Status of HSC Data Analysis Software

12.18.2012 HSC-AGN WS@Ehime-U
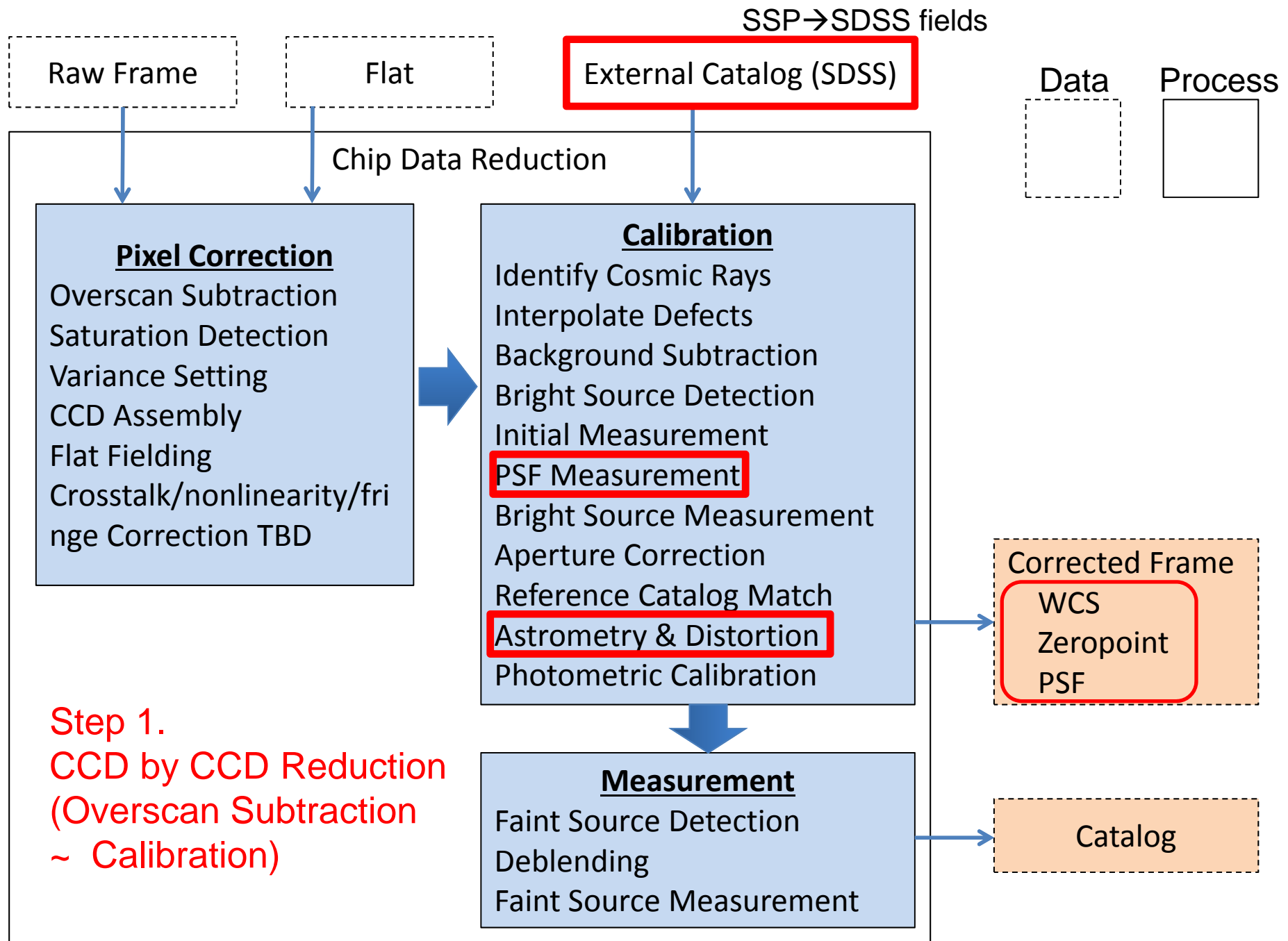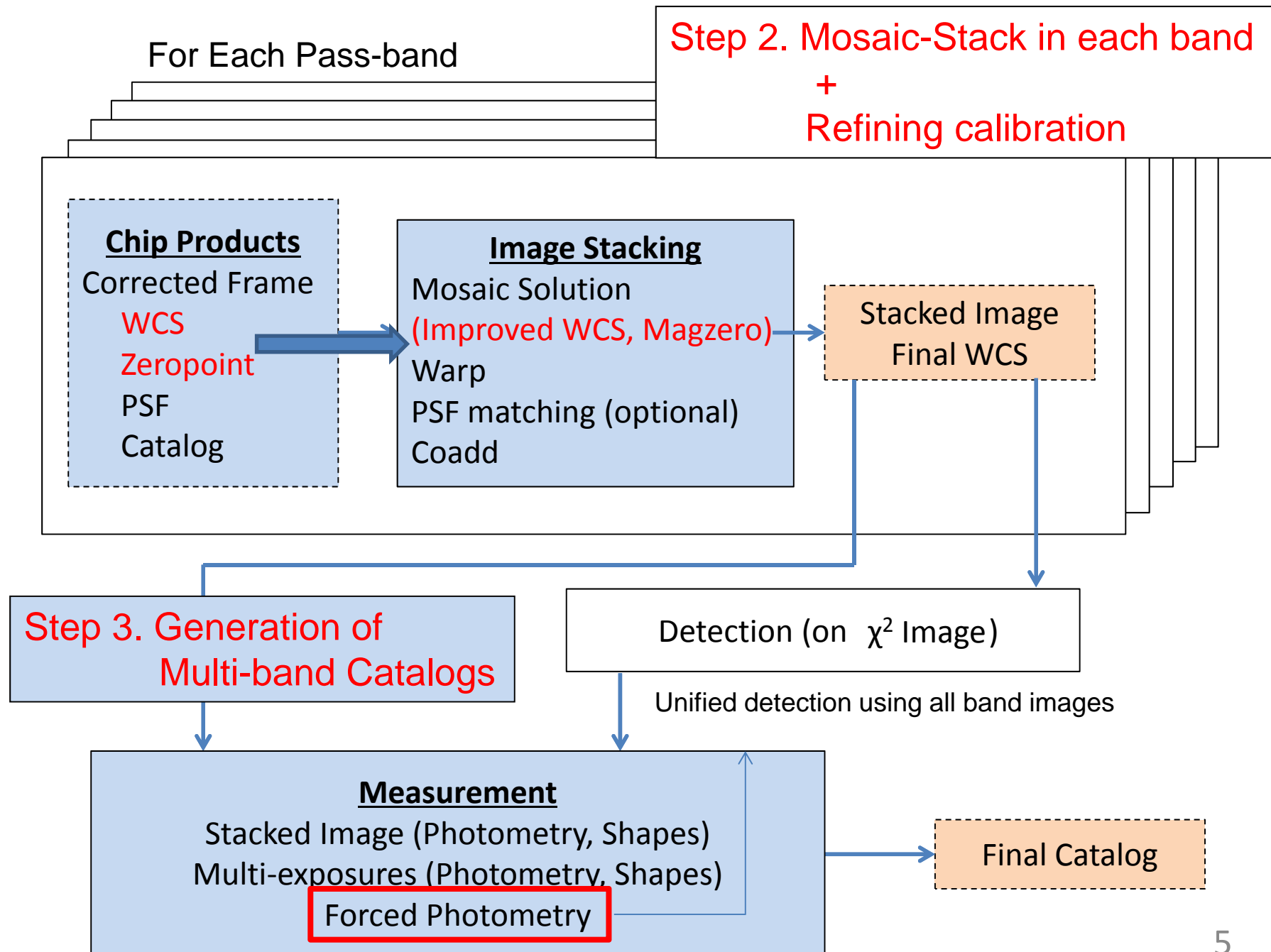
H. Furusawa (NAOJ)

1. HSCANA

- Scope
  - Production of calibrated images and catalogs
  - Distribution of data products to Co-I's
  - Assisting SSP observation
- Team
  - NAOJ, Princeton, K-IPMU/U-Tokyo, KEK
    - Takata, T., Yamada, Y., Koike, M., Okura, Y., Yamanoi, H., Lupton, R., Price, P., Bosch, J., Miyatake, H., Bickerton, S., Yasuda, N., Mineo, S., Furusawa. H, …

2. Software Components

- Platform, Framework (on CentOS Linux)
  - LSST stack
    - → classes/func for data access & manip. (c.f., IRAF)
  - Python, C++ (SWIG)
    - → building analysis tasks, pipeline
  - MPI, Torque (OpenPBS)
    - → parallelization
- HSC Pipelines
  - Reduction of CCD by CCD images
  - Mosaic-Stacking
  - Catalog Creation

3

SSP→SDSS fields

Raw Frame

Flat

External Catalog (SDSS)

Data   Process

Chip Data Reduction

Pixel Correction
Overscan Subtraction
Saturation Detection
Variance Setting
CCD Assembly
Flat Fielding
Crosstalk/nonlinearity/fri
nge Correction TBD

