Optical Cartography of the Northern Galactic Plane

Hywel John Farnhill

Submitted to the University of Hertfordshire in partial fulfilment
of the requirements of the degree of Doctor of Philosophy

November 2014

Abstract

Counting stars as a means of studying the structure of the Milky Way has a long history, which has progressed significantly with the undertaking of large-area surveys. Photographic surveys have been supplanted with the advent of CCD technology by digital surveys, which provide improved data quality allowing better calibration and fainter limits to be probed reliably.

The INT/WFC Photometric Hα Survey of the Northern Galactic Plane (IPHAS) provides broad-band r' and i' photometry down to 20^{th} magnitude at Galactic latitudes $|b| < 5°$. In this work I make use of the opportunity that IPHAS photometry provides to create stellar number density maps of the Northern Galactic Plane.

I produce preliminary maps which are used to identify and exclude poor quality data during the preparation of the second data release of the survey (DR2). By crossmatching IPHAS against the AAVSO Photometric All-Sky Survey (APASS), I derive transformations between the two photometric systems, and measure the per-IPHAS-field magnitude shifts needed to bring the two surveys in line before a global calibration can be applied. Repeating the crossmatching approach between IPHAS and the Sloan Digital Sky Survey (SDSS), I derive transformations between the two surveys and assess their agreement before and after global photometric calibration, in order to gauge the improvement achieved.

The effects of incompleteness begin to affect the fainter end of any photometric survey as a consequence of confusion and sensitivity limits. I present the application of artificial source insertion on every broad-band IPHAS DR2 image in order to measure the impact of incompleteness across the entire survey. These measurements are used to construct incompleteness-corrected density maps down to magnitude limits of $r' \lesssim 19$ and $i' \lesssim 18$ at an angular resolution of 1 arcminute. These maps represent a unique data product which has applications in studies of Galactic structure and extinction.

I perform a cluster search on the i'-band density map, which in addition to returning 71 known clusters, identifies 29 overdensities unassociated with any known clusters. I compare the stellar densities given by my maps to those in simulated versions of the Milky Way generated by models of Galactic population synthesis. I examine the Gaia Universe Model Snapshot (GUMS), a catalogue which predicts the sky as may be observed by the Gaia mission. In order to make meaningful comparisons between GUMS and IPHAS I determine transformations between the two photometric surveys. The results of the comparison are mixed. I also make use of the 2003 Besançon model of Galactic population synthesis to generate catalogues of synthetic photometry along three sightlines in the IPHAS footprint in order to test different 3D extinction prescriptions. The lowest Galactic longitudes ($\ell \approx 30°$) prove to be particularly challenging to emulate, suggesting 3D mapping of optical extinction in the Galactic Plane is not yet a mature art. The main problem appears to be one of underprediction of the obscuration.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 History of Galactic Astronomy

The field of Galactic Astronomy as it stands today began to take shape with the discovery by Galileo (1610) that the band of light that makes up the Milky Way in fact consists of an enormous number of stars, invisible to the naked eye, stating

> *for the Galaxy is nothing else but a mass of innumerable stars planted together in clusters. Upon whatever part of it you direct the telescope straightway a vast crowd of stars presents itself to view; many of them are tolerably large and extremely bright, but the number of small ones is quite beyond determination.*

Wright (1750) suggested that regions appearing as clouds to us were "*external creations*" bordering upon our own (i.e. that the Milky Way was just one of many large systems). Kant (1755) proposed that a system such as the Milky Way would, under the effect of its rotation and gravitational pull, form a disc-like structure analogous to our own Solar System.

In an attempt to map out the Galaxy, Herschel (1785) undertook the impressive task of counting stars along 683 sightlines, forming a circle running perpendicular to the Galactic plane and passing through the Galactic poles. The result of this endeavour is illustrated in Figure 1.1.

In the early 20th century, two conflicting models of the Galaxy were proposed: Kapteyn (1922), based on the analysis of star counts from photographic plates, placed the Sun near (≈ 650 pc from) the centre of the Galaxy, while Shapley (1918), measuring the distribution of globular clusters across the sky, estimated a distance of ≈ 15 kpc to the Galactic Centre, which he placed towards Sagittarius. Shapley & Curtis (1921) highlight the debate between the two authors over the nature of spiral nebulae; Shapley argued that they were located within the Milky Way while Curtis proposed that they were far beyond the extent of our Galaxy, pointing to the high concentration (and faint magnitude) of novae in the direction of M31. Hubble (1922) showed that these nebulae were indeed extragalactic, later using Cepheid variables to place M31 at a distance of 300 kpc.

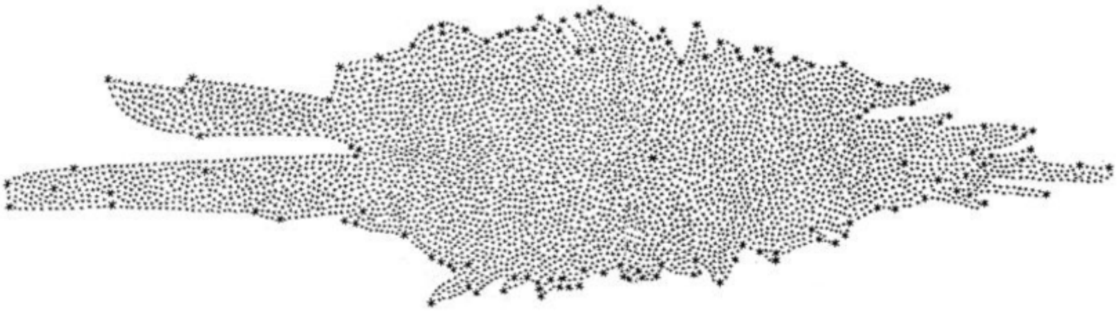


Figure 1.1: Map of the Milky Way from Herschel (1785). The bright star near the centre of the diagram denotes the position of the Sun. Vertical distance corresponds roughly to vertical distance above/below the Galactic midplane, while the horizontal direction corresponds to the line passing through the Sun, connecting Sirius and the Aquila region ($\ell \approx 227°$ and $\ell \approx 30°$).

Baade (1944) commented on the apparent difference in colour between the brightest stars in M31 compared to M32 and NGC 203, classifying the stellar population typified by that in the solar neighbourhood (and M31) *type I*, and the older, redder, population *type II*. Baade (1951) made the connection that type/population I stars appear preferentially along spiral arms, advocating the observation of O & B stars to trace out the structure of the Milky Way.

1.2 Parameterising the Milky Way

1.2.1 The Main Components of our Galaxy

The modern-day picture of the Milky Way, summarised by Turon et al. (2008), is of a barred spiral galaxy with either two or four spiral arms (see Section 1.2.8). Its baryonic content consists of stars, dust and gas, which are distributed across a number of components.

The oldest population of stars are distributed in a spherical halo with a radial density power-law profile of $\rho \propto r^{-2.8}$ (Jurić et al., 2008). This component is not perfectly smooth; $\approx$ 40% of its luminosity is contributed by substructure potentially consisting of debris of disrupted satellite galaxies (Bell et al., 2008).

The central region of the Milky Way is dominated by the bulge, a mainly old population that extends out to a Galactocentric radius $\approx$ 3 kpc (Binney & Dehnen, 1997) with ongoing star formation at its inner region. This region also contains a rotating central bar with a semi-major axis of $\approx$ 4 kpc (López-Corredoira et al., 2007).

The majority of the mass in the Milky Way is believed to be contained in a dark matter halo, with a mass constrained by radial velocity dispersions (of objects at large Galactocentric radii - globular clusters, giants, field horizontal branch stars) of $\approx 1 \times 10^{12}$ $M_\odot$ (Battaglia et al., 2005). These components do not impact on this work, hereafter the focus will be on the disc component of the Galaxy.

	l (pc)	h (pc)	R_{max} (kpc)
Thin	$2600 \pm 520^{\mathrm{a}}$	$300 \pm 60^{\mathrm{a}}$	$15 \pm 2^{\mathrm{b}}$, $> 25^{\mathrm{c}}$, $13.0 \pm 0.78^{\mathrm{d}}$
Thick	$3600 \pm 720^{\mathrm{a}}$	$900 \pm 180^{\mathrm{a}}$	–
HI	3150^{e}	150^{e}	$\approx 35^{\mathrm{e}}$
CO	1800^{f}	46.7^{g}	20^{f}

Table 1.1: Parameters of the galactic disc. (a) Jurić et al. (2008), (b) Ruphy et al. (1996), (c) Momany et al. (2006), (d) Sale et al. (2010), (e) Kalberla & Dedes (2008), (f) Wouterloot et al. (1990) (l_{CO} at $R_G > 13$ kpc), (g) Langer et al. (2014) (interpretation of Sanders et al. (1984) h_{CO} measurements at $3 < R_G < 8$ kpc).

The majority of the *baryonic* mass of the Milky Way is contained in the Galactic disc; a population of stars which shows an exponential decrease in density with both Galactocentric radius and vertical distance from the midplane of the disc. A seemingly older population within the disc, displaying a greater scaleheight, has been termed the 'thick disc', in contrast to the 'thin disc' which dominates at lower vertical distance from the plane (and therefore in the solar neighbourhood). This thesis works mainly in this domain ($|b| < 5°$); this section summarises the primary parameters and current understanding of the thin disc. A discussion of the thick disc and its origins can be found in Section 1.3.2.

Hereafter, the variable R_G refers to the radial distance of an object from the origin in the Galactocentric cylindrical coordinate system; the height above/below the origin and the azimuth will be denoted by z and ϕ respectively. The terms which determine the exponential drop off of stellar density with R_G and z (scalelength and scaleheight) will be denoted by l and h, modified to l_{thick} and h_{thick} when explicitly referring to the parameters of the thick disc. The galactocentric radius of the Solar Circle is denoted by R_0.

1.2.2 Disc Scalelength

It has long been known that the luminosity distributions in spiral galaxies display an exponential drop off with galactocentric radius (Patterson, 1940). A wide range of values have been obtained for the exponential scale-length of the Milky Way using a variety of tracers.

From measurements of the parameters of edge-on spiral galaxies, van der Kruit & Searle (1982) noted that if the disc model they obtain applies to the Milky Way, then based on the kinematics in the solar neighbourhood, the radial scale-length of the Galactic disc would be ≈5 kpc. Bahcall & Soneira (1984), observing five fields in different directions, obtained a lower limit of 2.5 kpc for l. Adopting a scaleheight of 325 pc (see Section 1.2.4), van der Kruit (1986) obtained a scalelength of 5.5 ± 1.0 kpc. Lewis & Freeman (1989) measured the kinematics of 600 disc K giants: on assuming a constant relation between velocity dispersions in the radial and vertical directions throughout the disc, they arrived at a value of $l = 4370 \pm 320$ pc.

Later studies saw shorter scalelengths. Kent et al. (1991) used the surface brightness profile of the Milky Way derived from a 2.4 μm map of the Northern Galactic Plane to fit an exponential scalelength of 3.0 kpc. Robin et al. (1992) observed three fields towards the Galactic Anticentre with long exposures down to $V = 23$, and used an exponential density law out to the furthest stars in their sample (assuming a density of zero beyond this R_G) to obtain a fit with $l = 2.5$ kpc. Ruphy et al. (1996) obtained a similar value, $l = 2.3 \pm 0.1$ kpc, using star counts from the DENIS project.

With the advent of large-scale sky surveys like DENIS, it became possible to study Galactic structure over large areas in a more comprehensive and uniform fashion. Robin et al. (2003) used DENIS data to arrive at a value of $l = 2.53 \pm 0.11$ kpc, which they used in their model of Galactic population synthesis. Jurić et al. (2008) made use of the large amount of data available from SDSS, looking mainly out of the Galactic Plane, to measure the number density distribution of stars. Using the distribution of M dwarfs in the solar neighbourhood (< 2 kpc), they obtained a best fit scalelength of 2.6 kpc. Sale et al. (2010) fit two exponential scalelengths to the A-type stellar density profile obtained from IPHAS, looking towards the Galactic Anticentre, finding an inner $l = 3.02 \pm 0.22$ kpc and an outer $l = 1.20 \pm 0.31$ kpc, transitioning at $R_G = 13.0 \pm 0.78$ kpc.

The scalelength of the thick disc has also been the subject of several studies. Robin et al. (1996) investigated the thick disc by fitting a number of model hypotheses to star counts compiled from a range of sources, finding a scalelength of 2.8 ± 0.8 kpc. Reylé & Robin (2001) obtained $l_{thick} = 2.5$ kpc using optical and infrared star counts. Robin et al. (2003) adopt this value in their model of Galactic population synthesis. Not all estimates agree so closely - Chiba & Beers (2000) obtain an estimate of $l_{thick} = 4.5 \pm 0.6$ kpc using the motions of metal poor stars in the solar neighbourhood. Jurić et al. (2008) obtained an estimate of $l_{thick} = 3.6$ kpc from SDSS data.

1.2.3 Disc Truncation

Studies of galaxies have modelled the observed surface brightness as a truncated exponential (van der Kruit & Searle, 1981) - i.e. assuming the brightness drops to zero beyond some galactocentric radius $\mathrm{R_{max}}$. van der Kruit & Searle (1982) studied four edge-on spirals to find a mean value of $\mathrm{R_{max}} = 4.2 \pm 0.6\ l$. Observing 20 face-on spirals, van der Kruit (1988) found that at $R = 4.5 \pm 0.1\ l$ the contours tracing out surface brightness become more closely spaced, indicating a reduction in scalelength. They proposed that this is the $\mathrm{R_{max}}$ observed in edge-on galaxies.

Truncations in the Galactic disc have since been fitted to observations of stellar populations. Robin et al. (1992) observed down to $V = 23$ in the anticentre, deducing the existence of a sudden drop off in stellar density at a distance between 5.5 and 6 kpc (corresponding to $R_G = 14$ kpc, or $5.6 \pm 0.6\ l$). Ruphy et al. (1996) found a similar cutoff at 15 ± 2 kpc (corresponding to $6.5 \pm 1.2\ l$). Reylé et al. (2009) found that their data were well fit by an ellipsoidal disc edge - i.e. that $\mathrm{R_{max}}$ varies with galactic azimuth: depending on the region considered, values of $12 \leq \mathrm{R_{max}} \leq 14$ gave the best fit to their 2MASS star counts.

Robin et al. (1992) noted that in the case of inside-out formation (see Section 1.3.3), younger objects would still be found at $R_G > 14$ kpc. Indeed, this has been observed: Carraro et al. (2010) detected O- & B- type stars out at $15 \leq R_G \leq 20$ kpc, finding a smooth decline in the stellar density profile. Momany et al. (2006) observed 2MASS red giants out to $R_G = 25$ kpc, and argued that their model of the galactic warp and flare could explain previously detected truncations. Lopez-Corredoira et al. (2002) also suggested that their flare could account for the apparent cutoff in the outer disc, and failed to find evidence for a sudden cutoff at $R_G < 15$ kpc.

Sale et al. (2010) compared the space density distribution of early A-type stars to the distribution predicted by the Besançon model (Robin et al., 2003), finding that rather than a sudden truncation in stellar density, that $\mathrm{R_{max}}$ is better viewed as a transition from an inner scalelength to a shorter outer scalelength at $R_G = 13.0 \pm 0.8$ kpc.

The molecular disc of the Galaxy extends to $R_G \approx 20$ kpc, while HI has been observed out to greater galactocentric radii (Wouterloot et al., 1990). This suggests a natural edge to the stellar content of the disc (Brand & Wouterloot, 2007, and references therein).

1.2.4 Disc Scaleheight

The stellar density falls off in an exponential fashion not only with galactocentric distance but also with vertical displacement above and below the Galactic midplane. As with scalelengths (Section 1.2.2), numerous studies have attempted to measure the Galactic scaleheight.

Faber et al. (1976) obtained a vertical scaleheight of 300 pc using the local space density of M dwarfs. Gilmore & Reid (1983) found an exponential scaleheight from the galactic midplane of ≈300 pc for main sequence stars between $100 \lesssim z \lesssim 1000$, while identifying a thick disc scaleheight of ≈1450 pc between $1000 \lesssim z \lesssim 5000$ pc. Bahcall & Soneira (1984) obtained a similar scaleheight of 350 ± 50 pc for disc dwarfs, and a shorter one of 250 pc from a dataset containing a disproportionately high fraction of giants. Kent et al. (1991) fit an exponential scaleheight of ≈247 pc in the solar neighbourhood, decreasing to ≈165 pc at $R_G = 5$ kpc. Haywood et al. (1997b) selected stellar populations of fixed ages towards the galactic poles, arguing for a smaller value, though they cast doubt on the choice of an exponential scaleheight.

Vallenari et al. (2006) fit thin disc components of different ages, finding a constant scalelength but varying scaleheight, increasing from 200 pc (for stars with ages of 1-3 Gyr) to 350 pc (for stars between 7-10 Gyr old). The increase in scaleheight with age agrees with the results of Bonatto et al. (2006), who obtained a scaleheight of 48 ± 3 pc for open clusters younger than 200 Myr, rising to 150 ± 27 pc between 200 Myr and 1 Gyr. Joshi (2007) find a similarly small value of $h = 61.4^{+2.7}_{-2.4}$ pc for their observed distribution of OB stars.

The thick disc, as its name suggests, must display larger scaleheights than the thin disc. Reid (1993) used faint star counts towards the Northern Galactic Pole to obtain a thick disc scaleheight of 1400-1600 pc. Reylé & Robin (2001) obtained a somewhat smaller $h_{thick} = 800$ pc, which Robin et al. (2003) used in their model of Galactic population synthesis. Jurić et al. (2008) found that a model with $h = 300 \pm 50$ pc and $h_{thick} = 900 \pm 180$ pc provided the best fit to SDSS data.

1.2.5 Disc Warp

Kerr (1957) observed an asymmetry in the distribution of HI gas at positive and negative Galactic latitudes, in an attempt to characterise the modulation of the plane of the Galaxy across the disc. Kerr (1957) commented that the Southern Galactic Plane slopes downwards towards the Magellanic Clouds. Westerhout (1957), also observing the distribution of neutral hydrogen, found deviations from the midplane of more than 300 pc in the outer disc. Henderson et al. (1982) presented maps of the distance of the centroid of the HI distribution from the Galactic Plane, showing a clear deviation of up to 2.4 kpc above the $b = 0°$ midplane in the North and up to -0.8 kpc in the South. Kerr (1983) found that the maximum warp lies along the line connecting Galactic azimuth $\phi = 80°$ to $\phi = 260°$.

Miyamoto et al. (1988) observed that O- & B-type stars in the Northern Hemisphere follow the warp as traced out by HI, inclined by $i \approx 3°$, beyond the Solar Circle but failed to detect systematic kinematic motions related to the warp. Carney & Seitzer (1993) made observations in the Southern Hemisphere, claiming to detect an excess of stars in the expected area of the warp, though warning against over interpretation of their limited sample. Porcel & Battaner (1995) also suggested that the distance above the Galactic midplane of their sample of nine OB stars in the warp region ($240° < \phi < 275°$), in addition to sixteen as a control group (at $\phi \approx 200°$), the expected direction of nodes in the warp), provided an unambiguous detection of the southern stellar warp. Reed (1996) improved on this detection by using nearly 1300 OB stars, finding that the warp reaches 1.5 kpc below the $b = 0°$ Galactic midplane at a galactocentric radius of 15 kpc in the direction of $\phi = 260°$.

Using low angular resolution maps of the Galactic plane in 10 wavelengths from DIRBE, Freudenreich et al. (1994) saw evidence for a warp in the dust distribution similar to that in HI, and noted that in the population of K and M giants at near-infrared wavelengths that the stellar warp demonstrated the same phase but smaller amplitude than the FIR-observed gaseous warp. Freudenreich (1998) fit a model of the Galactic Bulge to DIRBE observations, finding that the free parameter corresponding to the Galactic azimuth of the line of nodes lies at $\phi \approx 0°$. Drimmel & Spergel (2001) found similar results when fitting their model to DIRBE data, finding $\phi \approx 0°$ for the dust distribution while changing which model parameters were allowed to vary. The onset of the warp was found to occur at $R_G \approx 7$ kpc, with the dust warp amplitude given as 72.8 pc kpc^{-2} for dust and 27.4 pc kpc^{-2} for stars. Gyuk et al. (1999) modelled the absolute offset from the $b = 0°$ plane, z_{w}, as a quantity linearly increasing beyond the Solar Circle, dependent on the Galactocentric azimuth ϕ, i.e.

$$z_{\mathrm{w}} = \gamma_{\mathrm{w}}(R_G - R_0)\sin\phi \tag{1.1}$$

where the factor γ_{w} is found to be 0.214 from model comparisons with observed Hubble Space Telescope (HST) star counts in seven low-latitude fields. This description of the warp is used in the Besançon model of Galactic population synthesis (Robin et al., 2003), with a slightly revised $\gamma_w = 0.18$.

Lopez-Corredoira et al. (2002), modelling the Galactic warp as a series of tilted rings, studied the distribution of red giants in 2MASS data to find an azimuthal angle of $\phi = -5^\circ \pm 5^\circ$ for the line of nodes of the warp. They found a good agreement between the warp as traced by this old population and the gas warp. The asymmetry of the pulsar distribution in the Galactic Plane was studied by Yusifov (2004), whose fit to it implies a warp amplitude of $\approx$1 kpc at the outer edge of the Galaxy, and an angle of $\phi = 14.5^\circ$ for the line of nodes. Momany et al. (2006), also using 2MASS red giants, found the observations agreed well with the parameters obtained by Yusifov (2004) with the southern stellar warp amplitude twice that of the northern warp; this could be a consequence of the Sun not lying on the line of nodes of the warp. However Momany et al. (2006) also warn that the extent of the northern warp might be masked as a result of extinction, due to its location beyond the Norma-Cygnus (Outer) arm.

The Milky Way is not the only galaxy in which warping of the disc has been observed; Sánchez-Saavedra et al. (2003) found that 60% of their sample of 150 spiral galaxies displayed signatures of a warp and concluded that, taking into account detectability with respect to inclination of their galaxies, the vast majority of spiral galaxies are warped.

Several theories exist on the cause of galactic warps. Kerr (1957) observed that the Milky Way warp is oriented towards the Magellanic Clouds, but stated that their gravitational effects are insufficient to account for the degree of warping. Levine et al. (2006b) measured the HI warp out to 20 kpc, finding a symmetric warp out to a galactocentric radius of 15 kpc in both hemispheres - Weinberg & Blitz (2006) argued that these results could be explained by gravitational perturbations of the Magellanic Clouds. Intergalactic magnetic fields were proposed as the cause by Battaner et al. (1990), while López-Corredoira et al. (2002) argued for infall onto the disc from the intergalactic medium as a possible origin of the warp. Sanchez-Salcedo (2006) argued against both of these theories, on the basis that the magnetic warps would become spiral-shaped within a few galactic rotations, and that the required infall of matter with zero net angular momentum is far too high to explain the observed Galactic warp.

1.2.6 Disc Flare

While the warp affects the effective vertical centroid of the Galactic Plane at large galactocentric radii, the shape of the disc is also modified by the flaring of the disc; this results in increased scaleheight with galactocentric radius. Lozinskaya & Kardashev (1963) observed an increase in disc thickness in their 21 cm observations, measuring disc flaring up to a thickness of $\approx$1 kpc, and suggested this effect may be independent of the warp. Henderson et al. (1982) also observed the Galactic flare, measuring a thickness of nearly 2 kpc in the outer disc. Wouterloot et al. (1990) saw the flaring in both HI and CO observations, with the vertical thickness of the molecular distribution roughly doubling in the galactocentric radius range $10 < R_G < 17$ kpc.

In fitting profiles to the luminosity distribution observed by the Infrared Telescope, flown aboard the Space Shuttle in 1985, Kent et al. (1991) had to model the Milky Way with a constant scaleheight out to $R_G = 5.3$ kpc and a linearly increasing scaleheight thereafter. Gyuk et al. (1999), comparing HST star counts to models of Galactic structure, modelled the flare in the same way as Kent et al. (1991), multiplying the scaleheight in the Solar neighbourhood by the factor k_{flare}:

$$k_{\mathrm{flare}} = 1 + \gamma_{\mathrm{flare}}(R_G - R_0) \tag{1.2}$$

where γ_{flare} was determined to be 0.53 kpc^{-1}. A slightly modified version of this approach is used in the Besançon model of Galactic population synthesis (Robin et al., 2003), with R_0 replaced by R_{flare}, the galactocentric radius at the onset of the flare. A curious feature in the literature is that Robin et al. (2003) cite the Gyuk et al. (1999) value as $\gamma_{\mathrm{flare}} = 5.4 \times 10^{-4}$ kpc^{-1}, while Reylé et al. (2009) cite the same value as $\gamma_{\mathrm{flare}} = 5.4 \times 10^{-5}$ pc^{-1}; these values disagree with each other as well as with the value presented by Gyuk et al. (1999).

Drimmel & Spergel (2001) included a linear flare in their models, finding a flare of 14.8 pc kpc^{-1}, starting at $R_G = 4.4$ kpc. Lopez-Corredoira et al. (2002) saw a slower decrease of densities of 2MASS red giant at larger R_G, indicating the presence of the flare, choosing to model the increase in scaleheight with R_G as an exponential and measuring a scalelength of the flare (l_{flare}) of 3.4±0.4kpc. They found that the flaring begins inside the Solar Circle, but at a much longer flaring scalelength ($l_{flare} \approx 9$ kpc). Momany et al. (2006), however, found an almost constant scaleheight within the Solar Circle. Their three populations of 2MASS red giants at heliocentric distance (d) values of 2.8, 7.3 & 16.6 kpc sampled both the thick and thin discs. In the thin disc regime (the $d = 2.8$ kpc sample), their measurements agree with the model of Lopez-Corredoira et al. (2002).

1.2.7 A note on stellar metallicities

The metallicities of stellar populations serve as another tool with which to probe the structure and evolution of the Milky Way; McWilliam (1997) provides an overview of Galactic metallicity trends.

It has been observed in the solar neighbourhood that older stars tend to display lower metallicities (Twarog, 1980); this trend can be accounted for by successive rounds of star formation enriching their local interstellar medium. Rocha-Pinto et al. (2000a) derived such an age-metallicity relation (AMR) from late type dwarfs, measuring an increase of ≥ 0.56 dex from an initial disc metallicity of -0.70 dex.

The metallicity of an object is not determined solely by its age: Smartt & Rolleston (1997) observed a radial oxygen abundance gradient of -0.07 ± 0.01 dex kpc^{-1} in early B-type stars in the range $6 \lesssim R_G \lesssim 18$ kpc. Carraro et al. (1998) found a radial gradient of -0.09 dex kpc^{-1} and a vertical metallicity gradient of -0.25 dex kpc^{-1}. Nevertheless they stressed that their sample of open clusters is small, making it difficult to determine a vertical gradient for any one given Galactocentric radius.

Bovy et al. (2012b) used data from the SEGUE spectroscopic surveys to pick out stellar populations with given abundances, and determined that mono-abundance populations can be described by single-exponential density profiles in both R_G and z.

1.2.8 Spiral structure

The spiral structure of galaxies is known to correlate with star formation; concentrations of young star forming complexes trace out spiral arms (Morgan et al., 1953). Old stellar populations, observed in the near-infrared, show little variation in spiral structure with Hubble type, leading extragalactic studies of spiral arms to focus on younger populations (see e.g. Seigar & James, 1998).

The number of arms present in the Milky Way has been an area of some debate (Vallée, 1995, 2008, and references therein). Georgelin & Georgelin (1976) traced four spiral arms in the Milky Way using Hα emission (tracing HII), while Sofue & Fujimoto (1983) argued for a two-armed spiral structure based on Faraday rotation measures of extragalactic sources. Drimmel (2000) made use of DIRBE maps in both K and 240 μm bands to trace the stellar and gaseous content of the disc. He found that the stellar emission in the K-band traced out a two-arm structure, while the 240 μm emission from dust followed a four-arm structure.

Russeil (2003) created a catalogue of star forming complexes, determining the position, distances (kinematic and stellar where possible), and kinematics for each complex. She used the complexes with stellar distances to obtain the relation between velocity and distance, allowing more accurate determinations of kinematic distances. Two-, three-, and four-arm models were fit to the distribution of complexes, with results appearing to exclude the two-arm model and to favour slightly the four-arm model.

Reid et al. (2009a) measured the parallaxes and proper motions of 18 masers, which they found traced spiral arms. They favoured a four-arm model, and found that star forming regions rotate around the Galaxy at 15 km s^{-1} slower than would be expected in simple circular orbits. In order to reconcile their four-arm, 16° pitch angle model with the two arms visible in the Spitzer Glimpse results (Benjamin, 2008), Reid et al. (2009a) suggest the possibility of two dominant spiral arms populated by both young and old stars, and two weaker arms traced by young stars. Levine et al. (2006a) used data from the Leiden/Argentine/Bonn HI survey to trace spiral arms in the gas content of the disc out to $R_G = 25$ kpc.

The origin of spiral arms in galaxies has generated a great deal of interest. One proposed mechanism for the development of spiral structure is tidal excitation, either by an external companion or internally via a central bar (Kormendy & Norman, 1979; Bottema, 2003). Others advocate the idea that spirals can be excited by the disc itself, either by quasi-steady (Lin & Shu, 1964) or transient (Toomre, 1981) processes. Fujii et al. (2011) used N-body simulations to show that self-regulating mechanisms in a disc can maintain spontaneously formed spiral arms on a cosmological timescale. The currently mooted mechanisms potentially responsible for the occurrence of spiral arms are discussed at length by Dobbs & Baba (2014).

1.2.9 Gas and dust

The gas and dust content of the Milky Way is primarily concentrated in the thin disc. The content of the ISM is complex, encompassing a wide range of physical states (McKee & Ostriker, 1977). The atomic components occur in both cold and warm states (Heiles & Troland, 2003); the lines present in the cold neutral media (the diffuse clouds with a median $T = 70$ K) have provided the most insight into the nature of the Milky Way.

HI, detected through its 21-cm line (Ewen & Purcell, 1951), is an invaluable probe of Galactic structure and dynamics. Its natural line width is extremely narrow, allowing measures of its broadening to be used to study the velocity of HI gas along the line-of-sight (providing a probe of Galactic rotation). It has been used to study the midplane of highest gas concentrations (Westerhout, 1957) - indeed, determinations of the $b = 0°$ plane were in part informed by such HI maps (Gum & Pawsey, 1960). Thanks to the relative lack of obscuration due to dust at radio wavelengths, HI is a useful probe of structure; Oort et al. (1958) associated narrow regions of high HI density in their measurements with the spiral structure observed in external galaxies. Several efforts have been made to map the Galaxy at this wavelength (Heiles, 1975; Burton & Hartmann, 1994).

The entire sky was mapped in HI by the Leiden/Argentine/Bonn (LAB) survey, at an angular resolution of 0.6° (Kalberla et al., 2005). Using this data, Kalberla & Dedes (2008) measured HI in the disc out to ≈ 35 kpc, and estimated a radial scalelength of 3.15 kpc, and a scaleheight of 3.9 kpc. Levine et al. (2006a) used surface densities from the LAB survey to show that, coincident with spiral arms, increases in surface density correspond with regions of reduced scale height of the gas.

HII regions, which trace out regions of star formation in the disc, have been studied via recombination lines at radio wavelengths and via Hα emission.

In contrast to atomic hydrogen, its molecular state is comparatively difficult to observe. H_2 in the ISM is located in cold environments where molecules can form without being quickly destroyed. While H_2 is difficult to observe through a lack of suitable line emission, these regions contain other, more easily detectable molecules such as ^{12}CO (Dame et al., 2001). At 115 GHz the $J = 1 \rightarrow 0$ line traces out molecular clouds, indirectly providing an estimate of the H_2 column density. Almost all star formation takes place in these regions - tracing out molecular clouds is extremely important for understanding Galactic structure. Dame et al. (2001) produced a map in CO of the entire Milky Way which has been used in studies of extinction (e.g. Marshall et al., 2006) and spiral arm structure (e.g. Pettitt et al., 2014). The scaleheight of molecular clouds has been found to be less than that of atomic gas; Stark & Lee (2005) proposed that their result was a result of the molecular cloud formation process. Smaller clouds condense from atomic clouds, resulting in a slight reduction in scaleheight, while giant molecular clouds form rapidly due to the influence of passing spiral arms causing a reduction in velocity dispersion, thereby significantly reducing the Galactic latitudes at which these structures are observed.

Towards the inner Galaxy, the density of gas peaks at $R_G \approx 4$ kpc in a molecular ring, which Clemens et al. (1988) estimated contains 39% of the molecular content of the Milky Way. Portinari & Chiosi (1999) suggested that the peak at this Galactocentric radius is due to some dynamical process, preferring this explanation to an IMF which varies across the disc (see Section 1.4.1). They proposed that the Galactic bar sweeps up material which collects at the inner Lindblad resonance, thereby explaining the high density of gas in the region $3.5 \lesssim R_G \lesssim 4.5$ kpc.

The distribution of dust in the Milky Way resembles that of HI. Contributing ≈1% of the ISM mass, the dust has been mapped using infrared observations of its emission. Schlegel et al. (1998) generated all-sky maps of integrated extinction based on DIRBE and IRAS data. Drimmel & Spergel (2001) also used DIRBE data, but instead generated a three-dimensional distribution of dust at low Galactic latitudes. Marshall et al. (2006) used the colour of 2MASS stars combined with the Besançon model of Galactic population synthesis (Robin et al., 2003) to generate extinction curves for sightlines in the range ($|\ell| \leq 100°$, $|b| \leq 10°$). Relying solely on IPHAS photometry, Sale et al. (2014) constructed a 3D map of extinction which traces the dust distribution in the disc.

Dwek (2005) listed a number of observables that can help to constrain the properties of dust, the primary effect of dust impacting this work is the first item in their list: the extinction and reddening of starlight. This is discussed in more depth in Section 1.4.3. The two primary forms of dust grains are carbon- and oxygen-rich, identified by their spectral features (in both absorption and emission). Draine (2003) summarises the state of knowledge regarding constraints on dust properties, including composition.

1.2.10 Rotation law

The structure of the disc at any one point in time is not sufficient information for a full understanding of its history or to model its change over time.

In order to accurately describe non-circular motions (such as those brought on by the presence of spiral arms), the underlying rotation of the Galaxy must be accounted for. Oort (1927b) derived the Oort constants - quantities which are used to connect the observable relative motions of stars to the rotation curve of the Milky Way, assuming rotation about a central axis (i.e. the Galactic Centre) where the circular rotation speed is dependent on galactocentric radius. Oort (1927a) obtained values for the constants of $A = 19 \pm 3$ km s^{-1} kpc^{-1} and $B = -24 \pm 5$ km s^{-1} kpc^{-1} which have since been refined, e.g. by measurements of Hipparcos proper motions; Feast & Whitelock (1997) measured $A = 14.82 \pm 0.84$ km s^{-1} kpc^{-1} and $B = -12.37 \pm 0.64$ km s^{-1} kpc^{-1}, implying a slowly declining rotation curve at R_0.

Mapping the kinematics of the disc has primarily been achieved by measuring the large-scale distribution of HI and CO. Inside the solar circle, the HI tangent point method can be used, whereby only the radial velocity of HI is needed to determine the rotation curve. As a result, the kinematics of the inner disc have been well studied (Kwee et al., 1954; Shane & Bieger-Smith, 1966). Beyond the solar circle the tangent point method cannot be applied, thus independent measurements of both distances and velocities are required to characterise the rotation curve. For this reason the rotation law in the outer Galaxy is less well-constrained (Fich et al., 1989).

Fich et al. (1989) used CO velocities and spectrophotometric distances of HII regions to investigate the rotation curve out to a galactocentric radius of 17 kpc. They observed a rise in the rotation curve at large R_G, and attempted to produce a function that described this behaviour while simultaneously capturing the flat rotation curve of the inner Galaxy. They concluded that a flat curve with a circular rotation speed (Θ) of 220 km s^{-1} should be used to determine kinematic distances in the outer Galaxy, assuming $R_0 = 8.5$ kpc and $\Theta(R_0) = \Theta_0 = 220$ km s^{-1} (the current IAU recommended values). Brand & Blitz (1993) found a similar shape to the rotation law, but opted to fit a power law to retain the slight rise in rotational velocity at large R_G.

The distance to the Galactic Centre is a source of considerable uncertainty in studies of the rotation curve. Reid (1993) collected a number of estimates of R_0, and arrived at a best value of 8.0 ± 0.5 kpc by taking the weighted average of R_0 obtained through nearly independently calibrated methods. A range of values have continued to be put forward, ranging from 7.1 ± 0.4 kpc from considerations of local stellar kinematics (Olling & Merrifield, 1998) to $8.70^{+0.57}_{-0.43}$ kpc by comparing models to stellar populations observed with 2MASS and OGLE-II (Vanhollebeke et al., 2009). Reid et al. (2009b) obtained a distance of $7.9^{+0.8}_{-0.7}$ kpc from the trigonometric parallax of Sgr B2. From fitting the orbits of B-type stars within $0.5''$ of the Galactic Centre, Eisenhauer et al. (2005) obtained an R_0 of 7.62 ± 0.32 kpc, while Trippe et al. (2008) used the kinematics of evolved late-type stars close to the Galactic Centre to measure $R_0 = 8.07 \pm 0.35$ kpc. Ghez et al. (2008) measured $R_0 = 8.0 \pm 0.6$ kpc from the orbit of star S0-2, further constraining R_0 to 8.4 ± 0.4 kpc in the case that the central black hole has no massive companion.

The value of Θ is not independent of R_G, therefore the adopted value of R_0 will affect the determination of Θ_0. The ratio Θ_0/R_0 can be determined to a greater accuracy than either parameter alone (Reid et al., 2009a). Feast & Whitelock (1997) obtain a value of $\Theta_0/R_0 = 27.19 \pm 0.8$ km s^{-1} kpc^{-1} using Hipparcos proper motions of Galactic Cepheids, while Reid et al. (2009a) recovered a higher value of 30.3 ± 0.9 km s^{-1} kpc^{-1} using trigonometric parallaxes of star forming regions. Both are higher than the value recovered using the IAU recommended constants, which yield 25.9 km s^{-1} kpc^{-1}.

The result of Reid et al. (2009a) was recovered using their values of $\Theta_0 = 254 \pm 16$ km s^{-1} and $R_0 = 8.4 \pm 0.6$ kpc. They note that the resulting Θ_0/R_0 ratio agrees with the value obtained based on the proper motion of Sgr A* (Reid & Brunthaler, 2004). They also note that kinematic distances of objects determined using the IAU values tend to lie further than their distances determined by trigonometric parallax, arguing that their adopted values bring the kinematic distances in line with independently obtained distances. Bovy et al. (2009) reanalysed the data of Reid et al. (2009a) and obtained a lower estimate of $\Theta_0 = 244 \pm 13$ km s^{-1}. (McMillan & Binney, 2010) also reanalysed the same maser data, concluding that the solar velocity with respect to Θ_0 should be revised upwards to 11 km s^{-1} from previous estimates (e.g. 5.2 ± 0.6 km s^{-1} from Dehnen & Binney (1998)), resulting in a lower estimate of Θ_0 of 232 ± 24 km s^{-1} (for the case of a flat rotation curve). Bovy et al. (2012a) used the line-of-sight velocities of stars from the APOGEE survey to derive an approximately flat rotation curve with $\Theta_0 = 218 \pm 6$ km s^{-1}, and recommended a higher still value for the solar motion of 26 ± 3 km s^{-1}.

As the exact values of the parameters involved in describing the rotation curve are somewhat uncertain in the outer disc, it is clear that relying on kinematic distances can result in significant uncertainties - independent methods are required to determine the distances beyond the solar circle. Any improvement in the parameters involved would aid studies of Galactic motion and structure (e.g. studies of spiral arms). The astrometry provided by Gaia will provide the parallaxes of more than 100 million stars with accuracy better than 20% (Luri et al., 2014) which, along with radial velocities measured with the on-board radial velocity spectrometer, will enable the rotation curve to be studied to great precision.

1.3 Evolution of the Milky Way and its disc

1.3.1 Formation

The current understanding of galaxy evolution points to hierarchical formation processes (Cole et al., 2000, and references therein). In the early universe, fluctuations in density combined with gravitational instabilities to form the initial dark matter haloes, in which baryonic matter cooled into discs in their gravitational potential (White & Rees, 1978). These early galaxies merge to generate the bulges observed in spiral galaxies and bright elliptical galaxies (Cole et al., 2000). The dark halo of the Milky Way is quite extended, reaching beyond the distance of the Magellanic Clouds - Sakamoto et al. (2003) estimated that only a quarter of the mass of the halo ($\approx 5.4^{+0.1}_{-0.4} \times 10^{11}\ M_\odot$) lies within a radius of 50 kpc. Using N-body and smooth particle hydrodynamic simulations, Governato et al. (2007) reproduced the formation of disc galaxies inside dark matter haloes for a Λ cold dark matter universe, reproducing reasonable rotation curves and a dark matter distribution within the Solar Circle in agreement with observations.

The stellar halo of the Milky Way was the first baryonic component to form; some studies argue that the halo consists of tidally shredded globular clusters (e.g. Gnedin & Ostriker, 1997), while others favour accretion of satellite galaxies as the primary formation mechanism (e.g. Bell et al., 2008). The bulge, which is in fact a *pseudobulge* (classical bulges being those seen in early-type spirals which more closely resemble elliptical galaxies), is believed to have formed via secular processes in the disc; Kormendy & Kennicutt (2004) discuss these possible pseudobulge formation mechanisms in detail.

Evidence has been put forward for the tidal interactions of galaxies; Toomre & Toomre (1972) cited many such examples of seemingly interacting galaxies, and showed that their gravitational forces explain adequately the wide range of bridges and tails observed between them. Ibata et al. (1994) made the first detection of a dwarf satellite galaxy accreting onto the Milky Way. They observed a group of comoving stars on the far side of the Galaxy in the direction of Sagittarius in the process of being tidally disrupted.

1.3.2 The thick disc

While studying lenticular galaxies, Burstein (1979) & Tsikoudi (1980) both found that in addition to the typical thin disc and bulge necessary to account for the observed luminosity distributions, a thicker disc-like structure was needed to explain the fainter regions of the luminosity profiles. van der Kruit & Searle (1981) observed a similar thick disc in the vertical profiles of the spiral galaxies NGC 891 and 4565.

Gilmore & Reid (1983) found evidence for an analogous thick disc in the Milky Way, in a study of the luminosity function towards the South Galactic Pole. They fit two exponential density laws: one at $100 \lesssim z \lesssim 1000$ pc, and a second at $1000 \lesssim z \lesssim 5000$ pc, belonging to the newly discovered thick disc. They estimated that $\approx 2\%$ of stars in the Solar neighbourhood would belong to this thick disc. They also noted that stars belonging to the thick disc tended to be older and more metal-poor, implying a different origin for this population than for thin disc stars. Bahcall & Soneira (1984) re-analysed the work of Gilmore & Reid (1983) and suggested that their misclassification of giants as being on the main sequence was responsible for the apparent thick disc in the Milky Way.

Carney et al. (1989), based on considerations of the $|W|$ velocity distribution of their sample, picked out a thick disc population with metallicity and kinematics which lie between those of the thin disc and halo populations. They prefer a formation scenario for the thick disc involving a merger, occuring while the metallicity of the disc was still low ($[\mathrm{Fe/H}] \lesssim -0.8$). Gilmore et al. (1995) attempted to measure the vertical metallicity gradient in the thick disc, finding no gradient in the iron abundance distribution at $500 < z < 3000$ pc above the plane. They observed element ratios in the thick disc indicative of Fe enrichment from Type Ia supernovae, requiring a formation time scale of a few Gyr, with their data indicating an age of $\gtrsim 12$ Gyr. Several studies have reported that the thick disc rotation lags behind solar rotation, arguing for lags of $\approx$-20 km s^{-1} (Norris, 1987), -50 km s^{-1} (Soubiran et al., 2003), up to -100 km s^{-1} (Wyse & Gilmore, 1986).

The formation of a thick disc resulting from interaction with satellite galaxies is still one of several proposed mechanisms. Quinn et al. (1993) showed using numerical simulations that the observed thick disc could be accounted for by such mergers, heated from an ancient pre-existing thin disc. Quillen & Garnett (2001) found that stars with ages between 3 and 9 Gyr old show little sign of heating or dispersion, with a sudden increase at 9 Gyr, which they explain as the result of a minor merger. Brook et al. (2004), running an N-body chemodynamical simulation, found that the majority of their stars which formed at $8.5 < age < 10.5$ Gyr are thick disc stars with high velocity dispersion, consistent with the observations of Quillen & Garnett (2001).

Abadi et al. (2003), who found that more than half of their thick disc stars had their origin in tidal debris of satellites, argue for direct accretion of thick disc stars. This is also the favoured process of Yoachim & Dalcanton (2006), whose analysis of 34 late-type disc galaxies could not be reconciled with a kinematically heated thin disc; they propose a series of minor mergers in which the stellar components accreted form the thick disc while accreted gas cools and forms the thin disc. Bournaud et al. (2007) proposed a formation mechanism involving the infall of large amounts of gas into an unstable disc, consistent with observations of the kinematics of "chain galaxies" containing turbulent clumps of gas.

Meanwhile Roškar et al. (2013) demonstrated that the majority thickening of the Galactic disc could be accounted for by the effects of radial migration - however they require a certain amount of external perturbation to the Milky Way to fully match observations. Kroupa (2002) argued that direct accretion of material from satellite galaxies may not be necessary to explain the thick disc, and that massive star clusters could provide the necessary heating to add the observed kinematically hot components to the Milky Way.

In the provocatively titled "*The Milky Way has no distinct thick disk*", Bovy et al. (2012b) argued that while external accretion/heating mechanisms would produce distinct thick and thin discs, internal mechanisms should not lead to easily separable populations. They argued that previous studies contain a degree of circular reasoning by presupposing that the thin and thick discs are two different components of Galactic structure. They put forward the suggestion that future works should no longer define sub-samples of Galactic population using structural or dynamical markers - they advocate using elemental abundances as tags for sub-populations, and using this approach find no evidence of bi-modality that would describe distinct thick and thin discs.

The classification of stars as belonging to the thin or thick disc nevertheless continues to be used - Haywood et al. (2013) adopted the separations in the $[\alpha]/[\mathrm{Fe}]$ vs $[\mathrm{Fe}]/[\mathrm{H}]$ plane presented by Adibekyan et al. (2012). They pointed out, however, that their use of "thick" and "thin" disc was made for lack of better terminology. They concluded that their data were best accounted for by a thick disc that formed from a turbulent gaseous environment with a large vertical velocity dispersion, which decreased over several Gyr - i.e. star formation in the thick disc occurred in progressively thinner layers. The thin disc then began to form (at $R_G < 10$ kpc) in the environment left at the thick disc phase. They commented that the thin and thick disc at $R_G < 10$ kpc have more in common than the thin disc on either side of $R_G = 10$ kpc.

Robin et al. (2014) used their population synthesis approach to investigate Galactic populations and found that their thick disc was best simulated with an extended star formation period, which included a contraction. These results agree with the formation scenario presented by Bournaud et al. (2009), of thick disc formation during a period of high gas turbulence at high redshift. Schönrich & Binney (2009) developed a model of chemical evolution constrained by the metallicity distribution observed by the Geneva-Copenhagen Survey, and found that low-metallicity stars at large z above/below the plane at $\approx R_0$ originated at smaller Galactocentric radii, having migrated outwards.

1.3.3 Inside-out formation

Star formation continues in the thin disc, a timescale much longer than the traditionally mooted formation of the thick disc. Wood & Oswalt (1998) & Knox et al. (1999) used observations of white dwarfs to find a thin disc age of ≈ 10 Gyr.

Chiappini et al. (1997) emphasised that this extended timescale makes it unlikely that the thick disc is the primary source of gas feeding this ongoing formation, but rather that additional gas is accreted from the intergalactic medium. Their models assume that the formation of the thin disc occurs in an inside-out fashion - i.e. that star formation in the outer disc is fuelled by this accretion. The consequence of this model is that the rate of star formation would depend on Galactocentric radius. Trujillo & Pohlen (2005) searched for disc truncations in Hubble Ultra Deep Field data, and found that the average disk truncation radius had increased by $\approx 1-3$ kpc in the last 8 Gyr.

The observations by Yuan et al. (2011) of the metallicity of a face-on spiral at $z \approx 1.5$, showing a steep metallicity gradient, support the inside-out formation model in which the disc is enriched by the infall of material for the 9 Gyr after the formation of its central regions. This result agrees with simulations run by Pilkington et al. (2012), who find that inside-out evolution causes a variety of initial conditions to lead to present-day abundance gradients.

Evidence for inside-out growth has been seen in local galaxies; Williams et al. (2009) used the HST to observe four fields along the major axis of the disc of M33, finding an increasing scalelength with time, increasing from 1.0 ± 0.1 to 1.8 ± 0.1 kpc between 5 and 10 Gyr ago. Barker et al. (2007) made similar observations at larger R_G but found that the mean age increased with radius. Williams et al. (2009) explain this as being the result of radial mixing in the disc, which (beyond some break in the exponential disc profile) causes the opposite effect compared to that seen in the inner disc. Chiappini et al. (2001) predict an increase in abundance gradients in the outermost regions of the disc, due to the Galactic halo dominating at these large R_G rather than due to radial mixing.

Schönrich & Binney (2009) used models of chemical evolution of the Milky Way incorporating radial mixing, and obtained good correspondence to local stars without needing to invoke inside-out formation (i.e. they assumed simultaneous formation of the entire disc). Haywood et al. (2013) studied populations that spanned the metallicity regimes typically assigned to the thick and thin discs, and showed that the oldest outer thin disc stars are as old as the thick disc. They argued that an inside-out formation mechanism would lead to a radial gradient in metallicity, an effect that is not observed. Haywood et al. (2013) also point out that the predominantly 2 Gyr old population of the outer disc predicted by Roškar et al. (2013) is at odds with open clusters older than 4 Gyr seen out to $R_G = 15$ kpc.

1.3.4 The star formation rate

In order to model the present day populations of the Milky Way (the thin disc in particular), the rate at which star formation has occurred as a function of time needs to be understood.

This star formation rate (SFR) can be investigated using the current age distribution; Twarog (1980) studied the photometry of F dwarfs to find a roughly constant SFR over the lifetime of the thin disc. Barry (1988) used chromospheric ages of local dwarfs to find an increase in the SFR during the last 400 Myr.

Rocha-Pinto et al. (2000b) used simulations to investigate the effect of age-dependent spread in SFR, claiming to find with a high degree of certainty that the Milky Way could not have had a constant SFR during its history. They see several episodes of enhanced star formation, lasting 1-2 Gyr each. Just & Jahreiß (2010) assembled a Galactic model which preferred a SFR distribution which reaches a maximum ≈ 10 Gyr ago, declining to the present day. Aumer & Binney (2009) found that this description of SFR history agreed well with their approach of modelling the change in SFR as an exponential decline - their single exponential yielded a surprisingly high disc age between 13.5 and 11.5 Gyr, while a sum of two exponentials gave a lower limit of ≈ 10 Gyr.

1.3.5 Interpreting kinematics and abundances

In studies of Galactic dynamics, stars are largely treated as non-interacting; each star is assumed to accelerate uniformly due to the force exerted by the gravitational potential of the Galaxy, resulting in circular orbits around the Galactic Centre. Once small perturbations are introduced, the behaviour of stars becomes more complex, with perturbations in the radial direction inducing epicyclic behaviour. It has been shown that the rotational velocity of a population of stars lags the circular rotation speed at a given R_G, with the amount of lag proportional to the radial velocity of the population (asymmetric drift, see Section 4.8 of Binney & Tremaine, 2008). This agrees with kinematics of Hipparcos stars; Dehnen & Binney (1998) found a linear relation between the lag and their squared random velocity.

The kinematics of stars are known to correlate strongly with stellar type (Roman, 1950), where older populations display larger velocity dispersions than younger populations. Initially thought to be due to the kinematic imprint of the collapse of the Galactic disc during formation, with younger stars forming after the remaining gas reached a less turbulent state (Eggen et al., 1962), this dispersion is now known to increase from a small (≈ 10 km s^{-1}, (Wielen, 1977)) value over the lifetime of an object.

Dehnen & Binney (1998) used the proper motion and parallaxes of Hipparcos stars to investigate the kinematics of main sequence stars in the solar neighbourhood. They found that the velocity component in the direction of galactic motion correlates with $(B-V)$ colour up to $B-V=0.6$, redward of which the lifetime of main sequence stars is longer than the age of the solar neighbourhood. They also observed an increase in dispersion in all velocity components ($\sigma_U > \sigma_V > \sigma_W$ for all colour bins), which is attributed to scattering in the disc (Jenkins, 1992). Binney et al. (2000) used this increase in velocity dispersion to calculate the age of the solar neighbourhood. Finding that velocity dispersion ($\sigma(\tau) \propto \tau^{\beta}$) increases with age with $\beta = 0.33$, they estimated an age of 11.25 ± 0.75 Gyr, assuming a Salpeter IMF.

The exact mechanism of scattering has been the subject of some debate since early proposals of scattering off gas clouds (Spitzer & Schwarzschild, 1953). (Hanninen & Flynn, 2002) found that simulations of dispersions increased by clouds produced $\beta = 0.21 \pm 0.02$, which were ruled out by dispersion measurements of Aumer & Binney (2009). Spiral arms are an alternative source of heating (Carlberg & Sellwood, 1985); De Simone et al. (2004) found that their simulations produced a wide range of β values $(0.2 - 0.76)$, making it difficult to isolate the mechanism of disc heating. Binney & Lacey (1988) pointed out that molecular clouds tend to follow the spiral arms of the disc, blurring the distinction between the two modes of heating. They proposed that the spiral arms would increase dispersions in the Galactic Plane, while the molecular clouds would scatter stellar populations vertically.

Roškar et al. (2008) simulated the effect of spiral arms, finding that they scatter the Galactic rotation radius of stars, resulting in a surface density profile that transitions from a more gradually declining inner profile to a steeper outer profile. These results confirm the predictions of Sellwood & Binney (2002), who argued that this radial mixing is required in addition to the epicyclic motion about a fixed galactocentric orbit. Schönrich & Binney (2009) incorporated radial mixing into a model of chemical evolution, finding a good fit to star counts as functions of several parameters including [Fe/H].

A spread in the age-metallicity relation has been observed (Edvardsson et al., 1993), which cannot be explained if regions are enriched solely by the stars which form there - i.e. stars located in the solar neighbourhood either formed in an inhomogeneous medium or originated from different regions of the Galaxy. Carraro et al. (1998) also observed this scatter, arguing that chemical evolution models need to reproduce this effect in addition to the overall AMR. Feltzing et al. (2001) reexamined the analysis of previous studies, and pointed out that the observational errors of Edvardsson et al. (1993) were significantly less than the AMR scatter.

The simulations of Roškar et al. (2008) investigated the radial migration of stars over significant distances due to resonant scattering with the spiral arms. They found that the effect of migration provides an explanation for the observed spread in the age-metallicity relation - the age gradient in the Galaxy (see Section 1.3.3) is affected by radial migration, due to stars with a given metallicity indicative of galactocentric formation radius migrating to other R_G.

Radial migration has been proposed as the secular process responsible for the formation of the thick disc (Loebman et al., 2011); migration could explain the fact that thick disc stars in the solar neighbourhood can be found to have galactocentric formation radii of ≈ 4 kpc. Haywood et al. (2013) argued that the high velocity dispersions present in the thick disc (relative to the thin disc) could result in eccentric orbits which could explain the change in R_G of thick disc stars without the need for significant radial migration.

1.3.6 Substructure in the outer Galaxy

The understanding of the warp and flare of the Milky Way are crucial if a full understanding of Galactic structure is to be achieved. An example of the need for these parameters to be well-defined is evident in the form of the Canis Major (CMa) over-density. An apparent over-density of blue stars was first observed by Newberg et al. (2002) in SDSS data, who suggested this could be due to the tidal disruption of a dwarf galaxy within the Milky Way. Martin et al. (2004a), studying the asymmetry of the 2MASS M-giant population, agreed with this assessment. They ran N-body simulations of an interacting dwarf galaxy which successfully reproduced the observed over-density, including the Monoceros Ring structure highlighted by Yanny et al. (2003) (who also supported the idea of a tidally disrupted satellite).

Momany et al. (2004) compared the proper motions of the M-giants in the CMa over-density with surrounding stars in the Galactic disc, and found no peculiar signature indicative of an extragalactic origin. They succeeded in reproducing the over-density with 2MASS star count profiles taking the warp and flare into account. Using the Besançon model (Robin et al., 2003) which includes the warp and flare (see Equations 1.1 & 1.2), they generated synthetic colour-magnitude diagrams which contained features previously highlighted as the product of a tidally disrupted satellite (Martin et al., 2004a). They therefore argued that the CMa over-density is solely the result of Galactic warp and flare.

Martin et al. (2004b) responded to this by measuring the radial velocities of the over-density, claiming that the measured values were significantly higher than those expected at the position of the CMa over-density, with this population displaying a low velocity dispersion. In addition, they saw that the observed over-density of stars towards CMa ($\ell \approx 240°$) begins closer than the heliocentric distance corresponding to their assumed R_G of the warp onset. Martin et al. (2005) highlighted that the radial-velocity peak attributed to a population distinct from the expected Galactic population in Martin et al. (2004b) was due to an issue with template spectra, but maintained that the radial velocity distribution was indicative of a distinct population.

Momany et al. (2006) argued that the remaining radial-velocity evidence could be explained by galactic differential rotation, and that their thorough investigation of the Galactic warp and its effects on 2MASS red giants recover a warp signature that can account for the CMa over-density. While they did not discount the Monoceros Ring as being of extragalactic origin, they suggested that the combination of Galactic warp and flare could contribute significantly at its location. Hammersley & López-Corredoira (2011) reproduced stellar density distributions out to 25 kpc using F and G dwarf counts from SDSS, and found that the Galactic flare alone can account for the Monoceros Ring. However they found a galactocentric radius of 16 kpc for the onset of the flare, a significantly larger radius than found by previous studies.

This controversy highlights the necessity of a thorough understanding of the effect of the warp and flare on Galactic structure.

1.4 Important underpinning aspects of stellar astronomy

1.4.1 The stellar initial mass function

In order to understand and accurately model the stellar population that makes up any component of the Milky Way, the way in which mass is partitioned across the population needs to be understood. This is parameterised by the initial mass function (IMF):

$$\frac{dN}{dM} \propto M^{-\alpha} \tag{1.3}$$

where dN is the number of stars in the mass interval M to $M + dM$ in a given stellar population.

Several estimates of the IMF have been proposed. Salpeter (1955) first proposed an α of +2.3, a value which seems to serve well down to low stellar masses ($\approx 0.5M_{\odot}$). Haywood et al. (1997b) found for the range $1 < \frac{m}{M_{\odot}} < 3$ an IMF with $\alpha \approx +1.8$, dependent on the star formation rate (see Section 1.3.4). Kroupa (2000) proposed a three-part IMF slope, with α values of +0.3, +1.3, and +2.3 for ranges of $0.01 < \frac{m}{M_{\odot}} < 0.08$, $0.08 < \frac{m}{M_{\odot}} < 0.5$, and $0.5 < \frac{m}{M_{\odot}}$ respectively.

Studies have suggested that environment may modify the shape of the IMF; Blum et al. (2003) and Bartko et al. (2010) presented evidence for a top-heavy IMF close to the Galactic Centre. Nevertheless the IMF is often taken as being universal - Czekaj et al. (2014), for example, try fitting IMFs with several different slopes to data while updating the Besançon Galaxy model, assuming that they hold over the entire Milky Way. Bastian et al. (2010) provided an overview of potential variations in the IMF. Universality is most often assumed for the intermediate and higher-mass end of the IMF; at the low-mass/substellar regime the IMF becomes more uncertain (discussed in Section 3.2 of Luhman, 2012).

1.4.2 Multiplicity in stellar systems

The frequency with which multiple component systems occur is of particular interest to studies of galactic populations; in order to realistically model a population then its multiplicity fractions need to be constrained. An accurate picture of the IMF also requires knowledge of how mass is partitioned in stellar systems. Abt & Levy (1976) provided an early estimate, finding two thirds of the stars in their sample (type F3-G2, class IV & V) were in multiple systems. Duquennoy & Mayor (1991), who measured radial velocities for a slightly later-type sample (F7-G9), found 44% of systems had companions which resulted in a mass ratio $M_2/M_1 \geq 0.1$, rising to 57% after correcting for incompleteness effects. They estimated that up to an additional 20% of their sample could have very low mass companions ($0.01 - 0.10M_{\odot}$), leaving a third of local solar-type stars in single systems.

Lada (2006) compiled data from previous studies with the aim of investigating the variation in binary fraction against spectral type. He highlighted the apparent "brown dwarf desert" making it likely that Duquennoy & Mayor (1991) overestimated the low mass companion fraction, adopting their 57% of G-type stars as a primary in a multiple systems as an upper limit. Taking a binary frequency of 30 ± 12% for early M-type stars (Reid & Gizis, 1997) and 9 ± 4% for late M-types (Siegler et al., 2005), a trend of higher binary frequency with increasing mass emerged. The result of Mason et al. (2009), which placed 75% of O-type stars in clusters, is in agreement with these findings.

The mass ratio distribution (MRD) has also been seen to vary with spectral type (Section 1.2.1 Bastian et al., 2010). Determining the distribution is necessary to understand the IMF (Goodwin & Kouwenhoven, 2009); the shape of the observed IMF is determined by the MRD in addition to the binary fraction.

Robin et al. (2012a), using an updated version of the Besançon model, chose to generate multiple systems according to a probability that increased with object mass, with the separation also depending on mass (lower mass leading to smaller separations).

1.4.3 Extinction

The first estimate of the extinction suffered by starlight was made by Kapteyn (1904), who arrived at an estimate of 1.6 mag kpc^{-1}. Trumpler (1930), who estimated distances to open clusters by measuring their angular extent, found their light extinguished by 0.7 mag kpc^{-1} - a value still widely used (e.g. the local extinction normalisation applied by Robin et al. (2003)). He also found that the colour excess of the clusters increased with distance, proposing that the ISM contains dust particles which scatter light resulting in a wavelength dependence of the extinction.

As discussed in Section 1.2.9 the dust distribution in the Milky Way has been mapped using photometry from large-area surveys. Before the advent of these surveys, Burstein & Heiles (1982) generated maps of extinction using a combination of HI column densities and galaxy counts.

Launched in 1989, the Cosmic Background Explorer (COBE) was built with the primary goal of measuring the cosmic microwave background. On board were the Diffuse InfraRed Background Experiment (DIRBE) and the Far InfraRed Absolute Spectrophotometer (FIRAS), two instruments that were used to further the understanding of the gas and dust content of the Milky Way (Dwek et al., 1995; Reach et al., 1995). Schlegel et al. (1998) used DIRBE data to generate all-sky maps of integrated extinction - useful where total extinction along lines of sight are required. The extinction values provided by these maps have since been recalibrated; Schlafly & Finkbeiner (2011) used spectra from SDSS to confirm a 14% recalibration of the Schlegel et al. (1998) values.

For studies working in the Galactic regime, extinction against distance is highly desirable; several attempts were made, as summarised by Fitzgerald (1968), who noted that the available maps contradicted one another. More recently, three dimensional models of extinction have been published taking advantage of large-area surveys. Drimmel & Spergel (2001) used DIRBE data and a model of dust distribution (consisting of a warped and flared disc, spiral arms, and a local spiral arm) to estimate the extinction in three dimensions. Amôres & Lépine

(2005) produced a similar model incorporating spiral arms, using gas column densities and the assumption that dust and gas are well mixed. Marshall et al. (2006) used a combination of 2MASS data and idealised Milky Way predicted by the Besançon model (Robin et al., 2003) to produce extinction curves for sightlines towards the inner Galaxy. Sale et al. (2014) mapped the extinction in the Galactic Plane, applying a hierarchical Bayesian model in conjunction with IPHAS photometry. Lallement et al. (2014) produced maps of the ISM within 1 kpc in the Galactic Plane, and ≈ 300 pc vertically above/below, using colour excesses derived from a number of photometric surveys.

Extinction can lead to difficulty in determining stellar parameters; Bailer-Jones (2011) presented a technique combining multiband photometry and parallaxes, along with prior knowledge of the Hertzprung-Russell Diagram, in order to reduce the degeneracy between extinction and $T_{\rm eff}$. Such a tool will become extremely useful with the arrival of astrometry and photometry from Gaia (see Section 1.5.4).

The understanding of the wavelength dependence of attenuation caused by dust is of utmost importance to astronomy. The results of observations at different wavelengths can only be reconciled if suitable corrections for extinction can be applied. Early work showed that the variation of extinction with wavelength was dependent on the chosen sightline (Hiltner & Johnson, 1956, and references therein). The difference in this relation between sightlines becomes more extreme at shorter wavelengths (w.r.t. the Johnson V-band ($\lambda_{\rm eff} = 5500$Å)), and less significant towards IR wavelengths.

The extinction law of given sightlines were characterised by the ratio of total extinction A_λ to the observed colour excess (most commonly $E(B-V)$). Hiltner & Johnson (1956) found a ratio $R = \dfrac{A_V}{E(B-V)}$ of 3.0 ± 0.2, close to the value, $R = 3.1$, widely adopted for lines of sight that suffer from diffuse extinction only (i.e. do not pass through dense clouds) (Howarth, 1983). Cardelli et al. (1989) parameterised the extinction law between U- (360Å) and L- (3450Å) bands, providing $A(\lambda)/A(V)$ values for a number of common filters in this range.

In most studies extinction has usually been given as the extinction for a photometric band, which is averaged over a range of wavelengths; the spectral energy distribution of a star across this wavelength range can modify the actual extinction measured. To eliminate this effect, monochromatic extinction can be used, defining the amount of extinction at a single wavelength. Recent studies have begun to adopt $A_0 = A(5495\text{Å})$, the monochromatic equivalent to the V-band (Maíz-Apellániz, 2004; Drew et al., 2014; Sale et al., 2014).

1.5 Galactic surveys and IPHAS

In the past twenty years, the amount of data available to studies of the Milky Way has increased enormously, with the proliferation of large area surveys. Surveys focused on the Milky Way have been conducted since the mid 20th century. Amongst the earliest were attempts to study the gas content of the Galaxy (Weaver & Williams, 1973) (see Section

1.2.9). Photographic surveys have been used for over a century to survey large areas of the sky; Gill (1896), for example, compiled a catalogue of southern stars down to a magnitude of 9.5.

The photographic surveys still most heavily used today are those which form the USNO catalogues (Monet et al., 2003), which cover the whole sky. Over the course of 10 years, the Palomar Observatory Sky Survey (POSS) was completed, covering the range $-24° < \delta < 90°$ (Abell, 1959) - this survey covering all of the accessible northern sky and its successors would later be digitised as part of the Digitized Sky Survey (DSS) (Lasker et al., 1996) and included in USNO-B.

The first astronomical image recorded by a CCD was made in 1975 (Howell, 2006), and as CCD detector technology was more widely adopted, the possibility of surveying the sky over a wider dynamic range and with a linear response became feasible. The earliest digital surveys were all-sky; they endeavoured to observe all sources down to their detection limits, rather than focus on specific regions of the sky. The Galactic Plane, while of obvious interest to studies of the Milky Way, is extremely densely populated and hence far more liable to suffer from the effects of confusion. The instruments available for survey use have since improved in terms of resolution and area, allowing the Galactic Plane to be sensibly targeted as the sole focus of dedicated surveys.

1.5.1 The path to IPHAS

The INT/WFC Photometric Hα Survey of the Northern Galactic Plane (IPHAS) was designed to observe Galactic latitudes $|b| < 5°$ (Drew et al., 2005). The Hα emission line can pick out objects in the early and late stages of stellar evolution: for example HII regions, pre-main-sequence stars, Wolf-Rayet stars and PNe, among others. Also observing with broad-band r' and i' filters, the colours provided by IPHAS can be used to pick out main-sequence stars (discussed in Chapter 2).

Several surveys preceded IPHAS in imaging large regions of the sky in the narrow wavelength range covering the Hα emission line. Surveys such as the Virginia Tech Spectral-Line Survey (VTSS; Dennison et al., 1998) in the North and the Southern Hα Sky Survey Atlas (SHASSA; Gaustad et al., 2001) provided low ($0.8'$ and $1.6'$ respectively) resolution measures of Hα emission. Finkbeiner (2003) combined these datasets along with the 1° resolution spectral measurements obtained by the Wisconsin Hα Mapper Northern Sky Survey (WHAM; Haffner et al., 2003), to produce an all-sky map of Hα emission.

Beginning in 1997, the AAO/UKST Hα survey was a photographic survey covering the Southern sky, which was digitised to form SHS (Parker et al., 2005). SHS offered a far higher angular resolution ($1''$) than previous surveys, allowing it to be used to identify point-source Hα emitters in addition to studying Galactic structures on larger scales.

