

# AVHRR1B数据集的生成、格式及使用方法

范天锡 郭俊如

(卫星气象中心)

## 摘要

本文介绍了国家气象局卫星气象中心所生成的AVHRR(改进的甚高分辨辐射计)1B数据集的格式和使用方法，并简介其生成的原理和方法。

### 一、引言

美国TIROS-N系列极轨气象卫星实时发送数字化的HRPT(高分辨图象传输)资料，为了对这些资料作高精度的定量处理，首先要进行预处理。对HRPT资料进行分类编辑、质量检验、确定定标系数、定位数据计算及格式变换等工作，生成由美国地球卫星和信息服务局所取名的“1B”格式的数据集。

AVHRR 1B数据集中包含了时间码、数据质量检验码、定标系数、扫描点的地理经纬度和太阳高度角等数据。用1B数据集来进行各种定量处理是很方便的。因此，它是各种资料处理的数据来源。把它做成计算机兼容磁带(CCT)，用于卫星资料的原始存档，并提供给广大用户使用。

### 二、AVHRR 1B数据集生成方法

预处理的主要工作是质量检验、定标和定位。

#### 1. 质量检验

卫星资料在星上编排处理、发送、传输和地面接收过程中，都可能出现错误。因此，对卫星原始资料首先要做质量检验。检验的项目有时间码、主付帧计数、卫星标记、付帧误码率、奇偶校验等。质量检验后做出质量标记(具体格式及内容后面介绍)。

#### 2. 定标处理

包括斜坡校准处理和红外通道的定标计算。

(1) 斜坡标准处理 斜坡计数是用来反映每个通道的计数值与测量电压之间的关

系，斜坡处理就是检验这种关系是否为线性。AVHRR每个通道的斜坡计数，是按扫描线顺序依次放置的1024个阶梯电压的计数值。因此，它应与扫描线序号呈线性关系。一般取中间一段数据来进行处理，每条轨道只作一次。设扫描线序号为x，斜坡计数值为y，斜坡系数的斜率为a，截距为b，则：

$$a = s_{xy} / s_{xx}$$

$$b = \bar{y} - a\bar{x}$$

其中，

$$s_{xx} = \frac{1}{n} \sum_{i=1}^n (x_i^2 - \bar{x}^2),$$

$$s_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i y_i - \bar{x}\bar{y}),$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i,$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i.$$

n是所取的数据数。

线性关系的检验用F试验值来进行。

$$F = (n-2) R^2_{xy} / (1 - R^2_{xy})$$

其中 $R_{xy}$ 为相关系数。

$$R_{xy} = s_{xy} / (s_{xx} s_{yy})^{1/2}$$

$$s_{yy} = \frac{1}{n} \sum_{i=1}^n (y_i^2 - \bar{y}^2)$$

当F值小于某临界值时为线性，否则为非线性。由于目前星上仪器相当稳定，因此斜坡处理结果一般存放在1B数据集的头记录中，并打印出来，供资料处理人员检查，如发现问题再作处理。

(2) 定标系数的计算 所谓定标就是将仪器观测的计数值换算成物理量。可见光和近红外通道 ( $ch_1, ch_2$ ) 观测值可以换算成反射率值, 红外通道 ( $ch_3-ch_5$ ) 可以换算成辐射率值。红外通道在卫星运行过程中, 性能会有所变化, 且辐射的测量精度要求也高, 因此, 要不断更新定标系数。在预处理中, 将 5 条扫描线作为一个定标周期, 计算一次定标系数。对于每条轨道的数据来讲, 所有开始阶段的扫描线都使用第一次计算出来的定标系数。以后的每条扫描线使用前一定标周期做出的定标系数。但是, 若连续 3 个定标周期不能做出定标系数, 则将定标系数赋零。这时则不能对数据做定标处理。

定标公式为

$$N = GC + I$$

其中, G 为斜率, I 为截距, C 是计数值。对  $ch_1, ch_2$ , N 是反射率, 单位为%; 对于  $ch_3-ch_5$ , N 是辐射率, 单位是  $mW/(m^2 \cdot sr \cdot cm^{-1})$

### 3. 定位

所谓定位就是计算出卫星瞬时视场所扫描的地面观测点的地理经纬度和太阳高度角, 它们由观测时间的卫星位置、卫星姿态

和扫描角决定。首先, 根据卫星轨道根数和扫描点的观测时间, 计算出该时间的轨道参数(即卫星位置), 再由卫星位置、姿态和扫描角计算出该扫描点的地理经纬度。处理过程中, 从每条扫描线的 2048 点中选 51 个点进行定位计算, 即从第 25 点开始, 每 40 点选一点。而其它扫描点的定位数据则由内插产生。为了提高定位精度, 还可利用已知地理位置的地标, 对定位数据进行修正。这一工作称为地标定位。

