1} 的权重是 1, 过去的 Ψ_N 和 z_N 则必须加乘一个系数 ρ , 使其权重衰减, 当 Φ_N 向 Φ_{N+1} 递推时:

$$\Phi_{N+1} = \begin{bmatrix} \rho \Phi_N \\ \Psi_{N+1}^T \end{bmatrix}, \quad z_{N+1} = \begin{bmatrix} \rho z_N \\ z_{N+1} \end{bmatrix}$$

根据 P_{N+1} 的定义, 可以导出

$$P_{N+1} = \frac{1}{\rho^2} [P_N - P_N \Psi_{N+1} (\rho^2 + \Psi_{N+1}^T P_N \Psi_{N+1})^{-1} \Psi_{N+1}^T P_N] \quad (3-4-10)$$

$$\hat{\theta}_{N+1} = \hat{\theta}_N + P_N \Psi_{N+1} (\rho^2 + \Psi_{N+1}^T P_N \Psi_{N+1})^{-1} (z_{N+1} - \Psi_{N+1}^T \hat{\theta}_N) \quad (3-4-11)$$

式(3-4-10)和(3-4-11)就是衰减记忆最小二乘估计递推公式, $\lambda = \rho^2$ 称为遗忘因子。当 $\rho = 1$ 时, 其形式与无限记忆公式完全相同。

衰减记忆最小二乘算法由于它固有的特性, 仅适用于缓时变系统的参数识别, 对于非确定性动态水文系统模型参数跟踪乏力, 容易形成参数在线跟踪的滞后性^[36]。在衰减记忆最小二乘估计的基础上, T.R.Fortescue 在文献[37]中提出时变遗忘因子最小二乘法, 对系统模型时变参数进行在线识别。

时变遗忘因子最小二乘法是在式(3-4-10)和(3-4-11)的基础上添加了一个时变遗忘因子 $\lambda(N+1)$

$$\lambda(N+1) = 1 - \frac{(\hat{z}_{N+1} - z_{N+1})^2}{(1 + \Psi_{N+1}^T P_N \Psi_{N+1})R} \quad (3-4-12)$$

式中 $\hat{z}_{N+1}$ 为 $N+1$ 时刻的预报值, R 为误差的加权平方和。为防止意外干扰使 $\lambda(N+1)$ 变化过大, 必须对 $\lambda(N+1)$ 加以限制: 当 $\lambda(N+1) > \lambda_{\max}$ 时, 令 $\lambda(N+1) = \lambda_{\max}$, 当 $\lambda(N+1) < \lambda_{\min}$ 时, 令 $\lambda(N+1) = \lambda_{\min}$ 。模型试验结果表明, 这种递推算法能很快地跟踪时变参数的变化^[37]。

可见, 时变遗忘因子最小二乘法是各种最小二乘法中比较科学的一种方法。实时校正可以采用自回归法和时变遗忘因子最小二乘法耦合的方式, 根据各场实际洪水的变化自适应地调整其遗忘因子, 以达到最佳跟踪参数的效果, 提高洪水预报精度。

本文参数动态识别法拟采用自适应实时校正方法。

3.4 实时校正模型结构和参数的确定

3.4.1 实时校正模型结构的确定

在实时校正模型的类型基本确定之后, 有些模型是需要确定模型的结构,

即阶数，有些模型还需要估计参数。

反馈法是一种比较简单的方法，不需要确定模型的结构，也不需要参数进行估计，根据预报值和实测值就可以方便的进行校正。本节主要对如何确定自回归模型的结构进行了讨论（因自回归滑动平均模型与自回归模型相似，模型结构的确定可以依此类推）。该法同样适用于自适应实时校正模型中的自回归方程。

模型结构确定的方法很多，如：用样本自相关函数和偏相关函数的统计特性确定模型的阶数、准则损失函数法、残差方差分析法、残差的白色性检验法、最终预报误差准则法（FPE）、信息准则法（AIC）、BIC 法、汉恩—利萨宁线性定阶法等。

限于篇幅所限，这里着重介绍比较常用的最终预报误差准则法（FPE）和信息准则法（AIC）。

（1）最终预报误差准则法（FPE）

这一方法是由赤池提出来的，其基本精神是：当采用最小二乘法估计 AR 模型的参数时，以一步预报误差的均方值达到最小的阶数作为用 AR 模型拟合实测数据的最好阶数。

考虑一下单变量 $AR(p)$ 模型：

$$y(t) = \sum_{i=1}^p y(t-i)\theta(i) + \varepsilon(t) \quad (3-4-1)$$

用它做一步预报，设预报的误差为 $e(t)$ ，误差方差为 σ^2 ，第 K 步的预报误差方差为 σ_k^2 。赤池导出的最终预报误差估计的方法如下^[30]：

$$FPE(K) = (1 + \frac{K+1}{N})(1 - \frac{K+1}{N})^{-1} \sigma_k^2 \quad (3-4-2)$$

式中 N 为数据长度， k 为所选的自回归模型的阶数。

对 σ_k^2 的 Levinson-Durbin 渐进递推公式为：

$$\sigma_k^2 = [1 - \theta_k^2] \sigma_{k-1}^2 \quad (3-4-3)$$

$$\sigma_0^2 = r_0 \quad (3-4-4)$$

式中, $\hat{r}_k = E[y(t+k)y(t)]$ 。

这样, 在确定阶次时可取一正整数 M , 使之大于等于 AR 模型真正的阶次 p , 在 $0 \leq k \leq M$ 的范围内选择 k 和计算 FPE, 使之最小的 k 就是模型阶次 p 的估计值。

(2) 信息准则法 (AIC)

1974 年赤池提出了一种具有客观标准的判阶方法, 这一标准就是赤池信息准则 (AIC), 用它可以比较客观的确定模型的阶次。

考虑 (3-4-1) 式的线性模型, 为确定其阶数 N , 赤池引入以下准则:

$$AIC = -2 \log[L(\hat{\theta}_{ML})] + 2\hat{n} \quad (3-4-5)$$

式中, $\hat{\theta}_{ML}$ 为参数 θ 的最大似然估计值; $L(\hat{\theta}_{ML})$ 为在 $\hat{\theta}_{ML}$ 条件下的似然函数; $\hat{n}$ 是模型阶数的估计值。赤池证明使 AIC 最小的 $\hat{n}$ 就是相对合理的模型的阶数。

在确定模型阶次时应从两个方面考虑: 模型阶次升高, 则模型阶次逼近真实系统的准确性高, 预报残差下降, 这是有利的一面; 然而, 模型阶次的升高意味着模型参数增多, 可能导致计算误差增大, 这是不利的一面。因此, 综合两方面的影响, 应该存在一个较为合适的阶次, 阶次相应的模型就是适用模型, 这就是模型识别的准则^[38]。在平稳序列 AR 模型拟合的定阶分析中, 目前应用最广泛的是 Akaike 信息准则 (AIC)^[39]。

由 AIC 本身包含的信息内容可以知道, AIC 定阶法是最大似然函数的有一种应用。在确定模型阶次时, 可采用最大似然法做参数估计。对于每一个阶次, 都有一套用最大似然法估计的参数, 标准是使 AIC 最小为最佳。如果两个不同阶次的模型模拟程度一样, 那么阶数低者为佳。

然而, 在实际中洪水预报中, AIC 准则应用起来是非常困难的。

首先, 使用 AIC 准则计算量比较大。无论是定阶还是率定参数都需要大量

的洪水资料, 这些资料应该也是比较有代表性的洪水过程, 包括小型洪水、中型洪水、大型洪水。各场洪水历时是不同的, 小型洪水历时短, 大型洪水历时长。大江大河(如长江)的洪水一般持续时间比较长, 一场大洪水从起涨到最后回落到正常水位一般需要一周、十天甚至更长时间。如果采用 1 小时为计算时段, 持续一周的洪水就需要对一阶、二阶、三阶直至 168 阶回归模型分别用最大似然估计法估计参数, 然后确定使 AIC 最小的参数。如果洪水持续十天甚至半个月, 计算量更大。

其次, 使用 AIC 准则确定的模型结构不具有普遍性和外推性。即将发生的洪水, 一般不能准确预见其发生的量级和历时, 也就不能确定采用几阶的模型比较合适。

所以, AIC 准则对于洪水预报的自回归模型是不适用的。

人们在实际应用中发现, 模型一般采用二阶或者三阶即可满足要求^[40], AR 模型的阶数略高并无很大的害处, 实时上也并不存在什么真正的阶数^[28]。

根据以上研究成果, 在实际应用中可以分别利用洪水资料建立了 3 阶、6 阶、12 阶、36 阶、72 阶甚至更高阶数的自回归模型, 然后根据这些模型对多场洪水做一步预报, 即用第 k 步的预报值和实测值和自回归模型预报第 $k+1$ 步的值, 然后再利用第 $k+1$ 步的预报值和实测值预报第 $k+2$ 步的值。分别计算各回归模型预报的平均确定性系数, 通过比较确定最终需要使用的模型结构。

3.4.2 实时校正模型参数的确定

(1) 自回归模型参数的确定

在校正预报时如采用 6 阶一次的模型结构, 这样对自回归模型就需要确定七个参数(在实际计算中, 为了程序编写的方便性, 使用字母 b 代替回归参数 θ)。误差自回归的表达式为:

$$y_t = b_0 + b_1 y_{t-6} + b_2 y_{t-5} + b_3 y_{t-4} + b_4 y_{t-3} + b_5 y_{t-2} + b_6 y_{t-1} \quad (3-4-6)$$

$$Q(t) = Q_c(t) + y_t \quad (3-4-7)$$

式中, y_t 为 t 时刻的误差值, 是根据前六个时段的误差计算所得, $Q_c(t)$ 是 t 时刻的模型计算值, $Q(t)$ 是经过校正后 t 时刻的预报值。

