


The fifth data release of the Kilo Degree Survey: Multi-epoch optical/NIR imaging covering wide and legacy-calibration fields

Angus H. Wright¹ , Konrad Kuijken² , Hendrik Hildebrandt¹ , Mario Radovich³ , Maciej Bilicki⁴,
Andrej Dvornik¹, Fedor Getman⁵, Catherine Heymans^{1,6}, Henk Hoekstra², Shun-Sheng Li² , Lance Miller⁷,
Nicola R. Napolitano⁸, Qianli Xia⁶, Marika Asgari^{9,10}, Massimo Brescia¹¹ , Hugo Buddelmeijer¹² ,
Pierre Burger¹³ , Gianluca Castignani^{14,15}, Stefano Cavuoti⁵ , Jelte de Jong^{2,16} , Alastair Edge¹⁷ ,
Benjamin Giblin^{6,18}, Carlo Giocoli^{15,19} , Joachim Harnois-Déraps²⁰, Priyanka Jalan⁴ , Benjamin Joachimi²¹,
Anjitha John William⁴, Shahab Joudaki^{22,23}, Arun Kannawadi²⁴ , Gursharanjit Kaur⁴ , Francesco La Barbera¹⁵,
Laila Linke^{13,25} , Constance Mahony¹ , Matteo Maturi²⁶, Lauro Moscardini^{14,15,27} ,
Szymon J. Nakoneczny^{28,29} , Maurizio Paolillo¹¹, Lucas Porth¹³ , Emanuella Puddu⁵ , Robert Reischke¹,
Peter Schneider¹³, Mauro Sereno¹⁵ , HuanYuan Shan^{30,31}, Cristóbal Sifón³², Benjamin Stözlner¹,
Tilman Tröster³³, Edwin Valentijn¹⁶ , Jan Luca van den Busch¹, Gijs Verdoes Kleijn¹⁶ , Anna Wittje¹,
Ziang Yan¹ , Ji Yao³⁰ , Mijin Yoon¹, and Yun-Hao Zhang^{2,6} 

(Affiliations can be found after the references)

Received 24 April 2023 / Accepted 27 November 2023

ABSTRACT

We present the final data release of the Kilo-Degree Survey (KiDS-DR5), a public European Southern Observatory (ESO) wide-field imaging survey optimised for weak gravitational lensing studies. We combined matched-depth multi-wavelength observations from the VLT Survey Telescope and the VISTA Kilo-degree Infrared Galaxy (VIKING) survey to create a nine-band optical-to-near-infrared survey spanning 1347 deg². The median *r*-band 5 σ limiting magnitude is 24.8 with median seeing 0.7". The main survey footprint includes 4 deg² of overlap with existing deep spectroscopic surveys. We complemented these data in DR5 with a targeted campaign to secure an additional 23 deg² of KiDS- and VIKING-like imaging over a range of additional deep spectroscopic survey fields. From these fields, we extracted a catalogue of 126 085 sources with both spectroscopic and photometric redshift information, which enables the robust calibration of photometric redshifts across the full survey footprint. In comparison to previous releases, DR5 represents a 34% areal extension and includes an *i*-band re-observation of the full footprint, thereby increasing the effective *i*-band depth by 0.4 magnitudes and enabling multi-epoch science. Our processed nine-band imaging, single- and multi-band catalogues with masks, and homogenised photometry and photometric redshifts can be accessed through the ESO Archive Science Portal.

Key words. gravitation – gravitational lensing: weak – catalogues – surveys – galaxies: photometry – cosmology: observations

1. Introduction

Modern astronomical sky surveys are intended to be useful for a wide array of science objectives, extending both beyond the remit of their primary science goals and the lifetime of the survey's operation. Such extended science use is typically referred to as the 'legacy' value of a survey. Given the ever-increasing monetary and temporal costs required to undertake large programmes at state-of-the-art observatories, the allocation of large programmes is increasingly contingent on the inclusion of planning for legacy science.

In order to take advantage of the full legacy value of any modern astronomical dataset, however, said dataset ought to follow the 'FAIR' principle: being Findable, Accessible, Interoperable (i.e. able to be easily combined with other data), and Reusable. Conformity with these principles is generally achieved by the storage and release of the data in an easily interfaceable system, with relevant accompanying documentation, such as in the European Southern Observatory (ESO) archive.

While all aspects of the FAIR principle are important, the 'reusable' principle is particularly relevant here. This manuscript details the fifth data release of the Kilo Degree Survey (KiDS; de Jong et al. 2013), with a particular focus on ensuring that the legacy value of the KiDS dataset is preserved. This includes the survey's on-sky overlap with the VISTA Kilo-degree Infrared Galaxy Survey (VIKING; Edge et al. 2013) and coverage of wide and deep spectroscopic survey fields.

KiDS is most notable as a so-called stage-III imaging survey for cosmology (using nomenclature defined in the Dark Energy Task Force white paper; Albrecht et al. 2006). There are three principal stage-III imaging surveys for cosmology: KiDS, the Dark Energy Survey (DES; Sevilla-Noarbe et al. 2021), and the Hyper-Suprime Camera (HSC) survey (Aihara et al. 2018). These surveys predominantly differ in their complementary combinations of imaging depth, survey area, and wavelength coverage. This is illustrated in Fig. 1: the HSC survey is the deepest of the stage-III surveys, DES has the largest sky coverage, and KiDS spans the greatest range in wavelength.

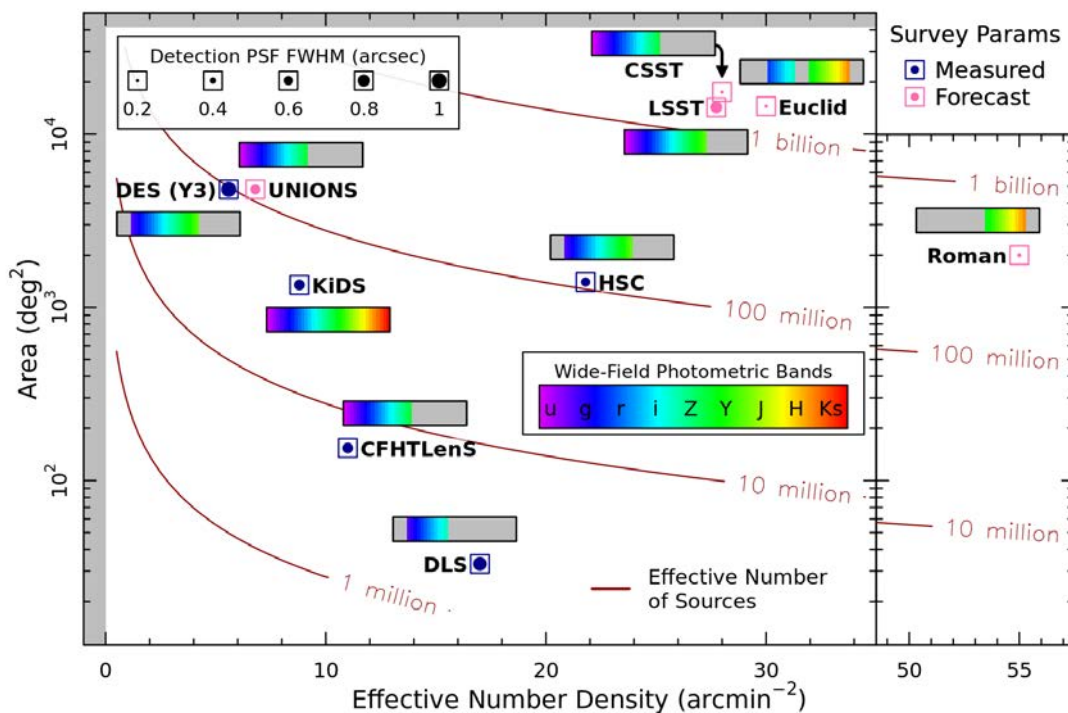


Fig. 1. Summary statistics for stage-II (Heymans et al. 2012; Jee et al. 2016), stage-III (this work; Sevilla-Noarbe et al. 2021; Hikage et al. 2018; Guinot et al. 2022), and stage-IV (Euclid Collaboration 2022; The LSST Dark Energy Science Collaboration 2018; Gong et al. 2019; Eifler et al. 2021) cosmological imaging surveys, comparing the survey area and effective number density of sources for weak lensing analyses (square markers). The red curves show the total effective number of sources, which principally determines the cosmological constraining power. The colour bars indicate each survey’s wavelength coverage, which principally determines the photometric redshift quality of a survey. The size of the circular point within each survey’s square marker shows the typical seeing in the lensing band, which principally determines the accuracy of shape measurements. The statistics presented for future surveys are forecasts, as indicated in pink. The final DES-Y6 data release will increase in depth, relative to the DES-Y3 statistics shown, raising the final effective number density of sources.

The stage-III surveys are directly comparable to earlier stage-II surveys, such as the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS; Heymans et al. 2012) and the Deep Lens Survey (DLS; Jee et al. 2016), which are notable due to their limited area but impressive photometric depth. Finally, the next generation of cosmological imaging surveys, stage-IV, stand out due to their combination of large areas and exceptional depths: *Euclid* (Laureijs et al. 2011; Euclid Collaboration 2022), the *Vera C. Rubin* Observatory Legacy Survey of Space and Time (LSST; Ivezić et al. 2019; The LSST Dark Energy Science Collaboration 2018), the Chinese Space Station Optical Survey (CSS-OS; Gong et al. 2019), and the *Nancy G. Roman* Telescope High Latitude Survey (*Roman*; Eifler et al. 2021). As with the stage-III surveys, each of these stage-IV surveys fills a particular niche combination of depth, resolution, wavelength coverage, and area on the sky. Figure 1 also includes the ground-based partner survey to *Euclid*, the Ultraviolet Near-Infrared Optical Northern Survey (UNIONS; Guinot et al. 2022), which has the capability to become a stage-III-like weak lensing survey in its own right.

As the final data production manuscript for the KiDS and VIKING surveys, this paper is designed as a one-stop reference for researchers and students who wish to use this dataset for their science. The manuscript therefore includes a full description of the production processes for all data products included in the fifth data release (DR5), including documentation of the input data, the analysis methods and settings, and the final data products.

The manuscript is arranged approximately in the order that the production itself takes place, which we describe in Sect. 2. Section 3 presents details of the optical reduction and

photometric processing. Section 4 presents details of the near-infrared (NIR) data reduction. Section 5 details the compilation of spectroscopic data in the KiDS DR5 fields. Section 6 presents details of the NIR photometric processing, the combination into the full multi-wavelength dataset, and the computation of photometric redshifts. Section 7 describes the shape measurements for weak gravitational lensing that are contained in the dataset, which will be used for cosmic shear analyses with the final KiDS DR5 weak-lensing sample, named the KiDS-Legacy sample. Finally, Sect. 8 outlines the contents of the full data release, detailing catalogues and imaging, important practical information for access and catalogue use, and summary statistics for the full dataset.

2. Production workflow and updates

To assist with the understanding of the production process (and with the navigation of this paper) we start by providing an overview of the production workflow, and direct readers to the relevant sections of the paper that describe the workflow stages. Details of surveys, technical concepts, algorithms, and other associated jargon are provided in the relevant sections, as are the appropriate references to the literature.

The fifth data release of KiDS (DR5) includes a collection of new observations, previously released data that have been re-reduced, and previously released data that are unchanged. As such, this release supersedes the previous releases of KiDS; the fourth data release (DR4) is not simply a subset of DR5. Of particular note in this release is the inclusion of an additional

30% of survey area, a complete second-pass of *i*-band observations, and 25 square degrees of imaging data over deep spectroscopic-survey fields (which are useful for photometric redshift calibration, and so are given the label ‘KiDZ’). Each of these additions has particular importance for scientific analyses with KiDS, as discussed in the following sections.

Figure 2 shows the updated data-production flowchart for DR5. The flowchart represents a slight development of a similar chart for DR4 presented in Fig. 1 of Kuijken et al. (2019), with added items for the processing of the KiDZ data, the NIR-reduction pipeline (KVPipe¹), the unified multi-band photometric processing pipeline (PHOTOPIPE²), and the final mosaic and compilation catalogue construction. Raw and input data products to the data-production pipeline are shown in yellow (accessible via e.g. the ESO archive), calibrated and released imaging data products are provided in green (accessible via the ESO archive and KiDS collaboration web pages), released per-tile catalogue data products are shown in pink (accessible via the ESO archive), and released mosaic data products are shown in blue (accessible via the KiDS collaboration web pages). The flowchart is annotated with the sections of the manuscript containing the description (and relevant changes with respect to previous work) related to each stage. We also outline the sections here for ease of reference.

Our production workflow starts, naturally, with raw observations made with the Very Large Telescope (VLT) Survey Telescope (VST; Sect. 3.1). Observations are then transferred to two separate reduction pipelines: the full set of photometric bands is sent to the ASTROWISE pipeline, and the *r*-band data are also sent to the THELI-Lens (for KiDS observations) or THELI-Phot (for KiDZ) pipelines. The THELI-Lens (Sect. 3.2) and THELI-Phot (Sect. 3.3) pipelines are responsible for producing the imaging (and associated products) that are used for source detection in the KiDS and KiDZ fields, respectively, while the THELI-Lens pipeline also produces calibrated individual detector images that are used by our fiducial shape measurement code *lensfit* (Sect. 7).

Within ASTROWISE (Sect. 3.5) we performed four main tasks: the primary KiDS and KiDZ source detection (Sect. 3.4), production of co-added images, weight-maps, and masks in each of the optical filters (Sect. 3.5), extraction of individual per-filter source catalogues (for stand-alone ‘single-band catalogues’, Sect. 3.5.5), and measurements of forced photometry using the Gaussian Aperture and PSF (GAAP) code on the optical images (Sect. 3.6). For the primary source extraction THELI-Lens and THELI-Phot *r*-band co-adds are ingested into ASTROWISE, and extraction is made on these images using SOURCE EXTRACTOR (Sect. 3.4). These catalogues form the basis of essentially all subsequent catalogues used in KiDS and KiDZ. For some use cases, however, it may be desirable to have individual source extractions in each of the optical filters (for example, for the detection of *r*-band drop-out sources). As such, the single band catalogues in each of the optical bands are also an important output of the reduction process. After the definition of the source lists, however, ASTROWISE falls back to the use of its own optical imaging for photometric measurements.

After the measurement of optical forced aperture photometry, the optical GAAP catalogues are passed to our unified multi-band photometry and photo-*z* measurement pipeline PHOTOPIPE (Sect. 6.1). This pipeline combines the optical data with

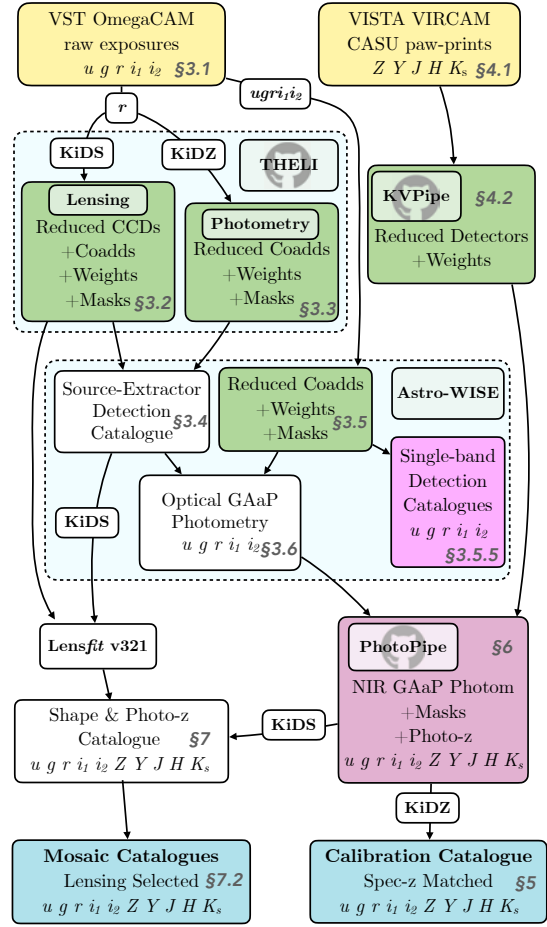


Fig. 2. Chart showing the data processing path that is taken for the KiDS and KiDZ data in DR5, from optical and NIR imaging (top) through to final mosaic catalogues (bottom). Each step outlined in this graph is discussed in the correspondingly annotated section. Yellow boxes show raw data products and data products input into our reduction pipelines. Green boxes contain imaging data products released as part of this data release. Pink boxes contain per-tile catalogue-level data products released as part of this data release. Blue boxes contain mosaic catalogue-level data products released as part of this data release.

the NIR images from the VIKING survey, starting with the ‘paw-prints’ provided by the Cambridge Astronomical Survey Unit (CASU). For all sources in KiDS and KiDZ, PHOTOPIPE performs the forced photometry of the NIR bands (Sect. 6.2), construction of NIR mosaics, combination of optical and NIR photometric catalogues, estimation of photometric redshifts (Sect. 6.3), and construction of combined multi-band masks (Sect. 6.4). These catalogues are the primary data product that is released to ESO (Sect. 8). For weak-lensing applications, however, we required robust shape measurements: for tiles in the main KiDS footprint, these are performed with the *lensfit* algorithm (Sect. 7).

Combined ‘mosaic catalogues’ are then prepared for both the KiDS and KiDZ datasets. These are the primary science catalogues that are used by the majority of the cosmology team in KiDS. In the KiDS (i.e. main survey) areas, these catalogues are constructed by combining all per-pointing catalogues after source selection based on masks, shape measurements, photometry, and more (Sect. 7.2). In the KiDZ fields, a final ‘calibration catalogue’ is constructed by combining all catalogues from the KiDZ fields and matching them to a master spectroscopic compilation (Sect. 5). These mosaic and calibration catalogues

¹ <https://github.com/AngusWright/VIKINGProcessingPipeline>

² <https://github.com/KiDS-WL/PhotoPipe>

Table 1. Summary of imaging data released in KiDS-DR5.

Telescope and Camera	Filter	λ_{cen} (Å)	Exp. time (s)	Mag. lim. (corr) (5σ 2'' AB)	Mag. lim. (uncorr) (5σ 2'' AB)	PSF $FWHM$ (")	Area (deg ²)
VST (OmegaCAM)	<i>u</i>	3 550	1000	24.17 ± 0.10	24.26 ± 0.10	1.01 ± 0.17	1 312.1
	<i>g</i>	4 775	900	24.96 ± 0.11	25.15 ± 0.12	0.88 ± 0.15	1 270.7
	<i>r</i>	6 230	1800	24.79 ± 0.13	25.07 ± 0.14	0.70 ± 0.12	1 233.8
	<i>i</i> ₁	7 630	1200	23.41 ± 0.26	23.66 ± 0.25	0.81 ± 0.18	1 246.2
	<i>i</i> ₂	7 630	1200	23.49 ± 0.28	23.73 ± 0.30	0.81 ± 0.18	1 264.4
VISTA (VIRCAM)	<i>Z</i>	8 770	480	23.47 ± 0.20	23.79 ± 0.20	0.90 ± 0.10	1 262.0
	<i>Y</i>	10 200	400	22.72 ± 0.18	23.02 ± 0.19	0.86 ± 0.09	1 261.4
	<i>J</i>	12 520	400	22.53 ± 0.20	22.72 ± 0.20	0.85 ± 0.07	1 262.7
	<i>H</i>	16 450	300	21.94 ± 0.24	22.27 ± 0.24	0.88 ± 0.09	1 258.3
	<i>K</i> _s	21 470	480	21.75 ± 0.19	22.02 ± 0.19	0.87 ± 0.08	1 258.1

Notes. Limiting magnitudes are constructed in two ways as described in Appendix A. One method accounts for pixel-to-pixel correlations in the noise background ('corr'), but is likely contaminated by contamination by source flux and therefore underestimates the true depth of the imaging. The second method does not account for pixel-to-pixel correlations in the noise background ('uncorr') but it is robust to contamination from sources, and therefore likely overestimates the true depth of the imaging. These estimates therefore bracket the true depth of the imaging released in KiDS-DR5. For quoted limiting magnitudes and PSF sizes, the uncertainties describe the variations in these metrics between the 1 deg² fields on-sky.

are released to the community via the KiDS database³ after the conclusion of the primary cosmological analyses.

Finally, for convenience of reference, we also provide here a summary of the data release in terms of imaging quality and other relevant statistics (Table 1).

3. Optical dataset and reduction

In this section we detail the optical portion of the KiDS DR5 dataset, its reduction, and its differences with the equivalent data released in DR4. Of particular note: new data have been observed with the VST (Sect. 3.1), THELI-Lens data have been fully re-reduced (Sect. 3.2), new KiDZ data have been reduced with the THELI-Phot pipeline (Sect. 3.3), source detection has been recomputed for the full dataset (Sect. 3.4), ASTROWISE data have been partially re-reduced (Sect. 3.5), and multi-band GAAP photometry has been re-extracted and the post-processing updated (Sect. 3.6). As a result, despite DR5 representing largely the union of existing (from Kuijken et al. 2019) and new ASTROWISE imaging, it is unlikely that source-lists, photometry, and/or higher-level data products will be identical (when comparing DR4 to DR5) for any individual source, population, or tile.

3.1. VST observations

The VST (Capaccioli et al. 2005) is a 2.6 m modified Ritchey-Chrétien telescope on an alt-az mount, located at the ESO's Cerro Paranal observatory in Chile. The VST hosts a single instrument at its Cassegrain focus: the 300-megapixel OmegaCAM CCD mosaic imager (Kuijken 2011). The 32 thinned CCDs that make up the 'science array' of the camera consist of 4102 × 2048-pixel e2v 44 – 82 devices, which uniformly sample the focal plane at a scale of 0'214 per 15 μm pixel and have very narrow chip-gaps (25'' and 85''). The telescope and camera were co-designed for optimal image quality, and as a result they are capable of producing imaging with a point-spread function (PSF) that is quite round and shows little variation across the focal plane: seeing ellipticities (defined as $e = 1 - b/a$, where a

Table 2. Run numbers of all VST observations taken in the KiDS and KiDZ fields.

