

数据论文

莱曼阿尔法太阳耀斑事件标准数据集

卢磊 封莉 黎辉

(中国科学院紫金山天文台 南京 210023)

摘要 太阳耀斑是太阳大气中最强烈的爆发现象,能够释放大量的能量并产生各种波长的电磁辐射。研究太阳耀斑对于理解太阳活动、空间天气预报以及保护地球空间环境至关重要。本数据集基于夸父一号 (ASO-S) 卫星搭载的全日面成像仪 (Solar Disk Imager, SDI) 在莱曼阿尔法波段 (121.6 ± 7.5) nm 采集的全日面图像数据,通过一套自主研发的太阳耀斑自动识别与关键参数提取算法,系统记录了 2024 年莱曼阿尔法太阳耀斑事件。该算法可有效避免宇宙线、粒子暴等事件的干扰,能够对不同强度级别的耀斑进行识别,并能对日面上同时发生的多个耀斑进行分别识别与追踪。本数据集收录了耀斑的起止时间、持续时间、发生位置、显著性等关键参数,包含耀斑识别过程记录文档、耀斑事件列表、耀斑峰值时刻快视图像和耀斑区域电影动画等数据。该数据集可为太阳物理学研究、空间天气预报以及相关领域提供重要的科学数据支持。

关键词 太阳耀斑, 莱曼阿尔法, 空间天气, 自动识别

中图分类号 P352

数据库(集)基本信息

数据库(集)名称	莱曼阿尔法太阳耀斑事件标准数据集
数据作者	卢磊
数据通信作者	卢磊 (leilu@pmo.ac.cn)
数据时间范围	2024年1月1日至2024年12月31日
数据类型	莱曼阿尔法太阳全日面观测数据
数据量	约2.3 TByte
数据格式	*.fits, *.txt, *.png, *.mp4
数据服务系统网址	DOI: https://doi.org/10.12176/01.01.00116-V01 CSTR: https://cstr.cn/14804.11.01.01.00116-V01
基金项目	国家重点研发计划项目 (2022YFF0711402, 2022YFF0503003), 中国科学院战略性先导科技专项 (XDB0560000) 和国家自然科学基金重点项目 (12233012)
数据库(集)组成	本数据集包含2024年全年由夸父一号卫星莱曼阿尔法成像仪采集的太阳耀斑事件数据。共记录了1937个太阳耀斑事件,包括耀斑识别过程记录文档、耀斑事件列表(记录耀斑的位置、时间及耀斑显著性等信息)、耀斑峰值时刻快视图像(显示耀斑峰值时刻区域图像与光变)以及耀斑区域电影动画。数据采用FITS, TXT, PNG和MP4等标准格式存储

2025-05-15 收到原稿, 2025-11-17 收到修定稿

©The Author(s) 2026. This is an open access article under the CC-BY 4.0 License
(<https://creativecommons.org/licenses/by/4.0/>)

Standard Dataset of Lyman-Alpha Solar Flare Events

LU Lei FENG Li LI Hui

(Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210023)

Abstract Solar flares are the most intense eruptive phenomena in the solar atmosphere, releasing large amounts of energy and producing electromagnetic radiation across various wavelengths. Research on solar flares is crucial for understanding solar activity, space weather forecasting, and protecting the Earth's space environment. Based on the full-disk solar images in the Lyman-alpha waveband (121.6 ± 7.5) nm collected by the Solar Disk Imager (SDI) carried by the China's ASO-S satellite, this dataset systematically records solar flare events observed in the Lyman-alpha band throughout 2024 using an independently developed automatic solar flare identification and key parameter extraction algorithm. This algorithm effectively avoids interference from cosmic rays and particle storms, identifies flares of different intensity levels, and can separately identify and track multiple flares occurring simultaneously on the solar disk. The dataset includes key parameters such as flare start and end times, duration, location, and significance, and contains data products including flare identification process documentation, flare event lists, quick-view images of flares at peak moments, and movies of flare regions. This dataset provides important scientific data support for solar physics research, space weather forecasting, and related fields.

