

1.3_Hawaii-HDFN

January 18, 2018

1 HDF-N master catalogue

1.1 Preparation of Hawaii-HDFN data

The catalogue comes from dmu0_Hawaii-HDFN.

It contains UBVR_Iz data.

In the catalogue, we keep:

- The identifier (it's unique in the catalogue);
- The position;
- The stellarity;
- The kron magnitude, there doesn't appear to be aperture magnitudes. This may mean the survey is unusable.

This notebook was run with herschelhelp_internal version:
33f5ec7 (Wed Dec 6 16:56:17 2017 +0000)

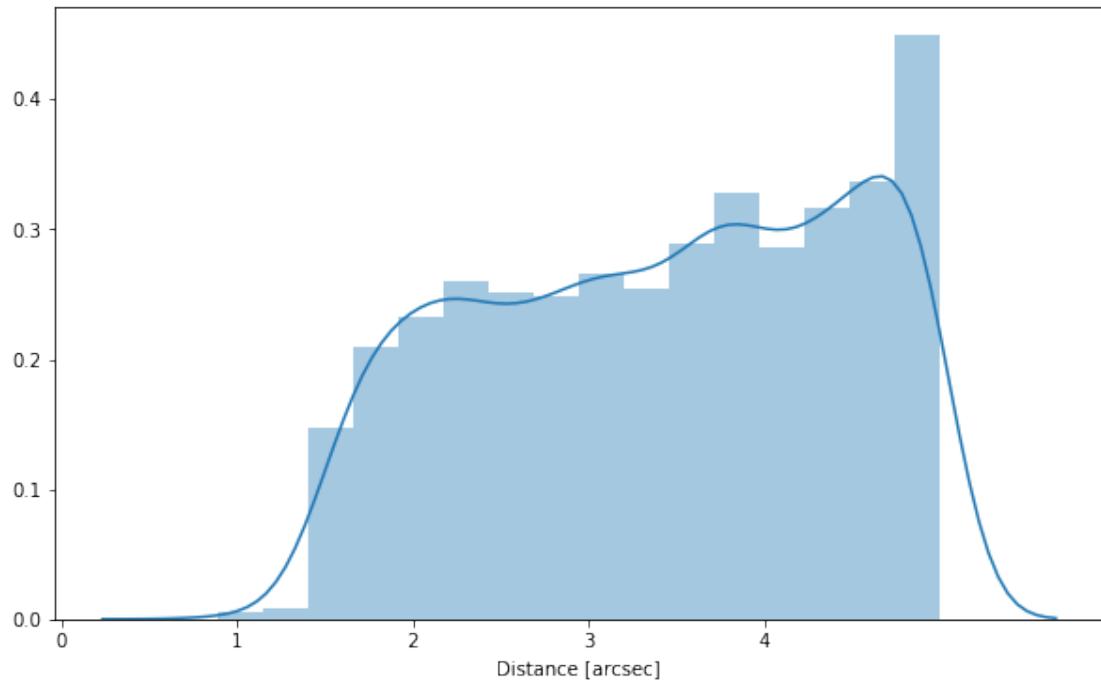
1.2 I - Column selection

Out[4]: <IPython.core.display.HTML object>

Out[7]: <IPython.core.display.HTML object>

Out[8]: <IPython.core.display.HTML object>

1.3 III - Merging



The catalogues appear to contain no cross matches. We therefore simply stack the catalogues. We need to understand why this is the case. Have cross matches already been removed from the Z selected catalogue?

Out [12]: <IPython.core.display.HTML object>

1.4 II - Removal of duplicated sources

We remove duplicated objects from the input catalogues.

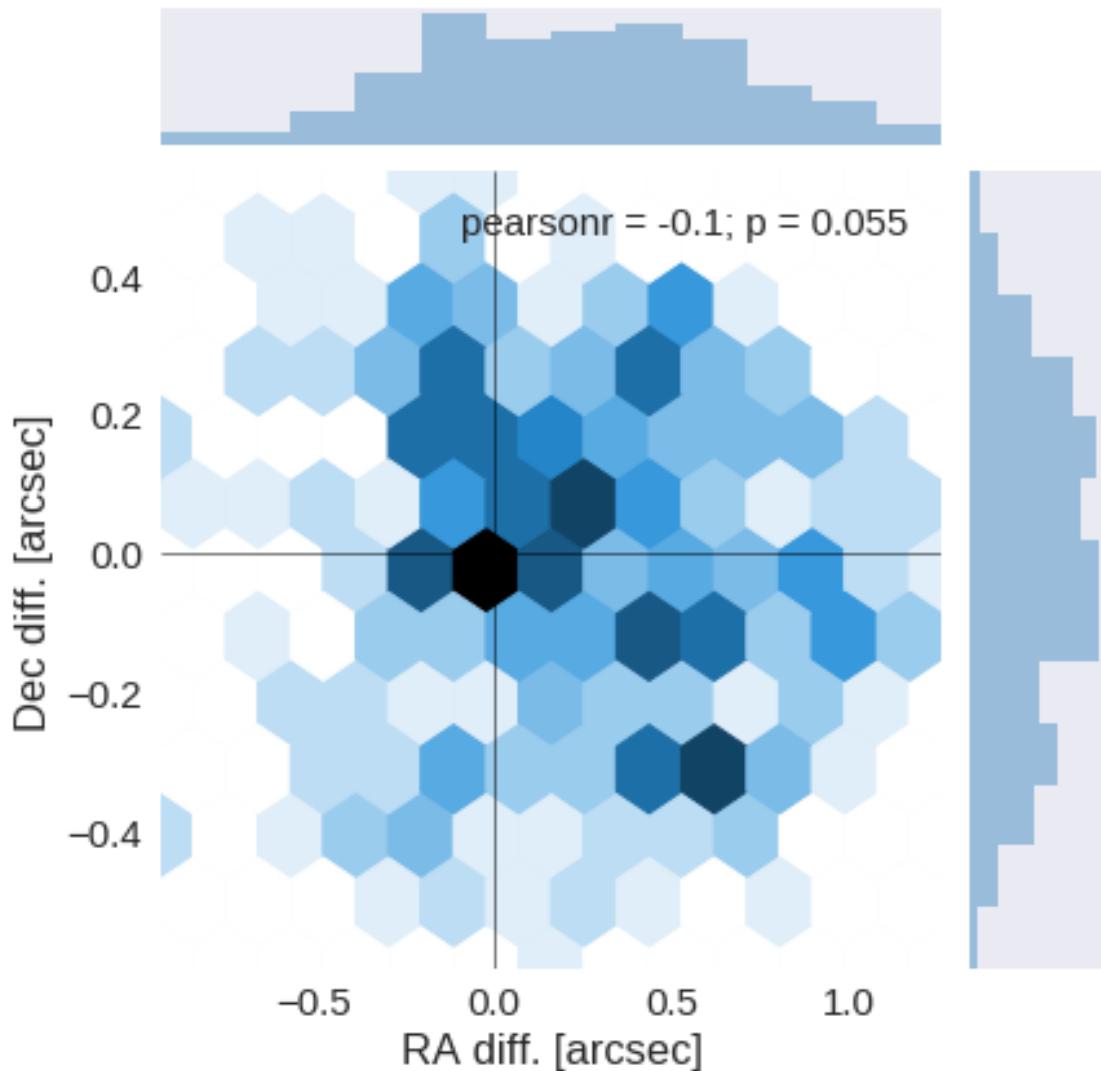
The initial catalogue had 48858 sources.

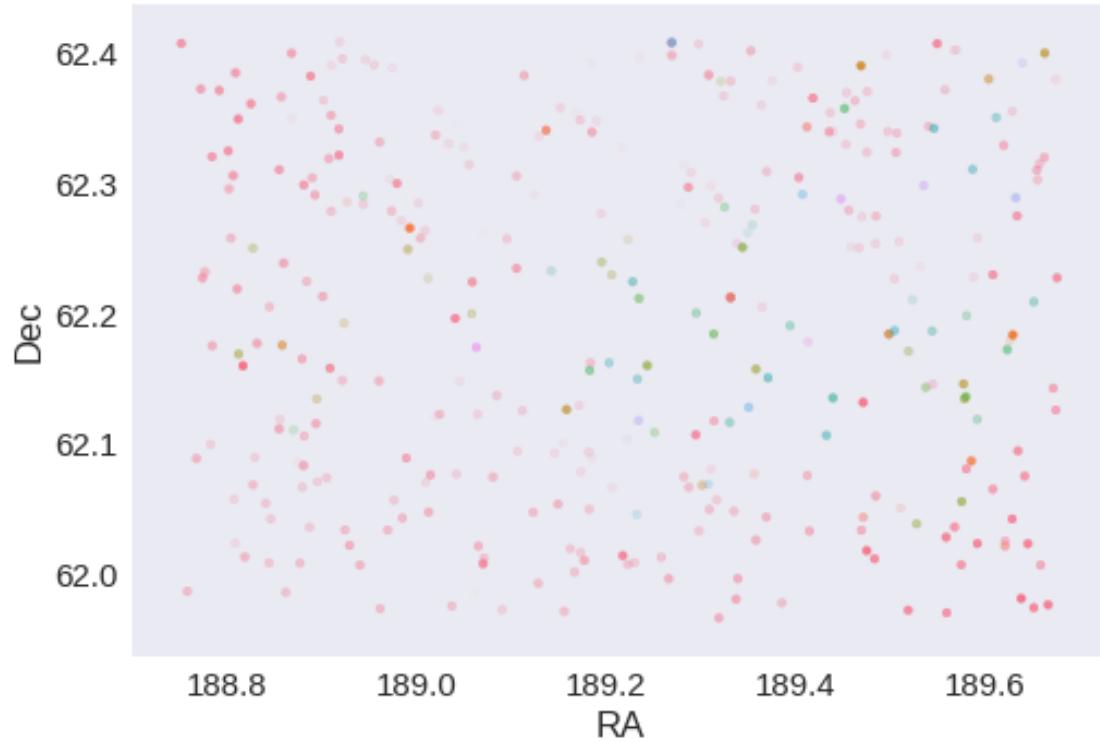
The cleaned catalogue has 48858 sources (0 removed).

The cleaned catalogue has 0 sources flagged as having been cleaned

1.5 III - Astrometry correction

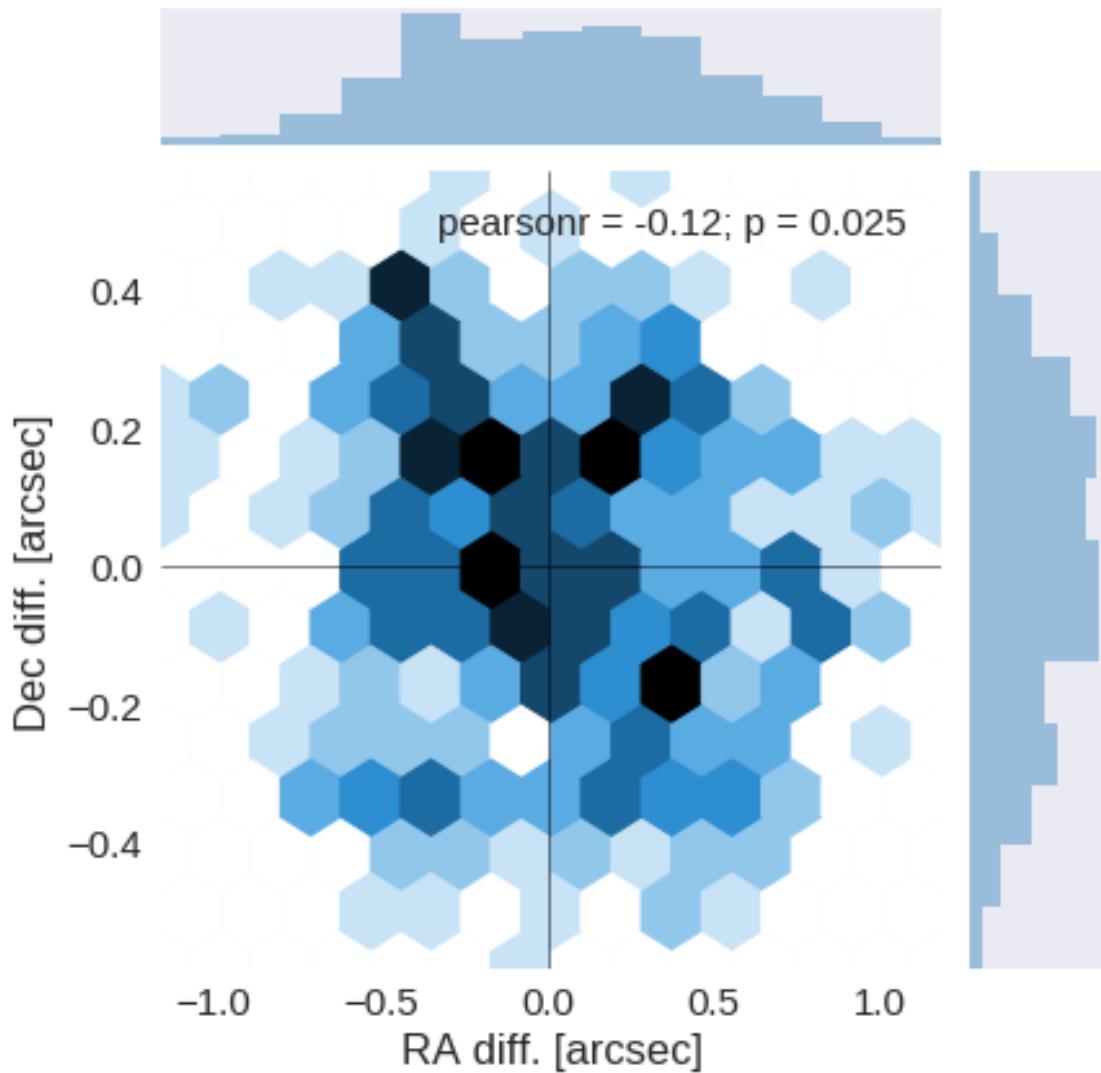
We match the astrometry to the Gaia one. We limit the Gaia catalogue to sources with a g band flux between the 30th and the 70th percentile. Some quick tests show that this give the lower dispersion in the results.

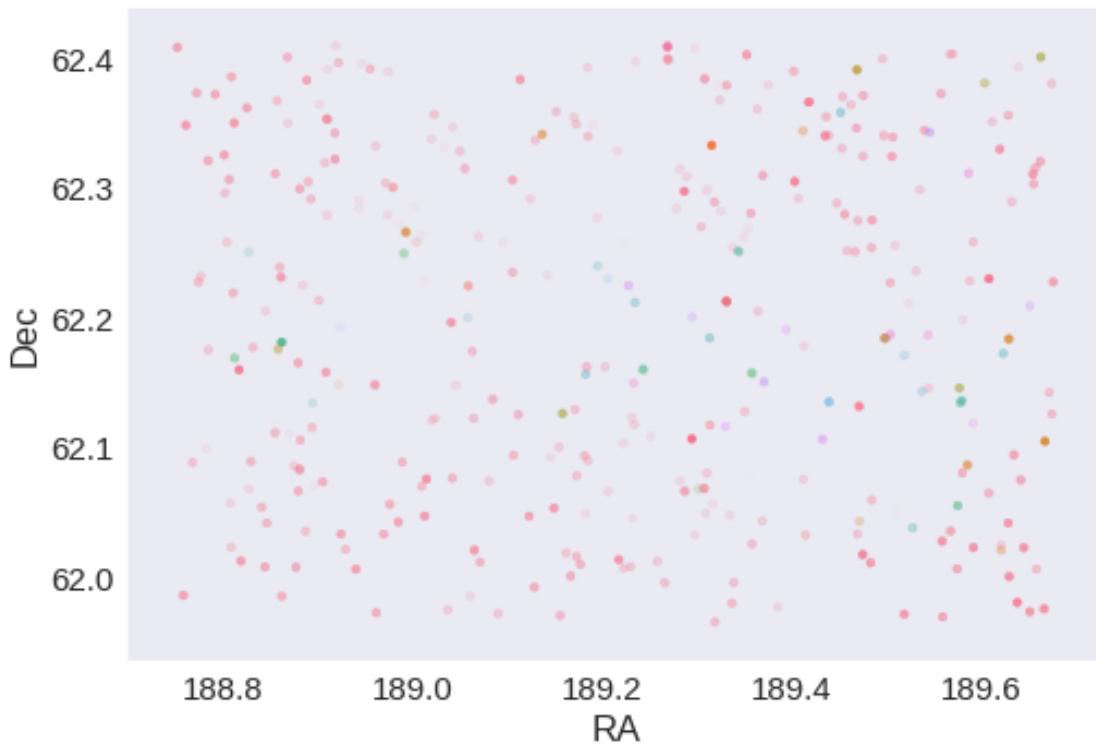




RA correction: -0.22843071726015296 arcsec

Dec correction: 0.014955391984017297 arcsec





1.6 IV - Flagging Gaia objects

431 sources flagged.

1.7 V - Flagging objects near bright stars

2 VI - Saving to disk

Out[21]: 48858

Out[22]: <IPython.core.display.HTML object>

1.4_Ultradeep-Ks-GOODS-N

January 18, 2018

1 HDF-N master catalogue

1.1 Preparation of Ultradeep-Ks-GOODS-N data

The catalogue comes from dmu0_Ultradeep-Ks-GOODS-N.

In the catalogue, we keep:

- The identifier (it's unique in the catalogue);
- The position;
- The total flux.

This notebook was run with herschelhelp_internal version:
33f5ec7 (Wed Dec 6 16:56:17 2017 +0000)

1.2 I - Column selection

Out [6]: <IPython.core.display.HTML object>

1.3 II - Removal of duplicated sources

We remove duplicated objects from the input catalogues.

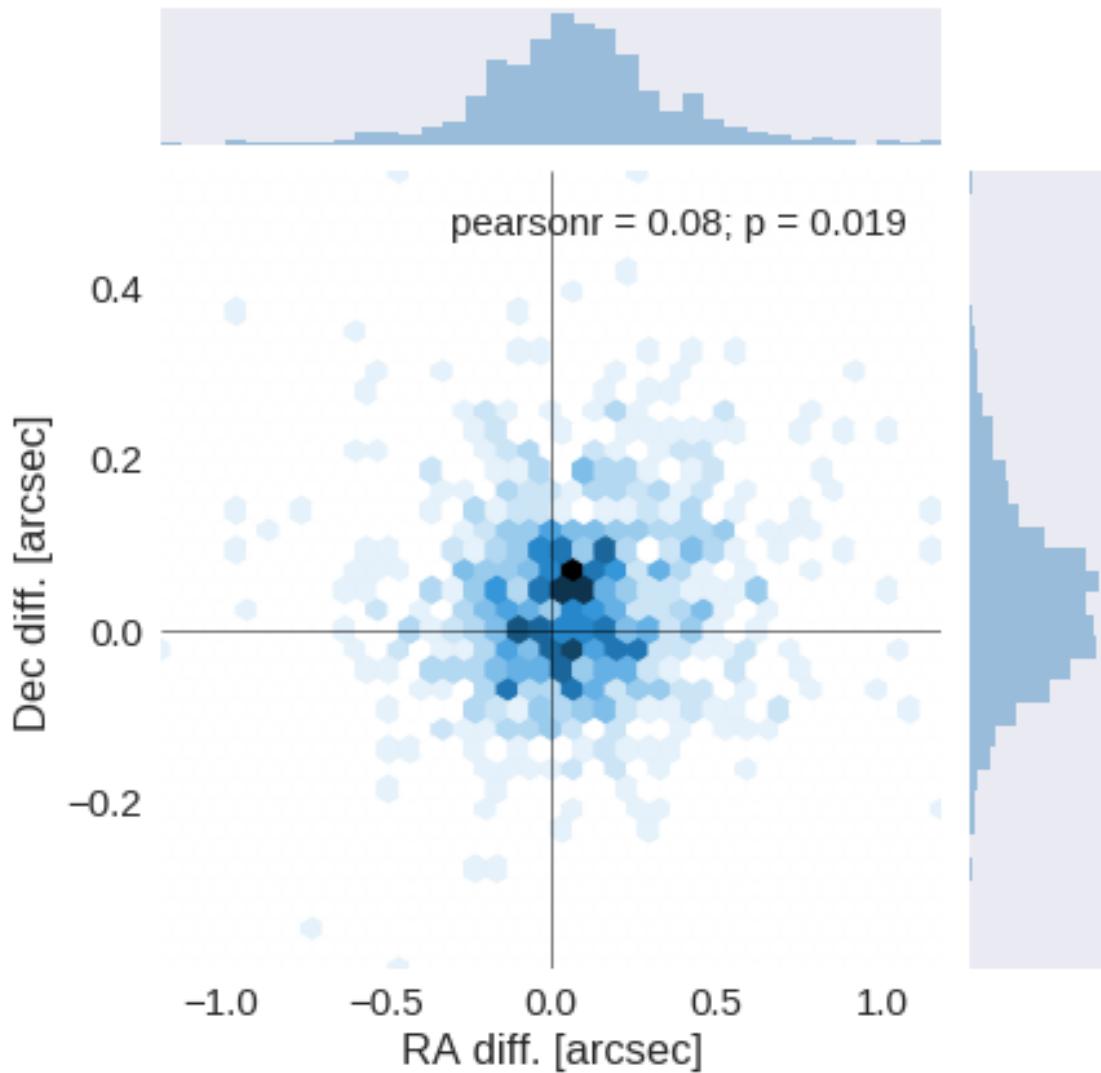
The initial catalogue had 94930 sources.

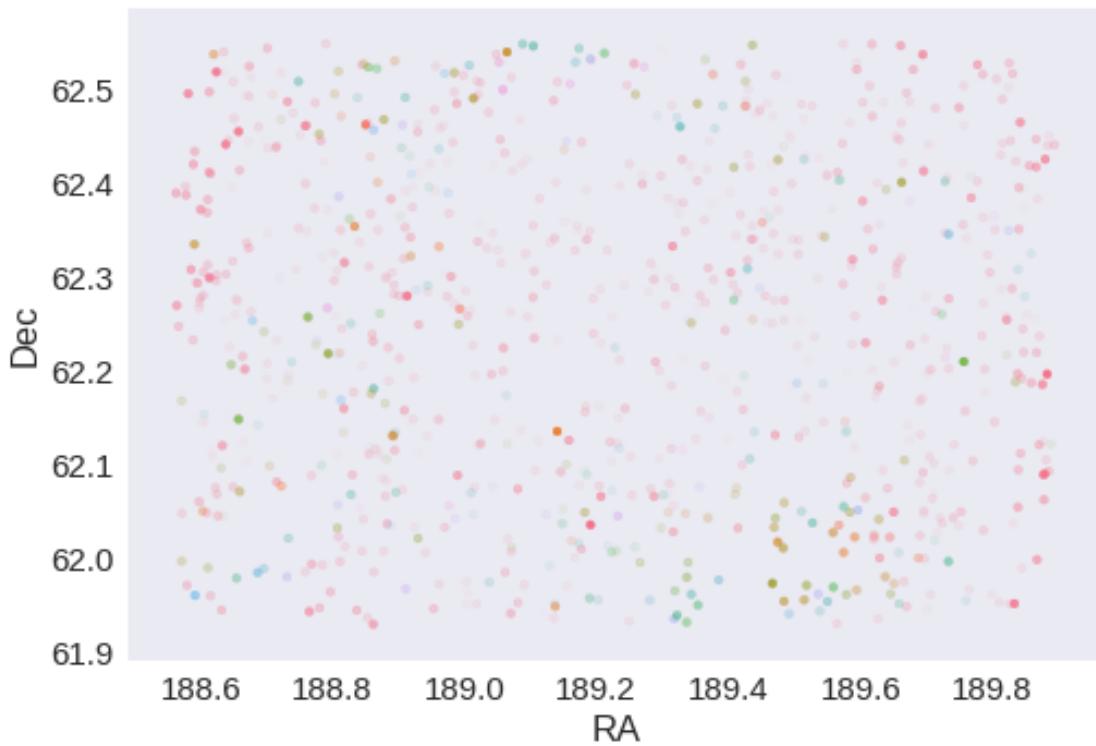
The cleaned catalogue has 94930 sources (0 removed).

The cleaned catalogue has 0 sources flagged as having been cleaned

1.4 III - Astrometry correction

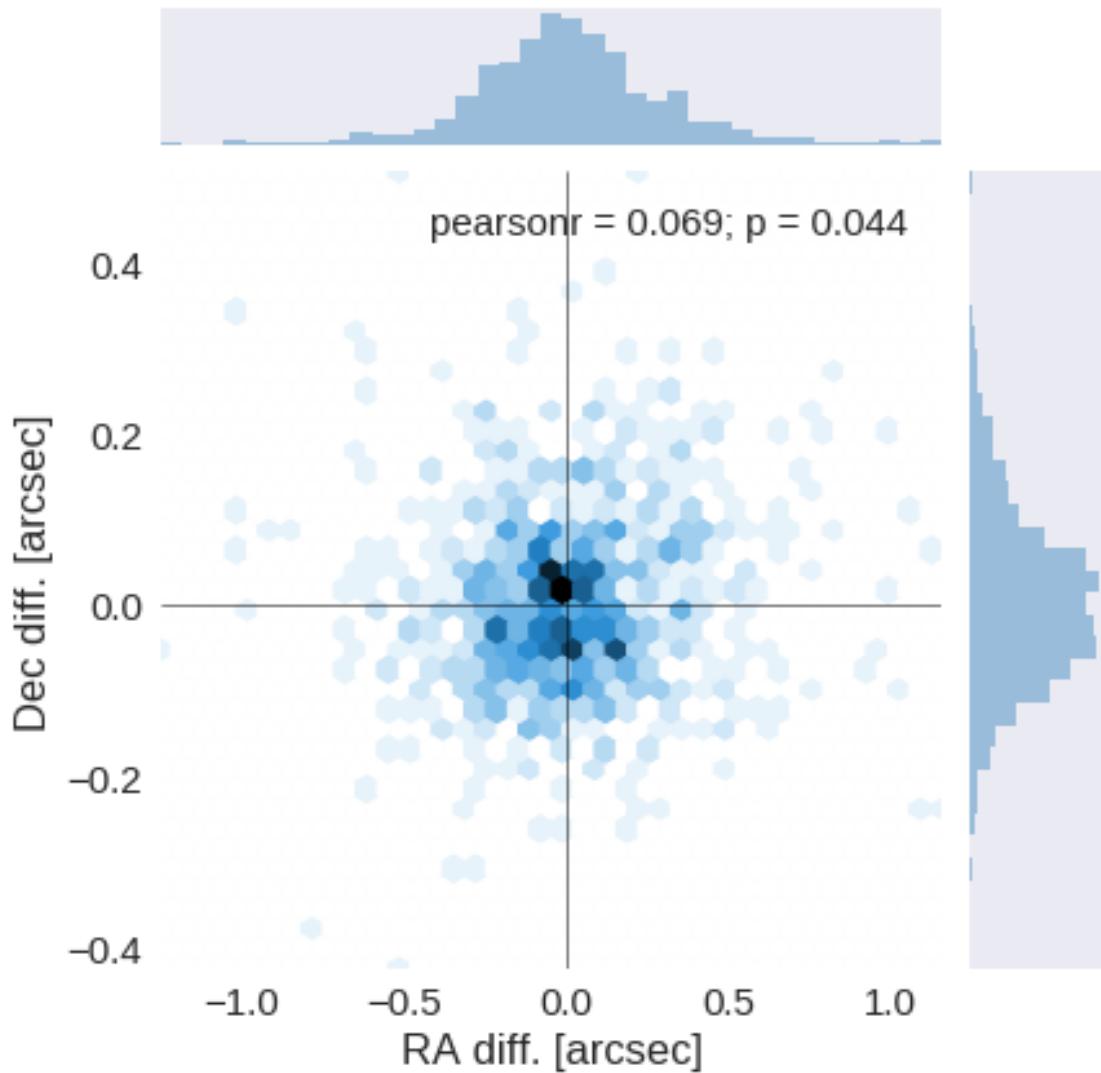
We match the astrometry to the Gaia one. We limit the Gaia catalogue to sources with a g band flux between the 30th and the 70th percentile. Some quick tests show that this give the lower dispersion in the results.

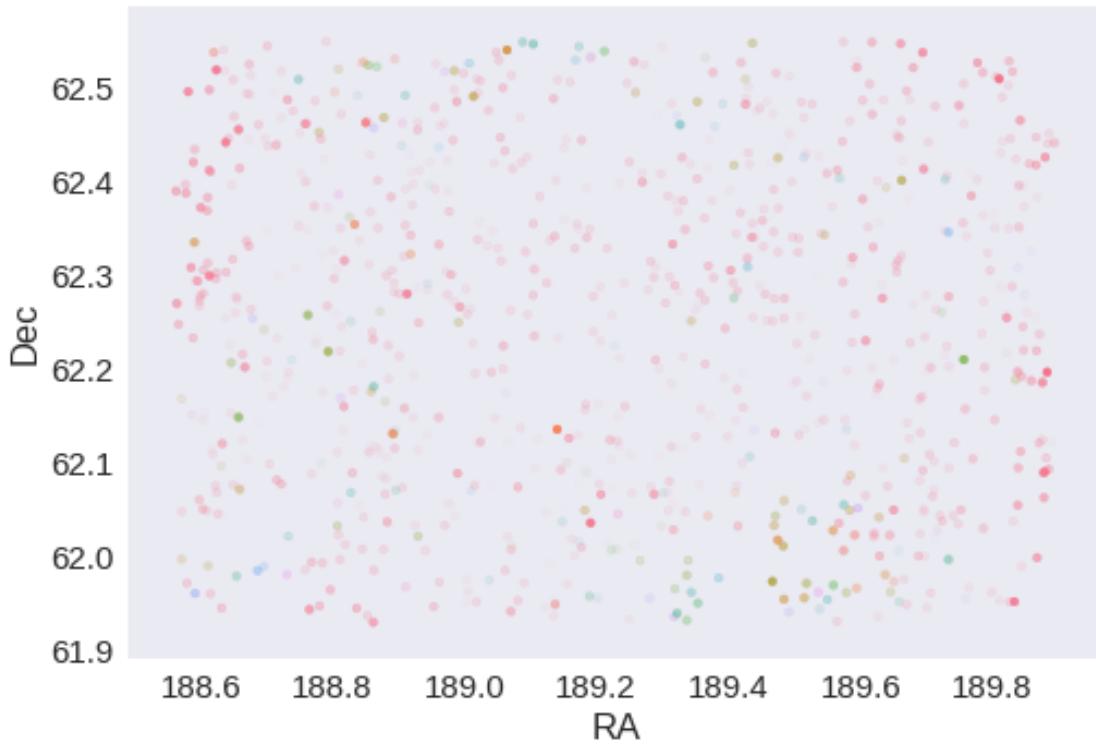




RA correction: -0.07018150315616367 arcsec

Dec correction: -0.02818903770247516 arcsec





1.5 IV - Flagging Gaia objects

888 sources flagged.

1.6 V - Flagging objects near bright stars

2 VI - Saving to disk

1.5_PanSTARRS1-3SS

January 18, 2018

1 HDF-N master catalogue

1.1 Preparation of Pan-STARRS1 - 3pi Steradian Survey (3SS) data

This catalogue comes from dm0_PanSTARRS1-3SS.

In the catalogue, we keep:

- The uniquePspSSTid as unique object identifier;
- The r-band position which is given for all the sources;
- The grizy <band>FApMag aperture magnitude (see below);
- The grizy <band>FKronMag as total magnitude.

The Pan-STARRS1-3SS catalogue provides for each band an aperture magnitude defined as “In PS1, an ‘optimal’ aperture radius is determined based on the local PSF. The wings of the same analytic PSF are then used to extrapolate the flux measured inside this aperture to a ‘total’ flux.”

The observations used for the catalogue were done between 2010 and 2015 ([ref](#)).

TODO: Check if the detection flag can be used to know in which bands an object was detected to construct the coverage maps.

TODO: Check for stellarity.

This notebook was run with herschelhelp_internal version:
33f5ec7 (Wed Dec 6 16:56:17 2017 +0000)

1.2 I - Column selection

```
/opt/anaconda3/envs/herschelhelp_internal/lib/python3.6/site-packages/astropy/table/column.py:10
Check the NumPy 1.11 release notes for more information.
    ma.MaskedArray.__setitem__(self, index, value)
```

Out [6]: <IPython.core.display.HTML object>

1.3 II - Removal of duplicated sources

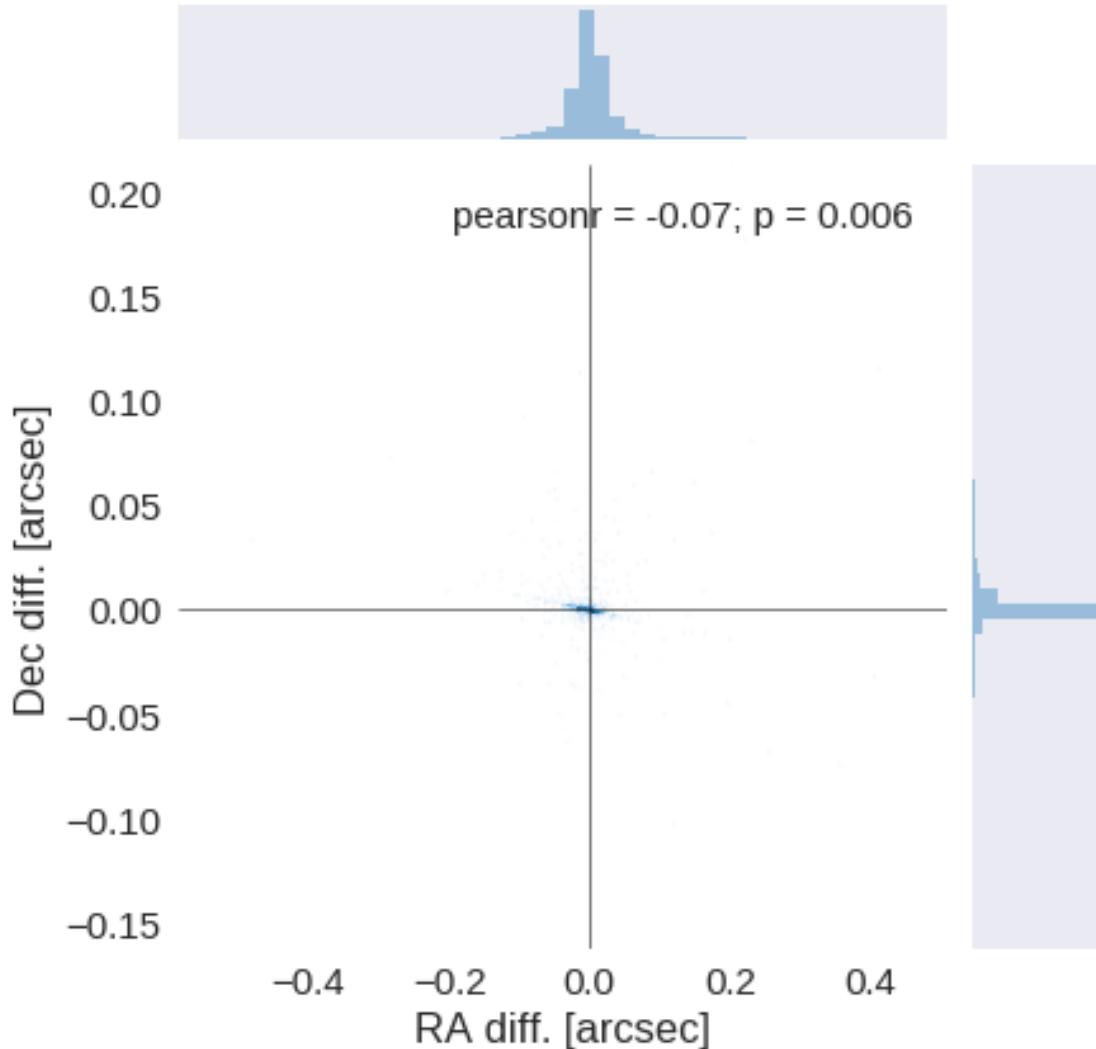
We remove duplicated objects from the input catalogues.