Survey	Run number	Start date	Number of OBs
KiDS	60.A-9038	2011-08-09	1
	177.A-3016	2011-08-26	6618
	177.A-3018	2011-09-19	63
	177.A-3017	2011-10-02	48
	094.B-0512	2014-11-20	4
	0103.A-0181	2019-05-03	1
KiDZ	177.A-3016	2012-11-25	105
	096.B-0501	2015-11-06	10
	098.B-0298	2016-11-21	6
	0100.B-0148	2017-11-24	4

Notes. Run numbers include observations taken as part of guaranteed time observations programmes.

and b are the major and minor axes of the PSF, respectively) across all KiDS VST observations are consistently less than $\langle e \rangle = 0.08$ per exposure.

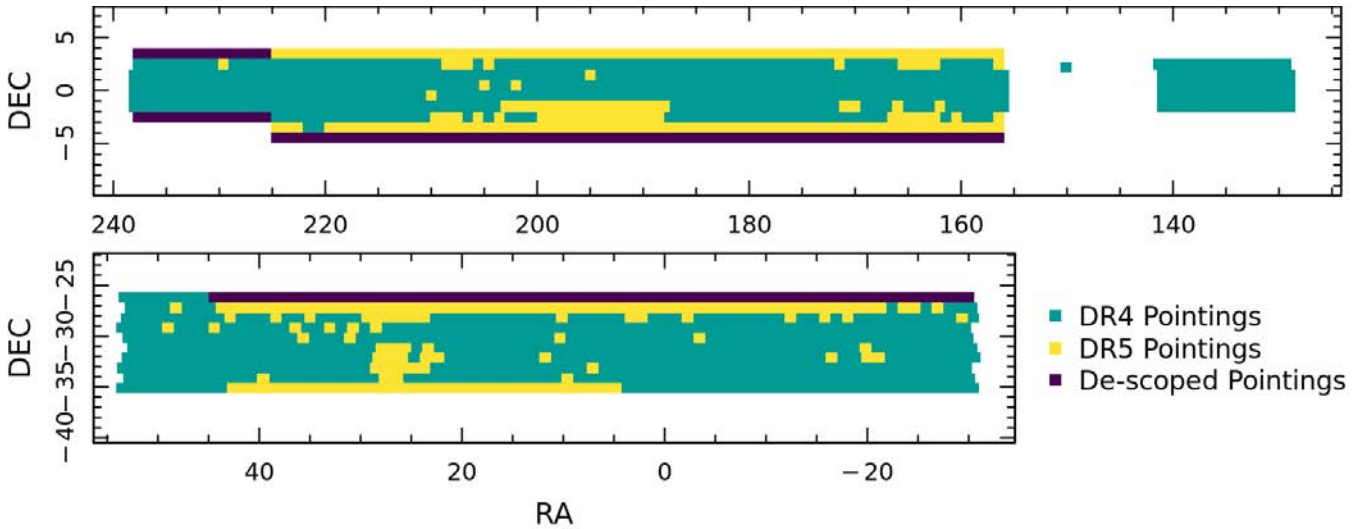
Observations with the VST are split into observing blocks (OBs), which each target a single tile with dithered observations using a particular filter, and which were queued in order to optimise the use of the VST (as it simultaneously carried out a suite of surveys). Observations for KiDS and KiDZ made with the VST were performed under the programmes listed in Table 2. Each tile was observed in four filters (*ugri*), with band-passes very similar to the *ugri* filters from the Sloan Digital Sky Survey (SDSS; York et al. 2000)⁴, and with various requirements regarding observational conditions. Table 3 summarises the observing constraints, number of OBs, number of dithers per OB, and total exposure time for each filter. *r*-band observations were made under only the best atmospheric conditions (seeing full-width at half-maximum of less than 0'8) during dark-time (maximum lunar illuminations of 40%). The *g*- and *u*-band observations similarly required dark-time, but with less stringent

⁴ The effective wavelengths of the VST and SDSS filters agree to approximately 2, 6, 100, and 50 Å in the *ugri*-bands, respectively.

³ <https://kids.strw.leidenuniv.nl/sciencedata.php>

Table 3. KiDS+KiDZ observing strategy: observing condition constraints and exposure times.

Filter	Max. lunar illumination	Min. Moon distance [deg]	Max. seeing [arcsec]	Max. airmass	Sky transp.	Dithers	Total exp. time per OB [s]	Num OBs
<i>u</i>	0.4	90	1.1	1.2	CLEAR	4	1000	1
<i>g</i>	0.4	80	0.9	1.6	CLEAR	5	900	1
<i>r</i>	0.4	60	0.8	1.3	CLEAR	5	1800	1
<i>i</i>	1.0	60	1.1	2.0	CLEAR	5	1200	2

**Fig. 3.** Distribution of KiDS DR5 pointings on-sky. The figure shows the distribution of DR4 pointings (dark cyan) and new DR5 pointings (yellow). Pointings that were originally included in the 1500 deg² KiDS footprint, but which were subsequently de-scoped due to the limited area observed by VIKING, are shown in dark purple. The combination of the green and yellow data therefore show the 1347 deg² of KiDS DR5 observations.

seeing requirements: allowing maximal seeing of $FWHM \leq 0''.9$ and $FWHM \leq 1''.1$, respectively. The *i*-band observations were required to be performed under the same seeing conditions as the *u*-band ($FWHM \leq 1''.1$), but, unlike the other bands, they could be obtained in grey or bright-time (i.e. with no restriction on maximum lunar illumination).

The observing strategy for KiDS and KiDZ can best be described as ‘depth first, area second’: each OB immediately achieves the full depth in a given filter on a tile, and that tile is not revisited with that filter unless the observations subsequently fail quality control. The one exception to this rule is the second *i*-band pass, which is revisited by design (see Sect. 3.1.1). There is no requirement that the observations with a different filter on a given tile be taken within a certain time period, although a bias towards observing partially completed tiles was built into the observing scheduler.

In November 2020, towards the end of the survey data taking and after the VST was recommissioned following an extended shutdown during the COVID-19 pandemic, one of the 32 CCDs in the OmegaCAM mosaic was found to have failed. This resulted in a rectangular gap near the centre of the mosaic for all observations taken after this date. This defect affects mostly the second *i*-band pass of the KiDZ fields, which were almost exclusively observed after this date (Sect. 3.1.2).

3.1.1. KiDS data acquisition

The KiDS survey footprint consists of 1347 tiles, of 1 deg × 1 deg each, corresponding to the field of view of the OmegaCAM CCD

mosaic camera. It covers two regions of the sky, roughly coincident with the main survey area of the 2dF Galaxy Redshift Survey (Colless et al. 2001): a 70 deg × 9 deg stripe (in RA × Dec) across the South Galactic Pole, and a 85 deg × 8 deg stripe across the celestial equator in the North Galactic Cap, with extensions to include the Galaxy and Mass Assembly (GAMA) 9hr field (12 deg × 5 deg, Driver et al. 2011) and the Cosmic Evolutions Survey field (COSMOS; 1 deg × 1 deg, Scoville et al. 2007). The layout of the KiDS survey fields is shown in Fig. 3.

The original plan for the KiDS survey, as endorsed by ESO’s Public Surveys Panel in 2005, was to cover 1500 square degrees in *ugri*-bands with the VST (de Jong et al. 2013). This proposed area is shown in Fig. 3 as the full span of coloured tiles. Subsequently, in 2006, ESO approved the VIKING survey (Edge et al. 2013), which would add five NIR filters to the same proposed survey area. The observations for KiDS and VIKING started in 2011 and 2009, respectively, soon after their corresponding telescopes were commissioned.

The initial VIKING time allocation resulted in about 90% of the planned area of sky being observed, at which point ESO decided to truncate that survey in favour of new projects (see Sect. 4.1.1). Given the slightly reduced area that VIKING would cover, it was decided also to truncate KiDS to the same total area, which resulted in the final footprint of 1347 tiles. Observations that were removed after this area truncation are shown in purple in Fig. 3. Some of the observing time on the VST originally allocated to the now-deprecated KiDS and VIKING tiles was reallocated to observations of deep spectroscopic fields, forming the basis of the ‘KiDZ’ dataset (Sect. 3.1.2).

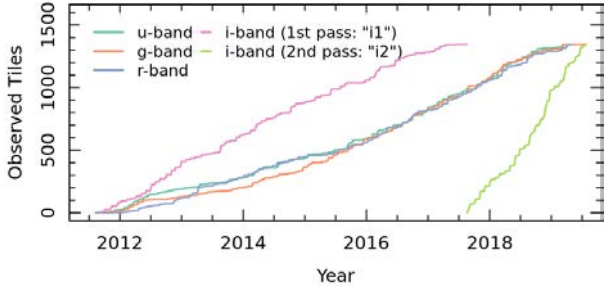


Fig. 4. Progression of optical observations over the full KiDS survey, for pointings within the final 1347 deg² footprint. The total number of DR5 pointings is shown as the dashed grey line. The figure demonstrates the slow initial progress made by the survey, whereby only ~5% of planned *r*-band observations were completed in the initial 18 months of data-taking, prompting changes in the queuing of (in particular) the dark-time observations. Conversely, at the end of the survey, the complete second-pass of *i*-band observations (designated *i*₂) was completed in roughly the same time span.

A second component of the original KiDS survey plan was to re-observe the entire footprint at the end of the survey in the *g*-band, for proper motion and variability studies. As the survey progressed, however, it became clear that bright-time was in considerably less demand than dark-time on the VST, with the bright-time *i*-band observations advancing much faster than those using the dark-time *u*, *g*, *r* filters. This can be seen in Fig. 4, which shows the progression of observations in each of the optical filters over the lifetime of the KiDS survey. Given the rapid progression of the *i*-band data, it was decided to change the filter for the repeat pass to *i*, rather than *g*. The second pass *i*-band observations are visible in the figure under the *i*₂ label. Similarly the *i*-band is unique in Table 3 as being the only filter with more than one OB per tile.

The choice to switch the filters used for the second pass comes with certain scientific benefits. First, the additional depth in the *i*-band improves the quality of photometric redshifts (particularly beyond $z = 0.9$, where the 4000 Å break redshifts into the *i*-band). Second, the repeat observations somewhat counteract the variability in the seeing and sky brightness that is inherent to the *i*-band observations, due to their less stringent constraints on observing conditions. Finally, all the second-pass observations were taken after significant improvements to the telescope baffling in 2015, much reducing scattered light that particularly affected bright-time observations. Due to the rapidity of the *i*-band progress during the first half of the survey, the scattered light effect is particularly prevalent in the *i*-band. Therefore, having the second pass be in this band is of particular benefit.

Given the primary focus of the repeated *i*-band observations on variability and transient science, it is worth noting the somewhat correlated nature of the cadence between *i*-band observations on-sky. As early KiDS data targeted the GAMA fields preferentially, the largest baselines in the *i*-band observation cadence are in those fields. This can be seen in Fig. 5, which shows the difference between *i*-band observation dates over the survey, with the rectangular GAMA patches (dashed outlines) clearly visible.

A second noteworthy feature of Fig. 4 is the particularly slow progress of observations taken in the *r*-band during the first 18 months of data-taking. This feature was traced to the queue scheduler, which was overly conservative for dark-time observations with strict seeing requirements. Changes to the scheduler in early 2013 addressed this problem and led to a speeding up of the dark-time KiDS observations.

3.1.2. KiDZ acquisition

As mentioned in Sects. 3.1 and 3.1.1, DR5 includes observations made by the VST over deep spectroscopic survey fields that significantly enhance KiDS science. Eight fields were identified for these observations, as having both rich existing (or ongoing) spectroscopic campaigns, and being visible from Paranal: *Chandra* Deep Field South (CDFS), the COSMOS field, the DEEP2 Galaxy Redshift Survey (DEEP2) 02 and 23 h fields, the GAMA G15-Deep field, the Visible Multi-object Spectrograph (VIMOS) Public Extragalactic Redshift Survey (VIPERS) W1 and W4 fields, and the VIMOS VLT Deep Survey (VVDS) 14hr field. Specific details of the spectroscopic samples available in these fields are provided in Sect. 5. Of these fields, two already resided completely within the KiDS DR5 footprint (CDFS, G15-Deep) and so required no additional observations. KiDZ observations with the VST were taken under the same strategy as was used in the main survey (Sect. 3.1 and Table 3), including the second *i*-band pass, and VIKING-like observations on these fields were taken with VISTA. The VST KiDZ observations were taken between November 2015 and September 2021, and can be identified in the ESO archive using the run numbers listed in Table 2. The full set of targeted fields are shown in Fig. 6, with VST OBs visible by the black outlines.

3.1.3. Optical data quality metrics

The two primary quality metrics that are relevant for our KiDS+KiDZ observations, made with the VST, are the size of the PSF and the brightness of the background. Combined, these two properties determine the effective depth of the imaging, and the selection of source galaxies. Figure 7 shows the distribution of PSF size (full-width at half-maximum) and limiting magnitude (5σ in a 2'' diameter circular aperture) for all images taken as part of the KiDS+KiDZ observing campaigns. Limiting magnitudes in each of the bands were calculated from randomly scattered apertures across the masked tiles using LAMBDA (Wright et al. 2016). The details of the magnitude limit calculation is given in Appendix A.

The figures show both the distribution functions and cumulative distribution functions for each parameter. The median seeing in each optical filter is $\{1''.01 \pm 0''.17, 0''.88 \pm 0''.15, 0''.70 \pm 0''.12, 0''.81 \pm 0''.18, 0''.81 \pm 0''.18\}$ for the $\{u, g, r, i_1, i_2\}$ -bands, respectively, and the corresponding median limiting magnitudes are $\{24.17 \pm 0.10, 24.96 \pm 0.11, 24.77 \pm 0.13, 23.41 \pm 0.26, 23.49 \pm 0.28\}$ ⁵ In both cases, the uncertainties describe the tile-to-tile scatter in these metrics, computed with the normalised median absolute deviation from the median (NMAD). The distribution of magnitude limits in the four bands can be seen to take a reasonably narrow range of values; the scatter of the limiting magnitudes in the *ugr*-bands is consistently less than 0.15 mag. The *i*-band, however, does show more variability in both seeing and limiting magnitude than is evident in the other bands, largely due to these observations being taken in poorer conditions (Sect. 3.1). However, these metrics keep the independent *i*-band passes separate, meaning that the effective limiting magnitude of the individual tiles in the *i*-band would in-fact be approximately 0.4 magnitudes deeper and have a 30% reduction in scatter, bringing it closer in line with the other bands (from $\sigma \approx 0.3$ mag to $\sigma \approx 0.2$ mag). We note that this approximation

⁵ These magnitude limits are slightly brighter than those presented in previous KiDS releases, due largely to their computation allowing for inclusion of correlated noise. For a direct comparison with previous releases, see Appendix A and Table 1.

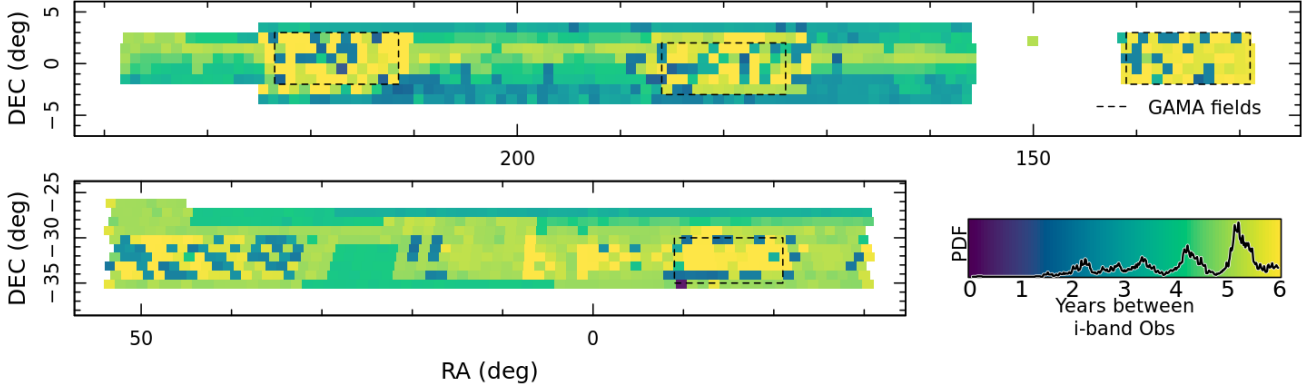


Fig. 5. Temporal separation between the two i -band passes, as a function of position on-sky. The figure shows all KiDS DR5 pointings, coloured by the number of years separating the two i -band passes. The overall distribution of these values is shown as a KDE within the colour bar, computed with a 0.1-year rectangular bandwidth. The largest temporal separations occur typically in the GAMA fields, where KiDS observations were focussed during early survey operations. The shortest temporal separation between any two passes is at RA = 350.3 deg, Dec = -35.1 deg, which was observed on 21 August 2017 and 7 October 2017 for the two i -band passes, respectively, for a separation of 50 nights.

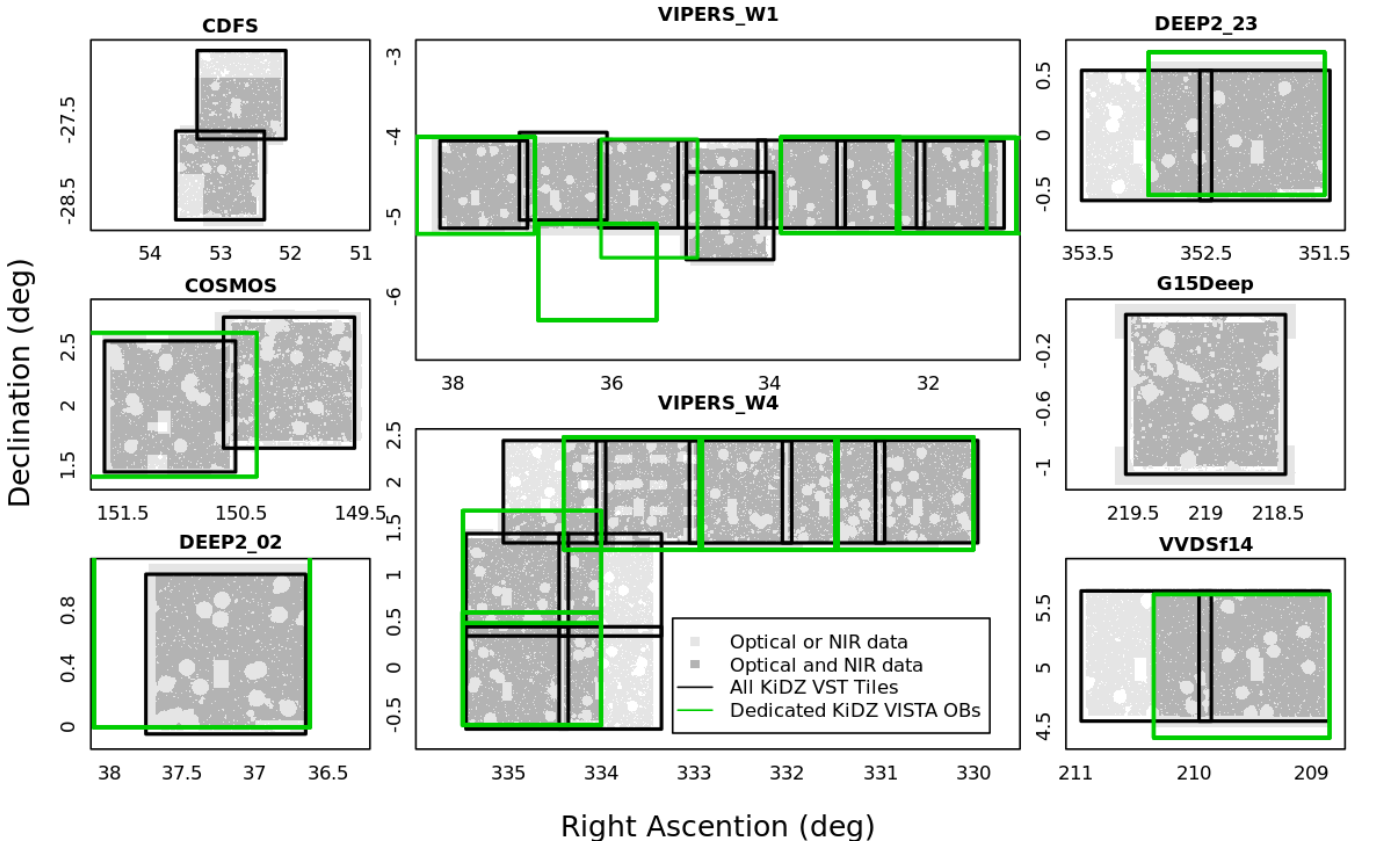


Fig. 6. Footprints of the KiDZ pointings. The figure shows the regions where there exist optical and/or NIR data in dark and light grey, respectively. The rectangles show the limits of all observations made with the VST (black) and dedicated observations made with VISTA (green). Dark grey regions that are not covered by our dedicated VISTA observations are either already contained within the VIKING footprint (CDFS, G15-Deep) or have VIKING-like observations constructed from existing deep observations (VIPERS, COSMOS; see Sect. 4.1.3).

is largely fair, as the individual distributions of PSF full-width at half-maximum (FWHM) and limiting magnitudes are extremely similar for the i_1 and i_2 passes.

