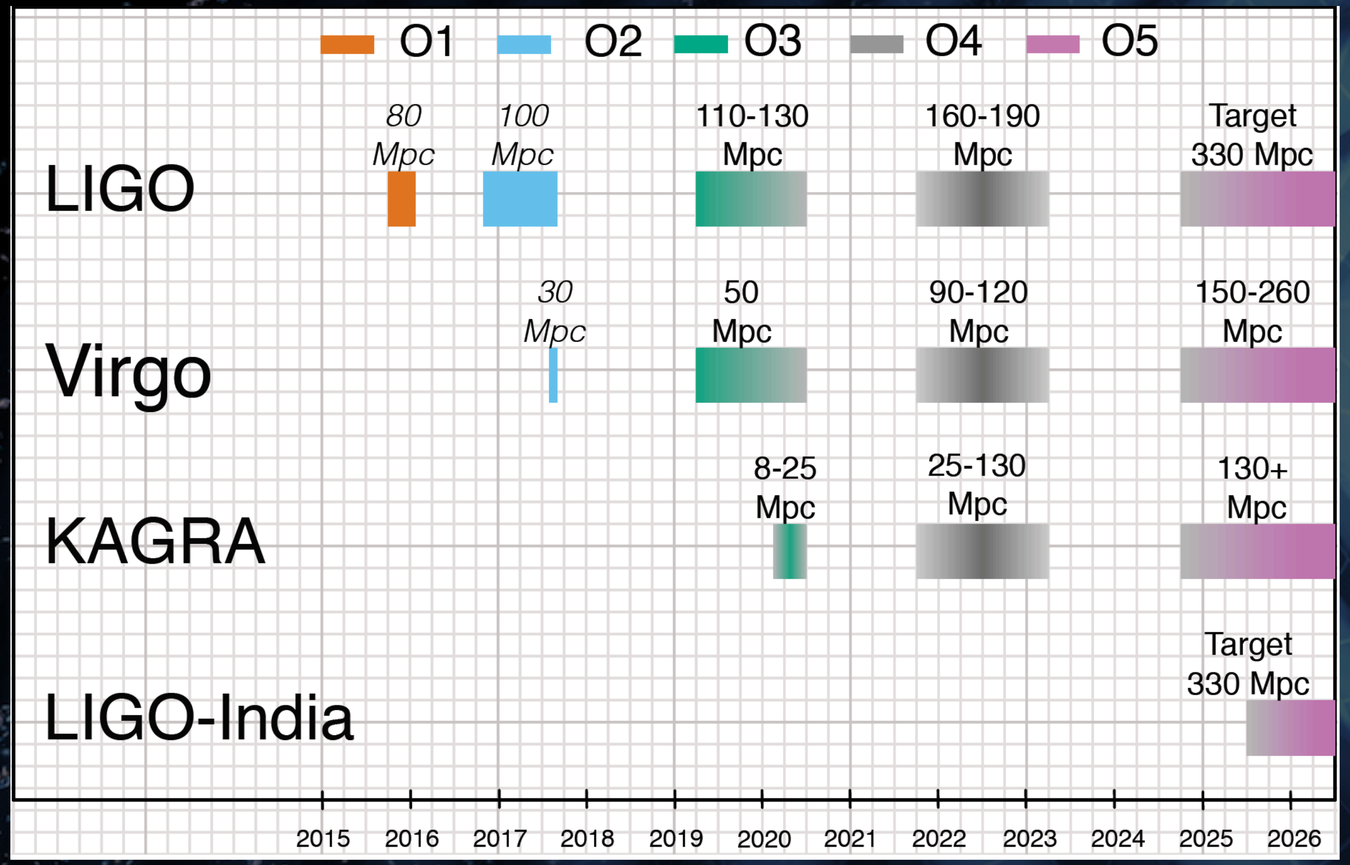
Accessing public gravitational data using the GWOSC website

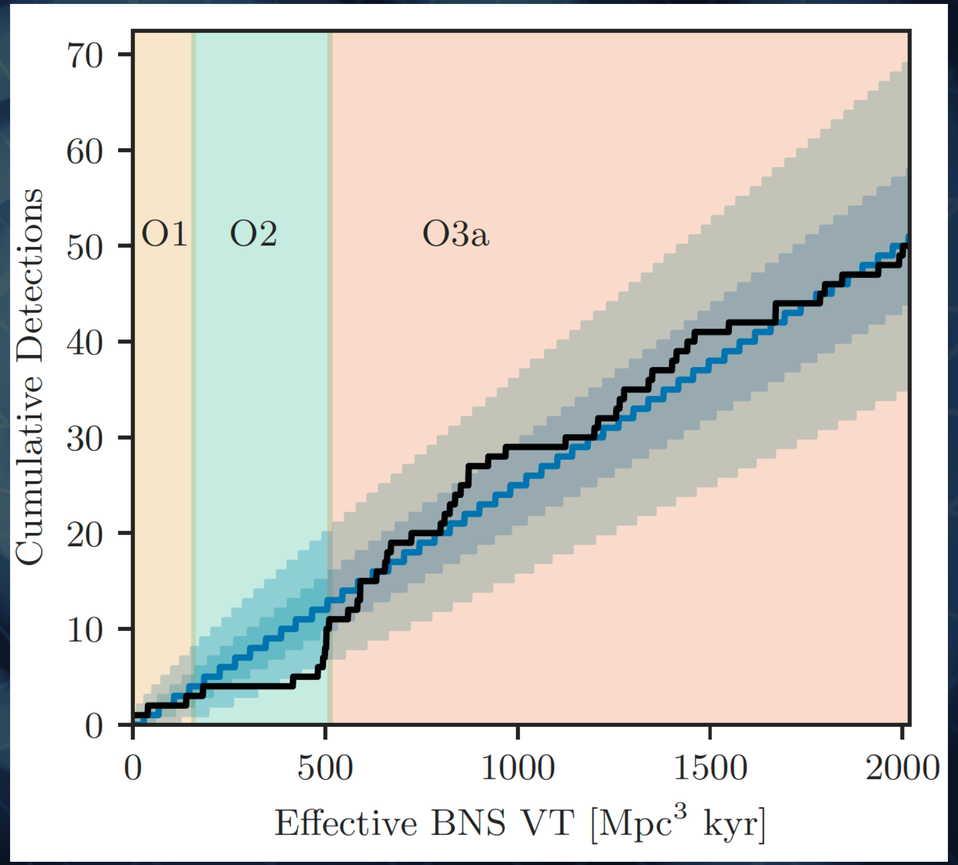
Agata Trovato on behalf of the LIGO and Virgo Collaborations
APC, CNRS/IN2P3, Univ. De Paris

Observing runs and detections

The detectors alternate periods of data taking and periods of upgrades to the machines



Cumulative number of transient GW events detected and published until now



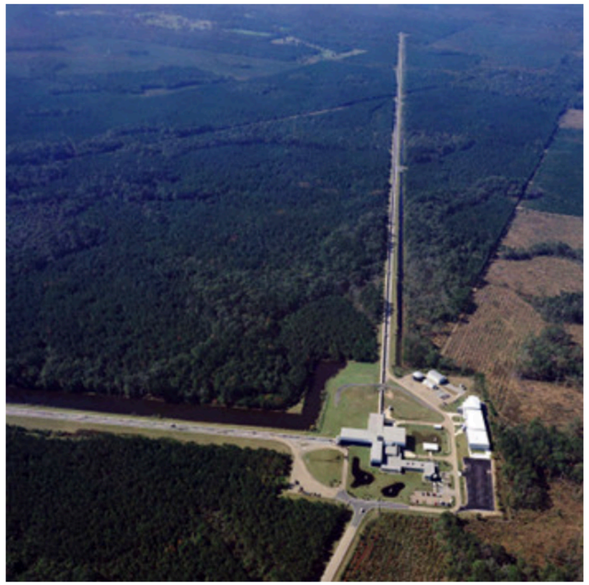
Gravitational Wave Open Science Center

Data Software Online Tools About GWOSC

The Gravitational Wave Open Science Center provides data from gravitational-wave observatories, along with access to tutorials and software tools.

