VESF Data Analysis School

Data Quality and vetos

Didier Verkindt

Virgo-LAPP
CNRS - Université de Savoie

Outline

- Why data analysis needs vetos
- Some preliminary definitions
- A simple data analysis tool: frequency band rms
- Examples of glitches
- Some useful tools
- How to develop a quality flag
- Performance of a DQ flag
- Example of DQ flag used as veto

Why data analysis needs vetos

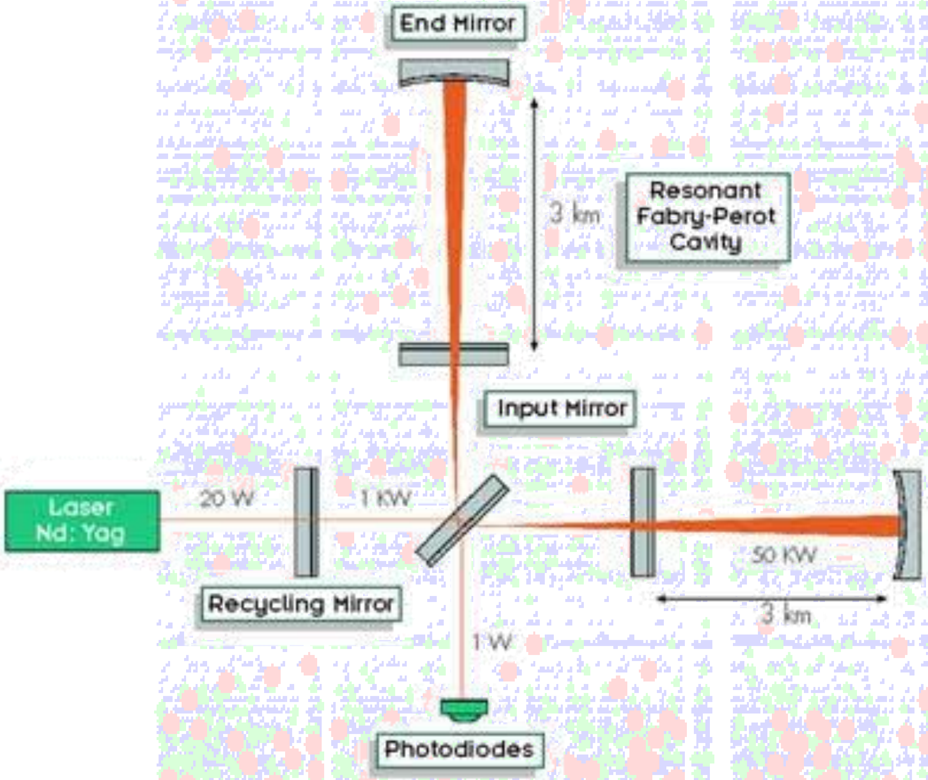
Why data analysis needs vetos

The Virgo interferometer is sensitive to several sources of noise:

Most of them are non-stationary with various time scales and couplings

→ short glitches or long transient noisy periods

- Seismic noise coupled to diffused light
- Magnetic noise, powerline
- Injection system
- TCS system
- OMC alignment
- Etc…



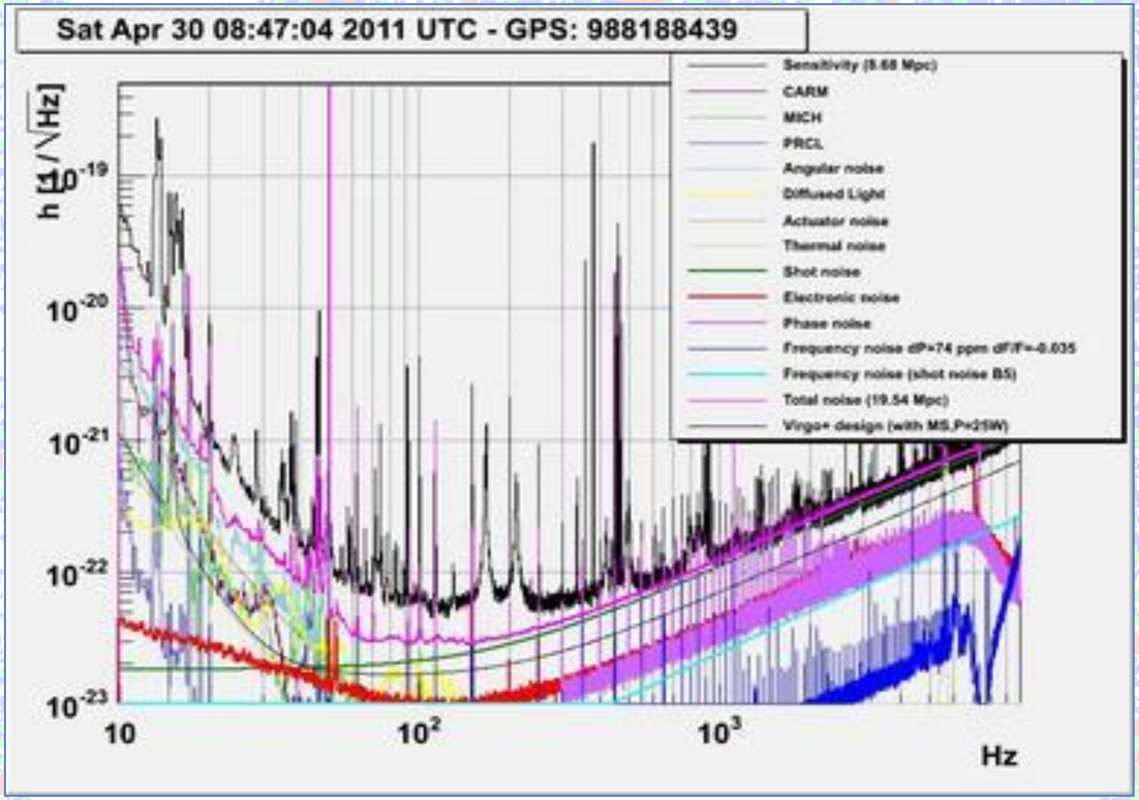
Why data analysis needs vetos

The Virgo interferometer is sensitive to several sources of noise:

Most of them are non-stationary with various time scales and couplings

→ short glitches or long transient noisy periods

- Seismic noise coupled to diffused light
- Magnetic noise, powerline
- Injection system
- TCS system
- OMC alignment
- Etc…

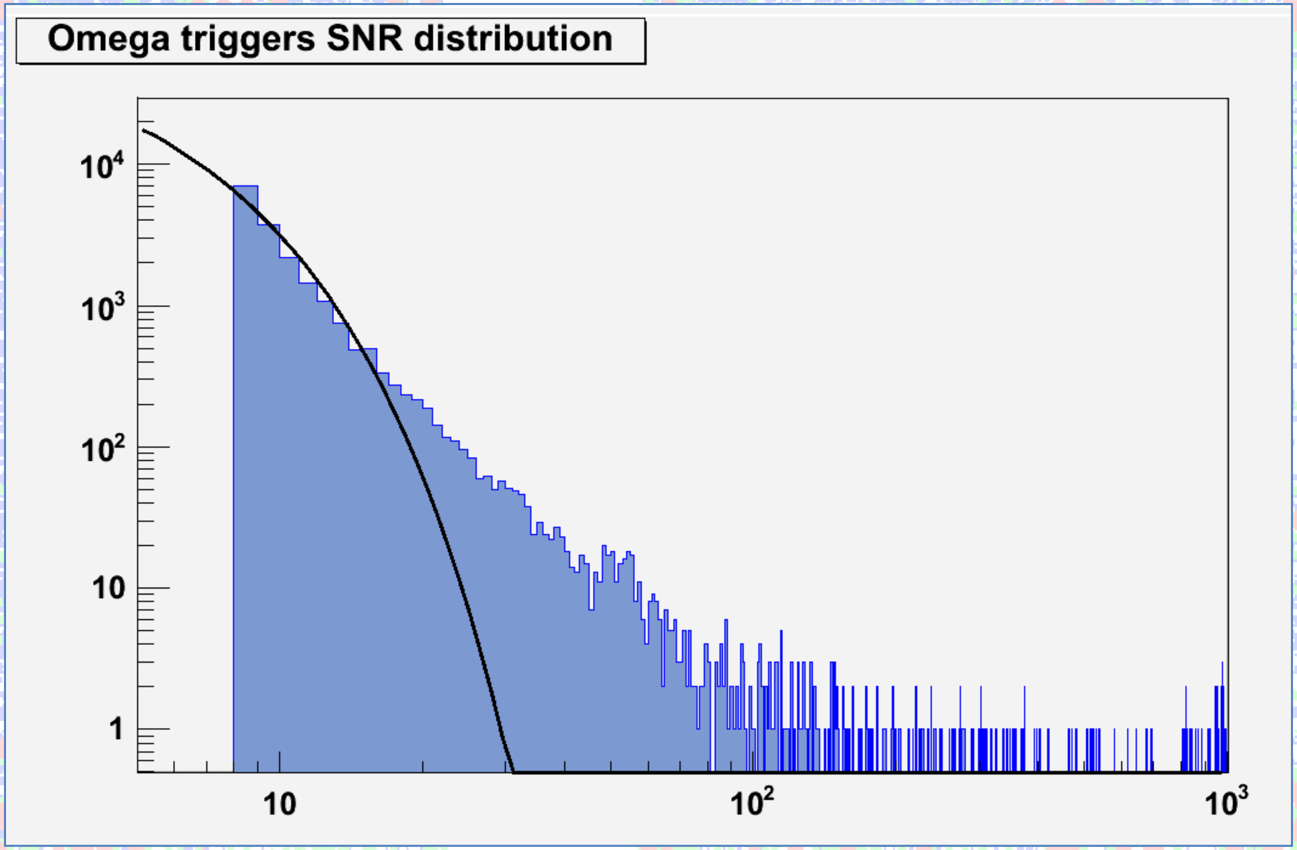


Why data analysis needs vetos

One interferometer like Virgo:

