

分类号：  
密 级：

单位代码：10019  
学 号：S05111110

中国农业大学

学位论文

干旱预测模型的设计与实现

**Design and Implementation of Forecasting Models for Drought  
Monitoring**

研 究 生：\_\_\_\_\_王 彦 集\_\_\_\_\_

指 导 教 师：\_\_\_\_\_朱德海 教授\_\_\_\_\_

申请学位门类级别：\_\_\_\_\_理学硕士\_\_\_\_\_

专 业 名 称：\_\_\_\_\_地图学与地理信息系统\_\_\_\_\_

研 究 方 向：\_\_\_\_\_地理信息系统技术\_\_\_\_\_

所 在 学 院：\_\_\_\_\_信息与电气工程学院\_\_\_\_\_

2007 年 6 月

# 独 创 性 声 明

本人声明所呈交的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，也不包含为获得中国农业大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志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。

研究生签名：

时间：

年 月 日

## 关于论文使用授权的说明

本人完全了解中国农业大学有关保留、使用学位论文的规定，即：学校有权保留送交论文的复印件和磁盘，允许论文被查阅和借阅，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。同意中国农业大学可以用不同方式在不同媒体上发表、传播学位论文的全部或部分内容。

(保密的学位论文在解密后应遵守此协议)

研究生签名：

时间：

年 月 日

导师签名：

时间：

年 月 日

分类号：

密 级：

单位代码：10019

学 号：S05111110

中国农业大学

学位论文

## 干旱预测模型的设计与实现

**Design and Implementation of Forecasting Models for Drought**

**Mormitoring**

国家自然科学基金项目（40571111、40371083）、教育部科学技术研究重点项目（105013）和“十一五”国家科技支撑计划课题（2006BAD10A01）资助

研 究 生： 王 彦 集

指 导 教 师： 朱德海 教授

申请学位门类级别： 理学硕士

专 业 名 称： 地图学与地理信息系统

研 究 方 向： 地理信息系统技术

所 在 学 院： 信息与电气工程学院

2007 年 6 月

## 摘要

干旱是我国主要的自然灾害之一，具有发生频率高、持续时间长、波及范围广的特点，对我国农业生产、能源需求、环境保护和经济建设的影响日趋增加。目前关于干旱灾害方面的研究虽然取得了很大进展，但如何有效地监测旱情的动态变化并进行准确预测，仍然是众多学者共同关心的研究课题。

本研究根据干旱预测的技术路线，以干旱预测系统的总体设计为框架，在已有的AVHRR遥感数据处理模块的基础上，针对干旱预测模块、系统主界面以及干旱监测数据处理模块中的标准化降水指数（SPI）计算功能进行设计和实现。对于干旱预测模块来说，根据加权马尔可夫模型、灰色预测以及基于条件植被温度指数（VTCI）的ARIMA预测的相关理论，一方面结合该模块和已有的AVHRR遥感数据处理相关功能中所涉及到的数据，对数据结构和访问接口进行设计；另一方面根据模块功能的要求，在相关算法设计的基础上，按照面向对象程序设计的思想进行类的设计和实现。

按照所选干旱预测模型的相关理论和面向对象程序设计的思想，实现了干旱预测模块。该模块不仅能够方便和高效地访问数据，而且通过加权马尔可夫模型、基于GM (1,1)模型的灰色预测和基于VTCI的ARIMA模型的相关功能，能够进行点上和面上的干旱预测。以陕西关中平原和渭北旱塬的降水量数据和AVHRR多时段卫星遥感数据，对该模块进行了应用，结果表明该模块在功能上能较好地满足干旱预测应用的要求，达到了预期系统开发的目标。

**关键词：**遥感，干旱指数，干旱预测，模块开发

## Abstract

Drought is one of the major natural disasters in China, with high frequency, long duration and large extent. Its influence on agricultural production, energy demand, protection of the ecological environment and economic development increases gradually. Researches related to drought monitoring have made great progresses, but how to effectively monitor drought situation and dynamic changes, and to accurately forecast are still concerned

According to the study of drought forecasting and the design of the drought forecasting system, this research is about design and implementation of drought forecasting modules, user interface and the function of Standardized Precipitation Index (SPI) calculation using both remotely sensed data and ground measured precipitation. According to the theory of the Markov Chain model with weights, Grey Forecast and ARIMA model, the data formats and interface, the algorithms of the models and the programs are designed and implemented.

According to the theory of drought forecasting and the ideas of object-oriented program design, the drought forecasting modules are implemented. The modules not only can access data conveniently and efficiently, but also can forecast drought through the Markov Chain model with weights, Grey Forecast and ARIMA model. The application results of the drought forecasting modules base on the precipitation data and remotely sensed data of Guanzhong Plain and Weibei Tablelands show that the modules can satisfy the application of the prediction of drought, and achieve the expected goals of the system development.

**Key Words:** remote sensing, drought index, drought forecasting, module development

目 录

摘要 ..... I

ABSTRACT ..... II

第一章 绪论 ..... 1

    1.1 研究背景 ..... 1

    1.2 干旱预测方法的研究进展 ..... 1

    1.3 系统设计和实现相关技术研究现状 ..... 4

    1.4 研究目标、内容和技术路线..... 7

    1.5 本文章节安排..... 8

第二章 干旱预测模型 ..... 10

    2.1 概述..... 10

    2.2 干旱监测指数..... 11

    2.3 干旱预测模型 ..... 14

    2.4 小结..... 19

第三章 干旱预测系统 ..... 20

    3.1 系统概述..... 20

    3.2 系统主界面的设计和实现 ..... 21

    3.3 小结 ..... 23

第四章 干旱预测模块的设计与实现 ..... 24

    4.1 数据接口的设计与实现 ..... 24

    4.2 干旱预测模块算法的设计 ..... 29

    4.3 干旱预测模块类的设计 ..... 31

    4.4 小结..... 38

第五章 干旱预测模块的应用 ..... 39

    5.1 数据准备和处理 ..... 39

    5.2 加权马尔可夫模型预测应用 ..... 41

    5.3 基于 VTCI 的 ARIMA 预测应用 ..... 45

    5.4 小结..... 46

第六章 结论与讨论 ..... 47

    6.1 结论 ..... 47

    6.2 讨论 ..... 47

参考文献 ..... 48

致谢 ..... 52

个人简历 ..... 53

## 第一章 绪论

### 1.1 研究背景

干旱是我国主要的自然灾害之一，具有发生频率高、持续时间长、波及范围广的特点。干旱的频繁发生和长期持续给我国的农业生产、能源需求、环境保护和经济建设带来了巨大的损失。近年来，我国受旱面积和旱灾损失呈不断上升趋势；干旱发生区域逐渐扩大；旱灾影响范围也从农村向城市蔓延。干旱已经不仅是困扰我国农业生产的主要因素，而且已成为严重制约我国经济发展和社会和谐的主要因素之一。

由于干旱发生范围广、影响因素多、成因复杂，因此至今还没有一个适用于任何条件下的干旱定义及其监测、预报方法。因此，探讨一套客观、动态、实时的干旱监测和预报方法，对于及时了解旱情的程度和分布，采取有效的防旱、抗旱措施具有重要的现实意义。

传统的干旱监测是用点上的数据来研究干旱的程度及范围，其中应用最多的是气象数据，此外还有水文、社会和经济等数据。但是采用点上的监测方法不仅无法实现大面积动态监测，而且费时、费力。而遥感技术具有宏观、快速、动态、经济的特点，特别是在可见光、近红外和热红外波段能够较为精确地提取一些地表参数，能够从面上进行干旱监测。利用遥感技术研究地理区域面上的干旱已成为干旱监测研究的热点。

由于干旱的物理成因很复杂，干旱的准确预测缺乏坚实的理论基础，所以过去直至现在主要是以统计模式预测为主。随着全球气候研究计划的实施，对干旱的物理成因有了较多的认识，应用动力学气候模式进行干旱预测成为新的研究热点，但尚未形成业务预测，此方法仍处于实验阶段。目前，在干旱预测研究中使用较多的方法是以干旱指标为基础，应用时间序列分析、多元回归分析、周期分析等数理统计方法来建立预测模型。近年来，依据各种指标所建立的干旱预测模型在防旱抗旱中也发挥了重要作用。

在干旱监测和预报过程中，不仅需要处理大量数据，而且有的预测模型比较复杂，用手工方式进行处理和分析的工作量相当大。特别是将遥感技术应用到干旱研究中，这种方式是根本不可取的。因此，需要以计算机为硬件支持，结合多种类型、多种指标以及多种数学方法建立干旱监测与预报系统，达到提供决策信息，降低干旱损失、促进农业生产的目的。

### 1.2 干旱预测方法的研究进展

随着与人类生存密切相关的干旱问题的日益突出，如何有效地对旱情进行预测，已成为国内外众多学者共同关心的研究课题<sup>[1]</sup>。美国、日本、俄罗斯以及澳大利亚等发达国家相继建立了气候监测及诊断分析业务，以加强对灌溉用水和干旱灾害的研究<sup>[2]</sup>。在国内，安徽、陕西以及西藏等地区也分别建立了各类指标的干旱预测模型<sup>[3]</sup>。但是，由于干旱灾害受到诸多长期天气过程的影响，其物理成因比较复杂，这些模型大多具有比较明显的区域性和时域性，尚未形成能广泛适用的预测模型。

根据旱情影响范围和研究角度的不同，干旱可以分为四种类型：气象干旱、农业干旱、水文干旱和经济干旱<sup>[4-6]</sup>。气象干旱是指在一个相对长的时期内，某一地区的蒸发量大于降水量，或降

水量异常偏少的现象<sup>[4]</sup>，常以降水量或降水量的统计量作为干旱指标<sup>[7-9]</sup>，例如帕尔默指数、降水量距平百分率和标准化降水指数。农业干旱是由环境因素造成农作物发生水分缺乏，正常生长发育受到影响，导致减产或失收的现象，一般采用降水量指标、土壤含水量指标、作物旱情指标及综合类指标等来评价<sup>[10,11]</sup>。水文干旱是指一种持续性、地区性、和广泛性的河川流量和蓄水量较常年偏少，难以满足需水要求的水文现象<sup>[7]</sup>，通常采用  $S = M \times D$  作为其干旱指标，其中  $S$  为干旱烈度， $M$  为干旱强度， $D$  为干旱历时<sup>[3,7]</sup>。经济干旱则是指社会、经济领域水分缺乏影响生产、消费活动等的干旱现象<sup>[4]</sup>，一般以干旱所造成的经济损失作为其研究指标。Ohlsson<sup>[12]</sup>提出 Social Water Scarcity Index (SWSI) 用于反映社会所面对的干旱胁迫程度。就干旱预测研究而言，目前研究的重点还是集中在与农业生产比较密切的气象干旱和农业干旱上，而对水文干旱和经济干旱的预测研究相对比较少，只建立了少量的预测模型<sup>[3,13]</sup>。

### 1.2.1 数值预测法

数值预测法是气象干旱预测中常采用的方法之一，它是根据气象学的原理建立预测模型，即建立一系列的偏微分方程，然后根据已知的初始场求得方程的解，再得出预测结果，其本质是以天气的数值预测作为模型基础<sup>[3]</sup>。数值预测法的最大优点是客观化和定量化，但是大气运动异常复杂，在目前计算机容量和速度有限的情况下，需要对预测方程组适当简化，而简化的方程组的预测结果与实际情况往往呈现一些差距，不可能预测的十分精确，而且只能反映大尺度系统的主要活动和演变，对中小尺度系统的活动和一些次要的过程预测不出来。数值预测的时间不能外延太长，延续时间越长，预测的结果与实际出入就越大。在气象数值预测中，应用最广的预测产品是模式输出统计量(Model Output Statistics)<sup>[14]</sup>。随着计算机技术的普及和天气预测技术的提高，数值预测法已经越来越多地被应用在气象干旱的预测中，是目前气象干旱预测中比较成熟的一种手段。

### 1.2.2 统计预测法

统计预测法是用数理统计的方法分析预测因子和预测量之间的数量关系，通过建立数学模式来预测未来的干旱程度。目前，使用较多的方法就是以各种干旱指标为基础，应用时间序列分析、多元回归分析、周期分析、谱分析等数理统计方法来建立预测模型，进行干旱预测<sup>[15]</sup>。这种方法在气象干旱和农业干旱预测中都有所应用。

时间序列分析是一门应用性很强的学科，凡是研究依时间顺序而获得的观测数据的规律都离不开时间序列分析这一统计工具。时间序列分析方法主要有博克斯-詹金斯法，又称 ARIMA (Autoregressive Integrated Moving Average) 模型，灰色预测方法、马尔可夫状态转移概率法、小波分析方法等。ARIMA 模型是时间序列分析方法中最有代表性的方法。其基本思想是依据有限长度的观测数据，研究序列的顺序性以及序列中观测值之间的依存关系，建立能够比较精确地反映时间序列中所包含的动态依存关系的数学模型，然后利用该模型预测系统未来的变化。灰色系统理论是 1982 年由我国学者邓聚龙教授提出的，在建模、预测和控制等方面都有广泛的应用<sup>[16]</sup>。其实质为一种曲线拟合过程，适合描述数据序列中所蕴含的确定性的指数规律。马尔可夫过程是 20 世纪初由前苏联学者 Markov 首先提出，随着该理论的逐步发展和完善，已经被广泛应用



于自然科学、工程技术以及经济管理等各个领域。马尔可夫预测法是将时间序列看成一个随机过程,通过对事物不同状态的初始概率和状态之间转移概率的研究,确定状态变化趋势,以预测事物的未来。小波分析是近几年比较前沿的研究领域,其实质上是时间序列在时间和频率两个方向上展开,进而对时间序列的时频结构作细致的分析,以提取有用的信息。

目前,将时间序列分析的各种方法应用于干旱预测的研究也有报道,例如 Modarres<sup>[17]</sup>应用季节性 ARIMA 模型预测伊朗伊斯法罕省 Zayandehrud 河的月径流,并选用不同的干旱阈值,对该地区的干旱预测进行了研究。Mishra<sup>[18]</sup>等也应用该模型来分析标准降水指数(SPI)时间序列,预测了印度 Kansabati 盆地未来 2 个月的干旱情况。Yurekli<sup>[19]</sup>等用 ARIMA 建模方法模拟 Kelki 河五年的月径流的观测数据。Fernando<sup>[20]</sup>等用不同的 ARMA 模型预测月降雨量情况。许建国<sup>[21]</sup>结合降水资料,分析了灰色理论中 GM(1,1)模型和马尔可夫模型在流域旱涝灾害预测中的运用。Lohani<sup>[22, 23]</sup>等以帕默尔干旱指数的干旱监测结果为基础,采用非线性马尔可夫链进行了干旱评估和早期预警。Paulo<sup>[24]</sup>等结合葡萄牙南部 7 个气象站 68 年的 SPI 数据,通过将马尔可夫模型和对数线性模型相结合来进行干旱的监测和早期预报。孙才志<sup>[25, 26]</sup>等也以降水量为数据采用加权马尔可夫模型对降水丰枯状况进行了预测。王希娟<sup>[27]</sup>等应用线性趋势分析和小波分析方法对青海东部地区 12 个气象台站 1961-2004 年春季降水的变化特征及时频特征进行了分析。

近年来,谱分析也被应用于干旱预测,但是由于谱分析过程都比较复杂,研究人员通常要具有很强的数学基础才能熟练地使用它,同时它的分析结果也比较抽象,不易于进行直观解释,导致谱分析方法的使用具有很大的局限性,如刘引鸽<sup>[28]</sup>等以降水量减蒸发量干旱指数,应用波谱分析与逐步自回归方法,分别建立了西北地区干旱指数的拟合回归预测模型。

### 1.2.3 干旱集成预测方法

随着计算机技术的迅速发展和普及,以及人们对于干旱预测精确性和实用性要求的提高,一种以计算机为硬件支持,结合多种类型、多种指标以及多种数学方法的干旱集成预测方法逐渐成为研究的热点。

近年来,在我国不仅开展了干旱集成预测的研究,而且建立了相应的预测业务系统。这些干旱预测应用系统虽然从系统平台、开发语言、系统功能和预测模型等方面来说都不尽相同,但是它们都是以计算机技术为支持,采用多种指标和方法来进行干旱预测。例如,李玉爱<sup>[29]</sup>等设计开发了大同市短期农业气候干旱预测系统,该系统运用最优子集回归、方差分析、平稳时间序列、概率转移预测、模糊聚类分析、曲线图模拟概率预测法、经验公式指标法、多因子相关相似指数法、综合决策集成预测等来进行干旱预测。系统具有资料采集、资料处理及各种预测方法模块、集成预测、结果输出等功能。李正明<sup>[30]</sup>等利用甘肃中部 13 个监测站历年 4~6 月的旬降水、温度资料,分析了干旱年、月、旬降水量、平均温度等气候变化特征及其空间分布,在此基础上建立了相应的春、夏干旱短期气候预测系统模型,并以 Windows NT 为操作系统,采用 TCP/IP 协议,应用 Internet 技术,以 SQL Server 数据库为核心,建成了集信息采集、预测产品制作加工、服务分发为一体的专业气象预报服务平台。李凤霞<sup>[31]</sup>等以地面常规监测、气候监测和空间遥感监测相结合为基础,设计和建立了“监测—评估—预测、预警”信息业务系统,可为研究干旱发生规律,及时监测和预测干旱的发生,评估干旱程度以及制定防灾、减灾、救灾提供科学依据。景毅刚<sup>[32]</sup>

等采用 Visual Basic 6.0 程序设计开发了陕西省干旱评价和预警系统,该系统结合相应气象水文资料、农作物发育期、50 cm 深土壤相对含水量、土壤物理参数和不同农作物在不同发育阶段需水分析,并采用气温、降水指标、农田缺水率指标及农田水分供需比指标等来进行干旱预测,能够预测未来出现某种程度干旱的日期等。杨太明<sup>[33]</sup>等在 Windows 环境下,利用分布式数据库以及 SQL Server 与其它高级语言开发的 OCX、COM 控件,实现气象遥感资料的处理、地面观测资料数据转换、土壤墒情监测站资料的接收分析,然后基于农田土壤水分平衡原理进行干旱预测,开发了安徽省干旱灾害监测及预警服务系统。刘治国<sup>[34]</sup>等开发了“西北干旱监测预测业务服务综合系统 v2.0”,该系统包括干旱资料库、干旱监测诊断、干旱气候预测、干旱影响评价、干旱预测评估、产品分发和辅助工具 7 个部分。该系统依托互联网和 MICAPS 系统能实现动态查询,资源共享以及资料的及时更新。Kung<sup>[35]</sup>等针对台湾近年来频繁发生的干旱灾害,提出并开发了一个干旱预警系统。该系统包括移动用户、生态监测传感器、集成服务服务器和智能干旱决策系统四个部分。此外,该系统结合无线传感器网络,嵌入式多媒体通讯和神经网络决策的相关技术能有效地实现干旱预测。

国外在这方面的研究比国内相对更成熟。Kogan<sup>[36]</sup>提到 80 年代末美国就开始用 NOAA 极轨气象卫星进行干旱监测,在全国建立了干旱监测系统网络。Mendicino<sup>[37]</sup>等指出从 2000 年开始意大利就应用干旱预测系统来预测其南部水分的空间变化。该系统是以一个嵌入式的地理信息系统和一个与数据接收系统相连接的模型为基础的。Sheffield<sup>[38]</sup>等提到一个实验性的实时干旱监测和季节预测系统已经在美国使用,并正在向全球推广。这个系统由三部分组成:首先是一个以 50 年数据为基础的干旱气候模型;其次是一个实时监测子系统,它能每天更新水文数据;最后是一个干旱预测子系统,它采用概率统计技术,对未来 9 个月土壤湿度进行季节性的预测。

对上述系统进行综合分析,不难发现它们都是以计算机技术为硬件支持,同时又集合了多种干旱指标和预测方法,并且在功能上也不仅仅局限于干旱预测,充分考虑了干旱研究和应用的功能需求。这也是干旱集成预测方法优势的体现,它与单类型和单指标的预测模型相比,在精确性和实用性方面有一定提高,因此干旱集成预测方法的研究可以促使干旱预测进入大面积的实际应用阶段,代表着未来干旱预测研究的一个发展趋势。