Key words Solar flares, Lyman-Alpha, Space weather, Automatic identification

Dataset Profile

Title	Standard Dataset of Lyman-alpha Solar Flare Events
Data Author	LU Lei
Data Corresponding Author	LU Lei (leilu@pmo.ac.cn)
Time Range	January 1, 2024 to December 31, 2024
Data Type	Lyman-alpha full-disk solar observation
Data Volume	Approximately 2.3 TByte
Data Format	*.fits, *.txt, *.png, *.mp4
Data Service System	DOI: https://doi.org/10.12176/01.01.00116-V01 CSTR: https://cstr.cn/14804.11.01.01.00116-V01
Sources of Funding	National Key Research and Development Program of China (2022YFF0711402, 2022YFF0503003), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB0560000), the National Natural Science Foundation of China (NSFC) (12233012)
Dataset Composition	This dataset contains solar flare events data collected by the Lyman-Alpha imager on the Kuafu-1 satellite throughout 2024. It has recorded 1937 solar flare events, including documentation of the flare identification process, lists of flare events (recording information such as the position, time, and significance of each flare), quick-view images at the flare peak moments (showing regional images and light curves of flares at their peak moments), and cinematic animations of flare regions. The data are stored in standard formats such as FITS, TXT, PNG, and MP4

0 引言

太阳耀斑是太阳大气中最强烈的爆发现象之一, 可在短时间内释放出巨大能量, 产生从伽马射线到无线电波的全波段电磁辐射增强^[1]. 太阳耀斑不仅是太阳物理学研究的重要对象, 也是影响地球空间环境的主要因素之一^[2]. 莱曼阿尔法 (Lyman-Alpha, Ly α) 位于极紫外波段 121.6 nm, 是氢原子的一个重要谱线, 对研究太阳大气低层色球活动具有重要价值, 能够提供其他波段难以获取的太阳活动信息^[3].

夸父一号星是中国自主发的首颗综合性太阳观测卫星^[4], 于 2022 年 10 月发射升空, 搭载了包括 Ly α 成像仪在内的多种先进探测设备, 能够全天候监测太阳活动与日地空间环境. 该卫星提供的全日面高时空分辨率观测数据, 为太阳活动与空间天气研究提供了新的观测窗口^[5].

近年来, 有关太阳耀斑自动识别与参数提取等研究不断推进^[5,6]. 但目前中国尚未有基于 Ly α 波段的太阳耀斑系统性观测数据集, 尤其缺乏针对 Ly α 波特性的耀斑自动识别与参数提取算法. 因此, 本数据集开发了一套适用于 Ly α 波段太阳耀斑自动识别与关键参数提取的算法^[7], 系统整理了 2024 年全年的太阳耀斑事件数据, 为太阳物理研究、空间天气预报以及相关领域提供科学数据支持.

1 数据采集与处理方法

1.1 数据来源

本数据集的原始数据来源于夸父一号卫星搭载的 Ly α 成像仪^[8-10]. 该成像仪具有 4 s 曝光时间、最快 6 s 采样间隔、空间分辨率 1.2'' (两倍像素尺寸) 的全日面观测能力, 波长覆盖 Ly α (121.6 $\pm$ 7.5) nm 波

段. 数据采集时间范围为 2024 年 1 月 1 日至 2024 年 12 月 31 日, 覆盖太阳活动上升期的重要观测窗口, 记录了不同强度和类型的太阳耀斑活动. 原始观测数据由中国科学院国家空间科学中心接收, 经过中国科学院紫金山天文台标定校正后形成基础数据^[11]. 本研究基于这些基础数据, 开发了专用的太阳耀斑自动识别与参数提取算法, 生成最终的标准数据集.

1.2 数据分类

本数据集根据 Ly α 太阳耀斑的强度特征与辐射特性, 结合 Ly α 波段的观测特点, 将太阳耀斑具体级别与划分依据列于表 1.

