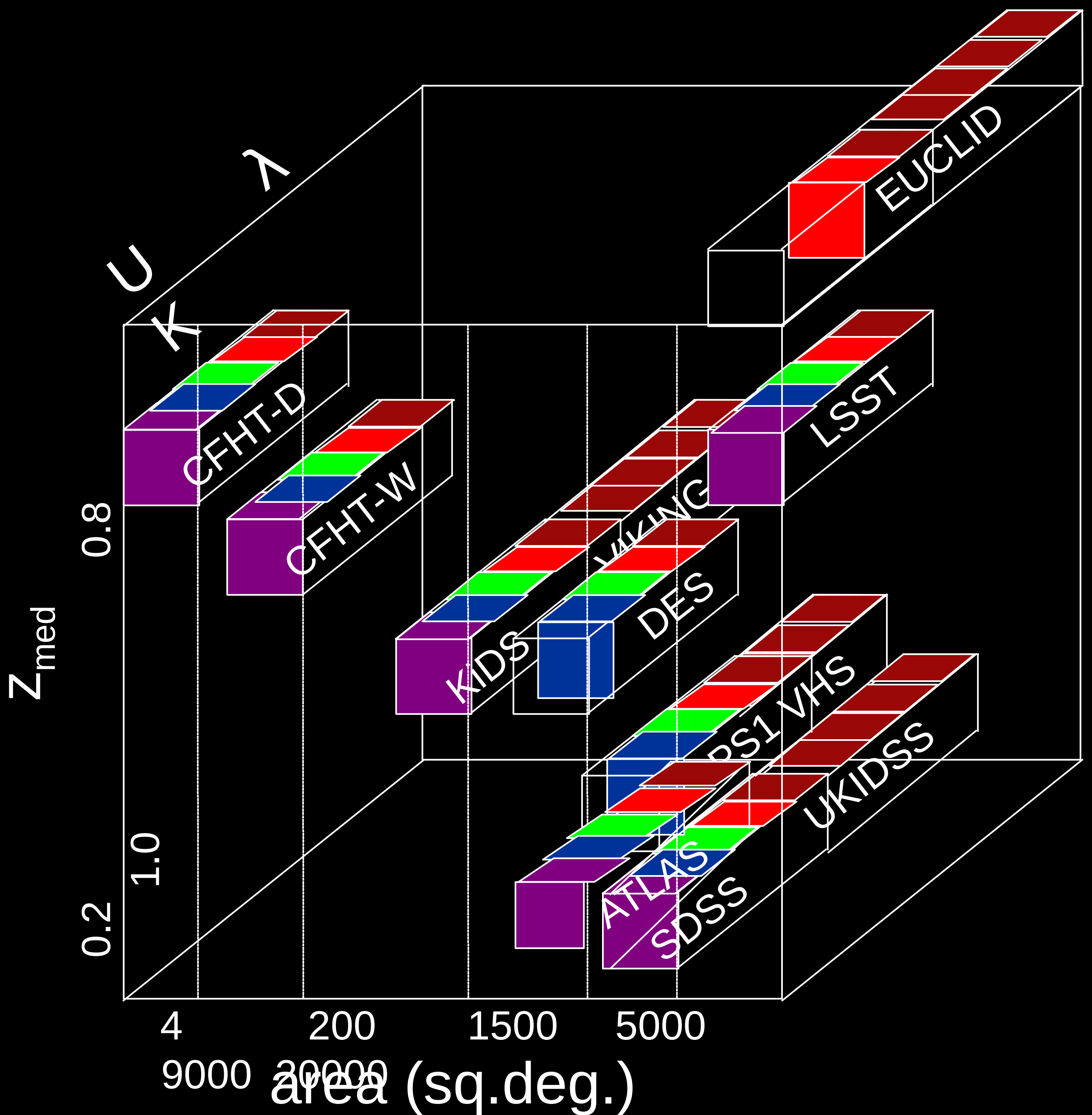
# Virtual Observations 2016
# -> Astronomical Data Science 2017

## Survey Photometry & Survey Astrometry
## enabled by
## Information systems

- Survey Photometry and Astro-WISE
  - Recap of photometry
  - KiDS survey photometry and Astro-WISE
- Global Astrometry and AGIS
  - Recap of astrometry
  - Gaia global astrometry and AGIS
- Conclusion

# Imaging surveys

- Area covered
- Median redshift
- Image quality
- Wavelength

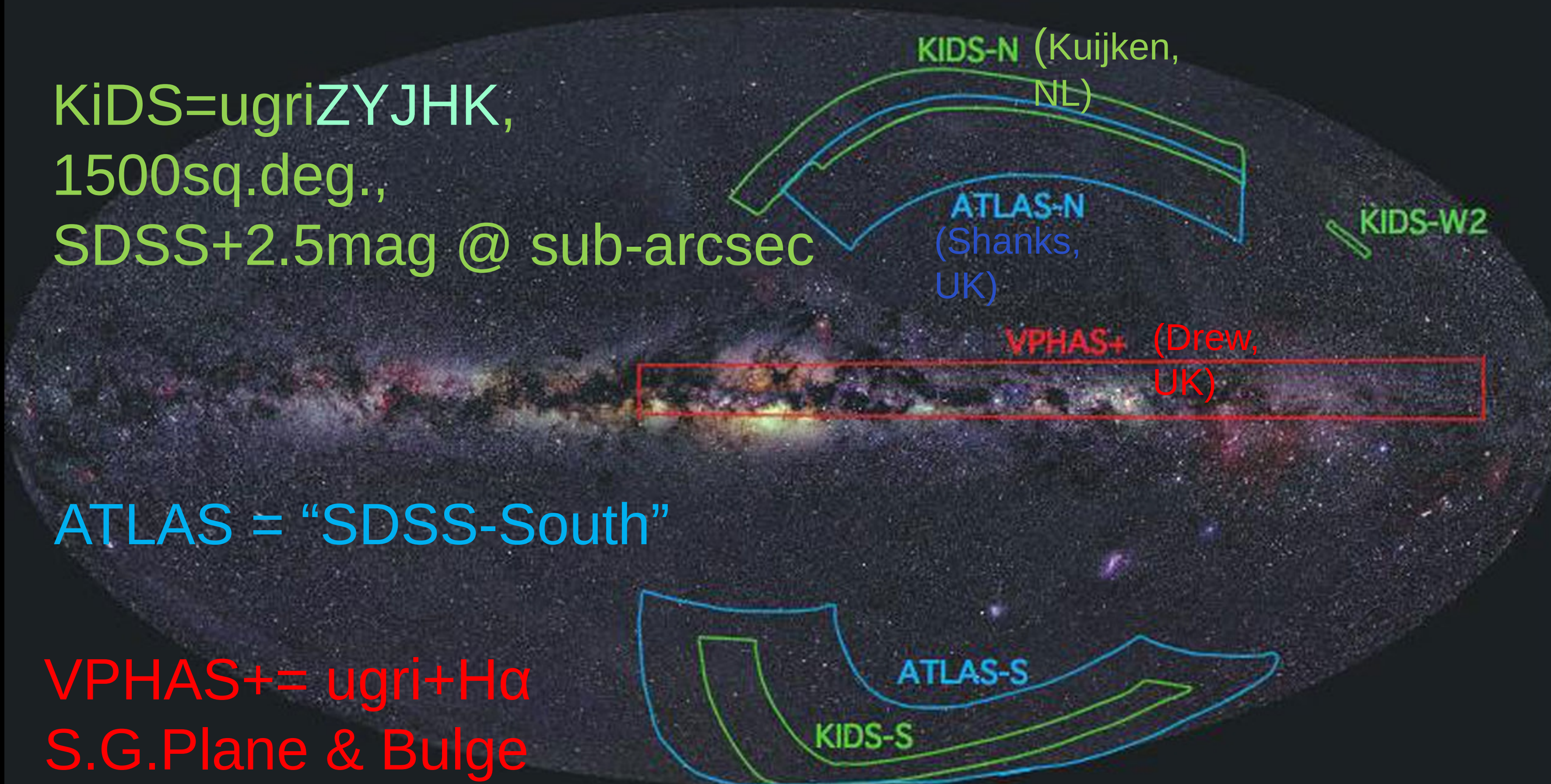


# 3 ESO public surveys (~1000 nights)

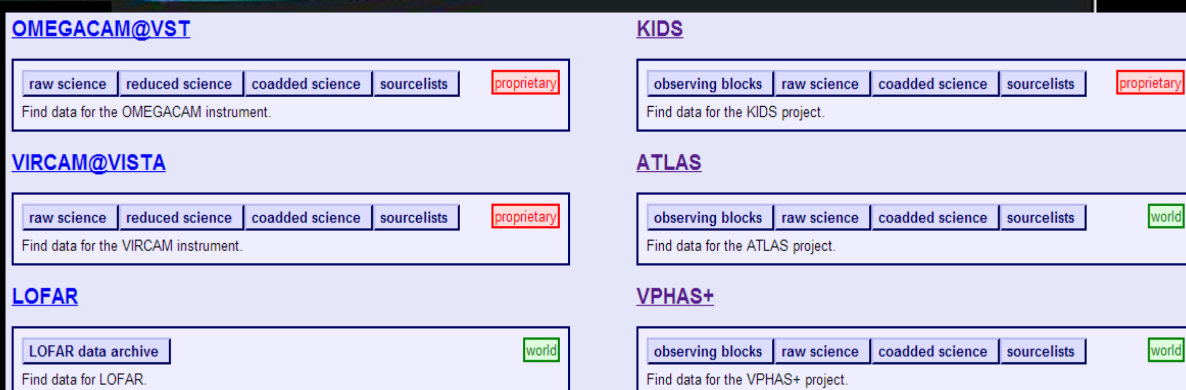
KiDS=ugriZYJHK,
1500sq.deg.,
SDSS+2.5mag @ sub-arcsec

ATLAS = "SDSS-South"

VPHAS+= ugri+Hα
S.G.Plane & Bulge



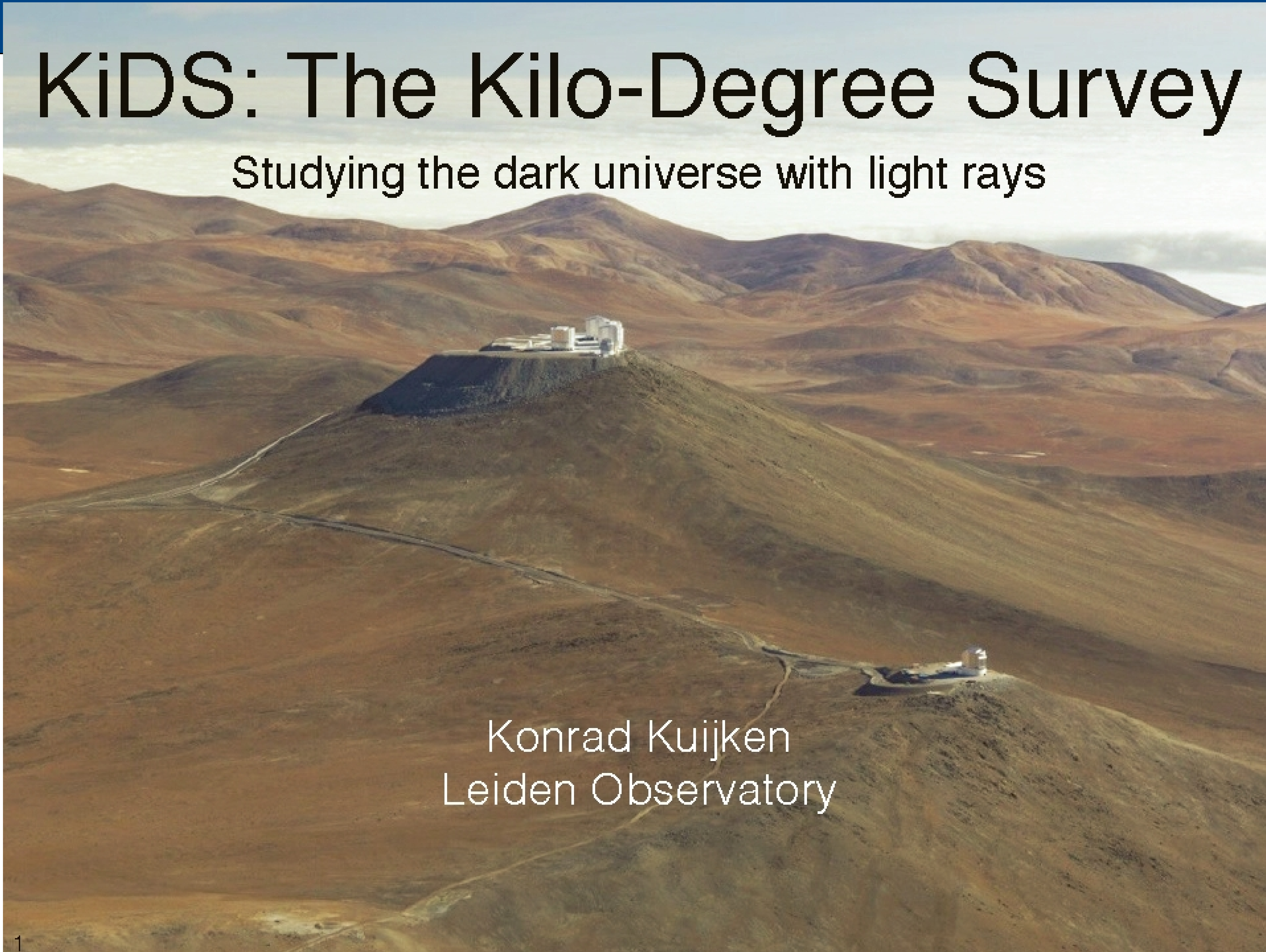

Available at each NOVA institute ASTRO WISE :




university of groningen / faculty of mathematics and natural sciences / kapteyn astronomical institute

# KiDS: The Kilo-Degree Survey

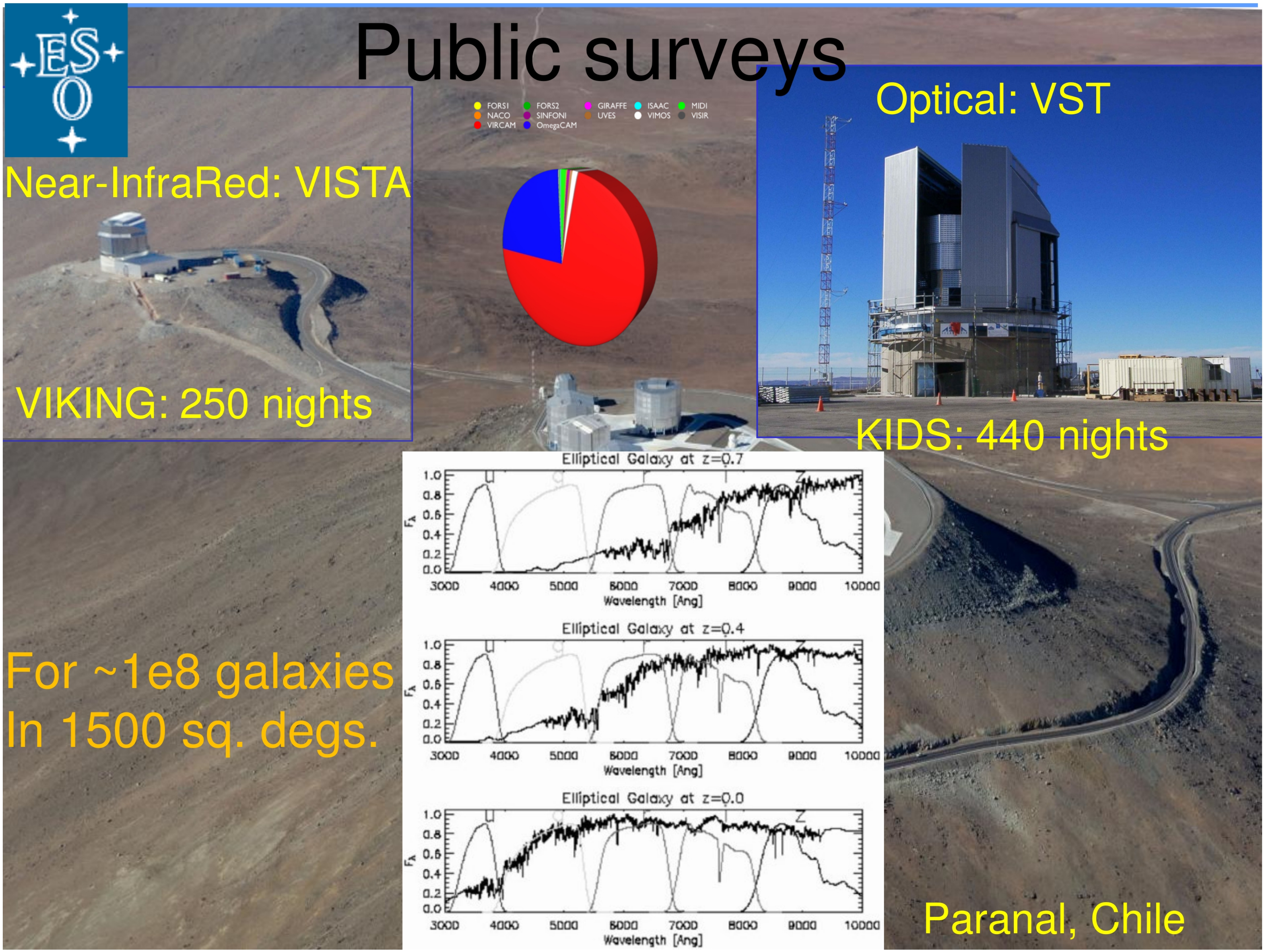
## Studying the dark universe with light rays