High glitchiness, large SNR distribution tail

→ difficulty to separate real GW events from noise background

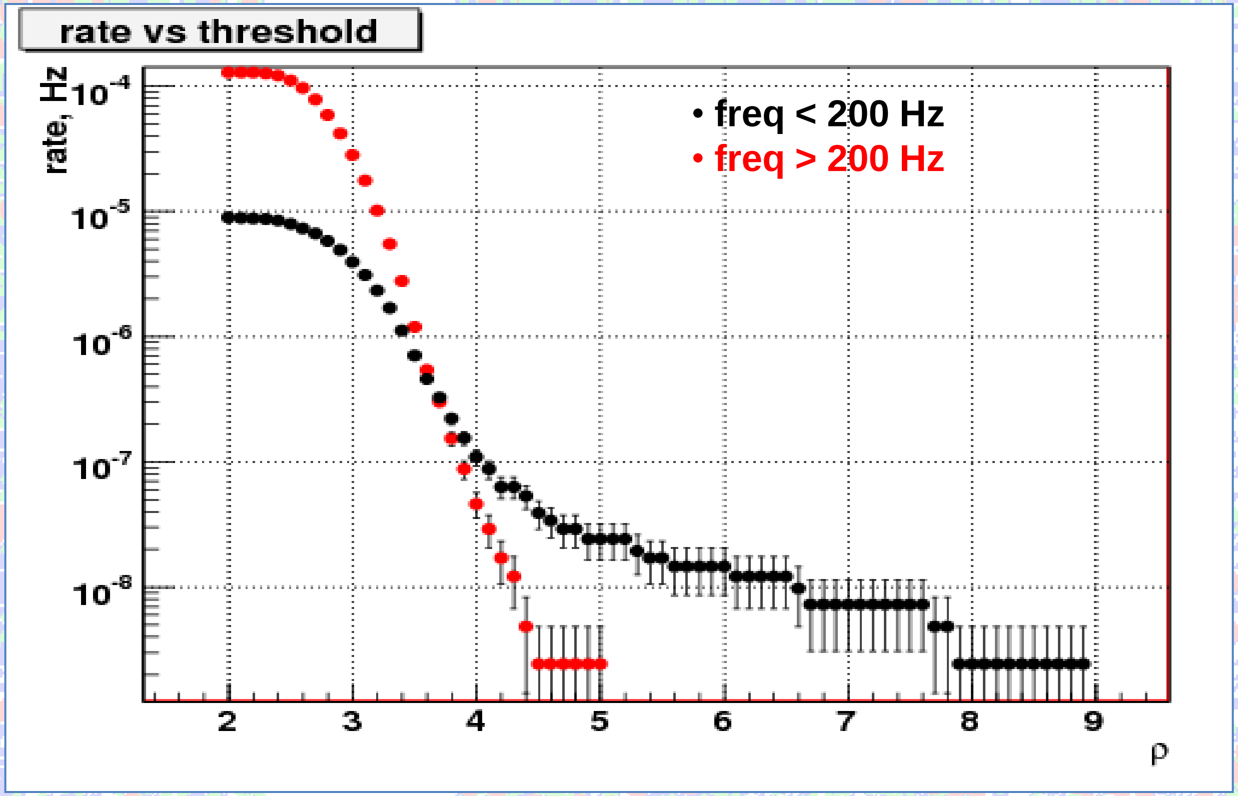


Why data analysis needs vetos

Several interferometers in coincidence (or in coherence):

Lower glitchiness, but still highly non-gaussian SNR distribution tail

→ Reduced statistical significance of GW events versus background



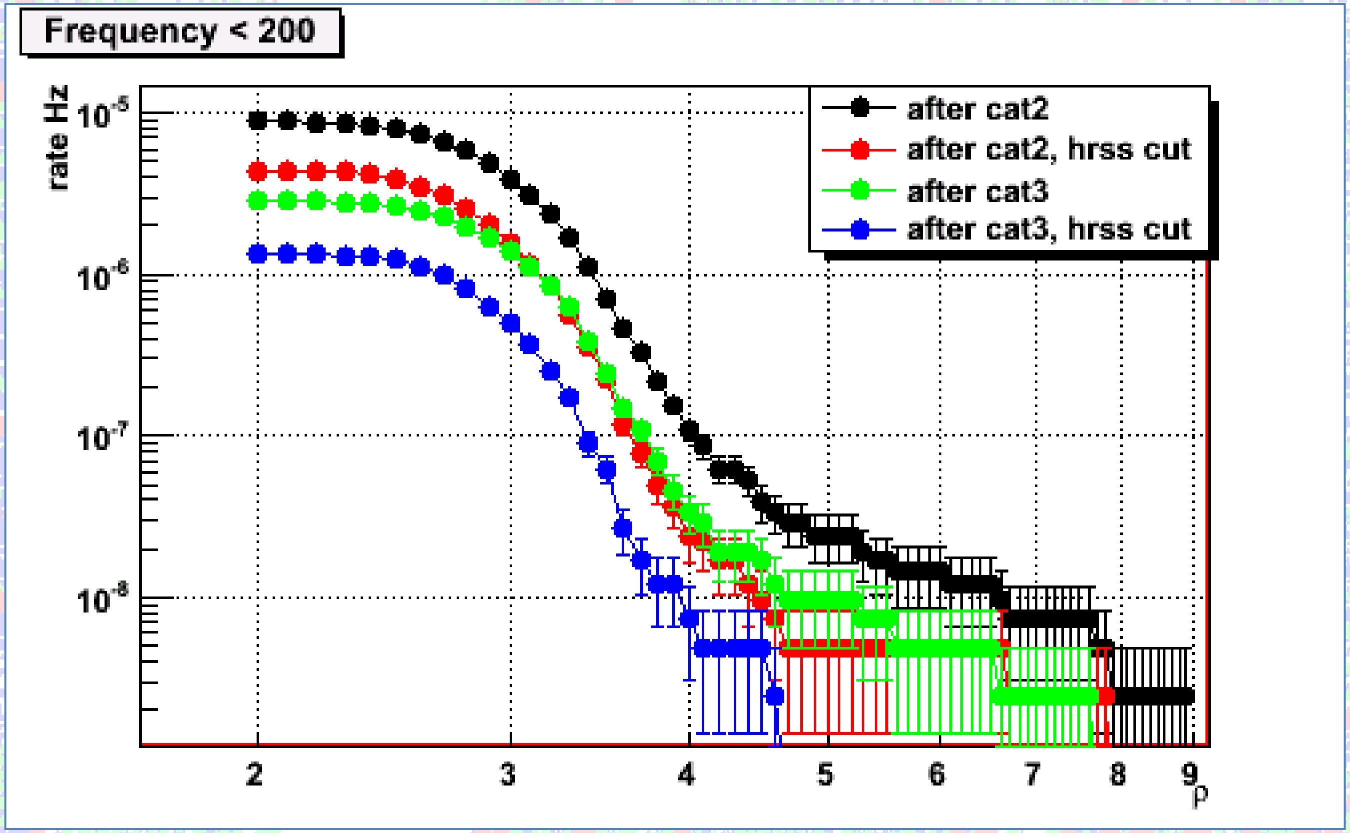
Why data analysis needs vetos

Several interferometers in coincidence (or in coherence):

Apply vetos in each detector's data set

→ reduce the distribution tail

→ improve statistical significance of remaining loud events



Some preliminary definitions

Some preliminary definitions

- **GPS time**

Atomic time counted in seconds from January 1st 1980 00:00:00 UTC
Example: 931220000 = July 10th 2009 00:13:05 UTC

- **Dark fringe / GW channel**

The channel which contains the GW signal (Pr_B1_Acp or h_16384Hz)

- **Auxiliary channel**

Any channel recorded which contains information about ITF environment or ITF control

- **SNR (Signal to Noise Ratio)**

Often defined by the power in the signal divided by the power in the noise background

- **Trigger**

Event in the analyzed signal that fulfill the criteria of the data analysis pipeline

- **Glitches**

Transient events in the analyzed signal whose origin is some noise source in the ITF

- **Veto**

Conditions applied to reject data or data analysis triggers

Some preliminary definitions

- **The environment channels**
  - **AC* : acoustic channels**
  - **SE* : seismic channels**
  - **MA* : magnetometers channels**

- **The ITF channels**
  - **Gx* : beam monitoring channels**
  - **Bs* : injection system channels (Beam source)**
  - **Sc_IB* : Suspension control of Injection Bench**
  - **Sc_NE* : Suspension control of North End mirror**

- **The detection photodiodes channels**
  - **Pr_B1* : Photodiodes signals for the B1 beam (output port after output MC)**
  - **Pr_B5* : Photodiodes signals for the B1 beam (output port at BS reflective face)**
  - **Pr_B7* : Photodiodes signals for the B1 beam (output port at North End)**
  - **Pr_B8* : Photodiodes signals for the B1 beam (output port at West End)**

- **Veto**

Conditions applied to reject data or data analysis triggers

Some preliminary definitions

• **GPS time**
Atomic time counted in seconds from January 1st 1980 00:00:00 UTC
Example: 93