建模过程跟踪参数估计的方法也很多, 视建模的要求可以采用初步估计法, 如相关矩估计; 也可以采用比较精确的方法, 如最小二乘估计, 极大似然估计等。考虑到实时校正预报的精确性和方便性, 参数离线和在线估计均可采用最小二乘法。

(2) 记忆因子初值的选择

由统计知识知, 当 $\rho=0.6$, 在最小二乘估计中起作用的旧数据不超过 9 组; 当 $\rho=0.9$, 则有 30-40 组旧数据起作用; 当 $\rho=0.99$ 可延长到 400 多组, $\rho=0.999$ 可记忆到 3000 多组, $\rho=0.9999$ 则可以记忆数据万组以上。

目前 ρ 的选择没有定量的方法, 选择大了不能突出新息的作用, 选择小了可能引起系统的不稳定, 导致校正值的跳动, 兼顾对于参数变化的跟踪能力和过滤平滑测噪声的能力, 是选取最佳 ρ 的原则。实际洪水系统一般选择 $\lambda=0.9\sim 1.0$ 之间^[32]。考虑到突出新息在参数估计中的作用, 记忆因子初值给定在 0.9, 然后在计算过程中时变的识别记忆因子, 记忆因子最大值设定为 1.0, 最小值可设定为 0.6, 如果识辨方程计算出的记忆因子超出这个范围, 就赋给遗忘因子最小或者最大值。

3.5 本章小结

本章首先介绍了目前国内常用的校正模型并讨论了各种模型的优缺点, 在大流域预报校正可以分别使用反馈法、自回归或自回归滑动平均法和参数动态识别方法中的自适应实时校正法三种方法进行预报校正, 再进行比较分析确定使用哪种模型比较合适。然后对校正模型结构和参数确定的方法进行了讨论, 完成了实时校正模型的选择、模型结构的确定和参数初估的工作。

第四章 大流域洪水预报实时校正方案构建

大流域各个对象间关系复杂。在前文中已经论述了怎样进行大流域模型类的抽象和流域拓扑关系的确定。如何将各模型集成、怎样进行大流域的校正、在校正中需要注意哪些方面、大流域中各个产汇流分区的演算顺序如何,只有解决好这些问题才能构建大流域洪水预报实时校正方案。

4.1 预报模型结构

4.1.1 大流域子系统的划分

如果将大流域看作一个大的系统来考虑预报校正问题,一般需要将大系统分解为若干个子系统,计算是以子系统为单位进行的。每一个子系统又包括产汇流分区(面对象)、河段(线对象)、测站(点对象)、水量交汇点(点对象)。这些子系统的级别是不同的,级别不同的子系统演算顺序也是不同的。

子系统的划分可以参照河系分级的划分原则,在对河系的分级法中,最简单的是Gravelius分级法。Gravelius规定:在任一河系中,最大的主流为一级河流,汇入主流的大支流为二级河流,汇入大支流的小支流为三级河流,依此类推^[41]。那么一级河流所在的子系统为一级子系统,二级河流所在的子系统为二级子系统,三级河流所在的子系统为三级子系统。各子系统中又包含子河段和单元面积。

任一预报流域,在实际应用中总可以找出一条干流河道,然后按照Gravelius规定对复杂流域的河系进行分级。那么一个大系统也可以划分为若干个子系统,然后在实际演算中由低级的子系统到高级的子系统逐级演算。

如图4-1所示,S5—S1所在河段为一级子系统,包括产汇流分区对象P1、P3,河道对象(R1、R2、R4三者合成河道对象),测站对象S1、S2、S4、S5,水量交汇点对象N1、N2、N3等。二级子系统和三级子系统分别在其所处的二级河流

和三级河流，各包括一些产汇流分区对象、河道对象、测站对象和水量交汇对象。

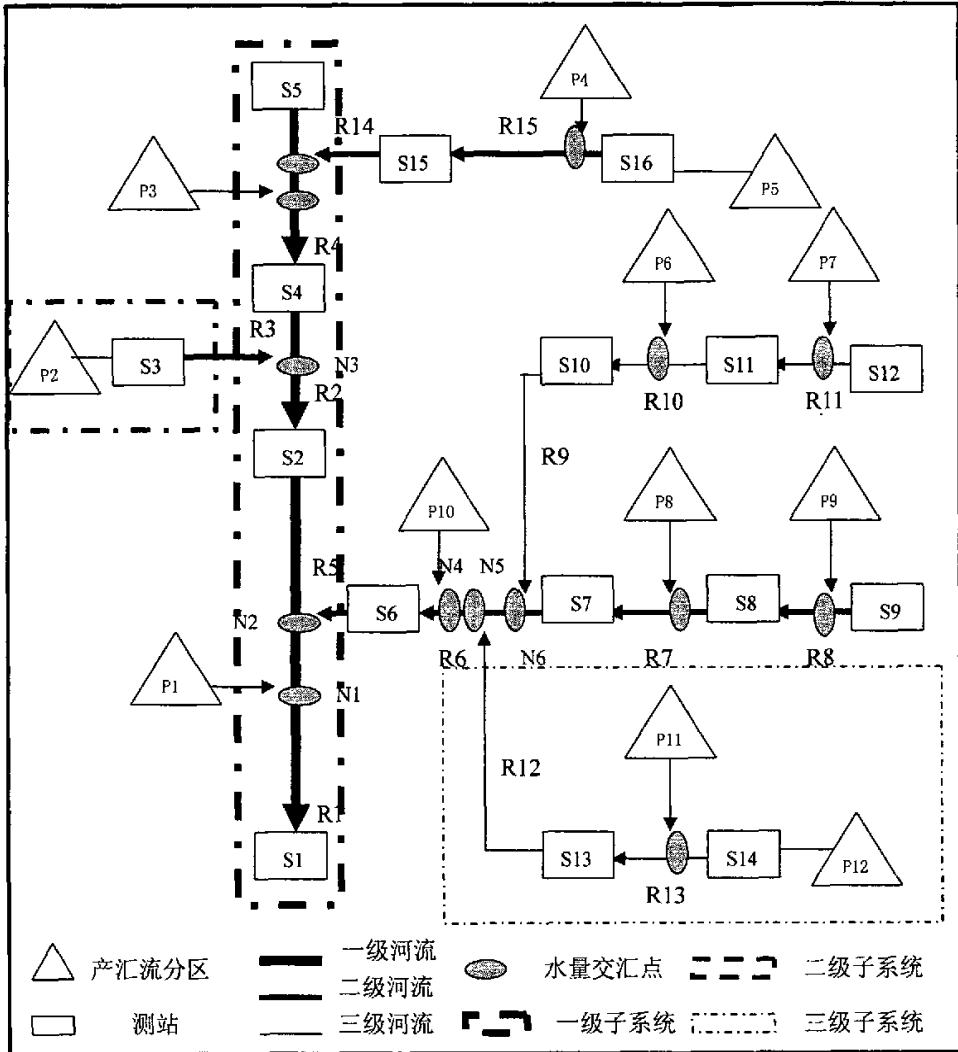


图4-1 复杂流域概化图

每一个子系统内的小流域可以单独进行产汇流计算，各个小流域间计算顺序是没有联系的，小流域还可以进一步细分为单元面积，具体划分方法参见第二章。为了充分利用河道内的水位流量信息做实时校正，河道对象常常需要划分为若干个子河段，河道内的流量站、水位站所在断面要做为划分子河段的依据。如图4-1中所示的一级子系统内河道对象就可以根据S1、S2、S4、S5四个测站分布划分为R1、R2、R4三个子河段，这些子河段所在的单元面积的和即为区间流域。在同一个级别子系统内，要先计算产汇流对象，再计算河道对象。

4.1.2 预报模型的选择

对于上游来水大，区间入流占来水量不多的情况下，预报精度主要取决于河道演算模型和实时校正方法。单元面积的降雨径流预报是次要的。当流域内有较多水文站时，应当注意把这些水位流量资料利用起来，因为这些资料量测方便比雨量资料误差小。

如果系统内水文站较少，难以控制水情，或者区间来水较多的情况下，降雨径流模型就显得比较重要了。一般应采用成熟简单的模型，如新安江模型国内用户较多，比较熟悉，在湿润半湿润地区效果较好，可以优先考虑使用。

在预报模型中还需要注意留有接口，可以方便连接指定的任何降雨径流模型

4.2 预报误差来源和实时校正原则

4.2.1 预报误差来源

目前在利用流域水文模型进行洪水实时预报时，采用的有效方法是用确定流域水文模型加上实时校正处理算法进行洪水实时校正预报。实时预报中不仅有实时雨量输入，更重要的是有实时流量（水位）信息的输入。如果能基于概念水文模型机理，又能将这些实时水文信息利用起来，很好的进行校正，就能够大大提高水文预报的精度。

预报误差主要来源于以下几个方面^[38]：

（1）模型结构误差。水文预报模型都是对实际流域系统进行概化后建立起来的，例如把空间分布过程表示为集总式的；水文循环的各个环节被人们概念的分离开，并做了不同程度的简化；单元面积之间的相互影响用简化了的函数系式描述。模型的这种不完善引起的误差称为模型结构误差。

（2）模型参数的误差。不论采用什么辨识方法（在线或是离线），也不用什么辨识技术，被辨识模型的参数对其真值来说总是存在误差的，即使是反流域自然地理特征的参数，如流域面积，流域长度等，由于测量误差存在，也能完全准确。

(3) 模型的输入误差。用概念性水文模型进行预报所要求的输入资料通常是降雨和流域蒸散发等。这些资料由实测得到,或是根据气象预报估算得到。前者存在测验预处理和平均误差,后者则存在相当大的预报误差。

