ISDC for the INTEGRAL USERS GROUP

Carlo Ferrigno (ISDC)

ESTEC, 26-27 November 2019

Organization status

- Swiss funding for 2020: approved with minimal direct funding (1 FTE). Operations guaranteed.
- Manpower allocation: ½ operator, ½ scientist (to be found). Contribution from infrastructure (CDCI project, led by S. Paltani) for web mastering, DB support and web-analysis. Contribution from ESA for operations.
- Synergies with other projects and past savings is essential. ½ scientist in 2019 from ESA; ½ scientist from FNRS.
- Operator retires in January 2020, substitute will start in December 2019 for training.

Routine tasks

- Updates of IC files provided by instrument teams
- monitoring of SPI gain at each revolution with automated procedure
- SPI gain coefficients updated last time in 2018
- Processing and archiving of CONS data

IBAS energy calibration with OSA11

- Outdated energy calibration in IBAS probably causes under detection
- Diego Gotz implemented an update to IBAS code with contribution by VS and ISDC
- Introduced in operations on 21 November 2019.
- Tuning threshold for too many weak alerts

ISDC Operations/data distribution

- NRT data are available within 3 hours. Smooth processing (monitoring issues more closely now for MMA).
- Page to distribute data since AO13, public for serendipitous science. Handled Russian peculiarity.
- Need for OSA energy reconstruction step both for JEM-X and ISGRI. NRT data for JEM-X2 are not always available due to difficult energy reconstruction.
- Occasional gaps in NRT telemetry due to hardware failure of the University infrastructure supporting the data transfer or interruption (switch).
- CONS data are obtained now from virtualized DVD transfer from MOC
- Service widely used for SPI-ACS data in NRT

INTEGRAL SPI-ACS public data service

In 2011, a public service was set up to promptly provide SPI-ACS data with the best timing accuracy

It was extensively used for years by IPN and Konus colleagues

Since 2015, Fermi/GBM team used the service to verify their detections and challenge SPI-ACS

Several other groups started to use it. In total >100 Gb has been served.

| IPN format SPI-ACS light curve | 2008-03-19T06:12:46 200 | Submit |
IPN format INTEGRAL ephemeris	2008-03-19T06:12:46	Submit
Plot SPI-ACS light curve	2008-03-19T06:12:46 200	Submit
INTEGRAL Attitude	2008-03-19T06:12:46	Submit
INTEGRAL HK light curves	SPI_VETOGATE 2008-03-1	Submit

Try using the script to access the lightcurves

RESTful service, providing various public INTEGRAL data as well as auxiliary information

# Quick look analysis of INTEGRAL data

- 3 GRB in the IBIS FOV in 2019, 7 in 2018
- ~200 GRB/year in SPI ACS. Used for IPN triangulation.
- Inform all PIs of data rights only in case of outstanding problems or relevant serendipitous sources.
- 18 ATeLs (5 lead by ISDC) and 71 GCNs related to INTEGRAL discoveries in 2019 (included GW and neutrino follow-up)
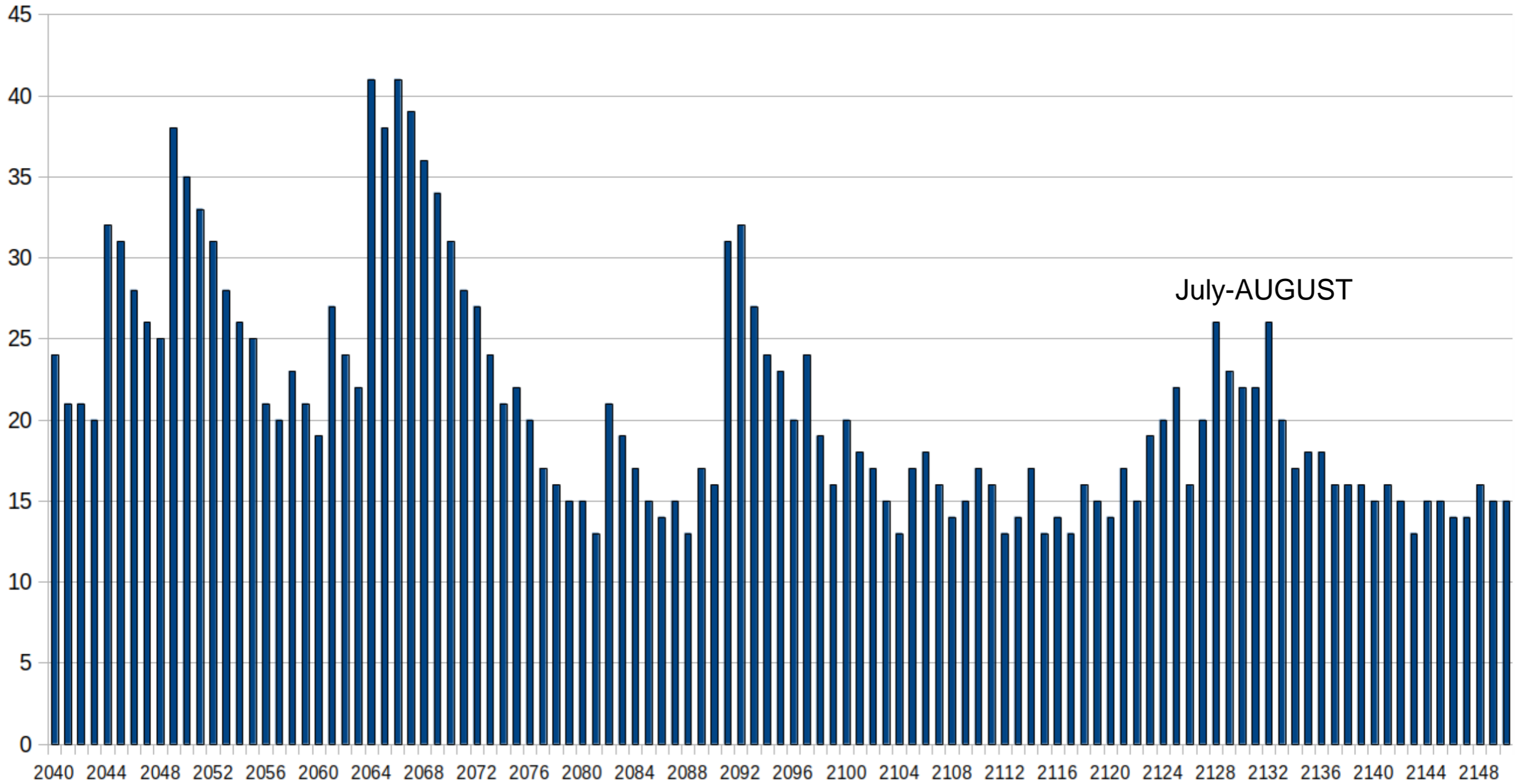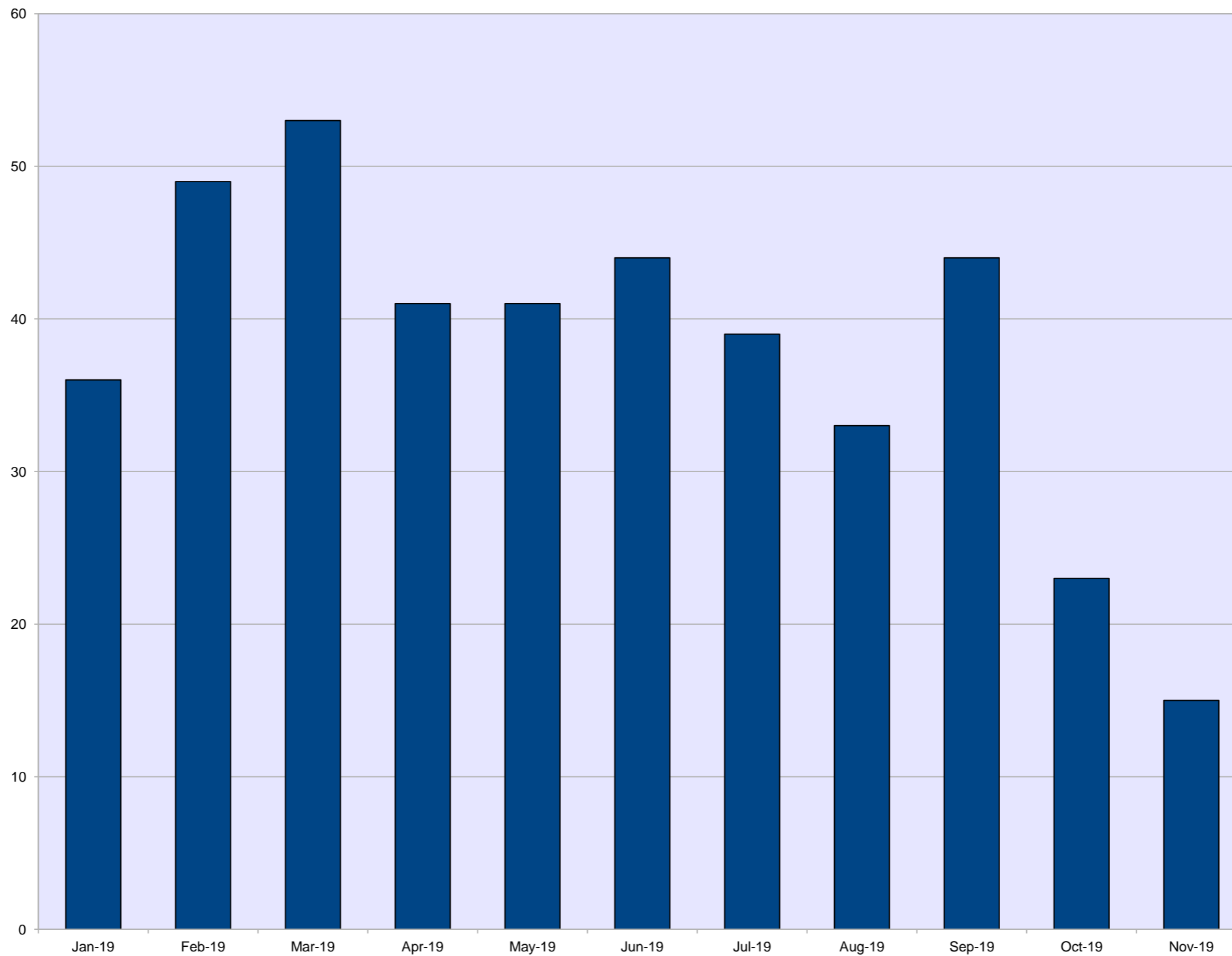
Telemetry