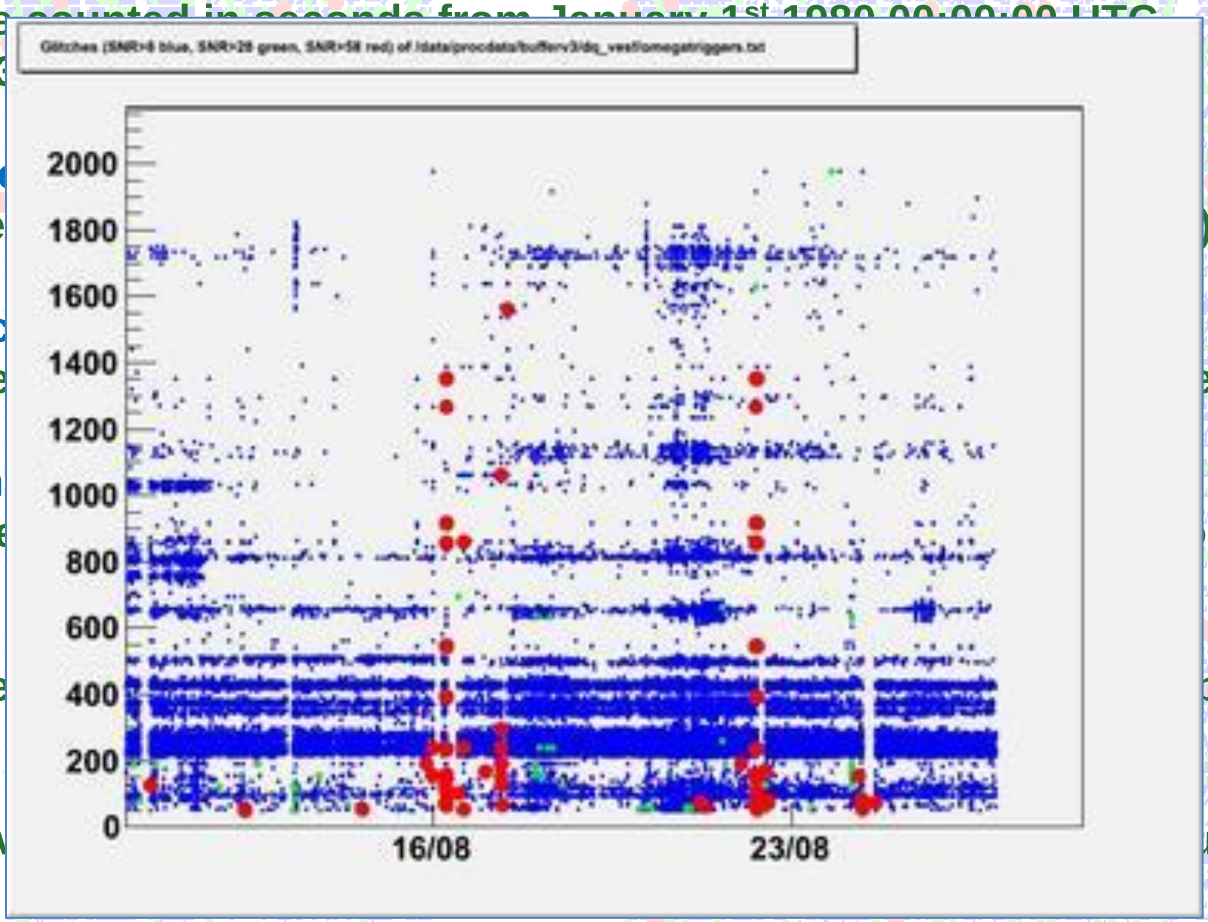
• **Dark fring**
The channe **)**

• **Auxiliary c**
Any channe **ent or ITF control**

• **SNR (Sign**
Often define **ise background**

• **Trigger**
Event in the **pipeline**

• **Glitches**
Transient ev **urce in the ITF**



• **Veto**
Conditions applied to reject data or data analysis triggers

Some preliminary definitions

- **Science mode segments**

Time periods where we consider the detector in good state for science data taking.

- **DQ quality flag segment**

Time periods where we consider that detector in science mode may be affected by some transient noise source

- **Padding**

Time extensions to be applied before and after each DQ flag segment.

- **Dead-time**

Percentage of science mode time vetoed because of a DQ flag

- **Efficiency**

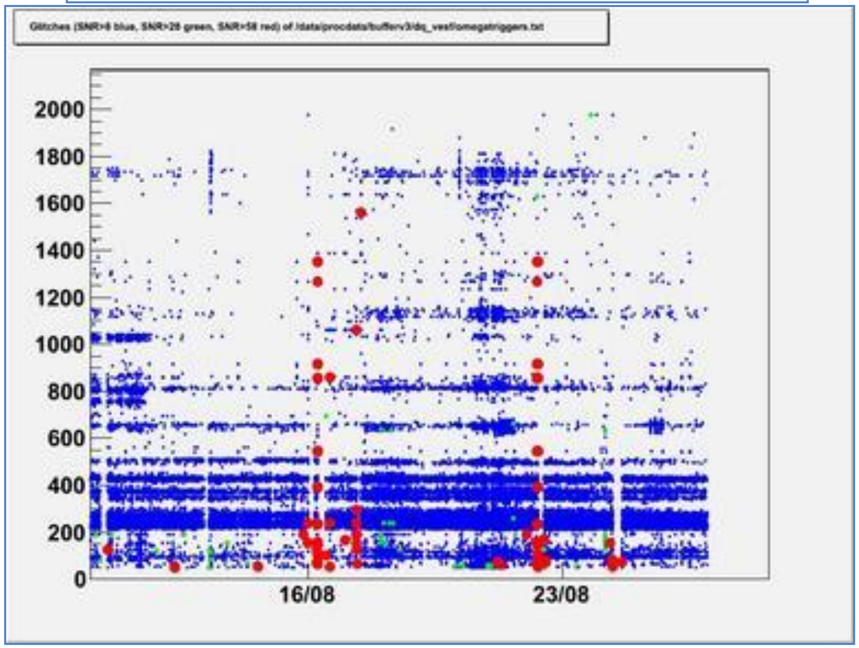
Percentage of data analysis triggers vetoed by a DQ flag

- **Use percentage**

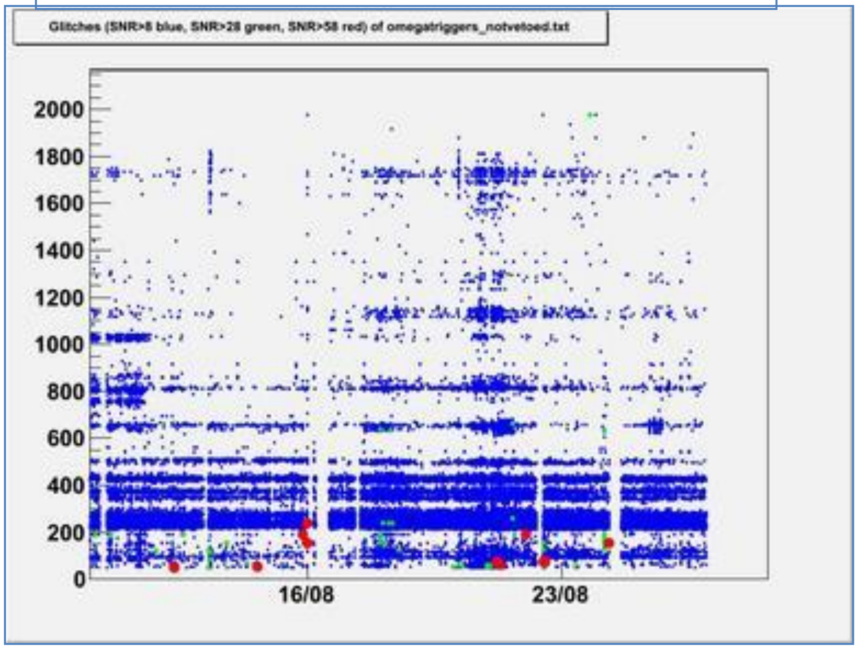
Percentage of veto segments which covered at least one data analysis trigger

Some preliminary definitions

All Omega triggers with SNR>8



Omega triggers with SNR>8
within science mode segments



Use percentage

Percentage of veto segments which covered at least one data analysis trigger

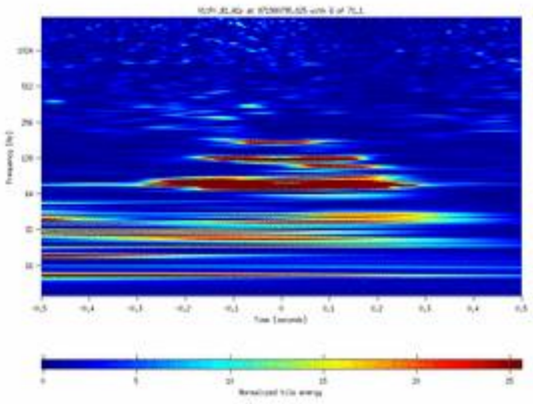
Some useful tools
for visualization and investigation

Some useful tools: wscan

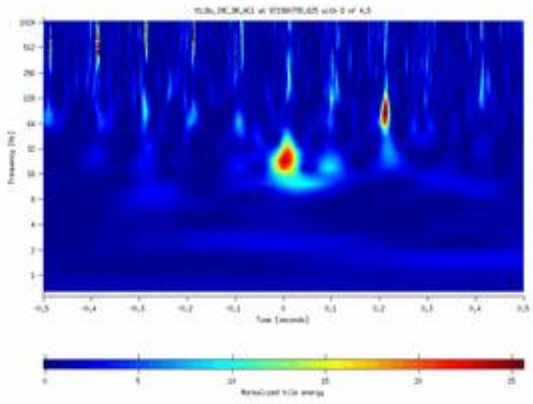
The wscans: for each loud Omega trigger, look for coincident glitches in a predefined list of auxiliary channels and show associated time-frequency plots