(4) 预见期降雨引起的误差。如果要使预报的遇见期尽可能长,必须借助于气象部门的短期定量降雨预报结果,降雨预报结果直接关系到面雨量的计算,进而对流域水文模型预报结果产生影响。一般来讲,现有的技术进行雨量的定量预报精度是比较低的。

(5) 系统初始状态的误差。水文预报的初始时刻,系统模型中往往有一些变量需要赋初值,如新安江模型中包气带各层的土壤含水量。不过水文预报是连续作业的,又因系统记忆长度有限,使得这种误差来源不重要。

对于提高水文预报精度而言,最主要的是模型结构和参数引起的误差,若这两者误差较大,单靠优选参数和改善模型输入是不能够解决问题的,仅仅靠实时校正也是不能解决问题的。在大流域预报系统中,模型校正效果的好坏与流域内水文(水位)站的多少密切相关。如果流域内一个水文(水位)站也没有,则毫无实时校正的能力。流域内大大小小的各个水文站(或水位站),就好像是一系列传感器,如果它们发现在什么地方出现了误差,预报模型就能自动进行判断、估计、在整个流域上对预报模型参数和状态变量进行全面校正,使整个预报系统总的误差最小是最理想的校正情况,这也就是第二章产汇流分区中区间流域划分时要求上下断面尽量有测站的原因。这种校正情况可以称为“面校正”。但是在目前情况下,因测站分布密度,资料精度,校正方法所限,进行面校正有很多困难,例如目前大多数实时校正方法对在系统运行过程中对预报模型参数和状态变量校正乏术,只能尽可能的利用测站信息,用误差序列或其他方法对这些点进行“点校正”,一般也能取得满意的效果。

4.2.2 实时校正中需要注意的问题

在上游来水占大多数,区间来水所占总水量比例不大的大流域预报中,考虑到大流域实时预报的时效性,可以对大江大河的主要断面控制点进行校正,在

校正过程中应采用逐段校正的原则，将上一子河段出流的校正结果作为下一子河段的入流，一段一段校正到最终出口断面。但是在校正过程中还要注意几方面的问题：

(1) 小流域如何校正。小流域采用降雨径流模型，处于山区的小流域洪水暴涨暴落，现有的校正方法一般不能够进行很好的校正，可将其误差叠加在子河道所在测站中一起校正。如图4-2中A点和B点需要进行降雨径流计算，计算结果可以用马斯京根法演算到C点，再根据C点的实测径流校正。

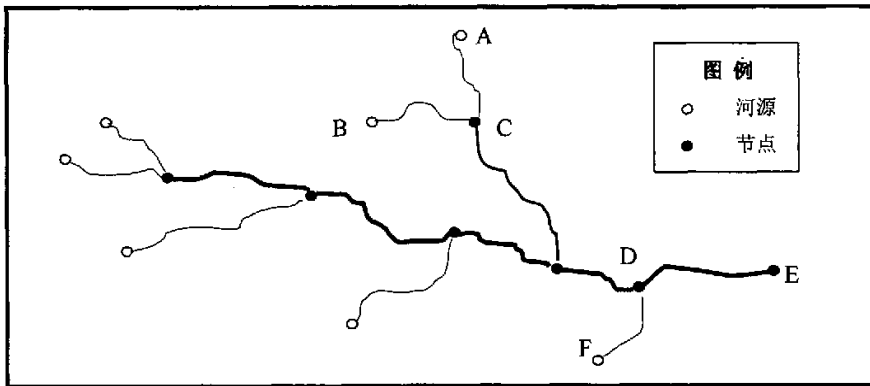


图4-2 某流域河道分布图

(2) 系统中支流如何校正。大流域中一般存在多条支流，对这些支流进行校正，即使校正结果比预报结果精度高，使用校正后的流量而不是预报流量作为区间入流进行计算也可能使下断面的预报结果精度下降。例如图4-2中，将C点校正后的结果作为区间入流，D点的预报结果可能比使用C点预报结果作为区间入流的精度低。所以在实时校正前一定要分析清楚，找到支流校正与干流校正的关系，那些校正后对下断面影响大的支流在校正时一定要十分小心，以免影响预报精度。

(3) 如何处理水量不平衡的问题。考虑一种理想情况，在一个封闭的河段，如果没有区间入流，那么上游断面总入流和下游断面的总出流量应当是相等的。考虑图4-2所示的DE河段，将D点的流量和F点的流量作为入流，那么在E点的出流与这两点的入流应当是相等的。马斯京根法是基于水量平衡方程的，那么使用D、F两点的入流总量推算出的E点出流也应当是符合水量平衡原理的。在实际情

况中却不是这样,由于模型结构误差和量测误差的存在,使用马斯京根法预报的下断面径流总量与上断面径流总量是符合水量平衡原理的,而实测的出流总量与入流总量却不能严格的遵守水量平衡原理。校正后的总出流量与总入流量也不可能严格的符合水量平衡。在校正后一定要认真分析,校正后的总出流量如果与总入流相差比较大的话,是否能采用校正后的流量作为下一子河段入流就值得商榷。这时不妨在该断面不进行校正,使用预报结果作为下一子河段入流,将所有误差放到下一子河段出流断面一起校正。

还有一种情况,就是上下断面校正误差刚好相抵消,如果对最终出口的预报结果影响不大,可以考虑逐段校正。

总之,大流域的误差来源是极为复杂的,校正也是比较复杂的,如果模型结构、参数选择正确,校正结果又达不到理想的情况,就可能是某一个断面的校正出了问题。所以在在大流域实时校正预报以前一定要对断面逐段检查,特别要防止出现水量严重不平衡的情况,以免造成全流域的预报错误。

4.2.3 流域预报的最大预见期

在大流域的预报实时校正中还需要注意最大预见期的选择。在预报中有预报的最大预见期,实时校正有实时校正的外推时长,大流域的预报实时校正应当取两者中较短的预见期。

在单纯的水文预报中,降雨径流预报的预见期一般指主要降雨基本停止到预报断面洪峰出现这个滞时;对于以上游来水为主的区间,预见期通常是上下断面之间的传播时间,即河道汇流时间。由于降雨中心经常变化,河道汇流也有非线性作用,故预见期也是变化的。对于一个具体流域具体断面而言,各次洪水的预见期都是客观存在的,决不能人为增加预见期,只能做到根据最新信息采用合适的模型不损失预见期。不同位置的防洪目标,有不同的预见期;不同的雨水情,又有不同的预见期,两场洪水由于洪量大小的区别,汇流时间绝对不会是一模一样的。一般来说,靠近上游的断面预见期短,靠近下游的断面预见期长;靠近降雨中心的断面预见期短,远离降雨中心的断面预见期长。

一个大流域系统的的预见期与遥测雨量站、水位站的布置以及预报精度密切相关。如果要求预见期长,意味着预报范围的扩大,河道必须向上游延伸,降雨径流模型还需要定量的面雨量预报,精度必然降低。预见期与预报精度是一对天然的矛盾,现有的方法很难解决。

对于一个具体的流域,根据历史水文资料,可以大致确定预见期有多长。但是在实际应用中预见期与河道汇流时间还不是完全相等的,预见期往往要比河道汇流时间更短一点。在区间流域中,如果上游断面知道洪峰的流量,峰现时间后6个小时以内的流量对下游断面的峰值预报还将产生一定影响,超过6个小时的洪峰流量对下游断面的洪峰计算基本上没有影响。例如某河道汇流时间为10个小时,某日8时上断面出现洪峰流量,按照理论8时就可以准确计算出下断面18时的洪峰流量,但是8时后面的8-14时的流量数据可能都会对18时的峰值产生影响,在滚动预报中,只有到14时才能基本准确的预报下游断面18时的洪峰流量,实际预见期只有4小时。

目前河道汇流演算普遍采用的马斯京根法,在处理这一问题时还有一些因素需要考虑。仍以上例为例,上断面在8时计算下断面18时的出流过程,上断面8-18时的流量是未知的,如何给出这一出流过程也是需要考虑的。这一出流过程可以人为给定为0(称为第一种出流处理方法),但是这样随着时间的推移,马斯京根法的计算结果表现为衰退。经过研究发现,将这一个过程给定为8时的流量过程精度将大大提高(称为第二种出流处理方法)。

再如某河道某场洪水汇流时间为30个小时,河道预见期可以认为是24小时,在下游断面洪峰出现前24小时使用马斯京根法进行预报。如果将这24个小时上断面的出流过程视为0,预报结果如图4-3所示:实测洪峰流量 $31100\text{m}^3/\text{s}$,预报的洪峰流量 $26300\text{m}^3/\text{s}$,洪峰值相对误差15.4%,峰现时间相差6个小时。如果将这24小时上断面出流过程全部视为上断面预报时刻的流量,预报结果如图4-4所示:实测洪峰流量 $31100\text{m}^3/\text{s}$,预报的洪峰流量 $28400\text{m}^3/\text{s}$,洪峰值相对误差8.7%,峰现时间几乎相等。哪种方法更优一目了然。

