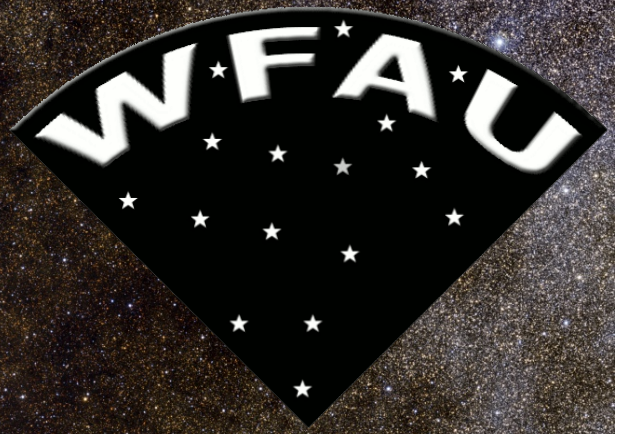


WFAU Science Archives: Facilitating Survey Astronomy

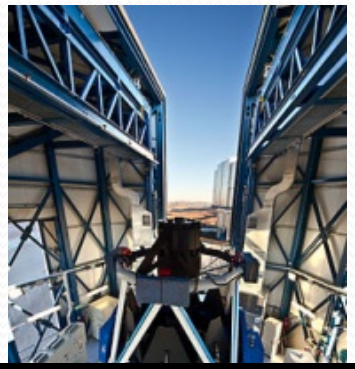
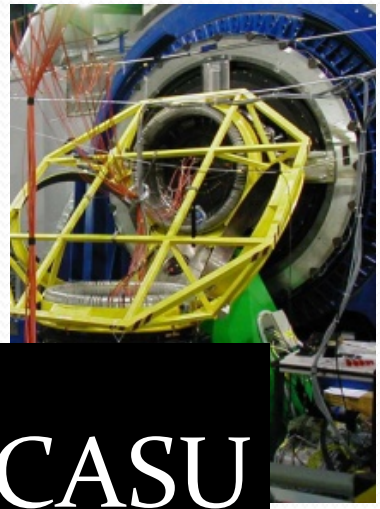
Nicholas Cross, Rob Blake, Ross Collins, Clive Davenhall, Mark Holliman, Mike Read, Eckhard Sutorius, Nigel Hambly, Dave Morris, Stellios Voustanis, Bob Mann, Keith Noddle

