Web-based HXMT Data-Analysis Platform

Yu Hu[1], Jianli Liu[1], Qingbao Hu[1], Shuang Wang[1], Hongmei Zhang[1] and Fazhi Qi[1],

[1]Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China

E-mail: huyu@ihep.ac.cn

Abstract.

The HXMT satellite is China's first space astronomy satellite. It is a space-based X-ray telescope capable of broadband and large-field X-ray sky surveys, as well as the study of high energy celestial objects such as black holes and neutron stars, focusing on short-term temporal variations and broadband energy spectra. HXMT data are mainly processed by HXMTDAS. However, the users, especially new users, often encounter difficulties in processing data with HXMTDAS. This work presents a user-friendly platform for processing and analysing HXMT data. It integrates the functionalities of the existing packages HXMTDAS and Xspec in the same toolbox. Its graphical interface, based on Jupyter Notebook widgets, combines an interactive approach for data analysis with a powerful environment designed to link scientists with large-scale facilities. Using virtualisation and cloud technology, the platform also allows multiple users to access and analyse data remotely. This platform can also be used for education and open science.

1 Introduction

The Hard X-ray Modulation Telescope (HXMT)[1], named "Insight", is China's first X-ray astronomy satellite, as shown in Figure 1. The mission provides high-resolution imagery across a wide spectrum of X-ray energies. It has been six years since Insight-HXMT was commissioned in January 2018. A number of exciting scientific results have been achieved in exploring the dynamic and energetic universe, such as the first identification of an X-ray burst from a magnetar as the counterpart of FRB 200428[2], the discovery of an accelerating and precessing jet via QPO detection and the spectral measurement in a broad energy band[3], the first evidence of observing a MAD via monitoring a complete magnetic transfer in accretion flow[4], the record-breaking discovery of the fundamental CRSF with the highest energy[5], detection of the brightest GRB in history[6], etc. Insight-HXMT is still operating in orbit normally and is expected to continue its scientific operations for several more years.

The HXMT User Data Analysis Software (HXMTDAS)[7] is primarily designed to analyse pointed observational data from the HXMT satellite, including on-axis and off-axis observations, to produce energy spectra, light curves, and energy response files. The HXMTDAS code is written in C++ and is primarily performed using the HEAsoft[8] and released as a part of HEAsoft. This means that users have to install the entire HEAsoft and the part of HXMTDAS. Then they have to configure complex running environments and resolve dependencies. This is very unfriendly to new users. Furthermore, the data should be downloaded locally. After the proprietary period, the data products will be open to the public. To make it easier for the public to participate in scientific exploration, we need to find ways to help the public access and analyse data of interest remotely. Jupyter(https://jupyter.org/) is particularly advantageous for this purpose. Several institutions have chosen Jupyter for open source or remote data access[9, 10]. In this paper, we present an interactive data analysis platform based on web technology and HXMTDAS. Using containers, virtualisation technology, and JupyterLab, this platform allows users to work locally or remotely and perform interactive data analysis via a web browser. The platform is out-of-the-box and operating system independent, eliminating the need for complex software installations and environment configuration steps.

