

# CSST 科学数据处理系统月报

第 36 期

主办方：中国空间站巡天望远镜科学工作联合中心

时间：2023 年 12 月 11 日

## 新闻

### CSST 科学数据处理系统召开线下组长例会

2023 年 11 月 30 日-12 月 1 日，CSST 科学数据处理系统在南京召开模块组长双周例会。本次会议对目前 Cycle8 阶段近两周研制进度进行汇报，并对开展过程中遇到的问题进行了专项讨论并给出推进方案，主要包括集成封装和集成测试规范讨论、系统测试团队的建立和测试规范、定标产品流水线与 CCDS 的讨论、各模块对仿真需求的进展、CTI 仿真与修正相关讨论、焦面 4 联合天测仿真、2 级数据产品定义和科学效能评估进展的相关讨论等；并对后期工作的开展做出计划安排。

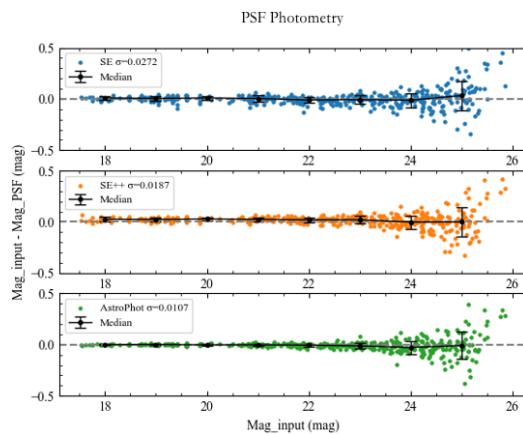


## 突出进展

### 主巡天多色成像模块研制工作进展

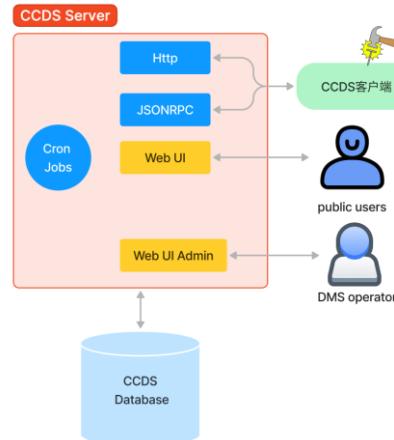
CSST 巡天具有大数据、高精度的特点，传统测光算法在效率和精度方面面临挑战，CPU/GPU 并行计算将是解决该问题有效途径之一。多色成像模块团队利用 Cycle6 仿真数据，对支持 CPU 并行的 SourceXtractor++（简称 SE++）、支持 GPU 并行的 AstroPhot（简称 AP）新型测光工具进行了测试，并与 SExtractor 结果进行比较。测试结果无论是 PSF 测光、还是星系模型（1 sersic+1 指数盘）测光，SE++ 和 AP 的精度都更高（见下表）。从运行效率上看，SE++速度最快，仅耗时 2 分钟，是 SExtractor 的 6.5 倍；AP 最慢，需数小时，可能原因是数据在内存和显存之间的通信瓶颈可能会阻碍 GPU 加速，尚需我们优化 I/O 性能。

	SExtractor	SE++	AP
PSF 测光相对测光误差	2.7%	1.9%	1.1%
模型测光相对测光误差	6.8%	6.1%	4%



### 在轨定标模块研制工作进展

定标参考数据系统是连接定标参考文件及科学数据处理流水线之间的重要纽带，用于将在轨定标生成的多类型多版本的定标参考文件，按照制定的规则以最佳的形式推荐给数据处理流水线使用，从而使得流水线的运转更加自动化且高效精确。在轨定标团队经过不懈努力和攻关，完成了 CSST 的定标参考数据系统软件的初步开发（简称 CCDS）。CCDS 的软件开发囊括了客户端、服务器端及网页，全面地为在轨定标科学家、流水线运行科学家、系统管理员等提供服务和支持。现阶段已实现了用户管理功能，锁定机制，参考文件提交的版本号优化和 Web 界面端等功能。并配合流水线的开发，在数据处理系统不同的测试平台上进行了部署测试，包括测试版和使用版。同时，五个后端模块还依照制定的在轨定标方案，进行了各类定标参考数据的定义；并依据 C6.2 版仿真软件进行各类定标数据仿真，生成和制作了供现阶段流水线使用的近三十类参考文件，正陆续接入 CCDS。这些都为在轨后定标参考数据系统的顺利运转奠定了基础。后续将继续全力优化定标参考数据系统软件，完成定标参考数据完备定义。