1.3 数据处理与耀斑事件提取

Ly α 太阳耀斑事件数据处理主要包括原始数据预处理、耀斑自动识别、追踪及关键参数提取等过程. 具体处理流程如图 1 所示.

数据集具体生产步骤如下.

步骤 1 原始观测数据预处理. 对夸父一号卫星 Ly α 成像仪获取的原始观测数据进行暗场校正、平场校正、坏点去除和噪声抑制等处理, 消除仪器效应和环境噪声, 生成标准化的 Ly α 全日面图像.

步骤 2 太阳耀斑自动识别与追踪. 传统方法多采用单阈值, 易导致耀斑起止时间误判. 本算法采用全局触发 (γ_1)+局部始末 (γ_2, γ_3) 的三重阈值设计, 能更精确地捕捉到耀斑完整的演化过程, 显著提升了对耀斑起始阶段弱信号的探测能力及计时精度. 首先, 使用一个较高的全局阈值 (γ_1) 初步判断耀斑是否发生; 然后, 在耀斑候选区内, 通过反向追溯至信号强度低于局部起始阈值 (γ_2) 的时刻来确定精确的起始时间; 最后, 在耀斑达到峰值后, 当信号强度持续低于局部结束阈值 (γ_3) 时, 判定耀斑结束. 通过区域聚类并与进行中事件列表进行实时位置比对, 本算法能够有

表 1 Ly α 太阳耀斑分类
Table 1 Classification of Ly α solar flares

强度级别	显著性特征	划分依据
微弱耀斑	$S \leq 10\%$	Ly α 辐射增强幅度不超过10%, 通常为局部小尺度能量释放
中等耀斑	$10\% < S \leq 20\%$	Ly α 辐射增强幅度介于10%~20%之间, 伴有明显的亮带结构
强烈耀斑	$20\% < S \leq 50\%$	Ly α 辐射增强幅度介于20%~50%之间, 通常伴有明显的色球蒸发和能量释放
极强耀斑	$S > 50\%$	Ly α 辐射增强幅度超过50%, 往往伴随日冕物质抛射等大尺度爆发现象

注 S 为耀斑显著性参数, 定义为耀斑区域在Ly α 波段的峰值辐射强度与背景辐射强度 (在耀斑爆发前, 耀斑所在区域的本地太阳辐射强度) 的比值减1.

效探测、分离,并独立追踪在时间或空间上紧密发生,甚至同时发生的多个耀斑.这是以往一些自动化工具难以解决的难点.针对空间观测中常见的宇宙线、高能粒子(例如南大西洋异常区影响)等噪声源,算法通过中值滤波预处理及设置触发像素数上限(NP_{max})等策略,能有效避免误报,鲁棒性强.算法在降采样的宏像素图像上进行初步探测,以保证高计算效率,同时在提取关键参数(例如位置)时,回归到原始高分辨率图像进行精化,实现了效率与精度的良好平衡.

步骤 3 耀斑关键参数提取.对识别的太阳耀斑区域,提取以下关键参数.时间参数:耀斑起始时间、峰值时间及结束时间.位置参数:耀斑中心位置(太阳表面坐标).强度参数:耀斑显著性、总辐射通量与时间积分通量.

步骤 4 数据生成与质量控制.根据提取的耀斑参数,生成标准格式的耀斑事件列表、耀斑峰值时刻

快视图图像及耀斑区域电影动画等数据.通过与其他波段观测数据(例如 GOES X 射线观测等)交叉验证,确保数据的准确性与可靠性.

2 数据样本描述

本数据集包括 2024 年全年太阳 Ly α 耀斑事件数据以及对应的数据,主要包括原始观测数据、处理后的耀斑事件列表、耀斑峰值时刻快视图图像及耀斑区域电影动画数据等.数据采用国际通用的 FITS, TXT, PNG 和 MP4 等标准格式存储,总数据量约 2.3 TByte.

