

新闻

CSST 科学数据处理系统召开线下组长会

2025 年 5 月 13 日-14 日，CSST 科学数据处理系统线下组长会在紫金山天文台顺利召开。会上，各模块组长依次介绍了本阶段的研制工作进展情况，会议还开展了专项讨论环节，针对存在交叉关联的工作进行了深入探讨，包括各模块观测需求编排、配置项测试、在轨定标初步方案、新增 0 级数据需求、流水线集成和数据系统配置管理等多方面内容；会议最后还讨论了主巡天无缝光谱 0 级像影响评估工作，以及之江与弱透镜科学团队协同工作进展情况。通过本次会议，全面了解了数据处理系统各模块的进展情况，借助各专题讨论，提出了一系列交叉关联问题的解决方案，为保障项目的顺利推进奠定了坚实基础。

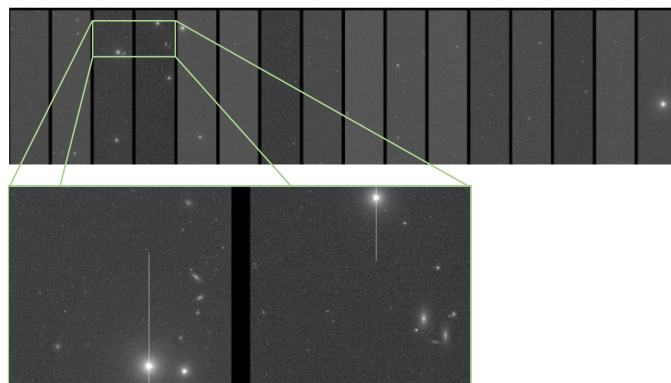


突出进展

观测数据仿真软件模块取得突出进展

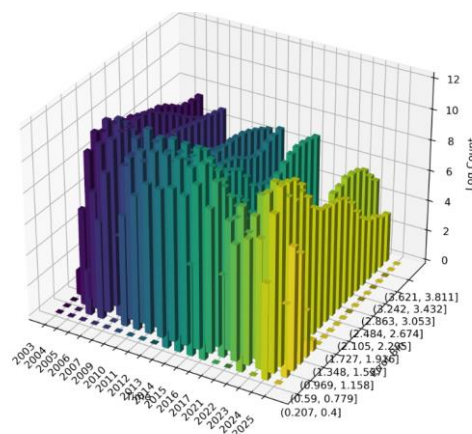
经过半年的更新迭代，CSST 主巡天仿真团队完成了观测仿真软件 3.2.0 的版本释放，在新版仿真软件中添加了探测器的通道间串扰和银河系消光的功能模块，升级亮星 PSF 的外插模块和头文件关键字定义，在无缝光谱仿真中添加电子弥散，改正了饱和溢出在通道间的拖尾截断和深场仿真的星等计算等模块。结合多个模块的内存优化，新版仿真软件在计算效率和功能实现上有了显著提升。基于最新的成像仿真软件，完成了 1 平方度的深场仿真，包含 636 个曝光指向的多色成像数据（见下图）。新版仿真软件的释放和 1 平方度的深场仿真的完成为后续 50 平方度

深场仿真和即将开展的 1000 平方度仿真奠定了良好基础。



天文算子模块取得突出进展

天文算子团队在 CTI 去除方面取得突出进展。CTI 去除的权威方法是哈勃 HST 采用的通过既定查表获规则逐像素得到改正后像素，并非通用可迁移算法，最大问题是强硬件依赖。相比下，Sota 方法“Arctic”更通用，基于 HST 的先验硬件缺陷模型，迭代计算“假拖尾”并去除，但是原理上并不鲁棒，且实验表明该方法的精度受先验缺陷参数的精度影响严重，以及速度慢等缺点。团队成员首先选取单一年份观测的 HST 的 CTI 改正前后的图像对，利用效应物理特性将图像单一列像素作为样本，训练有监督模型，在测试集上完成 CTI 去除功能。该方法在速度上只基于卷积层模型单卡推理相比于“Arctic”单核计算有 200 倍速度提升；在精度上本模型 CTI 去除率指标的直方图峰值位置显著小于 Arctic。下一步主要是提升模型精度和更鲁棒测试指标以及提升模型通用能力。



