Mon. Not. R. Astron. Soc. **370,** 1677–1698 (2006) doi:10.1111/j.1365-2966.2006.10675.x

# Panchromatic properties of 99 000 galaxies detected by SDSS, and (some by) *ROSAT*, GALEX, 2MASS, *IRAS*, GB6, FIRST, NVSS and WENSS surveys

M. Obrić,[1,2★] Ž. Ivezić,[2] P. N. Best,[3] R. H. Lupton,[4] C. Tremonti,[6] J. Brinchmann,[5] M. A. Agüeros,[2] G. R. Knapp,[4] J. E. Gunn,[4] C. M. Rockosi,[7] D. Schlegel,[4] D. Finkbeiner,[4] M. Gaćeša,[10,13] V. Smolčić,[10,12] S. F. Anderson,[2] W. Voges,[8] M. Jurić,[4] R. J. Siverd,[4] W. Steinhardt,[4] A. S. Jagoda,[4] M. R. Blanton,[9] D. P. Schneider[11]

[1]*Kapteyn Astronomical Institute, University of Groningen, PO Box 800, Groningen 9700 AV, the Netherlands*
[2]*University of Washington, Department of Astronomy, Box 351580, Seattle, WA 98195-1580, USA*
[3]*Institute for Astronomy, Royal Observatory Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*
[4]*Princeton University Observatory, Peton Hall, Princeton, NJ 08544-1001, USA*
[5]*Max-Planck-Institute für Astrophysik, D-85748, Garching, Germany*
[6]*Hubble Fellow, University of Arizona, Steward Observatory, 933 N. Cherry Avenue, Tucson, AZ 85721, USA*
[7]*University of California–Santa Cruz, 1156 High Street, Santa Cruz, CA 95060, USA*
[8]*Max-Planck-Institute für Extraterrestrische Physik, Karl-Schwarzschild-Str. 1, Postfach 1317, D-85741 Garching, Germany*
[9]*The New York University, Physics Department, 4 Washington Place, New York, NY 10003, USA*
[10]*University of Zagreb, Physics Department, Bijenička cesta 32, 10000 Zagreb, Croatia*
[11]*Department of Astronomy and Astrophysics, 525 Davey Laboratory, University Park, PA 16802, USA*
[12]*Max-Planck-Institut für Astronomie, Königstuhl 17, Heidelberg, D-69117, Germany*
[13]*University of Connecticut, Physics Department, 2152 Hillside Road, Storrs, CT 06269-3046, USA*

Accepted 2006 May 26. Received 2006 May 19; in original form 2005 December 23

**ABSTRACT**

We discuss the panchromatic properties of 99 088 galaxies selected from the Sloan Digital Sky Survey (SDSS) Data Release 1 'main' spectroscopic sample (a flux-limited sample for 1360 deg$^2$). These galaxies are positionally matched to sources detected by *ROSAT*, Galaxy Evolution Explorer (GALEX), two-Micron All-Sky Survey (2MASS), *Infrared Astronomical Satellite* (*IRAS*), Green Bank GB6 survey (GB6), Faint Images of the Radio Sky at Twenty-centimetres (FIRST), NRAO VLA Sky Survey (NVSS) and Westerbork Northern Sky Survey (WENSS). The matching fraction varies from $<1$ per cent for *ROSAT* and GB6 to $\sim$40 per cent for GALEX and 2MASS. In addition to its size, the advantages of this sample are well-controlled selection effects, faint flux limits and the wealth of measured parameters, including accurate X-ray to radio photometry, angular sizes and optical spectra. We find strong correlations between the detection fraction at other wavelengths and optical properties such as flux, colours and emission-line strengths. For example, $\sim 2/3$ of SDSS 'main' galaxies classified as active galactic nucleus (AGN) using emission-line strengths are detected by 2MASS, while the corresponding fraction for star-forming galaxies (SFs) is only $\sim 1/10$. Similarly, over 90 per cent of galaxies detected by *IRAS* display strong emission lines in their optical spectra, compared to $\sim$50 per cent for the whole SDSS sample. Using GALEX, SDSS and 2MASS data, we construct the ultraviolet–infrared (UV–IR) broad-band spectral energy distributions for various types of galaxies, and find that they form a nearly one-parameter family. For example, the SDSS *u*- and *r*-band data, supplemented with redshift, can be used to 'predict' *K*-band magnitudes measured by 2MASS with an rms scatter of only 0.2 mag. When a dust content estimate determined from SDSS spectra with the aid of models is also utilized, this scatter decreases to 0.1 mag and can be fully accounted for by measurement uncertainties. We demonstrate that this interstellar dust content, inferred from optical SDSS spectra by

★E-mail: mobric@astro.washington.edu

Kauffmann et al., is indeed higher for galaxies detected by *IRAS* and that it can be used to 'predict' measured *IRAS* 60 μm flux density within a factor of 2 using only SDSS data. We also show that the position of a galaxy in the emission-line-based Baldwin–Phillips–Terlevich diagram is correlated with the optical light concentration index and $u - r$ colour determined from the SDSS broad-band imaging data, and discuss changes in the morphology of this diagram induced by requiring detections at other wavelengths. Notably, we find that SDSS 'main' galaxies detected by GALEX include a non-negligible fraction (10–30 per cent) of AGNs, and hence do not represent a clean sample of starburst galaxies. We study the IR–radio correlation and find evidence that its slope may be different for AGN and SFs and related to the H$\alpha$/H$\beta$ line-strength ratio.

**Key words:** surveys – galaxies: active – galaxies: fundamental parameters – galaxies: starburst – infrared: galaxies – radio continuum: galaxies.



## 1 INTRODUCTION

The study of global galaxy properties has been recently invigorated by modern sensitive large-area surveys across a wide wavelength range. The Sloan Digital Sky Survey (SDSS; York et al. 2000, for more details see Section A1) stands out because it has already provided near-ultraviolet (NUV) to near-infrared (NIR) five-colour imaging data and high-quality spectra ($R \sim 1800$) for over 500 000 galaxies. The 'main' spectroscopic galaxy sample is defined by a simple *r*-band flux limit (Strauss et al. 2002), and will include close to 1 000 000 galaxies.

A number of detailed galaxy studies based on SDSS data have already been published. Strateva et al. (2001) and Shimasaku et al. (2001) demonstrated a tight correlation between the $u - r$ colour, concentration of the galaxy's light profile, and morphology. Blanton et al. (2003) presented the SDSS galaxy luminosity function, and Kauffmann et al. (2003a,b) determined and analysed stellar masses and star formation histories for 100 000 SDSS galaxies.

In addition to 'stand-alone' studies based on only SDSS data, SDSS can be used as a cornerstone for panchromatic studies of galaxies aided by recent surveys at wavelengths outside the optical range (0.3–1 μm). The special role of SDSS in such studies is due to its rich optical information, in particular high-quality spectra and photometry. Nevertheless, galaxies emit a substantial fraction of their bolometric flux outside the wavelength range accessible to SDSS. For example, in starburst and Seyfert 2 galaxies the mid-/far-IR wavelength range is the most important contributor to the bolometric flux, e.g. Schmitt et al. (1997). Information obtained by other surveys offers important observational constraints for models of galaxy formation and evolution.

Numerous studies that utilize SDSS and surveys at other wavelengths have already been published. For example, Finlator et al. (2000) analysed the properties of point sources detected by SDSS and two-Micron All-Sky Survey (2MASS), and Ivezić et al. (2001a) discussed the colours and counts of SDSS sources detected by SDSS, 2MASS and Faint Images of the Radio Sky at Twenty-centimetres (FIRST) surveys. Ivezić et al. (2002) cross-correlated SDSS and the FIRST survey, and analysed the optical and radio properties of quasars and galaxies. Bell et al. (2003) used SDSS and 2MASS data to estimate the baryonic mass functions of galaxies, and Anderson et al. (2003) studied the properties of active galactic nucleus (AGN) galaxies detected by SDSS and *ROSAT*. Best et al. (2005a,b) studied radio galaxies, Chang et al. (2006) analysed the SDSS–2MASS colours of elliptical galaxies, and Goto (2005) and Pasquali, Kauffmann & Heckman (2005) studied the optical properties of SDSS galaxies detected by *Infrared Astronomical Satellite* (*IRAS*). A detailed analysis of rest-frame colours in the Strömgren system synthesized from SDSS spectra was presented by Smolčić et al. (2006). They found that the galaxy distribution in the resulting colour–colour diagrams forms a very narrow locus with a width of only 0.03 mag. This finding agrees well with the conclusion by Yip et al. (2004), based on a principal component analysis of SDSS spectra, that galaxy spectra can be described by a small number of eigenspectra.

Here, we cross-correlate the catalogue of galaxies from SDSS Data Release 1 (DR1; Abazajian et al. 2003) with catalogues of sources detected by *ROSAT* (X-ray), Galaxy Evolution Explorer (GALEX; UV), 2MASS (NIR), *IRAS* (mid-/far-IR), Green Bank GB6 survey (GB6; 6 cm), FIRST (20 cm), NRAO VLA Sky Survey (NVSS; 20 cm) and Westerbork Northern Sky Survey (WENSS; 92 cm). References and a description of each survey are listed in Appendix A. The panchromatic galaxy samples discussed here are ∼10–100 times larger than those used in older pre-SDSS studies. In addition, they are selected by simple flux limits, and benefit from a wealth of accurately measured parameters including X-ray to radio photometry, angular sizes, and optical spectra. The main aim of this paper is to quantify the fraction and basic properties of SDSS 'main' galaxies detected by other surveys using a uniform approach for all analysed surveys. However, due to the size and quality of the resulting samples, even a simple, preliminary analysis presented here is sufficient to yield a wealth of additional results.

We describe our matching and analysis methods in Section 2. In Section 3 we discuss the detection fraction of SDSS galaxies by other surveys, and in Section 4 we present a preliminary analysis of some panchromatic properties of galaxies in our sample. We discuss and summarize our results in Section 5.

## 2 MATCHING AND ANALYSIS METHODS

There are 99 825 unique galaxies in the SDSS DR1 'main' spectroscopic sample,[1] a sample limited by Petrosian magnitude, $r_{\rm Pet} < 17.77$ and covering 1360 deg$^2$ (for more detailed description see Stoughton et al. 2002; Strauss et al. 2002). We further restrict the sample by requiring the redshifts to lie in the range $0.01 \leqslant z \leqslant 0.30$

[1] The recent SDSS DR4 contains spectra for 565 715 galaxies, see www.sdss.org.

**Table 1.** Catalogues, their wavelength range, matching radius, total number of matches (for SDSS total number of galaxies), false match probability for the adopted matching radius, and, in the bottom table, matching fractions for all galaxies and for each galaxy class. False match probabilities are computed from the source density in the matched catalogues, and are consistent with the random matching rate when SDSS positions are offset by 1° in declination. The matching fractions in the bottom table are corrected for the difference in area covered by each catalogue and the area covered by SDSS DR1. Emission/no emission tags in the bottom table refer to whether or not the emission lines were detected in a galaxy (see text). The bottom row lists the surface density of each galaxy subsample in the SDSS DR1 catalogue.

| Catalogue | Wavelength | Matching distance (arcsec) | All (absolute #) | False match probability (per cent) | False associations (per cent) |
|---|---|---|---|---|---|
| SDSS | NUV–NIR | – | 99 088 | – | – |
| *ROSAT* RASS | X-ray | 30 | 267 | 0.056 | 8.9 |
| GALEX | UV | 6 | 866 | 0.796 | 1.9 |
| 2MASS XSC | NIR | 1.5 | 19 184 | 0.002 | <1 |
| *IRAS* FSC | Far-IR | 30 | 1736 | 0.013 | <1 |
| GB6 | 6 cm | 20 | 132 | 0.003 | 14 |
| FIRST | 20 cm | 3 | 3402 | 0.037 | 1.0 |
| NVSS | 20 cm | 15 | 3478 | 0.291 | 8.3 |
| WENSS | 92 cm | 20 | 363 | 0.227 | 9.1 |

| Catalogue | All (per cent) | No emission (per cent) | Emission (per cent) | AGN (per cent) | SF (per cent) | Unknown (per cent) |
|---|---|---|---|---|---|---|
| *ROSAT* RASS | 0.63 | 0.63 | 0.46 | 0.87 | 0.24 | 0.23 |
| GALEX | 42.0 | 24.1 | 67.7 | 27.9 | 82.1 | 93.3 |
| 2MASS XSC | 38.1 | 35.7 | 39.6 | 63.8 | 10.7 | 34.0 |
| *IRAS* FSC | 1.77 | 0.14 | 2.95 | 3.01 | 1.73 | 3.54 |
| GB6 | 0.22 | 0.15 | 0.09 | 0.21 | 0.02 | 0.02 |
| FIRST | 3.86 | 2.76 | 4.68 | 8.06 | 0.98 | 3.40 |
| NVSS | 3.51 | 2.85 | 3.52 | 5.02 | 1.66 | 3.25 |
| WENSS | 2.50 | 3.23 | 1.40 | 2.46 | 0.49 | 1.01 |
| SDSS [#/deg$^2$] | 72.86 | 39.38 | 33.48 | 11.81 | 7.54 | 13.44 |

and obtain the sample of 99 088 galaxies analysed here. For each galaxy, SDSS provides numerous properties measured from five-colour imaging data, such as astrometry, photometry and morphological information, as well as high-quality spectra. In addition to standard spectroscopic parameters automatically measured by the spectroscopic pipeline, we also utilize emission-line measurements described by Kauffmann et al. (2003a). We emphasize that the SDSS astrometry is very accurate (∼0.1 arcsec, Pier et al. 2003), which significantly simplifies the matching algorithm.

For each SDSS galaxy, we search for the two nearest neighbours in each of the eight catalogues. We accept the nearest neighbour as a true association if its distance is smaller than the catalogue-dependent matching radius listed in Table 1. The matching radius for each catalogue was determined by analysing the distribution of distances between the quoted position in the catalogue and the SDSS position, and corresponds to a $\sim 3\sigma$ cut-off.[2] Due to either high astrometric accuracy of other catalogues (e.g. 2MASS), or their low source surface density (e.g. *IRAS*), the matching contamination rate (fraction of false associations) is typically very low (< 1 per cent) at non-radio wavelengths and ∼10 per cent for the four radio surveys, as implied by both the source density in the matched catalogues, and the matching rate when SDSS positions are offset by 1° in declination. The fraction of cases where two sources in other catalogues are found within the matching radius is typically small (<1 per cent); in these cases we simply take the nearest neighbour to represent the true association. This fraction is sufficiently low that none of the conclusions presented in this paper change when both neighbours are excluded.

Before we proceed with the discussion of matching rate for each catalogue, we describe our analysis methods in the next two sections.

### 2.1 The global optical properties of galaxies in the SDSS DR1 main spectroscopic sample

The first step in analysing galaxies detected at other wavelengths is to compare their distribution in the optical parameter space to that for the whole SDSS sample. The parameters measured by the SDSS photometric pipeline *photo* (Lupton et al. 2002) are numerous (∼100), and we limit our preliminary analysis to the distribution of galaxies in optical colour–magnitude–redshift space.

#### *2.1.1 Colour–magnitude–redshift distributions*

SDSS galaxies are not randomly distributed in the space spanned by apparent (or absolute) magnitude, colour and redshift. As shown by Strateva et al. (2001), and discernible in Fig. 1, galaxies show a bimodal $u - r$ colour distribution (hereafter, optical SDSS colours are constructed using so-called 'model' magnitudes; for details see Stoughton et al. 2002). Galaxies with $u - r < 2.22$ tend to be spiral galaxies, and those with $u - r > 2.22$ elliptical galaxies (see also Shimasaku et al. 2001; Baldry et al. 2004). Because the spiral/blue galaxies tend to have lower luminosities than elliptical/red galaxies (bottom right-hand panel), the former are typically found at lower redshifts in the flux-limited SDSS sample than are the latter. Blanton et al. (2003) give a detailed analysis of the dependence of luminosity function on galaxy type. Note that the 'features'

[2] The distance distribution for the SDSS–NVSS sample is better fit by a sum of two Gaussians. However, this behaviour has no significant consequence for the matching completeness and contamination.

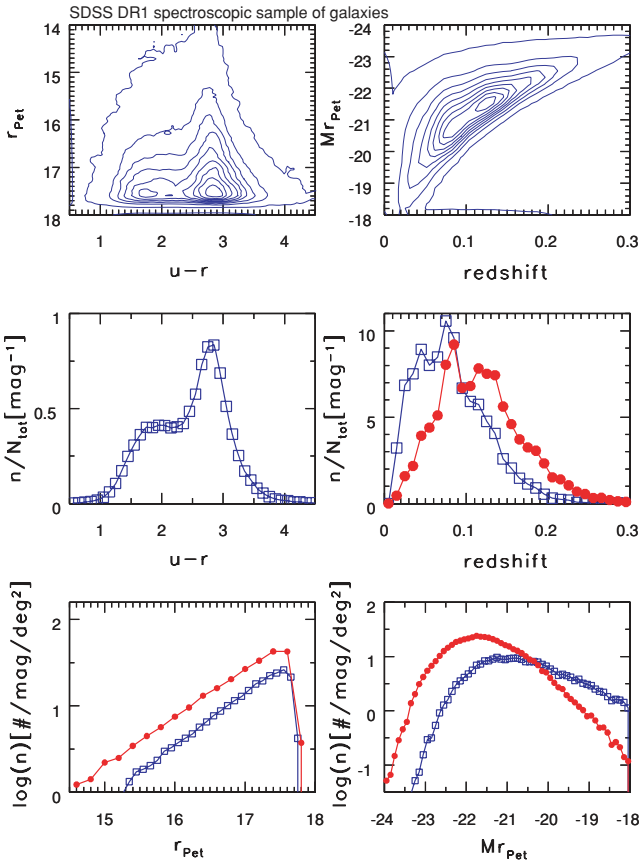


**Figure 1.** Colour–magnitude–redshift distributions for SDSS DR1 'main' galaxies ($r_{\rm Pet} < 17.77$). The top two panels show the distribution of galaxies using linearly spaced contours, in steps of 10 per cent. The middle left-hand panel shows the $u - r$ distribution of galaxies. Galaxies with $u - r < 2.22$ tend to be spiral/blue galaxies, and those with $u - r > 2.22$ elliptical/red galaxies (Strateva et al. 2001). Using this separation, the middle right-hand panel shows the redshift (probability) distributions for blue (open squares) and red (dots) subsets. The same symbols are used to display their differential apparent magnitude (bottom left-hand panel) and absolute magnitude (bottom right-hand panel) distributions.

