

# The Second Data Release of the INT Photometric H $\alpha$ Survey of the Northern Galactic Plane (IPHAS DR2)

Geert Barentsen<sup>1\*</sup>, H. J. Farnhill<sup>1</sup>, J. E. Drew<sup>1</sup>, E. A. González-Solares<sup>2</sup>, R. Greimel<sup>3</sup>, M. J. Irwin<sup>2</sup>, B. Mizalski<sup>4</sup>, C. Ruhland<sup>1</sup>, P. Groot<sup>5</sup>, A. Mampaso<sup>6,7</sup>, S. E. Sale<sup>8</sup>, M. J. Barlow<sup>9</sup>, P. J. Carter<sup>13</sup>, R. L. M Corradi<sup>6,7</sup>, J. J. Drake<sup>10</sup>, J. Eislöffel<sup>11</sup>, J. Fabregat<sup>12</sup>, B. T. Gänsicke<sup>13</sup>, N. P. Gentile Fusillo<sup>13</sup>, S. Greiss<sup>13</sup>, A. S. Hales<sup>14</sup>, A. A. Henden<sup>15</sup>, S. Hodgkin<sup>2</sup>, L. Huckvale<sup>16</sup>, J. Irwin<sup>9</sup>, C. Knigge<sup>17</sup>, T. Kupfer<sup>5</sup>, D. J. Lennon<sup>18</sup>, J. R. Lewis<sup>2</sup>, M. Mohr-Smith<sup>1</sup>, R. A. H. Morris<sup>19</sup>, T. Naylor<sup>20</sup>, Q. A. Parker<sup>21,22,23</sup>, S. Phillipps<sup>19</sup>, R. Raddi<sup>13</sup>, P. Rodríguez-Gil<sup>6,7</sup>, L. Sabin<sup>24</sup>, S. Scaringi<sup>25,26</sup>, D. Steeghs<sup>13</sup>, Y. C. Unruh<sup>27</sup>, J. van Roestel<sup>5</sup>, K. Viironen<sup>28</sup>, J. S. Vink<sup>29</sup>, N. A. Walton<sup>2</sup>, N. J. Wright<sup>1</sup>, A. A. Zijlstra<sup>15</sup>.

<sup>1</sup>School of Physics, Astronomy & Mathematics, University of Hertfordshire, College Lane, Hatfield, Hertfordshire, AL10 9AB, U.K.

<sup>2</sup>Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge, CB3 OHA, U.K.

<sup>3</sup>IGAM, Institute of Physics, University of Graz, Universitätsplatz 5, Graz, Austria.

<sup>4</sup>South African Astronomical Observatory, P.O. Box 9, Observatory, 7935 Cape Town, South Africa.

<sup>5</sup>Afdeling Sterrenkunde, Radboud Universiteit Nijmegen, Faculteit NWI, Postbus 9010, 6500 GL Nijmegen, The Netherlands.

<sup>6</sup>Instituto de Astrofísica de Canarias, Vía Láctea, s/n, La Laguna, E-38205, Santa Cruz de Tenerife, Spain.

<sup>7</sup>Departamento de Astrofísica, Universidad de La Laguna, La Laguna, E-38204, Santa Cruz de Tenerife, Spain.

<sup>8</sup>Rudolf Peierls Centre for Theoretical Physics, Keble Road, Oxford, OX1 3NP, U.K.

<sup>9</sup>University College London, Department of Physics & Astronomy, Gower Street, London WC1E 6BT, U.K.

<sup>10</sup>Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, U.S.A.

<sup>11</sup>Thüringer Landessternwarte, Sternwarte 5, D-07778, Tautenburg, Germany.

<sup>12</sup>Observatorio Astronómico, Universidad de Valencia, Catedrático José Beltrán 2, 46980 Paterna, Spain.

<sup>13</sup>Department of Physics, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, U.K.

<sup>14</sup>Joint ALMA Observatory, Alonso de Crdova 3107, Vitacura 763-0355, Santiago, Chile.

<sup>15</sup>AAVSO, 49 Bay State Road, Cambridge, MA 02138, USA.

<sup>16–29</sup> (affiliations 16 through 29 are given in the source file)

Current draft typeset 16 April 2014

## ABSTRACT

The INT/WFC Photometric H $\alpha$  Survey of the Northern Galactic Plane (IPHAS) is a 1800 deg<sup>2</sup> imaging survey covering Galactic latitudes  $|b| < 5^\circ$  and longitudes  $\ell = 30^\circ$  to  $215^\circ$  in the  $r$ ,  $i$  and H $\alpha$  filters using the Wide Field Camera (WFC) on the 2.5-metre Isaac Newton Telescope (INT) in La Palma. We present the first quality-controlled and globally calibrated source catalogue derived from the survey, providing single-epoch photometry for 219 million unique sources across 92 per cent of the footprint. The observations were carried out between 2003 and 2013 at a median seeing of 1.1 arcsec (sampled at 0.33 arcsec/pixel) and to a mean  $5\sigma$ -depth of 21.2 ( $r$ ), 20.0 ( $i$ ) and 20.3 (H $\alpha$ ) in the Vega magnitude system. We explain the data reduction and quality control procedures, describe and test the global re-calibration, and detail the construction of the new catalogue. We show that the new calibration is accurate to 0.03 mag (rms) and recommend a series of quality criteria to select the most reliable data from the catalogue. Finally, we demonstrate the ability of the catalogue's unique ( $r - \text{H}\alpha$ ,  $r - i$ ) diagram to (i) characterise stellar populations and extinction regimes towards different Galactic sight-lines and (ii) select H $\alpha$  emission-line objects. IPHAS is the first survey to offer comprehensive CCD photometry of point sources across the Galactic Plane at visible wavelengths, providing the much-needed counterpart to recent infrared surveys.

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

## 1 INTRODUCTION

The INT/WFC Photometric H $\alpha$  Survey of the Northern Galactic Plane (IPHAS; Drew et al. 2005) is providing new insights into the contents and structure of the disk of the Milky Way. This large-scale programme of observation – spanning almost a decade and using more than 300 nights in competitive open time at the Isaac Newton Telescope (INT) in La Palma – aims to provide the digital update to the photographic northern H $\alpha$  surveys of the mid-20th century (see Kohoutek & Wehmeyer 1999). By increasing the sensitivity with respect to these previous surveys by a factor  $\sim 1000$  (7 magnitudes), IPHAS can expand the limited bright samples of Galactic emission line objects previously available into larger, deeper, and more statistically-robust samples that will better inform our understanding of the early and late stages of stellar evolution. Since the publication of the Initial Data Release (IDR; González-Solares et al. 2008), these aims have begun to be realised through a range of published studies including: a preliminary catalogue of candidate emission-line objects (Witham et al. 2008); discoveries of new symbiotic stars (Corradi et al. 2008, 2010, 2011); new cataclysmic variables (Witham et al. 2007; Wesson et al. 2008; Aungwerojwit et al. 2012); new groups of young stellar objects (Vink et al. 2008; Barentsen et al. 2011; Wright et al. 2012); new classical Be stars (Raddi et al. 2013); along with discoveries of new supernova remnants (Sabin et al. 2013) and new and remarkable planetary nebulae (Mampaso et al. 2006; Viironen et al. 2009a,b; Sabin et al. 2010; Corradi et al. 2011; Viironen et al. 2011).

Over the years it has become apparent that the legacy of IPHAS will reach beyond these traditional H $\alpha$  applications of identifying emission-line stars and nebulae. Through the provision of  $r$  and  $i$  broadband photometry alongside narrowband H $\alpha$  data, IPHAS has created the opportunity to study Galactic Plane populations in a new way. For example, the survey's unique ( $r - \text{H}\alpha$ ,  $r - i$ ) colour-colour diagram has been shown to provide simultaneous constraints on intrinsic stellar colour and interstellar extinction (Drew et al. 2008). This has opened the door to a wide range of Galactic science applications, including the mapping of extinction across the Plane in three dimensions and the probabilistic inference of stellar properties (Sale et al. 2009, 2010; Giannanco et al. 2011; Sale 2012; Barentsen et al. 2013). In effect, the availability of narrowband H $\alpha$  alongside  $r$  and  $i$  magnitudes provides coarse spectral information for huge samples of stars which are otherwise too faint or numerous to be targeted by spectroscopic surveys. For such science applications to succeed, however, it is vital that the imaging data are transformed into a homogeneously-calibrated photometric catalogue, in which quality problems and duplicate detections are flagged.

When the initial data release was created in late 2007, just over half of the survey footprint was covered and the data were insufficiently complete to support a homogeneously calibrated source catalogue. The goal of this paper is to present the next release that takes the coverage up to 92 percent of the survey area and includes a global photometric re-calibration. In this work we (i) explain the data reduction and quality control procedures that were applied, (ii) describe and test the new photometric calibration,

and (iii) detail the construction of the source catalogue and demonstrate its use.

In §2 we start by recapitulating the key points of the survey observing strategy. In §3 we describe the data reduction and quality control procedures. In §4 we explain the global re-calibration, in which we draw upon the AAVSO Photometric All-Sky Survey (APASS) and test our results against the Sloan Digital Sky Survey (SDSS). In §5 we explain how the source catalogue was compiled. In §6 we discuss the properties of the catalogue and in §7 we demonstrate the scientific exploitation of the colour/magnitude diagrams. Finally, in §8 we discuss access to the catalogue, an online library of reduced images and relevant source code. The paper ends with conclusions in §9 where we also outline our future ambitions.

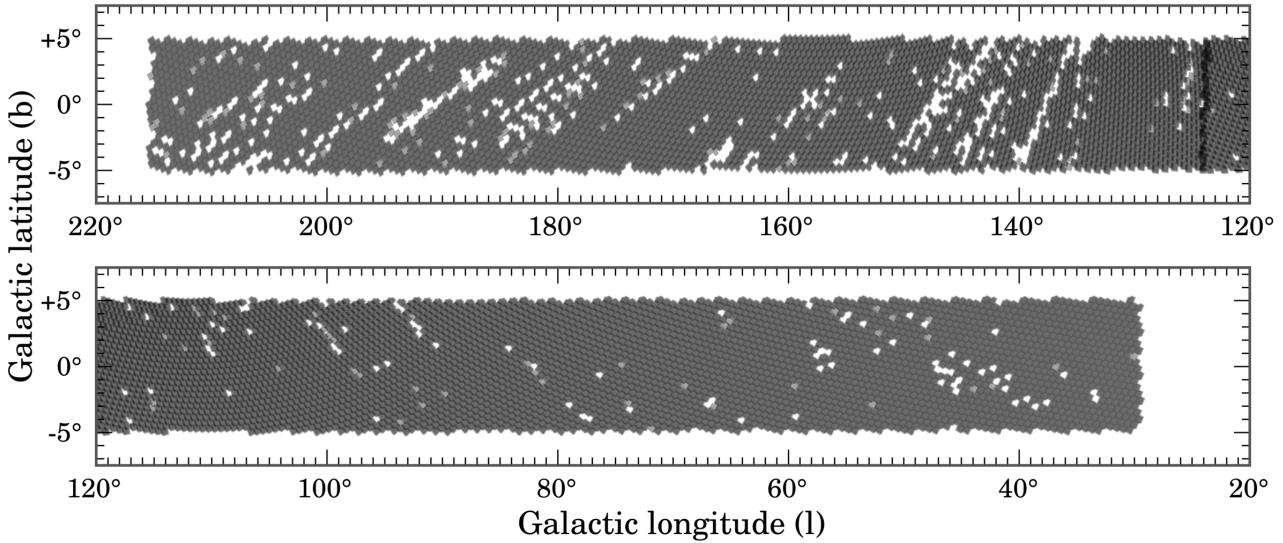
## 2 OBSERVATIONS

The detailed properties of the IPHAS observing programme have been presented before by Drew et al. (2005) and González-Solares et al. (2008). To set the stage for this release, we recap some key points. IPHAS is an imaging survey of the Galactic Plane north of the celestial equator, from which photometry in Sloan  $r$  and  $i$  is extracted along with narrowband H $\alpha$ . It is carried out using the Wide Field Camera (WFC) on the 2.5-metre INT in La Palma. It is the first digital survey to offer comprehensive optical CCD photometry of point sources in the Galactic Plane; the footprint spans a box of roughly 180 by 10 degrees, covering Galactic latitudes  $-5^\circ < b < +5^\circ$  and longitudes  $30^\circ < \ell < 215^\circ$ .

The WFC is a mosaic of 4 CCDs that captures a sky area of close to  $0.29 \text{ deg}^2$ . To cover the Northern Plane with some overlap, the survey area was divided into 7635 telescope pointings. Each of these pointings is accompanied by an offset position displaced by +5 arcmin in declination and +5 arcmin in right ascension, to deal with inter-CCD gaps, detector imperfections, and to enable quality checks. Hence, the basic unit of observation hence amounts to  $2 \times 3$  exposures, in which each of the 3 survey filters is exposed at 2 offset sky positions within, typically, an elapsed time of 10 minutes. We shall refer to the unit of 3 exposures at the same position as a *field*, and the combination of two fields at a small offset as a *field pair*. Altogether, the survey contains 15 270 fields grouped into 7 635 field pairs. To achieve the desired survey depth of 20th magnitude or fainter in each filter, the exposure times were set at 120 s (H $\alpha$ ), 30 s ( $r$ ) and 10 s ( $i$ ) in the vast majority of the survey observations.<sup>1</sup>

Data-taking began in the second half of 2003, and every field had been observed at least once by the end of 2008. At that time only 76 per cent of the field pairs satisfied our minimum quality criteria, however. The problems affecting the 24 per cent falling below survey standard were, most commonly: variable cloud cover; poor seeing; technical faults (the quality criteria will be detailed in the next section). Since then, a programme of repeat observations has been in place to improve data quality. As a result, 92 per cent of the

<sup>1</sup> The  $r$ -band exposure time was 10 s instead of 30 s in the first months of data taking. Since October 2010, the  $i$ -band exposure time has been increased from 10 s to 20 s to bypass a sporadic exposure timing bug that affects the WFC.



**Figure 1.** Survey area showing the footprints of all the quality-approved IPHAS fields which have been included in this data release. The area covered by each field has been coloured black with a semi-transparent opacity of 20 per cent, such that regions where fields overlap are darker. The IPHAS strategy is to observe each field twice with a small offset, and hence the vast majority of the area is covered twice (dominant grey colour). There are small overlaps between all the neighbouring fields which can be seen as a honeycomb pattern of dark grey lines across the survey area. Regions with incomplete data are apparent as white gaps (no data) or in light grey (denoting that only one offset is missing). The dark vertical strip near  $\ell \sim 125^\circ$  is an arbitrary consequence of the tiling pattern, which was populated starting from 0h in right ascension.

survey footprint now benefits from quality-approved data. The most recent observations included in this release were obtained in November 2012.

Fig. 1 shows the footprint of the quality-approved observations included in this work. The fields which remain missing – covering 8 per cent of the survey area – are predominantly located towards the Galactic anti-centre at  $\ell > 120^\circ$ . Fields at these longitudes are mainly accessed from La Palma in the months of November–December, which is when the La Palma weather and seeing conditions are often poor forcing many (unsuccessful) repeat observations. To enable the survey to be brought to completion, a decision was made recently to limit repeats in this area to individual fields requiring replacement, i.e. fresh observations in all 3 filters may only be obtained at one of the two offset positions. The catalogue is structured such that it is clear where a contemporaneous observation of both halves of a field pair is available.

### 3 DATA REDUCTION AND QUALITY CONTROL

#### 3.1 Initial pipeline processing

All raw IPHAS data were transferred to the Cambridge Astronomical Survey Unit (CASU) for initial processing and archival. The procedures used by CASU were originally devised for the INT Wide Field imaging Survey (WFS; McMahon et al. 2001; Irwin et al. 2005), which was a  $200 \text{ deg}^2$  survey programme carried out between 1998 and 2003 after the WFC was commissioned. Because IPHAS uses the same telescope and camera combination, we have been able to benefit from the existing WFS pipeline. A detailed de-

scription of the processing steps is found in Irwin & Lewis (2001). Its application to IPHAS has previously been described by Drew et al. (2005) and González-Solares et al. (2008), and the source code is available on-line<sup>2</sup>. In brief, the pipeline takes care of bias subtraction, linearity correction, flat-fielding, gain correction and de-fringing.

The reduced images are then stored in multi-extension FITS files with a primary header describing the characteristics (position, filter, exposure time, etc.) and four image extensions corresponding to each of the four CCDs. Source detection and characterisation is then carried out using the IMCORE tool (Irwin 1985, 1997). The flux of each source is measured using both the peak pixel height (i.e. a square  $0''.33 \times 0''.33$  aperture) and a series of circular apertures of increasing diameter ( $1''.2$ ,  $2''.3$ ,  $3''.3$ ,  $4''.6$  and  $6''.6$ ).

The local background levels are estimated by computing the sigma-clipped median flux in a grid of  $64 \times 64$  pixels ( $21'' \times 21''$ ) across the image, which is then interpolated to obtain an estimate of the background level at each pixel. These sky levels are subtracted from the aperture photometry and – when required – a deblending routine is applied to remove the contamination from any nearby sources. This approach works very well across the vast majority of the survey area. Nevertheless, the Galactic Plane contains crowded regions with large numbers of overlapping sources or rapidly spatially-varying nebulosity, where aperture photometry will be compromised by frequent blending or poor background subtraction. In §5 we will explain how affected sources are flagged.

Finally, an astrometric solution is determined based on

<sup>2</sup> <http://casu.ast.cam.ac.uk/surveys-projects/software-release>

the Two-Micron All Sky Survey (2MASS) point source catalogue (Skrutskie et al. 2006), which itself is calibrated in the International Celestial Reference System (ICRS). A provisional photometric calibration is also provided based on the average zeropoint determined from sets of standard stars observed within the same night. Sources are classified morphologically – stellar, extended or noise – based on the curve-of-growth determined from measuring the source intensity in a series of growing apertures. Finally, the resulting source detection tables are stored in multi-extension FITS files.

