

# An introduction to XMM-Newton data analysis and the SAS grand-scheme

*Matteo Guainazzi*

*XMM-Newton Science Operations Centre*

*European Space Astronomy Centre*

*Villafranca del Castillo, Spain*

- Basic principles of X-ray astronomy
- SAS grand-scheme
- Some basic principles of X-ray spectral analysis

X-ray detectors are **photon-counting** → two main consequences:

- X-ray astronomy is an **intrinsic Poissonian science**
  - Scientific products can have a few or even zero events in large ranges of their parameter spaces
- The “king” in the X-ray realm is the **event**, characterised by:
  - **position** on the detector
  - “**pulse height**”, which is related to the X-ray **energy** of the incoming photon in a complex and generally non-linear way
  - arrival **time** at the spacecraft
  - event “**shape**” (used to separate X-ray events from particles)
  - **CCD number**, and other secondary attributes (you don’t generally have to worry about)

# Even lists



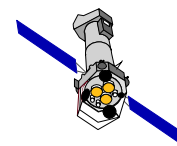
**When?**

**Where?**

**Who?**

**What?**

	<input type="checkbox"/> TIME D s	<input type="checkbox"/> X J 0.05 ARCSECONDS	<input type="checkbox"/> Y J 0.05 ARCSECONDS	<input type="checkbox"/> PHA I CHAN	<input type="checkbox"/> PI I CHAN	<input type="checkbox"/> PATTERN B	<input type="checkbox"/> CCDNR B
1	9.506202266412E+07	23743	21330	423	1447	2	1
2	9.506202266412E+07	28728	21990	25	98	0	1
3	9.506202527717E+07	28176	31623	25	97	0	1
4	9.506202527717E+07	29829	30841	327	1131	0	1
5	9.506202527717E+07	23686	19319	541	1854	0	1
6	9.506203046611E+07	25510	32711	1810	6171	0	1
7	9.506203566620E+07	29814	28823	102	360	0	1
8	9.506203826626E+07	26635	30601	2062	7028	0	1
9	9.506204346625E+07	26429	20314	443	1519	4	1
10	9.506204606629E+07	20691	28728	1608	5471	3	1
11	9.506204606629E+07	27989	29777	202	700	0	1
12	9.506204606629E+07	21937	25667	117	402	2	1
13	9.506204866632E+07	28132	32491	462	1589	0	1
14	9.506204866632E+07	27204	29741	904	3095	0	1
15	9.506205126638E+07	22124	20257	290	994	0	1
16	9.506205906643E+07	23193	18795	1398	4771	0	1
17	9.506206166646E+07	23224	19326	276	950	0	1
18	9.506206946653E+07	27755	28979	183	637	0	1
19	9.506207206939E+07	22533	29563	33	118	0	1