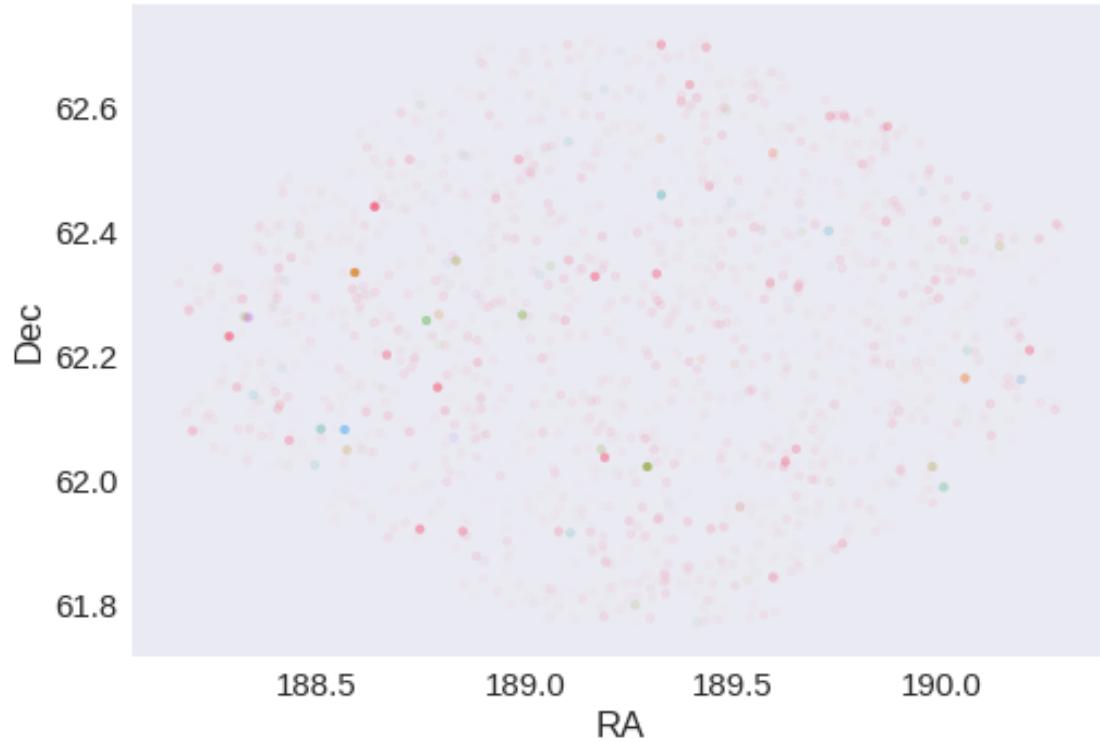
```
/opt/anaconda3/envs/herschelhelp_internal/lib/python3.6/site-packages/astropy/table/column.py:10
Check the NumPy 1.11 release notes for more information.
    ma.MaskedArray.__setitem__(self, index, value)
```

The initial catalogue had 10100 sources.
The cleaned catalogue has 10097 sources (3 removed).
The cleaned catalogue has 3 sources flagged as having been cleaned

1.4 III - Astrometry correction

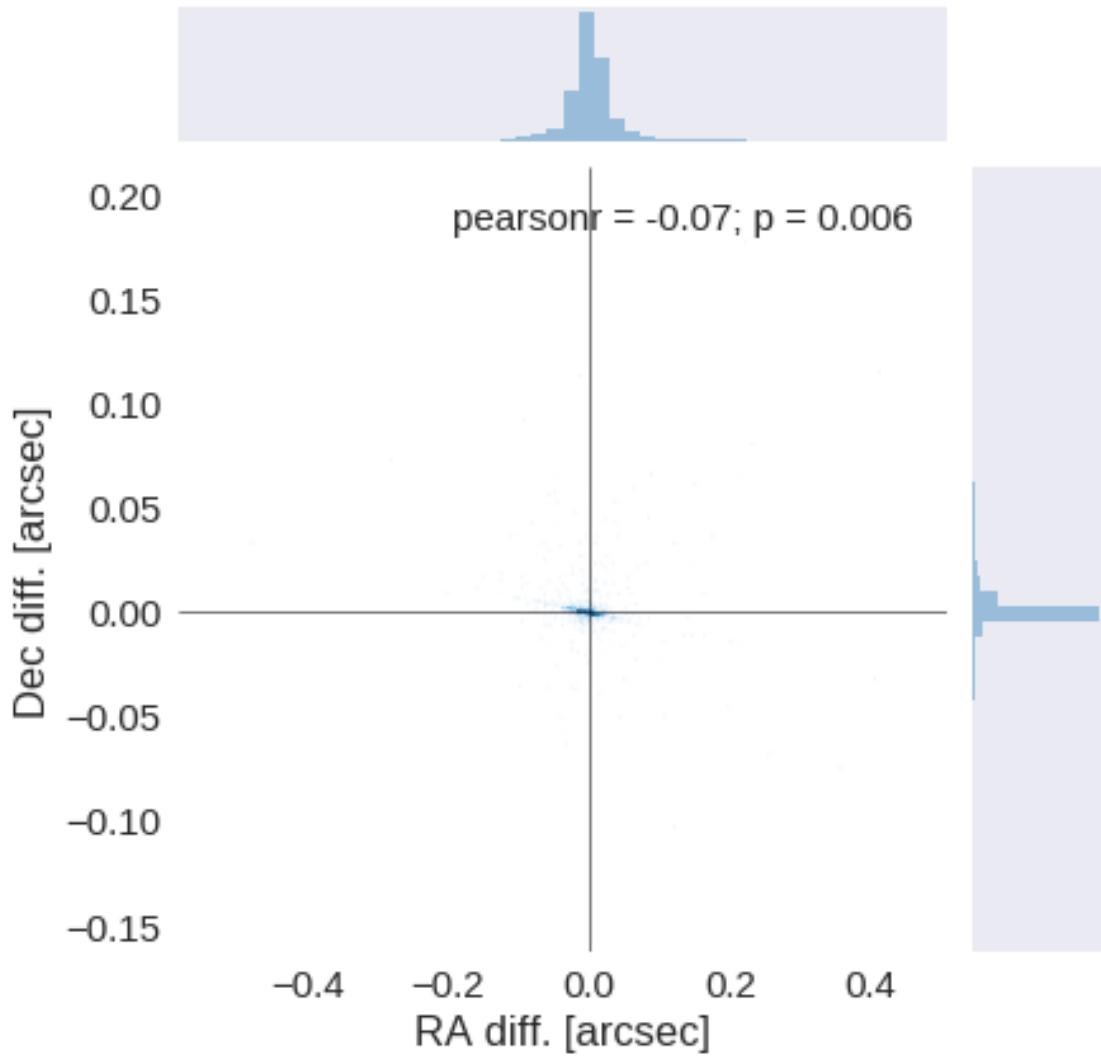
We match the astrometry to the Gaia one. We limit the Gaia catalogue to sources with a g band flux between the 30th and the 70th percentile. Some quick tests show that this give the lower dispersion in the results.

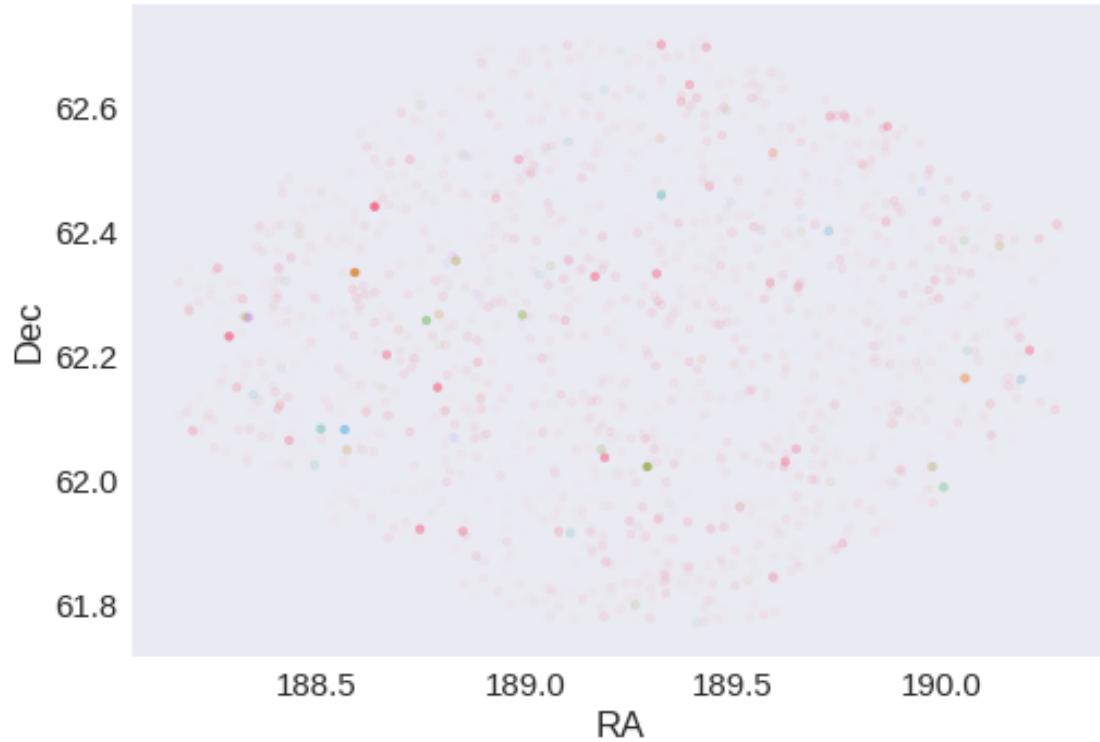




RA correction: 0.0009251711844626698 arcsec

Dec correction: -0.000312677103408987 arcsec





1.5 IV - Flagging Gaia objects

1548 sources flagged.

2 V - Saving to disk

1.6_CANDELS-GOODS-N

January 18, 2018

1 HDF-N master catalogue

1.1 Preparation of CANDELS-GOODS-N data

CANDELS-GOODS-N catalogue: the catalogue comes from `dmu0_CANDELS-GOODS-N`.

In the catalogue, we keep:

- The identifier (it's unique in the catalogue);
- The position;
- The stellarity;
- The total magnitude.

We don't know when the maps have been observed. We will use the year of the reference paper.

This notebook was run with `herschelhelp_internal` version:
33f5ec7 (Wed Dec 6 16:56:17 2017 +0000)

1.2 I - Column selection

```
WARNING: UnitsWarning: '0 use, gt 1 no-use' did not parse as fits unit: At col 2, Unit 'use' not
WARNING: UnitsWarning: '0 galaxy, 1 star' did not parse as fits unit: At col 2, Unit 'galaxy' no
```

```
/opt/herschelhelp_internal/herschelhelp_internal/utils.py:76: RuntimeWarning: divide by zero enco
    magnitudes = 2.5 * (23 - np.log10(fluxes)) - 48.6
/opt/herschelhelp_internal/herschelhelp_internal/utils.py:76: RuntimeWarning: invalid value enco
    magnitudes = 2.5 * (23 - np.log10(fluxes)) - 48.6
/opt/herschelhelp_internal/herschelhelp_internal/utils.py:80: RuntimeWarning: divide by zero enco
    errors = 2.5 / np.log(10) * errors_on_fluxes / fluxes
```

Out [6]: <IPython.core.display.HTML object>

1.3 II - Removal of duplicated sources

We remove duplicated objects from the input catalogues.

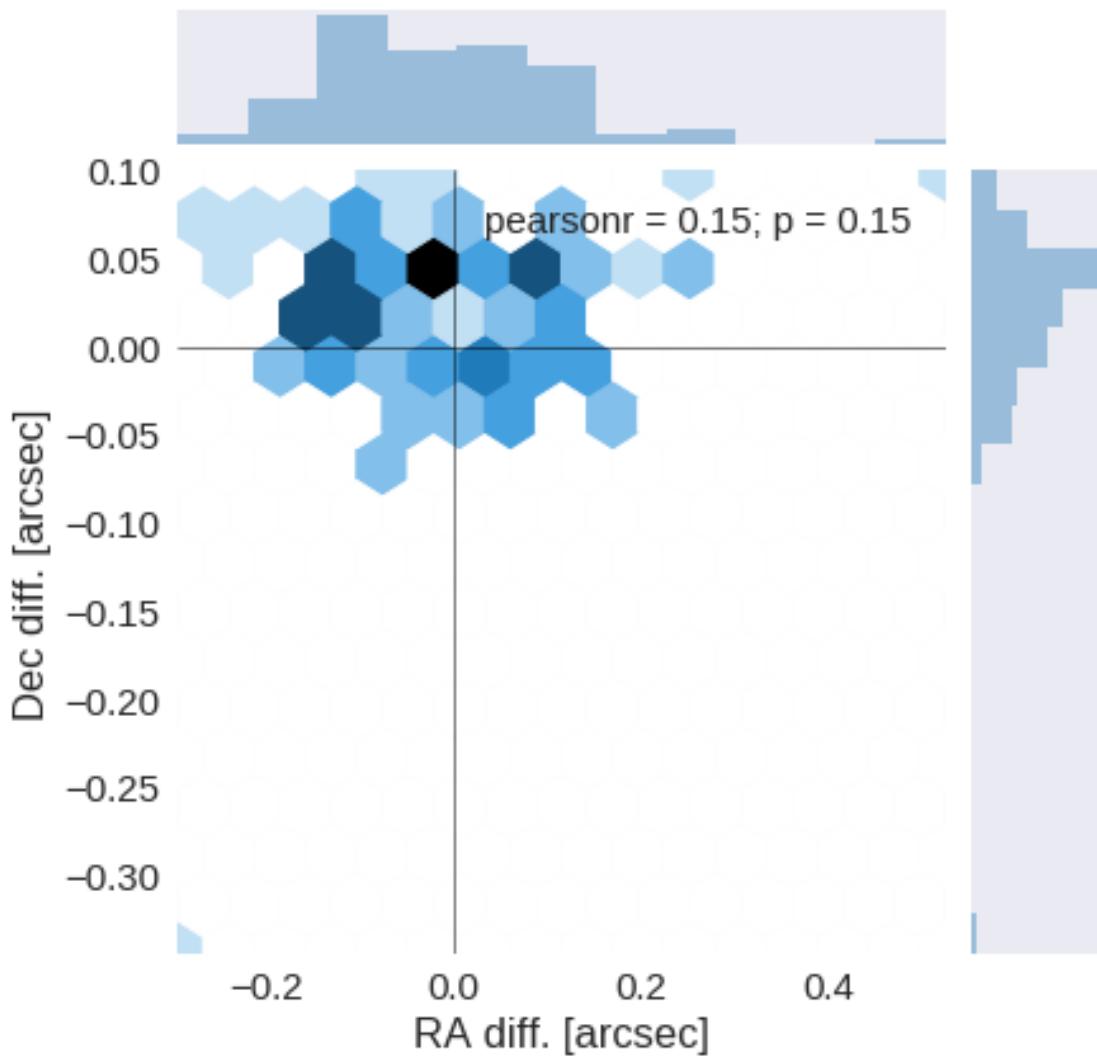
The initial catalogue had 35445 sources.

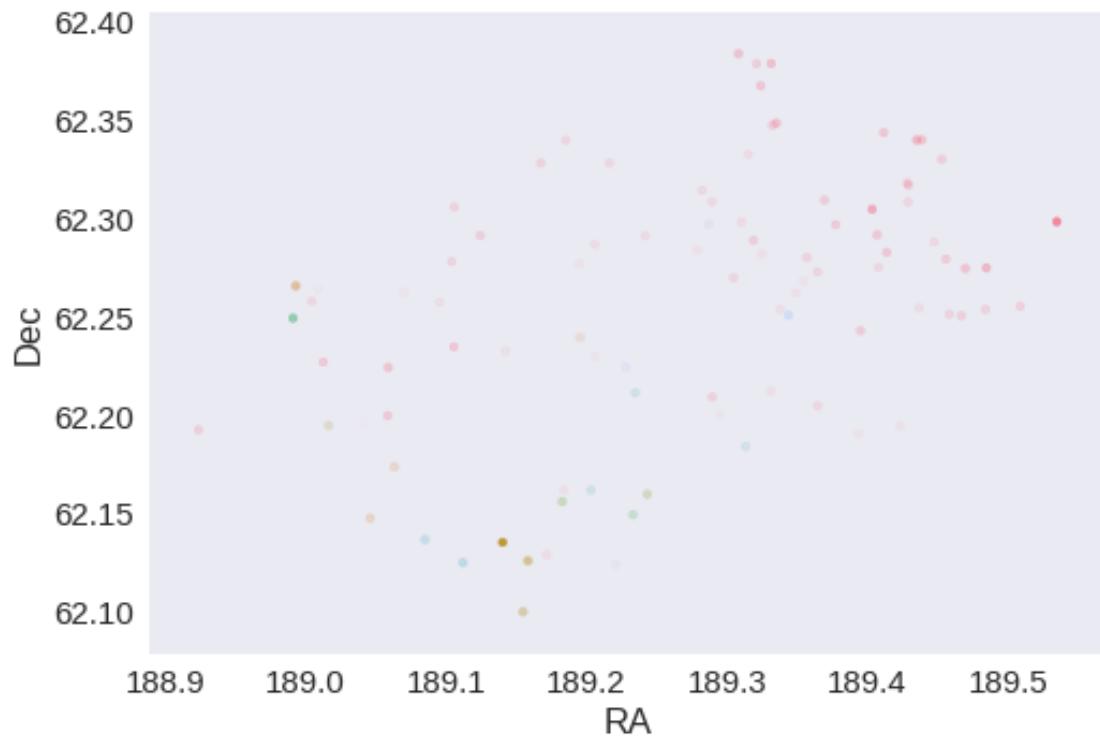
The cleaned catalogue has 35423 sources (22 removed).

The cleaned catalogue has 21 sources flagged as having been cleaned

1.4 III - Astrometry correction

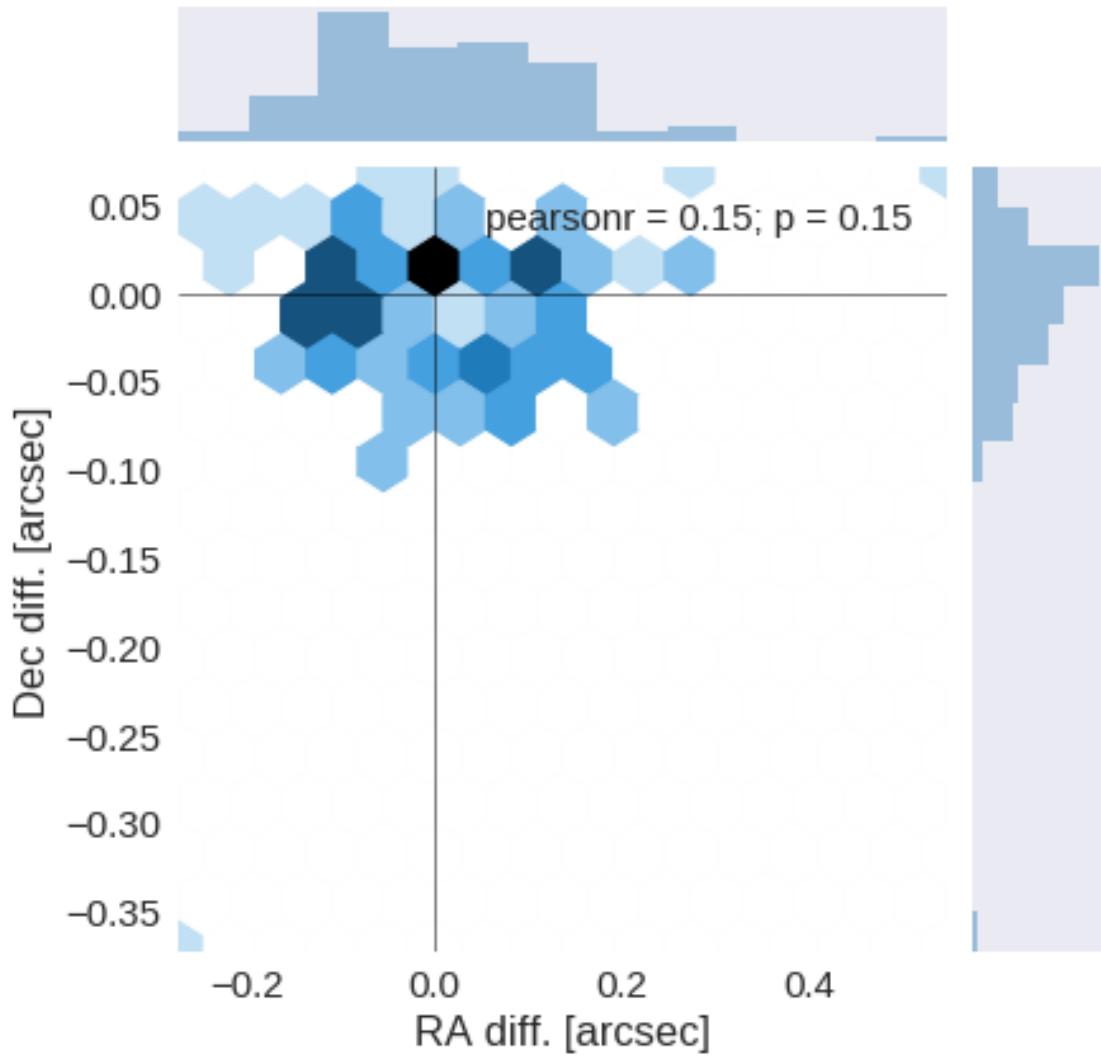
We match the astrometry to the Gaia one. We limit the Gaia catalogue to sources with a g band flux between the 30th and the 70th percentile. Some quick tests show that this give the lower dispersion in the results.

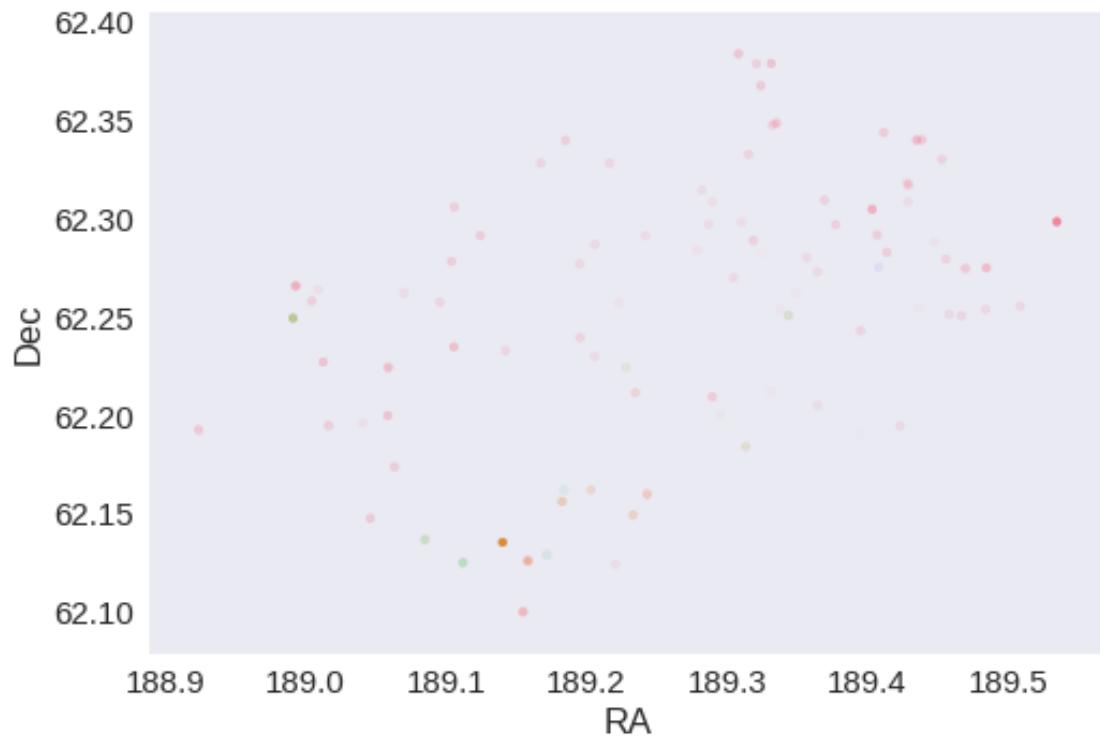




RA correction: 0.02082349079728374 arcsec

Dec correction: -0.02865639995377478 arcsec





1.5 IV - Flagging Gaia objects

102 sources flagged.

2 V - Saving to disk

2_Merging

January 18, 2018

1 HDF-N master catalogue

This notebook presents the merge of the various pristine catalogues to produce the HELP master catalogue on HDF-N.

This notebook was run with herschelhelp_internal version:
33f5ec7 (Wed Dec 6 16:56:17 2017 +0000)

1.1 I - Reading the prepared pristine catalogues

WARNING: UnitsWarning: '0 galaxy, 1 star' did not parse as fits unit: At col 2, Unit 'galaxy' no

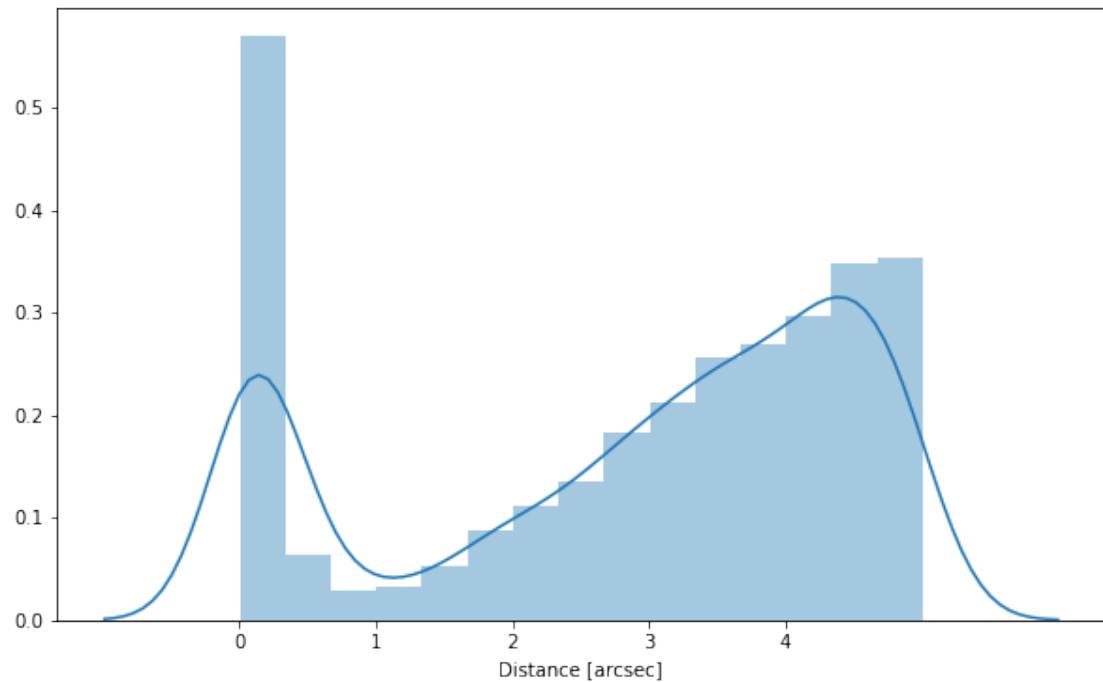
1.2 II - Merging tables

We first merge the optical catalogues and then add the infrared ones. We start with PanSTARRS because it covers the whole field.

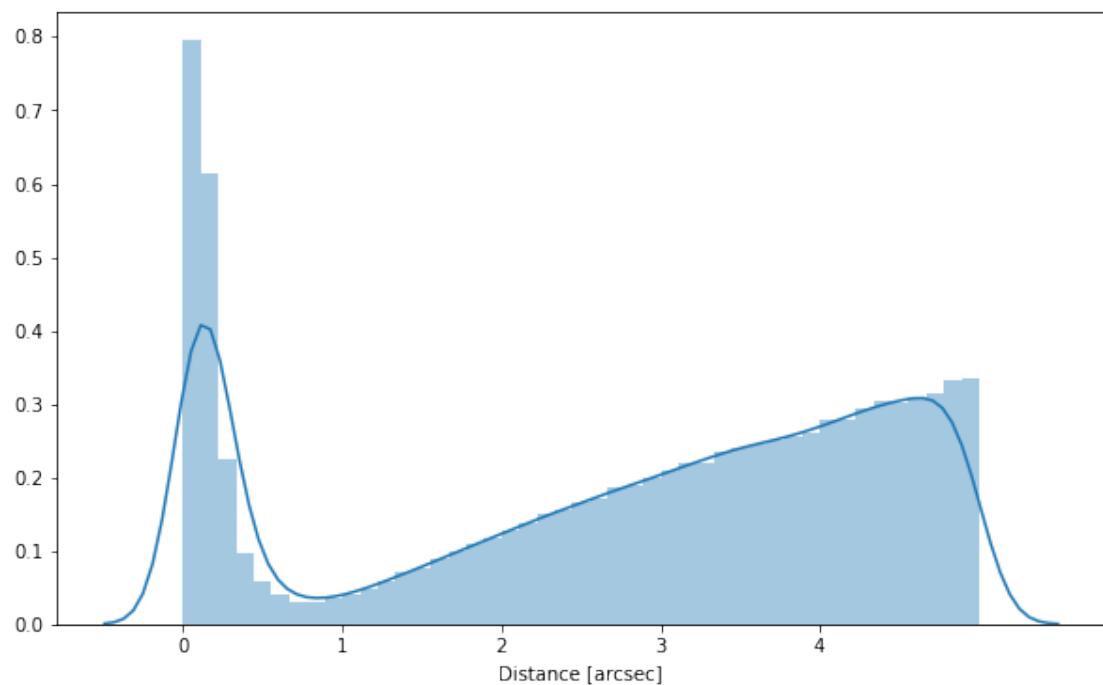
At every step, we look at the distribution of the distances to the nearest source in the merged catalogue to determine the best crossmatching radius.

1.2.1 PanSTARRS

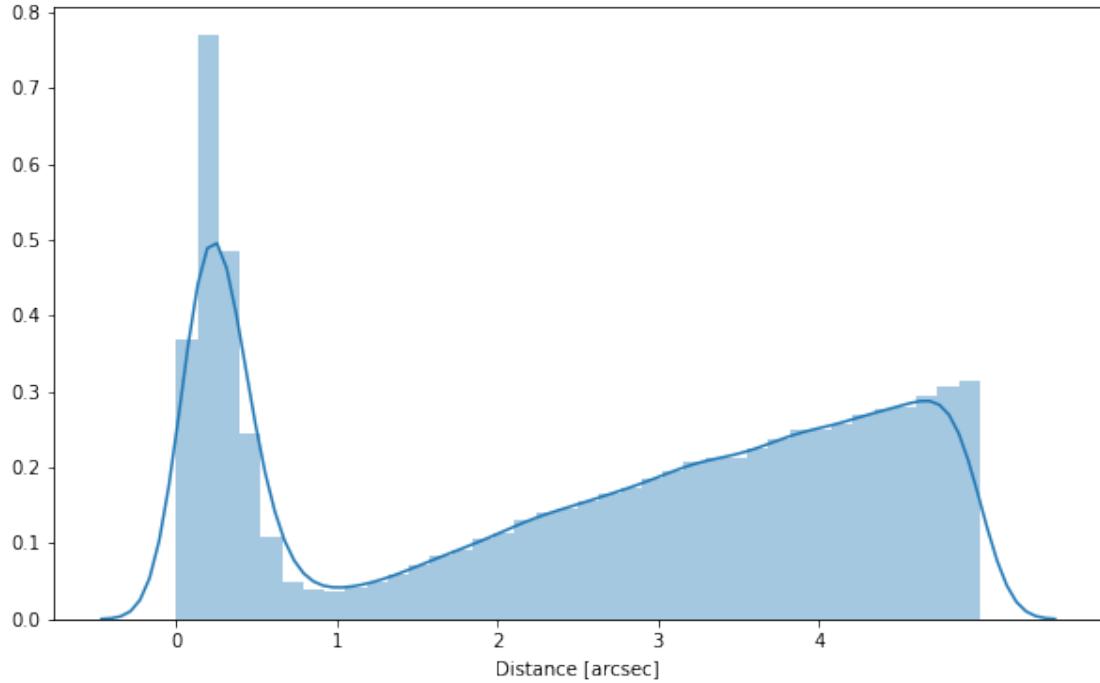
1.2.2 CANDELS-GOODS-N



1.2.3 Ultradeep



1.2.4 Hawaii



1.2.5 Cleaning

When we merge the catalogues, astropy masks the non-existent values (e.g. when a row comes only from a catalogue and has no counterparts in the other, the columns from the latest are masked for that row). We indicate to use NaN for masked values for floats columns, False for flag columns and -1 for ID columns.

Out [13]: <IPython.core.display.HTML object>

1.3 III - Merging flags and stellarity

Each pristine catalogue contains a flag indicating if the source was associated to a another nearby source that was removed during the cleaning process. We merge these flags in a single one.

combining the flag_merged column which contains information regarding multiple associations

Each pristine catalogue contains a flag indicating the probability of a source being a Gaia object (0: not a Gaia object, 1: possibly, 2: probably, 3: definitely). We merge these flags taking the highest value.

Each prisitine catalogue may contain one or several stellarity columns indicating the probability (0 to 1) of each source being a star. We merge these columns taking the highest value. We keep trace of the origin of the stellarity.

```
candels-gn_stellarity
```

1.4 IV - Adding E(B-V) column

1.5 V - Adding HELP unique identifiers and field columns

OK!

