Mon. Not. R. Astron. Soc. **000**, 1–16 (2005) Printed 10 February 2006 (MN LaTeX style file v2.2)

The UKIRT Infrared Deep Sky Survey (UKIDSS) Draft : 2006-02-10

A.Lawrence[1], S.J.Warren,[2] R.Jameson[3], N.C.Hambly[1], P.Lucas[4], A.Edge[5], O.Almaini[6], A.Adamson[7], S.Dye[9], S.Foucaud[6], P.Hewett[8], P.Hirst[7] S.Hodgkin[8], M.Irwin[8], N.Lodieu[3], C.Simpson[5], I.Smail[5]

[1] *Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ*
[2] *Blackett Laboratory, Imperial College of Science Technology and Medicine, Prince Consort Rd, London SW7 2AZ*
[3] *Leicester*
[4] *Herts*
[5] *Durham*
[6] *Nottingham*
[7] *JAC Hawaii*
[8] *IOA Cambridge*
[9] *School of Physics and Astronomy, Cardiff University, 5 The Parade, Cardiff, CF24 3YB, U.K.*

Accepted XXXX. Received XXXX ; in original form XXX

ABSTRACT

We describe the goals, design, and implementation of the UKIRT Infrared Deep Sky Survey (UKIDSS), which began in May 2005 and will take seven years to complete. UKIDSS is being carried out using the new UKIRT Wide Field Camera (WFCAM; Casali et al 2006), which has the largest grasp of any IR instrument to date. It is a portfolio of five survey components covering various combinations of the filter set ZYJHK. The Large Area Survey, the Galactic Cluster Survey, and the Galactic Plane Survey cover approximately 7000 square degrees to a depth of K$\sim$18.5; the Deep Extragalactic Survey covers 35 square degrees to K$\sim$21, and the Ultra Deep Survey covers 0.77 square degrees to K$\sim$23. Each of these five surveys independently is over an order of magnitude larger in effective volume than the 2MASS survey. The prime aim of UKIDSS is to provide a long term astronomical legacy database; the design is however driven by a series of specific goals – for example to find the nearest and faintest sub-stellar objects; to break the z=7 quasar barrier; to determine the epoch of re-ionisation; to determine the substellar mass function; to discover Population II brown dwarfs, if they exist; to measure the growth of structure from z=3 to the present day; to determine the epoch of spheroid formation; and to map the Milky Way through the dust, to several kpc. The survey data are being uniformly processed and both images and catalogues are being made available through a fully queryable user interface - the WFCAM Science Archive (WSA : http://surveys.roe.ac.uk/wsa). The data are being made available in a series of staged releases. The data are immediately public to astronomers in all ESO member states, and available to the world after eighteen months. Before the formal survey began, UKIRT and the UKIDSS consortium collaborated in obtaining and analysing a series of small science verification (SV) projects to complete the commissioning of the camera. We show some results from these SV projects in order to demonstrate the likely power of the eventual complete survey.

Key words: surveys, infrared: general

1 INTRODUCTION

The UKIRT Infrared Deep Sky Survey (UKIDSS) is the most significant step forward in JHK sky surveys since the Two Micron All Sky Survey Project (2MASS; ref), and can be considered the near-infrared counterpart of the Sloan Digital Sky Survey (SDSS; ref). It does not cover the whole sky, but is many times deeper than 2MASS. It is in fact not a single survey but a survey programme combining a set of five survey components of complementary combinations of depth and area, covering several thousand square degrees to $K \sim 18.5$, 35 square degrees to $K \sim 21$, and 0.77 square degrees to $K \sim 23$. The survey uses the new Wide Field Camera (WFCAM) on the 3.8m United Kingdom Infra-red Telescope (UKIRT). WFCAM has an instantaneous field of view of 0.21 square degrees, considerably larger than any previous IR camera, along with a pixel size of 0.4 arcsec. It is this combination of large telescope, large field of view, and good image quality, that makes such an ambitious survey possible. The various surveys employ up to five filters $ZYJHK$ covering the wavelength range $0.83 - 2.37\,\mu$m and extend over both high and low Galactic latitude regions of the sky. (The new Y band is explained below). The survey began on 2005 May 13, and is expected to take seven years to complete. The survey is being carried out by a private consortium but is intended to be fully public with no proprietary rights for the consortium. Data will be released quasi-continuously.

1.1 Origins and nature of project

The UKIDSS survey concept first emerged in 1998 while making the funding case for the WFCAM instrument itself, but eventually became a formal refereed proposal to the UKIRT Board in March 2001, submitted by a consortium of 61 UK astronomers. This included a commitment to making data available immediately to all UK astronomers, plus specified individual Japanese consortium members, and available to the world after a year or two . (See (**?**) for the final data release policy.) Later, during the UK's entry in to ESO, it was agreed that astronomers in all ESO member states would have the same data rights as UK astronomers, and at the same time, membership of the consortium was extended to any interested European astronomers. Consortium membership now stands at XXX. Note that individual astronomers are members, not their institutions.

The project is unusual compared to previous large survey projects, being neither private, nor conducted by a public body on behalf of the community. UKIDSS relies on the separate existence of three things. (i) The UKIRT observatory, operated as part of the UK's Joint Astronomy Centre (JAC). (ii) The WFCAM instrument, built at the Astronomy Technology Centre (ATC) at the Royal Observatory Edinburgh, as a funded PPARC project. Note that WFCAM was built as a common user instrument to be part of the UKIRT suite of instruments. The UKIDSS consortium is essentially the largest single user. (iii) The pipeline and archive development project, run by the Cambridge Astronomy Survey Unit (CASU) and the Edinburgh Wide Field Astronomy Unit (WFAU), and funded by several different PPARC grants. This data processing development is part of the VISTA Data Flow System (VDFS) project, with the WFCAM pipeline and archive being seen as an intermediate step. Note this data processing project deals with all the WFCAM data, not just the UKIDSS data.

The aim of the UKIDSS consortium is then to produce the scientific design for the survey, to win the telescope time necessary, to plan the implementation of the survey, liaising with the other bodies above, and finally to staff the observing implementation. A number of individuals in ATC, JAC, CASU and WFAU are also members of UKIDSS, so that the liaison with the camera construction and data processing projects, as well as telescope operations, has been well motivated. A clear relationship with UKIDSS has been built into each of these projects. For example the science requirements document for the pipeline and archive emerged from consultation with UKIDSS; and the commissioning schedule for WFCAM include a "science verification" phase following standard tests, in which the survey implementation described in (**?**) could be tested and refined.

1.2 Technical reference papers

This paper is one of a set of five which provide the reference technical documentation for UKIDSS. It summarises the scope, goals, and overall design of the survey, along with a brief discussion of implementation methods, and presentation of early "science verification" data. The other four papers, described briefly below, are Casali et al. (2006), Hewett et al. (2006), Irwin et al. (2006) and Hambly et al. (2006).

Casali et al. (2006) describe the survey instrument, WFCAM. The camera has four Rockwell Hawaii-II 2048^2 PACE arrays, with pixel scale $0.4''$, giving a solid angle of $0.21\,\mathrm{deg}^2$ per exposure. At the time of commissioning, 2004 November, the instrument *étendue*[1] of $2.38\,\mathrm{m}^2\,\mathrm{deg}^2$ was the largest of any near-infrared imager in the world. The Canada France Hawaii Telescope WIRCam instrument (Puget et al. 2004) covers a solid angle of $0.1\,\mathrm{deg}^2$ per exposure giving an *étendue* of $1.11\,\mathrm{m}^2\,\mathrm{deg}^2$. WFCAM is likely to remain as the near–infrared imager with the largest *étendue* in the world until completion of the near–infrared camera for VISTA (Dalton et al. 2004).

Irwin et al. (2006) and Hambly et al. (2006) describe respectively the data pipeline and science archive system. For typical observing programmes, WFCAM will accumulate data at the rate of approximately 1TB per week. All WFCAM data are processed by an automated pipeline, described by Irwin et al. (2006). The pipeline flat-fields the data, subtracts the counts from the background sky, detects and parameterises objects, and performs the photometric and astrometric calibrations. The reduced frames and catalogues are ingested into the WFCAM Science Archive (WSA), described by Hambly et al. (2006). A process of curation results in seamless images and catalogues, matched across bands. Access to the data is through a set of web pages that allow various flexible usage modes, including raw SQL commands for sophisticated searches.. It is expected that nearly all science analysis of UKIDSS will be initiated through the WSA.

The photometric system is described in Hewett et al (2006). The survey uses five broadband filters, $ZYJHK$. The JHK passbands are as close as possible to the MKO system; the Z passband is similar to the SDSS z' passband,

[1] product of telescope collecting area, and solid angle of instrument field of view, sometimes called *grasp*

but has a cleaner red tail. The Y passband is a new one centred at $0.97\,\mu$m aimed at discriminating substellar objects and high redshift quasars as cleanly as possible. Hewett et al present the measured passband transmisssions, and use synthetic colours of various classes of astronomical object to produce expected colour equations between certain WFCAM, SDSS, and 2MASS filters. A later paper (Hodgkin et al. 2006) will report on the colour equations determined on the sky from standard star observations.

1.3 Plan of paper

This paper begins with a description of the science goals of UKIDSS, and some illustrations of how the survey design will achieve them. We then describe the practical implementation of the survey - the tiling and jittering patterns, exposure times, calibration plan, and so on, in the context of the camera properties and the UKIRT operating procedures and software. Next we consider the detailed design of the individual survey components - areas, field selection, filters and scheduling. We also describe the staging of UKIDSS in a two year plan and final seven year plan. We then summarise the data processing arrangements, which as described above are pursued as a formally separate project, but which of course are crucial to the scientific success of UKIDSS. Following this we present some example data and simple analysis from the science verification phase of UKIDSS, and point towards the expected final data quality. Finally we describe the plan for publication of the data, and provide links to more detailed information about UKIDSS, WFCAM, and the science archive.