CONS / NRT, rev.   2040 - 2150
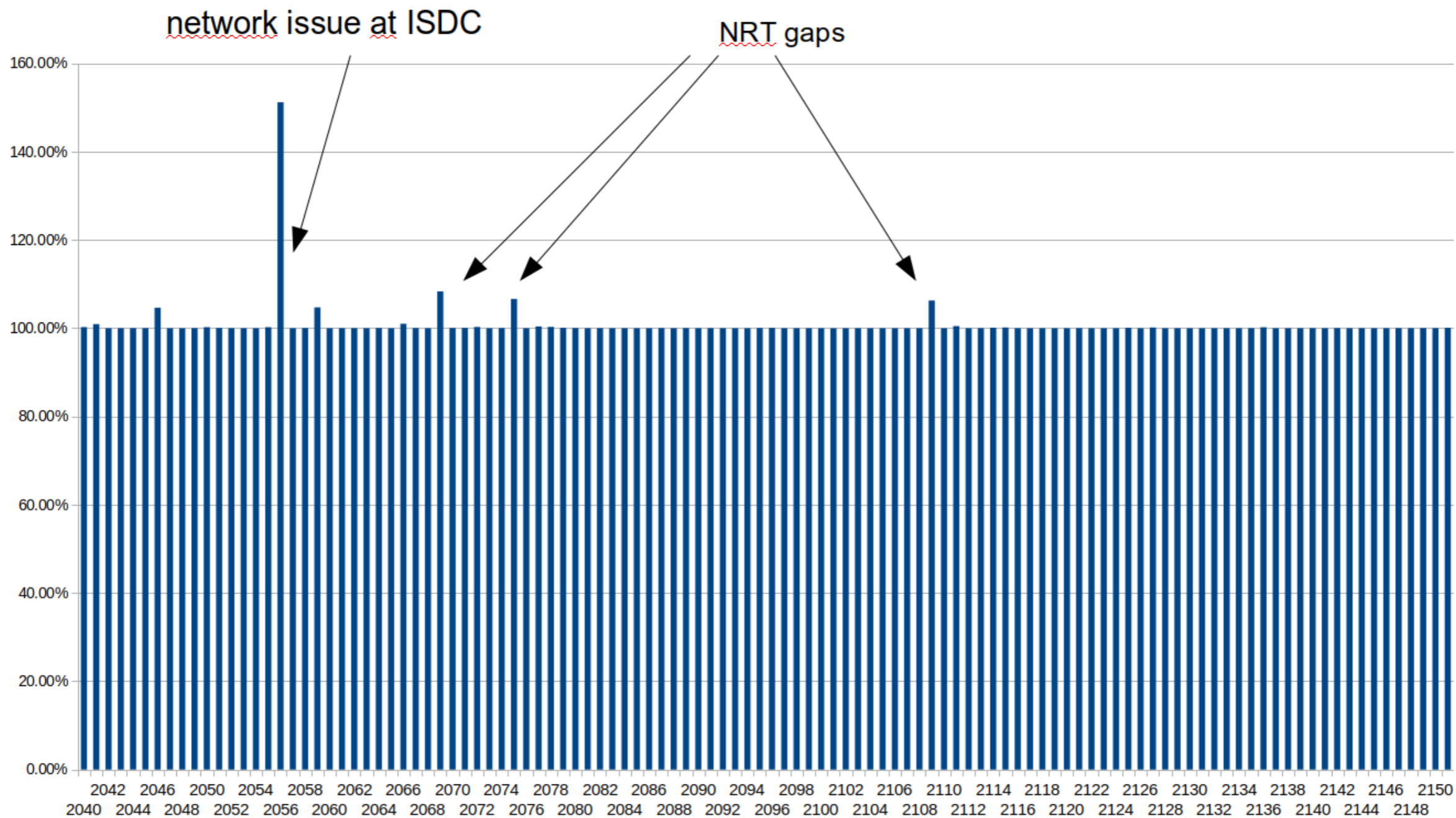
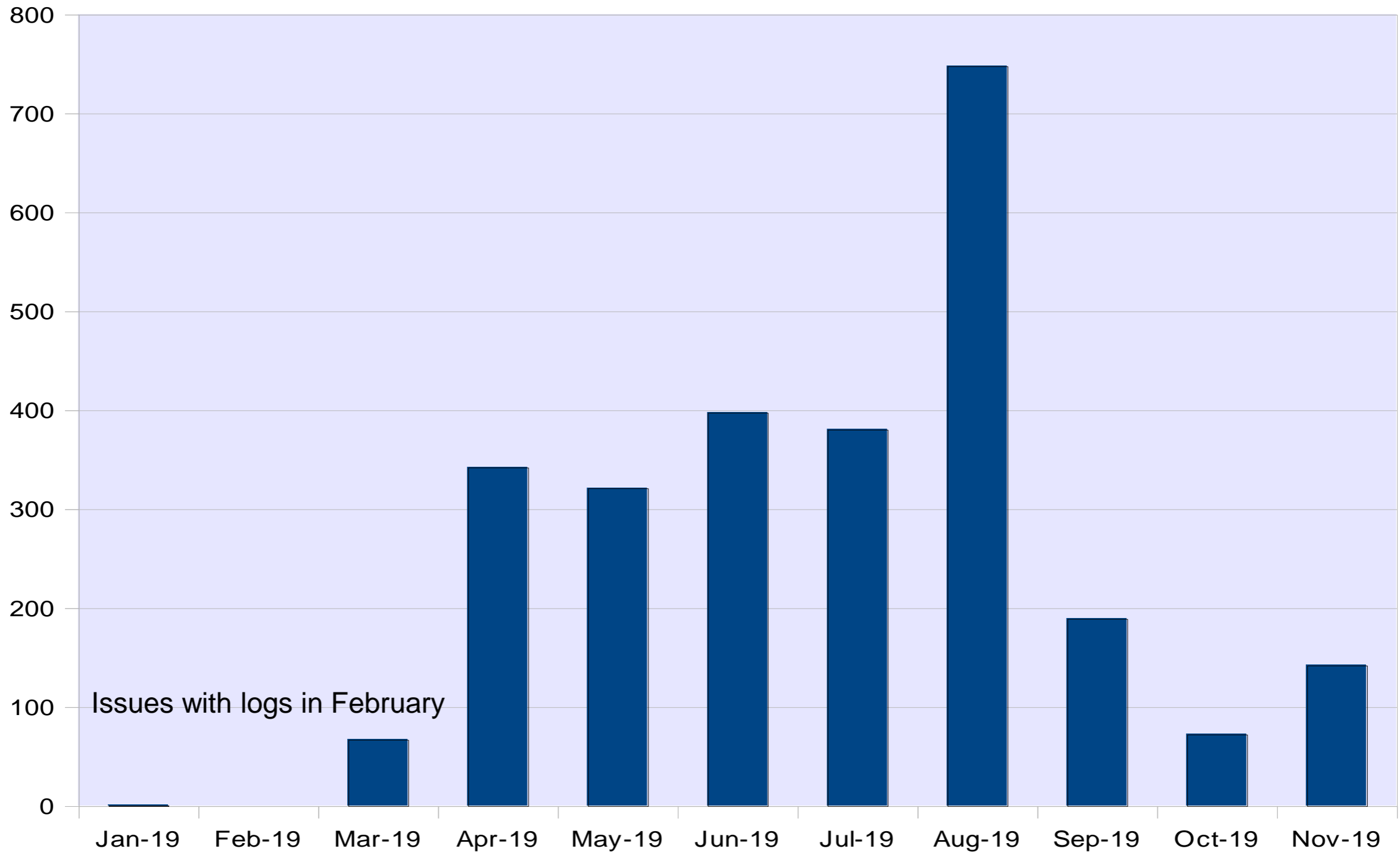Operations: Good