**XMM-Newton**

Matteo Guainazzi --  
SCI-SDX

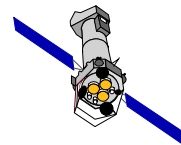
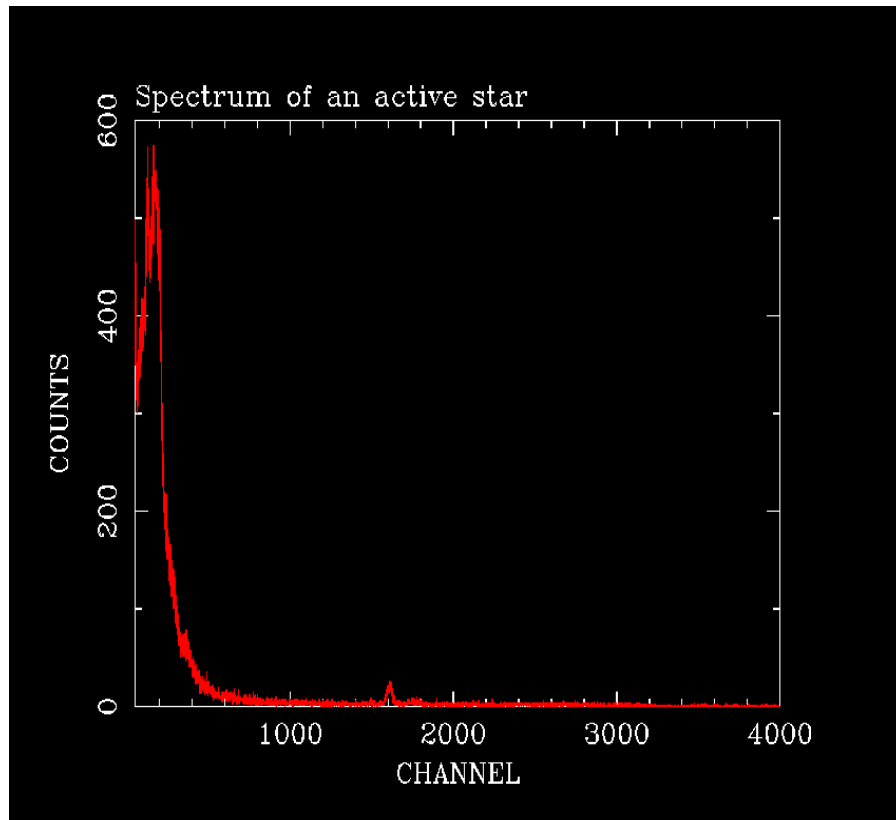
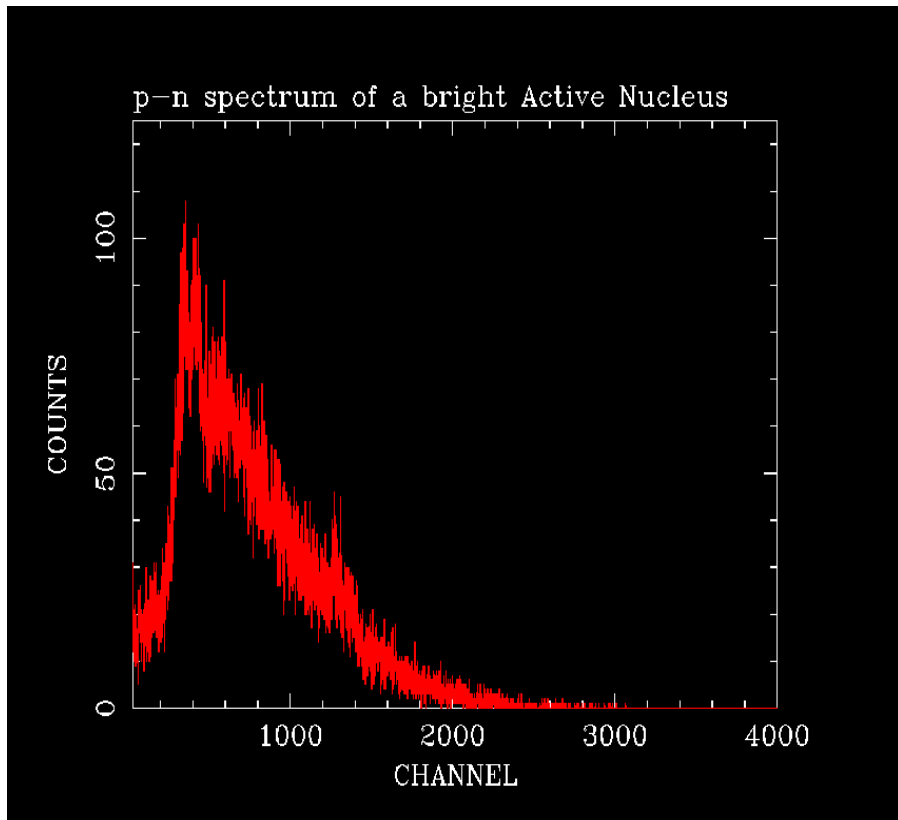
The X-ray scientific products can be seen as *projections* onto the sub-spaces defined by the event physical quantities

- By collapsing time and space, one gets an energy distribution function (*spectrum*) in units of *counts per energy bin*
- By collapsing time and energy, one gets a 2-D *image* in units of *counts per pixel*
- By collapsing space and energy, one gets an intensity *time series* in units of *counts per time bin*

These scientific products are expressed in units that are *indirectly* related to the intrinsic properties of celestial sources