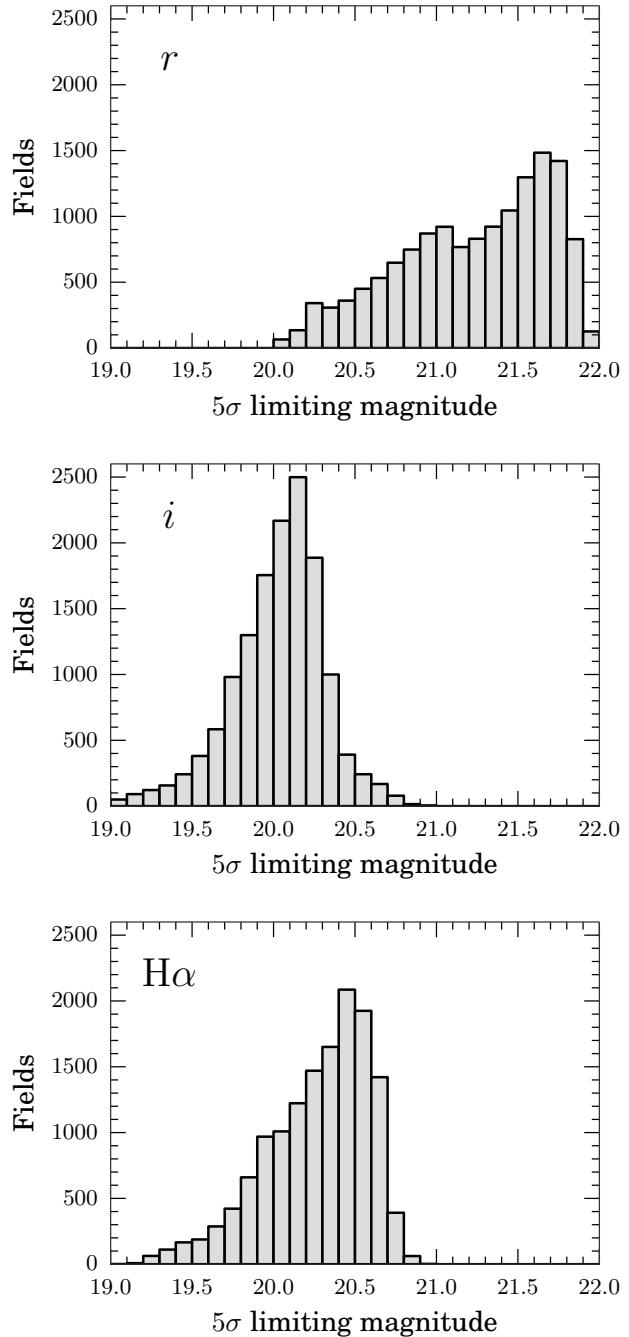
At the time of preparing DR2, the CASU pipeline had processed 74 195 single-band IPHAS exposures in which a total of 1.9 billion *candidate sources* were detected at the default detection level of  $1.25\sigma$ . This total inevitably includes spurious detections, artefacts and duplicate detections; in §5 we will explain how these have been removed or flagged in the final catalogue. The pipelined data set – comprising 2.5 terabyte of FITS files – was then transferred to the University of Hertfordshire for the purpose of transforming the raw detection tables into a reliable source catalogue which is (i) quality-controlled, (ii) homogeneously calibrated, and (iii) contains user-friendly columns and warning flags. It is these post-processing steps which are explained next.

### 3.2 Quality control

Observing time for IPHAS is obtained on a semester-by-semester basis through the open time allocation committees of the Isaac Newton Group of telescopes. The survey is allocated specific observing dates rather than particular observing conditions. In consequence, data were acquired under a large range of atmospheric conditions. Data taken under unsuitable conditions have been rejected using seven quality criteria, which ensure a reliable and homogeneous level of quality across the data release:

(1) *Depth*. We discarded any exposures for which the  $5\sigma$  limiting magnitude was brighter than 20th magnitude in the  $r$ -band or brighter than 19th in  $i$  or  $H\alpha$ . Such data were typically obtained during poor weather or full moon. Most observations were significantly better than these limits. Fig. 2 presents the distribution of limiting magnitudes for all quality-approved fields; the mean depths and standard deviations are  $21.2 \pm 0.5$  ( $r$ ),  $20.0 \pm 0.3$  ( $i$ ) and  $20.3 \pm 0.3$  ( $H\alpha$ ). The depth achieved depended most strongly on the presence of the moon, which was above the horizon during 62 per cent of the observations. The great range in sky brightness this produced is the main reason for the wide and bi-modal shape of the  $r$ -band limiting magnitude distribution (top panel in Fig. 2). In contrast, the depths attained in  $i$  and  $H\alpha$  are less sensitive to moonlight, leading to narrower magnitude limit distributions (middle and bottom panels in Fig. 2). To a lesser extent, the wide spread in the  $r$ -band depth is also explained by the shorter  $r$ -band exposure time that was used in the first year of data-taking.

(2) *Ellipticity*. The ellipticity of a point source, defined as  $e = 1 - b/a$  with  $b$  the semi-minor and  $a$  the semi-major axis, is a morphological measure of the elongation of the point spread function (PSF). It is expected to be zero (circular) across the field in a perfect imaging system, but it is slightly non-zero in any real telescope data due to optical distortions and tracking errors. Indeed, it is worth noting that IPHAS data have been collected from unguided expo-



**Figure 2.** Distribution of the  $5\sigma$  limiting magnitude across all quality-approved survey fields for  $r$  (top),  $i$  (middle) and  $H\alpha$  (bottom). Fields with a limiting magnitude brighter than 20th ( $r$ ) or 19th ( $H\alpha$ ,  $i$ ) were rejected from the data release. The  $r$ -band depth is most sensitive to the presence of the moon above the horizon: this is the main reason for the wide, bi-modal character of its distribution.

sures that rely entirely on the INT's tracking capability. The mean ellipticity measured in the data is  $0.09 \pm 0.04$ . There have been sporadic episodes with higher ellipticities due to mechanical glitches in the telescope tracking system. To exclude these, we rejected exposures in which the mean ellipticity *across the detectors* exceeded  $e > 0.3$ . We have also

found that, as this threshold is breached, the photometric measurements delivered by the pipeline become degraded.

(3) *Seeing*. The original survey goal was to obtain data at a resolution better than 1.7 arcsec, as evaluated by measuring the average PSF Full Width at Half Maximum (FWHM) across the detectors, which is a proxy for the combined effect of atmospheric and dome seeing. This target is currently attained across 86 per cent of the footprint. To increase the sky area offered by the data release slightly, we have decided to accept data obtained with seeing up to 2.5 arcsec. Fig. 3 presents the distribution of the PSF FWHM for the approved fields. In the  $r$ -band, 90 per cent is better than 1.5 arcsec, 50 per cent is better than 1.1 arcsec, and 10 per cent is better than 0.8 arcsec (the WFC pixel scale is 0.33 arcsec/px). In §5 we will explain that the photometry compiled in the source catalogue is normally derived from the field with the best-available seeing for a given object, and that the PSF FWHM measurement is included as a column in the catalogue.

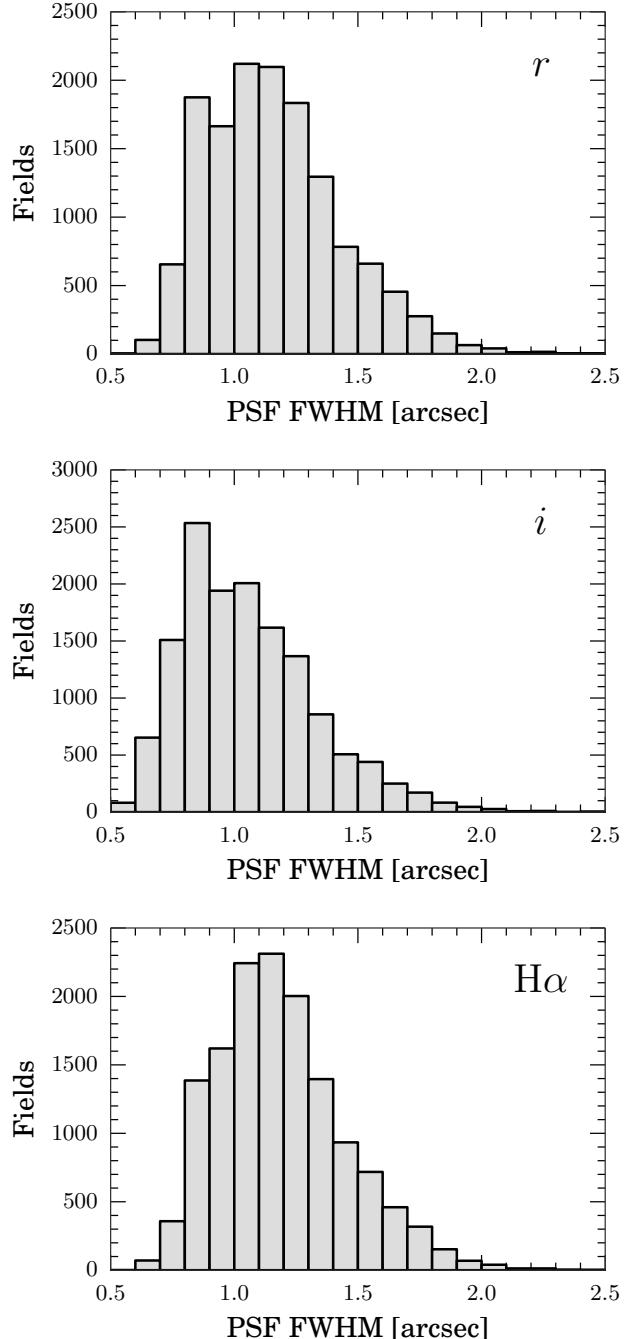
(4) *Photometric repeatability*. The IPHAS field-pair observing strategy normally ensures that every pointing is immediately followed by an offset pointing at a displacement of +5 arcmin in Dec and +5 arcmin in RA. This allows pairs of images to be checked for the presence of clouds or electronic noise. To exploit this information, the overlap regions of all field pairs were systematically cross-matched to verify the consistency of the photometry for stars observed in both pointings. We rejected field pairs in which more than 2 per cent of the stars showed an inconsistent measurement at the level of 0.2 mag, or more than 26 per cent were inconsistent at the level of 0.1 mag. These limits were set empirically after inspecting the images and photometry by eye.

(5) *Source density mapping*. Spatial maps showing the number density of the detected sources down to 20th magnitude were created to verify the health of the data and to check for unexpected artefacts. In particular, we created density maps which showed the number of *unique* sources obtained by cross-matching the detection tables of all three bands with a maximum matching distance of 1 arcsec. This was particularly effective for revealing fields with an inaccurate astrometric solution in one of the bands, which were subsequently corrected.

(6) *Visual examination*. Images, colour mosaics, and the associated photometric colour/magnitude diagrams were inspected by a team of 20 survey members, such that each image in the data release was looked at by at least three different pairs of eyes. Images affected by clouds or extreme levels of scattered moonlight were flagged, investigated, and excluded from the release by placing them on a black-list.

(7) *Contemporaneous field data*. Finally, only exposures which are part of a sequence of three consecutive images of the same field ( $\text{H}\alpha$ ,  $r$ ,  $i$ ) were considered for inclusion in the release. This ensures that the three bands for a given field are observed contemporaneously – normally within 5 minutes of each other. We note that the source catalogue details the exact epoch at the start of each exposure (columns  $r\text{MJD}$ ,  $i\text{MJD}$ ,  $\text{hMJD}$ ).

All of the above criteria were satisfied by at least one observing attempt for 14 115 out of the 15 270 fields (92 per cent). In some cases more than one successful attempt to observe a field was available due to stricter quality criteria being applied in the initial years of the survey. In such cases,



**Figure 3.** Distribution of the PSF FWHM for all the quality-approved fields included in the release, measured in  $r$  (top),  $i$  (middle) and  $\text{H}\alpha$  (bottom). The PSF FWHM measures the effective image resolution that arises from the combination of atmospheric and dome seeing.

only the attempt with the best seeing and depth has been selected for inclusion in the catalogue, in order to deliver the most reliable measurement at a single epoch.

We note that some of the excluded data may nevertheless be useful for e.g. time-domain studies of bright stars. The images and detection tables of the discarded data are made available through our website for this rea-

son ([www.iphas.org](http://www.iphas.org)), but will be ignored in the remainder of this work.

## 4 PHOTOMETRIC CALIBRATION

Having obtained a quality-approved set of observations, we now turn to the challenge of placing the data onto a uniform photometric scale.

### 4.1 Provisional nightly calibration

For the purpose of providing an initial calibration of the  $r$  and  $i$  broadband fluxes, photometric standard fields were observed every night. The standards were chosen from a list based on the Landolt (1992) and Stetson (<http://cadcwww.dao.nrc.ca/standards>) objects. Two or three standard fields were observed during the evening and morning twilight, and at intervals of 2–3 hours throughout the night. The CASU pipeline automatically identified the observed standards and used them to determine a sigma-clipped average zeropoint MAGZPT for each night and filter, such that the number counts  $DN$  in the pipeline-corrected CCD frames relate to a magnitude  $m$  as:

$$m = \text{MAGZPT} - 2.5 \log_{10}(DN/\text{EXPTIME}) - \text{EXTINCT} \cdot (\text{AIRMASS} - 1) - \text{APCOR} - \text{PERCORR}, \quad (1)$$

where EXPTIME is the exposure time in seconds, EXTINCT is the atmospheric extinction coefficient (set in the pipeline at 0.09 for  $r$  and 0.05 for  $i$  as representative averages for La Palma), AIRMASS is the normalised optical path length through the atmosphere and APCOR is a correction for the flux lost outside of the aperture used (we adopt a 2''.3-diameter circular aperture by default). Finally, PERCORR is a correction based on the difference between the median dark sky for a CCD against the median for all the CCDs, and as such is an ancillary correction to account for sporadic gain variations; it is zero for most exposures. All these quantities correspond to header keywords in the multi-extension FITS files produced by the CASU pipeline.

The broadband zeropoints were determined such that the resulting magnitude system refers to the spectral energy distribution (SED) of Vega as the zero colour object. Colour equations were used to transform between the IPHAS passbands and the Johnson-Cousins system of the published standard star photometry. The entire procedure has been found to deliver zeropoints which are accurate at the level of 1–2 per cent in photometric conditions (González-Solares et al. 2011).

Unlike the broadbands, standard-star photometry is not available in the literature for the  $\text{H}\alpha$  passband and hence there is no formally recognised flux scale for it. We can specify here, however, that the detected flux for Vega in the IPHAS  $\text{H}\alpha$  filter is 3.14 magnitudes less than the flux captured by the much broader  $r$  band which includes the  $\text{H}\alpha$  band. Hence to assure zero colour relative to the broadbands, we set the default zeropoint for the narrowband to be:

$$\text{MAGZPT}_{\text{H}\alpha} = \text{MAGZPT}_r - 3.14. \quad (2)$$

For reference, Table 1 details the flux of Vega in the IPHAS

**Table 1.** Mean monochromatic flux of Vega in the IPHAS filter system, defined as  $\langle f_\lambda \rangle = \int f_\lambda(\lambda)S(\lambda)\lambda d\lambda / \int S(\lambda)\lambda d\lambda$ , where  $S(\lambda)$  is the photon response function (which includes atmospheric transmission, filter transmission and CCD response) and  $f_\lambda(\lambda)$  is the model SED for Vega by Kurucz (<http://kurucz.harvard.edu/stars.html>). For reference, we also provide the filter equivalent width  $\text{EW} = \int S(\lambda)d\lambda / \int S(\lambda)d\lambda$ , the mean photon wavelength  $\lambda_0 = \int S(\lambda)\lambda d\lambda / \int S(\lambda)d\lambda$ , and the pivot wavelength  $\lambda_p = \sqrt{\int S(\lambda)\lambda d\lambda / \int \frac{S(\lambda)}{\lambda} d\lambda}$ . These notations follow the definitions by Bessell & Murphy (2012). After multiplying  $\langle f_\lambda \rangle$  by the EW, we find that the detected flux for Vega in  $\text{H}\alpha$  is 3.14 magnitudes less than that received in  $r$ ,

Filter	$\langle f_\lambda \rangle$ [erg cm $^{-2}$ s $^{-1}$ Å $^{-1}$ ]	EW [Å]	$\lambda_0$ [Å]	$\lambda_p$ [Å]
$r$	$2.47 \times 10^{-9}$	785.6	6223	6211
$\text{H}\alpha$	$1.81 \times 10^{-9}$	59.6	6568	6568
$i$	$1.30 \times 10^{-9}$	759.9	7674	7674

filter system. We refer to Drew et al. (2005) for a more comprehensive discussion on the throughput curves of the bands.

### 4.2 Global re-calibration

Despite the best efforts made to obtain a nightly calibration, large surveys naturally possess field-to-field variations due to atmospheric changes during the night and imperfections in the pipeline or the instrument (e.g. the WFC is known to suffer from sporadic errors in the timing of exposures). Such variations need to be corrected for during a global re-calibration procedure. Notable past examples include the global re-calibration of 2MASS (Nikolaev et al. 2000), SDSS (Padmanabhan et al. 2008) and the Panoramic Survey Telescope And Rapid Response System survey (Pan-STARRS; Schlafly et al. 2012), which all achieved photometry that is globally consistent to within 0.01–0.02 mag.

Surveys which observe identical stars at different epochs can use the repeat measurements to ensure a homogeneous calibration. For example, 2MASS attained its global calibration by observing six standard fields each hour, allowing zeropoint variations to be tracked over very short timescales (Nikolaev et al. 2000). Alternatively, the SDSS and PanSTARRS surveys could benefit from revisiting regions in their footprint to carry out a so-called *ubercalibration*<sup>3</sup> procedure, in which multiple measurements of stars at different epochs are used to fit the calibration parameters (e.g. Ivezić et al. 2007; Padmanabhan et al. 2008; Schlafly et al. 2012).

Unfortunately these schemes cannot be applied directly to IPHAS for two reasons. First, the survey was carried out in competitively-allocated observing time on a common-user telescope, rendering the 2MASS approach of observing standards at a very high frequency prohibitively expensive (it does not help that standard fields are very scarce within the

<sup>3</sup> ‘Ubercalibration’ refers to the name of the code used to re-calibrate SDSS photometry using repeat observations. It is an anglicised version of the German word ‘überkalibration’, which was reportedly chosen because the initial authors both had German-sounding names (Finkbeiner 2010).

Galactic Plane). Second, IPHAS is not specified as a variability survey, with the consequence that stars are not normally observed at more than one epoch, unless they happen to fall within a narrow overlap region between two neighbouring field pairs.