The desire for a northern equivalent to SHS was one of the factors that led to IPHAS (Drew et al., 2005). With its first observations made in 2003, the quality of IPHAS data was a vast improvement over the previous photographic and lower-resolution digital surveys. Figure

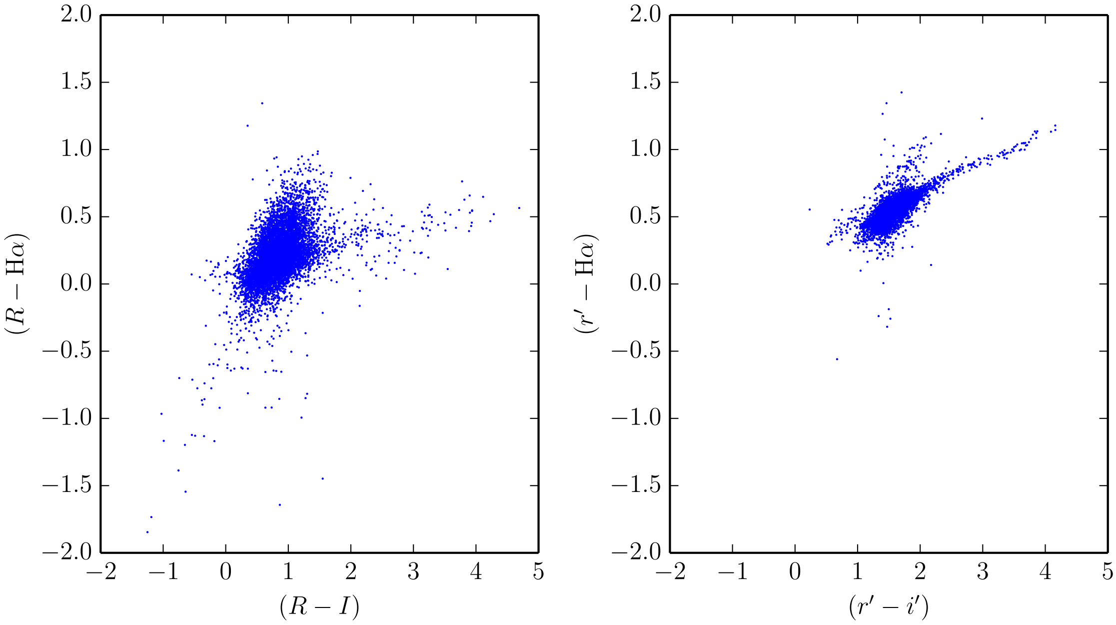


Figure 1.2: Comparison of the colour-colour diagrams of SHS (**left**) and IPHAS (**right**), for a 0.25 sq. deg. region centred on ($\ell = 32°, b = -4°$). The only sources plotted are those which appear in both surveys. The giant branch is more prominent in the IPHAS data, though can be seen in the SHS photometry, while the unreddened main sequence which is clearly defined in IPHAS is lost in the noisy SHS data.

1.2 compares IPHAS and SHS photometry from a region of overlap between the two surveys. Providing digital broad-band r' and i' photometry in addition to Hα, the combination of filters employed by IPHAS allows the effects of reddening and stellar type to be disentangled to a large degree; this is discussed in Chapter 2. Covering latitudes $|b| < 5°$, photometry obtained by IPHAS has enabled wide searches for emission line objects (Witham et al., 2008; Raddi et al., 2013), and the generation of a 3D extinction map (Sale et al., 2014).

Updating the photometry available for the Galactic Plane at shorter wavelengths, the UV-Excess Survey of the Northern Galactic Plane (UVEX) is the blue counterpart to IPHAS, following the same observing pattern but observing in $Ug'r'$ filters. HeI images were taken for half of the survey, with its long exposure times necessitating that the broad-band filters take precedence in order to complete the survey in a timely manner. Its blue filters facilitate studies of white dwarf populations (Verbeek et al., 2013) and together with IPHAS span the optical wavelength range down to $r' < 21$, forming the northern component of the European Galactic Plane Surveys (EGAPS).

1.5.2 Complementary optical and NIR northern surveys

A number of digital surveys complement EGAPS, with some covering Galactic latitudes higher than $|b| = 5°$ in the optical, while others provide a near-infrared view of the sky.

Proposed as a joint photometric and spectroscopic survey almost a decade before it began collecting data (Gunn & Knapp, 1993), Sloan Digital Sky Survey imaging commenced in 2000, observing with *ugriz* filters (Stoughton et al., 2002). The most recent update to its photometric coverage and/or calibration came with the release of SDSS DR9 - the survey now covers $\approx 14,500$ sq. deg and contains more than a billion objects, with approximately a quarter being unique stellar sources. The majority of SDSS photometry is concentrated on the Northern Galactic Cap (i.e. high Galactic latitudes), with the addition of a few additional strips, some of which pass through the Galactic Plane. These strips prove useful for testing the calibration of IPHAS, as will be discussed in Chapter 3.

With the eventual goal of covering the entire sky, the AAVSO Photometric All-Sky Survey (APASS) is an optical survey observing in $BVg'r'i'$ filters in the range $10 \leq V \leq 17$. Its magnitude range enables it to bridge the gap between bright (Tycho) and fainter optical surveys. This overlap with the IPHAS magnitude range makes APASS suitable for uniformly calibrating IPHAS, which will be done in Chapter 3.

Pan-STARRS1 is a survey using a dedicated telescope for a 7 sq. deg. field-of-view in broad-band $grizy$ filters, with plans to expand to a second telescope, currently undergoing commissioning, and to four telescopes thereafter. Pan-STARRS1 has already been used to probe the outer Galactic Disc - Slater et al. (2014) used a $(g - r)$ colour cut to select old ($\approx$ 9 Gyr) stars, then placed cuts on g-band magnitudes to pick out stars at increasing distances. These maps allowed the structure of the outer Galaxy, specifically tidal streams, to be compared with simulations. While Pan-STARRS data is not yet publicly accessible, its calibration (Magnier et al., 2013) presents a possible future cross-calibration tool for further IPHAS data releases.

The baseline present between recent optical surveys and the old photographic surveys make them ideal counterparts for proper motion studies. Munn et al. (2004) published a proper motion catalogue combining SDSS DR1 and USNO astrometry (the latter based on POSS plates), with a typical error of 3 mas yr^{-1}. Bond et al. (2010) used this catalogue to study Galactic dynamics, finding a decrease in rotational velocity and increase in velocity dispersion with distance from the Galactic Plane. In the plane, Deacon et al. (2009) used POSS data in conjunction with IPHAS to generate a catalogue of objects with significant proper motions.

The Two Micron All Sky Survey, a near-infrared (JHK) survey which ran from 1997 to 2001, provided photometry for 471 million point sources (Skrutskie et al., 2006). Detecting sources to a faint 10σ limit of $K_s = 14.3$, 2MASS has been used extensively in the study of red sources, for purposes such as estimating reddening (Cambrésy et al., 2002) and measuring Galactic structure (Lopez-Corredoira et al., 2002; Reylé et al., 2009). 2MASS became an even more powerful tool when combined with optical photometry, providing an SED over the range $0.35 - 2\ \mu$m when combined with SDSS. The two surveys were used together in searches for low mass objects (Metchev et al., 2008). Combining 2MASS with deeper IPHAS photometry has proved useful in studies of young stellar objects (Barentsen et al., 2013), symbiotic stars (Corradi et al., 2008), and planetary nebulae (Viironen et al., 2009).

Running from 2005, the UKIRT Infrared Deep Sky Survey (UKIDSS) encompassed five surveys, three of which have Galactic applications (Lawrence et al., 2007). The Galactic Plane

Survey (GPS) covers the Northern Galactic Plane with the same footprint as IPHAS ($|b| < 5°$) in JHK to a depth of $K = 19$ (Lucas et al., 2008). At low Galactic latitudes UKIDSS largely supersedes 2MASS, having improved spatial resolution and sensitivity. Lucas et al. (2008) showed the utility of combining GPS and IPHAS photometry, enabling spectrophotometric typing. A second epoch of K-band GPS observations followed, with a baseline of 2-8 years, to facilitate studies of proper motions (Smith et al., 2014).

1.5.3 Complementary southern surveys

In the Southern hemisphere, VPHAS+ serves as a combined IPHAS/UVEX counterpart, completing EGAPS coverage of the Galactic Plane (Drew et al., 2014). The footprint of VPHAS+ has been increased towards the Galactic Centre ($|\ell| < 10°$, $|b| < 10°$) to capture the bulge. The high resolution and Hα coverage of VPHAS+ has led to the discovery of nebulae at the arcsecond scale (Wright et al., 2013), and the large field-of-view (1 sq. deg.) proved useful for target selection and the investigation of extinction around Gaia-ESO Survey clusters (Cantat-Gaudin et al., 2014).

A series of surveys operating at VISTA in Chile are working to update the near-infrared view from the Southern hemisphere. The VISTA Hemisphere Survey (VHS) will cover the entire southern sky in J and $K_{\rm S}$, using data from other VISTA surveys to complete its coverage. VHS will provide photometry down to $K_{\rm S} = 18.1$, an increase in ≈ 30 in depth compared to 2MASS (McMahon et al., 2013). In the Galactic Plane in the range ($-65° < \ell < -10°$, $|b| < 2°$) and a 300 sq. deg. region encompassing the bulge, the VISTA Variables in the Vía Láctea (VVV) survey observes in $ZYJHK_{\rm S}$ filters (Minniti et al., 2010). Several studies of the inner Galaxy have already made use of VVV data, including measurements of the Galactic bar (González-Fernández et al., 2012), generation of metallicity maps (Gonzalez et al., 2012), mapping the density profile of the bulge (Wegg & Gerhard, 2013), among others.

SkyMapper is an Australian project surveying the southern sky over six epochs in a $griz$ filter set similar to that of SDSS, alongside uv filters that were chosen to enable easier determination of stellar parameters (Keller et al., 2007). In addition to the primary survey, which will reach $g = 22$, SkyMapper will run a secondary, brighter survey taking 5 second exposures covering 8th to 15th mag, with the overlap range allowing any non-photometric nights of the primary survey to be calibrated to brighter, photometric exposures. Among the primary science goals of SkyMapper is identifying extremely metal-poor stars, useful for the study of the earliest star formation in the Milky Way; a number of such objects have already been found and followed-up (Jacobson et al., 2014).

1.5.4 Future missions

The Hipparcos satellite, launched in 1989, provided an optical all-sky view. The mission generated two primary data products: the Hipparcos catalogue (Perryman et al., 1997) which provided astrometric measurements for an input catalogue of $\approx 180,000$ stars, and the Tycho catalogue (Høg et al., 1997) which provided photometry for more than one million stars brighter than $V = 11.5$, later updated to contain ≈ 2 million sources. These catalogues covered the closest few hundred parsecs, providing a detailed picture of the solar neighbourhood from which kinematics and galactic evolution were tested (Dehnen & Binney, 1998; Binney et al., 2000).

The successor to the Hipparcos and Tycho catalogues will be produced by the Gaia mission. Gaia, launched in 2013, is an astrometric, photometric, and spectroscopic mission designed to characterise more than one billion Galactic and Local Group sources (Perryman et al., 2001). It will provide photometry, parallaxes and proper motions down to a 20th magnitude faint limit and radial velocity information for stars brighter than $V \approx 16.5$.

The Large Synoptic Survey Telescope (LSST) is a planned 8 m telescope, whose camera will consist of 189 CCDs over a 9.6 sq. deg. field-of-view (LSST Science Collaboration, 2009). It will image the Southern sky from Chile every few days, providing multi-epoch *grizy* photometry. At this point it remains unclear how often LSST will visit the Southern Galactic Plane.

Both projects will generate enormous amounts of data; Gaia is estimated to generate 100 Tb over its lifetime (O'Mullane et al., 2007), while LSST will generate half a petabyte of imaging data per month (Becla et al., 2006). Such large datasets will need robust models of the Milky Way in order to extract the greatest amount of science. In preparation for the imminent Gaia datasets, for example, mock catalogues have been generated to test reduction pipelines and predict the likely contents of the real catalogues. This has been done by assuming a certain Galactic population model, namely an updated Besançon model (Robin et al., 2003, 2012a). The reliability of any results obtained by assuming such a model is dependent on how well it reproduces the real sky; validation of the model via the use of external datasets is therefore highly desirable.

1.6 Galactic models

Galactic population models can make testable predictions about the observable Milky Way; the properties of such models are usually motivated by an underlying dynamical model for the structure of the Galaxy, that are then refined empirically to reproduce a certain set of observations.

The model developed by Bahcall & Soneira (1980) is an early example of such a model. Using measured scaleheights and observed luminosity functions, combined with a model of the Galaxy consisting of a single disc component and a spheroid, the authors successfully reproduced star counts in the magnitude range $4 \lesssim V \lesssim 22$. Ortiz & Lepine (1993) modelled the Galactic stellar populations as the sum of four components: thick and thin discs, a spheroidal component (encompassing the bulge and halo), and spiral arms described by a thin disc of O- & B-type stars exhibiting overdensities along four spiral structures. This model was intended to reproduce star counts in several photometric bands; the authors claimed to closely reproduce $K-$band star counts, but required an enhancement in the spiral arm overdensities relative to their preferred parameters to achieve satisfactory match to 12 μm and 25 μm IRAS bands.

More recently models have been developed which incorporate the evolutionary history of the Milky Way, taking into account the IMF, AMR and star formation history among other parameters. Using evolutionary tracks, the stellar populations of the Milky Way are aged to the current time. Girardi et al. (2005) continue to develop their "TRIdimensional modeL of thE GALaxy" (TRILEGAL), which includes thin and thick discs and the halo. Their model manages to reproduce both faint 2MASS counts and the bright Hipparcos sample. However they warn against applying their model towards the Galactic Centre, due to their lack of bulge component, and at $|b| < 10^\circ$ due to the complex extinction in the Galactic Plane. TRILEGAL offers the ability to generate synthetic catalogues via a web interface[1], offering more flexibility in several areas compared to the Besançon implementation discussed below and in Chapter 6 (e.g. choice of IMF and disc scale lengths), and some limitations (e.g. no option to limit stellar types that are included).

1.6.1 The Besançon model

The most widely applied and fully-featured model for Galactic population synthesis is the Besançon model (Robin et al., 2003). Its history can be traced back to Robin & Creze (1986), who presented an algorithm to generate objects along sightlines defined by their Galactic coordinates. The model has since been improved, incorporating a dynamically self-consistent model of mass density (taking into account both baryonic and dark matter), constrained from star counts along several sightlines (Bienaymé et al., 1987). Further studies of the thick disc have also informed the model (Robin et al., 1996; Reylé & Robin, 2001); both scaleheights and ages of this component have been reduced from the values used by Robin & Creze (1986).

Similarly to TRILEGAL, the Besançon model of Robin et al. (2003) took into account the evolutionary history of the Milky Way, assuming an IMF, historical SFR and metallicity gradients for each of the components of the Galaxy. The model split the thin disc into seven age bins (and an eighth white dwarf population), each with their own local mass density and velocity dispersions. Disc kinematics were informed by Hipparcos constraints as determined by Gomez et al. (1997), and its warp and flare were characterised as discussed in Sections 1.2.5 and 1.2.6.

The Besançon model has been used extensively: applications include producing extinction maps when combined with 2MASS star counts (Marshall et al., 2006), testing the shape of the Galactic bulge (Robin et al., 2012b), and investigating the presence of stellar overdensities in the outer Galaxy (Momany et al., 2004; Williams et al., 2011).

The 2003 model has been updated by Czekaj et al. (2014) to allow the SFR and IMF of the stellar populations to vary as a free parameter, and replaced the stellar evolutionary model of Haywood et al. (1997a) with three sets of tracks for low-, high- and very high-mass stars. The best fit to Tycho-2 data requires a decreasing SFR, a departure from the constant SFR

---

[1] http://stev.oapd.inaf.it/cgi-bin/trilegal

of the Robin et al. (2003) thin disc. This updated approach also generates binary systems, assigning a star to be a single star or the primary of a binary system with a probability depending on its mass and class. These improvements were applied in the generation of a mock-Galaxy intended to simulate the Milky Way as may be seen by Gaia (Robin et al., 2012a).

1.7 This thesis

From Sections 1.2-1.4 it is apparent that the structure and history of the Milky Way is by no means certain. The current and upcoming generations of surveys discussed in Section 1.5 will help to constrain unknown Galactic parameters in conjunction with accurate models, as discussed in Section 1.6.

The broad-band photometry produced by IPHAS provides a deep optical view of the entire Northern Galactic Plane; the first time such a dataset has been made available. In order to maximise its utility it was brought onto a uniform photometric scale (Barentsen et al., 2014), at which point the star counts it provides can be used to test models.

This thesis will cover part of the effort to clean and prepare IPHAS photometry for release, and the challenges involved with deriving accurate stellar density maps from the resulting catalogues. Such a wide-area stellar density map represents a unique data product, which could potentially be used in a range of applications. Its utility in testing Galactic model predictions will be demonstrated in comparisons with Besançon model star counts. The layout of this work is as follows:

- Chapter 2 introduces IPHAS in depth, covering its observing strategy and detailing the photometric catalogues that are used in this work. IPHAS source counting is discussed, and the generation of a preliminary density map is described.
- Chapter 3 describes the effort that went into cleaning the IPHAS data during the production of IPHAS DR2 (Barentsen et al., 2014). The calibration of IPHAS against APASS, an as-yet incomplete all-sky photometric survey which has a degree of coverage in the Galactic Plane, is detailed, along with the validation of this calibration using SDSS photometry. The reliability of IPHAS is also considered.
- Chapter 4 details the calculation of incompleteness due to confusion and sensitivity limits across the IPHAS DR2 footprint. Methods of accounting for this appearing in the literature are discussed, before an approach of artificial source generation and detection is employed.
- Chapter 5 lays out the method by which the final completeness-corrected density maps have been produced and discusses its reliability, before using a search for new over-densities (potential clusters) as a demonstration of its utility.
- Chapter 6 introduces the application of the Besançon model, before describing the mock Gaia catalogues that use its predictions. Transformations between the Gaia and IPHAS photometric systems are derived to facilitate comparisons between the two catalogues.

The predictions of the Besançon model in terms of stellar densities are compared to IPHAS observations via the corrected density map.

- Chapter 7 provides concluding remarks and proposes potential directions for future work with IPHAS photometry.

Chapter 2

IPHAS and the first density maps as health diagnostics

In this chapter I will provide a more in-depth introduction to IPHAS and its data products, which will be used extensively in subsequent chapters. Section 2.1 will discuss the observing strategy and Section 2.2 will describe the catalogues generated from the reduced images.

Envisioned as a way to visualise the entire catalogue down to some faint limit prior to uniform calibration (introduced in Chapter 3), preliminary density maps were generated; in order to accomplish this the catalogues needed to be trimmed and combined to remove spurious sources and repeated observations. Section 2.3 discusses the adopted approach to source counting while Section 2.4 describes the steps involved in producing the first generation of density maps and discusses their limitations - thereby motivating the further development described in Chapters 4 and 5.

2.1 Introduction

IPHAS (The INT/WFC Photometric Hα Survey of the Northern Galactic Plane) is an optical survey of the Galactic midplane, acquiring data in broad-band r' and i' Sloan filters and in narrow-band Hα. IPHAS makes use of the Wide Field Camera (WFC), a device consisting of four 2k×4k CCDs arranged in an L-shape configuration. The camera covers a sky area of 0.29 sq. deg., delivering a scale of $0.33''$/pixel.

The completed survey will cover the region $29 < \ell < 215$, $-5 < b < +5$, consisting of a total of 15,270 individual pointings (see Section 2.1.1). Observing began in August 2003, with 99.14% of pointings observed by the commencement of this work (although not all fields have yet passed quality criteria, as laid out in Section 2.1.3).

2.1.1 Observing strategy

The IPHAS observing pattern ensures that the majority of the survey footprint is observed more than once. This is accomplished by observing in field pairs, with each field pointing having an associated offset pointing, positioned $+5'$ in both declination and right ascension.

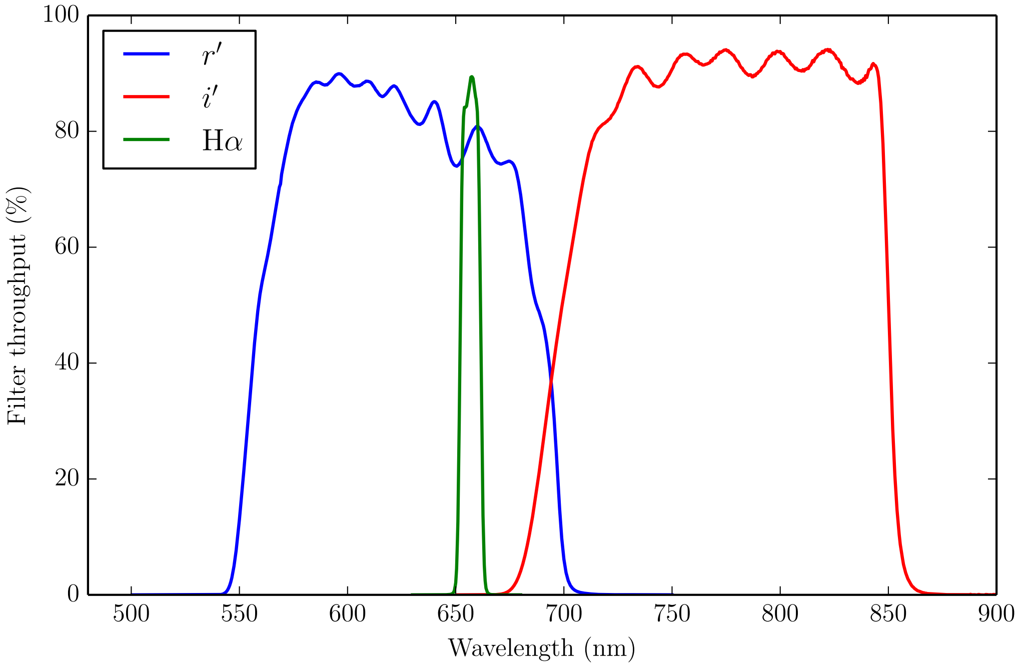


Figure 2.1: Filter profiles of IPHAS passbands. No modifications have been made for CCD efficiency or atmospheric transmission.

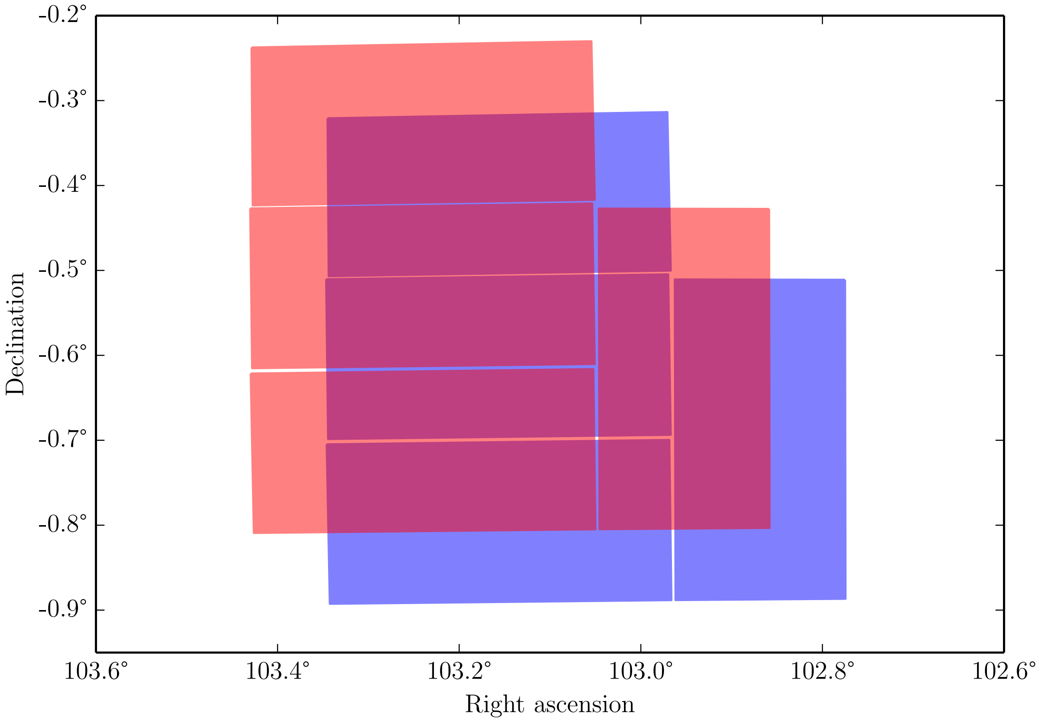


Figure 2.2: Pattern for IPHAS field pair layout, where offset (**red**) field is shifted from primary (**blue**) field coordinates by $+5'$ in both RA and DEC.

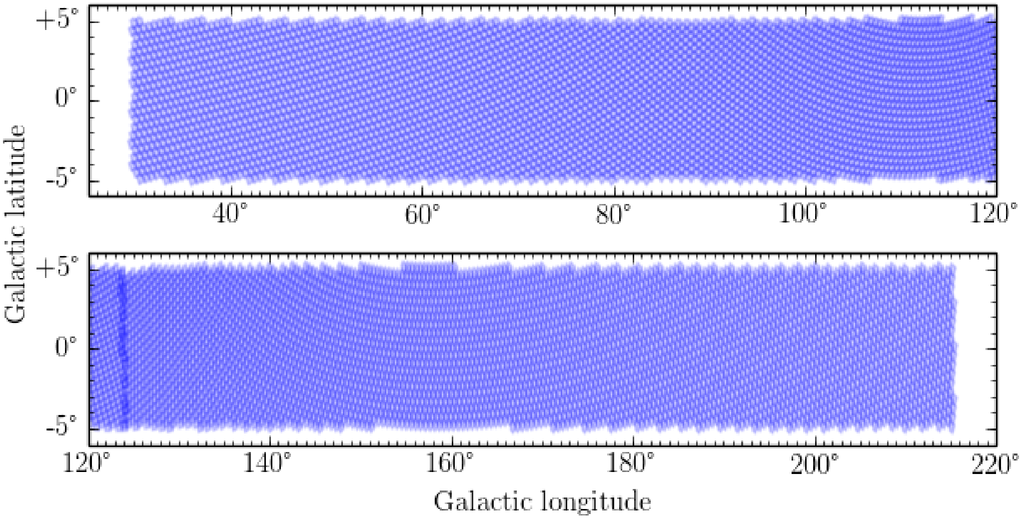


Figure 2.3: Pattern of IPHAS fields across the Northern Galactic Plane. Only primary fields are shown (i.e half of the total number of pointings that make up the survey). Darker regions are where neighbouring fields overlap; the area covered by neighbour overlaps is actually larger than showed here due to the offset fields. The darkest strip at $\ell \approx 124°$ is a result of the tiling pattern chosen for the fields in equatorial coordinates.

These field/pair offsets ensure that the innermost sections of the inter-CCD gaps present in a field are covered by its offset. Two small holes remain inside the footprint, where the gaps between CCDs 1 and 4, and 4 and 3 in the primary field intersect the gap between these three CCDs and the orthogonal CCD2 in the offset field, amounting to an area of 0.21 sq. arcmin. (or ∼0.01% of the area covered by the field pair). The inter-CCD gaps left towards the outer boundaries of field pairs are covered by neighbouring field pairs, which are placed such that they overlap.

Each field is observed in three filters sequentially: Hα with an integration time of 120s, r' (30s), i' (20s, increased from 10s in 2003). The field offset is then observed, leading to a field pair being observed in just over 10 minutes.

2.1.2 The IPHAS colour-colour plane

The combination of filters used in IPHAS provide two colours which together serve as a powerful tool with which to study the Galactic Plane. Figure 2.4 illustrates the $(r' - i')$, $(r' - \mathrm{H}\alpha)$ colour-colour plane. It can be seen that the tracks traced out by the main sequence at differing extinctions lie at a significant angle to the reddening vector, as shown for the example of early A-type stars.

The angle at which the main sequence reddens allows the effect of reddening to be disentangled from stellar type. Two example regions of IPHAS photometry are plotted in the lower panel: the first at $\approx (167.5°, 1.0°)$, displays a low $E(B - V)$ of ≈ 0.5 mag, while the other at $\approx (116.0°, 3.0°)$ shows an intermediate amount of extinction (≈2 mag).

Sources falling above the main sequence locus in a given field are those exhibiting Hα emission; Witham et al. (2008) has produced a catalogue of such emission line stars from

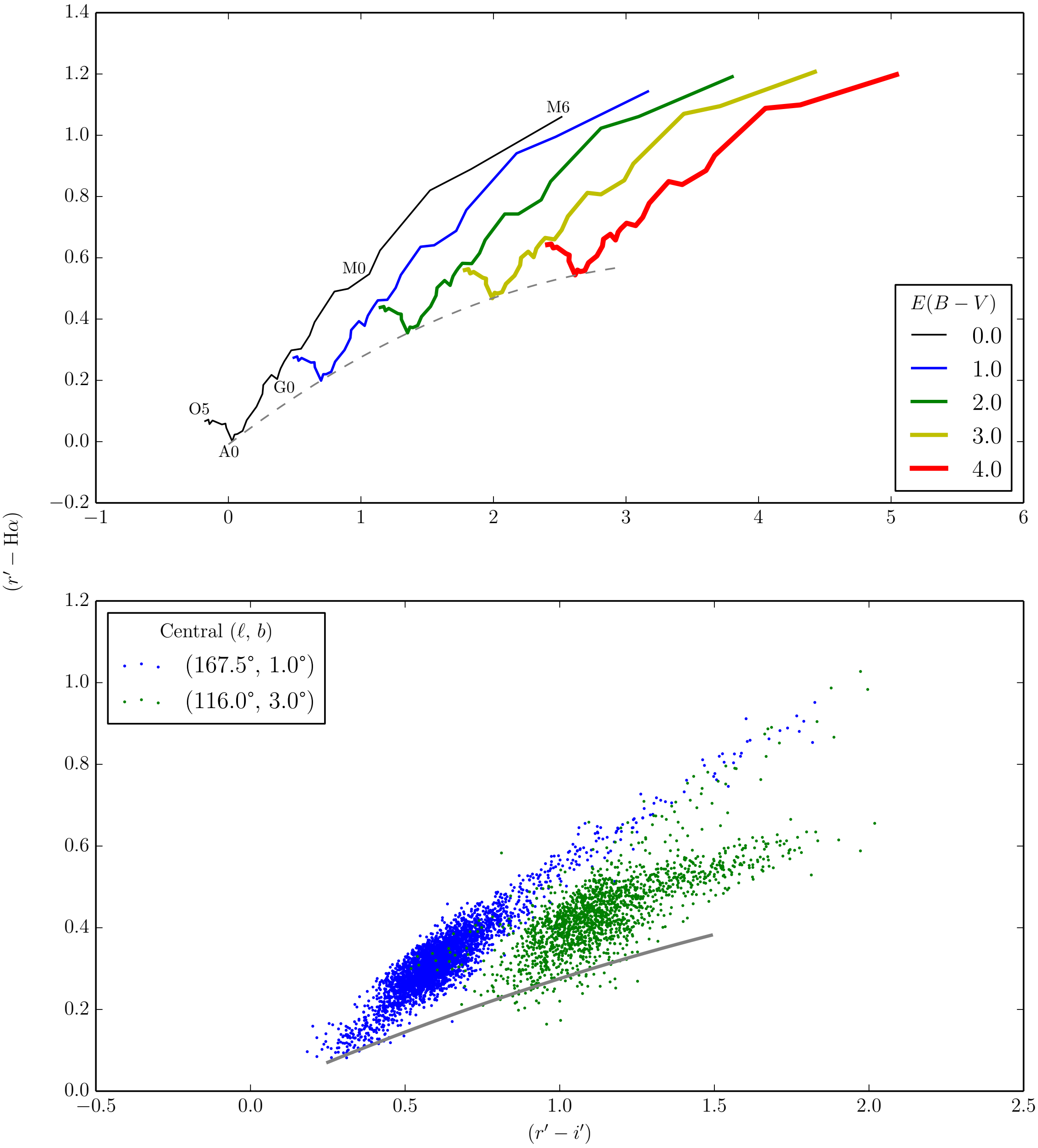


Figure 2.4: The IPHAS colour-colour plane. **Upper:** Tracks corresponding to the main sequence, suffering from increasing extinction up to $E(B-V) = 4.0$ in steps of 1 mag. The reddening line of early A-type stars is also shown (**dashed gray line**). **Lower:** The colour-colour planes of two regions taken from IPHAS DR2, showing stellar sources down to $r' < 19$. The region centred on (167.5°, 1.0°) has a typical $E(B-V)$ of ≈0.5 mag, while ≈2 mag is typical for the region centred on (116.0°, 3.0°).

IPHAS photometry.

2.1.3 Data quality

From the beginning of the survey, preliminary criteria were imposed on the data coming from the telescope; an ellipticity limit of 0.2 and a seeing threshold of $2.5''$ were required for fields to be processed (see Section 2.2.1).

Observing constraints were in place during data acquisition, in order to permit the use of bright time; fields were chosen to ensure no pointings were made within 30° of the moon to reduce sky background levels and scattered light effects. An airmass limit of 1.4 was generally applied but very occasionally ignored to finish field pairs contemporaneously.

An issue present in a number of fields is gain variation, whereby sections of CCD chips suffer from instabilities leading to substantial gain variations, affecting measured counts. These variations appeared intermittently from the start of IPHAS observations worsened in later years. These were often found to coincide with loose ground wiring in the WFC. The affected fields were eliminated from the survey by eyeballing colour-colour diagrams for misshapen stellar loci indicative of gain variation (see Section 3.1.1). The response of the WFC becomes non-linear at $\approx 55,000$ counts, and saturates at $\approx 64,000$, which corresponds to $r' \approx 12$ in a 30s exposure.

Stars more than a few magnitudes brighter than this limit create reflections from the telescope, as seen in Figure 2.5. The reflections caused by extremely bright ($\lesssim$ 5th magnitude) stars become significant enough to cause issues with catalogue generation. These issues can be seen by plotting the positions of detected stellar sources (shown in blue in the figure). The bright sources create zones in which they cause fainter sources to be missed. In addition to this, higher concentration of spurious sources can be seen in circular patterns tracing out the bright scattered light. Once faint limits are placed on detected sources, the fraction of spurious sources is lessened, though not entirely eliminated; sources with $i' < 18$ in field 4732_oct2005b (left) and $r' < 19$ in field 7080o_oct2012 (right) are shown in green. The presence of these false detections affects the overall reliability of the survey, as discussed in Section 3.4. The localised nature of this issue means that affected regions can be masked. How this is done is discussed in Section 5.1.6.

The brightest ($\lesssim$ 3rd magnitude) sources saturate entire columns of the detector and cause disruptions to the local background such that nearby sources are not picked up during catalogue generation (see Section 2.2). This issue tends to affect larger regions of the CCD than the spurious sources discussed above - corrections to account for this incompleteness at the CCD level are discussed and calculated in Chapter 4. This is not a widespread issue; the bright star catalogue of Hoffleit & Jaschek (1991) lists only 12 stars with $V < 3$ falling inside the footprint of IPHAS.

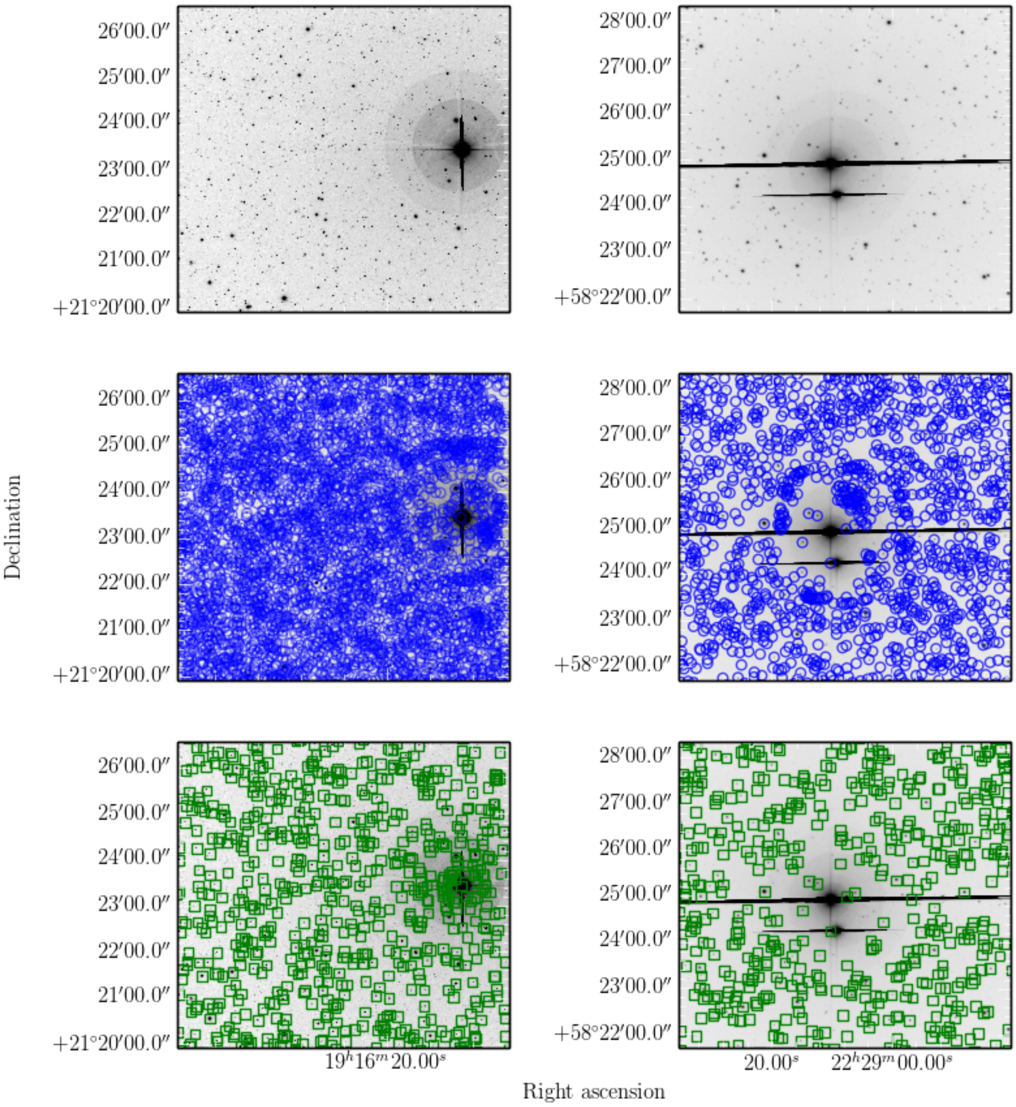


Figure 2.5: Cutouts of two regions of IPHAS, covered by fields 4732_oct2005b (**left**) and 7080o_oct2012 (**right**). The cutout regions show the stars 1 Vul ($V = 4.77$) and δ Cep (a resolved binary system, $V = 3.75$) respectively. The bright stars cause reflections which affects nearby photometry. **Upper:** i'-band image cutout of 4732_oct2005b and r'-band cutout of 7080o_oct2012. **Mid:** All i'/r'-band sources detected during generation of photometry (see Section 2.2) that were assigned morphological classifications -1 (stellar), -2 (probably stellar) or +1 (extended) (see Sections 2.2.2 & 2.3.1). **Lower:** $18 < i'/19 < r'$ sources.

Flag	Definition
-9	Saturated
-7	Contains bad pixels
-3	Compact but probably stellar
-2	Probably stellar
-1	Stellar
0	Noise
+1	Non-stellar
+5	Repeatedly non-stellar

Table 2.1: Definition of morphological flags in IPHAS catalogues. Note that +5 classification was introduced during density mapping (see Section 2.4.2).

2.2 Photometry

2.2.1 Catalogue generation

The reduction pipeline for the survey is located at the Cambridge Astronomical Survey Unit (CASU). It is a descendant of the system set up for the INT Wide Field Survey (McMahon et al., 2001). The catalogues derived from the reduced frames are produced using the CASUTools IMCORE program. The routine detects sources by first estimating a smoothly varying background over the field, then identifying contiguous groups of pixels with values 1.25σ higher than the local background level for a filtered version of the image. The sources at these coordinates are measured and parameterised using the original unfiltered data.

Photometry is generated by measuring the total flux enclosed by apertures of increasing diameter, with the default aperture radius set to 3.5 pixels (corresponding to a diameter of $2.3''$). Sources that overlap have their photometry refined at this point using isophote analysis, with the potential for additional sources to be introduced to account for some fraction of flux of the blend (Irwin, 1985).

The World Coordinate System (WCS) of the images are determined by crossmatching to 2MASS sources, after an initialisation using telescope pointing and camera setup. The final astrometric solutions achieve precisions better than $0.1''$ (Barentsen et al., 2014).

2.2.2 Morphological classifications

The catalogue generation by IMCORE supplies a morphological classification statistic for each detected object. The statistic is based primarily on the variation (curve of growth) in flux enclosed by apertures of increasing diameter, compared to the curve of growth for average stellar sources in the CCD. González-Solares et al. (2008) describe the method by which the 'stellar-ness' of sources is defined for a field, described by a distribution of the statistic centred on zero (perfectly stellar) with a deviation σ about this main stellar locus. The greater the deviation, the less stellar the source, with low valued sources $< 3\sigma$ appearing noise-like, and $> 3\sigma$ as extended non-stellar sources. Table 2.1 describes the possible morphology codes that can be assigned to objects.

For density mapping purposes, objects with classifications -9, -8, -7, -3 and 0 are ignored.

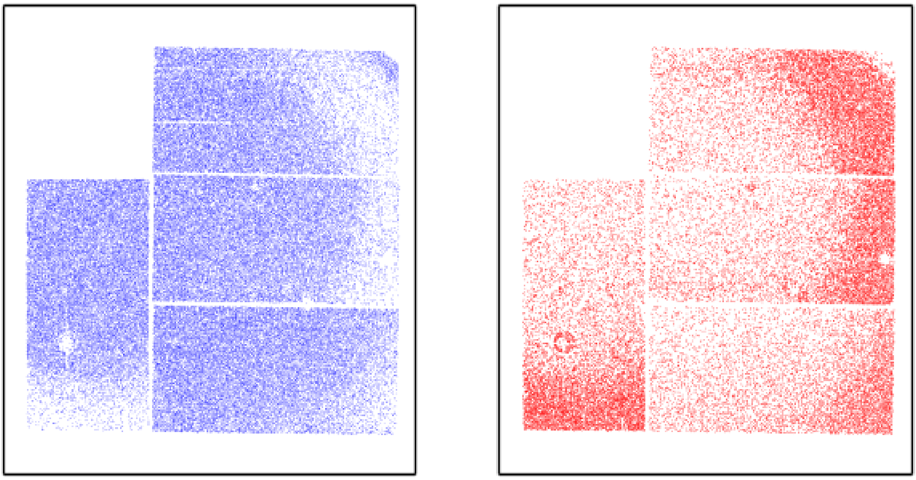


Figure 2.6: Distribution of objects detected in IPHAS field 4975o_aug2004a. **Left**: Objects with classification -1 or -2. **Right**: Objects with classification +1. The clear shift from objects classified as stellar to non-stellar at the field edges is due to the orientation of the WFC CCDs. The chips are known to tilt with respect to the focal plane of the instrument, an effect which becomes more pronounced with increasing distance from the optical axis. The point spread functions of objects falling in affected areas are degraded, resulting in a more extended shape and a +1 morphology code. This field was selected as a particularly badly affected example of this issue.

Initially non-stellar (+1) objects were also discarded as only non-extended point sources were likely to be Galactic stars. However it was found during the first attempts to generate a density map that a fraction of fields showed a lower number of stars than expected compared to neighbouring fields. Upon investigating these underdensities, it became apparent that the "missing" stars were in fact being misclassified as extended non-stellar sources by IMCORE (i.e. their classification statistic is significantly larger than is consistent with the main stellar locus).

Figure 2.6 shows the spatial distribution of this misclassification in two cases; corners of the detector furthest from the optical axis seem to be the most affected - in these corners the PSF varies such that they appear non-stellar to the classification scheme of IMCORE. This variation follows the non-flatness of the chips in the camera, and its occurrence in only a few runs points to the seating of the camera or the focus as the likely cause.

2.2.3 Fluxes & apertures

As mentioned in Section 2.2.2, IMCORE determines the curve-of-growth of flux for a source by measuring fluxes enclosed within apertures of increasing diameter - these fluxes form part of the output catalogues. The primary estimate is determined by an arbitrary "core" radius defined to be 3.5 pixels for WFC images. Fluxes are returned for apertures with 0.5×, 1×, $\sqrt{2}\times$, 2× and $2\sqrt{2}\times$ this aperture radius. In Chapter 2, the 2× aperture was initially used, corresponding to an aperture diameter of $3.3''$. This was later revised to the 1× aperture ($2.3''$ diameter), which is used from Chapter 3 onwards; in most cases the difference in flux between

Column	Definition
$\{r'\|i'\|ha\}$	Magnitude of relevant band derived from flux enclosed in a 2.3″ diameter aperture (see Section 2.2.3).
$\{r'\|i'\|ha\}$Err	Photometric error on magnitude measurement.
$\{r'\|i'\|ha\}_{3.3}$	Magnitude of relevant band derived from flux enclosed in a 3.3″ diameter aperture (see Section 2.2.3).
$\{r'\|i'\|ha\}$Class	Morphology code for source PSF in relevant band image (see Section 2.2.2).
RA/DEC	Coordinates of source, determined from first detection from $\{r'\|i'\|ha\}$

Table 2.2: Key columns output by IMCORE which will be used in this work, and the column names adopted.

the two apertures is minimal, however in regions with high nebulosity the 1× aperture is less affected by rapid variations in background counts, while still enclosing almost all of the flux of the 2× aperture.

The IMCORE output provides correction terms for flux missed by each aperture size, estimated from the average stellar curve-of-growth of each image. These corrections are necessary when computing magnitudes from measured fluxes, as discussed in Section 4.4.8.

2.2.4 Column definitions

The columns generated using IMCORE are assigned names according to CASU conventions. In preparing for the second IPHAS data release, column naming conventions shifted closer to UKIDSS. These are the column names which will be referenced throughout this work, as laid out in Table 2.2. Additional columns of use are detailed in Section 4.4.

The single-band catalogues generated by IMCORE are bandmerged before use, providing one set of equatorial coordinates per source. The coordinates are taken preferentially from the r'-band detection, or from the i'-band (or Hα) if the source is missing in r' (or r' and i'). Galactic coordinates were computed from the equatorial coordinates provided in the catalogues.

2.3 Source counting

Once the entire survey region is covered, the availability of catalogues allows properties of stellar populations to be studied over the entire Northern Galactic Plane. The existence of catalogues derived from the images allows objects to be easily counted and binned up. An obvious product of use would be a measure of object density across the whole survey region. This would allow the construction of density maps based on IPHAS photometry, facilitating direct comparisons to other studies of Galactic populations and surveys in other wavebands.

Fields observed in a variety of conditions across the survey have a range of depths; counting all objects detected regardless of magnitude would lead to field-to-field inconsistencies (see Section 2.4.4). Many fields reach fainter than $r' = 21$, though the scatter in the photometry of objects so faint is significant. Figure 2.7 shows the variation in photometric error

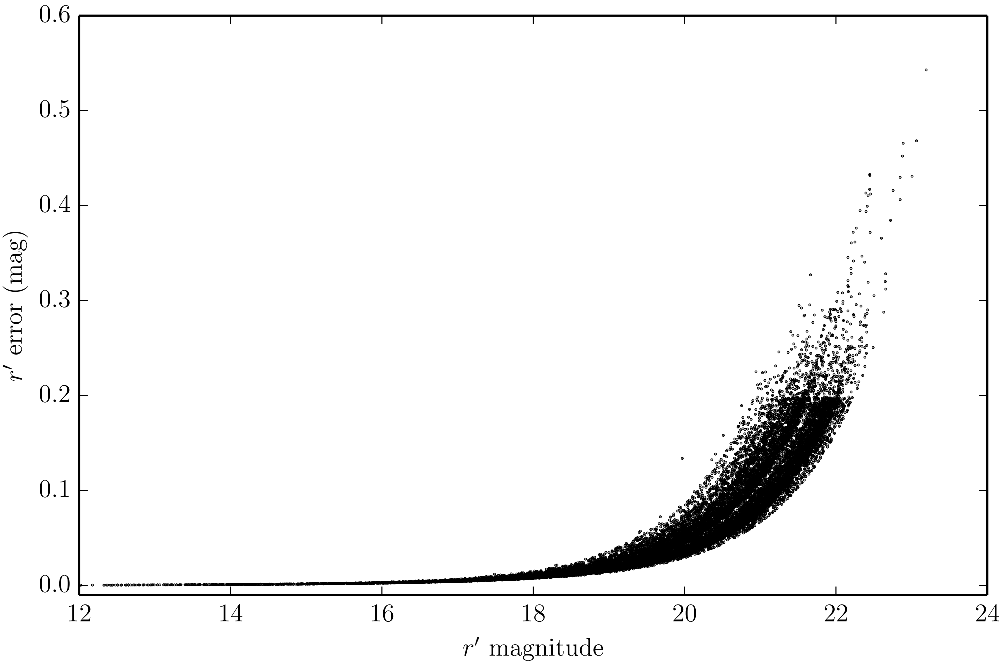


Figure 2.7: Variation of photometric error in r'-band with magnitude for a 5 sq. deg. region at $\ell \approx$110°.

with magnitude - at $r' \approx$19 the typical random error is 0.03 mag. At fainter magnitudes the errors increase substantially.

In producing preliminary density maps down to $r \gtrsim 20$, maps began to exceed the limiting magnitude of fields, leading to extreme field-to-field variations in measured source density. At limits brighter ($r \approx 19.5$) than the expected limiting magnitude, a degree of this variation was still visible - highlighting the need for a uniformly calibrated photometric scale for the entire survey.

Section 2.4 details the generation of these initial density maps with a preliminary global calibration, while Chapter 3 covers efforts made to obtain an improved global calibration by using the photometry of an external survey as a reference.

2.3.1 On the expected number density of extra-galactic sources

In the case where all point sources are accurately classified, the primary source of extended objects would be galaxies - only a smaller fraction would be contributed by Galactic sources with extended emission e.g. planetary nebulae and WR stars. Such Galactic contributions would be sufficiently small - such rare objects' number densities amount to an average of <1 object per IPHAS field (Sabin et al., 2014; Stock & Barlow, 2010) - that they would have negligible impact on density mapping efforts if included in stellar number counts. The inclusion of galaxies needs further thought, as the apparent density of galaxies varies considerably across the sky.

Due to the misclassification effect demonstrated in Figure 2.6, a certain fraction of objects

	Optical Axis			CCD3		
Field	-1	+1	% non-stellar	-1	+1	% non-stellar
4975o_aug2004a	437	139	24.13%	62	439	87.62%
0105_aug2004a	110	53	32.52%	6	152	96.20%
7474_dec2006	75	36	32.43%	15	80	84.21%
0310_dec2005	330	72	17.91%	227	67	22.79%

Table 2.3: Numbers of objects appearing as stellar ($rClass = -1$) and non-stellar ($rClass = +1$) in 5 sq. arcmin. regions at the optical axis of the field and towards the vignetted region in CCD3. The first three fields suffer from misclassification issue to varying degrees, while the fourth does not. The shift towards classifying objects as non-stellar at the outskirts of a pointing is illustrated in Figure. 2.6.

are rendered indistinguishable from galaxies on the basis of their morphology code alone. To aid in the decision of whether or not to exclude objects assigned the morphology class +1, the difference in this fraction between affected and non-affected (i.e. the optical axis) regions of the image plane was determined.

Table 2.3 shows the extent of the misclassification illustrated in Figure 2.6, along with other examples. The numbers quoted are for regions of area 5 sq. arcmin.. In the most extreme case investigated, shown in Figure 2.6, assuming that the $\frac{stellar}{non-stellar}$ ratio remains roughly constant across the field, an estimated ∼300 objects are construed as misclassified in the CCD3 region. If this density of objects is much greater than the expected density of genuine non-stellar sources, then the losses from eliminating $rClass = +1$ objects from counts would be more damaging than the uncertainties introduced by retaining them in a sample of point sources. Galaxies are the primary source of extended objects. Hence a comparison between the estimated misclassification rate and expected number densities of galaxies will aid in the decision as to whether or not to permit $rClass = +1$ objects to remain in our sample.

Yasuda et al. (2001) investigated galaxy number counts using early data from SDSS, making observations towards the Galactic caps. By visually inspecting galaxies towards the northern Galactic cap up to $r' = 16$, and then switching to automated object classification at fainter magnitudes, they counted 764,335 galaxies in the r'-band, and 716,099 in the i'-band. The values for each magnitude bin are listed in Table 2.4 and are plotted in Figure 2.8, after they have been scaled to the area of a WFC CCD. Over such an area, on average, ≈350 galaxies would be expected in both the r' and i'-band down to the limiting magnitude of the Yasuda sample (21.5). This would be the case if the extinction in the direction of IPHAS fields were the same as that of the observations by Yasuda et al. (2001). However their observations, towards the northern Galactic cap, experienced $E(B-V) < 0.1$ - significantly lower than the typical integrated extinction along sightlines in the Galactic plane.

The median $A_{r'}$ for all IPHAS fields from the Schlegel et al. (1998) map corrected according to Schlafly & Finkbeiner (2011) is 2.64 mag, and $A_{i'} = 1.97$ mag. Here the colour excess $E(B-V)$, derived from the Schlegel et al. (1998) map using the central coordinates of each IPHAS field, is taken to represent the extinction for entire fields. Taking the 5σ depths of IPHAS of $r' = 21.2$ and $i' = 20.0$ (Barentsen et al., 2014), the median limiting magnitudes

	r'		i'	
Mag range (mag)	Raw count	Norm. count	Raw count	Norm. count
11.0 - 11.5	4	0.018	3	0.013
11.5 - 12.0	3	0.013	3	0.013
12.0 - 12.5	6	0.026	16	0.070
12.5 - 13.0	23	0.101	28	0.123
13.0 - 13.5	28	0.123	61	0.267
13.5 - 14.0	73	0.320	97	0.425
14.0 - 14.5	118	0.517	244	1.07
14.5 - 15.0	315	1.38	523	2.29
15.0 - 15.5	577	2.53	924	4.05
15.5 - 16.0	1513	5.35	2597	9.20
16.0 - 16.5	3050	10.8	4806	17.0
16.5 - 17.0	5607	19.8	9337	33.0
17.0 - 17.5	10532	37.3	16474	58.3
17.5 - 18.0	18239	64.5	29553	104.6
18.0 - 18.5	16890	113.2	27232	182.5
18.5 - 19.0	29443	197.4	47046	315.4
19.0 - 19.5	50206	336.5	77322	518.3
19.5 - 20.0	80324	538.4	121858	816.0
20.0 - 20.5	123213	825.9	166314	1114
20.5 - 21.0	181428	1216	145706	976.7
21.0 - 21.5	242743	1627	65955	442.1

Table 2.4: Galaxy counts as reported by Yasuda et al. (2001). Normalised counts are given per 0.5 mag deg^{-2}.

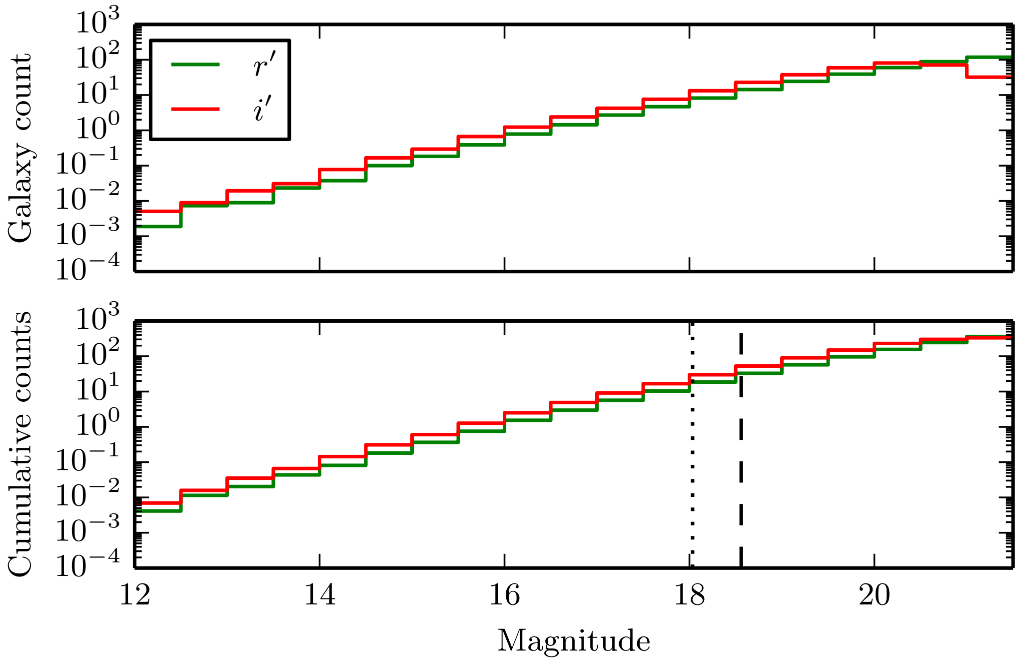


Figure 2.8: Galaxy counts vs. magnitude from Yasuda et al. (2001). Normalised counts (originally 0.5 mag deg^{-2}) have been scaled to effective WFC CCD area. **Vertical dashed line:** the 5σ limit of IPHAS r'-band detections (21.2 mag) less the median $A_{r'}$ of IPHAS fields taken from the Schlegel et al. (1998) extinction map (2.64 mag, after applying correction factor of Schlafly & Finkbeiner (2011)). **Dotted line:** the same but for i', with a 5σ limit of 20.0 mag and a median $A_{i'} = 1.97$ mag.

after applying the median $A_{r'}/A_{i'}$ are $r' = 18.56$ and $i' = 18.03$. The cumulative count of Yasuda et al. (2001) galaxy counts down to these limits are <20 in both r'- and i'- bands. For a region of 5 sq. arcmin. as considered in Table 2.3, the estimated galaxy count is $\lesssim$1 - insignificant compared to potentially misclassified point sources (see Table 2.3).

Based on these estimated extragalactic source counts it was obvious that in situations where accurate overall IPHAS source counts are needed, a far larger impact would be felt by excluding +1 classed sources (potentially losing hundreds of misclassified sources in the process) compared to admitting them, thereby counting a small number of galaxies.

2.4 Preliminary density mapping

Prior to the preparation and release of IPHAS DR2, the only catalogues in existence were bandmerged catalogues generated by CASU for each pointing (one field in three filters). A preliminary set of observations had been selected for release and had been used for initial tests of a method of global calibration using the approach of Glazebrook et al. (1994). These fields were gathered and were combined in a series of steps to produce a single catalogue which would only contain each object once. These steps are laid out in the following sections.

Class 1	Class 2	Combined Class
-1	-1	-1
-1	-2	-1
-2	-2	-2
+1	+1	+5
+1	-1	-1
+1	-2	-2

Table 2.5: Original ‘truth table’ for determining morphology flags for crossmatched objects in field pairs.

2.4.1 Trimming and calibrating the photometry

Before combining any catalogues, the data needed to be cleaned and brought onto a common photometric scale. An issue with the catalogues generated directly from the images was the lack of global photometric calibration. This is an issue due to the faint cutoff magnitude mentioned in Section 2.3 - a shift of ±0.1 *mag* in the photometry often brings about significant changes in the number of objects included or excluded in a magnitude limited map.

An interim global calibration was available using an early subset of IPHAS observations, different to those eventually released in DR2 (unpublished, prepared by Brent Mizalski). The photometric shifts estimated by this preliminary calibration were applied to each field before map generation.

A first step was to cut out detector boundaries in order to prevent erratic counting in these areas, affecting the final product. The pixel coordinates of each object had not been carried through into the bandmerged catalogues, which made this trimming difficult. Rather than regenerate files for the entire survey, the pixel coordinates were recomputed from the ICRS coordinates for each object and the WCS information contained in the catalogue headers. This conversion was of sufficient precision to cut out a 10 pixel border region around each CCD to exclude potentially truncated objects from the catalogues.

One of the corners of CCD3 suffers from significant vignetting, causing incorrect fluxes to be reported. The majority of the vignetted region is entirely obscured, with only a small area potentially contributing sources to the catalogues. To ensure no affected objects make it to the combined catalogue, a polynomial was fitted to the x,y pixel coordinates of sources closest to the vignetted corner, and sources within 10 pixels of this inferred edge excluded from the catalogues.

2.4.2 Combining field pairs

The next step in producing a mappable catalogue was to combine field pairs (contemporaneous on and off-field observations). This was achieved by using the STILTS package and its TMATCH2 routine to crossmatch the field pair. A matching radius of 0.5 arcsec was used as a sensible compromise between minimising spurious associations and accounting for poorer astrometry towards field edges. At this point all objects with a classification other than -1, -2 or +1 were dropped from the catalogues (see Section 2.3.1).

Matching object photometry was calculated from an error-weighted average of the pri-

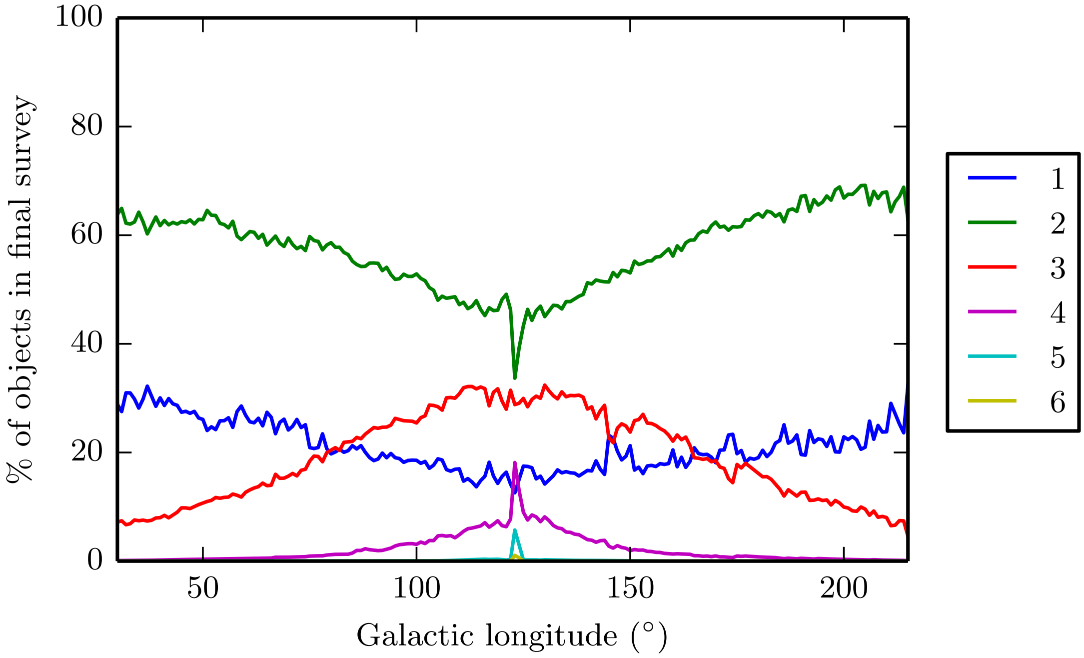


Figure 2.9: Percentage of objects in the final crossmatched catalogue with varying repeat frequencies. Note that the plot has not been corrected for missing fields; as such there are sudden jumps in relative fractions of frequencies - these jumps are due to objects appearing once as a result of field edges which would have been covered by neighbouring fields in a complete survey.

mary and offset field detection, and averaged coordinates calculated from the astrometry. Table 2.5 shows the adopted combined morphology class for all possible combinations of two observations of an object. A new class, +5, was adopted for objects which are repeatedly classified as non-stellar, as a way to differentiate genuine extended objects from those suffering the misclassification described in Section 2.2.2. Such objects are unlikely to have been misclassified twice, as the $5'$ offset of the fields in RA and DEC will ensure that at least one of the two detections will lie in an unaffected region of the image plane.

2.4.3 Combining further repeat observations

Figure 2.9 shows the variation in repeat observations computed across the survey via the cross-matching exercise. It can be seen that the highest frequency (objects observed five or six times by the survey) occurs only at $\ell \sim 124°$ - this is where the regular geometry of overlaps between fields breaks down where the two halves of the survey pattern meet (see Figure 2.3). At this point it is possible for up to four field pairs to overlap, while in the earliest and latest regions of the Galactic plane covered by IPHAS, a maximum of two field pairs can cover a given coordinate.

These numbers show the behaviour expected of IPHAS; around 60% of objects have been observed twice, while the remainder of the catalogue has mostly been observed either once or three times.

Once files were generated for all field pairs, the survey was split up into 1 deg wide strips

Class 1	2	3	Combined
-1	-1	-1	-1
-1	-1	-2	-1
-1	-1	+1	-1
-1	+1	+1	-1
-1	-2	+1	-1
-1	-1	+5	-1
-1	-2	+5	-1
-1	-2	-2	-2
-2	-2	-2	-2
-2	-2	+1	-2
-2	-2	+5	-2
+1	+1	+1	+5
-2	+1	+1	+5
-1	+5	+5	+5
-2	+5	+5	+5
+1	+1	+5	+5
+1	+5	+5	+5
+5	+5	+5	+5
-1	+1	+5	+5
-2	+1	+5	+5

Class 1	2	3	4	Combined
-1	-1	-1	-1	-1
-1	-1	-1	-2	-1
-1	-1	-2	-2	-1
-1	-1	-1	+1	-1
-1	-1	-2	+1	-1
-1	-2	+1	+1	-1
-1	-1	+1	+1	-1
-1	-1	+5	+5	-1
-1	-1	-2	+5	-1
-1	-1	+1	+5	-1
-2	-2	-2	-2	-2
-2	-2	-2	-1	-2
-2	-2	-2	+1	-2
-2	-2	-2	+5	-2
-2	-2	-1	+1	-2
-2	-2	+1	+1	-2
-2	-2	-1	+5	-2
All other combinations				+5

Table 2.6: Truth table for determining morphology flags for crossmatched objects **Left:** where objects appear three times. **Right:** where objects appear four times.

in Galactic longitude. Gathering all field pairs, any objects falling inside a strip were collected and crossmatched, allowing for up to four appearances per object. In theory this would allow eight individual contributions per object (four field pairs, each contributing two observations of the same object), however this is not observed as shown in Figure 2.9.

As in the case of merging the two detections of objects within a single field pair, an average position was adopted for each object, and a weighted average of the photometry was taken. The combining of morphology codes for objects appearing more than twice becomes quite complicated. Table 2.6 shows the scheme adopted for this step of the crossmatching process.

Following the consolidation of all objects in these longitude strips, the only possible remaining multiple appearances of objects were cases where each occurrence of an object lay on either side of one of these arbitrary strip boundaries. To complete the crossmatching process, the strip boundaries were shifted by 0.5 deg in Galactic longitude, and the objects crossmatched one final time. Objects were allowed to appear either once or twice, thereby eliminating remaining repeat observations.

2.4.4 Faint magnitude limits

Based on the increase in photometric uncertainty seen at fainter magnitudes (see Figure 2.7), a faint limit of r'=19.0 for the density map was initially adopted. Generating several iterations of the density map with different depths supported this decision - Figure 2.10 shows the effect of setting this limit to magnitudes of 19.0 and 20.4 in the r' band. The effect of

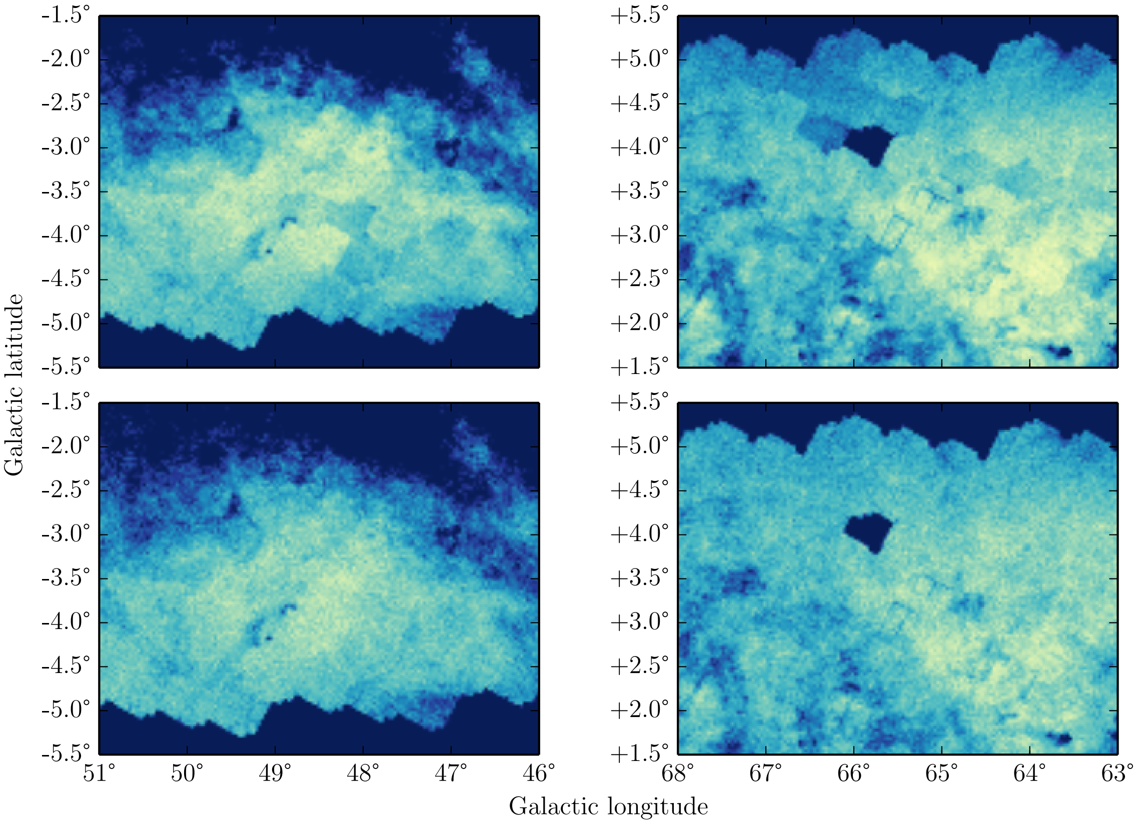


Figure 2.10: Two regions of an r'-band density map generated at a resolution of $2' \times 2'$. The two versions were generated by counting down to r'=20.4 (**upper**) and r'=19.0 (**lower**). Field-to-field depth variations with a WFC-shaped footprint are clearly apparent in the deeper maps.

setting the fainter limit results in field-to-field variations between observations carried out under different observing conditions appearing in the density map. These are not seen when the limit is set to $r' = 19$.

2.4.5 The map

Figures 2.11-2.14 show the r'- and i'-band density maps, limited to 19th and 18th magnitude respectively. Fine structure is immediately apparent, particularly at $\ell < 90°$. Some obvious clusters appear across the plane, most prominently towards the anticentre (e.g. NGC 2158 at (186.63°, +1.78°) and Trumpler 5 at (202.86°, +1.05°)).

However the maps are not without flaws; several sharply peaked overdensities appear at the coordinates of bright stars (e.g. Capella at (162.59°, +4.56°)). These are due to the spurious sources picked up around bright stars discussed in Section 2.1.3. The single-band nature of these maps (i.e. requiring that sources appear only in one band) make such sources hard to eliminate. These are dealt with in Chapter 5.1.6 by masking the regions around bright sources.