Konrad Kuijken
Leiden Observatory

# The KiDS Team

# ESO KiDS and VIKING surveys



# VLT Survey Telescope

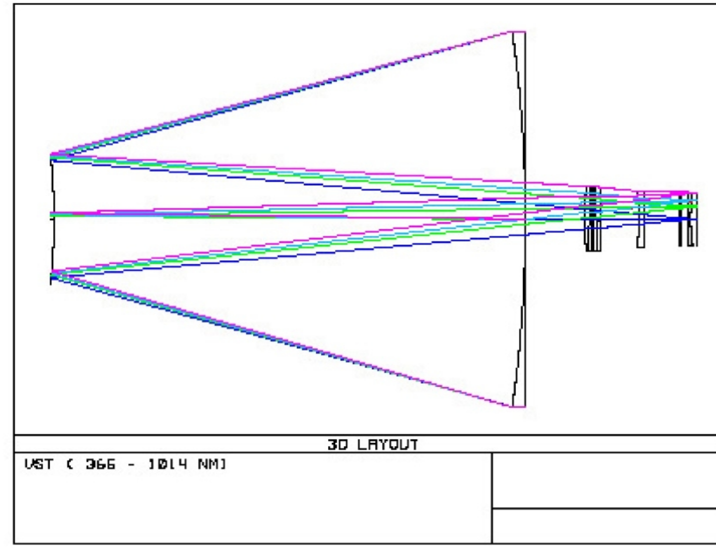




**Fig. 1 - VST complete optical layout of telescope with one lens and the ADC, with a curve dewar window**

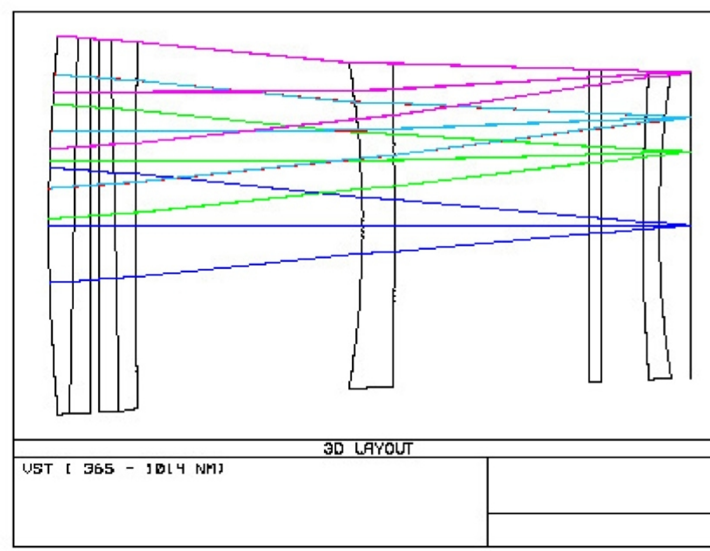
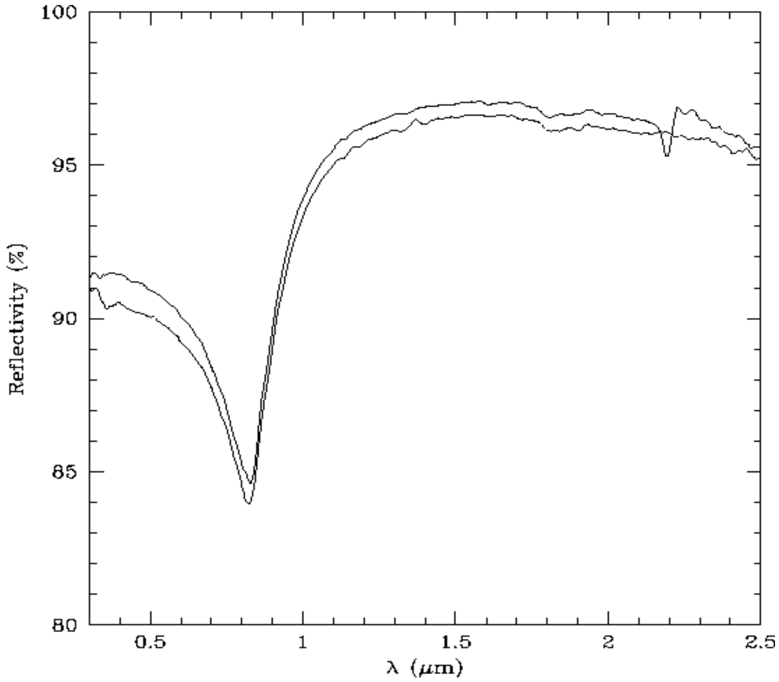
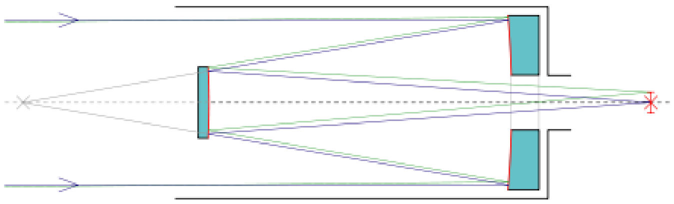


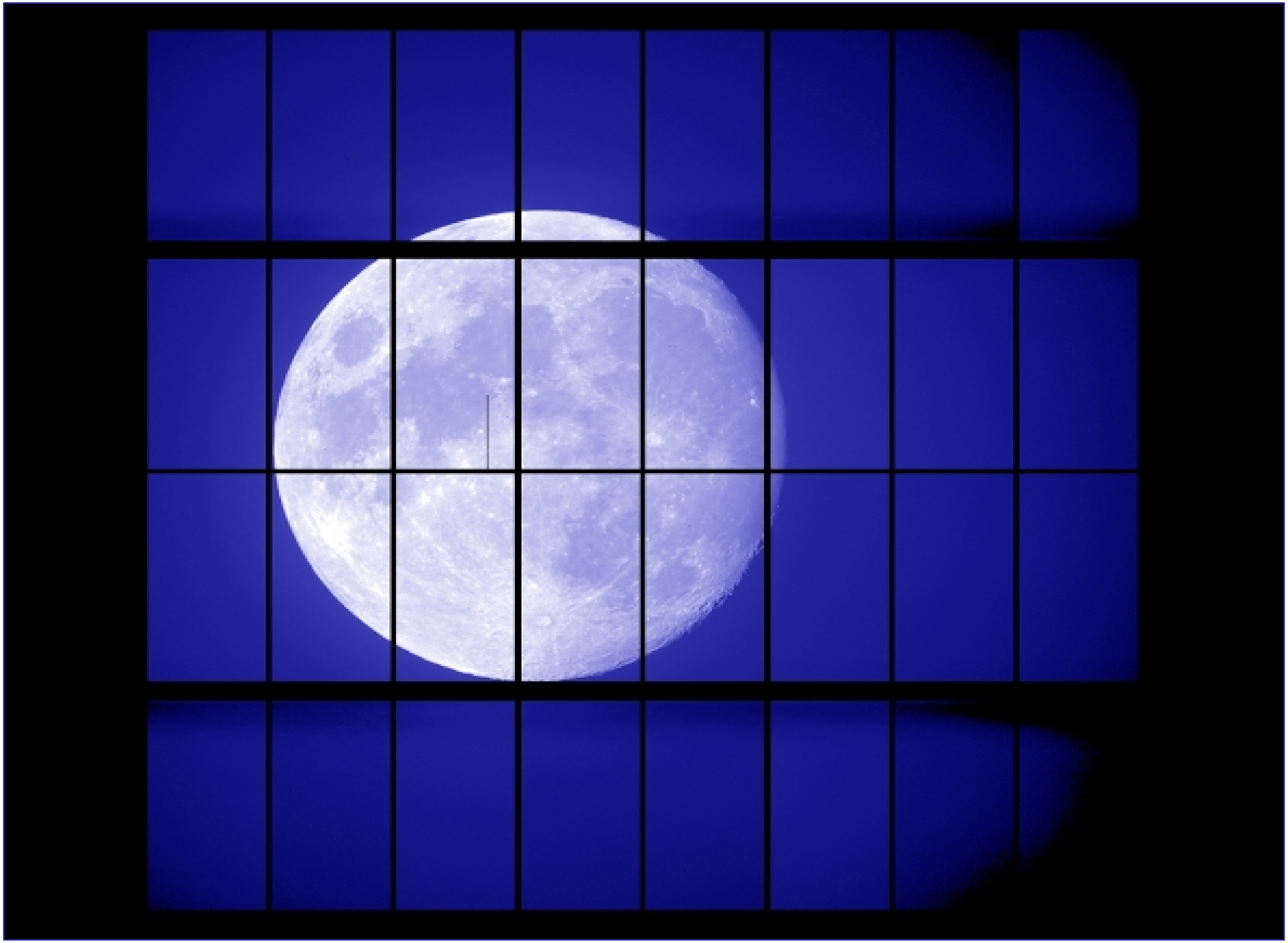

**Fig. 2 - VST zoom of the optical layout of the corrector with one lens and the ADC, with a curve dewar window**

# Optics



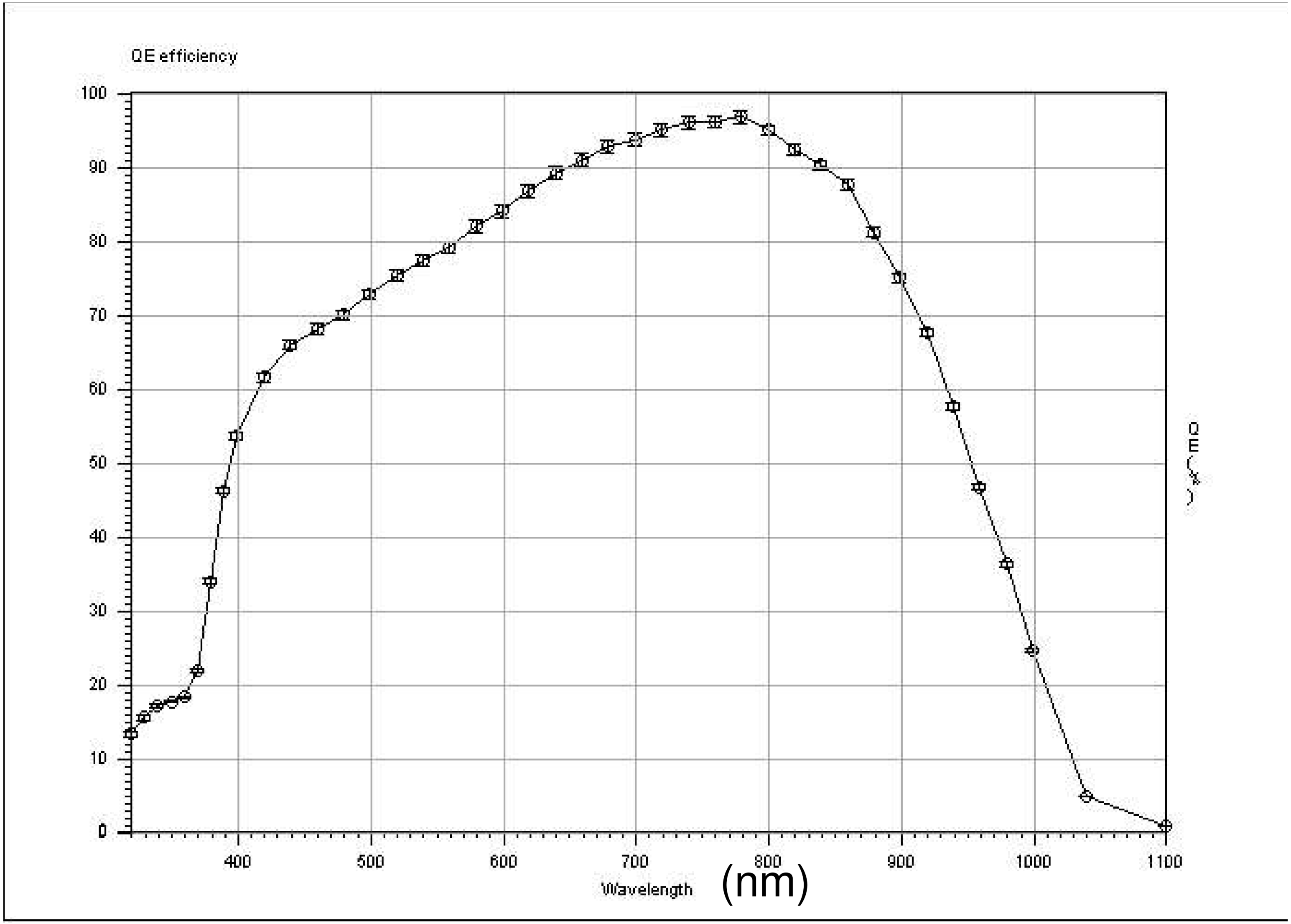



# OmegaCAM 32 detectors

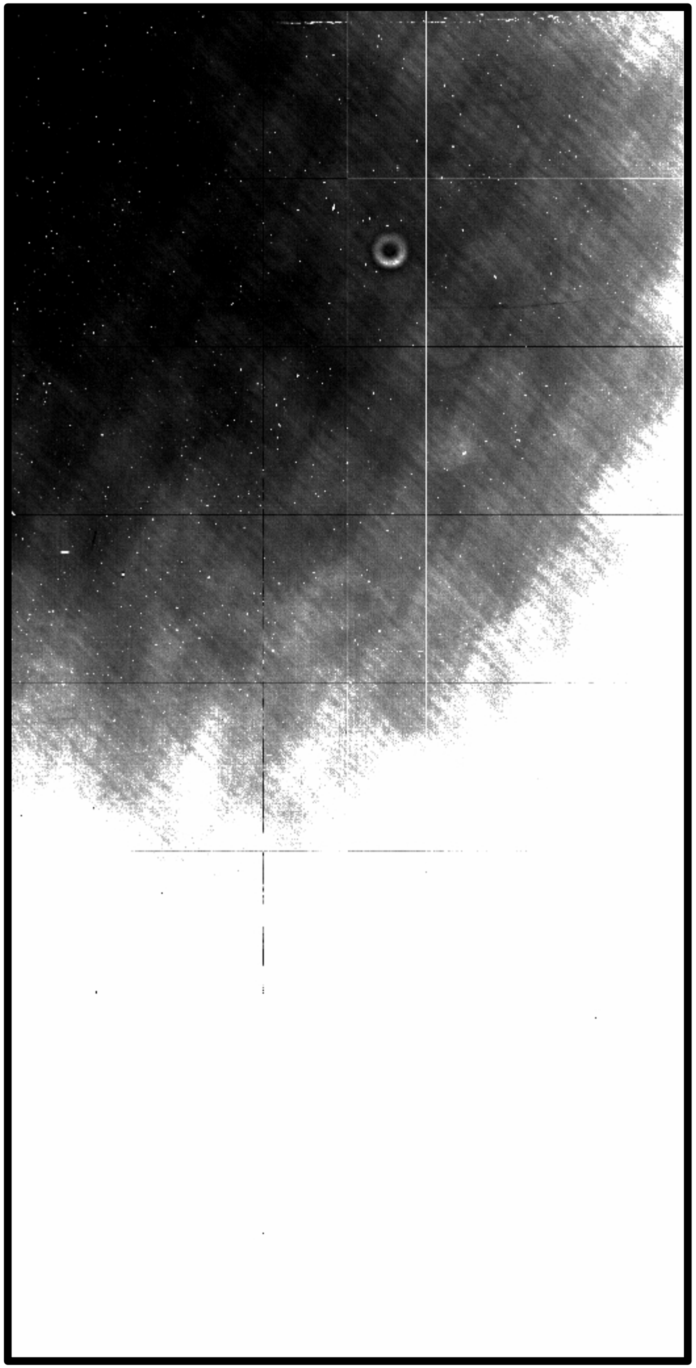


# Detector effects: quantum efficiency

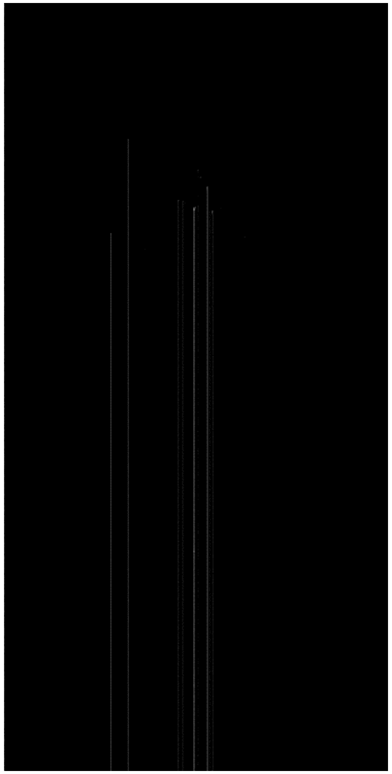
$$m_S = -2.5\ \log \frac{\int_{\lambda_1}^{\lambda_2} \lambda\, f_\lambda\, S_\lambda \mathrm{d}\lambda}{\int_{\lambda_1}^{\lambda_2} \lambda\, S_\lambda \mathrm{d}\lambda} + m_S^0.$$

