HAL open science

HELP: the Herschel Extragalactic Legacy Project

R Shirley, K Duncan, M Campos varillas, P Hurley, K Malek, Y Roehlly, M Smith, H Aussel, T Bakx, V Buat, et al.

To cite this version:

R Shirley, K Duncan, M Campos varillas, P Hurley, K Malek, et al.. HELP: the Herschel Extragalactic Legacy Project. Monthly Notices of the Royal Astronomical Society, 2021, 507 (1), pp.129-155. 10.1093/mnras/stab1526 . hal-03517356

HAL Id: hal-03517356

https://hal.science/hal-03517356v1

Submitted on 7 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **507,** 129–155 (2021)
Advance Access publication 2021 June 3

https://doi.org/10.1093/mnras/stab1526

HELP: the *Herschel* Extragalactic Legacy Project

R. Shirley,[1,2,3] K. Duncan,[4,5] M. C. Campos Varillas,[1] P. D. Hurley,[1] K. Małek,[6,7] Y. Roehlly,[1,7] M. W. L. Smith,[8] H. Aussel,[9] T. Bakx,[8,10,11] V. Buat,[7,12] D. Burgarella,[7] N. Christopher,[13]† S. Duivenvoorden,[1] S. Eales,[8] A. Efstathiou,[13] E. A. González Solares,[3] M. Griffin,[8] M. Jarvis,[14,15] B. Lo Faro,[7] L. Marchetti,[15,16,17] I. McCheyne,[1] A. Papadopoulos,[13] K. Penner,[9] E. Pons,[3,7] M. Prescott,[15,18] E. Rigby,[4] H. Rottgering,[4] A. Saxena,[4,19] J. Scudder,[1,20] M. Vaccari,[15,17,18] L. Wang[21,22] and S. J. Oliver[1]★

Affiliations are listed at the end of the paper

Accepted 2021 May 20. Received 2021 May 12; in original form 2021 March 26

ABSTRACT

We present the *Herschel* Extragalactic Legacy Project (HELP). This project collates, curates, homogenizes, and creates derived data products for most of the premium multiwavelength extragalactic data sets. The sky boundaries for the first data release cover 1270 deg^2 defined by the *Herschel* SPIRE extragalactic survey fields; notably the *Herschel* Multi-tiered Extragalactic Survey (HerMES) and the *Herschel* Atlas survey (H-ATLAS). Here, we describe the motivation and principal elements in the design of the project. Guiding principles are transparent or 'open' methodologies with care for reproducibility and identification of provenance. A key element of the design focuses around the homogenization of calibration, meta data, and the provision of information required to define the selection of the data for statistical analysis. We apply probabilistic methods that extract information directly from the images at long wavelengths, exploiting the prior information available at shorter wavelengths and providing full posterior distributions rather than maximum-likelihood estimates and associated uncertainties as in traditional catalogues. With this project definition paper, we provide full access to the first data release of HELP; Data Release 1 (DR1), including a monolithic map of the largest SPIRE extragalactic field at 385 deg^2 and 18 million measurements of PACS and SPIRE fluxes. We also provide tools to access and analyse the full HELP data base. This new data set includes far-infrared photometry, photometric redshifts, and derived physical properties estimated from modelling the spectral energy distributions over the full HELP sky. All the software and data presented is publicly available.

Key words: techniques: photometric – Astronomical databases: catalogues – surveys – galaxies: evolution – infrared: galaxies – submillimetre: galaxies.

1 INTRODUCTION

A fundamental requirement for rigorous testing of any theories of galaxy formation and evolution is a complete statistical audit or census of the stellar content and star-formation rates of galaxies in the Universe at different times and as a function of the mass of the dark matter haloes that host them.

This audit requires many elements. We need unbiased maps of large volumes of the Universe made with telescopes that probe different wavelengths at which different physical processes of interest manifest themselves. We need catalogues of the galaxies contained within these maps with photometry estimated uniformly from field-to-field, from telescope-to-telescope, and from wavelength-to-wavelength. We need to understand the probability of a galaxy of given properties appearing in our data sets. We need the machinery to bring together these various data sets and calculate the 'value-added' physical data of primary interest, e.g. the distances, stellar masses, star-formation rates, active galactic nuclei (AGNs) fractions, and the intrinsic number densities of the different galaxy populations.

For decades many teams have been undertaking ambitious coordinated multiwavelength programmes to study large volumes of the distant Universe. These surveys are becoming sufficiently complete that we are now able to undertake the necessary homogenizing and adding value, and thus provide the first representative and comprehensive census of the galaxy populations in the Universe.

ESA's *Herschel* (Pilbratt et al. 2010) mission has a unique role in these studies, probing the obscured star-formation activity, which at high redshifts forms about 80 per cent of all star formation. The *Herschel* extragalactic surveys were a major goal of *Herschel* and occupied around 10 per cent of the *Herschel* mission.

The *Herschel* Spectral and Photometric Imaging Receiver (SPIRE) instrument is sufficiently sensitive that the images can detect nearly all of the emission making up the Cosmic Infrared Background Radiation (CIRB) (Duivenvoorden et al. 2020), which itself makes up roughly half of the total background radiation from galaxies.

★ E-mail: s.oliver@sussex.ac.uk
† Deceased

© 2021 The Author(s).
Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

However, the large beam size means that the objects that can be clearly seen as individual sources only make up about 15 per cent of the CIRB. The *Herschel* Photoconductor Array Camera and Spectrometer (PACS) instrument complements the SPIRE observations with bands centred at 100 and 160 μm but at lower depths than SPIRE.

A particular focus of HELP is to employ new methods to learn from our large statistically meaningful samples. This requires harnessing the ancillary data and the *Herschel* data to unlock the full information from the *Herschel* images and then make that available as a legacy to the community.

The science possible with the *Herschel* data will be significantly enhanced with ongoing optical, NIR, and radio surveys. The VISTA near-infrared surveys detect the radiation from the old stellar population in galaxies, which accounts for most of the stellar mass, while the radio surveys being carried out over the next few years with LOFAR, MeerKAT, and ASKAP detect radiation associated with the young stellar population and with radioloud AGN.

The challenge for astronomers wishing to exploit these rich data sets is to collate the data, understand the selection effects, and derive physical parameters. Collation of multiwavelength data has been undertaken for very deep surveys over small areas (less than few deg^2) in particular COSMOS (Scoville et al. 2007; Ilbert et al. 2013; Laigle et al. 2016) and ASTRODEEP (Castellano et al. 2016; Merlin et al. 2016) and for wide nearby surveys (over 200–1000 deg^2) especially SDSS (Blanton et al. 2017) and GAMA (Driver et al. 2009, 2011). However, due to size of the data and complexity arising from the variety of observatories used, little concerted effort has been made to assemble the deep surveys over 10–1000 deg^2. These surveys are particularly important as they are large enough to probe representative ranges of environments and to provide large statistical samples to fully explore the range of galaxy phenomena in detail including rare and transitory phenomena.

Dealing with this complexity and volume of data is not trivial. It requires cross-matching hundreds of millions of objects observed at different bands, identifying spurious sources in a robust and reliable manner and this needs to be done consistently across all fields with varying depths and bands. Dealing with such volumes of data is also memory intensive and requires huge computing power to process the resulting far-infrared (FIR) photometry, photometric redshifts, and SED fitting.

This paper presents the *Herschel* Extragalactic Legacy Project (HELP) Data Release 1 (DR1) and details the pipelines and methods used to tackle the fore-mentioned challenges of complexity and volume size inherent to collating large, deep heterogeneous survey data. This paper follows specific HELP papers detailing specific parts of the project (e.g. Hurley et al. 2017; Duncan et al. 2018a, b; Małek et al. 2018; Shirley et al. 2019) and science results using data from DR1 (e.g. Duivenvoorden et al. 2016, 2020; Scudder et al. 2016; Lo Faro et al. 2017; Pearson et al. 2017b, 2018; Buat et al. 2018; Scudder et al. 2018; Donevski et al. 2020; Mountrichas et al. 2021). In Section 2, we define the HELP fields. In Section 3, we describe the overall HELP strategy. In Section 4, we describe the specific work-flow for DR1. In Section 5, we present some statistics of the data release. In Section 6, we discuss the uses of this data set and conclude.

2 THE HELP FIELDS

Many extragalactic surveys from different observatories and at different wavelengths have been coordinated in their planning and execution. Each survey had different motivations and factors constraining their choice of field locations, sizes, and thus their individual footprint on the sky. The superset of all survey footprints would be large and include many areas with only a few data sets. The primary motivation for HELP is the *Herschel* coverage so DR1 is limited to the main wide area extragalactic *Herschel* surveys.

Given that there is no imminent successor to *Herschel* the data from that mission provides a legacy benchmark. Within the *Herschel* observatory the SPIRE instrument (Griffin et al. 2010) mapped larger areas than the PACS instrument (Poglitsch et al. 2010). We thus define the boundaries of the project on the basis of the extragalactic surveys carried out with SPIRE. The specific *Herschel* OBSIDS chosen to define the project are listed in Appendix A. The footprint of these observations is conveniently captured in HEALPIX Multi Order Coverage maps, MOCs (Fernique et al. 2019) which are provided online.[1]

Some basic properties of the fields are tabulated in Table 1 and the footprints are illustrated on a map of the Galactic dust from Planck (Planck Collaboration XVI 2014) in Fig. 1. As expected from the requirement of the infrared surveys and alignment with other multiwavelength surveys, we can see that the HELP fields: avoid the emission of dust from our Galaxy; are distributed in right ascension; have some concentration at the celestial equator; and include fields near both ecliptic poles.

The *Herschel* Multi-tiered Extragalactic Survey (HerMES; Oliver et al. 2012) is a major survey conducted by the *Herschel* mission (Pilbratt et al. 2010) using the SPIRE (Griffin et al. 2010) and PACS (Poglitsch et al. 2010) instruments covering 380 deg^2. A number of important *Herschel* surveys are contained within the footprint of the SPIRE data in HerMES, notably the PACS evolutionary Probe (PEP; Lutz et al. 2011). The largest and shallowest of the HerMES SPIRE tiers is the HerMES Large Mode Survey, HeLMS, which adjoins the 70 deg^2 HerS survey (Viero et al. 2014b) to form the largest contiguous extragalactic SPIRE field, shown in Fig. 2, which we refer to as the *Herschel* Stripe 82 field. The largest SPIRE footprint comes from the *Herschel* Astrophysical Terahertz Large Area Survey' (H-ATLAS; Eales et al. 2010) which comprises 660 deg^2 (Smith et al. 2017). Additional SPIRE coverage comes from: the *Herschel*-AKARI NEP Deep Survey (Pearson et al. 2017a); and the SPIRE coverage of South Pole Telescope deep field (SSDF; Holder et al. 2013); and the SPIRE calibration field in the North Ecliptical Pole.

The multiwavelength data available in these fields is extremely rich. This is important scientifically through providing the key for basic properties of the objects such as their redshift and probing different physical emission processes. The wealth of data is partly because the choice of these fields by the *Herschel* teams was motivated by existing surveys. In addition, new surveys have been carried out through coordination between survey teams and an appreciation of the value of the accumulated data in these fields has encouraged independent surveys. There are also many very large area surveys that overlap with these fields by accident. A primary goal of HELP is to collate these data sets together. The number of overlapping surveys is continually expanding, so the current collation can only be a snapshot.

3 HELP STRATEGY

The area, depth, and wavelength coverage of the data in the surveys within the HELP fields have enormous potential for addressing important scientific questions, particularly addressing the questions

[1] http://hedam.lam.fr/HELP/dataproducts/dmu2/

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

Table 1. Names, locations, and areas of the individual HELP fields in alphabetical order. The total area is 1269.1 deg^2.

Name	RA (deg)	Dec (deg)	RA min (deg)	RA max (deg)	Dec min (deg)	Dec max (deg)	Area (deg^2)
AKARI-NEP	270.0	66.6	264.6	275.3	64.5	68.5	9.2
AKARI-SEP	70.8	– 53.9	66.2	75.4	– 55.9	– 51.7	8.7
Boötes	218.1	34.2	215.7	220.6	32.2	36.1	11.4
CDFS-SWIRE	53.1	– 28.2	50.8	55.4	– 30.4	– 26.0	13.0
COSMOS	150.1	2.2	148.7	151.6	0.8	3.6	5.1
EGS	215.0	52.7	212.4	217.5	51.2	54.2	3.6
ELAIS-N1	242.9	55.1	237.9	247.9	52.4	57.5	13.5
ELAIS-N2	249.2	41.1	246.1	252.3	39.1	43.0	9.2
ELAIS-S1	8.8	– 43.6	6.4	11.2	– 45.5	– 41.6	9.0
GAMA-09	134.7	0.5	127.2	142.2	– 2.5	3.5	62.0
GAMA-12	179.8	– 0.5	172.3	187.3	– 3.5	2.5	62.7
GAMA-15	217.6	0.5	210.0	225.2	– 2.5	3.4	61.7
HDF-N	189.2	62.2	188.1	190.4	61.8	62.7	0.67
Herschel-Stripe-82	14.3	0.0	348.4	36.2	– 9.1	8.9	363.4
Lockman-SWIRE	161.2	58.1	154.8	167.7	55.0	60.8	22.4
HATLAS-NGP	199.5	29.2	189.9	209.2	21.7	36.1	177.7
SA13	198.0	42.7	197.6	198.5	42.4	43.0	0.27
HATLAS-SGP	1.5	– 32.7	337.2	26.9	– 35.6	– 24.5	294.6
SPIRE-NEP	265.0	69.0	263.7	266.4	68.6	69.4	0.6
SSDF	– 8.1	– 55.1	– 357.8	– 18.5	– 60.5	– 48.5	110.4
xFLS	259.0	59.4	255.6	262.5	57.9	60.8	7.4
XMM–13hr	203.6	37.9	202.9	204.4	37.4	38.5	0.76
XMM–LSS	35.1	– 4.5	32.2	38.1	– 7.5	– 1.6	21.8

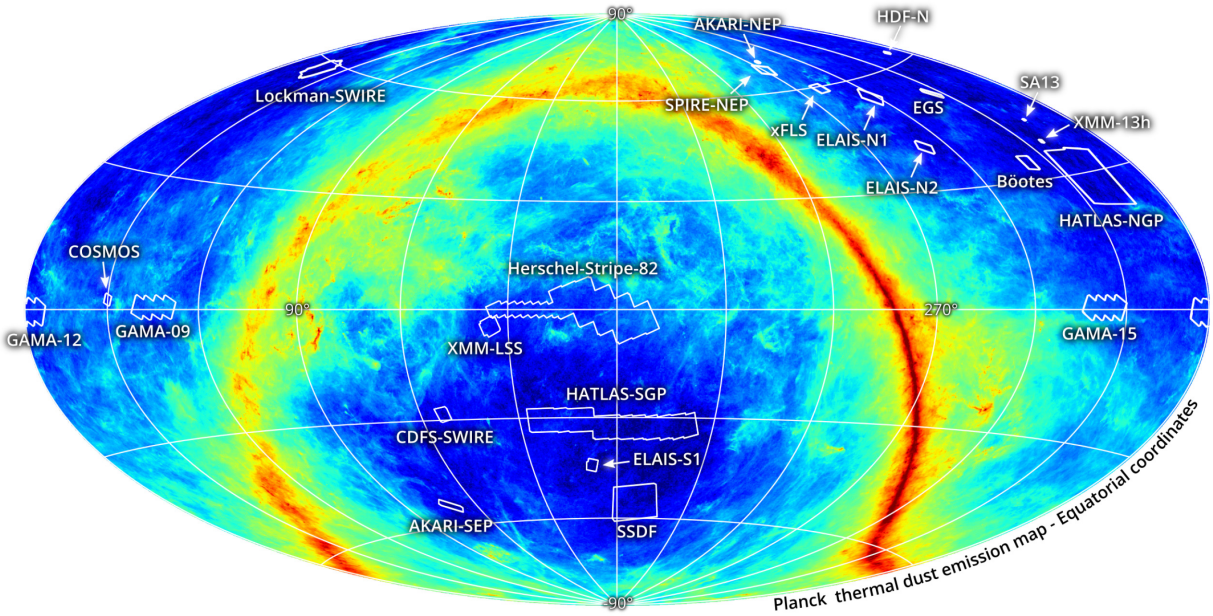


Figure 1. Projection of the HELP fields on to the dust emission from our own Galaxy. Reproduced from Shirley et al. (2019).

of galaxy evolution. The volume of the Universe probed is phenomenal allowing studies of rare or transitory phenomena. This volume also provides large samples of galaxies that can be divided into meaningful subsamples to test galaxy formation scenarios in more detail. The variety of areas and depth allows probes of faint and distant galaxies and enables comparison between distant and nearby samples, i.e. to study galaxy evolution. The volume also provides a complete sampling of the range of galaxy environments. The wealth of multiwavelength data allows for study of the different emission from all the important physical processes e.g. stellar mass, star formation, active galactic nuclei, and provides the basic information like positions and distances.

The challenge to realize this potential is that the information from different survey teams, from different wavelengths, from different facilities, and from different fields is not curated. This means that astronomers will tend to use a limited subset of the available data and also that basic analysis is unnecessarily repeated by many researchers.

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

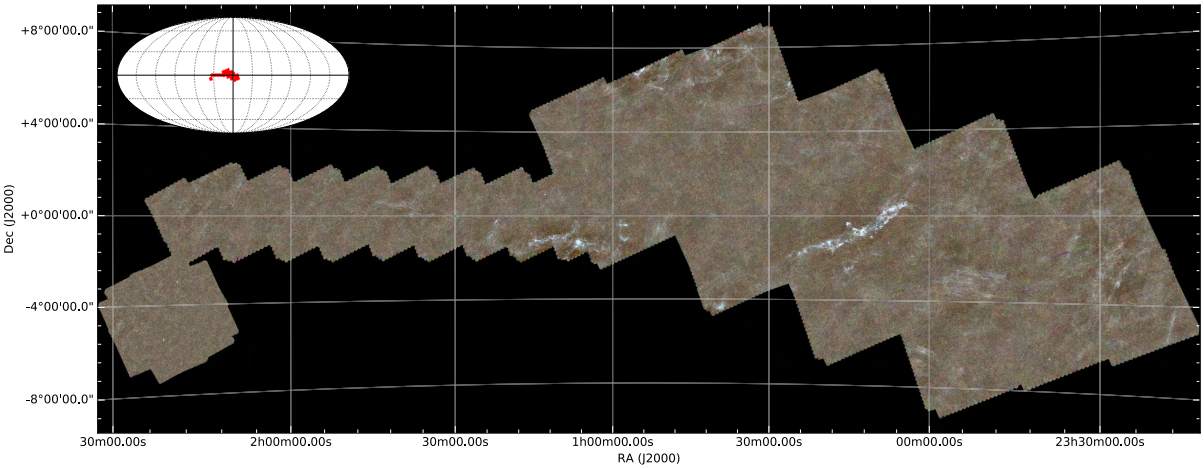


Figure 2. RGB representation of the *Herschel* Stripe 82 and *XMM–LSS* field, with 250, 350, and 500 μm represented by blue, green, and red, respectively. This is the largest contiguous extragalactic region observed by *Herschel*. The maximum scale of the field from the East to West tips is 50° and the separation from edge-to-edge (following the zig-zag, roughly North-to-South) is 11°. The inset indicates the location of this region on an all-sky equatorial projection. The total area of this field is 385 deg^2. Readily apparent is the strong cirrus structure throughout the map, including a 'seagull' like shape in the centre. The data comes from three different observations (*XMM–LSS*, HELMS from HerMES Oliver et al. 2012 and HerS). This image was built for HELP from the processed SPIRE time-lines using the HerMES SMAP processing.

