

# The *Herschel*/PACS Point Source Catalogue Explanatory Supplement

G. Marton<sup>1</sup>, L. Calzoletti<sup>2</sup>, A. M. Perez Garcia<sup>2</sup>, C. Kiss<sup>1</sup>, R. Paladini<sup>3</sup>, B. Altieri<sup>2</sup>, M. Sánchez Portal<sup>2</sup>, M. Kidger<sup>2</sup>, and  
the *Herschel* Point Source Catalogue Working Group

<sup>1</sup>Konkoly Observatory, Research Centre for Astronomy and Earth Sciences,  
Hungarian Academy of Sciences, H-1121 Budapest

<sup>2</sup>European Space Astronomy Centre (ESAC)/ESA, PO Box 78, 28690 Villanueva  
de la Cañada, Madrid, Spain

<sup>3</sup>Infrared Processing Analysis Center, California Institute of Technology, 770 South  
Wilson Ave., Pasadena, CA 91125, USA

May 17, 2017



## Abstract

The *Herschel* Space Observatory was the fourth cornerstone mission in the European Space Agency (ESA) science programme. It had excellent broad band imaging capabilities in the far-infrared (FIR) and sub-millimetre part of the electromagnetic spectrum. Although the spacecraft finished observing in 2013, it left a large legacy dataset that is far from having been fully explored and still has a great potential for new scientific discoveries. The PACS and SPIRE photometric cameras observed about 8% of the sky in six different wavebands. This document describes the *Herschel*/PACS Point Source Catalogue (HPPSC), a FIR catalogue based on the broad-band photometric observations of the PACS instrument with filters centred at 70, 100 and 160  $\mu\text{m}$ .

We analysed all combined, Level 2.5/Level 3 *Herschel*/PACS photometric observations including 682 Parallel Mode, 12932 nominal mode and 1644 SSO maps. The PACS photometer maps that we produced were generated by applying the JScanam task of the *Herschel* Interactive Processing Environment (HIPE) v13.0.0. Sources were identified with the HIPE implementation of SUSSExtractor, and the flux densities obtained by aperture photometry. We found a total of 108 319 point sources that are considered to be reliable in the 70  $\mu\text{m}$  maps, 131 322 at 100  $\mu\text{m}$  and 251 392 point sources in the 160  $\mu\text{m}$  maps. In addition, our quality control algorithm identified 546 587 candidate sources that were found to be extended and 7 185 160 features which did not pass the signal-to-noise and other criteria to be considered reliable sources. These sources were included in the Extended Source List and Rejected Source List of the HPPSC, respectively. The calculated completeness and photometric accuracy values are based on simulations, where artificial sources were injected into the observational timeline with well controlled flux density values. The actual completeness is a complex function of the source flux, photometric band and the background complexity.

---

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

# Contents

<b>1</b>	<b>General Information</b>	<b>6</b>
1.1	Purpose and Structure of the Document . . . . .	6
1.2	Purpose of the Catalogue . . . . .	6
1.3	Cautionary Notes . . . . .	6
1.3.1	Catalogue and additional material . . . . .	6
1.3.2	Photometry . . . . .	7
1.3.3	Coverage & Completeness . . . . .	7
1.3.4	Homogeneity . . . . .	7
1.3.5	Cross-matching . . . . .	7
1.3.6	Reliability . . . . .	7
1.3.7	Tracked maps of Solar system targets . . . . .	8
1.4	How to access the HPPSC data products . . . . .	8
1.5	Referencing the HPPSC . . . . .	8
<b>2</b>	<b>Introduction</b>	<b>9</b>
2.1	Mission overview . . . . .	9
2.2	The PACS Photometer . . . . .	9
2.3	The PACS Standard Products . . . . .	11
<b>3</b>	<b>Catalogue generation – data &amp; methods</b>	<b>13</b>
3.1	Map repository . . . . .	13
3.2	Structure noise maps . . . . .	14
3.3	Pipeline description . . . . .	16
3.3.1	Source detection . . . . .	17
3.3.2	Flux extraction . . . . .	17
3.3.3	2D Gaussian fit . . . . .	18
3.4	Quality assessment . . . . .	18
3.4.1	Structure noise based S/N . . . . .	19
3.4.2	Background RMS based S/N . . . . .	21
3.4.3	Fitted positions . . . . .	22
3.4.4	Fluxes . . . . .	23
3.4.5	FWHM . . . . .	23
3.5	Object consolidation . . . . .	24
3.5.1	Cleaning source tables . . . . .	25
3.5.2	Multiple detections . . . . .	27
<b>4</b>	<b>Point Source Catalogue</b>	<b>28</b>
4.1	Catalogue properties . . . . .	28
4.2	HPPSC columns . . . . .	31
4.2.1	Name ( <i>name</i> ) . . . . .	31
4.2.2	Band ( <i>band</i> ) . . . . .	31
4.2.3	R.A./Dec ( <i>ra/dec</i> ) . . . . .	31
4.2.4	RA <sub>err</sub> /Dec <sub>err</sub> ( <i>raerr/decerr</i> ) . . . . .	31
4.2.5	Flux ( <i>flux</i> ) . . . . .	32
4.2.6	S/N <sub>S</sub> ( <i>snr</i> ) . . . . .	32
4.2.7	N <sub>S</sub> based noise ( <i>snrnoise</i> ) . . . . .	32
4.2.8	S/N <sub>R</sub> ( <i>stn</i> ) . . . . .	32
4.2.9	RMS ( <i>rms</i> ) . . . . .	32
4.2.10	N <sub>S</sub> ( <i>strn</i> ) . . . . .	32
4.2.11	Flux ratio ( <i>fratio</i> ) . . . . .	32
4.2.12	FWHM <sub>X</sub> /FWHM <sub>Y</sub> <i>fwhmxfit/fwhmyfit</i> . . . . .	32
4.2.13	ELONGATION FLAG ( <i>flag_elong</i> ) . . . . .	32
4.2.14	EDGE FLAG ( <i>flag_edge</i> ) . . . . .	33
4.2.15	BLEND FLAG ( <i>flag_blend</i> ) . . . . .	33
4.2.16	WARM ATTITUDE ( <i>warmat</i> ) . . . . .	33
4.2.17	OBSERVATION ID ( <i>obsid</i> ) . . . . .	33
4.2.18	SOLAR SYSTEM MAP FLAG ( <i>ssomapflag</i> ) . . . . .	33

4.3	Additional material . . . . .	34
4.3.1	Extended Source List . . . . .	34
4.3.2	Rejected Source List . . . . .	35
4.3.3	Observation Table . . . . .	35
<b>5</b>	<b>Validation</b> . . . . .	<b>36</b>
5.1	Internal validation – Simulations . . . . .	36
5.1.1	Completeness . . . . .	37
5.1.2	Photometric accuracy . . . . .	38
5.1.3	Positional accuracy . . . . .	39
5.2	External Validation . . . . .	40
5.2.1	Comparison with point source flux calibration . . . . .	41
5.2.2	Faint star comparison . . . . .	43
5.2.3	PEP-COSMOS field . . . . .	45
5.2.4	PEP-GOODS field . . . . .	47
5.2.5	Goods- <i>Herschel</i> survey . . . . .	48
5.2.6	The Galactic Cold Cores OTKP . . . . .	50
5.2.7	Hi-GAL DR1 . . . . .	52
5.3	Conclusion . . . . .	54

## List of Acronyms

AOR	Astronomical Observation Request
ESA	European Space Agency
FIR	far-infrared
FWHM	Full Width at Half Maximum
GH	GOODS- <i>Herschel</i>
Hi-GAL	<i>Herschel</i> infrared Galactic Plane Survey
HIPE	<i>Herschel</i> Interactive Processing Environment
HPESL	<i>Herschel</i> /PACS Extended Source List
HPF	High-Pass Filtering
HPPSC	<i>Herschel</i> /PACS Point Source Catalogue
HPRSL	<i>Herschel</i> /PACS Rejected Source List
HSA	<i>Herschel</i> Science Archive
$N_S$	Structure noise
NASA	National Aeronautics and Space Administration
PACS	Photodetector Array Camera and Spectrometer
PEP	PACS Evolutionary Probe
PSF	Point Spread Function
ROI	Region Of Interest
$S/N_S$	Structure noise based Signal-to-Noise Ratio
$S/N_R$	Background RMS based Signal-to-Noise Ratio
mm	sub-millimetre
SPG	Standard Product Generation
SPIRE	Spectral and Photometric Imaging Receiver
SSO	Solar System Object
STRN	Structure noise
TOD	Time Ordered Data

## Document Change History

Date	Changed
12-April-2017	First version
24-April-2017	Language editing
05-May-2017	First public version

# 1 General Information

## 1.1 Purpose and Structure of the Document

The *Herschel*/PACS Point Source Catalogue (HPPSC) Explanatory Supplement is part of the first public version of the Catalogue. This Explanatory Supplement is intended to be a complete and self-contained description of the products generated for and the methods of generation used in the HPPSC. The first part is an introduction and overview of the mission, the PACS instrument and practical issues related to the work. The second part gives a detailed description of the procedure that was used in order to generate the HPPSC, including the products, the methods of detection and the extraction pipeline, as well as the quality assessment. The third part gives details of the resulting products, including the Catalogue itself and additional lists. The fourth part details the additional material, including the *Herschel*/PACS Extended Source List (HPESL), the *Herschel*/PACS Rejected Source List (HPRSL), and the observation table. The observation table contains information on the map that was used for the extraction of a given Catalogue object. The final chapter describes the validation procedures, including internal validation by using simulations and external validation by comparing our results with already published catalogues.

This document includes information that are essential in order to understand our products. However, more detail might be necessary for some users. Additional background information is found in the [PACS Observer's Manual](#). A deeper understanding of the map products can be obtained from the [PACS Map-Making Report](#). A detailed analysis of the PACS photometer Point Spread function is available [here](#). The Point Source Flux Calibration is described in [Balog et al., 2014](#) and confirmed with asteroids in [Müller et al., 2014](#). Documents of the colour corrections can be found in [Müller et al., 2011](#) and [here](#).

## 1.2 Purpose of the Catalogue

The *Herschel* Space Observatory ceased to take science observations on 29 April, 2013. It left a large legacy dataset that is far from having been fully explored and still has potential for new scientific discoveries. Some source catalogues have already been produced by individual observing programmes, which are available via the [Herschel User Provided Data Products](#). However, there are still many observations that have not yet been analysed for their full science content.

To maximise the science return of the PACS data sets, we built a homogeneous *Herschel*/PACS Point Source Catalogue (HPPSC) from all the scan map observations (see Section 3.1). Our source extraction enables a systematic and unbiased comparison of sensitivity across different *Herschel* fields that single programs will generally not be able to provide. The extracted point sources contain individual YSOs and unresolved YSO clusters in our Galaxy, as well as dusty extragalactic objects of the local and distant Universe. This rich dataset provides a pool of diverse celestial targets for scientists to study in more detail the early phases of star formation, galaxy properties, both locally and at large distances and galaxy evolution through time. In addition to enabling the possibility of statistical analysis of stellar and galaxy clusters, it also provides an excellent target list for follow-up observations with current state-of-art and future facilities such as ALMA and JWST.

While we did our best to deliver a catalogue that can be used for scientific analysis as it is, caveats are still present. We encourage all users of the HPPSC to read the Cautionary Notes below, and to do their science with these notes in mind.

## 1.3 Cautionary Notes

### 1.3.1 Catalogue and additional material

The primary goal of our work was to create a reliable Point Source Catalogue that provides information that can be easily and safely used for scientific study. Therefore, we provide detailed information of this work, including data description, calibration and quality assessment. However, *Herschel* did observe many other objects that could be potentially interesting for the users. Our additional source lists, the *Herschel*/PACS Extended Source List (HPESL) and the Rejected Source List (HPRSL) contain information that were extracted with the pipeline designed for the point sources, but that are too extended to be considered as point sources, and/or fail to pass one or more of our quality criteria (Figure 14). **These additional lists can be used at the user's**

own risk. We encourage strongly users to check the entries found in these lists and revise them using their own methods.

### 1.3.2 Photometry

Our photometric accuracy is within 1% in relative to the photometry of the PACS calibrator stars: see subsection 5.2.1. However, these values are based on measurements that were specifically designed for flux calibration, with multiple visits,  $20''\text{s}^{-1}$  scan speed and isolated sources on flat background. Their Spectral Energy Distribution is well known, therefore colour corrections were also possible. These conditions apply only for a very small number of sources observed by *Herschel*. Therefore, colour corrections were not made for the published values: it is up to the user to calculate them. Also, it is important to note, that the values listed in the Catalogue are **in-band** flux density values.

### 1.3.3 Coverage & Completeness

The *Herschel* Space Observatory, as its name indicates, was operated as an observatory, meaning that instead of an all-sky coverage it was pointed towards carefully selected targets/fields. The summed area of the observations, including overlaps, is  $\sim 7\%$  of the sky across the three PACS bands. However, not all observations could be used for Catalogue generation, as we required that the observations were processed at least up to Level 2.5 (see Section 3.1): 126 Parallel Mode observations and 4452 Scan map observations do not satisfy this requirement. Generally, observations reduced up to a lower processing level (Level 2) are spurious observations in which the sky fields are recovered by higher level products (Level 2.5 or Level 3). Nevertheless there are some large surveys the scanning strategies prevented the generation of standard Level 2.5. The most important of these are the North/South Galactic Poles fields (about 490 sqdg), belonging to [The Herschel Thousand Degree Survey \(H-ATLAS\)](#) Key Program and the [HERschel Inventory of the Agents of Galaxy Evolution \(HERITAGE\)](#) Key Program: this last consists of the LMC ( $8^\circ \times 8.5^\circ$ ), the SMC ( $5^\circ \times 5^\circ$ ), and a  $4^\circ \times 4^\circ$  region of the SMC Tail at two different epochs. So, the final estimated sky coverage of our Catalogue is  $\sim 4.5\%$ .

We made a considerable effort to understand how the completeness of our detections depends on the flux level of the sources and on the complexity of the environment. More details are found in Section 5.1.1. It is important to note that, especially in star forming regions, sources tend to be located in groups or clusters. In many cases the cluster members are so close together that they bias the S/N estimation, causing false rejection of sources despite our best effort to preserve them.

### 1.3.4 Homogeneity

*Herschel* carried out a multitude of observing programmes with different scientific goals. Due to the observing strategies and parameters of these programmes, the sky coverage is very arbitrary. One of the main goals with the *Herschel* Point Source Catalogues was to carry out a homogeneous source extraction, so the results can be compared, taking into account scan speed, sampling rate, repetition factor. We used the same pipeline for all maps, which was of course tailored for the different observing bands. Similarly, the input PSFs for the source detection algorithm (SUSSEXtractor) were chosen as a function of the observing mode and scan speed.

### 1.3.5 Cross-matching

The HPPSC is a blind catalogue, i.e., extraction of sources is not based on prior detections at different wavelengths. Extraction of sources in each band was carried out separately. Band merging and cross-identifications are out of the scope of this work.

### 1.3.6 Reliability

The unprecedented resolution of *Herschel* in the FIR and at sub-mm wavelengths allowed us to map the dusty ISM with a detail never before seen. The separation of the background fluctuation and the semi-extended, or point sources is mathematically challenging and triggered the development of many algorithms tailored for star forming regions observed by *Herschel*. These algorithms are very sophisticated and complex tools that are not meant for general use. We needed to select an algorithm that is both fast and easy to use for all kinds of fields, even if that comes at the cost of

reliability, or completeness. The [SUSSEXtractor](#) is mainly sensitive to point sources, but has also identified brightness fluctuations of the background as sources. We refer to the [IRAS Explanatory Supplement](#) and state that the sky at 60-200  $\mu\text{m}$  is **dominated by filaments termed "infrared cirrus" which, although concentrated in the Galactic plane, can be found almost all the way up to the Galactic poles. The primary, deleterious effects of the cirrus are that it can generate well-confirmed point and small extended sources that are actually pieces of degree- or arcminute-sized structures rather than isolated, discrete objects and that it can corrupt measurements of true point sources.**

Despite our conservative approach to filtering out spurious sources, based on various quality criteria, such objects remained occasionally in our catalogue. If the user is not confident that an astronomical object is present at the desired sky position, eyeballing of the map is encouraged. All the data (signal-to-noise, FWHM, flux ratio, etc., see Section 3.5) that we used in the process of source table cleaning and quality assessment, are included in the Point Source Catalogue files and provided for the users. Scientists are encouraged to create their own criteria, or improve ours if necessary, so the resulting source list is more suited for their scientific needs.

### 1.3.7 Tracked maps of Solar system targets

As it is detailed below, maps of Solar system objects have typically been observed in tracked mode, i.e. taking into account the apparent movement of the target. These maps were reprocessed in the 'rest' sky frame for the HPPSC. In the present version of the catalogue no identification of the moving Solar system targets is performed, but the users are encouraged to check against the possible solar system target contamination using publicly available solar system object databases (e.g. [Herschel Catalogue of Solar System Observations of Romero, Kidger & Rengel, in prep.](#)). The affected OBSIDs are marked by an 'SSOMAP' flag in the Observation Table.

## 1.4 How to access the HPPSC data products

The HPPSC data products are available in three different ways:

- Separate, compressed csv tables via the [ESA COSMOS website](#)
- [ESA Sky](#)
- [Catalog Search Tool for Herschel data](#) available from NASA's Infrared Science Archive (IRSA)

## 1.5 Referencing the HPPSC

The relevant rules of the [Herschel Publishing Rules and Guidelines](#) apply for scientific results based on the HPPSC. It has been agreed that in all papers using *Herschel* data a mandatory footnote on the first page should be included. The footnote serves multiple purposes, including providing credit and simplifying publication tracking. The footnote is: **"Herschel is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA."** The standard references to *Herschel* is Pilbratt, G.L., Riedinger, J.R., Passvogel, T. et al. 2010, *A&A*, 518, L1. The standard reference for the PACS instrument is Poglitsch, A., Waelkens, C., Geis, N. et al. 2010, *A&A*, 518, L2.

For the time being the HPPSC can be cited by means of the the arXiv publication of the Explanatory Supplement.



## 2 Introduction

### 2.1 Mission overview

The *Herschel* Space Observatory (Pilbratt et al. 2010) was the fourth cornerstone mission in the European Space Agency (ESA) science programme. It had a primary mirror of 3.5 m in diameter that allowed an unprecedented spatial resolution and sensitivity at far-infrared (FIR) and sub-millimetre (smm) wavelengths.

*Herschel* operated successfully from June 2009 to 29 April 2013 when it ran out of the liquid helium coolant required to maintain the operational temperatures for the instruments' detectors. The three instruments on-board covered the FIR and smm spectral range from 55 to 671  $\mu\text{m}$ . The Photodetector Array Camera and Spectrometer (PACS, Poglitsch et al. 2010) and the Spectral and Photometric Imaging REceiver (SPIRE, Griffin et al. 2010) were able to make both spectroscopic and photometric observations, while the Heterodyne Instrument for the Far Infrared (HIFI, de Graauw et al. 2010) was a purely spectroscopic instrument. Over 35 000 observations were made during the more than 25 000 hour long mission. A large legacy dataset was obtained that is far from having been fully analysed and which still has a great potential for new scientific discoveries.

The observing time was allocated to both Guaranteed and Open Time Programmes and some source catalogues have already been produced by these observing programs, however there are many observations that remained unexplored. To maximise the scientific return of the *Herschel* photometric observations, the *Herschel*/SPIRE Point Source Catalogue (HSPSC) and the *Herschel*/PACS Point Source Catalogue (HPPSC) were generated, the latter described in the present Explanatory Supplement.

Our homogeneous source extraction enables a systematic and unbiased comparison of sensitivity across the different *Herschel* fields that single programs will generally not be able to provide. The extracted point sources include mainly individual YSOs and unresolved YSO clusters of our Galaxy, as well as the dusty extragalactic objects of the local and distant Universe. Such a huge dataset can help scientists better to understand the early phases of star and galaxy formation in addition to the possibility of carrying out statistical analysis of stellar and galaxy clusters to unravel astrophysical evolution laws through time. It will also provide an excellent target list for future proposals.

### 2.2 The PACS Photometer

The PACS photometer was a dual-band instrument composed of two filled silicon bolometer arrays: the Blue camera consisted of  $32 \times 64$  pixels and the Red camera of  $16 \times 32$  pixels. The Blue camera could be used with two filters centered at 70  $\mu\text{m}$  and 100  $\mu\text{m}$  (blue and green band), while the Red camera acquired observations at 160  $\mu\text{m}$  nominal wavelength (red band). The filter transmission curves are shown in Figure 1. The two cameras allowed two-band simultaneous observations: one of the two bands from the Blue camera plus the Red camera, resulting in pairs of observations taken at 70/160 or 100/160  $\mu\text{m}$ . Both cameras had a field of view of  $\sim 1.75' \times 3.5'$ , with close to Nyquist beam sampling in each band.

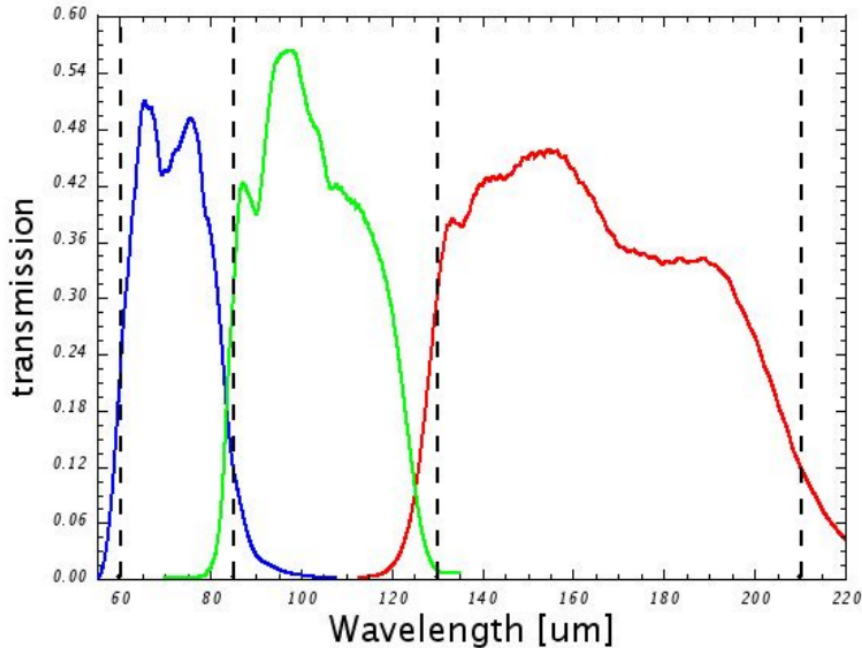


Figure 1: Filter transmissions of the PACS filters. The graph represents the overall transmission of the filters combined with the dichroic and the detector relative response in each of the three bands of the photometer. The dashed vertical lines mark the original intended (design values) of the band edges.