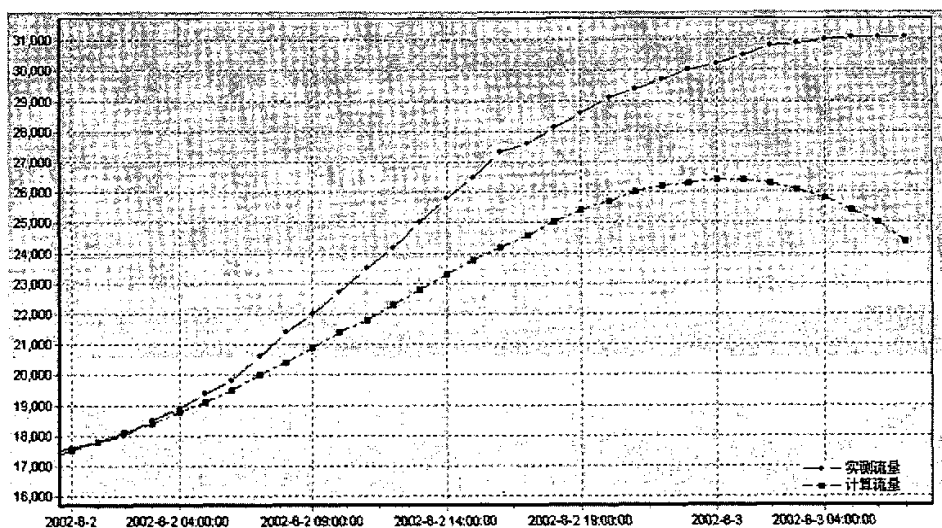


图4-3 第一种出流处理方法预报结果

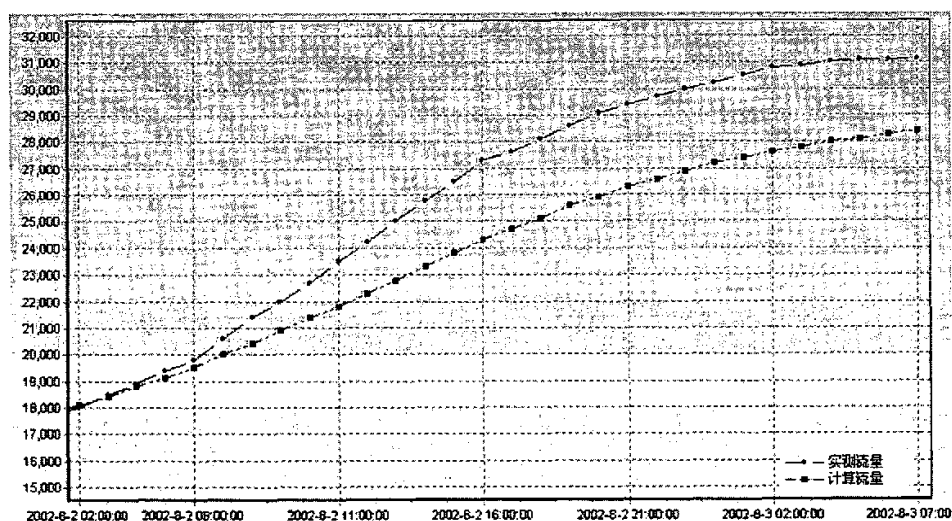


图4-4 第二种出流处理方法预报结果

实时校正也有一个合理的外推时长，过短将损失预见期，过长不仅效果越来越差，也没有理论支持。以误差自回归方法为例，误差序列绝不是想校正多长时间就校正多长时间的。这就需要对误差序列相关性进行分析，找出合适的校正外推时段长。分析多个流域流量误差序列相关性发现，24小时误差序列相关性在0.8327~0.9067之间(相关系数平均值0.8697)，48小时误差序列相关性在0.5623~0.7501之间(相关系数平均值0.6562)，72小时误差序列相关性在0.5039~0.7207

之间（相关系数平均值0.6123）。进一步对24小时内误差序列进行相关分析假设检验可以发现，这些流域流量误差序列线性相关程度都是显著的，少数流域线性相关是特别显著的，可以采用误差回归得方法进行校正。由于48小时、72小时预报误差相关系数不高，可认为基本是不相关的，因此误差自回归方法校正的外推时长应在24小时以内比较合适。

4.3 洪水预报实时校正方案构建

预报方案构建前还需要确定方案对象链和方案配置链。预报方案对象链是指预报断面以上在一定空间范围内的水文模型对象的集合，其主要包括产汇流分区对象和测站对象。预报方案配置链是记录预报方案的边界测站对象和预报方案中要进行实时校正的测站对象等信息^[11]。

图4-1是基于水文模型对象的复杂流域概化图，可利用第二章介绍地方法对预报对象编码，并建立拓扑关系：用图邻接链表表示概化图中线对象拓扑关系；用树链接表表示概化图中的点对象拓扑关系。最后可生成对象链：利用点对象拓扑关系表，搜寻水文模型对象。

例如图4-1中以S4、S10、S7、S13为边界的流域方案对象链和方案配置链为：方案对象链=P₁, P₂, P₁₀, S₁, S₂, S₄, S₃, S₆, S₇, S₁₀, S₁₃；方案配置链=S₄, S₇, S₁₀, S₁₃；S₂, S₆。这里数字使用下标来表示，下文将添加上标表示演算次序。

预报方案对象链包含了预报断面以上所有水文模型对象的集合，预报方案对象配置链记录了方案的边界信息等。这些信息都是散乱的集合，没有先后顺序，无法确定演算次序。在预报方案模型的结构中，子河段的划分是以河道站为界进行分割的，从而可将子系统分级的信息传递给河道站，作为河道站的计算分级属性，扩充到测站类中，如子系统级别为一级，那么子系统包含的河道站的计算分级也为一级，依此类推。

通过上述过程后，方案对象链就有了比较丰富的信息，如上述方案对象链和配置链可以写为：

方案对象链= $P_1, P_2, P_{10}, S_1^1, S_2^1, S_4^1, S_3^2, S_6^2, S_7^2, S_{10}^3, S_{13}^3$

方案配置链= $S_4^1, S_7^2, S_{10}^3, S_{13}^3; S_2, S_6$

测站对象S的上标表示计算分级，下标是编码。

预报方案对象链与配置链的生成及对象演算次序的确定，为进行预报方案的计算奠定了基础，那么在实际预报方案的计算过程中还需建立相关的属性表。

河段对象采用的汇流模型为马斯京根法、产汇流分区对象采用的产汇流模型为新安江模型。需要建立的属性表主要为三种：河段对象属性表，包括河段上下断面代码、河段分段信息、参数信息等；产汇流分区对象属性表，包括分区代码、出口控制站、分区面积、引用蒸发站、参数信息等；交汇点属性表主要包括所属河段，距离上断面信息等。交汇点属性表的建立主要是为了河道分段演算的方便，在已知河段长度和本交汇点与上断面距离的情况下，就可以知道交汇点在第几个河段入流。如河道在演算时分为4段，某产汇流分区距上断面30km，河段总长度40km，那么该分区可认为在第3段入流，这样在计算前两段时可以采用无区间入流的马斯京根法，在计算第3段时需要加入区间入流再进行河道演算。

利用方案对象链，再根据上述属性表，就可以进行预报方案的计算和实时校正：

(1) 首先计算方案对象链中的产汇流对象，即P类型点。从方案对象链“ $P_1, P_2, P_{10}, S_1^1, S_2^1, S_4^1, S_3^2, S_6^2, S_7^2, S_{10}^3, S_{13}^3$ ”中取出P类型点，在此链中为“ P_1, P_2, P_{10} ”。通过循环，分别从产汇流分区对象属性表中获取信息，计算出口控制断面的出流过程，并将值赋给点对象，包括测站对象或水量交汇点对象，在例中就赋给N1，S3，N4。由于S3有实测流量，可以对预报结果运用实时校正模型进行校正，将校正后的值赋给S3对象作为入流过程继续进行河道演算，在河段对象属性表中搜寻以“S3”为上断面的河段对象R3，并记录此河段对象的下断面N3，获取马法演算的参数，进行河道演算，计算出河段对象R3的下断面出流过程，并将值赋给下断面水量交汇点对象N3。所有的P类对象最终计算到水量交汇点，所有要赋给S对象的值必须经过校正后才能作为S对象的入流过程。

(2) 对S对象按所在河段等级逐级计算。对方案配置链中的边界信息“ $S_4^1, S_7^2,$

S_{10}^3, S_{13}^3 ”中的各个测站对象进行计算等级判断, 得出有多少级计算等级, 在此链中计算等级1级、2级、3级; 按计算级数进行循环, 先计算3级河段, 再计算2级河段, 最后再计算1级河段, 从配置链的边界信息中取出测站计算等级为3级的测站对象“ S_{10}^3, S_{13}^3 ”, 通过循环, 分别计算出至下一级河段入流点的入流过程, 与P对象相似, 最终计算到水量交汇点。具体过程如下:

- ① S_{10}^3 是三级河段的边界, 获取边界入流过程;
- ②搜寻“河段对象属性表”中以 S_{10}^3 上断面的河段对象为R9, 并记录此河段对象的下断面N6, 获取马法演算的参数, 进行河道演算, 计算出河段对象R9的下断面出流过程, 并将值赋给下断面水量交汇点对象N6;
- ③ S_{13}^3 也通过同样的方式, 最终计算到水量交汇点对象N5;
- ④通过上述过程已从3级河段计算到2级河段, 3级河段的出流过程, 作为2级河段的集中入流过程。从配置链的边界信息中取出测站计算等级为2级的测站对象“ S_7^2 ”, 同样获取边界入流过程, 在“河段对象属性表”中搜寻以 S_7^2 为上断面的河段对象为R6, 并记录此河段对象的下断面S6, 获取马法演算的参数, 进行河道演算, 在演算过程中将N4、N5、N6中记录的出流过程, 作为区间入流加入河段的河道演算中, 计算出河段对象R6的下断面出流过程, 并将值校正后赋给下断面水量交汇点对象S6, 然后重复上述过程, 最终计算到水量交汇点对象N2;
- ⑤完成3级河段、2级河段的计算, 同样可按上述过程计算1级河段, 最终计算得到出口断面S的流量过程, 在计算过程中所有要赋给S对象的值都要进行校正后才能赋给S对象。在整个计算校正过程中需要注意本章第二节中提到的大流域实时校正需要注意的问题, 以免造成错误。