The HELP strategy is to curate these data sets so that they can be used in their entirety by the whole astronomical community with the minimum of specialist knowledge and to add value to these data to enable more efficient and extensive scientific exploitation.

HELP is designed to create a framework for wide-area multiwavelength studies that can be continuously updated with new observations. The scope for Data Release 1, DR1, is to curate object catalogues and photometry at near-IR and optical wavelengths that have been provided by the survey teams from images at mid to far-IR wavelengths alongside spectroscopic redshifts. HELP also provides tools to access the original imaging for manual inspection of interesting sources identified in the final catalogues based on their far-infrared flux or physical properties.

The most fundamental element of the curation is by providing homogeneous data products with consistent measurements, units, and data formats. We provide comprehensive meta data describing the data, using Virtual Observatory (VO) standards. In particular, we provide the user with access to the original references and data from the survey teams, providing written descriptions of all the data in addition to machine-readable files with links to papers, summaries of coverage, and descriptions of instruments, including definitions of bands and links to transmission curves.

A key type of meta data for undertaking statistical studies of galaxy evolution is the selection function. The selection function is the probability of an object being detected and included in a given sample as a function of the galaxy properties. Determining the form of these functions is a major challenge for collated surveys. We therefore need to develop tools to reverse engineer the selection function and protocols to provide those to users. We provide the following for capturing selections functions at increasing levels of sophistication:

Binary coverage maps: These contain the basic information of where, on the sky, data exists. We choose Multi-Order Coverage maps, MOCs (Fernique et al. 2019) to capture this.

Depth maps: These capture a simple, scalar, estimate of the depth of data at any sky position in a given band. We use HEALPIX order 10 cells to provide a map of depths. This order can be changed and is chosen to be a compromise between attaining large enough samples to accurately measure depth and attaining a usefully high resolution.

Completeness maps: These capture the probability that an object of a given intrinsic flux at a particular sky position would be included in a catalogue. These are derived from the depth maps and separately provided for photometric redshift availability.

3.1 Open science

The project has been implemented using open science frameworks with the following general principles:

(i) All codes are publicly available through a version controlled git repository.

(ii) Production code is embedded in extensively annotated Jupyter Notebooks with integrated diagnostic plots.

(iii) Every version of each data product is associated with the git commit code for any code used at the time of production

These key principles enable rerunning of any section of the pipeline in order to facilitate both verification and extension of the work by external researchers. By using Jupyter notebooks to document all the processing on GitHub, all the information about data quality is readily available and the code can be rerun with future additional survey data.

3.2 Tools

The HELP philosophy is that astronomers can easily carry out their scientific investigations without a high degree of instrument or survey specific expertise. We have defined some specific scientific use-cases which should be achievable at the end of the project. Our target is that these recipes could be used by a postgraduate student to produce meaningful scientific results. Our intention is that all scientific results from the team are easily reproduced using these tools. Some of these tools are data base operations. Our data base is VO enabled with

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

ADQL interfaces. Some tools are traditional client/server interfaces. Other tools are developed to provide containers (e.g. Docker) that the user can download and run on their own CPU resources. We supply extensive examples and documentation to aid the uptake of all the tools developed and presented here.

4 THE DR1 WORKFLOW

In this section, we describe the HELP workflow, outlining the key data analysis steps, the decisions taken, and the outputs resulting from the workflow. Fig. 3 is a visual representation of the workflow which we summarize below, with additional details for each stage in the subsections that follow.

First, we create the *master list* of astronomical sources and collate photometry measurements for these sources at all wavelengths between 0.36 and 4.5 μm. Part of the photometry collation process involves determining the highest quality measurements available in a given field and wavelength region. In order for subsequent data processing to work effectively, there should be high quality photometry across a wide spread of wavelengths. This stage also allows us to investigate the depths available in a given area for a given band. Some of the fields in the HELP area have deeper surveys available and wider wavelength coverage than others. After the production of the *master list* which includes all the compiled spectroscopic redshifts, the catalogue is used to calculate photometric redshifts as described in Section 4.6. These are required for spectral energy distribution (SED) modelling. The next stage is to produce the *prior list* which is required for XID + forced FIR photometry. The forced photometry performed by XID + takes the *prior list* as a hard positional prior for objects that are most likely to be detectable in the FIR based on the optical to NIR photometry available. The exact selection of the *prior list* is defined in Section 4.3. Objects with fitted FIR fluxes are then fed through to the final stage where SED modelling is used to calculate galaxy properties.

The final merged catalogue contains all the objects from the *master list* and any subsequent quantities added by the HELP pipeline. The final catalogue can thus be broadly grouped into three hierarchical categories:

(i) The *master list*: Objects detected in an optical or NIR survey.
(ii) The *prior list*: Objects included in the XID + list of prior positions with FIR fluxes in any of the MIPS, PACS, and SPIRE bands available
(iii) The *A list*: Objects selected for SED modelling with XID + detections, a photometric redshift, an SED model, and physical properties estimated.

We will now describe the details of each stage of the pipeline.

4.1 Mid and far-infrared images

This is the first time that all *Herschel* extragalactic blank field survey images are presented together in an homogeneous form. We also provide the *Spitzer* Multiband Image Photometer (MIPS; Rieke et al. 2004) 24 μm band images that are also used for computing forced photometry as part of the general pipeline. There are a total of severn mid or far-infrared (FIR) imaging bands presented here and used to compute forced photometry for far-infrared fluxes. These are the MIPS 24 μm band, the PACS 70 μm, 100 μm bands, and 160 μm, and the SPIRE 250, 350, and 500 μm bands.

4.1.1 Spitzer MIPS 24 μm images

The MIPS images are from two different data sources depending on the field. The *Spitzer* Enhanced Imaging Products (SEIP) is a collection of Super Mosaics of *Spitzer* MIPS data. They are not presented in contiguous form but as individual sometimes overlapping images as originally provided in NASA/IPAC Infrared Science Archive.[2] We do not mosaic them here because each image may reuse data and taking account of this requires decision which reduce the general applicability of the data sets. The *Spitzer* Legacy Program was motivated by a desire to enable major science observing projects early in the *Spitzer* mission. Starting with 6 projects, the Legacy program was expanded to include 20 extragalactic projects over the cryogenic lifespan of *Spitzer*. The data products produced by the Legacy projects are monolithic images covering the full observed field and are used when available.

As the data come from different *Spitzer* projects, the point spread function (PSF) for MIPS is highly variable across fields, therefore we compute the MIPS PSF for each field independently. The process for computing the PSFs is described in Section 4.3.1.

4.1.2 PACS images

PACS observations are available for a subset of the SPIRE area. These observations were sometimes taken in parallel with SPIRE observations (typically larger fields, due to the offset between detectors), while some fields were taken with PACS alone. We have reprocessed all the PACS data, to create an optimized set of images across all fields. The timeline data are initially processed using the *Herschel* Data Processing System (HIPE; Ott 2010), and a basic 'PhotProject' image created per observation. We correct the individual observations for any shift in astrometry by stacking the images on the position of *WISE* sources. The measured RA and Dec shift are applied to the timelines and the data exported using the UniHipe plugin. The final map for the field is created using UNIMAP (Piazzo et al. 2012; Piazzo, Panuzzo & Pestalozzi 2015a; Piazzo et al. 2015b, 2016a, b; Piazzo 2017) combining all available observations. The PSF for PACS is made using the same procedure as for MIPS.

4.1.3 SPIRE images

Each of these projects had different processing pipelines but had similar procedures for producing images from the instrument timelines described in Oliver et al. (2012), Chapin et al. (2011), Levenson et al. (2010), Viero et al. (2013, 2014a), Smith et al. (2017). We have compared images produced by the H-ATLAS and HerMES pipelines using the same input data and found no significant differences. The images presented and used here are also all homogenized to the same units and storage format. The SPIRE images presented here have a number of layers containing the homogenized image, the error image, and *nebulized* image. The *nebulization* process removes large-scale structure caused by cirrus with the method presented in Smith et al. (2017) using the NEBULIZER algorithm developed by the Cambridge Astronomical Survey Unit.[3]

[2] https://irsa.ipac.caltech.edu/Missions/spitzer.html
[3] http://casu.ast.cam.ac.uk/surveys-projects/software-release/background-filtering

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

4.2 The *master list*

The HELP *master list* contains optical, near- and mid-infrared (e.g. *Spitzer* IRAC) catalogues. It includes every source with a measurement in any band. A positional cross-match is then used to combine the various wavelengths. Sources are flagged to indicate data coverage to discriminate lack of detection from lack of data. Full details of the cross-match criteria and mis-association fractions are given in Shirley et al. (2019). While cross-match radii are determined for each input catalogue depending on positional errors it is typically around 0.4 arcsec. We also provide a table of the original catalogue IDs and the original catalogues. This means that where additional useful information is included in the input catalogue, it can be quickly recovered using the table of cross identifiers. All these data are provided in a simple and well documented structure to facilitate independent validation and external use. Full details on the production of the *master list* are presented in Shirley et al. (2019) and through the code itself on Github.

The *master list* is central to the HELP pipeline and data products. As the *master list* progresses through the pipeline, some fraction of objects satisfy the criteria required for the next additional processing stage. For instance, if there are no sufficient optical photometry points then the photometric redshift (photo-z) calculation is likely to fail and the object will not have a photo-z, and so cannot be used for SED fitting. Likewise, when XID+ is run, some objects will have low signal to noise so will not be 'detected' and will not be used for SED fitting. We aim to model each of these selection effects so that the full selection function can be understood as accurately as possible. Shirley et al. (2019) provides depth maps in order to model the detection of objects in the original catalogues. Here we provide additional depth maps for both the photo $= z$ catalogues, which can significantly affect selection at the margins, and for the new photometry presented here.

4.2.1 *Star masks and artefacts*

Astronomical catalogues contains spurious artefacts resulting from instrument noise and dependent on the extraction methods. A key source of artefacts is bright stars, where the wings of the PSF or scattered light raise the background and spurious signals exceed the object detection threshold. We therefore developed a star masking pipeline which highlights regions likely to suffer from these artefacts. Our approach is to look at the excess number density of catalogued sources as a function of distance from bright stars.

For our bright star list we select all *Gaia* (Brown et al. 2016; Lindegren et al. 2016) stars with $g < 16$ mag. The reference band used on a given field is determined by which has the most impact on the *prior list* for XID+. In IRAC regions this is the IRAC 3.6 μm band, in other regions it is the deepest K band. Within magnitude bins we then determine an effective exclusion radius, r_{50}, at which the excess number density (above the background level) drops to fifty per cent of its peak. We choose fifty per cent because the decline is steep and taking the mid-point is a robust measure of its location. We then fit a linear relation:

$$\log_{10}(r_{50}) = A + BM_{\rm star}. \quad (1)$$

This function defines the radius of a circle around each *Gaia* star, within which all objects are excluded from the *prior list* and which should be excluded from all statistical analysis. This function typically reduces to zero for all objects below 14 mag. We fit the parameters A and B based on the magnitude bins of size 0.5 and generate star masks for each field and target band independently. The final star masks are provided in the DS9 and MOC formats.

4.3 XID+ : the probabilistic deblender for confusion dominated images

For many *Herschel* fields, in addition to the SPIRE images we also have *Spitzer* MIPS 24 μm and *Herschel* PACS 100 and 160 μm images that cover the mid to far-infrared part of the electromagnetic spectrum. However, due to the relatively large beam size of these images compared to the galaxy density (≈30 per SPIRE beam for optical sources with $B < 28$), multiple galaxies can be located within the same instrument beam. This is referred to as the problem of source confusion.

To obtain accurate photometry from these infrared images, accounting for source confusion is essential. One way to solve the problem is to use prior information to accurately distribute the flux in the images to the underlying astronomical objects. For example, if we know the location of a galaxy to a reasonable tolerance (e.g. from an optical image where resolution is much smaller than the *Herschel* beam), we may expect a galaxy to be found in the MIPS, PACS, and SPIRE images at the same location. Typically the position of known objects have errors significantly less than the FIR point spread function such that we assume the positions are known precisely.

As part of HELP, we have developed XID+ (Hurley et al. 2017) which uses a probabilistic Bayesian approach that provides a framework in which to include prior information and obtain the full posterior probability distribution on flux estimates. Obtaining the full posterior probability distribution is particularly important for getting accurate uncertainties on source flux. A given XID+ model is described as:

$$d_j = \sum_{i=1}^{S} P_{ij} f_i + N(0, \Sigma_{\rm inst}) + N(B, \Sigma_{\rm conf}), \quad (2)$$

where d_j is the model of the map pixel j, P_{ij} is the Point Response Function (PRF) for source i in pixel j, f_i is the flux density for source i, and two independent noise terms for instrumental and confusion noise: $N(0, \Sigma_{\rm inst})$ and $N(B, \Sigma_{\rm conf})$, respectively.

Rather than find just the flux values that maximizes the likelihood, XID+ maps out the entire posterior, $p(f|d)$, which can be defined as:

$$p(f|d) \quad \propto \quad p(d|f) \quad \times \quad p(f), \quad (3)$$

where $p(d|f)$ is the likelihood, the probability of the data given the flux densities, and $p(f)$ is the prior probability distribution on the fluxes. The method is fully described in Hurley et al. (2017).

4.3.1 *HELP XID+ pipeline*

HELP uses XID+ to carry out forced photometry on the *Spitzer* MIPS and the *Herschel* PACS and SPIRE images to produce catalogue fluxes for the HELP data base. Our prior source list for these images are constructed using two different pipelines, which we describe in detail in the following paragraphs. For flux priors, we use uninformative flux priors (i.e. uniform flux prior bounded with reasonable limits derived locally from the image) to enable an unrestricted range of analysis with the HELP data products. More informative prior information would be preferable for specific science projects and is a powerful approach to extract more information out of the data (e.g. Pearson et al. 2017b, 2018), however their use must be fully understood and taken into account,

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

such that they are more suited for bespoke projects than for a data product. If not then apparent results might reflect the ancilliary data more than the far-infrared maps directly. In the next paragraphs, we describe the steps followed to run XID+ across the HELP fields.

Our list of prior sources is constructed from the *master list*. Fitting all the sources in the *master list* to the source confused infrared images results in fluxes that are degenerate without using more informative flux priors. We therefore have to limit the number of sources that go into our prior source list to those that are most likely to be detectable in the images. This approach fits the Bayesian philosophy of model building; build a simple model, fit to the data, evaluate, and finally improve the model where necessary. The prior source list is an integral part of our model for the infrared images. Limiting the number of sources to those that are most likely to be detected simplifies the model and the Bayesian P value maps described later, which provide a data product to carry out model evaluation. They identify where additional sources are needed to model the images. We depict the two ways we have constructed the prior source lists in Fig. 3, one for fields where there is *Spitzer* and another for when there is no *Spitzer* coverage.

For fields covered by *Spitzer*, we use sources detected in any of the *Spitzer* IRAC bands as they are known to be a good tracer for the *Spitzer* MIPS images (Rodighiero et al. 2006). To remove any possible artefacts in the IRAC catalogues, we impose an additional constraint that sources must also have a detection in either the optical or NIR wavelength range. Sources that meet this criteria, constitute the XID+ *prior list* for MIPS images.

Once we have the output from XID+ on the *Spitzer* MIPS images, we use our definition of detection level to select sources to be used for the XID+ *prior list* for the *Herschel* PACS and SPIRE images, which are fit independently (we do not use PACS XID+ detections as a prior for SPIRE given that the PACS data tends to be shallower than SPIRE). Detection is determined by the MIPS flux level where the Gaussian approximation to uncertainties is valid. Below a certain flux level, the map uncertainty is too large to be able to constrain the source flux and the flux posterior probability distribution for sources becomes skewed as the uniform flux prior dominates over the likelihood. This point is determined by manual inspection and given in Table 2.

For areas that have not been observed by *Spitzer* IRAC, we compute a total dust luminosity ($L_{\rm dust} = \int_{8-1000\,\mu m} B_\lambda {\rm d}\lambda$) for each object using the CIGALE code, as described in Section 4.7. We used the relationship between the ratio $\frac{L_{\rm dust}}{f_{250}}$ and redshift, based on data from COSMOS field where we have both $L_{\rm dust}$ and f_{250}. We apply this relationship to the sources for which we have dust luminosity and redshift predictions to estimate f_{250}. Sources that have a predicted $f_{250} > 5$ mJy are added to the prior source list and XID+ is run on the *Herschel* PACS and SPIRE images. This flux cut was chosen after running XID+ on a small region within *Herschel* Stripe 82, using a range of flux cuts (e.g. using different *prior lists*), and comparing the Bayesian P-value maps described in Section 4.3 to check whether the cuts applied to the prior source lists provided a good fit to the map. This was to manually check whether bright sources were being missed at a given predicted $L_{\rm dust}$ cutoff value.

For a number of fields, the dust luminosity relationship varies slightly due to an early bug in the prediction. Upon investigation, this has the effect of missing out 17 per cent of the lowest flux sources that would otherwise have been included, whilst including 48 per cent of sources that otherwise would have been removed from the source list. In these data release this will have the effect of introducing a selection effect and increasing the effective flux cut in predicted 250 μm flux. This will be propagated into the effective forced photometry depth maps we provide and so will be automatically accounted for by modelling the selection using these maps.

As described in Section 4.2, star masks are used to define regions where bright stars cause large numbers of artefacts and spurious sources, resulting in these regions having no prior list sources. As XID+ is used for source de-blending rather than source detecting, it is not appropriate to apply XID+ to areas of the map where you have no prior knowledge of sources. We therefore exclude pixels from the area defined by the star masks from XID+ fitting using the Multi-Order Coverage maps (MOCs) built from the star masks.

