

一种用于气候相关资料管理和研究的新手段

L. V. Novak 等

1. 前 言

美国国家航空和宇宙航行管理局(NASA)的试验性气候资料库管理系统(Pilot Climate Data Base Management System, 缩写为PCDBMS)为研究人员利用气候相关资料提供了一种新的手段。PCDBMS的建立是为了对收集到的大量有价值的气候资料进行集中管理。这一系统可提供某些选定的NASA和非NASA资料集的统一分类目录、详细清单及其读取方法,研制了一些适用的资料管理程序,使科学工作者能得到这些资料集的图表化统计结果。

PCDBMS 原为美国国家气候计划的一个部分,目的是让外界用户方便地使用 NASA 的卫星资料。然而出于各种原因,PCDBMS 发展成了为本部门气候研究人员提供卫星和地面观测资料的系统。上述这一重要变化要求 PCDBMS 提供比原计划更多种类的资料和超出原设想的数据管理和显示实用程序。因而,PCDBMS 演变为具有灵活选择各类资料及分析方法的功能很强的软件包。

PCDBMS 为各类用户提供 服务。研究
人员可利用这一系统对所感兴趣的各
种气候参数进行扫描、分析、组织、
比较、显示和研究;资料加工人员可
利用这一系统确定资料

的有效性,列出资料的详细清单及进行统计加工;对于那些经费不多的一些非经常用户,本系统可为他们迅速提供所需的少量的资料。另外,管理人员也可利用这一系统得到制定行动计划所需要的概要性信息。

PCDBMS 由五部分组成,它们是,

(1) 一份内容广泛的联机分类目录, 对大量资料集做出规范化的描述;

(2) 一份联机的 PCDBMS 库存资料详细清单;

(3) 各类可供读取的资料集和一系列根据时间或地理区域进行资料选择的功能;

(4) 资料组织管理实用服务程序组;

(5) 资料显示实用服务程序组。

2. 分类目录

PCDBMS 的分类目录为用户集中提供有关以标准格式存储的各种资料集的主要信息。它可用来寻找资料和确定这一资料的可获性。

目录对气候相关资料集及所用感应器件性能做出描述。它所包括的信息远远超出了PCDBMS实际库存资料的项目。这种描述包括资料的特性、处理情况、可获性以及为进一步获取更详细信息所联系者的姓名等。它不仅对现有的资料集,而且对计划中的资料产品均做出描述,目前已有200多个资料集

列入分类目录。

目录是以硬拷贝形式存放的,但完全可用联机方式提取。联机目录分为摘要和详细说明两部分,前者利用关键词询问,后者同一本书一样可以一页一页地显示。摘要部分把每个资料集压缩为一行信息输出,也可在CRT终端上给每个资料集显示一屏(帧)较为详细的信息。详细说明部分可一页一页地为用户提供任一资料集的更详细信息。

查阅目录和清单时也可利用打印服务程序,可以节省荧光屏的幅面,供打印或联机阅览时参考。

分类目录现有的内容包括:

资料集:

风云卫星4号、5号、6号、7号卫星观测资料集;全球大气研究计划第一期全球试验(FGGE)观测资料集;海洋卫星观测资料集;平流层气溶胶和气体试验(SAGE)卫星观测资料集;OSTA 1号卫星观测资料集;地球静止卫星3号卫星观测资料集;气象地球静止卫星观测资料集;美国国家海洋大气管理局(NOAA)各种观测资料;美国国家气象中心分析资料。

参数:

云量、臭氧、辐射平衡、平流层气溶胶、风速、天气变量、冰盖、降水、海冰密集度、海面温度。

3. 详细清单

此清单为如何找到特定的资料而提供详细的信息。当用户提出有关资料类型、时段和地理位置等方面专门要求时,资料读取子系统就会利用这一信息,自动把从PCDBMS磁带库中的资料送到磁盘上的某一资料集中去。采用这种方法,一个资料子集的读取过程是不同用户发生关系的。用户也可以直接利用详细清单,得到有关某一资料集的资料量和时段的详细信息,也可用一个清单图象实用程序获得关于资料时段的图象信息。