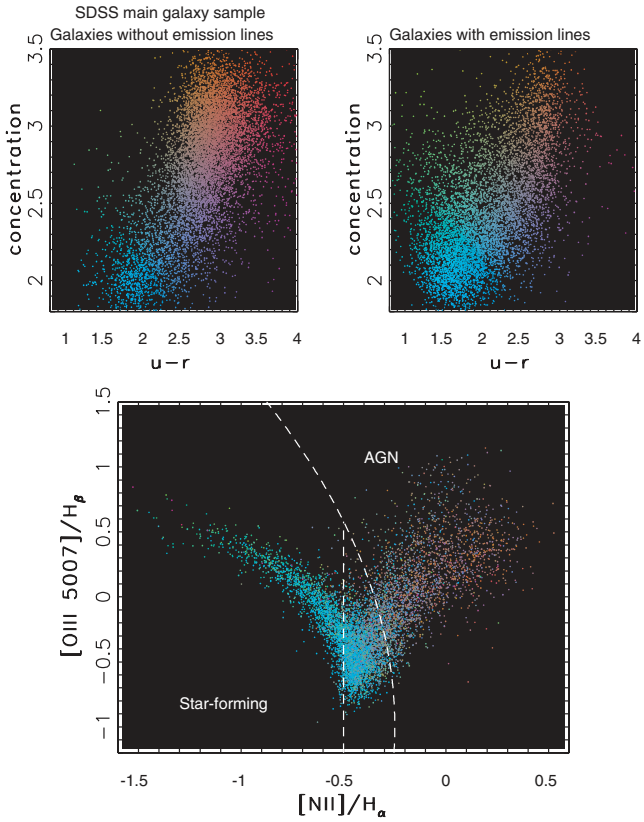


**Figure 2.** The top two panels compare the distribution of galaxies without (left-hand side) and with (right-hand side) emission lines in the concentration index versus $u - r$ diagram. The dots are two-dimensionally colour coded according to their concentration index and $u - r$ colour. The same colour-coding scheme is used in the bottom panel, which shows the BPT diagram for emission-line galaxies (note that the line flux ratios are expressed on a logarithmic scale). Emission-line galaxies can be separated into three groups according to their position in the BPT diagram: AGNs, star-forming and 'unknown', using the separation boundaries outlined by the dashed lines. Note the strong correlation between position in the BPT diagram and $u - r$.

in the middle right-hand panel of Fig. 1 are due to the large-scale structure of galaxies. The differential number counts of both colour types are well described by $\log(n) = C + 0.6\, r_{\rm Pet}$ (see also Yasuda et al. 2001). We will use diagrams such as that shown in Fig. 1 to compare the distributions of galaxies detected at other wavelengths to the distribution of all SDSS galaxies.

### *2.1.2 The distribution of emission-line galaxies in the Baldwin–Phillips–Terlevich diagram*

In addition to the overall comparison of galaxy distributions in diagrams shown in Fig. 1, we analyse the behaviour of three subsamples defined by their emission-line properties: galaxies without emission lines, and emission-line galaxies separated into star-forming and AGNs. To classify a galaxy as an emission-line galaxy, we follow Kauffmann et al. (2003a) and require a $3\sigma$ significant detection of the H$\alpha$, H$\beta$, [N II 6583] and [O III 5007] lines. To classify emission-line galaxies as star-forming or AGN, we use the standard BPT diagram (Baldwin, Phillips & Terlevich 1981).

The top two panels in Fig. 2 compare the distribution of galaxies without and with emission lines in the concentration index versus $u - r$ diagram. Galaxies without emission lines tend to have larger concentration index and redder $u - r$ than galaxies with emission lines. The dots in these two panels are two-dimensionally colour coded according to their concentration index and $u - r$ colour. The same colour-coding scheme is used in the bottom panel, which shows the BPT diagram for emission-line galaxies. *There is a strong correspondence between the position of a galaxy in the BPT diagram and its position in the concentration index versus $u - r$ diagram.* Galaxies in the 'star-forming branch' with small [N II]/H$\alpha$ ratio, for a given [O III 5007]/H$\beta$ ratio, have predominantly blue $u - r$ colours and small concentration index, while AGNs have redder $u - r$ colours and large concentration index. Furthermore, the distribution of *emission-line* galaxies in the BPT diagram is also correlated with $u - r$ and concentration index.

In the subsequent analysis, we separate emission-line galaxies into three groups according to their position in the BPT diagram: AGNs, star-forming and 'unknown'. The adopted separation boundaries are shown by the dashed lines, and are designed to produce robust clean samples of AGN and star-forming galaxies (SFs) (for alternative approaches see Hao et al. 2005, and references therein; for the aperture effects due to 3 arcsec fibre diameter see e.g. Kauffmann et al. 2003a; Kewley, Jansen & Geller 2005). The

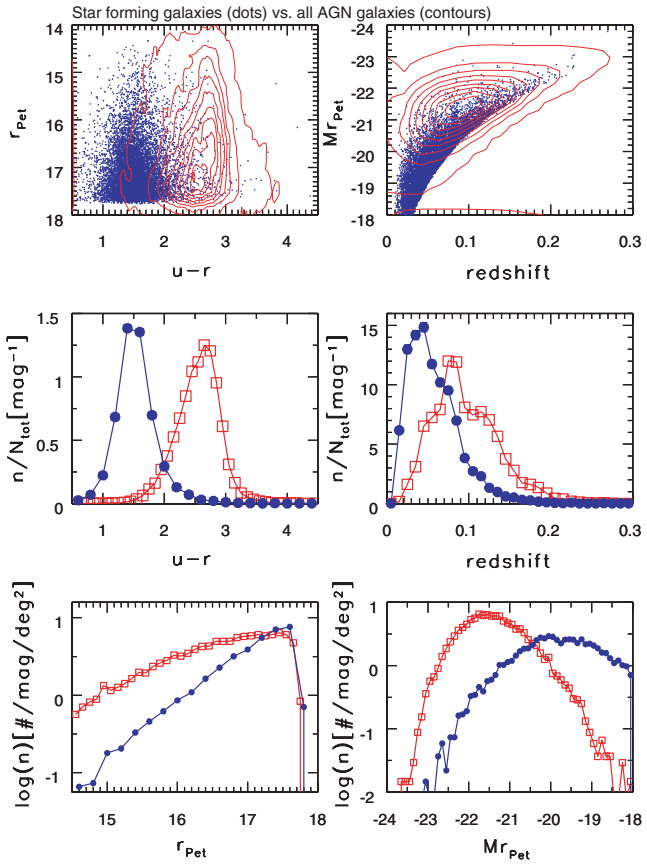


**Figure 3.** The comparison of distributions of AGN (contours, squares) and star-forming (dots, filled circles) emission-line galaxies in magnitude–colour–redshift space. Note remarkable differences in their broad-band imaging properties (especially $u - r$) showing the correlation of spectroscopic and photometric properties.

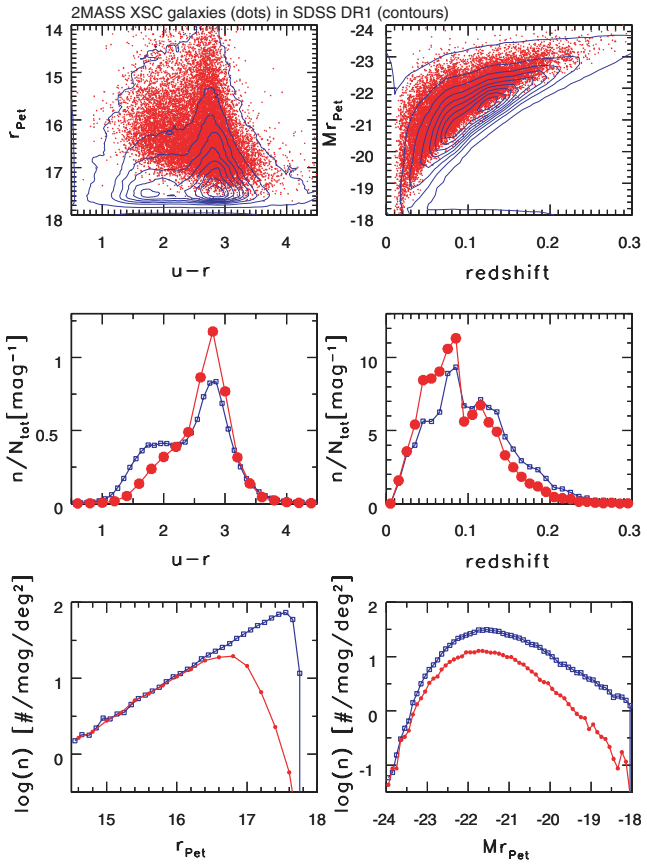


**Figure 4.** The comparison of colour–magnitude–redshift distributions for all SDSS DR1 'main' galaxies (contours, and histograms marked with small squares) and those listed in 2MASS XSC (dots, and histograms marked with large dots). 2MASS XSC galaxies are biased towards red galaxies, larger luminosities and smaller redshifts. The 2MASS XSC catalogue is essentially complete for galaxies with $r_{\rm Pet} < 16.3$ (see bottom left-hand panel).

'unknown' category is found at the join of the two branches, and it is not obvious from the displayed data to which class these galaxies belong. While their concentration index and $u - r$ colour indicate that they may be SFs, their IR colours and redshift distribution (see Section 4) suggest that they are more similar to AGN galaxies. Of course, it is possible that these objects are SFs that host an AGN.

The comparison of the distributions of AGN and SFs in magnitude–colour–redshift space is shown in Fig. 3. As discussed above, the two types of galaxies, classified using emission lines, have remarkably different $u - r$ distributions (see middle left-hand panel). Furthermore, SFs tend to have smaller luminosities than AGNs, and hence are observed at lower redshifts in the flux-limited SDSS sample. They also have very different differential number counts – the counts of AGN galaxies are flatter [$\mathrm{d}(\log N)/\mathrm{d}r \sim 0.3$] than those of SFs (and those of the whole SDSS sample). For detailed studies of the optical properties of star-forming and AGN galaxies, see e.g. Brinchmann et al. (2004), Tremonti et al. (2004) and Heckman et al. (2004), and references therein.

## 3 WHAT TYPES OF SDSS GALAXIES ARE DETECTED AT OTHER WAVELENGTHS?

The detection fraction of SDSS galaxies at other wavelengths is a strong function of optical properties such as flux, $u - r$ and emission-line strengths. In this analysis, we have taken into account the most important selection effects, namely: observational biases caused by varying survey depths, astrophysical effects such as intrinsically different colour distributions for different galaxy types, and $K$-correction (Gunn & Oke 1975) coupled with bias in redshift. The size and quality of our sample, in addition to its well-controlled selection criteria, allow us to separate observational and astrophysical effects, and to study intrinsic correlations amongst numerous measured galaxy properties.

The matching fraction varies from <1 per cent for *ROSAT* and GB6 to ∼40 per cent for GALEX and 2MASS (Table 1). We start the discussion with the 2MASS, continue toward longer wavelengths, and then proceed from optical toward shorter wavelengths.

### 3.1 2MASS survey

The distribution of SDSS galaxies detected by 2MASS [3] in the colour–magnitude–redshift space is compared to the distribution of all SDSS galaxies in Fig. 4. The requirement that a galaxy is detected and resolved by 2MASS (i.e. the XSC sample, see Section A4) introduces a bias towards red galaxies and lower redshift as shown in the middle two panels. The SDSS–2MASS XSC catalogue is

[3] We analyse only resolved 2MASS sources listed in the 2MASS XSC catalogue; see Section A4 for more details.

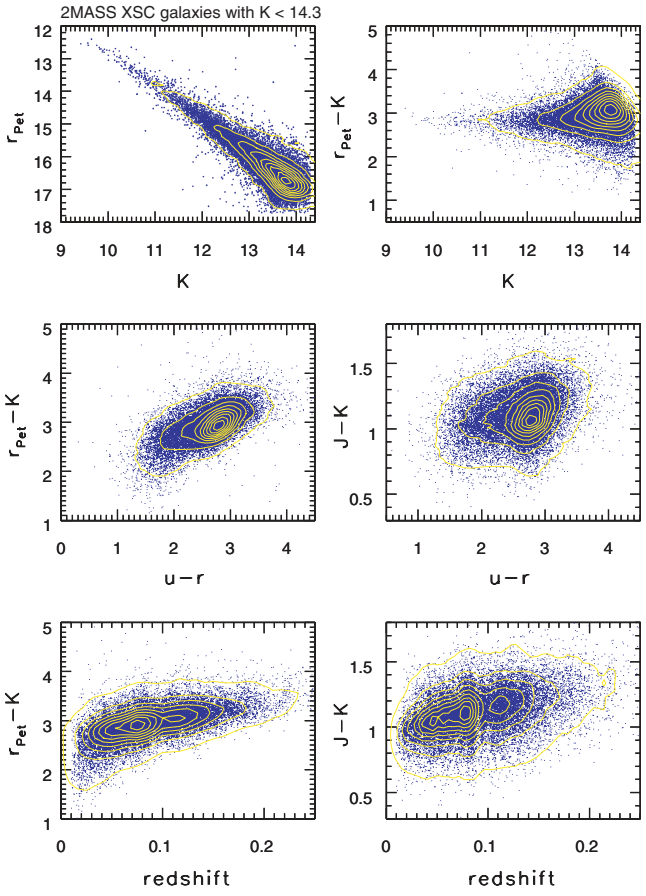


**Figure 5.** The optical and IR colours of SDSS galaxies listed in 2MASS XSC.

essentially complete[4] for galaxies brighter than $r_{\rm Pet} \sim 16.3$ (bottom left-hand panel, see also McIntosh et al. 2006).

The colour-dependent incompleteness of the 2MASS XSC catalogue for galaxies with $r_{\rm Pet} \gtrsim 16.3$ is due to the 2MASS faint limit, coupled with the optical/IR colour distribution of galaxies. We adopted $K = 14.3$, which corresponds to a $\sim 10\sigma$ detection; the $K$-band differential counts of 2MASS XSC galaxies indicate that the catalogue is complete to about $K = 13.5$. Fig. 5 shows the optical/IR magnitude and colour distributions of SDSS–2MASS galaxies. We use 2MASS 'default' magnitudes (see Jarrett et al. 2000) and do not correct for the differences between AB (SDSS) and Vega (2MASS) magnitudes.[5] These differences are $m_0 = m_{AB} - m_{\rm Vega}$, where $m_0(J) = 0.89$, $m_0(H) = 1.37$ and $m_0(K) = 1.84$ (see Finlator et al. 2000). Most galaxies have $r_{\rm Pet} - K$ in the range 2–3.5. Thus, the bluest galaxies are brighter than the 2MASS faint limit only if they have $r_{\rm Pet} < 16.3$. For $r_{\rm Pet} > 16.3$ only galaxies with $r_{\rm Pet} - K$ redder than $r_{\rm Pet} - 14.3$ are sufficiently bright in the $K$ band, and at $r_{\rm Pet} = 17.8$ practically no galaxies are listed in 2MASS XSC. This bias explains why the fraction of red galaxies in SDSS–2MASS sample is higher than among all SDSS galaxies (80 versus 66 per cent, with the blue/red separation defined by $u - r = 2.22$). Since red galaxies tend to be more luminous than blue galaxies (e.g. Blanton et al. 2003), this colour bias also explains why 2MASS–SDSS galaxies are biased towards larger luminosities.

As the bottom two panels in Fig. 5 demonstrate $r_{\rm Pet} - K$ and $J - K$ depend on redshift. This correlation ($K$-correction), coupled to the colour effects discussed above, introduces a dependence of the detection fraction on redshift. It is also an important effect to consider when comparing the colours of various subsamples that may have different redshift distributions, as we further discuss in Section 4.2.

### 3.1.1 *Predicting 2MASS K-band flux from UV/optical SDSS fluxes*

The optical–IR colour, $r_{\rm Pet} - K$, is correlated with the UV–optical $u - r$ colour, as shown in the middle left-hand panel in Fig. 5. This correlation indicates that it is possible to estimate the $K$-band flux using only SDSS data, and is consistent with the fact that galaxies form a nearly one-dimensional sequence in various optical colour–colour diagrams constructed with SDSS data. The correlation among colours is especially tight for optical rest-frame colours, with a scatter of only $\sim$0.03 mag perpendicular to the locus (Smolčić et al. 2006). The $r_{\rm Pet} - K$ versus $u - r$ correlation demonstrates that this one-dimensionality of broad-band galaxy spectral energy distributions (SEDs) extends to NIR wavelengths.

In order to quantitatively assess to what extent NIR flux is correlated with optical fluxes, we determine the $K$-band flux from $K_{\rm SDSS} = r_{\rm Pet} - (r - K)^*$, where $r_{\rm Pet}$ is the SDSS $r$-band Petrosian magnitude and $(r - K)^*$ is a best fit to the observed $r_{\rm Pet} - K$ colours for SDSS/2MASS galaxies sampled from SDSS. We use UV/visual fluxes ($u$ and $r$ bands) to fit the $r_{\rm Pet} - K$ colour because this is the 'hardest' wavelength combination with most astrophysical implications. According to 'common wisdom', such a relationship should not be very accurate due to the effects of starbursts and dust extinction. Predicting, for example, 2MASS $J$-band flux from SDSS $z$-band flux is trivial because these two bands are adjacent in wavelength space. We also take into account the $K$-correction (see the bottom left-hand panel in Fig. 5), and fit the following functional form

$$(r - K)^* = A + B\,(u - r) + C\,(u - r)^2 + D\,(u - r)^3 + E z_r + F z_r^2, \quad (1)$$

where $z_r$ is redshift.

The motivation for this functional form is the behaviour of the $r - K$ colour shown in the top panel in Fig. 6. The $r - K$ colour also seems to be a well-defined function of the position in the $g - r$ versus $u - g$ colour–colour diagram, shown in the bottom panel in Fig. 6, but the scatter around the median values in each bin is larger than for the $u - r$ versus redshift diagram (this is essentially due to larger uncertainty of photometric redshifts that are implied by the position of a galaxy in the $g - r$ versus $u - g$ diagram, in comparison to spectroscopic redshifts).

