

# Data Management for Ground-Based Science Surveys at the Cambridge Astronomical Survey Unit



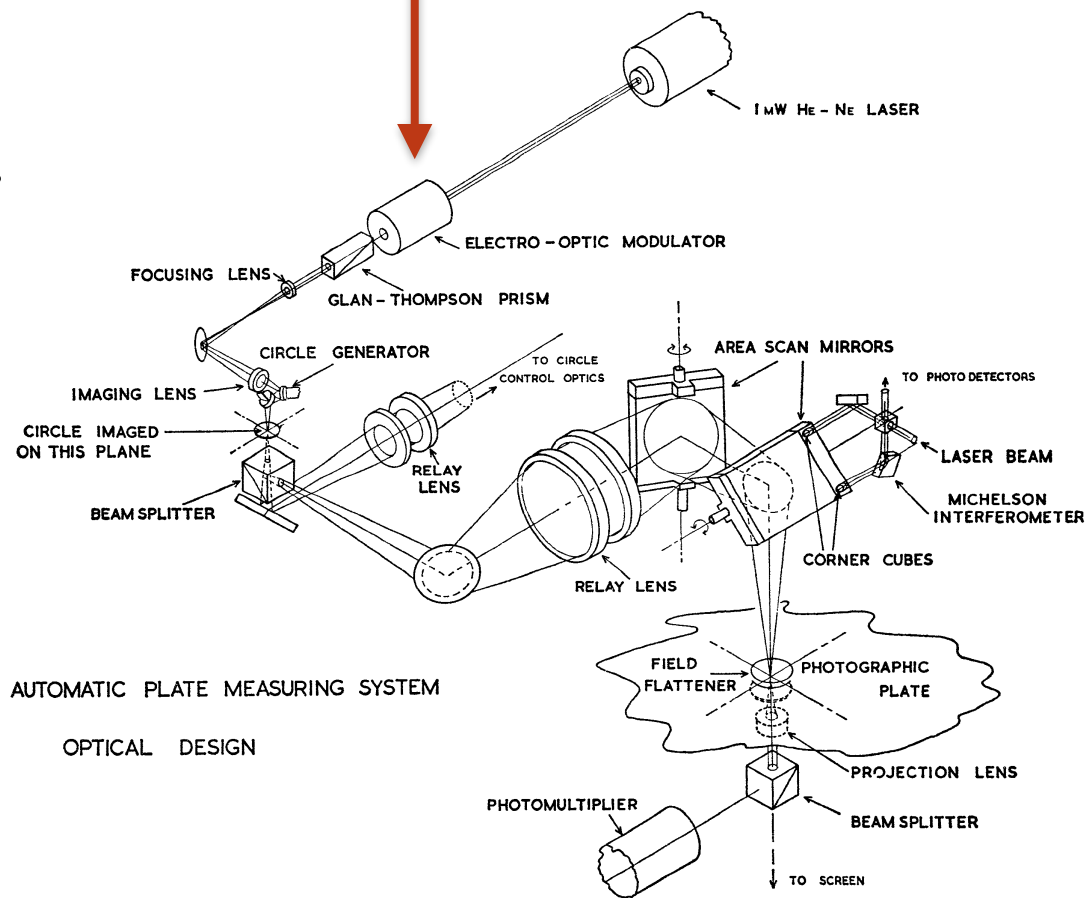
Mike Irwin  
Jim Lewis  
Eduardo Gonzalez-Solares  
Marco Riello  
Aybüke Küpcü Yoldaş  
Simon Hodgkin  
Nic Walton  
Carlos Gonzalez-Fernandez  
Greg Madsen

# The Cambridge Automatic Plate-Measuring Project

EDWARD J. KIBBLEWHITE  
The Observatories, Cambridge, U.K.

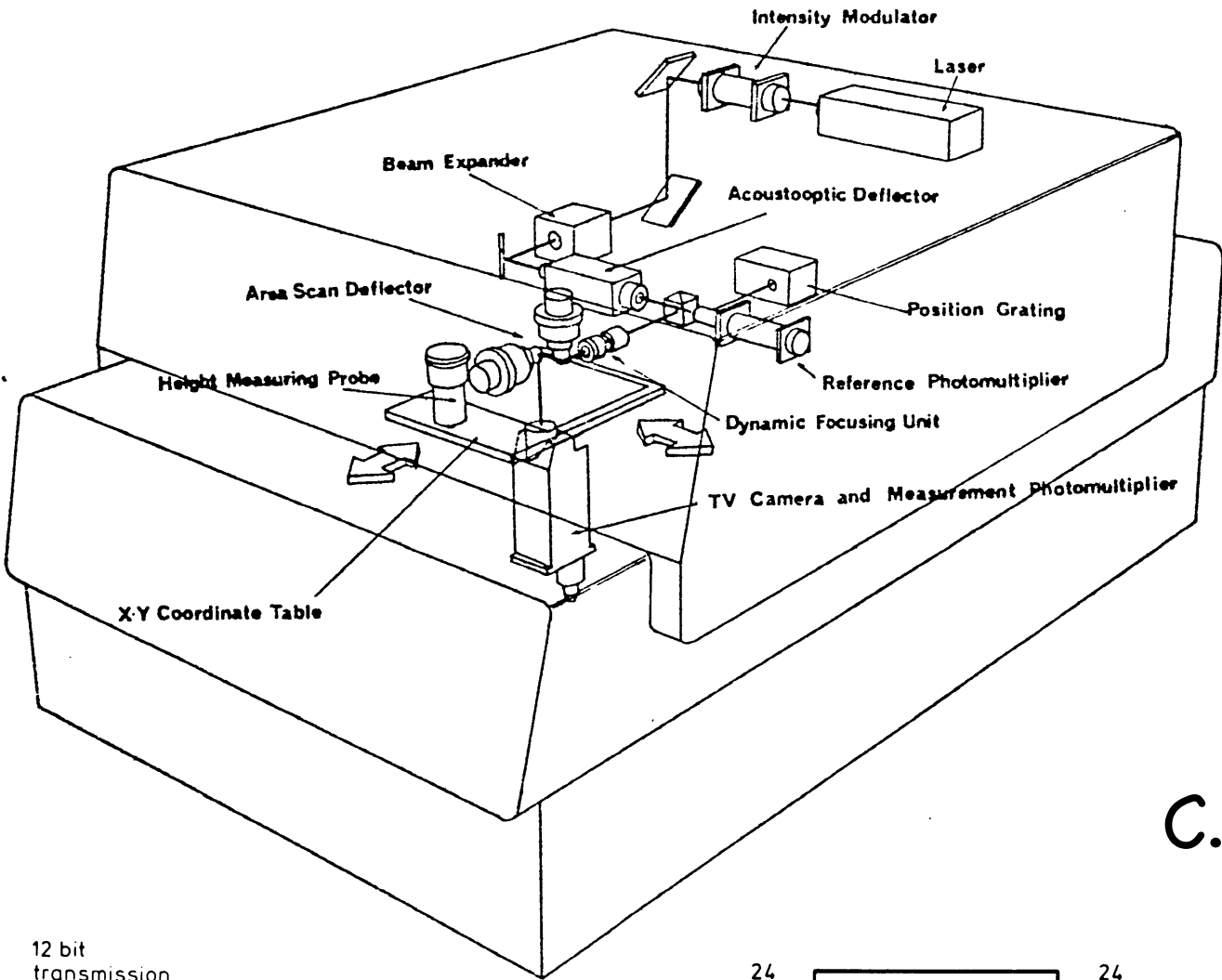
1971 → 1984

A.

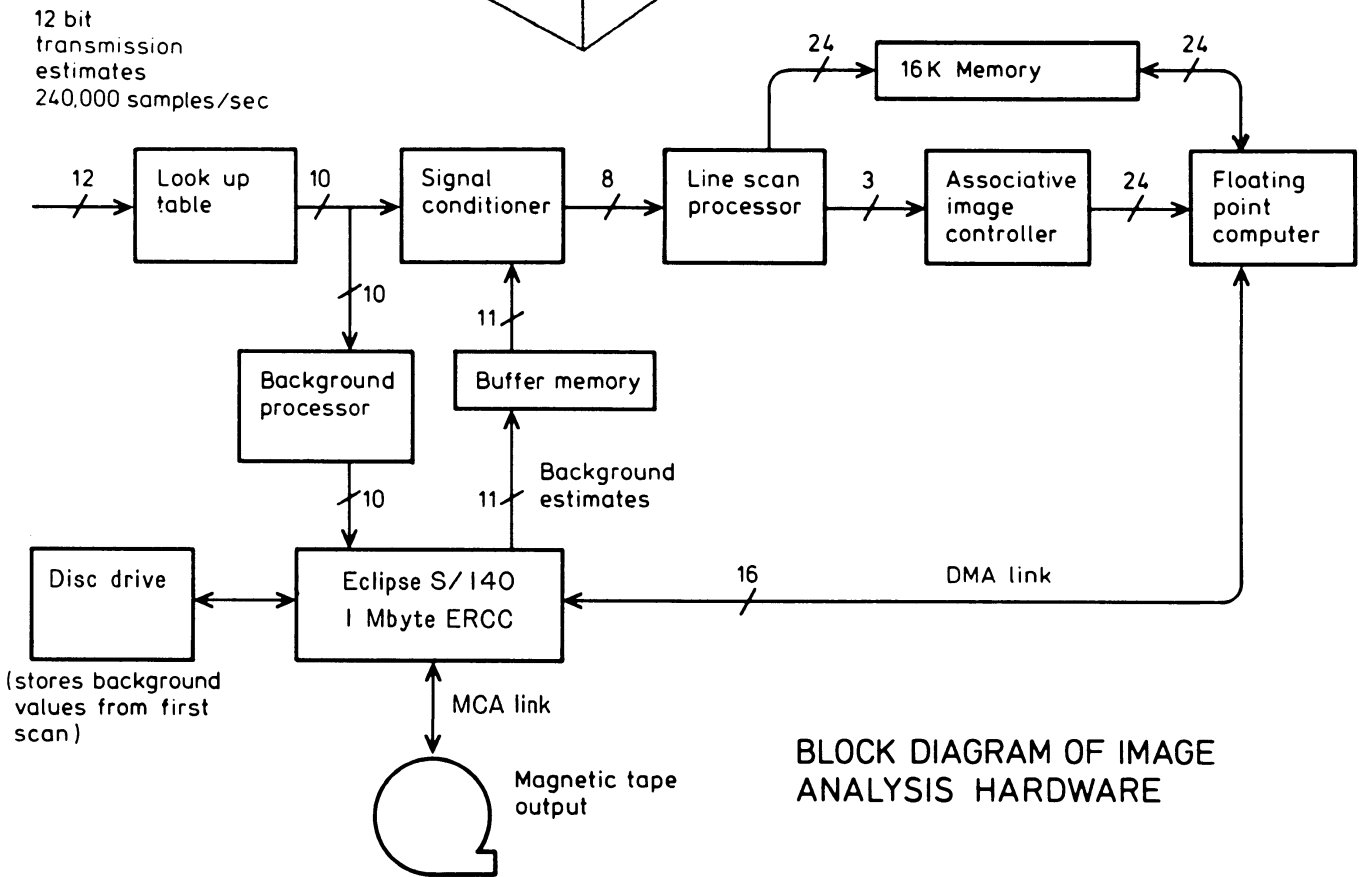


CAMBRIDGE AUTOMATIC PLATE MEASURING SYSTEM  
OPTICAL DESIGN

B.



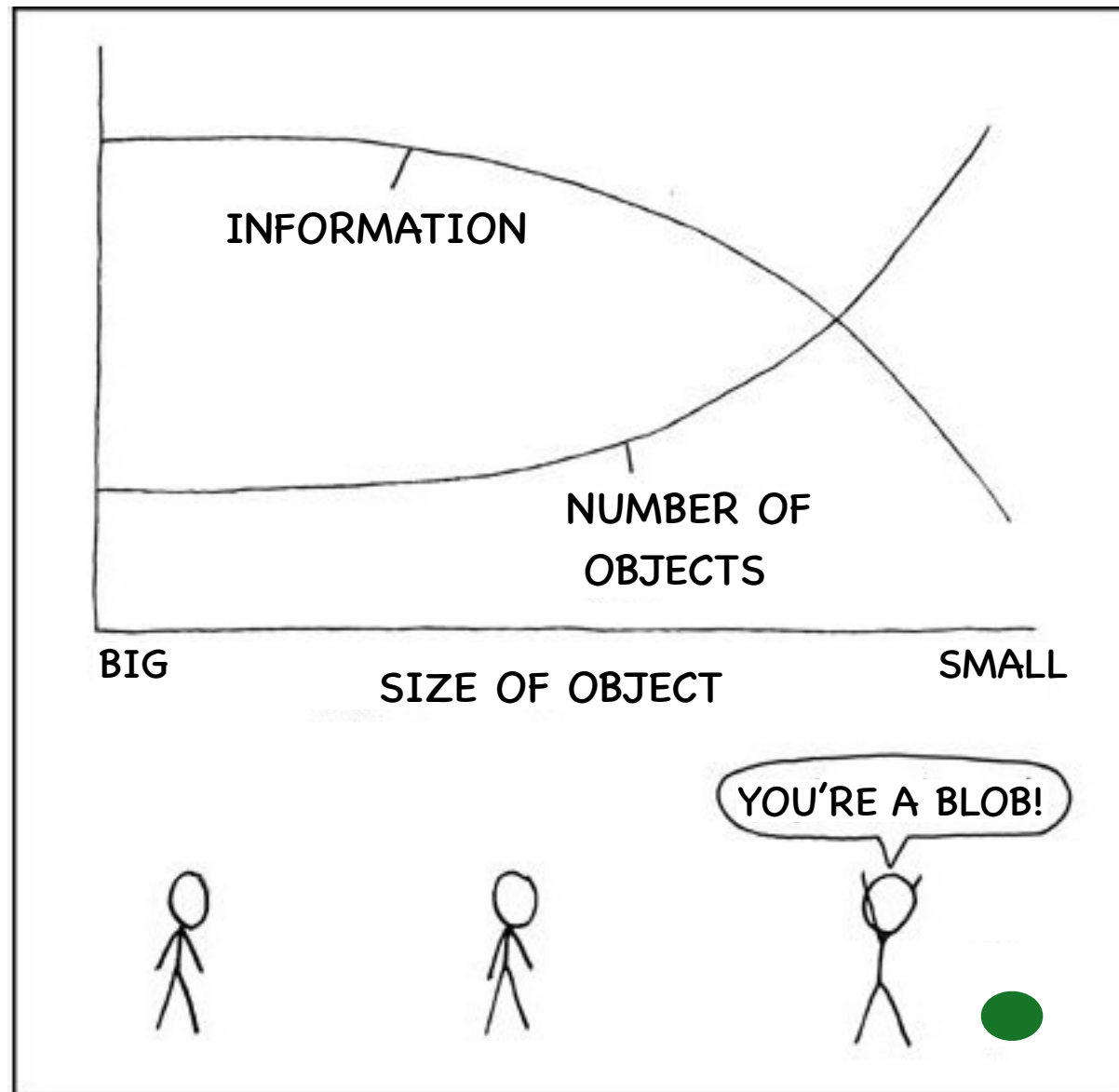
C.