See https://wwwcascina.virgo.infn.it/DataAnalysis/Burst/wscan/V1/2007/08/

Pr_B1_ACp

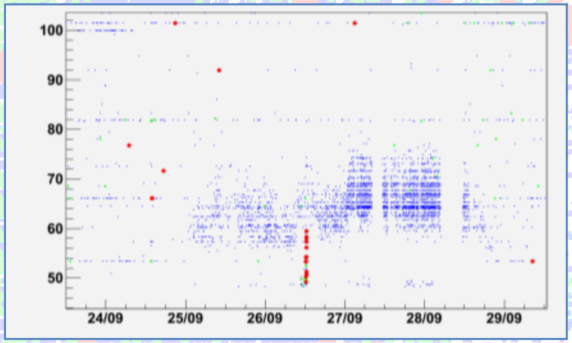


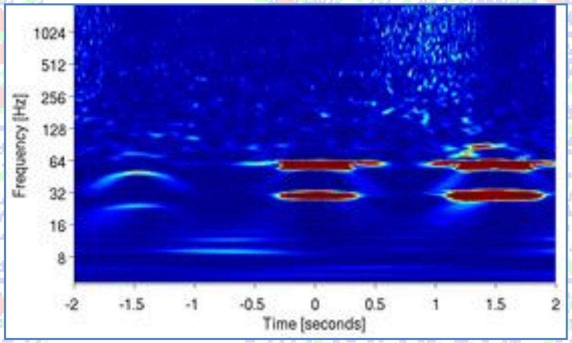
Bs_IMC_DR_AC1

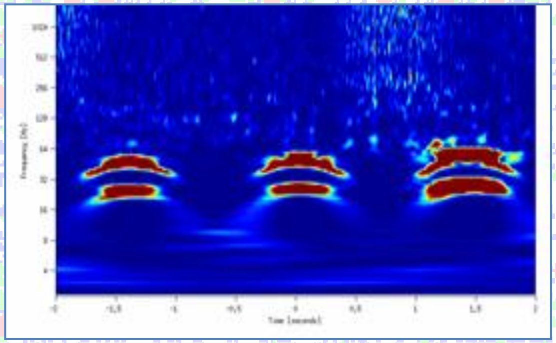


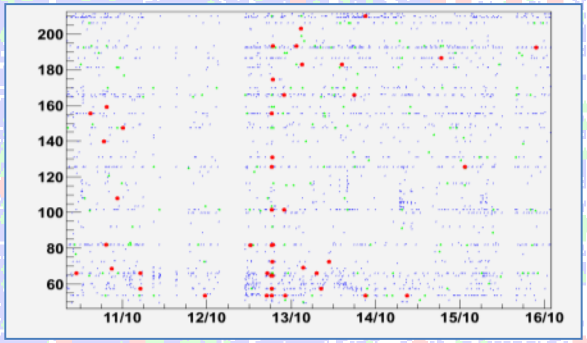
Some useful tools: wscan

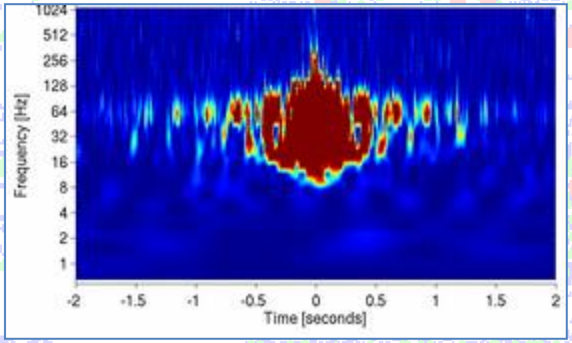
The wscans gives also visual indication about glitch families

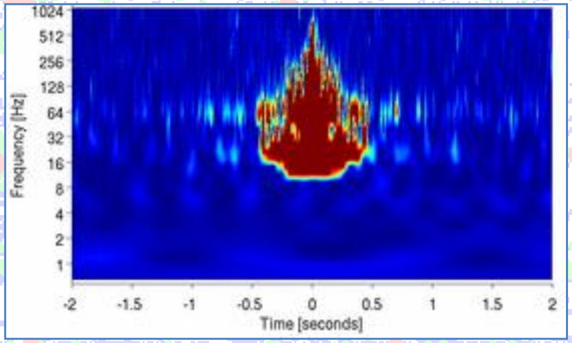








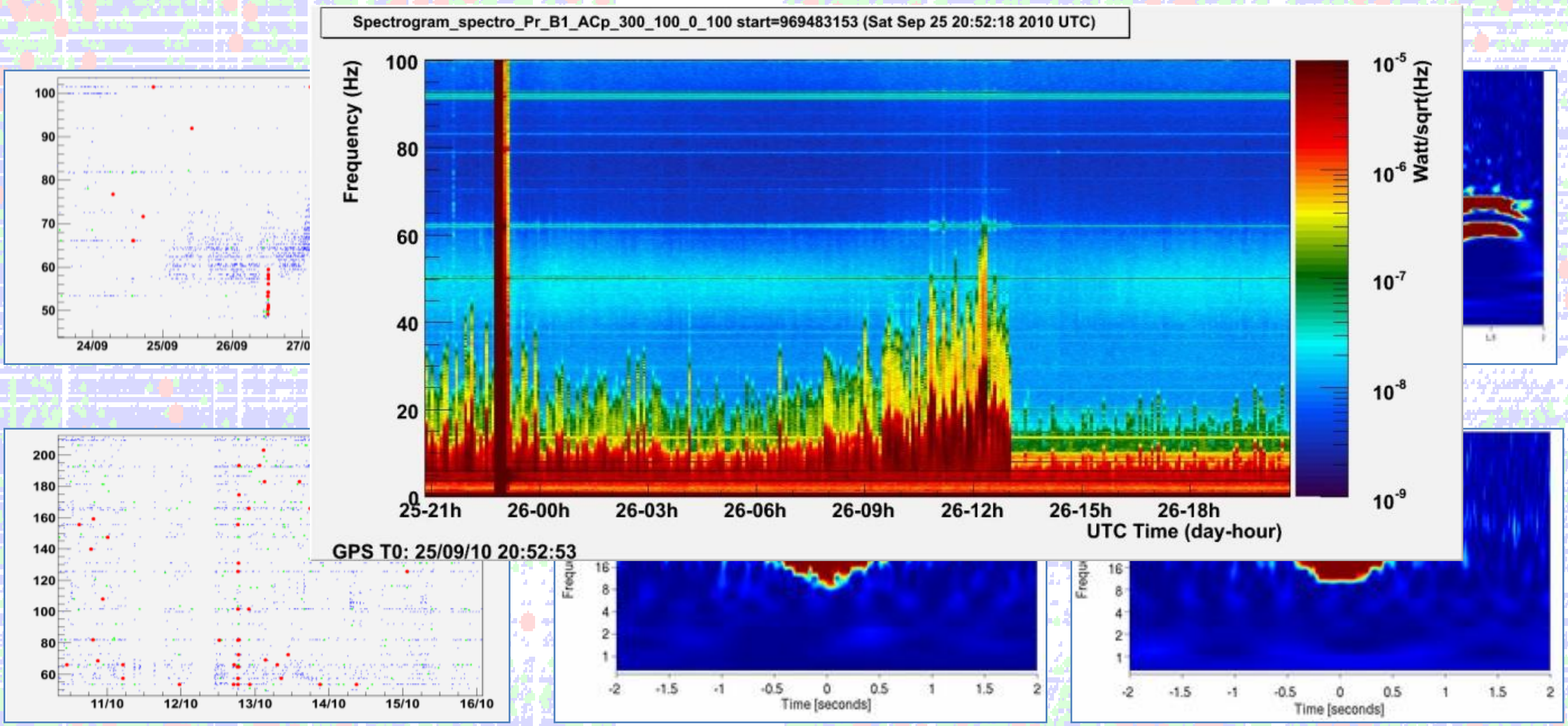




Some useful tools: spectrograms

Daily spectrograms gives also indications on the origin of some glitches

See http://wwwcascina.virgo.infn.it/MonitoringWeb/Spectro

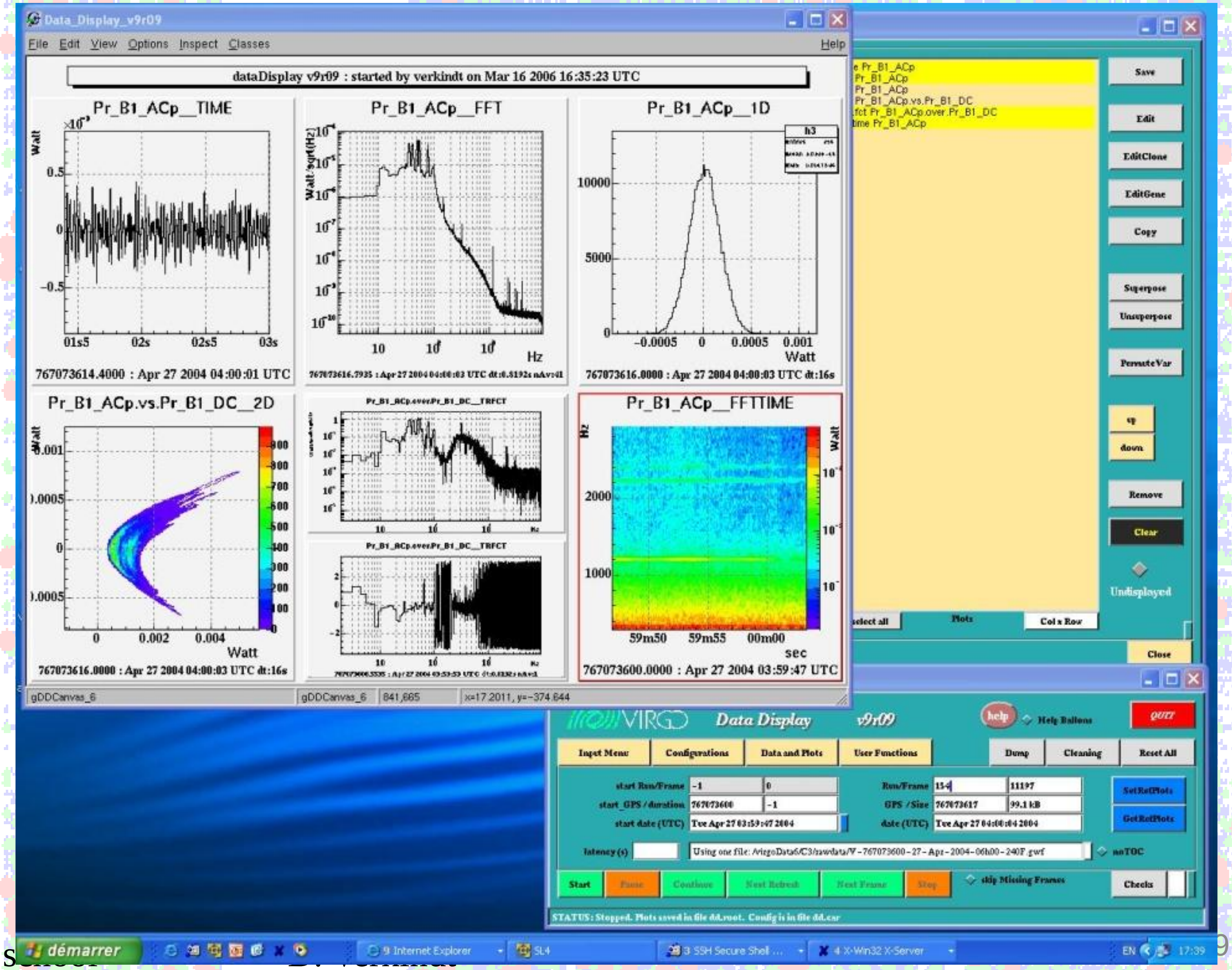


Some useful tools: dataDisplay

dataDisplay allows you to visualize the frame formated data (vesf_rds.ffl, trend.ffl or brmsmon.ffl)