# Transfer function

XMM-NEWTON

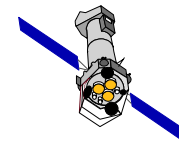
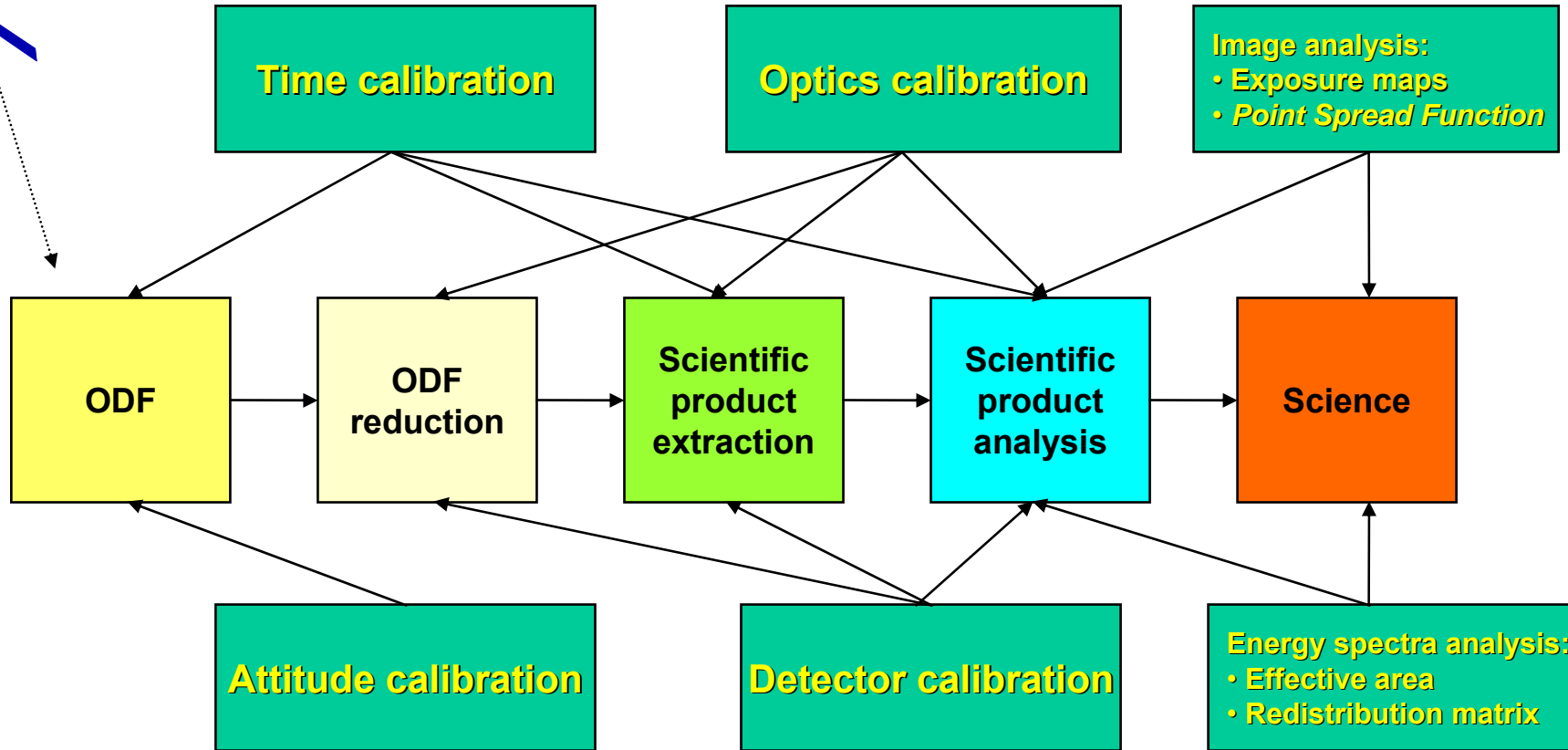
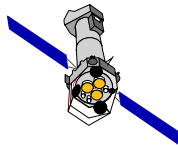