2 SCIENCE GOALS

2.1 General goals

The primary goal of UKIDSS is to produce IR sky atlases as a fundamental resource of lasting significance analogous to the various Schmidt photographic sky surveys of the 1970s and onwards (REFS), and to the SDSS survey of modern times (York et al 2000). None of our survey components covers the whole sky, but nonetheless each component deserves the term "atlas", as the volume surveyed, and the number of objects detected, are comparable to the above optical surveys, and each survey maps out some significant part of the universe - the solar neighbourhood, the Milky Way, the local extragalactic universe, the universe at z=1, and the universe at z=3. Each of the component surveys is many times larger than any existing IR survey at comparable depth.

The strength of an atlas is of course its potential for multiple use over many years, but this general aim does not fix the best combination of area, depth, and wavelength coverage. In a euclidean volume, for a given total time, a shallow survey always produces a larger sample size than a deep one; but specific science goals often require a given depth, for example to detect galaxies at a given redshift; and for relatively deep surveys, neither the Milky Way nor the universe at large are euclidean volumes. As we cannot predict all future uses of the UKIDSS atlases, our most general strategy is to pursue a "wedding cake" strategy, dividing the time between a large shallow survey, a medium sized fairly

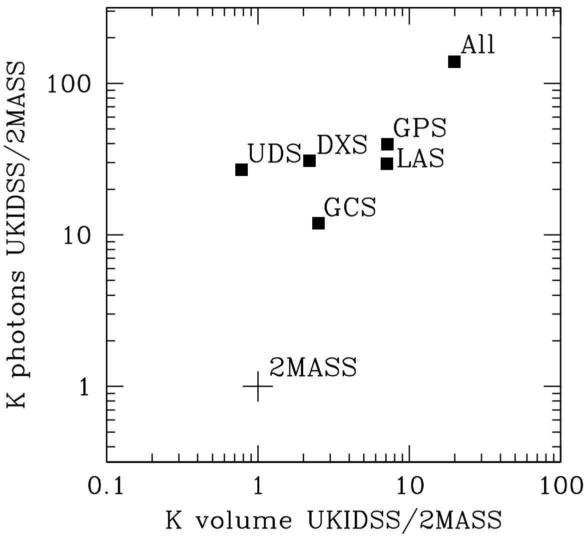

Figure 1. Illustration of the scope of the five UKIDSS survey components, and their sum, by comparison with 2MASS. The comparison is made in terms of expected number of photons and effective volume, for the K band, computed as described in the text

deep survey, and a small very deep survey. In the following subsections we summarise the scientific goals of each survey component; in section 4 we describe the design of each survey component - areas, depths, field locations, filters, implementation strategy - needed to achieve those goals.

The scope of the surveys is illustrated in an interesting way in Fig. 1. For sky-limited observations in the K band, say, the time to reach depth K is proportional to $10^{0.8K}$. One can therefore think of the quantity area$\times10^{0.8K}$ as being proportional to the number of photons collected. In a similar way, one can show that the quantity area$\times10^{0.6K}$ is proportional to the volume surveyed, for Euclidean space. The largest existing multiband near-ir survey in terms of both quantities is 2MASS. In Fig. 1 we have normalised the computed values for each of the five UKIDSS elements, to the 2MASS values, for the K band. Viewed in this way, each of the 5 surveys is between 10 and 30 times larger than 2MASS in terms of photons, and, except for the UDS, is a few to several times larger in volume. Summed over the whole programme, UKIDSS is 100 times larger than 2MASS in terms of photons, and 20 times larger in terms of volume.

2.2 Headline science goals

To develop a plan beyond this very general concept, we encouraged the formation of groupings within the consortium to promote distinct survey components and science goals. The whole consortium then debated the science goals and designs proposed by these groups, looked for overlaps, and made compromises until we felt we had a balanced strategy. (The resulting survey component designs are presented in Section 4.) As part of this scientific debate, we also agreed *headline science goals*, which were used to drive the specific designs of the survey componets. The most important of these specific goals are as follows :

- to find the nearest and faintest sub-stellar objects
- to break the z=7 quasar barrier
- to determine the epoch of re-ionisation
- to determine the substellar mass function
- to discover Pop II brown dwarfs, if they exist
- to construct a galaxy catalogue at z=1 as large as the SDSS catalogue
- to measure the growth of structure and bias from z=3 to the present day
- to determine the epoch of spheroid formation
- to clarify the relationship between quasars, ULIRGs, and galaxy formation
- to map the Milky Way through the dust, to several kpc
- to increase the number of known Young Stellar Objects by an order of magnitude, including rare types such as FU Orionis stars

2.3 Goals of the Large Area Survey (LAS)

The Large Area Survey (LAS) aims to map as large a fraction of the Northern Sky as feasible (4000 square degrees) within a few hundred nights, which when combined with the SDSS, produces an atlas covering almost an order of magnitude in wavelength. Furthermore a huge number of objects will already have spectroscopic data from the SDSS project, making an unparalleled dataset. The basic shallow survey reaches J=19.6, H=18.8, K=18.4, but we also plan a second pass in the J-band to detect proper motions of low mass objects and thus their kinematic distances, so that the final J depth is J=20.0. In addition we are using a new filter we have called Y, covering 0.97 to 1.07 microns, specifically designed to detect extremely high redshift (z=7) quasars, and to distinguish them from very low mass stars.

The Large Area Survey will produce a catalogue of a million galaxies with colours and spectra, and 10 million galaxies with photometric redshifts; will detect thousands of rich clusters out to z=1; will find ten times more brown dwarfs than 2MASS, will probe to much fainter objects, and can get statistical ages and masses from kinematics; and will produce a complete sample of 10,000 bright quasars, including reddened quasars, using the K excess method. Objects will be found using two colour diagrams such as the one shown in Figure 2

We are particularly driven however by three especially exciting prospects. (i) A search for the the nearest and smallest objects in the solar neighbourhood.Our survey is deep enough to detect brown dwarfs and young free floating planets with as little as 5 Jupiter masses out to distances of tens of parsecs.We can expect to find brown dwarfs even cooler than T dwarfs,T_e <700K,a new spectral class tentatively named Y dwarfs (Leggett et al 2005). (ii) The combination of IR and optical colours, and large expected proper motions, will allow us to find halo brown dwarfs if they exist, testing the universality of star formation processes, and the formation history of the Milky Way. (iii) We hope to find quasars at $z = 7$ and to detect the epoch of re-ionisation. SDSS have found z=5-6 quasars by "z′ drop-out". Beyond z=6 quasars become rapidly redder, indistinguishable from brown dwarfs in standard colours, and too faint to be in the SDSS z' survey. We therefore propose a survey in the new Y filter to match our JHK survey. Extrapolating popular evolution functions, we expect to find 10 quasars in the range z=6-7 and 4 in the range z=7-8. Figure 3 illustrates how the UKIDSS filter set can distinguish cleanly between very cool brown dwarfs and very high redshift quasars.

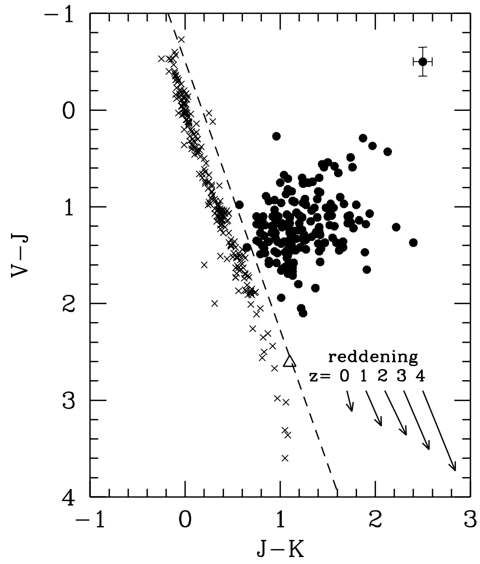

Figure 2. OLD VERSION OF FIGURE : TO BE REPLACED Illustration of the "K excess" quasar selection technique. In the VJK diagram, quasars stand out extremely clearly from the stellar sequence at any redshift, or with any reddening. We therefore expect to produce large clean samples of both naked and obscured quasars very efficiently.

2.4 Goals of the Galactic Plane Survey (GPS)

The Galactic Plane Survey aims to map half of the Milky Way to within a latitude of ±5°. Given the declination constraints of UKIRT, we can survey l=15-107 and l=141-230. Owing to interest in recent results from multi-waveband observations of the Galactic Centre region (eg. Wang et al.2002; Hasegawa et al.1998) the survey region has been extended south to include the l=-2 to 15 region in a narrow strip at $b = \pm 2$. Despite the large survey area it is possible to reach a 5−σ depth of J=20.0, H=19.1, K=19.0. This is deep enough to probe the IMF down to M$\sim$ 0.05 M$_\odot$ in star formation regions within 2 kpc of the sun, to detect stars below the main sequence turn off in the galactic bulge, and to detect luminous objects such as OB stars and post-AGB stars across the whole galaxy. The K band depth will be built up at three separate epochs (each with depth of K=18.4) in order to detect highly variable objects and locate nearby objects through their proper motions. In addition we will make a three epoch narrow band H_2 survey in a 300 square degree

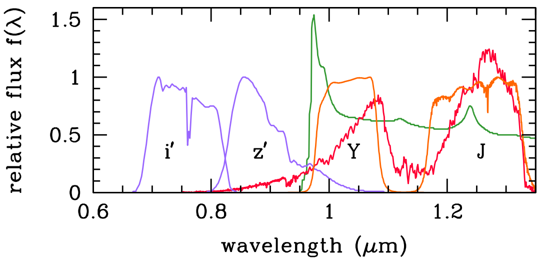

Figure 3. Plot illustrating the usefulness of the Y band for finding cool brown dwarfs and quasars of very-high redshift ($z > 6.4$). Filter curves are total system throughput (above atmosphere to detector), normalised to the peak, for the SDSS i' and z' bands (from Fan et al., 2001), and the WFCAM Y and J bands (from Hewett et al., 2006). The red curve is the spectrum of the T6 brown dwarf SDSS J162414.37+002915.6 (from Leggett et al., 2000), and the green curve is a model spectrum of a quasar at z=7. High-redshift quasars and brown dwarfs may be identified by the very sharp spectral discontinuity in moving from the optical (i', z') to the near-infrared (Y, J), while the quasars may be distinguished from the brown dwarfs, because they are somewhat bluer in Y-J colour.