Finally, systematic variation of observational conditions on-sky is an important nuisance in many scientific studies. As such, we include visualisations of the on-sky distribution of both magnitude limit and PSF FWHM, for each of the optical filters, in Appendix B.

3.2. *theli*-Lens r -band reduction

As with previous KiDS analyses, the primary imaging used for lensing science (specifically shape measurement) in DR5 has been produced with the THELI-Lens pipeline (version 1.3.0A). Since DR4 (which used THELI-Lens version 1.0.0A), THELI-Lens has undergone modifications to both the astrometric and photometric calibration routines. In this section we document, in

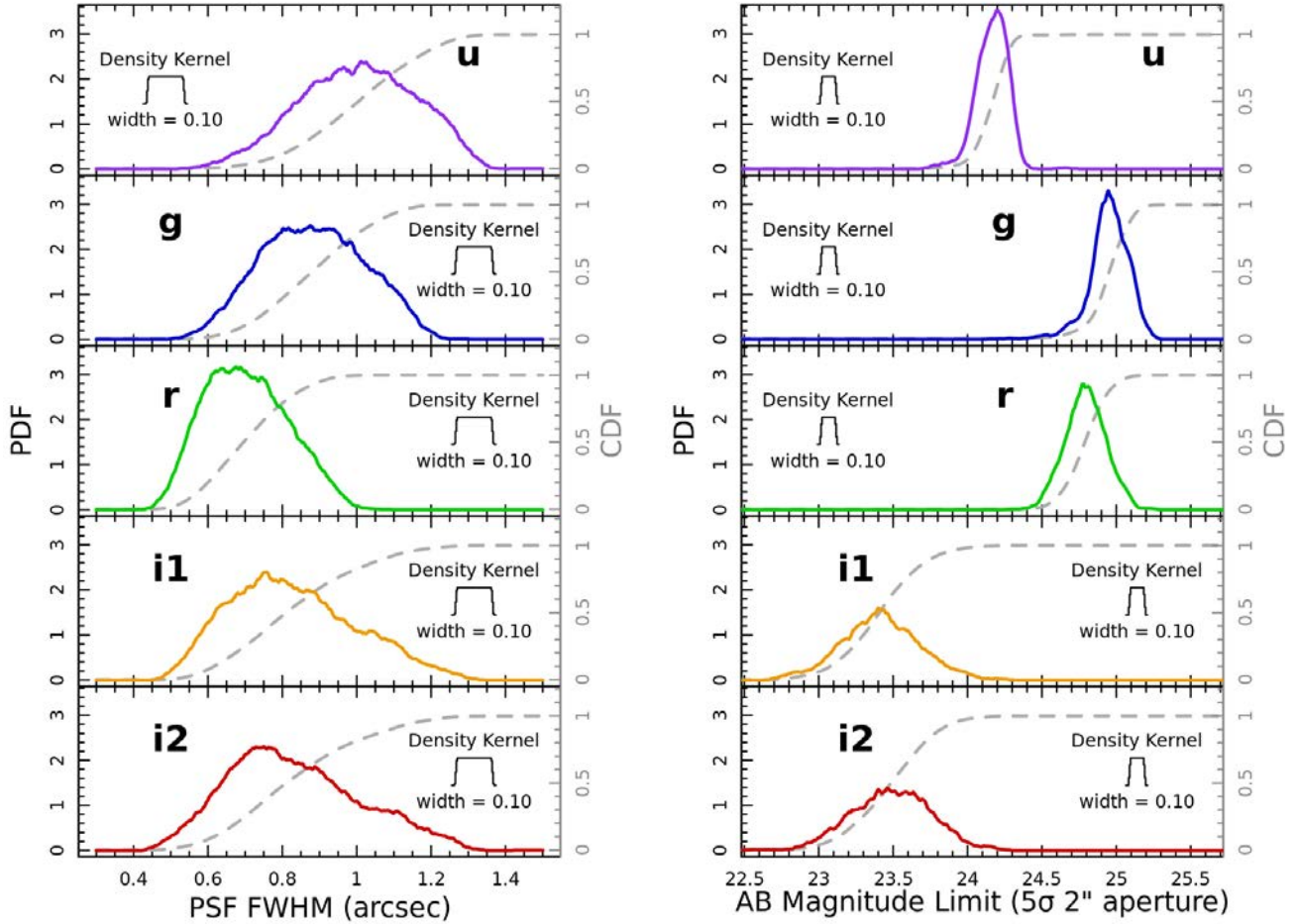


Fig. 7. Primary observational properties of pointings in KiDS+KiDZ for observations in the four optical bands. Observations in the *i*-band are split into the two epochs (labelled i_1 and i_2). Each row shows the distribution of average PSF sizes (as reported by ASTROWISE; left) and limiting magnitude (5σ in a $2''$ diameter circular aperture; right) determined with a KDE using the annotated kernel. The corresponding cumulative distribution functions for each panel are shown as grey dashed lines.

particular, the specific changes to THELI-Lens (with respect to the DR4 version) that are relevant to DR5 science.

Astrometric calibration with *Gaia*

Until version 1.3.0A, THELI-Lens utilised two separate astrometric calibration samples for the northern and southern patches of KiDS: in the north SDSS (York et al. 2000) was used as an absolute astrometric reference, whereas the southern fields were absolutely calibrated to 2MASS (Skrutskie et al. 2006). In DR5, KiDS utilises a single astrometric calibration to *Gaia* DR2 (Gaia Collaboration 2018) stars, with positions translated to epoch J2000 by Vizier⁶. We note that the choice to calibrate KiDS to a J2000 epoch can lead to some confusion when performing a direct match between KiDS stars and *Gaia* catalogues that do not contain the RAJ2000/DEJ2000 columns. For an explanation of this issue, we direct the interested reader to Appendix C.

Figure 8 shows the accuracy of the KiDS astrometric calibration to *Gaia* (here using *Gaia* DR3 stars, rather than the DR2 stars used to define the astrometric solutions), as a function of RA and Dec. Average absolute residuals are typically less than $0''.05$, demonstrating that the imaging has been successfully

tied to *Gaia*. The small residuals that remain are systematically induced by the choice to calibrate KiDS using *Gaia* stars at epoch J2000 positions: as KiDS observations were made after the year 2000, stellar proper motions introduce noise to the astrometric calibration. This is seen most clearly in the on-sky distribution of astrometric residuals, shown in Fig. 9: parts of KiDS that were observed earliest (i.e. the GAMA fields) clearly stand-out as having systematically smaller residuals than other (later) observations.

Finally, we note that the change to the *Gaia* astrometric calibration introduces another possible source of bias: differing astrometric solutions between our THELI-Lens and ASTROWISE reductions. As ASTROWISE astrometry is computed with 2MASS, it is possible that the locations of sources extracted from THELI-Lens images may have systematically different positions on-sky (as determined by our *Gaia* astrometry) to the same sources in ASTROWISE (as determined by the 2MASS astrometry). To verify the consistency of the astrometry between our THELI-Lens and ASTROWISE reductions, we matched sources extracted from the reduced ASTROWISE *r*-band images (see Sect. 3.5.5) to those in our master THELI-Lens source catalogue (Sect. 3.4). Figure 10 shows the astrometric agreement between stars in these two catalogues, thereby validating the use of THELI-Lens defined positions for photometric extractions on ASTROWISE images. The median residual between stars in the

⁶ Catalogue I/345/gaia2, columns RAJ2000 and DEJ2000.

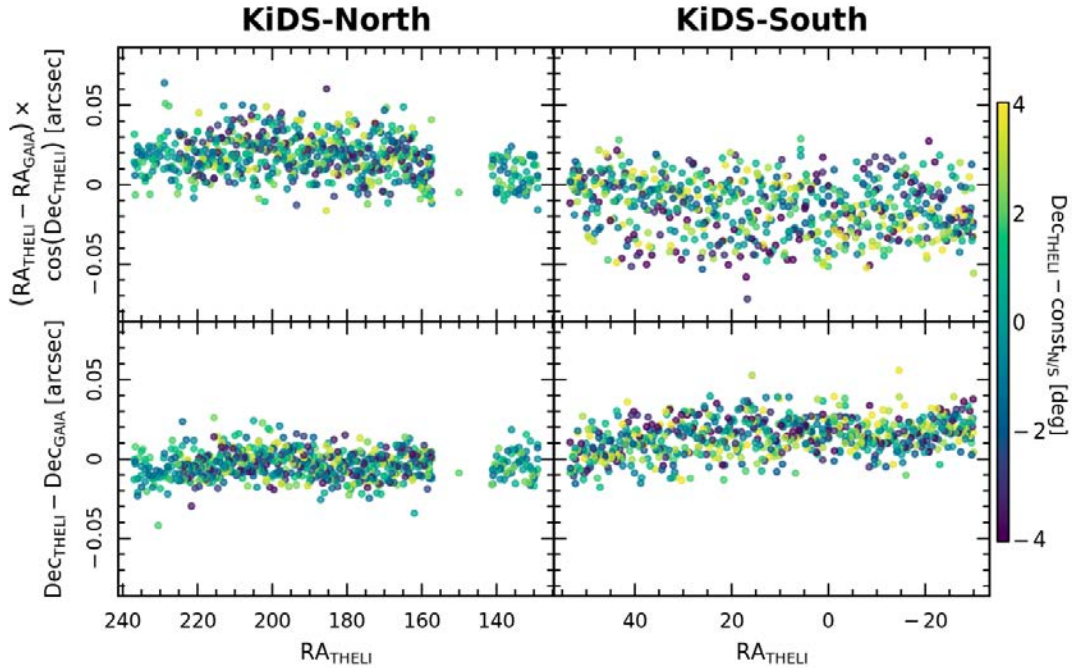


Fig. 8. Astrometric calibration of KiDS DR5, with respect to *Gaia* DR3 stars at epoch J2000. Values here show the median astrometric residual per pointing for stars in the magnitude range $16.5 \leq G \leq 19$. The colour bar shows the location of each field in the Dec direction, after subtracting a constant equal to the mean declination of all fields in the relevant hemisphere. Absolute residuals are typically less than the $0''.05$ level, and are therefore negligible. Residual offsets below this level are possibly attributable to barycentric motion between the J2000 and J2015 epochs (see Appendix C).

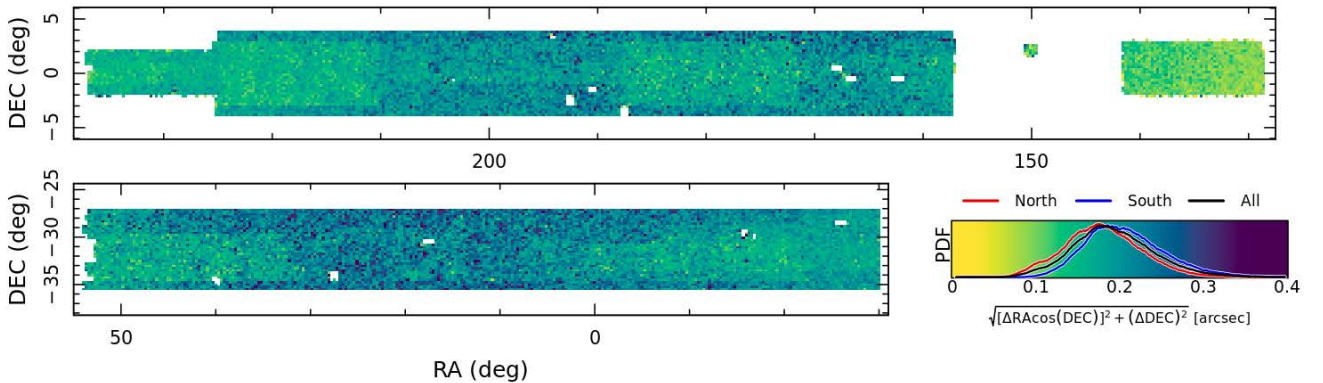


Fig. 9. On-sky variation in the KiDS DR5 astrometric calibration with respect to *Gaia* stars at the J2000 epoch. Residuals can be seen to fall around the earliest observations, due to the reduced proper motion differences between our observations and the assumed epoch.

two catalogues is $0''.014$ in the RA axis and $-0''.017$ in the Dec axis. The NMAD scatter in the RA and Dec directions is $0''.097$ and $0''.090$, respectively.

3.3. *Theli-Phot* *r*-band reduction

As mentioned in Sect. 3.1.2, VST observations in the *r*-band were made in the KiDZ fields between November 2015 and September 2018. These observations were reduced with a slightly different version of the THELI pipeline, as the observations were not planned to be used for lensing (rather they would only be used for photometry in the context of redshift distribution calibration). This different version, which we designate ‘THELI-Phot’, is notable for its different astrometric calibration sample (2MASS) and the lack of individual exposures (which are used for shape measurement).

In practical terms, the THELI-Phot pipeline is based on THELI-Lens version 1.2.0A, with changes made to the data products that are preserved at the end of each pointing’s reduction. As mentioned previously, the primary distinction between these V1.2.0A and V1.3.0A is the use of *Gaia* as an astrometric reference (in V1.3.0A), as opposed to the use of SDSS (north) and 2MASS (south) in V1.2.0A. This means that the astrometric solution for KiDZ and KiDS are nominally different, at the level of existing systematic differences between *Gaia* and SDSS or 2MASS. Figure 11 shows the residual difference between the astrometric solutions of KiDS stars (i.e. calibrated with *Gaia*) with respect to SDSS stars, demonstrating that there is no significant systematic bias in the two astrometric systems. As such, we conclude that there is unlikely to be any significant systematic effect imprinted on the KiDZ data through the use of a different astrometric basis in the THELI-Phot pipeline.

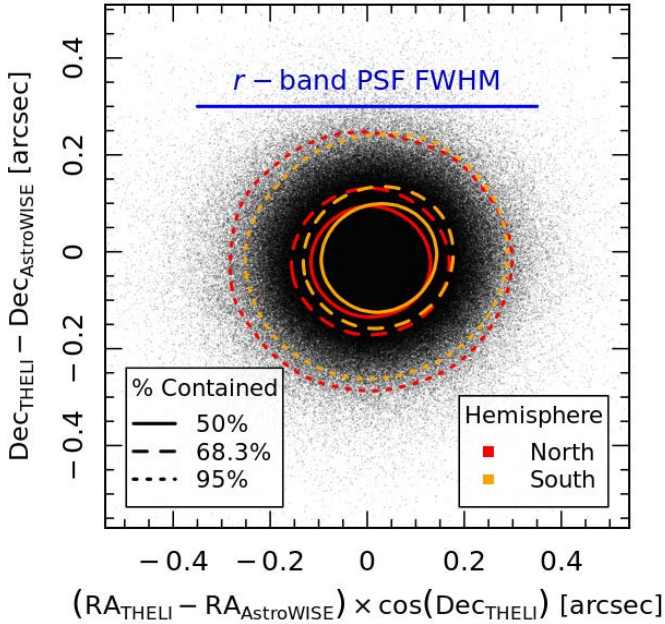


Fig. 10. Astrometric agreement between sources extracted from THELI-Lens and ASTROWISE images. The sample of all sources has a median offset of $0''.014$ and $-0''.017$ in the RA and Dec directions, respectively. The NMAD scatter in the RA and Dec directions is $0''.097$ and $0''.090$, respectively.

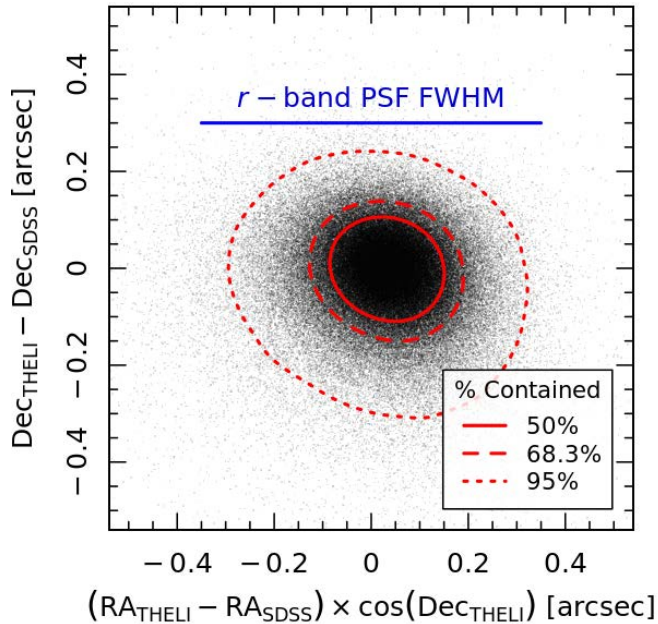


Fig. 11. Overall astrometric calibration of KiDS DR5, calibrated to *Gaia*, with respect to SDSS in the KiDS-N field. The sample of all sources has a median offset of $0''.031$ and $-0''.006$ in the RA and Dec directions, respectively. The NMAD scatter in the RA and Dec directions is $0''.094$ and $0''.084$, respectively.

3.4. Source detection

Source detection in KiDS+KiDZ is performed with SOURCE EXTRACTOR (Bertin & Arnouts 1996), within the ASTROWISE environment. The sources are extracted from the THELI-Lens and THELI-Phot *r*-band images, and these source lists are subsequently used for all primary science in KiDS. Source extraction is performed in a relatively ‘hot’ mode (i.e. significant

Table 4. Source extraction parameters used in the creation of the KiDS DR5 source lists within ASTROWISE.

Parameter	Value
ANALYSIS_THRESH	1.5
DEBLEND_MINCONT	0.001
DEBLEND_NTHRESH	32
DETECT_MINAREA	4
DETECT_THRESH	1.5
FILTER	Y
FILTER_NAME	default.conv
THRESH_TYPE	RELATIVE
CLEAN	Y
CLEAN_PARAM	3.0
BACK_TYPE	MANUAL
BACK_VALUE	[0.0, 0.0]

Notes. This list shows only parameters most relevant to the reproduction of the source catalogues, and so is not exhaustive.

fragmentation of sources), as the main targets of interest for KiDS lensing science are small, faint sources at intermediate-to-high redshift. This means, however, that there is likely to be fragmentation of the largest sources within the footprint (see e.g. the demonstration of the hot-mode shredding of galaxies in Figs. 5 and 6 of Andrews et al. 2017). The parameters used to perform the KiDS source extraction are provided in Table 4.

3.5. AstroWISE reduction

The ASTROWISE⁷ environment is a distributed database and data processing system, designed for calibrating and processing wide-field imaging data (McFarland et al. 2013). KiDS has utilised ASTROWISE for the production of reduced images, and forced optical photometry using GAAP (Kuijken 2008), in the *ugri*-bands since the beginning of the survey (de Jong et al. 2015).

For DR5 there have been a few minor modifications to the ASTROWISE imaging reduction pipeline compared to that described in Kuijken et al. (2019), which are important to document here.

3.5.1. Changes: Co-add production

There have been two changes to the co-add production for KiDS DR5: updates to the cross-talk coefficients (exclusively adding in new values), and a change to the polynomial order for defining the astrometric solution of the *u*-band co-adds.

Table 5 presents the cross-talk coefficients for the ASTROWISE co-add production. These coefficients have been updated to include values spanning the final observation window for KiDS and KiDZ, including the two lengthy shutdown periods of the VST as a result of the COVID-19 pandemic. As a result, the final set of cross-talk coefficients span a considerably longer period than the previous sets. We have not explored whether this difference introduces any systematic effect in the efficiency of the cleaning of cross-talk in the post-shutdown images obtained by the VST. However, as our source detection is performed on the THELI images (which were all observed prior to this period), we believe that this is unlikely to have a noticeable impact on our analyses.

⁷ <http://www.astro-wise.org>

Table 5. Applied cross-talk coefficients.

Period	CCD #95 to CCD #96 ^(a)		CCD #96 to CCD #95 ^(a)	
	<i>a</i>	<i>b</i> ($\times 10^{-3}$)	<i>a</i>	<i>b</i> ($\times 10^{-3}$)
2011.08.01 — 2011.09.17	-210.12	-2.5037	59.44	0.2739
2011.09.17 — 2011.12.23	-413.05	-6.8788	234.81	2.7275
2011.12.23 — 2012.01.05	-268.00	-5.1528	154.26	1.2251
2012.01.05 — 2012.07.14	-499.87	-7.8359	248.92	3.1095
2012.07.14 — 2012.11.24	-450.86	-6.9322	220.72	2.5348
2012.11.24 — 2013.01.09	-493.10	-7.2305	230.28	2.7224
2013.01.09 — 2013.01.31	-554.17	-7.5200	211.86	2.6099
2013.01.31 — 2013.05.10	-483.69	-7.0746	224.66	2.6283
2013.05.10 — 2013.06.24	-479.09	-6.9786	221.06	2.6380
2013.06.24 — 2013.07.14	-570.00	-7.7112	228.88	2.8397
2013.07.14 — 2014.01.01	-535.62	-7.4980	218.87	2.7008
2014.01.01 — 2014.03.08	-502.23	-7.1187	211.62	2.4288
2014.03.08 — 2014.04.12	-565.77	-7.5180	215.12	2.5782
2014.04.12 — 2014.08.12	-485.06	-6.8865	201.67	2.2365
2014.08.12 — 2014.01.09	-557.89	-7.5079	204.17	2.3036
2014.01.09 — 2015.05.01	-542.55	-7.5810	219.87	2.5347
2015.05.01 — 2015.07.25	-439.28	-6.9545	221.48	2.3951
2015.07.25 — 2015.08.25	-505.57	-7.5352	229.71	2.6048
2015.08.25 — 2015.11.10	-475.21	-7.3986	218.05	2.4449
2015.11.10 — 2016.06.17	-457.77	-6.8310	201.57	2.2117
2016.06.17 — 2016.06.25	-351.80	-4.9728	165.33	1.1678
2016.06.25 — 2016.09.08	-476.33	-6.9203	200.42	2.2018
2016.09.08 — 2017.08.01	-465.32	-6.5938	184.68	1.9801
2017.08.01 — 2018.11.01	-492.33	-6.4804	169.90	1.8016
2018.11.01 — 2021.10.01	-506.90	-6.4964	172.11	1.9174

Notes. ^(a)Correction factors *a* and *b* are applied to each pixel in the target CCD based on the pixel values in the source CCD:

$$I'_i = \begin{cases} I_i + a, & \text{if } I_j = I_{\text{sat.}}; \\ I_i + bI_j, & \text{if } I_j < I_{\text{sat.}} \end{cases}$$

where I_i and I_j are the pixel values in CCDs i and j , I'_i is the corrected pixel value in CCD i due to cross-talk from CCD j , and $I_{\text{sat.}}$ is the saturation pixel value.