## 1.3 系统设计和实现相关技术研究现状

在以遥感影像为主要数据源的干旱预测过程中,需要对大量的遥感影像进行处理,而遥感影像的数据量都比较大,并且在处理过程中有时可能涉及到多幅遥感影像,使数据的访问量成倍增长。因此,需要采用相应的计算机技术来提高对遥感数据处理的效率。目前,常用的方法有数据分块、内存映射文件等技术。

对于系统的设计和实现,也需要应用相关技术。UML (Unified Modeling Language) 作为一种定义良好、易于表达、功能强大且普遍适用的建模语言,主要用于系统的设计阶段。而软件工程的思想则贯穿整个系统的开发过程中。

### 1.3.1 数据分块

一般而言,遥感影像的数据量都比较大,而普通计算机的内存容量是有限的<sup>[39]</sup>。如果一次将

整幅影像数据调入来进行计算和处理的话,不仅效率低、执行速度缓慢,而且容易造成系统崩溃,因此采用数据分块的策略是进行遥感影像处理的有效方法。

遥感影像数据的自然组织形式是一个  $N \times M$  的像素矩阵,并且通常情况下是按行或按列存储的。从一幅影像中读取一小块影像,例如  $n \times m$ ,有很多种实现方式,其中比较有效的方式主要有两种<sup>[40]</sup>。一种是直接在磁盘上按行或按列定位读指针,然后读取需要的影像,这种方式需要移动读指针  $n$  次或  $m$  次。另一种是先把  $n \times M$  (按行顺序存储)或  $N \times m$  (按列顺序存储)大小的影像一次性读到内存,然后再在内容中提取需要的影像。我们可以根据  $n/N$  或  $m/M$  的值来决定采用哪种读取方式,如果遥感影像是按行顺序存储的,当  $m/M$  的值比较小时,前一种方式的效率比较高,当  $m/M$  的值比较大时,则后一种方式的效率比较高。

按行或列顺序存储是一种基本的存储方式,但不是一种高效的存储方式。为了提高对海量遥感影像的管理,一般都要对遥感影像进行分块处理,形成影像子块,然后以此为单位进行存储,并建立索引。目前关于影像分块已经有了很广泛的应用,特别是在影像数据库中几乎都要进行影像分块。使用分块策略需要对原始影像进行分割,再根据需要对各块影像编码和建立索引,这样用到哪块影像时就调用这一块,大大提高了效率。

分块技术在遥感影像处理中的应用十分广泛。利用分块技术可以提高影像数据显示等的效率。例如,蒋红成<sup>[41]</sup>等运用分块调度算法实现了SAR海量影像数据的快速浏览。余粉香<sup>[42]</sup>等针对大数据量影像数据的快速调度与显示问题,以影像分层、分块的金字塔处理技术为基础,提出了一种基于文件的数据组织与管理方法,实现了单幅和多幅大数据量影像的实时调度与快速显示。马照亭<sup>[43]</sup>等使用数据分块技术提高了地形漫游的效率和可视化效果。此外,分块技术在数据压缩和构建影像金字塔上也有所应用。例如,Kwok<sup>[44]</sup>等使用块自适应量化方法对SAR影像数据进行压缩处理。宋江洪<sup>[45]</sup>等通过建立遥感影像的分块分层结构和快速索引机制,分析了该结构在快速缩放、漫游、显示和灰度值修改等功能上的实用性与合理性。

### 1.3.2 内存映射文件

在进行数据处理的过程中,有时需要同时访问几幅甚至几十幅影像。如果同时调入多幅影像进行处理,计算机系统资源的消耗是相当大的。此外,在 Windows 系统中,超过 2GB 的数据不能一次性调入内存处理。因此,针对这一问题除了采用前面提到的分块技术外,还涉及到另一种提高大数据量处理效率的计算机方法:内存映射文件。

内存映射文件技术是Windows操作系统提供了一种文件数据存取机制。利用内存映射文件技术,系统可以在2GB的地址空间中为文件保留一部分空间,并将文件映射到这块保留空间。一旦文件被映射之后,操作系统将管理页映射、缓冲以及高速缓冲等任务,而不需要调用分配、释放内存块和文件输入/输出的API函数,也不需要自己提供任何缓冲算法。使用内存映射文件能给程序开发工作提供极大的方便,程序的运行效率也非常高。

内存映射文件分三种情况:第一种是可执行文件的内存映射;第二种是数据文件的内存映射;第三种是借助于页面交换文件的内存映射。可执行文件的内存映射主要由Windows自身使用,一般的应用程序主要使用后面两种内存映射文件。

关于该方法的相关研究也有报道,例如,杨宁学<sup>[46]</sup>等介绍了如何用Windows API 将文件映射

至内存,并通过实验数据得出,在对大数据量文件的处理中采用内存映射文件的效率远远高于传统的基于I/O方式的数据存取。胡伟忠<sup>[47]</sup>结合对海量遥感影像进行处理的要求,根据实际提出了一种快速读取海量图像的方法,而实现该方法的主要技术就包括内存映射文件和图像分块等相关技术。

### 1.3.3 UML

UML(Unified Modeling Language)由Booch、Rumbaugh和Jacobson三人于1996年提出,OMG(Object Technology Group)采纳UML1.1作为基于面向对象技术的统一建模语言。它具有定义良好、易于表达、功能强大且普遍适用的特点,已成为业务、应用和系统架构的标准可视化建模语言。

UML的出现,大大加快了建模技术在软件产业的应用和推广,它的发展现状可以归结于以下几点:

(1) UML已进入全面应用阶段的事实标准:从UML的早期版本开始,它便受到计算机产业界的重视,OMG的采纳和大公司的支持把它推上了实际上的工业标准地位,使它拥有越来越多的用户<sup>[48]</sup>。目前市场上各种建模工具中80%以上是基于UML的,在工业界、科技界和应用界UML获得了广泛支持。

(2) UML的应用领域正在逐渐扩展:OMG新批准的UML2.0版本中除了增强了基础设施、增加了新的建模能力、使模型交换更加简单之外,还增加了许多可扩展性。现在不仅在应用软件建模中广泛使用,还逐渐推广到嵌入式系统建模、业务建模、流程建模等多个领域。

(3) UML成为“产生式编程”的重要支持技术:可执行UML的出现为理想的自动产生代码的编程模式提供了许多新的发展基础,可以通过在UML的基础上进行形式化描述以生成可执行的代码和组件。

### 1.3.4 软件工程

软件工程是在克服上世纪60年代末出现的“软件危机”的过程中逐渐形成与发展起来的。自1968年在北大西洋公约组织(NATO)举行软件可靠性的学术会议上正式提出软件工程概念以来,在不到40年的时间里,软件工程在理论和实践两方面都取得了长足的进步。

软件工程是一门指导计算机软件开发和维护的科学。它包括两方面内容:软件开发技术和软件项目管理<sup>[49]</sup>。其中,软件开发技术包括软件开发方法学、软件工具和软件工程环境,软件项目管理包括软件度量、项目估算、进度控制、人员组织、配置管理、项目计划等。它的目标在于研究一套科学的工程化方法,并与之相适应发展一套方便的工具与环境,供软件开发者使用。

软件工程提出了软件开发遵循的基本过程,一般由需求分析、概要设计、详细设计、编码、测试、运行维护等环节构成。需求分析为软件开发的目标系统提出完整、准确、清晰和具体的要求。概要设计就是在需求明确的基础上,站在全局的高度,从较抽象的层次上分析对比多种可能的实现方案和软件结构,从中选出最佳方案和最合理的软件结构,从而用较低成本开发出较高质量的软件系统。详细设计是确定具体实现所要求的目标系统的方案,程序流程图、盒图等是完成详细设计的工具。程序编码是将设计思想转化为计算机可识别的指令的过程,一般需要借助一定

的开发工具和开发环境。这将形成系统的可执行文件、库文件、数据文件及相关的支持文件等。测试是软件开发过程中不可忽视的重要环节,只有通过合理的系统的测试,才能确保软件的质量。维护环节是在软件已经交付使用之后,为了改正错误或满足新的需要而修改软件的过程。

软件工程是一门非常实用的学科,它的发展是实际的软件开发所驱动的。随着我国信息化程度的提高,对软件的需求量和规模也会迅速扩大,使用软件工程的方法来开发计算机应用软件是提高软件开发效率和软件质量的重要保证。

## 1.4 研究目标、内容和技术路线

### 1.4.1 研究目标和内容

本研究以已有的遥感干旱监测数据处理的功能和相关干旱预测模型为基础,设计和实现干旱预测模块,并将数据处理相关功能和预测模块集成在同一系统中,使该系统能够集遥感数据预处理、干旱监测指标计算和干旱预测于一体,达到提高干旱监测数据处理能力和进行干旱预测的目的。具体研究内容包括:

- 根据干旱预测的技术路线,从干旱预测系统整体构架、模块和功能的划分等方面对系统进行总体上的设计。
- 根据遥感干旱监测数据处理相关功能所涉及到的数据,并结合干旱预测的需要,设计干旱预测模块的数据格式和数据接口,使两个模块间能够数据兼容。
- 根据干旱预测模型,采用 UML 进行模型设计,然后以 Visual C++6.0 为编程平台,实现干旱预测模块。该模块不仅需要能够高效率地进行数据分析、处理,而且预测结果也需要通过图形化的方式直观地显示出来。这一部分是本研究的重点。
- 以研究区域(关中平原和渭北旱塬)的相关数据为基础,应用本系统的相关功能对数据进行处理,并进行干旱预测的应用。

### 1.4.2 技术路线

根据以上的研究内容,本研究所采取的技术路线如图 1-1。

本研究大致分为四个阶段,第一阶段是对本研究所涉及到的相关理论的学习,包括遥感影像处理方法,干旱监测和预测等理论。第二阶段是相关模型的建立,主要有通过降水量数据计算 SPI 的模型和干旱预测模型。第三阶段是模型的设计和实现。第四阶段是以研究区域的数据为基础,采用系统提供的相关功能进行实验性应用。

第三阶段是整个研究过程的重点。该阶段基本按照软件工程的思想进行,首先需要进行需求分析,主要包括数据需求、功能需求和性能需求等几个方面。就数据需求而言,为了使经过遥感干旱监测数据处理模块处理后的数据不需要任何格式转换就能被干旱预测模块使用,因此干旱预测模块需要支持数据处理模块所使用的 RAW 格式遥感影像。就功能需求而言,根据干旱预测模型,尽可能地合理划分各个功能,既能实现干旱预测的目的,又使用户在使用上更加方便。就性能需求而言,考虑到一般图像数据量都比较大,而普通 PC 机的 CPU 速度和内存容量有限,如果一次全部调入所有的影像数据进行计算和处理,效率非常低,甚至会造成系统崩溃、死机。因此,

采用分块策略来进行处理。在通过需求分析对数据以及功能等需求了解的基础上, 下一步就是进行模块的总体设计, 包括对复杂功能的进一步分解以及设计模块结构等。接下来是详细设计, 主要是对模块涉及到的算法进行详细的描述, 从而在编码阶段可以把对这些算法的描述直接翻译成用某种程序设计语言书写的程序。对于总体设计和详细设计可以利用 UML 来辅助进行。设计完成后就要进行编码, 也就是功能模块的实现。本研究以 Visual C++6.0 为编程平台, 采用面向对象的 C++ 语言来实现相关模块和功能。编码完成后需要对整个模块进行系统地测试, 找出存在的漏洞, 并及时修改。到此, 干旱预测模块基本实现, 下面就是对遥感干旱监测数据处理模块和干旱预测模块进行集成。为了使集成后的系统功能更加完善, 以 ArcGIS 提供的 ArcGIS Engine 为二次开发工具进行 GIS 相关功能的开发。在对集成后的系统进行调试之后, 系统的开发基本完成。第三阶段中的各个步骤之间相互依赖, 有的步骤还需要反复进行才能达到预期的要求。最后就是系统的发布和试用。

## 1.5 本文章节安排

本章是论文的概述, 简述了研究的目的、意义以及研究内容和方案, 并作为引言概述了国内外的研究现状。第二章对本研究所涉及到的干旱指数以及干旱预测模型的相关理论作了较为详细的介绍。第三章从遥感干旱预测系统整体构架、模块划分和功能设计等方面对系统进行了概述, 明确了本研究所要实现的模块和功能包括系统主界面、干旱预测模块和 SPI 计算功能, 并对主界面进行了设计和实现。第四章是本研究的重点内容, 针对干旱预测模块的数据格式和访问接口、算法流程以及相关功能这三个方面的设计和实现作了详细的阐述。第五章主要介绍了以实验区域的相关数据对遥感干旱预测模块中的部分功能进行应用的情况。第六章是本研究的结论和展望, 提出了所得的结论和需要进一步研究的内容。

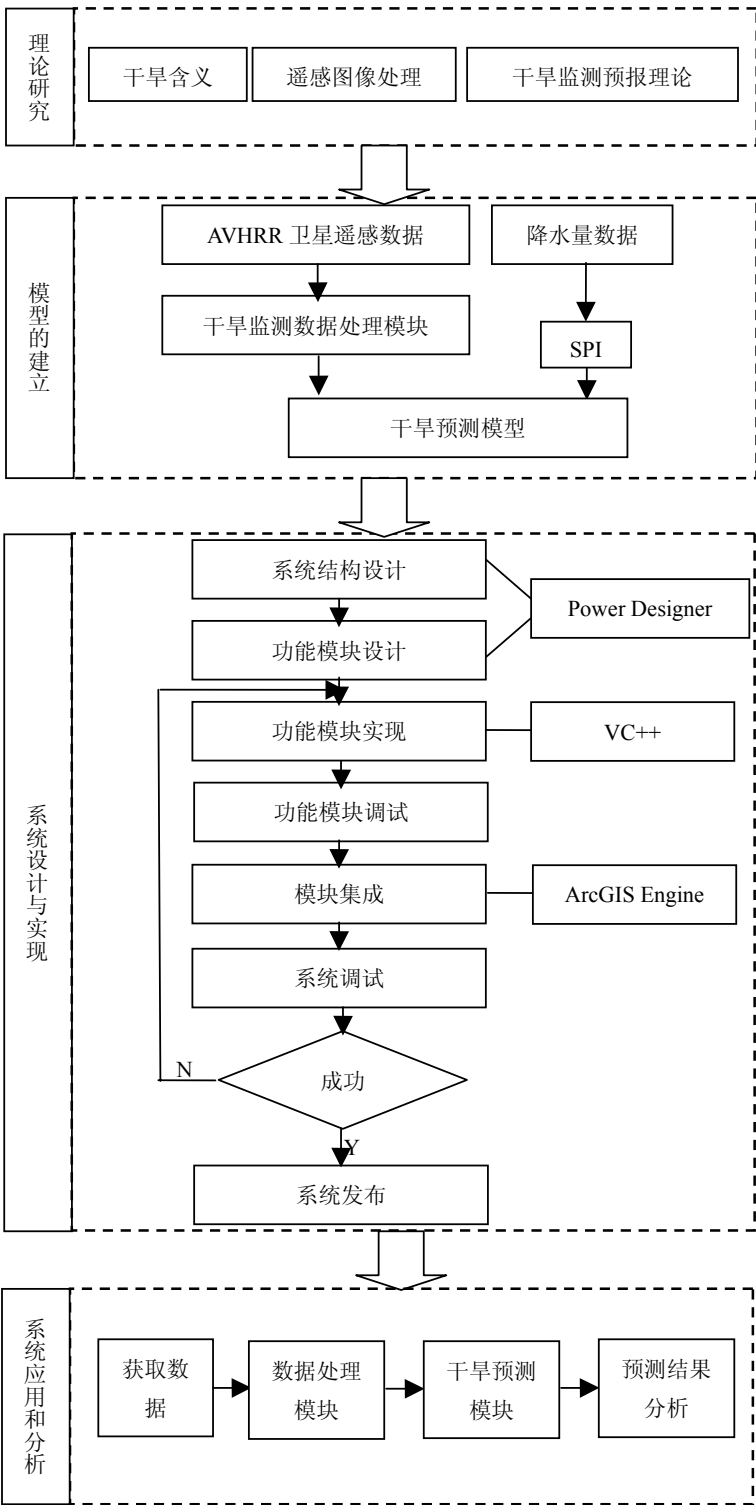


图 1-1 技术路线

## 第二章 干旱预测模型

### 2.1 概述

以气象数据和 AVHRR 多时相卫星遥感数据为主要数据源, 根据基于归一化植被指数和土地表面温度的干旱监测方法所获得的干旱监测指数, 采用时间序列分析的相关理论建立干旱预测模型, 实现面上或点上的干旱预测, 为农业决策者和生产者服务。在干旱预测的研究方案和技术路线 (如图 2-1) 中, 主要采用的干旱监测方法包括: 基于植被指数的方法、基于土地表面温度的方法、基于二者相结合的方法以及基于降水指数的方法。根据这些干旱监测方法, 可以得到相应的干旱监测指数。这些干旱监测指数在一定程度上反映了干旱状况, 是进行干旱预测的数据基础。对于干旱预测而言, 采用了加权马尔可夫模型、基于 GM (1, 1) 的灰色预测模型和基于 VTCI 的 ARIMA 预测模型。这些干旱预测模型都是以干旱监测指数为基础建立的。

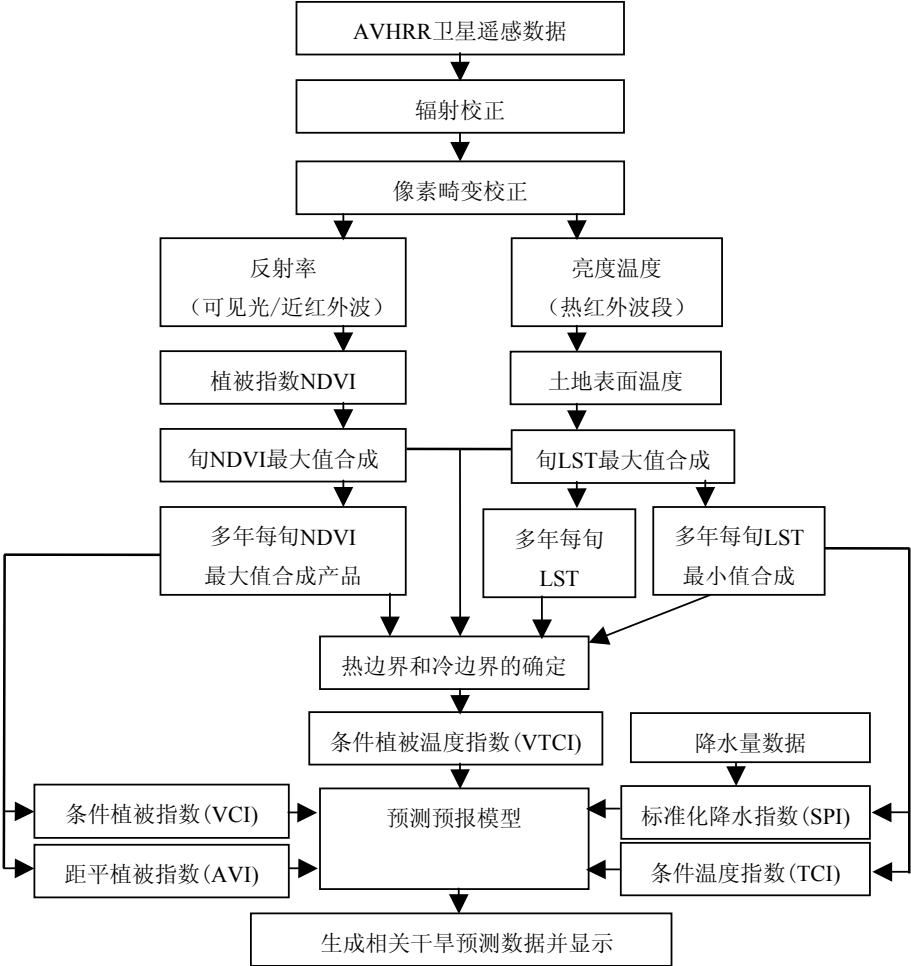


图 2-1 干旱预测技术路线



## 2.2 干旱监测指数

在遥感干旱预测技术路线中所涉及到的干旱监测指数包括：基于降水指数方法中常用的标准化降水指数；基于植被指数方法中的归一化植被指数、距平植被指数和条件植被指数；基于土地表面温度方法的条件温度指数；还有基于植被指数与土地表面温度相结合方法的条件植被温度指数。

