Web-based pipeline and survey progress monitoring

Marco Riello & Eduardo Gonzalez-Solares

September 29, 2006

1 WFCAM Data Quality Control Database

Every processed WFCAM data frame has a series of quality control (QC) measures automatically determined. These are designed to monitor the integrity of the instrument and also to provide measures of that greatly enhancing the usefulness of the data products for the users.

Over 30 Tbytes of data from 3 semesters of observations have been pipeline processed so far. From this figure alone it is clear that to make the best use of the QC measures a database is needed. The WFCAM QC database makes use of PostgreSQL and stores all the data quality information for every data product and input science frame. In addition several other relevant pieces of information are extracted from the FITS headers and ingested in the database: complete WCS calibration, weather conditions, temperature measures, etc. Such complementary information have indeed proven to be very useful to investigate systematic effects in data quality like, e.g. the dependence of sky brightness with temperature, humidity, etc. The first 3 semesters of data include over 329 nights and are made up of more than 600000 raw files, 632000 processed fits files, i.e. ≈ 2.5 million individual images (including the object catalogues, confidence maps and master calibrations) and populate several tables in the QC database for a total of more that 1100000 rows, not including object catalogues which include more than 765 million detected objects.

The ingestion of a pipeline-processed night into the QC database has proven to be very effective in spotting a number of problems that may occur during the reduction. The time required to ingest a night is proportional to the number of images in the night and is on average around 30-45 minutes for a typical night. Moreover the database is internally organised in such a way that when a new night in ingested the corresponding information is not immediately visible by the survey progress tool (see next section): this makes administration much easier by avoiding possible problems happening at ingestion, propagating all the way up to the end user looking at the web pages.

The QC database has a web-based front-end that is used internally by CASU members to rapidly search for images at a specific position on the sky and, optionally, satisying user-specified constraints for any QC parameter. The front-end is also complemented with an image cut-out service that is used to create on-the-fly object postage stamps and full-chip previews.

The QC database is used also to keep track of the night status within the processing data flow, to flag missing files when the tapes from JACH are ingested and the MEF files created, to record every single file that was transferred to ESO to be ingested in their archive ($\sim$ 400000 fits files have been transferred to date). Such information is used to generate user-friendly publicly-accessible web pages[1] showing for each night:

- the date of raw data ingestion at CASU (creation and validation of the MEF fits files).

- the current status of the night (*ingested*, *processing*, *processed* or *nodata* if the night does not contain any science data).

- the number of raw images (including calibration frames).

---

[1] Available at: `http://casu.ast.cam.ac.uk/surveys-projects/wfcam/data-processing/`.

« Back to WFCAM start page.