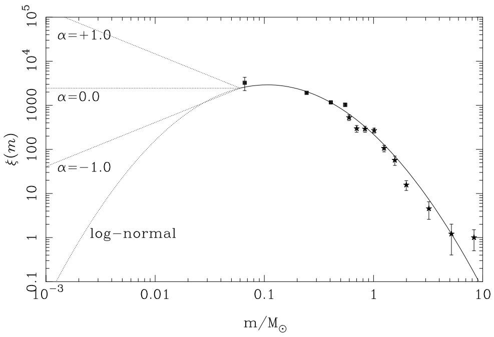

Figure 4. Various extrapolations (dotted lines) of the Pleiades mass function (after Hambly et al. 1999) illustrating uncertainties in the behaviour of the MF in the brown dwarf regime. The most recent surveys have probed the mass range 0.01 to 0.1 solar masses, using a variety of techniques, and have produced a range of different forms of the MF using heterogeneous datasets with varying degrees of completeness. The GCS aims to obtain maximal completeness in ten targets to settle the questions as to the form and universality (or otherwise) of the MF in the BD regime.

area of the Taurus-Auriga-Perseus molecular cloud complex (with JHK data also). This survey area closely follows the area molecular region detected by Ungerechts & Thaddeus (1986).

Like the high latitude LAS, the GPS has its prime importance as a fundamental resource for future astronomy. The survey depths are close to being confusion limited, so this survey is unlikely to be superseded until a high resolution wide angle camera is placed in space. We expect to detect 10^{8-9} sources in total. However, there are a number of immediately expected achievements, which will be achieved in combination with data from multiwaveband galactic surveys from many facilities. There will be particular benefit from surveys planned or in progress with the SPITZER space telescope (infrared), CHANDRA and XMM-Newton (X-ray), the VLA (radio, especially the 5 GHz CORNISH survey), HERSCHEL and SCUBA-2 (submm) and ASTRO-F (far infared). The following list illustrates some of the expected results. (1) An increase in the number of known Young Stellar Objects (YSOs) by an order of magnitude and measurement of the duration of the YSO phase as a function of mass and environment. (2) Star formation regions will be mapped throughout the Milky Way, measuring the environmental dependence of the IMF to low masses and estimating the overall star formation rate of the galaxy. (3) Rare or brief duration variables will be found in significant numbers, aiding the study of phenomena such as FU Orionis variability and helium flash events. (4) Thousands of evolved objects such as post-AGB stars and planetary nebulae will be found, a huge increase over previous samples. (5) Many stellar populations will be mapped to large distances through the Milky Way extinction, measuring the scale height versus stellar type and mapping poorly measured regions of the arms and warp. (6) The IR counterparts of hundreds of X-ray binaries and thousands of CVs will be identified and source lists provided for regions yet to be mapped by X-ray satellites.

2.5 Goals of the Galactic Cluster Survey (GCS)

The Galactic Cluster Survey (GCS) aims to survey ten large open star clusters and star formation associations, covering a total of 1600 sq.deg. using the standard single pass depth plus a second pass in K for proper motions, giving a depth of Z=19.9, Y=19.7, J=19.6, H=18.8, K=18.8. The targets are all relatively nearby, are at intermediate to low Galactic latitudes and are several degrees across.

The GCS is the most targeted of our surveys, being aimed at the crucial question of the sub-stellar initial mass function (IMF). Our current knowledge of the IMF is illustrated in Fig. 4, along with a variety of possible extrapolations. The stellar IMF is well determined down to the brown dwarf boundary but is much less well known below, and it is not known whether the IMF as a whole is universal or not (the current state of research into ultra low–mass star formation is described in Martín & Magazzù 2006). The mass limit reached varies somewhat from cluster to cluster, but is typically around $M_L \sim 30 M_J$. The number of objects expected to be detected in the range $M_L - -M_L + 10 M_J$ ranges from 100 to 3000 for the range of possible mass function models, showing how well we will constrain the IMF compared to current knowledge.

To find extreme objects — the very nearest examples, the lowest mass objects — the large area survey is better. But to measure the substellar IMF, one wants to target the 30 – 100 M_J region, and to obtain masses one needs both a distance and an age, for which mapping clusters are ideal. This approach has of course already been started (e.g. Moraux et al. 2006 and references therein). Our survey improves on current studies not by going deeper but by collecting much larger numbers, and examining objects formed in environments having a range of ages and metallicities, to examine the question of universaility.

2.6 Goals of the Deep Extragalactic Survey (DXS)

The Deep Extragalactic Survey (DXS) aims to map 35 sq.deg. of sky to a 5-σ point-source sensitivity of J=23 and K=21 in four carefully selected, multi-wavelength survey areas. The primary aim of the survey is to produce a photometric galaxy sample at a redshift of 1–2, within a volume comparable to that of the SDSS, selected in the same passband (rest frame optical). Figure 5 shows measured K magnitude versus redshift for galaxies in the Hubble Deep Field (REF). This shows that to achieve a sample such that the median redshift is $z \sim 1$ requires a depth of K$\sim$21. Such a sample will allow a direct test of the evolution of the galaxy population and determine how galaxies of different types (passive, star-forming, AGN) trace large scale structure (their bias). Each of these properties can be predicted from cosmological simulations so the DXS will set tight constraints on these models in volumes less susceptible to cosmic variance than previous, narrow-angle surveys at this redshift. The sample will also enable the selection of clusters of galaxies in this redshift range, where cosmological models predict numbers to be sensitive to the total mass density of the Universe, Ω_0.

The number of deep, multi-wavelength survey fields has increased dramatically in the past 5 years with the up-grade of existing facilites (e.g. Megacam on CFHT and VIMOS on the VLT) and new satellite missions (e.g. *Spitzer*, *GALEX*, *XMM-Newton* and *Chandra*). Each of these facilities has current surveys of 2–40 sq.deg. of contiguous area to levels where many of the counterparts are intrinsically faint in the optical (R>24) due to a combination of redshift and/or intrinsic dust obscuration but relatively red (R-K>4). Therefore deep NIR imaging (K$\sim$21) over tens of square degrees is required to fully characterise these dusty and/or distant objects. Looking ahead to the end of this decade and the completion of UKIDSS, there will be many more complementary surveys on these scales in other wavebands such as the far-infrared (*Herschel*), sub-mm (SCUBA-2), radio (EVLA and eMerlin) and Sunyaev-Zel'dovich (SZ telescope and AMI). The legacy potential of the DXS was a key driver to the science case and the field selection (see section 4.4).

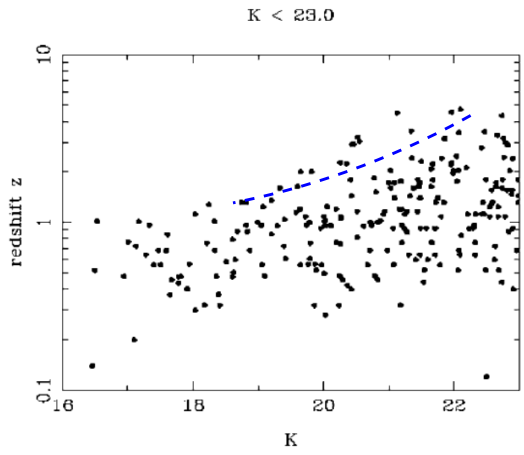

Figure 5. EXPECTING REPLACEMENT FIGURE ? K magnitudes versus redshift for galaxies in the Hubble Deep Field, measured by XXX.

2.7 Goals of the Ultra Deep Survey (UDS)

The Ultra Deep Survey (UDS) aims to map 0.77 sq. degrees to a 5σ point-source sensitivity of J=25, H=24, K=23. Such depths are required to reach typical L^* galaxies at $z = 3$. Covering an area one hundred times larger than any previous survey to these depths, this will provide the first large-volume map of the high-redshift Universe (30×30Mpc by 2 Gpc deep at $2 < z < 4$).

Deep near-infrared surveys are crucial for obtaining a more complete census of the Universe at these epochs. In particular, galaxies which are reddened by dust or those which appear red due to old stellar populations may be completely missed by standard optical surveys. From the UDS we anticipate over 10,000 galaxies at $z > 2$, allowing detailed studies of the luminosity functions, clustering and multi-wavelength SEDs over a large, representative volume. A major goal, together with the DXS and local surveys, is to measure clustering as a function of stellar mass and cosmic time, which will provide very powerful tests of models for biased galaxy formation and the growth of structure.

The UDS is also designed to address one of the major unsolved problems in modern astronomy, which is to understand when the massive elliptical galaxies are formed. A key test will be to determine the co-moving number density of the most massive galaxies at various epochs, particularly at $z > 2$. This requires a combination of both depth and area which has previously been impossible to achieve. If the density of massive galaxies (more massive than local L^* ellipticals) is similar to that of today, we should see ~ 1000 per square degree. Current semi-analytic models predict an order of magnitude fewer. Our goal is to directly measure the build-up of this population over cosmic time.

The survey field chosen for the UDS is the Subaru/XMM Deep Field, which has a wide range of multiwavelengh data available, including deep radio observations from the VLA, submm mapping from SCUBA, mid-IR photometry from Spitzer, deep optical imaging from Subaru Suprimecam and deep X-ray observations from XMM-Newton. When combined, these will enable detailed studies of the relationship between black hole activity, dust-dominated ULIRGs and IR-selected massive galaxies using an unprecedented high-redshift sample.