HPPSC data were taken from 430 unique observing programs, including major Guaranteed Time Key Programs and small, individual Open Time Programs. The key parameters that characterise the PACS observations are the following:

- Observing mode: the observing mode could either be the nominal Scan Map or Parallel Mode. In the latter the observations were acquired together with the SPIRE instrument, with a simultaneous five-band photometry of the same field of view. In both modes, the bolometer read-out frequency was 40 Hz, but due to data-rate limitations, an on-board averaging of 4 frames was performed, except for the blue and green bands in Parallel Mode where the averaging was increased to 8 frames. We emphasize that the HPPSC does not contain any extraction from chop-nod observations, as in the chop-nod mode the same source appears multiple times on the processed images, as described in [Nielbock et al. \(2013\)](#).
- Scan speed: the standard satellite scanning speed was  $20''\text{s}^{-1}$ . The highest speed of  $60''\text{s}^{-1}$  was dedicated to large Galactic surveys and Parallel Mode maps. We also included 47 calibration observations where the scan speed was  $10''\text{s}^{-1}$ .
- Coverage/ depth of the observations: depending on the science goals of the observations different scanning strategies were applied. The shallowest maps are those acquired in Parallel Mode observations, due to the high scan speed and the application of a single repetition. The deepest maps are the ones where low or medium scan speed was selected and the number of repetitions was high. These are typically dedicated observations of point sources. As mentioned above, in many cases Level 2.5 observations were combined into Level 3 images.

The shape of the PACS Point Spread Function (PSF) is dominated by a tri-lobate pattern in all three bands. A detailed description can be found in [Lutz \(2015\)](#). The PSF can be approximated by 2-dimensional Gaussian fits, as given in Table 1 below. For low and medium scan speeds the typical FWHMs are  $5.5''$ ,  $6.7''$  and  $11''$  in the 70, 100 and 160  $\mu\text{m}$  bands, respectively. These beams also define the achievable spatial resolution and set the limit on the separability of nearby sources.

Mode	AMA	blue 70 $\mu$ m		green 100 $\mu$ m		red 160 $\mu$ m	
		FWHM	PA	FWHM	PA	FWHM	PA
	°	"	°	"	°	"	°
Prime 10 "s <sup>-1</sup>	+63	5.20× 5.56		6.54× 6.78		10.38×11.97	6.1
Prime 20 "s <sup>-1</sup>	+63	5.41× 5.72		6.66× 6.89		10.55×12.08	9.1
Prime 60 "s <sup>-1</sup>	+63	5.70× 9.05	61.7	6.84× 9.81	61.8	11.39×13.37	41.2
Parallel 20 "s <sup>-1</sup>	+42	5.44× 6.51	30.8	6.62× 7.44	31.1	10.29×12.20	8.5
Parallel 20 "s <sup>-1</sup>	-42	5.31× 6.68	-26.5	6.53× 7.56	-27.0	10.37×12.27	-3.4
Parallel 60 "s <sup>-1</sup>	+42	5.85×12.58	43.7	6.99×13.15	43.9	10.90×14.09	27.7
Parallel 60 "s <sup>-1</sup>	-42	5.69×12.74	-36.9	6.87×13.41	-37.1	11.01×14.53	-23.7
Parallel 20 "s <sup>-1</sup>	+42,-42	5.74×6.26	0.4	6.98× 7.42	-2.9	10.46×12.27	3.1
Parallel 60 "s <sup>-1</sup>	+42,-42	8.80×9.60	-4.4	9.73×10.69	-3.8	11.51×13.65	5.3

Table 1: FWHM of the PACS PSF for several important cases. 2-dimensional Gaussian fits were used to derive the FWHM for the small and large axis. For noticeably non-round PSF cores, the position angle east of the spacecraft Z direction is noted. The array to map angle (AMA) of the scan is also specified. The maps used to derive the FWHM have been created by photProject with map pixel size 1" and pixfrac=1. Entries above the line refer to single direction scans, showing the in-scan elongation for fast scan and for Parallel Mode. Entries below the line refer to coadded Parallel Mode crossed scans, where a cross-like PSF emerges from co-adding the two elongated PSFs.

### 2.3 The PACS Standard Products

PACS observations were processed automatically by SPG (Standard Product Generator) and placed in the *Herschel* Science Archive (HSA). Several products belonging to different processing Levels are generated by SPG chain (see Figure 2); they are available in the HSA for users to download (see the PACS Observing Manual for a detailed description of the data product levels). Starting from the raw telemetry, Level 0 frames are generated and processed up to Level 1. In these steps, instrumental effects (1/f noise, common mode drift, cross-talk) are removed and frames are calibrated in Jy/pixel. Level 1 frames are the starting point for the mapmakers:

- Level 2 maps are generated by using the High-Pass Filtering (HPF) pipeline. Large scale structures (noise and extended emission) are removed by means of a sliding median filter on individual bolometer timeline. These maps are not suited to recovering extended emission.
- Level 2.5 maps are generated by combining scan and cross-scan observations of the same sky field (acquired in the same observing mode), by using three different mappers: HPF, [Unimap](#) and [JScanam](#). Level 2.5 HPF maps are the averages of the corresponding Level 2 HPF maps. The Unimap mapper exploits the Generalised Least Square approach with the pixel noise compensation for removing 1/f noise, while the JScanam mappers is a Java implementation of [Scanamorphos de-striper method](#) that does not rely on any noise model nor on filtering and exploits the redundancy to derive the drifts from the data. Unimap and JScanam maps are science ready, not absolutely calibrated and are reliable for recovering both extended emission and point-like sources.
- Level 3 maps are the average of Unimap and JScanam Level 2.5 maps of any overlapping areas. Additional criteria are that: the scan speed is uniform for all the observations; the observations combined are from the same observing program; they contain at least 180 seconds worth of data; there is at least one scan and cross-scan observation in all combined observations and that for the blue channel data only, the filter is the same. For the red channel, all observations are used without regard to the filter used in the blue channel.

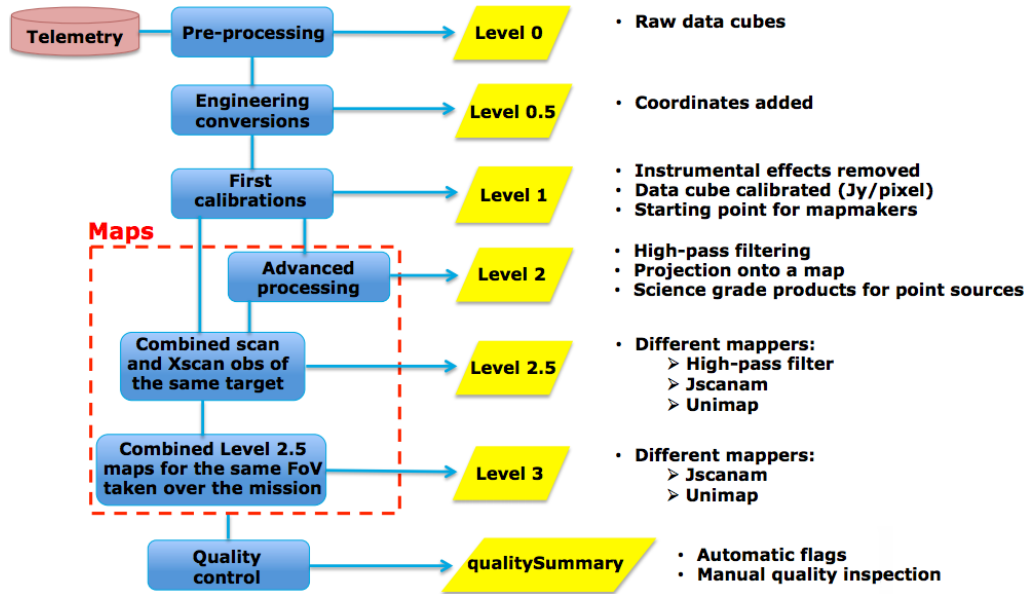


Figure 2: SPG chain from Level 0 to Level 3. The raw data are stored in Level 0 cubes. After numerous levels of noise removal and processing, the Level 2.5 and Level 3 maps are the final products that are used for the HPPSC generation.

### 3 Catalogue generation – data & methods

#### 3.1 Map repository

The HPPSC makes use of JScanam SPG 13.0.0 maps because they are considered to be science ready for catalogue purposes. The JScanam pipeline is stable and has been validated from SPG12 on. The updates produced up to the legacy version of the software (SPG14.2.1) do not introduce changes in the JScanam maps. That is also the reason why the JScanam Level 2.5 maps are selected as the standard maps in the HSA. Although in JScanam maps (as for all PACS maps) the zero-level of the background surface brightness is not known, it is not a concern when one performs photometry of point or compact sources, either as aperture or as synthetic photometry, because a certain background level is always subtracted in the source flux extraction procedure. The absolute calibration for point sources is very accurate, with a relative accuracy  $<5\%$  in all PACS bands [Balog et al. \(2014\)](#)).

We considered the most reliable and deepest maps available in the HSA for a specific field for our repository. When available, the JScanam Level 3 map is taken for each observed field, otherwise we use the related Level 2.5 maps. In some cases, HPF Level 2 maps are the highest level products available. These are not included in the repository, because of the low reliability of photometric analysis of Level 2 maps and the impossibility of generating structure noise ( $N_S$ ) maps (see Section 3.2) of high-pass filtered maps, given that high-pass filtering removes the large scale structures from the maps.

In some cases, large surveys made use of scanning strategies that prevented the generation of standard Level 2.5 products so these sky regions are not included in the HPPSC, nor, as mentioned earlier, are observations acquired with the chp/nod observing mode included in the catalogue, limiting the coverage of our Catalog to 4.5% of the sky (see Section 1.3.3).

The HPPSC map repository includes Level 2.5 observations of Solar System Objects (SSOs), re-processed in the fixed sky frame and not in the co-moving frame of the target, as it is performed in the standard processing. The rationale is to exclude SSOs from the catalogue, while including all the serendipitous detections of the SSO fields.

Pointing issues affecting the pointing were observed in SPG13 products during several ODs, introducing offsets for one of two reasons:

- Switch from the main star-tracker (STR1) to the backup unit (STR2) either due to the autonomous Failure Detection, Isolation and Recovery (FDIR) or commanded from ground for testing and calibrating the backup unit. This produces an offset of  $\sim 15''$  between the desired (nominal) and actually commanded coordinates.
- Incorrect upload of the Spacecraft Velocity Vector (SVV) to the star-tracker (STR) resulting in a incorrect light aberration correction in the coordinates of the STR guide stars. This produces a variable error in the range 0-20'' between the desired (nominal) and actually commanded coordinates.

The problem has been fixed in SPG14. The PACS photometric observations affected by these anomalies (151 individual maps, 1% of the repository) were updated with SPG14 products in our repository. The number of maps used for creating the HPPSC is detailed in Table 2.

SPG	Parallel Mode						Scan map					
	13.0.0		14.0.1		14.2.0		13.0.0		14.0.1		14.2.0	
Level	2.5	3	2.5	3	2.5	3	2.5	3	2.5	3	2.5	3
blue	270	4	7	0	0	0	3471	86	25	0	13	1
green	33	15	2	0	0	0	3320	63	18	0	10	2
red	292	25	9	0	0	0	3924	1558	33	5	19	7

Table 2: Number of maps in our repository in the different bands at different levels and different SPG versions.

	blue			green			red		
Scan speed [ $''\text{s}^{-1}$ ]	10	20	60	10	20	60	10	20	60
Parallel Mode	0	51	275	0	26	24	0	51	275
Scan map	14	3987	17	14	3782	10	24	6314	23

Table 3: Number of maps in our repository in the different bands and obtained at different scan speeds.

### 3.2 Structure noise maps

The confusion noise present in the FIR and smm photometric observations is a major limiting factor in sensitivity and photometric accuracy. Confusion noise comes either from the extragalactic background or from the cirrus emission of the Galaxy. Both are mainly due to the presence of the cold dust. It is essential to characterise the properties of the catalogue well. Completeness, flux boosting, and flux error at different sensitivity levels (depending on the depth of the observations) can be evaluated by creating simulations. We injected sources with well controlled flux into observational timelines and reprocessed the new data in the same way as the real data was. As expected we found that in extragalactic observations we are more complete than in the Galactic star forming regions. Also, the photometric accuracy is highly dependent on the complexity of the celestial environment. As our goal was to create a homogeneous catalogue, we had to find a way to treat all environments in the same way.

The structure noise ( $N_S$ ) measures the pixel fluctuation in the map (Kiss et al. 2005). It can be translated into the fluctuation power of the neighbouring areas, and gives a local information instead of a general (regional average) number. For each pixel of the map the  $N_S$  is calculated as the standard deviation of all flux differences between the pixel and all the surrounding pixels at a fixed distance (see Figure 3). It includes the spatial noise of the celestial environment as well as the noise contribution from the instrument.

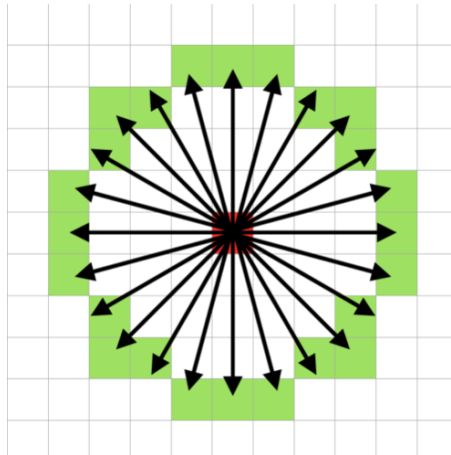


Figure 3: The circular configuration used for  $N_S$  calculation. The target pixel is shown with red square. The neighbouring pixels at the predefined angular distance are shown with green squares.

The  $N_S$  maps were generated with an IDL script. The code reads the same maps that were used to generate the catalogue. For each pixel of the map (target pixel, shown as a red square in Figure 3) we are looking for neighbouring pixels in 24 directions (green squares) at  $\sim 30''$  angular distance. If fewer than 24 pixels are found, the unique pixels are selected. The next step is to calculate the absolute difference between the unique neighbouring pixels and the target pixel. When the number of unique data points is greater than three, the standard deviation of these differences is stored as a pixel value in place of the target pixel. The pixel value is calculated according to :

$$\sigma_{strn} = \sqrt{\frac{1}{24} \sum_{i=1}^{24} (d_i - \mu)^2}, \quad (1)$$

where  $d_i$  are the differences of each of the fluxes of the 24 (or fewer) pixels and that of the central pixel, and  $\mu$  is the average of all  $d_i$ . The resulting  $N_S$  map is then stored as a standard FITS file with the header of the original map.

The distance between the target pixel and the neighbouring pixels corresponds to a spatial frequency, that had to be optimised. Choosing too small distance would mean that the  $N_S$  value includes the flux from the PSF wings causing a scaling with the source flux. To minimize this effect we created  $N_S$  maps of simulations where we injected sources with 20 Jy flux and increased the angular distance between the target pixel and the neighbouring pixels. On each  $N_S$  map the  $N_S$  value at the position of our artificial sources was calculated and checked as a function of the angular separation (see Figure 4). We found that at an angular distance of  $\sim 30''$  the  $N_S$  value decouples from the source flux and has a minimum value before the large scale structures start to dominate the fluctuation power. We decided to use this angular separation to create our  $N_S$  maps in all bands and to attach a  $N_S$  value to each of our detections. The average values where the minimum values were found for the different modes and bands are listed in Table 4.

	blue ["]	green ["]	red ["]
Parallel Mode	$32 \pm 3$	$31 \pm 3$	$31 \pm 7$
Scan map	$28 \pm 3$	$26 \pm 6$	$27 \pm 5$

Table 4: The smallest angular distance where the source flux decouples from the  $N_S$  value.

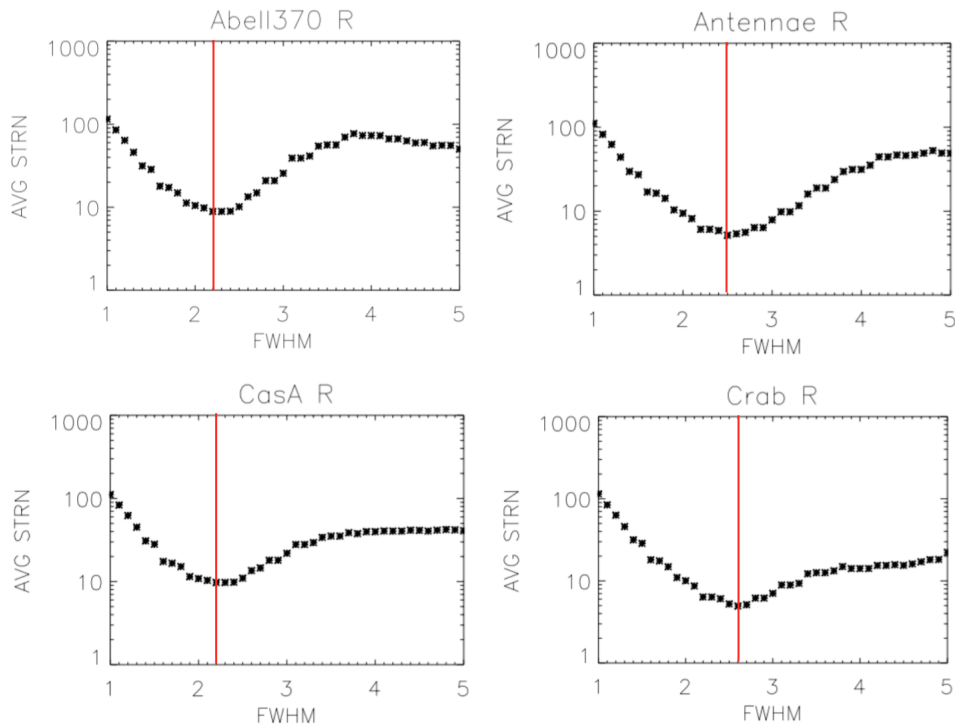


Figure 4: Examples showing the average  $N_S$  value of the injected sources in red Scan map simulations as a function of the angular distance between the target pixel and the neighbouring pixels. The minimum of the curves (indicated by red vertical lines) represent the angular separation where the source flux has the smallest impact on the  $N_S$ .

The pixel values have the same units as the input map. To avoid so-called NaN-Donuts, single NaNs in the input map are interpolated before the calculation.

The  $N_S$  value for a given source is calculated as follows: we place an aperture at the extracted position of each source, the diameter of which equals to the aperture size used for the photometry, i.e.  $6''$  for the blue camera and  $12''$  for the red camera. The total  $N_S$  inside the aperture is then

converted into units of  $\text{MJy sr}^{-1}$ . Eventually, the resulting  $N_S$  value is attached to each source in a separate database column.

### 3.3 Pipeline description

We created our pipeline using HIPE release 13.0.0. Our goal was to do as many tasks as possible within the environment to be consistent all the way from observational data processing to photometry.

The steps of the workflow are shown in Figure 5. Before executing the actual pipeline, the first step is to find the next unprocessed map. This step thus required a list of observations (Observation Table, see Section 4.3.3) that was used to check if a map was already processed. If the flag in the Observation Table allowed, the pipeline started by reading the corresponding L3/L2.5 map.

The  $N_S$  maps for the given map were created in IDL, as described in Section 3.2, and at this stage of the process they were read into HIPE.

The detection algorithm is applied on specific Region of Interest (ROI) of the maps, in order to identify the map regions with an appropriate coverage values. Slewing and turnaround regions are rejected and on the scan regions where scan and cross-scan coverage is guaranteed at constant speed, a threshold equal to 0.6 times the normalised coverage value is considered for the definition of the ROI. This threshold is not used on Parallel observations, because generally their scanning strategy does not exploit the homogeneous coverage on the field and the threshold approach generates holes in the ROI maps. For a Parallel map, the ROI is defined as the intersections of all the scanning regions of the observations that compose the map.

The JScanam map was then used to locate sources. This was done by the HIPE implementation of SUSSEXtractor, called sourceExtractorSussextractorTask (see Section 3.3.1). The provided sky coordinates are the J2000.0 Right Ascension (RA) and Declination (Dec) in units of degrees. The positions are used by four different processes:

1. Aperture photometry by `annularSkyAperturePhotometry()` task
2. `sourceFitter()` task
3. background aperture photometry for the background RMS calculation
4. aperture photometry on the  $N_S$  map to obtain the  $N_S$  value

All the resulting data are stored in a separate **comma separated value** (csv) table. Instead of handling all the individual csv tables, we created a relational PostgreSQL database, that allowed us to handle our data in a more efficient way.



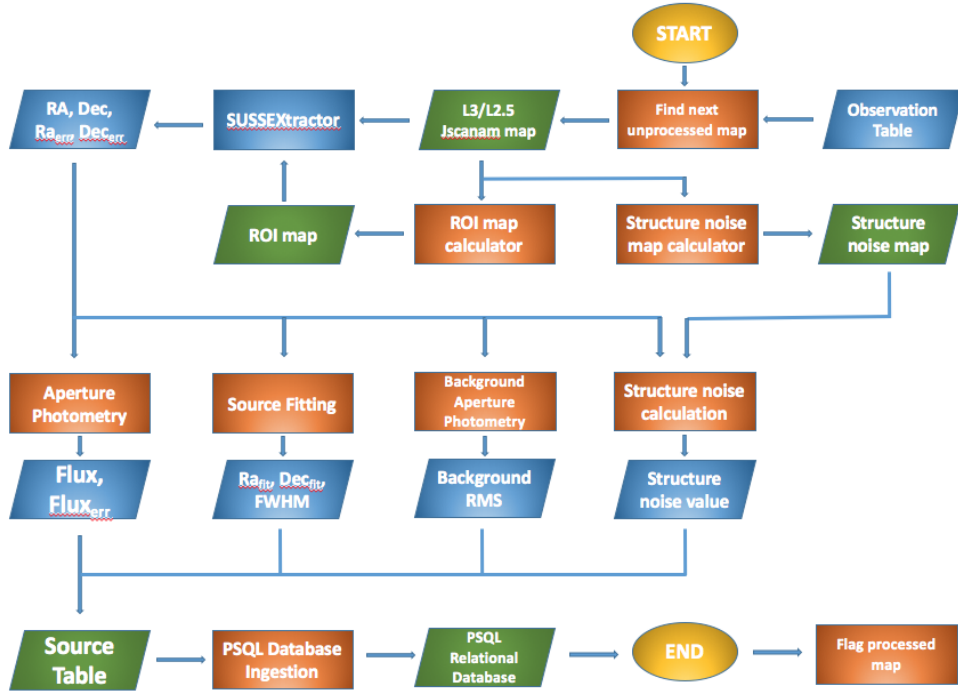


Figure 5: The HPPSC pipeline. The workflow shows the steps that were performed on each map from the START to the END. This workflow was repeated for all maps.

### 3.3.1 Source detection

Source detection in the presence of the cirrus emission is a non-trivial task. There are several tools specifically designed for identifying sources on top of filamentary structures. Such algorithms are *getsources* or *CuTE<sub>x</sub>*. For extragalactic/cosmological observations *Starfinder* is often used (e.g. PEP). These are very sophisticated tools that can be tuned for environments with various degrees of complexity. Their fine tuning and also their performance is, however, not sufficient for our purposes. An extensive test for source detection was carried out for the *Herschel/SPIRE Point Source Catalogue*. Our findings are detailed in "Testing Point Source Extractors" section of the *SPSC Explanatory Supplement*.

Following our findings for the SPSC, the `sourceExtractorSussextractor()` task of HIPE (hereafter simply `SUSSEXtractor`), which is an implementation of the `SUSSEXtractor` algorithm [Savage & Oliver 2007](#) was used by our pipeline. This task requires an error map for which we used the "stDev" map of the JScanam products. The detection threshold was set to 3. The algorithm can use a custom PRF, hence we used the PACS PSFs as PRFs according to the wavelength and to the observation mode. Despite using the proper PRFs, the algorithm also requires the FWHM of the sources that we are looking for. Our FWHM values in the blue, green and red bands were set to 5.5", 7.0" and 11.5". The most important information collected with the `SUSSEXtractor` was: RA, Dec, image X and Y pixel coordinates, RA error and Dec error.