Several overdensities that bear the imprint of the WFC detector mosaic appear in these maps (e.g. at ≈(75.6°, +3.61°) in the i'-band map), appearing to double the source count in the affected areas. These are due to astrometric problems in certain fields - the fact that these initial density maps highlights such issues was used to identify them in preparing DR2 (see Chapter 3).

During the crossmatching process, potentially problematic cases were noticed, in which poor seeing observations overlapped with good seeing data, often from an entirely different observing run. In this case source association could become difficult, particularly in regions of high source density. The densest regions reach number densities of ≈ 100 sources per sq. arcmin., which as discussed in Chapter 4 (Section 4.2), is in the regime where confusion significantly impacts photometry.

The approach taken in this chapter of averaging photometry and astrometry means that in regions of overlapping data, the observations with the best seeing will dictate the source counts; though fewer sources in poor data suffering from confusion may be detected and matched to the good data, the less confused (and therefore more complete) photometry will drive the source counts. However the averaging of good data with potentially poor data could negatively affect the resulting catalogues, and regions where only lower quality data is available will still suffer from incompleteness. Chapter 4 discusses the impact of incompleteness on IPHAS, and considers methods that have been employed by previous studies to quantify incompleteness in their data. A full treatment of incompleteness is presented and applied to the survey, allowing robust density maps to be generated, corrected for its effects.

The preliminary maps are of limited utility as information on pointing coverage has not been carried through: cells with zero occupancy could be genuinely devoid of sources while being covered by IPHAS, or they may simply fall outside the footprint of the survey. Chapter 5 introduces a method for producing the corrected density maps which carries through coverage information for each cell of the density map; this provides the ability not only to distinguish these unoccupied cells, but to correct the observed occupancies of partially covered cells and their associated uncertainties. These improved maps are produced are produced using the data selected for inclusion in DR2; the preparation of this dataset is presented next.

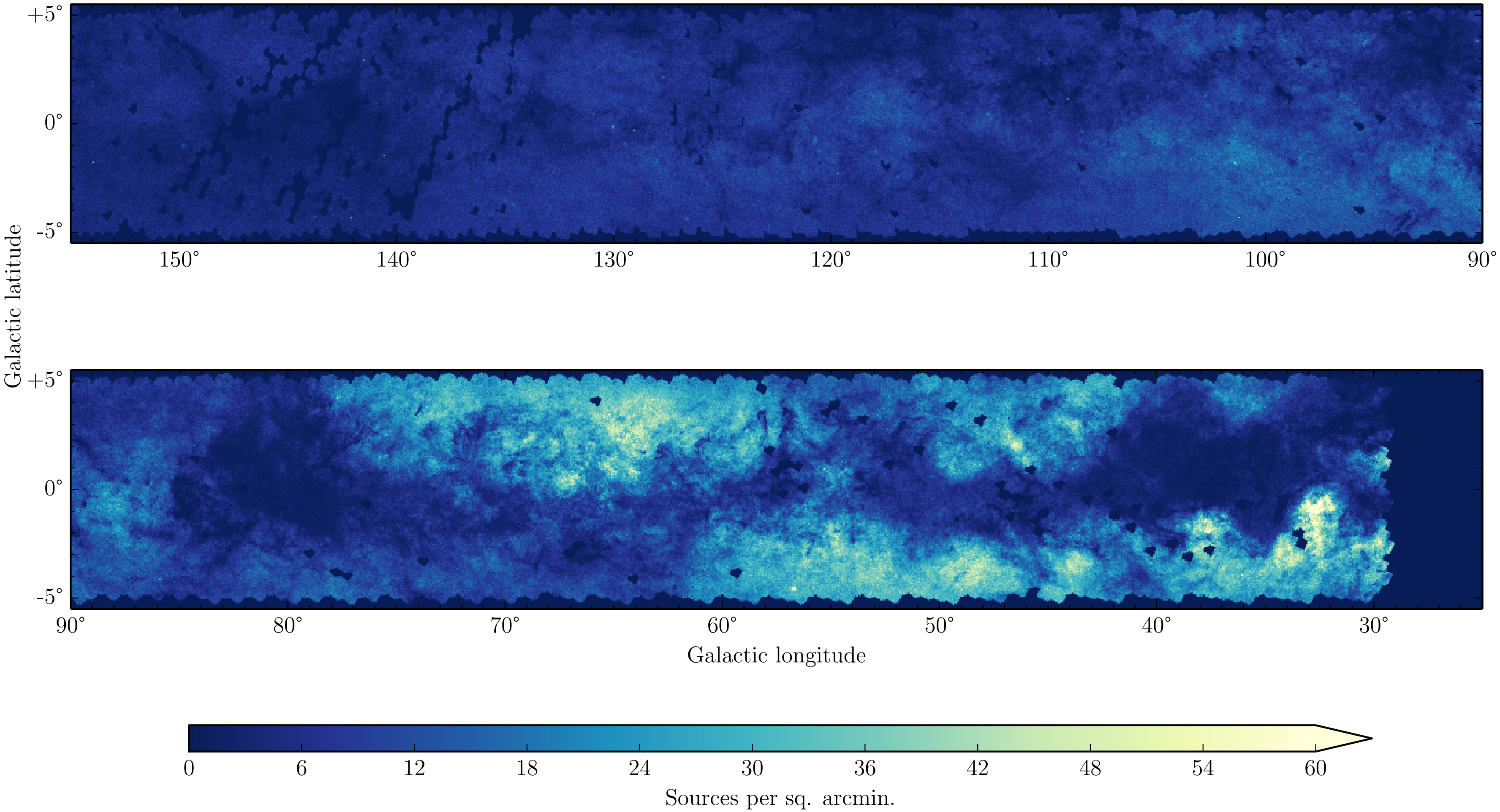


Figure 2.11: r'-band density map of sources down to 19th magnitude, at $2' \times 2'$ resolution, for $\ell < 155°$. Information on the fraction of each cell covered has not been carried through, leading to unreliable counts at survey edges, and no distinction between empty and uncovered cells.

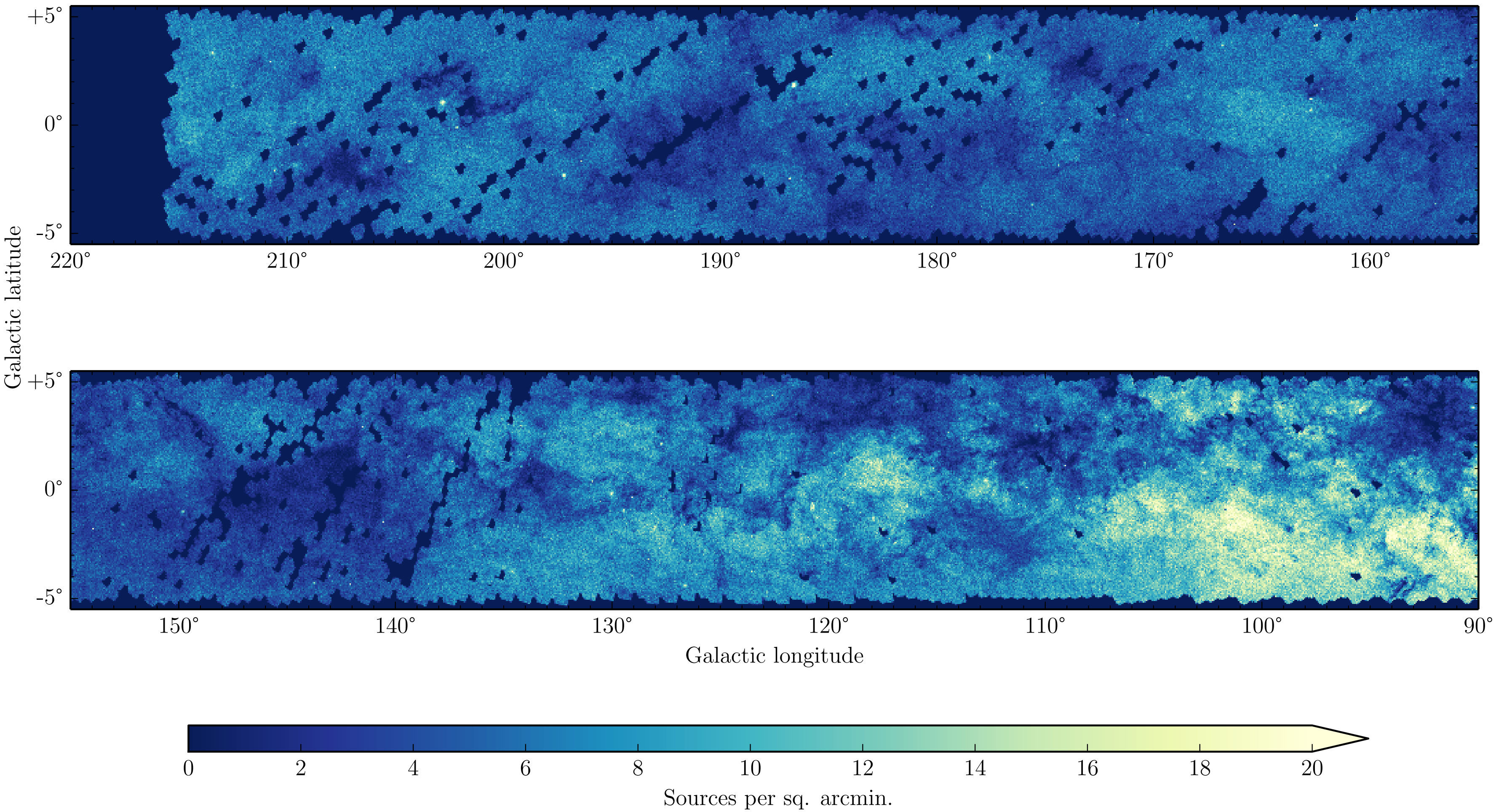


Figure 2.12: Same as Figure 2.11, only for $\ell > 90^\circ$, shown separately due to the larger range in density. The $90^\circ < \ell < 155^\circ$ panel has been reproduced in both Figures to provide a sense of scale of the variation in density across the plane.

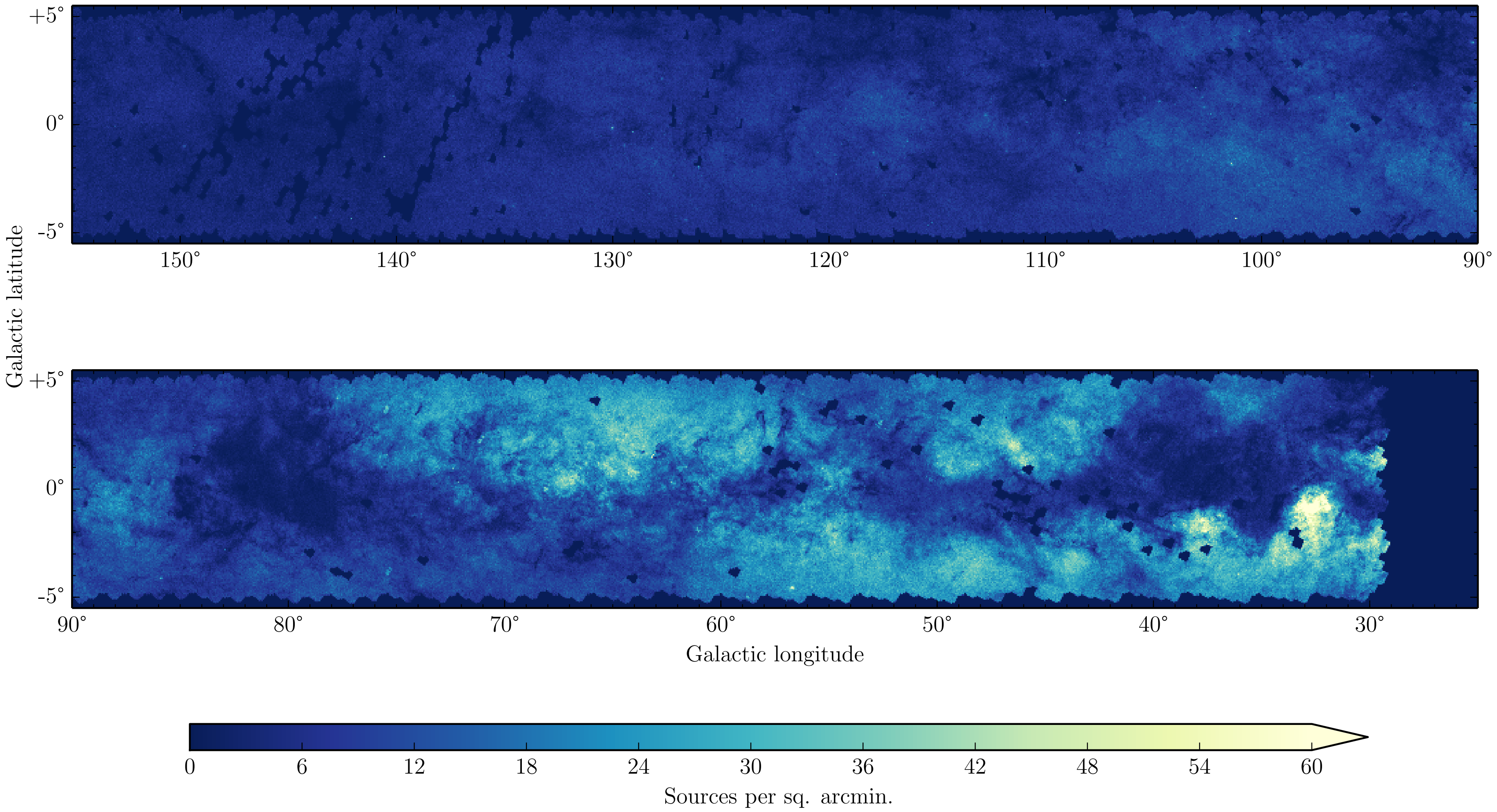


Figure 2.13: i'-band density map of sources down to 18th magnitude, at $2' \times 2'$ resolution, for $\ell < 155°$.

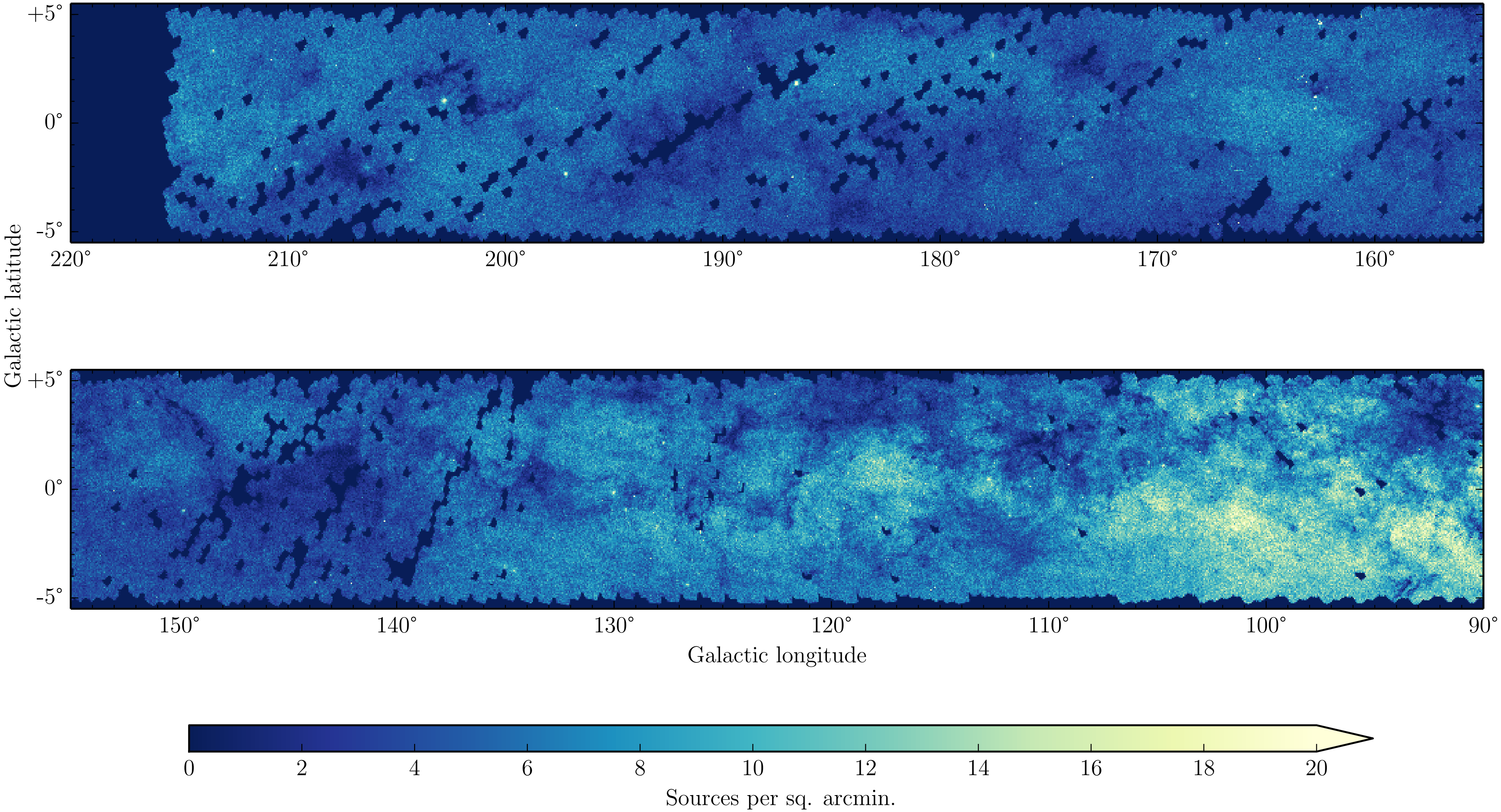


Figure 2.14: Same as Figure 2.13, only for $\ell > 90^\circ$.

Chapter 3

Calibrating IPHAS

Following the publication of the Initial Data Release (IDR) of IPHAS (González-Solares et al., 2008) and the commencement of this work, more of the survey footprint was observed and a programme of replacement of poor data began. This continued to be reduced at CASU, allowing a more uniform calibration to be applied to the photometry. During this time, a preliminary global calibration had been put together, that was already in place before producing the initial density map (see Section 2.4).

This calibration was known to be imperfect, and needed to be improved before any further data releases could go ahead. In this chapter the steps taken to improve the data are laid out. The totality of this work is summarised in Barentsen et al. (2014), which details all steps to the second data release (DR2) of IPHAS catalogues. This chapter highlights the particular tasks that I took on in preparing the photometry for release. As appropriate, I distinguish my work from that by the lead author, Geert Barentsen (GB hereafter), and IPHAS PI Janet Drew (JD hereafter).

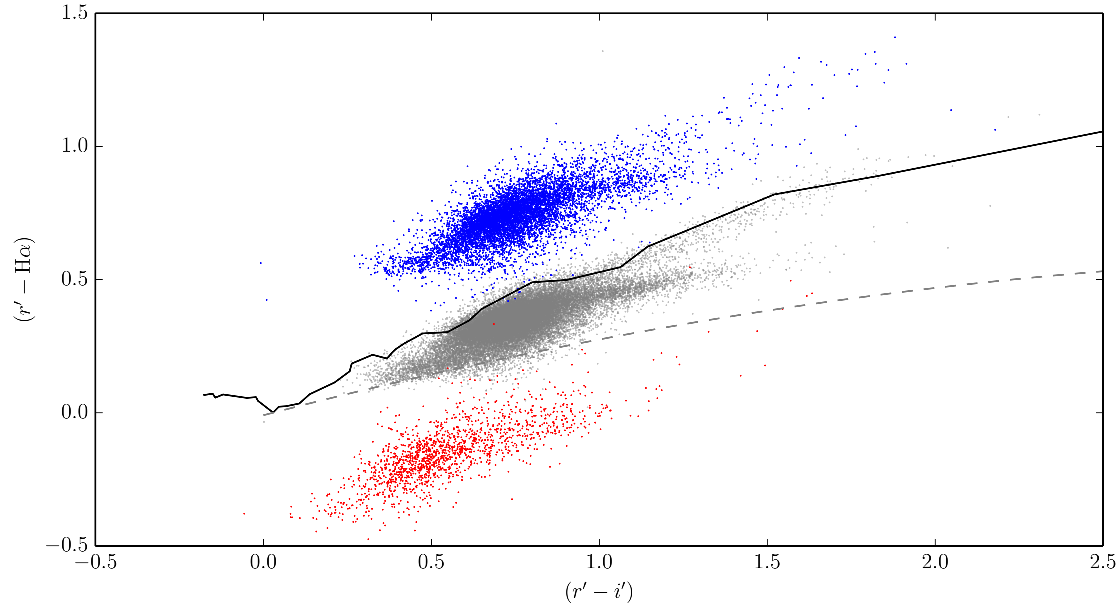
After excluding fields with poor photometry unrecoverable by recalibration alone (Section 3.1.1), the initial density map was used to identify errors in WCS parameters of a number of CCD frames (Section 3.1.2), before external surveys were used to calibrate and validate the IPHAS data (Sections 3.2 & 3.3). The reliability of sources detected by IPHAS was also assessed (Section 3.4).

3.1 Preparing IPHAS for release

3.1.1 Excluding poor fields

As seen in Section 2.1.2, the theoretical colour-colour plane is well defined by stars with a given type and reddening. In reality, not all photometry generates colour-colour diagrams as clean as those in Figure 2.4. All observations suffer from some degree of scatter in photometric measurements. However a few significant issues can degrade the quality of IPHAS photometry significantly (see Figure 3.1).

Gain variations adversely affect the count levels recorded over parts of CCDs, making them difficult to correct. These variations can modify magnitudes for one or more filters



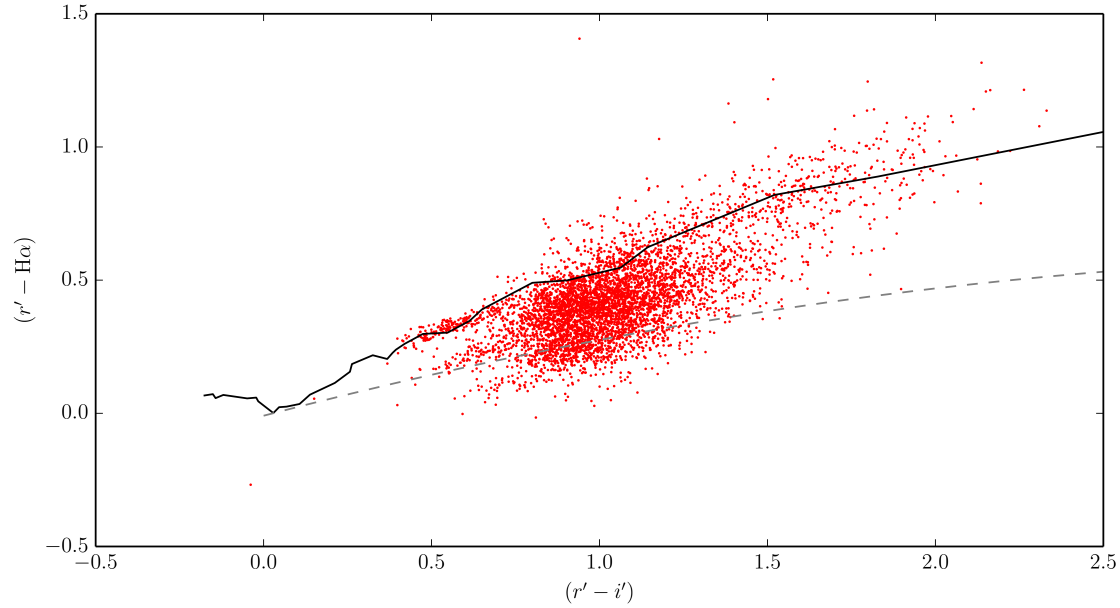


Figure 3.1: Examples of problematic photometry that was weeded out in the production of DR2. **Upper:** Two examples (in **red** and **blue**) of fields with one or more sets of magnitudes heavily affected (likely by transparency changes between filter observations) such that they lie far from the expected region of the colour-colour diagram. Objects within a 0.5 degree radius, selected from fields that were selected for DR2, are plotted in **gray**. The unreddened main sequence and early A-type reddening line are overplotted. **Lower:** Colour-colour diagram of field suffering from gain variation during readout; an effect which causes an increase in counts in a region of a CCD (or CCDs), shifting magnitudes in one or more filters.

by up to ≈20%, creating multiple overlapping stellar loci in the colour-colour plane. As a result this cannot be detected automatically - in order to identify fields with gain variations, all the fields were eyeballed. I eyeballed > 1000 fields myself, before putting together a series of webpages to distribute the necessary eyeballing across the IPHAS consortium. The responses to this campaign identified gain variations apparent in colour-colour diagrams, allowing affected fields to be removed or replaced by alternative observations in DR2.

A number of fields were affected by rapidly changing transparency during the observation of the three filters. Fields with this problem are typified by colour-colour diagrams lying far from the main stellar locus formed by the majority of surrounding fields (see Figure 3.1). The most extreme cases were identified by large median offsets in $(r' - i')$ or $(r' - \mathrm{H}\alpha)$, while the less serious cases were flagged during eyeballing (see above).

3.1.2 Identifying astrometric errors

The reduced image frames supplied by CASU contain an astrometric calibration determined by first initialising a WCS based on the pointing of the telescope and known properties of the WFC, including the change of scale with distance from the optical axis. The WCS parameters are improved in the pipeline by comparing the positions of objects detected in the frames to those present in the 2MASS (Skrutskie et al., 2006) catalogue. These astrometric solutions have been shown to result in accuracies better than $0.1''$ across the area of the WFC (González-Solares et al., 2008).

Though this approach performs well in the majority of cases, it will occasionally result in inaccurate WCS parameters for a given frame, particularly in busier fields where 2MASS astrometry is imperfect (due to confusion). The shifts in these WCS solutions are not large, at most a few arcseconds, most often causing inaccurate astrometry further from the optical axis. This fact allows the initial density map (see Section 2.3) to be used to diagnose these cases, exploiting the fact that sources in the affected regions also appear in overlapping fields, but in regions of the detector less or differently affected by the astrometry issues. Since the density map is generated by crossmatching fields with a match radius of $0.5''$, sources in these overlap regions may not be considered to be the same object, and as such contribute more than once to the object count. This results in regions of higher density appearing in the density map, conveniently conforming to the extent of the affected CCD. As a result these artificial overdensities are easily distinguished from genuine regions of high density (e.g. clusters) by their shapes - seen as sudden truncations of high density along CCD boundaries/corners.

These astrometric errors were identified by visually inspecting the density maps in all three bands. Figure 3.2 shows three cases - the first due to miscalibration affecting half of a CCD, the second displaying a problematic corner to a CCD, and the third example displaying a rarer case where sources falling at opposite ends of a CCD would be assigned incorrect coordinates, while objects laying in a strip across the centre were correctly assigned positions. After I had identified these problems, GB could alter the FITS header values corresponding to WCS quantities, correcting the astrometric solutions of the affected CCDs.

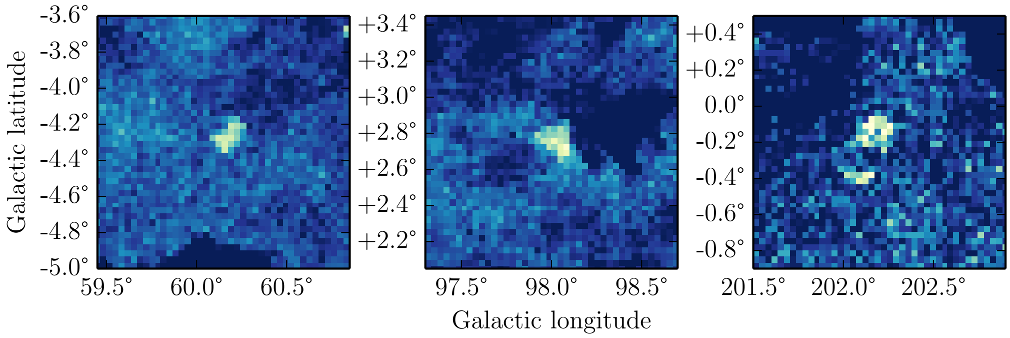


Figure 3.2: Examples of three regions of the density map displaying overdensities caused by improper astrometric calibration of WCS information.

3.2 Calibration using APASS

Attempts to globally calibrate IPHAS had been made prior to my involvement, that utilised the overlap method of Glazebrook et al. (1994) on the best set of IPHAS observations available, minimising the offsets between the magnitude zeropoints of neighbouring fields. This approach managed to reduce the scatter in photometry, although artificial gradients were introduced through this method, visible as waves when the zeropoint shift values were plotted in (ℓ, b). Anchor fields - observations deemed to originate from nights with the best photometric conditions - had been chosen by JD and set as fixed zeropoint values in an attempt to eliminate these variations. Insufficient numbers of anchor fields resulted in their being spread out too sparsely across the survey, thereby permitting artificial slopes in the corrections applied to appear in the stretches between anchors.

It was decided that introducing an external uniformly calibrated reference would allow fields to be corrected and fixed at a given zeropoint, allowing many more to serve as anchors; it was hoped this would eliminate slopes in the zeropoint corrections. The Sloan Digital Sky Survey (SDSS DR9) was initially considered, but dismissed for reasons of coverage - only a few SDSS stripes pass through the Galactic Plane (see Figure 3.3). These were instead used as a verification of the uniform calibration established (see Section 3.3). I took charge of obtaining transformations between photometric systems of APASS and IPHAS, and determining the shifts needed for IPHAS fields to be brought in line with APASS photometry.

3.2.1 The AAVSO Photometric All-Sky Survey

The AAVSO Photometric All-Sky Survey (APASS) is an effort to create a catalogue of all objects between $10.0 \leq V \leq 17.0$ in $BVg'r'i'$ filters (Henden & Munari, 2014). Observing from Dark Ridge Observatory in New Mexico for the Northern half of the survey, APASS uses a pair of 20cm f/3.6 astrographs giving a 8.41 sq. deg. field of view. Their detectors, with arrays of 4096×4096 pixels, achieve a 2.6 arcsec pixel^{-1} plate scale.

In a similar fashion to IPHAS, the APASS field pattern is designed with two sets of field

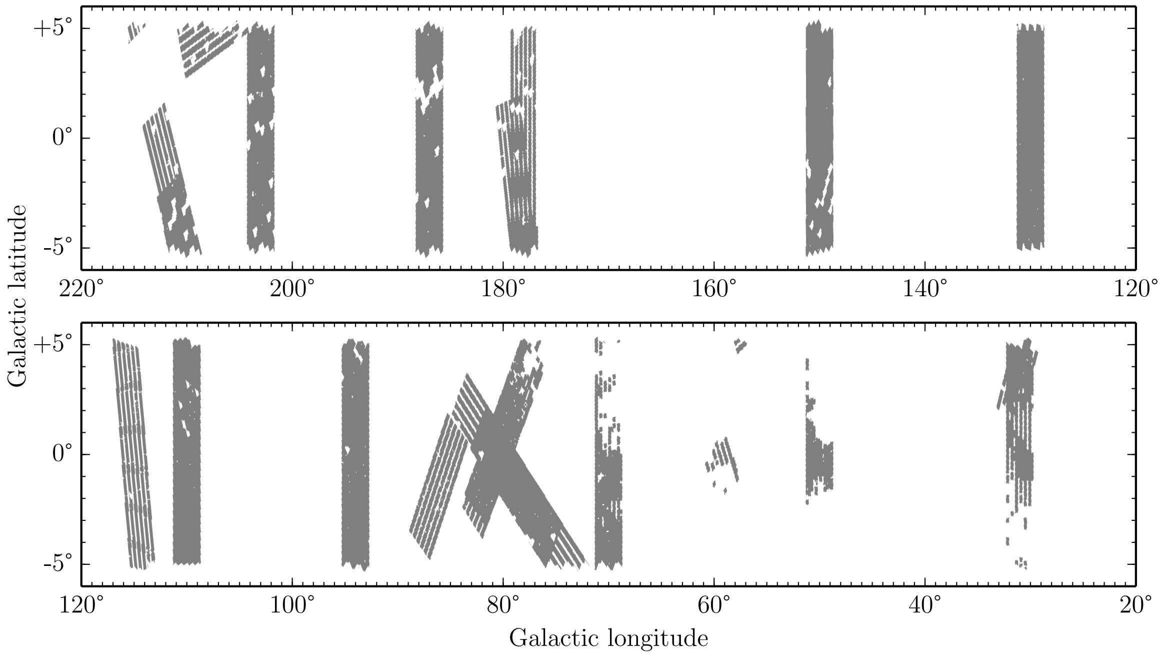


Figure 3.3: The intersection of SDSS DR9 and IPHAS DR2 footprints. This intersection is also valid for SDSS DR7 (see Section 3.4) as imaging data added between the two data releases were towards the Southern Galactic Cap.

centres offset from one another (in the case of APASS, offset by 1.45° in both RA and DEC), and each field overlapping its neighbour by 5%. APASS observes fields at least twice, ensuring coverage of every point in the sky more than four times in order to reduce observational errors. Exposures of 90s are taken in r' and i' with one of the two telescopes (along with the V, while the other telescope observes in B and g').

Figure 3.4 shows the coverage of the APASS DR7 catalogue over the whole sky and over the IPHAS footprint. APASS covers ≈50% of the Northern Galactic Plane, primarily at $\ell > 120°$ - a significantly larger region than SDSS DR9. The IPHAS data in this region are those obtained in the winter season - it is fortunate that this is the region which can be tied to APASS as it is the region most in need of calibration.

3.2.2 Crossmatching

The IPHAS DR2 fields were crossmatched against the APASS DR7 catalogue using STILTS, set to a 0.5″ match radius. The distribution of the separations between IPHAS and APASS sources (see Figure 3.5) suggests that this may be an overly conservative match radius, as the distribution reaches a maximum at ≈ 0.2″ and its tail is obviously truncated at 0.5″. The tighter distribution of offsets between IPHAS and SDSS sources suggests that APASS astrometry is the source of the large positional offsets.

Magnitude ranges of $13 < r'_{APASS} < 16.5$ and $12.5 < i'_{APASS} < 16.0$ were imposed for the crossmatch, to account for the faint limit of APASS and to exclude sources approaching

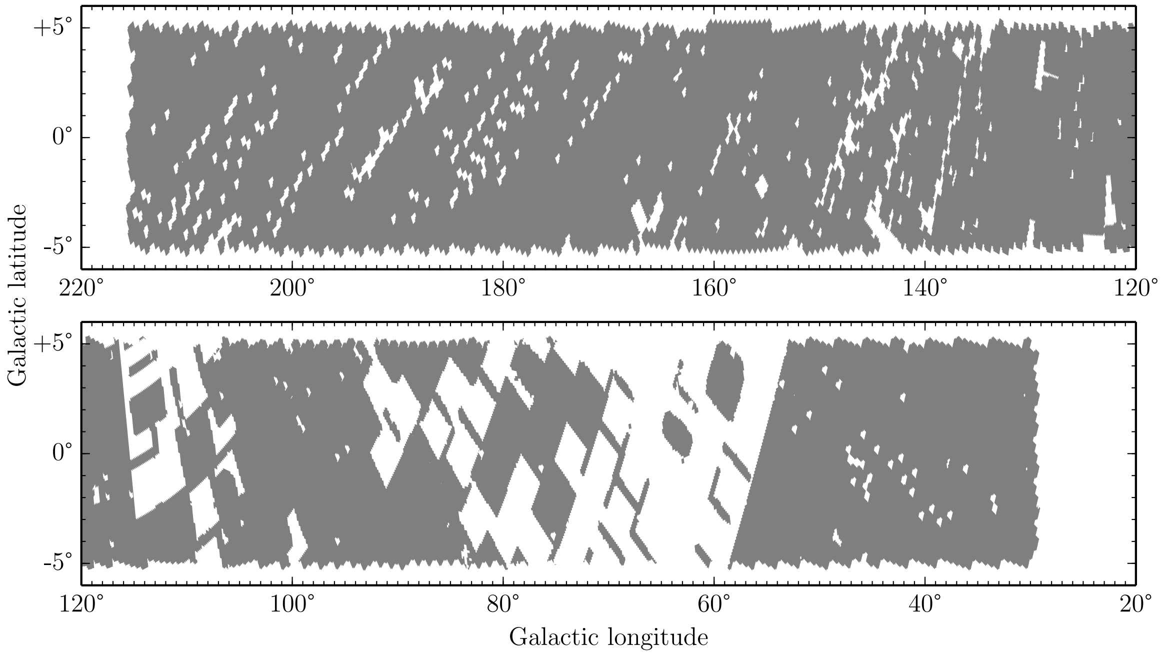


Figure 3.4: The intersection of APASS DR7 and IPHAS DR2 footprints. Large contiguous white regions inside the IPHAS DR2 footprint are due to missing APASS data - fields which have either not been observed or not observed a sufficient number of times in good conditions to be included in APASS DR7.

the saturation limit of IPHAS. Use of these match criteria resulted in ≈223,603 sources being associated across the two surveys - a respectable number of objects with which to obtain colour transformations, while ensuring that spurious crossmatches were minimised. A total of 8,199 fields (of the 14,415 in DR2) contained at least one source matched in APASS DR7.

3.2.3 Preliminary transformations and shifts

APASS uses Sloan filters for its $g'r'i'$ observations. There are no published curves or transformations to other systems for their filters, requiring the calculation of colour equations to convert from APASS to IPHAS magnitudes. Preliminary colour transformations were obtained by least squares fitting of the magnitude differences (APASS - IPHAS) of the crossmatched objects against their APASS $(r' - i')$ colour. These transformations were used to convert the APASS magnitudes to the IPHAS filter set equivalents. In the fitting of the data, an iterative sigma clipping was applied, removing objects greater than 3σ from the best linear fit to the data for five iterations.

Preliminary transformations were obtained of the form:

$$r'_{IPHAS} = r'_{APASS} - 0.122 + 0.041(r' - i')_{APASS} \tag{3.1}$$

$$i'_{IPHAS} = i'_{APASS} - 0.367 + 0.021(r' - i')_{APASS} \tag{3.2}$$

Based on these transformations, the APASS source magnitudes were transformed to the

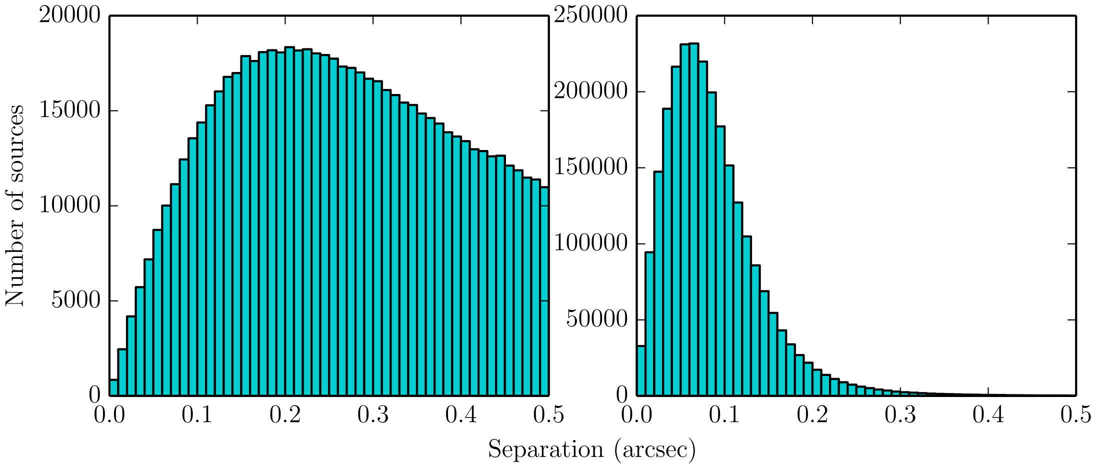


Figure 3.5: Distribution of separations between sources crossmatched between IPHAS DR2 (**left**) and APASS DR7, and between IPHAS DR2 and SDSS DR9 (**right**).

IPHAS system, allowing shifts to be calculated for IPHAS fields containing crossmatched sources. A cutoff of 30 crossmatched sources per field was imposed, excluding any fields with fewer matches as unreliable. Median shifts were calculated for each of the remaining 6,693 fields, and applied to fields displaying more than a 0.03 mag median shift with respect to APASS. Fields showing a smaller shift were not modified, as these fall within the 3% photometric accuracy aimed for with IPHAS DR2. The approach of shifting all fields (regardless of how closely they agreed with APASS) to zero offset with respect to APASS was avoided, as this would require an extremely high degree of confidence in the uniformity of APASS photometry. No survey can hope to be perfect, and inconsistencies between surveys will always be present at the low level (see Section 3.3.3 for a comparison between APASS and SDSS as an example).

The calculated offsets were applied to the data, resulting in a dataset more closely calibrated to APASS. The Glazebrook algorithm was run by GB on the roughly calibrated data, setting as anchors fields that either agreed with APASS to within 3% or had had offsets applied.

3.2.4 IPHAS-APASS colour equations

The previous colour transformations (Equations 3.1 and 3.2) were derived using uncalibrated IPHAS data, i.e. data containing a spread in measured photometry due to the observational scatter that the calibration is intended to correct. With the preliminary calibration generated as described in Section 3.2.3, this scatter is reduced, allowing more accurate transformations to be calculated. Using the same steps on the calibrated data produced the transforms:

$$r'_{IPHAS} = r'_{APASS} - 0.119 + 0.032(r' - i')_{APASS} \tag{3.3}$$

$$i'_{IPHAS} = i'_{APASS} - 0.360 + 0.003(r' - i')_{APASS} \tag{3.4}$$

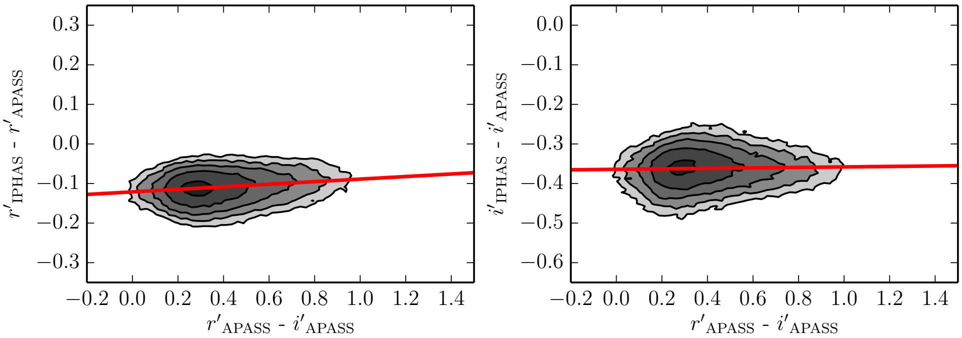


Figure 3.6: Best fit colour transformations from APASS to IPHAS magnitudes for r' (**left**) and i' (**right**), described by Equations 3.3 & 3.4.

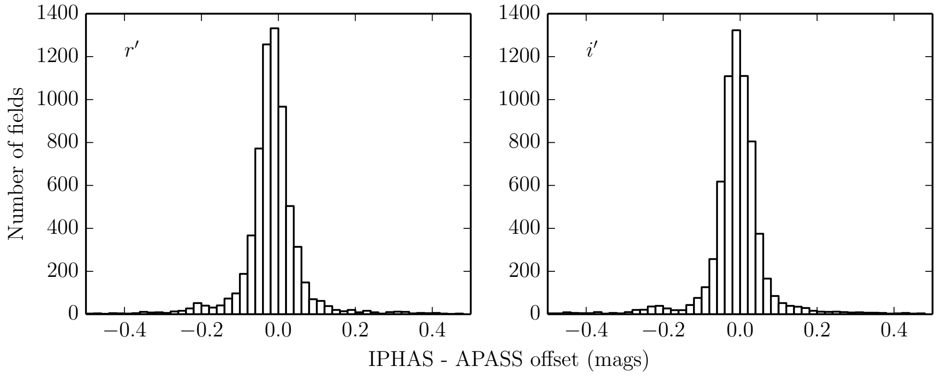


Figure 3.7: Distributions of median offsets to be applied to fields in IPHAS DR2 to bring the median shift per field to zero. Only IPHAS DR2 fields containing more than 30 sources matched to APASS DR7 are included here. **Left:** r'-band offsets. **Right:** i'-band offsets.

with an RMS of 0.041 and 0.051 respectively. It can be seen that the calibration only slightly modified the coefficients of the transforms. Figure 3.6 shows the transforms for the range $-0.2 < (r' - i')_{APASS} < 1.5$. While the colour terms are small, both filters require a large shift to bring the APASS magnitudes into the IPHAS system. This is the expected consequence of APASS magnitudes being given in the AB system (following the precedent set by SDSS) being transformed into magnitudes based on the Vega system used by IPHAS. The slightly larger colour term for the r'-band transformation indicates that the difference between APASS and IPHAS r' effective wavelengths is greater than for i' filters.

As the relatively bright magnitude range spanned by APASS will not capture a large number of distant reddened sources, there was no compelling reason to produce colour equations for regions of differing reddenings; in addition, the small RMS on the global transforms suggest that they are quite robust over the sky area in which both surveys intersect.

These corrected equations were used to repeat the calculation of offsets necessary to calibrate the DR2 fields, which can be seen in Figure 3.7. The tails of the distributions reach out beyond ±0.2 mag, and a small number of fields originating from a run suffering slightly poorer quality photometry can be seen at shifts between -0.3 and -0.2 mag.

Band	Pre-shifts		Post-shifts	
	Mean shift	σ	Mean shift	σ
r'	0.014	0.104	0.000	0.011
i'	0.007	0.108	0.000	0.011

Table 3.1: Magnitude shifts between IPHAS DR2 fields and APASS where more than 30 objects were crossmatched, before and after the shifts minimising differences between the two surveys were applied. The pre-shift distribution is shown in Figure 3.7.

Band	Pre-shifts		Post-shifts	
	Mean shift	σ	Mean shift	σ
r'	0.016	0.088	-0.001	0.029
i'	0.010	0.089	-0.002	0.032

Table 3.2: Magnitude shifts between calibrated IPHAS DR2 and SDSS DR9 data for the 2550 fields where more than 30 objects were crossmatched. The corresponding distributions are shown in Figure 3.10.

After applying the calibration, the median offsets were recalculated to ensure improved agreement between the two surveys. The improvement can be seen in Table 3.1. Figures 3.8 and 3.9 show the median offset of all fields distributed spatially, before and after the uniform calibration.

Following the calculation of the second set of shifts the recalibrated data were used to recalculate colour equations and shifts for further iterations; it was found that the calibration had converged to a sufficient degree with Equations 3.3 and 3.4.

3.3 Validation using SDSS

Figure 3.9 shows that the difference between APASS and IPHAS have been greatly reduced. This does not impart any new information - this improvement is the expected result of calibrating the survey against APASS as a reference. In order to verify that the calibration has indeed improved the photometry of DR2, an independent data set was needed. SDSS DR9 was chosen as its filters are similar to those of both IPHAS and APASS, and though its coverage is significantly less than either survey (see Figure 3.3), the SDSS stripes which run through the Galactic plane are well spaced, and also cover areas not directly calibrated to APASS. An SDSS comparison will therefore be able to serve as a check on what degree of improvement has been made by setting fields calibrated to APASS as anchors in the Glazebrook algorithm.

All objects in SDSS DR9 falling inside the IPHAS footprint were crossmatched with DR2 fields, with a crossmatch radius of 1″ and enforced magnitude ranges of $15 < r'_{SDSS} < 18.0$ and $14.5 < i'_{SDSS} < 17.5$. This resulted in a total of 1,214,577 sources being crossmatched over 2,657 IPHAS fields - roughly a third of all fields.

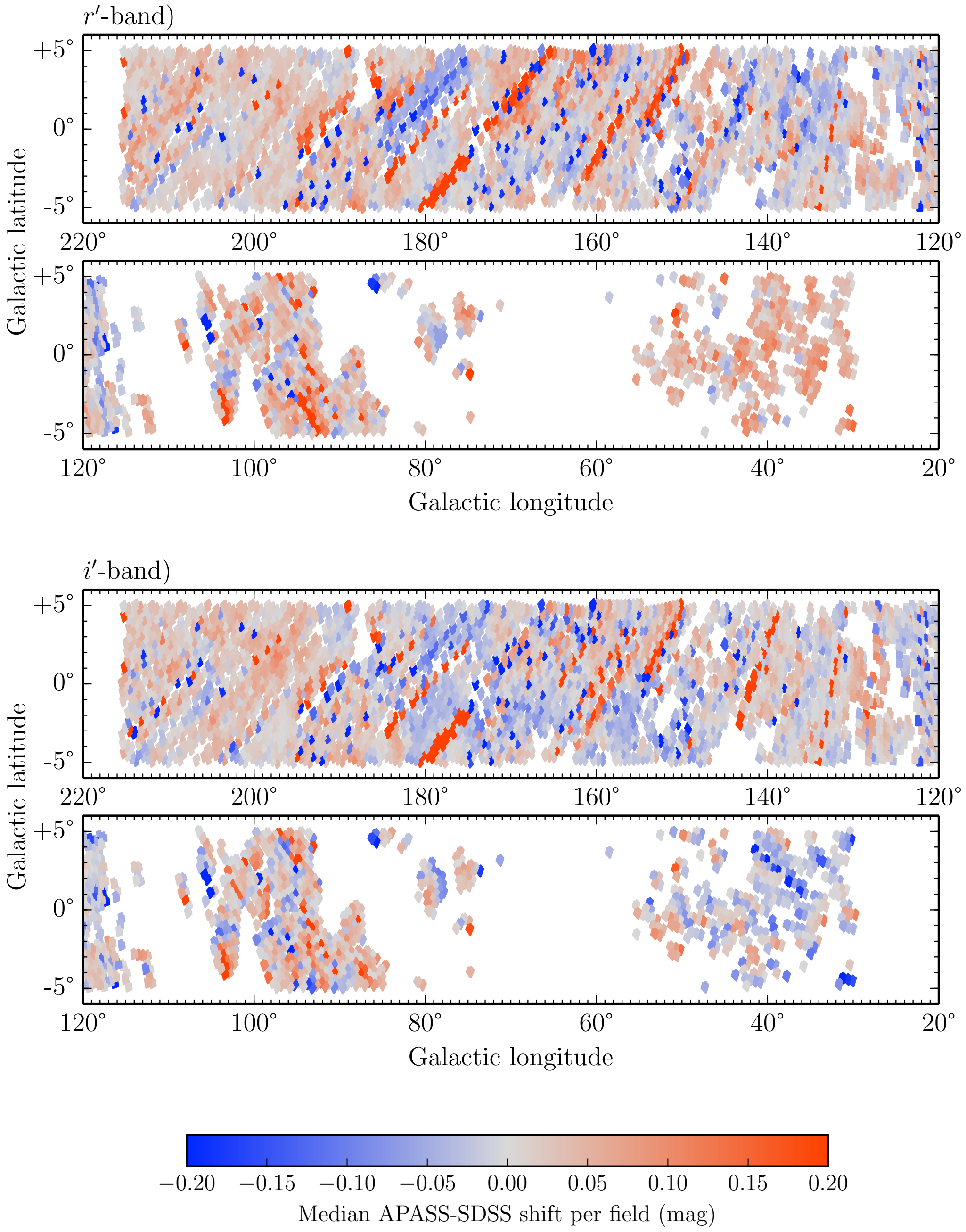


Figure 3.8: Spatial map of uncalibrated IPHAS fields with more than 30 objects crossmatched between IPHAS and APASS, with the colour of points denoting the median APASS-IPHAS shift - i.e. the shift that would need to be applied to all sources in an IPHAS field to bring it in line with the reported APASS magnitudes. **Upper:** r' band shifts. **Lower:** i' band shifts.

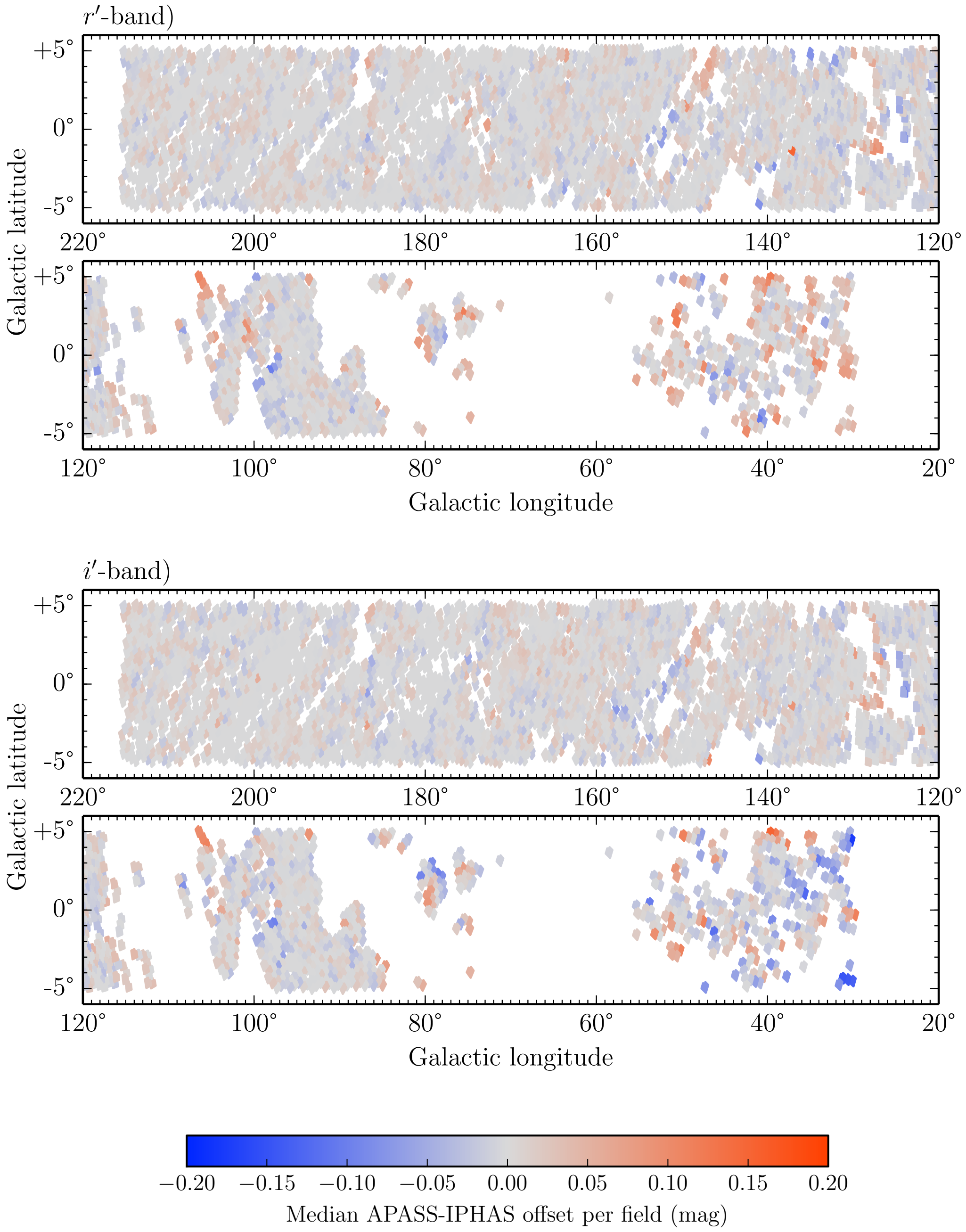


Figure 3.9: Same as Figure 3.8 but after calibration against APASS and GB running the Glazebrook algorithm.

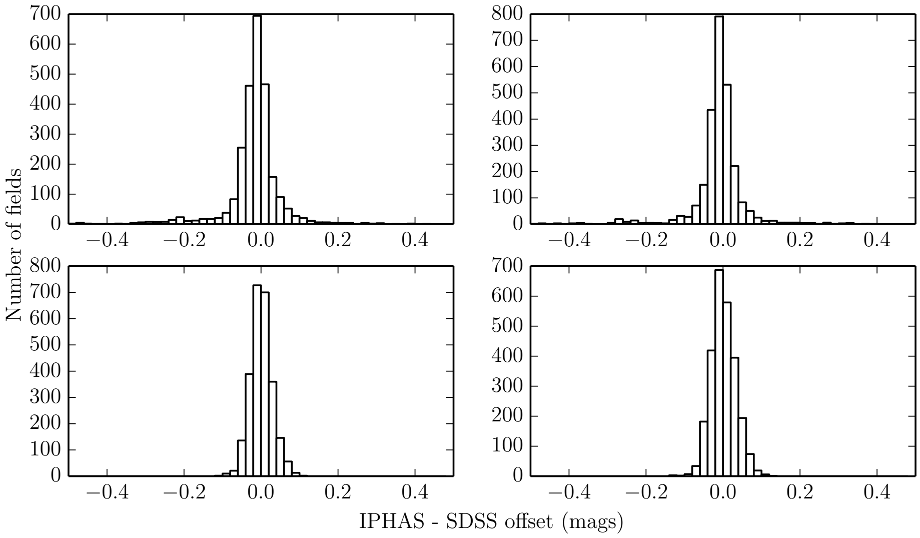


Figure 3.10: Distributions of median offsets between IPHAS DR2 and SDSS DR9 for IPHAS fields containing more than 30 crossmatched sources. **Upper:** Prior to calibration. **Lower:** Post-calibration. **Left:** r'-band. **Right:** i'-band.

3.3.1 IPHAS-SDSS colour equations

Colour transformations were obtained by least squares fitting of the magnitude differences of the crossmatched objects against their SDSS $(r' - i')$ colour. The calibrated data was used with no need for preliminary transformations (as was the case for APASS).

A global transformation (see Section 3.3.2) was used to transform all SDSS objects in the IPHAS footprint onto the IPHAS magnitude system. For the purposes of validating the calibration the relatively low number of intrinsically red objects (mainly M dwarfs) in low reddening areas were effectively neglected - the vast majority of red objects come from regions of moderate to high extinction, rendering the global transformation more suitable for describing these reddened objects (see Section 3.3.2 for a discussion of transformation variations with longitude). The global transformations were determined to be:

$$r'_{IPHAS} = r'_{SDSS} - 0.093 - 0.044(r' - i')_{SDSS} \tag{3.5}$$

$$i'_{IPHAS} = i'_{SDSS} - 0.318 - 0.095(r' - i')_{SDSS}. \tag{3.6}$$

with RMS values of 0.045 and 0.073 respectively. As with APASS, SDSS provides magnitudes in the AB system, requiring large offsets to transform magnitudes into the Vega system used by IPHAS. Figure 3.13 shows these transforms over the range $0 < (r' - i') < 1.5$.

The colour terms in Equations 3.5 and 3.6 are stronger than in APASS transformations (in particular for the i'-band), suggesting a greater difference in the filter profiles. Figure 3.14 compares both sets of filters, showing that both IPHAS filters capture greater flux towards redder wavelengths than their SDSS counterparts, with the i' filter extending an additional $\sim$+30 nm. Though no filter profiles have been published for APASS, our experience amounts

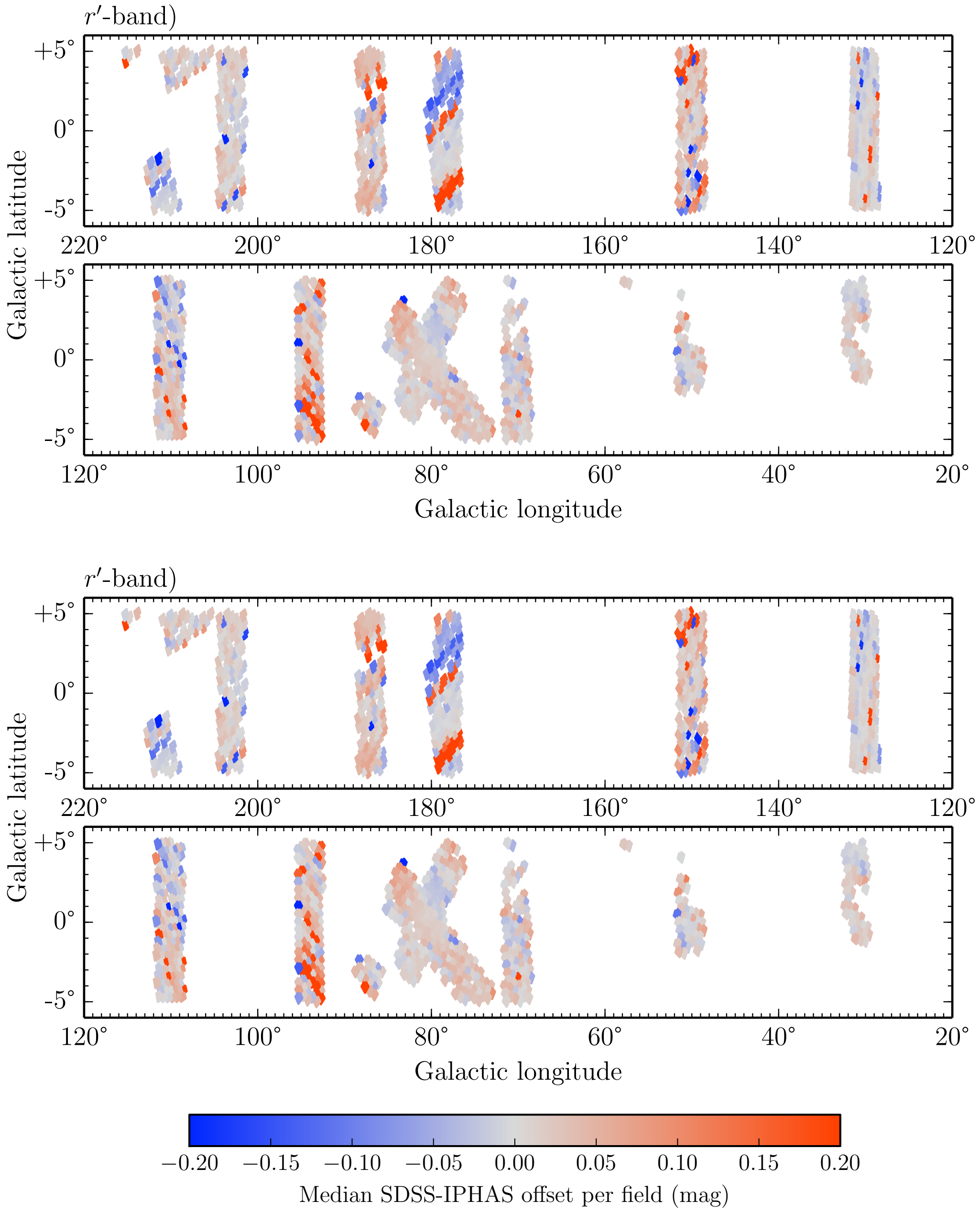


Figure 3.11: Spatial map of IPHAS fields (before uniform calibration) with more than 30 objects cross-matched between IPHAS DR2 and SDSS DR9, with the colour of fields denoting the median SDSS-IPHAS shift. **Upper:** r' band shifts. **Lower:** i' band shifts.

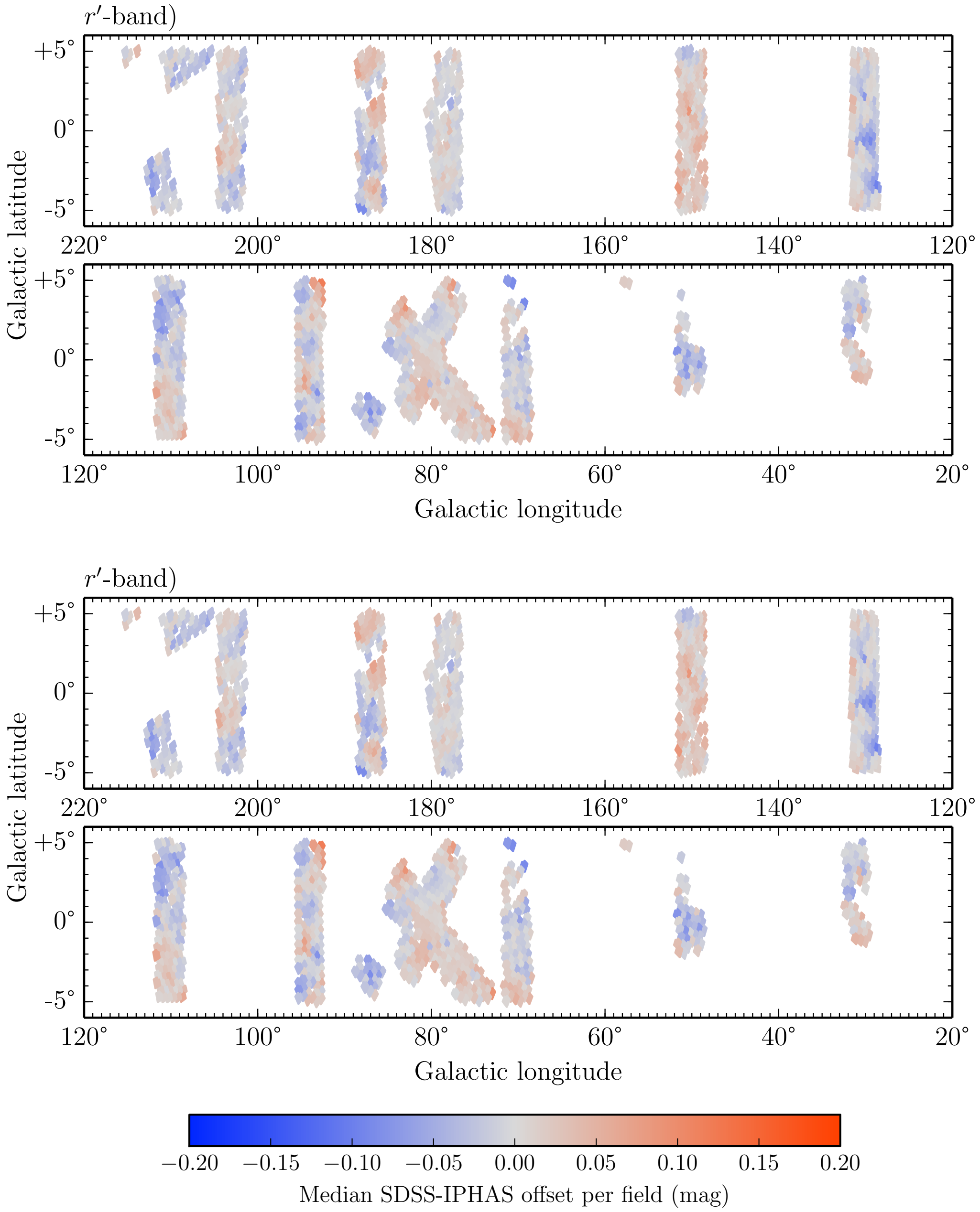


Figure 3.12: Same as Figure 3.11 but after calibrating IPHAS DR2 fields intersected by APASS DR7, and GB running the Glazebrook algorithm to calibrate the remainder of the IPHAS fields. The marked improvement is clearly seen.

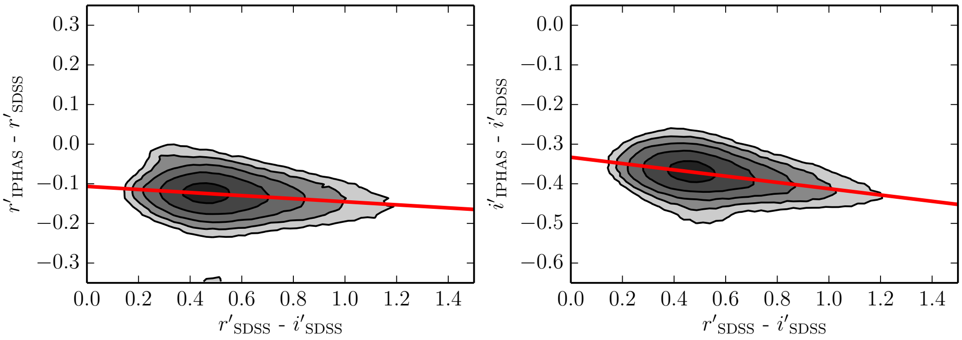


Figure 3.13: Best fit colour transformations from SDSS to IPHAS magnitudes for r' (**left**) and i' (**right**), described by Equations 3.5 & 3.6.

to evidence their filters more closely correspond to WFC $r'i'$ filters than do those used by SDSS.

Table 3.2 shows the improvement in agreement between IPHAS and SDSS after calibrating to APASS, and spatial maps of the shifts are presented in Figures 3.11 and 3.12. The distribution of residual shifts is consistent with a normally distributed population of field offsets with $\sigma \approx 0.03$ mag. A σ of 0.03 here is the result of the standard deviation of both surveys added in quadrature; the claimed accuracy of the SDSS calibration is much better than this (given as 0.008 mag in r', 0.007 mag for i' by Padmanabhan et al. (2008)) - on this basis, the majority of the combined σ is due to the calibration of IPHAS.

3.3.2 Variation with longitude

When deriving colour transformations to convert SDSS magnitudes to the IPHAS system (Equations 3.5 and 3.6), the fact that the magnitude range over which the two surveys overlap could probe a variety of populations was taken into account. While the APASS crossmatch faint limit was imposed at r'=16.5, the SDSS comparison reaches to r'=18.0; the effect of this deeper limit is evident in the number of sources used to derive the colour transformations: 223,603 for APASS compared to 1,214,577 for SDSS.

The steps detailed in Section 3.3.1 used to obtain the transformations were repeated over subsets of the IPHAS/SDSS crossmatch, roughly corresponding to a number of the SDSS strips (seen in Figure 3.3) - i.e. subsets separated in Galactic longitude. The i'-band transformations were used to investigate any effect as the differences between IPHAS and SDSS i' filters are more significant than for the r' filters (see Figure 3.14).

Figure 3.15 shows the fits made to each of the crossmatch subsets. Details of the fits are laid out in Table 3.3. It can be seen that the colour equations vary with position in the Galactic Plane: this is due to the variation of populations that are being probed owing to the change in typical reddening with ℓ. Despite these variations, it can be seen that the shape of the transformation is never modified such that it no longer resembles the global transformation given by Equation 3.6.

