Mon. Not. R. Astron. Soc. **311**, 707–718 (2000)

Galaxy number counts – IV. Surveying the Herschel Deep Field in the near-infrared

H. J. McCracken,[1★] N. Metcalfe,[1] T. Shanks,[1] A. Campos,[1,2] J. P. Gardner[1,3]† and R. Fong[1]

[1]*Department of Physics, University of Durham Science Laboratories, South Rd, Durham DH1 3LE*
[2]*Instituto de Matematicas y Fisica Fundamental (CSIC), Serrano 113 bis, E-28006 Madrid, Spain*
[3]*Laboratory for Astronomy and Solar Physics, Code 681, Goddard Space Flight Centre, Greenbelt, MD 20771, USA*

Accepted 1999 September 2. Received 1999 August 4; in original form 1998 September 28

ABSTRACT

We present results from two new near-infrared imaging surveys. One survey covers 47.2 arcmin^2 to $K(3\sigma) = 20$ mag whilst a second, deeper catalogue covers a subarea of 1.8 arcmin^2 to $K(3\sigma) = 22.75$ mag. Over the entire area we have extremely deep optical photometry in four bandpasses (*UBRI*), allowing us to track the colour evolution of galaxies to very faint magnitude limits.

Our K-band number counts are consistent with the predictions of non-evolving models with $0 \leqslant q_0 \leqslant 0.5$. The K-selected $(B-K)$ galaxy colour distributions from our surveys move sharply bluewards fainter than $K{\sim}20$. At brighter magnitudes ($K < 20$ mag) our K-selected $(B-K)$ distributions indicate a deficiency of red, early-type galaxies at $z \sim 1$ compared with the predictions of passively evolving models, which implies either a pure luminosity evolution (PLE) model, where star formation continues at a low level after an initial burst, or dynamical merging. At fainter magnitudes, the continuing bluewards trend observed in $(B-K)$ can be explained purely in terms of passively evolving PLE models. We detect 0.5 ± 0.1 galaxy arcmin^{-2} with $(I-K) > 4$ and $19 < K < 20$ mag. Although this is a factor of $\sim$3 (2σ) more objects than in the recent survey of Berger et al., this is still lower than the predictions of standard passively evolving models and more consistent with PLE models containing small amounts of ongoing star formation. Our observed numbers of $(I-K) > 4$ galaxies at $K \sim 20$ are lower than the predictions of passively evolving models or PLE models with a low level of continuing star formation, suggesting that at least part of the larger deficiency observed in $(B-K)$ at $K \sim 20$ may be caused by star formation rather than dynamical merging.

As we and others have noted, the number redshift distribution at $18 < K < 19$ of recent, deep K-selected redshift surveys is well fitted by non-evolving models, and passively evolving models with a Salpeter or Scalo initial mass functions predict too many galaxies with $z > 1$. Dynamical merging is one possible solution to reduce the numbers of these galaxies but (as we have suggested previously) a dwarf-dominated initial mass function for early-type galaxies could offer an alternative explanation; we show here that such a model reproduces well the optical–infrared colour distributions and K-band galaxy counts.

Key words: galaxies: evolution – galaxies: photometry – infrared: galaxies.

1 INTRODUCTION

In our previous papers [Metcalfe et al. 1991 (Paper II), Metcalfe et al. 1995a (Paper III) and Metcalfe et al. 1996] we have investigated the evolution of the counts and colours of faint field galaxies, culminating in extremely deep optical imaging of a representative region of sky we call the 'Herschel Deep Field'. In this paper we present new near-infrared observations of this field, and combine these data with our existing ultra-deep optical imagery. Near-infrared observations have many advantages which are often quoted in the literature; among them the insensitivity of

★ Present address: Laboratoire d'Astronomie Spatiale, 13376 Marseille Cedex 12, France.
† NOAO Research Associate.

the *k*-correction to type or star formation history at intermediate redshifts is perhaps the most relevant in a cosmological setting.

Previously we have considered our observations in the context of pure luminosity evolution (PLE) models. Recently, additional support for these models has come from fully complete, blue-selected, redshift surveys (Cowie, Songaila & Hu 1996) which have revealed the presence of an extended tail in the redshift distribution, partially resolving the paradox that non-evolving galaxy formation models could be used to fit the results of the early faint redshift surveys (Glazebrook et al. 1995b; Broadhurst, Ellis & Glazebrook 1992). Such models, as we have shown, are able to reproduce all the observable quantities of the faint field galaxy population (counts, colours, redshift distributions), at least for low-Ω_0 universes and within current observational uncertainties. High-Ω_0 universes can be accommodated by the model if we add an extra population of low-luminosity galaxies with constant star formation rates, which boosts the counts at faint ($B > 25$ mag) magnitude levels (Metcalfe et al. 1996; Campos 1997). In the PLE model, galaxies form monolithically at high redshift. Changes in luminosity and colour after the formation event are modelled with stellar population synthesis models. *K*-selected galaxy samples, at least until $K \sim 20$, are dominated by early-type galaxies; therefore, changes in optical–infrared galaxy colours are particularly sensitive to evolution of this galaxy class. By investigating near-infrared number counts and colour distributions we can potentially address questions concerning the formation and evolution of elliptical galaxies.

The two surveys presented here are complementary. The 'wide' survey covers a large, $\sim$50 arcmin^2 area to a shallow, $K(3\sigma) = 20$ mag, depth, while the 'deep' survey covers a much smaller area of $\sim$1.8 arcmin^2 to a much fainter $K(3\sigma) = 22.75$ mag limiting magnitude. In combination, both surveys provide a *K*-limited sample extending from $K \sim 16$ mag to $K = 22.75$ mag.

This paper is organized as follows: Section 2 describes the data reduction techniques used in both the wide and deep surveys; in Section 3 number counts and colour distributions for both our data sets are presented; in Section 4 we discuss these results in terms of our modelling procedure; and finally, in Section 5 we summarize the main conclusions we draw from our work.

2 OBSERVATIONS AND DATA REDUCTION

2.1 Reducing the 'wide' survey

The observations were made over four nights in 1995 July at the 3.8-m United Kingdom Infrared Telescope (UKIRT) using the IRCAM3 near-infrared detector with a standard *K* (2.2-μm) filter. Conditions were photometric on two of the four nights. The mean seeing was $\sim$1 arcsec FWHM. The centre of the region we imaged, RA $0^h19^m59^s$, Dec. $+0°4'0''$, corresponds to the centre of our deep optical fields, which are described in Metcalfe et al. (in preparation). For this region we have imaging data complete to the limiting magnitudes listed in Table 1. The optical observations were made at the 4.2-m William Herschel and 2.5-m Isaac Newton telescopes (WHT and INT hereafter) and cover approximately $\sim$50 arcmin^2.

Table 1. Photometric Limits of the Herschel Deep Field.

Filter	*U*	*B*	*R*	*I*	*K*
Limit (3σ)	26.8	27.9/28.2	26.3	25.6	20.0/22.75
Area (arcmin^2)	46.4	46.8/2.8	48.5	52.2	47.2/1.78

The IRCAM3 detector, a 256×256 element InSb chip, has a pixel scale of 0.286 arcsec and consequently each image covers only $\sim$1.5 arcmin^2. For this reason we observed a mosaic of 6×6 frames which completely overlapped the previously observed optical fields. The entire mosaic was observed in each pass for 2 min at each position, and offset slightly for each observation from the previous pointing. In total, the entire survey area was covered 15 times, giving a total integration time at each subraster of 30 min. The short exposures for each image is a consequence of the high sky brightness in the near-infrared and is fact the longest exposure time possible with IRCAM3 in the *K* band before sky background causes the detector to saturate. Our data reduction technique for both infrared surveys follows the methods outlined in Gardner (1995).

A two-pass procedure is adopted: for the first stage in the data reductions a dark frame is subtracted from each night's images, calculated from the median of the dark frames taken that particular night. Next, the images are sorted in order of time of exposure. Then, a *running sky* is constructed from the median of the six images nearest in time to the frame being processed. This is necessary as the sky background varies on exceedingly short timescales (i.e. under 30 min). This running sky is then subtracted from the current image and the procedure repeated for all the images in the mosaic. Next, for each night a 'superflat' is constructed by medianing together all the observations for that night. Each image is then divided by the superflat.

Once all the images have been dark- and sky-subtracted and flat-fielded, they are grouped into 36 separate pointings and each stack is averaged together using the IRAF IMCOMBINE task. Next, masks were constructed from each of the stacked images by flagging all pixels more than 3σ from the sky background. These masks were then used in recalculating the running sky frames. This was necessary as bright sources present in each of the individual exposures caused raised counts in the original running sky frames. The whole procedure outlined above was then repeated, using these new frames.

Finally, the 36 stacked subrasters were combined into a single image by matching each subraster with objects on the extremely deep *B*-band frame and using this to calculate the correct position of each subraster relative to the others. The final image, after trimming, covers an area of 47.2 arcmin^2, overlapping our WHT deep optical fields.

We calibrated this data by repeatedly ($\sim$30 times) observing the faint UKIRT standards (Casali & Hawarden 1992) each night at a range of airmasses, although fortuitously the mean airmass of the standards and of the programme field images was approximately the same, so accurate determination of the extinction coefficient was not necessary. The rms scatter between our measurements of standards was $\pm$0.03 mag.