1.6 VI - Cross-matching with spec-z catalogue

1.7 VII - Choosing between multiple values for the same filter

We have to choose between the various HST catalogues which may contains different objects depending on the prior catalogue. The CANDELS-GOODS-N catalogue is taken as a base and any missing wircam or IRAC fluxes are taken from the Ultradeep Ks selected catalogues

```
/opt/anaconda3/envs/herschelhelp_internal/lib/python3.6/site-packages/numpy/core/numeric.py:301:  
    format(shape, fill_value, array(fill_value).dtype), FutureWarning)
```

Out[29]: <IPython.core.display.HTML object>

Out[31]: <IPython.core.display.HTML object>

1.8 VIII.a Wavelength domain coverage

We add a binary flag_optnir_obs indicating that a source was observed in a given wavelength domain:

- 1 for observation in optical;
- 2 for observation in near-infrared;
- 4 for observation in mid-infrared (IRAC).

It's an integer binary flag, so a source observed both in optical and near-infrared by not in mid-infrared would have this flag at $1 + 2 = 3$.

Note 1: The observation flag is based on the creation of multi-order coverage maps from the catalogues, this may not be accurate, especially on the edges of the coverage.

Note 2: Being on the observation coverage does not mean having fluxes in that wavelength domain. For sources observed in one domain but having no flux in it, one must take into consideration the different depths in the catalogue we are using.

1.9 VIII.b Wavelength domain detection

We add a binary flag_optnir_det indicating that a source was detected in a given wavelength domain:

- 1 for detection in optical;
- 2 for detection in near-infrared;

- 4 for detection in mid-infrared (IRAC).

It's an integer binary flag, so a source detected both in optical and near-infrared by not in mid-infrared would have this flag at $1 + 2 = 3$.

Note 1: We use the total flux columns to know if the source has flux, in some catalogues, we may have aperture flux and no total flux.

To get rid of artefacts (chip edges, star flares, etc.) we consider that a source is detected in one wavelength domain when it has a flux value in **at least two bands**. That means that good sources will be excluded from this flag when they are on the coverage of only one band.

1.10 IX - Cross-identification table

We are producing a table associating to each HELP identifier, the identifiers of the sources in the pristine catalogues. This can be used to easily get additional information from them.

For convenience, we also cross-match the master list with the SDSS catalogue and add the objID associated with each source, if any. **TODO: should we correct the astrometry with respect to Gaia positions?**

8 master list rows had multiple associations.

```
['ps1_id', 'candels-gn_id', 'ultradeep_id', 'hawaii_id', 'help_id', 'sdss_id']
```

1.11 X - Adding HEALPix index

We are adding a column with a HEALPix index at order 13 associated with each source.

1.12 XI - Saving the catalogue

```
Missing columns: {'flag_acs_f140w', 'flag_acs_f606w', 'flag_gpc1_g', 'flag_acs_f435w', 'flag_ira
```

3_Checks_and_diagnostics

January 18, 2018

1 HDF-N master catalogue

1.1 Checks and diagnostics

This notebook was run with herschelhelp_internal version:
33f5ec7 (Wed Dec 6 16:56:17 2017 +0000)

Diagnostics done using: master_catalogue_hdf-n_20171212.fits

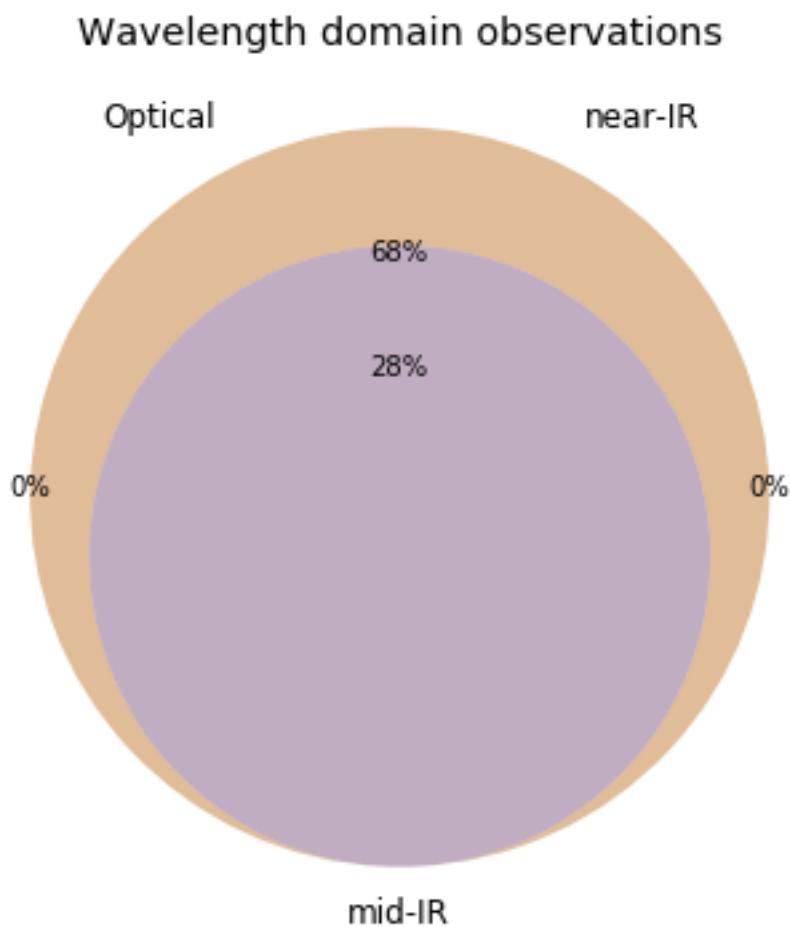
1.2 0 - Quick checks

```
/opt/anaconda3/envs/herschelhelp_internal/lib/python3.6/site-packages/numpy/core/numeric.py:301:
    format(shape, fill_value, array(fill_value).dtype), FutureWarning)
/opt/anaconda3/envs/herschelhelp_internal/lib/python3.6/site-packages/numpy/core/numeric.py:301:
    format(shape, fill_value, array(fill_value).dtype), FutureWarning)
```

Table shows only problematic columns.

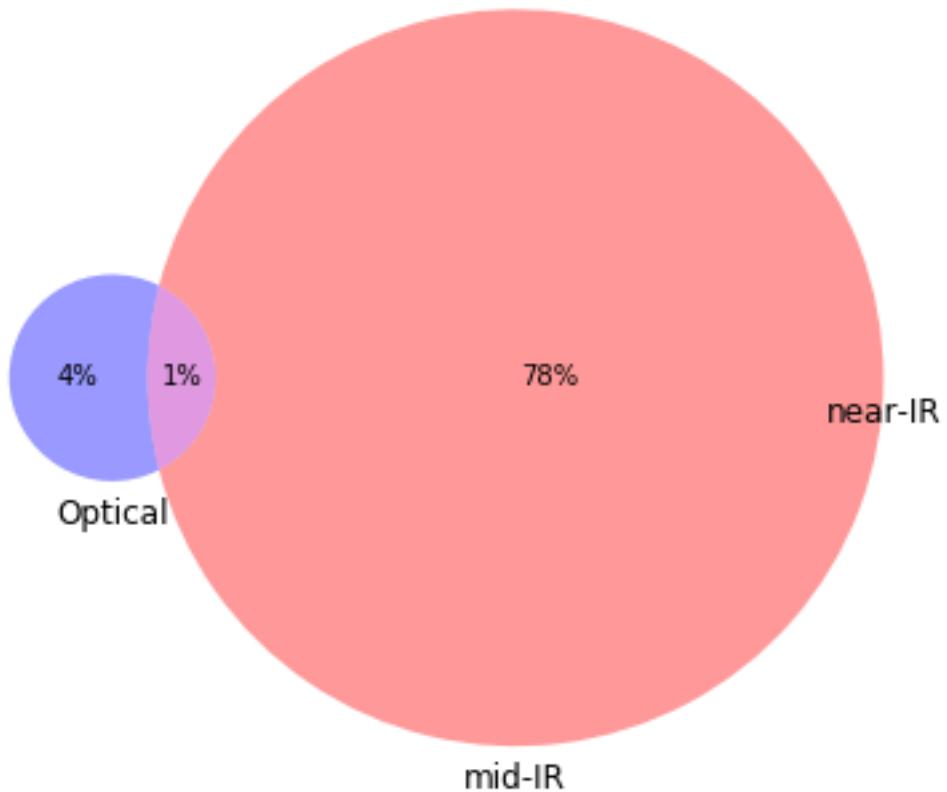
Out[4]: <IPython.core.display.HTML object>

1.3 I - Summary of wavelength domains



```
/opt/anaconda3/envs/herschelhelp_internal/lib/python3.6/site-packages/matplotlib_venn/_venn3.py:  
    warnings.warn("Circle B has zero area")
```

Detection of the 47,176 sources detected
in any wavelength domains (among 130,679 sources)

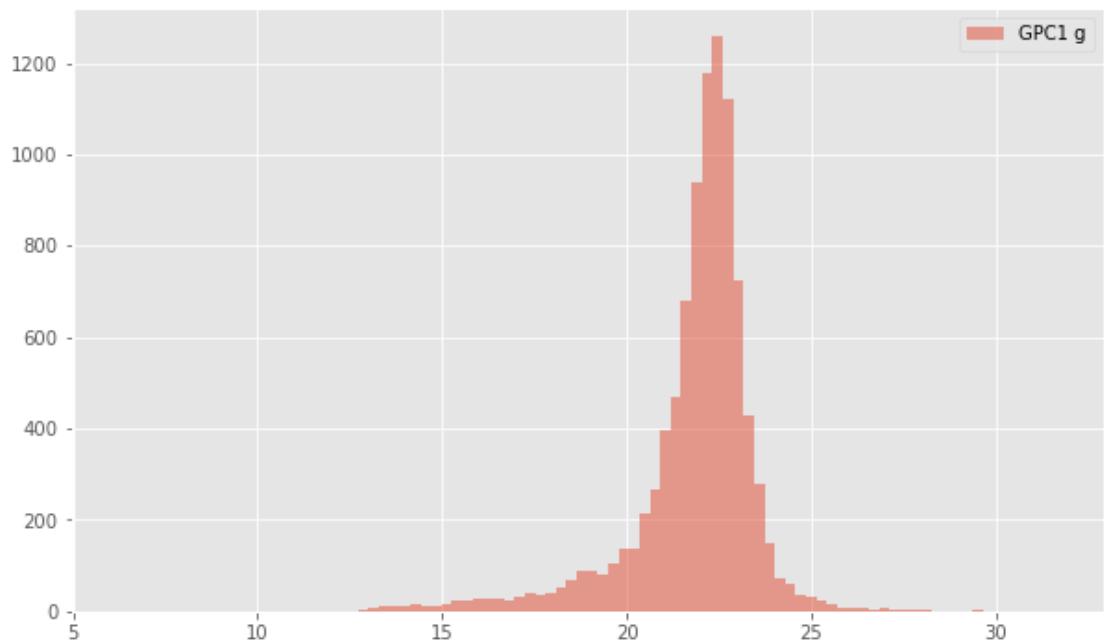
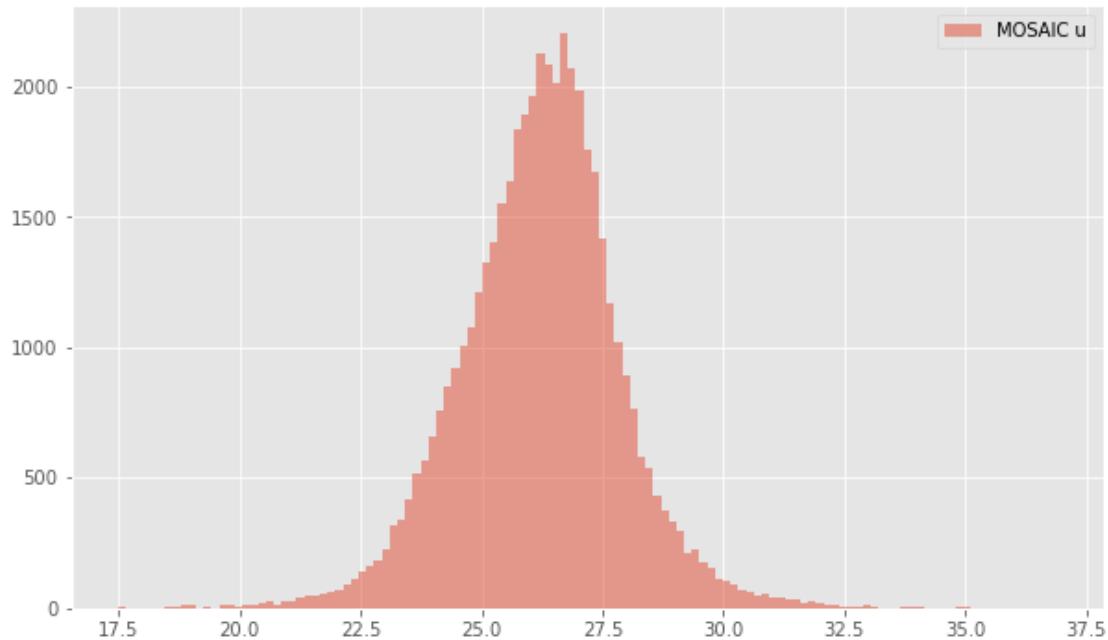


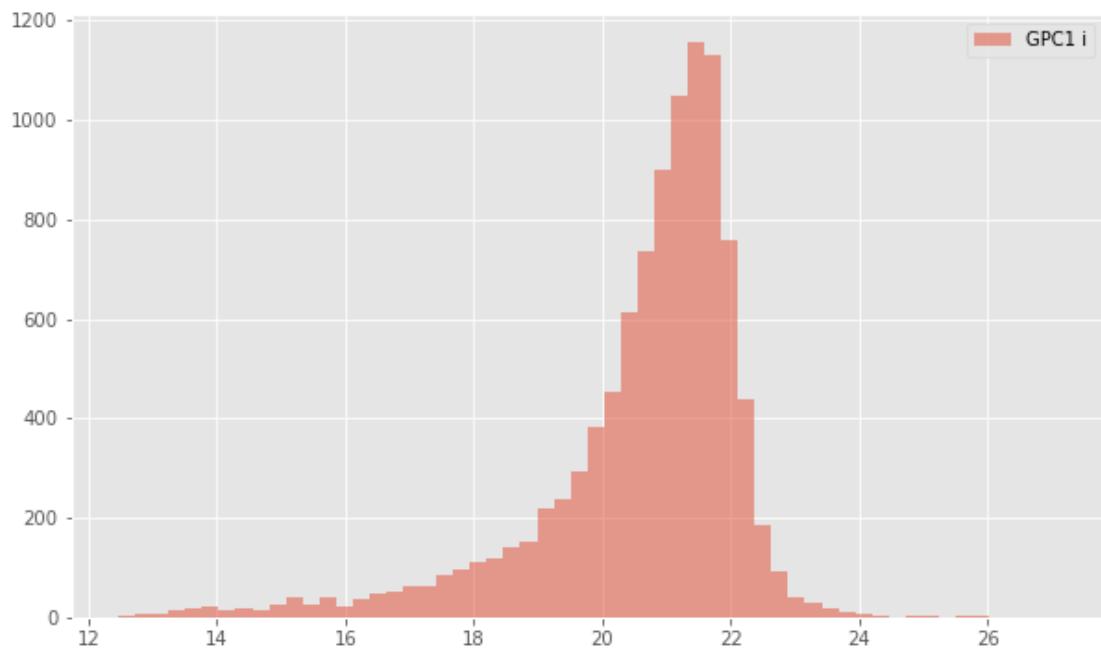
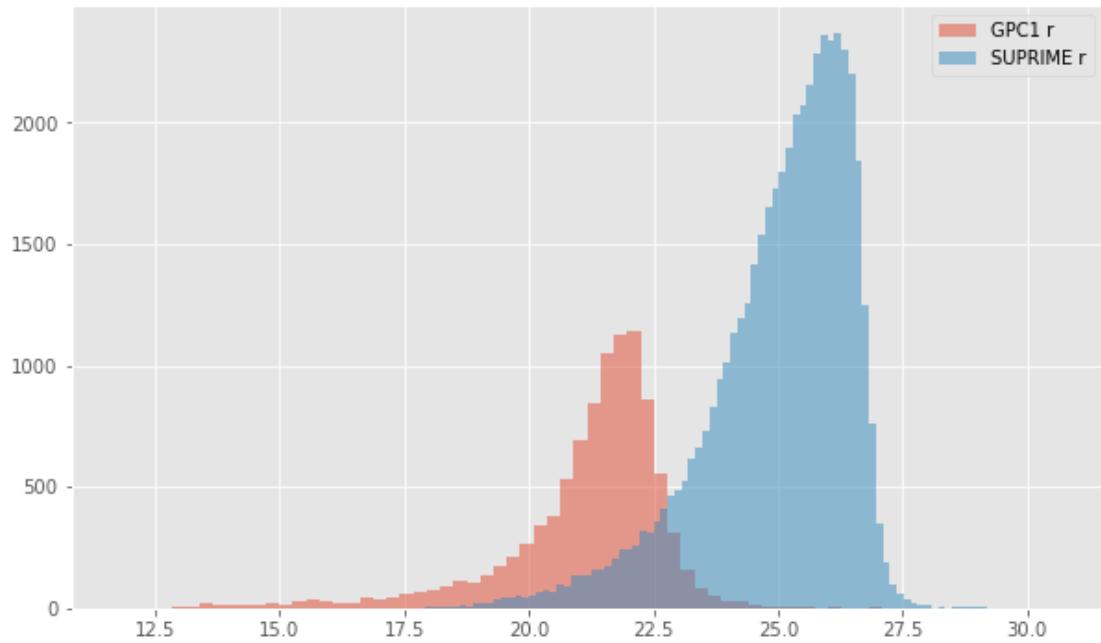
1.4 II - Comparing magnitudes in similar filters

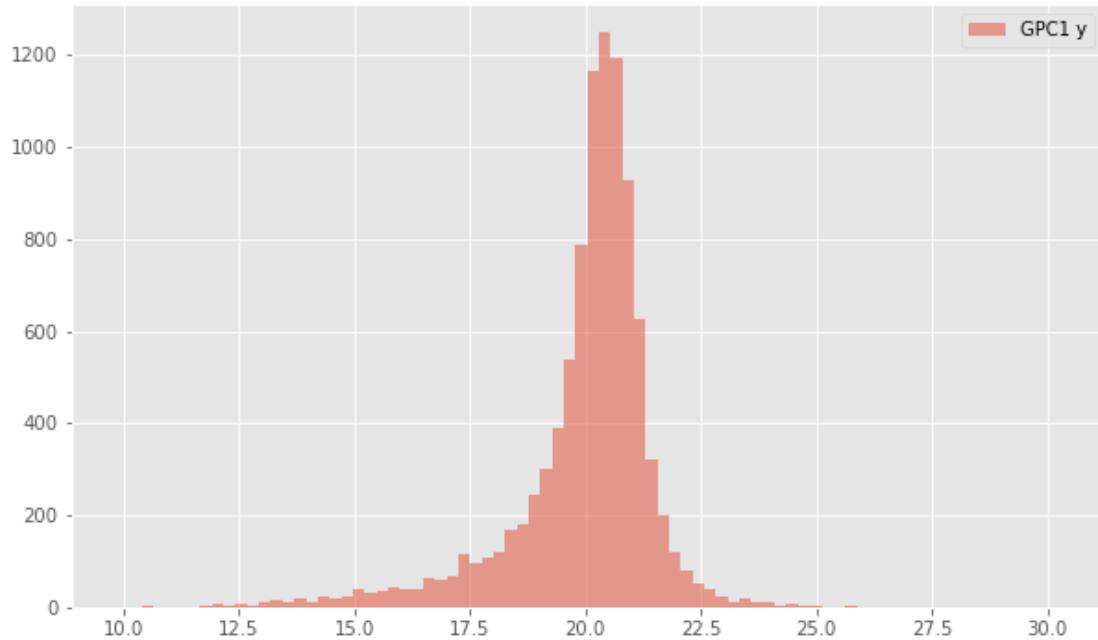
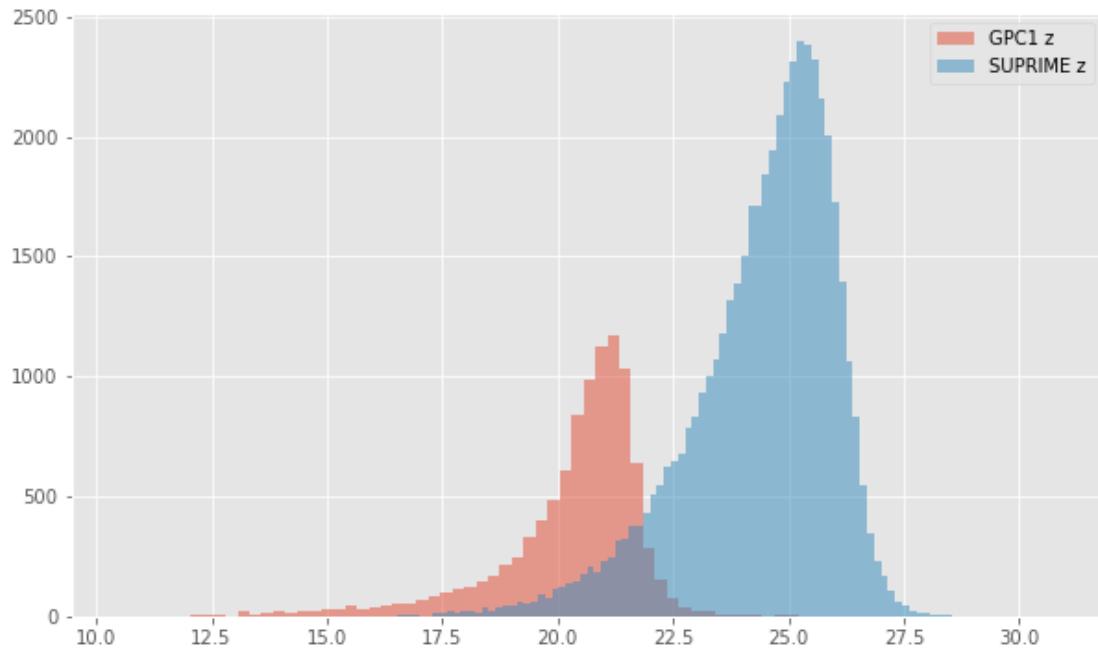
The master list is composed of several catalogues containing magnitudes in similar filters on different instruments. We are comparing the magnitudes in these corresponding filters.

1.4.1 II.a - Comparing depths

We compare the histograms of the total aperture magnitudes of similar bands.





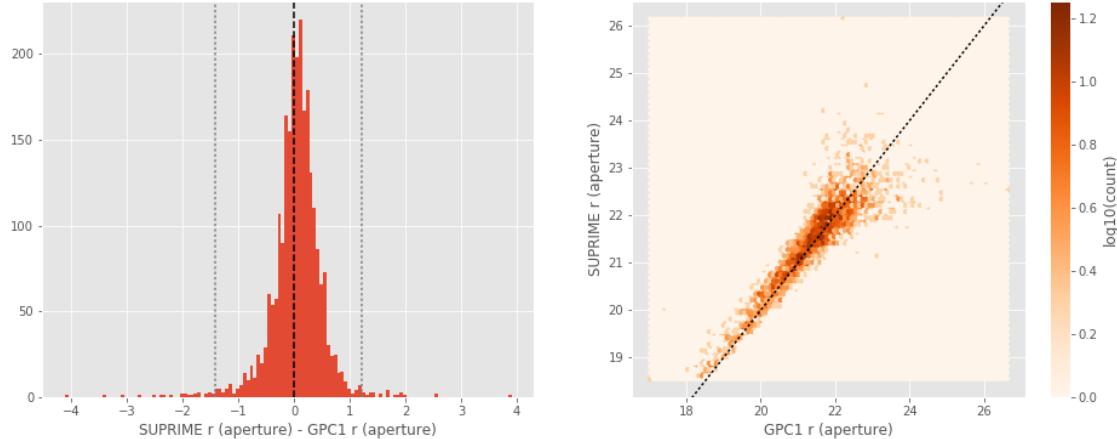


1.4.2 II.b - Comparing magnitudes

We compare one to one each magnitude in similar bands.

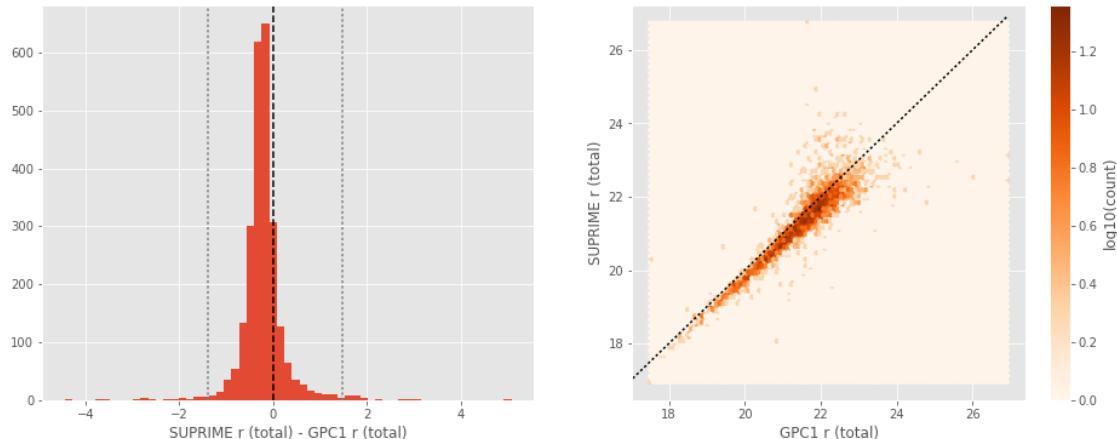
SUPRIME r (aperture) - GPC1 r (aperture):

- Median: 0.07
- Median Absolute Deviation: 0.21
- 1% percentile: -1.424673613403319
- 99% percentile: 1.2040346905273431



SUPRIME r (total) - GPC1 r (total):

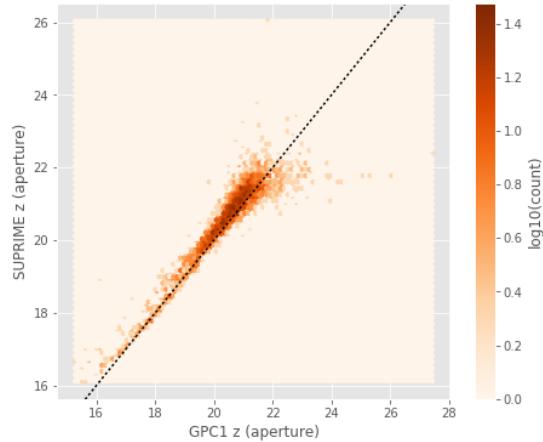
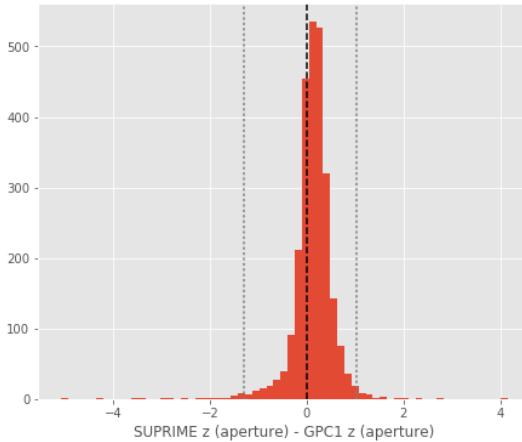
- Median: -0.24
- Median Absolute Deviation: 0.16
- 1% percentile: -1.4006093564038085
- 99% percentile: 1.4649504818432544



SUPRIME z (aperture) - GPC1 z (aperture):

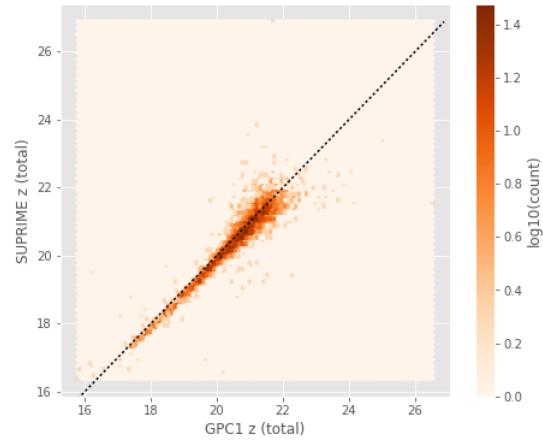
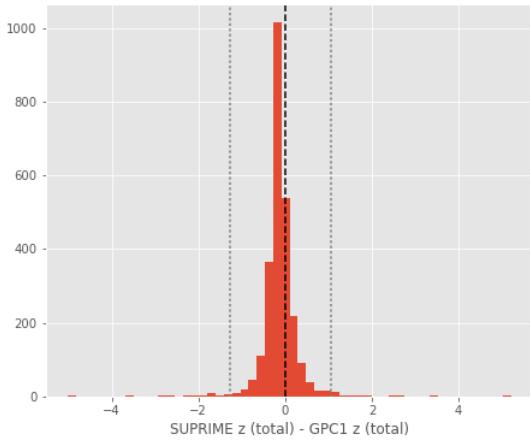
- Median: 0.15

- Median Absolute Deviation: 0.18
- 1% percentile: -1.3092136831054682
- 99% percentile: 1.036628628840331



SUPRIME z (total) - GPC1 z (total):

- Median: -0.14
- Median Absolute Deviation: 0.13
- 1% percentile: -1.280440952136231
- 99% percentile: 1.0462272840478488



1.5 III - Comparing magnitudes to reference bands

Cross-match the master list to SDSS and 2MASS to compare its magnitudes to SDSS and 2MASS ones.

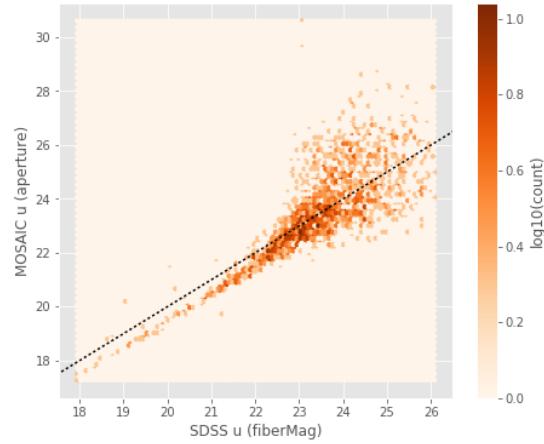
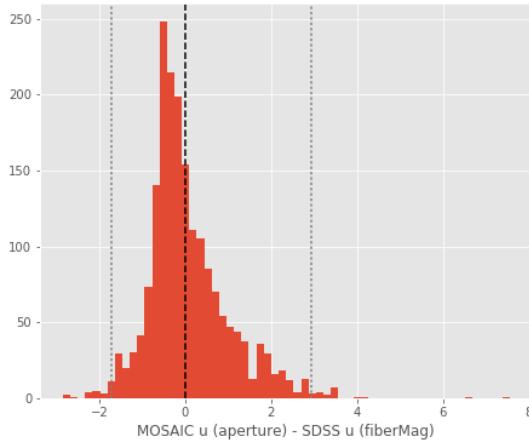
1.5.1 III.a - Comparing u, g, r, i, and z bands to SDSS

The catalogue is cross-matched to SDSS-DR13 withing 0.2 arcsecond.

We compare the u, g, r, i, and z magnitudes to those from SDSS using fiberMag for the aperture magnitude and petroMag for the total magnitude.