XID+ uses a Bayesian framework and so the flux parameters require a prior probability distribution. We use non-informative, uniform distributions with sensible limits. For *Spitzer* MIPS images, the upper and lower 24 micron flux limits are based on the longest wavelength IRAC flux available. For a lower limit, we take $f_{\rm IRAC}/500$ and for upper limit $f_{\rm IRAC} \times 500$. For *Herschel* SPIRE and PACS, we set the flux prior lower limit to zero and source specific upper limit equal to the local (as defined by the PRF) maximum pixel value plus the absolute value of the prior mean for background plus two times the standard deviation of the background prior. This combination of maximum pixel value alongside value and width of the background prior gives a conservative but not extreme upper limit on the flux.

An important part of the model is the Point Response Function (PRF). The PRF is the convolution of the point spread function and the transfer through to a pixel response function via the detection and map building process. This fully maps the contribution a point source makes to each pixel.

In *Herschel* SPIRE images the PSF is assumed to be a Gaussian, with full-width half-maximum (FWHM) of 18.15, 25.15, and 36.3 arcsec for 250, 350, and 500 μm, respectively (Griffin et al. 2010). This is convolved with the pixel space to produce the PRF. In the case of both *Spitzer* MIPS and *Herschel* PACS the PRF is calculated by stacking the flux of point-like sources from the astrometry corrected AllWISE catalogues, referenced into the *Gaia* reference frame. Morphological outliers (with a reduced $\chi^2 > 4$ in the AllWISE profile fit) were excluded before stacking.

The PRF obtained in the stacking will not be as high signal-to-noise as the instrumental PSF and will not track the extended wings of the PDF. To get the correct normalization, we match the curve-of-growth of the instrumental PSF to that determined from our PRF.

Having defined the PRF, we use it to populate the pointing matrix, which describes how much each source contributes to each pixel in the map. It is calculated by taking the PRF for each band, centring it on the position for each source and carrying out a nearest neighbour interpolation to establish the contribution each source makes to each pixel in the map.

Running XID+ on the full images simultaneously is computationally unfeasible. We therefore divide the image into equal areas using the Hierarchical Equal Area isoLatitude Pixelization of a sphere (HEALPIX). The resolution of the pixels are determined by the HEALPIX level, with optimum order for *Spitzer* MIPS and *Herschel* PACS set at 11, and 9 for *Herschel* SPIRE, which correspond to ≈ 1.718 arcmin, and ≈ 6.871 arcmin, respectively. When we fit each tile, the perimeter is extended by one HEALPIX pixel with two levels higher resolution (i.e. level 13 for MIPS/PACS, level 11 for SPIRE with a resolution ≈ 25.77 arcsec and 1.718 arcmin, respectively), so any sources that could contribute within the HEALPIX pixel of interest are taken into account.

As with other MCMC fitting, we need to run chains long enough to ensure we converge locally and with multiple chains to ensure we have found a global minimum. We use the default number of chains, four, and discard the first half of each chain as 'warm up' or 'burn

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

in'. In order to assess the convergence of each parameter we use the same diagnostics, $\hat{R}$ and $n_{\text{effective}}$ used in Hurley et al. (2017) and described in Gelman et al. (2013).

4.3.2 HELP XID+ data products

One of the key strengths of XID+ is that it maps out the posterior rather than just the maximum-likelihood estimate. For individual objects of interest the full posterior can be used to verify the quality of the fit. This also allows using the joint posterior probability distribution of two correlated sources, getting the full uncertainty on the fluxes. The full posterior is stored in a *'.pkl'* file for each HEALPIX tile. These data can be provided on request.

The posterior distribution also allows us to perform a probabilistic check of the XID+ fit. When examining goodness of fits, the traditional method is to look at the residuals. i.e. (data − model)/σ. Because we have many samples from the posterior, we can create a distribution of model images that cover all the possible images XID+ generates from the posterior parameter values. Having a distribution of model images that we can compare to the original data provides a more robust check than if we were to use one, best-fitting model map coming from likelihood. Using a distribution of models drawn from the posterior, and comparing to the original data is called a posterior predictive check (Gelman, Meng & Stern 1996).

For XID+, our approach to posterior predictive checks is to compare the observed flux of a pixel to the distribution realized in the model images. By calculating the fraction of model realizations that are above the observed value, we obtain the Bayesian P value. A P value of ∼0.5 means the model is consistent with the data. Values close to 0 shows that there is too much flux in the model compared to the map, whereas values close to 1 indicates there is flux in the map that our model cannot explain. We convert these probabilities to a typical 'σ' level. Fig. 4 illustrates this process, which is repeated for every pixel in the map to produce Bayesian P-value maps for each band and field.

By using the full statistical power of the posterior probability distribution, these maps are more robust and less noisy than a traditional residual image and can be used to identify where the XID+ model assumptions break down. Examples of where XID+ might provide a bad fit are extended sources, which will appear in the Bayesian P value maps as a negative centre and positive rings. Another more interesting example would be missing sources: i.e sources in the map that are not in the *prior list* which will appear in the Bayesian P-value maps as positive peaks. Obtaining a catalogue of sources from the Bayesian P-value maps can be added to the prior source list for XID+ to obtain updated catalogues or to find interesting objects which are drop outs in the lower wavelength images.

For the final catalogues, we summarize the posterior flux distribution in the form of 16th, 50th, and 84th percentile. This is equivalent to mean, and mean $\pm\sigma$ if the posterior distribution is Gaussian. These values are used for SED fitting and most science cases. We use the skewness level of the catalogue to determine a detection level, as described in Hurley et al. (2017) and flag sources that are below this level. We also carry out model checking to identify whether the local (defined by the PRF) area of the map for each source is a good fit. To quantify this check, we define a Bayesian P-value residual statistic as follows; from the XID+ posterior probability samples, we use the same model images as our Bayesian P-value maps and calculate weighted residuals for the local pixels for every sample. We then calculate what percentage of our model images have a χ^2 statistic greater than we would expect. This value is a probability, with zero indicating our inferred model will always provide a fit deemed good given the uncertainties, and 1 indicating our inferred model would always provide a fit deemed poor given the uncertainties. We provide this Bayesian P-value residual statistic for each source and each band. Sources with a value >0.5 are flagged as unreliable.

The final product consists of a catalogue with the flux percentiles, the median background, and the convergence statistics, Bayesian P-value residual statistic and a flag for sources that are below detection level or have a high Bayesian P- value residual statistic. In addition to the final products, the XID+ example notebooks also provides some visualization tools to conduct a comprehensive analysis of the posterior. It is possible to create posterior replicated images and animations, create the marginalized posterior plots or reproduce the Bayesian P-value maps. There are examples available in the XID+ user guide.

4.4 Blind catalogues

An important additional step for providing a legacy data set, suitable for community exploitation is to construct a catalogue of objects detected in the SPIRE images without reference to any other data and with fluxes extracted at the SPIRE wavelengths (a 'blind' catalogue). These catalogues give a perspective of the sub-mm sky unaffected by any prior prejudice. Again, the most significant challenge is the large SPIRE beam, leading to source confusion (e.g. Nguyen et al. 2010) which requires careful de-blending and the resultant catalogues of sources do not necessarily correspond one-to-one to individual galaxies. To enable statistical studies there are a number of key metrics required: positional and flux biases and accuracy; completeness and reliability. Similar catalogues and metrics have been produced and made public for the other HerMES fields (Smith et al. 2012; Wang et al. 2014), H-ATLAS fields (Maddox et al. 2018), and for all SPIRE fields in the *Herschel* SPIRE Point source catalogue (ESA 2017). We have produced new blind catalogues for all the HELP fields using a similar method to Chapin et al. (2011), and described below:

4.4.1 Peak finding in match filtered images

The blind sources are identified by searching for peaks in the matched filtered (MF) images, as they are optimized for identifying sources in source-confused images (Chapin et al. 2011). We require peaks to have a flux density above the 85 per cent completeness level for each SPIRE band, where completeness is quantified as $1 - \frac{N_{\text{spurious}}}{N_{\text{real}}}$ as a function of flux density. N_{real} is the number of identified peaks in a given flux bin and N_{spurious} are the number of identified peaks in the negative version of a map (we assume the noise in the map is symmetrical about zero, thereby identifying peaks in the negative map will quantify the number of peaks that are from random fluctuations in the noise, as a function of flux).

4.4.2 Determining accurate positions

Having found the peaks for each band independently, we determine accurate centres for each source by determining the best-fitting flux density for the three SPIRE bands using inverse variance weighting and for subpixel steps around each peak. For each subpixel position, we calculate the Pearson correlation coefficient between best-fitting

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

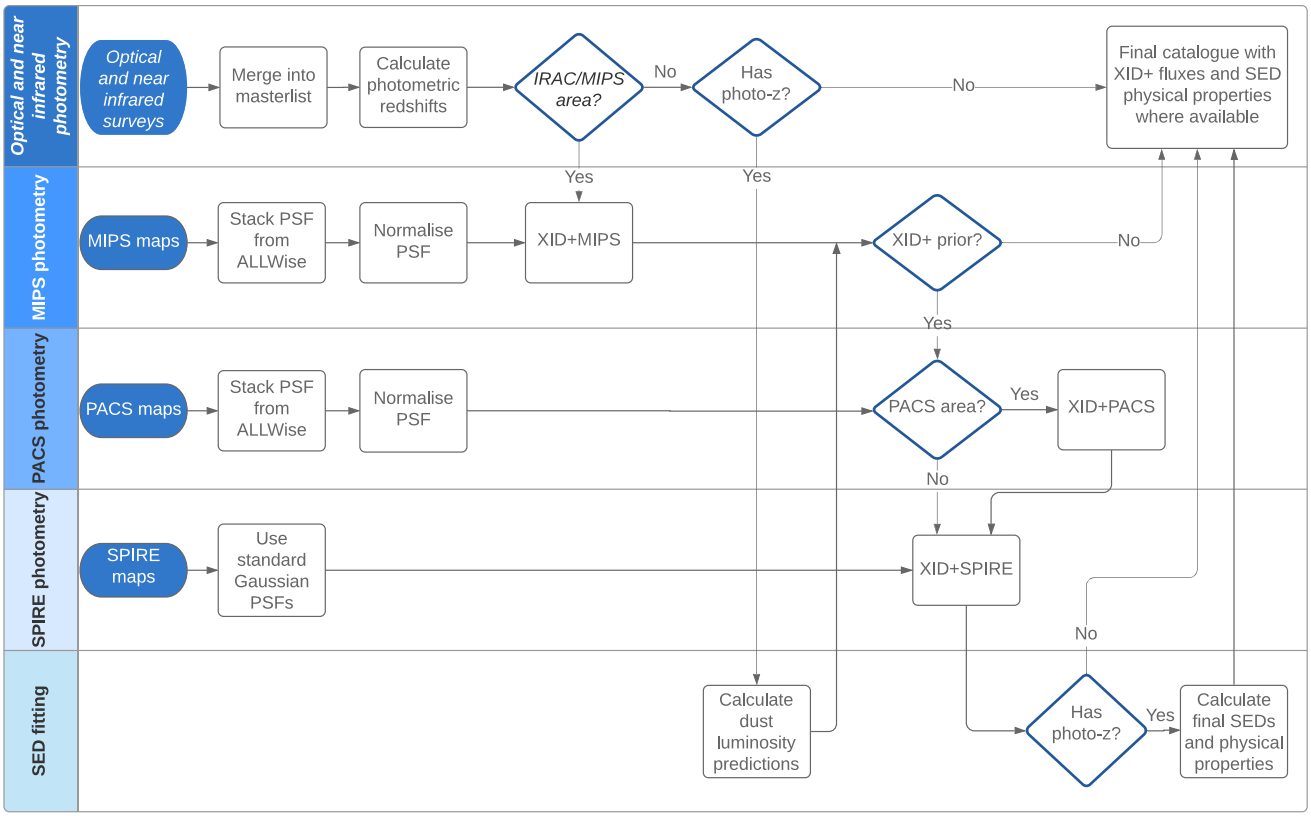


Figure 3. Overview of the full workflow. All objects in the original *master list* make it through to the final list regardless of which added value derived quantities are available. A given astronomical object from an input optical or near-infrared survey can be traced through the pipeline using this high level schematic. For each of the XID + runs and final CIGALE SED run there are further criteria applied to each object that are not shown here for simplicity but are described in the relevant sections.

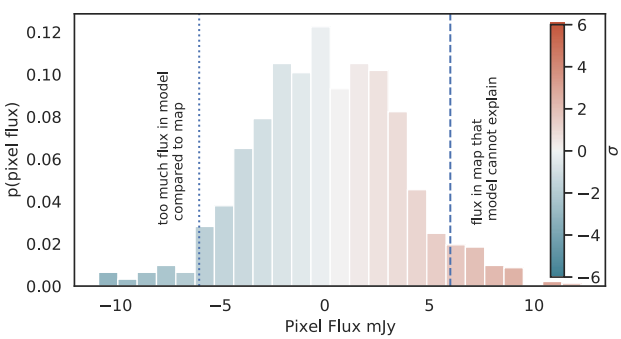


Figure 4. An example distribution of posterior flux estimates for a single pixel. Comparison with the observation gives a measure of how unlikely the observed value is, given the posterior model, and is comparable to a residual. The resulting probability is converted to an equivalent Gaussian σ level. A high value of σ can be an indication of a missing source i.e. there is flux in the map that the model cannot explain. This can be a means to find interesting new objects not seen in our prior catalogues.

and map. The position with largest correlation is taken as the new position. We combine the three resulting SPIRE catalogues by removing duplicates at 350 and 500 μm using a nearest neighbour matching algorithm with 12 and 18 arcsec radius, respectively, adopting the position in the shortest wavelength available for each merged source.

4.4.3 Determining fluxes with XID+

Having obtained accurate positions and removed duplicates, XID+ is applied to the standard image map,[4] using the merged, blind source matched filtered catalogue as the prior source list. The resulting XID+ catalogue, as described in Section 4.3 is merged with the blind source matched filtered catalogue (so as to preserve information about the sources used as priors for XID+) to produce the HELP blind source catalogue.

4.5 Spectroscopic redshifts

As part of the HELP project we have collected spectroscopic redshifts from 101 different origins across all HELP fields, and created one merged catalogue for each field. Merging catalogues can be problematic due to the varying degrees of information and data quality available. Here we briefly describe the process used to make the matched catalogues.

4.5.1 Spectroscopic redshift catalogue homogenization

The first step is to homogenize the individual redshift catalogues into a 'standard format', where we extract the RA, Dec, redshift, and if available whether the spectra classify the object as a QSO or

[4]We apply XID + to the standard map rather than the nebulized map since it simultaneously fits for the background alongside the individual source

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

AGN. We also assign a 'Quality Flag' (Q) to each spectra, using the information provided by the individual survey. For HELP we adopted the same Q definition as used by the 2dF survey (Colless et al. 2001), and GAMA team (Hopkins et al. 2013), where Q characterizes confidence level on a five point scale.

In reality, assigning an exact Q particularly for small surveys where reliability information is not given is difficult, therefore if a survey does not give an estimated reliability and claims the redshift is 'good' or 'reliable' it is normally assigned to $Q = 3$. If available we also record the exact reliability given for the individual survey, as many large surveys have reliabilities higher than 90 per cent but not high enough to be assigned $Q = 4$; this enables the user to set a individual reliability threshold (for example >95 per cent). For science studies, we recommend using any redshifts with $Q \geq 3$. We also search the individual catalogue to ensure there are no duplicate entries by checking for no self-matches within 0.4 arcsec. Any duplicates were manually investigated and the best redshift kept following the procedure outlined below.

Before the catalogues for each field are merged, each catalogue is given a unique binary identifier (i.e. 1, 2, 4, 8, etc...). This means if the same object is observed by multiple surveys the source identifier numbers are added together, and the new source identifier can be used to see each catalogue that provided the corresponding redshift. For example a redshift with a source identifier of 11 would have been observed by surveys 1, 2, and 8 (but not 4). All individual redshift catalogues and their identifier are listed in Table C2.

4.5.2 Merging of spectroscopic redshift catalogues

To merge the individual catalogues, we match each catalogue sequentially by using STILTS (Taylor 2006a) to perform a sky match in a radius of 1–2 arcsec and checking up to 5 arcsec, where the radius is chosen to give the optimum matches. By performing a manual check of objects close to the matching radius the code can be modified so that any above/below the matching threshold should be merged/split.

For any redshifts that have been found to be a match between the merged catalogue and the new individual catalogue the redshift with the higher Q flag is kept. For the case where both Q flags are $\geq$3 a check is performed to see if the redshifts have a $\Delta z < 0.01$ and <5 per cent (for lower redshifts). If the two reliable redshifts disagree, a manual choice is made to decide the best redshift based on available information (i.e. if multiple sources, quality of data etc...). The fraction of conflicts between reliable redshifts is very small, for example in the SGP field we have 16 conflicts out of 47 213 redshifts. For redshifts from PRIMUS (Coil et al. 2011) due its lower resolution we increased the matching to $\Delta z < 0.03$ and used the higher resolution spectra values. For any merged galaxies the QSO/AGN flag is switched on if any of the QSO/AGN flags is set.

In total for HELP, we collected 891 317 spectroscopic redshifts, of which 713 660 are considered reliable ($Q \geq 3$), and of these 621 407 are unique sources. Table 2 gives the number of reliable spectroscopic redshifts available for each field.

4.6 Photometric redshifts

For the majority of HELP fields, where extensive multiwavelength photometry is available, photo-zs have been calculated and are presented here for the first time. For 8 of the small HELP fields, where the best available optical photometry is provided only by all-sky photometric surveys and therefore is not improved by combining multiple surveys, we make use of photo-zs presented in the literature (Zou et al. 2019). These account for under 1 per cent of HELP DR1 photo-zs.

4.6.1 HELP derived photometric redshifts

Photo-zs for the prime HELP optical data sets are estimated based on the method presented by Duncan et al. (2018a, b). The approach combines multiple templates and machine-learning estimates to produce a hybrid consensus photo-z estimate with accurately calibrated uncertainties. This method is only possible on fields with sufficient spectroscopic redshift samples as described below. We refer the reader to the original papers for a detailed discussion on the motivation and implementation. Below we summarize the implementation of this method for HELP.

As in Duncan et al. (2018a), three different template-based estimations are calculated using the EAZY software (Brammer, van Dokkum & Coppi 2008) with three different template sets: one set of stellar-only templates, the EAZY default library (Brammer et al. 2008), and two sets including both stellar and AGN/QSO contributions, the XMM-COSMOS templates (Salvato et al. 2008, 2011) and the Atlas of Galaxy SEDs (Brown et al. 2014). The individual template fitting results are optimized using zero-point offsets calculated from the spectroscopic redshift sample[5] and the posterior redshift predictions calibrated such that they accurately represent the uncertainties in the estimates. When sufficient spectroscopic training sets are available for a given field, additional Gaussian process photo-z estimates (GPz; Almosallam et al. 2016a; Almosallam, Jarvis & Roberts 2016b) are produced by training on one or more subsets of the *master list* photometry. The multiple individual photo-z estimates are then combined following the Hierarchical Bayesian combination method first presented in Dahlen et al. (2013), incorporating the additional improvements outlined in Duncan et al. (2018a, b).