3 IMPLEMENTATION WITH THE UKIRT WIDE FIELD CAMERA

3.1 General characteristics of telescope and camera

UKIDSS is implemented using the new Wide Field Camera (WFCAM) on the United Kingdom Infrared Telescope (UKIRT), which is operated by the Joint Astronomy Centre (JAC), an establishment of the UK's Particle Physics and Astrophysics Research Council (PPARC). General technical details for UKIRT are given on the JAC website (http://www.jach.hawaii.edu/UKIRT/). It is an infra-red dedicated 3.8m telescope operating at the summit of Mauna

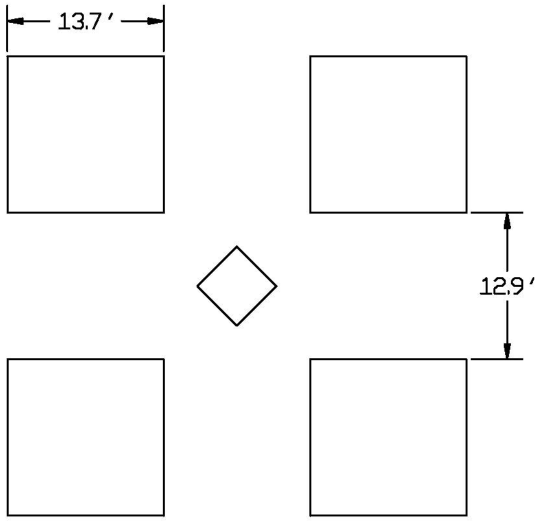

Figure 6. The WFCAM focal plane. The spacing between detectors is 94% of width of each detector. A sequence of four pointings therefore produces complete coverage for one "tile", plus a small overlap regions. The central diamond is the autoguider CCD. For details see Casali et al (2006)

Kea on Hawaii. Of particular importance is a tip-tilt secondary, which primarily removes dome and windshake effects on seeing, delivering close to free-atmosphere seeing (half arcsecond on many nights) across the whole WFCAM field of view. UKIRT operates as an open access telescope offering time through peer-reviewed competition. UKIDSS is essentially a single extremely large open access proposal. WFCAM is currently scheduled for 60% of UKIRT time, 220 nights per year. After removal of engineering time, and time allocated to the University of Hawaii, and Japan, an average of 167 nights per year is left, 80% of which, 134 nights per year, is devoted to UKIDSS, and the remaining time to other peer-reviewed programmes. (The latter are selected by the UK PATT system, open to world wide proposals).

WFCAM is described in detail by Casali et al. (2006). Here we summarise some key characteristics. WFCAM has an unusual design, with an array of IR detectors inside a long tube mounted above Cassegrain focus. The forward-Cassegrain Schmidt-like camera design makes possible a very wide field of view (40 arcmin) on a telescope not originally designed for this purpose. The camera has four 2048×2048 Rockwell Hawaii-II PACE arrays. The arrays have a physical pixel size of $0.4''$, which gives an instantaneous exposed field of view of $0.207\,\mathrm{deg}^2$ per exposure. The arrays are spaced by 0.94 detector widths. The focal plane coverage is illustrated in Figure 6

The WFCAM filters, the optical performance, and the detector efficiency are presented in Casali et al (2006). Hewett et al (2006) show the total system performance, derive synthetic colours for various types of astrophysical object, and provide links to pasband data. The JHK passbands are close to the standard MKO system; the Z passband is similar to the SDSS z' passband; the Y passband is new and was specially designed to optimise the selection of very high redshift quasars and very cool brown dwarfs. Final calibrations and colour equations for transformation to other systems will be derived from on-sky observations and wil be shown in a future paper (Hodgkin et al 2006).

3.2 Observing with WFCAM

WFCAM makes very short exposures on the sky. Coverage of the sky is then built up in several stages - by small jittering patterns at a fixed telescope pointing position; by macrostepping to make a filled in "tile"; and by accumulating sets of such tiles to gradually cover the sky, or revisiting tiles to build up depth.

Exposure and Readout. Each WFCAM array has an independent SDSU controller, reading out 32 channels (8 segments in each of 4 quadrants). The total readout time, and the reset time are each approximately 0.6 seconds, making an *exposure overhead* of 1.2 seconds. Each exposure makes a set of four data files which in later processing are later linked together as a *multi-frame*. Exposures of 5-10 seconds are long enough to be reasonably efficient, and for sky background to be larger than the readout noise. The main exception is in the Y and Z bands, where exposures of 20 seconds are needed for the sky noise to exceed readout noise. Exposures longer than this are not normally used, as an increasing number of stars in the field are saturated.

Jittering and micro-stepping. Small accurate telescope offset patterns at a fixed pointing position are used to improve WFCAM data. The first method is to use a jitter sequence with offsets equal to whole numbers of pixels, resulting in frames which can be co-added. The aim of such a *jitter pattern* is minimise the effects of bad pixels and other flat-fielding complications. A variety of jitter patterns can be used. The second kind of pattern is to offset by N+0.5 pixels, typically in a 2×2 pattern. The data are then interlaced (i.e. keeping the pixels independent) into a grid of pixel spacing 0.2 arcsecond, producing an image of size 4096×4096 pixels for each array. The aim of such a *microstepping pattern* is to improve image sampling - the WFCAM pixel size of 0.4 arcsecond is adequate for moderate seeing, but undersamples the expected seeing a significant fraction of the time.

Integration efficiency. The sequence of offset pattern exposures at a fixed telescope pointing makes up an integration on a particular piece of sky. The time to make a small offset is $\sim 2-3$ seconds. Together with the exposure overhead, the total *integration overhead* is therefore around 4 seconds. The minimum sensible exposure time is therefore $\sim$ 5 seconds, with an efficiency of $\sim$ 56%. For UKIDSS surveys we normally use exposures of 10 seconds, with an efficiency of $\sim$70%, beyond which there are diminishing returns in efficiency, but increasing disadvantages in for example saturating stars in the field. For sampling and/or cosmetic reasons, all UKIDSS surveys use at least 4 offset positions. The "standard shallow integration" is therefore 40 seconds. The data from offset sequences is normally interleaved and/or stacked to make a single multi-frame data file, which is the basic unit of the archived data.

Tiling. The WFCAM arrays are spaced by 0.94 detector widths. The sky could be potentially be covered in a variety of mosaic patterns, but a typical procedure would be to expose in a pattern of four macro-steps to make a complete filled-in "tile". Allowing for overlaps with adjacent tiles, the width of a single tile is then 3.88 detector widths, i.e. 0.883°, giving a solid angle of 0.78 sq.degs. Because of the overlaps, the effective *coverage efficiency* is 83%. The time between macrostep integrations (slew, stabilise, guide star lock) is ~ 15 seconds. For shallow surveys, where each pointing has typically 4 offset positions each with a 10 second exposure and a 4 second overhead, this means that a tile with 4 pointings spends 160 seconds exposing out of an elapsed time of 284 seconds, making a total *observing efficiency* of 56%. For deeper surveys, where many exposures are made at fixed pointing, the macro-step overhead is negligible, and the efficiency tends towards $\sim$70%.

Schedule Blocks and Survey Definition. UKIRT operates an automated flexible queuing system. Observers specify in advance a precise sequence of exposures, offset patterns, and filters at each of a list of pointing positions, which can be thought of as grouped into "tiles" as appropriate. A suitable list of such sequences is entered as a single "Minimum Schedulable Block (MSB)" intended to occupy roughly 20-60 minutes of telescope time. The MSBs also contains constraints that determine whether they can be observed - required seeing, sky brightness, etc. Calibration observations - twlight flats, standard stars, etc - are entered as independent MSBs. The MSBs are entered into a database which is queried during telescope operations in order to determine a priority for observation. Each MSB is then either completed or not, and if completed removed from the database. For UKIDSS, a Survey Definition Tool (SDT) is used to design the list of pointing positions for each survey, which are then grouped into MSBs, and likewise into smaller "projects" which help planning and monitoring of survey completion.

3.3 Survey calibration

The fundamental calibration for all WFCAM observations uses observations of UKIRT faint standard fields (REF). The WFCAM passbands are described in Hewett et al (2005). In Hodgkin et al (2006) we will describe empirically derived colour equations for transforming between the WFCAM natural system and other magnitude systems, and in particular the 2MASS system. Given the derived colour equations, calibration of each survey field using 2MASS stars in the field has proved very successful, and this is what we describe below. Note however that current observing practice is to observe UKIRT faint standard fields frequently through the night, providing an independent check when necessary.

All WFCAM data processed by the CASU pipeline are calibrated into the Vega system, using 2MASS stars as calibrators (but with photometry converted into the WFCAM natural system). A few reasons for basing a calibration on 2MASS are itemized :

- The 2MASS global calibration is accurate to better than 2% across the entire sky (Nikolaev et al. 2000).
- There are plenty of 2MASS stars in each WFCAM exposure which are not saturated.
- Measuring the calibrators at exactly the same time as the targets enables photometric calibration even during non-photometric conditions (limited by the uniformity of the cloud cover).

The step-by-step procedure for the conversion of measured fluxes to magnitudes is outlined here. For each WFCAM frame, 2MASS Point Source Catalogue stars (with error $\sigma_m(\mathrm{2MASS}) \leqslant 0.1$) which fall within the pawprint are identified. The photometry for each 2MASS source is transformed into the WFCAM ZYJHK system using empirically derived colour terms[2]. The WFCAM sources, with photometry extracted from a 1 arcsecond radius aperture, are matched against the 2MASS calibrators. A per-star zeropoint is calculated from the WFCAM aperture-corrected flux and the 2MASS magnitude as

$$ZP_i = m_i(\mathrm{2MASS}) + 2.5 \times \log10(\mathrm{counts_i}/\mathrm{exptime}) \qquad (1)$$

Also note that the WFCAM counts are corrected to account for the radial variation in the pixel scale[3]. The photometric zeropoint for the pawprint is then simply the median zeropoint over all unsaturated stars. Note that detector-to-detector differences in the zeropoint are largely removed by the gain correction in 2D processing (based on twilight flatfields), and residuals are measured to be $\leqslant$ 1%. A typical WFCAM pointing contains in the range 60-1000 2MASS calibrators, dependendent on Galactic Latitude, with a median of around 150 unsaturated 2MASS stars with $\sigma_{JHK}(2MASS) \leqslant 0.1$. We would therefore expect this 2MASS-based approach to calibrate the WFCAM data to $\sim$ 1%, and to be limited by the global systematics in 2MASS (see Nikolaev 2000), and the accuracy and robustness of the transformations between the filter systems (in particular the WFCAM J-band filter is somewhat narrower and redder than the 2MASS J-band filter), and accounting for the significantly different altitudes of the respective observatories.