Calibration
Identify Cosmic Rays
Interpolate Defects
Background Subtraction
Bright Source Detection
Initial Measurement
PSF Measurement
Bright Source Measurement
Aperture Correction
Reference Catalog Match
Astrometry & Distortion
Photometric Calibration

Corrected Frame
WCS
Zeropoint
PSF

Step 1.
CCD by CCD Reduction
(Overscan Subtraction
~ Calibration)

Measurement
Faint Source Detection
Deblending
Faint Source Measurement

Catalog

4

For Each Pass-band

Step 2. Mosaic-Stack in each band
+
Refining calibration

Chip Products
Corrected Frame
WCS
Zeropoint
PSF
Catalog

Image Stacking
Mosaic Solution
(Improved WCS, Magzero)
Warp
PSF matching (optional)
Coadd

Stacked Image
Final WCS

Step 3. Generation of Multi-band Catalogs

Detection (on χ^2 Image)

Unified detection using all band images

Measurement
Stacked Image (Photometry, Shapes)
Multi-exposures (Photometry, Shapes)
Forced Photometry

Final Catalog

3. Deliverables

Products Sizing

- Images
  - Corrected Single CCD Images —————— 85MB / CCD
    - WCS, Magnitude zeropoint
    - Variance, Mask Map
    - PSF Map
  - Mosaic-Stacked Images per Unit-area ——— 10GB / Stack / 1.5^2 sq.deg
    - WCS, Magnitude zeropoint
    - Variance, Mask Map
    - PSF Map

- Catalogs
  - Single CCD Catalogs → catalog db —— 3 million row / sq.deg
  - Multi-band Catalogs（HSC）for Stacked Images per Unit-area of sky → catalog-db ———— 0.4 million row / sq.deg
  - Multi epoch?

6

4. Snapshots from CCD-by-CCD Analysis

- PSF mapping
- CR / crosstalk correction & Mask values
- Photometric calibration

Processed Image Format

HDU0: Header only
HDU1: Science Image
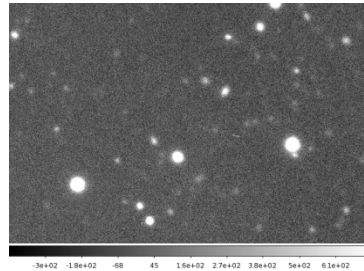
Calibrated astrometry and photometry

16bit integer
Assigned for pixel attributes

HDU2: Mask Image
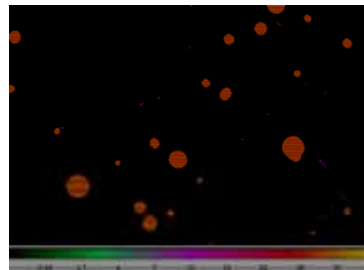
Plane 0 -> BAD
Plane 1 -> SAT
Plane 2 -> INTRP
Plane 3 -> CR
Plane 4 -> EDGE
Plane 5 -> DETECTED
Plane 6 -> DETECTED_NEGATIVE
Plane 7 -> CROSSTALK

HDU3: Variance Image
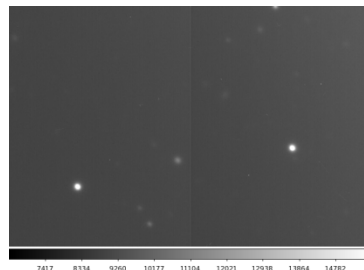
ADU count / Gain

Example of Reduction & Mask Values for real SCam data
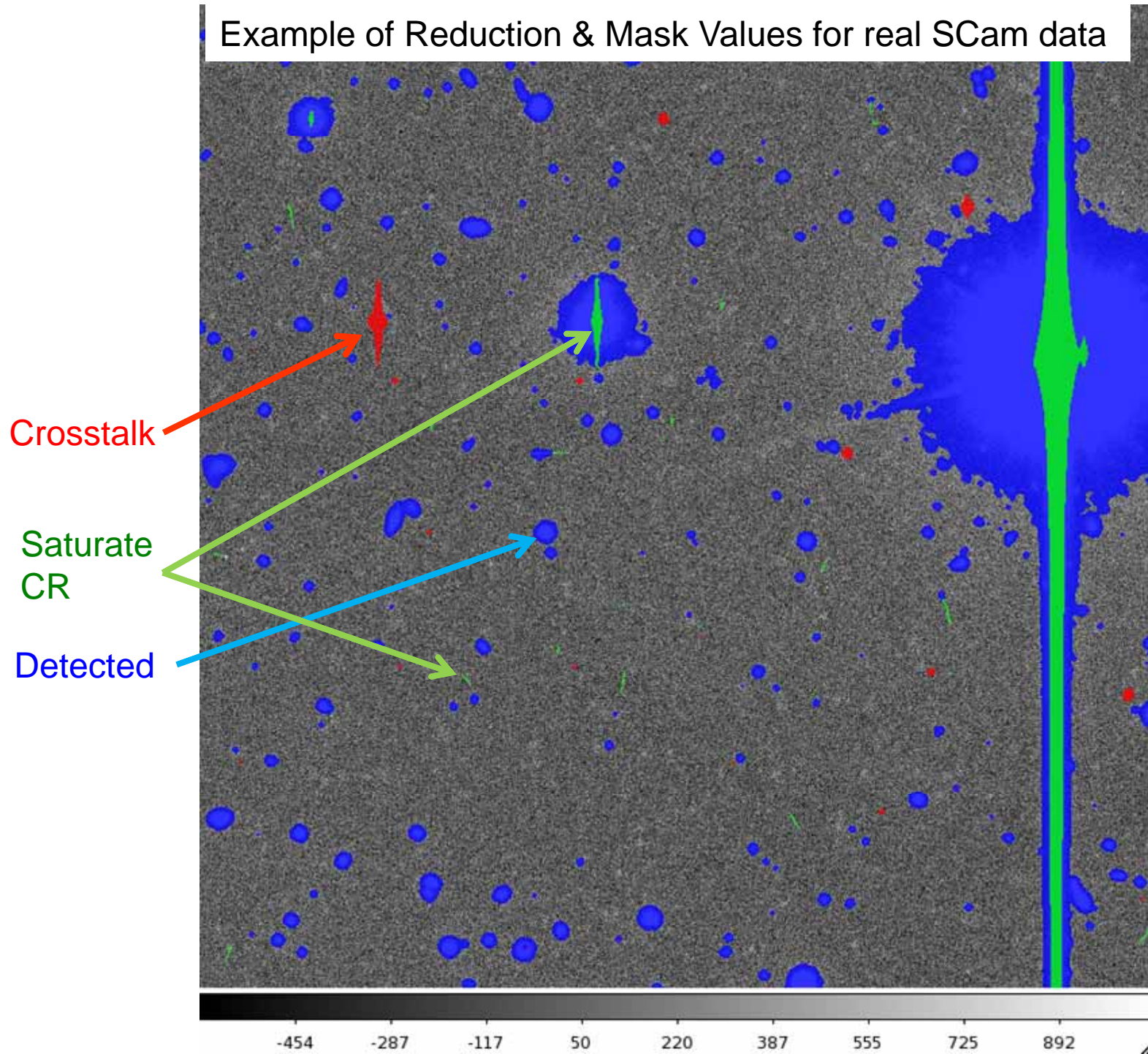
Mapping Point Spread Functions

- SDSS approach of PSF determination by PCA

- Spatial variation in PSF fitted by polynomials

Eigen PSF images

Reconstructed PSFs

Hironao Miyatake
Paul Price
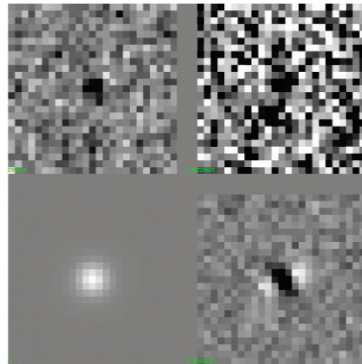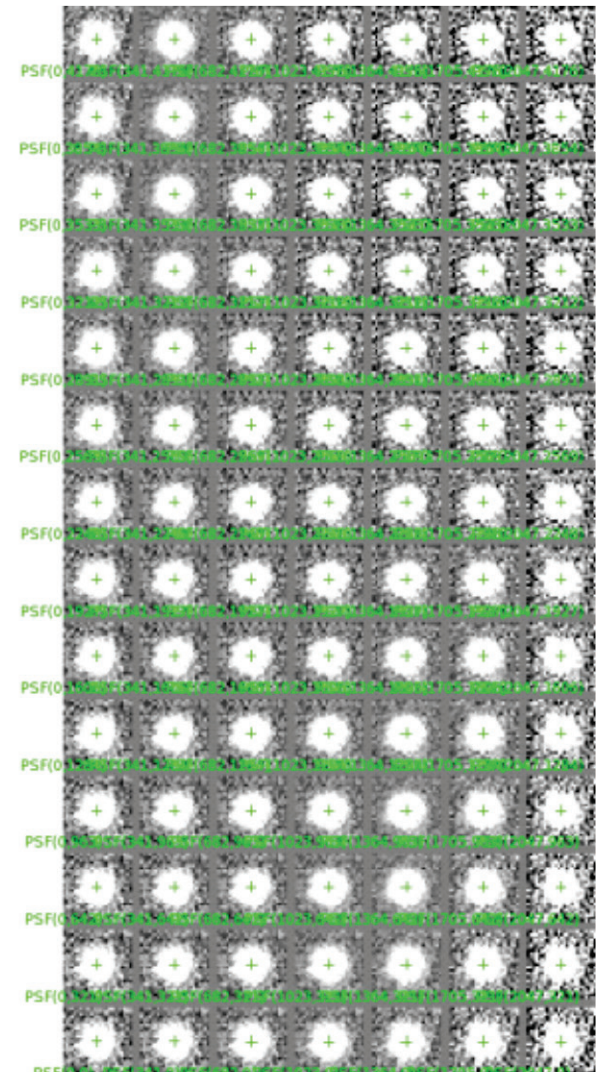
Photometric Calibration