在这里测站对象和水量交汇点对象进行实测的或计算的水文要素的存储, 河段对象仅进行模型计算和建立测站对象间的链接。产汇流对象同样仅进行模型计算。

4.4 本章小结

本章首先介绍了大流域预报模型结构，将大系统划分为若干个子系统，各子系统又根据所处河流的级别划分为不同的级别，子系统河道对象可以根据水文（水位）站的位置划分为不同的子河段，以最大限度的利用流域中的信息进行预报和校正。然后分析了大流域预报误差的来源以及在校正中需要注意的问题，论述了如何取得预报的最大预见期。最后根据流域拓扑结构建立方案对象链和方案配置链，确定了预报和实时校正的次序，构建了预报校正方案。

第五章 应用实例

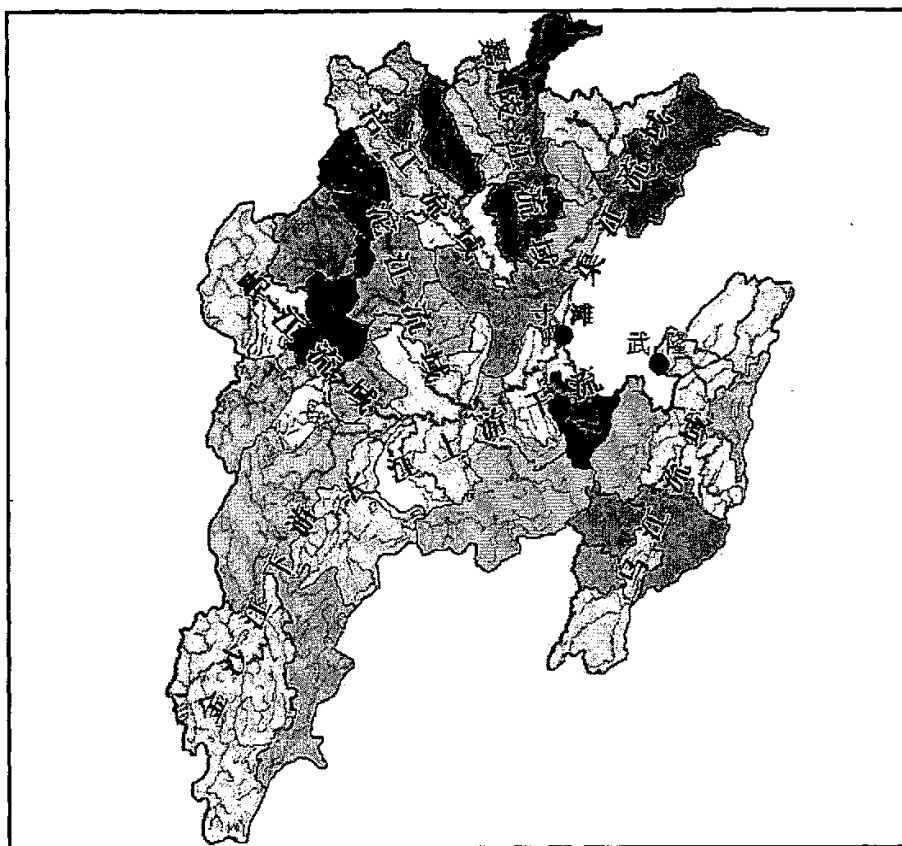
长江是我国第一大江，历史上曾多次发生大洪水，给两岸人民带来了无穷的灾难。为了提高长江的防洪能力，同时兼顾水力发电、改善航运、供水灌溉的需要，我国于 1994 年底开始修建长江三峡工程。

修建三峡工程的主要目的是防洪和发电，要实现防洪与兴利相结合，科学的水库调度是必不可少的。这就必须充分掌握其上游的水情信息，并对洪水未来的变化趋势做出准确的预报。长江干流寸滩以上和支流乌江流域武隆以上的洪水是三峡水库入库洪水的主体部分，做好这两个控制站的洪水预报至关重要。寸滩以上集水面积 866559km^2 ，武隆以上集水面积 83035km^2 ，占入库径流总量的 90% 以上。

5.1 流域洪水预报实时校正方案构建

该流域区域范围是从长江干流金沙江龙街以下至寸滩（以下简称为龙寸流域，包括嘉陵江流域、岷江流域、沱江流域、金沙江下游、长江上游干流区）和乌江流域（如图 5-1）。长江干流自江源至湖北宜昌为长江上游，包括江源段、通天河段、金沙江段和川江段。长江干流上游河段长 4504km ，流域面积约 100万 km^2 。乌江流域没有大的支流汇入，流域拓扑关系简单，限于时间和篇幅所限，主要对龙寸流域的预报校正方法进行了研究。

要进行预报方案的构建，首先应当使用前面介绍的方法对龙寸流域进行子系统的划分，在子系统中进行单元面积的划分，划分好的单元面积和产汇流分区可以参见前文第二章图 2-4 和图 2-5。然后对模型对象进行分析，确定面对象、线对象和点对象的属性表格。根据这些对象确定流域拓扑关系和计算次序，进行预报校正。这些方法在上文已经有过详细介绍，这里不再赘述。本节将重点放在分析流域特性上，以实例来说明如何找出适合本流域的预报校正方法。



5-1 龙寸流域和乌江流域示意图

对龙寸流域进一步研究发现,龙寸流域以上游来水和岷江、嘉陵江入流为主,需要做降雨径流预报的小流域产流量不大,一般只有每秒几十至几百立方米,这对于长江干流每秒几万立方米的流量来说是可以忽略不计的。整个流域应当以河道演算为主,预报精度直接取决于河道演算的精度。对这些小流域和比较小的支流进行实时校正后发现,这些小流域小支流对整个长江的洪水影响甚微,可以不考虑校正。

长江干流李庄—泸州段,李庄入流占主要部分,支流沱江富顺站入流量很少,几乎可以忽略不计,李庄—泸州030830次洪过程和来水组成分布图如图5-2和图5-3所示。在李庄—泸州段,对沱江流域预报结果没有进行校正,主要用上断面李庄和区间沱江流域入流采用有区间入流的河道演算模型计算泸州的流量,然后对在泸州断面对预报结果进行了校正。

龙寸流域比较大的支流从东到西有:岷江、沱江、涪江、嘉陵江、渠江,其

中岷江和嘉陵江入流比较大。岷江入流占19%左右；沱江入流较小，仅为6%左

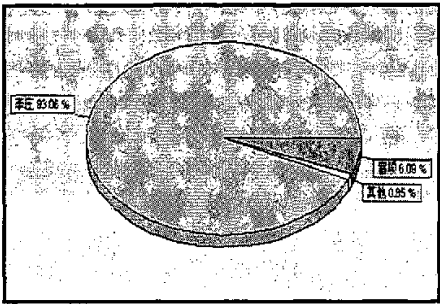
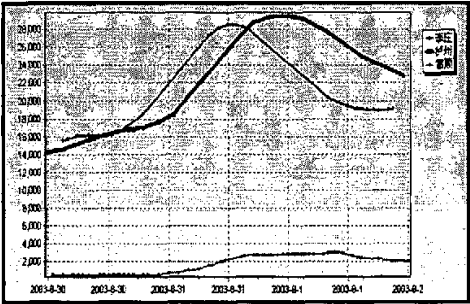


图 5-2 李庄—泸州 030830 次洪水过程图 图 5-3 泸州站 030830 次洪水组成分布图

右；涪江渠江都汇入嘉陵江，嘉陵江入流占 16%左右。考虑到入流量太小的流域即使校对整个大流域预报结果影响也不大，该流域在支流校正方面可采用四种方法校正：两条支流均不实时校正，直接使用预报结果进行计算；只对岷江干流实时校正；只对嘉陵江干流实时校正；两条支流均进行校正，使用校正后的结果作为区间入流计算。通过这四种方法对比发现：两条支流均进行校正的情况下预报结果最好。如长江干流 020801 次洪水预报过程，如果两条支流不进行校正，流量峰值为 $27900\text{m}^3/\text{s}$ ；对一条支流校正后流量峰值在 $28100\text{m}^3/\text{s}$ 左右；两条支流都校正后峰值可以提高到 $28400\text{m}^3/\text{s}$ ，精度有一定提高。

对这些支流的主要支流也要进行类似的分析，经过分析决定：对岷江流域的高场，五通桥两站的预报结果进行校正；对嘉陵江流域的罗渡溪、风滩的预报结果进行校正。长江干流主要的控制站华弹、屏山、李庄、泸州、朱沱、寸滩预报结果进行逐段校正，本子河段出流作为下一子河段入流。整个大流域预报校正过程中水量基本平衡。

对本流域研究发现，使用自回归模型时采用流量和雨量序列作为回归项或滑动平均项效果均不理想，目前比较普遍的处理方法是利用水文模型的预报流量序列与实时观测流量序列差值的残差序列进行自回归分析，建立残差预报模型，将预报的残差叠加到预报流量序列上，完成预报流域洪水实时校正，从而提高洪水预报精度。

两种回归模型，误差自回归模型和自回归滑动平均模型使用起来是有差别的，一般认为自回归滑动平均模型比单纯的自回归模型精确。在需要进行河道演

算的河段上分别建立三阶误差自回归模型和自回归滑动平均模型，自回归滑动平均模型用本断面的流量误差序列作为自回归项，上断面的流量误差序列作为滑动平均项。两种方法对比后发现，采用误差自回归模型与自回归滑动平均模型做一步预报时，预报结果相差很小，由表 5-1 可以看出：AR(3)模型平均确定性系数为 0.9027，ARMA(3,3)为 0.9044。本文研究方法将应用于长江上游干流，河道水量变化相对平缓，区间入流不大。考虑到方法的简便性和预报的时效性，采用误差自回归模型就已经足够。时间序列模型使用误差自回归方法。