清单只包括PCDBMS或与其密切相关

的其它系统的现有资料信息。也可以通过关键词方法来使用清单。

这份清单目前可用于:

- (1) 列出现有全部气候参数;
 - (2) 列出现有的资料类型和相应气候参数;
 - (3) 列出各可供选择的资料逻辑单位(如图、廓线等)的信息;
 - (4) 列出现有资料类型(如占用磁带数量和时段等)的信息摘要;
 - (5) 列出某一资料类型每盘磁带的信息;
 - (6) 列出某一磁带每份文件的信息;
 - (7) 列出各盘磁带在此系统清单上的收录史;
 - (8) 有一个与分类目录中所采用的打印实用服务程序相同的实用服务程序;
 - (9) 利用作为时间函数的资料率点线图象或通过显示在世界地图上的资料覆盖情况(见图1)对清单上的内容进行图象显示。
- 清单现可提供自下列试验中选择出的资料集:
- (1) 风云4号卫星后向散射紫外线的测量;
 - (2) 风云7号卫星太阳后向散射紫外线的测量;
 - (3) 风云7号卫星臭氧总量测绘光谱仪的测量;
 - (4) 风云7号温湿红外辐射仪的测量;
 - (5) 风云7号地球辐射平衡的测量;
 - (6) 风云7号平流层气溶胶测量(第二期);
 - (7) 平流层气溶胶和气体试验;
 - (8) 国家气象中心分析;
 - (9) 全球月平均地面站气候资料;
 - (10) 全球大气研究计划第一期全球试验;
 - (11) 美国国家海洋大气局热平衡研究;
 - (12) 风云5号卫星电动扫描微波辐射仪测量。

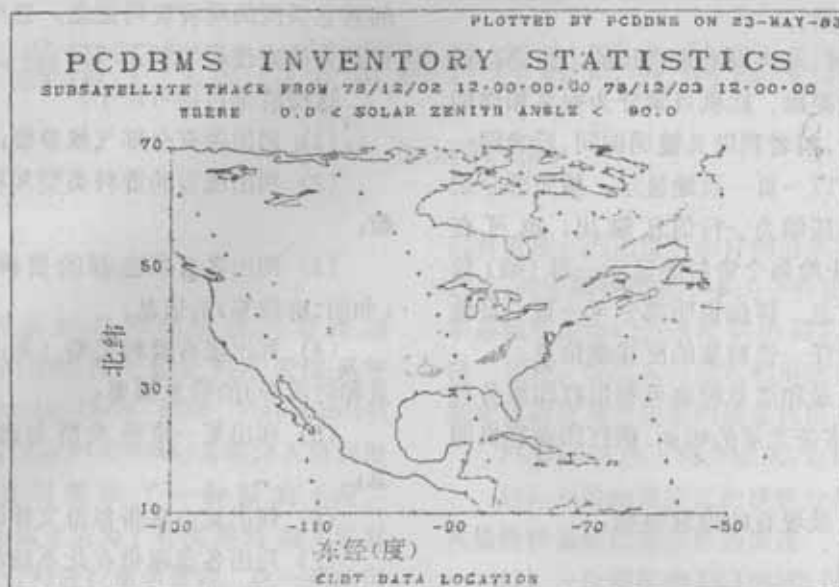


图1 清单的资料覆盖情况

4. 资料读取

资料读取子系统主要用于根据时间和地理位置选择资料子集。资料集存放在PCDBMS的磁带库中。这些磁带是以资料集的编制者当初所提供的原格式存放的。用户可利用资料读取子系统任意选择所需任一资料集的子集，并得到以某种格式提供的输出产品。

用户可以把资料子集复制到另一磁带或磁盘上，也可只要一份打印清单。输入资料可以是PCDBMS磁带库所提供的标准磁带，这种磁带是使用这个资料读取实用程序建立；也可以是一个早先存在磁盘上的资料子集。

本系统最有用途的输出方式选择之一是可以把资料子集转换为一种有描述信息的独立资料，随后通过PCDBMS的数据组织和图表显示分系统而得到利用。这种格式的独立资料称之为气候资料文件(CDF)。