- Calibration by SDSS stars
- To 1-2% scatter in ideal data (dense stellar field with expt = 30s) after color term correction
- Can be to 4-5% in deep image (in z band)
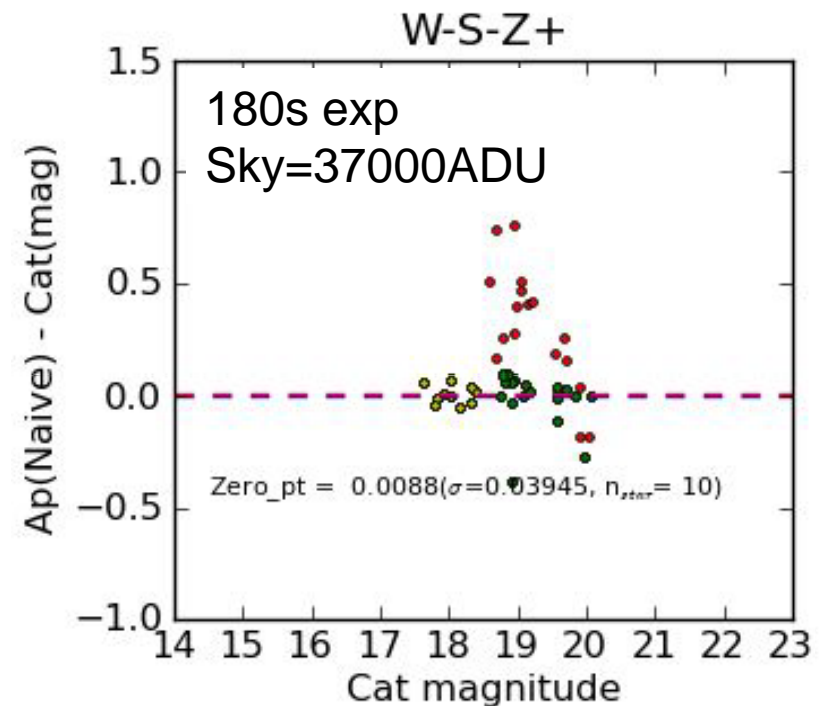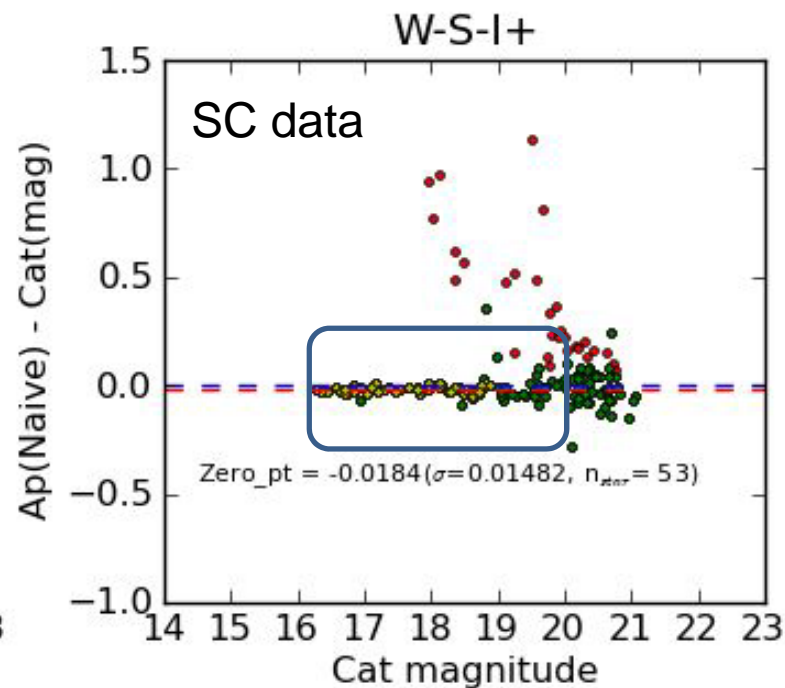
5. Mosaic-Stacking

- Astrometric Calibration
  - Improves WCS of each CCD image

Source $(x, y)_i$ on CCD

Set of source positions
Incl. SDSS stars

Least-square fitting

CCD (X, Y, θ)

TAN-SIP WCS coefficients
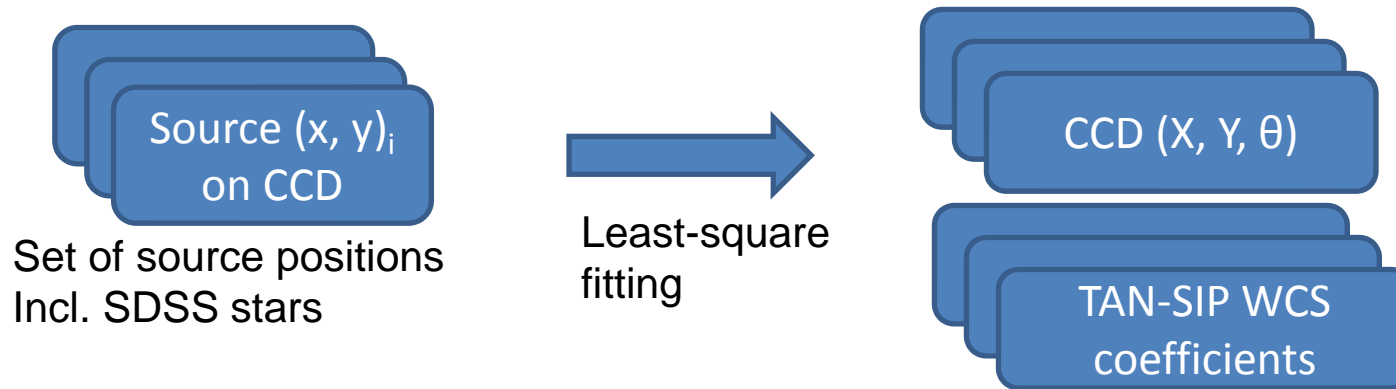
- Photometric Calibration
  - Determines flux scaling (magzero) of each CCD image