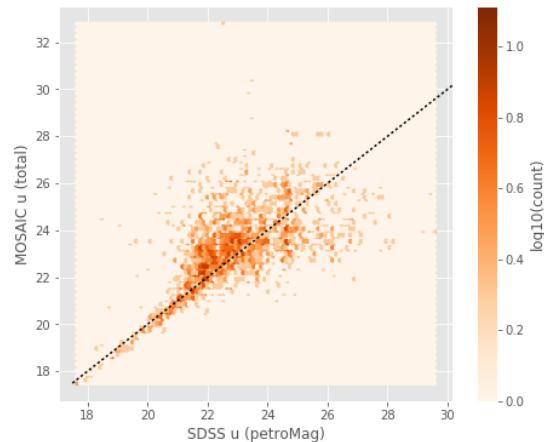
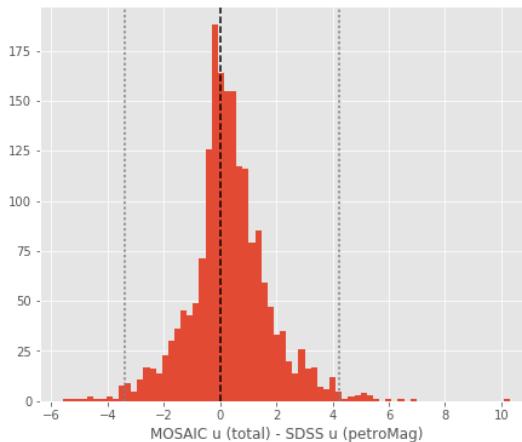
MOSAIC u (aperture) - SDSS u (fiberMag) :

- Median: -0.15
- Median Absolute Deviation: 0.44
- 1% percentile: -1.7168668045751936
- 99% percentile: 2.942735320161134



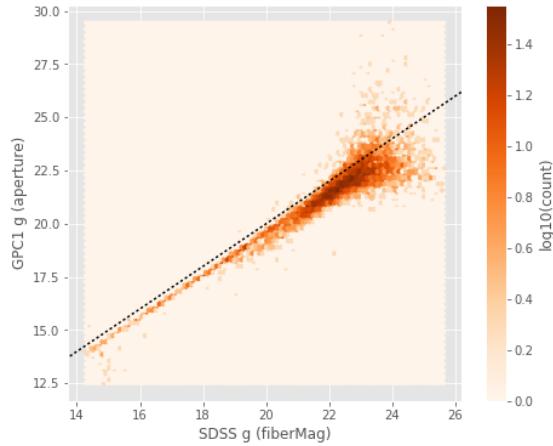
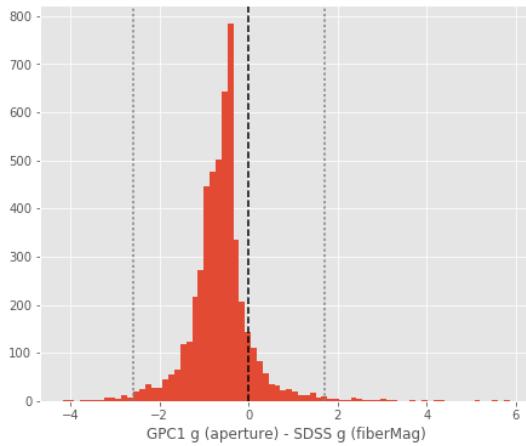
MOSAIC u (total) - SDSS u (petroMag) :

- Median: 0.26
- Median Absolute Deviation: 0.67
- 1% percentile: -3.3910033163061515
- 99% percentile: 4.251866573155514



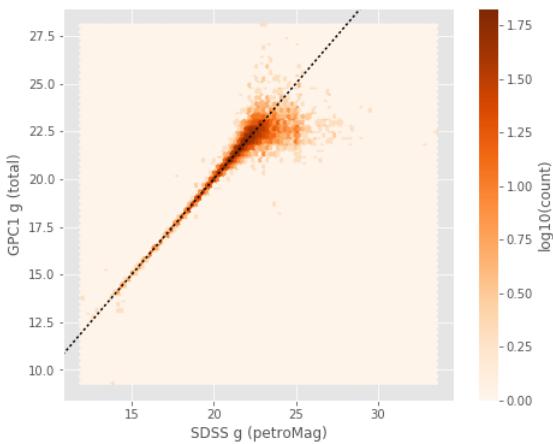
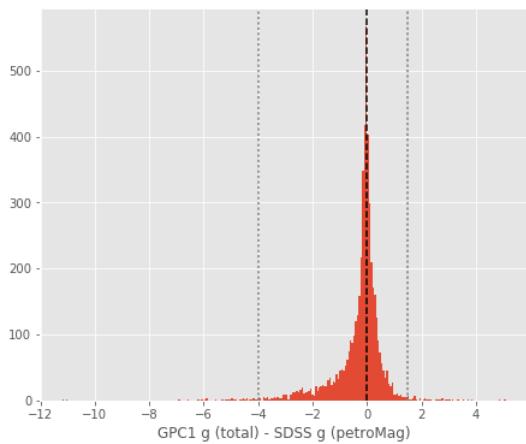
GPC1 g (aperture) - SDSS g (fiberMag):

- Median: -0.60
- Median Absolute Deviation: 0.29
- 1% percentile: -2.5760694885253907
- 99% percentile: 1.7048368072509783



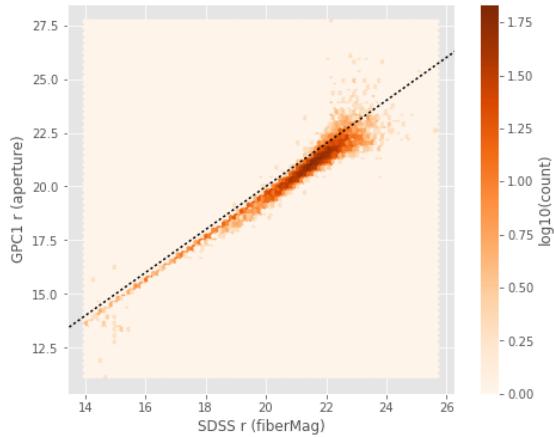
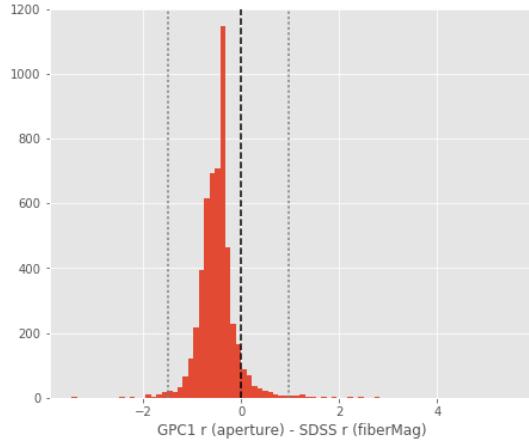
GPC1 g (total) - SDSS g (petroMag):

- Median: -0.08
- Median Absolute Deviation: 0.25
- 1% percentile: -3.9923860931396487
- 99% percentile: 1.4683370971679683



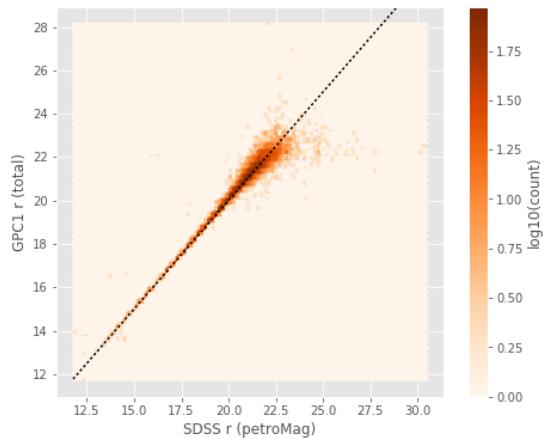
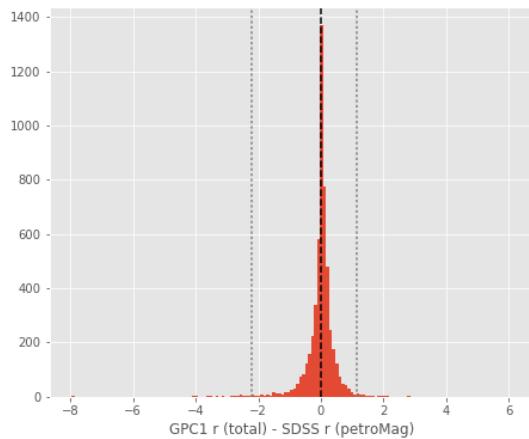
GPC1 r (aperture) - SDSS r (fiberMag):

- Median: -0.47
- Median Absolute Deviation: 0.17
- 1% percentile: -1.4966218566894531
- 99% percentile: 0.9670242309570285



GPC1 r (total) - SDSS r (petroMag):

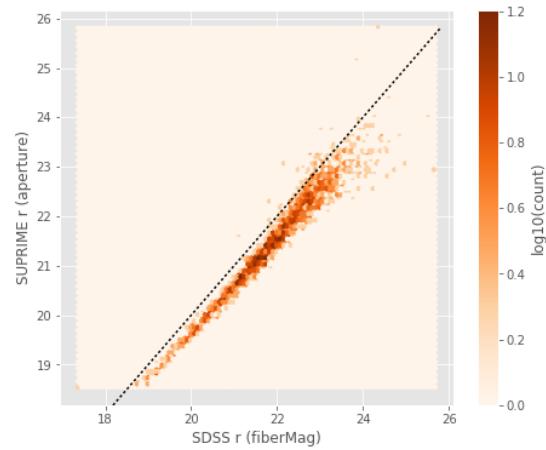
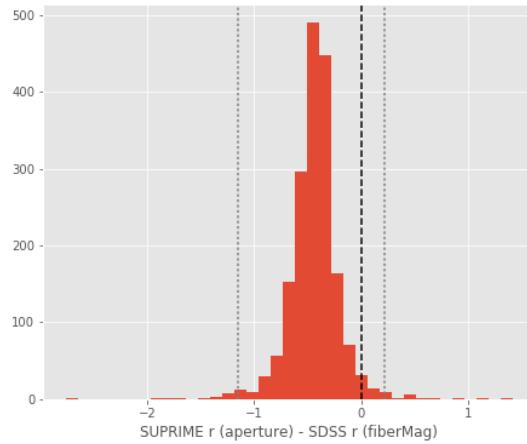
- Median: 0.03
- Median Absolute Deviation: 0.14
- 1% percentile: -2.213793106079102
- 99% percentile: 1.1340791320800772



SUPRIME r (aperture) - SDSS r (fiberMag):

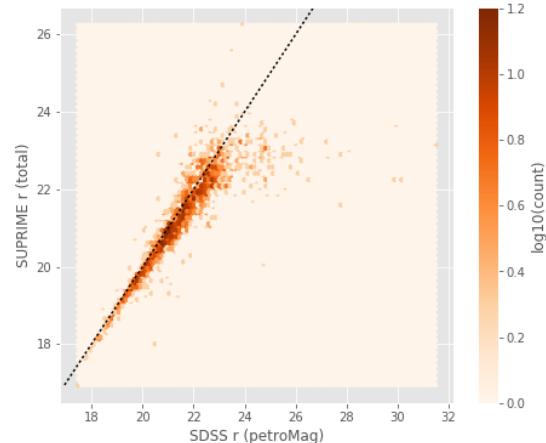
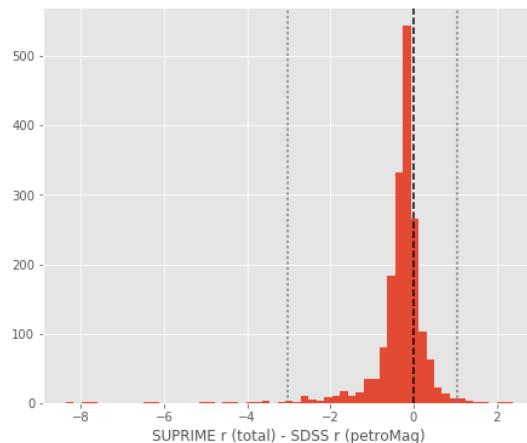
- Median: -0.42

- Median Absolute Deviation: 0.11
- 1% percentile: -1.1524517989306626
- 99% percentile: 0.2232965563378855



SUPRIME r (total) - SDSS r (petroMag):

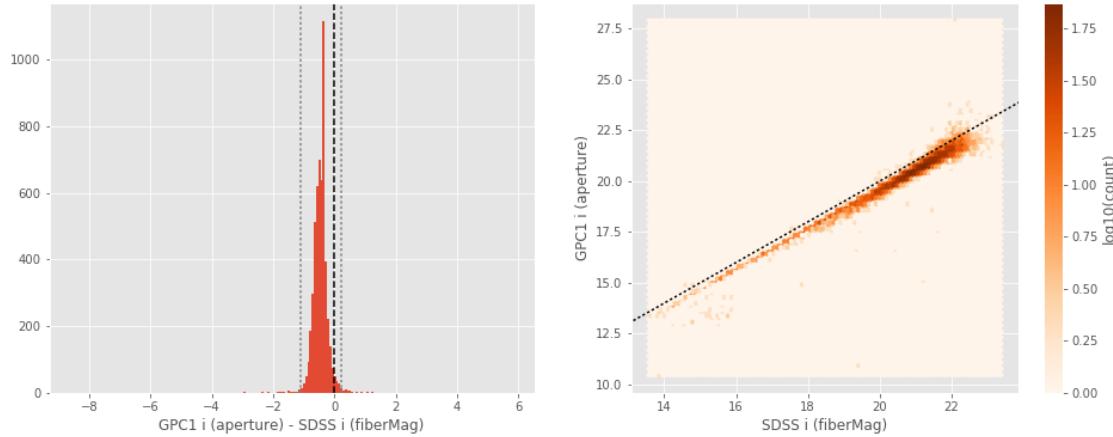
- Median: -0.20
- Median Absolute Deviation: 0.19
- 1% percentile: -3.034374487626954
- 99% percentile: 1.0472558652038568



GPC1 i (aperture) - SDSS i (fiberMag):

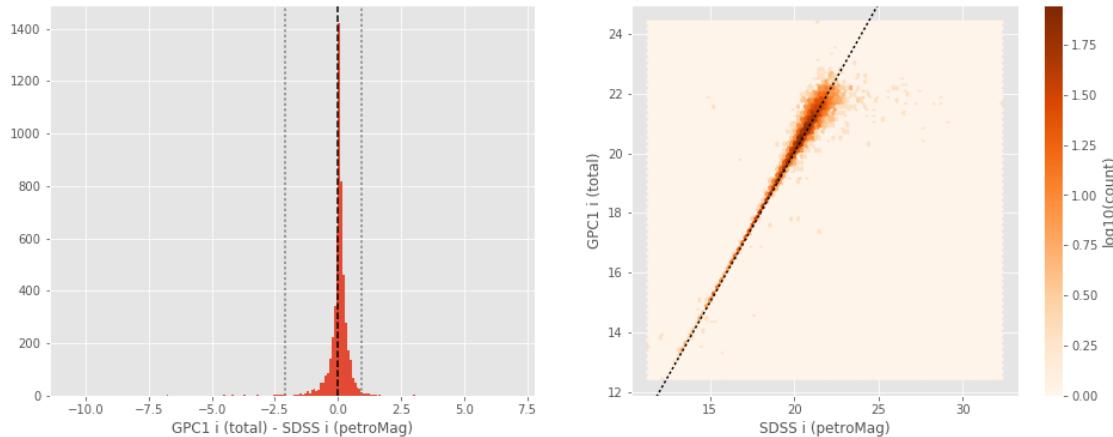
- Median: -0.44
- Median Absolute Deviation: 0.12

- 1% percentile: -1.1171233367919922
- 99% percentile: 0.2286124420166025



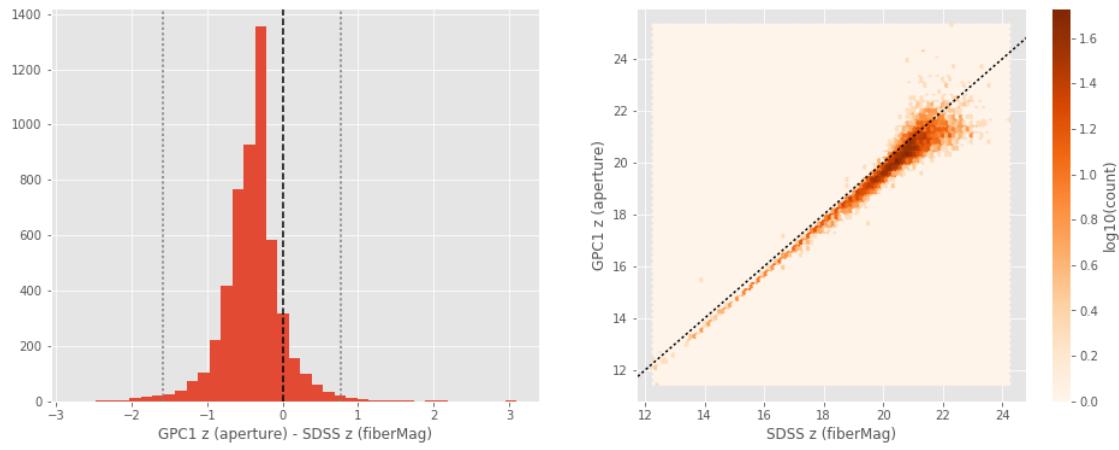
GPC1 i (total) - SDSS i (petroMag):

- Median: 0.06
- Median Absolute Deviation: 0.12
- 1% percentile: -2.0888433837890625
- 99% percentile: 0.9435255050659161



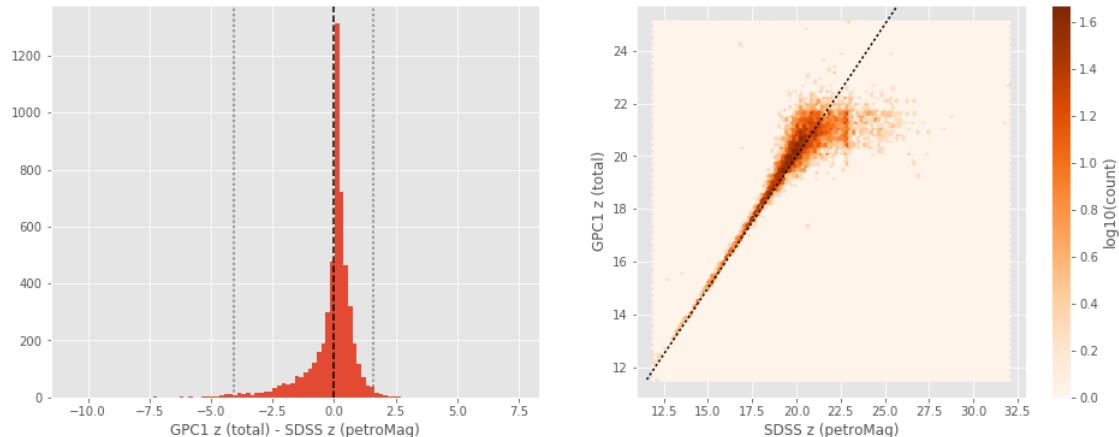
GPC1 z (aperture) - SDSS z (fiberMag):

- Median: -0.36
- Median Absolute Deviation: 0.18
- 1% percentile: -1.5850920677185059
- 99% percentile: 0.7624429321289085



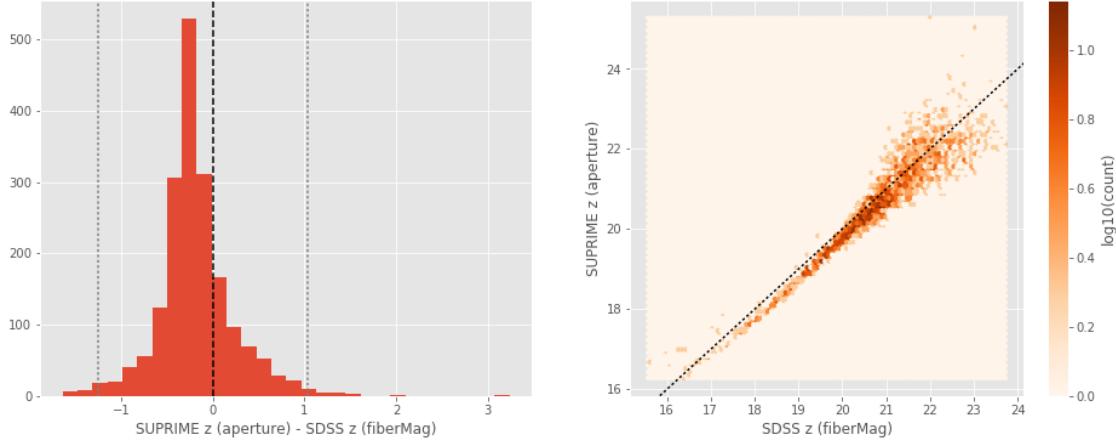
GPC1 z (total) - SDSS z (petroMag):

- Median: 0.10
- Median Absolute Deviation: 0.30
- 1% percentile: -4.0788907623291015
- 99% percentile: 1.606949920654297



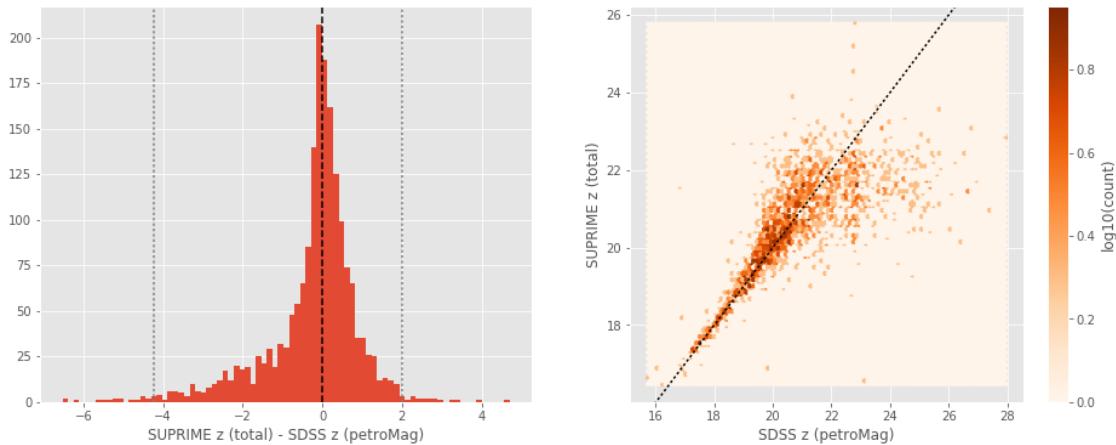
SUPRIME z (aperture) - SDSS z (fiberMag):

- Median: -0.23
- Median Absolute Deviation: 0.17
- 1% percentile: -1.2519832500915538
- 99% percentile: 1.0297534400939927



SUPRIME z (total) - SDSS z (petroMag):

- Median: -0.02
- Median Absolute Deviation: 0.43
- 1% percentile: -4.218913514050293
- 99% percentile: 2.0048278384594718



1.5.2 III.b - Comparing J and K bands to 2MASS

The catalogue is cross-matched to 2MASS-PSC within 0.2 arcsecond. We compare the UKIDSS total J and K magnitudes to those from 2MASS.

The 2MASS magnitudes are “Vega-like” and we have to convert them to AB magnitudes using the zero points provided on [this page](#):

Band	F - 0 mag (Jy)
J	1594
H	1024
Ks	666.7

In addition, UKIDSS uses a K band whereas 2MASS uses a Ks (“short”) band, [this page](#) give a correction to convert the K band in a Ks band with the formula:

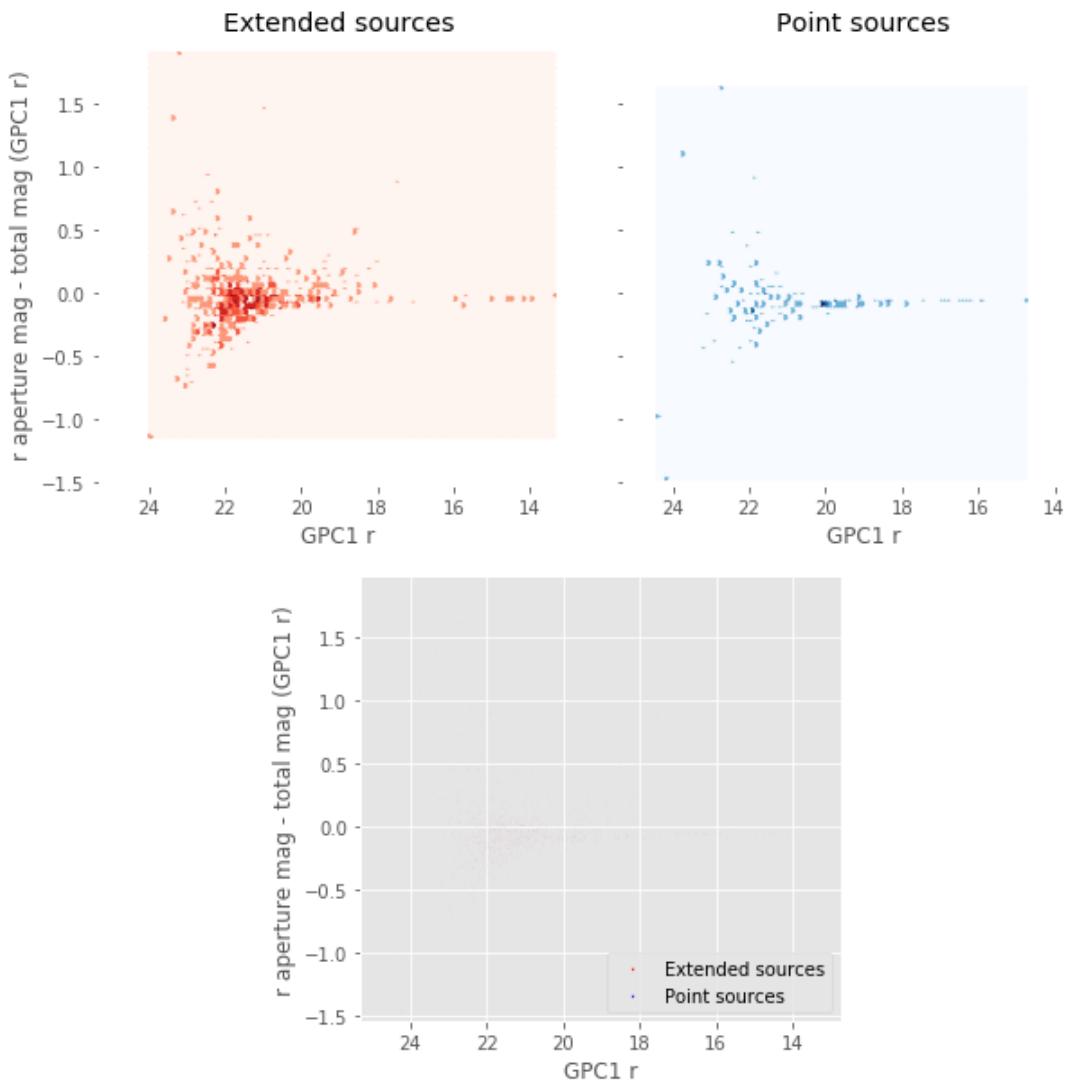
$$K_{s(2MASS)} = K_{UKIRT} + 0.003 + 0.004 * (JK)_{UKIRT}$$

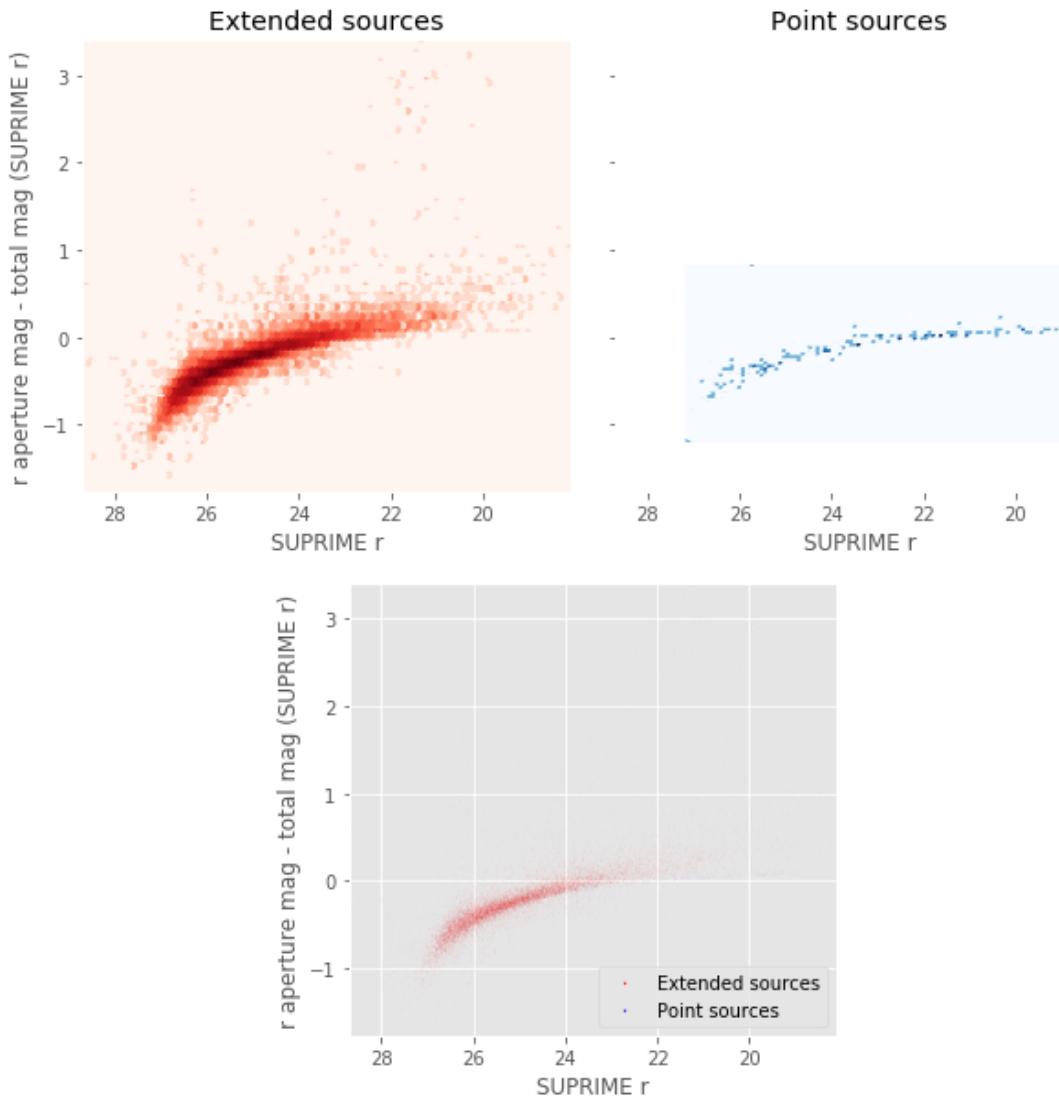
No sources have both 2MASS Ks and WIRCAM Ks (total) values.

1.6 IV - Comparing aperture magnitudes to total ones.

Number of source used: 682 / 130679 (0.52%)

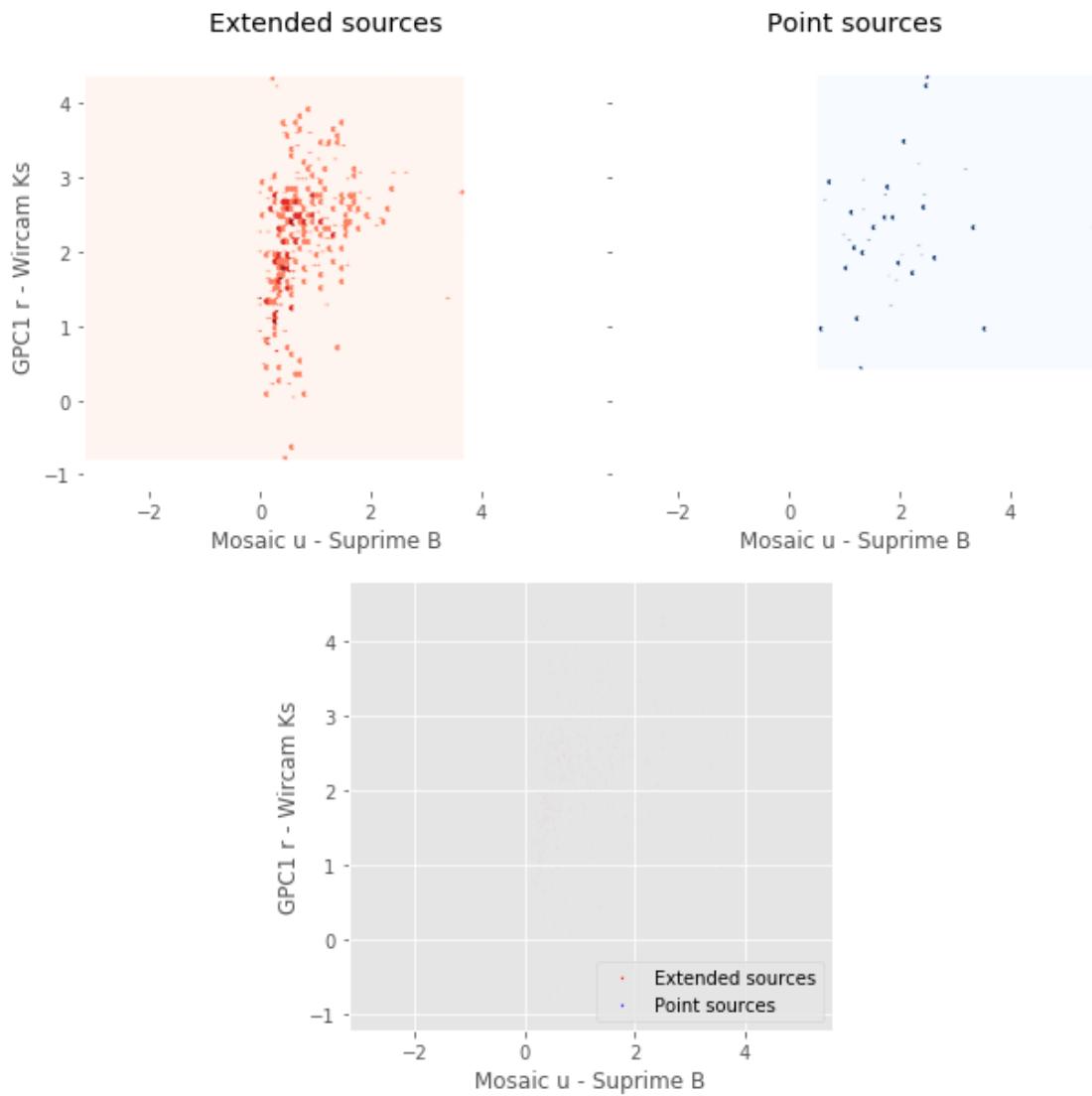
Number of source used: 12017 / 130679 (9.20%)



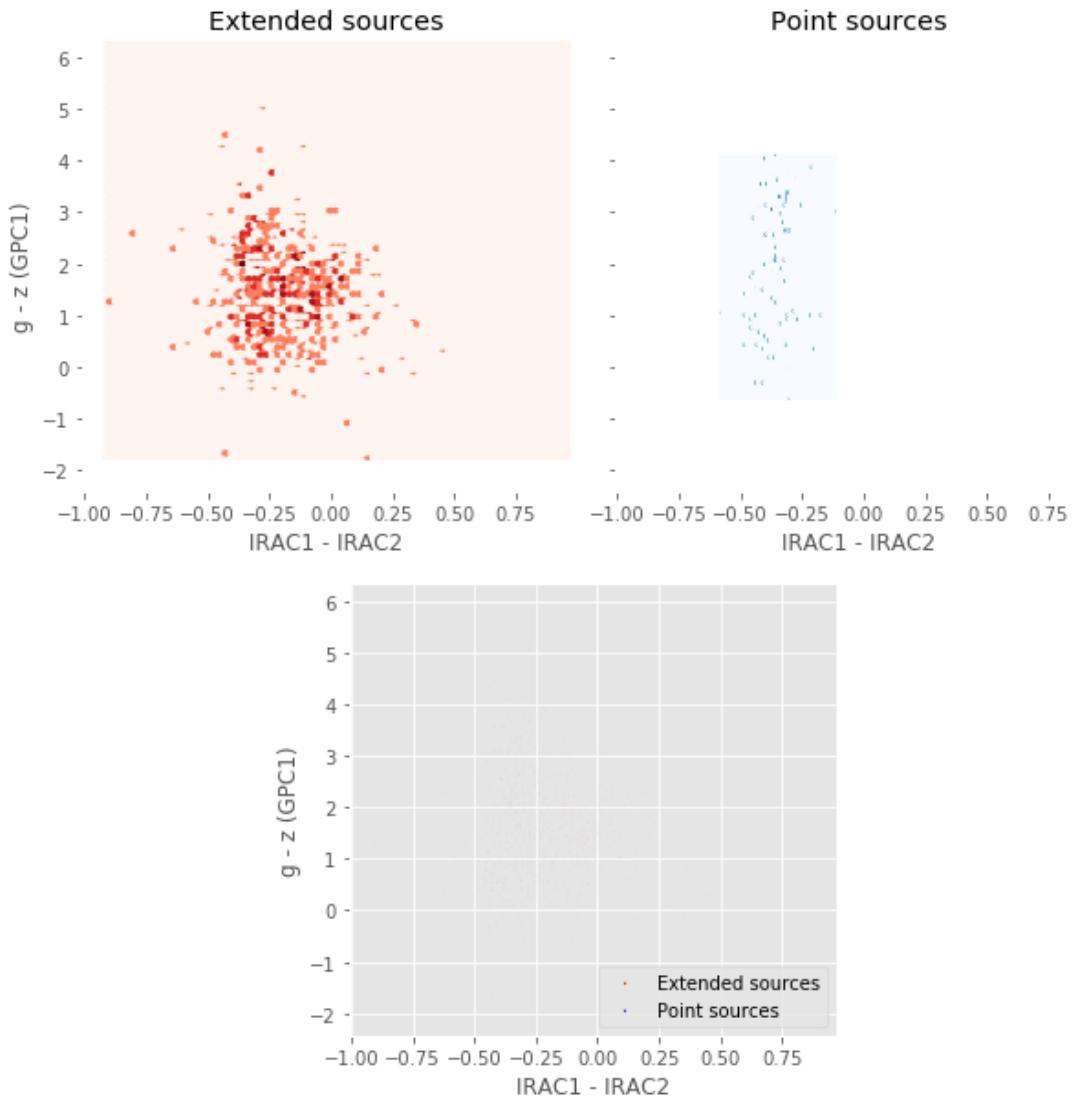


1.7 V - Color-color and magnitude-color plots

Number of source used: 374 / 130679 (0.29%)



Number of source used: 640 / 130679 (0.49%)



2_Merging

January 18, 2018

1 HDF-N master catalogue

This notebook presents the merge of the various pristine catalogues to produce the HELP master catalogue on HDF-N.