LIGO Hanford Observatory, Washington

(Credits: C. Gray)



LIGO Livingston Observatory, Louisiana

(Credits: J. Giaime)

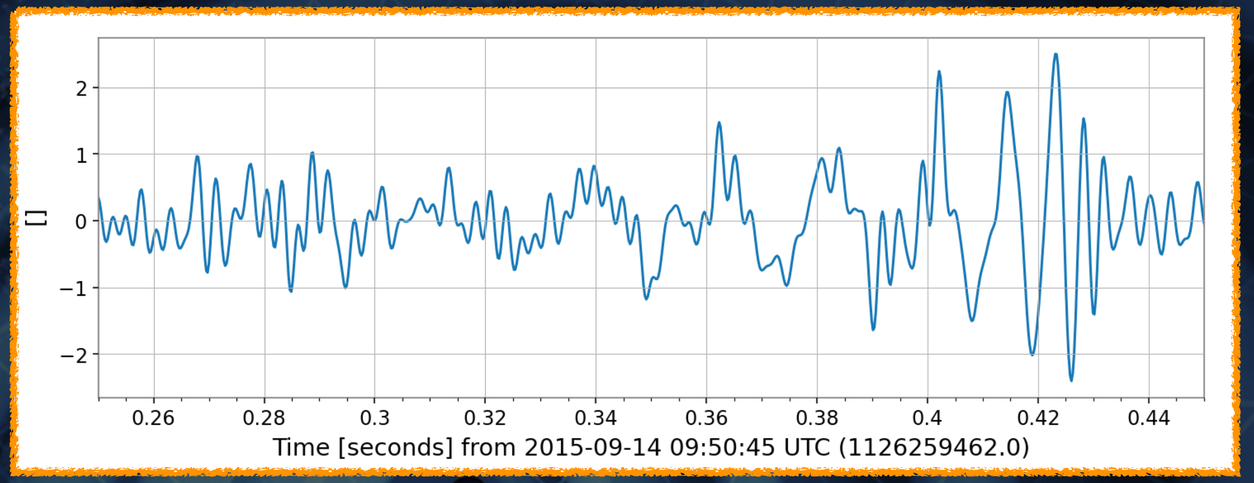
Virgo detector, Italy

(Credits: Virgo Collaboration)

https://www.gw-openscience.org/

Available data products

Strain Data

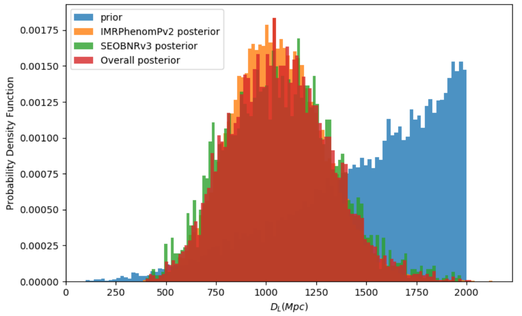


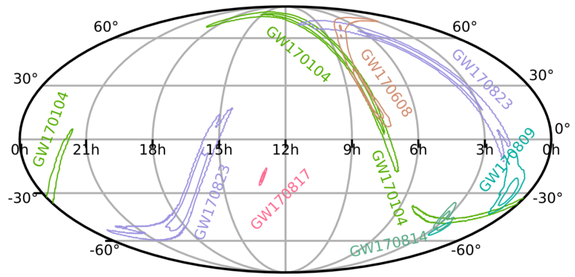
Segments (Timelines)

Start	Stop	Duration
1164559440	1164559654	214
1164560599	1164561392	793
1164562093	1164569775	7682

Analysis Results

(mainly related to events or event catalogs)

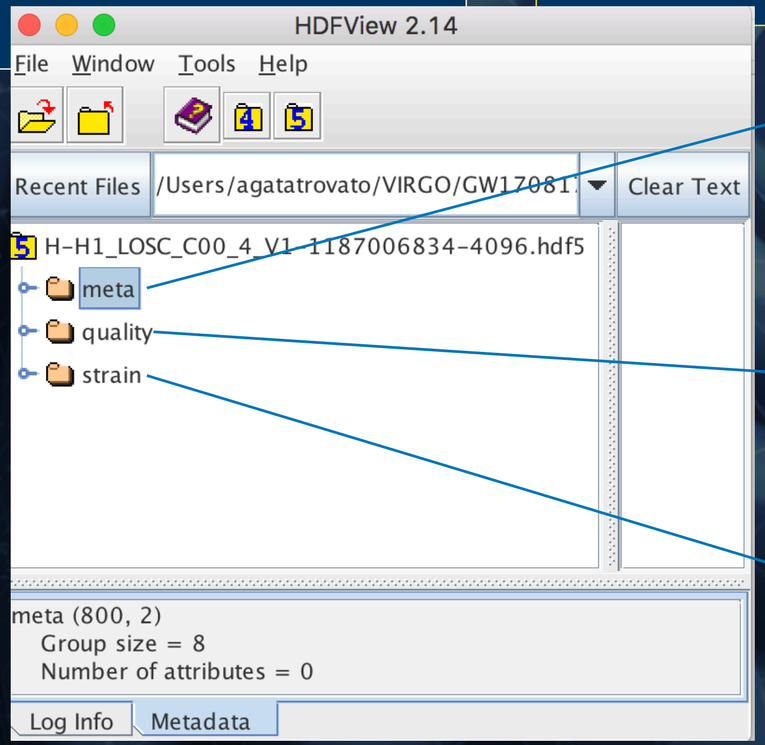




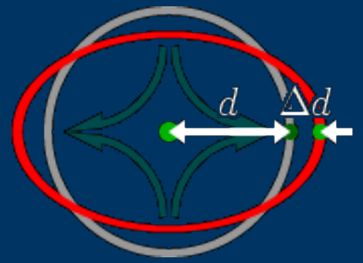
LIGO/Virgo data in GWOSC files

- LIGO/Virgo data: strain, data quality and hardware injections
- LIGO/Virgo data are arranged in files provided in different formats:
  - HDF5: easily readable in python, MATLAB, C/C++, and IDL
  - Frame format (.gwf)
  - Text file

You can use HDFView to quickly see what is inside the file



Reminder: strain



$$h = \frac{\Delta d}{d} = \frac{change\ in\ relative\ position}{separation}$$

Meta-data for the file. This is basic information such as the GPS times covered, which instrument, etc.

Refers to data quality. The main item here is a 1 Hz time series describing the data quality for each second of data.

Strain data from the interferometer. In some sense, this is "the data", the main measurement performed by LIGO/Virgo.

GWOSC releases

- Two different types of data release:
  - **Gravitational wave data surrounding discoveries**
  - **Data taken during a whole observation run**

For details on LVC open data see: SoftwareX 13 (2021) 100658

Data	Date of release
Special events	WIth publication of the corresponding paper
First Observing run, **O1** (Sep 2015 - Jan 2016)	Jan 2018
GWTC-1 Catalog (O1 + O2 detections)	Dec 2018
Second Observing run, **O2** (Nov 2016 - Aug 2017)	Feb 2019
GWTC-2 Catalog (O3a detections)	Oct 2020
First block of the third observing run **O3a** (Apr 2019 - Sep 2019)	Apr 2021

Data management plan at: https://dcc.ligo.org/LIGO-M1000066/public

Finding GWOSC data

Data | Software | On

- Strain Data
- Event Portal
- Timelines
- Auxiliary Channels
- Low Latency Alerts

Events and Catalogs

Event Portal

Large Data Sets

For users of computing clusters or if accessing large amounts of data, CernVM-FS is the preferred method to access public data.

CVMFS Docs

Distributed filesystem that will allow you to mount the data local to your computer

Auxiliary Data Release

Time Range: 3 hours around event GW170814 (August 14, 2017)
Detectors: H1 and L1
Description: Around 1,000 channels that monitor the LIGO instruments and surrounding enviornment.