Source mag $(x, y)_i$ on CCD

Set of source magnituds
Incl. SDSS stars

Least-square fitting

magzero

+ Flux Correction

12
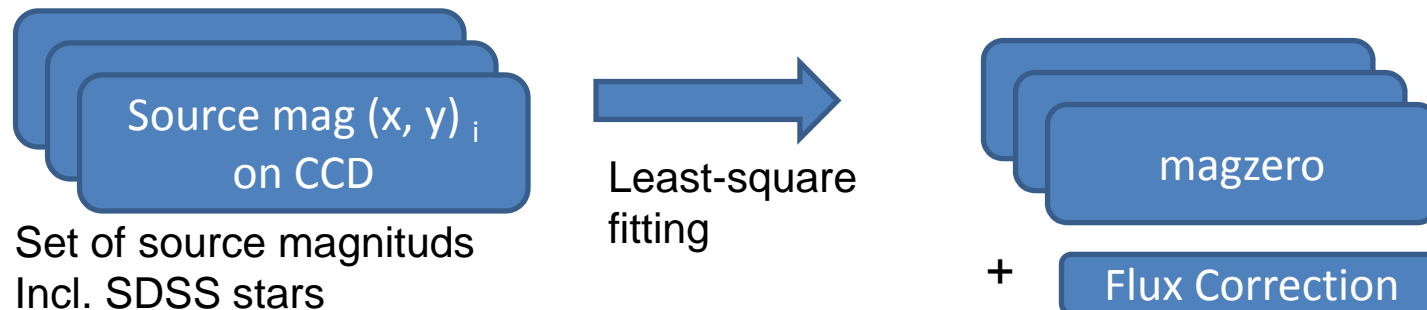
Diagnostics of Mosaicing Result

(1) Fitting scatter of astrometric alignment

0.01 arcsec rms (Internal)
0.03-0.04 arcsec rms (External to sdss)
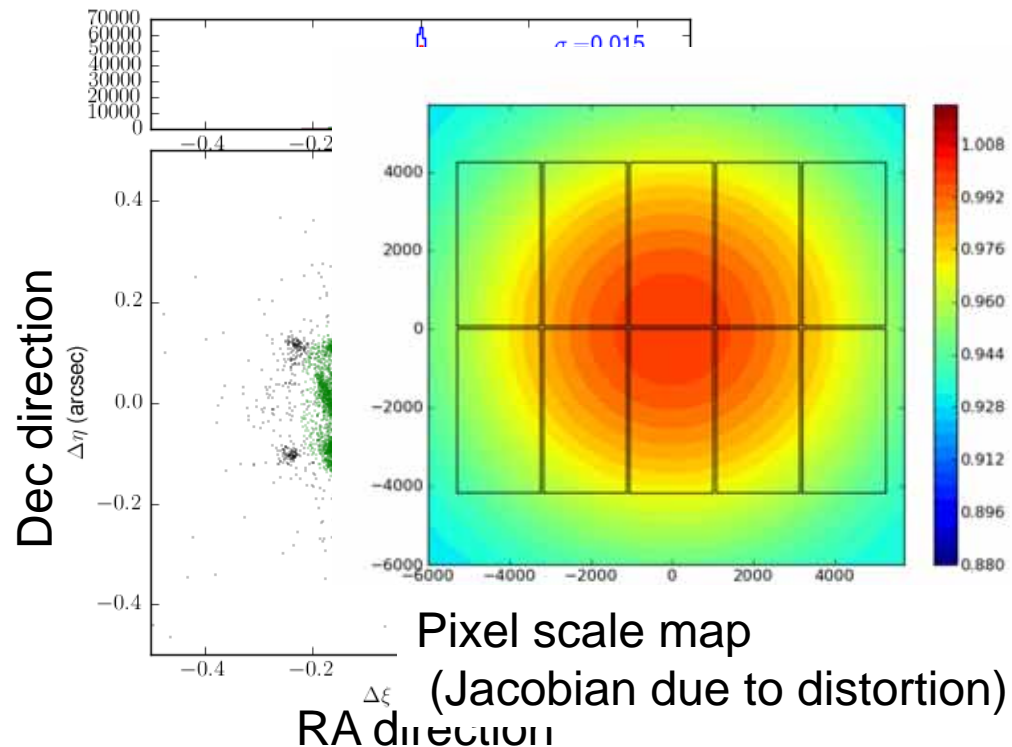
(2) Fitting scatter of flux scaling

0.01 mag (Internal)
0.03-0.04 mag rms (External to sdss)
after flux correction applied

Pixel scale map
(Jacobian due to distortion)

Dec direction

RA direction

magnitude
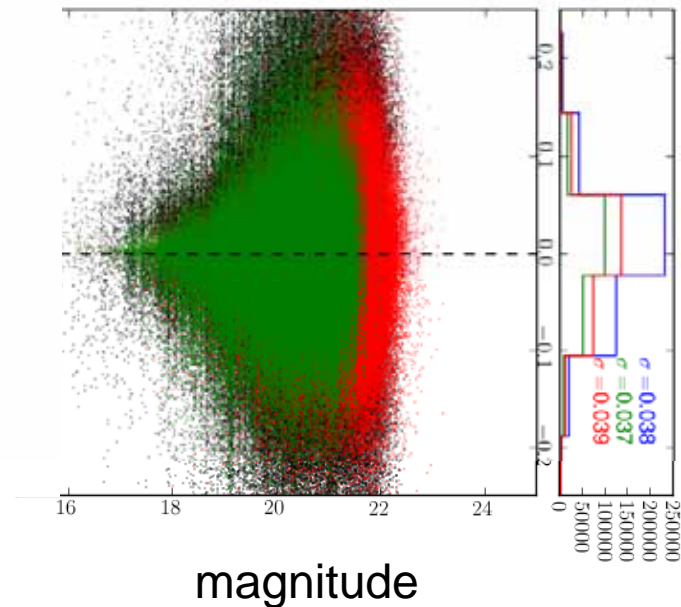
Flux Correction Pattern by Mosaicing

hscMosaic can provide additional flux correction to magzero in each CCD

SCam
engineering data
30sec x 18 exposures
at stellar dense field
(~600stars / chip)
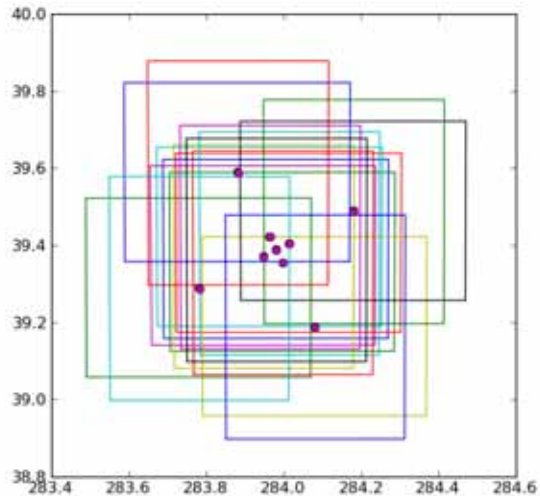
Experiment by Yasuda-san

Remaining scatter (Internal) dm ~ 0.01 rms
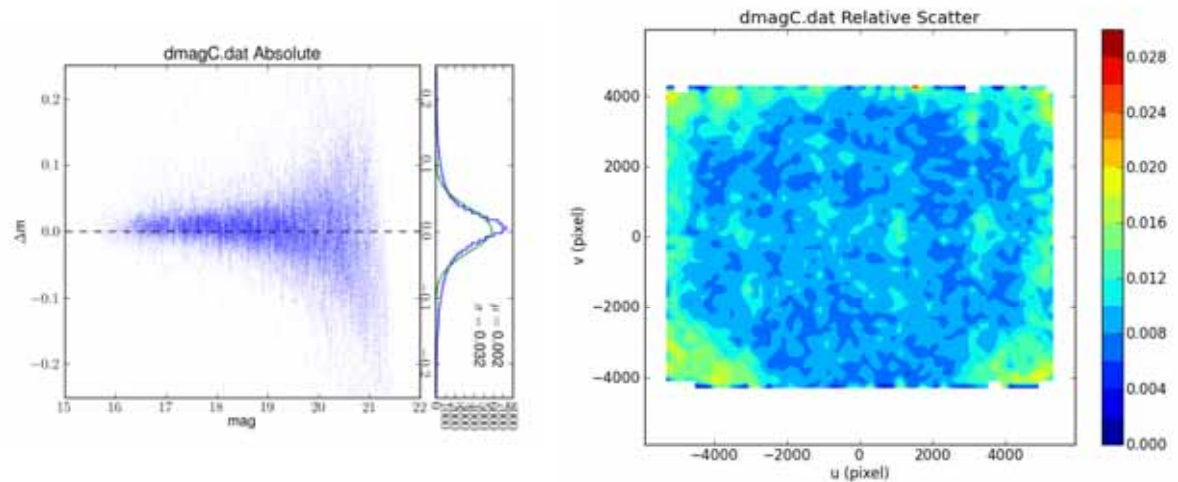
Remaining scatter (SDSS sources) dm ~ 0.03 rms
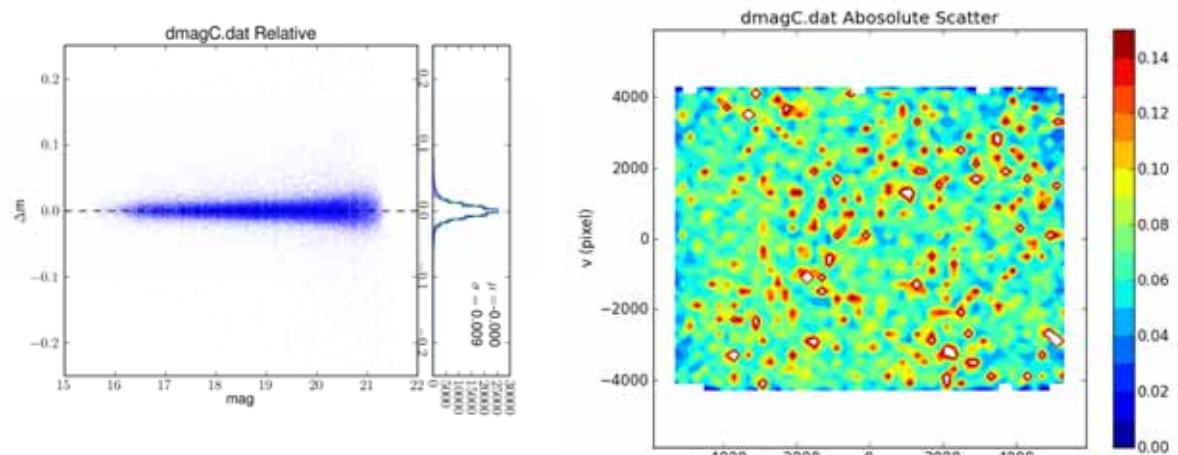
Stacked Image

- Suprime-Cam z-band
- 575 shots
  - 16 hrs (2FOVs), 8hrs (2FOVs)
- Time
  - 2hrs / mosaic
  - 5days / stack (single thread)