表 5-1 AR(3)与 ARMA(3,3)模型预报结果确定性系数对比表

ID	流域名	AR(3)	ARMA(3,3)	ID	流域名	AR(3)	ARMA(3,3)
1	三汇	0.9777	0.98	12	福录镇	0.8545	0.8461
2	罗渡溪	0.9295	0.9145	13	金溪	0.9401	0.9527
3	武胜	0.8058	0.7941	14	李庄	0.9989	0.9989
4	涪江桥	0.7105	0.7281	15	高场	0.9624	0.9697
5	射洪	0.6961	0.7024	16	泸州	0.9334	0.9452
6	小河坝	0.8694	0.8594	17	风滩	0.8116	0.8112
7	北碚	0.831	0.805	18	五通桥	0.9777	0.9853
8	富顺	0.9473	0.9274	19	昭化	0.9505	0.9652
9	阆中	0.9321	0.9423	20	屏山	0.9751	0.9734
10	朱沱	0.9655	0.9798	21	华蓥	0.9736	0.9867
11	寸滩	0.9136	0.9245				

在校正模型结构选择方面，本文利用洪水资料分别建立了 3 阶、6 阶、12 阶、36 阶的自回归模型，然后根据这些模型对多场洪水做一步预报，计算各回归模型预报的平均确定性系数。使用 3 阶误差自回归模型进行计算，总体情况可参照表 5-1，平均确定性系数 0.9027。6 阶的预报模型精度有一定提高，平均确定性系数达到 0.9329。而 6 阶、12 阶和 36 阶精度相差不多，6 阶模型确定性系数比 12 阶还略高，拟采用 6 阶模型进行预报。

5.2 结果分析和精度评定

长江流域中大江大河的洪水一般历时比较长，有的洪水过程长达一周甚至十天。根据本文第四章的研究成果，如果使用误差项作为自回归序列，外推时长在 24 小时以内比较合适。该流域河道汇流时间为 70 小时左右，因此洪水实时预报校正预见期应在 24 小时。因为洪水预报主要预报峰值，所以本文主要使用洪水

起涨阶段的资料预报洪峰值，在预报中为消除模型初始状态对预报的影响，在峰现时间前 3—4 天的进行预报预热计算，自适应实时校正模型使用这三四天的资料同时进行在线参数估计，在峰现时间前 24 小时进行洪水预报实时校正计算。

5.2.1 反馈法计算结果

使用反馈法对龙寸流域的十场洪水进行 24 小时预见期实时预报校正发现，大部分的洪水预报精度都有一定的提高，但是提高有限，而且随着时间的推移，校正作用逐渐消失。反馈法对龙寸流域 020613 次洪水和 020801 次洪水实时预报校正结果分别见图 5-4 和图 5-5，其他洪水评定结果见表 5-2。

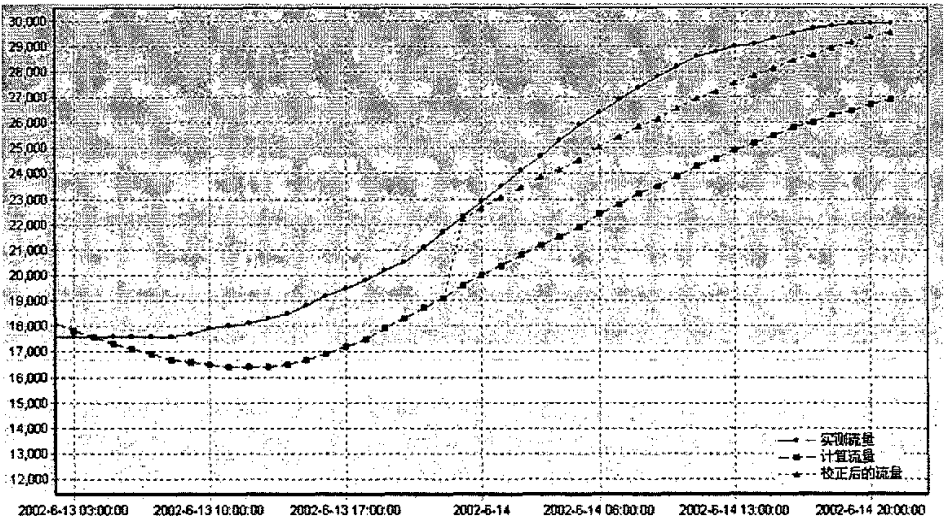


图 5-4 寸滩站反馈法 020613 次洪水预报结果

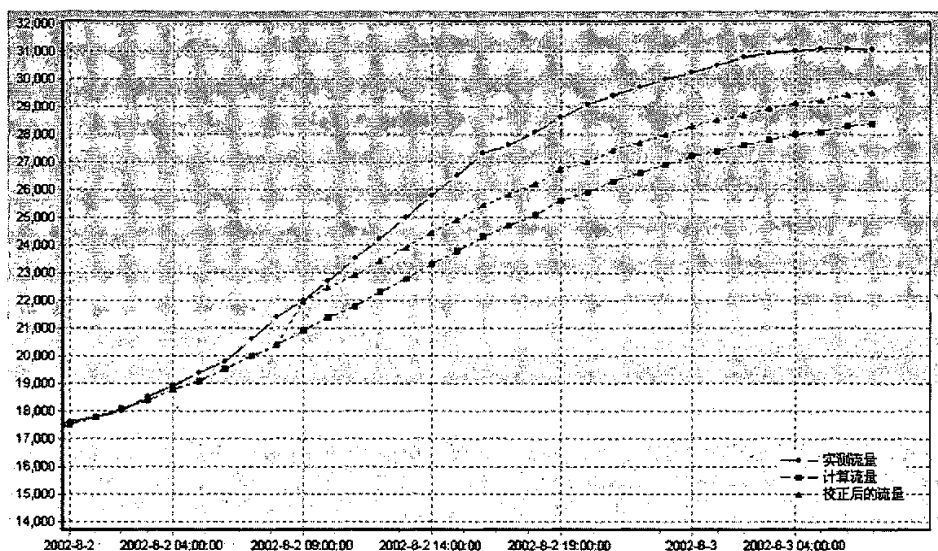


图 5-5 寸滩站反馈法 020801 次洪水预报结果

表 5-2 龙寸流域反馈法洪水预报结果评定分析表

洪号	确定性系数		洪峰流量(m^3/s)			洪峰相对误差(%)		峰现时间	
	校正前	校正后	实测值	校正前	校正后	校正前	校正后	校正前	校正后
810824	0.84	0.88	47500	45700	42800	3.79%	9.89%	-4h	-4h
020507	0.81	0.84	10000	10900	10400	9.00%	4.00%	-1h	-1h
020613	0.79	0.89	29900	26900	29600	10.03%	1.00%	2h	2h
020702	0.80	0.82	31600	28100	31300	11.08%	0.95%	-1h	-1h
020716	0.79	0.84	35800	36700	36400	2.51%	1.68%	0h	2h
020801	0.85	0.87	31100	28400	29500	8.68%	5.14%	2h	2h
030728	0.80	0.81	21500	19200	19700	10.70%	8.37%	1h	3h
030810	0.83	0.83	22400	23300	22100	4.02%	1.34%	4h	3h
030920	0.87	0.87	32200	32900	32200	2.17%	0.00%	0h	2h
031002	0.77	0.78	22800	23100	23000	1.32%	0.88%	3h	2h

由表 5-2 可以看出,寸滩流域用反馈法进行校正后,精度有了一定提高,确定性系数平均值由 0.82 提高到 0.84;洪峰相对误差由 6.33% 下降到 3.33%。反馈法虽然比较简单,但是校正能力有限,有时候还会降低预报精度。

