

The PSCz catalogue

W. Saunders,^{1★} W. J. Sutherland,² S. J. Maddox,³ O. Keeble,⁴ S. J. Oliver,⁴
M. Rowan-Robinson,⁴ R. G. McMahon,³ G. P. Efstathiou,³ H. Tadros,²
S. D. M. White,⁵ C. S. Frenk,⁶ A. Carramiñana⁷ and M. R. S. Hawkins¹

¹*Institute for Astronomy, Blackford Hill, Edinburgh EH9 3HJ*

²*Nuclear and Astrophysics Laboratory, Keble Road, Oxford OX1 3RH*

³*Institute of Astronomy, Madingley Road, Cambridge CB3 0HA*

⁴*Blackett Laboratory, Imperial College, Prince Consort Road, London SW7 2BZ*

⁵*MPI-Astrophysik, Karl-Schwarzschild-Strasse 1, D-85740 Garching bei Munchen, Germany*

⁶*Department of Physics, University of Durham, Durham DH1 3LE*

⁷*Instituto Nacional de Astrofísica Óptica y Electrónica, Apartado Postal 51 y 216, 72000 Puebla, Mexico*

Accepted 2000 March 8. Received 2000 March 8; in original form 1999 October 26

ABSTRACT

We present the catalogue, mask, redshift data and selection function for the PSCz survey of 15 411 *IRAS* galaxies across 84 per cent of the sky. Most of the *IRAS* data are taken from the Point Source Catalog, but this has been supplemented and corrected in various ways to improve the completeness and uniformity. We quantify the known imperfections in the catalogue, and we assess the overall uniformity, completeness and data quality. We find that overall the catalogue is complete and uniform to within a few per cent at high latitudes and 10 per cent at low latitudes. Ancillary information, access details, guidelines and caveats for using the catalogue are given.

Key words: catalogues – surveys – galaxies: clusters: general – galaxies: distances and redshifts – large-scale structure of Universe.

1 INTRODUCTION

Data from *IRAS* (the infrared astronomical satellite) allow unparalleled uniformity, sky coverage and depth for mapping the local galaxy density field. In 1992, with completion of the QDOT and 1.2-Jy surveys (Rowan-Robinson et al. 1990a; Strauss et al. 1990, 1992; Fisher et al. 1995; Lawrence et al. 1999), and with other large redshift surveys in progress, it became clear that a complete redshift survey of the *IRAS* Point Source Catalog (the PSC, Beichman et al. 1988, henceforth ES) had become feasible.

Our specific targets for the PSCz survey were twofold: (i) we wanted to maximize sky coverage in order to predict the gravity field, and (ii) we wanted to obtain the best possible completeness and flux uniformity within well-defined area and redshift ranges, for statistical studies of the *IRAS* galaxy population and its distribution. The availability of digitized optical information allowed us to relax the *IRAS* selection criteria used in the QIGC (Rowan-Robinson et al. 1990b), and use optical identification as an essential part of the selection process. This allowed greater sky coverage, being essentially limited only by requiring that optical extinctions be small enough to allow complete identifications. The PSC was used as starting material, because of its superior sky coverage and treatment of confused and extended sources as compared with the Faint Source Survey (Moshir et al. 1989). The

depth of the survey (0.6 Jy) is derived from the depth to which the PSC is complete over most of the sky.

The topology of the survey is analysed and presented by Canavezes et al. (1998); the inferred velocity field by Branchini et al. (1999); the real-space power spectrum and its distortion in redshift space by Tadros et al. (1999); the redshift-space power spectrum by Sutherland et al. (1999). The direction and convergence of the dipole have been investigated by Rowan-Robinson et al. (2000), and their implications for cosmological models have been investigated by Schmoldt et al. (1999). Sharpe et al. (2000) presented a least-action reconstruction of the local velocity field, while an optical/*IRAS* clustering comparison was presented by Seaborne et al. (1999). Many of these results are summarized in Saunders et al. (1999a).

2 CONSTRUCTION OF THE CATALOGUE

2.1 Sky coverage

We aimed to include in the survey all areas of sky with (i) reliable and complete *IRAS* data, and (ii) optical extinction small enough to allow reliable galaxy identifications and spectroscopic follow-up. We defined a mask from those parts of the sky excluded from the survey; the final mask was constructed from the union of the following areas.

- (i) Areas failing to get two hours-confirming coverages

★ E-mail: will@roe.ac.uk

(HCONs, ES III.C.1). In these areas either no data at all or the data do not allow adequate source confirmation.

(ii) Areas flagged as being of high source density (HSD) at 12, 25 or 60 μm . HSD at 12 or 25 μm implies an impossibly high stellar density for galaxy identifications. Areas flagged as HSD at 60 μm were processed differently in the PSC, with completeness sacrificed for the sake of reliability.

(iii) Areas with $I_{100} > 25 \text{ MJy sr}^{-1}$. The 100- μm intensity values are those of Rowan-Robinson et al. (1991). Above this value, we found the fraction of sources that were identifiable as galaxies dropped dramatically.

(iv) Areas with extinctions $A_B > 2 \text{ mag}$ where secure optical identifications become increasingly difficult. Our extinction estimates were based on the I_{100} values above, and incorporated a simple model for the variation in dust temperature across the Galaxy; they are discussed further in Section 4.3. Except towards the centre and anticentre, this criterion and the previous one are almost equivalent.

(v) Small patches covering the Large Magellanic Cloud (LMC) and Small Magellanic Clouds (SMC), defined by $I_{100} > 10$ and 5 MJy sr^{-1} respectively. In these areas there are large numbers of H II regions in the clouds themselves, with similar optical and *IRAS* properties to the background galaxies, making identifications very uncertain.

The mask is specified as a list of excluded 1-deg² ‘lune-bins’ (ES Ap.X.1), so these values are necessary averages over 1 deg².

The overall area outside the mask is 84 per cent of the sky. For statistical studies of the *IRAS* galaxy population and its distribution, where uniformity is more important than sky coverage, we made a ‘high- l ’ mask, as above but included all areas with $A_B > 1 \text{ mag}$, leaving 72 per cent of the sky. In practice this criterion is almost identical to $I_{100} > 12.5 \text{ MJy sr}^{-1}$. Henceforth, when we say ‘high latitude’, we simply mean outside this mask. Both masks are shown in Fig. 1.