Auxiliary Data

O3a Data Release

O3 Time Range: April 1, 2019 through October 1, 2019
Detectors: H1, L1 and V1

~1.1 TB | ~3.5 TB

4 kHz Data | 16 kHz Data | Documents

Timeline

O2 Data Release

O2 Time Range: November 30, 2016 through August 25, 2017
Detectors: H1, L1 and V1

~0.8 TB | ~3.2 TB

4 kHz Data | 16 kHz Data | Documents

Timeline

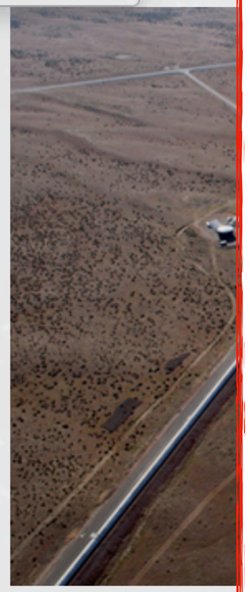
O1 Data Release

O1 Time Range: September 12, 2015 through January 19, 2016
Detectors: H1 and L1

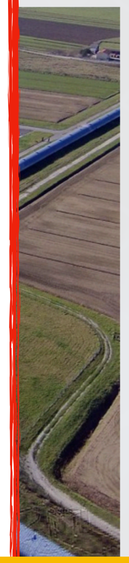
~0.3 TB | ~1.4 TB

4 kHz Data | 16 kHz Data

Timeline



LIGO Hanfo



Scrolling down you get the data for S5 and S6

Example: O3a page

Archive for O3a_4KHZ_R1 dataset

Each data file corresponds to 4096 seconds of GPS time, and may contain up to half a GB. The file may be downloaded in either HDF5 or Frame format. For documentation, see the tutorials.

O3a_4KHZ_R1 start GPS: 1238166018 UTC: 2019-04-01T15:00:00

O3a_4KHZ_R1 end GPS: 1253977218 UTC: 2019-10-01T15:00:00

Next choose your gravitational wave detector:

- V1
- H1
- L1

Now choose the start and end time of the data that you want, either Universal time or GPS. Change either side and the other responds immediately.

	Universal Time (ISO8601)		GPS Time	
Start Time	2019-04-01T15:00:00	⇔	1238166018	OK
End Time	2019-10-01T15:00:00	⇔	1253977218	OK

Choose your output format:

- Time series data in HDF5 and Frame files
- Time series data in HDF5 and Frame files, with data quality guide
- Includes statistics of each file: min/max, band-limited RMS, etc.
- JSON formatted table of files and data quality

Click the button to continue [Continue]

Archive

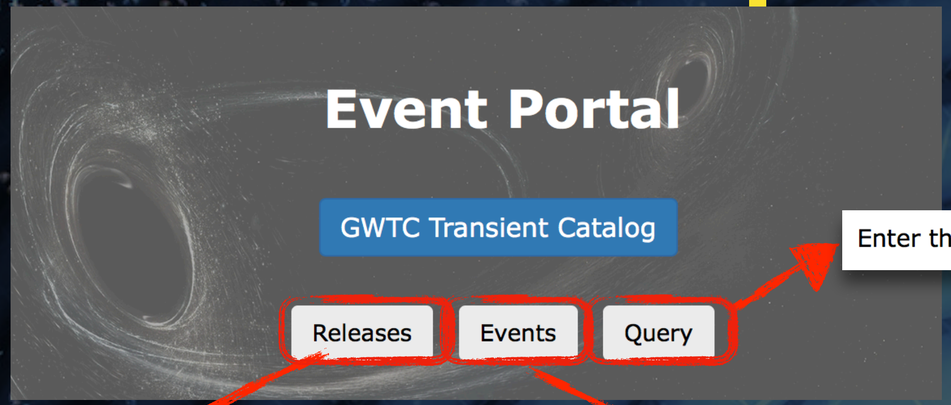
This page lists the files that are available from the **O3a_4KHZ_R1** dataset, between times 1238166018 and 1253977218, from the **L1** detector. Each file covers a 4096-second period, with either full (16 kHz) strain data or downsampled to 4 kHz, and is shown with its size in megabytes. The time of the beginning of the tile is shown as 'GPS start time', and is linked to a timeline showing which parts of the tile have science-mode data available.

For each file, the following statistics are shown:

- The fraction of time when data is available (duty cycle)
- The sky and position averaged distance which the detector would detect a binary neutron star merger with SNR 8 (BNS Range)
- The number of DATA segments in the file (notes)
- 2 band-limited RMS values, representing the root-mean-square average of the amplitude spectral density of h(t). RMS values are shown for the 100-200 Hz band and the 200-1000 Hz band.
- The minimum and maximum values of the dimensionless strain h(t) time series.
- The standard deviation of the dimensionless strain h(t) time series.

Timeline	UTC	Mbytes	HDF5	Frame	Duty Cycle	BNS Range (Mpc)	Notes	ASD RMS 100-200Hz	ASD RMS 200-1kHz	h(t) (Min,Max)	h(t) Stdev
1238163456	2019-04-01T14:17:18	49 MB	HDF5	Frame	37%	124.9	1 DATA segments	4.57e-24	3.43e-23	(-1.6e-18, 1.6e-18)	2.62e-19
1238167552	2019-04-01T15:25:34	87 MB	HDF5	Frame	67%	122.8	1 DATA segments	4.60e-24	3.50e-23	(-1.9e-18, 1.9e-18)	2.49e-19
1238171648	2019-04-01T16:33:50	11 MB	HDF5	Frame	8%	123.3	1 DATA segments	4.88e-24	6.93e-23	(-2.1e-18, 2.1e-18)	3.52e-19
1238175744	2019-04-01T17:42:06	130 MB	HDF5	Frame	100%	122.2	1 DATA segments	4.63e-24	3.41e-23	(-1.9e-18, 1.8e-18)	2.68e-19
1238179840	2019-04-01T18:50:22	130 MB	HDF5	Frame	100%	123.8	1 DATA segments	4.59e-24	3.27e-23	(-3.1e-18, 3.1e-18)	3.14e-19
1238183936	2019-04-01T19:58:38	130 MB	HDF5	Frame	100%	121.0	1 DATA segments	4.65e-24	3.26e-23	(-2.5e-18, 2.5e-18)	2.74e-19

Event portal



Enter the **name or partial name of the event**: (GPS ↔ UTC converter)

Release List

Release Name	Description
GWTC-1-confident	Confident detections from "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs." Additional data products, including PE samples and skymaps, are linked from the documentation at https://doi.org/10.7935/82H3-HH23
GWTC-1-marginal	Marginal triggers from "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs." Additional data products are linked from the documentation at https://doi.org/10.7935/82H3-HH23
GWTC-2	Events from the O3a observation run of LIGO and Virgo, as described in the GWTC-2 catalog paper. These events are also included in a cumulative list of all GWTC events published to date. Details and additional data products are linked from the documentation page.
Initial_LIGO_Virgo	Event data releases from initial LIGO and Virgo, 2005 - 2010. No astrophysical detections were made during this period.
O1_O2-Preliminary	Notable events in O1 and O2 initially published before the GWTC-1 catalog. These data releases may contain preliminary versions of data quality segments and calibration. For additional documentation released at the time of publication, see the "reference" link for each event. Updated information for these events may be found in the GWTC-1 catalog.
O3_Discovery_Papers	Notable events in O3 initially published outside of main catalogs. Associated data releases may contain preliminary versions of data quality segments and calibration. See documentation page for additional notes.