### 2.2.1 标准化降水指数

McKee<sup>[50]</sup>等在评估美国科罗拉多干旱状况时提出标准化降水指数(Standardized Precipitation Index, SPI)。该指数能够较好地反映干旱强度和持续时间，而且具有多时间尺度应用的特性，使得用同一个干旱指数反映不同时间尺度和不同方面的水资源状况成为可能，因而得到广泛应用<sup>[51, 52]</sup>。Bussay<sup>[53]</sup>等利用 SPI 研究了匈牙利的干旱，得出 SPI 适合于不同类型的干旱定量化研究，时间尺度为 5-24 月的 SPI 可用于反映地下水位的变化；2-3 月的 SPI 可反映农业干旱情况。袁文平<sup>[54, 55]</sup>分析了 SPI 的应用特性及其在我国的适用性，发现 SPI 计算简单可行，在我国不同气候区域都具有良好的计算稳定特性，优于降水 Z 指数，其多时间尺度应用特性可以满足不同地区、不同应用的需求，可以为我国的水资源评估和不同时间尺度的干旱监测服务。

标准化降水指数 (SPI) 是将某一时间尺度的降水量时间序列看作服从  $\Gamma$  分布，通过降水量的  $\Gamma$  分布概率密度函数求累积概率，再将累积概率正态标准化而得<sup>[54]</sup>。正态标准化处理的目的是消除降水量在时空分布上的差异，使 SPI 能够适用于反映不同地区、不同时间尺度的旱涝情况。其计算方法如下<sup>[18, 54]</sup>：

假设某一时间尺度下的降水量为  $x$ ，则其  $\Gamma$  分布的概率密度函数为：

$$g(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} (x > 0) \quad (2-1)$$

式中， $\alpha > 0$ ， $\beta > 0$ ，分别为形状参数和尺度参数， $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$  为 Gamma 函数。

最佳的  $\alpha$ 、 $\beta$  估计值可用极大似然估计法求得：

$$\hat{\alpha} = \frac{1 + \sqrt{1 + 4A/3}}{4A} \quad (2-2)$$

$$\hat{\beta} = \bar{x} / \hat{\alpha} \quad (2-3)$$

式中， $A = \ln(\bar{x}) - \frac{\sum \ln(x)}{n}$ ； $n$  为降水量序列的长度。给定时间尺度的累积概率计算方法为：

$$H(x) = u + (1-u)G(x) \quad (2-4)$$

式中,  $G(x) = \frac{1}{\Gamma(\hat{\alpha})} \int_0^x t^{\hat{\alpha}-1} e^{-t} dt$ ,  $t = \frac{x}{\hat{\beta}}$ ;  $u$  是降水量为 0 的概率,  $u = m/n$ ,  $m$  为序列中降水量为 0 的数量。

将累积概率  $H(x)$  转换为标准正态分布函数, 即可获得  $SPI$  的计算方法:

$$SPI = -(k - \frac{2.52 + 0.80k + 0.01k^2}{1 + 1.43k + 0.19k^2 + 0.001k^3}) \quad 0 < H(x) \leq 0.5 \quad (2-5)$$

$$SPI = (k - \frac{2.52 + 0.80k + 0.01k^2}{1 + 1.43k + 0.19k^2 + 0.001k^3}) \quad 0.5 < H(x) < 1 \quad (2-6)$$

式中,  $k = \sqrt{\ln(1/(H(x))^2)}$ ,  $0 < H(x) \leq 0.5$ ;  $k = \sqrt{\ln(1/(1-H(x))^2)}$ ,  $0.5 < H(x) < 1$ 。

### 2.2.2 归一化植被指数

归一化植被指数 (Normalized Difference Vegetation Index, NDVI) 的定义为:

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}} \quad (2-7)$$

式中,  $\rho_{NIR}$  和  $\rho_{RED}$  分别为地表在近红外波段和红光波段的反射率。

由于 NDVI 可以消除大部分与仪器定标、太阳角、地形、云阴影和大气条件有关辐照度的变化, 增强了对植被的响应能力, 是目前已有的 40 多种植被指数中应用最广的一种<sup>[56]</sup>。虽然归一化植被指数能直接反映作物的长势, 但一般不直接用于干旱监测, 而是作为建立其它干旱监测指数的基础。

### 2.2.3 距平植被指数

距平植被指数 (Anomaly Vegetation Index, AVI) 的定义为<sup>[57]</sup>:

$$AVI = NDVI_i - \overline{NDVI} \quad (2-8)$$

式中,  $NDVI_i$  为某一特定年某一时期(如旬、月等) NDVI 的值,  $\overline{NDVI}$  为多年该时期 NDVI 的平均值。引入距平植被指数概念的目的是将 NDVI 的变化与天气、气候研究中“距平”的概念联系起来, 对比分析 NDVI 的变化与短期的气候变化之间的关系。对选定的某个地区, AVI 是 NDVI 的当年值与参考年(多年)平均值的差异。若 AVI 为正, 则表示植被生长相对于平均情况较好; 若 AVI 为负, 则表示植被生长较平均情况差。

### 2.2.4 条件植被指数

条件植被指数 (Vegetation Condition Index, VCI) 的定义为<sup>[58]</sup>:

$$VCI = \frac{NDVI_i - NDVI_{\min}}{NDVI_{\max} - NDVI_{\min}} \times 100 \quad (2-9)$$

式中,  $NDVI_i$  为某一特定年第  $i$  个时期的 NDVI 值,  $NDVI_{\max}$  和  $NDVI_{\min}$  为研究年限内第  $i$  个时期 NDVI 的最大值和最小值。

上式分母部分在一定意义上代表了 NDVI 的最大变化范围, 反映了当地植被在研究年限内的最好和最差的生长情况; 分子部分则表示植被当年长势与历史最坏情况的差别, 若  $NDVI_i$  和  $NDVI_{\min}$  之间的差值较小, 则表明该时段作物长势很差。

### 2.2.5 条件温度指数

条件温度指数 (Temperature Condition Index, TCI), 其定义为<sup>[59]</sup>:

$$TCI = \frac{BT_{\max} - BT_i}{BT_{\max} - BT_{\min}} \times 100 \quad (2-10)$$

式中,  $BT_i$  为某一特定年第  $i$  个时期的 AVHRR 第 4 波段亮度温度(Brightness Temperature)的值;  $BT_{\max}$  和  $BT_{\min}$  分别为所研究年限内第  $i$  个时期亮度温度的最大值和最小值。TCI 愈小, 表示愈干旱。

### 2.2.6 条件植被温度指数

距平植被指数、条件植被指数和条件温度指数等方法适用于年度间相对干旱程度的监测, 对某一研究区域来说, 这些方法均可得出区域内某一时期的干旱程度及范围。然而由于干旱发生的时间和地点存在着时空变异, 所以在像素水平上, 这些方法所使用的指标就有可能不同, 进而造成某一特定时期内不同像素间监测结果的可比性较差。王鹏新等<sup>[60]</sup>提出了条件植被温度指数 (Vegetation Temperature Condition Index, VTCI), 其定义为:

$$VTCI = \frac{LST_{NDVI_i \cdot \max} - LST_{NDVI_i}}{LST_{NDVI_i \cdot \max} - LST_{NDVI_i \cdot \min}} \quad (2-11)$$

其中,  $LST_{NDVI_i \cdot \max} = a + b \cdot NDVI_i$ ;  $LST_{NDVI_i \cdot \min} = a' + b' \cdot NDVI_i$

式中,  $LST_{NDVI_i \cdot \max}$  和  $LST_{NDVI_i \cdot \min}$  分别表示在研究区域内, 当  $NDVI_i$  值等于某一特定值时的土地表面温度的最大值和最小值;  $a$ 、 $b$ 、 $a'$ 、 $b'$  为待定系数, 可通过绘制研究区域的 NDVI 和 LST 的散点图近似地获得。如何确定四个待定系数是应用 VTCI 干旱监测方法的关键, 可以使用最大—最小值合成方法, 即应用多年的旬 NDVI 最大值合成产品和基于每年的旬 LST 最大值合成产品再合成的多年 LST 最小值产品确定冷边界的系数  $a'$  和  $b'$ ; 使用最大值合成技术, 即应用多年的旬 NDVI 和 LST 最大值合成产品确定热边界的系数  $a$  和  $b$ 。

VTCI 既考虑了某一区域内归一化植被指数的变化, 又考虑了在归一化植被指数值相同条件下土地表面温度的变化, 更适合于监测某一特定年内某一时期区域级的干旱程度, 具有地域专一性和时域专一性的特点, 可用于研究干旱程度的空间变化特征。王鹏新等<sup>[61]</sup>应用 MODIS 数据对 VTCI 干旱监测模型进行了验证, 证实 VTCI 是一种近实时的遥感干旱监测模型。

## 2.3 干旱预测模型

### 2.3.1 加权马尔可夫模型

马尔可夫过程的最基本特征是在已知过程“现在”的条件下，其“将来”的状态不依赖于“过去”的状态，即“无后效性”，其数学表述如下<sup>[62]</sup>：

设随机序列  $\{X(n), n = 0, 1, 2, \dots\}$  的离散状态空间为  $E = \{1, 2, \dots, N\}$ 。若对于任意  $m$  个非负整数  $n_1, n_2, \dots, n_m$  ( $0 \leq n_1 < n_2 < \dots < n_m$ ) 和任意自然数  $k$ ，以及任意  $i_1, i_2, \dots, i_m$ ， $j \in E$ ，满足

$$P\{X(n_m+k)=j | X(n_1)=i_1, X(n_2)=i_2, \dots, X(n_m)=i_m\} = P\{X(n_m+k)=j | X(n_m)=i_m\} \quad (2-12)$$

则称  $\{X(n), n = 0, 1, 2, \dots\}$  为马尔可夫链。

加权马尔可夫链预测的基本思想是考虑到一系列相依的随机变量，其各阶自相关系数刻画了各种滞时状态间相关关系的强弱，分别依其以前若干时段指数值的状态进行预测，然后按前面各时段与该时段相依关系的强弱进行加权求和，最后根据各状态加权求和的结果来预测该时段可能所处的状态。在本研究中采用加权马尔可夫链进行预测的步骤如下：

(1) 选择干旱监测指数，并确定其马尔可夫链的状态空间 ( $E = \{1, 2, \dots, N\}$ )。对于有明确干旱等级划分的指数，如标准化降水指数，其状态空间的确定就采用对该指数的干旱等级划分。而对于目前没有明确等级划分的指数，但其指数值在一定取值范围内，通过将该取值范围平分为若干小段来确定其状态空间。对于分段数目的确定，如果分段数目太少，则预测出的结果比较粗，难以满足要求；但如果分段数目太多，计算需要的时间比较长，而预测结果也并不是分段数目越多越好。因此，要根据所选的具体干旱监测指数来确定分段数目，以达到较好的预测效果。

(2) 根据状态空间确定该指数序列中各值所处的状态，然后计算不同滞时的马尔可夫链的转移概率矩阵。

(3) 对指数序列进行马尔可夫性检验，检验通过的序列才能作为马尔可夫链来处理。对离散序列的马氏性检验通常可用  $\chi^2$  统计量来进行，具体方法如下<sup>[62, 63]</sup>：

用  $f_{ij}$  表示指数值序列中从状态  $i$  经过一步转移到达状态  $j$  的频数， $i, j \in E$ 。将转移频数矩阵的第  $j$  列之和除以各行各列的总和所得的值称为边际概率，记为  $p_{\bullet j}$ ：

$$p_{\bullet j} = \sum_{i=1}^m f_{ij} / \sum_{i=1}^m \sum_{j=1}^m f_{ij}, \text{ 且 } p_{ij} = f_{ij} / \sum_{j=1}^m f_{ij} \quad (2-13)$$

当  $n$  充分大时，统计量

$$\chi^2 = 2 \sum_{i=1}^m \sum_{j=1}^m f_{ij} \left| \log \frac{p_{ij}}{p_{\bullet j}} \right| \quad (2-14)$$

服从自由度为  $(m-1)^2$  的  $\chi^2$  分布。

给定显著性水平  $\alpha$ ，查表可得分位点  $\chi_{\alpha}^2((m-1)^2)$  的值，计算后得统计量  $\chi^2$  值。若  $\chi^2 > \chi_{\alpha}^2((m-1)^2)$ ，则可认为序列  $\{x_i\}$  符合马尔可夫性，否则可认为该序列不可作为马尔可夫链来处理。

(4) 计算各阶自相关系数  $r_k$ ：

$$r_k = \sum_{i=1}^{n-k} (x_i - \bar{x})(x_{i+k} - \bar{x}) / \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2-15)$$

$k$  为滞时(步长),  $k=1,2,\dots,m$ ,  $r_k$  为第  $k$  阶滞时的自相关系数,  $x_i$  为指数序列中第  $i$  时段的价值,  $\bar{x}$  为序列均值,  $n$  为序列的长度。

规范化各阶自相关系数, 得到各种滞时的马尔可夫链的权重  $w_k$ :

$$w_k = |r_k| / \sum_{k=1}^m |r_k| \quad (2-16)$$

(5) 分别以预测时段前面若干滞时的指数所处状态为初始值, 结合其相应的转移概率矩阵, 预测该时段指数的状态概率  $p_i^{(k)} (i \in E)$ 。

(6) 将同一状态的各预测概率加权和作为该指数处于该状态的预测概率  $p_i$ :

$$p_i = \sum_{k=1}^m w_k p_i^{(k)} \quad i \in E \quad (2-17)$$

(7) 计算预测时段的状态特征值  $S^{[25]}$ :

$$S = \sum_{i=1}^N i \times (p_i^\eta / \sum_{i=1}^N p_i^\eta) \quad i \in E \quad (2-18)$$

式中,  $\eta$  为最大概率的作用系数, 其值越大, 越突出最大概率的作用。状态特征值的最大优点是最大概率和其它概率的影响进行综合考虑, 能够较好地反映预测时段所处的状态。确定  $\eta$  取值的方法为: 从步骤(5)计算出的各状态预测概率值中找出最大( $p_{Max}$ )和次大( $p_{SecondMax}$ )的概率值, 计算二者的比值  $A = p_{Max} / p_{SecondMax}$ , 再根据  $A$  的值确定  $\eta$  的值。

$$\eta = \begin{cases} 4 & A > 4 \\ A & 1 \leq A \leq 4 \\ 1 & A < 1 \end{cases} \quad (2-19)$$

(8) 根据  $|S - i| < 0.5$ ,  $i \in E$  确定该时段指数所处的预测状态为  $i$ 。待该时段的指数值确定后, 将其加入原序列, 重复步骤(2)-(8)进行下一时段的预测。

### 2.3.2 基于 GM (1,1)模型的灰色预测

灰色系统理论是1982由我国学者邓聚龙教授提出的<sup>[64]</sup>。它把一般系统论、信息论、控制论的观点和方法延伸到社会、经济、生态等抽象系统, 并结合数学方法, 发展成为一套解决信息不完备系统即灰色系统的理论和方法<sup>[65]</sup>。

灰色预测作为灰色系统理论中的一部分, 是通过对原始数据的处理和灰色模型的建立发现、掌握系统发展规律, 对系统的未来状态做出科学的定量预测。按照灰色预测的功能和特征将其分为数列预测、灾变预测、拓扑预测和系统预测四类。

本研究主要针对数列预测和灰色灾变预测进行模型实现。数列预测是对系统变量的未来行为进行预测, 常用的数列预测模型是GM (1,1)模型。而灾变预测是指通过对灾变日期的研究, 寻找其规律性, 预测以后灾变发生的日期, 它是通过对灾变日期序列建立GM (1,1)模型来实现的<sup>[66]</sup>。因此, GM (1,1)模型是数列预测和灾变预测的基础, 以下就对GM (1,1)模型的建立和检验作进一

步的介绍。

GM (1, 1)模型是最常用的一种灰色模型, G 表示 Grey(灰色), M 表示 Model(模型), 第一个 1 表示 1 阶导数, 第二个 1 表示 1 个变量, 即它是由一个只包含单变量的一阶微分方程构成的模型。该模型要求数据较少, 原理简单, 计算量适中, 但是不能用于较长时间的预测, 否则会产生较大的误差。GM (1, 1)模型的建模过程如下:

(1) 设原始数据序列为  $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$ , 对此序列依次累加, 累加公式为  $X^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i) \quad k=1, 2, \dots, n$ , 得到新序列为  $X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$ , 序列  $X^{(1)}$  称为一次累加生成数列。

(2) 利用序列  $X^{(1)}$  构成 GM (1, 1)模型相应的微分方程:

$$\frac{dx^{(1)}}{dt} + \alpha x^{(1)} = \mu \quad (2-20)$$

式中,  $\alpha$  为发展系数, 它反映了一次累加生成数列和原始数列的发展趋势。 $\mu$  为灰色作用量, 它是从背景值挖掘出来的数据, 它反映数据变化的关系。

利用最小二乘法来求解参数  $\alpha$ 、 $\mu$ , 即

$$\begin{bmatrix} \alpha \\ \mu \end{bmatrix}^T = (B^T B)^{-1} B^T Y_n \quad (2-21)$$

式中,

$$B = \begin{bmatrix} -\frac{1}{2}[X^{(1)}(1) + X^{(1)}(2)] & 1 \\ -\frac{1}{2}[X^{(1)}(2) + X^{(1)}(3)] & 1 \\ \dots & \dots \\ -\frac{1}{2}[X^{(1)}(n-1) + X^{(1)}(n)] & 1 \end{bmatrix}$$

$$Y_n = [X^{(0)}(2), X^{(0)}(3), \dots, X^{(0)}(n)]^T$$

(3) 将求解出的参数  $\alpha$ 、 $\mu$  的值带入下式即可计算出序列  $X^{(1)}$  的灰预测值  $\hat{X}^{(1)}$ 。

$$\hat{X}^{(1)}(k+1) = \left[ x^{(0)}(1) - \frac{\mu}{\alpha} \right] e^{-\alpha k} + \frac{\mu}{\alpha}, \quad k=1, 2, \dots, n \quad (2-22)$$

(4) 通过下面的累减还原公式, 即可得到灰预测值  $\hat{X}^{(1)}$  的还原值  $\hat{X}^{(0)}$ ,  $\hat{X}^{(0)}$  也就是序列  $X^{(0)}$  的预测值。

$$\hat{X}^{(0)}(k+1) = \hat{X}^{(1)}(k+1) - \hat{X}^{(1)}(k) = (1 - e^\alpha) \left( x^{(0)}(1) - \frac{\mu}{\alpha} \right) e^{-\alpha k} \quad k=1, 2, \dots, n \quad (2-23)$$

(5) 模型检验: 按照上述方法确定模型后, 需要对模型进行检验, 判定其是否合理, 是否合格。只有通过检验的模型才能用于预测。灰色预测检验一般考虑相对误差、均方差比值、小误

差概率和关联度四个方面。

设原始序列  $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$  相应的预测模型模拟序列为  $\hat{X}^{(0)} = \{\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(n)\}$ ，则残差序列为：

$$\varepsilon^{(0)} = (\varepsilon(1), \varepsilon(2), \dots, \varepsilon(n)) = (x^{(0)}(1) - \hat{x}^{(0)}(1), x^{(0)}(2) - \hat{x}^{(0)}(2), \dots, x^{(0)}(n) - \hat{x}^{(0)}(n)) \quad (2-24)$$

相对误差序列为：

$$\Delta = (\Delta_1, \Delta_2, \dots, \Delta_n) = \left( \left| \frac{\varepsilon(1)}{x^{(0)}(1)} \right|, \left| \frac{\varepsilon(2)}{x^{(0)}(2)} \right|, \dots, \left| \frac{\varepsilon(n)}{x^{(0)}(n)} \right| \right) \quad (2-25)$$

则称  $\bar{\Delta} = \frac{1}{n} \sum_{k=1}^n \Delta_k$  为平均相对误差。

设  $\bar{x} = \frac{1}{n} \sum_{k=1}^n x^{(0)}(k)$ ， $S_1^2 = \frac{1}{n} \sum_{k=1}^n (x^{(0)}(k) - \bar{x})^2$  分别为序列  $X^{(0)}$  的均值和方差，

$\bar{\varepsilon} = \frac{1}{n} \sum_{k=1}^n \varepsilon(k)$ ， $S_2^2 = \frac{1}{n} \sum_{k=1}^n (\varepsilon(k) - \bar{\varepsilon})^2$  分别为残差的均值和方差，则称  $C = \frac{S_2}{S_1}$  为均方差比值，

