

CSST 科学数据处理系统月报

第 15 期

主办方：中国空间站巡天望远镜科学工作联合中心

时间：2022 年 1 月 10 日

简讯

中国空间站巡天空间望远镜科学工作委员会 在国家天文台组织召开

2021 年 12 月 29 日，中国空间站巡天空间望远镜（简称 CSST）科学工作委员会采用线上和线下结合的方式召开了 2021 年度第四次会议，会务工作由 CSST 科学工作联合中心具体承担。来自全国 15 家单位的 20 名科学工作委员会成员参加了本次会议。

会议首先由科学工作委员会主任常进院士讲话，希望科学工作委员会、联合中心、科学中心通力合作把 CSST 的科学研究工作做好。之后，光学设施、科学工作联合中心及国家天文台科学中心、北京大学科学中心、长三角地区科学中心、粤港澳大湾区科学中心分别汇报了工作进展。会上还对科学数据处理系统进行了工作进展介绍；就《CSST 科学数据政策》当前的编制进展进行审议和讨论，并要求联合中心进行进一步修订；讨论了第二批二类课题的整体策划工作，需要联合中心深入了解并联合四个科学中心确定好课题方向。科学工作委员会副主任韩占文、何子山对 CSST 科学产出、科学活动及科学影响力等议题做出发言和建议。

载人航天工程空间科学首席专家、科学工作委员会顾问顾逸东院士介绍了第一批科学课题经费管理情况，提出了科研进展情况和第二批课题组织的问题，并提出相应具体建议，对联合中心的组织协调等工作提出要求。

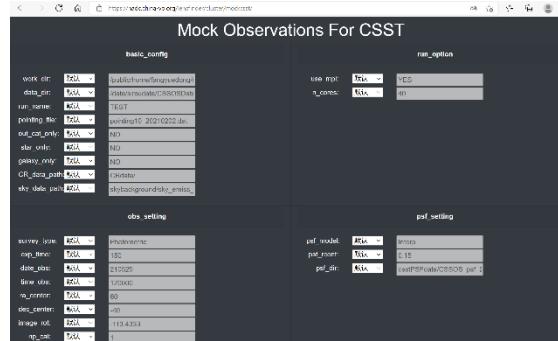
最后，科学工作委员会主任常进院士讲话，指出联合中心要做好联系与衔接工作，数据系统的软件要和科学中心反复细致沟通讨论，在疫情以后也开展充分的国际交流。

突出进展

观测数据仿真模块研制工作进展

CSST 仿真软件发布方面取得进展。CSST 仿真软件发布包括两种场景：一是源码发布，整个仿真软件的源代码和所需数据将面向全国天文学家发布，允许用户直接下载、使用和修改 CSST 仿真软件，其面向的是对仿真软件有强烈定制需求的用户；二是提供仿真数据生成服务，允许用户通过交互式图形界面生成仿真软件的配置文件，并使用国家天文台提供的计算服务器生成所需数据，其面向的是定制需求不高且生成数据量不大的科学用户。

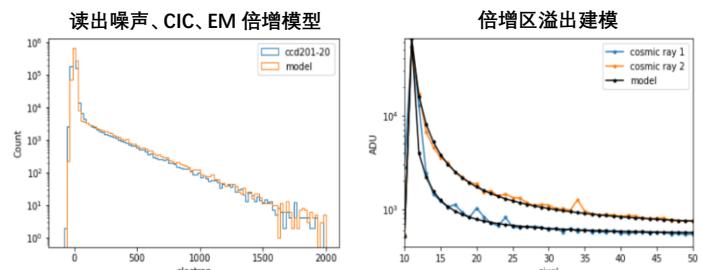
目前，场景一已经基本完成，并在实施最终测试阶段；场景二所需的用户界面已经完成开发，正在开展用户界面和仿真软件的集成测试以及计算服务器的压力测试。



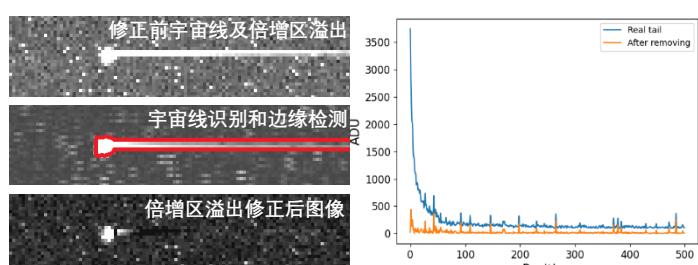
星冕仪数据处理模块研制工作进展

星冕仪科学数据处理团队对搭载星冕仪使用的 E2V CCD201-20 芯片的 Andor IXON 888 相机进行测试。测试结果为图像仿真和数据处理提供了重要参考。

在观测图像仿真方面，模块团队结合实测结果，对 EMCCD 芯片成像过程中转移、倍增、读出过程进行建模，完成了 EMCCD 图像的仿真工作。仿真图像包括读出噪声、暗电流、CIC 噪声、EM 增益、倍增区图像溢出、坏点、CTE、量子效率等效应。下图为读出噪声、CIC 噪声、EM 增益、倍增区溢出等效应仿真程序和实测图像的比较。



