

文章编号: 1006-4354 (2009) 02-0035-04

陕西省级气象科学数据共享数据库设计与实现

夏巧利, 陈高峰, 邸永强, 高宇

(陕西省气象信息中心, 西安 710014)

摘要: 介绍陕西省级气象科学数据共享数据库建设的内容、技术方案、设计参考标准、结构、信息流程等。元数据的基本概念、作用和使用方法。介绍了元数据库的表名、表结构及基于元数据导航的数据发布; 数据实体库设计的基本原则, 数据集及数据库表; 数据库管理系统的设计及维护。

关键词: 气象科学数据; 共享数据库; 设计开发

中图分类号: P409

文献标识码: A

陕西省级气象科学数据共享数据库建设是中国气象局承担的国家科技部专项资金项目“气象科学数据共享中心”下“省级气象数据资源建设与共享服务”课题中“陕西省气象科学数据共享系统建设”专题的开发建设内容之一, 包括陕西省气象科学数据共享元数据库、实体数据库和数据库管理系统的开发建设, 于2006年底建成并投

入业务运行。

1 技术方案及设计参考标准

1.1 技术方案

建立以省级气象资料数据集为基本数据、以商用数据库为平台、对结构化和非结构化数据为用户提供透明访问存取。与现有业务系统结合, 以定时启动或人工控制方式, 建立数据集的实时或

收稿日期: 2008-09-25

作者简介: 夏巧利 (1963—), 女, 陕西户县人, 高工, 从事气象信息管理、开发建设。

基金项目: 国家科技基础性工作专项资金项目“气象科学数据共享中心”(2005DKA31700)

坡轻。向阳坡地因气候温暖, 作物生育期偏早, 同样的天气条件下, 处于盛花期的花椒造成的损失更严重。春季霜冻已成为严重制约凤县花椒产业发展主要瓶颈, 充分利用当地小气候资源, 因地制宜地发展花椒种植产业, 最大限度规避和减轻花椒霜冻危害。

4 花椒春季霜冻的防御

4.1 凤县大红袍品质好, 种植面积大, 但发育期早, 抗冻性差, 受霜冻危害重, 用现代育种技术选育抗冻性强的品种。

4.2 凤县特殊的气候、地形条件已决定大红袍花椒适宜种植范围, 大规模推广种植时, 要重视当地气候保证率, 盲目扩大种植造成的经济损失是巨大的。调查表明: 平木是退耕还林大规模发展花椒种植最早的乡镇之一, 大规模化种植已有10 a多历史, 几乎每年都有不同程度的霜冻发生, 其

中5 a出现严重霜冻, 损失很大。

4.3 充分开发利用当地小气候资源, 因地制宜发展花椒。调查表明: 双石铺的西山、张家窑一带, 大红袍花椒种植面积大, 近年一直保持稳产, 霜冻危害甚微, 与该区域独特的地形所形成的小气候有着极为密切的关系。

4.4 花椒花期霜冻对产量影响最大。通过对凤州王家山农场(海拔1700 m)大红袍花椒种植的实地调查, 花椒近几年未发生明显因霜冻受害而减产的现象, 原因是海拔高、生育期晚, 特别是花期在5月上旬, 有效地避开了花期霜冻。使用生长抑制剂延缓作物生长, 使花期避开霜冻高发期不失为一项有效的避霜的措施。

4.5 开展物理、化学方法防治花椒霜冻试验研究, 探索适合当地防冻害的技术措施及霜冻后的补救措施。

准实时追加自动化业务流程。操作系统为 Microsoft windows server 2003, 数据库平台为 SQL Server 2000。

1.1.1 数据的存储与管理 将气象资料统一分类和编码,根据数据格式进行数据处理和拆分,设计数据库和数据文件目录结构和框架,建立数据库管理系统,对结构化数据和元数据由数据库存储管理,非结构化数据以文件方式分级、分类存储。建立数据追加更新流程,实现数据的实时或准实时的追加更新。

1.1.2 数据的检索与分发 利用数据检索策略,根据数据检索关键字,由查询语句访问元数据库,通过元数据库找到基础数据库或文件最终获取数据。对检索获得的数据采用灵活的分发方式,可由页面浏览或以FTP方式下载,显示以表格为基本方式,并辅助以图形等方式。