2.2 PSC selection criteria

Our aim was to relax the criteria of the QIGC sufficiently to detect

virtually all galaxies, even at low latitude, purely from their *IRAS* properties; simultaneously, we wanted to keep contamination by Galactic sources to a reasonable level. Our actual colour selection criteria were almost the same as the QIGC:

$$\log_{10}(S_{60}/S_{25}) > -0.3,$$

$$\log_{10}(S_{25}/S_{12}) < 1.0,$$

$$\log_{10}(S_{100}/S_{25}) > -0.3,$$

$$\log_{10}(S_{60}/S_{12}) > 0.0,$$

$$\log_{10}(S_{100}/S_{60}) < 0.75.$$

However, unlike the QIGC, upper limits were used only where they guaranteed inclusion or exclusion. We did not at this stage exclude any source solely on the basis of an identification with a Galactic source. The PSC_z evolved from the QIGC, and consequently three galaxies remain in the catalogue that actually fail the new selection criteria. In total, we selected 16 422 sources from the PSC.

Many very nearby galaxies have multiple PSC sources, associated with individual star-forming regions. 70 such sources were pruned to leave, for each such galaxy, the single brightest PSC source. Also at this stage, Local Group galaxies were excised from the catalogue, and a separate catalogue of far-infrared properties for Local Group galaxies was made.

2.3 Extended sources

All *IRAS* surveys are bedevilled by the question of how to deal with galaxies that are multiple or extended with respect to the *IRAS* 60- μm beam. The size of most 60- μm detectors was 1.5 arcmin in-scan by 4.75 arcmin cross-scan. The raw data were taken every 0.5 arcmin, and for the PSC this was then filtered with an eight-point zero-sum linear filter, of the form $--++++--$. Hence, the fluxes of galaxies with in-scan far-infrared sizes larger than about 1.5 arcmin will be underestimated in the PSC. There is a lesser sensitivity to the cross-scan diameter, caused by some of the scans only partially crossing the full-width extended sources.

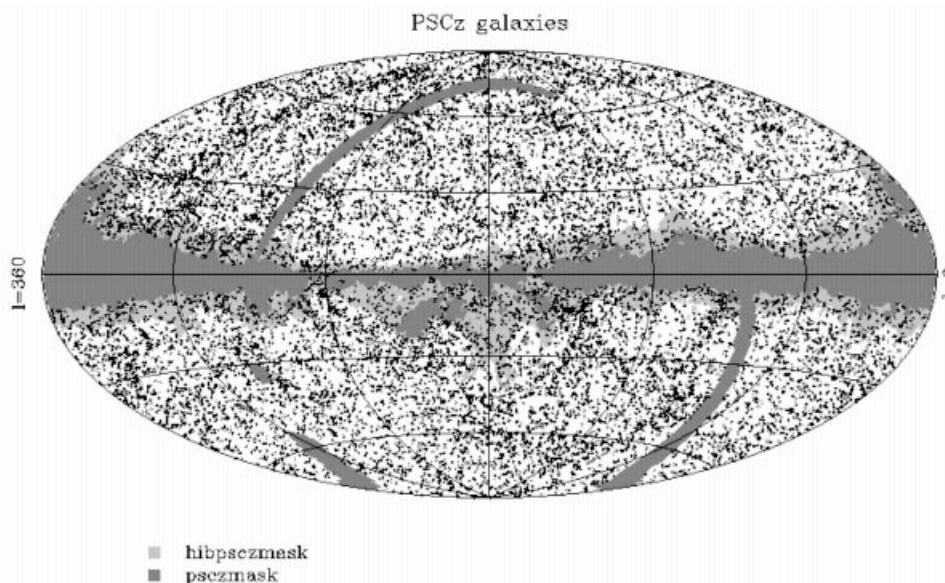


Figure 1. 15 431 PSC_z sources identified as possible or definite galaxies, high l and default masks in galactic coordinates.

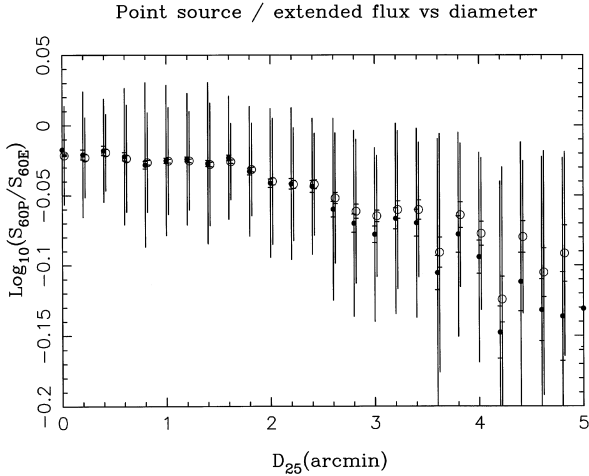


Figure 2. Ratio of extended flux to PSC (filled circles) or point-source-filtered ADDSCAN flux (open circles), as a function of optical diameter. The open circles are displaced slightly to the right for clarity. The inner error bars are the uncertainty in the ratio, the outer error bars are the population scatter. Where D_{25} is not known it is set to 0.

The approach we have settled on is to use PSC fluxes preferentially, except for sources identified with individual galaxies with large diameters that are likely to lead to significant flux underestimation in the PSC. The PSC flux is based on a least chi-squared template fit to the point-source-filtered data stream, so that the flux is correctly measured for slightly extended sources with diameters smaller than about 1.5 arcmin.¹ As far-infrared diameters are typically half of those in the optical (e.g. Rice et al. 1988), and optical D_{25} diameters are typically several times the FWHM for spirals, we can expect that isolated galaxies must have extinction-corrected diameters larger than several arcmin for their fluxes to be badly underestimated by the PSC.

Fluxes for extended sources can be derived using the ADDSCAN (or SCANPI) software provided at IPAC, which co-adds all detector scans passing over a given position. We have tested the way in which the ratio of PSC to extended flux depends on D_{25} , and, as expected, we find that the ratio is almost constant for galaxies up to 2.5 arcmin (Fig. 2), where the ADDSCAN flux is typically 10 per cent larger. However, we also find that ADDSCAN fluxes are systematically 5 per cent larger than PSC fluxes, even for much smaller galaxies. This discrepancy is not caused by any difference in calibration (we have tested this by extracting a point-source-filtered flux from the ADDSCAN data and substituting this for the PSC flux, but the discrepancy in Fig. 2 remains). The reason is the large number of multiple and/or interacting galaxies seen by *IRAS*, for which the ADDSCAN picks up the entire combined flux. Using such combined fluxes throughout is not a consistent approach, as nearby interacting galaxies will always be resolved into two or more sources, while more distant ones will not be.² The effect of the PSC filter is more subtle: a close pair of galaxies unresolved by the *IRAS* beam will in general have unequal fluxes, and in a

¹By comparison, the Faint Source Survey (Moshir et al. 1989) fluxes are peak amplitudes of the point-source-filtered data, so are much less tolerant of slightly extended sources.