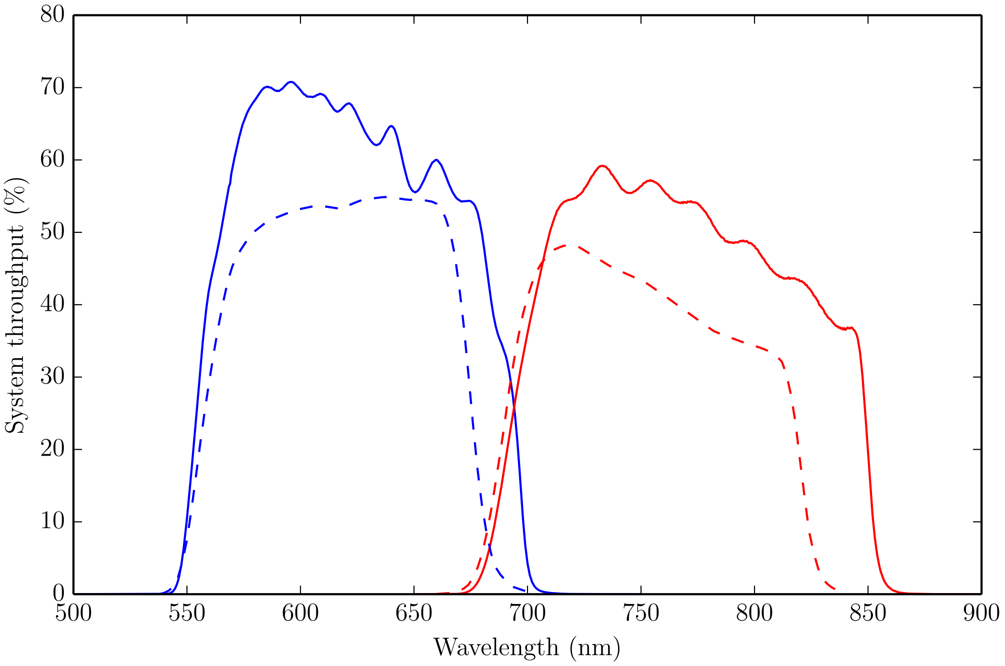


Figure 3.14: Comparison of r' (**blue**) and i' (**red**) filter profiles for IPHAS (**solid**) and SDSS (**dashed**). Quantum efficiency of the WFC detector has been accounted for, while atmospheric transmission has been left out for each survey. SDSS filter profiles were taken from Doi et al. (2010).

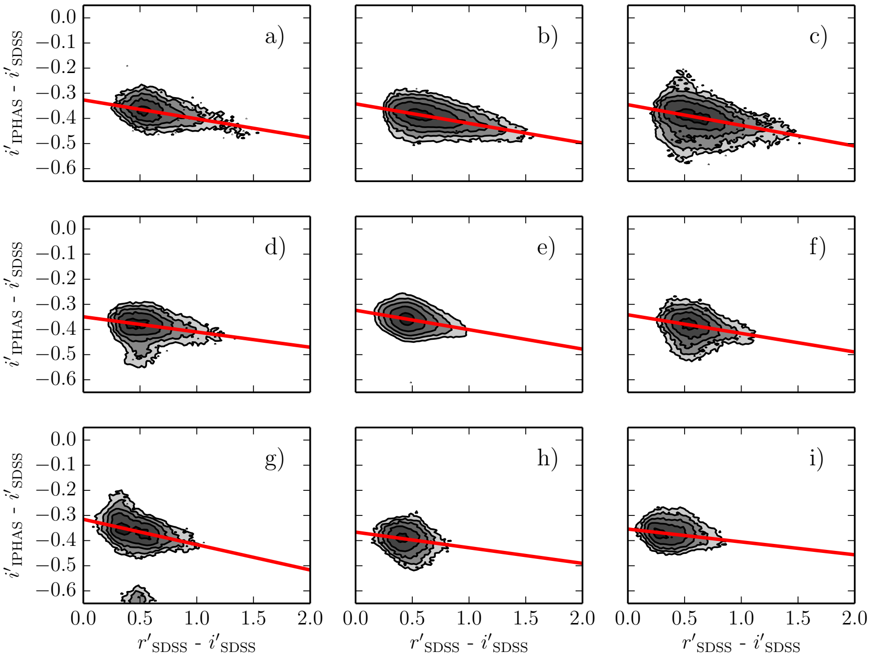


Figure 3.15: Best fit transformations between i'-band magnitudes of SDSS and IPHAS, as in Figure 3.13, separated roughly into SDSS strips. Table 3.3 provides the coordinates, number of sources per sample, and the coefficients of the best fit equations.

subset	ℓ bounds min	 max	source count	offset	colour term
a	65°	72.5°	39822	-0.326	-0.075
b	72.5°	90°	188799	-0.342	-0.078
c	90°	100°	120557	-0.346	- 0.082
d	100°	120°	91966	-0.349	-0.061
e	120°	140°	413242	-0.324	-0.077
f	140°	160°	81433	-0.342	-0.074
g	160°	182.5°	101928	-0.315	-0.101
h	182.5°	190°	63120	-0.366	-0.062
i	200°	220°	101587	-0.354	-0.051

Table 3.3: Details of fits to data made in Figure 3.15. Offset and colour terms are the coefficients of the fits, which were found to be -0.318 and -0.095 respectively over the whole crossmatched dataset.

3.3.3 Comparing SDSS to APASS

Upon comparing the calibrated DR2 data to APASS and SDSS, it appeared that IPHAS had been successfully calibrated to 3% accuracy with respect to both external references. The occasional outlying field, with a large (>0.03 mag) median magnitude offset between crossmatched IPHAS and APASS/SDSS sources, is expected assuming a distribution of offsets with σ=0.03 mag.

However, these external surveys have been treated as reporting ‘the truth’, i.e. that they report perfectly the magnitude of all sources. As this clearly cannot be the case, a comparison between APASS and SDSS is now offered, to close the loop.

Far fewer sources were crossmatched between the two surveys, owing to the requirements that

- $13 < r'_{APASS} < 16.5$ and $12.5 < i'_{APASS} < 16.0$
- $15 < r'_{SDSS} < 18.0$ and $14.5 < i'_{SDSS} < 17.5$
- $20 < \ell < 220$ and $-6 < b < 6$.

The last constraints were included to avoid including any regions of the sky outside of the IPHAS footprint (as shown in Figure 2.3) which could introduce unexpected populations of objects that would bias the colour equations away from those applicable in the Galactic Plane. These constraints led to 18,602 sources being crossmatched across the two surveys.

Figure 3.16 shows the best fit transformations fit to the data, given by:

$$r'_{APASS} = r'_{SDSS} + 0.024 - 0.053(r' - i')_{SDSS} \tag{3.7}$$

$$i'_{APASS} = i'_{SDSS} + 0.042 - 0.074(r' - i')_{SDSS}. \tag{3.8}$$

These transformations differ from those involving IPHAS in that there is no transformation between AB and Vega magnitude systems - both APASS and SDSS work in AB magnitudes, resulting in a small absolute offset between the two systems.

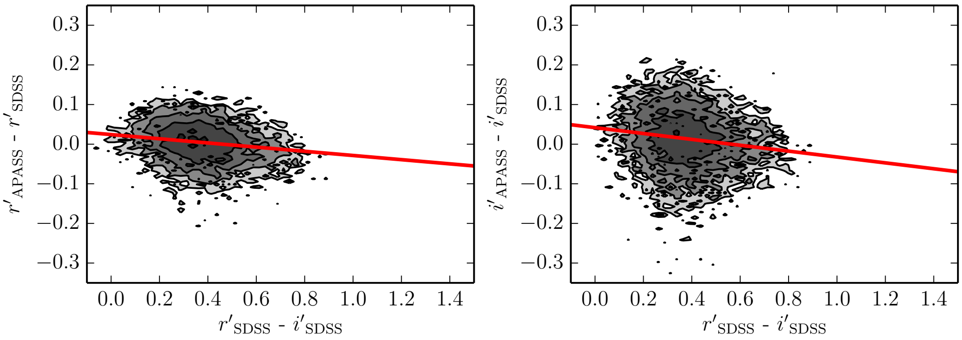


Figure 3.16: Best fit transformations from SDSS to APASS magnitudes based on the limited matched objects over the Northern Galactic Plane.

These transformations were used to convert SDSS magnitudes into APASS magnitudes, with the offsets between the two surveys once brought onto the common APASS system illustrated in Figure 3.17. This cannot be directly compared to the spatial maps comparing IPHAS to SDSS or APASS, as these were split into units of IPHAS fields, with the median of each field being plotted. Here each object is plotted with its colour representing its APASS-SDSS offset. It can be seen that sources with similar offsets are grouped together across the survey - this shows as anticipated that both surveys are not always consistent with one another; regions (e.g. ≈(187°, -4.5°) in the i'-band) can show inconsistencies of >0.03 mag between APASS and SDSS.

These inconsistencies influenced the decision not to tie the calibration of IPHAS too tightly to APASS (as discussed in Section 3.2.3), as this could propagate errors due to APASS into IPHAS photometry. The approach adopted has been to use APASS lightly, providing a guide to a mean level that minimisation via the Glazebrook algorithm otherwise struggles to find.

3.4 Survey reliability

If a photometric survey is to take its place as a quantitative resource then it needs to be accompanied with statistics describing the quality of its data products.

Characterisation of a survey's *completeness* provides an estimate of the fraction of objects not included in the survey - either due to non-detection from an image or not making it through the reduction pipeline. *Confusion* is the effect of losing sources due to the blending of objects, a source of incompleteness more likely to affect dense fields or those in poor seeing. The determination of these quantities will be discussed in Chapter 4.

Reliability is the estimate of what fraction of detections are repeatable. Skrutskie et al. (2006) used a practical definition of the ratio of actual astronomical sources to the total number of sources in their 2MASS catalogue, comparing their detections both internally across multiple observations and externally against SDSS. The reliability for IPHAS is dealt with next.

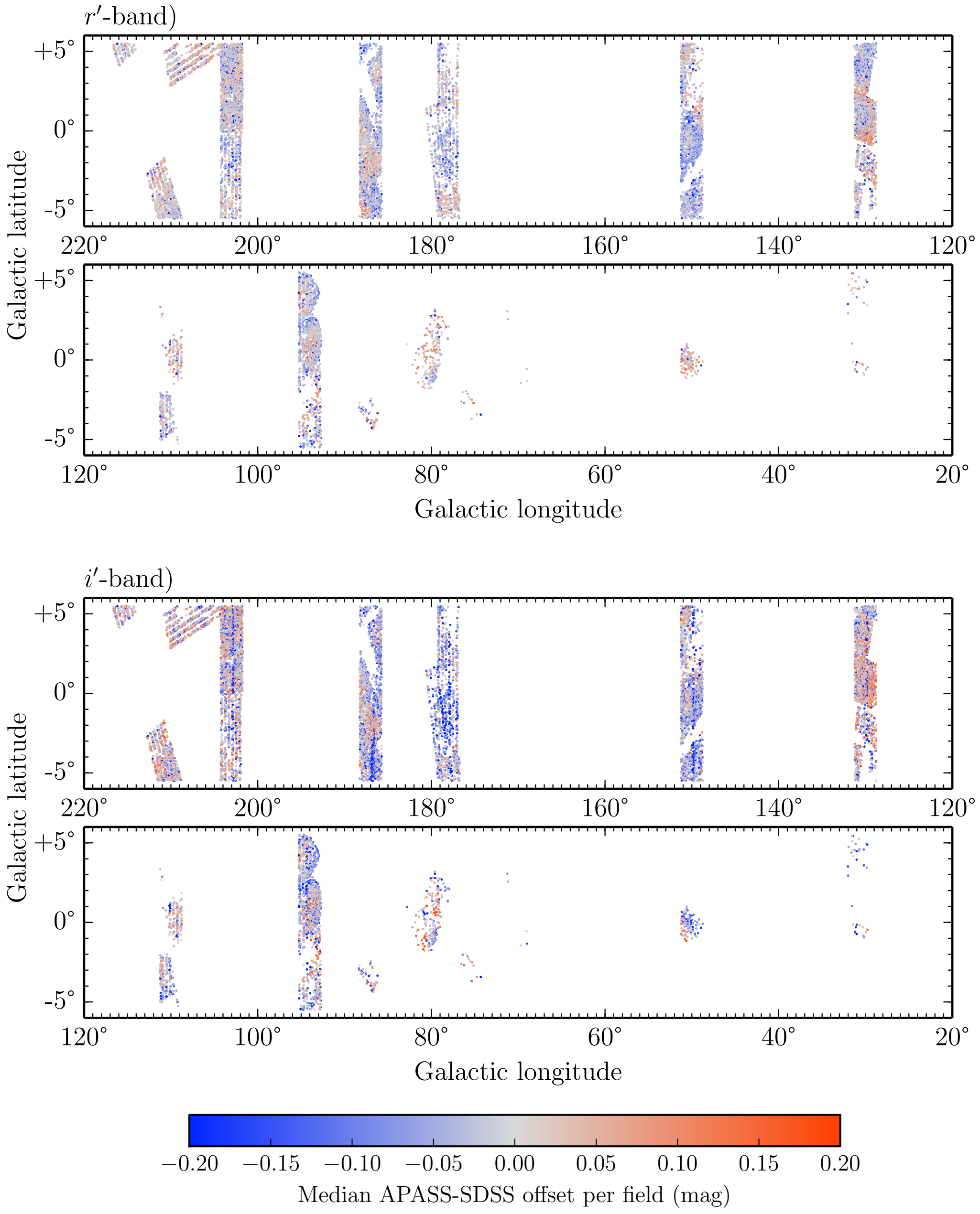


Figure 3.17: Spatial map of relatively bright sources crossmatched between APASS DR7 and SDSS DR9, coloured to denote the APASS-SDSS magnitude difference. **Upper:** r' band shifts. **Lower:** i' band shifts.

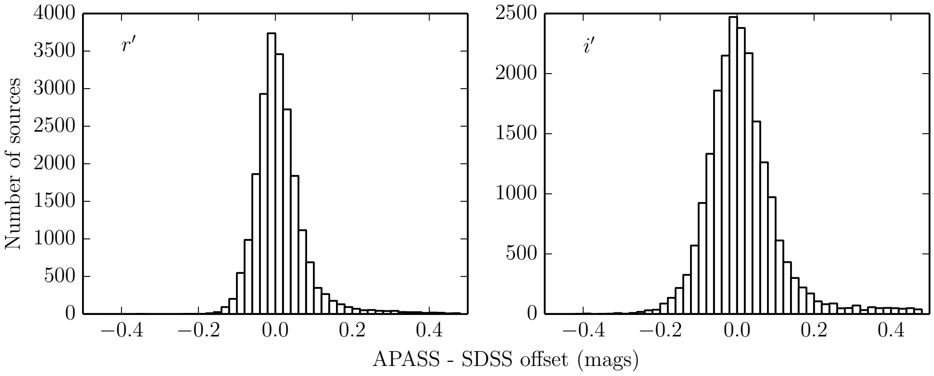


Figure 3.18: Distributions of magnitude offsets between APASS DR7 and SDSS DR9 for all sources cross-matched between the two surveys. SDSS magnitudes have been shifted onto the APASS magnitude system using Equations 3.7 & 3.8. **Left:** r'-band. **Right:** i'-band. The disagreement between the two surveys is greater in the i'-band, consistent with a larger scatter between APASS-calibrated IPHAS and SDSS magnitudes.

Region (ℓ, b)	Raw	nBands == 3	rErr < 0.1	iErr < 0.1	a10	a10point
(149.39, 4.06)	92.95	98.83	98.53	98.54	99.67	99.82
(186.59, -2.50)	89.99	98.82	98.34	98.44	99.36	99.57
(202.70, -1.75)	89.42	98.46	97.8	98.11	99.4	99.65

Table 3.4: Reliability percentages for three regions of IPHAS, where a source is considered reliable if it also appears in SDSS DR7 within a matching radius of 0.5″. **Column definitions:** 'Raw' sources counts every row appearing in the IPHAS DR2 catalogue in the 1 deg^2 region. 'nBands==3' means sources were detected in r', i' and Hα images. $rErr/iErr < 0.1$ means that the source has an error less than 0.1 mag in the given filter. 'a10' and 'a10point' are quality flags provided in DR2 (Barentsen et al., 2014). The former requires consistency between the magnitude measured in the default 2.3″ aperture and the smaller 1.2″ aperture, essentially rejecting sources with poor background subtraction. The latter requires in addition that sources not be deblended and not lie in the vicinity of an extremely bright source.

3.4.1 Cross-matching to test reliability

Regions of contiguous coverage in both IPHAS and SDSS DR9 were chosen for comparisons - the strips were located in the direction of the outer Galaxy, as early plane SDSS strips are not complete as of DR9. The two surveys were crossmatched, and IPHAS sources deemed reliable if a match was found in SDSS.

The estimated reliability recovered when comparing IPHAS DR2 against SDSS DR9 initially seemed low (∼95%), which upon closer investigation was due to the effect of stars in the range $8 < r' < 12$. Stars brighter than this range create an optional exclusion zone in the IPHAS DR2 catalogues comparable to the impacted area in SDSS DR9, leading to good agreements of object counts with the exclusion zone applied to both catalogues. Stars in this range, however, are not assigned an exclusion zone in IPHAS DR2, while still causing objects to be missed (or at least not appear) in SDSS DR9. It was found that these 'missing' objects were present in previous data releases of SDSS, such that crossmatching IPHAS against SDSS DR7 provided much more reassuring reliability estimates. Figure 3.19 compares the point source catalogues of IPHAS to SDSS DR7 and DR9, showing the missing sources in DR9.

Table 3.4 shows the reliability for three 1 deg^2 regions of IPHAS. The cross-match radius

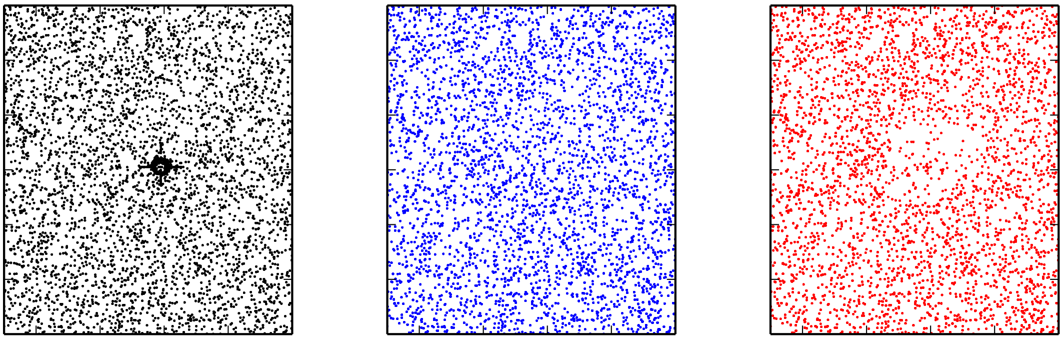


Figure 3.19: Cut out of point source catalogues for region centred on a bright ($V = 8.25$) star for **Left:** IPHAS DR2, **Centre:** SDSS DR7, and **Right:** SDSS DR9. It can be seen that in SDSS DR9, sources previously included in DR7 around the bright star are missing.

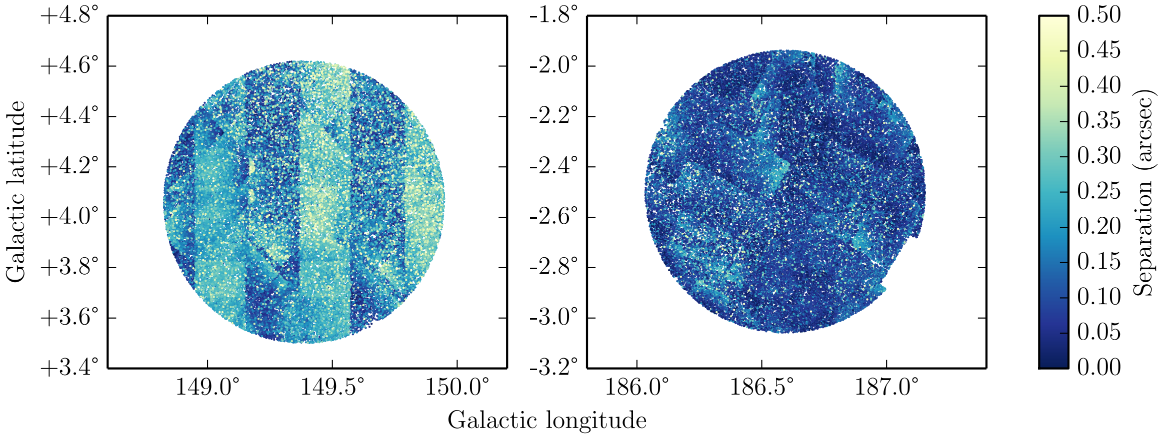


Figure 3.20: Spatial plots for two cross-matched regions coloured according to the separation between sources in IPHAS DR2 and SDSS DR7. The left-hand plot shows a larger ($\approx 0.25''$) separation for alternating strips, indicating that astrometric shifts are originating from SDSS. The right-hand plot shows a pattern that indicates that the largest astrometric shifts originate from IPHAS (specifically at corners of IPHAS fields).

was set at $0.5''$ after testing the typical separation between IPHAS and SDSS sources (see Figure 3.5). Figure 3.20 shows the separations for two of the considered regions - one displaying astrometric shifts at the $\approx 0.25''$ level in alternating strips of SDSS observations, and the other where the only shifts approaching this level are at the very corners of IPHAS fields. $0.5''$ was chosen as the cross-match radius to ensure that all sources in the regions suffering from astrometric errors (from either survey) were given sufficient tolerance to be matched, without being so large as to present an unduly optimistic estimate of reliability by causing incorrect matches.

An overall reliability fraction is useful when considering the DR2 catalogue as a whole, though it is necessary to perform more specific tests if the DR2 is to be used rigorously. The reliability fraction will not remain constant for all objects; it is likely that there are more spurious detections at faint magnitudes as noise is more prevalent in this regime, and as such more likely to mimic faint objects.

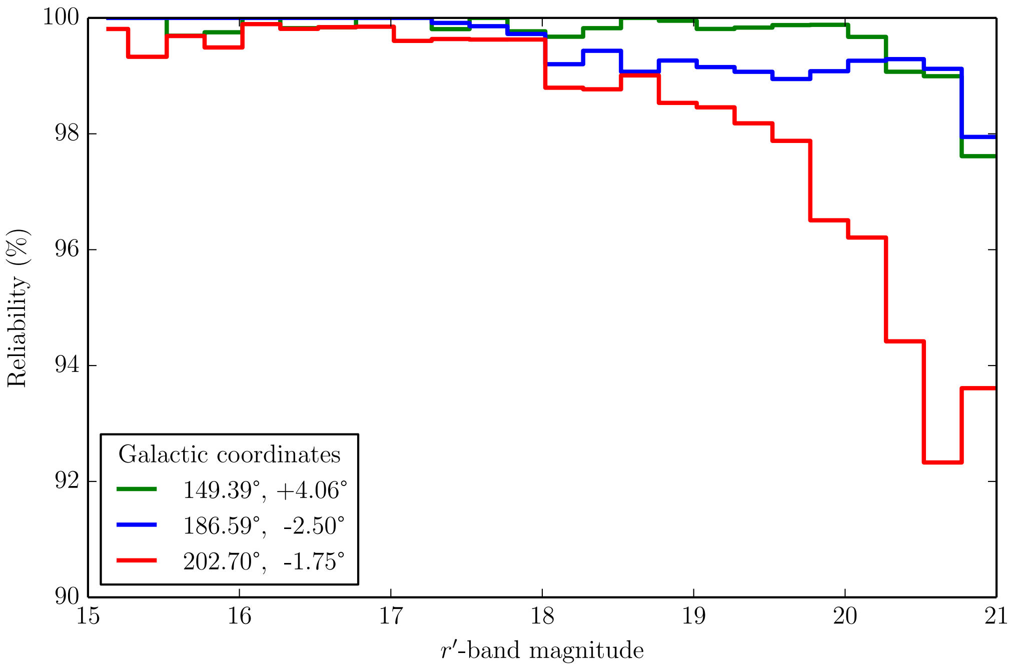


Figure 3.21: Reliability as a function of r'-band magnitude of sources for the three regions referenced in Table 3.4. All sources were considered here (i.e. the reliabilities presented here are magnitude binned versions of the 'raw' column in Table 3.4.)

3.4.2 Variation with magnitude

Binning objects according to their r'-band magnitudes when assessing reliability confirms that fainter objects are less likely to be reliable (see Figure 3.21). Down to $r' = 18$ the reliabilities for the three regions are comparable, staying above 99%. The reliability curves begin to deviate at $18 < r' < 19$, where the reliability of one region ($\ell = 202.70°$, $b = -1.75°$) declines to ∼98%. In all three regions, objects brighter than r'=19 composed less than 10% of all unreliable sources. At $r' > 19$ the lower curve deviates more significantly, dropping to ∼93% at $r' = 21$, with the other regions dipping to ∼98%. At these fainter magnitudes the majority of unreliable sources are the result of spurious detections surrounding bright stars.

At r' <19, unreliable sources were seen to be distributed uniformly across the regions under consideration. The one exception is in the $\ell = 186.59°$, $b = -2.50°$ region. Half of the unreliable sources at these magnitudes are found in one grouping, the coordinates of which match those of the open cluster Berkeley 21 (Tosi et al., 1998). This implies that an attempt to correct the number count of stars for estimated reliability would have overcorrected in the vicinity of this cluster due to IPHAS outperforming SDSS at picking out its members. As such no corrections for reliability have been made to any density maps produced; in addition such a spatially varying correction would need a reference with greater overlap than that between IPHAS and SDSS DR7 (see Figure 3.3).

Figure 3.21 supports the decision to set the default magnitude limit on any produced r'-band density map to 19th magnitude - not only due to varying field depths (see Section

2.4.4) but due to the fact that reliability is still good (>98%) at this limit.

Chapter 4

Completeness and confusion

The density map generated in Chapter 2 could be extremely useful in studies of Galactic populations. Its ability to provide star counts down to fine resolution enables immediate comparisons with Galactic stellar densities as measured by other surveys in other wavebands, and provides a tool to test Galactic population synthesis models.

The final density map, envisioned as a tool for studying the stellar population of the Milky Way, needs to be as free of observational artefacts and bias as possible to be used in this way.

IPHAS sources can be lost (i.e. fail to reach the final catalogue) due to a number of causes. For example, detector issues may eliminate sources before the raw images are even created; bad columns, hot/cold pixels and vignetting are all known problems with the WFC. In the majority of cases, any sources lost due to these effects will be picked up by the offset partner pointing or by an overlapping neighbour field. Due to this self-correcting behaviour built into IPHAS, sources lost to these detector effects do not need to be corrected for in generating the density map. It is the losses that are carried through into the final catalogue that need to be dealt with. This degree of incompleteness is important to understand for any survey - without it any estimate of stellar density will inevitably be an underestimate with unknown error.

One such cause of incompleteness, the blending of sources in dense fields, is discussed in Section 4.2. A straightforward approach to characterising completeness which has been taken by previous surveys is presented in Section 4.3, where it is applied to IPHAS and its shortcomings discussed. A more rigorous approach, involving the simulation of sources in IPHAS fields, is presented and described in depth in Section 4.4.

4.1 Correcting the density map

4.2 Measuring confusion loss - nearest neighbours

Confusion, the effect of background noise caused by unresolved sources, has been shown to scale with source density (Condon, 1974). Hogg (2001) quote a rule-of-thumb that states *"confusion becomes important at 1/30 of a source per beam"*. In the language of optical observations, a beam corresponds to the area traced out by the typical PSF of an image, taken by Hogg (2001) to be the area defined by a circle with a radius of 1σ of the Gaussian PSF (of FWHM θ_{FWHM}) of an image. Using their derived expression for the solid angle of a "beam"

$$\Omega_{\mathrm{beam}} = \frac{\pi\theta^2_{\mathrm{FWHM}}}{8\ln(2)} \tag{4.1}$$

and a seeing of $1.8''$ - slightly above average for IPHAS observations - gives an Ω_{beam} of 1.83 sq. arcsec.. Taking the highest densities seen in the maps produced in Chapter 2 (≈200 sources / sq. arcmin.), this yields ≈0.1 sources per "beam" - higher than the 1/30 figure at which confusion starts to become important.

Therefore the effect of confusion cannot be ignored. Hogg (2001) investigated the impact of confusion on determination of astrometry and association of sources at different wavelengths. For the single-band density maps generated here, these are not critical issues, however confusion in the most densely populated regions of the plane could cause source losses that would affect the final density maps. In this section the impact of confusion is assessed and potential approach to corrections discussed.

4.2.1 Expected Poissonian behaviour

The effect of source confusion (particularly a problem in the redder i'-band filter and in dense fields) can be estimated based on the fact that for stars randomly distributed on an image, the distance of their nearest neighbours can be described by a Poisson distribution. The probability of finding x neighbours within an angular distance θ of a star can be estimated to be

$$P(x) = \frac{\mu^x}{x!}\mathrm{e}^{-\mu} \tag{4.2}$$

where $\mu = \rho\pi\theta^2$ (the expected number of sources within the area $\pi\theta^2$, with ρ giving the density of stars per unit solid angle). The probability of finding a source within an angular distance θ is given by

$$\begin{aligned} P(x \geq 1) &= 1 - P(x = 0) \\ &= 1 - \mathrm{e}^{-\mu} \\ &= 1 - \mathrm{e}^{-\rho\pi\theta^2} \end{aligned} \tag{4.3}$$

and the probability of finding a nearest neighbour *at* an angular distance θ is given by

$$f(\theta) = \frac{dP(x \geq 1)}{d\theta} = 2\rho\pi\theta\mathrm{e}^{-\rho\pi\theta^2} \tag{4.4}$$

Multiplying this by the number of stars in the area of interest ($\rho \times \Omega$) yields the nearest neighbour distribution

$$n(\theta) = 2\rho^2\Omega\pi\theta\mathrm{e}^{-\rho\pi\theta^2} \tag{4.5}$$

as presented in Bahcall & Soneira (1983).

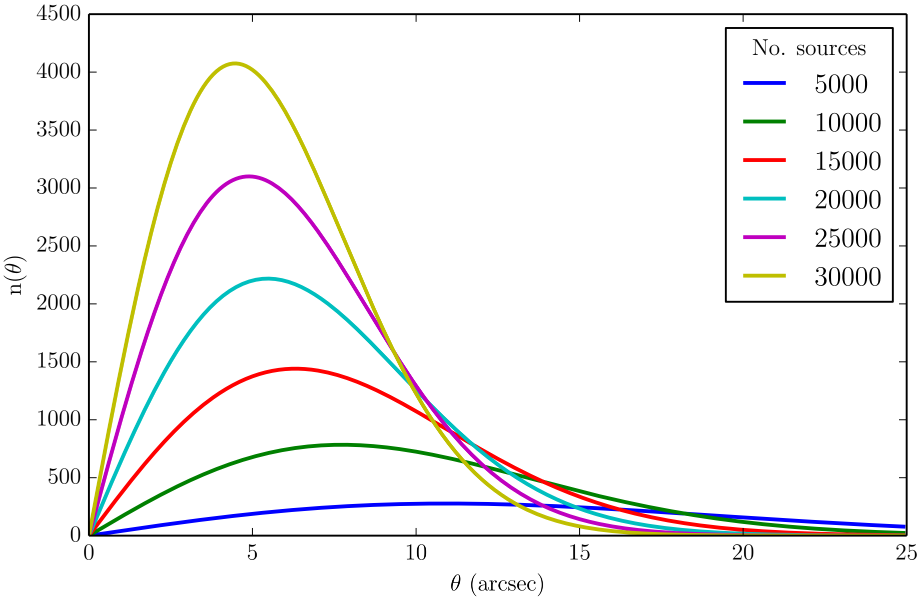


Figure 4.1: Theoretical distribution (see Equation 4.5) of nearest neighbour distances for an area with the effective area of a WFC image, for fields containing varying numbers of sources.

Figure 4.1 shows this distribution for an area with the effective area of the WFC field of view (0.29 sq. deg.) containing varying numbers of sources. It can be seen that increasing the density of sources increases the number of nearest neighbours at small separations, while suppressing the number at greater separations. This has the effect of decreasing the peak separation value, θ_{max}.

This distribution holds for a population of sources randomly distributed regardless of brightness. In practice, where the distribution of source brightnesses is described by a power law, the susceptibility of a source to confusion increases with decreasing brightness; a limiting faint magnitude must be chosen in order to define a meaningful confusion rate. For this reason, the treatment of nearest neighbour distances adopted in this discussion excludes sources fainter than 19th magnitude - i.e. ρ is in fact $\rho(\mathrm{m}_0)$, the density of sources down to this faint limit. This eliminates the complexity that would be introduced by comparing confusion of all sources across fields with different depths.

4.2.2 Applying to IPHAS fields

In order to reproduce the theoretical distribution (Equation 4.5), all sources in the area Ω would need to be recovered. This will never be the case in dense fields, which will modify the observed distribution. The loss of a source to confusion is more likely once its PSF overlaps with that of its nearest neighbour (or indeed neighbours). As such the smallest separations will be under-reported in the distributions of dense IPHAS fields, where the average inter-source distance is at its lowest - sources at small separations will not be distinguished from one another. The apparent nearest neighbour recovered will therefore lie at a greater distance,

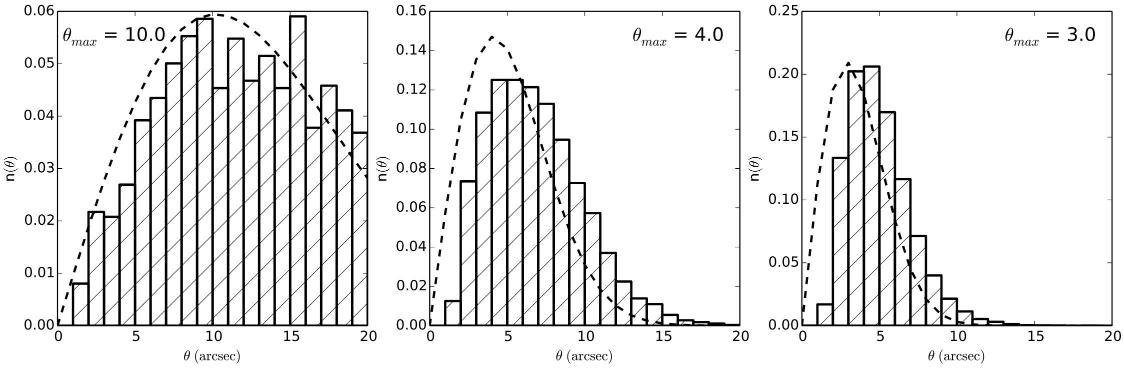


Figure 4.2: Nearest neighbour distributions for three IPHAS fields, normalised to allow the overplotting of theoretical distribution (dashed line). Density of fields increase from left to right.

boosting the numbers at larger values of θ.

Three test cases can be seen in Figure 4.2: these IPHAS fields, located around $\ell \sim 32^\circ$, were chosen as they lay in a region of the survey containing fields spanning a wide range of densities. The $n(\theta)$ distributions were generated by selecting subregions of the CCDs such that a border of width $30''$ was excluded, and the nearest neighbour was identified for each of the sources in the central subregion. This approach was necessary to avoid the case where a source close to a CCD edge has its nearest neighbour assigned to a source in the CCD area, when in fact its nearest neighbour on the sky is beyond the border of the image. When identifying the nearest neighbours for each source in the central region, matches from the border region were allowed.

The fields are displayed in order of increasing density. The left-most field, 4198_jul2004, contains 3,966 sources - a number sufficiently low that confusion would not be expected to make a significant contribution to incompleteness. Fields 4450o_jul2009 and 4285_jun2004, containing 24,291 and 49,150 sources respectively, suffer from increasing levels of confusion. The overplotted theoretical distribution gives an idea of the amount of confusion; at low field densities, the observed and theoretical distributions agree quite well, while at higher densities the theoretical distributions clearly predict closer nearest neighbours than are observed.

4.2.3 Quantifying source losses due to confusion

While the observed nearest neighbour distribution is affected significantly by confusion most significantly at smaller ($\lesssim 10''$) separations, the observed number of nearest neighbours at larger separations is less affected by confusion. The theoretical distribution which best fits the tail at large θ values should describe the field as if confusion were not an issue.

Distributions generated from Equation 4.5 were fitted to the observed distributions at $> 10''$ separation, varying ρ such that the number of sources in a field were allowed to range from half its observed value up to twice the observed value. By calculating χ^2 for each fit at large separations, the best fitting value of ρ was found for each field, identifying the correction

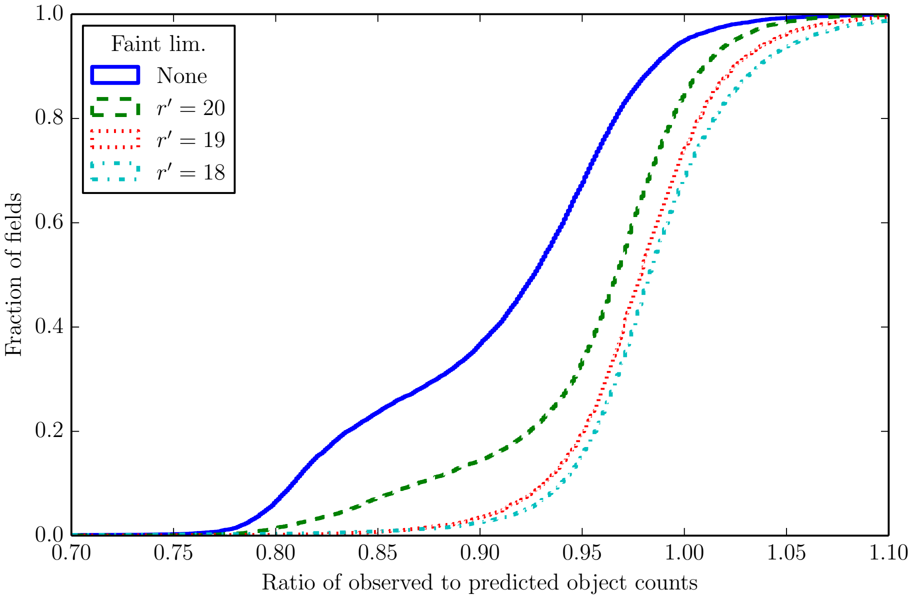


Figure 4.3: Cumulative histogram of $\frac{observed}{predicted}$ source counts, where the predicted counts were estimated by fitting Equation 4.5 to the tail of the nearest neighbour distribution. The four curves represent the ratio as determined for sources down to three different limiting magnitudes, and an instance where no limit was placed on the magnitude of sources.

that should be applied for confusion.

Figure 4.3 shows the effect of confusion on the survey as a whole. Best fitting n(θ) distribution (corresponding to a theoretical unconfused source count) was obtained for each IPHAS field, allowing a picture to be built up of the impact of confusion. The statistic chosen to represent the magnitude of the correction is the ratio of $\frac{observed}{predicted}$ sources, where the "predicted" source count is the count corresponding to the best fit theoretical n(θ) distribution. In order to understand the variation in confusion at different limiting magnitudes m_0, fits were performed on sources down to 18th, 19th and 20th magnitudes in r', in addition to fits across all sources.

In a perfect survey capable of deblending overlapping sources, all fields would exhibit zero confusion, with their curves jumping from zero to 100% of all fields at a ratio of 1.0. In reality some sources will always be lost to confusion, with the impact increasing the higher the density of objects. Fainter objects suffer more from confusion losses as their PSFs are more likely to be lost in the vicinity of brighter sources, in addition to their intrinsically higher densities (see Figure 4.5 for examples of the magnitude distributions of IPHAS fields). Figure 4.3 demonstrates this behaviour as the incompleteness takes hold for a greater number of fields as the limiting magnitude m_0 is increased.

For a fraction of fields at all values of m_0, the best fit to their nearest neighbour distributions correspond to a lower predicted number of sources than actually observed sources. This is possible since the allowed minimum value when searching for the best fit is half the observed count. The maximum ratio returned is ~1.1, for fields which can be seen to lie in

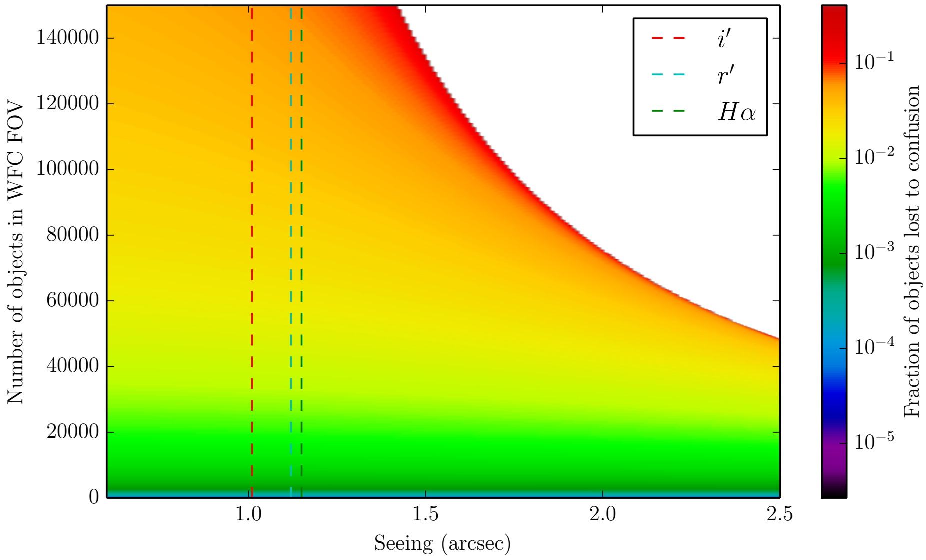


Figure 4.4: Correction term in Equation 4.6 for the range of $n_{src} = \rho' \times \Omega_{WFC}$ and θ_{FWHM}. Median seeings for each filter are marked by vertical lines. White region denotes domain where Equation 4.7 is true and Equation 4.6 breaks down.

regions of lowest density when plotted spatially. These fields would be expected to return a corresponding best fit ratio of 1.0, suggesting that the fitting to the n(θ) distribution tails has an associated uncertainty of ~10%; such an uncertainty renders this approach for correcting source counts unreliable.

4.2.4 Breakdown of theoretical approach

There exists a formalism for correcting for the number of sources lost to confusion, based upon Equation 4.5, given by

$$\rho = -\rho' \frac{\log\left(1 - 4\rho'\pi\theta^2_{FWHM}\right)}{4\pi\theta^2_{FWHM}} \tag{4.6}$$

where ρ is the actual source density, ρ' is the observed density, and θ^2_{FWHM} is the seeing in which the field was observed. The expression was derived in Irwin & Trimble (1984), and used to correct the observed n(θ) distributions in González-Solares et al. (2008). Clearly fields with poorer seeing and higher densities will have their observed densities corrected by a significant fraction; (González-Solares et al., 2008) reported that the IPHAS Initial Data Release fields suffering from greatest confusion were missing 41% of their sources.

However Equation 4.6 does not hold for for the entire [ρ', θ_{FWHM}] space covered by all IPHAS DR2 fields - it breaks down in dense fields, where the seeing is worse than the median. Figure 4.4 shows the variation of the correction term for the range of parameters encountered by IPHAS. The white region shows the parameter space in which Equation 4.6 is not applicable, where

$$4\pi\rho\theta^2 > 1 \tag{4.7}$$

Although this condition will not be met very often, the fact it is possible means that applying it across the survey is not viable.

4.3 Assessing completeness using magnitude distributions

4.3.1 The turnover approach

An approach used in several previous surveys (Ruphy et al., 1997; Cambrésy et al., 2002; Lucas et al., 2008) is to estimate the completeness limit by taking the magnitude distribution of sources and identifying at which magnitude the distribution begins to drop off.

As part of a study of stellar populations in the Galactic Plane, Ruphy et al. (1997) commented on the need for simulated images for a quantitative study of completeness in their frames, although they opted instead to report a completeness limit derived from the star count histograms of uncrowded fields. This approach may serve well in studies of limited sky regions, or if sources fainter than this putative completeness limit can be excluded without causing problems. Ruphy et al. (1997) mentioned that confusion due to crowding was the only source of incompleteness for which they account, choosing to ignore the "slight variations due to the observing conditions" (an approach unfeasible for IPHAS).

Cambrésy et al. (2002) estimate their completeness from the turnover of 2MASS magnitude distributions, and argue that their density maps would show imprints of individual observations had they overestimated their limiting magnitudes. This would certainly be the case for IPHAS - without any limiting magnitudes in place, the field-to-field variation of densities is extremely obvious.

In a similar treatment of incompleteness, Lucas et al. (2008) quote the 90% completeness limit of UKIDSS, mentioning that in uncrowded fields the modal depths vary by 0.25 mag due to observing conditions. This method still depends heavily on the turnover in the magnitude distribution, assuming that sensitivity is the only cause of turnover, and relies on "visually extrapolating" the histogram.

A single completeness statistic as provided by Lucas et al. (2008) could be used to apply a uniform completeness correction, although this assumes that all fields in a survey suffer from the same degree of incompleteness. Even if this were indeed the case, only the count of sources brighter than the given completeness limit could be corrected; it is not possible with this information to correct the count for any other magnitude range.

Figure 4.5 shows the magnitude distributions for three IPHAS fields. A visual inspection would suggest that incompleteness sets in at ∼20th mag as a crude estimate across the three bands. Attempts to automate the determination of turnover magnitude for all fields are hampered by the facts that the mode of these distributions are dependent on the chosen binning, and that the magnitude distributions of many fields plateau before dropping off.

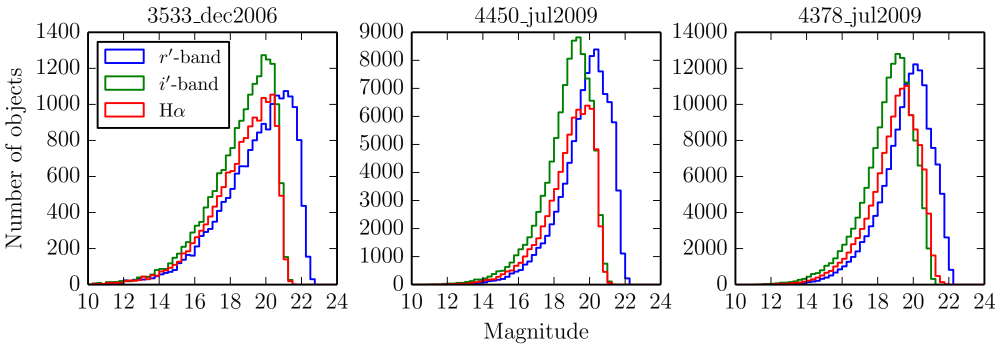


Figure 4.5: Magnitude distributions for three IPHAS fields of increasing density. From left to right, fields have r'-band source counts of 18,113, 99,061 and 153,430. A requirement that sources counted here have $nBands > 1$ places the greatest constraint on i'-band counts, as redder sources are picked up in the i'-band only.

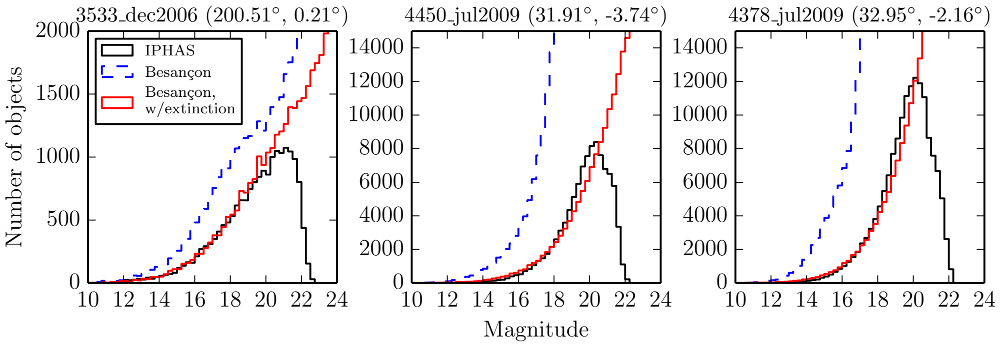


Figure 4.6: r'-band magnitude distributions for three IPHAS fields of increasing density. **Black**: IPHAS star counts. **Blue**: Besançon prediction for equivalent sightline. **Red**: Besançon prediction scaled to bright ($r' = 17$) IPHAS counts.

4.3.2 Comparisons with Besançon star counts

A possible external source of information that could be used to check completeness with magnitude distributions is the Besançon model of galactic populations (Robin et al., 2003). By querying the web interface of the Besançon group, synthetic catalogues can be generated based on supplied parameters.

Catalogues of synthetic sources along the same sightlines as IPHAS fields were generated, corresponding to the area of the WFC footprint, and retaining the default normalisation of 0.7 mag kpc^{-1} for the extinction gradient. This is unrealistic for most sightlines in the Galactic plane, where a thorough treatment would require varying the extinction to replicate the distances to spiral arms, known molecular clouds, and other known large-scale perturbations to the density of the ISM. As a first order estimate, the default normalisation was retained, and scaled to match bright (in IPHAS, $r < 17$) star counts between the Besancon prediction and IPHAS counts. As the Besancon approach models diffuse extinction with an exponential decay, the extinction can be scaled by an arbitrary factor rather than needing to recompute the result for each normalisation.

Figure 4.6 shows the r'-band distribution from IPHAS fields against the distributions expected by the Besançon model, as well as the Besançon model scaled to the bright IPHAS star counts. The agreement is not perfect, owing to the approximation of reddening as a simple exponential (in addition to any potential shortcomings of the Besançon model). It can be seen that for the three sample fields, the IPHAS and scaled Besançon counts appear to diverge at around $r' = 19 - 20$. This is in agreement with the visual inspection of the r'-band magnitude distribution turnovers.

This approach is still rather crude; an exact value at which the two distributions diverge would be needed in order to derive a completeness factor for each field. In cases where the properties of the stellar population varies over length scales smaller than those of an IPHAS field, it would be necessary to reduce the area considered for each correction, perhaps to a CCD-by-CCD basis. Since one of the immediate uses of a corrected density map would be to test the utility of the Besançon model at low Galactic latitudes, using its predictions of Galactic populations to correct the map would be self-defeating. Given these various issues and constraints on characterising the incompleteness of the survey, a fuller treatment is desirable.

4.4 Generating artificial sources

A thorough treatment of incompleteness in any survey requires measuring its sensitivity to sources over the entire magnitude range of interest - this can be achieved by simulating observations and then processing these simulated frames in the same way as the original data. This method was used (in addition to the magnitude distribution method discussed in Section 4.3.1) by Harvey et al. (2006) as part of a study of interstellar clouds observed by Spitzer.

Such an approach is much more powerful, as statistics can be returned on magnitude bins, rather than on the entire magnitude distribution down to a specified faint limit. This requires the simulation of sources of all magnitudes, resembling real data as closely as possible, without being too costly to generate in either computing power or time.

For the purposes of correcting the density map, the completeness of each field was assessed using the images and catalogues of CCD 4 only. Using only one of the WFC CCDs per field cuts the processing time by a factor of four, bringing the total time necessary to compute the completeness of the entire survey down to a little over one week. Due to the fact that DR2 preferentially selects sources from the centre of fields (closer to the optical axis), CCD 4 was selected as the chip to represent each field (see Section 4.5.5 for further discussion).

This approach to completeness calculation was divided into steps. First the typical properties of an image were determined, and used to generate synthetic sources that closely resemble genuine stellar sources (Section 4.4.1). The parameters of these synthetic sources were recorded and then catalogues generated from the new frames containing synthetic objects. The tables of synthetic source parameters were cross-matched against the newly generated catalogues, and the rate of recovery measured (Section 4.5). The fraction recovered at differ-

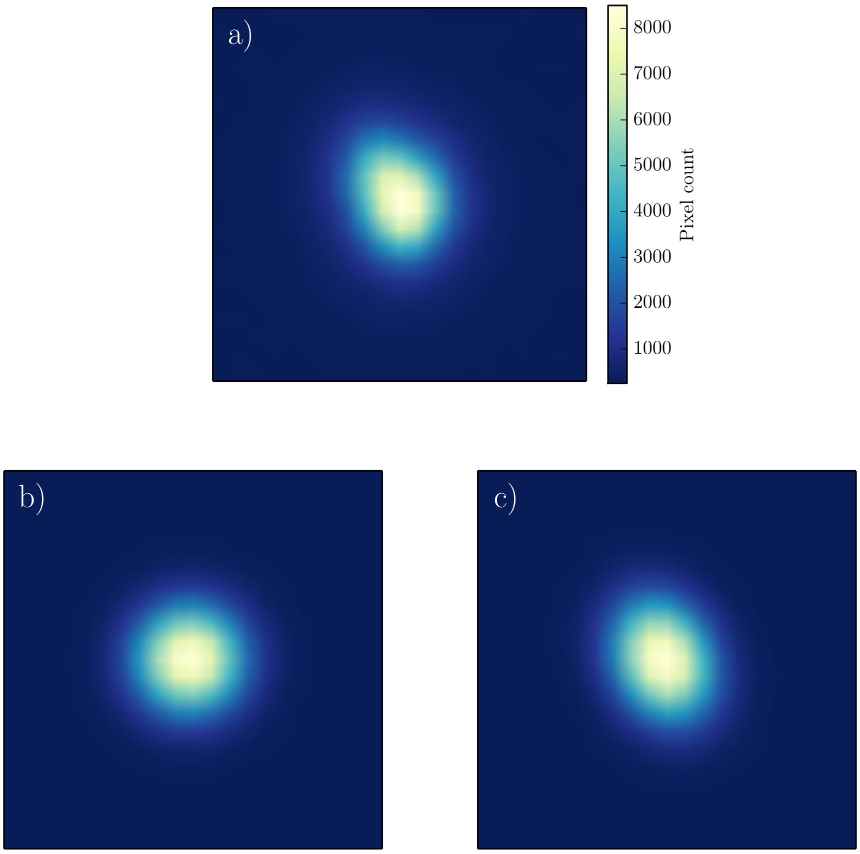


Figure 4.7: a) 16th magnitude source from an IPHAS field. **b)** Simulated source generated using best-fit parameters of a 2D circular Gaussian. **c)** source generated using best-fit parameters of a 2D elliptical Gaussian.

ent magnitudes will allow a completeness curve to be built up for each field, and uniformly computed corrections to be applied across the entire survey.

4.4.1 Simulating stellar sources

The complexity of simulating stellar sources depends on the accuracy with which the real sources are to be recreated. A first order approximation might be a circular PSF characterised by a two-dimensional Gaussian, requiring only three parameters - FWHM, peak height and centroid coordinates. A cursory glance at the reduced IPHAS frames reveals that a perfectly circular PSF will not reproduce realistic stellar sources - a more complex approximation is needed (see Figure 4.7).

The array of parameters returned from the aperture photometry performed by IMCORE suggest what might have to be included to accurately recreate a stellar source:

- Peak flux height
- Total measured flux

- FWHM
- Ellipticity
- Position angle
- Sky background level
- Coordinates (or pixel position on detector)

As the IMCORE routine is quite complex, it was unclear that inserting PSFs parameterised solely by its estimates of these parameters would return the expected values once the modified image was run through IMCORE. While it was possible to simply take the output catalogue from the original image, it was deemed necessary to ensure that estimates of PSF parameters and the PSFs inserted into images were estimated and generated respectively using similar code that was fully understood.

Following an examination of the number and magnitude distribution of sources to be inserted (Section 4.4.2), each of these parameters is discussed, clarifying how measurements from the original image and catalogue were used to generate realistic artificial PSFs. Section 4.4.9 provides a summary on how all these are combined to generate artificial sources (illustrated in Figure 4.14).

4.4.2 Synthetic source counts and magnitude distribution

Ideally measuring the completeness using artificial sources would be done by adding a single source at a time, verifying whether or not it is detected. This process would be repeated as many times as required to obtain decent statistics, for the entire magnitude range of interest. This would be the route taken in the absence of any limitations on either time or computing resources.

The limiting factor in the completeness investigation is the time taken by the IMCORE routine, which takes ~20 seconds to read in an IPHAS image, run its background estimation routines and generate a catalogue. Adding a single source at a time would require ~20 s, multiplied by both the number of magnitude bins that are being considered, and the number of sources needed in each bin for sufficient statistics. This would push the time required to generate the number of sources detailed in Table 4.1 to ~ 65 hours for a single field.

To reduce the cost, multiple sources needed to be inserted into each image simultaneously. It was necessary to ensure that not too many sources were inserted at any one time; inserting too high a number would modify the intrinsic properties of the image - for example, probing the completeness of an originally sparse image with a high number of artificial sources inserted would not return statistics useful for understanding the original image. The value of $\frac{\delta n}{n_{image}}$ needed to be kept sufficiently low, where δn is the number of artificial sources added to the image, and n_{image} is the number of sources present in the original image. The quality control information available for DR2 fields report that the most sparsely populated fields contain more than 1000 stars. This is not an extremely constraining limit; a maximum of up to 50 stars was chosen as a value that would keep $\frac{\delta n}{n_{image}}$ below 0.05 for all fields.

Magnitude bin				
Start	End	N	M	*No. sources*
12.0	12.25	10	10	100
12.25	12.5	10	10	100
⋮	⋮	10	10	100
14.5	14.75	10	10	100
14.75	15.0	10	10	100
15.0	15.25	20	10	200
⋮	⋮	20	10	200
16.25	16.5	20	10	200
16.5	16.75	30	10	300
⋮	⋮	30	10	300
17.75	18.0	30	10	300
18.0	18.25	40	10	400
⋮	⋮	40	10	400
18.75	19.0	40	10	400
19.0	19.25	50	10	500
19.25	19.5	50	10	500
19.5	19.75	50	15	750
19.75	20.0	50	15	750
20.0	20.25	50	20	1000
⋮	⋮	50	20	1000
20.75	21.0	50	20	1000
		Total:	410	12300

Table 4.1: Number of artificial sources added per magnitude bin. *No. sources* denotes total number that will be generated over M runs. N is the number of sources that will be added to each image, which will be repeated M times. A total of 12300 are added per field, across 410 images containing artificial sources.

Allowing up to 50 sources per artificial source injection increases efficiency by reducing the number of IMCORE runs needed. The low value of $\frac{\delta n}{n_{image}}$ renders unworkable an approach of inserting sources based on the pre-existing magnitude distribution of the sources in the image - the number of runs required to inject artificial sources in this fashion would be significantly higher then the approach actually adopted; the requirement that $\frac{\delta n}{n_{image}} < 0.05$ would mean a very small number of bright (~ 12 mag) sources being injected in each run, requiring a large number of runs to obtain usable statistics at brighter magnitudes. The approach adopted was to split the magnitude range of interest into magnitude bins 0.25 mag wide, inserting sources from only one bin at a time, a method previously used by Harvey et al. (2006).

Table 4.1 details the number of sources inserted per image, and the number of artificial images that were generated for each magnitude bin. The total numbers per magnitude bin were increased as sources become progressively fainter. This was achieved by increasing the number of artificial sources injected per run until the limit of 50 sources was reached. At this point the number of runs was increased to raise the total number of sources injected in the faintest bins, minimising the noise at the faint end of the completeness curve, where the most significant corrections will be necessary.

An alternative program available as part of CASUTools is IMCORE_LIST, a version of

IMCORE which generates list driven photometry based on supplied coordinates. It is a significantly shorter process, since the locations of apertures at which fluxes are measured are fixed by the supplied list. However the artificial sources need to be given the same treatment as genuine sources during the production of the DR2 catalogues - unless this approach is taken, no artificial sources will be lost, as fluxes will be returned at all supplied coordinates. The originally used IMCORE needs to be applied so that artificial sources that are lost (e.g. due to blending) can be missed entirely by IMCORE. It is only by applying the same approach consistently that useful completeness statistics can be returned.

The setup detailed in Table 4.1 requires around 2.2 hours (∼20 s × 410 IMCORE runs). For the 14,115 fields that make up DR2, this results in a total computing time of ∼31000 hours. Using the computer cluster available within the Science & Technology Research Institute at the University of Hertfordshire, a large number of CPUs are available to run such jobs. Utilising 128 CPUs simultaneously brings the total time to estimate completeness for the entire survey in a single band to ∼10 days.

4.4.3 Sky background level

In order to determine if the sky background level and its variation across an image led to a noticeable uncertainty in parameterising sources, the PSF fitting algorithm was tested on both sky-subtracted and unsubtracted pixel counts. In order to obtain an estimate of the background level across the reduced images an approach was adopted based on that of CASU in its IMCORE background estimation routines.

First a CCD is split into a grid of squares of arbitrary scale, and the median value of all non-saturated pixels is estimated for each sub-frame after applying iterative upper sigma clipping . The resulting grid of median values is first passed through a median filter to remove unrealistically high values, and then smoothed using a Hanning window. The sky pixel values for the entire CCD are then bilinearly interpolated from the sub-frame sky medians. An example of such interpolated sky backgrounds is shown in Figure 4.8.

Varying the size of the grid squares used to split up a CCD showed that 64×64 pixels (the CASU default) is indeed the optimum choice. Figure 4.9 shows the effect of setting the grid scale to 32, 64 and 128 pixels. At 32 pixels the scale is small enough that the interpolated backgrounds are affected by bright stars, although nebulosity is well traced. At 128 pixels nebulosity at small scales is mostly smoothed away, and the effects of bright stars are mostly lost. A reasonable compromise is achieved at 64×64 pixels.

Before refining the background estimation further, tests were conducted to determine the overall effect of background subtraction on the fitting of two-dimensional Gaussians to stellar sources. Positions of the sources were obtained from catalogues generated by running IMCORE on the image. Using the SCIPY module in Python, elliptical two-dimensional Gaussians were fitted, allowing the peak value, orientation and σ along both semi-major and semi-minor axes to vary.

It was found that for the magnitude range of interest ($r' < 21$) that fitting a Gaus-

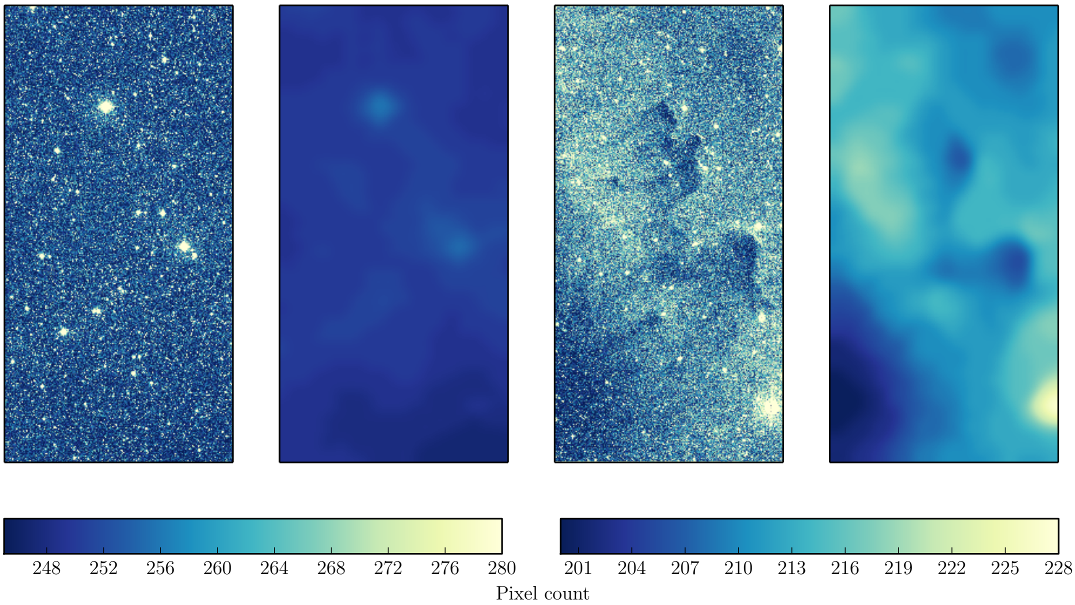


Figure 4.8: Examples of interpolated sky backgrounds. **Left:** Reduced image of CCD 1 for IPHAS field 4121o_jul2009 alongside interpolated background generated with grid scale of 64 pixels. **Right:** Same for CCD 2 of IPHAS field 6662o_aug2004a.

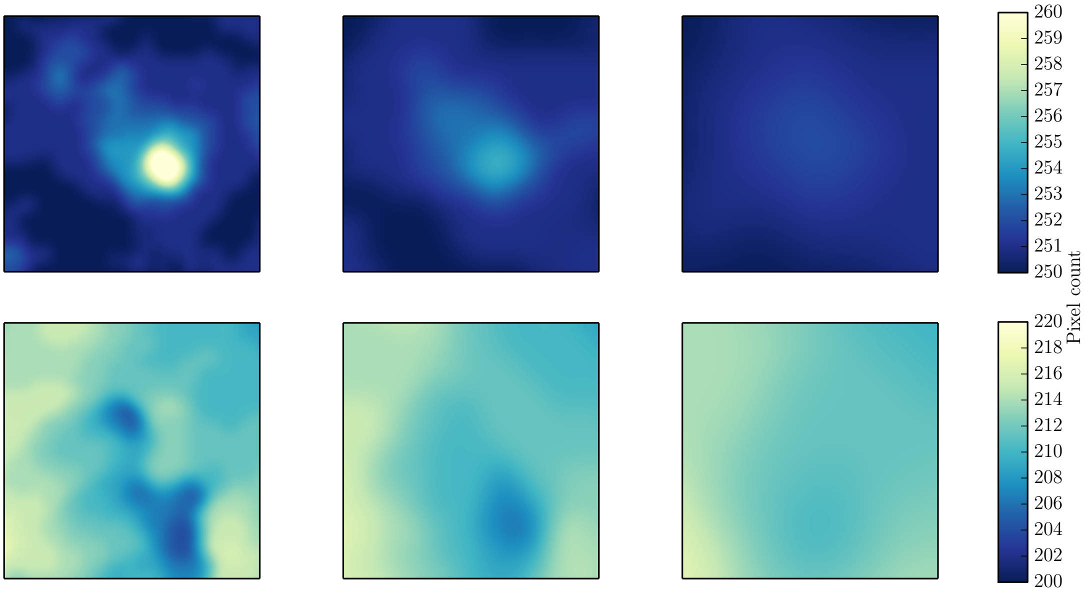


Figure 4.9: Demonstration of the effects of grid cell size on background estimation for fields 4121o_jul2009 (**upper**) and 6662o_aug2004a (**lower**) (close-ups of the fields shown in Figure 4.8). A region of size 1000×1000 pixels is shown. From left to right: grid scale sizes of 32×32, 64×64 and 128×128 pixels.

sian to the original data with a floor value that is allowed to vary, and fitting to the sky-subtracted data with no underlying background, produced similar best-fit values for the peak and FWHM. These tests were conducted on well-behaved fields (i.e. with no nebulosity or notable artefacts), for which it was concluded that sky-subtraction was unnecessary. For the cases with high nebulosity, background levels can change so rapidly that this background estimation method would prove insufficient. For these reasons, sky subtraction of frames was avoided, with the estimation of optimal parameters for artificial sources proceeding on the reduced frames as provided by CASU.

4.4.4 Full-width half-maximum

Initially the parameters of genuine sources in IPHAS frames were determined by fitting elliptical two-dimensional Gaussian profiles to each source detected by IMCORE, building up lists of best fit parameters for the purpose of generating realistic artificial sources.

Parameters of sources were excluded from the lists if they exhibited one of the following behaviours:

- Source position returned further than 5 pixels from the position reported by IMCORE
  - these fits are likely to have been disturbed by nearby sources
- Object where the peak value of the best-fit Gaussian is > 55000 counts
  - this is the regime for the WFC where bright sources saturate, distorting the PSF.
- If the position angle is measured to be exactly zero
  - in these cases the fit is assumed to have failed.

The best fit FWHM values returned for the semi-major and semi-minor axes of the sources were gathered, and Gaussians fit to their distributions. Figure 4.10 shows the distributions for four IPHAS fields, in addition to the central positions of the best fit Gaussians.

This would work across all fields, but the fitting of 2D Gaussians to each source in the frame took too long. Since the FWHM values obtained in this way were in good agreement with the median FWHM as determined by IMCORE, it was decided that taking this value as representative of the entire field was sufficient. Figure 4.10 shows how the median FWHM values fall within 1σ of the best fit Gaussian distributions for four randomly selected IPHAS DR2 fields.

4.4.5 Ellipticity

The treatment of ellipticity followed that of FWHM values; although a reliable measure of field-wide ellipticity was determined from the median of the best-fit FWHM values across the field and the relation

$$e = 1 - \frac{\mathrm{FWHM}_{min}}{\mathrm{FWHM}_{maj}} \tag{4.8}$$

(where FWHM_{min} and FWHM_{maj} are the full-width half-maxima along sources' semi-minor and semi-major axes respectively).

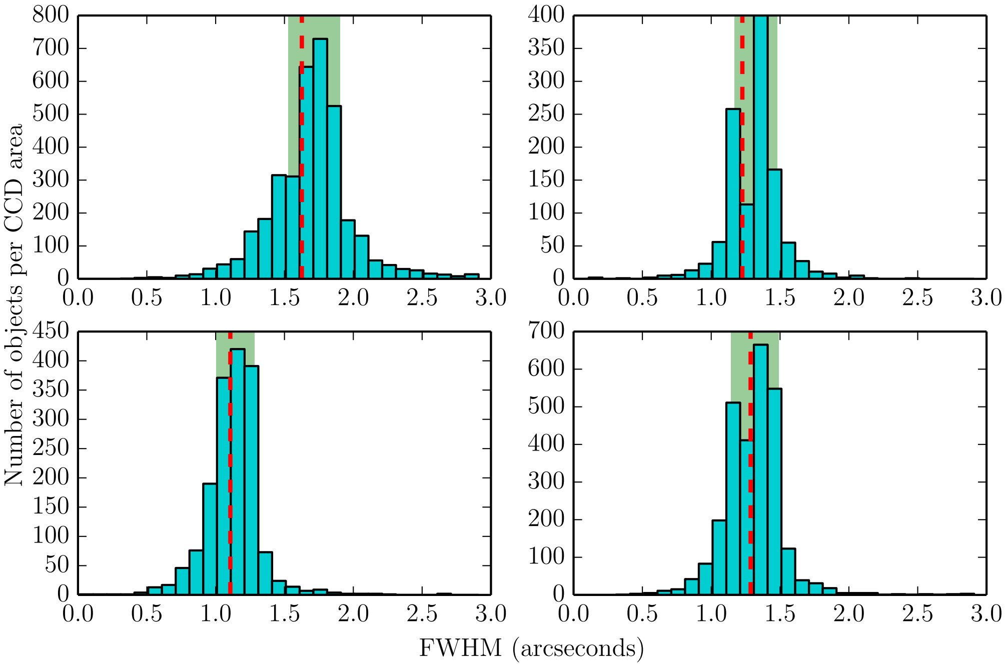


Figure 4.10: Distribution of best fitting FWHM values in arcseconds for sources from four randomly selected IPHAS DR2 fields. Values were obtained by fitting elliptical Gaussians to every stellar source in CCD4 of the IPHAS field. FWHM values reported here are for the Gaussian profile along the semi-major axes of the sources. **Red lines** denote the median FWHM as determined by IMCORE for CCD4. **Green shaded areas** highlight the regions centred on the best fit FWHM for the distributions, and encompass the area $\pm\sigma$.