### 3.3.2 Flux extraction

Unlike for SPIRE sources, photometry of PACS sources can at present only be obtained by aperture photometry. Therefore flux extraction was carried out using HIPE's `annularSkyAperturePhotometry()` task. This task performs aperture photometry of the target, enclosed by a circular aperture. The sky is estimated from a concentric annular aperture, with configurable inner and outer radii. In our case the inner and outer radii were set to 25" and 35", respectively. The task is able to perform five possible sky estimation. We used the default `daophot` estimation, which is an implementation of the IDL `mmm.pro` routine.

The flux errors provided by the task are calculated in the same way as in the IDL `aper.pro` routine. The error for the target flux including the background is defined as the square-root of the (absolute value of the) total flux in the target aperture (including the background). The error for

the sky is defined as the standard deviation of the sky value. Note that this standard deviation is calculated only with the sky intensities that are used to calculate the sky intensity (as there are sky estimation algorithms that reject sky values). The squared error for the target flux is thus:

- the absolute value of the target flux from which the sky contribution has been subtracted
- the standard deviation of the sky value \* the surface of the target aperture
- the squared standard deviation on the sky value

The `annularSkyAperturePhotometry()` task was used for three different purposes.

- First of all, this task was responsible for the accurate source flux density measurement of the detected sources. The input coordinates were provided by `SUSSEXtractor`. We used 11 different aperture sizes to measure the flux. The aperture radii in the red band were 9,10,11,12,13,14,15,16,17,18 & 22 arcseconds and 4,5,6,7,8,9,10,11,12,13 & 18 arcseconds in the blue and green bands. The inner and outer edge of the sky annulus was set in 25'' and 35'' in all cases. After applying the aperture correction, we were able to set quality criteria based on the flux values measured in the different apertures.
- The  $N_S$  was also obtained with aperture photometry from the  $N_S$  maps. We placed 6'' and 12'' apertures (depending on the camera) at the position of the source, on the  $N_S$  maps. The  $N_S$  inside the aperture was converted then to units of  $\text{MJy sr}^{-1}$ .
- It was also used to measure the background RMS

### 3.3.3 2D Gaussian fit

The `HIPE sourceFitting` task is a task to fit a two-dimensional Gaussian to a source in a specified rectangular region on an image. It can use either an elongated or a circular Gaussian, and can fit a tilted background depending on whether or not there is a slope in the background. For the circular fit the model is a `Gauss2DModel + PolySurfaceModel`, the initial parameters are a) the maximum value in the cropped region, which provides an initial value for the centre of the Gaussian, and its peak value, and b) the background, which is estimated via the `Daophot` algorithm (as for aperture photometry) on the whole rectangle. In the elongated case the initial parameter a) is the fit that would be made if the source were circular, which yields an initial value for the centre and the peak value of the Gaussian and for the background, while b) the sigma's and the rotation angle of the Gaussian are estimated from profile plotting. For all straight lines through the estimated centre with an angle between 0 and 180 degrees w.r.t. the x-axis (in steps of 15 degrees), a Gaussian combined with a constant function is fitted (deriving the initial parameters from the circular 2D-fit) and the largest sigma is taken to be the initial value for the width of the Gaussian in the x-direction and the corresponding angle as the initial value for the rotation angle of the Gaussian. The sigma in the y-direction is estimated from the fit of the line perpendicular to the previous one. The fitter used in all cases is the `LevenbergMarquardtFitter`.

The `sourceFitting` is applied on the maps at the positions provided by the `SUSSEXtractor` task. The assumption is that the objects are always slightly elongated, since the PSF is also not circular, thus the pipeline first tries to fit an elongated Gaussian on a rectangle in a square of  $11 \times 11$  pixels. If the fit did not converge in  $10^4$  iterations, then a circular fit was applied instead. The task provides a set of output parameters suitable for quality assessment. From all parameters derived of 2D Gaussian fitting (see Section 3.3.2), we use the Gaussian peak (RA, DEC) and the  $\sigma_X$  and  $\sigma_Y$  to set the requirements that HPPSC sources have to be met (more details in Section 3.5). The  $\sigma$  values of the fitted Gaussians were converted to FWHM values, by multiplying them by  $2\sqrt{2 \ln 2}$ .

## 3.4 Quality assessment

The pipeline collected a large amount of data that were essential to analyse further the detections and to create a process that is able to qualify the listed sources. First of all, the flux error provided by the `annularSkyAperturePhotometry` task was not intended to calculate the noise for bolometers in the presence of a strongly fluctuating background. Therefore we had to find methods that are able to give us a proper estimate of the noise. The workflow for quality assessment is presented in Figure 6. Our process used internal flags to indicate if an entry from the database matches our quality criteria, which we detail below.

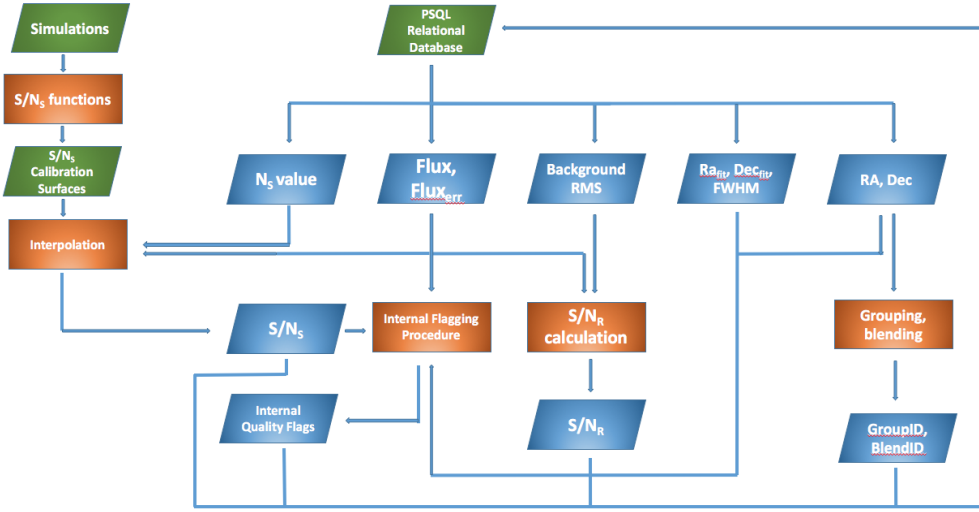


Figure 6: Workflow diagram for quality assessment.

### 3.4.1 Structure noise based S/N

As it was described in the [SPSC Explanatory Supplement](#), a method was developed to give a proper statistical error of the photometry. As described in the Section 9 below, we injected sources into fields with various complexity. The same pipeline that we used to detect our sources and to collect photometry from real observations was used to detect our simulated sources and to measure their flux. The  $N_S$  values were collected also in the same way. This procedure allowed us to compare the input flux to the measured flux as a function of the  $N_S$ .

The signal to noise value depends on the source flux and on the noise, in our case the  $N_S$  that is the combination of the instrument noise and the sky confusion noise. The theoretical flux ( $F_{in}$ ) for each injected source is known and we have also the  $N_S$  values for each. For a given  $F_{in}$  level the  $N_S$  was binned. The bin size was  $25 \text{ MJy sr}^{-1}$ . In each bin we calculated the ratio of the measured flux ( $F_{out}$ ) and  $F_{in}$ . The standard deviation of the  $F_{out}/F_{in}$  ratio ( $\sigma_F$ ) is the photometric uncertainty, that we were looking for. By dividing  $F_{in}$  by  $\sigma_F$  we obtain the S/N. This S/N value is just for any given  $N_S$  values. To calculate it for any  $N_S$  levels, we fitted the following function to the data:

$$S/N = a \times N_S^b \quad (2)$$

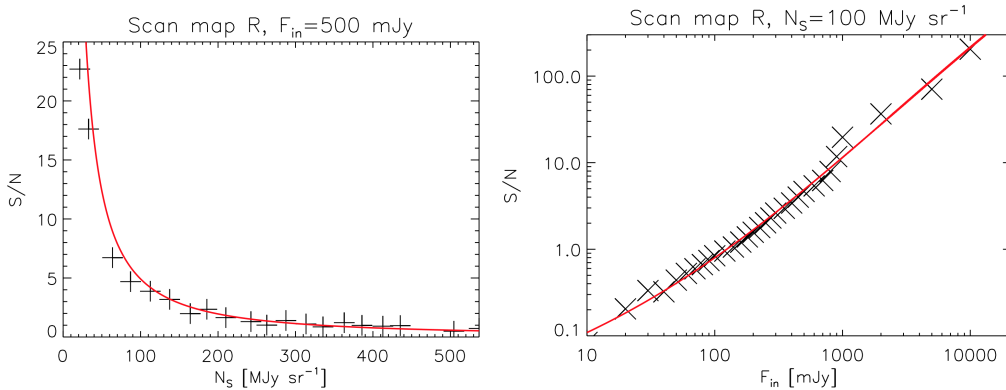


Figure 7: Left: The photometric uncertainty ( $\sigma_F$ ) as a function of the  $N_S$  for sources with 500 mJy  $F_{in}$  (black plus signs). Right: The photometric uncertainty ( $\sigma_F$ ) as a function of the  $F_{in}$  for 100  $\text{MJy sr}^{-1}$   $N_S$  (black plus signs).

The left panel on Figure 7 shows the function fitted to the data points, allowing us to calculate the S/N value for any  $N_S$  value. The data points were weighted by the number of elements in each bin.

To have a grid sufficient sampled that can be used for interpolation, we created a one with steps of  $1 \text{ MJy } sr^{-1}$  in the  $N_S$  direction and  $1 \text{ mJy}$  in the  $F_{in}$  direction. The above function was used to calculate the S/N values for all  $N_S$  values on the grid, but only for the 32 input fluxlevels. To fill up the grid, we had to find a function that could be used to calculate the S/N for all flux levels at give  $N_S$  level. The following equation was used:

$$S/N = a + \frac{b}{c + \exp\left(\frac{x-d}{e}\right)} \quad (3)$$

The right panel of Figure 7 shows the fitted function. After repeating the fit for all desired  $N_S$  levels, the result is an array that can be used to calculate the S/N for the catalogue sources. Figure 8 shows the resulting arrays for each band and for the two observing modes. The S/N value for the objects in the catalogue was calculated by interpolation/extrapolation using the corresponding S/N array.

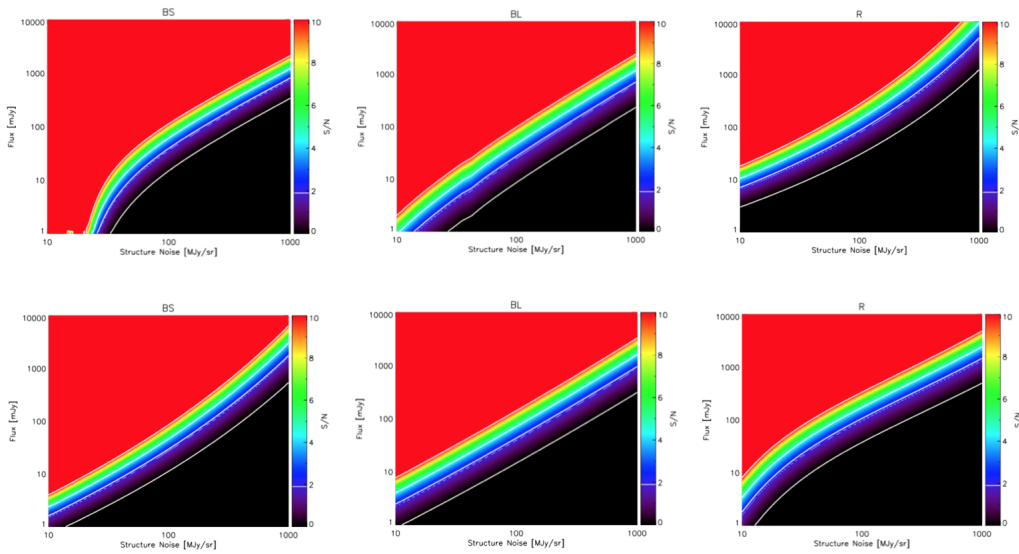


Figure 8: The signal to noise ratio (S/N) surfaces for the different bands and observing modes. The upper row shows the S/N as a function of  $N_S$  and flux in the  $70\mu m$ ,  $100\mu m$  and  $160\mu m$  bands (from left to right) for Parallel Mode observations. Similarly, the bottom row shows the SNR surfaces for scan map observations. Solid white lines indicate S/N levels 1, 3, 5 and 10.

Although we created a large number of simulated maps and injected several thousand sources at each flux levels, our simulations cover only a part of the Flux– $N_S$  parameter space. Figure 9 shows where the recovered simulated sources are located on the plane. Our experience shows that the calculated S/N is reliable only between 0 and  $2000 \text{ MJy } sr^{-1}$   $N_S$  and between 0 and  $20 \text{ Jy}$  source flux. Although these ranges cover  $\sim 98\%$  of our sources, we used additional methods to estimate a proper S/N for each source, including the brightest ones that are important quality indicators for such a catalogue.

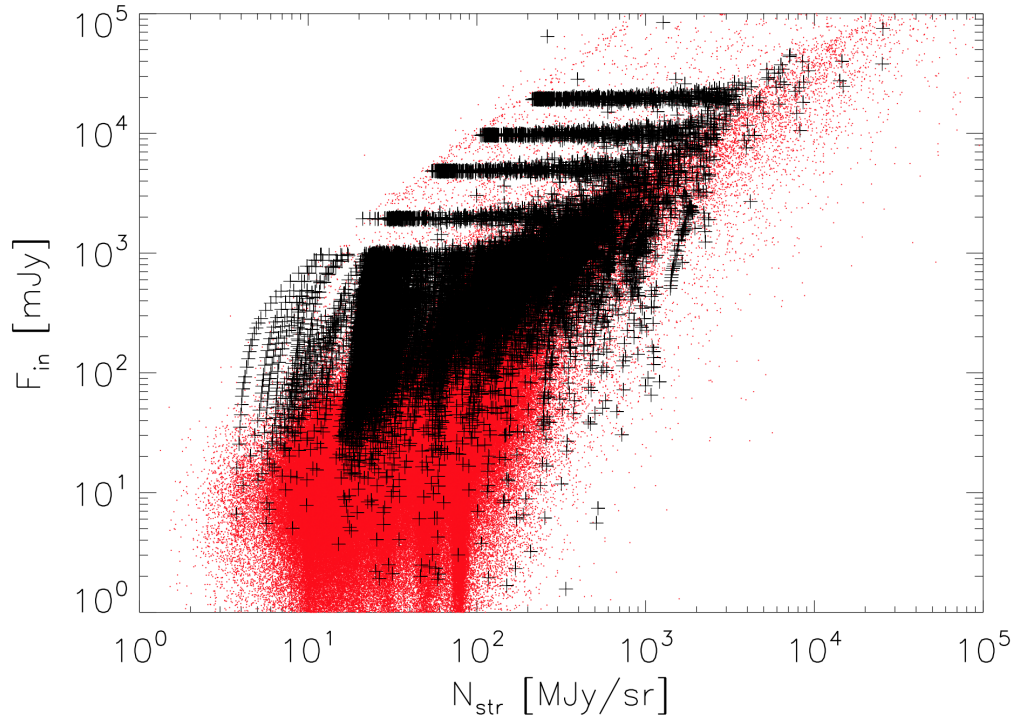


Figure 9: The simulated (black plus sign) and Catalogue (red dots) sources on the  $N_s$ -Flux diagram.

### 3.4.2 Background RMS based S/N

A different way to estimate the properties of the noise around the source in an observation is to place various apertures around the source and to measure the flux in these apertures. The HIPE pipeline "scanmap Pointsources Photproject" follows this idea. The script creates high-pass filtered maps and then calculates the photometry at any given input coordinates.

In order to speed up the up the procedure we implemented the photometric part as an IDL script. The core of the algorithm is to place apertures on a grid of  $7 \times 7$  points with an interval of 5 pixels, centred on the source. The flux was measured in each aperture and corrected for the EEF. An initial mean and standard deviation were calculated. Then, those apertures, in which the flux was higher than the mean plus twice the standard deviation, were excluded. The remaining apertures we term the blank apertures. The RMS of the background was calculated as the standard deviation of the flux measured in the blank apertures.

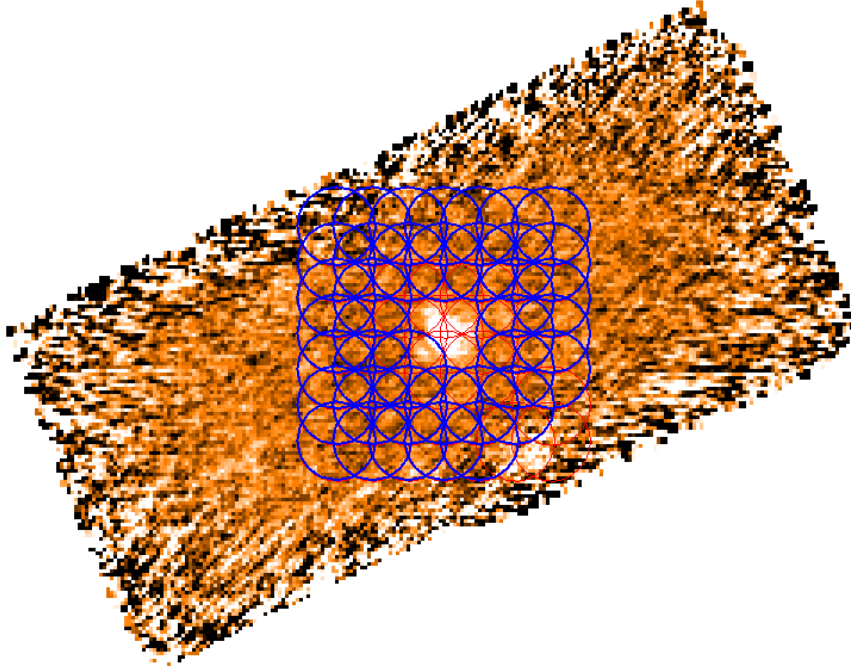


Figure 10: A red band image of  $\gamma$  Dra. The background RMS is calculated as the standard deviation of the flux measured in the blank apertures around the source (blue circles).

### 3.4.3 Fitted positions

A good quality indicator for our detections is the angular distance between the sky position provided by SUSSExtractor and the position of the 2D Gaussian fit performed by the sourceFitting() task. We found that the distance between the two positions is very small for the reliable extractions. Figure 11 shows the distribution of the separation for scan map observations (left panel) and for Parallel Mode observations (right panel).

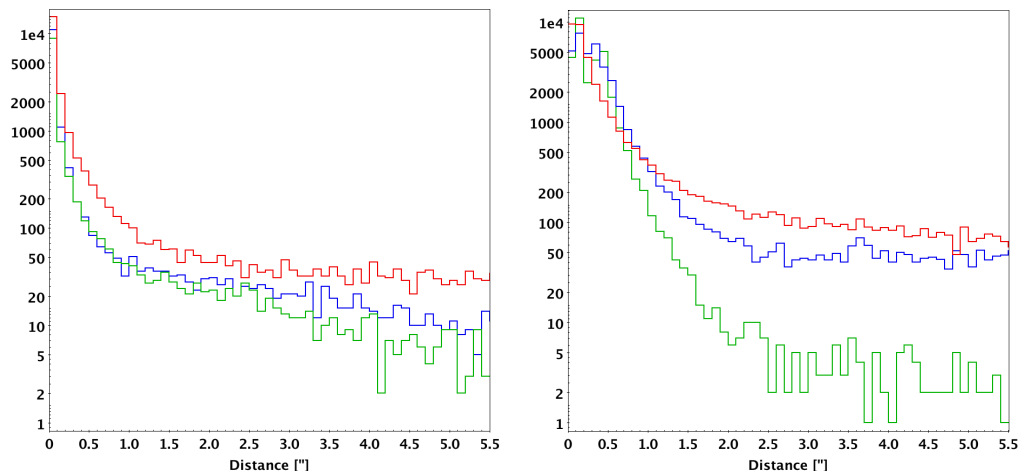


Figure 11: Position uncertainty based on simulations for the scan map (left) and Parallel Mode (right). The uncertainty is measured as the angular distance between the SUSSExtractor position and the 2D Gaussian fit source position. The blue, green and red bands are indicated with the corresponding colours.

### 3.4.4 Fluxes

As described in Section 3.3.2 the flux of our sources was measured with 11 apertures of different sizes, and the aperture correction applied to each. As shown for the calibration star HD18884 ( $\alpha$ Cet) in Figure 12, the flux measured in the different apertures is stable after applying the aperture correction, as only a few per cent difference can be seen. It is valid to assume that all sources having a PSF-like shape and flux distribution have a similar behaviour.

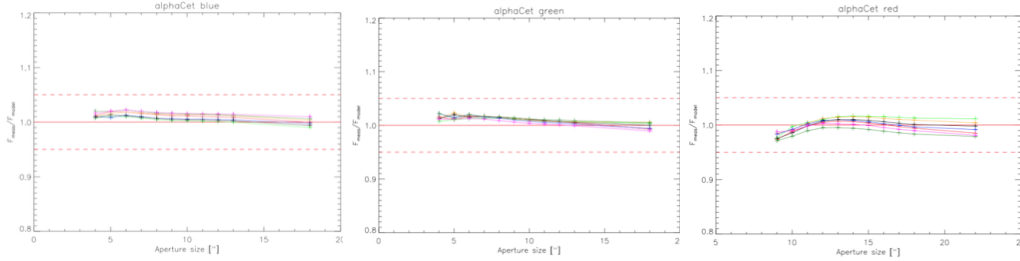


Figure 12: Measured flux over the model flux for the calibration star HD18884 for different aperture size, plotted along the horizontal axis in the blue (left), green (middle) and red (right) bands. Observations made at different epochs are indicated with different colours.

### 3.4.5 FWHM

The `sourceFitting()` task of HIPE fitted a 2D Gaussian to our detections. If an ellipsoidal fit was possible, the fit resulted in two FWHM values, measured along the major and the minor axis. If only a circular Gaussian could be fitted then only one FWHM value was added to our database. These FWHM values were compared to the PSF FWHM values. Figure 13 shows the FWHM value distribution for our artificial sources in different band and at different scan speeds in scan map and Parallel Mode observations. Our simulations showed that both the shape of the PSF and the measured FWHM can be affected by the background emission, especially at faint flux levels.

