

## 补充说明：ADC值校正局限性

- 理想情况下，ADC值校正需要考虑探测器在事件发生时（具体到那一秒）的温度和电压信息。
- 由于卫星公司在卫星数据下传方面的特殊机制，同一时间点的科学数据和housekeeping数据可能不会在同一下传文件中，甚至可能在不同日期下传。例如，一个事件对应的温度和电压数据可能出现在几天后的文件中。
- 如果遍历所有housekeeping数据文件去查找匹配时间戳的温度偏压数据，效率太低。
- 目前程序是按照每个下传文件为单位去处理数据（每个 .dat 文件作为一个处理单元），无法对时间戳进行跨文件搜索，导致无法直接获取事件发生时的温度和电压信息。
- 由于无法实时匹配温度和电压，`calibrate_adc` 方法当前仅返回简单的ADC值（对于通道3返回NaN，其他通道直接返回输入值，未进行温度-偏压校正）。

## 按天切分数据以优化查找

- 当前程序按照下传文件（.dat）处理数据，每个文件可能包含多天的科学数据和housekeeping数据。
- 由于下传机制的随机性，同一时间点（UTC）的科学数据和housekeeping数据可能分散在不同文件中。
- 如果将数据按天切分并存储，文件名可以直接反映数据的日期范围，从而快速定位对应时间点的housekeeping数据。
- 将科学数据（EVENTS、SPECTRA）和housekeeping数据（HOUSEKEEPING）按天（以UTC时间为基准）切分，存储为独立的FITS文件。
- 文件名中嵌入日期信息（如 YYYYMMDD），这样每次搜索就只需要在对应那一天的一个文件里去搜索，极大减小了搜索量。
- 在一天的文件里，使用二分查找与事件UTC最近的housekeeping记录，获取 SIPM\_TEMP 和 VMON 值。找到了之后就直接进行标准。

```
def calibrate_adc(self, adc_value, channel, temp, bias):  
    if channel == 3:  
        return np.nan  
    adc_calibrated = self.adc_mapping.adcmap(channel, adc_value, temp, bias)  
    return adc_calibrated
```

## 使用时序数据库Apache IoTDB

- IoTDB以时间戳为核心组织数据，支持高效的时间范围查询、最近点查询和插值操作，非常适合用于处理科学数据和housekeeping数据的时间戳匹配。

- IoTDB的 NEAREST 查询功能可直接返回与给定时间戳（如事件UTC）最接近的housekeeping记录，取代按天FITS文件中的线性搜索或二分查找。
- 将所有科学数据和housekeeping数据存储在IoTDB中，按时间戳和通道组织，无需按文件切分。
- IoTDB仅作为中间存储和查询工具。

## 补充说明：添加高通量模式数据到最终FITS文件

- 当前最终FITS文件（ G11\_evt\_YYYYMMDD.fits ）仅包含特征量模式（EVENTS）数据，即单个光子事件的精确时间（ TIME ）和能量通道（ PI ）。
- 探测器经常在特征量模式和高通量模式（SPECTRA）之间切换，导致时间序列上存在空缺（即高通量模式时间段无事件数据）。
- 这些空缺影响时间序列分析（如光变曲线生成），因为无法完整覆盖观测时间。
- 高通量模式以固定时间间隔（通常每秒）记录能谱数据，包括短周期能谱（ SHORT\_SPECTRA ， 20个时间bin×8个能量通道）和长周期能谱（ LONG\_SPECTRA ， 82个能量通道）。
- 能谱数据基于ADC值统计到固定区间（ adc\_bounds ），但ADC值需通过温度和偏压校正转换为能量值。由于温度和偏压每秒变化，能量边界（ ENERGY\_MIN 和 ENERGY\_MAX ）也需每秒重新计算。

- **当前FITS结构：**

- **主HDU (PRIMARY)：**空。
- **扩展HDU**（不再存储特征量模式以外的信息）：
  - **HDU EBOUNDS：**
    - 列：
      - Channel （整数）：能量通道编号（1到N）。
      - E\_MIN （浮点数，单位：keV）：通道能量下限。
      - E\_MAX （浮点数，单位：keV）：通道能量上限。
    - 用途：定义能量通道的边界。
  - **HDU GTI**（Good Time Intervals）：
    - 列：
      - START （浮点数，单位：秒）：观测开始时间（相对于2018-01-01T00:00:00）。
      - STOP （浮点数，单位：秒）：观测结束时间。

- 用途：观测计划指定的有效观测时间区间，用于时间筛选。对于11B来说，由于24小时开机，实际上有效观测时间就是从文件开始到文件结束的全部时间。
- HDU **EVENTS{channel}**（每个通道一个，如 EVENTS0、EVENTS1、EVENTS2）：
  - 列：
    - TIME（浮点数，单位：秒）：事件时间（相对于2018-01-01 T 00:00:00）。
    - PI（整数）：能量通道编号（基于 EBOUNDS 的能量分区的编号）。
    - DEAD\_TIME（字节，单位：微秒）：死时间（固定为4微秒）。
    - EVT\_TYPE（字节）：事件类型，用于判断事件的能量是否超出了 EBOUNDS 定义的所有能量范围（0：能量低于最低通道；1：正常事件；2：能量高于最高通道）。
  - 用途：提供光子事件的时间和能量通道信息，适合能谱和光变曲线分析。
- 头信息（Header）：
  - 包含元数据：DATE（文件创建时间）、FILE\_VER（版本，如 PRELIMINARY）、MISSION（GRID）、CUBESAT（GRID-11B）、DATE\_OBS（观测开始时间）、DATE\_END（观测结束时间）、DATE\_REF（参考时间2018-01-01 T 00:00:00）。