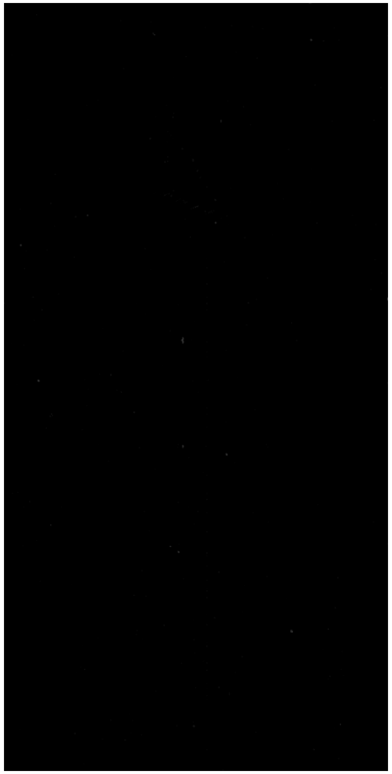


# Detector effects: flat field

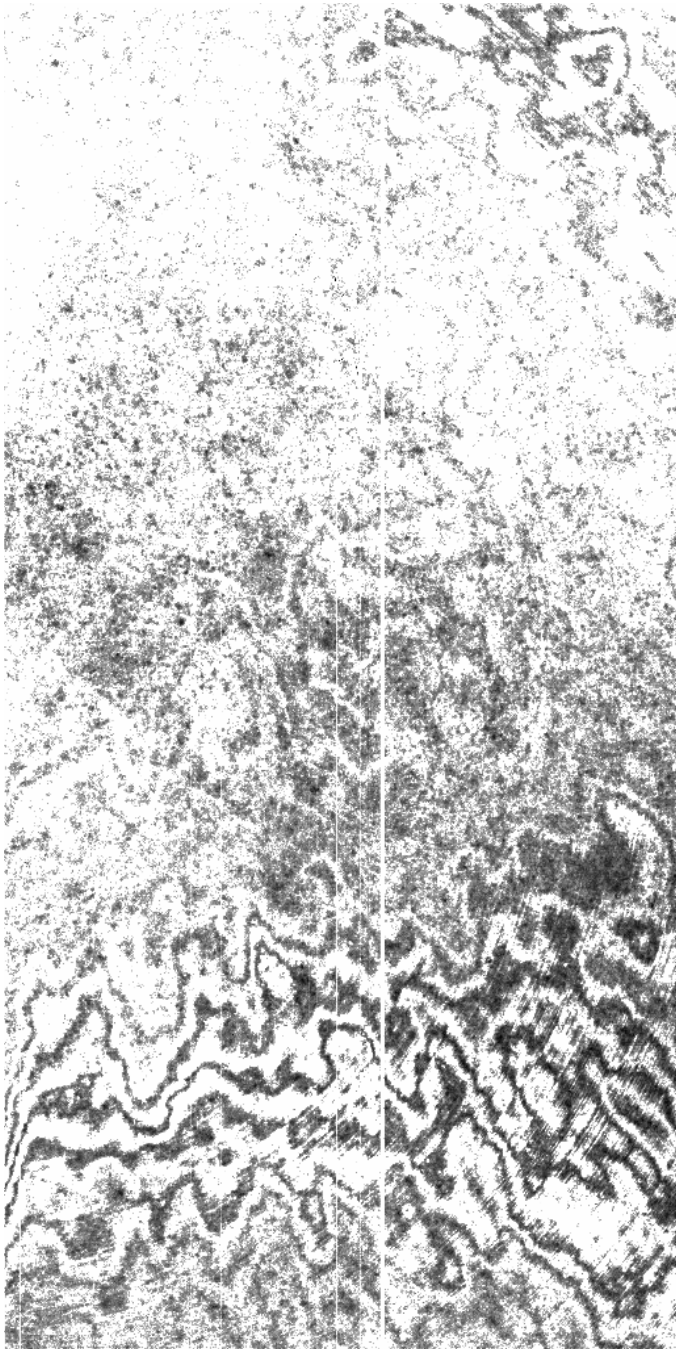


- QE variation per pixel
- Dome-, twilight



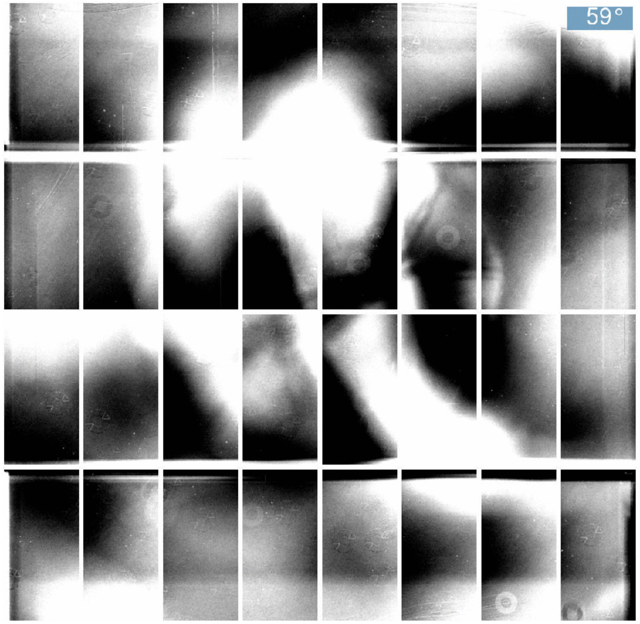


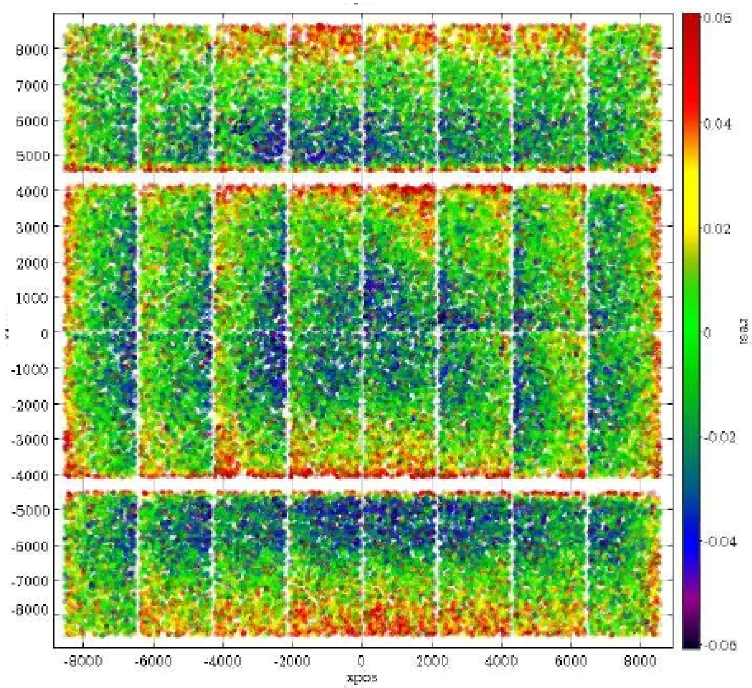
# Detector effects: fringing

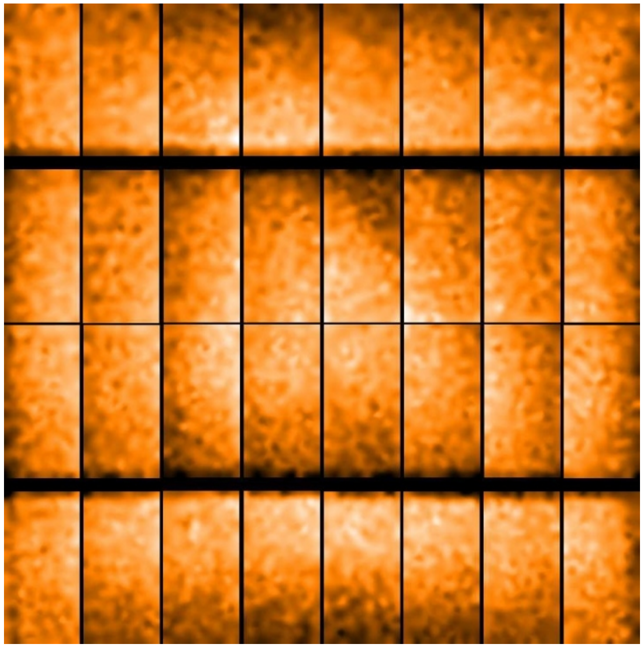


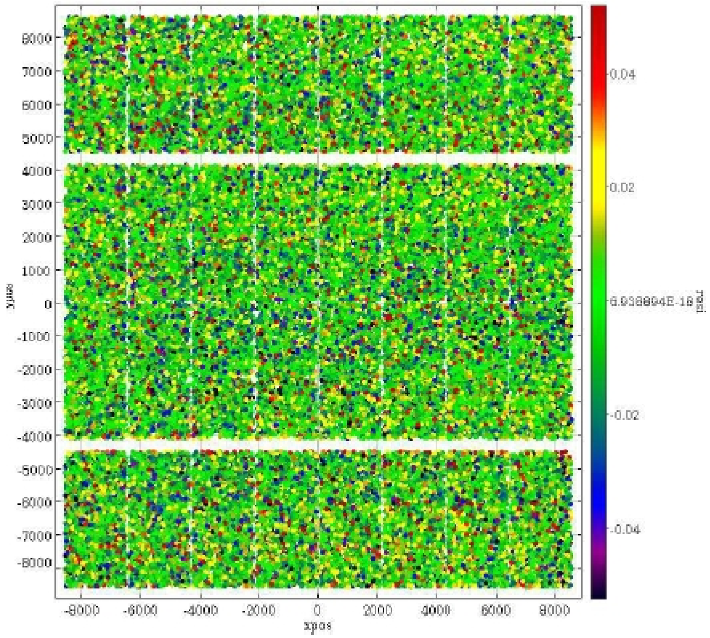
- Interference of nightsky lines with CCD substrate
- Red and IR

# Detector effects: illumination correction

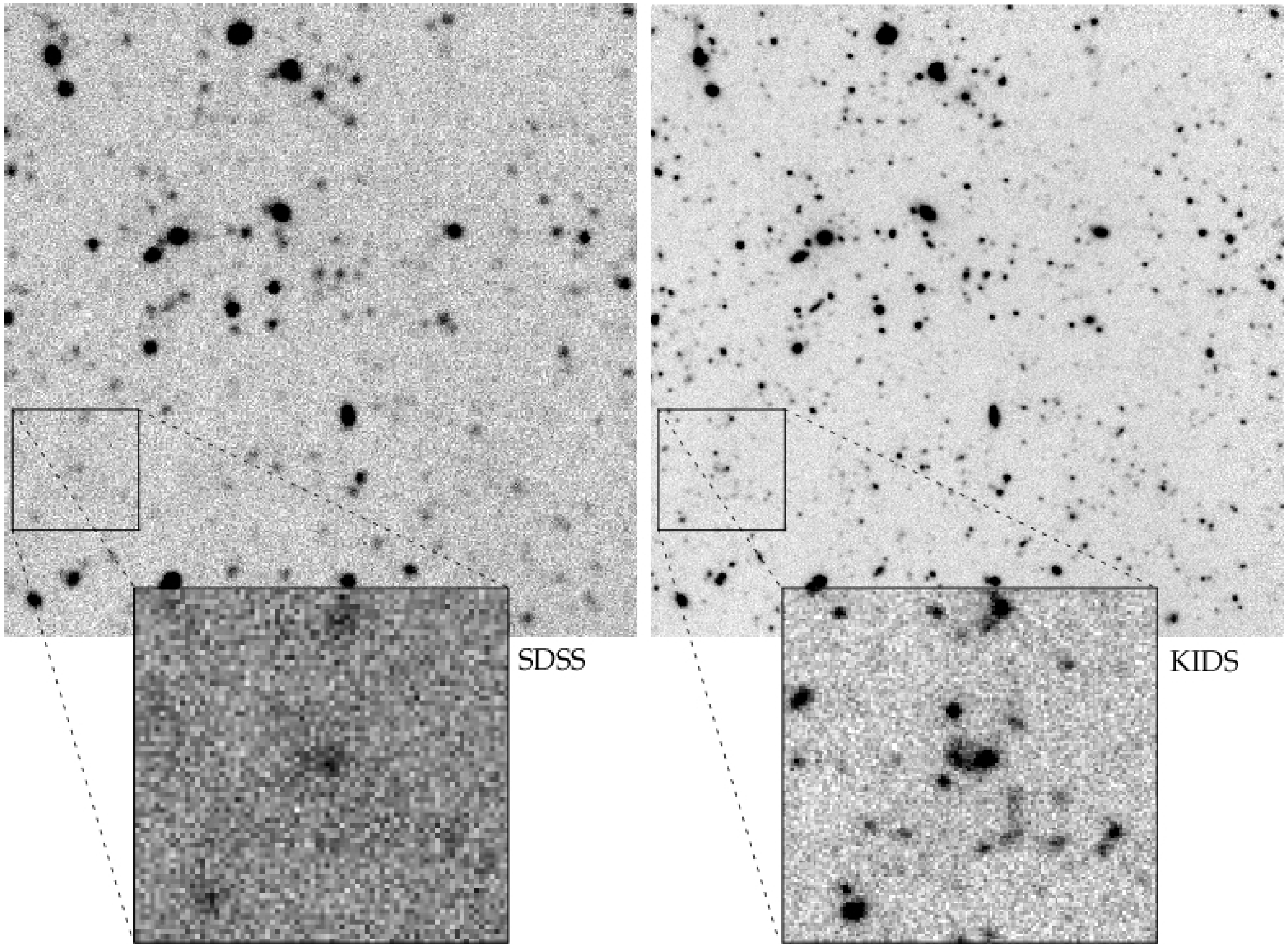




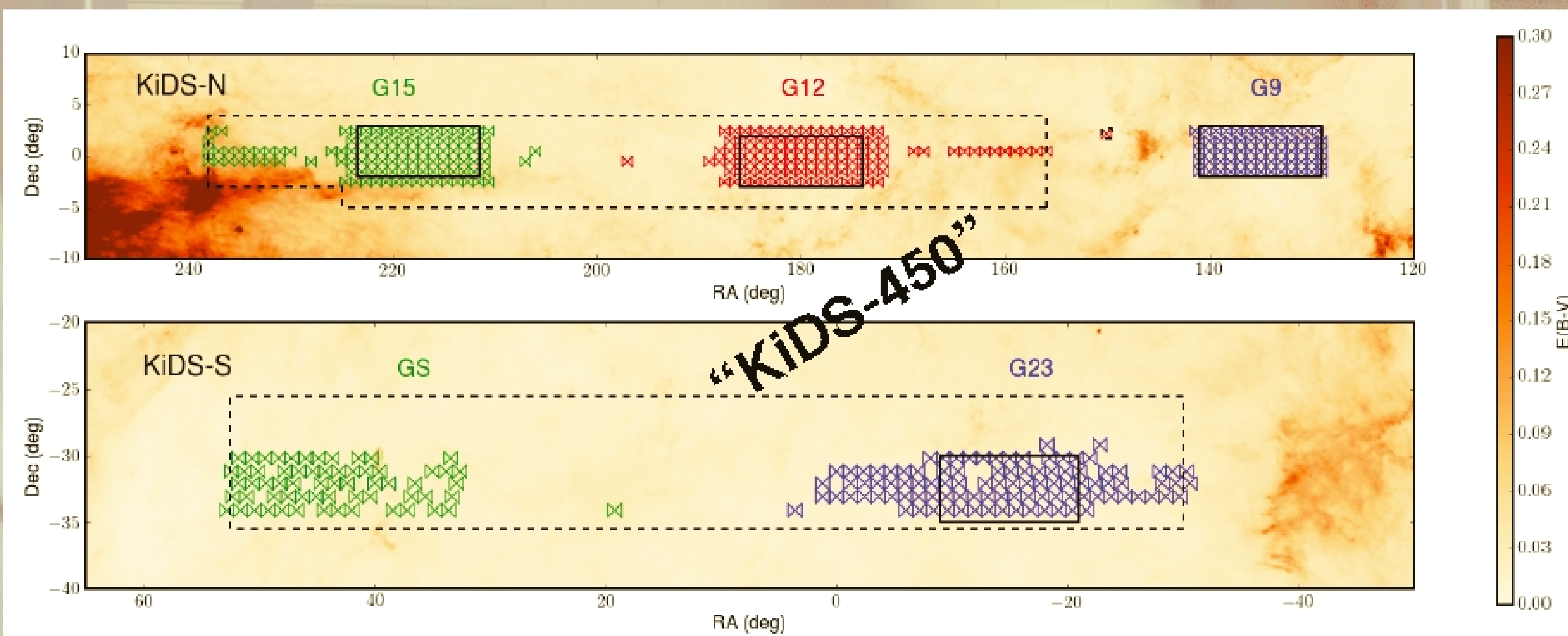
# Kilo Degree Survey with instrumental fingerprint removed



