Initial data release from the INT Photometric Hα Survey of the Northern Galactic Plane (IPHAS)

E. A. González-Solares,[1]⋆ N. A. Walton,[1] R. Greimel,[2,3] J. E. Drew,[4,5] M. J. Irwin,[1] S. E. Sale,[4] K. Andrews,[6] A. Aungwerojwit,[7,8] M. J. Barlow,[9] E. van den Besselaar,[10] R. L. M. Corradi,[2,11] B. T. Gänsicke,[7] P. J. Groot,[10] A. S. Hales,[9,12] E. C. Hopewell,[4] H. Hu,[10] J. Irwin,[1,13] C. Knigge,[14] E. Lagadec,[15] P. Leisy,[2,11] J. R. Lewis,[1] A. Mampaso,[11] M. Matsuura,[15] B. Moont,[4] L. Morales-Rueda,[10] R. A. H. Morris,[16] T. Naylor,[17] Q. A. Parker,[18,19] P. Prema,[1] S. Pyrzas,[7] G. T. Rixon,[1] P. Rodríguez-Gil,[7,11] G. Roelofs,[10,13] L. Sabin,[15] I. Skillen,[2] J. Suso,[20] R. Tata,[21] K. Viironen,[11] J. S. Vink,[22] A. Witham,[14] N. J. Wright,[9] A. A. Zijlstra,[15] A. Zurita,[23] J. Drake,[13] J. Fabregat,[20] D. J. Lennon,[2,11] P. W. Lucas,[5] E. L. Martín,[11,21] S. Phillipps,[16] D. Steeghs[7,13] and Y. C. Unruh[4]

[1]*Institute of Astronomy, Madingley Road, Cambridge CB3 0HA*
[2]*Isaac Newton Group of Telescopes, Apartado de correos 321, E38700 Santa Cruz de La Palma, Tenerife, Spain*
[3]*Institut für Physik, Karl-Franzen Universität Graz, Universitätsplatz 5, 8010 Graz, Austria*
[4]*Imperial College London, Blackett Laboratory, Exhibition Road, London SW7 2AZ*
[5]*Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield AL10 9AB*
[6]*Royal Observatory Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*
[7]*Department of Physics, University of Warwick, Coventry CV4 7AL*
[8]*Department of Physics, Faculty of Science, Naresuan University, Phitsanulok 65000, Thailand*
[9]*University College London, Department of Physics & Astronomy, Gower Street, London WC1E 6BT*
[10]*Department of Astrophysics/IMAPP, Radboud University Nijmegen, PO Box 9010, 6500 GL, Nijmegen, the Netherlands*
[11]*Instituto de Astrofísica de Canarias, Via Láctea s/n, E38200 La Laguna, Santa Cruz de Tenerife, Spain*
[12]*National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, Virginia 22903-2475, USA*
[13]*Harvard–Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*
[14]*School of Physics & Astronomy, University of Southampton, Southampton SO17 1BJ*
[15]*School of Physics and Astronomy, University of Manchester, Sackville Street, PO Box 88, Manchester M60 1QD*
[16]*Astrophysics Group, Department of Physics, Bristol University, Tyndall Avenue, Bristol BS8 1TL*
[17]*School of Physics, University of Exeter, Stocker Road, Exeter EX4 4QL*
[18]*Department of Physics, Macquarie University, NSW 2109, Australia*
[19]*Anglo-Australian Observatory, PO Box 296, Epping, NSW 1710, Australia*
[20]*Observatorio Astronómico de Valencia, Universidad de Valencia, 46071 Paterna-Valencia, Spain*
[21]*Physics Department, University of Central Florida, Orlando, FL 32816, USA*
[22]*Armagh Observatory, College Hill, Armagh BT61 9DG*
[23]*Departamento de Física Teorica y del Cosmos, Facultad de Ciencias, Av. Fuentenueva s/n, E18071 Granada, Spain*

ABSTRACT

The INT/WFC Photometric Hα Survey of the Northern Galactic Plane (IPHAS) is an imaging survey being carried out in Hα, r' and i' filters, with the Wide Field Camera (WFC) on the 2.5-m Isaac Newton Telescope (INT) to a depth of $r' = 20$ (10σ). The survey is aimed at revealing the large scale organization of the Milky Way and can be applied to identifying a range of stellar populations within it. Mapping emission line objects enables a particular focus on objects in the young and old stages of stellar evolution ranging from early T-Tauri stars to late planetary nebulae. In this paper we present the IPHAS Initial Data Release, primarily a photometric catalogue of about 200 million unique objects, coupled with associated image data covering about 1600 deg^2 in three passbands. We note how access to the primary data products has

⋆E-mail: eglez@ast.cam.ac.uk

been implemented through use of standard virtual observatory publishing interfaces. Simple traditional web access is provided to the main IPHAS photometric catalogue, in addition to a number of common catalogues (such as 2MASS) which are of immediate relevance. Access through the AstroGrid VO Desktop opens up the full range of analysis options, and allows full integration with the wider range of data and services available through the Virtual Observatory. The IDR represents the largest data set published primarily through VO interfaces to date, and so stands as an exemplar of the future of survey data mining. Examples of data access are given, including a cross-matching of IPHAS photometry with sources in the UKIDSS Galactic Plane Survey that validates the existing calibration of the best data.

Key words: catalogues – surveys – stars: emission line, Be – Galaxy: stellar content.

1 INTRODUCTION

The Isaac Newton Telescope (INT) Photometric Hα Survey of the Northern Galactic Plane (IPHAS) is a 1800 deg^2 CCD survey of the northern Milky Way ($|b| < 5°$) using the r', i' and Hα passbands (Table 1) down to $r' = 20$ (Vega, 10σ for a point-like source in an aperture of 1.2 arcsec). The execution of IPHAS, the properties of its characteristic colour–colour plane, and the likely scientific scope of the survey were discussed by Drew et al. (2005)

This paper describes the Initial Data Release (IDR) of the IPHAS survey, containing observations for about 200 million objects and comprising 2.4 Tb of processed imaging data. The survey is being carried out using the Wide Field Camera (WFC) on the INT telescope. The WFC consists of four 2048 × 4096 pixel CCDs with a pixel scale of 0.33 arcsec and a field of view of 0.29 deg^2. The camera CCD detectors are configured in an L shape and CCD number 3 is slightly vignetted in one corner.

The data have already been used to carry out a series of investigations: the interaction between the planetary nebula Sh2−188 and the interstellar medium (Wareing et al. 2006); the discovery of a new quadrupolar nebula (Mampaso et al. 2006); the study of the properties of a sample of cataclysmic variables (Witham et al. 2006); the interaction of the jet of Cygnus X-1 with the surrounding medium (Russell et al. 2007); brown dwarf searches (Martín et al. 2008); detection of nova progenitors (Steeghs et al. 2006, 2007); the study of the properties of symbiotic stars and first discoveries (Corradi et al. 2008), the study of young pre-main-sequence stars in CygOB2 (Vink et al. 2008), the presentation of a sample of about 5000 Hα emitting stars (Witham et al. 2008) and the description of a method for selecting A stars and exploiting them to estimate distances and reddening (Drew et al. 2008). Several other investigations are ongoing and we hope the data release described here will lead to many data mining programs by the community.

In addition to the optical data provided by IPHAS, the Galactic Plane Survey (GPS; Lucas et al. 2007) of the UKIDSS (Dye et al.

2006) is also observing a similar area in *J*, *H* and *K*. This will be followed in the next few years by a survey of the southern Galactic plane, VPHAS+, on the VLT Survey Telescope (VST) using OmegaCam (Kuijken et al. 2004). It serves as a digital successor to the UK Schmidt Hα Survey (Parker et al. 2005) and will reach about 2 mag deeper. The UVEX project is targeting the same area in the *U* and *g* bands. The *Herschel* observatory (Pilbratt 2005) will also target the Galactic plane. Together, these surveys will provide a better understanding of the stellar content and properties of our Galaxy.

The aim of this paper is to provide a guide to the IDR of the IPHAS survey, describe the data products and the data properties and usage. Section 2 summarizes the observations, sky coverage and data quality. The data processing is briefly described in Section 3. The data products contained in this release are described in Section 4. Section 5 shows an example usage of the catalogue to extract sources detected in a large area of sky and cross-match with the GPS survey. Finally in Section 6 we describe the different user access interfaces to the data.

2 SKY COVERAGE AND DATA QUALITY

The IDR contains 1600 deg^2 of imaging data in three bands covering the northern Galactic plane at latitudes $|b| < 5°$ and longitudes $30° \lesssim \ell \lesssim 220°$. Fig. 1 shows the distribution in the sky. The data have been obtained between 2003 August and 2005 December during a total of 212 nights (about 26 per cent of this observing time has been completely lost due to bad weather).

The area to be observed has been divided into 'fields' each of them corresponding to the coverage of the WFC. Fields have been defined in a diagonal rather than rectangular geometry in order to take advantage of the L shape of the detector configuration allowing for a 5 per cent overlap between adjacent pointings. Fig. 2 provides an example of the observation configuration.

Gaps between detectors are typically 20 arcsec. Therefore the area not observed in each pointing due to the gaps is about 12 arcmin^2 (i.e. ∼1 per cent). In order to cover the chip gaps and bad columns each 'field' is observed in pairs which are referred to as 'on' and 'off' positions with an offset of 5 arcmin west and 5 arcmin south between the pointings. Most of the objects are thus observed twice. Field pairs are observed consecutively in the three bands before moving to the next field pair in order to minimize the effects of photometric variability. Fields with mean stellar seeing values larger than 2 arcsec, ellipticity larger than 0.2, r'-band sky counts larger than 2000 ADUs (which indicates observations too close to the moon or with clouds such that sensitivity is reduced to a significantly lower magnitude limit) or anomalous star ratios between the filters

Table 1. Filter parameters. The second column lists the effective wavelength of the filter and the third its width. The last column lists the value to be added to the Vega magnitude to convert it to AB. Filter profiles are available from the ING web site (www.ing.iac.es) and discussed in Drew et al. (2005).