A key step in the Hierarchical Bayesian photo-z method outlined in Duncan et al. (2018b) is the separate treatment of priors and uncertainty calibration for known AGN. For HELP, we identified known AGN as follows:

(i) *Optical AGN*: We identify known optical AGN through cross-matching of the *master list* with the Million Quasar Catalogue compilation of optical AGN (Flesch 2015). Sources which have been spectroscopically classified as AGN are also flagged (see 4.5).

(ii) *X-ray AGN:* In HELP fields that have been targeted by deep pointed X-ray surveys, we make use of any publicly available X-ray catalogues and associated optical IDs to identify known X-ray AGN in the *master list*. Outside of the publicly available deep X-ray surveys, we make use of the Second Rosat all-sky survey (2RXS; Boller et al. 2016) and the *XMM–Newton* slew survey (XMMSL2)[6] all-sky surveys. X-ray sources were matched to their HELP *master list* optical counterparts using the published AllWISE cross-matches of Salvato et al. (2018). AGN and star-forming (or stellar) X-ray source populations are identified based on the colour criteria presented in Salvato et al. (2018):

$$[W1] > -1.625 \times \log_{10}(F_{0.5-2\mathrm{keV}}) - 8.8, \quad (4)$$

where [$W1$] is the AllWISE W1 magnitude *in Vega* magnitudes and $F_{0.5-2\,\mathrm{keV}}$ the 2RXS or XMMSL2 flux in units of erg^{-1} s^{-1} cm^{-2}.

[5]The zero-point offsets derived for each template set are stored and made available to the user for reference

[6]https://www.cosmos.esa.int/web/xmm-newton/xmmsl2-ug

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

Table 2. Flux cuts where Gaussian approximation to uncertainties is valid.

field	MIPS	SPIRE (250, 350, 500 μm)	PACS (100, 160 μm)
AKARI-NEP	30 μJy	5, 5, 6 mJy	12.5, 17.5 mJy
AKARI-SEP	40 μJy	–, –, –	–, –
Bootes	20 μJy	5, 5, 10 mJy	10, 12.5 mJy
CDFS-SWIRE	20 μJy	4, 4, 6 mJy	30, 30 mJy
COSMOS	μJy	mJy	mJy
EGS[a]	–	–, –, –	10, 10 mJy
ELAIS-N1	20 μJy	4, 4, 4 mJy	12.5, 17.5 mJy
ELAIS-N2	20 μJy	4, 4, 4 mJy	12.5, 17.5 mJy
ELAIS-S1	30 μJy	4, 4, 6 mJy	20, 30 mJy
GAMA-09[a]	–	4, 4, 6 mJy	20, 30 mJy
GAMA-12[a]	–	4, 4, 6 mJy	20, 30 mJy
GAMA-15[a]	–	4, 6, 10 mJy	20, 30 mJy
HDF-N[a]	–	4, 4, 4 mJy	–
Herschel-Stripe-82[a]	–	10, 10, 12 mJy	–
Lockman-SWIRE	20 μJy	4, 4, 6 mJy	16, 25 mJy
NGP[a]	–	6, 6, 10mJy	25, 25 mJy
SA13[a]	–	4, 4, 4 mJy	–
SGP[a]	–	6, 6, 9 mJy	–
SPIRE-NEP	20 μJy	6, 6, 6 mJy	–
SSDF[a]	–	10, 10, 10 mJy	–
xFLS	20 μJy	4, 4, 4 mJy	–
XMM–13hr[a]	–	4, 4, 4 mJy	–
XMM–LSS	SWIRE: 20 μJy; SPUDS: 10 μJy	4, 4, 4 mJy	12.5, 17.5 mJy

Note. [a]$L_{\rm dust}$ was used as prior.

(iii) *Infrared AGN:* When *Spitzer* photometry is available for a given field, IR AGNs are also identified using the updated IR colour criteria presented in Donley et al. (2012).

Based on these criteria, *master list* sources classified as AGN are flagged and processed following the AGN (Duncan et al. 2018a, b). We note that by design these selections are not intended to provide complete samples of the AGN population within the HELP *master list*.

For each HELP field, we provide documentation outlining the specific *master list* and spectroscopic redshift compilation versions that were used for the photo-z estimation, calibration, and validation. We also document the precise set of optical filters included in the template fitting along with the lists of filter combinations used for GPz machine-learning estimates (where available).

The photo-z estimates produced by HELP are provided in two ways. First, we provide the full calibrated photo-z posterior for all sources for which the hierarchical Bayesian procedure outlined above could be performed (as well as working examples of how to extract and use this information). Secondly, we provide summary values for the posteriors in a format suitable for catalogues and single-value based quality statistics. We follow the approach outlined in Duncan et al. (2019), which aims to provide an accurate representation of the redshift posteriors, regardless of whether the posterior is uni- or multimodal (as is often the case for photo-zs). In summary, for each calibrated redshift posterior the primary (and secondary if present) peak above the 80 per cent highest probability density (HPD) credible interval (CI) is identified based on the redshifts at which the redshift posterior, $P(z)$, crosses this threshold. For each peak, the median redshift within the boundaries of the 80 per cent HPD CI is calculated to produce our point-estimate of the photo-z (hereafter $z_{1,\,\rm median}$ or $z_{2,\,\rm median}$). To present a measure of redshift uncertainty within the HELP catalogues we also then present the lower and upper boundaries of the 80 per cent HPD CI peaks (i.e. where the $P(z)$ crosses the threshold). In the HELP catalogues, data base and the subsequent analysis (e.g. Section 4.7), photo-zs values are taken to be $z_{1,\,\rm median}$.

4.6.2 *Literature photometric redshifts*

For the fields AKARI-NEP, AKARI-SEP, ELAIS-N2, HDF-N, SA13, SPIRE-NEP, xFLS, and XMM-13hr we use the photometric redshifts presented in Zou et al. (2019) based on Legacy Survey *grz* fluxes and Wise W1 and W2. These are fields without additional data sets to those presented there and therefore where recalculating them was of little additional benefit. These are matched into the *master list* using a positional cross-match with a radius of 0.4 arcsec. These redshifts are subject to a cut of $r < 23$ mag. After processing they also impose a redshift cut of $z < 1$ and stellar mass cut $8.4 < \log(M_*) < 11.9$. These cuts impose limits on studies that can be conducted with these areas but lead to a well-defined selection function. This redshift selection function can be automatically handled using the photometric redshift depth maps.

4.6.3 *Photo-z validation*

Due to the range in photometric data quality and spectroscopic training samples available in each field, there is significant variation in the photo-z quality across the HELP footprint. Fig. 5 presents a qualitative illustration of the accuracy of photo-z in fields that demonstrate the dynamic range in parameter space probed by HELP; the deep but small COSMOS field and the *Herschel* Stripe 82 field that spans over 360 deg^2. In both fields, we limit the sample to sources with reliable spectroscopic redshifts and detections in at least the optical and NIR regimes. With additional selection criteria (such as magnitude selection and redshift limits), samples with reliable and precise photo-z can be easily produced for each field. To facilitate

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

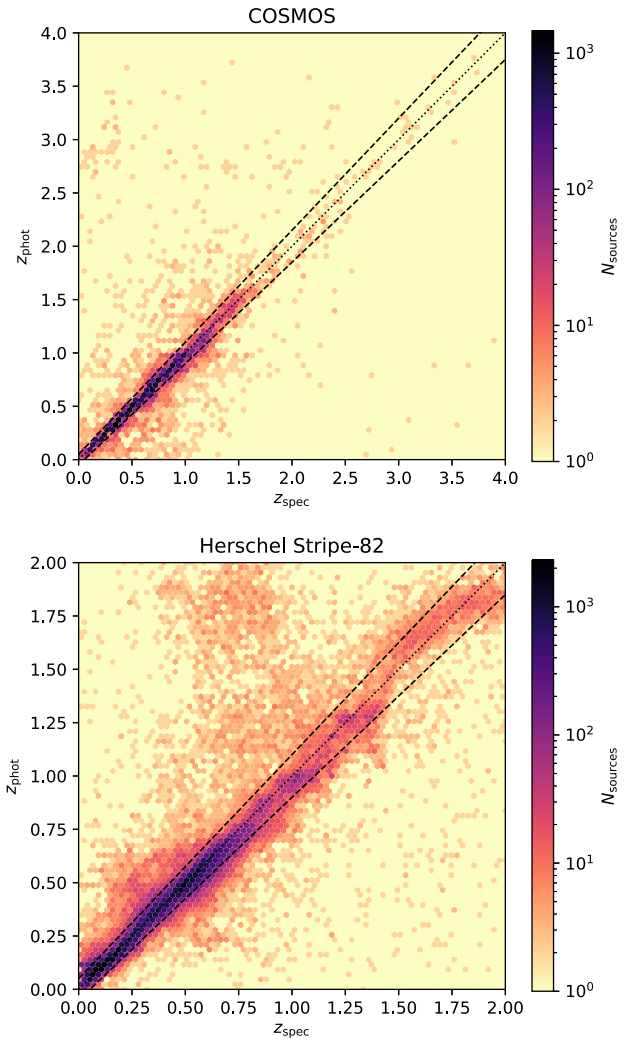


Figure 5. Consensus photometric redshift estimates as a function of spectroscopic redshift for two HELP fields at scales of $\sim 10^0$ and $\sim 10^2$ deg^2 (COSMOS and *Herschel* Stripe 82, respectively). The dotted black line corresponds to the desired 1:1 relation while the dashed lines correspond to $\pm 0.05 \times (1 + z_{spec})$, illustrating the typical scatter within the samples.

this, as part of the photo-z validation steps we generate a number of diagnostic and illustrative plots to enable assessment of the photo-z quality within each field. However, we note that given the limited availability of spectroscopic data in some fields (and the overall variation in spectroscopic coverage), it is not possible to provide homogeneous and complete assessment across the full HELP photo-z sample.

4.6.4 *Photo-z selection functions*

Following the HELP goals and philosophy, we have also endeavoured to provide informative data products and tools for understanding and accounting for both the explicit and implicit photo-z selection functions.

Given the in-homogeneity across both the full HELP sky and across each individual survey field, quantifying the spatially varying selection function is critical. However, due to the complicated nature of the optical selection function within deep fields that typically have very heterogeneous depth and filter coverage the selection will be highly multidimensional. Additionally, as the exact selection function corresponding to a given sample are user and science-case dependent (i.e. depending on required redshift or photometric quality criteria), a novel and flexible approach is required.

Building upon the HEALPIX based optical depth maps produced in the production of the optical *master list* (Shirley et al. 2019), we provide a set of tools to calculate photo-z completeness across a given HELP field as a function of any desired *master list* magnitude within the field. Specifically, these tools allow simple calculation of the area within a field where desired photo-z completeness is met (given the measured photometric quality as a function of magnitude). Or alternatively, the same tools can be used to calculate the magnitude selection at every position in the field that is required to meet a desired photo-z selection criteria – accounting for variable photo-z quality across a field due to heterogeneous photometric coverage. Full details of the motivation and method will be presented in Duncan et al. (in preparation). As part of HELP Data Release 1, we provide working example notebooks for both the generation and exploitation of these HEALPIX based photo-z selection functions.

4.6.5 *Future HELP photo-z plans*

The photo-z estimation methodology applied to majority of HELP fields and sources (i.e. Duncan et al. 2018a, b) was developed with the aim of providing the optimum photo-z estimate given the available data, regardless of source type. However, the requirement for template fitting (and machine learning), hierarchical Bayesian combination and detailed posterior. This difficulty is particularly acute in the largest and most complicated data sets (e.g. *Herschel* Stripe 82) where significant high-performance computing resources were required for the runs.

Going forward, we will therefore aim to move to a more scaleable and automated approach that can exploit the infrastructure provided by HELP. The ingestion of the homogenized optical data sets into the HELP virtual observatory now opens up possibilities for optimized machine-learning based photo-zs that combine all spectroscopic redshifts available for a given combination of optical to near-IR photometry, regardless of field. By combining this unified resource with, for example, updated GPz algorithm that naturally allow for redshift predictions in the case of missing data (Almosallam et al., in preparation), it will be possible to provide comparable high quality photo-z estimates in a more scalable manner. Additionally, on-demand computation within the data base could be provided following the approach presented in Beck et al. (2017).

4.7 Physical modelling

Physical modelling of *master list* sources with both a redshift estimate and FIR data was carried out using Code Investigating GALaxy Emission[7] (CIGALE; Boquien et al. 2019). We refer to Burgarella, Buat & Iglesias-Páramo (2005), Noll et al. (2009), Boquien et al. (2019) for a detailed description of the code, and Małek et al. (2018) for a detailed description of the range of parameters used within the HELP fits but describe the key features here.

CIGALE conserves the energy balance between dust absorption in the UV to near-infrared domain and emission in the mid and far IR when generating SEDs. The stellar emission is constructed from composite stellar populations from simple stellar populations (SSP) combined with flexible star formation histories (SFH). Attenuation

[7]https: //cigale.lam.fr

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

curves are then applied to estimate the fraction of energy from stars and gas absorbed and re-emitted by the dust using a dust emission template. CIGALE is a highly flexible code, containing multiple different modules for SSP, SFH, dust attenuation, dust emission and AGN component for users to apply or they can add their own modules.

For the HELP project, we selected low resolution (Bruzual & Charlot 2003) single stellar population models, assuming a Chabrier (2003) initial mass function. SFHs are chosen to have the so-called delayed τ parametrization ($\propto t\, e^{-t/\tau}$) form, with additional bursts to allow the SFH to model starburst populations on top of older (potentially quiescent) populations. The SFH is defined as

$$\mathrm{SFR(t)} \propto \begin{cases} SFR_{\mathrm{delayed}}(t) & \text{if } t < t_0. \\ SFR_{\mathrm{delayed}}(t) + SFR_{\mathrm{burst}}(t), & \text{if } t \geq t_0. \end{cases} \quad (5)$$

with t_0 being the age of onset of the second episode of star formation.

We perform SED fitting runs with Charlot & Fall (2000) model as the dust attenuation recipe. The Charlot & Fall (2000) law assumed that stars are formed in interstellar birth clouds (BC), and after 10^7 yr, young stars disrupt their 'nursery' and migrate into the ambient interstellar medium (ISM). Both regions, BC and ISM, are characterized by a different power law – one representing dust attenuation in the BC and the second in the ISM for older stars. As a result, the emission from young stellar population is attenuated first in the BC and then it goes through the dust in the ISM. Stars older than 10^7 yr are attenuated only in the ISM. Different power slopes for BC and ISM can be used, but here we chose to keep both power-law slopes of the attenuation fixed at -0.7, as in the original Charlot & Fall (2000) work.

To model the IR SEDs of the HELP galaxies, we use a dust emission model where the majority of the dust is heated by a radiation field from the diffuse interstellar medium, while a much smaller fraction of dust is illuminated by the starlight. AGNs can substantially contribute to the mid-IR emission. To improve the accuracy of derived galaxy properties and because we have data coverage from optical to FIR, we add an AGN component to the SED modelling, using the dusty torus models of Fritz, Franceschini & Hatziminaoglou (2006).

Based on the five components a large grid of models is fitted to the data. The number of parameter values depends on the properties of the field containing between 50 and 100 million individual models. The physical properties published as a part of HELP DR1; dust luminosity (L_{dust}), stellar mass (M_{star}), and SFR, are built from the probability distribution function and for each parameter, the likelihood-weighted mean and standard deviations are calculated. These measurements have already been used for science purposes providing validation of the method against the literature (Ocran et al. 2021).

For each fitted galaxy, we calculate two values of the reduced χ^2 for the best-fitting template and photometric measurements: OPT_{χ^2} (for wavelengths lower than 8 μm restframe) and IR_{χ^2} (for wavelengths larger than 8 μm restframe). Using these two χ^2 values enables identification of SED fitting failures or peculiar objects (see Małek et al. 2018). An example CIGALE fit is presented in Fig. 6.

CIGALE is run for all HELP galaxies with at least two optical and at least two near-IR detections. Additionally, to select the sample for which physical properties are estimated, we keep only sources with at least 2 far-IR measurements (signal-to-noise ratio ≥2). As the redshift is essential for the physical modelling process we used the photo-zs (as presented in Section 4.6). As SED fitting codes are sensitive to having one wavelength region overly weighted due to the presence of multiple measurements on a single passband. Constraining power from multiple measurements at similar passbands will dominate over other bands. When multiple measurements are available in similar passbands we take the deepest only as determined by the signal-to-noise ratio.

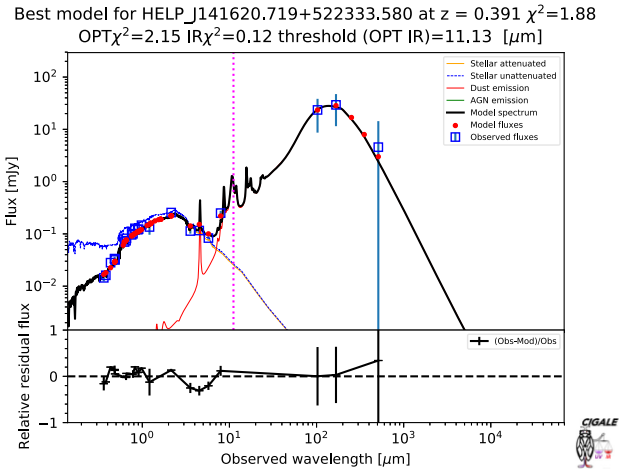


Figure 6. An exemplary fit of a galaxy from the HELP-EGS field. Observed fluxes are plotted with open blue squares. Filled red circles correspond to the model fluxes. The final model is plotted as a solid black line. Attenuated stellar component is plotted as a solid yellow line, while the unattenuated – as the dashed blue line. Red line mimics the dust emission. The relative residual fluxes, calculated as (observed flux − best model flux)/observed flux, are plotted at the bottom of each spectrum. Magenta dotted line represents 8 μm rest frame wavelength (here 11.13 μm for redshift 0.391). Obtained physical properties of the presented galaxy are following: $\log(M_{\mathrm{star}}) = 10.69 \pm 0.14[\mathrm{M}_\odot]$, $\log(L_{\mathrm{dust}}) = 11.15 \pm 0.14[\mathrm{L}_\odot]$, and $\log(\mathrm{SFR}) = 1.27 \pm 0.17[M_\odot\ \mathrm{yr}^{-1}]$.

On average, based on the selection described above, we estimated physical properties of 1.7 million galaxies or 1 per cent off all objects in the *master list*. Table 3 shows the summary of the catalogues in each field, including the number of sources for which M_{star}, L_{dust}, and SFR were estimated. To make HELP's SED fitting easily reproducible, we published a dedicated version of CIGALE, `cigalon`[8] which contains the modules and parameters used to fit HELP's galaxies. This version can be found on the main CIGALE page.