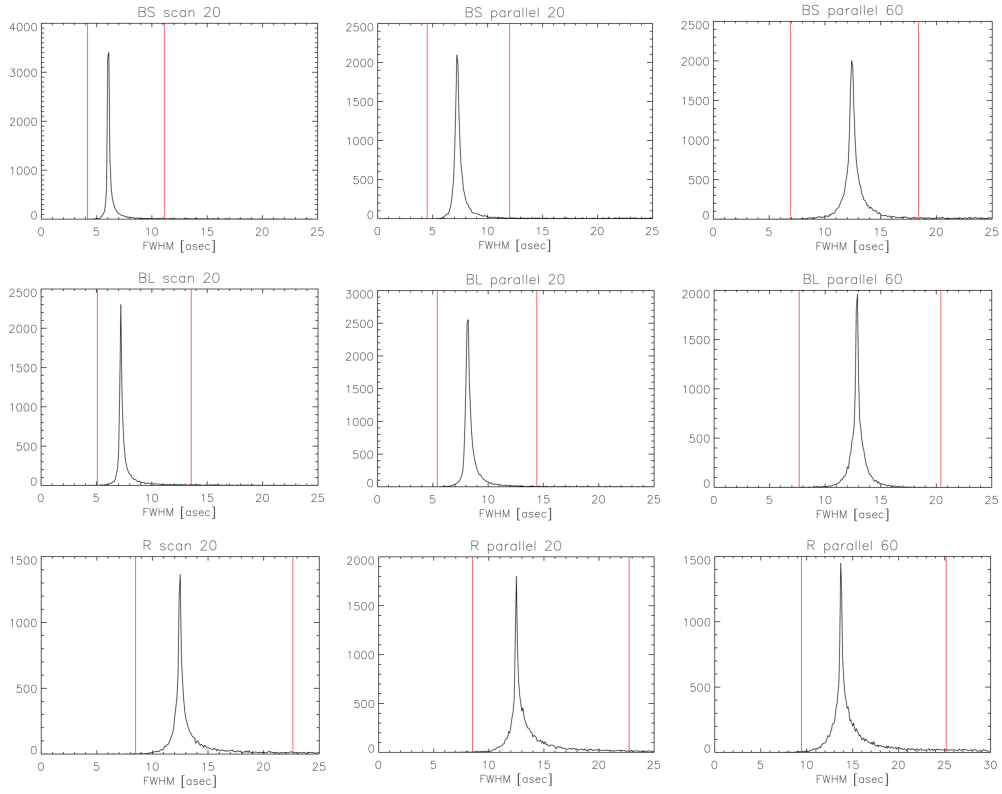


Figure 13: Distribution of the FWHM values of our simulated sources in the different bands and at different scan speeds and observing modes. 0.75 and 2.0 times the PSF FWHM values are marked with red solid vertical lines.

### 3.5 Object consolidation

In order to increase the reliability of our products, and to consolidate the final catalogue, we went through a process of objects consolidation. The first part was the cleaning of our source tables. We analysed the parameters derived from the 2D Gaussian fit (centroid, FWHM, elongation), the signal-to-noise ratios of the sources, and their flux distribution. The second part was to identify the multiple detections of the same astronomical object and to decide which entry should appear in the HPPSC.



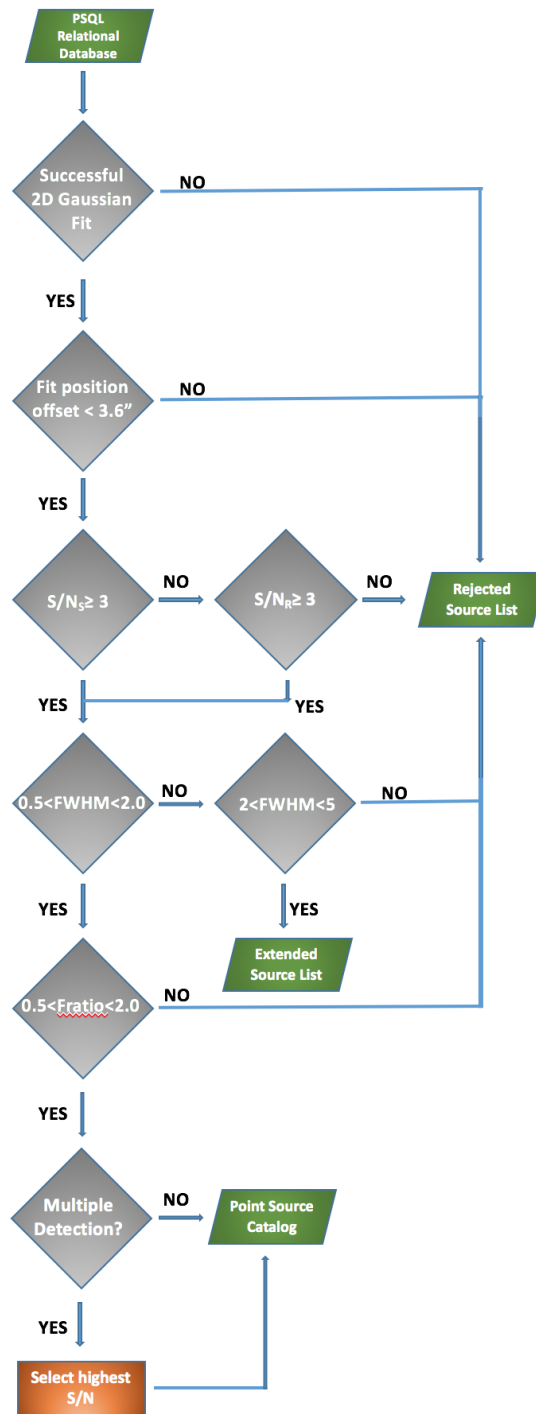


Figure 14: The workflow of object consolidation. Criteria based on different properties of the sources were applied and sources classified accordingly as point source, extended source, or rejected source.

### 3.5.1 Cleaning source tables

As stated previously, the cleaning process consists of imposing different conditions on the parameters obtained in previous steps (pipeline procedure and quality assessment) to classify all detected objects in our list as point, extended or rejected sources (see the workflow of object consolidation in Figure 14).

The first step is to confirm that a 2D Gaussian fit is achieved. If the fit does not converge,

the object is automatically rejected, and add to the HPRSL. Of the sources in our initial list (8 254 064), 1 041 616 objects were discarded at this point.

The second condition is to examine the difference between the coordinates derived by the SUS-SExtractor and the sourceFitting task. The distance between them has to be less than  $3.6''$ , i.e, three times the mean *Herschel* pointing accuracy. As shown Figure 11, this limit is supported by the positional uncertainty from simulations. Again, if this criterion is not satisfied, we do not consider this source as a good candidate and we move it to the HPRSL.

As described in Section 3.4, we provide two different methods for the estimation of the signal-to-noise ratio. Our condition is fulfilled if at least one of them gives  $S/N > 3$ . All sources with both  $S/N$  ratios below 3 are rejected and are included in the HPRSL. Check the columns  $S/N_S$  and  $S/N_R$  in the HPPSC and HPESL to identify which criteria are met.

As the next step, we analyse the FWHM of the 2D Gaussian fit in order to reject other possible spurious sources, although mainly to classify the source as "point" or "extended". An object is considered a point source if the value of FWHM satisfies the criterion that:

$$0.75 \times PSF < FWHM < 2 \times PSF \quad (4)$$

The thresholds were determined using a FWHM based on simulations (see Figure 13). As explained previously, the 2D Gaussian fit provides values for the FWHMs in both the X and the Y direction. If the 2D Gaussian is elliptical, the FWHM values in both axes need to meet this criterion. The PSFs were taken from Table 1. Alternatively, if a source is not included in the HPPSC and the FWHM is between 2 and 5 times the tabulated PSF, it is added to the Extended Source List, all other sources are rejected. After this step, the HPESL is complete. The final number of extended sources identified is 546 587.

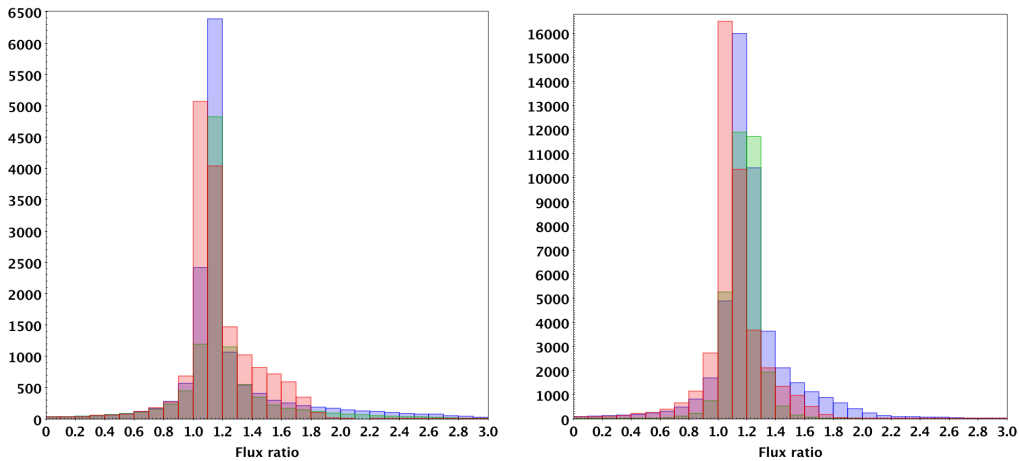


Figure 15: Histograms of ratio fluxes (aperture6/ aperture1, see text) for the scan map (left) and Parallel Mode (right). The blue, green and red bands are indicated with the corresponding colors

The final condition takes for the flux distribution of the source. To do this, we check the ratio between two fluxes measured from two different apertures, aperture 6 ( $9''$  for blue and green bands,  $14''$  for the red) and aperture 1 ( $4''$  for blue and green bands,  $9''$  for the red band). To include an object in our final point source catalogue, this flux ratio must be between 0.5 and 2. These limits are established using results from simulations. Figure 15 shows the flux ratios obtained to simulated sources, for both Scan and Parallel mode. Note that this condition does not apply to extended sources. At this point, the number of consolidated point sources is 522317. However, the catalogue still contains duplicated sources. In the next section, we explain the procedure used to group these duplicated objects to a single source detection in the HPPSC.

### 3.5.2 Multiple detections

Once the Source List is consolidated using the quality criteria described above, a further step is necessary to generate the final catalogue table. The source extraction is performed on a map repository that includes, the highest processing Level available within the HSA for every sky region, and does not prevent multiple detection to the same source appearing in different sky maps. To identify single objects with multiple detections in overlapping maps, sources belonging to different observations and separated by a smaller distance on the sky than within the PSF FWHM are grouped. Within each group, only the source with the highest S/N (the average between RMS S/N and Structure Noise S/N is taken - see Sections 3.4.1, 3.4.2) is included in the final catalogue table. The PSF values are those given in Table 1, they depend strongly on the observing mode, while the absolute pointing error for SPG13 (1.2'') is adopted as uncertainty.

## 4 Point Source Catalogue

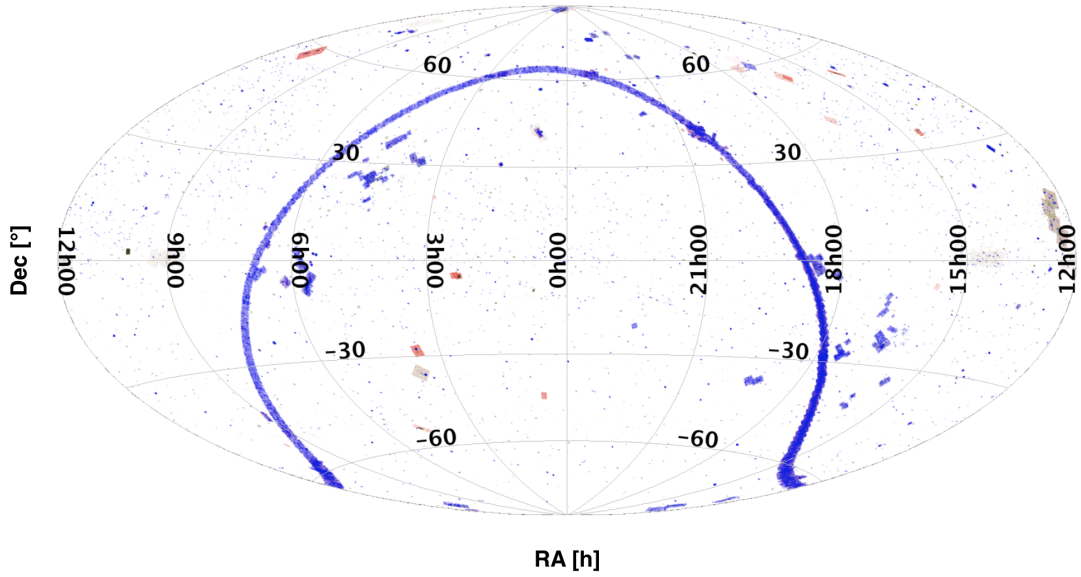


Figure 16: All-sky Aitoff projection of the HPPSC sources across the three PACS bands.

### 4.1 Catalogue properties

The objects of the *Herschel*/PACS Point Source Catalogue has been identified selecting carefully the parameters that best define the point sources and applying our filtering criteria to the raw source table based on positional uncertainties, FWHM, S/N, flux ratio and elongation to eliminate spurious sources and extended sources. The HPPSC is presented in three comma-separated value (CSV) files with headers that identify the columns. We include 251 392, 131 322 and 108 319 sources in the red, green and blue bands, respectively. Detailed numbers of sources included in the different tables of the HPPSC are found in Table 5. The sky coverage of the extracted sources is shown in Figure 16, using the J2000 equatorial system in an Aitoff projection. The Galactic Plane is clearly visible as a blue stripe on the image, indicating area covered by the observations of the [Hi-GAL](#) Key Program.

	blue			green			red		
	HPPSC	HPESL	HPRSL	HPPSC	HPESL	HPRSL	HPPSC	HPESL	HPRSL
Scan map $10'' s^{-1}$	319	143	12728	391	72	9532	435	219	9810
Scan map $20'' s^{-1}$	17010	10573	329657	87278	23859	1211037	46231	37597	896410
Scan map $60'' s^{-1}$	177	175	5874	270	42	5682	441	578	6884
Parallel Mode $20'' s^{-1}$	3343	11534	118274	7497	6043	63870	44325	48438	832131
Parallel Mode $60'' s^{-1}$	87470	145634	1680660	35886	22525	479784	159960	239155	1554111

Table 5: The number of sources at different scan speeds and in different observing modes in the products.

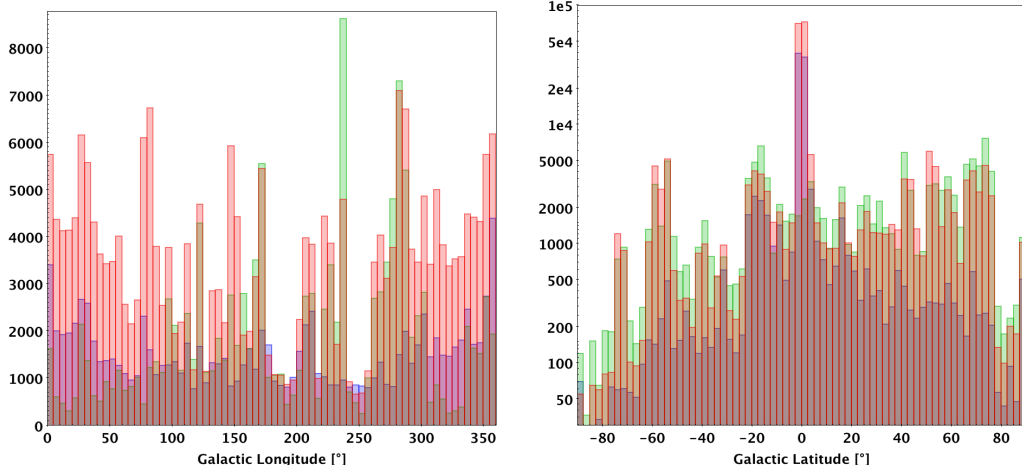


Figure 17: The HPPSC source distribution in the blue, green and red bands, indicated with the corresponding bar colours, as a function of the Galactic latitude and longitude.

Figure 17 shows the distribution of our HPPSC sources as a function of Galactic Longitude (left) and Latitude (right). On the Galactic latitude figure the peaks in the blue and red bands are caused by the Hi-GAL sources. An excess of source counts is visible in the green and red bands between  $\sim 40^\circ$  and  $\sim 60^\circ$ , caused by the large extragalactic fields observed for HeViCS. Other large surveys can be also identified based on the source distribution.

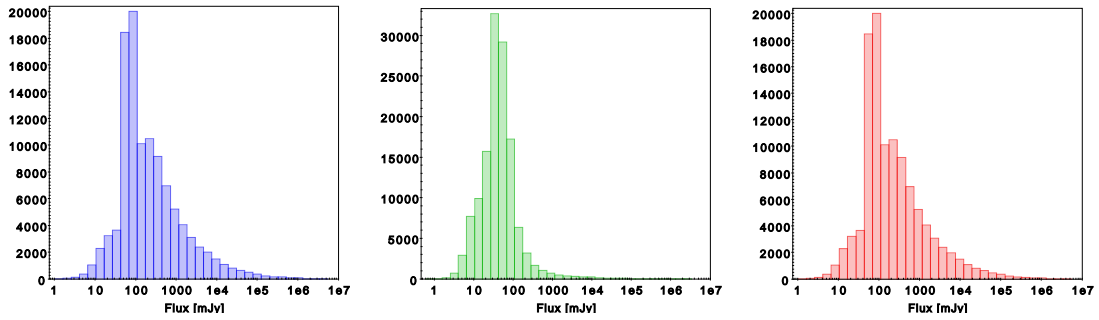


Figure 18: Flux distribution in the blue, green and red bands, indicated with the corresponding colour from left to right.

The flux distribution of our Catalogue objects is shown on Figure 18. In all bands the flux distribution has a peak at  $\sim 90$  mJy. The blue and the red bands have an excess of bright sources compared to the green band. This is due to the fact that observations using the green/red filter combinations were performed in significantly smaller numbers close to the Galactic Plane, where sources have to be bright to be detected and reliably extracted against the bright background.

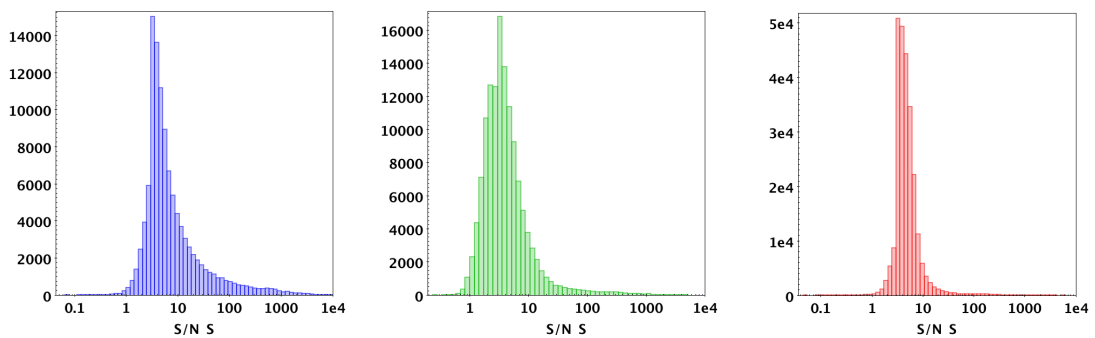


Figure 19:  $S/N_S$  distribution in the blue, green and red bands, from left to right.

The distribution of the  $S/N_S$  values (i.e. the signal-to-noise ratios estimated from the sim-

ulations, calculated as a function of the  $N_S$  and the source flux) have a peak between 3 and 5. Apparently the  $S/N_S$  values in the blue band are more distributed towards higher values than in the green band. The red band  $S/N_S$  values have the narrowest distribution, see Figure 19.

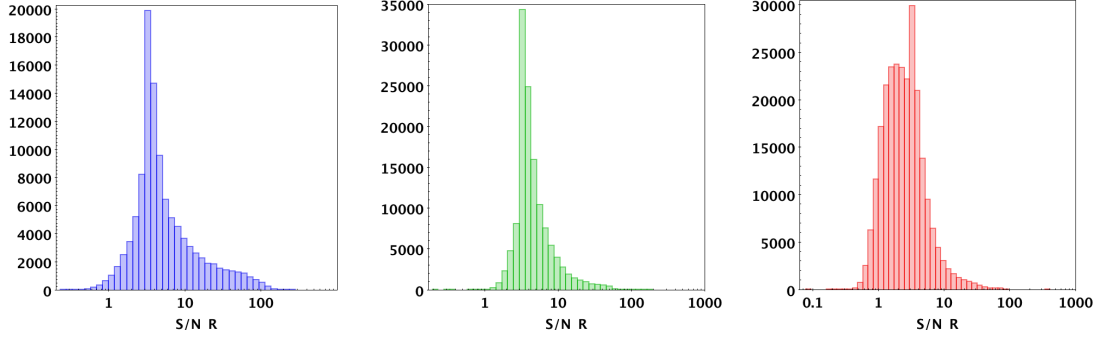


Figure 20:  $S/N_R$  distribution in the blue, green and red bands, from left to right.

Figure 20 shows the distribution of the background RMS based  $S/N_R$  values. It shows similarities with the  $S/N_S$  distribution as values in the blue band occupy a larger interval, and have an excess towards high values, while in the green and the red bands the  $S/N_R$  distribution is more concentrated.

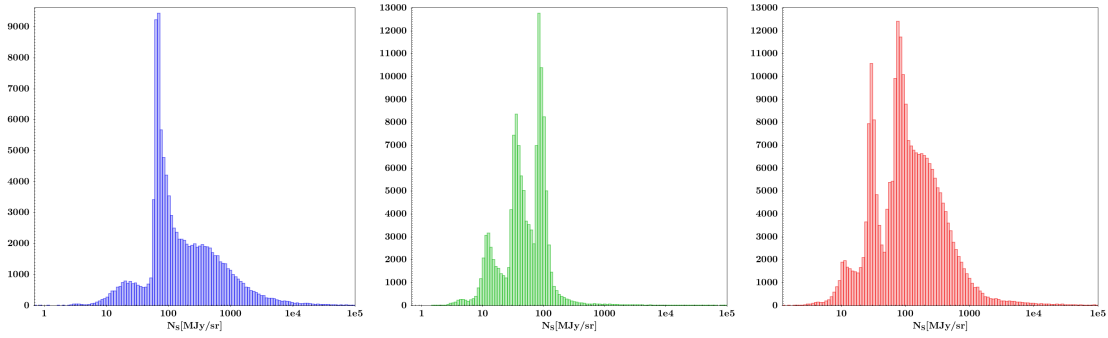


Figure 21: Distribution, from left to right, of the  $N_S$  values in the blue, green and red bands.

The distribution of the  $N_S$  (structure noise) values is plotted on Figure 21. In the blue band the distribution peaks at  $\sim 20$ , 90 and 400  $\text{MJy sr}^{-1}$ . The same behaviour is present in the green and red bands. These values identify clearly regions of different complexity. Figure 22 shows the  $N_S$  values projected in the sky in Galactic coordinate system. The high values are concentrated along the Galactic plane and around the Galactic Centre.

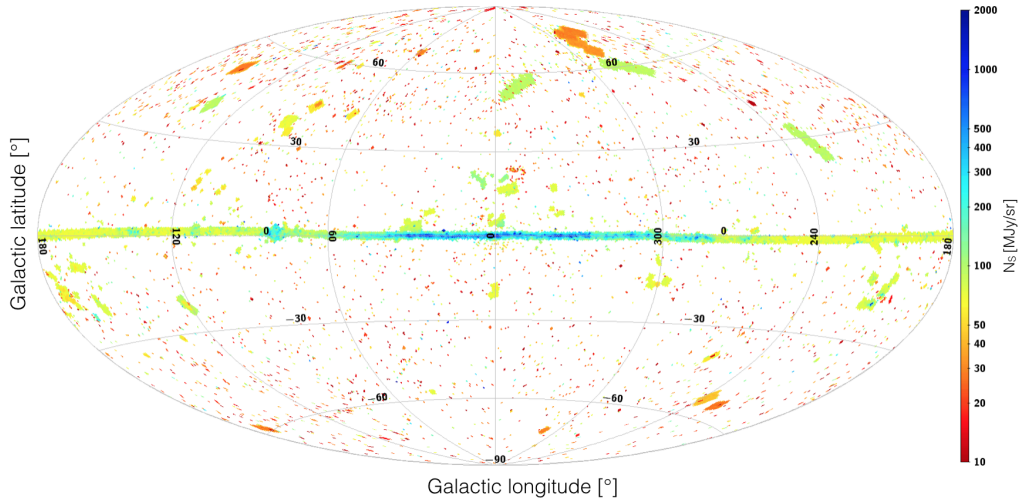


Figure 22: Distribution of the red band  $N_S$  values on the sky in Galactic coordinates, using an Aitoff projection. Colours from red to blue indicate the increasing structure noise.

