

BNU Observational Astronomy Course



Lecture 4: **Data Reduction and Photometry**

CONTENTS

- **PART 1:** The basics of Data Reduction
-
- **PART 2:** The basics of Photometry

The basics of Data Reduction

Data Reduction

- In Lecture 2, we saw that when talking about photometric observations, the term **calibration frames** refers collectively to **bias**, **dark** and **flat** frames.
- These frames are used to remove unwanted signals and correct imperfections on the science data. Collectively, this process is called **data reduction**.
- The data reduction process involves first taking a series of each type of frame – we will soon see why we need a series of frames and not a single one of each type.
- The individual frames are then combined to create **master** frames, i.e. masterbias, masterdark and masterflat frames. These master frames are the ones we use to reduce our scientific data.

An actual astronomical image...

140	155	157	158	187	180	187	190	218	172	207	199	218	221	205	216	193	196	185	172	146	148	159	154	161
139	160	142	176	190	212	198	228	240	254	247	247	251	255	258	240	216	185	203	174	159	167	167	170	142
144	161	162	172	204	209	236	264	303	285	290	320	333	311	312	297	257	262	220	242	194	190	164	158	150
161	160	190	186	223	252	283	312	351	382	405	429	454	421	392	346	323	297	279	233	204	185	177	182	152
180	154	197	198	245	271	332	393	473	522	569	612	582	532	500	511	431	361	335	291	261	209	208	164	182
169	183	197	236	290	328	396	522	590	775	844	820	937	873	757	690	604	506	416	313	265	240	211	197	183
164	190	217	257	321	418	548	735	893	1141	1330	1448	1501	1399	1263	1037	758	688	523	390	335	283	225	202	192
191	202	224	293	402	535	733	998	1450	1833	2189	2460	2472	2265	1998	1530	1201	879	667	496	382	335	277	247	207
199	197	287	336	476	691	1008	1453	2188	3088	3948	4377	4343	3889	3233	2459	1747	1282	902	660	517	390	286	247	203
186	245	320	408	557	845	1342	2155	3338	5079	6892	7940	7792	6805	5460	3869	2610	1734	1103	842	599	409	317	247	208
203	253	339	450	689	1035	1669	2972	5157	8275	11696	13777	13464	11300	8419	5766	3725	2320	1460	966	701	456	346	266	221
228	246	360	481	699	1218	2085	3885	7044	11843	17738	21638	20549	16644	11974	7783	4915	2851	1725	1140	757	525	367	286	253
220	278	352	487	775	1316	2326	4305	8404	14919	22477	26194	24628	19709	13871	9370	5602	3300	1994	1262	783	568	360	289	221
206	292	337	501	766	1295	2328	4286	8225	14797	22034	25366	23374	18267	13222	9199	5574	3346	1995	1287	840	573	391	326	235
226	250	323	498	678	1218	2024	3741	6852	11673	16734	18833	17846	14049	10527	7184	4837	2898	1847	1118	724	500	364	313	249
206	236	299	410	589	974	1606	2798	4721	7392	10245	11673	11018	9130	6894	5029	3454	2256	1431	960	645	453	356	232	198
209	240	275	383	524	806	1178	1937	2909	4376	5662	6448	6295	5345	4307	3190	2283	1608	1132	786	512	388	308	257	205
179	216	268	326	425	586	832	1241	1785	2429	3039	3500	3452	3032	2495	2007	1540	1072	783	589	463	325	255	220	227
175	213	232	272	352	454	591	768	1066	1369	1671	1823	1835	1792	1537	1255	1011	773	606	445	364	298	240	177	183
163	181	217	286	278	349	467	549	665	865	1011	1023	1108	1072	917	788	665	572	480	353	283	248	217	195	163
170	188	219	207	260	265	336	415	491	546	612	683	675	610	608	550	474	419	344	309	251	216	196	164	166
132	182	182	230	215	255	274	289	325	393	408	439	439	451	431	385	348	318	289	249	220	198	176	177	157
176	163	157	175	208	228	225	248	289	284	322	305	338	340	312	322	251	271	250	207	174	188	178	165	155
146	171	151	165	173	167	192	199	210	229	254	240	248	283	255	247	252	242	193	221	184	179	177	156	165
146	128	143	179	194	169	180	201	197	217	201	223	229	233	207	205	216	176	214	187	191	163	154	153	146

Combining frames

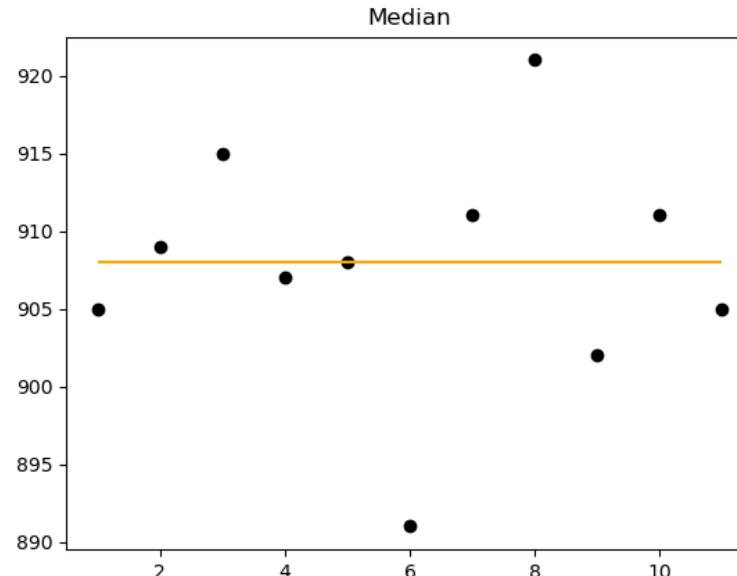
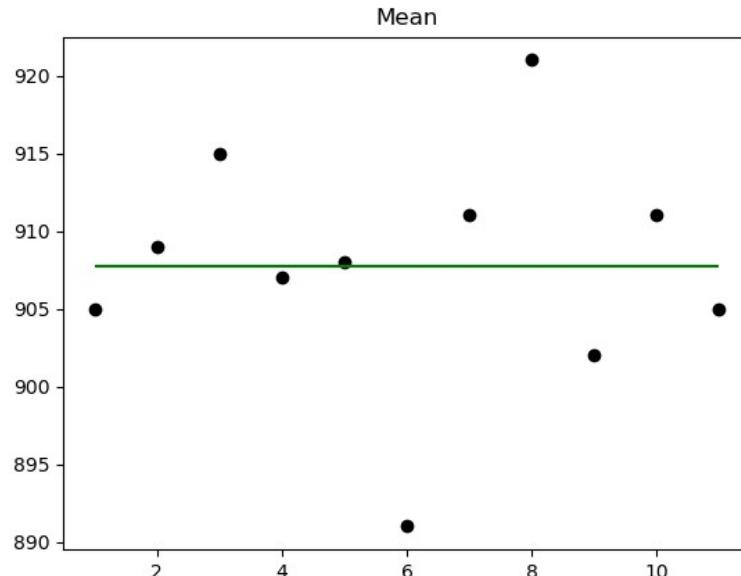
- Remember that astronomical data are basically just numbers. The same is true for the calibration frames!
- When we combine frames to create the corresponding master frame, we perform simple mathematical operations with these numbers, like for example calculating the *average*, the *median*, the *standard deviation* etc

Combining Frames – an example

- Imagine we have 11 frames and the values for pixel (1,1) in each of these 11 frames are the following:

905, 909, 915, 907, 908, 891, 911, 921, 902, 911, 905

- If we combine these values with the **average/mean** or the **median**, the output pixel (1,1) in the master frame will have a value of **907.73** (average) or **908** (median)

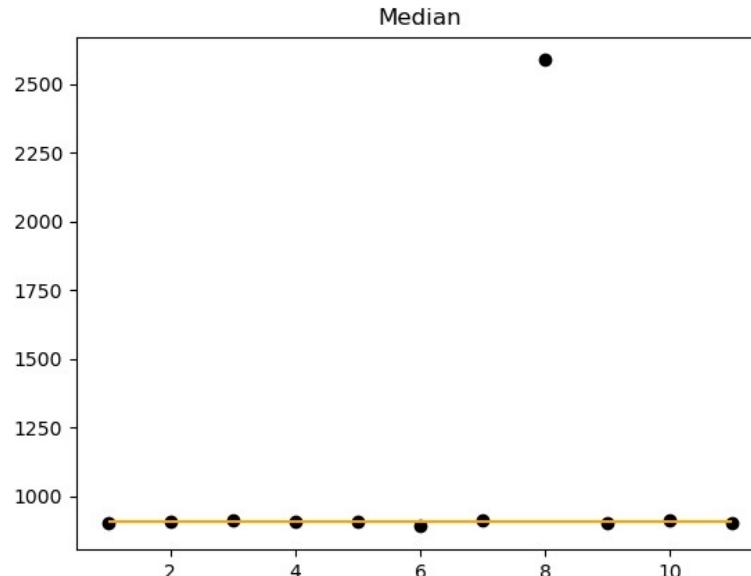
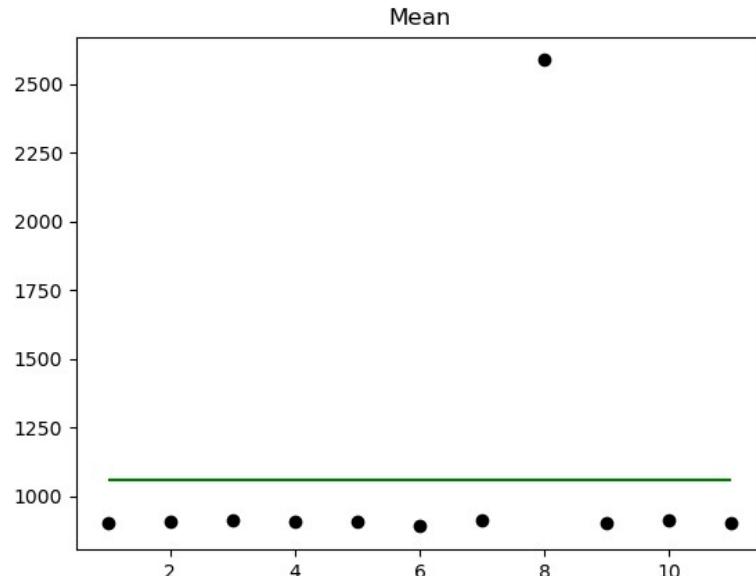


Combining Frames – outliers

- Now imagine that, for some reason, one of the values is very different from the others. Such a value is usually called an **outlier**:

905, 909, 915, 907, 908, 891, 911, **2587**, 902, 911, 905

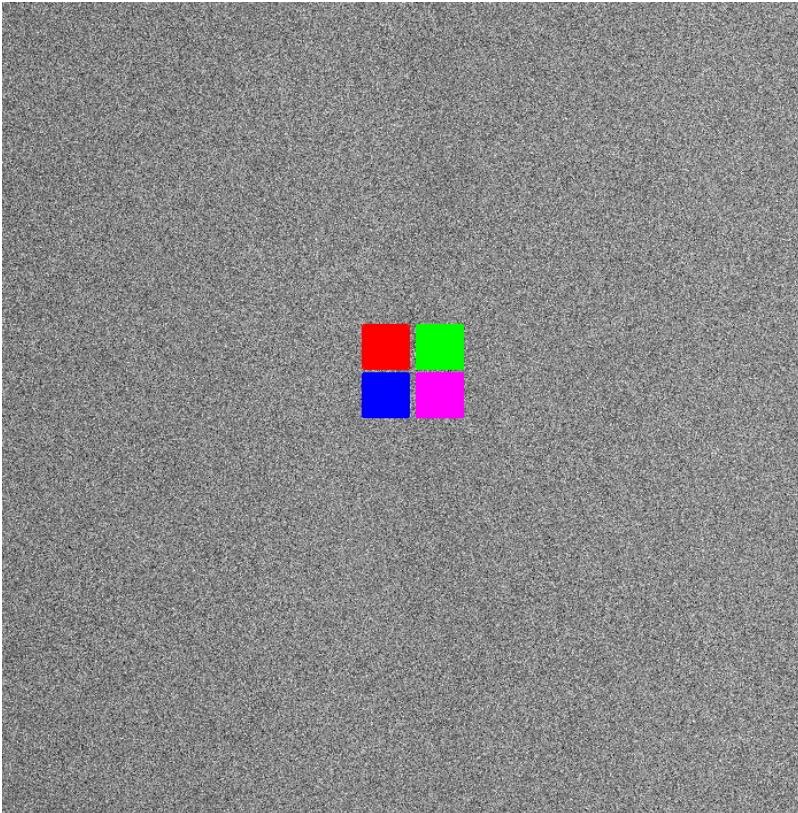
- Combining these values with the **average/mean** now gives **1059.18**, whereas using the **median**, we still get a value of **908**.



BIAS FRAMES

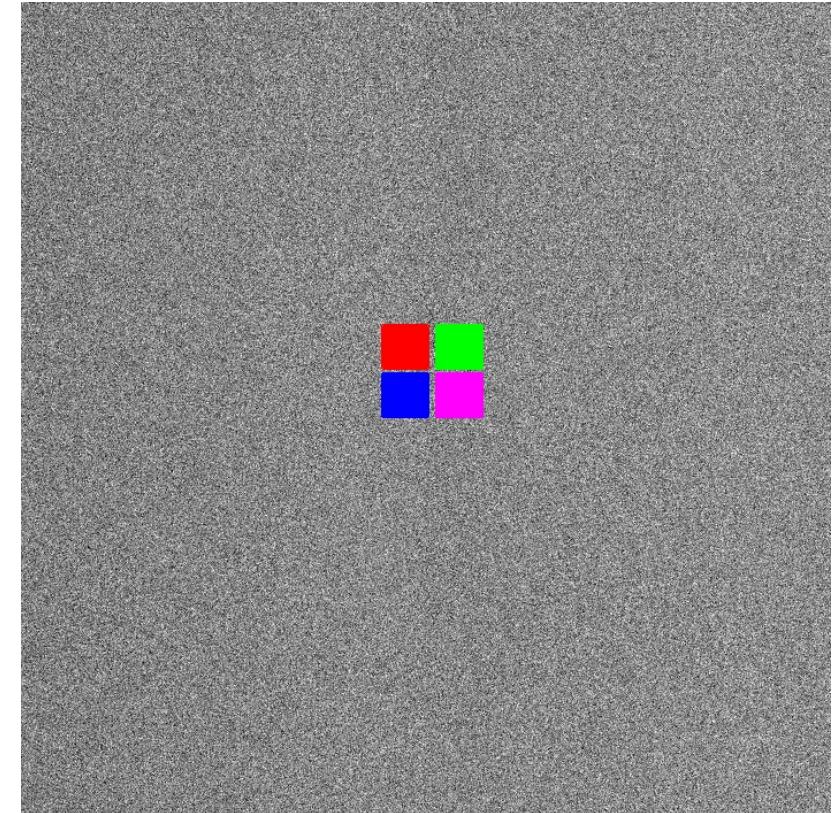
BIAS

Let's take a series of bias frames and focus on 4 pixels (always the same ones)



Bias frame #1

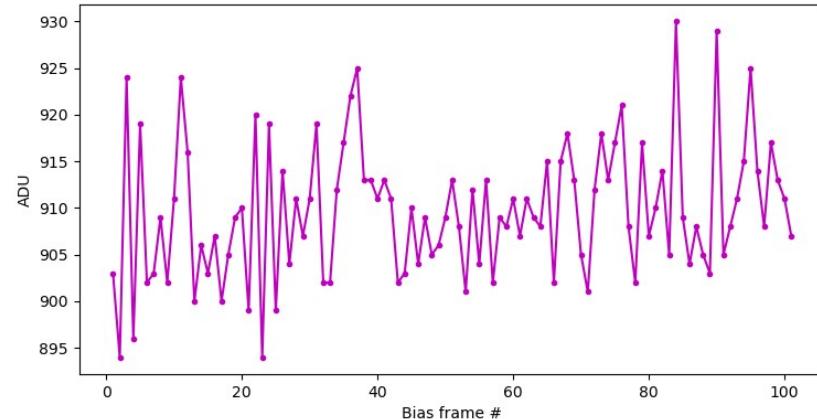
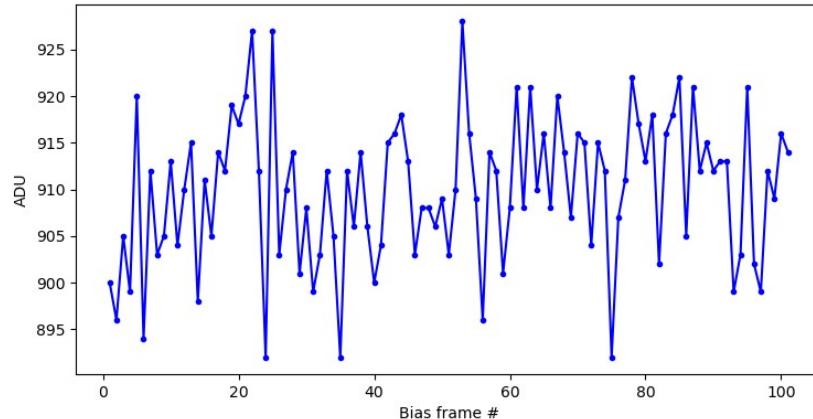
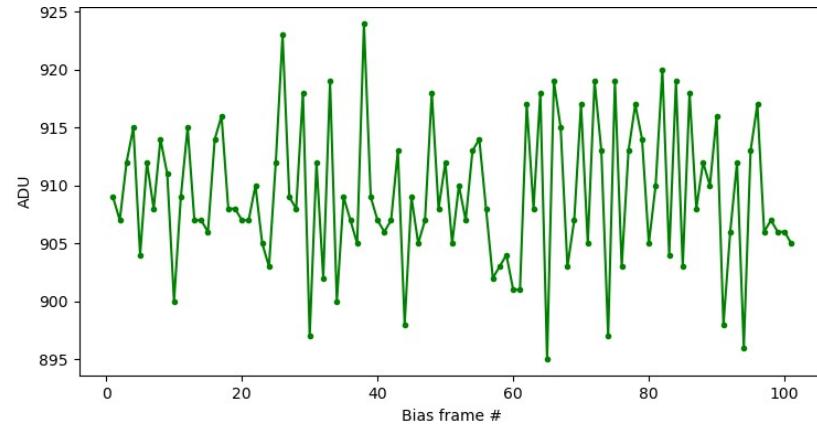
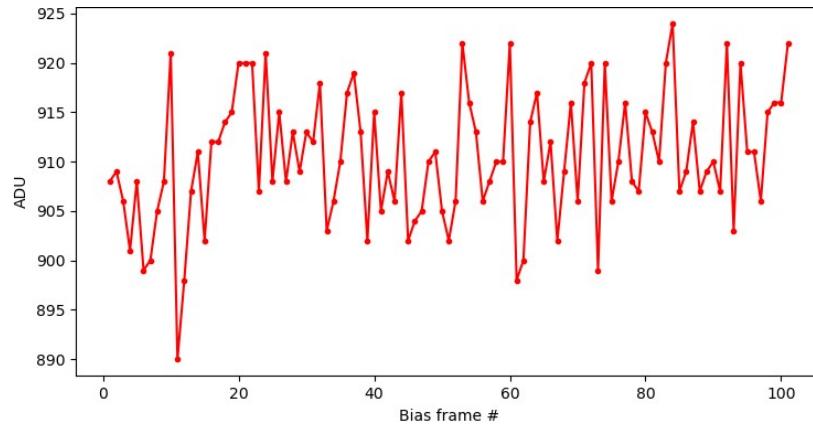
• • •



Bias frame #101

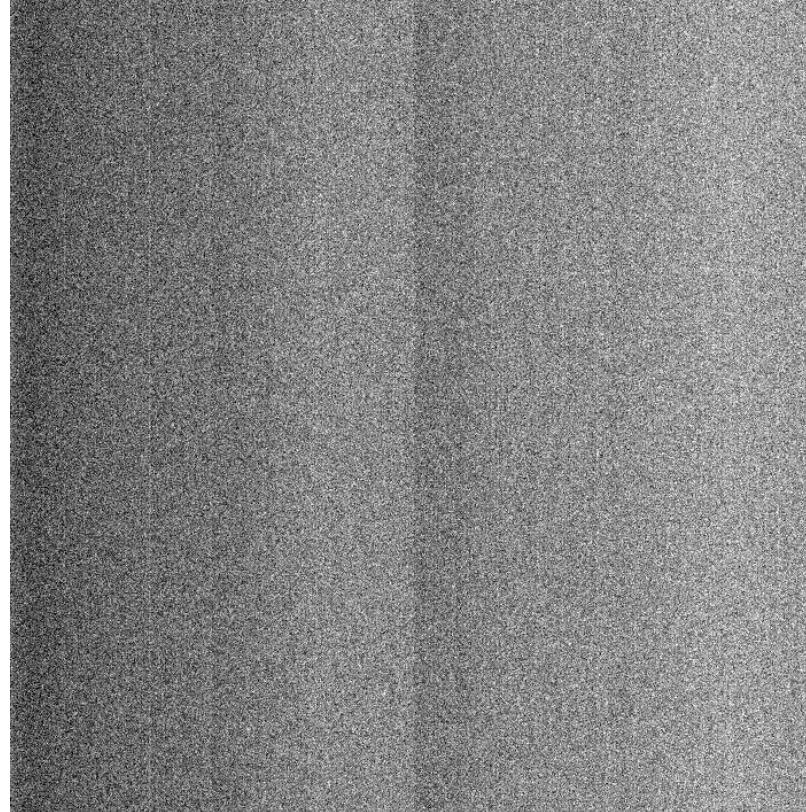
BIAS

What are the values of each of the 4 pixels in the series of 101 bias frames?



