

简讯

CSST 科学数据处理系统召开 2022 年度第二次调度会

8 月 30 日，CSST 科学数据处理系统 2022 年度第二次调度会以线上线下相结合的方式在中国科学院国家天文台召开。来自国家天文台、上海天文台、紫金山天文台和南京天光所的指挥线领导和技术骨干成员共 37 人参加本次会议。

会上，刘超研究员首先对上次调度会的闭环情况进行了介绍；然后四家单位的技术负责人介绍了各单位开展的研制进展情况，并提出了工作中遇到的技术难题和需要协调的问题，主要包括仿真数据使用政策问题和硬件与软件的协同，重新梳理了时间节点与硬件研制节点协调同步。各单位指挥线领导表示严格按照项目研制计划积极配合并推进工作的开展。本次调度会按计划顺利召开，达到调度会的目的，共同推进 CSST 科学数据处理系统研制项目的顺利进行。

突出进展

一级流水线封装研制工作进展

Cycle 5 的 CSST 一级流水线封装任务已经顺利完成，封装了包括主巡天多色成像（MBI）的仪器效应改正、位置定标、像场畸变改正、流量定标以及测光共 5 个功能模块。封装后的主巡天多色成像流水线已经在紫金山实验室提供的计算节点上完成了 2 平方度仿真图像的处理。Cycle 6 的封装工作已于 8 月 23 日全面开展，对包括主巡天多色成像（MBI）和无缝光谱（SLS）、多通道成像仪（MCI）、积分光谱仪（IFS）、星冕仪（CPIC）、太赫兹模块（THZ）在内的 6 个设备共 20 余个功能模块以及数据可用性标记（QC）模块等进行封装，将于近期完成文档和单元测试等工作。预计封装工作将于 10 月底完成，整个一级流水线将在 11 月底完成与 DFS 的对接。

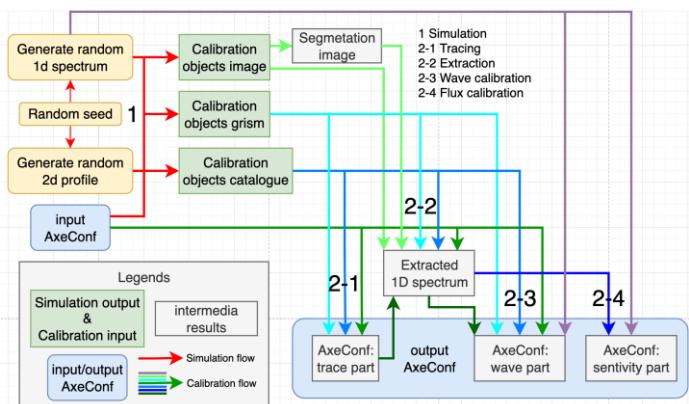
CSST 一级流水线采用 Python 作为主要编程语言，PEP8 作为代码风格标准，所有 Python 代码最终需要通过 PyFlake8 进行缺陷（警告和错误）检测。所有团队的

代码均在 gitlab 上进行管理，以 https://csst-bao.ac.cn/code/csst-l1/csst_proto 为封装原型，每个流水线功能模块单独进行打包封装。除提供 Python 源代码和打包所需数据、文件以外，所有包都需要定义一个 `top_level_interface` 模块用于导入需要接入流水线中的类和方法，并将调用案例/使用说明写在 `README` 中。最终所有功能模块由 `csst_l1` 进行集成和调度。

无缝光谱数据处理模块研制工作进展

无缝光谱数据处理模块团队实现了基于图像数据的色散流程，并新开发了基于 IFU 数据的色散流程。这种方法额外考虑到了色散函数和目标源光谱随图像像素（或物理位置）的波动，从而可以正确的对于大尺度、弥散源进行无缝光谱建模。测试结果表明，新方法对于图像数据不完整（例如图像过亮、受到星芒影响）的源，可以成功进行拟合，并得到比 HST 无缝光谱正向建模软件 Grizli 残差更小的建模结果。

基于自主开发的色散流程，我们首先实现对无缝光谱色散函数的正向计算和反向拟合，并最终完成 CSST 无缝光谱色散与定标框架的搭建（如下图）。整个框架实现了从已知色散函数生成模拟数据，进行完整的定标流程，并与输入的色散函数进行比较的完整过程。首先随机生成定标源星表和其二维轮廓，然后根据输入的色散函数，生成仿真模拟定标数据。随后对此数据进行完整的定标流程，依次得到轨迹项（trace calibration）、波长项（wavelength calibration）和响应曲线（flux calibration）的信息，最后输出得到定标后的色散函数。该框架搭建完成为未来 CSST 色散函数的大规模计算和反向定标奠定了坚实基础。