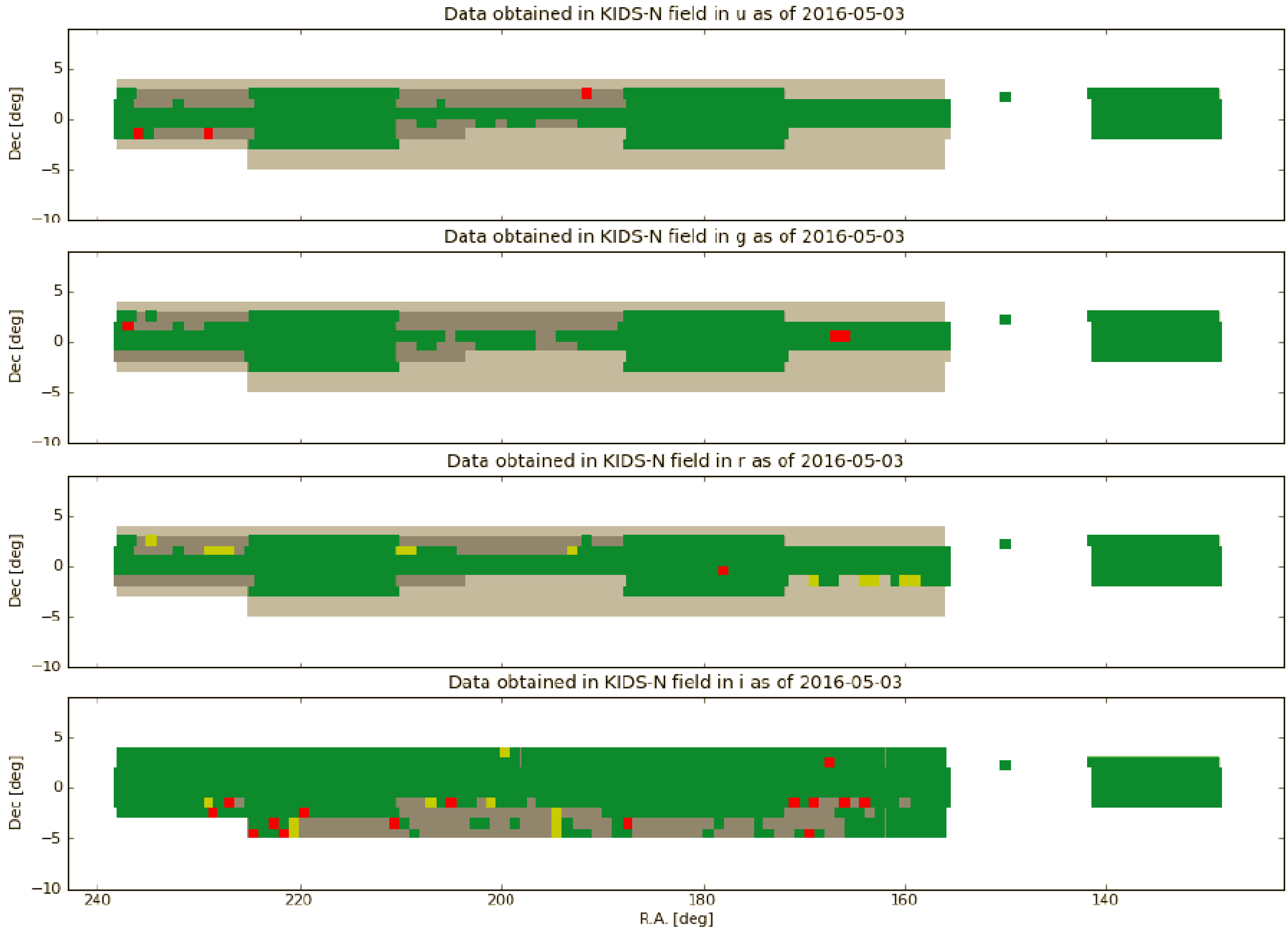
KiDS

# Survey footprint

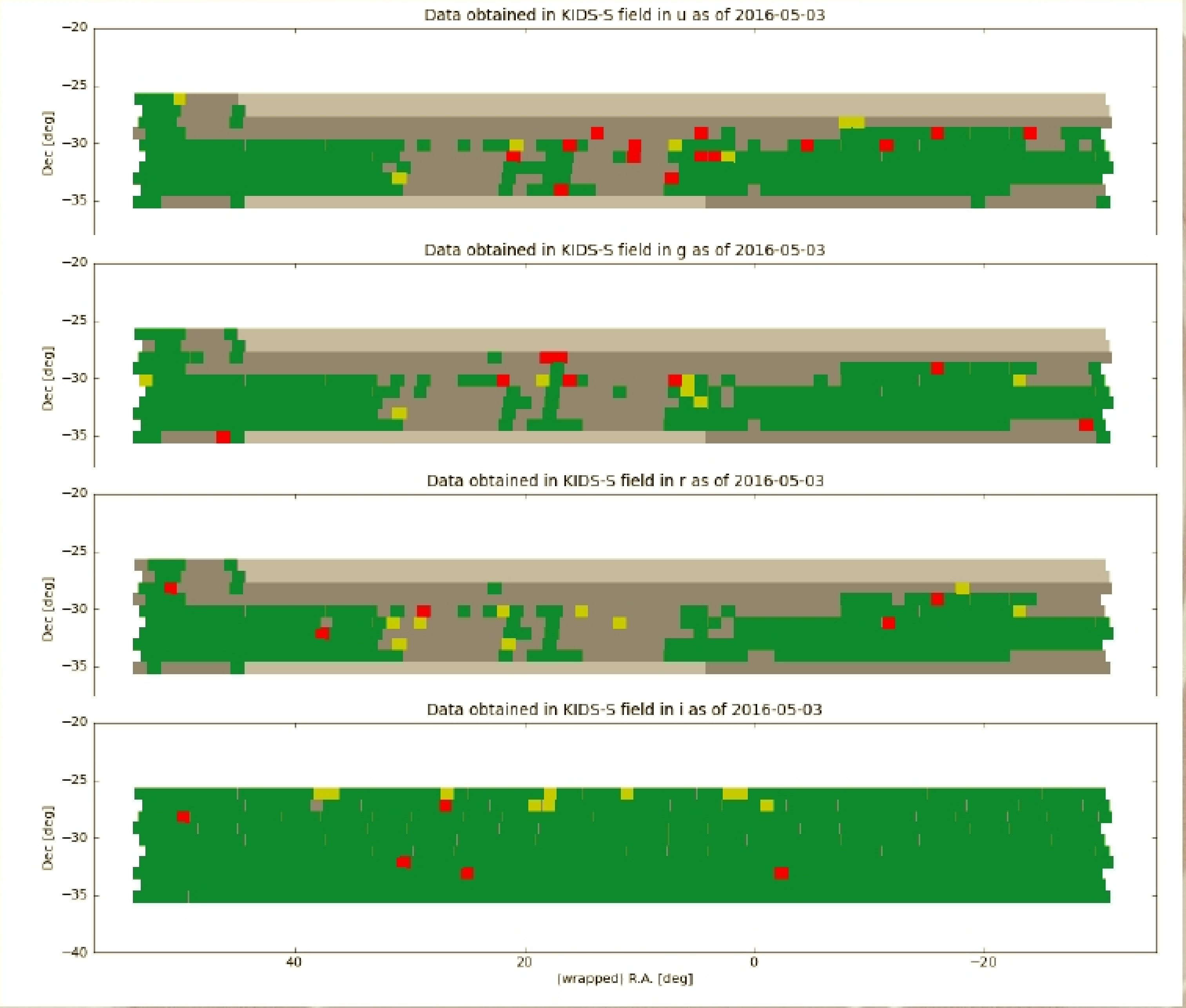
- Two patches, observe year-round
- 1500 sq.deg. total area (677 in hand)
- **matches VIKING and includes GAMA (4/5 patches)**



# Observations per band: KiDS-N



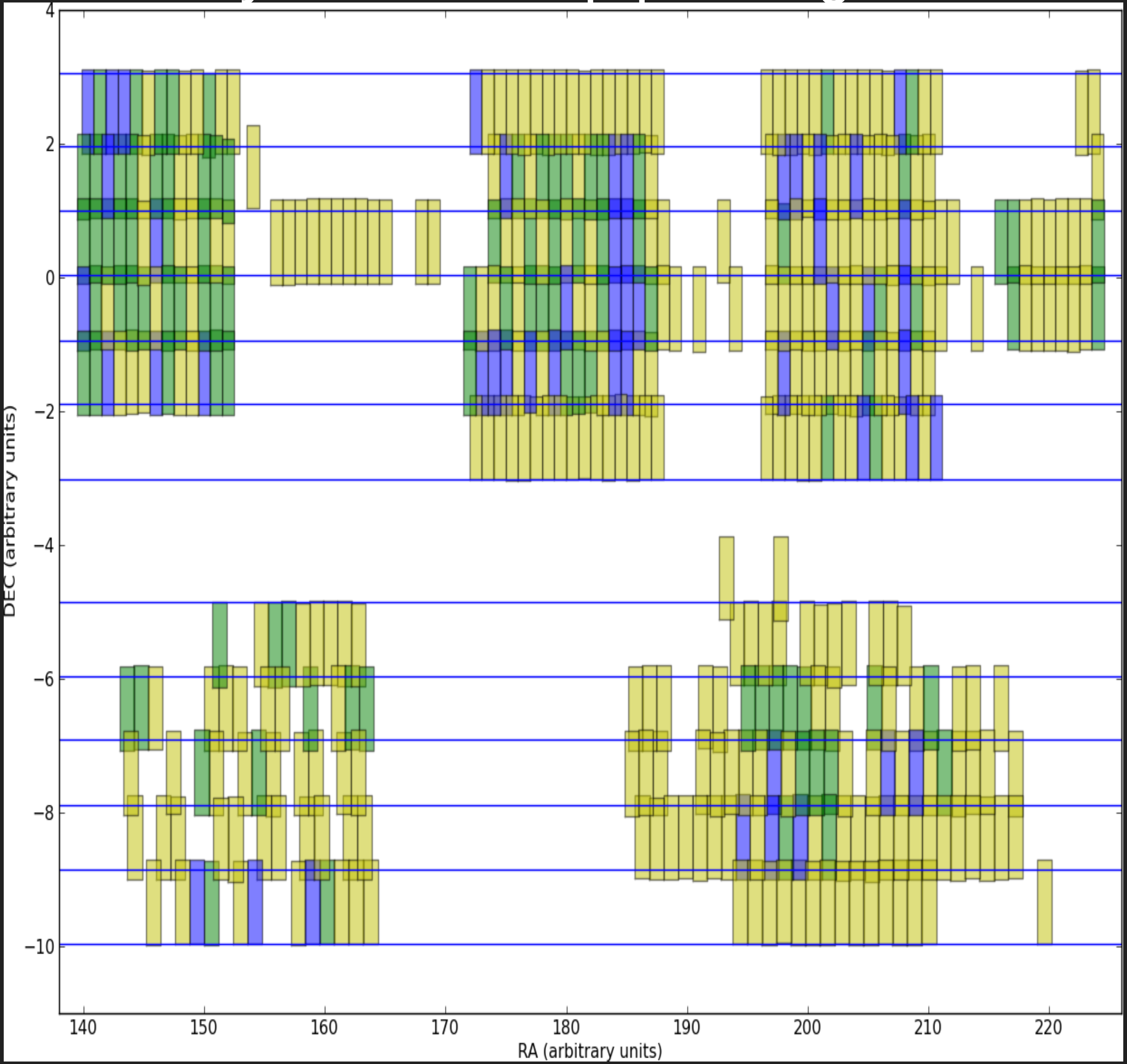
# Observations per band: KiDS-S



# KiDS Survey Photometry

Gert Sikkema
Gijs Verdoes Kleijn
Kapteyn Institute
+
KiDS Team

# Survey Photometry with overlap pointings



# Survey Photometry with overlaps

Define Best Observations == Anchors:

-) no atmospheric extinction variations

-) no “forevers” photometry

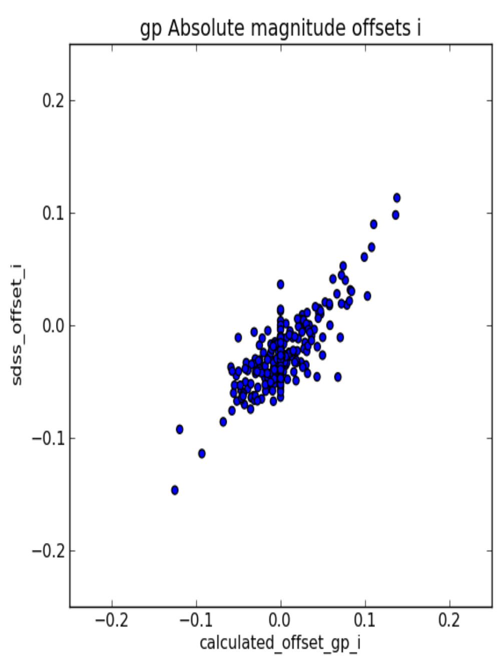
-) no data prior to april 2012 (CCD82 problem)

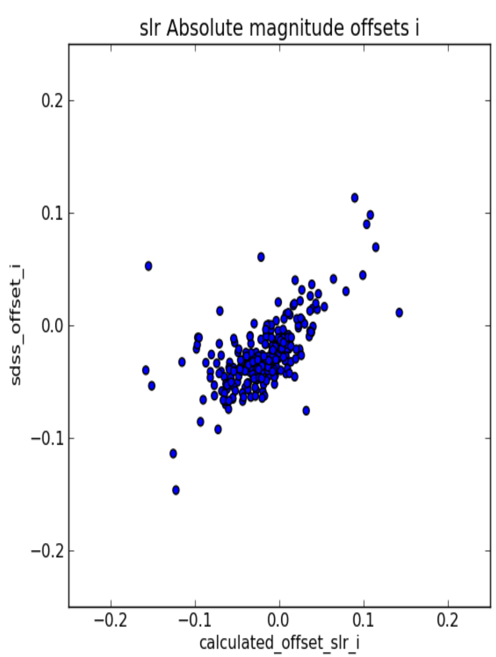
-) no large differences between PSF_zeropoint - PSF_science

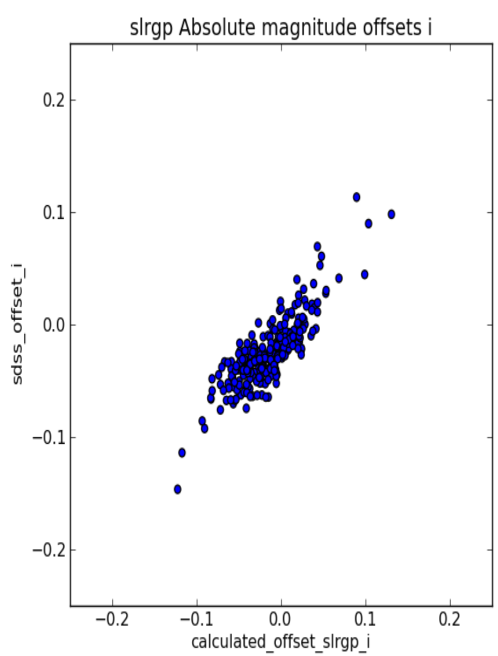
Anchors about 30% in R band.

Tie all other tiles to anchors.

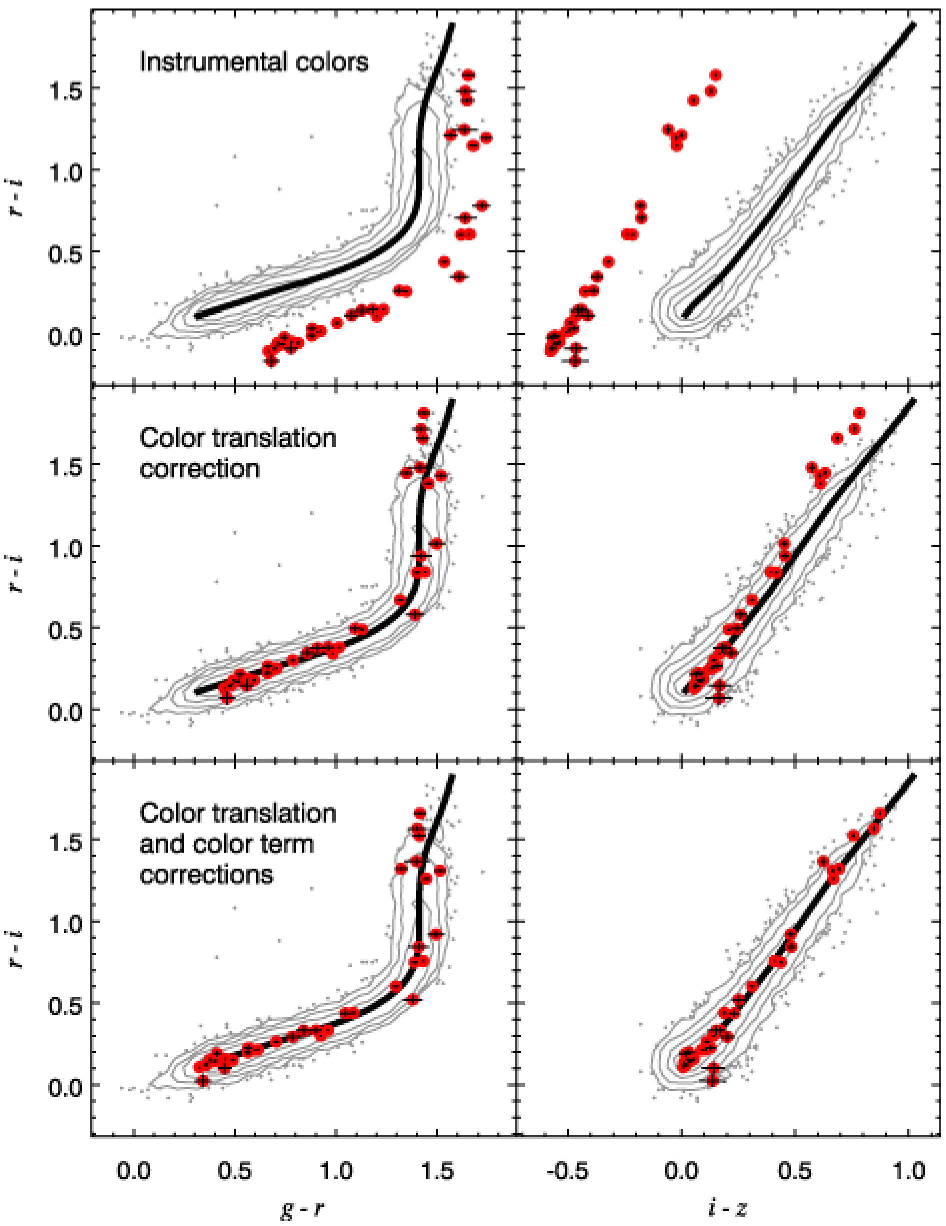
# Result compared to previous surveys (SDSS)