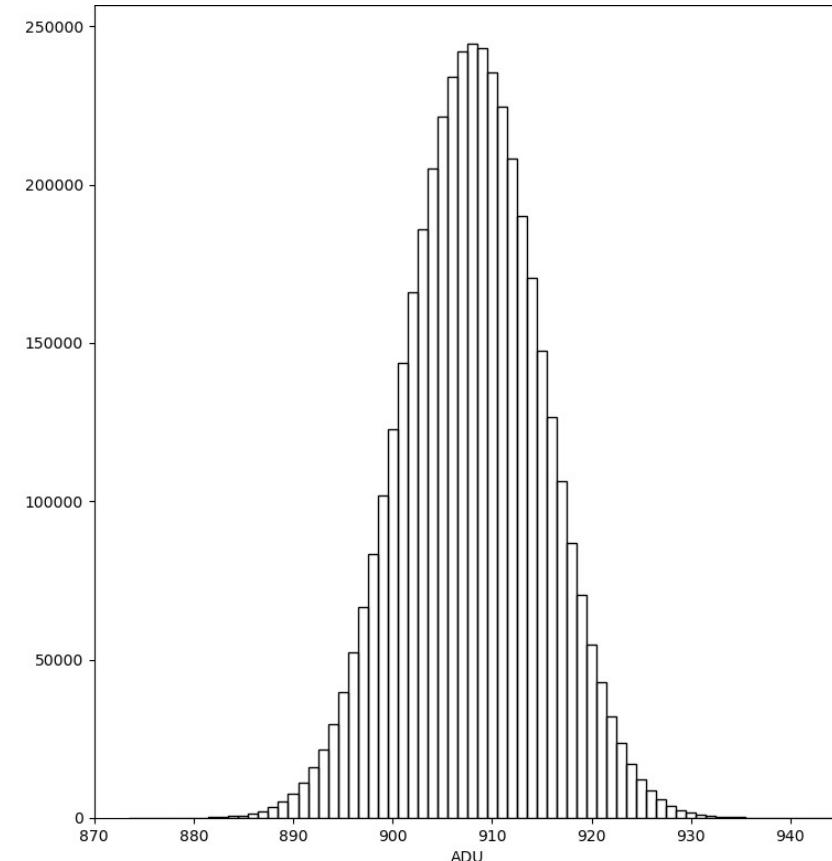
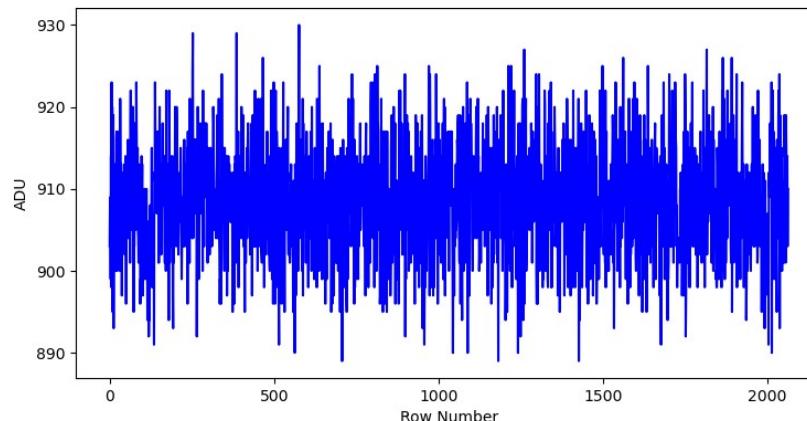
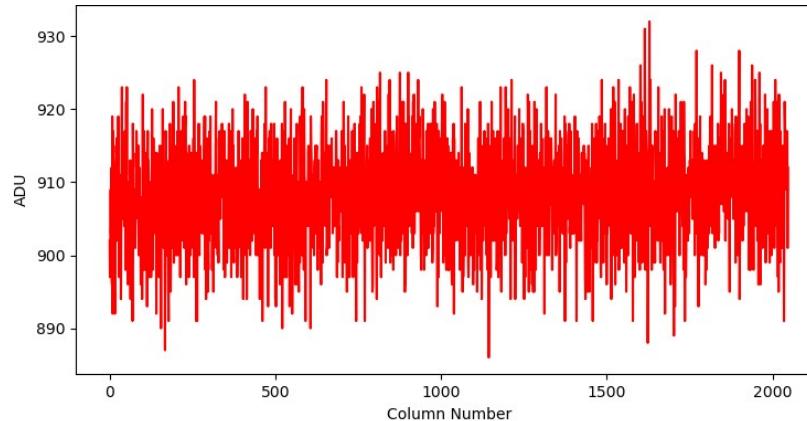
BIAS

We can combine a series of bias frames to create a *master* bias frame



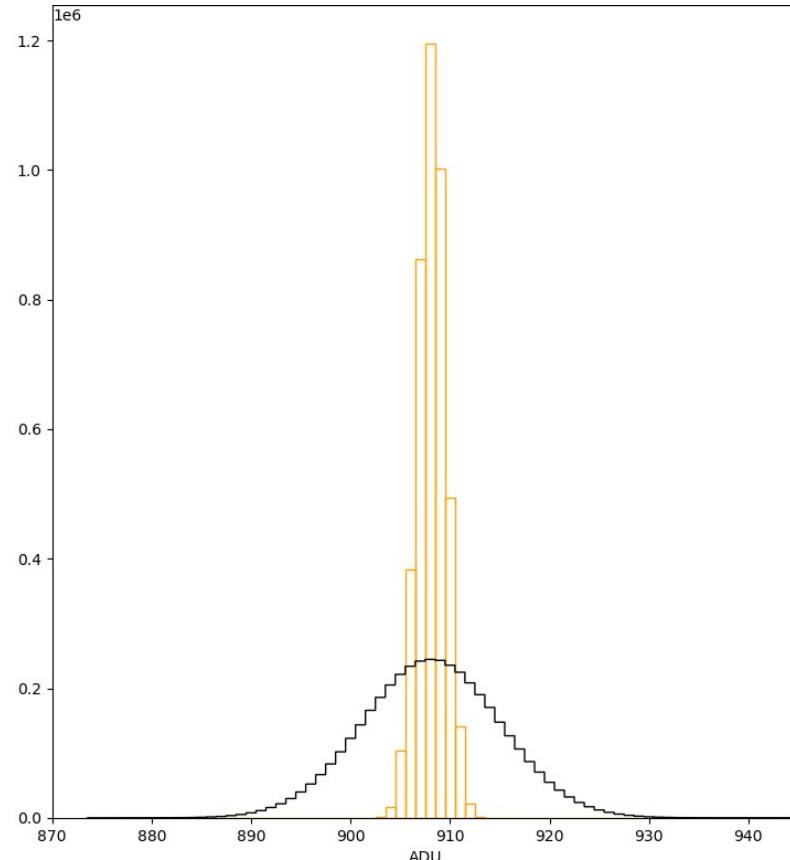
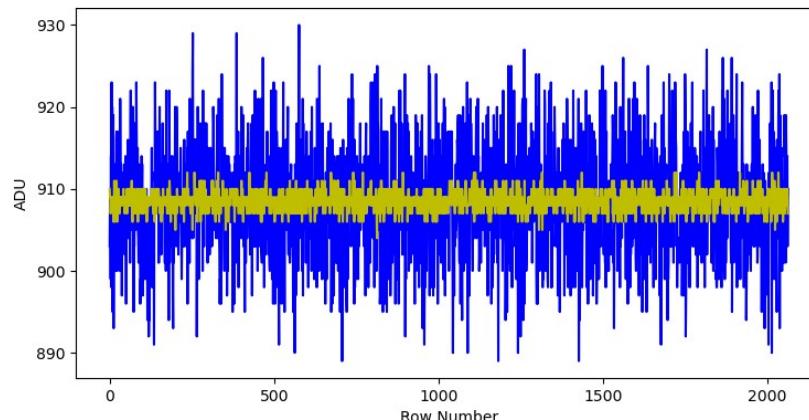
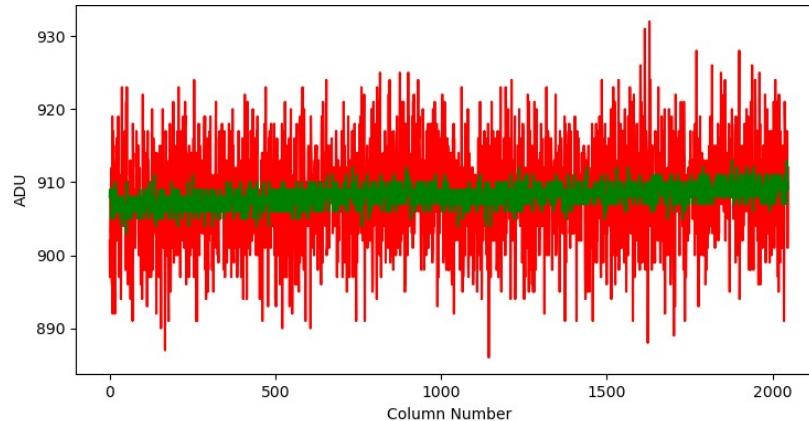
BIAS

Quick reminder: these were the results from a single bias frame



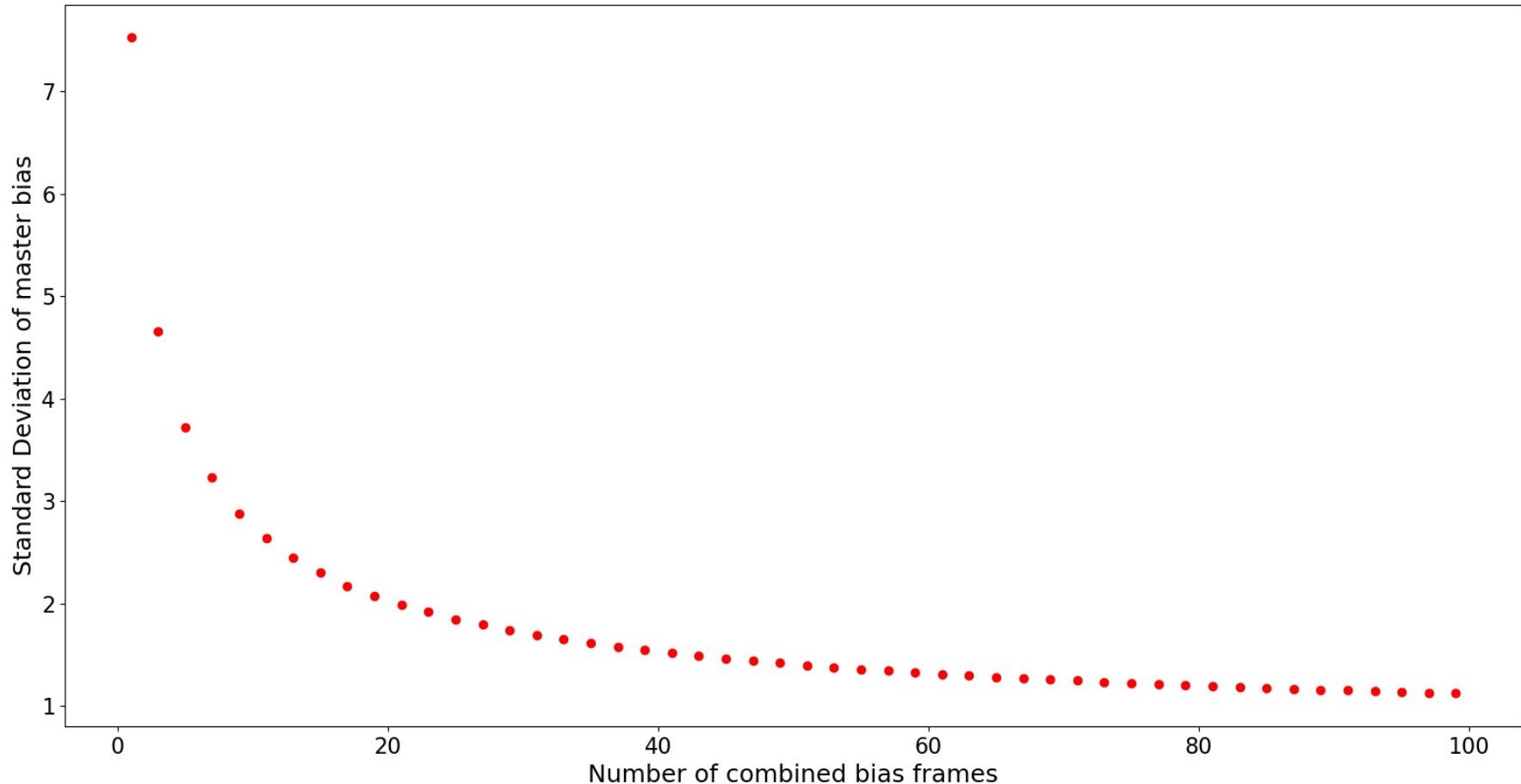
BIAS

How do they compare with the (same) results from the master bias frame?



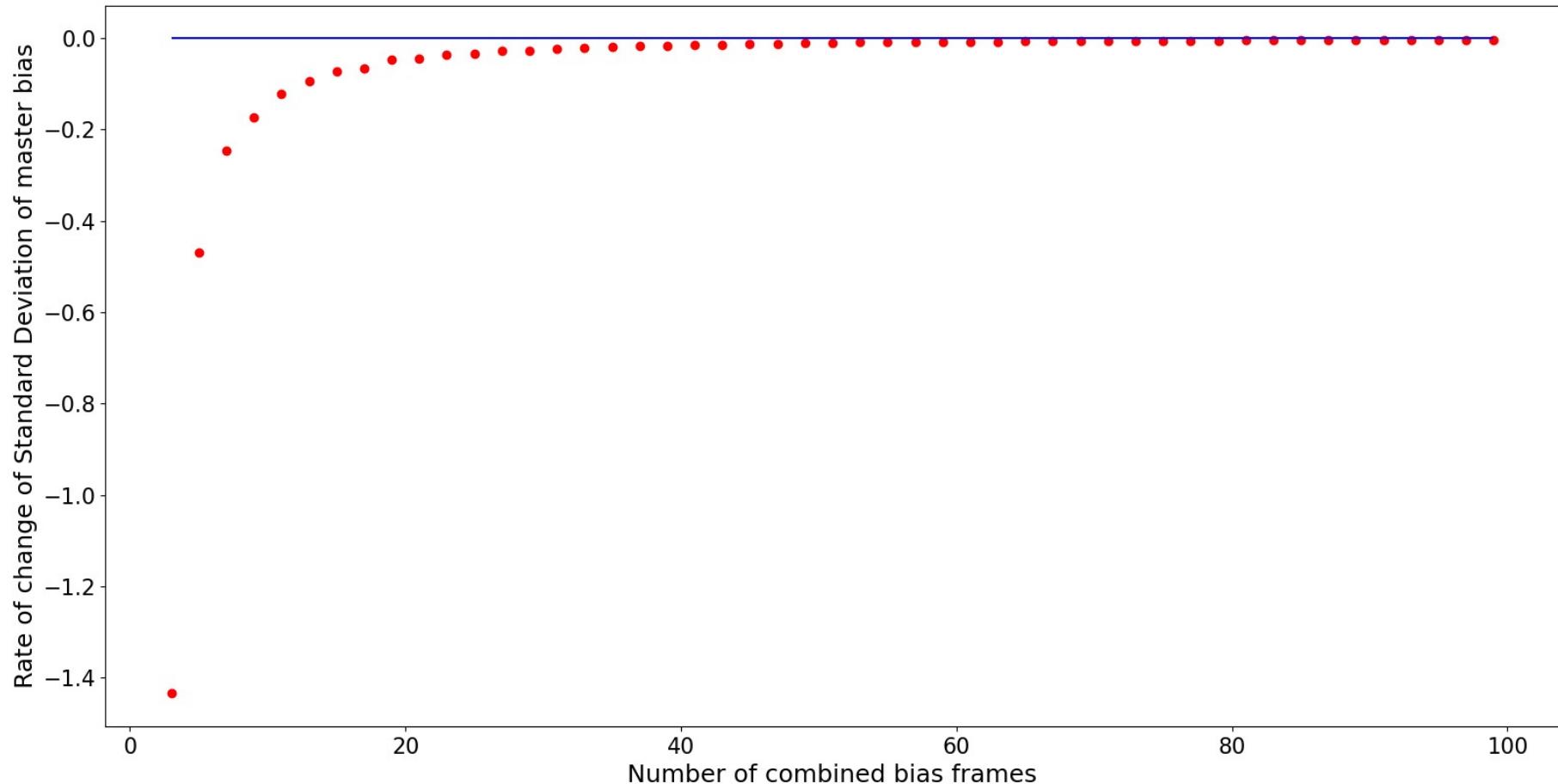
BIAS

How many bias frames do we need to combine? How does the width of the master change?



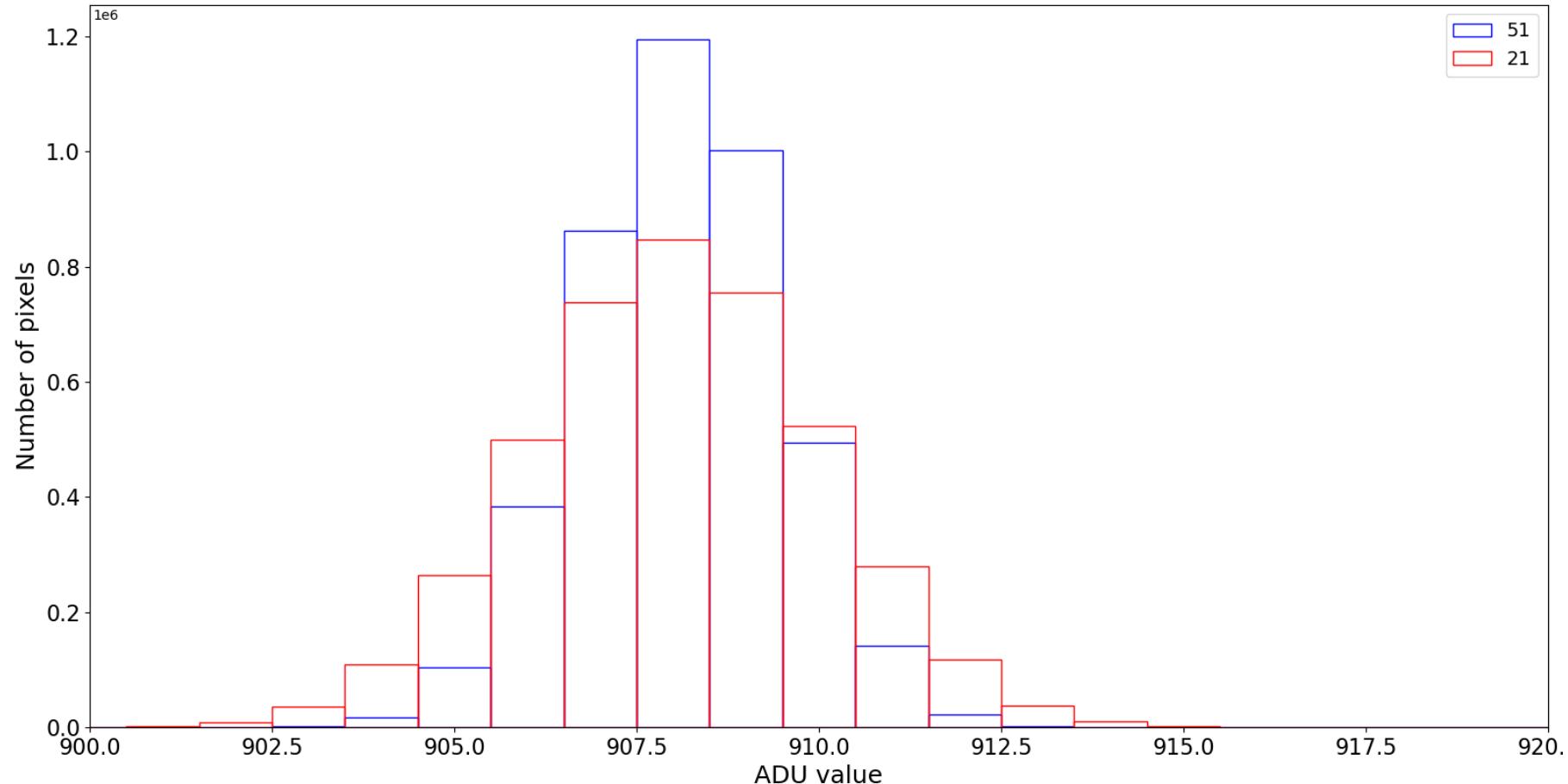
BIAS

Another way of looking at it: how does the slope change (~ “derivative”)?



BIAS

Comparing two different master bias frames



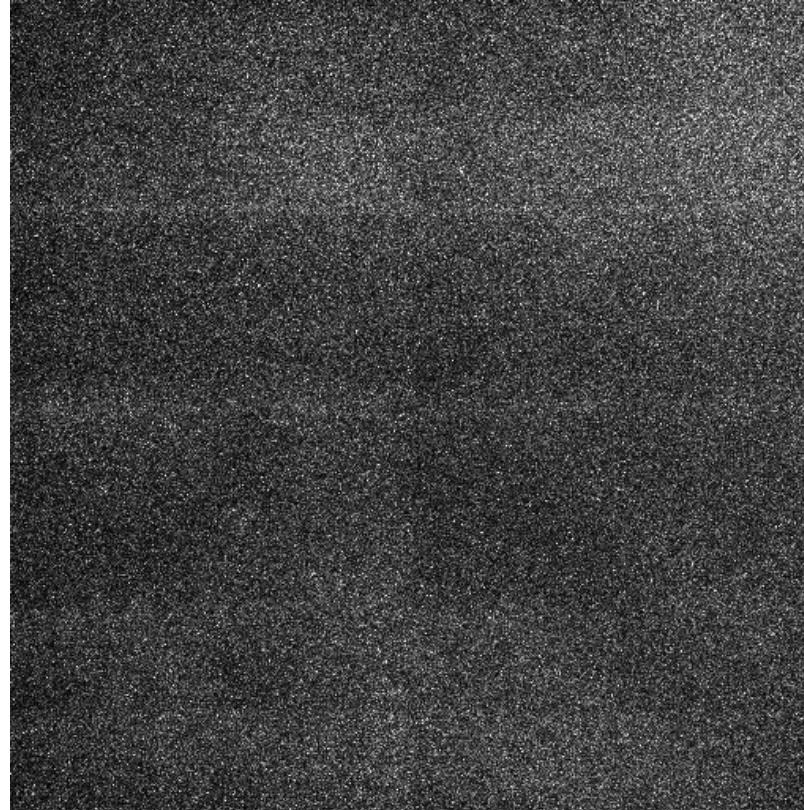
A small, dark, abstract graphic element consisting of a cluster of small dots or a smudge, positioned next to the word "GALACTICA".

