

省级大气环流资料多用户管理检索系统

胡厚麟 程 煜

(浙江省气象科学研究所)

一、概 况

在天气预报和研究中,大气环流资料具有很重要的作用。我们所去年从国外引进的一台 PDP-11/23 计算机,拥有较丰富的外存容量,操作系统具有很强的数据文件管理功能。在这个基础上,我们着手建立一个省级大气环流资料多用户管理检索系统。系统集合了省级气象单位所需的基本环流资料,以约定的数据文件形式存放在磁盘或磁带上。整个系统分为三个部分,即资料的预处理、数据的组织和管理、资料的检索和应用。其中前面两个部分是面向系统管理员的,第三部分则是面向用户的。系统管理员和用户通过系统提供的不同种类的命令和工具,实施数据文件的建立、维护和管理,或者通过各种终端设备去访问数据文件中的数据。

系统除个别子程序采用 PDP-11 汇编语言外,全部采用 FORTRAN 语言编写,具有较强的通用性,凡配有 FORTRAN 语言的机种,都能移植该系统。系统除现已集合的资料外,视资料来源情况,可以不断得到扩充。

二、资料的预处理

系统目前处理的大气环流原始资料包括以下几个方面:

1. 北半球逐月月平均 500 mb 高度网格点 (1951~1983 年)

2. 北半球逐月月平均 100 mb 高度网格点 (1956~1983 年)

3. 北半球逐日逐次海平面气压网格点 (1951~1979 年)

4. 北半球逐日逐次 500 mb 高度网格点 (1951~1980 年)

5. 大气环流特征值*

6. 其他资料**

由于以上资料的来源及数据的存放格式不同,因此在把资料组织成约定格式的数据文件的预处理中,采用了不同的方法。

1. 可显示字符信息的处理

这类信息是以 ASCII 码、EBCDIC 码或国际电传码等代码形式构成的。由于是字符信息,可直接显示和打印,因此处理比较方便,只需进行代码转换,滤去控制符、间隔符,以及纠错处理等。如第 1, 2 项资料是属于这一类的。

2. 二进制压缩码信息的处理

第 3, 4 项资料是由二进制压缩码构成的。这类代码属非直接显示型代码,必须首先对压缩的二进制码按一定的规则进行适当的二进制位 (bit) 截取操作。由于 PDP-11 计算机没有配备对二进制位进行操作的系统子程序, FORTRAN 语言也不具备这样的处理功能,所以这种码处理是由我们编写的 PDP-11 汇编子程序来完成的。

以上信息经过预处理后,就成为 FORTRAN 数据文件能够接受的、本系统管理检索的最基本元素——数据项。

三、数据的组织和管理

1. 数据的组织

一个大气环流资料的管理检索系统,首先必须强调系统中的数据应反映气象业务所要处理的实体对象及其相互关系,其目的在于找出适合气象业务处理的数据的逻辑结构。其次必须强调系统中的数据应按它们所依附的存储介质的特点来加以表现,其目的在于确定数据的物理结构。因此,我们在全面衡量了所处理的大气环流资料的属性、气象业务和科研应用的具体情况,以及计算机系统的设备配置情况后,确定了数据的逻辑结构和物理结构。

1) 数据的逻辑结构 大气环流资料中,一个格点的数据,除了有其自身确定的数值外,还有在其等压面上的物理位置、时间、空间的定义。我们所确定的逻辑结构包含了四种基本的数据结构,即数据项、

* 大气环流特征值资料包括有该内容的全部资料 (1951~1983 年)。

** 包括有美国海军全球月平均海表水温和逐日海表水温等 (1951~1983 年)。

记录、文件和文件子集。

数据项反映了相应实体的某种属性,在这里,一个数据项的数值、位置、时间及空间的定义都是确定了的。

记录是由一些数据项组成的集合,它反映某一特定的实体,即某一格点资料在时间上的延伸。

文件是由记录组成的。在我们的定义下,它反映出了大气环流在历史年代中,同一时刻、同一空间面上的情况。

文件子集是由各种文件所组成的集合,它反映了各种实体和实体间的关系。这种文件的集合通过约定可以有不同的组合,因此它既可以反映出环流形势在同一时刻不同空间的演变,又可体现出同一空间面随时间变化的发展情况。

2) 数据的物理结构 它与信息所依附的存贮介质的特点有关。大气环流资料信息之大是完全可以预知的。它们中除了一部分存贮在快速外存磁盘上外,还有大部分须存放于磁带上,所以本系统对所有文件均构造为顺序结构的文件,因为只有这种结构形式的数据文件,才能共存于磁盘和磁带这两种不同类型的介质上。顺序结构的优点是能够合理地、有效地利用存贮空间,可以顺序加工处理数据,也能按某一特定的记录号随机存取。

2. 数据的管理

系统对数据的管理是通过系统对数据文件的各种定义,以及由系统提供给系统管理员的命令和工具来实现的。

由于本系统是一个多用户系统,因此数据的管理是围绕以下几个方面来设计的:

1) 数据的共享性 系统中所有数据文件中的数据可以同时为多个终端用户或应用程序所使用,这样可充分利用存贮空间,避免数据的重复冗余。

2) 数据的安全性 由于数据是共享的,系统对数据的存取权加以控制,以防止用户对数据的非法存取或有意无意的损坏。除系统管理员拥有存贮权和修改权外,其他用户仅有数据获取权。

3) 数据的独立性 文件中的数据是按照上面阐述的逻辑结构和物理结构来存放的。但应用程序和用户不依赖于数据的组织方式和存取位置,不同的应用程序均可按自己的需要,以一定的检索命令去索取文件中的数据。而当文件中数据的内部组织发生变化时,不需要重新编制或改写已有的应用程序。