We have found the information contained in these overlap regions to be insufficient to constrain the calibration parameters. This is because photometry at the extreme edges of the WFC – where neighbouring fields overlap – is prone to systematics at the level of 1–2 per cent. The cause of these errors is thought to include the use of twilight sky flats in the pipeline, which are known to be imperfect for calibrating stellar photometry due to stray light and vignetting (e.g. Manfroid 1995). Moreover, the illumination correction in the overlap regions is more affected by a radial geometric distortion in the WFC, which causes the pixel scale to increase as the edges are approached (González-Solares et al. 2011). Although these systematics are reasonably small within a single field, they can combine, during the global calibration routine, to cause artificial zeropoint gradients across the survey unless controlled by other external constraints.

For these reasons, we have decided not to depend on an ubercalibration-type scheme alone, and have opted to involve an external reference survey – where available – to bring the majority of our data onto a homogeneous calibration.

#### 4.2.1 Correcting zeropoints using APASS

We have been able to benefit from the AAVSO Photometric All-Sky Survey (APASS; <http://www.aavso.org/apass>) to bring most of the survey onto a uniform scale. Since 2009, APASS has been using two 20 cm-astrographs to survey the entire sky down to  $\sim$ 17th magnitude in five filters which include Sloan  $r$  and  $i$  (Henden et al. 2012). The most recent catalogue available at the time of preparing this work was APASS DR7, which provides a good coverage across  $\sim$ half of the IPHAS footprint. The overlap regions are shown in Fig. 4. The photometric accuracy of APASS is currently estimated to be at the level of 3 per cent (Henden, private communication), which is significantly better than the original nightly calibrations of IPHAS which are accurate to  $\sim$ 10 per cent (Tab. 2). APASS achieves its uniform accuracy by measuring each star at least two times in photometric conditions, along with ample standard fields, benefiting from the large 3-by-3 degrees field of view of its detectors.

With the aim of bringing IPHAS to a similar accuracy of 3 per cent, we used the APASS catalogue to identify and adjust the calibration of all IPHAS fields which showed a magnitude offset larger than 0.03 mag against APASS. Experience of re-running the calibration and testing the results showed us that it was inadvisable to more finely tune the match for IPHAS data obtained in what were generally the most photometric nights. To this end, the  $r$ - and  $i$ -band detection tables of each IPHAS field were cross-matched against the APASS DR7 catalogue using a maximum matching distance of 1 arcsec. The magnitude range was limited to  $13 < r_{\text{APASS}} < 16.5$  and  $12.5 < i_{\text{APASS}} < 16.0$  in order to avoid sources brighter than the IPHAS saturation limit on one hand, and to avoid sources near the faint detection limit of APASS on the other.

The resulting set of  $\sim$ 220 000 cross-matched stars were

then used to derive APASS-to-IPHAS magnitude transformations using a linear least-squares fitting routine, which iteratively removed  $3\sigma$ -outliers to improve the fit. The solution converged to:

$$r_{\text{IPHAS}} = r_{\text{APASS}} - 0.121 + 0.032(r - i)_{\text{APASS}} \quad (3)$$

$$i_{\text{IPHAS}} = i_{\text{APASS}} - 0.364 + 0.006(r - i)_{\text{APASS}} \quad (4)$$

The root mean square (rms) residuals of these transformations are 0.041 and 0.051, respectively. The small colour terms in the equations indicate that the IPHAS and APASS broadband filters are very similar. The transformations include a large fixed offset, but this is simply due to the fact that APASS magnitudes are given in the AB system and IPHAS uses Vega-based magnitudes. Separate transformations were derived for sight-lines with varying extinction properties to investigate the robustness of the transformations with respect to different reddening regimes. The variations at these different sight-lines were found to be insignificant. This is not surprising because heavily-reddened objects are naturally scarce at  $r < 16$ .

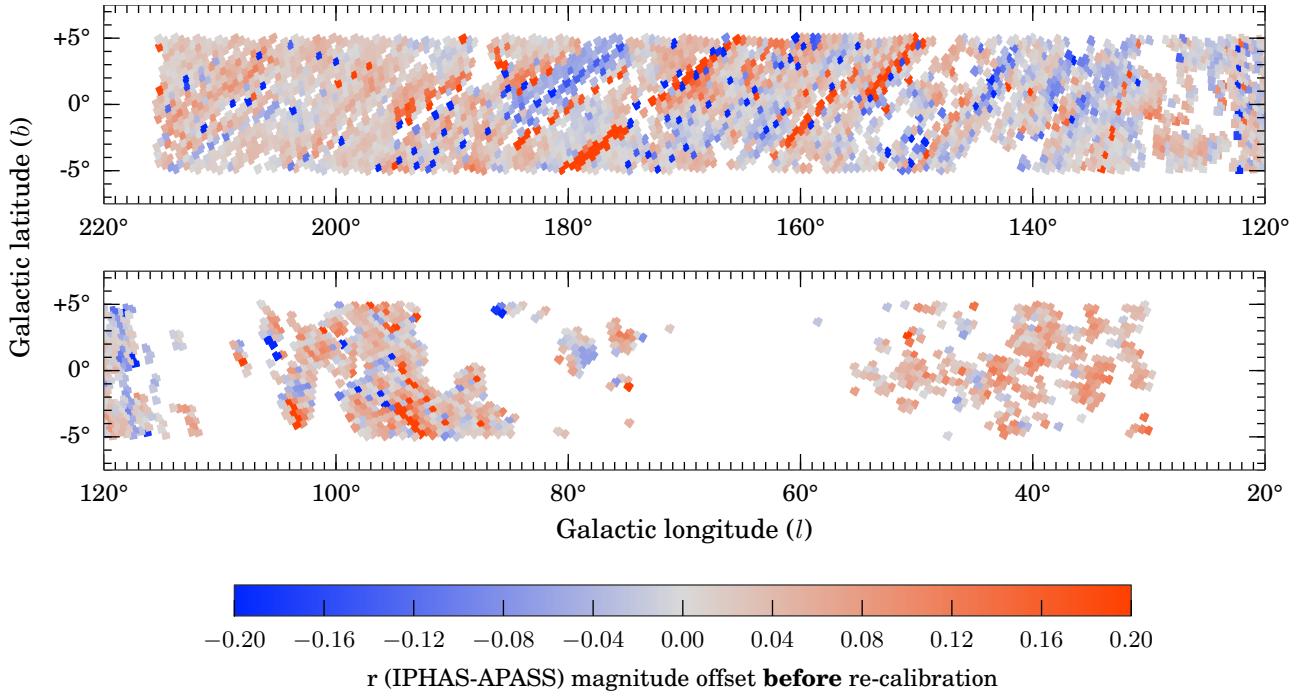
Having transformed APASS magnitudes into the IPHAS system, we then computed the median magnitude offset for each field which contained at least 30 cross-matched stars. This was achieved for 48 per cent of our fields (shown in Figs. 4 and 5). The offsets follow a near-Gaussian distribution centred around  $0.014 \pm 0.104$  mag in  $r$  and  $0.007 \pm 0.108$  mag in  $i$  (Table 2). A total of 4596 fields showed a systematic offset exceeding  $\pm 0.03$  mag in either  $r$  or  $i$ .

We then applied the most important step in our calibration scheme, which is to adjust the provisional zeropoints of these 4596 aberrant fields such that their offset is brought to zero. This allowed the mean IPHAS-to-APASS offset to be brought down to  $0.000 \pm 0.011$  mag in both  $r$  and  $i$  (Table 3). The procedure of fitting magnitude transformations and correcting the IPHAS zeropoints was repeated a few times to ensure convergence, which was essentially reached after the first iteration.

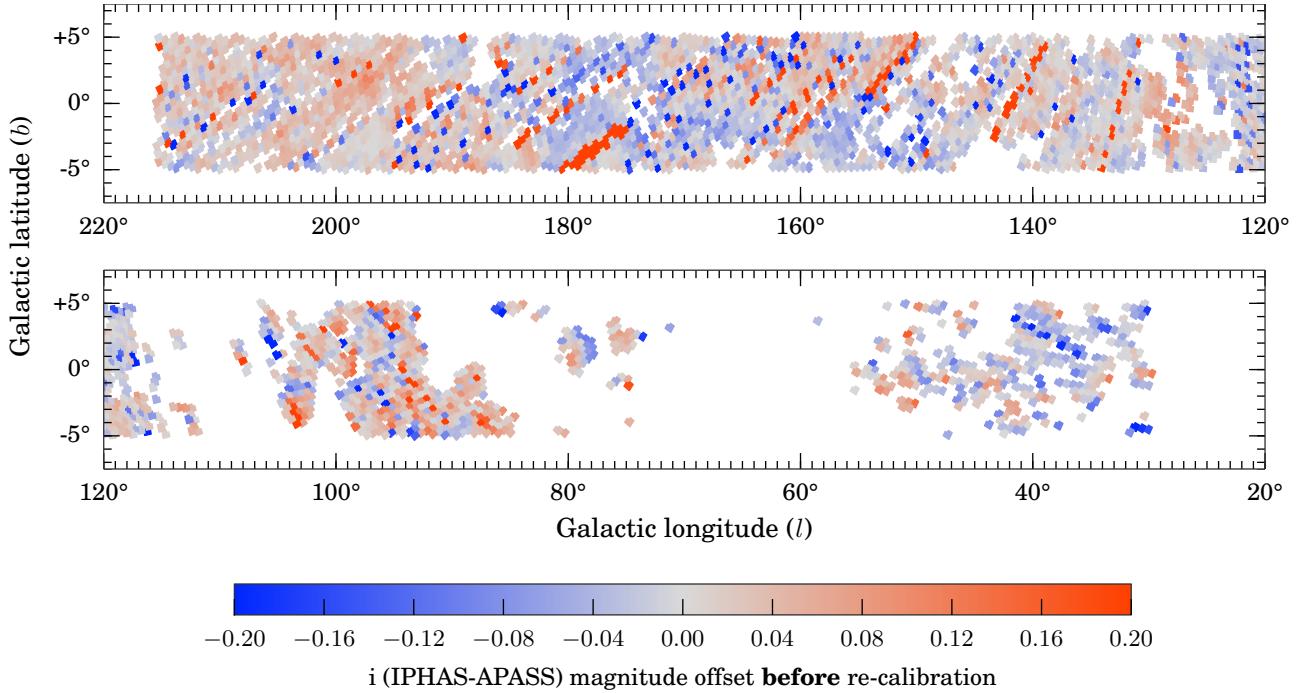
#### 4.2.2 Adjusting fields not covered by APASS

At the time of writing, the APASS catalogue did not provide sufficient coverage for 7359 of our fields. Fortunately, these fields are located mainly at low Galactic longitudes (cf. Figs. 4 and 5), which were typically observed during the summer months when photometric conditions are more prevalent at the telescope. These remaining fields have nevertheless been brought onto the same uniform scale by employing an ubercalibration-style scheme, which minimises the magnitude offsets between stars located in the overlap regions between neighbouring fields.

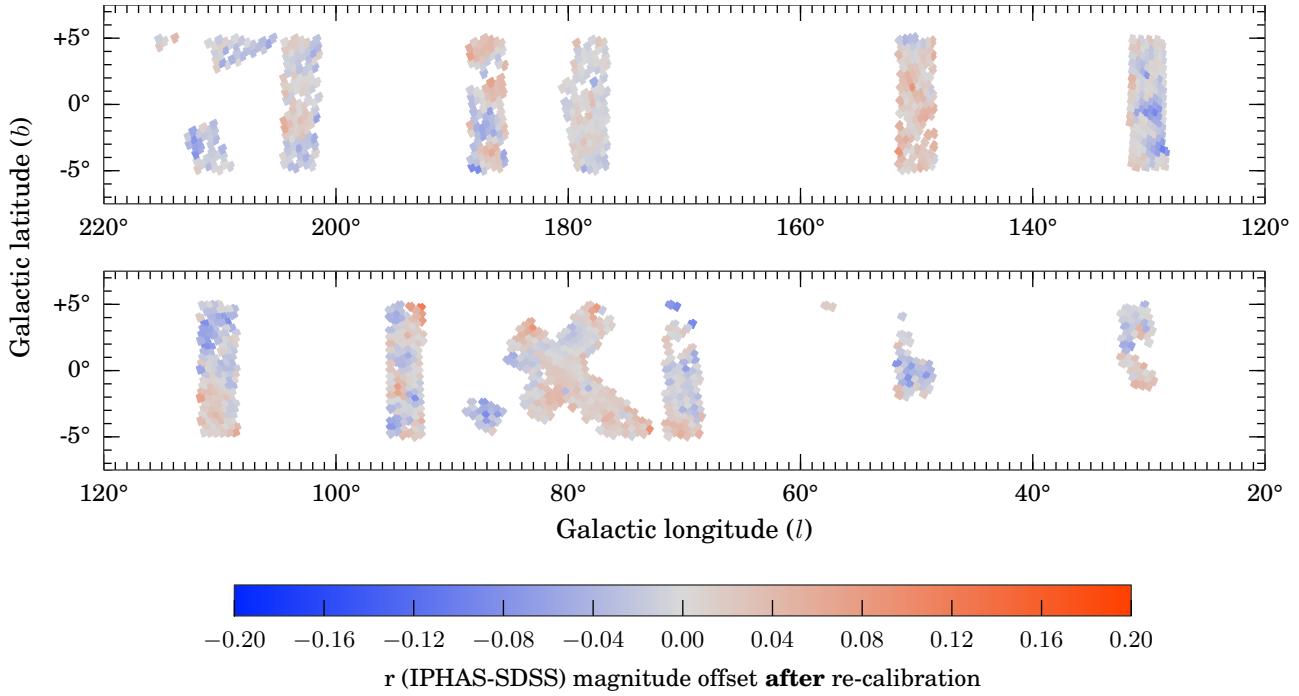
An algorithm for achieving this minimisation has previously been described by Glazebrook et al. (1994). In brief, there are two fundamental quantities to be minimised between each pair of overlapping exposures, denoted by the indices  $i$  and  $j$ . Firstly, the mean magnitude difference between stars in the overlap region  $\Delta_{ij} = \langle m_i - m_j \rangle = -\Delta_{ji}$  is a local constraint. Secondly, to ensure the solution does not stray far from the existing calibration, the difference in zeropoints  $\Delta ZP_{ij} = -\Delta ZP_{ji}$  between each pair of exposures must also be minimised.



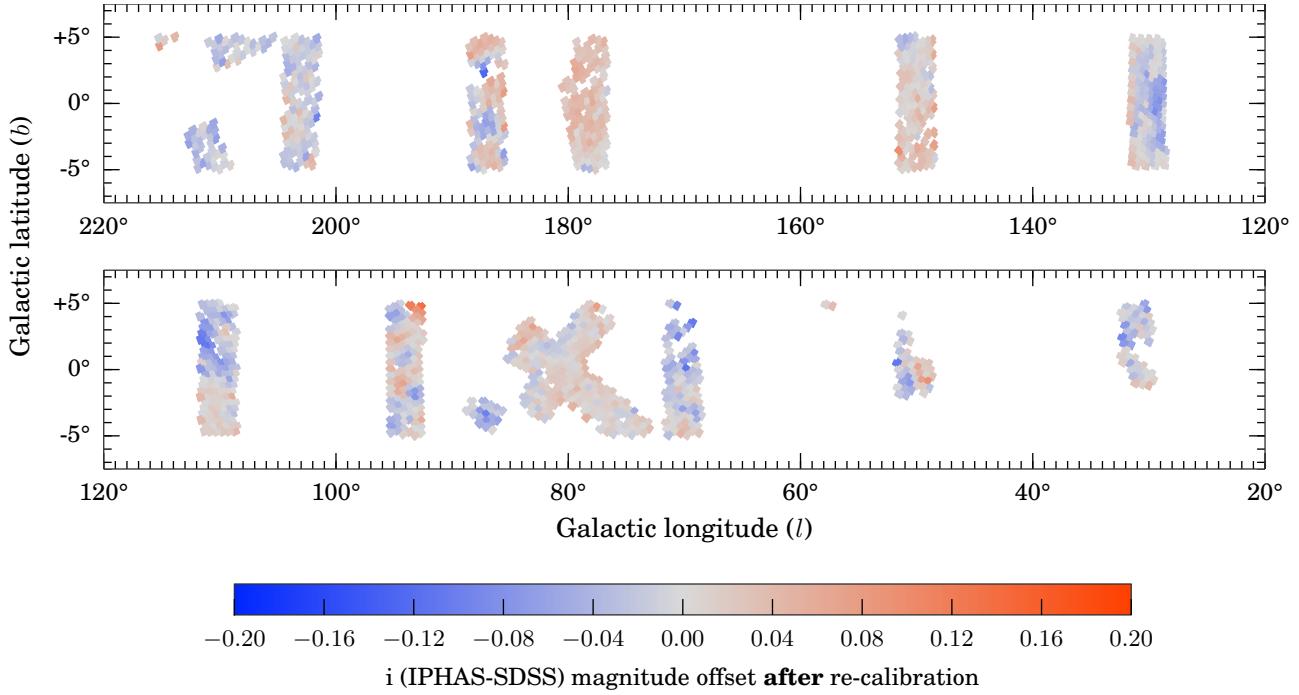
**Figure 4.** Median magnitude offset in the  $r$  band between IPHAS and APASS, plotted on a field-by-field basis prior to the re-calibration procedure. Each square represents the footprint of an IPHAS field which contains at least 30 stars with a counterpart in the APASS DR7 catalogue. The colours denote the median IPHAS-APASS magnitude offset in each field, which was computed after applying the APASS-to-IPHAS transformation to the APASS magnitudes (Eqn. 3). For clarity, we do not show the fields at the offset positions.



**Figure 5.** Same as Fig. 4 for the  $i$ -band.



**Figure 6.** Median magnitude offset between IPHAS and SDSS in the  $r$  band *after* the re-calibration procedure using APASS was applied. Each square represents the footprint of an IPHAS field which contains at least 30 stars with a counterpart in the SDSS DR9 catalogue. The colours denote the median IPHAS-SDSS magnitude offset in each field, which was computed after applying the SDSS-to-IPHAS transformation to the SDSS magnitudes (Eqn. 9).



