

CSST 科学数据处理系统月报

第 44 期

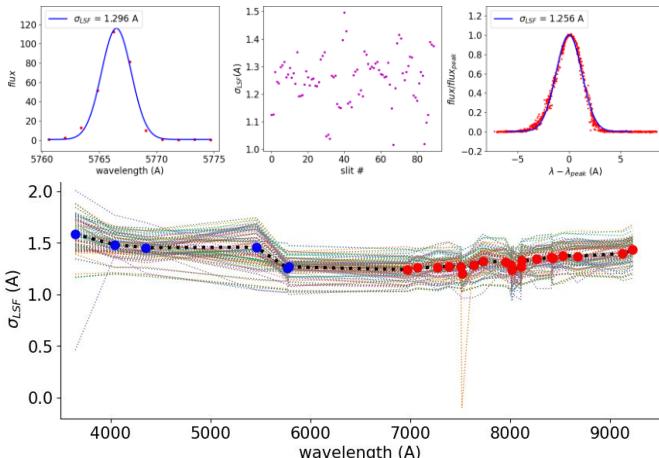
主办方：中国空间站巡天望远镜科学工作联合中心

时间：2024 年 10 月 14 日

突出进展

积分视场光谱仪模块研制工作进展

IFS 数据处理团队显著提高了 IFS 仪器的线扩展函数 (LSF) 的测量精度。IFS 数据处理团队基于实验室实测得到的 IFS 灯谱发射线，首先对不同位置处的发射线光谱进行追迹和波长定标，在此基础上采用亚像素平移对不同位置处的同一发射线进行波长对齐和叠加，从而获得高精度的谱线展宽轮廓，即线扩展函数 (line spread function)。以 5767Å 波长谱线为例，相比于单条谱线的线扩展函数的拟合（图上左），不同位置处该谱线（图上中）叠加后，其线扩展函数的拟合精度可以提高 9 倍以上（图上右）。通过该方法，IFS 仪器在不同波长的 LSF 的宽度都可以获得高精度测量和标定（图下），这对精确测量科学观测目标中其它物理因素造成的谱线展宽效应具有重要意义。



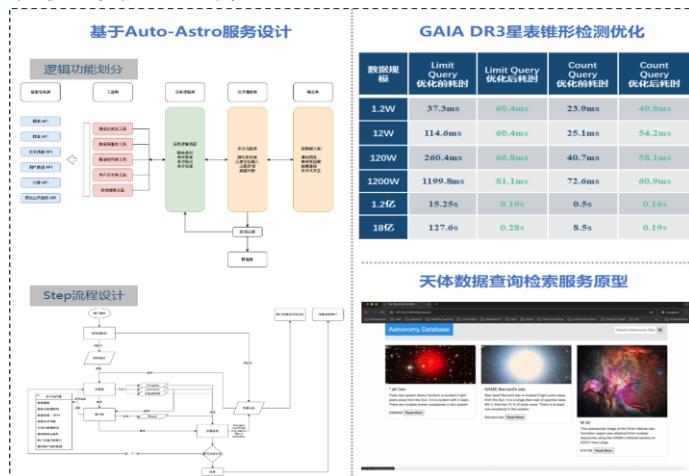
仿真软件配置项测试取得突出进展

CSST 科学数据处理系统仿真软件是 CSST 科学数据处理系统三大配置项之一，包括主巡天、多通道成像仪、积分视场光谱仪和星冕仪四个后端模块的仿真软件。配置项测试是贯穿于软件开发全生命周期的

重要阶段。C9 中后期仿真测试团队正式启动了仿真软件配置项测试工作。团队特邀了行业测试专家，对测试人员进行了专业培训。随后，经过一系列精心筹备，包括测试人员的明确分工、测试方案的深度剖析、测试细项讨论，再到测试的每一个用例，经过多次线上线下沟通和协作，团队 C10 初期（8-9 月）依次确定了四个后端模块的配置项测试方案，对测试用例进行逐一讨论，以确保测试方案的科学性、完整性和可操作性，为形成正式的《软件配置项测试说明》文档奠定了基础，也为 CSST 科学数据处理系统仿真软件的高质量运行提供了保障。

数据融合模块研制工作进展

多波段数据融合分析服务构建星表元数据知识集，统一数据与任务特征描述；基于 LLM 设计 Auto-Astro 服务系统架构，并搭建原型系统；在之江云 ODPS、ADB 等云原生软件服务基础上设计面向大数据星表的智能层级自适应混合索引，将 GAIA DR3 等星表数据的 20 亿分页查询耗时从 7.05s 优化提升到 0.28s，并在百亿星表交叉匹配分钟级基础上实现星表数据融合服务的 200+ 交叉匹配方法，设计、实现 5 星表联合交叉匹配方法。