With the best-fitting values (A, B, C, D, E, F) = (1.115, 0.940, −0.165, 0.008 51, 4.92, −9.10), this relation predicts 2MASS $K$-band magnitudes with an rms scatter of only 0.20 mag. The residuals between the fitted and measured values depend on neither colour nor redshift, and are nearly Gaussian (see the two middle panels in Fig. 7). However, there is a correlation between the residuals and the optical galaxy size, parametrized by $R_{50}^z$, the radius enclosing 50 per cent of the Petrosian flux in the $z$ band (for details see Stoughton et al. 2002; Strauss et al. 2002). To correct for these aperture and resolution effects, which presumably depend on galaxy

[4] Here 'essentially complete' implies a completeness of ∼99 per cent, as demonstrated by the direct comparison of the full SDSS and 2MASS overlap (see Finlator et al. 2000; Ivezić et al. 2001b).

[5] For completeness, the AB-to-Vega offsets for SDSS bands are $m_0(u) = 0.94$, $m_0(g) = -0.08$, $m_0(r) = 0.17$, $m_0(i) = 0.40$ and $m_0(z) = 0.57$ (see http://www.sdss.org/ for details on the transformations).

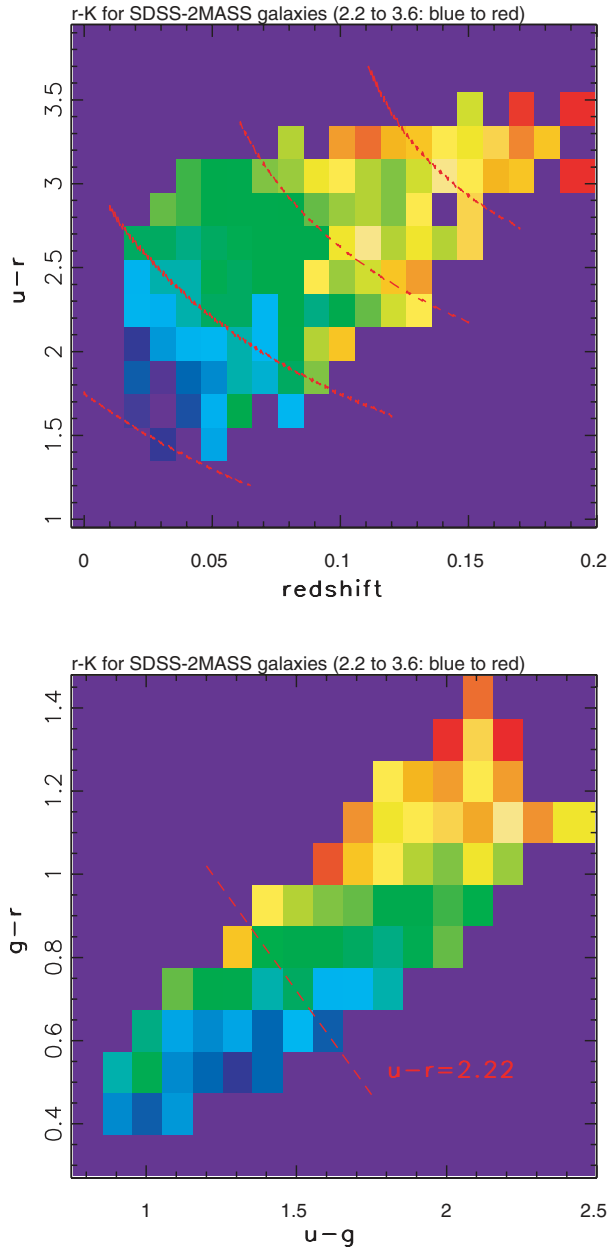


**Figure 6.** The top panel shows the median $r - K$ colour in small $u - r$ versus redshift bins. The median colour is colour-coded blue to red in the 2.2–3.6 range with a linear stretch. Bins with fewer than three galaxies are shown with magenta colour. The distribution of the rms of $r - K$ colour per bin has a median of 0.2 mag and a width of 0.08 mag. The dashed lines represent contours of constant $r - K$ (2.3, 2.7, 3.0 and 3.2) given by equation (1). The bottom panel shows the median $r - K$ colour in small $g - r$ versus $u - g$ colour bins, with analogous colour coding. The distribution of the rms of $r - K$ colour per bin has a median of 0.23 mag and a width of 0.10 mag. The dashed line shows the boundary between blue and red galaxies ($u - r = 2.22$) proposed by Strateva et al. (2001).

profile,[6] or nearly equivalently on galaxy colour (Strateva et al. 2001), we add to the right-hand side of equation (1) $\Delta(r - K)^* = (0.496 - 0.154\,R^z_{50})$ for galaxies with $u - r < 2.22$ and $\Delta(r - K)^* = (0.107 - 0.045R^z_{50})$ for redder galaxies. This correction has

[6] This assumption was recently verified by Chang et al. (2006).

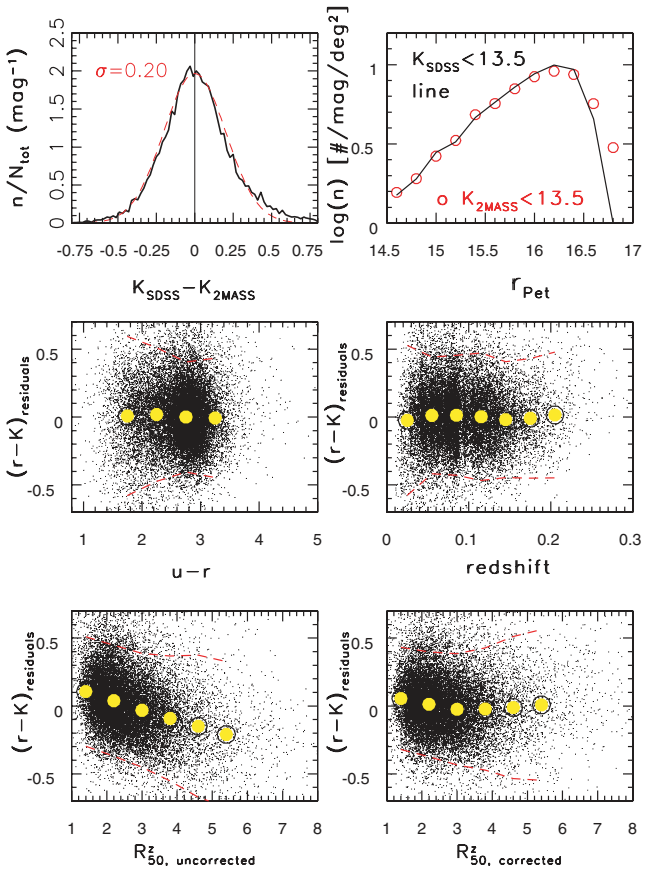


**Figure 7.** The solid line in the top left-hand panel shows the distribution of differences between SDSS-predicted and 2MASS-measured $K$-band flux, using equation (1). The dashed line is a Gaussian with $\sigma = 0.20$ mag. The right-hand panel shows the $r$-band counts of SDSS-predicted (line) and 2MASS-measured (symbols) galaxies with $K < 13.5$. The two middle panels show the dependence of residuals between predicted and measured $K$-band magnitudes on the $u - r$ colour (left-hand side) and redshift (right-hand side). The bottom panels show the dependence of the same residuals on $R^z_{50}$, the radius enclosing 50 per cent of the Petrosian flux in the $z$ band, before (left-hand side) and after (right-hand side) correcting for this effect. In the lower four panels, large yellow circles represent medians and $2\sigma$ envelope is given by red dashed curves.

a negligible effect on the rms scatter in the predicted $K$ magnitude, and only removes a correlation of $K_{\rm SDSS} - K_{\rm 2MASS}$ residuals with galaxy size (see the two bottom panels in Fig. 7).

The distribution of differences between the predicted and measured $K$-band magnitudes is shown in Fig. 7 (top left-hand panel). The median residuals, as a function of $u - r$ and $z_r$, do not exceed 0.03 mag, and the rms scatter decreases to 0.15 mag at the bright end ($K < 12$). The top right-hand panel in Fig. 7 compares the differential number counts as a function of $r_{\rm Pet}$ for galaxies with $K < 13.5$, where the latter condition is imposed using measured and predicted values. The good agreement shows that predicted $K$-band flux is *not overestimated* for galaxies that are not in 2MASS XSC, and indicates that the proposed relations may be applicable for galaxies fainter than the 2MASS faint cut-off.

Given typical measurement errors in $u$, $r$, $R^z_{50}$ and $K$, we conservatively conclude that the true astrophysical scatter of $K$-band magnitudes predicted from the blue part of the SED is not larger than ∼0.1 mag. Similarly, the relation $(J - K) = 2.172\,z_r + 0.966$, where $z_r$ is redshift (see the bottom right-hand panel in Fig. 5),

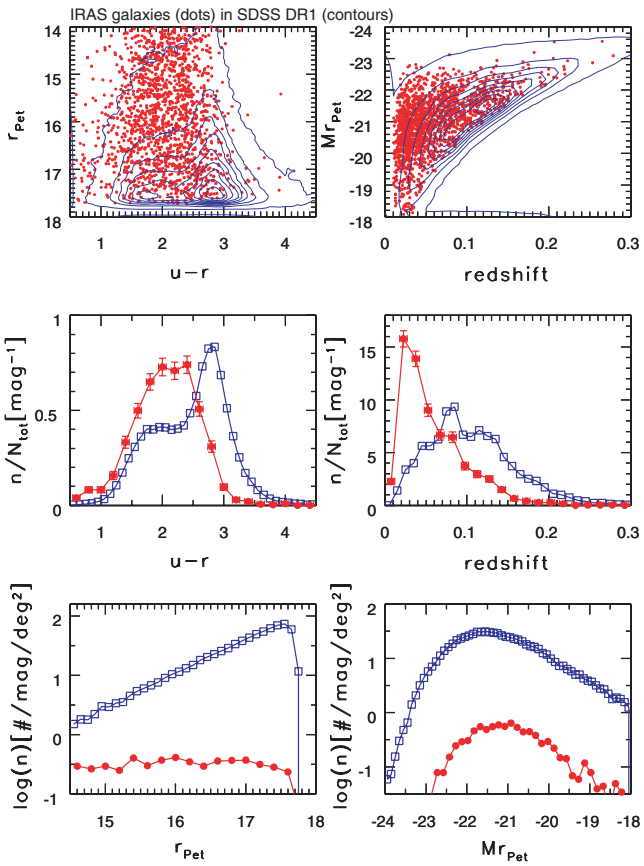


**Figure 8.** Analogous to Fig. 4, except that here galaxies listed in the *IRAS* FSC (dots, filled circles) are compared to the whole SDSS sample (contours, squares). Galaxies detected by *IRAS* are strongly biased towards blue galaxies and lower redshifts. The fraction of galaxies detected by *IRAS* decreases from 22 per cent for $r_{\rm Pet} = 14.5$ to 1 per cent for $r_{\rm Pet} = 17.5$.

predicts the $J - K$ measured by 2MASS with an rms scatter of 0.11 mag (0.07 mag at the bright end), and no significant residuals with respect to $K$, $u - r$ and redshift; that is, the rest-frame $J - K$ colour distribution of *all* low-redshift galaxies is very narrow: ∼0.1 mag. These tight correlations demonstrate the remarkable one-dimensionality of galaxy SEDs from UV to IR wavelengths. We further discuss the SED of galaxies in Section 4.1, and an improvement to the $K$-band flux prediction given by equation (1) in Section 4.4.

### 3.2 *IRAS* FSC

The distribution of SDSS galaxies detected by *IRAS*[7] in colour–magnitude–redshift space is compared to the distribution of all SDSS galaxies in Fig. 8. The requirement that a galaxy is detected by *IRAS* introduces a strong bias towards optically blue galaxies and lower redshift (the middle two panels). The majority of these galaxies have emission lines and include both star-forming and AGN galaxies, as we discuss in more detail in Section 4.2. The completeness of the *IRAS* FSC catalogue depends strongly on $r_{\rm Pet}$, and varies from 22 per cent for $r_{\rm Pet} \sim 14.5$ to 1 per cent for $r_{\rm Pet} \sim 17.5$.

[7] We analyse the sources listed in the *IRAS* FSC catalogue; see Section A5 for more details.

#### 3.2.1 *The correlation between $u - m_{60}$ and galaxy dust content*

Interstellar dust absorbs UV and optical radiation and re-emits it at mid- and far-IR wavelengths. Hence, some degree of correlation is expected between the far-IR–optical/UV colours and the amount of dust in a galaxy.

Kauffmann et al. (2003a) used the distribution of galaxies in the plane spanned by the strength of the H$\delta$ line and the 4000 Å break ($D_{4000}$) to obtain model-dependent estimates of stellar masses and dust content for SDSS galaxies. Given the position of a galaxy in the H $\delta$ –$D_{4000}$ plane, the most probable mass-to-light ratio is drawn from a model library. With the measured luminosity, this ratio then yields stellar mass. The observed luminosity is corrected for the dust extinction determined by comparing observed imaging $g - r$ and $r - i$ colours to model-predicted colours (the latter do not include the effects of dust reddening). The reddening correction needed to make models agree with data is interpreted as an effective optical depth in the SDSS $z$ band, $A_z$, due to a galaxy's interstellar dust. Here we find, using the measured properties of galaxies detected by *IRAS*, *independent* support for the notion that these model-dependent estimates of $A_z$ are indeed related to the galaxy dust content.

The top panel in Fig. 9 compares the distributions of $A_z$, determined by Kauffmann et al. (2003a), for all SDSS galaxies (pluses) and for the subset detected by *IRAS* (squares with error bars). Galaxies detected by *IRAS* have systematically higher values of $A_z$ than the full SDSS 'main' galaxy sample. If values of $A_z$, determined using *only* SDSS data, were not related to the dust content, there would be no systematic difference induced by requiring a detection by the fully *independent* *IRAS* survey.

Furthermore, we find a correlation between $u - m_{60}$ colour ($m_\lambda$ are *IRAS* measurements expressed as AB magnitudes, with $\lambda =$ 12, 25, 60 and 100 μm) and $A_z$. The small symbols in the middle panel in Fig. 9 show $u - m_{60}$ colour as a function of $A_z$ for 1200 highly probable SDSS–*IRAS* identifications, selected from the full SDSS–*IRAS* sample by limiting the maximum SDSS–*IRAS* distance to 20 arcsec. We first determine median values of $u - m_{60}$ in $A_z$ bins of range 0.4–2.0, and then fit a linear relation to obtain

$$u - m_{60} = (6.0 \pm 0.2) + (1.64 \pm 0.2)\, A_z. \qquad (2)$$

We do not find significant differences in the best-fitting relations fitted separately to AGN and star-forming subsamples (classified using emission-line strengths). The adopted $A_z$ range excludes ∼3 per cent of the sample that has very small $A_z$ and $u - m_{60}$ about 2.5 mag redder than predicted by the above relation (triangles in the upper left-hand corner in the middle panel in Fig. 8). It is not clear whether this 60 μm excess is physical, or due to random matches. In any case, the fraction of the excluded sources is sufficiently small to have no effect on the overall correlation.

Using this relation and the measured SDSS $u$-band fluxes, we estimate the 60 μm flux, and compare it to the measured values in the bottom panel in Fig. 9. The *IRAS* 60 μm flux can be predicted within a factor of ∼2 (rms, or 0.8 mag) using *only* SDSS data. If, instead, the $A_z$ estimates are ignored, and the 60 μm flux is estimated by assuming $u - m_{60} = 7.6$ (the median value) for all galaxies, the rms scatter between the predicted and measured values becomes 1.41 mag. Hence, the $A_z$ estimates do contain information about the dust content.

We have also attempted to use the H$\alpha$/H$\beta$ line strength ratio as a proxy for effective dust extinction (e.g. see Moustakas, Kennicutt & Tremonti 2006, and references therein). The values of H$\alpha$/H$\beta$ and $A_z$ determined by Kauffmann et al. (2003a) are well correlated. We find that H$\alpha$/H$\beta$ for emission-line galaxies can be determined from

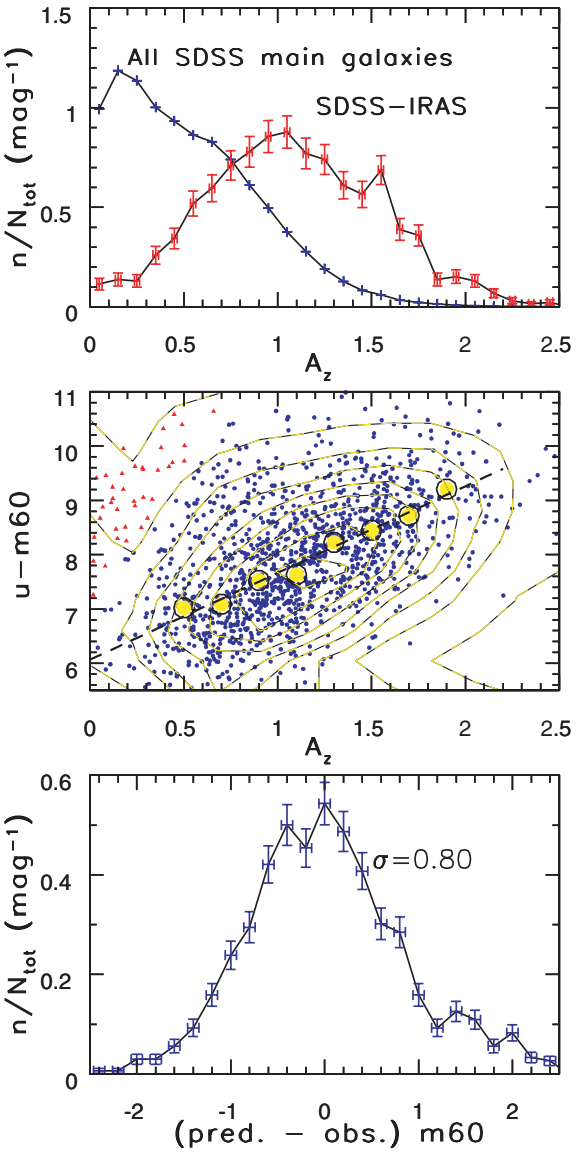


**Figure 9.** The top panel compares the distributions of the $z$-band dust extinction, $A_z$, inferred from SDSS spectra by Kauffmann et al. (2003a), for all SDSS galaxies (pluses) and for the subset detected by *IRAS* (squares with error bars). The middle panel shows the UV-far-IR colour, $u - m_{60}$, as a function of $A_z$ (both symbols and contours). The large symbols mark median values of $u - m_{60}$ in $A_z$ bins in the 0.4–2.0 range. The dashed line is the best linear fit to these medians. The triangles, concentrated in the upper left corner, mark the ∼3 per cent of galaxies that apparently do not follow this relation. The bottom panel shows the distribution of differences between the 60 μm flux measured by *IRAS*, and predicted values using the best-fitting $u - m_{60}$ versus $A_z$ relation, and the $u$-band fluxes measured by SDSS. The rms of the distribution is marked in the panel ($\sigma$, in magnitudes).