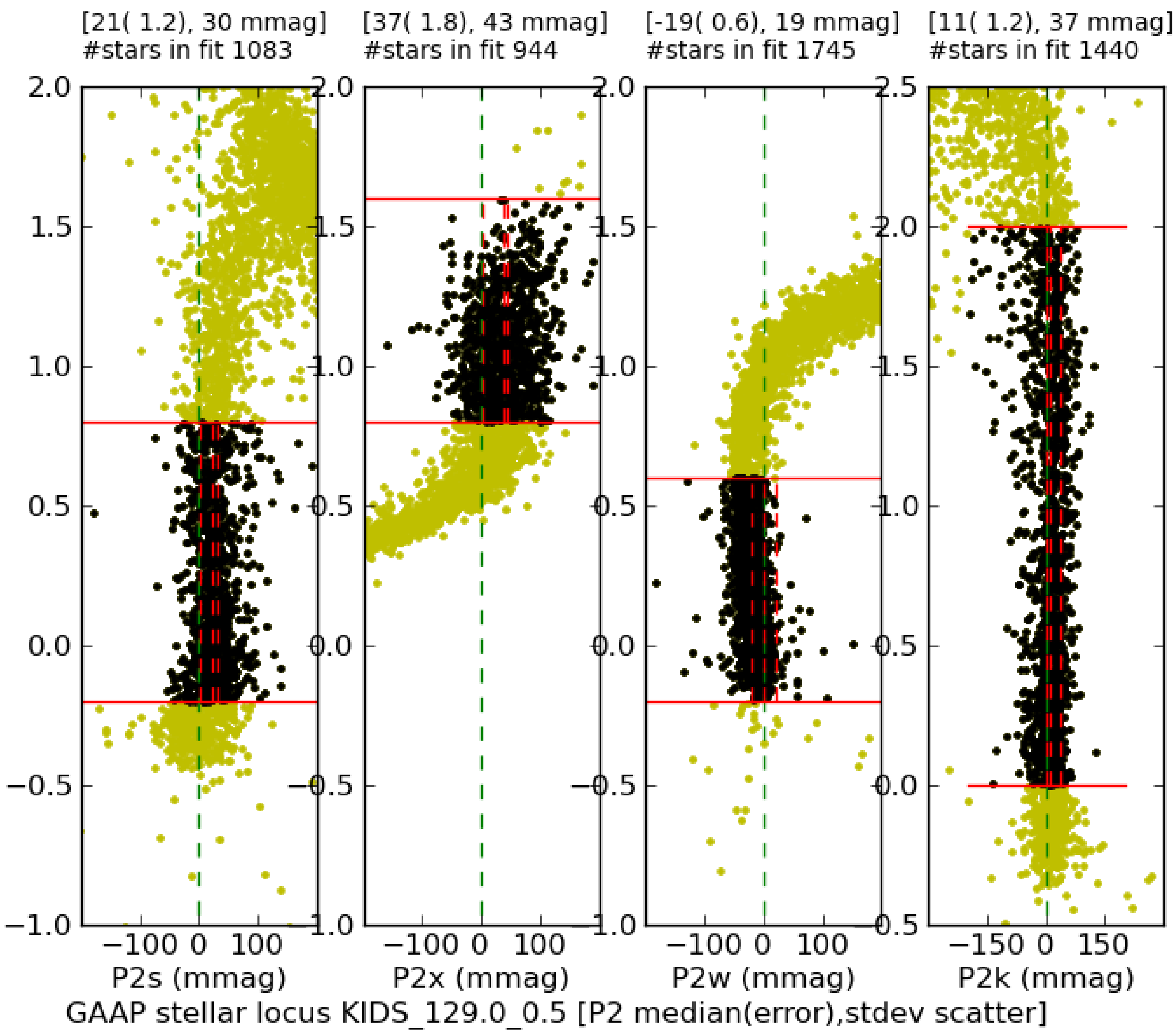



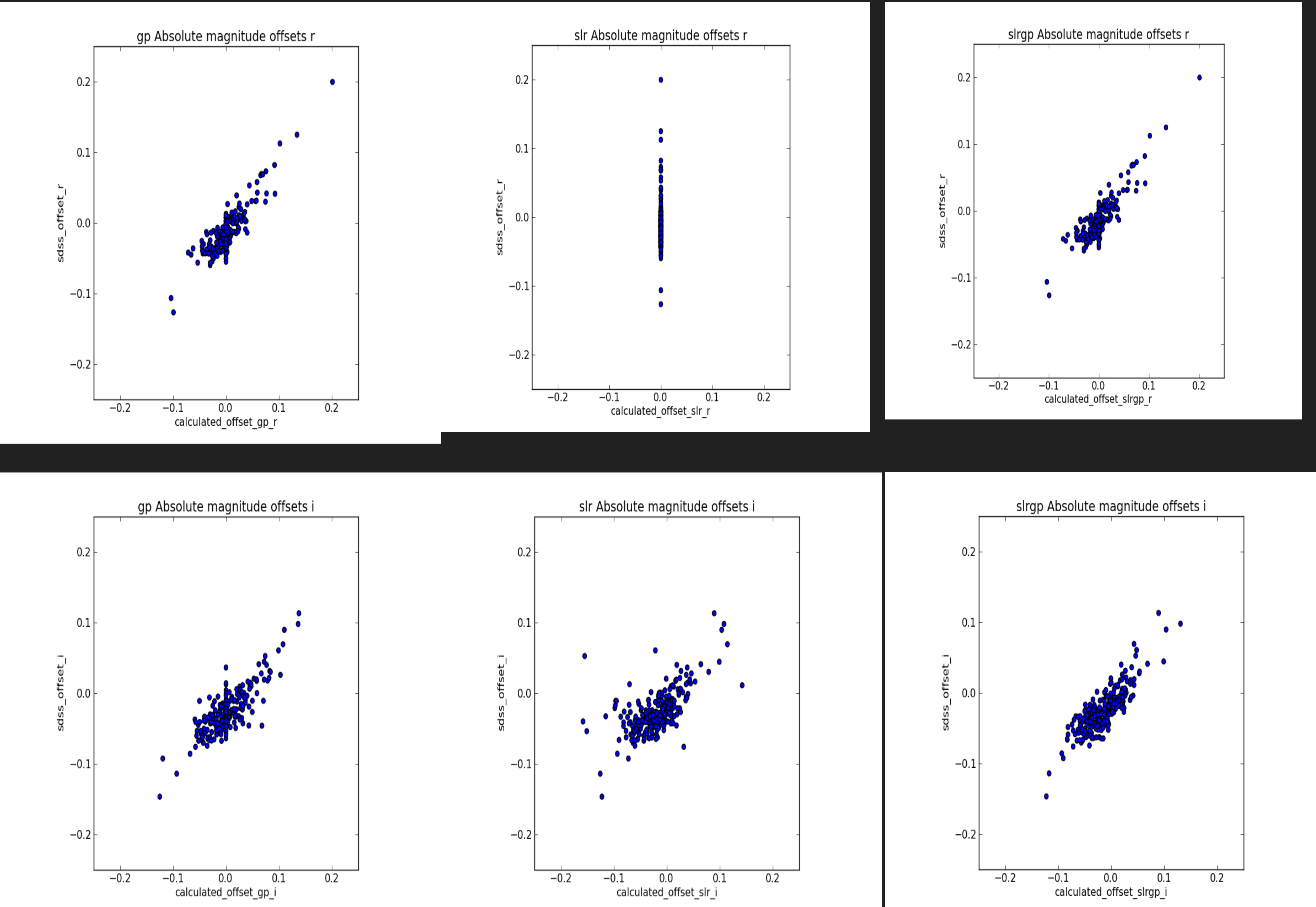



# Stellar colors



# Stellar Locus Regression



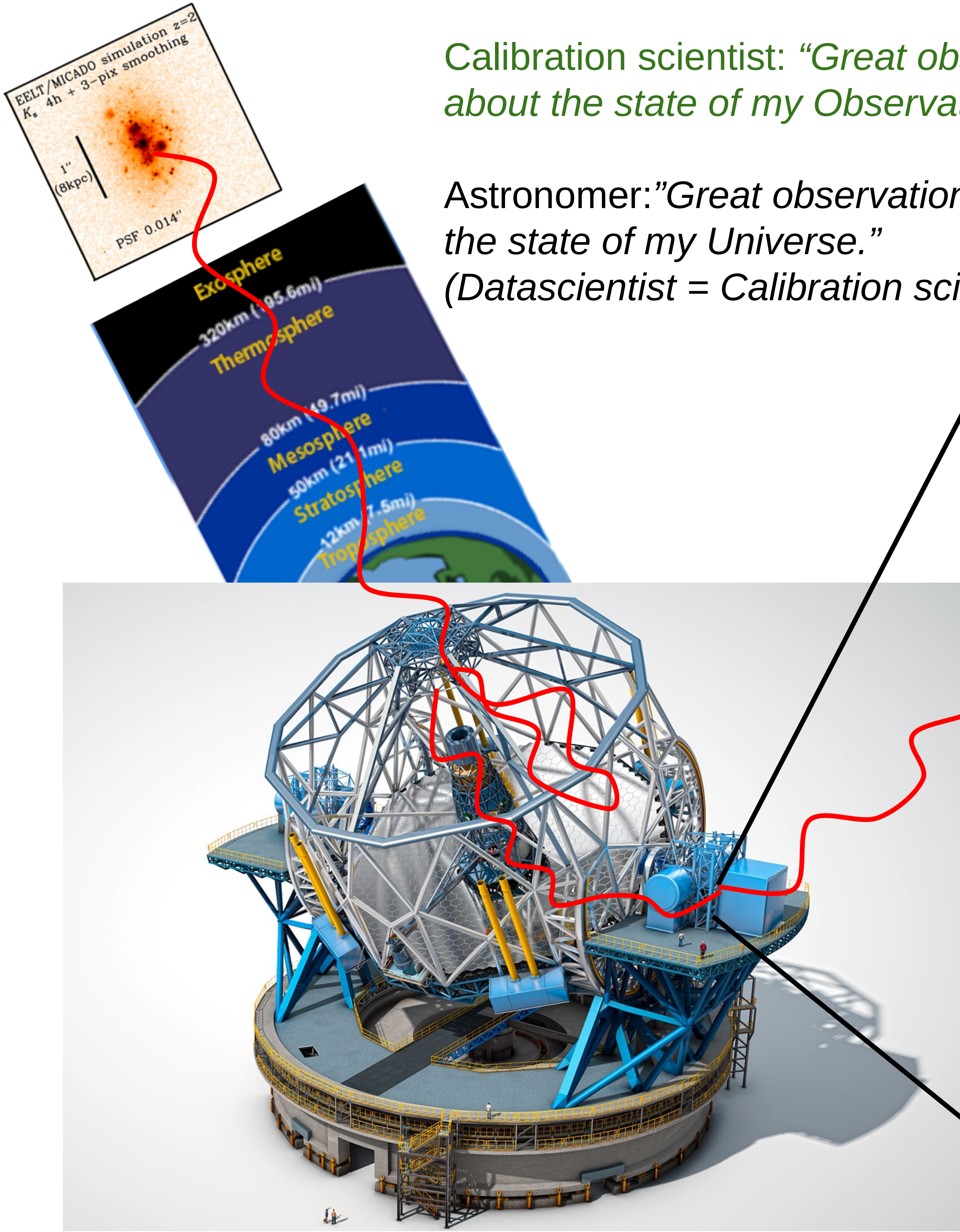

gp Absolute magnitude offsets r
sdss_offset_r
calculated_offset_gp_r
slr Absolute magnitude offsets r
sdss_offset_r
calculated_offset_slr_r
slrgp Absolute magnitude offsets r
sdss_offset_r
calculated_offset_slrgp_r
gp Absolute magnitude offsets i
sdss_offset_i
calculated_offset_gp_i
slr Absolute magnitude offsets i
sdss_offset_i
calculated_offset_slr_i
slrgp Absolute magnitude offsets i
sdss_offset_i
calculated_offset_slrgp_i

# Surveys photometry and Astro-WISE

Astro-WISE: lecture next week

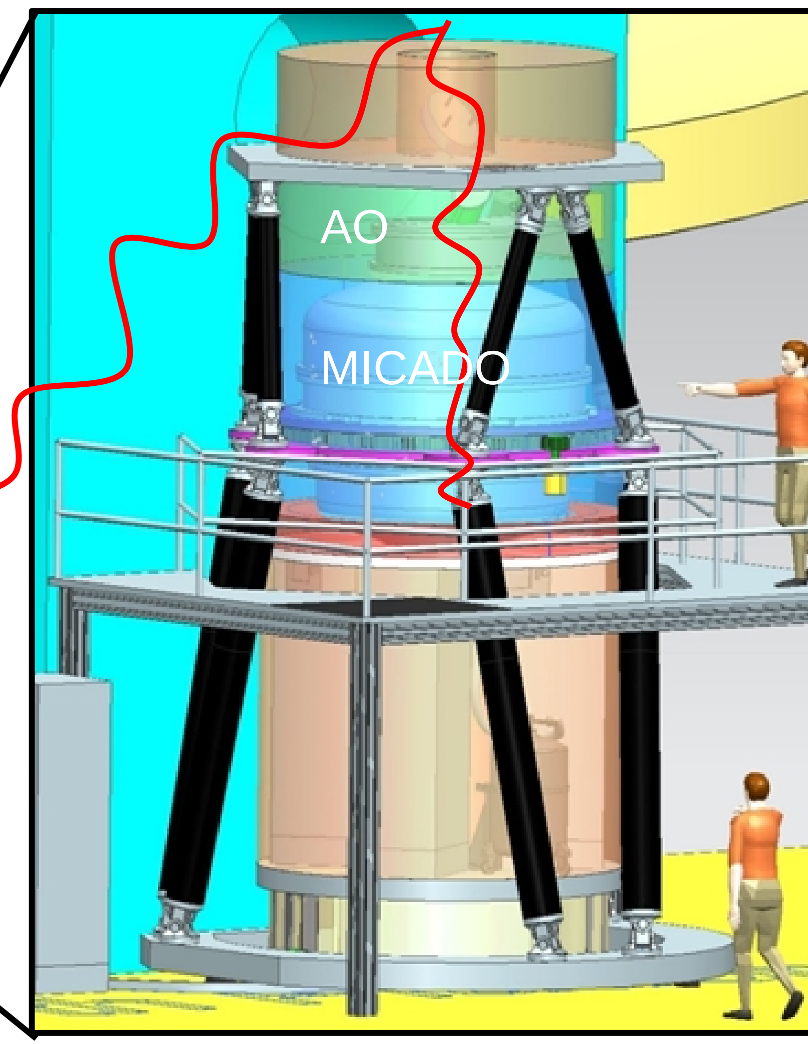
- Data Bookkeeping
- Archive
- Processing
- Quality Assessment & Control
- Science analysis
- Collaboration

# Observing and Observatory




Calibration scientist: *"Great observations! They will tell me everything about the state of my Observatory"*

Astronomer:*"Great observations! They will tell me everything about the state of my Universe."*
*(Datascientist = Calibration scientist + astronomer)*



# Design phase: the mission

| | Outerspace | Atmosphere | Telescope | AO | MICADO |
|---|---|---|---|---|---|
| **Physical components** | Science object | Composition model (thermal / non-thermal) | M1 | SCAO | Entrance window |
| | Galactic extinction | Kinematical model for ~32 layers (AO) | M2 | MCAO | Derotator |
| | Cosmic rays | | M3 | | Collimator |
| | Moon, planets, bright stars | | M4 | | Filter |
| | Zodiacal light | | M5 | | ADC |
| | | | M7 | | Imager optics, reimager |
| | | | | | Pedestal, pixelsensitivity, persistence, crosstalk, dark current, read-outl |
| | | | | | |
| **Observables** | **Astrophysics** | **On-site measurements (T, r, H, P, v..), models of layers** | **DM telemetry** | **WFS telemetry** | **Detector exposures** |

# Design phase: the mission

| | Outerspace | Atmosphere | Telescope | AO | MICADO |
|---|---|---|---|---|---|
| **Physical components MODEL** | Science object | Composition model (thermal / non-thermal) | M1 | SCAO | Entrance window |
| | Galactic extinction | Kinematical model for ~32 layers (AO) | M2 | MCAO | Derotator |
| | Cosmic rays | | M3 | | Collimator |
| | Moon, planets, bright stars | | M4 | | Filter |
| | Zodiacal light | | M5 | | ADC |
| | | | M7 | | Imager optics, reimager |
| | | | | | Pedestal, pixelsensitivity, persistence, crosstalk, dark current, read-outl |
| ↓ | | | | | |
| **Observables MODEL** | **Astrophysics** | **On-site measurements (T, r, H, P, v..), models of layers** | **DM telemetry** | **WFS telemetry** | **Detector exposures** |

# Observing an observatory
# Observing a Universe

( [science-pixdata, observatory-pixdata] | Observables)

# Observing an observatory
# Observing a Universe

P( [science-pixdata, observatory-pixdata] | Observables)

# Data scientist's view:

| | **Outerspace** | **Atmosphere** | **Telescope** | **AO** | **MICADO** |
|---|---|---|---|---|---|
| **Pixel data MODEL** | Science object | Composition model (thermal / non-thermal) | M1 | SCAO | Entrance window |
| | Galactic extinction | Kinematical model for ~32 layers (AO) | M2 | MCAO | Derotator |
| | Cosmic rays | | M3 | | Collimator |
| | Moon, planets, bright stars | | M4 | | Filter |
| | Zodiacal light | | M5 | | ADC |
| | | | M7 | | Imager optics, reimager |
| | | | | | Pedestal, pixelsensitivity, persistence, crosstalk, dark current, read-outl |
| ↑ | | | | | |
| **Observables MODEL** | **Astrophysical standards** | **On-site measurements (T, r, H, P, v..), models of layers** | **DM telemetry** | **WFS telemetry** | **Detector exposures** |

