

February 2, 2023

The data center for the X-ray spectrometer/imager STIX onboard Solar Orbiter

Hualin Xiao¹, Shane Maloney², Säm Krucker¹, Ewan Dickson⁴, Paolo Massa⁶, Erica Lastufka¹, Andrea Francesco Battaglia^{1,3}, László Etesi¹, Nicky Hochmuth⁸, Frédéric Schuller⁷, Daniel F. Ryan¹, Olivier Limousin⁵, Hannah Collier^{1,3}, Alexander Warmuth⁷,
and Michele Piana^{9,10}

¹ University of Applied Sciences and Arts Northwestern Switzerland (FHNW), 5200 Windisch, Switzerland

e-mail: hualin.xiao@fhnw.ch

² Astronomy and Astrophysics Section, School of Cosmic Physics, Dublin Institute of Advanced Studies, 31 Fitzwilliam Place Dublin 2, D02XF86, Ireland

³ ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland

⁴ University of Graz, Universitätspl. 3, 8010 Graz, Austria

⁵ Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191 Gif-sur-Yvette, France

⁶ Department of Physics & Astronomy, Western Kentucky University, Bowling Green, KY 42101, USA

⁷ Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany

⁸ Ateleris, Badenerstrasse 13, CH-5200 Brugg Switzerland

⁹ MIDA, Dipartimento di Matematica, Università di Genova, via Dodecaneso 35, I-16146 Genova, Italy

¹⁰ Istituto Nazionale di Astrofisica, Osservatorio Astrofisico di Torino, Via Osservatorio 20, I-10025 Pino Torinese, Italy

February 2, 2023

ABSTRACT

Context. The Spectrometer/Telescope for Imaging X-rays (STIX) onboard Solar Orbiter observes solar X-ray emission in the range of 4–150 keV and produces spectra and images of solar flares over a wide range of flare magnitudes. During nominal operation, STIX continuously generates data. A constant data flow requires fully automated data processing pipelines to process and

analyze the data, and a data platform to manage, visualize, and distribute the data products to the scientific community.

Aims. The STIX Data Center has been built to fulfill these needs. In this paper, we outline its main components to help the community better understand the tools and data it provides.

Methods. The STIX Data Center is operated at the University of Applied Sciences and Arts Northwestern Switzerland (FHNW) and consists of automated processing pipelines and a data platform. The pipelines process STIX telemetry data, perform common analysis tasks and generate data products at different processing levels. They have been designed to operate fully automatically with minimal human intervention. The data platform provides web-based user interfaces and application programmable interfaces for searching and downloading STIX data products.

Results. The STIX Data Center has been operating successfully for more than two years. The platform facilitates instrument operations and provides vital support to STIX data users.

Key words. Solar flares – Data platform – STIX data products – X-ray imaging – Data processing pipeline

1. Introduction

Solar Orbiter is a Sun-observing mission led by ESA with contributions from NASA that addresses the interaction between the Sun and the heliosphere. It was launched on 10 February 2020, with a nominal mission duration of seven years and a planned extension of three years (Müller et al. 2020). It carries ten instruments for comprehensive remote sensing and in-situ observations. Solar Orbiter’s closest approach to the Sun is 0.28 AU and as the mission progresses, its orbit will shift out of the ecliptic, providing better views of the solar poles (Forveille & Shore 2020). Solar Orbiter’s Spectrometer Telescope for Imaging X-rays (STIX; Krucker et al. 2020) observes X-rays from 4 to 150 keV. Its main scientific objective is to study the extremely hot solar plasma and accelerated electrons produced during solar flares. STIX consists of 32 sub-collimators, each with a pair of slightly offset grids that encode spatial information in a Moiré pattern which is detected by a pixelated cadmium telluride (CdTe) detector. Using this method, images with a few arcsec angular resolution can be produced on the ground using indirect Fourier-based imaging techniques. STIX’s spectroscopic detectors enable these images to be produced in any energy range between 4 to 150 keV as well as simultaneously providing spectra of the emission. Hence, STIX provides observations of the intensity, spectrum, dynamics, and location of accelerated electrons and heated plasma during solar flares. For more information on STIX instrumentation and its scientific capabilities, see Krucker et al. (2020).

STIX observes continuously during normal operations. The data are compressed and binned in time and energy to save onboard storage. Low-latency telemetry data are sent to Earth during passes of the ground station (typically one pass a day) and are processed automatically by the STIX Data Center into housekeeping and quicklook data products. These are crucial for monitoring the

health and performance of the instrument as well as requesting higher telemetry science-quality data.

The STIX Data Center runs the quicklook data through a flare detection algorithm, which identifies periods of scientific interest for which science-quality data should be requested from the instrument. To save telemetry these data requests can be made at coarser time, energy, and pixel binning than is used natively onboard. The severity of the binning is chosen depending on the apparent scientific importance of the observations. Human operators can manually review and edit the data requests compiled by the STIX Data Center and can include requests for additional periods of interest. Despite this human involvement, the automatic pipelines of the STIX Data Center carry a large fraction of the work and enable STIX operations to be conducted by only a small number of people. Higher-telemetry science data are sent back less frequently (almost weekly) and only for periods of interest (e.g. flares) that are identified on the ground via the low-telemetry quicklook data.

Once science data are received, the STIX Data Center automatically processes them into higher-level data products. To help users identify and understand the data products, the STIX Data Center performs common analyses for flaring data, such as the determination of coarse flare locations, imaging, and spectroscopy. This is particularly helpful for STIX as its indirect imaging technique is not intuitive, and hence its data can be confusing to new users. The STIX data center provides various web-based interfaces for viewing data products, metadata created by data processing pipelines, and results from common analyses.

In this paper, we describe the processing pipelines, core analysis algorithms, data products, and other tools used and provided by the STIX Data Center. The aim is to help users better understand the Data Center and hence how to best utilize it and the STIX instrument.

2. STIX telemetry data products

STIX generates many different types of raw telemetry packets. However, only three categories are relevant to scientific users: housekeeping, quicklook, and science data. Housekeeping and quicklook data are directed to the low-latency data stored in the spacecraft's solid-state mass memory (SSMM) and sent to the ground with the highest priority. Except for the raw pixel data product, all count- and trigger-based values in quicklook and science data products are compressed with an integer compression algorithm (Krucker et al. 2020).