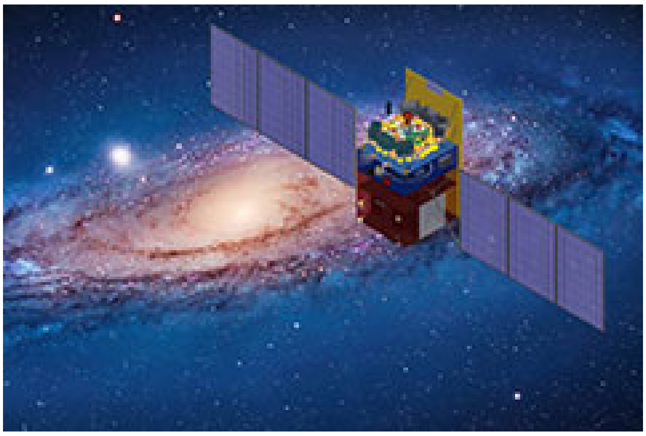

Figure 1: Insight-HXMT

2 The Insight-HXMT and data analysis software package

The HXMT satellite was launched on 15 June 2017. It is a low Earth orbit telescope with an altitude of 550 km and an inclination of 43 degrees. In order to fulfill the requirements of the broadband spectral and fast variability observations, three payloads are configured on board the HXMT satellite. The high energy X-ray telescope (HE; [11]; 20-250 keV, 5100 cm^2) contains three detector boxes with Swept Charge Devices (SCD), which work in a continuous read-out mode, in contrast to more traditional CCD detectors with fixed exposure times, giving it greater time resolution. The medium energy X-ray telescope (ME; [12]; 5-30 keV, 952 cm^2) uses 1728 Silicon-PIN diodes for detecting X-rays. The low energy X-ray telescope (LE; [13]; 1-15 keV, 384 cm^2) consists of two concentric rings of a total of 18 NaI/CsI phosphor sandwich (phoswich) scintillation detectors with collimators for 1.41° x 5.71° FOV in different orientations. The three payloads are seamlessly integrated onto a shared support structure, allowing them to maintain the same pointing direction. This arrangement allows for the simultaneous observation of the same source. Each payload has a collimator to define its field of view (FOV). The main scientific objectives of Insight-HXMT are: (1) to scan the Galactic Plane to find new transient sources and to monitor the known variable sources, (2) to observe X-ray binaries to study the dynamics and emission mechanism in strong gravitational or magnetic fields, and (3) to monitor and study gamma-ray bursts (GRBs) and gravitational wave electromagnetic counterparts (GWEM).

The HXMTDAS package is designed for processing HXMT data and extracting scientific products such as energy spectra, light curves, redistribution matrix files (RMF), and background files. It comprises several tasks, each dedicated to a specific step in data analysis. These tasks are crafted in the style of ftools[14] and are fully compatible with HEASoft.

The HXMTDAS package processing procedures are structured into three distinct stages: calibration, screening, and the extraction of high-level scientific data products. The input for these processes is an observation (referred to as the exposure ID) derived from the HXMT Level 1 (1L) data products, which are in FITS format. Files of L1 data are arranged in a hierarchical directory tree with the cue of Observation ID - Exposure ID. The root directory of every product is named by the Observation ID. There are three types of catalogues: List Files, Exposures Directory, and Auxiliary Data for the whole observation. The List Files contain all file information of the archived observation and offer a detailed list of exposures specific to that observation. Each exposure has a separate directory to store the scientific and engineering data collected in the exposure period. The Auxiliary Data Directory stores attitude data, orbit data, and auxiliary data originating from the ground segment.

3 Design and implementation of the platform

The HXMTDAS package contains comprehensive data processing tools for X-ray astronomical data analysis, but it is difficult to install and can only be used via the command line, which is not conducive to