**XMM-Newton**

Matteo Guainazzi --  
SCI-SDX

# Removing transfer function = calibration

XMM-NEWTON



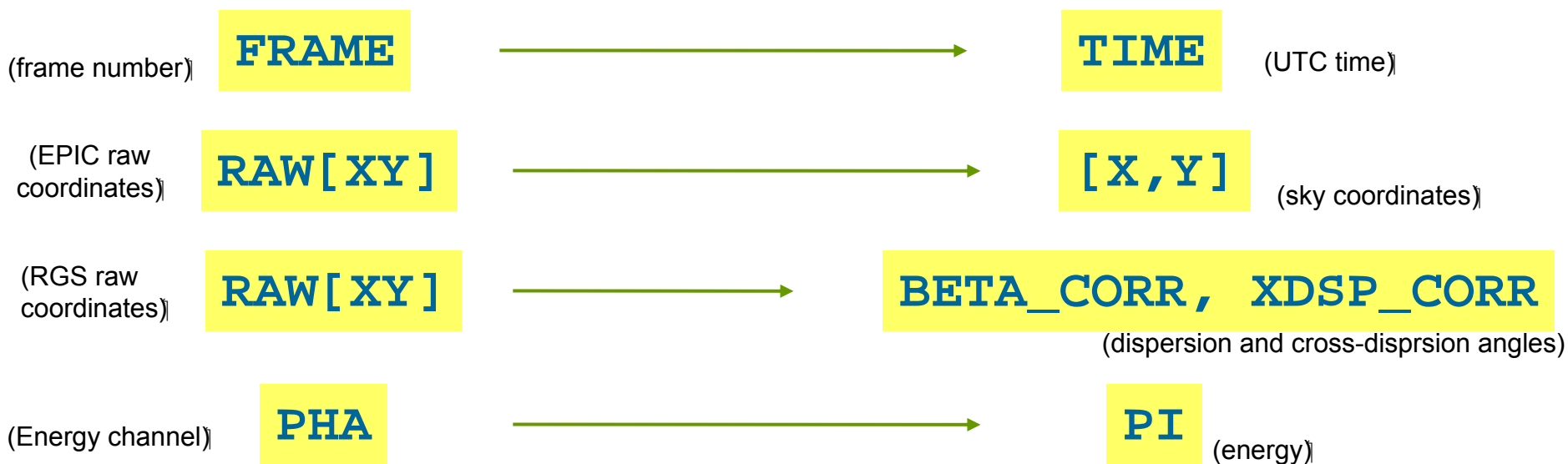
**XMM-Newton**

Matteo Guainazzi --  
SCI-SDX

No! Yet another software!  
Why do I need SAS?