²In any survey with unresolved sources, a consistent approach to multiple sources is not possible, except by artificially degrading the resolution for each source to make it the same in physical units as for the most distant source.

flux-limited survey at least one will be below the flux limit. In such a situation the zero-sum linear PSC filter still overestimates the flux of the brightest source on average (because of clustering), but by much less than the ADDSCAN.

Very large galaxies will have fluxes underestimated by the ADDSCAN because their cross-scan extent is large compared with the detector size. Fluxes for galaxies with optical diameters $D_{25} > 8$ arcmin have been taken from Rice et al. (1988), who made co-added maps for each such galaxy. In principal, galaxies somewhat smaller than this may have underestimated fluxes from the ADDSCANS. However, we find that, for the 19 PSCz galaxies in Rice et al. with quoted diameters 8–10 arcmin, the flux difference between ADDSCAN and Rice et al. fluxes is $\log_{10}(S_{60A}/S_{60R}) = 0.044 \pm 0.034$, that is the ADDSCAN fluxes are slightly, but not significantly, larger; we therefore do not expect smaller galaxies to have underestimated ADDSCAN fluxes.

Bearing these points in mind, we proceeded as follows: we constructed a catalogue, with the same sky coverage as the PSCz, of optically selected galaxies from the Lyon–Meudon Extragalactic Data Base (LEDA) (Paturel 1989) with extinction-corrected D_{25} diameters larger than 2.25 arcmin, where the extinctions were estimated as will be described in Section 4.3, and the consequent corrections to the diameters are as given by Cameron (1990). Saunders et al. (1995) argued that LEDA is reasonably complete to this limit. For the largest sources, we used the positions and fluxes of Rice et al. (1988). IPAC kindly ADDSCANNED all the remaining sources for us to provide co-added data. We then used software supplied by Amos Yahil to extract positions, fluxes and diameters from this data, on the assumption that the galaxies have exponential profiles. Where these ADDSCAN fluxes are used in the catalogue, they have been arbitrarily decreased by 10 per cent to bring them statistically into line with the PSC fluxes at the 2.25-arcmin switchover. We ended up with 1402 sources associated with large galaxies entering the catalogue, and 1290 PSC sources associated with large galaxies flagged for deletion. The latter are kept in the catalogue so that a purely *IRAS*-derived catalogue can be extracted if required.

2.4 Other problem sources

In the QIGC, sources flagged as associated with confirmed extended sources or a poor correlation coefficient (CC), or flagged as confused, were ADDSCANNED. The above prescription should have dealt with the extended sources; the only other reasons for a galaxy source to have a poor CC are that (i) it is multiple, and as argued above we preferred PSC fluxes in these cases, or (ii) it has low signal-to-noise ratio (S/N), in which case the PSC is an unbiased flux estimator, while ADDSCANNING risks noise-dependent biases. The confusion flag is set very conservatively (i.e. most sources flagged as being confused have sensible PSC fluxes), and in any case the PSC deals better, in general, with confusion than the ADDSCANS in both cross-scan and in-scan directions. For these reasons, and to avoid introducing any latitude-dependent biases into the catalogue, we opted to use PSC fluxes throughout for sources not identified with optical galaxies as above. ADDSCAN fluxes are included in the catalogue for all sources for those wishing to experiment with them.

2.5 Supplementary sources

IRAS scanned the sky in a series of hours–confirming coverages (HCONs). Sources needed at least two hours confirmed detections

(also HCONs) to enter the PSC (ES V.D.6). Thus, at low S/N, areas of the sky with only two HCONs are inevitably less complete than those with three or more; the estimated completeness at 0.6–0.65 Jy in 2-HCON areas is only 82 per cent (ES XII.A.4). To improve this completeness, we supplemented the catalogue with 1-HCON sources satisfying our colour criteria in the Point Source Catalog reject file (ES VII.E.1), where there is a corresponding source in the Faint Source Survey (FSS). We demanded that the PSC and FSS sources be within each other's 2σ error ellipse, and that the fluxes agree to within a factor of 1.5. This revealed many sources in the reject file where two individual HCON detections had failed to be merged in the PSC processing (ES XII.A.3), as well as sources that failed at least one HCON for whatever reason. New sources were created or merged with existing ones, and the flux overestimation parameters (ES XII.A.1) assigned or recalculated accordingly. As well as these 1-HCON sources, we also searched in the reject file for additional sources with flux quality flags 1122 and 1121,³ neither of which had been included in the PSC. Finally, we looked for sources in the PSC itself with flux quality flags 1113, but with a correlation coefficient at 60 μm equal to A, B or C indicating a meaningful detection. Altogether we found an additional 323 galaxies, mostly in 2-HCON regions, and made 143 deletions as a result of merging individual HCON detections.

2.6 Optical identifications

Optical material for virtually all sources was obtained from COSMOS or APM scans, including new APM scans taken of 150 low-latitude POSS I *E* plates. In general, we used red plates at $|b| < 10^\circ$ and blue otherwise. The actual categories of optical material are the following.

(i) Uncalibrated POSS *O* plates scanned with APM. Here the magnitudes are typically good to 0.5 mag for faint ($B \sim 19$ mag) galaxies.

(ii) SRC *J* plates scanned with APM. These plates were scanned and matched for the APM survey, and have b_J mag typically good to 0.25 mag for faint galaxies.

(iii,iv,v) Uncalibrated SRC *J* or *EJ* plates scanned with COSMOS, giving b_J accurate to about 0.5 mag.

(vi) SRC *SR* plates scanned with COSMOS. The quoted magnitudes are r magnitudes, good to 0.5 mag.

(vii) POSS *E* plates scanned with APM. For these plates, the standard POSS *O* calibration was assumed. Since the *O* plates are about 1 mag deeper than the *E* plates, while $B - R \sim 1$ for *IRAS* galaxies, this gives a reasonable approximation of the *B* magnitude. When there is extinction, this procedure gives systematically too bright a magnitude by about $0.43A_B$ (assuming Mathis 1990). The scatter is still dominated by the 0.5-mag zero-point error.

At bright magnitudes, photographic photometry for galaxies becomes increasingly uncertain because of non-linearity and saturation in the emulsion.