数据系统研制进展

序号	模块名称	月度进展
1	多色成像	配合流水线集成完成多色成像数据处理各算法模块的更新工作，添加位置和流量定标的模版更新工作并纳入 CCDS 管理；依据 0 级数据改动修改 1 级数据产品的相应 keywords，更新 toml 和模版文件；与仿真团队合作完成 250 秒*8 深场数据仿真并进行初步检查；完成 C9.2 版本仿真数据 W5 天区 mosaic 测试，具备深场图像合并流水线开发能力。
2	无缝光谱	针对新版透过率曲线导致的 GI 波段零级像强度减弱问题，进一步深入分析了其对无缝光谱图像 WCS 定标的影响，并详细研究了零级像形态的变化及其引发的定心误差；继续配合开展流水线的集成与配置项测试相关工作。
3	数据流管理	根据新的定标参考文件定义修改 CCDS，完成测试系统的重置，使用新版 Mapping 的定义；根据新版编排数据定义修改 DFS 中的相关部分。
4	流水线运行管理	开展数据重处理的讨论，调整 gRPC 接口，新增获取“所有执行实例”接口；开展 scalebox 平台持续优化工作。
5	多通道成像仪	进行多轮编排定义和在轨定标定义的迭代和讨论；完成编排 json 文件并上传；根据编排文件完成仿真数据的生成工作；对新一批仿真数据进行 toml 格式检查并通过；开展单元测试，完成仪器效应改正模块本地测试并进行下一步测试；使用仿真数据完成参考文件的生成。
6	积分视场光谱仪	完成 IFS 仪器仿真文章并投稿；完成 IFS 二级数据处理中宽发射线拟合程序及相关文档；协助完成多份早期科学观测提案和数据竞赛方案并提交。
7	星冕仪	开展星冕仪数据处理流水线初步测试，并结合测试结果对数据处理流水线开展性能和稳定性的优化工作；开展单元测试工作；确定编排输出文件接口；优化观测仿真程序，增加新的宇宙线模型，并增加了成像暗区设置功能；完成观测编排接口文件的修订工作。
8	太赫兹	进一步细化太赫兹编排数据文件定义，敲定 Cycle11 研制阶段太赫兹数据处理流水线的编排文件具体内容；开展新版太赫兹 0 级仿真数据和编排数据文件生成工作，推进流水线测试级配置项测试工作；完成太赫兹仿真文章编辑反馈意见修改。
9	天体测量	进行多轮编排定义和在轨定标定义的迭代和讨论；根据最新情况完成编排 json 文件并上传；根据编排文件完成仿真数据的生成工作。
10	观测数据仿真软件	持续优化仿真软件；完成黄道光背景光的参数文件更新工作；完成 1 平方度的深场仿真数据，生成 1 平方度强透镜像星表并开展仿真接口的适配测试。
11	数据可用性标记	QC0 按照新版 0 级数据定义更新完成并集成到流水线中。
12	在轨定标	完成定标参考数据定义 Cycle11 版的修改完善和最终会签；进行在轨定标方案的讨论并根据意见进行修改完善；依据新版定标参考数据定义完成新版参考文件的更新并提交 CCDS；依据新的 0 级定义更新各类定标仿真数据及相应编排文件，更新定标产品生成流水线；完成定标产品生成工具包测光零点、颜色项改正、测光平场、超级天光平场、像场畸变的封装工作并提交 gitlab，其他工具包也依据新的数据定义进行了更新。
13	观测需求编排	以打通流水线集成全过程为牵引，联合制订主巡天及各精测模块的定标及观测参数列表，并与 0 级数据格式和 DFS 进行参数名称统一，给出了各种观测模式的完整例子；完成一般观测申请前后端重构工作，实现申请填写-提交-评审-反馈-观测的全链路模块化，可根据实际运行的需求灵活调整，提高了整个系统的可复用性。
14	数据处理软件集成	推动编排-0 级数据定义-DFS 的关键字统一；推动编排参数确定和格式更新，补齐各个测试数据集对应的编排数据；推动 0、1 级数据定义调整和测试数据集的更新。
15	存储与计算	完成 z100 集群网络插件升级、扩展 A40 GPU 节点，后续可提供更多 GPU 资源；探索高性能 POSIX FileSystem 方案。
16	天文算子	利用效应物理特性将图像单列像素作为样本，训练有监督模型，在测试集上完成 CTI 去除功能，在速度和精度上得到提升，将继续开展对本模型精度和更鲁棒测试指标的提升，进一步提升模型通用能力。
17	CSST 计算工作台	在线协作开发功能联调，提测完成，待发布；流水线重处理任务发起、DAG 日志管理功能需求讨论，并对 PRD 和 UI 设计稿进行了评审，待开发排期
18	数据融合	参考 Joint Survey Process (JSP) 设计多波段图像数据融合服务功能模块，包括图像搜索，图像校准，生成 PSF 估计等。设计实现面向内存优化的无缝光谱图像的动态切割算法。

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
6 月 20 日	完成 1 级流水线的全部开发和版本固化	8 月 30 日	完成 2 级流水线的全部开发和版本固化
6 月 30 日	完成定标产品生成流水线和定标产品的生成工具包固化	8 月 30 日	完成数据处理软件 1 级流水线配置项测试
6 月 30 日	在轨定标方案评审	9 月 30 日	完成数据处理软件 2 级流水线配置项测试
7 月 30 日	基本完成 1000 平方度仿真	8 月 30 日	完成数据处理软件新功能开发
7 月 30 日	完成 50 平方度深场仿真	12 月 30 日	完成编排软件配置项测试
7 月 30 日	完成《软件设计说明》评审	26 年 1 月	实现共有云环境下的部署，开展系统测试