### 三、1B 数据集的格式及内容

经过以上处理后, 再经格式变换, 生成固定的 AVHRR 1B 数据集格式。由于 AVHRR 一条扫描线记录较长, 在做成 CCT 时, 分其为两个记录, 每个记录长度为 7400 字节。它的第一个记录为数据集名称, 记录 2 为 1B 数据集头记录, 存放此数据集的统计信息; 记录 3 备用; 从记录 4 开始为数据记录。

#### 1. 数据集名称

数据集名称由 122 个 ASCII 码组成, 实际只有 44 个字符, 由一串有关此数据集的信息组成。另外还有 8 个字符的时间码。具体格式及内容见表 1。其中,

表 1 记录 1 在格式及内容

字节数	30	44	14	8	6	16	4	7218
内容	‘□’	数据集名称	‘TALLALLALL□ALL’	时间码	‘Y11111’	‘.’	‘□’	备份

44 个字符的分配

字符数	4	5	3	7	6	6	9	4
内容	NSS.	数据类型	卫星标记	年·日	开始时间	结束时间	轨道序号	数据来源

注: ‘ ’ 代表要充的内容, □—空格

数据类型: HRPT.;

卫星标记: TN. = TIROS-N, NA. = NOAA-A, ..., NG. = NOAA-G;

年·日: 如 “D85104.”, 其中 “D” 为标识符, “85” 即 1985 年, “104” 表示年初开

始计起, 一年中的日计数;

开始时间: 如 “S1355.”, “S” 为标识符, “1355” 表示世界时 13 时 55 分;

结束时间: 例如 “E1410.”, 其中 “E” 为标识符, “1410” 表示世界时 14 时 10 分;

轨道序号：如“B0016464.”，“B”为标识符，“00164”表示开始时间的轨道序号，“64”为结束时间轨道序号的十位和个位数；

数据来源：表示数据由何处得到，如北

京站接收的资料可自定标符“TP”；

时间码：第一个字节是空格，第2—5字节是开始时间（时和分），第6—8字节是开始至结束时间的时间间隔（分）。

### 2.1B数据集头记录格式（见表2）

表2 记录2的格式及内容

字节数	1	1	6	2	6	7	1	2	6	2	1	7366	
内 容	卫星识别记	数据类型	开始时间	扫描线数	结束时间	轨道区 域序号	及识 别	斜坡校准标记	数据脱漏数	质量信息 (DASS)	定标参数集识别码 (DASS)	状态信息 (DASS)	备 份

具体内容如下：

卫星识别标记：0=备份，1=TIROS-N，2=NOAA-A，……，8=NOAA-G。

数据类型：前4位为数据类型码，3=HRPT；后4位为TIROS-N信息处理机(TIP)数据来源，目前不考虑TIP数据，赋零。

开始时间：是此数据集第一条扫描线的卫星时间码（世界时）。前2个字节的前7位为年（例85=1985年），后9位为一年中的日计数；后4个字节的前5位赋零，后27位为一日中的毫秒计数。

扫描线数：此数据集的扫描线总数，不包括脱漏的扫描线。

结束时间：是此数据集最后一条扫描线的卫星时间码（世界时），格式同开始时间。

轨道序号：ASCII码。如“0016465”中的“00164”表示数据集开始时的卫星轨道序号，“65”为数据集结束时的卫星轨道序号的个位和十位。

斜坡校准标记：第4—8位分别为ch<sub>2</sub>—ch<sub>5</sub>的斜坡校准判识码，0=线性，1=非线性。

数据脱漏计数：此数据集中脱漏数据的扫描线数。

DASS（数据接收和控制子系统）质量信息：第1—2字节为非比特同步失锁或帧同步失锁错的扫描线数；第3—4字节为具有TIP奇偶错的扫描线数；第5—6字节为辅助同步码错的扫描线数。

定标参数数据集识别码：定标计算时，选择参数数据集用的标识码，用户一般不用。

DACS状态：第5位赋“1”，表示接收卫星飞行数据，其它位赋零。

### 3.1B数据记录

1条扫描线=2个记录。

注意在1B数据记录与头记录之间有一个空记录为记录3（表略），字节长是7400。

表3、表4分别为1B数据集的记录4和记录5。

表3 记录4的格式与内容

字节数	2	6	4	40	1	51	204	104	6952
内 容	扫描线序号	时间码	质量指示码	定标系数	赋地理经纬度和太阳高度角的扫描点数	太阳天顶角	地理经度	遥测数据	AVHRR数据