BLOCK DIAGRAM OF IMAGE ANALYSIS HARDWARE



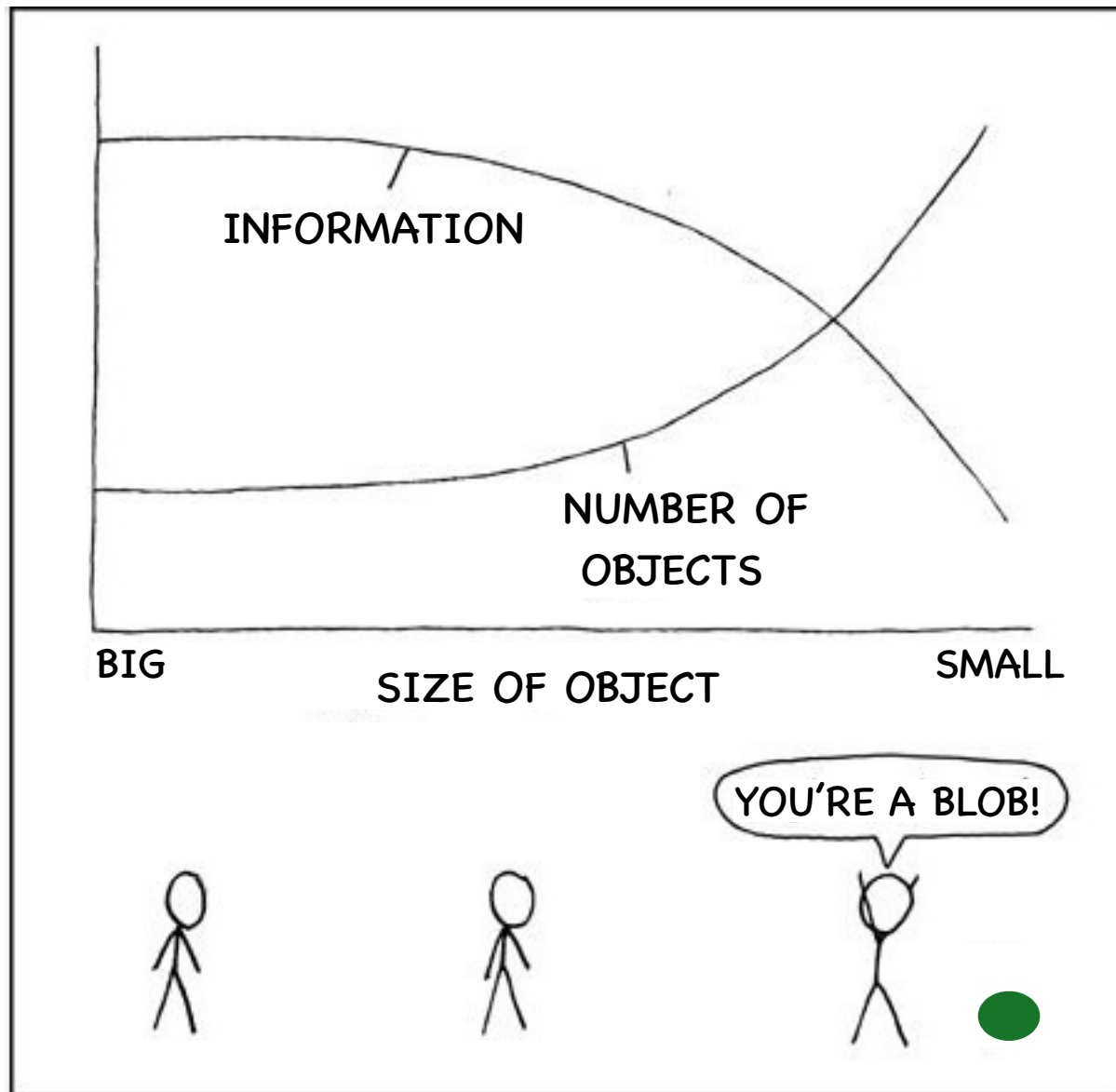
# Kibblewhite Diagram



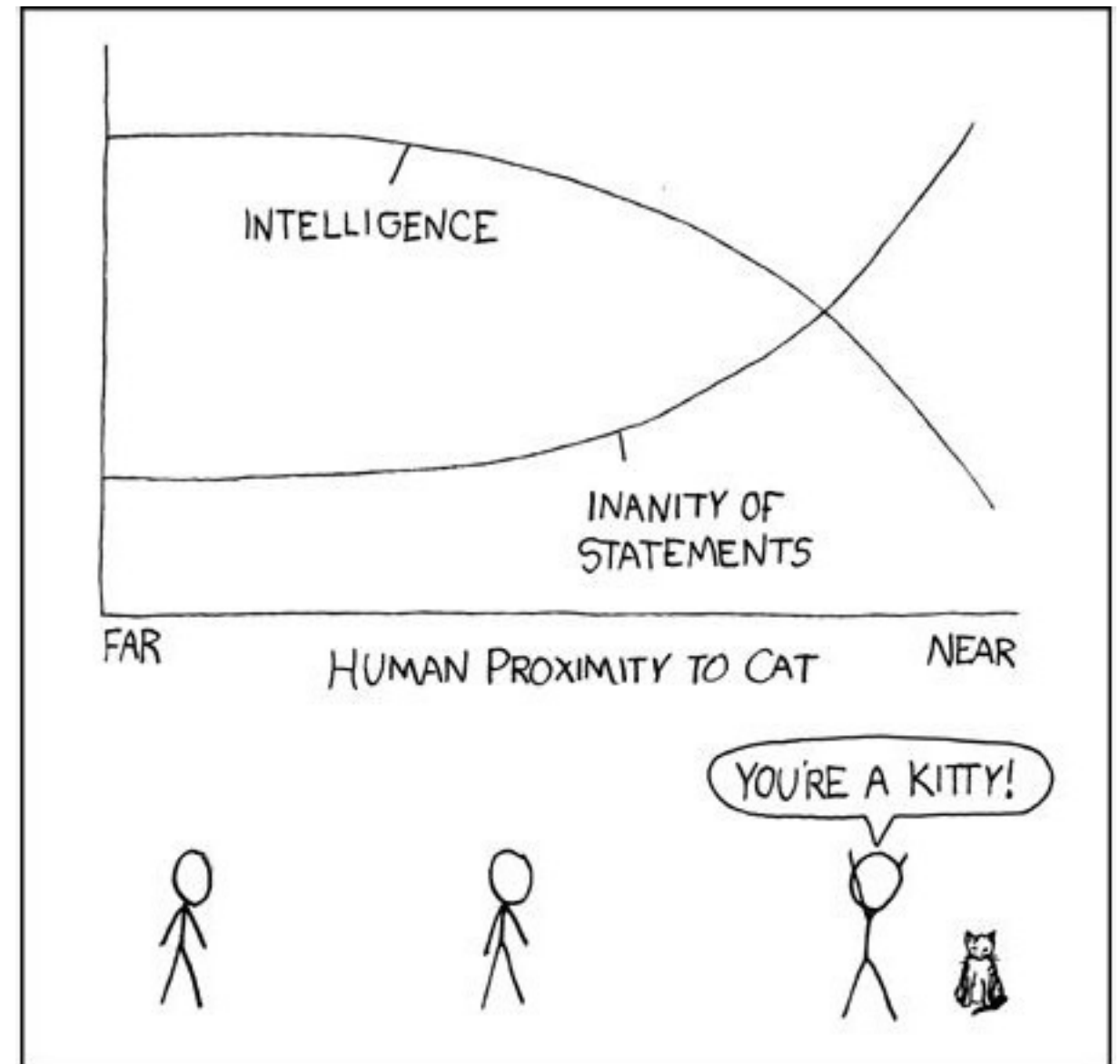
APM processing philosophy:

- design from ML perspective
- decide relevant parameters
- compute MVB error bounds
- tradeoff complexity -v- results
- allow for diversity of objects
- aim for majority science user
- do science with data products

## Kibblewhite Diagram



## Proximity to cats







Mainly used to digitise  
14"x14"  
UKST & POSSI/II  
sky survey plates  $7.5\ \mu$   
sampling  
 $\Rightarrow 43k \times 43k$  pixels  
1.8 Gpixels per plate

### Operational modes:

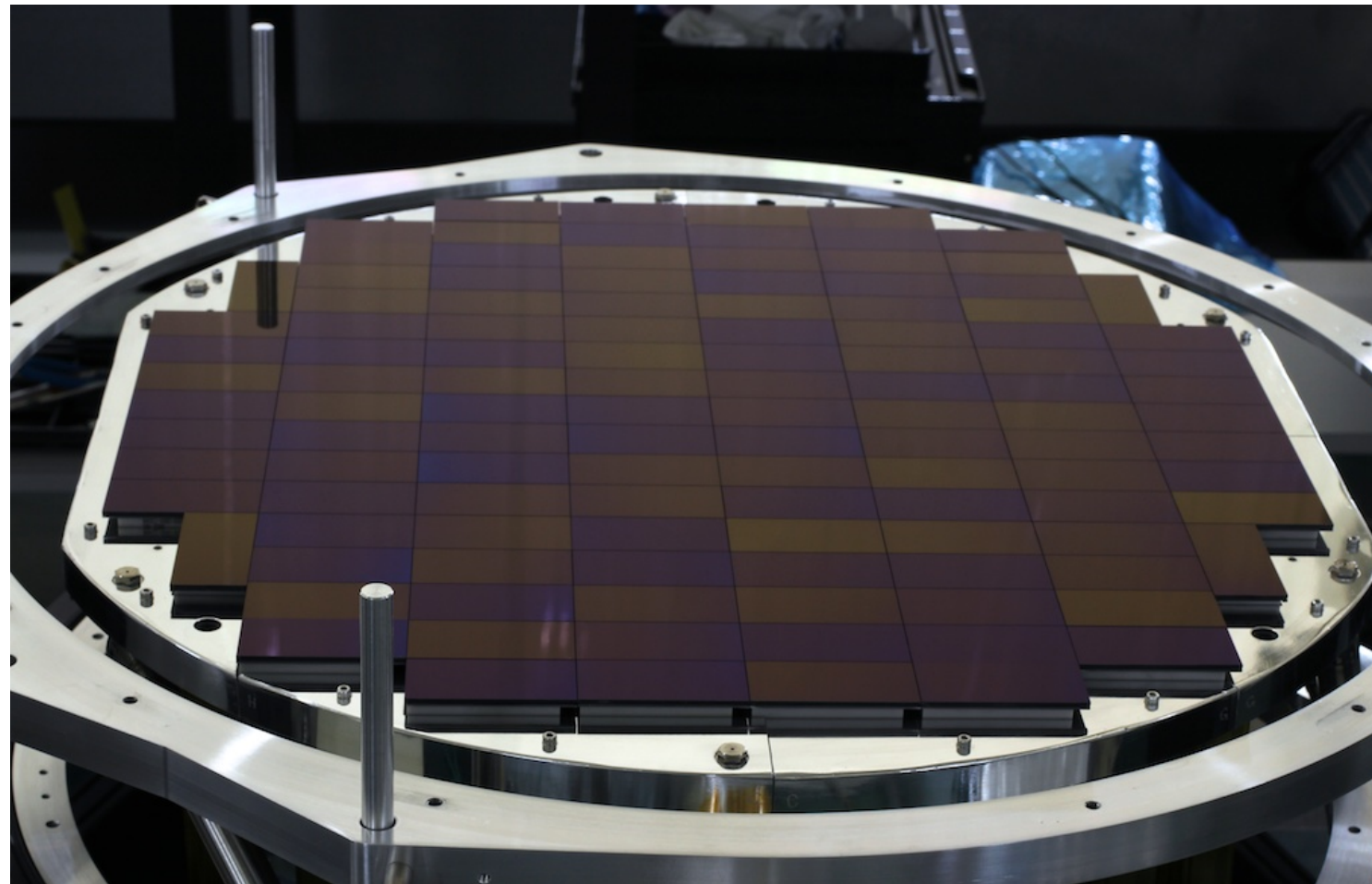
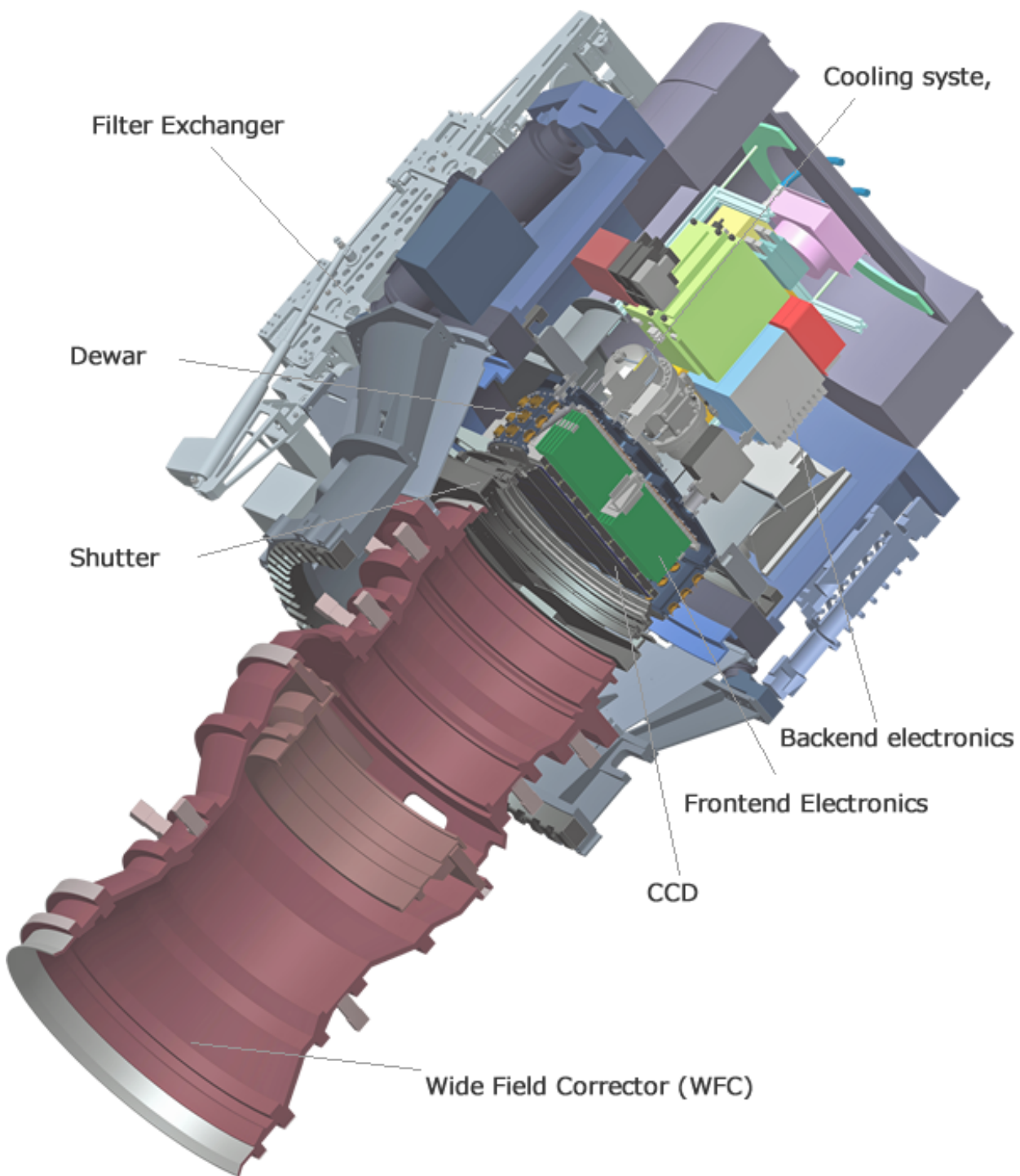
object cataloguing  
objective prism spectra  
scanning images  
candidate verification

