Synoptic Data in the WFCAM & VISTA Science Archive

Nicholas Cross, Nigel Hambly, Ross Collins, Johann Bryant, Bob Mann,
Mike Read, Eckhard Sutorius and Peredur Williams
Wide Field Astronomy Unit (WFAU), Institute for Astronomy, University of Edinburgh

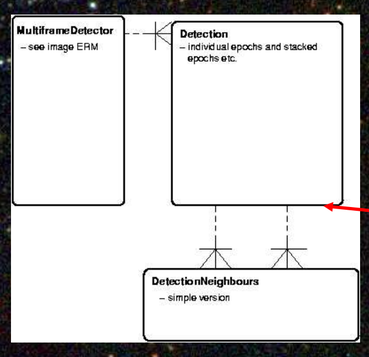
Abstract

The VISTA Data Flow System (VDFS) comprises the pipeline and archive system for infra-red data from WFCAM on UKIRT and the forthcoming VISTA telescope. These data include the UKIDSS surveys, the VISTA public surveys and all PI data taken on WFCAM. UKIDSS are the largest near-IR surveys to date and the VISTA Public Surveys will be the largest infrared surveys in the near-future.

The WFCAM/VISTA Science Archives (WSA/VSA) are designed by the Wide Field Astronomy Unit at the Institute for Astronomy, Edinburgh. Here we present the data model that we use for synoptic data in the WSA/VSA. We show how WSA/VSA data can be used with calibration data to calibrate the surveys and to select good candidates as future standard stars. We review a few different synoptic data sets and how best to use them. Finally we discuss future enhancements to the archive with respect to synoptic data.

Surveys and Synoptic Data

UKIRT and VISTA are the main near-IR survey instruments for the ESO community and will be the fastest near-IR survey instruments for the next decade, producing 10s-100s of Tb of data a year. While most of the surveys are designed are to produce scientific results from the deepest combination of data, there are some surveys designed to use time variable data. These include the the UKIRT Planetary Transit Survey (PTS:U/06a52) and the VISTA Public Survey Proposals 'VISTA Variables in Via Lactea' (VVV) and 'VISTA Magellanic Survey' (VMS). In addition all the calibration data can be processed in the same way.



Initial Database Design

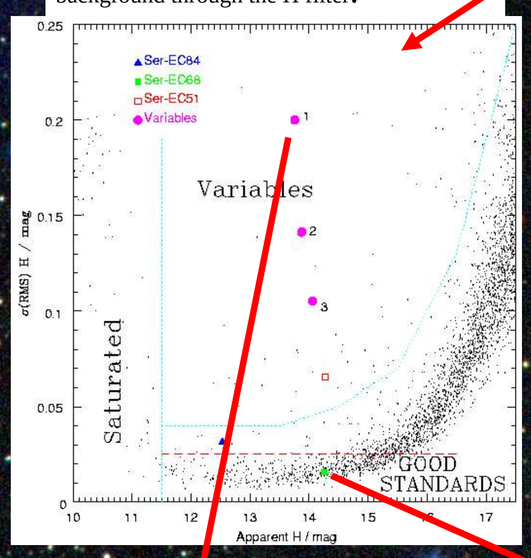
On the left is the **current** entity relationship model for the synoptic data. Each image frame has a set of image extensions, one for each detector in the camera (4 for WFCAM, 16 for VISTA).

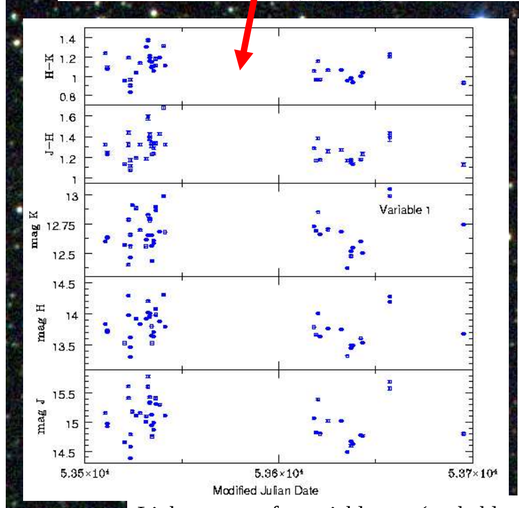
- Metadata for the each image extension are stored in the MultiframeDetector entity.
- Sources are extracted from each image extension and are stored in the Detection entity.
- The DetectionNeighbours entity stores pairs of matched with in a few arcseconds, usually the same object observed multiple times.
- This is very simple way to generate light curves and other plots to identify variable stars.