数据系统研制进展

序号	模块名称	月度进展
1	多色成像	完成多色成像单次曝光流水线的封装，正在开展代码规范和注释更新；完成 2 平方度数据处理，流水线开发人员开展了基本的结果检查，目前除零点未达标外，定标和极限星等都接近或到达计划指标；正在利用新数据开展合并星表，PSF 重构等算法测试。
2	无缝光谱	开展一级流水线调试和封装工作；协助梳理无缝光谱在轨定标内容；与无缝光谱对应天区的 mosaic 构建取得实质进展；进一步检验基于发射线测量的星系红移精度。
3	数据流管理	在紫金山实验室提供的 3 节点搭建完成 DFS 测试环境，初步满足后续交叉测试需要；测试并验证并行化方式提高星表导入性能；交叉协作进展：优化图像基础处理代码，满足管线开发要求；提高千万级星表合并性能。
4	流水线运行管理	在紫金山实验室测试环境进行了流水线主程序与流水线运行系统的 python 接口 pyscalebox 的测试部署；完成二级流水线消息转发主模块的研发，并在紫金山实验室测试环境部署，通过各业务模块进行测试。
5	多通道成像仪	MCI 仿真加入天体测量误差模块；开展一级流水线封装。
6	积分视场光谱仪	启动搭建 IFS 一级流水线，改动之前流水线中各模块相对独立的运行模式，模块之间接口由流水线统一调度；RSS 流水线需要将 IDL 程序转为 python，CUBE 流水线进行中；完成 IFS 2 级数据文件定义修订稿。
7	星冕仪	完成星冕仪一级流水线进行了封装并上传至 gitlab；对地面测试需求报告闭环并参加评审；开展了星冕仪与 DFS 系统的对接工作。
8	太赫兹	初步完成太赫兹模块 1 级数据文件格式和内容定义；持续优化太赫兹模块数据处理 pipeline，对相关代码按照总体制定一级流水线原型和封装形式开展更新与优化工作。
9	天体测量	开展天体测量一级流水线的封装对接工作；将太阳系移动天体仿真模块添加到 CSST 图像仿真软件中；对接 CSST 1 级数据产品定义。
10	观测数据仿真软件	完成地气光光、黄道光与哈勃数据的对比；植入平场灯简单模块，着手数据搜集；图像读入仿真已实现 fits 格式读入，并应用于仿真；宽、深场仿真已完成 10 组 PSF 数据，半解析星表完成过半。
11	数据可用性标记	完成第一次精密稳像小天区的分析，稳像系统对于密集星场的适配性需要特殊考虑；正在进行 QC 模块程序的 bug 修复，目前已经更新为可安装的情况；正在进行 1 级数据定义文档的修改，之后会进行第二版更新。QC2 完成第一版 2 级数据定义；讨论了多色成像和无缝光谱模拟数据的情况，准备开始 QC2 模块的程序编写。
12	在轨定标	基于仿真的定标参考数据进行 bias 图像的宇宙线探测试验，后续仿真时定标参考图像仿真 5 幅即可满足测试需求；对 3 个星场的输入星表进行了历元转换，使用新版（1.0.4）仿真软件生成了定标参考图像随时间变化的仿真数据；基于 HST 规则文件和定标参考文件，测试实现了在本地重构生成新版规则文件；正在修改 CRDS 代码以构建 CSST 规则文件；基于 Gaia DR3 进行早型星定标星场的选取，后续还要考虑使用其他星表进行选取；通过对传统辐射定标星表的综合整理，形成定标星库和查询分析软件。
13	观测需求编排	讨论并形成基于 JSON 的编排格式初稿；基于 QUASAR+VUE 搭建 CSST 编排可视化前端框架；正在进行警报响应审批流程开发。

近期节点和计划

2022 年下半年 Cycle6 节点：

时间节点	计划任务安排	时间节点	计划任务安排
22 年 7 月中旬	系统分析设计说明外审闭环	22 年 9 月	协调与数据中心的外部接口
22 年 7 月中旬	Cycle5 任务总结/Cycle6 计划	22 年 10 月	配置项需求规格说明书评审
22 年 7 月中旬	V1.1.0 仿真软件发布	22 年 10 月	协调与运控系统的外部接口
22 年 7 月-12 月	不同探测器地面测试	22 年 12 月	配置项软件设计评审
22 年 7 月-12 月	观测需求编排外部接口定义	22 年 12 月	实现所有模块统一程序框架并完成主要功能封装
22 年 7 月-12 月	开展算法科学验证	22 年 12 月	二级数据流水线应用接口发布
22 年 8 月底	1 级数据定义	22 年 12 月	Cycle6 任务总结