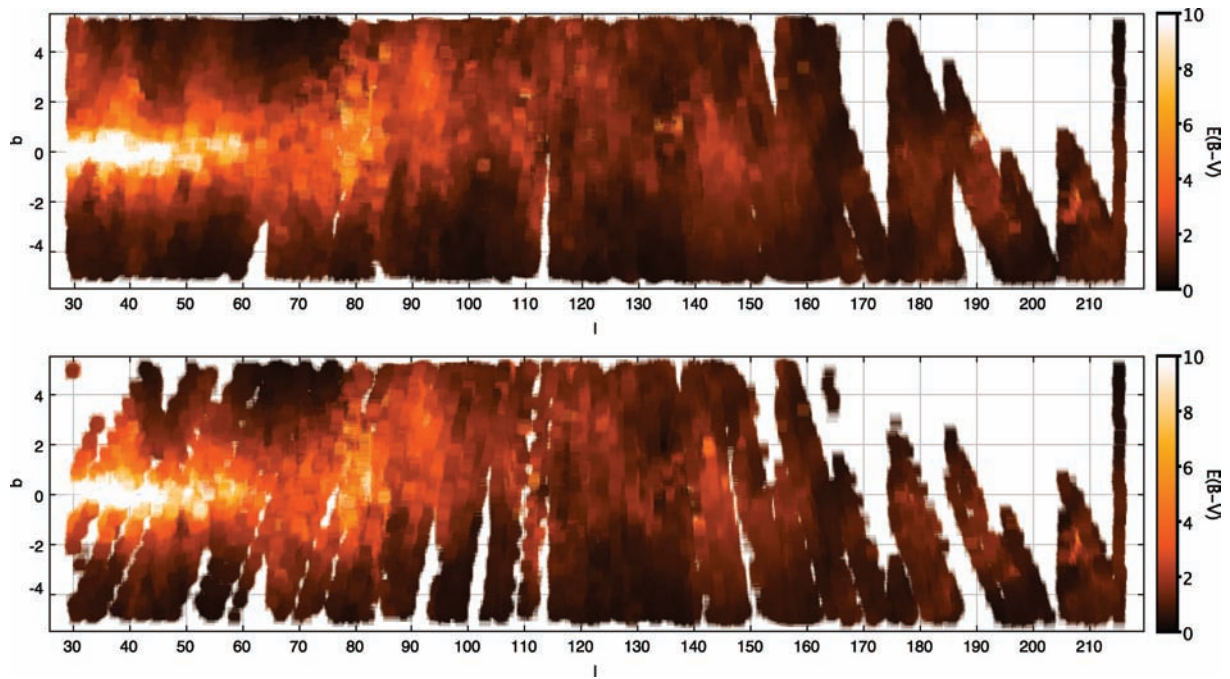| Filter	Cen wave (Å)	Width (Å)	AB (mag)
r'	6254.1	1350.0	0.164
Hα	6568.3	96.0	0.352
i'	7771.9	1490.0	0.413

Figure 1. IPHAS IDR sky coverage in Galactic coordinates (degrees). The top figure displays the coverage of the full release and the bottom one the coverage of the PhotoObjBest best observations subset. Brightness indicates the total line-of-sight reddening as computed from the Schlegel et al. (1998) dust maps using 4 pixel interpolation.
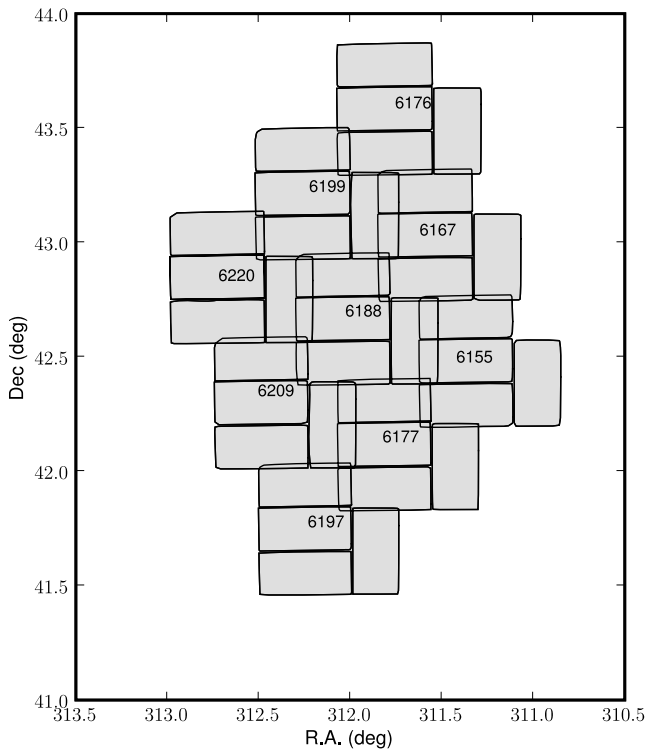

Figure 2. Tiling configuration of 3 × 3 pointings covering an area of ∼2 deg^2. Chips are outlined in black. Field numbers are also shown. An additional pattern with an offset of 5 × 5 arcmin^2 is not included in order to make the figure clearer.
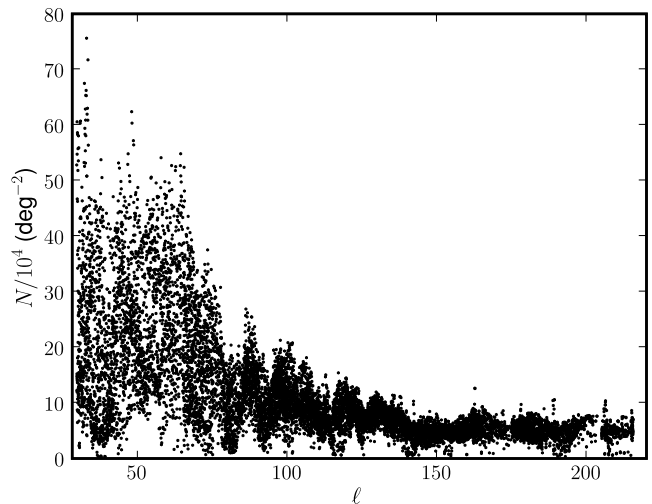

Figure 3. Source counts as a function of Galactic longitude (one data point per IPHAS field).

(which indicates variable clouds during observations) are scheduled for re-observation in the three filters. Each field pair is assigned an identification which we will refer to as 'field number'. The average stellar counts is ∼20 000 deg^{-2} although this value is highly dependent on the position along the Galactic plane as seen in Fig. 3 which shows the number of sources detected in each field as a function of Galactic longitude.

The overall survey properties and quality of the data are summarized in Table 2. The exposure time for each filter is 30 s in r' (except in the first year of observations, 2003, when it was 10 s), 10 s in i' and 120 s in Hα. All data obtained up to 2005 December have been included in the IDR regardless of seeing or photometric conditions.

Table 2. Survey parameters and IDR quality summary.

| Parameter	Value
Telescope	2.5-m Isaac Newton Telescope, La Palma
Instrument	Wide Field Camera
Field of view	0.29 deg^2 (per pointing)
Detector scale	0.33 arcsec pixel^{-1}
Total area of survey	~1800 deg^2
Total area in IDR	~1600 deg^2
Area included in the PhotoObjBest catalogue	~1000 deg^2
Number of unique objects in the full catalogue	~200 000 000
Number of unique objects in the PhotoObjBest catalogue	~90 000 000
Filters	r', i', Hα
Exposure time	30 s (r')†, 10 s (i'), 120 s (Hα)
FWHM (median)	1.3 arcsec (r'), 1.2 arcsec (i'), 1.3 arcsec (Hα)
Magnitude limit (Vega, 5σ)	21.8 (r'), 20.2 (i'), 20.6 (Hα)
Astrometry	<0.1 arcsec (based on 2MASS).

† 10 s in 2003.
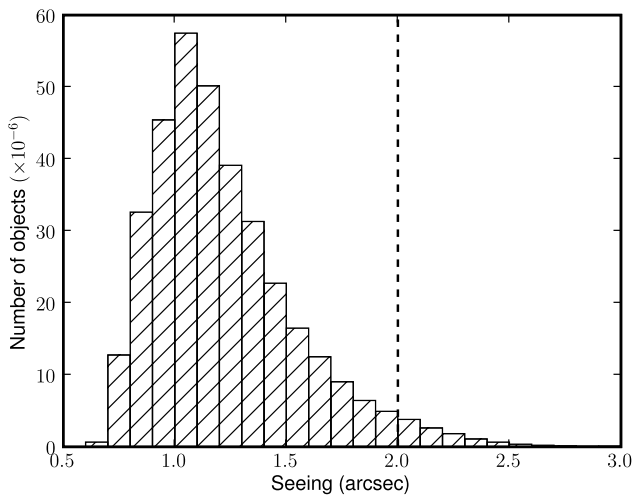
Figure 4. Seeing distribution in the *r'* band for the whole IDR. One of the criteria for inclusion in the PhotoObjBest catalogue is that objects have to be observed with seeing conditions better than 2.0 in the three bands (left-hand side of the dashed line).

Figure 5. Magnitude limit distribution for the IDR in the three survey bands. The *r'*-band distribution has been split in two, the leftmost corresponding to the 2003 observations and the rightmost corresponding to post-2003 observations in which the exposure time was increased from 10 to 30 s.

Fig. 4 shows the seeing values for these data in the *r'* band. Although the mean value is about 1.26 arcsec (and the median 1.18 arcsec) there are clearly observations obtained in poorer conditions. The maximum acceptable seeing value for the data to be accepted as final survey quality is 2 arcsec in the three bands. Data with larger seeing values have been included in this release but they do not appear by default to the user. For this purpose we are releasing a full catalogue, PhotoObj, with all the observations obtained and a PhotoObjBest subset which only contains the observations which pass our scientific criteria for this release (see Section 4.2 for more details). As noted in Table 2 the coverage of this latter catalogue is about 1000 deg^2 and contains about 90 million unique objects. The Final Data Release of the data will extend the defined scientific criteria to the whole 1800 deg^2. Fig. 5 shows the magnitude limit (5σ) for all the objects in the three survey bands. Note that n*r'* observations obtained in 2003 are shallower due to the smaller integration time used.
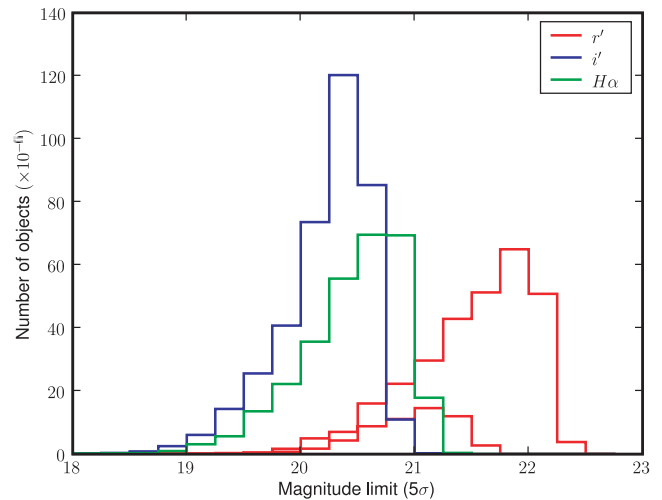
3 DATA PROCESSING

All the observations obtained with the INT telescope at La Palma are transmitted in pseudo-real time to the Cambridge Astronomy Survey Unit (CASU) for data archival via the internet. Although quasi-real-time data processing is possible it is preferable to perform the data processing on a run by run basis. The calibration is then performed using the whole run master flats and biases. For each observation the data are stored in a multi-extension FITS file (MEF) with a primary header describing the overall characteristics of the observation (pointing, filter, exposure time, etc.) and four image extensions corresponding to each of the CCD detectors.

The data are processed by CASU as described in Irwin & Lewis (2001) following the same steps as for the processing of the Wide Field Survey (McMahon et al. 2001; Irwin et al. 2005) which used the same telescope and camera combination. We provide here a short description of the reduction steps; additional details are available from Drew et al. (2005).

The data are first debiased (full 2D bias removal is necessary). Bad pixels and columns are then flagged and recorded in confidence maps, which are used during catalogue generation. Linearity tests using sequences of dome flats revealed that the CCDs have significant non-linearities at the level of 1–2 per cent depending on the detector and period of observation. Due to several changes to the WFC controllers the linearity values change from time to time so linearity measurements are updated periodically. A linearity correction using look-up tables is then applied to all data. Flat-field images in each band are constructed by combining a series of sky flats obtained in bright sky conditions during the twilight for each run.

The standard catalogue generation software (Irwin 1985) makes direct use of the confidence maps previously generated (described in Section 4.1.1) for a variety of tasks. These include object detection and parametrization, producing quality control information, standard object descriptors and detected object overlay files. The possibly varying sky background is estimated automatically, prior to object detection, using a combination of robust iteratively clipped estimators. The image catalogues are then further processed to yield morphological classification for detected objects and used to generate astrometric and photometric calibration information.