times vs scheduled 2041 – 2150
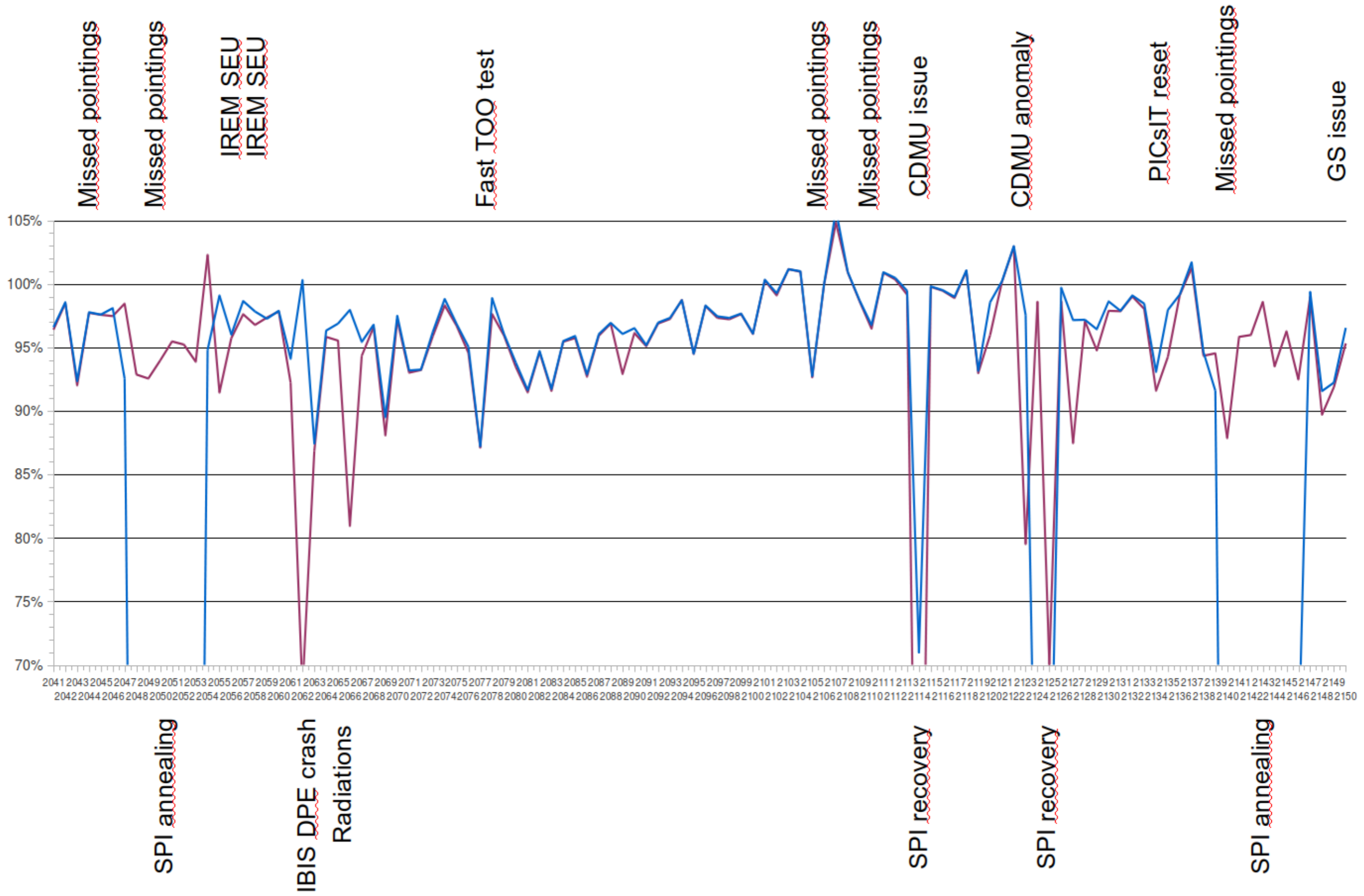
Delay between observation and distribution
rev. 2040 - 2150