$p = P(|\varepsilon(k) - \bar{\varepsilon}| < 0.6745 S_1)$  为小误差概率。

上述给出的检验模型的方法都是通过对残差的考察来判断模型的精度，其中平均相对误差  $\bar{\Delta}$  要求越小越好，均方差比值  $C$  越小越好以及小误差概率  $p$  越大越好。常用的精度等级见表 2-1<sup>[16]</sup>。

表 2-1 精度检验等级参照表

精度等级	指标临界值		
	平均相对误差	均方差比值	小误差概率
一级	0.01	0.35	0.95
二级	0.05	0.50	0.80
三级	0.10	0.65	0.70
四级	0.20	0.80	0.60

关联度分析是用来分析系统中各因素关联程度的方法。系统中母序列  $X_0(k)$  与影响因子子序列  $X_i(k)$  存在关联，则它们之间的关联度的计算公式为：

$$r = \frac{1}{n} \sum_{k=1}^n n(k) \quad (2-26)$$

其中，

$$n(k) = \frac{\min_i \min_k |X_0(k) - X_i(k)| + \rho \min_i \min_k |X_0(k) - X_i(k)|}{\max_i \max_k |X_0(k) - X_i(k)| + \rho \min_i \min_k |X_0(k) - X_i(k)|} \quad (2-27)$$

式中， $\rho$  被称为分辨率， $0 < \rho < 1$ ，一般取  $\rho = 0.5$ 。 $\min_i \min_k |X_0(k) - X_i(k)|$  为两级最小差。其

中,  $\min_k |X_0(k) - X_i(k)|$  是第一级最小差, 表示在  $X_{(0)}(k)$  序列中找各点与序列  $X_i(k)$  的最小差。  
 $\min_i \min_k |X_0(k) - X_i(k)|$  为第二级最小差, 表示在各序列找出的最小差基础上找所有序列中的最小差。  
 $\max_i \max_k |X_0(k) - X_i(k)|$  是二级最大差, 其含义与最小差相似。

GM(1,1)模型中可运用关联度来分析预测值  $\hat{X}^{(0)}$  和原始值  $X^{(0)}$  之间的相关关系。关联系数越大, 说明预测值和实际值越接近。因为 GM(1,1)模型只有两个序列, 所以式 (2-27) 中取  $X_0(k) = \hat{X}^{(0)}(k)$ ,  $X_i(k) = X^{(0)}(k)$ , 并且  $i=1$ , 也就是说只有第一级最小值和最大值。

(6) 经过模型检验并且合格的模型就可以应用到预测中, 将  $k$  ( $k \geq n$ ) 带入公式 (2-23) 就可得到序列  $X^{(0)}$  第  $k+1$  个数据的预测值。

数列预测和灰色灾变预测的建模方法如上所述, 它们的区别主要是需要的数据不同。数列预测只需要一时间序列的数据即可, 而灰色灾变预测需要先对时间序列的数据给定一个异常值 ( $\xi$ ), 构建灾变日期序列, 再对灾变日期序列进行建模、预测。构建灾变日期序列的方法为: 设原始序列为  $X = (x[q(1)], x[q(2)], \dots, x[q(n)])$ , 则灾变日期序列由  $q(i)$  ( $1 \leq i \leq n$ ,  $x[q(i)] \geq \xi$  或  $x[q(i)] \leq \xi$ ) 组成。GM (1,1)灾变预测模型使用方便, 尤其对呈指数上升或下降的序列有较好的预测结果, 但对随机性强的序列若不做好相应的数据分析或数据预处理工作, 则会造成预测结果精度不高<sup>[67]</sup>。

### 2.3.3 基于 VTCI 的 ARIMA 预测

VTCI 是一种近实时的遥感干旱监测模型, 通过区域的 VTCI 就可以对干旱进行面上的预测。要对时空变化的 VTCI 序列进行预测, 就要找到每个像点 VTCI 序列随时间的变化规律, 建立合适的数学模型。但是对一个区域成千上万的像点分析其变化规律, 建立适合的预测模型的方法是行不通的。一个可行的方法是选择该区域内的若干像点的 VTCI 时间序列, 找出其共有的特征和变化规律, 建立统一的预测模型, 再对整个区域各个像点的 VTCI 时间序列进行预测。

基于这种思想, 以关中平原 22 个气象站点的 VTCI 时间序列为分析数据, 采用 ARIMA 模型对它们进行建模分析, 来寻找它们共有的变化规律。ARIMA 模型建模通常需要以下几个步骤:

#### (1) 平稳性检验

根据时间序列的时序图, 自相关函数 (ACF) 图, 或通过 ADF 单位根检验, 判断该序列的平稳性。平稳序列的时序图显示序列无明显的上升和下降趋势, 各观测值围绕其均值上下波动; ACF 图中, 如果随着延迟阶数的增加, 自相关系数很快衰减到零, 则为平稳时间序列。若为非平稳时间序列, 需进行差分处理转化为平稳序列, 再进行 ARMA 建模。差分的次数即为 ARIMA( $p, d, q$ ) 中的阶数  $d$ 。

#### (2) 模型定阶

平稳时间序列的 ARMA 模型识别可通过分析序列自相关和偏自相关函数图进行, 如果时间序列的自相关拖尾, 而偏自相关在  $p$  步截尾, 则序列可建立 AR ( $p$ ) 模型; 如果时间序列的偏自相关拖尾, 而自相关在  $q$  步截尾, 则序列可建立 MA ( $q$ ) 模型; 如果时间序列的偏自相关和自相关都拖尾, 则序列可建立 ARMA ( $p, q$ ) 模型。



### (3) 参数估计

参数的估计方法有三种：矩估计、极大似然估计和最小二乘估计。其中矩估计为粗估计，极大似然估计和最小二乘估计为精估计。

### (4) 模型检验

模型检验主要是检验模型的有效性，即检验残差序列是否为白噪声序列。如果模型通不过检验，说明所建模型对信息的提取不充分，转向步骤2，重新选择模型进行拟合。

### (5) 模型的优化

通过模型检验的模型不止一个，用Akaike的AIC(Akaike Information Criterion)准则和Schwartz的SBC(Schwarz Bayesian Criterion)准则可评判模型的相对优劣，计算不同模型的AIC和SBC，两个值达到最小的模型为相对最优模型。

按照 ARIMA 模型的建模方法，对 22 个气象站点的 VTCI 时间序列进行分析表明，这些序列是平稳的，都适合 AR(1)模型，均值为 $\mu$ 的时间序列的 AR(1)模型形式为：

$$(Y_t - \mu) - \phi_1(Y_{t-1} - \mu) = a_t \quad (2-28)$$

其中  $Y_t$  为某一像点  $t$  时刻的序列值； $\mu$  为该时间序列的均值； $a_t$  为随机扰动白噪声序列； $\phi_1$  为 AR(1)模型的唯一参数， $\phi_1$  的最小二乘参数估计为：

$$\phi_1 = \frac{\sum_{t=1}^{N-1} (Y_t - \bar{Y})(Y_{t+1} - \bar{Y})}{\sum_{t=1}^{N-1} (Y_t - \bar{Y})^2} \quad (2-29)$$

其中， $\bar{Y}$  为时间序列的均值； $N$  为时间序列的长度； $Y_t$  为  $t$  时刻的序列值。

$Y_t$  的  $\ell$  步预测模型为：

$$\hat{Y}_{t+l} - \mu = \phi_1^\ell (Y_t - \mu) \quad (2-30)$$

其中， $\ell$  为预测步长， $\hat{Y}_{t+l}$  为  $t+l$  时刻的预测值。

因此，公式(2-30)就是针对关中平原的 VTCI 时间序列建立的统一预测模型。应用该模型就可以预测整个区域的 VTCI 情况，从而实现对于干旱的预测。

## 2.4 小结

干旱预测是以干旱指数为基础来建立预测模型。本研究涉及的干旱指数主要包括标准化降水指数和条件植被温度指数等。在干旱预测模型上选择了加权马尔可夫模型、基于 GM(1,1)的灰色预测模型和基于 VTCI 的 ARIMA 预测模型。对这三种模型的理论 and 预测流程的理解是进行模型实现的基础和关键。

### 第三章 干旱预测系统

#### 3.1 系统概述

干旱预测系统以遥感干旱预测的技术路线为基础，根据基于归一化植被指数和土地表面温度的干旱监测方法所得到的干旱监测指数，采用时间序列分析的相关模型进行区域旱情的预测。

本系统主要包括三大部分：系统主界面、干旱监测数据处理模块和干旱预测模块，三者之间的相互关系如图 3-1 所示。用户通过主界面提供的功能接口访问干旱监测数据处理模块。该模块一方面针对 NOAA-AVHRR 数据的特点，从 AVHRR 卫星遥感数据生成的 NDVI 和 LST 数据产品中提取出干旱监测信息，实现 AVHRR 数据的最大、最小值合成，遥感干旱监测指数 VCI、AVI、TCI 和 VTCI 的计算等；另一方面针对降水量数据实现标准化降水指数的计算。同样，用户也可以通过主界面访问干旱预测模块，以干旱监测数据处理模块计算得到的各种干旱监测指数为数据源，采用加权马尔可夫模型、基于 GM (1,1)的灰色预测模型和基于 VTCI 的 ARIMA 预测模型进行干旱预测。本系统的用例图如图 3-2 所示。

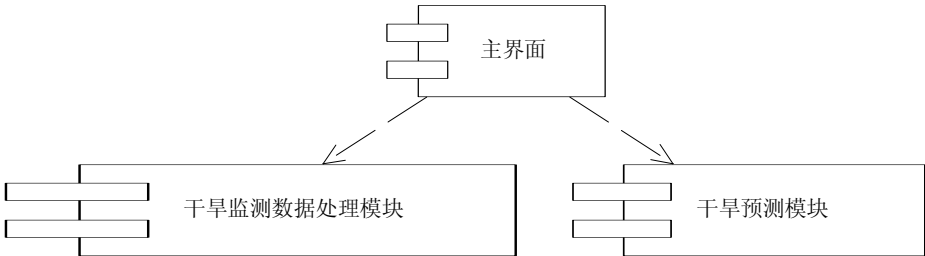


图 3-1 系统模块关系图

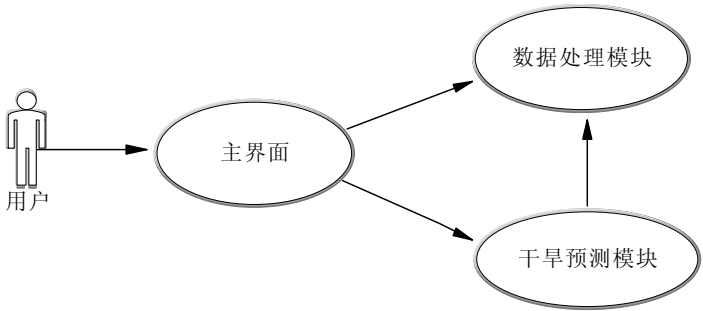


图 3-2 系统用例图

系统的功能模块如图 3-3 所示。图中虚线框内对 AVHRR 遥感数据处理的相关功能已经基本实现，本研究主要针对干旱预测模块、系统主界面以及干旱监测数据处理模块中的 SPI 计算功能进行设计和实现。对于干旱预测模块来说，一方面要根据已有的 AVHRR 遥感数据处理相关功能对数据的要求，并结合该模块中所涉及到的数据，对数据结构和访问接口进行设计；另一方面根据模块功能的要求，在相关算法设计的基础上，按照面向对象程序设计思想进行类的设计和实现。鉴于系统中部分功能已经完成，从系统开发模块化、代码重用等方面考虑，采用动态链接库（Dynamic Linkable Library, DLL）技术来进行模块的开发。将干旱监测数据处理模块和干旱预

测模块分别封装在不同的 DLL 文件中，通过主界面模块调用它们，从而将两个模块整合到同一系统中。

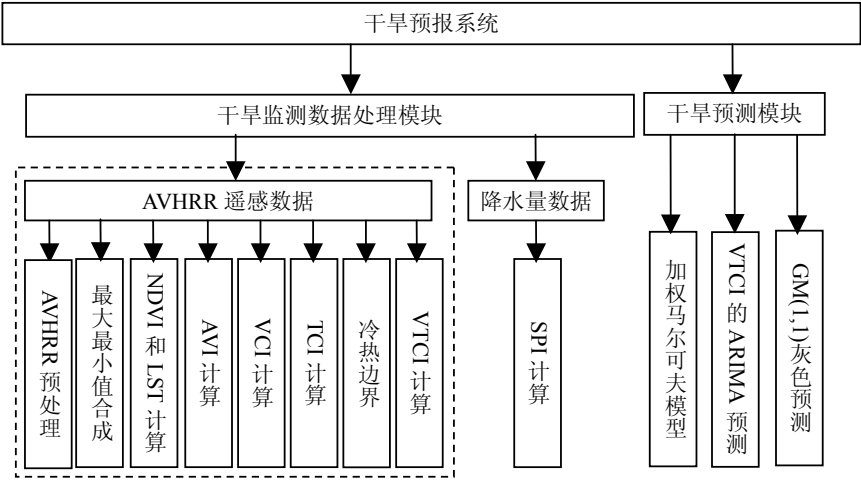


图 3-3 系统功能模块图

3.2 系统主界面的设计和实现

根据对系统总体结构的设计，系统是由主界面、数据处理模块和干旱预测模块组成。主界面是系统的框架，为各个模块提供一个运行的平台。按照这些要求，以微软基础类（Microsoft Foundation Classes，MFC）提供的文档视图结构（Document/View Architecture）为基础进行系统框架的设计和实现。

系统框架主要由 9 个类组成，其类对象结构如图 3-4 所示，图中虚线箭头表示依赖关系，实线箭头表示继承关系。其中应用程序类（CDroughtApp）、视图类（CDroughtView）、框架类（CMainFrame）和文档类（CDroughtDoc）是实现文档视图结构所必须的四个类，由 Visual C++6.0 的应用程序向导（AppWizard）向导生成<sup>[68]</sup>。此外，对话框类（CDialog）和控制条类（CControlBar）是 MFC 所提供的 Windows 应用程序开发的基础类。关于系统框架相关类的详细说明见表 3-1。

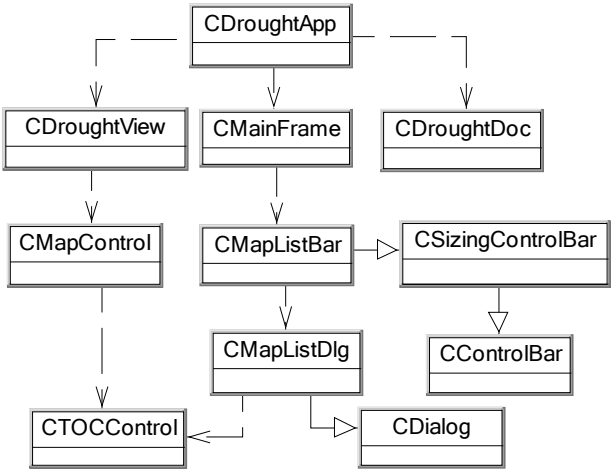


图 3-4 系统框架类对象结构图

表 3-1 系统框架相关类说明表

类名	功能说明
CDroughtApp	应用程序类，提供管理整个程序及初始化程序等功能
CDroughtView	文档/视图应用程序的基本视图，也是包含控件的视图
CMainFrame	单文档应用程序的主框架窗口
CDroughtDoc	文档类，提供保存应用程序的数据和磁盘文件操作
CMapControl	ArcGIS 提供的地图控件，实现对地图对象的相关操作
CTOCControl	ArcGIS 提供的目录树控件，与地图控件“绑定”后可以实现地图图层树状视图显示
CSizingControlBar	继承于控制条类（CControlBar），实现可在框架内浮动、改变大小的的控制窗口
CMapListBar	继承于 CSizingControlBar，实现停靠在界面左侧的浮动窗口
CMapListDlg	继承于对话框类（CDialog），作为目录树控件的“容器”“嵌入”到界面左侧浮动窗口内

系统主界面遵循简洁明快、美观和使用方便的原则，按照上述对系统框架的设计，采用 C++ 语言进行程序编写，从而实现系统的主界面如图 3-5 所示。它主要分为三个部分：第一是界面顶部的菜单工作区和工具栏。通过菜单可以实现数据处理和预测等相关功能，而通过工具栏可以方便地对地图进行一些常用的操作。第二是界面左侧的图层管理窗口。通过它可以显示地图、图层或符号的状况，对图层进行相关控制。第三是界面中部的地图显示窗口。该部分主要用于显示矢量和栅格图形，是数据处理模块和预测模块处理结果的一个可视化窗口，便于用户更直观地了解相关信息。

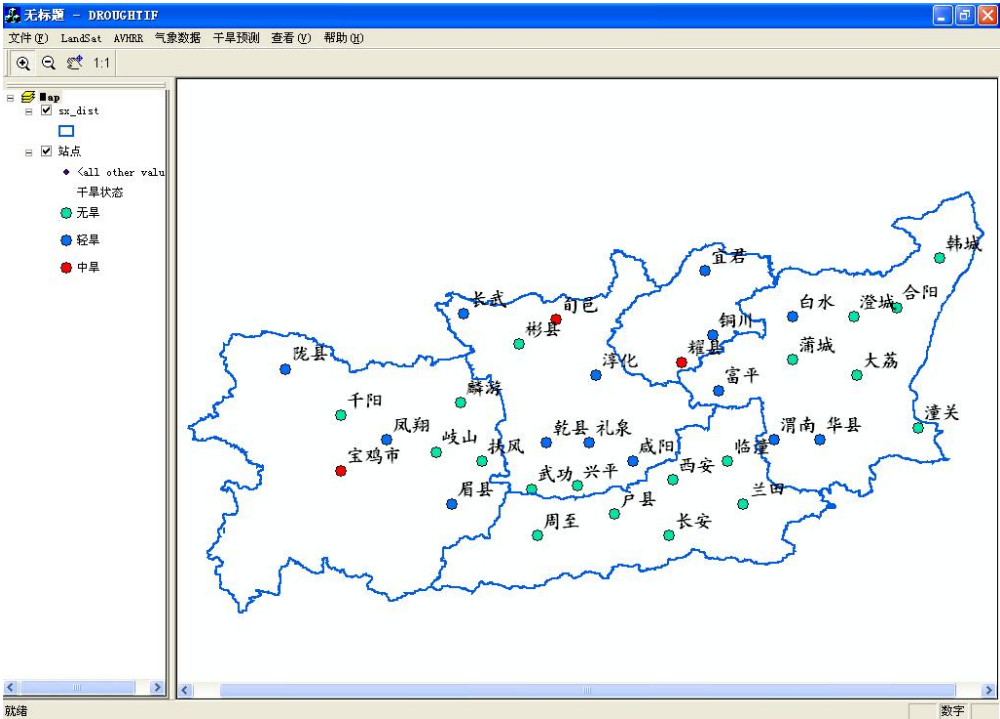


图 3-5 系统主界面

### 3.3 小结

本章从遥感干旱预测系统整体构架、模块划分和功能设计等方面对系统进行了概述，明确了本研究所要实现的模块和功能包括系统主界面、干旱预测模块和 SPI 计算功能。并根据系统总体结构的设计和主界面在应用上的需求，以微软基础类提供的文档视图结构为基础，对主界面进行了设计和实现，从而为系统搭建起了一个框架。

## 第四章 干旱预测模块的设计与实现

本研究的重点是干旱预测模块的设计和实现。将该模块采用的加权马尔可夫模型、基于 VTCI 的 ARIMA 预测和 GM(1, 1) 预测模型分别独立为一个功能，其用例图如图 4-1。此外，在干旱监测数据处理模块中针对降水量数据添加了标准化降水指数计算功能。在整个模块的设计和实现过程中，首先根据已有的 AVHRR 遥感数据处理相关功能对数据的要求，并结合干旱预测模块中所涉及到的数据，对相关数据的数据结构和访问方式进行了设计和实现。其次，在第二章所介绍的三种干旱预测模型理论的基础上，设计了相关算法的流程。最后，按照面向对象程序设计思想，根据需要将模块抽象为不同的类，这些类各负其责共同实现干旱预测模块。