GALACTICA

DARK FRAMES

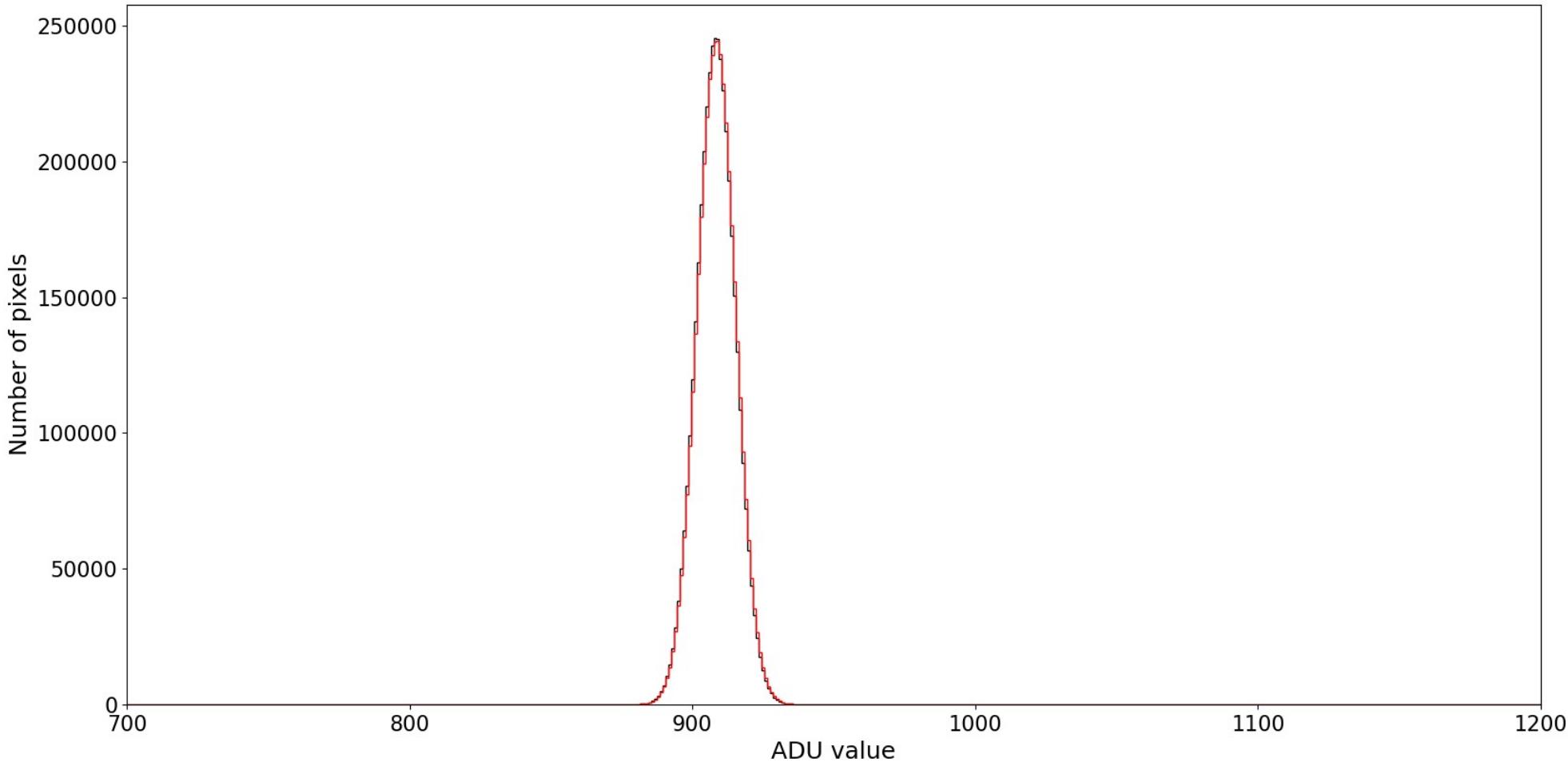
DARK

We can combine a series of dark frames to create a *master* dark frame



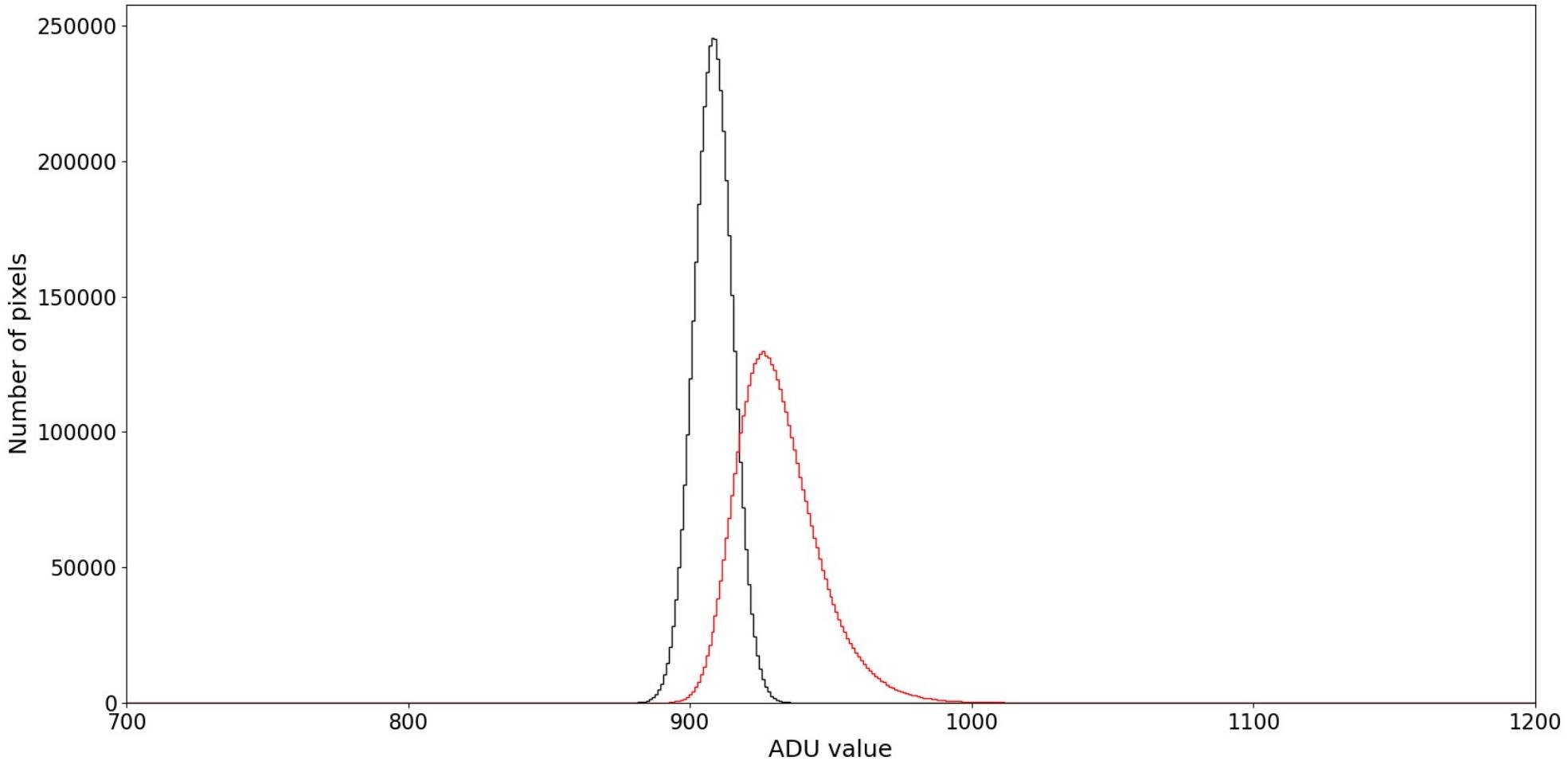
DARK

Do we need dark frames? Comparing a bias with a 1 sec dark



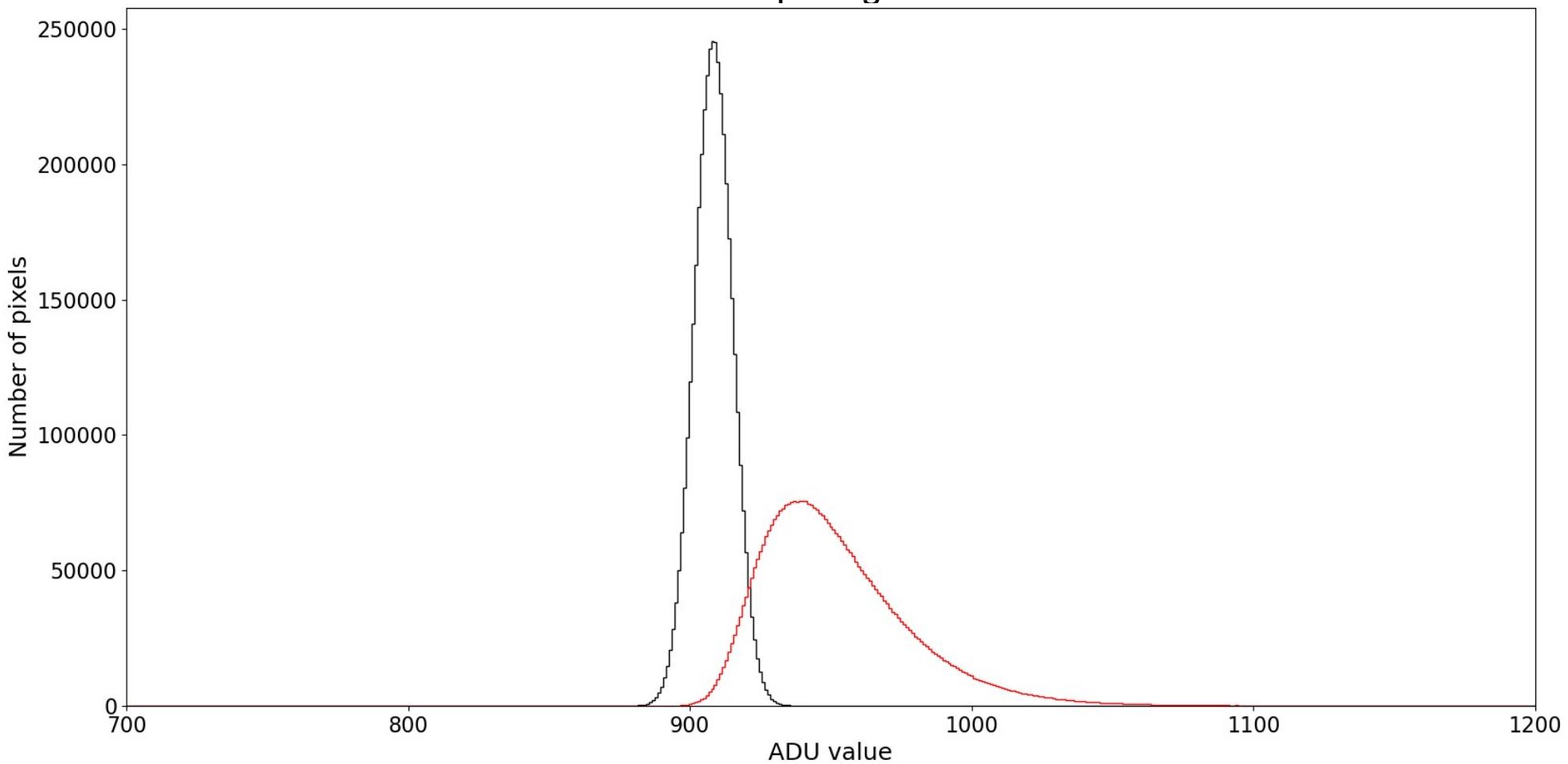
DARK

Do we need dark frames? Comparing a bias with a 30 sec dark



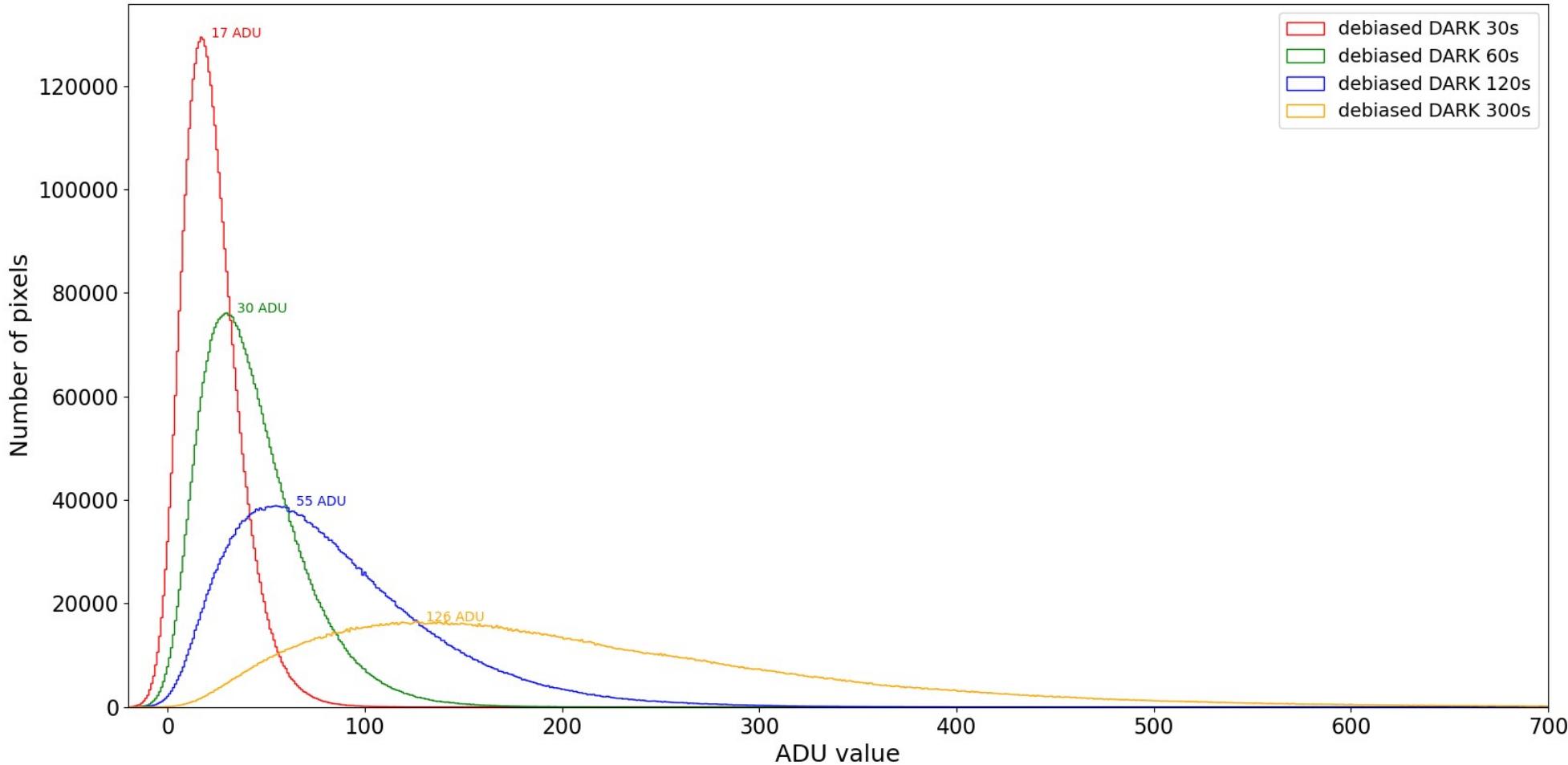
DARK

Do we need dark frames? Comparing a bias with a 60 sec dark



DARK

Do we need dark frames? Comparing various de-biased dark frames

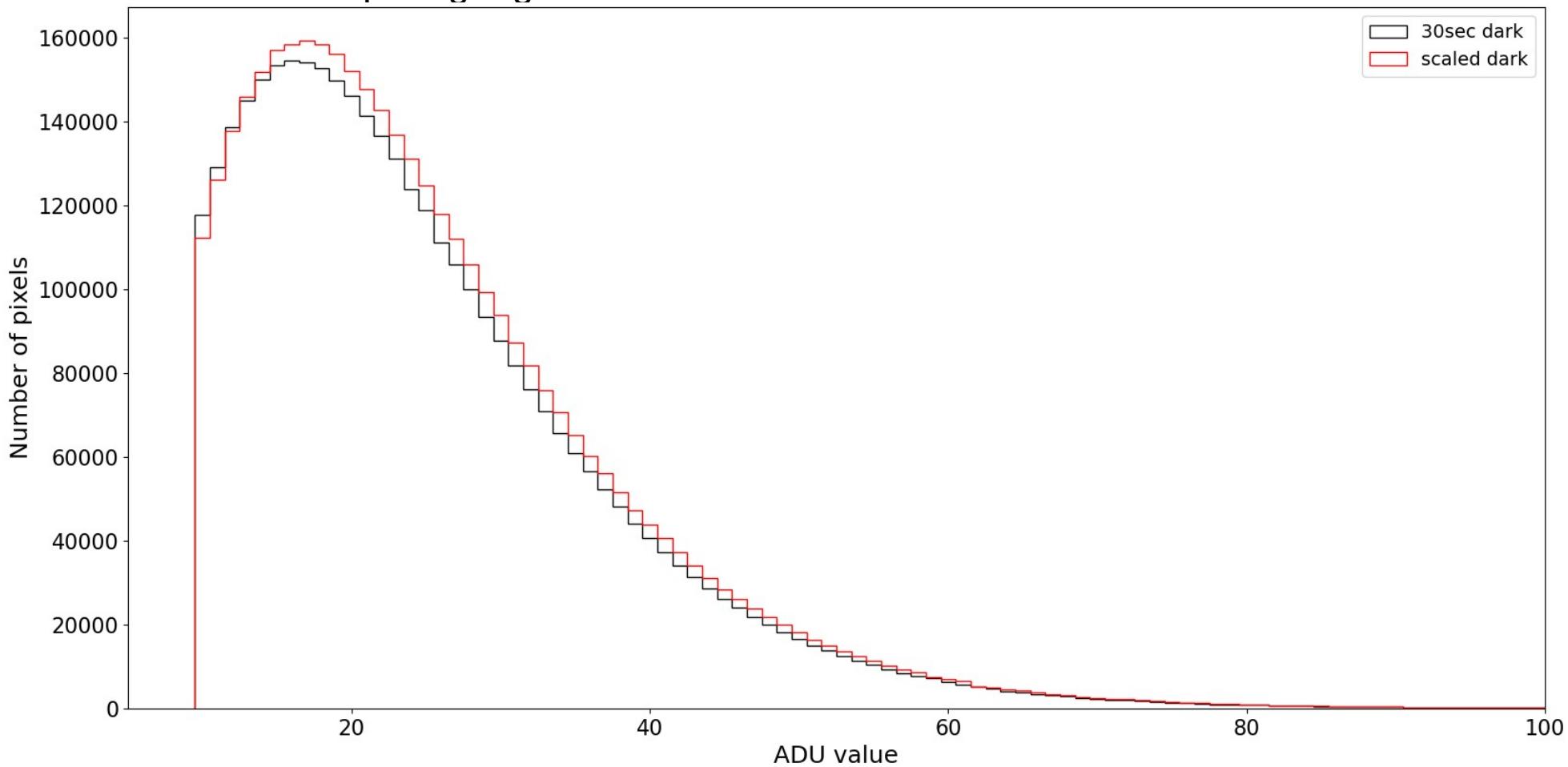


Scaled Dark frames

- Sometimes, it can be difficult to obtain all the necessary dark frames, for example when we have various exposure times due to observing multiple objects in multiple filters.
- We saw that the dark current increases linearly with time, $C_D(T) * t_{exp,dark}$, so we can use a technique called **dark scaling**
- For this technique, we obtain dark frames only in one $t_{exp,dark}$ (usually, a long one). We then subtract the *masterbias* from each of these dark frames, combine them into a *masterdark*, and then **normalise** the masterdark, i.e. divide it by $t_{exp,dark}$.
- So, we have a representation of pure dark signal (as we removed the bias) in 1sec. Then, we can **scale** this normalised masterdark to each of our science exposure times: $MDARK_{science} = t_{exp,science} * MDARK_{norm}$

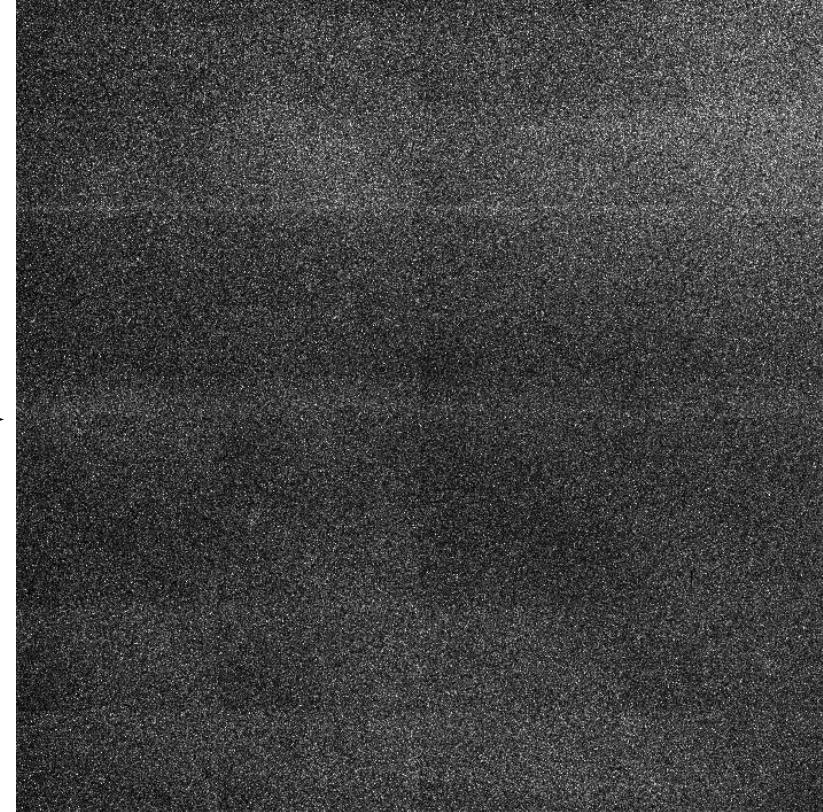
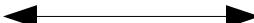
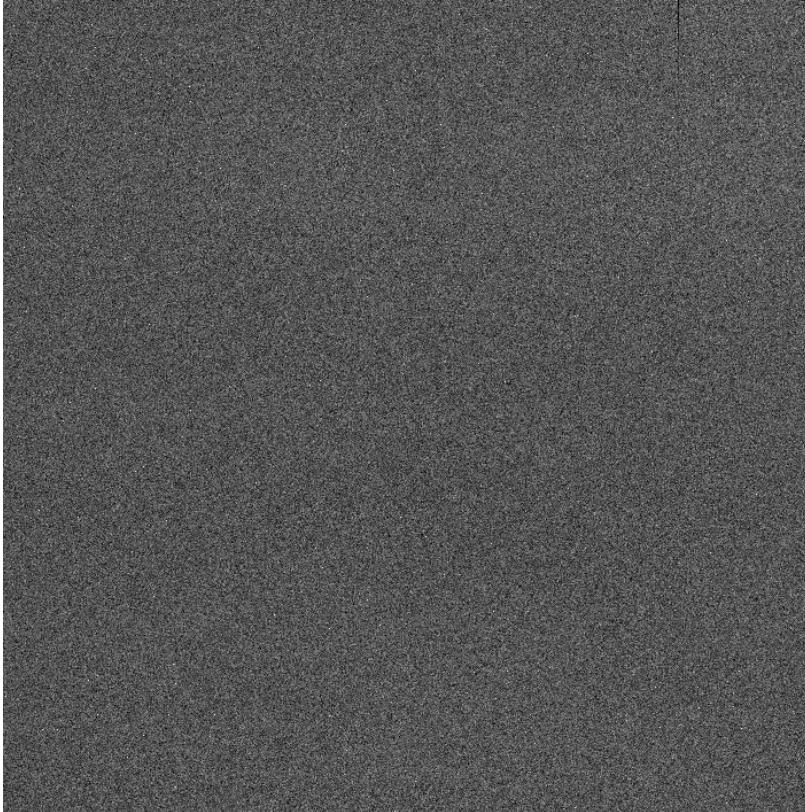
DARK

Comparing a genuine 30sec dark with a scaled 30 sec dark



DARK

A 1 sec dark frame compared to a scaled 1 sec dark frame



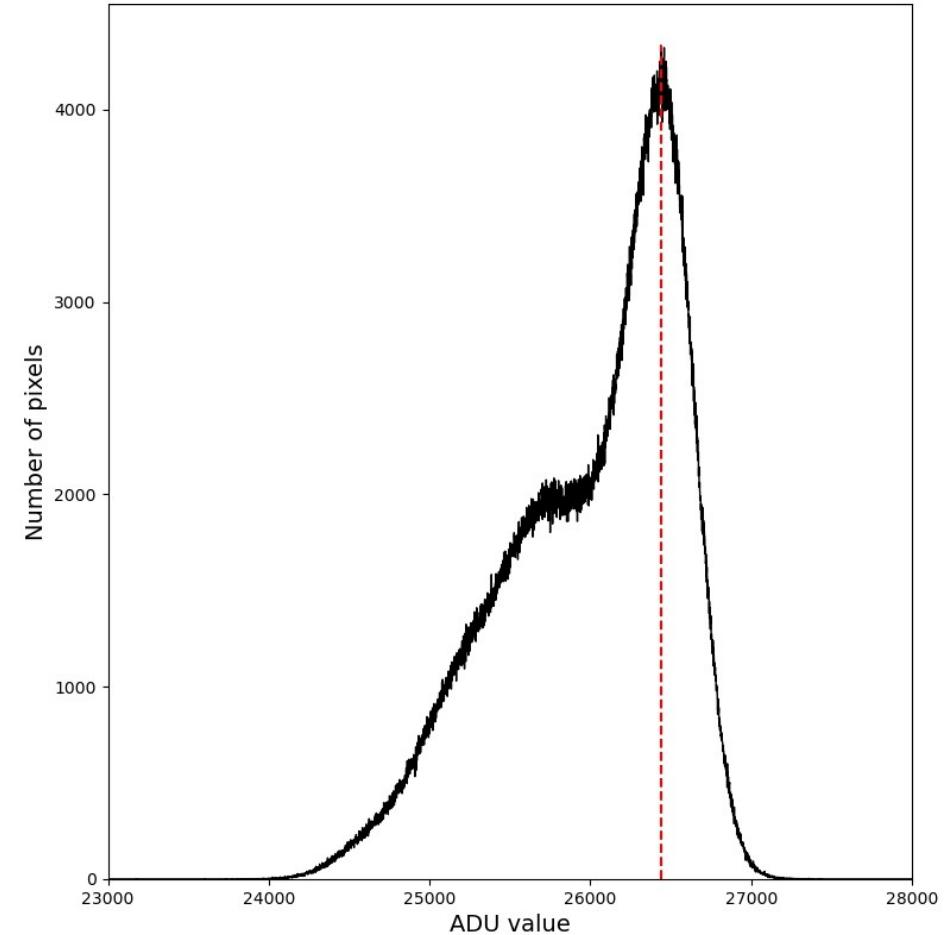
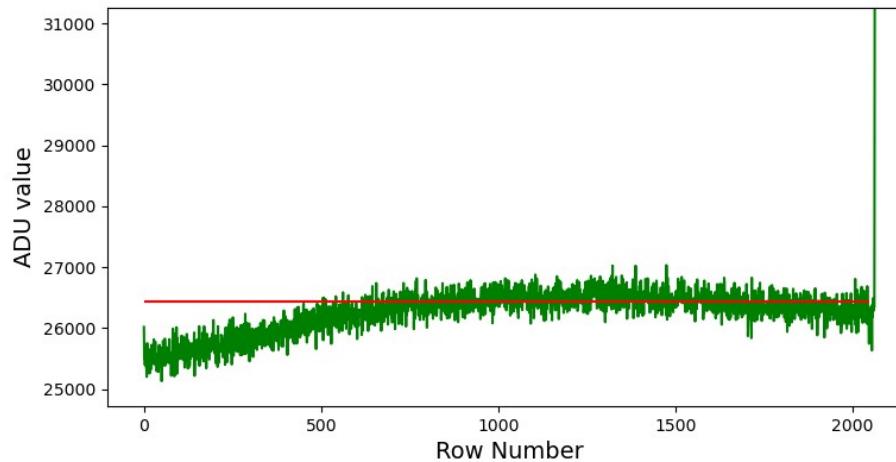
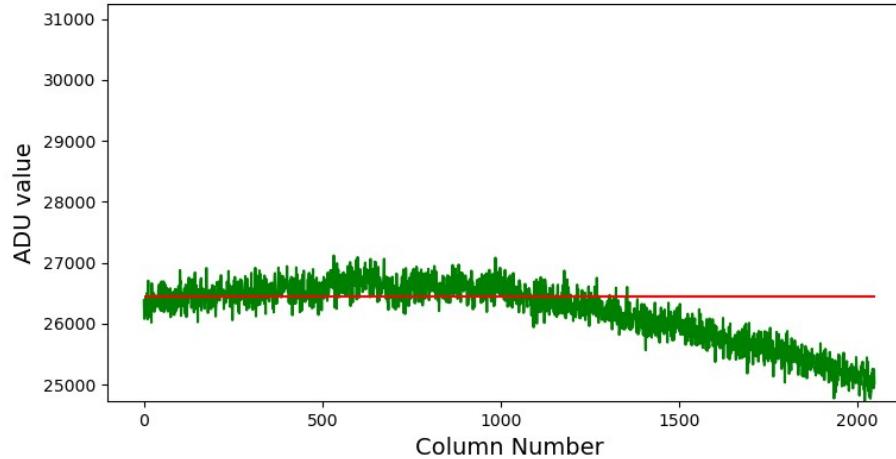