数据系统研制进展

序号	模块名称	月度进展
1	多色成像	完成 W2 天区仿真数据单次曝光数据处理；对 W2 天区数据处理定标结果进行评估；与之江实验室讨论了星表融合、结果评估等相关工作和计划。
2	无缝光谱	在 W2 天区的无缝光谱仿真数据中发现了一些问题，并已提供反馈。稳步推进无缝光谱实测仿真的生成工作。在这一过程中，对输入星表的物理性质进行了细致的检查，以确保其准确性。同时，对红移测量程序进行了初步的测试。此外，持续参与流水线集成的相关工作，以确保整个流程的顺畅和高效。
3	数据流管理	优化 2 级数据服务，2 级数据写入查询支持指定版本；配合主巡天二级流水线的开发工作；参与讨论 DFS L2 相关讨论；正在开展 CCDS 文件存储优化工作。

序号	模块名称	月度进展
4	流水线运行管理	配合数据处理软件集成模块在国台环境进行一级流水线测试；完成流水线平台动态资源适配优化；参加 CSST 运行管线关键技术研讨会。
5	多通道成像仪	完成 C10 计划的确定；撰写《CSST 数据处理系统科学数据仿真软件配置项测试说明-多通道成像仪软件》文档并开展配置项测试；调整 PSF 重构工具包，分别加入图像对减和图像叠加模块。
6	积分视场光谱仪	完成仿真软件配置项测试中 IFS 模块的测试计划；开展 0 级数据校验和二级流水线的集成；配合完成 IFS 长光所集成测试数据的分析。
7	星冕仪	完成星冕仪科学目标仿真和仪器仿真两篇文章初稿；完成星冕仪仿真软件测试说明文档的编写；正在开展数据处理流水线程序的整理重构和性能分析工作。
8	太赫兹	开展太赫兹模块科学数据处理二级流水线之 OTF 算法模块研制并初步实现 OTF 算法；开展太赫兹模块数据处理仿真文章撰写，目前已完成初稿。
9	天体测量	基于新的仿真数据完成了天体测量二级流水线单元测试；完成 5 平方度数据天体测量一级流水线的数据初步处理。
10	观测数据仿真软件	持续优化仿真软件，添加了配置参数的网页接口工具；完成 50 平方度仿真数据更新；完成 1000 平方度仿真的星系星表；完成配置项测试计划的初步编写并着手搭建配置项测试环境和测试数据。
11	数据可用性标记	完成初版天区密度分析功能，正在扩大模型的适用范围。
12	在轨定标	开展 OOC 流水线新的结构调整，完成代码修改工作，计划开展跟 DFS 和 CCDS 的联合测试；开展可观测性及定标星场选取约束研究，并计算运行 10 年的可观测天数；开展星团列表梳理工作，综合优化定标星场的选取；初步完成 C10 阶段波长定标源和流量定标源的仿真并基于波长定标源仿真数据优化了波长定标算法，生成 12 个无缝光谱 chips 的波长定标参考文件；使用加了 Av 的最新输入星表进行 NGC2298 的仿真。
13	观测需求编排	持续集成编排仿真程序并优化程序流程；实现对一般观测申请提案合作者的提案状态和数据通知功能；完成一般观测申请保存型操作的单元测试代码编写，通过 CI/CD 自动测试；进行观测天区可视化前端优化，调整天区显示方式，实现并测试时间轴功能。
14	数据处理软件集成	讨论并确认数据处理软件平台的框架，明确了各个开发团队的任务；完成多色成像星表合并流水线调试；完成 v092 版主巡天仿真数据 W2 天区的星表合并处理。
15	存储与计算	支持天文算子和 CSST 计算工作台的开发；计算引擎方面，完成 slurm on k8s 的研制，以支持 HPC 类型的作业提交；存储引擎方面，尝试在 z100 集群上自建分布式存储，用以支持块存储、对象存储和 FileSystem 类型存储。
16	天文算子	完成 4 条一级流水线(sls、mbi、photmix、sls-extraction-zero)的天文算子拆分及上云改造，各算子实现在之江云环境中独立运行；对各拆分后的算子进行了初步的性能测试；instrument 模块在 GPU 上跑通。
17	CSST 计算工作台	科研工作台增加用户体系模块，完成需求评审、技术评审，整体开发进度完成 60%；完成天文 AI 助手 2.0 需求评审及技术调研；针对科研工作台中计划的流水线开发测试环境，开展了多次讨论来收集需求并实现初版的 PRD。
18	数据融合	设计面向大数据星表的智能层级自适应混合索引，星表数据的 20 亿分页查询耗时从 7.05s 优化提升到 0.28s，并实现星表数据融合服务的 200+ 交叉匹配方法。



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
2024 年 9 月	完成 2 级数据定义会签（已完成）	2024 年 12 月	完成《软件设计说明》评审
2024 年 10 月	完成《仿真软件配置项测试说明》评审	2025 年 12 月	基本完成 1000 平方度仿真
2024 年 10 月	完成数据系统 0、1 数据定义和定标数据定义版本更新	2025 年 1 月	完成编排软件 v1.0
2024 年 10 月	完成 50 平方度深场仿真	2025 年 1 月	完成《CSST 数据处理系统技术报告》
2024 年 11 月	完成《仿真软件配置项测试报告》评审	2025 年 1 月	CSST 科学数据处理系统 v0.8 版纳入受控库
2024 年 12 月	完成 C10 流水线集成测试		