As it was shown in Małek et al. (2018), CIGALE's implementation of the energy balance enables predictions of L_{dust} for standard IR galaxies, which preserve energy budget, based on the UV to near-infrared data only. Here we would like to stress that CIGALE cannot estimate monochromatic fluxes with reliable uncertainties, but only the total value of L_{dust}. We used the predictions as priors for the IR extraction pipeline XID+ . A similar method was used by Pearson et al. (2018) for the star forming galaxies beyond the *Herschel* confusion limit. To predict the L_{dust} for galaxies without IR measurements, we run `cigalon` using the same parameters and methodology as described above but without the AGN module as without mid-IR data we are not able to constrain a reliable AGN component.

4.8 Data base structure and access

HELP data are distributed through the *Herschel Database in Marseille*[9] (HeDaM) in addition to the *Virtual Observatory at*

[8]https://gitlab.lam.fr/cigale/cigale.git
[9]https://hedam.lam.fr

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

Table 3. Summary of the HELP catalogue numbers on each field. The blind sources are measured completely independently of the *master list*. HDF-N, a very small field with extremely deep priors, has no objects selected for CIGALE fitting due to no objects being sufficiently bright in SPIRE bands to pass the signal to noise thresholds. This is reflected in the absence of a single blind detection on the field. Specialist fields such as this are subject to further development of more sophisticated priors. Starred photometric redshifts are from Zou et al. (2019).

Field	Objects	Area deg^2	XID +	Photo-z	CIGALE	Blind	spec-z
AKARI-NEP	531 746	9.2	31 441	∗107 228	1239	9848	1243
AKARI-SEP	844 172	8.7	108 119	∗139 059	566	20 169	362
Boötes	3 398 098	11.4	495 159	1 570 512	38 980	30 566	23 424
CDFS-SWIRE	2 171 051	13.0	283 406	136 944	9308	40 880	29 063
COSMOS	2 599 374	5.1	25 898	691 502	15 747	12 603	36 686
EGS	1 412 613	3.6	223 598	1 182 503	4159	9551	19 799
ELAIS-N1	4 026 292	13.5	269 611	2 714 686	49 985	34 501	4619
ELAIS-N2	1 783 240	9.2	86 591	∗120 723	6798	19 483	2471
ELAIS-S1	1 655 564	9.0	194 276	1 013 582	25 393	22 743	10 396
GAMA-09	12 937 982	62.0	1 386 659	8 833 874	130 293	112 461	38 407
GAMA-12	12 369 415	62.7	1 099 477	8 569 951	108 139	112 471	41 149
GAMA-15	14 232 880	61.7	1 236 395	10 083 210	117 234	116 436	81 413
HATLAS-NGP	6 759 591	177.7	1 233 547	3 166 952	185 290	344 635	58 476
HATLAS-SGP	29 790 690	294.6	3 511 594	17 054 138	352 804	497 501	47 213
HDF-N	130 679	0.67	834	∗7 435	0	0	3 360
Herschel-Stripe-82	50 196 455	363.2	2 976 447	21 509 448	250 644	232 589	132 358
Lockman-SWIRE	4 366 298	22.4	242 065	1 377 139	46 719	54 106	7243
SA13	9799	0.27	812	∗2884	70	315	188
SPIRE-NEP	2674	0.13	562	∗935	71	374	1
SSDF	12 661 903	111.1	4 395 253	9 250 727	305 576	196 895	1417
XMM–13hr	38 629	0.76	3 563	∗10 773	670	1 218	365
XMM–LSS	8 705 837	21.8	360 500	6 124 027	61 888	50 362	78 192
xFLS	977 148	7.4	52 187	∗100 993	5944	19 757	3562
Totals:	171 602 130	1269.1	18 217 994	93 769 225	1 717 517	1 939 464	621 407
Percentages:	–	–	10.6 per cent	54.6 per cent	1.0 per cent	–	0.4 per cent

susseX[10] (VOX). The former allows access to all raw data for reprocessing or direct handling. The latter permits fast queries over the full HELP area to generate samples for scientific analysis. Data accessed by code on GitHub can be found via its relative link on HeDaM such that the user can download the entire data base and perform a full rerun. We also include meta data files in the Git repository with links to the corresponding data files.

4.8.1 HeDaM catalogues and images

On HeDaM each data product is organized by field. For each field there is a final catalogue containing all the information from the optical fluxes to the infrared, the redshifts, and the physical parameters derived with SED fitting. The HELP *Herschel* SPIRE and PACS images are also present in addition to the blind sources associated with each. The data base is designed to run across the entire HELP sky. Many scientific users will be interested in an individual field of interest. We have therefore provided overviews of each field to help a new user become familiar with the data presented there.

HeDaM also provides everything to reprocess the data exactly as described here. This facilitates rerunning the HELP work while changing some parameters or, for instance, adding a new optical catalogue to the process. As we have emphasized throughout, all our code is available on GitHub as Python modules and Jupyter notebooks. For storage reasons, the data files are not included in the Github repository. HeDaM contains a file storage with the exact same structure but with the data files within. For many science cases, the final merged catalogues with summary information on every galaxy will be sufficient.

4.8.2 Virtual observatory

The Virtual Observatory at susseX (VOX) is a virtual observatory server built using the German Astrophysical Virtual Observatory (GAVO) DaCHS software: *Data Center Helper Suite*[11] (Demleitner et al. 2014; Demleitner 2018). VOX contains both the images and the catalogue data.

Images are available through the *Simple Image Access Protocol* (SIAP). In particular, VOX makes is possible to get image cutouts at a given position.

The catalogue data are gathered into a single table across all the coverage that users can query using the *Table Access Protocol* (TAP) with compliant software like TOPCAT (Taylor 2005), STILTS (Taylor 2006b), or PYVO.[12] This allows users to make sophisticated queries or to remotely cross-match their catalogues with HELP data.

The total catalogue containing all photometry measurements across all fields has around 500 columns where each source may only have flux information in a small subset of these bands. We therefore also provide a 'best' photometry catalogue which contains the lowest error measurement in each $ugrizyJHKK_s$ band in addition to the far-infrared fluxes. This reduces the number of columns to around 50. Due to the reduction in size it is also possible to index every column allowing fast queries. If the user requires the full

[10] https://herschel-vos.phys.sussex.ac.uk

[11] https://dachs-doc.readthedocs.io

[12] https://pyvo.readthedocs.io

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

photometry measurement they can then join their selection to the main table. For people working on a specific field they might prefer to download the full catalogue table on the field from HeDaM. VOX is particularly helpful when looking at samples scattered across a large area where it is unfeasible to download the full catalogue to perform cross-matching.

5 RESULTS

In this section we summarize the quality and sensitivities of the DR1 catalogue products. Table 3 gives an overview of the numbers of processed objects and areas associated with each field. All fields have been fully processed through the HELP pipeline. Each field has different features which determine the depths and quality of forced far-infrared fluxes and calculated physical parameters. Our aim is to allow these features to be modelled automatically across the whole HELP area as much as possible when users are constructing samples. Depending on the scientific question at hand, the desired sample properties can range from complete but heterogeneous samples over large areas/multiple fields to homogeneous samples within individual fields (and any permutation in between). Constructing the selection functions associated with these varying samples can be facilitated using the depth maps described and presented in Shirley et al. (2019) and here with the additional far-infrared bands.

Fig. 7 shows the differential number counts in the *Herschel* bands PACS 100, PACS 160, SPIRE 250, SPIRE 350, and SPIRE 500. Fig. 8 shows the relative fractions and numbers of each type of object in the final catalogue. Together these figures give an overview of the galaxy numbers as they pass through the pipeline.

5.1 Summary of *master list* and depths

The *master list* number counts as a function of the optical and near-infrared band magnitude on each field are summarized in Shirley et al. (2019). We also show the basic catalogue statistics here in Table 3. Defining depth for the forced photometry is dependent on the various factors contributing to the selection function. Here we take a similar approach to the depth maps described in Shirley et al. (2019) and compute mean errors for objects with signal to noise above 2. This gives a metric analogous to the traditional notion of depth for optical and NIR detection images and allows us to compare limits on the faintness of objects across the full wavelength range. Nevertheless, the effective far-infrared flux depth is in turn dependent on the depths of all the bands contributing to the determination of the *prior list*.

Fig. 9 shows the cumulative depths of MIPS, PACS, and SPIRE coverage. Fig. 10 shows an overview of depths for each band compared to a typical HELP ULIRG galaxy SED. Figs 11 to 13 show areas available to a given depth or deeper as a function of various optical to FIR band depths to illustrate how band depth are correlated with each other. These figures show the complexity of selection effects and its high dimensionality, often being dependent on over five band depths in addition to the shape of the SED. The additional data products presented here facilitate selection function modelling. The method used to model selection effects must be targeted to a given science goq given physical property of interest will be impacted by different selection effects depending on how it is correlated with the relevant detection bands and requirements made by each processing stage.

5.2 Summary of XID+ catalogues

The number of objects that have been deblended with XID+ (for MIPS, PACS, and or SPIRE) in each field can be seen in Table 3. We provide the number counts in the PACS and SPIRE bands produced by the forced photometry presented here on each field in Fig. 7. These number counts are jointly determined by the depth of the images and the *prior list*. These numbers go beyond the confusion limit that determine the number counts for blind source extracted catalogues. The flux cuts for each field, described in Section 4.3.1 can be found in Table 2.

5.3 Summary of blind catalogues

The number of blind SPIRE objects in each field can be seen in Table 3. In total there are 1.9 million blind sources across the HELP fields. We have compared these objects to previous blind catalogues (Smith et al. 2012; Wang et al. 2014; ESA 2017) and find over 90 per cent overlap between them. The previous blind catalogues from HerMES and HATLAS have been used to identify numerous high star-forming high redshift galaxies (e.g. Riechers 2013; Asboth et al. 2015) and candidates for follow-up programs (Duivenvoorden et al. 2018). The HELP blind catalogues, across all 1270 deg^2 are a data product for which many more of the rare, highly star forming, high redshift candidates can be identified. They are also a useful comparison with the forced XID+ catalogues in understanding the multiplicity of SPIRE sources (Scudder et al. 2016) and the impact of the prior catalogues used for forced photometry.

5.4 Summary of CIGALE physical properties

The main physical properties were estimated for all galaxies with known redshift (spectroscopic or photometric), with at least two detections in the optical range, two detections in NIR range, and at least two FIR measurements (PACS and/or SPIRE) with SNR>2. We call this the *A list*. The CIGALE code was used to estimate dust luminosity, star formation rate, stellar mass, and the AGN contribution to dust luminosity. Alongside the physical properties obtained from the full UV to far-infrared SED fitting, three different values of χ^2s are provided: reduced χ^2_r, quantifying the global quality of the SED fit for each galaxy, and OPT_{χ^2} and IR_{χ^2}, described in Section 4.7 (detailed description can be found in Małek et al. 2018). Those quantities can help us to identify possible interesting sources or modelling failures.

Fig. 14 shows the distribution of estimated physical properties: dust luminosity, star formation rate, and stellar mass as a function of redshift. These all-sky samples are drawn from vastly different areas of sky in terms of depths and areas. To illustrate this variation in dynamic range, we overlay contours between the wide and deep *Herschel* Stripe 82 field and the narrow deep field Boötes in Fig. 14. Those figures show the uniqueness of HELP in providing data for extremely bright IR sources at high redshift, and normal star forming galaxies at low redshift. Around 90 per cent of HELP galaxies with SED fits have dust luminosity larger than $10^{11}\,\mathrm{L}_\odot$. The majority of IR galaxies belong to Luminous Infra Red Galaxies (75 per cent), 15 per cent are classified as Ultra Luminous Infra Red Galaxies. Using our SED fitting procedure, we found more than 3500 (0.2 per cent) IR galaxies with dust luminosity larger than $10^{13}\,\mathrm{L}_\odot$ – those extraordinarily bright and active objects are still rare and

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

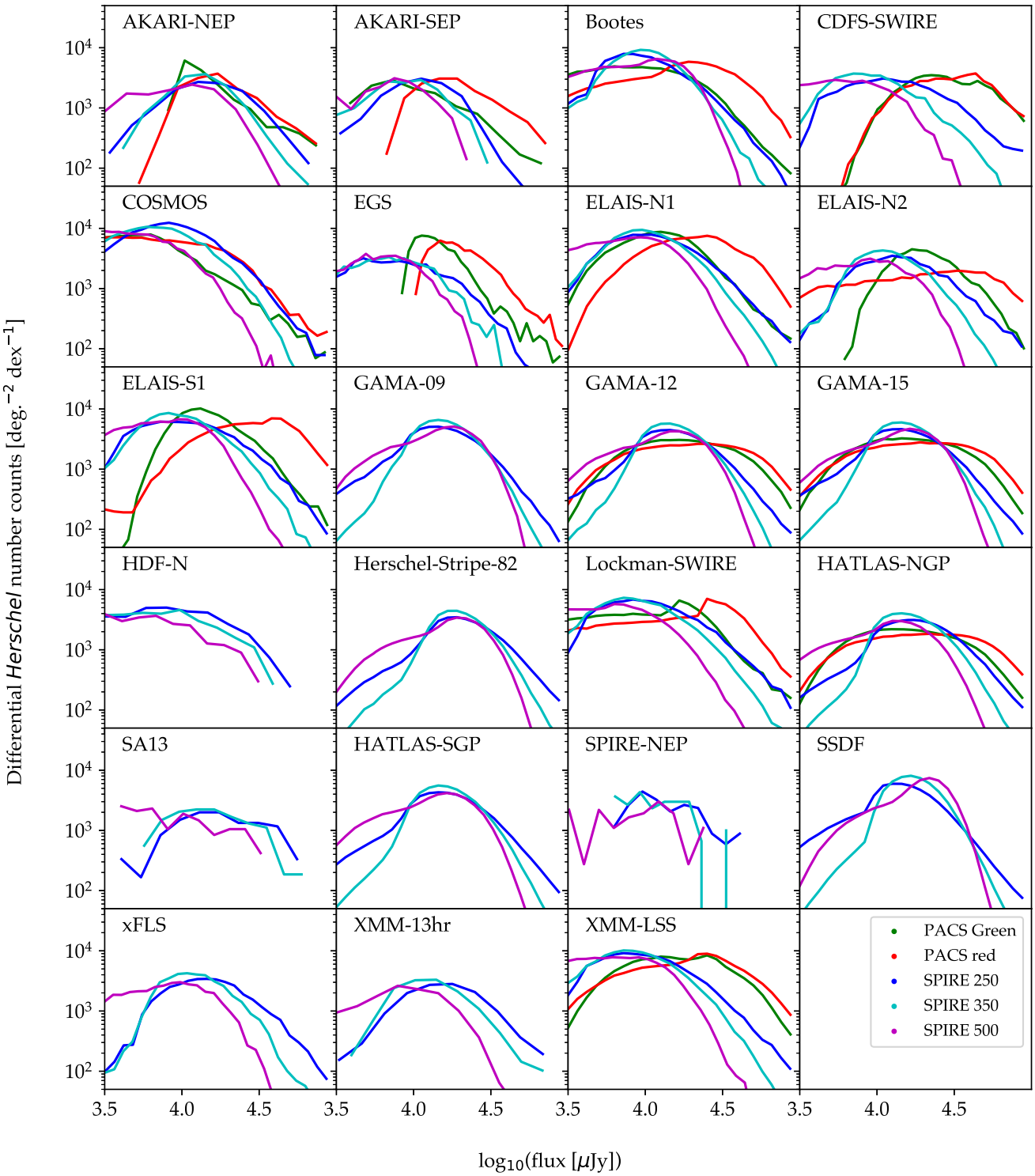


Figure 7. The differential number counts in *Herschel* bands on each field. Some of the small fields suffer from small number statistics. The normalization factors for each field are computed based on the area over which there are any XID+ priors. The size of this area is dependent on what XID+ prior used for each field, and can cover less than half of the field depending (e.g. IRAC based priors).

not well understood (e.g. Rowan-Robinson 2000; Farrah et al. 2002; Wang et al. 2020). The most extreme IR sources are also very active in star formation processes, especially at high redshift. This selection effect can also be seen in Fig. 14(a).

6 DISCUSSION AND CONCLUSION

Understanding galaxy formation and evolution requires measurements from many facilities to trace the different physical processes such as star formation and AGN activity in galaxies over cosmic

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

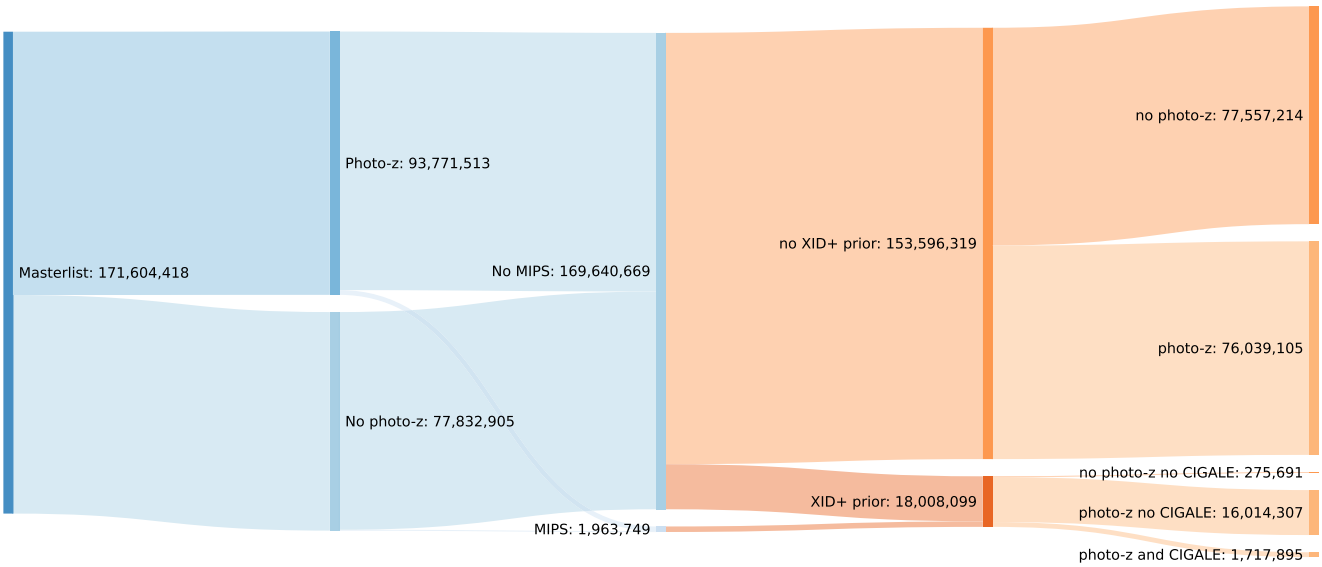


Figure 8. Overview of the final numbers of objects showing the fraction that have a given measurement. There are broadly two types of object that are in the XID + SPIRE prior; those with a MIPS detection in the MIPS area and those with an $L_{\rm dust}$ prediction prior in the regions without MIPS. The final stage shows the relative numbers which will constitute typical samples.