**Figure 7.** Same as Fig. 6 for the  $i$ -band.

**Table 2.** Magnitude offsets for objects cross-matched between IPHAS and APASS/SDSS *before* the global re-calibration was carried out. We characterise the distribution of the offsets, which is approximately Gaussian in each case, by listing the mean and the standard deviation values. We remind the reader that transformations were applied to the APASS and SDSS magnitudes to bring them into the Vega-based IPHAS system prior to computing the offsets.

Before re-calibration	Mean	$\sigma$
<i>r</i> (IPHAS - APASS)	+0.014	0.104
<i>i</i> (IPHAS - APASS)	+0.007	0.108
<i>r</i> (IPHAS - SDSS)	+0.016	0.088
<i>i</i> (IPHAS - SDSS)	+0.010	0.089

**Table 3.** Same as Table 2, but computed *after* the global re-calibration was carried out. The mean and standard deviation values of the offsets have improved significantly.

After re-calibration	Mean	$\sigma$
<i>r</i> (IPHAS - APASS)	+0.000	0.011
<i>i</i> (IPHAS - APASS)	+0.000	0.011
<i>r</i> (IPHAS - SDSS)	-0.001	0.029
<i>i</i> (IPHAS - SDSS)	-0.002	0.032

Minimisation of these two quantities is a linear least squares problem because the magnitude  $m$  depends linearly on the ZP (Eqn. 11). Hence we can find the ZP shift to be applied to each field by minimising the sum:

$$S = \sum_{i=1}^N \sum_{j=1}^N w_{ij} \theta_{ij} (\Delta_{ij} + a_i - a_j)^2, \quad (5)$$

where  $i$  denotes an exposure,  $j$  an overlapping exposure,  $N$  the number of exposures,  $a_i$  the ZP to solve for and  $a_j$  the ZP of an overlapping field ( $\Delta ZP_{ij} = a_i - a_j$ ).  $w_{ij}$  are weights set to the inverse square of the uncertainty in  $\Delta_{ij}$ , and  $\theta_{ij}$  is an overlap function equal to either 1 if exposures  $i$  and  $j$  overlap or 0 otherwise. Solving for  $a_i$  is equivalent to solving  $\partial S / \partial a_i = 0$ , which gives the matrix equation:

$$\sum_{j=1}^N A_{ij} a_j = b_j, \quad (6)$$

where

$$A_{ij} = \delta_{ij} \sum_{k=1}^N w_{jk} \theta_{jk} - w_{ij} \theta_{ij}, \quad (7)$$

$$b_i = \sum_{j=1}^N w_{ij} \theta_{ij} \Delta_{ji} = - \sum_{j=1}^N w_{ij} \theta_{ij} \Delta_{ij}. \quad (8)$$

We enforce a strong external constraint on the solution by keeping the zeropoint fixed for the fields which have already been compared and calibrated against APASS. We hereafter refer to these fields as *anchors*. It is asserted that the zeropoints  $a_i$  of the anchor fields are known and not solved for. However they do appear in the vector  $b_j$  as constraints. In addition to the APASS-based anchors, we selected 3273 additional anchor fields by hand to provide additional constraints in regions not covered by APASS. These

extra anchors were deemed to have reliable zeropoints based on (i) the information contained in the observing logs, (ii) the stability of the standard star zeropoints during the night, and (iii) photometricity statistics provided by the Carlsberg Meridian Telescope, which is located  $\sim$ 500 m from the INT.

We then solved Eqn. 6 for the  $r$  and  $i$  bands separately using the least-squares routine in Python's SCIPY.SPARSE module for sparse matrix algebra. This provided us with corrected zeropoints for the remaining fields, which were shifted on average by  $+0.02 \pm 0.11$  in  $r$  and  $+0.01 \pm 0.12$  in  $i$  compared to their provisional calibration. The outcome is validated in the next section.

We then turned to the global calibration of the  $H\alpha$  data. It is not possible to re-calibrate the narrowband in the same way as the broadbands, because the APASS survey does not offer  $H\alpha$  photometry. We can reasonably assume however, that the corrections required for  $r$  and  $H\alpha$  are identical, much of the time, because the IPHAS data-taking pattern ensured that a field's  $H\alpha$  and  $r$ -band exposures were taken consecutively, albeit separated by the  $\sim$ 30 s read-out time. Hence, we have corrected the  $H\alpha$  zeropoints by re-using the zeropoint adjustments that were derived for the  $r$  band in the earlier steps. An exception was made for a small subset of fields for which our quality-control routines revealed strong zeropoint variations during the night, suggesting non-photometric conditions. In these cases, the  $H\alpha$  zeropoints were adjusted by solving Eqn. 6 rather than being linked to the  $r$ -band shift.

#### 4.3 Testing the calibration against SDSS

Having re-calibrated all fields to the expected APASS accuracy of 3 per cent, we turned to a different survey, SDSS Data Release 9 (Ahn et al. 2012), to validate the results. SDSS DR9 includes several strips at low Galactic latitudes, providing data across 18 per cent of the fields in our data release. We cross-matched the IPHAS fields against the subset of objects marked as reliable stars in the SDSS catalogue<sup>4</sup> cross-matching in much the same way as for APASS, with the difference of selecting from the fainter magnitude ranges of  $15 < r_{SDSS} < 18.0$  and  $14.5 < i_{SDSS} < 17.5$ . This provided us with a set of 1.2 million cross-matched stars.

Colour transformations were again obtained using a sigma-clipped linear least squares fit:

$$r_{IPHAS} = r_{SDSS} - 0.093 - 0.044(r - i)_{SDSS} \quad (9)$$

$$i_{IPHAS} = i_{SDSS} - 0.318 - 0.095(r - i)_{SDSS}. \quad (10)$$

The rms residuals of these transformations are 0.045 and 0.073, respectively. The equations are similar to the ones previously determined for APASS, although the colour terms are slightly larger. The throughput curve of the SDSS  $i$ -band filter appears to be somewhat more sensitive at longer wavelengths than both the IPHAS and APASS filters.

These global transformations were deemed adequate for the purpose of validating our calibration in a statistical sense. Separate equations were derived towards differ-

<sup>4</sup> We used the CasJobs facility located at <http://skyserver.sdss3.org/CasJobs> to obtain photometry from the SDSS PHOTOPRIMARY table with criteria TYPE = STAR, CLEAN = 1 and SCORE > 0.7.

ent sight-lines to investigate the effects of varying reddening regimes. The colour term was found to show some variation towards weakly reddened areas, where different stellar populations are observed. The vast majority of red objects in the global sample are those in highly reddened areas, however, which agree well with the global transformations and dominate the statistical appraisal of our calibration.

Having transformed SDSS magnitudes into the IPHAS system, we then computed the median magnitude offset for each IPHAS field which contained at least 30 objects with a cross-matched counterpart in SDSS. This was the case for 2602 fields. The median offsets for each of these fields are shown in Figs. 6 and 7. Importantly, the mean offset and standard deviation found is  $-0.001 \pm 0.029$  mag in  $r$  and  $-0.002 \pm 0.032$  mag in  $i$  (Table 3). In comparison, offsets computed in the identical way *before* carrying out our re-calibration showed means of  $+0.016 \pm 0.088$  mag in  $r$  and  $+0.010 \pm 0.089$  mag in  $i$  (Table 3). We conclude that our re-calibration procedure has been successful in improving the uniformity of the calibration by a factor three and has achieved our aim of bringing the accuracy down to the level of 0.03 mag (rms).

The SDSS comparison revealed a number of fields where the offsets exceeded 0.05 mag (523 fields) or even 0.1 mag (18 fields). This pattern of outliers is consistent with the tails of a Gaussian distribution with  $\sigma = 0.03$ . Furthermore, it should not be forgotten that both the SDSS and APASS calibrations are approximations to perfection and will not be entirely free of anomalies. Indeed as we worked, we noticed the occasional unsurprising examples of inconsistency between these two surveys.

## 5 SOURCE CATALOGUE GENERATION

Having obtained a quality-checked and re-calibrated data set, we now turn to the challenge of transforming the observations into a user-friendly catalogue. The aim is to present the best-available information for each unique source in a convenient format, including flags to warn about quality issues such as source blending and saturation. Compiling the catalogue involved four steps:

- (i) the single-band detection tables produced by the CASU pipeline were augmented with new columns and warning flags;
- (ii) the detection tables were merged into multi-band field catalogues;
- (iii) the overlap regions of the field catalogues were cross-matched to flag duplicate (repeat) detections and identify the primary (best) detection of each unique source; and
- (iv) these primary detections were compiled into the final source catalogue.

Each of these four steps are explained next.

### 5.1 User-friendly columns and warning flags

Enhancement of the detection tables by creating new columns is the necessary first step because the tables generated by the CASU pipeline refer e.g. to source positions in pixel coordinates, to photometric measures in number counts etc., rather than in directly useful astronomical

terms. To transform these data into user-friendly quantities, we have largely adopted the units and naming conventions which are in use at the WFCAM Science Archive (WSA; Hambly et al. 2008) and the VISTA Science Archive (VSA; Cross et al. 2012). These archives curate the high-resolution near-infrared photometry from both the UKIRT Infrared Deep Sky Survey (UKIDSS; Lawrence et al. 2007) and the VISTA Variables in the Via Lactea survey (VVV; Minniti et al. 2010). There is a significant degree of overlap between the footprints of UKIDSS Galactic Plane Survey (GPS; Lucas et al. 2008) and IPHAS, and hence by adopting a similar catalogue format we hope to facilitate scientific applications which combine both data sets.

A detailed description of each column in our source catalogue is given in Appendix A. In the remainder of this section we highlight the main features.

First, we note that each source is uniquely identified by an IAU-style designation of the form ‘IPHAS2 JHH-MMSS.ss+DDMMSS.s’ (cf. column *name* in Appendix A), where ‘IPHAS2’ refers to the present data release and the remainder of the string denotes the J2000 coordinates in sexagesimal format. For convenience, the coordinates are also included in decimal degrees (columns *ra* and *dec*) and in Galactic coordinates (columns *l* and *b*). We have also included an internal object identifier string of the form ‘#run-#ccd-#detection’ (e.g. ‘64738-3-6473’), which documents the INT exposure number (#run), the CCD number (#ccd), and the row number in the CASU detection table (#detection). These columns are named *rDetectionID*, *iDetectionID*, *haDetectionID*.

Photometry is provided based on the 2.3-arcsec diameter circular aperture by default (columns *r*, *i*, *ha*). The choice of this aperture size as the default is based on a trade-off between concerns about small number statistics and centroiding errors for small apertures on one hand, and diminishing signal-to-noise ratios and source confusion for large apertures on the other hand. The user is not restricted to this choice, because the catalogue also provides magnitudes using three alternatives: the peak pixel height (columns *rPeakMag*, *iPeakMag*, *haPeakMag*), the circular 1.2-arcsec-diameter aperture (*rAperMag1*, *iAperMag1*, *haAperMag1*) and the 3.3-arcsec-diameter aperture (*rAperMag3*, *iAperMag3*, *haAperMag3*).

Each of these magnitude measurements have been corrected for the flux lost outside of their respective apertures, using a correction term which is inferred from the mean shape of the PSF measured locally in the CCD frame. In the case of a point source, the four alternative magnitudes are expected to be consistent with each other within the photon noise uncertainty (which is given in columns *rErr*, *rPeakMagErr*, *rAperMag1Err*, *rAperMag3Err*, etc.). When this is not the case, it is likely that the source is either an extended object for which the aperture correction is invalid, or that the object has been incorrectly measured as a result of source blending or a rapidly spatially-varying nebulous background. In §6.1 we will explain that the consistency of the different-aperture magnitude measurements can be used as a criterion for selecting stellar objects with reliable photometry from the catalogue.

The brightness of each object as a function of increasing aperture size is also used by the CASU pipeline

to provide a discrete star/galaxy/noise<sup>5</sup> classification flag (*rClass*, *iClass*, *haClass*) and a continuous stellarness-of-profile statistic (*rClassStat*, *iClassStat*, *haClassStat*). For convenience, we have combined these single-band morphological measures into band-merged class probabilities and flags using the merging scheme in use at the WSA<sup>6</sup> (*pStar*, *pGalaxy*, *pNoise*, *mergedClass*, *mergedClassStat*).

Information on the quality of each detection is included in a series of additional columns. We draw attention to three useful flags which warn about the likely presence of a systematic error:

(i) The *saturated* column is used to flag sources for which the peak pixel height exceeds 55000 counts, which is typically the case for stars brighter than 12-13th magnitude in *r*. Although the pipeline attempts to extrapolate the brightness of saturated stars based on the shape of the PSF, such extrapolation is prone to error, and should be viewed as indicative rather than as precise measurement (systematic uncertainties as a function of magnitude will be discussed in §6.2).

(ii) The *deblend* column is used to flag sources which partially overlap with a nearby neighbour. Although the pipeline applies a deblending procedure to such objects, the procedure is currently applied separately in each band, and hence the (*r - i*) and (*r - Ha*) colours may be inaccurate if the deblending proceeded differently in each band.

(iii) The *brightNeigh* column is used to flag sources which are located within a radius of 5 arcmin from an object brighter than  $V = 7$  according to the Bright Star Catalogue (BSC; Hoffleit et al. 1991), or within 10 arcmin if the neighbour is brighter than  $V = 4$ . These brightest stars are known to cause systematic errors and spurious detections as a result of stray light and diffraction spikes.

In addition to the above, we also created warning flags for internal bookkeeping. For example, we flagged detections which fell in the strongly vignetted regions of the focal plane, which were truncated by CCD edges, or which were otherwise affected by bad pixels in the detector. No such detections have had to be included in the catalogue, as alternative detections were available in essentially all these situations thanks to the IPHAS field pair strategy. In turn, this has meant that it has not proved necessary to include these internal warning flags in the published source catalogue.

Finally, we note that basic information on the observing conditions is included (*fieldID*, *fieldGrade*, *night*, *seeing*). A table containing more detailed quality control information, indexed by *fieldID*, is made available on our website.

## 5.2 Band-merging the detection tables

The second step in compiling the source catalogue is to merge the contemporaneous trios of *r*, *i*,  $\text{Ha}$  detection tables into multi-band field catalogues. This required a position matching procedure to link sources between the three

<sup>5</sup> For consistency with the terminology used in the CASU pipeline and the WSA archive, extended objects are classified as ‘galaxies’ in the catalogue. This class may however include other types of resolved objects.

<sup>6</sup> Explained at [http://surveys.roe.ac.uk/wsa/www/gloss\\_m.html#gpssource\\_mergedclass](http://surveys.roe.ac.uk/wsa/www/gloss_m.html#gpssource_mergedclass)

bands. We used the TMATCHN function of the STILTS software for this purpose, which allows rows from multiple tables to be matched (Taylor 2006). The result of the procedure is a band-merged catalogue in which each row corresponds to a group of linked *r*, *i*, and  $\text{Ha}$  detections which satisfy a maximum matching distance criterion in a pair-wise sense. Sources for which no counterpart was identified are retained in the catalogue as single-band detections.

We employed a maximum matching distance of 1 arcsec, which was chosen, trading off completeness against reliability. On the one hand, a matching distance larger than 1 arcsec was found to allow too many spurious and unrelated sources to be linked. On the other, a value smaller than 1 arcsec would pose problems for very faint sources with large centroiding errors, and would occasionally fail near CCD corners, where the astrometric solution can show local systematic errors which exceed 0.5 arcsec (even though the nominal rms accuracy – estimated against 2MASS across the focal plane – tends to be better than 0.1 arcsec). The position offsets between the *r* detection and detections in *i*, and/or  $\text{Ha}$  have been included in the catalogue, giving the user the option to tighten them further if necessary (columns *iXi*, *iEta*, *haXi*, *haEta*). We note that UKIDSS/GPS adopted the same maximum matching distance of 1 arcsec for similar reasons (Hambly et al. 2008), and warn that applications which require absolute astrometry with sub-arcsecond accuracy must use these catalogues with caution.

The resulting band-merged catalogues were found to be reliable for the vast majority of fields. We do warn that source blending and confusion is unavoidable for faint objects in the most crowded regions of the Galactic Plane; in §6 we will show that 19 per cent of the sources in our catalogue are flagged as blended objects (column *deblend*), and their band-merged data should be treated with care because they may have fallen victim to source confusion during the band-merging step.

## 5.3 Selecting the primary detections

We explained earlier that the survey contains repeat observations of identical sources as a result of field offsetting and overlaps. Amongst all sources in the magnitude range  $13 < r < 19$ , we find that 65 per cent were detected twice and 25 per cent were detected three times or more. Only 9 per cent were detected once.

Since the principal aim of this data release is to provide reliable photometry at a single epoch, we have focused on providing the magnitudes and coordinates from the *best-available* detection of each object – hereafter referred to as the *primary* detection. Although overlapping fields could have been co-added to gain a small improvement in depth, we have decided against this for two reasons. Firstly, combining the information from multiple epochs would make the photometry of variable stars difficult to interpret. Secondly, co-adding would cause the image quality to degrade towards the mean, which is particularly a drawback for crowded fields.

Anyone interested in the alternative detections of a source – hereafter called the *secondary* detections – can nevertheless obtain this information in two ways. To begin with, whenever a secondary detection was collected within 10 min-

utes of the primary, we have included the identifier and the photometry of that secondary detection in the catalogue for convenience (columns *sourceID2*, *fieldID2*, *r2*, *i2*, *ha2*, *rErr2*, *iErr2*, *haErr2*, *errBits2*). Second, detections not included in the catalogue are made available on our website.

Primary detections have been selected from all available detections using a so-called *seaming* procedure, which we adapted from the algorithm developed for the WSA<sup>7</sup>. In brief, the first step is to identify all the duplicate detections by cross-matching the overlap regions of all field catalogues, again using a maximum matching distance of 1''. The duplicate detections for each unique source are then ranked according to (i) filter coverage, (ii) quality score (column *errBits*), and (iii) the average seeing of stars in the CCD frame rounded to 0.2 arcsec. If this ranking scheme reveals multiple ‘winners’ of seemingly identical quality, then the one that was observed closest to the optical axis of the camera is chosen.