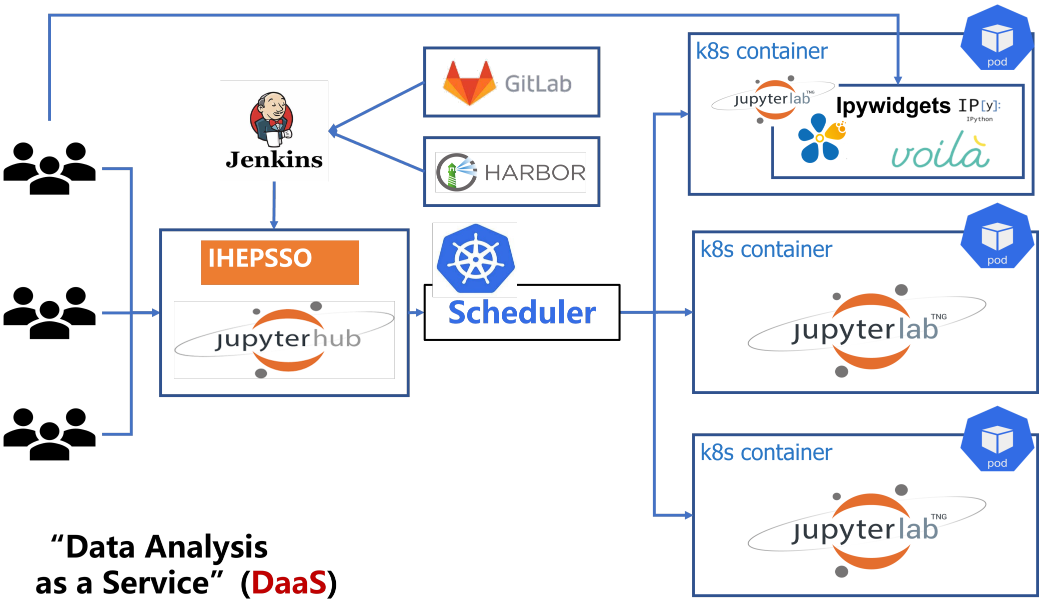

Figure 2: The architecture of the platform

new users and open science. We therefore designed and implemented the web-based interactive Insight-HXMT data analysis platform, based on cloud technology, which allows users to perform step-by-step data analysis with the help of a guide, as shown in Figure 2.

The platform visualises the data processing flow of HXMT and provides interactive services to users in the form of web pages. The front-end of the platform is based on the Jupyter ecosystem, which provides users with software, algorithm development, and data processing environments through a web browser. The platform is also out-of-the-box, so there is no need to download and install software or configure the computing environment. Services are always up to date, and new features are available to users as soon as they are released. The resources of the platform are flexible and scalable, and can be configured according to the computing tasks to be performed. The platform removes the operating system dependency of HXMTDAS and can be accessed through a browser on common operating systems. The platform provides both a terminal and a friendly user interface suitable for users of different skill levels.

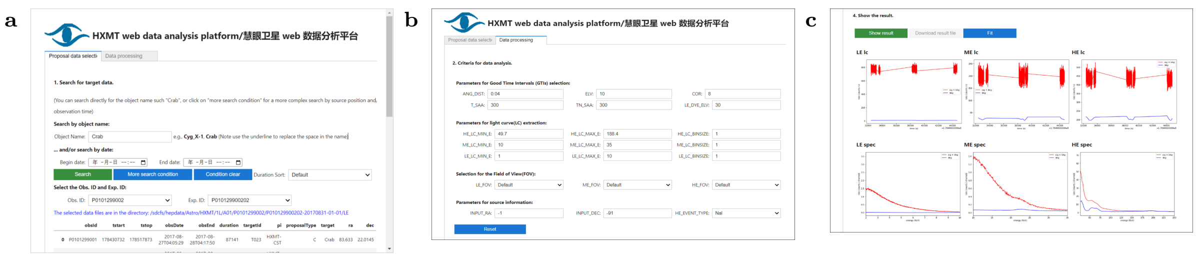

Figure 3: The Graphical User interface of the platform

3.1 The architecture

This platform works with Python 3.8 and is licensed under the GNU General Public License v3.0. The package is dependent on and built around HXMTDAS. The platform needs to support multiple users processing data simultaneously. The JupyterHub (https://jupyter.org/hub), a multi-user Hub that spawns, manages, and proxies multiple instances of the single-user Jupyter notebook server, was used to serve Jupyter notebooks for multiple users. JupyterHub can provide a separate Jupyter environment for each user. The platform's user authentication is based on IHEPSSO (IHEP single sign-on interface), which is an authentication system based on OAuth2.0[15]. Users can log in to the platform with IHEPSSO through the JupyterHub page. To deploy dynamic servers on the cloud and support more users, the JupyterHub runs on Kubernetes, which is an open source system for automating deployment, scaling, and management of containerized applications. Docker (https://www.docker.com) containers are used to encapsulate computing environments such as HXMTDAS and Xspec. Kubernetes (https://kubernetes.io) is responsible for centralized management of heterogeneous resources, on-demand scaling of resources, and container orchestration. JupyterLab provides the main user interface, which is a web-based integrated development environment that can be used to write notebooks, operate terminals, edit markdown text, open interactive modes, and other functions. The data analysis application was implemented by the IPywidgets package (https://github.com/jupyter-widgets/ipywidgets). The graphical user interface can be rendered into a web application using Voila, a Jupyter extension that provides an interactive Dashboard. In addition, GitLab (https://gitlab.com) was used for code version control, Harbor (https://goharbor.io/) was used for container image management, and Jenkins (https://www.jenkins.io) accelerated the entire development and deployment process.