Wide Field Astronomy Unit, Institute for Astronomy, University of Edinburgh, UK

njc@roe.ac.uk

WFAU Science Archives



EARLIER PROCESSING AT CASU

See Gonzalez-Solares talk

WFCAM Science Archive

VISTA Science Archive

OmegaCAM Science Archive

Gaia ESO Survey

WFAU

Roles of the Science Archive

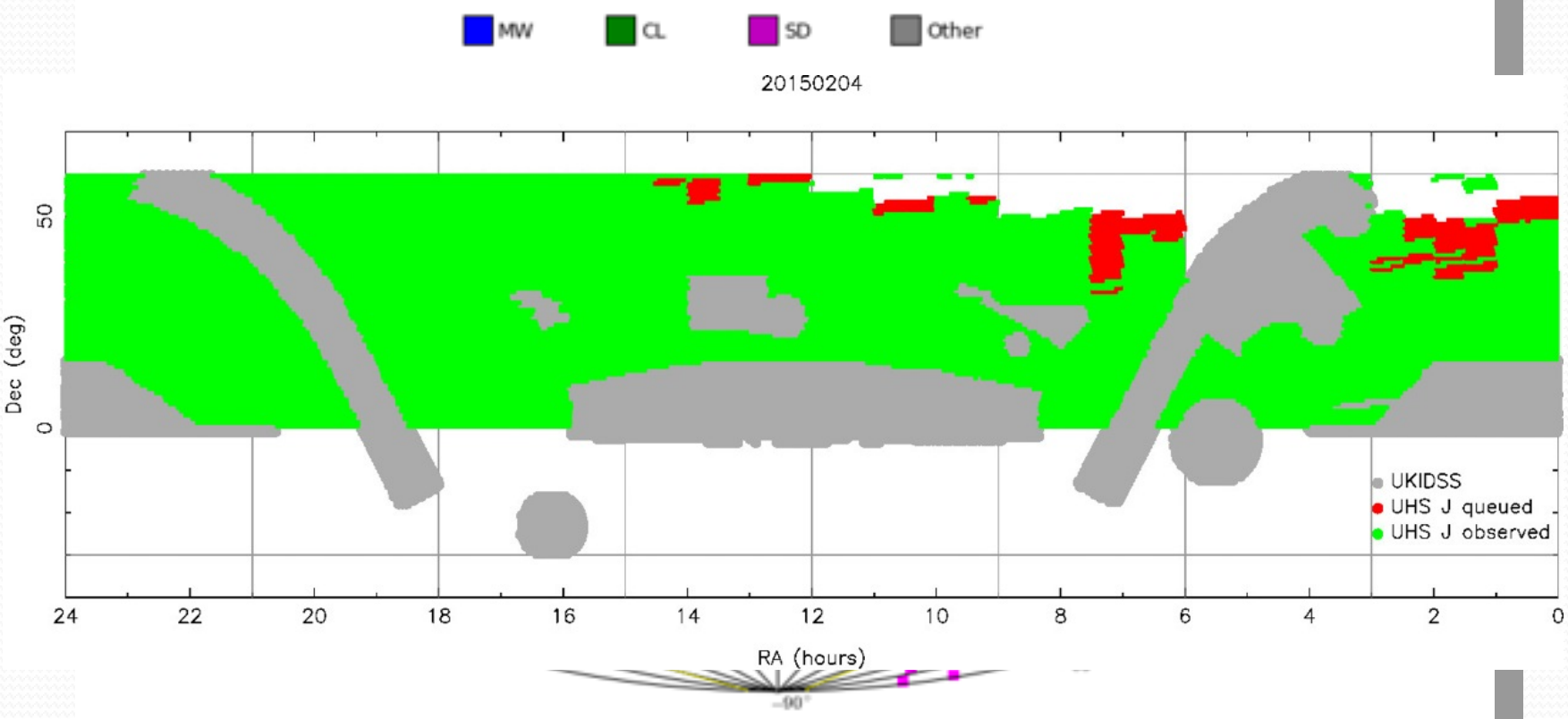
- Repository of (e.g. VISTA) data from reduced images to complex, catalogue products.
- Interface to do science – select data.
- Interface for teams to check data for quality control.
- Requires both:
  - a continuously updated dynamic main-DB: new data, better calibration, reprocessing, quality control, higher order products.
  - Static, well documented release-DBs that can be referred to in publications.

Common Design

- WSA, VSA, OSA developed through VDFS – VISTA Data Flow System
- MS-SQL database system based on SDSS SkyServer
- Similar inputs: FITS files generated by CASU
- Common software architecture, data pipeline, user interfaces, hardware architecture.
- GES is significantly different: spectroscopic, wider range of data from multiple groups, different interfaces

Surveys

WF VST: ATLAS, [VPHAS+] [ING, VIDEO



Archives: http://surveys.roe.ac.uk/

Imaging Surveys Pipeline

- Ingest from CASU (and teams)
- Provenance - link related images
- Quality Control: Automated + input from teams.
- Process data for semester - done per programme:
  - Produce and ingest deep stacks/tiles/mosaics + catalogues
  - Merge pass-band catalogues to create source tables
  - Create neighbour tables to link external catalogues
  - Link multi-epoch data and calculate variability statistics
- Release a documented, static data product to users
- Convert data to ESO format and copy to ESO (VISTA, VST)

Gaia-ESO Spectroscopic Archive

- Spectroscopic survey of ~100,000 stars, covering all major components of MW using VLT-FLAMES
- Archive includes:
  - list of targets with photometry used in the target selection
  - spectra processed by CASU
  - atomic and molecular line lists
  - Recommended astrophysical parameters derived by various working groups:
    - T_{eff}, log(g), abundances …
  - Neighbour tables to other surveys: SuperCOSMOS, VST-ATLAS, 2MASS, VISTA-VHS
- No public release yet. 3rd internal data release.