This notebook was run with herschelhelp_internal version:
33f5ec7 (Wed Dec 6 16:56:17 2017 +0000)

1.1 I - Reading the prepared pristine catalogues

WARNING: UnitsWarning: '0 galaxy, 1 star' did not parse as fits unit: At col 2, Unit 'galaxy' no

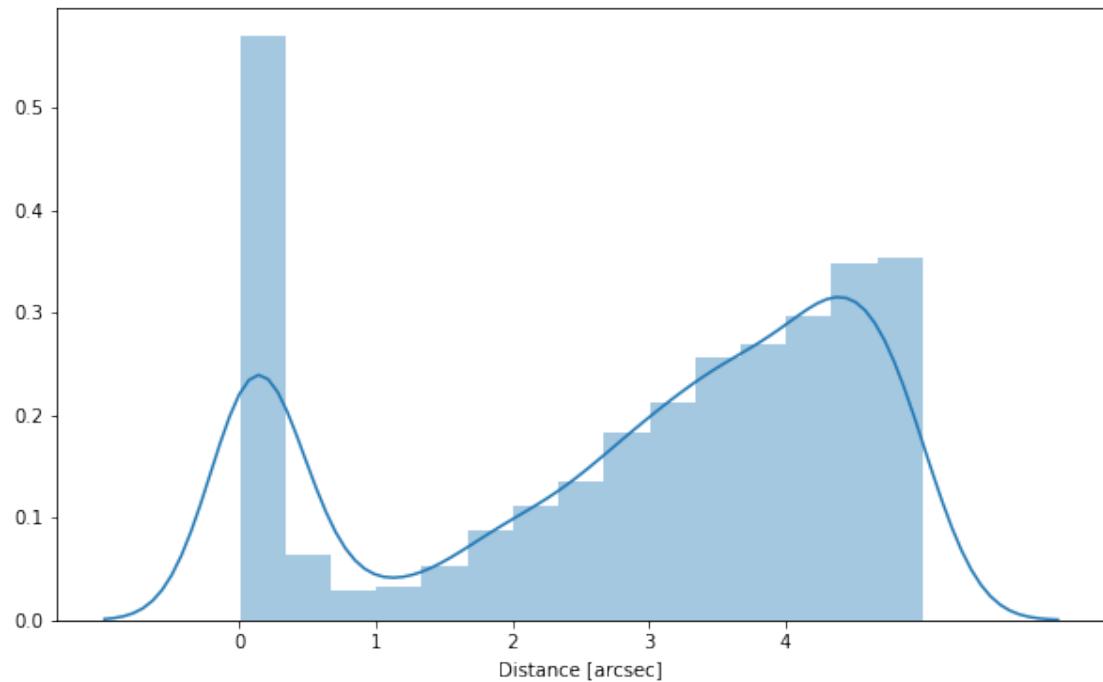
1.2 II - Merging tables

We first merge the optical catalogues and then add the infrared ones. We start with PanSTARRS because it covers the whole field.

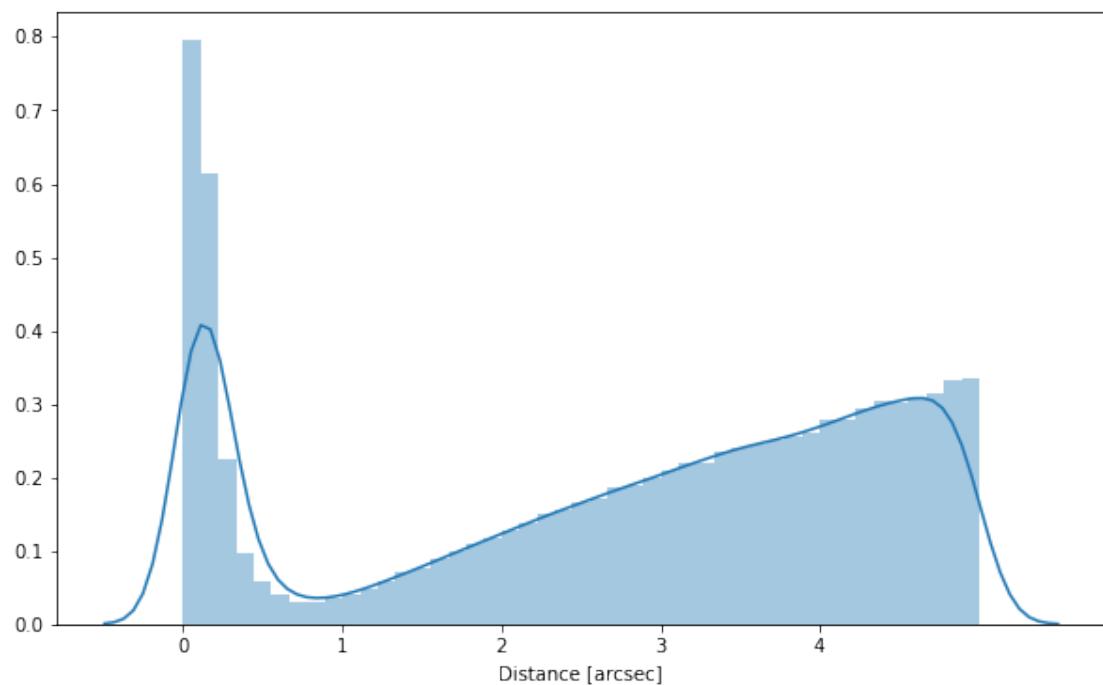
At every step, we look at the distribution of the distances to the nearest source in the merged catalogue to determine the best crossmatching radius.

1.2.1 PanSTARRS

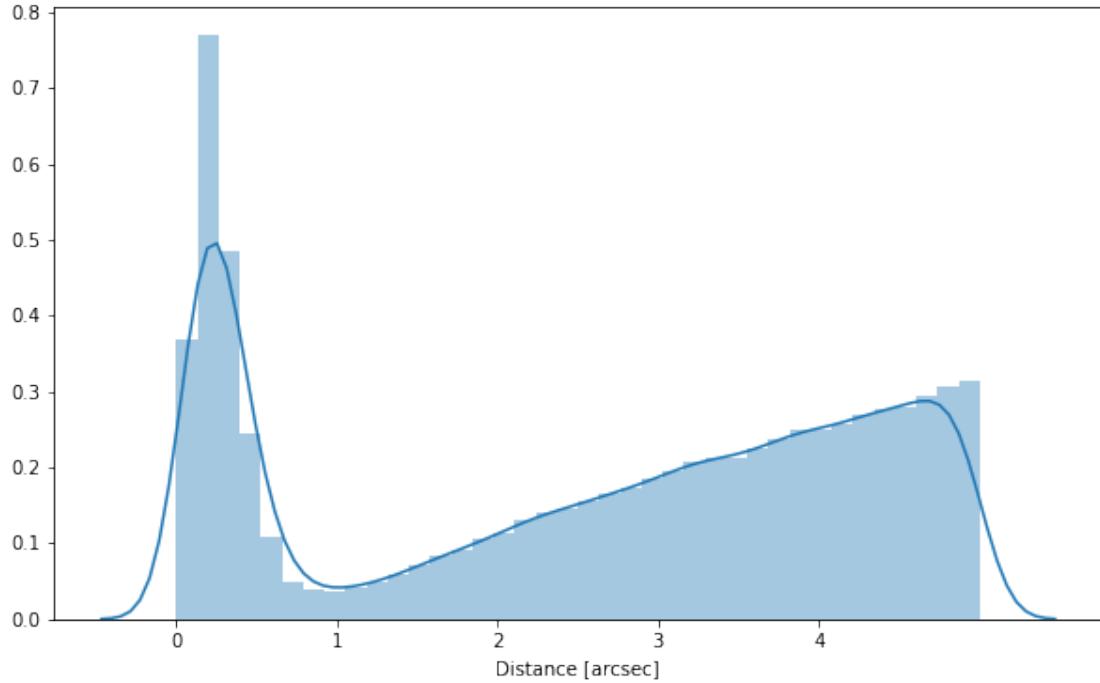
1.2.2 CANDELS-GOODS-N



1.2.3 Ultradeep



1.2.4 Hawaii



1.2.5 Cleaning

When we merge the catalogues, astropy masks the non-existent values (e.g. when a row comes only from a catalogue and has no counterparts in the other, the columns from the latest are masked for that row). We indicate to use NaN for masked values for floats columns, False for flag columns and -1 for ID columns.

Out [13]: <IPython.core.display.HTML object>

1.3 III - Merging flags and stellarity

Each pristine catalogue contains a flag indicating if the source was associated to a another nearby source that was removed during the cleaning process. We merge these flags in a single one.

combining the flag_merged column which contains information regarding multiple associations

Each pristine catalogue contains a flag indicating the probability of a source being a Gaia object (0: not a Gaia object, 1: possibly, 2: probably, 3: definitely). We merge these flags taking the highest value.

Each prisitine catalogue may contain one or several stellarity columns indicating the probability (0 to 1) of each source being a star. We merge these columns taking the highest value. We keep trace of the origin of the stellarity.

```
candels-gn_stellarity
```

1.4 IV - Adding E(B-V) column

1.5 V - Adding HELP unique identifiers and field columns

OK!

1.6 VI - Cross-matching with spec-z catalogue

1.7 VII - Choosing between multiple values for the same filter

We have to choose between the various HST catalogues which may contains different objects depending on the prior catalogue. The CANDELS-GOODS-N catalogue is taken as a base and any missing wircam or IRAC fluxes are taken from the Ultradeep Ks selected catalogues

```
/opt/anaconda3/envs/herschelhelp_internal/lib/python3.6/site-packages/numpy/core/numeric.py:301:  
    format(shape, fill_value, array(fill_value).dtype), FutureWarning)
```

Out[29]: <IPython.core.display.HTML object>

Out[31]: <IPython.core.display.HTML object>

1.8 VIII.a Wavelength domain coverage

We add a binary flag_optnir_obs indicating that a source was observed in a given wavelength domain:

- 1 for observation in optical;
- 2 for observation in near-infrared;
- 4 for observation in mid-infrared (IRAC).

It's an integer binary flag, so a source observed both in optical and near-infrared by not in mid-infrared would have this flag at $1 + 2 = 3$.

Note 1: The observation flag is based on the creation of multi-order coverage maps from the catalogues, this may not be accurate, especially on the edges of the coverage.

Note 2: Being on the observation coverage does not mean having fluxes in that wavelength domain. For sources observed in one domain but having no flux in it, one must take into consideration the different depths in the catalogue we are using.

1.9 VIII.b Wavelength domain detection

We add a binary flag_optnir_det indicating that a source was detected in a given wavelength domain:

- 1 for detection in optical;
- 2 for detection in near-infrared;

- 4 for detection in mid-infrared (IRAC).

It's an integer binary flag, so a source detected both in optical and near-infrared by not in mid-infrared would have this flag at $1 + 2 = 3$.

Note 1: We use the total flux columns to know if the source has flux, in some catalogues, we may have aperture flux and no total flux.

To get rid of artefacts (chip edges, star flares, etc.) we consider that a source is detected in one wavelength domain when it has a flux value in **at least two bands**. That means that good sources will be excluded from this flag when they are on the coverage of only one band.

1.10 IX - Cross-identification table

We are producing a table associating to each HELP identifier, the identifiers of the sources in the pristine catalogues. This can be used to easily get additional information from them.

For convenience, we also cross-match the master list with the SDSS catalogue and add the objID associated with each source, if any. **TODO: should we correct the astrometry with respect to Gaia positions?**

8 master list rows had multiple associations.

```
['ps1_id', 'candels-gn_id', 'ultradeep_id', 'hawaii_id', 'help_id', 'sdss_id']
```

1.11 X - Adding HEALPix index

We are adding a column with a HEALPix index at order 13 associated with each source.

1.12 XI - Saving the catalogue

```
Missing columns: {'flag_acs_f140w', 'flag_acs_f606w', 'flag_gpc1_g', 'flag_acs_f435w', 'flag_ira
```

1.6_CANDELS-GOODS-N

January 18, 2018

1 HDF-N master catalogue

1.1 Preparation of CANDELS-GOODS-N data

CANDELS-GOODS-N catalogue: the catalogue comes from `dmu0_CANDELS-GOODS-N`.

In the catalogue, we keep:

- The identifier (it's unique in the catalogue);
- The position;
- The stellarity;
- The total magnitude.

We don't know when the maps have been observed. We will use the year of the reference paper.

This notebook was run with `herschelhelp_internal` version:
33f5ec7 (Wed Dec 6 16:56:17 2017 +0000)

1.2 I - Column selection

```
WARNING: UnitsWarning: '0 use, gt 1 no-use' did not parse as fits unit: At col 2, Unit 'use' not
WARNING: UnitsWarning: '0 galaxy, 1 star' did not parse as fits unit: At col 2, Unit 'galaxy' no
```

```
/opt/herschelhelp_internal/herschelhelp_internal/utils.py:76: RuntimeWarning: divide by zero enco
    magnitudes = 2.5 * (23 - np.log10(fluxes)) - 48.6
/opt/herschelhelp_internal/herschelhelp_internal/utils.py:76: RuntimeWarning: invalid value enco
    magnitudes = 2.5 * (23 - np.log10(fluxes)) - 48.6
/opt/herschelhelp_internal/herschelhelp_internal/utils.py:80: RuntimeWarning: divide by zero enco
    errors = 2.5 / np.log(10) * errors_on_fluxes / fluxes
```

Out [6]: <IPython.core.display.HTML object>

1.3 II - Removal of duplicated sources

We remove duplicated objects from the input catalogues.

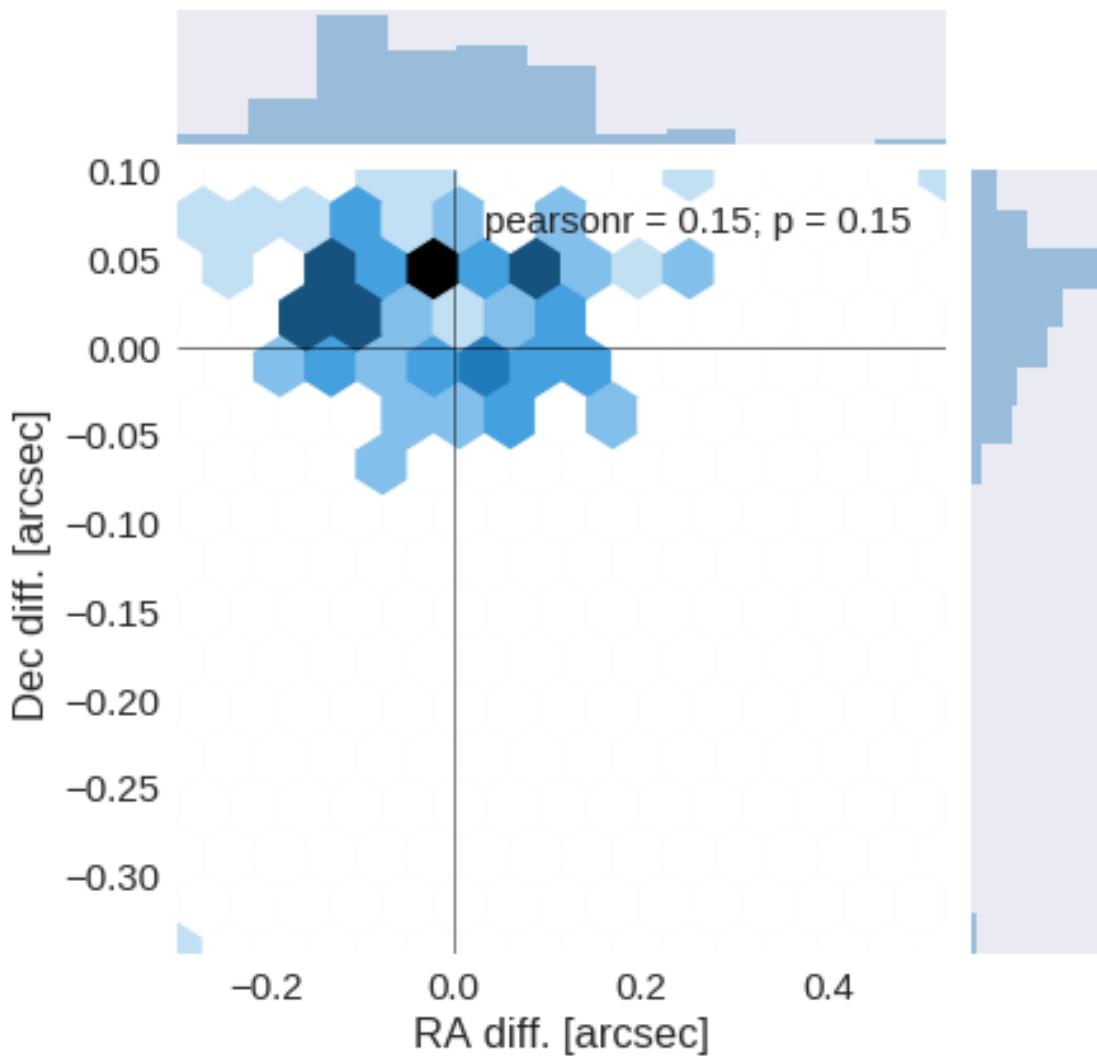
The initial catalogue had 35445 sources.

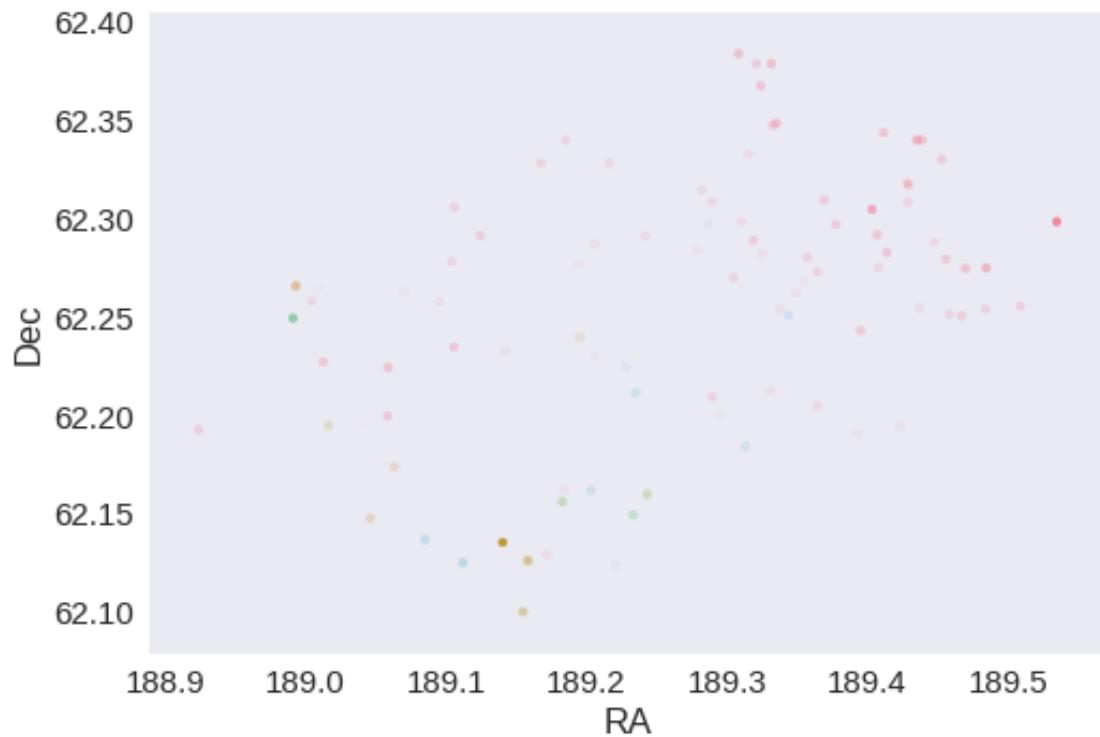
The cleaned catalogue has 35423 sources (22 removed).

The cleaned catalogue has 21 sources flagged as having been cleaned

1.4 III - Astrometry correction

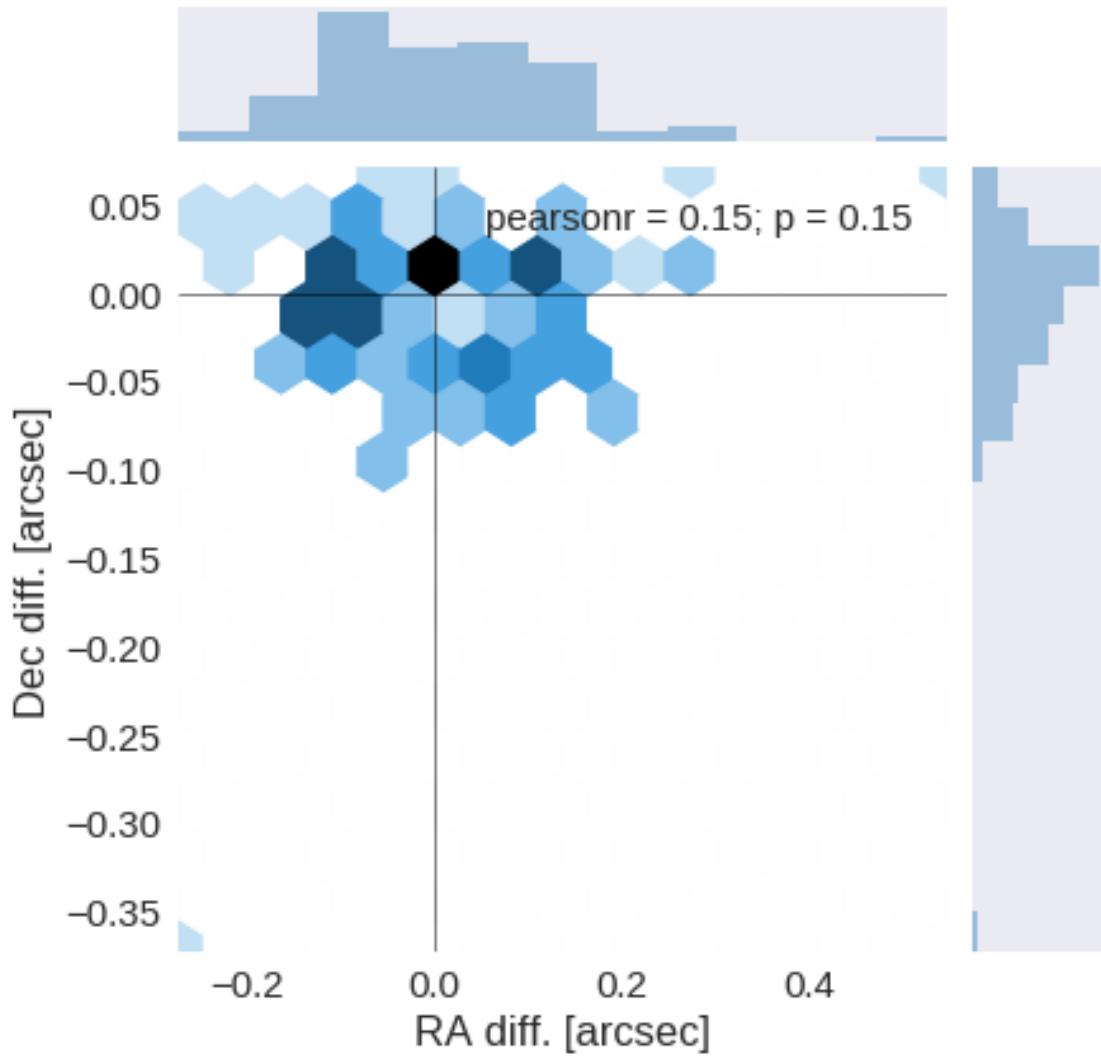
We match the astrometry to the Gaia one. We limit the Gaia catalogue to sources with a g band flux between the 30th and the 70th percentile. Some quick tests show that this give the lower dispersion in the results.

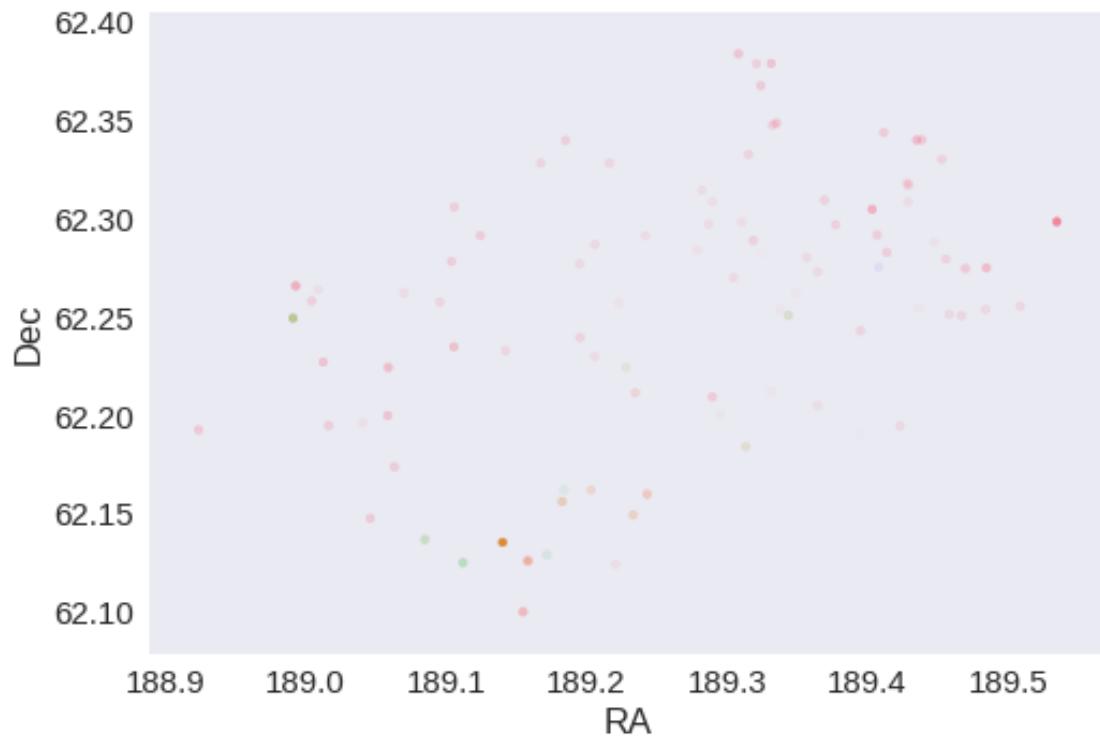




RA correction: 0.02082349079728374 arcsec

Dec correction: -0.02865639995377478 arcsec





1.5 IV - Flagging Gaia objects

102 sources flagged.

2 V - Saving to disk

3_Checks_and_diagnostics

January 18, 2018

1 HDF-N master catalogue

1.1 Checks and diagnostics

This notebook was run with herschelhelp_internal version:
33f5ec7 (Wed Dec 6 16:56:17 2017 +0000)

Diagnostics done using: master_catalogue_hdf-n_20171212.fits

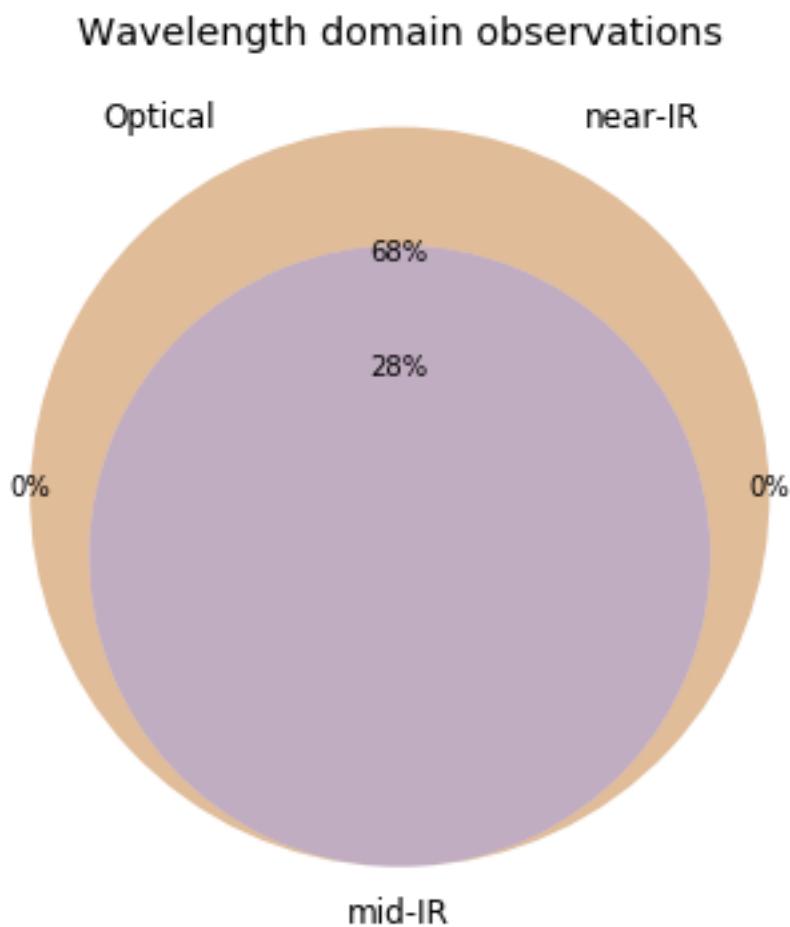
1.2 0 - Quick checks

```
/opt/anaconda3/envs/herschelhelp_internal/lib/python3.6/site-packages/numpy/core/numeric.py:301:
    format(shape, fill_value, array(fill_value).dtype), FutureWarning)
/opt/anaconda3/envs/herschelhelp_internal/lib/python3.6/site-packages/numpy/core/numeric.py:301:
    format(shape, fill_value, array(fill_value).dtype), FutureWarning)
```

Table shows only problematic columns.