$A_z$ with an rms scatter of 0.07 using the relationship

$$\frac{\mathrm{H}\alpha}{\mathrm{H}\beta} = 0.49 + 0.143 * A_z. \quad (3)$$

This relationship maintains its accuracy when only radio-selected, IR-selected and subsamples separated into AGN and SFs are considered. Thus, the $u - m_{60}$ versus $A_z$ and H$\alpha$/H$\beta$ versus $A_z$ correlations imply the existence of a $u - m_{60}$ versus H$\alpha$/H$\beta$ correlation.

As expected, we find an overall correlation between $u - m_{60}$ and H$\alpha$/H$\beta$ ratio for SDSS–*IRAS* galaxies. However, it is not as strong as the $u - m_{60}$ versus $A_z$ relation discussed above. In terms of the rms scatter between predicted and measured 60 μm flux, it predicts 60 μm magnitude within 1.2 mag, that is, not as well as when using $A_z$. This implies that $A_z$ values determined using SDSS spectra and sophisticated stellar population models may be a better estimator of effective dust content than the straightforward application of the H$\alpha$/H$\beta$ line-strength ratio.

Although it is hard to estimate the errors in *IRAS* flux measurements without an independent data set, the radio–IR correlation discussed in Section 4.2.3 suggests that they are not larger than ∼0.4 mag, and therefore smaller than the rms scatter of 0.82 mag between predicted and measured 60 μm fluxes (the quoted formal *IRAS* FSC photometric errors are ∼0.2 mag). Thus, it may be possible to further improve the prediction for far-IR flux by using additional SDSS measurements such as sizes and UV/optical colours. For example, the differences between predicted and measured 60 μm fluxes are somewhat correlated with $u - r$ colour: the median value is −0.1 mag for galaxies with $u - r < 2.22$ and 0.2 mag for redder galaxies. A similar effect is seen when the sample is separated into AGN and SFs. We postpone such an analysis until the larger samples needed for robust quantitative multidimensional analyses are constructed.

### 3.3 Radio surveys (GB6, FIRST, NVSS, WENSS)

The advent of modern sensitive large-area radio surveys (see Section A6 for brief descriptions and references), combined with an optical survey such as SDSS, offers significantly larger, more diverse and accurate samples of radio sources with optical identifications than available until recently. Detailed studies of SDSS sources detected by the FIRST and NVSS 20 cm surveys was presented by Ivezić et al. (2002) and Best et al. (2005a,b). Here, we extend their analysis to multiwavelength radio observations by including data from the GB6 (6 cm) and WENSS (92 cm).

A summary of optical properties of SDSS galaxies detected by GB6 and WENSS is shown in Figs 10 and 11. Analogous diagrams for galaxies detected by FIRST and NVSS can be found in Ivezić et al. (2002). The matching rate is the smallest for the GB6 catalogue (0.22 per cent, see Table 1), and the highest for FIRST (3.86 per cent). The difference in matching fractions for the two 20 cm surveys (FIRST and NVSS) is due to their different faint flux limits and angular resolution. All four radio catalogues show similar distributions in colour–magnitude–redshift space, despite the relatively large wavelength coverage and varying angular resolution. Radio galaxies are biased towards red, luminous galaxies and higher redshifts. Even when red galaxies ($u - r > 2.22$) are considered separately, their median $u - r$ colour is redder by about 0.3 mag for the radio-detected subsample than for the whole red sample. However, this is simply a consequence of a bias in redshift induced by requiring radio-detection coupled with the $K$-correction, as discussed by Ivezić et al. (2002). When compared in a small redshift range, *the radio-detected red galaxies have the same $u - r$ colour distribution as red galaxies without radio detections*.

This sample is sufficiently large to test whether the radio spectral slope is correlated with optical properties, such as $u - r$. Using NVSS and WENSS measurements, we compute the radio spectral slope between 20 and 92 cm, and find no correlation with the $u - r$ colour (see the top right-hand panel in Fig. 12). We find that the distribution of this spectral slope for 'main' SDSS galaxies is different from the distribution for the full multiwavelength radio

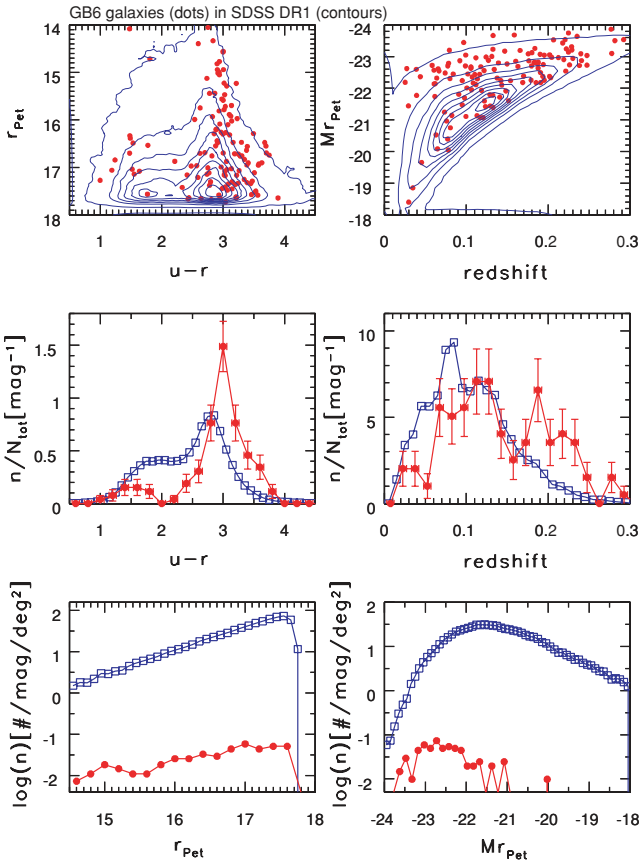


**Figure 10.** Analogous to Fig. 4, except that here galaxies detected by GB6 (dots, filled circles) are compared to the whole SDSS sample (contours, squares). Galaxies detected by GB6 are biased towards luminous red galaxies. The fraction of galaxies detected by GB6 is 0.22 per cent.

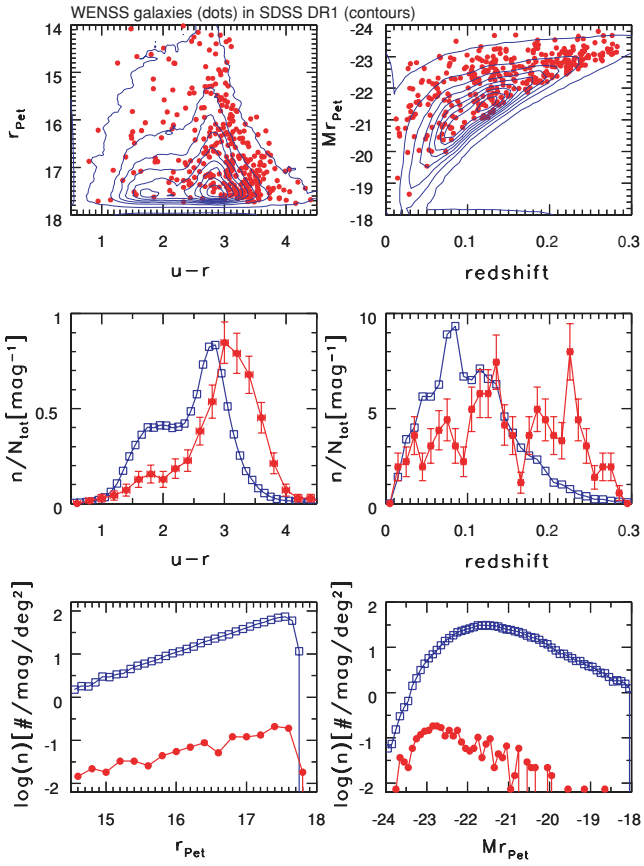


**Figure 11.** Analogous to Fig. 4, except that here galaxies detected by WENSS (dots, filled circles) are compared to the whole SDSS sample (contours, squares). Galaxies detected by WENSS are biased towards luminous red galaxies. The fraction of galaxies detected by WENSS is 2.5 per cent.

sample (bottom right-hand panel in Fig. 12); the latter have a larger fraction of sources with 'steep' spectra ($\alpha \sim -1$). This difference is probably caused by distant radio galaxies that are not present in the 'main' SDSS sample, and by quasars. For further discussion of the distribution of galaxies and quasars in radio 'colour–colour' diagrams, we refer the reader to Ivezić et al. (2004a).

We also analysed the radio-to-optical flux ratio as a function of $u - r$. The top left-hand panel in Fig. 12 shows the radio-optical colour[8] $z - t_{\rm NVSS}$ as a function of $u - r$ colour for a subsample of SDSS–NVSS–WENSS galaxies with redshift in the range $0.10 \leqslant z \leqslant 0.14$. We use the $z$ band because the dependence of the bolometric correction for galaxies on colour is the smallest in this band (see Section 4.1), and restrict the redshift range to minimize the effects of $K$-correction. There is no discernible correlation between the radio-optical and $u - r$ colour. However, the measured distribution of the radio-optical colour, shown in the bottom left-hand panel, is subject to numerous selection effects (such as multiple faint flux limits), and it is hard to uncover the intrinsic distribution without detailed simulations, which will be attempted elsewhere.

[8] Following Ivezić et al. (2002), we express all radio fluxes on AB magnitude scale. We would like to apologize to radio astronomers, as this seemed less of a problem than expressing SDSS and the UV-to-IR measurements in Janskys.

### 3.4 GALEX

The distribution of SDSS galaxies detected by GALEX[9] in colour–magnitude–redshift space is compared to the distribution of all SDSS galaxies in Fig. 13. The requirement that a galaxy is detected by GALEX introduces a bias towards blue galaxies and lower redshift (the middle two panels). The fraction of SDSS 'main' galaxies detected by GALEX[10]is ∼42 per cent, and approaches 100 per cent at the bright end (the bottom left-hand panel). The comparison of their $u - r$ distribution with those shown in Fig. 3 suggests that they are dominated by star-forming ($u - r \lesssim 2.2$) galaxies, but also include AGN ($2 < u - r < 3$) galaxies. The UV colours measured by GALEX support this conclusion (Agüeros et al. 2005). Furthermore, the majority (70 per cent) of these galaxies have emission lines, and their distribution in the BPT diagram (discussed in more detail in Section 4.3) confirms that AGN emission, rather than starbursts, is the origin of UV flux in at least 10 per cent of SDSS–GALEX galaxies.

[9] We analyse sources listed in the GALEX Early Release Observations; see Section A3 for more details.

[10] For an analysis of SDSS sources detected by GALEX that is not limited to SDSS 'main' galaxies, we refer the reader to Agüeros et al. (2005) and references therein.

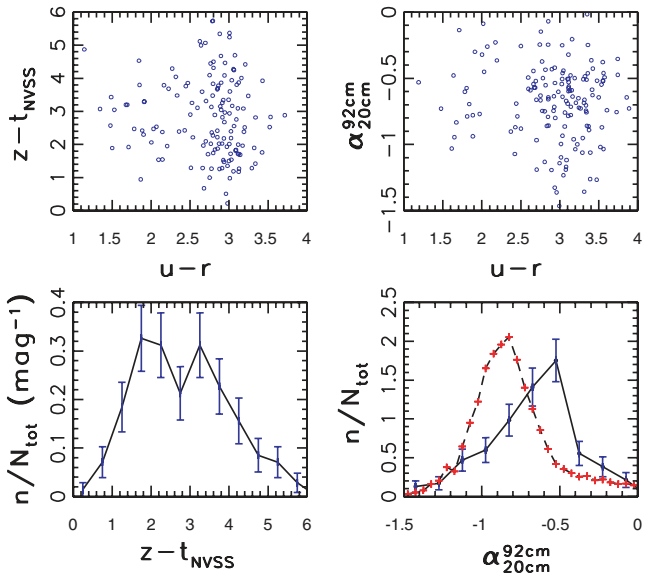


**Figure 12.** The top left-hand panel shows the radio-optical colour (20 cm and the $z$ band) versus $u - r$ colour for a subsample of these galaxies with redshift in the range $0.10 \leqslant z \leqslant 0.14$, and the bottom left-hand panel shows the distribution of their radio-optical colour. The top right-hand panel shows the radio spectral slope ($\alpha$ for $F_\nu \propto \nu^\alpha$ between 20 and 92 cm) versus $u - r$ colour for SDSS galaxies detected by NVSS and WENSS surveys. The bottom right-hand panel compares the distribution of the radio spectral slope for this sample (histogram with error bars) to the distribution of the radio spectral slope for all radio sources detected by the NVSS and WENSS surveys (histogram with pluses).

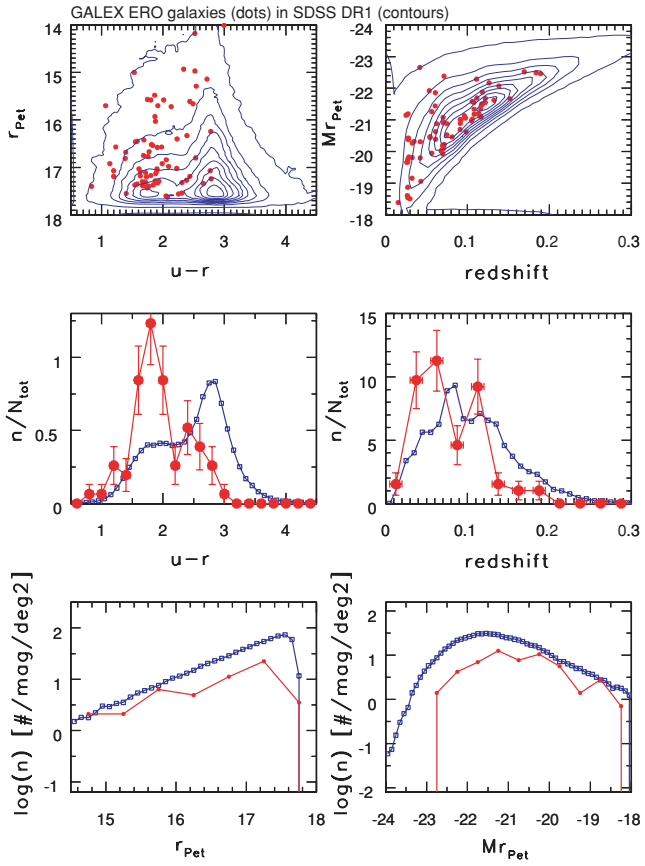


**Figure 13.** Analogous to Fig. 4, except that here galaxies detected by GALEX (dots, filled circles) are compared to the whole SDSS sample (contours, squares). Galaxies detected by GALEX are biased towards blue galaxies, and are dominated (70 per cent) by emission-line galaxies. The latter include both star-forming and AGN (∼10 per cent) galaxies (compare the middle left-hand panel to Fig. 3). The fraction of SDSS 'main' galaxies detected by GALEX is ∼42 per cent, and approaches 100 per cent at the bright end.

### 3.5 *ROSAT* FSC

Matching *ROSAT* X-ray detections (including both hard and soft X-ray data) to SDSS DR1 optical counterparts produced a low (0.63 per cent) matching fraction. Colour–magnitude–redshift diagrams (Fig. 14) reveal a bias toward red galaxies, similar to that seen for radio surveys, but without a redshift bias. The small sample size prevents more detailed analysis. We refer the reader to Anderson et al. (2003) for an analysis of SDSS sources detected by *ROSAT* that is not limited to SDSS 'main' galaxies.

## 4 PANCHROMATIC PROPERTIES OF SDSS GALAXIES

In this section, we combine the data from multiple surveys to construct and compare the UV–IR SEDs for various subsamples of galaxies, and analyse the changes in the BPT diagram induced by requiring detection at different wavelengths spanning the X-ray to radio range.

### 4.1 Dependence of colours on redshift and mean SEDs

The broad-band colours (SED) of a galaxy depend both on its type and redshift (*K*-correction). The dependence of optical and infrared colours on redshift is illustrated in Fig. 15 for SDSS 'main' galaxies listed in the 2MASS XSC and *IRAS* FSC. We used these diagrams to select a narrow redshift range with a sufficient number of galaxies to construct the median SEDs for various subsamples (i.e. we use the median of each colour to construct the overall SED). The median SEDs constructed with GALEX, SDSS and 2MASS photometry for two subsamples of galaxies separated by $u - r$ colour following Strateva et al. (2001) and with redshifts in the $0.03 \leqslant z \leqslant 0.05$ range (small enough that the colours are essentially rest frame) are shown in the top panel in Fig. 16 (when constructing SEDs as a function of wavelength, we use Vega to AB conversion for 2MASS magnitudes from Finlator et al. 2000).

Galaxies with blue $u - r$ have all other colours, in the plotted wavelength range, bluer than galaxies with red $u - r$. Equivalently, the galaxy SEDs constructed with GALEX, SDSS and 2MASS data are a nearly one parameter family [GALEX far-UV (FUV) measurements do provide some additional information which cannot be extracted from SDSS and 2MASS broad-band measurements, see Agüeros et al. 2005]. In particular, we demonstrated in Section 3.1.1 that the 2MASS *K*-band flux can be predicted within 0.2 mag using SDSS $u$ and $r$ fluxes. Smolčić et al. (2006) discuss an even tighter one-dimensional behaviour of galaxies at wavelengths probed by SDSS.