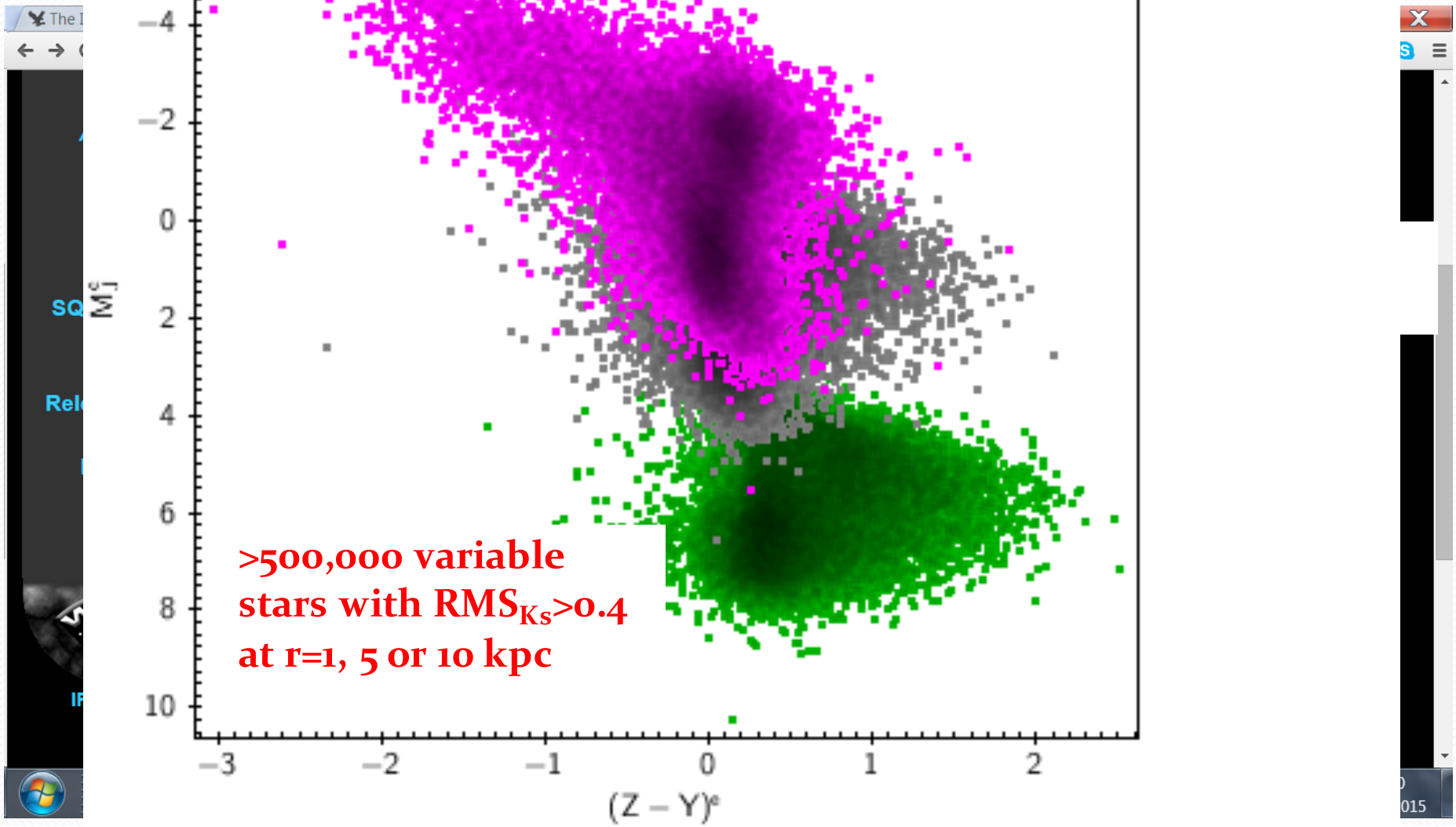
Value Added Data

- Data created by science teams:
- VMC :
  - PSF catalogues
  - Variable types (Cepheid, Eclipsing Binary)
- VVV:
  - 3D extinction maps
  - PSF catalogues (multi-band 1 epoch) + DIA light-curves [Discussions with Javier Alonso-Garcia and Eamonn Kerins]
  - Proper motion + Parallax [Not yet delivered]
- GES:
  - Wide range of parameters from multiple working groups.
- Work to improve connections of team generated data with CASU/WFAU generated to maximise usefulness.