We have summarised the distributions of the above quantities in Table 6.

	blue			green			red		
	5%	Median	95%	5%	Median	95%	5%	Median	95%
Flux [mJy]	21.2	138.8	8057.1	7.9	40.7	231.7	43.4	271.5	4393.4
$S/N_S$	2.1	5.1	127.7	1.4	3.5	19.1	2.6	4.4	11.1
$S/N_R$	1.6	4.1	42.9	2.2	3.9	14.6	0.9	2.4	8.6
$N_S$ [MJy sr <sup>-1</sup> ]	19.8	110.3	1940.3	10.1	50.0	129.5	14.1	98.9	790.8

Table 6: 5%, 50% and 95% quantiles of some quantities characterising our Catalogue.

## 4.2 HPPSC columns

### 4.2.1 Name (*name*)

The object name is an identifier consisting of two parts, acronym and sequence, and is formatted as 'HPPSCxxxA\_Jhhmmss.s±ddmmss'. 'HPPSC' stands for *Herschel* PACS Point Source Catalogue, xxx is the central wavelength of the band in which the source was observed, identifying the wavelength in  $\mu\text{m}$ , and the letter A indicates the first version of the catalogue. The second part has the format 'Jhhmmss.s±ddmmss' derived from the source coordinates with 'J' indicating a J2000 reference system, 'hhmmss.s' is the Right Ascension in hours, minutes and seconds accuracy of one digit after the decimal, and '+ddmmss' is the Declination in signed degrees, minutes and seconds to an accuracy of one digit after the decimal.

### 4.2.2 Band (*band*)

This is one of the strings "blue", "green", "red" identifying the PACS detector array names for the filter bands centred at 70, 100, 160  $\mu\text{m}$ , respectively. The wavelength is also specified in the name of the object.

### 4.2.3 R.A./Dec (*ra/dec*)

Right Ascension/ Declination coordinate in a J2000.0 reference frame. These are the peak position of the SUSSEXtractor detection. The coordinates are degrees written in double precision format.

### 4.2.4 $RA_{err}/Dec_{err}$ (*raerr/decerr*)

Positional uncertainties of each object as provided by SUSSEXtractor. Units are degrees, the data type is double precision.

#### 4.2.5 Flux (*flux*)

This is the flux as provided by the AnnularSkyAperturePhotometry task. For objects that were identified several times across different maps, this flux corresponds to the extraction with the best average S/N, where the  $N_S$  based S/N ( $S/N_S$ ) and the background RMS based S/N ( $S/N_R$ ) values are taken into account. For the blue and green bands the aperture radius was  $6''$ , for the red band  $12''$  (see Section 3.3 for more details). For all cases the raw flux density values were corrected for the aperture size, based on Lutz (2015). Colour corrections were not applied. The units are mJy, written as double precision.

#### 4.2.6 $S/N_S$ (*snr*)

Estimated S/N based on a statistical approach, using the  $S/N_S$  calibration surfaces. It is calculated as a function of the source flux and the  $N_S$  value. The method is described in details in Section 3.4.1. The data type is double precision.

#### 4.2.7 $N_S$ based noise (*snrnoise*)

It is calculated from the source flux and the estimated  $S/N_S$ , by dividing the flux by the  $S/N_S$ . This noise includes both the sky confusion and the instrument noise.

#### 4.2.8 $S/N_R$ (*stn*)

Calculated as the ratio of the source flux and the RMS. Data type is double precision.

#### 4.2.9 RMS (*rms*)

RMS noise as measured in the blank sky apertures placed around the source (see Section 3.4.2). Double precision. Units are mJy.

#### 4.2.10 $N_S$ (*strn*)

Structure noise value obtained from the structure noise maps as described in Section 3.4.1. Double precision, in units of  $\text{MJy sr}^{-1}$ .

#### 4.2.11 Flux ratio (*fratio*)

The ratio of the flux density values measured in the 6th and the 1st aperture, corresponding to  $4''$  and  $9''$  in the blue and green bands, and  $9''$  and  $14''$  in the red band.

#### 4.2.12 $\text{FWHM}_X/\text{FWHM}_Y$ *fwhmxfit/fwhmyfit*

This column contains the FWHM values measured by the 2D Gaussian model in X and Y directions, fitted by the sourceFitter() task. If the elongated Gaussian fit was possible then both  $\text{FWHM}_X$  and  $\text{FWHM}_Y$  are listed. If only a circular Gaussian could be fitted then  $\text{FWHM}_Y$  is equal to -99.9. It is written in double precision format, in units of second of arc.

#### 4.2.13 ELONGATION FLAG (*flag\_elong*)

If an elliptical Gaussian model was fitted in the source, this column contains the ratio of the  $\text{FWHM}_Y$  and  $\text{FWHM}_X$ . A value of 0 means that the source is circular, 0.5 means that  $\text{FWHM}_Y = 0.5 \times \text{FWHM}_X$ . If  $\text{FWHM}_Y = -99.9$  (circular detection),  $\text{FLAG\_ELONG} = 0$ .  $\text{FLAG\_ELONG}$  is also set to 0 if the source eccentricity  $(\text{FWHM}_X - \text{FWHM}_Y)/\text{FWHM}_X$  defined in a similar way with the FWHM values of the PACS PSF reported in Table 1, is smaller than the PSF eccentricity. Double precision format.



#### 4.2.14 EDGE FLAG (*flag\_edge*)

The position of a source with respect to the edges of the image affects the quality of the photometry, as the extraction procedure needs a large enough area to compute properly both the flux density and the background to be subtracted. Also, due to the scan speed inhomogeneities during the turnarounds, the map edges may have different noise properties than the rest of the image.

The edge flag for each source is obtained using a Jython procedure that defines a subimage from the corresponding ROI map using the background external aperture (35'') centred on the source. We count the total number of pixels and the number of good values (1) in this subimage, obtaining the ratio between the two quantities. The conditions to set the edge flag are:

- The external radius is off the map

- The ratio is less than 0.9

The map edge flag is set when the source shows  $\geq 10\%$  of pixels within the outer radius of the background annulus, with 0 value in the ROI maps.

The data type is Boolean.

#### 4.2.15 BLEND FLAG (*flag\_blend*)

Objects with distance within the PSF FWHM have the same value ( $>0$ ) of FLAG\_BLEND and they can be considered as blended sources. FLAG\_BLEND = 0 means that the object is not blended with other objects. The PACS PSFs depend on the observing mode and the adopted values are the ones reported in Table 1. The absolute pointing error for SPG13 (1.2'') is adopted as uncertainty.

#### 4.2.16 WARM ATTITUDE (*warmat*)

Flag indicating that the observation was affected by "warm" attitude. The term refers to spacecraft attitudes in which the star-tracker support structure is subject to thermal distortions as was not completely shaded from sunlight. This has an impact on the pointing accuracy. Data type is Boolean.

#### 4.2.17 OBSERVATION ID (*obsid*)

The term is the observation identifier for the *Herschel* observations. It is a 10 digit integer, starting always with 1342. The column lists the first OBSID from the list that built the L2.5/L3 map. It corresponds to the meta keyword *obsid001*.

#### 4.2.18 SOLAR SYSTEM MAP FLAG (*ssomapflag*)

Maps of solar system objects have typically been observed in tracking mode, but were reprocessed in the 'rest' sky frame for the HPPSC, and may be contaminated by the moving targets. The affected OBSIDs are marked with an 'ssomap' flag.

Column	Format	Unit	Description
name	String		IAU compatible object identifier
band	String		PACS detector array blue=70 $\mu$ m green=100 $\mu$ m red=160 $\mu$ m
RA	Double Precision	deg	Right Ascension (J2000)
Dec	Double Precision	deg	Declination (J2000)
raerr	Double Precision	deg	Positional uncertainty in RA
decerr	Double Precision	deg	Positional uncertainty in Dec
flux	Double Precision	mJy	non-color corrected flux density
snr	Double Precision		Structure noise based signal-to-noise ratio
snrnoise	Double Precision	mJy	Noise derived from S/N <sub>S</sub>
stn	Double Precision		Background RMS based signal-to-noise
rms	Double Precision	mJy	Background RMS
strn	Double Precision	MJy sr <sup>-1</sup>	Structure noise
fratio	Double Precision		Flux ratio measured in apertures 6 and 1
fwhmxfit	Double Precision	asec	FWHM of the fitted Gaussian along the major axis
fwhmyfit	Double Precision	asec	FWHM of the fitted Gaussian along the minor axis
flag_elong	Double Precision		ratio of the FWHM <sub>Y</sub> and FWHM <sub>X</sub>
flag_edge	Boolean		T = at least one contributing source shows $\geq 10\%$ of map pixels as NaN within the outer background annulus
flag_blend	Integer		Sources with the same numbers are blended
warmat	Boolean		T=the observation was affected by the "warm" attitude
obsid	Integer		Observation identifier
ssomapflag	Boolean		T=SSO map reprocessed in the <i>rest</i> sky frame

Table 7: Columns in the HPPSC tables. Column names, format, unit and a short description are included.

### 4.3 Additional material

#### 4.3.1 Extended Source List

The *Herschel*/PACS Extended Source List (HESL) contains slightly extended objects, extracted from database source table. Quality criteria are similar to those for point sources with two exceptions. We do not impose any condition on the ratio of flux values measured in different apertures, and also the FWHM of the extended source must satisfy:

$$2 \times PSF < FWHM < 5 \times PSF \quad (5)$$

The detailed workflow of object consolidation, including the HPESL compilation is presented in Figure 14.

The columns of the Extended Source Table are similar to the ones derived for HPPSC, with a few exceptions:

- The column NAME is similar to HPPSC name, but the acronyms for extended sources is HPESLxxxA. HPESL means *Herschel*/PACS Extended Source List, xxx the wavelength in units of  $\mu$ m, and the letter A indicates the version of the catalogue - currently A, indicating the first version.
- The column *flux* contains the flux value measured in aperture 11 (18'' in the blue/green, 22'' in the red band). It is important to note, that unlike for point sources, for extended sources the aperture correction was **not** applied.
- The S/N<sub>S</sub> and S/N<sub>R</sub> values were calibrated for point sources, but were used as a quality indicator, so they were kept in the HPESL. However, the noise values derived for them are not valid for extended sources, so they were removed.
- The WARMAT and FLAG\_BLEND flags are also not included.

**We want to emphasise that our pipeline was tuned for sources that have a shape and flux distribution similar to the instrument PSF. The flux density values listed in the Source List are only approximations. We recommend strongly at users retrieve**

the observations from the HSA and extract fluxes and source parameters using their a preferred method. Our values are not suitable for use in scientific publications.

### 4.3.2 Rejected Source List

This list contains all sources that are neither included in the HPPSC nor in the HPESL, because they did not verify our quality criteria (see Section 3.4), and were thus rejected. The columns of the HPRSL are Source ID, RA, Dec,  $RA_{err}$ ,  $Dec_{err}$ , OBSID and band. **Intentionally, other data are not provided. The detections listed here are completely unreliable and any further investigation of them has to be done by the user.**

### 4.3.3 Observation Table

The source extraction technique for the generation of the HPPSC was applied to the map repository as described in the Section 3.1. The *Observation Table*, available for download, reports the main parameters that characterise the maps on which the source extraction was performed. Most of them correspond to Meta-keywords of the Observation Context. The *level* parameter states which map level was adopted for each map and the *nobsids* parameter specifies how many observations (obsID) contribute to the generation of the map. The *obsid* parameter always refers to the first observation identifier in the Observation Context (obsid001). For every *obsid* value, there are two entries in the table, corresponding to the two filters used.

Parameter	Meta-keyword	Description
obsid	obsid001	Observation identifier
startod	odNumber	Operational day at the beginning of the first observation
targetname	object	target name, as given by the observer
proposal	proposal	proposal name
obsmode	obsMode	Parallel Mode/Scan map/Point-source photometry <sup>1</sup>
ra	ra	RA of the map centre
dec	dec	Dec of the map centre
starttime	startTime	Start date/time of the first observation <sup>2</sup>
endtime	endTime	End date/time of the last observation <sup>2</sup>
aorlabel	aorLabel	AOR label as centred in HSPOT
nrep	repFactor	Repetition factor <sup>3</sup> . 0 value for Parallel Mode
mapx	naxis1 from WCS	X map dimension in pixels
mapy	naxis2 from WCS	Y map dimension in pixels
mapscanspeed	mapScanVelocity	10, 20 or 60 $''s^{-1}$
band	blue	green, blue or red filter <sup>4</sup>
level	-	SPG level
nobsids	-	number of obsids
swversion	creator	SPG version

Table 8: Description of the columns (Parameter) of the *Observation Table* and of the Observational Context meta-keywords to which they relate

<sup>1</sup> *Point-source photometry* observations are in chop/nod mode and they are discarded for the catalogue generation.

<sup>2</sup> StartTime and endTime are in expressed TAI by using the format YYYY-MM-DDTHH:MM:SS.

<sup>3</sup> The repFactor keyword doesn't exist for *Parallel Mode* and nrep is set to 0.

<sup>4</sup> PACS is a dual-band photometer: an image is always acquired by the Red detector. The *blue* meta-keyword value specifies which filter is applied for the Blue camera (blue1 = 70  $\mu m$ , blue2 = 100  $\mu m$ )

## 5 Validation

### 5.1 Internal validation – Simulations

Simulations played an essential role in the validation of the HPPSC. One way to test our photometry is to measure the flux of sources that are injected artificially into the observational timelines with well controlled flux values. The completeness can also be characterised by simulations: as the number of injected sources per brightness bin is known we can compare these values with the source counts obtained through our pipeline.

For PACS the PSF shape is significantly different from a simple 2D Gaussian function, especially in the blue band at high scan speed. Unlike for the SPIRE PSC, we had to use actual PSF images and add them to the Level1 TOD. The process of creating the simulations consisted of the following steps:

- **Random position generation:** A number of random positions were generated with the criteria of a minimum separation of  $35''$ . The number of randomly generated sky positions is determined by the extent of the map. In general, for large Parallel Mode observations we were able to generate more positions.
- **PSF back-projection:** in order to simulate the real sources, we used the official Vesta PSFs that are specific to the observing mode, band and scan speed. After the correct PSF was selected, it was rotated according to the scan angle of the observation in which the PSF was injected. Then, the Level1 TOD of the observations was duplicated and emptied. The back-projection from the L25 PSF image to the Level TOD was carried out with a dedicated task in HIPE, called `Map2signalCubeTask()`. We used PSFs that were scaled so that source flux was exactly 1 Jy. As a result of the process we were left with a TOD that was of the same length as the observation, but contained only 1 Jy flux sources at the randomly generated positions (source-only timelines).
- **Map projection:** The source-only timeline has the advantage that it can be scaled, multiplying it by 10 will result in 10 Jy sources. Similarly dividing by 10 results in sources with 100 mJy source flux. After scaling to the desired flux level, we added the PSF data to the actual observational TOD. The combined data was then processed with JScanam SPG13.0.0, so the same algorithm was used as for the vast majority of the maps in our repository.

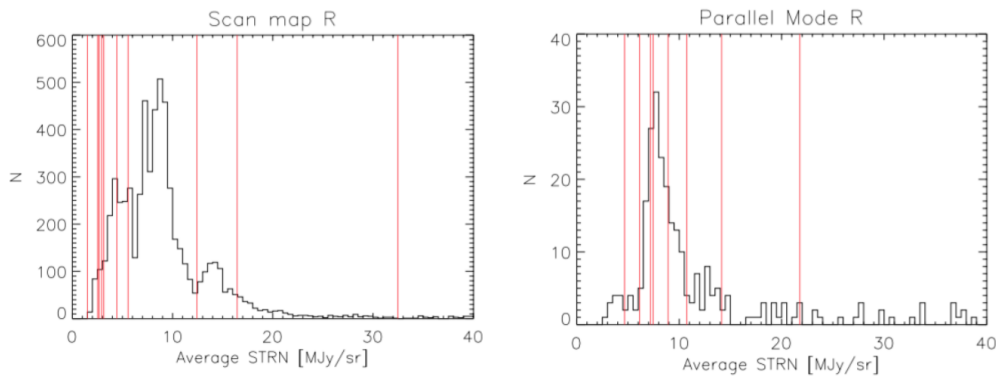


Figure 23: Distribution of the average structure noise values calculated for Scan maps (left) and for Parallel Mode maps in the red band. The structure noise values of the simulated fields are marked with red lines.

Field name	OBSID	RA, Dec [°]	size [']	Mode	Speed [″s <sup>-1</sup> ]	Band	$N_S(\text{red})$ [MJy sr <sup>-1</sup> ]	$N_{inj}$
Atlas	1342189661	212.46,0.55	331×345	parallel	60	BL+R	7.45	300
Field 2_0	1342204104	267.87,-27.05	208×210	parallel	60	BS+R	91.10	200
Field 171_0	1342250342	81.10,36.15	208×210	parallel	60	BS+R	8.92	200
G159.23	1342239263	44.37,19.751	99×103	parallel	20	BL	7.21	200
Globule2	1342188683	187.54,-63.70	80×84	parallel	60	BS+R	14.15	200
GP field-1	1342183070	266.61,-28.93	125×126	parallel	60	BS+R	576.69	200
L1489	1342202088	61.00,26.28	81×81	parallel	60	BS+R	10.74	200
LDN1780	1342224989	235.10,-7.22	153×156	parallel	20	BL+R	6.14	300
M31	1342211604	10.99,41.36	186×224	parallel	20	BS+BL+R	4.67	200
N6357	1342204847	261.71,-34.43	136×139	parallel	20	BS	172.50	300
Rosette	1342186121	98.22,4.27	121×127	parallel	20	BS+R	21.77	200
Abell370	1342223332	39.97,-1.58	20×21	scan	20	BL+R	2.53	20
A851	1342271023	145.75,46.99	20×21	scan	20	BS	2.65	20
Arp 244	1342187836	180.47,-18.87	22×26	scan	20	BS+BL+R	16.45	20
Cas A	1342188205	350.85,58.81	66×68	scan	20	BL+R	12.44	150
Crab	1342204442	83.63,22.01	66×70	scan	20	BS+BL+R	5.56	50
IRAS16132	1342216487	244.25,-50.79	63×64	scan	20	BS+R	82.33	100
IRDC010.70	1342191803	272.46,-19.73	37×40	scan	20	BL+R	95.68	50
M81	1342186085	148.90,69.07	103×103	scan	20	BS+R	2.91	100
NGC1365	1342183021	53.40,-36.14	75×78	scan	20	BL	1.50	100
NGC6946	1342191947	308.72,60.16	47×47	scan	20	BS+BL+R	3.14	30
NGC7026	1342223919	316.58,47.85	24×23	scan	20	BS+R	4.45	30
RCW79	1342188880	205.06,-61.73	88×92	scan	20	BL+R	32.48	100
RCW120	1342216585	258.09,-38.47	88×91	scan	20	BS+R	48.93	150

Table 9: Observations used to create simulations.

### 5.1.1 Completeness

The completeness is calculated as a ratio between the number of sources injected and the number of sources detected by our pipeline. We present it here for two different fields. The Abell 370 field is a deep cosmological observation, mainly free of extended emission, therefore it has a low structure noise. The RCW120 field is a HII region with ongoing star formation, close to the Galactic Plane, and therefore very structured. As one can see from the curves in Figure 24, 90% completeness is reached at  $\sim 20$  mJy, while the same completeness in the Galactic field is reached at  $\sim 1000$  mJy.

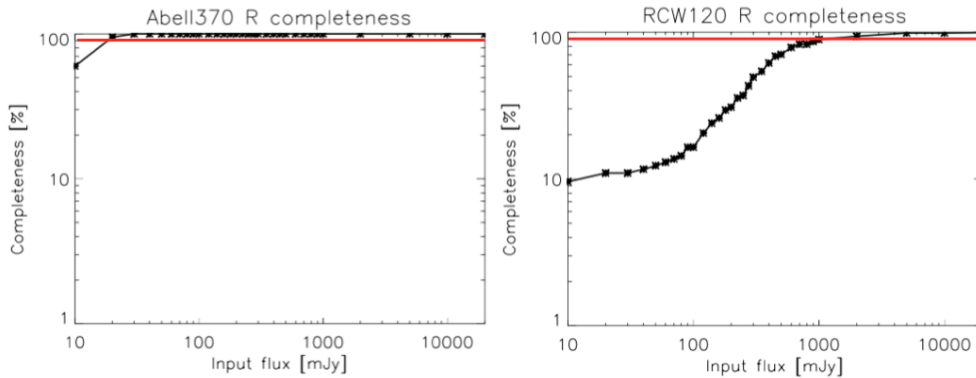


Figure 24: Completeness in two of the simulated fields. Abell370 (left) is a deep cosmological observation. RCW120 (right) is a very complex HII region.

As it is described in the Section 3.2, the structure noise includes noise both from the instrument and the sky confusion. Figure 24 clearly shows that the completeness highly depends on the complexity of the image. To characterise the relation between completeness and structure noise for each injected flux level we created  $N_S$  bins, and counted the number of sources that were injected in the given  $N_S$  interval and the number of sources what we detected. The same criteria that were used to clean the source tables were used here, hence the results reflect the true completeness (see

Section 3.5). A logarithmic scale was used to set size of the  $N_S$  bins. The bins were centred on  $N_S$  values of  $10^x$ , where  $x$  varied between 1 and 3 with an interval of 0.05.

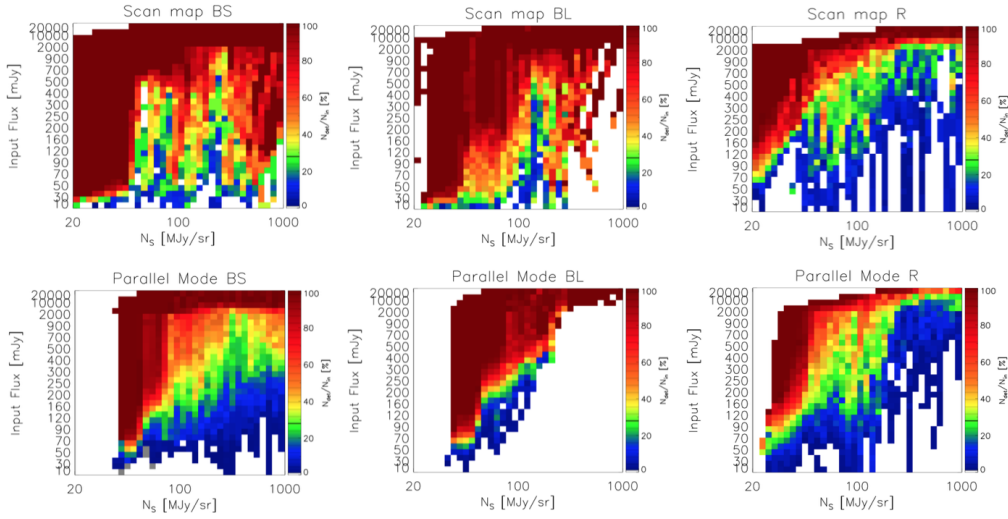


Figure 25: The simulation based completeness in the blue, green and red bands (column 1,2 and 3, respectively) for the Scan map (row 1) and parallel (row 2) observing modes. The colour corresponding to from blue to red means low to high completeness.