2.2 Reducing the 'deep' survey

Observations for the 'deep' survey were made over six nights in 1994 September at UKIRT using IRCAM3 with a *K*-band filter. Conditions were photometric during the first two nights, and the remainder of the data were scaled to the photometry of bright objects in the field made on those nights. The seeing ranged from 0.8 to 1.4 arcsec. The total observing time on the field was just over 100 ks (27.9 h), and was made up of 838 individual exposures of 120 s each. Each individual exposure was made up of 12 coadds of 10 s each. The telescope was dithered using the offset guide camera in a diamond pattern of 13 positions, each position

Table 2. Detection and measurement parameters used in both surveys.

Frame	Limiting isophote (mag arcsec^{-2})	Detection limit (K mag)	Min. r (arcsec)	Kron multiplying factor	Correction to total (mag)
Deep K	24.50	23.50	1.25	1.40	0.32
Wide K	21.50	22.00	1.35	1.50	0.25

separated by 5.25 arcsec in both right ascension and declination. Data reduction proceeded generally in the manner described in Section 2.1. Dark frames were made several times each night, and were subtracted from the images. A running sky was constructed, using the median of 12 images (six before and six after). The running sky was subtracted (after applying the two-pass object masking system described above), and a 'superflat' was constructed for each night and divided into the data. The input dither pattern was recovered and adjusted by a centroid of the brightest object in the field. The data taken each night were shifted and coadded. The photometry of each night's data were scaled to that of the photometric first two nights (which were themselves consistent). The final image covers an area of 1.78 arcmin^2, and has a limiting depth of $3\sigma = 22.75$. The deep field is centred at RA $00^{\rm h}\ 19^{\rm m}\ 59\overset{\rm s}{.}6$, Dec.$00°02'56''$ (1950.0) and covers 81 arcsec square (140×140) 0.58-arcsec pixels.

2.3 Object detection and photometry

For both the deep and wide images we use the same prescription for object detection and photometry as in our optical studies (Metcalfe et al. 1995a). Note that to improve the reliability of image detection, both images were binned 2×2 (the 0.286-arcsec pixel size of IRCAM3, combined with median seeing of $\sim$1 arcsec FWHM, means that our raw images are oversampled). We fit the sky background approximately with a third-order polynomial and then subtract this from the frame.

Object selection is initially performed with an isophotal detection routine which searches for connected pixels above a threshold of $K = 21.50$ mag arcsec^{-2}, the integrated isophotal magnitude of which is brighter than $K_{21.50} = 22$ and the isophotal radius of which is greater than 0.5 arcsec (these values are for the 'wide' survey; values used for the 'deep' survey are given in Table 2). These objects are removed from the frame and replaced by a local sky value. The resulting image is heavily smoothed (with several passes with box smoothing up to 10 pixels on a side) and then subtracted from the original. The result is an extremely flat image on which the isophotal detection procedure outlined above is then rerun.

At this stage, in order to reduce false detections, any images within a few pixels of one another are recombined into one. The resulting catalogue deliberately samples deep into the noise. To extract reliable images we then calculate a Kron-type aperture magnitude (Kron 1980) at each image coordinate in the isophotal catalogue, using a recalculated local sky value. This is essentially an aperture magnitude calculated to $mR_{\rm k}$, where $R_{\rm k}$ is the Kron radius and m is the multiplying factor given in Table 2. It is necessary to set a minimum aperture equivalent to that for an unresolved object, and apply a correction to total magnitude. These are also listed in Table 2. A more detailed discussion of this technique is given in Metcalfe et al. (1991). Our optical–infrared colours are measured inside fixed 1.5-arcsec radius apertures. Note that to form the 'deep' $(B-K)$ colours, the K-band magnitudes are matched with the 2-arcmin diameter, stacked, ultra-deep B-band image described in Metcalfe et al. (1998) (see also Table 1). All other colours are formed from matches with the 7×7 arcmin^2 Herschel Deep Field optical images. Star–galaxy separation is taken from Metcalfe et al. (in preparation).

Table 3. Results of adding artificial stars to the deep K data.

True magnitude	Measured magnitude	Detection rate (per cent)
20.75	20.70 ± 0.17	90
21.25	21.22 ± 0.21	88
21.75	21.76 ± 0.26	85
22.25	22.36 ± 0.36	76
22.75	22.83 ± 0.37	51

To remove most of the spurious noise detections, we then restrict our final catalogue to all objects with integrated magnitude inside their Kron radius greater than the magnitude of a positive 3σ noise variation inside an aperture of the same diameter. Metcalfe et al. (1995a) have shown that this restricts spurious images to be at most around $\sim$10 per cent of the differential count at the magnitude limit.

2.4 Confusion corrections

The deeper the ground-based number-counts are extended, the more important the corrections for confusion become. This is of particular relevance for the deep K-band data. To make an estimate of the completeness of our counts on this frame, we have added numerous artificial stars of various known magnitudes to the real data frame and then reanalysed the resulting images using the normal data reduction procedure. We then compare the measured magnitudes with those that we input. Table 3 gives the mean magnitudes, scatter and detection rate for these stars. Note that an image is considered undetected if it is merged with another image and the combined brightness is a factor of 2 or more greater than the true magnitude, or if it is not found within ± 2 pixels of its true position. As expected, the detection rates in the real data drop as the magnitude falls. This is almost entirely the result of objects being merged with other, brighter objects – isolated objects are recovered with almost 100 per cent efficiency. Note, however, that the mean recovered magnitudes are close to the true values. Although most objects at the limit of our data are galaxies, they are in general close to being unresolved, and so we correct our counts using the detection rates implied by our artificial stars.

Owing to the much wider mean separation between objects, the shallower K-band data do not suffer from confusion, and no correction is made.

3 RESULTS

The final 'wide' catalogue contains 298 galaxies to $K(3\sigma) = 20.0$ mag over an area of 47.2 arcmin^2. The deep catalogue contains 86 galaxies to $K(3\sigma) = 22.75$ mag over an area of 1.8 arcmin^2.

3.1 *K*-band galaxy number counts

Number counts derived from both our catalogues are presented in Fig. 1, which shows our data (with error bars calculated using Poisson counting statistics) and a compilation of published number counts. At bright magnitudes (16 mag $< K <$ 20 mag) our results agree well with the counts from the literature, whereas at fainter magnitudes (20 mag $< K <$ 22 mag) our data favour the higher counts found by Bershady, Lowenthal & Koo (1998) as opposed to the lower counts measured by Djorgovski et al. (1995). Bershady et al. claim that the aperture corrections employed by Djorgovski et al. (1995) resulted in an overestimation of the depths of the latter survey by ∼0.5 mag; our results support this interpretation.

It is worth noting that in our last bin, as in Bershady et al., our counts are only approximately 50 per cent complete and therefore subject to large corrections. Furthermore, given the small areas of surveys at this magnitude level (where areas covered are typically ∼1 arcmin^2) it is also possible that cosmic variance could account for the large observed field-to-field variations between the different groups.

Considering the slope of our number counts, we derive a value of $\mathrm{d}(\log N)/\mathrm{d}m = 0.37 \pm 0.03$ (1σ errors) over the magnitude range covered by both wide and deep catalogues (namely, $K \sim 16$–23). This value is in good agreement with $\mathrm{d}(\log N)/\mathrm{d}m = 0.36 \pm 0.02$ found by Bershady et al. We do not find the count slope for the deeper survey to be significantly different from the shallower one [however, as pointed out by Gardner, Cowie & Wainscoat (1993) the compiled *K*-band number counts show a slope change at $K \sim 17$].

In Fig. 1 we also plot the predictions of two non-evolving galaxy count models: one with deceleration parameter $q_0 = 0.5$ (solid line) and one with $q_0 = 0.05$ (dotted line). Both models give a good fit to the observations from $K \sim 14$ to $K \sim 22$. These non-evolving model were computed from the *k*-corrections of Bruzual & Charlot (1993) and the $z = 0$ luminosity function parameters given in Table 4. We compute M_K^* by combining the tabulated M_B^* with the rest-frame $(B - K)$ colours. In our models we adopt the same normalization as in our previous papers. As we and others have noted before (Metcalfe et al. 1991; Shanks 1989), normalizing galaxy counts at brighter ($B \sim 16$ mag) magnitudes requires that substantial amounts of *B*-band evolution must take

Table 4. Luminosity function parameters and rest-frame colours used in all models.