The image parameters from COSMOS and APM data were also used to obtain arcsec positions and offsets from nearby stars, and to make 'cartoon' representations of the 4×4 arcmin² field centred on each *IRAS* source. The identifications were made using the likelihood methods of Sutherland & Saunders (1992). In general, and always at low latitudes, these cartoons were supplemented by

grey-scale images from the Digital Sky Survey, APM scans and/or inspection of copy plates, and any change in the best identification was noted.

At this stage, non-galaxies were weeded out by a combination of optical appearance, *IRAS* colours and ADDSCAN profiles, VLA 20-cm maps from the NVSS survey (Condon et al. 1998), millimetre data from Wouterlout & Brand (1989), SIMBAD and other literature data. Several hundred low-latitude sources were also imaged at K' , as part of the extension of the PSC_z to lower latitudes (Saunders et al. 1999b). We found a total of 1376 confirmed non-galaxies. We are left with 15 332 confirmed galaxies, and a further 79 sources where no optical identification is known but there is no clear Galactic identification either. The distribution of identified galaxies and the mask are shown in Fig. 1.

The Galactic sources are dominated by infrared cirrus (774 sources). However, there are also 88 planetary nebulae, 140 emission-line stars, 24 sources identified with Galactic H II regions, 212 sources identified with bright stars or reflection nebulae and 138 sources associated with young stellar objects.

3 THE REDSHIFT SURVEY

An essential part of the project was the maintenance of a large data base of redshift information from the literature, from data bases such as the NASA Extragalactic Data Base (NED), LEDA and Huchra's ZCAT, and also from work in progress for other surveys. Redshifts were accepted on the basis that their claimed error was better than any other available for that source, including our own.

Of the 15411 galaxies in the sample, about 8000 had known redshift at the inception of the project, with about another 2000 expected to be observed as part of ongoing projects such as the CfA2 and SSRS surveys. The project was fortunate enough to be allocated six weeks of INT + FOS time, one week INT + IDS time, six nights AAT + FORS time, 18 nights on the CTIO 1.5-m telescope, two weeks at the INAOE 2.1-m telescope and 120 h at the Nançay radio telescopes, over a total of four years. 4600 redshifts were obtained in this time. For the INT + IDS and CTIO spectra, redshift determinations were made from the average observed wavelengths of the H α , N II and S II features. The rest of the data were taken with low dispersion spectrographs with

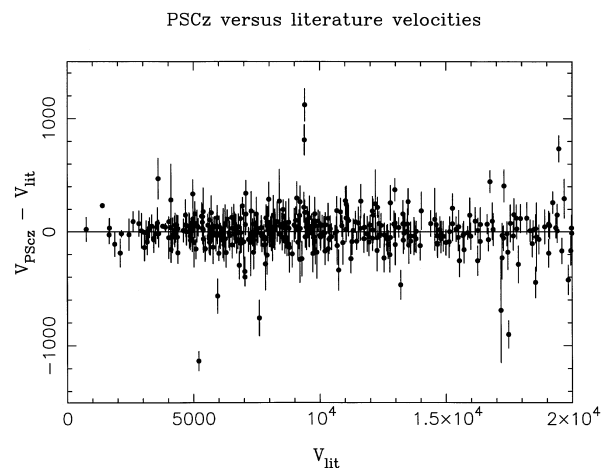


Figure 3. PSC_z versus literature redshifts. Error bars show combined errors.

³ 1 – upper limit only, 2 – moderate quality detection, 3 – good detection, in each of the four bands.

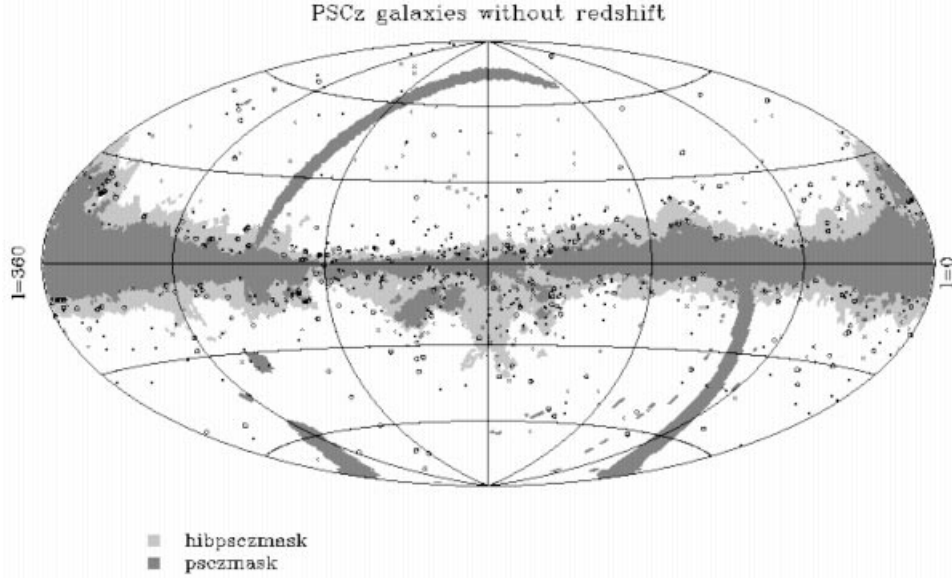


Figure 4. Sky distribution of PSCz sources without redshift, with galaxy identifications $b_J < 19.5$ mag (open circles, 192 sources) and fainter than this or with no identification (crosses, 542 sources).

resolutions of around $15\text{--}20 \text{ \AA}$, and the $\text{H}\alpha/\text{N II}$ lines are blended together, as are the two S II lines. We modelled each continuum-subtracted spectrum as a linear combination of $\text{H}\alpha$, N II and S II features, with the redshift as a free parameter. The model giving the smallest χ^2 when compared with the data gave the redshift, its error, and also the various line strengths and a goodness of fit. Spectra with poor χ^2 , large redshift uncertainty or unphysical line ratios were checked by hand and where necessary refitted with different FWHM or initial guesses. Wavelength calibration included cross-correlation of each sky spectrum with a well-calibrated template. The final derived errors average 120 km s^{-1} , as compared with 300 km s^{-1} for the QDOT survey. Further details are presented in Keeble (1996).