3D Extinction Maps

- Included VVV extinction maps of bulge (Chen et al. 2013)
- Match to vvvSource.
- Filter conversions for VISTA, UKIRT, VST, Spitzer and WISE filters
- Will add in more (IPHAS, PanSTARRS …)

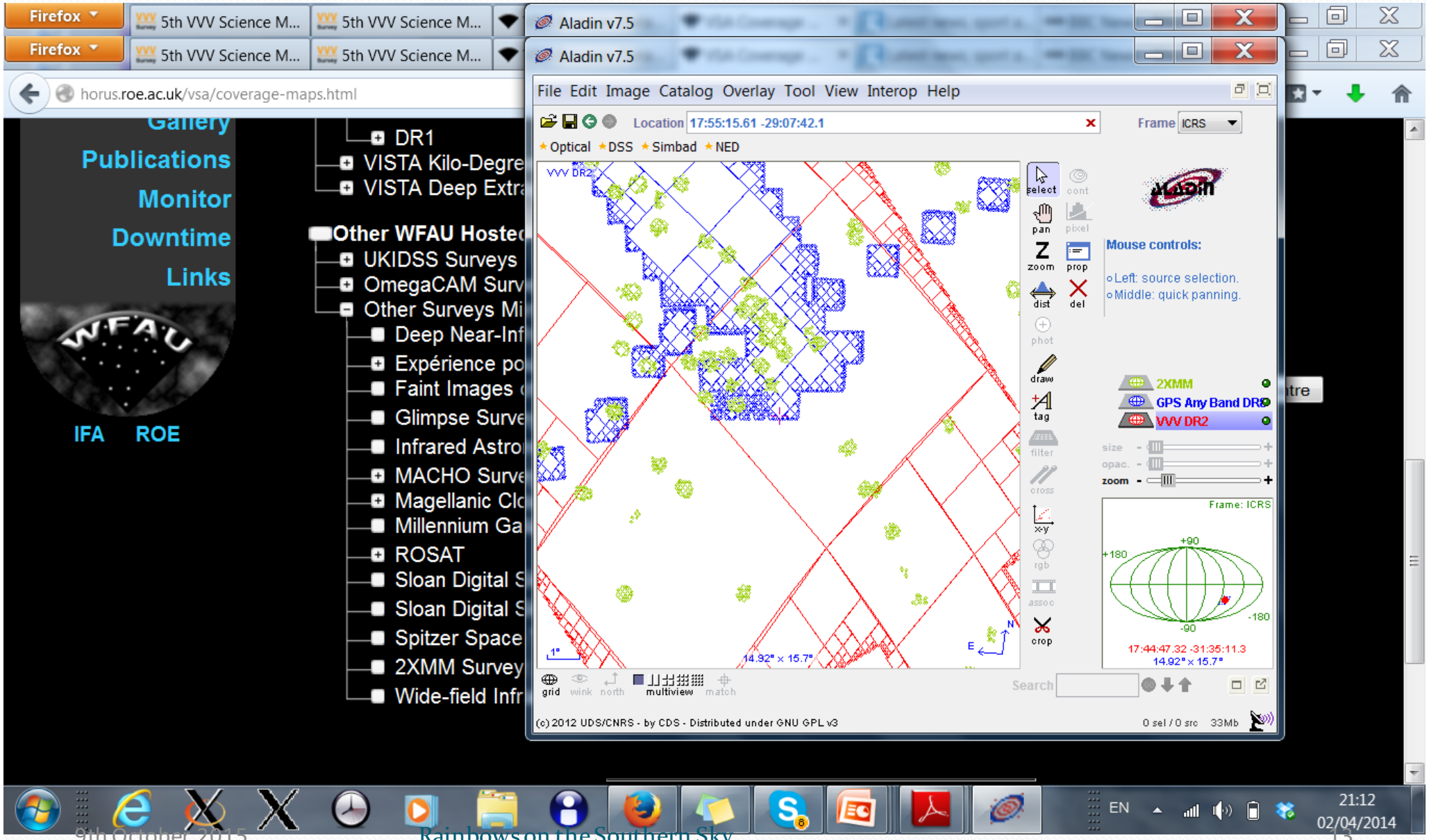
3D


1 : r = 1.
1 : r = 5.
1 : r = 10.
−6
−4
−2
0
2
4
6
8
10
M_J^e
−3
−2
−1
0
1
2
$(Z - Y)^e$
>500,000 variable
stars with RMS_{Ks}>0.4
at r=1, 5 or 10 kpc