Each detected object has an attached set of descriptors, forming the columns of the binary table and summarizing derived position, shape and intensity information. During further processing stages ancillary information such as the sky properties, seeing and mean stellar image ellipticity, are derived from the catalogues and stored in the FITS headers attached to each catalogue extension. In addition to being the primary astronomical products from the pipeline processing, the catalogues and associated derived summary information form the basis for astrometric and photometric calibration and quality control monitoring.

Finally an astrometric solution starts with a rough World Coordinate System (WCS) based on the known telescope and camera geometry and is then progressively refined using the Two Micron All Sky Survey (2MASS) catalogue. The WFC field distortion is modelled using a zenithal equidistant projection with polynomial radial distortion (ZPN; Greisen & Calabretta 2002). The resulting internal astrometric precision is better than 100 mas over the whole WFC array (based on intercomparison of overlap regions; see Section 3.3). The object detection is performed in each band separately using a standard APM-style object detection and parametrization algorithm (Irwin 1985). The curve of growth, defined as the flux inside an aperture as a function of its radius, is calculated measuring the flux in a set of apertures of radius $r/2, r, \sqrt{2}r, 2r, 2\sqrt{2}r$ where $r = 3.5$ pixel (1.2 arcsec) and an automatic aperture correction (based on the average curve of growth for stellar images) is applied to all detected objects. Finally, a distortion correction is applied to the photometry to take into account the change of scale with respect to the optical axis distance (see Section 4.1).

It is straightforward to show that an aperture of radius approximately equal to the full width at half-maximum (FWHM) delivers optimal (i.e. 90 per cent of the total flux) photometry compared to more detailed point spread function (PSF) fitting (Naylor 1998) in non-crowded areas. The default photometry is then done in an aperture of 1.2 arcsec and corrected for light loses outside the aperture. After adding some extra refinements to deal with the most crowded regions we find that there is little to be gained from PSF fitting (Irwin et al., in preparation). The difference between crowded and non-crowded areas is shown in Fig. 6 which represents the nearest neighbour source distribution for two different fields: one with high source density to illustrate the worst case scenario of a crowded
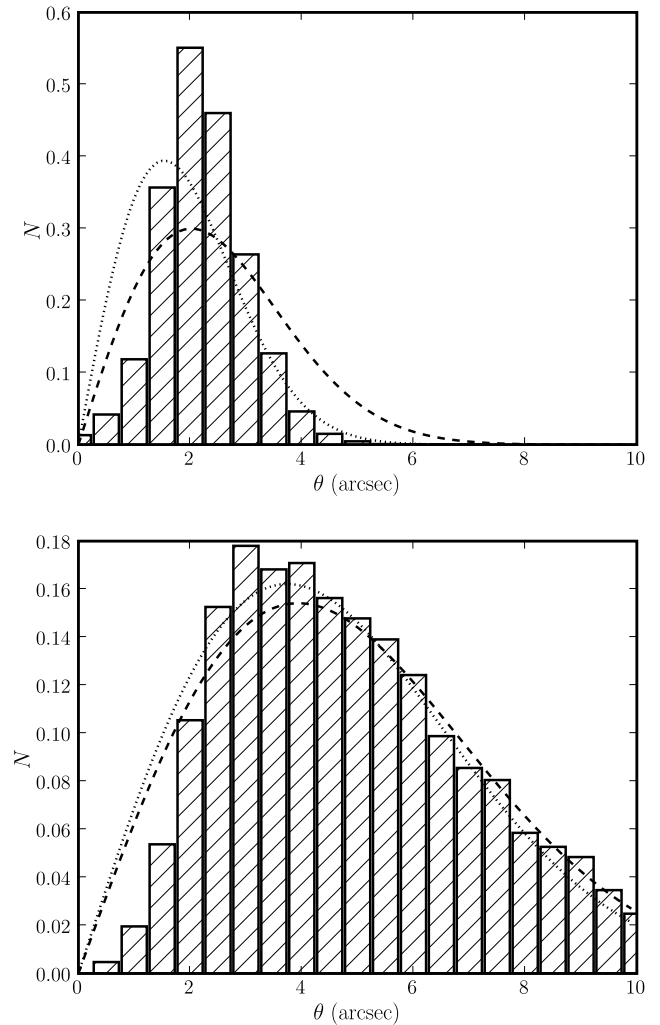

Figure 6. Normalized nearest neighbour source distribution for a crowded field (top; 140 sources per arcmin^2, 146 160 sources in total) and for an average source density field (bottom; 40 sources arcmin^{-2}; 40 760 sources in total) both observed with seeing of 1.2 arcsec. The dashed line shows the distribution for a random sample (equation 1) using the observed source density while the dotted line shows the expectation when the average source density is corrected for incompleteness. Close pairs are missing due to the seeing and this is reflected in the smaller number of pairs with respect to the Poisson expectation at seeing smaller than 2 arcsec. This effect is larger in the crowded field leading to a large difference between the observed and expected normalized distributions.

field and another with an average source density. For comparison we use the nearest neighbour probability density distribution (PDF) expected for a random sample which can be written as (Bahcall & Soneira 1981)

$$P_{NN}(\theta) = 2\pi\rho\theta\, e^{-\pi\rho\theta^2}, \tag{1}$$

where ρ is the average density and $\pi\rho\theta^2$ is the number of neighbours in the interval $(0, \pi\theta^2)$. The expected mean nearest neighbour separation can then be calculated as (Scott & Tout 1989)

$$\langle\theta\rangle = \int_0^\infty \theta\, P_{NN}(\theta)\,\mathrm{d}\theta = \frac{1}{2\sqrt{\rho}}. \tag{2}$$

For the space densities of the fields plotted in Fig. 6 we obtain $\langle\theta\rangle = 2.5$ arcsec for the crowded field and $\langle\theta\rangle = 4.7$ arcsec for the average density field. The maximum of the P_{NN} distribution is given

by $\theta_m = 1/\sqrt{2\pi\rho}$ and corresponds to $\theta_m = 2.0$ arcsec and $\theta_m = 3.8$ arcsec in the crowded and average density fields, respectively. The large discrepancy between the random and observed distributions for the high density field at $\theta > 2$ arcsec is due to missing close pairs which produces an overestimation of the nearest neighbours at $\theta \simeq 2$ arcsec. We can calculate the fraction of missed objects due to crowding using the following expression (Irwin & Trimble 1984)

$$\rho_{\mathrm{corr}} = -\frac{\log(1 - 4\pi\rho\,\theta_{\mathrm{FWHM}}^2)}{4\pi\,\theta_{\mathrm{FWHM}}^2}, \qquad (3)$$

where ρ_{corr} is the corrected average density and θ_{FWHM} the FWHM of the images (i.e. 1.2 arcsec). For the average density field the fraction of sources lost due to close pairs is 9 per cent while in the crowded field this percentage increases to 41 per cent. Note as well that non-resolved objects pair, on average, with objects at a distance $\sim\langle\theta\rangle$ occupying, and overestimating, the central bins of the nearest neighbour distribution (Fig. 6, top). We also plot in Fig. 6 the corrected nearest neighbour PDF as a dotted line using the corrected average density given by equation (3).

Note that the crowded field case represents an extreme case which has source density more than three times higher than the median source density of the survey.

The values obtained for the mean nearest neighbour separations are clearly well above our seeing values in both cases indicating that aperture photometry is appropriate for our case.

3.1 Morphological classification

The morphological classification is based primarily on the aperture fluxes and the discrete curve of growth for each detected object defined by them. Ancillary information such as object ellipticity derived from intensity-weighted second moments is also used but only as a secondary indicator. The curve of growth is a direct measure of the integral of the PSF out to various radii and is independent of magnitude if the data are properly linearized, and if saturated images are excluded. In using this property the classifier further assumes that the effective PSF for stellar objects is constant over each detector,[1] although individual detectors are allowed to have different PSFs since all detectors, pointings and passbands are treated independently.

The average stellar locus on each detector in these parameter spaces is generally well defined and is used as the basis for a null hypothesis stellarness test. Fig. 7 shows an example of this comparing the magnitude difference recorded in two apertures (1.2 and 2.4 arcsec diameter) as a function of the derived overall magnitude and illustrates how this relates to the final classification for this frame.

In practice, the aperture with radius $r = 1.2$ arcsec is used as a fixed reference and also defines the internal magnitude (flux) scale. The linearity of the system implies that the position of the stellar locus for *any function* of the aperture fluxes is independent of magnitude (at least until images saturate). Therefore marginalizing the flux ratios over magnitude yields 1D distributions that can be used to greatly simplify locating the stellar locus using an iteratively clipped median. With the location estimated, the median of the absolute deviation from the median (MAD) provides a robust measure of the scatter about this locus as a function of magnitude, at least until galaxies dominate in number. This process is repeated iteratively

---

[1] In practice we find that the effects of the spatial variation of the PSF on the aperture fluxes at the detector level are generally negligible.
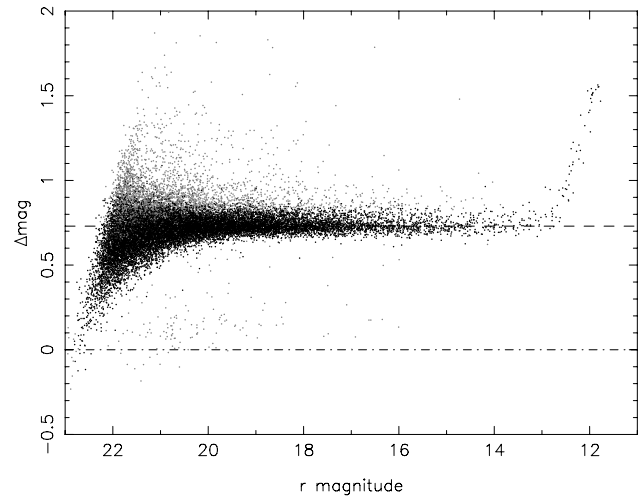
Figure 7. An example of how the difference in flux recorded between pairs of apertures is used to provide information for morphological classification. The stellar locus (black points) is self-defining at a constant offset (here 0.73 mag) while the spread of points about this locus, after iterative clipping, defines the average scatter of stellar objects as a function of magnitude. Grey points denote noise-like objects that lie well below the stellar locus ridge, or non-stellar or blended stellar objects which lie above the stellar population. Saturated stars form the upturned sequence brighter than 13th magnitude. The offset of the top dashed line defines the average differential aperture correction for stellar objects for this pair of apertures and in combination with other pairs of apertures defines the overall stellar curve of growth and hence individual stellar aperture corrections.

for each distribution, using 3σ clipping to remove non-stellar outliers, until satisfactory convergence is reached. After convergence the equivalent Gaussian sigma is estimated using $\sigma_{\mathrm{gauss}} = 1.48$ MAD and by this means each of the image shape descriptors (in this case several different flux ratios) can be renormalized to follow a zero-median, unit variance Gaussian-like $N(0,1)$ distribution.

These measures are then directly combined to form the final classification statistic. The combination (an unweighted sum of the normalized signed distributions) is designed to preserve information on the 'sharpness' of the object profile and is finally renormalized, as a function of magnitude, to produce the equivalent of an overall $N(0,1)$ classification statistic.