4) 数据的准确性 从原始资料到成为文件中

信息的整个过程中,需经过多个步骤,错误可能来自某个环节,或由于硬设备发生故障而导致信息的出错或丢失。系统具有随机查错和纠错的功能,并可利用转贮、修复等手段以保证资料的可靠性。

5) 数据的完整性 本系统中的所有数据文件,都有随时增加新的资料的可能,系统对新引入的资料经加工鉴别,装入相应的文件中,随时保证资料的完整。

四、资料的检索和应用

本节将介绍用户如何通过系统提供的各种命令去同数据打交道。作为系统的这一部分,是用户访问数据的界面,它接收、分析、并解释用户提出的命令请求,然后转到相应的处理程序去操纵数据文件中的数据。

1. 检索特点

和通常的数据管理系统一样,用户对数据的检索是通过系统提供的终端操纵语言来完成的。这种语言的表现形式有多种,本系统所采用的是关键字命令方式,即由系统给出参数名字和参数值的选择范围。用户根据给定的范围,选取适当的参数值,由终端键盘输入。系统再根据用户键盘输入的关键字命令逐一处理,实行人机会话,交互作用。例如:

```
系统输出:  LAYER OPTION:
              500 MB/F      ;
              100 MB/O      ;
              SEA LEVER/S   ;
              :
              :
              :
用户输入:   F <CR>
              :
```

例中,用户从层次选择开始,随后逐步由系统给出的参数,加细选择范围,直至检索出所需的数据为止。这种检索方法不仅简化了各种条件的逻辑运算和逻辑判断,加快了检索时间,而且它还具有很强的直观性,特别适合于非程序员用户,因为他们是运用大气环流资料用户中最多的一类用户。这类用户并不习惯具有严格语法约定的语言或操作语句。而面对这样的方法,即使是初次使用系统的用户,也完全能根据系统的提示,作出快速的响应。

2. 派生资料

所谓派生资料就是在系统集合的资料基础上,根据用户的需要,加工处理成的新资料。这也是气象业务要求所必须具备的一种功能。如中期天气预报需要候平均或旬平均资料;环流形势的分析需要

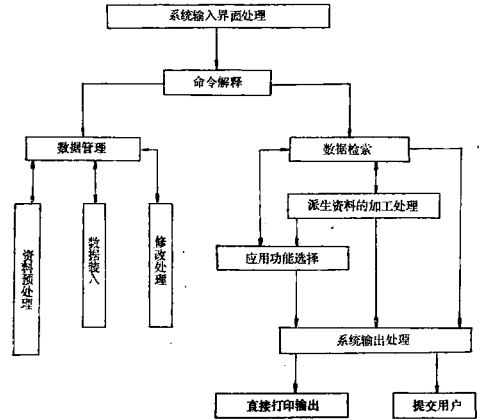
多年准平均值或距平值,以及N天滑动平均值等等。这种资料加工处理的功能,实质上也可以说是检索功能的扩充。虽然系统为此增设了一定的处理模块,但系统在处理时,采用边检索边处理的方法,对产生的新资料,一般处理成临时文件形式,这种文件在达到用户或用户程序的使用目的以后,在系统关闭时自行删除。这样它不受到存贮空间的限制,而其灵活性和可塑性也是十分显著的。它完全可以根据用户的要求,加工出各种形式的资料。例如,用户可选取任意历史年份的多种准平均值、N天滑动平均值等等。

派生资料的处理,完全符合气象业务中对资料的习惯处理要求。如候平均资料中每月最后一候的天数,根据月大月小统计为六天或五天。如果当月是二月份则为三天或四天。

3. 资料的应用

检索出的任何一种资料,可直接为用户程序所使用。但在气象业务和科研业务中,使用大气环流资料的许多应用程序是常用的或具有通用性。如在中、长期预报中,常用统计方法对环流资料进行统计普查,选取预报因子。为了减少不必要的重复劳动,同时又能使系统充分地为广大非程序员用户服务,开拓系统的应用面,我们设计了一个环流资料应用程序子系统。这是本系统的一个可选的子系统,它具有和系统一致的程序接口和内存的分配。但并不依赖于本系统,可单独运行。就拿统计资料一项来说,如果用户在开启系统后,选择了这一运用目的,那么用户只需输入预报的对象,确定资料的统计范围,就能很快以打印文件的形式,取得统计结果。又如:用户常希望将统计取得的预报因子样本和预报对象,用逐步回归法求得预报方程。那么,用户只需将选择好的预报因子序号(即格点序号)回送给系统,系统即把组织好的因子样本以临时文件的形式存在磁盘上,然后调用逐步回归程序进行运算,直至取得预报方程。这种从数据资料的检索到取得某一特定目的的运转结果的配套方式,使得较单纯的数据管理检索系统具有较强的实用性。

附图是系统功能模块的示意框图。



附图 系统功能模块示意框图

五、结 束 语

整个系统通过使用,用户的反映是操作简便、检索快、又很实用。特别是在统计工作中,减少了资料的收集、整理、统计等工序,大大缩短了研制一个预报方法的周期。在天气分析中,能灵活地提供各种形势图,带来了很大方便。

目前系统处理的资料都属非实时资料,如果解决了通信等方面的问题,只要增加相应的实时资料处理模块,就可以将实时资料接收处理后,直接装入系统数据文件中。

由于我们的计算机系统还未配备数据库管理系统(DBMS),因此本系统还不可能具备数据库的全部功能。一旦配备了DBMS,本系统即可组织成气象数据库的一部分。

上述工作得到了我科研所钱国荣同志的直接指导和支持。在收集资料及系统的设计中,得到了国家气象局北京气象中心许建民同志、长期预报组和气象科学研究院何家骅同志的大力支持,特此表示感谢。