GALACTICA

FLAT FRAMES

FLAT

Reminder of how a flat frame looks like



FLAT

- Ideally, the *same* amount of light falling on *any* pixel on the CCD chip would produce the same counts

26.000	26.000
26.000	26.000

- In reality, because of the reasons we mentioned, this is *not* the case

26.694	25.589
27.079	25.058

FLAT

- In order to be able to correct the imperfections we discussed, we are not really interested in the *absolute* value of the ADU counts

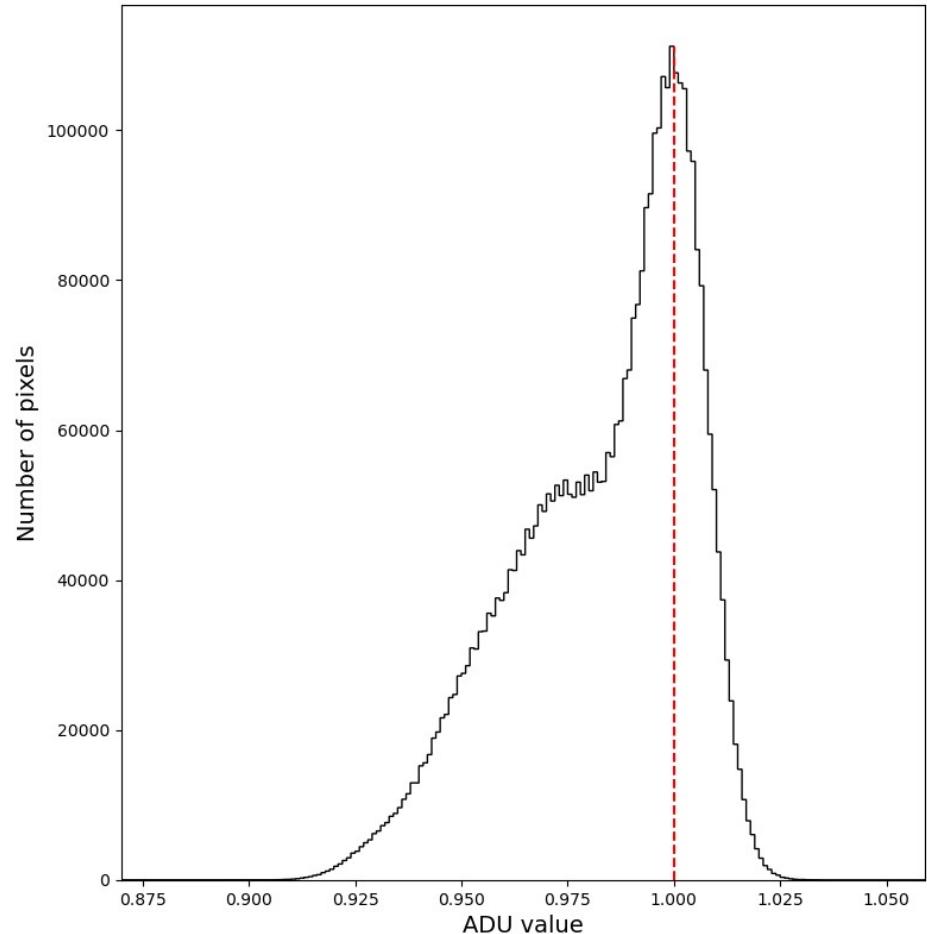
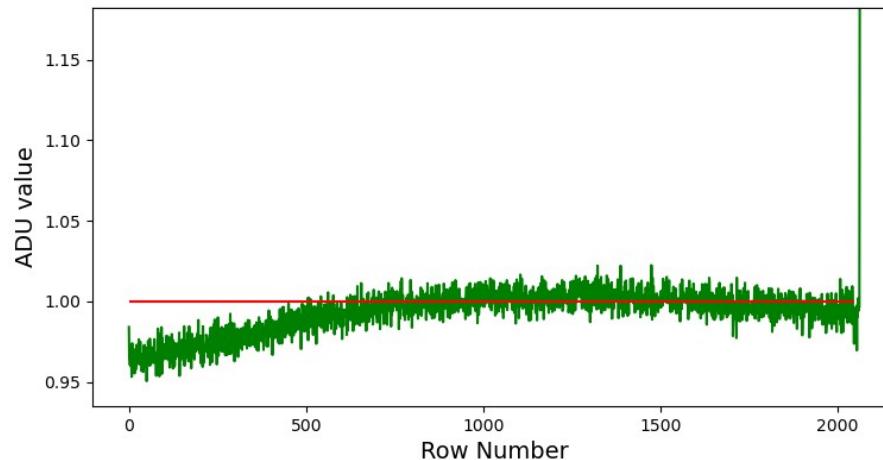
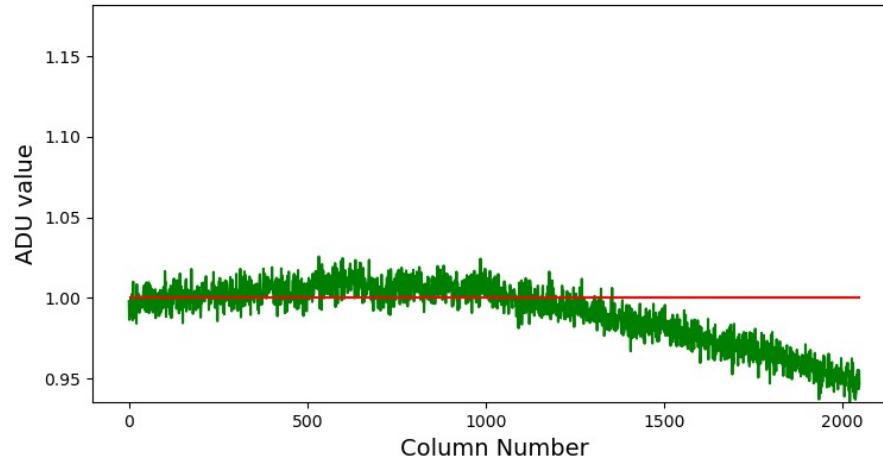
26.694	25.589
27.079	25.058

- But rather, we want to know the *relative* amount registered in each pixel, with respect to some reference value. Therefore, we *normalise* the flat frames with the value we choose as reference

1.0267	0.9842
1.0415	0.9638

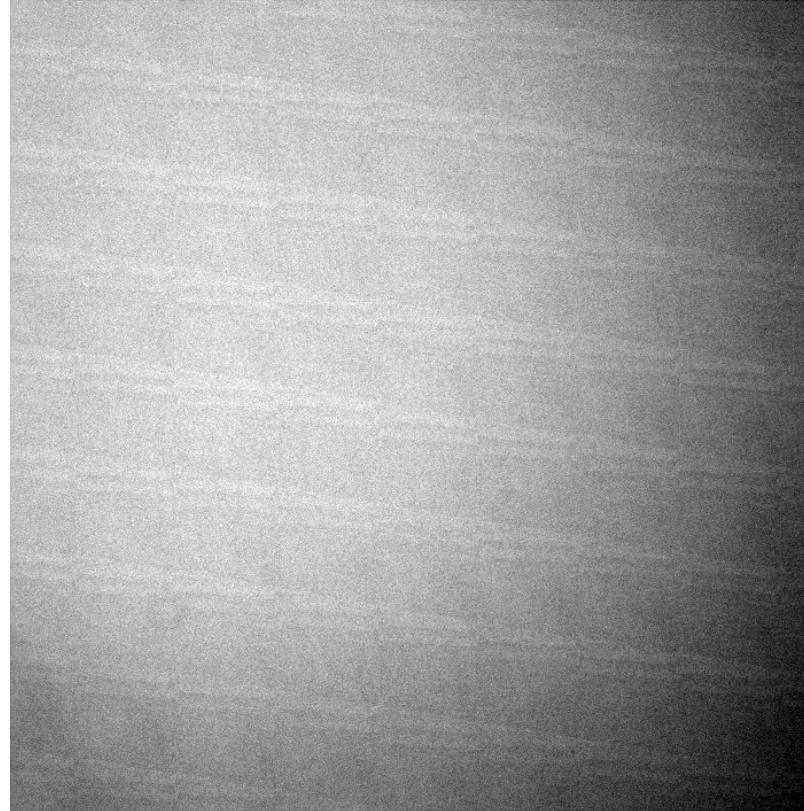
FLAT

Normalising a flat frame by its mode value



FLAT

We can combine a series of normalised flat frames to create a *master flat*



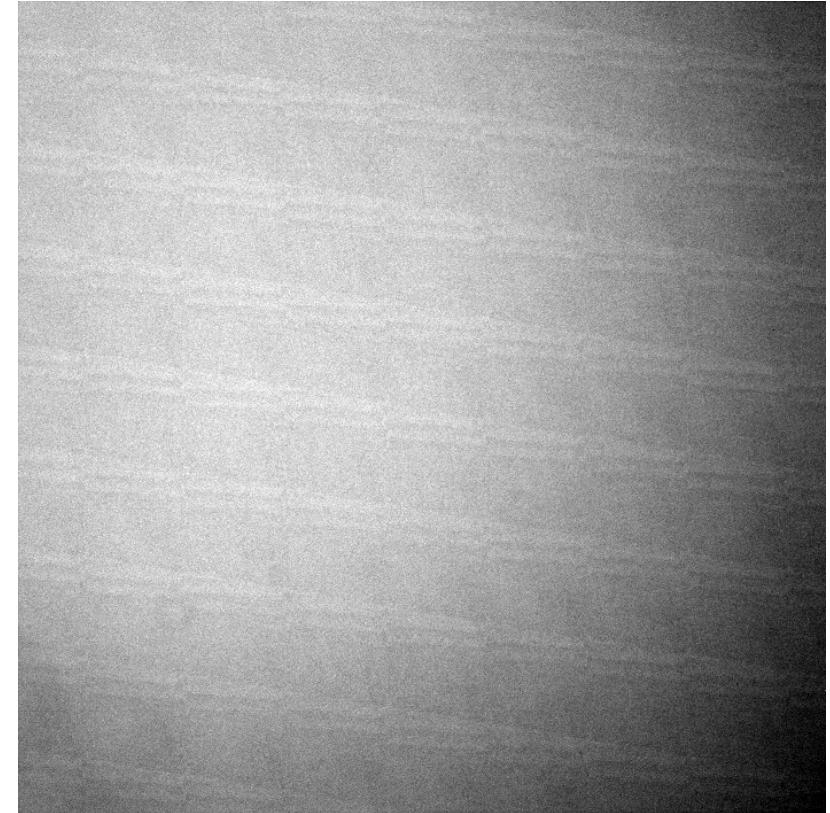
FLAT

We can apply the master flat on a single flat frame by division



Single flat frame (original)

/



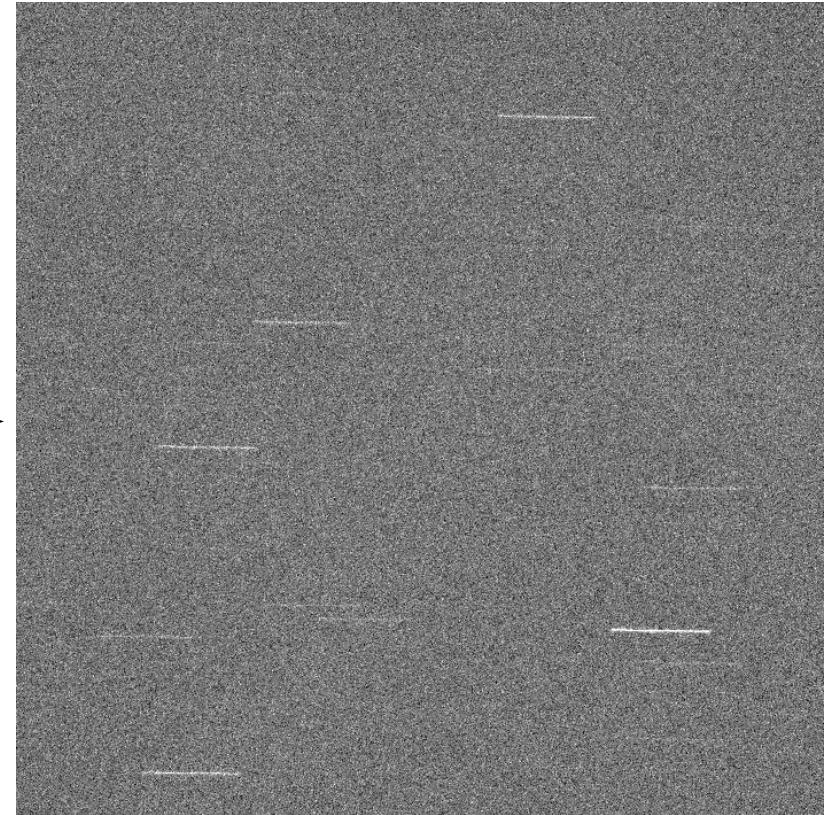
Master flat frame

FLAT

The difference before and after the correction is very obvious!



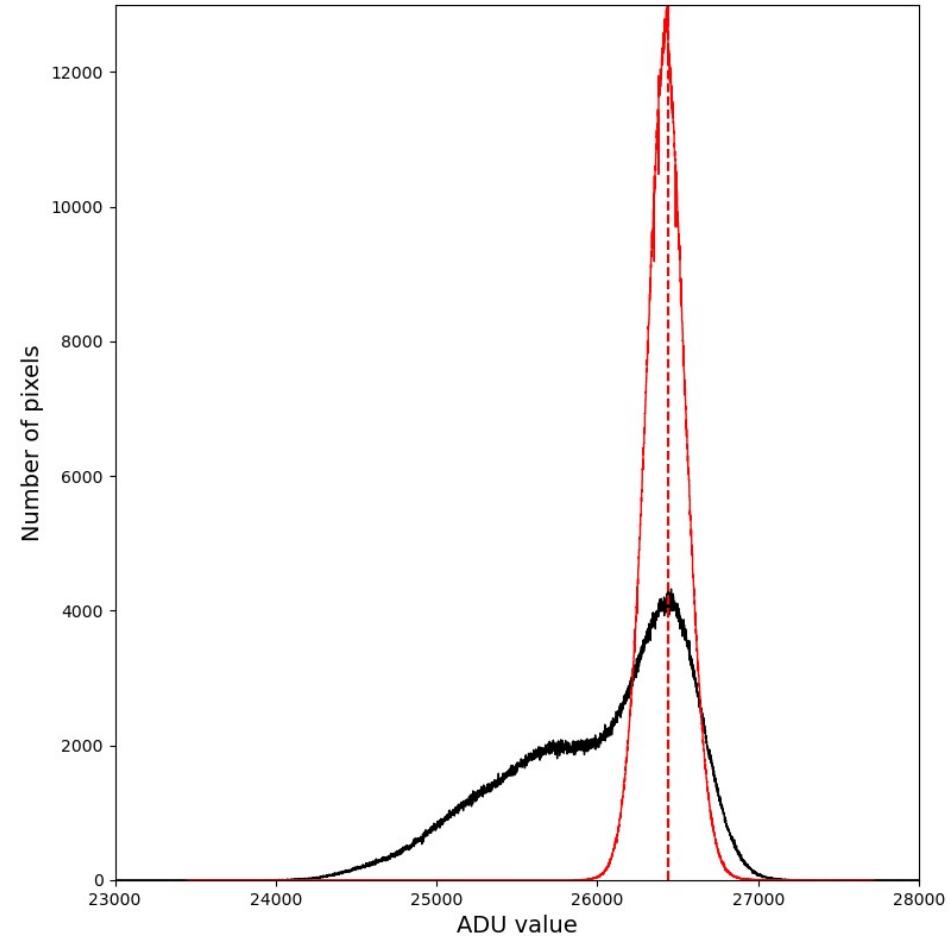
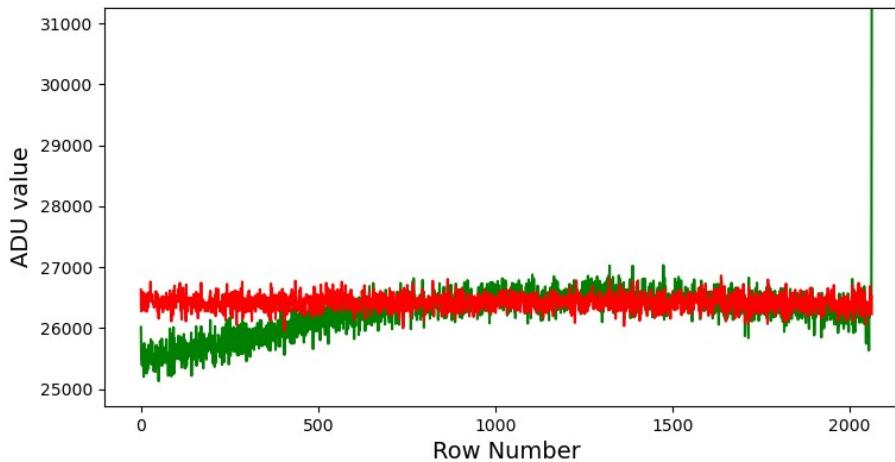
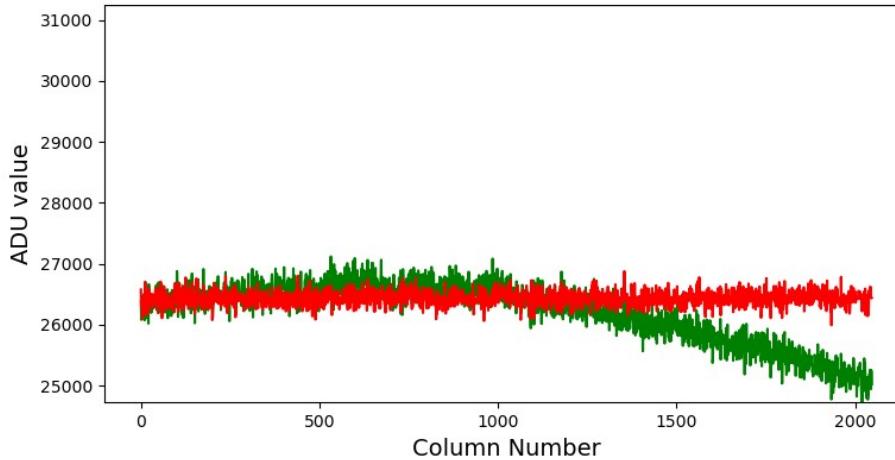
Single flat frame (original)



Single flat frame (corrected)

FLAT

Comparing the original and the corrected flat frame



SUMMARY:

DATA REDUCTION FLOW

Data Reduction – if no darks needed

- Set of BIAS frames → **combine** → masterbias
- Set of FLAT frames → **subtract** masterbias → **normalise** by reference value
→ **combine** → **masterflat**
- SCIENCE data → **subtract** masterbias → **divide by masterflat**

Data Reduction – if darks are needed

Method 1: **ONLY IF** $t_{\text{exp,science}} = t_{\text{exp,dark}}$

- Set of BIAS frames → **combine** → masterbias
- Set of DARK frames → **combine** → masterdark
- Set of FLAT frames → **subtract** masterbias → **normalise** by reference value
→ **combine** → masterflat
- SCIENCE data → **subtract** masterdark → **divide by** masterflat

Data Reduction – if darks are needed

Method 2: **ONLY IF** $t_{\text{exp,science}} = t_{\text{exp,dark}}$

- Set of BIAS frames → **combine** → masterbias
- Set of DARK frames → **subtract** masterbias → **combine** → masterdark
- Set of FLAT frames → **subtract** masterbias → **normalise** by reference value
→ **combine** → masterflat
- SCIENCE data → **subtract** masterbias → **subtract** masterdark
→ **divide by** masterflat

Data Reduction – if darks are needed

With scaled darks

- Set of BIAS frames → **combine** → masterbias
- Set of DARK frames → **subtract** masterbias → **combine** → **masterdark_{temp}**
→ **normalise** by $t_{\text{exp,dark}}$ → **masterdark_{norm}** → **scale by** $t_{\text{exp,science}} * \text{masterdark}_{\text{norm}}$
→ **masterdark_{science}**
- Set of FLAT frames → **subtract** masterbias → **normalise** by reference value
→ **combine** → **masterflat**
- SCIENCE data → **subtract** masterbias → **subtract** masterdark_{science}
→ **divide by** masterflat

The basics of Photometry

Photometry

The technique called **photometry** allows us to measure the amount of light that has reached our CCD from a celestial object during an exposure.

In our case, this amount of light is represented by the number of ADU counts in the image (from photons to photoelectrons to ADUs).

The goal of photometry is to measure the **total** amount of ADU counts of the celestial object we are observing.

There are a few different ways that we can perform photometry on a celestial object.

For the purposes of this course, we will go over one technique called **aperture photometry** and we will mainly apply it to stars.