July-AUGUST

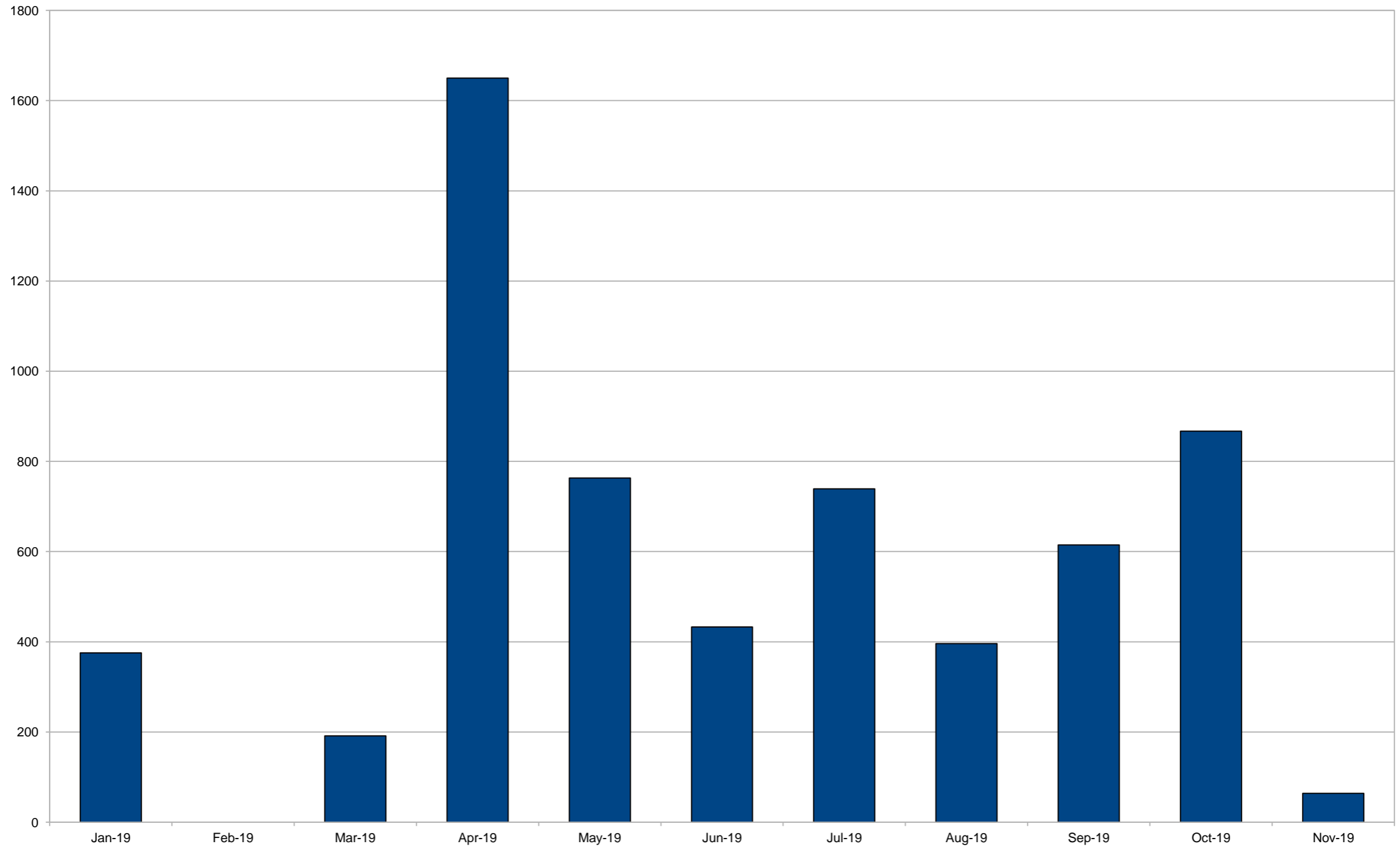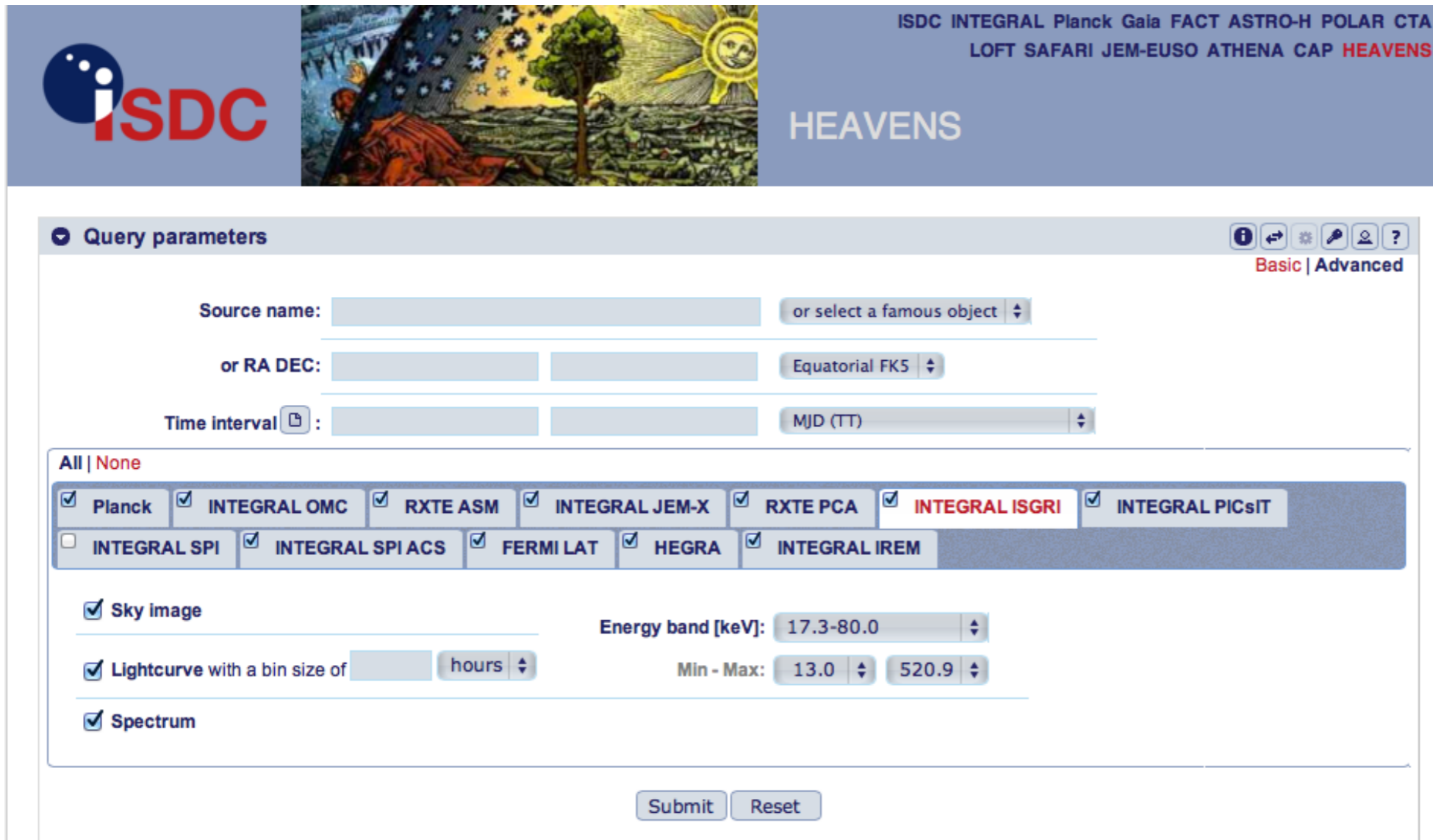# Browse unique visitors