耀斑事件列表采用 TXT 格式,包含每个耀斑事件的详细信息,属性字段如表 2 所示.耀斑峰值时刻快视图图像采用 PNG 格式,展示耀斑峰值时刻的区域图像与光变曲线.耀斑区域电影动画采用 MP4 格式,记录耀斑从开始到结束的完整演化过程.

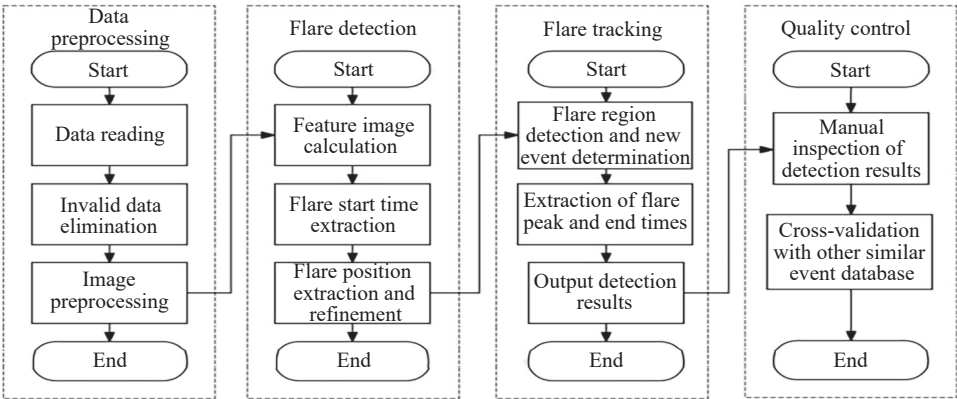


图 1 Ly α 太阳耀斑事件数据处理技术路线

Fig. 1 Technology roadmap of Ly α solar flare event data processing

表 2 Ly α 太阳耀斑事件属性说明
Table 2 Description of Ly α solar flare event

属性名称	数据类型	属性说明
Date	日期型	耀斑发生日期
Start_Time	日期时间型	耀斑起始时间 (UTC)
Peak_Time	日期时间型	耀斑峰值时间 (UTC)
End_Time	日期时间型	耀斑结束时间 (UTC)
Duration	数值型 (min)	耀斑持续时间
Position	数值型 (")	耀斑中心坐标 (日心坐标)
Peak_Flux	字符型 (DN)	耀斑峰值流量
Significance	数值型	耀斑显著性参数

截至 2024 年 12 月 31 日, 本数据集共记录了 1937 个 Ly α 太阳耀斑事件, 耀斑在太阳表面的空间位置分布如图 2 所示. 依据本数据集分类标准 (参见表 2), 数据统计结果如表 3 所示. 经统计, 2024 年发生的 Ly α 太阳耀斑中, 各强度级别占比由大到小依次为: 微弱耀斑 (69.0%)、中等耀斑 (20.5%)、强烈耀斑 (8.7%) 和极强耀斑 (1.8%).

3 数据质量控制

为了确保夸父一号卫星 Ly α 太阳耀斑事件数据集的准确性与可靠性, 实施了严格的数据质量控制措施. 质量控制过程主要包括数据预处理、质量检验方法、数据验证与校准、数据误差范围四个方面.

3.1 数据预处理

原始观测数据经过一系列预处理步骤, 包括中值

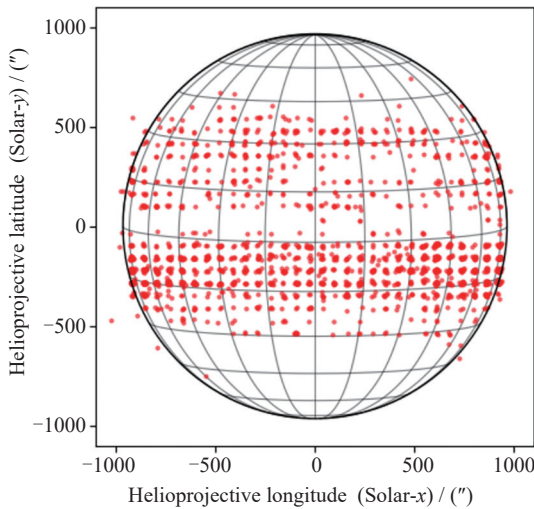


图 2 2024 年 Ly α 太阳耀斑在太阳表面的空间分布
Fig. 2 Spatial distribution of Ly α flares on the solar disk in 2024