即使在用户不了解详细信息的情况下，资料读取子系统也可用PCDBMS清单自动找寻资料。用户无须花很大气力去了解磁带、文件、资料格式等方面的详细情况。清单和

专门软件负责管理所有用于找寻、读出和解释各类资料的详细信息。在这种情况下，资料读取子系统只提供PCDBMS库中现存的资料集。目前所能读取的资料范围与上一节(即3.0节)所列的内容相同。(然而，要提请注意的是，任何资料集都能转换为CDF格式，并能用于数据组织和图表显示子系统。假如一个资料集不能用于数据组织和图表显示子系统，它就无须存入PCDBMS库)。

5. 数据组织

资料加工子系统为用户提供了一组实用服务程序，它们可用于组织管理一切CDF格式的资料集，在使用图表显示子系统之前，或者是为下步工作之用而把资料集复制到另一磁带上并送到其它系统之前，可利用这些实用服务程序对资料集进行加工处理，使其便于用户使用。所有这些实用服务程序均可输入、输出CDF文件，这样可把它们组配起来完成各种多功能处理。

现可用的实用程序如下：

(1) 制作CDF子集服务程序——把资料分为更小的子集。PCDBMS提供的最为有用的实用程序之一就是制作CDF子集服

务程序。它能按用户选定某些 CDF 字段或在 CDF 某些字段组合基础上,经编辑加工后输出用户所需的资料。

(2) 合并 CDF 文件服务程序——把两个资料子集合并在一起。该服务程序的一个重要用法是建立映衬型资料集以供对比分析之用。

(3) CDF 打印服务程序——可打印任何资料子集。要想改变打印产品的格式只需通过选定 CDF 字段即可。

(4) 网格资料集转换服务程序——把网格资料还原。把网格资料转换为通用格式,例如,用于同非网格资料比较。

(5) 对 CDF 字段进行各种运算的服务程序。根据用户规定的参加运算的字段和运算要求,产生运算结果的输出产品。

(6) 对 CDF 字段进行统计学运算。利用这一程序,可求出 CDF 允许的任一时段的平均值和方差。

6. 图象显示

图象显示子系统提供各种图象显示实用程序,可用于以 CDF 格式存放的任何资料,而不管它们是否保存在 PCDBMS 库。所有经这一子系统处理的资料均成为以 CDF 格式存放和使用的独立资料。这一子系统可用以对资料组进行查阅,或为下一步分析或者出版之需提供打印输出。可提供二维和三维的图象输出。

二维图象实用程序可用于提供标有一个以上变量的标准化 x - y 点绘图。还可提供分布点图、矢量图和直方图。这些图用直角坐标系或极坐标系,坐标或以线性划分或以对数划分。绘图前,还可预先对资料集进行线性、多项式或样条曲线拟合处理。有几种统计学方法可供采用,如数据平滑和求资料的平均值和方差等等。输出的字形有好多种,其中包括可供出版的字形。图 2 和图 3 为二维图象输出的实例。

利用三维图象实用程序可得到等值线

图、立体等值面图和假彩色图象。在这些点绘图上可以添上各种地图投影的并有海岸线的格点,这些实用程序还可提供许多选择项,以根据用户对格式、质量和内容的要求输出相应的产品。

另外,还有一个图象后处理(编排)实用程序,它可以把上述各过程的输出转送到各种硬拷贝装置去。根据用户选择提供下列服务,如提供适合于出版要求的输出,或把几份图合印在一页打印纸上等等。图 4 和图 5 是典型的三维图象输出产品。

7. PCDBMS 的实现和完善

PCDBMS 已在马里兰州的格林贝尔特的 VAX-11/780 计算机系统上实现。它利用可传输应用执行系统(TAE)作为用户接口。TAE 也是在戈达德宇宙飞行中心研制出来的,这一系统为编制程序和其它操作提供了方便。它可为新用户提供服务菜单驱动式接口,而为富有经验的用户提供命令语言接口。