FTP access statistics (bandwidth GB)

Rsync

Bandwidth (GB)

Issues with logs in February

High level quick-look products, HEAVENS

- Development on hold
- Used OSA9
- Updated IREM on request

~60 single accesses per month

INTEGRAL conference 2019

https://www.astro.unige.ch/integral2019/

- Proceedings in the final phase of production in Italy

ISDC catalog

- Released catalog v. 41 in June 2018 (major revision)
- Released catalog v. 42 in December 2018 (+11 sources)
- Released catalog v. 43 in June 2019 (+5 sources)

OSA11 downloads

(since release and until 25 November)

- 255 downloads (76 downloads of OSA 10.2 after OSA11 release)
- 3 source code (often the same user as for binaries)
- 79 test data
- 221 catalog bundles
- >2300 pulls of the docker image (not possible to trace provenance on dockerhub)

OSA future activities

- Release of OSA 11.1 with j_ima_iros solving several issues (January –February 2020)
- Compton imaging as delivered now cannot be distributed, because
- it assumes the local filesystem of the developer
- It does not use dal libraries for IC files (so not possible to access and update them)
- It does use the parameter files
- It does not use the logging system

- ISGRI energy calibration updates do not require any update to OSA

ISGRI energy calibration

- A plan has been agreed between ISDC and ESA to perform ISGRI calibration in 2020.
- Work is supervised by VS, but actually performed by an ESA trainee: Lionel Metrailler (TBC)
1. Identify sources for cross-calibration with SPI and possibly NuSTAR
2. Preparing basic calibration files using existing pipeline
3. Verification and adjustment

Time-domain astronomy

- MoU with Antares
- MoU with IceCube for non-public alerts
- LVC issues public notices

- Implemented the real-time dump of SPI-ACS stream to be possibly used in fast triangulation with GBM and IPN satellites.
- System of "burst advocates"

Transient event dashboard

- Every input (GCN notice) is automatically processed and it generates results.

- Link to the workflow is private

- Need to improve on dissemination, as results are in Zenodo sandbox

The private dashboard to work on data
https://analyse.reproducible.online/transients/dashboard/

INTEGRAL status

Snapshot at 2019-06-09T18:16:25 UTC

Orbit 2100, 151.93 Mm to Earth

	State/last	Latency	RA	Dec
Real-time	ONLINE	66.0 s	4.7	59.6
NRT	210000250010	1.2 h	4.3	61.7
CONS	209000990010	25.3 d	320.0	-45.0

INTEGRAL status   Schedule

next break in data in 29 hr: 2019-06-10T23:32:59, for 9.6 hr

Gravitational Wave Detector Network

Operational Snapshot as of Jun 09, 18:16 UTC

Detector	Status	Duration
GEO 600	Observing	2:40
LIGO Hanford	Observing	15:40
LIGO Livingston	Observing	9:08
Virgo	Troubleshooting	2:53
KAGRA	Future addition	