2.1. Housekeeping data

The housekeeping data monitor the instrument's status and performance in order to ensure it functions properly. STIX generates housekeeping packets continuously while being powered. The housekeeping packets include details such as the instrument's temperature, voltages, currents, CPU usage, the status of the attenuator switches, trigger rates, file system information, and aspect sys-

tem readouts (Warmuth et al. 2020; Krucker et al. 2020). STIX typically generates a housekeeping packet every 64 seconds.

2.2. Quicklook data

The quicklook data are only generated when STIX is in the nominal observation mode, as they require the detectors to be powered and operating. There are five types of quicklook data:

- **Quicklook lightcurves** contain time series of 4-second-integrated, detector-summed counts in five energy bands. Furthermore, they include the corresponding detector-summed triggers, and the rate control regime state (i.e., attenuator state. See Krucker et al. 2020). Note that the quicklook lightcurves do not include time-series from STIX’s two special detectors, the background monitor (BKG) and the coarse flare locator (CFL). It is worth mentioning that STIX does not correct for detector dead time, absorption of X-rays by the entrance and grids, impacts of the presence of the attenuator, or rate control regime states onboard. Hence the effects of these phenomena are present in the quicklook lightcurves.
- **Quicklook background lightcurves:** STIX uses the background monitor, a detector with a special grid aperture, to monitor both the X-ray background and the intense unattenuated X-ray fluxes from large solar flares. It consists of an open front grid window and a rear grid window that is fully opaque except for six small openings. Quicklook background lightcurves contain counts and triggers of all the detector pixels integrated over 8 seconds in the same five energy bands as the quicklook lightcurves.
- **Quicklook variance data** are the on-board computed variance of 40 successive detector-summed count rates based on 0.1-second integration.
- **Quicklook spectra** are produced for each detector by integrating over all its pixels and accumulated over 32 seconds. During nominal operations, a quicklook spectrum for each detector is generated every 1024 seconds.
- **Calibration spectra** are high energy resolution, per pixel, energy spectra in ADC units (ADU) accumulated for photons emitted by ^{133}Ba radioactive sources located within the STIX instrument. STIX generates a calibration spectrum for each pixel every 24 hours during normal operations. These are used to calibrate the gain of each pixel on the ground by observing the locations of the well-known spectral line peaks. If the gain is seen to shift, a new ADU to keV conversion table is uploaded for each pixel from the ground.

2.3. Science data

Science data products are only generated and down-linked in response to a data request from the ground. They do not present a continuous record of solar activity because of the limited bandwidth and onboard storage. Instead, they cover scientifically interesting periods identified on the ground from the low-telemetry quicklook data. Science data are generated from the pixel data stored in the STIX onboard archive memory. They are rebinned in time, energy, pixels, and detectors in

accordance with the parameters of the data request. These parameters are chosen on the ground on a case-by-case basis to optimize both telemetry efficiency and scientific return. Due to the limited storage available onboard, raw pixel data is overwritten after a few months. Therefore, bulk science data (BSD) requests must be made within the following weeks of observations being made.

STIX can generate six different types of science data products:

- **Raw pixel data** are the least processed data product. Provides time-binned pixel counts in the highest available time and energy resolution, generally with a dynamic accumulation time between 0.5 s and 20 s depending on solar activity. The raw pixel values are not integer-compressed, and hence require a large amount of telemetry. Therefore, they are rarely requested and are used primarily for testing and verification.
- **Compressed pixel data** are based on the raw pixel data product. They combine the highly-resolved pixel counts into larger user-defined time and energy bins. In addition to the re-binning, they are integer-compressed onboard before being sent to the ground. This is the most common BSD product and is well suited for scientific analysis.
- **Summed and compressed pixel data** are also based on the raw pixel data product. Like the compressed pixel data, they combine the raw pixel counts into larger user-defined time and energy bins but also sum the counts over multiple pixels. The pixel sums can be configured by telecommand. By default, the top and bottom pixels in the same column are summed, reducing the apparent number of pixels per detector to 4 (Krucker et al. 2020). Therefore, it is a suitable option when telemetry is limited. In addition, the summed pixel data values are integer-compressed.
- **Visibility data**: Each STIX subcollimator measures a spatial Fourier component on the sky which can be combined into images on the ground using indirect imaging algorithms. Each Fourier component is encoded into an intensity sine wave which is sampled at 4 locations by the subcollimator’s pixel columns. This information can be further encoded into a complex number which gives the amplitude and phase of the sine wave (Massa et al. 2022). This format, known as visibility, is twice as data efficient as the summed pixel column values. Because STIX has been given a larger telemetry budget than originally anticipated, this data product is rarely requested because the visibility calculation can be done slightly more accurately on the ground. However, should telemetry be significantly reduced in the future, this product can be requested with a limited loss of scientific value.
- **Spectrogram data** are summed over selected pixels, detectors, and time-energy bins. The pixel and detector summing does not lead to a significant loss of spectral resolution but it does eliminate the spatial information. In almost all cases, however, spectrogram data are equally useful for spectral analysis as compressed pixel data but require two orders of magnitude less telemetry. This means that spectral analysis can be performed at high resolution and time cadence for longer time periods than would otherwise be possible. Hence, along with compressed pixel data, spectrogram data are the most requested science data product. It is worth mentioning that

the spectrogram data are available at the highest resolution possible at all times that STIX is in the nominal observation mode since January 2022.

- **High-time-resolution aspect data:** STIX stores readouts from the four photodiodes in the aspect system in the onboard archive memory at a resolution of 16 ms. So-called "burst-mode" aspect data can be later requested at full resolution, or binned to user-specified resolution. This is especially useful to accurately measure the direction where STIX is pointing, even when the spacecraft attitude is changing rapidly.

Science data are transmitted to the ground in raw binary packets. These are labeled level-0 at the STIX data center.