By construction, we have minimized the overlap with other redshift surveys, but there are to date 448 galaxies for which we have both our own measurement and higher resolution measurements by other workers (Fig. 3). There are 31 sources where the velocity difference is more than 3σ (errors combined in quadrature); if these are clipped out the remaining 417 sources give an average offset $V_{\text{PSCz}} - V_{\text{lit}} = -6.1 \pm 6.1 \text{ km s}^{-1}$, with a scatter of 124 km s^{-1} and a reduced $\chi^2_{\nu} = 0.89$. The scatter includes the contribution from the non-PSCz redshift, and suggests that our quoted errors are good estimates of the genuine external error. Of the 31 discrepant redshifts, five involve H I observations with unrealistically small quoted errors and with absolute differences less than 250 km s^{-1} , a further 11 are between 100 and 600 km s^{-1} and $3\text{--}5\sigma$, and suggest a non-Gaussian error distribution, and 15 are greater than 700 km s^{-1} and 5σ and must therefore represent different galaxies or incorrect line identification or calibration. We thus have 15 redshifts seriously in error out of a total of 896 recent measurements (both ours and other workers), giving an error rate of 1.7 per cent.

We did not pursue very faint ($b_J > 19.5$ mag) galaxies, on the basis that (i) they were certain to be at distances too large to be included for either statistical or dynamical studies, and (ii) very large amounts of telescope time would be needed to achieve useful completeness for these galaxies. There are 438 sources without known redshift with a clear or probable faint optical galaxy

$n(z)/\text{ster}$ for high $|b|$ PSCz

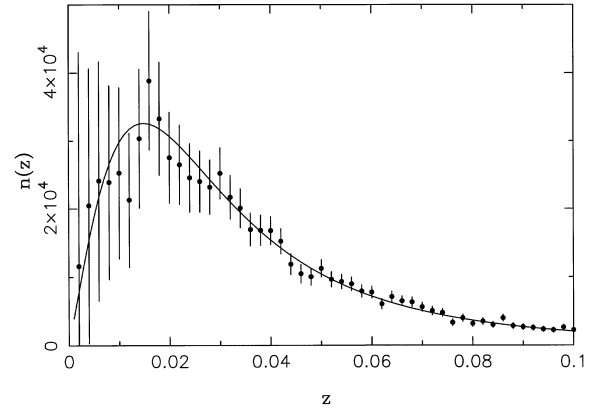


Figure 5. $n(z)$ distribution for the high-latitude PSCz survey. Error bars are J_3 -weighted. The line is the prediction from the selection function.

identification (with $b_J > 19.5$ mag), and a further 127 with no obvious optical identification but no secure identification as a Galactic source either. At the time of writing we are still lacking redshifts for 189 brighter ($b_J < 19.5$ mag) galaxies. The sky distribution of these categories is shown in Fig. 4. Redshifts are available for 14 677 sources.

3.1 $n(z)$ and selection function

The $n(z)$ distribution is shown in Fig. 5. The source density amounts to $1460 \text{ galaxy sr}^{-1}$ at high latitudes and the median redshift is 8500 km s^{-1} . To account for the effect of the flux limit on the observed number density of galaxies as a function of distance, we need to know the *selection function* $\psi(r)$, here defined as the expected number density of galaxies in the survey as a function of distance in the absence of clustering. We have derived this both non-parametrically and parametrically using the methods of Mann, Saunders & Taylor (1996). This method is almost completely insensitive to the assumed cosmology, in the

sense that the derived expected $n(z)$ is invariant. The resulting selection function can be transformed to other cosmologies or definitions of distance simply by mapping the volume element or distance, keeping $n(z) dz$ invariant.

For simplicity, and to allow comparison with simulations, we assume for derivation purposes a Euclidean universe without relativistic effects, and with distance r equal to $V/100 h \text{ km s}^{-1}$. We derive the result both non-parametrically and parametrically using the double power-law form:

$$\psi(r) = \psi_* \left(\frac{r}{r_*}\right)^{1-\alpha} \left[1 + \left(\frac{r}{r_*}\right)^\gamma\right]^{-(\beta/\gamma)}, \quad (1)$$

which describes the non-parametric results very well. The parameters ψ , α , r_* , γ and β respectively describe the normalization, the nearby slope, the break distance in $h^{-1} \text{ Mpc}$, its sharpness and the additional slope beyond it.

Using the high-latitude PSCz to derive the selection function, and correcting redshifts only for our motion with respect to the centroid of the local group, $V = V_{\text{hel}} + 300 \sin l \cos b$, we obtain

$$\psi_* = 0.0077, \quad \alpha = 1.82, \quad r_* = 86.4, \quad \gamma = 1.56, \quad \beta = 4.43.$$

Both non-parametric and parametric results are shown in Fig. 6. The uncertainty in the selection function is less than 5 per cent for distances of 30–200 $h^{-1} \text{ Mpc}$, and 10 per cent for 10–300 $h^{-1} \text{ Mpc}$, although the selection function drops by four decades over this range.

In Fig. 6, we also show the non-parametric selection function derived by the same method for the QDOT and 1.2-Jy surveys.

4 RELIABILITY, COMPLETENESS, UNIFORMITY AND FLUX ACCURACY

The utility of the PSCz for cosmological investigation depends on its reliability, completeness and uniformity.

4.1 Reliability

The major sources of unreliability in the catalogue are (i) incorrect identifications with galaxies nearby in angular position, and (ii) incorrect identification of spectral features. Later cross-referencing used increasingly sophisticated methods based on Sutherland & Saunders (1992), but many older identifications depend on simple 2-arcmin proximity. The most recent cross-correlation with ZCAT provided 1975 updated velocities, of which 84 were altered by more than 500 km s^{-1} and 50 by more than 1000 km s^{-1} , suggesting an error rate of 2 per cent, in agreement with the analysis of Section 3.

The identifications made for our own redshift follow-up were performed much more carefully than those taken from the literature, using a likelihood approach including full knowledge of the error ellipse, optical galaxy source counts etc., and checked by eye from a grey-scale image. The number of ambiguous identifications is negligible, though the PSC source is often confused between two galaxies separated by a few arcmin in the cross-scan direction. Where there was real ambiguity, we took redshifts for both galaxies and used the one with larger $H\alpha$ flux. It is also possible that Galactic sources were occasionally identified as background galaxies, but we believe this to be very rare.