Detector status summary pages   LVC links

GOES Xray Flux (1−minute data)   Begin: 2019 Jun 7 0000 UTC

Updated 2019 Jun 9 18:14:12 UTC   NOAA/SWPC Boulder, CO USA

⊞ Events ▼ Observations ▼ LIGO/Virgo ▼ AMON/IceCube ▼ INTEGRAL ▼ SPI-ACS ▼ Fermi ↻ All

Event	Origin	Role	UTC	Sky Location	Orientation (θ, φ)	FoV exposure	ScW	Data	Visibility	Planning urgency	Raw Notice
▼ S190602aq	LIGO Virgo	observation	2019-06-02T17:59:27.0	73.39 -7.03	bottom (127.8, -20.2)	0.0 ks	209700520010	NRT	1.3%		VOEvent JSON ⓘ
▼ S190602aq	LIGO Virgo	observation	2019-06-02T17:59:27.0	73.39 -7.03	bottom (127.8, -20.2)	0.0 ks	209700520010	NRT	1.3%		VOEvent JSON ⓘ
▼ S190602aq	LIGO Virgo	observation	2019-06-02T17:59:27.0	73.39 -7.03	bottom (127.8, -20.2)	0.0 ks	209700520010	NRT	1.3%		VOEvent JSON

# "Standard" INTEGRAL transient analysis: end to end

Research, development environment lets **experts develop standardized, test, and integrate:**
- data reduction (close to data)
- GRB spectral models (linked to literature)
- statistical methods (as portable as possible)
- visibility planning tools (remote ESAC service)

- **Find combinations** of data, adapters, statistical methods, publishers, planners
- **suggest follow-up**
- **distribute** standard results with public data, uploads to zenodo sandbox.

experts
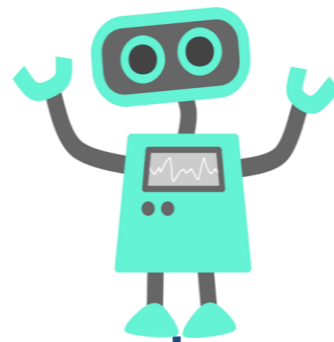
Shift (24/7)

```
In [1]:   # Parameters
          gcn_number = 88888
          name = "MC1AUTO"
          t0_utc = "2019-05-17T20:33:36"
          ra = 94.76600985221675
          dec = -52.22292534493076
          radius = 5
          event_kind = "IceCubeEvent"
          test = 0
          healpix_url = "https://gracedb.ligo.org/api/superevents/S190503bf/files/bayes"
          datasource = 'rt'
```
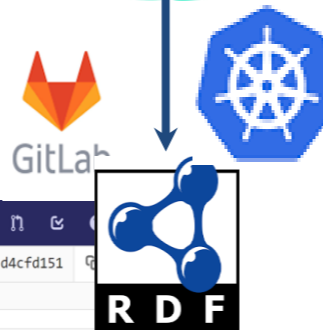
TITLE: GCN CIRCULAR
NUMBER: 25505
SUBJECT: LIGO/Virgo S190828l:
in INTEGRAL SPI-ACS prompt ob
DATE: 19/08/28 08:59:07 GM

VOEvent, GCN, ATel, Kafka, etc

INTEGRAL visibility at 2019-11-19T01:01:29
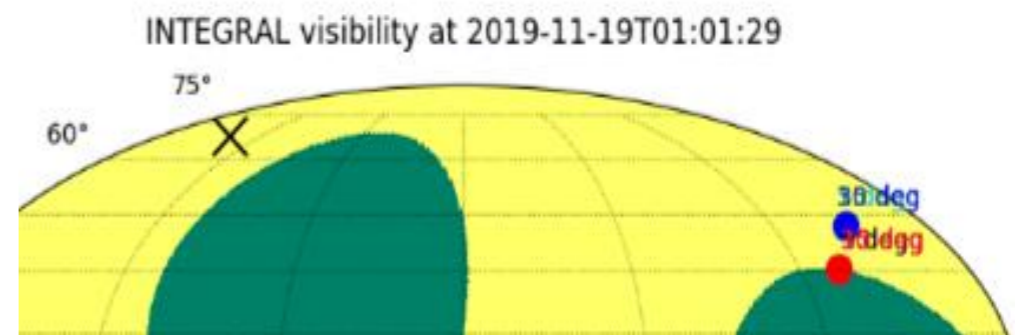
From ISDC to CDCI

- Raw data are not enough, we need to have the ability to run a stramlined analysis and easily access high-level data.
- Unige is working on financial support for a common data center infrastructure (CDCI)
- As part of this, we are making a pilot project for an online tool for INTEGRAL data analysis and long-term preservation of S/W and archive
- It was extended also to another missions at UNIGE: Polar
- Reproducible workflows-> OSA simplified, calibration workflows

Cross calibration activities

- We are working on integrating the dashboard concept to a living cross calibration archive (sketched at IACHEC)
- We will aim at an automatic fetching of calibration observations from IACHEC and provide comparison with INTEGRAL
- We plan to perform automated tests to check the cross-calibration of INTEGRAL instruments
- This is very relevant also for the development of ISGRI calibration

CROSS-calibration
Her X-1 checks for cyclotron line energy

- We collected all simultaneous observation of INTEGRAL and NuSTAR and compare them (preliminary results)
- We are using SPI or NuSTAR data to benchmark ISGRI energy calibration (systematic differences in OSA11 to be understood)