3. Data reception

During nominal science operations, low-latency data are transmitted to Earth during every ground station pass regardless of orbital geometry, whereas science data are only downlinked when bandwidth permits. Every instrument team is responsible for keeping the transmitted data volume within pre-allocated boundaries. ESA's Ground Operation System (EGOS) processes telemetry data received by ground stations at the Solar Orbiter mission control center. Then the telemetry data are distributed to the instrument teams regularly via the EGOS Data Dissemination System (EDDS) (Peccia 2005). Low-latency telemetry data generally arrive at the STIX data center within a few days, depending on the next ground station pass. However, science data may be delayed several weeks after being generated onboard due to possibly low bandwidth or priority. In addition to telemetry data, the STIX data center receives auxiliary data (Acton 1996; Acton et al. 2018) from the Science Operations Center (SOC), which contain information on spacecraft ephemeris, attitude, and calibration factors required for the conversion of onboard timestamps to UTC times.

4. Data processing pipelines

4.1. Data processing pipeline overview

STIX telemetry data are processed immediately upon reception at the STIX data center by the data processing pipeline shown in Fig. 1. Raw packets are processed to level-1 which includes parsing, integer decompressing, and timestamp conversion. L1 packets are stored in a tree-like structure and written to a NoSQL database. Then they are selected and processed in four different paths. In the first, quicklook packets are selected for the identification of solar flares. The second path performs common analyses of science data for flares. The third path selects calibration data for the determination of energy calibration factors. In the fourth, the housekeeping, quicklook, and science data are successively selected and used to create level-1 FITS files.

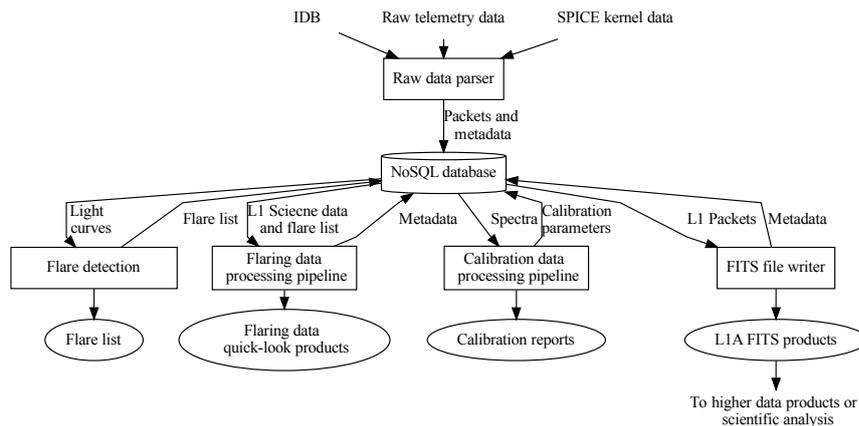


Fig. 1: Telemetry processing pipelines at STIX data center.

4.2. Raw Packet Parsing

Raw telemetry data at the STIX data center are stored as binary packets. Each packet contains a fixed-length header and a list of parameters that vary with the type of packet. The parsing of parameters is based on information in the Mission Information Base (MIB), which contains the name and length of each parameter for each type of packet. The parsed packets contain raw values of parameters, which need to be converted to physical values. Raw values of spacecraft-local times are converted to UTC times with the latest version of SPICE kernels (Acton 1996; Acton et al. 2018; and 2019). Raw values of housekeeping parameters are converted to physical values using the ground-calibrated conversion factors stored in the MIB. Compressed counts in science data are decompressed using a look-up table. After the above processing steps, packets are organized in a tree-like structure. They are considered level-1 packets and written to a collection in the NoSQL database. The NoSQL database is schemaless, which means that the data format of each record can be different and a pre-definition of the data format is not required. This makes it very suitable for storing the tree-like structure level-1 data packets. The NoSQL database provides great convenience for tasks such as searching, sorting, organizing, merging data packets, and verifying data integrity.

In addition to level-1 packets, other metadata, such as filenames, SPICE kernel version, and MIB version are written to another collection in the NoSQL database. This allows fast querying of the associated raw telemetry data.

4.2.1. FITS Products

The Flexible Image Transport System (FITS) is a portable file standard widely used in astronomy to store, transmit and manipulate scientific images, tables, and associated data (Pence, W. D. et al. 2010). Therefore, the FITS format is adopted by the STIX data center to store the standard data products. After parsing each new raw telemetry file, housekeeping, quick look, and science packets are sequentially selected from the NoSQL database and merged after passing checks for data in-

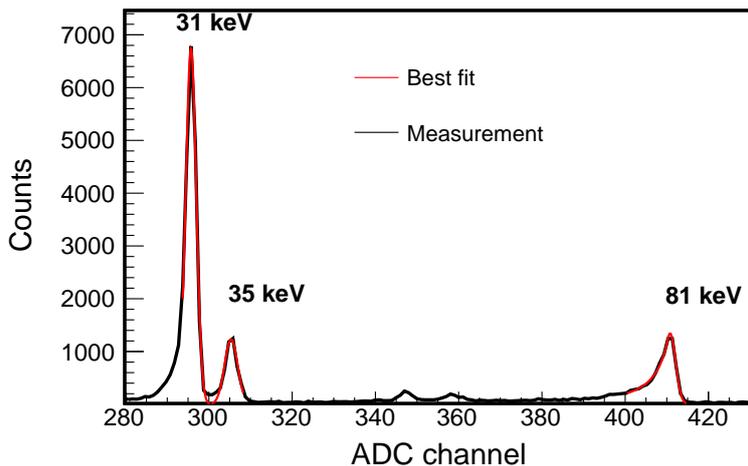


Fig. 2: An example of STIX in-flight calibration spectrum recorded by a single pixel. The most prominent peaks, from left to right, are the 31, 35, and 81 keV photo-peaks. The first two are fitted with a double-Gaussian function, and the high energy peak with a crystal-ball function.

tegrity and consistency. The merged data as well as the associated metadata and auxiliary data are written to FITS files. In the meantime, their metadata are written to a collection in NoSQL, which allows fast querying of the products. The FITS files, created from level-1 packets immediately upon the parsing process, are defined as level-1A (pre-released level-1) products. They are used by some subsequent data processing pipelines.

FITS files are recreated in a similar pipeline after a few days to weeks after all inputs are validated manually. The created FITS files are regarded as the formal level-1 products at the STIX data center. In most cases, the FITS files at level-1A and level-1 are almost identical, except that the level-1A FITS files may use predicted ephemeris data. In addition, level-1A FITS files may not contain all header keywords required by the ESA standard. As such, level-1 FITS files are recommended for STIX data users whenever they are available.