Parameter / type	E/S0	Sab	Sbc	Scd	Sdm
ϕ^* (Mpc$^{-3}\times10^{-4}$)	10.2	5.09	6.82	3.0	1.5
α	−0.7	−0.7	−1.1	−1.1	−1.5
M_B^*	−21.0	−21.0	−21.32	−21.44	−21.45
$(B - K)$	3.93	3.78	3.51	2.90	2.26
$(I - K)$	1.57	1.63	1.7	1.47	1.0

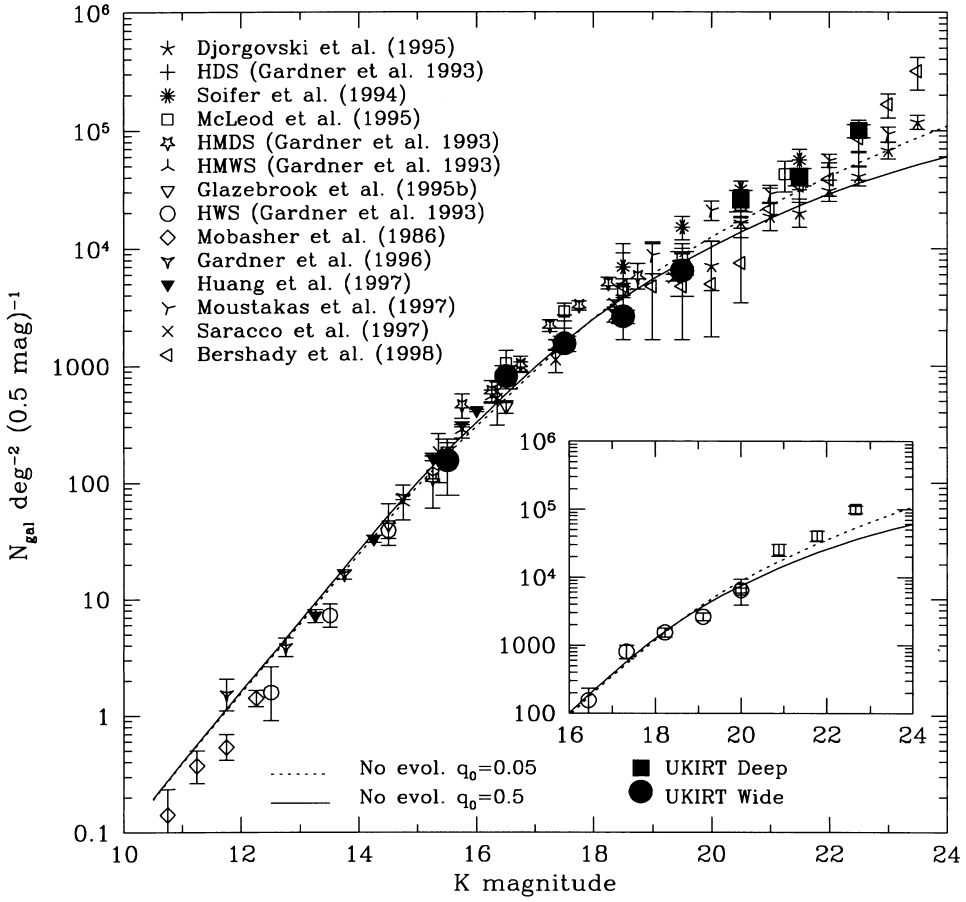

Figure 1. Differential galaxy number counts in half-magnitude intervals as a function of *K*-limiting magnitude as measured from our UKIRT wide survey (filled circles) and our UKIRT deep survey (filled squares). Also plotted is a compilation of published *K*-band number counts and the predictions of the two non-evolving models described in the text for $q_0 = 0.05$ (dotted line) and $q_0 = 0.5$ (solid line). Error bars are calculated from the number of galaxies in each magnitude bin using Poisson $\sqrt{n}$ counting statistics. The inset shows the number counts from our surveys alone (this time with open symbols) compared with the non-evolving predictions.

place at relatively low ($z < 0.1$) redshift. Conversely, by normalizing at $18\,\mathrm{mag} < B < 20\,\mathrm{mag}$ we allow a better fit at intermediate magnitudes and reduce the amount of evolution required to fit the counts at fainter magnitudes, but underpredict counts at brighter magnitudes. Support for this high normalization has recently come from the 2-μm all-sky survey (2MASS, Kleinmann 1992), which finds a significant underdensity in the galaxy counts at $10\,\mathrm{mag} < K < 13\,\mathrm{mag}$ and is discussed further in a forthcoming paper (Shanks et al., in preparation).

Could substantial galaxy luminosity evolution be occurring at these low redshifts? Bright, $K < 20\,\mathrm{mag}$ samples are dominated by early-type galaxies. Morphologically segregated number counts from the Medium Deep Survey (Griffiths et al. 1994) and the *Hubble Deep Field* (HDF) (Williams et al. 1996) indicate that counts for elliptical galaxies at intermediate magnitude levels ($18\,\mathrm{mag} < I_{816\mathrm{W}} < 22\,\mathrm{mag}$) follow the predictions of a non-evolving model (Driver et al. 1998; Driver, Windhorst & Griffiths 1995; Glazebrook et al. 1995a). This would argue against an evolutionary explanation for the steep slope observed in the range $10\,\mathrm{mag} < K < 14\,\mathrm{mag}$. We are then faced with finding an alternate explanation for the number counts underdensity; one suggestion comes from the findings of Metcalfe, Fong & Shanks (1995b) and Bertin & Dennefeld (1997), who both show how scale errors (i.e. non-linearities) in the photographic Automatic Plate Measuring (APM) photometry of Maddox et al. (1990) could result in anomalously low galaxy counts at brighter magnitudes in *B*-selected surveys; however, the *K*-band surveys are carried out with electronic detectors, and such effects are likely to be less significant (although the Mobasher, Ellis & Sharples (1986) survey was based on objects selected from *B*-band photographic plates). Another suggestion is that our galaxy may reside in a locally underdense region of the Universe (Shanks 1989); however, to produce the observed discrepancy in the number counts would require this void to be $\sim 150\,h^{-1}$ Mpc in size, and to have an underdensity of ~30 per cent; fluctuations on this scale are difficult to understand in terms of current cosmological models and also measured large-scale bulk flows. Some evidence for a local void has come from studies of peculiar velocities of type Ia supernovae (Zehavi et al. 1998), although the size of the void they detect ($\sim 70\,h^{-1}$ Mpc) is not large enough to explain the low number counts. For time being, resolution of this issue awaits a better determination of the bright end of the *B* and *K* number counts by much larger surveys.

3.2 Optical–infrared colours

Fig. 2 shows the $(B-K)$ versus *K* colour–magnitude diagram for both samples for galaxies in both the 'wide' (open triangles) and 'deep' (open circles) surveys. Additionally, the median $(B-K)$ colour computed in one-magnitude bins is shown (filled squares) and compared with the values given in Gardner et al. (1993) (filled triangles). Error bars on the median colours were calculated using a bootstrap resampling technique.

Incompleteness is represented by the dashed line; objects to the right of this line have $B > 28\,\mathrm{mag}$. Several objects have been identified in both surveys that are undetected in *B* but are detected in *K*, and these are plotted in the figure as right-pointing arrows and are discussed in Section 4.3. Considering the median $(B-K)$ colour, we see that it increases steadily until around $K \sim 18\,\mathrm{mag}$, where it reaches a maximum value of ~6. After this, it sharply becomes bluer and this blueward trend continues to the faintest limits we have measured. Our median colours agree well with

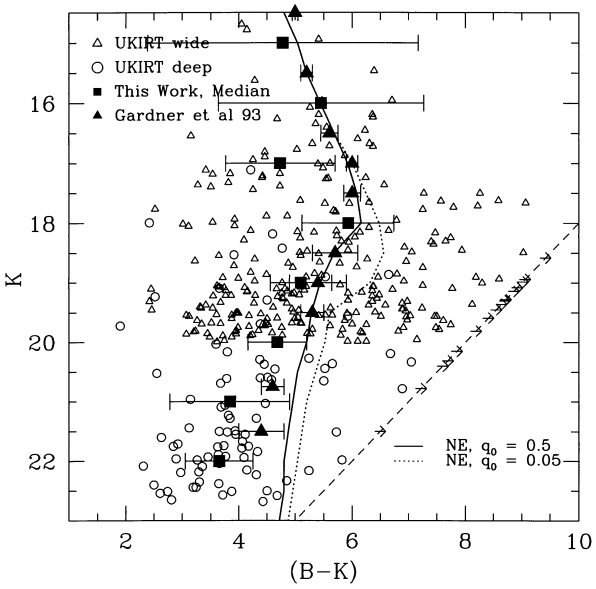

Figure 2. *K* magnitude against $(B-K)$ colour for galaxies in the wide survey (open triangles) and the deep survey (open circles). Also plotted are the median $(B-K)$ colours in 1-mag bins for our survey (filled squares), with 1σ error bars calculated by a bootstrap resampling technique. Triangles show median colours from the compilation of Gardner et al. (1993). The dotted line represents the region of incompleteness; galaxies below this line have $B < 28\,\mathrm{mag}$. Dropouts are plotted at the upper limit of their colours as right-pointing arrows. For comparison, predictions from two non-evolving models are plotted; one for $q_0 = 0.05$ (dotted line) and one for $q_0 = 0.5$ (solid line).

those in Gardner et al. (1993), even for brighter bins where the number of galaxies in our sample is much smaller. We have also found our $K < 18$ $(B-K)$ distributions to be in good agreement with the brighter survey of Saracco et al. (1997). Finally, it is worth noting the extremely large spread in colour in this diagram – at $K \sim 20$, $(B-K)$ ranges from 2 to 9.

Additionally, the predictions of the two non-evolving models shown in Fig. 1 are plotted. In general, the $q_0 = 0.5$ model gives a slightly better fit to the observed colours, at least until $K \sim 20\,\mathrm{mag}$; at fainter magnitudes, galaxies are significantly bluer than both model predictions.