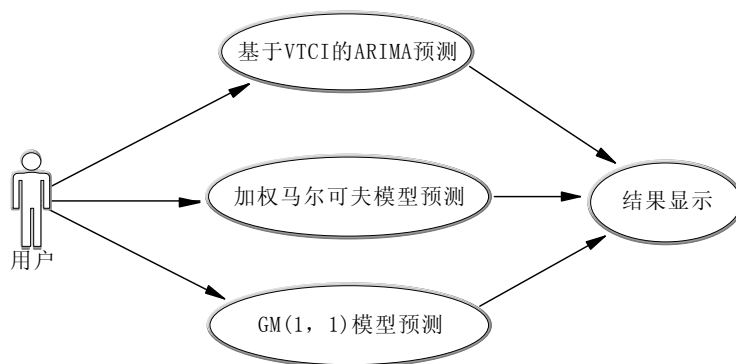


图 4-1 干旱预测模块用例图

### 4.1 数据接口的设计与实现

干旱预测以一定的数据为基础，在预测的过程中，涉及到数据格式和内容不同的输入数据和输出数据。对这些数据存储方式和数据结构的合理设计，是干旱预测模块，甚至整个系统设计的基础。对数据格式的设计既要满足相关模型计算以及用户应用的需求，又要尽可能使其格式简单，便于数据的访问，减少数据的冗余。而数据接口的实现又是以合理的数据格式设计为基础，它将对这些数据的相关操作封装在独立的接口中。通过对这些接口的调用，就可以既方便又高效地对数据进行访问，从而实现了数据访问代码的重用。

#### 4.1.1 数据格式的设计

数据结构和文件格式是整个系统的基础，对它们的设计关系到整个系统在开发和使用过程中数据访问的方式和效率。本系统涉及到的数据主要包括：RAW 格式的遥感影像数据、气象站点的降水量数据、标准化降水指数数据、模型计算所需参数数据以及预测结果数据等。

遥感干旱监测数据处理模块支持的是 RAW 格式的遥感影像数据，而本系统的开发是在它的基础上进行的，因此遥感影像的数据格式仍然选择 RAW 格式。RAW 格式是不经过压缩处理而直接存储的影像文件格式，没有影像相关信息的头文件。与其它数据格式相比，RAW 格式比较简

单，对它的读写操作比较容易，这也是本系统选择 RAW 格式的原因。但是由于 RAW 格式的影像文件没有存储与影像相关的头文件信息，如行列数，数据类型等，用户在使用过程中需要手工输入相关信息才能读取或存储影像文件，给用户的使用带来不方便。针对这一问题，为 RAW 格式的影像文件设计了相匹配的头文件来存储数据处理过程中需要的相关信息。用户在第一次使用一幅 RAW 格式的影像文件时，将该文件的相关信息（影像行数、影像列数、波段数和数据类型）输入后，系统自动在该影像文件所在目录下生成与它同名的头文件，并将这些信息存储在该头文件中。用户在以后使用该影像文件时，系统会自动去匹配该文件的头文件，并读取相关信息，而不需要用户再手工输入，这样就一定程度方便了用户的使用。RAW 格式影像文件的头文件描述具体见表 4-1。

表 4-1 RAW 格式影像文件头文件字段描述表

数据项	数据类型	说明
Lines	Int	影像行数
Samples	Int	影像列数
Band	Int	影像波段数
Data Type	Int	影像数据类型（1 表示 Float，2 表示 Unsigned Char）

除了遥感影像文件外，在干旱预测中，计算标准化降水指数时还要涉及到对气象站点降水量数据的处理。对每个气象站一般涉及到三十年以上的月降水量数据，而且在进行计算时需要同时处理多个气象站，因此有必要对这些数据的格式和组织方式进行设计。为了便于管理和计算，以单个气象站点为一个文件来存储和组织众多的降水量数据。每一个文件都分为文件头和数据存储块两个部分。文件头（Head）主要用于存储气象站点的相关信息，如站点编号，名称，经度和纬度，而降水量数据则全部存储在数据存储块中。对于各年逐月的降水量数据又以数据记录（Record）为单位进行组织，每一条数据记录包含三个独立的字段：年份，月份和降水量。其具体数据格式结构如图 4-2 所示，每个数据项的详细描述见表 4-2。对于计算出的标准化降水指数也需要文件进行存储，对它的文件结构设计与降水量文件结构的设计基本相同，区别在于将数据记录中的降水量字段改为 SPI 字段，用于存储计算出的 SPI 值。

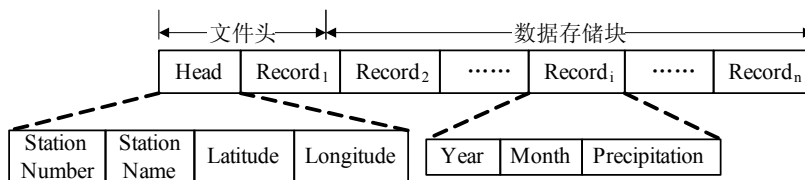


图 4-2 降水量文件数据格式结构图

表 4-2 降水量文件字段描述表

数据项	数据类型	说明
Station Number	Char	气象站点编号
Station Name	Char	气象站点名称
Latitude	Float	气象站点纬度
Longitude	Float	气象站点经度
Year	Int	年份
Month	Int	月份，取值范围 1 至 12
Precipitation	Float	降水量，缺失数据记为 Missing

对于加权马尔可夫模型中状态空间的确定，有一种情况是根据已有的干旱等级划分标准来设置状态空间。如果将这些参数写在程序中，就不便于参数的修改，降低了该预测功能的灵活性。为了使该功能适用于更广的范围和使用户更方便地设置参数，将这些参数以文件的形式存储。系统通过读取参数文件来完成模型参数的设置，用户也可以通过修改或新建这些参数文件来满足不同的需求。通过对涉及到的干旱等级划分标准的分析，发现这些划分标准主要由干旱等级、属于该等级的指数上下限值和干旱类型组成，对该文件数据格式的详细描述见表 4-3。

表 4-3 SPI 干旱等级划分文件字段描述表

数据项	数据类型	说明
Scale	Int	等级
Drought Category	Char	干旱类型
Upper Limit	Float	上限值
Lower Limit	Float	下限值

使用加权马尔可夫模型和基于 GM (1,1)的灰色预测模型来对单一时间序列进行预测时，它们的输入数据包含的字段如表 4-4 所示，每一条数据中对数据的附加说明是由用户根据具体需要来设置的，它不参与模型的计算。对于灰色预测模型预测结果文件包括预测值、相对误差、均方差比值、小误差概率和关联度这五项。第一项根据用户选择的预测步长的不同，其数目也不同。后四项是对模型检验的标准，用户可以通过它们的值来了解模型的优劣。加权马尔可夫模型只需要保存其预测结果，因此它的预测结果文件也只包含预测结果一个字段。

表 4-4 输入数据文件字段描述表

数据项	数据类型	说明
Description	Char	对数据的附加说明
Data	Int 或 Float	具体的时间序列数据

4.1.2 数据接口的实现

在数据格式设计的基础上，需要实现相应的数据接口才能方便地访问各种数据。通过上节对数据文件格式的分析，采用面向对象程序设计的方法，将数据文件抽象为不同的类。因为降水量



文件和标准化降水指数文件都是由文件头和数据存储块两部分组成，而数据存储块又包含若干条数据记录，因此将它抽象为两个类：文件头类(CHead)和数据记录类(CRecord)。文件头类负责对文件头进行相关操作。数据记录类对应存储在文件中的一条数据记录，通过该类提供的方法 Set\*\*\*()函数和 Get\*\*\*()函数来设置和获取相关字段的值，从而实现对相关数据的读取和存储。

仅有关于数据格式的相关类还不能完全满足数据文件操作的要求，还需要对相应数据文件的类进行设计和实现。根据上节对涉及到的几种数据文件的分析，将它们抽象为四个独立的类，包括：RAW 格式影像头文件类(CRawHeadFile)、模型参数文件类(CParameterFile)、降水量数据文件类(CPrecipitationFile)和 SPI 文件类(CSPIFile)。通过对数据文件操作需求的分析，发现它们都具有一些相同的操作，如打开文件(Open)、关闭文件(Close)、读文件(Read)和写文件(Write)等，这些操作都可以通过继承微软基础类(Microsoft Foundation Classes, MFC)中的 CFile 类来实现。而对于各个数据文件特殊的操作需求，就需要单独在相应的类中定义并实现，各个函数的具体说明见表 4-5。

表 4-5 函数说明表

类名	函数名	说明
CRawHeadFile	FindExistHead()	搜索影像文件的头文件是否存在
	GetHeadInfo()	获取影像头文件中的相关信息
	SetHeadInfo()	设置影像头文件中的相关信息
CParameterFile	GetParameter()	获取模型所需要相关参数
	ModifyParameter()	修改模型的相关参数
CPrecipitationFile 或 CSPIFile	GetRecordSet()	获取文件中所有降水量的记录
	GetRecord()	获取某一条降水量的记录
	GetHeadInfo()	获取降水量文件中站点相关信息

按照上述设计数据访问的接口基本实现，各个类的具体定义和类之间的关系如图 4-3 所示。各个类之间分工明确，既彼此独立又相互依赖，共同完成系统数据访问的任务。对于 RAW 格式影像文件的访问，首先需要获取该影像的行列数，数据类型等信息，才能进行相关处理，其具体过程如图 4-4 所示。标准化降水指数的计算是通过读取用户输入的降水量数据文件，获得降水量数据，经过对数据的处理，最后计算出相应的 SPI 并存储在 SPI 文件中，其具体过程如图 4-5 所示。

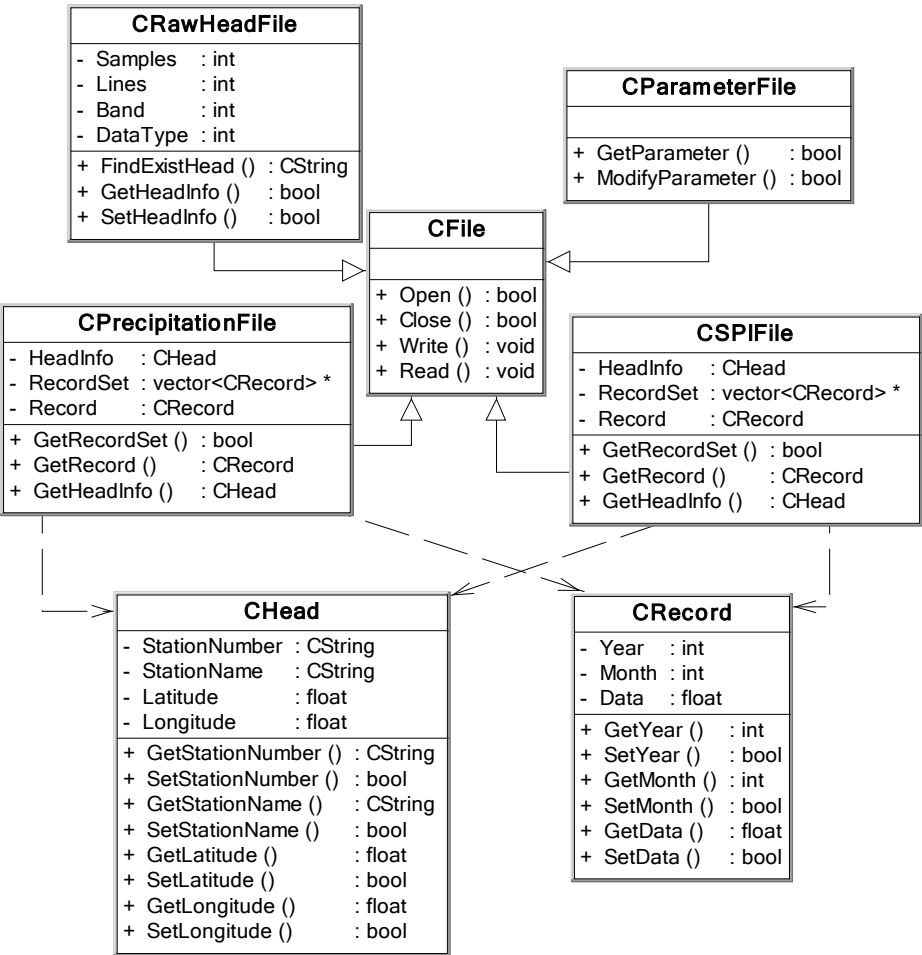


图 4-3 数据接口各类间的关系图

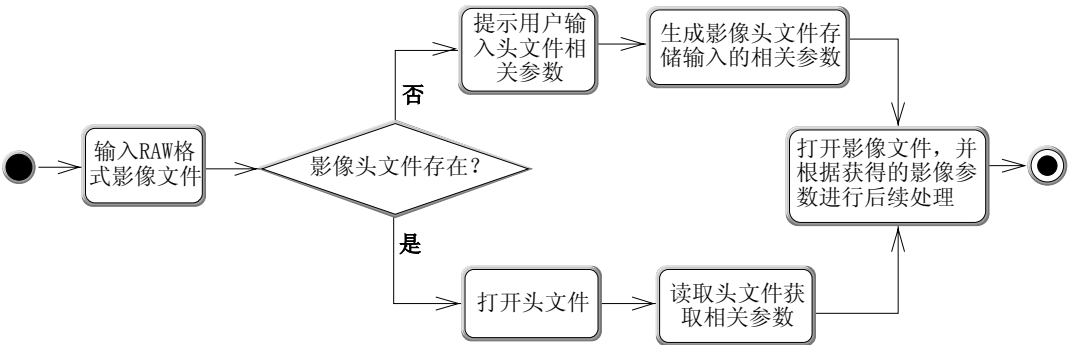


图 4-4 访问 RAW 格式影像文件过程

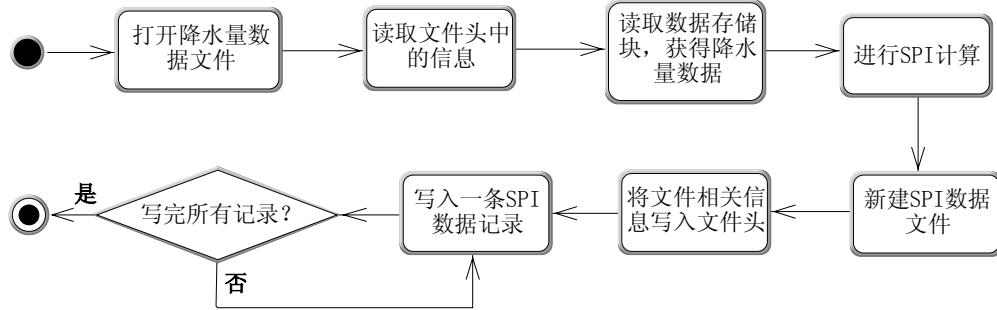


图 4-5 访问降水量和 SPI 文件过程

## 4.2 干旱预测模块算法的设计

第二章中对所采用的干旱预测模型作了详细的描述, 根据这些描述来实现对相关模型算法的设计。对于加权马尔可夫模型预测, 用户输入的数据可能是单一时间序列数据, 也可能是时间序列的若干影像数据。在进行模型计算前, 需要对一些参数进行设置, 例如该时间序列的周期, 确定状态空间需要的等级划分参数文件或者分段数目等。根据输入数据的不同, 在文件读取和相关参数设置上也有所区别, 例如如果是影像数据还需要设置影像行列数, 处理范围等参数。其具体的程序流程图如图 4-6。

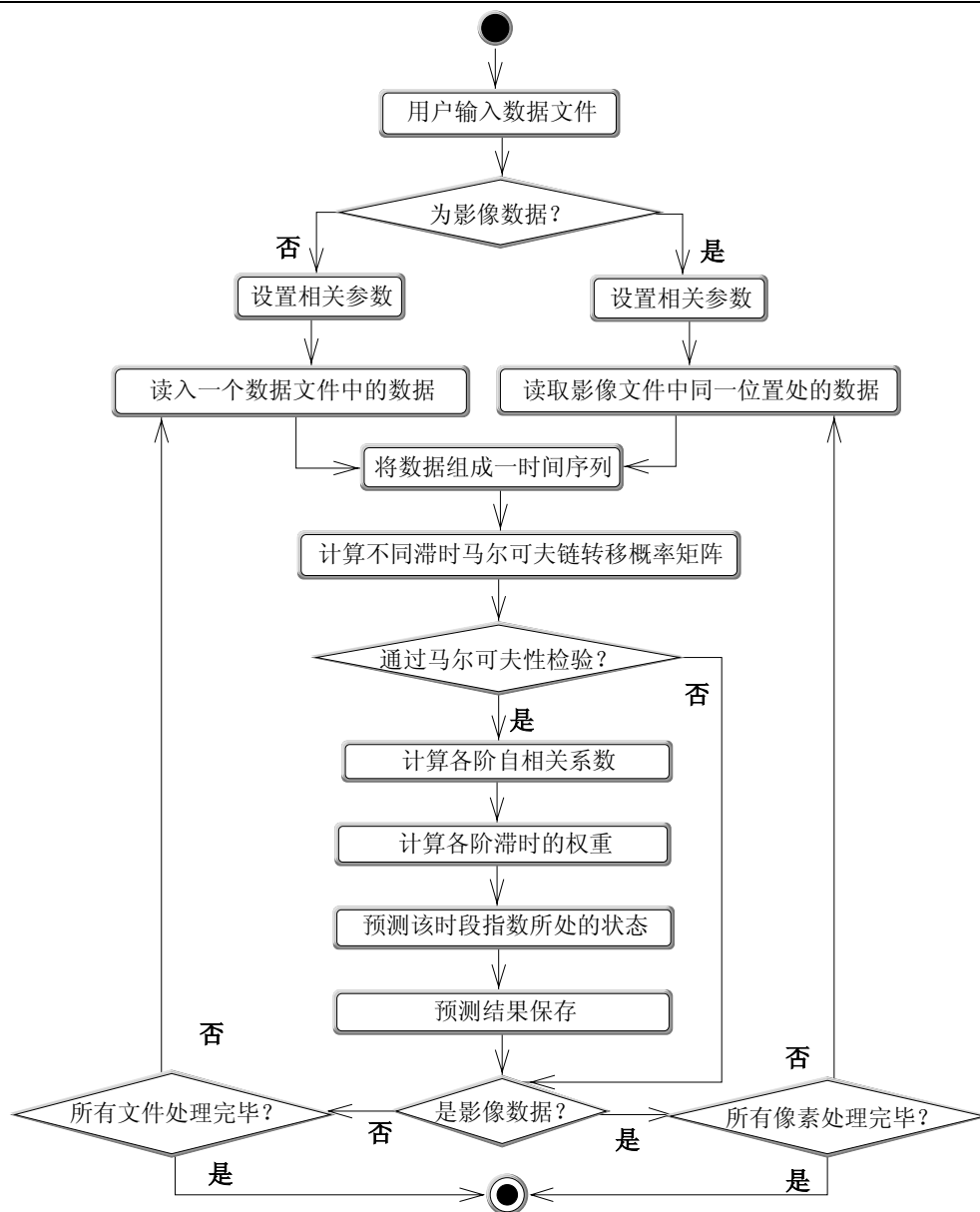


图 4-6 加权马尔可夫模型预测活动图

根据第二章对灰色预测模型的介绍，灰色预测分为建模和模型检验两个大步骤。建模采用 GM(1,1)模型，其中发展系数 $\alpha$ 和灰色作用量 $\mu$ 的计算是建模的关键部分。在程序设计上主要通过矩阵运算来实现参数的计算。灰色预测检验主要有相对误差、均方差比值、小误差概率和关联度四种方法，前三者方法是通过对残差的考察来判断模型的精度，因此程序设计上计算出数据的残差是进行模型检验的基础。灰色预测方法详细程序流程图如图 4-7。

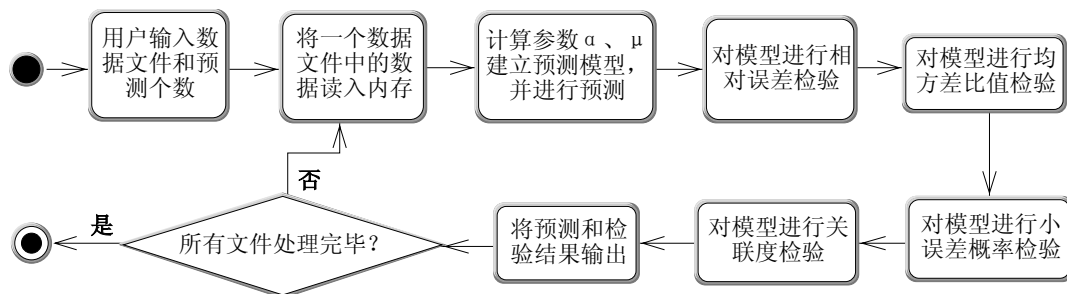


图 4-7 灰色预测活动图

基于 VTCI 的 ARIMA 预测是以 VTCI 时间序列影像为数据，在计算时需要逐像素地进行处理。对每一个像素而言，首先要根据所有影像该像素的 VTCI 值，采用最小二乘参数估计得到参数  $\varphi_1$  的值，然后根据预测步长，按照预测公式计算出该像素的预测值，最后将其保存。其详细程序流程图如图 4-8。