The second change to the co-add production is related to the astrometry in the u -band. For all co-adds in DR4, the polynomial order for the distortion was set to three. In DR5, we found that the paucity of stars (per chip) detected in the u -band can occasionally lead to poorly constrained polynomial fits, leading to unphysical distortion solutions for some chips. Figure 12 shows the distribution of maximal astrometric residuals between the corners of all detectors in each pointing, assuming that all chips can be simply translated onto a common RA and Dec centroid. The figure shows that there is a clear systematic difference between the u -band and the other bands, whereby a significant number of pointings have chip corners that differ by more than $1''$; up to $24''$ in the worst case. Investigation of this effect led to the determination that failures in the u -band distortion parameters could be resolved by restricting the distortion polynomials to linear order. We opted to fit all u -band exposures with more than $1''$ chip-corner-residuals with a linear order polynomial fit instead. The $1''$ threshold was chosen because this is the median size of the PSF in the u -band, and (as the threshold is determined using the maximal distortion per pointing) this ensures any remaining systematic effects are below the seeing level for typical u -band images. The u -band co-adds that have been reprocessed in this way are listed in Appendix D.

3.5.2. Changes: Zero-point calibration

One particular complexity that arises in DR5 is the treatment of the zero-point calibration in the presence of the two i -band passes. In previous KiDS data reductions, the calibration process has utilised a stellar locus regression (SLR) between the four

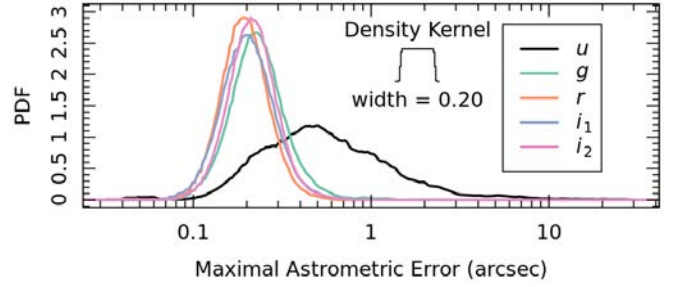


Fig. 12. Maximal residual between chip corners within a single pointing, assuming all chips can be shifted to a common RA and Dec centroid. Cases where the astrometric distortion parameters are poorly constrained (and so result in unphysical chip distortions) manifest as large residuals. The distribution in the u -band when using third-order distortion polynomials is clearly systematically larger than the other bands, with some chips having significant biases (greater than $10''$), due to the lower number of stars creating instability in the polynomial fits. To minimise the effect of this bias, we reprocessed all u -band tiles with maximal residuals greater than $1''$ using a linear polynomial distortion order.

optical filters. We implemented the same fundamental procedure, outlined in Sect. 3.1.3 of Kuijken et al. (2019) with further details in de Jong et al. (2017). Briefly, this process of SLR calibration within ASTROWISE proceeds as follows. First, so-called ‘principal colours’ are constructed from the measured GAAP fluxes for bright, unsaturated stars in each tile (Ivezić et al. 2004). Second, these colours are shifted so that straight sections of the colour-colour diagrams align with a set of fiducial templates. Finally, an overall zero-point correction is applied to all bands, in order to match the de-reddened ($r - G, g - i$) diagram to a template constructed from the SDSS survey. For this purpose, G band photometry is taken from *Gaia* DR2.

Given the additional i -band observation, we are required to run the SLR calibration process multiple times (i.e. running with the individual i -band passes separately). As a result, we have four SLR calibration estimates per tile: two for the GAAP minimum apertures of $0''.7$ and $1''.0$ (see Sect. 3.6), each computed with either the i_1 or i_2 imaging as the reference i -band. Performing the SLR in this way allows us to estimate the systematic uncertainty introduced in the calibration of the ugr -bands by variations in the quality of the i -band data, and to flag potentially bad calibrations.

The difference in the zero-point calibration offsets as estimated with the i_1 and i_2 fluxes, and with the two GAAP minimum aperture sizes, are presented in Fig. 13. Looking firstly at the accuracy of the g - and r -band calibrations, we can see that the zero-points are consistently reproduced to ± 0.02 magnitudes for all but a few tiles in the g -band. Similarly, looking at the calibration measured between the different GAAP settings, we see that, for all but a handful of outliers, the g - and r -band zero points are very consistently estimated. However, we of course cannot overlook the outliers in the gr -band comparisons, nor the significantly larger dispersion between zero-points estimated in the u -band.

In the figure we can see that the u -band zero-point calibration is somewhat unreliable, as has been previously documented (see e.g. Kuijken et al. 2019). In DR4, this behaviour prompted a recalibration effort that utilised *Gaia* and SDSS photometry to improve the u -band zero-point accuracies. For DR5, we adopted a slightly different approach to zero-point recalibration than was utilised in DR4 (and presented in Kuijken et al. 2019), who based zero-point corrections solely on the *Gaia* DR2 ‘white-light’ G -band. Our zero-point corrections are instead now based on a

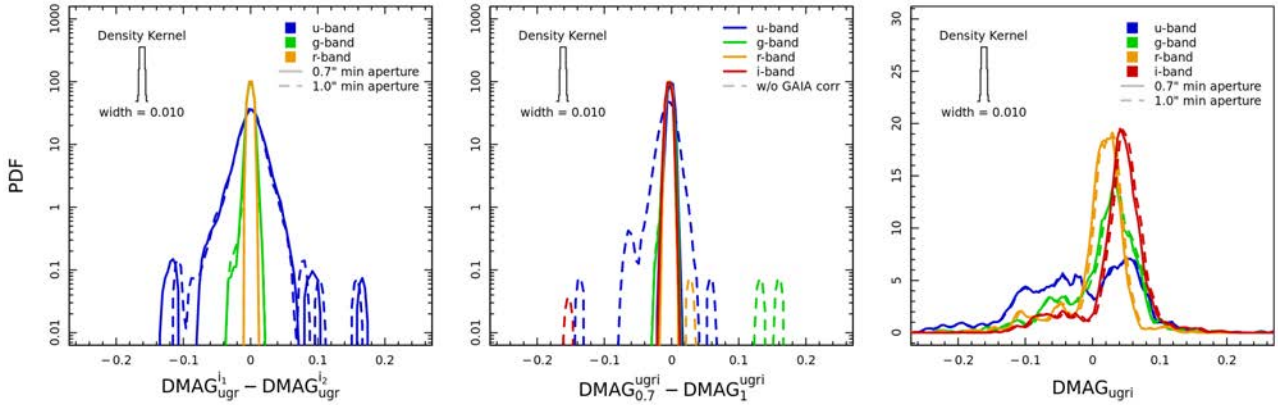


Fig. 13. Zero-point corrections estimated in KiDS+KiDZ using SLR. Left: difference between SLR offsets estimated using i_1 photometry and i_2 photometry, prior to recalibration of the u -band zero-points. Centre: differences between SLR offsets estimated using $0.7''$ minimum-radius aperture fluxes and $1.0''$ minimum-radius aperture fluxes (see Sect. 3.6). Right: distribution of final SLR offsets used in DR5 (i.e. corrections to the nightly zero points derived from photometric standards), including *Gaia* recalibration.

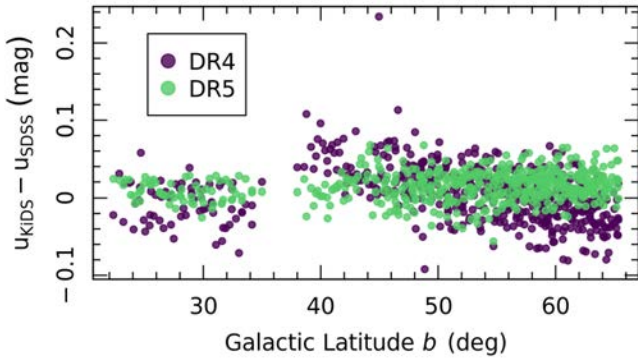


Fig. 14. Comparison between the u -band zero-point calibration in KiDS DR4 (green) and KiDS DR5 (purple). The updated DR5 calibration procedure produces zero-points that show considerably less systematic and random variation, when compared to fluxes from SDSS.

combination of the *Gaia* white-light G -band, blue G_{BP} -band, and red G_{RP} -band magnitudes, available in *Gaia* Early Data Release 3 (eDR3; [Gaia Collaboration 2021](#)). Use of this information in our zero-point calibration requires knowledge of the colour transformations between the *Gaia* and SDSS photometric systems, for which we use relationships derived by [Riello et al. \(2021\)](#). However, the transformations in [Riello et al. \(2021\)](#) were only computed for the gri -bands; as such, we introduce a predictor of the SDSS u -band from available *Gaia* magnitudes, \mathcal{U} , in a manner similar to [Kuijken et al. \(2019\)](#):

$$\mathcal{U} = f(G, G_{BP}, G_{RP}) + g(E_{B-V}, b), \quad (1)$$

where the first term is the *Gaia* internal colour transformation,

$$f(G, G_{BP}, G_{RP}) = G + 2.8 (G_{BP} - G_{RP}) - 0.6, \quad (2)$$

and the second term is the u -band extinction, as a function of Galactic latitude b ,

$$g(E_{B-V}, b) = -1.9 E_{B-V} - 0.10 |\sin(b)| + 0.09. \quad (3)$$

The predictor \mathcal{U} is calculated for sources that are bright ($G < 19$) and occupy the heart of the stellar locus $0.7 < G_{BP} - G_{RP} < 1.1$, in an effort to suppress non-linear terms in the preceding two equations. The dependence on Galactic latitude b was added

as discussed in [Kuijken et al. \(2019\)](#), with the implicit assumption that the Galactic latitude correction derived in KiDS-North (i.e. where there is overlap with SDSS) can be extrapolated to KiDS-S. The coefficients contained in Eqs. (2) and (3) are derived with two separate fits between u -band PSF magnitudes from the SDSS DR16 ([Ahumada et al. 2020](#)) and *Gaia* eDR3, selected within 710 KiDS-North tiles that overlap with the SDSS. The u -band offset for each tile in both KiDS-North and KiDS-South were then computed as the median of the differences $\Delta u = \mathcal{U} - u_{\text{KiDS}}$, measured for *Gaia* stars. A comparison between the quality of the zero-point calibration in the u -band in DR5 and DR4 is presented in Fig. 14. The figure demonstrates that there is considerably less variability (systematic and random) in the DR5 u -band zero-points compared to DR4. The scatter in the u -band zero-point reduces from ~ 0.035 in DR4 to ~ 0.018 in DR5: we notice that a similar scatter (0.016) was obtained by [Liang & von der Linden \(2023\)](#), where the calibration of u -band data in KiDS and other photometric surveys was improved using observed colours of blue Galactic halo stars.

The distribution of zero-point calibration offset difference between the two i -band passes (and the two minimum aperture sizes) is presented in Fig. 13. Somewhat by construction, the various calibrations now all agree with one-another, having been pegged to the external *Gaia* data. This is most apparent in the u -band, where we have applied this recalibration to every pointing. However, equally relevant is the choice to apply this recalibration process to the handful of tiles that show disagreement in the gr -bands. A similar recalibration approach can also be applied to the gri -bands, where we verified that robust results were obtained using the colour transformations available in [Riello et al. \(2021\)](#). However, given the accuracy that we observe already in the gr -bands, we decided to follow this option only for the few tiles where the SLR+*Gaia* approach described above failed to produce reliable offsets.

After the definition of final zero point corrections, individual corrections are applied to the catalogues. As with DR4, the corrections computed with the $0.7''$ minimum aperture are applied to fluxes computed with the same aperture; the same is done for the $1.0''$ minimum aperture corrections and sources. In DR5, though, we must combine the corrections computed using the i_1 and i_2 bands in the other (non- i) bands. In this case, we opted to take the straight arithmetic mean of the corrections estimated with the two i -band images for the ugr -bands.

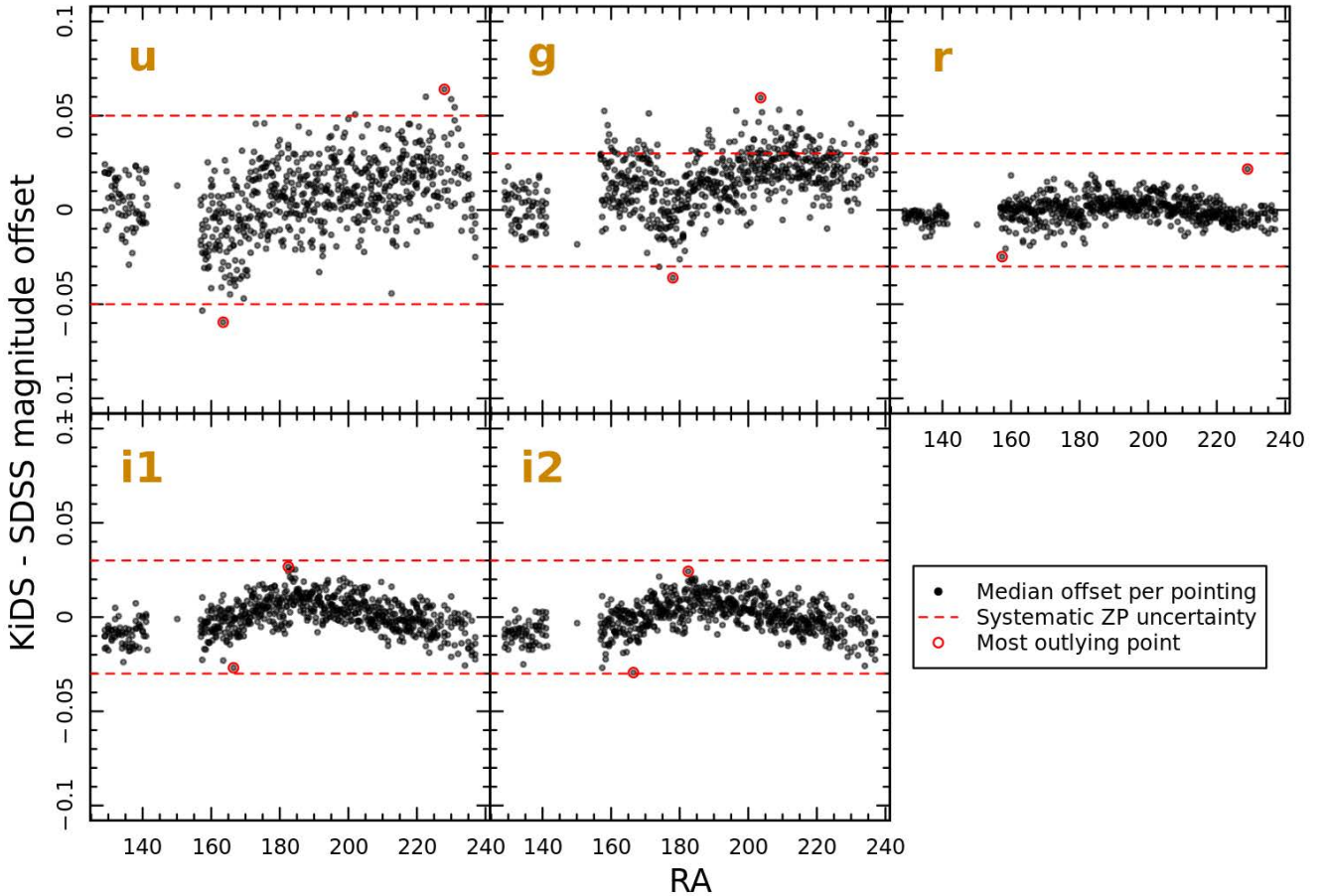


Fig. 15. Distribution of the median offsets between stars in KiDS DR5 tiles (in the magnitude range $16.5 < r < 19$) and their counterparts from SDSS imaging, for tiles with 10% or more unmasked data. The offsets here were calculated after SLR corrections and *Gaia* recalibration, and therefore represent the final quality of photometry in the survey. Horizontal dashed lines demonstrate the systematic zero-point uncertainty that is included per-band in our scientific analyses (such as in the computation of photo- z), which are designed to encapsulate any residual systematic variation in the photometry (such as can be seen as a function of RA, and whose origins are unclear).

Figure 15 shows the residual median offsets between magnitudes for stars in KiDS and SDSS, after the full zero-point calibration and recalibration process. Systematic residuals are below 0.05 magnitudes in the u -band for the majority of the tiles (<0.07 mag for the two most outlying points). In the gri -bands, residuals are smaller: consistently less than 0.03 magnitudes in the r - and i -bands, and typically less than 0.03 in the g -band. For our subsequent analyses, we implemented systematic error floors of $\sigma_u = 0.05$, and $\sigma_X = 0.03 \forall X \in \{g, r, i_1, i_2\}$, which are designed to encapsulate any residual systematic variation in the photometry (such as can be seen as a function of RA).

3.5.3. Changes: r -band PULECENELLA masks

The ASTROWISE pipeline includes the automatic masking of image artefacts around bright, saturated stars. This masking is performed with the PULECENELLA software, which separates the artefacts into different components (de Jong et al. 2015): saturated pixels in the cores of the stars, spikes caused by diffraction of the mirror supports, spikes caused by the readout of saturated pixels, and up to three families of wide annular ‘ghost’ reflection halos with spatially dependent offsets around bright stars.

The properties of these components (the size of saturation cores, the size and offset of reflection halos, the orientation of diffraction spikes) all depend on the brightness and position of

the stars in the focal plane, in a way that is stable in time for each photometric band. It is therefore possible to configure the modelling of each of these components by PULECENELLA at the beginning of the survey (de Jong et al. 2015). In DR4, however, it was noted that some of these parameters were perhaps too conservative (particularly in the r -band), producing masks that were in better agreement with the THELI-Lens ‘conservative’ masks rather than the fiducial masks. As such, in DR5 we updated the settings of the PULECENELLA software, related particularly to the size of reflection halos and orientation of diffraction spikes, to produce masks that are in better agreement with the fiducial masks generated by THELI-Lens. These changes are shown in Table 6, and their effect on a sample KiDS tile is displayed in Fig. 16.

3.5.4. Changes: $ugri_1i_2$ -band manual masking

As a final step of the optical reduction process, compressed 1000×1000 pixel images of all 1347×5 ASTROWISE co-adds were visually inspected for artefacts, and manually masked by means of a polygon region file. The main contaminating features discovered during this manual masking process were scattered light from bright sources outside the field of view (primarily in the pre-mid-2015 data, when the telescope was still poorly baffled). These artefacts can appear as fairly sharp-edged bright

Table 6. Updated parameters for the r -band PULECENELLA masking in KiDS DR5.

Parameter	DR4 value(s)	DR5 value(s)
halo1_area_min	850	1300
halo2_area_min	950	1250
halo3_area_min	950	950
spkd_profile [1]	13.2025	13.2025
spkd_profile [2]	34.9697	34.9697
spkd_profile [3]	4.5	4.0
spkd_profile [1]	102.386	102.386
spkd_profile [2]	133.923	133.923
spkd_profile [3]	4.5	4.0
spkd_profile [1]	170.512	170.512
spkd_profile [2]	192.542	192.542
spkd_profile [3]	3.5	3.0
spkd_profile [1]	195.332	195.332
spkd_profile [2]	216.286	216.286
spkd_profile [3]	4.0	3.5
spkd_profile [1]	291.721	291.721
spkd_profile [2]	309.5	309.5
spkd_profile [3]	3.0	2.5
spkd_profile [1]	310.561	310.561
spkd_profile [2]	331.917	331.917
spkd_profile [3]	3.5	3.0
spkd_profile [1]	351.588	351.588
spkd_profile [2]	369.336	369.336
spkd_profile [3]	3.5	3.0

Notes. These parameters define the minimum area of each halo (halo_area_min) and describe the positioning of the diffraction spikes (spkd_profile): starting angle ([1]), end angle ([2]), and spike length ([3]). Changed values are highlighted in bold face.

arcs across the focal plane (due to stars or planets), or as diffuse patches (due to the Moon). Other artefacts discovered during this masking included satellite flares, airplanes, and higher-order reflections from very bright stars.

Figure 17 shows an example of a heavily masked field in KiDS DR5. The field contains bright reflections from the out-of-field star Fomalhaut (α PsA, $V = 1.16$), which resides ~ 1.5 deg NNW of the centre of the focal plane. The field has been manually masked to remove the majority of bright contaminants. Faint residual fluctuations are only visible after heavily smoothing (with a $5''$ Gaussian filter) and scaling the image to emphasise the background variations.

3.5.5. Single-band catalogues

As with previous KiDS data releases, we provide source catalogues extracted from individual ASTROWISE co-adds in each band. These individual extractions are of particular interest for transient and variability studies, where one expects source lists to vary intrinsically between bands and/or observations. Of particular additional interest in DR5 is the distribution of single-band extractions in the two i -band passes, as sources with significant flux differences in the two i -band passes may indicate the presence of stationary transient features, such as extra-galactic supernova.