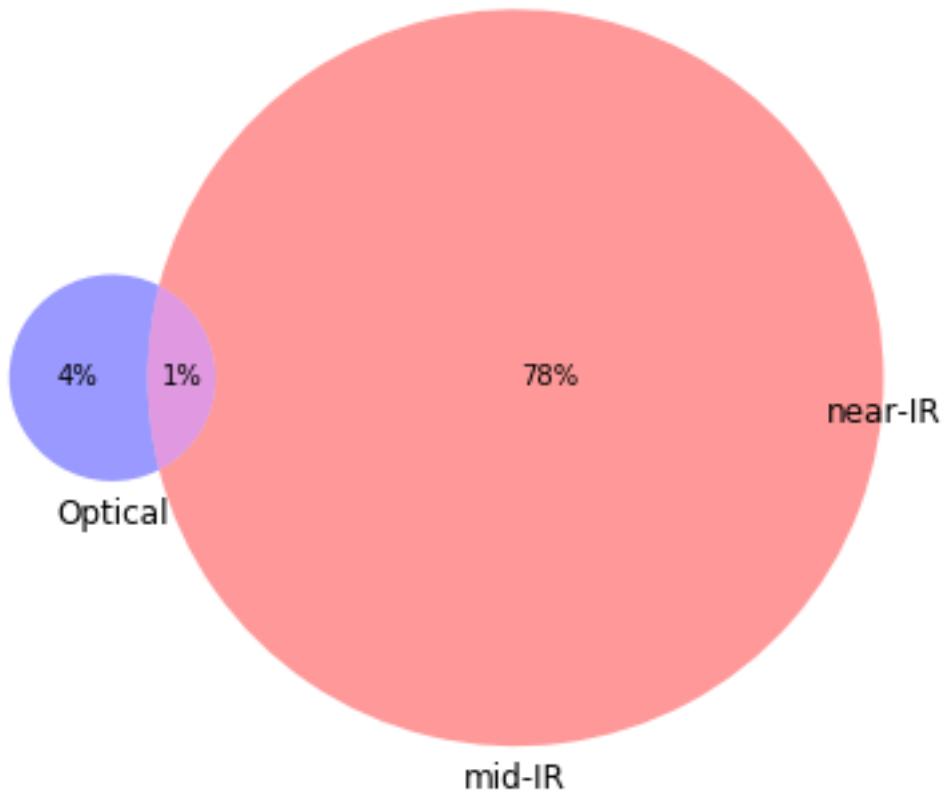
Out[4]: <IPython.core.display.HTML object>

1.3 I - Summary of wavelength domains



```
/opt/anaconda3/envs/herschelhelp_internal/lib/python3.6/site-packages/matplotlib_venn/_venn3.py:  
    warnings.warn("Circle B has zero area")
```

Detection of the 47,176 sources detected
in any wavelength domains (among 130,679 sources)

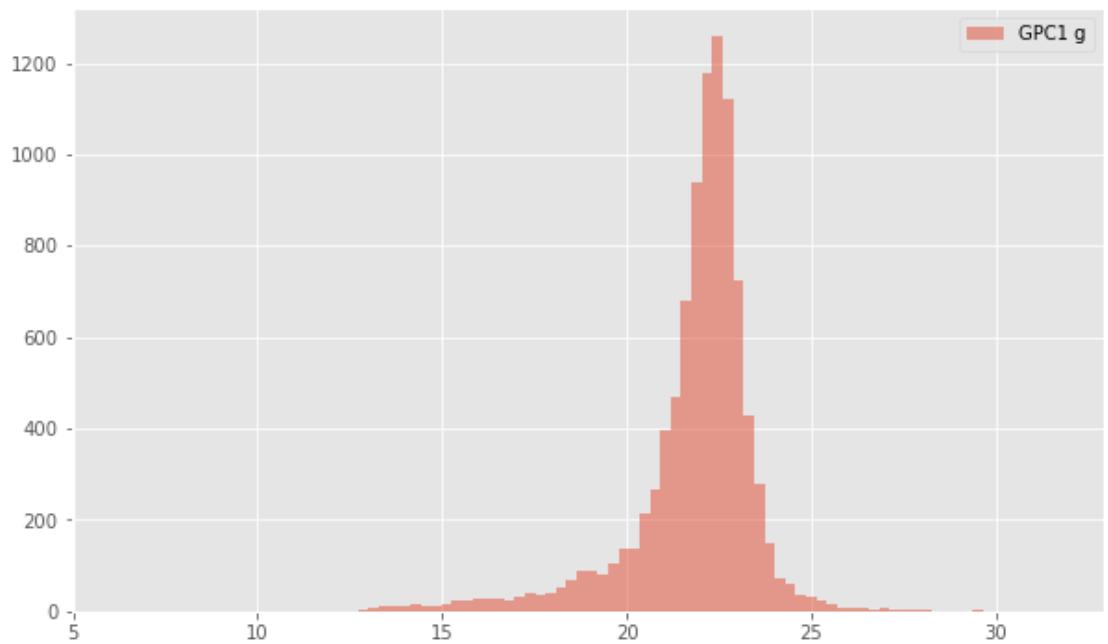
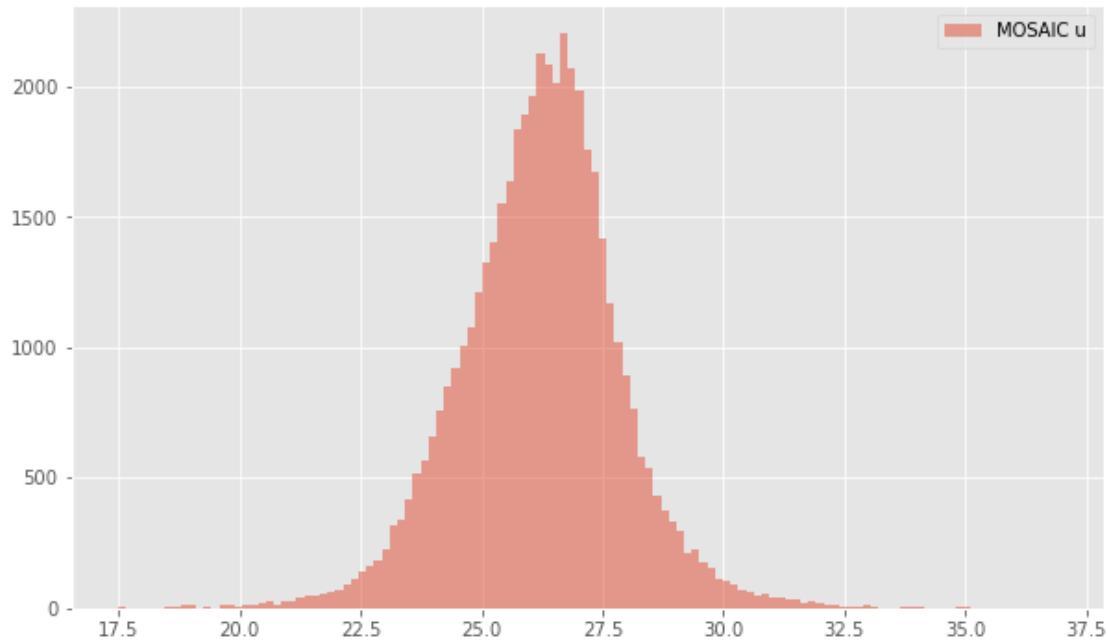


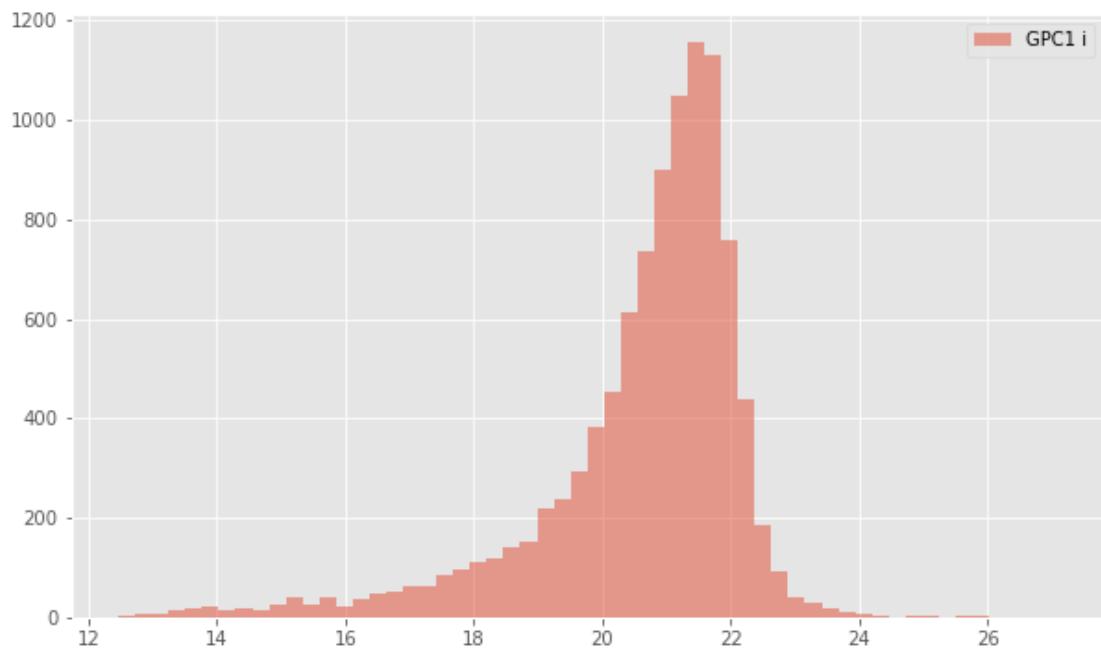
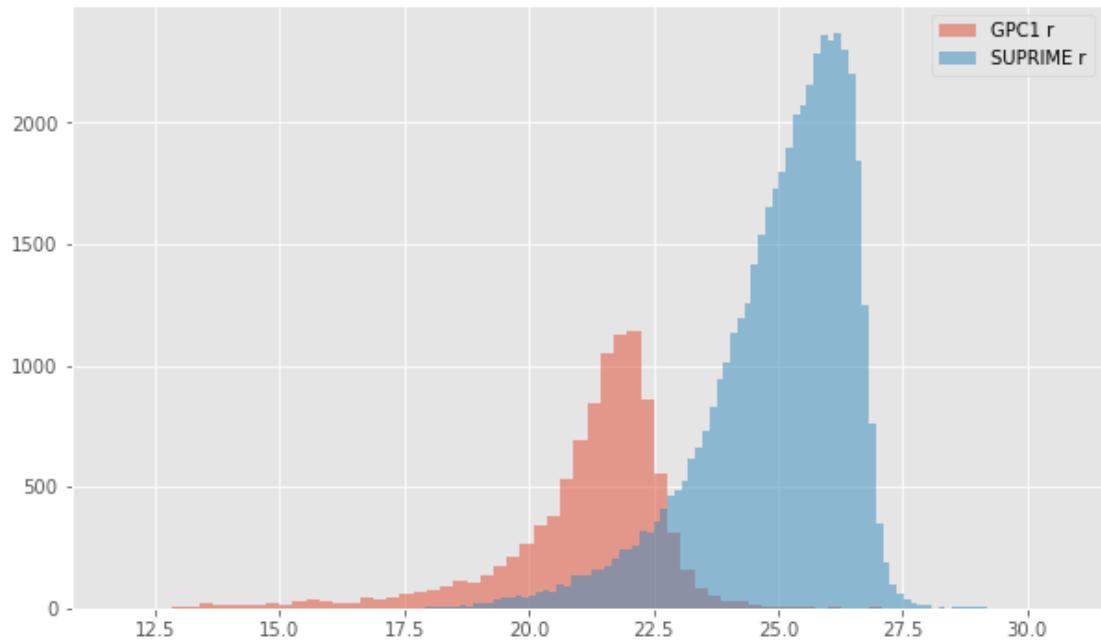
1.4 II - Comparing magnitudes in similar filters

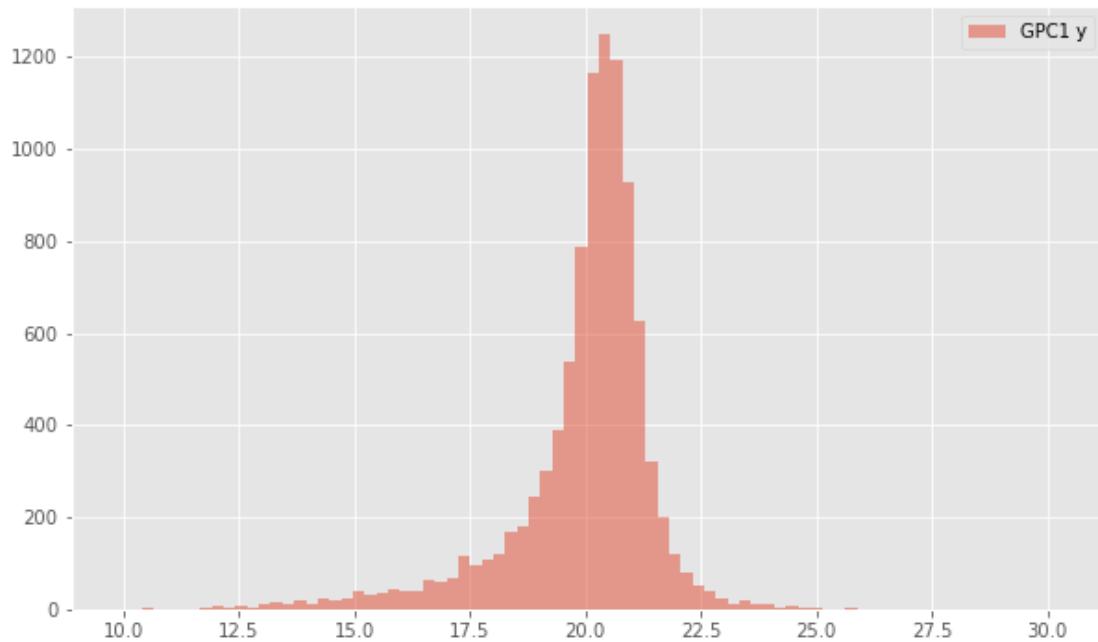
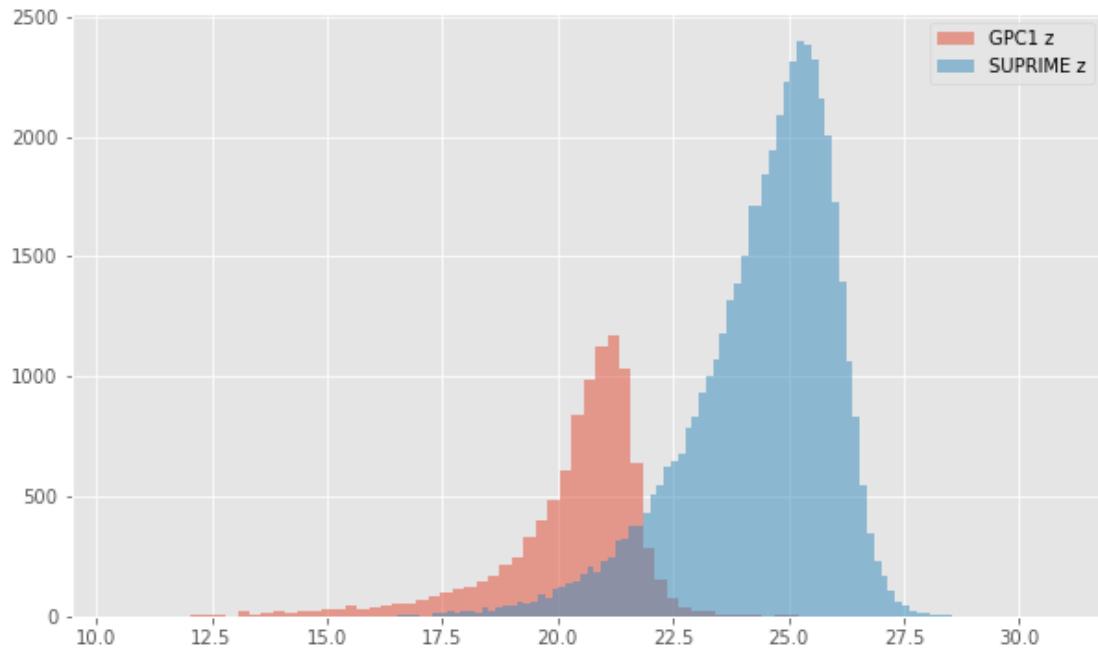
The master list is composed of several catalogues containing magnitudes in similar filters on different instruments. We are comparing the magnitudes in these corresponding filters.

1.4.1 II.a - Comparing depths

We compare the histograms of the total aperture magnitudes of similar bands.





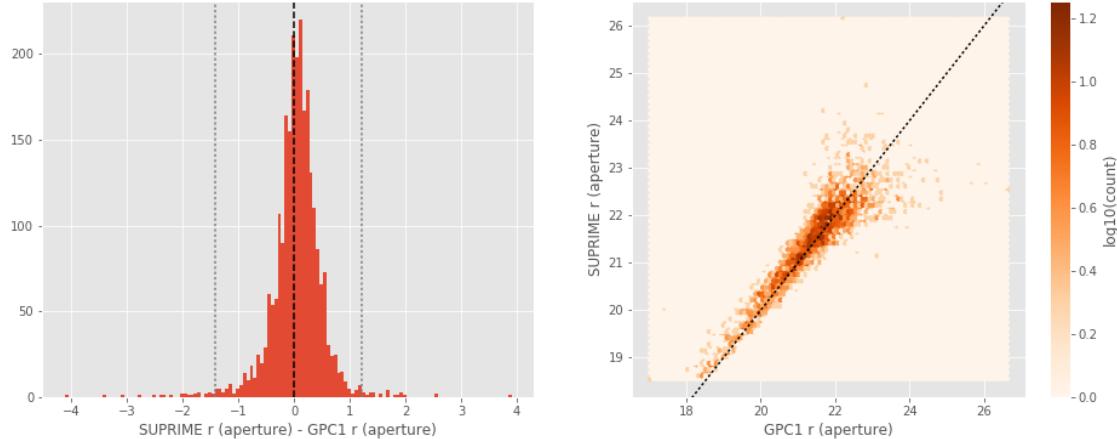


1.4.2 II.b - Comparing magnitudes

We compare one to one each magnitude in similar bands.

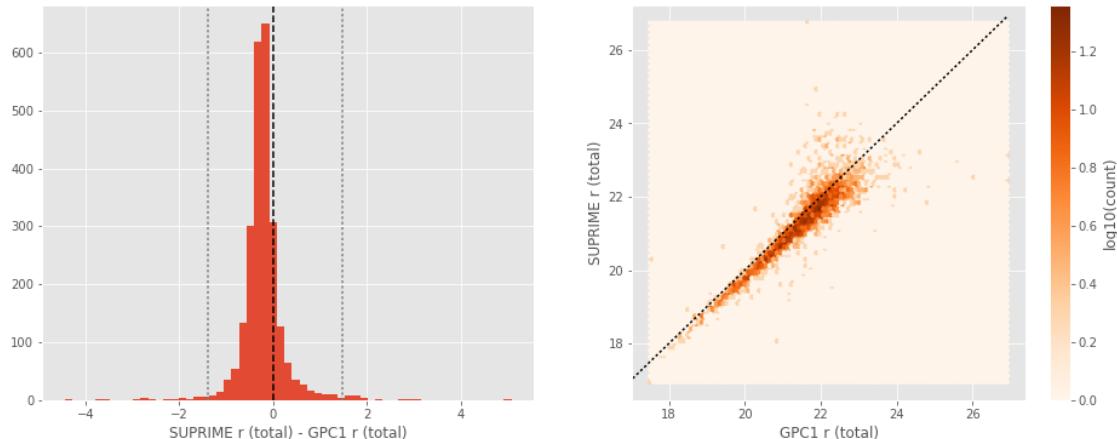
SUPRIME r (aperture) - GPC1 r (aperture):

- Median: 0.07
- Median Absolute Deviation: 0.21
- 1% percentile: -1.424673613403319
- 99% percentile: 1.2040346905273431



SUPRIME r (total) - GPC1 r (total):

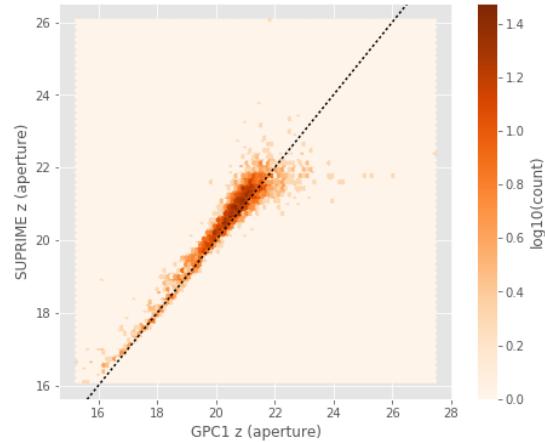
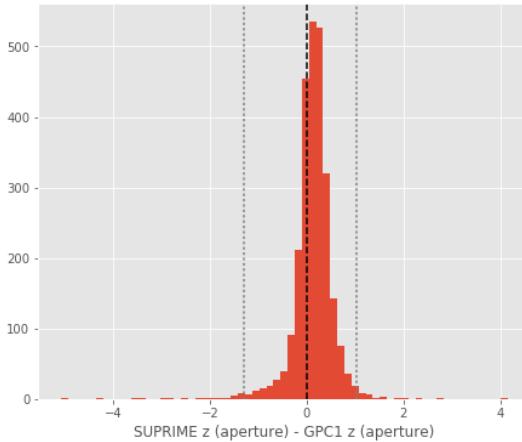
- Median: -0.24
- Median Absolute Deviation: 0.16
- 1% percentile: -1.4006093564038085
- 99% percentile: 1.4649504818432544



SUPRIME z (aperture) - GPC1 z (aperture):

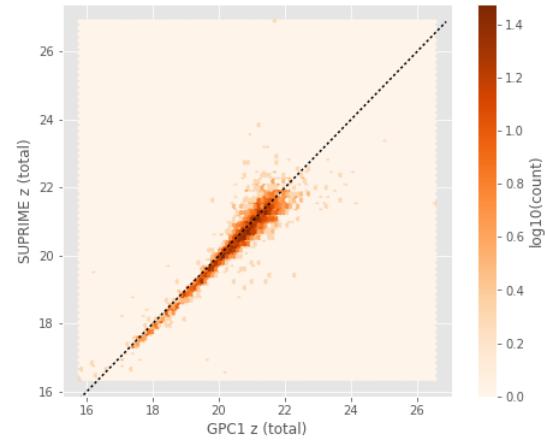
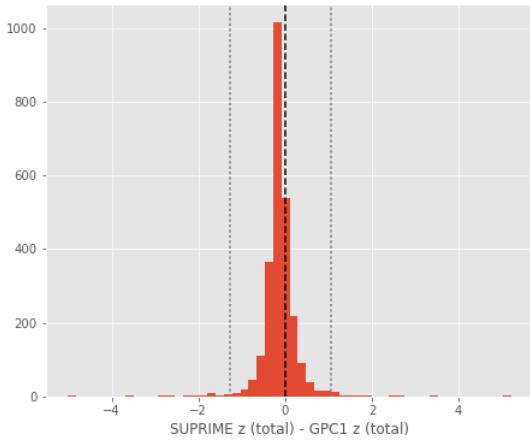
- Median: 0.15

- Median Absolute Deviation: 0.18
- 1% percentile: -1.3092136831054682
- 99% percentile: 1.036628628840331



SUPRIME z (total) - GPC1 z (total):

- Median: -0.14
- Median Absolute Deviation: 0.13
- 1% percentile: -1.280440952136231
- 99% percentile: 1.0462272840478488



1.5 III - Comparing magnitudes to reference bands

Cross-match the master list to SDSS and 2MASS to compare its magnitudes to SDSS and 2MASS ones.

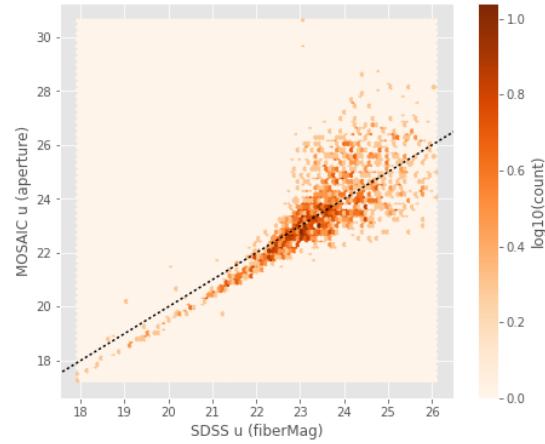
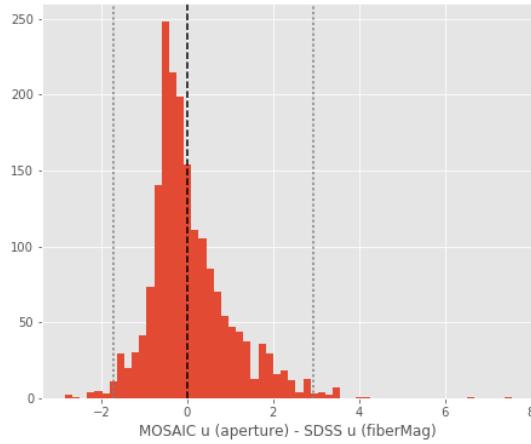
1.5.1 III.a - Comparing u, g, r, i, and z bands to SDSS

The catalogue is cross-matched to SDSS-DR13 withing 0.2 arcsecond.

We compare the u, g, r, i, and z magnitudes to those from SDSS using fiberMag for the aperture magnitude and petroMag for the total magnitude.

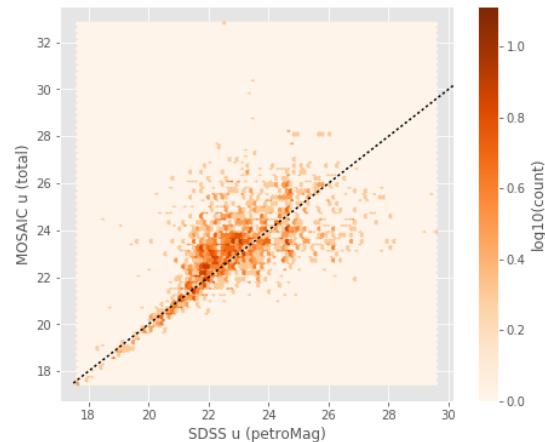
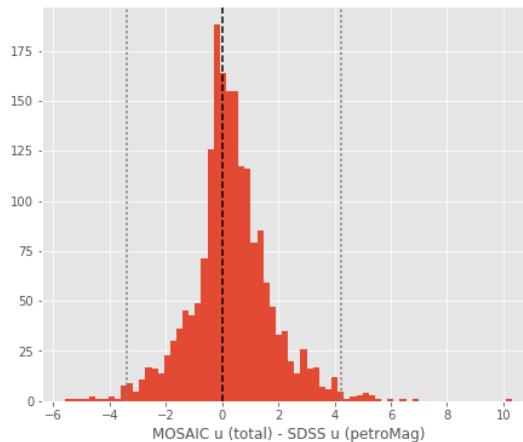
MOSAIC u (aperture) - SDSS u (fiberMag) :

- Median: -0.15
- Median Absolute Deviation: 0.44
- 1% percentile: -1.7168668045751936
- 99% percentile: 2.942735320161134



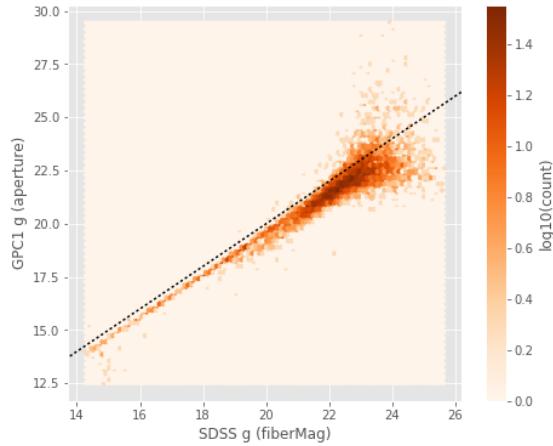
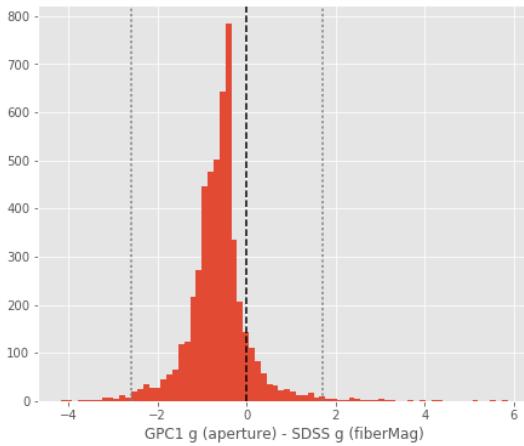
MOSAIC u (total) - SDSS u (petroMag) :

- Median: 0.26
- Median Absolute Deviation: 0.67
- 1% percentile: -3.3910033163061515
- 99% percentile: 4.251866573155514



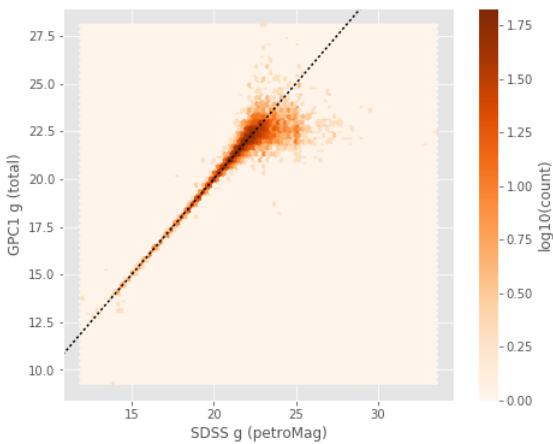
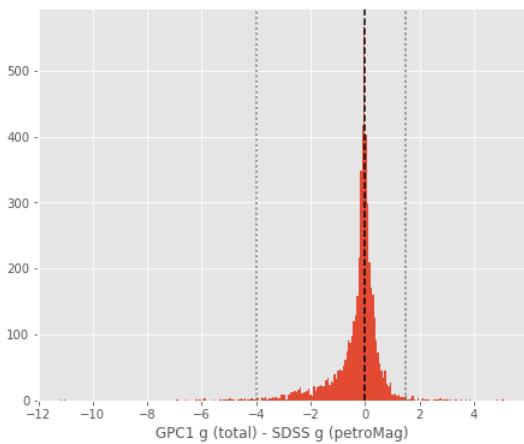
GPC1 g (aperture) - SDSS g (fiberMag):

- Median: -0.60
- Median Absolute Deviation: 0.29
- 1% percentile: -2.5760694885253907
- 99% percentile: 1.7048368072509783



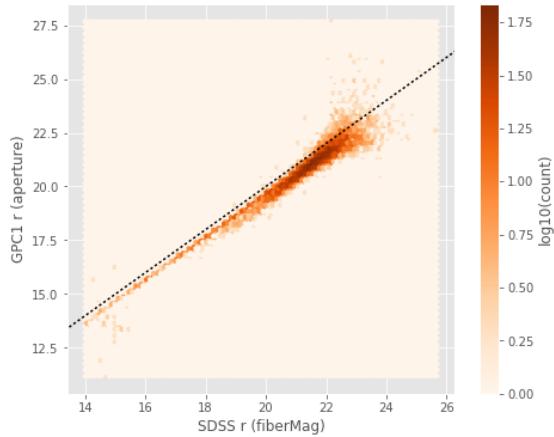
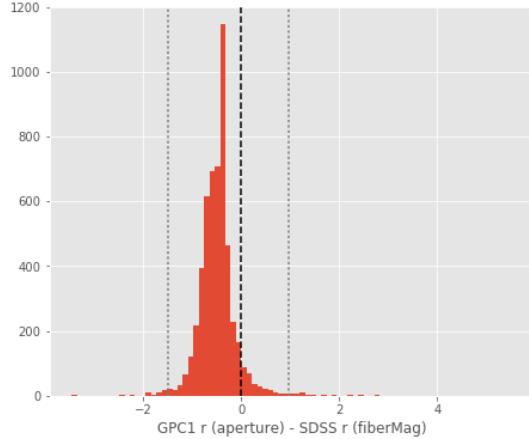
GPC1 g (total) - SDSS g (petroMag):

- Median: -0.08
- Median Absolute Deviation: 0.25
- 1% percentile: -3.9923860931396487
- 99% percentile: 1.4683370971679683



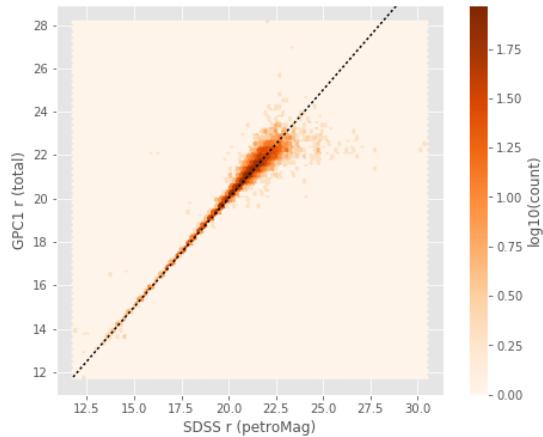
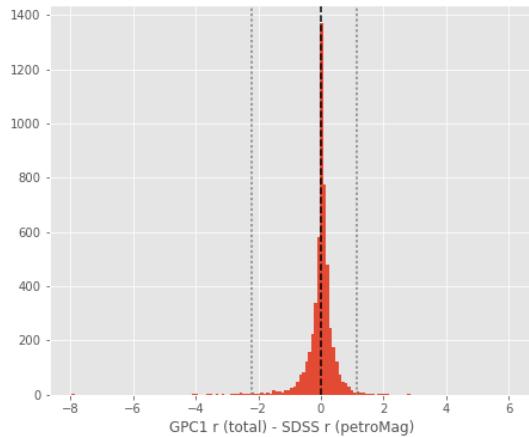
GPC1 r (aperture) - SDSS r (fiberMag):

- Median: -0.47
- Median Absolute Deviation: 0.17
- 1% percentile: -1.4966218566894531
- 99% percentile: 0.9670242309570285



GPC1 r (total) - SDSS r (petroMag):

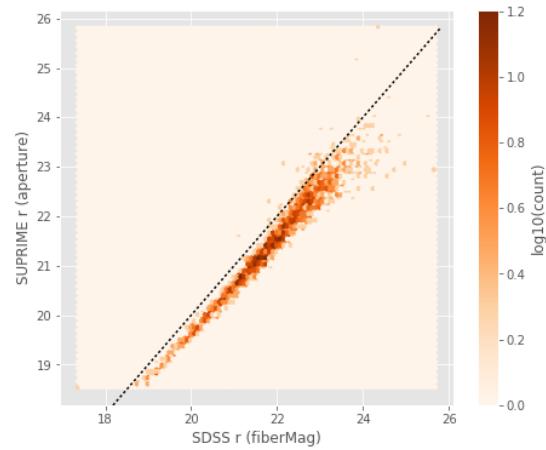
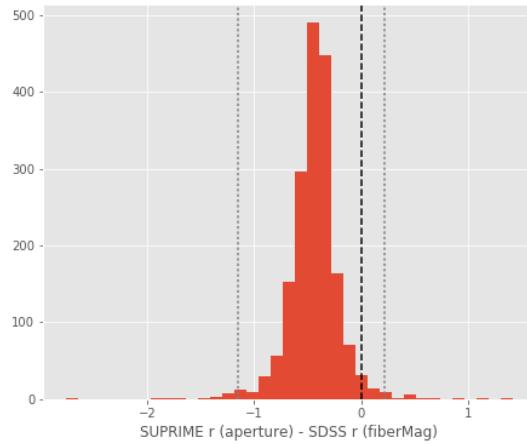
- Median: 0.03
- Median Absolute Deviation: 0.14
- 1% percentile: -2.213793106079102
- 99% percentile: 1.1340791320800772



SUPRIME r (aperture) - SDSS r (fiberMag):

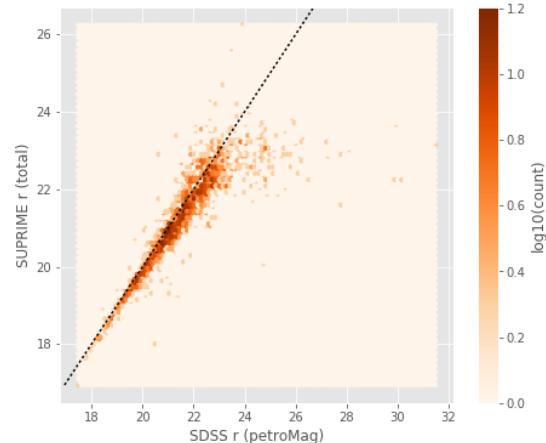
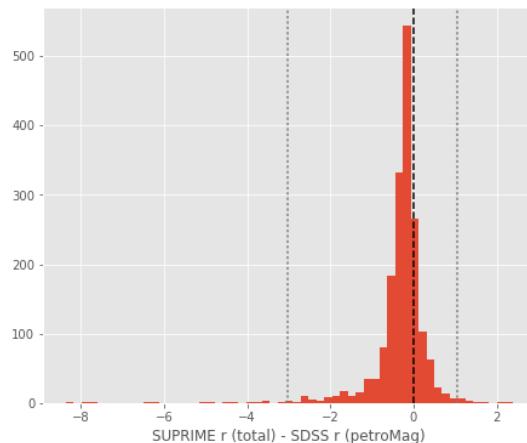
- Median: -0.42

- Median Absolute Deviation: 0.11
- 1% percentile: -1.1524517989306626
- 99% percentile: 0.2232965563378855



SUPRIME r (total) - SDSS r (petroMag):

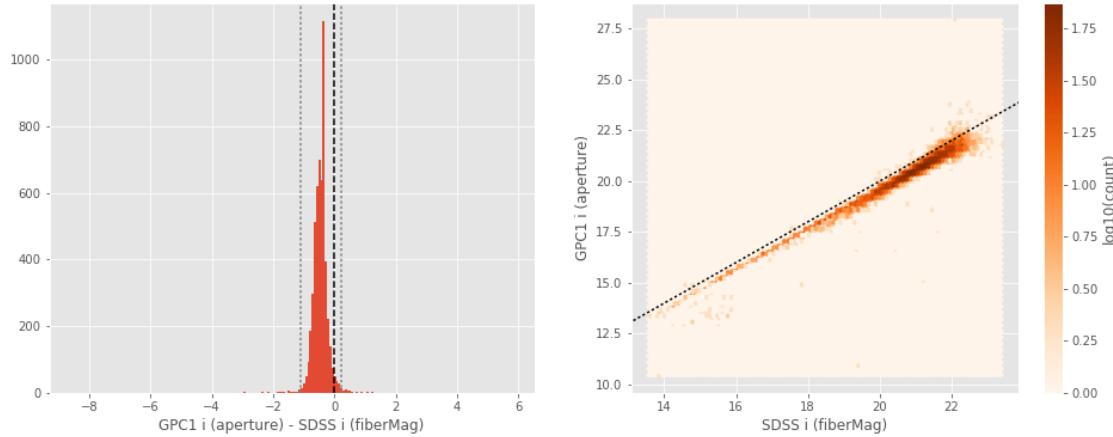
- Median: -0.20
- Median Absolute Deviation: 0.19
- 1% percentile: -3.034374487626954
- 99% percentile: 1.0472558652038568



GPC1 i (aperture) - SDSS i (fiberMag):