Night	Status	N_raw	N_ESO	Checked	Transferred by WFAU	Version	Summary Plots	Photometry Plots	Summary Info	Observation Log	Size raw [Gb]	Size red [Gb]	N stacks All	N stacks UKIDSS	N stacks LAS	N stacks DXS	N stacks UDS	N stacks GPS	N stacks GCS
2006/04/28	reduced	476	361	23 Aug 2006	06 Sep 2006	1	GIF1 GIF2	GIF	summary.list	obs_index	8.86	16.13	39	30	0	0	0	30	0
2006/04/29	reduced	411	366	23 Aug 2006	06 Sep 2006	1	GIF1 GIF2	GIF	summary.list	obs_index	8.60	2.58	22	0	0	0	0	0	0
2006/04/30	reduced	145	133	23 Aug 2006	06 Sep 2006	1	GIF1 GIF2	GIF	summary.list	obs_index	2.01	1.68	10	0	0	0	0	0	0
2006/05/01	reduced	1055	794	23 Aug 2006	06 Sep 2006	1	GIF1 GIF2	GIF	summary.list	obs_index	20.46	35.09	126	69	42	0	0	27	0
2006/05/02	reduced	1290	139	23 Aug 2006	06 Sep 2006	1	GIF1 GIF2	GIF	summary.list	obs_index	24.99	50.84	47	0	0	0	0	0	0
2006/05/03	reduced	2402	640	23 Aug 2006	06 Sep 2006	1	GIF1 GIF2	GIF	summary.list	obs_index	32.26	47.46	44	0	0	0	0	0	0
2006/05/04	reduced	1564	182	23 Aug 2006	06 Sep 2006	1	GIF1 GIF2	GIF	summary.list	obs_index	29.42	60.45	60	0	0	0	0	0	0
2006/05/05	reduced	1639	116	23 Aug 2006	06 Sep 2006	1	GIF1 GIF2	GIF	summary.list	obs_index	25.64	44.36	37	0	0	0	0	0	0
2006/05/06	nodata	66	55							obs_index	0.69								
2006/05/07	nodata	0																	
2006/05/08	nodata	554								obs_index	5.57								
2006/05/09	nodata	768	84							obs_index	7.09								
2006/05/10	reduced	2516	1918	25 Aug 2006	27 Sep 2006	1	GIF1 GIF2	GIF	summary.list	obs_index	42.30	76.66	257	226	0	0	0	0	226
2006/05/11	reduced	2603	2553	25 Aug 2006	23 Sep 2006	1	GIF1 GIF2	GIF	summary.list	obs_index	49.60	128.29	357	330	2	0	0	107	221
2006/05/12	reduced	1841	1823	25 Aug 2006	27 Sep 2006	1	GIF1 GIF2	GIF	summary.list	obs_index	35.96	57.71	229	178	89	0	0	52	37
2006/05/13	reduced	1129	563	19 Sep 2006	23 Sep 2006	1	GIF1 GIF2	GIF	summary.list	obs_index	18.18	60.40	60	0	0	0	0	0	0
2006/05/14	reduced	2402	1424	19 Sep 2006	23 Sep 2006	1	GIF1 GIF2	GIF	summary.list	obs_index	46.56	95.81	111	30	1	17	0	12	0
2006/05/15	reduced	2674	1540	19 Sep 2006	23 Sep 2006	1	GIF1 GIF2	GIF	summary.list	obs_index	53.59	105.73	103	20	0	20	0	0	0
2006/05/16	reduced	148	145	19 Sep 2006	23 Sep 2006	1	GIF1 GIF2	GIF	summary.list	obs_index	1.85	1.93	15	0	0	0	0	0	0

Figure 1: A snapshot of the reduction progress page for 06A. See the text for a description of the columns.
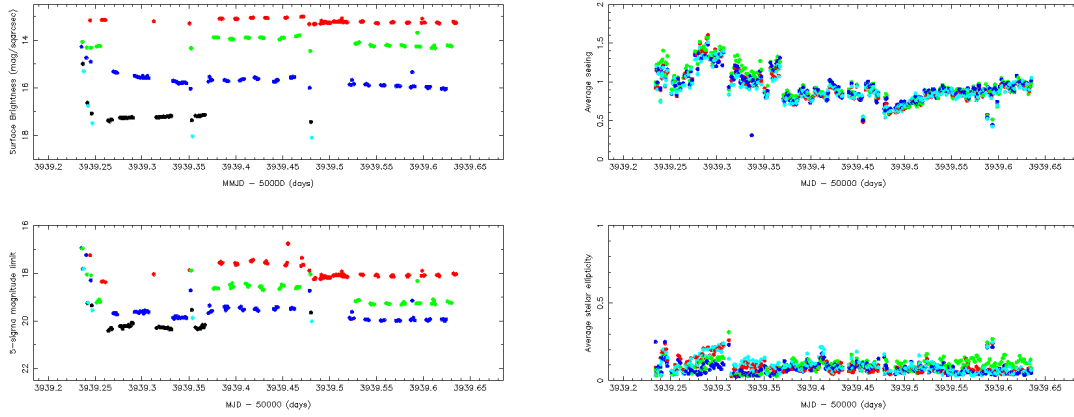

Figure 2: An example of the QC plots showing the average sky brightness, magnitude limit, seeing (FWHM) and stellar object ellipticity for each stacked frame of a night.

- the date when the data was flagged by CASU as ready to be transferred by WFAU (after being processed).

- the date when the data was actually transferred by WFAU.

- three QC summary plots showing, for each stacked frame, the photometric zero points, average seeing (FWHM), average stellar ellipticity, the sky brightness and the $5 - \sigma$ magnitude limit deduce from the sky noise.

- the total size in Gbytes of the raw and reduced data directories (both raw and reduced images are stored in compressed format).

- once the night has been processed and fully ingested in the QC database, it is possible to show also: the total number of stacked images, the total number of stacked images for the UKIDSS surveys altogether and by survey.