3.2 The application

The web-based HXMT data analysis platform provides users with a single user-friendly and interactive application, combining the above packages while providing browser-based data visualisation tools, as shown in Figure 3. The application follows the Model View Controller (MVC) pattern. In the pattern, Model is the backend that contains all the data logic, View is the frontend or graphical user interface, and Controller is the brain of the application that controls how data are processed and displayed. This pattern allows the application to be scalable, maintainable, and easy to expand. In our application, HXMTDAS and Xspec are integrated into the Model and defined as methods that can be called. The data model is also defined inside the Model. The Controller is responsible for responding to user actions on the interface, controlling the flow of data to the model object, calling the model to process the data, and updating the view when the data changes. The View defines the GUI and presents the data. The GUI of the application is divided into two tabs (Fig. 4), separated into six groups: search for target data, criteria for data analysis, select the commands to execute, perform the data processing, show the result, and fit the result.

3.3 The analysis workflow

For scientific users, data analysis starts from the L1 data product. An L1 data product includes all the event data, auxiliary data, and other related data during an observation proposal period. The data in the product is divided based on exposure time. Each payload has its own data processing independent pipeline. Typically, the data analysis process from HXMT 1L data products to level 2 (2L) data products consists of four steps, as shown in Figure 4:

- Step 1: Search for data according to specific criteria using requests.
  Users can filter the data they want to analyse from the database. The filtering criteria can be one or a combination of source name, observation date, source coordinates, proposal type, proposal number, or any combination thereof. Data that meets the filter criteria will be listed in the table below.

- Step 2: Set the parameters for data reduction.
  Before analysing the data, the users need to set up a number of analysis parameters, including Good Time Intervals selection conditions, parameters for generating light curves, field of view selection parameters, and source-related parameters.

- Step 3: Confirm Good Time Intervals and perform data reduction.
  After setting the parameters, users can start the first stage of data processing and confirm the Good Time Intervals. After confirming the Good Time Intervals, users can proceed to the second stage of data processing and obtain the final data products.

- Step 4: Show and download the data reduction results.
  Users can visualise the generated data products, perform simple power law fitting, and download the data products.

The 2L data product contains several types of data, such as energy spectra, light curves, redistribution matrix files (RMF), and background files, which can be used for further scientific analysis.