To quantify this, we have analysed J, H and K photometry for a total of 600 measurements of 46 fiducial stars (drawn from the UKIRT faint standards, (Hawarden et al. 2001) measured on 19 photometric nights. We find that the RMS scatter for these measurements is better than 2% which meets the UKIDSS requirements for photometric accuracy (see Hodgkin et al. in prep). The quality of the Z-, Y- and narrowband-filter calibration has not yet been quantified.

4 SURVEY DESIGN

The design of the UKIDSS survey components was driven by a mixture of the legacy ambition, practical limitations, and specific science goals. The total size of the project was decided by a decision to continue long enough (seven years) to achieve a product of international significance and lasting value. The total time available was driven by UKIRT/WFCAM scheduling constraints. We thus arrived at a *seven year plan* totalling approximately a thousand nights. Seven years is however a long time to wait for science results; we therefore also designed an initial *two year plan* that would produce a self contained product and valuable science.

Table 1 summarises the design parameters of each of the five UKIDSS survey components. Figure 7 shows the location of the survey fields on the sky. The logic behind

[2] http://www.ast.cam.ac.uk/vdfs/docs/reports/sv/index.html
[3] http://www.ast.cam.ac.uk/vdfs/docs/reports/wfcam_photcal_report/index.html

Survey	Area	Filter	limit	t_{int}	t_{tot}	Nights
LAS	4028	Y	20.5	40s	367h	262 nts
	4028	J×2	20.0	80s	734h	
	4028	H	18.8	40s	367h	
	4028	K	18.4	40s	367h	
GPS	1800	J	20.0	80s	286h	186 nts
	1800	H	19.1	80s	286h	
	1800	K×3	19.0	120s	495h	
	300	$H_2\times$ 3	—	450s	237h	
GCS	933	Z	20.4	40s	86h	74 nts
	933	Y	20.5	40s	86h	
	933	J	19.7	40s	86h	
	933	H	18.8	40s	86h	
	933	K×2	18.7	80s	172h	
DXS	35	J	22.3	3.8h	415h	118 nts
	35	K	20.8	2.5h	287h	
UDS	0.77	J	24.8	209h	983h	296 nts
	0.77	H	23.8	174h	818h	
	0.77	K	22.8	58h	271h	
TOTAL						936 nts

Table 1. *Summary of expected final parameters for each survey.* Notes : *(i) Area is in square degrees. (ii) "J×2" implies that two passes of the whole area are made in that filter (iii) "Limit" is the Vega magnitude of a point source predicted to be detected at 5σ (iv)* t_{int} *is the accumulated integration time at each sky position. (v)* t_{tot} *is the number of hours required on-sky to copmplete the survey, allowing for expected exposure efficiency and mosaic efficiency. (vi) "Nights" is the number of nights required, allowing for calibration and average fraction of useable UKIRT time (70% of dark hours).*

this design, and some more detail about how the surveys are implemented, is described below for each survey component in turn. The LAS is described in a little more detail, to get some general points across.

4.1 Design of the Large Area Survey (LAS)

The main science goals of the LAS require as large a volume as possible, with increased area being a more efficient use of time than increased depth. However the survey rate is constrained by our requirement for multiple pass bands, by the need for exposures long enough to avoid inefficient observing, and by the need for jittering in order to improve cosmetic quality and/or spatial sampling.

The detailed implementation also depends on the weather constraints used. Because LAS uses a large fraction of the UKIDSS time, it would obviously not demand the best seeing. (This is partly why micro-stepping was not chosen, except for J-band). Colours are very important, and some objects are variable, which argues for doing all four bands when each sky position is visited; on the other hand, the Y and J observations require darker sky than the H and K observations. MSBs were then grouped so that H and K would be done together, and Y and J done separately. However, the queue is monitored and adjusted to try to make sure the YJ and HK observations are not too far apart. At the time of writing the detailed implementation strategy (see Dye et al 2006) has J-band observations using 5 seconds per exposure, 2×2 microstepping, and a 2 position jitter; H and K band observations using a 2×2 jitter and 10 seconds per exposure; and Y band using a 2 position jitter and 20 seconds per exposure.

Field selection for the LAS was designed to maximise the overlap with the Sloan Digital Sky Survey (SDDS), but also to keep below the UKIRT Declination limit (+60), and to have a reasonable amount of sky coverage at lower declinations, for follow–up on ESO telescopes. There are three sub-areas, shown in Fig **??**. The detailed field co-ordinates are defined on the UKIDSS web pages.

(i)**The LAS equatorial block : 1907.6 sq.deg**. This includes most of SDSS stripes 9 to 16.

(ii) **The LAS northern block : 1907.6 sq.deg** This includes most of SDSS stripes 26 to 33.

(iii) The **LAS southern stripe : 212.5 sq.deg.** The idea is to survey a section of SDSS stripe 82, extending over -25°¡RA¡+60°, -1.25°¡Dec¡+1.25°. Stripe 82 has been repeatedly scanned by SDSS, and this section is the region of highest quality.

4.2 Design of the Galactic Plane Survey (GPS)

The GPS aims to map as much of the Galactic Plane as possible to a latitude of ±5°. The Galactic Latitude limit is chosen to match other surveys, for example the MSX survey (ref). The survey area is then largely dictated by the UKIRT Declination limit to the North, and by the latitude of UKIRT to the South. Within a reasonable length of time, we can then afford to go roughly a factor of two deeper than the minimum "shallow depth" described in the LAS section above. This depth is good enough to see all of the IMF in quite distant clusters, to see AGB stars all the way through the Galaxy, and to see ordinary G-M stars to several kpc. The tradeoff to consider is then between depth, colours, and repeat coverage. At least three bands, and preferably, four are needed, in order to estimate both spectral type and extinction. However, extinction is large enough in much of the Plane that Y band observations are not very useful; we therefore settled on a JHK survey. Measurement of both variability and proper motions is a goal of the GPS. As a bare minimum to achieve this, we plan three epochs in one band spread across $\sim$ 5 years. We choose the K-band to make these repeats, because, given extinction, this is the sensible band in which to build up depth - it is the K-band that allows us to see clean through the Galaxy. In summary then, we plan five passes - in J and H, with integrations of 80 seconds each; and three passes in K-band, each of 40 second integration.

One of the goals of GPS is the discovery and study of Young Stellar Objects, and in particular molecular outflows. We therefore plan in addition to the above a survey of a single large star forming region in a narrow band H_2 filter.

The area to be mapped is shown in Fig. **??**. The main area is defined by the Galactic latitude range $b = \pm5°$, Dec¡60°, and Dec ¿-15°. These constraints define two sections of Galactic longitude, which are $15° < l < 107°$, and $142° < l < 230°$. In addition we will map a narrow extension through the Galactic centre, within $b = \pm2°$, covering Galactic latitudes $-2° < l < 15°$. The Galactic bulge will also be explored by surveying a thin stripe extending upwards

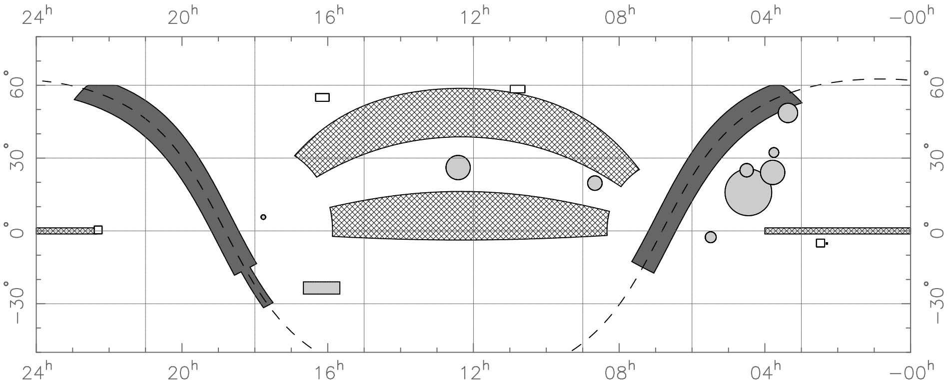

Figure 7. Location on the sky of the fields comprising the various survey components. Cross-hatch : Large Area Survey. Dark Grey : Galactic Plane Survey. Light Grey : Galactic Clusters Survey. Open rectangles : Deep Extragalactic Survey. Note that the Ultra Deep Survey. lies inside the DXS field at "GC" marks the Galactic Centre. The dashed line marks the Galactic plane, and the dotted line marks the ecliptic. Note that UKIRT lies at latitude +20°.

in latitude from the Galactic centre. Finally, the molecular hydrogen survey maps the Taurus-Auriga-Perseus complex, defined by ???

For the GPS, object separation is important and so 2×2 microstepping is employed, typically also with a 2 point jitter. In current implementation, single exposures are 10 seconds at J and H, 5 seconds at K, and 20 seconds in the narrow band H_2 filter. Normal procedure is to do all bands in one visit, as colours are important and many objects are variable.

4.3 Design of the Galactic Clusters Survey (GCS)

The design of the GCS is relatively simple. It needs to target several separate clusters, in order to examine the substellar IMF over a range of ages and metallicities. The depth requirement is set by the need to detect objects in the 30-100 M_J range. Assuming the standard "shallow survey" depth, this means that clusters have to be fairly close and/or young, i.e. within a few hundred parsecs and/or less than a few hundred million years old. There are relatively few such objects, and they are several degrees across. A natural strategy therefore emerges using the standard shallow depth and surveying ten nearby clusters, covering 933 sq. deg. in total. To distinguish cluster members, all five passbands are needed (ZYJHK), plus a measurement of a proper motion using a second pass in the K band. Full colour information provides cluster sequence discrimination in multi–colour space, reddening estimates using shorter versus longer wavelength colour indices, breaking the degeneracy between reddening due to instellar extinction and that due to the presence of circumstellar disks. Six passes are therefore needed. Observations use two point jitters in Z and Y with 20 second exposures, four point jitters in J and H with 10 second exposures, and finally 2 point jitters plus 2–by–2 microstepping in K with 5 second exposures. Microstepping is performed in K only since we anticipate making precision relative astrometric measurements in that passband for proper motion determination, and full profile sampling in the K band is advantageous.