## SAS does two things, that no other tool can do for you:

- applies **calibrations** to raw data

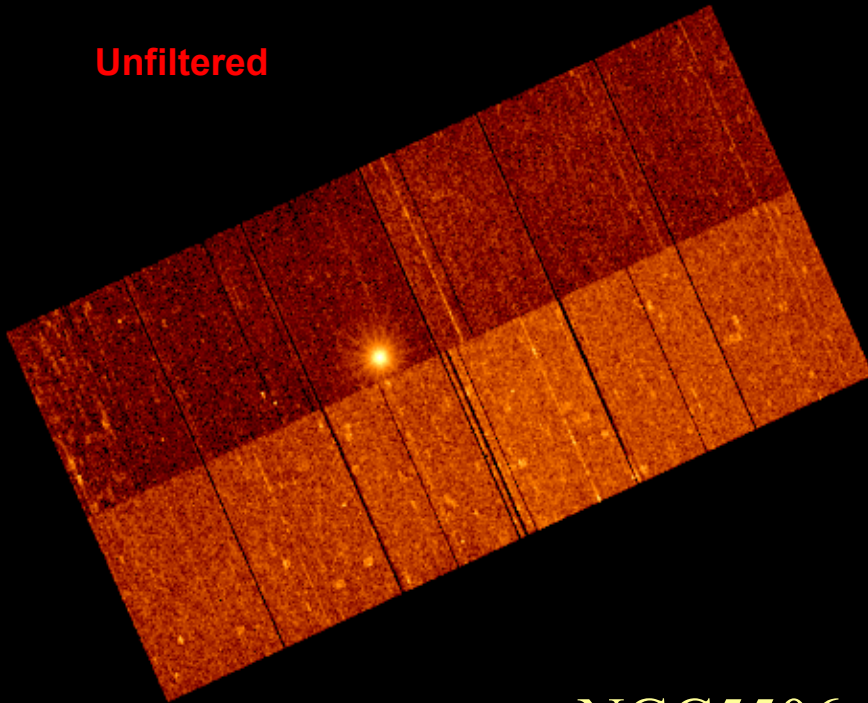


- optimally **screen** your data

# Importance of screening

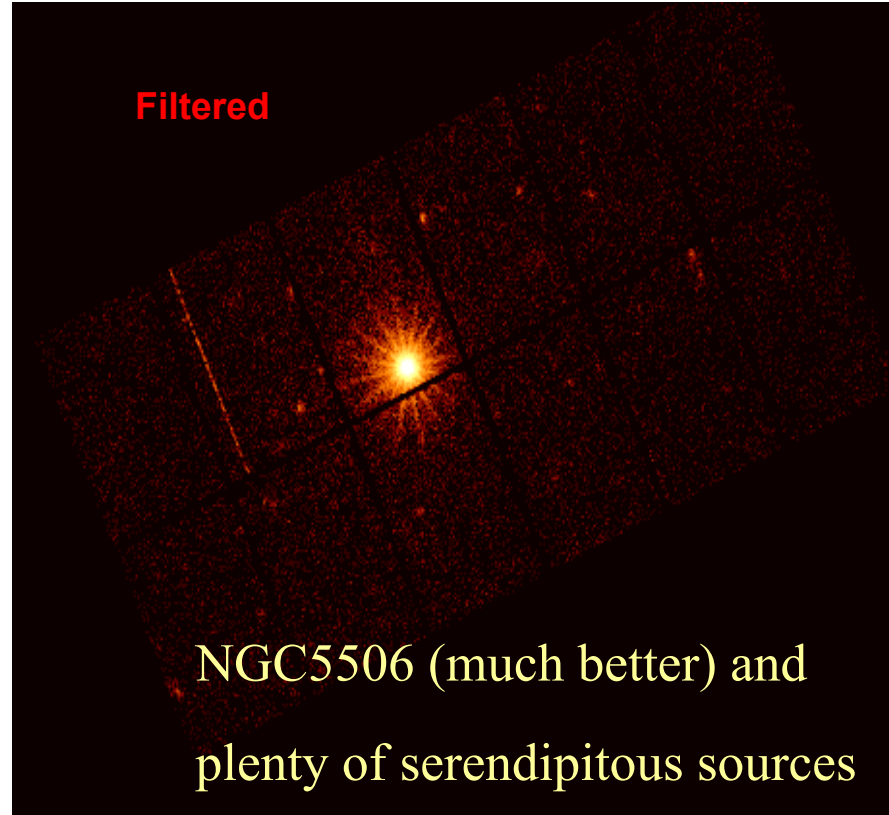


Unfiltered

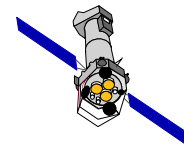


NGC5506

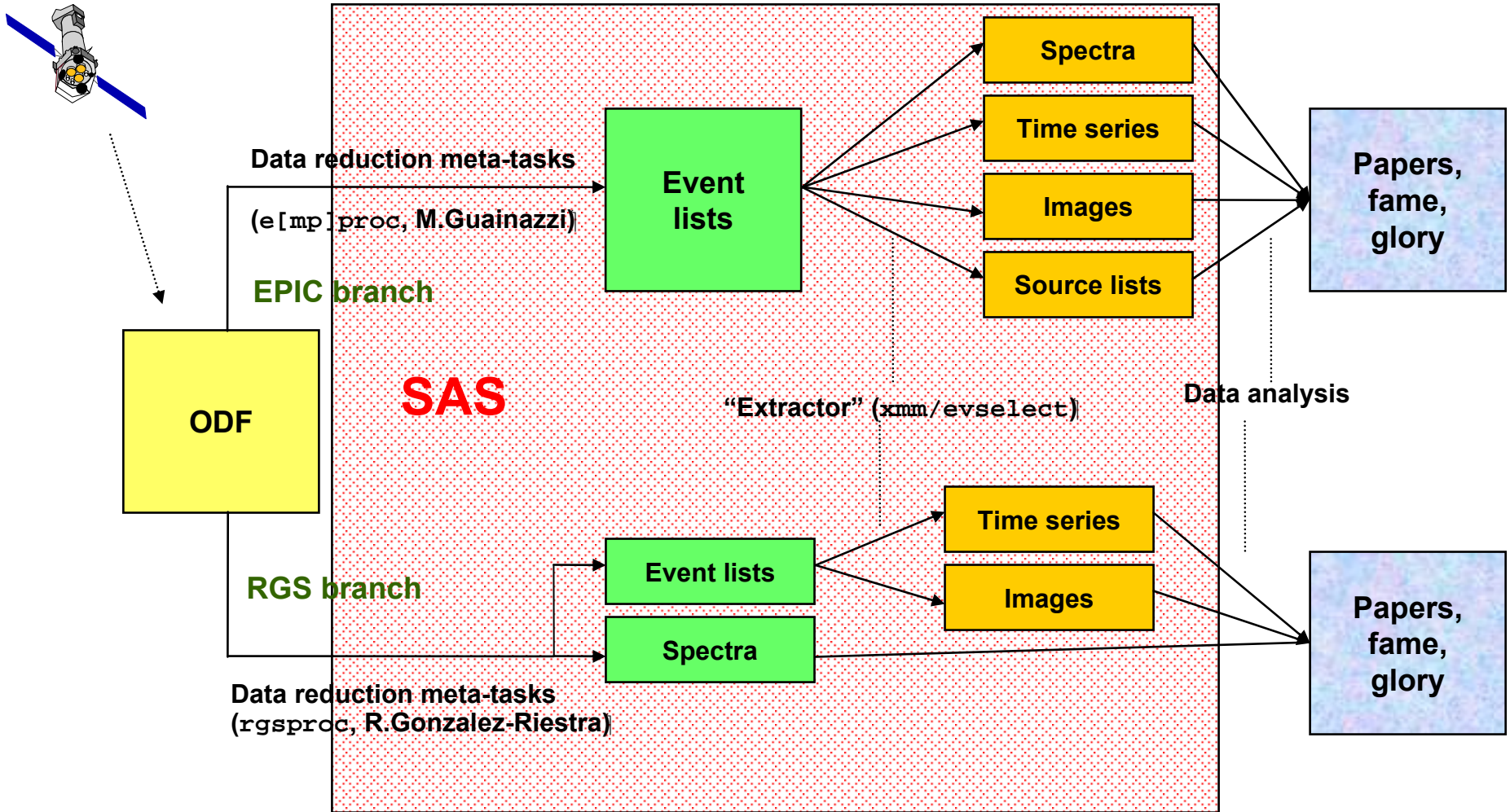
Filtered



NGC5506 (much better) and  
plenty of serendipitous sources



- You have to calculate (via SAS) and apply some more calibrations during the science analysis phase
- For **image analysis** (more in M.Ehle's talk):
  - **Exposure maps**: the mirror and detector sensitivity across the field-of-view, taking into account any changes in the spacecraft pointing direction
  - **Background maps**: smoothed maps of the field-of-view background (both of instrumental and cosmic origin)
- For **energy spectral analysis** (more in R.Saxton's and R.Gonzalez-Riestra's talks):
  - **Effective area**: transfer function of optics+detector as a function of energy and position
  - **Redistribution matrix**: probability that a photon of a given energy is registered in a given channel

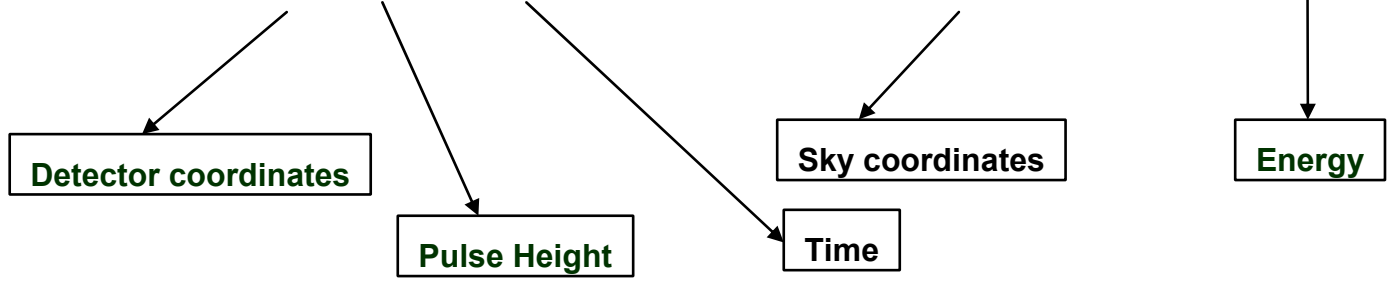


- **CCD-based event lists, containing uncalibrated data**
- **Auxiliary and Housekeeping files, p-n/RGS diagnostic images**
- **Spacecraft housekeeping** **FITS**
- **Spacecraft attitude showing the satellite star tracker pointing**
- **Time correlation file (onboard frame counter versus UTC)**
- **ODF summary file** **ASCII**

- **SAS reduction pipelines are actually already run for you by the **Science Survey Centre** in **Leicester (U.K.)****
- **The XMM-Newton Science Archive (XSA) contains:**
  - **ODF** data
  - **PPS** data, *i.e.* high-level, quick-look scientific products created by these pipelines:
    - **event lists**
    - detector and sky images
    - source lists and cross-correlation products
    - exposure maps
    - dispersion/cross-dispersion & dispersion/energy RGS images
    - RGS background-subtracted spectra for the brightest source in the field
    - attitude time series
    - background time series
  - EPIC serendipitous source catalogue (2XMM)
  - EPIC serendipitous Slew Source catalogue
  - XID (optical identification program) images



$$C(\underline{d}, \text{PHA}, t) = T(\underline{d}, t) \int dE \int dr R(\underline{d}, \text{PHA}, E, r, t) \times S(E, r, t)$$