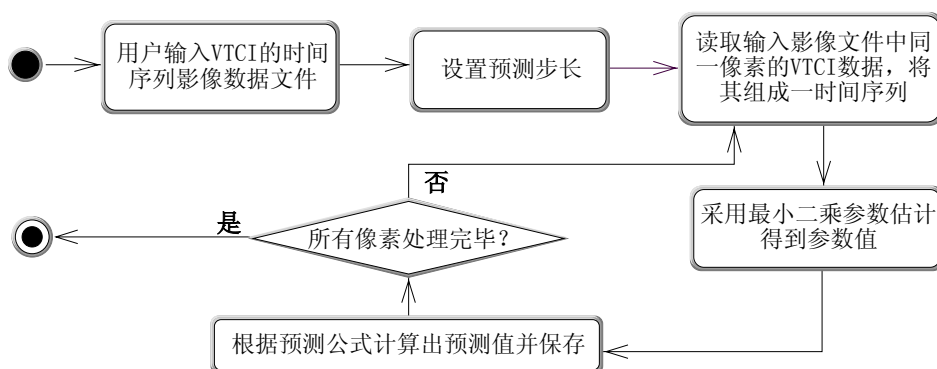


图 4-8 基于 VTCI 的 ARIMA 预测活动图

以上是对干旱预测模块所涉及到的模型的算法设计，在本研究中对数据处理模块添加了标准化降水指数计算功能。其程序流程图如图 4-9。

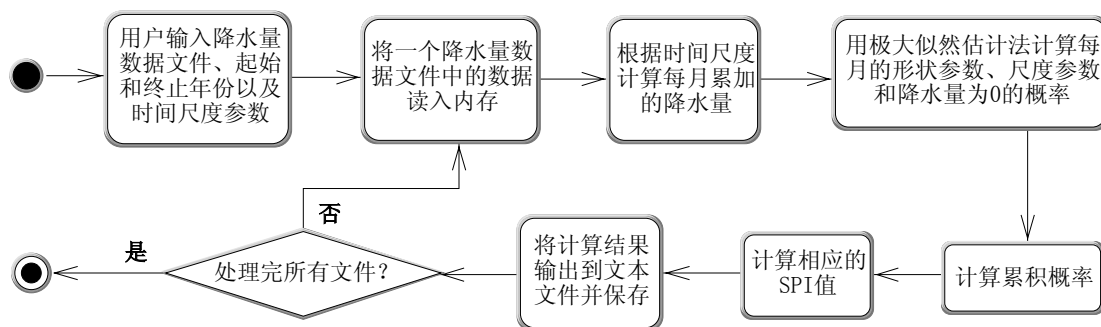


图 4-9 标准化降水指数计算活动图

### 4.3 干旱预测模块类的设计

干旱预测模块中主要包括三个功能：加权马尔可夫模型预测、基于 GM(1,1)模型的灰色预测和基于 VTCI 的 ARIMA 预测。此外，对数据处理模块添加了标准化降水指数计算功能。对这些功能而言，它们大致都可分为“前台”和“后台”两个部分。所谓“前台”就是用户所使用的功能界面

部分,该部分主要负责与用户的交互,主要包括相关数据及参数的输入和处理结果的输出。而“后台”部分是整个功能的核心,是相关模型的具体实现。该部分程序设计和编写的好坏直接关系到模型运行的效率。

按照“前台”和“后台”划分的思想,将每个功能都抽象为界面类和控制类两部分。它们各负其责,互不影响,界面类负责获取数据,并将数据传给控制类进行处理,处理结束后控制类再将处理结果返回给界面类加以显示或输出。这样进行类的抽象符合面向对象程序设计的思想,不仅易于理解,而且便于程序的实现、代码的维护和模型的扩展。因此,按照该思想将干旱预测模块的三个功能抽象为以下六个类:加权 Markov 预测界面类(CMarkovDlg),加权 Markov 预测控制类(CMarkovModle);灰色预测界面类(CGreyForecastDlg),灰色预测控制类(CGreyForecastModle);基于 VTCI 的 ARIMA 预测界面类(CVtciARIMADlg),基于 VTCI 的 ARIMA 预测控制类(CVtciARIMA)。此外,对 SPI 计算也同样抽象为 SPI 计算界面类(CSPIDlg),SPI 计算控制类(CSPICalculation)。各个类对象之间的结构关系见图 4-10,其中类 CDroughtView 是整个系统的视图类,主要与系统界面的实现相关;类 CGridCtrl 是表格类,主要负责将处理结果以表格的形式输出;类 CImageCtrl 是图像类,当处理的数据是遥感影像时,对影像的显示主要就是由该类来实现;虚线箭头表示类对象之间存在依赖关系。

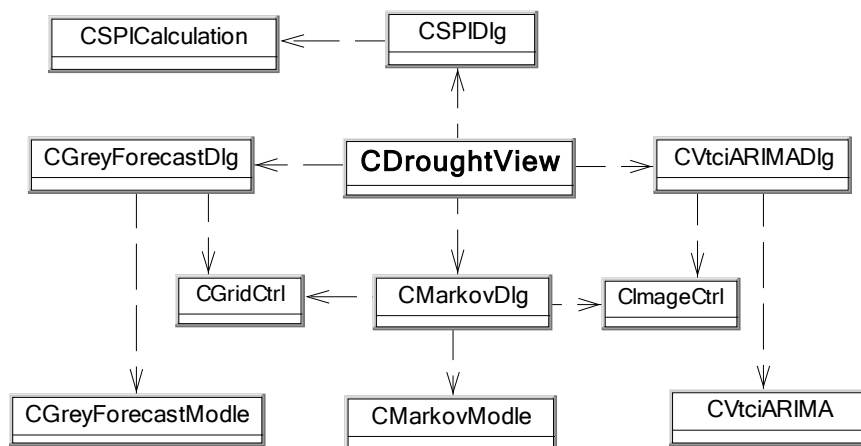


图 4-10 干旱预测模块类对象结构图

界面类主要负责用户输入数据、初始化相关控件和处理结果的显示,其相关函数也是围绕这三方面的需要来设计的,这里就不作详细介绍。而控制类作为整个功能的关键部分,其函数的设计不仅要满足实现功能的要求,而且应该尽可能使函数的划分合理,避免代码的冗余。对于干旱预测模型中的三个控制类,按照上节相关算法的设计,将其划分为一系列的函数,每个函数实现算法中的一部分功能,函数之间相对独立,互不影响。三个控制类主要函数说明见表 4-6、表 4-7 和表 4-8。此外, SPI 计算控制类的主要函数说明见表 4-9。

表 4-6 加权 Markov 预测控制类部分函数说明表

名称	说明
ComputerProbabilityMatrix()	计算转移概率矩阵
MarkovCheck()	进行马尔可夫性检验
ComputerCorrelationCoefficient()	计算各阶自相关系数
ComputerWeight()	计算各种滞时的马尔可夫链的权重
Forecast()	进行预测

表 4-7 灰色预测控制类部分函数说明表

名称	说明
ModleComputer()	计算相关参数建立模型
RelativeErrorTest()	实现相对误差检验
MeanVarianceRatioTest()	实现均方差比值检验
Errorprobability Test()	实现小误差概率检验
RelationshipTest()	实现关联度检验
Forecast()	进行预测

表 4-8 基于 VTCI 的 ARIMA 预测控制类部分函数说明表

名称	说明
ComputerAverage()	计算某一像素的序列值的平均值
ComputerParameter()	计算某一像素的模型参数
Forecast()	进行预测

表 4-9 SPI 计算控制类部分函数说明表

名称	说明
GetPrecipitation()	获取用户输入的降水量数据
ComputerParameter()	计算形状参数和尺度参数
ComputerProbability()	计算累加概率
ComputerSPI()	根据累加概率计算 SPI 值

对于加权马尔可夫模型预测和基于 VTCI 的 ARIMA 预测而言，它们处理的数据和预测结果也可以是遥感影像数据，因此需要实现对影像数据的显示。图 4-11 中类 CImageCtrl 正是实现这一功能。此外，考虑到用户的需求，还需要支持矢量数据的添加并与影像数据进行叠加。为了实现这两个功能，利用 ArcGIS Engine 提供的工具来进行开发。ArcGIS 支持多种矢量和栅格数据类型，矢量数据如：ESRI Shapfile，ArcInfo Coverage 等，栅格数据如：BMP，JPEG 等。针对本系统需要显示矢量和栅格数据，使用 ArcGIS Engine 提供的相关接口可以很容易地实现，使用的接口主要有 IMap 接口和 IMapControl 接口。IMap 接口主要控制地图数据和相关元素，通过它提供

的 AddLayer 方法就可以实现图层的添加,包括栅格数据图层。IMapControl 接口提供了访问地图控件 (MapControl) 的一系列方法,通过 AddLayerFromFile 和 AddShapeFile 方法就可以很方便地实现 Layerfile 文件和 Shapfile 文件的添加和显示。

但是 ArcGIS 不支持 RAW 格式的图像文件,而本系统使用的图像文件格式基本都是 RAW 格式的,因此为了能显示 RAW 格式的图像,需要将其转换为 ArcGIS 所支持的图像格式。BMP 作为被广泛使用的图像格式,不仅是 ArcGIS 所支持的栅格数据格式,而且它与 RAW 格式相比,主要区别在于 BMP 格式存储了关于位图相关信息描述,图像数据部分区别不大,因此将 RAW 格式转换为 BMP 格式相对更容易,只需要按照 BMP 格式的要求为 RAW 格式添加相关位图描述信息即可。

位图文件 (Bitmap File, BMP) 由四个部分组成。下面对各个部分结构的定义和 RAW 格式转换为 BMP 格式相关参数的设置做简单介绍:

#### (1) 位图文件头:

位图文件头包含文件类型、文件大小、存放位置等信息。在 Windows 3.0 以上版本的位图文件中用 BITMAPFILEHEADER 结构来定义,该结构中各参数的含义和将 RAW 格式转换为 BMP 格式各参数的取值见表 4-10。

表 4-10 位图文件头参数含义及取值说明表

名称	说明	取值
bftype	文件类型	“BM”
bfsz	文件大小	按公式 3-1 和 3-2 计算
bfReserved1	保留值	0
bfReserved2	保留值	0
bfOffBits	从位图文件头结构开始到实际的图像数据之间的字节偏移量	1078

文件大小的计算公式为:

$$\text{FileSize} = \text{ImageWidth}' \times \text{ImageHeight} + 1078 \quad (4-1)$$

$$\text{ImageWidth}' = \text{Int}((\text{ImageWidth} + 3) \div 4) \times 4 \quad (4-2)$$

其中, FileSize 为文件的大小,单位是字节; ImageWidth 和 ImageHeight 分别为 RAW 格式原始图像的宽度和高度,以像素为单位;由于 BMP 格式规定每一行的字节数必须是 4 的整倍数,如果不是则需要补齐,因此根据公式 4-2 将 ImageWidth 换算为 4 的整数倍;1078 是将 RAW 格式转换为 BMP 格式后其文件头的大小,其中位图文件头占 14 字节,位图信息头占 40 字节,调色板占  $256 \times 4$  字节。

#### (2) 位图信息头:

位图信息头包含位图的大小、压缩类型和颜色格式等,用 BITMAPINFOHEADER 结构来定义,该结构中各参数的含义和将 RAW 格式转换为 BMP 格式各参数的取值见表 4-11。



表 4-11 位图信息头参数含义及取值说明表

名称	说明	取值
biSize	位图信息头结构所需字节数	40 字节
biWidth	图像宽度	RAW 格式原始图像的宽度
biHeight	图像高度	RAW 格式原始图像的高度
biPlanes	为目标设备说明位面数 指定表示颜色时要用到的位数，常用	通常其值都设为 1
biBitCount	的值为 1（黑白二色图），4（16 色图）， 8（256 色图），24（真彩色图） 压缩说明，常用的值为 BI_RGB：无压	针对本系统使用的图像设为 8
biCompression	缩； BI_RLE8：8 位 RLE 压缩； BI_RLE4：4 位 RLE 压缩	针对本系统使用的图像设为 BI_RGB
biSizeImage	指定实际的位图数据占用的字节数	文件大小与其文件头大小之差
biXPelsPerMeter	水平分辨率	0
biYPelsPerMeter	垂直分辨率	0
biClrUsed	指定位图实际用到的颜色数，如果该 值为零，则用到的颜色数为 2 的 biBitCount 次方	0
biImportant	说明对图像显示有重要影响的颜色索引的数目，如果是 0，表示都重要	0

（3）调色板

调色板实际上是一个数组，共有 biClrUsed 个元素（如果该值为零，则有 2 的 biBitCount 次方个元素）。数组中每个元素的类型是一个 RGBQUAD 结构，占 4 个字节，该结构各参数的含义和将 RAW 格式转换为 BMP 格式各参数的取值见表 4-12。

表 4-12 RGBQUAD 结构参数含义及取值说明表

名称	说明	取值
rgbBlue	指定蓝色强度	0-255
rgbGreen	指定绿色强度	0-255
rgbRed	指定红色强度	0-255
rgbReserved	保留值	0

针对本系统使用的图像，由于 biClrUsed 取值为 0，因此调色板数组共有 256 个元素，也就是说有 256 个 RGBQUAD 结构。为调色板赋值的方法如图 4-11。

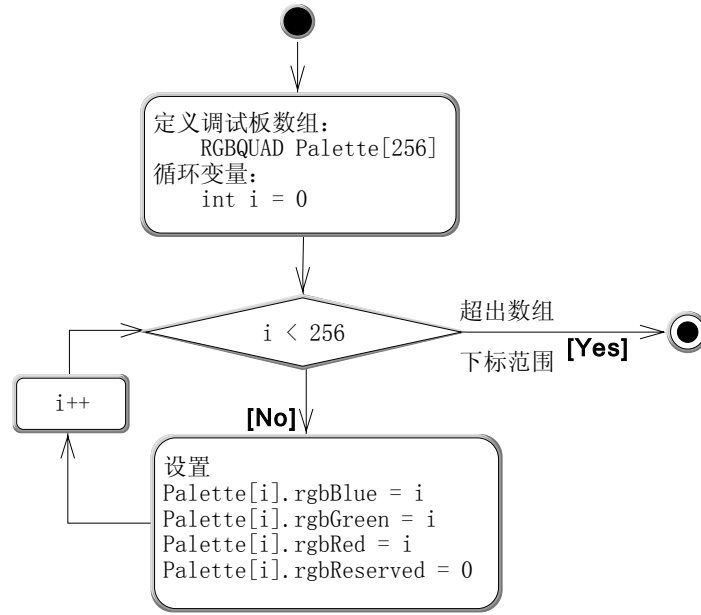


图 4-11 调色板赋值活动图

#### (4) 位图数据

紧跟在调色板之后的是图像数据字节阵列。图像的每一扫描行由表示图像像素的连续字节组成，每一行的字节数取决于图像的颜色数目和用像素表示的图像宽度。对于用到调色板的位图，图像数据就是该像素颜色在调色板中的索引值。对于真彩色图，图像数据就是实际的 R、G、B 值。值得注意的是 BMP 文件是从下到上，从左到右存储的，也就是说，阵列中的第一个字节表示位图左下角的像素，而最后一个字节表示位图右上角的像素。而 RAW 格式图像的存储方式是从上到下，从左到右。因此，在将 RAW 格式图像转换为 BMP 格式时，要按照 BMP 存储方式的要求写入数据。此外，对于没有进行线性拉伸的 RAW 格式图像，在写入数据的同时还要进行线性拉伸处理，将数据值转换到 0 至 255 之间，计算方法如下：

$$\text{Value}' = \left( \frac{\text{Value} - \text{Value}_{\max}}{\text{Value}_{\max} - \text{Value}_{\min}} \right) \times 255 \quad (4-3)$$

式中，Value 为图像某像素的值；Value' 为拉伸到 0 至 255 该像素的值；Value<sub>max</sub> 为该图像所有像素中的最大值；Value<sub>min</sub> 为该图像所有像素中的最小值。

根据上述对 BMP 文件格式和相关参数设置的分析，按照图 4-12 提供的流程就可以将 RAW 格式图像转换为 BMP 格式，从而实现 RAW 格式图像在系统中的显示，如图 4-13 所示。

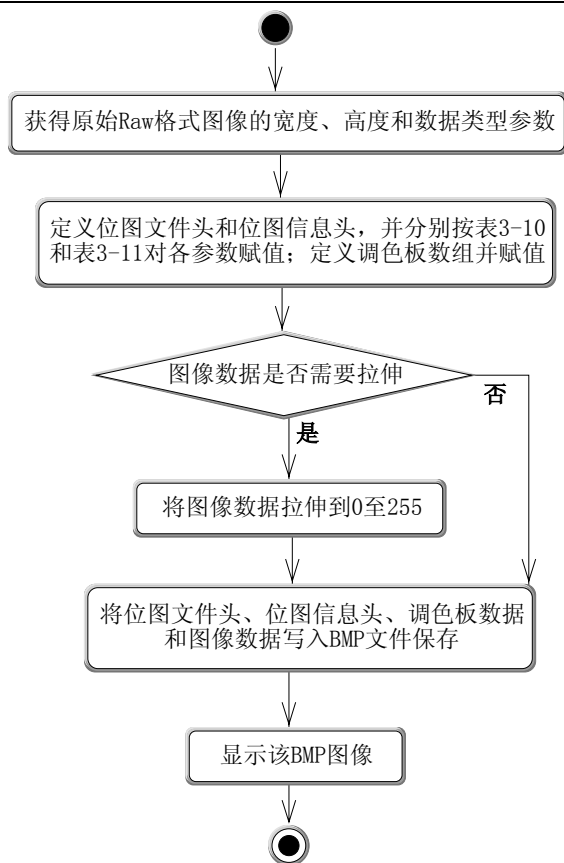


图 4-12 Raw 格式图像转换为 BMP 格式活动图

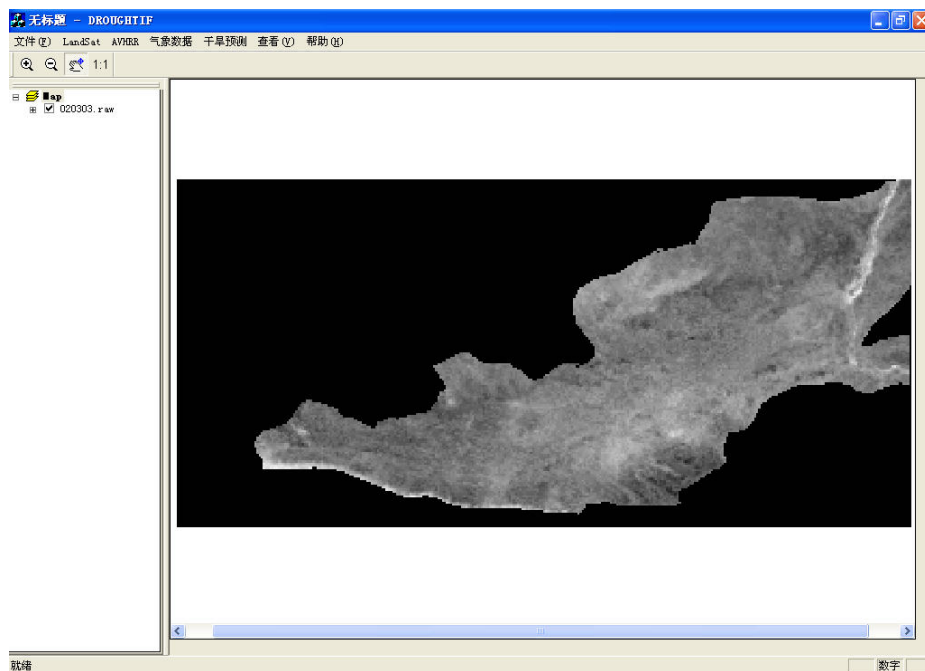


图 4-13 RAW 格式图像在系统中的显示

## 4.4 小结

本章针对干旱预测模块的设计和实现作了详细的阐述。首先通过对于干旱预测在应用过程中数据要求的分析，对数据存储方式和数据结构进行了设计，并以此为基础，采用面向对象程序设计的思想设计和实现了数据访问的接口，从而实现对数据方便、高效地访问。然后，设计出三种干旱预测模型相关算法的流程，这是进行模型程序实现的基础。最后，结合干旱预测模块功能的要求，将每个功能都抽象为界面类和控制类两部分，并以 Visual C++6.0 为开发平台，实现了该模块中的相关功能。