In practice measures derived from real images do not exactly follow Gaussian distributions. However, by combining multiple normalized distributions (with well-defined first and second moments), the Central Limit Theorem works in our favour such that the resulting overall statistic is Gaussian-like to a reasonable approximation and hence can be used with due care as the likelihood component of a Bayesian Classification scheme, making optional use of prior knowledge.

Objects lying within $2-3\sigma$ of the stellar locus (i.'. of zero) are generally flagged as stellar images, those below -3 to -5σ (i.e. sharper) as noise-like, and those above $2-3\sigma$ (i.e. more diffuse) as non-stellar. Although the discrete classification scheme is based on the $N(0,1)$ measure of stellar appearance it also has several overrides built in to attempt to make it more reliable. For example, adjustments to the boundaries at the faint end (to cope with increased rms noise in the statistic) and at the bright end (to cope with saturation effects) are also made, while the overall image ellipticity provides a further check.

A by-product of the curve-of-growth analysis and the classification is an estimate of the average PSF aperture correction for

each detector for those apertures (up to and including 4*r*, which includes typically ∼99 per cent, or more, of the total stellar flux) used in deriving the classification statistic. Accurate assessment of the aperture correction to place the (stellar) fluxes on a total flux scale is a crucial component of the overall calibration. We find that this method of deriving aperture corrections contributes ≤ ±1 per cent to the overall photometry error budget and also provides a useful first order seeing correction for non-stellar sources. Further by-products of the morphological classification process are improved estimates of the seeing and average PSF ellipticity from making better use of well-defined stellar-only sources. These parameters are required for quality control monitoring of telescope performance and 'atmospheric' seeing.

3.2 Photometric calibration

Photometric calibration is done using series of Landolt standard stars (Landolt 1992) with photometry in the Sloan Digital Sky Survey (SDSS) system with additional stars from Stetson (2000). For each night a zero-point and error estimate using the observations of all the standard fields in each filter is derived. For photometric nights the calibration over the whole mosaic has an accuracy of 1–2 per cent. For the purpose of the photometric calibration, standard star observations have been obtained each night at an interval of 2 h and have been used to calibrate the *r'* and *i'* frames. The Hα frames have been calibrated using a fixed zero-point offset of 3.14 mag with respect to *r'*, corresponding to the magnitude difference in *r'* − Hα for Vega. This value has been calculated by convolving the *Hubble Space Telescope* Vega spectrum (Bohlin 2007) with the atmospheric throughput, INT optical throughput (primary mirror and field corrector), CCD efficiency and the WFC filter curves (available from the ING web pages at www.ing.iac.es) so that *r'* − Hα = 0 for Vega.

All calibration is by default corrected during pipeline processing for the mean atmospheric extinction at La Palma (0.09 in *r'* and Hα and 0.05 in *i'* per unit airmass).

During non-photometric nights, in otherwise acceptable observing conditions, we find that the derived zero-point systematic errors can be up to 10 per cent or more. Although the pipeline usually successfully flags such nights as non-photometric (see Fig. 8) it still leaves open the problem of what to do about tracking the varying extinction during these nights. For this IDR we have not included a global photometric solution for the whole survey. While it is possible to perform an offset to a common reference system for a small subset of the observations using overlaps, as demonstrated below in Section 5, the application of such methods to the whole survey area is more complicated since one has to carefully select good reference frames and allow for variation in parameters like extinction or zero-point over the whole survey period. Availability of good photometric observations to tie the solution as much as possible is crucial. Clearly this is a major task which we are currently investigating and is best accomplished when the survey is complete. Data from non-photometric nights have been included in the release but flagged as such so that they do not appear in our PhotoObjBest catalogue. Fig. 9 shows the mean magnitude rms for the objects included in this catalogue, zero-point uncertainties are at the level of 2 per cent for *r'* and *i'* and 3 per cent for Hα. We also plot in Fig. 10 the stellar counts in the three bands along two different line of sights in the Galactic plane with different reddening.

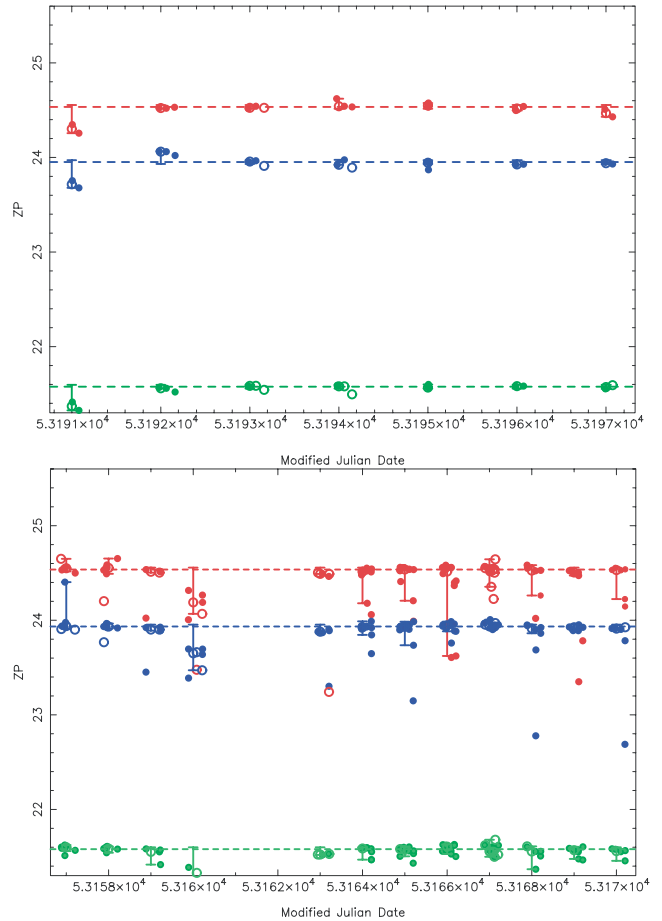All magnitudes quoted are in the Vega system but we give the conversion factors to the AB system in Table 1.

Figure 8. Photometric zero-point (ZP) versus date for two runs. The dashed lines are the expected values for a good photometric night in *r'* (red), *i'* (blue) and Hα (green). Dots represent the individual measurements for each frame. The top figure shows a good photometric run in which all the zero-points are consistent with the expected values. The bottom one instead shows a variable run in which a large percentage of observations have been obtained in bad photometric conditions and therefore not included in the PhotoObjBest catalogue. These plots are used to flag problematic frames and runs.

3.3 Astrometric calibration

Astrometric calibration is a multistage process and aims to provide each image, and any derived catalogues, with a WCS to convert between pixel and celestial coordinates. This happens in the pipeline in two generic stages.

An initial WCS based on knowledge of the instrument, e.g. orientation, field-scale, telescope pointing, is embedded in the FITS headers, with telescope-specific information in the primary header and detector-specific information in the secondary headers. This serves to locate each detector image to within a few to several arcseconds, depending on the pointing accuracy of the telescope and model parameters. The essential information required is the right ascension (RA) and declination (Dec.) of the pointing, a (stable) reference point on the detector grid for those coordinates (e.g. the optical axis of the instrument), the central pixel scale, the rotation of the camera, the relative orientation of each detector and the geometrical distortion of the telescope and camera optics, which defines the astrometric projection to use.
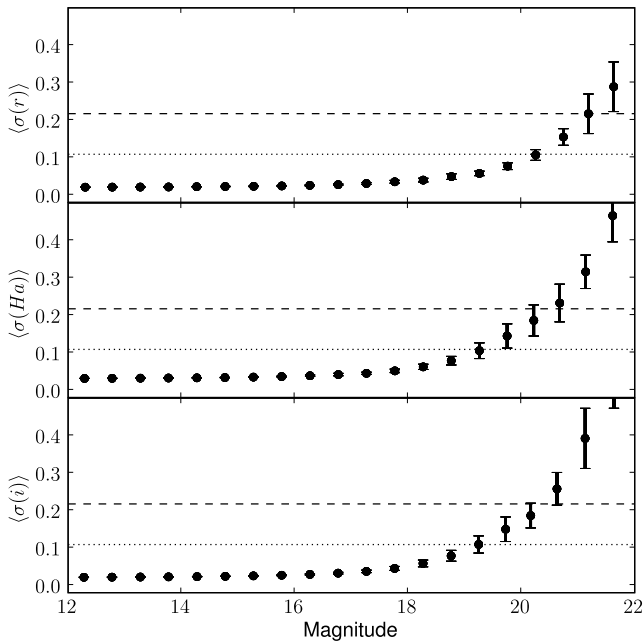
Figure 9. Photometric measurement uncertainties as a function of magnitude in 0.5 mag bins for all the objects included in PhotoObjBest. The vertical error bars show the rms dispersion in the mean uncertainties in each brightness bin and the horizontal lines in each panel show the S/N = 10 (dotted) and S/N = 5 (dashed) values.

Given a rough WCS for the processed frames, a more accurate WCS can be defined using astrometric standards. We have based our calibration on the 2MASS point source catalogue (Skrutskie et al. 2006) for several reasons: it is an all-sky near-infrared (NIR) survey; it is calibrated on the International Celestial Reference System (ICRS); it provides at least 100 or more suitable standards per pointing; it is a relatively recent epoch (mid-1990s) minimizing proper motion problems; the global systematics are better than 100 mas over the entire sky (Zacharias et al. 2003); and for 2MASS point sources with signal-to-noise ratio (S/N) > 10:1 the rms accuracy per source is $\lesssim$100 mas.

By stacking the astrometric residuals from a series of independent pointings and CCD WCS solutions it is possible to assess the accuracy of the simple INT distortion model. This is illustrated in the Fig. 11 using the average residuals from a stack of a one week run. The end product of the full pipeline currently has an astrometric precision better than 100 mas over the whole array (i.e. across CCDs), as determined by analysis of independently calibrated adjacent overlapping pointings.

4 DATA PRODUCTS

4.1 Images

There are two types of imaging data products: (1) object images and (2) confidence maps. At this point we do not provide stacked images from the overlapping frames. We describe each of them separately.

As stated before, each image is a MEF file containing the four CCD frames. Each CCD is photometrically and astrometrically calibrated as described in the previous section. Photometry can be extracted from the images using the following expression in the instrumental system:

$$m = \mathrm{ZP} - \kappa\,(X - 1) - 2.5\log_{10}(f/t) + c, \qquad (4)$$

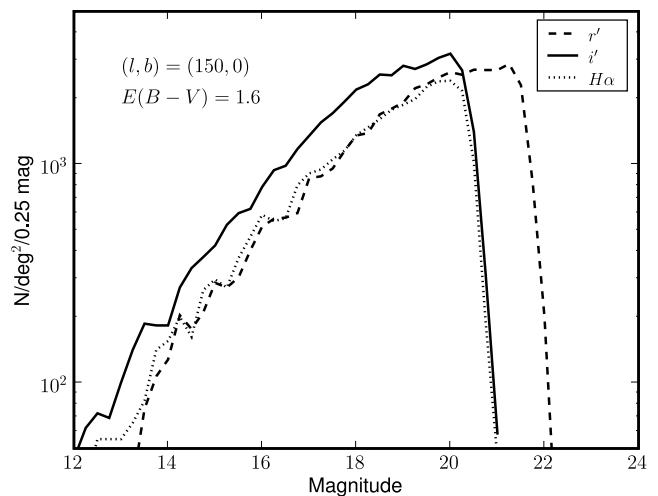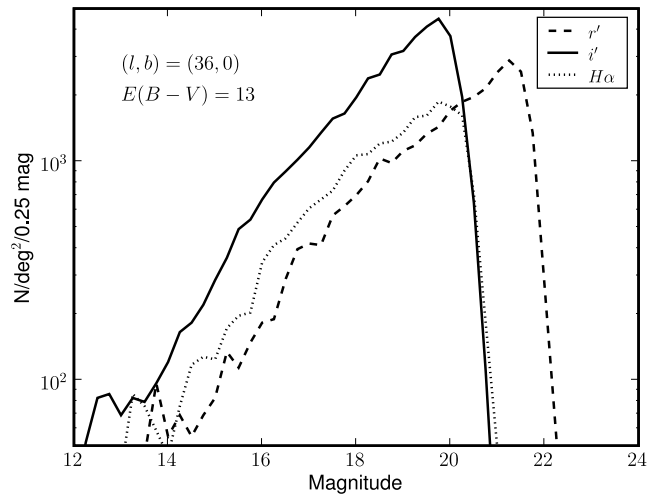**Figure 10.** Star counts along two contrasted lines of sight with star densities of $\sim$50 000 deg^2 and different reddening.

where m is the calibrated magnitude, ZP is the zero-point at unit airmass as extracted from the image headers, κ is the extinction coefficient, X is the airmass, f is the flux in counts and t is the exposure time. The c correction is based on the median dark sky recorded in science frames compared to the median for all the CCDs and as such is an ancillary correction to the gain correction derived from the flat-field (usually twilight flats). This correction is in the header keyword PERCORR and is typically smaller than 1 per cent and also corrects for the small quantum efficiency differences between the detectors.