As expected, the completeness is greater at low  $N_S$  values. Figure 25 also shows that at high flux levels we are always almost complete, while the faint sources are undetectable on the complex and bright background.

### 5.1.2 Photometric accuracy

The photometric accuracy is characterised as the ratio of the input flux to the measured flux. The photometric error is the standard deviation of the measured flux values at a given injected flux levels. For example if the injected flux is 60 mJy and the measured flux is  $58.74 \pm 9.44$  mJy, the photometric accuracy is  $58.74/60.0 = 0.979$ , the error is 9.44, and the S/N is  $58.74/9.44 = 6.22$ .

As shown in Figure 26, the photometric accuracy depends strongly on the environment. On the left the photometric accuracy was studied in the same Abell370 field used for the completeness study. On the right we used the observation of the RCW120 field. Both datasets were obtained in the red band. The dashed green lines show perfect photometry. As expected, our photometry performs better in a less complex environment. The error bars present the photometric errors, and are also lower in the cosmological observation.

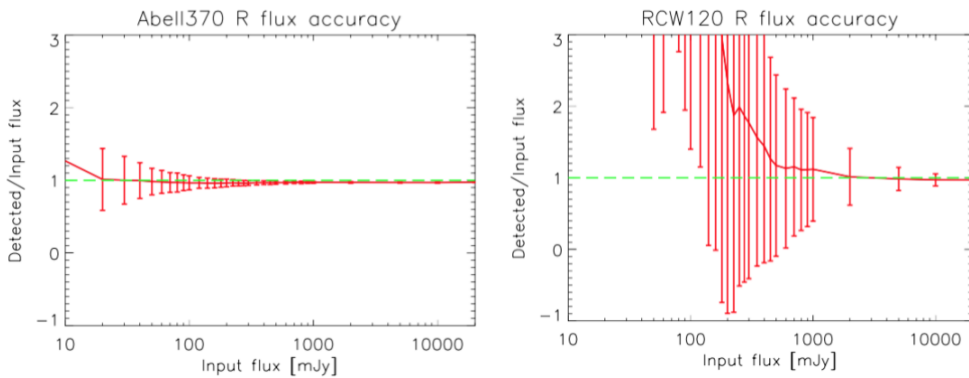


Figure 26: Photometric accuracy in two of the simulated fields. RCW120 (right) is a very complex HII region. Abell370 (left) is a deep cosmological observation.

These tests also led us to the conclusion that our photometry in areas in the Galactic Plane and in star forming regions (i.e. high complexity) is less accurate. The same strategy of  $N_S$  binning was applied here, so we were able to determine the flux accuracy as a function of the  $N_S$  and the input flux. In Figure 27 we show that at high flux levels our photometric performance is very good, with flux ratios close to unity. However, at low flux levels and in environments of high complexity the flux determination is quite unreliable.

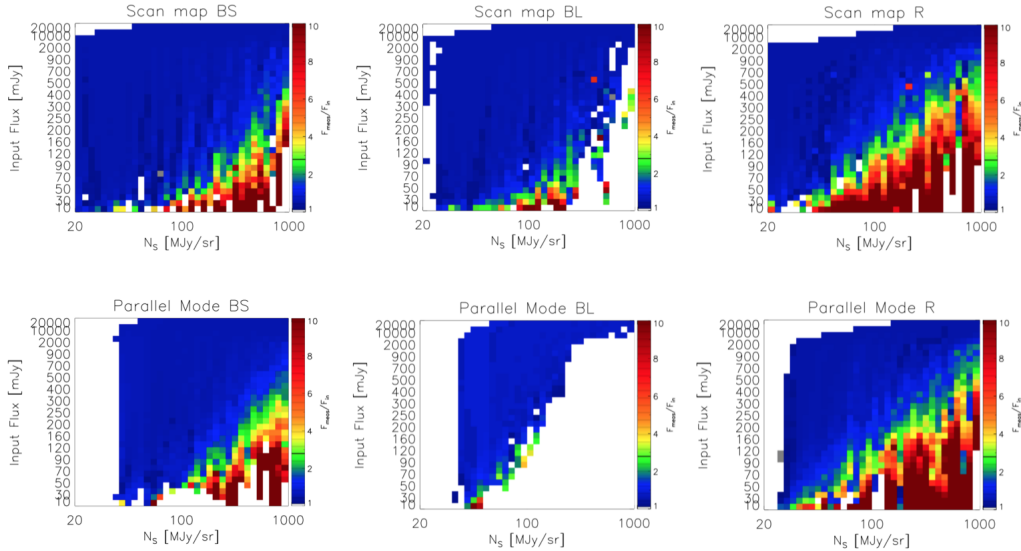


Figure 27: The photometric accuracy in the blue, green and red bands (column 1,2 and 3, respectively) for Scan map (row 1) and Parallel (row 2) observing modes. The colour coding from blue to red means high to low photometric accuracy.

### 5.1.3 Positional accuracy

We used the fitted coordinates of standard stars and compared them with their known sky positions. To characterise the positional accuracy of the catalogue. We considered the same sample for this purpose as was used for the photometric accuracy tests, i.e. objects in Balog et al. (2014) for the bright standard stars and in Klaas et al. (2017) for the faint standard stars. The median positional difference was found to be  $1.5''$  for the blue PACS detector (70 and  $100 \mu\text{m}$  filters) and  $1.7''$  for the red detector ( $160 \mu\text{m}$  filter). This accuracy is comparable with the typical  $\sim 2''$  pointing accuracy of the Herschel Space Observatory (see Sanchez-Portal et al. 2014).

A second test that we performed was to compare the recovered positions of the sources that we injected into the simulations. The difference between the input coordinates and the output coordinates calculated by the SUSSEXtractor algorithm includes the errors from the source back-projection to the observational timeline, the error introduced by the map re-projection and the error introduced by SUSSEXtractor. However, as shown on Figure 28 the peak difference is between  $1.0''$  and  $1.5''$  for all bands and observing modes. This value is still well below the FWHM values of the instrument PSFs.

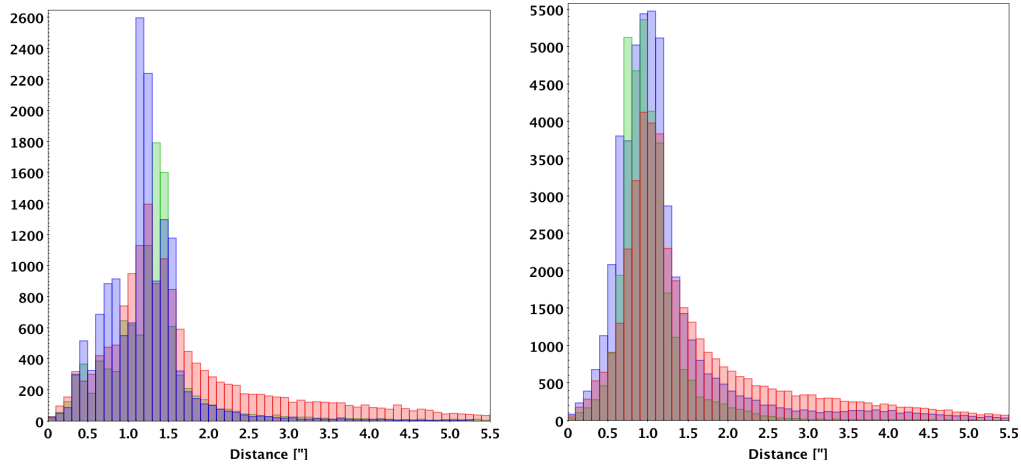


Figure 28: Position uncertainty based on simulations for the Scan map (left) and Parallel Mode (right). The uncertainty is measured as the angular distance between the injected source position and the SUSSEXtractor source position. The blue, green and red bands are indicated with the corresponding colours.

As an additional test, we compared the source positions from the Hi-GAL project with those of the corresponding sources in the HPPSC (see Fig. 29). The distances between the Hi-GAL and the HPPSC positions show similar median values as those obtained in the previous tests.

While the positional accuracy of the HPPSC is in general comparable to the pointing accuracy of the observatory and this positional accuracy is notably better than the spatial resolution achievable due to the finite beams in the PACS bands there are some situations in which the absolute positional accuracy of the telescope may be degraded. This may happen in the case of those measurements which were taken at ‘hot’ solar aspect angles ( $> 110^\circ$ , see [Sanchez-Portal et al. 2014](#), for a detailed discussion). To warn the users of the HPPSC of these conditions an additional flag ‘WARMAT’ was introduced for sources extracted from maps with  $SAA > 110^\circ$ . About 3% of all maps used for HPPSC source extraction may be affected.

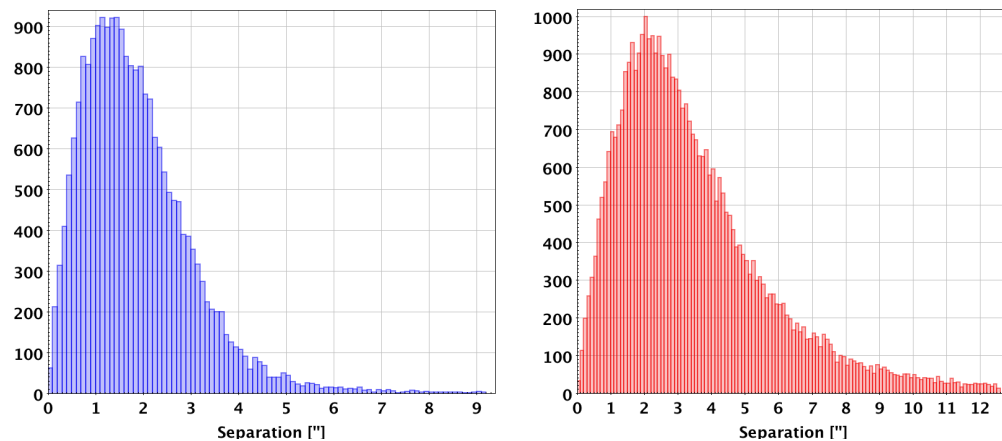


Figure 29: Distribution of the distances between the HiGal and the HPPSC source positions for the blue (left) and red (right) PACS detectors.

## 5.2 External Validation

While our extracted sources show internally consistent properties in the framework that we set for the PACS Point Source Catalogue, it is important to check our results against already existing catalogues of large surveys based on the same data. These calibration samples and large programmes used diverse approaches and data reduction/evaluation techniques – typically significantly different to ours – to reach their goals, therefore comparison with these results gives information about



the robustness of our techniques, and the general quality and reliability of our data evaluation algorithms.

### 5.2.1 Comparison with point source flux calibration

Balog et al. (2014) provided an overview of the PACS photometer flux calibration concept, in particular for the principal observation mode, the scan map. The absolute flux calibration was tied to the photospheric models of five fiducial stellar standards. We selected the same objects from our catalogue and compared our colour corrected flux values to their reported flux density values. Our photometry was carried out using 12'' and 24'' aperture for the blue and red camera and applying aperture correction. For the comparison colour correction was also applied, assuming 5000K black body temperature. The colour correction values based on the PACS report by Müller et al. (2011) was 1.016, 1.033 and 1.074 for the blue, green and red bands. Table 10 lists the name of the calibrators, the model flux, the average measured flux and the average ratio of the model and catalogue flux. The left panels of Figures 30, 31 and 32 show the catalogue flux as a function of the model flux. The 1:1 line is shown in green. Taking into account the 5% error of the model, our flux values agree very well with the predicted values. The right panels show the flux ratios. These figures highlight the fact that in most of the cases our flux values are within 5% of the model flux.

Name	band	$F_{model}$ [mJy]	$F_{catalogue}$ [mJy]	$F_{model}/F_{catalogue}$
HD6860	blue	5594	$5718.506 \pm 45.171$	$0.978 \pm 0.008$
	green	2737	$2803.471 \pm 11.109$	$0.976 \pm 0.004$
	red	1062	$1070.047 \pm 9.263$	$0.993 \pm 0.009$
HD18884	blue	4889	$4969.804 \pm 28.160$	$0.984 \pm 0.006$
	green	2393	$2431.604 \pm 7.849$	$0.984 \pm 0.003$
	red	928	$933.292 \pm 5.058$	$0.9943 \pm 0.005$
HD29139	blue	14131	$13897.883 \pm 134.419$	$1.017 \pm 0.010$
	green	6909	$6771.814 \pm 34.735$	$1.020 \pm 0.005$
	red	2677	$2595.375 \pm 12.287$	$1.031 \pm 0.005$
HD124897	blue	15434	$15099.168 \pm 348.331$	$1.023 \pm 0.023$
	green	7509	$7381.912 \pm 105.521$	$1.018 \pm 0.014$
	red	2891	$2914.236 \pm 24.985$	$0.992 \pm 0.008$
HD164058	blue	3283	$3247.424 \pm 22.110$	$1.011 \pm 0.007$
	green	1604	$1568.252 \pm 9.618$	$1.023 \pm 0.006$
	red	621	$619.558 \pm 12.230$	$1.003 \pm 0.020$

Table 10: Comparison of the flux predicted by photospheric models and the flux value measured by our pipeline. The names of the calibrators are listed in the first column. The second column gives the band. The predicted flux value is shown in the third column. The average measured flux and the corresponding  $1\sigma$  uncertainty are given in the fourth column. Column five shows the flux ratio.

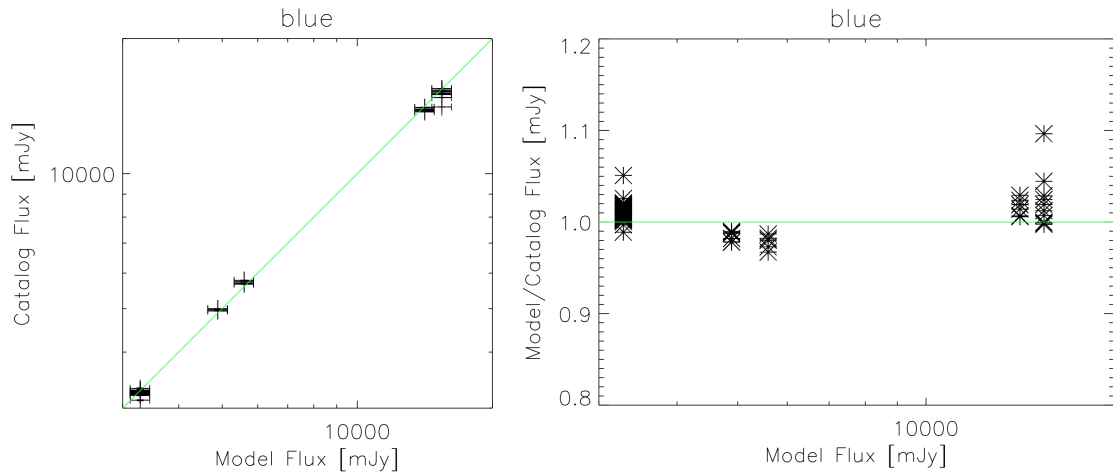


Figure 30: Left: Catalogue flux density values as a function of the flux density predicted by photospheric models in the blue ( $70\mu\text{m}$ ) band. The solid green line presents the 1:1 line. Right: Model flux to catalogue flux ratio as a function of the model flux in the blue band. The green horizontal line presents ratio of one.

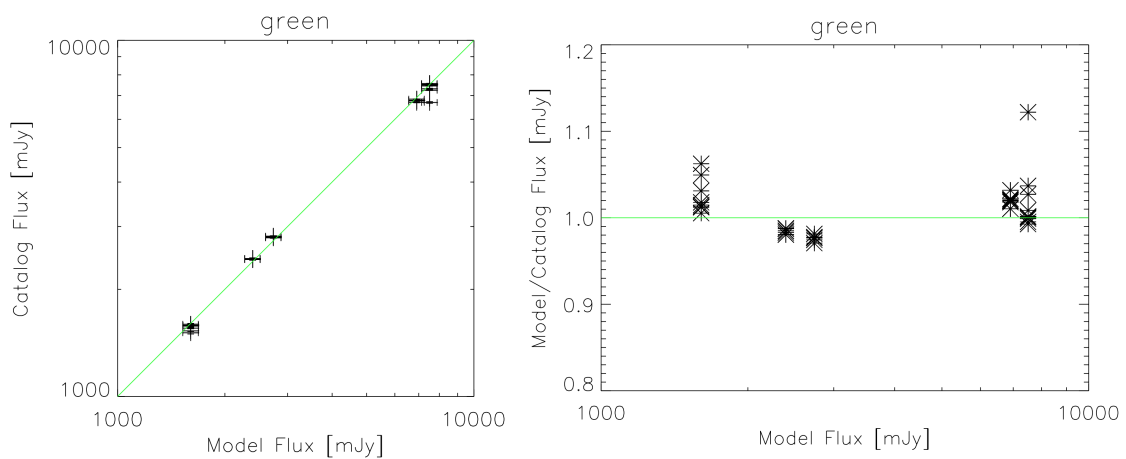


Figure 31: Left: Catalogue flux density values as a function of the flux density predicted by photospheric models in the green ( $100\mu\text{m}$ ) band. The solid green line presents the 1:1 line. Right: Model flux over the catalogue flux as a function of the model flux in the green band. The green horizontal line presents ratio of one.

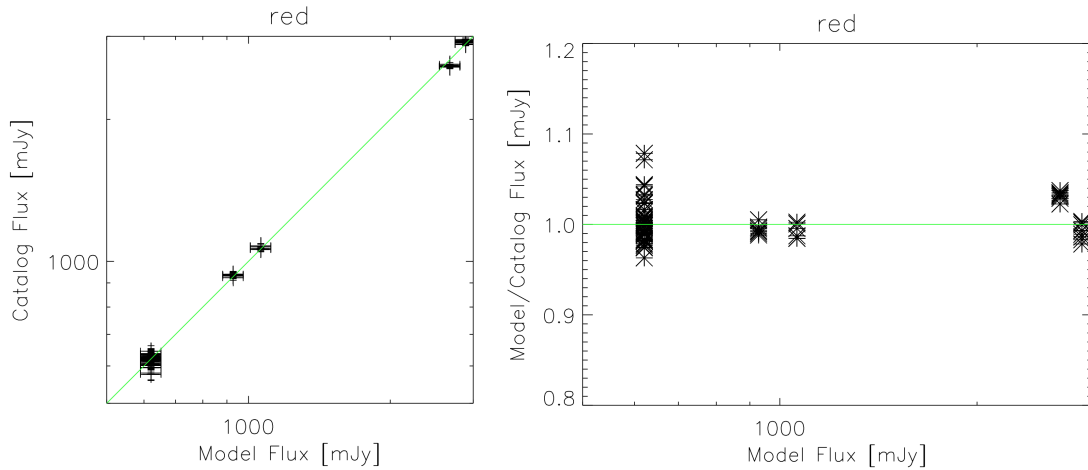


Figure 32: Left: Catalogue flux density values as a function of the flux density predicted by photospheric models in the red ( $160\mu\text{m}$ ) band. The solid green line presents the 1:1 line. Right: Model flux to catalogue flux ratio as a function of the model flux in the red band. The green horizontal line presents ratio of one.

We conclude from the comparison that our flux density values are in very good agreement with those predicted by the photospheric models. The average flux ratios are  $1.0026 \pm 0.01$ ,  $1.0042 \pm 0.006$  and  $1.0026 \pm 0.04$  in the blue, green and red bands, respectively. These are also in almost perfect agreement with the values reported by Balog et al. (2014).

### 5.2.2 Faint star comparison

Klaas et al. (in prep.) determined flux densities and photometric accuracy for a set of 17 stars that range in flux from intermediate brightness ( $\sim 2.5$  Jy) to faint ( $\sim 5$  mJy). Their data processing followed that of Balog et al. (2014). It should be noted here that Klaas et al. performed the photometry on Level 2 high-pass filtered maps, while in the preparation of the HPPSC we used maps generated by the JScanam method. Therefore, the comparison of the faint standard star fluxes from the two different sources tests also the compatibility of the two mapping algorithms.

Name	band	$F_{model}$ [mJy]	$F_{catalogue}$ [mJy]	$F_{model}/F_{catalogue}$	$F_{Klaas}$ [mJy]	$F_{catalogue}/F_{Klaas}$
HD62509	blue	2457	2702.094±3.835	0.909±0.001	2649.400	1.020± 0.001
	green	1190	1334.990±3.963	0.891±0.003	1287.100	0.835± 0.004
	red	456	526.971±6.069	0.865±0.010	497.400	1.059± 0.012
HD12929	blue	1707	1704.437± 13.264	1.002± 0.008	1668.300	1.022± 0.008
	green	831	851.174± 7.679	0.977± 0.009	831.600	1.024± 0.009
	red	321	343.872± 2.094	0.934± 0.006	336.000	1.023± 0.006
HD32887	blue	1182	1194.326± 6.118	0.990± 0.005	1157.100	1.032± 0.005
	green	576	592.220± 2.975	0.973± 0.005	568.800	1.041± 0.005
	red	222	234.146± 4.685	0.951± 0.019	217.500	1.077± 0.022
HD198542	blue	857	868.183± 3.489	0.988± 0.004	839.000	1.035± 0.004
	green	418	428.658± 1.955	0.975± 0.004	414.300	1.035± 0.005
	red	161	170.364± 6.439	0.949± 0.037	168.600	1.010± 0.038
HD148387	blue	479	519.299± -	0.923± -	517.800	1.003±-
	green	232	256.012± -	0.909± -	237.900	1.076±-
	red	89	86.323± -	1.047± -	116.400	0.742±-
HD180711	blue	428	443.526± 3.371	0.968± 0.007	433.600	1.023± 0.008
	green	207	214.613± 3.839	0.972± 0.016	214.000	1.003± 0.018
	red	79	81.978± 4.573	1.007± 0.047	89.000	0.921± 0.051
HD139669	blue	286	291.876± -	0.981± -	278.700	1.047±
	green	140	-±-	-±-		
	red	53	55.507± -	0.971± -	72.800	0.762±
HD41047	blue	196	-±-	-±-		
	green	95	101.709± -	0.938± -	99.400	1.023± 0.000
	red	36	33.121± -	1.114± -	52.800	0.627± 0.000
HD170693	blue	153	150.013± -	1.025± -	146.500	1.024±-
	green	75	75.222± -	1.001± -75.100	1.002±-	
	red	29	32.628± -	0.902± -31.200	1.046±-	
HD138265	blue	115	114.789± -	1.010± -	109.200	1.051± -
	green	56	56.997± -	0.997± -	57.200	0.996±-
	red	22	29.010± -	0.766± -	31.200	0.930±-
HD159330	blue	64	64.754± -	0.991± -	60.700	1.067±
	green	31	33.194± -	0.951± -32.900	1.009±-	
	red	12	17.104± -	0.719± -	-	-±-
HD152222	blue	39	40.308± -	0.977± -	35.100	1.148±-
	green	19	21.103± -	0.915± -22.400	0.942±-	
	red	7	11.406± 4.415	0.726± 0.388	6.200	1.840± 0.712
HD39608	blue	30	32.308± -	0.956± -	29.800	1.084± -
	green	15	-±-	-±-		
	red	5	17.459± -	0.338± -	18.200	0.959± -
HD15008	blue	22	24.930± -	0.919± -	7.700	3.238±-
	green	11	12.118± -	0.924± -	-	-±-
	red	4	-±-	-±-		

Table 11: Comparison of the flux predicted by photospheric models and the flux value measured by our pipeline. The names of the calibrators are listed in the first column. The second column shows the band. The predicted flux value is listed in the third column. The average measured flux and the corresponding  $1\sigma$  uncertainty are in the fourth column. Column five shows the flux ratio.