And to see various channels in time or frequency domain

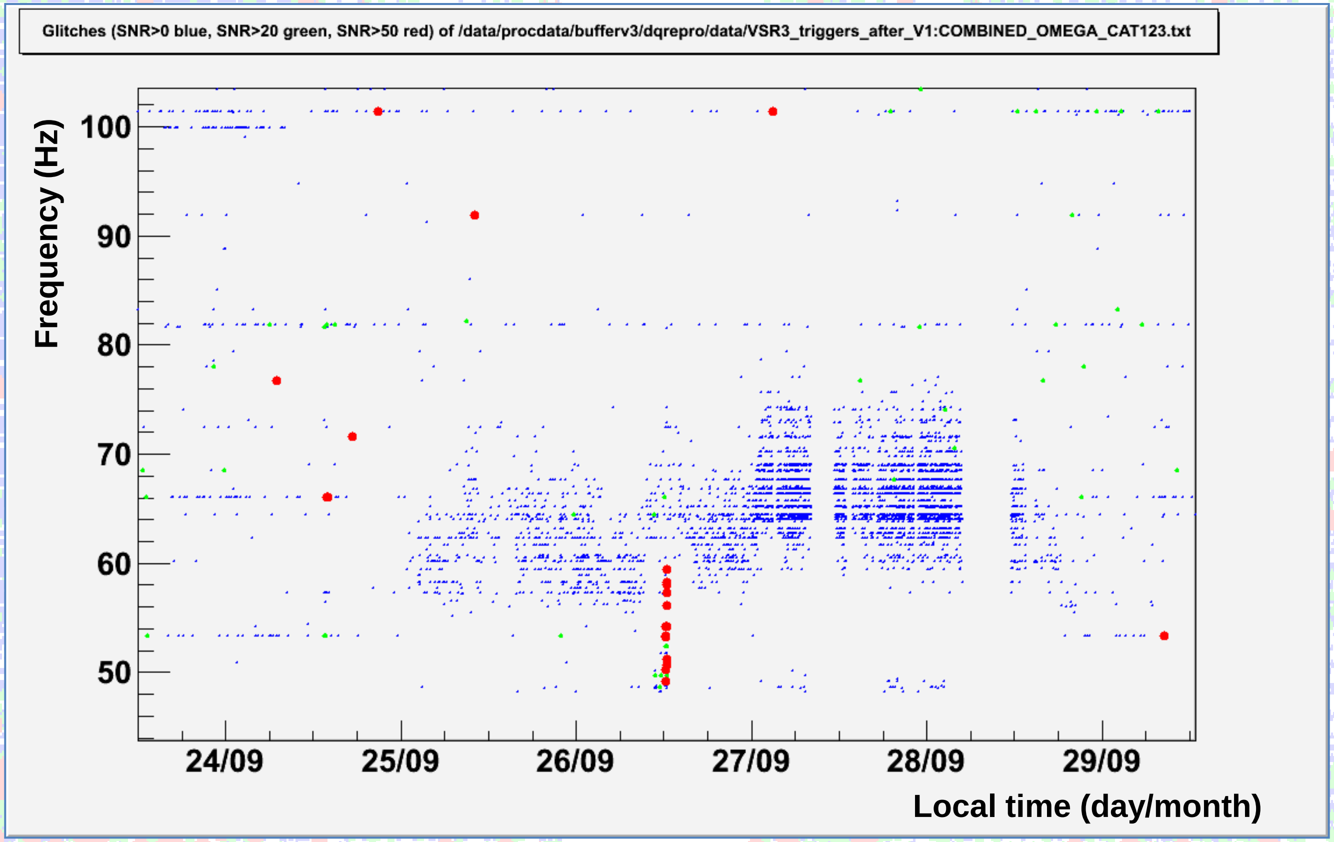
Type dataDisplay



Examples of glitches

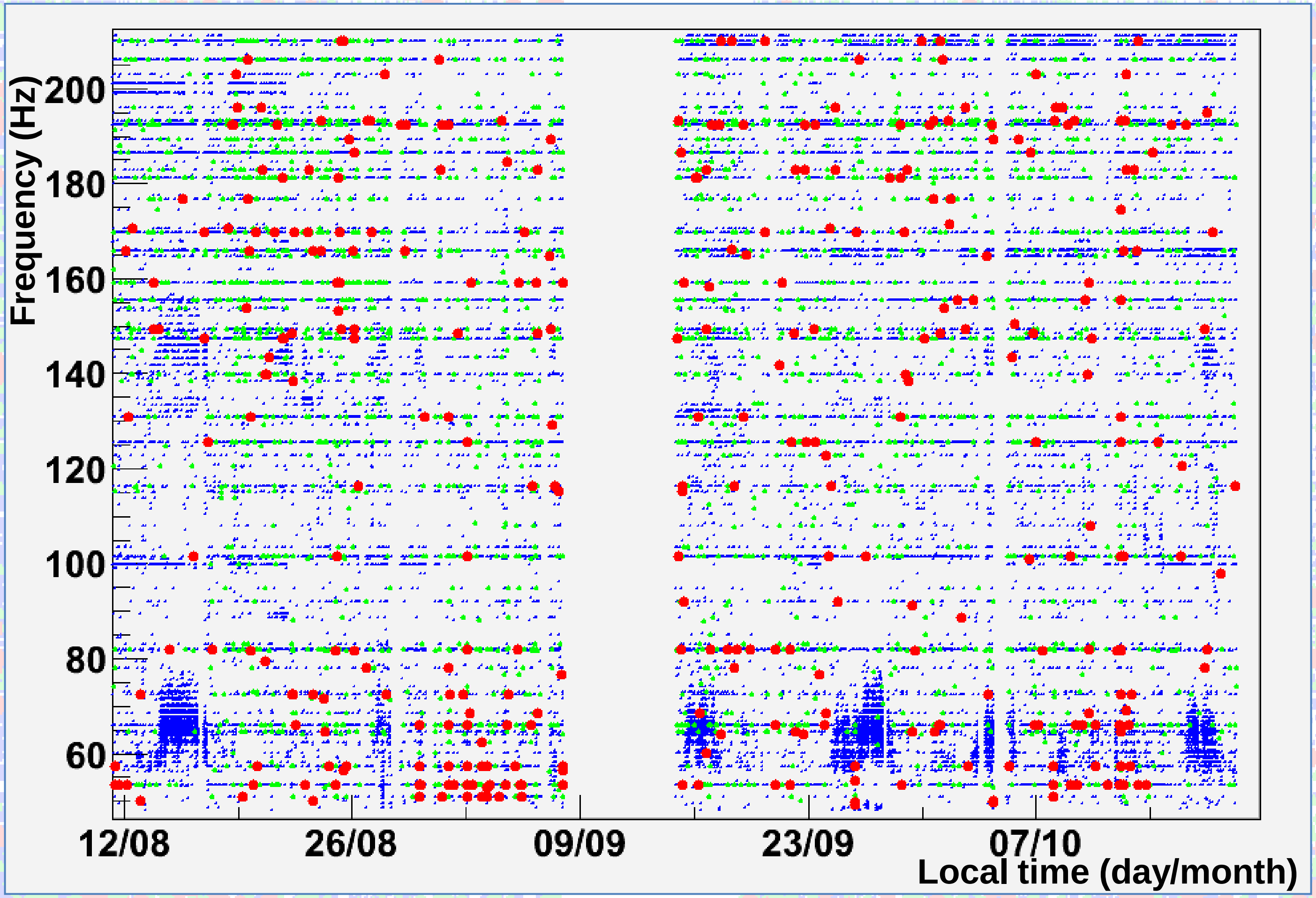
Example of glitches

- Some loud glitches periods well localized in time
- Many low SNR glitches associated to non-stationary frequency lines