1.2 设计参考标准

《气象数据分级方案》,《气象数据集元数据格式标准》,《气象资料的分类编码及命名规范》。

1.3 框架

共享数据库的结构、维护以及信息流程如图1。

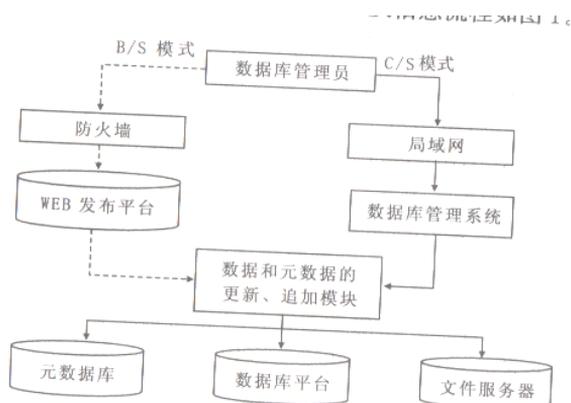


图1 共享数据库的结构、维护以及信息流程

2 元数据的基本概念和作用

分布在全国各地的气象数据种类繁多、格式复杂、存储方式各不相同,常规的文档管理方法无法面向用户需求提供有效的检索和处理,迫切需要一种方便有效的方法实现气象科学数据的共享管理,以发挥气象科学数据社会效益和实现其科学价值,元数据为实现数据的实时在线、可

扩展的共享管理提供强有力的手段。

2.1 元数据的定义

元数据就是“关于数据的数据”。元数据为各种形态的数字化信息单元和资源集合提供规范、普遍的描述方法和检索工具。元数据为分布的,由多种数字化资源有机构成的信息体系(如分布式气象科学资料数据库)提供整合的工具与纽带。

在气象资料领域,元数据主要是指说明气象资料数据集的一些信息(包括内容、质量、状况和其他有关特征)和检索工具,主要包括:数据集编码、数据集名称、资料类别、要素或项目名称、观测次数和时间、观测时制、统计特征、数据集内容关键词、数据起止时间、数据空间范围、数据表达方式、数据质量状况、数据存储介质、数据存储格式、数据量、数据来源、数据集所使用语种、数据集作者信息、元数据作者信息、数据集存放地点、数据集索取方式以及数据尺度、数据更新周期和数据空间参考系及坐标等。

2.2 元数据的作用

元数据在数据的共享发布过程中具有重要的衔接作用。利用元数据对实体数据的详细、全面的记录描述,协助管理员管理实体数据。

3 元数据库的设计与实现

3.1 元数据库表

元数据库共建有16个数据库表,其表名及相关说明见表1。

3.2 元数据库表结构

不同的表其结构不同,以地面台站信息表为例:

```

if exists (select * from sysobjects where id =
object_id ('STN _ CHN _ SURF') and
OBJECTPROPERTY (id, 'IsUserTable') = 1)
drop table STN _ CHN _ SURF
Create table STN _ CHN _ SURF
(

```

```

STA _ ID varchar (5), 一区站号
STA _ NAME varchar (30), 一站名
STA _ CLASS numeric (2), 一台站共享级别

```

表 1 元数据库表的表名及相关说明

序号	数据表名	中文名称	用途	备注
1	Data_Catalog	数据集分类	描述数据集的分类信息	检索用表, 字段均不为空
2	DS_DESC_Info	基本描述信息	描述数据集基本元数据信息	
3	DS_File_Info	数据集与文件对应	主要描述以文件格式存储的数据基本的信息	检索用表, 字段均不为空
4	DS_STOR_Info	数据集存储策略(方式)描述	描述数据集的存储方式: t 为以表存放、f 为以文件存放	检索用表, 字段均不为空
5	DS_Table_Info	数据集与表对应	数据集与表对应	检索用表, 字段均不为空
6	ELE_Info	要素元数据		
7	Meta_ Table_ Fields	字段描述表, 描述各数据库表结构	描述表字段信息	检索用表, 字段均不为空
8	PROV_Info	州/省编码表	省编码	
9	Search_Way_Info	检索策略信息描述		
10	STN_CHN_RADI	太阳辐射台站信息表	辐射台站信息	检索用表
11	STN_CHN_SURF	地面台站信息表	地面台站信息	
12	STN_CHN_UPAR	高空台站信息表	高空台站信息	检索用表
13	STN_Info	台站类型说明表	台站类型说明	
14	Used_Data_Info	用户使用数据信息记录	主要描述以文件格式存储的数据基本的信息	
15	User_Info	用户基本信息描述	描述数据库用户(非应用检索用户)的基本信息属性	检索用表, 字段均不为空
16	UserRight_Info	权限定义表	描述用户权限分配基本信息	

PROVCode varchar (3), 一省代码
 Latitude numeric (5, 2), 一纬度
 Longitude numeric (6, 2), 一经度
 ELEV1 numeric (6, 2), 一水银槽高度
 ELEV2 numeric (6, 2), 一拔海高度
 TYPE varchar (20), 一台站类型
 NOTES varchar (1024), 一台站沿革说明
 Begin_Year numeric (5), 一起始年
 Begin_Mon numeric (3), 一起始月
 End_Year numeric (5), 一终止年
 End_Mon numeric (3) 一终止月
 constraint pk_STN_CHN_SURF primary
 key (STA_ID)