Another interesting feature in Fig. 2 is the apparent deficit of galaxies with magnitudes in the range $K > 21\,\mathrm{mag}$ and colours from $(B-K) > 5$. The different areal coverages and depths of the two surveys doubtless exacerbates this effect, but its presence is consistent with a population of galaxies that turns rapidly blueward faintwards of $K \sim 20$. We discuss the origin of this feature in Section 4.2.

In Fig. 3 we plot $(B-R)$ versus $(R-K)$ colour for all the galaxies in the wide survey (open circles) and the deep survey (filled circles). Fainter galaxies are plotted with smaller symbols. The 'deep', $K < 22.75\,\mathrm{mag}$ sample is much bluer in both $(B-R)$ and $(R-K)$ than the 'wide' than the brighter, $K < 20\,\mathrm{mag}$, sample, which reaches $(R-K) \sim 6$. We also plot colour–colour tracks representing the different galaxy classes in our evolutionary models. We defer discussion of these tracks until the following section.

4 INTERPRETATION AND MODELLING

4.1 Outline of the models

To interpret our results we investigate variants of a pure

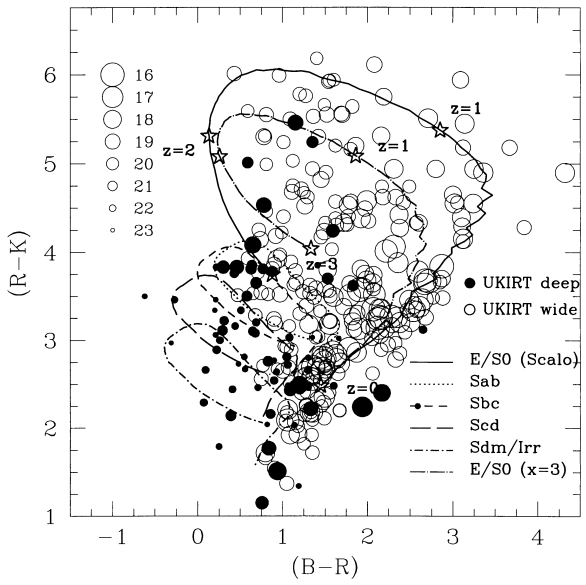

Figure 3. $(B-R)$ colour against $(R-K)$ colour for galaxies in the wide survey (open circles) and the deep survey (filled circles). Fainter galaxies are plotted with smaller symbols. Also shown are evolutionary tracks for the five classes of galaxy in our model, as computed from the population synthesis code of Bruzual & Charlot (1993). The solid line represents type E/S0, the dotted line Sab and the dashed, long-dashed and short dash–dotted lines types Sbc, Scd and Sdm respectively. The long dot–dashed line shows the path of of an elliptical galaxy which has an IMF slope, x, of 3. For the elliptical tracks, the labelled stars indicate the redshift.

luminosity evolution (PLE) model in which star formation increases exponentially with look-back time. Earlier versions of these models are discussed in our previous papers (Metcalfe et al. 1991, 1995a, 1996). In this paper we consider cosmologies in which $\Lambda = 0$ and take $H_0 = 50\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, although changing the value of H_0 does not markedly affect any of the conclusions of this paper. Two values of the deceleration parameter, $q_0 = 0.05$ and $q_0 = 0.5$, are adopted, corresponding to open and flat cosmologies respectively.[1] The input parameters to our models (given in Table 4) consist of observed *local* galaxy parameters (namely, rest-frame colours and luminosity functions) for each of the five morphological types (E/S0, Sab, Sbc, Scd and Sdm) that we consider in our models. These morphological types are divided into elliptical (E/S0) and spiral (Sbc/Scd/Sdm), and these two classes are each given a separate star formation history. We could, in principle, subdivide the spirals into different morphological types, each with different star formation histories, but for simplicity we do not; $(k+e)$ corrections for the different types are fairly similar to each other in these models in any case. Instead, taking a Sbc model as representative of all types, we produce the other types by normalizing the Sbc track to the observed rest-frame $(B-K)$ colours from Table 4. As we have already explained in Section 3.1, our ϕ^* values are chosen to match the galaxy counts at $B \sim 18$–20 and we seek to explain the low number counts at bright magnitudes from a combination of photometric errors and anomalous galaxy clustering, rather than substantial evolution at low redshift. Our models also include the effects of the Lyman α forest, and, for spiral types, dust extinction corresponding to the Large Magellanic Cloud as described in Pei (1992).

PLE models have difficulty in correctly reproducing the observed K-band redshift distribution. Even essentially passive evolution over-predicts the numbers of galaxies seen at $1 < z < 2$, and this fact has been used by several authors (most recently Kauffmann & Charlot (1998)) to argue for the hierarchical, merger-driven galaxy formation models in order to reduce the mean redshift of K-selected galaxy redshift distributions. However, in Metcalfe et al. (1996) we showed how PLE models *could* produce a K-band redshift distribution compatible with the observations of Cowie et al. (1996) by assuming a very steep slope ($x = 3$) for the stellar initial mass function (IMF) in E/S0 types. This steep, low-mass-star dominated IMF reduces the amount of passive evolution observed in K and allows us to reproduce the observed redshift distribution, which is close to the predictions of the non-evolving model. It should be noted however, that if the incompleteness in the Cowie et al. (1996) $n(z)$ were entirely caused by $z > 1$ galaxies, then the deficit relative to the PLE model (with standard IMF) is only $\sim 3\sigma$ from $\sqrt{N}$ statistics alone and is likely to be less significant when galaxy clustering is taken into account. Conversely, the B-band redshift distribution from the same survey does have a large fraction of high-redshift ($z > 1$), high-luminosity ($L > L^*$) galaxies. However, this is not a problem for our models. Our steep luminosity function slope for spiral types combined with a small amount of internal dust extinction (corresponding to $A_B = 0.3$ mag at $z = 0$) allows us to reproduce this result.

Although the resolution of the data set discussed in this paper is too low to extract useful size information for the faint galaxy population, we mention in passing that observed faint galaxy sizes are compatible with these models, models, as we have demonstrated in a recent paper (Shanks et al. 1998). We have used a modified version of Howard Yee's PPP code to produce a simulated HDF frame which contains galaxies with parameters based on the output of our PLE models. For this image, spiral galaxies were generated assuming that luminosity evolution only affects disc surface brightness and not size, combined with Freeman's (1970) law to relate disc size to absolute magnitude at $z = 0$. For bulges, we adopted the diameter–magnitude relation of Sandage & Perelmuter (1990). By applying our photometry package to this image, we account for selection effects in a realistic manner. We find that the galaxies in the simulation have sizes not dissimilar from those measured in the HDF, in part because many are intrinsically small galaxies seen well down the luminosity function (LF).

In Fig. 4, the effect that changing the IMF slope for early types has on predicted redshift distributions is illustrated. The dotted line shows a model computed using $\tau = 1.0$ Gyr and a Scalo (Scalo 1986) IMF, whereas the solid line shows the $x = 3$ model, which has $\tau = 2.5$ Gyr. Also plotted are the redshift distributions taken from table 1 of Cowie et al. (1996). Objects that were marked as unidentified in each band in this table are shown in their respective incompleteness boxes. We have normalized both model predictions to the total number of *observed* objects (i.e. unidentified objects are included). We take $q_0 = 0.05$ and $z_f = 6.4$ in all our models. The star formation e-folding time, τ, is 9.0 Gyr for all the spiral types. The $x = 3$ model, as it is dominated by dwarf stars, produces *much smaller* amounts of passive evolution in the K band and as a consequence the median redshift of galaxies

[1] Zero-curvature cosmologies with $\Omega = 0.3$ and $\lambda_0 = 0.7$ are almost identical to our $q_0 = 0.05$ model because comoving distance and look-back time as a function of redshift are almost identical to that found in the zero-Λ models, at least to $z \sim 1.5$, and are within ~ 20 per cent to $z \sim 2.2$ (Peebles 1984).

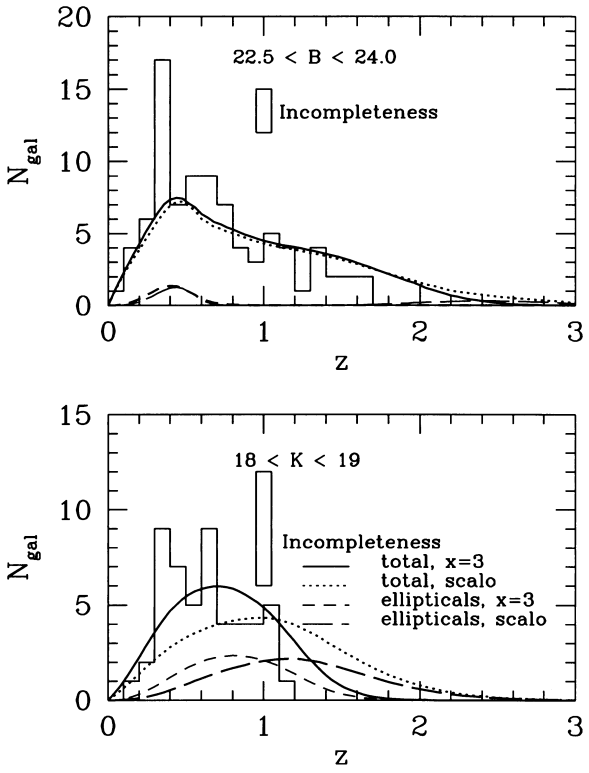

Figure 4. Predicted redshift distributions for both 'Scalo' ($\tau = 1.0$ Gyr) and '$x = 3$' ($\tau = 2.5$ Gyr) models (dotted lines and solid lines) compared with the results of the K (lower panel) and B (upper panel) redshift distributions taken from table 1 of Cowie et al. (1996). Incompleteness is represented by the open boxes. Also plotted are the predicted redshift distributions for elliptical types (long-dashed and short-dashed lines) for the 'Scalo' and '$x = 3$' models respectively.