## 第五章 干旱预测模块的应用

系统应用是系统开发的主要目的。本研究利用陕西关中平原和渭北旱塬的降水量和关中平原的 AVHRR 多时段卫星遥感数据,通过数据处理模块计算出 SPI 和 VTCI 指数,采用干旱预测模块中的加权马尔可夫模型和基于 VTCI 的 ARIMA 模型进行了实验性应用,并检验系统运行效果。

关中平原位于陕西省中部的渭河流域,北起北山,南止秦岭,西起宝鸡,东达潼关,年平均降水量540~750 mm,是陕西境内秦岭以北最暖湿的部分。李俊霖<sup>[69]</sup>等在对关中平原气候干旱化程度进行比较分析时指出该地区属于温带大陆性季风气候区,降水量在时、空尺度上,尤其是时间(年际、季际)上存在着明显的波动性,是旱涝灾害的多发区。而渭北旱塬也是一个生态脆弱区,近40多年来年降水量明显减少,气候具有明显的暖干化趋势<sup>[70]</sup>。因此,选用关中平原和渭北旱塬作为实验区域进行干旱预测具有可行性和代表性。

### 5.1 数据准备和处理

本次应用采用的是关中平原和渭北旱塬36个气象站39年(1966-2004年)的逐月降水量数据和关中平原1999-2003年3-6月的AVHRR多时段卫星遥感数据,以标准化降水指数(SPI)和条件植被温度指数(VTCI)为干旱监测指数进行干旱预测应用。

通过对降水量数据的处理,获得标准化降水指数。首先,按照SPI计算对降水量数据格式的要求,组织36个气象站39年的逐月降水量数据。每个气象站对应一个数据文件,每一个数据文件由一个文件头和若干降水量记录组成。文件头记录站点名称等信息。降水量记录按时间先后排列,每一条记录包含年份,月份和降水量,缺失降水量的月份其值设置为Missing。按此方法组织好所有降水量数据后,利用数据处理模块的SPI计算功能,计算出36个气象站时间尺度为3个月的SPI数据。SPI计算界面如图5-1。



图 5-1 SPI 计算界面

通过对 AVHRR 多时段卫星遥感数据的处理,获得条件植被温度指数,其处理过程大致分为两个步骤。第一个步骤是先通过 AVHRR 数据预处理功能对 1999-2003 年 3-6 月单日的 1a5 格式的 HRPT(High Resolution Picture Transmission)数据进行粗几何纠正。然后利用 ERDAS 软件进行精几何纠正,提取研究区域(关中平原),并分波段存储。最后利用 NDVI 和 LST 计算功能计算出单日的归一化植被指数(NDVI)和土地表面温度(LST)。其具体的处理过程如图 5-2。第二个步骤是在计算出单日 NDVI 和 LST 基础上,首先利用最大值合成功能,以各年 3-6 月每旬内的

多日 NDVI 和 LST 为数据，得到相应旬的 NDVI 和 LST 最大值合成数据。然后，对各年同月同旬的 NDVI 和 LST 最大值合成数据也进行最大值合成，得到年间各月各旬 NDVI 和 LST 最大值合成数据。此外，还对各年同月同旬的 LST 最大值合成数据进行最小值合成，得到年间各月各旬 LST 最大-最小值合成数据。接下来利用冷热边界计算功能，根据年间各月各旬 NDVI 和 LST 最大值合成数据以及年间各月各旬 LST 最小值合成数据，确定出各月各旬的冷热边界系数。最后，根据各年 3-6 月各旬的 NDVI 和 LST 最大值合成数据，以及该月该旬的冷热边界系数，利用 VTCI 计算功能得到相应旬的 VTCI 数据。其具体的处理过程如图 5-3。按照此过程进行处理，得到 1999-2003 年 3-6 月各旬的 VTCI 数据，作为进行预测的基础数据。

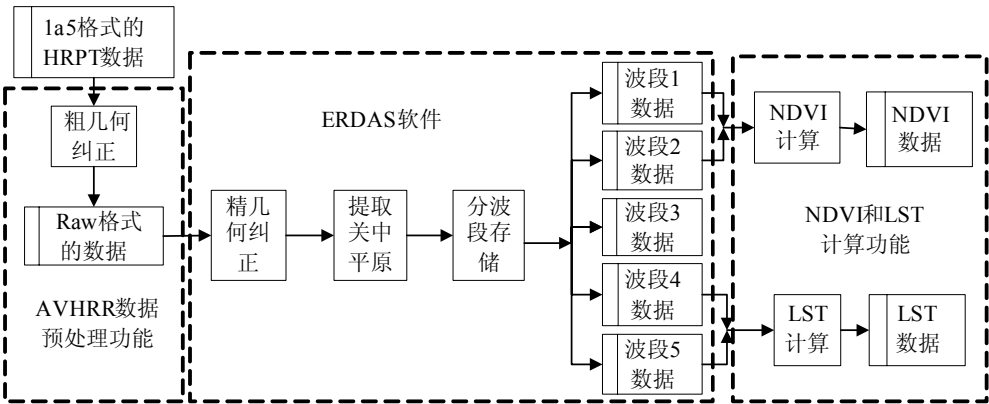


图 5-2 NDVI 和 LST 计算流程

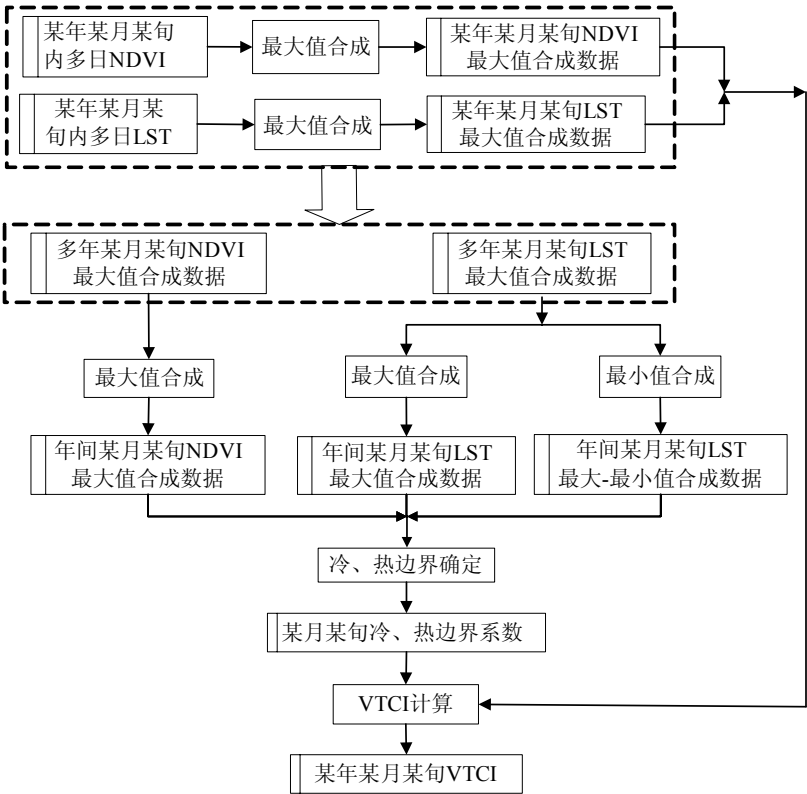


图 5-3 VTCI 计算流程

5.2 加权马尔可夫模型预测应用

5.2.1 基于标准化降水指数的预测

以计算得到的 36 个气象站时间尺度为 3 个月的 SPI 数据为基础，采用干旱预测模块中的加权马尔可夫模型预测功能对这些站点 2004 年 7 月份的干旱情况进行预测。对 SPI 等级的划分采用的是《气象干旱等级》国家标准中的划分标准<sup>[71]</sup>（表 5-1）。该功能的界面和相关参数的设置如图 5-4，计算完成后点击文件列表中的站点名称就可以查看该站点的相关预测结果，其界面如图 5-5。为了更为直观地了解各个站点的干旱预测情况，根据站点的经纬度和预测结果将其图形化，并与研究区域的边界进行叠加，如图 5-6 所示，其中圆形表示无旱状态，三角形表示轻旱状态，矩形表示中旱状态。

表 5-1 标准化降水指数干旱等级划分表

等级	类型	SPI 值
1	无旱	$-0.5 < \text{SPI}$
2	轻旱	$-1.0 < \text{SPI} \leq -0.5$
3	中旱	$-1.5 < \text{SPI} \leq -1.0$
4	重旱	$-2.0 < \text{SPI} \leq -1.5$
5	特旱	$\text{SPI} \leq -2.0$

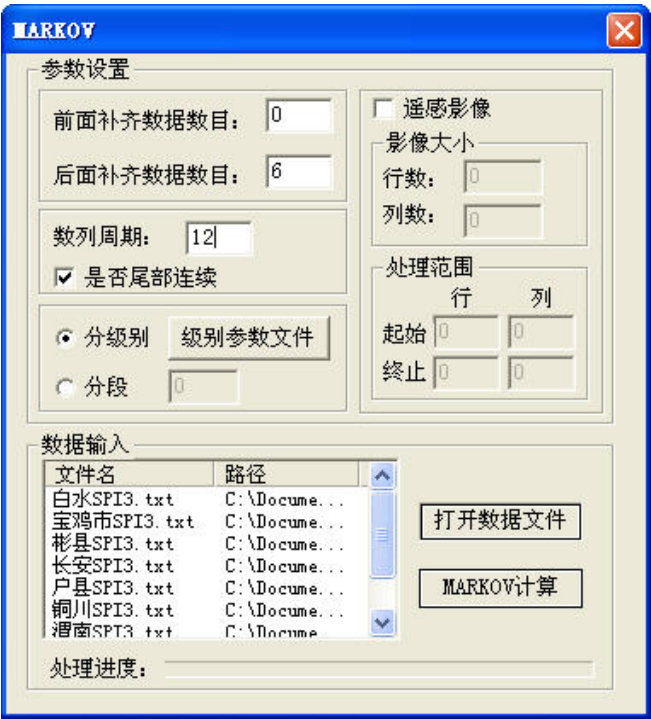


图 5-4 加权马尔可夫模型预测界面



图 5-5 单个站点预测结果显示界面

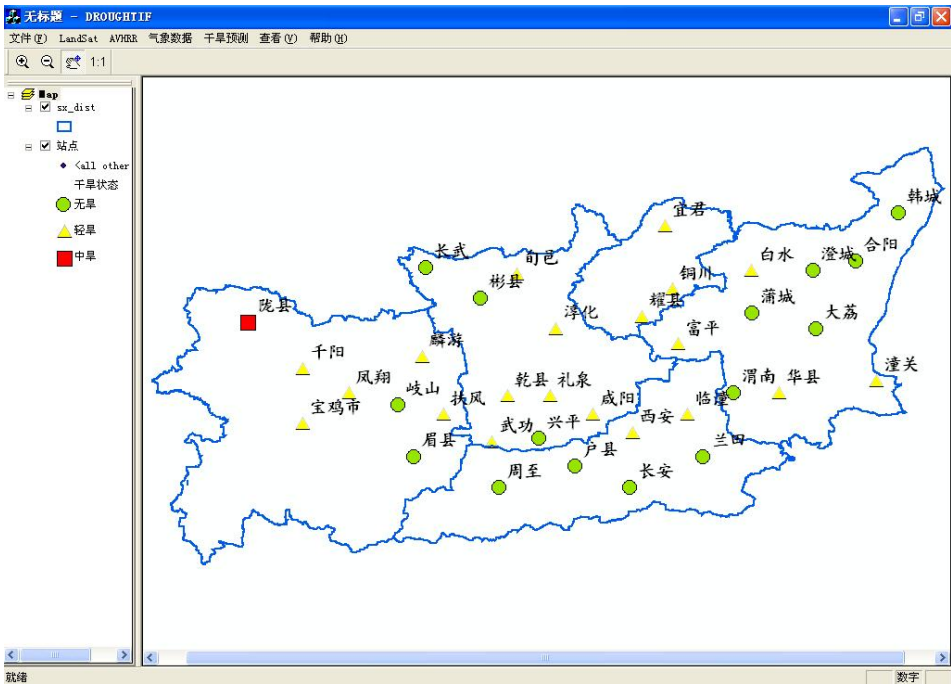


图 5-6 预测结果及站点分布图

为了对基于 SPI 的加权马尔可夫模型预测方法的预测能力进行进一步分析,采用同样的过程,根据这 36 个气象站的 5 个时间尺度(1 个月、3 个月、6 个月、9 个月和 12 个月)上的 SPI 值,应用加权马尔可夫模型对它们 2004 年 12 个月的干旱状态进行了预测。对预测结果分别从总预测正确率和五种状态各自的预测正确率两方面进行统计分析,结果见表 5-2。

从表 5-2 中可知,从时间尺度为 3 个月开始,预测正确率随着时间尺度的增加明显提高,而对同一时间尺度而言,随着干旱等级的增加,即干旱程度的加重,预测正确率逐渐降低。在表 5-2 中无旱的预测正确率在 5 种时间尺度上都超过了 90%,而轻旱的预测正确率最高的也只有 59%,其余三个等级的预测正确率则更低。针对这一情况,对所有错误预测进行了分析,发现其中存在



着不少预测出了干旱但未能正确预测干旱程度的情况，例如宝鸡市 2004 年 7 月干旱等级的预测值是 2，但实际干旱等级为 3。对于这种情况，虽然没有预测准确干旱程度，但它预测出了干旱的发生，从干旱预测的角度而言，仍然具有一定的价值。因此，将这种情况考虑到预测正确的范畴内对预测结果重新进行统计，发现时间尺度为 3 个月、6 个月、9 个月和 12 个月的总预测正确率都有所提高，分别为 76%、78%、89%和 93%，而在 5 种干旱等级中，除了无旱和轻旱的预测正确率与表 5-2 相比基本持平外，其余三个等级的预测正确率基本上都提高了 50%。

表 5-2 预测结果统计表

时间尺度		无旱	轻旱	中旱	重旱	特旱	合计
1 个月	实际数	321	61	23	14	13	432
	预测正确数	321	0	0	0	0	321
	正确率(%)	100	0	0	0	0	74
3 个月	实际数	300	65	48	17	2	432
	预测正确数	285	16	1	0	0	302
	正确率(%)	95	25	2	0	0	70
6 个月	实际数	286	85	51	9	1	432
	预测正确数	262	44	5	0	0	311
	正确率(%)	92	52	10	0	0	72
9 个月	实际数	320	80	30	2	—	432
	预测正确数	318	47	3	0	—	368
	正确率(%)	99	59	10	0	—	85
12 个月	实际数	383	41	8	—	—	432
	预测正确数	383	15	0	—	—	398
	正确率(%)	100	37	0	—	—	92

为了进一步分析该方法存在的问题，对所有错误预测的结果分两种情况进行统计，一种是漏分，即实际情况是 A 等级而预测情况却是 B 等级；另一种是错分，即实际情况不是 A 等级而被预测成了 A 等级。漏分和错分的统计结果分别见表 5-3 和表 5-4。从表 5-3 中可以看出，漏分在 5 个时间尺度的各个干旱等级基本都存在，而且在轻旱和中旱中比较严重。通过对漏分结果的进一步分析发现其预测值都偏向比实际等级低的干旱等级，这也是在表 5-4 中错分的情况主要集中在无旱和轻旱两个等级上的原因。以上分析结果表明干旱程度越严重，该方法预测的精度也越低，因此，要提高该方法的预测能力就必须提高其对中旱、重旱和特旱的预测精度。

表 5-3 漏分结果统计表

时间尺度	无旱	轻旱	中旱	重旱	特旱
1 个月	0	61	23	14	13
3 个月	15	49	47	17	2
6 个月	24	41	46	9	1
9 个月	2	33	27	2	0
12 个月	0	26	8	0	0

表 5-4 错分结果统计表

时间尺度	无旱	轻旱	中旱	重旱	特旱
1 个月	111	0	0	0	0
3 个月	89	40	1	0	0
6 个月	69	49	3	0	0
9 个月	44	18	2	0	0
12 个月	28	6	0	0	0

在对预测结果进行分析时，还发现当干旱状态发生转变时，预测的结果往往也不准确，而对于连续的相对平稳的状态预测效果则更好。同样在对预测正确率较高的结果进行分析后，发现它们中大部分干旱状态变化频率都比较小，而且连续处于同一状态的时间越长，预测效果也越好。

根据以上分析可知该方法在不同时间尺度上都具有一定的预测能力，并且随着时间尺度的增加预测效果也明显提高。因此，该方法适合应用于中长期的干旱预测。同时，该方法对无旱的预测比较准确，对干旱的发生也有一定预测能力，可以作为早期干旱预警的参考。但该方法也存在着一些不足：（1）其预测能力随着干旱程度的加重逐渐降低；（2）该方法虽然对持续干旱状态有一定的预测能力，但对干旱状态突变的预测能力较弱。针对这些问题，该方法还需要做进一步的改进，以提高其预测的精度。

5.2.2 基于 VTCI 的加权马尔可夫模型预测

从 1999-2003 年 3-6 月各旬的 VTCI 数据中选择 1999 年 5 月上旬至 2003 年 5 月中旬的 VTCI 数据，使用干旱预测模块中的加权马尔可夫模型预测 2003 年 5 月下旬的 VTCI，由此预测出该旬的干旱情况。该功能的界面、相关参数的设置以及预测结果界面如图 5-7。VTCI 的取值范围为  $[0,1]$ ，其值越小，在影像中的颜色越暗，相对干旱程度越严重。将 2003 年 5 月下旬实际 VTCI 的情况和预测情况进行比较，如图 5-8 所示，发现在原始图中位于晋陕豫三省交界处的黄河(影像右上角)在预测图中也界线清晰。虽然预测图对部分细节的反映与原始图还有差异，但从干旱总体分布来看，预测图与原始图也有相似之处，就是关中平原北部的干旱程度较南部更严重；黄河两岸存在着相对较重的干旱。



图 5-7 基于 VTCI 的加权马尔可夫模型预测

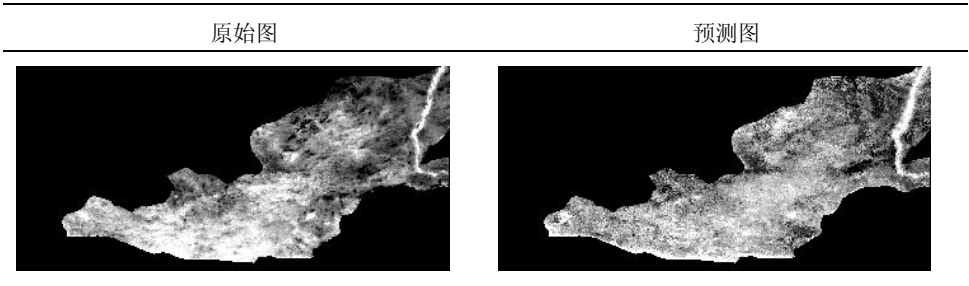


图 5-8 2003 年 5 月下旬 VTCI 影像

### 5.3 基于 VTCI 的 ARIMA 预测应用

基于 VTCI 的 ARIMA 预测也是采用的 1999 年 5 月上旬至 2003 年 5 月中旬的 VTCI 数据，使用干旱预测模块中的基于 VTCI 的 ARIMA 预测功能进行了 1 步预测，即预测了 2003 年 5 月下

旬的 VTCI 情况。该功能的界面和预测结果界面如图 5-9。将 2003 年 5 月下旬 VTCI 的实际情况和预测情况进行对比,如图 5-10 所示,发现对 2003 年 5 月下旬 VTCI 的预测从干旱整体分布来看,预测情况与实际情况基本相符,并且对其中部分细节在预测图中也有所反映。