Astrometric information is written in the FITS headers using a ZPN projection which allows for a description of the non-uniform scale over the field of view without performing pixel resampling (necessary e.g. in the case of describing the distortions using a more usual tangential projection).

The aim of conventional flat-fielding is to create a flat background by normalizing out perceived variations from uniformly illuminated frames. If the sky area per pixel changes then this is reflected in a systematic error in the derived photometry. However, the change of scale, i.e. the change of sky area per pixel, also creates photometric effects. The magnitude of the multiplicative correction to be applied
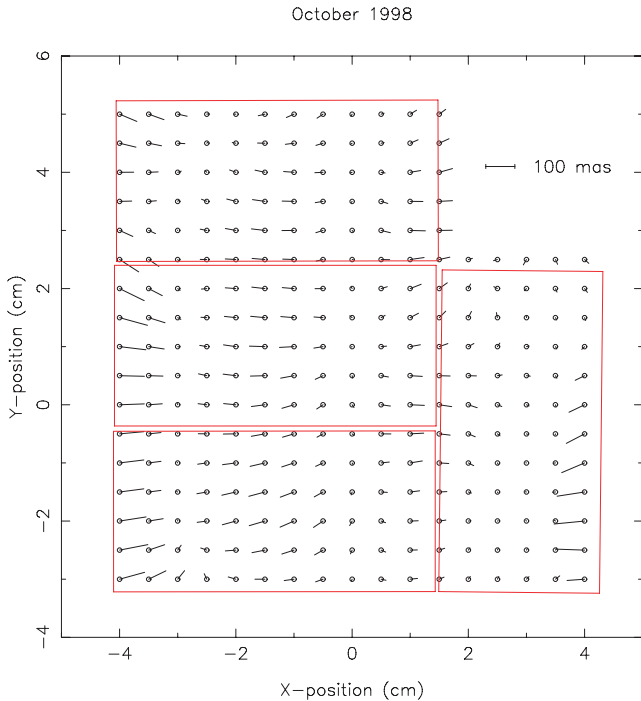
October 1998

Figure 11. Astrometric residuals derived from a series of independent WFC pointings. The residuals are shown in physical coordinates with respect to the centre of the optical axis.

to the measured flux can be modelled as (Irwin et al., in preparation)

$$d = \left(1 + \frac{3\,P_3\,R^2}{P_1}\right)\left(1 + \frac{P_3\,R^2}{P_1}\right), \tag{5}$$

where R is the distance from the optical axis and for the WFC the coefficients are $P_1 = 1$ and $P_3 = 220$ (corresponding to the PV2_1 and PV2_3 WCS keywords as defined in Greisen & Calabretta 2002). Fig. 12 shows the effect of this term in the photometry which is at a maximum level of only 2 per cent at the outer parts of the frames. The value displayed in figure has to be subtracted from the magnitude calculated in equation (4) and is automatically included in the catalogue data products but has to be applied by the user when measuring fluxes from the images. Table 3 lists the main
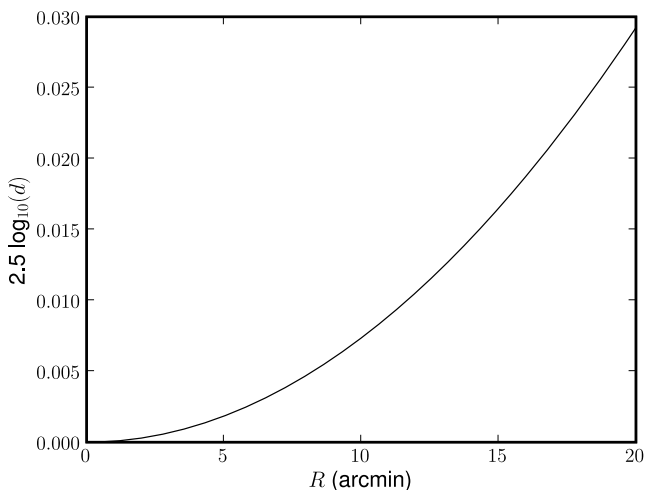

Figure 12. Effect of scale change in photometry (magnitude versus distance to field centre).

Table 3. Subset of relevant photometric and astrometric keywords available from the image headers after pipeline processing.

Keyword	Description
MAGZPT	(mag) Magnitude zero-point for default extinction measured at unit airmass
MAGZRR	(mag) Error in magnitude zero-point
AIRMASS	Airmass at start of observation
EXPTIME	(s) Exposure time
WFFBAND	Filter bandpass
SKYLEVEL	(counts pixel^{-1}) Median sky brightness
SKYNOISE	(counts) Pixel noise at sky level
SEEING	(pixels) Mean stellar FWHM
APCOR	(mag) Stellar aperture correction
PERCORR	(mag) Sky calibration correction
NUMBRMS	Number of objects used when computing the astrometric fit
STDCRMS	(arcsec) Astrometric fit error

photometric keywords present in the image headers and needed to convert the fluxes to magnitudes as explained in this section.

4.1.1 Confidence maps

We define a confidence c_{ij} map, where j refers to pixel j of frame i, that is an inverse variance weight map in units of relative variance denoting the 'confidence' associated with the flux value in each pixel. The c_{ij} are normalized to a median level of 100 per cent (i.e. $\langle c_{ij}\rangle = 1$). This has the advantage that the same map can also be used to encode for hot, bad or dead pixels, by assigning zero confidence.

The confidence map for each frame is derived from analysis of the flat-fields and is unique for each filter/detector combination. It encodes individual pixel sensitivities and allows for vignetted or bad regions to be masked. These confidence maps are used in the object detection phase of the processing.

Use of the confidence maps for weighted co-addition of frames, or for object detection, simply requires an overall estimate of the average noise properties of the frame. This can be readily derived from the measured sky noise, in the Poisson noise-limited case, or from a combination of this and the known system characteristics (e.g. gain, readout noise). All processed frames have an associated derived confidence map which is propagated through the processing chain. A detailed description of the process is given in Irwin et al. (in preparation).

4.2 Catalogues

The object catalogue is made available using a relational data base system. We have followed the philosophy of reducing the number of tables presented to the user as much as possible. The data base contains four tables which are summarized below.

(i) FieldsAll (Table 4): information about the field centres and number of objects detected. Each field pair has associated a unique field number (fieldno) and each of those an on/off position (onoff where a value of 0 has been assigned to the off position and 1 to the on position). These two are combined into the fieldID to provide a unique identifier for each pointing. Equatorial and Galactic coordinates for each pointing as well as number of detections are also provided.

Table 4. Column names, types and description of FieldsAll table. Columns marked with an asterisk are indexed.

Column name	Type	Description
fieldID*	long	Field ID
onoff*	int	On/off observation flag
fieldno*	int	Field number
ra*	float	RA of centre of field (deg, J2000)
dec*	float	Dec. of centre of field (deg, J2000)
glon	float	Galactic longitude of centre of field (deg)
glat	float	Galactic latitude of centre of field (deg)
nobj	int	Number of objects detected

Table 5. Column names, types and description of data quality control table ChipsAll. Columns marked with an asterisk are indexed.

Column name	Type	Description
chipID*	long	Chip ID
runno*	int	Run number
ra*	float	RA of the chip centre (deg, J2000)
dec*	float	Dec. of the chip centre (deg, J2000)
glon	float	Galactic longitude (deg)
glat	float	Galactic latitude (deg)
ebv	real	$E(B - V)$ from Schlegel dust map
naxis1	int	Chip size along x-axis (pixels)
naxis2	int	Chip size along y-axis (pixels)
mjd	float	MJD of observation
magzpt*	real	Photometric zero-point (Vega)
magzrr	real	Zero-point error (Vega)
extinct	real	Atmospheric extinction
airmass*	real	Airmass
exptime	real	Exposure time (s)
night*	int	Night of observation (YYYYMMDD)
dateobs*	int	Date of observation (YYYYMMDD)
band*	char	Filter name
maglim*	real	Magnitude limit (5σ, Vega)
seeing*	real	Mean seeing (arcsec)
elliptic*	real	Mean stellar ellipticity
cd1_1	float	CD matrix
cd1_1	float	CD matrix
cd1_1	float	CD matrix
cd1_1	float	CD matrix
crval1	float	RA of reference pixel (deg)
crval2	float	Dec. of reference pixel (deg)
crpix1	float	x coordinate of reference pixel
crpix2	float	y coordinate of reference pixel
percorr	real	Sky correction (Vega)
apcor	real	Aperture correction (Vega)
skylevel	real	Sky level (counts)
skynoise	real	Sky noise (counts)
nobj	real	Number of detected objects
field_id*	long	Link to FieldsAll table

Table 6. Column names, types and description of main photometric tables PhotoObj and PhotoObjBest. Columns marked with an asterisk are indexed. Columns marked with †*ger* are not included in the PhotoObjBest table.

Column name	Type	Description
objID*	long	Object ID
objname	char	Object name
ra*	float	RA in degrees (deg, J2000)
dec*	float	Dec. in degrees (deg, J2000)
glon	float	Galactic longitude (deg)
glat	float	Galactic latitude (deg)
coremag_r	float	r' aperture corrected magnitude
coremag_i	float	i' aperture corrected magnitude
coremag_ha	float	Hα aperture corrected magnitude
coremagerr_r	float	r' magnitude error
coremagerr_i	float	i' magnitude error
coremagerr_ha	float	Hα magnitude error
rmi*	float	$r'-i'$ colour
rmha*	float	$r'-$Hα colour
rmierr	float	$r'-i'$ colour error
rmhaerr	float	$r'-$Hα colour error
class_r	int	r' star/galaxy classification index
class_i	int	i' star/galaxy classification index
class_ha	int	Hα star/galaxy classification index
ellipt_r	real	Ellipticity as measured in r'
ellipt_i	real	Ellipticity as measured in i'
ellipt_ha	real	Ellipticity as measured in Hα
posang_r	real	Position angle as measured in r'
posang_i	real	Position angle as measured in i'
posang_ha	real	Position angle as measured in Hα
seeing_flag*	int	Flag sources with seeing <1.7
photom_flag†*ger*	int	Zero if object detected in all bands
class_flag*	int	Combined probably of stellar object
chip_r_id*	long	Link to r' chip table
chip_i_id*	long	Link to i' chip table
chip_ha_id*	long	Link to Hα chip table
cx†	float	Spherical coordinate $x = \cos(\alpha)\sin(\delta)$
cy†	float	Spherical coordinate $y = \sin(\alpha)\sin(\delta)$
cz†	float	Spherical coordinate $z = \cos(\delta)$
htmid*	long	20-deep HTM index of this object
field_id*	long	Link to field table

(ii) ChipsAll (Table 5): this table contains parameters about the observation (airmass, exposure time, WCS) as well as the Data Quality Control attributes for each chip, i.e. mean seeing and ellipticity, zero-point, sky level, etc. Also for each pointing the Galactic coordinates are available as well as the Galactic extinction as measured from the Schlegel maps (Schlegel, Finkbeiner & Davis 1998) using a 4 pixel linear interpolation for each position.