Name	Version	Release	GPS ↓	Mass 1 ($M_\odot$)	Mass 2 ($M_\odot$)	Network SNR	Distance (Mpc)
GW190930_133541	v1	GWTC-2	1253885759.2	$12.3^{+12.5}_{-2.3}$	$7.8^{+1.7}_{-3.3}$	9.8	780^{+370}_{-330}
GW190929_012149	v1	GWTC-2	1253755327.5	$64.7^{+22.4}_{-18.9}$	$25.7^{+14.4}_{-9.7}$	9.9	3680^{+2980}_{-1680}
GW190924_021846	v1	GWTC-2	1253326744.8	$8.8^{+7.0}_{-2.0}$	$5.0^{+1.3}_{-1.9}$	13.2	570^{+220}_{-220}
GW190915_235702	v1	GWTC-2	1252627040.7	$34.9^{+9.5}_{-6.2}$	$24.4^{+5.5}_{-6.0}$	13.1	1700^{+710}_{-640}
GW190910_112807	v1	GWTC-2	1252150105.3	$43.5^{+7.6}_{-6.2}$	$35.1^{+6.3}_{-7.0}$	13.4	1570^{+1070}_{-640}
GW190909_114149	v1	GWTC-2	1252064527.7	$43.2^{+50.7}_{-12.2}$	$27.6^{+13.0}_{-10.9}$	8.5	4770^{+3700}_{-2660}
GW190828_065509	v1	GWTC-2	1251010527.9	$23.8^{+7.2}_{-7.0}$	$10.2^{+3.5}_{-2.1}$	11.1	1660^{+630}_{-610}
GW190828_063405	v1	GWTC-2	1251009263.8	$31.8^{+5.8}_{-3.9}$	$25.9^{+4.4}_{-4.6}$	16.0	2220^{+630}_{-950}
GW190814	v2	GWTC-2	1249852257.0	$23.2^{+1.1}_{-1.0}$	$2.6^{+8.0e-02}_{-9.0e-02}$	22.2	240^{+40}_{-50}
GW190814	v1	O3_Discovery_Papers	1249852257.0	$23.2^{+1.1}_{-1.0}$	$2.59^{+0.08}_{-0.09}$	$25.0^{+0.1}_{-0.2}$	241^{+41}_{-45}
GW190803_022701	v1	GWTC-2	1248834439.9	$36.1^{+10.2}_{-6.7}$	$26.7^{+7.1}_{-7.6}$	8.6	3690^{+2040}_{-1690}
GW190731_140936	v1	GWTC-2	1248617394.6	$39.3^{+11.8}_{-8.2}$	$28.0^{+8.9}_{-8.4}$	8.5	3970^{+2560}_{-2070}
GW190728_064510	v1	GWTC-2	1248331528.5	$12.2^{+7.1}_{-2.2}$	$8.1^{+1.7}_{-2.6}$	13.6	890^{+250}_{-370}

Example of an event page

GW190412

Documentation

Version: v3

All Versions: v1 v2 v3

GPS: 1239082262.2

UTC Time: 2019-04-12 05:30

Release: GWTC-2

GraceDB: S190412m

GCN: Notices • Circulars

Timeline: Query for segments

DOI: https://doi.org/10.7935/99gf-ax93

Data sourced from frame channels.

FrameChannels: [H1:DCS-CALIB_STRAIN_CLEAN_C01, L1:DCS-CALIB_STRAIN_CLEAN_C01, V1:Hrec_hoft_16384Hz]

Data sourced from frame types:

FrameTypes: [H1_HOFT_C01, L1_HOFT_C01, V1Online]

To open GWF files, use channels names as shown for GWTC-1: https://doi.org/10.7935/82H3-HH23

GstLAL Search Pipeline

show / hide parameters

Default SEARCH

GWTC-2 PE for GW190412

show / hide parameters

chi_eff	$0.25^{+0.08}_{-0.11}$
chirp_mass (M_sun)	$15.2^{+0.2}_{-0.2}$
chirp_mass_source (M_sun)	$13.3^{+0.4}_{-0.3}$
final_mass_source (M_sun)	$37.3^{+3.9}_{-3.9}$
luminosity_distance (Mpc)	740^{+140}_{-170}
mass_1_source (M_sun)	$30.0^{+4.7}_{-5.1}$
mass_2_source (M_sun)	$8.3^{+1.6}_{-0.9}$
redshift	$0.15^{+0.03}_{-0.03}$
total_mass_source (M_sun)	$38.4^{+3.8}_{-3.7}$