#### 5.4 Compiling the final source catalogue

As the final step, the primary detections selected above were compiled into the 98-column source catalogue that is described in Appendix A and is made available on our website and through Vizier. The original unweeded list of sources naturally included a significant number of spurious entries as a result of the very sensitive default detection settings that are employed by the CASU pipeline. To limit the size of the source catalogue, we have decided to enforce three basic criteria which must be met for a candidate source to be included in the catalogue:

- (i) the source must have been detected at  $S/N > 5$  in at least one of the bands, i.e. it is required that at least one of *rErr*, *iErr* or *haErr* is smaller than 0.2 mag;
- (ii) the shape of the source must not be an obvious cosmic ray or noise artefact, i.e. we require either *pStar* or *pGalaxy* to be greater than 20 per cent;
- (iii) the source must not have been detected in one of the strongly vignetted corners of the detector, not have had any known bad pixels in the aperture, and not have been on the edge of one of the CCDs (this information is contained in the *errBits* quality score column, which is used for internal bookkeeping).

A total of 219 million primary detections satisfied the above criteria and have been included in the catalogue.

Table 4 details the number of bands in which the catalogue entries are captured; 159 million sources are detected in all three filters (73 per cent), 30 million are detected in two filters (14 per cent), and the remaining 30 million are single-band detections. Table 4 also presents the average magnitudes and the fraction of ‘confirmed’ objects, which we define as those sources which have been detected in more than one field (recall that the majority of the survey area is observed twice due to the field pair strategy). We find that the single-band detections tend to show low confirmation rates, which suggests that a significant fraction of these

**Table 4.** Breakdown of catalogue entries as a function of the band(s) in which the source was detected. We also show the fraction of ‘confirmed’ objects, which we define as those sources detected in more than one field (usually the field pair partner). We also show the average magnitudes in each class.

Band(s)	Sources [10 <sup>6</sup> ]	Confirmed [nObs>1]	$\bar{r}$ [mag]	$\bar{i}$ [mag]	$\bar{H}\alpha$ [mag]
<i>r, i, H</i> $\alpha$	159	91%	19.3	18.2	18.8
<i>r, i</i>	25	77%	21.2	19.9	-
<i>i, H</i> $\alpha$	3	73%	-	19.6	20.2
<i>r, H</i> $\alpha$	2	65%	20.6	-	20.1
<i>i</i>	15	43%	-	19.8	-
<i>r</i>	9	27%	21.1	-	-
$H\alpha$	6	12%	-	-	19.9
Total	219	81%	19.6	18.6	18.9

entries may be spurious detections. Hence, we strongly recommend users not to rely on single-band objects without inspecting the image data by eye.

Not all the single-band detections are spurious sources, however. We note that the confirmation rate for *i*-band detections is markedly better than for *r* and  $H\alpha$ , which is likely explained by the fact that *i* is least affected by interstellar extinction, and so the survey can occasionally pick up highly-reddened objects in *i* which are otherwise lost in *r* and  $H\alpha$ . Moreover, objects which are intrinsically very red may also be picked up in *i* alone, while faint objects with strong Balmer emission may appear only in  $H\alpha$ .

## 6 DISCUSSION

We now offer an overview of the properties of the catalogue by discussing (i) the recommended quality criteria, (ii) the photometric uncertainties and repeatability, and (iii) the source densities and the frequency of source blending.

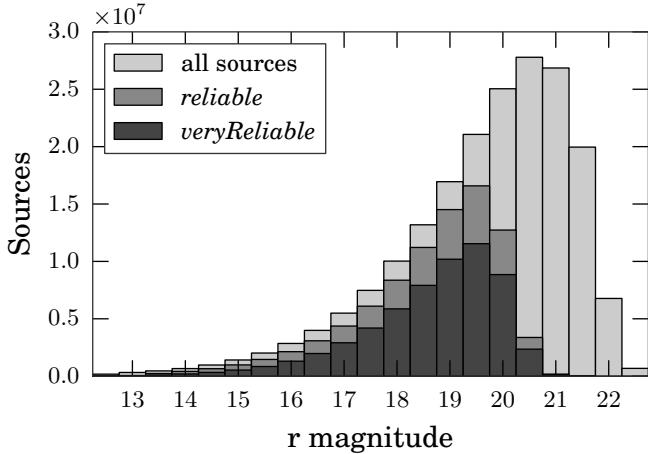
### 6.1 Recommended quality criteria

Like any other photometric survey, the majority of the objects in our catalogue are faint sources observed near the detection limits; 55 per cent of the entries in the catalogue are fainter than  $r > 20$ . The measurements of faint objects are naturally prone to larger random and systematic uncertainties: for example, an inaccurately-subtracted background will introduce a proportionally larger systematic error for a faint object. Most scientific applications will hence require a set of quality criteria to be enforced for the purpose of removing lower-quality objects.

The choice of quality criteria will always tension completeness against accuracy. To aid users we have listed two sets of recommended quality criteria in Tables 5 and 6.

Table 5 specifies a set of minimum quality criteria which should benefit most applications which desire reliable colours as well as completeness. The listed criteria are designed to (i) remove low-S/N sources, (ii) remove saturated sources, and (iii) remove objects for which the 2''.3-diameter aperture magnitude is inconsistent with its alternative 1''.2-diameter measurement within each filter. This last criterion is a proxy for identifying objects which are affected by poor

<sup>7</sup> <http://surveys.roe.ac.uk/wsa/dboverview.html#merge>



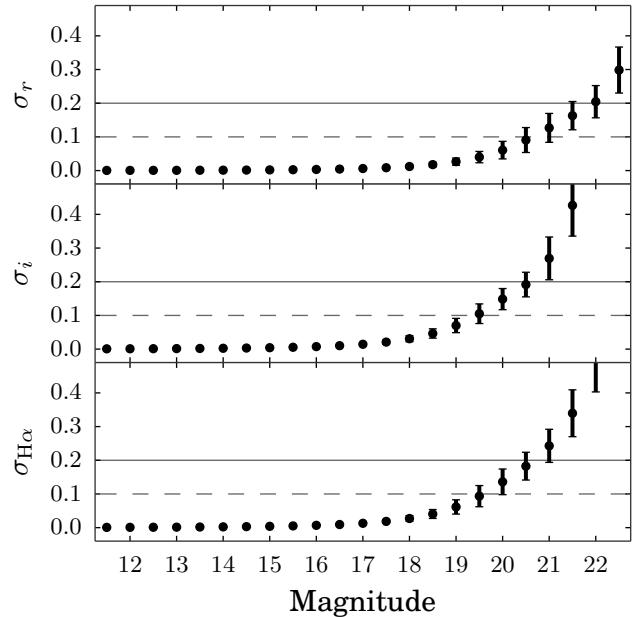
**Figure 8.** r-band magnitude distribution for all objects in the catalogue (light grey), for objects flagged as *reliable* according to the criteria set out in Table 5 (grey), and for objects flagged as *veryReliable* following Table 6 (dark grey). The magnitude distributions for *i* and H $\alpha$  look identical, apart from being shifted by about 1 and 0.5 mag towards brighter magnitudes, respectively.

background subtraction or failed source deblending. A total of 86 million sources out of 219 million (39 per cent) satisfy all the criteria listed in Table 5 and are hereafter referred to as ‘reliable’. For convenience, the catalogue contains a boolean column named *reliable* that directly flags these objects.

For applications which require a higher standard of reliability at the expense of completeness, a further set of additional quality criteria are suggested in Table 6. These criteria are designed to ensure that (i) the object appeared as a perfect point source, (ii) the object was not blended with a nearby neighbour, and (iii) the object was not located near a very bright star. 59 million sources (27%) satisfy these stricter criteria and are hereafter referred to as ‘very reliable’. Again, the catalogue contains a boolean column named *veryReliable* which flags these objects.

Fig. 8 compares the *r*-band magnitude distributions for *reliable*, *veryReliable* and unfiltered objects. We find that 81% of the sources are considered *reliable* and 54% are *veryReliable* in the magnitude range  $13 < r < 19$ . We will explain below that the *veryReliable* category is least complete at low Galactic longitudes, where source blending can affect up to a quarter of the objects. The *veryReliable* flag should hence only be used in applications in which completeness is not a concern.

It is easy to see how the quality criteria may be adapted to be more tolerant. For example, by raising the allowed photometric uncertainties from 0.1 mag to 0.2 mag, 42 million candidate sources would be added to the 109 million satisfying the tighter error bound. Our choice to adopt 0.1 mag as the cut-off uncertainty for the *reliable* category is a pragmatic trade-off which we found to suit many science applications, but certainly not all. Users of the data are encouraged to revise the quality criteria according to their needs.



**Figure 9.** Mean photometric uncertainties for *r* (top), *i* (middle) and H $\alpha$  (bottom). Data points shown are the average values of columns *rErr*, *iErr* and *haErr* in the catalogue and the error-bars show the standard deviations. The dashed and solid lines indicate the  $10\sigma$  and  $5\sigma$  limits, respectively. These statistics show the average level of the Poissonian photon noise and do not include systematic uncertainties.

## 6.2 Random and systematic uncertainties

Fig. 9 shows the mean photometric uncertainties (*rErr*, *iErr*, *haErr*) as a function of magnitude. We find the typical uncertainty to reach 0.1 mag near  $r = 20.5$  and  $i/\text{H}\alpha = 19.5$ . We note that the fainter depth in *r* is compensated by the fact that most stars have brighter magnitudes in *i* and H $\alpha$ ; the average colours in the catalogue are  $(r - i) = 1.06 \pm 0.12$  and  $(r - \text{H}\alpha) = 0.44 \pm 0.03$ . We warn that the statistics shown in Fig. 9 are the random errors based on the expected Poissonian photon noise. Systematics, such as calibration and deblending errors, are not included.

To appraise the level at which our photometry is affected by such systematics, we can exploit the secondary measurements which are present in the catalogue (i.e. *r* vs *r2*, *i* vs *i2*, *ha* vs *ha2*). In Fig. 10a we show the mean absolute residuals between these primary and secondary magnitudes as a function of magnitude (black dots). We also plot the Poissonian uncertainties for comparison (solid red line). We find the mean residual to equal  $0.03 \pm 0.04$  mag across the magnitude ranges 13 to 18 (*r*) and 12 to 17 (*i*, H $\alpha$ ), consistent with the accuracy of our calibration estimated above. Stars fainter than this range appear to be dominated by photon noise (red line), while stars at the bright end appear to suffer from large systematic errors due to saturation effects.

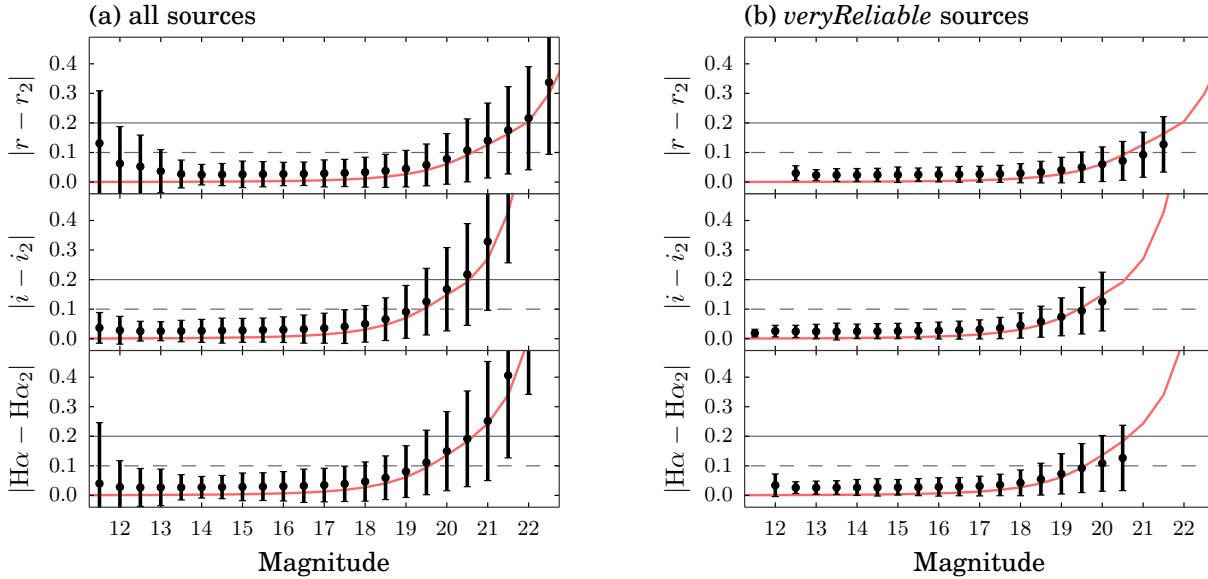
In Fig. 10b we show a similar comparison between the primary and secondary detections, but this time we have only included sources which are flagged as *veryReliable* in the catalogue. We do not observe an improvement in the average residuals as a function of magnitude, but the number of outliers has decreased markedly (evidenced by the

**Table 5.** Recommended minimum quality criteria for selecting objects with reliable colours from the IPHAS DR2 source catalogue. 86 million entries in the catalogue (39%) satisfy all the criteria listed in this table. For convenience, these have been flagged in the catalogue using the column named *reliable*.

Quality criterion	Rows passed	Description
$rErr < 0.1 \text{ AND } iErr < 0.1 \text{ AND } haErr < 0.1$	109 million (50%)	Require the photon noise to be less than 0.1 mag in all bands (i.e. S/N>10). This implicitly requires a detection in all three bands.
$r > 13 \text{ AND } i > 12 \text{ AND } H\alpha > 12.5 \text{ AND NOT saturated}$	158 million (72%)	The brightness must not exceed the nominal saturation limit and the peak pixel height must not exceed 55 000 counts. Again, this implicitly requires a detection in all three bands.
$ r - rAperMag1  < 3\sqrt{rErr^2 + rAperMag1Err^2} + 0.03$	176 million (80%)	Require the $r$ magnitude measured in the default 2''.3-diameter aperture to be consistent with the measurement made in the smaller 1''.2 aperture, albeit tolerating a 0.03 mag systematic error. This will reject sources for which the background subtraction or the deblending procedure was not performed reliably.
$ i - iAperMag1  < 3\sqrt{iErr^2 + iAperMag1Err^2} + 0.03$	183 million (84%)	Same as above for $i$ .
$ ha - haAperMag1  < 3\sqrt{haErr^2 + haAperMag1Err^2} + 0.03$	158 million (72%)	Same as above for $H\alpha$ .
All of the above (flagged as <i>reliable</i> )	86 million (39%)	

**Table 6.** Additional quality criteria which are recommended for applications which require very reliable colours at the expense of completeness. For convenience, the sources which satisfy the criteria listed in this table have been flagged in the catalogue using the column named *veryReliable*.

Quality criterion	Rows passed	Description
<i>reliable</i>	86 million (39%)	The object must satisfy the criteria listed in Table 5.
$pStar > 0.9$	145 million (66%)	The object must appear as a perfect point source, as inferred from comparing its PSF with the average PSF measured in the same CCD.
NOT <i>deblend</i>	177 million (81%)	The source must appear as a single, unconfused object.
NOT <i>brightNeighb</i>	216 million (99%)	There is no star brighter than $V < 4$ within 10 arcmin, or brighter than $V < 7$ within 5 arcmin. Such very bright stars cause scattered light and diffraction spikes, which may add systematic errors to the photometry or even trigger spurious detections.
All of the above (flagged as <i>veryReliable</i> )	59 million (27%)	



**Figure 10.** Photometric repeatability as a function of magnitude for all sources in the catalogue (left panel) and for the *veryReliable* sources alone (right panel). Black dots show the mean absolute residuals between the primary and the secondary detections. The error-bars show the standard deviations. The red trend line shows the average Poissonian uncertainties we derived in Fig. 9. We find that the *veryReliable* quality criteria are successful at removing objects with large uncertainties.

shorter error bars which denote the standard deviation of the absolute residuals). We conclude that the *veryReliable* quality criteria are effective at reducing the level of the systematic errors, while also removing the unreliable data at the bright and faint end.

### 6.3 Source counts and blending

Fig. 11 shows the number of sources in the catalogue counted in  $1^\circ$ -wide strips as a function of Galactic longitude (thick blue line). Unsurprisingly, we find the number of sources to increase towards the Galactic centre. For example, the average source density near  $\ell \simeq 30^\circ$  is roughly 300 000 objects per square degree, which is six times more than the density found near  $\ell \simeq 180^\circ$ . In addition to the global trend, variations are also apparent on smaller scales. For example, we find a significant drop near the constellations of Aquila ( $\ell \simeq 40^\circ$ ) and Cygnus ( $\ell \simeq 80^\circ$  and  $\ell \simeq 90^\circ$ ), which are regions known to be affected by high levels of foreground extinction (the extremities of ‘the Great Rift’, e.g. Bok & Bok 1941). However, we warn that the source counts shown have not been corrected for field pairs that have yet to be released or for variations in the depth across the included fields. For example, the dip in the density near  $\ell \simeq 140^\circ$  is an artificial feature caused by gaps in the footprint coverage (seen in Fig. 1).

In a forthcoming paper, properly calibrated maps of stellar density of the northern Galactic Plane will be presented (Farnhill et al., in preparation). This will incorporate completeness corrections based on the statistics of artificial source recovery. Such maps are of interest as tests of detailed models of our Galaxy.

Fig. 11 also shows the number counts for non-blended

sources (thin red line). These are sources for which the *deblend* flag is FALSE, i.e. sources for which the CASU pipeline did not have to apply a deblending procedure to separate the flux originating from two or more overlapping objects. This provides some insight into how the fraction of blend-affected sources correlates with the local source density. In hard numbers, only 11% of the sources are blended at  $\ell > 90^\circ$ , whereas 24% are blended at  $\ell < 90^\circ$ .