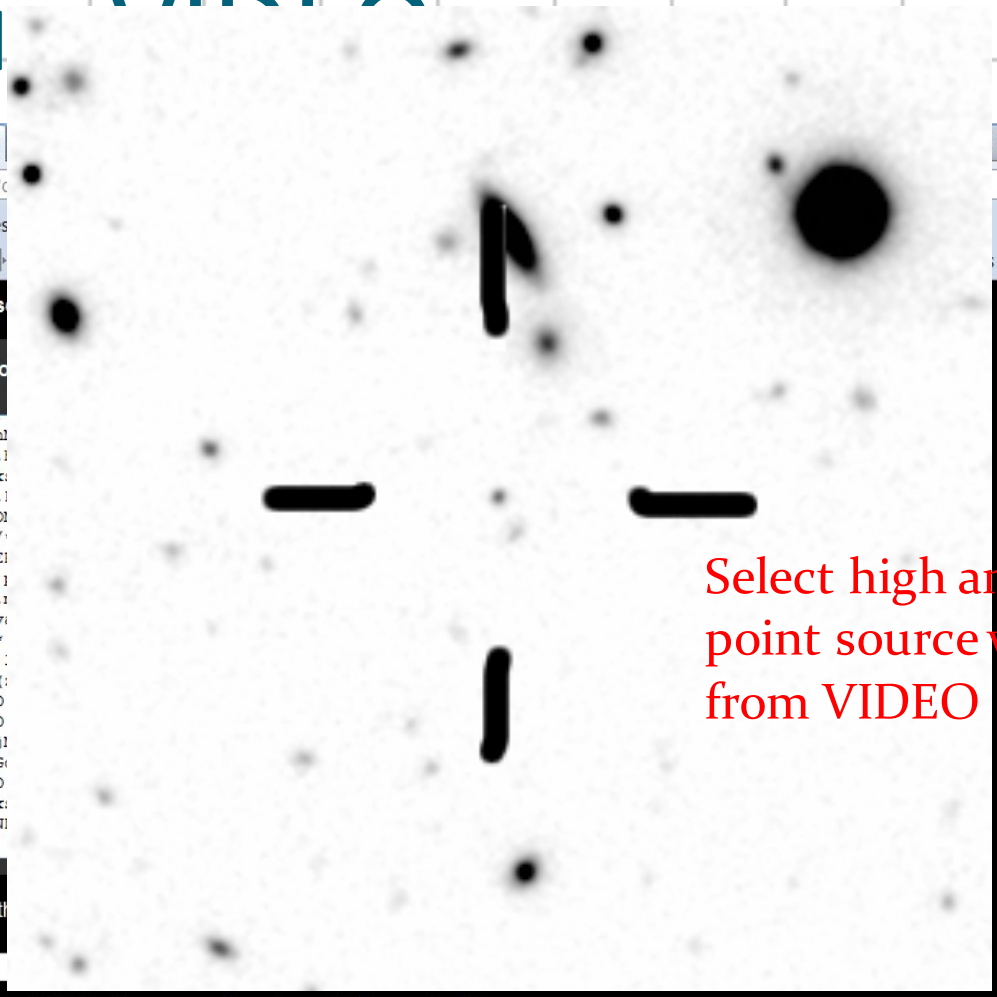
Survey Interfaces

- Coverage Maps – Multi-Order Coverage
- Archive Listing - Image/Flat file download
- GetImage / Multi-GetImage – Image cutouts
- Region – Cone search
- CrossID – Catalogue matching
- Freeform SQL – Most powerful catalogue query tool
- ADQL (OSA only so far) – Astronomical Data Query Language. To do distributed queries between archives (e.g. Gaia, PanSTARRS, LSST)
- Firethorn – MyDB works with ADQL for distributed multi-wavelength work (Under development)