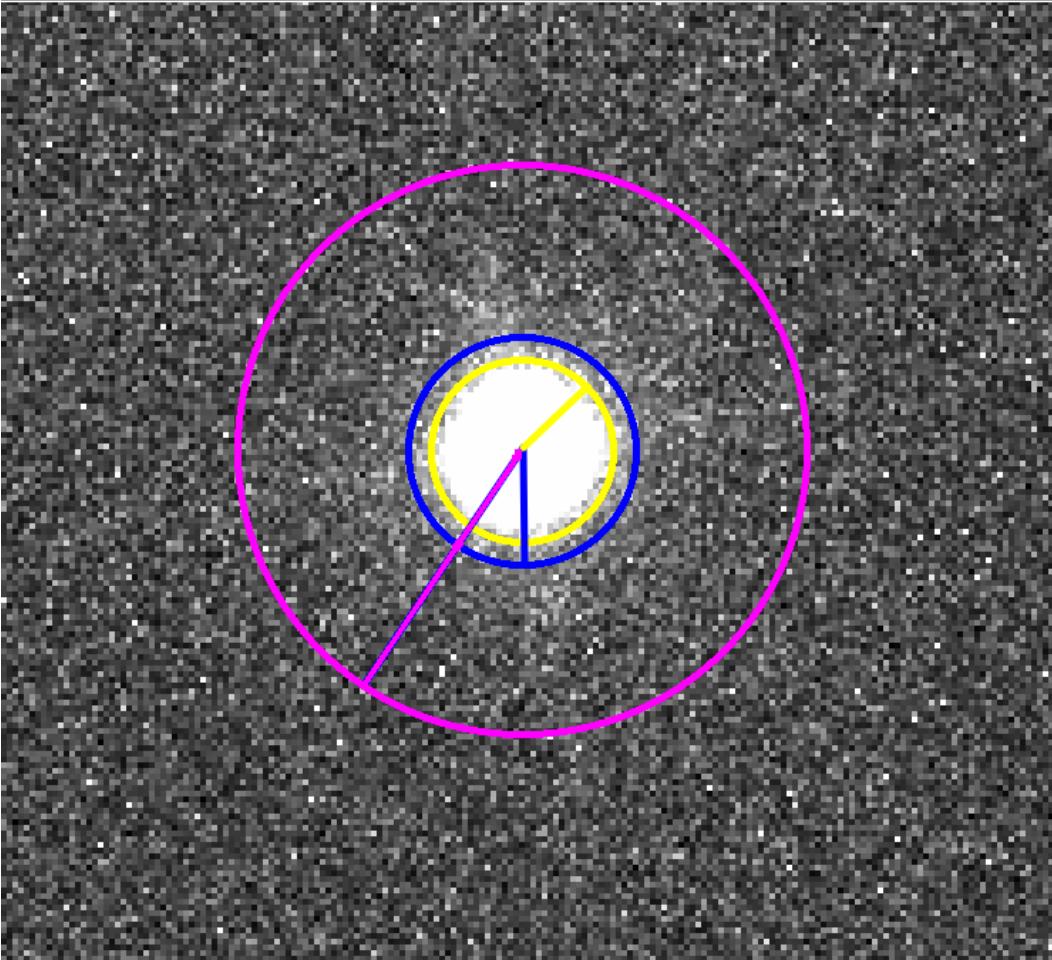
Aperture photometry



This is how a star looks like
on an astronomical image.

How do we perform aperture
photometry?

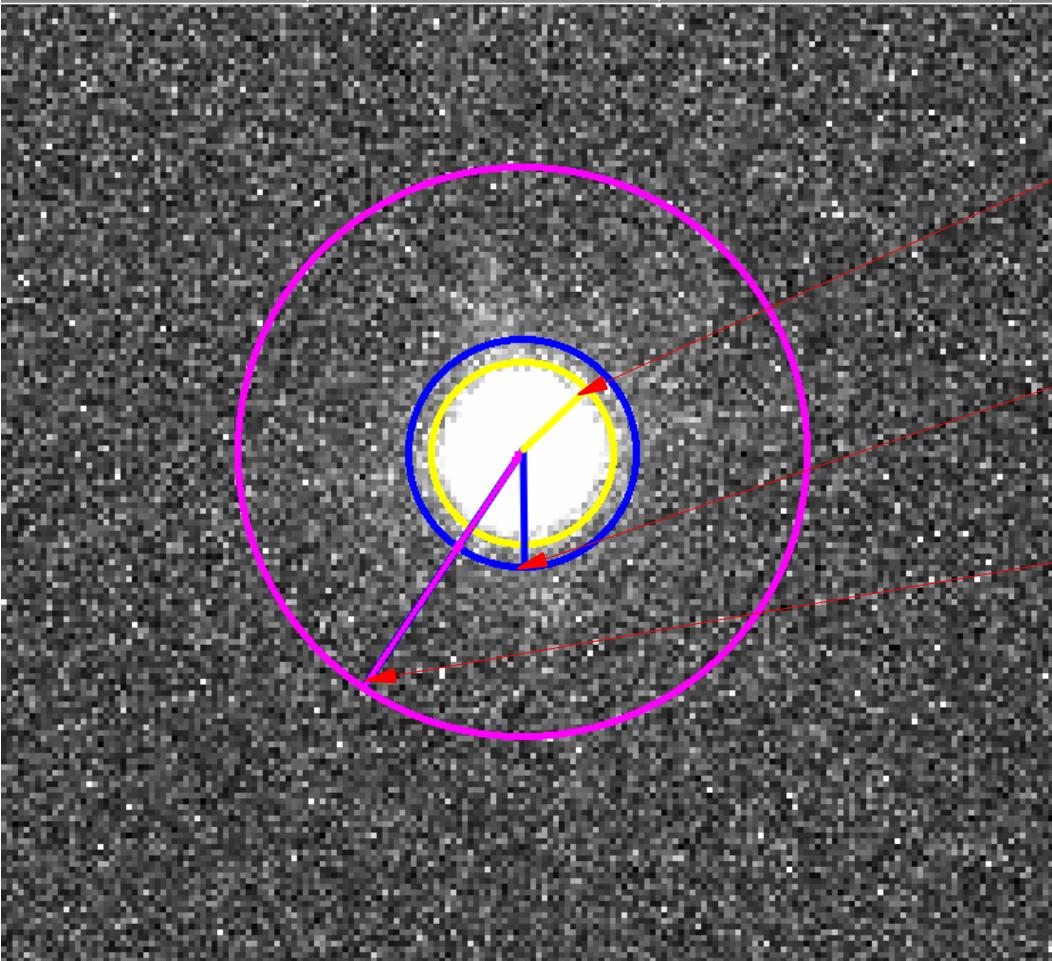
Aperture photometry



First of all, we need to define the apertures; three of them to be precise.

When applied to stars, apertures are, usually, circular.

Aperture photometry



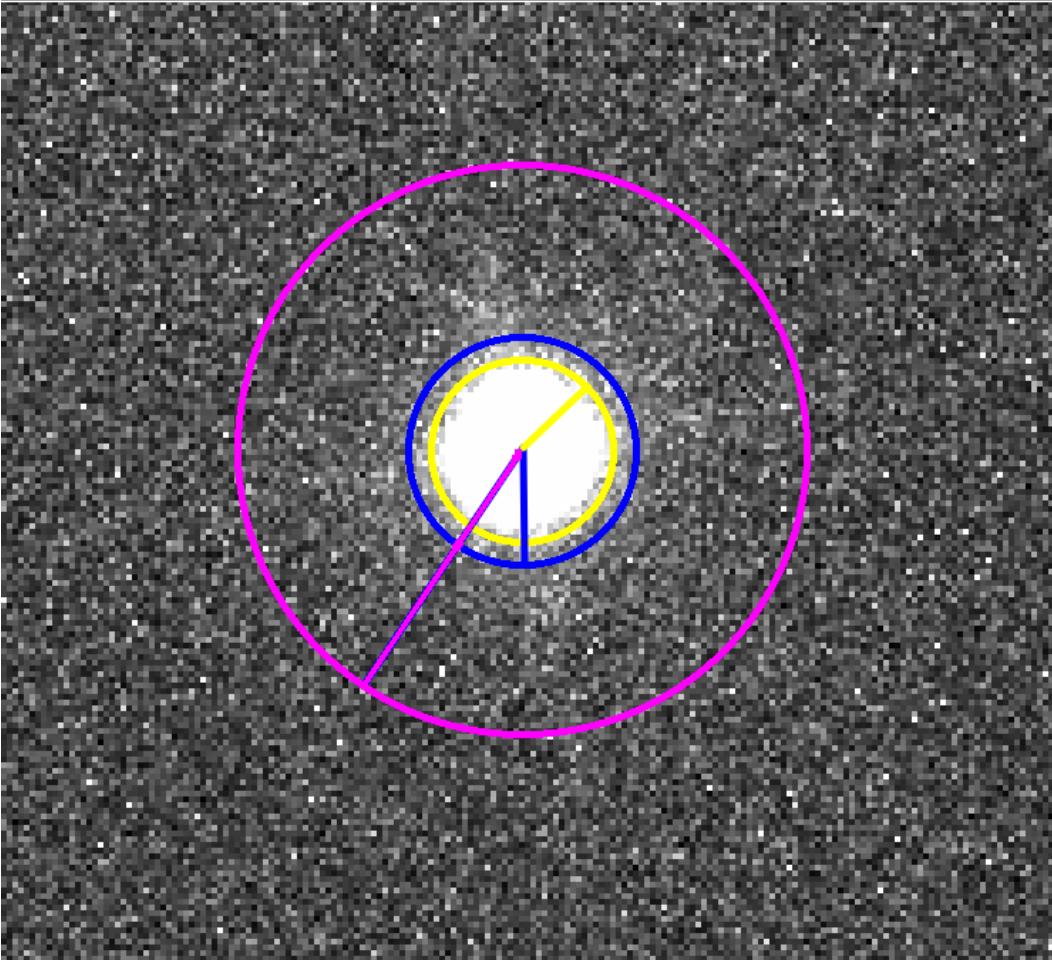
Object aperture around our actual object of interest

Inner sky annulus

Outer sky annulus

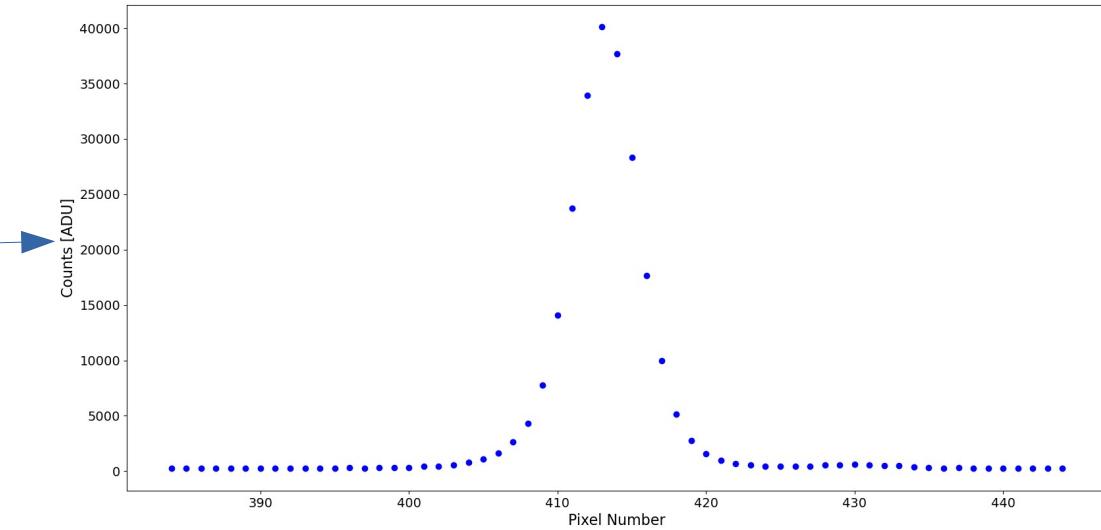
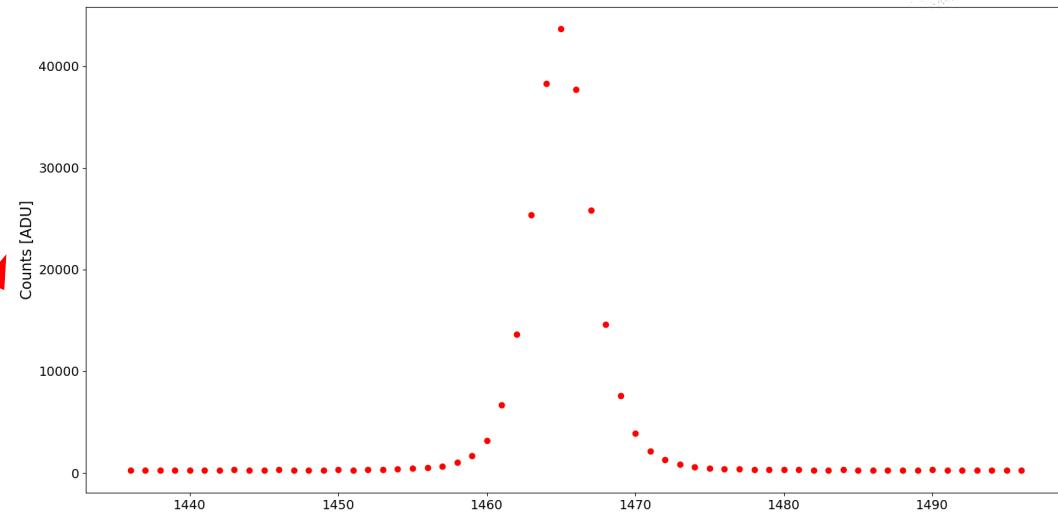
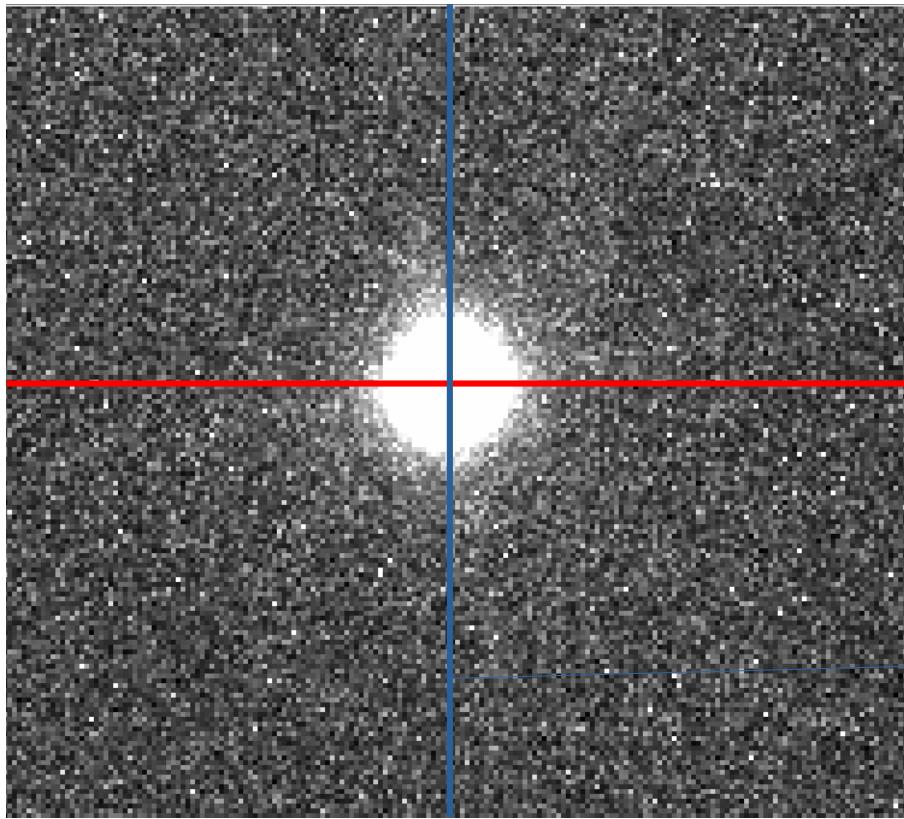
The two annuli define a ring, where we can estimate the contribution of the sky background

Aperture photometry

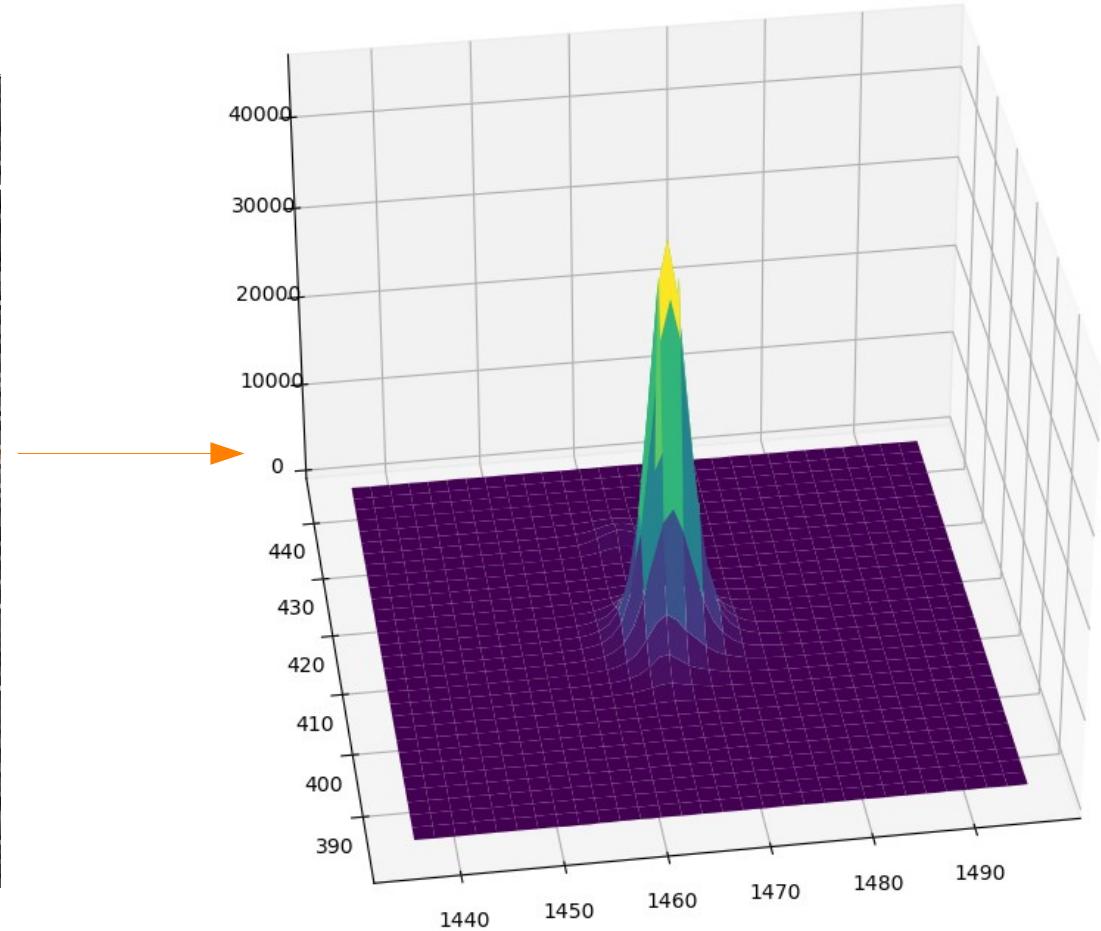
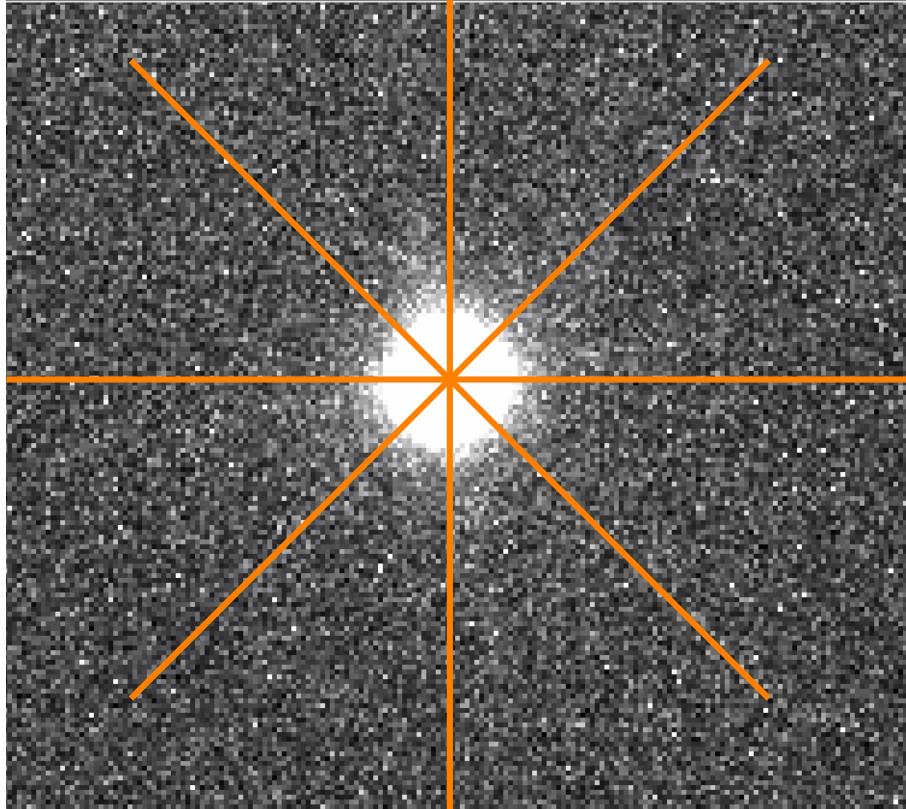


But how can we best define the apertures?