The bottom panel in Fig. 16 shows the same SEDs as in the top panel, except that a linear scale is used instead of a logarithmic scale, and the SEDs are normalized by the bolometric flux. The bolometric flux is determined by integrating a spline fit to the nine data points provided by GALEX, SDSS and 2MASS, and using Rayleigh–Jeans

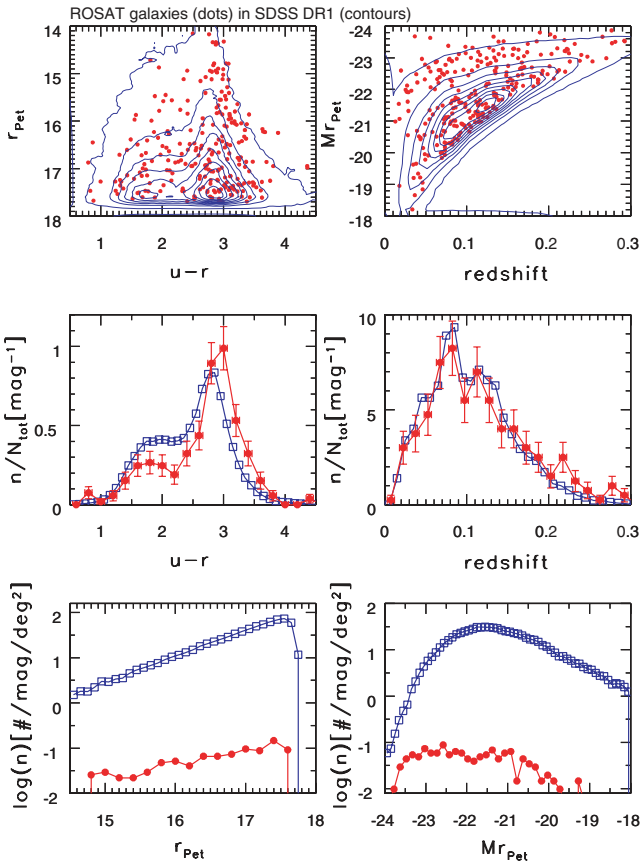


**Figure 14.** Analogous to Fig. 4, except that here galaxies detected by *ROSAT* (dots, filled circles) are compared to the whole SDSS sample (contours, squares). Galaxies detected by *ROSAT* are biased towards red luminous galaxies. The fraction of galaxies detected by *ROSAT* decreases from 0.6 per cent for $r_{\rm Pet} = 14.5$ to 0.2 per cent for $r_{\rm Pet} = 17.5$.

extrapolation at wavelengths longer than 2.2 μm. The data values shown in the figure are also listed in Table 2.

The two normalized SEDs cross around the SDSS $z$ band (∼0.9 μm); showing that the dependence of the bolometric correction for galaxies on colour in the SDSS photometric system is the smallest in the $z$ band. Hence, the flux measured in this band provides the best approximation to the bolometric flux (up to a constant; the important feature is the absence of colour dependence). This band also has the smallest $K$-correction for the redshifts probed by the SDSS 'main' galaxy sample (since the SED slope is the smallest around this wavelength range), and is less sensitive to dust extinction than other SDSS bands. Thus, a good colour-independent estimate of the bolometric flux (in the 0.2–2.2 μm wavelength range) can be simply obtained from the expression

$$(\nu F_\nu)_z = 0.58\, F_{\rm bol}, \tag{4}$$

or, equivalently,

$$L_{\rm bol} = 20.2\, 10^{-0.4\, M_z}\, {\rm L}_\odot, \tag{5}$$

where $M_z$ is the absolute SDSS $z$-band magnitude. The uncertainty of these estimates is of order 5–10 per cent (including calibration and SED integration errors, but not the individual $z$-magnitude measurement error, which can exceed 10 per cent for faint galaxies).

We caution that the $z$ band should be used as a proxy for bolometric flux only for galaxies at redshifts $\lesssim$0.2. For galaxies with larger redshifts the 2MASS $J$-band measurement should be used instead

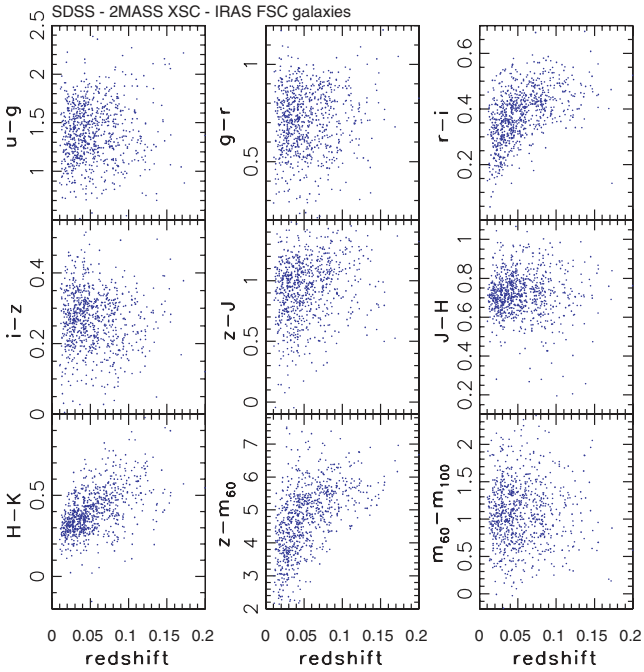


**Figure 15.** The dependence of optical and infrared colours on redshift for SDSS galaxies listed in 2MASS XSC and *IRAS* FSC catalogues. colours are constructed with SDSS 'model' magnitudes (AB system) and 2MASS XSC 'default' magnitudes (Vega system). *IRAS* measurements at 60 μm and 100 μm are expressed on AB system ($m_{60}$ and $m_{100}$).

(and the $H$ band for galaxies with redshifts beyond 0.6, though the number of such galaxies detected by 2MASS may be extremely small).

### 4.2 Do star-forming and AGN galaxies have different UV, IR and radio properties?

In Section 2.1.2 we demonstrated that star-forming and AGN galaxies, classified using *only* emission-line strengths, have different broad-band optical properties such as $u - r$ and concentration index. Here we extend that analysis and compare their UV, IR and radio properties.

#### *4.2.1 The UV-colour difference between star-forming and AGN galaxies*

The SDSS–GALEX sample discussed here is fairly small, and we only determined the median FUV-to-NUV colours,[11] $f - n$ ($f$ and $n$ are AB magnitudes measured in the FUV and NUV GALEX bands), for the two classes. We find that SFs are bluer (median $f - n$ is 0.1 ± 0.1) than AGNs (median $f - n$ is 0.5 ± 0.1). This is similar to the difference observed for $u - r$ (i.e. SFs are bluer than AGNs). However, the analysis by Agüeros et al. (2005) suggests that the origin of the colour differences is different for the $u - r$

[11] For GALEX detections we require $n < 21$ or $f < 21$ and correct magnitudes for interstellar extinction using $A_f = 2.97A_r$ and $A_n = 3.23A_r$, where $A_r$ is the $r$-band extinction from the maps of Schlegel, Finkbeiner & Davis (1998) distributed with SDSS data. These coefficients were evaluated using the standard interstellar extinction law from Cardelli, Clayton & Mathis (1989; M. Seibert private communication). The median $A_r$ for the three AIS fields is 0.12, with an rms scatter of 0.02 mag.

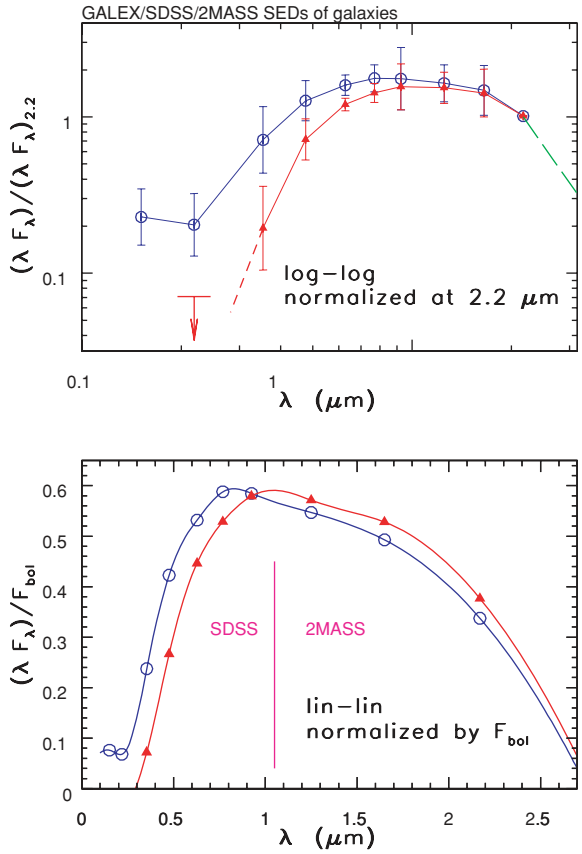


**Figure 16.** The top panel compares the UV-to-NIR SEDs for blue ($u - r < 2.22$, circles) and red ($u - r > 2.22$, triangles) galaxies detected by the GALEX, SDSS and 2MASS surveys, and with redshifts in the range $0.03 \leqslant z \leqslant 0.05$. The error bars indicate the rms colour scatter (determined from the interquartile range). The data points are connected to guide the eye. The dashed line extending redwards from the *K*-band point (2.2 μm) is a Rayleigh–Jean extrapolation. The bottom panel shows the same data on a linear scale, normalized by the bolometric flux, with points connected by a spline fit. The dependence of the bolometric correction on colour is the smallest in the SDSS *z* band (∼0.9 μm).

**Table 2.** The median UV-to-NIR SEDs, normalized by the bolometric flux ($\nu F_\nu / F_{\rm bol} \equiv \lambda F_\lambda / F_{\rm bol}$), for blue ($u - r < 2.22$) and red ($u - r > 2.22$) galaxies with redshifts in the range $0.03 \leqslant z \leqslant 0.05$.

| Bandpass | λ (μm) | Blue | Red |
|---|---|---|---|
| *f* | 0.15 | 0.08 | <0.01 |
| *n* | 0.22 | 0.07 | <0.02 |
| *u* | 0.35 | 0.24 | 0.07 |
| *g* | 0.48 | 0.42 | 0.27 |
| *r* | 0.63 | 0.53 | 0.45 |
| *i* | 0.77 | 0.59 | 0.53 |
| *z* | 0.93 | 0.58 | 0.58 |
| *J* | 1.25 | 0.55 | 0.57 |
| *H* | 1.65 | 0.49 | 0.53 |
| *K* | 2.17 | 0.34 | 0.38 |

and $f - n$ colours, because the latter are much more sensitive to the presence of starbursts and AGNs, while the $u - r$ colour reflects the bolometrically dominant stellar population, as suggested by the analysis described in Section 3.1.1.

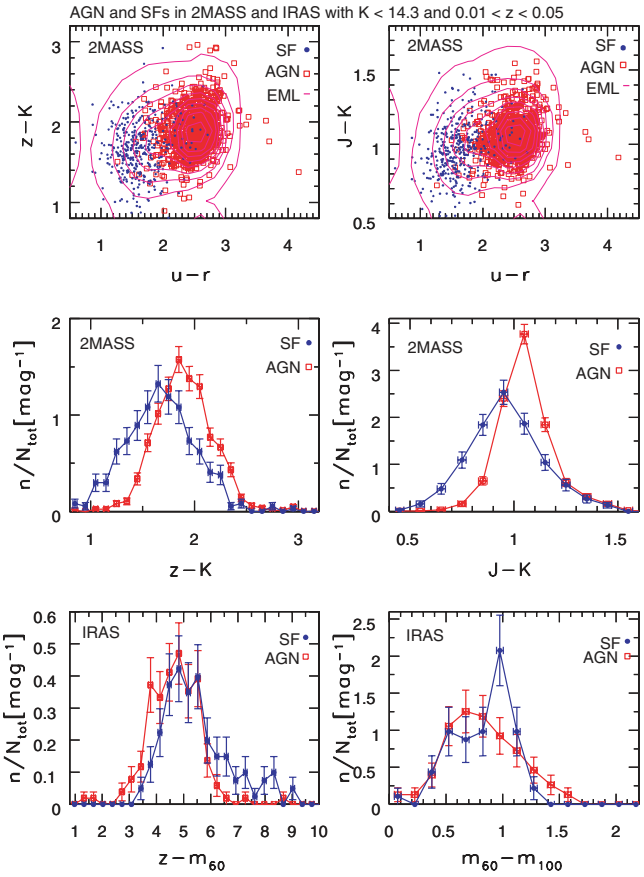


**Figure 17.** The comparison of optical and infrared colours for galaxies detected by SDSS, 2MASS and *IRAS* surveys, in a restricted redshift range ($0.01 \leqslant z \leqslant 0.05$) to avoid *K*-correction effects. Star-forming galaxies (designated by 'SF') show more far-IR emission, relative to optical/IR, than AGN galaxies, but bluer optical and NIR colours.

### *4.2.2 The IR colour difference between star-forming and AGN galaxies*

In order to study differences in infrared properties between star-forming and AGN galaxies, we use the SDSS–2MASS and SDSS–*IRAS* samples, and restrict the redshift range to $0.01 < z < 0.05$ to avoid *K*-correction effects (further discussed below). The distribution of these galaxies in various optical/infrared colour–colour diagrams is shown in Fig. 17. In general, for any combination of optical and NIR bands, SFs are bluer than AGNs. For example, the $z - K$ and $J - K$ mean colours differ by 0.2 and 0.1 mag, respectively. According to the Kolmogorov–Smirnov test (KS hereafter), these differences are highly statistically significant, as is easily discernible from the two middle panels. However, this relationship reverses when using far-IR bands; SFs have *redder* $z - m_{60}$ colours by ∼1 mag than AGNs (the KS probability that the two subsamples are drawn from the same parent sample is $\sim 10^{-8}$) and have a much higher fraction of sources with $z - m_{60} \gtrsim 6$ (see the bottom left-hand panel). This reversal is illustrated in Fig. 18 which compares the optical-to-far-IR SEDs, normalized to the SDSS *z* band, for star-forming and AGN galaxies selected from a narrow redshift range. Since the *z*-band flux is a good measure of the 0.2–2.2 μm bolometric luminosity, this implies that SFs emit more IR radiation as a fraction of their 0.2–2.2 μm bolometric output, than do AGN galaxies. This difference could be partially caused by a selection effect;

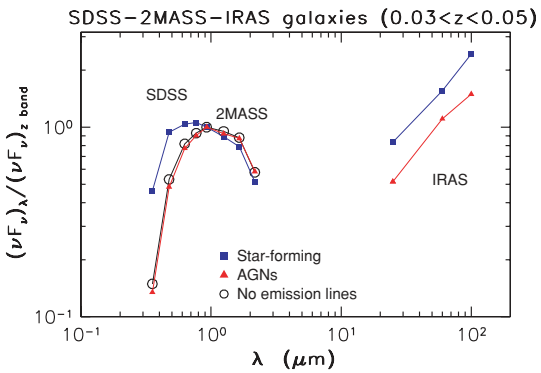


**Figure 18.** The optical-to-far-IR SEDs, normalized to the SDSS $z$ band, shown separately for galaxies without emission lines and AGN and star-forming emission-line galaxies detected by SDSS, 2MASS and *IRAS*, and with redshifts in the $0.03 \leqslant z \leqslant 0.05$ range. Only the latter two subsamples have a sufficient number of *IRAS* detections to characterize the far-IR SED.

since the apparent $z$-band magnitudes of SFs tend to be somewhat fainter that those of AGN galaxies due to differences in luminosity functions and sampled redshift range, they will be detected by *IRAS* with the same probability only if they have somewhat redder $z - m_{60}$ colours. However, this effect does not seem to be quantitatively sufficient to explain the observed difference in median $z - m_{60}$ colours.

There are several plausible explanations for the different median $z - m_{60}$ colours for AGN and SFs: SFs could have warmer dust, significantly more dust, or more UV radiation that is absorbed by dust and re-emitted in the far-IR[12] than AGN galaxies (of course, these possibilities are not mutually exclusive). Different dust temperatures seems an unlikely explanation because two types have very similar $m_{60} - m_{100}$ colour distributions; this colour is much more sensitive to dust temperature than to the amount of dust; for a detailed discussion see e.g. Ivezić & Elitzur 1997), as shown in the bottom right-hand panel in Fig. 17. In order to test the hypothesis that the difference is caused by a different amount of dust, we compared the distributions of extinction estimates, $A_z$, discussed in Section 3.2.1. As discernible from the bottom panel in Fig. 19, SFs detected by *IRAS* typically have *smaller* $A_z$ than AGN galaxies, thus ruling out this hypothesis.

Thus, it appears that SFs emit more far-IR radiation (as a fraction of bolometric flux in the 0.2–2.2 μm wavelength range) than AGN galaxies, despite having smaller $A_z$ of dust at similar temperatures, because they have significantly more UV radiation (again, as a fraction of bolometric flux in the 0.2–2.2 μm wavelength range) that is processed into far-IR wavelength range. This conclusion is supported by GALEX data discussed by Agüeros et al. (2005), who find a correlation between $f - n$ and contribution of the UV flux to the bolometric flux. Since SFs tend to have bluer $f - n$ colours than AGNs (see Section 4.2.1), this implies that their UV flux contributes more to the bolometric flux than for AGN galaxies. Indeed, once a larger SDSS–GALEX–*IRAS* sample is available, the expected strong correlation between $f - m_{60}$ and $A_z$ (similar to, and perhaps stronger than, the correlation shown in Fig. 9) can be directly tested.[13]

[12] The observed difference could also be due to different dust optical properties, i.e. different ratio of far-IR to UV/optical opacity. While this possibility is not excluded by our analysis, it is not necessary in order to explain the observed trends. Similarly, we ignore the possible effects of dust geometry, e.g. the dust around AGN could have significantly different distribution than interstellar dust (e.g. Nenkova, Ivezić & Elitzur 2002).

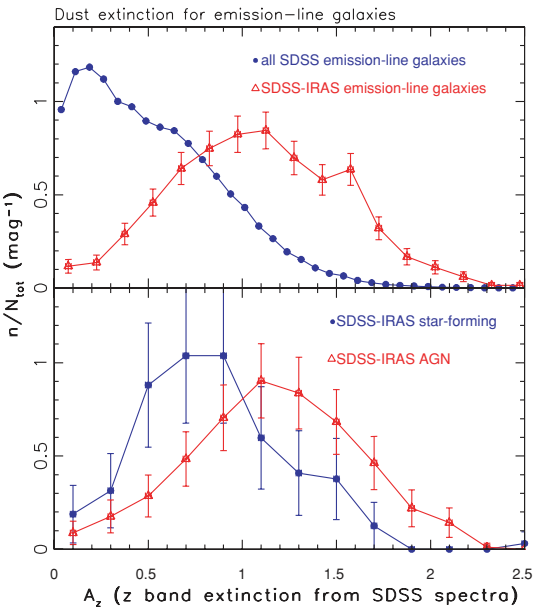


**Figure 19.** The top panel compares the $z$-band dust extinction inferred from SDSS spectra by Kauffmann et al. (2003a), $A_z$, for all SDSS emission-line galaxies (dots) and for those also detected by *IRAS* (triangles; analogous to the top panel in Fig. 9, which also includes galaxies without emission lines). The bottom panel compares the $A_z$ distributions for AGN (triangles) and SFs (dots) detected by *IRAS*. Note that SFs detected by *IRAS* typically have *smaller* $A_z$ than AGN galaxies.