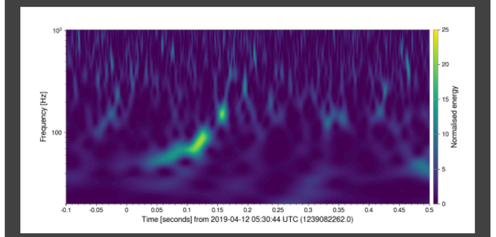
Source File

Posterior Samples DCC Entry

Skymap

Default PE

H1 strain



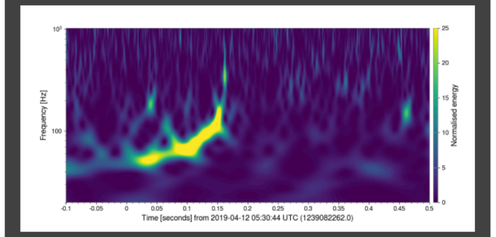
32sec • 16KHz: GWF HDF TXT

32sec • 4KHz: GWF HDF TXT

4096sec • 16KHz: GWF HDF TXT

4096sec • 4KHz: GWF HDF TXT

L1 strain



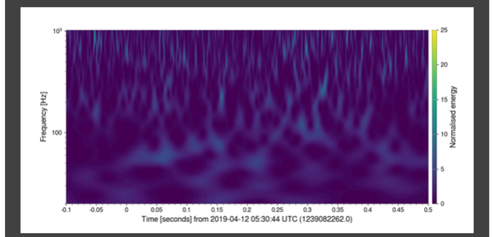
32sec • 16KHz: GWF HDF TXT

32sec • 4KHz: GWF HDF TXT

4096sec • 16KHz: GWF HDF TXT

4096sec • 4KHz: GWF HDF TXT

V1 strain



32sec • 16KHz: GWF HDF TXT

32sec • 4KHz: GWF HDF TXT

4096sec • 16KHz: GWF HDF TXT

4096sec • 4KHz: GWF HDF TXT

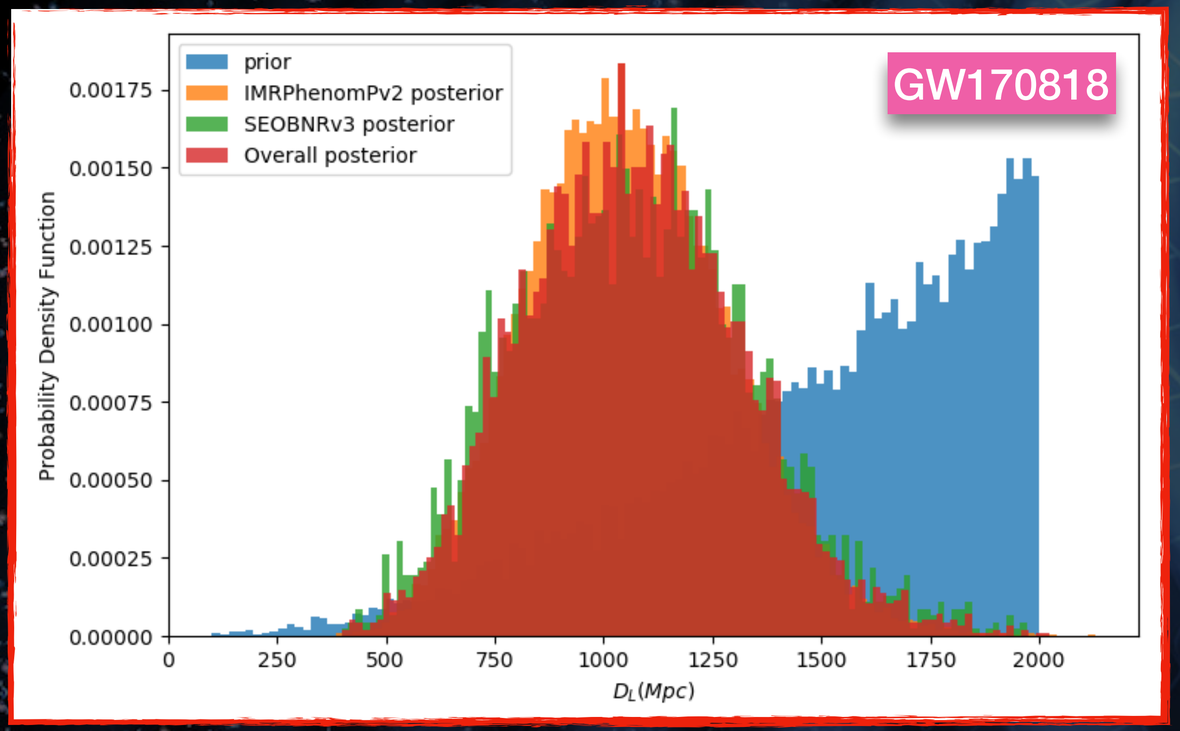
Other analysis products

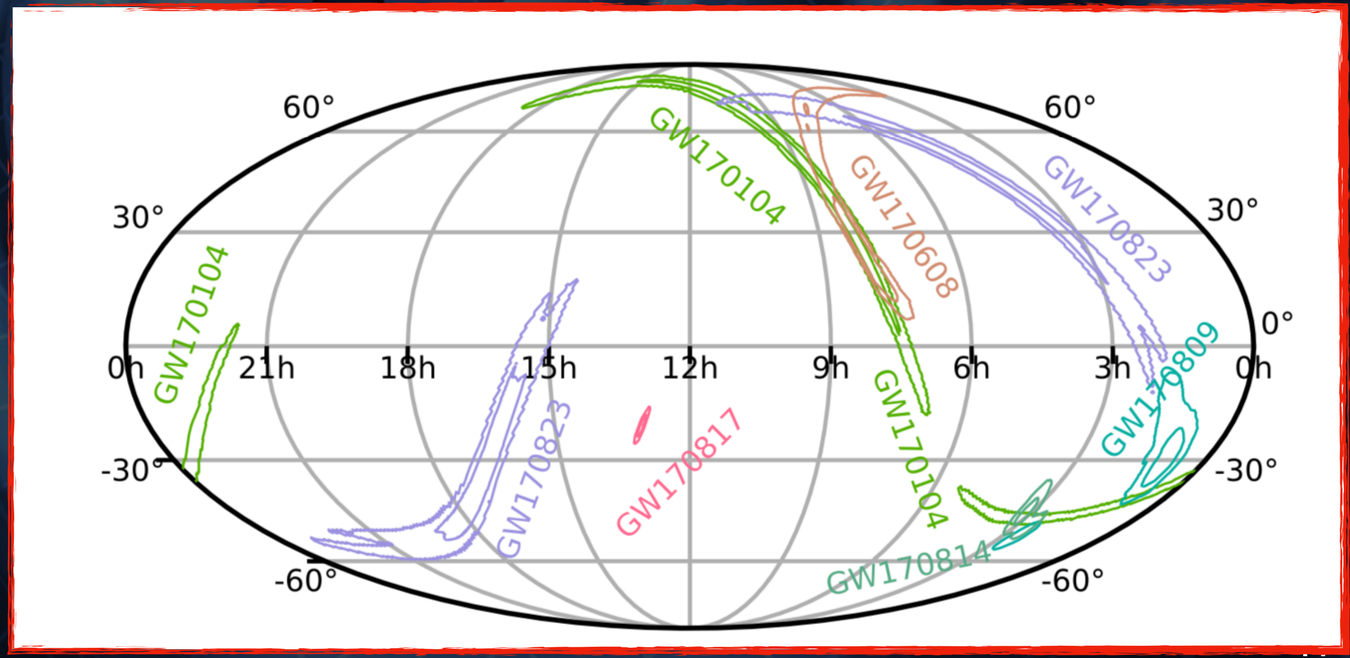
- The event portal contains links to:
  - ✓ Posterior samples
  - ✓ Confidence intervals
  - ✓ Skymaps

Data Products and Publications

GWTC-2 documentation page

- **Catalog Paper and Figures:** P2000061
- **Strain Data:** Event Portal
- **Parameter Estimation Samples & Skymaps**: P2000223
- **Tests of General Relativity:** P2000091
- **Population Properties:** P2000077
- **Search Sensitivity**: P2000217
- **Glitch Models**: P2000289
- **Low-Latency Alerts**: GraceDB