5.2.2 误差自回归法计算结果

使用误差自回归方法对龙寸流域 020613 次洪水和 020801 次洪水实时预报校正结果分别见图 5-6 和图 5-7，其他洪水评定结果见表 5-3。

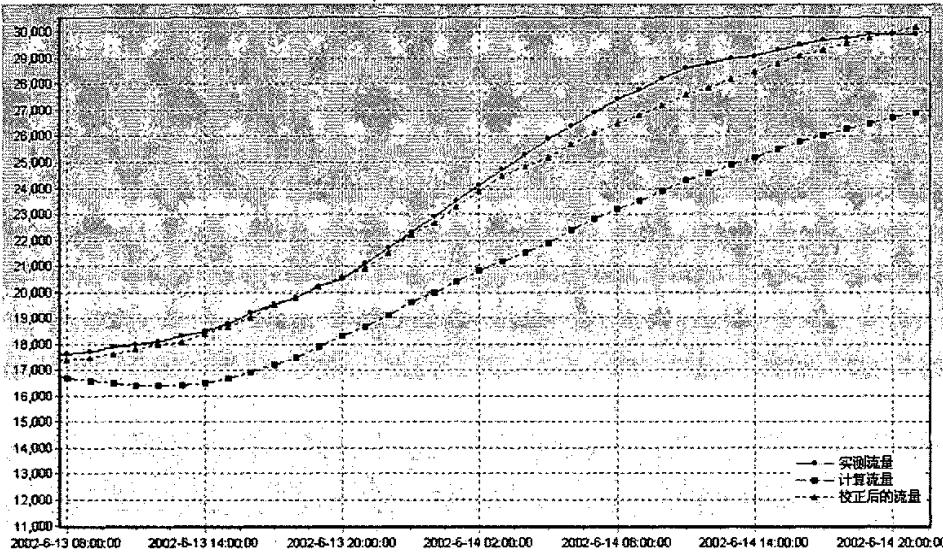


图 5-6 寸滩站误差自回归法 020613 次洪水预报结果

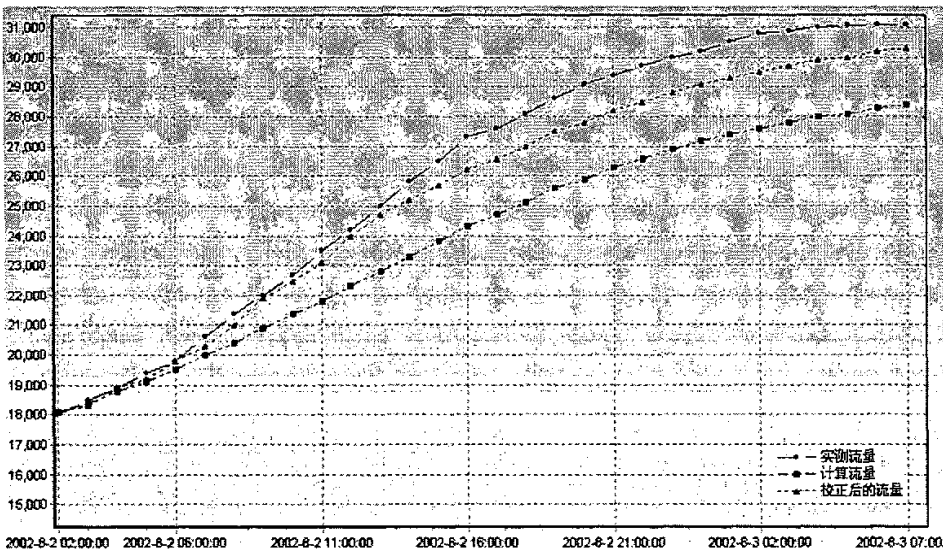


图 5-7 寸滩站误差自回归法 020801 次洪水预报结果

表 5-3 龙寸流域误差自回归法洪水预报结果评定分析表

洪号	确定性系数		洪峰流量 (m ³ /s)			洪峰相对误差 (%)		峰现时间	
	校正前	校正后	实测值	校正前	校正后	校正前	校正后	校正前	校正后
810824	0.84	0.90	47500	45700	46100	3.79%	2.95%	-4h	-4h
020507	0.81	0.86	10000	10900	10700	9.00%	7.00%	-1h	-1h
020613	0.79	0.94	29900	26900	30100	10.03%	0.67%	2h	2h
020702	0.80	0.85	31600	28100	30000	11.08%	5.06%	-1h	-1h
020716	0.79	0.83	35800	36700	36500	2.51%	1.96%	0h	2h
020801	0.85	0.91	31100	28400	30200	8.68%	2.89%	2h	2h
030728	0.80	0.89	21500	19200	19900	10.70%	7.44%	1h	3h
030810	0.83	0.91	22400	23300	23100	4.02%	3.13%	4h	3h
030920	0.87	0.90	32200	32900	32200	2.17%	0.00%	0h	2h
031002	0.77	0.85	22800	23100	23000	1.32%	0.88%	3h	2h

由表 5-3 可以看出，寸滩流域用误差自回归法进行校正后，校正能力有限，确定性系数平均值由 0.82 提高到 0.88，洪峰相对误差由 6.33% 下降到 3.20%。

5.2.3 自适应实时校正法计算结果

由于卡尔曼滤波法在水文领域应用的局限性，参数动态识别法主要采用自适应实时校正方法。龙寸流域 020613 次洪水和 020801 次洪水预报结果分别见图 5-8 和图 5-9，其他洪水评定结果见表 5-4。

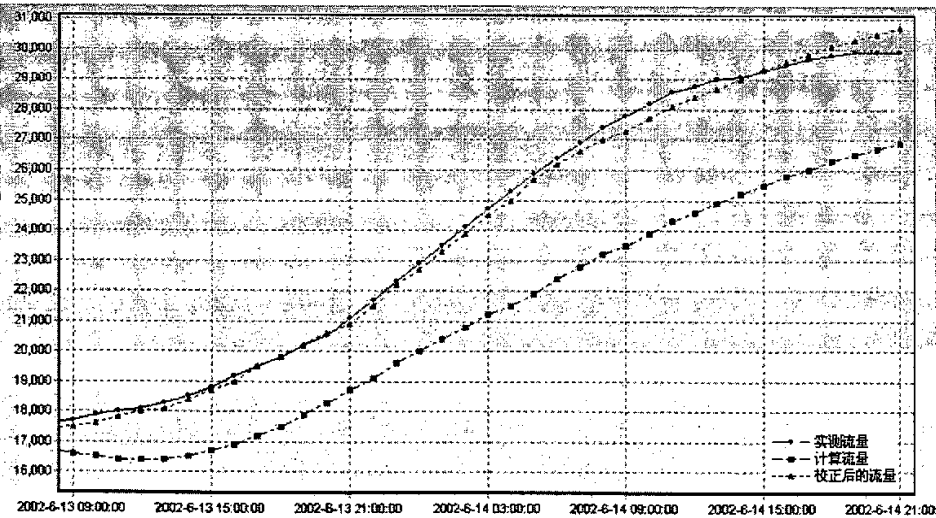


图 5-8 寸滩站自适应实时校正算法 020613 次洪水预报结果

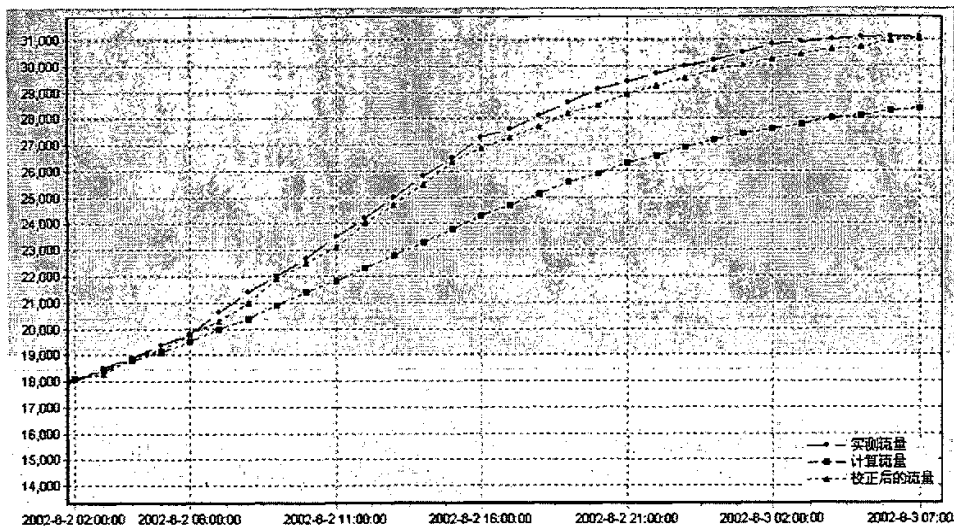


图 5-9 寸滩站自适应实时校正算法 020801 次洪水预报结果

5-4 表龙寸流域自适应实时校正算法洪水预报结果评定分析表

洪号	确定性系数		洪峰流量 (m³/s)			洪峰相对误差 (%)		峰现时间	
	校正前	校正后	实测值	校正前	校正后	校正前	校正后	校正前	校正后
810824	0.84	0.91	47500	45700	46300	3.79%	2.53%	-4h	-4h
020507	0.81	0.86	10000	10900	10700	9.00%	7.00%	-1h	-1h
020613	0.79	0.95	29900	26900	30800	10.03%	3.01%	2h	2h
020702	0.80	0.86	31600	28100	30300	11.08%	4.11%	-1h	-1h
020716	0.79	0.92	35800	36700	36400	2.51%	1.68%	0h	2h
020801	0.85	0.96	31100	28400	31000	8.68%	0.32%	2h	2h
030728	0.80	0.93	21500	19200	20200	10.70%	6.05%	1h	3h
030810	0.83	0.94	22400	23300	23100	4.02%	3.13%	4h	3h
030920	0.87	0.91	32200	32900	32700	2.17%	1.55%	0h	2h
031002	0.77	0.87	22800	23100	23000	1.32%	0.88%	3h	2h

由表 5-11 可以看出，寸滩流域用自适应实时校正算法进行校正后，精度提高很多，确定性系数平均值由 0.82 提高到 0.91，洪峰相对误差由 6.47% 下降到 3.03%。

5.2.4 实时校正效果分析

以上对三种方法校正的洪水进行了误差评定，下面对三种方法做一个比较，如表 5-5 所示。表中的方法一指反馈法，方法二指误差自回归法，方法三指自适

应实时校正法。

表 5-5 三种方法比较

流域名	洪号	确定性系数			洪峰相对误差 (%)		
		方法一	方法二	方法三	方法一	方法二	方法三
龙 寸 流 域	810824	0.88	0.90	0.91	9.89%	2.95%	2.53%
	020507	0.84	0.86	0.86	4.00%	7.00%	7.00%
	020613	0.89	0.94	0.95	1.00%	0.67%	3.01%
	020702	0.82	0.85	0.86	0.95%	5.06%	4.11%
	020716	0.84	0.83	0.92	1.68%	1.96%	1.68%
	020801	0.87	0.91	0.96	5.14%	2.89%	0.32%
	030728	0.81	0.89	0.93	8.37%	7.44%	6.05%
	030810	0.83	0.91	0.94	1.34%	3.13%	3.13%
	030920	0.87	0.90	0.91	0.00%	0.00%	1.55%
	031002	0.78	0.85	0.87	0.88%	0.88%	0.88%
	平均	0.84	0.88	0.91	3.33%	3.20%	3.03%

由表 5-5 可以看出，三种方法中，反馈法精度不高而且不够稳定，误差自回归法校正稳定但是精度不高，自适应实时校正法精度是最高也是最稳定的。自适应实时校正法可以很好的进行预报，预报结果都在甲等以上，可以进行大流域的预报校正。

5.3 本章小结

本章用前文介绍的方法在寸滩流域进行了预报方案的构建，并使用三种实时校正方法分别进行校正运算，进行比较后得出结论：在现阶段资料有限的情况下，如果能够很好的分析流域特性，采用合适的预报模型和校正方法，利用简单的实时校正法即可很好的校正模拟洪水，实现洪水实时预报。