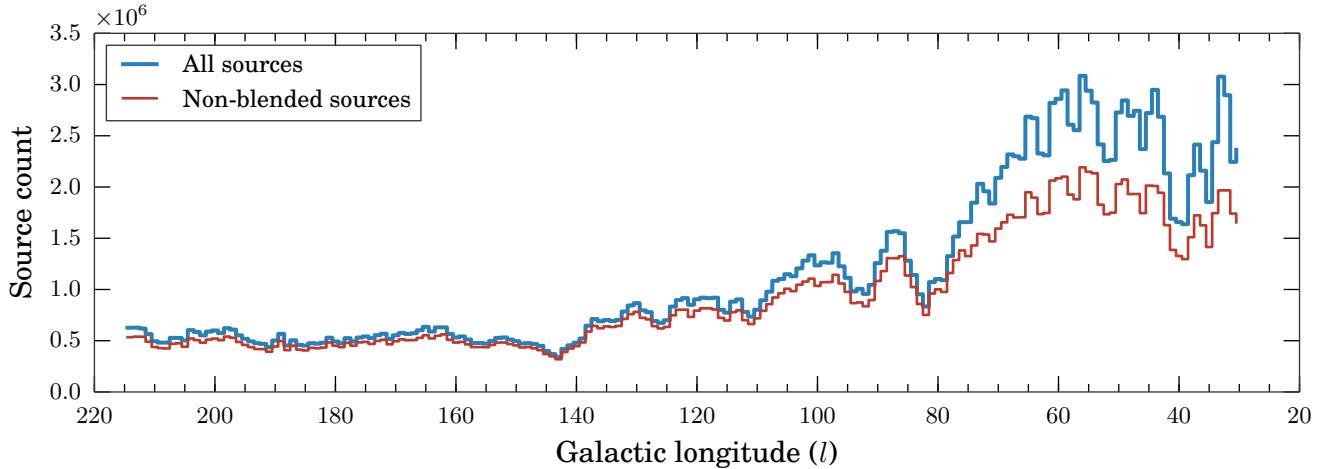
Finally, we warn that blended objects are more likely than unblended objects to have fallen victim to source confusion during the band-merging procedure. It is important to bear this in mind when appraising stars of seemingly unusual colours (such as candidate emission line stars) – if blending is flagged, the probability that the unusual colour is spurious is greatly increased.

## 7 DEMONSTRATION

We conclude this paper by demonstrating how the unique  $(r - H\alpha, r - i)$  colour-colour diagram offered by this catalogue can readily be used to (i) characterise the extinction regime along different sight-lines, and (ii) identify  $H\alpha$  emission-line objects.

### 7.1 Colour-colour and colour-magnitude diagrams

The survey’s unique  $(r - H\alpha)$  colour, when combined with  $(r - i)$ , has been shown to provide simultaneous constraints on intrinsic stellar colour and interstellar extinction (Drew et al. 2008). Put differently, the main sequence in the  $(r - H\alpha, r - i)$  diagram runs in a direction that is at a large angle relative to the reddening vector, because the  $(r - H\alpha)$  colour tends to act as a coarse proxy for spectral type and



**Figure 11.** Number of entries in the IPHAS DR2 source catalogue as a function of Galactic longitude. The upper blue line shows the number of sources counted in  $1^\circ$ -wide longitude bins. The lower red line uses the same binning but includes only those sources for which the *deblend* flag is FALSE, i.e. unconfused sources for which the CASU pipeline did not have to apply a deblending procedure. In both cases we counted only those sources in the latitude range  $-5^\circ < b < +5^\circ$ , such that one may obtain approximate source densities by dividing the counts by  $10 \text{ deg}^2$ . The global variations in the source counts traces the structure of the Galaxy and the distribution of foreground extinction, but is also affected by instrumental effects such as variations in the observed depth and completeness (see text).

is less sensitive to reddening than  $(r - i)$ . As a result, the distribution of a stellar population in the IPHAS colour-colour diagram can offer a handle on the properties of the population and the extinction along a line of sight.

This is demonstrated in Fig. 12, 13 & 14, where we present three sets of IPHAS colour/magnitude diagrams towards three distinct sight-lines located at Galactic longitudes  $180^\circ$ ,  $45^\circ$  and  $30^\circ$ , respectively, which were chosen because they show stellar populations with different characteristics. Each figure contains all the sources flagged as *veryReliable* within a region of one square degree centred on the coordinates indicated in the diagram (i.e. within a radius of  $0.564^\circ$  from the indicated sight-line). For clarity, we have imposed the additional criterion that the photometric uncertainties must be smaller than 0.05 mag in each band, corresponding to a cut-off near 19th magnitude.

Each of the diagrams reveals a well-defined locus, which helps to demonstrate the health of the catalogue and the global calibration for investigating stellar populations across wide areas. We have annotated the colour-colour diagrams by showing the position of the unreddened main sequence (thin solid line), the unreddened giant branch (thick solid line), and the reddening track for an A0V-type star (dashed line) – all three are based on the Pickles (1998) library of empirical spectra transformed into the Vega-based IPHAS system by Drew et al. (2005). In the colour-magnitude diagrams we only show the reddening vector together with the unreddened 1 Gyr isochrone due to Bressan et al. (2012), which is made available for the IPHAS system through the on-line tool hosted at the Observatory of Padova (<http://stev.oapd.inaf.it/cmd>). The isochrone and reddening vector have been placed at an arbitrary distance of 2 kpc.

Each of the sightlines reveals a stellar population with distinct characteristics. Towards the Galactic anti-centre at  $\ell = 180^\circ$  (Fig. 12) we find a population dominated by lowly-reddened main sequence stars. This is consistent with the estimated total sight-line extinction of  $E(B - V) = 0.49$

given by Schlegel et al. (1998), following the correction of Schlafly & Finkbeiner (2011). Looking in more detail we can see that the stellar locus is narrower for M-type dwarfs than for earlier types; we do not observe M dwarfs experiencing the strongest reddening possible for this sight-line. This implies that extinction is still increasing at distances of  $\sim 1\text{-}2$  kpc, where M dwarfs become too faint to be contained in the IPHAS catalogue. It is also clear that there are no unreddened stars earlier than  $\sim K0$  visible; such stars would be saturated if within a few hundred parsecs. This therefore suggests that there is a measurable increase in extinction locally. We also note a relative absence of late type giants which, due to the relative brevity of the corresponding phase of stellar evolution, would only account for a small proportion of the more nearly *volume-limited* sample seen in the Anticentre direction.

In contrast, closer towards the galactic centre, we encounter samples which are *magnitude-limited* instead. For example, at  $\ell = 45^\circ$  (Fig. 13), there is a wealth of reddened objects and late type giants. In the colour-magnitude diagram, it is clear that the stars are split into two distinct groups, with one significantly redder than the other. The bluer group is composed of main sequence stars, with the slope of this group in the colour-magnitude diagram attributable to the significantly increasing extinction. Meanwhile the redder group is principally composed of red giant stars (see Wright et al. 2008). As these stars are intrinsically brighter, they will be substantially further away than their main sequence counterparts at the same apparent magnitude. Given that extinction continues to increase with distance, along this sightline, the red giants we observe will be subject to appreciably more reddening than the main sequence stars, pushing them to  $(r - i) \sim 1.5$ .

Finally, in one of our earliest sight-lines at  $\ell = 30^\circ$ , we find a very high number of extremely reddened giants in addition to an unreddened population of foreground dwarfs. In contrast to the sight-line at  $\ell = 45^\circ$ , there is no clear

group of giant stars visible in the colour-magnitude diagram of Fig. 14, although the red clump stars are manifest as a track of slight over-density sitting roughly 0.4 mag redder than the A0V reddening track. At  $(l, b) = (45^\circ, +2^\circ)$  the giant stars observed exhibit a relatively narrow range of reddenings as they lie beyond most or essentially all of the Galactic dust. At  $(l, b) = (30^\circ, 0^\circ)$  this is not the case: even at the substantial distances at which we can observe reddened giant stars, extinction is continuing to rise within the Galactic mid-plane. It is also apparent that the  $(r - i)$  width of both the M dwarfs and early A dwarfs is greater than that in Fig. 13. This is indicative of a steeper rise in reddening, both within several hundred parsecs (M dwarfs) and within a few kpc (early A dwarfs).

These are just qualitative vignettes of the information obtainable from IPHAS colour-colour and colour-magnitude plots. A more rigorous quantitative analysis of the IPHAS catalogue can be undertaken to estimate both the stellar density distribution in the Milky Way (Sale et al. 2010) and to create detailed three-dimensional maps of the extinction across several kpc (Sale et al. 2009; Sale 2012). A 3-D extinction map based on the DR2 catalogue is being released in a separate paper (Sale et al., in preparation).

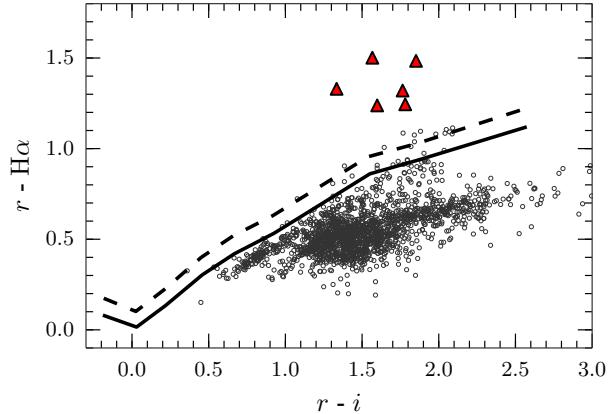
## 7.2 Identifying $\text{H}\alpha$ emission-line objects

An aim of IPHAS is to enable the discovery of new fainter emission-line objects across the Galactic Plane.  $\text{H}\alpha$  emission is a well-known tracer for stars in the short-lived pre- or post-main sequence stages of their evolution, and hence IPHAS allows larger, deeper and more statistically robust samples of such rare objects to be established. The survey-wide identification and analysis of emission-line objects is beyond the scope of the present work and will be the focus of a forthcoming paper (Barentsen et al, in preparation). In this section we merely aim to demonstrate a use of the catalogue for this purpose.

An initial list of candidate  $\text{H}\alpha$ -emitters based on the first IPHAS data release was previously presented by Witham et al. (2008). Because no global calibration was available at the time, Witham et al. employed a sigma-clipping technique to select objects with large, outlying  $(r - \text{H}\alpha)$  colours. In contrast, the new catalogue allows objects to be picked out from the  $(r - \text{H}\alpha, r - i)$  colour-colour diagram using model-based colour criteria rather than a statistical procedure. In what follows we demonstrate this ability by selecting candidate emission-line objects towards a small region in the sky.

The target of our demonstration is Sh 2-82: a 5 arcmin-wide HII region located near  $(l, b) = (53.55^\circ, 0.00^\circ)$  in the constellation of Sagitta. Nicknamed by amateur astronomers as the ‘Little Cocoon Nebula’, Sh 2-82 is ionised by the  $\sim 10$ th magnitude star HD 231616 with spectral type B0V/III (Georgelin et al. 1973; Mayer & Macák 1973; Hunter & Massey 1990). This ionising star has been placed at a likely distance of 1.5–1.7 kpc based on its photometric parallax (Mayer & Macák 1973; Lahulla 1985; Hunter & Massey 1990).

Fig. 15 shows a 20-by-15 arcmin colour mosaic centred on Sh 2-82, composed of our  $\text{H}\alpha$  (red channel),  $r$  (green channel), and  $i$  (blue channel) images. The ionising star can be seen as the bright object in the centre of the HII region,

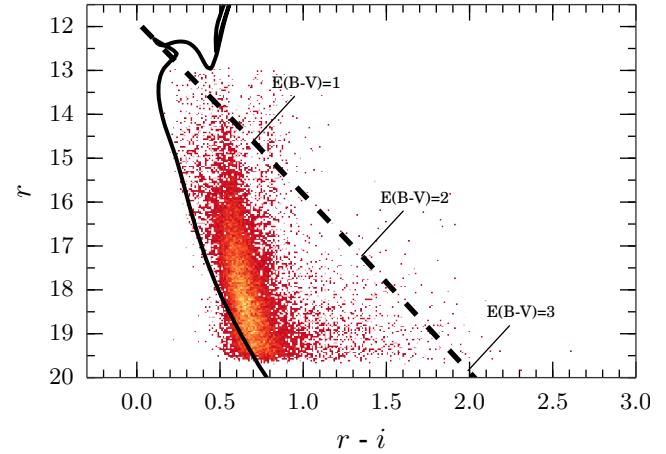
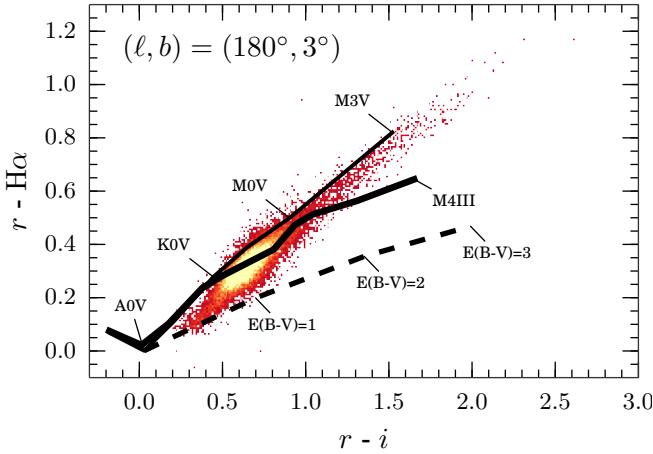


**Figure 17.**  $(r - \text{H}\alpha, r - i)$  diagram for the rectangular region of 20-by-15 arcmin centred on the HII region Sh 2-82, which is the area shown in Fig. 15. The diagram shows all objects in the catalogue which have been flagged as *reliable* and are brighter than  $r < 20$  (grey circles). The unreddened main sequence is indicated by a solid line, while the main sequence for stars with an  $\text{H}\alpha$  emission line strength of  $-10 \text{ \AA EW}$  is indicated by a dashed line (both based on the colour simulations by Barentsen et al. 2011). Red triangles indicate objects which have been identified as likely  $\text{H}\alpha$ -emitters.

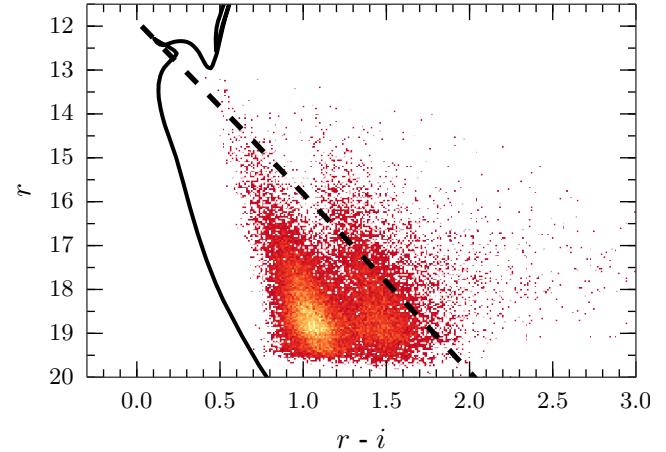
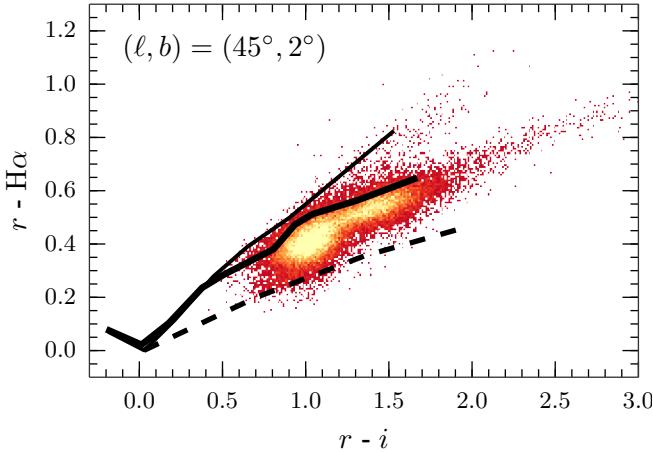
which is surrounded by a faint reflection nebula and several dark cloud filaments. For comparison, Fig. 16 shows a mosaic of the same region as seen in the mid-infrared by the Spitzer Space Telescope (Benjamin et al. 2003; Churchwell et al. 2009). The infrared image reveals an enclosing fuzzy bubble (appearing green in Fig. 16) which is thought to originate from the mid-infrared emission of Polycyclic Aromatic Hydrocarbons (PAHs) – i.e. warm dust – which is frequently observed at the interface between neutral regions of interstellar material and the ionising radiation from early-type stars (Churchwell et al. 2006). Yu & Wang (2012) recently noted that the warm dust that surrounds Sh 2-82 appears to contain infrared-bright Young Stellar Objects (YSOs). Many of these young objects appear as red- and pink-coloured stars in Fig. 16, located predominantly in the top-left part of the bubble.

Fig. 17 presents the IPHAS colour-colour diagram for the 20-by-15 arcmin region shown in the mosaics. Grey circles show all objects which are brighter than  $r < 20$  and have been flagged as *reliable* in our catalogue. The diagram also shows the unreddened main sequence (solid line) and the expected position of unreddened main-sequence stars with  $\text{H}\alpha$  in emission at a strength of  $\text{EW} = -10 \text{ \AA}$  (dashed line). Six stars are found to lie above the dashed line at the level of  $3\sigma$ , i.e. the distance between the objects and the dashed line is larger than three times the uncertainty in their  $(r - \text{H}\alpha)$  colour. These candidate  $\text{H}\alpha$ -emitters are marked by red triangles in the colour-colour diagram, and yellow triangles in the image mosaic (Fig. 15). Their details are listed in Table 7.