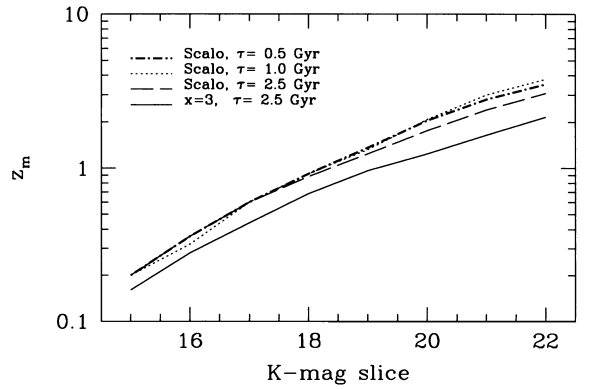

Figure 5. The median redshift, $z_{\rm med}$, of elliptical galaxies in our model populations for the magnitude slices used in Figs 8 and 9.

in this model is much lower than in the Scalo model, which produces an extended tail in the redshift distribution. Notice that the effect of changing the IMF slope in the B band is negligible.

This point is further illustrated in Fig. 5, where we show the median redshift of elliptical galaxies with a range of star formation histories for several K-selected magnitude slices. This panel shows that, in all all magnitude ranges, elliptical galaxies in the 'Scalo' model have a higher median redshift than in the '$x = 3$' model, irrespective of τ (demonstrating that the differences in

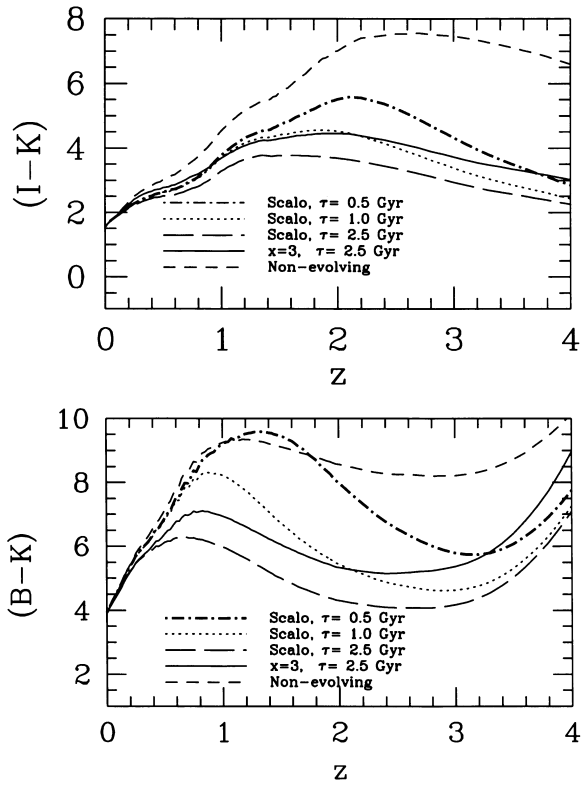

Figure 6. Lower panel: $(B - K)$ colour against redshift, z, for an elliptical galaxy with an $x = 3$ IMF and $\tau = 2.5$ Gyr (solid line), and for the Scalo IMF with three values of τ: $\tau = 0.5$ Gyr (dot–dashed line), 1.0 Gyr (dotted line) and 2.5 Gyr (long-dashed line). Also shown is a non-evolving model (short-dashed line). Upper panel: as for the lower panel, but with $(I - K)$ colour.

Fig. 4 are caused by changing the IMF and not τ). This point is discussed further in Section 4.2.

In Fig. 6 we illustrate the effect that varying the IMF slope and τ has on the E/S0 colour–redshift relation. The dot–dashed, dotted and long-dashed lines show Scalo IMF tracks with $\tau = 0.5$, 1.0 and 2.5 Gyr respectively, whereas the solid line shows an $x = 3$ model with $\tau = 2.5$ Gyr. As is apparent from the plot, at $z \sim 1$ the $(B - K)$ colour depends *very sensitively* on the assumed value of τ and the slope of the IMF. For a given IMF, longer τ values cause the peak in the $(B - K)$–z relation to shift to progressively higher redshifts, a point we will return to in Section 4.2. Finally, the $x = 3$ IMF model track, as it is dominated by low-mass stars, is in general redder (for a given value of τ) than the Scalo IMF track. We note that the single-burst models of Worthey (1994) have shown that the amount of passive evolution in the rest frame I and R bands is sensitive to metallicity variations; consequently at $z =$ 1–2 the predicted K magnitude from our essentially single-burst model is also expected to be approximately independent of metallicity. Therefore reasonable changes in metallicity seem unlikely to change the rate of passive evolution in the K band, which, in the context of these models, only appears to depend on the slope of the IMF.

We now return to Fig. 3, where we have overplotted on our $(B - R)$–$(R - K)$ diagram evolutionary tracks for each of the galaxy types in our model. The dot–dashed and solid lines show the path of an elliptical galaxy, whereas the remaining tracks show the four spiral types in the models. On the elliptical tracks, the labelled stars indicate the redshift at selected intervals. Generally,

there is good agreement between the model tracks and the observed colours. The low-redshift ($0.5 < z < 0.0$) elliptical track is populated by the brightest galaxies in our sample. Generally, we do not find a that large fraction of objects have colours that cannot be reproduced by the tracks [in contrast with the findings of Moustakas et al. (1997), who claimed there was a significant population of objects with colours that could not be reproduced by the models]. There are *some* objects that appear to be redder than the tracks, but these do not constitute a significant fraction of the observed galaxy population. The broad range of $(B-K)$ colours of objects with $(R-K) > 4$ (which according to the models should all be E/S0 galaxies) is a real effect which we interpret as evidence of the diverse range of star formation and metallicity histories present in the elliptical population.

4.2 Counts and colour distributions

In Fig. 7, we present a compilation of all published galaxy number counts in B and K bandpasses, as well as the predictions from our two evolutionary models.

What is immediately apparent is that both the models discussed provide a good fit to the number counts over a wide magnitude range for the low q_0 case. Furthermore, in both B and K bandpasses at faint magnitudes the differences between the Scalo and '$x = 3$' models is *much smaller* than the differences between the high- and low-q_0 models. It is also apparent, as has been noted by many others (Metcalfe et al. 1996; Campos 1997; Babul & Rees 1992) that, unless another population (such as bursting dwarfs) is added to the models then the $q_0 = 0.5$ model underpredicts the observed B-band galaxy counts. In the K band this model is marginally excluded at the faintest magnitudes.

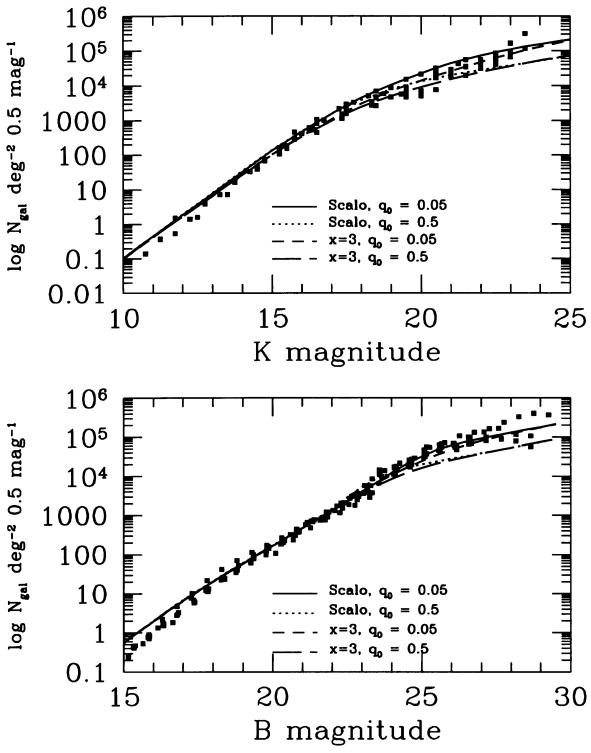

Figure 7. Galaxy number counts for two evolutionary models – the 'Scalo' model (solid and dotted lines) and the '$x = 3$' model (short-dashed and long-dashed lines) – for low and high values of q_0 respectively. K-band counts are shown in the upper panel; B-band ones in the lower one.

Number counts are a coarse test of any model; the extremely deep optical photometry covering the Herschel Field is one of the main virtues of this work. To test our models more stringently we therefore now investigate our observed optical–infrared colour distributions. We confine our discussion to the $q_0 = 0.05$ case, as these models provide a better fit to the counts, although the colour distribution for the $q_0 = 0.5$ models will be very similar.