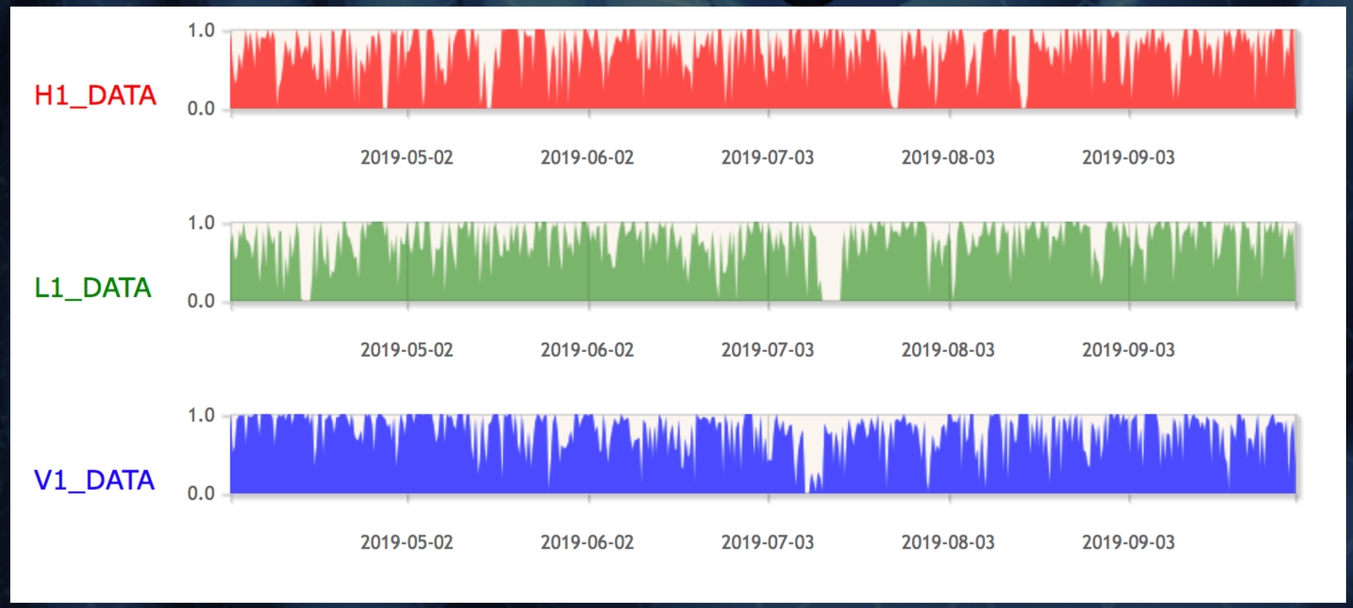




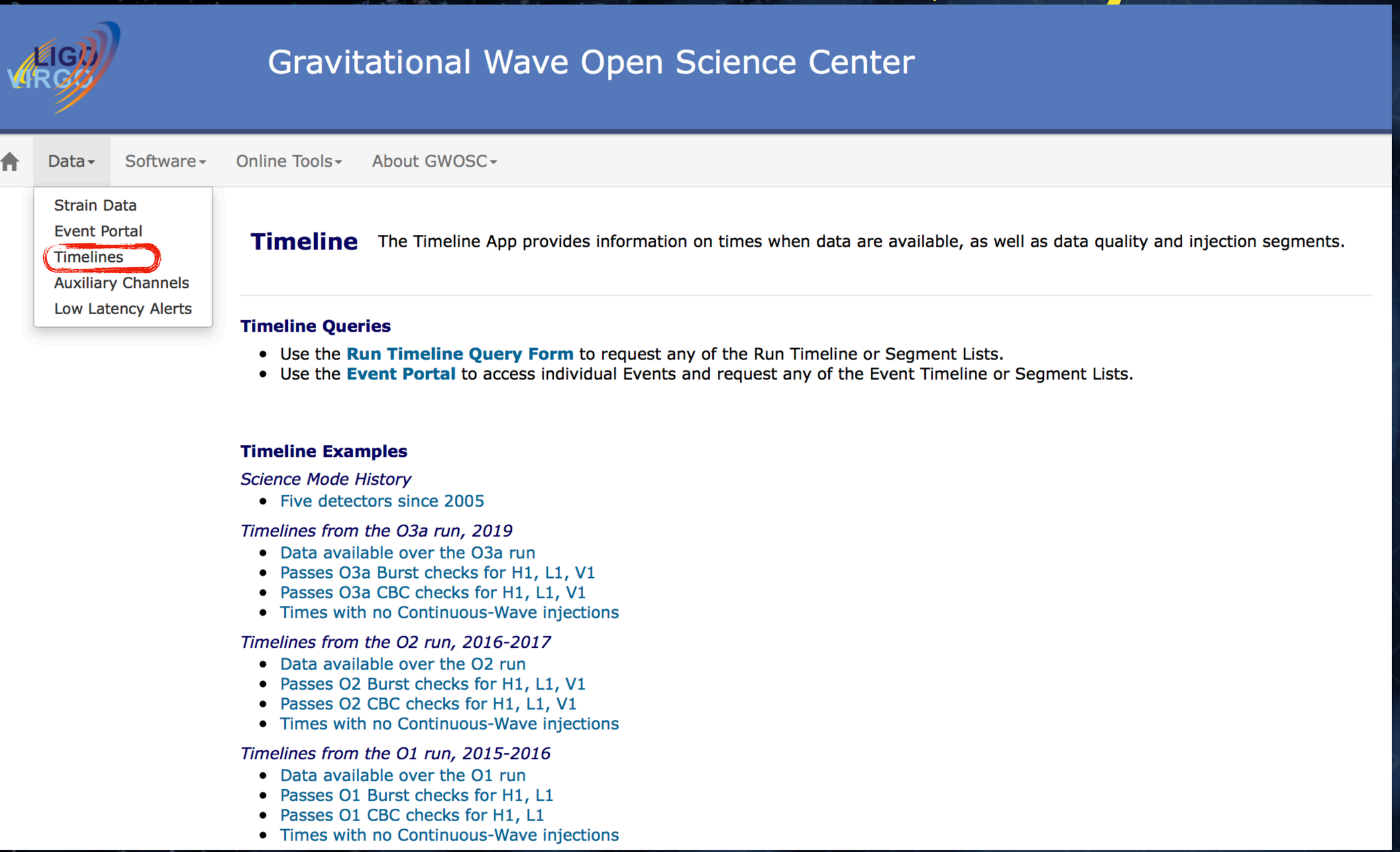
Segment Lists

- LIGO/Virgo detectors are not always 'on' (in science mode)
- Data quality may not meet basic requirements
- Consequence : GW data analysis is applied to data segments (different in each detector)
- The “Timeline” tool of the GWOSC website allows you to select segments that fulfill specific data quality checks in a time period that can be selected specifying GPS time and duration

Timelines for O3a



Run Timeline Query Form



Run Timeline Query Form

Gravitational Wave C

Data Software Online Tools About GWOSC

- Strain Data
- Event Portal
- Timelines
- Auxiliary Channels
- Low Latency Alerts

Timeline The Timeline App provides

Timeline Queries

- Use the **Run Timeline Query Form** t
- Use the **Event Portal** to access individ

Timeline Examples

Science Mode History

- Five detectors since 2005

Timelines from the O3a run, 2019

- Data available over the O3a run
- Passes O3a Burst checks for H1, L1, V1
- Passes O3a CBC checks for H1, L1, V1
- Times with no Continuous-Wave injecti

Timelines from the O2 run, 2016-2017

- Data available over the O2 run
- Passes O2 Burst checks for H1, L1, V1
- Passes O2 CBC checks for H1, L1, V1
- Times with no Continuous-Wave injecti

Timelines from the O1 run, 2015-2016

- Data available over the O1 run
- Passes O1 Burst checks for H1, L1
- Passes O1 CBC checks for H1, L1
- Times with no Continuous-Wave injections

Select Data Set: S5 S6 O1 O2 O3a

Dataset: O3a

GPS start: 1238166018 (2019-04-01T15:00:00)

GPS end: 1253977218 (2019-10-01T15:00:00)

Enter the **starting GPS**: 1238166018 (GPS ↔ UTC converter)

Enter the **end GPS or duration** in seconds: 1253977218

Select display type:

- **Plot**: Plot fraction of time that passes the chosen DQ level in each time bin
- **Segment List**: ASCII segment list (tutorial on segments)
- **JSON Single Segment List**: Segment list for single category in JSON format
- **JSON All Segment Lists**: Segment lists for all categories in JSON format (Takes several minutes to load.)

Select flag names. For plots and JSON, multiple selections can be made, but for segments, only one can be selected. There is more information about these flags here.

☑ H1_BURST_CAT1	☐ L1_BURST_CAT1	☐ V1_BURST_CAT1
☐ H1_BURST_CAT2	☐ L1_BURST_CAT2	☐ V1_BURST_CAT2
☐ H1_BURST_CAT3	☐ L1_BURST_CAT3	☐ V1_BURST_CAT3
☐ H1_CBC_CAT1	☐ L1_CBC_CAT1	☐ V1_CBC_CAT1
☐ H1_CBC_CAT2	☐ L1_CBC_CAT2	☐ V1_CBC_CAT2
☐ H1_CBC_CAT3	☐ L1_CBC_CAT3	☐ V1_CBC_CAT3
☐ H1_DATA	☐ L1_DATA	☐ V1_DATA
☐ H1_NO_BURST_HW_INJ	☐ L1_NO_BURST_HW_INJ	☐ V1_NO_BURST_HW_INJ
☐ H1_NO_CBC_HW_INJ	☐ L1_NO_CBC_HW_INJ	☐ V1_NO_CBC_HW_INJ
☐ H1_NO_CW_HW_INJ	☐ L1_NO_CW_HW_INJ	☐ V1_NO_CW_HW_INJ
☐ H1_NO_DETCHAR_HW_INJ	☐ L1_NO_DETCHAR_HW_INJ	☐ V1_NO_DETCHAR_HW_INJ
☐ H1_NO_STOCH_HW_INJ	☐ L1_NO_STOCH_HW_INJ	☐ V1_NO_STOCH_HW_INJ

Get your Timeline at:

https://www.gw-osc.org/timeline/show/O3a_16KHZ_R1/H1_BURST_CAT1/1238166018/15811200/

Example of a segment

Timeline The vertical axis indicates the fraction of time a flag is on during each "Sample time".