These source extractions were performed using the same source detection parameters as used in the main r -band extraction on the THELI-Lens images, except they were applied to the ASTROWISE reduced co-adds. As a result, there are differences

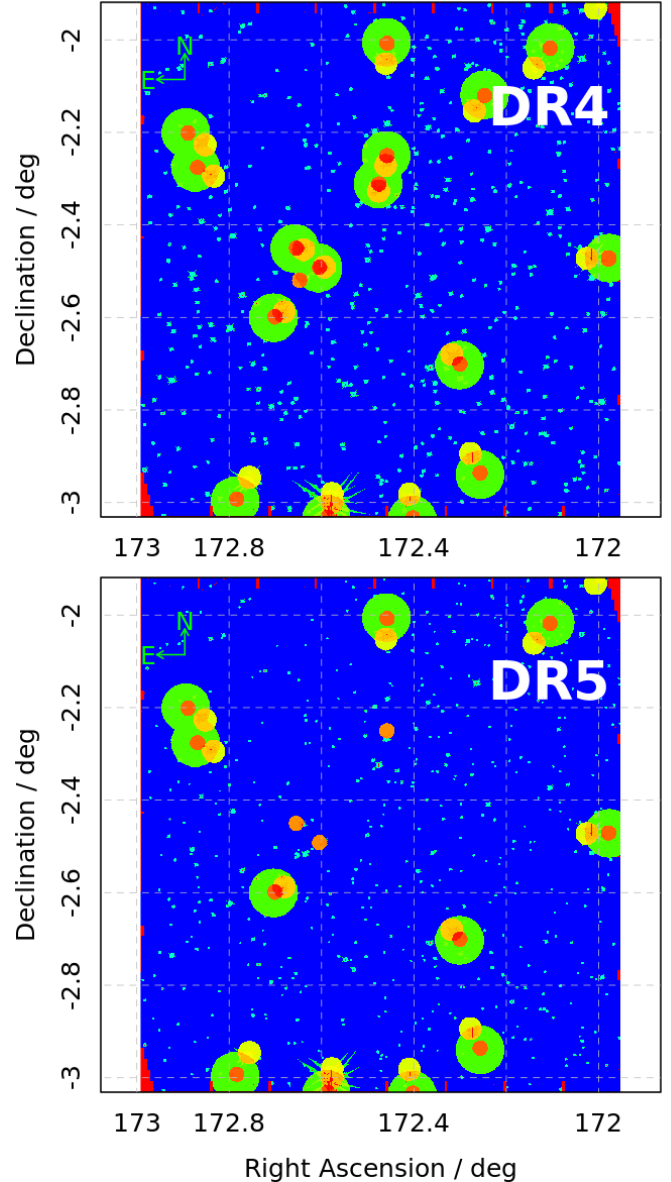


Fig. 16. Comparison between the ASTROWISE PULECENELLA masks for a selected KiDS pointing. The images show the masks on a consistent linear colour scale, so pixels masked with the same bit(s) are shown with the same colour in each image. The DR5 implementation of these masks was refined to use slightly different parameters (see Table 6), which more closely reproduce the masking behaviour of THELI-Lens in the ASTROWISE r -band. The primary effect is a reduction (in DR5) of the size of stellar masks (cyan) and the frequency of masking of large reflection halos (green).

in the numbers of detected sources (and their observed properties), even when comparing the single-band r -band catalogues to the primary one. Also, it is important to note that cross-matching between single-band catalogues is unlikely to produce multi-band photometry that is accurate (at the same level as the dedicated multi-band photometry, Sect. 3.6), because there is no guaranteed consistency of apertures and/or deblend solutions for the independent extractions. As such, multi-band photometry obtained from position-matching single-band catalogues should be treated with caution.

Finally, as with previous releases, single-band catalogues are provided to the public largely ‘as-is’; they have not gone through

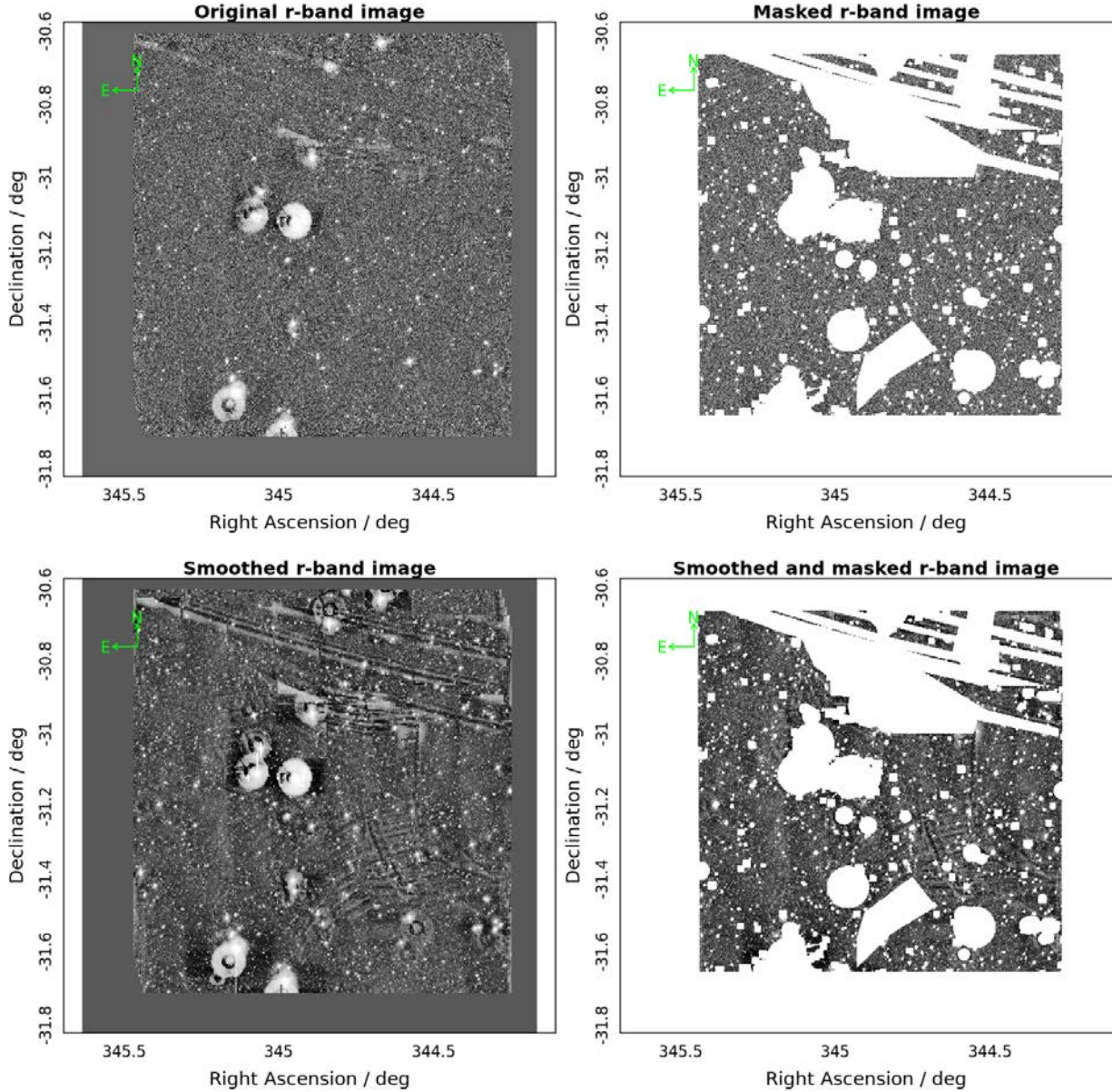


Fig. 17. Examples of the new ASTROWISE manual masking implemented for KiDS DR5. The figure shows a heavily contaminated field, caused by scattered light from Formalhaut (α PsA), a visible-magnitude star ($V = 1.16$) that is ~ 1.5 deg from the centre of the focal plane. The left column shows the tile before masking, while the right column shows the tile after application of the manual and PULECENELLA masks. The upper row shows the tile at native resolution, while the bottom row shows the tile after smoothing with a $5''$ Gaussian filter.

the same rigorous quality control and testing as the lensing catalogues. More details of the contents of these catalogues are provided in Sect. 8.

3.6. Optical multi-band GAAP photometry

One of the most important products released by the KiDS consortium in each release is the multi-band photometry measured using the GAAP code (Kuijken 2011). For a range of images (spanning various photometric bandpasses and PSFs), GAAP computes a (typically non-total) intrinsically consistent flux (i.e. probing the same pre-convolution extent in each image). This is performed as follows. Given an intrinsic source flux distribution, $f(x, y)$, an estimate of the flux can be made with a weighted aperture, using a Gaussian kernel. In GAAP, the Gaussian kernel is defined using an estimate of the source's major and minor axis

lengths (A and B , respectively), as measured in the r -band, and the source's orientation angle θ . The flux estimated by GAAP is therefore

$$F_{\text{GAAP}} = \iint dx dy f(x, y) \exp \left\{ \frac{-[(x'/A)^2 + (y'/B)^2]}{2} \right\}, \quad (4)$$

where x' and y' are the x and y coordinates shifted to the centre of the source and rotated into the galaxy frame using the position angle θ . In the presence of an arbitrary Gaussian PSF with standard deviation σ_{PSF} , this intrinsic flux can be shown to be related directly to a Gaussian-weighted flux measured on the observed image, using a modified Gaussian kernel with

$$A' \mapsto \sqrt{A^2 - \sigma_{\text{PSF}}^2}; \quad (5)$$

$$B' \mapsto \sqrt{B^2 - \sigma_{\text{PSF}}^2}. \quad (6)$$

This therefore provides a simple way to calculate fluxes across multiple images that probe the same intrinsic scales of a galaxy, provided that the image PSF is Gaussian (and assuming that sources are unblended).

To ensure that image PSFs are Gaussian, we performed a step of Gaussianisation within ASTROWISE. This process involves accurately measuring the native PSF of each image, and using a shapelets expansion (Refregier 2003; Kuijken 2006) to compute the (spatially varying) kernel required to convert the native PSF to a Gaussian (with minimal information loss). The native images are then convolved with the Gaussianisation kernel, producing images with a Gaussian PSF (at the cost of an increased correlation of the noise profile). The image PSF size is essentially unchanged in this process (as quantified via the FWHM), and the correlated noise caused by the convolution is propagated to the estimate of the flux uncertainty.

As is clear from Eq. (5), there is a physical limitation to the GAAP formalism when $\sigma_{\text{PSF}} \geq [A, B]$. In this limit, the effective aperture sizes (A', B') become imaginary in one or both axes, and flux measurement is not possible. It is therefore sensible to define a minimum intrinsic aperture size for every source, that is larger than the expected PSF size in all bands. In KiDS DR5, as in DR4, we measured all sources with two intrinsic aperture sizes, which are a combination of the source intrinsic size (as the RMS of the flux distribution along the major and minor axes, in arcseconds, estimated from the THELI-Lens r -band imaging by SOURCE EXTRACTOR), a minimum aperture size r_{min} , and a maximum aperture size of $2''$ (to limit blending effects). As such, the intrinsic aperture sizes for all sources in DR5 are defined as

$$A_{\text{GAaP}} = \min\left(\sqrt{A^2 + r_{\text{min}}^2}, 2''\right); \quad (7)$$

$$B_{\text{GAaP}} = \min\left(\sqrt{B^2 + r_{\text{min}}^2}, 2''\right). \quad (8)$$

The two distinct sets of apertures come from the use of two distinct minimum aperture radii: $r_{\text{min}} \in [0''.7, 1''.0]$. While valid GAAP fluxes across all bands can be obtained with the lower r_{min} value for most sources, data with poorer seeing require the larger aperture (this point is further discussed in Sect. 6.2).

4. NIR observations and reduction

One of the unique features of KiDS is the complementary five-band NIR survey VIKING. The utility of this overlap lies particularly in estimation of photometric redshifts and redshift distribution calibration, resulting in higher-quality measurements across a larger redshift baseline, as well as in classification studies generally (see e.g. Nakoneczny et al. 2021). With the extension of KiDS to include the KiDZ observations in DR5, we also required new NIR observations over these fields. In this section we detail the existing and recently acquired NIR observations in the KiDS+KiDZ fields, their reduction, and their quality.

4.1. VISTA paw-print observations

The Visible and Infrared Survey Telescope for Astronomy (VISTA) is a 4m ESO Telescope, is also located at ESO's Cerro Paranal observatory (albeit on a separate peak, roughly 1500 m from the VST), and is serviced only by the Visible

Table 7. Run numbers of all VISTA observations taken in the KiDS and KiDZ fields.

Survey	Run number	Start date	N_{paw}
VIKING	179.A-2004	2009-11-13	40 741
	179.A-2006	2009-11-05	168
	179.A-2005	2010-02-18	24
KiDZ	298.A-5015	2016-12-05	97
	0100.A-0613	2017-09-17	204
	0102.A-0047	2018-10-02	252

Notes. Run numbers include observations taken as part of guaranteed time observations programmes and under director's discretionary time.

and Infrared Camera (VIRCAM). VIRCAM consists of 16 individual HgCdTe detectors, each with a 0.2×0.2 square degree angular size, but which jointly span a 1×1.2 deg² field of view. A single exposure of the sky therefore contains (considerable) gaps between the detectors, in what is referred to as the 'paw-print' pattern, and standard VIRCAM observations combine six dithers designed to fill in these gaps (see Dalton et al. 2006). Additionally, observations of a single paw-print consist of a number of small jitters, taken in quick succession, which allow the paw-prints to have reliably estimated backgrounds and to sample over detector defects. These exposures are stacked into a single 'stacked paw-print', and six of these are combined to form a single contiguous ~ 1.5 deg² 'tile'. The task of reducing the raw images, particularly producing stacked paw-prints, was carried out by the Cambridge Astronomy Survey Unit (CASU; González-Fernández et al. 2018; Lewis et al. 2010).

VISTA observations used in DR5 span both the KiDS survey area (with observations from VIKING, Sect. 4.1.1), and the KiDZ fields (with a combination of new, dedicated VIKING-like observations, and reconstructed VIKING-like observations made from pre-existing deep VISTA observations). The relevant observing programmes are listed in Table 7.

4.1.1. VIKING observations within KiDS

As mentioned previously, the VIKING imaging survey conducted on VISTA was designed in combination with KiDS, to produce well-matched optical and NIR data. As such, the surveys share an almost identical footprint on-sky: Fig. 18 shows the coverage of VIKING observations within the northern and southern patches of KiDS. We note the excellent overlap between the two surveys, whereby only relatively small areas at the extreme edges of the northern and southern patches are without complete ($ZYJHK_s$) information.

Initial VIKING observations were taken between 13 November 2009 and 24 August 2016, defining the final footprint. Subsequently, some observations were repeated (primarily of data with instrumental problems), with the final observation dating from 16 February 2018. The data rate of observations is shown in Fig. 19. All VIKING observations can be found in the ESO archive under programme ID 179.A-2004, and a summary of the observational requirements is provided in Table 8. A detailed description of the survey design and observing strategy is given in Edge et al. (2013) and Venemans et al. (2015).

4.1.2. KiDZ VISTA observing programme

In order to produce NIR observations in the KiDZ fields that are consistent with those existing in the KiDS fields, we undertook a

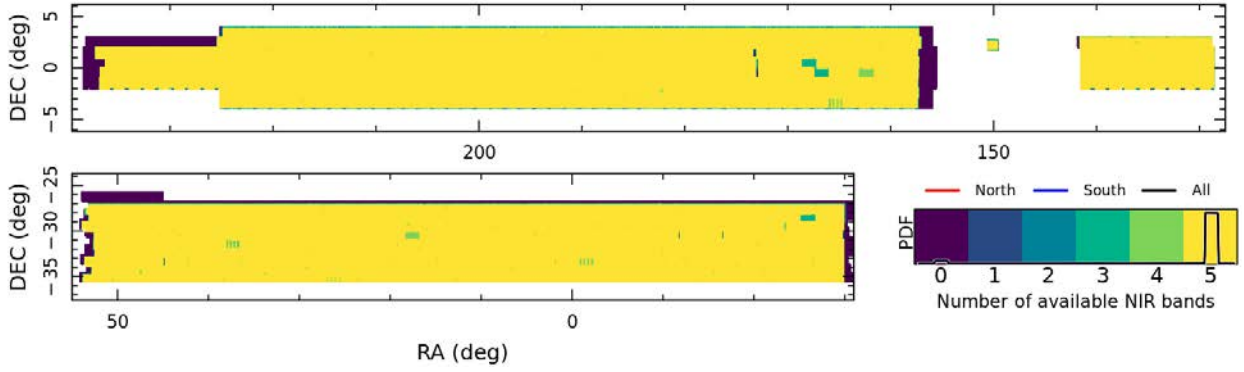


Fig. 18. Coverage of VISTA VIKING data in the KiDS fields, demonstrating that nearly the entire KiDS DR5 footprint is covered by complete ZYJHK_s-band VIKING observations.

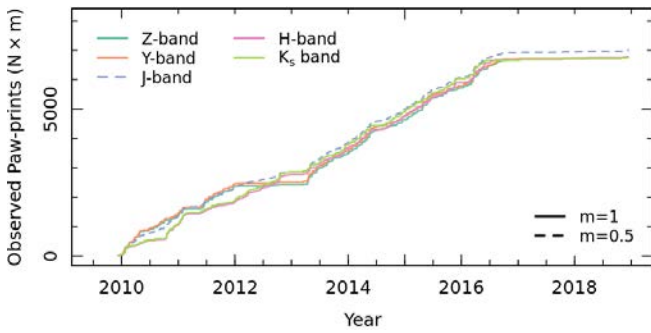


Fig. 19. Progression of the NIR observations by the VIKING survey. Note that the *J*-band line has been multiplied by a factor of 0.5, to account for the fact that it is observed with twice the frequency of the other bands. The observing strategy, combining observations of ZYJ-bands and JHK_s-bands, can be seen in the correlated increase in observed paws in the various filters.

Table 8. Requirements and settings for VIKING observations with VIRCAM on VISTA.

Filter	t_{exp} (s)	N_{exp}	N_{jitter}	$\langle N_{\text{dither}} \rangle$ (per pix)	$\langle t_{\text{tot}} \rangle$ (s)
Z	60	1	4	2	480
Y	25	2	4	2	400
J	25	2	2	4	400
H	10	5	3	2	300
K _s	10	6	4	2	480

Notes. The number of dithers per pixel (N_{dither}), and therefore also the total exposure time per pixel (t_{tot}), varies over the contiguous 1.2 deg²tile. As such, in this table we show the N_{dither} and t_{tot} values for ‘typical’ pixels, and emphasise this point by annotating these columns with ‘average’ brackets ($\langle \dots \rangle$).

campaign to obtain VIKING-like data over the KiDZ fields from 5 December 2016 to 2 October 2018 (see Table 7), with the same observing constraints and settings as the VIKING survey itself (i.e. Table 8). Figure 6 shows the distribution of these dedicated VISTA observations in each of the KiDZ fields, as green boxes. The areas of the KiDZ fields that are not covered by these dedicated observations, but which nonetheless contain both optical and NIR data (shown by the dark grey), consist of VIKING-like data that was reconstructed from pre-existing deep VISTA observations in these fields (Sect. 4.1.3).

Table 9. Requirements and settings for deep VIRCAM with UltraVISTA and VIDEO on VISTA.

Filter	t_{exp} (s)	N_{exp}	N_{jitter}	$\langle N_{\text{dither}} \rangle$ (per pix)	$\langle t_{\text{tot}} \rangle$ (s)
VIDEO					
Z	50	1	10	1	500
Y	30	2	8	1	480
J	30	2	8	1	480
H	10	6	7	1	420
K _s	10	6	7	2	840
UltraVISTA					
Y	60	2	30	1	3600
J	30	4	30	1	3600
H	10	6	30	1	1800
K _s	10	6	30	1	1800

Notes. Observation settings for each paw-print (t_{exp} , N_{exp} , N_{jitter}) were used to compute the average number of dithers per-pixel on-sky ($\langle N_{\text{dither}} \rangle$) that are needed to reach at least VIKING depth. We then randomly sampled $\langle N_{\text{dither}} \rangle$ dithers per pixel on-sky to construct our VIKING-like data in these deep fields. Subsequent further degradation of the photometry (in catalogue space) is needed to reach equivalent VIKING depths.

4.1.3. Constructed VIKING-like data in KiDZ deep fields

For a number of the KiDZ fields, there were pre-existing extremely deep VIRCAM observations obtained by the UltraVISTA (McCracken et al. 2012) and VIDEO (Jarvis et al. 2013) surveys. Rather than re-observing these fields, we constructed VIKING-quality data by selecting observations such that the total exposure time per pixel was at least as deep as in VIKING, and which contain paw-prints of similar seeing. The observation settings for UltraVISTA and VIDEO (and how many of such observations were selected to reach at least VIKING depth, $\langle N_{\text{dither}} \rangle$) are given in Table 9. We note that there are no Z-band observations within UltraVISTA; here new Z-band imaging was obtained with VISTA using VIKING observation parameters (included in Table 7). Given the very different observing strategies, it proved impossible to match the depth of the VIKING data exactly, and we therefore further degraded the photometry in these bands to the expected depth of VIKING (a process described in Hildebrandt et al. 2017 as ‘magnitude adaption’). This step is clearly particularly relevant in UltraVISTA, where typical depths are equivalent to the deepest parts of VIKING (i.e. where dither overlaps create $N_{\text{dither}} = 6$).