Coverage Maps



An Example: Selecting point source variables i

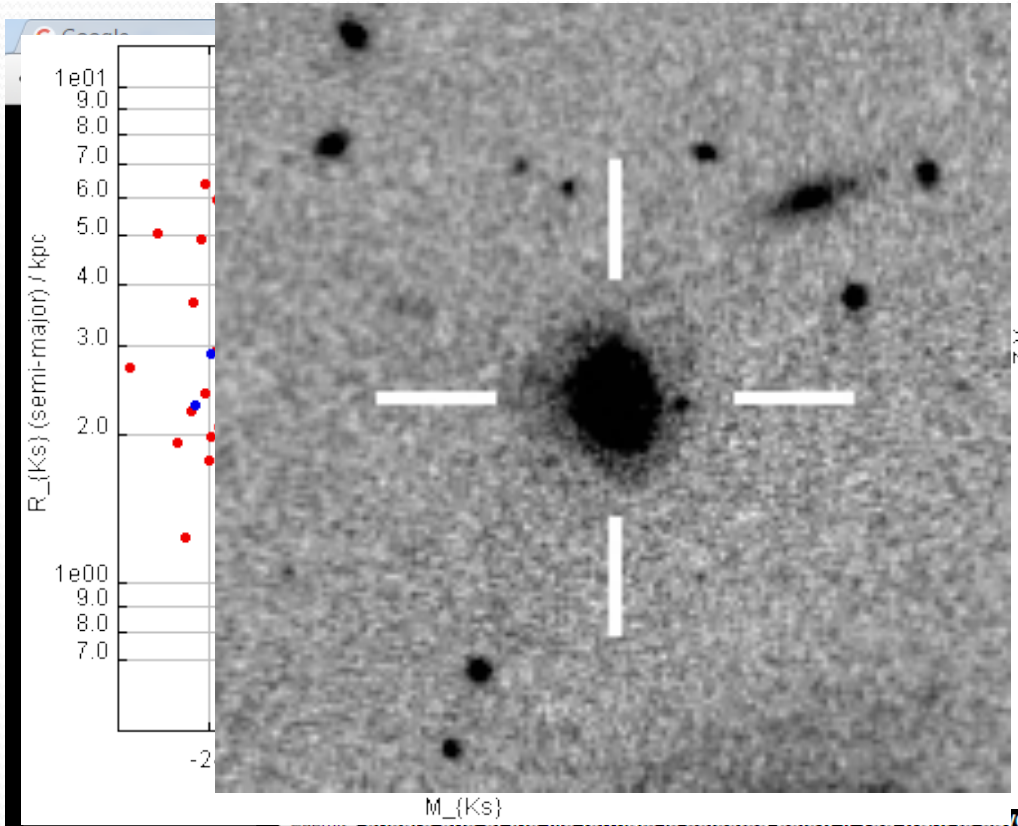


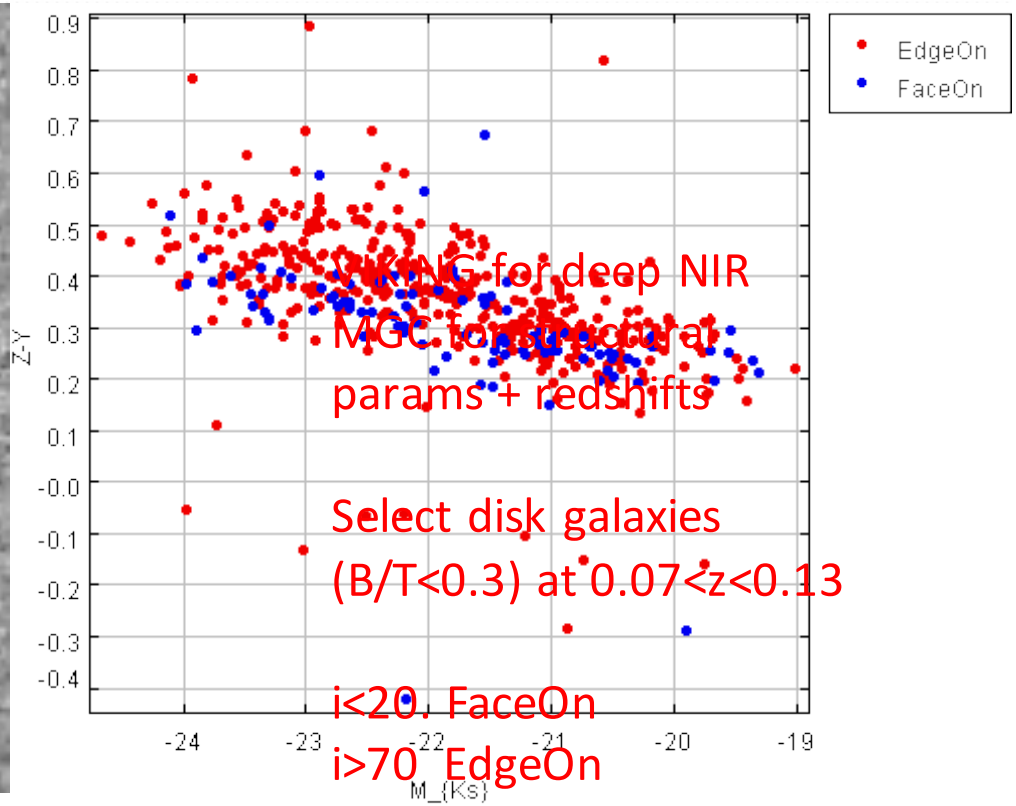
Select high amplitude point source variables from VIDEO