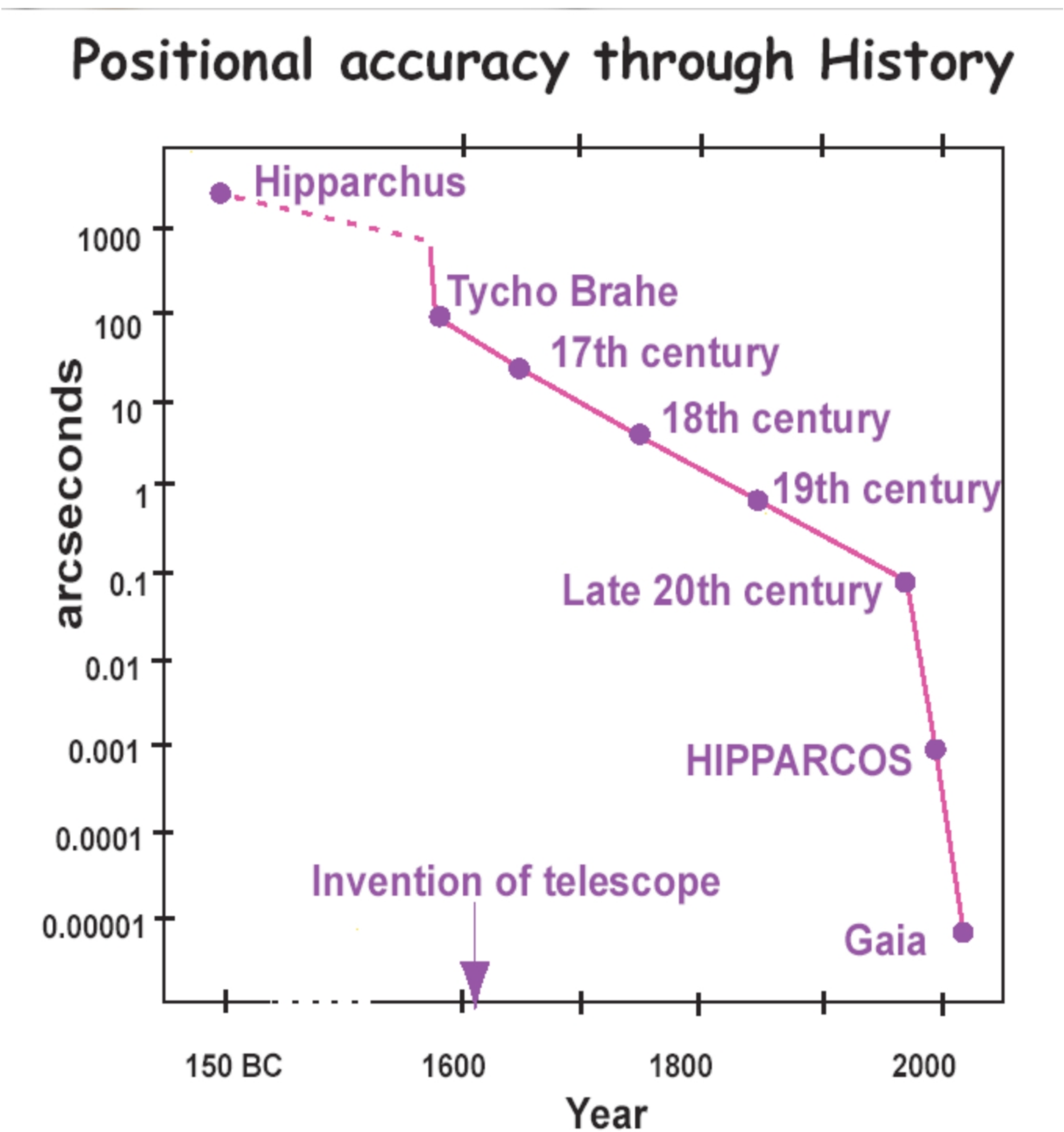
# Astronomer's view

| | | Atmosphere | Telescope | AO | MICADO |
|---|---|---|---|---|---|
| **Pixel data MODEL** | Science object | Composition model (thermal / non-thermal) | M1 | SCAO | Entrance window |
| | | Kinematical model for ~32 layers (AO) | M2 | MCAO | Derotator |
| | | | M3 | | |
| | | | M4 | | |
| | | | M5 | | |
| | | | M7 | | |
| | | | | | |
| | | | | | |
| **Observables MODEL** | **Astrophysical science objects** | | | | |

# Survey astrometry

1. Astrometric science revival: Gaia & MICADO
2. Gaia global astrometry math formalization
3. AGIS: Gaia’s astrometric information system
4. The science <-> Information System symbiosis

# Astrometry in History




- Standards
- Astrometric catalogs
- Surveys & catalogs

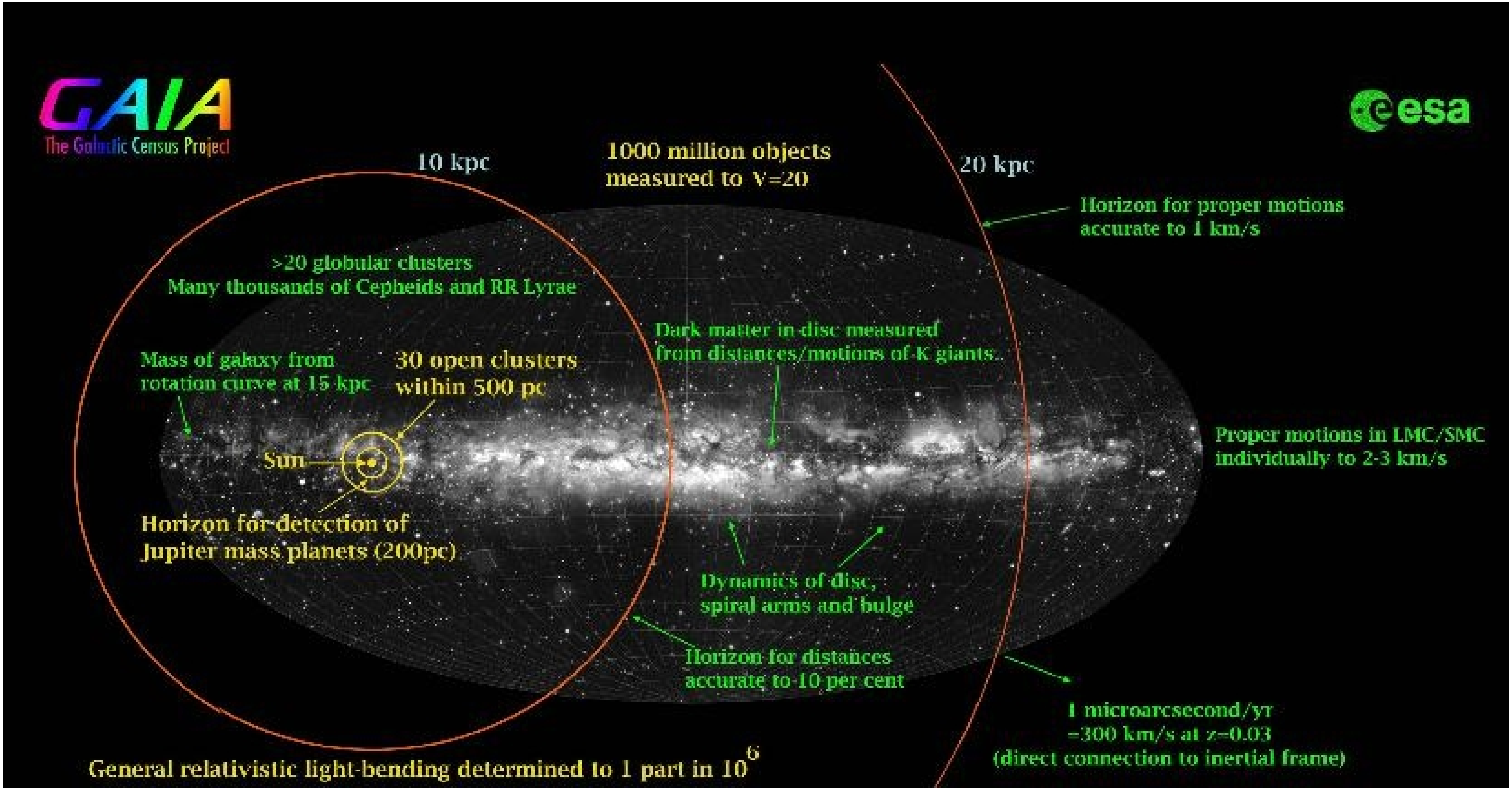
# IT challenges of Gaia's Astrometric Global Iterative Solution