r/o 0.5 MByte/s





## HyperSuprime 870 Mpixel camera

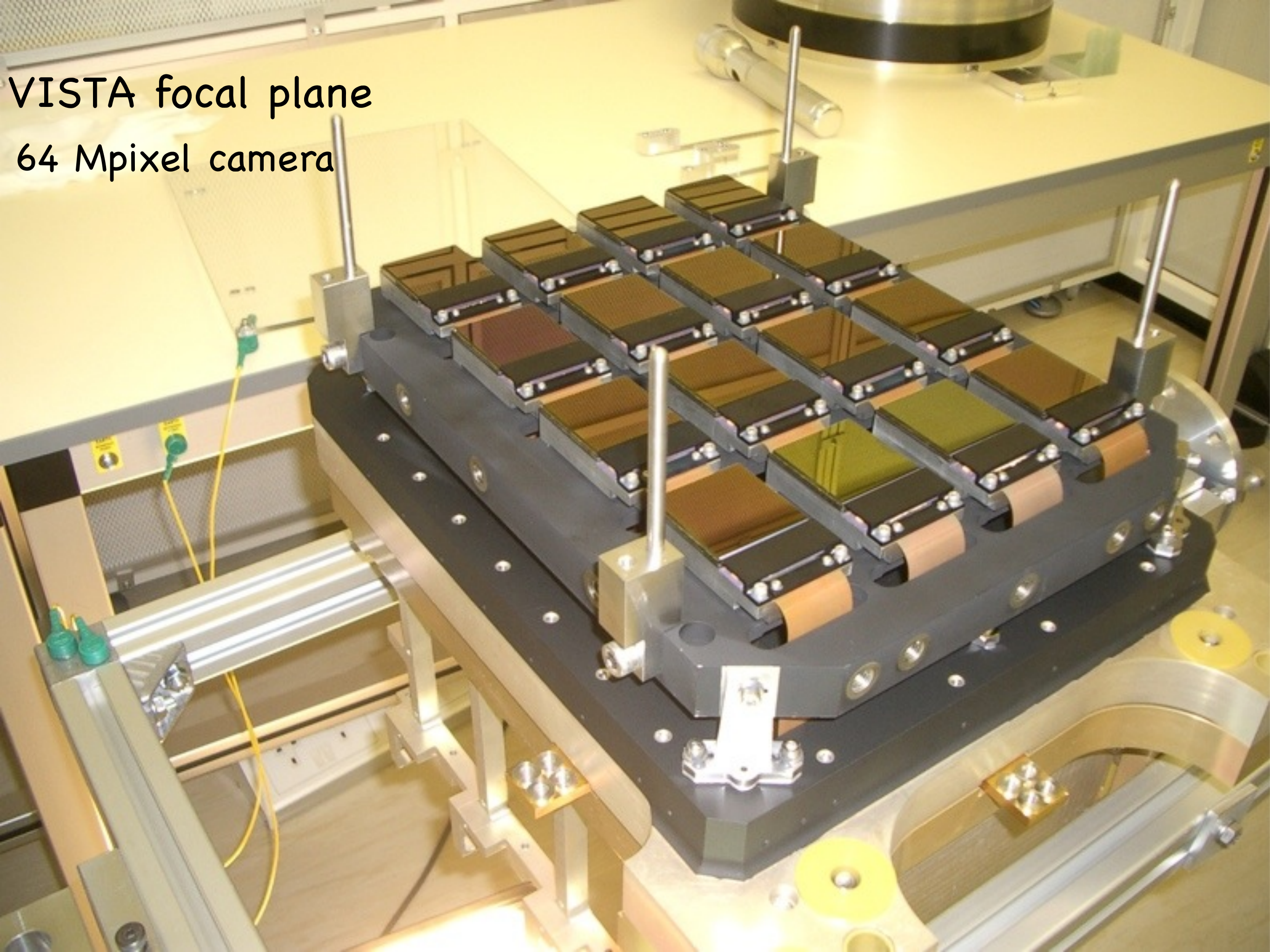








VISTA focal plane  
64 Mpixel camera

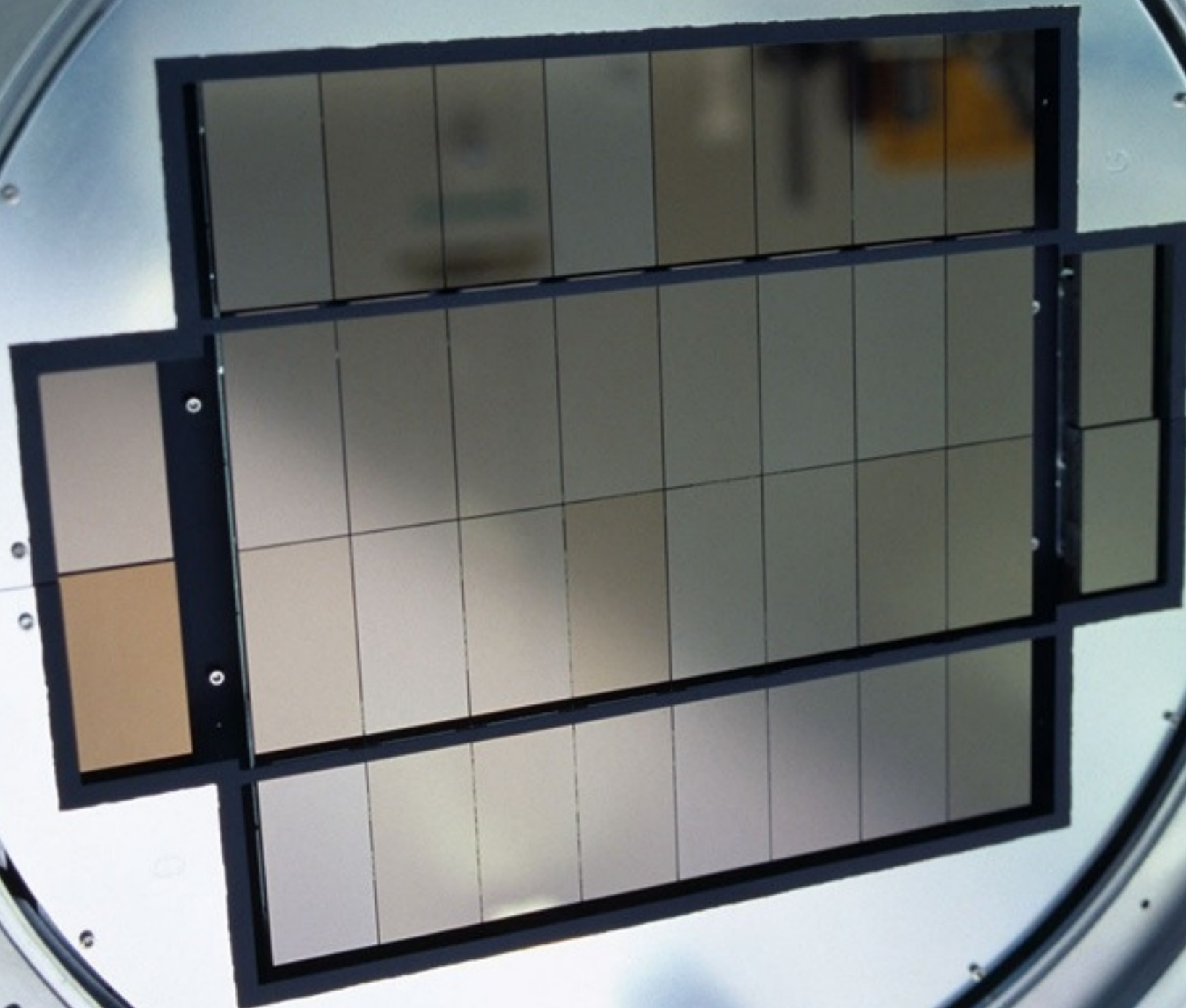








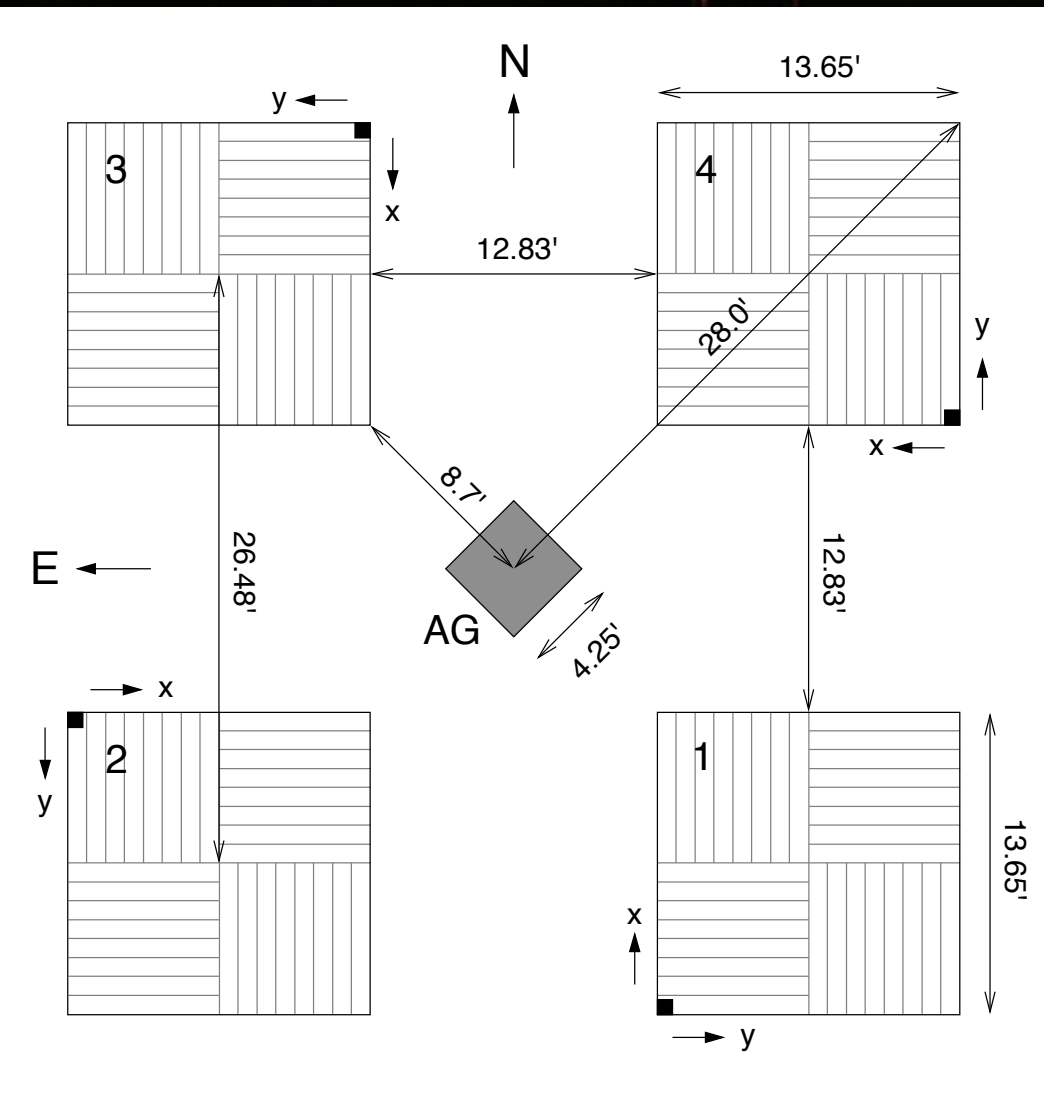
VST focal plane  
256 Mpixel camera







# UKIRT and WFCAM







UKIRT @ Mauna Kea



VISTA @ Paranal



VST @ Paranal

Internet transfer

~ 500-1000GB/night



Processing



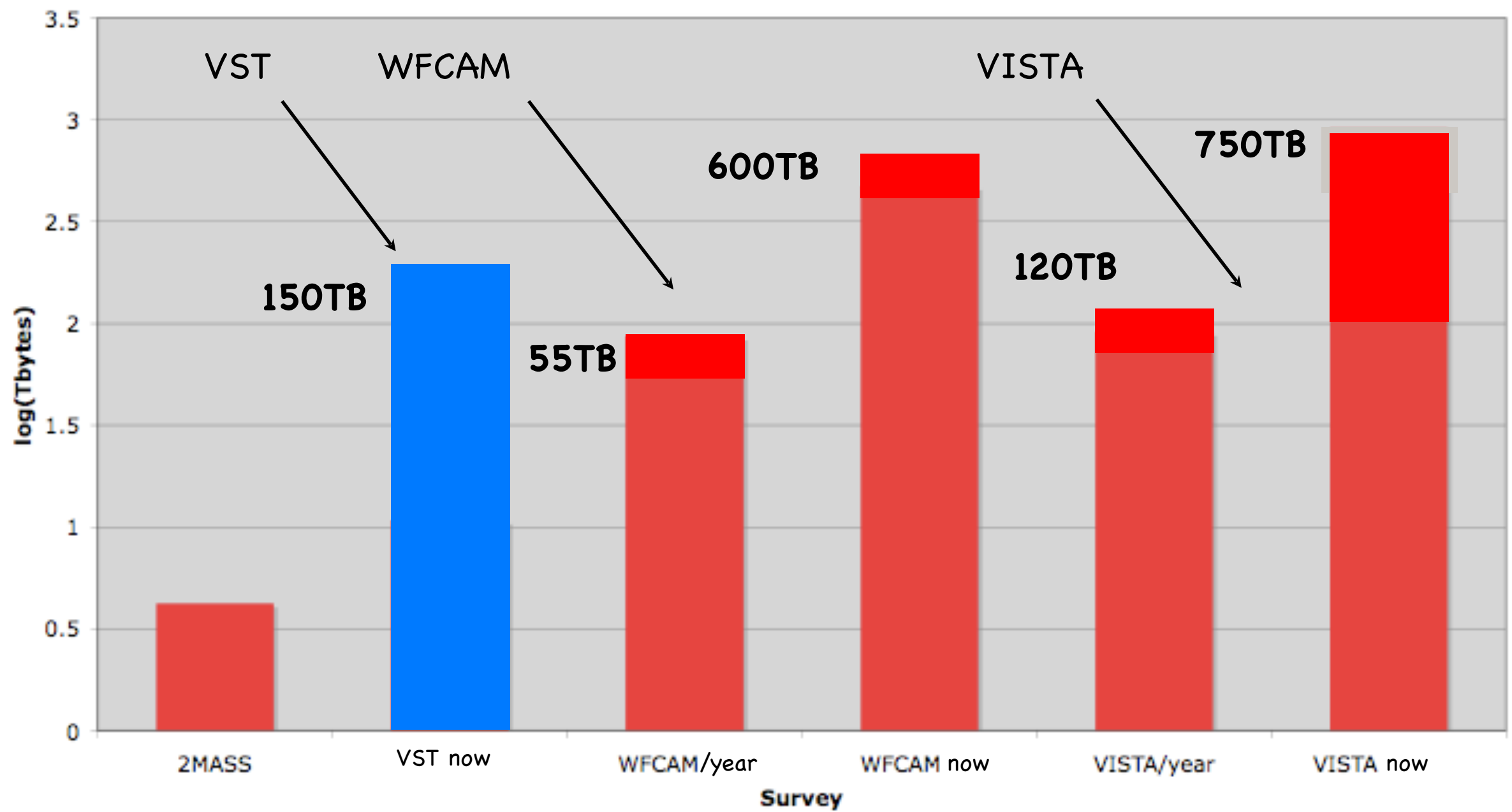
Database ingestion  
and QC



Data serving

CASU

## Data products volume





- ground-based imaging surveys -> *VISTA Data Flow System*
  - all data taken on UKIRT WFCAM (Z,Y,J,H,K.....)
  - all data taken on VISTA VIRCAM (Z,Y,J,H,Ks.....)
  - all VST public survey data (ATLAS, VPHAS, KiDS - u,g,r,i,z,H $_{\alpha}$ )
- processing for other large area/deep imaging instruments
  - CFH, CTIO, INT WFC, Subaru .....
  - all UT4 HAWKI data 2008 ->
  - all UT3 VIMOS imaging data 2002 ->
- processing for wide-field multi-object spectroscopic surveys
  - Gaia-ESO survey
  - data management for WEAVE, 4MOST, MOONS?

# Astro-imaging surveys -> science

- image processing to remove instrumental effects  
aim is a linear photon noise-limited image  
systematic + instrument specific left-overs  
deep stacking with rejection (CRs, bad pixels)
- detect and parameterise objects  
complex background variations (local/global)  
optimal detection via matched filters  
image segmentation into objects  
parameter estimation: position, flux, shape
- external calibration and object classification  
astrometric registration & photometric fluxing  
morphological typing from shape descriptors
- quality control from automated measures
- matching across bands or as time series

# CASU mantra

- MEFs as container -> simplifies bookkeeping
- use lossless Rice-compression -> (x 2-4 less space)
- FITS images and catalogue binary tables (CFITSIO)
- FITS headers record processing details
  - derived QC parameters
  - WCS astrometric calibration
  - photometric calibration
  - table/image fluxes in ADU, x,y positions in pixels
  - versioning and software details
- modular software -> C & perl/python scripts
- minimise external software dependencies





# Cambridge Astronomy Survey Unit

Home

Surveys & Projects

Documents

Publications



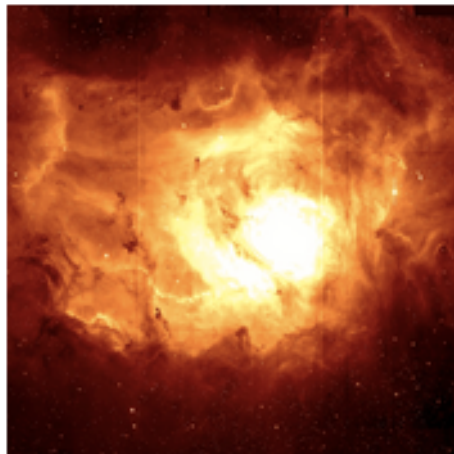
## VISTA

[Technical Documentation](#)

[Data Processing Status](#)

[Survey Progress and Data Access](#)

[more...](#)



## VST

[Technical Documentation](#)

[Data Processing Status](#)

[Survey Progress and Data Access](#)

[more...](#)

## WFCAM

[Technical Documentation](#)

[Data Processing Status](#)

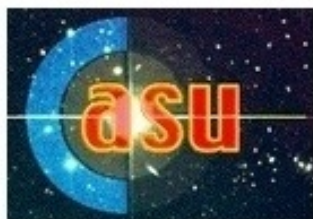
[Survey Progress and Data Access](#)

[more...](#)



<http://casu.ast.cam.ac.uk>

NEW: CASUTools Software Release ([more information](#))



# Cambridge Astronomy Survey Unit

[Home](#)[Surveys & Projects](#)[Documents](#)[Publications](#)

You are here: [Home](#) / [Surveys & Projects](#) / [VST](#) / [Data Processing](#)

[Back](#)