4.2. KiDS-VIKING pipeline reduction

The NIR data in KiDS+KiDZ were reduced with the same reduction pipeline used in previous KiDS data releases: KPIPE (Wright et al. 2018), based on the processing pipeline developed in Driver et al. (2016) for GAMA. We opted to start from the reduced CASU paw-prints, because of the complex observing strategy and dither pattern of VISTA observations: within one tile there is a wide range of total exposure times (between $1\times$ and $6\times$ the individual exposure times, t_{exp}) and observational properties (up to 96 different PSFs). This variability motivated us to utilise individual stacked chips, extracted from the CASU paw-prints, as the basis for our reduction and forced photometry, rather than co-added tiles.

The processing pipeline first corrects the images for atmospheric extinction (τ) given the observation airmass ($\sec\chi$), removes the exposure time (t , in seconds) from the image units, and converts the images from various Vega zero-points (Z_v) to a standard AB zero-point of 30 (using Vega to AB corrections, X_{AB} , from González-Fernández et al. 2018). These various corrections and transformations are performed using a single multiplicative recalibration factor \mathcal{F} per VISTA detector that is applied to all pixels in the detector image I :

$$I_{\text{new}} = I_{\text{orig}} \times \mathcal{F}. \quad (9)$$

The factor \mathcal{F} is constructed as

$$\log_{10}(\mathcal{F}) = -0.4[Z_v - 2.5 \log_{10}(1/t) - \tau(\sec\chi - 1) + X_{\text{AB}} - 30]. \quad (10)$$

The processing pipeline also performs a re-orientation of the individual paw-print detector stacks using SWARP (Bertin 2010) and a background subtraction (again using SWARP). The SWARP background subtraction is computed using a 256×256 pixel mesh, and a 3×3 filter for the bicubic spline. This additional background subtraction was considered optimal by Driver et al. (2016), who demonstrated that these settings allowed for maximal removal of backgrounds with minimal impact on the photometry of extended sources.

The distribution of the recalibration factor \mathcal{F} per band is shown in Fig. 20. Following Driver et al. (2016); Wright et al. (2018) we applied a blanket quality-control selection of $\mathcal{F} \leq 30.0$ to remove detectors with strong persistence or other artefacts.

4.3. NIR quality metrics

We quantify the quality of the NIR reduction using the same metrics as for the optical portion of the survey: PSF FWHM and depth (as determined in blank apertures scattered randomly over the detectors, Appendix A). Figure 21 presents the summary of PSF sizes and limiting magnitudes in each of the five NIR bands.

The PSF sizes in the NIR bands are all exceptionally consistent. This is due (at least in part) to the rapid variability of the NIR sky: PSFs vary on a second-to-second timescale, as do backgrounds. This necessitates that the ‘typical’ PSF of any one observation approaches the mean of the seeing conditions (i.e. following the central limit theorem). However, the distributions also all show considerable tails to high seeing; again, a demonstration of the difficulties in observing the NIR sky from the ground. This tail of poor seeing observations has some important consequences for our analysis and sample selection further down the line in our processing pipeline (Sects. 6.2 and 6.4).

Limiting magnitudes in the NIR filters were determined on individual detectors. However, as there are possibly many individual detectors per source, we need a reasonable method for

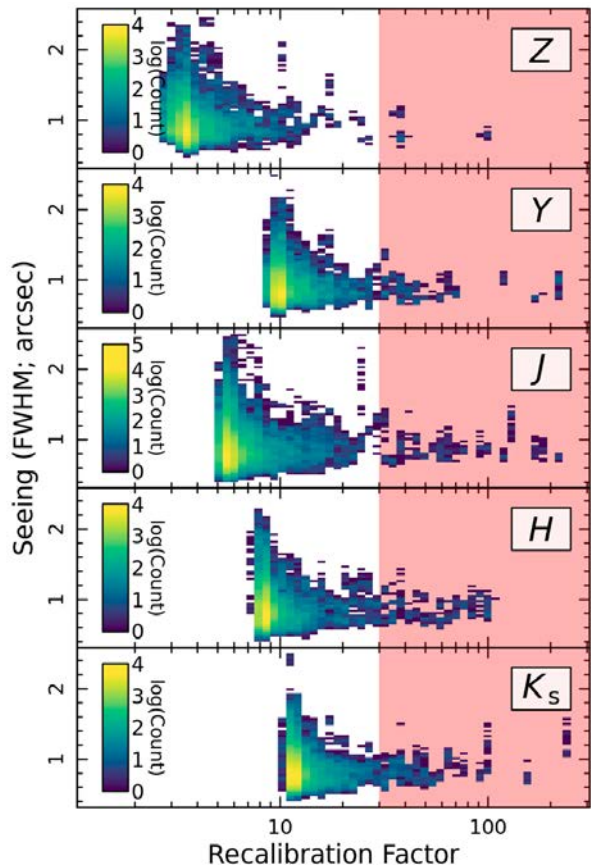


Fig. 20. Distribution of recalibration factors \mathcal{F} , derived by KPIPE using the parameters provided by CASU, for VIKING and VIKING-like observations in the KiDS+KiDZ fields. Following Driver et al. (2016) and Wright et al. (2019), we applied a rejection of detectors with recalibration factors $\mathcal{F} \geq 30$. There are additional indirect selections, however, that occur for fields with large PSF sizes (see Sect. 6.4).

combining the estimated magnitude limits into a final representative distribution of depths (i.e. that reflects the approximate depth that we achieve per source, in the correct proportion). We achieved this by measuring the background variance of the individual chips at a number of locations within the chip. Each of these background estimates is tagged with the central RA and Dec used for the estimate, and estimates (from a single band) that are within $\Delta\{\text{RA}, \text{Dec}\} \leq 0.02$ deg of each other are pooled into a single estimate, assuming that fluxes are combined using a simple average of their individual measurements (which is not strictly the case; see Sect. 6.2), and that the noise in overlapping chips is uncorrelated. This means that the variance of the mean stack of chips at each location on sky is simply

$$\sigma_{\text{final}}^2 = \frac{\sum_i^N \sigma_i^2}{N^2}, \quad (11)$$

where N is the number of detectors overlapping with each other at this point on-sky, and σ_i^2 is the variance of the i^{th} detector. From this calculation, we found that the median limiting magnitudes in the ZYJHK_s-bands are {23.49, 22.74, 22.55, 21.96, 21.77}. The distribution of these combined limiting magnitude estimates on-sky is given in Appendix E.

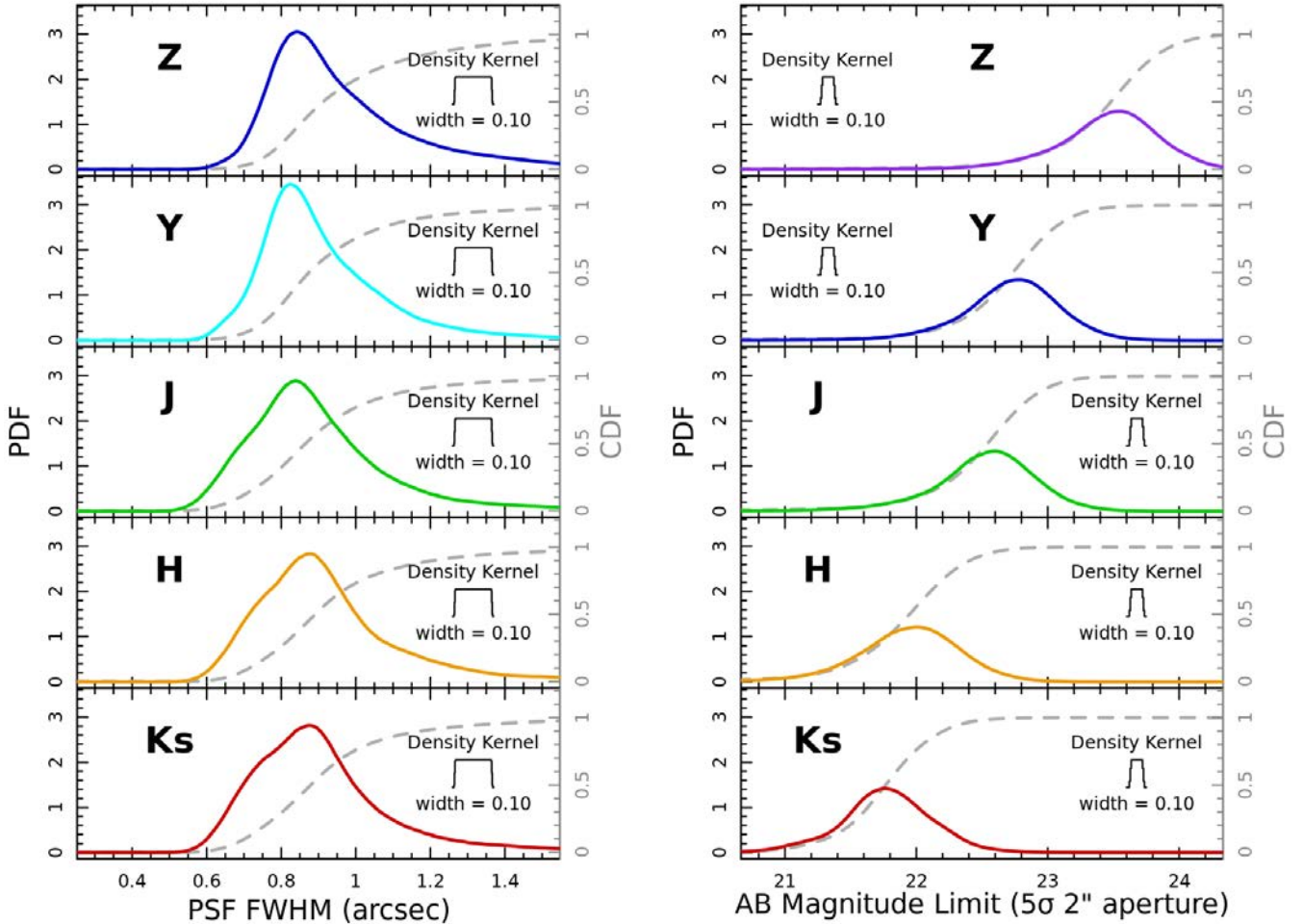


Fig. 21. Primary observational properties of pointings in KiDS+KiDZ for observations in the five NIR bands. Each row shows the distribution of PSF sizes (as measured on each VISTA chip; left) and limiting magnitude (as determined by the magnitude of a 5σ source in a $2''$ circular aperture; right) determined with a KDE using the annotated kernel. The corresponding cumulative distribution functions for each panel are shown as grey lines.

5. KiDZ spectroscopic compilation

As discussed in Sects. 3.1.2 and 4.1.2, we undertook a targeted campaign to observe eight spectroscopic calibration fields, which make up the ‘KiDZ’ fields. The associated spectroscopic redshift data used in DR5 is a compilation of public and proprietary measurements from a range of surveys, which we refer to here simply as ‘the full compilation’. The list of spectroscopic surveys that we draw from to create the full compilation is provided in Table 10.

The compilation includes two main components. The first are redshifts from wide-angle spectroscopic surveys that mainly cover lower redshifts and brighter magnitudes, and which overlap at least partly with the KiDS main survey area. Surveys that contribute to this portion of the dataset include SDSS (York et al. 2000), GAMA (Driver et al. 2022), 2dFLens (Blake et al. 2016), and WiggleZ (Parkinson et al. 2012). The second component comes from several deep (often pencil-beam) spectroscopic surveys that intersect the KiDZ fields, including VVDS (Ilbert et al. 2006), VIPERS (Scodreggio et al. 2018), DEEP2 (Newman et al. 2013), C3R2 (Euclid Collaboration 2021), and zCOSMOS (Lilly et al. 2009).

A number of these datasets have sources in common (due, in part, to some of the samples being compilations themselves:

e.g. G10_COSMOS, zCOSMOS, and GOODS). Furthermore, some datasets include duplicate and/or multiple redshift estimates, with some appearing more than twice in their respective datasets. To remove duplicated spectroscopic redshift estimates, we adopted a strategy designed to retain only the highest-quality redshift estimates per source. Before performing this duplicate removal, however, we first homogenised the redshift quality flags among the different surveys (which are defined using various inconsistent criteria). This procedure resulted in a single redshift quality flag, based on the GAMA NQ nomenclature, that indicates the confidence in a redshift estimate’s accuracy as

$$NQ = \begin{cases} 4 & \text{certain } (\geq 98\% \text{ confidence}) \\ 3 & \text{high } (95\text{--}98\% \text{ confidence}) \\ 2 & \text{moderate confidence} \\ 1 & \text{low confidence} \end{cases} \quad (12)$$

We note that in what follows we only use objects with $NQ \geq 3$.

This spectroscopic compilation is nominally the same as used previously in van den Busch et al. (2022), although the overlap with photometric data has increased considerably due to our dedicated KiDZ imaging. van den Busch et al. (2022) describe the homogenisation process used in the construction of the compilation, and the results of the merger in the spectroscopic fields

Table 10. Statistics for the KiDZ spectroscopic sample.

Survey	Full compilation				Matched to KiDZ sources					Key reference
	N_{spec}	$[z_{\text{min}}, z_{\text{max}}]$	$\langle z \rangle$	NMAD(z)	N_{spec}	$[z_{\text{min}}, z_{\text{max}}]$	$\langle z \rangle$	NMAD(z)	$\langle r_{\text{auto}} \rangle$	
VIPERS	76 519	[0.008, 2.148]	0.712	0.181	43 382	[0.045, 2.148]	0.713	0.177	22.6	Scodreggio et al. (2018)
G10_COSMOS	23 545	[0.007, 4.596]	0.582	0.322	20 214	[0.007, 4.596]	0.583	0.324	22.4	Davies et al. (2015)
DEVILS	18 213	[0.004, 4.786]	0.581	0.333	15 103	[0.005, 4.786]	0.569	0.333	21.4	Davies et al. (2018)
VVDS	15 059	[0.003, 4.540]	0.609	0.295	12 598	[0.010, 4.116]	0.609	0.295	22.6	Le Fèvre et al. (2013)
DEEP2	15 888	[0.009, 2.161]	0.957	0.227	11 889	[0.009, 2.161]	0.962	0.233	23.7	Newman et al. (2013)
ACES	4 276	[0.037, 2.838]	0.592	0.250	4 157	[0.037, 2.838]	0.591	0.250	22.4	Cooper et al. (2012)
C3R2	4 348	[0.030, 4.503]	1.014	0.517	3 132	[0.030, 4.175]	0.967	0.538	24.0	Euclid Collaboration (2021)
zCOSMOS	3 496	[0.006, 4.445]	1.396	1.155	2 585	[0.006, 4.445]	1.490	1.110	23.6	priv. comm. (M. Salvato)
DEIMOS	4 197	[0.009, 6.604]	1.272	0.421	2 457	[0.009, 4.654]	1.003	0.381	23.5	Hasinger et al. (2018)
OZDES	12 511	[0.002, 4.541]	0.574	0.352	2 288	[0.008, 4.541]	0.697	0.372	21.2	Lidman et al. (2020)
G15_Deep	1 840	[0.006, 1.193]	0.357	0.163	1 840	[0.006, 1.193]	0.357	0.163	21.2	Kafle et al. (2018)
GAMA	241 130	[0.003, 3.870]	0.234	0.115	1 755	[0.006, 0.729]	0.217	0.103	19.0	Driver et al. (2022)
GOODS	2 517	[0.002, 6.277]	1.382	0.765	1 703	[0.002, 4.029]	1.187	0.644	23.8	ESO compilation
hCOSMOS	1 081	[0.012, 1.265]	0.302	0.132	876	[0.012, 1.265]	0.303	0.132	20.1	Damjanov et al. (2018)
SDSS	109 459	[0.002, 1.002]	0.369	0.251	525	[0.004, 0.817]	0.355	0.285	19.0	Abolfathi et al. (2018)
VANDELS	1 701	[0.309, 5.784]	3.086	0.765	458	[0.309, 4.537]	2.346	0.875	24.5	Garilli et al. (2021)
FMOS_COSMOS	383	[0.969, 2.486]	1.568	0.083	325	[1.313, 2.486]	1.568	0.080	24.1	Silverman et al. (2015)
WIGGLEZ	55 076	[0.002, 6.159]	0.598	0.204	308	[0.006, 1.776]	0.586	0.190	21.6	Parkinson et al. (2012)
LEGA_C	352	[0.346, 1.085]	0.798	0.155	302	[0.519, 1.050]	0.806	0.153	22.9	van der Wel et al. (2016)
VUDS	331	[0.098, 6.091]	2.219	1.488	186	[0.098, 4.161]	2.003	1.294	24.7	Le Fèvre et al. (2015)
2dFLens	43 177	[0.004, 4.688]	0.361	0.249	2	[0.273, 0.336]	0.304	0.046	18.4	Blake et al. (2016)

Notes. Sources and statistics of the spectroscopic redshift estimates contained within the full spectroscopic compilation used in KiDZ (‘Full Compilation’) and after matching to sources extracted from KiDZ imaging (including masking, ‘Matched to KiDZ sources’). The numbers in the N_{spec} column refer to samples with duplicates removed (both internally and between overlapping surveys). Statistics for each sample shown are the minimum and maximum redshift in the sample ($z_{\text{min}}, z_{\text{max}}$), the mean redshift of the sample ($\langle z \rangle$), and the NMAD scatter of the redshifts in the sample (NMAD(z)). For each survey we provide a single key reference; additional relevant papers are provided in the main text.

available during DR4. This process is unchanged here, but is applied to the larger KiDZ sample that is now available. Here we summarise the selection and quality flag assignment per input survey.

- Arizona CDFS Environment Survey (ACES; Cooper et al. 2012): galaxies were selected with $Z_{\text{QUALITY}} \geq 3$ and $z_{\text{Err}}/z < 0.01$. We assigned $\text{NQ} = Z_{\text{QUALITY}}$;
- Deep Extragalactic Visible Legacy Survey (DEVILS; Davies et al. 2018): galaxies were selected with spectroscopic redshifts (i.e. $z_{\text{BestType}} = \text{spec}$), and with flags $\text{starFlag} = 0$, $\text{mask} = 0$, and $\text{artefactFlag} = 0$. We assigned $\text{NQ} = 4$ if $z_{\text{BestSource}} = \text{DEVILS}$ and $\text{NQ} = 3$ otherwise;
- Complete Calibration of the Colour-Redshift Relation (C3R2) survey: galaxies were selected from the combination of four C3R2 public datasets: DR1 (Masters et al. 2017), DR2 (Masters et al. 2019), DR3 (Euclid Collaboration 2021), and KMOS (Euclid Collaboration 2020). We required $\text{QFLAG} \geq 3$ for galaxies in this sample, and assigned $\text{NQ} = \text{QFLAG}$;
- Deep Imaging Multi-Object Spectrograph survey (DEIMOS; Hasinger et al. 2018): galaxies were selected with quality flag $Q = 2$. We assigned $\text{NQ} = 4$ for $Q_{\text{f}} \in [4, 14]$, and $\text{NQ} = 3$ otherwise;
- DEEP2 (Newman et al. 2013): as in the previous KiDS papers (Hildebrandt et al. 2017, 2020, 2021) we selected galaxies from two equatorial fields (0226 & 2330), with $Z_{\text{QUALITY}} \geq 3$ and $z_{\text{Err}}/z < 0.01$. We assigned $\text{NQ} = Z_{\text{QUALITY}}$;
- Fiber Multi-Object Spectrograph COSMOS survey (FMOS-COSMOS; Silverman et al. 2015): galaxies were selected with quality flag $q_z \geq 2$. We assigned $\text{NQ} = 4$ if $q_z = 4$ and $\text{NQ} = 3$ otherwise;

- GAMA (Driver et al. 2022): galaxies from the 4th Data Release were selected with redshift quality $\text{NQ} \geq 3$ and with $z > 0.002$ to avoid stellar contamination. We propagated NQ from GAMA;
- GAMA-G15Deep (Kafle et al. 2018; Driver et al. 2022): galaxies were selected with input redshift quality $Z_{\text{QUAL}} \geq 3$ and with redshifts $z > 0.001$ to avoid stellar contamination. We assigned $\text{NQ} = Z_{\text{QUAL}}$;
- G10-COSMOS (Davies et al. 2015): galaxies were selected with Z_{BEST} as the redshift value, and otherwise following the documentation for selecting galaxy redshifts: $Z_{\text{BEST}} > 0.0001$, $Z_{\text{USE}} < 3$, and $\text{STAR_GALAXY_CLASS} = 0$. All sources are assigned $\text{NQ} = 3.5$.
- Great Observatories Origins Deep Survey (GOODS CDFS): galaxies were selected from the public ESO compilation of spectroscopy in the CDFS field⁸ (Popesso et al. 2009; Balestra et al. 2010). Following the recommendations in the dataset description⁹ we selected ‘secure’ redshifts (assigning $\text{NQ} = 4$ to them) and ‘likely’ redshifts ($\text{NQ} = 3$);
- Hectospec COSMOS survey (hCOSMOS; Damjanov et al. 2018): all galaxies from the published dataset were selected, and assigned them redshift quality $\text{NQ} = 4$;
- The Large Early Galaxy Astrophysics Census (LEGA-C; van der Wel et al. 2016): galaxies were selected with $f_{\text{use}} = 1$. We assigned $\text{NQ} = 4$ to all sources;
- Australian Dark Energy Survey (OzDES; Lidman et al. 2020): galaxies were selected in two patches partly

⁸ Available from <https://www.eso.org/sci/activities/garching/projects/goods/MasterSpectroscopy.html>

⁹ https://www.eso.org/sci/activities/garching/projects/goods/MASTERCAT_v3.0.dat

overlapping with the KiDZ around CDFS and the VVDS 2h field, with required quality $qop \in \{3, 4\}$. An additional selection of $z > 0.002$ was made to exclude stellar contaminants. We assigned $NQ = qop$;

- SDSS (Abolfathi et al. 2018): galaxies from the 14th Data Release were selected with $zWarning = 0$ and $zErr > 0$. We furthermore required that $zErr < 0.001$, $zErr/z < 0.01$, and $z > 0.001$. We assigned $NQ = 4$ to all such selected galaxies;
- VANDELS survey (Garilli et al. 2021): galaxies were selected with $(zflg \bmod 10) \in \{2, 3, 4\}$. We assigned $NQ = 4$ if $(zflg \bmod 10) \in \{3, 4\}$, and $NQ = 3$ otherwise. The reassignment of the quality flags here was motivated by the reportedly high redshift confidence of objects with flag values of two and three.
- VIPERS (Scodreggio et al. 2018): galaxies were selected with $2 \leq zflg < 10$ or $22 \leq zflg < 30$. We assigned $NQ = 4$ if $3 \leq zflg < 5$ or $23 \leq zflg < 25$, and $NQ = 3$ otherwise;
- VIMOS Ultra Deep Survey (VUDS; Le Fèvre et al. 2015): galaxies were selected with flag $zflags$ ending with $\{3, 4, 9\}$ (reliability $\geq 80\%$) and assigned $NQ = 4$ if $3 \leq zflags < 5$ or $13 \leq zflags < 25$, and $NQ = 3$ otherwise;
- VVDS (Le Fèvre et al. 2005, 2013): galaxies were selected from the combined WIDE, DEEP, and UDEEP sub-samples, with $ZFLAGS \in \{3, 4, 23, 24\}$. We assigned $NQ = 4$ to all sources.
- zCOSMOS: galaxies were selected from a compilation of public (Trump et al. 2009; Comparat et al. 2015; Lilly et al. 2009) and proprietary¹⁰ spectra in the COSMOS field, kindly provided to us by Mara Salvato, updated as of 1 September 2017. That dataset includes some of the surveys already included in our compilation, but also provides redshifts from various other campaigns. We used the provided quality flag and selected galaxies with $3 \leq Q_f \leq 5$, or $13 \leq Q_f \leq 15$, or $23 \leq Q_f \leq 25$, or $Q_f \in \{6, 10\}$. We rejected sources with low-confidence redshift estimate (e.g. from grism spectroscopy), and limit the galaxies to $z_{spec} > 0.002$ to avoid stellar contamination. We assigned redshift quality as $NQ = \min((Q_f \bmod 10), 4)$;

When combining the above spectroscopic samples, we removed both internal (i.e. within the same input catalogue) and external (i.e. in different input catalogues) duplicates. In the latter case, we assigned the most reliable measurement per source based on a specific ‘hierarchy’. Namely, we joined the catalogues by cross-matching objects within $1''$ radius and apply the following order of preference:

- GAMA takes precedence over others, followed by SDSS; then:
- COSMOS field: G10-COSMOS > DEIMOS > hCOSMOS > VVDS > Lega-C > FMOS > VUDS > C3R2 > DEVILS > zCOSMOS;
- CDFS field: ACES > VANDELS > VVDS > VUDS > GOODS CDFS > DEVILS > OzDES;
- VIPERS_W1 field: VIPERS > VVDS > C3R2 > DEVILS > OzDES.