PCDBMS 的设计考虑到发挥最大效益,因而采用各种新技术,并将这些技术综合应用构成一个功能很强的数据库管理系统。除 TAE 外,PCDBMS 还使用一个商用数据库管理系统(ORACLE)、图象处理程序包(TEMPLATE)、统计程序包(IMSL/PROTRAN)和网络通讯软件(DECNET)。通过适当地配置这一系统,保证能从它建立开始就不断采用新技术。由于该系统便于扩充和采用新技术,PCDBMS 不愧为一个集当代数据库管理技术的大成者,而不必浪费时间和财力来研制专用硬件和软件。

8. 结 束 语

NASA 试验性气候数据库管理系统为气候研究人员提供了多种的资料管理和分析软件系统。这些软件系统组成了一个完整系统,使用很方便,且适应性极强。使用 PCDBMS 无须熟悉资料格式和编程语言,也无须为某项应用专门编制程序。

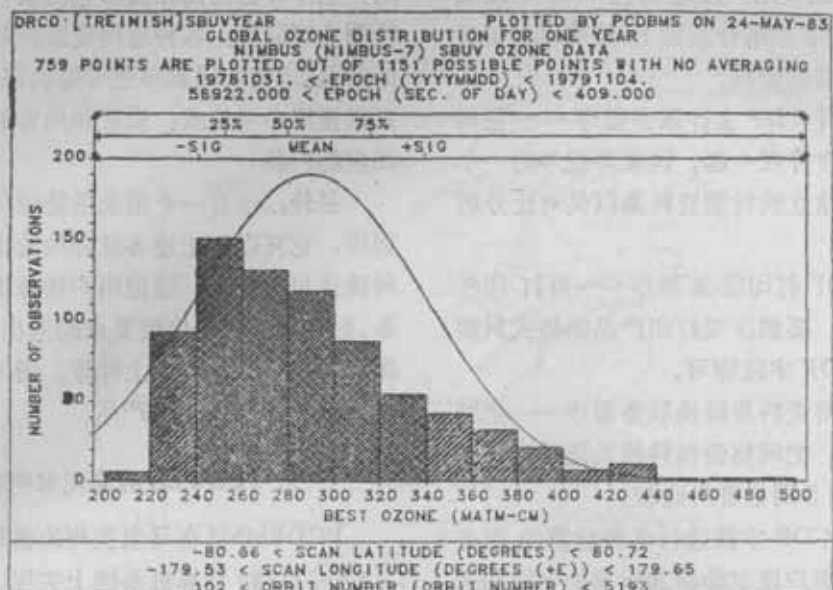


图 2 样本统计结果的直方图

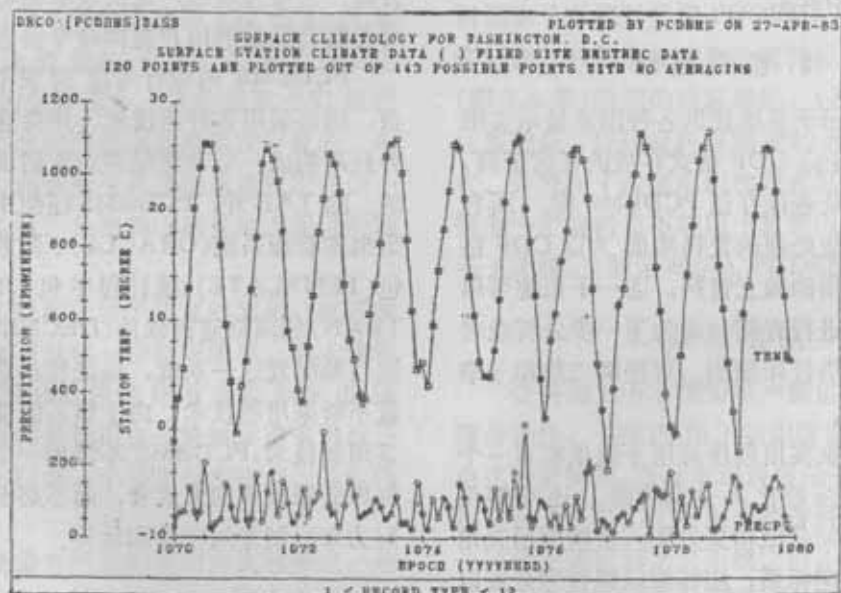