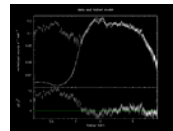
If R (as usually happens) is not diagonal the equation cannot be inverted. We need an alternative: *assuming* a physical model, *convolving* it with the response, and *comparing* the result with the observed counts, using an appropriate *statistical indicator of good fit quality*

**Forward-folding approach**

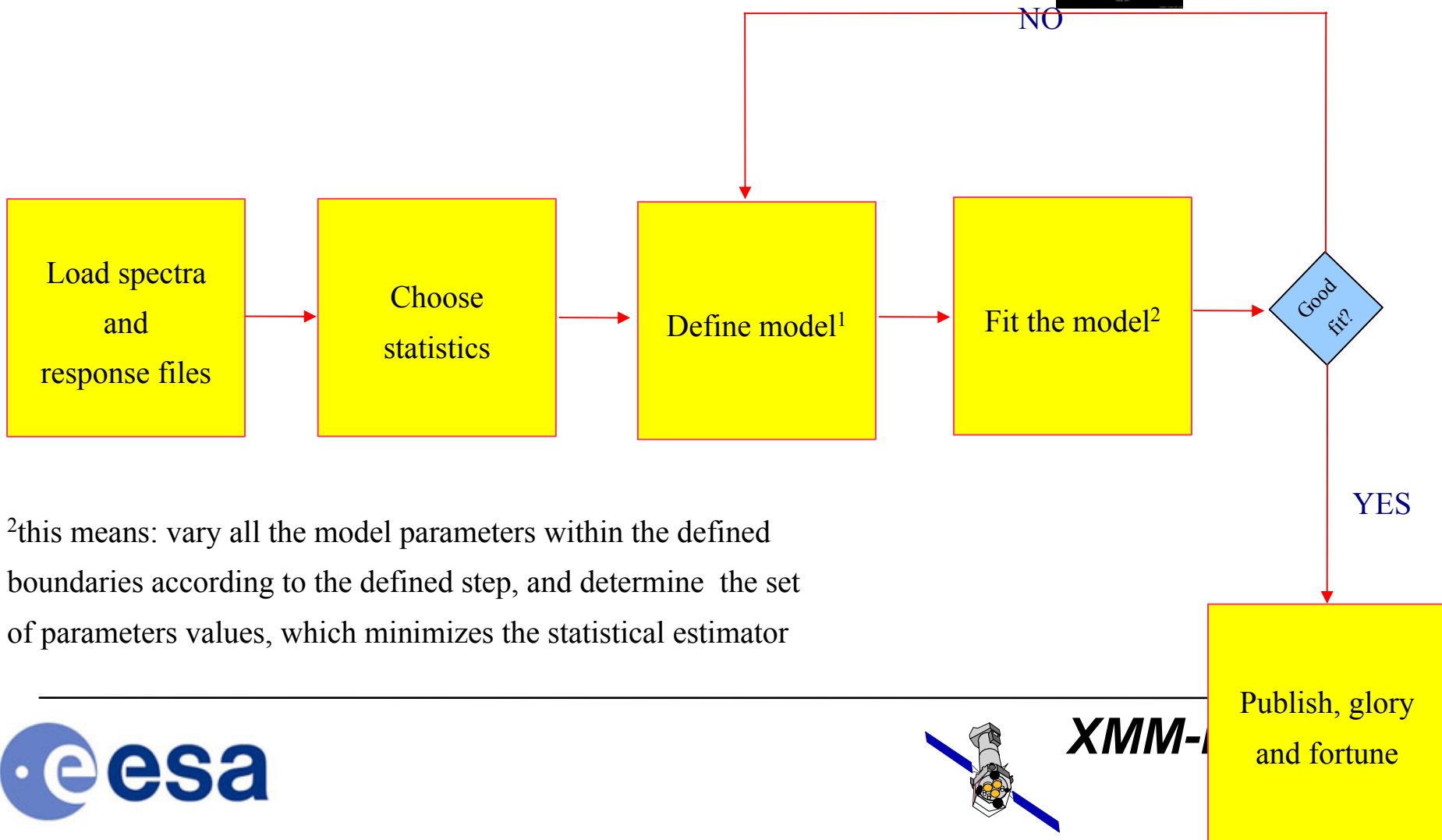


- **Actually **two** spectra need to be specified:**
  - Source+background spectrum
  - Background spectrum
  