Mask
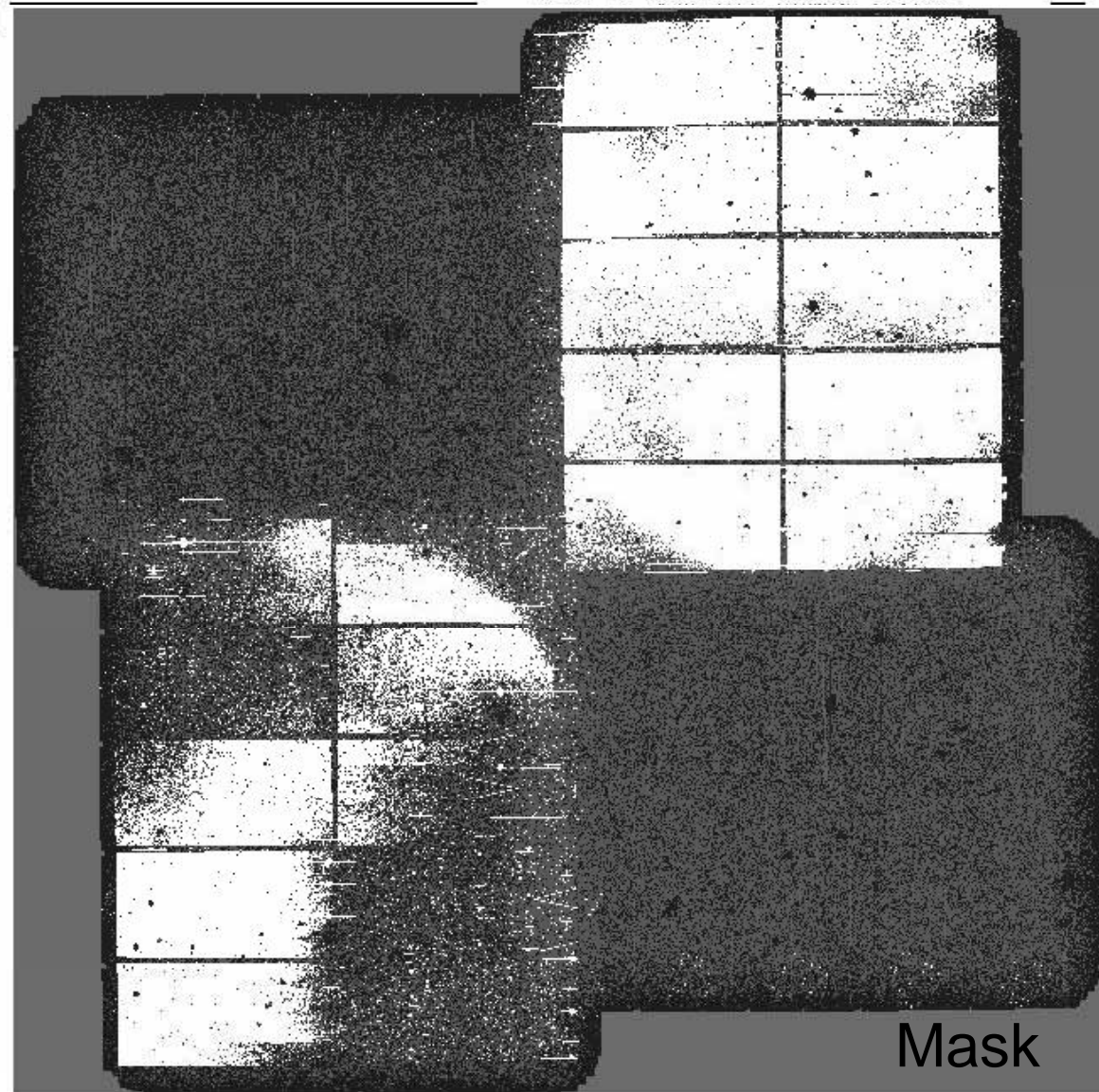
Comparison

- Exam of results just started, comparing with SDFRED
  - COSMOS, UDS, HDF …
  - Depth comparable, with noise less correlated?
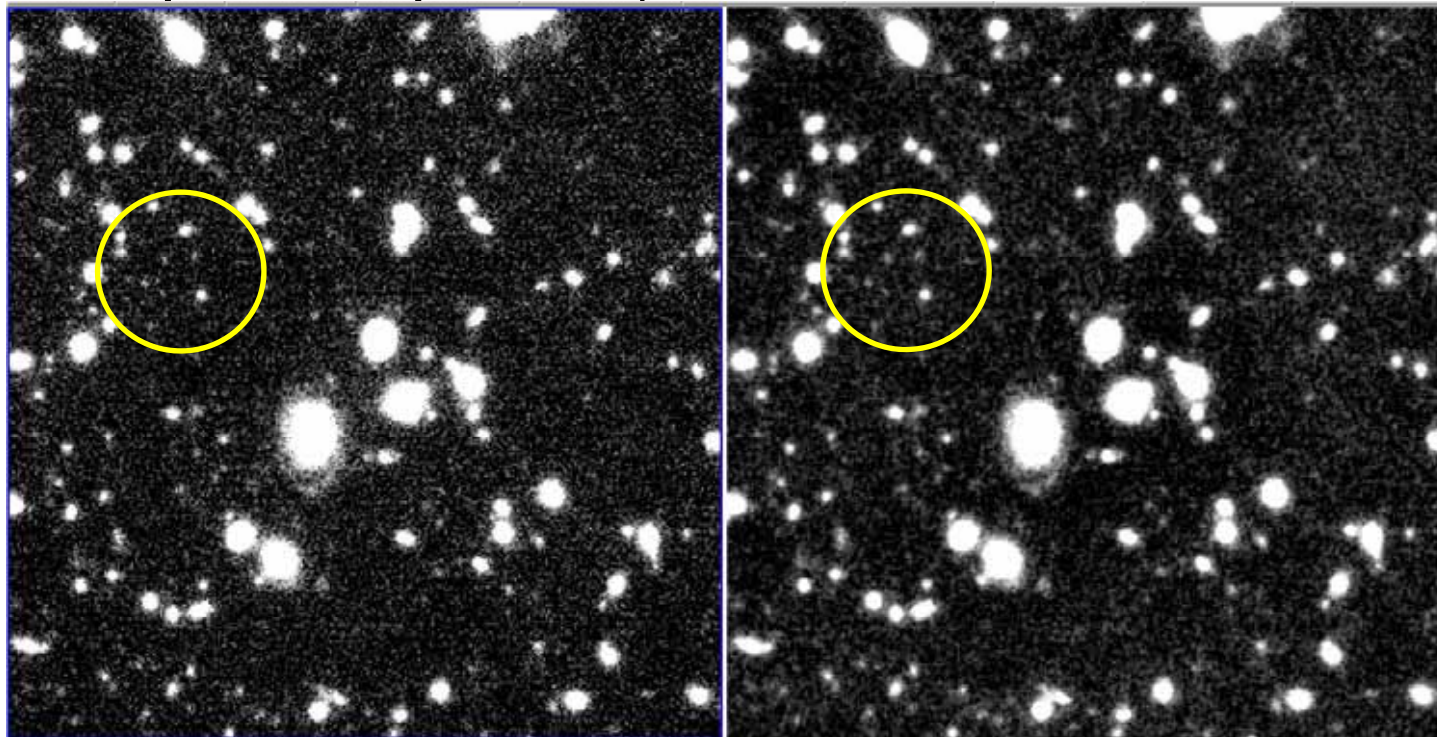
hscMosaic   SDFRED

6. Catalog Creation
Forced Photometry

- Detection on one image & Measurements in each image → Single multi-band catalog

Each-band Image

Measurement

Detection Image
→ χ^2 Image
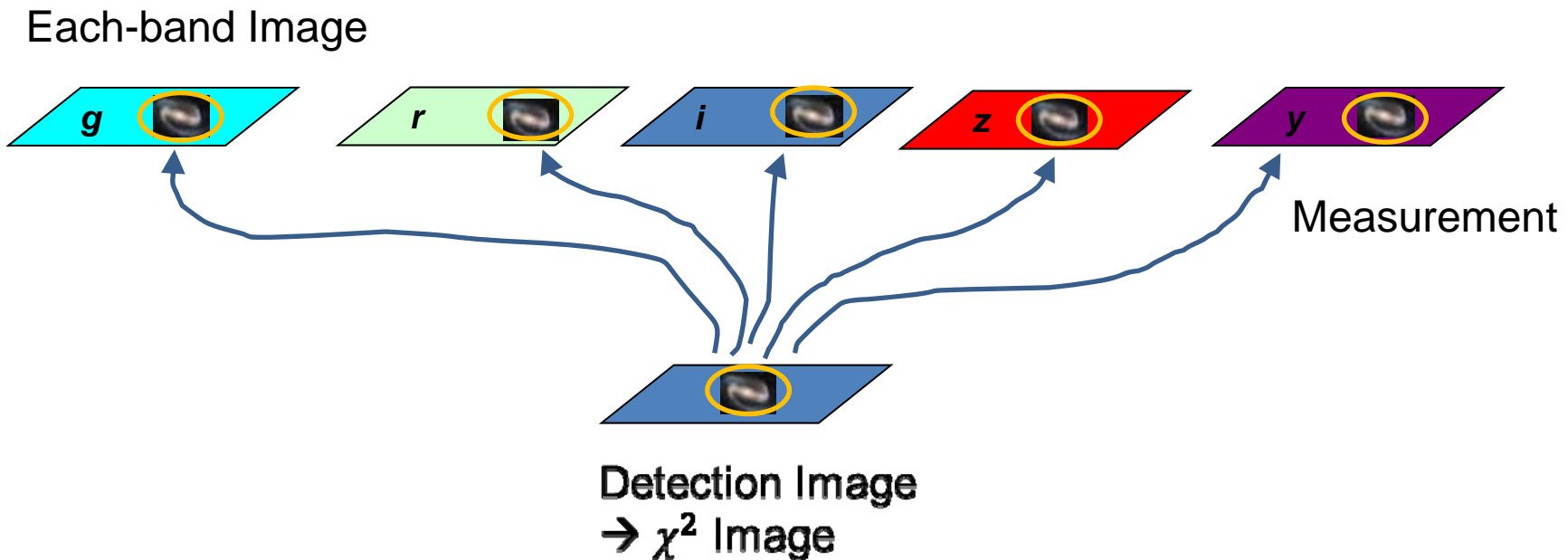
$$y = \sum_{i=1}^{N_{\text{band}}} g_i^2$$

g_i = Image – sky : Sum of Bkg counts of all bands
→ Statistical detects significant objects