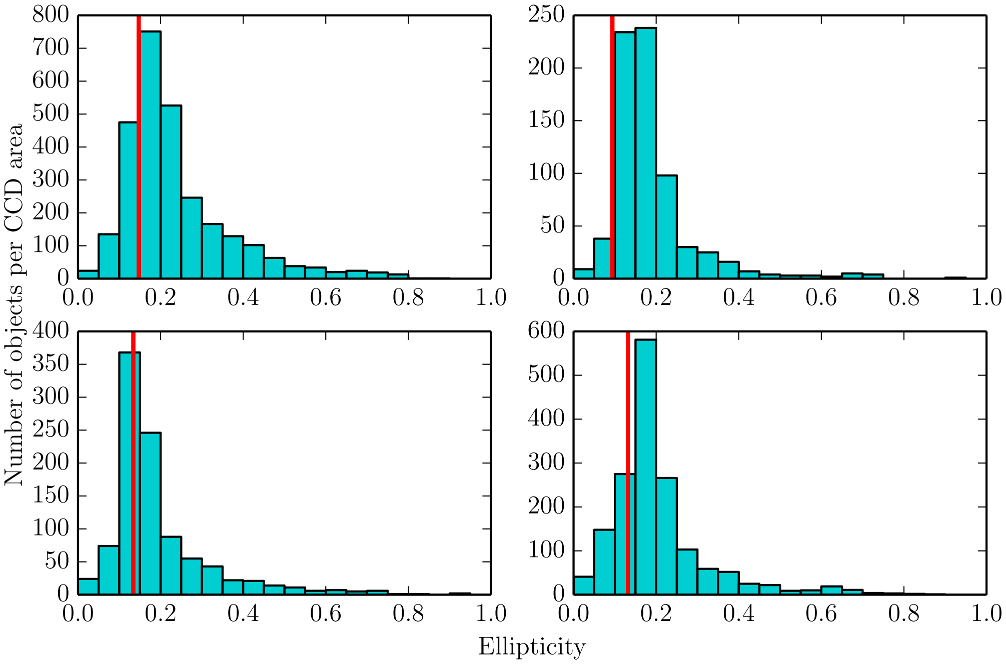


Figure 4.11: Distribution of ellipticities for all stellar sources in four randomly selected IPHAS fields. Ellipticities were computed from the profiles of the elliptical Gaussians best fitting each object along their semi-major and -minor axes. **Red lines** denote the median ellipticities for CCD4 of each field as determined by IMCORE.

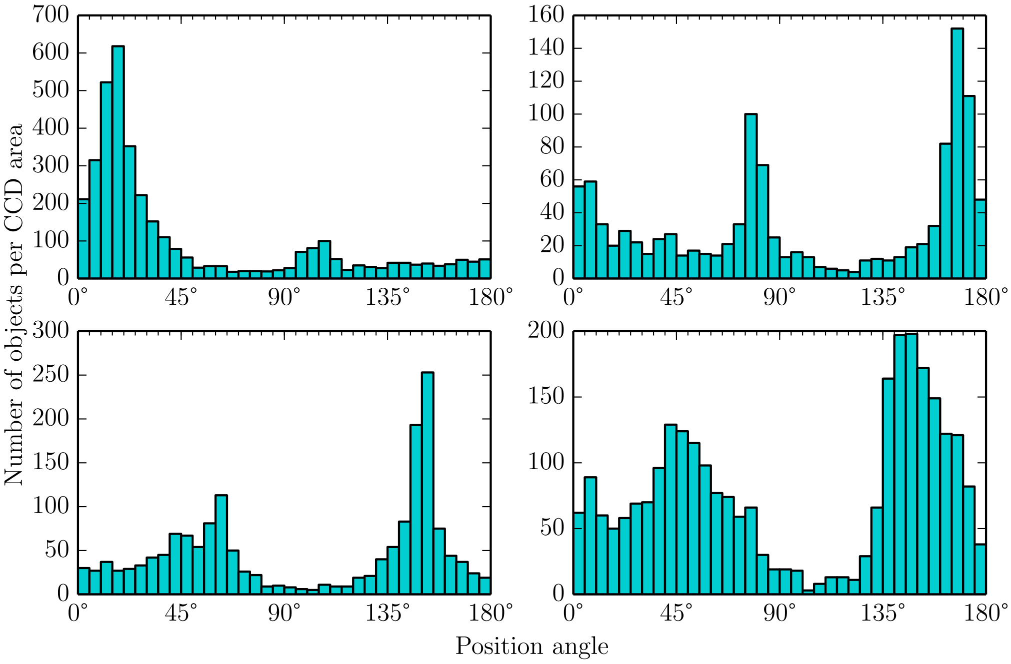


Figure 4.12: Distribution of position angles for all stellar sources in four randomly selected IPHAS DR2 fields. Position angles were obtained by including a rotation term into the elliptical Gaussian profile fit to each source in the field.

Figure 4.11 shows the distributions of ellipticities as determined by Equation 4.8 and the FWHM values determined in Section 4.4.4. The median ellipticity value as determined by IMCORE is denoted by a red line. It can be seen that the median values for three of the four test fields are slightly optimistic (i.e. underestimate the ellipticities measured from 2D Gaussian fits). Owing to the fact fields displaying ellipticities greater than 0.3 were excluded from release in DR2, the slight mismatch between the median of the fitted ellipticities and the comparatively optimistic median calculated by IMCORE will never be so large that it will significantly affect the recovery of sources.

4.4.6 Position angle

Figure 4.12 shows the distribution of position angles for sources in four IPHAS fields. It can be seen that the optimum position angle is not as easily identifiable as the previously determined FWHM and ellipticity values. By running several tests it was found that varying the position angle of artificial sources did not significantly impact its measured flux. As such the determination of the optimal position angle was not of equal priority to the FWHM and ellipticities.

The median parameters estimated by IMCORE do not include position angle, as such if a value was to be obtained it would have to be based on the distribution shown in Figure 4.12. Rather than attempt a complex fitting procedure which would slow down the completeness estimation, a basic Gaussian was fit to the distribution of each field. In many cases no good fit could be obtained, in which case no rotation was assigned to the artificial objects (i.e. the semi-major axis of the source would lie parallel to the y-axis of the CCD. In the few cases that a good Gaussian fit could be made to the position angle distribution (e.g. the top left of Figure 4.12), the central position of the peak was assigned as the artificial objects' rotation.

4.4.7 Position

Object x and y pixel positions were drawn randomly from uniform distributions across the 2048×4096 pixels of the WFC CCDs.

As the completeness calculated for CCD 4 of each field was taken as representative of the completeness over all four chips, no adjustment needed to be made for the placement of sources, as the entirety of this detector is used in generating the density map. However should other CCDs have their completenesses estimated, care should be taken not to place sources in regions that are excluded from contributing to the density map for reasons of source loss (e.g. the vignetted corner of CCD 3) - otherwise the correction for completeness obtained will over-correct the density for these CCDs. See Section 5.1.1 for a discussion of usable CCD area.

4.4.8 Flux

As discussed in Section 4.4.2, pre-defined numbers of sources are inserted in each magnitude bin. In order to insert a source of a given magnitude m, it is necessary to understand what flux that source would have to recover a measurement of m. The magnitude of an object is determined by

$$m = ZP - 2.5 \times \log_{10} \frac{c}{t} - APCOR - PERCORR \tag{4.9}$$

where ZP is the photometric zeropoint of the image, c is the measured counts within the defined aperture, t is exposure time in seconds, $APCOR$ and $PERCORR$ are small correction terms.

The ZP of a given field is calculated based on an average of photometric standards observed on the night of observation. $APCOR$ is an aperture correction term calculated by IMCORE, which uses the curve-of-growth of stellar sources to determine the correction required to transform the chosen aperture measurement to total flux. $PERCORR$ is a sky calibration correction, obtained by comparing dark sky regions with the median across each CCD. ZP, $APCOR$ and $PERCORR$ are all provided in the catalogue headers.

As shown in Figure 4.7, an elliptical 2D Gaussian approximates observed stellar sources quite well. Inserting sources with this profile into an image requires five parameters (discussed in Section 4.4.4 to 4.4.7) and a relation to determine the correspondence between peak flux height and source magnitude. This relation will vary from observation to observation; an example of a power law fit to peak flux against magnitude is shown in Figure 4.13. The fit was carried out in the range $14 < r' < 19$. It can be seen that at $r' > 19$ that the data deviate from the fit, and significant scatter around the fit sets in at $r' < 16$.

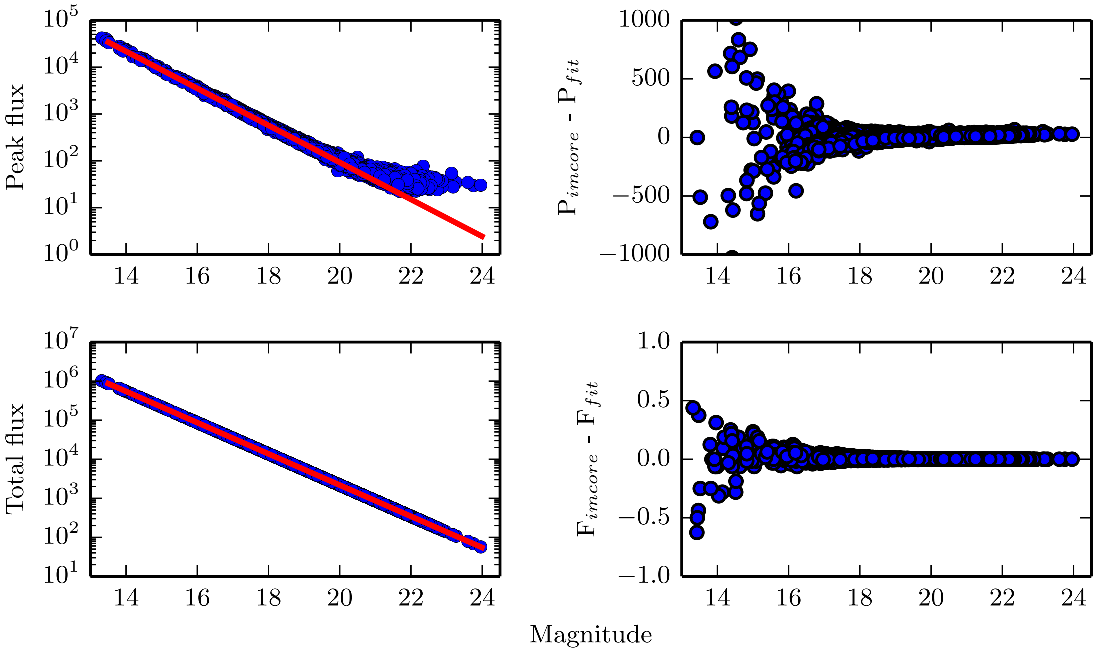


Figure 4.13: Upper left: Relation between peak flux height and reported magnitude of sources in IPHAS field 4121o_jul2009. **Upper right:** Difference between peak height measured by IMCORE (P_{imcore}) and the peak height predicted for a given magnitude from a power law fit (P_{fit}). **Lower:** The same as above only for total flux enclosed in the aperture as measured by IMCORE. Note the difference in y axis scales in the plots on the right.

Figure 4.13 also shows that the relation between total flux and magnitude is well behaved over a much larger range than the peak height relation. This is unsurprising, as IMCORE uses the total flux inside the aperture to calculate the magnitude according to Equation 4.9; the correspondence between total enclosed flux and magnitude is a direct consequence of the aperture photometry employed. Nevertheless investigating the results represented by Figure 4.13 was useful in determining how much use the peak flux would serve in generating realistic objects of a given magnitude.

Initially the Gaussian profiles used to simulate sources were informed by the peak height vs. magnitude relation. The generation was modified to use the total flux vs magnitude relation after a number of trial runs (see Section 4.5.2). The synthetic sources were generated and then had their counts scaled to match the total expected flux for its intended magnitude.

4.4.9 Summary

With a defined path to establish the parameters of artificial sources (illustrated in Figure 4.14), all IPHAS fields can be processed, each having the distribution of objects laid out in Table 4.1 inserted.

As objects were inserted, the quantities that have an element of randomness to their generation (x,y pixel coordinates and magnitude drawn from a given bin) were stored in a table, with one table stored for each run (the M quantity in Table 4.1). These records were

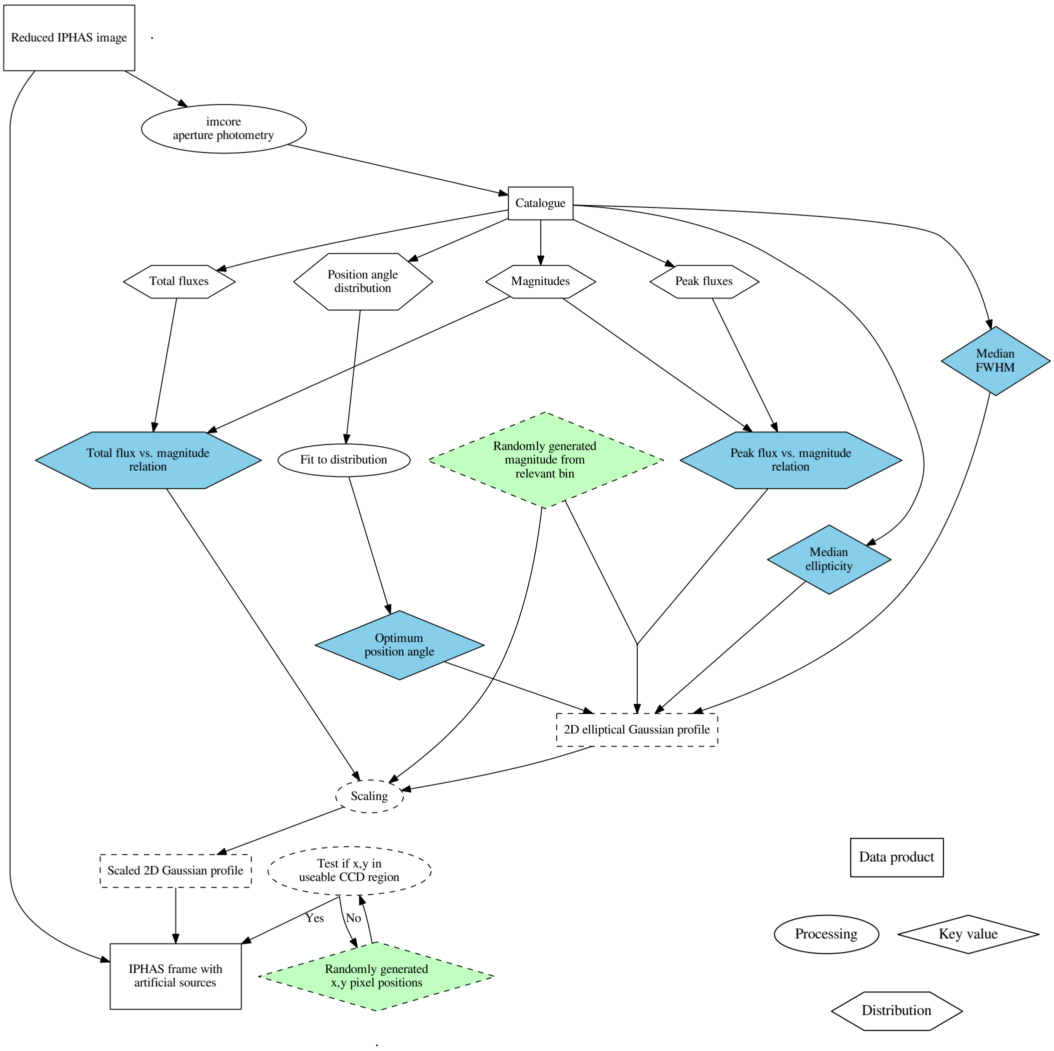


Figure 4.14: Steps taken to insert synthetic photometry into IPHAS frames. **Dashed** polygons denote steps that will differ for each artificial source inserted into an image. **Green** polygons denote quantities that are drawn from relevant uniform distributions for each artificial source. **Blue** polygons are values/relations that are determined per IPHAS field, based on its properties as measured during the aperture photometry step. The only external inputs are the original reduced IPHAS frame and the values assigned per artificial source.

required for the next stage of completeness assessment, where the fraction of recoverable artificial sources was determined.

4.5 Recovery of artificial sources

4.5.1 Processing the modified images

In order to obtain completeness fractions for each field, the fraction of sources that could be identified needed to be determined. Frames containing artificial sources added to the original data need to be processed in a manner that as closely resembles the original processing pipeline as possible. This meant running IMCORE on the modified frames with the same settings as were used to generate the original catalogues.

Cross-matching the recovered sources to the artificial source records (see Section 4.4.9) required the imposing of matching thresholds in both position and brightness. While a recovered source should lie centred on the exact x,y pixel position where it was added, there is no guarantee that IMCORE will report its centroid to lie at these coordinates. Background variations in the original image and pixel binning may cause the reported centroid to shift by a small amount; a generous threshold was adopted of 5 pixels from the recorded position of inserted sources.

Simply cross-referencing the recorded position of an artificial source with the new set of catalogues was not sufficient to ensure that a source had been successfully recovered. Including the magnitude of the source inserted in the comparison added an additional constraint which helped to reduce spurious crossmatches between the artificial source and pre-existing nearby sources of different magnitudes.

Choosing a threshold on the offset in magnitude space (Δ_m) was more involved than in pixel position. Previous attempts to use artificial photometry in this way have faced similar issues; Mateo & Hodge (1986), as part of a study of a globular cluster in the LMC, added artificial stars to their images using DAOPHOT. They considered a source recovered if a source was returned at the same coordinates within 0.5 mag. Harvey et al. (2006) used this approach in a study of stellar sources in the Serpens molecular cloud, and plotted the average difference ($| \, mag_{obs} - mag_{injected} \, |$) for artificial sources injected into their data, finding a range from 0.1 mag at ~10-13 mag to more than 0.5 mag at ~15 mag.

In order to decide on the best Δ_m tolerance to adopt, a list of detected sources within 5 pixels of the expected coordinates of each artificial source was generated, and the source with the smallest Δ_m was chosen as the best match available. This best Δ_m was recorded for each artificial source, allowing the Δ_m tolerance to be tweaked later by selecting on this information.

Figure 4.15 shows the effect of decreasing the acceptable Δ_m tolerance on the returned r'-band completeness fraction for IPHAS field 4121o_jul2009, across the range $17 < r' < 21$. The fraction reported by the blue distribution shows the completeness before scaling artificial PSFs to the total flux vs magnitude distribution (see Section 4.4.8). At a tolerance range

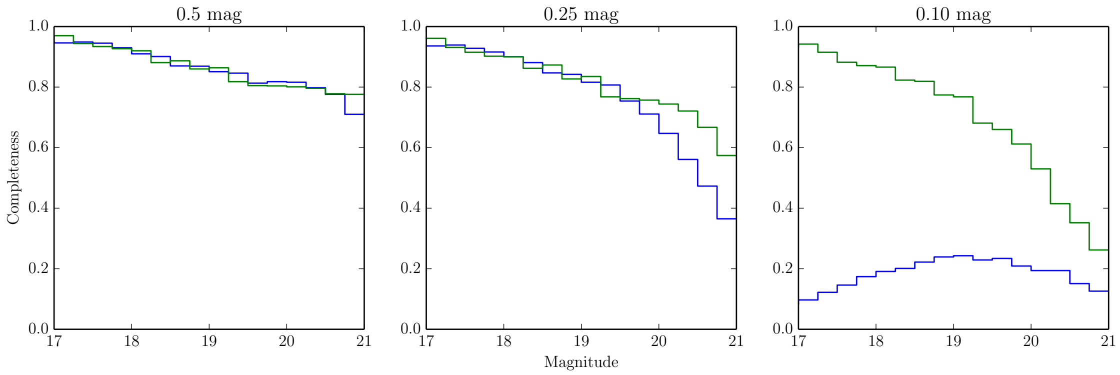


Figure 4.15: r'-band completeness fractions (**green**) for IPHAS field 4121o_jul2009, with offset tolerances between inserted and recovered magnitudes of 0.5, 0.25, and 0.1 mag (from left to right). **Blue:** Completeness fraction before implementing the scaling to total flux vs. magnitude relation (see Section 4.4.8) and the shift to account for the effect of seeing (see Section 4.5.2). **Green:** Completeness fraction after applying these corrections.

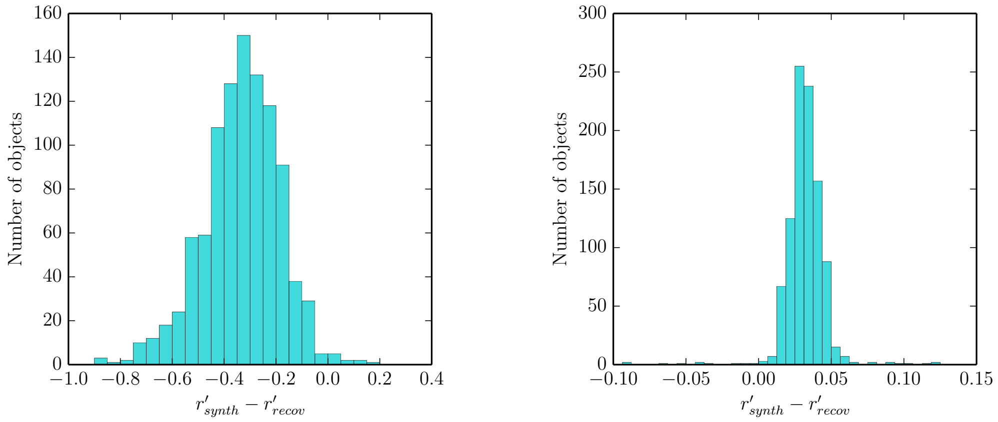


Figure 4.16: Histograms of the difference between synthesised and recovered r'-band magnitudes of 1000 sources inserted into IPHAS field 3624_dec2003 in 20 separate batches of 50 stars. All sources synthesised were drawn from the magnitude range $17 < r' < 17.5$. **Left:** Based on sources synthesised using only the peak flux vs. magnitude relation to inform PSF shape. **Right:** Scaling the PSF by a correcting factor based on the total flux vs magnitude relation.

of 0.5 mag, the completeness returned decreases slowly across the magnitude range under consideration. At a tolerance of 0.25 mag the completeness falls off faster at $r' \approx 18$, dropping below 50% at $r' \approx 20.5$. At a tolerance of 0.1 mag, the completeness fraction sits below 50% for the whole magnitude range, increasing in the range $17 < r' < 19$ and falling thereafter. While a decreasing completeness fraction was expected with a stricter Δ_m tolerance, the basic change of shape was not; this effect warranted further investigation.

4.5.2 Corrections to source generation and recovery

In order to investigate the prevalence of the unusual distribution of completeness fractions observed at smaller Δ_m tolerances, the completeness fractions of a number of fields were calculated. A random selection of 300 fields were given the same treatment, and the distribution

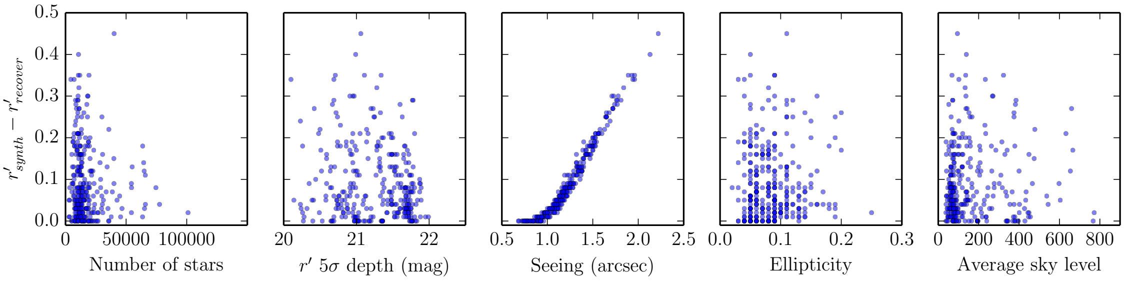


Figure 4.17: Modal difference between artificial inserted and recovered magnitude difference per field, plotted against field parameters.

of Δ_m values considered. These distributions, an example of which is shown in Figure 4.16, seemed consistent with Gaussian distributions.

By inspecting the distributions of Δ_m values it was clear that two issues were affecting the determination of completeness - the spread of the offset, in addition to the peak position, which would ideally sit at zero. The spread visible in the left hand panel in Figure 4.16 would suggest that a Δ_m tolerance of $\approx$0.5 mag would be the minimum needed to recover the majority of artificial sources. These Δ_m values were obtained while the peak flux/magnitude relation was still being used to generate the PSF of artificial sources. The upper panel of Figure 4.13 goes some way to explaining the behaviour seen in Figure 4.15; the large scatter on the peak flux of bright objects will cause large Δ_m values. The decrease in scatter between $14 < r' < 19$ leads to a decrease in Δ_m, causing the increase in reported completeness fraction when a Δ_m tolerance of 0.1 mag is adopted. The deviation from the peak flux/magnitude relation at $r' > 20$ suggests that completeness will likely be overestimated at these faint magnitudes.

It was at this point that it was decided to scale the PSF of artificial sources according to its total flux, using the total flux vs. magnitude relation for the field into which it was inserted (see Section 4.4.8). The right-hand panel of Figure 4.16 shows the effect that this scaling had in reducing Δ_m. The spread has reduced significantly, suggesting that for this field, a Δ_m tolerance of $\approx$0.1 mag would suffice to recover all detected artificial sources. However the shift of the distribution from zero, while greatly reduced, is still present. Such a small shift could be ignored, though repeating these tests on the other randomly selected 300 fields showed that this shift was not always negligible.

For each field, the distribution of Δ_m values was binned (as in Figure 4.16), and the modal value determined. For a number of fields, the modal value reached $\approx$0.4 mag - an effect that would require a Δ_m tolerance of >0.5 mag. Such a large tolerance would likely result in many spurious cross-matches and hence an unduly optimistic estimate of completeness fractions, especially at fainter magnitudes.

In the course of investigating whether field conditions correlated with the offset of the Δ_m distribution, the modal shift value for each field was plotted against a number of field parameters. The resulting plots, seen in Figure 4.17, clearly show that the remaining offset

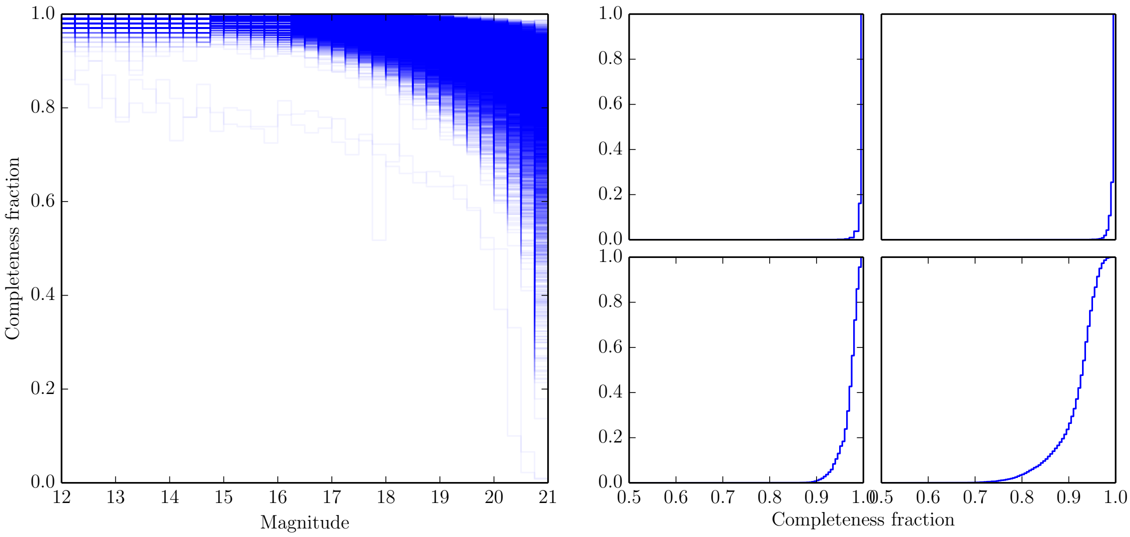


Figure 4.18: Left: Plot of every completeness curve in DR2, for magnitude bins of width 0.25 in range $12.0 < r' < 21.0$. **Right:** Cumulative histograms of field completeness for magnitude bins **a)** 12.0-12.25, **b)** 15.0-15.25, **c)** 18.0-18.25, **d)** 20.0-20.25.

to the recovered magnitudes correlates with increasingly poor seeing conditions.

Poor seeing was found to affect the measured PSF such that the returned magnitudes from an image were offset by a uniform amount rather than having their spread increased - this is illustrated by the tightness of the correlation between Δ_m values and seeing, up to the highest seeing values encountered in the survey. This justified the application of a uniform shift to all measured magnitudes from an image, applied to bring the median Δ_m of a field to zero.

Having reduced the systematic shifts to Δ_m values, a threshold value needed to be chosen for considering an artificial source recovered. As illustrated in Figure 4.15, the completeness curves remain relatively unchanged at $r' > 19$ as the Δ_m threshold is varied between 0.5 and 0.1 mag (the magnitude range of interest for correcting the density map). A threshold value of 0.25 mag was chosen as a compromise between avoiding admitting a large number of spurious detections at the faint end of the magnitude distribution, and ensuring that few objects are missed due to the occasional large Δ_m value.

4.5.3 Completeness fractions

Plotting the completeness curve for every field in the survey (Figure 4.18) shows that all fields share a rough curve shape. Only two fields deviate from the main group (whose completeness at bright magnitudes, as expected, reaches ≈ 1.0. These two fields are the pair 2298_nov2005, which cover the 0th magnitude star Capella on CCD4. The images exhibit saturation over many columns, leading to the low completeness fractions seen in Figure 4.18.

It has been noted since the first visual examination of IPHAS catalogues that bright stars create ghosts on the images, affecting their background levels, and therefore potentially

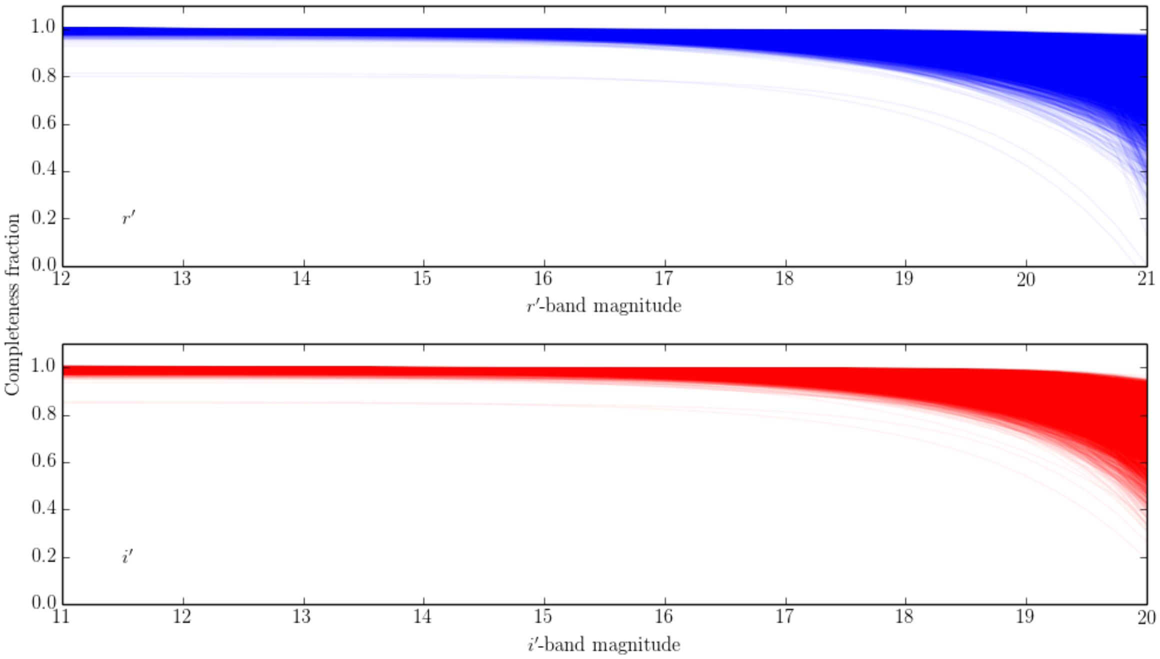


Figure 4.19: Best fits to completeness curves (as illustrates in Figure 4.18) for both r' (**upper**) and i' (**lower**) catalogues.

impacting the quality of the photometry. This is the first time this impact can be assessed quantitatively, in the form of the fraction of stars that are either not detected, or at the very least have their magnitudes measured incorrectly by $> \pm 0.25$ mag.

The discrete nature of the completeness fractions for each magnitude bin are due to the total number of sources simulated in each, as given in Table 4.1. The upper two cumulative histograms shown in 4.18 show that incompleteness remains low at bright magnitudes, with a median completeness of 99.5% at r'=15. At 18th magnitude the median completeness of IPHAS is 97.7% complete, falling to a median completeness of 93.1% at 20th magnitude.

4.5.4 Fitting to completeness fraction distributions

The aim of generating a correction factor for each source in the survey requires some form of lookup table. However the binned nature of the recovered completeness curves is an issue; for a given field, sources of magnitude 19.0 and 19.24 would be assigned the same correction factor. The noise present in these curves is also problematic - it makes little physical sense for the completeness fraction of a field to increase as source brightness decreases, yet this is seen in numerous fields due to fluctuations in the recovered completeness curves. Generating more sources to both allow reduced bin size and fewer fluctuations would require far more computing time than would be feasible (see Section 4.4.2).

It was decided to fit a function to the completeness curve of each field, allowing an appropriate fraction to be quickly obtained for every source in the survey. From visual inspection, it could be seen that an exponential function would fit well the curves in Figure

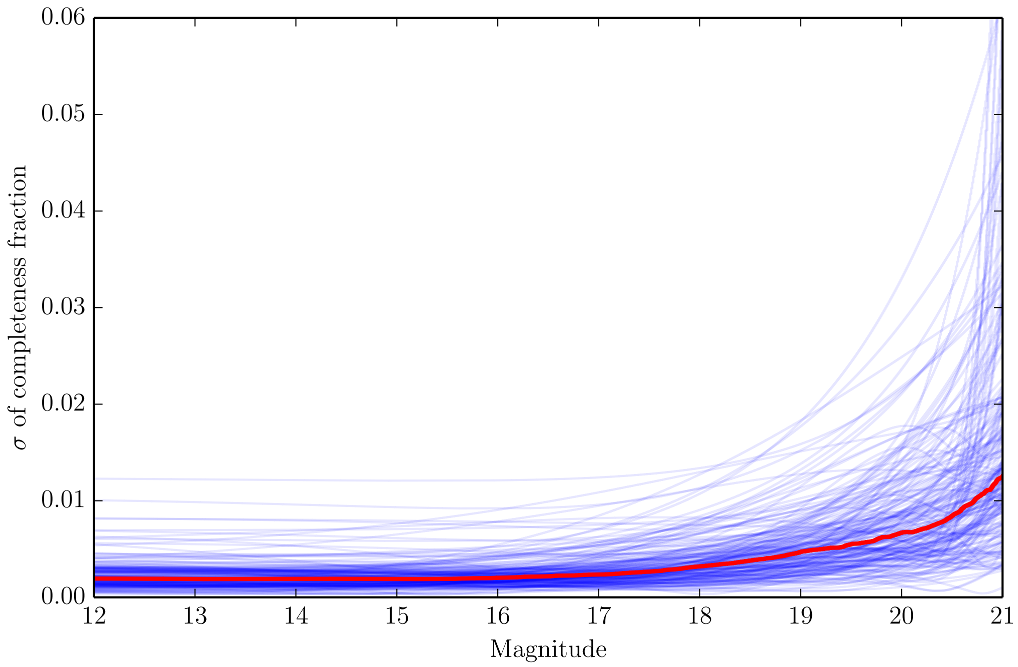


Figure 4.20: Variation of σ of completeness fraction across all four CCDs for a randomly selected sample of 200 IPHAS DR2 fields.

4.18. The final fits were of the form

$$C(r') = \alpha - \gamma \times \mathrm{e}^{\frac{r'}{\beta}} \tag{4.10}$$

where α, β, and γ are parameters allowed to vary to find the best fit for each field. These parameters were collected and formed a lookup table for use in correcting the density map. The curves of this form can be seen plotted for every field in Figure 4.19.

4.5.5 Inter-CCD differences

As mentioned in Section 4.4, the completeness curves generated from CCD 4 of each field were used to represent those of the field as a whole. To ensure that this approach was sensible, a random set of fields had completeness curves generated for all four CCDs in order to compare completeness curves and ensure that the variation between chips was acceptably low. For the magnitude range $12 < r' < 21$, standard deviations in the completeness corrections were calculated, and can be seen in Figure 4.20. At $r' = 19$, the deviation in completeness corrections between CCDs reaches as high as 0.025 - these cases occur where a bright star appears in one or more CCD of the field. These cases are rare; the median σ at $r' = 19$ is 0.005.

4.6 Correction map

Figure 4.21 shows the extent of the completeness corrections for an r'-band density map containing sources down to 19th magnitude. For this visualisation the completeness curve of the CCD that best covers each cell was used to calculate the correction for each cell. Cells that are less than $\frac{1}{3}$ covered by a CCD were excluded from this map; cells with such low coverage will have large uncertainties in the density map and will be easily excluded from the final product by placing a threshold on the estimated error.

The outer plane ($\ell > 120$) needs very little correcting; stellar densities in this region never reach sufficiently high densities to suffer badly from incompleteness. Some fields here can be seen to require larger corrections than neighbouring fields - this is due to bright ($r' < 10$) stars creating scattered light resulting in more source loss. Even in such cases, the required correction to fields in the outer plane is rarely more than one. An exception can be seen at ($\ell \approx 162.5, b \approx 4.5$), where the 0th magnitude star Capella has caused a significant amount of incompleteness - in close vicinity to Capella the required correction exceeds ten objects per $1' \times 1'$ cell. The impact of Capella on the field pair 2298_nov2005 can be seen in Figure 4.18 - these two images correspond to the two curves that sit below all others, never attaining a completeness above 80% over the entire magnitude range. Prior to publishing the density maps as a data product for public consumption, these fields will be excised from the dataset. At $\ell < 120$ the variation in corrections is more extreme, reaching up to ≈10 sources per cell in the densest regions of the map.

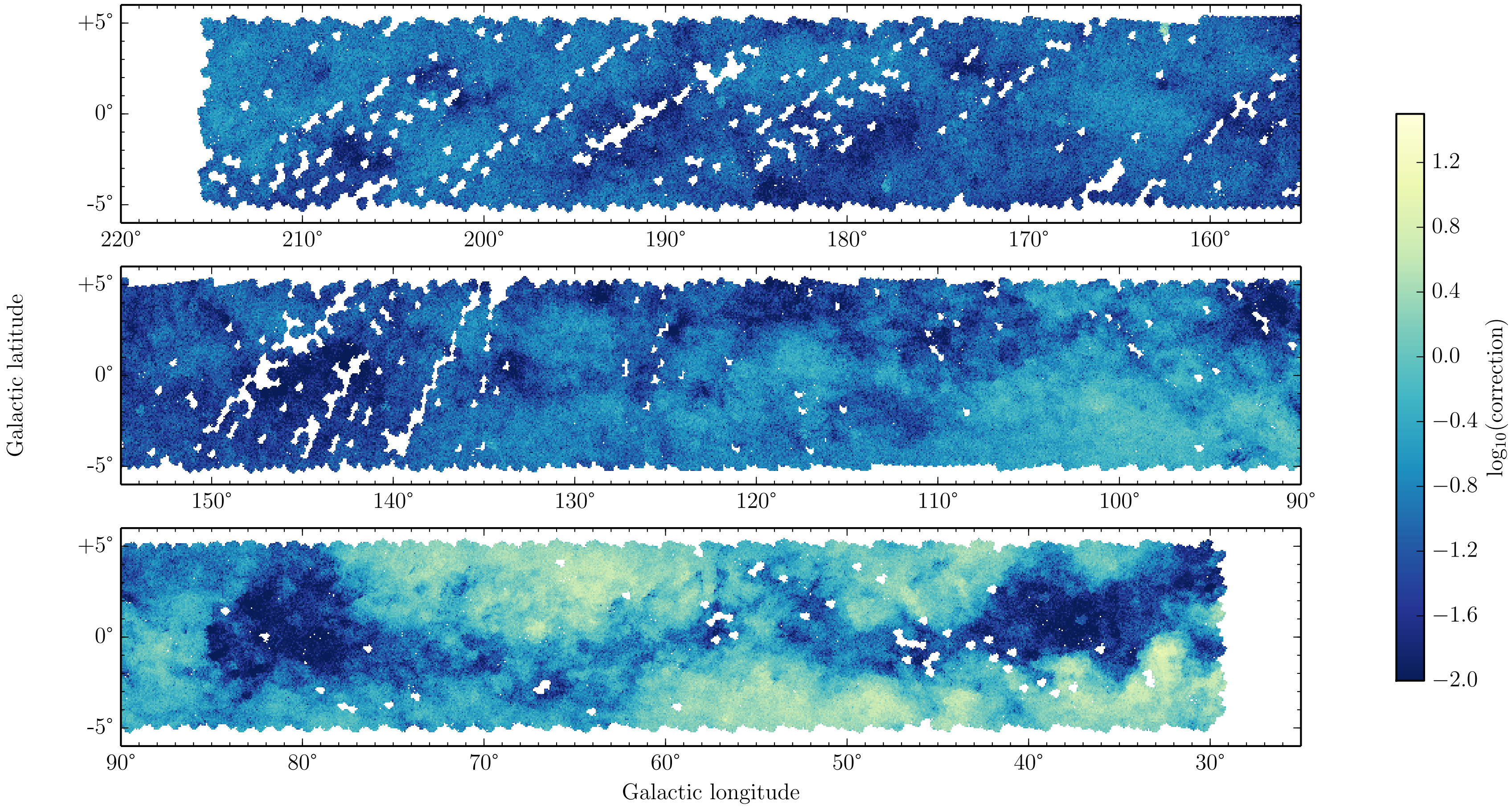


Figure 4.21: Map of corrections to r'-band density map down to $r' = 19$ - i.e. the logarithm of the number of additional sources added to each $1' \times 1'$ cell. Corrections are obtained from the CCD that best covers each cell, by taking the completeness curve (see Section 4.5.4) of that CCD and estimating the correction factor for each source present. Only cells that are at least $\frac{1}{3}$ covered by a CCD are included in this map.

Chapter 5

Density mapping DR2

Following a global photometric calibration and elimination of problematic data (see Chapter 3), the second data release (DR2) of IPHAS data was published (Barentsen et al., 2014). To maintain consistency and repeatability across IPHAS data products, the method of density map generation was modified to exploit this calibrated dataset.

Previously (see Section 2.4) the density map was assembled by crossmatching all overlapping fields and averaging the astrometry and photometry of sources determined to be observed multiple times. While this approach generated a usable density map, it did not take into account corrections for incompleteness, nor did it carry any survey footprint information - this information was lost once field pairs were merged. A new method, detailed here, was developed which measures independently the source density from each observation (accepted into DR2) of each point inside the IPHAS DR2 footprint.

5.1 Method of generation

5.1.1 Calculating cell coverage

The IPHAS footprint was split into cells of the desired resolution (the final cell size was $1' \times 1'$), and for each cell, a table containing the coordinates of every image contributing to DR2 (at the CCD level, i.e. $4 \times$ sub-images distinguished per field) was queried to identify which intersect the cell (either completely or in part). In order to calculate the coverage of each CCD, the pixel coordinates of each corner were determined, with an unusable border region taken into account for each of the four CCDs in order to exclude sources detected far from the WFC optical axis (source counts from these regions would be unreliable). Figure 5.1 shows the extents of the CCDs that were included when calculating coverage.

In order to correct the occupancy of density map cells, accurate coverage information was essential. The area of intersection between a cell and a CCD was calculated first using the CCD limits obtained during the generation of DR2 data. These limits were calculated by taking the maximum and minimum values of right ascension and declination from sources in the field. The sources used in this estimation of field boundaries were not limited to point sources - extended sources and even noise were used to determine detector edges, allowing

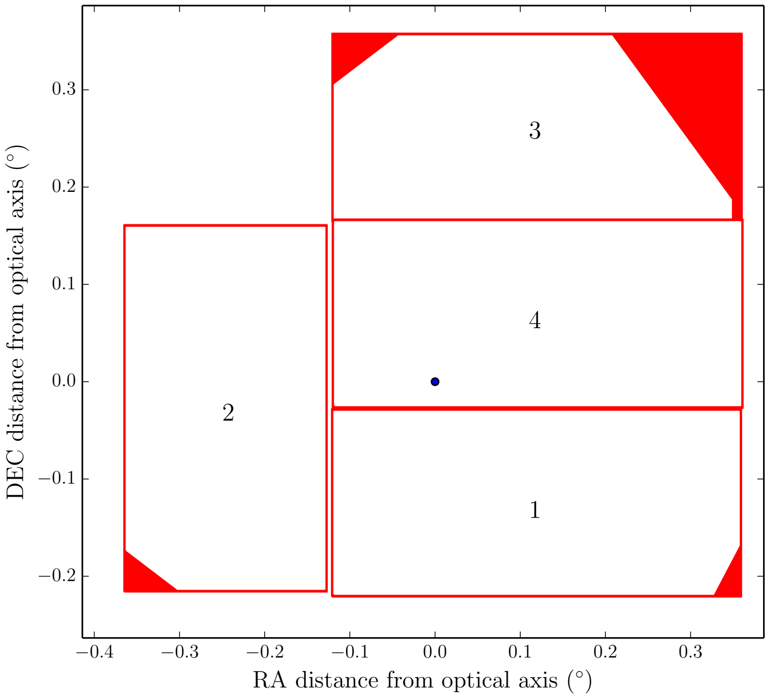


Figure 5.1: Area discarded when calculating coverage of density map cells by IPHAS CCDs contributing to DR2. Blue point denotes optical axis of the WFC. CCD chips are numbered.

this method to be applied to fields with a low density of detected point sources.

Upon the generation of a $1' \times 1'$ density map, a regular pattern of lower density stripes presented themselves. Closer inspection revealed these to appear at CCD boundaries with (almost) constant declination (i.e. boundaries between CCDs 1&4, and 3&4). A sample of fields were chosen for investigation, and the positions of the detectors visualised. It became evident that in several cases, the boundaries of detectors appeared to overlap on the sky.

Examination of the documentation of the WFC to determine the actual size of inter-CCD gaps revealed the lack of rectilinearity in the equatorial system present in the arrangement of WFC CCDs. This feature means that taking maximum/minimum values of RA & DEC from sources detected in the image will return a CCD footprint larger than the true footprint.

The alternative method employed was to calculate the corners of each CCD using their pixel coordinates along with the world coordinate system recorded in the header of the image. Figure 5.2 shows the difference in the CCD boundaries determined using the two methods, where the difference (the red area not actually covered by the CCDs) accounts for the stripes of low values emerging in the density map. The filling factors of regularly spaced cells were being reported as higher than in reality, resulting in underdensities.

Determination of overlaps between CCDs and cells were determined using the Shapely module for Python, which allows representations and manipulations of points and polygons, including the combination of polygons, determination of their intersections and areas, and identification of points lying within polygons.

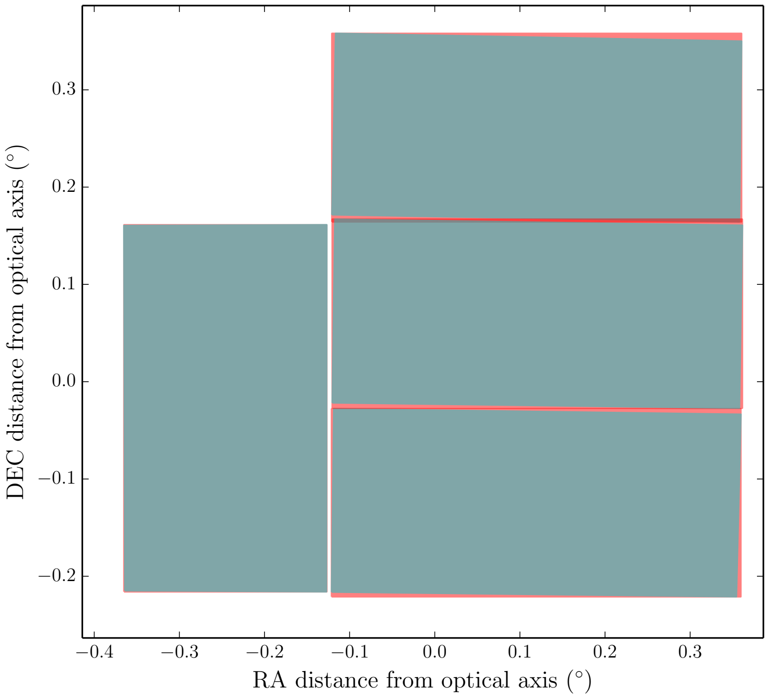


Figure 5.2: Area covered on sky computed from coordinates of sources falling in CCD (**red**) and from calculating the corner coordinates using the WCS header information of each CCD (**blue**).

For each CCD covering a cell (either completely or partially), the fraction of the cell covered was computed, building up a table which was used to access relevant field catalogues for retrieving source counts.

5.1.2 Populating the density map

For each cell, the relevant files identified in the table generated in Section 5.1.1 was accessed, and all sources within the cell boundaries selected from the intersecting fields. Sources meeting the following criteria were retained:

- Morphology classification -1, -2 or +1
- Brighter than faint limit (default $r' = 19$)
- Flag errBits < 64 for band of interest

where the errBits criterion eliminates sources with bad pixels within their PSF, which are truncated, or which are vignetted. The exclusion zone shown in Figure 5.1 will have eliminated the majority of such cases, while this pass removed sources affected by issues such as bad columns and hot pixels.

For each remaining source, a *corrected contribution* to the number count of the cell was computed from the completeness curves generated in Section 4.5.3. The table containing α, β, and γ values (defining Equation 4.10) for each field was read in, the values relevant to the current field identified, and the magnitude of the source under consideration used to calculate the corrected contribution to consider. The original (ϕ) and corrected (Φ) source counts were recorded, along with the area of the intersection between the cell and the IPHAS field. Each estimate of total cell occupation, Φ' is given by

$$\Phi' = \frac{\Phi}{coverage fraction} \tag{5.1}$$

and the estimate of uncertainty is simply the Poisson noise of the observed area scaled to the cell:

$$\Phi'_{err} = \frac{\sqrt{\Phi}}{coverage fraction}. \tag{5.2}$$

This is repeated for each CCD covering the cell, and for each cell in the density map.

This method of populating density map cells provides additional information compared to the averaging of repeated sources. At $1' \times 1'$ resolution, the density map contains 6,537,051 cells overlapping with IPHAS photometry. Of these, 92.9% have an overlap with a second IPHAS CCD, with 38.3% having a third. These cells provide an estimate of the variance in source counts between observations, which in turn can be compared with the Poisson uncertainty calculated for these cells (see Section 5.2.2).

5.1.3 Empty cells and the extent of IPHAS

In the final density map, there exist two forms of 'empty' cell. The first are the cells which are considered 'uncovered', i.e. beyond the current footprint of IPHAS observations. These are to be expected beyond the defined edges of the survey ($29 < \ell < 215, -5 < b < 5$). Several fields at survey boundaries extend beyond $|b| > 5°$, in places reaching $|b| \approx 5.4°$. Figure 5.3 shows a multi-order coverage (MOC) representation of the IPHAS DR2 coverage footprint, generated at HEALPix order 12, corresponding to a pixel size of $51.5''$ (chosen for its similarity to the resolution of the density map).

Cells beyond the defined bounds of the survey are not the only source of uncovered cells - those which fall in the regions covered solely by observations which have been deemed unfit for release in DR2 are also uncovered. Such cells are stored in the density map as NaN, in order to avoid confusion by the user with astronomically unpopulated cells.

Genuinely empty cells are those in which no objects have been detected, despite being covered, at least partially, by IPHAS observations. These cells are most likely in regions of low source density, however in some cases these can be caused by cells covered only once by a partially overlapping CCD that results in a very small intersection area. For empty cells, uncertainties as computed for occupied cells (see Section 5.1.2) are undefined. The approach adopted for estimating an uncertainty on unoccupied cells was to attempt to estimate the upper limit on sources in the area of cell/CCD intersection. This was achieved by taking the the curve parameterising the completeness of the relevant IPHAS observation (see Section 4.5.4), and calculating the completeness fraction at the faint magnitude limit being applied. The upper limit was defined as a single object at the limiting magnitude, corrected by this

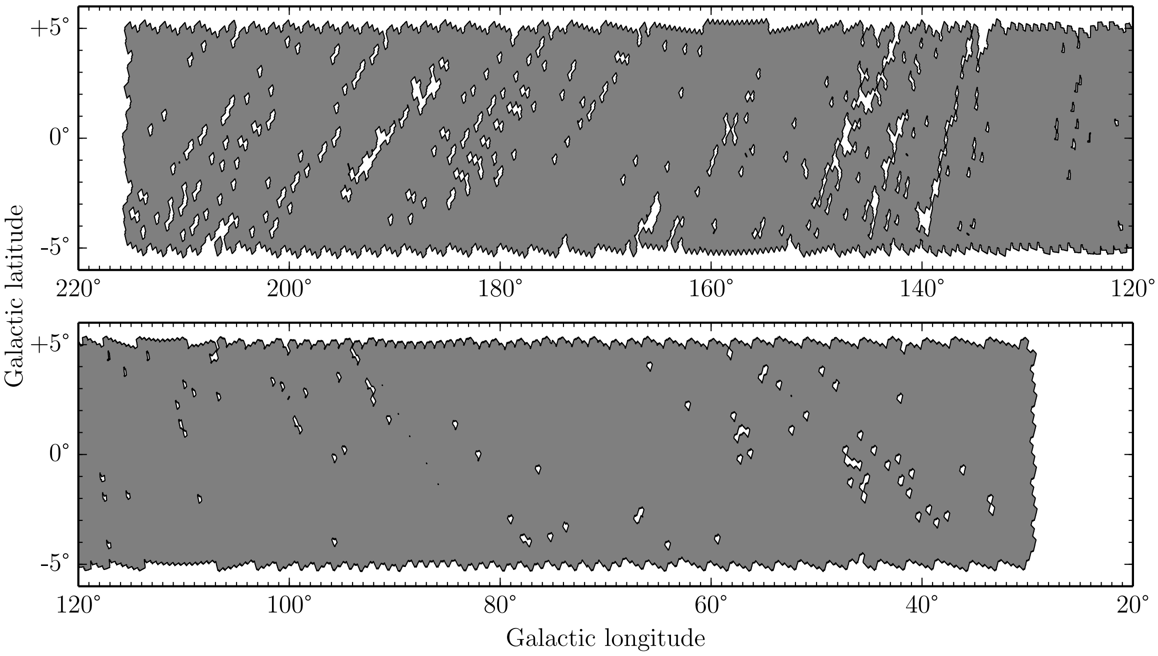


Figure 5.3: Multi-Order Coverage (MOC) representation of the area covered by DR2, generated by reading in the world coordinate system information from each CCD contributing to DR2. MOC files use the HEALPix system, with this file computed to 12th order (equivalent to an angular resolution of 51.5″).

completeness fraction; for example, in the case where the density map cut at $r' = 19$ contains an empty cell, and the completeness fraction at this magnitude for the intersecting CCD is 0.85, the upper limit on the uncertainty was recorded as $\frac{1}{0.85}$.

This uncertainty was scaled by the area covered by the cell/CCD intersection area, such that empty intersections accounting for small fractions of cells will return large uncertainties.

5.1.4 Cell coverage

The approach described above improves over the naive treatment laid out in Section 2.4 by including information on cell coverage. The fraction of each cell intersected by a CCD was recorded, and the measured occupancy of the cell scaled by this coverage factor. The Poisson uncertainty was increased proportionally, allowing cells whose densities have been extrapolated from a small coverage fraction to be excluded. The improvement this delivers can be seen in Figure 5.4 where in the left-hand panel, the dearth of sources in cells only partially covered (at survey boundaries and between CCDs in fields with only one of the primary/offset pointing pair) has been ignored. In the right-hand panel, the fields at the survey edges have been excluded due to low coverage (see Section 5.2.3), and the remaining cells have been suitably corrected. As a result the inter-CCD gaps visible in the left-hand panel are no longer present.

Figure 5.5 shows the distribution of coverage fractions for cells in the $1' \times 1'$ resolution density map, where it can be seen that more than 96% of cells are entirely covered. The

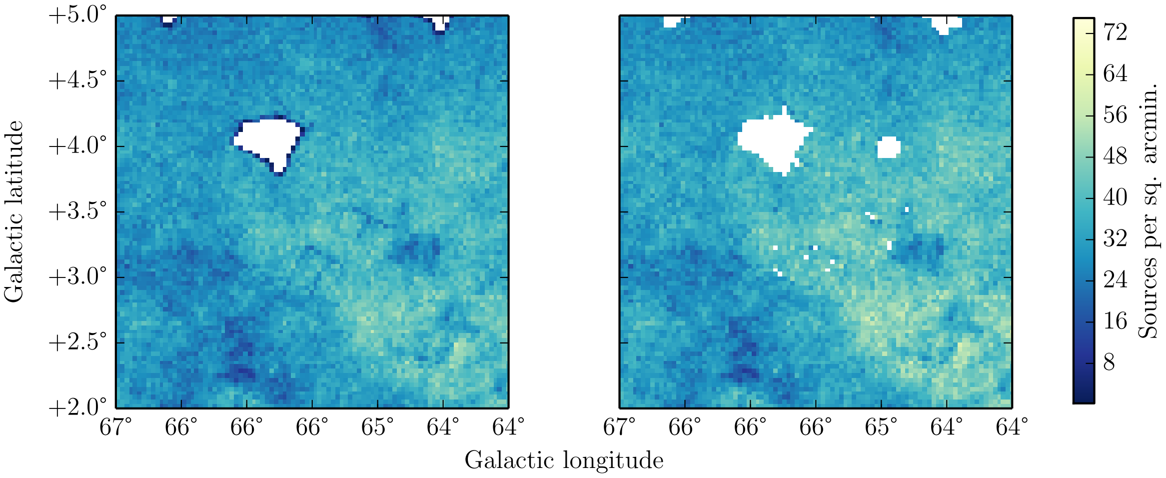


Figure 5.4: Two versions of the density map showing the same region. **Left:** Generated using the method laid out in Section 2.4. **Right:** Generated using the method laid out in Section 5.1 after applying completeness corrections. The new map no longer shows the outlines of WFC CCDs, caused by cells only partially covered by an IPHAS field. It does, however, have a number of empty cells where the coverage is zero (at a resolution of $2' \times 2'$, chosen for ease of comparison). The additional hole is due to the $5'$ exclusion zone around bright stars (see Section 5.1.6).

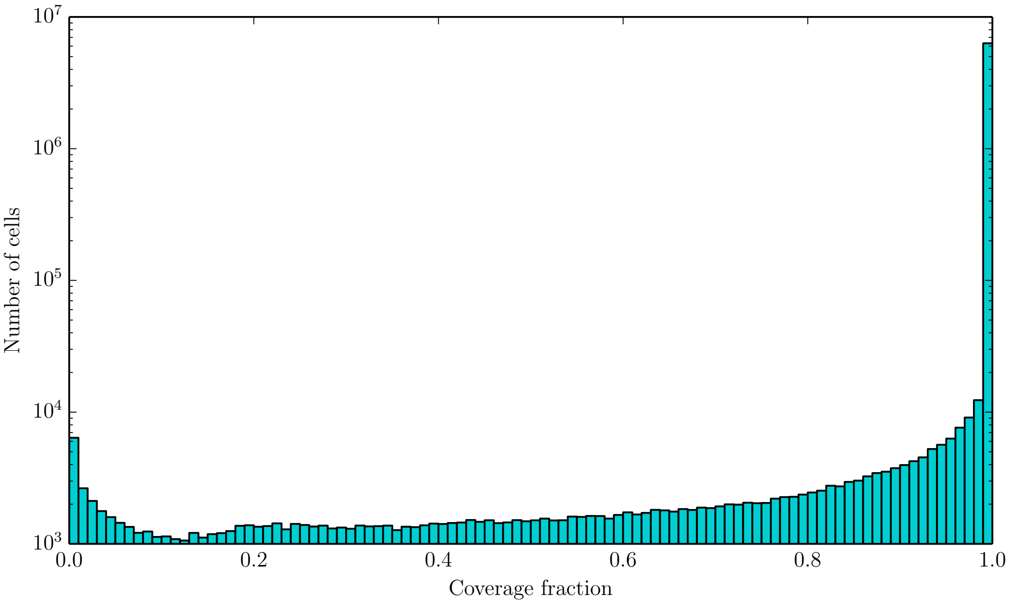


Figure 5.5: Histogram of cell coverage fraction across the entire survey when cells of $1' \times 1'$ resolution in (ℓ, b) are used. Only the highest coverage fraction per cell was taken in cases where multiple CCDs cover a cell.

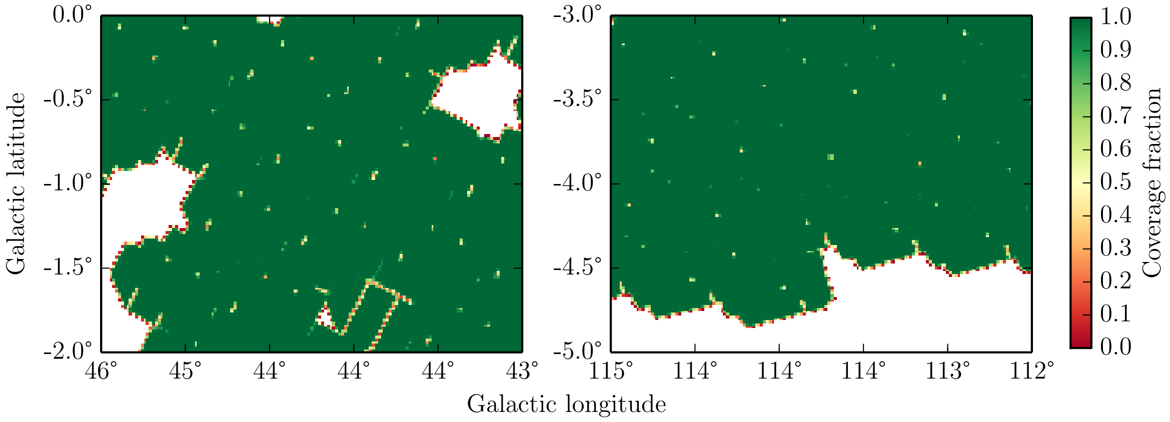


Figure 5.6: Maps showing the variation in cell coverage across two regions of the survey at $1' \times 1'$ resolution. The lines of low coverage inside the survey footprint are inter-CCD gaps present due to missing field partners (i.e. areas where only one of a primary/offset pointing have been included in DR2).

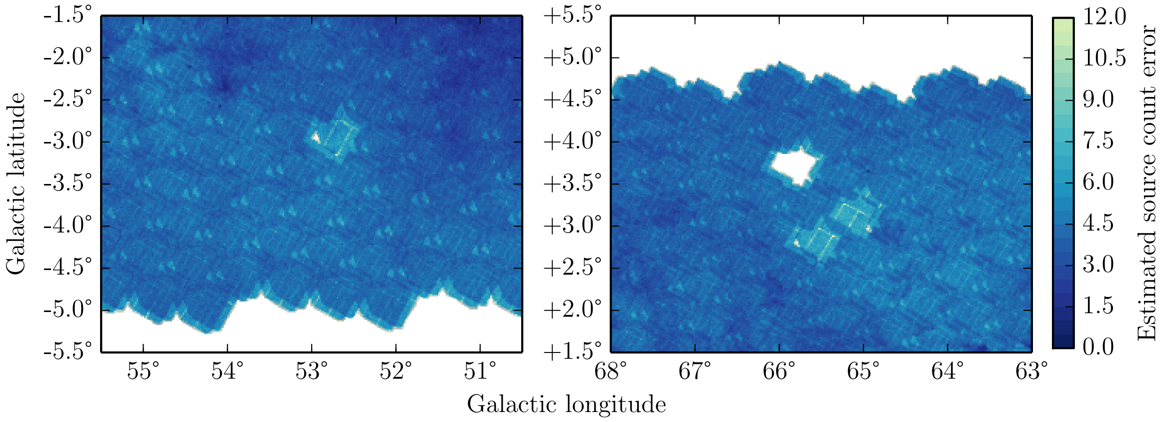


Figure 5.7: Cutouts of the r'-band density map at $1' \times 1'$ resolution showing the variation in estimated error per cell.

spatial distribution of coverages is illustrated for two regions of the survey in Figure 5.6. In addition to the expected lower cell coverage at DR2 footprint edges, a repeated pattern of lower coverage is apparent - these less covered cells are those which partially fall into the zones of avoidance defined in Section 5.1.1, and between CCDs of fields where only one of a primary/offset pair has been included in DR2. The size of these regions change across the survey footprint due to the increasing overlap between fields towards $\ell \approx 124°$ (0h RA), an effect apparent between the two regions in Figure 5.6.

5.1.5 Error

The primary estimate of error accompanying the density map comes from the Poisson error in counts per cell/CCD intersection. These are scaled up to correct for the area covered by the intersection, before the counts from multiple intersecting CCDs are averaged, weighted according to their scaled errors. Figure 5.7 shows two region cutouts highlighting the variation in error estimates. The subtle pattern visible corresponds to the pattern of cell coverage

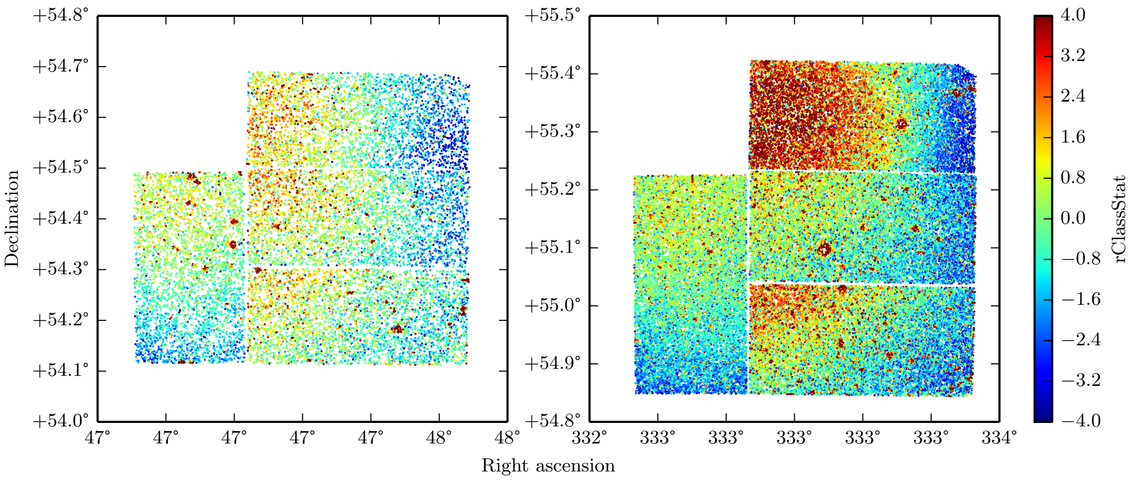


Figure 5.8: Plot of sources detected in the r'-band image for fields 1080_oct2004 (**left**) and 6970_oct2004 (**right**), with point colour mapped to the *rClass* statistic (see Section 2.2.2).

fractions shown in Figure 5.6. Survey edges are lighter due to small cell/CCD area intersections. Lighter regions inside the survey bounds are due to regions where fewer CCDs cover individual cells; this occurs at CCD edges (repeating grid pattern) and where only one of a primary/offset field was included in DR2.

5.1.6 Remaining data issues

Once the density map was generated, remaining issues with the DR2 data were identified. The most prominent was a strip of low density holes running in the direction of an IPHAS run. Upon investigation it emerged that all fields covering these holes were observed on one night during November 2004 (20041103).

Figure 5.8 shows the variation of the *rClassStat* (the statistic which determines the morphology classification *rClass*) for two of the affected fields. Section 2.3.1 discusses the misclassification of stellar sources as extended, as illustrated in Figure 2.6. The problem evident in fields observed during the afflicted November 2004 night is a more extreme manifestation of this problem - rather than PSF distortion only causing larger classification statistic values (i.e. apparently extended/non-stellar) further from the optical axis, distortions across all CCDs cause variations from low (noise-like) to high (non-stellar) *rClassStat* values.

The reason for the uncharacteristic misclassification of sources as noise is that the telescope was poorly focused on the affected night. This led to a variation of the PSF across the image that generated a pessimistic median PSF by which all objects were judged. The regions in which objects were better focused, therefore, were considered unrealistically sharply peaked by IMCORE, leading to a noise-like *rClassStat*.