(iii) PhotoObj (Table 6): main catalogue containing all detections with image derived parameters and calibrated magnitudes. Each row

has a set of keys (chip_r_id, chip_i_id, chip_ha_id, field_id) which allow the user to enquire about the photometric conditions (via the ChipsAll table) under which the measurements of a particular object have been done or to which field it belongs. Additional flags (e.g. seeing_flag) allow the selection of objects observed in particular (seeing) conditions. We have also included the hierarchical triangular mesh index (htmid) as defined by the Sloan Digital Survey Science Archive.[2] The total number of detections in this catalogue is 411 million including duplicate observations. Due to the on/off observation strategy which leads to every source being observed twice this means about 200 million unique objects.

(iv) PhotoObjBest (Table 6): subset of PhotoObj with additional constraints used to select only those observations obtained in good photometric conditions and thus will vary little between the IDR and the Final Data Release. The data obtained in the 2003 August and 2004 December observing runs have been discarded due to poor photometric conditions. This catalogue contains only observations for those nights for which the mean magnitude zero-point rms in
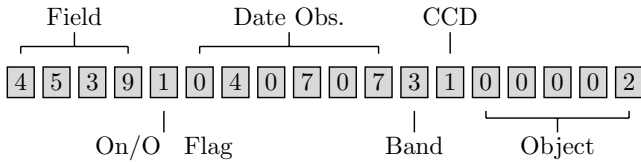
---

[2] See http://skyserver.org/htm/.

Figure 13. The unique object ID encodes the field number, on/off flag (1 = on, 0 = off), date of observation (in the format YYMMDD), band (1 = *r'*; 2 = *i'*; 3 = Hα) and CCD number.

the *r'* band is better than 5 per cent, the mean *r'* zero-point is greater than 24.2 and the individual frames have been obtained in seeing conditions better than 2 arcsec and ellipticity better than 0.2. Additionally only sources detected in all three bands have been included. The total number of detections in this catalogue is 181 million (approximately 90 million unique objects).

(v) tGetNearbyObj: this is a store procedure to perform a cone search without having to use the complete haversine formula explicitly to calculate angular distances. Given the coordinates (ra, dec) in degrees and a radius (r) in arcmin it returns a list of identifiers for the objects in such defined cone. The output can then be joined with one of the PhotoObj/Best tables to return the object parameters (see example in Appendix A).

Tables 4–6 list the column names and their description. Columns marked with an asterisk are indexed by the data base software, i.e. searches in those columns are particularly fast. Users should use when possible at least one indexed column in order to speed up the queries.

Each entry in each table is assigned a unique ID number used for cross-reference. Fig. 13 shows an object ID in the PhotoObj/Best tables and how it is decoded. Each entry in the ChipsAll table is encoded in the same way but without the last five figures which represent the object number.

Individual detection star/galaxy classifications are combined into one single value following a similar approach to the WFCAM Science Archive (WSA; Hambly et al. 2007). A classification table is defined which assigns probability values for each classification code (Table 7) and then each classification is combined for a merged source using Bayesian classification rules assuming each value is independent:

$$P(c_k) = \prod_j P(c_k)_j \bigg/ \sum_k \prod_j P(c_k)_j, \qquad (6)$$

where c_k is the classification flag and *i* denotes the *i*th single detection passband measurement available. Decision thresholds for the resulting discrete classification flag are 90 per cent for definitive and 70 per cent for probable. An additional decision rule enforces class_flag=255 (saturated) when any individual classification flag indicates saturation.

5 CROSS-MATCHING WITH UKIDSS GPS

In Drew et al. (2005) we describe the importance of the IPHAS data by themselves and the several papers published to date using just IPHAS data are clearly a demonstration of that. In this section however, we want to focus in the multiwavelength aspect which the present IPHAS release will facilitate. We provide an example of cross-matching the sources extracted from the IPHAS catalogue with those from the UKDSS GPS. NIR data are required to probe the many parts of the galaxy which suffer high extinction and an

Table 7. Meaning of classification index and probability assigned to each of them.

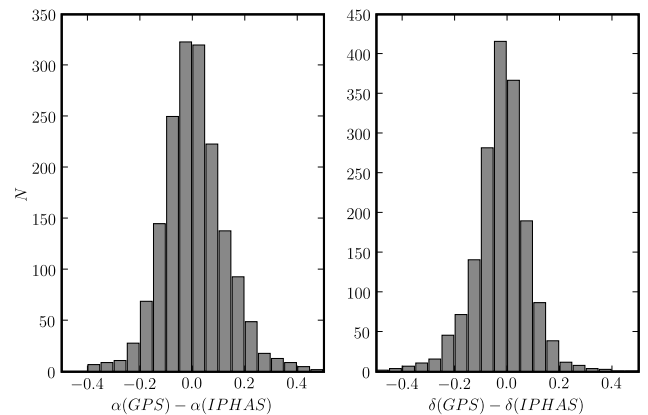| Index	Meaning	Probability (per cent)			
		Star	Extended	Noise	Saturated
−9	Saturated	0	0	5	95
−3	Probable extended	25	70	5	0
−2	Probable star	70	25	5	0
−1	Star	90	5	5	0
0	Noise	5	5	90	0
1	Extended	5	90	5	0

Figure 14. Positional offsets in RA (left-hand panel) and Dec. (right-hand panel) between GPS and IPHAS positions (arcsec) in one of the fields included in the area considered in the text.

ESO-public NIR survey of the northern plane, the UKIDSS Galactic Plane Survey, has recently begun.

We have selected a contiguous area of 2 deg² centred at $\alpha = 20^h48^m$, $\delta = 42°45'$ (*l, b* = 83.0, −0.5) in the Cygnus-X region and extracted all objects from the PhotoObjBest catalogue. From the Schlegel et al. (1998) maps the extinction is $E(B - V) = 3$. The total number of objects in that area is 235 231. In order to assess the photometry we have grouped all the objects by field and used the field overlap algorithm as described in Glazebrook et al. (1994) in order to offset the photometry to a common reference system.

For this exercise we have just set one frame as the reference calibrated frame and allowed the others to vary and solved the matrix equation for each band. The resultant magnitude offsets between the different fields is smaller than ∼0.03 in the three bands confirming the good photometric calibration.

For the same area we extract also the NIR magnitudes from the UKIDSS GPS survey. The GPS is carrying out a complementary survey of the Galactic plane in *J, H* and *K*. Using AstroGrid access to the WFCAM Science Archive (WSA) we download the NIR catalogue of objects detected in the area covered by the IPHAS observations. We then cross-match the GPS and IPHAS catalogues selecting the nearest neighbour in a radius of 1.2 arcsec using STILTS.[3] Fig. 14 shows the positional offsets between the IPHAS and GPS catalogues for one field. The offset is effectively zero and the rms is 0.13 and 0.12 arcsec in Δα and Δδ, respectively. This good agreement is not coincidence since both surveys use 2MASS as the reference

---

[3] STILTS is a set of command line tools to perform operations with tables like cross-matching (http://www.star.bris.ac.uk/~mbt/stilts/).

Table 8. Synthetic tracks in the $(r' - J, J - K)$ plane, for main-sequence dwarfs and giants, calculated for a range of reddening. A similar table with optical colours only can be found in table 2 of Drew et al. (2005).