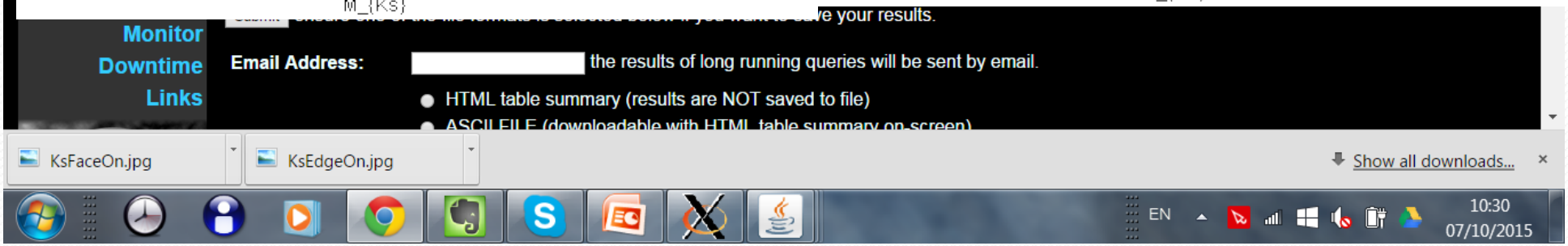
Surveys are not independent entities

- Importance of multi-wavelength astronomy
- Common survey fields
  - VIKING and KiDS
  - VST-ATLAS and VHS
  - UKIDSS LAS/UHS and SDSS
  - VVV, GPS, IPHAS, VPHAS+
- Need for data integration
  - Cross-neighbours tables, publishing to VO
    - ADQL + Firethorn to be able to link to any published.
  - Matched aperture photometry

External Datasets







Data Volumes

- Images:
  - WFCAM 1720 raw image frames a day (4 2kx2k)
  - VISTA 580 raw image frames a day (16 2kx2k)
  - VST 100 raw image frames a day (32 4kx2k)
  - ~1PB of image data
- Catalogues:
  - VVV completely dominates: combination of high source density, short exposure time.
  - Next VVV release will be 45TB (Few weeks away)
  - 55 billion detections.
  - Largest astronomical SQL database?
  - Final VVV+VVVX (?) >100TB

WFAU VDFS Publications

Hambly et al. 2008, MNRAS, 384, 637 (WSA)

Cross et al. 2009, MNRAS, 399, 1730 (Multi-Epoch processing)

Cross et al. 2012, A&A, 548A, 119 (VSA)

Various other technical documents:
http://surveys.roe.ac.uk/vsa/publications

Future work – some highlights

- Matched Aperture Pipeline:
  - imcore_list & SExtractor
  - Between surveys, inter survey, user defined
- Variable Star Selection:
  - Add in improved indices for variable selection
- Much more value added data
- MOC driven cone-searches
- MOONS?

Archive Usage