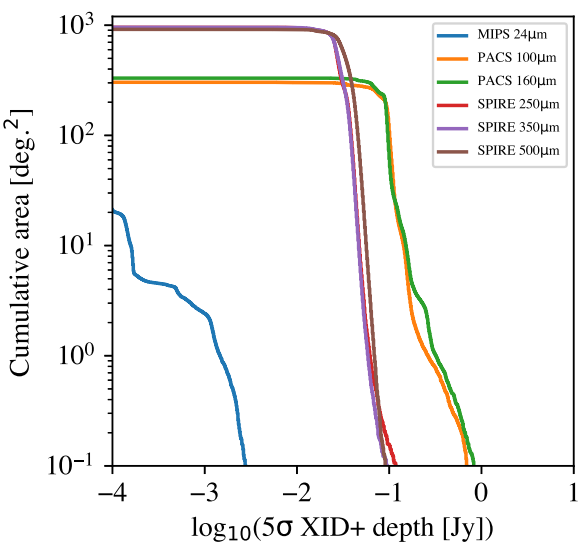


Figure 9. Cumulative area to a given depth or deeper for the XID+ forced photometry for far-infrared fluxes in the MIPS 24, PACS Green, Red, SPIRE 250, 350, and 500 bands.

time. These data need to cover large areas of the sky to obtain the large samples to characterize the population properties. These data sets need to be homogeneous and well understood to enable the statistical studies required.

Construction of such a data set has been the main goal of the HELP project. The corresponding data release documented in this paper describes HELP's first attempt at solving the challenges in the collation and merging of large, disparate, complex multiwavelength data sets. Focusing on the *Herschel* fields for our first data release, DR1, HELP collates and homogeneous optical and near-infrared catalogues over 1270 deg^2 and applies novel methods to provide reliable photometry from the inherently source-confused far-infrared images. Making as much use as possible of the *Herschel* data remains an important consideration, given this data captures most of the CIRB over a wavelength range for which no new instrument is planned for the foreseeable future.

6.1 Enabling science

The ultimate validation of the HELP data will be through its utilization for new scientific analysis.

The collation of data over many fields enables scientific investigations with larger statistical samples. Duivenvoorden et al. (2020) used the collated *master list* and *Herschel* SPIRE images to investigate how the contribution to the CIRB varies with galaxies selected at different wavelengths.

The novel methods of extracting information from low resolution far-infrared data also enables new investigations. Scudder et al. (2016, 2018) used the full posterior probability distribution provided by XID+ to show how the multiplicity of *Herschel* SPIRE sources changes with flux.

The large-area means that rare objects are included, and the wide variety of data enables these to be discovered. An example can be found in the hyper luminous obscured quasar recently discovered at $z \sim 4$ (Efstathiou et al. 2021).

The far-infrared data, uniquely available in the DR1 fields, permits studies of massive dusty star-forming galaxies (DSFGs, e.g. Weiß et al. 2013; Casey et al. 2014). These galaxies play an important role in understanding galaxy evolution because the dust-obscured star formation activity becomes more important at higher redshift (Donevski et al. 2020). They are also crucial to understand how massive galaxies assembled (Hamed et al. 2021). Furthermore, a large sample of infrared galaxies spanning a wide redshift range allow us to study the effect of dust attenuation on the physical properties of galaxies (Lo Faro et al. 2017; Buat et al. 2018, 2019)

The HELP data products cover many significant extragalactic survey fields, they are a valuable resource for existing surveys such as the *Spitzer* Extragalactic Representative Volume Survey (SERVS)

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

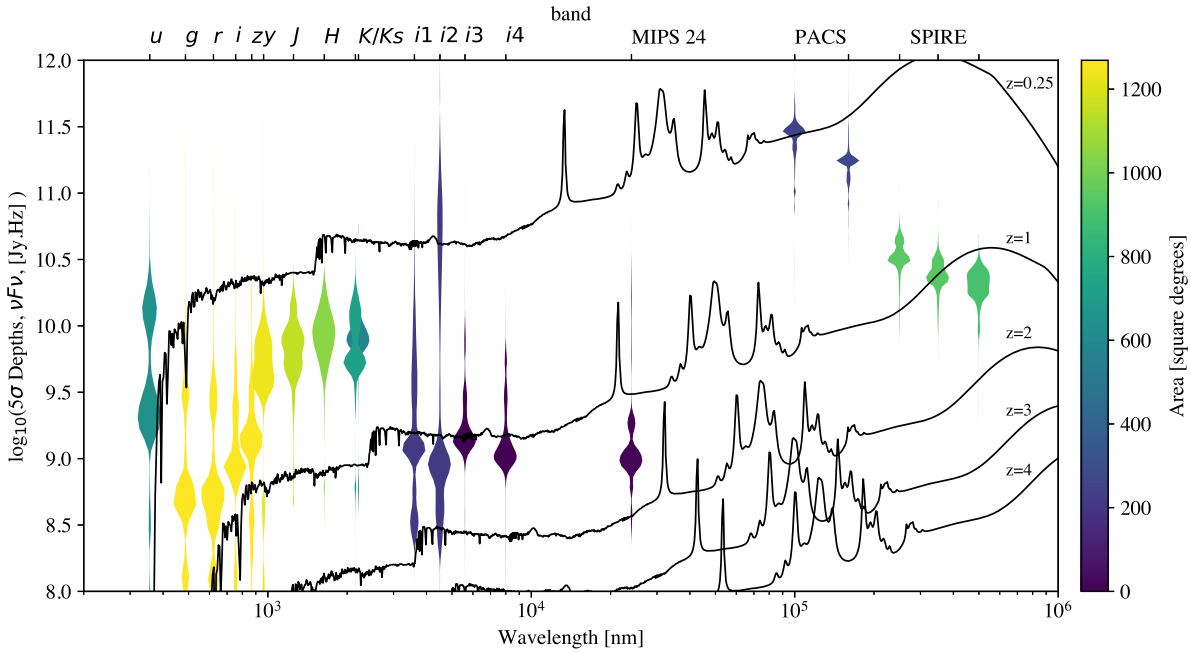


Figure 10. The distribution of area on the sky to a given depth, $\nu F\nu$, shown via a violin plot, for each broad-band type (taking the deepest specific band available in a given HEALPIX cell). Optical and NIR depths are 5σ depths in a 2 arcsec aperture. MIPS, PACS, and SPIRE depths are 5σ depths for the XID+ forced photometry values. The colour of each area is determined by the total area that data for that band is available. We also plot a HELP spectral energy distribution for a ULIRG galaxy with star formation rate of 200 $M_\odot$ yr^{-1} and a stellar mass of 10^{10} $M_\odot$ at various redshifts.

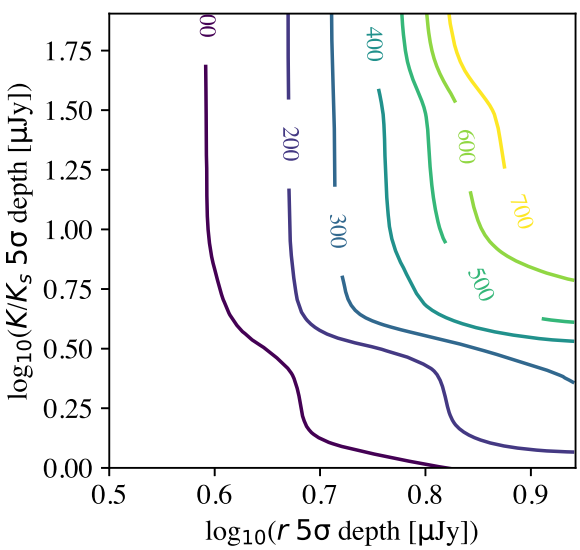


Figure 11. Area on the sky available to a given depth or deeper as a function of r band and K/K_s band depth.

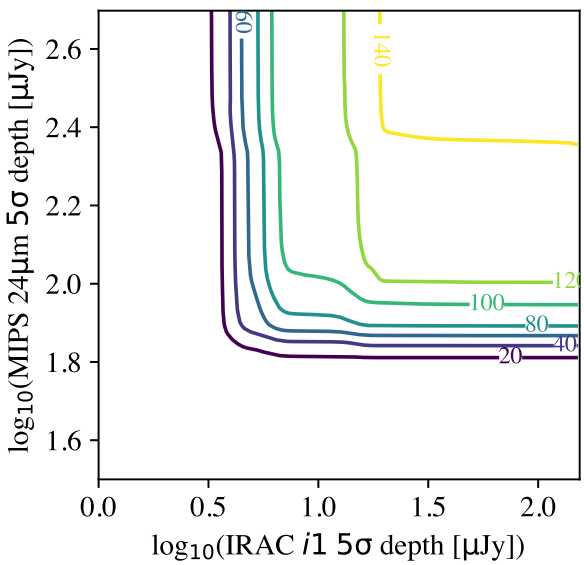


Figure 12. Area on the sky available to a given depth or deeper as a function of IRAC $i1$ band and MIPS 24 μm band depth.

and the DeepDrill survey (Mauduit et al. 2012). Ongoing and future surveys will also be able to exploit and build on the HELP data products. The LOFAR deep field team have used the HELP products directly (e.g. Gloudemans et al. 2020; Smith et al. 2020; Wang et al. 2020) and together with HELP team, exploring the far-infrared radio correlation (McCheyne, in preparation). The MeerKAT International GHz Tiered Extragalactic Exploration (MIGHTEE; Jarvis et al. 2016), is another ongoing radio survey with fields that overlaps with HELP and so products like the deblended FIR XID+ fluxes will be a valuable resource e.g. in deeper exploration of the far-infrared-radio correlation.

The wide spectrum of physical properties can be also used to design or simulate upcoming large surveys such as the Legacy Survey of Space and Time (LSST) (Riccio, in preparation). The selected LSST Deep Drilling fields (ELAIS S1, *XMM–LSS*, Extended Chandra Deep Field-South and COSMOS) are also covered by HELP and so data products these fields could provide immediate inputs to planning and direct contributions to early science.

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

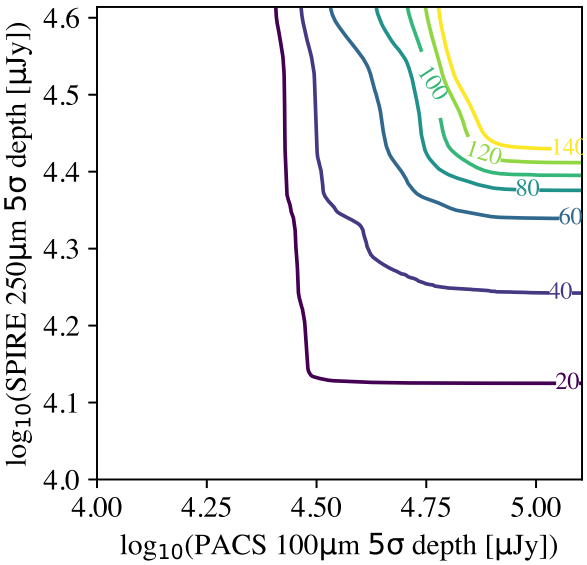


Figure 13. Area on the sky available to a given depth or deeper as a function of PACS 100 μm band and SPIRE 250 μm band depth.

The large area and depth of HELP can enable more robust analysis of the fundamental statistical relations. However such analyses must take into account the associated selection functions. Modelling the full selection functions for physical properties is now the focus of HELP research. Modelling how fluxes propagate through to redshifts and other physical properties. We have aimed to provide the necessary information throughout the HELP pipeline to make this more tractable. Campos Varillas (in preparation) will develop empirical methods to estimate the HELP selection function in order to probe the bright end of the stellar mass function of galaxies in *Herschel* Stripe 82.

6.2 Future data releases and how to contribute

HELP's open approach means decisions made during the collation and production of DR1 are transparent and can be reproduced by other teams. It also provides a pathway for any astronomers to contribute and add to the HELP data set in the future. As optical, infrared, and radio surveys produce new data sets, these will need to be combined with legacy surveys. The methods, pipelines, and tools described in this paper provide a platform to enable their ingestion.

As further data sets are incorporated it will also be possible to improve the *prior list* and the subsequent photometry and the SED fitting. The open source nature of the HELP pipelines and tools means these improvements can be done by anyone and are not dependent on the current HELP team. For DR1 we decided to generate the most versatile data products, with a non-informative prior, but specific science may benefit from using the full optical to mid-infrared photometry and SED fitting to use more informative flux priors.

For HELP to continue being a valuable resource there will need to be a transition from a centrally managed project to becoming a distributed community endeavour. Using version control systems such as GitHub which have enabled many collaborators to contribute to open source projects will be one way in which such a community can be fostered. We are working with other survey teams and training them in the use of the HELP pipelines and tools. We welcome new teams interested in combining their observational data, or value-added related data sets (e.g. alternative physical parameter catalogues, galaxy cluster catalogues etc) to contribute and integrate their data.

6.3 Summary

We present the HELP project which collates extragalactic surveys from the optical to *Herschel* far-infrared bands. This includes an open source software pipeline as well as the resultant data products. This first data release, DR1, can be used to study an unprecedented wide area of *Herschel* sky. Some key highlights of this new data set are:

(i) We have collated data for 170 million objects from optical to far-infrared over 1270 deg^2 of the prime extragalactic fields, with boundaries defined by the mapping of the *Herschel* Space observatory.

(ii) We present far-infrared photometry for 18 million objects in an optical to mid-infrared selected *prior list* chosen to be tightly correlated with far-infrared bright objects. We calculate a posterior distribution on the flux for all objects using Bayesian inference.

(iii) We publish the main physical properties; stellar mass, dust luminosity, and star formation rate based on spectral energy distribution modelling with the CIGALE code. This is done for all galaxies with at least two detections in each wavelength region for a total of 1.7 million objects, being 1 per cent of the total HELP catalogue.

(iv) The new catalogue is presented alongside an array of other data products, including newly homogenized images, supplementary catalogues, and extensive tools for accessing and analysing these new data sets.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Manda Banerji, Michelle Cluver, Duncan Farrah, Peter Hatfield, Tom Jarrett, Emeric Le Floc'h, Richard McMahon, Mark Sargent, Roberto Scipioni, Dan Smith, Elisabetta Valiante, Ivan Valtchanov, Stephen Wilkins, and Louise Winters for useful conversations or important contributions to the project that do not form a specific part of this paper. We are grateful to Stephen Serjeant for useful comments on the manuscript while under review.

The research leading to these results has received funding from the European Union Seventh Framework Programme FP7/2007-2013/ under grant agreement No. 607254 (The *Herschel* Extragalactic Legacy Project). This publication reflects only the authors' view and the European Union is not responsible for any use that may be made of the information contained therein.

SO acknowledges support from the Science and Technology Facilities Council (grant number ST/L000652/1)

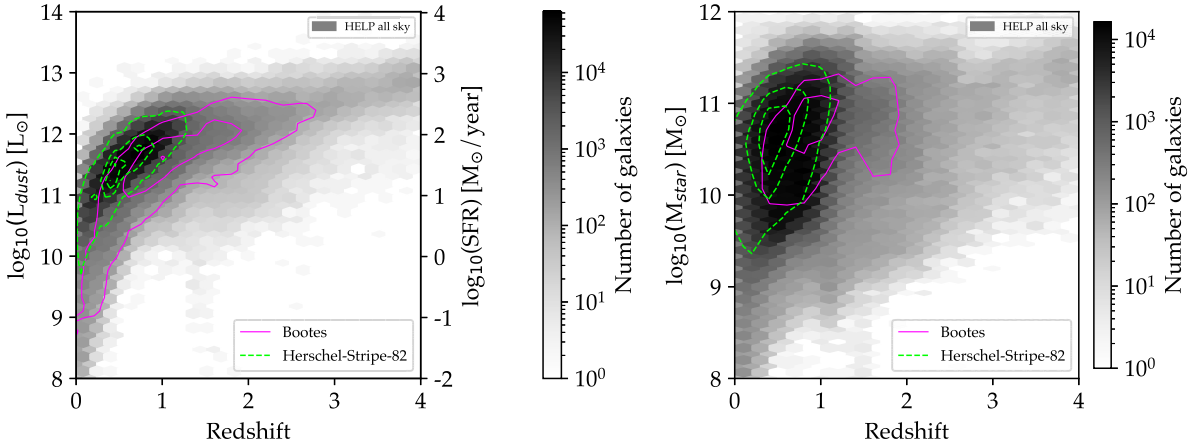
KM has been supported by the Polish National Science Centre (UMO-2018/30/E/ST9/00082).

MJ, LM, MP, and MV acknowledge support from the South African Department of Science and Technology (DST/CON 0134/2014).

MP and MV acknowledge financial support from the Inter-University Institute for Data Intensive Astronomy (IDIA). IDIA is a partnership of the University of Cape Town, the University of Pretoria, the University of the Western Cape, and the South African Radio Astronomy Observatory.

VB has received funding from the Excellence Initiative of Aix-Marseille University-AMIDEX, a French 'Investissement d'Avenir' programme.

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022



(a) Dust luminosity and star formation rate (b) Stellar mass

Figure 14. Overview of dust luminosity, star formation rate, and stellar mass estimated using CIGALE SED fitting tool as a function of redshift across the full HELP sky. As an example, we overlay contours between the wide *Herschel* Stripe 82 field and the narrow deep field Boötes. Star formation rates are scaled to dust luminosity using the median ratio.

HCSS / HSpot / HIPE are joint developments by the *Herschel* Science Ground Segment Consortium, consisting of ESA, the NASA *Herschel* Science Center, and the HIFI, PACS, and SPIRE consortia.

This research made extensive use of TOPCAT (Taylor 2005), STILTS (Taylor 2006b), and ASTROPY (Astropy Collaboration 2018).

This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

This work has made use of the COSMOS2015 catalogue (Scoville et al. 2007) based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under ESO programme ID 179.A-2005 and on data products produced by TERAPIX and the Cambridge Astronomy Survey Unit on behalf of the UltraVISTA consortium.

SPIRE has been developed by a consortium of institutes led by Cardiff Univ. (UK) and including Univ. Lethbridge (Canada); NAOC (China); CEA, LAM (France); IFSI, Univ. Padua (Italy); IAC (Spain); Stockholm Observatory (Sweden); Imperial College London, RAL, UCL-MSSL, UKATC, Univ. Sussex (UK); Caltech, JPL, NHSC, Univ. Colorado (USA). This development has been supported by national funding agencies: CSA (Canada); NAOC (China); CEA, CNES, CNRS (France); ASI (Italy); MCINN (Spain); SNSB (Sweden); STFC, UKSA (UK); and NASA (USA).