In Fig. 8 we compare in detail our model predictions with our observations for K-selected $(B-K)$ colour distributions and compare these with the data (shown as the solid histograms) in six 1-mag slices from $K = 16$ mag to $K = 22$ mag. The 'Scalo' model is shown as dotted and dot–dashed lines and the '$x = 3$' model is represented by a solid line. All models have been convolved with a conservative 0.3-mag colour measurement error (overestimating errors in the brighter magnitude slices). For the non-evolving model, as for the evolving models, k-corrections are calculated from the models of Bruzual & Charlot (1993).

Incompleteness in the histograms was determined by counting how many galaxies had K magnitudes but were not detected in B in each magnitude slice (and consequently have undetermined colours). The colour of the *reddest* object that could be detected is represented by the downwards-pointing arrow, which is plotted at the colour corresponding to the K magnitude at the centre of each magnitude slice (these objects have $B > 28$ mag). The number of non-detections is shown by the size of the box. Models have been normalized to the total number of galaxies in each slice; in the latter panels this includes B-band non-detections.

The non-evolving model predictions (dashed line) rapidly diverge from the data. Faintwards of the $K = 17$–18 mag magnitude slice a bump appears in the model colour distributions, corresponding to the unevolved, red, $(B-K) \sim 8$ elliptical population. At approximately the same magnitude limit and at bluer colours [this time at $(B-K) \sim 4$] a second peak becomes apparent. This corresponds to the model spiral population, and at fainter magnitudes this peak becomes the more prominent of the two. This is a consequence of our adopted luminosity function parameters and morphological mix.

Qualitatively, the evolutionary models follow the pattern of the non-evolving model, though the location and amplitude of the E/S0 peaks depend sensitively on the choice of IMF and τ. On the blueward side of the distributions, the spiral model colours appear to be a good fit to the observations. Referring to Fig. 6, it is apparent that beyond $z \sim 0.3$, the $(B-K)$ colour for the $\tau = 0.5$ Gyr E/S0 model is redder than that for the $\tau = 1.0$ Gyr model; furthermore, as the median redshift of both populations is almost the same (Fig. 5), changes in the colour distributions are almost completely a product of differences in the $(B-K)$–z relation. In general, Bruzual & Charlot (1993) type models for elliptical galaxies with short τ values predict a $(B-K)$–z relation in which galaxies are at their reddest at $z \sim 1$, and become bluer thereafter until $z \sim 3$, at which point they turn redder once more as a result of Lyman α forest reddening. The position of this turn-over depends on the assumed τ, the IMF and the redshift of formation z_f.

If we now we consider the '$x = 3$', $\tau = 2.5$ Gyr model, we see that its colour distributions are in general *bluer* than those predicted from the 'Scalo' $\tau = 0.5$ or $\tau = 1.0$ Gyr models. This is partly a consequence of the longer τ adopted in this model, but also of the fact that (as we have seen in Fig. 5) galaxies lie at a lower redshift than the 'Scalo' model and consequently are sampling a bluer region of the $(B-K)$–z relation. Therefore, despite the $x = 3$ model being low-mass star dominated, the larger

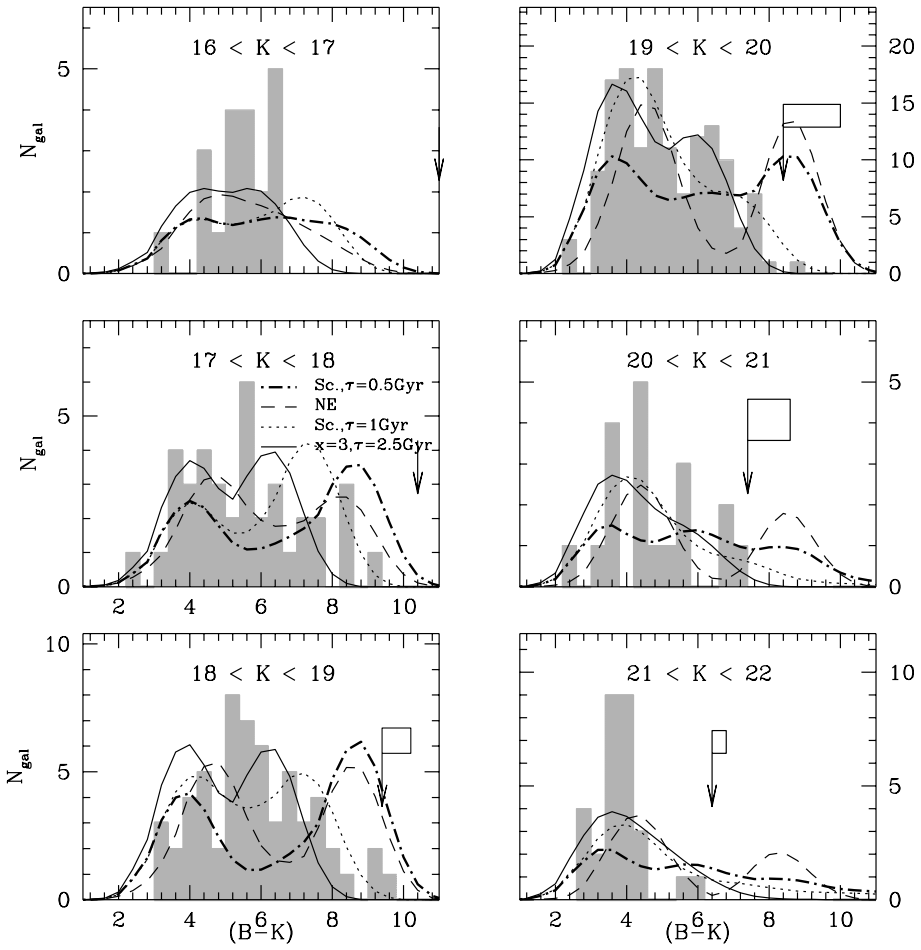

Figure 8. *K*-selected $(B-K)$ colour distributions (solid histogram) in six magnitude slices for both the wide (brightest four slices) and deep (faintest two slices) surveys. Also shown are predictions from four models: '$x=3$' (solid line), 'Scalo' (dotted and dot–dashed lines) and 'non-evolving' (long-dashed line). The two 'Scalo' models have differing star formation histories: the dotted and heavy-dashed line has $\tau=1.0$ Gyr while the heavy dot–dashed line has $\tau=0.5$ Gyr. The $x=3$ model has $\tau=2.5$ Gyr. The model predictions have been normalized to the *total* number of galaxies in each slice. In the fainter ($K>18$ mag) bins this includes the *B*-band non-detections discussed in Section 4.3 and the number of such objects in each magnitude slice is represented by the incompleteness boxes. The position of the downward-pointing arrow represents the colour of the *reddest* object at the centre of each magnitude slice that could be detected; objects in the incompleteness boxes must lie rightwards of this arrow. For all models in this plot we adopt $q_0=0.05$.

value of τ allows us to produce colours that are blue enough to match the observations.

Returning to Fig. 8, we can now understand the origin of the differences between the $(B-K)$ colour distributions predicted by the models. Considering first the brighter magnitude slices in Fig. 8, at $K=17.0$–18.0 mag we see that the 'Scalo' $\tau=1.0$ Gyr model predicts a peak at $(B-K)\sim 7$, which appears in the '$x=3$' model at $(B-K)\sim 6$. At this magnitude, elliptical galaxies in the 'Scalo' model have $z_{\rm med}\sim 0.6$, whereas for the '$x=3$' model $z_{\rm med}\sim 0.4$ (from Fig. 5); as a consequence 'Scalo' model colours are redder than '$x=3$' ones. Furthermore, the essentially passively evolving Scalo $\tau=0.5$ Gyr model contains insufficient recent star formation to shift the peak at $(B-K)\sim 9$ blueward; as we have seen from our considerations of the non-evolving model, this peak is from unevolved ellipticals. At fainter ($K>20$ mag) magnitudes, both the $x=3$ model and the Scalo $\tau=1.0$ Gyr model turn sharply blueward; the Scalo $\tau=0.5$ Gyr model colour distribution, however, has an extended tail which reaches $(B-K)\sim 9$–10.

It is clear from Fig. 8 that either a Scalo IMF with a $\tau=1.0$ Gyr or an $x=3$ IMF with $\tau=2.5$ Gyr model reproduces the main features of the data histograms, including the broadening of the colour distribution at the $K<18$ mag slice, and its subsequent narrowing and blueward trend in the fainter slices. Furthermore, as we will see in the following Section, these models are consistent with the redwards limits of our data. Of these two models, the $x=3$, $\tau=2.5$ Gyr model, as commented in Section 4.1, gives a much better fit to the observed *K*-band redshift distributions. As is apparent from Figs 5 and 6, distinguishing between the effects of changes in IMF slope or star formation history on the colour–redshift relation is difficult. Making this choice requires additional information such as the redshift distribution of *K*-selected galaxies.

Finally, these models also allow us to understand the deficit of galaxies in the colour-magnitude relation at $(B-K)>5$ and $K>21$ mag, which we commented upon in Section 3.2. First, as a consequence of our choice of luminosity function parameters, elliptical counts turn over at $K\sim 20$ and spiral types dominate faintwards of this. Secondly, from Fig. 5 we see that at $K\sim 21$ E/S0s are at $z\sim 1$; from these redshifts to $z>3$, Fig. 6. indicates that elliptical colours turn sharply blueward. In combination, these factors produce the rapid blueward trend observed at $K>21$ mag

and the absence of redder objects, both of which happen in all our models regardless of our choice of IMF or τ.