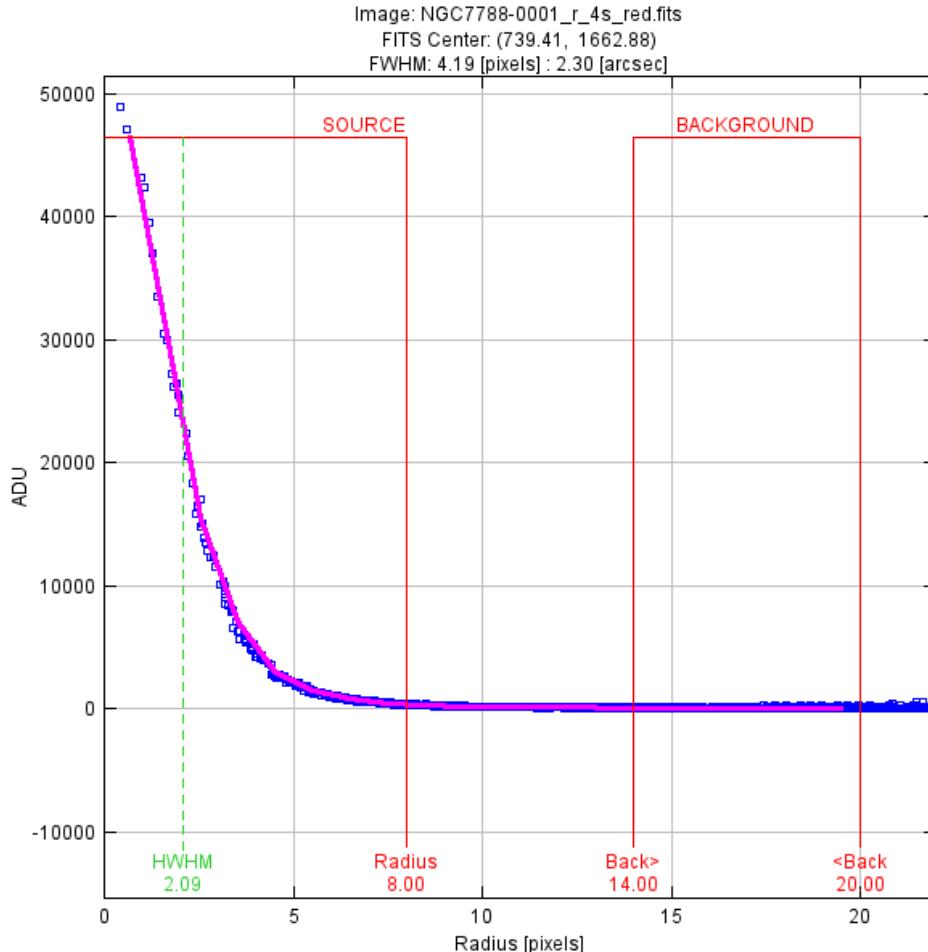
Stellar profiles - 2D



Stellar profiles - 3D

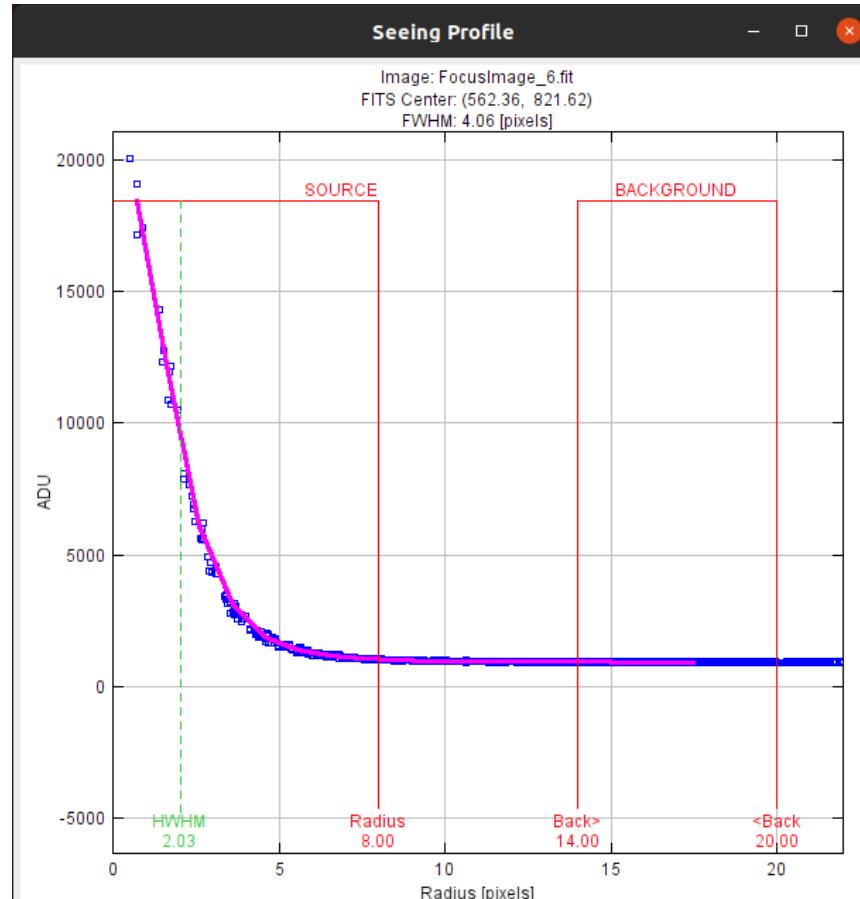
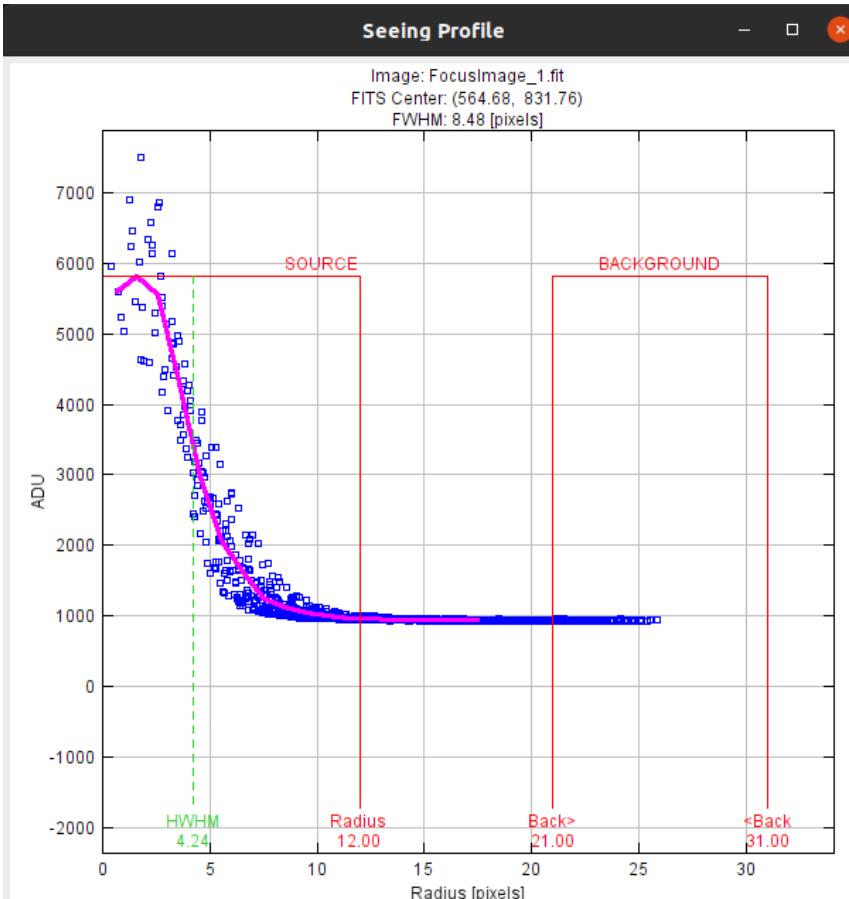


Stellar profiles - radial



We can use the profile of stars on our science image to better define the required apertures.

Stellar profiles - radial



The Airy disc

Stars are point sources,
but on an image they ■
clearly have “dimensions”

Diffraction

Diffraction by the telescope optics causes the “point” to be spread out to a disc, called the Airy disc.

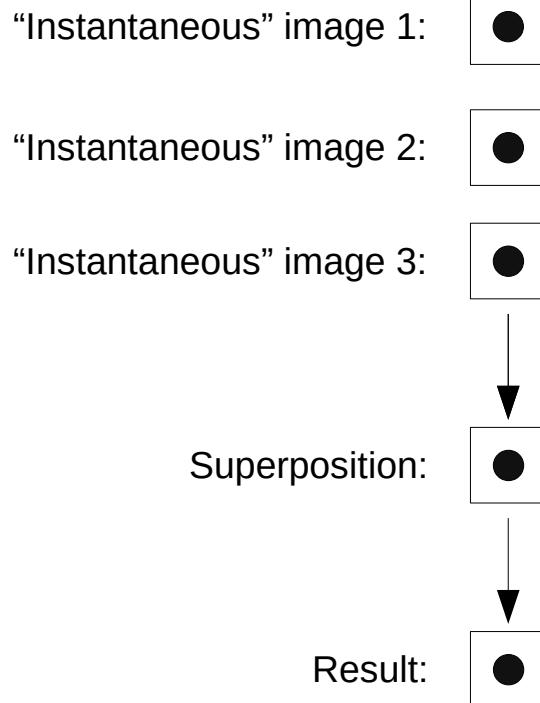


The radius of the Airy disc depends on the wavelength λ } of the incoming light and the diameter of the telescope D } $a = 1.22 \frac{\lambda}{D}$ [rad]

- › For a telescope with $D = 0.8\text{m}$ and $\lambda = 400\text{ nm}$ —► $\alpha = 6.1 \cdot 10^{-7} [\text{rad}] = 0.126 [\text{arcsec}]$
- › For a telescope with $D = 0.8\text{m}$ and $\lambda = 900\text{ nm}$ —► $\alpha = 1.4 \cdot 10^{-6} [\text{rad}] = 0.283 [\text{arcsec}]$

Astronomical seeing

NO ATMOSPHERE

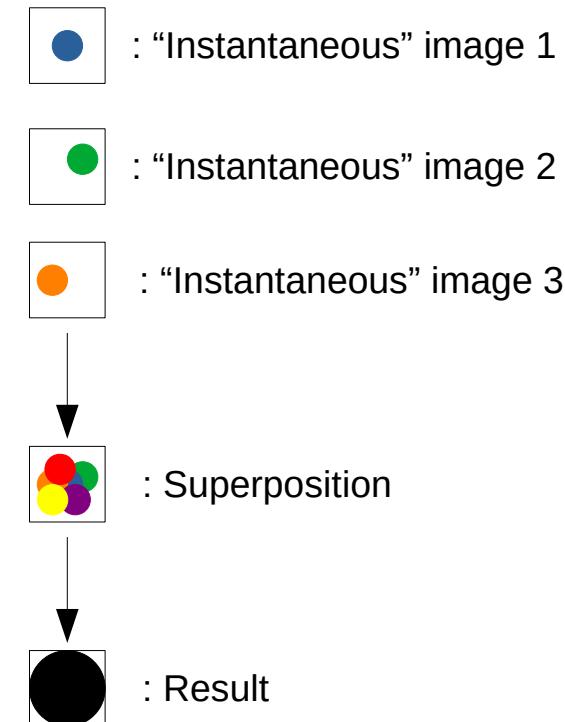


As the star light passes through the atmospheric layers, turbulence in these layers causes the Airy disc to shift position and move around.

The end result of an exposure is a larger stellar profile on our image.

We usually represent the amount of turbulence in the atmosphere by a parameter called ***seeing***, measured in arcseconds.

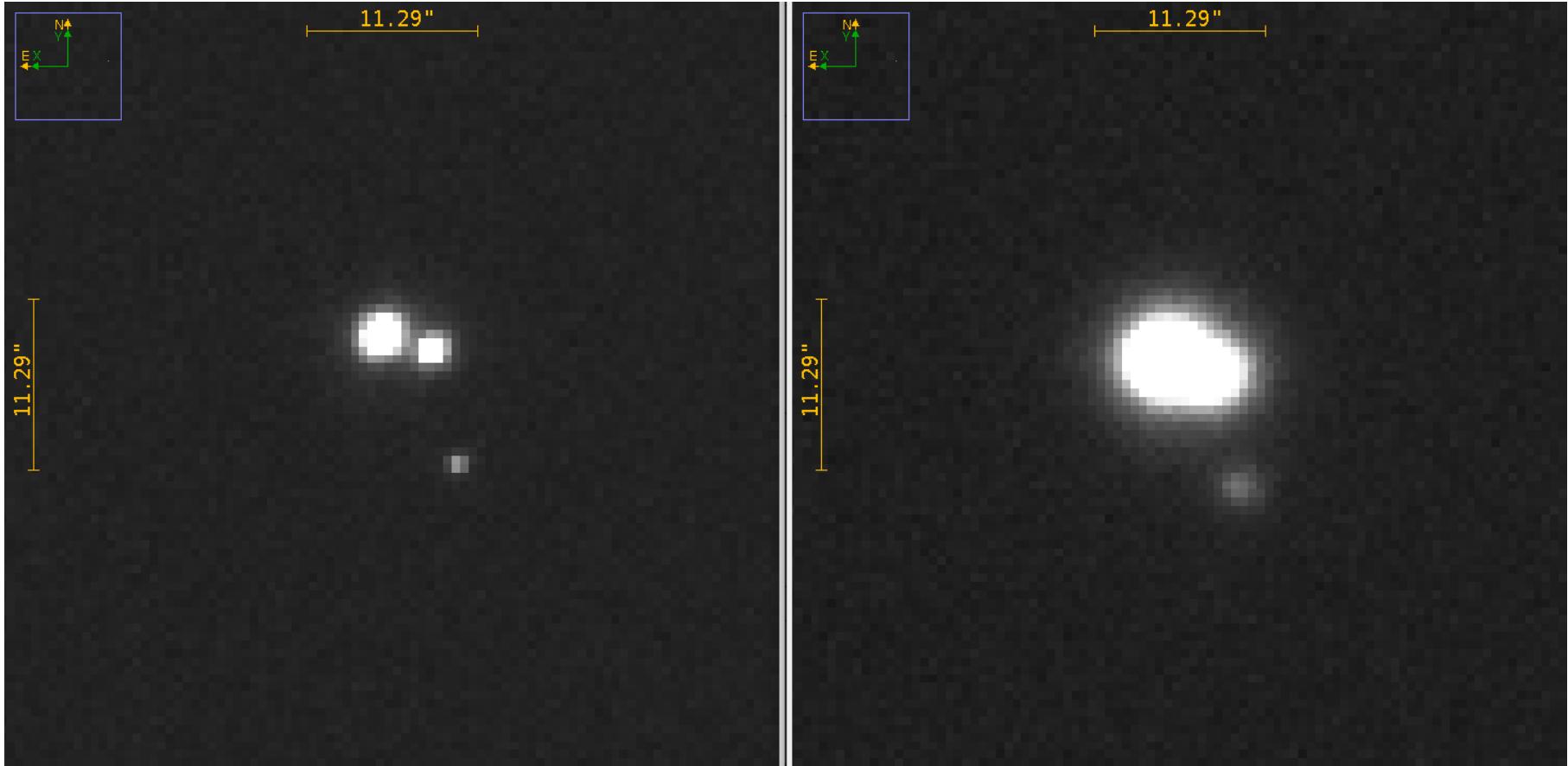
WITH ATMOSPHERE



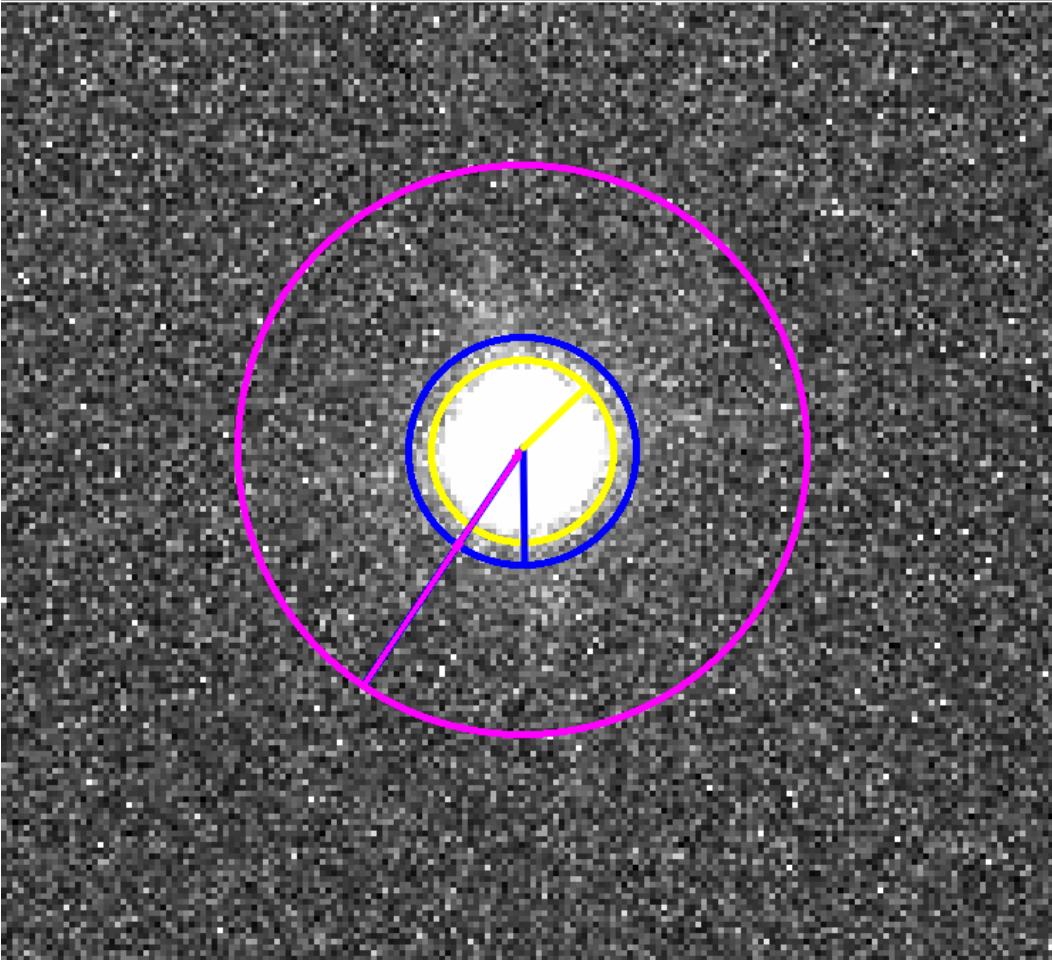


GALACTICA

Astronomical seeing

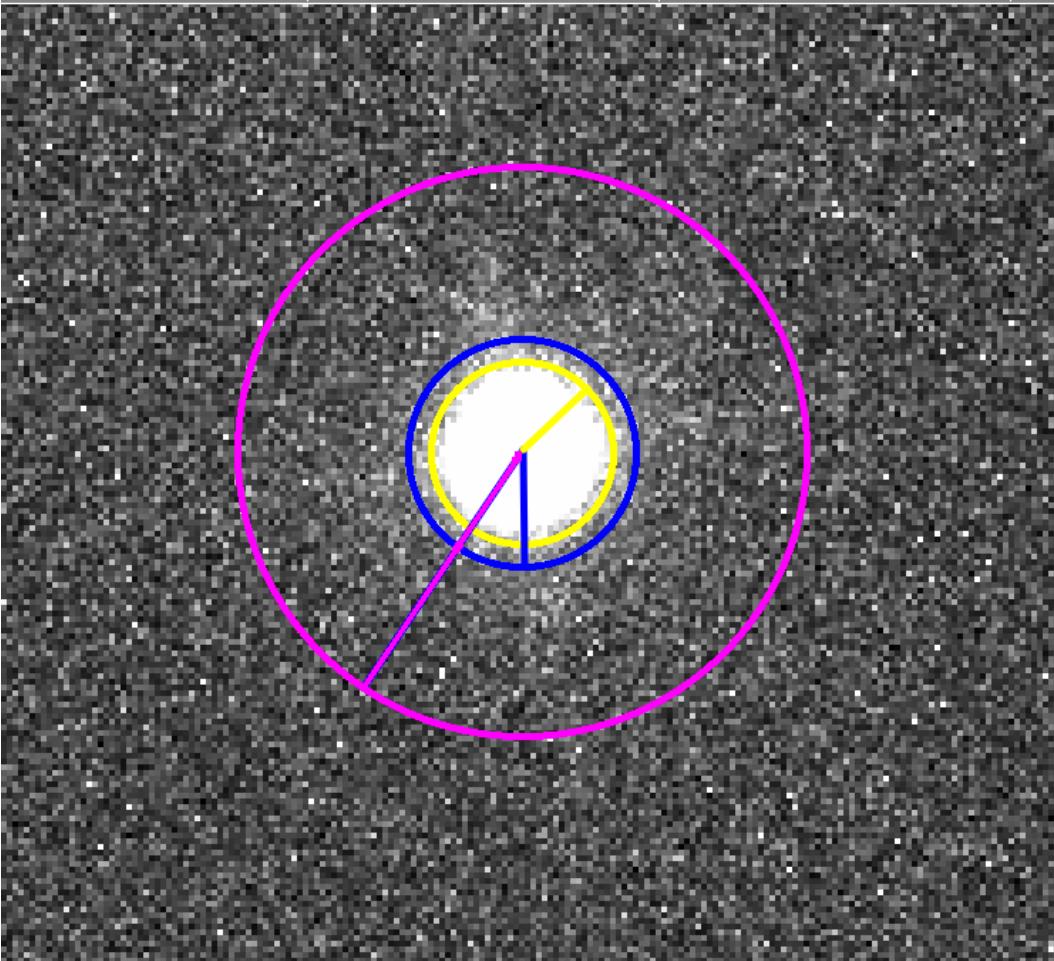


Aperture photometry



Once the apertures have been defined,
we can proceed with the photometry

Aperture photometry



STEP 1: Raw object flux, F_{raw}

The sum of the ADU count values of all the pixels contained within the object aperture

STEP 2: Sky flux, F_{sky} [pix⁻¹]

The sky flux is estimated per pixel – it's the sum of the ADU values of all the pixels contained between the inner and outer sky annuli, divided by the total number of these same pixels

STEP 3: Final object flux, F_{obj}

$$F_{\text{obj}} = F_{\text{raw}} - N_{\text{pix}} * F_{\text{sky}}$$

where N_{pix} is the number of pixels contained in the object aperture

Signal to Noise Ratio

Every measurement comes with an associated error!

A simple way to quantify the quality of our photometry, is to use the ***signal to noise ratio***

$$SNR = \frac{\text{Signal}}{\text{Noise}}$$

For CCD photometry, the fastest way to calculate the SNR is to use the expression:

$$SNR = \frac{\text{Total counts}}{\sqrt{\text{Total counts}}}$$

Note that the above expression is *only* a quick estimate and, technically, it is *not exactly* correct!

Actual Signal to Noise Ratio (Informative only)

$$SNR = \frac{S_{obj} \cdot t_{exp} \cdot g_{CCD}}{\sqrt{S_{obj} \cdot t_{exp} \cdot g_{CCD} + S_{sky} \cdot t_{exp} \cdot g_{CCD} \cdot n_{pix} + S_{dark} \cdot t_{exp} \cdot n_{pix} + R^2 \cdot n_{pix}}}$$

- S_{obj} = object signal in ADU/sec
- S_{sky} = sky signal in ADU/sec/pix
- S_{dark} = dark current signal in e⁻/sec/pix
- R = readout noise in e⁻/pix
- n_{pix} = number of pixels our star occupies
- t_{exp} = exposure time in sec
- g_{CCD} = the CCD gain in e⁻/ADU

It is necessary to use the gain and convert from ADU to electrons, because (photo)electrons follow the Poisson statistics, and ADUs do not! With Poisson statistics we can express the noise N as the square root of the signal S:

$$SNR = \frac{S}{N} = \frac{S}{\sqrt{S}}$$

From counts to magnitudes

Instrumental magnitudes

At the end of the process, we have the total sum of the ADU counts of our object, F_{obj}

How do we convert these ADU counts into something more meaningful, e.g. magnitudes?

- First, we convert them into an internal scale, called *instrumental magnitude*, m_{inst}

$$m_{\text{inst}} = - 2.5 * \log_{10} \left(\frac{F_{\text{obj}}}{t_{\text{exp}}} \right)$$

Example 1: $F_{\text{obj}} = 10,000 \text{ ADU}$ & $t_{\text{exp}} = 10 \text{ sec} \rightarrow m_{\text{inst}} = -7.5$

Example 2: $F_{\text{obj}} = 1,000,000 \text{ ADU}$ & $t_{\text{exp}} = 10 \text{ sec} \rightarrow m_{\text{inst}} = -12.5$

Zeropoint magnitude

The instrumental magnitudes are on an arbitrary scale that depends on the instrument and the observing system in general. We **cannot** compare instrumental magnitudes of different instruments/telescopes!

This means that we need to convert our instrumental magnitudes into apparent magnitudes, for our measurements to be useful to other astronomers as well!