**数据系统研制进展**

序号	模块名称	月度进展
1	多色成像	基本完成所有多色成像模块单元测试；主巡天公共模块部分功能开发完成并通过单元测试；开展地面测试代码编写工作，初步完成部分功能 docker 封装。
2	无缝光谱	配合开展流水线-接口-测试；初步完成单元测试代码更新；对伪直接成像开展 PSF 重构测试并得到初步结果；为系统测试工作进行准备。
3	数据流管理	正在开展数据处理系统云环境部署；2 级科学数据文件的服务实现；完成 2 级科学数据导入和数据查询，创建 Web 前端辅助生成 SQL。
4	流水线运行管理	安装空间应用中心测试环境的流水线运行管理系统，在运行端实现对算法模块权限管理；流水线引擎功能完善，以结构化形式管理用户程序的 stdout/stderr，解决后期生产运行中的问题。
5	多通道成像仪	MCI 仿真接入天光背景杂散光模块，可模拟随观测时间及空间轨道变化的杂散光、黄道光、地气光效应；完成 1 级 2 级流水线代码优化工作，根据修改意见进行修改中，计划 12 月开展单元测试工作；完成地面测试用 bias、dark 和外部 flat 图像采集。
6	积分视场光谱仪	一级流水线代码 QC, RSS, CUBE 模块通过单元测试，WCS 模块单元测试中；正在进行二级流水线代码规范化修改；完成焦面 4 联合定标第一轮任务布置；正在分析仪器测试数据。
7	星冕仪	完成数据处理流水线代码优化并开展单元测试；完成可见光相机仿真中的制冷温度和电子倍增漂移效应的仿真和建模。
8	太赫兹	继续开展一级流水线规范化工作，新版本进入新要求下的单元测试，对流水线新的启动机制初步形成方案；继续开展太赫兹模块 CCDS 系统接入一级流水线工作，推进太赫兹模块定标产品制定工作。
9	天体测量	恒星部分、太阳系近地天体部分代码已在本地跑通；完成单元测试准备数据，完成单元测试代码的编写工作。
10	观测数据仿真软件	基本完成软件改版模块的功能集成和测试，优化 CTI 及平场、postflash、无缝光谱多级像 PSF 等模块实现和数据接口，测试分析了光学系统中鬼像杂散光效应，完成了摆镜偏移对成像质量影响的分析报告，检查优化恒星星表和半解析星系星表，讨论并推进各模块对仿真的需求和任务细化工作，完成了 Jenkins 平台的仿真单元测试样例示范并逐步开展其他仿真模块的单元测试。
11	数据可用性标记	正在进行 QCO 的代码优化，计划 12 月中旬能够完成优化并进行单元测试。
12	在轨定标	整理了发射线天体（PNe 及 Green Pea 发射线星系）的 catalog 并尝试仿真；完成定标产品生成流水线初步的代码修改及审核，正在进一步修改优化中；进行新版的定标参考数据定义修改；解决 CCDS 参考文件提交方面遇到的问题；配合流水线方面的需求，调整用户名密码为环境变量；测试了二进制模型文件接入 CCDS 的可行性；继续分析 MCI、IFS 实验室测试数据，计划与模块研制方讨论数据细节；MCI 安排近期利用新疆的望远镜开展部分在轨定标星的观测计划；IFS 整理了部分 WR 和 Be 发射线天体，准备增加 PNe 目标获取光谱并开展仿真分析。
13	观测需求编排	修正 UWS 后台任务管理的一些 BUG，初步完成后端编排程序到新框架的迁移并与编排仿真软件讨论确认新的 Python 接口；维护警报监测系统，优化与完善事件响应程序；尝试动态生成 CSST 观测天区图层文件以支持 WWT 在平面及球幕模式环境中进行展示；丰富一般观测申请中评审专家的相关界面以便于提案分发与评审。
14	数据处理软件集成	完成流水线算法开发规范和单元测试规范的拟定和讲解；推动数据统一，方便多版本数据的部署和测试；正在调整优化 Jenkins 上各个算法模块的格式检查和单元测试脚本；更新流水线网站内容；对接空间应用中心公有云环境；部署国台集成测试环境并计划开展集成测试。

**近期节点和计划**

时间节点	计划任务安排	时间节点	计划任务安排
23 年 8 月	数据处理系统完成中期评审	23 年 11 月	在轨定标方案评审
23 年 8 月-23 年 12 月	仿真软件更新版本	23 年 11 月	各模块完成单元测试
23 年 8 月-24 年 1 月	数据处理软件已有算法完善和优化	23 年 12 月	完成观测需求编排科学专家咨询
23 年 8 月-24 年 1 月	数据处理软件 2 级流水线开发	23 年 12 月	流水线完成集成测试
23 年 8 月-24 年 1 月	配合开展地面测试	23 年 12 月	根据鉴定件定标参数完成仿真软件的修改和小天区仿真数据生成
23 年 8 月-24 年 1 月	定标产品生成算法迭代和更新	23 年 12 月	完成第二次科学效能评估
23 年 8 月-24 年 1 月	观测需求编排软件的更新迭代	23 年 12 月	完成配置项软件设计评审
23 年 8 月-24 年 1 月	配合完成地面测试	23 年 12 月	完成 2 级数据定义专家咨询
23 年 10 月	流水线设计优化	24 年 1 月	完成 C8 阶段结束，完成总结和任务验收
23 年 11 月初	单元测试平台完善和优化		