## VST Data Reduction Progress: P93

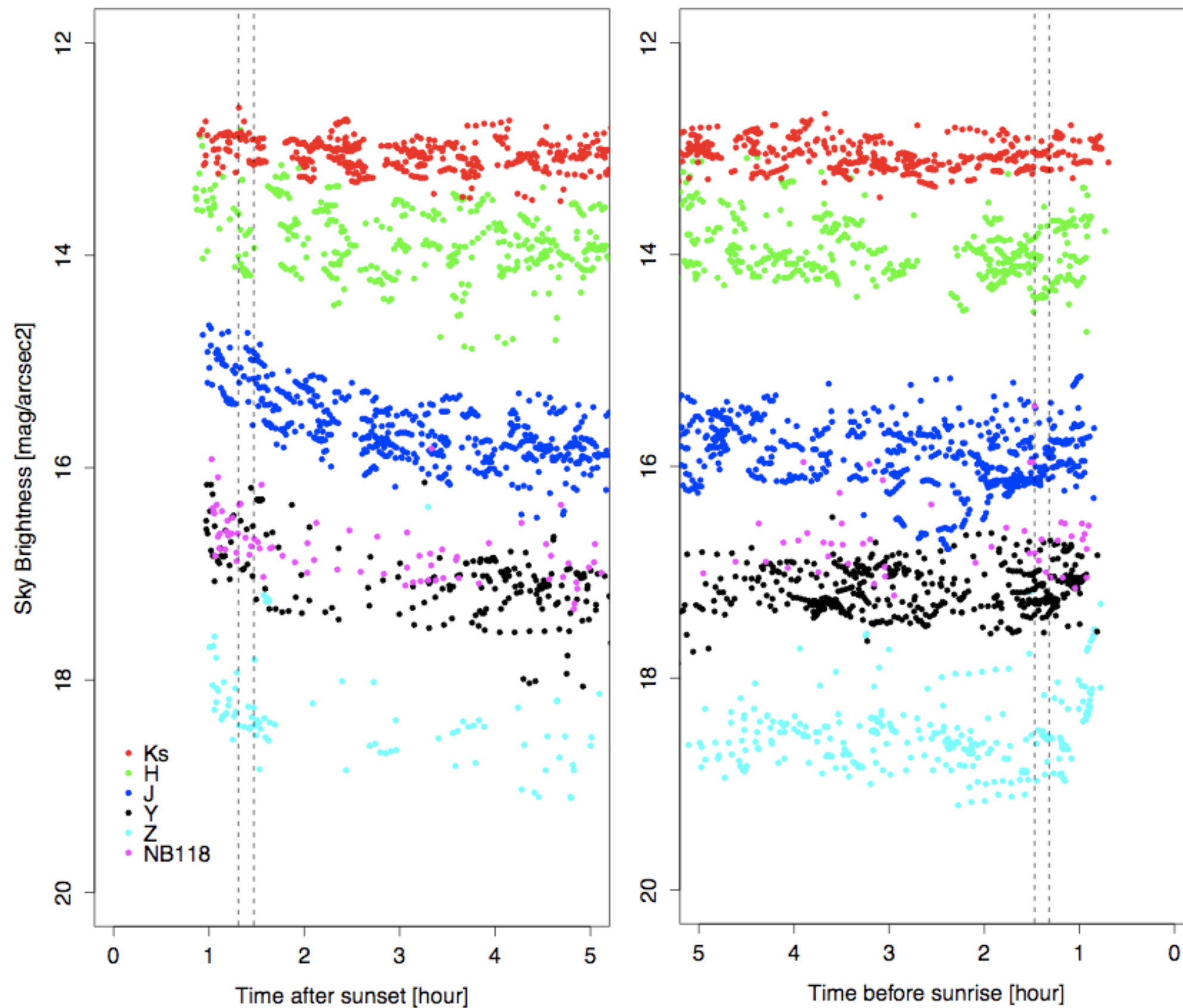
01 Apr 2014 to 30 Sep 2014

This page displays the reduction progress of VST data. Information is automatically updated hourly.

Night	Status	Data	N <sub>raw</sub>	Version	Released	Summary Plots	Photom Plots	Summary Info	Observation Log	Paranal ambient conditions	Size red [GB]	ATLAS	VPHAS	KIDS
20140401	REDUCED	<a href="#">D</a>	231	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	60.29	68	69	5
20140402	REDUCED	<a href="#">D</a>	193	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	50.71	68	3	34
20140403	REDUCED	<a href="#">D</a>	167	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	49.27	68	0	39
20140404	UNPROCESSED		16						<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>				
20140405	UNPROCESSED		14						<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>				
20140406	REDUCED	<a href="#">D</a>	132	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	33.92	0	0	51
20140407	REDUCED	<a href="#">D</a>	172	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	45.94	0	62	25
20140408	REDUCED	<a href="#">D</a>	145	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	52.77	0	77	31
20140409	REDUCED	<a href="#">D</a>	170	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	42.70	0	0	0
20140410	REDUCED	<a href="#">D</a>	198	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	61.62	68	56	15
20140411	REDUCED	<a href="#">D</a>	177	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	40.65	24	0	0
20140412	REDUCED	<a href="#">D</a>	191	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	36.90	0	0	10
20140413	REDUCED	<a href="#">D</a>	106	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	40.84	0	0	5
20140414	REDUCED	<a href="#">D</a>	144	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	57.90	0	63	0
20140415	REDUCED	<a href="#">D</a>	131	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	47.13	0	0	15
20140416	REDUCED	<a href="#">D</a>	151	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	50.59	0	0	15
20140417	REDUCED	<a href="#">D</a>	177	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	48.57	0	0	30
20140418	REDUCED	<a href="#">D</a>	149	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	39.35	0	0	52
20140419	REDUCED	<a href="#">D</a>	159	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	56.11	68	0	34
20140420	REDUCED	<a href="#">D</a>	168	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	57.96	0	105	9
20140421	REDUCED	<a href="#">D</a>	114	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	28.02	0	18	30
20140422	REDUCED	<a href="#">D</a>	175	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	55.20	0	93	14
20140423	NODATA									<a href="#">nightmon</a>				
20140424	UNPROCESSED		17						<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>				
20140425	UNPROCESSED		15						<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>				
20140426	REDUCED	<a href="#">D</a>	162	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	38.28	6	40	16
20140427	REDUCED	<a href="#">D</a>	198	1.0	20140711	<a href="#">GIF1</a> <a href="#">GIF2</a>	<a href="#">GIF</a>	<a href="#">summary</a>	<a href="#">index</a> ∴ <a href="#">eso</a>	<a href="#">nightmon</a>	37.17	0	0	60



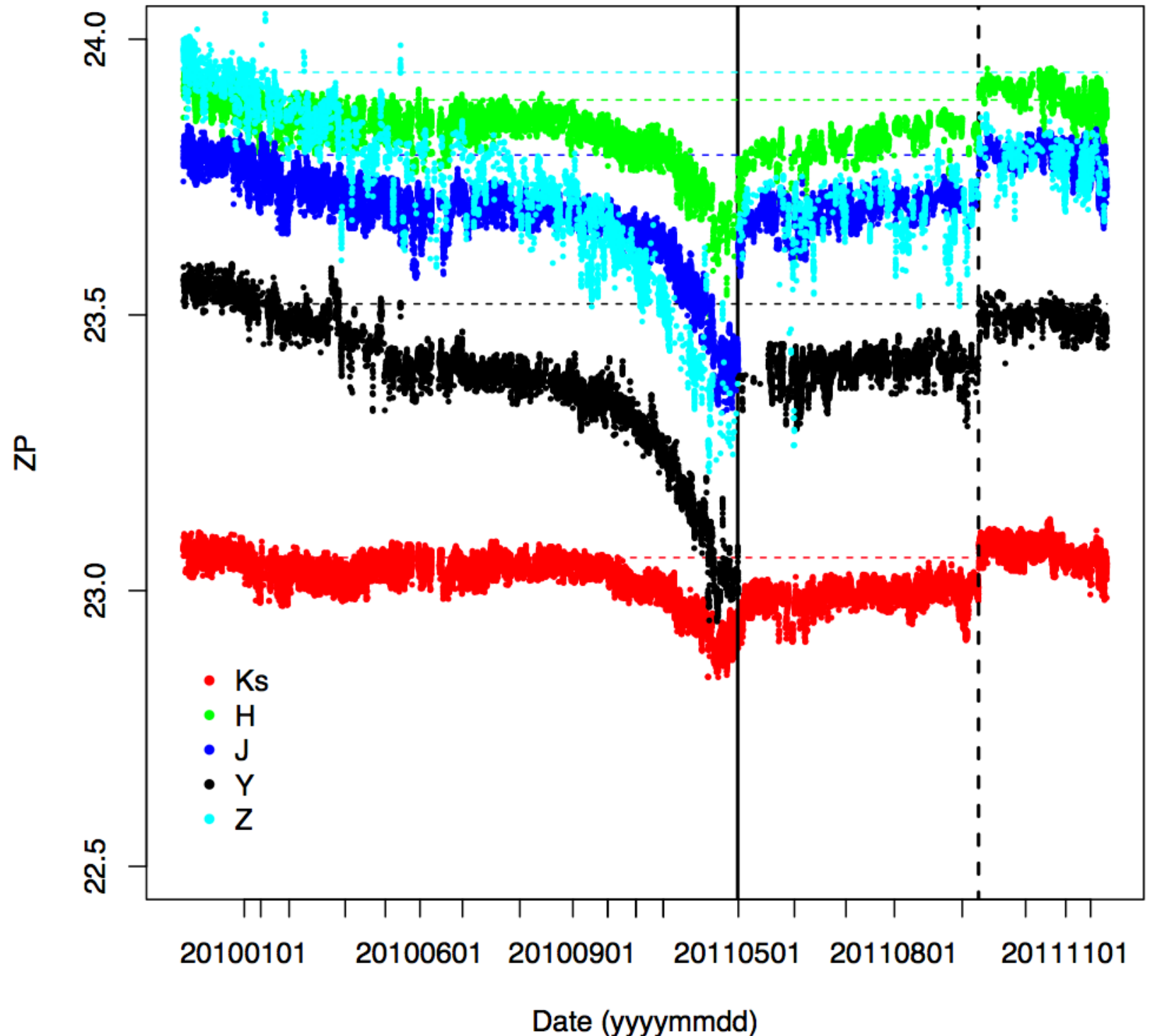
# Monitoring Paranal NIR sky surface brightness



# An example of QC trend analysis VISTA/VIRCAM

Variable zero-point  
(and hence survey  
depth) with time

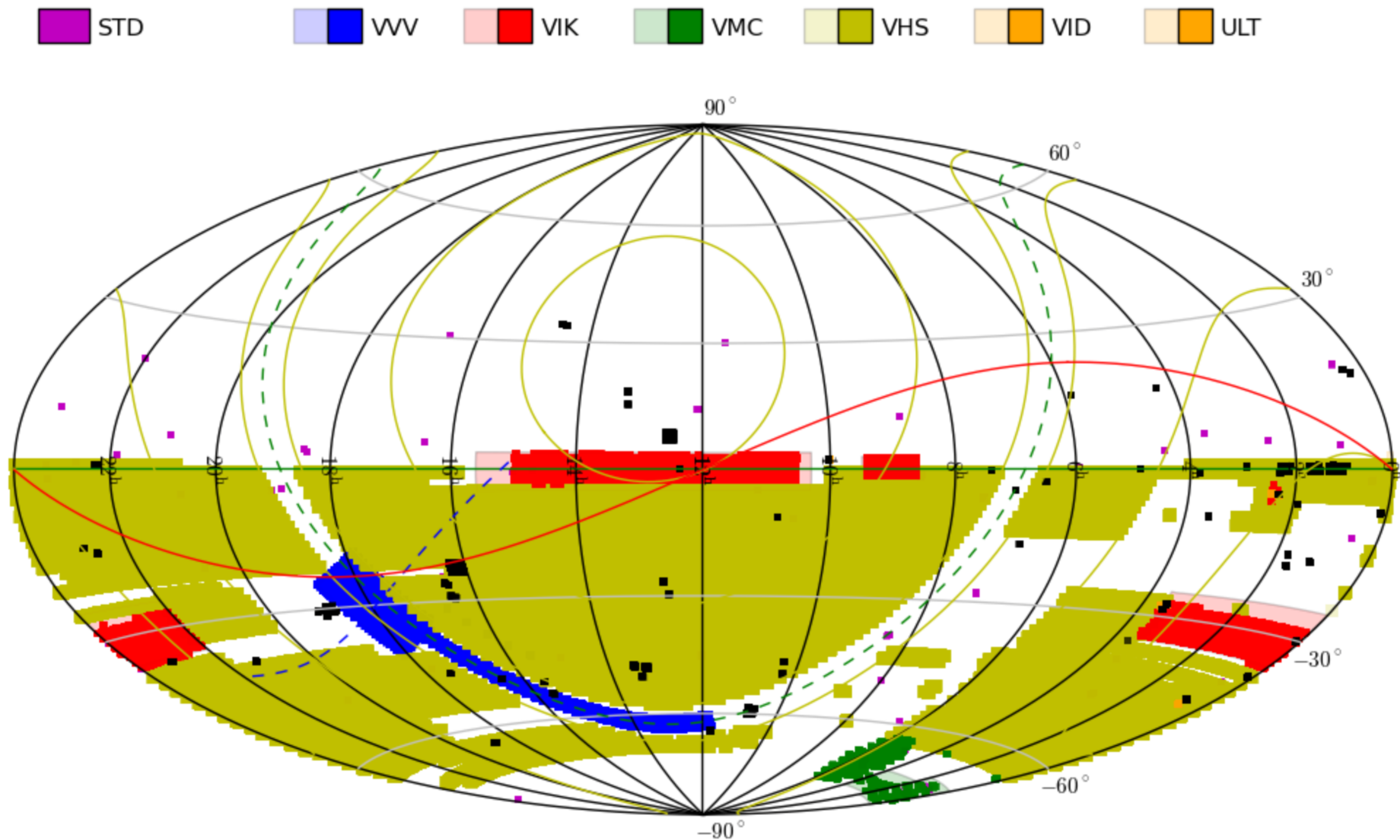
- silvered mirror degradation
- aluminised mirror
- a bit of window cleaning





# VISTA survey progress

Distribution of all VISTA observations in the sky using three different projections, Aitoff, Zenithal Equal Area and Cartesian.



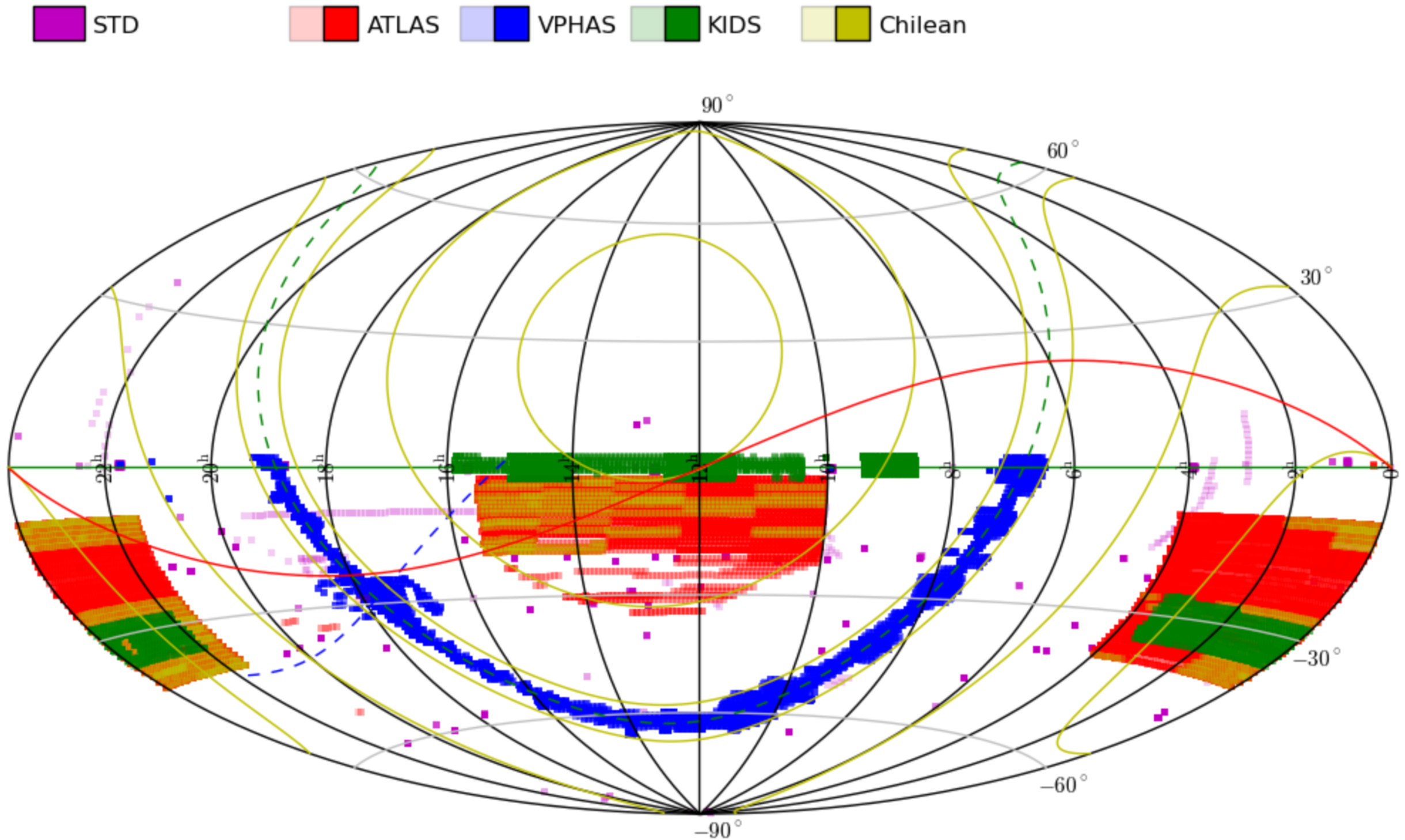
Observing dates: 20091015 - 20150531

Cambridge Astronomy Survey Unit

Last Updated: 05/10/2015

# VST survey progress

Distribution of all VST observations in the sky using three different projections, Aitoff, Zenithal Equal Area and Cartesian.