HELP would like to thank the HELP Scientific Advisory Board members past and present for invaluable advice in defining the project: Simon Driver (chair), Loretta Dunne, Carol Lonsdale, Mark Lacy, Peter Capak, Takashi Onaka, Mara Salvato, Brent Groves, Göran Pilbratt, and David Elbaz.

The data presented in this paper is released through the *Herschel* Data base in Marseille HeDaM (http://hedam.lam.fr/HELP)

Huge thanks also to our Project Manager, Louise Winters, for keeping us on-track and on-time with good humour.

Finally, we would like to take this opportunity to remember Natalie Christopher who passed away in 2019 August. Natalie worked on HELP from the start and her contribution to astronomy is sorely missed. Many have described how Natalie's unique vision of making astronomy accessible to all, coupled to her recognition of the power of astronomy to bring people together through their shared common interest, have had a profound impact.

Herschel is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA.

Author contributions: RS made the largest contribution to the data products, analysis, and led the writing of text in this paper and is thus identified as first author. SJO defined the project, secured the funding, and led the team in the development and execution of the project and is identified in the senior author (last) position. KD, PDH, KM, MCC-V, YR, and MWLS contributed very significantly and equally to the production of the data products, the analysis, and the writing. The other authors all contributed significantly to either the design of the project, some data products, validation, analysis, or writing of the paper.

DATA AVAILABILITY

The data underlying this article are available in HeDaM at http://hedam.lam.fr/HELP/, and can be accessed publicly.

REFERENCES

Almosallam I. A., Lindsay S. N., Jarvis M. J., Roberts S. J., 2016a, MNRAS, 455, 2387
Almosallam I. A., Jarvis M. J., Roberts S. J., 2016b, MNRAS, 462, 726
Asboth V. et al., 2015, MNRAS, 462, 1989
Astropy Collaboration, 2018, AJ, 156, 123
Baldry I. K. et al., 2018, MNRAS, 474, 3875
Balestra I. et al., 2010, A&A, 512, A12
Barger A. J., Cowie L. L., Wang W.-H., 2008, ApJ, 689, 687
Beck R., Dobos L., Budavári T., Szalay A. S., Csabai I., 2017, Astron. Comput., 19, 34
Berta S. et al., 2007, A&A, 467, 565
Blanton M. R. et al., 2017, AJ, 154, 28
Boller T., Freyberg M. J., Trümper J., Haberl F., Voges W., Nandra K., 2016, A&A, 588, A103
Boquien M., Burgarella D., Roehlly Y., Buat V., Ciesla L., Corre D., Inoue A. K., Salas H., 2019, A&A, 622, A103
Bradshaw E. J. et al., 2013, MNRAS, 433, 194

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503
Brown M. J. I. et al., 2014, ApJS, 212, 18
Brown A. G. et al., 2016, A&A, 595, A2
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Buat V., Boquien M., Małek K., Corre D., Salas H., Roehlly Y., Shirley R., Efstathiou A., 2018, A&A, 619, A135
Buat V., Ciesla L., Boquien M., Małek K., Burgarella D., 2019, A&A, 632, A79
Burgarella D., Buat V., Iglesias-Páramo J., 2005, MNRAS, 360, 1413
Cardamone C. N. et al., 2010, ApJS, 189, 270
Casey C. M. et al., 2014, ApJ, 796, 95
Castellano M. et al., 2016, A&A, 590, A31
Chabrier G., 2003, PASP, 115, 763
Chapin E. L. et al., 2011, MNRAS, 411, 505
Chapman S. C., Smail I., Blain A. W., Ivison R. J., 2004, ApJ, 614, 671
Chapman S. C., Blain A. W., Smail I., Ivison R. J., 2005, ApJ, 622, 772
Charlot S., Fall S. M., 2000, ApJ, 539, 718
Childress M. J. et al., 2017, MNRAS, 472, 273
Chincarini G., Tarenghi M., Sol H., Crane P., Manousoyannaki I., Materne J., 1984, A&AS, 57, 1
Cohen J. G., Hogg D. W., Blandford R., Cowie L. L., Hu E., Songaila A., Shopbell P., Richberg K., 2000, ApJ, 538, 29
Coil A. L. et al., 2011, ApJ, 741, 8
Colless M. et al., 2001, MNRAS, 328, 1039
Comparat J. et al., 2015, A&A, 575, A40
Cooper M. C. et al., 2011, ApJS, 193, 14
Cooper M. C. et al., 2012, MNRAS, 425, 2116
Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, AJ, 112, 839
Cowie L. L., Barger A. J., Hu E. M., Capak P., Songaila A., 2004, AJ, 127, 3137
Cristiani S., D'Odorico V., 2000, AJ, 120, 1648
Croom S. M., Warren S. J., Glazebrook K., 2001, MNRAS, 328, 150
Croom S. M., Smith R. J., Boyle B. J., Shanks T., Miller L., Outram P. J., Loaring N. S., 2004, MNRAS, 349, 1397
Croom S. M. et al., 2009, MNRAS, 392, 19
da Costa L. N. et al., 1998, AJ, 116, 1
Dahlen T. et al., 2013, ApJ, 775, 93
Dawson S., Stern D., Bunker A. J., Spinrad H., Dey A., 2001, AJ, 122, 598
Demleitner M., 2018, DaCHS: Data Center Helper Suite, record ascl:1804.005
Demleitner M., Neves M. C., Rothmaier F., Wambsganss J., 2014, Astron. Comput., 7, 27
Dickinson M. et al., 2004, ApJ, 600, L99
Donevski D. et al., 2020, A&A, 644, A144
Donley J. L. et al., 2012, ApJ, 748, 142
Dressler A., Smail I., Poggianti B. M., Butcher H., Couch W. J., Ellis R. S., Oemler A., Jr, 1999, ApJS, 122, 51
Drinkwater M. J. et al., 2010, MNRAS, 401, 1429
Driver S. P. et al., 2009, Astron. Geophys., 50, 5.12
Driver S. P. et al., 2011, MNRAS, 413, 971
Duivenvoorden S. et al., 2016, MNRAS, 462, 277
Duivenvoorden S. et al., 2018, MNRAS, 477, 1099
Duivenvoorden S. et al., 2020, MNRAS, 491, 1355
Duncan K. J. et al., 2018a, MNRAS, 473, 2655
Duncan K. J., Jarvis M. J., Brown M. J., Röttgering H. J., 2018b, MNRAS, 477, 5177
Duncan K. J. et al., 2019, A&A, 622, A3
Eales S. et al., 2010, PASP, 122, 499
Efstathiou A. et al., 2021, MNRAS, 503, L11
ESA, 2017, Herschel-SPIRE Point Source catalogue. Available at: http://archives.esac.esa.int/hsa/legacy/HPDP/SPIRE/SPIRE-P/SPSC/SPIREPointSourceCatalogExplanatorySupplementFull20170203.pdf
Falco E. E. et al., 1999, PASP, 111, 438
Farrah D., Verma A., Oliver S., Rowan-Robinson M., McMahon R., 2002, MNRAS, 329, 605
Fernique P., Boch T., Donaldson T., Durand D., O'Mullane W., Reinecke M., Taylor M., 2019, Technical report, MOC - HEALPix Multi-Order Coverage map, Version 1.1. Available at: http://www.ivoa.net/Documents/
Feruglio C. et al., 2008, A&A, 488, 417
Flesch E. W., 2015, Publ. Astron. Soc. Aust., 32, e010
Fritz J., Franceschini A., Hatziminaoglou E., 2006, MNRAS, 366, 767
Garcet O. et al., 2007, A&A, 474, 473
Gelman A., Meng X.-L., Stern H., 1996, Stat. Sin., 6 , 733
Gelman A., Carlin J. B., Stern H. S., Dunson D. B., Vehtari A., Rubin D. B., 2013, Bayesian Data Analysis. CRC Press, Boca Raton
Gloudemans A. J. et al., 2020, A&A, 648, A7
Griffin M. J. et al., 2010, A&A, 518, L3
Hamed M., Ciesla L., Béthermin M., Małek K., Daddi E., Sargent M. T., Gobat R., 2021, A&A, 646, A127
Herenz E. C. et al., 2017, A&A, 606, A12
Holder G. P. et al., 2013, ApJ, 771, L16
Hopkins A. M. et al., 2013, MNRAS, 430, 2047
Houck J. R. et al., 2005, ApJ, 622, L105
Huang J. S. et al., 2009, ApJ, 700, 183
Huchra J. P. et al., 2012, ApJS, 199, 26
Hurley P. D. et al., 2017, MNRAS, 464, 885
Ilbert O. et al., 2013, A&A, 556, A55
Jarvis M. et al., 2016, in MeerKAT Science: On the Pathway to the SKA. p. 6, preprint (arXiv:1709.01901)
Jeltema T. E., Mulchaey J. S., Lubin L. M., Fassnacht C. D., 2007, ApJ, 658, 865
Jones D. H. et al., 2004, MNRAS, 355, 747
Kochanek C. S. et al., 2012, ApJS, 200, 8
Kriek M. et al., 2008, ApJ, 677, 219
Kurk J. et al., 2013, A&A, 549, A63
Lacy M., Petric A. O., Sajina A., Canalizo G., Storrie-Lombardi L. J., Armus L., Fadda D., Marleau F. R., 2007, AJ, 133, 186
Lacy M. et al., 2013, ApJS, 208, 24
Laigle C. et al., 2016, ApJS, 224, 24
Le Fèvre O. et al., 2013, A&A, 559, A14
Levenson L. et al., 2010, MNRAS, 409, 83
Lidman C. et al., 2013, Publ. Astron. Soc. Austr., 30, e001
Lindegren L. et al., 2016, A&A, 595, A4
Liske J., Lemon D. J., Driver S. P., Cross N. J. G., Couch W. J., 2003, MNRAS, 344, 307
Liu C. T., Petry C. E., Impey C. D., Foltz C. B., 1999, AJ, 118, 1912
Lo Faro B., Buat V., Roehlly Y., Alvarez-Marquez J., Burgarella D., Silva L., Efstathiou A., 2017, MNRAS, 472, 1372
Loveday J., Peterson B. A., Maddox S. J., Efstathiou G., 1996, ApJS, 107, 201
Lutz D. et al., 2011, A&A, 532, A90
Maddox S. J. et al., 2018, ApJS, 236, 30
Małek K. et al., 2018, A&A, 620, A50
Mao M. Y. et al., 2012, MNRAS, 426, 3334
Marleau F. R., Fadda D., Appleton P. N., Noriega-Crespo A., Im M., Clancy D., 2007, ApJ, 663, 218
Masters D. C. et al., 2019, ApJ, 877, 81
Mauduit J. C. et al., 2012, PASP, 124, 714
Maza J., Ortiz P. F., Wischnjewsky M., Antezana R., González L. E., 1995, Rev. Mex. Astron. Astrofis., 31, 159
McLure R. J. et al., 2013, MNRAS, 428, 1088
McLure R. J. et al., 2018, MNRAS, 479, 25
Merlin E. et al., 2016, A&A, 590, A30
Mignoli M. et al., 2005, A&A, 437, 883
Momcheva I. G. et al., 2016, ApJS, 225, 27
Mountrichas G., Buat V., Yang G., Boquien M., Burgarella D., Ciesla L., 2021, A&A, 646, A29
Nandra K. et al., 2015, ApJS, 220, 10
Newman J. A. et al., 2013, ApJS, 208, 5
Nguyen H. T. et al., 2010, A&A, 518, L5
Noll S., Burgarella D., Giovannoli E., Buat V., Marcillac D., Muñoz-Mateos J. C., 2009, A&A, 507, 1793
Ocran E. F., Taylor A. R., Vaccari M., Ishwara-Chandra C. H., Prandoni I., Prescott M., Mancuso C., 2021, MNRAS, 500, 4685

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

Oliver S. J. et al., 2012, MNRAS, 424, 1614
Ono Y., Ouchi M., Shimasaku K., Dunlop J., Farrah D., McLure R., Okamura S., 2010, ApJ, 724, 1524
Onodera M. et al., 2015, ApJ, 808, 161
Ott S., 2010, in Mizumoto Y., Morita K. I., Ohishi M., eds, ASP Conf. Ser. Vol. 434, Astronomical Data Analysis Software and Systems XIX. p. Astron. Soc. Pac., San Francisco, 139
Ouchi M. et al., 2008, ApJS, 176, 301
Page M. J. et al., 2006, MNRAS, 369, 156
Papovich C. et al., 2006, AJ, 132, 231
Pearson C. et al., 2017a, Publ. Korean Astron. Soc., 32, 219
Pearson W. J., Wang L., van der Tak F. F. S., Hurley P. D., Burgarella D., Oliver S. J., 2017b, A&A, 603, A102
Pearson W. J. et al., 2018, A&A, 615, A146
Piazzo L., 2017, IEEE Trans. Image Process., 26, 5232
Piazzo L., Ikhenaode D., Natoli P., Pestalozzi M., Piacentini F., Traficante A., 2012, IEEE Trans. Image Process. , 21, 3687
Piazzo L., Panuzzo P., Pestalozzi M., 2015a, Signal Process., 108, 430
Piazzo L., Calzoletti L., Faustini F., Pestalozzi M., Pezzuto S., Elia D., di Giorgio A., Molinari S., 2015b, MNRAS, 447, 1471
Piazzo L., Raguso M. C., Carpio J. G., Altieri B., 2016a, 2016 24th European Signal Processing Conference (EUSIPCO). Budapest, Hungary,p. 1553
Piazzo L., Raguso M. C., Calzoletti L., Seu R., Altieri B., 2016b, IEEE Trans. Image Process., 25, 4458
Piffaretti R., Arnaud M., Pratt G. W., Pointecouteau E., Melin J. B., 2011, A&A, 534, A109
Pilbratt G. L. et al., 2010, A&A, 518, L1
Planck Collaboration XVI, 2014, A&A, 571, A12
Poglitsch A. et al., 2010, A&A, 518, L2
Pope A. et al., 2008, ApJ, 675, 1171
Ravikumar C. D. et al., 2007, A&A, 465, 1099
Reddy N. A., Steidel C. C., Erb D. K., Shapley A. E., Pettini M., 2006, ApJ, 653, 1004
Riechers D. A., 2013, Nature, 502, 459
Rieke G. H. et al., 2004, ApJS, 154, 25
Rodighiero G. et al., 2006, MNRAS, 371, 1891
Roseboom I. G. et al., 2012, MNRAS, 426, 1782
Rowan-Robinson M., 2000, MNRAS, 316, 885
Rowan-Robinson M., Gonzalez-Solares E., Vaccari M., Marchetti L., 2013, MNRAS, 428, 1958
Sacchi N. et al., 2009, ApJ, 703, 1778
Salvato M. et al., 2008, ApJ, 690, 1250
Salvato M. et al., 2011, ApJ, 742, 61
Salvato M. et al., 2018, MNRAS, 473, 4937
Sargsyan L. A., Weedman D. W., 2009, ApJ, 701, 1398
Saunders W. et al., 2000, MNRAS, 317, 55
Scodeggio M. et al., 2018, A&A, 609, A84
Scoville N. et al., 2007, ApJS, 172, 1
Scudder J. M., Oliver S., Hurley P. D., Griffin M., Sargent M. T., Scott D., Wang L., Wardlow J. L., 2016, MNRAS, 460, 1119
Scudder J. M., Oliver S., Hurley P. D., Wardlow J. L., Wang L., Farrah D., 2018, MNRAS, 480, 4124
Sedgwick C. et al., 2017, Publ. Korean Astron. Soc. , 32, 281
Shim H. et al., 2013, ApJS, 207, 37
Shirley R. et al., 2019, MNRAS, 490, 634
Silverman J. D. et al., 2010, ApJS, 191, 124
Smail I., Chapman S. C., Blain A. W., Ivison R. J., 2004, ApJ, 616, 71
Smith A. J. et al., 2012, MNRAS, 419, 377
Smith M. W. L. et al., 2017, ApJS, 233, 26
Smith D. J. B. et al., 2020, A&A, 648, A6
Stalin C. S., Petitjean P., Srianand R., Fox A. J., Coppolani F., Schwope A., 2010, MNRAS, 401, 294
Steffen A. T., Barger A. J., Capak P., Cowie L. L., Mushotzky R. F., Yang Y., 2004, AJ, 128, 1483
Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., Giavalisco M., 2003, ApJ, 592, 728
Strolger L.-G. et al., 2004, ApJ, 613, 200
Swinbank A. M., Smail I., Chapman S. C., Blain A. W., Ivison R. J., Keel W. C., 2004, ApJ, 617, 64
Swinbank A. M. et al., 2005, MNRAS, 359, 401
Swinbank A. M. et al., 2007, MNRAS, 379, 1343
Szokoly G. P. et al., 2004, ApJS, 155, 271
Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, ASP Conf. Ser. Vol. 347, Astronomical Data Analysis Software and Systems XIV. Astron. Soc. Pac., San Francisco, p. 29
Taylor M. B., 2006a, in Gabriel C., Arviset C., Ponz D., Solano E. , eds, ASP Conf. Ser. Vol. 351, STILTS - A Package for Command-Line Processing of Tabular Data. Astron. Soc. Pac., San Francisco, p. 666
Taylor M. B., 2006b, in Gabriel C., Arviset C., Ponz D., Enrique S., eds, ASP Conf. Ser. Vol. 351, Astronomical Data Analysis Software and Systems XV. Astron. Soc. Pac. , San Francisco, p. 666
Treister E. et al., 2009, ApJ, 693, 1713
Trichas M. et al., 2010, MNRAS, 405, 2243
Tully R. B., Shaya E. J., Karachentsev I. D., Courtois H. M., Kocevski D. D., Rizzi L., Peel A., 2008, ApJ, 676, 184
Vaccari M., 2015, in Proceedings of 'The many facets of extragalactic radio surveys: towards new scientific challenges' (EXTRA-RADSUR2015). 20-23 October 2015. Bologna, p. 27,
van der Wel A., Franx M., van Dokkum P. G., Rix H. W., 2004, ApJ, 601, L5
Vanden Berk D. E., Stoughton C., Crotts A. P. S., Tytler D., Kirkman D., 2000, AJ, 119, 2571
Vanzella E. et al., 2008, A&A, 478, 83
Viero M. P. et al., 2013, ApJ, 772, 77
Viero M. P. et al., 2014a, ApJS, 210, 22
Viero M. P. et al., 2014b, ApJS, 210, 22
Wang L. et al., 2014, MNRAS, 444, 2870
Wang L. et al., 2020, A&A, 648, A8
Weedman D. W., Houck J. R., 2009, ApJ, 693, 370
Weiß A. et al., 2013, ApJ, 767, 88
Wirth G. D. et al., 2004, AJ, 127, 3121
Yamada T. et al., 2005, ApJ, 634, 861
Yan L. et al., 2007, ApJ, 658, 778
York D. G. et al., 2000, AJ, 120, 1579
Zou H., Gao J., Zhou X., Kong X., 2019, ApJS, 242, 8

APPENDIX A: OBSIDS

All the *Herschel* observation identification numbers or OBSIDs are available on the GitHub pages for PACS here:

https://github.com/H-E-L-P/dmu_products/blob/master/dmu18/dmu18_HELP-PACS-maps/pacs_obsid.csv

and for SPIRE here:

https://github.com/H-E-L-P/dmu_products/blob/master/dmu19/dmu19_HELP-SPIRE-maps/spire_obsids.csv

All products available through the HELP www pages herschel.sussex.ac.uk.