17

Catalog Creation

- **Single-CCD Catalogs**   (*) Advanced / Challenging Items
  - Direct measurements per each band, each frame
  - (*) Forced photometry on each frame
- **Multi-band catalogs for stacks**
  - Forced photometry on stacked images
    - (*) χ^2-Image detection
  - (*) Consideration of multi-epoch data
  - HSC standard catalogs for HSC bands
    - First, broad bands only, then, narrow bands

7. Production of Mosaic-Stack + Catalogs in SSP

- On each pre-defined 'SkyTile' (~ 1.5 deg wide)

- Corrected frames TAN projected onto SkyTile's

- Full stack & catalog production in 6 – 12 months interval (TBD)? in stable operation phase

Each HSC shot

SkyTile

Mapped with spatial indices (e.g., Healpix) for data handling with DB

SkyTile divided into sub-patches for efficient stacking/cataloging
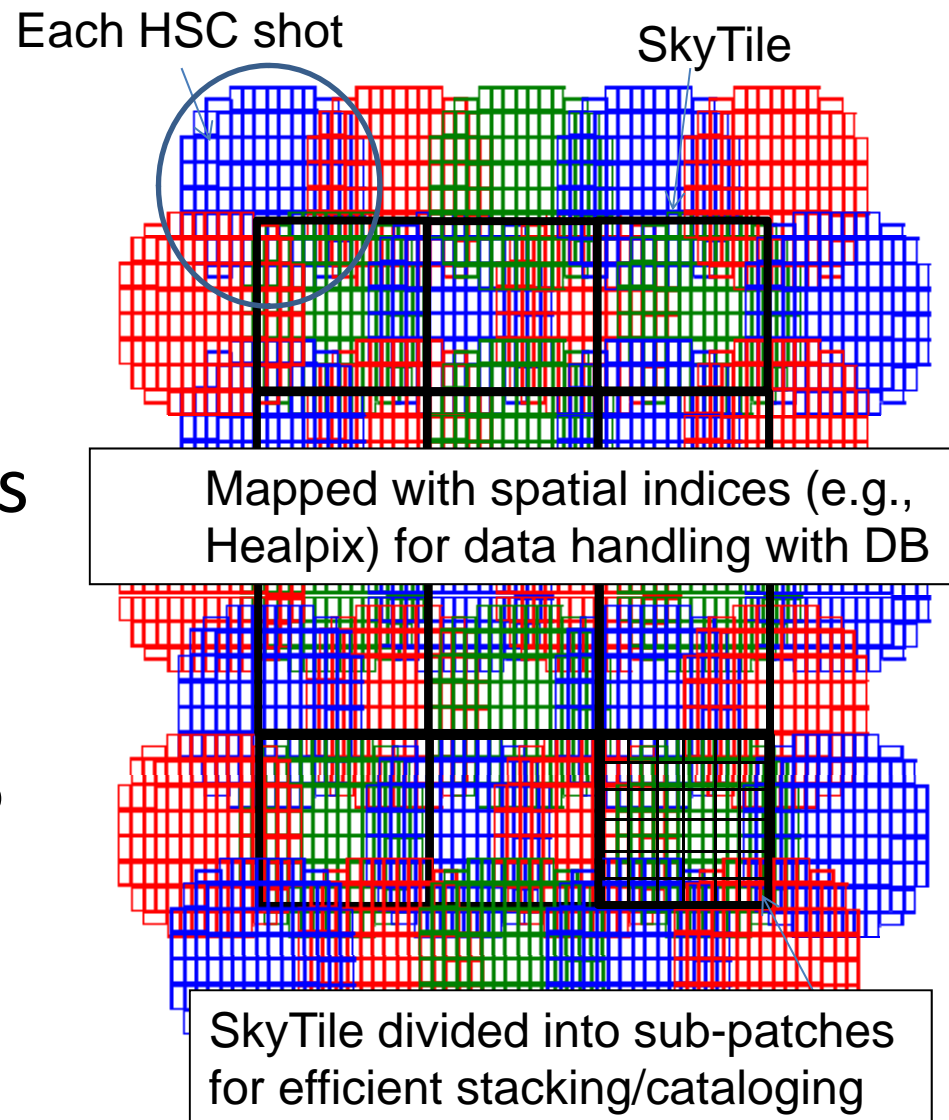
8. Ongoing/Remaining Concerns

- Measurement parameters
  - e.g., Kron-mag, Petrosian-mag, Shape-fit parameters (r_eff, etc..)
- Reliable sky subtraction
- Optimization of mosaic, stacking (especially for deep, many frames)
- Multi-band catalog creation with forced photometry & chi-square detection
- Exam of images & catalogs (calib., accuracy of measurements)
- Procedure of SSP data release production
- Support of general observation

9. Assisting SSP observation

- Overview of Hilo onsite system
  - realtime quick analysis for Data evaluation (QA) & visualization
  - data screening by QA database
- Prototype of progress map

HSC Onsite-Analysis Cluster at Hilo Facility

Network Switch

- 10G
- 10G
- 1GbE

File server (NFS) — FC8, IB — 28TB — Rawdata

File server (NFS) — FC8, IB — 30TB — Processed data

x3

Master node Torque, pBASF — FC8, IB

IB-QDR 40

Slave node Frame analysis — FC8, IB

x5 — Data processing

Mosaic Interactive node — IB

Mosaic analysis

Interactive analysis tools (IRAF etc)

Database — IB

fileInfo, analysireResults

Web/SurveyInfo — B
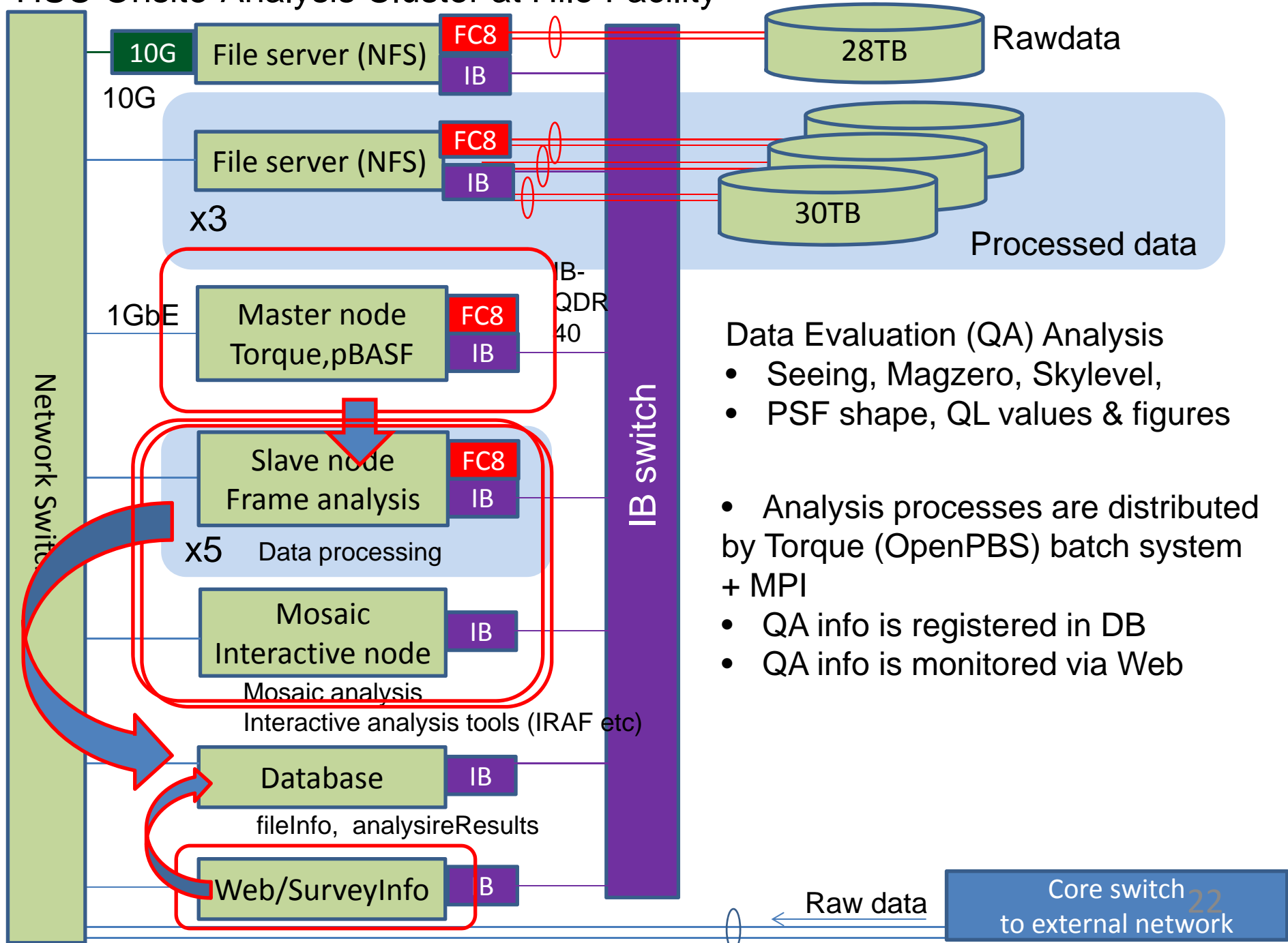
IB switch

Data Evaluation (QA) Analysis
- Seeing, Magzero, Skylevel,
- PSF shape, QL values & figures