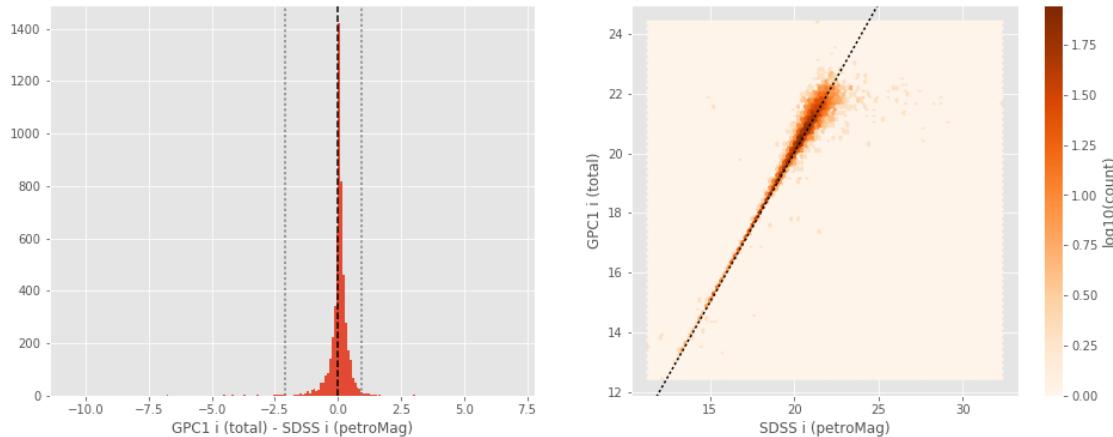
- Median: -0.44
- Median Absolute Deviation: 0.12

- 1% percentile: -1.1171233367919922
- 99% percentile: 0.2286124420166025



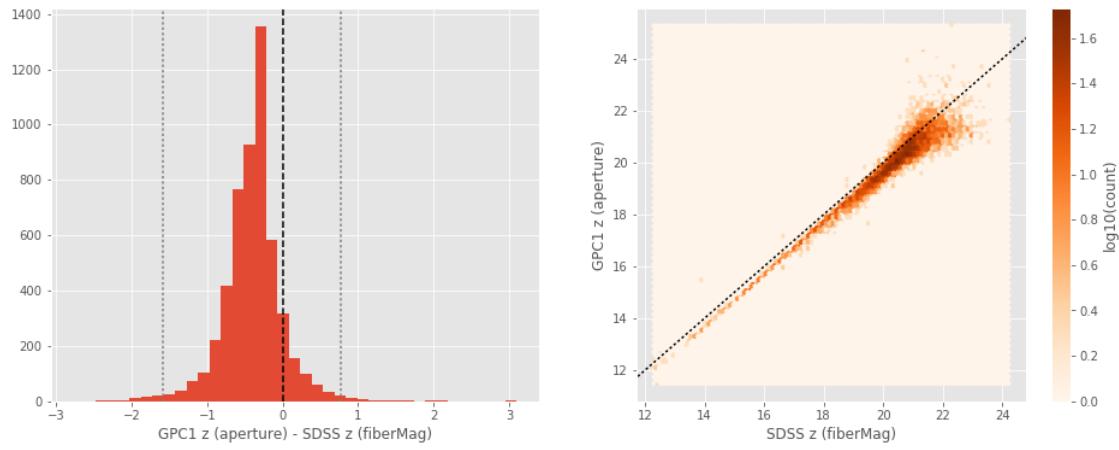
GPC1 i (total) - SDSS i (petroMag):

- Median: 0.06
- Median Absolute Deviation: 0.12
- 1% percentile: -2.0888433837890625
- 99% percentile: 0.9435255050659161



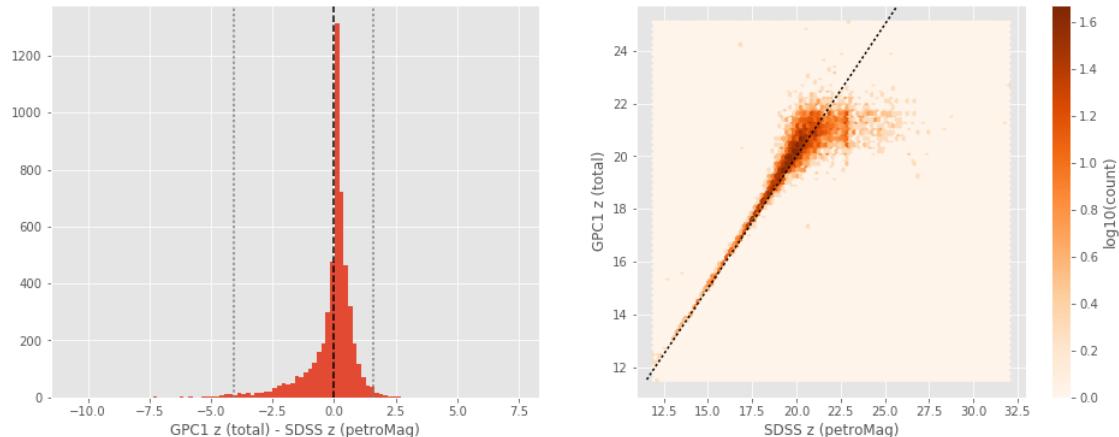
GPC1 z (aperture) - SDSS z (fiberMag):

- Median: -0.36
- Median Absolute Deviation: 0.18
- 1% percentile: -1.5850920677185059
- 99% percentile: 0.7624429321289085



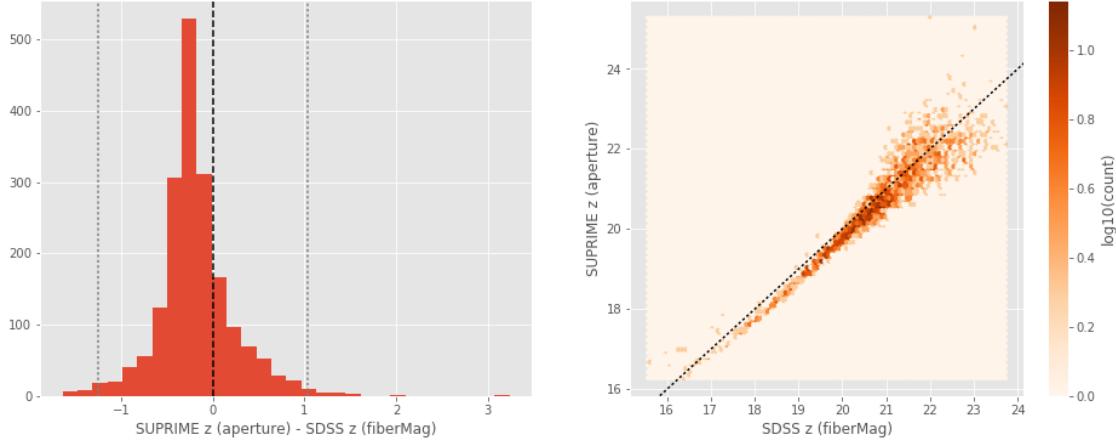
GPC1 z (total) - SDSS z (petroMag):

- Median: 0.10
- Median Absolute Deviation: 0.30
- 1% percentile: -4.0788907623291015
- 99% percentile: 1.606949920654297



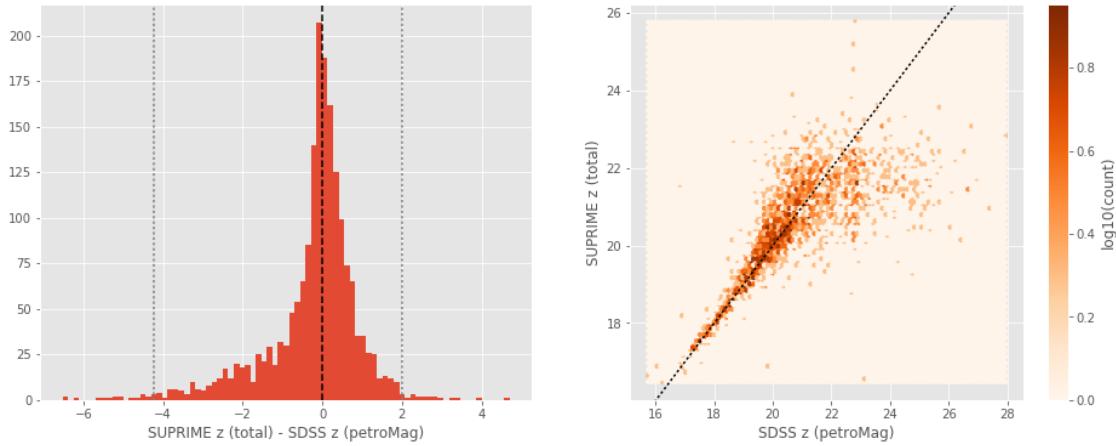
SUPRIME z (aperture) - SDSS z (fiberMag):

- Median: -0.23
- Median Absolute Deviation: 0.17
- 1% percentile: -1.2519832500915538
- 99% percentile: 1.0297534400939927



SUPRIME z (total) - SDSS z (petroMag):

- Median: -0.02
- Median Absolute Deviation: 0.43
- 1% percentile: -4.218913514050293
- 99% percentile: 2.0048278384594718



1.5.2 III.b - Comparing J and K bands to 2MASS

The catalogue is cross-matched to 2MASS-PSC within 0.2 arcsecond. We compare the UKIDSS total J and K magnitudes to those from 2MASS.

The 2MASS magnitudes are “Vega-like” and we have to convert them to AB magnitudes using the zero points provided on [this page](#):

Band	F - 0 mag (Jy)
J	1594
H	1024
Ks	666.7

In addition, UKIDSS uses a K band whereas 2MASS uses a Ks (“short”) band, [this page](#) give a correction to convert the K band in a Ks band with the formula:

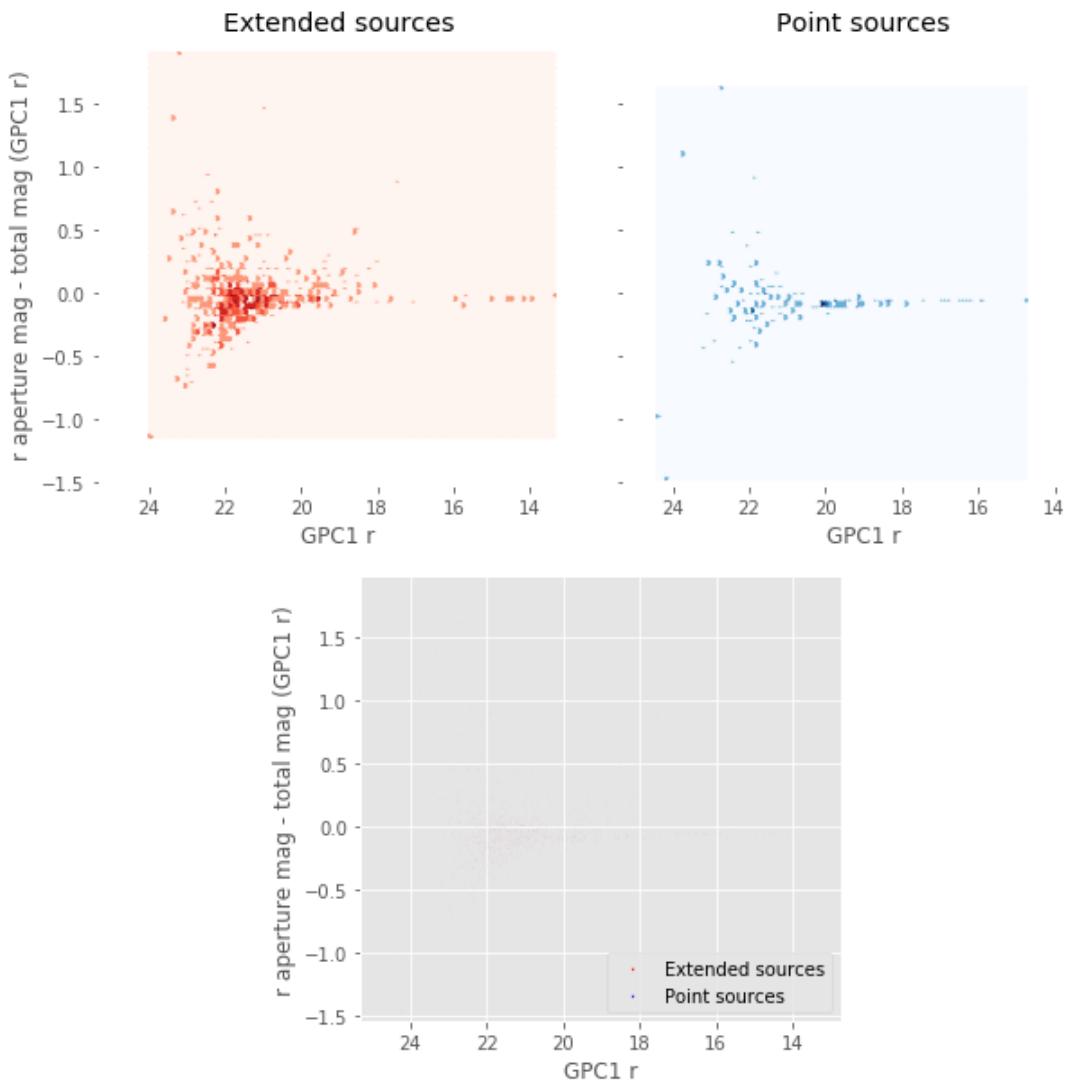
$$K_{s(2MASS)} = K_{UKIRT} + 0.003 + 0.004 * (JK)_{UKIRT}$$

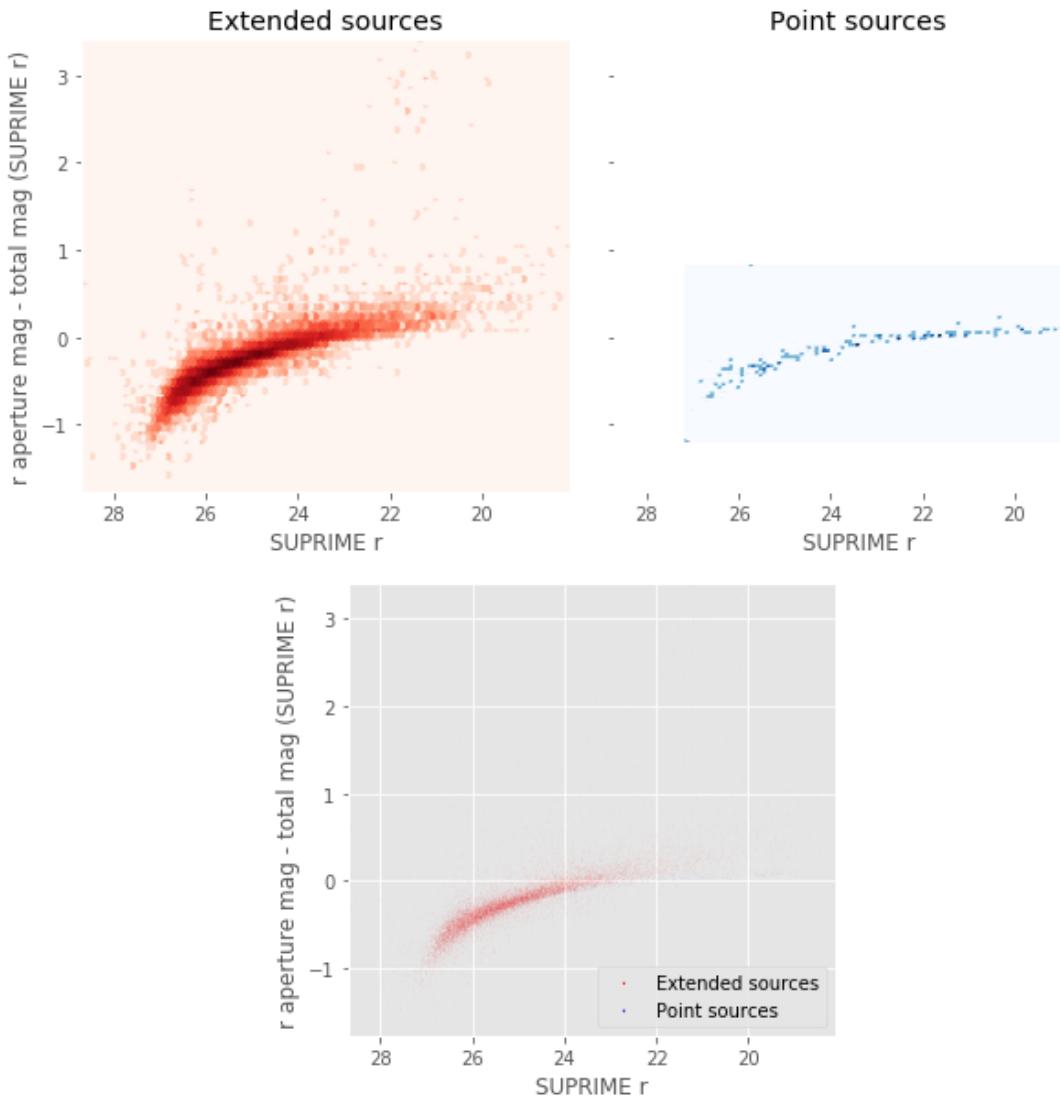
No sources have both 2MASS Ks and WIRCAM Ks (total) values.

1.6 IV - Comparing aperture magnitudes to total ones.

Number of source used: 682 / 130679 (0.52%)

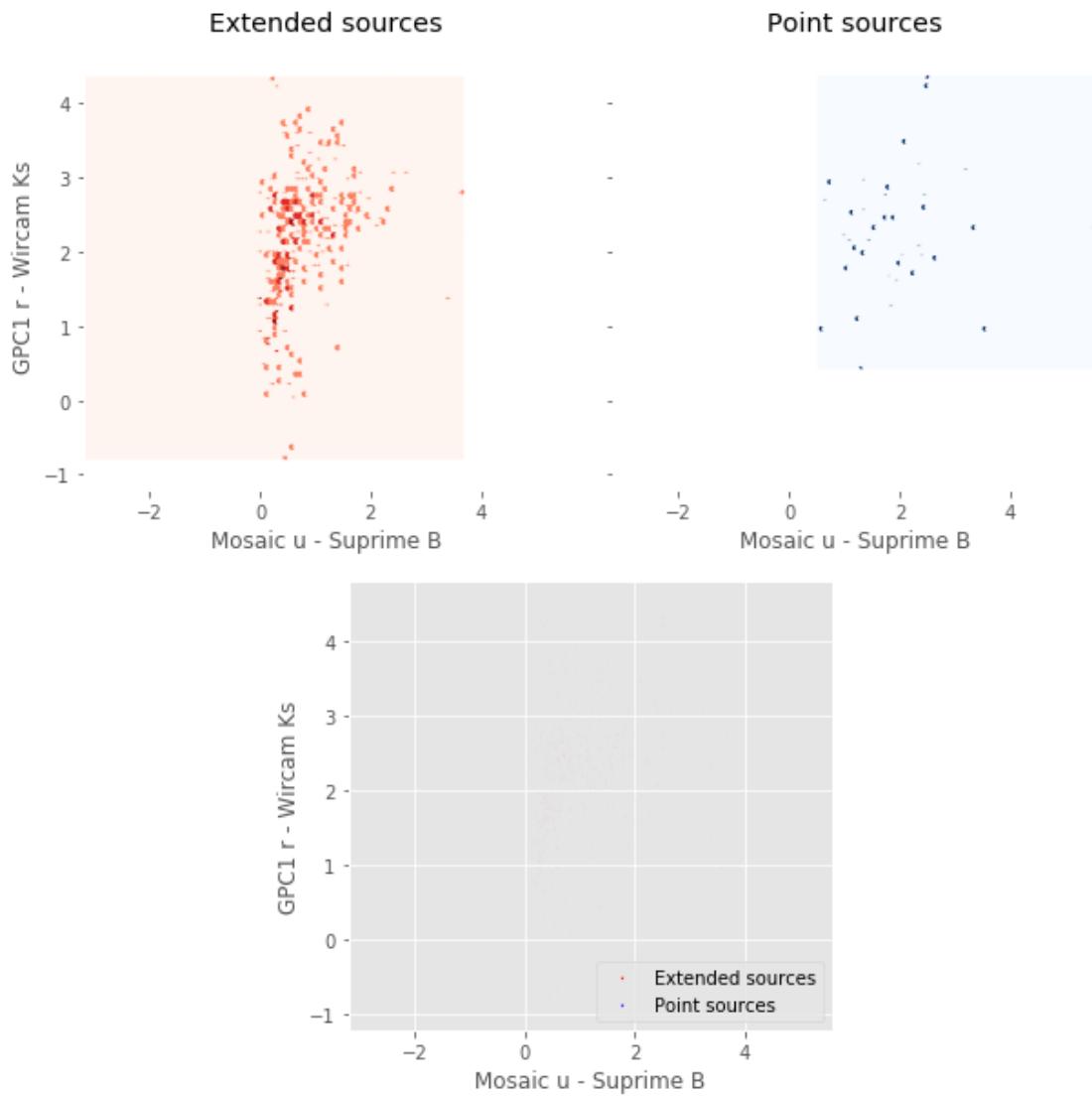
Number of source used: 12017 / 130679 (9.20%)



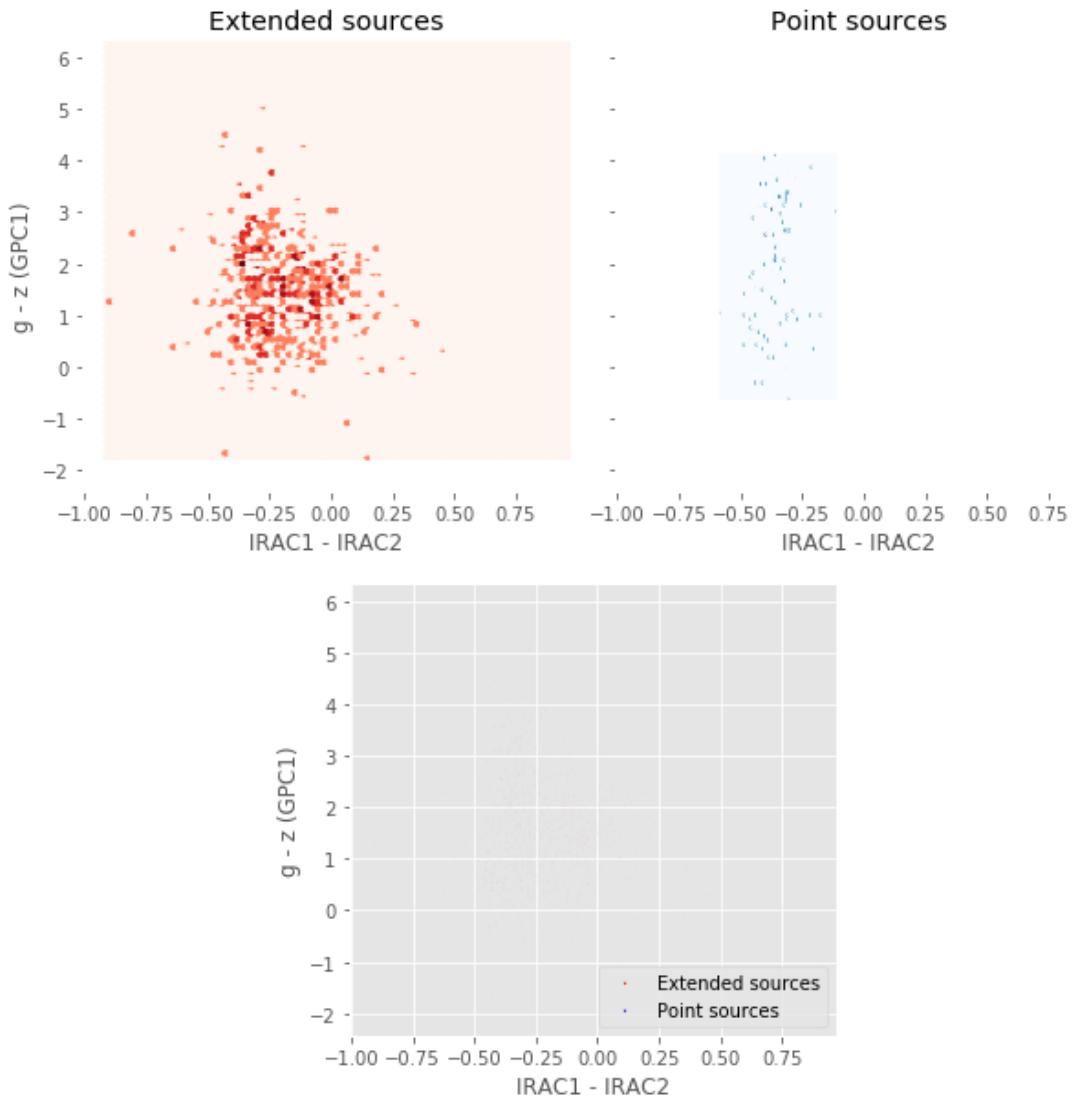


1.7 V - Color-color and magnitude-color plots

Number of source used: 374 / 130679 (0.29%)



Number of source used: 640 / 130679 (0.49%)



1.5_PanSTARRS1-3SS

January 18, 2018

1 HDF-N master catalogue

1.1 Preparation of Pan-STARRS1 - 3pi Steradian Survey (3SS) data

This catalogue comes from dm0_PanSTARRS1-3SS.

In the catalogue, we keep:

- The uniquePspSSTid as unique object identifier;
- The r-band position which is given for all the sources;
- The grizy <band>FApMag aperture magnitude (see below);
- The grizy <band>FKronMag as total magnitude.

The Pan-STARRS1-3SS catalogue provides for each band an aperture magnitude defined as “In PS1, an ‘optimal’ aperture radius is determined based on the local PSF. The wings of the same analytic PSF are then used to extrapolate the flux measured inside this aperture to a ‘total’ flux.”

The observations used for the catalogue were done between 2010 and 2015 ([ref](#)).

TODO: Check if the detection flag can be used to know in which bands an object was detected to construct the coverage maps.

TODO: Check for stellarity.

This notebook was run with herschelhelp_internal version:
33f5ec7 (Wed Dec 6 16:56:17 2017 +0000)

1.2 I - Column selection

```
/opt/anaconda3/envs/herschelhelp_internal/lib/python3.6/site-packages/astropy/table/column.py:10
Check the NumPy 1.11 release notes for more information.
    ma.MaskedArray.__setitem__(self, index, value)
```

Out [6]: <IPython.core.display.HTML object>

1.3 II - Removal of duplicated sources

We remove duplicated objects from the input catalogues.

```
/opt/anaconda3/envs/herschelhelp_internal/lib/python3.6/site-packages/astropy/table/column.py:10
Check the NumPy 1.11 release notes for more information.
    ma.MaskedArray.__setitem__(self, index, value)
```

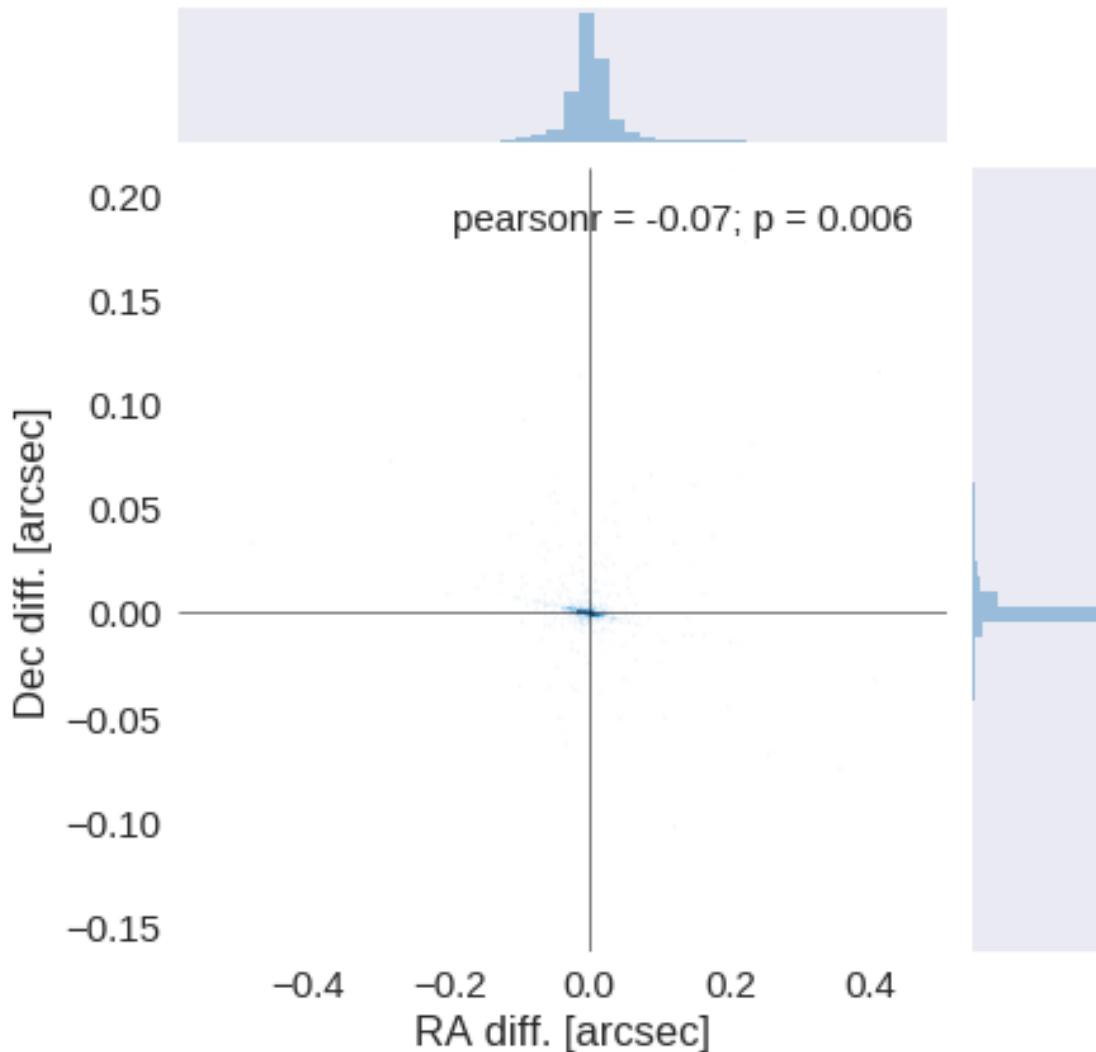
The initial catalogue had 10100 sources.

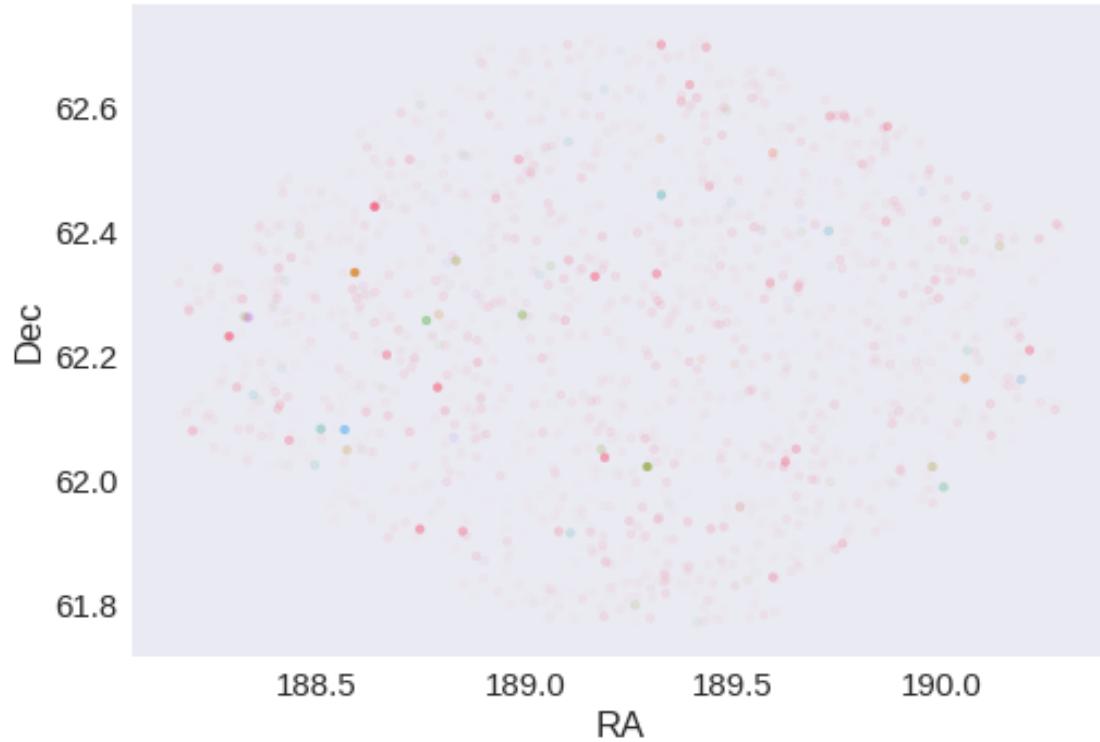
The cleaned catalogue has 10097 sources (3 removed).

The cleaned catalogue has 3 sources flagged as having been cleaned

1.4 III - Astrometry correction

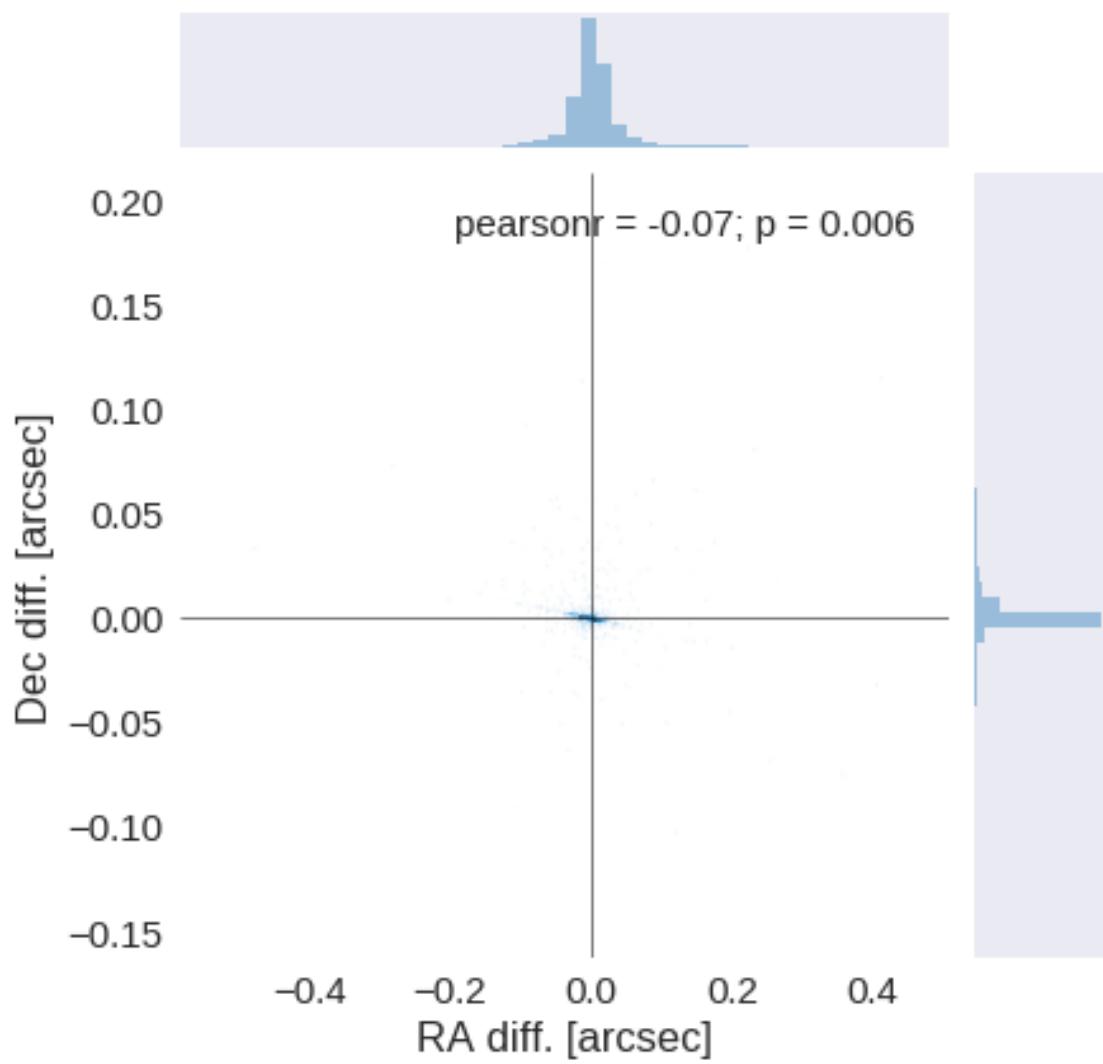
We match the astrometry to the Gaia one. We limit the Gaia catalogue to sources with a g band flux between the 30th and the 70th percentile. Some quick tests show that this give the lower dispersion in the results.