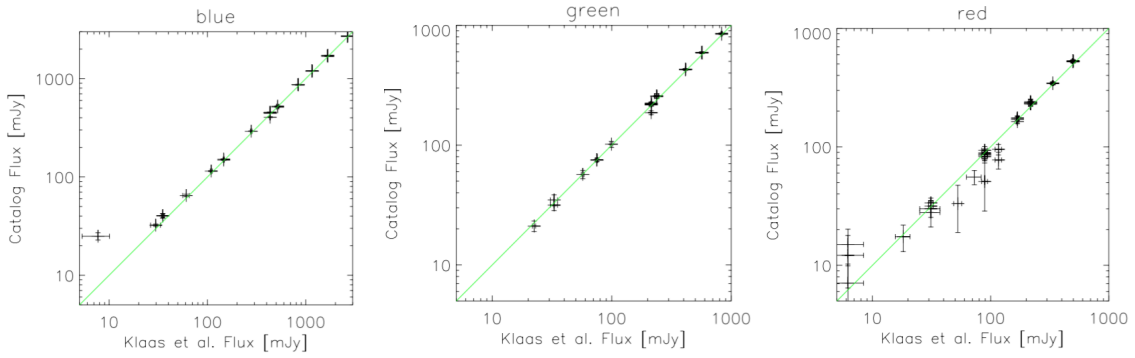


Figure 33: From left to right, respectively, HPPSC flux densities for faint stars as a function of the values reported by Klaas et al. in the blue, green and red bands.

### 5.2.3 PEP-COSMOS field

The COSMOS field (RA:150.11917 Dec:2.20583) has been observed by [Lutz et al \(2011\)](#) for the Guaranteed Time Key Programme PACS Evolutionary Probe (PEP). COSMOS is the largest field observed by this programme (2 square degrees, 200h). We compared the PEP blind catalogues with the HPPSC catalogue. PEP sources have been extracted via the Starfinder code ([Diolaiti et al 2000](#)). Only sources with a S/N higher than three have been included in DR1 PEP catalogues.

We have carried out the cross-match applying TOPCAT between our Point Source Catalogue and the PEP-COSMOS list, using a search radius of 4 and 6'' for the 100 and 160 $\mu$ m catalogues, respectively (these values are near to the corresponding HWHM in each band).

The HPPSC catalogue has 5607 and 2663 sources in the COSMOS region. It must be noted that there are other observations, corresponding to the [CANDELS](#) proposal (not included in PEP catalogues), that contribute to populate our catalogue. In this comparison, we do not use the slightly extended sources in the field (i.e., the HPESL) hence the percentages of cross-matches will be smaller than true values. The PEP catalogue has 7443 objects at 100 $\mu$ m and 7047 sources at 160 $\mu$ m in the COSMOS field. The number of matching sources is 3866 and 2335 respectively for the green and the red bands. Hence 65%-83% of HPPSC sources match PEP objects, while 52%-33% of PEP sources found in the HPPSC catalogue.

- Separation

The position of matched sources is consistent, with most of them agreeing to a few arcsec (see [Figure 34](#)). Mean values for the separation between objects are 1.9 and 2.3'' (for 100 and 160 $\mu$ m, respectively), values much lower than the selected search radius in each band.

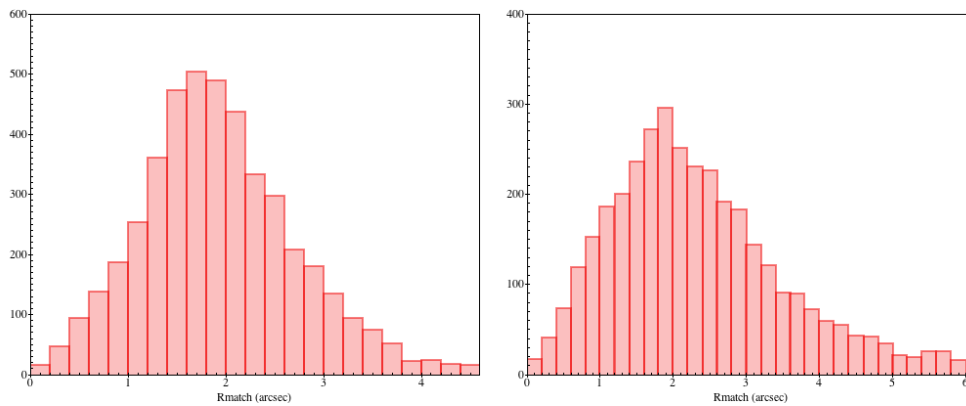


Figure 34: From left to right: 100 $\mu$ m and 160 $\mu$ m separation between HPPSC and PEP-COSMOS sources.

- Flux ratios

We have compared the starfinder PEP-COSMOS fluxes to our aperture fluxes in the matched catalogue. The results are shown in Figures 35 and 36. The fluxes are in good agreement, with resulting mean differences between fluxes of around 3% for the green band and 8% for the red band.

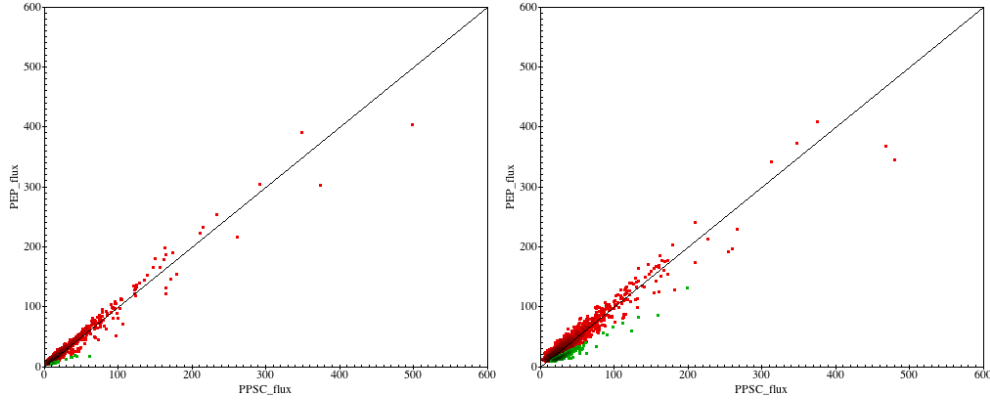


Figure 35: From left to right:  $100\mu\text{m}$  and  $160\mu\text{m}$  HPPSC vs PEP-COSMOS fluxes. All matches (red points) and outliers (green points) are plotted.

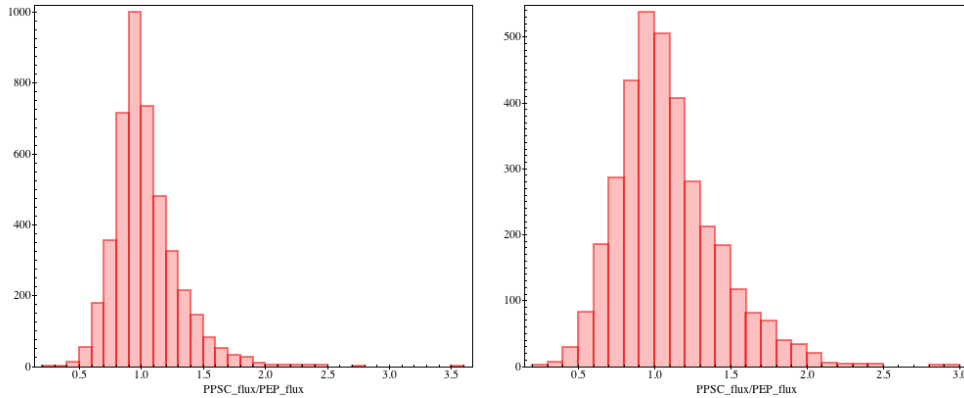


Figure 36: From left to right:  $100\mu\text{m}$  and  $160\mu\text{m}$  HPPSC/PEP-COSMOS ratio fluxes.

- Outliers

As shown Figure 35, there are objects that can be considered as outliers in both bands. There are 17 objects in the green band (see Figure 37) for which PEP fluxes are systematically lower than HPPSC fluxes. Of these, 6 HPPSC sources have a double match in the PEP-COSMOS catalogues, and the sum of the values from the two PEP sources in the match radius is similar to the HPPSC flux. The other 12 sources have a unique match. For these objects, the differences between fluxes can be explained because they are weak sources (PEP fluxes are  $\sim 5\text{mJy}$ ), the distance between peaks is larger than the median value of  $r_{\text{match}}$  and by the use of two distinct procedures (aperture and starfinder) to obtain the fluxes.

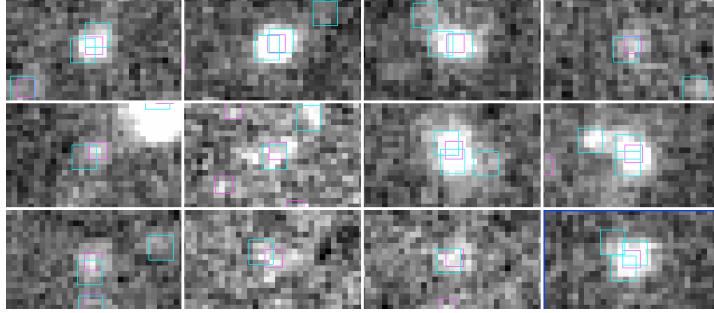


Figure 37: Thumbnails of some outliers found in the HPPSC/PEP-COSMOS match at  $100\mu\text{m}$  (19). Magenta squares: HPPSC sources. Cyan squares: PEP sources.

In the red band, there are 39 outliers. As well as the green band, 9 are not unique detections, and the flux of HPPSC source is similar to the sum of PEP fluxes of the contributing sources. For

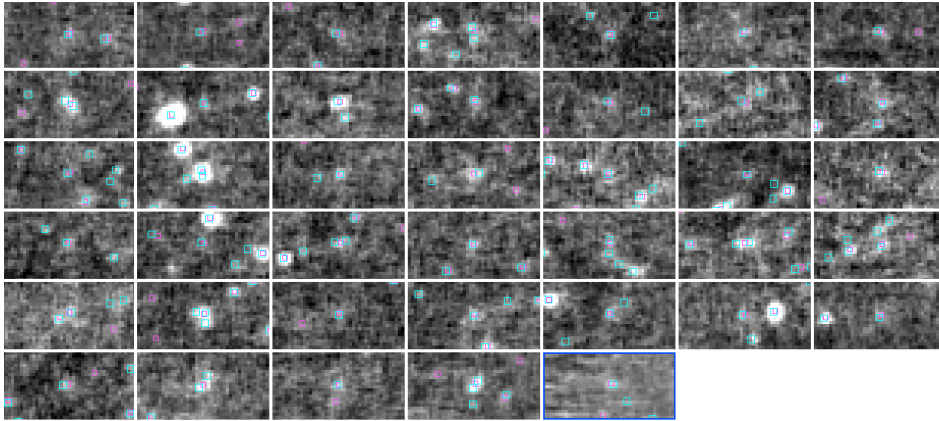


Figure 38: Thumbnails of some outliers found in HPPSC/PEP-COSMOS match at  $160\mu\text{m}$ . (40). Magenta squares: HPPSC sources. Cyan squares: PEP sources.

#### 5.2.4 PEP-GOODS field

GOODS-South (RA: 53.12654 Dec: -27.80467) is the only field observed by PEP at  $70\mu\text{m}$ , covering  $17 \times 11$  arcmin<sup>2</sup>. The number of sources in this field of the PPS catalogue is 199, while the PEP catalogue has 486 sources. We have matched both catalogues with a search radius of  $3''$ . About 83% of HPPSC sources are seen by PEP catalogue. The positions of matches agree well (mean  $2.4''$ ). Fluxes are consistent with a scatter of 10% (see Figure ??).

There is a subgroup of matches that have a flux that is considerably smaller in the PEP catalogue (5 sources). One of them has a double match in PEP catalogue, with a sum of individual fluxes very close to the HPPSC flux. The other four objects are weak (see Figure 39) so the differences can be explained by different method to extract fluxes.

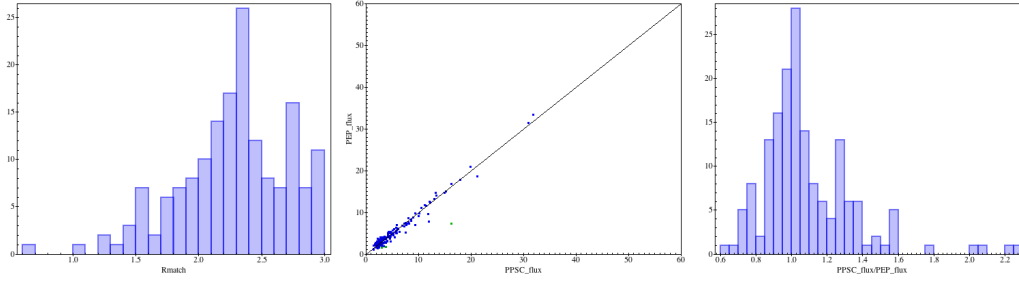


Figure 39: From left to right:  $70\mu\text{m}$  separation between HPPSC and PEP-GOODS-South sources, HPPSC vs. PEP-GOODS-South fluxes and HPPSC/PEP-GOODS-South flux ratio.

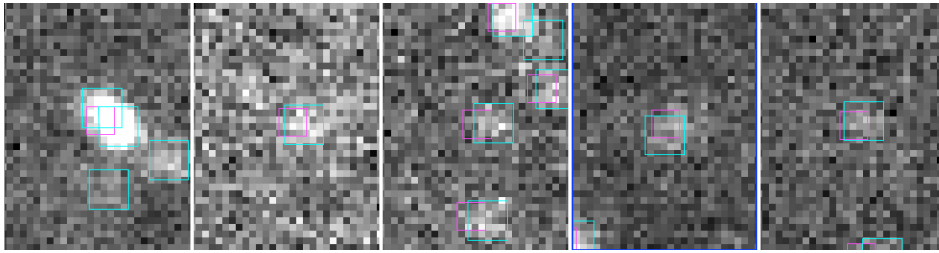


Figure 40: Thumbnails of outliers found in the HPPSC/PEP-GOODS match at  $70\mu\text{m}$ . Magenta squares: HPPSC sources. Cyan squares: PEP-GOODS sources.

## Conclusions

The positions of matches generally agree well and the scatter is consistent with the absolute pointing uncertainty of *Herschel*.

The fluxes of matching sources are generally consistent with a scatter of about 10% or better. However, there are some sources where PEP fluxes are considerably smaller than the SPSC fluxes. These are cases in which the source is multiple and the constituents are closer together than the FWHM at the wavelength and thus are not distinguished by the HSPSC extraction algorithm or weak sources for which the difference between the peaks are larger than the mean difference. We conclude that the agreement between the two catalogues is quite good.

### 5.2.5 Goods-*Herschel* survey

GOODS-*Herschel* [Elbaz et al, 2011](#) is an ESA Open Time Key Project consisting of the deepest *Herschel* observations of the two Great Observatories Origins Deep Survey (GOODS) fields in the Northern and Southern hemispheres. Flux densities and their associated uncertainties were obtained from point source fitting using prior  $24\mu\text{m}$  positions.

For PACS- $100\mu\text{m}$ , they have used MIPS- $24\mu\text{m}$  priors down to the  $3\sigma$  limit, imposing a minimum flux density of  $20\mu\text{Jy}$ . For PACS- $160\mu\text{m}$  and SPIRE- $250\mu\text{m}$ , they restricted the  $24\mu\text{m}$  priors to the  $5\sigma$  ( $30\mu\text{Jy}$ ) limit (reducing the number of priors by about 35%). GH-GOODS sources are flagged with a "clean\_index". This flag marks contamination by nearby sources. Only sources with a clean\_index = 0 have a reliable flux in the GH catalogue.



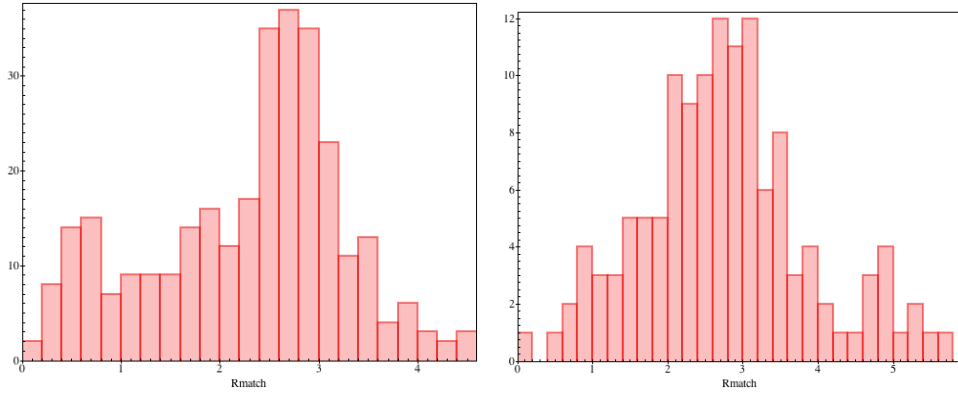


Figure 41: From left to right:  $100\mu\text{m}$  and  $160\mu\text{m}$  separation between HPPSC and GH-GOODS-South sources.

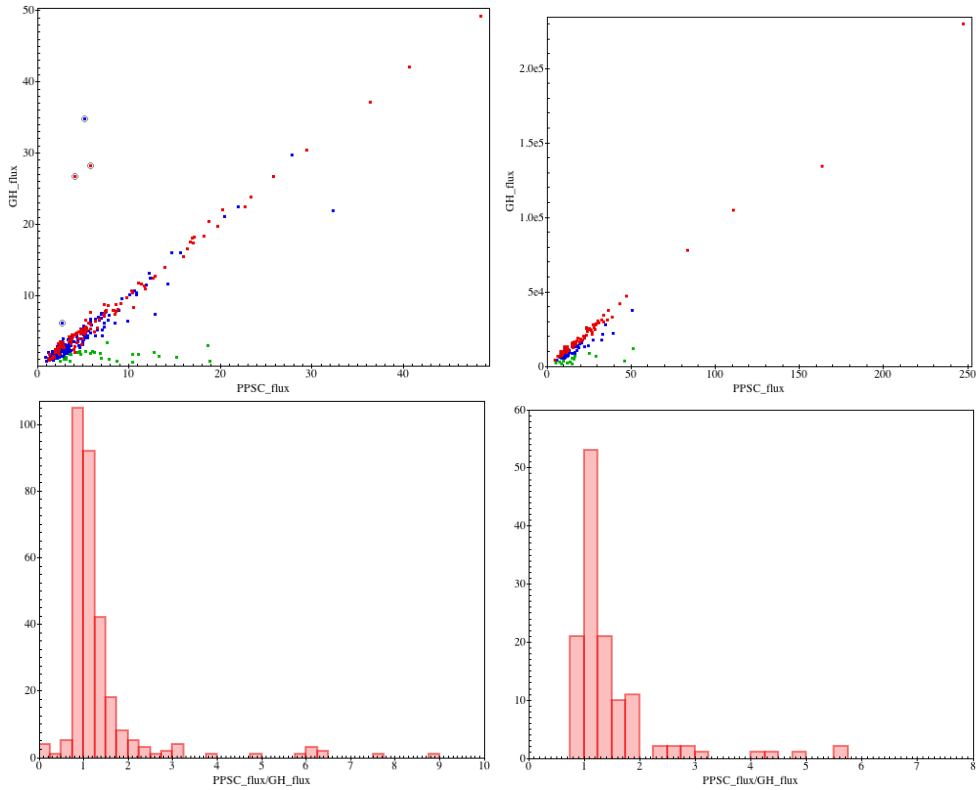


Figure 42: From left to right and top to bottom:  $100\mu\text{m}$  and  $160\mu\text{m}$  HPPSC vs GH-GOODS-South fluxes and HPPSC/GH-GOODS-South fluxes. Red points: all matches, Green Points: outliers with HPPSC/GH fluxes  $> 2$ , Blue points: clean sources, Grey circles: outliers with GH/HPPSC fluxes  $> 2$

The number of matches is 304 in the green band, and 130 at the red band. Match radii have a mean below  $2''$  (green band) and  $3''$  (red band). If we consider all matches, the ratio between the HPPSC and the GH-GOODS-S fluxes is 1.6 in two bands. We have found 29 (green) and 14 (red) outliers (with flux ratio bigger than 2), but if we constrain our comparison to clean sources, the flux ratios are 1.01 and 1.03, for the green and the red bands, respectively. Furthermore, none of the outliers are clean sources. At  $100\mu\text{m}$ , we detect also 5 sources which have considerably larger fluxes in the GH-GOODS catalogue. We show these objects in Figure 43. All of them are at the edge of the field, such that the GH-GOODS fluxes are not accurate.

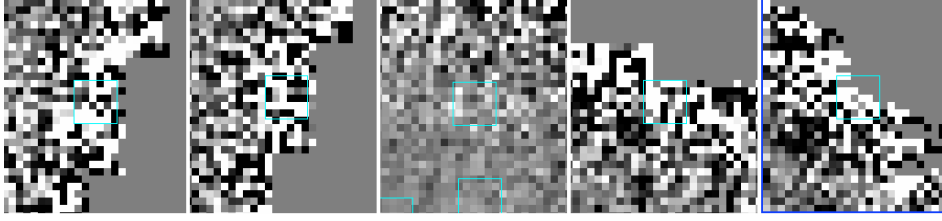


Figure 43: Thumbnails of 5 objects found in HPPSC/GH-GOODS-South match with GH-GOODS-S/HPPSC fluxes ratio larger than 2. Cyan squares: GH-GOODS-South sources.

In conclusion, the agreement (both in position and in flux) between the two catalogues is really excellent, if we consider only 'clean sources' (sources without contamination by other objects) in the GH-catalogue, despite the use of different methods to detect and extract the flux of sources. The flux discrepancies between clean sources and our sources are explained due to them being at the edge of the GH images.

### 5.2.6 The Galactic Cold Cores OTKP

The "Galactic Cold Cores: A *Herschel* survey of the source populations revealed by Planck" was an Open Time Key Programme (Proposal ID: KPOT\_mjuvela\_1, PI:Mika Juvela). The goal of the project was to study starless cores and the initial conditions of star formation using the *Herschel* satellite and the combined power of its PACS and SPIRE instruments. Their catalogue was published in [Montillaud et al. \(2015\)](#). The *getsources* multiwavelength source extraction algorithm was employed in their work to build a catalogue of several thousand cold sources, extracted from 116 fields. *getsources* is a powerful multi-scale, multi-wavelength source extraction algorithm, designed primarily for use in FIR surveys of star-forming regions observed with *Herschel*.

The GCC catalogue contains 4466 rows, listing band-merged data. Most of their targets were very cold ( $<10\text{K}$ ) sources, therefore their PACS data is not very reliable, and only  $160\mu\text{m}$  flux densities are reported. To match their findings with our catalogue we used the source matching tool of TOPCAT. The maximum separation allowed was  $12''$ . 400 cases were found in which a GCC object was found within a  $12''$  radius of our HPPSC sources. The distribution of the angular separations between the matching sources is shown on Figure 44. The distribution is rather uniform, suggesting that the matches are more coincidental than systematic.