Selection function for PSCz, QDOT and 1.2Jy surveys

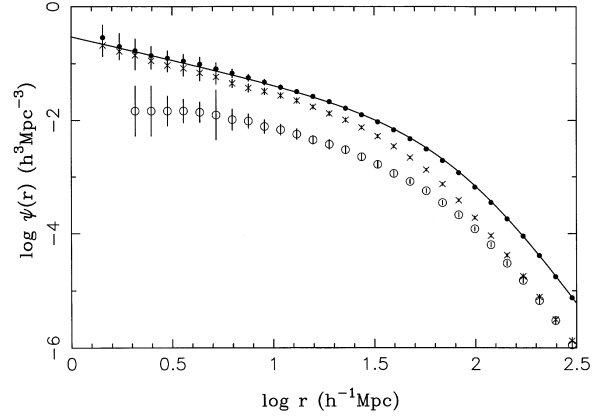


Figure 6. Parametric and non-parametric selection functions for the high-latitude PSCz survey (line and filled circles). Also shown are the QDOT (open circles) and 1.2-Jy (crosses) non-parametric selection functions.

PSCz Source Counts

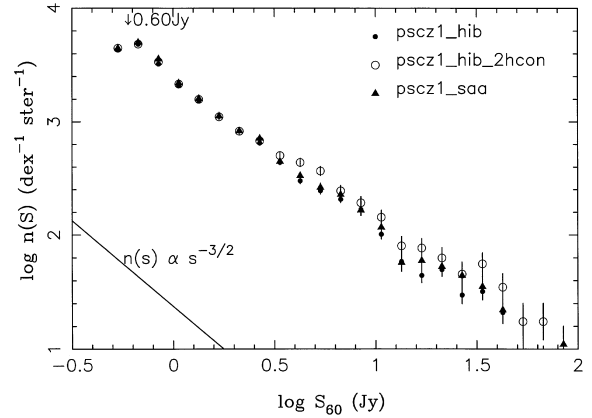


Figure 7. Source counts for all high-latitude sky (filled circles), high-latitude 2-HCON sky (open circles) and an area with potential SAA problems (triangles).

Galaxy source counts as function of I_{100}

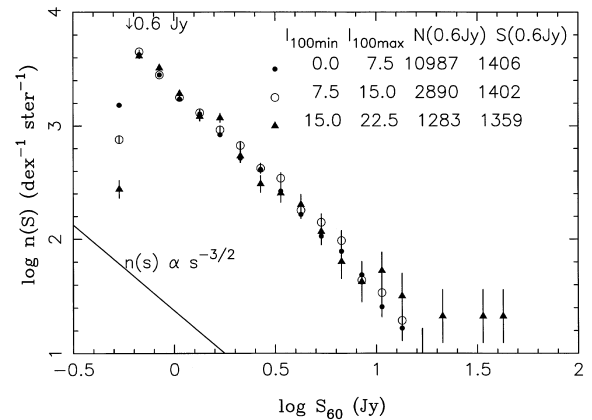


Figure 8. Source counts for identified PSCz galaxies as a function of 100- μm background, $I_{100} < 7.5 \text{ MJy sr}^{-1}$ (filled circles), 7.5–15 (open circles) and 15–22.5 (triangles).

4.2 Completeness

Incompleteness enters into the catalogue for many reasons.

(i) There will be *IRAS* incompleteness where sources with true 60- μm flux greater than 0.6 Jy fail to appear in the catalogue. The largest source of this incompleteness is the areas of sky with only two HCONs, where the PSC incompleteness is estimated as 20 per cent (differential) and 5 per cent (cumulative) at 0.6 Jy (ES XII.A.4). Based on source counts, we estimate that our recovery procedure (Section 2.5) for these sources has reduced the overall incompleteness in these areas from 5 to 1.5 per cent; Fig. 7 shows the source counts for the high-latitude 2-HCON sky. Our recovery procedure was impossible for $|b| < 10^\circ$, so the lower latitude 2-HCON sky (2 per cent of the catalogue area) retains the PSC incompleteness.

(ii) The PSC is confusion-limited in the Galactic plane. However, by construction, the PSCz does not cover any areas with high source density at 60 μm (as defined in ES V.H.6), and the mask defined at 100 μm effectively limits the number of confusing sources. Also, at low latitudes, there are known problems in the PSC with noise lagging. However, the source counts shown in Fig. 8 show low-latitude incompleteness to be small down to 0.6 Jy.

(iii) Some galaxies are excluded by our colour criteria. Cool, nearby galaxies may fail the 100/60 μm condition, but will normally be included separately as extended sources. From the comparison with the 1.2-Jy survey, we estimate that about 50 galaxies from the PSC have been excluded (see Section 5).

(iv) No attempt was made to obtain redshifts systematically for galaxies with $b_J > 19.5$ mag. At high latitudes, the work of e.g. Clements et al. (1996) shows that these will in general be further than $z > 0.1$. At lower latitudes, incompleteness cuts in at lower redshifts. It was originally hoped that the PSCz would be everywhere complete to $z = 0.05$, but it is clear from the 3D density distributions that there is significant (10 per cent or more) incompleteness down to $z = 0.04$ towards the anticentre. The reasons for this are discussed in the next subsection.

Patchy extinction can lead to higher local extinction than our values, which are averages over 1-deg² bins. We have obtained K' images of most low-latitude sources without obvious galaxy or Galactic identification; many faint galaxies are revealed, but only a handful of nearby ones. Spectroscopy of the missing low-latitude galaxies is continuing as part of the behind-the-plane extension to the PSCz survey; for the time being, we estimate that good completeness has been achieved for $z < 0.1/10^{[0.2(A_B + A_B^2/10)]}$. The incompleteness is strongly concentrated towards the anticentre and at low latitudes; for $|b| > 10^\circ$, the survey is estimated to be useably complete to $z = 0.05$ everywhere.

The number of galaxies with $b_J < 19.5$ mag for which redshifts are still unknown is 192, or 1.2 per cent of the sample. A further 85 have only marginal redshift determinations and some of these will be incorrect. Both unknown and marginal redshifts are well distributed round the sky, but with some preference for lower latitudes.

4.3 Extinction maps

Our extinction maps started from the I_{100} intensity maps of Rowan-Robinson et al. (1991), binned into lune bins, with the I_{100}/A_V ratio as given by Boulanger & Perault (1988), and the A_B/A_V ratio as given by Mathis (1990). The dust temperature declines

with galactic radius and consequently this procedure overestimates extinctions towards the Galactic Centre, and underestimates them towards the anticentre. We attempted to correct for this dependence of dust temperature on position by devising a simple model for the dust and starlight in our Galaxy. We assumed a doubly exponential, optically thin distribution of dust and stars, with standard IAU scalelengths for the solar radius and stellar distribution ($r_0 = 8.5$, $r_{sc} = 3.5$ and $z_{sc} = 0.35$ kpc), and assumed the dust to have the same radial scalelength and half the vertical scalelength. The dust properties were assumed to be uniform, and we assumed a dust temperature of 20 K at the solar radius, and proportional to the one fifth power of the local stellar density elsewhere. We then found, for each position on the celestial sphere, the ratio of 100- μm emission to column density of dust, and normalized this ratio by the Boulanger & Perault value for the NGP.