The strategy is to cover the entire 933 sq. degs in a single pass in K, to provide the proper motion baseline, within the initial two year plan. In the same timeframe, ZYJH imaging of 230 sq. degs will be completed. Following this, the remainder of the area will be covered in ZYJH, followed by the second K pass. Table X lists the parameters of the chosen clusters.

Priority/ Name	**Type**	**RA (2000)**	**Dec (2000)**	**Area (sq.deg.)**
(1) Pleiades	open cl.	03 47	+24 07	79
(2) Alpha Per	open cl.	03 22	+48 37	50
(3) Praesepe	open cl.	08 40	+19 40	28
(4) IC 4665	open cl.	17 46	+05 43	0.8
(5) Taurus-Auriga	SF assoc.	04 30	+25 00	386
(6) Orion	SF assoc.	05 29	-02 36	314.2
(7) Sco	SF assoc.	16 10	-23 00	154
(8) Per-OB2	SF assoc.	03 45	+32 17	12.6
(9) Hyades	open cl.	04 27	+15 52	292
(10) Coma-Ber	open cl.	12 25	+26 06	78.5

Table 2. *Clusters targeted for the Galactic Clusters Survey*

4.4 Design of the Deep Extragalactic Survey (DXS)

The DXS is primarily motivated by the desire to detect galaxies at redshifts of 1–1.5 to determine their number density and clustering, identify clusters of galaxies and trace large scale structure. To avoid selecting only the brightest and hence most massive galaxies, this requires the detection of galaxies close to the break in the galaxy luminosity function, M^*_K which is -22.6 locally (Bell et al. 2003). At z=1 this corresponds to an evolution corrected, total K magnitude of 20.7 and 21.8 at z=1.5. Therefore, taking into account aperture effects, our target depth of K=21 will reach to within 0.5–0.7 magnitudes of M^*_K and hence sample a representative galaxy population at z=1. The NIR galaxy colours at this redshift lie in the range J-K=1.5–1.8 so to provide a photometric constraint on the galaxy redshift we also require observations to J=22.8 to ensure matched J and K detections for the target galaxies.

The survey area was driven by the aim to sample large scale structure at z=1 on scales and volume comparable to that measured locally (≈100 Mpc and 0.2 Gpc^3

Name	**Survey**	**Area**	**RA(2000)**	Dec (2000)
XMM-LSS	DXS	8.75	02 18 00	-05 00 00
XMM-Subaru	UDS/DXS	0.77	02 18 00	-05 00 00
Lockman Hole	DXS	8.75	11 00 00	+57 00 00
ELAIS N1	DXS	8.75	16 00 00	+57 00 00
VIMOS 4	DXS	8.75	22 17 50	+00 24 27

Table 3. *Fields targeted for the Deep Extragalatic Survey and the Ultra Deep Survey. Note that the UDS field is at the centre of the DXS XMM-LSS field*

respectively). At z=1, our assumed cosmology implies that 100 Mpc corresponds to 3.5° and a 0.2 Gpc^3 volume in the range z=1-1.5 requires 40 sq.deg. Therefore a minimum combination of 3x3 WFCAM fields will span these scales and a total of 54 WFCAM fields would be required to cover that area (0.75 sq.deg. per field).

The number and position of the DXS survey fields were chosen to provide the best combination of quality and coverage of supporting, multi-wavelength data, to maximise the spatial scale sampled by each individual field ($\sim$ 100 Mpc) for clustering studies and allow a uniform coverage in right ascension. Balancing these factors resulted in the selection of four survey fields: 1) XMM-LSS (centre: 2h45m -04d30m) - a SWIRE, CFHTLS, VVDS, GALEX and XMM survey field adjacent to the UKIDSS UDS area; 2) the Lockman Hole (centre: 10h57m +57d40m) - centred on the SHADES survey area but within the SWIRE and GALEX survey areas with extensive radio coverage; 3) ELIAS-N1 (centre: 16h10m +54d00m) - a SWIRE and GALEX field with additional radio, optical and X-ray data; 4) SA22 (centre: 22h17m +00d20m) - centred on VVDS-4 but the least well surveyed area included to ensure a uniform demand with RA. Other survey fields were considered (COSMOS, NOAO-DWFS, Groth Strip, Spitzer-LFS) but most were either not sufficiently large or comprehensive to justify inclusion. The chosen fields are listed in Table 3

The total area covered by the DXS within the full 7 year span of UKIDSS will dependent on weather and competetion from other surveys (most notably the UDS) but our goal is 35 sq.deg. or 12 WFCAM fields in each DXS survey area.

For DXS, star-galaxy separation at faint magnitudes will be very important, so 2×2 micro-stepping is employed to give good sampling. Achieving the required depth will require reliable stacking, and so minimising any systematic effects in detector structure that do not flat-field out. The DXS strategy therefore employs substantial jittering. Each visit to a given tile position uses 10 second exposures, a sixteen point jitter, and 2×2 microstepping at each of these jitter positions. Each such visit therefore has an integration of 640 second at each sky point.To reach the intended depth requires a total exposure of 3.8 and 2.5 hours in J and K respectively or 21 and 14 visits each. Given that the DXS observations do not require photometric conditions or the very best seeing, then the final number of visits for each field may be higher to compensate for these poorer conditions.

4.5 Design of the Ultra Deep Survey (UDS)

The UDS aims to go as deep as possible in a single contiguous WFCAM tile. The depth is set by the aim of detecting giant ellipticals at z=3 if they exist. The total magnitude of such objects is expected to be K$\sim$ 21 but they will significantly extended, so that we need to reach a point source depth of K=23. Three bands are needed, to get photometric redshifts and discriminate objects. To effectively separate ellipticals and starbursts, we need to be able to detect colours J-K$\sim$ 2 and H-K$\sim$1, otherwise most of our detections may be K-band only. This sets limits of J=25 and H=24, which are in fact more demanding in time than the K-band observations.

As with the DXS, we need good sampling to enable star-galaxy separation at faint magnitudes, and multiple jitters to overcome detector systematics when stacking. These issues are even more demanding however for the UDS; at each visit we use 3×3 microstepping and a 9 point jitter, and repeated visits are not at precisely the same position, but in a carefully arranged pattern - a kind of super-jitter.

The field chosen for the UDS is just to the side of one the DXS fields, and is also a Subaru deep field, guaranteeing deep optical data. Considerable data also exists at other wavelengths.

4.6 Two Year Goals

We aim to complete self contained and scientifically valuable datasets on a two year timescale. The detailed plan is set out in Dye et al (2006), but briefly is as follows. The shallow surveys (LAS, GCS, and GPS) are accelerated compared to the deep stacked surveys (DXS and UDS). In addition, the LAS concentrates on southern latitudes in the first two years, in order to maximise VLT follow-up. For LAS, roughly half the area - the equatorial block, and the southern stripe - will get complete YJHK coverage. (Second epoch J for the same areas will come later). Additional J only coverage in the Northern block will be achieved as time permits. For GPS, the prime aim is to obtain the first of three K epochs over the whole survey area, with J, H, and H_2 coverage over a sub-area. For GCS the aim is to get complete filter coverage for five of the ten target clusters, and the central regions of three, plus K only coverage of the remaining two. For DXS, the two year aim is to reach the full depth in J and K for a subset of the area - four tiles (3.1 square degrees) in each of the four fields. The UDS is of course a single tile. The two year goal is to achieve K=22.8 (full depth) and J=23.8 (one magnitude short), with no H coverage.

5 DATA PROCESSING AND DATA PRODUCTS

The commitment to making a public survey requires the construction of complete, reliable, and documented products from the raw data. The very large volume of UKIDSS data (200 GBytes/night) means that to achieve these goals requires a uniform and automated approach to data processing. Likewise the large accumulated volume of products, many tens of Terabytes, means that as well as providing public data access, we need to provide online querying and analysis facilities as a service. These ambitions are met for all WFCAM data (both UKIDSS survey data and PATT PI data) by the VISTA Data Flow System (VDFS). VDFS is a PPARC-funded project involving QMUL, Cambridge and

Edinburgh, aimed at handling the data from first WFCAM and then the VISTA telescope. (The Science Archive was also prototyped on the SuperCosmos Science Archive : see http://surveys.roe.ac.uk/ssa/ and Hambly et al 2004). The system aims at (i) removing instrumental signature; (ii) extracting source catalogues on a frame by frame basis; (iii) constructing survey level products - stacked pixel mosaics and merged catalogues; (iv) providing users with both data access and methods for querying and analysing WFCAM data.

5.1 The WFCAM Pipeline

The general philosophy behind the pipeline processing is that all fundamental data products are FITS multiextension files with headers describing the data taking protocols in sufficient detail to trigger the appropriate pipeline processing components, and that all derived information, quality control measures, photometric and astrometric calibration and processing details, are also incorporated within the FITS headers. Generated object catalogues are stored as multiextension FITS binary tables. These FITS files thereby provide the basis for ingest into databases both for archiving and for real time monitoring of survey progress and hence survey planning.

After conversion at the summit to FITS files, to reduce the data storage, I/O overheads and transport requirements, we make use of lossless Rice tile compression (eg. Sabbey et al. 1998). For this type of data (32 bit integer) the Rice compression algorithm typically gives an overall factor of 3–4 reduction in file size. Data are shipped roughly weekly from JAC using LTO tapes, one per detector channel, and combined to create the raw archived multiextension FITS files on ingest in Cambridge.

The data processing strategy attempts to minimise the use of on-sky science data to form "calibration" images for removing the instrumental signature. By doing this we also minimise the creation of data-related artefacts introduced in the image processing phase. To achieve this we have departed somewhat from the usual NIR processing strategies by, in particular, making extensive use of twilight flats, rather than dark-sky flats (which potentially can be corrupted by thermal glow, fringing, large objects and so on) and by attempting to decouple, insofar as is possible, sky estimation/correction from the science images.

Each night of data is pipeline processed independently using the master calibration twilight flats (updated at least monthly) and a series of nightly generated dark frames covering the range of exposure times and readout modes used during that night. A running sky "average" in each passband is used for sky artefact correction. After removing the basic instrumental signature the pipeline then uses the header control keywords to produce interleaved and/or combined (stacked) image frames for further analysis. This includes generation of detected object catalogues, and astrometric and photometric calibration based on 2MASS. A more detailed description of the WFCAM processing is given in Irwin et al. (2006).