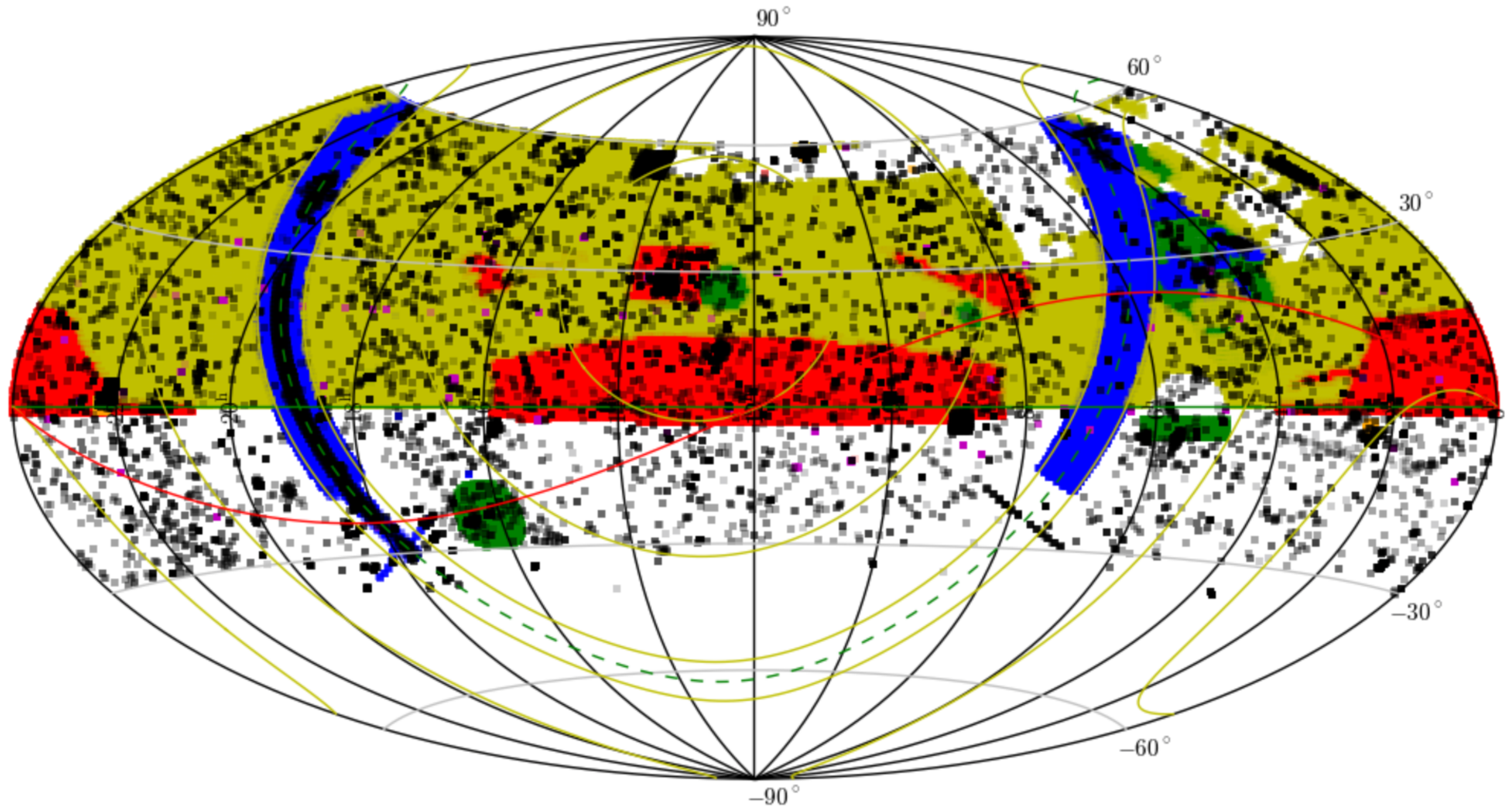


Observing dates: 20110606 - 20150331  
Cambridge Astronomy Survey Unit

Last Updated: 23/04/2015



# WFCAM survey progress

[All](#)[Z](#)[Y](#)[J](#)[H](#)[K](#) CAL GPS LAS GCS UHS DXS, UDS PPI

Date Range: 20050401 - 20140731

Last Updated: 20150223



# VST Archive At CASU

[Example manager] logout

Home

Manager

My requests

Search

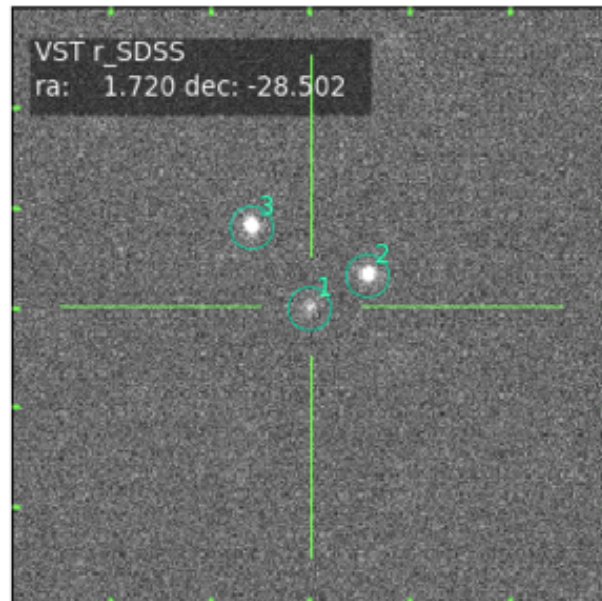
Data processing

Technical Information

QC Tables

Help

o20110926\_00214\_st.fit[11] - r\_SDSS



Obs date 2011-09-27 07:00:52  
Airmass 1.211  
Exposure Time [sec] 45.0  
Average seeing [arcsec] 1.37  
WCS fit rms 0.09  
Ellipticity 0.04  
Magnitude limit [Vega] 22.56  
Programme 177.A-3011 (ATLAS)

Current cutout size 60 arcsec

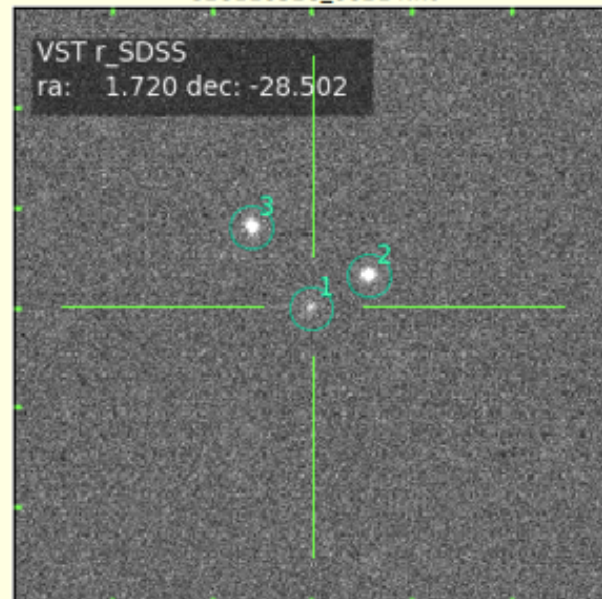
30 arcsec 60 arcsec 90 arcsec 120 arcsec

A search by position returns images that contain that position and allows preview of postage stamps, catalogue sources and a view of provenance images.

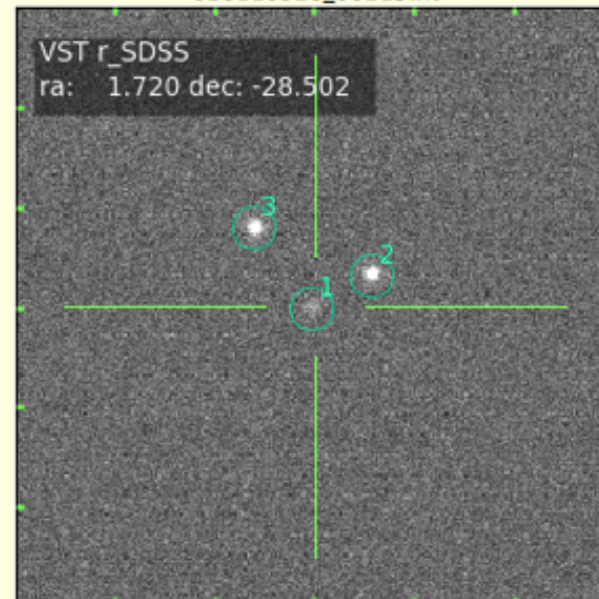
ID	Coords (J2000)	Apermag3	Class	Ellipt	Pos Ang	X	Y	AvConf	ErrBit
1	00:06:52.766 -28:30:07.622	21.168 (0.062)	extended	0.11	-11.95	356.714	1180.97	100.606	0.0
2	00:06:52.325 -28:30:04.273	19.982 (0.025)	pointlike	0.09	79.33	329.41	1196.59	100.618	0.0
3	00:06:53.211 -28:29:59.505	19.650 (0.020)	pointlike	0.10	25.62	384.081	1219.12	100.582	0.0
4	00:06:53.409 -28:29:33.547	22.443 (0.182)	pointlike	0.09	50.23	395.974	1340.87	100.817	0.0
5	00:06:53.223 -28:30:56.738	22.085 (0.135)	extended	0.50	6.14	385.536	950.749	100.76	0.0

## Individual Images

o20110926\_00214.fit



o20110926\_00215.fit



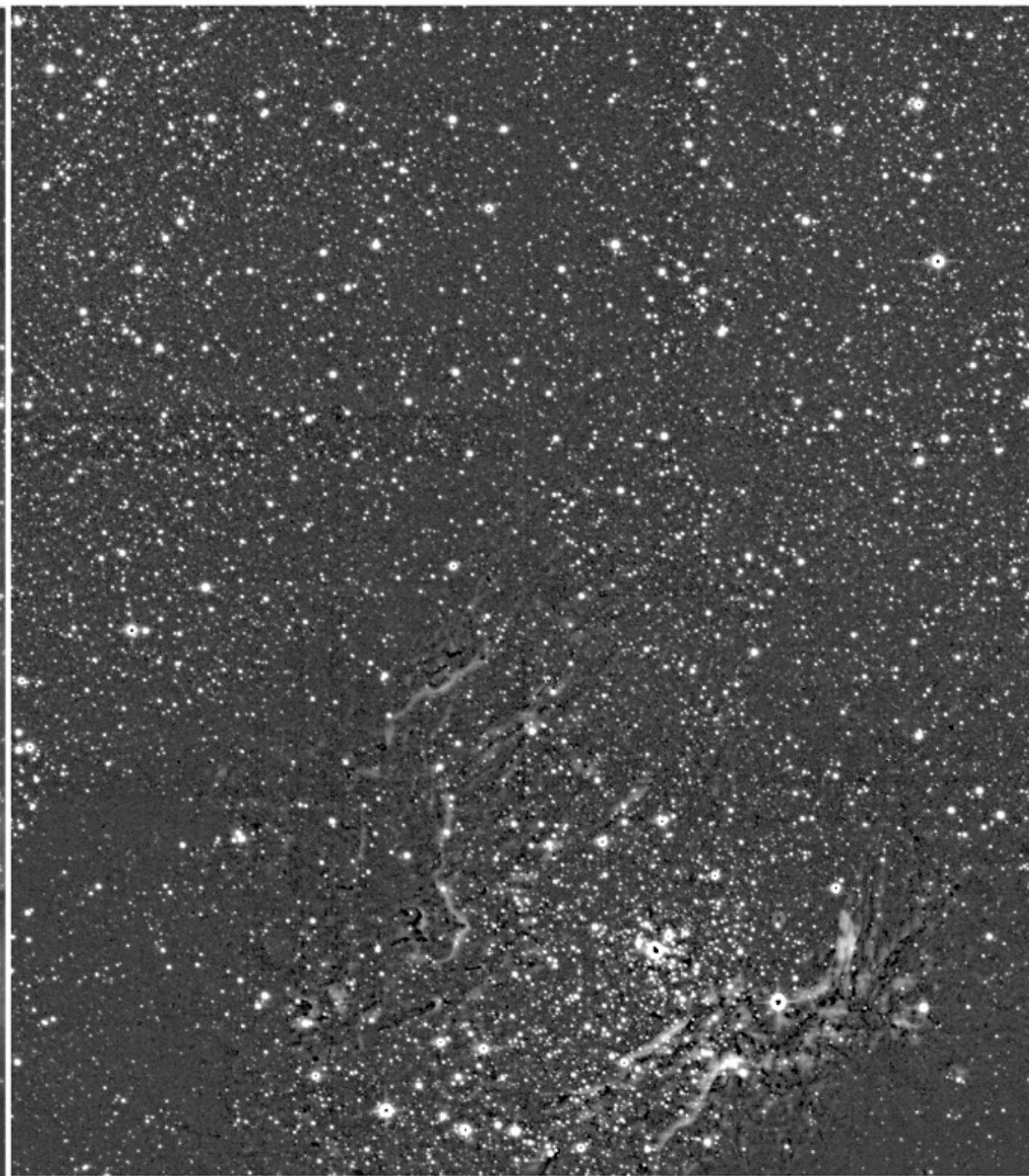


# Some software solutions for common issues

- nebuliser
  - removes complex background variations
  - enhanced object detection & parameterisation
- despiker
  - removes diffraction spikes, charge bleeding artefacts, and saturated stellar cores, and even satellite tracks
- artefact/noise pickup removal
  - crosstalk intra- and inter-detectors
  - noise pickup from other systems
- point-spread-function estimation
  - automatically generates detector-level PSFs for QC
  - and for performing PSF photometry



# Nebuliser -> M17 K-band WFCAM



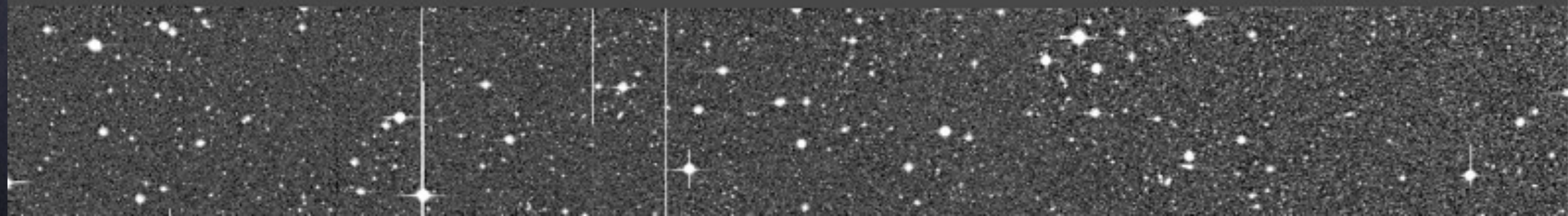


Nebuliser -> M31 field 23 MegaCam





Nebuliser -> M31 field 23 MegaCam





Example of  
VST pickup  
20130605





imcombine  
of pickup  
pattern  
using all  
32 CCDs





scaled  
removal  
of pickup

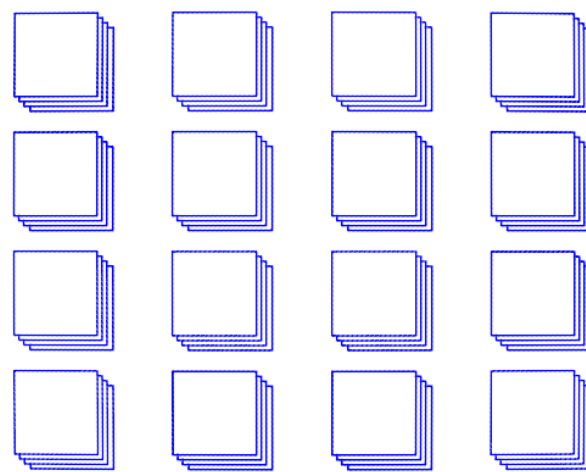




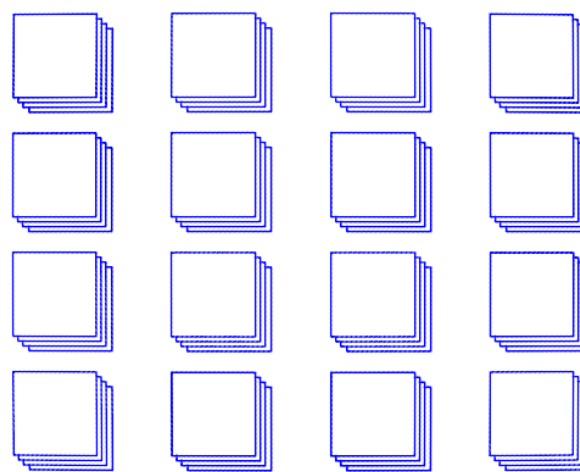
[illegible]