4.3. Energy calibration

STIX converts ADC channels to "science energy channels" onboard by the FPGA using an energy lookup table (ELUT) which defines the ADC channel edges for each science channel for each pixel. An ELUT can be constructed using energy conversion factors determined from calibration runs. STIX continuously accumulates an energy spectrum (in ADC units) for each pixel separately for events from the onboard Ba^{133} sources and formats a spectrum typically every 24 hours. Fig. 2 shows an example of such a spectrum. The three most prominent peaks are produced by photons of 31 keV, 35 keV, and 81 keV energy from the calibration sources. To determine the positions of the photo-peaks, the first two peaks are fitted with a double-Gaussian function, and the third peak with a crystal-ball function (Skwarnicki 1986), which consists of a Gaussian core and a power-law low-end tail, below a certain threshold. Then a linear fit is made to the positions and keV energies whose slope and intercept give the gain (i.e. the ADC to energy conversion factor) and baseline,

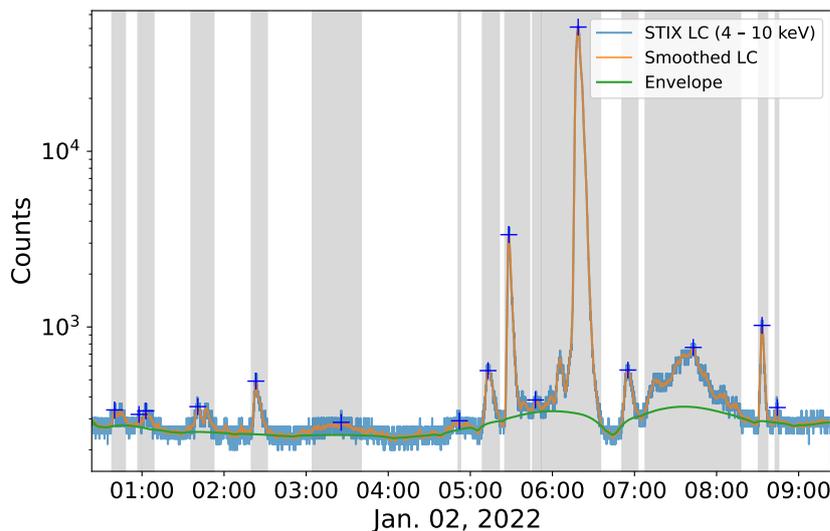


Fig. 3: STIX 4–10 keV quicklook lightcurve for 2022-01-02T00:30:00 – 2022-01-02T09:30:00 (UTC) and identified flares. The orange curve is the smoothed lightcurve using a moving average filter with a time window of 1 minute. The green curve is the estimated signal envelope using the SNIP algorithm. The identified peaks are marked with plus signs, and the gray ranges show their time ranges.

respectively. The ECC method (see Maier & Limousin 2016; Maier et al. 2020) is another method often used by the STIX team to determine the calibration factors. We found that the results of the two methods are consistent within 1σ .

The above steps are performed for each calibration spectrum once the data are available at the STIX data center. The calibration factors are written to a collection in the NoSQL database and used for further correction of energy bins in offline data analysis. Once significant changes in the calibration factors are observed, the STIX operations team creates a new ELUT and uploads it to STIX.

4.4. Solar flare identification

STIX identifies solar flares onboard based on detector count rates, and the results are included in the quicklook data (Krucker et al. 2020). However, the data only provide limited information on flares due to the constraints of the telemetry budget and onboard computing resources. Moreover, microflares are not reported due to the relatively high trigger threshold. It is necessary to maintain a flare list on the ground for requesting science data and also for users to find events of interest.

Using quicklook lightcurves, solar flares can be identified in greater detail on the ground. The ground identification procedure includes the following steps:

- Lightcurve smoothing: The selected lightcurve is filtered using an unweighted moving average filter with a time window of 1 min. This can smooth out statistical fluctuations and electric surge spikes.
- Envelope subtraction: A flare may last hours, and there may be short-duration pulses lying on the envelope (the main pulse) in the light curve. To facilitate the identification of those short-

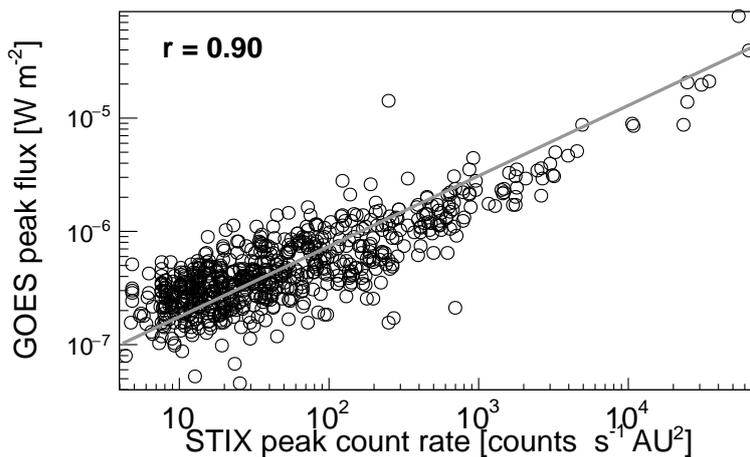


Fig. 4: Scatter plot of GOES/XRS low channel peak flux with respect to STIX 4 – 10 keV count rates scaled to 1 AU for 717 solar flares observed by both GOES and STIX during the cruise phase. The solid line is a linear fit to the log-log graph. The Pearson’s correlation coefficient in the log-log scale is 0.91. From the fit, we get the GOES flux estimation formula as follows: $f = 10^{0.622 - 7.376 \log_{10}(X')}$ (in units of W/m^2), where X' is the STIX peak count rate corrected for the distance variations between the Sun and Solar Orbiter.

duration pulses, the envelope is subtracted from the smoothed light curve, which is estimated using the SNIP algorithm (Ryan et al. 1988).

- Identification of flare peaks: We consider that a flare is detected if the peak count rate after envelope subtraction lies beyond two standard deviations of the mean count rate during quiet Sun periods. The flare start and stop times are given at the times exceeding this threshold.
- Merging of flare peaks: Two peaks are considered from one single flare if the peak times differ by less than 5 min. This can reduce the number of reported flares and simplify data analysis. Note that this can also affect any statistical flare distribution derived from this list.