5.2 The WFCAM Science Archive

Data processing delivers standard nightly pipeline processed images and associated single passband catalogues, complete with astrometric and first–pass photometric calibrations and all associated 'meta' (descriptive) data in flat FITS files. Image data volume is typically $\sim$100 GBytes per night, with catalogue and descriptive data being typically $\sim$ 10% of that figure. Hence, over the course of several years of observations it is anticipated that 10s/100s of Tbytes of catalogue/image data will be produced by survey operations with WFCAM. In order to enable science exploitation of these datasets, the concept of a 'science archive' has been developed as the final stage in the systems–engineered data flow system from instrument to end–user (Hambly et al. 2004).

The WFCAM Science Archive (WSA) is much more than a simple repository of the basic data products described previously. A commercial relational database management system (RDBMS) deployed on a mid–range, scalable hardware platform is used as the online storage into which all catalogue and meta data are ingested. This RDBMS acts as the backing store for a set of curation applications that produce enhanced database driven data products (both image products, e.g. broad–band/narrow–band difference images; and catalogue products, e.g. merged multi–colour, multi–epoch source lists). Moreover, the same relational data model is exposed to the users through a set of web–interface applications that provide extremely flexible user access to the enhanced database driven data products via a Structured Query Language interface. The primary purpose of the WSA is to provide user access to UKIDSS datasets, and a full description, along with typical usage examples, is given in Hambly et al. (2006).

The WSA can be accessed at http://surveys.roe.ac.uk/wsa. Step-by-step examples of WSA useage are also included in Dye et al (2006).

6 SCIENCE VERIFICATION PROGRAMME

Following the technical commissioning of WFCAM, and before the commencement of formal survey operations in May 2005, the UKIDSS consortium undertook a modest set of test observations, as a "Science Verification (SV)" programme. These observations were aimed primarily at further technical commissioning, testing and tuning the implementation strategy, and exercising the data flow system. However the data collected have clearly demonstrated the scientific power of UKIDSS, and confirm the efficacy of the survey design. In this section we show some examples of science results from these SV data.

6.1 Science Verification results for the LAS

Final text to come from Jameson and Warren

The LAS SV programme covered X square degrees to the standard shallow depth in filters A, B, C. One scientific aim was to test the likely recovery of cool brown dwarfs. The success of this is illustrated in Fig. 8, which shows data from a tile aimed at a known T3 dwarf (REF). This object was indeed detected, and the colours found are consistent with those previously published. Within this figure, there are a

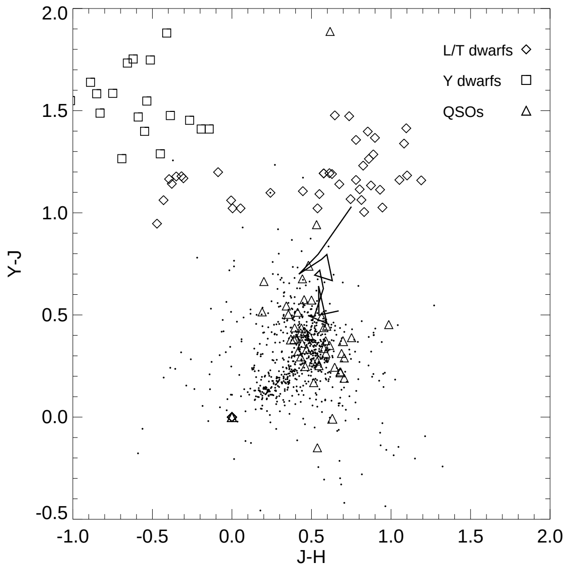

Figure 8. The Y-J,J-H two colour diagram for a single tile observed in the LAS SV programme. Black dots show the data for stellar sources detected in the WFCAM data. Also shown are the expected positions of QSOs (triangles) L/T dwarfs (squares), and Y dwarfs (diamonds). The observation was targeted at a known 2MASS T3 dwarf which is recovered with Y-J=1.1 and J-H=0.25.

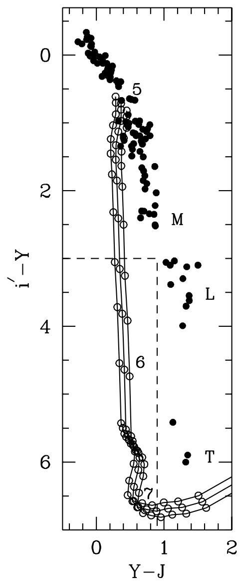

Figure 9. OLD DIAG : TO BE REPLACED. Separating T dwarfs and high-z quasars in iYJ diagram. Caption to come.

handful of other objects with colours in the same region, Y-J$\sim$1.2 It actually seems unlikely that these are genuine T-dwarfs, unless they are much more common than thought. Their nature will not be clear until spectroscopic follow-up is obtained. The SV exercise therefore shows both the power and the problems associated with such a large volume of data.

A second aim was to test the location of quasars, using the new Y filter and combining with SDSS data. Figure 9 shows an optical-IR two colour diagram for stellar sources over an area of X square degrees. (Figure and more words to come)

6.2 Science Verification results for the GPS

The science verification data for the Galactic Plane Survey included a 0.75 deg^2 tile centred on M17 (a high mass star formation region) and several tiles running across the plane at Dec=-1°. This latter region at $l \approx 30$ has a high source density since it lie close to the tangent point of a spiral arm. The M17 region has been well studied (eg. Jiang et al.2002) and it provided a good test of the photometric reliability of the data in a nebulous region.

The central 25% of the M17 tile is shown in Figure 10 This illustrates the sensitivity of the WFCAM to the structure and stellar population of distant star formation regions. Within the brightest nebulosity the archival source lists become seriously incomplete. Very few spurious sources are detected however. While this situation may be improved by profile fitting photometry, this suggests that more complete luminosity functions may be derived in nebulous regions by performing independent photometry on the reduced data, using the zero points provided in the image headers.

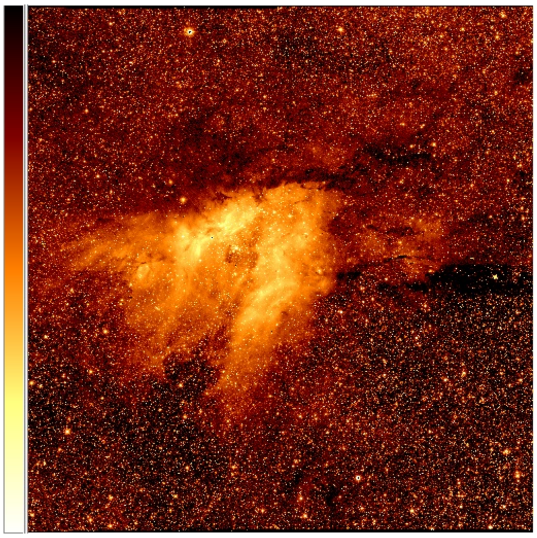

Figure 10. Central part of M17 tile.

At the present time only aperture photometry is available in the WFCAM Science Archive. As might be expected, these data suffer considerably from the effects of crowding. This is illustrated in the two colour diagram in Figure 11, which shows the colours of the stellar sources within a tile at l=28, b=5 that are listed as having well measured photometry (errors $<$ 0.02 mag at J, H and K) in the WSA. The great majority of these relatively bright sources can be dereddened to lie on the of the two colour sequences indicated by the curved lines. These curves are for luminosity Class V main sequence stars (lower curve) and luminosity class III giants (upper curve). However a noticeable minority of sources are scattered throughout the plotted diagram and beyond. Profile fitting photometry is planned for a future release of UKIDSS data and this is expected to significantly improve the completeness and reliability of the

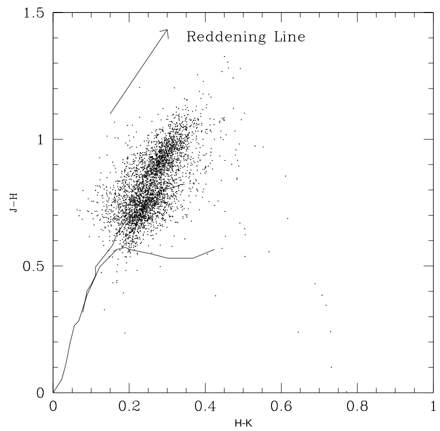

Figure 11. coour-colour diag from M17. Caption to come.

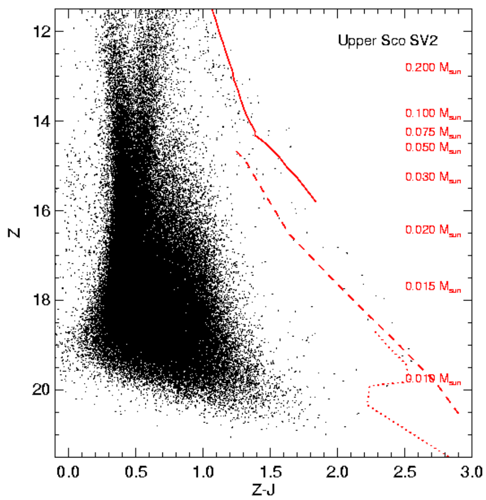

Figure 12. Z vs J–Z diagram for the Upper Sco field, showing only those sources with J>10.5 and Z>11.5. Model isochrones are described in the text. The cluster sequence can be clearly seen following the models well.

photometry. In relatively uncrowded regions the aperture photometry appears to have a sensitivity within 0.5 mag of the desired depths at J, H and K. The image quality was relatively poor during the science verification phase so we expect that the bulk of the actual survey data will be very close to the sensitivity limits quoted in section 2.4. Even in the aperture photometry in Figure X there are relatively few sources with colours consistent with those of late M dwarfs near the end of the class V sequence. Hence bona fide late M and L dwarfs should be detectable by follow observations with a reasonable success rate, especially proper motion information becomes available with the 2nd and 3rd epoch data.