This hierarchy was followed regardless of the relative quality flags between different surveys. For objects with multiple spectroscopic measurements within a particular survey, we either selected the redshift with the highest quality flag or, if various entries for the same source have the same quality flag and the

reported redshifts differ by no more than 0.005, we used the average of the provided redshift estimates. If the reported redshifts have the same quality flag but differ by more than 0.005, we excluded the source from the compilation.

The final spectroscopic sample consists of 635 099 spectroscopic redshift estimates taken from 22 samples. Table 10 lists the statistics for this full compilation, prior to matching with sources detected in our KiDZ imaging. The table presents the number of spectroscopic redshift estimates, redshift range, mean redshift, and redshift scatter (computed using the NMAD) for each sample. Figure 22 shows the distribution of spectroscopic redshift estimates from the full compilation (blue and red) in the context of the individual KiDZ fields on-sky. The spectroscopic data can be seen to extend beyond the spatial extent of the KiDZ optical and NIR data (grey scale), to maximise the cross-match between available spectroscopic redshift estimates and extracted KiDZ photometric sources.

Finally, the spectroscopic redshift estimates were cross-matched to the sources detected in the KiDZ THELI-Phot-reduced images using a simple sky match at $1''$ radial tolerance¹¹. The matched sources are shown in Fig. 22 as blue points. The final sample of spectroscopic redshift estimates after matching to the KiDZ photometric data contains 126 085 sources; the statistics for these redshift estimates are also provided in Table 10.

Figure 23 shows the cross-matched KiDZ spectroscopic redshift estimates in the RA- z plane, demonstrating the relative depth of the various fields, and large-scale structures contained within them. For comparison, the spectroscopic redshift estimates that were available for previous KiDS analyses are shown in orange.

6. Optical and NIR

At this stage in the reduction process we have calibrated optical imaging and masks in the $ugri_1i_2$ -bands from THELI and ASTROWISE, source catalogues extracted with SOURCE EXTRACTOR from THELI imaging, forced GAAP photometry in the $ugri_1i_2$ -bands from ASTROWISE, and calibrated $ZYJHK_s$ -band detectors from KVIPIE. These data products are prepared for all pointings in KiDS and KiDZ. The remaining tasks are therefore the measurement of GAAP photometry for all sources in the $ZYJHK_s$ -bands (on individual VIKING chips), combination of per-chip flux estimates for individual sources (per band), correction of fluxes for Galactic extinction, estimation of photometric redshifts, and construction of final ten-band masks per tile. Each of these tasks is performed within our new post-processing pipeline PHOTOPIPE.

6.1. The PhotoPipe repository

The PHOTOPIPE repository was constructed from scripts used for KiDS DR4, and is publicly available on GitHub¹². The pipeline serves one primary function: to ensure that reduction is performed consistently across a wide array of complex steps and processes, even under significant changes to underlying datasets and/or methodology choices.

With regard to the scripts used for DR4, significant changes implemented in PHOTOPIPE include updated selection of source

¹⁰ Sources from this sample that are proprietary are tagged with an additional flag, and are only provided to members of the community with allowed access (or after the public release of these data).

¹¹ In practice, the matching of the spectroscopic compilation to the KiDZ sources happens after the computation of multi-band information for the KiDZ fields, including masks, discussed in Sect. 6. Nonetheless we include this step here for simplicity.

¹² <https://github.com/KiDS-WL/PhotoPipe>

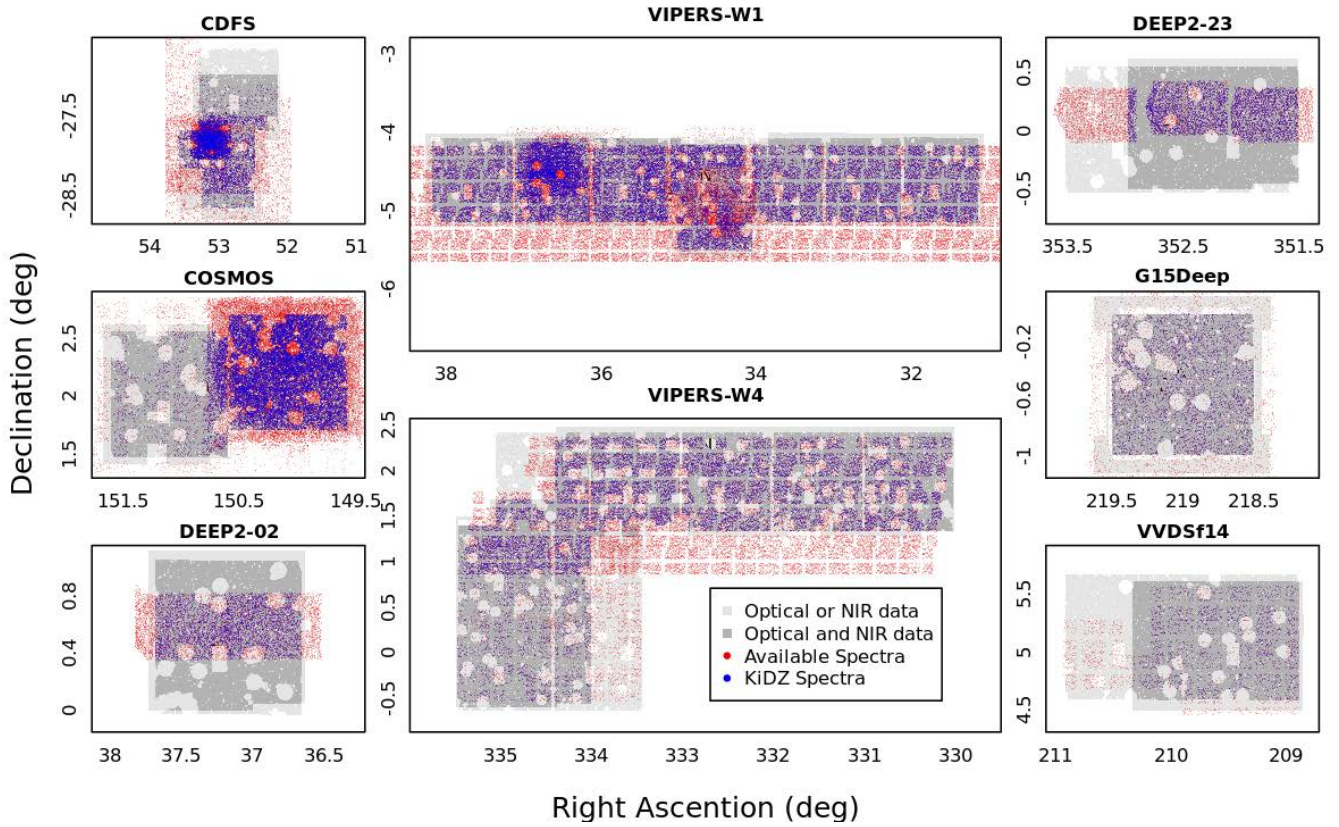


Fig. 22. Distribution of KiDZ spectroscopic redshift estimates on-sky. The figure shows the distribution of all available spectroscopic redshift estimates from the full spectroscopic compilation (red) and those that are matched to unmasked KiDZ sources (blue). The available footprint of the KiDZ imaging is shown in grey scale beneath the points, demonstrating where we have imaging but no available spectra (and vice versa).

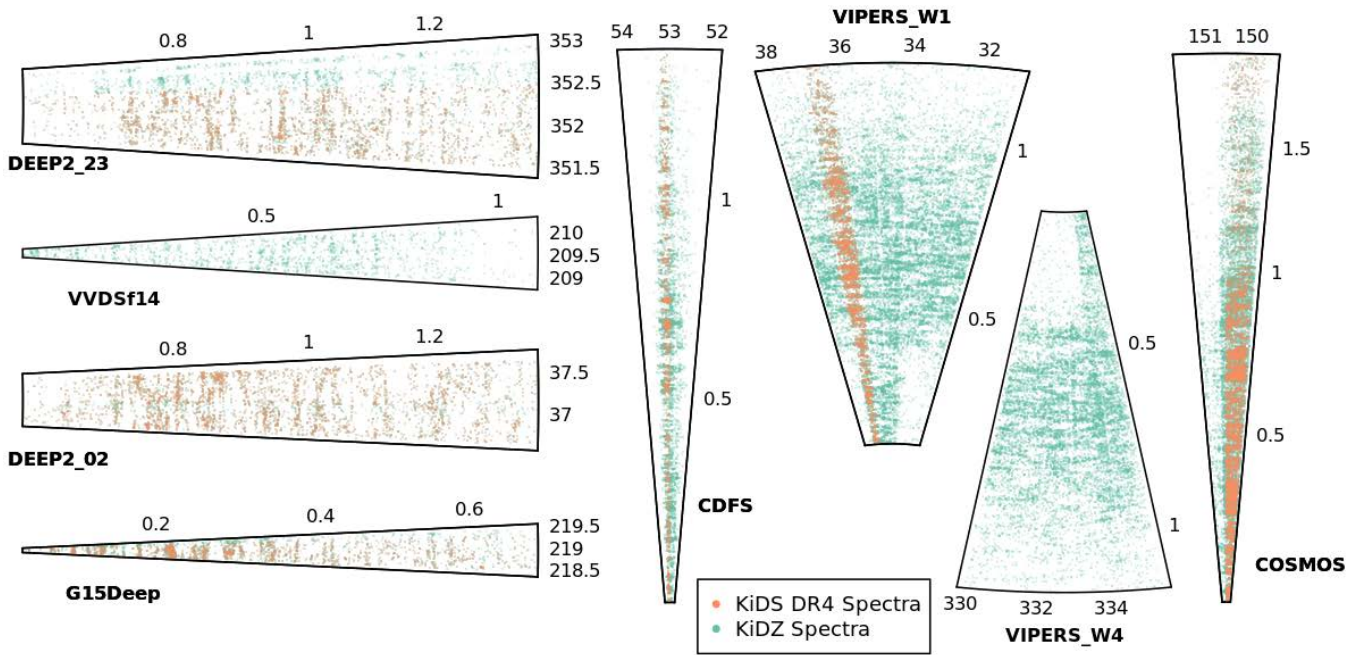


Fig. 23. Distribution of KiDZ spectroscopic redshift estimates in RA- z space. Sources in orange indicate estimates that were available for previous KiDS analyses, and sources in cyan show new estimates added here.

apertures used for photometry (Sect. 6.2), new masking of pathological photometric failures (Sect. 6.4), and updated calibration for THELI-Lens ‘auto’ magnitudes.

6.2. NIR GAAP photometry

The VIKING NIR data are treated differently than the optical data due to the inherently different dither patterns of KiDS and VIKING. As discussed in Sect. 3.1, KiDS was designed with small dithers that ensure similar PSF properties are combined at each point in the focal plane, thereby improving the stability of the final co-added images and resulting in stacks that mostly have smoothly varying PSF. In contrast, as discussed in Sect. 4, VIKING tiles require dithering by a fair fraction of the field of view, resulting in large discontinuous variations in the PSF across the mosaicked tile. Rather than modelling these complex PSF patterns directly (required for the Gaussianisation process performed by GAAP), we opted to extract the NIR photometry from the individual chips of each exposure, and perform an optimal averaging of the flux measurements at the catalogue level; that is, we performed our mosaic construction in catalogue space.

GAAP yields a flux measurement for each object and each exposure on which it appears, typically resulting in two observations for each object in the $ZYHK_s$ -bands, and four observations in the J -band. These fluxes are then combined using an approximately¹³ optimal inverse variance weighting, based on the flux error reported by GAAP.

One complication in this process, already described in Kuijken et al. (2019), is the fact that GAAP will fail to produce a flux estimate in band X if the aperture size, which is set by the r -band seeing, is smaller than the X -band seeing. Furthermore, short of complete failure, the GAAP flux measurements also become extremely noisy when the aperture size is comparable to the PSF size residual, as flux information is being determined by few data pixels. Such circumstances are infrequent but nonetheless bothersome: inspection of the PSF distributions in Figs. 7 and 21 demonstrate that the NIR PSFs consistently exceed the median r -band seeing by more than 0.7 for roughly 10% of the observations (closer to 15% in the Z -band). To remedy this behaviour the pipeline runs GAAP twice, implementing two different choices of minimum aperture sizes: 0.7 and 1.0. This forces the smallest sources to have larger apertures in the 1.0 case, in an effort to suppress the bias and failure rate (the NIR PSF sizes are only more than 1.0 larger than the median r -band seeing in less than 1% of observations). Once all flux measurements are available in all bands, the choice of which aperture to use for flux measurement is taken on an object-by-object basis, using the following criterion:

$$\begin{cases} [\min(R) < 1/\max(R)] \parallel [\max(R) < 0] & \text{use } r_{\min} = 1.0 \\ \text{otherwise} & \text{use } r_{\min} = 0.7 \end{cases}, \quad (13)$$

where R is the vector of flux error ratios for the two apertures in each band: $R = \{\sigma_X^{0.7}/\sigma_X^{1.0}\} \forall X \in \{ugri_1i_2ZYJHK_s\}$ (a negative R identifies cases where all 0.7" apertures fail to yield a flux.) This approach maximises the number of objects with high-S/N GAAP flux measurements in all bands, and yields the best possible photo- z from the combined KiDS+VIKING data.

¹³ Insofar as the optimal estimator would require knowledge of the true measurement variance of the sources, whereas the flux error reported by GAAP is only a partial reflection of the true variance, as it does not account for all sources of flux uncertainty.

To demonstrate the quality of the ten-band photometric catalogues and the cross-survey calibration, in Fig. 24 we present an optical and NIR colour-colour diagram for KiDS DR5 sources. The figure shows the $u - g$ and $g - i$ optical colours versus the $J - K_s$ colour from VIKING. The figure shows the standard stellar locus and galaxy colour distributions, with the sources that match to known SDSS stars highlighted with red contours. We also overlay the colours of all Pickles main sequence stellar templates (Pickles 1998), as observed through the relevant optical and NIR filters. We note the excellent agreement between the Pickles templates and the observed stellar locus: the only potential discrepancy is a small offset ($\Delta u \leq 0.05$) in the u -band, where (uncertain) modelling of metallicity in the stellar population can cause significant bias. Nonetheless, the u -band discrepancy is consistent with our systematic uncertainty of $\Delta u_{\text{sys}} = 0.05$ (Sect. 3.5.2).

6.3. Photometric redshift estimation

Photometric redshifts were estimated in a very similar way as in KiDS DR4, with the primary difference being the incorporation of the second pass i -band information. We ran the Bayesian photo- z (BPZ; Benítez 2000) code on the ten-band GAAP photometry using the Bayesian prior presented in Raichoor et al. (2014), a maximum redshift of $z = 7$, and a redshift stepping of $\Delta z = 0.01$. The two i -band measurements are treated independently, relying on BPZ to optimally use the information based on the two GAAP magnitude errors. Providing the two fluxes independently in this way is mathematically equivalent to providing BPZ with a single variance-weighted average of the two i -band flux measurements.

The resulting photo- z point-estimates were compared to the available KiDZ spec- z (Sect. 5), and summary statistics for the various fields and full compilation are presented in Fig. 25. The figure contains both summary statistics computed for the entire compilation of spec- z present in all of the KiDZ fields and for statistics computed in the individual KiDZ fields. In all cases the samples are binned into quasi-equal- N bins, which contain at most 2000 sources per bin. In the event that a field has fewer than 4000 spectra, the sample is split into two equal- N bins. This has the effect of producing short straight lines for some fields in the figures.

We quantify the photo- z quality via the statistics of $\Delta z = (z_B - z_{\text{spec}})/(1 + z_{\text{spec}})$. In particular, we report in Fig. 25 the median, NMAD scatter, and rate of outliers (objects with $|\Delta z| > 0.15$) as a function of r -band magnitude, spectroscopic redshift, and photometric redshift. The well-known dependence of photo- z quality on photometric S/N, here parametrised by the r -band magnitude, is clearly visible for $r > 22$. For $r < 22$ the photo- z scatter and outlier rate are roughly constant, at $\sigma_{\Delta z} = 0.05$ and $\eta_{\Delta z}^{0.15} = 0.1$, respectively, indicating that other sources of error limit the precision of the photo- z at bright magnitudes. Of further note is the increase in photo- z bias seen a magnitudes brighter than $r = 20$; here the photometric pipeline, and photo- z priors, used for the KiDS-DR5 sample are not optimal, having been chosen to optimise performance at fainter magnitudes (see e.g. Fig. 10 of Wright et al. 2019).

The dependence of photo- z quality on spectroscopic and photometric redshift is more complicated, showing several features that we attribute to the non-trivial interplay of the specific filter set, depth in the different bands, the spectral energy distribution templates, and the real galaxy population. Only at high redshift and/or photo- z can the behaviour be easily attributed to S/N

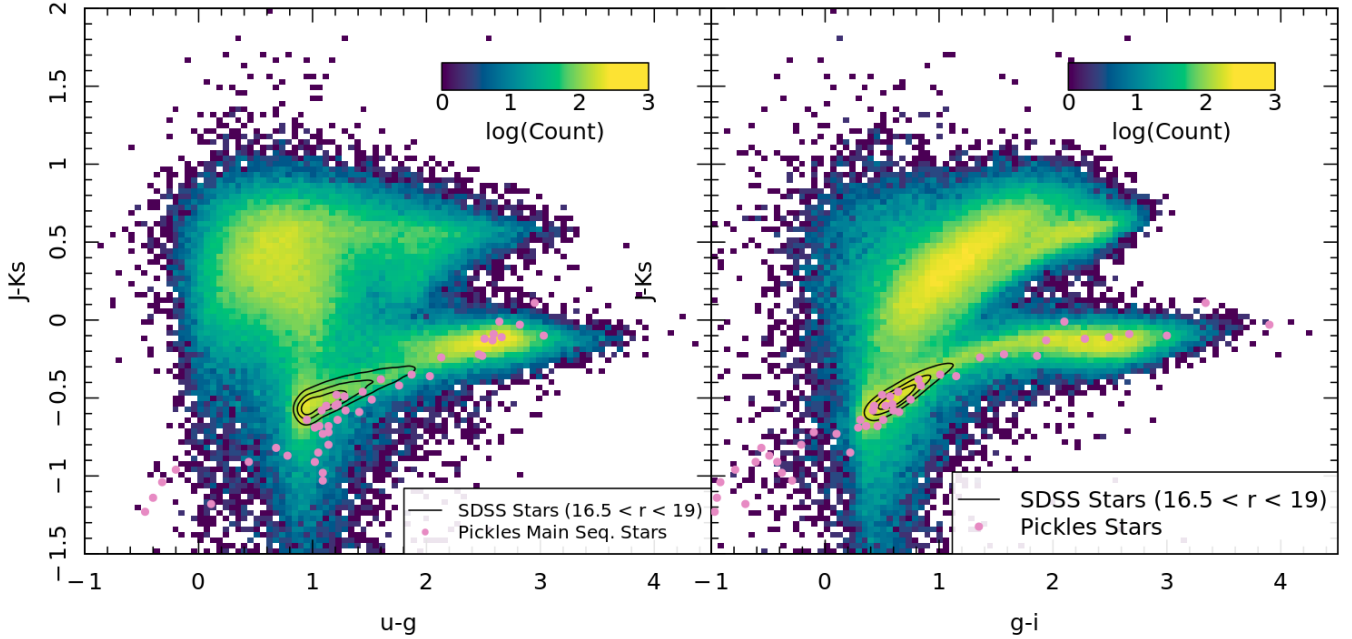


Fig. 24. Colour-colour diagrams for KiDS DR5 sources. The stellar locus can be identified by the distinct clouds of data that are coincident with both SDSS stars (black contours) and Pickles stellar templates (pink dots).

effects, visible by the consistent degradation of quality in all KiDZ fields.