- To do that, we need to establish the **zeropoint magnitude**, m_{zp}

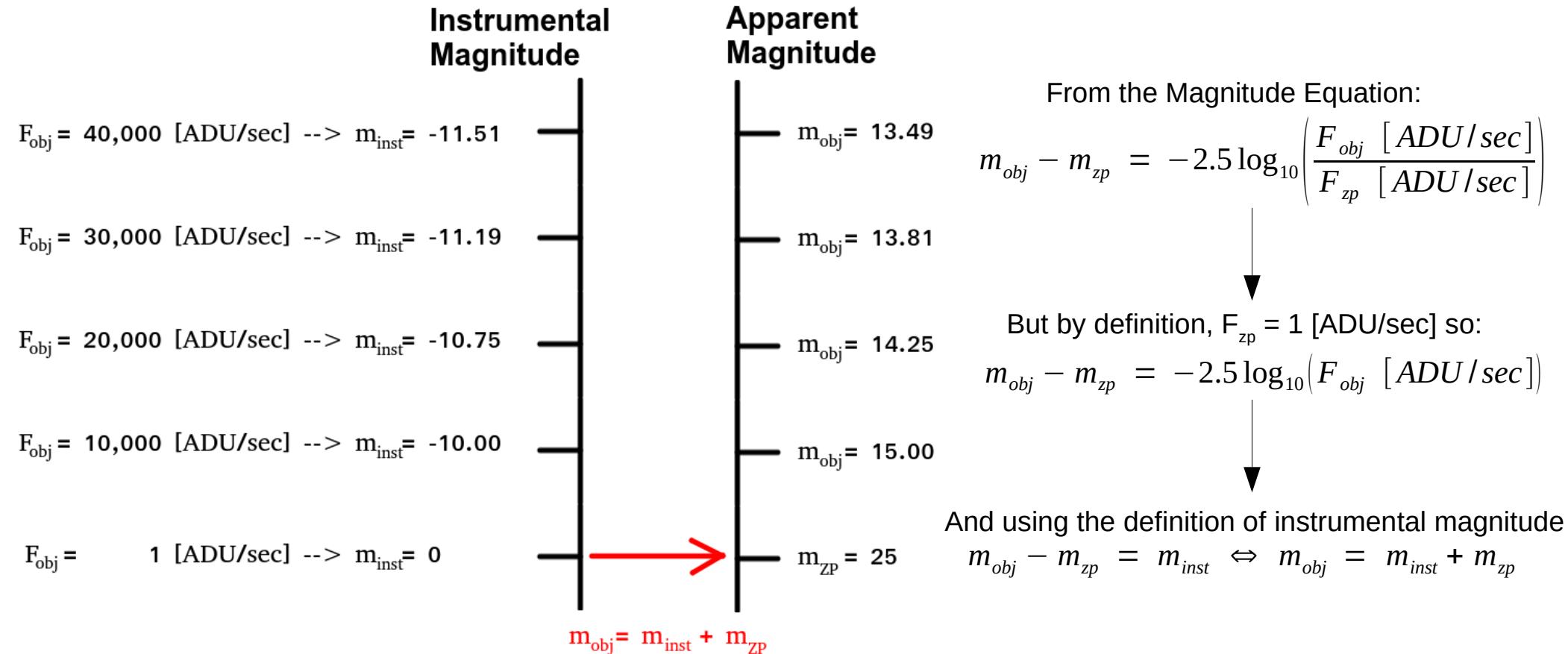
Definition:

m_{zp} is the apparent magnitude of a star that, when observed with our CCD, gives **1 ADU/sec.**

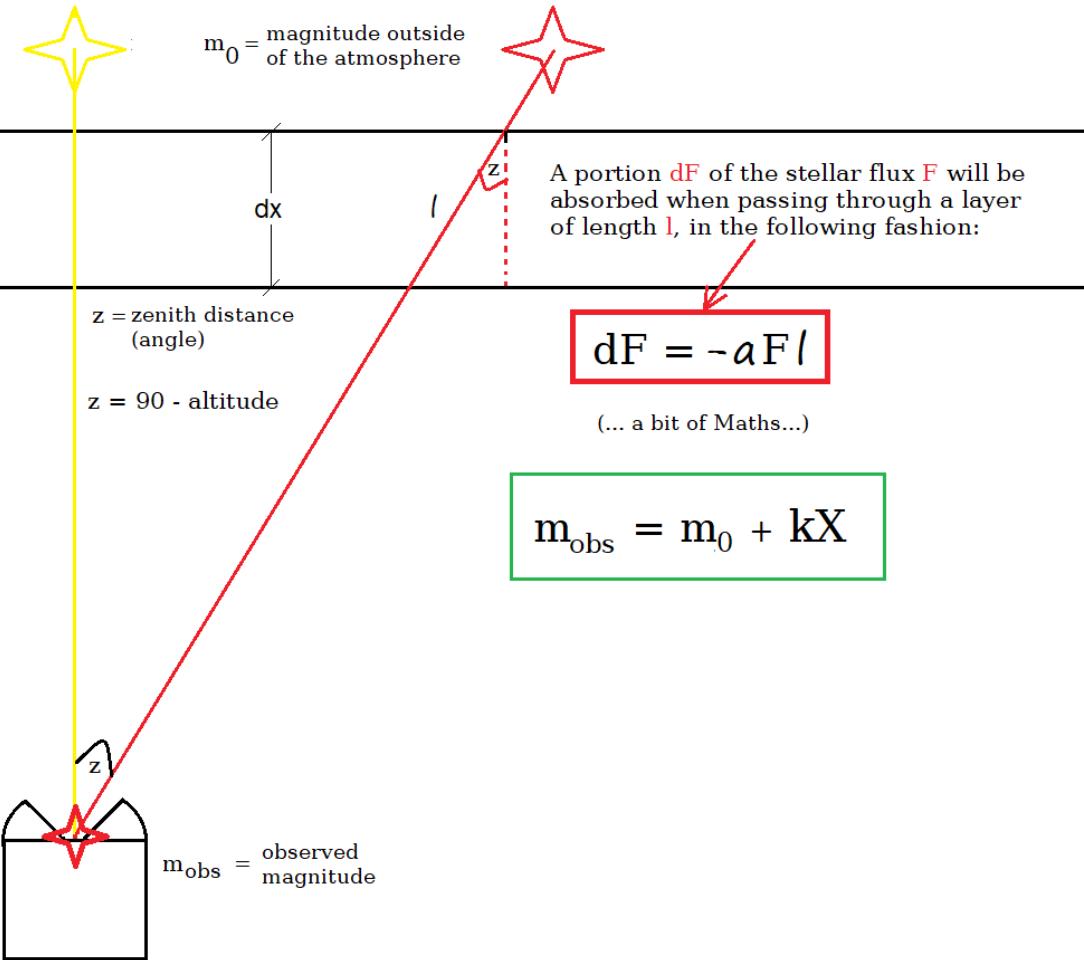
It is the apparent magnitude of a star that, when observed with our CCD, has instrumental magnitude $m_{\text{inst}} = 0$.

Converting instrumental to apparent magnitude

Once the zeropoint magnitude is known, we convert with the equation $m_{\text{obj}} = m_{\text{inst}} + m_{\text{zp}}$



Measuring the zero point - Definitions



Proof

$$dF = -aFl \Leftrightarrow \frac{dF}{F} = -a \frac{dx}{\cos(z)} \Leftrightarrow \int_0^{obs} \frac{dF}{F} = -\frac{1}{\cos(z)} \int_0^{obs} a dx \Leftrightarrow$$

① $\cos(z) = \frac{dx}{l}$

$$[\ln(F)]_0^{obs} = -\frac{1}{\cos(z)} \int_0^{obs} a dx \Leftrightarrow \ln(F_{obs}) - \ln(F_0) = -\frac{1}{\cos(z)} \int_0^{obs} a dx \quad \textcircled{2}$$

② $\ln(A) - \ln(B) = \ln\left(\frac{A}{B}\right)$

$$\ln\left(\frac{F_{obs}}{F_0}\right) = -\frac{1}{\cos(z)} \int_0^{obs} a dx \quad \textcircled{3} \quad \frac{\log_{10}\left(\frac{F_{obs}}{F_0}\right)}{\log_{10}(e)} = -\frac{1}{\cos(z)} \int_0^{obs} a dx \Leftrightarrow$$

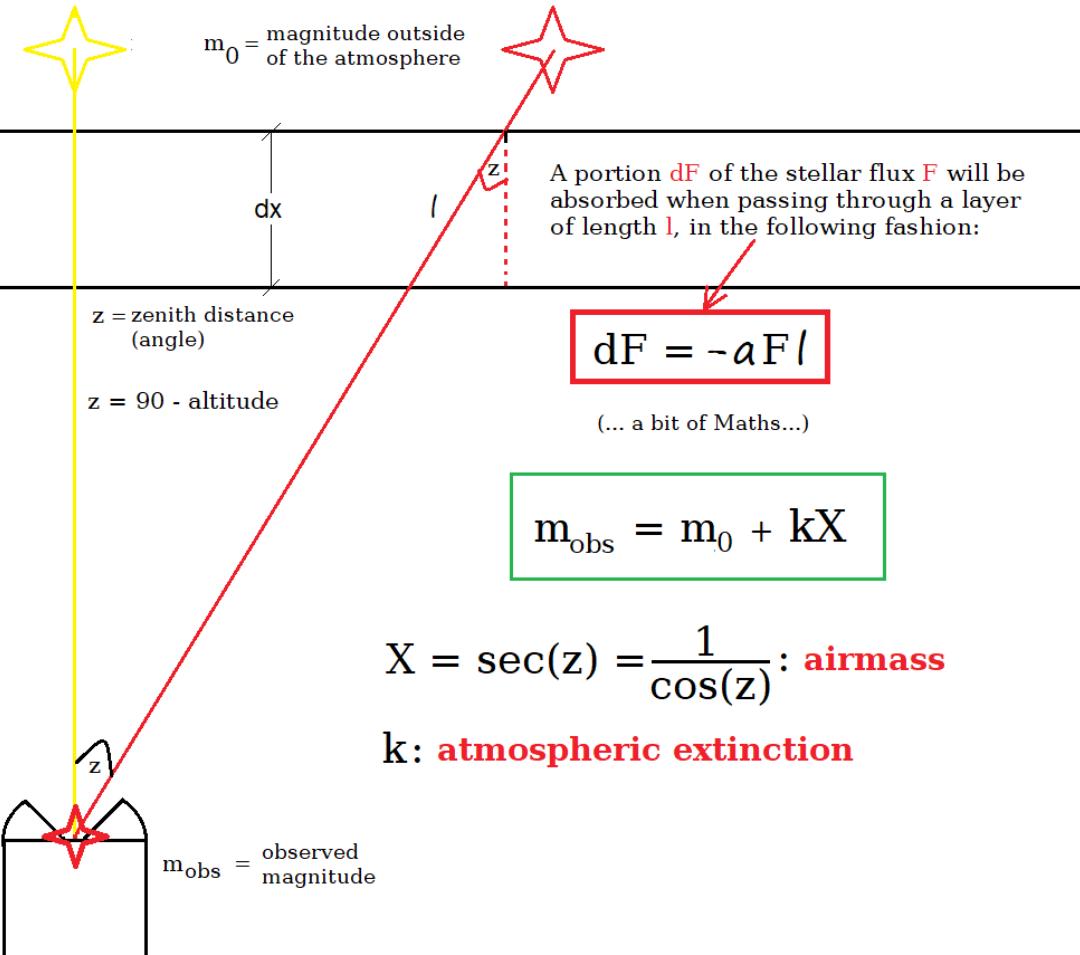
③ $\ln(A) = \frac{\log_{10}(A)}{\log_{10}(e)}$

$$\log_{10}\left(\frac{F_{obs}}{F_0}\right) = -\log_{10}(e) \frac{1}{\cos(z)} \int_0^{obs} a dx \quad \textcircled{4} \quad \begin{array}{c} \text{Magnitude Equation} \\ -2.5 \log_{10}\left(\frac{F_{obs}}{F_0}\right) = \end{array} \quad \begin{array}{c} \text{airmass} \\ \frac{1}{\cos(z)} * \end{array} \quad \begin{array}{c} \text{atmospheric extinction} \\ 2.5 \log_{10}(e) \int_0^{obs} a dx \end{array} \Leftrightarrow$$

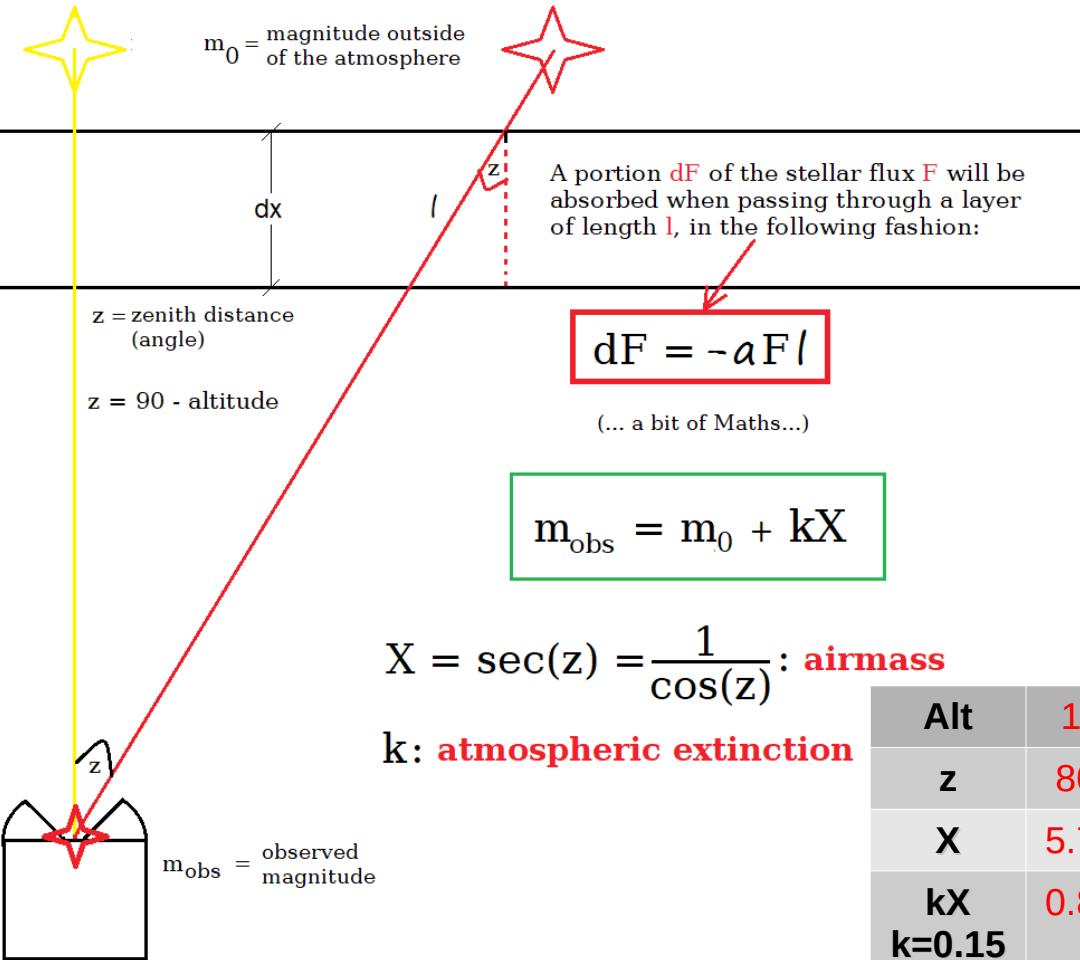
④ Multiply both sides by -2.5

$$m_{obs} - m_0 = k * X$$

Measuring the zero point - Definitions



Measuring the zero point - Definitions



Alt	10	20	30	40	50	60	70	80
z	80	70	60	50	40	30	20	10
X	5.76	2.92	2.00	1.56	1.31	1.15	1.06	1.02
kX $k=0.15$	0.86	0.44	0.30	0.23	0.20	0.17	0.16	0.15

Measuring the zeropoint

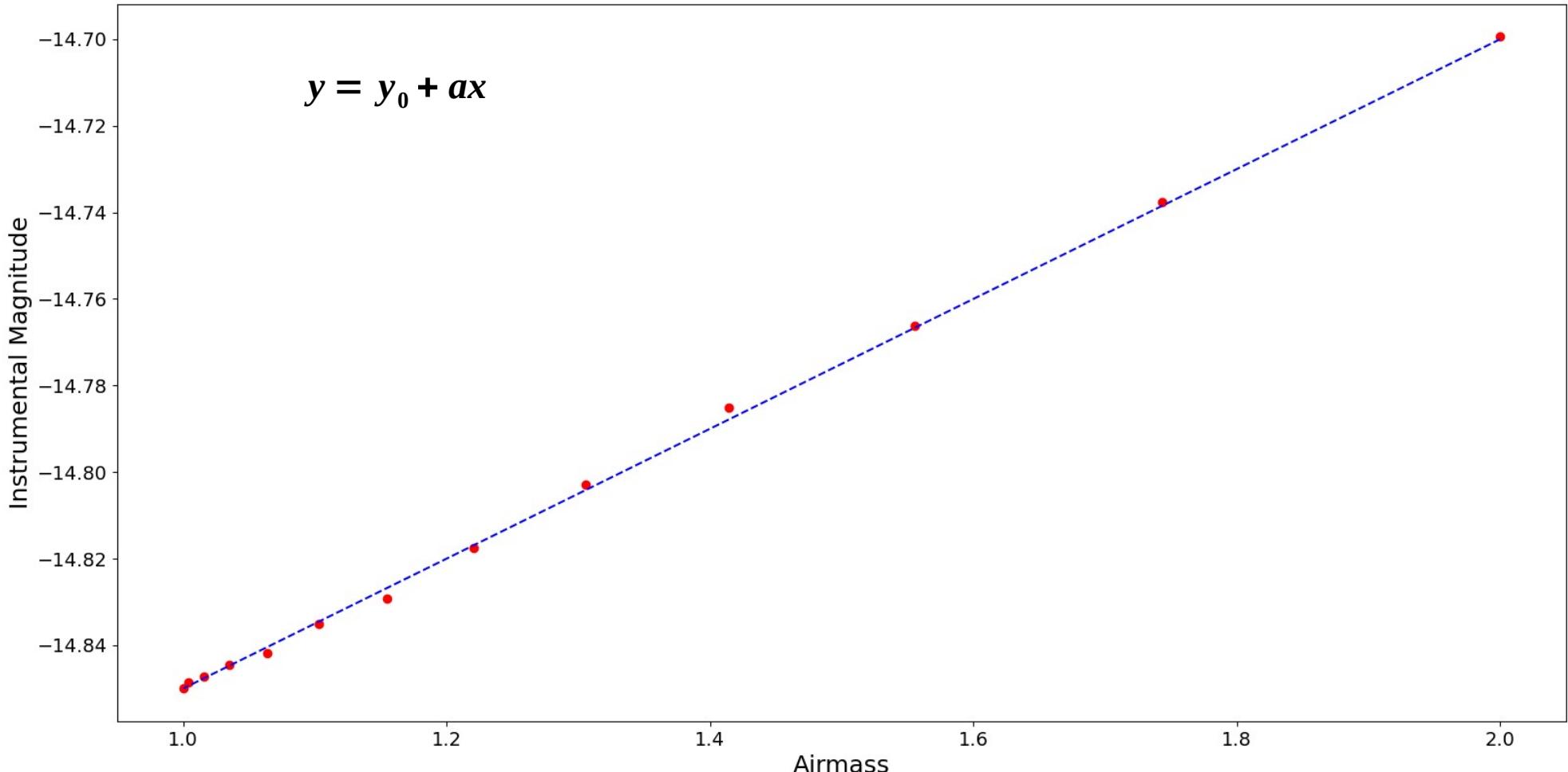
In the past, to measure the zeropoint magnitude, astronomers used ***standard stars*** (also called *photometric* and/or *spectrophotometric* standard stars, SPS stars).

Standard stars are non-variable stars, which have their magnitudes (and fluxes) measured with extremely high precision in most photometric systems (e.g. UBVRI, SDSS, etc...)

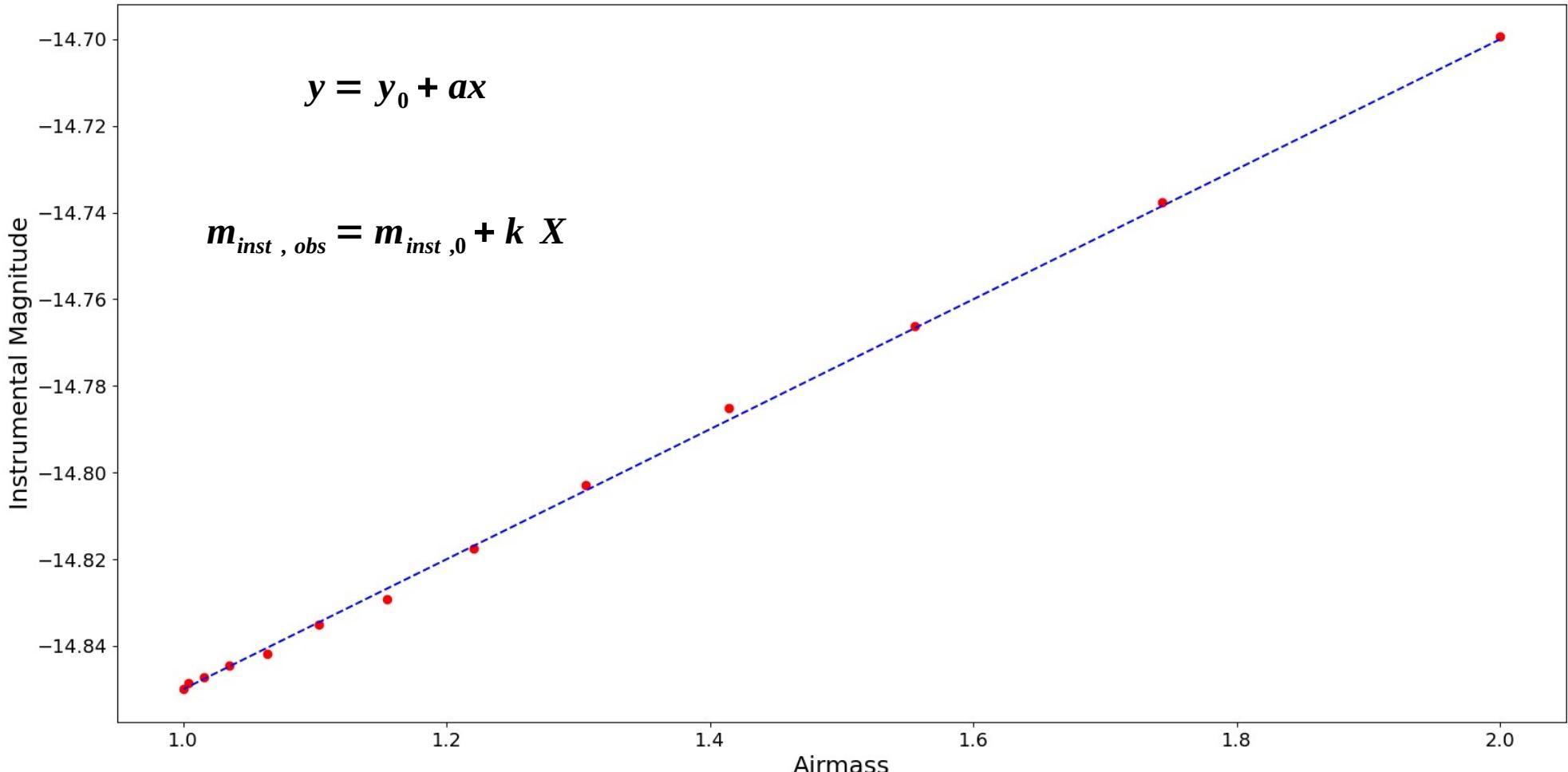
The process included the following steps:

- Observe a standard star throughout the night, at different airmass (= different altitudes)
- Perform photometry and obtain instrumental magnitudes
- Plot airmass versus instrumental magnitude and fit a straight line