Subsequent to our definition of the PSCz catalogue and mask, Schlegel, Finkbeiner & Davis (1998) used the *IRAS* ISSA and *COBE* DIRBE data to make beautiful high-resolution maps of the dust emission and extinction in our Galaxy. Comparison of our own maps and those of Schlegel et al. show that (i) our temperature corrections across the Galaxy are too small, and (ii) there are several areas where cooler (16-K) dust extends to $|b| \sim 30^\circ$.

Overall, our extinction estimates may be in error by a factor of 1.5–2, in the sense of being too low towards the anticentre and too high towards the Galactic Centre. This has two principal effects.

(i) Towards the anticentre, galaxies may fall below the $b_J = 19.5$ mag limit because of extinction. A subsequent programme of K' imaging has revealed a handful of extra nearby galaxies, and a significant number at redshifts of 0.05 and greater.

(ii) The definition of the optically selected catalogue in Section 2.3 depends on the extinction corrections. We will have selected too many optical galaxies towards the Centre and too few towards the anticentre.

However, our matching of ADDSCAN and PSC fluxes was designed to be robust to errors exactly of this type. To test for any effect, we have compared the simple, number-weighted dipole of the surface distribution of PSCz galaxies: (i) using the normal catalogue and (ii) using purely PSC-derived fluxes as described in Section 2.3. The dipole changes by 1° in direction and 2.5 per cent in amplitude when we do this, showing that any bias caused by incorrect extinctions is negligible.

4.4 Flux accuracy and uniformity

The error quoted for PSC 60- μm fluxes for genuine point sources is just 11 per cent (ES VII.D.2). We are more concerned with any non-fractional random error component, as this may lead to Malmquist-type biases. The analysis in Lawrence et al. (1999), based on the 12/60 μm colours of bright stars, finds an absolute error of 0.059 ± 0.007 Jy in addition to a fractional error of 10 per cent. Stars are better point sources than galaxies, so this may underestimate the absolute error for our sample. We have made an independent estimate, by investigating the scatter between our PSC and ADDSCAN fluxes. Of course PSC and ADDSCAN fluxes start from the same raw data, but the processing and background estimation are entirely different, while the actual photon noise is negligible. The absolute component of the scatter is 0.06–0.07 Jy, of which an estimated 20 per cent comes from the error in the ADDSCAN flux – confirming that the Lawrence et al. value is a reasonable estimate of the PSC absolute error component.

This error estimate leads to Malmquist biases in the source densities of the order of 5–6 per cent (differential, at the flux limit) and 2–3 per cent (cumulative) (Murdoch, Crawford & Jauncey 1973). 2-HCON sky may have biases as large as 10 per cent (differential) and 4 per cent (cumulative), so the non-uniformity introduced into the catalogue should be no worse than 5 per cent (differential) and 2 per cent (cumulative). This is borne out by the source counts in Figs 7 and 8, but note that in any noise-limited catalogue such as the PSC, where the noise varies across the catalogue, there will always be a regime where Malmquist biases are masked by incompleteness.

Lawrence et al. (1999) find evidence for slight non-linearity in the PSC flux scale, but this should not affect any analysis except evolutionary studies. The effect on evolution was considered by Saunders et al. (1997).

Non-uniformity can also come about as a result of simple changes in the absolute flux scale. A small fractional error in the calibration will on average lead to an error half again as large in the source density. There are three obvious reasons for flux scale variations.

(i) The absolute calibration of the third HCON was revised by a few per cent after the release of the PSC (and FSS). The effect of this revision on PSC fluxes would be to change those in the 75 per cent of the sky covered by 3 HCONs by 1 per cent.

(ii) Whenever the satellite entered the South Atlantic Anomaly (SAA), radiation hits altered the sensitivity of the detectors, and data taken during such times were discarded (ES III.C.4). However, this leaves the possibility of Malmquist effects, and also data taken near the boundaries of the SAA potentially suffer residual effects. We have investigated this by checking the source counts for declinations $-40^\circ < \delta < 10^\circ$, where data taken are most likely to be affected. We see no evidence for any variation (Fig. 7) above the few per cent level in source counts.

(iii) Whenever the satellite crossed the Galactic plane, or other very bright sources, the detectors suffered from hysteresis. This effect was investigated by Strauss et al. (1990). They found that the likely error is typically less than 1 per cent and always less than 2.2 per cent. This is confirmed by the constancy, to within a few per cent, of our galaxy source counts for identified galaxies in the PSC_z as a function of I_{100} .

Overall, differential source densities across the sky as a result of incompleteness, Malmquist effects and sensitivity variations are not believed to be greater than a few per cent anywhere at high latitudes for $z < 0.1$. Tadros et al. (1999) found an upper limit to the rms amplitude of large-scale, high-latitude spherical harmonic components to the density field of the PSC_z of 3 per cent. As this is close to the expected variations resulting from clustering, the variations arising from non-uniformity in the catalogue must be smaller than this. At lower latitudes, variations are estimated to be no greater than 10 per cent for $z < 0.05$.

5 COMPARISON WITH THE 1.2-JY SURVEY

The 1.2-Jy survey of 5500 *IRAS* galaxies (Fisher et al. 1995) used looser colour criteria than the PSC_z, and as such acts as a valuable check on the efficacy of our selection procedure from the PSC. We find that the 1.2-Jy survey contains 11 galaxies that have been excluded by our selection criteria. It contains a further five galaxies that satisfy the PSC_z criteria, but are not included as a consequence of programming and editing errors, during construction either of the QIGC survey or its various extensions and

supplements to form the PSC_z. Extrapolating these numbers to lower fluxes, we can expect that about 50 PSC galaxies are missing altogether from the catalogue.