As an example, Fig. 3 shows STIX quicklook light curve in the energy range 4 – 10 keV, recorded from 2022-01-02T00:30:00 to 2022-01-02T09:30:00 (UTC). The orange curve is the smoothed light curve. The identified peaks are marked with the plus signs, and the colored ranges show the time ranges.

The above steps are repeated for quicklook lightcurves of the other four higher energies, which provide information on the upper limit of the X-ray energy of the flare. The time ranges, peak count rate, and total counts as well as the corresponding ephemeris data of the identified flares are stored in a collection called flare list in the NoSQL database, which is later used for the creation of data requests (see 5).

4.5. Solar flare standard analysis pipeline

4.5.1. Estimation of solar flare GOES class

Solar Orbiter is often far from the Sun-Earth line and hence a considerable number of flares observed by STIX are not observed by GOES satellites (and vice versa). In order to estimate the

GOES classes of such flares, we selected 717 solar flares observed by both GOES/XRS and STIX. Fig. 4 shows the scatter plot of the peak fluxes measured by GOES satellites with respect to the STIX background-subtracted count rates at the peaks, in the energy range of 4 to 10 keV. STIX count rates x have been corrected for the different distances between the Solar Orbiter and the Sun using $X' = xr^2$, where x is the count rate after background subtraction and r is the distance between the Sun and Solar Orbiter in units of AU. A clear correlation (The correlation coefficient $r = 0.90$) is seen in Fig. 4. The wide spread at low fluxes can be explained by the difference in the energy response of the two instruments and the variation in flare temperatures. The correlation can be fitted with a linear fit in the log-log scale. From the fit, we get the GOES flux estimation formula as follows: $f = 10^{0.622-7.376\log_{10}(X')}$ (in units of W/m^2), where X' is the STIX peak count rate corrected for the distance variations between the Sun and Solar Orbiter. It is currently used to estimate the GOES classes of flares that are not directly observed by GOES satellites. The estimated GOES fluxes are stored in the flare list collection in the database.

4.5.2. Estimation of coarse flare locations using CFL data

STIX estimates approximate flare locations onboard by maximizing the correlation between observed counts recorded by the 12 pixels in its coarse flare locator (CFL) subcollimator with expected counts using a lookup table (Krucker et al. 2020). With the requested science data, coarse flare locations can be reconstructed more accurately, as it allows for more sophisticated algorithms, and greater flexibility in selecting time and energy range to be integrated.

When the pixel data of a flare are available at the STIX data center, its flare location is estimated. The steps are as follows:

1. Integrate counts around the peak for each CFL pixel.
2. Subtract the background using a background file.
3. Estimate the illuminated area on each CFL pixel. This is based on two assumptions: the illuminated area of a pixel is proportional to its relative count rate, and the total illuminated area of the imaging detectors is independent of the source location.
4. Estimate the flare location by minimizing the weighted sum of squared deviations (i.e., weighted chi-squares) between the calculated illuminated areas and expectations simulated for potential flare locations in a 400×400 grid.

As an example, the left panel of Fig. 5 shows the calculated and best-fit illuminated areas of the CFL pixels for the flare observed at 2021-05-07T19:00:00 (GOES class M3.9); the middle panel shows the best-fit flare centroid location, as well as its 1σ , 2σ , and 3σ contours. The simulated CFL shadow pattern is shown in the right panel. The estimated flare locations are stored in the flare list in the database.

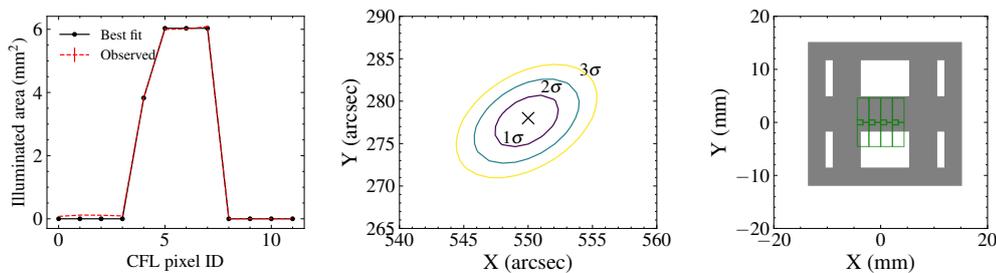


Fig. 5: Left: Calculated areas of illuminated regions of the 12 CFL pixels and the best-fit simulated pattern for the flare location at (550, 278) arcsec. Pixels 0 to 3 are the top big pixels, Pixels 4 to 7 are the bottom pixels, and 8 to 11 are the small pixels as shown in the right panel. Middle: Best-fit flare centroid location (marked by x) and its 1σ , 2σ , and 3σ confidence contours. Right: Projection of CFL sub-collimator (the gray shaded regions) on the detector simulated for the best-fit flare location.

4.5.3. Imaging and spectroscopy pipeline

To help in finding flares of interest, we developed an imaging and spectroscopy pipeline that automatically reconstructs images and performs spectral analysis for each flare with detector summed counts greater than 10,000 after receiving its pixel data. The pipeline first selects and integrates counts for 60 seconds around the peak of the flare for each pixel, and then subtracts the background from the integrated counts using the pixel data acquired during a quiet solar period. It is worth mentioning that quiet solar periods are determined on the basis of counts in the quick-look lightcurves. Subsequently, the transmission and dead-time corrections are performed on the background-subtracted counts. Then the corrected counts are further converted into the visibilities of two energy bands of 4 – 10 keV (thermal energy) and 16 – 28 keV (nonthermal energy). Then the visibilities are used for image reconstruction using four different algorithms: Back-Projection (BP) (Massa et al. 2022), CLEAN (Högbom 1974), MEM_GE (Cornwell & Evans 1985; Massa et al. 2020) and VIS_FWDFIT (Volpara et al. 2022). Finally, the reconstructed images are corrected for spacecraft off-pointing and rotations. As an example, the first panel of Fig. 6 shows the lightcurves and time range selected for a flare that occurred at 2022-10-07T13:50:39 (UTC). The rest of the panels show the images reconstructed with the algorithms. In addition to image reconstruction, the transmission and the dead-time corrected counts are used for spectral analysis. Fig. 7 shows the results of the spectral fitting for the same flare as in Fig. 6. The spectrum is fitted with a thermal component and a non-thermal component.