- **Spectra programs behave in two ways with respect to the background:**
  - *either* subtract the background spectrum  $[B(E)]$  from the source+background spectrum  $[S(E)]$ , to get a **background-subtracted spectrum  $[C(E)]$** 
    - $C(E) = S(E)/t_{S(E)} - b_{S(E)}/b_{B(E)} \cdot B(E)/t_{B(E)}$
  - *or require independent models to be simultaneously fit to  $S(E)$  and  $B(E)$*
  - *we'll follow the former approach*

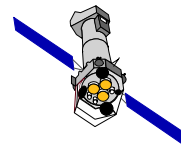
# How does it work in practice?



<sup>1</sup>including start values, upper/lower boundaries, variation steps for all parameters



<sup>2</sup>this means: vary all the model parameters within the defined boundaries according to the defined step, and determine the set of parameters values, which minimizes the statistical estimator





- The fit procedure requires a “goodness-of-fit” test, *i.e.* a statistical test, which allows us to tell “if” and “how well” a given model describes the data

- Two “goodness-of-fit” tests are commonly used

- $\chi^2 = \sum [C(E_i) - M(E_i)] / \sigma_i^2(E)$

$C(E)$  → Observed counts at energy  $E$

$\sigma(E)$  → Errors on the observed counts

$M(E)$  → Predicted source counts at energy  $E$

- Simple and well-known distribution:  $\chi^2/\nu \approx 1$  if the fit is good
    - Requires that the counts distribution in each spectral channel is Gaussian. *This may imply spectral rebinning potential loss of spectral resolution*
    - Requires that the estimate of the variance is uncorrelated with the counts

- C-statistics:  $C = 2 \sum [m(E_i) - S(E_i) \times \log(m(E_i)) + \log(S(E_i))]$  (Cash 1976)

- It can be used with low-counts spectra without rebinnig

$m(E)$  → Predicted source+background counts at energy  $E$

- It cannot be applied to background-subtracted spectra

$s(E)$  → Observed source+background counts at energy  $E$

- It does not provide us with an absolute estimate of the quality of the fit

SAS User's Guide

Problem reports

Threads

Watchout and evergreen tips

SAS task-by-task documentation

Calibration links

SAS workshops

XMM-Newton Science Analysis System - Mozilla

esa.es/external/xmm\_sw\_cal/sas\_frame.shtml

XMM XXX Virtual Observatory

esa

XMM-Newton Science Operations Centre

XMM-Newton Revolution 1185

Home Watchouts Software requirements Installation Download

User support  
Software & Calib.  
Data Access  
Science  
Links

Within these pages you may find all the information needed to download, install and start using SAS. To proceed, click on the specific link as shown on top of the page.

Before you start surfing through them, you may want to have a look at a very [concise SAS description](#).

The current version is **6.5.0**, released **August 17, 2005**.

A [patch](#) has been released.

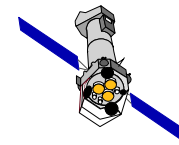
**Credits**

The XMM-Newton SAS is developed by the [Survey Science Centre](#) at the University of Leicester, Leicester, UK. Most of the code for the Reflector Grating Spectrometer is contributed by the [Columbia Astrophysics Laboratory](#).

[Webmaster](#)

[Designed with Netscape in mind] [Please report any problems to the [Webmaster](#)]

http://xmm.esac.esa.int/



**XMM-Newton**

Matteo Guainazzi --  
SCI-SDX

## ... before starting analysing data of an XMM-Newton observation:

- 1. Verify the quality of the pre-processed scientific products (PPS), produced by the automatic SSC processing**
- 2. Check the expected accuracy of the XMM-Newton calibrations, through:**
  1. Instrument calibration status reports
  2. SAS Science Validation Reports
  3. Current Calibration File (CCF) Release Notes
- 3. Compare your own set of calibration files with the latest available**
  1. Reduce the data again if a calibration file has changed, which may affect your scientific conclusions. Always stay on the safe side!
- 4. Once you have the SAS installed, your job is not finished ...**
  1. Check the SAS “watchout and evergreen” SAS pages, which contain known caveats or bugs
  2. Subscribe to the calibration mailing list
  3. Install an automatic mirror of the calibration files
  4. Make use the threads, would you like to learn something new
  5. Contact the HelpDesk, if everything else fails: [xmmhelp@xmm.vilspa.esa.es](mailto:xmmhelp@xmm.vilspa.esa.es)