表、字段命名方式: 根据《气象资料分类编码和命名规范》数据表名称、字段名称的确定。

4.2 数据集及数据库表

4.2.1 数据集 常规资料数据集: 包括全省 99 个台站 1951—2005 年 8 个气象要素(气压、气温、水汽压、相对湿度、降水量、蒸发、风、日照)的时、日、旬、月、年值资料, 数据集分别是: 陕西省地面气候资料定时值数据集, 陕西省地面气候资料日值数据集, 陕西省地面气候资料旬值数据集, 陕西省地面气候资料月值数据集, 陕西省地面气候资料年值数据集。全省 20 个原国家基本基准站 1991—2006 年的逐小时降水数据集和逐小时风数据集。全省 66 个地面自动气象站 2004—2006 年 8 个气象要素(气压、气温、水汽压、相对湿度、降水量、蒸发、风和日照)的逐小时数据集。

特色资料数据集: 选择近几年用户使用较多, 且发生频率较高的天气现象。本课题选取全省 96 个台站 1981—2005 年雷暴、雾、沙尘天气现象资料, 由 W 文件加工制作。包括 4 个数据集, 分别

4 实体数据库的设计与实现

4.1 设计的基本原则

尽量减少数据库维护的工作量: 与数据集对应, 一个数据集对应一个数据表。检索条件确定、外键关联: 与元数据库相关描述一致, 用户级别与检索条件(主键字段的对应)。统一规范的数据

是：陕西省沙尘暴日值数据集，陕西省地面气候资料沙尘日数月值数据集，陕西省雷暴数据集，陕西省雾数据集。

4.2.2 数据库表 数据库表的建立与数据集对应，一个数据库表对应一个数据集。

陕西省地面气候资料定时值数据库表，表名为 SURF_CLI_SN_MUL_FTM；

陕西省地面气候资料日值数据库表，表名为 SURF_CLI_SN_MUL_DAY；

陕西省地面气候资料旬值数据库表，表名为 SURF_CLI_SN_MUL_TEN；

陕西省地面气候资料月值数据库表，表名为 SURF_CLI_SN_MUL_MON；

陕西省地面气候资料年值数据库表，表名为 SURF_CLI_SN_MUL_YER。

陕西省沙尘暴日值数据库表，表名为 SURF_CLI_SN_WEP_DUS_DAY；

陕西省地面气候资料沙尘日数月值数据库表，表名为 SURF_CLI_SN_WEP_DSS_MON；

陕西省雷暴日值数据库表，表名为 SURF_CLI_SN_WEP_THU_DAY；

陕西省雾日值数据库表，表名为 SURF_CLI_SN_WEP_FOG_DAY。

4.2.3 表结构 实体数据库中不同的表其结构不同，每个表分别由不同的字段、数据类型和单位组成。以陕西省雷暴日值数据表为例简要介绍表的结构，在雷暴日值数据表中区站号、年、月、日、天气现象代码和天气现象描述分别对应的字段描述为 V01000、V04001、V04002、V04003、C20064_001 和 C20064_002，其数据类型除天气现象描述为 Nvarchar 外，其余均为 Numeric。

5 数据库管理系统设计与实现

5.1 基本要求和设计原则

最大限度地获取数据库的当前状态、基本信息；迅速初始化数据库、实现标准批量资料迅速入库；自动和手工的资料入库。简单、实用，人机

交互、友好，可靠、稳定。

5.2 结构框架

省级共享数据库管理系统的结构框架如图 2 所示。

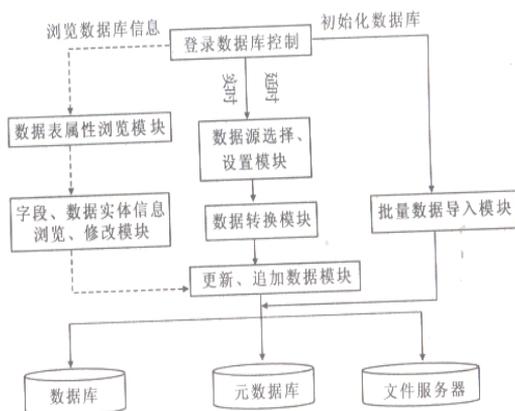


图 2 省级共享数据库管理系统的结构框架

5.3 各组成部分功能要求

5.3.1 数据库信息浏览 数据表名、字段名、字段属性、数据记录长度查看；数据表名、字段名、字段属性修改。

5.3.2 数据实时、延时更新 数据源选择模块：数据源类型、路径、时空属性设置；数据转换模块：根据设定的数据源属性转换数据，使之与数据库相应表对应。更新、追加数据模块：修改、删除已有记录、追加新的数据记录，完成数据更新，生成数据更新日志。

5.3.3 初始化数据库 数据初始化模块：建库、生成数据表。批量数据导入模块：将标准格式批量数据导入到指定的数据表。

6 结束语

陕西省级气象科学数据共享数据库是“陕西省气象科学数据共享服务系统”的基础，随着我省气象综合观测系统的发展，会有大量的资料需要共享服务，气象科学数据共享数据库入库的资料种类和数量会越来越多，功能也会越来越齐全。