To make the pages easier to use, the nights are grouped by semester. The web pages fetch the relevant information from the QC database, which is kept up to date by a script running every hour.
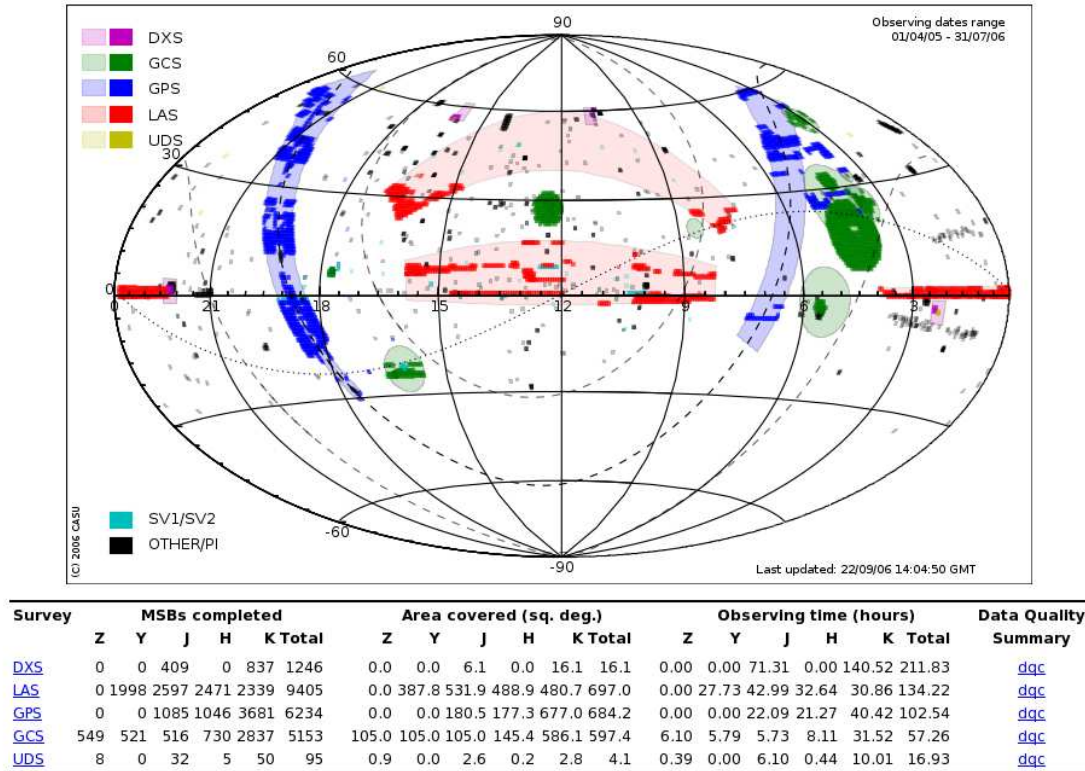
2

Figure 3: All sky Aitoff projection showing the areas defined in the UKIDSS survey (light colours) and the areas actually observed and processed (dark colours).

Survey	MSBs completed						Area covered (sq. deg.)						Observing time (hours)						Data Quality
	Z	Y	J	H	K	Total	Z	Y	J	H	K	Total	Z	Y	J	H	K	Total	Summary
DXS	0	0	409	0	837	1246	0.0	0.0	6.1	0.0	16.1	16.1	0.00	0.00	71.31	0.00	140.52	211.83	dqc
LAS	0	1998	2597	2471	2339	9405	0.0	387.8	531.9	488.9	480.7	697.0	0.00	27.73	42.99	32.64	30.86	134.22	dqc
GPS	0	0	1085	1046	3681	6234	0.0	0.0	180.5	177.3	677.0	684.2	0.00	0.00	22.09	21.27	40.42	102.54	dqc
GCS	549	521	516	730	2837	5153	105.0	105.0	105.0	145.4	586.1	597.4	6.10	5.79	5.73	8.11	31.52	57.26	dqc
UDS	8	0	32	5	50	95	0.9	0.0	2.6	0.2	2.8	4.1	0.39	0.00	6.10	0.44	10.01	16.93	dqc

2 Survey Progress Tool

The purpose of the WFCAM Survey Progress Tool is to monitor the progress of the observations carried out by WFCAM and in particular of the UKIDSS surveys. This tool uses the QC database described in the previous section.

The entry page[2] provides a whole sky overview of the planned UKIDSS observations and of the observations actually carried out, with relevant statistics like the number of MSBs completed, area covered and observing time (see figure 3).