第六章 结论与展望

6.1 结论

大流域由于其地域广、支流多的特点,洪水预报实时校正一直是比较难的问题。随着科技的发展,虽然现在出现了一些新方法,但是由于资料的原因,特别是下垫面资料精度不足等问题,很多方法都不能够应用到实际工作中去。

本文在前人研究的基础上,首先对大流域中的各种研究对象进行了水文解析,将整个流域抽象为一个个点、线、面,并将其以表格的形式储存在数据库中,方便计算。

大流域内模型对象众多,各模型对象采用的预报模型是不同的,各对象计算先后也有严格的顺序。本文将大流域分成各级子系统,子系统的计算顺序可以根据河道的分级由低到高逐级计算。本文还运用拓扑学的知识将子系统中各模型对象很好的集成起来,解决了模型之间的数据传递和计算顺序的问题,构建了洪水预报实时校正方案。

大流域中还包含很多支流,实时校正是一个比较复杂的问题。本文分析了误差来源,提出了大流域中实时校正的原则和需要注意的问题。本文还对流域预报预见期的问题进行了讨论,确定使用流域水文模型预报的预见期和实时校正模型外推期的交集作为流域的预见期。在大流域中要仔细分析流域特性,选择合适的预报校正模型,如果支流的校正结果对整个流域影响比较大,对这种支流的实时校正结果的运用是要极为小心的。

通过本文研究表明,如果仔细分析大流域的流域特点,选择合适的模型,用简便的方法也能取得比较好的效果。

6.2 展望

(1) 概念性水文模型的研究需要加强

现阶段的水文模型有些过于复杂,资料要求较高;有些对资料要求低但是在模型结构等方面还存在一定的误差,如何完善这些模型结构,探求一种结构简单,预报精度高的模型是努力的方向。要加强概念性水文模型的研究,除了在水文学方面积极研究流域特性外,还要结合其他领域的发展,综合考虑,完善预报模型。

(2) 实时校正方法需要进一步改进

目前流行的实时校正或者滤波算法都存在着以下几个问题:滞后一步校正,即利用前一阶段新息校正前一阶段发布预测的模型变量或参数,并且仅适用于缓变条件。还应加强模型参数或状态变量在突变环境下预报校正方法的研究。

另外目前的校正模型多数是数学领域的模型,应当寻求一种适合水文领域的实时校正模型,使其在能够实现模型参数和状态变量的全面校正,尽可能的延长大流域预报的预见期,提高大流域预报精度。

参考文献

- [1]大连理工大学等. 水库防洪调度方法及应用[M]. 北京: 中国水利水电出版社, 1996.
- [2]李少华. 冯家山水库洪水预报调度系统的研究与开发[D]. 西安: 西安理工大学, 2002.
- [3]J-E·Nash, B-I·Barsi. 大流域径流预报的混合模型[J]. 长江工程职业技术学院学报, 1992(4):71-75.
- [4]梁庚辰, J.E.Nash. 大流域汇流演算的线性扰动模型[J]. 水力发电学报, 1991(2).
- [5]何少华. 水文系统模型在大流域汇流演进计算中的应用[J]. 东北水利水电, 1995(8):21-30.
- [6]李致家, 董增川, 梁忠民, 等. 大流域洪水预报与洪水调度管理方法研究[J]. 水利发电, 2004(1):12-15.
- [7]刘昌明, 李道峰, 田英, 等. 基于DEM 的分布式水文模型在大尺度流域应用研究[J]. 地理科学进展, 2003 (9) 437—447.
- [8]陈利群, 刘昌明, 袁飞. 大尺度资料稀缺地区水文模拟可行性研究[J]. 水电资源科学, 2006 (1) 87—92.
- [9]张雪刚. 故县水库实时洪水预报模型研究[D]. 南京: 河海大学, 2004.
- [10]宋星原, 叶守泽. 洪水实时预报系统模型时变参数识别与校正技术. 现代水科学不确定性研究与进展. 成都: 成都科技大学出版社, 1994.
- [11]夏达忠. 流域洪水预报系统关键技术研究[D]. 南京: 河海大学水资源环境学院, 2005.
- [12]张行南、齐晶、张丽. 流域流水网推导方法[J]. 河海大学学报, 2000(1):26—31.
- [13]Beven, K J and Kirkby, M J. 1979 A physically based variable contributing area model of basin hydrology, Hydrol.Sci.Bull.,24(1),43-69.
- [14]Wharton, G. Progress in the use of drainage network indices for rainfall-runoff modeling and runoff prediction. Progress in Physical Geography, 1994, 18: 539-557.
- [15]Hydrologic Modeling Tools---ARC/INFO user's guide.Environmental System

Reserch Institute,Inc.

- [16] 万洪涛, 万庆, 周成虎. 流域水文模型研究的进展[J]. 地球信息科学, 2000 (4) 46—50.
- [17] 赵人俊. 新安江模型与陕北模型[M]. 北京: 水利电力出版社, 1984.
- [18] 李景玉, 孙德树, 张忠国. 多支流汇入的河道洪水演算[J]. 东北水利水电, 2006 (2) 25—27.
- [19] R.K.林斯雷等, 刘光文等译. 工程水文学[M]. 北京: 水利电力出版社, 1981.
- [20] 翟家瑞. 常用水文预报算法和计算程序[M]. 黄河水利出版社, 1995.
- [21] 朱华. 水情自动测报系统[M]. 北京: 水利电力出版社, 1993
- [22] 陈述彭等. 地理信息系统导论[M]. 北京: 科学出版社, 1999
- [23] 井立阳. 数字化流域水文模拟研究[D]. 南京: 河海大学, 2005.
- [24] 陈异植, 庄希澄. 马斯京根法评述[J]. 水文水资源.
- [25] 庄一鹤, 林三益. 水文预报[M]. 河海大学出版社.
- [26] 唐勇, 胡和平, 田富强. 流域洪水计算模型系统的拓扑关系分析[J]. 水利学报, 2002 (5) 81—85.
- [27] 潘道才. 数据结构[M]. 成都: 电讯工程学院出版社, 1988
- [28] 王凤君, 刘忠, 韩玉梅. 洪水误差时序分析实时校正法[J]. 黑龙江水专学报, 1995 (1) 34—37.
- [29] 孙桂华. 实用水文预报中反馈模拟实时校正的应用[J]. 水文, 1991 (1) 24—29.
- [30] 李纪人, 刘德平. 水文时间序列模型及预报方法[M]. 南京: 河海大学出版社, 1991
- [31] 宋文尧, 张牙. 卡尔曼滤波[M]. 北京: 水利电力出版社, 1993
- [32] 葛守西. 现代洪水预报技术[M]. 北京: 中国水利电力出版社, 1999
- [33] 芮孝芳. 洪水预报理论的新进展及现行方法的适用性[J]. 水利水电科技进展, 2001 (10) 1—5.
- [34] 包为民. 洪水预报信息利用问题研究与讨论[J]. 水文, 2006 (4) 18-21.
- [35] 瞿思敏, 包为民. 实时洪水预报综合修正方法初探[J]. 水科学进展, 2003 (3) 167—171.
- [36] 夏军. 河川径流预报的水文系统理论与方法[J]. 水文科技信息, 1994.
- [37] T.R.Fortescue. et al. Implementation of self-tuning Regulalors with Variable Forgetling Factors, Automatica, Vol.17, No.6, 1981.
- [38] 覃光华, 王建华, 赵英林. 实时洪水预报研究综述及展望[J]. 城市道桥与防洪, 1999 (4) 35—39.

- [39] 胡峰. 自回归模型参数的递阶辨识[J]. 自动化学报, 1994 (7) 464—469.
- [40] 何少华. 递推最小二乘和误差自回归联合校正方法[J]. 水电能源科学, 1996 (6) 78—83.
- [41] 芮孝芳, 陈界仁河流水文学[M]. 南京: 河海大学出版社, 2003.

致 谢

本文的研究与写作是在导师及诸多师长、同学、朋友的支持与关心下完成的。

首先向尊敬的导师张行南教授表示真诚的感谢和崇高的敬意。本论文自开题到完成，整个过程都得到了导师张行南教授的悉心指导和帮助。在老师的细心指导下，我对论文的选题、研究思路和工作内容形成了清楚、具体的工作计划。论文开题后，导师经常督促我，使我在时间进度上控制得比较好，能够按时完成论文，论文完成的整个过程，都离不开导师在各方面给我的帮助。

另外，夏达忠老师的在论文的选题和进行过程中给了我很大的帮助，在研究思路给了我很多建议，让我的思维也开阔很多，论文的完成与夏老师的关心是密不可分的。非常感谢夏老师给予我的帮助！

感谢河海大学水资源环境学院、研究生院的各位领导和老师给予我的支持、关心和帮助，在此深表感谢！

感谢刘建芬博士、井立阳博士、李发文博士、周爱霞博士、杜迎燕硕士、王随霞硕士、刘永志硕士、张文婷硕士、李威硕士、赵海伟硕士、王晓航硕士、夏兰娣硕士、慈慧硕士、许进硕士、夏希硕士、许君君硕士、张晓丽硕士、王珂硕士、商胜波硕士、徐燕硕士等同门对我的帮助！

感谢同一寝室的张文明硕士、尹志杰硕士、刘猛硕士对我的帮助，是他们给我创造了良好的生活环境，三年的生活给我留下了极为难忘的记忆。

感谢一直默默关心和照顾我的父母和亲友们！感谢我的女友孙元元硕士多年来的给予我的关怀、支持与帮助，特别是在我意志消沉的时候对我的鼓励和照顾，给我提供了巨大的精神动力和支持，使我能够顺利完成毕业论文！

长江水利委员会水文局为本课题的研究提供了大量资料与工作上的支持，在此表示衷心的感谢。

段唯鑫

2006年5月