Science detectors,  
controlled by IRACE

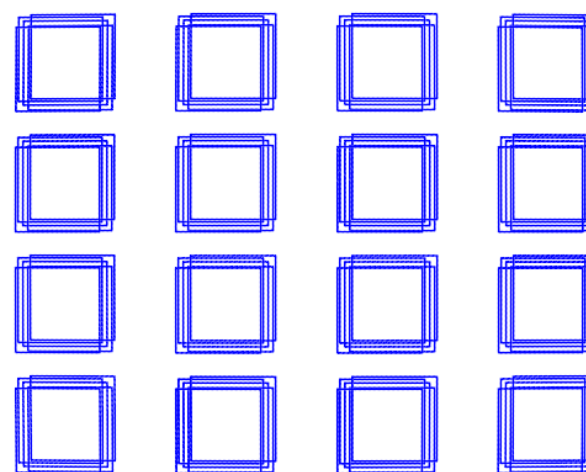




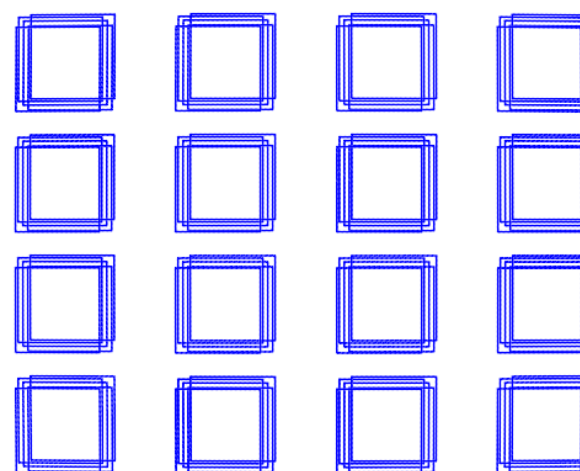
VISTA  
pawprints



x6  
.....

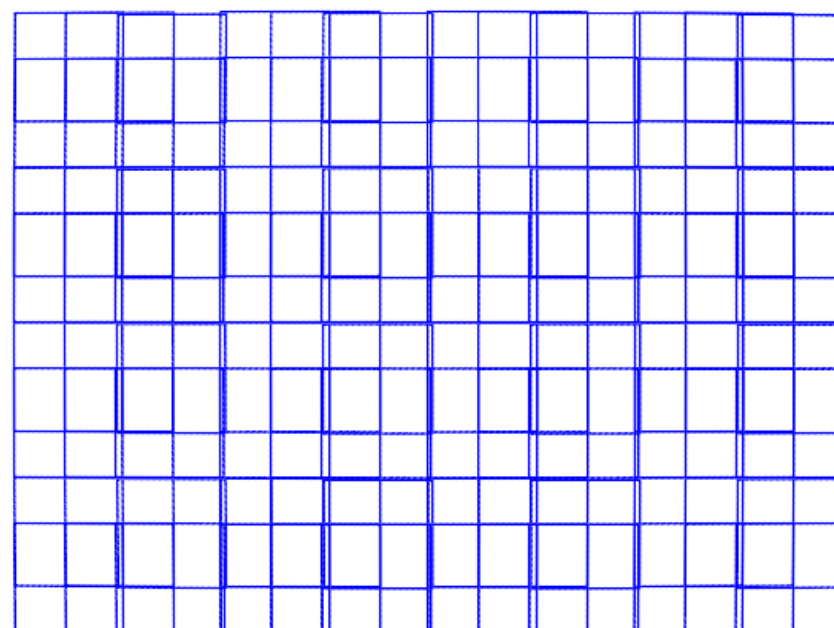


dither  
stacked



+ confidence map  
+ catalogue

mosaic  
tile



96 chips

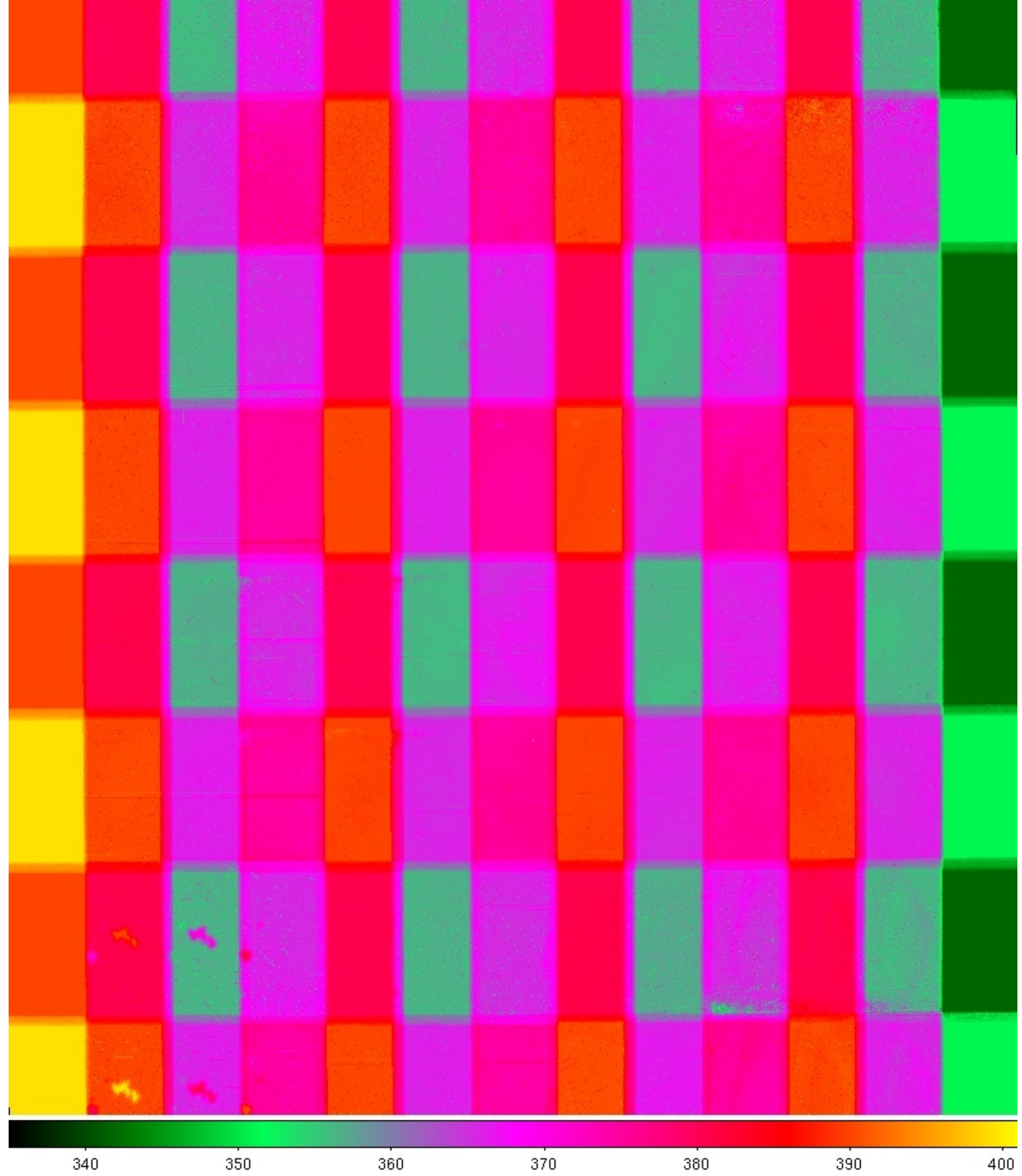
+ confidence map  
+ catalogue

# Challenges when dealing with VIRCAM tiles

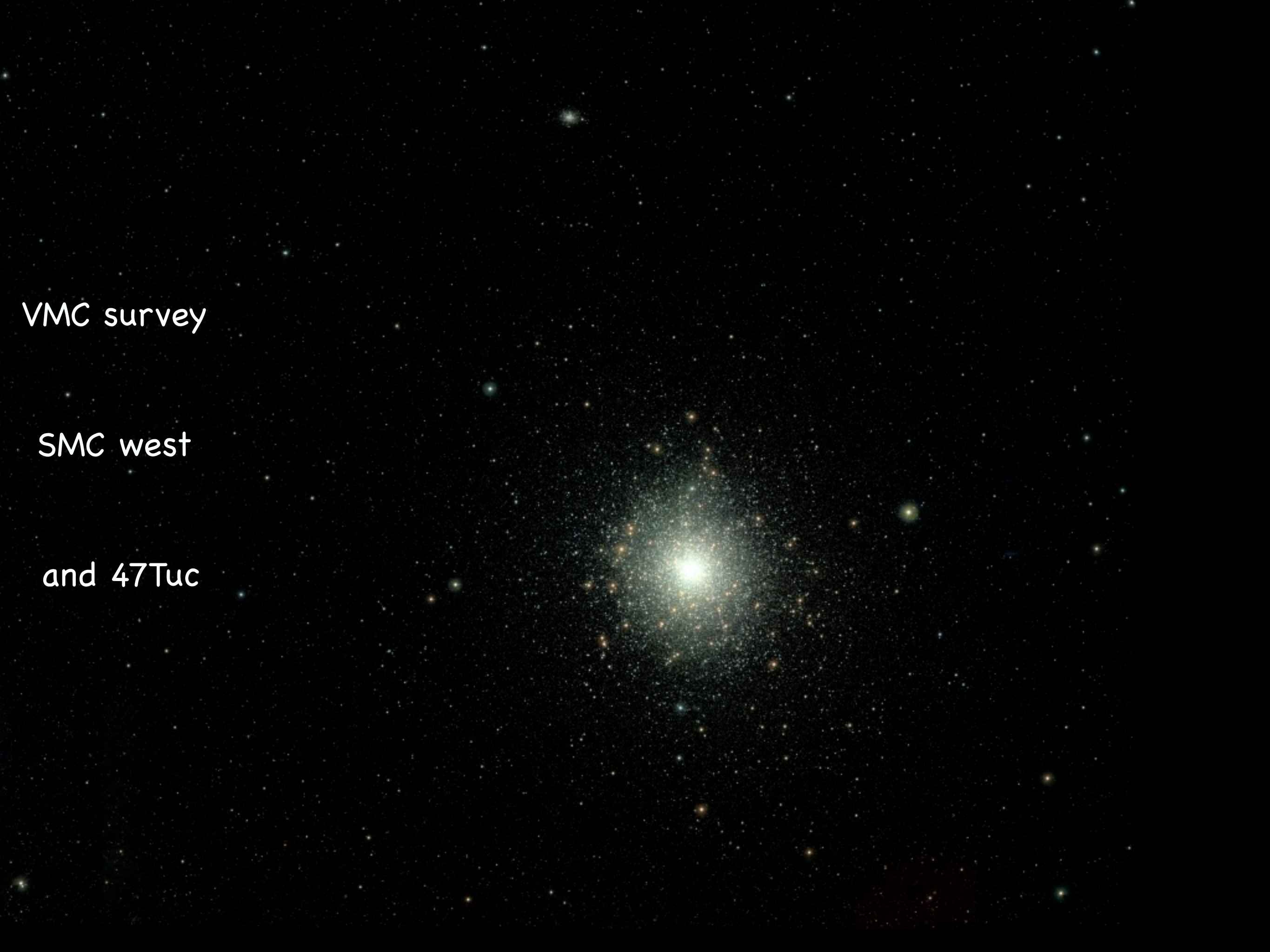
- imperfect sky subtraction pawprint matching + artefacts
- variable PSF across pawprint detectors
- variable seeing conditions for entire pawprint
- variable saturation levels of each detector
- variable extinction (ZP) during tile observations
- astrometric distortion = photometric distortion
- interpolation = varying correlated noise patterns
- “interesting” Modified Julian Date pattern