In a recent paper based on similar data, Pasquali et al. (2005, hereafter PKH) found that the 'AGN exhibit a significant excess in far-IR emission relative to SFs', which apparently contradicts the results presented here. However, it should be pointed out that here we discuss the amount of IR radiation (at 60 μm) as a fraction of bolometric output in the 0.2–2.2 μm range (i.e. $z - m_{60}$ colour), while PKH discuss the IR luminosity *per unit stellar mass*. Furthermore, their statement is valid for carefully selected pairs of AGN and SFs that have similar physical characteristics such as stellar mass, colour, size, etc. (it is also noteworthy that the distributions of stellar mass and stellar mass-to-light ratio are very different for AGN and SFs, as shown by Kauffmann et al. 2003a). Hence, these are two independent, rather than contradictory, findings.

### *4.2.3 The radio–IR correlation*

The large number of SDSS–*IRAS*–NVSS emission-line galaxies allows us to examine whether the well-known narrow distribution of the far-IR-to-radio flux ratio (van der Kruit 1971) is the same for (optically classified) star-forming and AGN galaxies. This correlation was interpreted by Helou, Soifer & Rowan-Robinson (1985) as

[13] The analysis of Buat et al. (2005) appears to support this expectation: they find that interstellar extinction for a far-IR-flux-limited sample is higher than for a UV-flux-limited sample. However, their estimates of interstellar extinction are derived from the UV–IR colours (rather than from independent data), and thus are strongly correlated with the sample flux limits. That is, the same effect would be obtained even if there were no correlation between the amount of dust in a galaxy and the UV–IR colours, and hence their analysis *cannot* be used to support our interpretation.

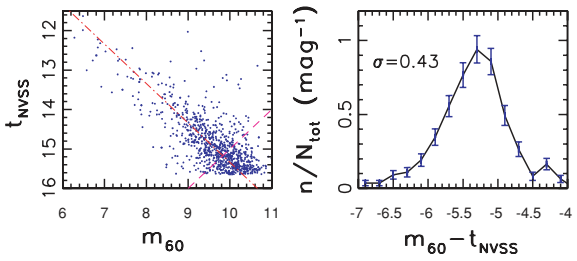


**Figure 20.** The left-hand panel shows the radio 20 cm magnitude measured by NVSS as a function of 60-μm flux measured by *IRAS*, for 948 SDSS galaxies with a counterpart in *IRAS* FSC within 30 arcsec and in the NVSS catalogue within 15 arcsec. The right-hand panel shows the distribution of far-IR–radio colour, $m_{60} - t_{\rm NVSS}$, for 568 galaxies with $m_{60} + t_{\rm NVSS} < 25$ (upper left-hand region, away from the dashed line, shown in the left-hand panel). The median value of $m_{60} - t_{\rm NVSS}(-5.31)$ is shown as the diagonal dot–dashed line in the left-hand panel, and the rms distribution width determined from the interquartile range is marked in the right-hand panel.

a consequence of coupling between infrared thermal dust emission and radio non-thermal synchrotron emission, and it is not known whether the details of this coupling are the same for star-forming and AGN galaxies (e.g. see Bell 2003). In addition, both far-IR and radio emission are used as probes for ongoing and recent star formation (Hopkins et al. 2003 and references therein), and thus a robust measurement of this correlation is important for the comparison of studies based on different data sets.

The left-hand panel in Fig. 20 shows the radio 20 cm magnitude measured by the NVSS as a function of 60 μm magnitude measured by *IRAS* for 948 SDSS 'main' galaxies with a counterpart in *IRAS* FSC within 30 arcsec, and in the NVSS catalogue within 15 arcsec. The strong correlation between the two fluxes[14] is evident. Only about 1 per cent of sources have anomalously bright radio emission (by about ∼2 mag), a fraction that is consistent with contamination by random associations (the same fractions of AGN and SFs are found among those 1 per cent of sources, as in the whole sample; they are all bright point sources at 20 cm), but could also be due to radio-loud objects. In order to avoid the effect of faint flux limits on the distribution of the far-IR-to-radio flux ratio (i.e. the $m_{60} - t_{\rm NVSS}$ colour), we restrict the sample to 568 galaxies with $m_{60} + t_{\rm NVSS} < 25$ (a condition perpendicular to the $m_{60} - t_{\rm NVSS} =$ constant lines in the $t_{\rm NVSS}$ versus $m_{60}$ plane, see the dashed line in the left-hand panel in Fig. 20), and show the distribution of the $m_{60} - t_{\rm NVSS}$ colour in the right-hand panel. The median value of the $m_{60} - t_{\rm NVSS}$ colour is $-5.31 \pm 0.02$ (statistical errors only), with an equivalent Gaussian width (determined from interquartile range) of 0.43 mag. Such a narrow width is quite remarkable given so different wavelengths and survey technologies, and represents a strong constraint for the theories of coupled radiation mechanisms (e.g. Helou et al. 1985). It also places an upper limit of 0.4 mag on the *IRAS* photometric error for galaxies.

A very similar analysis of IR–radio correlation based on a smaller sample (176 UGC galaxies) was presented by Condon & Broderick (1988), who used a slightly different parameter $u = -0.4(m_{60} - t_{\rm NVSS})$ (sometimes also called the $q$ parameter, e.g. Bell 2003). They find a peaked distribution similar to that shown in Fig. 20, with 63 galaxies in the peak. The position of that peak corresponds to $m_{60} - t_{\rm NVSS} = -5.05 \pm 0.05$, in good agreement with our analysis based on a ∼10 times larger sample. The width of the $m_{60} - t_{\rm NVSS}$ distribution determined here (0.43 mag) corresponds to a width of 0.17 for the $u$ (or $q$) distribution, somewhat smaller than ∼0.26 obtained in previous studies (e.g. Yun, Reddy & Condon 2001; Bell 2003). Condon & Broderick proposed that galaxies in this peak ($m_{60} - t_{\rm NVSS} < -4$ ) are dominated by starbursts, while AGNs ('monsters' in their terminology) have too weak IR emission to be detected by *IRAS*, implying $m_{60} - t_{\rm NVSS} > -4$.

The high-quality SDSS spectra and corresponding BPT-diagram based separation of emission-line galaxies into AGN and SFs allows us to examine IR–radio correlation in detail. We find that over 1/3 of SDSS–NVSS–*IRAS* galaxies with emission lines can be reliably classified as an AGN based on their position in the BPT diagram, with an additional 50 per cent in the transition region ('unknown' sources, using classification described in Section 2.1.2). That is, in addition to SFs, optically classified AGN galaxies also follow the tight IR–radio correlation. Hence, not all the 'monsters' are confined to $m_{60} - t_{\rm NVSS} > -4$ (assuming that they are correctly recognized by BPT analysis).

We note that the AGN/SF separation in the BPT diagram adopted here is fairly conservative (see Section 2.1.2). Fig. 21 demonstrates that the majority of SDSS–*IRAS*–NVSS emission-line galaxies have ambiguous classification in the BPT diagram. However, it is easily discernible that a substantial fraction of these galaxies have optical emission line-strength ratios fully consistent with an AGN classification. The figure also shows that the position of a galaxy in the BPT diagram and the $m_{60} - t_{\rm NVSS}$ colour do not appear to be correlated. Visual inspection of SDSS $g$, $r$, $i$ colour composite images convincingly shows that *both* AGN and star-forming subsamples are dominated by galaxies with spiral morphology, as already pointed out by Condon & Broderick (1988). About half of them show disturbed morphology and nearby companions.

Using classification based on the BPT diagram, we compute and compare the slope of IR–radio correlation separately for AGN and

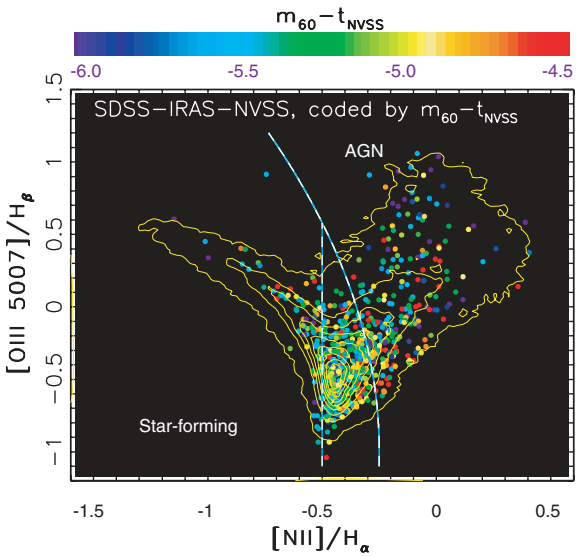


**Figure 21.** The distributions of SDSS–*IRAS*–NVSS emission-line galaxies (dots) and all SDSS emission-line galaxies (contours) in the BPT diagram. The dots are coloured according to their $m_{60} - t_{\rm NVSS}$ colour, as shown in the legend on top. *Both* AGN and SFs are found in this sample, and they all follow the well-known radio–IR correlation.

[14] Sometimes this correlation is shown in the luminosity versus luminosity form, which boosts the impression of correlation strength. However, this is not a good practice since even uncorrelated measurements could give such an impression if the redshift distribution is sufficiently broad.

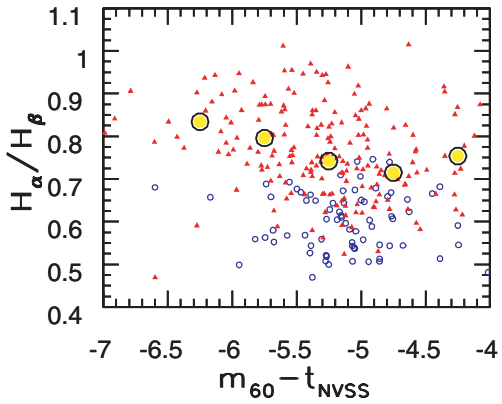


**Figure 22.** The distribution of AGN and SFs in the radio–IR colour, $m_{60} - t_{\rm NVSS}$, versus optical emission line-strength ratio, H$\alpha$/H$\beta$. Small blue circles show SFs and red triangles AGN galaxies. The large circles are the median values of H$\alpha$/H$\beta$ in narrow $m_{60} - t_{\rm NVSS}$ bins for AGN galaxies, and demonstrate that the strength of IR emission for AGN galaxies, relative to their radio emission, increases with the H$\alpha$/H$\beta$ line-strength ratio.

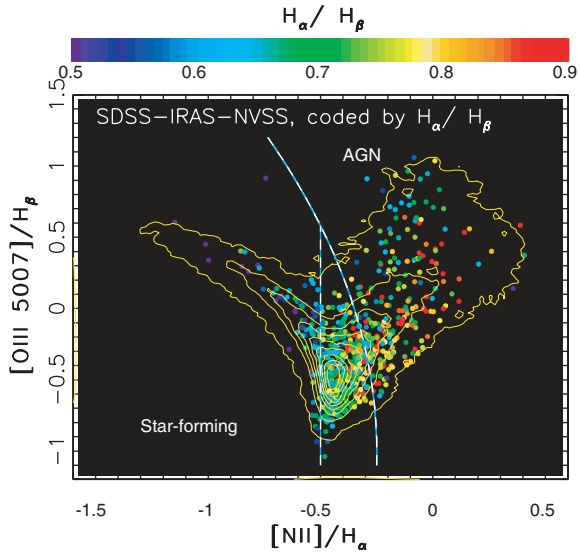


**Figure 23.** The distributions of SDSS–*IRAS*–NVSS emission-line galaxies (dots) and all SDSS emission-line galaxies (contours) in the BPT diagram. The dots are coloured according to their H$\alpha$/H$\beta$ line-strength ratio, as shown in the legend on top. Note that objects with the largest values of H$\alpha$/H$\beta$ ratio (red) are found almost exclusively in AGN region.

SFs. Two subsamples with 128 AGN and 46 SFs yield median $m_{60} - t_{\rm NVSS}$ values of $-5.35 \pm 0.05$ and $-5.14 \pm 0.05$, and widths of 0.61 and 0.38 mag, respectively. [15] The difference between the medians is $\sim 3\sigma$ significant, and suggests that the details of radiation coupling mechanisms may be different for star-forming and AGN galaxies (i.e. among galaxies that follow IR–radio correlation, SFs appear to show ∼20 per cent more radio emission, relative to far-IR, than AGN galaxies). We note that this result should be considered somewhat tentative because there may be systematic effects that are not included when estimating uncertainties in the medians,[16] and thus its significance could be overestimated. It is also noteworthy that none of SFs have $m_{60} - t_{\rm NVSS} > -4$, while this is true for 4 per cent of AGNs (i.e., an opposite trend than for medians).

We searched for possible correlations between $m_{60} - t_{\rm NVSS}$ and other observables that could perhaps explain the different median values of this colour for AGN and star-forming subsamples (either as selection effects, bad data, bad analysis method or astrophysics). We analysed quantities such as colours, redshift, stellar mass, luminosity, angular size, dust extinction estimate $A_z$, etc., and the only quantity that appears to have an effect on $m_{60} - t_{\rm NVSS}$ is the H$\alpha$/H$\beta$ ratio, as shown in Fig. 22 (see also Fig. 23 which illustrates correlation between the H$\alpha$/H$\beta$ ratio and the position in BPT diagram). SFs are confined to the H$\alpha$/H$\beta$ < 0.7 region, while AGN galaxies span the whole observed range of H$\alpha$/H$\beta$. Furthermore, as the H$\alpha$/H$\beta$ line-strength ratio increases, the strength of IR emission for AGN galaxies, relative to radio emission, also increases (i.e. the $m_{60} - t_{\rm NVSS}$ colour becomes bluer). Since the measurement of $m_{60} - t_{\rm NVSS}$ is fully independent of the H$\alpha$/H$\beta$ measurement, this behaviour provides additional support for the $3\sigma$ significant difference in the slopes of IR–radio correlation for star-forming and AGN galaxies (note, however, that the AGN presence could have an effect on the measured H$\alpha$/H$\beta$ ratio). It is noteworthy that there is no evidence for a similar behaviour in the $m_{60} - t_{\rm NVSS}$ versus $A_z$ diagram, although some degree of correlation exists between H$\alpha$/H$\beta$ and $A_z$. This may mean that the effects of gas and dust on infrared and radio emission are more complex than implied by a simple linear IR–radio correlation (for a detailed discussion of this possibility see Bell 2003).

Given that both AGN and star-forming subsamples of SDSS–NVSS–*IRAS* galaxies follow very similar, if not identical, radio–IR correlation, it is interesting to investigate what fraction of the full SDSS–NVSS and SDSS–*IRAS* subsamples could follow this correlation (that is, galaxies that are detected by only two, instead of all three, surveys). We perform this analysis by using the observed radio–IR correlation to predict $t_{\rm NVSS}$ for SDSS–*IRAS* galaxies, or $m_{60}$ for SDSS–NVSS galaxies. In the second step, we select galaxies with predicted fluxes 1 mag brighter than the faint limit of the corresponding third catalogue (to account for the scatter due to photometric errors), and then determine what fraction of these galaxies are actually detected. For example, if *every IRAS* galaxy follows IR–radio correlation, then we estimate that > 90 per cent of SDSS–*IRAS* galaxies with predicted $t_{\rm NVSS} < 15.5$ should be detected by NVSS. We find that indeed 92 per cent of SDSS–*IRAS* galaxies with $t_{\rm NVSS}^{\rm predicted} = m_{60} + 5.3 < 14.5$ are detected by NVSS.

The converse is not true. Only about 1/3 of SDSS–NVSS *emission-line* galaxies with $m_{60}^{\rm predicted} = t_{\rm NVSS} - 5.3 < 9$ are detected by *IRAS*. This is consistent with a hypothesis that there is another source of radio emission, in addition to the component that is correlated with IR emission. The elevated radio emission is then responsible for the NVSS detection, but IR emission is too weak for an *IRAS* detection. Presumably, those AGNs that are detected by both NVSS and *IRAS* are mostly radio-quiet, while those with elevated radio emission are mostly radio-loud. Not surprisingly, SDSS–NVSS emission-line galaxies without *IRAS* detection have typically redder $u - r$ colours than those detected by both NVSS and *IRAS* (the

[15] The $m_{60} - t_{\rm NVSS}$ colour measurement should not be interpreted as implying a power-law SED between 60 μm and 20 cm with the power-law index of ∼0.6 ($F_\nu \propto \nu^\alpha$). For example, galaxies have $\alpha \sim -0.5$ between 6 and 20 cm (Ivezić et al. 2004a), and $m_{60} - m_{100} \sim 1$, implying $\alpha \gtrsim 1$ between 100 μm and 6 cm, that is, a steeper decrease of flux with wavelength, than implied by the $m_{60} - t_{\rm NVSS}$ colours.

[16] Some of this difference could be due to systematically different radio morphology, an interesting possibility that is beyond the scope of this work. It is noteworthy that we did not find a correlation between the $m_{60} - t_{\rm NVSS}$ colour and the radio spectral slope between 20 and 92 cm.

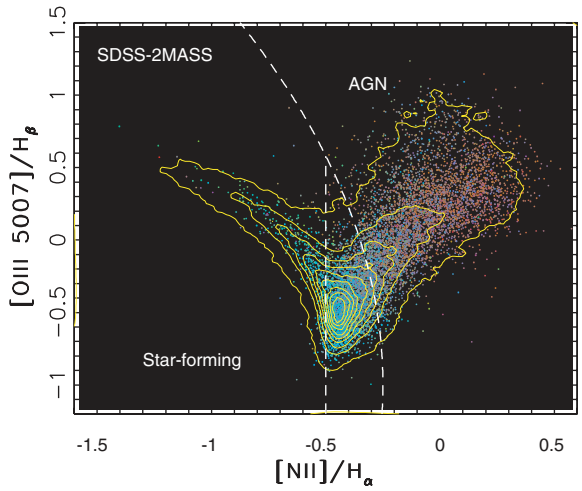


**Figure 24.** The comparison of the distributions of SDSS–2MASS emission-line galaxies (dots) and all SDSS emission-line galaxies (contours) in the BPT diagram. The dots are coloured according to their position in the concentration index versus $u - r$ diagram, shown in the top right-hand panel in Fig. 2. The AGN-to-SF number ratio for SDSS–2MASS galaxies is 10.1, while it is 1.6 for the whole SDSS sample.