The results of the pipeline are saved to files in both FITS and PNG formats. At the same time, the indexing information from the file, the parameter values of the spectral analysis, and the auxiliary data are written to a collection in the database.

5. Science data request strategy

As mentioned previously, high-resolution pixel data are only downlinked in response to requests from the ground. They are stored in the onboard archive memory for a few months before being overwritten by new data. They are processed and downlinked after receiving data request telecom-

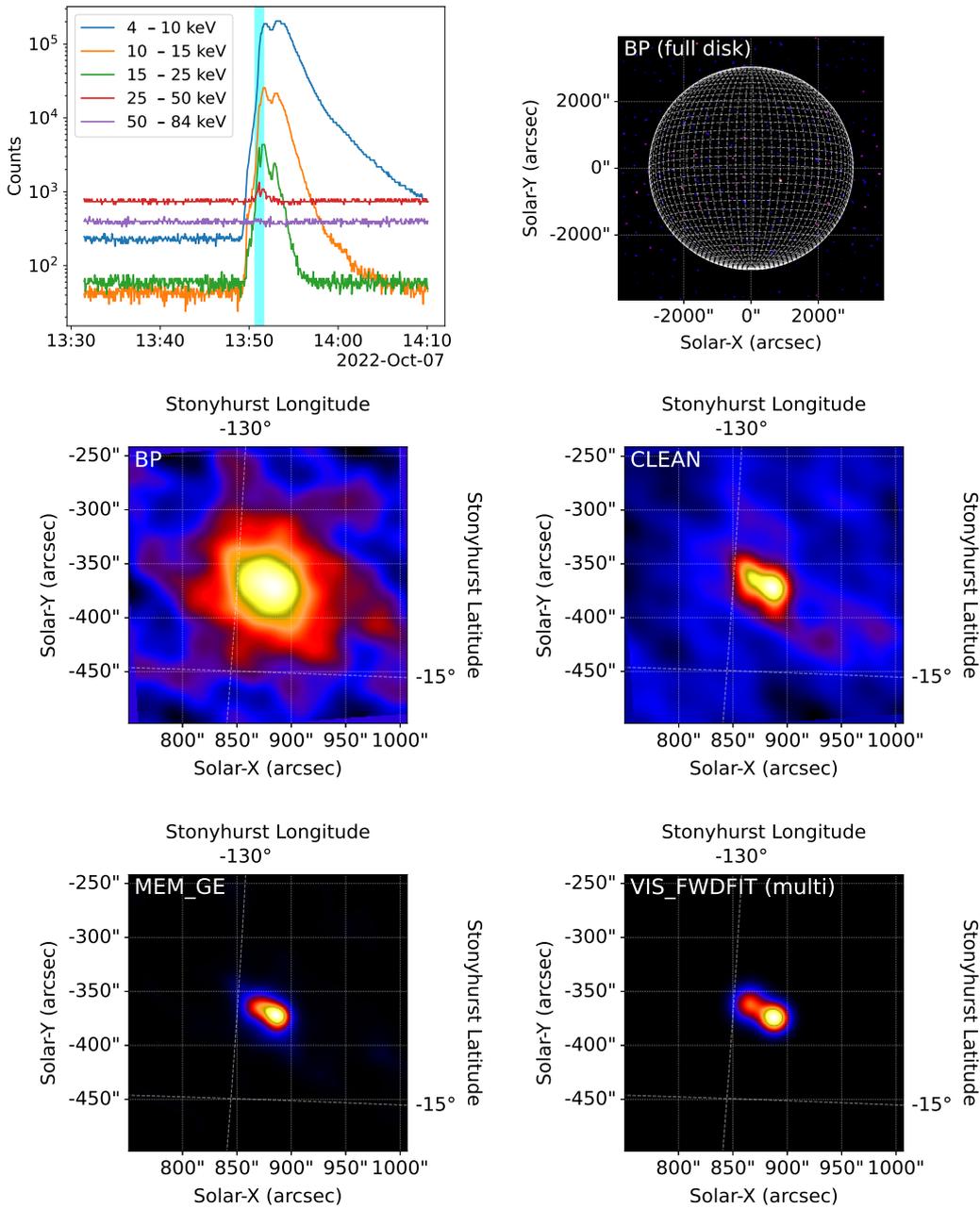


Fig. 6: STIX Quicklook lightcurves and reconstructed images of the solar flare observed at 2022-10-07T13:50:39 (UTC), created by the image reconstruction pipeline. A 1-min integration time around the peak was selected and the images were reconstructed using four different algorithms: Back-Projection (BP), CLEAN, MEM_GE and VIS_FWDFIT. The full-disk Back-Projection image shown in the top-row right panel has been used to identify the source location. The other reconstructions were performed around the location of the flare. Note that the multiple circular Gaussian shapes have been selected for reconstructing the flaring source by means of VIS_FWDFIT (bottom-right panel).

mands from the ground. A data request telecommand contains information about the selected data and values of parameters required to process the data onboard, including the data compression level, time range, minimal time bin, energy bin width, and masks indicating detectors and pixels to be selected. STIX detects thousands of flares per year; therefore, selecting data is a tedious task, as many factors must be considered, such as count rates, time binning of data, statistics of selected