图 5-9 基于 VTCI 的 ARIMA 预测

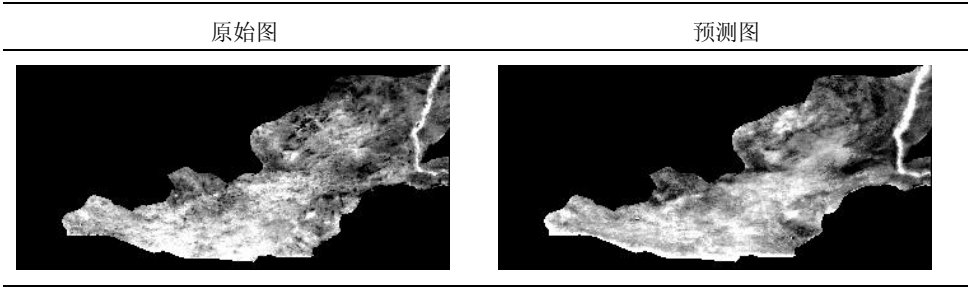


图 5-10 2003 年 5 月下旬 VTCI 影像

5.4 小结

为了对模块的运行情况进行检验,分别选用了两种干旱指标和干旱预测模型进行了实验性应用。以陕西关中平原和渭北旱塬的降水量为数据,先利用 SPI 计算功能得到其 SPI 指数,然后利用加权马尔可夫模型进行预测,并对预测结果进行了统计和分析。除了应用点上的气象数据外,还使用关中平原的 AVHRR 多时段卫星遥感数据,通过 VTCI 计算功能获得该区域的 VTCI 干旱监测指数,并采用了加权马尔可夫模型和 ARIMA 预测模型对 VTCI 的情况进行了短期预测。通过以上应用说明系统在功能上能较好地满足干旱预测应用的要求,完成和达到了预期系统开发的目标。

## 第六章 结论与讨论

### 6.1 结论

本文根据干旱预测应用的实际需要,在已有的遥感干旱监测数据处理相关功能的基础上,完成了一套集遥感干旱监测指标计算和干旱预测于一体的系统。结论如下:

- 针对本系统,特别是干旱预测模块所涉及到的数据,对它们的数据格式进行了设计,并采用面向对象程序设计的思想,实现了数据访问的接口,将对这些数据的相关操作封装起来。这样做不仅使对数据的访问更加方便和高效,而且也为系统日后的扩展,在数据访问方面打下了良好的基础。
- 在干旱预测模块方面,采用面向对象程序设计的思想和 Visual C++6.0 开发平台,设计和实现了加权马尔可夫模型、基于 GM (1,1)模型的灰色预测和基于 VTCI 的 ARIMA 模型,从而可以对干旱进行点上和面上的预测。
- 以陕西关中平原和渭北旱塬气象站点的降水量数据计算得到 SPI,采用干旱预测模块中的加权马尔可夫模型,对该地区的干旱情况进行了点上的预测应用;以关中平原和渭北旱塬 AVHRR 多时段卫星遥感数据计算得到的 VTCI,采用加权马尔可夫模型和 ARIMA 模型,对干旱情况进行了面上的预测应用。通过这些应用表明该系统在功能上能较好地满足干旱预测应用的要求,达到了预期系统开发的目标。

### 6.2 讨论

干旱预测研究的复杂性决定了该系统设计和实现的困难,再加上对系统的完善是一个循序渐进的过程,限于本人技术能力和时间的有限,许多研究还不够深入,本人认为以下几个方面的研究有利于本研究的进一步发展:

- 在干旱预测模型方面,目前的方法主要侧重于统计预测的方法,在进一步的研究中应该尝试一些其它模型。此外,对基于遥感的干旱预测模型也只是初步研究,还需要进一步验证和完善。随着研究的深入和需求的提高,需要不断完善系统的相关功能,使系统真正能够应用到防旱、抗旱的工作中。
- 在遥感影像数据格式方面,目前本系统只支持 RAW 格式的数据处理和 RAW 格式、BMP 格式的数据显示,这样就给用户在使用过程中带来不便。因此,在下一步的系统改进中需要增加对一些常用图像格式的支持,使系统能够处理多种遥感影像。
- 在系统功能方面,只有数据处理和干旱预测这些功能是不够的,需要在预测的基础上实现干旱预警,才能更好地为应用服务。此外,本系统 GIS 方面的功能还比较弱,只能完成一些常规的操作。随着研究的深入,GIS 相关功能的扩充势在必行。在 GIS 强大的数据存储、空间分析和信息查询的支持下,使系统成为一个综合干旱服务平台。

尽管本系统目前还存在一些不足,但作为一个专业干旱监测预测软件还是有其特点和优势。随着系统的不断完善,相信该系统能更好地发挥其作用。

## 参考文献

- [1] Palmer W C. Meteorological drought[J]. U.S Weather Bureau Research Paper, 1965, 45: 58
- [2] 王密侠, 胡彦华, 熊运章. 陕西省作物旱情预报系统的研究[J]. 西北水资源与水工程, 1996, 7(2): 52- 561
- [3] 邹仁爱, 陈俊鸿. 干旱预报的研究进展评述[J]. 灾害学, 2005, 20(3): 114-118
- [4] 耿鸿江. 干旱定义述评[J]. 灾害学, 1993, 8(1): 21- 24
- [5] American Meteorological Society. Meteorological drought-Policystatement[J]. Bulletin of American Meteorological Society, 1997, 78: 847-849
- [6] 王密侠, 马成军, 蔡焕杰. 农业干旱指标研究与进展[J]. 干旱地区农业研究, 1998, 16(3): 119-1241
- [7] 孙荣强. 干旱定义及其指标评述[J]. 灾害学, 1994, 9 (1): 17-21
- [8] Wang Mixia, Ma Chengjun, Cai Huanjie. Research progress in agricultural drought index[J]. Agricult ural Research in the Arid Areas, 1998, 16(3): 119-124
- [9] 冯平, 朱元. 干旱灾害的识别途径[J]. 自然灾害学报, 1997, 6(3): 43- 49
- [10] 朱自玺, 牛现增, 候建新. 冬小麦水分动态分析和干旱预报[A]. 国际旱地农业学术讨论会论文集[C], 陕西, 杨陵, 1987
- [11] 张叶, 罗怀良. 农业气象干旱指标研究综述[J]. 资源开发与市场, 2006, 22(1): 52-54.
- [12] Ohlsson. Water conflict and social resource scarcity[J]. Physics and Chemistry of The Earth (B), 2000, 25(3): 213-220
- [13] 许炳南, 武文辉. 贵州春旱短期气候预测信号及预测模型研究[J]. 灾害学, 2001, 16(3): 34-38
- [14] 周家斌, 黄嘉佑. 近年来中国统计气象学的新进展[J]. 气象学报, 1997, 55 (3): 42-50
- [15] Drosdowsky W. Analog (nonlinear) forecasts of the Southern scillation Index time series[J]. Weather and Forecasting, 1994, 9(1): 78-84
- [16] 刘思峰, 党耀国, 方志耕等著. 灰色系统理论及其应用[M]. 北京: 科学出版社, 2005. 126-164
- [17] Modarres R. Streamflow drought time series forecasting[J]. Stochastic Environmental Research and Risk Assessment, 2007, 21(3): 223-233
- [18] A. K. Mishra, V. R. Desai. Drought forecasting using stochastic models[J]. Stochastic Environmental Research and Risk Assessment, 2005, 19(5): 326-339
- [19] Yurekli k, Kurune A, Ozturk F. Application of linear stochastic models to monthly flow data of Kelkit Steam[J]. Ecological Modeling, 2005, 183: 67-75
- [20] Fernando, D A K, Jayawardena, A W. Generation and forecasting of monsoon rainfall data[A]. Proceedings of the 20th WEDC Conference[C]. Colombo: Sri Lanka, 1994, 310-313
- [21] 许建国. GM(1,1)模型和马尔柯夫模型在流域旱涝灾害预测中的运用[J]. 南京师大学报(自然科学版), 2002, 25(4): 121-124

- [22] Lohani V K, Loganathan G V. An early warning system for drought management using the Palmer drought index [J]. Journal of the American Water Resources Association, 1997, 33 (6): 1375-1386
- [23] Lohani V K, Loganathan G V, Mostaghimi S. Long-term analysis and short-term forecasting of dry spells by the Palmer drought severity index[J]. Nordic hydrology, 1998, 29(1): 21-40
- [24] Paulo A. A, Ferreira E, Coelho C, et al. Drought class transition analysis through Markov and Loglinear models, an approach to early warning[J]. Agricultural Water Management, 2005, 77(1-3): 59-81
- [25] 孙才志, 林学钰. 降水预测的模糊权马尔可夫模型及应用[J]. 系统工程学报, 2003, 18(4): 8-13
- [26] 孙才志, 张戈, 林学钰. 加权马尔可夫模型在降水丰枯状况预测中的应用[J]. 系统工程理论与实践, 2003, (4): 101-106
- [27] 王希娟, 唐红玉, 张景华. 近 40 年青海东部春季降水变化特征及小波分析[J]. 干旱地区农业研究, 2006, 24(3): 21-41
- [28] 刘引鸽, 缪启龙. 西北地区农业旱灾与预测研究[J]. 干旱区地理, 2004, 27(4): 564-569
- [29] 李玉爱, 郭志梅, 栗永忠, 等. 大同市短期农业气候干旱预测系统[J]. 山西气象, 2001, 54(1): 38-42
- [30] 李正明, 刘洪. 甘肃中部春夏干旱气候分析与预测系统[J]. 甘肃气象, 2003, 21(2): 1-3
- [31] 李凤霞, 伏洋, 张国胜, 等. 青海省干旱预警服务系统设计与建立[J]. 干旱地区农业研究, 2004, 22(1): 4-8
- [32] 景毅刚, 刘安麟, 张树誉, 等. 陕西省干旱评价和预警系统[J]. 陕西气象, 2004, (6): 23-25
- [33] 杨太明, 陈金华, 李龙澍. 安徽省干旱灾害监测及预警服务系统研究[J]. 气象, 2006, 32(3): 113-117
- [34] 刘治国, 王遂缠, 林纾, 等. 西北干旱监测预测业务服务综合系统的开发与应用[J]. 气象科技, 2006, 34(4): 485-489
- [35] Hsu-Yang Kung, Jing-Shiuan Hua, Chaur-Tzuhn Chen. Drought forecast model and framework using wireless sensor networks[J]. Journal of Information Science and Engineering, 2006, 22(4): 751-769
- [36] Kogan F. Droughts of the late 1980s in the United States as derived from NOAA polar orbiting satellite data[J]. Bulletin of the American Meteorological Society, 1995, 76(5): 655-668
- [37] Mendicino G, Senatore A, Versace P. Drought forecasting monitoring system, Geophysical Research Abstracts [J]. European Geophysical Society, 2003, 5: 06262
- [38] Sheffield J, Luo L, Wood E F. High resolution drought monitoring and seasonal forecasting for the USA, Geophysical Research Abstracts[J]. European Geosciences Union, 2006, 8: 08910
- [39] 许涌, 万国龙. 海量图像数据快速显示技术[J]. 计算机工程与设计, 2003, 24(6): 36-38
- [40] Gonzalez R. C, Woods R. E 著, 阮秋琦, 阮宇智等译. 数字图像处理(第二版) [M]. 北京: 电子工业出版社, 2004
- [41] 蒋红成, 赵忠明, 朱海青. 用分块调度置换算法快速浏览 SAR 带状海量影像[J]. 计算机

工程与应用, 2004, (28): 50-51

[42] 余粉香, 王光霞, 万刚. 大数据量遥感影像的快速调度与显示[J]. 海洋测绘, 2006, 26(2): 27~30

[43] 马照亭, 潘懋, 胡金星, 等. 一种基于数据分块的海量地形快速漫游方法[J]. 北京大学学报(自然科学版), 2004, 40(4): 619-625

[44] Kwok R, Johnson W T K. Block adaptive quantization of Magellan SAR data[J]. IEEE Transactions on Geoscience and Remote Sensing, 1989, 27(4): 375-383

[45] 宋江洪, 赵忠明. 图像分块分层结构在海量数据处理中的应用[J]. 计算机工程与应用, 2004, (33): 31-33

[46] 杨宁学, 诸昌铃, 聂爱丽. 内存映射文件及其在大数据量文件快速存取中的应用[J]. 计算机应用研究, 2004, (8): 187-188

[47] 胡伟忠, 刘南, 刘仁义. 基于内存映射文件技术的海量影像数据快速读取方法[J]. 计算机应用研究, 2005, (2): 115-116

[48] 陈涵生, 郑明华著. 基于 UML 的面向对象建模技术[M]. 北京: 科学出版社, 2006. 67-70

[49] 郑人杰, 殷人昆, 陶永雷著. 实用软件工程[M]. 北京: 清华大学出版社, 2003. 10-15

[50] McKee T B, Doesken N J, Kleist J. The relationship of drought frequency and duration to time Scales [A]. American Meteorological Society, Boston, Massachusetts, 1993, 179-184

[51] 姚玉璧, 张存杰, 邓振镛. 气象、农业干旱指标综述[J]. 干旱地区农业研究, 2007, 25(1): 191-195

[52] Hayes M J, Svoboda M D, Wilhite D A, et al. Monitoring the 1996 drought using the standardized precipitation index [J]. Bulletin of the American Meteorological Society, 1999, 80(3): 429-438

[53] Bussay A, Szinell C, Szentimery T. Investigation and measurements of droughts in Hungary[A]. Hungarian Meteorological Service, Budapest, 1999

[54] 袁文平, 周广胜. 标准化降水指标与 Z 指数在我国应用的对比分析[J]. 植物生态学报. 2004, 28(4): 523-529

[55] 袁文平, 周广胜. 干旱指数的理论分析与研究展望[J]. 地球科学进展, 2004, 19(6): 892-991

[56] 郭锐. 植被指数及其研究进展[J]. 干旱气象, 2003, 21(4): 71-75

[57] Chen Wenying, Xiao Qianguan, Sheng Yongwei. Application of the anomaly vegetation index to monitoring heavy drought in 1992[J]. China Remote Sensing of Environment, 1994, 9: 106-112

[58] Kogan F N. Remote sensing of weather impacts on vegetation in non-homogeneous areas [J]. International Journal of Remote Sensing, 1990, 11: 1405-1419

[59] Kogan F N. Application of vegetation index and brightness temperature for drought detection[J]. Advances in Space Research, 1995, 15: 91-100

[60] 王鹏新, 龚健雅, 李小文. 条件植被温度指数及其在干旱监测中的应用[J]. 武汉大学学报(信息科学版), 2001, 26(5): 412~418

[61] Wan Zhengming, Wang Pengxin, Li Xiaowen. Using MODIS land surface temperature and



Normalized Difference Vegetation Index products for monitoring drought in the southern Great Plains, USA [J]. International Journal of Remote Sensing, 2004, 25(1): 61-72

[62] 汪荣鑫. 随机过程[M]. 西安: 西安交通大学出版社, 1988. 185-186

[63] 冉景江, 赵燮京, 梁川. 基于加权马尔可夫链的降水预测应用研究[J]. 人民黄河, 2006, 28(4): 34-36

[64] 邓聚龙. 灰预测和灰决策[M]. 武汉: 华中科技大学出版社, 2002. 111-174

[65] 徐国祥. 统计预测和决策[M]. 上海: 上海财经大学出版社, 2005. 202-225

[66] 和宛琳, 徐宗学. 渭河流域近 40 年降水量变化规律及干旱预测. 人民黄河, 2007, 29(1): 38-39

[67] 陆殿梅, 于堃, 高琴. GM(1,1)灾变模型在奇台县旱灾年份预测中的应用研究[J]. 中国科技信息, 2006, (1): 78-79

[68] 陈天华. 面向对象程序设计与 Visual C++6.0 教程[M]. 北京: 清华大学出版社, 2006. 311-332

[69] 李俊霖, 延军平, 孙虎, 等. 关中平原东、中、西部气候干旱化程度比较分析[J]. 干旱区资源与环境, 2005, 19(1): 132-135

[70] 殷淑燕, 张佩丽. 近 46 年 El Nino/ La Nina 事件与渭北旱塬气候相关性分析[J]. 干旱区研究, 2003, 20(2): 131-134

[71] 中华人民共和国国家质量监督检验检疫总局, 中国国家标准化管理委员会. GB/T 20481-2006. 气象干旱等级. 北京: 中国标准出版社, 2006-11-01

## 致谢

本论文是在导师朱德海教授、刘峻明副教授的悉心指导下完成的。两年来，三位老师不仅在学业上给予我无微不至的精心指导，还在治学态度和治学方法上使我受益匪浅。本论文从选题、构思到最终完成倾注了三位老师的大量心血。在研究和论文撰写过程中，刘峻明老师在编程和论文构思方面给予了大量的指点和帮助。本论文的选题来自于王鹏新副教授主持的国家自然科学基金项目和教育部科学技术研究重点项目，王鹏新老师在遥感技术方面给予了很大的指导，使我在研究上少走了很多弯路，并且在论文写作方面提出了很多宝贵意见。

两年的时间虽然短暂，但是在导师的严格要求下，本人在专业理论知识、科研工作能力以及写作水平等方面都有了较大的提高。我不仅学会了该如何做学问，而且深刻地领会到了做人的道理。对导师的教诲我将铭记于心，终生难忘。

本研究得到国家自然科学基金项目（40571111、40371083）、教育部科学技术研究重点项目（105013）和“十一五”国家科技支撑计划课题（2006BAD10A01）的资助，使用了课题组成员提供的相关资料和数据，在此表示衷心地感谢。

在研究和论文的写作的过程中，得到了孙威师姐、黄建师兄以及韩萍老师的大力支持。孙威师姐提供了研究所需要的数据资料。黄建师兄在程序设计方面提供了宝贵的经验。韩萍老师在模型算法上提供了很多帮助。

感谢本系严泰来老师、张晓东老师、杨建宇老师。他们在我的学习上、工作上和生活上都给予了很大的指导和关心。

感谢信息管理系的全体同学，他们在学习上和生活中都给予了我真诚的帮助和关心。

感谢信电学院所有关心和帮助过我的老师、同学。

最后，特别感谢我的父母，他们无私而伟大的爱是我精神的支柱，激励着我不断前进！

## 个人简历

### 个人简介:

王彦集: 男, 1982 年 10 月生, 重庆铜梁县人, 硕士研究生, 毕业于中国农业大学。

### 在校期间参与科研项目:

1. 国家自然科学基金项目: 近实时量化遥感干旱监测方法的研究;
2. 北京市 1: 5 万土地利用图缩编项目;
3. 山东德州市土地定级估价信息系统开发项目。

### 发表论文:

1. 王彦集, 刘峻明, 王鹏新, 韩萍, 朱德海, 张树誉. 基于加权马尔可夫模型的标准化降水指数干旱预测研究. 干旱地区农业研究, 待刊.
2. 王彦集, 杨建宇, 樊雅婷, 朱德海. 基于 MapGis 的城镇土地定级系统的实现. 微计算机信息, 待刊.
3. 樊雅婷, 杨建宇, 朱德海, 王彦集. 基于阈值的城镇土地定级距离衰减模型. 武汉大学学报(信息科学版), 待刊.

### 软件著作权:

1. 计算机软件著作权登记证书, 登记号: 2006SRBJ3031; 软件名称: 城镇土地定级系统 V1.0; 著作权人: 中国农业大学; 权利获得方式: 原始取得; 权利范围: 全部权利; 首次发表日期: 2006 年 8 月 15 日, 授权日期: 2006 年 12 月 30 日。
2. 计算机软件著作权登记证书, 登记号: 2006SRBJ3032; 软件名称: 城镇土地估价系统 V1.0; 著作权人: 中国农业大学; 权利获得方式: 原始取得; 权利范围: 全部权利; 首次发表日期: 2006 年 8 月 16 日, 授权日期: 2006 年 12 月 30 日。
3. 计算机软件著作权登记证书, 登记号: 2006SRBJ3033; 软件名称: 城镇土地监测与地价管理系统 V1.0; 著作权人: 中国农业大学; 权利获得方式: 原始取得; 权利范围: 全部权利; 首次发表日期: 2006 年 8 月 15 日, 授权日期: 2006 年 12 月 30 日。