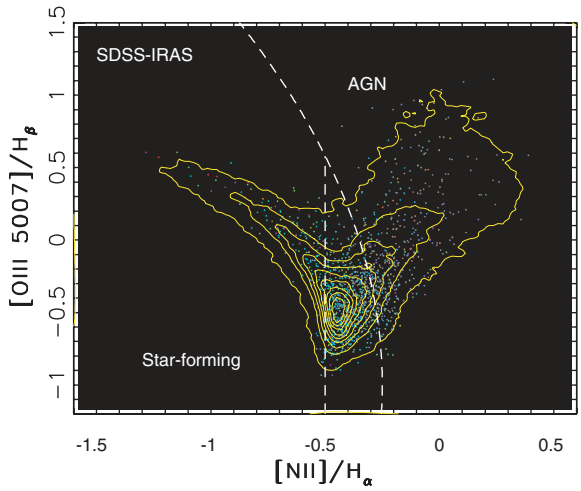


**Figure 25.** Analogous to Fig. 24, except for SDSS–*IRAS* galaxies. The AGN-to-SF number ratio for SDSS–*IRAS* galaxies is 2.7.

difference in medians is 0.6 mag). The detection fraction by *IRAS* is different for AGN and star-forming subsamples of SDSS–NVSS sample: while 50 per cent of star-forming SDSS–NVSS galaxies are detected by *IRAS*, this is true for only 20 per cent of AGNs (for galaxies with $m_{60}^{\rm predicted} < 9$). These statistics indicate that the *majority* (80 per cent) of emission-line galaxies classified *optically* as AGNs emit more flux at radio wavelengths than implied by their IR fluxes and the mean IR–radio correlation. Again, practically all of those remaining 20 per cent AGNs from SDSS–NVSS subsample that are also detected by *IRAS* do follow a very tight IR–radio correlation.

### 4.3 Adding panchromatic information to the BPT diagram

In preceding Sections we have discussed the broad-band panchromatic properties of galaxy samples selected by their emission-line properties with the aid of the BPT diagram (see Section 2.1.2). In this section, we study how the morphology of galaxy distribution in the BPT diagram changes when requiring detections at other wavelengths (see Figs 24–28).

The AGN-to-SF number ratio is 1.6 for the whole SDSS sample. Although we exclude a large number of galaxies with uncertain classification (see Section 2.1.2), this ratio is a good relative measure of the changes in the BPT diagram.[17] We find that the AGN-to-SF number ratio is systematically larger for subsamples with detection at other wavelengths, except for SDSS–GALEX sample, and it is the largest for SDSS–2MASS sample (10.1). Such a high ratio for SDSS–2MASS sample is a consequence of fairly bright *K*-band flux limit, and the fact that AGN galaxies have redder optical-to-NIR SEDs than SFs.

The AGN-to-SF number ratio is the smallest for SDSS–GALEX sample (0.1, but it could be as high as 0.3 if the unclassified galaxies

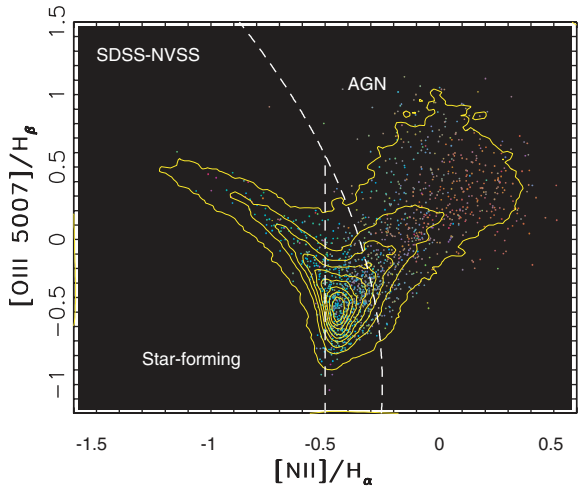


**Figure 26.** Analogous to Fig. 24, except for SDSS–NVSS galaxies. The AGN-to-SF number ratio for SDSS–NVSS galaxies is 4.7.

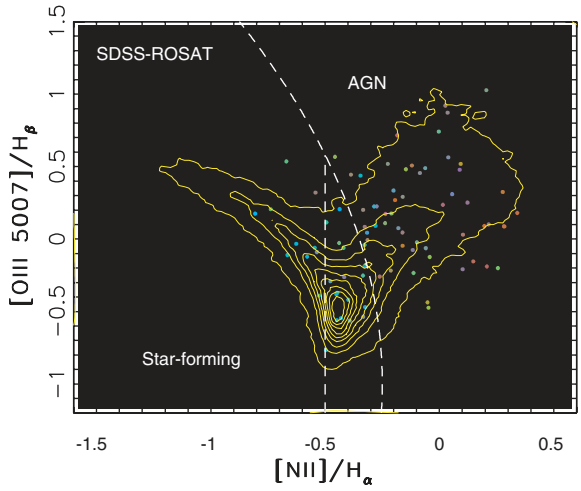


**Figure 27.** Analogous to Fig. 24, except for SDSS–*ROSAT* galaxies. The AGN-to-SF number ratio for SDSS–*ROSAT* galaxies is 5.7.

[17] Note that the AGN-to-SF number ratio depends on the adopted cut-off for the emission-line detection significance. Higher values than the $3\sigma$ adopted here (see Section 2.1.2) would result in a lower AGN-to-SF number ratio because many AGNs are very weak-lined LINERs (Heckman et al. 2004).

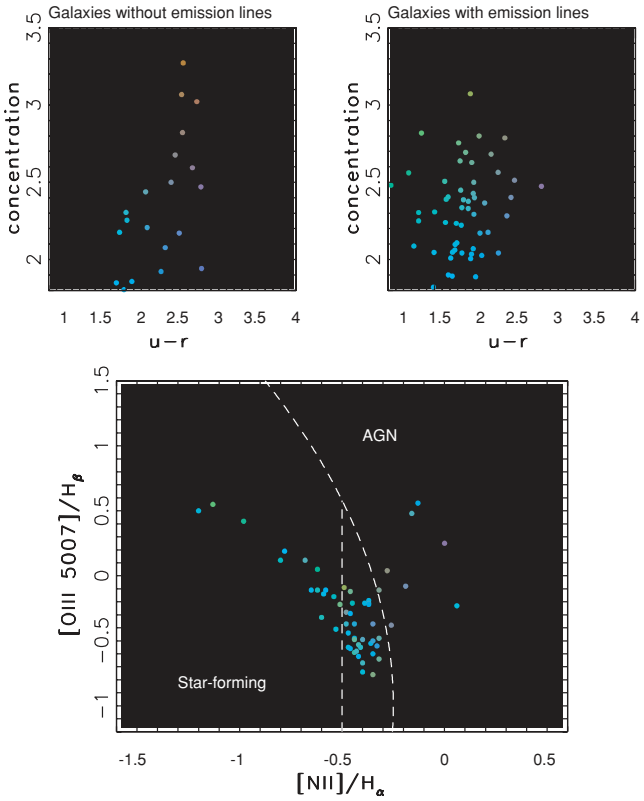


**Figure 28.** Analogous to Fig. 2, except for SDSS–GALEX galaxies. The lower limit on AGN-to-SF number ratio for SDSS–GALEX galaxies is at least 0.1, and it could be as high as 0.3.

are dominated by AGN galaxies, see Fig. 28). The lower limit on this ratio is sufficiently high to exclude the possibility that SDSS–GALEX galaxies represent a clean sample of starburst galaxies. In order to present further evidence for this claim, we have visually inspected SDSS *g*, *r*, *i* colour composite images of these galaxies (a total of 55) and found that the classification based on emission-line strengths is well correlated with morphology. SDSS images of random subsamples of AGN, star-forming, and unclassified galaxies are shown in Fig. 29. Clear morphological differences between galaxies classified as star-forming and as AGN are easily discernible, with the latter being more centrally concentrated. This further demonstrates that *at least some GALEX/SDSS galaxies are more likely to be AGN than star-forming*.

### 4.4 An improvement of the *K*-band flux prediction

In Section 3.1.1, we showed that it is possible to estimate the *K*-band magnitude with a scatter as small as ∼0.2 mag using only SDSS data. In this section, we explore whether the residuals between predicted and 2MASS *K*-band magnitudes correlate with several model-dependent quantities determined by Kauffmann et al. (2003a), and whether the residuals show the same behaviour when AGN and SFs are treated separately.

The strongest correlation between the *K*-band SDSS–2MASS residuals and another quantity is found for $A_z$, the galaxy dust content discussed in detail in Section 3.2.1. The top panel in Fig. 31 illustrates this correlation. A best straight line fit is given by

$$\Delta(r-K)^* = 0.213A_z - 0.033. \quad (6)$$

This correlation may be interpreted as the effect of dust on the observed *r*-band flux (however, note that the simple extinction screen

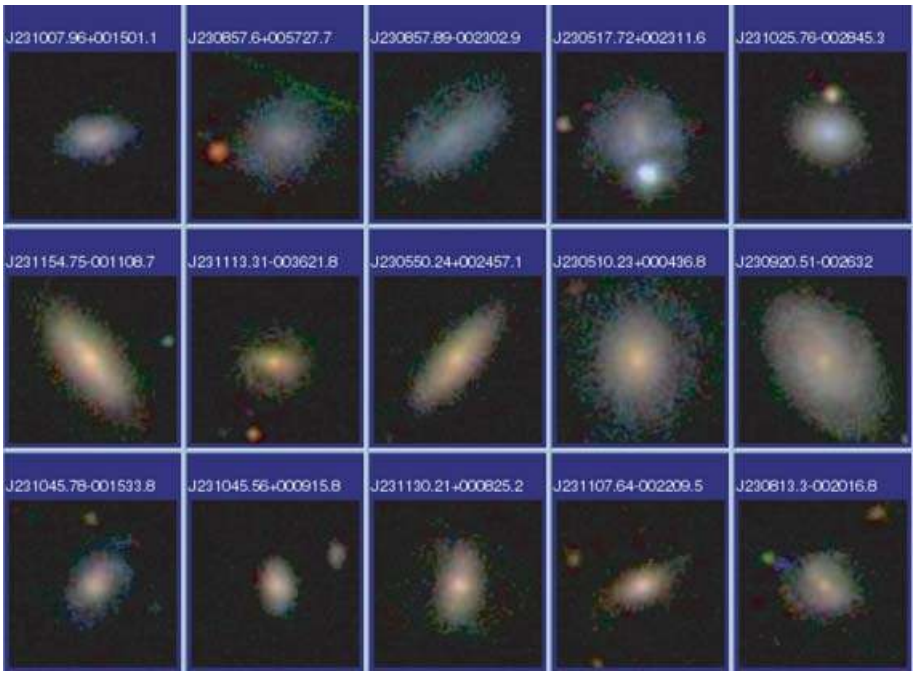


**Figure 29.** *g*, *r*, *i* composite SDSS images of SDSS–GALEX galaxies randomly chosen from three subsamples classified using emission-line strengths measured from SDSS spectra (Kauffmann et al. 2003a). The first row shows images of star–forming galaxies, the second of AGN; and galaxies in the third row have uncertain classifications. North is up, and the images are roughly 25 arcsec on a side.

approximation is probably not appropriate). When this correlation is subtracted from $(r-K)^*$ given by equation (1), the width of the residuals distribution decreases by a factor of 2, to 0.1 mag! In other words, given the $u-r$ colour, redshift and $A_z$ determined using SDSS data, the 2MASS $K$-band measurements can be predicted with a scatter of only 0.1 mag. Most of this scatter can be attributed to the measurement errors. For example, assuming conservative lower limits for errors in $u-r$ (0.03 mag), $K$ (0.03 mag; Jarrett et al. 2000), and $R_{50}^z$ (2 per cent), the expected scatter due to measurement errors is 0.09 mag. Hence, the observed residual scatter of 0.1 mag is likely dominated by measurement errors. We note that the final distribution of the $r-K$ residuals shown in the bottom panel in Fig. 30 is skewed. This could be due to the fact that the dependence of the $r-K$ residuals on $A_z$ was fit by a straight line, while the data display some curvature. Also, it could be that there are two subpopulations of galaxies that have slightly different SEDs.

Since the residual astrophysical scatter is apparently much smaller than 0.1 mag, one is tempted to conclude that 2MASS measurements are not required to study SDSS galaxies. This would not be a valid conclusion for at least two reasons. First, the measurement error for 2MASS $K$-band magnitudes is ∼0.03 mag, which is smaller than the residual scatter for predicted $K$-band magnitudes (0.1 mag). Secondly, we have not investigated morphological properties of galaxies in the NIR, where smaller dust extinction could reveal features not visible in the optical wavelength range.

We have also studied the correlation between the $K$-band SDSS–2MASS residuals and the 4000 Å break ($D_{4000}$), the strength of H$\delta$ line, and stellar mass (Fig. 31). None of the correlations is as strong as the correlation with $A_z$. As shown in Fig. 31, the overall behaviour of AGN and star-forming subsamples are similar to each other, though not identical.

### 4.4.1 *Comparison with Bruzual & Charlot models*

Motivated by the surprisingly small residual scatter in the $K$-band flux prediction (0.1 mag), we have investigated the Bruzual & Charlot (1993, 2003) model predictions for the distribution of optical and NIR galaxy colours. The top left-hand panel in Fig. 32 shows that for a given galaxy type, the $r-K$ colour is a function of redshift, and at a given redshift the $r-K$ colour is a function of galaxy type. This behaviour is in agreement with observations, although model galaxies have the $u-r$ colours too blue by several tenths of a magnitude. The models are also in agreement with a $J-K$ versus redshift relation that is independent of galaxy type, as discussed in Section 3.1.1 (see top right-hand panel).

The behaviour of simple stellar populations (as opposed to synthetic galaxies) in the same diagrams suggests that both the $r-K$ and $J-K$ colours are sensitive to metallicity, with about 0.7 and 0.2 mag redward shifts in the $r-K$ and $J-K$ colours as the metallicity increases by a factor of 10. The observed small residual scatter in the $K$-band flux prediction thus implies that the galaxy metallicity distribution is fairly narrow: about 0.2 dex around the median value. This is in agreement with Tremonti et al. (2004) who found a very tight mass-metallicity relation (0.1 dex metallicity scatter at a given mass).

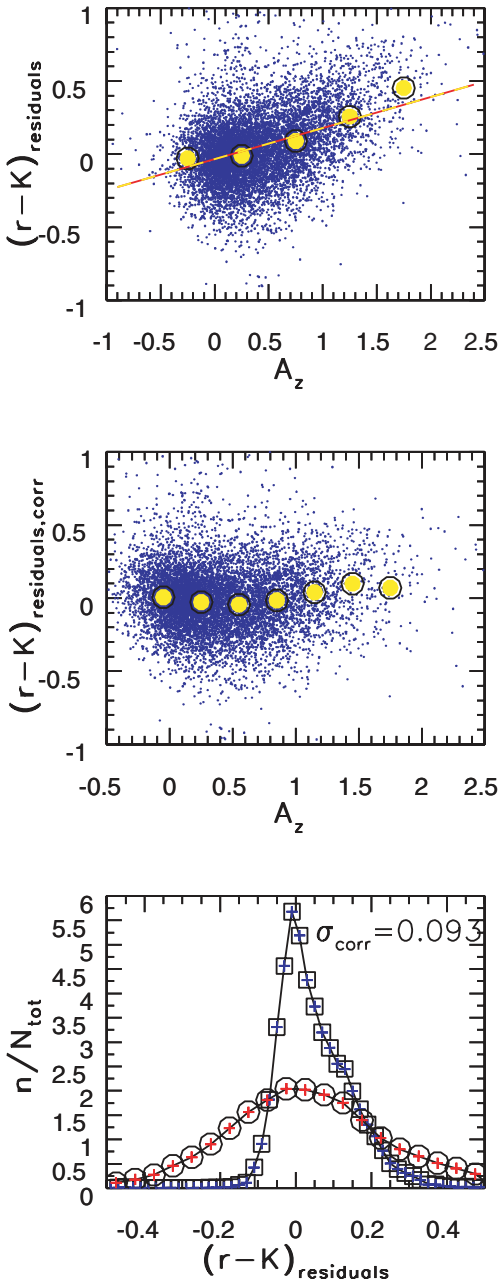


**Figure 30.** The top panel shows the dependence of residuals between predicted and measured $K$-band magnitudes on the $z$-band dust extinction, $A_z$, inferred from SDSS spectra by Kauffmann et al. (2003a). Individual galaxies are shown by small symbols, and large circles show the median values in $A_z$ bins. The best-fitting straight line to these medians is also shown. The residuals corrected for this median trend are shown in the middle panel. The bottom panel compares the distributions of uncorrected residuals (circles) and the corrected ones (squares). The distribution width for the latter is only about one half of that for the former.

## 5 SUMMARY AND DISCUSSION

This study indicates the enormous potential of modern massive sensitive large–scale surveys, and emphasizes the added value obtained by combining data from different wavelengths. While qualitatively our study is in agreement with previous work (e.g. galaxies detected by *IRAS* tend to be blue), the sample size and the wealth of measured parameters allowed us to obtain some qualitatively and quantitatively new results.

Galaxy SEDs form a nearly one-dimensional sequence in the optical-to-NIR range. For example, the SDSS $u$- and $r$-band data, supplemented with redshift and dust content estimate, can be used to

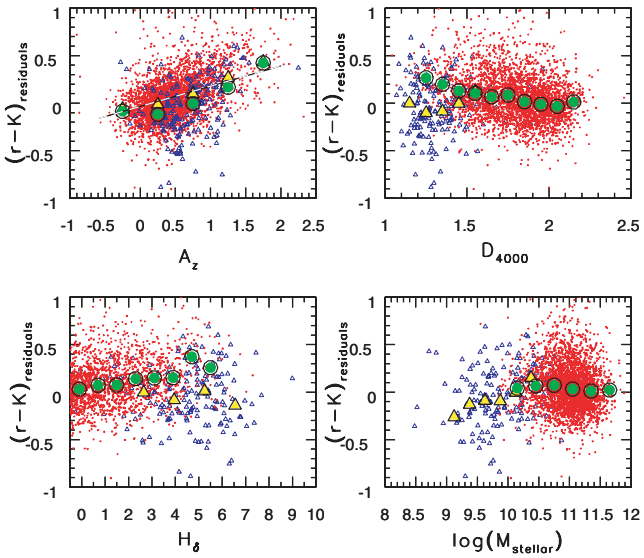


**Figure 31.** The top left-hand panel is analogous to the top panel in Fig. 30, except that here AGN (dots and circles) and SFs (small and large triangles) are treated separately. The straight line is the same line as in the top panel in Fig. 30. The other three panels are analogous to the top left-hand panel, except that they show the dependence of residuals on the 4000 Å break ($D_{4000}$), the strength of H$\delta$ line, and stellar mass.