- Analysis processes are distributed by Torque (OpenPBS) batch system + MPI
- QA info is registered in DB
- QA info is monitored via Web

Raw data ← Core switch to external network

22

Plots from Seeing Measurements
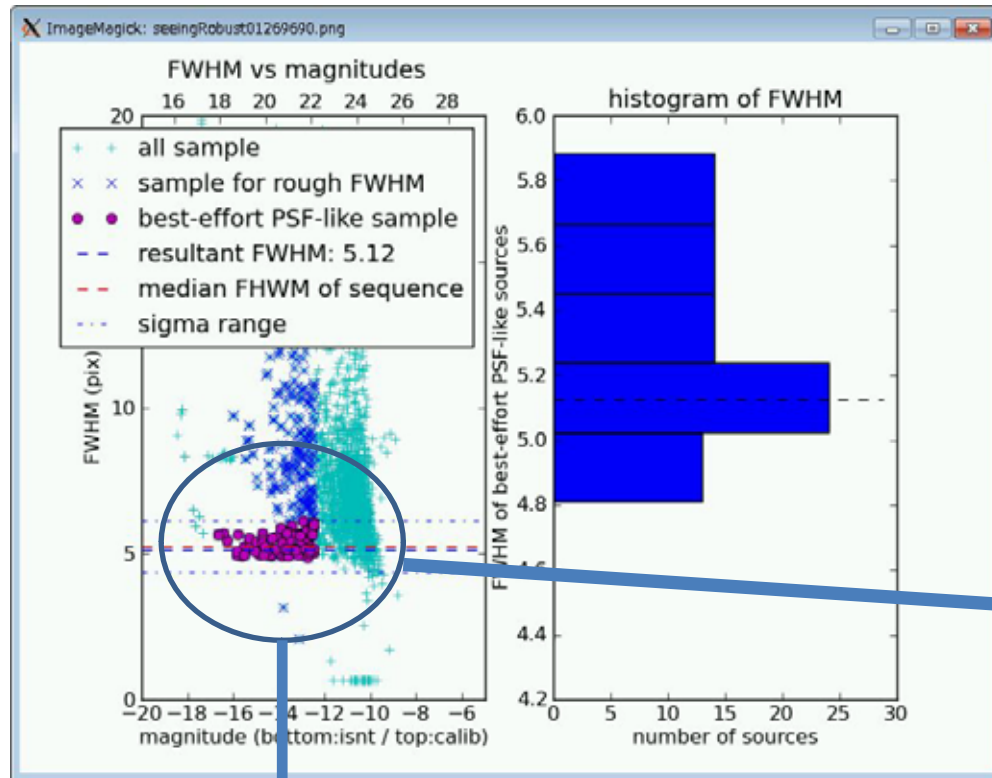
PCA modeled PSF

Selected PSF candidates
Mode is adopted for robust FWHM
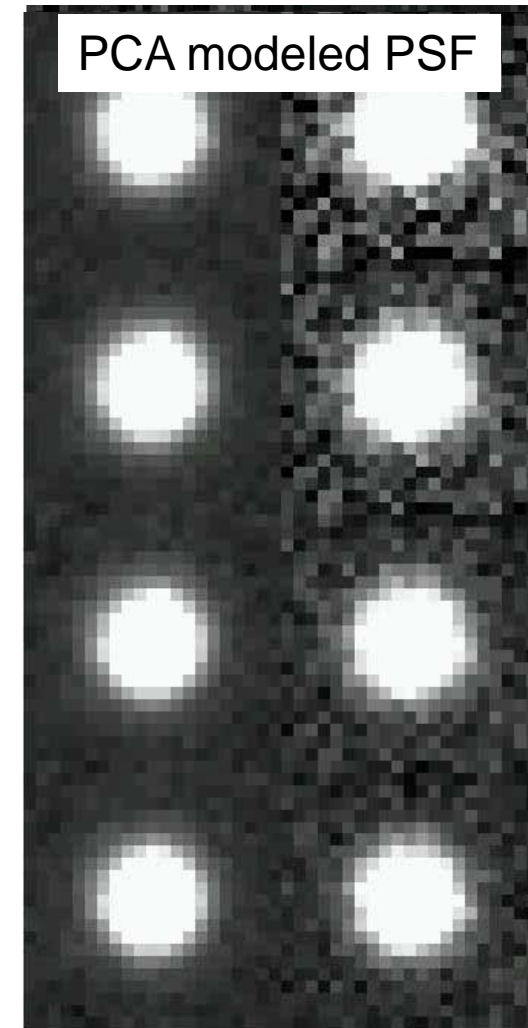
PSF mapped in Grids
Both with stack of sources & with PCA models

Tiled QA Figures

- Demonstration with HSC simulation data
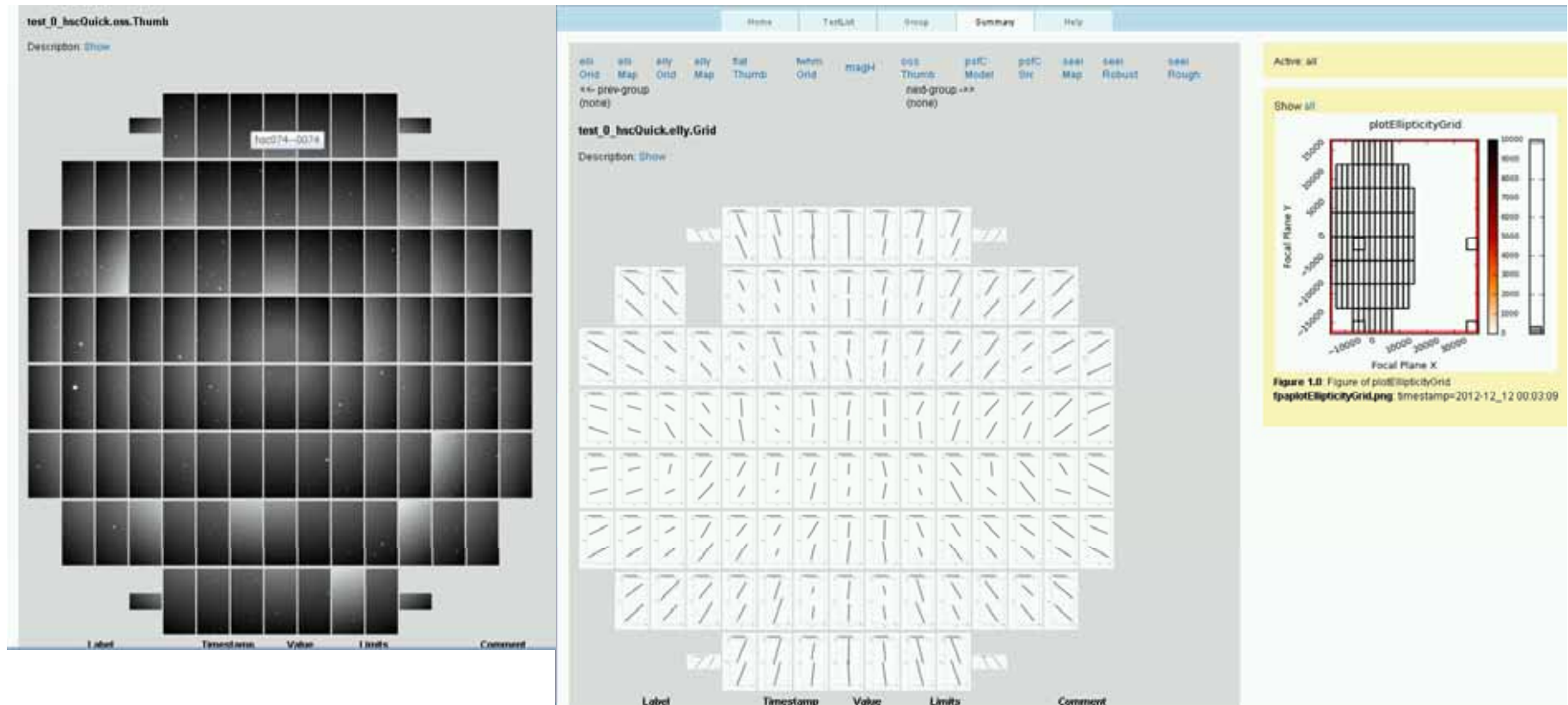
QL of pverscan-subtracted images