Spectral type	Reddening									
	$E(B-V)=0.0$		$E(B-V)=1.0$		$E(B-V)=2.0$		$E(B-V)=3.0$		$E(B-V)=4.0$	
	$(r'-J)$	$(J-K)$	$(r'-J)$	$(J-K)$	$(r'-J)$	$(J-K)$	$(r'-J)$	$(J-K)$	$(r'-J)$	$(J-K)$
O5 V	−0.955	−0.331	1.415	0.246	3.742	0.820	6.028	1.391	8.275	1.959
O9 V	−0.826	−0.252	1.543	0.323	3.869	0.895	6.154	1.465	8.401	2.032
B0 V	−0.720	−0.139	1.649	0.437	3.976	1.010	6.262	1.580	8.509	2.147
B1 V	−0.674	−0.107	1.690	0.467	4.011	1.038	6.291	1.607	8.533	2.172
B3 V	−0.706	−0.283	1.658	0.294	3.980	0.868	6.262	1.439	8.505	2.008
B8 V	−0.249	0.000	2.030	0.531	4.423	1.144	6.806	1.715	9.044	2.281
B9 V	−0.145	0.003	2.108	0.573	4.530	1.147	6.697	1.712	8.933	2.278
A0 V	0.005	−0.005	2.366	0.569	4.684	1.140	6.961	1.708	9.200	2.274
A2 V	−0.139	−0.088	2.220	0.489	4.536	1.063	6.812	1.634	9.050	2.204
A3 V	−0.155	−0.131	2.202	0.446	4.517	1.020	6.791	1.591	9.028	2.160
A5 V	0.324	0.085	2.678	0.658	4.990	1.228	7.261	1.796	9.494	2.361
A7 V	0.390	0.132	2.739	0.705	5.047	1.275	7.314	1.842	9.544	2.407
F0 V	0.543	0.159	2.888	0.731	5.192	1.300	7.455	1.867	9.683	2.432
F2 V	0.475	0.057	2.819	0.632	5.122	1.205	7.384	1.775	9.612	2.343
F5 V	0.859	0.221	3.202	0.793	5.503	1.362	7.765	1.929	9.991	2.493
F6 V	0.909	0.218	3.249	0.788	5.547	1.357	7.806	1.922	10.029	2.486
F8 V	0.995	0.230	3.333	0.802	5.630	1.372	7.887	1.939	10.109	2.504
G0 V	1.042	0.275	3.378	0.847	5.674	1.416	7.930	1.982	10.150	2.546
G2 V	1.271	0.329	3.606	0.899	5.899	1.467	8.154	2.031	10.373	2.594
G5 V	1.210	0.310	3.543	0.882	5.836	1.451	8.089	2.018	10.308	2.582
G8 V	1.415	0.398	3.746	0.968	6.035	1.536	8.286	2.101	10.502	2.664
K0 V	1.528	0.499	3.853	1.068	6.138	1.635	8.385	2.198	10.598	2.760
K2 V	1.790	0.550	4.114	1.118	6.398	1.683	8.643	2.246	10.855	2.807
K3 V	1.899	0.556	4.220	1.125	6.501	1.691	8.744	2.254	10.954	2.815
K4 V	2.096	0.648	4.411	1.216	6.686	1.781	8.924	2.344	11.128	2.904
K5 V	2.245	0.674	4.560	1.241	6.835	1.806	9.074	2.368	11.280	2.928
M0 V	2.992	0.781	5.296	1.348	7.561	1.912	9.789	2.474	11.985	3.033
M2 V	3.418	0.807	5.961	1.375	7.971	1.942	10.194	2.506	12.387	3.067
M3 V	3.953	0.820	6.237	1.388	8.482	1.953	10.695	2.516	12.878	3.076
M4 V	4.560	0.832	6.830	1.399	9.062	1.965	11.261	2.527	13.433	3.088
O8 III	−0.813	−0.076	1.550	0.496	3.870	1.065	6.149	1.631	8.390	2.195
B1-2 III	−0.603	−0.198	1.766	0.376	4.092	0.947	6.377	1.516	8.623	2.082
B3 III	−0.427	0.018	1.935	0.591	4.253	1.161	6.530	1.729	8.769	2.295
B9 III	−0.206	−0.070	2.153	0.504	4.469	1.076	6.744	1.644	8.982	2.210
B5 III	−0.361	−0.078	2.000	0.497	4.318	1.068	6.595	1.637	8.835	2.203
A0 III	−0.030	0.020	2.326	0.593	4.641	1.164	6.915	1.732	9.151	2.297
A3 III	−0.119	−0.139	2.238	0.437	4.553	1.010	6.828	1.581	9.065	2.150
A5 III	0.323	0.101	2.675	0.673	4.985	1.243	7.254	1.810	9.486	2.374
A7 III	0.285	−0.003	2.638	0.572	4.948	1.144	7.218	1.713	9.451	2.280
F0 III	0.565	0.199	2.911	0.770	5.216	1.339	7.481	1.905	9.709	2.468
F2 III	0.595	0.100	2.936	0.673	5.236	1.242	7.496	1.809	9.720	2.373
F5 III	0.876	0.291	3.214	0.861	5.510	1.428	7.766	1.993	9.987	2.556
G0 III	1.458	0.424	3.791	0.992	6.083	1.559	8.337	2.123	10.555	2.684
G5 III	1.681	0.477	4.008	1.046	6.294	1.612	8.542	2.175	10.756	2.736
G8 III	1.727	0.516	4.053	1.086	6.339	1.653	8.586	2.217	10.799	2.779
K0 III	1.844	0.541	4.167	1.110	6.449	1.676	8.694	2.239	10.904	2.800
K1 III	1.961	0.571	4.283	1.140	6.564	1.707	8.808	2.271	11.018	2.832
K2 III	2.188	0.671	4.506	1.239	6.783	1.805	9.024	2.368	11.230	2.929
K5 III	2.948	0.864	5.253	1.431	7.519	1.995	9.748	2.557	11.945	3.116
M0 III	3.190	0.968	5.493	1.534	7.756	2.097	9.983	2.657	12.177	3.215
M1 III	3.364	0.987	5.664	1.553	7.925	2.116	10.149	2.676	12.341	3.234
M2 III	3.586	1.068	5.881	1.633	8.135	2.196	10.354	2.757	12.541	3.315
M3 III	3.910	1.062	6.200	1.627	8.450	2.190	10.664	2.750	12.846	3.307
M4 III	4.464	1.086	6.744	1.652	8.983	2.215	11.187	2.775	13.361	3.333
M5 III	5.249	1.142	7.515	1.707	9.742	2.270	11.934	2.830	14.098	3.387

astrometric system and the WFCAM processing pipeline utilizes the same algorithms developed at CASU for the optical processing. It is however worth noting the success of the astrometric pipeline in such different data products.

The importance of the $(r' - H\alpha, r' - i')$ plane has been extensively discussed in Drew et al. (2005). The mean sequence is clearly defined and sources above it are emission line objects (see Fig. 16). We explore here the $(r' - H\alpha,$
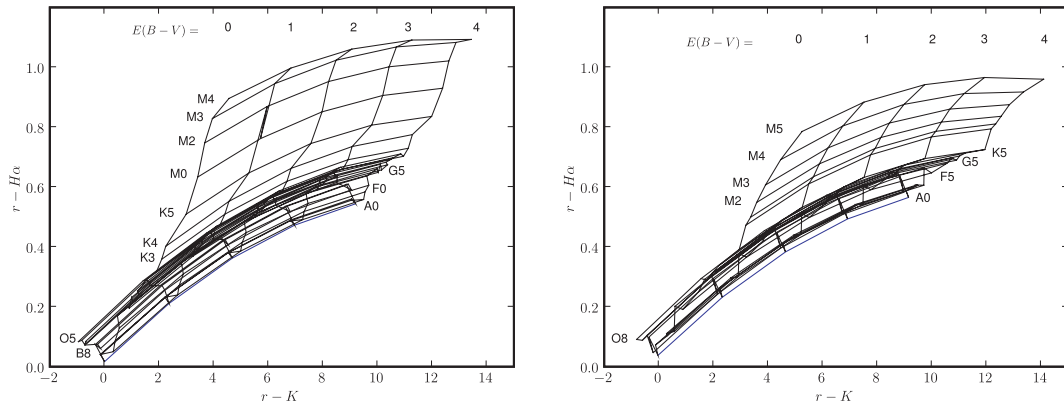
Figure 15. Reddening tracks in the $r - H\alpha$ versus $r - K$ plane for main-sequence dwarfs (left-hand panel) and giants (right-hand panel).
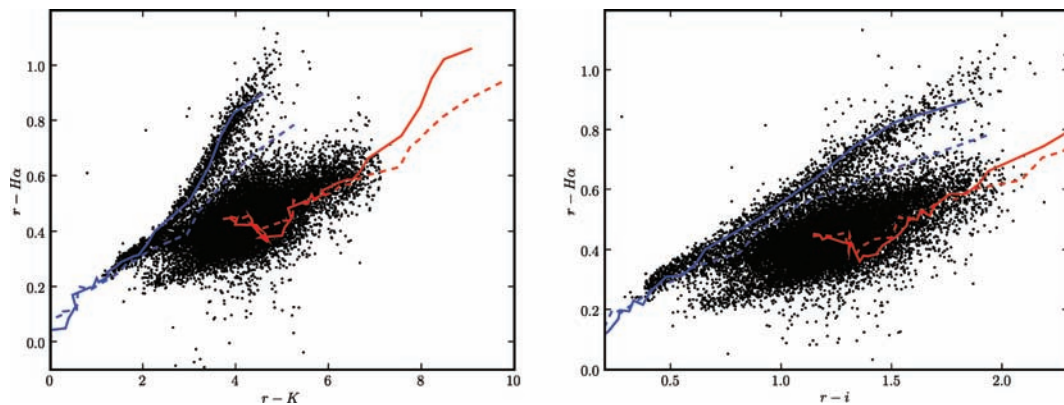
Figure 16. Colour–colour plots for point like sources detected in IPHAS and GPS with main-sequence tracks overlayed (continuous line: main-sequence dwarfs; dashed-line: giants) for reddening: $E(B - V) = 0$ (blue) and 2 (red)).

$r' - K$) plane provided by the combination of IPHAS and GPS.

Using the SED model library from Pickles (1998) and the WF-CAM filter profiles we have simulated the colours of normal stars. The detailed process is described in Drew et al. (2005) and Sale et al. (in preparation) and we list in Table 8 the synthetic colours for dwarfs and giants only. The tracks are plotted in Fig. 15 for extinction between 0 and 4. Fig. 16 shows the $(r' - H\alpha, r' - K)$ and $(r' - H\alpha, r' - i')$ colour–colour diagrams with $E(B - V) = 0$ and $E(B - V) = 2$ reddening tracks for dwarfs and giants overplotted. We have included in these plots only sources classified as point-like in the optical and NIR data (class_r $= -1, -2$, pStar > 0.7) with magnitude errors smaller than 0.04 in r' and K. We have also excluded objects brighter than $r' = 13$ and $K = 12$ to avoid saturation and finally we have removed all sources within 10 arcsec of any WFCAM chip edge (although problematic sources in the edges are anyway flagged by the pipeline classifier).

The main sequence of objects at zero reddening is well defined in the $(r' - H\alpha, r' - K)$ plane with a large population of objects up to the M stars. These data are for fields in the Cygnus-X part of the plane where there is a local essentially unreddened stellar population to a few hundred parsecs, and then a decline in stellar density and sharp rise in molecular gas and extinction to a distance of a kpc or so. Beyond this, the stellar density recovers as distant associations are traversed but of course it presents as a much more heavily reddened population. Furthermore stars later than spectral type mid-K are intrinsically faint and are only detectable at $E(B -$

$V) < 0.8$ in the Fig. 16 where our limit is at $r' - K < 7$. This explains the lack of K and M dwarfs reddened objects in these diagrams.

Note that no artificial offsets have been introduced to the synthetic tracks to make them fit with the observed colours. This is further confirmation of the good photometric calibration of both the IPHAS and GPS catalogues.

6 USER ACCESS

User access to the catalogues and images has been provided by the AstroGrid project. There are two separate user interfaces to the IPHAS IDR. Users who only want to extract fluxes or images around a small position of sky (there is a limit on the radius of $1°$) will find the web interface useful.[4] Given a position or an object name the user will obtain a cone search of the PhotoObjBest catalogue in a variety of output formats (VOTable, Binary VOTable, Comma Separated Value or HTML). Input lists of objects are also supported. Together with the IPHAS catalogue we provide access from the same interface to other useful VO compliant archives like SDSS DR5 (Adelman-McCarthy et al. 2007), 2MASS Point Source Catalogue (Skrutskie et al. 2006), NED and the Carlsberg Meridian Astrometric Catalogue (Evans, Irwin & Helmer 2002). Postage stamps generation for a particular input position is also available using the finding chart utility (see example output in Fig. 17).

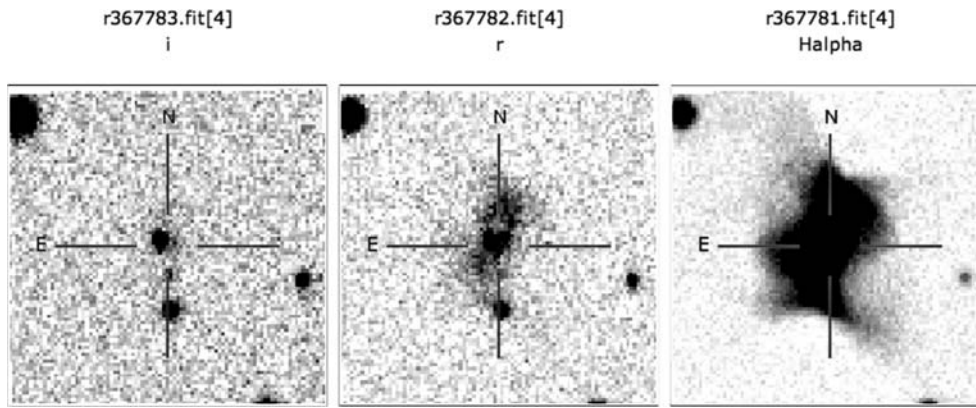---

[4] Available from http://www.iphas.org/idr.

Figure 17. Example postage stamps obtained using the finding chart utility. Each stamp is labelled with the run number, CCD number and filter. The object corresponds to the Principe de Asturias quadrupolar nebula presented in Mampaso et al. (2006).