From: **2019-04-01T15:16:12** = GPS 1238166990 Plot width: **4.42 minutes** = 265 s [Zoom out all the way] [Zoom out]

To: **2019-04-01T15:20:37** = GPS 1238167255 Sample time: **1.00 secs** [Coarser resolution] [Finer resolution]

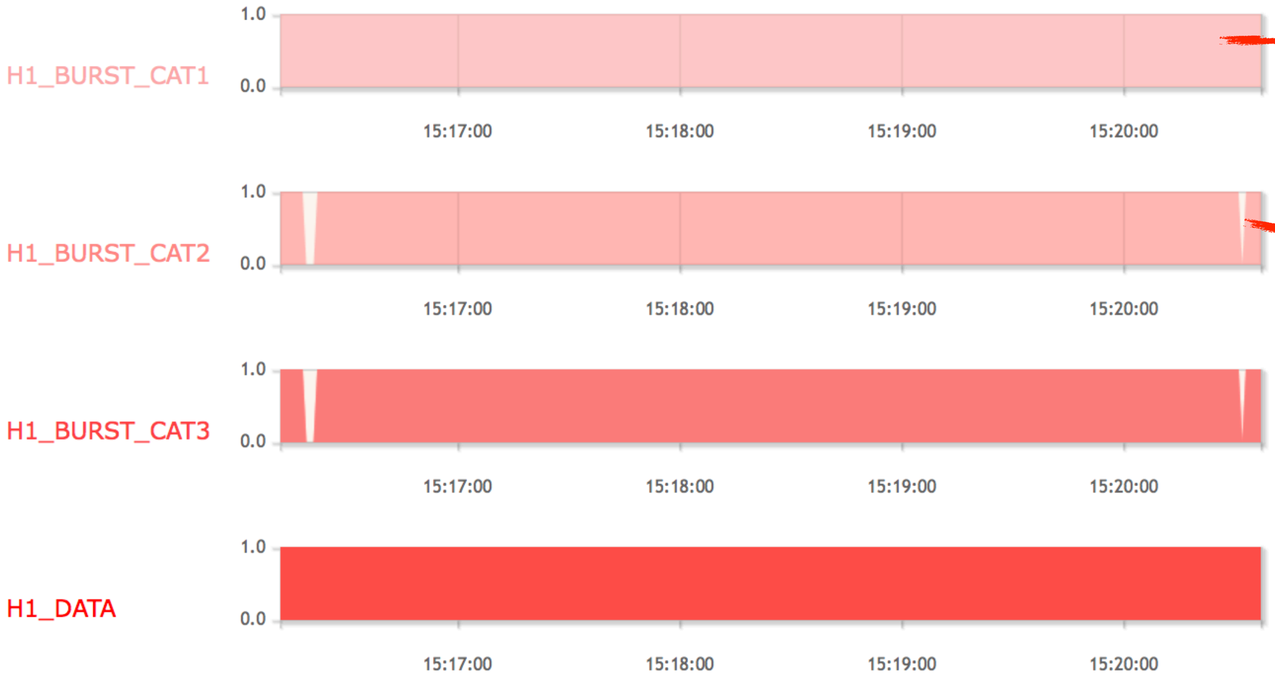
URL for this view | Download these data

Download Segment Lists [?]: H1_BURST_CAT1 H1_BURST_CAT2 H1_BURST_CAT3 H1_DATA

This plot as JSON plot0 plot1 plot2 plot3

Get strain data [?]: H1

To zoom by factor 2, click in any panel.



Start	Stop	Duration
1238166990	1238167255	265

Start	Stop	Duration
1238166990	1238166997	7
1238167000	1238167250	250
1238167251	1238167255	4

Software for GW data

- Software for working with Gravitational Wave Data available to the public: https://www.gw-openscience.org/software/
- API on the GWOSC web site + python client (called “gwosc”)
- Part of the software developed by LIGO/Virgo and open-source

Gwpy

a python package for gravitational-wave astrophysics.

PyCBC

Free and open software to study gravitational waves.

At this workshop

ligo.skymap

The ligo.skymap package provides tools for reading, writing, generating, and visualizing gravitational-wave probability maps from LIGO and Virgo. It includes the rapid sky localization code BAYESTAR, tools for making sky maps from MCMC samples, observation planning utilities, and tools for making beautiful astronomical maps.

Bilby

Bilby: a user-friendly Bayesian inference library.

GstLAL

gstlal provides a suite of GStreamer elements that expose gravitational-wave data analysis tools from the LALSuite library for use in GStreamer signal-processing pipelines.

LALSuite

The LSC Algorithm Library Suite (LALSuite) is a collection of component packages, each of which is tagged, packaged, and released separately.

Readligo.py

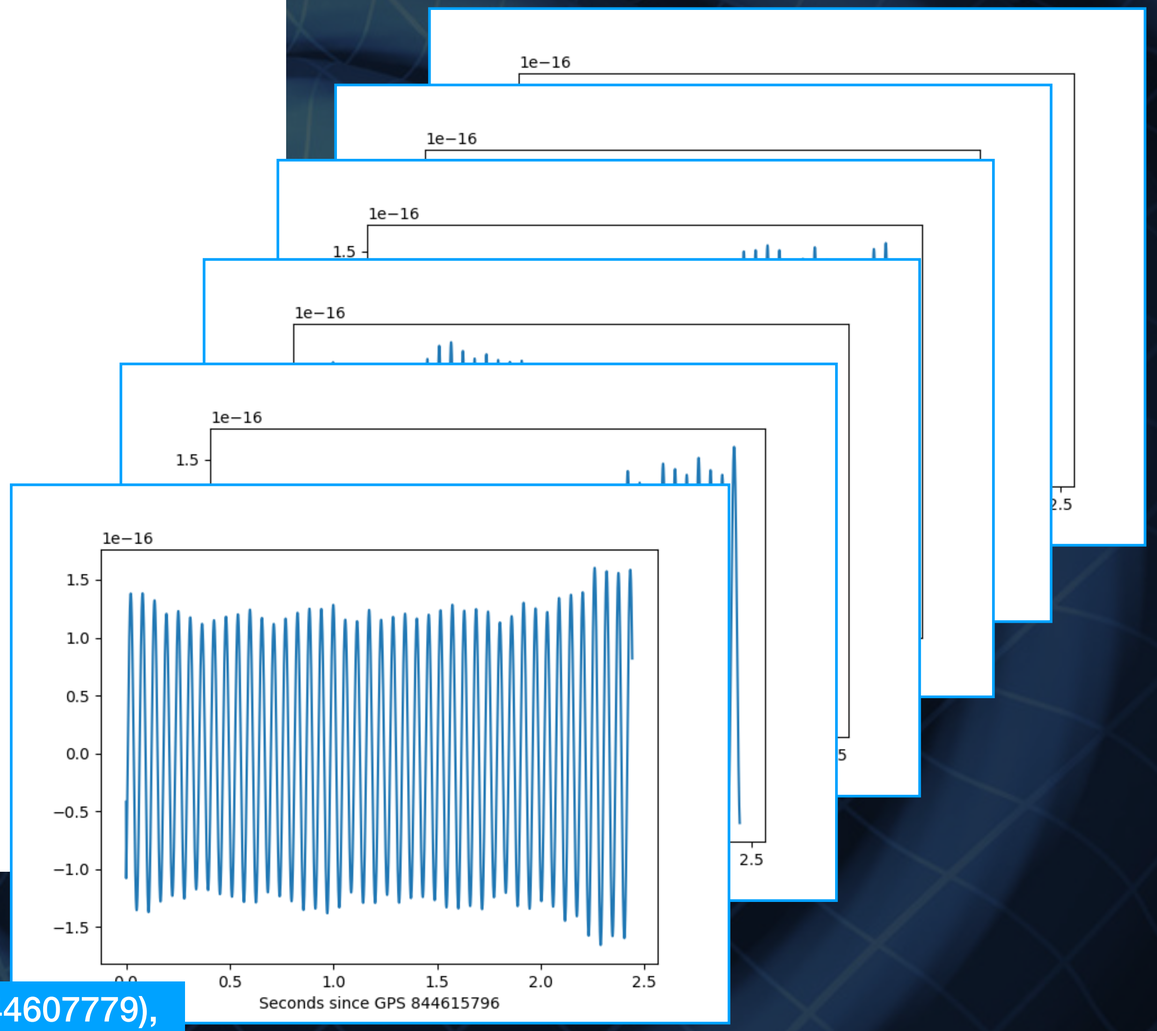
```
import numpy as np
import matplotlib.pyplot as plt
import readligo as rl

start = 844605900
stop =  844615900

segList = rl.getsegs(start, stop, 'H1', flag='BURST_CAT2')
print (segList)
#-----------------------------------------
# Plot a few seconds of each "good" segment
#-----------------------------------------
N = 10000
for (begin, end) in segList:
    # -- Use the getstrain() method to load the data
    strain, meta, dq = rl.getstrain(begin, end, 'H1')

    # -- Make a plot
    plt.figure()
    ts = meta['dt']
    rel_time = np.arange(0, end-begin, meta['dt'])
    plt.plot(rel_time[0:N], strain[0:N])
    plt.xlabel('Seconds since GPS ' + str(begin) )
plt.show()
```

https://www.gw-openscience.org/s/sample_code/use_readligo.py



Output:

SegmentList([(844605900, 844606294), (844606594, 844606649), (844606759, 844607779), (844608784, 844609581), (844612612, 844615722), (844615796, 844615900)])

Tutorials and workshops

Gravitational Wave Open Data Workshop #4 (2021)

Remote workshop, May 10-14, 2021

Workshop web site

This workshop

Gravitational Wave Open Data Workshop #3 (2020)

Lecture videos and tutorials from 2020 workshop

Course Material

Gravitational Wave Open Data Workshop #2 (2019)

Lecture videos and tutorials from 2019 workshop

Course Material

Gravitational Wave Open Data Workshop #1 (2018)

Lecture videos and tutorials from 2018 workshop

Course Material

Recorded lectures

Day 1: May 26 - Recording

08:00 - Workshop Welcome
08:15 - Introduction to LIGO detectors: Gregg Harry (slides)
09:00 - Public LIGO/Virgo data: Jonah Kanner (slides)
09:30 - Coffee Break
09:45 - Data quality and GWpy: Marissa Walker (slides)
10:30 - Session end
--- Break ---
A hands-on session will follow lectures (see below)

Day 2: May 27 - Recording

08:00 - Introduction to CBC: Alan Weinstein (slides)
08:45 - Searches with PyCBC: Derek Davis (slides)
09:30 - Coffee Break
09:45 - Parameter estimation with bilby: Sylvia Biscoveanu (slides)
10:30 - Session end
--- Break ---
A hands-on session will follow lectures (see below)

Hands-sessions

Gravitational Wave Open Data Workshop #3

Tutorial 1.1: Discovering open data from GW observatories

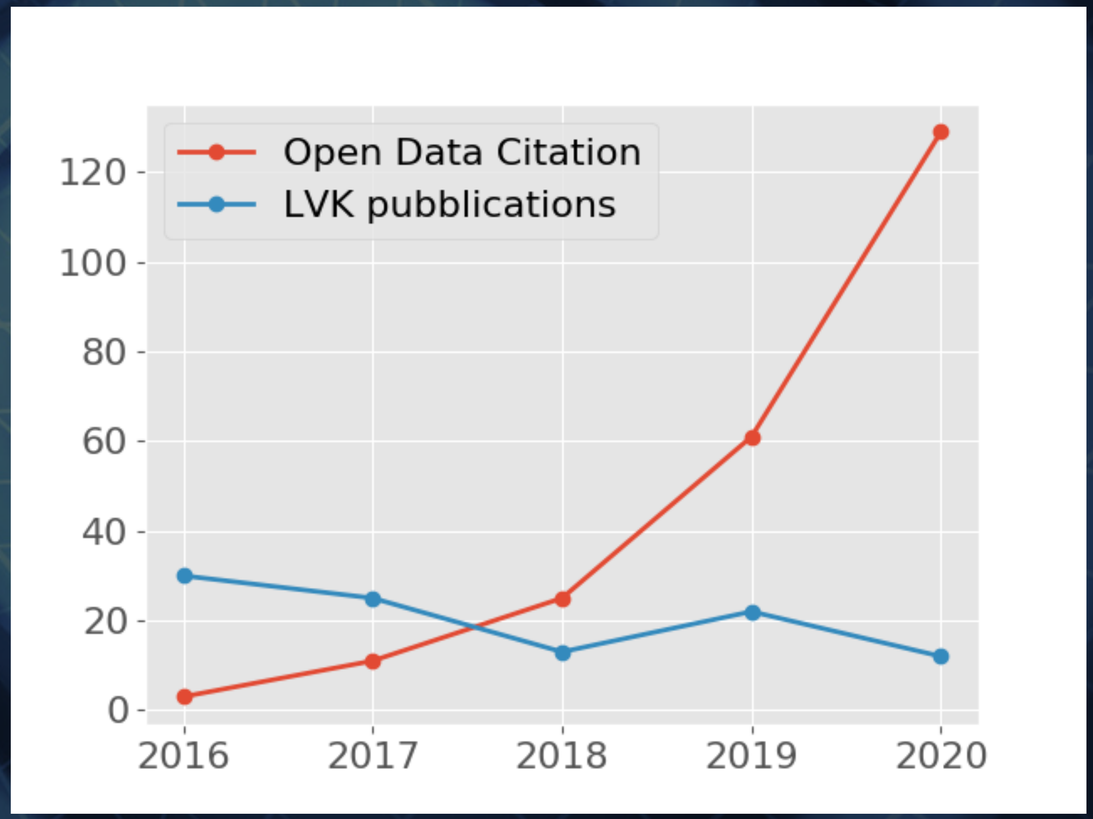
This notebook describes how to discover what data are available from the Gravitational-Wave Open Science Center (GWOSC).

Click this link to view this tutorial in Google Colaboratory

GWOSC Impact

- Examples of projects using GWOSC data: gw-openscience.org/projects/
  - ✓ Professional research, student projects, classroom activities, text books, art projects, workshops, training
  - ✓ Around 3000 visitors each month (unique IP) and thousands of strain file downloads

About 250 published papers acknowledge use of GWOSC (INSPIRE-HEP)



New program for students

Learning Path

Teachers & Students

Objectives

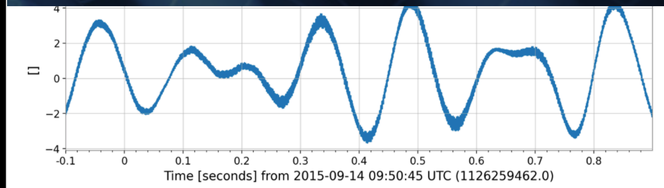
Step 1: Watch Introductory Video (1)

7 Minutes

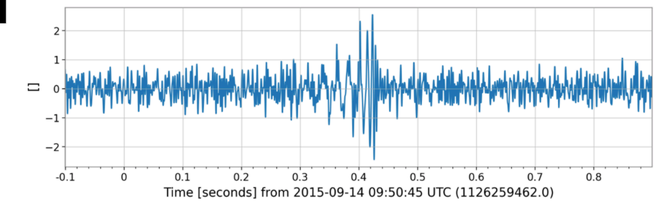
- How are gravitational waves created?
- How are gravitational waves measured?

Footnote: Construction of the KAGRA detector in Japan has been completed since the production of this video.

https://www.gw-openscience.org/path



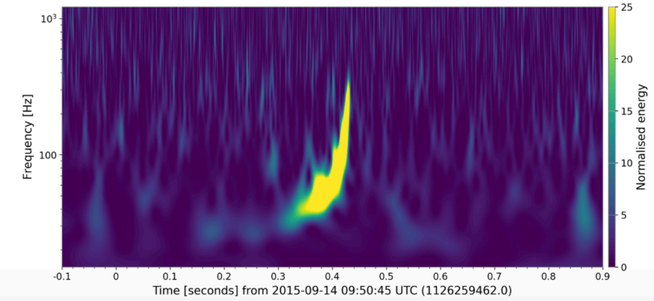
Whitened and Band-passed Data



Download Data as CSV File

See notes +

Q-transform



Detector

H1

Set Plot Parameters

Time Range (seconds)

1.00

0.10 8.00

Whitened and band-passed data

Whiten?

Band-pass frequency range (Hz)

30 400

10 2000

Q-tranform plot

Colorbar Max Energy

25

10 500

Q-value

5

5 120

Overview

- GWOSC is a successful open science project
  - ✓ A large and growing community of users
  - ✓ Data used in a good fraction of the publications related to GW astronomy

- Reproducibility of LVC results (see e. g. arXiv:2010.07244)
- O3a just released; major upcoming data releases include: O3b (expected Oct 2021), catalog for O3b
- A wealth of science to come and to share!

A wealth of science to come and to share!