4.3 $(I-K)$ distributions and extremely red galaxies

In the previous section we discussed the broad characteristics of the K-selected $(B-K)$ colour distribution and showed how these features can be understood in terms of our models; in this section we turn our attention to very red objects, which are at the outskirts of the colour distributions. The red colours of these objects imply they are E/S0 galaxies at moderate ($z>1$) redshift and consequently it is interesting to see whether their observed number density is consistent with current galaxy formation scenarios. As previous authors (Cowie et al. 1994; Moustakas et al. 1997; Zepf 1997; Barger et al. 1998) who focused their attentions upon these galaxies discussed them in terms of K-selected $(I-K)$ distributions, here we introduce our observations in these bandpasses. Following these workers, we define an 'extremely red galaxy' as an object with $(I-K)>4$; from Fig. 6 we see that at $(I-K)\sim 4$, $z\sim 1$. In this section we will make comparisons with models generated using the same set of parameters as the K-selected $(B-K)$ distributions discussed previously. In interpreting $(I-K)$ colours, one faces some significant differences from $(B-K)$ observations. As is apparent from Fig. 6, $(I-K)$ colour is relatively insensitive to the amount of star formation; at $z\sim 1$ the difference in $(I-K)$ between elliptical galaxies with a Scalo IMF and $\tau=0.5, 1, 2.5$ Gyr is 0.5; by comparison, at this redshift $(B-K)$ ranges from ~3 to ~5. This difference is reflected in the tightness of the K-selected $(I-K)$ distributions plotted in Fig. 9. This insensitivity makes $(I-K)$ distributions relatively poor probes of the elliptical star formation history. However, it is precisely this behaviour that makes $(I-K)$ colours useful in detecting high-redshift, evolved ellipticals; Fig. 6 shows that, until $z\sim 2$, $(I-K)$ colour is approximately proportional to z, and is relatively insensitive to variations in τ or the IMF.

Turning to work from the literature, Cowie et al. (1994), in a 5.9 arcmin^2 survey, detected 13 objects with $(I-K)>4$ to $K<20.9$ mag, and concluded that these galaxies do not dominate the faint $K>20$ mag population. They also concluded that less that 10 per cent of present-day ellipticals could have formed in single-burst events. In contrast, Moustakas et al. (1997) found 8 galaxies with $(I-K)>4$ and $K<22$ mag over a small, 2 arcmin^2 area. They also isolated a population of objects with blue optical $[(V-I)<2.5]$ and red near-infrared colours and argued that these colours could not be reproduced with Bruzual & Charlot (1993)-type models. Although we do not have V-band photometry, the top middle and top left of our Fig. 3 roughly correspond to the regions highlighted in the colour–colour plot shown in fig. 9 of Moustakas et al. These 'red outlier' objects, as Moustakas et al. describe them, viewed in the context of the models discussed here, are most likely elliptical galaxies at $1<z<2$. Finally, both Zepf (1997) and Barger et al. (1998) investigated extremely red objects in the *Hubble Deep Field*

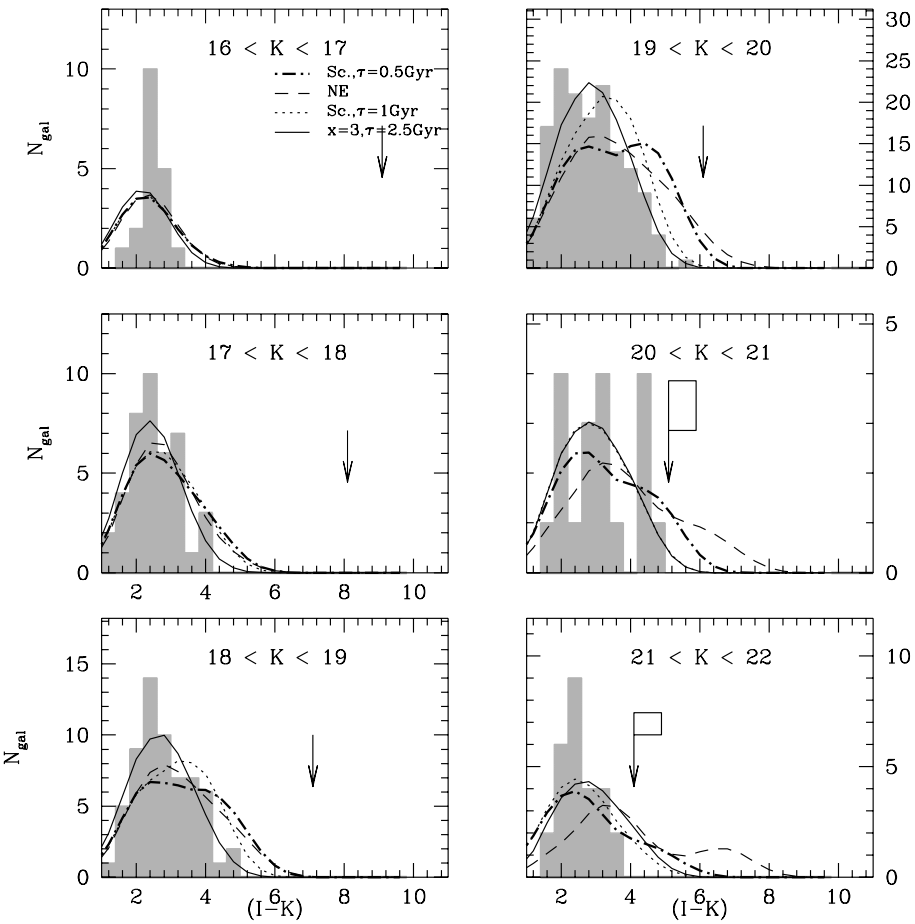

Figure 9. K-selected $(I-K)$ colour distributions for both the wide and deep surveys (solid histogram). All symbols are as in Fig. 8.

Table 5. Numbers of objects with $(I-K)>4$ per arcmin^2 with $\pm1\sigma$ errors. The survey area in arcmin^2 is shown in parentheses. In the deep survey ($20<K<22$) there are seven objects that are undetected in I, and therefore only lower limits can be placed on their colours. They are included in their respective bins. The number quoted from Moustakas et al. (1997) covers the magnitude range $20<K<22$.

Author	$19<K<20$		$20<K<21$		$21<K<22$	
Model ($x=3$)	0.3		1.4		1.8	
Model (Scalo, $\tau=1.0$ Gyr)	0.6		1.4		1.2	
Model (Scalo, $\tau=0.5$ Gyr)	1.1		3.1		2.8	
This work	$0.5^{+0.1}_{-0.1}$	(47)	$3.9^{+2.1}_{-1.4}$	(1.8)	$1.1^{+1.5}_{-0.7}$	(1.8)
Barger et al. (1999)	$0.2^{+0.1}_{-0.1}$	(62)	$1.2^{+0.5}_{-0.4}$	(7.8)	–	
Cowie et al. (1994)	$0.3^{+0.4}_{-0.3}$	(7.1)	$1.5^{+0.7}_{-0.5}$	(5.9)	$0.3^{+0.5}_{-0.2}$	(5.9)
Moustakas et al. (1997)	–		$4.0^{+2.0}_{-1.4}$	(2)		

(HDF). Zepf (1997), using publicly available Kitt Peak near-infrared imaging data covering the 5 arcmin^2 of the HDF, found that to a 50 per cent completeness limit of $K<22$ mag there were $\sim$2 objects with $(V_{606}-K)>7$. Barger et al. (1998), in a much larger $\sim$60 arcmin^2 infrared survey comprising both the HDF and the HDF flanking fields, found 12 galaxies with $(I-K)>4$ and $19<K<20$.

Table 5 presents a compilation of observations from these papers as well as our current work. We also show the predictions of the $x=3$ and Scalo $\tau=0.5$ model (our model that most closely approximates a single burst), normalized to the total numbers of objects in each magnitude slice. Generally, in the brighter bin ($19<K<20$) the only substantial comparison that we can make is with Barger et al. We see a factor of $\sim$3 more objects than they do, although this represents only a 2σ discrepancy. In the fainter bin $20<K<22$, our numbers of red objects agree well with those of Moustakas et al. Although our numbers are higher than Cowie et al. (1994) in this range, there is still no significant discrepancy, owing principally to our small area.

Barger et al. compared their observations with the predictions of a passively evolving $\tau=0.1$ Gyr model elliptical population, and concluded that the numbers of $(I-K)>4$ objects observed disagreed with the predictions of this model. Our $\tau=0.5$ Gyr distribution (solid dot–dashed line) in Fig. 9 is quite similar to this model, and in the $19<K<20$ magnitude slice we can see from Table 5 that it does indeed produce a large number of objects with $(I-K)>4$, which are not seen in our observations. We therefore conclude, as Barger et al. did, that the low numbers of objects with $(I-K)>4$ observed at $19<K<20$ disfavour non-evolving and passively evolving models (the constraints on the models that can be obtained from the distributions at fainter magnitudes are less significant). However, we note that the $x=3$, $\tau=2.5$ Gyr and $\tau=1.0$ Gyr Scalo models do reproduce the observed colour distributions correctly. This is because both these models contain enough ongoing star formation to move galaxy colours sufficiently blueward to match the observations.