The sources that end up classified as noise-like in these fields were not included in the first density map generated with DR2 calibrated data; once this issue was identified, a clause was added to to the source counting code to include objects with morphology code 0 for fields

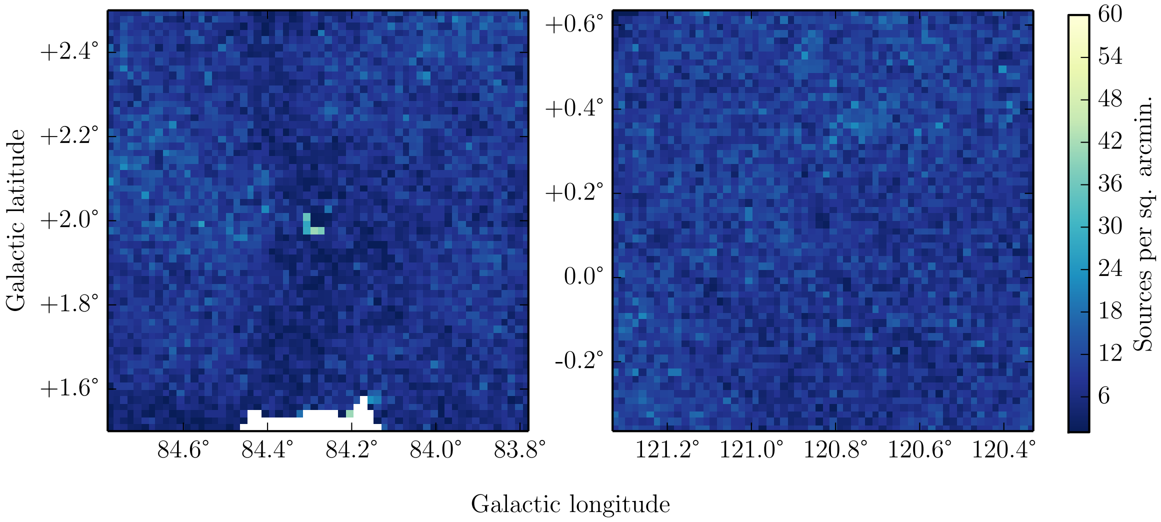


Figure 5.9: Cutout of the r'-band density map at $1' \times 1'$ resolution, centred on two bright stars: Deneb (**left**) and kap Cas (**right**). An excess of sources is caused by the presence of Deneb (V=1.25) due to scattered light, while an absence of sources can be seen at the coordinates of kap Cas (V=4.16).

observed on the affected night.

Another issue with the density maps, generated from single-band catalogues, is that around bright ($V \lesssim 3$) sources, the scattered light produced can generate a significant number of spurious detections that can appear stellar (or extended - see Section 2.3.1). Cross-matching to other bands to ensure repeated detections is not an option for these single-band catalogues - while it might be feasible in the r'-band, it fails in the i'-band because a large number of red sources are observed with no counterpart detected in r'. Requiring i'-band sources to have an r'-band counterpart would severely limit the number of genuine sources contributing to the i'-band map.

Around stars fainter than this but still ($\lesssim 5$) far brighter than the saturated limit of IPHAS, the effect of saturation around the sources is to reduce the number of sources observed in a much more localised fashion than taken into account in Chapter 4, where the completeness of entire fields was estimated.

Figure 5.9 shows the region around two bright stars, Deneb and kap Cas. The scattered light around Deneb (a V=1.25 object) causes an excess of sources in the r'-band catalogues, while the fainter (V=4.16) kap Cas does not produce such a source excess, but rather a dearth of sources due to a high degree of localised incompleteness. In order to avoid small scale issues in the final density map, cells within $5'$ of stars brighter than $V = 5$ appearing in the catalogue of Hoffleit & Jaschek (1991) have simply been removed. Figure 5.10 shows the distribution of these bright stars across the Northern Galactic Plane. A $5'$ zone of exclusion results in ≈0.2% of the density map cells being discarded.

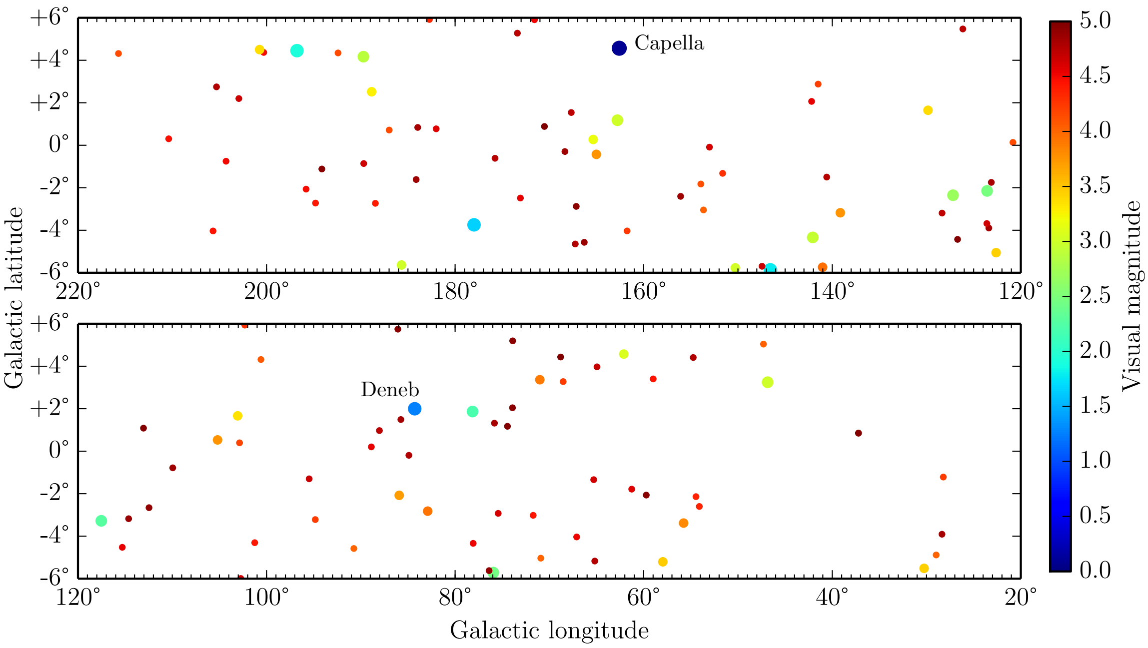


Figure 5.10: Distribution of bright ($V < 5$) stars in the Galactic Plane, based on the catalogue of Hoffleit & Jaschek (1991). Symbol size, as well as colour, is scaled to visual magnitude.

5.2 Using the density map

5.2.1 Choosing the resolution of the density map

The previous (see Chapter 2) method of generating density maps produced a $2' \times 2'$ resolution view of the Galactic Plane. This was an arbitrarily chosen scale which balanced well the computing time required to regenerate the map while modifying the scripts involved, and the detail visible from the resulting maps.

In preparing the density maps now generated from DR2, the final scale was reconsidered. Initially a $1' \times 1'$ resolution was generated, with the intention of rebinning the density map to the previously used $2' \times 2'$ resolution. However the additional detail visible in the $1' \times 1'$ resolution map was a noticeable improvement, to the extent that it was adopted as the default resolution for the density map to be offered in. Figure 5.11 shows the two resolutions side-by-side, illustrating the increased definition of finer structures brought out by the $1' \times 1'$ resolution. Although the $1' \times 1'$ map appears less complete (due to missing cells), the $2' \times 2'$ map was generated by reducing the resolution of the $1' \times 1'$, interpolating over missing cells; i.e. no new information is present.

Coarser resolutions have also been generated, for use in situations where the density over larger regions would be required - e.g. to support planning pointings for multi-fibre spectrographs with a large (≈1° diameter) field-of-view.

A resolution of $1' \times 1'$ compares favourably with density maps presented in previous studies of stellar number densities. In previous works density maps tend to be intermediate

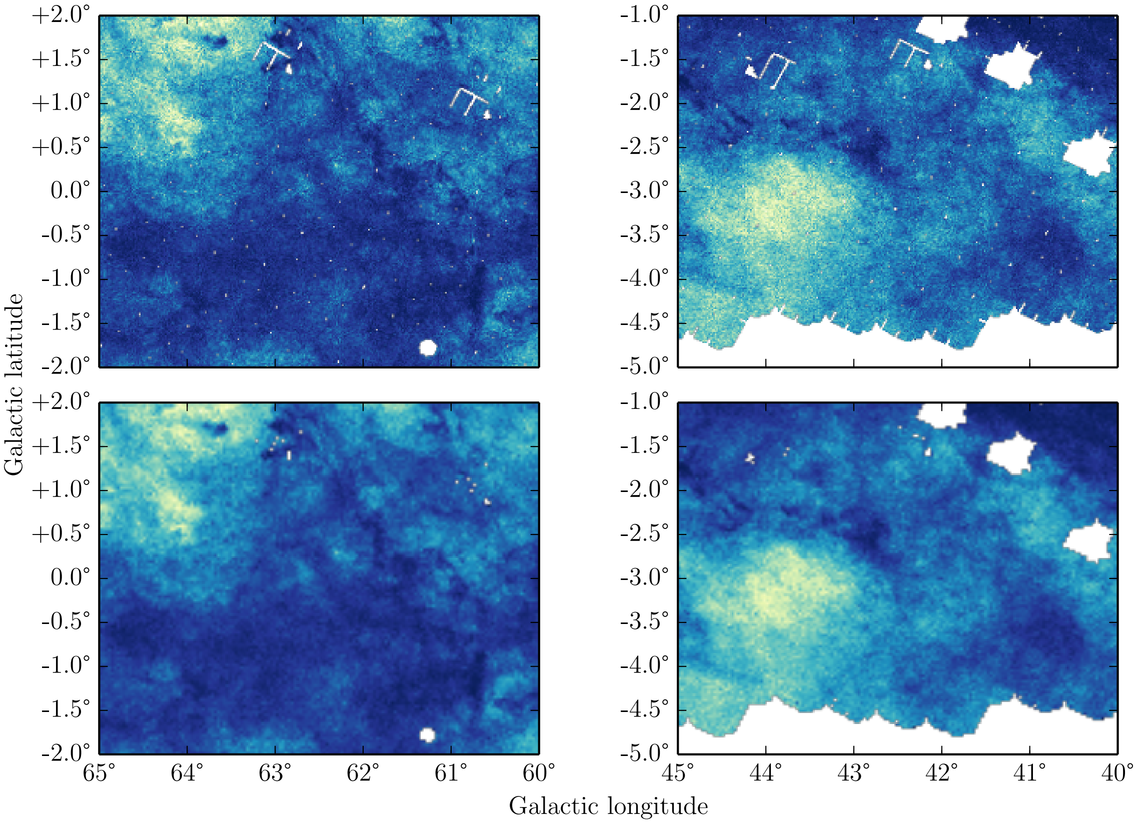


Figure 5.11: Two subsections of the density map, displayed at $1' \times 1'$ (**upper**) and $2' \times 2'$ (**lower**) resolutions.

data products, generally produced ad hoc for small regions under investigation (e.g. the study of foreground stars towards the LMC by Zaritsky et al. (1999), who chose a binning of 84″). While this approach has the advantage that the grid used for binning can be chosen and modified (Froebrich et al., 2007, , who used a 3.5″ binning), the fact that such maps are produced using final data products (i.e. released survey catalogues) means that corrections for incompleteness cannot be made in the way presented in Chapter 4. The scale and completeness-corrected nature of the maps presented here therefore present a unique data product of use to studies requiring stellar number densities across the Northern Galactic Plane.

5.2.2 Deviation in repeated cell counts

The ability to measure the source counts in a cell independently - due to multiple overlapping CCDs - allows the effect of the corrections determined in Chapter 4 to be assessed. In the case that the applied corrections are inconsistent with one another (i.e. result in significantly different estimates of corrected source count), then the standard deviation of the count estimates will be large - significantly larger than the averaged Poisson uncertainty on the cell. In this case, using the Poisson uncertainty to estimate the error of the density map will underestimate the error.

Figure 5.12 shows a two dimensional histogram comparing the distribution of the standard deviation of the repeated count estimates and averaged Poisson uncertainties for all cells. The averaged Poisson uncertainties do not exceed ≈10 sources - consistent with an upper limit of cell density at $\approx 100 - 200$ (Figure 5.13 shows the distribution of cell counts in the $1' \times 1'$ r'-band density map). It can be seen that for the majority of cells, the averaged Poisson uncertainty is greater (i.e. slightly more pessimistic) than the standard deviation of repeated corrected count estimates. The tail of cells with higher σ values are those where CCDs covered a small fraction of the cell, leading to large corrections (and therefore large errors, which translate to an increased deviation between repeated cell counts). A lower σ compared to the Poisson uncertainty for the majority of cells suggests that the corrections applied in Chapter 4 are consistent between CCDs.

5.2.3 Usability

The distribution of cell densities (Φ') shown in Figure 5.13 includes cells that have been scaled up from small cell/CCD intersection areas. As discussed in 5.2.2, the Poisson uncertainty on the unscaled density ($\sqrt{\Phi}$) in the intersecting area is also scaled to the full cell area (Φ'_{err} - see Equation 5.2), meaning that for cells where there are no other intersecting CCDs with which to generate a weighted average, the final uncertainty on the cell will be large.

To determine the degree to which the density in each cell can be considered significant, the ratio

$$\nu = \frac{\sqrt{\Phi'}}{\Phi'_{err}} \tag{5.3}$$

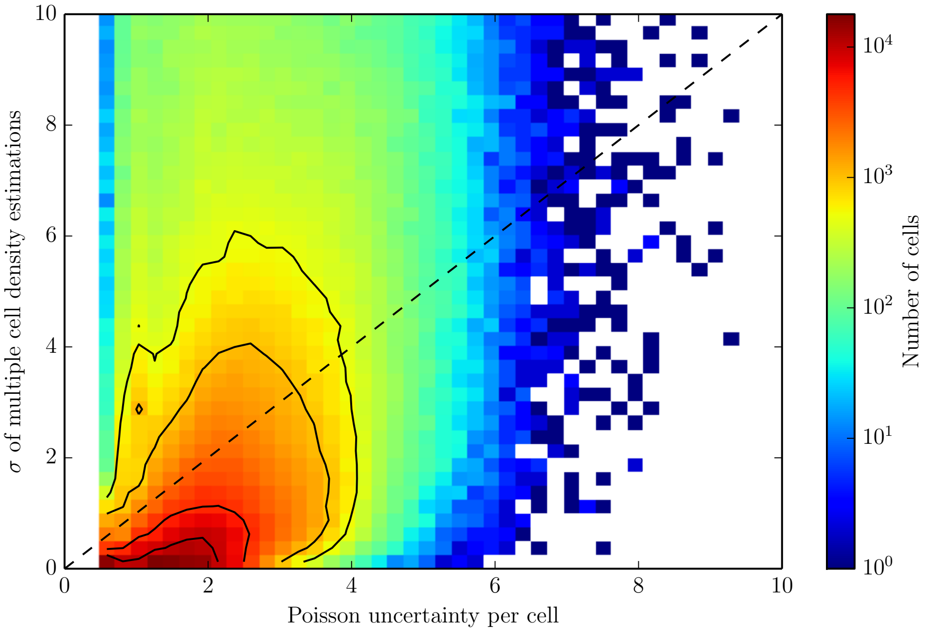


Figure 5.12: Density plot of cells which are more than 20% covered by at least three CCDs, according to the deviation in sources counted in each cell/CCD intersection (scaled to 100% cell coverage) and Poisson uncertainty resulting from weighted averages. The majority of cells show greater Poisson uncertainty than deviation between repeated observations, while the tail of higher count deviations is due to cells with low coverage by IPHAS CCDs.

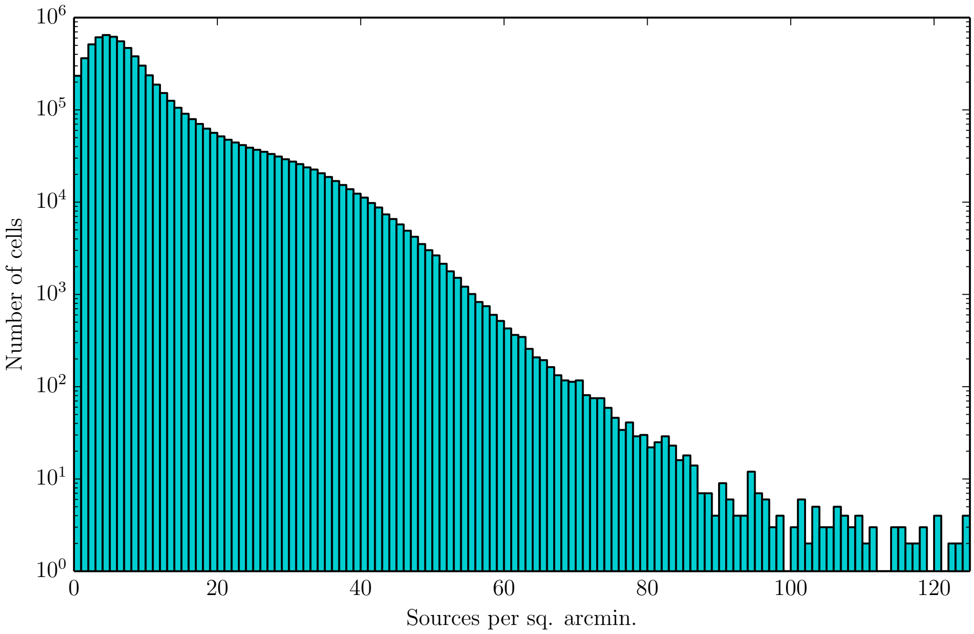


Figure 5.13: Histogram of source densities for all cells in the r'-band density map at $1' \times 1'$ resolution. The histogram is truncated at 125 sources per sq. arcmin., although a small number of cells have higher occupancies.

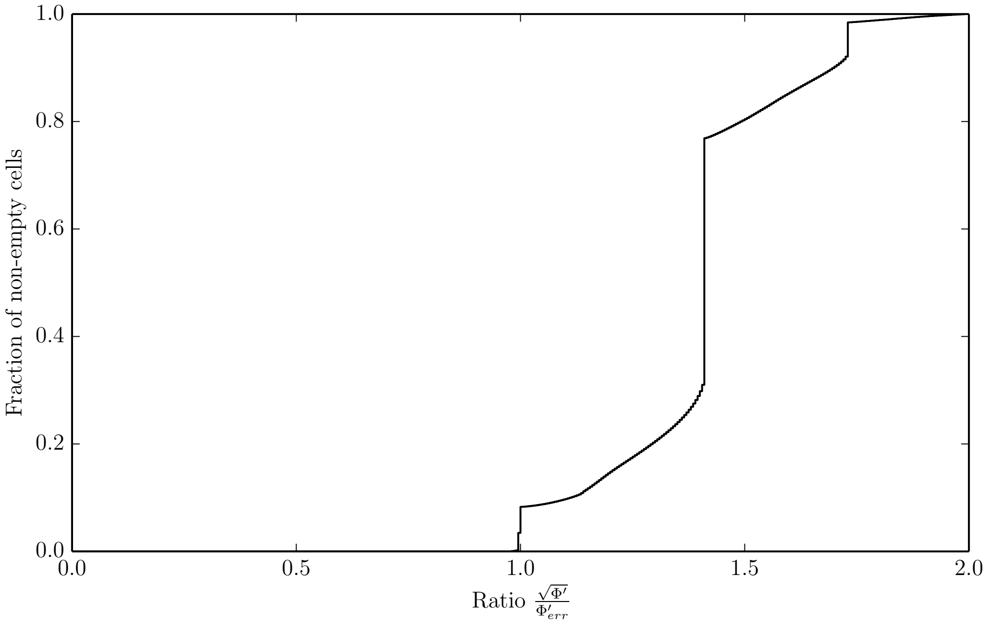


Figure 5.14: Cumulative histogram of $\frac{\sqrt{\Phi'}}{\Phi'_{err}}$ (see Equation 5.3) for all non-empty ($\Phi' > 0$) cells.

was used. In the case that a cell is intersected by only one CCD, which covers it entirely, the final uncertainty will be the Poisson uncertainty on the density of the cell ($\Phi'_{err} = \sqrt{\Phi}$) - resulting in a ratio $\frac{\sqrt{\Phi'}}{\Phi'_{err}}$ of 1. This ratio will drop below 1 in cases where the measured densities and associated uncertainties were scaled up to the area of a cell.

Figure 5.14 shows the cumulative fraction of non-empty ($\Phi' > 0$) cells against the value of this ratio - these cells make up 97% of the total density map. Only 2% of these non-empty cells display a ratio below 1.0 - these are located at survey boundaries and in inter-CCD gaps of fields with no primary/offset partner. The remaining 98% of non-empty cells are either covered well once (the cells at a $\frac{\sqrt{\Phi'}}{\Phi'_{err}}$ of 1.0) or more (displaying higher values). These are the occupied cells for which statistically significant measures of density have been obtained, and therefore included in the final density map.

These are only the cells in which at least one source was counted; 3% of cells contained no sources (either due to falling in genuinely sparsely populated regions of the Galactic plane or due to low intersection area between the cell and covering CCDs). The uncertainty adopted for these cells is the equivalent to detecting a single source at the faint limit of the density map (see Section 5.1.3), prior to scaling for area. For the $r' = 19$ limited density map, these uncertainties will be $\lesssim$1.17 - equivalent to an incompleteness of $\approx$0.85 (see Section 4.5.4). As these uncertainties will be scaled by area of the cell/CCD intersection to produce ν, the magnitude of ν can be used to determine which empty cells should be included in the density map as being statistically significant, and which should be discarded as being unreliable.

The median source count across all r'-band density map cells (the distribution shown in Figure 5.13) is $\approx$6.18, corresponding to $\approx$20,000 sources per sq. deg.. Assuming that over the $1' \times 1'$ cells that source density is uniformly distributed, on average 16% of the area of any one cell would need to be covered in order to detect a single source - if less than this area is covered , the source count would be unreliable. Taking the expected uncertainty (≈1.2) on empty cells, and rounding up the minimum acceptable cell coverage fraction to 20%, I obtain a scaled error of 6 sources for an empty cell's uncertainty. Above this value, the cell coverage fraction is below the reliable threshold.

At this point the maps were deemed sufficiently cleaned to be exploited, with ≈98% of the original cells passing the above criteria. Figures 5.17 - 5.20 show the final r'- and i'-band density maps at $1' \times 1'$ resolution. These figures show that the excluded cells are not only around the borders of missing data, but are also inside the survey footprint - this occurs most often where only one field of a primary/offset pair has been observed, and cells falling between the CCDs of the observed pointing are not well covered.

5.2.4 Example: cluster search

As an example of potential uses for the density map, a naive search for clusters was run on the $1' \times 1'$ i'-band map (limited to $i' < 18$). A $20' \times 20'$ window was moved across the map, with missing data being interpolated provided no more than 20 cells (5%) in the window fell outside DR2 coverage. An elliptical Gaussian was fit to the cells inside the window, and if successfully fit, its parameters were recorded provided it displayed

- a positive amplitude greater than 5 counts
- centroid position (x_0, y_0) in central $10' \times 10'$ window
- both σ values non-zero and less than $10'$
- floor (i.e. background) value positive, less than 150 sources per sq. arcmin.

All fitted 2-D Gaussians meeting these criteria were collected and the table internally cross-matched, associating all fits with (x_0, y_0) positions within $5'$ of one another. Median fit parameters were calculated, and fits not meeting the following empirically derived criteria were excluded.

Groupings of fewer than 30 successful fits were dropped as likely spurious chance associations. An upper limit of $4'$ was placed on both median σ values - this places a constraint on the size of clusters that can be detected. The most prominent clusters visible by eye in the density maps are larger than this extent, however their number is low; most detectable clusters which are not immediately obvious by eye are likely compact. There is also a possibility that clusters cause only a small enhancement in stellar density, however this method of picking out clusters is likely to pick up noise if this regime is probed. For this reason a constraint was placed requiring that the median amplitude of the grouped fits be greater than twice the Poisson uncertainty in median background counts.

Using the above criteria, 100 candidate overdensities were identified, 67 of which were found to correspond to clusters in SIMBAD (reduced from 201/1962 without any selection criteria). Figure 5.15 shows several examples of known clusters picked up using this method.

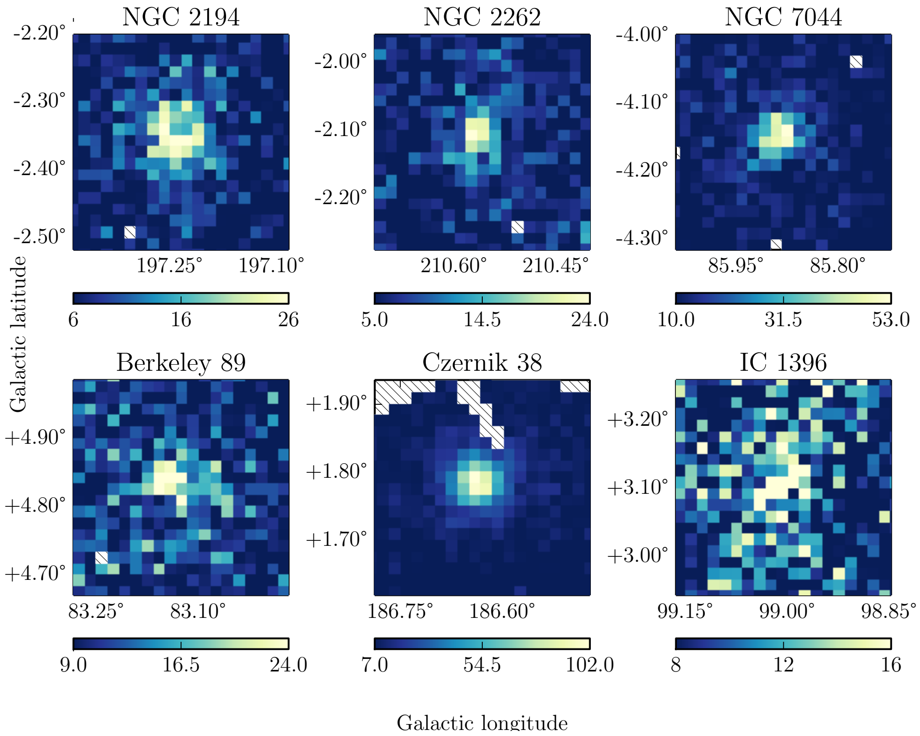


Figure 5.15: Example of six clusters found in the $1' \times 1'$ i'-band density map, limited to $i' < 18$. Regions not covered by the density map are hatched to distinguish between no data and high counts.

An identical treatment was applied to the r'-band density map (limited to $r' < 19$), in order to ascertain what fraction of clusters picked out in the i'-band also appeared in the bluer r'-band.

A number of overdensities detected in the r'-band did not correspond to known clusters. Upon extending the SIMBAD match radius around these overdensities, three were found to lie between $5'$ and $10'$ away from known cluster coordinates. These are listed as having angular diameters greater than $5'$ in the literature: Berkeley 14 ($18'$ (Kharchenko et al., 2005)), Berkeley 53 ($12'$ (Maciejewski et al., 2009)) and IC 1311 ($6'$ (Delgado et al., 1994)). An additional comparison against WEBDA database of stellar clusters[1] found an additional match, as its coordinates listed for Berkeley 101 place it almost a degree from its coordinates reported by SIMBAD.

Of the 71 SIMBAD-matched clusters in the i'-band, 45 were detected in the r'-band within $5'$ of the redder detection, with 26 not matching any r'-band feature.

Of the 29 remaining overdensities, six also appeared in the r'-band density map, leaving 23 detected only in the i'-band. This is significantly different to the overdensities matched to known clusters; 63% of matched clusters were detected in both i'- and r'-bands, while only 21% of unmatched overdensities are detected in both. Table 5.1 summarises the number of overdensities that are matched and unmatched, and in which filters they are detected.

---

[1] http://www.univie.ac.at/webda/

	r' & i'	i'
Matched	45	26
Unmatched	6	23

Table 5.1: Number of overdensities matched and unmatched to cluster coordinates from crossmatches with SIMBAD and WEBDA databases, which appear in either both r'-band and i'-band density maps, or in the i'-band map only.

A certain fraction of these unmatched overdensities are likely to be non-clusters - asterisms potentially caused by variations in extinction on small scales. However the larger extinction values observed in the r'-band also means that redder clusters are more likely missed compared to the i'-band.

Table 5.2 shows the median parameters for the overdensities unmatched to any SIMBAD clusters. Shown are the galactic coordinates, amplitude of overdensity and the background density level, along with semi-major and -minor axes of the fitted overdensities. Figure 5.16 shows six such overdensities.

5.3 Density map files

The density maps generated as part of this work, shown at $1' \times 1'$ resolution in Figures 5.17 - 5.20, are attached along with this thesis in Flexible Image Transport System (FITS) format. At each available resolution the r'- and i'-band maps are stored as a multi-extension FITS file, with extensions corresponding to different faint limiting magnitudes. Table 5.3 gives the extension that corresponds to each faint limiting magnitude for both band maps.

The maps are supplied with cells of side $1'$, $2'$, $10'$, $30'$, and $60'$. The files are named "*dmap-b-x.fits*", where b denotes the band and x denotes the resolution.

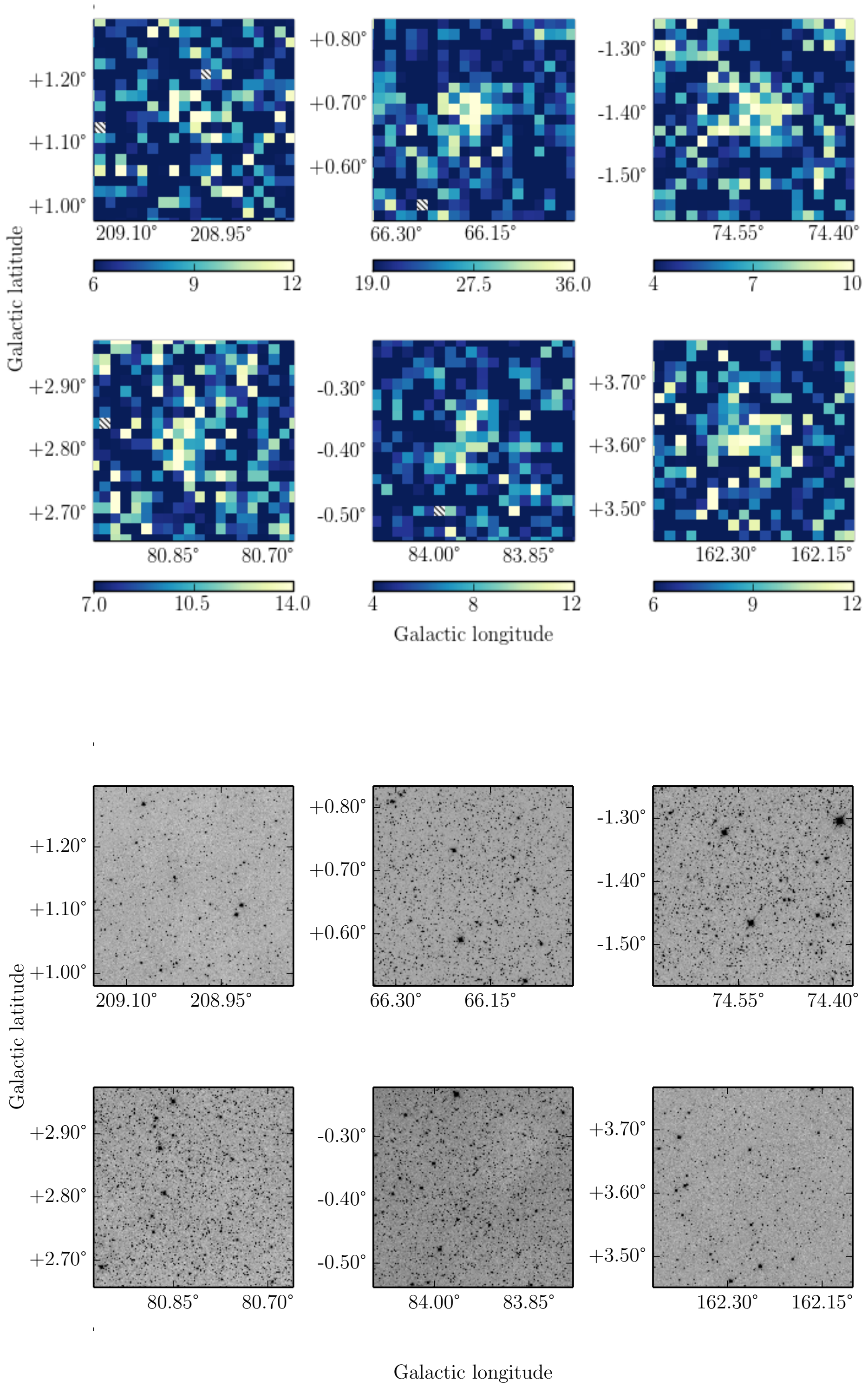


Figure 5.16: **Upper:** Example of six overdensities found in the $1' \times 1'$ i'-band density map, limited to $i' < 18$, with no corresponding cluster within $5'$ in the SIMBAD database. Regions not covered by the density map are hatched to distinguish between no data and high counts. **Lower:** 2MASS K_s cutouts for the same regions.

ℓ	b	A	A_{err}	B	B_{err}	σ_1	σ_{1err}	σ_2	σ_{2err}
42.24	-3.10	10.78	0.66	18.29	0.45	2.52	0.80	1.98	0.46
43.89	0.91	9.70	1.33	20.38	1.68	3.65	2.29	2.29	1.31
44.75	1.52	14.21	2.02	28.15	2.64	2.63	2.80	1.31	2.10
57.23	-1.40	10.07	0.13	17.62	0.30	2.97	0.09	2.42	0.27
89.06	2.57	7.90	0.41	9.17	0.17	3.00	0.54	1.51	0.10
97.58	1.64	6.52	0.28	7.19	0.30	3.34	0.27	1.44	0.16
32.90	1.37	7.16	0.34	8.46	0.52	3.32	0.16	1.58	0.28
35.12	-2.06	5.85	0.68	5.90	0.32	2.72	0.56	2.27	0.46
35.34	0.90	6.09	0.19	2.41	0.13	3.55	0.10	1.33	0.21
41.08	0.94	6.27	0.50	5.25	0.28	3.11	0.79	1.67	0.22
42.81	2.15	8.65	0.65	16.85	0.25	3.12	1.29	2.29	0.12
55.38	-0.40	8.44	0.55	17.17	0.76	3.73	0.42	3.34	0.50
61.86	-1.04	6.89	0.22	10.21	0.64	3.37	0.64	2.96	0.57
65.08	-0.15	9.71	0.41	17.87	0.88	3.49	1.00	3.02	0.69
66.18	0.68	15.77	0.39	19.62	0.50	2.34	0.25	1.59	0.11
66.36	-0.64	7.21	0.81	10.87	0.68	4.00	1.13	2.53	0.94
68.17	-1.37	12.58	1.68	12.94	0.58	3.52	1.93	1.61	0.22
68.30	0.24	9.31	0.33	19.62	0.59	3.64	0.73	2.68	0.62
68.73	-3.32	7.30	0.83	5.89	1.29	3.55	1.86	2.49	1.59
74.53	-1.41	5.30	0.11	4.26	0.23	3.10	0.15	1.85	0.16
77.62	2.12	5.71	0.11	5.37	1.23	3.03	2.38	1.68	2.23
80.82	2.82	6.04	0.03	7.84	0.10	2.61	0.07	1.23	0.05
83.94	-0.38	7.27	0.02	4.40	0.06	2.56	0.03	1.03	0.03
101.65	1.13	7.98	0.86	10.78	1.13	2.86	1.88	1.46	1.31
121.58	-1.27	6.88	1.16	8.57	0.78	2.90	1.61	1.57	1.35
162.26	3.61	5.65	0.22	6.04	0.11	2.71	0.15	1.59	0.17
164.71	-0.57	5.42	0.28	6.46	0.34	2.80	0.34	1.84	0.22
194.25	2.47	7.16	0.02	5.68	0.09	1.78	0.05	1.16	0.03
208.99	1.14	5.33	0.13	6.02	0.16	2.52	0.11	1.21	0.04

Table 5.2: Median parameters of overdensities found in i'-band density map (limited to $i' < 18$), which do not appear in literature searches for known clusters.The first six entries are located within $5'$ of an overdensity detected in the $r' < 19$ density map. **Column definitions:** A - median amplitude of overdensity, B - median background counts at overdensity coordinates, σ_1,σ_1 - median standard deviations along the semi-major and semi-minor axes of the overdensity. All errors quoted are 1σ.

	17.0	17.5	18.0	18.5	19.0	19.5	20.0
r'	-	-	1	2	3	4	5
i'	1	2	3	4	5	-	-

Table 5.3: The FITS extension corresponding to faint limiting magnitudes for both r'- and i'- density maps.

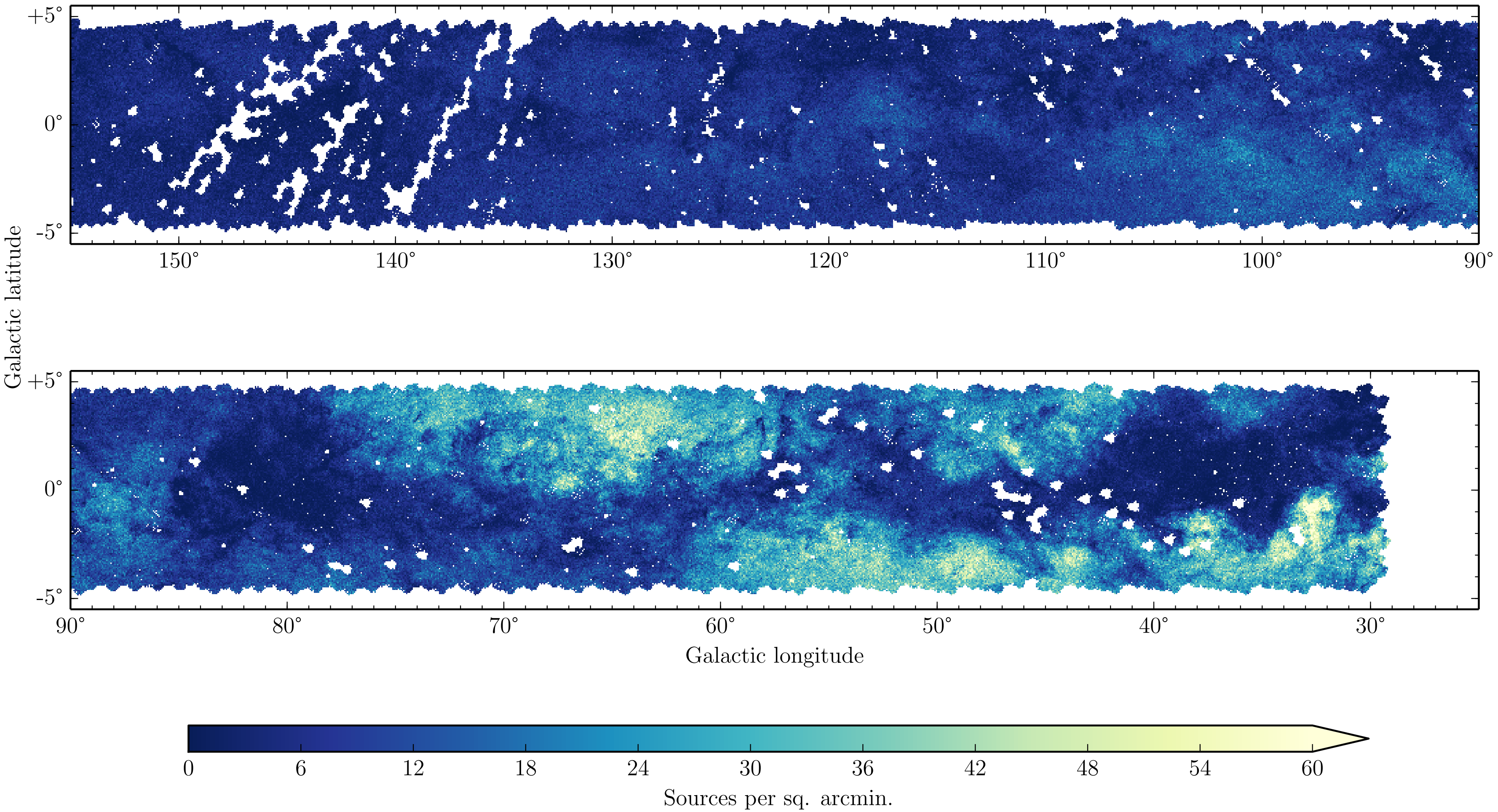


Figure 5.17: r'-band density map of sources down to 19th magnitude, at $1' \times 1'$ resolution, for $\ell < 155°$.

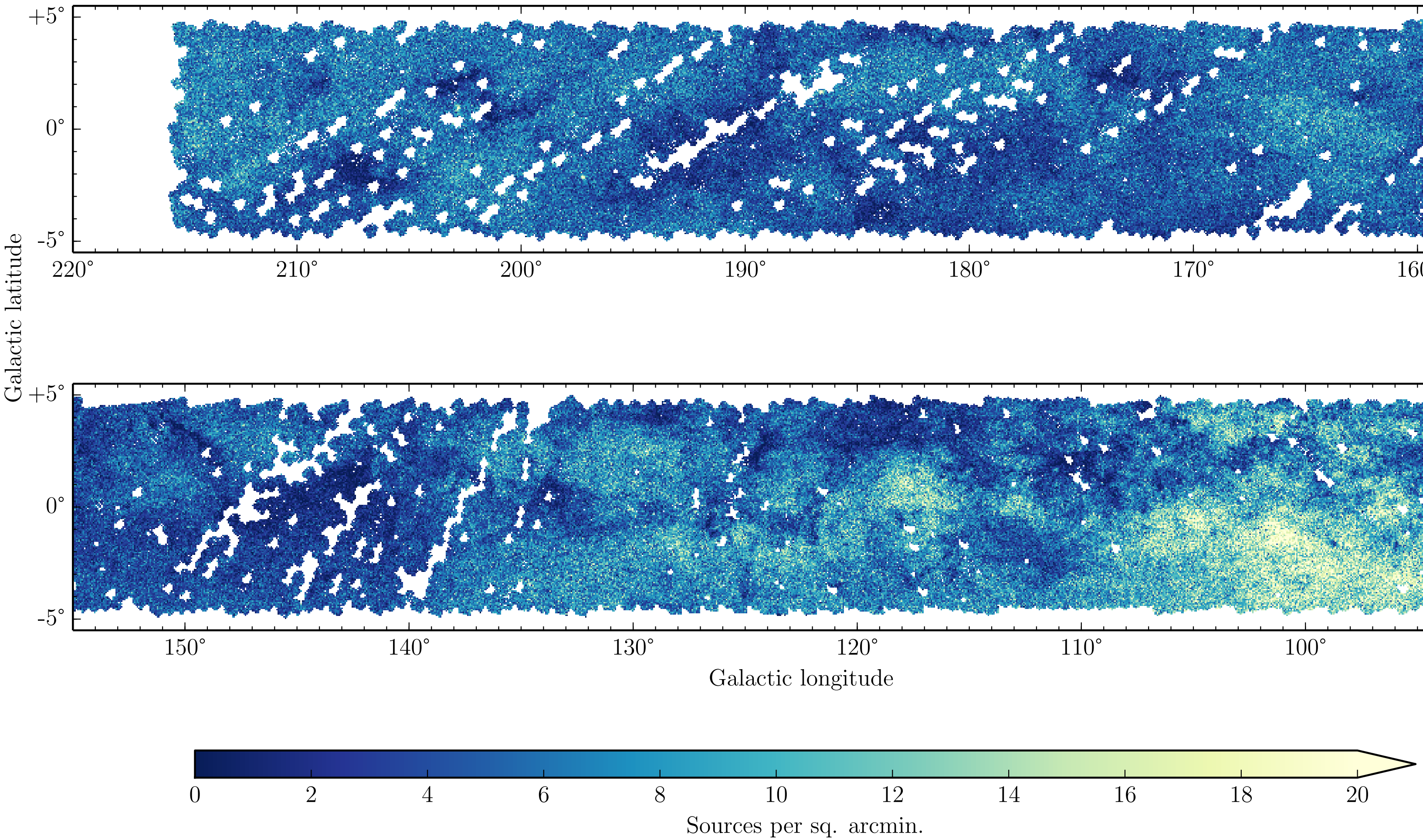


Figure 5.18: r'-band density map of sources down to 19th magnitude, at $1' \times 1'$ resolution, for $\ell > 90°$.

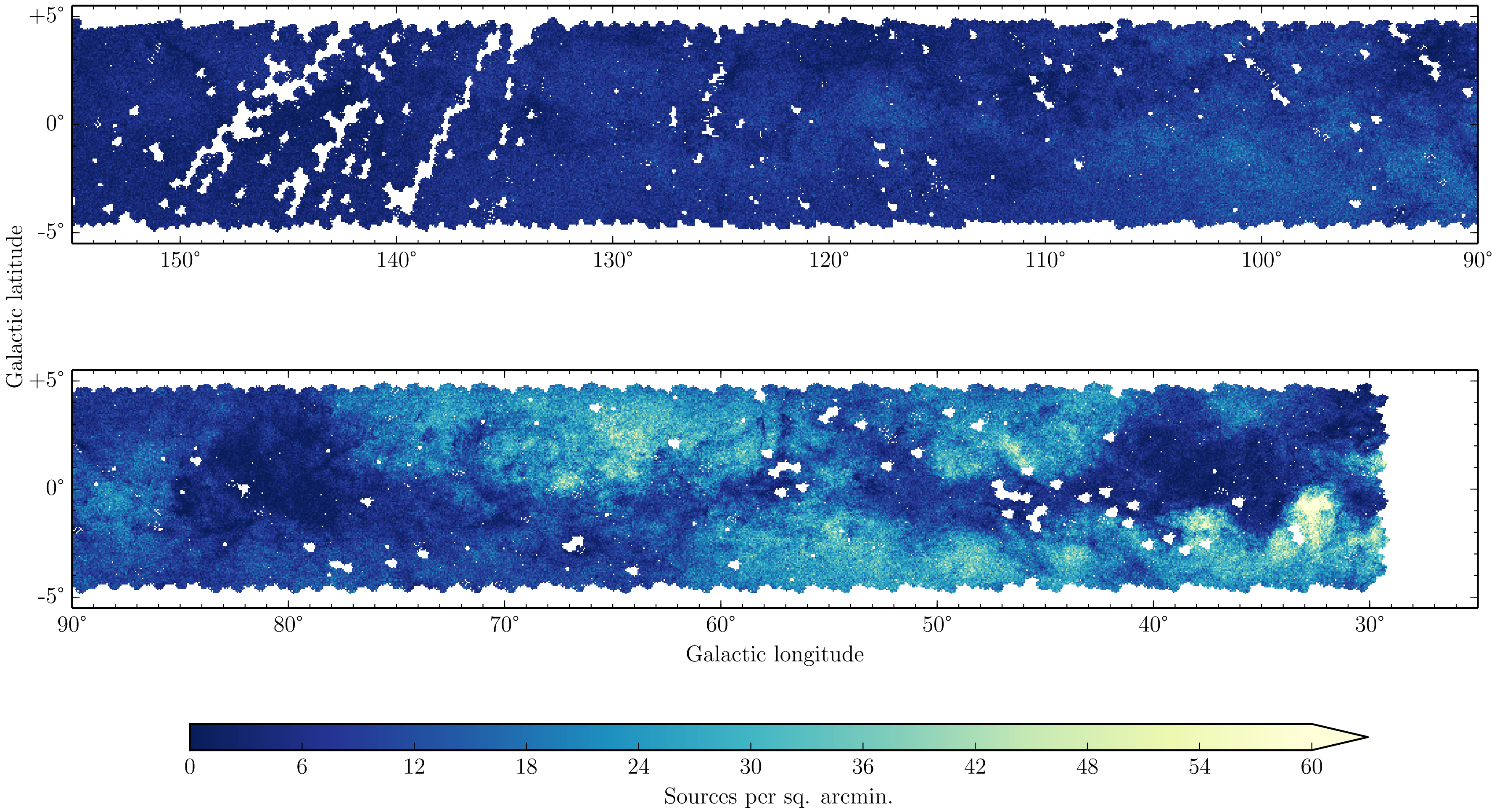


Figure 5.19: i'-band density map of sources down to 18th magnitude, at $1' \times 1'$ resolution, for $\ell < 155°$.

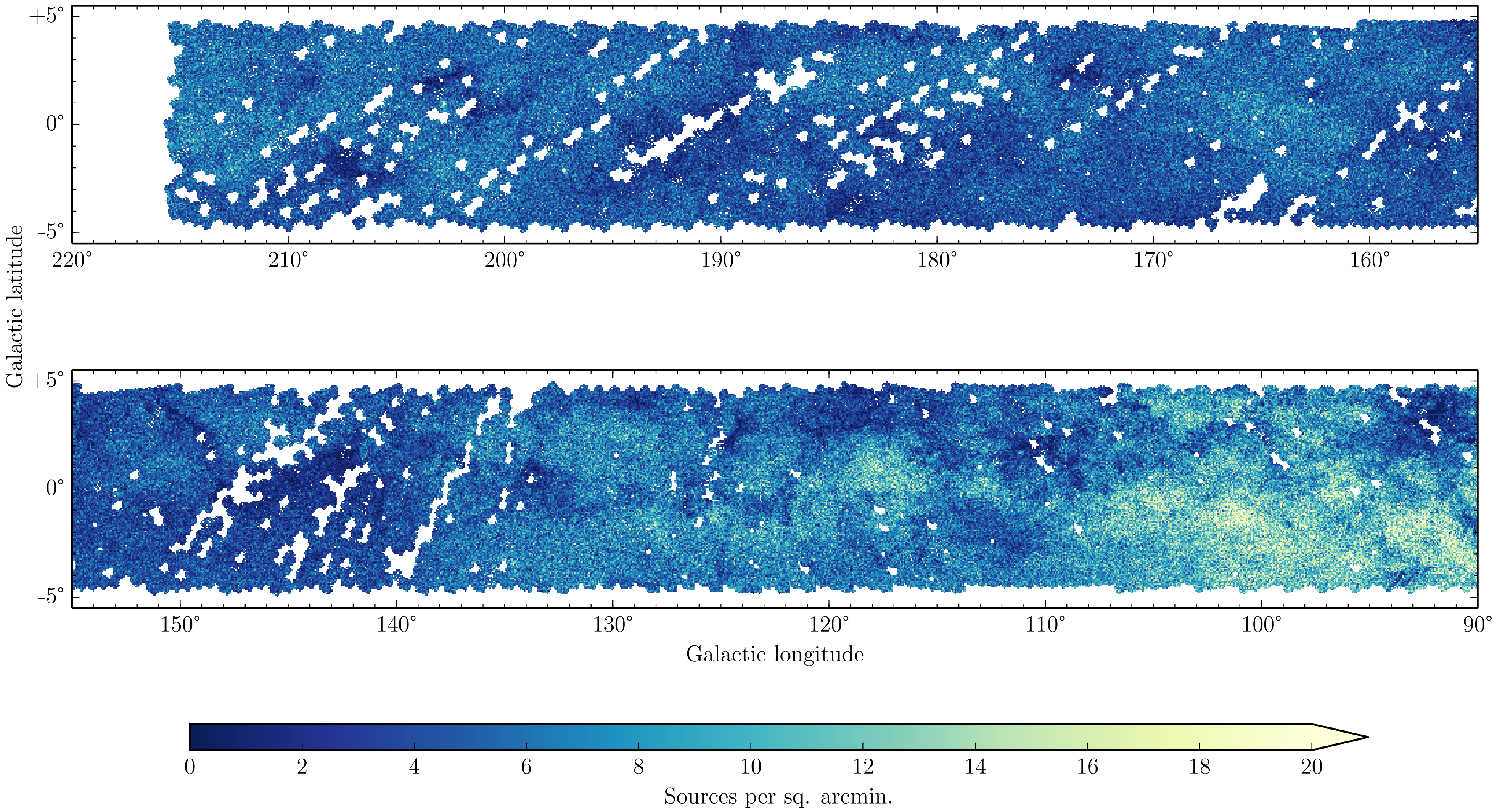


Figure 5.20: i'-band density map of sources down to 18th magnitude, at $1' \times 1'$ resolution, for $\ell > 90°$.

Chapter 6

Comparisons with the Besançon Model

As described in Chapter 1, realistic models of the Milky Way are in demand as aids to studies of Galactic structure: the Besançon model of Galactic population synthesis (Robin et al., 2003) has been used extensively (Momany et al., 2004; Sale et al., 2010; Williams et al., 2011).

In Section 6.1 the Milky Way as simulated in anticipation of Gaia data is discussed (the GUMS model). In order to make a comparison with this model, it is necessary to map the Gaia magnitude system onto the IPHAS system. To achieve this, transforms between the Gaia and IPHAS passbands are derived.

Attention then turns in Section 6.2 to a set of comparisons that explores the impact of different prescriptions for Galactic extinction on stellar number densities as predicted by the widely used implementation of Robin et al. (2003).

6.1 The Besançon model as prepared for Gaia

6.1.1 The Gaia Universe Model Snapshot (GUMS)

In anticipation of Gaia data, Robin et al. (2012a) simulated the contents of the Milky Way (and the Magellanic Clouds) in great detail, using slightly modified version of the Robin et al. (2003) Besançon model, creating a catalogue they named the Gaia Universe Model Snapshot (GUMS). This approach involves modelling populations separated in age, making use of a series of density laws for each component (thin/thick discs, spheroid and bulge) of the Milky Way, extrapolating from stellar densities in the solar neighbourhood. Sources are assigned key parameters (M_V, T_{eff}, and age) according to the distribution of these parameters as calculated using an assumed IMF and SFR history passed through the evolutionary model of Haywood et al. (1997b).

The extinction model of Drimmel et al. (2003) is used to redden GUMS sources, which provides the mean extinction per $0.35° \times 0.35°$ pixel. The same parameter warp and flare as described in Robin et al. (2003) are applied to the thin disc populations. Robin et al. (2012a) introduces a two-arm spiral structure, which generate enhancements in the number density of stars. The resulting catalogue is designed to be analysed as part of the Gaia Object Generator (GOG), a framework for simulating Gaia data (Luri et al., 2014) taking into account simulated errors relating to known instrument performance.

The goal of testing the performance of the Besançon model in reproducing the Galactic plane as observed by IPHAS requires that the GUMS be used for comparison, rather than the simulated Gaia output of GOG. A large number of parameters are available for each simulated source: stellar parameters (e.g. age, $\log g$, $\mathrm{T_{eff}}$), information on multiple systems, and even information on simulated exoplanets. For comparison with IPHAS stellar densities, the only parameters required are position and magnitude, allowing a much smaller subset of the whole simulation to be downloaded (see Section 6.1.3).

6.1.2 The Gaia photometric system and IPHAS

Gaia will provide four magnitudes per source, corresponding to the passbands of the instruments aboard Gaia (Jordi et al., 2010). The primary magnitude, denoted by G, is the result of a (very) broad unfiltered passband defined by mirror coatings and CCD quantum efficiency (Jordi et al., 2006), covering a wavelength range of 330-1050 nm. Two further broad-band magnitudes, overlapping the blue and red halves of the G passband, are measured by Gaia using two low resolution prism spectrophotometers; These are the G_{BP} and G_{RP} magnitudes in the ranges 350-680 nm and 640-1060 nm respectively. A high resolution spectrograph will observe the CaII triplet at 847-874 nm, providing the magnitude G_{RVS} and measuring radial velocities for stars as faint as G_{RVS}=17.

Figure 6.1 compares the normalised Gaia passbands to those of IPHAS, where it can be seen that G_{BP} extends to much shorter wavelengths than IPHAS r' while G_{RP} spans the coverage of i' to both longer and shorter wavelengths.

Without the CCD response curves of each of the Gaia instruments, a full treatment of transformations between Gaia and IPHAS photometric systems cannot be performed. However Jordi et al. (2010) provide well defined transformations between Gaia and SDSS photometric systems. The transformations with the smallest error involve using two SDSS colours. Here the transformations making use of $(g-r)$ (rather than $(g-i)$) and $(r-i)$ are used (both sets of transformations have the same scatter):

$$\begin{aligned} G - g_{\mathrm{SDSS}} = -0.0992 &- 0.5749(r-i)_{\mathrm{SDSS}} - 0.2427(r-i)^2_{\mathrm{SDSS}} + 0.0365(r-i)^3_{\mathrm{SDSS}} \\ &- 0.5277(g-r)_{\mathrm{SDSS}} - 0.1158(g-r)^2_{\mathrm{SDSS}} + 0.0086(g-r)^3_{\mathrm{SDSS}} \\ &- 0.0337(g-r)_{\mathrm{SDSS}}(r-i)_{\mathrm{SDSS}} \end{aligned} \tag{6.1}$$

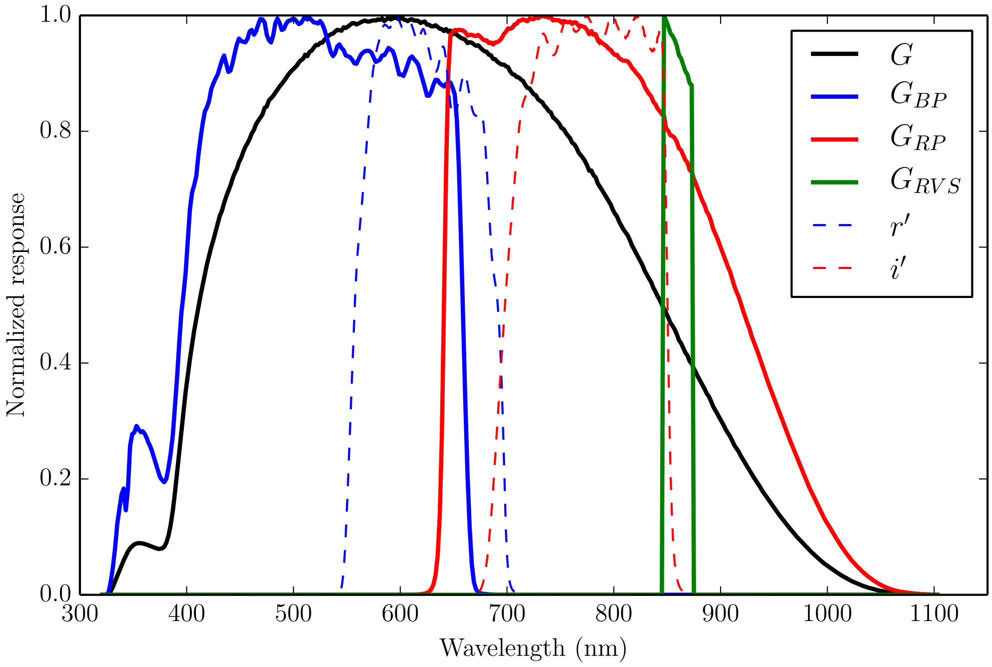


Figure 6.1: Response of the four Gaia passbands alongside IPHAS r' & i'.

$$\begin{aligned} g_{\mathrm{SDSS}} - G_{BP} = {} & -0.0438 - 0.0176(r-i)_{\mathrm{SDSS}} + 0.0207(r-i)^2_{\mathrm{SDSS}} + 0.0001(r-i)^3_{\mathrm{SDSS}} \\ & + 0.2755(g-r)_{\mathrm{SDSS}} + 0.1356(g-r)^2_{\mathrm{SDSS}} - 0.0071(g-r)^3_{\mathrm{SDSS}} \\ & - 0.0114(g-r)_{\mathrm{SDSS}}(r-i)_{\mathrm{SDSS}} \end{aligned} \tag{6.2}$$

$$\begin{aligned} g_{\mathrm{SDSS}} - G_{RP} = {} & 0.3725 + 1.3600(r-i)_{\mathrm{SDSS}} + 0.0434(r-i)^2_{\mathrm{SDSS}} - 0.0206(r-i)^3_{\mathrm{SDSS}} \\ & + 0.9022(g-r)_{\mathrm{SDSS}} - 0.0199(g-r)^2_{\mathrm{SDSS}} + 0.0062(g-r)^3_{\mathrm{SDSS}} \\ & + 0.0394(g-r)_{\mathrm{SDSS}}(r-i)_{\mathrm{SDSS}} \end{aligned} \tag{6.3}$$

Combined with the transformations between IPHAS and SDSS derived in Section 3.3, these can be used to predict how stars with given IPHAS colours will appear as observed by Gaia. Combining Equations 3.5 and 3.6 yields

$$(r-i)_{\mathrm{SDSS}} = 0.951(r'-i')_{\mathrm{IPHAS}} - 0.214. \tag{6.4}$$

In order to convert the WFC $(g'-r')$ colour to SDSS, the colour equations of González-Solares et al. (2011) were combined to obtain the transformations

$$(g-r)_{\mathrm{SDSS}} = 1.166(g'-r')_{\mathrm{IPHAS}} - 0.238 \tag{6.5}$$

$$g_{\mathrm{SDSS}} = g'_{\mathrm{IPHAS}} - 0.092 + 0.159(g'-r')_{\mathrm{IPHAS}}. \tag{6.6}$$

IPHAS ($r' - i'$) and UVEX ($g' - r'$) colours were taken from Drew et al. (2005) and Groot et al. (2009) for a range of extinctions ($E(B - V)$ values of 0, 1, 2, & 3) and class V stellar types. These were transformed into Gaia G, G_{BP} and G_{RP} magnitudes to identify which filters (in both IPHAS and Gaia systems) would together provide the best comparison between survey and GUMS output.

Figure 6.2 shows the result of transforming the colours of a range of dwarfs with varying E(B-V) excesses while keeping their r' and i' magnitudes fixed. Unsurprisingly, the largest spread occurs when converting r' to G_{RP} and i' to G_{BP} - the effective wavelength changes in both cases are very large. The next largest spread occurs when converting from r' to G. Here the effective wavelengths of the filters are similar. This is why unreddened objects up to type ≈K0V show very little change in $G-r'$. However, for redder sources, the offset becomes more pronounced, owing to the G filter extending to ≈300 nm redder wavelengths than r'.

The smallest ($\lesssim$ 1 mag) spreads in Figure 6.2 are seen in the $G_{BP} - r'$, $G_{RP} - i'$ and $G - i'$ shifts. This is unsurprising in the case of $G_{RP} - i'$, where the i' filter profile sits to the centre of that of the G_{RP} passband. The spread in $G_{BP} - r'$ shifts reduces for redder sources, whose fluxes will mostly fall in the same wavelength range for both filters.

The G and i' passbands were chosen to set limits on the GUMS and IPHAS datasets; the $G-i'$ shows the smallest shift - ≈0.4 mag - as seen in Figure 6.2. By applying an appropriate shift to either set of magnitudes, roughly the same stellar types can be captured. Determining this shift is discussed in 6.1.4.

6.1.3 Querying

The GUMS catalogue simulating the stellar content of the Milky Way contains 2,143,475,885 sources, each with more than forty parameters, the majority being stellar parameters of no use in comparing spatial number densities of sources.

In order to retrieve the relevant rows in the catalogue for comparing with the IPHAS-generated density map, the TAP service provided by the German Astrophysical Virtual Observatory was queried using an ADQL query as detailed in Listing 6.1. The GUMS catalogue lists positions in equatorial coordinates, requiring a conversion to Galactic coordinates in order to cut out the Galactic Plane. The expression included in Listing 6.1 requires that $-6° < b < 6°$. The query returned 864,477,040 sources within 6° of the Galactic equator.

Listing 6.1: ADQL query used to retrieve GUMS data covering the Galactic Plane

```
SELECT
    alpha, delta, magg, maggrp, maggbp, ag, av, rv
    FROM gums.mw
    WHERE
        gums.mw.magg BETWEEN 12 AND 22
        AND
        abs(degrees(asin(sin(radians(27.12825))*sin(radians(gums.mw.
```

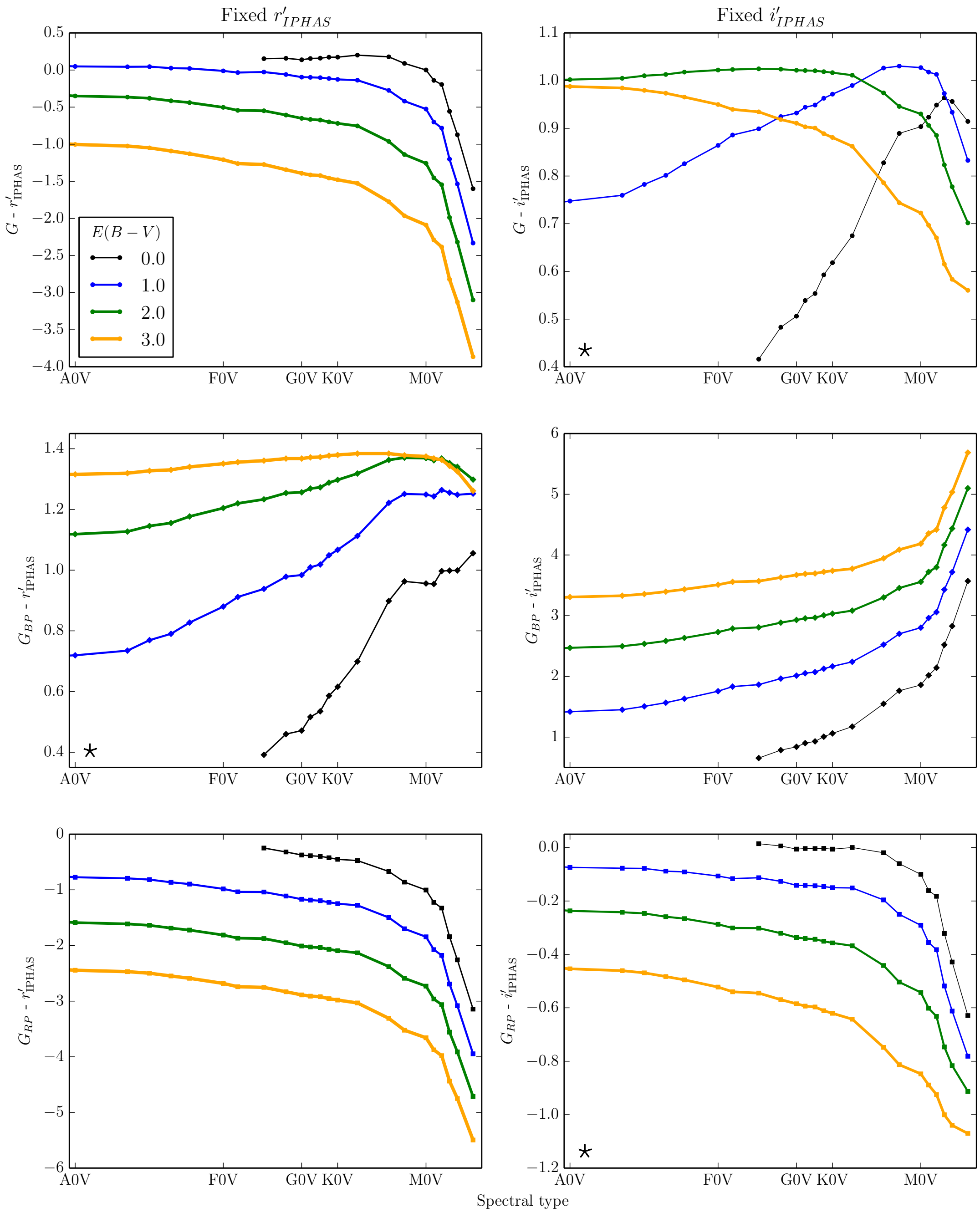


Figure 6.2: Conversion into Gaia G, G_{BP} & G_{RP} magnitudes for IPHAS sources fixed at r'=19 (**left**) and i'=18 (**right**). Colours for a range of spectral types and $E(B-V)$ values were taken from Drew et al. (2005) & Groot et al. (2009), using transformations between SDSS colours and Gaia magnitudes taken from Jordi et al. (2010). IPHAS colours were transformed into SDSS using Equations 6.4 & 6.5 (the latter taken from González-Solares et al. (2011)). Note that the vertical scale in the panels marked by the asterisks is much narrower than the other panels - these Gaia/IPHAS passband combinations will permit more useful comparisons. The $E(B-V) = 0$ main sequence begins at F5V - earlier types will most likely be saturated in IPHAS photometry.

```
↪ delta))+cos(radians(27.12825))*cos(radians(gums.mw.
↪ delta))*cos(radians(gums.mw.alpha-192.85948)))))<6
```

6.1.4 Comparisons

As described in Section 6.1.2, Gaia G and IPHAS i' were chosen as the filters on which to make magnitude cuts for generating density maps. Using comparisons plotted in Figure 6.2,

$$G - i' = 0.8 \tag{6.7}$$

was chosen as the offset to compare roughly equivalent populations. Figure 6.3 shows magnitude tracks for the main sequence suffering from increasing extinction ($E(B-V)$ of 0, 1 and 2). Sources were placed at distances corresponding to the amount of extinction experienced, assuming a diffuse ISM with an exponential profile and a local normalisation of $A_V = 1.0$ mag kpc^{-1}. This normalisation was chosen to be slightly higher than the default 0.7 mag kpc^{-1} used by the Besançon model for intermediate and high latitude sightlines, to account for higher typical extinctions in the Galactic Plane. Five sightlines were plotted at $E(B-V)$ of 1 and 2: $\ell = 28°$, $45°$, $90°$, $135°$, and $180°$ (a distance of 500 pc was adopted for zero extinction).

The default i'-band density map faint limit of 18th magnitude was chosen for the comparison between IPHAS and GUMS. Applying Equation 6.7, limits on GUMS of $12.8 < G < 18.8$ were imposed. These ranges are presented as green shaded regions in Figure 6.3 - it can be seen that these regions generally capture the same populations in IPHAS and GUMS data. As such the objects falling into the green bands can be sensibly compared, testing the ability of GUMS to simulate the Milky Way as seen by IPHAS.

6.1.5 Limitations of GUMS

Figures 6.4-6.6 show the comparisons between the IPHAS i'-band density map, limited to 18th magnitude, with the stellar densities predicted by GUMS, limited to $G < 18.8$.

Upon mapping the density of GUMS sources, an obvious issue presented itself. In the range $30° < \ell < 60°$ a band is evident across the Galactic equator, at latitudes $|b| < 1°$, showing less variation than at higher/lower Galactic latitudes (see Figure 6.4). This band appears distinct from regions of low extinction caused by clouds in the Galactic plane, which should be visible in the region covered by this band. The band flares up to encompass the region $|b| \lesssim 6°$ at $\ell > 70°$. The band cannot be explained by the application of the Drimmel et al. (2003) extinction map adopted by Robin et al. (2012a), as the band remains evident when the provided A_G per simulated source is subtracted from apparent G magnitudes.