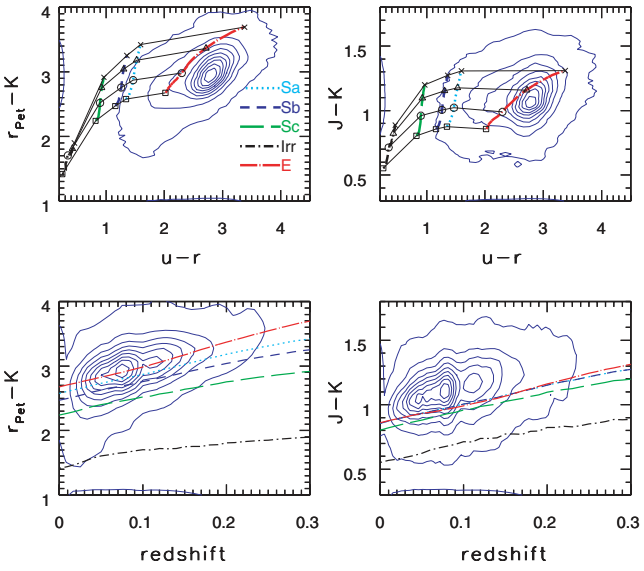


**Figure 32.** A comparison of the observed galaxy distribution (contours) in representative colour–colour and colour–redshift diagrams and the Bruzual & Charlot (1993, 2003) model predictions. The SEDs for five representative synthetic galaxies are convolved with SDSS and 2MASS bandpasses for a grid of redshifts in the 0–0.3 range (thick lines). The positions along these lines that correspond to redshifts of 0, 0.1, 0.2 and 0.3 are marked by squares, circles, triangles and crosses, respectively, and are connected by thin lines.

predict $K$-band magnitudes measured by 2MASS with an rms scatter of only 0.1 mag and the intrinsic astrophysical scatter probably significantly smaller. Within the restricted wavelength range probed by SDSS, this scatter is even smaller. Smolčić et al. (2006) show that the rest-frame $r - i$ colour (5500–8500 Å wavelength range) can be predicted with an rms of only 0.05 mag using Strömgren colours evaluated in the 4000–5800 Å wavelength range. Smolčić et al. (2006) also find a strong correlation between Strömgren colours and the position of a galaxy in the BPT diagram, a result that is confirmed here using the $u - r$ colour. As shown by Smolčić et al. (2006), other parameters, such as those determined by Kauffmann et al. (2003a), can also be used to parametrize the position of a galaxy on this one-dimensional sequence in the multidimensional colour space. Or equivalently, using the terminology from Yip et al. (2004), most of the variance in galaxy SEDs is already absorbed in the first few principal components.

We integrate the broad-band UV-to-NIR SEDs of dominant galaxy types and find that the $z$-band flux is the closest, colour-independent, proxy for bolometric flux measurement in the 0.2–2.2 μm range for galaxies with redshifts smaller than ∼0.2.

We find that galaxies detected by GALEX include a non-negligible fraction (10–30 per cent) of AGNs, and hence do not represent a clean sample of starburst galaxies. This conclusion is supported by their $u - r$ colour distribution, position in the BPT diagram, and morphological appearance (see Agüeros et al. 2005, for more details).

We demonstrate that interstellar dust content inferred from optical spectra by Kauffmann et al. (2003a) is indeed higher for galaxies detected by *IRAS* confirming the reliability of the $A_z$ measurement. This represents a dramatic *independent* support for the notion that these model-based *optical* estimates of $A_z$ are related to the galaxy dust content. Furthermore, $A_z$ can be used with the $u$-band measurements to predict *IRAS* 60-μm flux within a factor of ∼2, which suggests that the Bruzual & Charlot (1993, 2003) models used to derive $A_z$ provide fairly good description of the relevant physics. Even more detailed and robust analysis along these lines will be possible with the advent of GALEX and Spitzer data.

We find that SFs tend to be bluer than AGN galaxies for all photometric bands bluewards from the $K$ band, while they have redder far-IR-to-optical colours. We conclude that SFs have redder far-IR-to-optical colours than AGN galaxies because they have more UV light that is processed to far-IR range, as a fraction of the 0.2–2.2 μm bolometric flux. We emphasize that a variety of different data for the *same* galaxy sample was required to reach this conclusion: GALEX, SDSS, 2MASS and *IRAS* photometry, as well as the interstellar dust content inferred from SDSS optical spectra with the aid of sophisticated models.

A large sample of galaxies that have SDSS spectra and are detected by *IRAS* and NVSS allowed us to study the IR–radio correlation separately for star-forming and AGN galaxies. We confirm that both galaxy types follow a tight correlation, and find that a large fraction (80 per cent) of optically classified SDSS–NVSS AGN galaxies show significantly more radio emission than expected from their IR flux (technically, we show that IR fluxes predicted from observed radio emission for galaxies that are not detected by *IRAS* are higher than the corresponding *IRAS* upper flux limits, see Section 4.2.3). We also find marginal evidence for different slopes of IR–radio correlation for AGN and star-forming subsamples, an effect that seems to be related to the H$\alpha$/H$\beta$ line-strength ratio.

Perhaps the most important conclusion of our study is that little more than a single datum can be learned about galaxies from photometric data whose accuracy is not demonstrably better than ∼0.1 mag. Fortunately, all three major modern galaxy surveys, GALEX, SDSS and 2MASS, appear to have achieved this goal and, together with surveys such as FIRST and NVSS, opened unprecedented opportunities for detailed studies of galaxies.

## ACKNOWLEDGMENTS

MO is grateful to Alex Szalay, Mercedes Filho and Peter Barthel for their insights, and to Princeton University and University of Washington for financial support.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the US Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society and the Higher Education Funding Council for England. The SDSS Web Site is http://www.sdss.org/.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, Cambridge University, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPA), the Max-Planck-Institute for Astrophysics (MPIA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory and the University of Washington.

The FIRST Survey is supported in part under the auspices of the Department of Energy by Lawrence Livermore National Laboratory under contract no. W-7405-ENG-48 and the Institute for Geophysics and Planetary Physics.

This publication makes use of data products from the 2MASS, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

The Galaxy Evolution Explorer (GALEX) is a NASA Small Explorer. The mission was developed in cooperation with the Centre National d'Etudes Spatiales of France and the Korean Ministry of Science and Technology.

## REFERENCES

Abazajian K. et al., 2003, AJ, 126, 2081
Agüeros M. A. et al., 2005, AJ, 130, 1022
Anderson S. F. et al., 2003, AJ, 126, 2209
Baldry I. K., Glazebrook K., Brinkmann J., Ivezić Ž., Lupton R. H., Nichol R. C., Szalay A. S., 2004, ApJ, 600, 681
Baldwin J., Phillips M., Terlevich R., 1981, PASP, 93, 5
Becker R. H., White R. L., Helfand D. J., 1995, ApJ, 450, 559
Beichman C. A., Neugebauer G., Habing H. J., Clegg P. E., Chester T. J., 1985, *IRAS* Catalogs and Atlases. US GPO, Washington, DC
Bell E. F., 2003, ApJ, 586, 794
Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, ApJ, 585, L117
Best P. N., Kauffmann G., Heckman T. M., Ivezić Ž., 2005a, MNRAS, 362, 9
Best P. N., Kauffmann G., Heckman T. M., Brinchmann J., Charlot S., Ivezić Ž., White S. D. M., 2005b, MNRAS, 362, 25
Blanton M. R. et al., 2003, ApJ, 592, 819
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Bruzual G., Charlot S., 1993, ApJ, 405, 538
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Buat V. et al., 2005, ApJ, 619, 51
Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245
Chang R., Gallazzi A., Kauffmann G., Charlot S., Ivezić Ž., Brinchmann J., Heckman T. M., 2006, MNRAS, 366, 1717
Condon J. J., Broderick J. J., 1988, AJ, 96, 30
Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, AJ, 115, 1693
Eisenstein D. J. et al., 2003, ApJ, 585, 694
Finlator K. et al., 2000, AJ, 120, 2615
Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, AJ, 111, 1748
Goto T., 2005, MNRAS, 360, 322
Gregory P. C., Scott W. K., Douglas K., Condon J. J., 1996, ApJS, 103, 427
Gunn J. E., Oke J. B., 1975, ApJ, 195, 255
Gunn J. E. et al., 1998, AJ, 116, 3040
Hao L. et al., 2005, AJ, 129, 1783
Heckman T. M., Kauffmann G., Brinchmann J., Charlot S., Tremonti C., White S. D. M., 2004, ApJ, 613, 109
Helou G., Soifer B. T., Rowan-Robinson M., 1985, ApJ, 298, 7
Hogg D. W., Finkbeiner D. P., Schlegel D. J., Gunn J. E., 2002, AJ, 122, 2129
Hopkins A. M. et al., 2003, ApJ, 599, 971
Ivezić Ž., Elitzur M., 1997, MNRAS 287, 799
Ivezić Ž., Becker R. H., Blanton M. et al., 2001a, in Green R. F., Khachikian E. Ye., Sanders D. B., eds, IAU Colloq. 184, AGN Surveys. Astron. Soc. Pac., San Francisco, p. 137 (astro-ph/0111024)
Ivezić Ž. et al., 2001b, AJ, 122, 2749
Ivezić Ž. et al., 2002, AJ, 124, 2364
Ivezić Ž. et al., 2004a, in Mújica R., Maiolino R., eds, Multiwavelength AGN Surveys, Guillermo Haro Conference. Conzumel, Mexico, p. 53 (astro-ph/0403314)
Ivezić Ž. et al., 2004b, Astron. Nachr., 325, 583
Jarrett T. H., Chester T., Cutri R., Schneider S., Skrutskie M., Huchra J. P., 2000, AJ, 119, 2498
Kauffmann G. et al., 2003a, MNRAS, 341, 33
Kauffmann G. et al., 2003b, MNRAS, 341, 54
Kauffmann G. et al., 2003c, MNRAS, 346, 1055
Kewley L., Jansen R. A., Geller M. J., 2005, PASP 117, 227
Lupton R. H., Ivezić Ž., Gunn J. E., Knapp G. R., Strauss M. A., Yasuda N., 2002, in Tyson J. A., Wolff S., eds, Proc. SPIE Vol. 4836, Survey and Other Telescope Technologies and Discoveries. SPIE, Bellingham WA, p. 350
McIntosh D. H., Bell E. F., Weinberg M. D., Katz N., 2006, MNRAS, submitted (astro-ph/0511737)
Moustakas J., Kennicutt R. C., Tremonti C. A., 2006, ApJ, 642, 775
Nenkova M., Ivezić Ž., Elitzur M., 2002, ApJ, 570, L9
Pasquali A., Kauffmann G., Heckman T. M., 2005, MNRAS, 361, 1121 (PKH)
Petrosian V., 1976, AJ, 209, L1
Pier J. R., Munn J. A., Hindsley R. B., Hennesy G. S., Kent S. M., Lupton R. H., Ivezić Ž., 2003, AJ, 125, 1559
Rengelink R. B., Tang Y., de Bruyn A. G., Miley G. K., Bremer M. N., Roettgering H. J. A., Bremer M. A. R., 1997, A&AS, 124, 259
Richards G. T. et al., 2002, AJ, 123, 2945
Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
Schmitt H. R., Kinney A. L., Calzetti D., Storchi Bergman T., 1997, AJ, 114, 592
Shimasaku K. et al., 2001, AJ, 122, 1238
Smith J. A. et al., 2002, AJ, 123, 2121
Smolčić V. et al. 2006, MNRAS, submitted
Stoughton C. et al., 2002, AJ, 123, 485
Strateva I. et al., 2001, AJ, 122, 1861
Strauss M. A. et al., 2002, AJ, 124, 1810
Tremonti C. A. et al., 2004, ApJ, 613, 898
van der Kruit P., 1971, A&A, 15, 110
Voges W. H. et al., 1999, A&A, 349, 389
Voges W. H. et al., 2000, IAU Circ., 7432, 1
Yasuda N. et al., 2001, AJ, 122, 1104
Yip C. W. et al., 2004, AJ, 128, 585

York D. G. et al., 2000, AJ, 120, 1579
Yun M. S., Reddy N. A., Condon J. J., 2001, ApJ, 554, 803

## APPENDIX A: TECHNICAL SUMMARY OF THE ANALYSED SURVEYS

### A1 Sloan Digital Sky Survey

SDSS (www.sdss.org) is a digital photometric and spectroscopic survey that will cover one quarter of the celestial sphere in the North Galactic cap and produce a smaller area ($\sim$225 deg$^2$), but much deeper, survey in the Southern Galactic hemisphere (York et al. 2000; Stoughton et al. 2002; Abazajian et al. 2003, and references therein). The flux densities of detected objects are measured almost simultaneously in five bands ($u$, $g$, $r$, $i$ and $z$, Fukugita et al. 1996; Hogg et al. 2002; Smith et al. 2002) with effective wavelengths of 3551, 4686, 6166, 7480 and 8932 Å (Gunn et al. 1998), accurate to 0.02 mag (rms scatter for sources not limited by photon statistics, and also for zero-points, Ivezić et al. 2004b). The survey will result in photometric measurements for close to 100 million stars and a similar number of galaxies. Astrometric positions are accurate to better than 0.1 arcsec per coordinate (rms) for sources with $r < 20.5^{\rm m}$ (Pier et al. 2003), and the morphological information from the images allows reliable star–galaxy separation to $r \sim 21.5$ mag (Lupton et al. 2002). The imaging data are used to select sources for follow-up spectroscopic observations, which will result in over a million spectra. The spectra have a resolution of 1800–2000 in the wavelength range from 3800 to 9200 Å. Extragalactic sources targeted in the SDSS spectroscopic survey include a flux-limited 'main' galaxy sample ($r < 17.77$, Strauss et al. 2002), the luminous red galaxy sample (Eisenstein et al. 2003), and quasars (Richards et al. 2002).

### A2 *ROSAT* survey

The ROentgen SATellit (*ROSAT*, 1990–1999) was an X-ray observatory which included the X-Ray Telescope (XRT) with its 2.4 m focal length mirror assembly consisting of four nested Wolter-I mirrors. The focal plane instrumentation consisted of the Position Sensitive Proportional Counter (PSPC) and the High Resolution Imager (HRI). The Wide Field Camera (WFC) with its 0.525 m focal length mirror assembly consisting of three nested Wolter–Schwarzschild mirrors (co-aligned with the XRT). XRT covered $\sim$6–100 Å ($\sim$2.4–0.12 keV) band, and the WFC covered the $\sim$60–300 Å ($\sim$0.21–0.05 keV) band. *ROSAT* provided a $\sim$2 degree diameter field of view with the PSPC in the focal plane, and $\sim$40 arcmin diameter field of view with the HRI in the focal plane. The main aim of the *ROSAT* mission was the first all-sky survey with imaging X-ray and XUV telescopes; its X-ray sensitivity was about a factor of 1000 higher than that of the UHURU satellite. About 100 000 sources have been detected in the survey, an order of magnitude more than were known before *ROSAT* (Voges et al. 1999, 2000).

### A3 GALEX survey

The Galaxy Evolution Explorer (GALEX) was launched in 2003 April, and will eventually map the entire sky in two bands: the NUV (1750–2800 Å) and the FUV (1350–1750 Å), and to faint flux levels ($m$ = 20–25, AB). GALEX's 0.5-m telescope and 1.°2 field of view will also be used to make deep observations (> tens of ks) of individual interesting fields (such as the Lockman Hole and the *Chandra Deep Field*-South). The primary science goal of the mission is to observe SFs and to track galaxy evolution. The GALEX Early Release Observations used here include three AIS fields (see www.galex.caltech.edu) which overlap with the SDSS footprint.

### A4 2MASS survey

2MASS used two 1.3-m telescopes, one at Mt Hopkins, AZ, and one at CTIO, Chile, to survey the entire sky in NIR light (see www.ipac.caltech.edu/2mass). Each telescope's camera was equipped with three 256 × 256 arrays (the pixel size is 2 arcsec) of HgCdTe detectors which simultaneously observed in the $J$ band (1.25 μm), $H$ band (1.65 μm) and $K_s$ band (2.17 μm). The detectors were sensitive to point sources brighter than about 1 mJy at the 10$\sigma$ level, corresponding to limiting (Vega-based) magnitudes of 15.8, 15.1 and 14.3, respectively. Point-source photometry is repeatable to better than 10 per cent precision at this level, and the astrometric uncertainty for these sources is less than 0.2 arcsec. The 2MASS catalogues contain positional and photometric information for 470 992 970 point sources (2MASS PSC) and 1 647 599 extended sources (2MASS XSC). Details about 2MASS photometry of galaxies can be found in Jarrett et al. (2000).

### A5 *IRAS* Survey

The Infrared Astronomical Satellite (*IRAS*; Beichman et al. 1985) produced an almost all-sky survey ($\sim$98 per cent of the sky) at 12, 25, 60 and 100 μm, with the resulting *IRAS* point-source catalogue (*IRAS* PSC) containing over 250 000 sources, and the Faint Source Catalogue additional 173 000 sources. While the *IRAS* faint limits are of order 1 Jy, it remains a valuable resource due to its important wavelength range and nearly full sky coverage.

### A6 Radio surveys

The basic properties of the radio surveys considered here are as follows.

**GB6:** The Green Bank GB6 survey (GB6; Gregory et al. 1996) is at 4850 MHz (6 cm), with 3-arcmin resolution, and covers the declination band between 0° and 75°. The completeness limit of the GB6 catalogue is 18 mJy, and it includes 75 000 sources.

**NVSS:** The NRAO VLA Sky Survey (NVSS; Condon et al. 1998) is a 1.4-GHz (20-cm) survey with 45-arcsec resolution, and covers the sky north of $-40°$ declination. The completeness limit of the NVSS catalogue is about 2.5 mJy, and it includes 1.8 million sources.

**FIRST:** The Faint Images of the Radio Sky at Twenty-centimetres (FIRST; Becker, White & Helfand 1995) is a 1.4-GHz (20-cm) survey with 5 arcsec resolution, and will cover a quarter of the sky matched to the SDSS footprint. The completeness limit of the FIRST catalogue is 1 mJy, and it will include about 1 million sources. The FIRST survey provides the highest resolution and most accurate radio positions among the large radio surveys. It also has the highest source density of about 90 deg$^{-2}$.

**WENSS:** The Westerbork Northern Sky Survey (WENSS; Rengelink et al. 1997) is a 326-MHz (92-cm) survey with 54 × 54 arcsec$^2$ cos ($\delta$) resolution, and covers the sky north of $+30°$ declination. The completeness limit of the WENSS catalogue is 18 mJy, and it includes 230 000 sources.

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