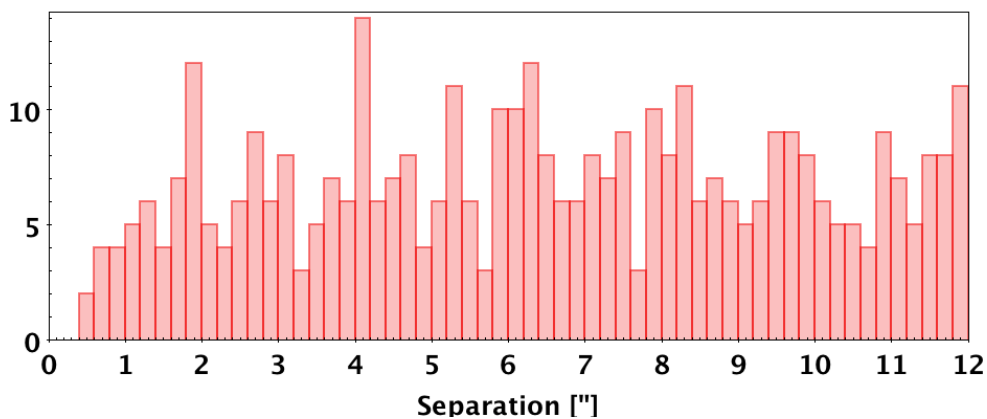


Figure 44: Angular separation between the overlapping catalogue sources.

*getsources* calculates two different flux values. One of them is the peak flux, while the other is the integrated flux measured inside the footprint of the source. We compared both with our HPPSC flux values. Figure 45 shows the GCC peak (left) and the GCC total (right) flux as a function of our values. A perfect match is shown by the dashed red line. The corresponding errors for our values are the ones calculated from the structure noise based  $S/N_S$  value. The figures show that their peak flux is in better agreement with our fluxes, especially at the bright end of

the diagram. The integrated GCC fluxes are systematically higher than the ones we calculate by aperture photometry.

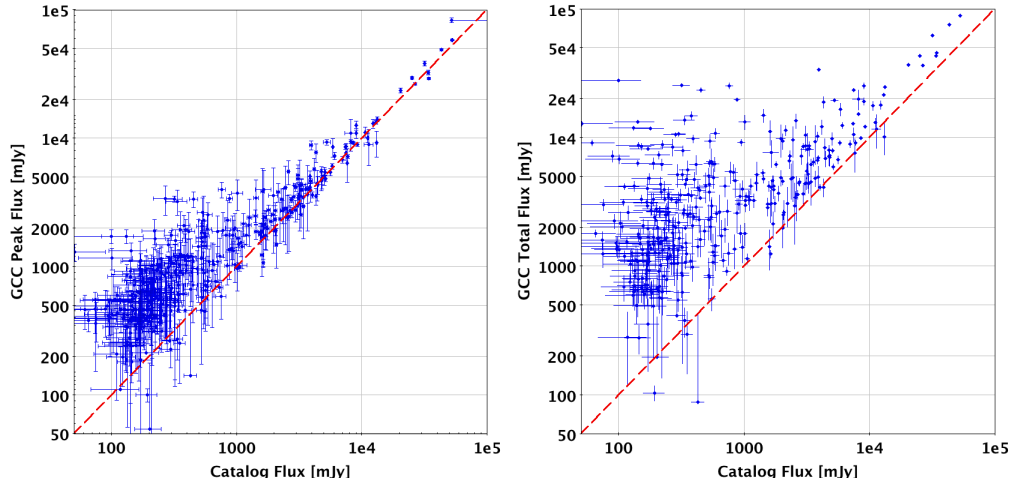


Figure 45: Left: The GCC peak flux as measured by *getsources* as a function of our HPPSC flux. Right The total GCC flux as integrated by *getsources* as a function of the HPPSC flux. Error bars represent the corresponding flux uncertainty reported by the GCC catalogue and the structure noise based statistical error for our measurements.

The size of the difference can be seen if we plot the flux ratios for both cases. Figure 46 shows the HPPSC flux over the GCC peak flux (left) and the HPPSC flux divided by the GCC integrated flux (right). A secondary maximum is seen in the left image at  $\sim 1$ , meaning that there are sources for which flux the agreement is good. In most cases, the ratio is below 1, meaning that the GCC peak flux is larger than the corresponding HPPSC value. Using the integrated GCC flux we found that in almost all cases the flux reported in the GCC catalogue is higher than our values.

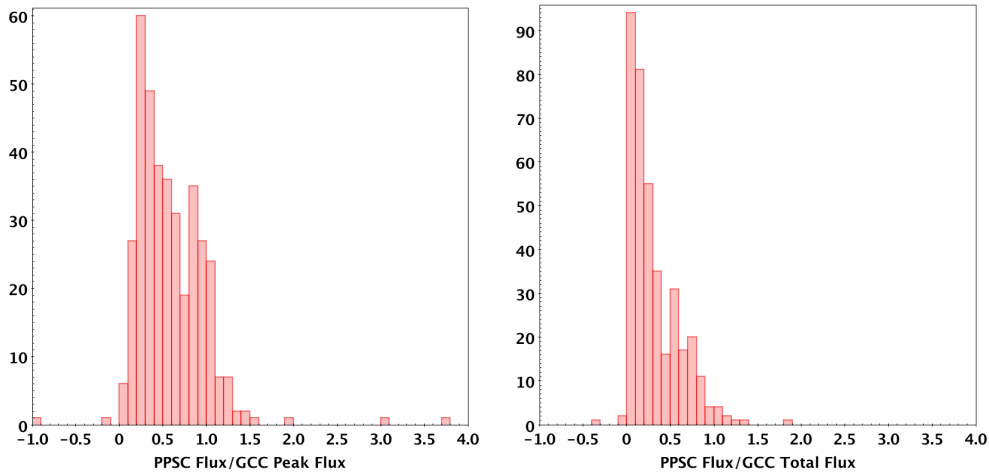


Figure 46: Left: The HPPSC flux over the GCC peak flux. Right: Flux ratios obtained by using the integrated GCC flux.

The large difference between the GCC integrated flux and the HPPSC flux can easily be understood by looking at the size of the GCC source footprints. While our definition of point sources does not allow the fitted FWHM to be larger than 2.0 times the corresponding PSF FWHM ( $< 25''$ ), the GCC FWHM is  $68''.28 \pm 20''.78$ .

### 5.2.7 Hi-GAL DR1

The **Hi-GAL** (*Herschel*-infrared Galactic Plane Survey) project carried out a full survey of the Galactic Plane in 5 bands, including the blue and red PACS band in Parallel Mode. The **first data release (DR1)** from the Hi-GAL team covered the inner Milky Way in the longitude range  $68^\circ \geq l \geq -70^\circ$  in a  $|b| \leq 1^\circ$  latitude strip. The Hi-GAL project used different data processing pipelines to the HPPSC in every aspect. The data reduction was carried out using the ROMAGAL data processing software described in details in [Traficante et al. \(2011\)](#). Source extraction was carried out with the **CuTEx** algorithm, standing for Curvature Thresholding Extractor. The algorithm looks for the pixels in the map with the highest curvature by computing the second derivative of the map. All the “clumps” of pixels above a defined threshold are analysed and those larger than a certain area are kept as candidate detections. The output fluxes and sizes are determined by fitting simultaneously elliptical Gaussian functions plus a 2<sup>nd</sup>-order 2D surface for the background.

The released catalogues contain information on 120 581 and 291 858 sources in the blue and red bands, respectively. We found 22 684 and 40 373 matching sources in the two bands. Figure 47 shows a part of Hi-GAL image Field 297\_0 in the blue band. The picture demonstrates well how our criteria for source shape classification work. The HPPSC sources are well defined point sources on the image. We can also see clearly, how differently the CuTEx algorithm works. Many of the larger, extended objects are divided into individual peaks by the Hi-GAL DR1 pipeline. Similarly, there are many detections by SUSSEXtractor that are not recognised by CuTEx as sources. Our quick comparison with the HPESL showed that many objects from the Hi-GAL DR1 can be also found in our product. The number of matches is  $\sim 2.4 \times 10^4$  and  $\sim 6.8 \times 10^4$  in the blue and red bands, respectively, dependent on the radius used for matching the sources.

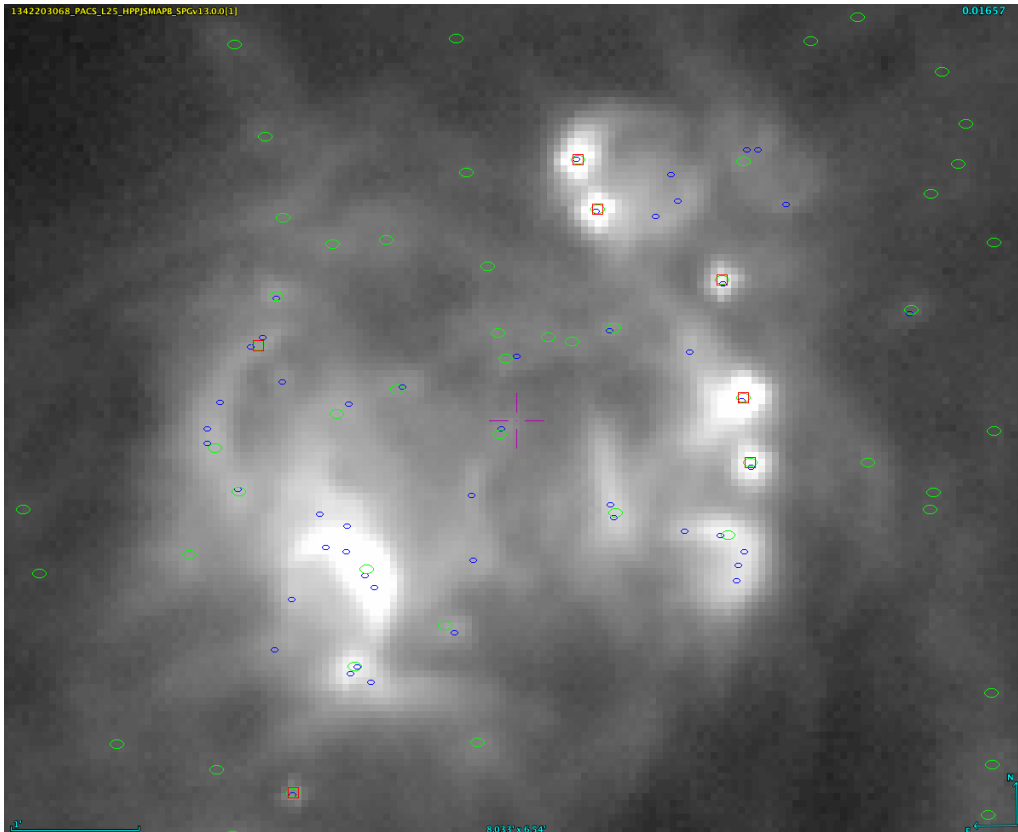


Figure 47: Part of the Hi-GAL tile Field 297\_0 in the blue band. Blue circles indicate the sources from the Hi-GAL DR1 catalogue. The green ovals show all of our extractions from the region. The red squares show the objects in the final HPPSC.

Our flux comparison shows a good agreement with the Hi-GAL extractions, especially for the blue band, see the left panel in Figure 48. The average flux ratio was found to be  $0.91 \pm 0.24$ . The

distribution of the flux ratios is plotted on the right panel of Figure 48

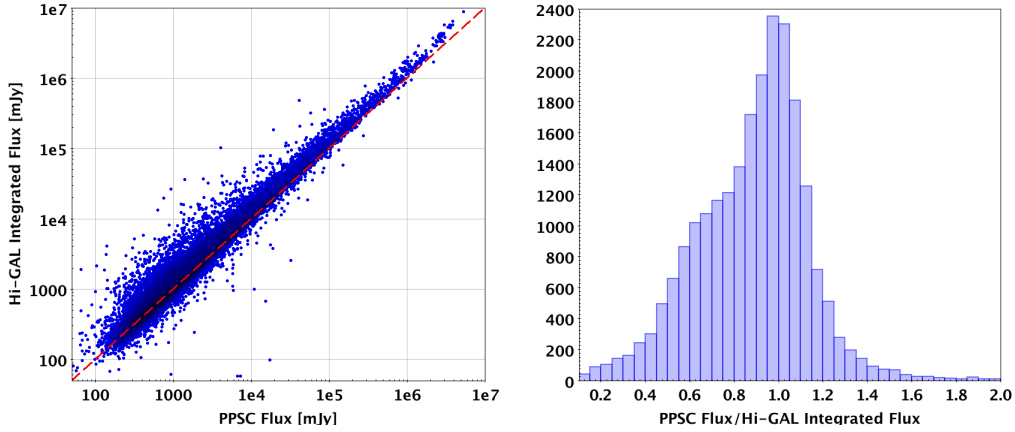


Figure 48: Left: The HPPSC flux in the blue band over the Hi-GAL integrated flux. Right: Flux ratios by using the same flux values.

The correlation is still very good, although the difference between the fluxes is slightly larger (see left panel on Figure 49). The average flux ratio was found to be  $0.88 \pm 0.36$  (right panel).

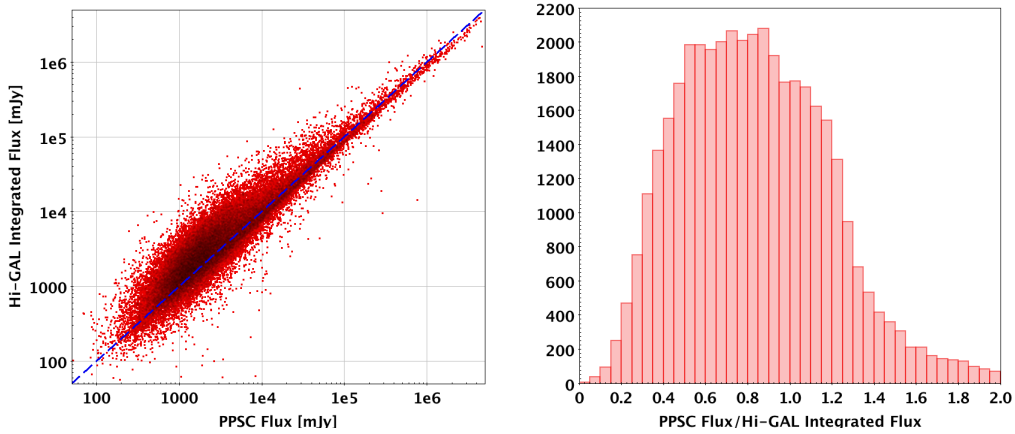


Figure 49: Left: The HPPSC flux in the red band over the Hi-GAL integrated flux. Right: Flux ratios found by using the same flux values.

We have also compared our estimated S/N values to the S/N values calculated by the Hi-GAL team. As seen in Figure 50 for the blue band (left panel) the  $S/N_R$  values (blue dots) occupy an interval of values that is similar to Hi-GAL. The Hi-GAL noise is also a background RMS but, instead of calculating it around the source, it is calculated in the position of the source, after source subtraction from the image. Our  $S/N_S$  values cover a wider interval in the blue band, but they are similar to the Hi-GAL ones in the red band (right panel). The Hi-GAL S/N values are systematically higher than our  $S/N_R$  values.

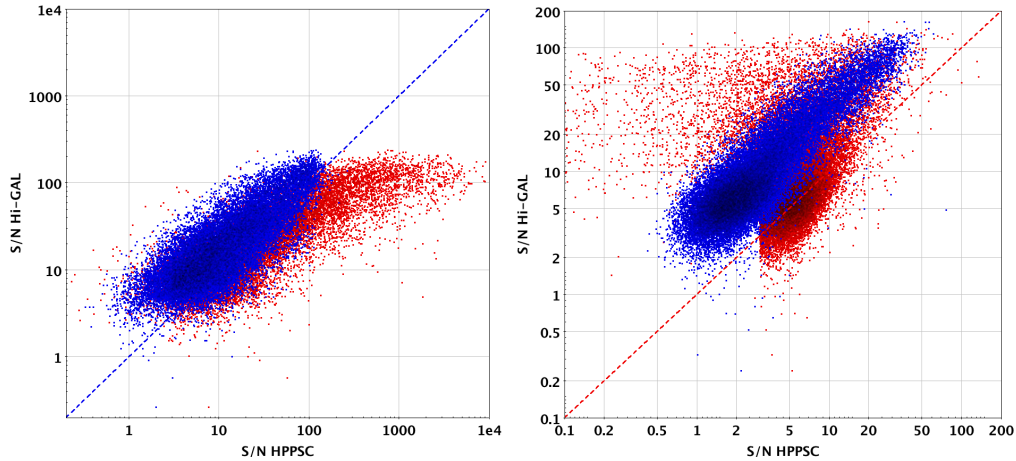


Figure 50: Comparison of the HPPSC  $S/N_S$  (red dots) and  $S/N_R$  (blue dots) with the  $S/N$  values provided by Hi-GAL in the blue band (left panel) and in the red band (right panel).

### 5.3 Conclusion

Our simulations show that the photometry is most reliable in case of environments with low complexity and for sources brighter than  $\sim 100$  mJy. External validation by calibration stars show that our photometry is very accurate and deviates from previous results by less than 1%, especially in the case of the bright targets. Comparison with cosmological catalogues suggest that fluxes of matching sources are generally consistent with a scatter of about 10% or better. In the case of catalogues compiled from observations of star forming regions (therefore containing more complex extended emission) the agreement is less good. The deviation of our fluxes from the Hi-GAL values is  $\sim 10\%$ , with a scatter of  $\sim 25\text{-}30\%$ .

Our completeness is good in fields with low  $N_S$  values, but in complex regions the bright extended emission does not allow the detection of faint sources. In deep extragalactic observations 90% completeness can be reached at a source flux  $\sim 30$  mJy. In complex regions this level of completeness can be reached only at around 1 Jy based on our simulated data. Comparison with existing extragalactic catalogues shows that we are able to recover  $\sim 50\%$  of their sources, and we have detections that are not listed in their work. The overlap with the Galactic Plane catalogues is 20-30%, which may be caused by our criteria for point source classification. Most of the sources catalogued in other works are slightly extended.

Recovered positions are in good agreement with the input positions in our simulations. Comparison with other works suggests that the positional accuracy is within the nominal pointing error, and definitely within the PSF FWHM.

## Acknowledgement

A project of this magnitude has always many more contributors than can be found on the list of authors and small things can make a big difference. For that reason we would like to specifically acknowledge: Bernhard Schulz and the SPIRE Point Source Catalogue team for experimenting the right path through the HPSC project and helping us on our way along.

We would like to acknowledge Erika Varga-Verébelyi and Róbert Szakáts for their contribution in tests and SSO related work. David Shupe has helped us in the exploration of our dataset. John Rector has helped us to create our database schema. Göran Pilbratt, Anthony Marston, David Teyssier, and Pedro Garcia Lario helped the project from the beginning.

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

PACS has been developed by a consortium of institutes led by MPE (Germany) and including UVIE (Austria); KU Leuven, CSL, IMEC (Belgium); CEA, LAM (France); MPIA (Germany); INAF-IFSI/OAA/OAP/OAT, LENS, SISSA (Italy); IAC (Spain). This development has been supported by the funding agencies BMVIT (Austria), ESA-PRODEX (Belgium), CEA/CNES (France), DLR (Germany), ASI/INAF (Italy), and CICYT/MCYT (Spain).

HIPE is a joint development by the *Herschel* Science Ground Segment Consortium, consisting of ESA, the NASA *Herschel* Science Centre, and the HIFI, PACS and SPIRE consortia.

This research has made use of data from the PEP (a guaranteed time key programme of *Herschel*) and the GOODS- *Herschel* Survey (an Open Time Key Programme of *Herschel*). PEP data was downloaded from the NASA/IPAC Science Archive ([irsa.ipac.caltech.edu](http://irsa.ipac.caltech.edu)), and GOODS-*Herschel* data was accessed through the *Herschel* Database in Marseille (<http://hedam.lam.fr/>).

The production of the PACS Point Source Catalogue was supported by the European Space Agency and the Hungarian Space Office through grant PECS 4000109997/13/NL/KML.

## References

- De Graauw, Th., et al., "The Herschel-Heterodyne Instrument for the Far-Infrared (HIFI)", 2010, A&A, 518, L6
- Diolaiti, E. et al. "Analysis of isoplanatic high resolution stellar fields by the StarFinder code", 2000, A&AS, 147, 335
- Elbaz D., et al, "GOODS–*Herschel*: an infrared main sequence for star-forming galaxies", 2011, A&A, 533, 119
- Griffin, M.J., et al., "The Herschel-SPIRE instrument and its in-flight performance", 2010, A&A, 518, L3
- Kiss, Cs., et al, " Determination of confusion noise for far-infrared measurements", 2005, A&A, 430, 343
- Lutz, D. et al, "PACS Evolutionary Probe (PEP) - A *Herschel* key program", 2011, A&A, 532, 90
- Lutz, D. "PACS photometer point spread function", 2012
- Men'shchikov, A., et al., "A multi-scale, multi-wavelength source extraction method: get-sources", 2012, A&A, 542, 81
- Molinari, S. et al., "Hi-GAL: The Herschel Infrared Galactic Plane Survey", 2010, PASP, 122, 314
- Molinari, S. et al., "Source extraction and photometry for the far-infrared and sub-millimeter continuum in the presence of complex backgrounds", 2011, A&A, 530, 133
- Molinari, S. et al., "Hi-GAL, the Herschel Infrared Galactic Plane Survey: photometric maps and compact source catalogues. First data release for Inner Milky Way:  $+68 \geq l \geq -70$ ", 2016, A&A, 591, 33
- Montillaud, J. et al. , "Galactic cold cores. IV. Cold submillimetre sources: catalogue and statistical analysis", 2015 A&A, 584, 92
- Müller, T. et al, " PACS Photometer Passbands and Colour Correction Factors for Various Source SEDs", 2011
- Neugebauer, B. & Beichman, C. A., "INFRARED ASTRONOMICAL SATELLITE (IRAS) CATALOGS AND ATLASES Explanatory Supplement", 1987
- Nielbock, M., et al., "The Herschel PACS photometer calibration. A time dependent flux calibration for the PACS chopped point-source photometry AOT mode", 2013, ExA, 36, 631
- Ott, S.: "The *Herschel* data processing system HIPE and pipelines". ASP Conference Series 434, 139 (2010)
- Pilbratt, G., et al., " *Herschel* space observatory. An ESA facility for far-infrared and submillimetre astronomy". A&A, 518, L1 (2010)
- Poglitsch, A. et al., "The Photodetector Array Camera and Spectrometer (PACS) on the Herschel Space Observatory", 2010, A&A, 518, L2
- Savage, R.S. & Oliver, S., " Bayesian Methods of Astronomical Source Extraction" , ApJ 661, 1339 (2007)
- Stetson, P.B., "DAOPHOT - a computer program for crowded-field stellar photometry". PASP 99, 191 (1987)
- Schultz, B. et al, "SPIRE Point Source Catalog Explanatory Supplement", 2017
- Taylor, M.B., "TOPCAT & STILT: Starlink Table/VOTable Processing Software", in Astronomical Data Analysis Software and Systems XIV, eds. P Shopbell , M. Britton, and R. Ebert, ASP Conf. Ser. 347, p. 29 (2005)
- Traficante, A. et al., "Data reduction pipeline for the Hi-GAL survey", 2011, MNRAS, 416, 2932