Ignoring the region covered by the band in the inner $\ell < 80°$ Galaxy, the regions further from the midplane can be compared. The structure visible in the IPHAS density map is well-reproduced by GUMS source densities. Both dark clouds and regions of higher densities

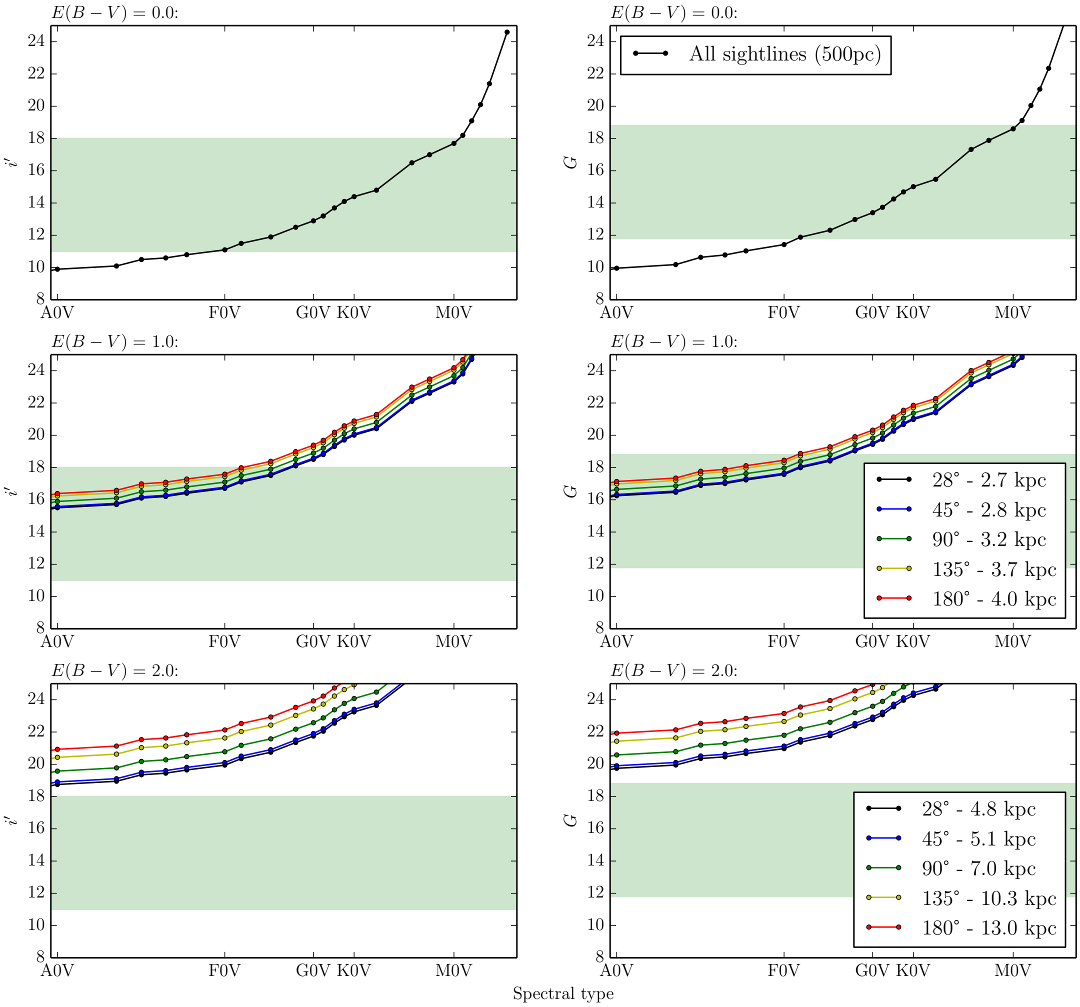


Figure 6.3: IPHAS i' (**left**) and Gaia G (**right**) magnitudes for a range of stellar types. **Upper:** Magnitudes for unreddened sources at 500 pc. **Middle:** Magnitudes for sources with $E(B-V)=1$, placed at a distance corresponding to this colour excess for five sightlines (assuming an exponential distribution of extinguishing ISM, as described in the text). **Lower:** Same as middle panels, only for $E(B-V)=2.0$. **Green band:** Limits placed on IPHAS and GUMS data as discussed in Section 6.1.4, where a shift of 0.8 mag is applied between IPHAS and GUMS limits. The stellar types falling into these bands are very similar for both surveys, indicating that comparing the two datasets with these limits makes sense.

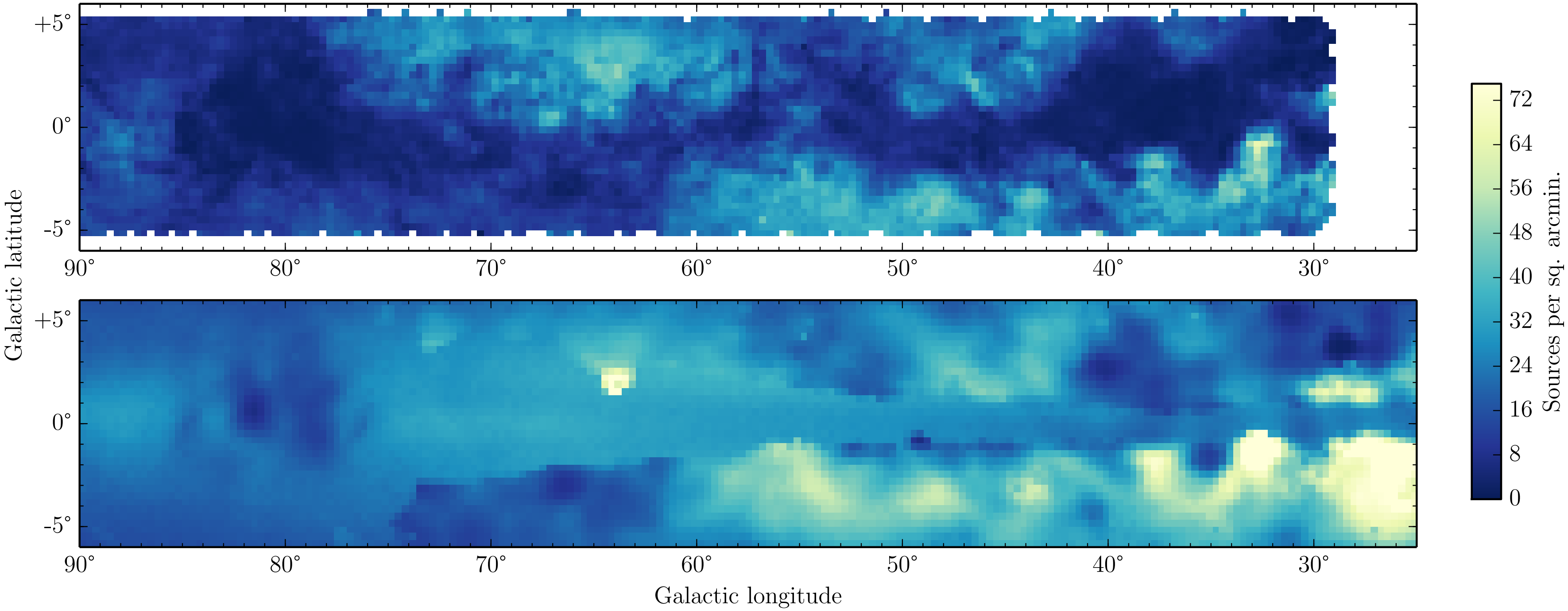


Figure 6.4: Comparison of density maps: IPHAS single-band i' with a faint limit of 18th magnitude (**upper**) and GUMS sources where $G < 18.8$ (**lower**). Both are displayed at $20' \times 20'$ resolution, in the region $25^\circ < \ell < 90^\circ$.

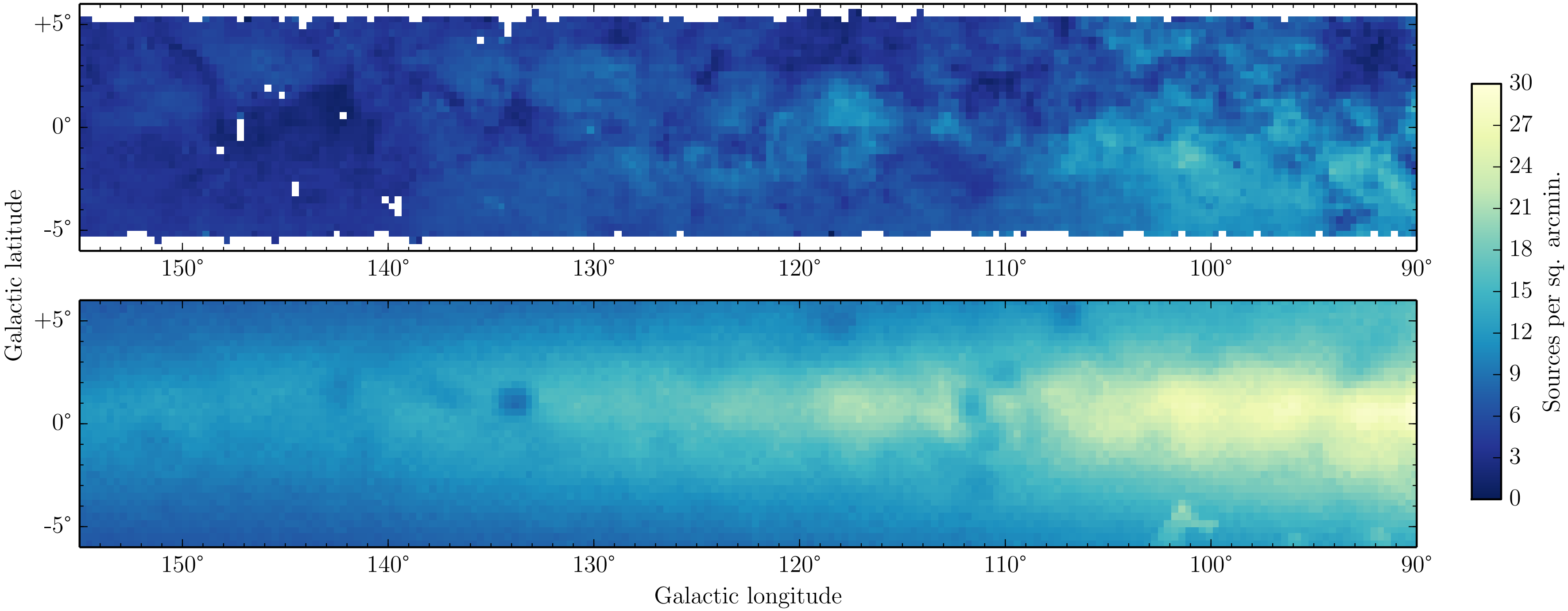


Figure 6.5: Same as Figure 6.4, but for $90° < \ell < 155°$.

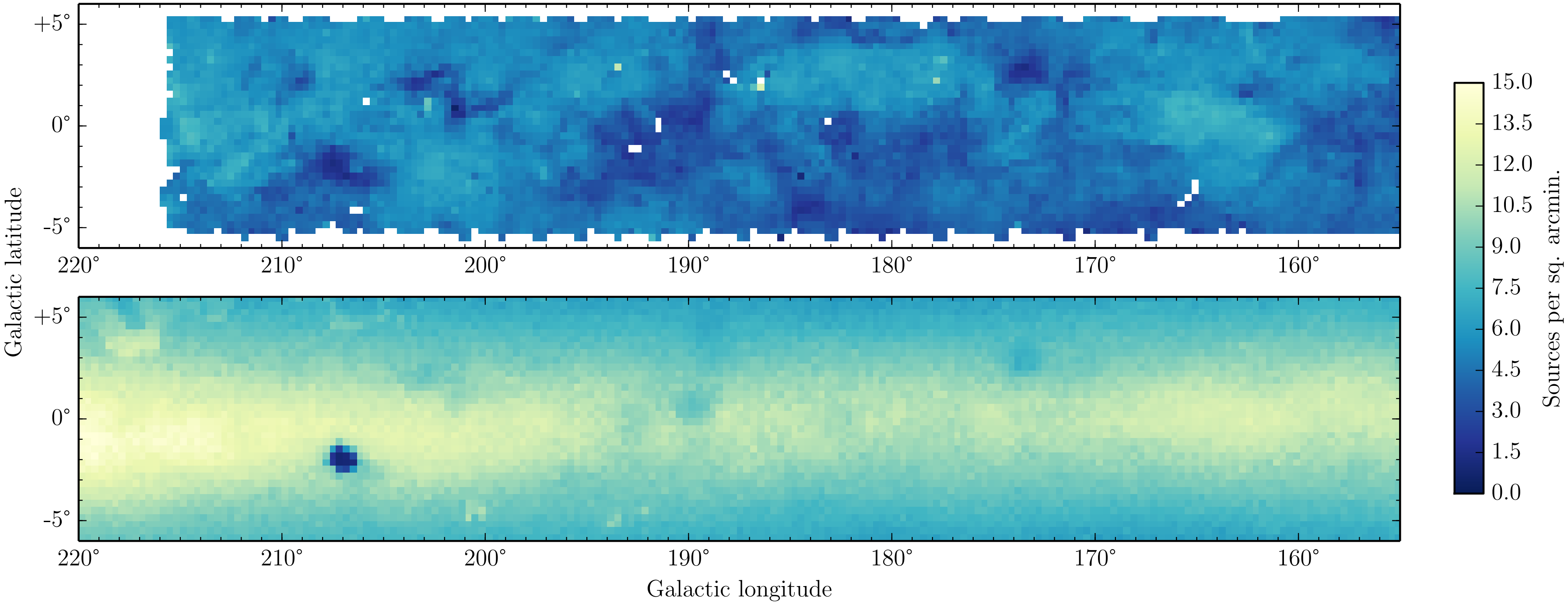


Figure 6.6: Same as Figure 6.4, but for $155° < \ell < 220°$.

are predicted: at $\ell \lesssim 60°$ the correspondence between the two maps is striking, while at positive latitudes and $60° \lesssim \ell \lesssim 70°$, the highest densities seem to occur at higher values and slightly closer to the plane in the GUMS map relative to the IPHAS density map. This could potentially be explained by the extinction map used. Figure 10 of Drimmel et al. (2003) shows the extinction curves along lines-of-sight in the inner disc - the sightline closest to the highest GUMS stellar density in this region, (60°, 0°), shows that the extinction curve underestimates the extinction to OB stars closer than 3 kpc. A steeper rise in extinction at small distances in this region might go some way to bringing the GUMS prediction closer in line with IPHAS.

The extinction map used is not the only possibility that could account for mismatches between the maps; as $\ell \rightarrow 90°$, the outer Galaxy contributes increasingly to the line-of-sight. As discussed in Chapter 1, the parameters describing the outer disc are not well-constrained (more so than the inner disc). The performance of models in the outer disc are highly dependent on its chosen paramaterisation; consequently models should be expected to be less capable at reproducing sightlines in this regime.

The expectation of the model faring less well in the outer disc is borne out in Figures 6.5 and 6.6. While GUMS counts towards the inner Galaxy may appear slightly higher than seen by IPHAS, at $\ell > 90°$ the GUMS counts appear significantly higher than IPHAS counts, particularly in the midplane. At these longitudes the smooth exponential distribution of the simulated disc is highly evident in GUMS predictions, with dark regions reproduced less realistically than at $\ell < 70°$. The exceptions are the regions of high extinction corresponding to the star forming region W3 ($\approx 140°$, $+1°$) and the Rosette Nebula ($\approx 206°$, $-2°$).

Due to the presence of the band towards the inner Galaxy, quantatively comparing cross-sections through the plane could not be sensibly performed. For this reason it was decided to revert to the original Besançon model predictions for quantitative comparisons against the density maps.

6.2 The 2003 Besançon model

In order to compare the Robin et al. (2003) Besançon model with the IPHAS density map, extinction needs to be added to the sources in the generated synthetic catalogues. The default extinction model, which assumes an exponential disc of obscuring material, is highly simplistic and needs the local normalisation (in mag kpc^{-1}) as an input. Rather than rely solely on the default model, several extinction prescriptions were applied to the returned catalogues, to see how they fare.

6.2.1 Querying the Besançon model

The Besançon model as presented in Robin et al. (2003) is made available for use through a set of web pages[1], which can be used to generate synthetic catalogues based on a series of

---

[1] `http://model.obs-besancon.fr/`

Parameter	Value
FOV	Large field
Coordinate system	Galactic
Steps	0.2 deg
Diffuse extinction	0 mag kpc^{-1}
Clouds	None
Absolute magnitudes	$-7 < M_v < 20$
Earliest type	O0
Latest type	DA5
Luminosity classes	I - V & WD & T Tauri
Populations	1 - 10
Apparent mag. passband	r'
Intervals (u^*, g', i', z')	$-99 < m < 99$
r' interval	$8 < r' < 25$

Table 6.1: Settings used to obtain synthetic catalogues via the Besançon web interface.

criteria. The region over which objects are simulated can be specified in both heliocentric distance and Galactic coordinates, extinction parameters can be modified, spectral types and absolute magnitude ranges can be selected, and limits based on apparent magnitudes imposed. Two output photometric systems are available: Johnson-Cousins and CFHTLS-Megacam. Fortunately the Megacam filters were designed to closely replicate the SDSS filter set, meaning catalogues output by the Besançon website can be sensibly compared with IPHAS. The transformations derived in Chapter 3 were applied to bring the returned synthetic magnitudes into the IPHAS photometric system.

The web interface places limits on the region of sky that can be simulated in a single simulation, preventing the filesizes of resulting catalogues from becoming too large. In order to simulate a significant ($>$ a few sq. deg.) region of the Galactic Plane, a script was used to make repeated small ($\ll$1 sq. deg.) requests, moving across the area of interest, thereby building up a larger catalogue. The diffuse extinction parameter was set to zero, allowing custom distributions to be applied later (see Section 6.2.4). Querying small regions of the Galactic Plane produced results so quickly (in a minute, roughly) that all populations of sources were included, and the default absolute magnitude range ($-7 < M < 20$) retained. An apparent magnitude range of $8 < r' < 25$ was imposed to exclude any objects falling far outside IPHAS detection limits. Options for including kinematics are also available but were not used here.

6.2.2 IPHAS densities

Regions of 20 sq. deg. were chosen for the comparison based on the coverage of the density map generated in Chapter 5; areas at $\ell \approx 30°$, $90°$, and $175°$ were simulated via the Besançon web interface, and the stellar densities of the corresponding regions in the IPHAS i'-band density maps were extracted.

Figure 6.7 shows the variation in stellar densities with Galactic latitude in these regions,

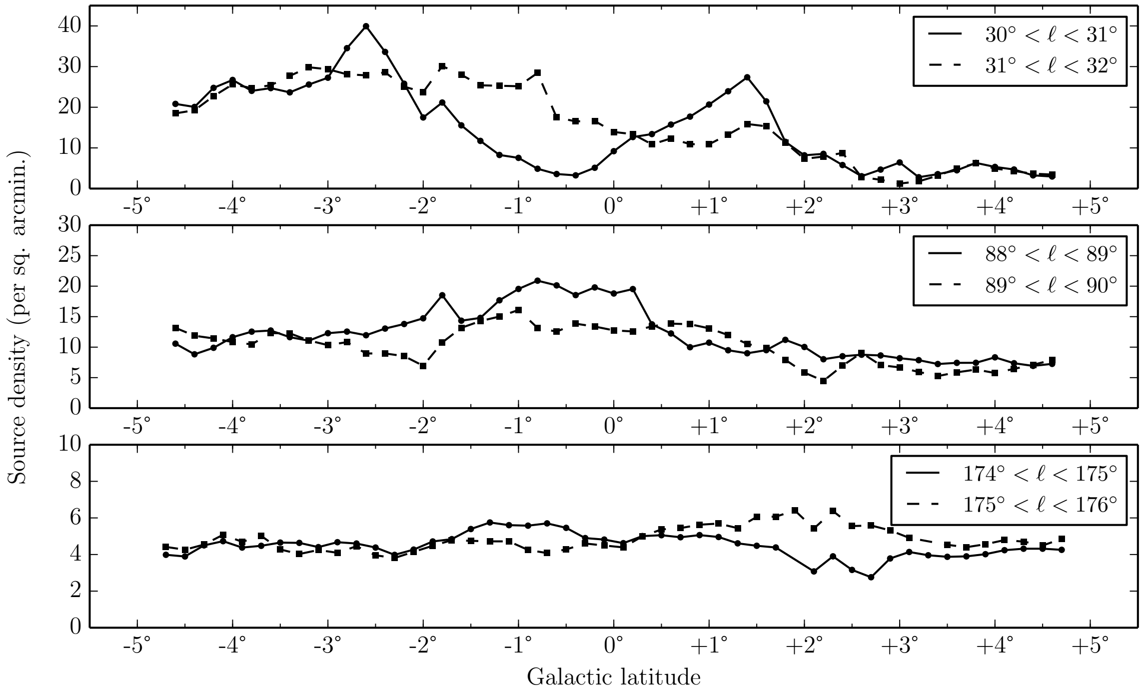


Figure 6.7: Stellar density profiles in Galactic latitude, averaged over 1° wide strips in Galactic longitude, as given by the $20' \times 20'$ resolution IPHAS i'-band density map. Each panel shows two profiles for each sightline to be compared against star counts predicted by the Besançon model.

as retrieved from the $20' \times 20'$ resolution IPHAS i'-band density map. Each sightline was split into two 1°-wide strips in Galactic longitude to prevent the blurring of structure in the disc; variation on these scales can be seen by comparing the densities in the two strips. Figure 6.8 shows the sightlines and how they relate to spiral arms according to the model of Vallée (2008).

The inner Galaxy sightline shows the largest variation with latitude and the largest variation between the two separated 1°-wide strips. The Aquila region ($\ell \lesssim 45°$) contains both the highest and lowest valued cells of the entire IPHAS density map. A dark cloud at ($\approx 30°$, $-1°$) is responsible for the reduction in stellar density (of the amplitude $\approx 10 - 20$ sources per sq. arcmin.) in the $30° < \ell < 31°$ profile relative to the $31° < \ell < 32°$ profile. The Aquila Rift is responsible for the drop-off in stellar density at higher latitudes in both strips. Taking the spiral arm positions of Vallée (2008), it can be seen (in Figure 6.8) that this sightline will pass through the Sagittarius-Carina arm twice by a distance of 10 kpc, and passes though the tangent of the Scutum-Crux arm; such spiral arm crossings are likely to render the extinction distribution in this direction quite complex.

The sightline towards $\ell = 90°$, bordering the Cygnus region ($60° \lesssim \ell \lesssim 90°$), shows more pronounced variation with latitude, and greater variation between the two strips. Further from the Galactic midplane the two seem to be in good agreement. The spiral arm predictions of Vallée (2008) place the Perseus Arm at ≈ 4 kpc in this direction.

The anticentre sightline shows the least variation- at positive latitudes showing a difference of ≈ 2 per sq. arcmin. between the two strips. The flatness of the distributions highlight the relative lack of structure apparent in the density map at these longitudes; this suggests that the Besançon model is likely to perform well along this sightline - an exponentially decreasing stellar density with galactocentric radius has the potential to capture the observed smooth behaviour. A potential deviation from a smooth distribution comes from the Perseus Arm, predicted by Vallée (2008) to lie at a distance of ≈ 2 kpc at this longitude.

The Galactic warp is not pronounced along any of the three sightlines; it might have been expected towards $\ell \approx 90°$, which lies along the direction of maximum warp. It would seem that extinction in the plane is sufficient in the i'-band to counteract the expected enhanced densities due to the warp in the Galactic midplane. The variations on the scale of $\approx 1°$ in Figure 6.7 suggest that regions of enhanced and suppressed stellar densities are primarily due to substructure in the disc.

6.2.3 The synthetic catalogues

The most obvious feature in the unreddened Besançon catalogues were several regions that displayed anomalously high star counts with respect to their neighbouring results. Figure 6.9 shows an example of the star counts returned by the Besancon web interface at $\ell \approx 90°$ - across a longitude range of interest ($88° < \ell < 90°$): cells returned at specific latitudes ($b = -2°, -1.2°, -0.6°, 0.8°$, and $1.8°$) show significantly higher star counts. The affected latitudes are not constant over the entire Galactic Plane - each set of results at a different longitude display high counts at different latitudes. This effect seemed to scale with count - i.e. peaks in more densely populated regions were higher, with the opposite effect observed in low density areas. An attempt to bring the number count in these cells into line with their surroundings revealed that the counts were apparently doubled; hence halving their counts brought them in line. Latitude strips where twice the number of stars are clearly returned had their counts halved for the following analysis.

6.2.4 Adding extinction

With the diffuse extinction parameter set to 0 mag kpc^{-1}, the synthetic catalogues returned provided an unreddened view of the Milky Way. As the catalogues provide the distance to each generated object, a custom extinction profile could be imposed on the catalogued sources. Four different extinction prescriptions were applied to the catalogues, as detailed below.

Marshall extinction map

The extinction map produced by Marshall et al. (2006) was applied to the catalogues, taking the sightline closest to the synthetic catalogue coordinates - the Marshall et al. (2006) sightlines are binned in 0.25° increments in Galactic longitude and latitude. They provide curves only for sightlines with $\ell < 100°$; comparisons in the outer disc cannot make use of

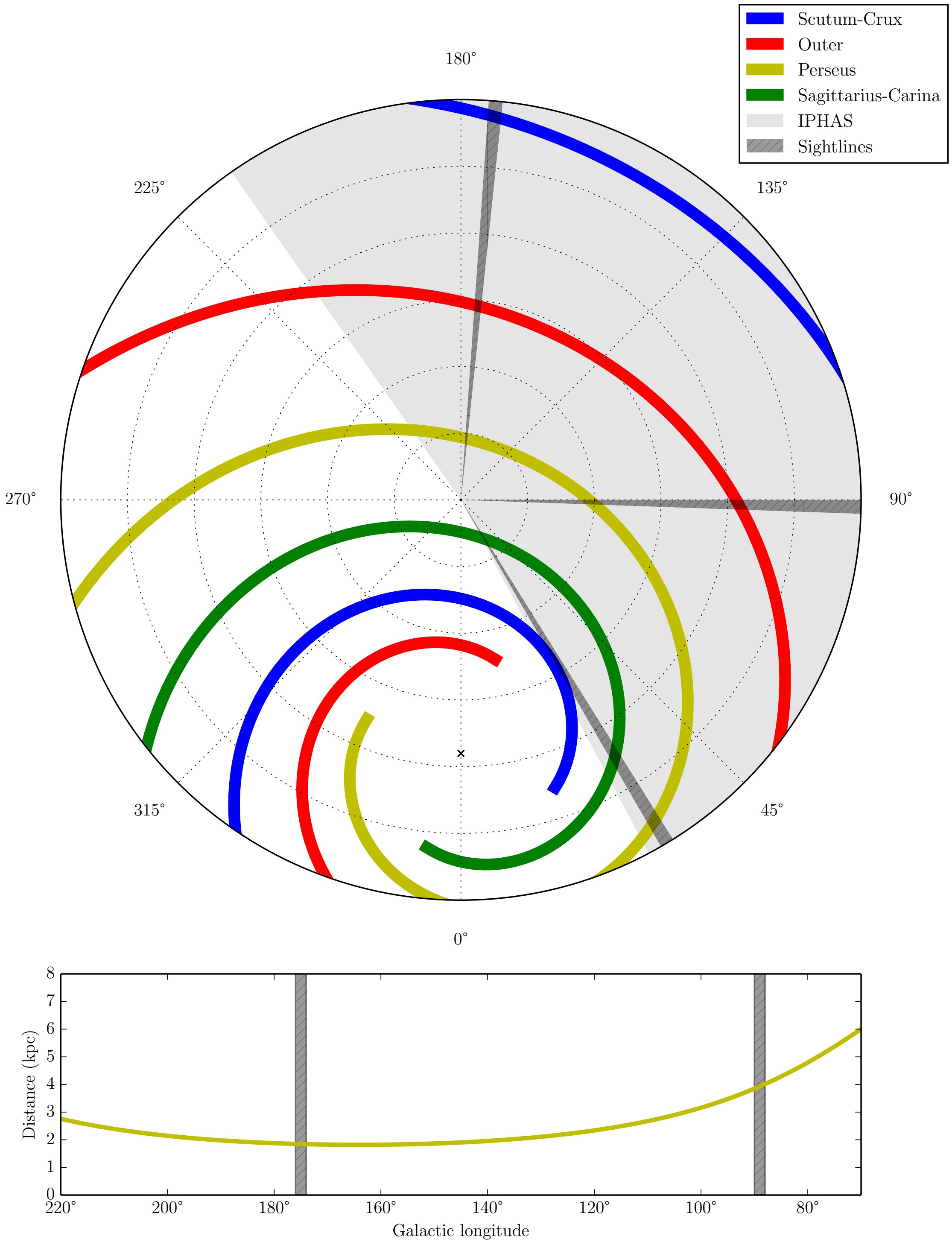


Figure 6.8: Upper: Distribution of spiral arms in Galactic coordinates, with distance from the Sun to the spiral arms based on the parameters obtained by Vallée (2008). Dotted circles denote distances in increments of 2 kpc. The Galactic Centre is shown as a black cross. The region covered by IPHAS is highlighted in gray. The sightlines compared with Besançon model predictions are denoted by darker hatched regions. **Lower:** Distance to the Perseus Arm in the range $70^\circ < \ell < 220^\circ$ based on the parameters of Vallée (2008); this spiral arm is expected to be the main discrete feature likely to influence the observed star counts at $\ell \gtrsim 60^\circ$. The two IPHAS sightlines compared to Besançon model predictions in this region are marked by gray hatched regions.

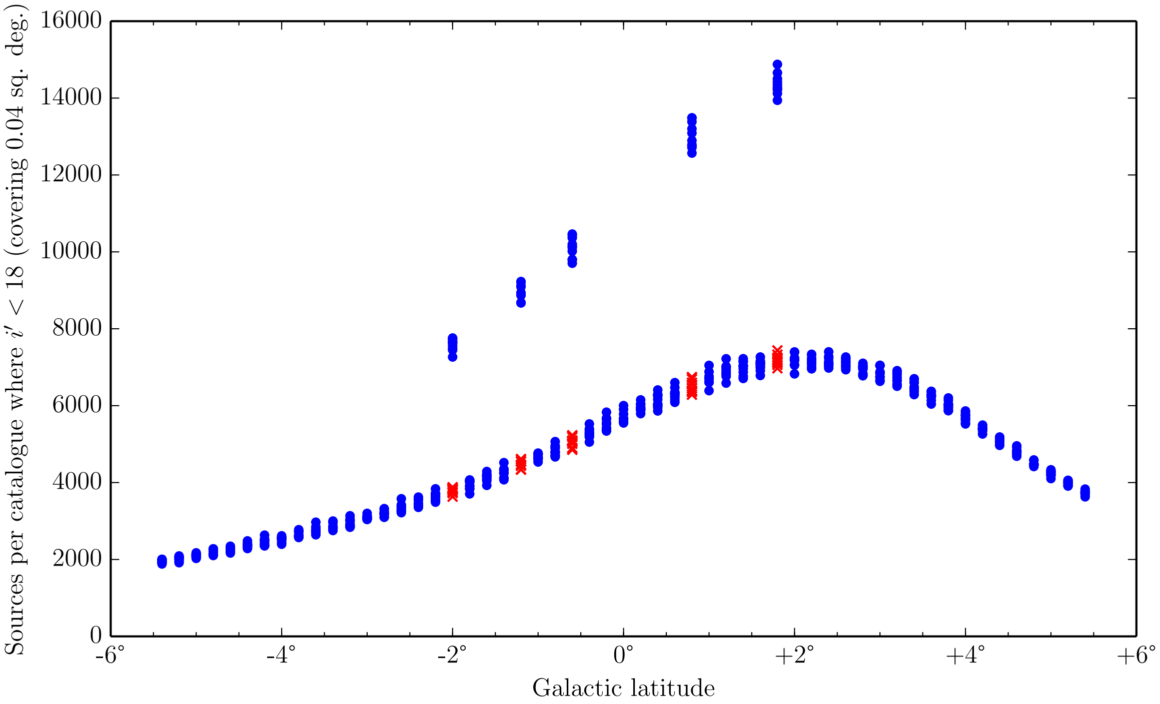


Figure 6.9: The number of sources brighter than $i' = 18$ returned in each Besançon catalogue in a $2^\circ \times 11^\circ$ region ($88^\circ < \ell < 90^\circ, -5.5^\circ < b < +5.5^\circ$). Note that this sightline is in the direction of maximum Galactic warp (see Section 1.2.5), hence the Galactic midplane, showing the greatest stellar density, has shifted upwards to $b \approx 2^\circ$. Each catalogue covers an area of 0.04 sq. deg.. **Blue circles:** Returned number of sources. **Red crosses:** Half the returned number of sources at latitudes with anomalously high counts.

this model. The extinction in this model is given for the K_s band (A_{K_s}) as a function of distance. These values were converted to $A_{i'}$ using the extinction law of Cardelli et al. (1989) (using the conversion factors for K and I filters, giving $A_{i'} = 4.2A_{K_s}$).

Sale extinction map

The 3D map of Sale et al. (2014) are the result of combining IPHAS DR2 photometry with the hierarchical Bayesian model developed in Sale (2012), which estimates the distance-extinction relationship along a given sightline, along with estimates of stellar parameters of the stars sampled. The priors adopted by this approach assume all sightlines consist entirely of thin disc stars, with a scalelength of 3 kpc.

The map provides a typical angular resolution of 10′; the nearest sightline to the Besançon catalogue under consideration was used to extinguish the synthetic photometry. While each sightline is provided with a maximum reliable distance, assuming the final value beyond these limits would result in underestimations of extinction at larger distances. For this reason, entire sightlines were used. Being based on IPHAS photometry, all the sightlines considered fall into the area covered by the map.

Perseus model

The extinction values provided by Schlegel et al. (1998) are integrated values along the line of sight - no radial information is provided that would allow the extinguishing material to be placed along a given sightline. As a simple model of the ISM for sightlines passing through the Perseus Arm, the majority of the extinction was attributed to its position, after assigning an amount of local extinction ($d < 600$ pc) based on the measurements of Lallement et al. (2014). The location of the Perseus Arm was determined from the parameters of the spiral arms obtained by Vallée (2008), as shown in Figure 6.8. Taking a spiral arm half-width of 400 pc (Vallée, 2014) and modelling the dust distribution across the arm as a Gaussian, the total extinction of Schlegel et al. (1998) (recalibrated by Schlafly & Finkbeiner (2011)), less the extinction accounted for locally, was assigned across this pseudo-Perseus Arm.

As illustrated by Figure 6.8 sightlines in the IPHAS footprint become more complicated at $\ell \lesssim 60°$, as they pass through multiple spiral arms. For this reason the Perseus model was limited to sightlines passing through the Perseus (and potentially Outer) Arm only.

Exponential disc model

The default extinction model when using the Besançon web interface is an exponential disc of obscuring material, with a local extinction normalisation of 0.7 mag kpc^{-1} - it is noted that this value should be treated with caution in the Galactic Plane. In order to reproduce a similar extinction model, a disc of obscuring material was imposed on the Besançon catalogue. The local A$_0$ (monochromatic extinction at 5495Å) normalisation was determined by imposing discs with normalisations ranging from 0 to 3 mag kpc^{-1} in steps of 0.05 mag kpc^{-1}, to the Besançon catalogue along each sightline. The magnitude distribution of bright ($12 < r' < 15.5$) stars in the resulting extinguished synthetic catalogue was compared to the corresponding IPHAS DR2 distribution in the same area. The local normalisation that resulted in the best agreement between the two magnitude distributions was adopted for the sightline.

Discussion

Figures 6.10 - 6.12 show comparisons between the four extinction distributions (where available) for sightlines at three longitudes, with sightlines close to and further from the midplane, for which Besançon catalogues were generated. The exponential disc model and Sale et al. (2014) predictions display a certain thickness due to the relation between A_0 and A_i depending on the $T_{\rm eff}$ of sources (a parameter returned by the Besançon catalogues), based on transformations tabulated by Drew et al. (2014) based on the extinction law of Fitzpatrick & Massa (2007).

It should be emphasised that each of these sightlines' curves was applied to one $0.2° \times 0.2°$ region - up to four extinction curves, depending on the region coverage of the models, were applied to each returned Besançon catalogue (480 of which were used to cover each of the three sightline under consideration).

The Marshall et al. (2006) sightlines are more coarsely sampled than the extinction curves of Sale et al. (2014), with more than 90% of sightlines intersecting the IPHAS footprint being defined by fewer than 20 distance points. The extinction in between specified values was linearly interpolated; in addition the extinction at distances closer than the first bin was interpolated assuming $A_i(0\text{kpc}) = 0$, ignoring any localised extinction. In the four sightlines illustrated in Figures 6.11 and 6.12, the Sale et al. (2014) curves include an the amount of local extinction comparable to the estimates of Lallement et al. (2014) (included in the Perseus model) - suggesting that in some cases, linearly interpolating the Marshall et al. (2006) curves to zero will underestimate the contribution of local extinction.

The total extinction attributed to the Perseus Arm in the Perseus model, illustrated in the bottom two panels of Figures 6.11 and 6.12, is in line with the amounts given by the Sale et al. (2014) curves further from the plane; while in the midplane the Sale et al. (2014) curve achieves extinctions both much lower ($\ell = 90°$) and higher ($\ell = 175°$) than the Schlegel et al. (1998) value. The distance range over which the increase in extinction occurs is much narrower in the Perseus model. Towards the anticentre the arm is placed at ≈ 2 kpc (as prescribed by Vallée (2008)), in a region of steady increase of extinction in the Sale et al. (2014) curve - a wider spiral arm could bring the two into better agreement in the examples away from the midplane. At smaller latitudes a wider arm might fare slightly better, though the discrepancy between the extinction beyond 6 kpc would not be reconciled by this adjustment.

In the Perseus extinction prescription, the Perseus Arm was placed at a heliocentric distance corresponding to the midplane distance from Vallée (2008), regardless of the Galactic latitude of the sightline under consideration. At $\ell = 175°$, an arm at 2 kpc at $b = 5°$ corresponds to an out-of-plane distance of ≈ 175 pc; while at 4 kpc along the $\ell = 90°$ sightline, this corresponds to ≈ 350 pc. The extinction at these putative distances, particularly along the $\ell = 90°$ sightline, is likely to be slightly overestimated - in Chapter 1 the distributions of gas and obscuring dust were seen to correlate, and Table 1.1 gives the scaleheight of CO as ≈ 50 pc. However the total extinction along the sightline is still constrained by the Schlegel et al. (1998)/Schlafly & Finkbeiner (2011) value.

The exponential disc of extinguishing material generally agrees with the Sale et al. (2014) curves at $\ell = 90°$, and with the low latitude field at $\ell = 175°$. However it appears to underestimate the extinction in the midplane towards the Galactic anticentre, which is likely to lead to higher than observed numbers of sources in the Besançon catalogues once the model is applied. The opposite is true of the inner disc examples where the larger local normalisation, and line-of-sight through smaller galactocentric radii, result in an extinction curve which rapidly increases beyond the Sale et al. (2014) and Marshall et al. (2006) expectations.

The four extinction models described above were applied (coverage permitting) to each set of synthetic photometry obtained through the Besançon web interface, producing curves relevant to the line-of-sight of each catalogue.

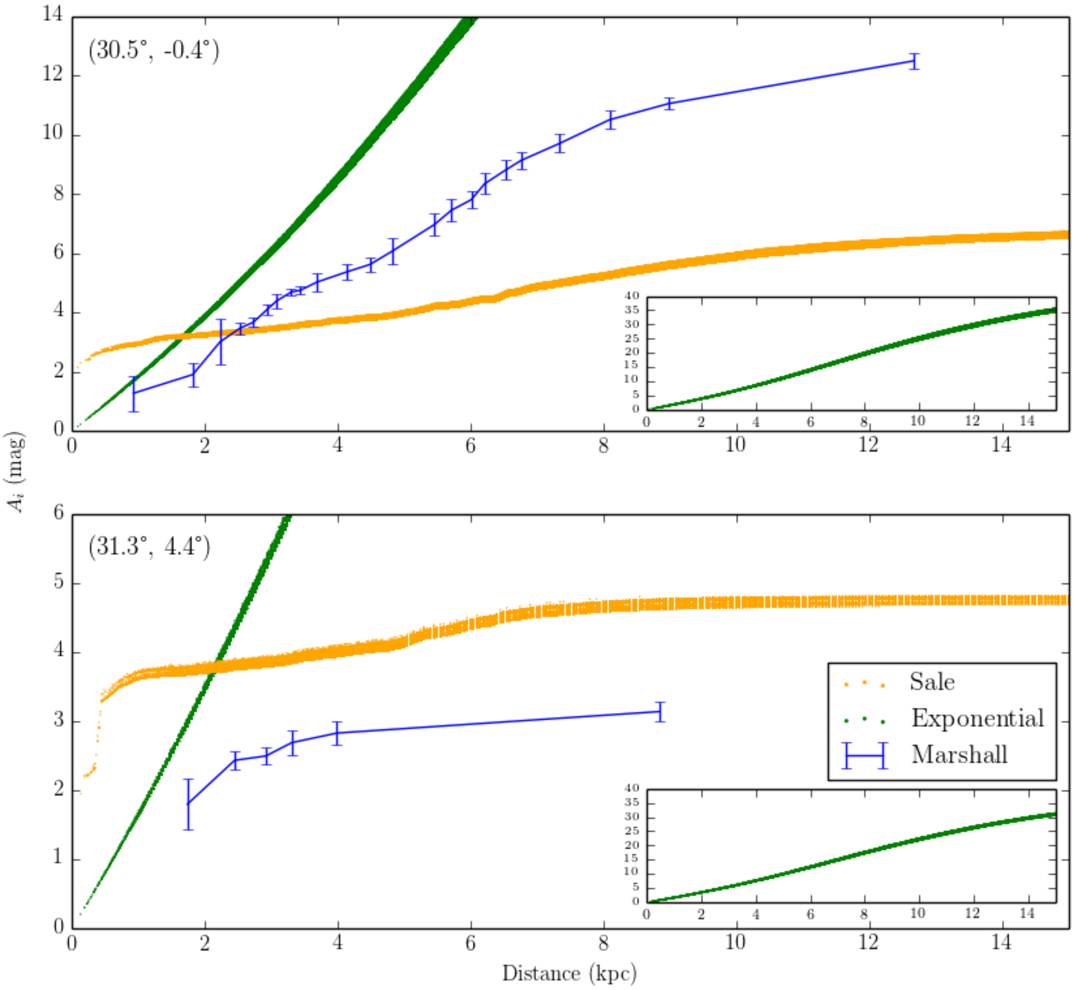


Figure 6.10: A_i, the extinction in the i'-band, against distance for two sightlines at $\ell \approx 30°$, near the Galactic midplane (**upper**) and at larger $|b|$ (**lower**). The curves given by the 3-D extinction map of Sale et al. (2014) are shown in **orange**. The value determined by adding an extinguishing disc of material, with the local normalisation determined by fitting bright Besançon star counts to the observed IPHAS magnitude distribution, is shown in **green**. The Marshall et al. (2006) extinction curve is shown in **blue**. For the first two models, the extinction to every object in the relevant Besançon catalogue is displayed, resulting in a certain thickness due to the T_{eff} dependence involved in converting from A_0 to A_i. The inset plot shows the large extinction towards the inner plane in the case of the exponential extinguishing disc.

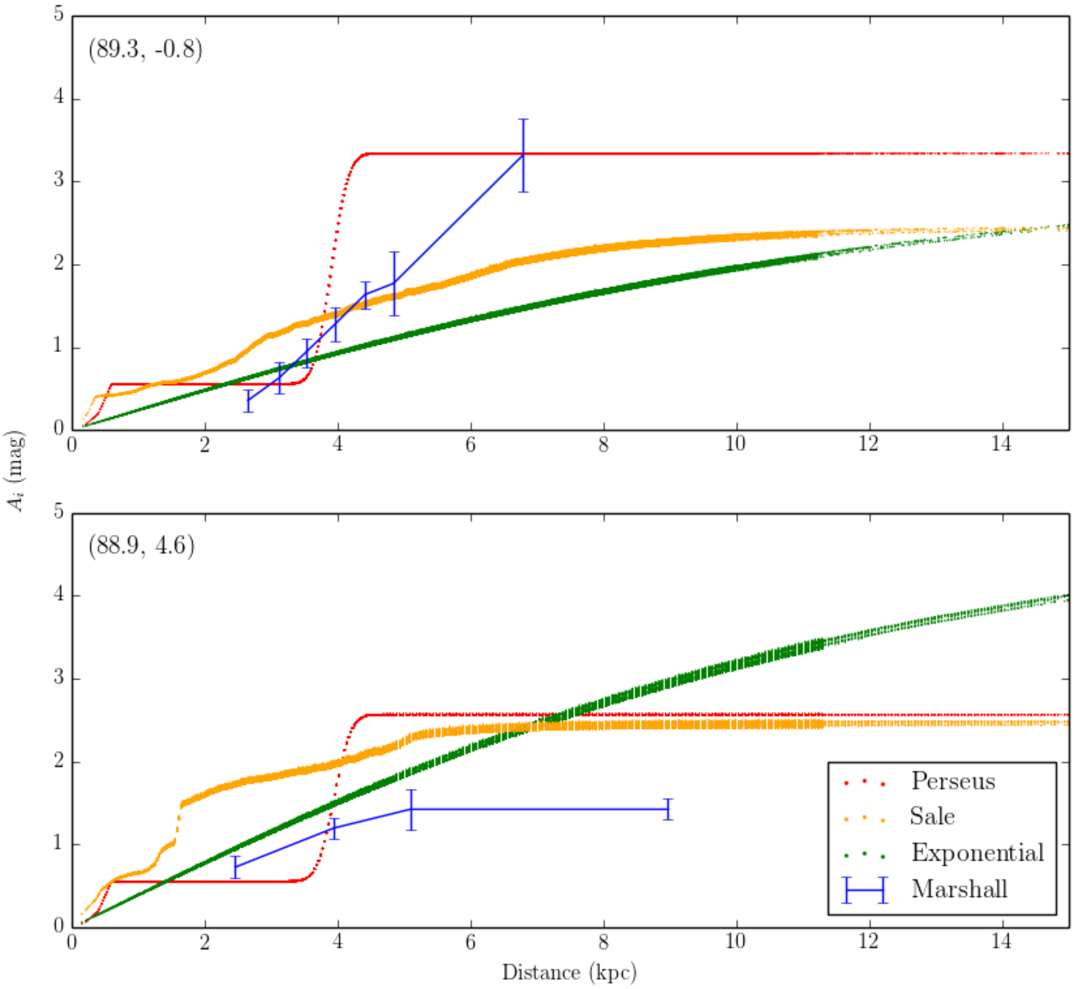


Figure 6.11: The same as Figure 6.10, only for $\ell \approx 90^\circ$. At this longitude the Perseus model is also applied, in which the integrated extinction of Schlegel et al. (1998) (corrected according to Schlafly & Finkbeiner (2011)) is distributed as a degree of local extinction (determined by Lallement et al. (2014)) and a pseudo-Perseus Arm, shown in **red**.

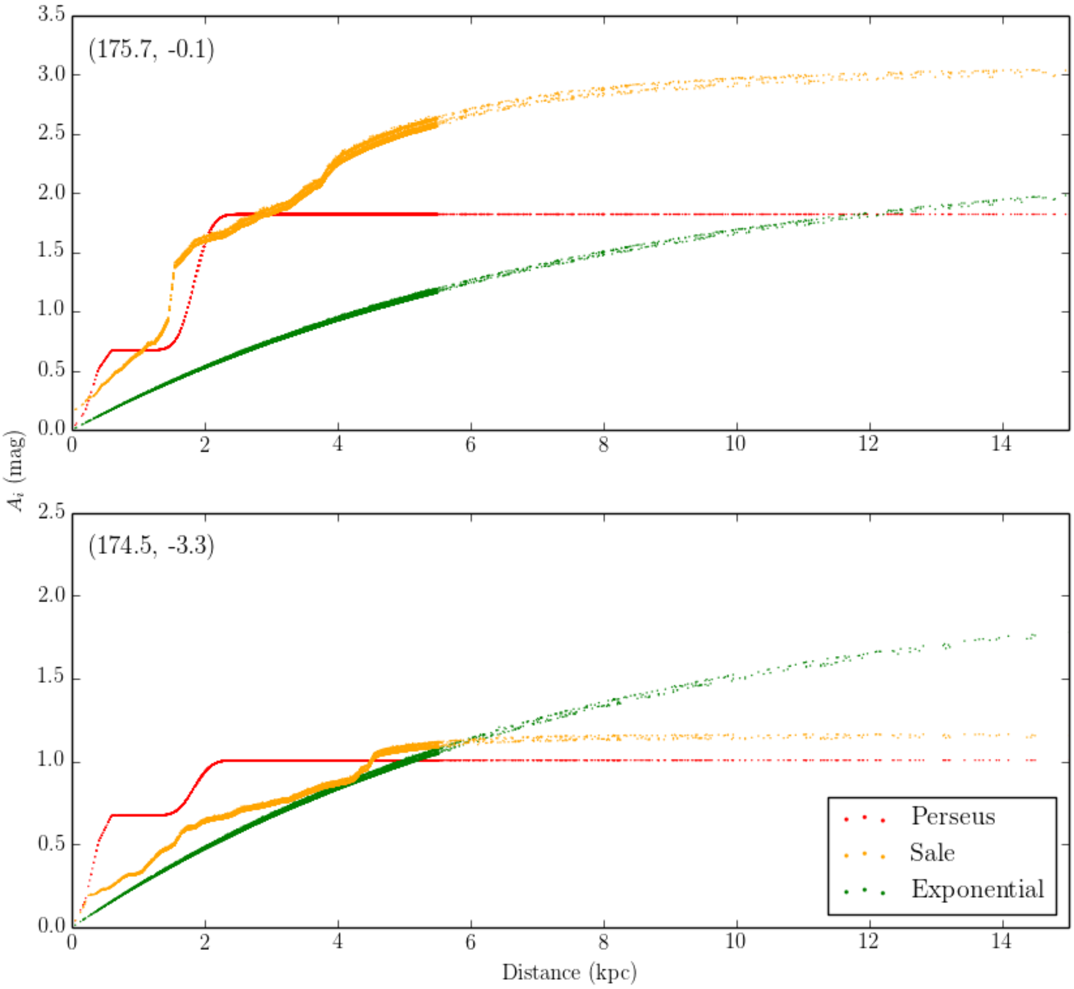


Figure 6.12: Same as Figure 6.10, only for $\ell \approx 175°$. At this longitude the Perseus model is also applied, in which the integrated extinction of Schlegel et al. (1998) (corrected according to Schlafly & Finkbeiner (2011)) is distributed as a degree of local extinction (determined by Lallement et al. (2014)) and a pseudo-Perseus Arm, shown in **red**. With no sightlines at $\ell > 100°$, no Marshall extinction curve is available for this longitude.

6.2.5 Comparing IPHAS and Besançon densities

Figures 6.13 - 6.15 compare IPHAS stellar density distributions presented in Figure 6.7 to the Besançon model predictions, extinguished by the prescriptions detailed in Section 6.2.4.

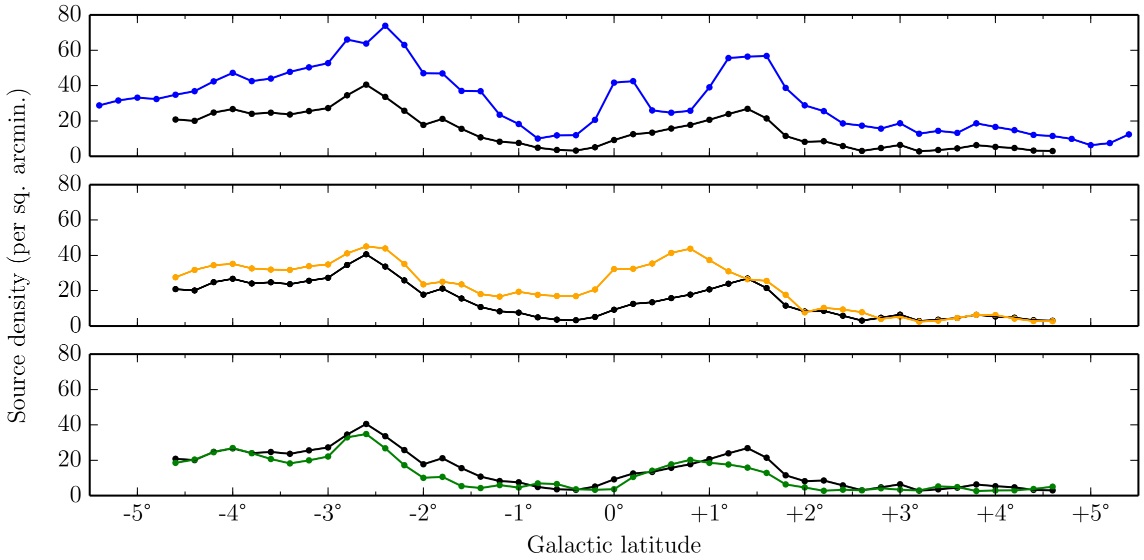
$\ell = 30°$

Towards the inner Galaxy, applying the Marshall et al. (2006) extinction curves to the Besançon star counts results in stellar densities far greater than those observed by IPHAS. This is likely, at least in part, due to the significant variations over small angular scales in the environments probed by sightlines passing through the inner disc; such variations likely require variations in the reddening law assumed. The fixed conversion factor between A_{K_s} and $A_{i'}$ may well be responsible for the variation in the mismatch between IPHAS and reddened Besançon stellar densities across this region.

The predictions of the Besançon model reddened by the Sale et al. (2014) extinction curves also fail to reproduce the observed stellar densities, yielding far higher stellar densities than observed towards the Galactic midplane. This is somewhat counter-intuitive; the Sale et al. (2014) approach models only a thin disc, and therefore is likely to place higher probabilities on stars being closer than in a multi-component Milky Way. Closer stars would require a more rapidly increasing extinction, resulting in overpredictions of extinction and therefore underpredictions of stellar densities. Since the opposite is observed, this implies that the Besançon model significantly overpredicts unreddened star counts towards $\ell = 30°$.

The exponential model applied to the Besançon catalogues results in lower densities than observed by IPHAS. This is explained by the large normalisations adopted along these sightlines- Figure 6.10 shows steep extinction curves (green) which will result in a more heavily extinguished catalogue than from applying the Marshall et al. (2006) and Sale et al. (2014) curves.

The overprediction of densities by the Besançon model could also be due to its adoption of a constant star formation rate (SFR) across the history of the Milky Way. The current understanding of Galactic evolution (see Section 1.3) suggests a SFR declining with time. A higher SFR in the past would serve to relatively enhance the faint end of the stellar luminosity function, with the result that fewer bright-end stars are available to be detected through the higher extinctions encountered at lower Galactic longitudes. If implemented, this could bring predictions more in line with observed IPHAS stellar densities. Recent results presented by Czekaj et al. (2014) rely on an updated Besançon model taking into account a declining SFR. Their Figure 15 shows that even after this adjustment, there is still a failure to reproduce the much brighter ($V < 11.5$ mag) Tycho star counts in this part of the Galactic Plane (by up to a factor of four). Given this, it is unsurprising that the 2003 model overpredicts the IPHAS counts by up to a factor of three.



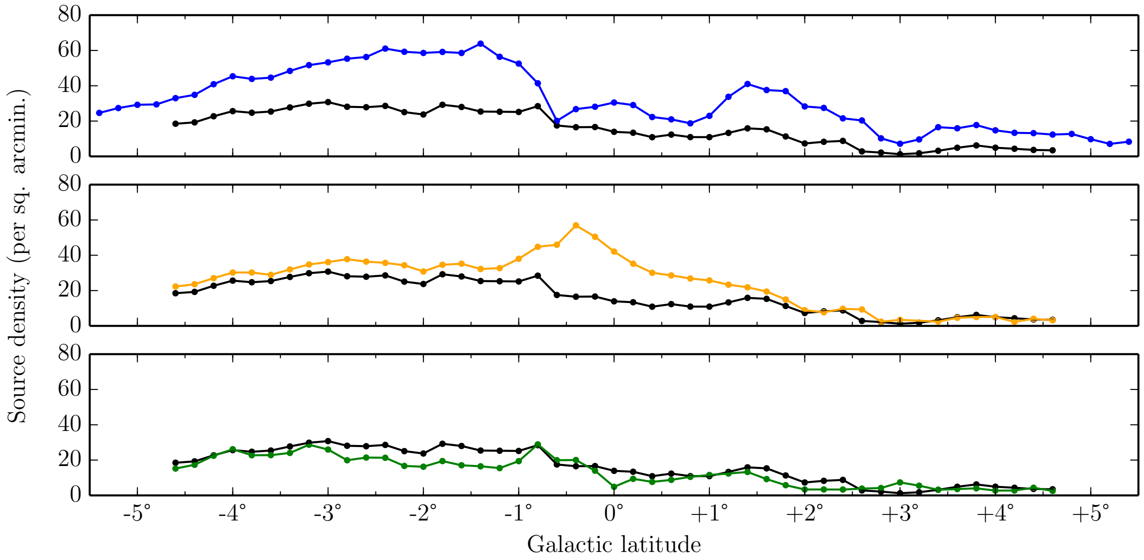


Figure 6.13: Comparisons of i'-band stellar number densities predicted by the Besançon model (combined with various extinction models) and observed IPHAS densities, for $30° < \ell < 31°$ (**upper**) and $31° < \ell < 32°$ (**lower**). Counts from catalogues of the same latitude have been averaged; the area was split into two longitude regions to avoid blurring any substructure with angular size $\approx 1°$. **Black:** Observed IPHAS stellar number densities derived from the $12' \times 12'$ resolution i'-band density map. **Blue:** Besançon stellar densities after Marshall et al. (2006) extinction model has been applied **Orange:** Besançon stellar densities after the Sale et al. (2014) 3D extinction map has been applied. **Green:** Besançon stellar densities after diffuse extinction has been applied, with the local normalisation determined as described in the text.

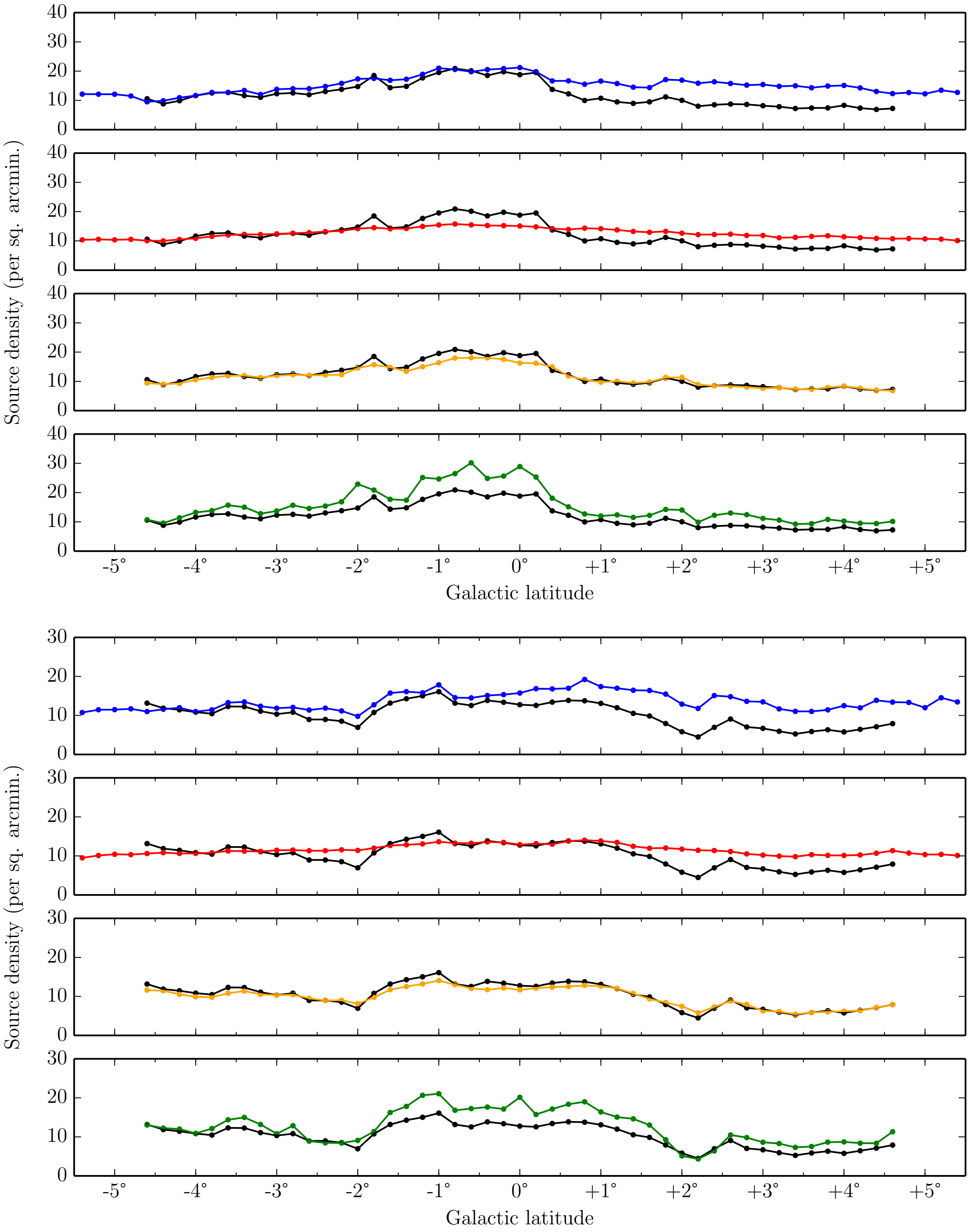
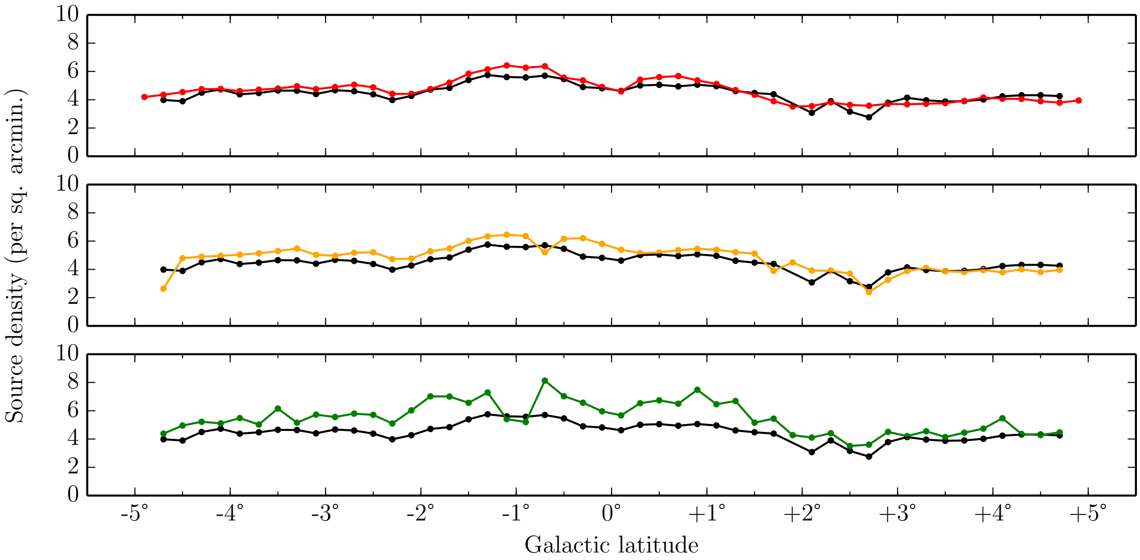


Figure 6.14: Same as Figure 6.13 only for 88° $< \ell <$ 89° (**upper**) and 89° $< \ell <$ 90° (**lower**). An additional extinction model is shown in **red**: these values are the Besançon stellar densities after the "pseudo-Perseus Arm" described in the text has been applied.



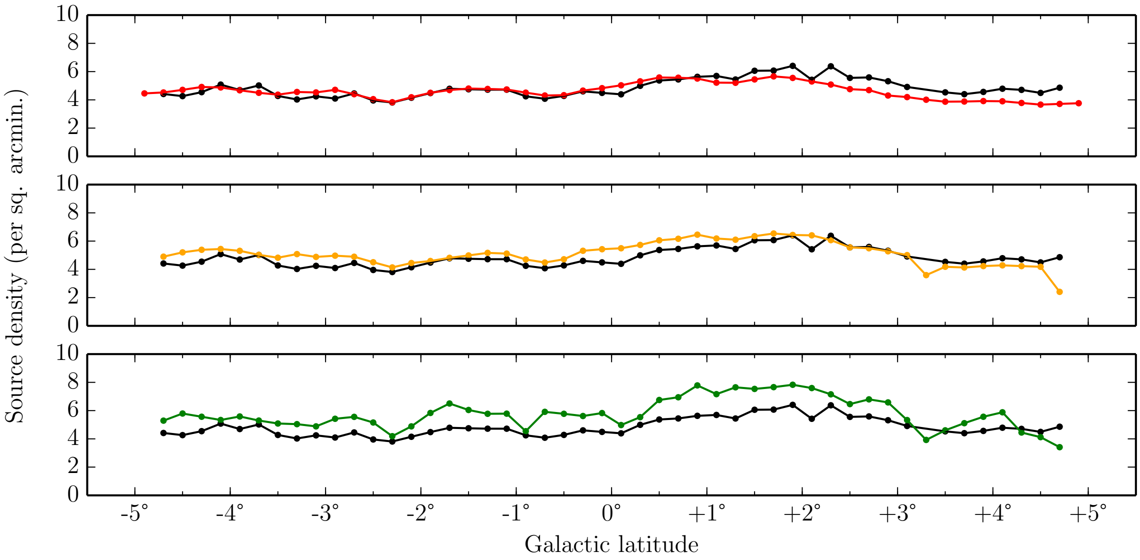


Figure 6.15: Same as Figure 6.13 only for $174^\circ < \ell < 175^\circ$ (**upper**) and $175^\circ < \ell < 176^\circ$ (**lower**). An additional extinction model is shown in **red**: these values are the Besançon stellar densities after the "pseudo-Perseus Arm" described in the text has been applied. No Marshall et al. (2006) extinction curves are available at these Galactic longitudes.

$\ell = 90°$

Along the $\ell = 90°$ sightline, the Besançon counts extinguished according to the Marshall et al. (2006) curves lead to densities in good agreement with IPHAS observations at negative latitudes, but increasingly overestimate the stellar density above the plane. At these Galactic longitudes the effect of the warp is maximised (see Figure 6.9); the overpredictions at positive latitudes could be the result of incorrect model of the Galactic warp. However the ability of the Sale et al. (2014) model to extinguish the Besançon predictions such that they are in good agreement with IPHAS densities suggest that the treatment of the warp is actually quite good. Alternatively, the assumption of a universal extinction law could be the cause of the increasingly poor predictions with increasing latitude; if this is the case the conversion from A_{K_s} to A_i, based on the Cardelli et al. (1989) extinction law, should be allowed to vary based on sightline.

At $\ell = 90°$, the Perseus model does not result in Besançon stellar number densities that agree with IPHAS over the entire area - in fact the resulting distribution of stellar densities is quite flat. The slight increase towards the midplane may be due to the warp, however the smaller scale variations are not reproduced.

$\ell = 175°$

The Perseus Arm model performs quite well at towards the anticentre sightline, following the variations in observed density across the latitude range, as does the Sale et al. (2014) map.

The exponential obscuring disc results in higher densities than observed over the whole latitude range. Though an overprediction of densities towards the anticentre is not entirely unexpected, the fact that the Perseus model outperforms the exponential model suggests that the default 2003 Besançon reddening model of an exponential disc (even after modifying the local normalisation) is not to be trusted even in the outer disc. This part of the Galactic Plane ought to be the best behaved - there is not much visible substructure in the density maps, and the line of sight is along the general exponential decline in stellar density of the disc. Nevertheless the better performance by the Perseus model highlights the need for an increase in extinction more rapid than predicted by a simple exponential disc, consistent with a spiral arm at ≈ 2 kpc (and local extinction of ≈ 0.6 mag).

6.2.6 Summary

The comparisons presented in Section 6.2.5 provide the following insights into the Besançon model and the extinction prescriptions laid out in Section 6.2.4.

- The ability of the Sale et al. (2014) extinction estimates to extinguish the Besançon predicted catalogue, such that it agrees with IPHAS densities in the direction of maximum Galactic warp (shown in Figure 6.14), suggests that the treatment of the warp in the Besançon model is highly realistic.

- Since the warp appears realistic, the overprediction of the Marshall et al. (2006) extinguished Besançon densities relative to IPHAS at positive latitudes (seen in Figure 6.14) may be a consequence of the assumption of a common extinction law (Cardelli et al. (1989)) for the entire Galactic Plane.
- The performance of the Perseus Arm model exceeding the diffuse exponential model in the anticentre direction highlights the inability of the default Besançon extinction model to reproduce the in-plane extinction, even in the outer disc.
- Towards the inner Galaxy, the extinction maps of Marshall et al. (2006) and Sale et al. (2014) fail to extinguish the synthetic Besançon photometry to bring the resulting stellar densities in line with those seen in the IPHAS density map. Again the assumption of a common extinction law across the extent of this sightline may be partly responsible.
- The inability of the Sale et al. (2014) extinction curves to bring the predicted Besançon densities in line with IPHAS towards the inner Galaxy, where an overprediction of extinction might be expected - but the opposite is observed - suggests that the Besançon model is predicting an excess of stellar mass along the $\ell = 30^\circ$ sightline.
- The inner Galaxy sightline is the most difficult to interpret - its intersection with several Vallée (2008) spiral arms (seen in Figure 6.8) alone mean that the Besançon model must reproduce structure that is not entirely understood (as discussed in Chapter 1). The use of additional galactic population synthesis and extinction models along inner disc sightlines may be required to determine which produce the most reliable results.
- The assumption of a constant SFR in the Robin et al. (2003) Besançon model could contribute to the disagreement with observed stellar densities, particularly along the inner Galaxy sightline. However the Czekaj et al. (2014) model which adopts a declining SFR fails to reproduce Tycho star counts in the same region ($\approx(30^\circ, 0^\circ)$), suggesting comparisons with deeper IPHAS counts would still result in overpredictions.