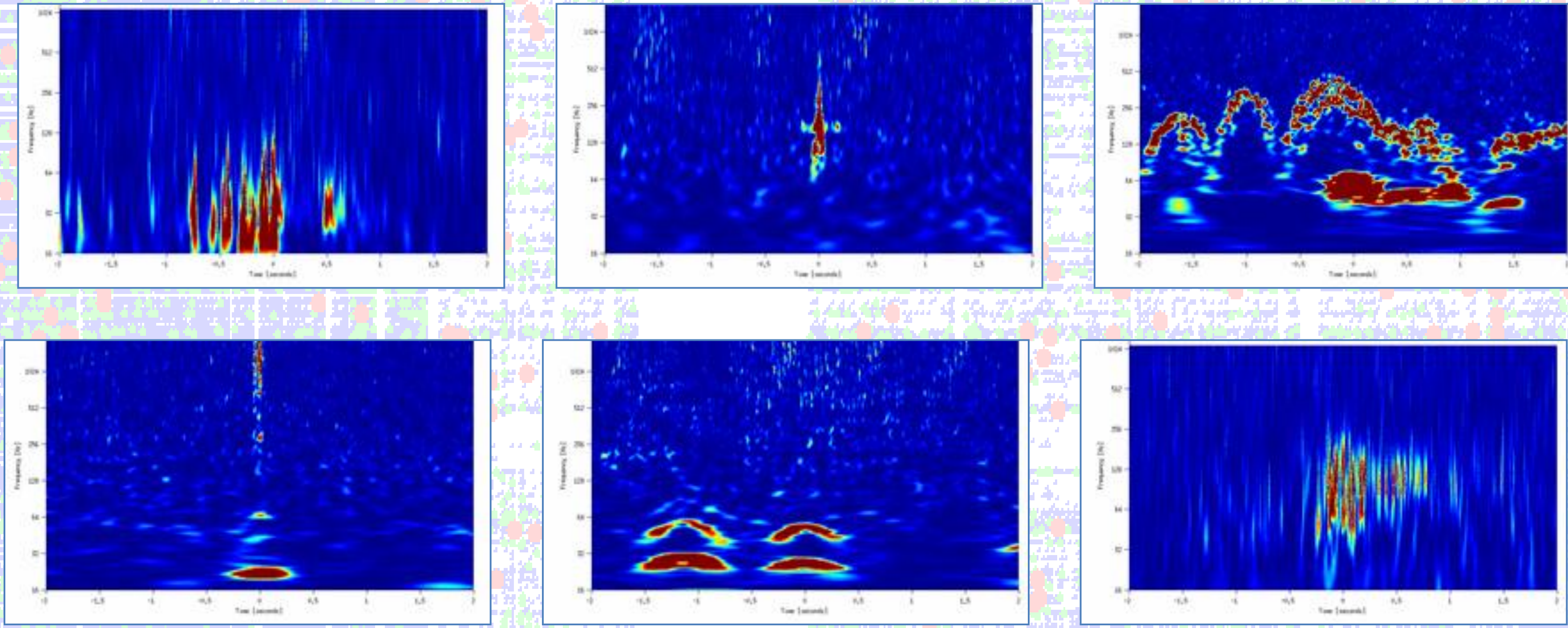
Example of glitches

And more various types of glitches…



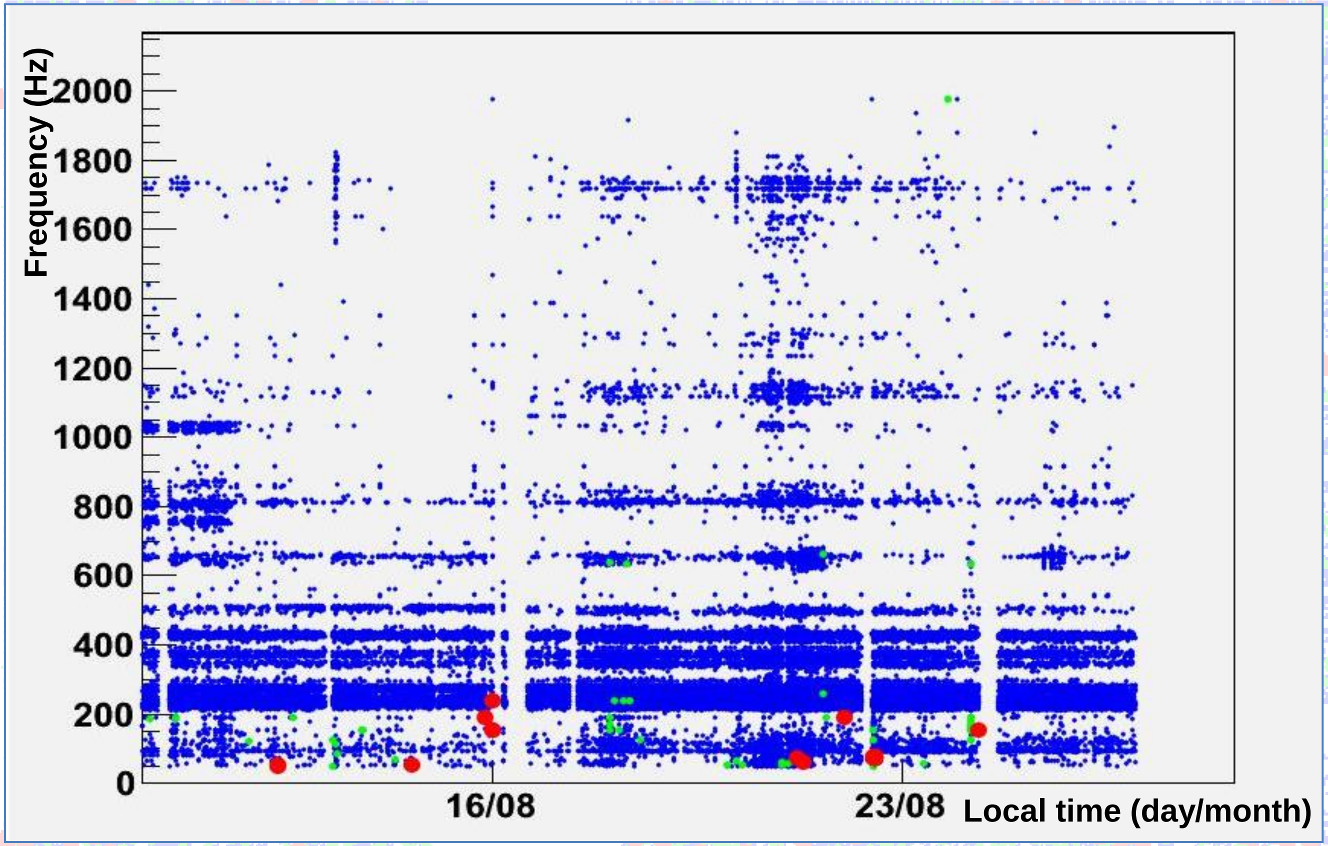
Example of glitches

Some glitches vetoed by seismic DQ flags



Example of glitches

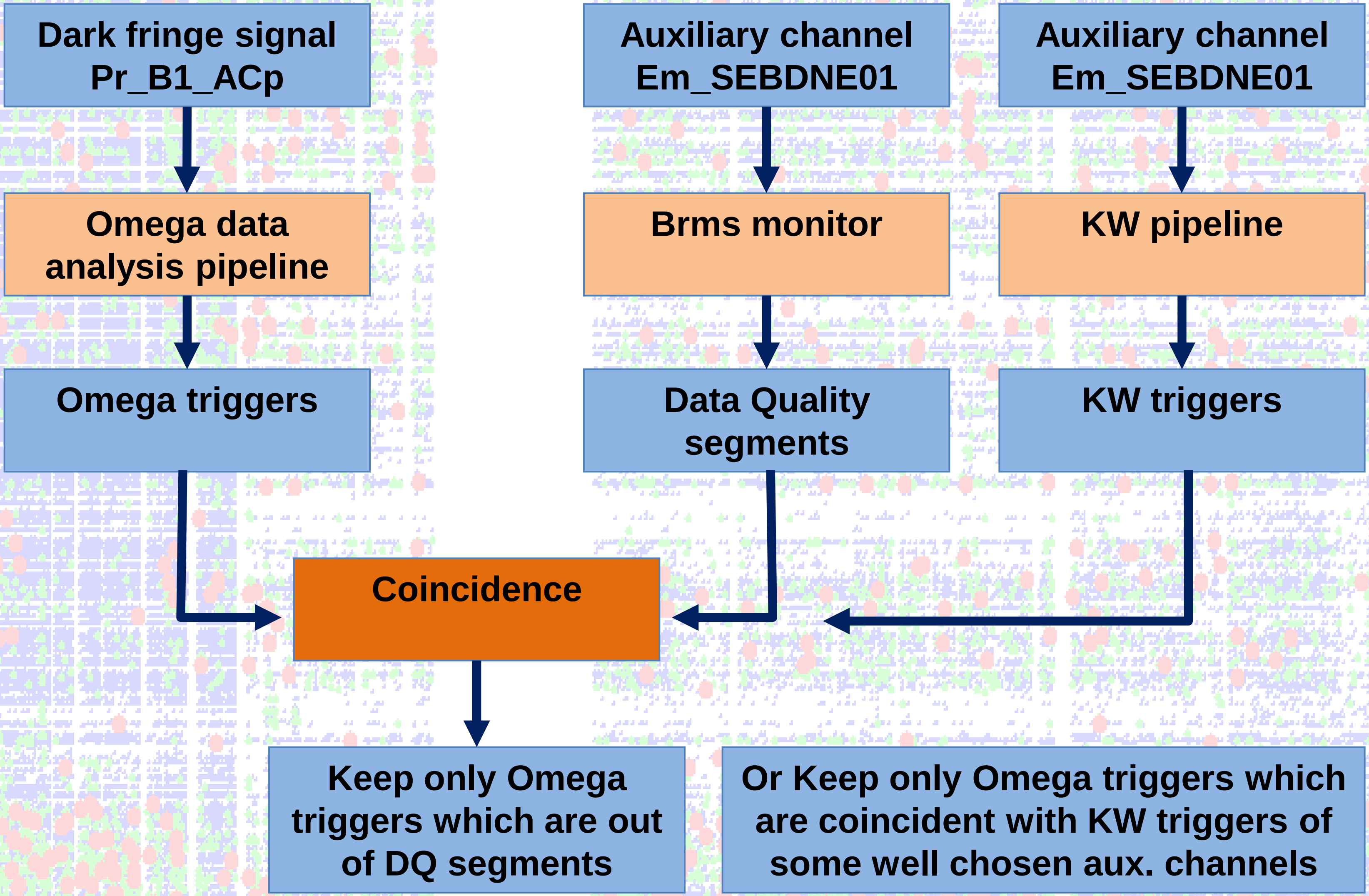
Some glitches on which you will work



DQ flag production

How is made a veto: general scheme

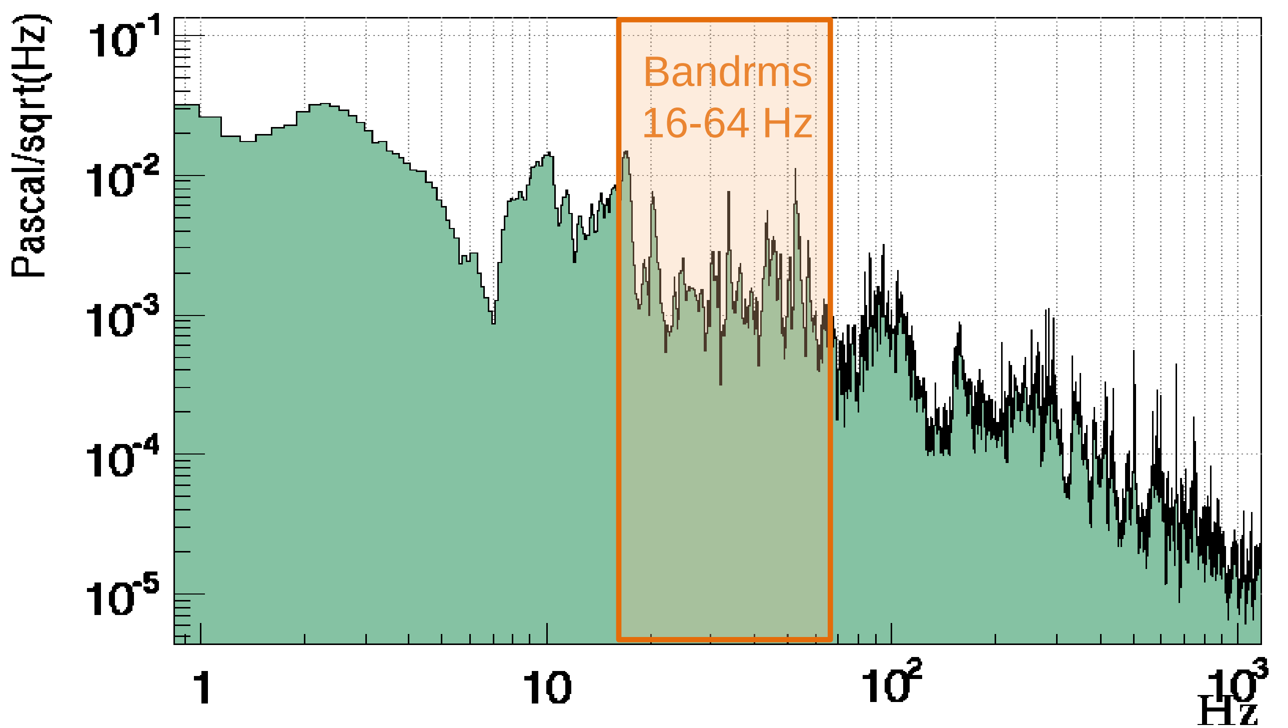
Two examples: Brms monitor and KW pipeline