This page provides links to more detailed views for each UKIDSS survey (figure 4). This is composed basically of zoomed areas displaying only the information of that particular survey. The user can select several to constrains on the frames to display:

- Bands of the frames to display and how to combine mutliple observations. The user can then select to display e.g. only those fields observed in the three J, H and K bands and observed more than once in each of them.

- Display frames with seeing within a range.

- Display frames observed within a specific date range.
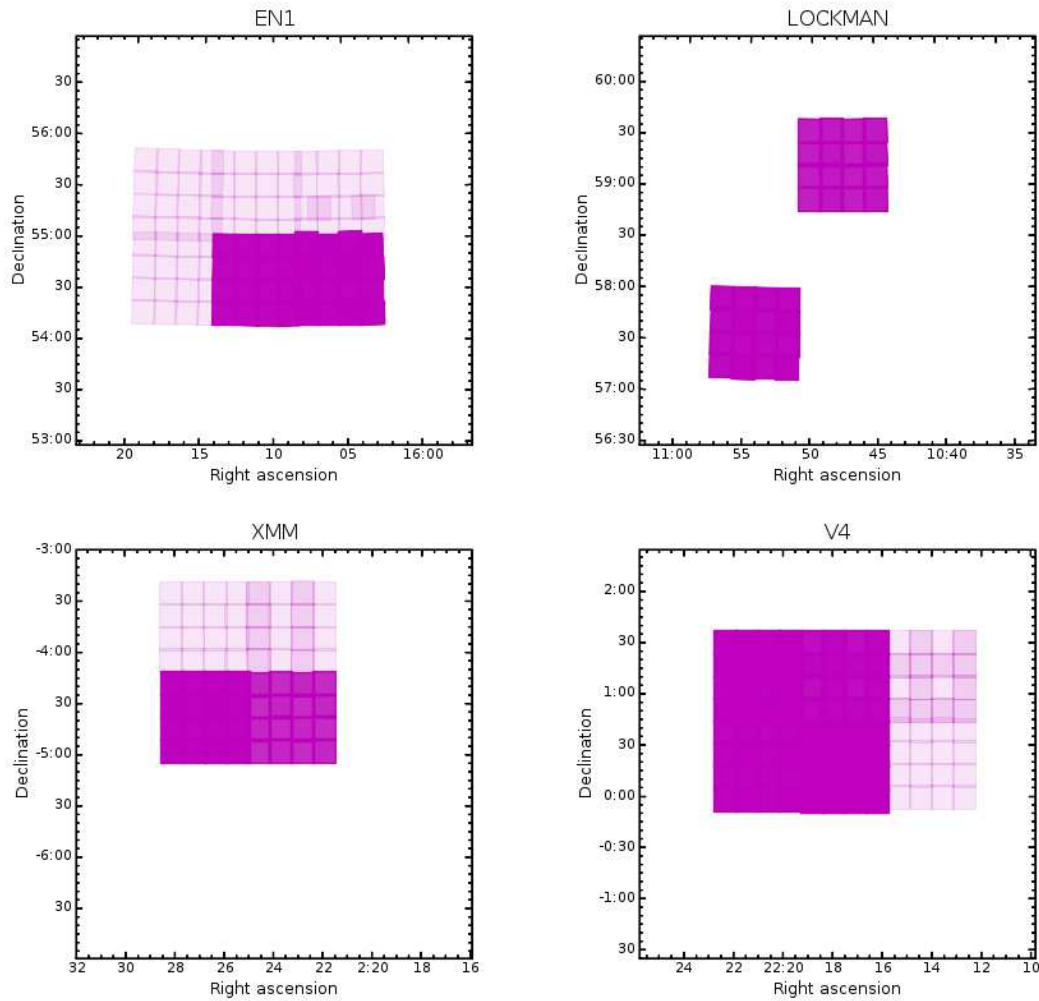
or a combination of the above.

Also linked from the main page the DQC view provides access to the distribution and evolution with time of different quality control parameters: seeing, ellipticity, magnitude zero point, magnitude limit and sky brightness (see figure 5) for a example. As before the user has the option to display a range of dates, zoom in the plots or specify additional constraints using SQL language.

---

[2]Available at `http://casu.ast.cam.ac.uk/survey-progress/wfcam/`.