Chapter 7

Summary and future work

7.1 Summary

IPHAS provides a view of the Galactic Plane in unprecedented detail in broad-band r' and i' filters in addition to narrow-band Hα. In this work its broad-band photometry has been utilised to produce a view of the Northern Galactic Plane previously unavailable in the form of completeness-corrected stellar density maps.

Chapter 2 introduced the photometry generated by IPHAS, and showed how simple stellar density maps could be generated from its catalogues; also discussed were some of the pitfalls to be avoided when measuring star counts in IPHAS data. The density maps produced evidently suffered from incompleteness, and revealed issues remaining in the data.

Chapter 3 described how the maps produced in Chapter 2 were used to identify astrometric errors in the dataset, and the exclusion of poor quality IPHAS fields. Crossmatches were performed between IPHAS and APASS photometry in regions where the two surveys intersect, and the median offsets per IPHAS field were measured. These shifts were applied to IPHAS to bring the difference between the two surveys in line, leading the way for the uniform photometric calibration presented by Barentsen et al. (2014). The calibration was verified by performing the same exercise between IPHAS and SDSS photometry; the IPHAS/SDSS offsets were reduced after calibrating to APASS, with residual per-field offsets indicating consistency at a level of 3%. The reliability (i.e. repeatability of sources) of IPHAS was shown to be $> 98\%$ at $r' < 19$.

In Chapter 4, the method by which previous studies assessed their own completeness limits were investigated. After showing that confusion effects can be predicted for a uniformly distributed population of a given number density, it was demonstrated that in some extreme (high number density) cases that the formalism of this approach breaks down, requiring an alternate method be employed. A full treatment of incompleteness due to confusion and sensitivity limits was performed across the survey by inserting artificial sources into every field, and measuring the fraction of recovered sources as a function of magnitude. Examples of similar techniques are present in the literature, however they are rarely described in detail.

Using the corrected, calibrated dataset resulting from the efforts detailed in Chapter 3, and applying the completeness corrections obtained in Chapter 4, corrected density maps were produced as presented in Chapter 5; the maps, which constitute a unique form of data product, are included alongside this thesis. The method by which the maps were produced is detailed, and the uncertainty in their counts discussed. The default resolution of the maps is chosen as $1' \times 1'$, with several lower resolutions also produced; low resolution maps could potentially be of use in the preparation for programmes on multi-fibre spectrographs with large fields-of-view. A potential use of the high-resolution maps is presented in the form of a search for clusters in the i'-band. 70 known clusters were recovered, with 30 further overdensities detected - six of which were also picked up in the r'-band map, while 24 were only seen in the redder i'-band.

In Chapter 6, IPHAS stellar number densities were compared to results from the Besançon model of Galactic population synthesis. The mock catalogue of sources produced by Robin et al. (2012a) (GUMS) was queried to retrieve more than half a billion sources in the Galactic Plane. By combining the SDSS/Gaia transformations of Jordi et al. (2010) and the IPHAS/SDSS transformations obtained in Chapter 3, Gaia magnitudes were calculated for main sequence stars with known IPHAS colours. The optimum combination of filters for comparing the two catalogues was shown to be i' and G, with the Gaia G limits set 0.8 mag fainter. In generating a density map from the GUMS catalogue, it was found that source counts could not be trusted at small Galactic latitudes; the source of the problem remains undetermined.

In order to make meaningful comparisons between Besançon model predictions and the density maps, the web form returning results of Robin et al. (2003) was used to obtain synthetic catalogues along three sightlines spanning the Galactic latitude range of IPHAS. Four treatments of extinction were applied to the unreddened Besançon synthetic Galaxy, with varying degrees of success in replicating IPHAS stellar densities. The Sale et al. (2014) extinction map combined with the Besançon model was found to reproduce IPHAS stellar densities at $\ell = 90°$ and $175°$, suggesting the Besançon model characterisation of the Galactic warp is realistic. However this approach failed towards the inner Galaxy - as did the Marshall et al. (2006) map - suggesting that the Besançon model itself may fail to reproduce the complexity present in the disc within the Solar Circle.

7.2 Future work

7.2.1 Selecting sources from the colour-colour diagram

The properties of the IPHAS colour-colour plane can be exploited (see Section 2) to disentangle the effects of reddening and distance along a sightline. The variation in $(r' - i')$,$(r' - \mathrm{H}\alpha)$ space with spectral type for a given amount of reddening permits the selection of a specified range of stellar types - for example, Sale et al. (2010) picked out early A-type stars from IPHAS photometry. This presents possible avenues for future work, including an obvious extension to the density mapping work that was the focus of this work.

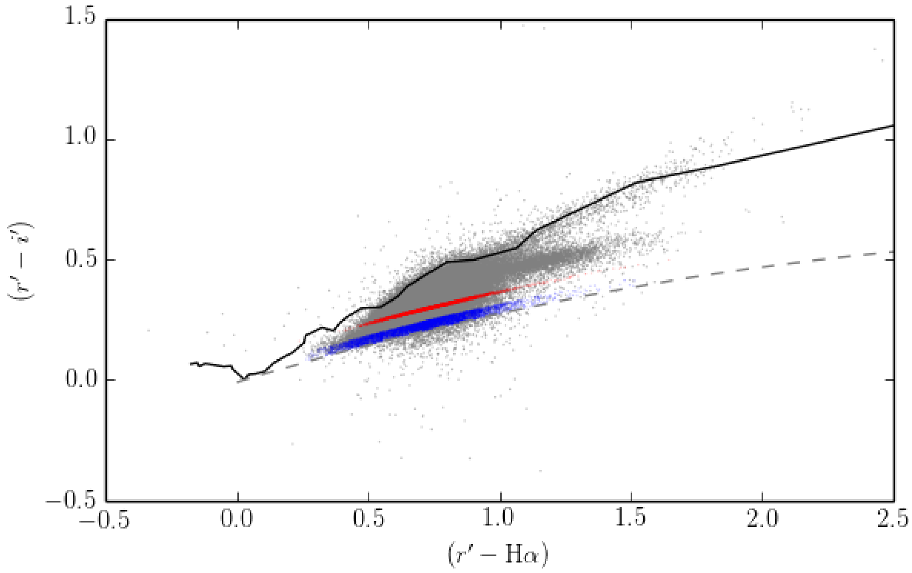


Figure 7.1: The IPHAS colour-colour diagram for the 1 sq. deg. region around the primary sightline of (118°, +2°). **Gray:** All sources with $r' < 19$. **Blue:** Stars selected in a band corresponding to early A-type stars. **Red:** Stars selected in a band corresponding to early F-type stars. **Solid line:** The unreddened main sequence track. **Dotted line:** The A0 reddening line.

7.2.2 An example - object selection for follow-up spectroscopy

As discussed in Chapter 1, the Galactic rotation curve is less certain beyond the Solar Circle. The ability of IPHAS to pick out spectral types was used to select populations in the outer disc for spectroscopic follow up, whose radial velocity distribution could be used to study the rotation curve.

A proposal for time using the Hectospec multi-fibre spectrograph at the MMT was successful, in which a carefully selected sample of A-type and F-type stars were targeted along two sightlines in the outer disc. Figure 7.1 shows the selections made for the primary sightline. A-type stars were selected in a strip of 0.04 mag above the early A-type reddening line, while F-types were picked out in a strip 0.01 mag wide ($0.08 - 0.09$ mag above the early A-type reddening line). The targets were chosen with a range of apparent magnitudes to ensure the heliocentric distance distribution along the line-of-sight was well-sampled. A-type stars are well suited to the aim of studying the rotation law as their velocity dispersions are lower than for objects of later type (Aumer & Binney, 2009), while F-types sample a closer distance range suitable for investigating any perturbation to kinematics of stars due to the presence of the Perseus Arm.

The primary sightline, at $(118°, +2°)$, was chosen for its relatively high stellar density in the direction of maximum radial velocity shear, which occurs at $110° \lesssim \ell \lesssim 120°$. The second sightline at $\ell = 180°$ was chosen as a control, where Galactic rotation will not generate a radial component; any radial motion observed will be due to intrinsic scatter which can inform the analysis of the primary sightline. By building up a kinematic model which incorporates sources of uncertainty such as velocity dispersion and binarity, fine tuned according to the results of the anticentre control sightline, the rotation curve in the outer Galaxy can be

Configuration ID	i'-band mag.	Exposure (min)	Spectra	Useful
118-1	< 16.5	75	259	259
118-2a	$16.5 - 17.5$	140	266	254
118-2b	$16.5 - 17.5$	135	259	259
118-3	$17.5 - 18.5$	265	264	207
118-4	< 17.5	135	256	254
118-5	< 16.5	75	253	253
180-1	< 17.5	120	262	257
180-2	< 18	165	252	252

Table 7.1: The fibre configurations observed using Hectospec. The number of spectra and the number of those which were deemed useful are noted - 'useful' was defined as a median count level above 1000 in the Ca II triplet region.

measured.

The Hectospec instrument offers 300 fibres across a 1°-diameter field of view, of which $\approx$ 30 were reserved per pointing for obtaining contemporaneous sky spectra. Using the high-dispersion grating providing $\approx$ 3Å resolution, spectra for $\approx$ 2000 stars were obtained. The high-dispersion grating was chosen to allow more accurate radial velocity measurements; the spectral range covered by Hectospec at $\approx$ 3Å resolution (6550 − 9045Å) spans the Ca II triplet region. In this region the spectral lines are well defined for A- and F-type stars (Paschen and Ca II respectively), useful for measuring radial velocities.

Table 7.1 shows the fibre configurations for which observations were made, along with their magnitude ranges, exposure times, and the number of object spectra (i.e. not counting sky spectra) that were obtained. Also shown is the number of useful spectra, defined as those whose median count level in the Ca II region is above 1000, corresponding to $S/N \gtrsim 30$.

Visual inspection of the spectra suggest that the spectral typing achieved by using IPHAS colours has worked well; either Ca II triplet or Paschen lines are visible in the vast majority of obtained spectra. The next step will involve obtaining spectral types by fitting template spectra to the data, likely using MCMC code which treat not only stellar parameters as free parameters, but also vary the associated sky spectrum (or combination of sky-fibre spectra) subtracted from the object spectrum to produce the best fit.

Once the spectral types are known to better accuracy than simply 'early A/F-type', distances can be estimated, which along with measurements of radial velocity, will provide a powerful dataset for comparisons against kinematic models. It should be emphasised that the selection for this study was performed prior to the calibration discussed in Chapter 3 - calibrated DR2 photometry will render future selections even more accurate. A further example of the potential applications of colour-cuts is presented next.

7.2.3 Colour-cut density maps

Slater et al. (2014) showed the utility of presenting a density map with colour cuts imposed. They plotted the density of stars derived from SDSS with $0.2 < (g - r)_0 < 0.3$ to filter out most foreground (nearby disc) and background (halo) sources, in order to study the

Monoceros Ring. They then presented three density maps, selecting stars with increasingly faint g_0 magnitudes, in order to select more distant objects, selecting stars with magnitudes in the ranges $17.8 < g_0 < 18.4$, $18.8 < g_0 < 19.6$ and $20.2 < g_0 < 20.6$, corresponding to distances of $14.4-17.4$, $7.6-11.0$ and $14.4-17.4$ kpc. Due to high extinction in the Galactic Plane, the maps of Slater et al. (2014) are missing significant coverage at $|b| < 10°$.

The generation of single-band density maps (Chapter 5) could be modified to count objects based on their $(r' - i')$, $(r' - \mathrm{H}\alpha)$ colours. The visualisation of such counts would map out populations based on stellar type; however this endeavour would require the density map correction method presented in Chapter 4 be heavily modified. Rather than measuring the completeness of 3× frames (r', i' & Hα) independently for each field, the completeness *of regions of the colour-colour plane* would need to be evaluated. This would require artificial sources to be generated with three magnitudes simultaneously which result in sensible colours, then inserted into the three frames of the field at the same position on the sky. The varying recovery rate of sources across the colour-colour plane would be determined by a threshold in the difference between simulated and recovered colours, analogous to the $r'_{synth} - r'_{recover}$ magnitude threshold applied in Chapter 4.

The ability of the IPHAS colour-colour plane to deliver relatively accurate estimates of spectral type, as demonstrated in Section 7.2.2, makes it possible to pick out populations of a specified range of stellar types for density mapping purposes.

7.2.4 Further model tests

Due to the challenges discussed above, the immediate future is arguably best spent developing the use of star counts as tests of ongoing efforts to map Galactic extinction. As highlighted in the summary of Chapter 6, the comparisons of galactic population and extinction models with the IPHAS density maps produced in this work allow their ability to reproduce observations to be assessed. Comparisons with additional models, along additional sightlines, will improve the ability to determine which models perform best, in which regions of the Galactic Plane. Adopting the revised Besançon model (Czekaj et al., 2014) would reduce the effect of outdated assumptions (e.g. a constant SFR, as discussed in Section 6.2.5). Until the revised model is made available publicly, decomposing the results of the Robin et al. (2003) model into stellar populations could help to determine the source of discrepancies encountered in Chapter 6.

The approach of Drimmel et al. (2003) of rescaling their model predictions using COBE data before testing against 2MASS data could be adapted; using the more recent Herschel/-Planck estimates of column densities to determine potential rescaling factors for models of extinction, and the increased i'-band sensitivity to reddening relative to 2MASS, the performance of the models could be more accurately tested.

Personal acknowledgements

Firstly I must thank my supervisor Janet Drew whose guidance, support, and insight over the last four years have been invaluable resources, without which this thesis would not exist.

Thanks to Geert for many useful conversations on astronomy, and for frequent passionate rants regarding best coding practices. Thanks also to Nick, who offered me a place to stay during my visit to CfA, and has since become a good friend.

During my time at UH I have been lucky enough to form friendships with exceptionally kind and intelligent people, both within and outside the department. In what is by no means an exhaustive list, I would like to thank Kieran, Silvia, James F, Maria, Carlos, Noemí, Beatriz, Joana, João, Carla, Gülay, Emrah, Nicole, Ged, Ylvi, Liz, Ian, Lientjie, Lisette, Manuel, Hugo, ZengHua, James C, Neil, Mike, Emma, Gaius, Nancy, Graham, Giulio, Stuart, Christine, Andrés, Noé, Ed, George, Majd, Lauryanne, and of course Ariane.

The experience of living in Hatfield has been improved immensely thanks to the residents of Horsa Gardens, with whom I shared a roof and an extremely unreliable boiler. In particular I would like to thank Roberto, Federico and Andy, whose enthusiasm for throwing parties has ensured I will never forget my time here.

Last but certainly not least, I would like to thank my family. It is thanks to their support that I have arrived at this point, I could not have done this without them.

Further acknowledgements

This work has made use of the University of Hertfordshire Science and Technology Research Institute high-performance computing facility. Many thanks to John Atkinson and Martin Hardcastle for their administration of the IT infrastructure at the STRI.

IPHAS is carried out at the Isaac Newton Telescope (INT). 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 Astrofisica de Canarias. All IPHAS data are processed by the Cambridge Astronomical Survey Unit, at the Institute of Astronomy in Cambridge.

This thesis makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

This research has made use of the SIMBAD database, the VizieR catalogue access tool, and the Aladin service, operated at CDS, Strasbourg, France. Also used was the WEBDA database, operated at the Department of Theoretical Physics and Astrophysics of the Masaryk University, and NASA's Astrophysics Data System.

During the data reduction, analysis and visualisation steps of this work, the software packages TOPCAT and STILTS were invaluable resources. I have also made extensive use of Python, in particular IPython, during the course of this work. Packages that have been of particular use are MATPLOTLIB, ASTROPY, KAPTEYN, NUMPY and SCIPY.

The comparison with GUMS in Chapter 6 uses the results of simulations run at the MareNostrum supercomputer.

I would like to thank the STFC for providing the studentship which funded my studies.

Bibliography

Abadi, M. G., Navarro, J. F., Steinmetz, M., & Eke, V. R. 2003, ApJ, 597, 21

Abell, G. O. 1959, Astronomical Society of the Pacific Leaflets, 8, 121

Abt, H. A. & Levy, S. G. 1976, ApJS, 30, 273

Adibekyan, V. Z., Delgado Mena, E., Sousa, S. G., Santos, N. C., Israelian, G., González Hernández, J. I., Mayor, M., & Hakobyan, A. A. 2012, AA, 547, A36

Amôres, E. B. & Lépine, J. R. D. 2005, AJ, 130, 659

Aumer, M. & Binney, J. 2009, MNRAS, 397, 1286

Baade, W. 1944, ApJ, 100, 137

—. 1951, Publications of the Observatory of Michigan, 10, 7

Bahcall, J. N. & Soneira, R. M. 1980, ApJS, 44, 73

—. 1984, ApJS, 55, 67

Bahcall, N. A. & Soneira, R. M. 1983, ApJ, 270, 20

Bailer-Jones, C. A. L. 2011, MNRAS, 411, 435

Barentsen, G., Farnhill, H. J., Drew, J. E., González-Solares, E. A., Greimel, R., Irwin, M. J., Miszalski, B., Ruhland, C., Groot, P., Mampaso, A., Sale, S. E., Henden, A. A., Aungwerojwit, A., Barlow, M. J., Carter, P. J., Corradi, R. L. M., Drake, J. J., Eislöffel, J., Fabregat, J., Gänsicke, B. T., Gentile Fusillo, N. P., Greiss, S., Hales, A. S., Hodgkin, S., Huckvale, L., Irwin, J., King, R., Knigge, C., Kupfer, T., Lagadec, E., Lennon, D. J., Lewis, J. R., Mohr-Smith, M., Morris, R. A. H., Naylor, T., Parker, Q. A., Phillipps, S., Pyrzas, S., Raddi, R., Roelofs, G. H. A., Rodríguez-Gil, P., Sabin, L., Scaringi, S., Steeghs, D., Suso, J., Tata, R., Unruh, Y. C., van Roestel, J., Viironen, K., Vink, J. S., Walton, N. A., Wright, N. J., & Zijlstra, A. A. 2014, MNRAS, 444, 3230

Barentsen, G., Vink, J. S., Drew, J. E., & Sale, S. E. 2013, MNRAS, 429, 1981

Barker, M. K., Sarajedini, A., Geisler, D., Harding, P., & Schommer, R. 2007, AJ, 133, 1138

Barry, D. C. 1988, ApJ, 334, 436

Bartko, H., Martins, F., Trippe, S., Fritz, T. K., Genzel, R., Ott, T., Eisenhauer, F., Gillessen, S., Paumard, T., Alexander, T., Dodds-Eden, K., Gerhard, O., Levin, Y., Mascetti, L., Nayakshin, S., Perets, H. B., Perrin, G., Pfuhl, O., Reid, M. J., Rouan, D., Zilka, M., & Sternberg, A. 2010, ApJ, 708, 834

Bastian, N., Covey, K. R., & Meyer, M. R. 2010, ARAA, 48, 339

Battaglia, G., Helmi, A., Morrison, H., Harding, P., Olszewski, E. W., Mateo, M., Freeman, K. C., Norris, J., & Shectman, S. A. 2005, MNRAS, 364, 433

Battaner, E., Florido, E., & Sanchez-Saavedra, M. L. 1990, AA, 236, 1

Becla, J., Hanushevsky, A., Nikolaev, S., Abdulla, G., Szalay, A., Nieto-Santisteban, M., Thakar, A., & Gray, J. 2006, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 6270

Bell, E. F., Zucker, D. B., Belokurov, V., Sharma, S., Johnston, K. V., Bullock, J. S., Hogg, D. W., Jahnke, K., de Jong, J. T. A., Beers, T. C., Evans, N. W., Grebel, E. K., Ivezić, v., Koposov, S. E., Rix, H.-W., Schneider, D. P., Steinmetz, M., & Zolotov, A. 2008, ApJ, 680, 295

Benjamin, R. A. 2008, Massive Star Formation: Observations Confront Theory, ASP Conference Series, 387, 375

Bienaymé, O., Robin, A. C., & Creze, M. 1987, AA, 180, 94

Binney, J. & Dehnen, W. 1997, MNRAS, 287

Binney, J., Dehnen, W., & Bertelli, G. 2000, MNRAS, 318, 658

Binney, J. & Lacey, C. 1988, MNRAS, 230, 597

Binney, J. & Tremaine, S. 2008, Galactic Dynamics: Second Edition (Princeton University Press),

Blum, R. D., Ramirez, S. V., Sellgren, K., & Olsen, K. 2003, ApJ, 597, 323

Bonatto, C., Kerber, L. O., Bica, E., & Santiago, B. X. 2006, AA, 446, 121

Bond, N. A., Ivezić, v., Sesar, B., Jurić, M., Munn, J. A., Kowalski, A., Loebman, S., Roškar, R., Beers, T. C., Dalcanton, J., Rockosi, C. M., Yanny, B., Newberg, H. J., Allende Prieto, C., Wilhelm, R., Lee, Y. S., Sivarani, T., Majewski, S. R., Norris, J. E., Bailer-Jones, C. A. L., Fiorentin, P. R., Schlegel, D. J., Uomoto, A., Lupton, R. H., Knapp, G. R., Gunn, J. E., Covey, K. R., Smith, J. A., Miknaitis, G., Doi, M., Tanaka, M., Fukugita, M., Kent, S., Finkbeiner, D., Quinn, T. R., Hawley, S., Anderson, S., Kiuchi, F., Chen, A., Bushong, J., Sohi, H., Haggard, D., Kimball, A., McGurk, R., Barentine, J., Brewington, H., Harvanek, M., Kleinman, S., Krzesinski, J., Long, D., Nitta, A., Snedden, S., Lee, B., Pier, J. R., Harris, H., Brinkmann, J., & Schneider, D. P. 2010, ApJ, 716, 1

Bottema, R. 2003, MNRAS, 344, 358

Bournaud, F., Elmegreen, B. G., & Elmegreen, D. M. 2007, ApJ, 670, 237

Bournaud, F., Elmegreen, B. G., & Martig, M. 2009, ApJ, 707, L1

Bovy, J., Allende Prieto, C., Beers, T. C., Bizyaev, D., da Costa, L. N., Cunha, K., Ebelke, G. L., Eisenstein, D. J., Frinchaboy, P. M., García Pérez, A. E., Girardi, L., Hearty, F. R., Hogg, D. W., Holtzman, J., Maia, M. A. G., Majewski, S. R., Malanushenko, E., Malanushenko, V., Mészáros, S., Nidever, D. L., O'Connell, R. W., O'Donnell, C., Oravetz, A., Pan, K., Rocha-Pinto, H. J., Schiavon, R. P., Schneider, D. P., Schultheis, M., Skrutskie, M. F., Smith, V. V., Weinberg, D. H., Wilson, J. C., & Zasowski, G. 2012a, ApJ, 759, 131

Bovy, J., Hogg, D. W., & Rix, H.-W. 2009, ApJ, 704, 1704

Bovy, J., Rix, H.-W., & Hogg, D. W. 2012b, ApJ, 751, 131

Brand, J. & Blitz, L. 1993, AA, 275, 67

Brand, J. & Wouterloot, J. G. A. 2007, AA, 464, 909

Brook, C. B., Kawata, D., Gibson, B. K., & Freeman, K. C. 2004, ApJ, 612, 894

Burstein, D. 1979, ApJ, 234, 829

Burstein, D. & Heiles, C. 1982, AJ, 87, 1165

Burton, W. B. & Hartmann, D. 1994, Astrophysics and Space Science, 217, 189

Cambrésy, L., Beichman, C. A., Jarrett, T. H., & Cutri, R. M. 2002, AJ, 123, 2559

Cantat-Gaudin, T., Vallenari, A., Zaggia, S., Bragaglia, A., Sordo, R., Drew, J. E., Eisloeffel, J., Farnhill, H. J., Gonzalez-Solares, E., Greimel, R., Irwin, M. J., Kupcu Yoldas, A., Jordi, C., Blomme, R., Sampedro, L., Costado, M. T., Alfaro, E., Smiljanic, R., Magrini, L., Donati, P., Friel, E. D., Jacobson, H., Abbas, U., Hatzidimitriou, D., Spagna, A., Vecchiato, A., Balaguer-Nunez, L., Lardo, C., Tosi, M., Pancino, E., Klutsch, A., Tautvaisiene, G., Drazdauskas, A., Puzeras, E., Jiménez-Esteban, F., Maiorca, E., Geisler, D., San Roman, I., Villanova, S., Gilmore, G., Randich, S., Bensby, T., Flaccomio, E., Lanzafame, A., Recio-Blanco, A., Damiani, F., Hourihane, A., Jofré, P., de Laverny, P., Masseron, T., Morbidelli, L., Prisinzano, L., Sacco, G. G., Sbordone, L., & Worley, C. C. 2014, AA, 569, A17

Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245

Carlberg, R. G. & Sellwood, J. A. 1985, ApJ, 292, 79

Carney, B. W., Latham, D. W., & Laird, J. B. 1989, AJ, 97, 423

Carney, B. W. & Seitzer, P. 1993, AJ, 105, 2127

Carraro, G., Ng, Y. K., & Portinari, L. 1998, MNRAS, 296, 1045

Carraro, G., Vázquez, R. A., Costa, E., Perren, G., & Moitinho, A. 2010, ApJ, 718, 683

Chiappini, C., Matteucci, F., & Gratton, R. 1997, ApJ, 477, 765

Chiappini, C., Matteucci, F., & Romano, D. 2001, ApJ, 554, 1044

Chiba, M. & Beers, T. C. 2000, AJ, 119, 2843

Clemens, D. P., Sanders, D. B., & Scoville, N. Z. 1988, ApJ, 327, 139

Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, MNRAS, 319, 168

Condon, J. J. 1974, ApJ, 188, 279

Corradi, R. L. M., RodríguezFlores, E. R., Mampaso, A., Greimel, R., Viironen, K., Drew, J. E., Lennon, D. J., Mikolajewska, J., Sabin, L., & Sokoloski, J. L. 2008, AA, 480, 409

Czekaj, M. A., Robin, A. C., Figueras, F., Luri, X., & Haywood, M. 2014, AA, 564, A102

Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, ApJ, 547, 792

De Simone, R., Wu, X., & Tremaine, S. 2004, MNRAS, 350, 627

Deacon, N. R., Groot, P. J., Drew, J. E., Greimel, R., Hambly, N. C., Irwin, M. J., Aungwerojwit, A., Drake, J., & Steeghs, D. 2009, MNRAS, 397, 1685

Dehnen, W. & Binney, J. 1998, MNRAS, 298, 387

Delgado, A. J., Alfaro, E. J., Aparicio, A., & Cabrera-Cano, J. 1994, AJ, 108, 2193

Dennison, B., Simonetti, J. H., & Topasna, G. A. 1998, PASA, 15, 147

Dobbs, C. & Baba, J. 2014, PASA, 31, 35

Doi, M., Tanaka, M., Fukugita, M., Gunn, J. E., Yasuda, N., Ivezić, v., Brinkmann, J., de Haars, E., Kleinman, S. J., Krzesinski, J., & Leger, R. F. 2010, AJ, 139, 1628

Draine, B. 2003, ARAA, 41, 241

Drew, J. E., González-Solares, E. A., Greimel, R., Irwin, M. J., Kupcu Yoldas, A., Lewis, J., Barentsen, G., Eisloffel, J., Farnhill, H. J., Martin, W. E., Walsh, J. R., Walton, N. A., Mohr-Smith, M., Raddi, R., Sale, S. E., Wright, N. J., Groot, P., Barlow, M. J., Corradi, R. L. M., Drake, J. J., Fabregat, J., Frew, D. J., Gansicke, B. T., Knigge, C., Mampaso, A., Morris, R. A. H., Naylor, T., Parker, Q. A., Phillipps, S., Ruhland, C., Steeghs, D., Unruh, Y. C., Vink, J. S., Wesson, R., & Zijlstra, A. A. 2014, MNRAS, 440, 2036

Drew, J. E., Greimel, R., Irwin, M. J., Aungwerojwit, A., Barlow, M. J., Corradi, R. L. M., Drake, J. J., Gansicke, B. T., Groot, P., Hales, A., Hopewell, E. C., Irwin, J., Knigge, C., Leisy, P., Lennon, D. J., Mampaso, A., Masheder, M. R. W., Matsuura, M., Morales-Rueda, L., Morris, R. A. H., Parker, Q. A., Phillipps, S., Rodriguez-Gil, P., Roelofs, G., Skillen, I., Sokoloski, J. L., Steeghs, D., Unruh, Y. C., Viironen, K., Vink, J. S., Walton, N. A., Witham, A., Wright, N. J., Zijlstra, A. A., & Zurita, A. 2005, MNRAS, 362, 753

Drimmel, R. 2000, AA,

Drimmel, R., Cabrera-Lavers, A., & López-Corredoira, M. 2003, AA, 409, 205

Drimmel, R. & Spergel, D. N. 2001, ApJ, 556, 181

Duquennoy, . & Mayor, . 1991, AA, 248, 485

Dwek, E. 2005, AIP Conference Proceedings, 761, 103

Dwek, E., Arendt, R. G., Hauser, M. G., Kelsall, T., Lisse, C. M., Moseley, S. H., Silverberg, R. F., Sodroski, T. J., & Weiland, J. L. 1995, ApJ, 445, 716

Edvardsson, B., Andersen, J., Gustafsson, B., Lambert, D. L., Nissen, P. E., & Tomkin, J. 1993, AA,

Eggen, O. J., Lynden-Bell, D., & Sandage, A. R. 1962, ApJ, 136, 748

Eisenhauer, F., Genzel, R., Alexander, T., Abuter, R., Paumard, T., Ott, T., Gilbert, A., Gillessen, S., Horrobin, M., Trippe, S., Bonnet, H., Dumas, C., Hubin, N., Kaufer, A., KisslerPatig, M., Monnet, G., Strobele, S., Szeifert, T., Eckart, A., Schodel, R., & Zucker, S. 2005, ApJ, 628, 246

Ewen, H. I. & Purcell, E. M. 1951, Nature, 168, 356

Faber, S. M., Burstein, D., Tinsley, B., & King, I. R. 1976, AJ, 81, 45

Feast, M. & Whitelock, P. 1997, MNRAS, 291, 683

Feltzing, S., Holmberg, J., & Hurley, J. R. 2001, AA, 377, 911

Fich, M., Blitz, L., & Stark, A. A. 1989, ApJ, 342, 272

Finkbeiner, D. P. 2003, ApJS, 146, 407

Fitzgerald, M. P. 1968, AJ, 73, 983

Fitzpatrick, E. L. & Massa, D. 2007, ApJ, 663, 320

Freudenreich, H. T. 1998, ApJ, 492, 495

Freudenreich, H. T., Berriman, G. B., Dwek, E., Hauser, M. G., Kelsall, T., Moseley, S. H., Silverberg, R. F., Sodroski, T. J., Toller, G. N., & Weiland, J. L. 1994, ApJ, 429

Froebrich, D., Scholz, A., & Raftery, C. L. 2007, MNRAS, 374, 399

Fujii, M. S., Baba, J., Saitoh, T. R., Makino, J., Kokubo, E., & Wada, K. 2011, ApJ, 730, 109

Galileo, G. 1610, Sidereus Nuncius (Venice: Thomas Baglioni)

Gaustad, J., McCullough, P., Rosing, W., & VanBuren, D. 2001, PASP, 113, 1326

Georgelin, Y. M. & Georgelin, Y. P. 1976, AA, 49, 57

Ghez, A. M., Salim, S., Weinberg, N. N., Lu, J. R., Do, T., Dunn, J. K., Matthews, K., Morris, M. R., Yelda, S., Becklin, E. E., Kremenek, T., Milosavljevic, M., & Naiman, J. 2008, ApJ, 689, 1044

Gill, D. 1896, Annals of the Cape Observatory,

Gilmore, G. & Reid, N. 1983, MNRAS, 202, 1025

Gilmore, G., Wyse, R. F. G., & Jones, B. J. 1995, AJ, 109, 1095

Girardi, L., Groenewegen, M. A. T., Hatziminaoglou, E., & da Costa, L. 2005, AA, 436, 895

Glazebrook, K., Peacock, J. A., Collins, C. A., & Miller, L. 1994, MNRAS, 266

Gnedin, O. Y. & Ostriker, J. P. 1997, ApJ, 474, 223

Gomez, A. E., Grenier, S., Udry, S., Haywood, M., Meillon, L., Sabas, V., Sellier, A., & Morin, D. 1997, Proceedings of the ESA Symposium ‘Hipparcos - Venice ’97’, , 621

Gonzalez, O. A., Rejkuba, M., Zoccali, M., Valenti, E., Minniti, D., Schultheis, M., Tobar, R., & Chen, B. 2012, AA, 543, A13

González-Fernández, C., López-Corredoira, M., Amôres, E. B., Minniti, D., Lucas, P., & Toledo, I. 2012, AA, 546, A107

González-Solares, E. A., Irwin, M. J., McMahon, R. G., Hodgkin, S., Lewis, J. R., Walton, N. A., Jarvis, M., Marchetti, L., Oliver, S., Pérez-Fournon, I., Siana, B., Surace, J., & Vaccari, M. 2011, MNRAS, 416, 927

González-Solares, E. A., Walton, N. A., Greimel, R., Drew, J. E., Irwin, M. J., Sale, S. E., Andrews, K., Aungwerojwit, A., Barlow, M. J., van den Besselaar, E., Corradi, R. L. M., Gänsicke, B. T., Groot, P. J., Hales, A. S., Hopewell, E. C., Hu, H., Irwin, J., Knigge, C., Lagadec, E., Leisy, P., Lewis, J. R., Mampaso, A., Matsuura, M., Moont, B., Morales-Rueda, L., Morris, R. A. H., Naylor, T., Parker, Q. A., Prema, P., Pyrzas, S., Rixon, G. T., Rodriguez-Gil, P., Roelofs, G., Sabin, L., Skillen, I., Suso, J., Tata, R., Viironen, K., Vink, J. S., Witham, A., Wright, N. J., Zijlstra, A. A., Zurita, A., Drake, J., Fabregat, J., Lennon, D. J., Lucas, P. W., Martn, E. L., Phillipps, S., Steeghs, D., & Unruh, Y. C. 2008, MNRAS, 388, 89

Goodwin, S. P. & Kouwenhoven, M. B. N. 2009, MNRAS, 397, L36

Governato, F., Willman, B., Mayer, L., Brooks, A., Stinson, G., Valenzuela, O., Wadsley, J., & Quinn, T. 2007, MNRAS, 374, 1479

Groot, P. J., Verbeek, K., Greimel, R., Irwin, M. J., González-Solares, E. A., Gänsicke, B. T., de Groot, E., Drew, J. E., Augusteijn, T., Aungwerojwit, A., Barlow, M. J., Barros, S., van den Besselaar, E., Casares, J., Corradi, R., Corral-Santana, J., Deacon, N., van Ham, W., Hu, H., Heber, U., Jonker, P. G., King, R., Knigge, C., Mampaso, A., Marsh, T., Morales-Rueda, L., Napiwotzki, R., Naylor, T., Nelemans, G., Oosting, T., Pyrzas, S., Pretorius, M., Rodriguez-Gil, P., Roelofs, G., Sale, S. E., Schellart, P., Steeghs, D., Szyszka, C., Unruh, Y. C., Walton, N. A., Weston, S., Witham, A. R., Woudt, P., & Zijlstra, A. A. 2009, MNRAS, 399, 323

Gum, C. S. & Pawsey, J. L. 1960, MNRAS, 121

Gunn, J. E. & Knapp, G. R. 1993, Sky Surveys. Protostars to Protogalaxies. Proceedings of a Conference in Honor of Gerry Neugebauer, 43

Gyuk, G., Flynn, C., & Evans, N. W. 1999, ApJ, 521, 190

Haffner, L. M., Reynolds, R. J., Tufte, S. L., Madsen, G. J., Jaehnig, K. P., & Percival, J. W. 2003, ApJS, 149, 405

Hammersley, P. & López-Corredoira, M. 2011, AA, 527, A6

Hanninen, J. & Flynn, C. 2002, MNRAS, 337, 731

Harvey, P. M., Chapman, N., Lai, S., Evans, II, N. J., Allen, L. E., Jorgensen, J. K., Mundy, L. G., Huard, T. L., Porras, A., Cieza, L., Myers, P. C., Merin, B., van Dishoeck, E. F., Young, K. E., Spiesman, W., Blake, G. A., Koerner, D. W., Padgett, D. L., Sargent, A. I., & Stapelfeldt, K. R. 2006, ApJ, 644, 307

Haywood, M., Di Matteo, P., Lehnert, M. D., Katz, D., & Gómez, A. 2013, AA, 560, A109

Haywood, M., Robin, A. C., & Creze, M. 1997a, AA, 320, 428

—. 1997b, AA, 320, 440

Heiles, C. & Troland, T. H. 2003, ApJ, 586, 1067

Heiles, . 1975, AAS, 20

Henden, A. & Munari, U. 2014, Contributions of the Astronomical Observatory Skalnaté Pleso, 43, 518

Henderson, A. P., Jackson, P. D., & Kerr, F. J. 1982, ApJ, 263, 116

Herschel, W. 1785, Philosophical Transactions of the Royal Society of London, 75, 213

Hiltner, W. A. & Johnson, H. L. 1956, ApJ, 124, 367

Hoffleit, D. & Jaschek, C. V. 1991, The Bright Star Catalogue (New Haven, Conn.: Yale Univ. Obs.) [LINK]

Høg, E., Bässgen, G., Bastian, U., Egret, D., Fabricius, C., Groß mann, V., Halbwachs, J. L., Makarov, V. V., Perryman, M. A. C., Schwekendiek, P., Wagner, K., & Wicenec, A. 1997, AA, 323, 57

Hogg, D. W. 2001, AJ, 121, 1207

Howarth, I. D. 1983, MNRAS, 203, 301

Howell, S. B. 2006, Handbook of CCD Astronomy (Cambridge University Press) [LINK]

Hubble, E. P. 1922, ApJ, 56, 162

Ibata, R. A., Gilmore, G., & Irwin, M. J. 1994, Nature, 370, 194

Irwin, M. J. & Trimble, V. 1984, AJ, 89, 83

Irwin, . 1985, MNRAS, 214, 575

Jacobson, H. R., Asplund, M., Bessell, M. S., Casey, A. R., Da Costa, G. S., Frebel, A., Keller, S. C., Lind, K., Norris, J. E., Schmidt, B. P., Tisserand, P., & Yong, D. 2014, Memorie della Societa Astronomica Italiana, 85, 227

Jenkins, A. 1992, MNRAS, 257, 620

Jordi, C., Gebran, M., Carrasco, J. M., de Bruijne, J., Voss, H., Fabricius, C., Knude, J., Vallenari, A., Kohley, R., & Mora, A. 2010, AA, 523, A48

Jordi, C., Hog, E., Brown, A. G. A., Lindegren, L., Bailer-Jones, C. A. L., Carrasco, J. M., Knude, J., Straizys, V., de Bruijne, J., Claeskens, J.-F., Drimmel, R., Figueras, F., Grenon, M., Kolka, I., Perryman, M. A. C., Tautvaisiene, G., Vansevicius, V., Willemsen, P. G., Bridzius, A., Evans, D. W., Fabricius, C., Fiorucci, M., Heiter, U., Kaempf, T. A., Kazlauskas, A., Kucinskas, A., Malyuto, V., Munari, U., Reylé, C., Torra, J., Vallenari, A., Zdanavicius, K., Korakitis, R., Malkov, O., & Smette, A. 2006, MNRAS, 367, 290

Joshi, Y. C. 2007, MNRAS, 378, 768

Jurić, M., Ivezić, v., Brooks, A., Lupton, R. H., Schlegel, D. J., Finkbeiner, D., Padmanabhan, N., Bond, N., Sesar, B., Rockosi, C. M., Knapp, G. R., Gunn, J. E., Sumi, T., Schneider, D. P., Barentine, J. C., Brewington, H. J., Brinkmann, J., Fukugita, M., Harvanek, M., Kleinman, S. J., Krzesinski, J., Long, D., Neilsen, J., Nitta, A., Snedden, S., & York, D. G. 2008, ApJ, 673, 864

Just, A. & Jahreiß, H. 2010, MNRAS, 402, 461

Kalberla, P. M. W., Burton, W. B., Hartmann, D., Arnal, E. M., Bajaja, E., Morras, R., & Pöppel, W. G. L. 2005, AA, 440, 775

Kalberla, P. M. W. & Dedes, L. 2008, AA, 487, 951

Kant, I. 1755, Allgemeine Naturgeschichte und Theorie des Himmels (Konigsburg und Leipzig: Johann Friederich Petersen) [LINK]

Kapteyn, J. C. 1904, AJ, 24, 115

—. 1922, ApJ, 55, 302

Keller, S. C., Schmidt, B. P., Bessell, M. S., Conroy, P. G., Francis, P., Granlund, A., Kowald, E., Oates, A. P., Martin-Jones, T., Preston, T., Tisserand, P., Vaccarella, A., & Waterson, M. F. 2007, PASA, 24, 1

Kent, S. M., Dame, T. M., & Fazio, G. 1991, ApJ, 378, 131

Kerr, F. J. 1957, AJ, 62, 93

Kerr, F. J. 1983in Kinematics, dynamics and structure of the Milky Way; Proceedings of the Workshop on the Milky Way, 91 [LINK]

Kharchenko, N. V., Piskunov, A. E., Röser, S., Schilbach, E., & Scholz, R.-D. 2005, AA, 438, 1163

Knox, R. A., Hawkins, M. R. S., & Hambly, N. 1999, MNRAS, 306, 736

Kormendy, J. & Kennicutt, R. C. 2004, ARAA, 42, 603

Kormendy, J. & Norman, C. A. 1979, ApJ, 233, 539

Kroupa, P. 2000, Astronomische Gesellschaft Meeting Abstracts, 16, 11

—. 2002, MNRAS, 330, 707

Kwee, K. K., Muller, C. A., & Westerhout, G. 1954, Bull. Astron. Inst. Netherlands, 12, 211

Lada, C. J. 2006, ApJ, 640, L63

Lallement, R., Vergely, J.-L., Valette, B., Puspitarini, L., Eyer, L., & Casagrande, L. 2014, AA, 561, A91

Langer, W. D., Pineda, J. L., & Velusamy, T. 2014, AA, 564, A101

Lasker, B. M., Doggett, J., McLean, B., Sturch, C., Djorgovski, S., de Carvalho, R. R., & Reid, I. N. 1996, Astronomical Data Analysis Software and Systems V, 101

Lawrence, A., Warren, S. J., Almaini, O., Edge, A. C., Hambly, N. C., Jameson, R. F., Lucas, P., Casali, M., Adamson, A., Dye, S., Emerson, J. P., Foucaud, S., Hewett, P., Hirst, P., Hodgkin, S. T., Irwin, M. J., Lodieu, N., McMahon, R. G., Simpson, C., Smail, I., Mortlock, D., & Folger, M. 2007, MNRAS, 379, 1599

Levine, E. S., Blitz, L., & Heiles, C. 2006a, Science, 312, 1773

—. 2006b, ApJ, 643, 881

Lewis, J. R. & Freeman, K. 1989, AJ, 97, 139

Lin, C. C. & Shu, F. H. 1964, ApJ, 140, 646

Loebman, S. R., Roškar, R., Debattista, V. P., Ivezić, v., Quinn, T. R., & Wadsley, J. 2011, ApJ, 737, 8

López-Corredoira, M., Betancort-Rijo, J., & Beckman, J. E. 2002, AA, 386, 169

Lopez-Corredoira, M., Cabrera-Lavers, A., Garzon, F., & Hammersley, P. 2002, AA, 394, 883

López-Corredoira, M., Cabrera-Lavers, A., Mahoney, T. J., Hammersley, P. L., Garzón, F., & González-Fernández, C. 2007, AJ, 133, 154

Lozinskaya, T. A. & Kardashev, N. S. 1963, Soviet Astronomy, 7, 161

LSST Science Collaboration. 2009, LSST Science Book, Version 2.0,

Lucas, P. W., Hoare, M. G., Longmore, A., Schröder, A. C., Davis, C. J., Adamson, A., Bandyopadhyay, R. M., de Grijs, R., Smith, M., Gosling, A., Mitchison, S., Gáspár, A., Coe, M., Tamura, M., Parker, Q., Irwin, M. J., Hambly, N., Bryant, J., Collins, R. S., Cross, N., Evans, D. W., González-Solares, E. A., Hodgkin, S., Lewis, J., Read, M., Riello, M., Sutorius, E. T. W., Lawrence, A., Drew, J. E., Dye, S., & Thompson, M. A. 2008, MNRAS, 391, 136

Luhman, K. L. 2012, ARAA, 50, 65

Luri, X., Palmer, M., Arenou, F., Masana, E., de Bruijne, J., Antiche, E., Babusiaux, C., Borrachero, R., Sartoretti, P., Julbe, F., Isasi, Y., Martinez, O., Robin, A. C., Reylé, C., Jordi, C., & Carrasco, J. M. 2014, AA, 566, A119

Maciejewski, G., Mihov, B., & Georgiev, T. 2009, Astronomische Nachrichten, 330, 851

Magnier, E. A., Schlafly, E., Finkbeiner, D., Juric, M., Tonry, J. L., Burgett, W. S., Chambers, K. C., Flewelling, H. A., Kaiser, N., Kudritzki, R.-P., Morgan, J. S., Price, P. A., Sweeney, W. E., & Stubbs, C. W. 2013, ApJS, 205, 20

Maíz-Apellániz, J. 2004, PASP, 116, 859

Marshall, D. J., Robin, A. C., Reylé, C., Schultheis, M., & Picaud, S. 2006, AA, 453, 635

Martin, N. F., Ibata, R. A., Bellazzini, M., Irwin, M. J., Lewis, G. F., & Dehnen, W. 2004a, MNRAS, 348, 12

Martin, N. F., Ibata, R. A., Conn, B. C., Lewis, G. F., Bellazzini, M., & Irwin, M. J. 2005, MNRAS, 362, 906

Martin, N. F., Ibata, R. A., Conn, B. C., Lewis, G. F., Bellazzini, M., Irwin, M. J., & McConnachie, A. W. 2004b, MNRAS, 355, L33

Mason, B. D., Hartkopf, W. I., Gies, D. R., Henry, T. J., & Helsel, J. W. 2009, AJ, 137, 3358

Mateo, M. & Hodge, P. 1986, ApJS, 60, 893

McKee, C. F. & Ostriker, J. P. 1977, ApJ, 218, 148

McMahon, R., Walton, N., Irwin, M., Lewis, J., Bunclark, P., & Jones, D. 2001, New Astronomy Reviews, 45, 97

McMahon, R. G., Banerji, M., Gonzalez, E., Koposov, S. E., Bejar, V. J., Lodieu, N., Rebolo, R., & VHS Collaboration. 2013, The Messenger, 154, 35

McMillan, P. J. & Binney, J. 2010, MNRAS, 402, 934

McWilliam, A. 1997, ARAA, 35, 503

Metchev, S. A., Kirkpatrick, J. D., Berriman, G. B., & Looper, D. 2008, ApJ, 676, 1281

Minniti, D., Lucas, P., Emerson, J., Saito, R., Hempel, M., Pietrukowicz, P., Ahumada, A., Alonso, M., Alonso-García, J., & Arias, J. 2010, New Astronomy, 15, 433

Miyamoto, M., Yoshizawa, M., & Suzuki, S. 1988, AA, 194, 107

Momany, Y., Zaggia, S., Bonifacio, P., Piotto, G., De Angeli, F., Bedin, L. R., & Carraro, G. 2004, AA, 421, L29

Momany, Y., Zaggia, S., Gilmore, G., Piotto, G., Carraro, G., Bedin, L. R., & de Angeli, F. 2006, AA, 451, 515

Monet, D. G., Levine, S. E., Canzian, B., Ables, H. D., Bird, A. R., Dahn, C. C., Guetter, H. H., Harris, H. C., Henden, A. A., Leggett, S. K., Levison, H. F., Luginbuhl, C. B., Martini, J., Monet, A. K. B., Munn, J. A., Pier, J. R., Rhodes, A. R., Riepe, B., Sell, S., Stone, R. C., Vrba, F. J., Walker, R. L., Westerhout, G., Brucato, R. J., Reid, I. N., Schoening, W., Hartley, M., Read, M. A., & Tritton, S. B. 2003, AJ, 125, 984

Morgan, W. W., Whitford, A. E., & Code, A. D. 1953, ApJ, 118, 318

Munn, J. A., Monet, D. G., Levine, S. E., Canzian, B., Pier, J. R., Harris, H. C., Lupton, R. H., Ivezić, v., Hindsley, R. B., Hennessy, G. S., Schneider, D. P., & Brinkmann, J. 2004, AJ, 127, 3034

Newberg, H. J., Yanny, B., Rockosi, C., Grebel, E. K., Rix, H.-W., Brinkmann, J., Csabai, I., Hennessy, G., Hindsley, R. B., Ibata, R., Ivezić, v., Lamb, D., Nash, E. T., Odenkirchen, M., Rave, H. A., Schneider, D. P., Smith, J. A., Stolte, A., & York, D. G. 2002, ApJ, 569, 245

Norris, J. 1987, ApJ, 314, L39

Olling, R. P. & Merrifield, M. R. 1998, MNRAS, 297, 943

O'Mullane, W., Lammers, U., Bailer-Jones, C., Bastian, U., Brown, A. G. A., Drimmel, R., Eyer, L., Huc, C., Katz, D., Lindegren, L., Pourbaix, D., Luri, X., Torra, J., Mignard, F., & van Leeuwen, F. 2007, Astronomical Data Analysis Software and Systems XVI ASP Conference Series, 376

Oort, J. H. 1927a, Bull. Astron. Inst. Netherlands, 4, 79

—. 1927b, Bull. Astron. Inst. Netherlands, 3, 275

Oort, J. H., Kerr, F. J., & Westerhout, G. 1958, MNRAS, 118, 379

Ortiz, R. & Lepine, J. R. D. 1993, AA, 279, 90

Padmanabhan, N., Schlegel, D. J., Finkbeiner, D. P., Barentine, J. C., Blanton, M. R., Brewington, H. J., Gunn, J. E., Harvanek, M., Hogg, D. W., Ivezić, v., Johnston, D., Kent, S. M., Kleinman, S. J., Knapp, G. R., Krzesinski, J., Long, D., Neilsen, J., Nitta, A., Loomis, C., Lupton, R. H., Roweis, S., Snedden, S., Strauss, M. A., & Tucker, D. L. 2008, AJ, 674, 1217

Parker, Q. A., Phillipps, S., Pierce, M. J., Hartley, M., Hambly, N. C., Read, M. A., MacGillivray, H. T., Tritton, S. B., Cass, C. P., Cannon, R. D., Cohen, M., Drew, J. E., Frew, D. J., Hopewell, E., Mader, S., Malin, D. F., Masheder, M. R. W., Morgan, D. H., Morris, R. A. H., Russeil, D., Russell, K. S., & Walker, R. N. F. 2005, MNRAS, 362, 689

Patterson, F. S. 1940, Harvard College Observatory Bulletin, 914

Perryman, M. A. C., de Boer, K. S., Gilmore, G., Hø g, E., Lattanzi, M. G., Lindegren, L., Luri, X., Mignard, F., Pace, O., & de Zeeuw, P. T. 2001, AA, 369, 339

Perryman, M. A. C., Lindegren, L., Kovalevsky, J., Hoeg, E., Bastian, U., Bernacca, P. L., Crézé, M., Donati, F., Grenon, M., Grewing, M., van Leeuwen, F., van der Marel, H., Mignard, F., Murray, C. A., Le Poole, R. S., Schrijver, H., Turon, C., Arenou, F., Froeschlé, M., & Petersen, C. S. 1997, AA, 323, L49

Pettitt, A. R., Dobbs, C. L., Acreman, D. M., & Price, D. J. 2014, MNRAS, 444, 919

Pilkington, K., Few, C. G., Gibson, B. K., Calura, F., Michel-Dansac, L., Thacker, R. J., Mollá, M., Matteucci, F., Rahimi, A., Kawata, D., Kobayashi, C., Brook, C. B., Stinson, G. S., Couchman, H. M. P., Bailin, J., & Wadsley, J. 2012, AA, 540, A56

Porcel, C. & Battaner, E. 1995, MNRAS, 274, 1153

Portinari, L. & Chiosi, C. 1999, AA, 350, 827

Quillen, A. C. & Garnett, D. R. 2001, in Galaxy Disks and Disk Galaxies, 230, 87

Quinn, P. J., Hernquist, L., & Fullagar, D. P. 1993, ApJ, 403, 74

Raddi, R., Drew, J. E., Fabregat, J., Steeghs, D., Wright, N. J., Sale, S. E., Farnhill, H. J., Barlow, M. J., Greimel, R., Sabin, L., Corradi, R. M. L., & Drake, J. J. 2013, MNRAS, 430, 2169

Reach, W. T., Dwek, E., Fixsen, D. J., Hewagama, T., Mather, J. C., Shafer, R. A., Banday, A. J., Bennett, C. L., Cheng, E. S., Eplee, R. E., J., Leisawitz, D., Lubin, P. M., Read, S. M., Rosen, L. P., Shuman, F. G. D., Smoot, G. F., Sodroski, T. J., & Wright, E. L. 1995, ApJ, 451, 188

Reed, B. C. 1996, AJ, 111, 804

Reid, I. N. & Gizis, J. E. 1997, AJ, 113, 2246

Reid, M. J. 1993, ARAA, 31, 345

Reid, M. J. & Brunthaler, A. 2004, ApJ, 616, 872

Reid, M. J., Menten, K. M., Zheng, X. W., Brunthaler, A., Moscadelli, L., Xu, Y., Zhang, B., Sato, M., Honma, M., Hirota, T., Hachisuka, K., Choi, Y. K., Moellenbrock, G. A., & Bartkiewicz, A. 2009a, ApJ, 700, 137

Reid, M. J., Menten, K. M., Zheng, X. W., Brunthaler, A., & Xu, Y. 2009b, ApJ, 705, 1548

Reylé, C., Marshall, D. J., Robin, A. C., & Schultheis, M. 2009, AA, 495, 819

Reylé, C. & Robin, A. C. 2001, AA, 373, 886

Robin, A. & Creze, M. 1986, AA, 157, 71

Robin, A. C., Creze, M., & Mohan, V. 1992, AA, 265, 32

Robin, A. C., Haywood, M., Creze, M., Ojha, D. K., & Bienaymé, O. 1996, AA, 305, 125

Robin, A. C., Luri, X., Reylé, C., Isasi, Y., Grux, E., Blanco-Cuaresma, S., Arenou, F., Babusiaux, C., Belcheva, M., Drimmel, R., Jordi, C., Krone-Martins, A., Masana, E., Mauduit, J. C., Mignard, F., Mowlavi, N., Rocca-Volmerange, B., Sartoretti, P., Slezak, E., & Sozzetti, A. 2012a, AA, 543, A100

Robin, A. C., Marshall, D. J., Schultheis, M., & Reylé, C. 2012b, AA, 538, A106

Robin, A. C., Reylé, C., Derriere, S., & Picaud, S. 2003, AA, 409, 523

Robin, A. C., Reylé, C., Fliri, J., Czekaj, M., Robert, C. P., & Martins, A. M. M. 2014, AA, 569, A13

Rocha-Pinto, H. J., Maciel, W. J., Scalo, J., & Flynn, C. 2000a, AA, 358, 850

Rocha-Pinto, H. J., Scalo, J., Maciel, W. J., & Flynn, C. 2000b, AA, 358, 869

Roman, N. G. 1950, ApJ, 112, 554

Roškar, R., Debattista, V. P., & Loebman, S. R. 2013, MNRAS, 433, 976

Roškar, R., Debattista, V. P., Quinn, T. R., Stinson, G. S., & Wadsley, J. 2008, ApJ, 684, L79

Ruphy, S., Epchtein, N., Cohen, M., Copet, E., de Batz, B., Borsenberger, J., Fouque, P., Kimeswenger, S., Lacombe, F., Le Bertre, T., Rouan, D., & Tiphene, D. 1997, AA, 326, 597

Ruphy, S., Robin, A. C., Epchtein, N., Copet, E., Bertin, E., Fouque, P., & Guglielmo, F. 1996, AA, 313, L21

Russeil, D. 2003, AA, 397, 133

Sabin, L., Parker, Q. A., Corradi, R. L. M., Guzman-Ramirez, L., Morris, R. A. H., Zijlstra, A. A., Bojii, I. S., Frew, D. J., Guerrero, M., Stupar, M., Barlow, M. J., Mora, F. C., Drew, J. E., Greimel, R., Groot, P., Irwin, M. J., Mampaso, A., Miszalski, B., Olguin, L., Phillipps, S., Garcia, M. S., Viironen, K., & Wright, N. J. 2014, MNRAS, 443, 3388

Sakamoto, T., Chiba, M., & Beers, T. C. 2003, AA, 397, 899

Sale, S. E. 2012, MNRAS, 427, 2119

Sale, S. E., Drew, J. E., Barentsen, G., Farnhill, H. J., Raddi, R., Barlow, M. J., Eisloffel, J., Vink, J. S., Rodriguez-Gil, P., & Wright, N. J. 2014, MNRAS, 443, 2907

Sale, S. E., Drew, J. E., Knigge, C., Zijlstra, A. A., Irwin, M. J., Morris, R. A. H., Phillipps, S., Drake, J. J., Greimel, R., Unruh, Y. C., Groot, P. J., Mampaso, A., & Walton, N. A. 2010, MNRAS, 402, 713

Salpeter, E. E. 1955, ApJ, 121, 161

Sánchez-Saavedra, M. L., Battaner, E., Guijarro, A., López-Corredoira, M., & Castro-Rodríguez, N. 2003, AA, 399, 457

Sanchez-Salcedo, F. J. 2006, MNRAS, 365, 555

Sanders, D. B., Solomon, P. M., & Scoville, N. Z. 1984, ApJ, 276, 182

Schlafly, E. F. & Finkbeiner, D. P. 2011, ApJ, 737, 103

Schlegel, D. J., Finkbeiner, D., & Davis, M. 1998, ApJ, 500, 525

Schönrich, R. & Binney, J. 2009, MNRAS, 396, 203

Seigar, M. S. & James, P. A. 1998, MNRAS, 299, 685

Sellwood, J. A. & Binney, J. 2002, MNRAS, 336, 785

Shane, W. W. & Bieger-Smith, G. P. 1966, Bull. Astron. Inst. Netherlands, 18, 263

Shapley, H. 1918, ApJ, 48, 154

Shapley, H. & Curtis, H. D. 1921, Bulletin of the National Research Council, 2, 171

Siegler, N., Close, L. M., Cruz, K. L., Martin, E. L., & Reid, I. N. 2005, ApJ, 621, 1023

Skrutskie, M. F., Cutri, R. M., Stiening, R., Weinberg, M. D., Schneider, S., Carpenter, J. M., Beichman, C., Capps, R., Chester, T., Elias, J., Huchra, J., Liebert, J., Lonsdale, C., Monet, D. G., Price, S., Seitzer, P., Jarrett, T., Kirkpatrick, J. D., Gizis, J. E., Howard, E., Evans, T., Fowler, J., Fullmer, L., Hurt, R., Light, R., Kopan, E. L., Marsh, K. A., McCallon, H. L., Tam, R., Van Dyk, S., & Wheelock, S. 2006, AJ, 131, 1163

Slater, C. T., Bell, E. F., Schlafly, E. F., Morganson, E., Martin, N. F., Rix, H.-W., Peñarrubia, J., Bernard, E. J., Ferguson, A. M. N., Martinez-Delgado, D., Wyse, R. F. G., Burgett, W. S., Chambers, K. C., Draper, P. W., Hodapp, K. W., Kaiser, N., Magnier, E. A., Metcalfe, N., Price, P. A., Tonry, J. L., Wainscoat, R. J., & Waters, C. 2014, ApJ, 791, 9

Smartt, S. J. & Rolleston, W. R. J. 1997, ApJ, 481, L47

Smith, L., Lucas, P. W., Bunce, R., Burningham, B., Jones, H. R. A., Smart, R. L., Skrzypek, N., Rodriguez, D. R., Faherty, J., Barentsen, G., Drew, J. E., Andrei, A. H., Catalan, S., Pinfield, D. J., & Redburn, D. 2014, MNRAS, 443, 2327

Sofue, Y. & Fujimoto, M. 1983, ApJ, 265, 722

Soubiran, C., Bienaymé, O., & Siebert, A. 2003, AA, 398, 141

Spitzer, Lyman, J. & Schwarzschild, M. 1953, ApJ, 118, 106

Stark, A. A. & Lee, Y. 2005, ApJ, 619, L159

Stock, D. J. & Barlow, M. J. 2010, MNRAS, 409, 1429

Stoughton, C., Lupton, R. H., Bernardi, M., Blanton, M. R., Burles, S., Castander, F. J., Connolly, A. J., Eisenstein, D. J., Frieman, J. A., Hennessy, G. S., Hindsley, R. B., Ivezić, v., Kent, S., Kunszt, P. Z., Lee, B. C., Meiksin, A., Munn, J. A., Newberg, H. J., Nichol, R. C., Nicinski, T., Pier, J. R., Richards, G. T., Richmond, M. W., Schlegel, D. J., Smith, J. A., Strauss, M. A., SubbaRao, M., Szalay, A. S., Thakar, A. R., Tucker, D. L., Vanden Berk, D. E., Yanny, B., Adelman, J. K., Anderson, Jr., J. E., Anderson, S. F., Annis, J., Bahcall, N. A., Bakken, J. A., Bartelmann, M., Bastian, S., Bauer, A., Berman, E., Böhringer, H., Boroski, W. N., Bracker, S., Briegel, C., Briggs, J. W., Brinkmann, J., Brunner, R., Carey, L., Carr, M. A., Chen, B., Christian, D., Colestock, P. L., Crocker, J. H., Csabai, I., Czarapata, P. C., Dalcanton, J., Davidsen, A. F., Davis, J. E., Dehnen, W., Dodelson, S., Doi, M., Dombeck, T., Donahue, M., Ellman, N., Elms, B. R., Evans, M. L., Eyer, L., Fan, X., Federwitz, G. R., Friedman, S., Fukugita, M., Gal, R., Gillespie, B., Glazebrook, K., Gray, J., Grebel, E. K., Greenawalt, B., Greene, G., Gunn, J. E., de Haas, E., Haiman, Z., Haldeman, M., Hall, P. B., Hamabe, M., Hansen, B., Harris, F. H., Harris, H., Harvanek, M., Hawley, S. L., Hayes, J. J. E., Heckman, T. M., Helmi, A., Henden, A., Hogan, C. J., Hogg, D. W., Holmgren, D. J., Holtzman, J., Huang, C.-H., Hull, C., Ichikawa, S.-I., Ichikawa, T., Johnston, D. E., Kauffmann, G., Kim, R. S. J., Kimball, T., Kinney, E., Klaene, M., Kleinman, S. J., Klypin, A., Knapp, G. R., Korienek, J., Krolik, J., Kron, R. G., Krzesinski, J., Lamb, D. Q., Leger, R. F., Limmongkol, S., Lindenmeyer, C., Long, D. C., Loomis, C., Loveday, J., MacKinnon, B., Mannery, E. J., Mantsch, P. M., Margon, B., McGehee, P., McKay, T. A., McLean, B., Menou, K., Merelli, A., Mo, H. J., Monet, D. G., Nakamura, O., Narayanan, V. K., Nash, T., Neilsen, Jr., E. H., Newman, P. R., Nitta, A., Odenkirchen, M., Okada, N., Okamura, S., Ostriker, J. P., Owen, R., Pauls, A. G., Peoples, J., Peterson, R. S., Petravick, D., Pope, A., Pordes, R., Postman, M., Prosapio, A., Quinn, T. R., Rechenmacher, R., Rivetta, C. H., Rix, H.-W., Rockosi, C. M., Rosner, R., Ruthmansdorfer, K., Sandford, D., Schneider, D. P., Scranton, R., Sekiguchi, M., Sergey, G., Sheth, R., Shimasaku, K., Smee, S., Snedden, S., Stebbins, A., Stubbs, C., Szapudi, I., Szkody, P., Szokoly, G. P., Tabachnik, S., Tsvetanov, Z., Uomoto, A., Vogeley, M. S., Voges, W., Waddell, P., Walterbos, R., Wang, S.-i., Watanabe, M., Weinberg, D. H., White, R. L., White, S. D. M., Wilhite, B., Wolfe, D., Yasuda, N., York, D. G., Zehavi, I., & Zheng, W. 2002, AJ, 123, 485

Toomre, A. 1981, In: The structure and evolution of normal galaxies; Proceedings of the Advanced Study Institute, , 111

Toomre, A. & Toomre, J. 1972, ApJ, 178, 623

Tosi, M., Pulone, L., Marconi, G., & Bragaglia, A. 1998, MNRAS, 299, 834

Trippe, S., Gillessen, S., Gerhard, O. E., Bartko, H., Fritz, T. K., Maness, H. L., Eisenhauer, F., Martins, F., Ott, T., Dodds-Eden, K., & Genzel, R. 2008, AA, 492, 419

Trujillo, I. & Pohlen, M. 2005, ApJ, 630, L17

Trumpler, R. J. 1930, PASP, 42, 214

Tsikoudi, V. 1980, ApJS, 43, 365

Turon, C., Primas, F., Binney, J., Chiappini, C., Drew, J. E., Helmi, A., Robin, A. C., & Ryan, S. G. 2008, ESA-ESO Working Group reports,

Twarog, B. A. 1980, ApJ, 242, 242

Vallée, J. P. 1995, ApJ, 454, 119

—. 2008, AJ, 135, 1301

—. 2014, AJ, 148, 5

Vallenari, A., Pasetto, S., Bertelli, G., Chiosi, C., Spagna, A., & Lattanzi, M. 2006, AA, 451, 125

van der Kruit, P. C. 1986, AA, 157, 230

—. 1988, AA, 192, 117

van der Kruit, P. C. & Searle, L. 1981, AA, 95, 116

—. 1982, AA, 110, 61

Vanhollebeke, E., Groenewegen, M. A. T., & Girardi, L. 2009, AA, 498, 95

Verbeek, K., Groot, P. J., Nelemans, G., Scaringi, S., Napiwotzki, R., Drew, J. E., Steeghs, D., Casares, J., Corral-Santana, J. M., Gansicke, B. T., Gonzalez-Solares, E., Greimel, R., Heber, U., Irwin, M. J., Knigge, C., Wright, N. J., & Zijlstra, A. A. 2013, MNRAS, 434, 2727

Viironen, K., Mampaso, A., Corradi, R. L. M., Rodríguez, M., Greimel, R., Sabin, L., Sale, S. E., Unruh, Y., Delgado-Inglada, G., Drew, J. E., Giammanco, C., Groot, P., Parker, Q. A., Sokoloski, J., & Zijlstra, A. 2009, AA, 502, 113

Weaver, H. & Williams, D. R. W. 1973, AAS, 8

Wegg, C. & Gerhard, O. 2013, MNRAS, 435, 1874

Weinberg, M. D. & Blitz, L. 2006, ApJ, 641, L33

Westerhout, G. 1957, Bull. Astron. Inst. Netherlands, 13, 201

White, S. D. M. & Rees, M. J. 1978, MNRAS, 183, 341

Wielen, R. 1977, AA, 60, 263

Williams, B. F., Dalcanton, J. J., Dolphin, A. E., Holtzman, J. A., & Sarajedini, A. 2009, ApJ, 695, L15

Williams, M. E. K., Steinmetz, M., Sharma, S., Bland-Hawthorn, J., de Jong, R. S., Seabroke, G. M., Helmi, A., Freeman, K. C., Binney, J., Minchev, I., Bienaymé, O., Campbell, R., Fulbright, J. P., Gibson, B. K., Gilmore, G. F., Grebel, E. K., Munari, U., Navarro, J. F., Parker, Q. A., Reid, W., Siebert, A., Siviero, A., Watson, F. G., Wyse, R. F. G., & Zwitter, T. 2011, ApJ, 728, 102

Witham, A. R., Knigge, C., Drew, J. E., Greimel, R., Steeghs, D., Gänsicke, B. T., Groot, P. J., & Mampaso, A. 2008, MNRAS, 384, 1277

Wood, M. A. & Oswalt, T. D. 1998, ApJ, 497, 870

Wouterloot, J. G. A., Brand, J., Burton, W. B., & Kwee, K. K. 1990, AA, 230, 21

Wright, N. J., Wesson, R., Drew, J. E., Barentsen, G., Barlow, M. J., Walsh, J. R., Zijlstra, A., Drake, J. J., Eisloffel, J., & Farnhill, H. J. 2013, MNRAS, 437, L1

Wright, T. 1750, An Original Theory Or New Hypothesis of the Universe, Founded Upon the Laws of Nature, and Solving by Mathematical Principles the General Phænomena of the Visible Creation; and Particularly the Via Lactea. (London: Chapelle)

Wyse, R. F. G. & Gilmore, G. 1986, AJ, 91, 855

Yanny, B., Newberg, H. J., Grebel, E. K., Kent, S., Odenkirchen, M., Rockosi, C. M., Schlegel, D. J., Subbarao, M., Brinkmann, J., Fukugita, M., Ivezić, v., Lamb, D. Q., Schneider, D. P., & York, D. G. 2003, ApJ, 588, 824

Yasuda, N., Fukugita, M., Narayanan, V. K., Lupton, R. H., Strateva, I., Strauss, M. A., Ivezić, v., Kim, R. S. J., Hogg, D. W., Weinberg, D. H., Shimasaku, K., Loveday, J., Annis, J., Bahcall, N. A., Blanton, M., Brinkmann, J., Brunner, R. J., Connolly, A. J., Csabai, I., Doi, M., Hamabe, M., Ichikawa, S.-I., Ichikawa, T., Johnston, D. E., Knapp, G. R., Kunszt, P. Z., Lamb, D. Q., McKay, T. A., Munn, J. A., Nichol, R. C., Okamura, S., Schneider, D. P., Szokoly, G. P., Vogeley, M. S., Watanabe, M., & York, D. G. 2001, AJ, 122, 1104

Yoachim, P. & Dalcanton, J. J. 2006, AJ, 131, 226

Yuan, T.-T., Kewley, L. J., Swinbank, A. M., Richard, J., & Livermore, R. C. 2011, ApJ, 732, L14

Yusifov, I. 2004, The Magnetized Interstellar Medium, , 165

Zaritsky, D., Shectman, S. A., Thompson, I., Harris, J., & Lin, D. N. C. 1999, AJ, 117, 2268