

CSST 科学数据处理系统月报

第 28 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 04 月 13 日



CSST 科学数据处理系统参加 CSST 科学年会

2023 年 3 月 27 日-4 月 1 日，CSST 科学数据处理系统团队相关代表参加了在北京怀柔召开的中国空间站工程巡天望远镜科学年会，并在年会上做了系列报告，其中有 6 个大会报告，主要是关于 CSST 科学数据处理系统总体结构和进展、CSST 在轨定标初步方案、CSST 观测需求编排以及各后端模块的数据处理研制和仿真软件的相关进展等，其中主巡天模块无缝光谱数据处理与科学研究开设了专门的分会场，进行深入交流讨论。通过此次科学大会，让更多科学家了解了 CSST 科学数据处理系统具体的研制内容和研制进展情况，同时通过与科学家的面对面讨论，也让科学数据处理系统了解到科学家对科学数据处理系统的需求，为后续科学数据系统的研制提供了帮助，有着积极的推动作用。

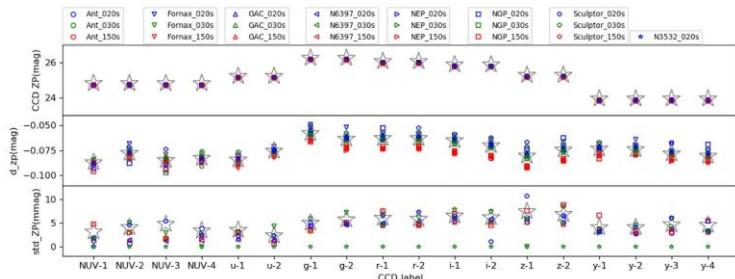


在轨定标模块研制工作进展

在轨定标团队完成了最新在轨定标仿真数据的处理和分析，形成初版的“定标星场选取及仿真分析”报告文档并讨论下一步需要对 NUV 波段增选或补充仿真的地方。详细比较了 9 个不同特征天区星场 (NGP、GAC、NEP、Bulge、Ant、N6397、N3532、Fornax、Sculptor) 的单指向和多指向仿真数据的基本数据质量，包括 PSF、天光背景、位置定标、流量定标等信息。

分析结果表明：1) 天体数密度过高的密集星场天区（如 Galactic Bulge）会对数据处理产生限制，不适合作为定标星场使用；2) 不同曝光时间/不同数密度的天区的 PSF 基本一致，其弥散与天区数密度相关；3) 不同曝光时间不同天区测得的流量定标星等零点基本一致，但曝光时间越长其弥散越小，且用于流量定标的参考源数目与天区数密度强相关。（见图中上图为星等零点，五角星为理论值；中图为实测值与理论值偏差；下图为不同曝光之间的零点弥散；中、下图中五角星为均值。）

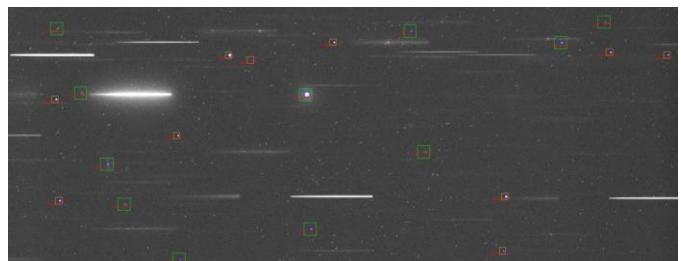
同时，利用 30 指向的仿真数据，我们改进了超级天光平场和测光平场的拟合流程，对各天区 g、r、i、z 波段可以得到正常的天光平场，但 NUV、u 和 y 波段因天光背景值过低尚无法合成有效的天光平场，下一步我们将对其进行针对性的补充仿真和分析。



主巡天无缝光谱模块研制工作进展

CSST 无缝光谱的硬件设计和观测模式不同于其他空间望远镜 (HST、JWST、Roman 和 Euclid)，无法开展对应观测得到直接成像 (direct image)。为了采用传统方法即以直接成像的中心作为光谱提取的参考点，无缝光谱处理团队克服了重重困难与挑战，有条不紊地展开“伪直接成像”的构建工作。同时，经过团结协作，不懈努力，已在基于“零级像”作为参考点的光谱提取算法上获得了重要进展。

以零级像作为参考点，首先要对目标源的零级像进行正确探测，再根据配套的定标文件提取光谱。目前开发了两种探测零级像的方案：一种是从全局图像探测零级像，另一种是基于伪直接像的星表作为初始位置，叠加零级像与伪直接像的偏移量来构建 search box，并在每个 box 内开展零级像的探测 (见 GI 图像区域示例。其中橙色框表示恒星的 search box，绿色框表示星系的 search box，蓝色圆表示探测出的零级像的轮廓)。该算法已应用在 C6 仿真数据上进行测试，初步测试结果表明：对于 g 波段星等 MAG<21 (22) 的点源和亮星系，在 GI 与 GV 波段的综合探测成功率高达~97% (90%)，效果较好，但在 GU 波段的综合探测成功率~20%，仍需作进一步测试 (下表为测试统计表)。



图像编号	g 波段星等	星表	零级像数目	Search box	全局探测
GI(01)	MAG<=22	175	172	162	152
GI(01)	MAG<=21	124	121	120	119
GV(02)	MAG<=22	160	157	137	113
GV(02)	MAG<=21	104	103	101	100
GU(03)	MAG<=21	69	69	15	10