- 包含 EBOUNDS（静态的能量边界，适用于特征量模式）、GTI（根据观测计划拟定的有效观测时间区间）和 EVENTS{channel}（4个通道的事例数据）。
- 缺乏高通量模式数据。

## 新的FITS文件结构

为满足上述需求，应该修改最终FITS文件（SAT-A\_evt\_YYYYMMDD.fits）的结构，包含以下扩展HDU：

### 主HDU（PRIMARY）

- 内容：空。

### 扩展HDU 1：模式切换时间表（MODES）

- 替换原GTI扩展：
- 新扩展 MODES 记录特征量模式（EVENTS）和高通量模式（SPECTRA）的切换时间段，描述探测器在每个时间段的工作模式。
- 列：
  - START（浮点数，单位：秒）：时间段开始时间（相对于2018-01-01T00:00:00）。
  - STOP（浮点数，单位：秒）：时间段结束时间。

- **MODE** (字符串): 工作模式 ( **EVENTS** 表示特征量模式, **SPECTRA** 表示高通量模式)。

## 扩展HDU 2: **EBOUNDS** (特征量模式的能量边界)

- **内容:** 与当前一致, 定义特征量模式的静态能量通道边界。
- **列:**
  - **Channel** (整数): 能量通道编号 (1到N)。
  - **E\_MIN** (浮点数, 单位: keV): 通道能量下限。
  - **E\_MAX** (浮点数, 单位: keV): 通道能量上限。

## 扩展HDU 3: **EVENTS{channel}** (特征量模式的事例数据)

- **内容:** 与当前一致, 按通道存储特征量模式的事件数据。
- **列:**
  - **TIME** (浮点数, 单位: 秒): 事件时间 (相对于2018-01-01 T 00:00:00)。
  - **PI** (整数): 能量通道编号 (基于 **EBOUNDS**)。
  - **DEAD\_TIME** (字节, 单位: 微秒): 死时间 (固定为4微秒)。
  - **EVT\_TYPE** (字节): 事件类型 (0: 低于最低通道; 1: 正常事件; 2: 高于最高通道)。

## 扩展HDU 4: **SPECTRA{channel}** (高通量模式能谱数据)

- **内容:** 新增扩展, 存储高通量模式的能谱数据和每秒的动态能量边界。
- **列:**
  - **TIME** (浮点数, 单位: 秒): 能谱数据的采样时间 (每秒一个时间点, 相对于2018-01-01T00:00:00)。
  - **SHORT\_SPECTRA** (20×8数组, 整数): 短周期能谱 (20个时间bin × 8个能量通道)。
  - **LONG\_SPECTRA** (82元素数组, 整数): 长周期能谱 (82个能量通道)。
  - **LONG\_EBOUNDS\_MIN** (82元素数组, 浮点数, 单位: keV): 长周期能谱的每秒动态能量下限 (对应 **LONG\_SPECTRA** 的82个通道)。
  - **LONG\_EBOUNDS\_MAX** (82元素数组, 浮点数, 单位: keV): 长周期能谱的每秒动态能量上限 (对应 **LONG\_SPECTRA** 的82个通道)。
  - **SHORT\_EBOUNDS\_MIN** (8元素数组, 浮点数, 单位: keV): 短周期能谱的每秒动态能量下限 (对应 **SHORT\_SPECTRA** 的8个通道)。
  - **SHORT\_EBOUNDS\_MAX** (8元素数组, 浮点数, 单位: keV): 短周期能谱的每秒动态能量上限 (对应 **SHORT\_SPECTRA** 的8个通道)。
  - **说明:**
    - 动态 **EBOUNDS\_MIN** 和 **EBOUNDS\_MAX** 支持每秒变化的能量校正, 反映温度和偏压的影响。
- **头信息 (Header):**
  - 包含元数据: **DATE** (文件创建时间)、**FILE\_VER** (版本, 如 **PRELIMINARY**)、**MISSION** (**GRID**)、**CUBESAT** (**GRID-11B**)、**DATE\_OBS** (观测开始时

间)、 DATE\_END (观测结束时间)、 DATE\_REF (参考时间2018-01-01 T 00:00:00)。

## 时间覆盖

- **EVENTS{channel}和SPECTRA{channel}的 TIME 字段与MODES**扩展的 START 和 STOP 对应，确保整个文件的观测时间（从文件开始到结束）被完全覆盖。
- 例如， MODES 扩展可能包含：

```
START: 1734567890.0, STOP: 1734567990.0, MODE: EVENTS  
START: 1734567990.0, STOP: 1734568090.0, MODE: SPECTRA
```

- **EVENTS**数据覆盖 MODE=EVENTS 的时间段。
- **SPECTRA**数据覆盖 MODE=SPECTRA 的时间段。
- 两者拼接后形成连续时间序列，无空缺。