滤波处理、图像分割处理以及背景噪声抑制. 这些预处理步骤有效提高了计算效率, 同时显著增强了耀斑特征的可见性. 特别是采用中值滤波技术, 可以消除宇宙射线干扰, 有效抵抗太空环境中宇宙射线或高能粒子流的轰击影响, 避免因此产生的孤立亮点被误判为太阳耀斑.

3.2 质量检验方法

数据质量控制采用了多种技术手段, 具体如下.

暗场校正: 消除探测器本底噪声, 提高图像信噪比.

平场校正: 校正探测器像素灵敏度的不均匀性, 确保观测数据的一致性.

坏像素去除: 识别并修复探测器中的坏像素点, 防止数据异常.

噪声抑制: 应用特定算法抑制图像噪声, 提高信号质量.

3.3 数据验证与校准

为验证数据的准确性, 采用多种方法进行交叉检验.

多波段数据交叉验证是利用不同仪器在其他波段 (例如软 X 射线) 的同步观测数据进行的交叉检验. 特别是与 GOES 卫星提供的软 X 射线耀斑列表进行比较, 在有同步观测的 36 个 GOES M 级及以上耀斑中, 本文算法成功地识别出了其中的 29 个 (具有显著 Ly α 波段增亮响应, 占比 81%).

人工审核是由太阳物理学家对检测结果进行人工审核的过程, 特别是对边缘案例进行详细分析.

3.4 数据误差范围

辐射强度的绝对误差. 仪器的绝对辐射定标是一个复杂过程, 依赖于发射前的地面标定与在轨期间的监测校准, 其不确定性会受到仪器老化、在轨环境变化等多种因素影响. 根据夸父一号卫星载荷团队的在

表 3 Ly α 太阳耀斑事件统计
Table 3 Statistics of Ly α solar flares

序号	强度级别	事件数量	数量占比/(%)
1	微弱耀斑	1337	69.0
2	中等耀斑	396	20.5
3	强烈耀斑	169	8.7
4	极强耀斑	35	1.8
合计	—	1937	100

轨标定结果,全日面成像仪(SDI)在经过平场、暗场等一系列校正后,其相对辐射定标不确定度在5%~10%的水平.因此,数据中给出的峰值流量(Peak_Flux)等绝对物理量的误差预计也在该范围内.

耀斑位置精度.本数据集中的耀斑位置精度与图像的空间分辨率和定位算法密切相关.该算法采用两步定位法:首先在降采样的宏像素图像上进行快速粗定位,然后回归到原始分辨率(像素大小约为0.6'')的图像上,通过质心计算或二维高斯拟合的方法进行精确定位.理论上,该方法可以达到亚像素级的定位精度.综合考虑卫星的指向误差等系统性因素,评估本数据集中耀斑的最终定位精度优于2个像素,即约1.2''的水平.

4 数据价值

夸父一号卫星Ly α 太阳耀斑事件标准数据集具有重要的科学价值及应用价值,主要体现在科学研究价值、实际应用价值及数据共享价值三个方面.

4.1 科学研究价值

太阳活动研究的新视角. Ly α 辐射主要源于太阳的色球层与过渡区,是研究太阳低层大气活动最有力的探针之一. 太阳耀斑是一个多层次、多波段的复杂现象, X射线主要揭示日冕中的高温等离子体行为,而其他远紫外(EUV)波段则反映不同温度层次的大气响应. Ly α 观测恰好填补了对耀斑初始能量释放关键区域——色球层的直接成像观测的空白,为理解耀斑能量释放、传输与耗散过程提供了独特视角,是对传统X射线与EUV观测的重要补充.