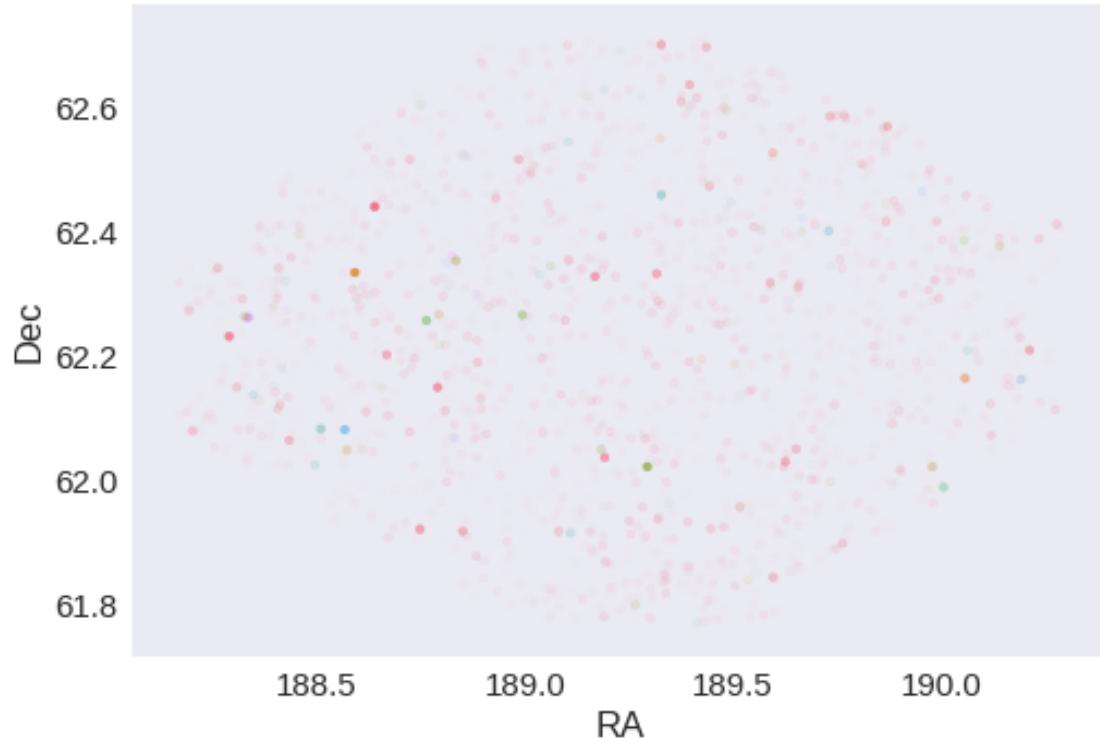




RA correction: 0.0009251711844626698 arcsec

Dec correction: -0.000312677103408987 arcsec





1.5 IV - Flagging Gaia objects

1548 sources flagged.

2 V - Saving to disk

1.4_Ultradeep-Ks-GOODS-N

January 18, 2018

1 HDF-N master catalogue

1.1 Preparation of Ultradeep-Ks-GOODS-N data

The catalogue comes from dmu0_Ultradeep-Ks-GOODS-N.

In the catalogue, we keep:

- The identifier (it's unique in the catalogue);
- The position;
- The total flux.

This notebook was run with herschelhelp_internal version:
33f5ec7 (Wed Dec 6 16:56:17 2017 +0000)

1.2 I - Column selection

Out [6]: <IPython.core.display.HTML object>

1.3 II - Removal of duplicated sources

We remove duplicated objects from the input catalogues.

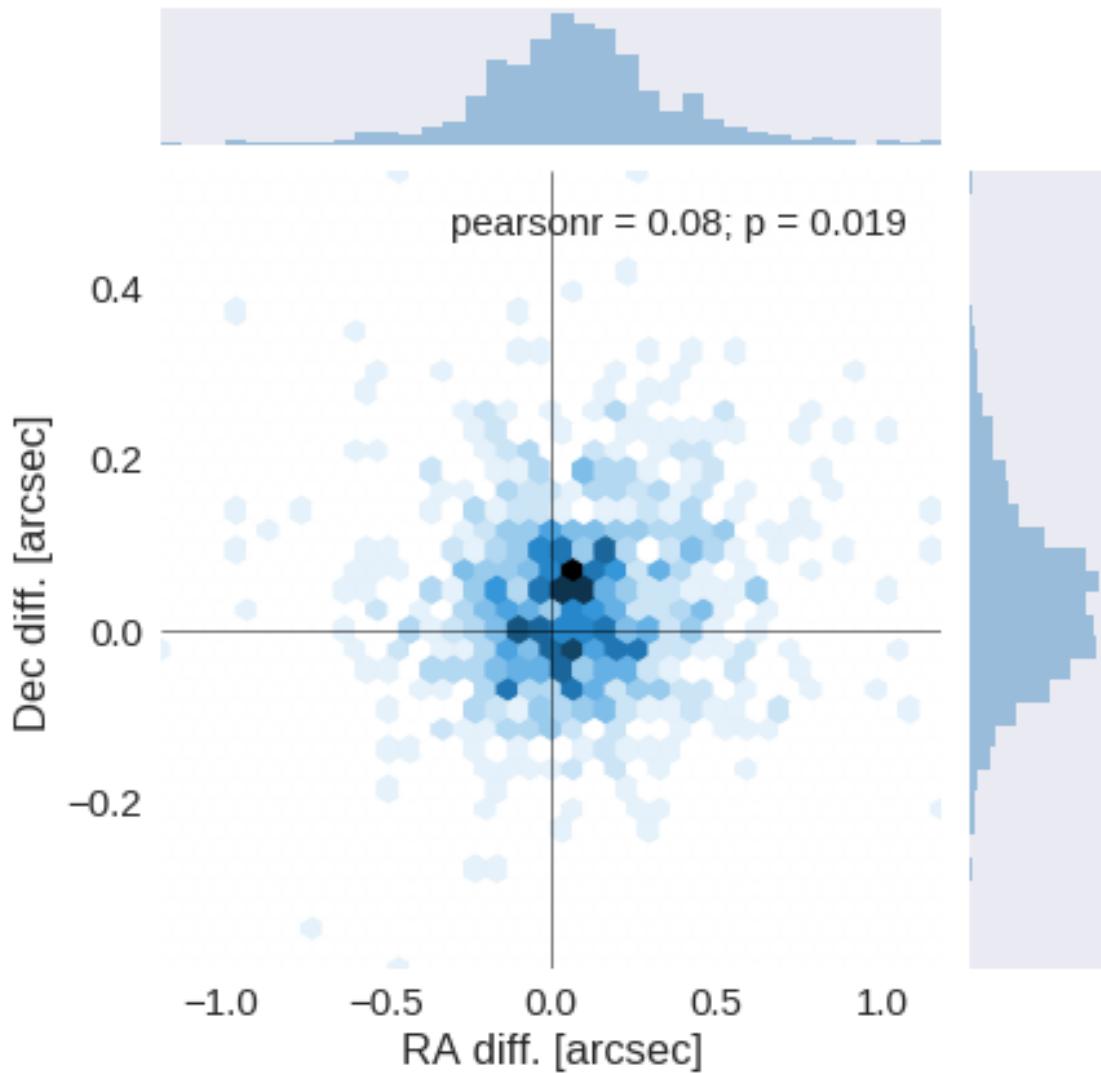
The initial catalogue had 94930 sources.

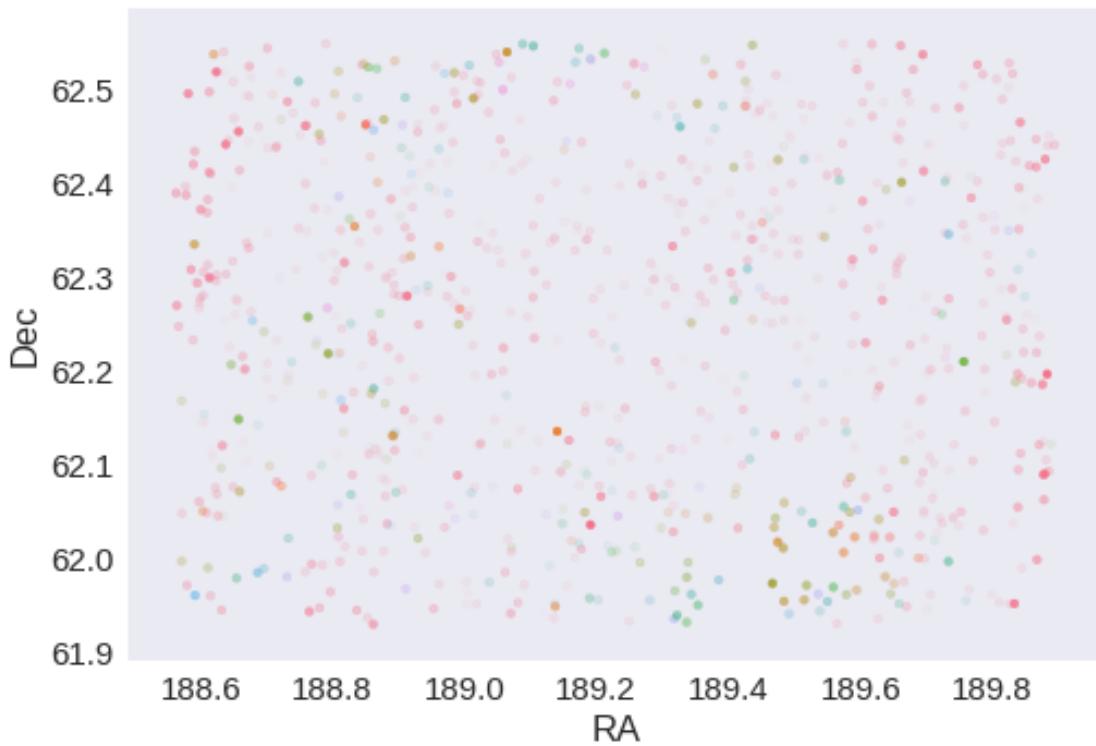
The cleaned catalogue has 94930 sources (0 removed).

The cleaned catalogue has 0 sources flagged as having been cleaned

1.4 III - Astrometry correction

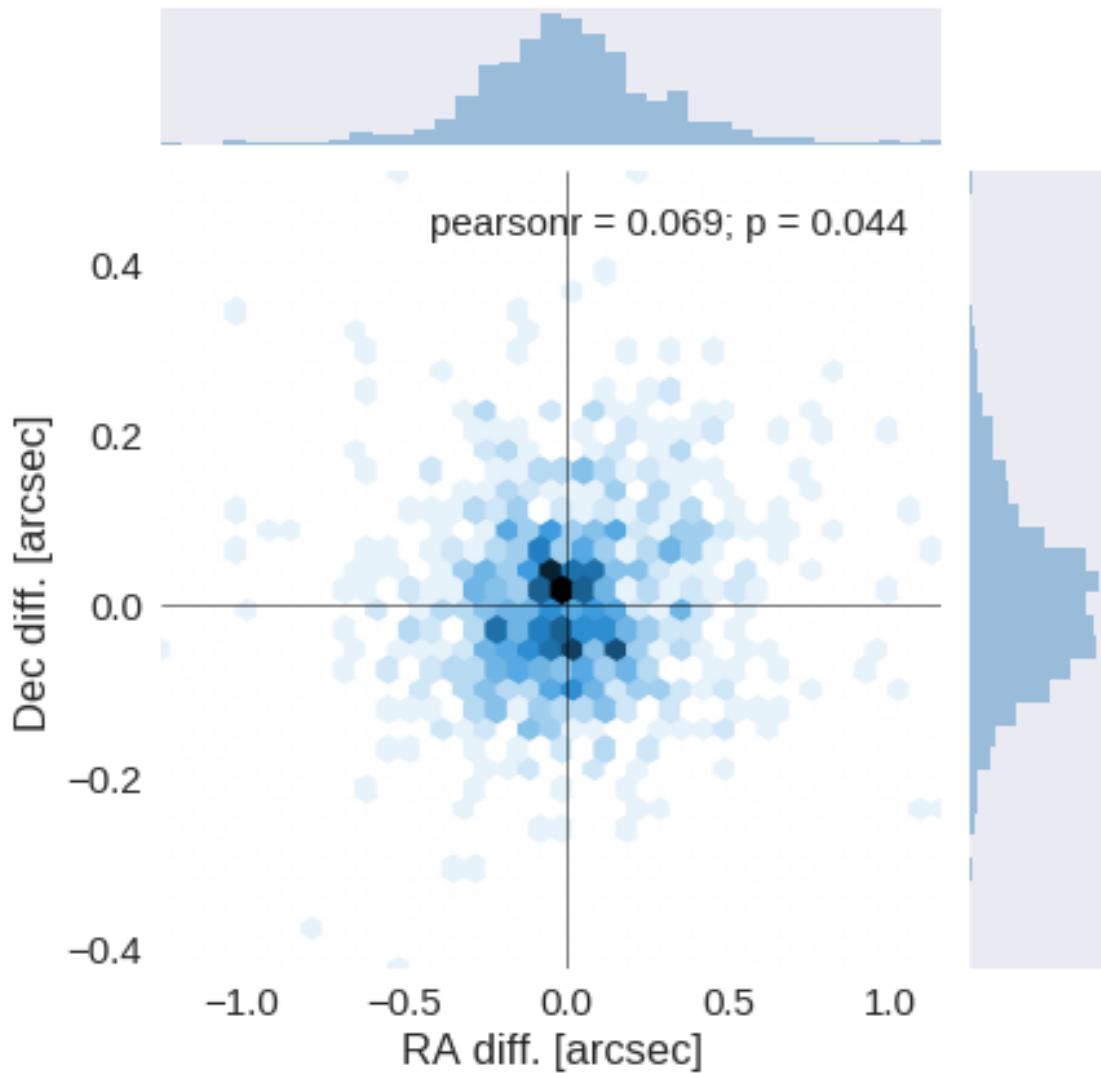
We match the astrometry to the Gaia one. We limit the Gaia catalogue to sources with a g band flux between the 30th and the 70th percentile. Some quick tests show that this give the lower dispersion in the results.

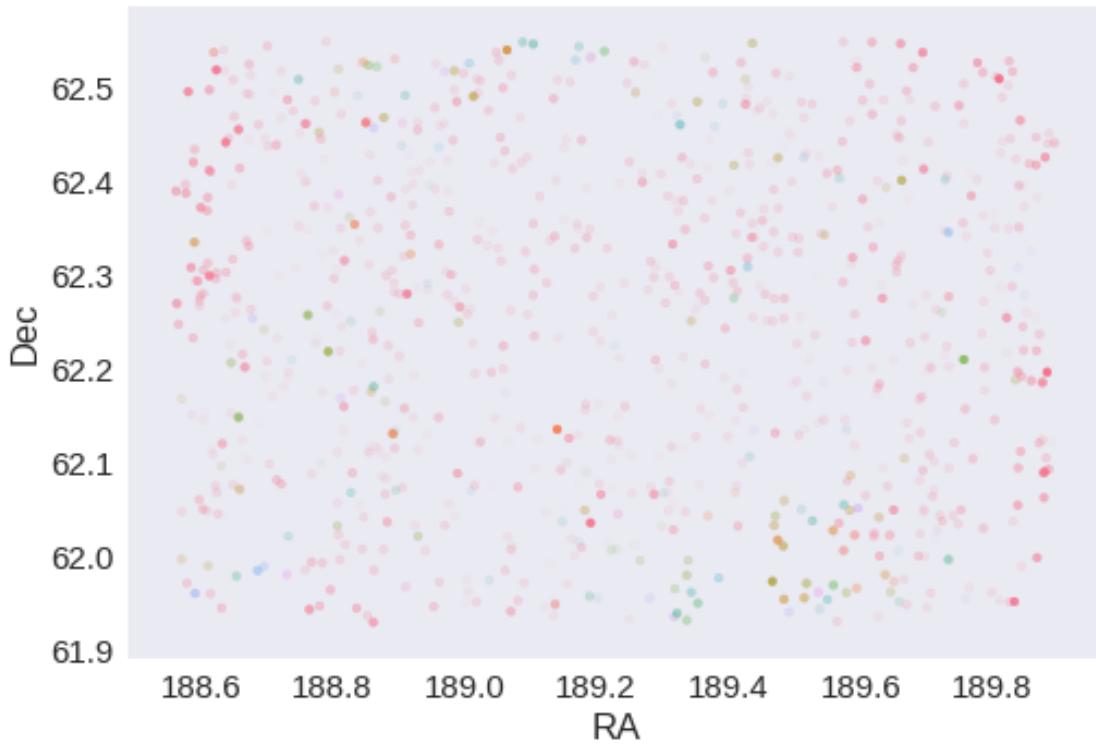




RA correction: -0.07018150315616367 arcsec

Dec correction: -0.02818903770247516 arcsec





1.5 IV - Flagging Gaia objects

888 sources flagged.

1.6 V - Flagging objects near bright stars

2 VI - Saving to disk

1.3_Hawaii-HDFN

January 18, 2018

1 HDF-N master catalogue

1.1 Preparation of Hawaii-HDFN data

The catalogue comes from dmu0_Hawaii-HDFN.

It contains UBVR_Iz data.

In the catalogue, we keep:

- The identifier (it's unique in the catalogue);
- The position;
- The stellarity;
- The kron magnitude, there doesn't appear to be aperture magnitudes. This may mean the survey is unusable.

This notebook was run with herschelhelp_internal version:
33f5ec7 (Wed Dec 6 16:56:17 2017 +0000)

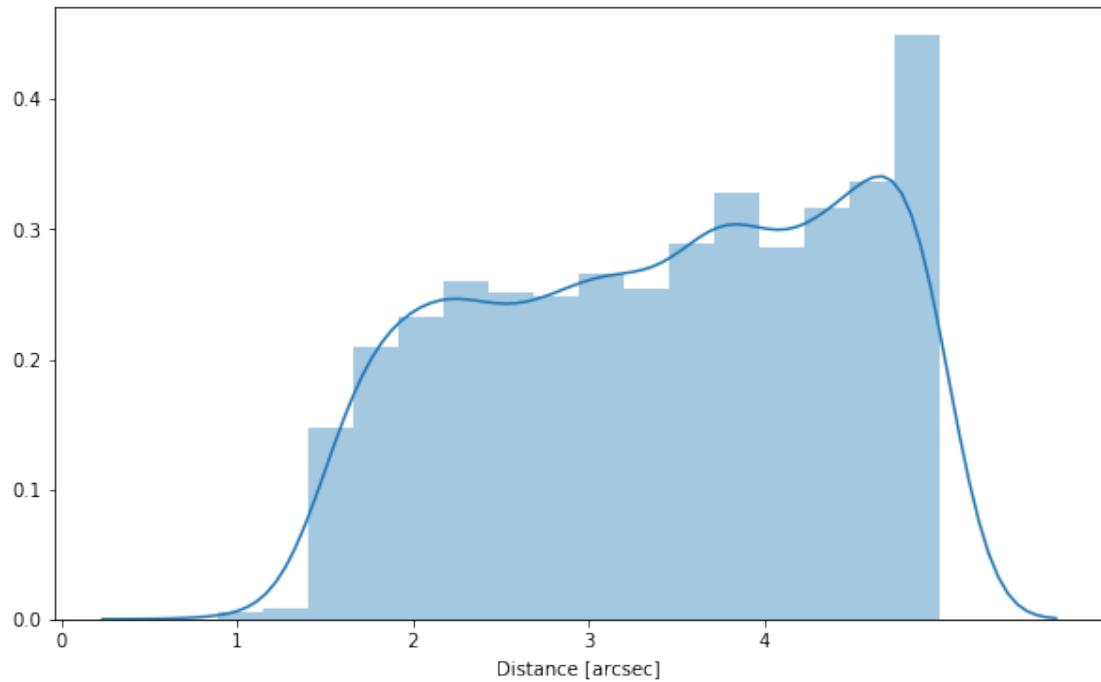
1.2 I - Column selection

Out[4]: <IPython.core.display.HTML object>

Out[7]: <IPython.core.display.HTML object>

Out[8]: <IPython.core.display.HTML object>

1.3 III - Merging



The catalogues appear to contain no cross matches. We therefore simply stack the catalogues. We need to understand why this is the case. Have cross matches already been removed from the Z selected catalogue?

Out [12]: <IPython.core.display.HTML object>

1.4 II - Removal of duplicated sources

We remove duplicated objects from the input catalogues.

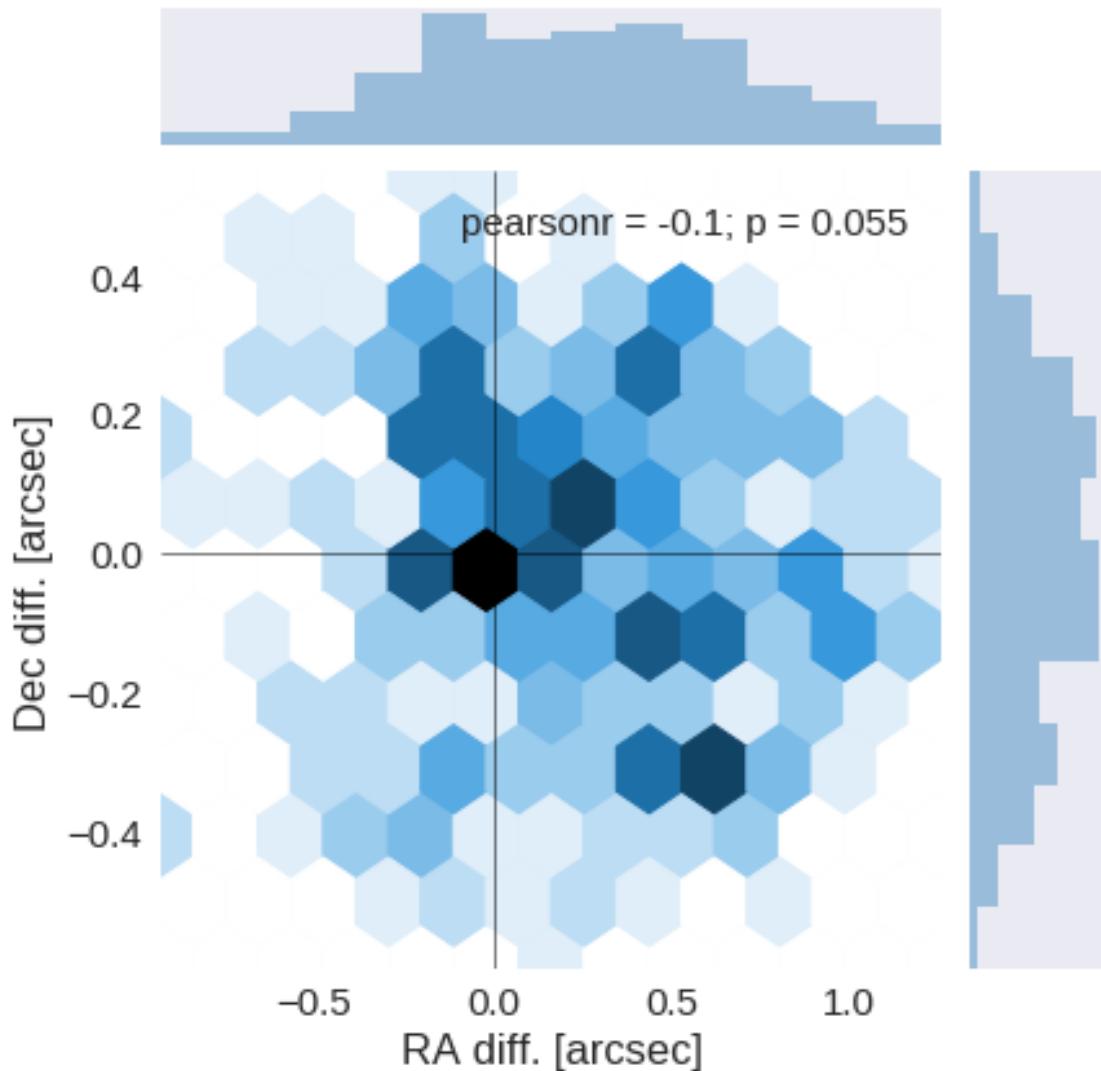
The initial catalogue had 48858 sources.

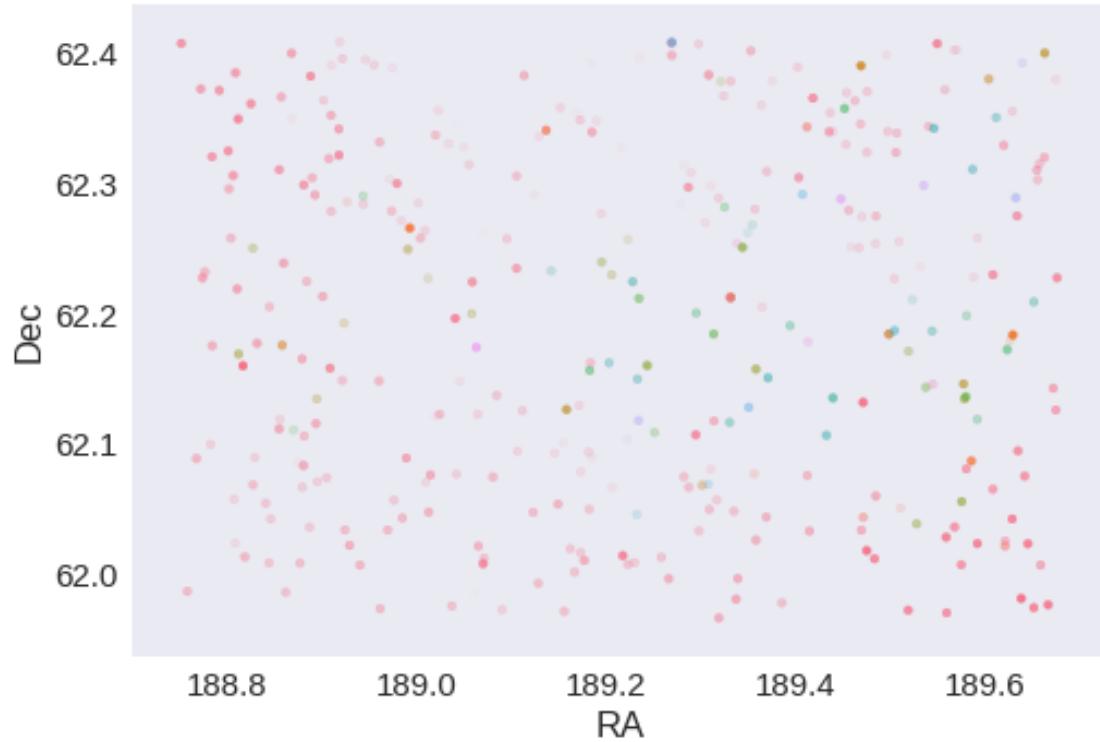
The cleaned catalogue has 48858 sources (0 removed).

The cleaned catalogue has 0 sources flagged as having been cleaned

1.5 III - Astrometry correction

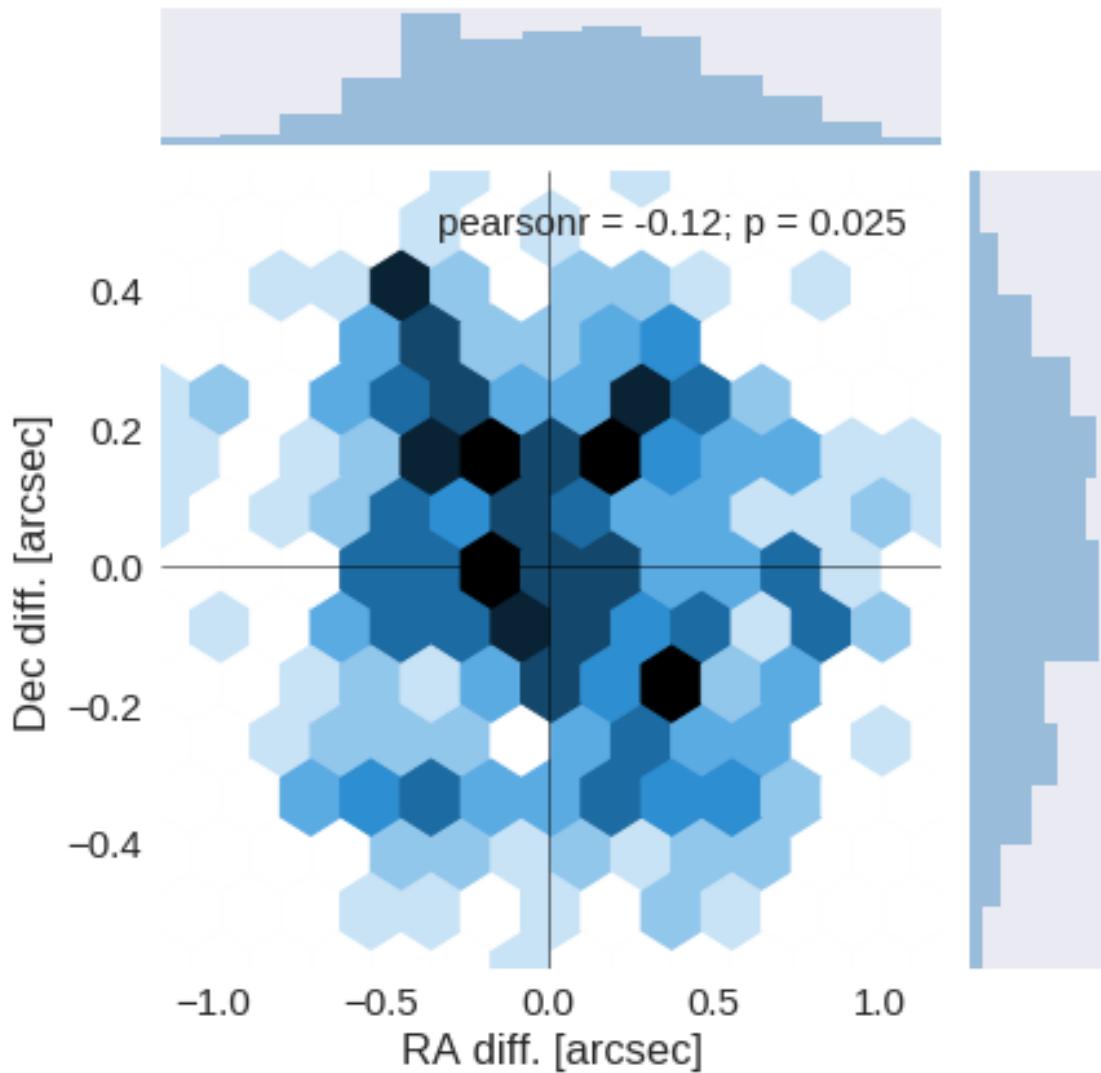
We match the astrometry to the Gaia one. We limit the Gaia catalogue to sources with a g band flux between the 30th and the 70th percentile. Some quick tests show that this give the lower dispersion in the results.

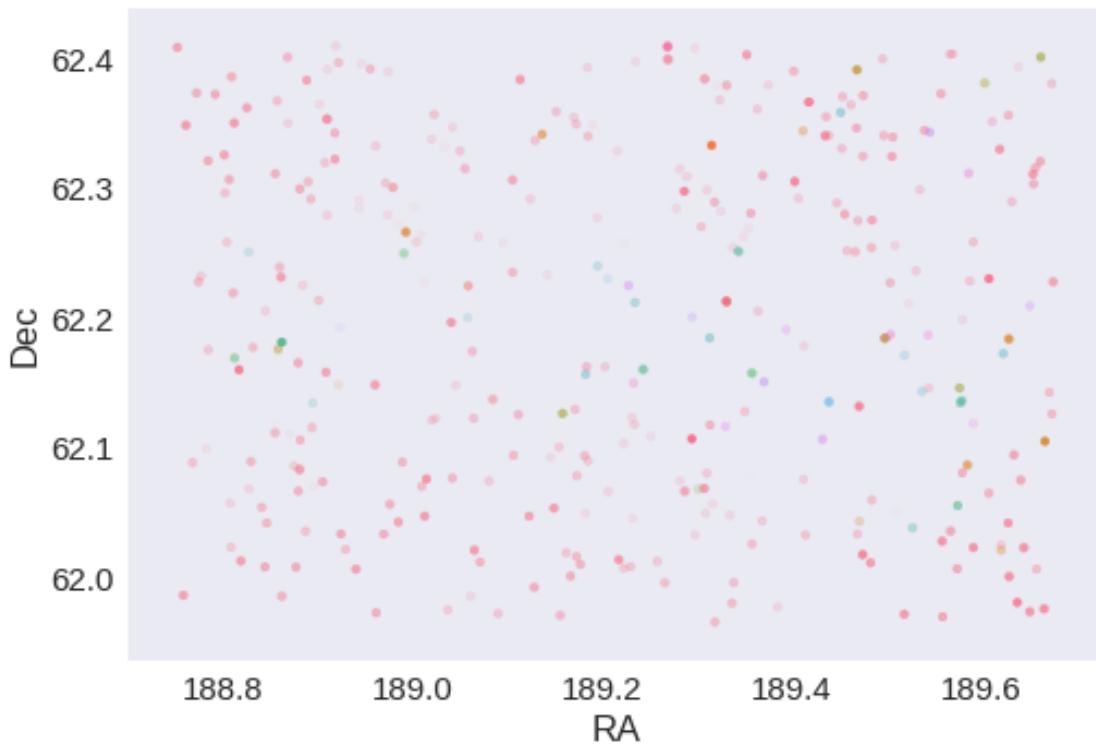




RA correction: -0.22843071726015296 arcsec

Dec correction: 0.014955391984017297 arcsec





1.6 IV - Flagging Gaia objects

431 sources flagged.

1.7 V - Flagging objects near bright stars

2 VI - Saving to disk

Out[21]: 48858

Out[22]: <IPython.core.display.HTML object>