注：51个扫描点的选取：每条线2048点，每40点作一个取样。从第25点开始，序号为25,65,……,2025。共51个点。

表4 记录5的格式与内容

字节数	6704		696
内 容	AVHRR数据		备份

表中内容具体说明如下：

扫描线序号：从1—n。

时间码：该扫描线的开始时间（世界时），格式与头记录中的开始时间相同。

质量指示码：格式如表5所示。发生定

表 5 质量指示码格式

比特序号	1	2	3	4	5	6	7	8	1	2	3	4—8	1—5	6—8	1—5	6	7—8
内 容	严 重 错	时 序 错	本 帧 以 前 的 数 据 脱 漏	数 据 脱 漏	发 生 重 步	定 标 系 数	没 有 定 位 数 据	升 降 轨 标 记“1”为 升“0”为降	备 备 份	比 特 同 步 失 锁	备 份	帧 同 步 失 锁	备 份	1—5个 TIP付帧 奇偶错	备 份	误 码 率 大 于 0	备 份

义中的质量问题时赋“1”，没有问题时赋零。

其中，时序错(亦无定位数据)和没有足够定标数据为严重错误，此时该付帧数据不能使用。

在预处理过程中，质量检验有多种用途，例如，用以确定每一条轨道的数据中的第一条好扫描线；发现扫描线脱漏和确定扫描线序号；当付帧的时间和顺序码有错误时，则该帧数据不参加定标系数计算，对此付帧数据不定位等。

**定标系数：5个通道的斜率和截距的长度和格式如表6所示。**

表 6 5个通道的斜率和截距对应的字节数

Ch <sub>1</sub>	Ch <sub>2</sub>	Ch <sub>3</sub>	Ch <sub>4</sub>	Ch <sub>5</sub>
4 4	4 4	4 4	4 4	4 4
斜 率	截 率	斜 率	截 率	斜 率
距 离	距 离	距 离	距 离	距 离

有太阳天顶角和地理经纬度值的扫描点数：最大值为51。

**太阳天顶角：**51个扫描点的太阳天顶角数据，单位为0.5度。每个数据占1个字节，其中前7位为整数，后一位为小数。

**地理位置：**51个扫描点的地理经纬度，单位为1/128度。每个数据占2个字节，按扫描点1纬度，扫描点1经度……，扫描点51经度排列。

**遥测数据：**复制HRPT付帧中的遥测数据，用户一般不用。

**AVHRR数据：**按扫描点顺序存放，即按扫描点1(ch<sub>1</sub>—ch<sub>5</sub>)，……，扫描点2048(ch<sub>1</sub>—ch<sub>5</sub>)排列，共10240个数据。每个数据10位，每3个数据放在1个字(4个字节)中，每个字前2位充零，共占用3414字(13656字节)。

#### 四、1B数据集的使用方法

##### 1. 计算机兼容磁带(CCT)

AVHRR 1B数据集CCT，记录密度可用6250 BPI或1600 BPI。如果采用记录密度为1600 BPI，每条轨道的接收时间为

12—15分钟，需两盘磁带。第一盘记2500条线，第二盘由轨道接收时间长短确定。记录采用无标号记录形式(NL)。

##### 2. 读取方法

可用FORTRAN语言写的程序读取和进行数据处理，读取记录时，可使用READ语句或使用专门的FFIO(FORTRAN快速输入输出)实用程序。

3. 第一个记录为数据集名称，当磁带内容为已知时，一般可以跳过去。第二个记录为头记录，可以提取用户所需要的参数。第三个记录无具体内容，空读过去。

4. 从第4个记录起，每次读2个记录(一条扫描线数据)放入一个数组中，再按其格式提取所需信息。

5. 定标系数、太阳天顶角和地理经纬度都是整型数，使用时要化成实型数。转换方法为：

斜率值须除以 $2^{30}$ ，截距值须除以 $2^{22}$ ；

太阳天顶角须除以2；

地理经纬度须除以128。

6. 辐射计数值在1B数据集中是经过装配的，使用时首先要拆开，把每个计数值放入16位的半字中，这一工作耗时较多，最好只对所需要的数据作拆卸。

#### 欢迎订阅《贵州气象》

贵州省气象局主办的《贵州气象》是目前国内唯一公开发行的省级气象科技刊物。本刊将给读者提供应用气象方面的科技成果，论文、译文、技术经验，介绍计算机应用成果，并举办天气学、气候学、气象基础理论等讲座。不定期开设介绍中学地理课本中有关天气、气候知识的栏目。同时还酌量登载短小的气象小说、散文、诗歌等。

读者对象是广大气象台站、农村科技工作者，大学有关气象专业的师生、中学地理教师及广大气象爱好者。

《贵州气象》16开本48页，逢双月20日出版，定价0.40元，全年2.40元。可破月订阅。凡订阅本刊者，请直接将订阅款汇至：贵阳市《贵州气象》编辑部收。