A simple data analysis tool: Frequency band rms

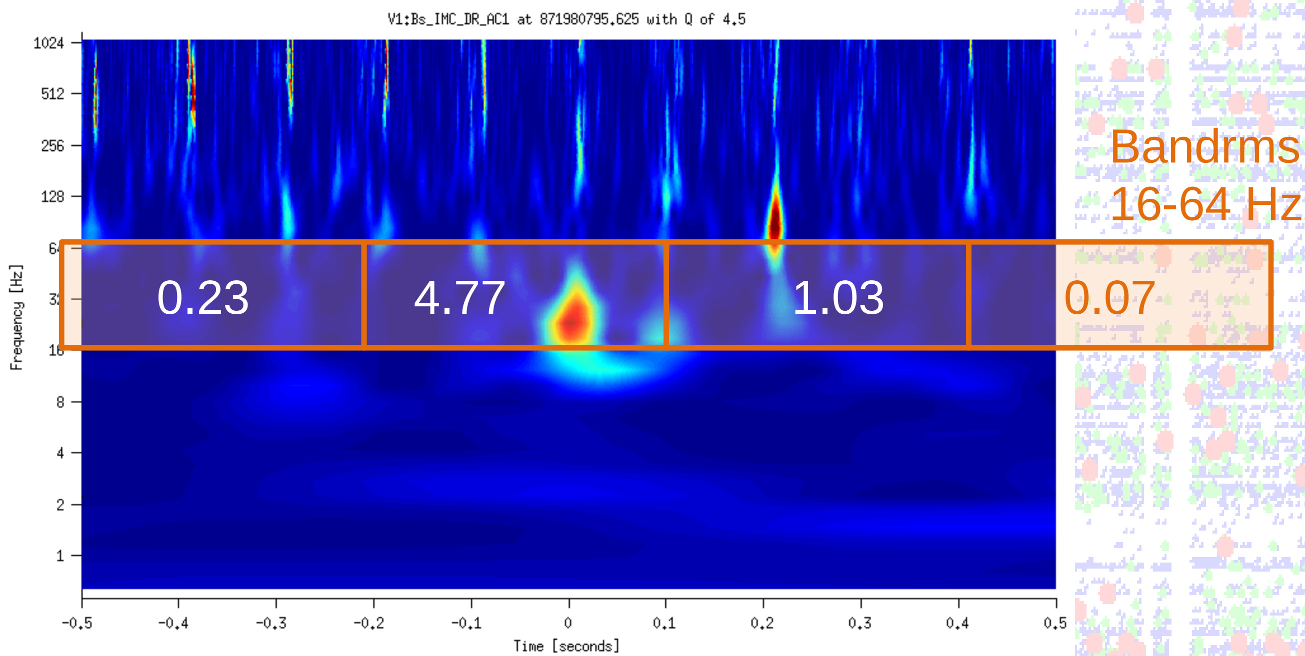
Frequency band RMS

- Compute FFT of a signal
- Compute bandrms = FFT^2 integrated from fmin to fmax
- Compare bandrms to a threshold (fixed or adaptative)



Brmsmon algorithm

- Compute FFT of a signal
- Compute bandrms = FFT^2 integrated from fmin to fmax
- Compare bandrms to a threshold (fixed or adaptative)
- Do this in a sliding time window



How to develop a DQ flag

- **Using the various tools previously presented, you choose some auxiliary channels and frequency bands useful for veto**

- **Now, you need to setup the parameters (frequency band and threshold) used by the brmsmon algorithm for each channel**

- **And to define what is the best way to use the results of each channel to define a DQ flag**

brmsmon algorithm

Keyword	input file	start GPS time	duration
FDIN_FILE	/virgoData/ffl/vesf_rds.ffl	931220000	540000

Keyword	list of input channels needed
FDIN_TAG	"Em_SEBDNE* Alp_Main_LOCK_STEP_STATUS Qc_Moni_ScienceMode“

Keyword	output files	nbr of frames per file	channels to write in output file
FDOUT_FILE	data/V-BRSMonRepro	10000	"V1:BRMSMon*"

Keyword	flag name	number of signals needed over th
DQ_NAME	SE_NE_16_256	2

Keyword	channel name	FFTlength	fmin	fmax	th	adaptative th time scale
BRMS_CHANNEL	Em_SEBDNE01	1	16	256	8	1000
BRMS_CHANNEL	Em_SEBDNE02	1	16	256	8	1000
BRMS_CHANNEL	Em_SEBDNE03	1	16	256	8	1000

How to develop a DQ flag

- **You have produced the DQ flag values**
- **Now, you need to build the DQ segments (time periods where DQ flag is 1)**

segonline algorithm

Keyword	input file	start GPS time	duration
FDIN_FILE	brmsmon.ffl	931220000	540000

Keyword	list of input channels needed
FDIN_TAG	“BRMSMon*“

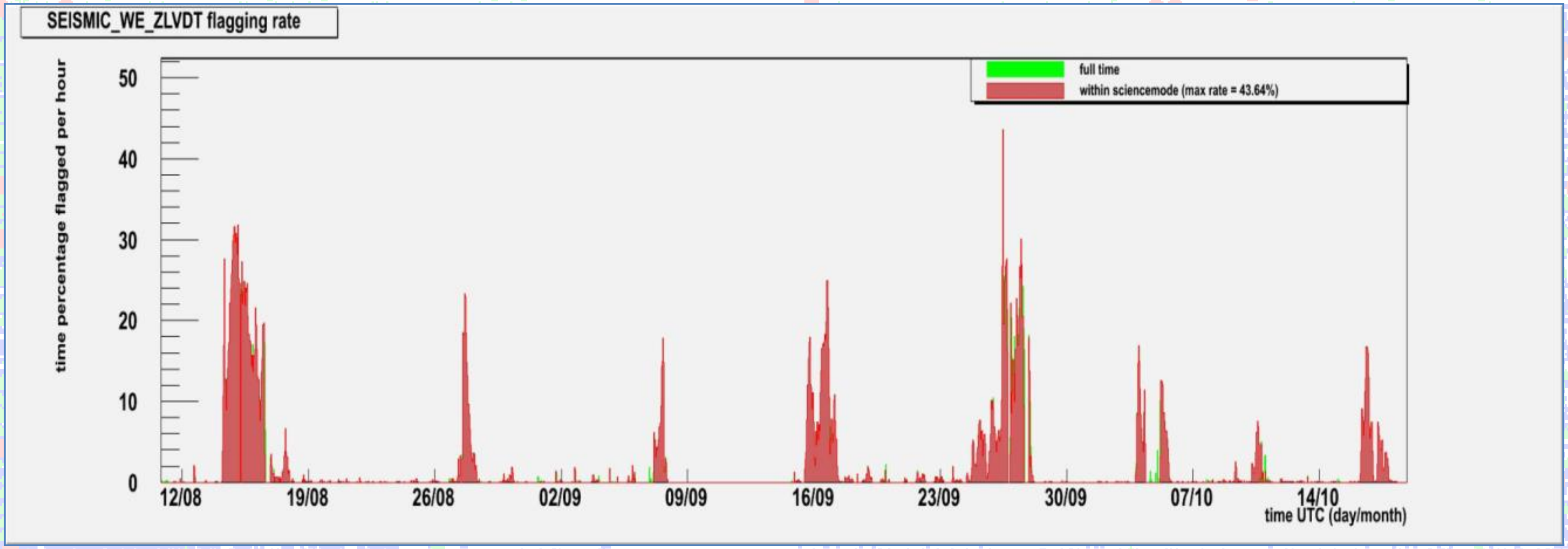
Keyword	output dir	period of segments file update
SEG_ASCII	/users/vesfuser	30

Keyword	flag name	segments file name	comment
SEG_FLAG	BRMSMon_SE_NE_16_256	V1:SE_NE_16_256	“seismic activity"

DQ flag performances

DQ flag performances

First information about a DQ flag: its flagging rate
It shows sometimes that quality flags are associated to well defined periods (for instance: seismic activity)



DQ flag performances

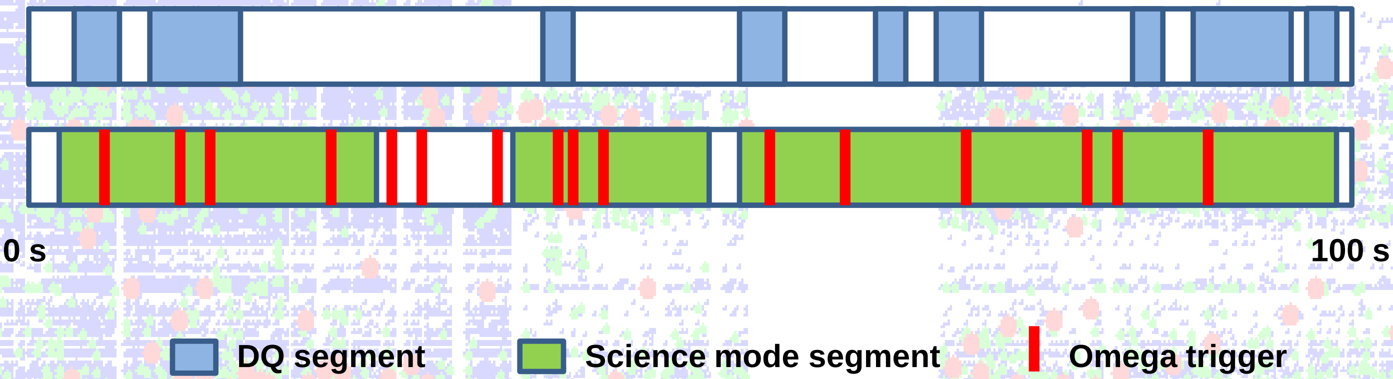
For each DQ flag produced, we then need to know some figure of merits:

Deadtime: Does it veto too much data?