Conversely, the conservative selection criteria for the 1.2-Jy survey led to much greater levels of contamination by cirrus and other Galactic sources than in the PSC_z. In the 1.2-Jy survey, these were eliminated by visual inspection of sky survey plates; inevitably real galaxies occasionally got thrown out by mistake, especially at low latitudes. We have found, within the PSC_z area, 110 galaxies that are misclassified as Galactic in the 1.2-Jy survey, and we have obtained redshifts for 67 of them. Most of the remainder are fainter than our $b_j = 19.5$ mag cut-off. There are also known to be, to date, a further 117 galaxies classified as Galactic in the 1.2-Jy survey outside the PSC_z area. These were found as part of the ongoing ‘behind-the-plane’ extension of the PSC_z to lower latitudes described in more detail in Saunders et al. (1999b).

6 ANCILLARY INFORMATION

Along with the PSC data, the catalogue also contains the following information for each source: POSS/SRC plate and position on that plate; the RA, δ , offset, diameters and magnitude of the best major/minor match from the digitized sky survey plates; where available, name, magnitude and diameter from UGC/ESO/MCG, PGC name, de Vaucouleurs type and H I width; most accurate available redshift and our own redshift measurement; classification as galaxy/cirrus/etc; estimated I_{100} and extinction; ADDSCAN fluxes and width when treated as an extended source, and point source filtered ADDSCAN fluxes.

7 ACCESSING THE CATALOGUE

The data are available from the CDS catalogue service (<http://cdsweb.u-strasbg.fr/Cats.html>). Full and short versions of the catalogue, mask files, description files, format statements and notes are also available via the PSC_z web site (<http://www-astro.physics.ox.ac.uk/~wjs/PSCz.html>), or by anonymous ftp from <ftp://ftp-astro.physics.ox.ac.uk/pub/users/wjs/PSCz/>.

ACKNOWLEDGMENTS

The PSC_z survey has only been made possible because of the generous assistance from many people in the astronomical community. We are particularly grateful to John Huchra, Tony Fairall, Karl Fisher, Michael Strauss, Marc Davis, Raj Visvanathan, Luis DaCosta, Riccardo Giovanelli, Nanyao Lu, Carmen Pantoja, Tadafumi Takata, Kouichiro Nakanishi, Toru Yamada, Tim Conrow, Delphine Hardin, Mick Bridgeland, Renee Kraan-Kortweg, Amos Yahil, Ron Beck, Esperanza Carrasco, Pierre Chamaraux, Lucette Bottinelli, Gary Wegner, Roger Clowes and Brent Tully, for the provision of redshifts prior to publication or other software or data. We are also grateful to the staff at IPAC and the INT, AAT, CTIO and INOAE telescopes. We have made very extensive use of the ZCAT, NED, LEDA, SIMBAD, VLA NVSS and STSCI DSS data bases.

REFERENCES

- Beichman C. A., Neugebauer G., Haling H. J., Clegg P. E., Chester T. J., 1988, *IRAS* Catalogs and Atlases, Vol 1: Explanatory Supplement (ES). JPL, Pasadena

- Boulanger F., Perault M., 1988, *ApJ*, 330, 964
 Branchini E. et al., 1999, *MNRAS*, 308, 1
 Cameron L. M., 1990, *A&A*, 233, 16
 Canavezes A. et al., 1998, *MNRAS*, 297, 777
 Clements D. L., Sutherland W. J., Saunders W., Efstathiou G., McMahon R. G., Maddox S. J., Lawrence A., Rowan-Robinson M., 1996, *MNRAS*, 279, 459
 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
 Fisher K. B., Huchra J. P., Strauss M. A., Davis M., Yahil A., Schlegel D., 1995, *ApJS*, 100, 103
 Keeble O. J., 1996, PhD thesis, Imperial College, Univ. London
 Lawrence A. et al., 1999, *MNRAS*, 308, 897
 Mann R. G., Saunders W., Taylor A. N., 1996, *MNRAS*, 279, 636
 Mathis J. S., 1990, *ARA&A*, 28, 37
 Moshir M. et al., 1989, Explanatory Supplement to the *IRAS* Faint Source Survey. JPL, Pasadena
 Murdoch H. S., Crawford D. F., Jauncey D. L., 1973, *ApJ*, 183, 1
 Paturel G., 1989, *A&AS*, 80, 299
 Rice W. L., Lonsdale C. J., Soifer B. T., Neugebauer G., Kopan E. L., Lloyd L. A., de Jong T., Habing H. J., 1988, *ApJS*, 68, 91
 Rowan-Robinson M. et al., 1990a, *MNRAS*, 247, 1
 Rowan-Robinson M., Saunders W., Lawrence A., Leech K. J., 1990b, *MNRAS*, 253, 485
 Rowan-Robinson M., Hughes J., Jones M., Leech K., Veda K., Walker D., 1991, *MNRAS*, 249, 729
 Rowan-Robinson M. et al., 2000, *MNRAS*, 314, 375
 Saunders W. et al., 1995, in Maddox S. J., Aragon-Salamanca A., eds, 35th Herstmonceux Conference. World Scientific Press, Singapore
 Saunders W. et al., 1997, in Manon C. A., Trinh vân Thuan X., Trân Than Vân J., eds, XVIIth Moriond Astrophysics Meeting: Extragalactic Astronomy in the Infrared. Editions Frontières, Gif-sur-Yvette
 Saunders W. et al., 1999a, in Courteau S., Strauss M., Willick J., eds, Towards an Understanding of Cosmic Flows of Large-Scale Structure, Victoria. July 1999, *PASP*, in press (astro-ph/9909191)
 Saunders W. et al., 1999b, in Courteau S., Strauss M., Willick J., eds, Towards an Understanding of Cosmic Flows of Large-Scale Structure, Victoria. July 1999, *PASP*, in press (astro-ph/9909174)
 Schlegel D. J., Finkbeiner D. P., Davis M., 1998, *ApJ*, 500, 525
 Schmoltdt I. et al., 1999, *MNRAS*, 304, 893
 Seaborn M. D. et al., 1999, *MNRAS*, 309, 89
 Sharpe J. et al., 2000, *MNRAS*, submitted
 Strauss M. A., Davis M., Yahil A., Huchra J. P., 1990, *ApJ*, 361, 49
 Strauss M. A., Huchra J. P., Davis M., Yahil A., Fisher K., Tonry J., 1992, *ApJS*, 83, 29
 Sutherland W. J., Saunders W., 1992, *MNRAS*, 259, 413
 Sutherland W. J. et al., 1999, *MNRAS*, 308, 289
 Tadros H. et al., 1999, *MNRAS*, 305, 527
 Wouterlout J. G. A., Brand J., 1989, *A&AS*, 80, 149

This paper has been typeset from a \TeX/L\AA\TeX file prepared by the author.