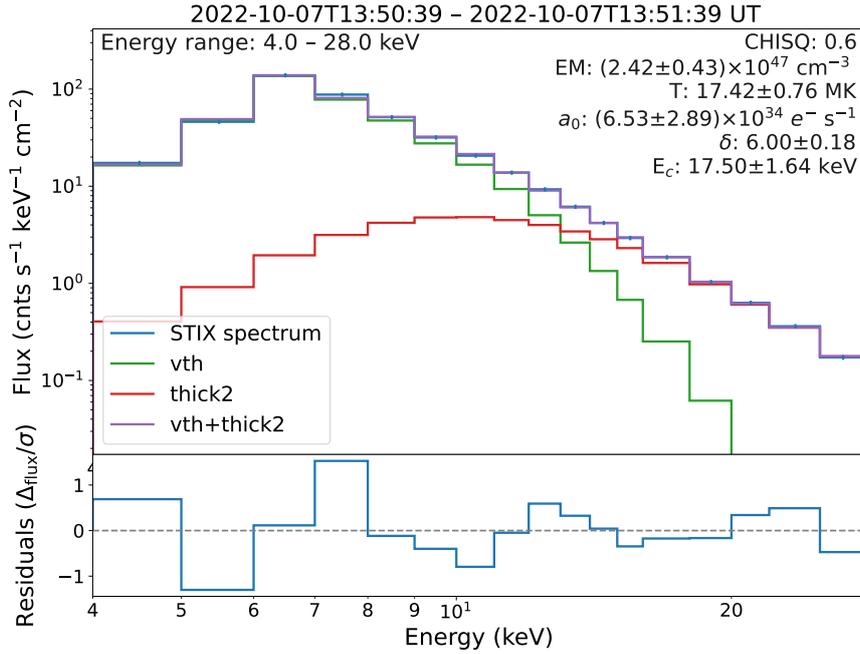


Fig. 7: Spectral fitting results for the same flare as in Fig. 6. The spectrum is fitted with a thermal component and a non-thermal component.

data, and also the telemetry budget. The data selection strategy has been continuously optimized over the past two years. The current strategy is as follows:

- Compressed level-1 pixel data are requested for each of the ground-identified flares with a total number of signal counts in the quicklook lightcurves greater than 10,000, which is approximately the minimum counts to reconstruct an X-ray image. The requested energy range is chosen to be the range in which obvious signal counts are seen, whereas, the requested time resolution is adjusted based on the amount of data allocated and the scientific relevance of the flare. If the peak count rate is above 125 counts/sec (approximately equivalent to the count rate observed for a B3 flare at 1 au), pixel counts with high time resolution are requested. Otherwise, pixel counts are integrated over the whole flaring time to reduce the telemetry data volume. These parameters may be customized for individual flares by human operators during review.
- Spectrograms with the highest time and energy resolution are requested for all periods when STIX is in the observation mode.
- Time-integrated pixel data for background subtraction: Time-integrated pixel data with durations of one or two detector temperature cycles (each cycle lasts about 40 minutes), which are acquired during quiet-Sun periods, are requested.

The data are used for background subtraction when performing spectroscopy and imaging.

- High time-resolution aspect data are requested, for example, for periods when the spacecraft’s attitude changes drastically. Such periods can be known from the SPICE kernels or the aspect system readouts in housekeeping data.

The selection of science data of the above types is done automatically using a program (except for aspect data). The information of the selected data is written into a collection in the NoSQL database. In addition, the STIX operations team also selects data for special needs. After being checked and adjusted by the STIX operations team, groups of new data requests that meet the operations requirements are selected from the database and then compiled into instrument operation requests (IORs), which are used to create final telecommands at the mission operations center (MOC). These telecommands are then uploaded to the spacecraft and executed by STIX typically two to three weeks later.

6. STIX data center user interfaces

6.1. Interactive web pages

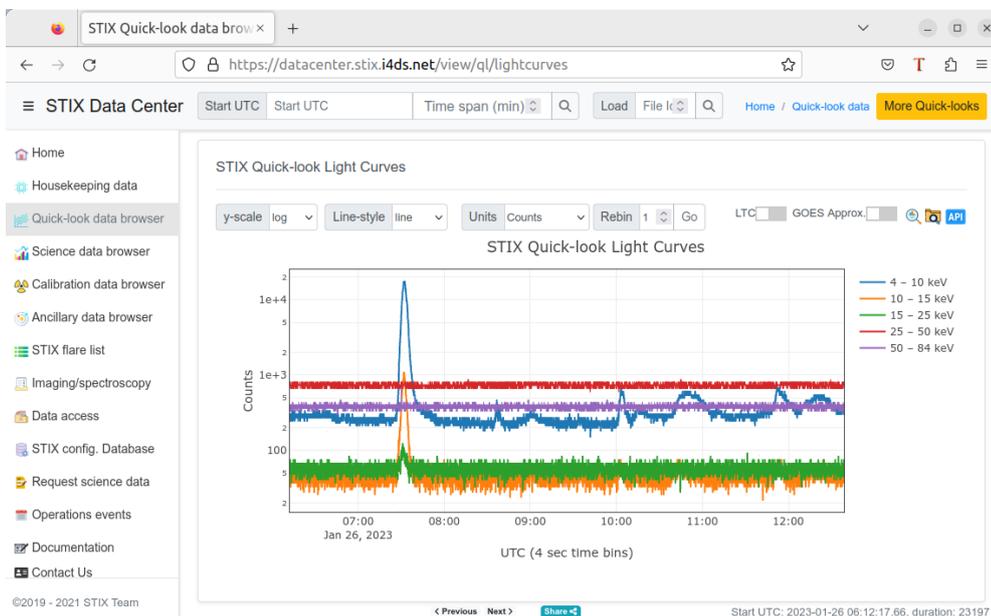


Fig. 8: Interactive web-based STIX Quicklook data browser. In addition to STIX quicklooks, it can also display quicklooks of simultaneous measurements performed by other instruments, such as GOES/XRS, SDO/AIA, and EUI onboard Solar Orbiter. The browser is available at the link <https://datacenter.stix.i4ds.net/view/ql/lightcurves>.

The data center platform provides various HTTP interfaces (APIs) that allow access to STIX data products and the NoSQL database via HTTP requests. We have built dozens of web applications to manage and browse STIX data based on these APIs. Web techniques are chosen because they offer many advantages, such as clear cross-platform usability, broad access through browsers, rapid development, and easy maintenance.