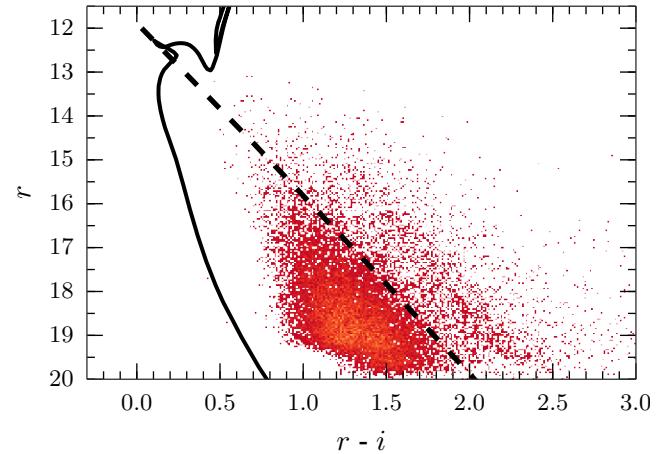
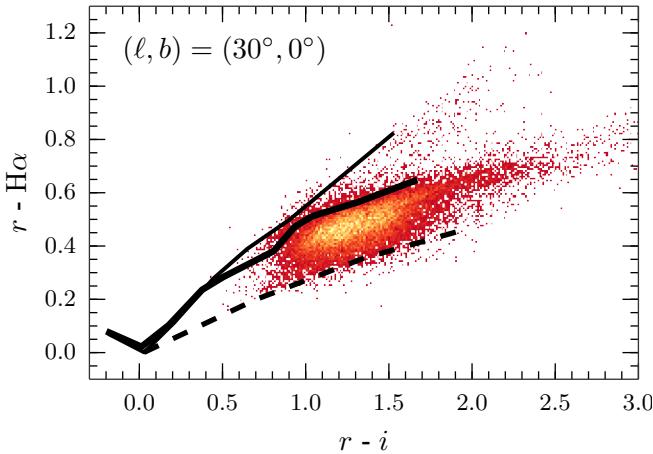
In previous work, we have shown that the majority of  $\text{H}\alpha$ -emitters seen by IPHAS towards an HII region are likely to be Classical T Tauri Stars (Barentsen et al. 2011). These are young objects which are thought to show  $\text{H}\alpha$  in emission due to the presence of hot, infalling gas which is accreting onto the star from a circumstellar disk. This is likely to be



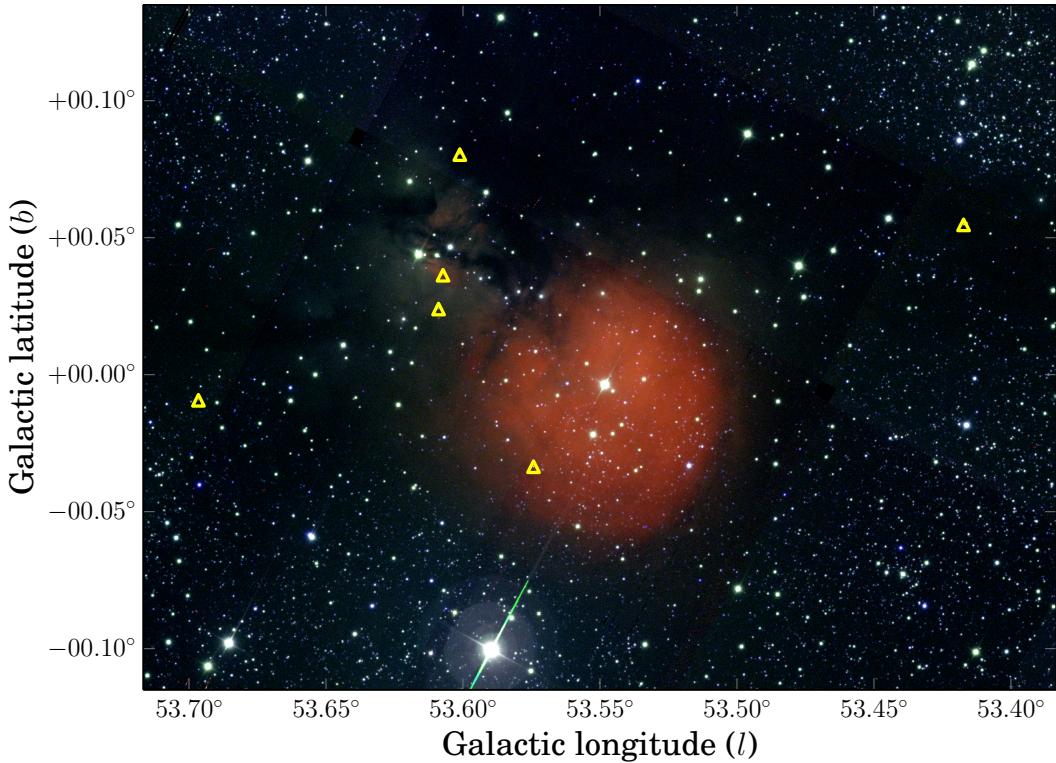
**Figure 12.** Colour-colour and colour-magnitude diagrams (left and right panel) showing sources flagged as *veryReliable* located in an area of  $1 \text{ deg}^2$  centred near the Galactic anti-centre at  $(l, b) = (180^\circ, +3^\circ)$ . The diagrams are plotted as 2D-histograms which show the density of objects in bins of 0.01-by-0.01 mag; bins containing 1 to 10 objects are coloured red, bins with 10 to 20 objects are orange, and bins with 20 to 25 objects are yellow. The left panel is annotated with the position of the main sequence (thin solid line), giant stars (thick solid line) and the reddening track for an A0V-type star (dashed line). The right panel shows the unreddened 1 Gyr isochrone from the models by Bressan et al. (2012, solid line) along with the reddening vector for an A0V-type (dashed line), both placed at an arbitrary distance of 2 kpc. This is one of the least reddened sight-lines in the survey and hence the observed stellar population appears to be dominated by lowly reddened main sequence stars (see text).



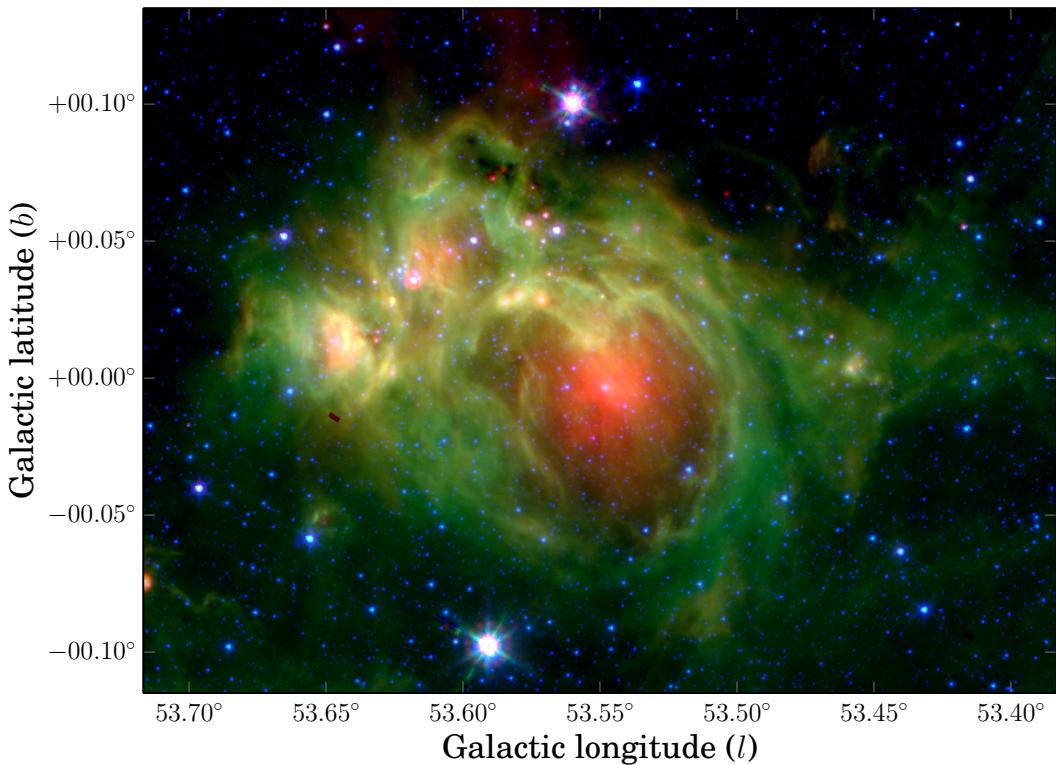
**Figure 13.** Same as above for  $(l, b) = (45^\circ, +2^\circ)$ , which is one of the highest-density sight-lines in the survey, revealing two groups of stars in colour-magnitude space.



**Figure 14.** Same as above for  $(l, b) = (30^\circ, 0^\circ)$ , showing one of the most reddened sight-lines in the survey.



**Figure 15.** IPHAS-based mosaic of HII region Sh 2-82, composed of H $\alpha$  (red channel),  $r$  (green channel) and  $i$  (blue channel). Yellow triangles show the position of candidate H $\alpha$ -emitters which have been selected from the colour-colour diagram in Fig. 17.



**Figure 16.** Star-forming region Sh 2-82 as seen in the mid-infrared by the Spitzer Space Telescope. The mosaic is composed of the 24  $\mu\text{m}$  (red), 8.0  $\mu\text{m}$  (green) and 4.5  $\mu\text{m}$  (blue) bands. The image reveals a bubble-shaped structure which surrounds the HII region that is seen in the IPHAS mosaic which spans the same region (Fig. 15). This structure has previously been labelled as N115 in the catalogue of Churchwell et al. (2006), and could be a possible site of triggered star formation (Thompson et al. 2012; Kendrew et al. 2012).

**Table 7.** Candidate H $\alpha$ -emitters towards Sh 2-82.

#	Name [IPHAS2 ...]	<i>r</i>	<i>i</i>	H $\alpha$
1	J192954.40+181026.1	17.69 ± 0.01	16.12 ± 0.01	16.19 ± 0.01
2	J193011.01+182051.2	18.55 ± 0.02	16.95 ± 0.02	17.31 ± 0.02
3	J193021.52+181954.5	19.72 ± 0.05	17.94 ± 0.03	18.47 ± 0.04
4	J193024.45+181938.3	19.31 ± 0.04	17.55 ± 0.02	17.99 ± 0.03
5	J193033.00+181609.3	18.25 ± 0.01	16.91 ± 0.01	16.92 ± 0.01
6	J193042.48+182317.4	19.96 ± 0.03	18.11 ± 0.03	18.48 ± 0.03

the case for the candidate H $\alpha$ -emitters we discovered towards Sh 2-82 as well. Two of our candidates, #1 and #4 in Tab. 7, have previously been identified as candidate YSOs by Robitaille et al. (2008) and Yu & Wang (2012), respectively. In these studies, the authors used Spitzer data to find intrinsically red objects, which is consistent with the presence of a circumstellar disk. Although the other four candidate emitters in our sample have not previously appeared in the literature, we note that all four are detected in the Spitzer 8.0  $\mu\text{m}$  image at S/N>5. They are likely to be YSOs exhibiting a mild infrared excess.

Sh 2-82 is one of a large population of poorly-studied star-forming regions located at low Galactic latitudes, which have only recently started to become revealed by efforts to catalogue the wealth of ‘bubbles’ detected at mid-infrared wavelengths (Churchwell et al. 2006; Simpson et al. 2012), and by efforts to catalogue previously unknown clusters seen in the near-infrared (e.g. Bica et al. 2003). IPHAS data can offer a handle on the extinction, distance and stellar contents of these unexplored regions.

## 8 CATALOGUE AND IMAGE ACCESS

The catalogue is made available through the Vizier service (<http://vizier.u-strasbg.fr>), where it can be queried using a web interface and using Virtual Observatory (VO) protocols. In addition, the catalogue can be downloaded in its entirety from our website ([www.iphas.org](http://www.iphas.org)) as a collection of binary FITS tables, each covering a  $5^\circ \times 5^\circ$  tile of the footprint.

We do not recommend using the catalogue to study extended objects which are larger than the aperture diameters used in this work. To enable the analysis of diffuse sources, our website provides access to the pipeline-processed imaging data. These have been updated to include a new FITS header keyword, called PHOTZP, which contains the re-calibrated zeropoint. This keyword can be used to convert the number counts  $DN$ , i.e. the pixel values in the images, into Vega-based magnitudes  $m$  using:

$$m = \text{PHOTZP} - 2.5 \log_{10}(DN). \quad (11)$$

The PHOTZP value has been computed such that it absorbs the required corrections for atmospheric extinction, gain variations, exposure time, and the re-calibration shift. Users will only need to apply an aperture correction, if necessary.

To estimate absolute narrow-band H $\alpha$  fluxes from the image data, we note that the integrated in-band energy flux for Vega in the IPHAS H $\alpha$  filter is  $1.52 \times 10^{-7}$  erg cm $^{-2}$  s $^{-1}$  at the top of the Earth’s atmosphere, which is the flux obtained by folding the Vega model SED from Kurucz with the filter transmission curve (uncorrected for the atmosphere and detector quantum efficiency, which would otherwise scale down the narrowband flux by 0.707). This implies that the in-band flux corresponding to zero magnitude

is  $1.56^{-7}$  erg cm $^{-2}$  s $^{-1}$ , when the H $\alpha$  magnitude for Vega is set by convention to 0.03 (Fukugita et al. 1996).

The image repository on our website is accompanied by an archive of detection tables – one for each image – which may be used to retrieve secondary detections that are not included in the DR2 source catalogue. We warn that this set of original images and detection tables includes data that did not pass quality control, and has not been globally re-calibrated. Such data are flagged and must be used with great caution.

In the spirit of reproducibility, the source code that was used to generate the catalogue is made available at <https://github.com/barentsen/iphas-dr2>

## 9 CONCLUDING REMARKS AND FUTURE WORK

A new catalogue of 219 million sources has been derived from the INT/WFC Photometric H $\alpha$  Survey of the Northern Galactic Plane. It is the first catalogue to offer comprehensive CCD photometry of point sources across the Northern Galactic Plane at visible wavelengths, taking in the Galactic latitude range  $|b| < 5^\circ$  at longitudes  $\ell = 30$  to  $215^\circ$ . The new 98-column catalogue provides single-epoch photometry across 92% of the survey area, and is the first quality-controlled and globally calibrated catalogue to have been constructed from the imaging data. This now means that there is H $\alpha$  coverage, accessible online, of the entire Galactic Plane – given that the southern Plane is already available thanks to the UK Schmidt H $\alpha$  Survey (SHS, Parker et al. 2005), the last of the photographic surveys carried out by that telescope.

The observations included in this release achieve a median seeing of 1.1 arcsec and 5 $\sigma$ -depths of  $r = 21.2 \pm 0.5$ ,  $i = 20.0 \pm 0.3$ , and H $\alpha = 20.3 \pm 0.3$ . The global calibration and photometric repeatability are found to be accurate at the level of 0.03 mag (rms), providing a significant improvement over the previous data release. The source catalogue specifies the best-available single-epoch astrometry and photometry for 219 million unique sources. To support its exploitation, we provide a list of recommended quality criteria that will permit the selection of objects with reliable colours from the catalogue. The closing demonstrations highlight the use of the survey’s unique  $(r - \text{H}\alpha, r - i)$  diagram for characterising stellar populations and selecting emission-line objects. More comprehensive applications of IPHAS can be found in the works of Sale et al (submitted), which applies DR2 to the problem of 3D extinction mapping, and of Sabin et al (in preparation), where the results of a search of the image database for new planetary nebulae is presented.

The current plan is to work toward one further major IPHAS source catalogue, in which the remaining gaps in sky coverage will have been eliminated – observations aimed at replacing data not meeting the IPHAS minimum quality requirements are continuing. We will also examine options to further improve the global calibration, perhaps tightening the accuracy to better than 2%. For example, we have in mind investigating the use of the PanSTARRS photometric ladder (Magnier et al 2013) as a reference set, when it becomes available and we will explore improving source recovery in the most dense fields via the implementation of

PSF fitting in place of aperture photometry. Finally, the next catalogue will detail all the secondary detections to aid time-domain studies.

The data-taking strategy developed for IPHAS has since been reapplied to carry out a companion INT/WFC Galactic Plane survey called UVEX in  $U, g, r, \text{He I}$  (Groot et al. 2009), a survey of the Kepler field in  $U, g, r, i, \text{H}\alpha$  (Greiss et al. 2012), and a survey of the Southern Galactic Plane and Bulge in  $u, g, r, i, \text{H}\alpha$  called VPHAS+ (Drew et al. 2014). The last of these incorporates the digital update of the SHS, offering all the advantages of calibrated photometry across a little over half the SHS footprint. The work presented here stands as a potential template for the catalogues that remain to be generated from these sibling surveys. In prospect from them, whether they are mined separately or together, are the means to ask seamless questions on the contents and structure of the most highly-populated components of the Milky Way.

## ACKNOWLEDGMENTS

The INT is operated on the island of La Palma by the Isaac Newton Group in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias. We are deeply indebted to the ING staff and students for their ongoing support. All data were processed by the Cambridge Astronomical Survey Unit at the Institute of Astronomy in Cambridge. The catalogue presented in this work was assembled at the Centre for Astrophysics Research, University of Hertfordshire, supported by a grant from the Science & Technology Facilities Council of the UK (STFC, ref ST/J001335/1).

Preparation of the catalogue was eased greatly by a number of software packages, including the POSTGRESQL database software, the TOPCAT and STILTS packages (Taylor 2005, 2006), and the Python modules ASTROPY (Astropy Collaboration et al. 2013), NUMPY and SCIPY (Oliphant 2007), MATPLOTLIB (Hunter 2007), IPYTHON (Pérez & Granger 2007), and APLPY. We also made use of the MONTAGE software maintained by NASA/IPAC, and the SIMBAD, VIZIER and ALADIN services operated at CDS, Strasbourg, France (Bonnarel et al. 2000).

Our work made extensive use of several complementary photometric surveys. Our global calibration was aided by the AAVSO Photometric All-Sky Survey (APASS), funded by the Robert Martin Ayers Sciences Fund. The calibration was tested against the Sloan Digitized Sky Survey (SDSS), funded by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The astrometric pipeline reduction made significant use of the Two Micron All Sky Survey (2MASS), which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/ California Institute of Technology, funded by NASA and the NSF. This work includes observations made with the Spitzer Space Telescope, which is operated by the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA.

GB, JED, SES and BTG acknowledge support from the Science & Technology Facilities Council of the United Kingdom (grants: GB and JED ST/J001333/1, SES ST/K00106X/1, BTG ST/I001719/1). HJF and MM-S both acknowledge STFC postgraduate studentships. JED would also like to convey her thanks to the Physics Department of Imperial College London that hosted this project from its inception to 2007 and supported her via a sabbatical year in 2003-4. JF is supported by the Spanish Plan Nacional de I+D+i and FEDER under contract AYA2010-18352. BTG acknowledges funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 320964 (WDTracer). PRG is supported by a Ramón y Cajal fellowship (RYC-2010-05762), and acknowledges support provided by the Spanish MINECO AYA2012-38700 grant. NJW is in receipt of a Fellowship funded by the Royal Astronomical Society of the United Kingdom.