Ellipticity of PSF-like sources mapped on FOV

Prototype of Realtime Quality Monitor
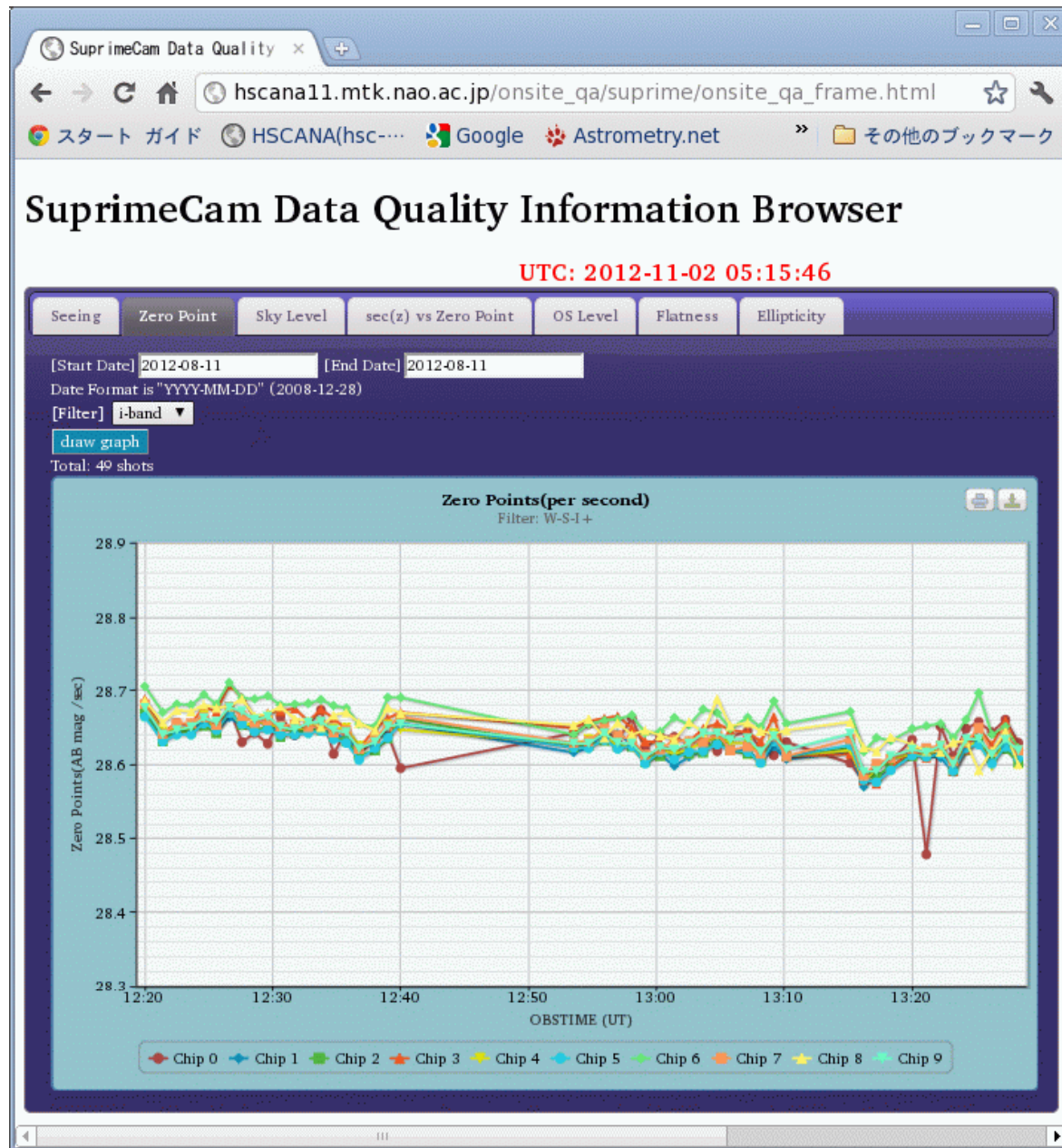
Temporal change in magnitude zeropoint, seeing, skylevel …

Timing of Onsite QA

- Analysis cycle of 2.5 minutes after data arrival
- Minimalist pipeline for QA, with localdisk writes
  - → approaching targeted ~ <2min

Detection	Disk Write	FrameAna (CCD-by-CCD)	ExpAna (PostProcess)	Total
Bright + Deep Sources (10sgm, 2sgm)	NFS	03m45s	00m25s	04m45s
Bright Sources (10sgm)	Local	01m45s	00m25s	02m30s

Tools for Survey Progress Check

- Visualizing achievements in each survey area
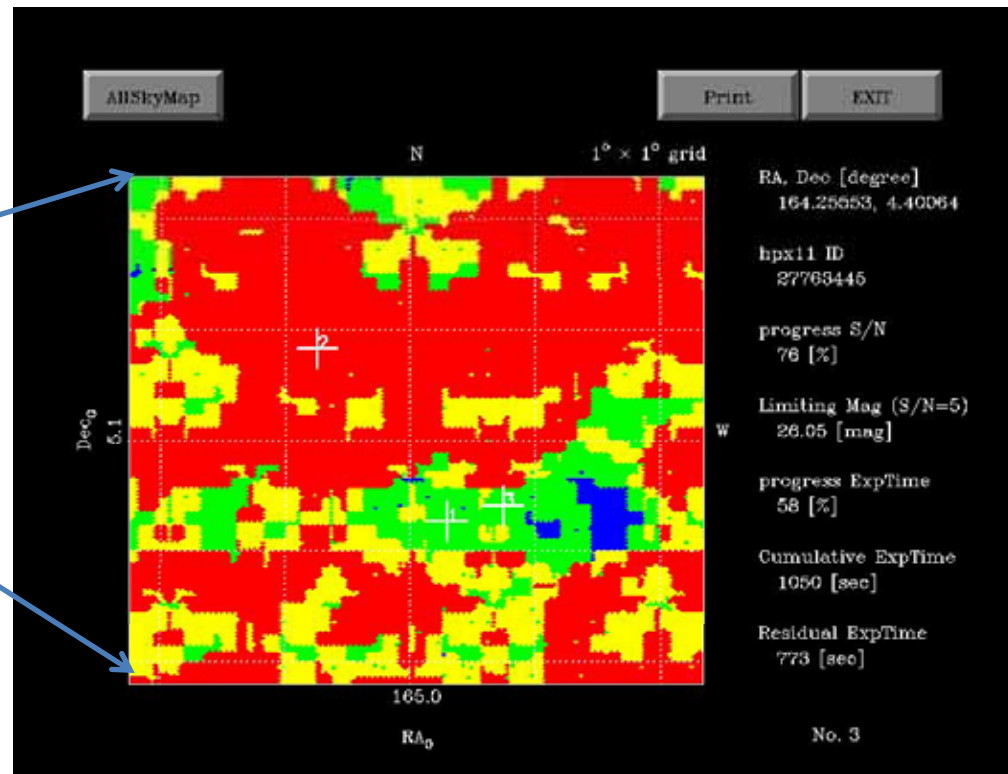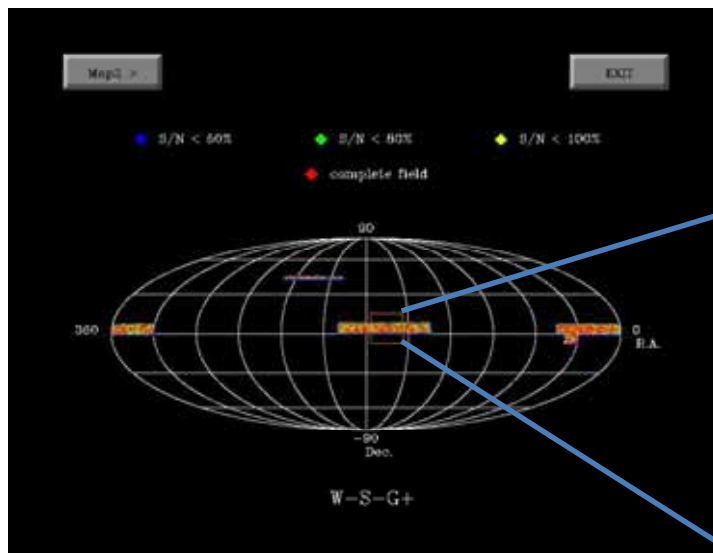
Map of Expected S/N
based on
seeing, magzero, & sky sigma

per HEALPix index

10. Products Distribution

- Archive Database
  - Shares analysis DB & onsite-QA DB
  - All or part of products data & database will be mirrored at core institutes
- Portal web, user interface to access products
  - Image data through DataArchiveServer (DAS)
  - Catalogs through CatalogArchiveServer (CAS)

User Portal for Data Retrieval (DAS)

Prototype by H. Yamanoi