Figure 4: Survey progress for the DXS showing the area completed in each of the four survey areas plus number of MSBs completed in each filter, exposure time. The user is able to interact with the information presented so he can choose e.g. to display only those areas covered multiple times in J and K or/and within a specific seeing range.
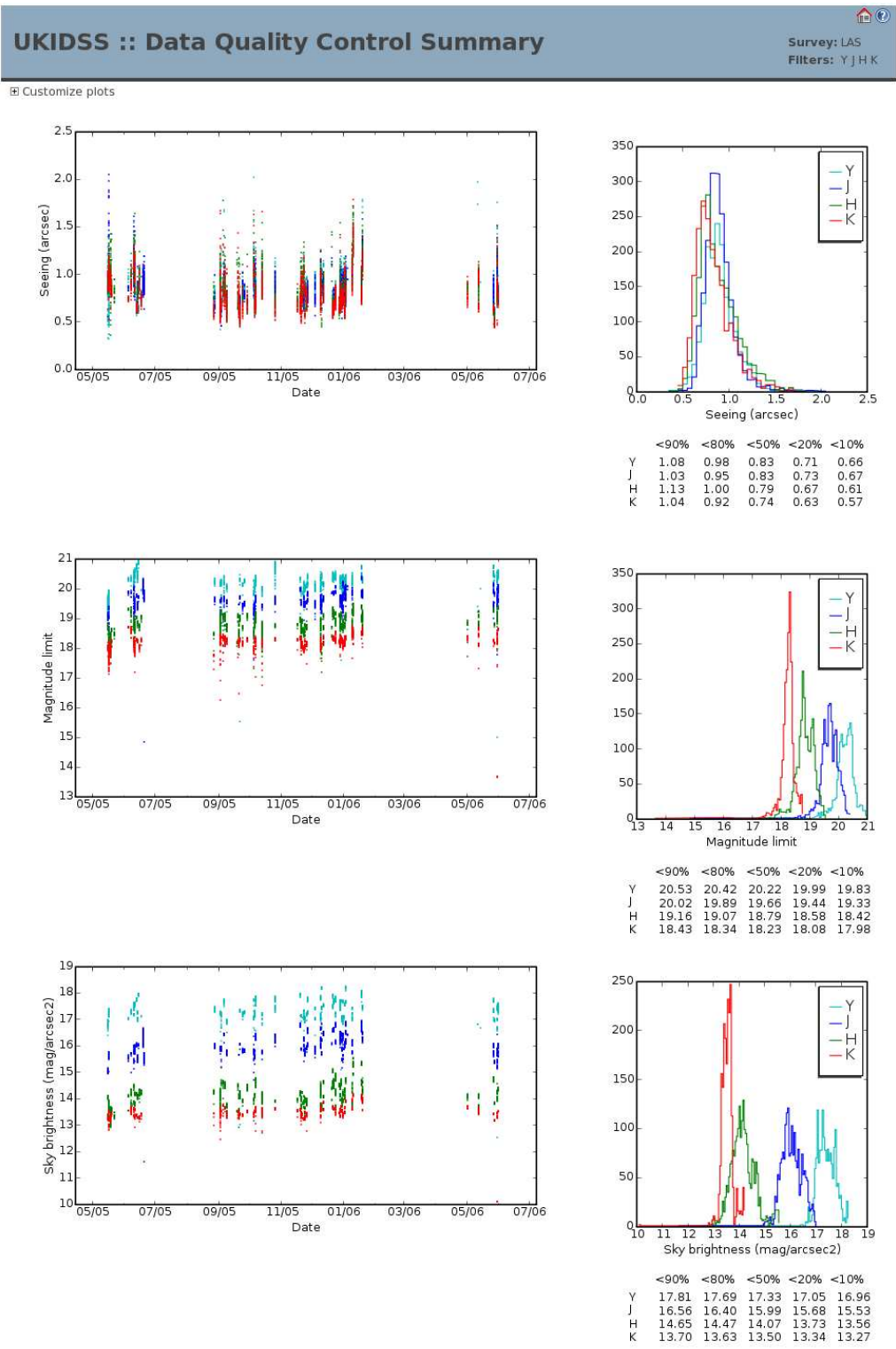
Figure 5: DQC plots for the LAS survey showing the values of seeing and sky brightness showing e.g. that 90% of the frames obtained in K band have seeing better than 0.94 arcsec.