与非热电子注入的直接关联. 大量研究表明,耀斑中的Ly α 辐射增强与非热电子束直接相关. 在耀斑的脉冲相,高能电子从日冕的磁重联区域加速,沿着磁力线高速撞向稠密的色球层,通过碰撞加热色球等离子体,从而产生强烈的Ly α 辐射. 因此, Ly α 光变曲线及空间形态的演化,能够直接反映耀斑中非热电子的注入过程与能量沉积区域,这对于理解耀斑最核心的能量释放与粒子加速机制至关重要. 相比之下,那些由热传导主导的热耀斑可能只产生较弱的Ly α 辐射.

更早的响应时间(Neupert效应). 由于Ly α 辐射与非热电子直接关联,其响应通常早于以热辐射为主

的软X射线. 这一现象符合著名的Neupert效应,即非热电子束在色球层产生硬X射线与Ly α 等紫外辐射(瞬时响应),随后被加热的色球物质“蒸发”到日冕环中,逐渐累积并以软X射线形式辐射出去(积分响应). 因此, Ly α 辐射的峰值通常出现在软X射线耀斑的上升阶段,是耀斑爆发更早期的信号之一. 本研究开发的算法也充分考虑了这一点,通过三重阈值方案精确追溯耀斑的真实起始时间,而非简单依赖于峰值信号,这对于捕捉耀斑的早期触发过程至关重要.

揭示耀斑发展过程.通过对Ly α 辐射的成像观测,可以清晰看到耀斑双带(two-ribbon)结构在色球层的形成与分离过程,这些亮带的形态和移动轨迹直接描绘了耀斑磁环的足点位置及其演化,为研究磁场重联过程提供了直观的证据.

4.2 实际应用价值

空间天气服务. 数据集可直接应用于空间天气监测服务,为卫星操作、航天活动等提供决策支持. 减少空间天气对技术系统的影响.

辐射环境评估. 耀斑事件数据可用于评估近地空间和行星际空间的辐射环境,为航天设计及航天员辐射防护提供依据.

算法开发与验证. 该套标准数据集为太阳物理研究提供了算法开发与验证的基准数据,促进人工智能与机器学习技术在太阳物理领域的应用.

4.3 数据共享价值

本数据集以标准化格式整理,并通过国家空间科学数据中心开放共享,具有以下价值.

多学科研究平台. 数据集可以支持太阳物理、空间物理、大气物理等多个学科的交叉研究,促进学科融合.

国际合作基础. 标准化的数据格式与详细的元数据描述方便了国际合作研究,可与其他空间任务数据进行联合分析.

教育科普资源. 数据中的耀斑可视化图像与动画可作为优质的科学教育和科普资源,提升公众对太阳活动的认识.

5 讨论

在2024年共记录了1937个莱曼阿尔法太阳耀斑,其强度分布呈现出显著的幂律分布特征:微弱耀

斑占据了绝大多数 (69.0%), 而强烈与极强耀斑则相对稀少 (合计占比 10.5%). 这一统计规律深刻揭示了太阳磁能释放的基本物理特性.

能量释放的尺度不变性. 耀斑的发生频率与其能量 (或峰值亮度) 大致成反比, 这是一种在太阳物理学中被广泛观测到的现象. 这表明无论是在活跃的太阳活动峰年, 还是在相对平静的年份, 驱动耀斑的磁重联过程在物理机制上是统一的. 小尺度的、局域性的磁能释放事件在数量上远远超过需要巨大、复杂磁场结构才能触发的大尺度能量释放事件.