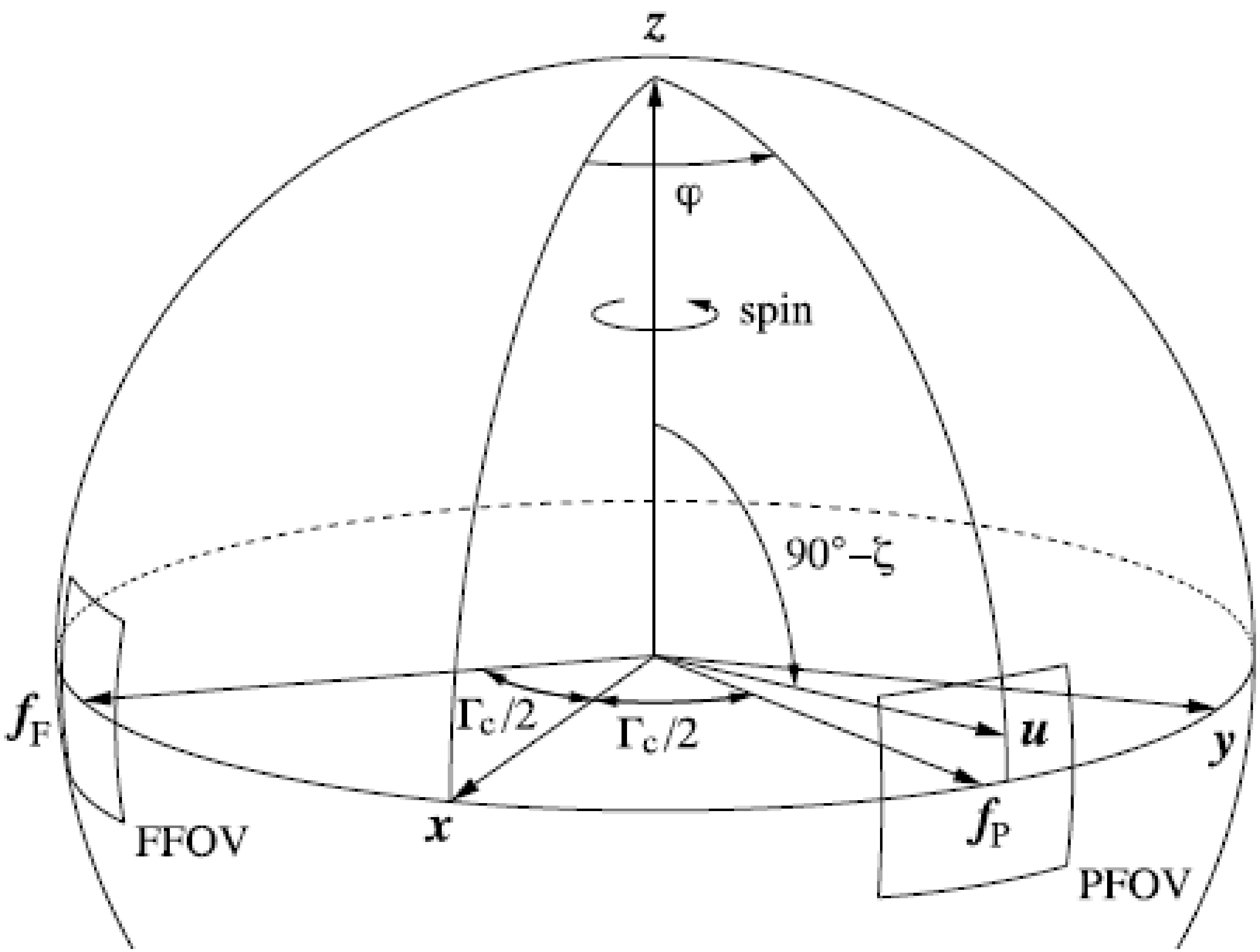
William O'Mullane for Jose Hernandez

Gaia Science Operations Centre
European Space Astronomy Centre
Madrid, Spain

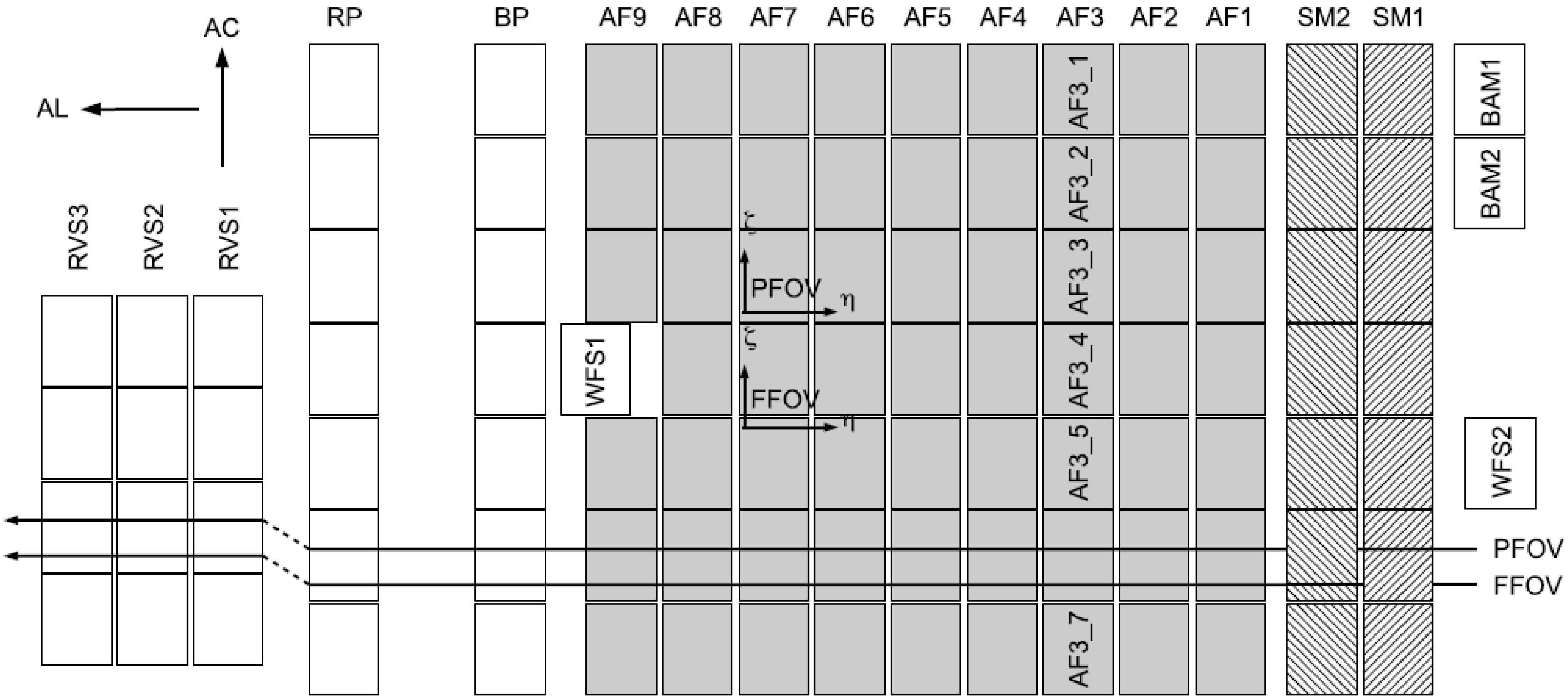
November 25th 2015
ESO Garching

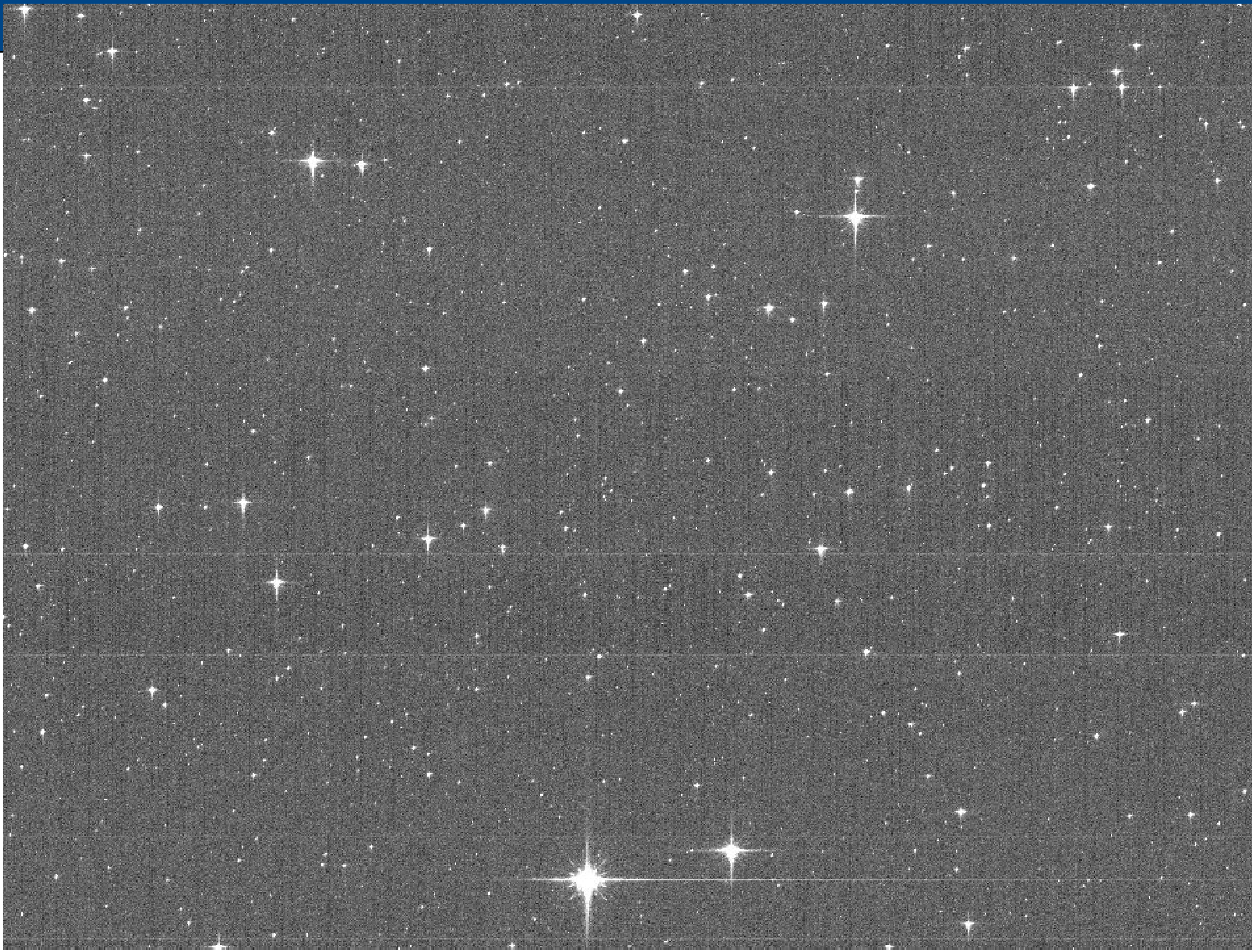
# Astrometric science

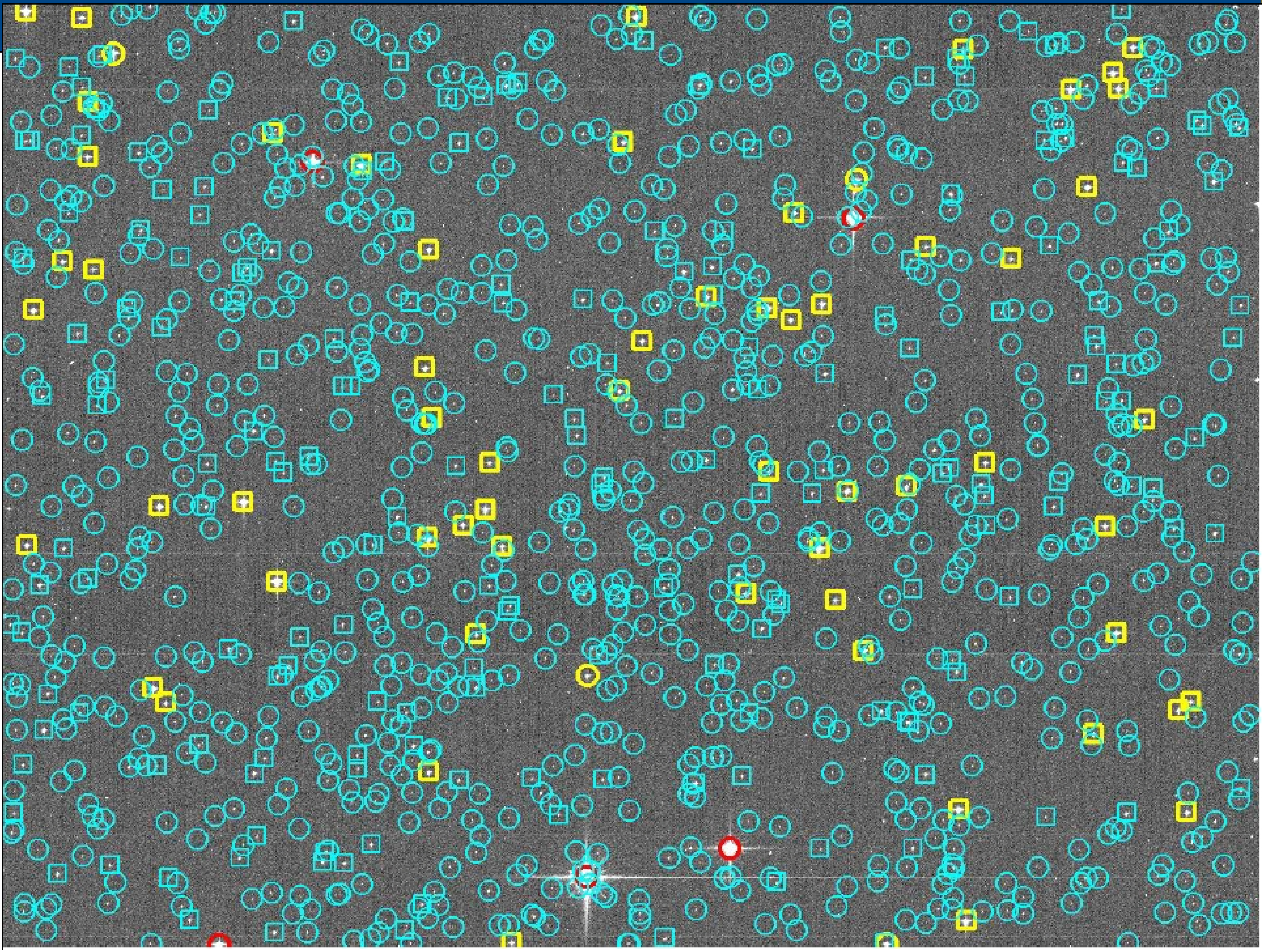




z
φ
spin
90°–ζ
$f_F$
$\Gamma_c/2$
$\Gamma_c/2$
u
y
x
FFOV
$f_P$
PFOV

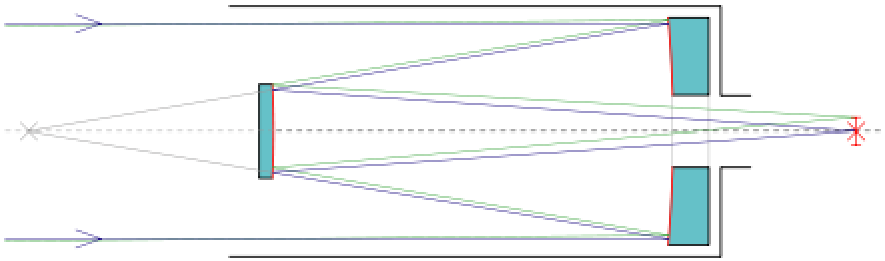

AC
AL
RP
BP
AF9
AF8
AF7
AF6
AF5
AF4
AF3
AF2
AF1
SM2
SM1
BAM1
BAM2
AF3_1
AF3_2
AF3_3
AF3_4
AF3_5
AF3_7
RVS3
RVS2
RVS1
WFS1
WFS2
ζ
PFOV
η
ζ
FFOV
η
PFOV
FFOV

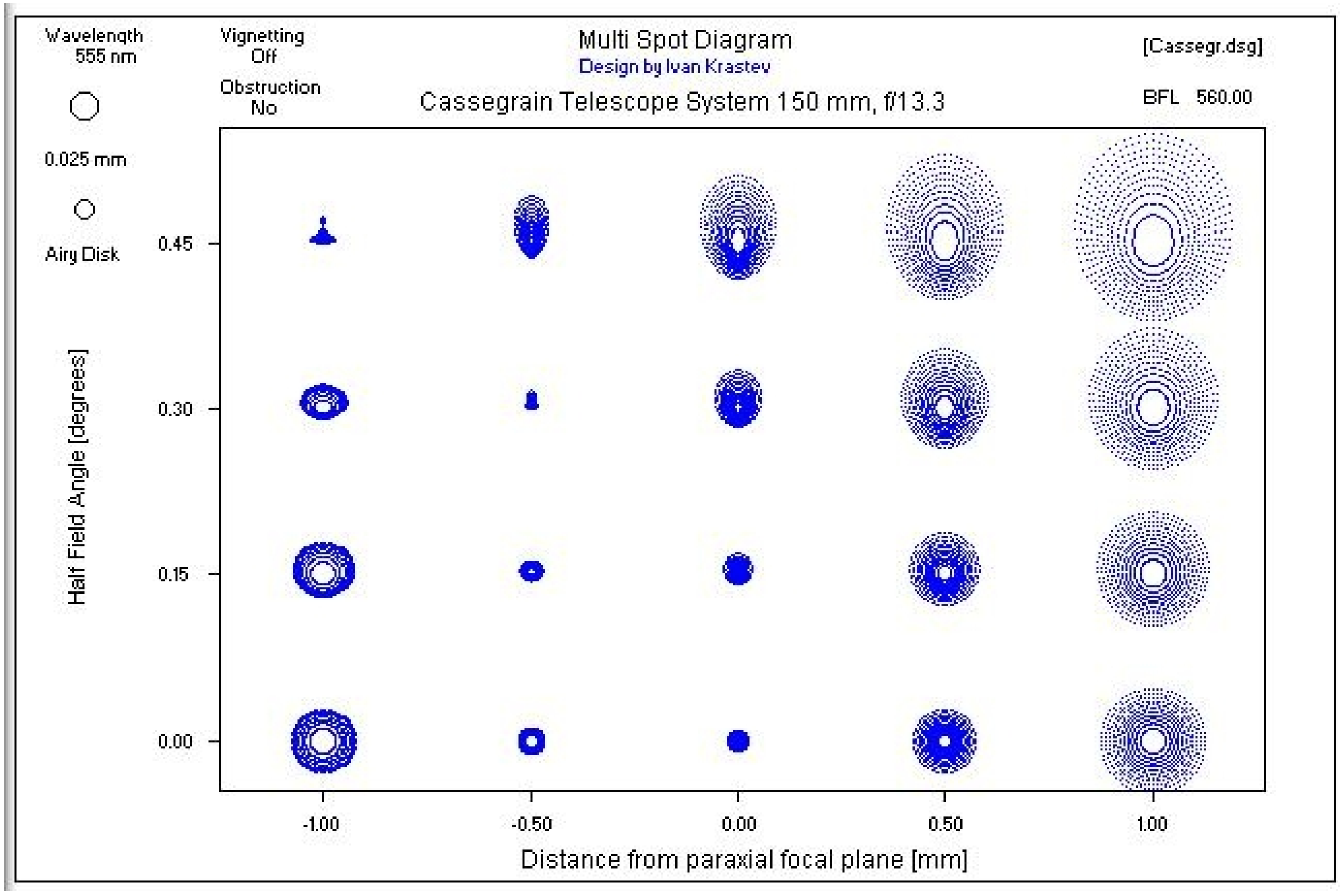




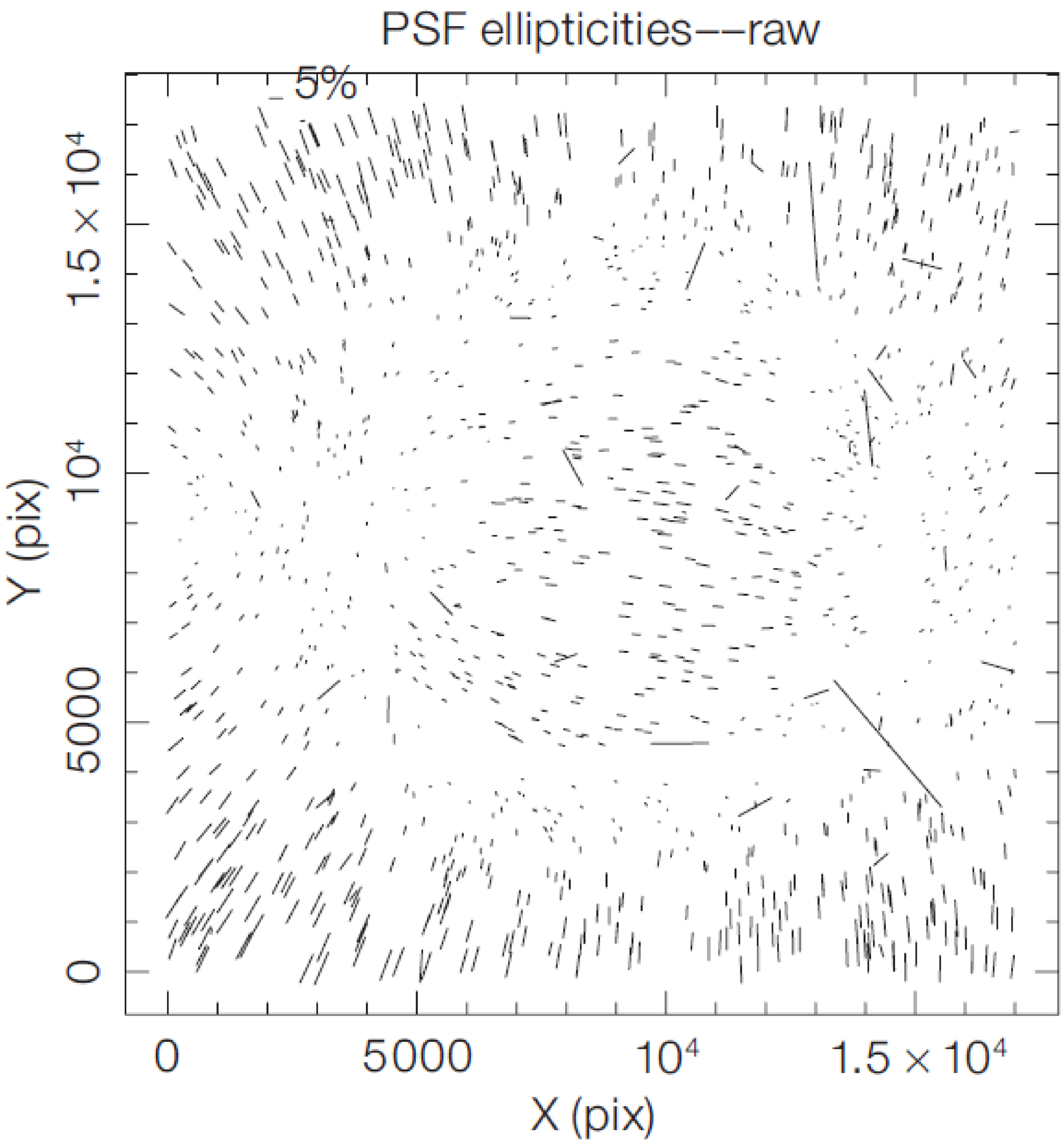
# Astrometric Calibrations

- Atmospheric refraction (from 0 in zenith to 5' at 10o)
- Aberration

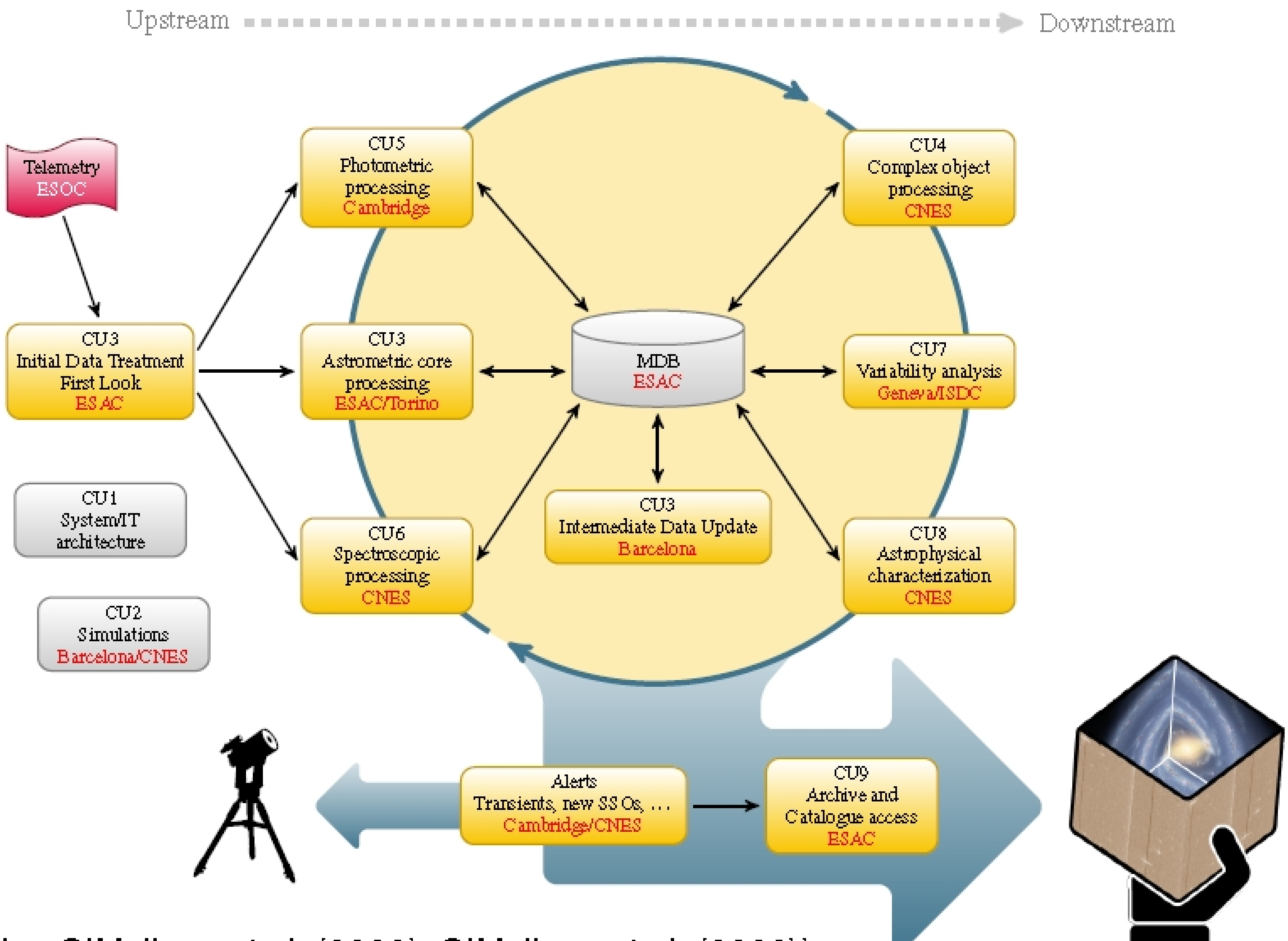




# PSF Homogenization



# Simplified processing overview




(see also O'Mullane et al. (2006), O'Mullane et al. (2009))