As an example, Fig. 8 shows a screenshot of the STIX Quicklook data browsing tool. It allows users to browse available Quicklook data interactively. It interacts with the server through APIs. The data exchange mechanism between the server side and the client side is shown in Fig. 9. After receiving a request from the client side, the server retrieves quicklook counts from the NoSQL database for the user-specified time range. After excluding duplicates and merging, the server sends

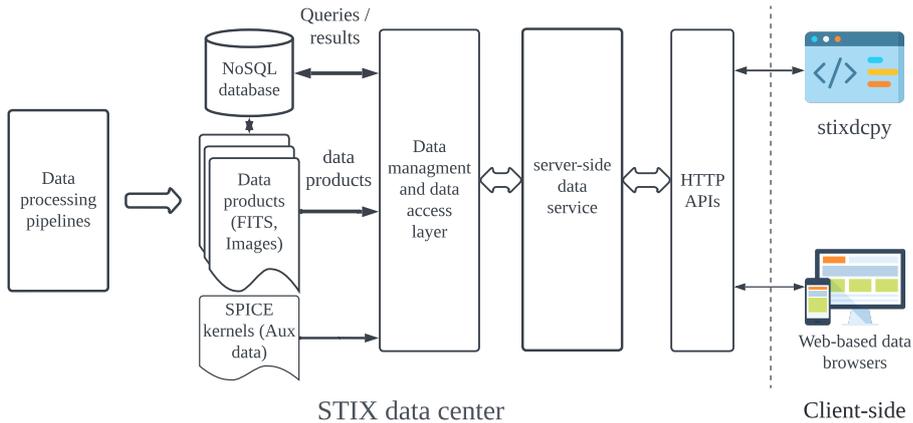


Fig. 9: Client side and server side data exchange mechanism.

quicklook counts and metadata in JSON format back to the client. The data are then used to create interactive light curve plots using JavaScript on the client side. The interactive plot uses state-of-the-art web technologies that enable users to perform a range of operations, such as rebinning the integration time, correcting the light travel time between the spacecraft and the Earth, and exporting plot data to a local file. In addition, quicklooks from other solar-observing instruments can also be displayed on the same page after users' activation, making it easier to find events of interest for joint analysis.

Based on similar concepts, tens of web tools have been developed to browse other STIX data products. The four most commonly used tools are listed below:

- **The science data manager and browser** provides users with tools for searching, downloading, visualizing, and analyzing science data. The interactive analysis tools allow users to select data of interest for standard analysis tasks, such as background subtraction and energy rebinning, without installing additional software. The algorithms are implemented client side using JavaScript. In addition, users can submit imaging and spectroscopy tasks to the server and view the results on the same page. This reduces the barriers to exploring STIX data for new users and provides convenience for experienced users.
- **The preview images and spectroscopy product viewer** is a web-based tool for managing and viewing the imaging and spectroscopy results. The viewer also provides tools for plotting the time evolution of emission measures and temperatures, creating animations of X-ray images for the selected runs, generating IDL or Python templates that allow reproducing the same results on local machines, and so on.
- **The auxiliary data viewer** allows the user to view auxiliary data, such as the spacecraft locations, velocity, and attitude, using data derived from the SPICE kernel and pointing solutions computed from aspect system data. The viewer also provides tools to calculate the looking angles of flares and the coordinates of solar limbs within the STIX field of view.

- **The housekeeping data browser** enables users to view time series of all STIX housekeeping parameters, including the temperature, voltage, operation mode, memory status, etc. It provides great convenience for the instrument operations team to monitor the instrument status.
- **The STIX data access page** offers users a variety of tools to search and download STIX data products. It also provides links to web tools to preview the products. As soon as they are generated at the STIX data center, STIX data products are immediately available for access on the page.

6.2. STIX data center interface via *stixdcp*

*stixdcp*¹ is a python package that facilitates accessing and analyzing STIX data. With *stixdcp*, users can easily query and download the data products available at the STIX data center. Similarly to the web tools, *stixdcp* also provides tools to perform some standard analysis of STIX data, such as dead time correction, transmission correction, data clipping, and merging. *stixdcp* is still under active development. As a result, its features and capabilities may change over time.

7. Summary

STIX is one of ten instruments onboard the Solar Orbiter, which was launched into space on February 10, 2020. STIX measures intensity and spectrum of hard X-rays emitted during solar flares in the energy range of 4 – 150 keV. During nominal operations, STIX continuously generates telemetry data. To process and archive the data as well as to support the operations of the instrument and scientific activities using STIX data, dedicated data processing pipelines and a data platform have been developed at the STIX data center. The pipelines generate telemetry at different levels and perform standard scientific analyses. The data center platform distributes STIX data products and also provides users with various web-based tools for searching and browsing STIX data products. The data center is designed to work in a fully automatic mode with minimal human intervention. The concept has proven successful and has been running continuously for more than two years. The data center not only facilitates the operations of the instrument but also provides great support to STIX data users.

References

- Acton, C., Bachman, N., Semenov, B., & Wright, E. 2018, *Planetary and Space Science*, 150, 9, enabling Open and Interoperable Access to Planetary Science and Heliophysics Databases and Tools
- Acton, C. H. 1996, *Planetary and Space Science*, 44, 65, planetary data system
- and. 2019, Solar Orbiter SPICE Kernel Dataset
- Cornwell, T. J. & Evans, K. F. 1985, *A&A*, 143, 77
- Forveille, T. & Shore, S. 2020, *A&A*, 642, E1
- Högbom, J. A. 1974, *A&AS*, 15, 417
- Krucker, Hurford, G. J., Grimm, O., et al. 2020, *A&A*, 642, A15

¹ *stixdcp* source code is hosted on the GitHub repository at <https://github.com/i4Ds/stixdcp>.

- Maier, D. & Limousin, O. 2016, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 812, 43
- Maier, D., Limousin, O., & Daniel, G. 2020, EPJ Web Conf., 225, 01003
- Massa, P., Battaglia, A. F., Volpara, A., et al. 2022, Sol. Phys., 297, 93
- Massa, P., Schwartz, R., Tolbert, A. K., et al. 2020, ApJ, 894, 46
- Müller, D., St. Cyr, O. C., Zouganelis, I., et al. 2020, A&A, 642, A1
- Peccia, N. 2005, in 2005 IEEE Aerospace Conference, 3988–3995
- Pence, W. D., Chiappetti, L., Page, C. G., Shaw, R. A., & Stobie, E. 2010, A&A, 524, A42
- Ryan, C., Clayton, E., Griffin, W., Sie, S., & Cousens, D. 1988, Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 34, 396
- Skwarnicki, T. 1986, PhD thesis, Cracow, INP
- Volpara, A., Massa, P., Perracchione, E., et al. 2022, A&A, 668, A145
- Warmuth, A., Önel, H., Mann, G., et al. 2020, Sol. Phys., 295, 90