APPENDIX B: DATA BASE STRUCTURE AND ACCESS

The data presented here is all publicly available including all input data sets that were used to produce it. All the code used for the processing is also publicly available on GitHub here:

https://github.com/H-E-L-P/dmu_products/

The raw data as fits tables and images is available at the *Herschel* database at Marseille here:

http://hedam.lam.fr/HELP/dataproducts/

Those large files can be inconvenient so we also supply a Virtual Observatory server for automated querying of the full data set here:

https://herschel-vos.phys.sussex.ac.uk/

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

All the code used to produce the figures presented in this paper is available here:

https://github.com/H-E-L-P/dmu_products/tree/master/dmu31/dmu31_Examples

In order to access the flat files you will need to navigate the structure of Data Management Units products defined in Table C1. At each stage of the pipeline the working files are saved with the final merged catalogues stored in DMU32. We have produced and continue to work on extensive documentation to aid navigating the data base. Starting at the front page of the GitHub repository should allow the reader to locate detailed descriptions of each section of the data base. Alternatively we also provide per field summaries in addition to imaging of each field in order to inspect individual areas or objects.

APPENDIX C: SPECTROSCOPIC REDSHIFT SOURCES

Spectroscopic redshifts are compiled from numerous sources. Table C2 gives all the relevant references and can be used to find the source for a given object in the *masterlist*.

Table C1. Overview of the Data Management Unit (DMU) data base structure which is used in processing and web data bases.

Data Management Unit (DMU) name	DMU content
DMU0	Pristine catalogues
DMU1	Masterlist data
DMU2	Field definitions
DMU3	Morphologies (under development)
DMU4	Bright Star Mask
DMU5	Known Star Flag
DMU6	Optical photometry validation
DMU7	Optical photometry (under development)
DMU8	Radio data - LOFAR & FIRST/NVSS/TGSS
DMU9	Radio data - JVLA-DEEP & GMRT-DEEP
DMU10	Data Fusion
DMU11	Cross matching MIPS/PACS/SPIRE
DMU12	Cross Matching LOFAR & FIRST/NVSS/TGSS
DMU13	Cross Matching JVLA-DEEP & GMRT-DEEP
DMU14	*GALEX* data
DMU15	X-Ray data (under development)
DMU16	*WISE* Photometry
DMU17	MIPS images
DMU18	PACS images
DMU19	SPIRE images
DMU20	MIPS blind photometry (under development)
DMU21	PACS blind photometry (under development)
DMU22	SPIRE blind photometry
DMU23	Spec-z data
DMU24	Photo-z
DMU25	Prior model
DMU26	XID +
DMU27	Empirical models / templates
DMU28	SED fitting / CIGALE
DMU29	Radiative transfer models (under development)
DMU30	Missing (supplementary) Sources
DMU31	Tools
DMU32	Final merged catalogues

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

Table C2. The individual spectroscopic catalogues used for each field.

Field	Identifier	Source
	1	Shim et al. (2013)
AKARI-NEP	2	2MASS Redshift Survey (2MRS; Huchra et al. 2012)
	4	IRAS Point Source Catalog Redshift Catalog (IRASPSCZ; Saunders et al. 2000)
	8	Updated Zwicky Catalog (UZC; Falco et al. 1999)
	16	NED sources compiled by M. Vaccari (Vaccari 2015)
	1	Sedgwick et al. (2017)
AKARI-SEP	2	6dF Galaxy Survey (Jones et al. 2004)
	4	2MASS Redshift Survey (2MRS; Huchra et al. 2012)
	8	IRAS Point Source Catalog Redshift Catalog (IRASPSCZ; Saunders et al. 2000)
	16	Chincarini et al. (1984)
	32	Dressler et al. (1999)
	64	Loveday et al. (1996)
	128	Maza et al. (1995)
	256	Tully et al. (2008)
	1	AGES: The AGN and Galaxy Evolution Survey (Kochanek et al. 2012)
Bo'otes	2	SDSS DR12 (York et al. 2000)
	4	Houck et al. (2005)
	8	Weedman & Houck (2009)
	1	2dF Galaxy Redshift Survey (Colless et al. 2001)
CDFS-SWIRE	2	6dF Galaxy Survey (Jones et al. 2004)
	4	VIMOS VLT Deep Survey (Le Fèvre et al. 2013)
	8	VLT/FORS2 spectroscopy in the GOODS-South Field (Vanzella et al. 2008)
	16	GOODS - VLT/VIMOS Spectroscopy DR 2.0.1 (Balestra et al. 2010)
	32	IMAGES spectroscopy (Ravikumar et al. 2007)
	64	GMASS Ultradeep Spectroscopy (Kurk et al. 2013)
	128	Szokoly et al. (2004), via MUSYC catalogue (Cardamone et al. 2010)
	256	Croom, Warren & Glazebrook (2001), via MUSYC catalogue (Cardamone et al. 2010)
	512	van der Wel et al. (2004), via2004 MUSYC catalogue (Cardamone et al. 2010)
	1024	Cristiani & D'Odorico (2000), via MUSYC catalogue (Cardamone et al. 2010)
	2048	Strolger et al. (2004), via MUSYC catalogue (Cardamone et al. 2010)
	4096	Lira et al. (in preparation), via MUSYC catalogue (Cardamone et al. 2010)
	8192	Treister et al. (2009), via MUSYC catalogue (Cardamone et al. 2010)
	16384	Kriek et al. (2008), via MUSYC catalogue (Cardamone et al. 2010)
	32768	K20 Survey (Mignoli et al. 2005)
	65536	Silverman et al. (2010)
	131072	Dickinson et al. (2004)
	262144	PRIMUS (Coil et al. 2011)
	524288	2MASS Redshift Survey (2MRS; Huchra et al. 2012)
	1048576	IRAS Point Source Catalogue Redshift Catalogue (IRASPSCZ; Saunders et al. 2000)
	2097152	The Arizona CDFS Environment Survey (ACES; Cooper et al. 2012)
	4194304	Lacy et al. (2013)
	8388608	OzDES DR1 (Childress et al. 2017)
	16777216	MUSE-Wide Survey (Herenz et al. 2017)
	33554432	VANDELS DR2 (McLure et al. 2018)
	1	SDSS DR15 (York et al. 2000)
COSMOS	2	2MASS Redshift Survey (2MRS; Huchra et al. 2012)
	4	PRIMUS (Coil et al. 2011)
	8	COSMOS spec-z catalogue[a] (public redshifts), source 3DHST (Momcheva et al. 2016)
	16	COSMOS spec-z catalogue[a] (public redshifts), source Onodera et al. (2015)
	32	COSMOS spec-z catalogue[a] (public redshifts), source FAST (N.Wright, F Civano)
	64	COSMOS spec-z catalogue[a] (public redshifts), source AZTEC (N.Wright, F Civano)
	128	COSMOS spec-z catalogue[a] (public redshifts), source (Roseboom et al. 2012)
	256	COSMOS spec-z catalogue[a] (public redshifts), source (Comparat et al. 2015)
	512	COSMOS spec-z catalogue[a] (public redshifts), source 2MRS (Huchra et al. 2012)
	1024	COSMOS spec-z catalogue[a] (public redshifts), source GEMINI-S (M. Balogh)
	2048	COSMOS spec-z catalogue[a] (public redshifts), source *HST* GRISM (K. Kartaltepe, M. Brusa)
	1	SDSS DR15 (York et al. 2000)
EGS	2	DEEP 2 & 3 (Cooper et al. 2011; Newman et al. 2013)
	4	3DHST (Momcheva et al. 2016)
	8	Steidel et al. (2003)
	16	AEGIS-X (Nandra et al. 2015)

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

Table C2 – *continued*

Field	Identifier	Source
	32	Huang et al. (2009)
	64	C3R2 DR1 & DR2 (Masters et al. 2019)
	1	Berta et al. (2007)
ELAIS-N1	2	SDSS DR13 (York et al. 2000)
	4	Trichas et al. (2010)
	8	Swinbank et al. (2007)
	16	Rowan-Robinson et al. (2013) (WIYN, Keck and Gemini sources)
	32	Rowan-Robinson et al. (2013) sources taken from NED)
	64	Updated Zwicky Catalog (UZC Falco et al. 1999)
	128	Lacy et al. (2013)
	1	SDSS DR13 (York et al. 2000)
ELAIS-N2	2	Berta et al. (2007)
	4	Swinbank et al. (2005)
	8	Updated Zwicky Catalogue (UZC; Falco et al. 1999)
	16	IRAS Point Source Catalog Redshift Catalog (IRASPSCZ; Saunders et al. 2000)
	32	Rowan-Robinson et al. (2013) (WIYN, Keck and Gemini sources)
	64	Rowan-Robinson et al. (2013) sources taken from NED)
	128	2MASS Redshift Survey (2MRS; Huchra et al. 2012)
	256	Lacy et al. (2013)
	1	Australian Telescope Large Area Survey (Mao et al. 2012)
ELAIS-S1	2	Sacchi et al. (2009)
	4	Feruglio et al. (2008)
	8	PRIMUS (Coil et al. 2011)
	16	2dF Galaxy Redshift Survey (Colless et al. 2001)
	32	6dF Galaxy Survey (Jones et al. 2004)
	64	IRAS Point Source Catalog Redshift Catalogue (IRASPSCZ; Saunders et al. 2000)
	128	Lacy et al. (2013)
	256	OzDES DR1 (Childress et al. 2017)
	1	GAMA I - DR3 Baldry et al. (2018)
GAMA-09,12,15	2	SDSS DR13 (York et al. 2000)
	4	WiggleZ DR1 (Drinkwater et al. 2010)
	8	6dF Galaxy Survey (Jones et al. 2004)
	16	2dF Galaxy Redshift Survey (Colless et al. 2001)
	32	2SLAQ-QSO (Croom et al. 2009)
	64	2SLAQ-LRG (Croom et al. 2009)
	128	2MASS Redshift Survey (2MRS; Huchra et al. 2012)
	256	2dF QSO Redshift Survey (Croom et al. 2004)
	512	Millennium Galaxy Catalogue (MGC; Liske et al. 2003)
	1024	Updated Zwicky Catalogue (UZC; Falco et al. 1999)
	2048	NED sources
	4096	GAMA I (Liverpool Telescope) Baldry et al. (2018)
	8192	IRAS Point Source Catalog Redshift Catalog (IRASPSCZ; Saunders et al. 2000) *(Two proprietary sources are stored in HELP database for new releases of GAMA and WiggleZ for when they are made public)*
	1	SDSS DR14 (York et al. 2000)
HDF-N	2	DEEP 3 (Cooper et al. 2011)
	4	3DHST (Momcheva et al. 2016)
	8	Steidel et al. (2003)
	16	Liu et al. (1999)
	32	Team Keck Treasury Redshift Survey (Wirth et al. 2004)
	64	Reddy et al. (2006)
	128	Caltech Faint Galaxy Redshift Survey (Cohen et al. 2000, and references there in)
	256	Chapman et al. (2004)
	512	Chapman et al. (2005)
	1024	Swinbank et al. (2004)
	2048	Dawson et al. (2001)
	4096	Pope et al. (2008)
	8192	Barger, Cowie & Wang (2008, their sources 1, 3, 7, 12, and 14)
	16384	Vanden Berk et al. (2000)
	32768	Cowie et al. (2004)
	1	2dF Galaxy Redshift Survey (Colless et al. 2001)
Herschel-Stripe-82	2	2SLAQ-QSO (Croom et al. 2009)
	4	2SLAQ-LRG (Croom et al. 2009)

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

Table C2 – *continued*

Field	Identifier	Source
	8	6dF Galaxy Survey (Jones et al. 2004)
	16	PRIMUS (Coil et al. 2011)
	32	SDSS DR12 (York et al. 2000)
	64	DEEP 2 (Newman et al. 2013)
	128	WiggleZ DR1 (Drinkwater et al. 2010)
	256	2MASS Redshift Survey (2MRS; Huchra et al. 2012)
	1	Steffen et al. (2004)
Lockman-SWIRE	2	Berta et al. (2007)
	4	SDSS DR13 (York et al. 2000)
	8	Rowan-Robinson et al. (2013) sources taken from NED
	16	Rowan-Robinson et al. (2013) (WIYN, Keck and Gemini sources)
	32	2MASS Redshift Survey (2MRS; Huchra et al. 2012)
	64	IRAS Point Source Catalogue Redshift Catalog (IRASPSCZ; Saunders et al. 2000)
	128	Updated Zwicky Catalogue (UZC; Falco et al. 1999)
	256	Lacy et al. (2013)
NGP	1	SDSS DR12 (York et al. 2000)
	2	2MASS Redshift Survey (2MRS; Huchra et al. 2012)
	1	SDSS DR13 (York et al. 2000)
SA13	2	Chapman et al. (2005)
	4	Smail et al. (2004)
	8	Cowie et al. (1996)
	1	2dF Galaxy Redshift Survey (Colless et al. 2001)
SGP	2	6dF Galaxy Survey (Jones et al. 2004)
	4	2MASS Redshift Survey (2MRS; Huchra et al. 2012)
	8	Southern Sky Redshift Survey (SRSS; da Costa et al. 1998)
SPIRE-NEP	1	2MASS Redshift Survey (2MRS; Huchra et al. 2012)
	1	2dF Galaxy Redshift Survey (Colless et al. 2001)
SSDF	2	6dF Galaxy Survey (Jones et al. 2004)
	4	IRAS Point Source Catalog Redshift Catalog (IRASPSCZ; Saunders et al. 2000)
	8	2MASS Redshift Survey (2MRS; Huchra et al. 2012)
	1	SDSS (York et al. 2000)
xFLS	2	Updated Zwicky Catalogue (UZC; Falco et al. 1999)
	4	Papovich et al. (2006)
	8	Marleau et al. (2007)
	16	Lacy et al. (2007)
	32	Lacy et al. (2013)
	64	Yan et al. (2007)
	128	NED sources compiled by M. Vaccari (Vaccari 2015)
	1	SDSS DR14 (York et al. 2000)
XMM–13hr	2	Page et al. (2006)
	4	Jeltema et al. (2007)
	8	MCXC (Piffaretti et al. 2011)
	1	VIMOS VLT Deep Survey (Le Fèvre et al. 2013)
XMM–LSS	2	Garcet et al. (2007)
	4	Lacy et al. (2007)
	8	Stalin et al. (2010)
	16	Lidman et al. (2013)
	32	Yamada et al. (2005)
	64	Ouchi et al. (2008)
	128	Ono et al. (2010)
	256	Sargsyan & Weedman (2009)
	512	UDSz (Bradshaw et al. 2013; McLure et al. 2013)
	1024	UDS catalogue[a] source marked 'CJS-VIMOS'
	2048	UDS catalogue[a] source marked 'CJS-ISIS'
	4096	UDS catalogue[a] source marked 'CVB-DEIMOS'
	8192	UDS catalogue[a] source marked 'JEG-LDSS2'
	16384	UDS catalogue[a] source marked 'Doi-FOCASS'
	32768	UDS catalogue[a] source marked 'SJC-AAOmega'
	65536	UDS catalogue[a] source marked 'IRS-AAomega'
	131072	UDS catalogue[a] source marked 'Aki-FOCAS'

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022

Table C2 – *continued*

Field	Identifier	Source
	262144	UDS catalogue[a] source marked 'Aki-2df'
	524288	SDSS DR15 (York et al. 2000)
	1048576	6dF Galaxy Survey (Jones et al. 2004)
	2097152	VIPERS PDR2 (Scodeggio et al. 2018)
	4194304	PRIMUS (Coil et al. 2011)
	8388608	Magellan/IMACS catalogue[b]
	16777216	Lacy et al. (2013)
	33554432	OzDES DR1 (Childress et al. 2017)
	67108864	C3R2 DR1 & DR2 (Masters et al. 2019)
	134217728	VANDELS DR2 (McLure et al. 2018)

Notes. [a]Available from https://www.nottingham.ac.uk/astronomy/UDS/data/data.html.
[b]Available from http://localgroup.ps.uci.edu/cooper/IMACS/zcatalog.html.

[1]*Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK*
[2]*Astronomy Centre, Department of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK*
[3]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[4]*Sterrewacht Leiden, Universiteit Leiden, NL-2300 RA Leiden, the Netherlands*
[5]*Royal Observatory, Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK*
[6]*National Centre for Nuclear Research, ul. Pasteura 7, PL-02-093 Warszawa, Poland*
[7]*Aix Marseille Univ, CNRS, CNES, LAM, F-13004 Marseille, France*
[8]*School of Physics and Astronomy, Cardiff University, Queens Buildings, The Parade, Cardiff CF24 3AA, UK*
[9]*Laboratoire AIM-Paris-Saclay, CEA/DSM/Irfu - CNRS - Université Paris, CE-Saclay, F-91190 Saint-Aubin, France*
[10]*Division of Particle and Astrophysical Science, Graduate School of Science, Nagoya University, Aichi 464-8602, Japan*
[11]*National Astronomical Observatory of Japan, 2-21-1, Osawa, Mitaka, Tokyo 181-8588, Japan*
[12]*Institut Universitaire de France (IUF), F-75231 Paris, France*
[13]*School of Sciences, European University Cyprus, Engomi, 1516 Nicosia, Cyprus*
[14]*Astrophysics, Department of Physics, University of Oxford, Keble Road, Oxford OX1 3RH, UK*
[15]*Department of Physics and Astronomy, University of the Western Cape, 7535 Bellville, Cape Town, South Africa*
[16]*Department of Astronomy, University of Cape Town, 7701 Rondebosch, Cape Town, South Africa*
[17]*INAF – Istituto di Radioastronomia, via Gobetti 101, I-40129 Bologna, Italy*
[18]*Inter-University Institute for Data Intensive Astronomy, University of the Western Cape, 7535 Bellville, Cape Town, South Africa*
[19]*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK*
[20]*Department of Physics and Astronomy, Oberlin College, Oberlin, OH 44074, US*
[21]*SRON Netherlands Institute for Space Research, Landleven 12, NL-9747 AD Groningen, the Netherlands*
[22]*Kapteyn Astronomical Institute, University of Groningen, Postbus 800, NL-9700 AV Groningen, the Netherlands*

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

Downloaded from https://academic.oup.com/mnras/article/507/1/129/6291723 by guest on 07 January 2022