在数据处理方面，项目团队使用相机测试数据开展了宇宙线识别及修正研究。开发了通过宇宙线边缘及倍增区溢出效应的宇宙线识别程序，并通过正向建模完成了溢出效应的修正。



数据系统研制进展

序号	模块名称	月度进展
1	多色成像	在阿里云和云栖服务器上完成多色成像单次曝光流水线的布置和运行测试，达成了单幅图像 0 到 2 级数据处理的全过程；使用新仿真宇宙线数据进行机器学习，目前识别率已经超过 98%；与在轨定标对 2 级星表数据进行了检测；完成 Cycle4 阶段总结。
2	无缝光谱	开展 Cycle4 阶段总结工作；完成地面无缝光谱数据处理自动化管线的开发；基本完成使用正向建模方式对 CSST 仿真数据进行抽谱；初步完成一维无缝光谱分类与红移算法模块的开发；改进分类与红移测量工作包，并采用无缝光谱处理团队独立构建的仿真模拟光谱数据进行了测试。
3	数据流管理	完成无缝光谱初步数据库设计，接口访问和后端服务的实现；完成主巡天 2 级星表数据的数据库设计，写入和查询服务。
4	流水线运行管理	完成新版流水线引擎的集成调试；基于新流水线引擎，完成 Cycle3 原型的部署测试，简化流水线定义，提高了运行效率；梳理 Cycle4 原型的逻辑模块关系，设计并实现了支持分支结构的辅助模块，完成 Cycle4 原型的总体结构设计，待业务模块接入后，就可以进行 Cycle4 原型的流水线集成测试。
5	多通道成像仪	超深场图像仿真加入 SED；完成 Cycle4 总结报告；完成仿真单元测试文档。
6	积分视场光谱仪	完成 IFS 多次曝光叠加过程中的误差协变矩阵的完整计算和资源需求评估；完成 IFS Cycle4 任务总结。
7	星冕仪	开展宇宙线识别及修正研究；开发了通过宇宙线边缘及倍增区溢出效应的宇宙线识别程序，并通过正向建模完成溢出效应的修正。
8	太赫兹	完成基于遗传算法的太赫兹模块观测任务初步编排工作，在以观测目标完成度最大化为目标的情况下，获得了高达 20% 的性能提升；完成 Cycle4 阶段总结报告。
9	天体测量	精细创建天体测量数据处理软件模块工作，完成各个工作模块的软件详细设计；正在进行密集星场图像目标检测算法及定心算法对比测试工作。
10	观测数据仿真软件	为确保 CSST 主巡天仿真程序的顺利发布，本月对仿真程序各模块的功能实现做了进一步的梳理和测试工作，完成了包含暗端星表改正、像场畸变建模、饱和溢出和无缝光谱分光等各模块的更新升级，添加了自定义星表、主巡天时间截和自行等模块，优化了发布程序的配置接口，开展了仿真程序的稳定性测试；同时，仿真团队正积极组织开发仿真程序在线发布的用户接口和编写用户使用手册，为仿真程序的发布和 CSST 科学预研究的顺利开展提供保证。
11	数据可用性标记	进行 Cycle4 流水线测试工作；完成 Cycle4 总结报告及 QC1 的检测参数报告。
12	在轨定标	针对第一个仿真出的定标星场的不同曝光时间的数据处理结果进行检测分析，包括多色成像数据的定标零点、天光、PSF 等，以及抽取的无缝光谱，形成相应技术报告；增选了两个较为密集的星场，生成了包括天测信息在内的输入星表，目前正在使用最新版的仿真程序由仿真团队开展星场的图像仿真；讨论大尺度平场改正的问题；进一步打磨完善在轨定标方案书；完成 Cycle4 阶段总结。
13	观测需求编排	编排文件生成模块设计并初步实现编排流程的原型系统；初步确定编排权重接口形式；根据数据库观测记录表结构（v1.4）和编排软件输出结果，调研并初步设计上传到运控的观测编排文件格式；一般观测申请模块更新了一般观测申请用户界面，修正了发现的 bug；警报模块通过设计 COMET 的消息转发处理插件，完成 KAFAK 与 COMET 的整合；观测计划可视化部分使用 flex 布局、Less、js+jquery 等进行数据大屏可视化的原型初步设计，包括可视化内容选择、屏幕区域划分等。

近期节点和计划

2021 年下半年 Cycle4 节点：

时间节点	计划任务安排	时间节点	计划任务安排
8 月中	完成多色成像流水线原型评估	8-12 月	各数据处理模块流水线搭建，实现无缝光谱流水线原型
8 月底	完成 0 级数据结构定义文档评审	8-12 月	公共模块的定义和开发
9 月	完成阿里云环境测试	8-12 月	开展系统软件设计
9 月底	完成系统分析设计方案评审	10 月	完成系统功能需求评审（顺延）
8-10 月	完成 1-2 级数据定义并评审	12 月	完成在轨定标策略方案书评审
8-10 月	完成在轨定标流水线搭建	12 月底	完成仿真软件 V1.0 版本的发布
8-12 月	完成科学需求编排软件原型搭建		