data and folded model
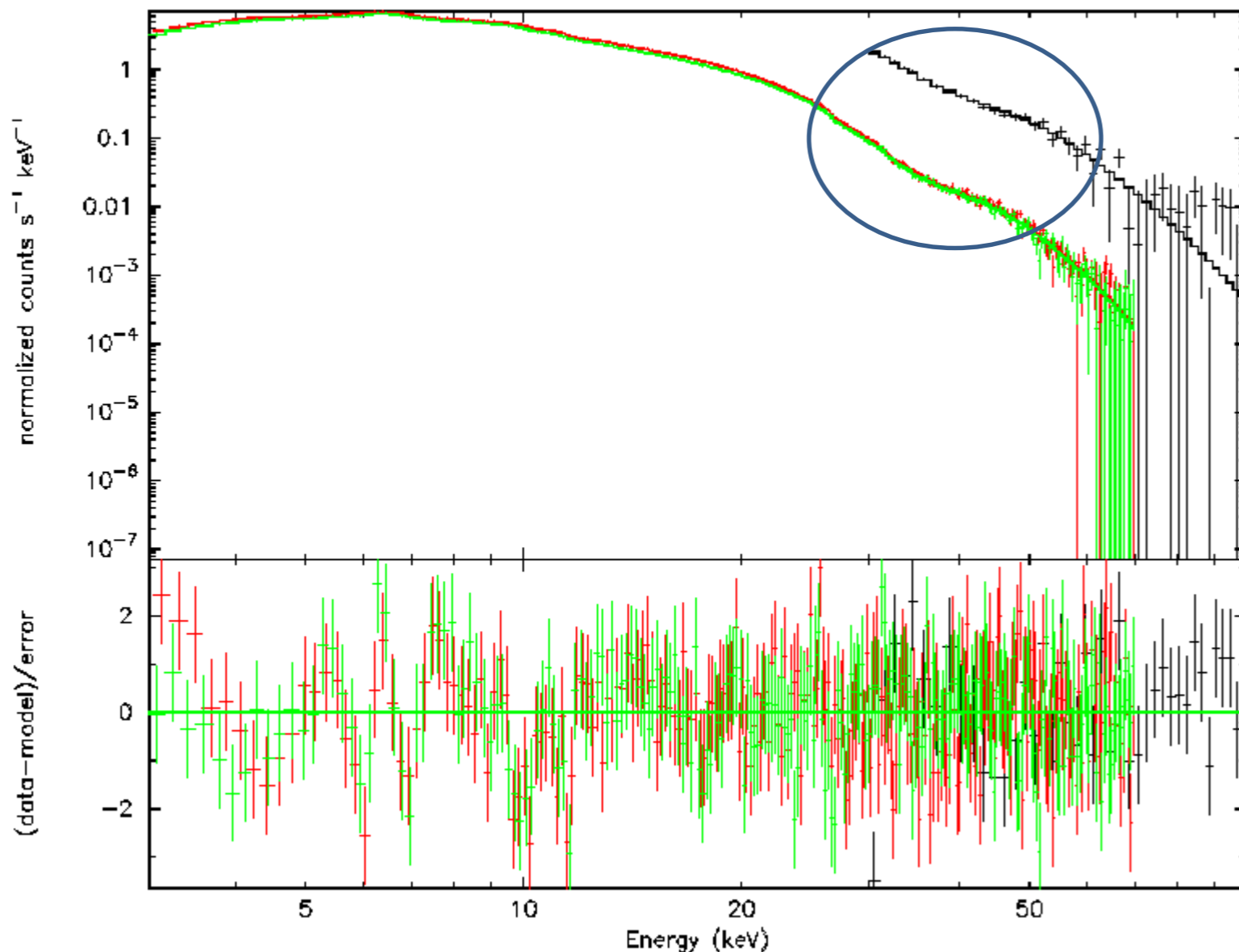
REV 2105
Par ISGRI Reference (sigma)
LineE 36.02 +/- 0.39 ; 35.37 +/- 0.16 ; 1.6
lg10Flux -7.89 +/- 0.01 ; -7.89 +/- 0.00 ; 0.4

'status': 'OK',

General infrastructure: a game changer

- Need to define standards for data and for data mining
- There is a deep discussion on this point: challenge is the need to be flexible, but also rigorous
- Need to define standards for data manipulation and standardized workflows.
- Availability of docker environment and easily accessible tools as notebooks is a game changer: it is possible to ship s/w and methods in a portable manner. Challenge is to make this searchable and explorable.
- VS developed a way to execute python notebooks having inputs and outputs and call them as functions. Translated in common workflow language (standard for other projects, e.g. Reana at CERN)
- Building an infrastructure to expose the results of analysis

Offline Data Analysis (from OSA to ODA)

- We also run OSA executable from a web tool or python API
- We will have a public version for public data (~1-year old)
- We have a private internal version with access to NRT data for operations and transients.

A python API for OSA

- It is possible to call OSA runs using python
- However, processing is done at ISDC with insufficient computing resources
- There is no current solution to this issue, but ESA project on **Astro-Lab** could help
- Backend is portable, but it requires work to be ported

```python
jemx2_spectrum = disp.get_product(instrument="jemx",
                jemx_num='2',
                product="jemx_spectrum",
                product_type="Real",
                osa_version='OSA10.2',
                E1_keV=3.0,
                E2_keV=20.0,
                scw_list=scw_list_str)
```

171500630010.001,171500870010.001,171500650010.001
waiting for remote response, please wait run_analysis https://www.astro.unige.ch/cdci/astrooda/dispatch-data
the job has been submitted on the remote server
/ the job is working remotely, please wait status=done - job_id=-6756293937004601253 53

query done succesfully!

Examples

https://github.com/cdcihub/oda_api_benchmark/tree/master/examples
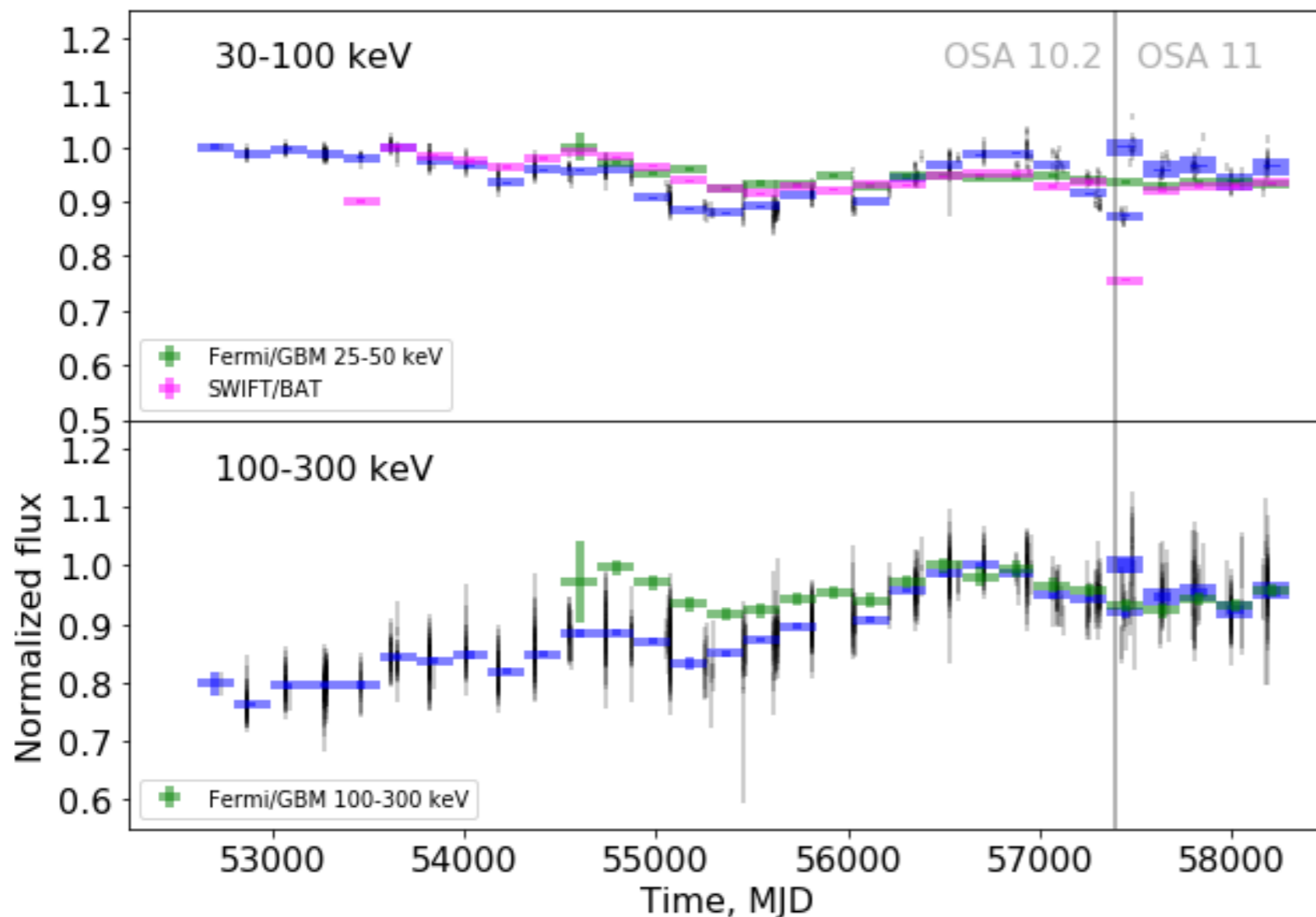

```
data=disp.get_product(instrument
='isgri',
product='isgri_lc',
T1=T1_utc,
T2=T2_utc,
T_format=T_format,
E1_keV=E1_keV,
E2_keV=E2_keV,
query_type='Real',
osa_version='OSA10.2',
            RA=ra,
            DEC=dec,
product_type='Real',
time_bin=time_bin,
selected_catalog=api_cat)
```