上升期的活动特征. 2024 年作为太阳活动高年, 其耀斑总数显著增加. 本数据集观测到的大量微弱耀斑, 正是太阳大气中无处不在的、持续不断的小尺度能量释放过程的直接体现. 这些看似微不足道的事件, 其累积效应对加热太阳大气 (特别是色球层和日冕) 可能起至关重要的作用. 而少数极端耀斑事件的发生, 则标志着活动区磁场能量积累与演化达到了临界点, 是太阳活动峰年最主要的特征之一.

因此, 本数据集提供的莱曼阿尔法波段耀斑强度分布统计结果, 不仅为研究太阳活动周期间的能量释放谱提供了宝贵的观测样本, 也从一个独特的观测窗口 (色球层) 验证了耀斑能量释放的尺度不变性这一基本物理规律.

总体来说, 这套涵盖全年的 Ly α 太阳耀斑事件数据集不仅丰富了中国自主太阳观测数据, 也为太阳活动研究和空间天气预报提供了宝贵的科学数据支撑, 具有重要的科学价值与应用价值. 通过这一数据的公开共享及广泛应用, 将进一步促进太阳物理领域的科学研究与技术创新, 同时为太阳活动对地球环境影响的研究和预测提供了数据基础.

利益冲突 所有作者均声明不存在利益冲突.

参考文献

- [1] BENZ A O. Flare observations[J]. *Living Reviews in Solar Physics*, 2008, **5**(1): 1
- [2] CURTO J J. Geomagnetic solar flare effects: a review[J]. *Journal of Space Weather and Space Climate*, 2020, **10**: 27
- [3] WAUTERS L, DOMINIQUE M, MILLIGAN R, *et al.* Observation of a flare and filament eruption in Lyman- α on 8 September 2011 by the PROject for OnBoard Autonomy/Large Yield Radiometer (PROBA2/LYRA)[J]. *Solar Physics*, 2022, **297**(3): 1-12
- [4] GAN W Q, FENG L, SU Y. A Chinese solar observatory in space[J]. *Nature Astronomy*, 2022, **6**(1): 165
- [5] FERNANDEZ BORDA R A, MININNI P D, MANDRINI C H, *et al.* Automatic solar flare detection using neural network techniques[J]. *Solar Physics*, 2002, **206**(2): 347-357
- [6] BONTE K, BERGHMANS D, DE GROOF A, *et al.* SOFAST: automated flare detection with the PROBA2/SWAP EUV imager[J]. *Solar Physics*, 2013, **286**(1): 185-199
- [7] LU L, TIAN Z Y, FENG L, *et al.* Automatic solar flare detection using the solar disk imager onboard the ASO-S mission[J]. *Solar Physics*, 2024, **299**(5): 72
- [8] LI H, CHEN B, FENG L, *et al.* The Lyman-alpha Solar Telescope (LST) for the ASO-S mission—I. scientific objectives and overview[J]. *Research in Astronomy and Astrophysics*, 2019, **19**(11): 158
- [9] CHEN B, LI H, SONG K F, *et al.* The Lyman-alpha Solar Telescope (LST) for the ASO-S mission—II. design of LST[J]. *Research in Astronomy and Astrophysics*, 2019, **19**(11): 159
- [10] FENG L, LI H, CHEN B, *et al.* The Lyman-alpha Solar Telescope (LST) for the ASO-S mission—III. data and potential diagnostics[J]. *Research in Astronomy and Astrophysics*, 2019, **19**(11): 162
- [11] HUANG Y, LI Y P, LIU S, *et al.* Science operation and data analysis center of the Advanced Space-based Solar Observatory (ASO-S) mission[J]. *Solar Physics*, 2024, **299**(9): 1-10

作者简介



卢 磊 (通信作者) 男, 1987 年 6 月出生于安徽省蒙城县, 现为中国科学院紫金山天文台副研究员, 主要从事太阳活动的多波段观测与诊断, 利用地基和空间观测数据研究太阳磁场、耀斑与日冕物质抛射等活动现象及其物理机制.
E-mail: leilu@pmo.ac.cn