Lastly, could a merging model, like those adopted by Kauffmann & Charlot (1998), produce these results? It is not possible to rule this model out from the redshift data and it may even be said that the difficulty that models with standard Scalo IMFs have in fitting the $K<19$ redshift distribution is evidence in favour of the merging scenario. We therefore consider whether the K-selected $(B-K)$ and $(I-K)$ colour distributions presented here can discriminate between a PLE ($x=3$) model and a merging model. The K-selected $(B-K)$ distributions in Fig. 8 can be understood in terms of either scenario; the red galaxies that are missing at $17<K<22$ may be either because of their becoming blue at high redshift, owing to continuing star formation or passive evolution, or because they are fainter than expected owing to demerging (although our lack of red galaxies even at the faintest limits implies either several magnitudes of fading or that the pre-merger components are blue). However, as we have already noted, the $(I-K)$–z relation for E/S0 galaxies is insensitive to changes in τ or the IMF. Furthermore, in the range $18<K<20$ some galaxies with the $I-K>4$ colour expected of $z=1$ elliptical galaxies are detected (Fig. 9). Indeed, the number of these galaxies we see is within a factor of $\sim$2 of the value predicted on the basis of the Scalo, $\tau=0.5$ Gyr model and in even better agreement with the $x=3$, $\tau=2.5$ Gyr models or even a Scalo $\tau=1.0$ Gyr model. If the merging model is the explanation of the large deficiency of red galaxies in $(B-K)$ in comparison with passively evolving models, then it might be expected that a similar deficiency should be seen in $(I-K)$. As the deficiency in $(I-K)$ is less, then this might be taken to be an argument *against* merging and *for* the $x=3$ PLE model.

5 CONCLUSIONS AND SUMMARY

In this paper we have presented the results of two near-infrared surveys to $K\sim20$ and $K\sim23$ which cover our ultra-deep ($B\sim28$) optical fields. We draw the following conclusions from this work.

(i) Our K number counts are consistent with the predictions of non-evolving models with $0\leqslant q_0\leqslant0.5$.

(ii) As previously noted by Cowie et al. (1996), Metcalfe et al. (1996) and Kauffmann & Charlot (1998), the $18<K<19$ $n(z)$ of Cowie et al. (1996) is also well fitted by non-evolving models. However, passively evolving models with a Salpeter/Scalo IMF predict too many galaxies with $z>1$. Dynamical merging is one possible solution to reduce the numbers of these galaxies, but we have shown that a dwarf-dominated IMF for early-types could offer an alternative explanation.

(iii) Our K-selected $(B-K)$ colour distributions display a strong bluewards trend for galaxies fainter than $K\sim20$, confirming results previously observed in the shallower surveys of Gardner et al. (1993).

(iv) At brighter magnitudes ($K<20$ mag) our K-selected $(B-K)$ distributions indicate a deficiency of red, early-type galaxies at $z\sim1$ compared with the predictions of passively evolving models. This implies either a PLE model where star formation continues at a low level after an initial burst or dynamical merging.

(v) At fainter magnitudes ($20<K<22$) the continuing

bluewards trend observed in $(B-K)$ can be explained purely in terms of passively evolving PLE models with no need to invoke any additional mechanisms.

(vi) Our observed numbers of $(I-K)>4$ galaxies at $K\sim 20$ are lower than the predictions of passively evolving models or PLE models with a low level of continuing star formation, suggesting that at least part of the larger deficiency observed in $(B-K)$ at $K\sim 20$ may be caused by star formation rather than dynamical merging.

(vii) In the range $19<K<20$, where our statistical uncertainties are lowest, we detect 0.5 ± 0.1 red galaxies per arcmin^2 with $(I-K)>4$. We see a factor of $\sim$3 more objects than Barger et al. do, although this represents only a 2σ discrepancy. The PLE models discussed here suggest that these galaxies will have redshifts $1<z<2$. The numbers of these galaxies are consistent with the predictions of PLE models with small amounts of ongoing star formation.

6 FUTURE WORK

We have recently completed a *H*-band survey of the Herschel field using the 1024×1024 Rockwell HgCdTe array at the Calar Alto 3.5-m telescope. This survey is complete to $H<22.6$ mag and covers the entire area of the Herschel Deep Field. Analysis of this data set will be presented in future papers.

ACKNOWLEDGMENTS

HjMcC acknowledges financial support from PPARC and the hospitality and generosity of Dr Mariano Moles at IMAFF, Madrid, where an earlier version of this paper was written. Support for this trip was also provided by the University of Durham Rolling Grant for Extragalactic Astronomy. This work was supported in part by a British Council grant within the British/Spanish Joint Research Programme (Acciones Integradas). NM acknowledges partial PPARC support.

REFERENCES

Babul A., Rees M. J., 1992, MNRAS, 255, 364
Barger A., Cowie L. L., Trentham N., Fulton E., Hu E., Songaila A., 1999, AJ, 119, 102
Bershady M. A., Lowenthal J. D., Koo D. C., 1998, ApJ, 505, 50
Bertin E., Dennefeld M., 1997, A&A, 317, 43
Broadhurst T. J., Ellis R. S., Glazebrook K., 1992, Nat, 355, 55
Bruzual G., Charlot S., 1993, ApJ, 405, 538
Campos A., 1997, ApJ, 488, 606
Casali M. M., Hawarden T., 1992, JCMT-UKIRT Newsl. 3, 33
Cowie L. L., Gardner J. P., Hu E. M., Songaila A., Hodapp K. W., Wainscoat R. J., 1994, ApJ, 434, 114
Cowie L. L., Songaila A., Hu E. M., 1996, AJ, 112, 839
Djorgovski S. et al., 1995, ApJ, 438, L13
Driver S. P., Fernandez-Soto A., Couch W. J., Odewahn S. C., Windhorst R. A., Phillips S., Lanzetta K., Yahil A., 1998, ApJ, 496, L93
Driver S. P., Windhorst R. A., Griffiths R. E., 1995, ApJ, 453, 48
Freeman K. C., 1970, ApJ, 160, 811
Gardner J. P., 1995, ApJS, 98, 441
Gardner J., Cowie L., Wainscoat R., 1993, ApJ, 415, L9
Gardner J. P., Sharples R. M., Carrasco B. E., Frenk C. S., 1996, MNRAS, 282, L1
Glazebrook K., Ellis R., Santiago B., Griffiths R., 1995a, MNRAS, 275, L19
Glazebrook K., Peacock J. A., Miller L., Collins C. A., 1995b, MNRAS, 275, 169
Griffiths R. E. et al., 1994, ApJ, 437, 67
Huang J., Cowie L. L., Gardner J. P., Hu E. M., Songaila A., Wainscoat R. J., 1997, ApJ, 476, 12
Kauffmann G., Charlot S., 1998, MNRAS, 297, L23
Kleinmann S. G., 1992, in Filippenko A.V., ed., ASP Conf. Ser. Vol. 34, Robotic Telescopes in the 1990s. Astron Soc. Pac., San Francisco, p. 203
Kron R. G., 1980, ApJS, 43, 305
McLeod B. A., Bernstein G. M., Rieke M. J., Tollestrup E. V., Fazio G. G., 1995, ApJS, 96, 117
Maddox S. J., Sutherland W. J., Efstathiou G., Loveday J., Peterson B. A., 1990, MNRAS, 247, 1p
Metcalfe N., Shanks T., Fong R., Jones L. R., 1991, MNRAS, 249, 498 (Paper II)
Metcalfe N., Shanks T., Fong R., Roche N., 1995a, MNRAS, 273, 257 (Paper III)
Metcalfe N., Fong R., Shanks T., 1995b, MNRAS, 274, 769
Metcalfe N., Shanks T., Campos A., Fong R., Gardner J. P., 1996, Nat, 383, 263
Mobasher B., Ellis R. S., Sharples R. M., 1986, MNRAS, 223, 11
Moustakas L. A., Davis M., Graham J. R., Silk J., Peterson B. A., Yoshii Y., 1997, ApJ, 475, 445
Peebles P. J. E., 1984, ApJ, 284, 439
Pei Y., 1992, ApJ, 395, 130
Sandage A., Perelmuter J. M., 1990, ApJ, 361, 1
Saracco P., Iovino A., Garilli B., Maccagni D., Chincarini G., 1997, AJ, 114, 887
Scalo J. M., 1986, Fund. Cosmic Phys., 11, 1
Shanks T., 1989, The Galactic and Extragalactic Background Radiation, IAU
Shanks T., 1990, in Bowyer S., Leinert C., eds, Galactic and Extragalactic Background Radiations. Kluwer, Dordrecht, p. 269
Shanks T., Metcalfe N., Fong R., McCracken H. J., Campos A., Gardner J. P., 1998, in D'Odoricos., Fontana A., Giallengo E., eds, ASP Conf. Ser. Vol. 146, The Young Universe: Galaxy Formation and Evolution at Intermediate and High Redshift. Astron. Soc. Pac., San Francisco, p. 102
Williams R. E. et al., 1996, AJ, 112, 1335
Worthey G., 1994, ApJS, 95, 107
Zehavi I., Riess A. G., Kirshner R. P., Dekel A., 1998, ApJ, 503, 483
Zepf S. E., 1997, Nat, 390, 377

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