Efficiency: Does it veto a good fraction of the glitches?

Use percentage: Does it uses more segments than what would be useful?

Eff/deadtime: Does it veto suitably or randomly?
(low use percentage and efficiency/deadtime ~ 1 means that it vetoes randomly)



Efficiency=8/16=50%
Use percentage=6/9=66%
Deadtime=20s/90s=22%
Eff./deadtime=50/22=2.3

DQ flag performances

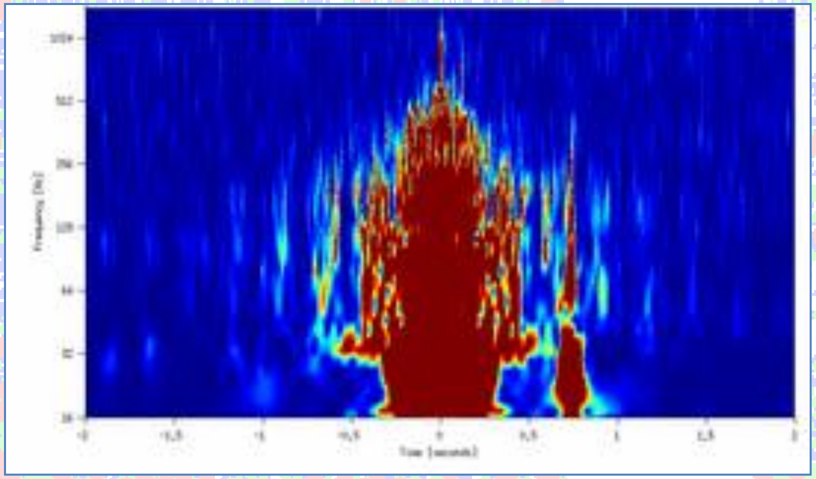
Padding of DQ segments:

- According to the type of data analysis triggers you want submit to veto
- Or according to the type of glitches you expect to veto

- You need to search for an optimal padding of the DQ flag segments

- The padding is an extension of each DQ segment.

- For instance padding [-3,+2] applied to segment 931200010 931200015
  produces segment 931200007 931200017



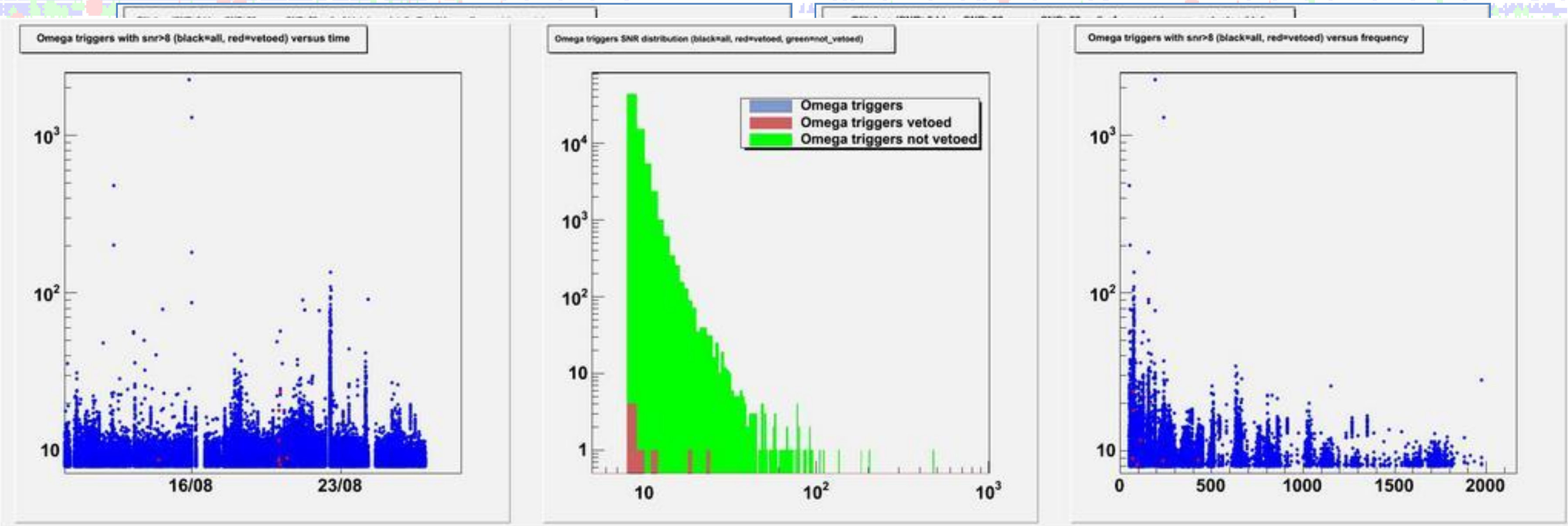
DQ flag performances

One DQ flag result:

- Several loud triggers removed
- Good use percentage
- Low deadtime

	Efficiency	Use Percentage	Eff/deadtime
SNR > 5	45/2736874=0.00164421%	11/12=91.6667%	1.02367
SNR > 8	8/68389 =0.0116978%	1/12=8.33333%	7.28293
SNR > 15	2/1058 =0.189036%	0/12=0%	117.692
SNR > 30	0/124 =0%	0/12=0%	0

Deadtime: 21/1307440 = 0.00160619%



DQ flag performances

Combine DQ flags:

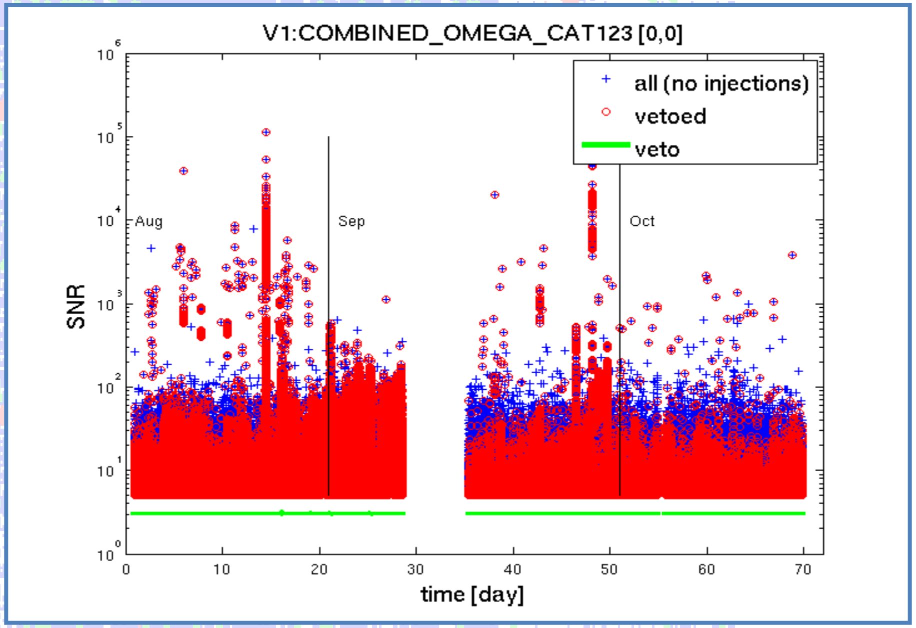
- **You can then choose the DQ flags of interest for your data analysis triggers and apply a OR of them.**
- **Resulting deadtime must not be too high (<10%)**
- **Resulting efficiency for loud triggers (SNR>15) must be above 50%**

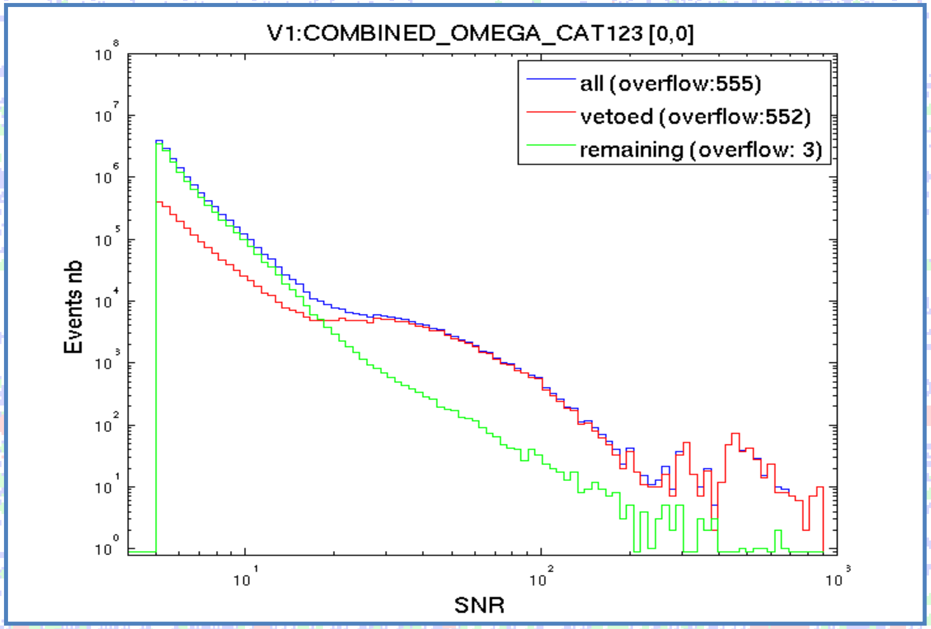
VSR3 DQ flags performances

Example of combined DQ flags performances

	Efficiency	Use percentage	Efficiency/deadtime
SNR> 5	13.793 %	91.55 %	1.56
SNR> 8	27.163 %	40.22 %	3.08
SNR> 15	70.745 %	10.93 %	8.01
SNR> 30	92.720 %	4.64 %	10.50

Deadtime: 8.833 %





Last words

- **Other types of vetoes:**
  - **Event y event vetoes (like KW vetoes)**
  - **Signal based vetoes (like waveform consistency check)**
  - **Followup tests (check of detector status, check event's pattern, coherence with aux. channels, etc…)**

- **The veto developments are a long-term (never-ending) activity**

- **It needs strong updates each time the interferometer's configuration is changed.**

- **Vetos are not of marginal use even for coincidence/coherent analyses**