MJD  
variation  
across  
tiles





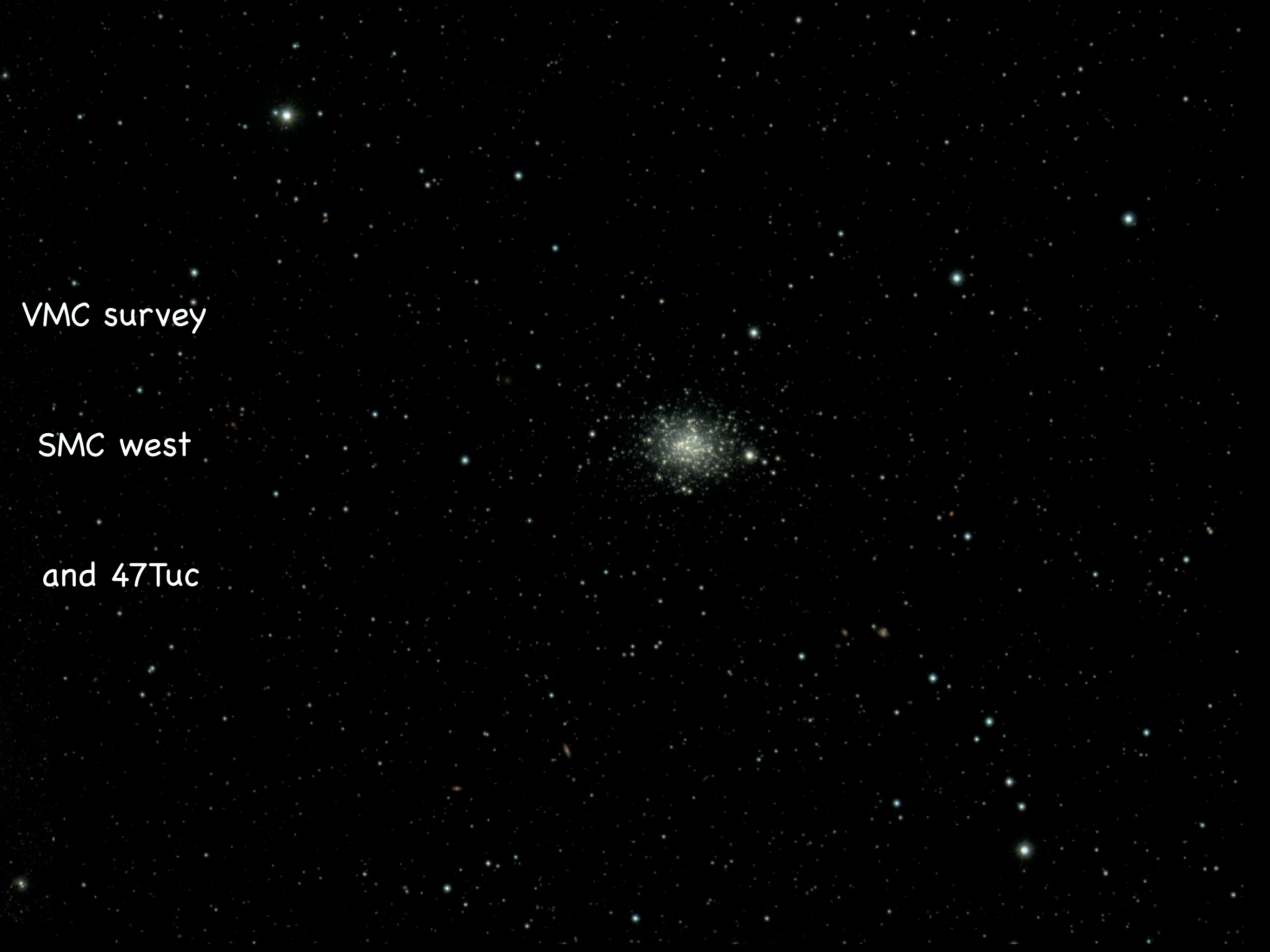


VMC survey

SMC west

and 47Tuc



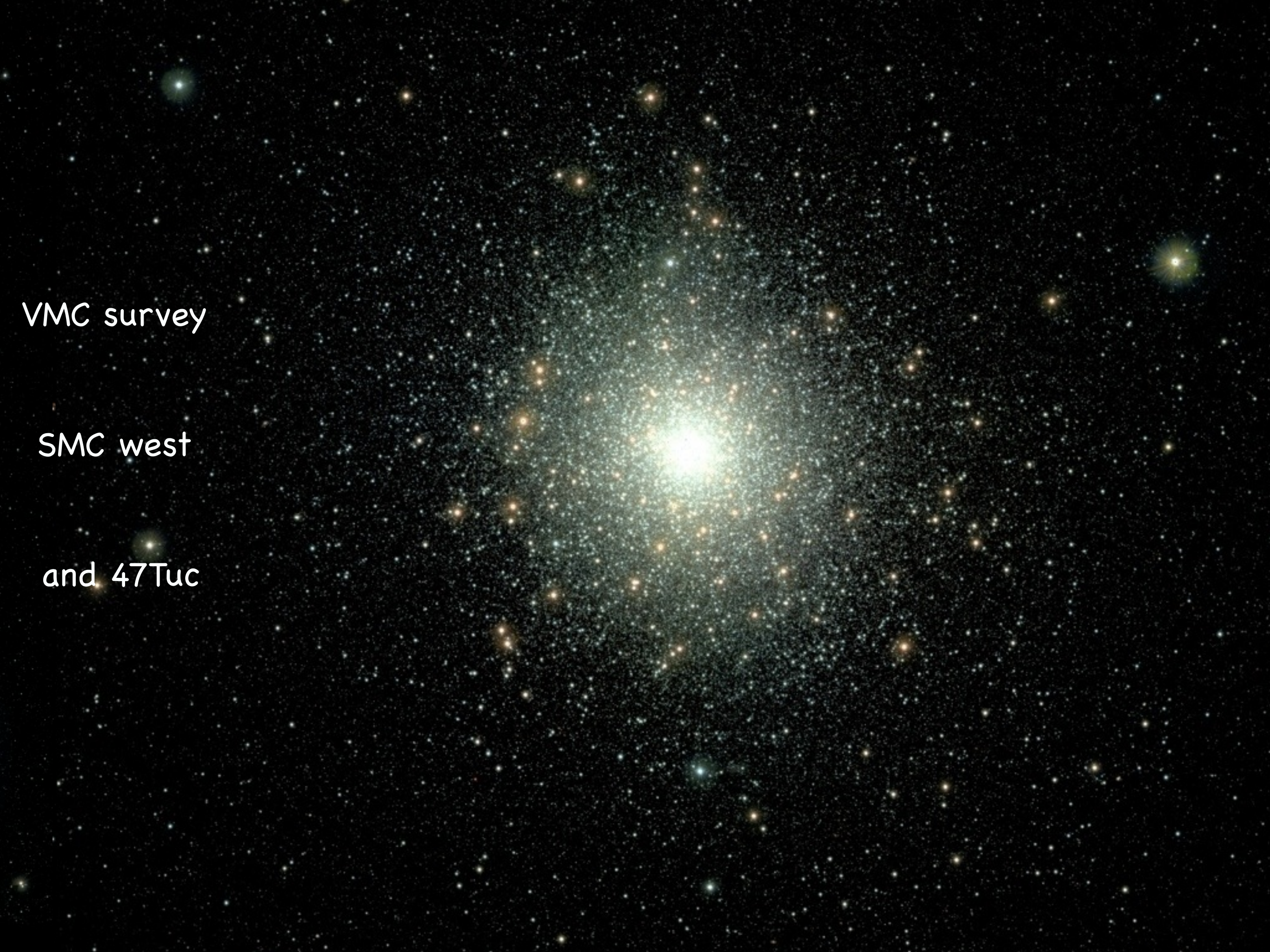


VMC survey

SMC west

and 47Tuc





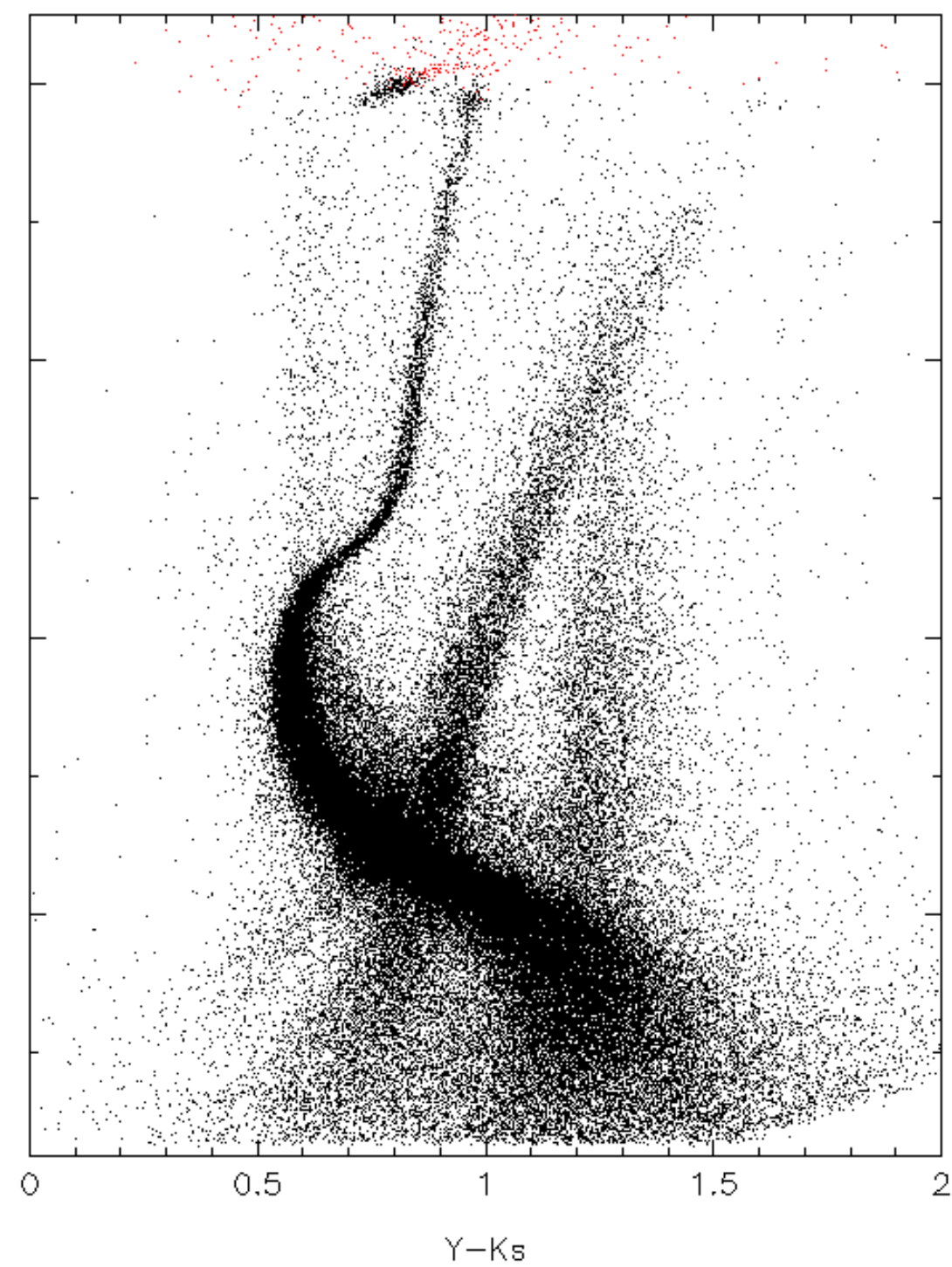
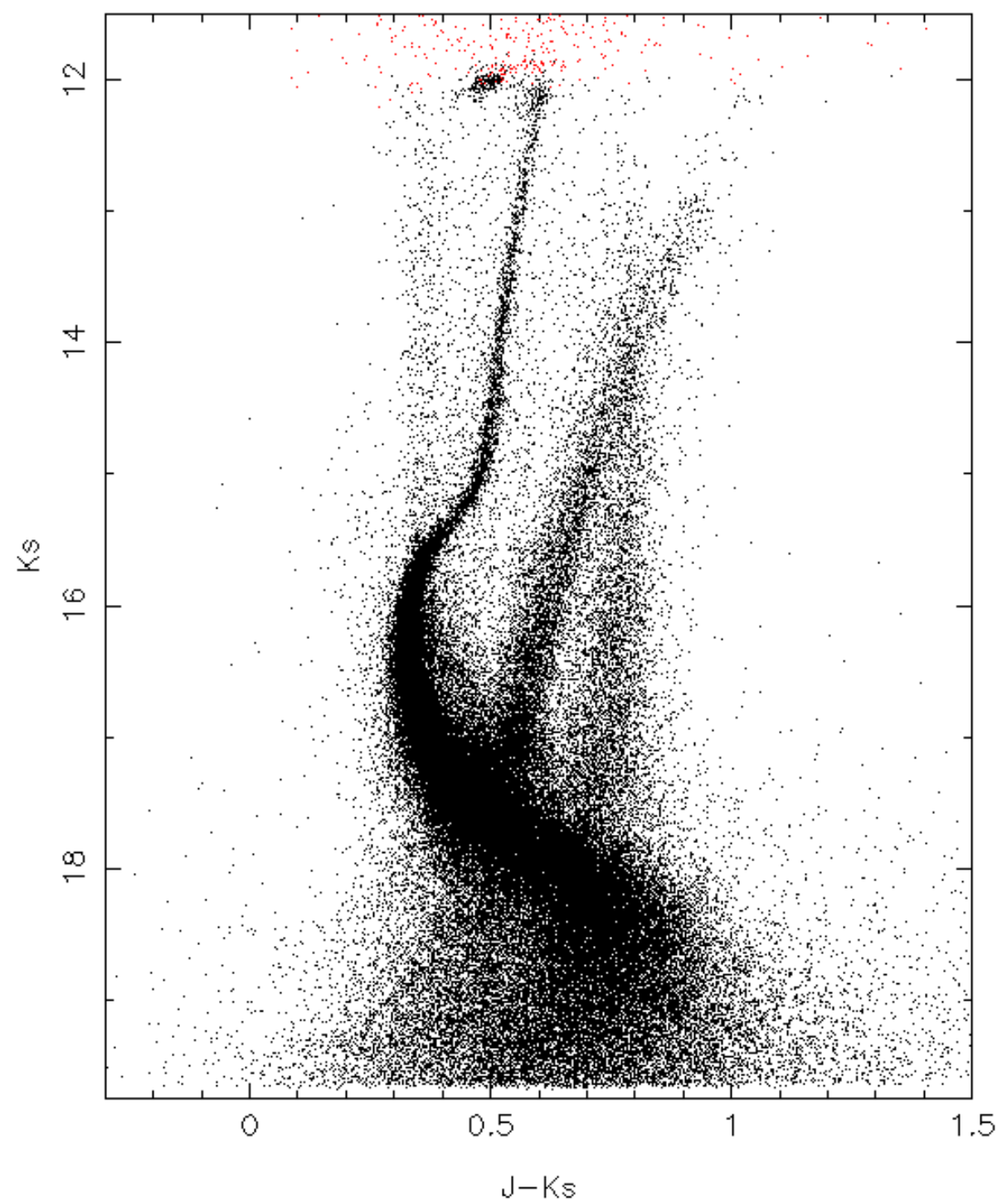
VMC survey

SMC west

and 47Tuc



# VMC survey SMC west & 47Tuc -> pipeline CMD





# CASU processing for large scale spectroscopic survey facilities

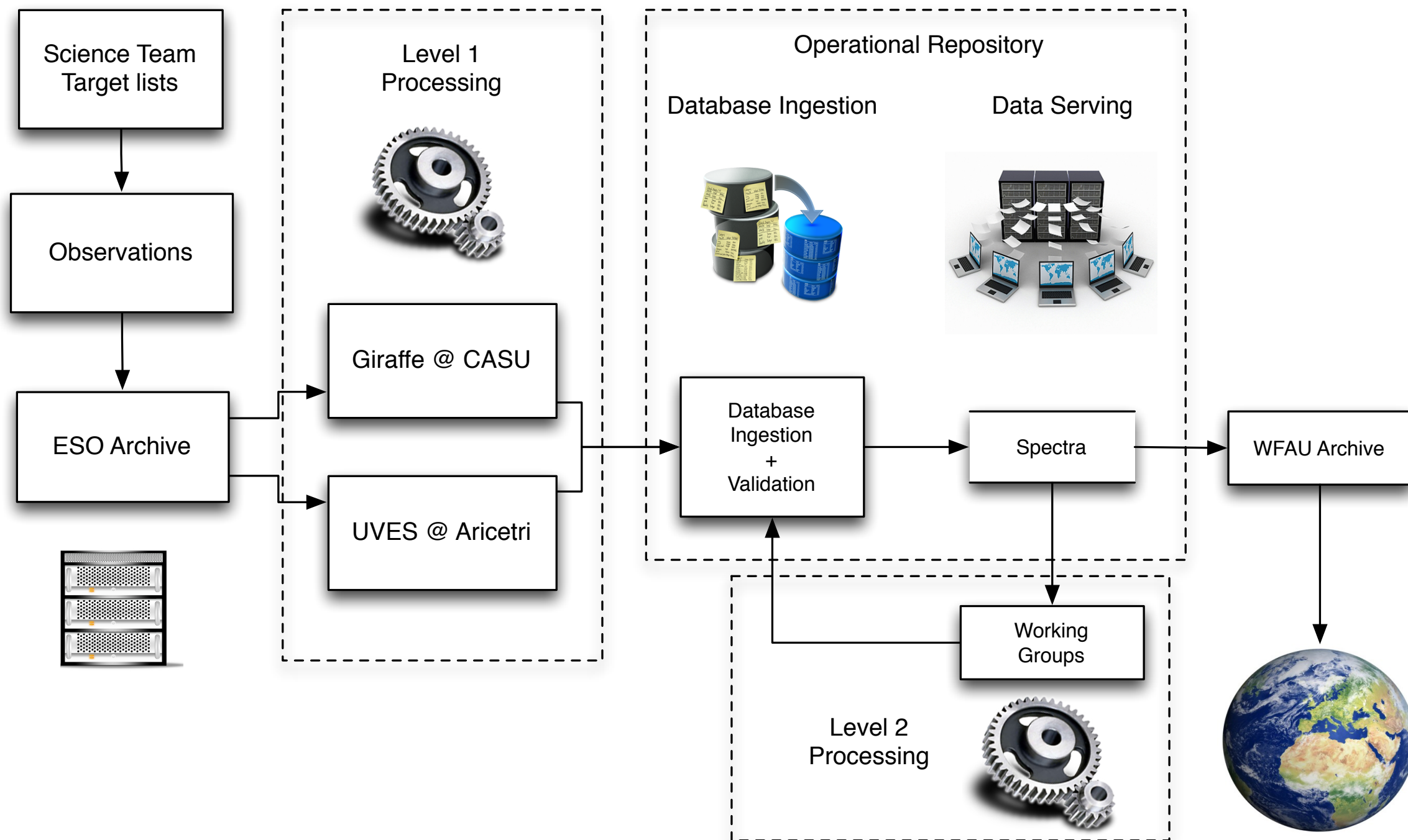


Telescope, Diameter	WHT, 4.2m
Field of View	2 degrees
Number of fibres	960
Fibre size	1.3"
Small IFUs	20 x 11" x 12" (1.3" spaxels)
Large IFU	1.3' x 1.5' (2.6" spaxels)
Low resolution	5750 (4000 - 7250)
Low resolution coverage	3660 - 9590
High resolution	20000 (15000 - 25000)
High resolution coverage	4040-4650, 4730-5420, 5950-6850
Spectra / night	~ 20000 - 30000

Telescope, Diameter	VISTA, 3.7m
Field of View	2.5 degrees
Number of fibres	1600 (LR) + 800 (HR)
Fibre size	1.45"
Low resolution	5000
Low resolution coverage	3950 - 8950
High resolution	18000
High resolution coverage	3950-4400, 5000-5550, 6050-6750
Spectra / night	~ 50000