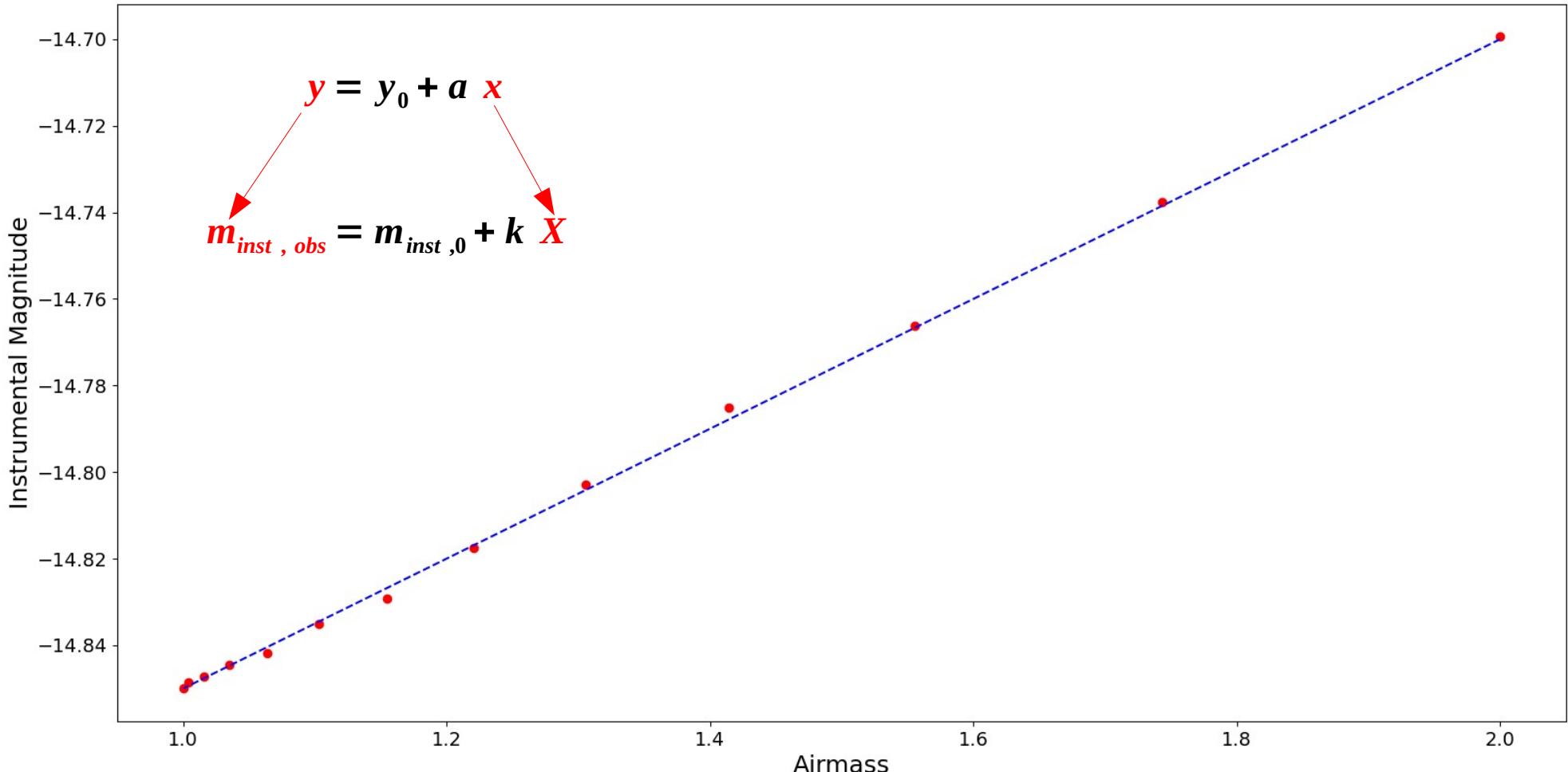
Measuring the zero point



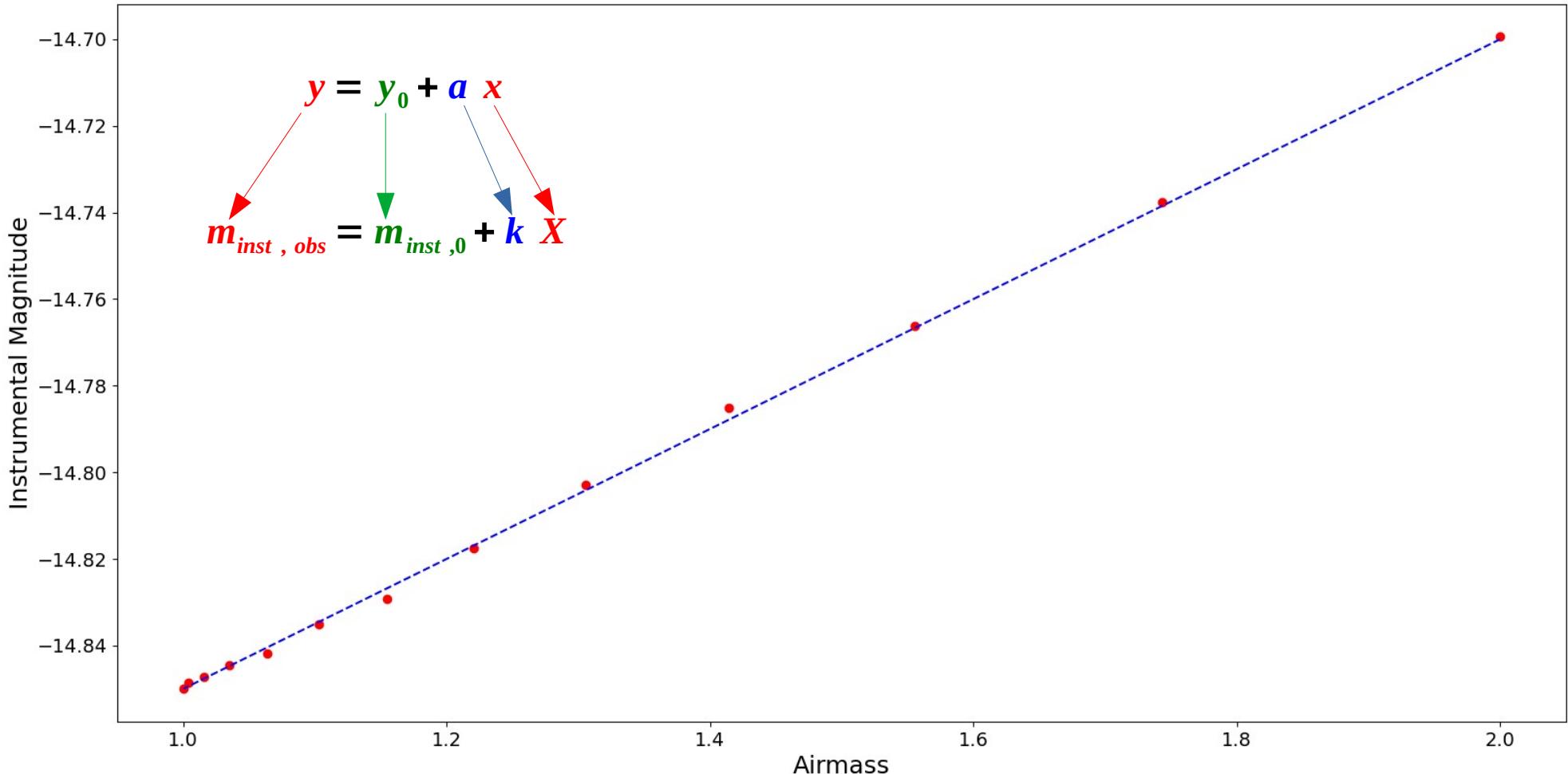
Measuring the zero point



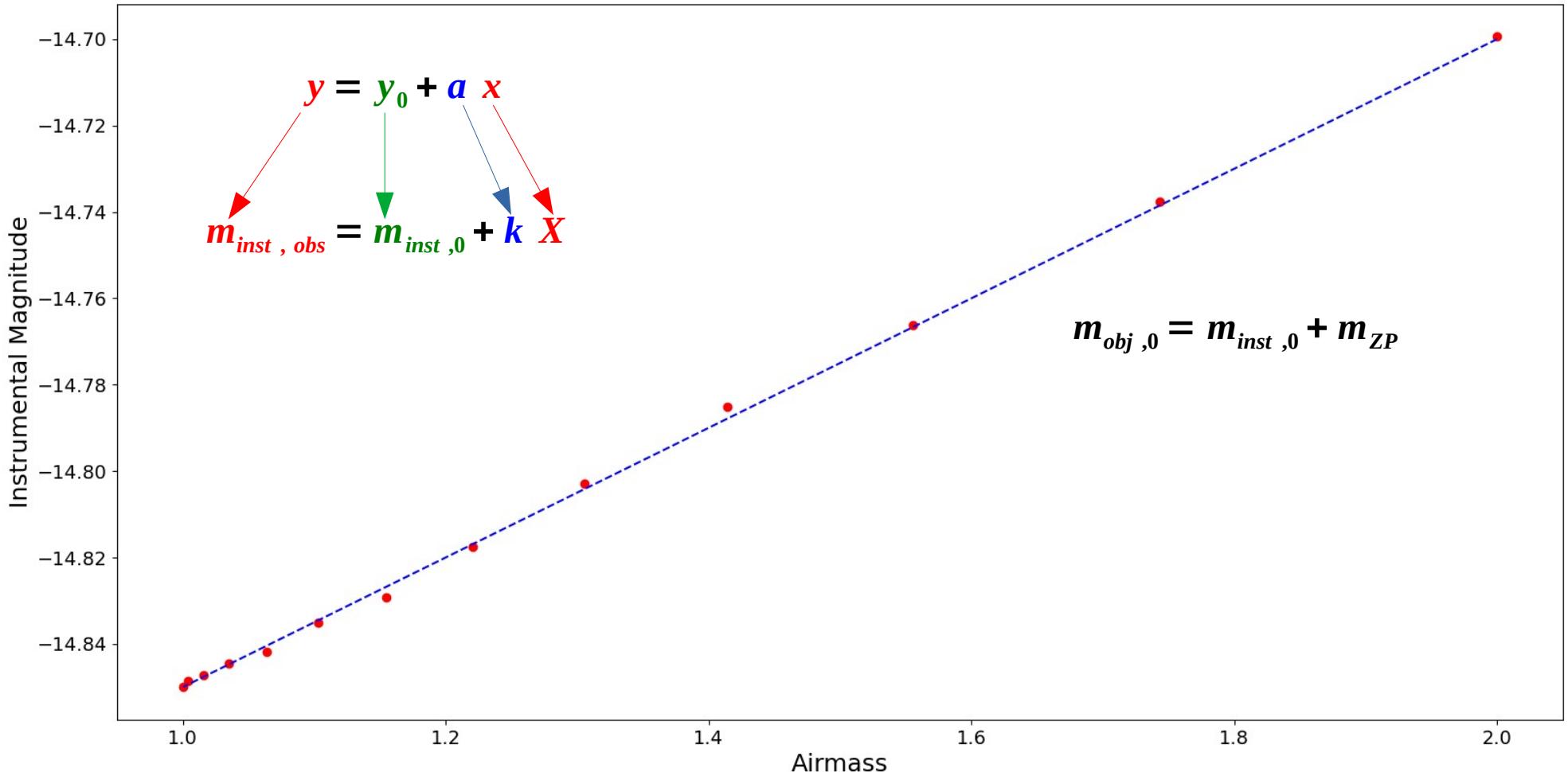
Measuring the zero point



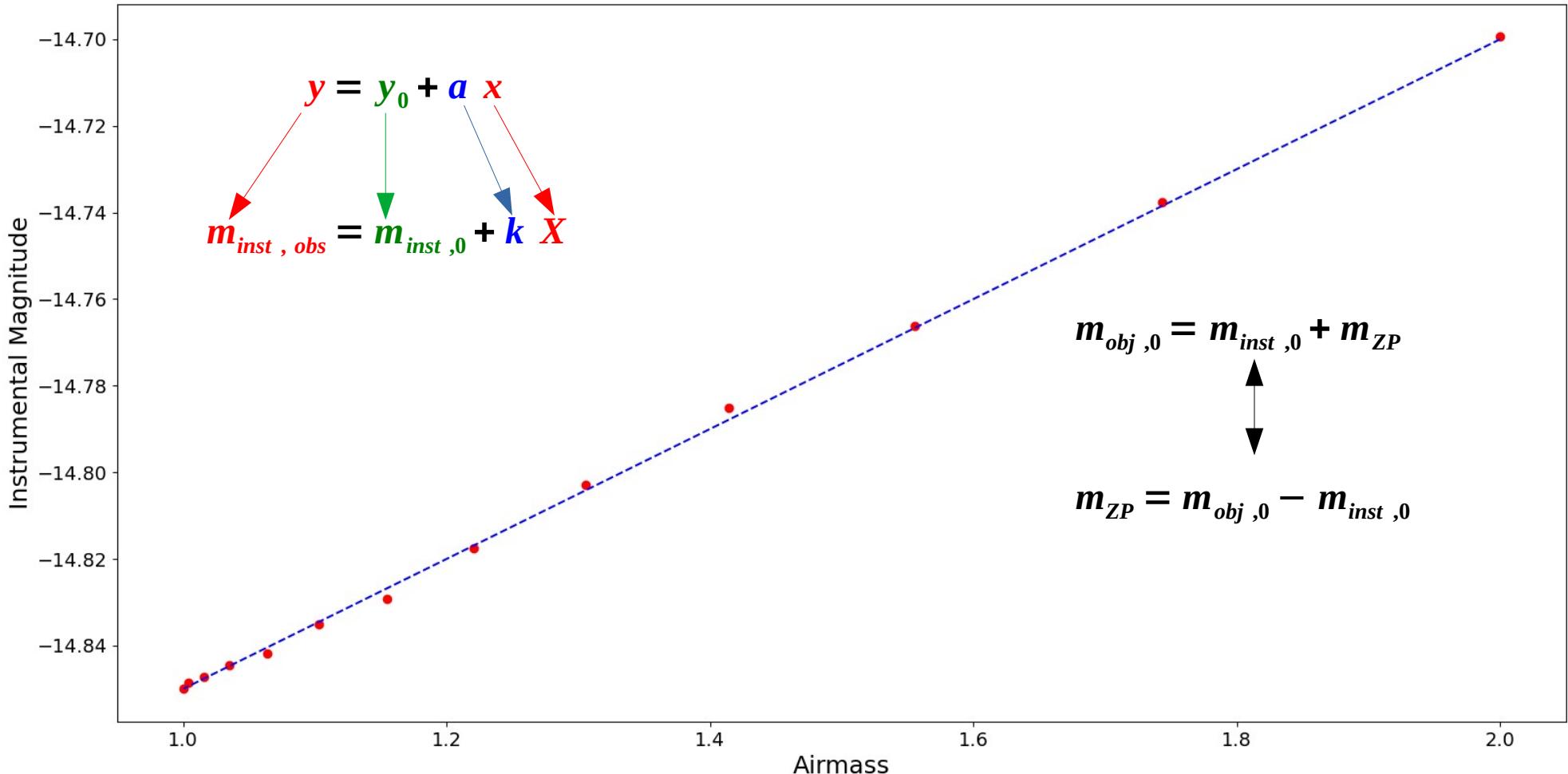
Measuring the zero point



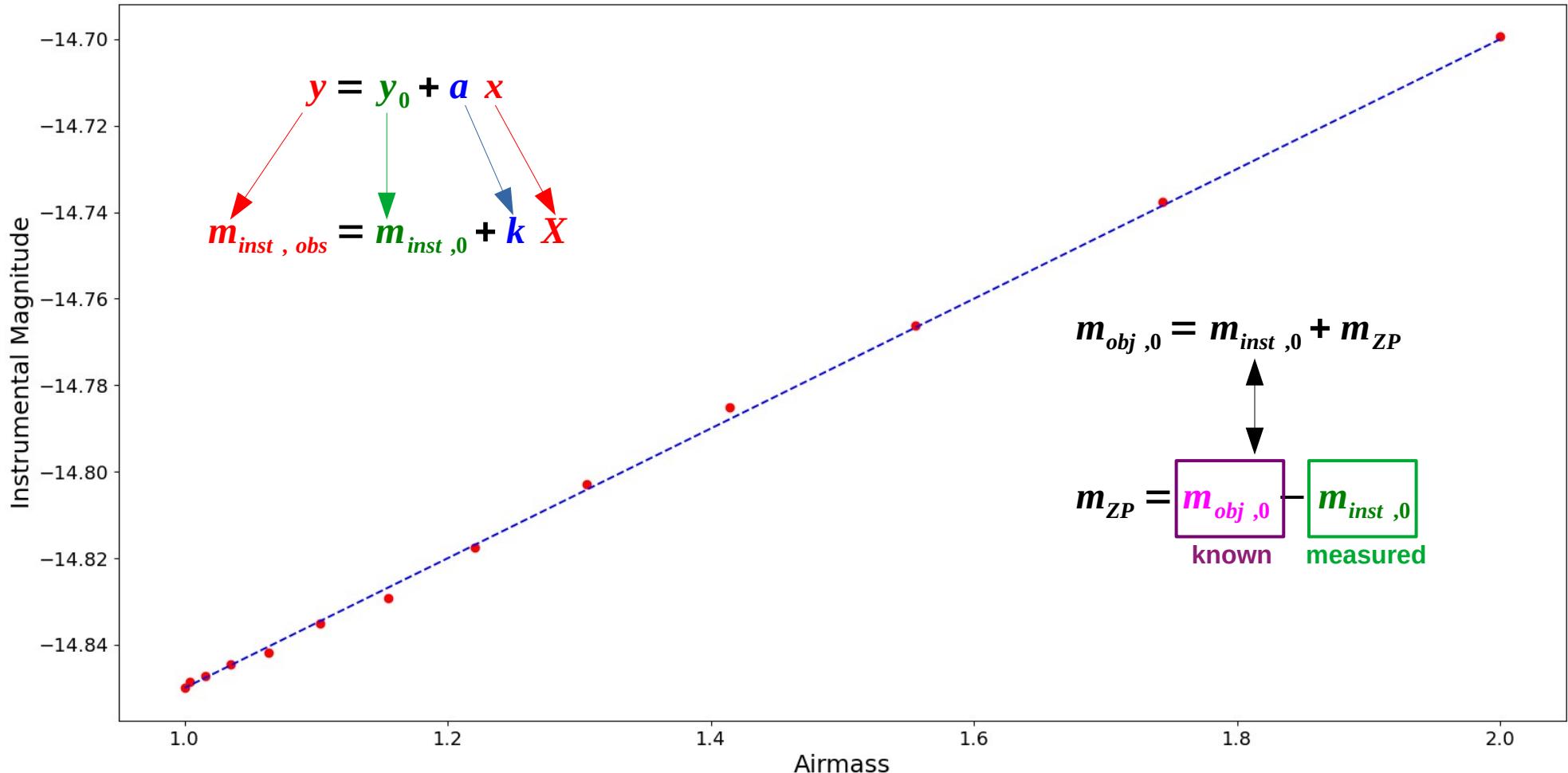
Measuring the zero point



Measuring the zero point



Measuring the zero point



Estimating the zeropoint magnitude

Luckily, we live in the era of *large-scale, all-sky* photometric surveys (e.g. SDSS, PanSTARRS, J-PAS, Gaia etc).

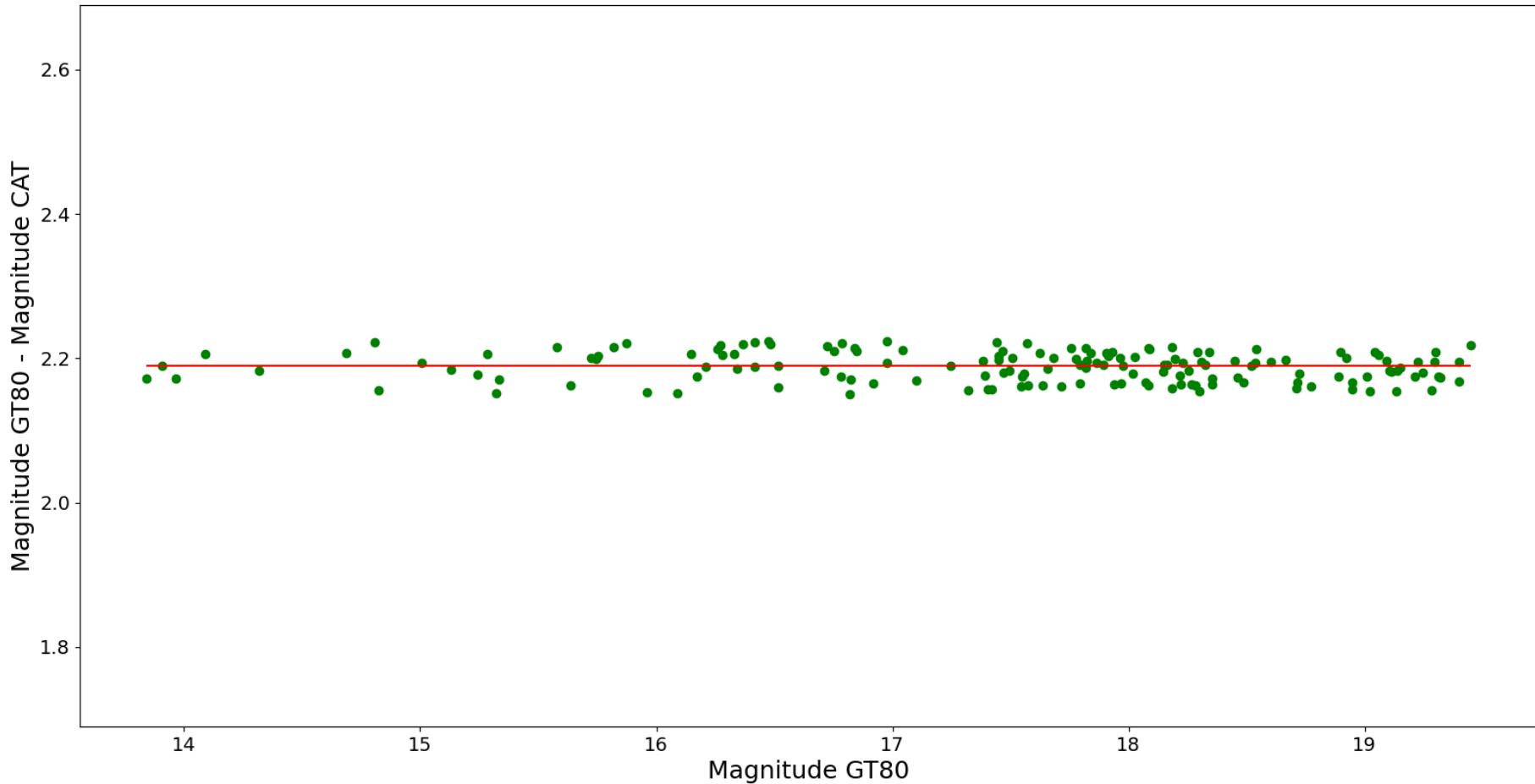
All of these surveys do a very careful job of properly calibrating their photometry. In addition, and most importantly, their data are *open* and *accessible* to everyone!

The idea is simple: use these surveys to calibrate our own photometry, in the following fashion:

- **Step 1:** we perform photometry in the way we have seen and use an **arbitrary** value for m_{zp} , e.g. $m_{zp} = 25$, to obtain apparent magnitudes m_{GT80}
- **Step 2:** we **crossmatch** our observations with one of the surveys and obtain their calibrated apparent magnitudes, m_{CAT}
- **Step 3:** We calculate the difference $\Delta mag = m_{GT80} - m_{CAT}$
- **Step 4:** We repeat the process for many stars, in order to have good statistics, and obtain an average (or median) $\langle \Delta mag \rangle$

Estimating the zeropoint magnitude

We obtain our calibrated photometry with the equation: $m_{\text{GT80,calib}} = m_{\text{GT80,org}} - \langle \Delta \text{mag} \rangle$



SUMMARY

- We saw how we can combine our calibration frames to create the master frames. We went over a few considerations for each type of calibration frame (bias, dark, flat). We also saw the series of steps needed to reduce our data.
- We saw how to define apertures and perform aperture photometry.
- We saw that, in order to convert our measurements from ADU counts to apparent magnitudes we need to know the zeropoint magnitude m_{zp} . We also saw two methods of getting the m_{zp} : (i) directly, with observations of standard stars; and (ii) indirectly, by using external catalogues.



GALACTICA

APPENDIX

Averaging Frames with rejection #1 - MinMax

- Usually, an average combination also uses some form of rejection algorithm.
- One such algorithm is **MinMax [n1,n2]**:
 - Original values: 905, 909, 915, 907, 908, 891, 911, 2587, 902, 911, 905
 - Ordered values: 891, 902, 905, 905, 907, 908, 909, 911, 911, 915, 2587
 - MinMax [0,1]: **891, 902, 905, 905, 907, 908, 909, 911, 911, 915, 2587** → **906.4**
 - MinMax [1,1]: **891, 902, 905, 905, 907, 908, 909, 911, 911, 915, 2587** → **908.1**

Averaging Frames with rejection #2 - SigmaClip

- Usually, an average combination also uses some form of rejection algorithm.
- Another such algorithm is **SigmaClip [sLow,sHigh]**:
 - Original values: 905, 909, 915, 907, 908, 891, 911, 2587, 902, 911, 905
 - Calculate mean and standard deviation: 1059.18, 483.17
 - SigmaClip [1,1]: only accept $\text{Mean} - 1 \times \text{StdDev} < \text{values} < \text{Mean} + 1 \times \text{StdDev}$, which in our case is: **576.01 < values < 1542.36**, so

905, 909, 915, 907, 908, 891, 911, 2587, 902, 911, 905 → 906.4