# Astrometric Solution

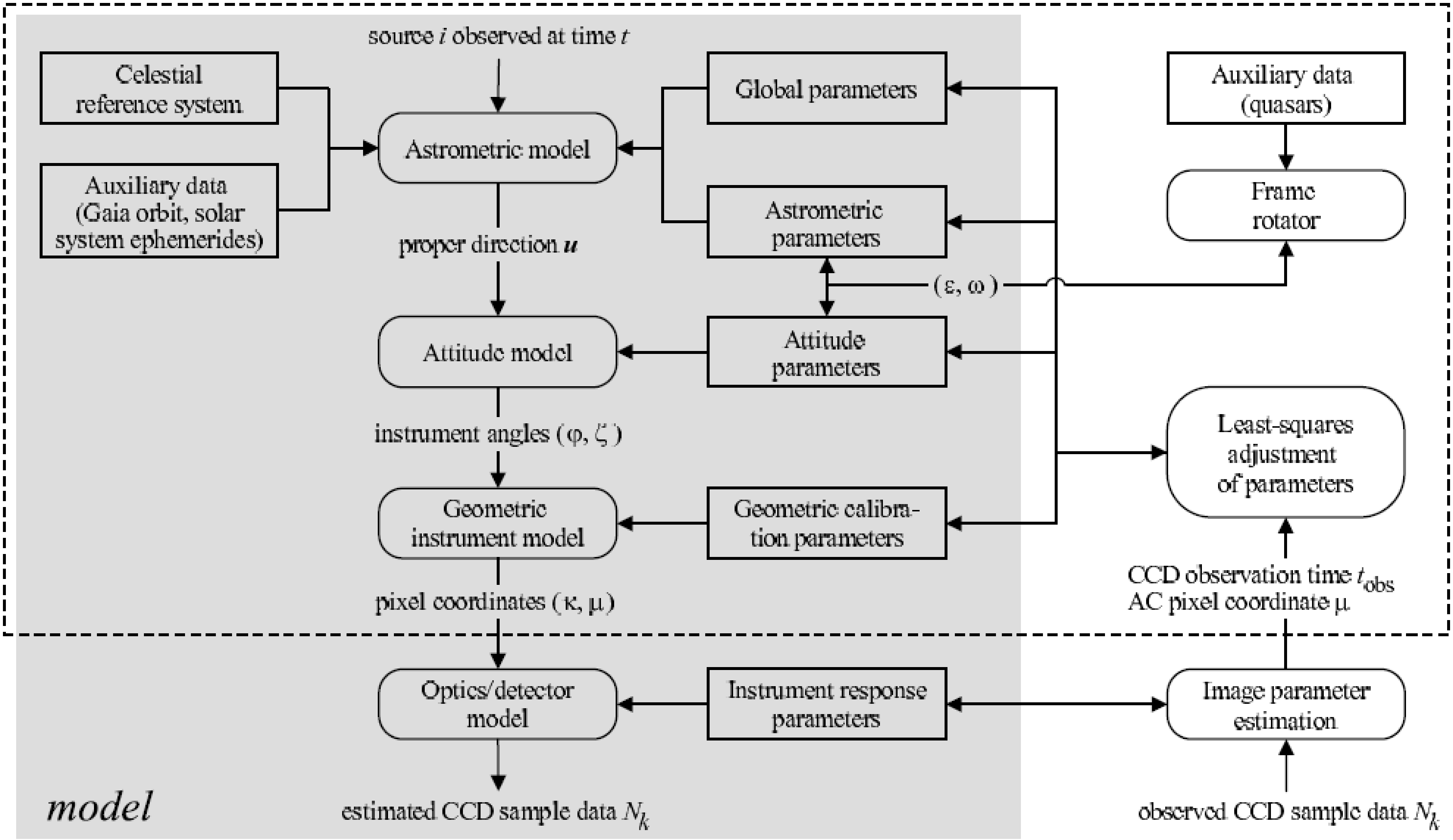
Just one part of the processing !
From the Hipparcos catalogue (ESA, 1997, Volume 3 Chapter 23).

## Minimisation problem for astrometry

$$\min_{\mathbf{a},\mathbf{n}} \|\mathbf{g}^{\mathbf{obs}} - \mathbf{g}^{\mathbf{calc}}(\mathbf{a},\mathbf{n})\|_M \quad (1)$$

- $\mathbf{a}$ is the vector of unknowns describing a star's barycentric motion represented by the measurement vector $\mathbf{g_k} = (G_k, H_k)'$ and associated statistics.
- $\mathbf{g}^{\mathbf{obs}}$ represents the vector of all measurements
- $\mathbf{g}^{\mathbf{calc}}$ represents the vector of detector coordinates calculated from the astrometric parameters.
- $\mathbf{n}$ is a vector of nuisance parameters - required for realistic modelling (e.g. attitude, instrument calibration)
- $M$ metric defined by the statistics of the data, (error weighting)

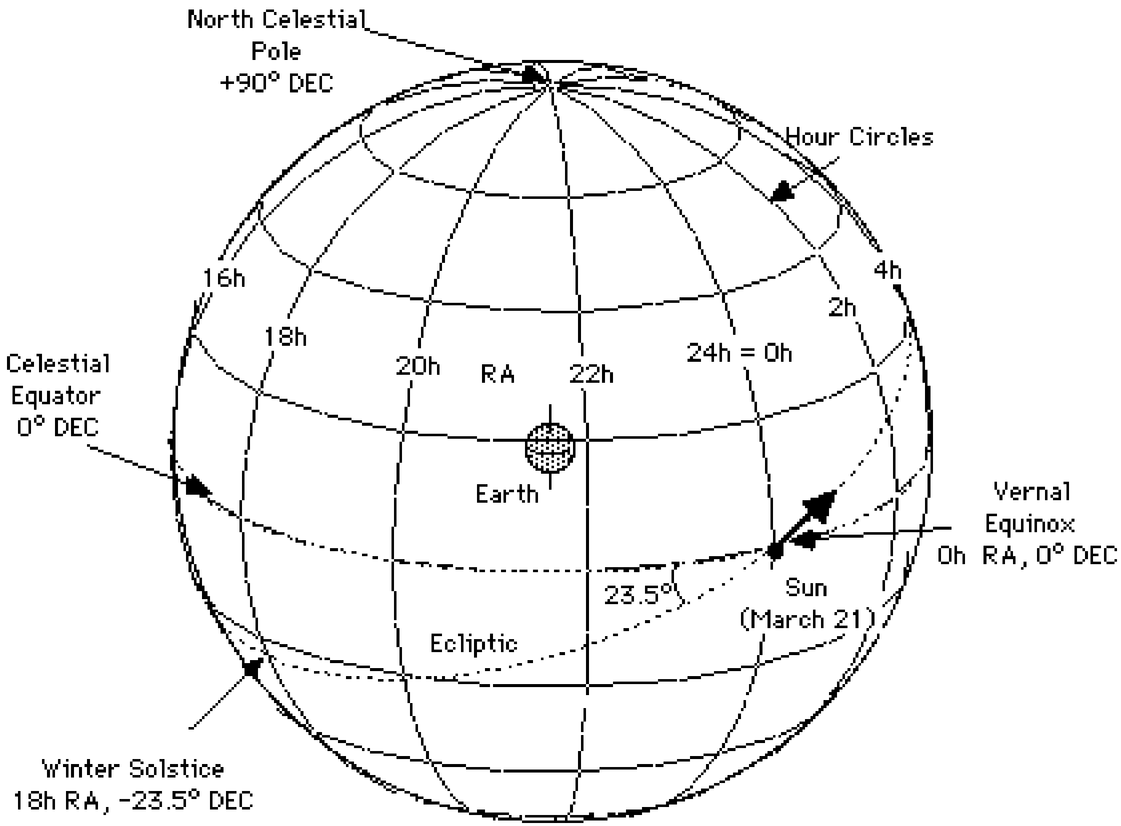
The complete new formulation for Gaia is in (Lindegren et al., 2012).


source $i$ observed at time $t$
Celestial reference system
Auxiliary data (Gaia orbit, solar system ephemerides)
Astrometric model
Global parameters
Astrometric parameters
Auxiliary data (quasars)
Frame rotator
proper direction $\boldsymbol{u}$
(ε, ω)
Attitude model
Attitude parameters
instrument angles (φ, ζ)
Least-squares adjustment of parameters
Geometric instrument model
Geometric calibration parameters
CCD observation time $t_{obs}$
AC pixel coordinate μ
pixel coordinates (κ, μ)
Optics/detector model
Instrument response parameters
Image parameter estimation
model
estimated CCD sample data $N_k$
observed CCD sample data $N_k$

# Data processing - challenging

- Astrometric centroid of the CCD image to be determined to an accuracy of 1% of the pixel size!
  - There will be $10^{12}$ images $\approx 100$TB downlink need to handle $\approx 1$PB
  - Processing estimate remains $\approx 10^{20}$ FLOP
- Reconstructed attitude is required to order 10 $\mu$arcsec
  - Path of light through instrument needed to nanometre level
  - System must be extremely stable
  - Must consider relativistic light bending from solar system objects.
- Attitude and Geometric calibration can only by done using Gaia's own observational data. (AGIS)
- Testing and verification is very difficult - still running Operational Rehearsals

# Gaia astrometric Reference Frames

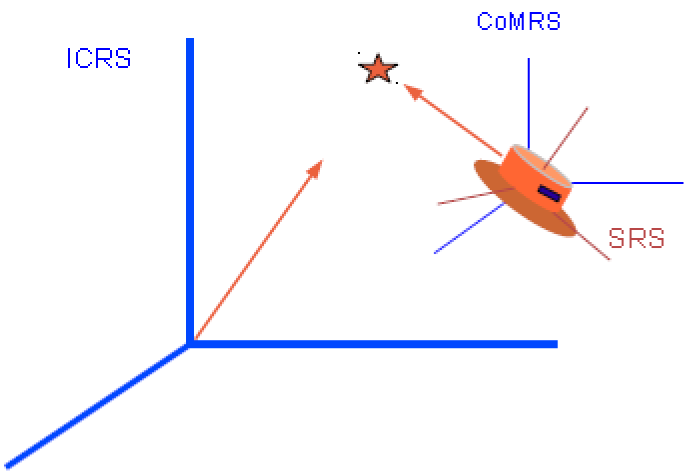


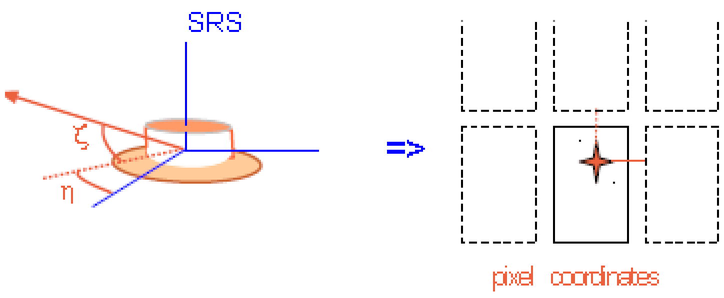

- **Relativity-based systems**
- **Solar system barycentric**
  - Barycentric Celestial Reference System based on ICRS
  - Barycentric Time Coordinate (TCB)
- **Gaia-centric**
  - CoMRS based on ICRS
  - Gaia-time

# Look at one block: Source modelling

The mapping or modelling of the observables **g** is done by three successive transformations:

1. from astrometric parameters to the celestial directions of a star at the instant of observation, using an astrometric model **S**
2. from celestial to instrument frame directions using an attitude model **A**
3. and finally from instrument directions to detector coordinates using an instrument model **C**



# Astrometric Global Iterative Solution

AGIS solves Equation (1) block iteratively to determine $5 \times 10^9$ **S**ource par meters from the $> 10^{12}$ measurements
Three models are needed:

- **S**ource (star) model ⟶ wanted params
- **A**ttitude(pointing) model
- **C**alibration (focal plane geometry) model

} nuisance params

Nuisance parameters couple wanted params in complex way.

- This is by design - both good and bad!

- Each model solved as a *block* (Cholesky decomposition)
- Theoretically the order of the blocks is unimportant.
  - In fact we do S and then A,C in parallel (outer iteration)
- Conjugate Gradient Solver used to drive the outer iterations.
- Matrix inversion not computable (Bombrun et al., 2010)

# Source Update

We fit the model to the observations:

**Least squares for source update**

$$\mathbf{A}\mathbf{x} \sim \mathbf{b} \pm \sigma \quad (2)$$

$$\text{where } \mathbf{b}_i = \mathbf{y}_i - f_i(\mathbf{a}, \mathbf{q}) \quad (3)$$

Here $\mathbf{y}_i$ are the observed field angles $f_i$ is a function to calculates field angles from the the current model.
In java (SourceUpdateCalculatorWrapper):

```
// get calculated angles + derivatives ...
ExtendedFieldAngles ecfa = angleCalc.getCalculatedEtaZeta(ae, origSrc, UpdateBlock.Source.getId());

// ... from those, get just calculated angles
double[][] calcEtaZeta = ecfa.getEtaZeta();

// ... and the observed ones from the angle calculator
double[][] obsEtaZeta = angleCalc.getObservedEtaZeta(ae);

// compute residuals [rad] and attach those to the Elementary
double[][] etaZetaRes = AgisUtils.subtractArrays(obsEtaZeta, calcEtaZeta);
```

Yes it is ALL in Java.

$$\min_{\boldsymbol{s},\boldsymbol{a},\boldsymbol{c},\boldsymbol{g}} \; Q = \sum_{l \in \mathrm{AL}} \frac{(R_l^{\mathrm{AL}})^2 \, w_l^{\mathrm{AL}}}{(\sigma_l^{\mathrm{AL}})^2 + (\epsilon_l^{\mathrm{AL}})^2} + \sum_{l \in \mathrm{AC}} \frac{(R_l^{\mathrm{AC}})^2 \, w_l^{\mathrm{AC}}}{(\sigma_l^{\mathrm{AC}})^2 + (\epsilon_l^{\mathrm{AC}})^2}, \tag{24}$$

where

$$R_l^{\mathrm{AL}}(\boldsymbol{s},\boldsymbol{a},\boldsymbol{c},\boldsymbol{g}) = \eta_{fng}(\mu_l, t_l \,|\, \boldsymbol{c}) - \eta(t_l \,|\, \boldsymbol{s},\boldsymbol{a},\boldsymbol{g}), \tag{25}$$

$$R_l^{\mathrm{AC}}(\boldsymbol{s},\boldsymbol{a},\boldsymbol{c},\boldsymbol{g}) = \zeta_{fng}(\mu_l, t_l \,|\, \boldsymbol{c}) - \zeta(t_l \,|\, \boldsymbol{s},\boldsymbol{a},\boldsymbol{g}) \tag{26}$$

# AGIS implementation in Java

(O'Mullane et al., 2011)

- In 2013 - 218276 lines → 140305 code + 77971 comments
- In 2015 - 246566 lines → 160587 code + 85979 comments
- Started 2005 with O'Mullane and Lammers
- Algorithms provided by Lindegren (*close science/engineering collaboration*
  - usually as tech notes e.g. Lindegren (LL-072), Lindegren (LL-065), Bombrun et al. (LL-096)..
  - and sometimes Java code
- Guiding principle has been to take a minimalist approach (Datatrain (O'Mullane et al., 2006))
  - access data as little as possible - once per iteration
  - distribute: take advantage of multi core distributed systems
  - try to cut down on single point bottlenecks
  - try to keep the algorithm isolated from the framework
    - use Derby for test, Intersystems Cache for production (Oracle to 2012)

$$\begin{bmatrix} \boldsymbol{N}_{ss} & \boldsymbol{N}_{sa} & \boldsymbol{N}_{sc} & \boldsymbol{N}_{sg} \\ \boldsymbol{N}_{as} & \boldsymbol{N}_{aa} & \boldsymbol{N}_{ac} & \boldsymbol{N}_{ag} \\ \boldsymbol{N}_{cs} & \boldsymbol{N}_{ca} & \boldsymbol{N}_{cc} & \boldsymbol{N}_{cg} \\ \boldsymbol{N}_{gs} & \boldsymbol{N}_{ga} & \boldsymbol{N}_{gc} & \boldsymbol{N}_{gg} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_{s} \\ \boldsymbol{x}_{a} \\ \boldsymbol{x}_{c} \\ \boldsymbol{x}_{g} \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_{s} \\ \boldsymbol{b}_{a} \\ \boldsymbol{b}_{c} \\ \boldsymbol{b}_{g} \end{bmatrix}$$

# The astrometry <-> Information System relation

- The astrometric solution has been developed over an extended period of time (could say it started with Hipparcos)
- Only the combination of computing and astrometry expertise has made the process achievable (science + engineering)
- It has been tested with simulations to prove the method works
- It has been extended to use new priors to provide early results with one year of data. A full solution requires at least 18 months of mission data.
- First catalogue coming in Summer 2016 !

# Final ponderings

- Ignore the observatory
- Model the observatory
- Information systems for
  - Bookkeeping
  - “Everything changes”
  - Associations

- Superhuman Intelligence (Watson, AlphaGo)

# END