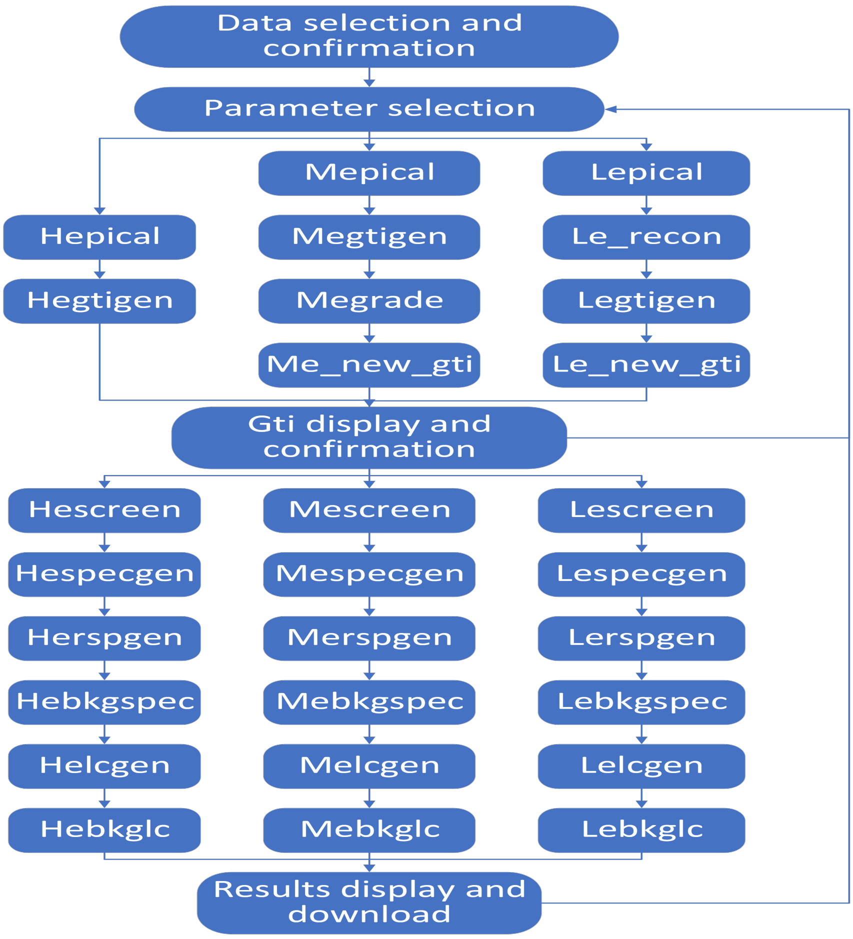

Figure 4: The analysis workflow of Insight-HXMT

4 Conclusions and Future Work

The web-based HXMT data analysis platform is a versatile tool for HXMT satellite data analysis. It puts HXMTDAS and other X-ray astronomy data analysis tools at researchers' fingertips in an intuitive interface and exploits the Jupyter framework to provide either local or remote operation seamlessly. This platform can host multiple users and perform data analysis from signal or multiple exposure IDs to take advantage of the computing power of cluster resources.

This platform provides a way of overcoming several problems by bridging remote access, cluster computing, and a user-friendly interface, consequently lowering the barrier of access for new users. The public can also use the platform to analyse open data, thereby promoting open science.

5 Acknowledgements

This work was supported by IHEP, CAS, through the Technological Innovation Program of the Institute of High Energy Physics of the Chinese Academy of Sciences under Grant No. E3545JU210. This research used resources of the National High Energy Physics Science Data Center (NHEPSDC).

References

[1] Zhang, SN., Li, T., Lu, F. et al. Overview to the Hard X-ray Modulation Telescope (Insight-HXMT) Satellite. Sci. China Phys. Mech. Astron. 63, 249502 (2020).

[2] Li, C.K., Lin, L., Xiong, S.L. et al. HXMT identification of a non-thermal X-ray burst from SGR J1935+2154 and with FRB 200428. Nat Astron 5, 378–384 (2021).

[3] Ma, X., Tao, L., Zhang, SN. et al. Discovery of oscillations above 200 keV in a black hole X-ray binary with Insight-HXMT. Nat Astron 5, 94–102 (2021). https://doi.org/10.1038/s41550-020-1192-2

[4] Bei You et al. ,Observations of a black hole x-ray binary indicate formation of a magnetically arrested disk.Science381,961-964(2023).DOI:10.1126/science.abo4504

[5] Ling-Da Kong et al. Insight-HXMT Discovery of the Highest-energy CRSF from the First Galactic Ultraluminous X-Ray Pulsar Swift J0243.6+6124 . ApJL 933 L3 (2022)

[6] An Zheng-Hua, Antier S; Bi Xing-Zi et al. Insight-HXMT and GECAM-C observations of the brightest-of-all-time GRB 221009A . arXiv230301203A (2023).

[7] Zhao, H. 2020, Astronomical Data Analysis Software and Systems XXIX, 527, 469

[8] Blackburn, J. K. 1995, in ASP Conf. Ser., Vol. 77, Astronomical Data Analysis Software and Systems IV, ed. R. A. Shaw, H. E. Payne, and J. J. E. Hayes (San Francisco: ASP), 367.

[9] B. M. Randles, I. V. Pasquetto, M. S. Golshan and C. L. Borgman, "Using the Jupyter Notebook as a Tool for Open Science: An Empirical Study," 2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL), Toronto, ON, Canada, 2017, pp. 1-2, doi: 10.1109/JCDL.2017.7991618.

[10] Lasser, J. Creating an executable paper is a journey through Open Science. Commun Phys 3, 143 (2020). https://doi.org/10.1038/s42005-020-00403-4

[11] C. Liu, Y. Zhang, X. Li et al., Sci. China Phys. Mech. Astron. 63, 249503 (2020)

[12] X. Cao, W. Jiang, B. Meng et al., Sci. China Phys. Mech. Astron. 63, 249504 (2020)

[13] Y. Chen, W. Cui, W. Li et al., Sci. China Phys. Mech. Astron. 63, 249505 (2020)

[14] Blackburn, J. K. et al. FTOOLS: A general package of software to manipulate FITS files. 1999ascl.soft12002B

[15] B. Leiba, "OAuth Web Authorization Protocol," in IEEE Internet Computing, vol. 16,no. 1, pp. 74-77, Jan.-Feb. 2012, doi: 10.1109/MIC.2012.11.