## REFERENCES

- Ahn C. P. et al., 2012, ApJS, 203, 21
- Astropy Collaboration et al., 2013, A&A, 558, A33
- Aungwerojwit A., Gänsicke B. T., Wheatley P. J., Pyrzas S., Staels B., Krajci T., Rodríguez-Gil P., 2012, ApJ, 758, 79
- Barentsen G. et al., 2011, MNRAS, 415, 103
- Barentsen G., Vink J. S., Drew J. E., Sale S. E., 2013, MNRAS, 429, 1981
- Benjamin R. A. et al., 2003, PASP, 115, 953
- Bessell M., Murphy S., 2012, PASP, 124, 140
- Bica E., Dutra C. M., Soares J., Barbuy B., 2003, A&A, 404, 223
- Bok B. J., Bok P., 1941, The Milky way
- Bonnarel F. et al., 2000, A&AS, 143, 33
- Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubelle S., Nanni A., 2012, MNRAS, 427, 127
- Churchwell E. et al., 2009, PASP, 121, 213
- Churchwell E. et al., 2006, ApJ, 649, 759
- Corradi R. L. M. et al., 2008, A&A, 480, 409
- Corradi R. L. M. et al., 2011, MNRAS, 410, 1349
- Corradi R. L. M. et al., 2010, A&A, 509, A41
- Cross N. J. G. et al., 2012, A&A, 548, A119
- Drew J. E. et al., 2014, ArXiv e-prints
- Drew J. E. et al., 2005, MNRAS, 362, 753
- Drew J. E., Greimel R., Irwin M. J., Sale S. E., 2008, MNRAS, 386, 1761
- Finkbeiner A., 2010, A Grand and Bold Thing: An Extraordinary New Map of the Universe Ushering. Free Press
- Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, AJ, 111, 1748
- Georgelin Y. M., Georgelin Y. P., Roux S., 1973, A&A, 25, 337
- Giammanco C. et al., 2011, A&A, 525, A58
- Glazebrook K., Peacock J. A., Collins C. A., Miller L., 1994, MNRAS, 266
- González-Solares E. A. et al., 2011, MNRAS, 416, 927
- González-Solares E. A. et al., 2008, MNRAS, 388, 89
- Greiss S. et al., 2012, AJ, 144, 24
- Groot P. J. et al., 2009, MNRAS, 399, 323
- Hambly N. C. et al., 2008, MNRAS, 384, 637

- Henden A. A., Levine S. E., Terrell D., Smith T. C., Welch D., 2012, Journal of the American Association of Variable Star Observers (JAAVSO), 40, 430
- Hunter D. A., Massey P., 1990, AJ, 99, 846
- Hunter J. D., 2007, Computing In Science & Engineering, 9, 90
- Irwin M., Lewis J., 2001, New Astronomy Reviews, 45, 105
- Irwin M., McMahon R., Walton N., González-Solares E., Hodgkin S., Irwin J., Lewis J., 2005, The Newsletter of the Isaac Newton Group of Telescopes, 9, 8
- Irwin M. J., 1985, MNRAS, 214, 575
- Irwin M. J., 1997, Detectors and data analysis techniques for wide field optical imaging., Rodríguez Espinosa J. M., Herrero A., Sánchez F., eds., pp. 35–74
- Ivezić Ž. et al., 2007, AJ, 134, 973
- Kendrew S. et al., 2012, ApJ, 755, 71
- Kohoutek L., Wehmeyer R., 1999, A&AS, 134, 255
- Lahulla J. F., 1985, A&AS, 61, 537
- Landolt A. U., 1992, AJ, 104, 340
- Lawrence A. et al., 2007, MNRAS, 379, 1599
- Lucas P. W. et al., 2008, MNRAS, 391, 136
- Mampaso A. et al., 2006, A&A, 458, 203
- Manfroid J., 1995, A&AS, 113, 587
- Mayer P., Macák P., 1973, Bulletin of the Astronomical Institutes of Czechoslovakia, 24, 50
- McMahon R. G., Walton N. A., Irwin M. J., Lewis J. R., Bunclark P. S., Jones D. H., 2001, New Astronomy Reviews, 45, 97
- Minniti D. et al., 2010, New A, 15, 433
- Nikolaev S., Weinberg M. D., Skrutskie M. F., Cutri R. M., Wheelock S. L., Gizis J. E., Howard E. M., 2000, AJ, 120, 3340
- Oliphant T. E., 2007, Computing in Science & Engineering, 9, 10
- Padmanabhan N. et al., 2008, AJ, 674, 1217
- Parker Q. A. et al., 2005, MNRAS, 362, 689
- Pérez F., Granger B. E., 2007, Computing in Science and Engineering, 9, 21
- Pickles A. J., 1998, PASP, 110, 863
- Raddi R. et al., 2013, MNRAS, 430, 2169
- Robitaille T. P. et al., 2008, AJ, 136, 2413
- Sabin L. et al., 2013, MNRAS, 431, 279
- Sabin L., Zijlstra A. A., Wareing C., Corradi R. L. M., Mampaso A., Viironen K., Wright N. J., Parker Q. A., 2010, PASA, 27, 166
- Sale S. E., 2012, MNRAS, 427, 2119
- Sale S. E. et al., 2010, MNRAS, 402, 713
- Sale S. E. et al., 2009, MNRAS, 392, 497
- Schlafly E. F., Finkbeiner D. P., 2011, ApJ, 737, 103
- Schlafly E. F. et al., 2012, ApJ, 756, 158
- Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
- Simpson R. J. et al., 2012, MNRAS, 424, 2442
- Skrutskie M. F. et al., 2006, AJ, 131, 1163
- Taylor M. B., 2005, in Astronomical Society of the Pacific Conference Series, Vol. 347, Astronomical Data Analysis Software and Systems XIV, Shopbell P., Britton M., Ebert R., eds., p. 29
- Taylor M. B., 2006, in Astronomical Society of the Pacific Conference Series, Vol. 351, Astronomical Data Analysis Software and Systems XV, Gabriel C., Arviset C., Ponz D., Enrique S., eds., p. 666
- Thompson M. A., Urquhart J. S., Moore T. J. T., Morgan L. K., 2012, MNRAS, 421, 408
- Viironen K. et al., 2009a, A&A, 504, 291
- Viironen K. et al., 2011, A&A, 530, A107
- Viironen K. et al., 2009b, A&A, 502, 113
- Vink J. S., Drew J. E., Steeghs D., Wright N. J., Martin E. L., Gänsicke B. T., Greimel R., Drake J., 2008, MNRAS, 387, 308
- Wesson R. et al., 2008, ApJ, 688, L21
- Witham A. R. et al., 2007, MNRAS, 382, 1158
- Witham A. R., Knigge C., Drew J. E., Greimel R., Steeghs D., Gänsicke B. T., Groot P. J., Mampaso A., 2008, MNRAS, 384, 1277
- Wright N. J., Drake J. J., Drew J. E., Guarcello M. G., Gutermuth R. A., Hora J. L., Kraemer K. E., 2012, ApJ, 746, L21
- Wright N. J. et al., 2008, MNRAS, 390, 929
- Yu N.-P., Wang J.-J., 2012, Research in Astronomy and Astrophysics, 12, 651

## APPENDIX A: CATALOGUE FORMAT

Table A1: Definition of columns in the IPHAS DR2 source catalogue.

#	Column	Type	Unit	Description
1	name	string		Sexagesimal, equatorial position-based source name in the form: JHH-MMSS.ss+DDMMSS.s. The full naming convention for IPHAS DR2 sources has the form "IPHAS2 JHHMMSS.ss+DDMMSS.s", where "J" indicates the position is J2000. The "IPHAS2" prefix is not included in the column.
2	ra	double	degrees	J2000 Right Ascension with respect to the 2MASS PSC reference frame (which is consistent with ICRS to within 0.1 arcsec). The coordinate given is obtained from the astrometric measurement in the r'-band exposure. If the source is undetected in r', then the i' or H $\alpha$ -band coordinate is given.
3	dec	double	degrees	J2000 Declination. See comments above.
4	sourceID	string		Unique identification number of the detection. Identical to rDetectionID if the source was detected in the r-band; identical to iDetectionID or haDetectionID otherwise.
5	posErr	float	arcsec	Astrometric fit error (RMS). Be aware that the error might be significantly larger than the RMS near CCD edges.
6	l	double	degrees	Galactic longitude $\ell$ converted from ra/dec (IAU 1958 system).
7	b	double	degrees	Galactic latitude $b$ converted from ra/dec (IAU 1958 system).
8	mergedClass	short		Image classification flag based on all bands (1=galaxy, 0=noise, -1=star, -2=probableStar, -3=probableGalaxy, -9=saturated). Computed using the UKIDSS scheme.
9	mergedClassStat	float		Merged N(0,1) stellarness-of-profile statistic. Computed using the UKIDSS scheme.
10	pStar	float		Probability that the source is a star (value between 0 and 1).
11	pGalaxy	float		Probability that the source is a galaxy (value between 0 and 1).
12	pNoise	float		Probability that the source is noise (value between 0 and 1).
13	pSaturated	float		Probability that the source is saturated (value between 0 and 1).
14	rmi	float	mag	( $r - i$ ) colour, formed by subtracting columns r and i. Included in the catalogue for convenience only. To obtain the uncertainty, take the root of the sum of the squares of columns rErr and iErr.
15	rmha	float	mag	( $r - H\alpha$ ) colour, formed by subtracting columns r and ha. See comments above.
16	r	float	mag	Default r-band magnitude using a 2.3 arcsec diameter aperture. Calibrated in the Vega system.
17	rErr	float	mag	Uncertainty for r. Does not include systematic errors.
18	rPeakMag	float	mag	Alternative r-band magnitude derived from the peak pixel height (i.e. a 0.3x0.3 arcsec square aperture). Calibrated in the Vega system.
19	rPeakMagErr	float	mag	Uncertainty in rPeakMag. Does not include systematics.
20	rAperMag1	float	mag	Alternative r-band magnitude using a 1.2 arcsec diameter aperture. Calibrated in the Vega system.
21	rAperMag1err	float	mag	Uncertainty in rAperMag1. Does not include systematics.
22	rAperMag3	float	mag	Alternative r-band magnitude using a 3.3 arcsec diameter aperture. Calibrated in the Vega system.
23	rAperMag3err	float	mag	Uncertainty in rAperMag3. Does not include systematics.
24	rGauSig	float	pixels	RMS of axes of ellipse fit in r.
25	rEll	float		Ellipticity in the r-band.
26	rPA	float	degrees	Position angle in the r-band.
27	rClass	short		Discrete image classification flag (1=galaxy, 0=noise, -1=star, -2=probableStar, -3=probableGalaxy, -9=saturated).
28	rClassStat	float		N(0,1) stellarness-of-profile statistic.
29	rErrBits	short		Bitmask used to flag a bright neighbour (1), source blending (2) and saturation (8).
30	rMJD	double	days	Modified Julian Date at the start of the r-band exposure.
31	rSeeing	float	arcsec	Average Full Width at Half Maximum (FWHM) of stars in the same CCD frame.
32	rDetectionID	string		Unique identifier of the r-band detection in the format "#run-#ccd-number", i.e. composed of the INT telescope run number, the CCD number and a sequential source detection number.
33	rX	float	pixels	Pixel coordinate of the source in the r-band exposure, in the coordinate system of the CCD.
34	rY	float	pixels	Pixel coordinate of the source in the r-band exposure, in the coordinate system of the CCD.
35	i	float	mag	Default i-band magnitude using a 2.3 arcsec diameter aperture. Calibrated in the Vega system.
36	iErr	float	mag	Uncertainty for i. Does not include systematic errors.

**Table A1 – continued**

#	Column	Type	Unit	Description
37	iPeakMag	float	mag	Alternative i-band magnitude derived from the peak pixel height (i.e. a 0.3x0.3 arcsec square aperture). Calibrated in the Vega system.
38	iPeakMagErr	float	mag	Uncertainty in iPeakMag. Does not include systematics.
39	iAperMag1	float	mag	Alternative i-band magnitude using a 1.2 arcsec diameter aperture. Calibrated in the Vega system.
40	iAperMag1err	float	mag	Uncertainty in iAperMag1. Does not include systematics.
41	iAperMag3	float	mag	Alternative i-band magnitude using a 3.3 arcsec diameter aperture. Calibrated in the Vega system.
42	iAperMag3err	float	mag	Uncertainty in iAperMag3. Does not include systematics.
43	iGauSig	float	pixels	RMS of axes of ellipse fit.
44	iEll	float		Ellipticity.
45	iPA	float	degrees	Position angle.
46	iClass	short		Discrete image classification flag (1=galaxy, 0=noise, -1=star, -2=probableStar, -3=probableGalaxy, -9=saturated).
47	iClassStat	float		$N(0,1)$ stellarness-of-profile statistic.
48	iErrBits	short		Bitmask used to flag a bright neighbour (1), source blending (2) and saturation (8).
49	iMJD	double	days	Modified Julian Date at the start of the single-band exposure.
50	iSeeing	float	arcsec	Average Full Width at Half Maximum (FWHM) of stars in the same CCD frame.
51	iDetectionID	string		Unique identifier of the i-band detection in the format "#run-#ccd-#number", i.e. composed of the INT telescope run number, the CCD number and a sequential source detection number.
52	iX	float	pixels	Pixel coordinate of the source, in the coordinate system of the CCD.
53	iY	float	pixels	Pixel coordinate of the source, in the coordinate system of the CCD.
54	iXi	float	arcsec	Position offset of the i-band detection relative to the ra column. The original i-band coordinates can be obtained by computing (ra+iXi/3600, dec+iEta/3600).
55	iEta	float	arcsec	Position offset of the i-band detection relative to the dec column. See comments above.
56	ha	float	mag	Default H-alpha magnitude using a 2.3 arcsec aperture. Calibrated in the Vega system.
57	haErr	float	mag	Uncertainty for ha. Does not include systematic errors.
58	haPeakMag	float	mag	Alternative H-alpha magnitude derived from the peak pixel height (i.e. a 0.3x0.3 arcsec square aperture). Calibrated in the Vega system.
59	haPeakMagErr	float	mag	Uncertainty in haPeakMag. Does not include systematics.
60	haAperMag1	float	mag	Alternative H-alpha magnitude using a 1.2 arcsec diameter aperture. Calibrated in the Vega system.
61	haAperMag1err	float	mag	Uncertainty in haAperMag1. Does not include systematics.
62	haAperMag3	float	mag	Alternative H-alpha magnitude using a 3.3 arcsec diameter aperture. Calibrated in the Vega system.
63	haAperMag3err	float	mag	Uncertainty in haAperMag3. Does not include systematics.
64	haGauSig	float	pixels	RMS of axes of ellipse fit.
65	haEll	float		Ellipticity
66	haPA	float	degrees	Position angle.
67	haClass	short		Discrete image classification flag (1=galaxy, 0=noise, -1=star, -2=probableStar, -3=probableGalaxy, -9=saturated).
68	haClassStat	float		$N(0,1)$ stellarness-of-profile statistic.
69	haErrBits	short		Bitmask used to flag a bright neighbour (1), source blending (2) and saturation (8).
70	haMJD	double	days	Modified Julian Date at the start of the single-band exposure.
71	haSeeing	float	arcsec	Average Full Width at Half Maximum (FWHM) of stars in the same CCD frame.
72	haDetectionID	string		Unique identifier of the H-alpha detection in the format "#run-#ccd-#number", i.e. composed of the INT telescope run number, the CCD number and a sequential source detection number.
73	haX	float	pixels	Pixel coordinate of the source, in the coordinate system of the CCD.
74	haY	float	pixels	Pixel coordinate of the source, in the coordinate system of the CCD.
75	haXi	float	arcsec	Position offset of the H-alpha detection relative to the ra column. The original Ha-band coordinates can be obtained by computing (ra+haXi/3600, dec+haEta/3600).
76	haEta	float	arcsec	Position offset of the H-alpha relative to the ra column. See comments above.
77	brightNeighb	boolean		True if a very bright star is nearby (i.e. errBits $\geq 1$ ). This indicates that the source might be spurious, or the photometry unreliable.

**Table A1 – continued**

#	Column	Type	Unit	Description
78	deblend	boolean		True if the source was blended with a nearby neighbour due to crowding (i.e. errBits $\geq 2$ ). Although a deblending procedure is applied when measuring the photometry, the result may be unreliable (e.g. colours should not be trusted).
79	saturated	boolean		True if the source is saturated (i.e. peak pixel $> 55000$ counts) in one or more bands (i.e. errBits $\geq 8$ ). The photometry of saturated stars is affected by systematic errors.
80	errBits	short		Maximum value of (rErrBits, iErrBits, haErrBits).
81	nBands	short		Number of bands in which the source is detected (equals 1, 2 or 3).
82	reliable	boolean		True if: errBits $\leq 2$ & nBands == 3 & r > 13 & i > 12 & ha > 12.5 & rErr < 0.1 & iErr < 0.1 & haErr < 0.1 & (abs(r-rAperMag1) < 3*hypot(rErr,rAperMag1Err)+0.03) & (abs(i-iAperMag1) < 3*hypot(iErr,iAperMag1Err)+0.03) & (abs(ha-haAperMag1) < 3*hypot(haErr,haAperMag1Err)+0.03).
83	veryReliable	boolean		True if: reliable & pStar > 0.9 & errBits = 0.
84	fieldID	string		Human-readable IPHAS field number and observing run (e.g. 0001o_aug2003).
85	fieldGrade	string		Internal quality control score of the field. One of A, B, C or D.
86	night	integer		Night of the observation (YYYYMMDD). Refers to the UT date at the start of the night.
87	seeing	float	arcsec	Maximum value of (rSeeing, iSeeing, haSeeing).
88	ccd	short		CCD-chip number on the Wide Field Camera (WFC) of the Isaac Newton Telescope (INT). 1, 2, 3 or 4.
89	nObs	short		Number of repeat observations of this source in the survey.
90	sourceID2	string		SourceID of the object in the partner exposure (if obtained within 10 minutes of the primary detection).
91	fieldID2	string		FieldID of the partner detection (e.g. 0001o_aug2003).
92	r2	float	mag	r-band magnitude in the dithered partner field, i.e. the dithered repeat measurement obtained within 10 minutes (if available).
93	rErr2	float	mag	Uncertainty for r2.
94	i2	float	mag	i-band magnitude in the dithered partner field, i.e. the dithered repeat measurement obtained within 10 minutes (if available).
95	iErr2	float	mag	Uncertainty for i2.
96	ha2	float	mag	H-alpha magnitude in the dithered partner field, i.e. the dithered repeat measurement obtained within 10 minutes (if available).
97	haErr2	float	mag	Uncertainty for ha2.
98	errBits2	integer		Error bitmask for the partner detection. Used to flag a bright neighbour (1), source blending (2), saturation (8), vignetting (64), truncation (128) and bad pixels (32768).