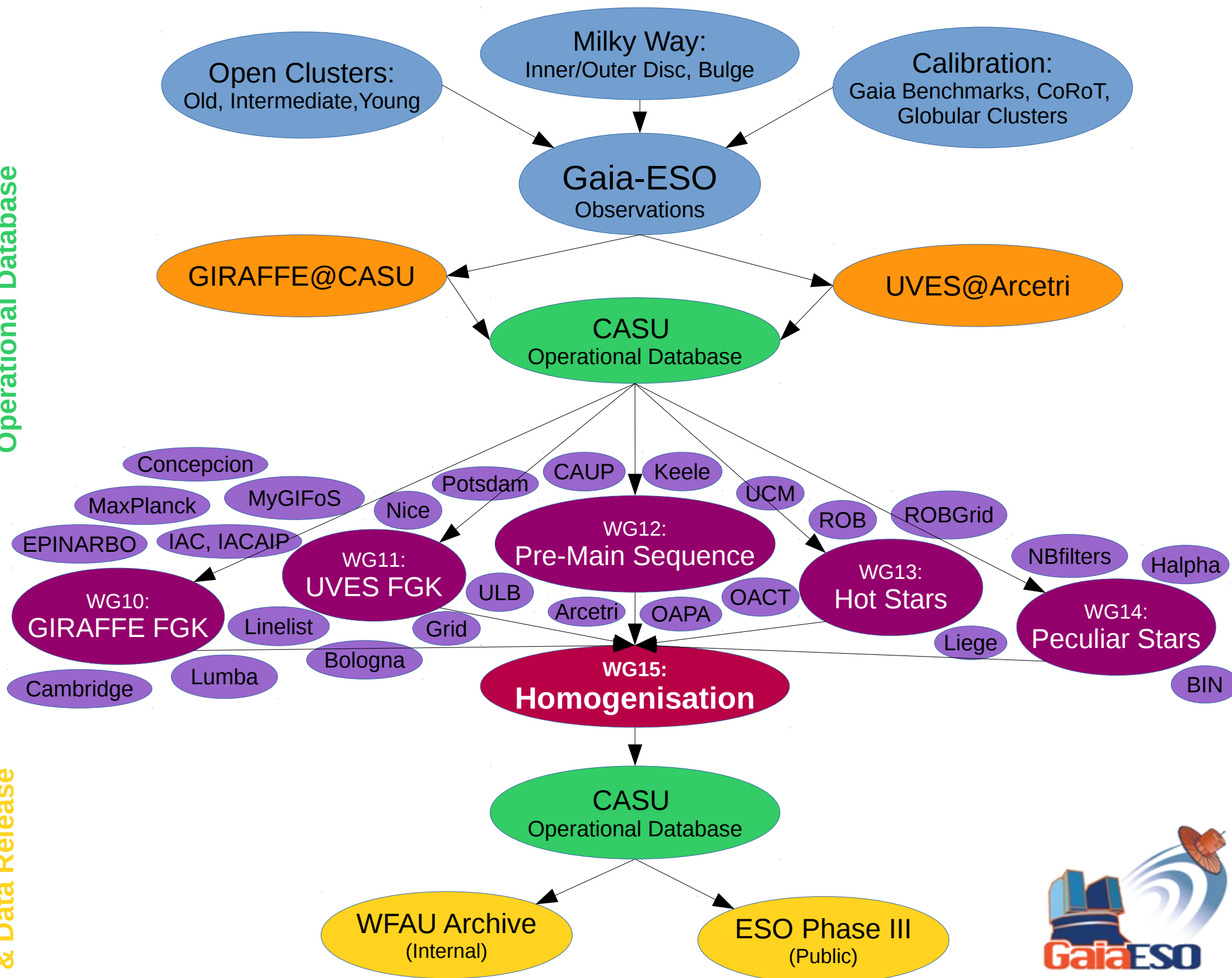


# GAIA-ESO Data Flow





Target Selection  
Spectral Reduction, Radial Velocities & Operational Database  
Stellar Parameters & Abundances  
Operational Database & Data Release





# MOS calibration overview



- data characterisation
  - biases, linearity, darks, fibreflats, arcs, twilight flats
  - detector (2D) flats, “salsa” flats & arcs → PSFs
- science calibration
  - wavelength calibration → arcs, skylines, telluric lines
  - relative flux calibration → fibreflats; spectral types
  - absolute flux calibration → Gaia BP/RP; photometry
- survey verification & calibration
  - Gaia benchmark stars → fairly infrequent
  - RV standards → ditto
  - standard fields e.g. Kepler, open clusters ...

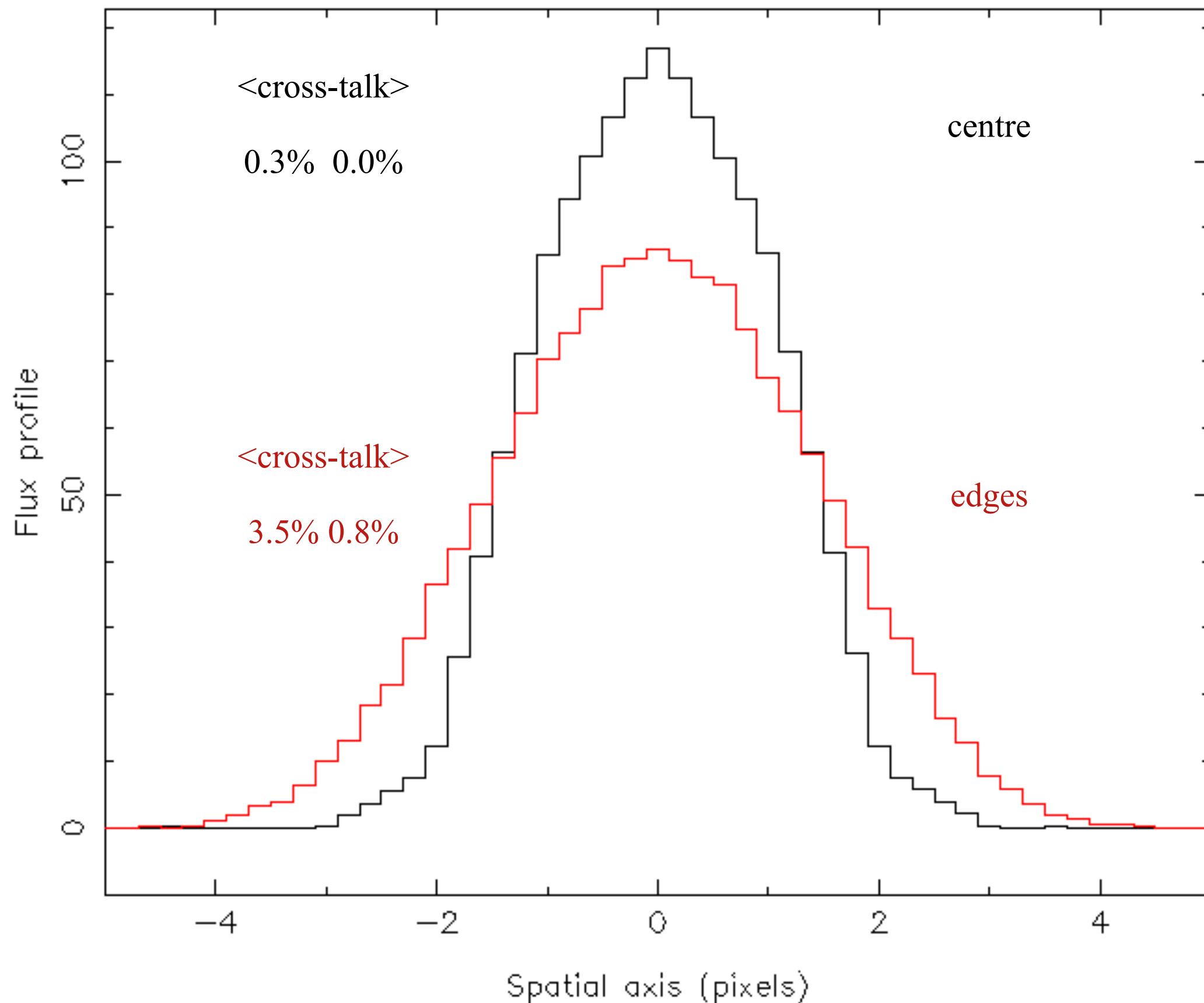


# MOS quality issues overview

- extraction issues
  - what is “optimal” extraction in a close-packed MOS
  - how to characterise (PSFs) and deal with crosstalk
  - handling cosmic-rays single-v-multiple exposures
- spectrum issues
  - telluric lines → modelling MOLECFIT; reference stars
  - sky emission line removal
  - sky continuum and scattered light removal
- quality measures
  - quality of various processing/extraction stages
  - estimates of s:n, flux levels continuum/emission lines
  - standard fields for inter-survey comparison



# WEAVE simulations - derived subsampled PSF

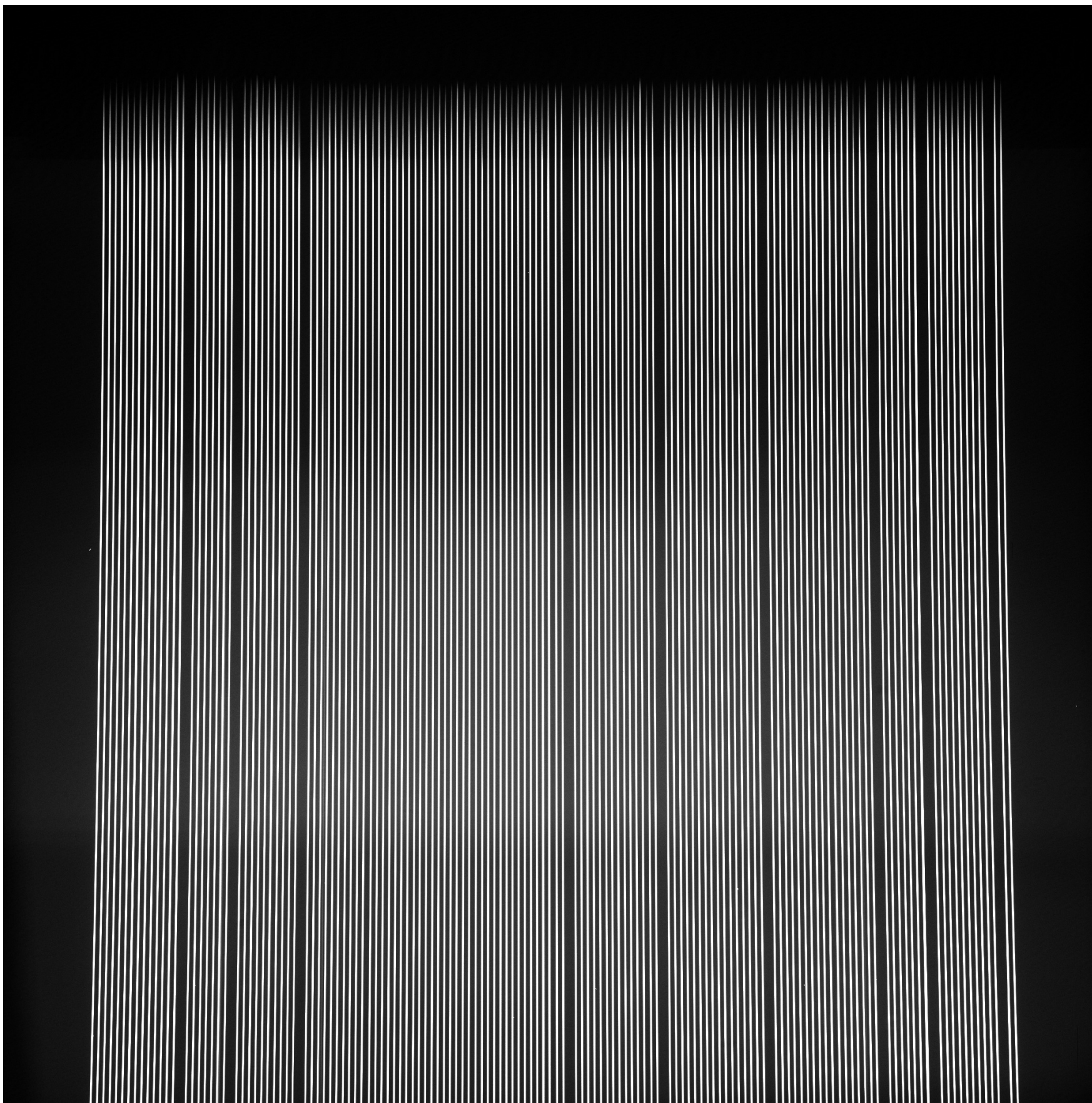




WHT+AF2  
WYFFOS

fibre profiles

$$\psi_k(x, y)$$





# Fibre crosstalk and extraction

## (original cunning plan)

data model 
$$d(x, y) = \sum_{k=1}^m S_k(y) \psi_k(x, y) + \epsilon(x, y)$$
 fibre profiles  $\psi_k(x, y)$

LS solution 
$$\begin{bmatrix} C_{11}(y) & C_{12}(y) & C_{13}(y) & \cdots & C_{1m}(y) \\ C_{21}(y) & C_{22}(y) & C_{23}(y) & \cdots & C_{2m}(y) \\ C_{31}(y) & C_{32}(y) & C_{33}(y) & \cdots & C_{3m}(y) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ C_{m1}(y) & C_{m2}(y) & C_{m3}(y) & \cdots & C_{mm}(y) \end{bmatrix} \begin{bmatrix} S_1(y) \\ S_2(y) \\ S_3(y) \\ \cdots \\ \cdots \\ S_m(y) \end{bmatrix} = \begin{bmatrix} D_1(y) \\ D_2(y) \\ D_3(y) \\ \cdots \\ \cdots \\ D_m(y) \end{bmatrix}$$

fibre-fibre  
overlap integral

$$C_{ij}(y) = \langle \psi_i(x, y) | w(x, y) | \psi_j(x, y) \rangle_{\mathbf{x}}$$

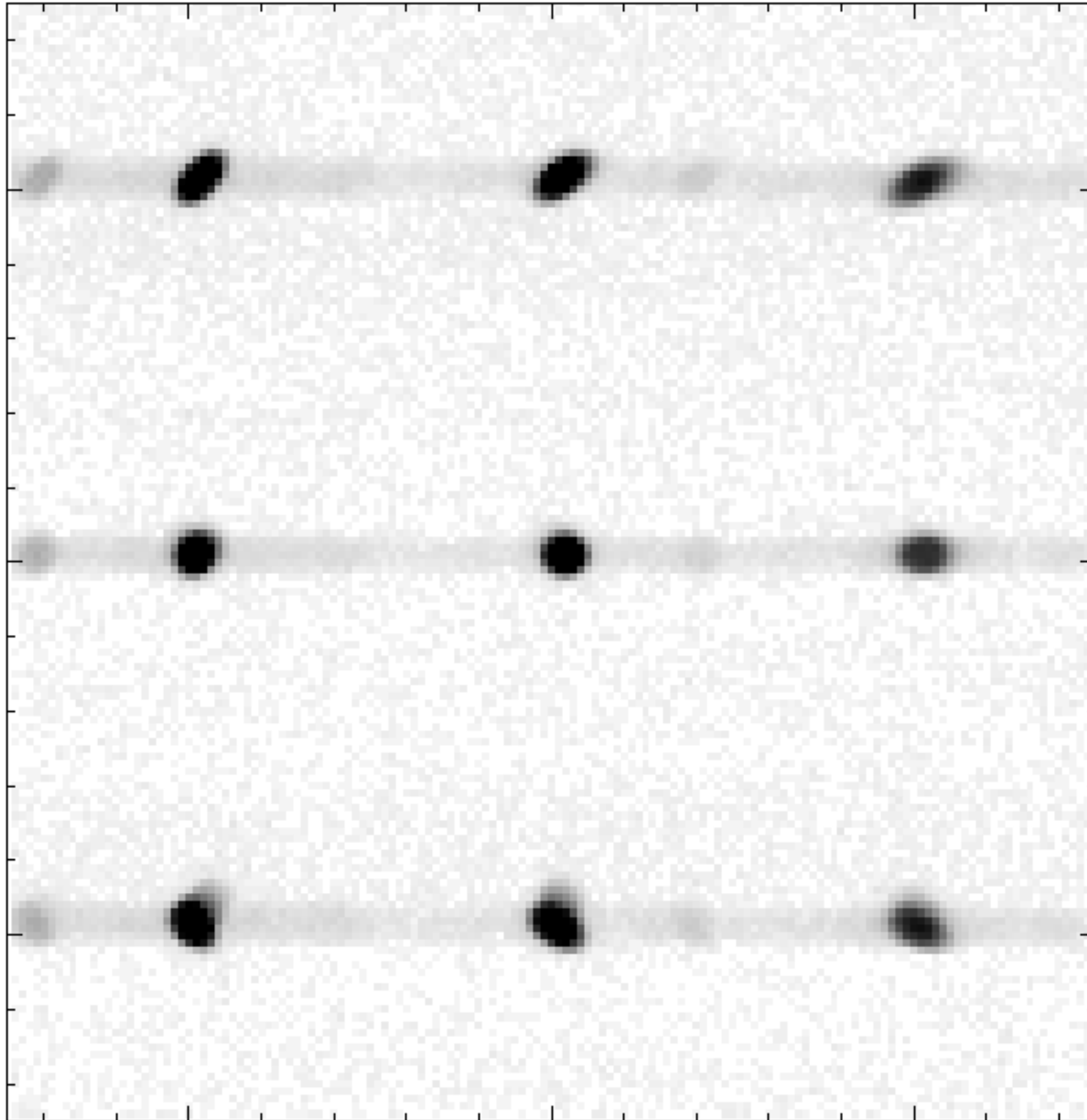
data-fibre  
overlap integral

$$D_i(y) = \langle d(x, y) | w(x, y) | \psi_i(x, y) \rangle_{\mathbf{x}}$$

(but probably need to go full 2D!)



# HERMES – arc line PSFs #2



fin