Reproducible and storable

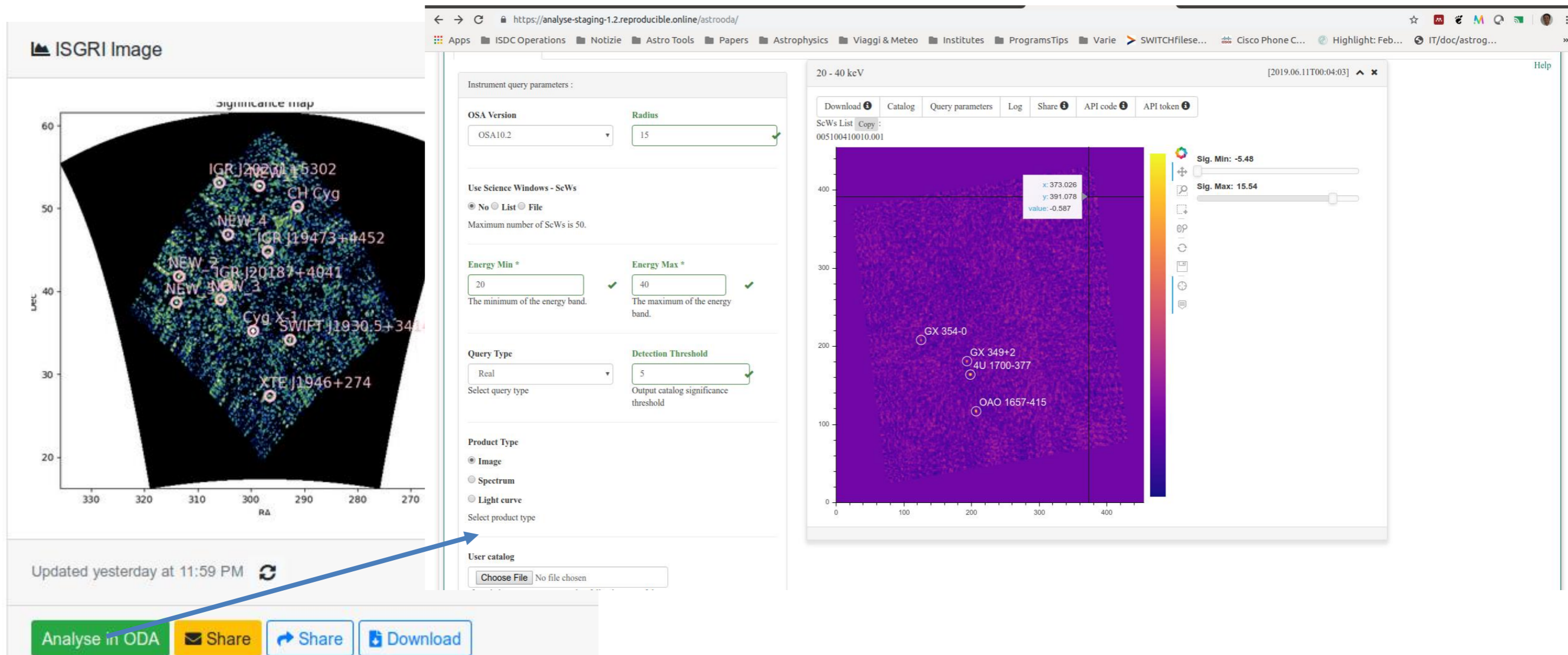- The system is built with internal cache to save intermediate products.
- The second time you make the same query, results are faster.
- Backend can be deployed virtually anywhere, because it is based on a "singularity" cluster, which runs science windows in parallel.
- The extension to other queue systems (slurm) is not done.
- It needs a system called k8s that listens to requests and dispatches them.

ODA current limitations

- We have very limited computing resources and virtually no sysadmin supports (self administrated cluster)
- We need to limit science window number to 50 per run to avoid overcharge (it can be changed)
- We have very limited human resources (~2 FTEs)
- JEM-X is not fully developed.
- Working towards possible DataLabs integration at ESA

Future plans and wishes

- The current archive of cached data is constructed with a lineage of data and dependencies as in a arborescence, a noSQL database.
- User requests will effectively create an archive of results accessible with a limited latency.
- For a *legacy archive*, we should also create a SQL data base of pre-computed results with active links to the processing. This will allow an easy access to spectra, light curves and images.

A prototype of data gallery

- We wish to populate an archive of relevant results with links to the online analysis results (images, spectra lightcurve), but above all to the workflow having generated it.

- With a simple click, we can pass from the image to the workflow producing it. Access source files and in case modify the analysis.