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 data model 和算法文档；开展 Cycle6 的 2 平方度数据处理；和 QC2 团队共同开展测光星表检验；完成密集星场恒星匹配开发。
2	无缝光谱	完成一级流水线集成的单元测试、报告和算法文档；进一步优化 CDE 抽谱程序，测试对基于 MaNGA 多光纤光谱构建的近邻星系的无缝光谱抽取；测试基于 0 级像进行无缝光谱提取。
3	数据流管理	进行定标数据更新后的数据导入及集成 DFS API 辅助工作；实现 CRDS Server 的删除 reference 和 bestref 功能。
4	流水线运行管理	完善流水线集成中的新增功能研发，包括规范化的定制时间戳，定制日志输出的截取、容器时区的定制支持等；完成一级流水线的定时触发模块的研发；参与 MCI/IFS 一级流水线集成的设计；为主巡天流水线研发中的数据处理提供支持。
5	多通道成像仪	利用 MCI 仿真生成用于一级流水线仿真数据，继续生成用于二级流水线仿真数据；基于上海台流水线环境配置开展一级流水线数据处理测试；二级流水线中将暂现源模拟数据放入图像数据中。
6	积分视场光谱仪	完成 IFS 仿真和数据处理软件的单元测试、算法文档和 data model；完成 0 级数据和 1 级数据定义文档修订；完成 data manager 和 data model 类的原型开发。
7	星冕仪	完成数据处理与轨定标的算法文档和 data model 文档；编写星冕仪流水线配置项测试计划文档；完成观测仿真程序 2.0 版本的编写。
8	太赫兹	完成太赫兹模块一级数据处理算法文档、0 级数据和 1 级数据 data model 文档；初步完成太赫兹模块在轨定标中算法文档编写，对位置修正与频率修正的开展调研和讨论，初步形成方案；针对最新定义的 0 级以及 1 级数据 data model，开展一级流水线的更新工作，持续推进。
9	天体测量	利用 WISE 数据，对编号为 00475、00933、00159 的三颗小行星进行计算，给出每颗小行星分别属于七类轨道类型的概率；进行一级数据处理和二级数据处理模块算法收集；开展 1 级数据处理和 2 级数据各个处理模块独立测试。
10	观测数据仿真软件	结合三款不同探测器的实验室数据，开展成像仿真任务，为探测器选型的科学测试提供仿真数据；完成导星模块的焦面配置，开发精密导星仪的仿真程序；完成无缝光谱 0、1 级像的 PSF 仿真采样，开展杂散光模块的软件集成和测试，优化仿真程序和星表，开展仿真软件的单元测试，为新版仿真软件的发布做准备。
11	数据可用性标记	基本完成 1 级数据定义最新稿并同步至 TeamBition；针对 CRDS 流水线对 QC0 的部分需求更新，重新进行开发。
12	在轨定标	进行 9 个定标星场的各类仿真数据处理结果分析，形成初版的“定标星场选取及仿真分析”报告文档并讨论对 NUV 波段增选或补充仿真的问题；定标产品生成流水线初步实现与 DFS 的接口；根据新版 0 级数据定义，更新了定标产品头文件关键词，更新了 CRDS 规则所用的 parkey 并测试成功，重新构建了 tpn 文件及初版 rmap 用于服务器端的测试；完成不同曝光时间的 dark 仿真，用于分析短 dark 与 bias 的对比、长 dark 的必要性等；在 gitlab doc 文档库上撰写定标参考文件方面的算法文档和 data model；召开 MCI、IFS 地面定标测试讨论会，讨论了地面测试细节、地面测试与在轨测试的相关关系；针对 NGC6397 中 IFS 视场内亮源目标进行仿真，IFS pipeline 处理形成 data cube，其中部分亮目标可以作为定标源用于辐射和位置定标；利用 M67 和 NGC188 分别对中窄带和中宽带滤光片进行了测光仿真分析，利用 Gaia G 星等和 BP-RP 颜色和 LAMOST 光谱估计 CSST 不同波段的星等、颜色信息。
13	观测需求编排	完成编排仿真软件部分核心功能的 Python 化，可在 windows 及 ubuntu 等平台下正常安装使用，后续将在此基础上提供一系列在线工具，包括望远镜坐标预测、目标可见性计算器、目标源初步编排等功能；基于各后端提供的观测申请字段信息，更新用户界面设计以支持可配置的在线申请字段调整与提交；实现望远镜轨道、望远镜观测天区的不同可视化方式及其灵活切换。


近期节点和计划

2023年上半年Cycle7 节点：

时间节点	计划任务安排	时间节点	计划任务安排
2023年2月底	与运控系统签订接口控制单	2023年7月底	模拟云环境下数据处理软件可运行
2023年3月底	与数据中心签订接口控制单	2023年7月底	数据处理系统软件 C7 原型
2023年3月底	配置项需求分析评审	2023年7月底	编排软件具备初步功能
2023年4月底	数据仿真软件 2.0.0 版本发布	2023年7月底	完成各模块一、二级流水线及定标流水线集成和运行
2023年5月底	配置项设计评审	2023年7月底	中期评估内审
2023年2-7月	C7 阶段各个软件配置项算法优化和功能完善	2023年7月底	收集鉴定件测试标定参数数据
2023年7月中	完成在轨定标方案评审		

编辑：王清

审核：高廷

签发：刘维海