图 3 样条曲线拟合的散点图

目前, PCDBMS 只为戈达德宇宙飞行中心的气候研究人员提供帮助。根据这些用户的需求对该系统不断改进。同时, 增加处理用户要求的新型资料集的功能以及某些通

用功能。另外, 还致力于研究和开发新的数据库管理技术, 并将成果应用于 PCDBMS 的改进之中。

PCDBMS 表明: 现代数据库管理技术

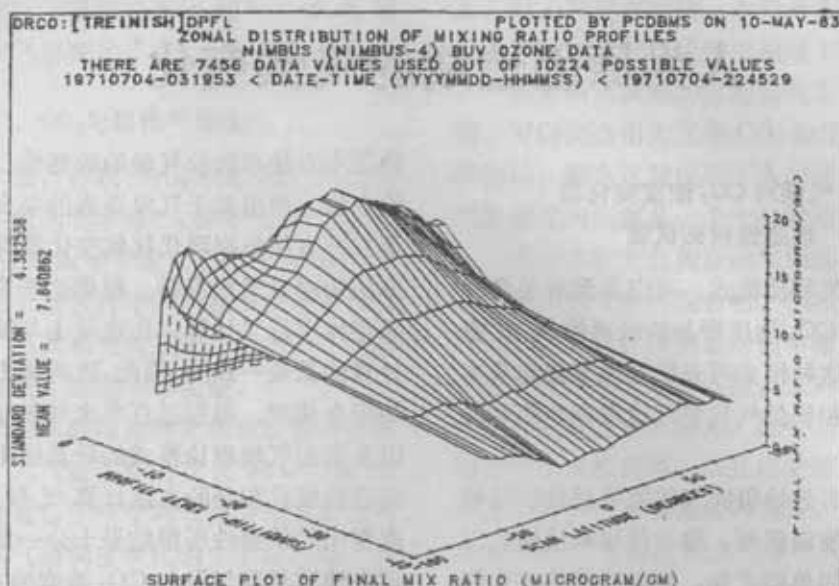


图 4 三维立体等值面图

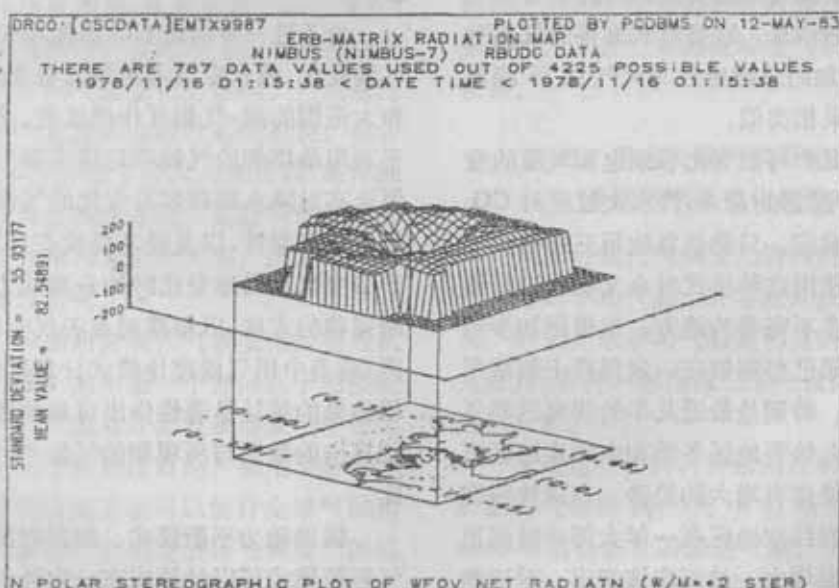


图 5 附有地图的立体等值面图

正用于科学研究，而科学研究所用的资料集的特点是容量大、资料来源多种多样，其意义不仅在于当前的各种应用，这一系统可作今后更大规模的数据库管理系统的样板，

但也还只是更大型系统软件的一个部分，而这种大型系统软件将能用于多学科目的。

李丽译自《第八次大气科学概率论和统计学国际会议》高良成校