Users who want more sophisticated queries on flags and algebraic combinations of parameters (e.g. colours) will find the AstroGrid Query Builder more suited to their needs. The Query Builder is an application inside the AstroGrid VO Desktop which allows us to submit an arbitrary query to the data base system. The Query Builder facilitates the task of building complex queries written in Astronomical Dataset Query Language (ADQL,[5] an international agreed standard based on the Structured Query Language – SQL), save the queries and submit them asynchronously to the server (this facilitates sending long running queries and retrieve the results later).

6.1 Image access

The Simple Image Access Protocol (SIAP) defines a prototype standard for retrieving image data from a variety of astronomical image repositories through a uniform interface.[6] The IPHAS images are available through a SIAP (ivo://uk.ac.cam.ast/IPHAS/ images/SIAP) interface. The result of a query is, given a box centred at coordinates RA, Dec. and a box size in degrees, a table of image CCDs which overlap the defined box. The table contains links to the processed image file itself. Alternatively the same list of images can be obtained with an appropriate SQL query (see below). At the moment only full image retrieval is supported. Although the graphical interface does not allow for a list of positions this is possible to accomplish from the command line interface.

6.2 Full catalogue access

The ingestion of all the catalogue data into a data base provides the capability of using SQL in order to query the catalogue with any user defined constraint. The AstroGrid Data Set Access (DSA) provides a layer on top of any data base system which makes it readily VO compliant and accessible from any VO compliant application. The examples in appendix A provide ADQL queries that can be used in the Query Builder to retrieve a variety of data.

6.3 Command line access

The AstroGrid Run Time (AR) provides a bridge between the user command line tools or programs and the VO. Catalogues can be accessed using this bridge from programming languages like PYTHON, PERL, C or JAVA. There are several reasons to provide a command line interface to the VO. The more important one is the ability to perform batch queries for a list of objects and also to implement a call to a VO service inside a user written script that performs some analysis afterwards.

7 SUMMARY

We have presented the IDR of the IPHAS project. The total release contains photometry in r', i' and Hα for about 200 million objects in the northern Galactic plane comprising the largest photometric sample available imaged in narrowband Hα. We have also made available a subset of this catalogue, PhotoObjBest, which contains the best observations in terms of final survey quality as defined in the text and which comprises about 60 per cent of the final release. The legacy of the survey is increased by the extensive coverage appearing from other projects spanning the wavelength range from X-rays to NIR and submillimetre.

The data have been made available through a VO compliant interface provided by AstroGrid and accessible to all VO projects.

This initial release will be supplemented with data obtained from 2006 January onwards at a later date to be announced in the project web pages. At the time of writing, data taking continues. There will also be a separate Final Release once the survey is entirely complete and all finally accepted data have been corrected on to a fully interlaced uniform photometric scale.

More information and examples are available from the IPHAS IDR and AstroGrid websites at http://www.iphas.org/idr and http://www.astrogrid.org.

[5] http://www.ivoa.net/Documents/latest/ADQL.html.
[6] See http://www.ivoa.net/Documents/latest/SIA.html.

telescope: Andrew Cardwell, Jesus Corral-Santana, Pasqual Diago, James Furness, Antonio García, Wilbert van Ham, Marie Hrudkova, Avon Huxor, Mansura Jaigirdar, Rob King, Richard Parker, Mark Pringle, Aaron Robotham, Ernesto Rodriguez-Flores, Miguel Santander, Rachel Smith and Carolin Winkworth.

REFERENCES

Adelman-McCarthy J. K. et al., 2007, ApJS, 172, 634
Bahcall J. N., Soneira R. M., 1981, ApJ, 246, 122
Bohlin R. C., 2007, in Sterken C., ed., ASP Conf. Ser. Vol. 364, The Future of Photometric, Spectrophotometric and Polarimetric Standardization. Astron. Soc. Pac., San Francisco, p. 315
Corradi R. L. M., et al., 2008, A&A, 480, 409
Drew J. E. et al., 2005, MNRAS, 362, 753
Drew J. E., Greimel R., Irwin M. J., Sale S. E., 2008, MNRAS, 386, 1761
Dye S. et al., 2006, MNRAS, 372, 1227
Evans D. W., Irwin M. J., Helmer L., 2002, A&A, 395, 347
Glazebrook K., Peacock J. A., Collins C. A., Miller L., 1994, MNRAS, 266, 65
Greisen E. W., Calabretta M. R., 2002, A&A, 395, 1061
Hambly N. C. et al., 2008, MNRAS, 384, 637
Irwin M. J., 1985, MNRAS, 214, 575
Irwin M., Lewis J., 2001, New Astron. Rev., 45, 105
Irwin M. J., Trimble V., 1984, AJ, 89, 83
Irwin M., McMahon R., Walton N., González-Solares E., Hodgkin S., Irwin J., Lewis J., 2005, ING Newsl., 9, 8
Kuijken K. et al., 2004, in Moorwood A. F. M., Masanori I., eds, Proc. SPIE Vol. 5492, Ground-based Instrumentation for Astronomy. SPIE, Bellingham, p. 484
Landolt A. U., 1992, AJ, 104, 340
Lucas P. W. et al., 2007, preprint (arXiv:0712.0100)
McMahon R. G., Walton N. A., Irwin M. J., Lewis J. R., Bunclark P. S., Jones D. H., 2001, New Astron. Rev., 45, 97
Mampaso A. et al., 2006, A&A, 458, 203
Martin E. L., Solano E., Cabrera J., Valdivieso L., Rodrigo C., 2008, A&A, submitted
Naylor T., 1998, MNRAS, 296, 339
Parker Q. A. et al., 2005, MNRAS, 362, 689
Pickles A. J., 1998, PASP, 110, 863
Pilbratt G. L., 2005, ESA-SP, 577, 33
Russell D. M., Fender R. P., Gallo E., Kaiser C. R., 2007, MNRAS, 376, 1341
Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
Scott D., Tout C. A., 1989, MNRAS, 241, 109
Skrutskie M. F. et al., 2006, AJ, 131, 1163
Steeghs D., Greimel R., Drew J., Irwin M., Gaensicke B., Groot P., Knigge C., 2006, The Astronomer's Telegram, 795, 1
Steeghs D., Drew J., Greimel R., Barlow M., Gaensicke B., Drake J., Witham A., 2007, The Astronomer's Telegram, 1031, 1
Stetson P. B., 2000, PASP, 112, 925
Vink J. S., Drew J., Steeghs D., Wright N. J., Martin E. L., Gaensicke B. T., Greimel R., Drake J., 2008, MNRAS, 387, 308
Wareing C. J., O'Brien T. J., Zijlstra A. A., Kwitter K. B., Irwin J., Wright N., Greimel R., Drew J. E., 2006, MNRAS, 366, 387
Witham A. R. et al., 2006, MNRAS, 369, 581
Witham A. R., Knigge C., Drew J. E., Greimel R., Steeghs D., Gänsicke B. T., Groot P. J., Mampaso A., 2008, MNRAS, 384, 1277
Zacharias N., McCallon H. L., Kopan E., Cutri R. M., 2003, IAU Joint Discussion, 16, 43

APPENDIX A: EXAMPLE USAGE ACCESS

Example 1: Return coordinates, magnitudes and magnitude errors of objects in a rectangular area.

This is the most basic query returning all objects in a specific box with RA between $300°$ and $301°$ and Dec. between $30°$ and $31°$.

```
SELECT P.ra, P.dec,
  P.coremag_r, P.coremag_i, P.coremag_ha,
  P.coremagerr_r, P.coremagerr_i, P.coremagerr_ha
FROM PhotoObjBest as P
WHERE P.ra BETWEEN 300.0 AND 301.0
  AND P.dec BETWEEN 30.0 and 31.0
```

Example 2: Return coordinates and magnitudes of objects in a cone defined by its centre at RA = 300.0 and Dec. = 30.0 and radius r = 10 arcmin.

In this case instead of coding the whole trigonometry in the query we use the auxiliary table tGetNearbyObj which returns the object IDs which satisfy a cone search specified by the centre coordinates (ra, dec) and radius (r) in minutes of arc:

```
SELECT P.ra, P.dec,
  P.coremag_r, P.coremag_i, P.coremag_ha,
  P.coremagerr_r, P.coremagerr_i, P.coremagerr_ha
FROM PhotoObjBest as P, tGetNearbyObj as G
WHERE G.ra=300.0 AND G.dec=30.0 AND G.r=10.0
  AND P.objID = G.objID
```

Example 3: As previous example but only return objects classified as point-like in r and observed in conditions of seeing best than 1.5 arcsec as measured in the r band.

In this query we are asking for information about the quality of observations and this is stored in the ChipsAll table. We need therefore to join the results with this table in order to select only observation with good seeing conditions.

```
SELECT P.ra, P.dec,
  P.coremag_r, P.coremag_i, P.coremag_ha,
  P.coremagerr_r, P.coremagerr_i, P.coremagerr_ha
FROM PhotoObjBest as P, tGetNearbyObj as G,
  ChipsAll as CR
WHERE G.ra=300.0 AND G.dec=30.0 AND G.r=10.0
  AND CR.seeing<1.5 AND (P.class_r= −1 OR
P.class_r= −2)
  AND P.objID = G.objID AND
P.chips_r_id=CR.chipID
```

Adding additional constraints on e.g. object colours, magnitude limits, observation dates are trivially implemented in the same way, e.g.:

```
SELECT P.ra, P.dec,
  P.coremag_r, P.coremag_i, P.coremag_ha,
  P.coremagerr_r, P.coremagerr_i, P.coremagerr_ha
FROM PhotoObjBest as P, tGetNearbyObj as G,
  ChipsAll as CR
WHERE G._ra=300.0 AND G._dec=30.0 AND G._r=10.0
  AND P.coremag_r-P.coremag_ha>0.8
  AND P.coremag_r-P.coremag_i>1.5 AND
P.coremagerr_r<0.02
```

Example 4: Select all objects observed in a particular offset field.

The following example returns all objects detected in a particular observation from the (not cleaned) PhotoObj catalogue.

```
SELECT * FROM PhotoObj as P
  JOIN FieldsAll as F ON (F.fieldID = P.field_id)
WHERE F.fieldno=7012 AND F.onoff=0
```

Example 5: Return the number and position of fields observed in the best photometric conditions.

Section 4.2 explains the constraints used to define the catalogue of best observations. The following query returns the number, position and photometric characteristics of the fields which satisfy those conditions.

```
SELECT
  CC.night, F.fieldID, F.fieldno, F.onoff,
  F.ra, F.dec, F.glon, F.glat,
  T.magzpt_mean, T.magzrr_mean,
  CC.seeing, CC.elliptic
FROM
  (SELECT C.night,
sum(C.magzpt)/count(C.magzpt) as magzpt_mean,
sum(C.magzrr)/count(C.magzrr) as magzrr_mean
  FROM FieldsAll F, ChipsAll C
WHERE
    C.field_id = F.fieldID AND C.band='r'
    AND C.chipno=4
  GROUP BY night) AS T,
  ChipsAll CC, FieldsAll F
WHERE
  T.night = CC.night AND CC.field_id = F.fieldID
  AND CC.chipno=4 AND CC.band='r' AND CC.seeing<2
  AND CC.elliptic<0.2 AND T.magzpt_mean>24.2
  AND T.magzrr_mean<0.05
ORDER BY night
```

The IDR web site (http://idr.iphas.org) provides further information as well as examples on accessing the catalogue from the command line.

This paper has been typeset from a TEX/LATEX file prepared by the author.