A full analysis of the data quality in the GPS and some results from the SV data will be presented in a future paper (Lucas et al., in prep) after profile fitting photometry becomes available.

6.3 Science Verification results for the GCS

Science verification observations for the GCS yielded 8 tiles in each of the three targets observable at that time: IC 4665, Upper Scorpius and Coma Berenices. In the case of IC 4665, the SV observations complete the required survey for that cluster. In Figure 12 we show a Z versus Z–J colour–magnitude diagram for Upper Scorpius. The observations cover 6 square degrees and approximately 100,000 point sources are detected. Nearly all of these are foreground stars. In Fig 12 we show only those sources with J>10.5 and Z>11.5. The main sequence and giant branch are clearly seen. In addition one can notice a clean sequence to the right of the diagram running from (Z-J,Z)=(1.0,12.0) to (Z-J,Z)=(2.5,19.0), which must be the cluster sequence.

Overplotted are 5 Myr theoretical isochrones shifted at a distance of 140 pc (Upper Sco is 5+/-3 Myr old and located at 140 pc): BCAH98 or NextGen models (solid line; Baraffe et al. 1998), DUSTY or BCAH00 (dashed line; Chabrier et al. 2000), and COND03 (dotted line; Baraffe et al. 2003). These isochrones were specifically computed for the WF-CAM filters. On these isochrones we also indicate stellar masses. We can clearly see therefore that we are indeed locating brown dwarfs within the cluster and will be able to derive a substellar mass function.

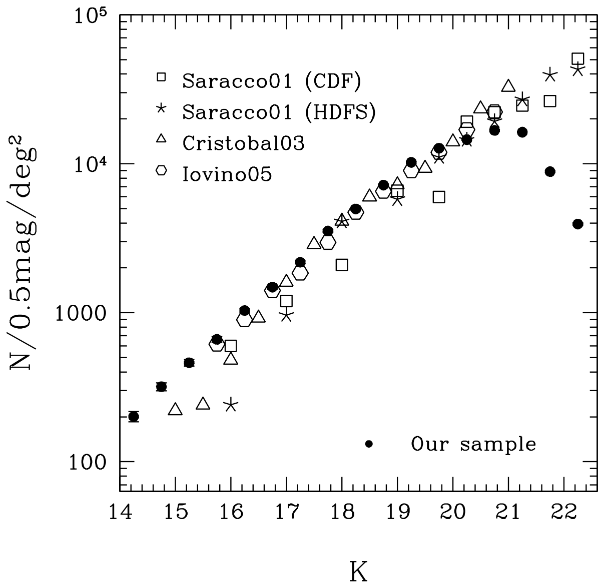

Figure 13. K counts. Caption to come.

6.4 Science Verification results for the DXS and UDS

Final text to come from Edge and Almaini.

Full depth for UDS will of course not be achieved until the end of the survey programme. The UDS and DXS and therefore undertook a joint science verification programme. (Size, area)

Fig 13 shows the K-band number counts from an accumulated exposure of 1.5 hours (6 hours of data) on a single tile in the ELAIS-N1 field. Simulations give a 50% complete-

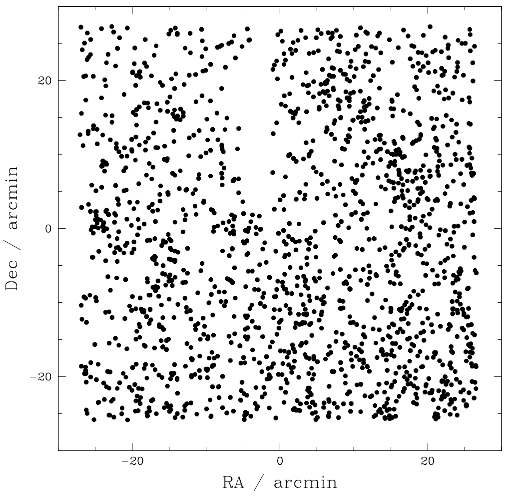

Figure 14. EROs. Caption to come.

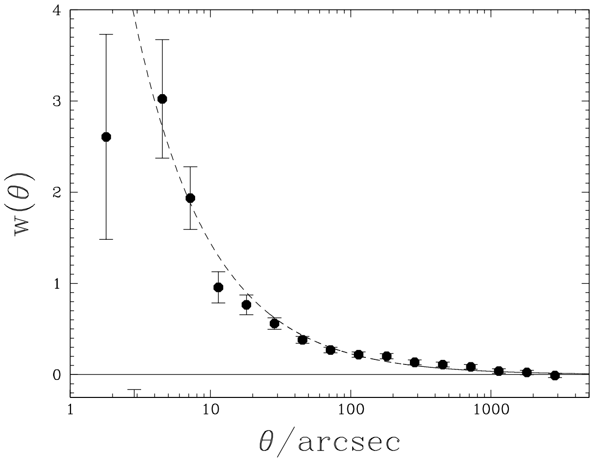

Figure 15. K counts. Caption to come.

ness limit of around K=20.3, roughly in line with expectations for 0.9 arcsec FWHM image quality. The final DXS survey will be approximately half a magnitude deeper than this. Note that this figure includes both stars and galaxies. For comparison, we show number counts from several other surveys. This figure illustrates the power of UKIDSS, as we improve on the results of many other surveys in a single tile.

One of the scientific goals of DXS and UDS is to locate Extremely Red Objects (EROs). Again, with a tiny fraction of the eventual survey data we are already competitive with all previous studies. Using the same UKIDSS SV data as above, we cross-match with publicly available optical data from the INT Wide Field Camera Survey (http://www.ast.cam.ac.uk/ wfcsur/), and select EROs as those with R-K>5 down to a limit of K=19. This produces a sample of 1660 EROs. The reliability of this sample is limited by the optical data at R=24, not by the IR data at K=19. Fig. 14 shows the sky distribution of these objects, and Fig. 15 shows the angular correlation function derived from this sky distribution, along with the best fitting from Roche et al. (2002). Note that we have extrapolated this model by a factor five in scale compared to the range over which it was fitted in that paper.

7 SURVEY RELEASES

Data access policy for UKIDSS is set by the UKIRT Board, and is set out on the JAC web pages. (http://www.jach.hawaii.edu/UKIRT/surveys/UKIDSSdatapolicies.html) Observations obtained through the standard UK PATT time allocation system are proprietary to the PI for one year after the time of observation, and public after that. Data access is through the WSA. While in the proprietary period this requires a username and password for the PI; afterwards anonymous access becomes possible.

UKIDSS is intended to produce multi-use data of general benefit to astronomers worldwide, but with a temporary advantage for the communities that developed the camera and surveys. Initially this meant UK astronomers, but now means any astronomer currently working in an ESO member state. The general principle is that the data is freely available to any such astronomer from the point of release, and available worldwide eighteen months later. (Note that individual members of the consortium have no privileged data access.) During the ESO-restricted phase, data access requires registration with the WSA. This is organised through a set of "community contacts" at astronomical institutions in ESO member states, who maintain their own databases of local users through the WSA system. Any reader who is not yet registered who believes they are eligible should contact their local community contact, or if necessary ask for a new community contact to be established. Fuller instructions and a list of current contacts is on the UKIDSS web page at http://www.ukidss.org/archive/archive.html.

It is intended to make UKIDSS data avaialable in a series of well defined staged releases. As well as involving incrementally more data, each release will correspond to a distinct processing history, with updated correction of artefacts and so on. Each release will therefore be documented by a paper describing the contents and limitations of that release. The first preliminary release, avalable from Feb 10th 2006, has a relatively small amount of data (about 1% of the expected total), and several known imperfections in the data processing. This is therefore being labelled an "Early Data Release (EDR)". It is described in more detail in Dye et al (2006). Even though the EDR is a small fraction of the eventual complete UKIDSS, we estimate that it already contains as many photons as the entire 2MASS survey. Likewise, although it contains some known processing imperfections, these are well documented and the data are easily good enough to do some exciting new science.

The first full data release (DR1) is currently scheduled for summer 2006, and is expected to contain approximately 10% of the expected UKIDSS total. Further releases are likely to take place thereafter every six months or so.

Raw WFCAM data are available through the ESO archive system (URL). All of the UKIDSS processed images and catalogues are accessible and queryable through the web-based WFCAM Science Archive (WSA : http://surveys.roe.ac.uk/wsa)

8 ACKNOWLEDGEMENTS

This paper is written on behalf of the entire UKIDSS consortium. The membership list can be found at http://www.ukidss.org. The formal authorship includes the heads of the various UKIDSS working groups, plus a small number of individuals who worked particularly hard during the design or SV phases. In addition to the bulk of the consortium, there are others we would like to thank. The UKIDSS enterprise would be impossible without the staff of UKIRT, the staff at UKATC who built WFCAM, and the staff of CASU and WFAU who built the data processing system.

9 REFERENCES

Bell, E.F., McIntosh, D.H., Katz, N., Weinberg, M.D. 2003 ApJS 149, 289

Fan, X., et al., 1999, 2001, AJ, 121, 31.

Hambly, N.C., et al., 1999, MNRAS, 303, 835.

Hambly, N.C., et al., 2004, In: Optimizing Scientific Return for Astronomy through Information Technologies, Proceedings of the SPIE, 5493, 423 (2004)

Hambly, N.C., et al., 2004, MNRAS, xxx, yyy.

Hambly, N.C., et al., 2006, in preparation

Hasegawa T., Oka T., Sato F., Tsuboi M., Yamazaki A. 1998, IAU Symp. 184, 171

Jiang Z., Yao Y., Yang J., Ando M., Kato D., Kawai T., Kurita M., Nagata T., Nagayama T., Nakajima Y., Nagashima C., Sato S., Tamura M., Nakaya H., Sugitani K. 2002, AJ 577, 245

Legget, S. K, et al, 2000, ApJ, 536, L35.

Legget, S., et al, astro-ph/0409389.

Martí, E.L., Magazzù, A., 2006, AN, in press

Moraux, E., et al,. 2006, AN, in press

Sabbey et al. 1998, PASP, 110, 1067

Ungerechts H, Thaddeus P, 1987 ApJS 63, 645

Wang Q., Gotthelf E., & Lang C. 2002, Nature, 415, 148