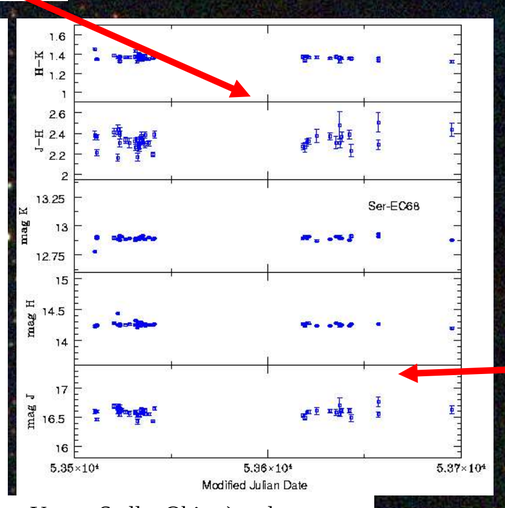
Finding standards and variables.

- Select all observations of each source from Detection and DetectionNeighbours
- Calibrate each observation compared to the others
- Measuring the rms of all observations of each object.
- Use RMS plot to classify objects- standards (like known ones marked), or variables, in magenta.

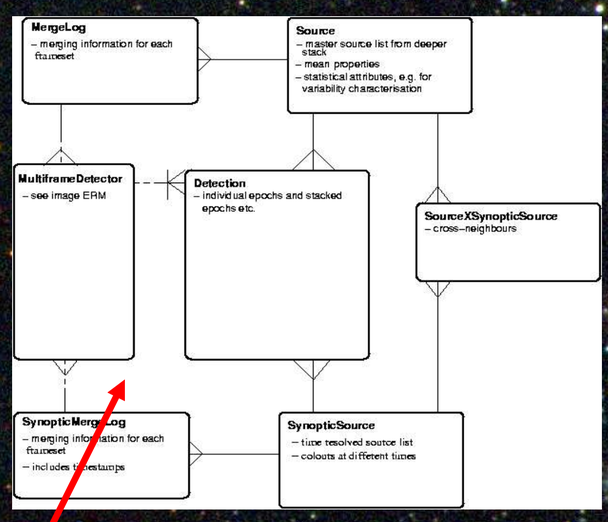
RMS vs magnitude for all stars in the background through the H filter.







Light curves of a variable star (probably a Young Stellar Object) and a UKIRT faint standard star, in J, H, K, (J-H) and (H-K).



New Relational Design

The synoptic data curation procedure and data model has been enhanced to improve the ease of access and to give more options to users. The pipeline will do the following:

- Stack the 5 best (seeing) images in each band to create a deep catalogue of all real sources
- Merge all band passes to create a **master source entity** to select from.
- **Recalibrate** each observation to get good **differential photometry**. Improves by ~0.005 mag.
- If multiple passbands at each visit create synopticSource entity and merge log. These contain **colour information at each visit** and the **mean time** of each visit.
- Create cross-neighbour table between Source and synoptic Source (or Detections if one pass only)
- Calculate **variability statistics** for each object in Source for users to search on.

Advantages:

- this method gives a **single unique source list** which can be used to **calculate statistics**
- the X-neighbour table is much smaller (upto 1000x in the case of the calibration data) and
- this **significantly speeds up selection of data**
- **colours are automatically generated at each visit**, which helps remove bad data.

The light curves on the left were generated from data in the current WFCAM database using the original database design. This required both selection in SQL and much post processing to associate passbands and recalibrate the data. With the new database design, all users could reproduce the same plots with simple SQL queries. **WSA: http://surveys.roe.ac.uk/wsa.**

The background image shows the Serpens EC cloud core, one of the calibration fields observed by WFCAM. The image is a composite of deep stacks (see new database design) in Z, J and K.. The stacks were produced using a tool from the Cambridge Astronomy Survey Unit. The 3 colour image was produced using SWARP and STIFF (Terapix, IAP). Any queries, email me: **njc@roe.ac.uk**