We find photo- z biases at the few percent level over all magnitudes, but which exacerbate at the extremes in photo- z ($\mu_{\Delta z} > 0.05$) and redshift ($\mu_{\Delta z} < -0.05$). These biases are typically accompanied by an increase in photo- z scatter, shown in the second row of the figure. Beyond spectroscopic and photometric redshifts (i.e. at $z_B \geq 1$ and $z_{\text{spec}} \geq 1$), the scatter in the photo- z can be seen to deteriorate for all available surveys due to the increased photometric noise of these distant objects. Of particular note, though, is the increased scatter and outlier rate observed in the COSMOS sample at $z_B \approx 0.6$ and $z_{\text{spec}} \approx 0.7$. Indeed, in this region the COSMOS sample of sources display roughly twice the number of outliers ($\eta_{\Delta z}^{0.15} \approx 0.25$) as other samples ($\eta_{\Delta z}^{0.15} \lesssim 0.13$). This may simply be indicative of different selection effects in the COSMOS field spectroscopic sample, compared to the other fields (as the COSMOS sample is made up of many individual spectroscopic samples with a range of targeting criteria). However, the increase in scatter is coincident with an overdensity in the large-scale-structure, visible as a spike in the COSMOS number counts in the bottom row (and visually in the redshift-axis diagrams shown in Fig. 23). This increase in scatter may therefore be caused by the overdensity, which causes sources at this redshift to over-sample a particular subset of colour redshift degeneracies (specifically those degenerate with red galaxy spectra at redshift $z \approx 0.7$).

Regarding the inclusion of the additional i -band information in the computation of the photo- z , we find that the summary statistics presented here change only slightly when moving between nine and ten bands. We find, however, that this is primarily due to the strong selection effects present in the sample that we use to construct this data-side photo- z statistics. As a demonstration of the expected performance improvements in the photo- z for the wide-field photometric, we can utilise simulations. Using the simulated KiDS galaxy sample presented in van den Busch et al. (2020), we computed our BPZ photo- z with and without an additional i -band photometric realisation.

These results are presented in Fig. 26. We find a consistent 5–10% reduction in photometric scatter and outlier rate at magnitudes fainter than $r = 22$, where the additional noise realisation provides useful information. More significantly, however, the improvement seen in the photo- z as a function of true redshift: we see a consistent 10–20% reduction in scatter at $z > 0.7$, coupled with 5–20% reduction in outlier rate in the same redshift range. This is understandable given the propagation of the major spectral features as a function of redshift, mainly the 4000Å and Balmer breaks: these features enter the i -band at approximately $z = 0.7$, and beyond this redshift the additional photometric information allows us to better localise the break (and its absence, for $z > 1.2$).

We stress here that these biases are not reflective of the bias imprinted on typical cosmological measurements from weak lensing with KiDS. The template-based photo- z shown here are typically only used to define subsamples of galaxies with redshift distributions that are largely localised and distinct along the line of sight. The redshift distributions themselves are determined via empirical methods making direct use of the spec- z compilation, which here is only employed for validation. Furthermore, the underlying magnitude, colour, and morphology distributions of the sources that Fig. 25 is based on are inherently different from typical weak lensing source samples, due to selection effects and weighting. As a result, the comparison of photo- z and spec- z shown in Fig. 25 does not directly translate into the photo- z quality of galaxy samples used for cosmological measurements, but simply illustrates the raw performance on the spec- z without accounting for these differences.

Finally, we note that there are ongoing efforts to produce additional photo- z estimates within the KiDS collaboration, aimed at improving photo- z performance via sample selection (see e.g. Vakili et al. 2023), using machine learning of fluxes (see e.g. Bilicki et al. 2021), and using machine learning directly on images (see e.g. Li et al. 2022). These efforts, however, will not produce photo- z estimates that are included in the formal DR5 ESO release. Nonetheless, such additional photo- z estimates will

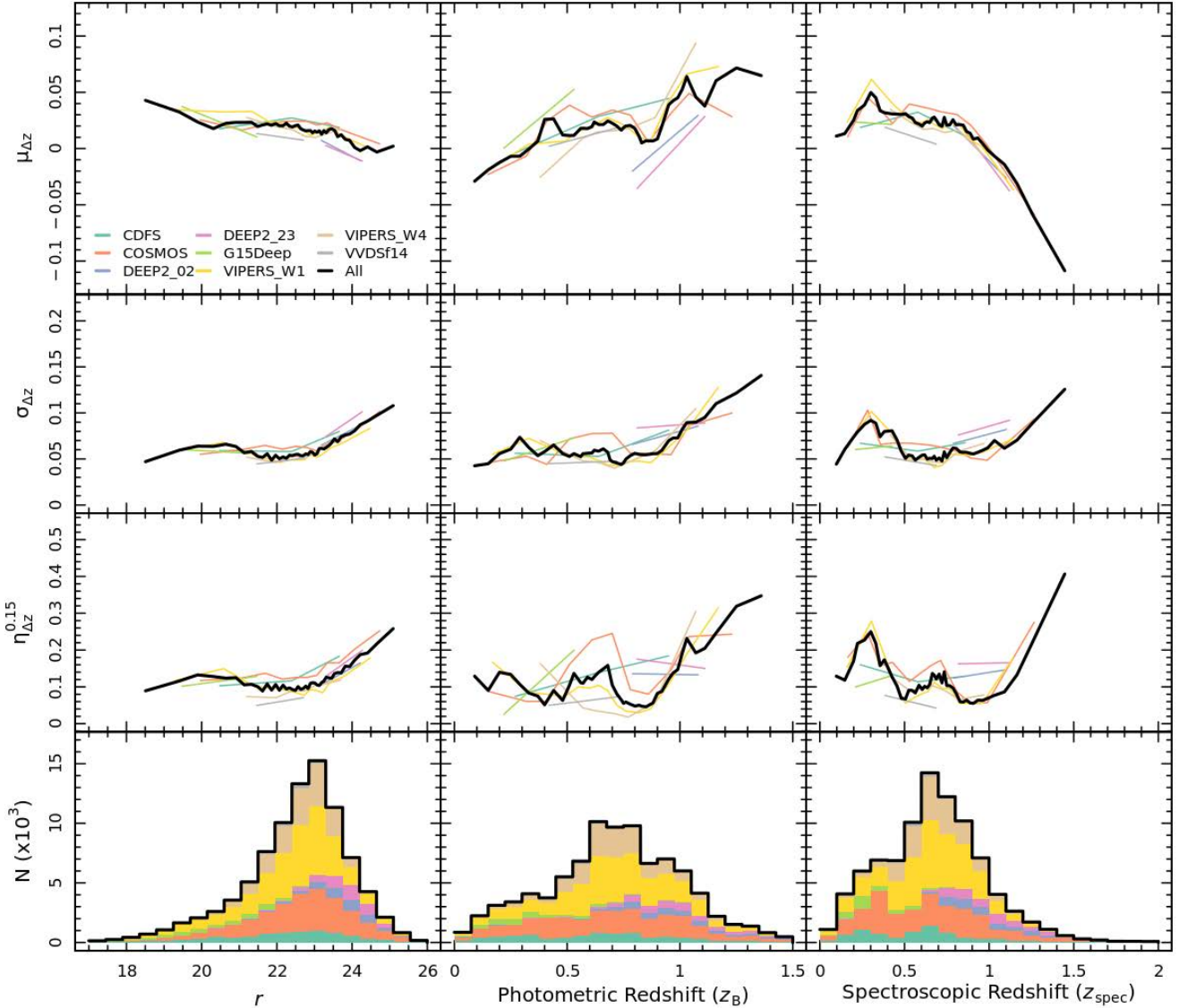


Fig. 25. Quality metrics for photo- z point-estimates produced by BPZ in PHOTPIPE. Each row shows one quality metric, computed from the distribution of $\Delta z = (z_B - z_{\text{spec}})/(1 + z_{\text{spec}})$: the running median (‘bias’, μ), the running normalised median-absolute-deviation (‘scatter’, σ), and the fraction of sources with $|\Delta z| > 0.15$ (‘outlier rate’, $\eta^{0.15}$). The columns show these statistics computed as a function of r -band magnitude (left), photo- z point-estimate (z_B , centre), and spectroscopic redshift (z_{spec} , right).

undoubtedly be of value to the community as they are designed to produce higher accuracy and precision for (in particular) samples brighter than $r = 20$.

6.4. Ten-band mask construction

We adopt a bit-masking scheme for the FITS masks that encodes the availability and quality of data at a given sky position. The 16 bits corresponding to different image defects are described in Table 11. Compared to previous KiDS data releases, we reorder the bits to facilitate the masking of a fiducial sample for typical weak lensing applications by selecting objects with

$$\text{MASK} \leq 1. \quad (14)$$

For compatibility with previous data releases and codes, the fiducial mask as a bitwise-logical statement is

$$\text{MASK} \& 32766 \quad (15)$$

(corresponding to 0111 1111 1111 1110₂). The first four bits encode objects and defects that are present in the THELI r -band detection image. The zeroth bit (MASK = 1) corresponds to faint halos created by reflected starlight that typically do not significantly bias flux and shape measurements, and is therefore excluded in the fiducial masking scheme given in Eq. (14). The first bit (MASK = 2) corresponds to brighter halos from reflections and the direct bright starlight. The second bit (MASK = 4) encodes areas that are masked out manually due to defects that are not captured by any of the automatic algorithms, and also rejects area that is affected by resolved dwarf galaxies or globular clusters, which confuse our automatic star-galaxy separation and subsequent PSF measurement algorithms due to the abnormally high stellar density. The third bit (MASK = 8) encodes areas that are removed from consideration by the THELI masking of regions with abnormally low number densities (the ‘void mask’), sources automatically flagged as being asteroids, and areas of THELI imaging with zero weight (caused by, for example, chip gaps

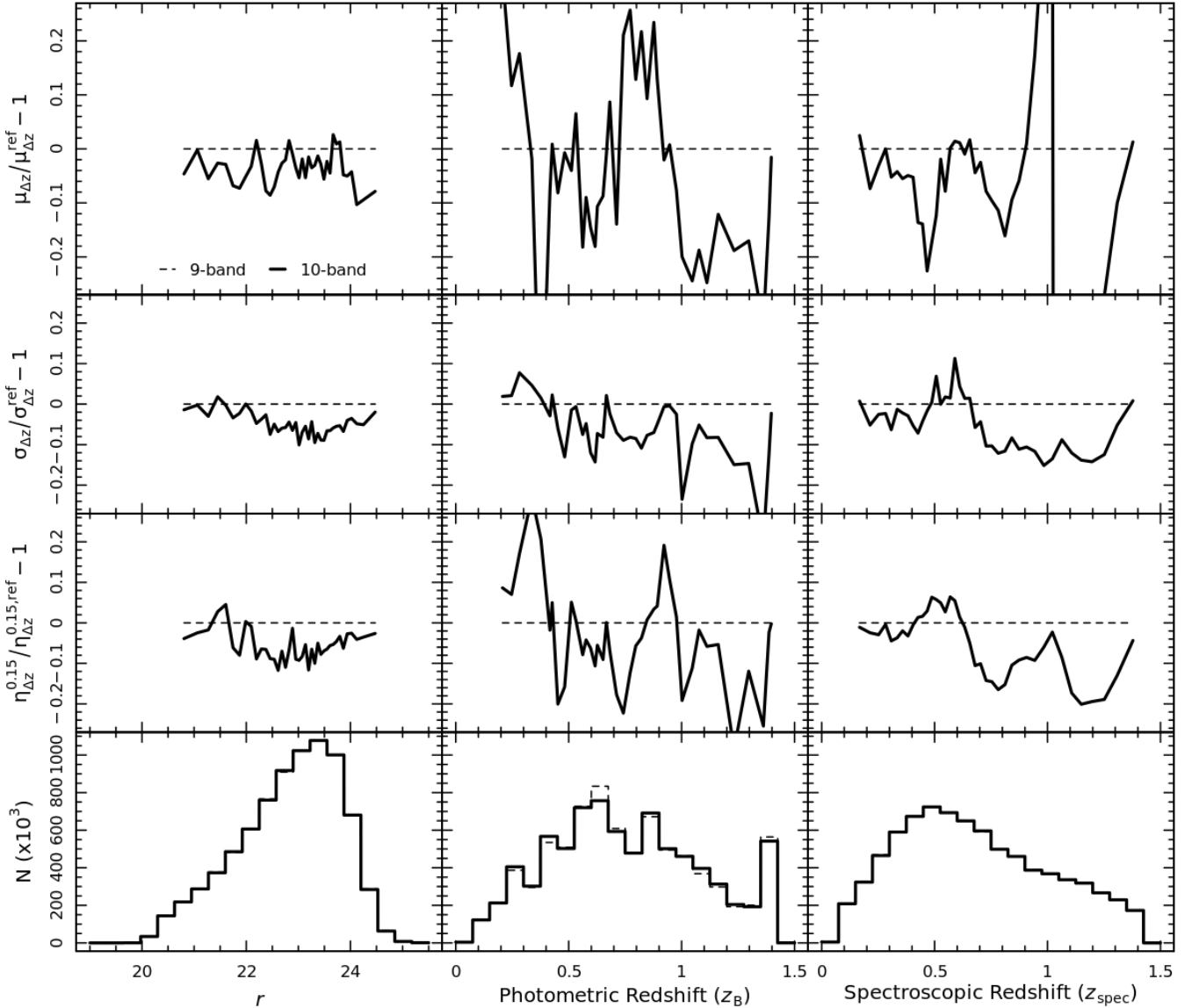


Fig. 26. Differences between photo- z point-estimates produced by BPZ in PHOTOPIPE computed without the second pass i -band (‘9-band’) and with the second pass i -band (‘10-band’), using a full simulated KiDS wide-field sample described in van den Busch et al. (2020). Metrics here are not directly comparable to Fig. 25 due to this sample being simulated, having a different redshift baseline, and representing a full wide-field sample (unmatched to spectroscopy). Therefore, only relative differences between the nine- and ten-band cases are relevant. Metrics are computed as in Fig. 25, except metrics here are shown as percentage difference with respect to the nine-band case.

and detector saturation). The following ten bits (MASK = 16 to MASK = 8192) correspond to defects in the ten $ugri_1i_2ZYJHK_s$ -bands that the GAAP photometry is extracted from. The final, 14th bit that is used (MASK = 16384) corresponds to area that is outside the limits of each pointing, chosen to knit together the overlapping VST tiles without source duplication. These are referred to as the ‘WCS’ cuts, as they are defined using constant limits of RA and Dec in the World Coordinate System (WCS) of the individual tiles.

One significant change in the masks from previous KiDS releases is the removal of redundancy between the r -band mask created for the THELI detection images (bits 0–3) and the $ugri_1i_2$ -band masks created for the ASTROWISE images that are used for the optical GAAP photometry (bits 4–8). Previous releases ignored the faint stellar reflection halos from the ASTROWISE r -band masks (bit 5), instead relying on the ability of the THELI r -band mask to capture those effects in bits 1 and

0 (MASK = {1, 2}). This choice was made because the default ASTROWISE masking was slightly too conservative, correlating more with the THELI-Lens conservative mask (MASK = 1) than with the standard THELI mask (MASK = 2). This effect was largely resolved with the update of the PULECENELLA mask parameters (Sect. 3.5.3), but nonetheless persists at low levels. In DR5, we apply all ASTROWISE mask bits in our fiducial mask. If we were to remove the ASTROWISE r -band masks from the fiducial masking, as in DR4, we would free up 16.0 deg² of data over the full survey area ($\sim 1.6\%$). The final correlations between the various bits of the MASK are provided in Appendix F.

Another addition to the masks for DR5 is the inclusion of a ‘strong-selection’ mask, which is designed to remove areas of the survey where there have been pathological photometric measurement failures within GAAP. These failures are related to the difference in the PSF sizes between the detection r -band images and the other bands, as discussed in Sect. 6.2. As discussed,

Table 11. KiDS+KiDZ MASK bits, their names, and additional information.

Bit	(value)	Band	Name	Additional info
0	(1)	THELI r -band	starhalo (cons.)	$10.5 < r < 11.5$ of UCAC4 and GSC1 stellar catalogue
1	(2)	THELI r -band	star & starhalo	$r < 10.5$ (14.0) of UCAC4 and GSC1 for stars (starhalos)
2	(4)	THELI r -band	manual mask	
3	(8)	THELI r -band	void & asteroid	contains weight==0 (saturation, chip gap, etc) selection
4	(16)	ASTROWISE u -band	auto masks	star & starhalo & weight & manual
5	(32)	ASTROWISE g -band	auto masks	star & starhalo & weight & manual
6	(64)	ASTROWISE r -band	auto masks	star & starhalo & weight & manual
7	(128)	ASTROWISE i_1	auto masks	star & starhalo & weight & manual
8	(256)	ASTROWISE i_2	auto masks	star & starhalo & weight & manual
9	(512)	VISTA Z -band	footprint mask	contains strong-selection mask
10	(1024)	VISTA Y -band	footprint mask	contains strong-selection mask
11	(2048)	VISTA J -band	footprint mask	contains strong-selection mask
12	(4096)	VISTA H -band	footprint mask	contains strong-selection mask
13	(8192)	VISTA K_s -band	footprint mask	contains strong-selection mask
14	(16384)	THELI r -band	WCS tiling cuts	
15	(32768)	Reserved, unused		

this effect led initially to the requirement of measuring fluxes in apertures with two minimum radii (see Sect. 3.6), but in some cases even the use of the larger aperture is not sufficient to stop GAAP from being unable to measure a flux. As these failures are related to the size of the object aperture, when such failures occur, they preferentially affect the smallest objects first. This can imprint a redshift-dependent bias on the source distribution that is localised in particular areas on-sky. Such a selection bias is problematic for (primarily) photometric clustering studies, but also for cosmic shear (albeit at higher order) as it exacerbates variable depth (see e.g. Heydenreich et al. 2020).

We therefore opted to remove areas of the survey where such pathological failures on-sky are detected. We quantify this failure using a 2D kernel-density estimate (KDE) measurement of the source density on-sky before and after selection due to GAAP photometric failures. If the ratio of these two KDEs is less than 0.8 (i.e. more than 20% of the sources, per square arcminute, have been removed due to photometric failures), we flag these regions and remove them from the survey. This masking can be seen in Fig. 27, where we show the effect of this masking for one of the KiDZ tiles in VIPERS. The figure shows the distribution of all sources in the tile, without any preselection. The sources are coloured by whether or not they have a photometric failure in the Y -band: blue have successful photometric measurements, and red have failed measurements. In the background of the image we show the Y -band sum-image for this field, after the application of the WCS cuts (bit 14). This shows the regular tile pattern formed by the dithering of many paw-prints (Sect. 4). It is apparent from the photometric failures that there are two paw-prints that have particularly poor seeing, which causes sources covered only by these paws to have failed photometric measurements. These areas of the sky have been identified by the strong selection algorithm, and these areas have been masked (leading to the sum-image being zero, white, in these areas). We note that the sources have not all failed in these regions; however, the pathological nature of the selection means that we have nonetheless removed these areas from consideration. It should also be stressed that this strong-selection masking does not correlate with cosmic large-scale structure, as it is driven by observing conditions. As such, it is unlikely to introduce any bias in cosmological measurements from weak lensing.

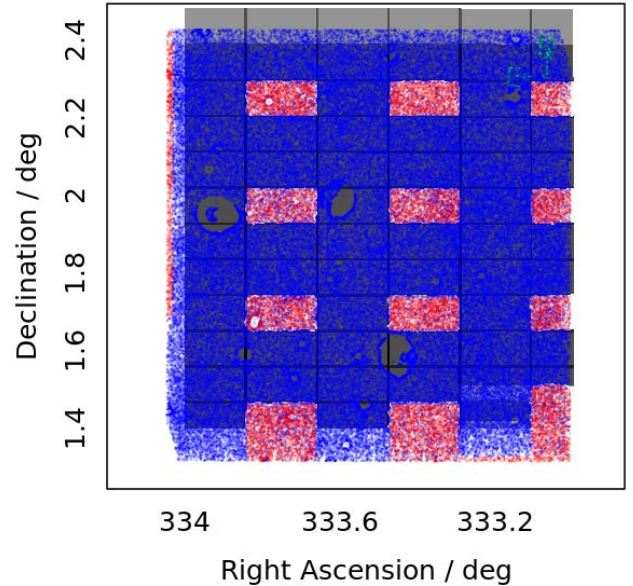


Fig. 27. Demonstration of the strong-selection masking in the KiDS+KiDZ catalogues for KiDZ pointing KIDZ_333p0_1p9 in the Y -band. All sources detected by SOURCE EXTRACTOR are shown coloured by the number of missing photometric bands as determined by GAAP (0: blue, 1: red). The mask is determined by the fraction of sources per unit area on sky that are missing one or more bands, as determined using a ratio of KDEs constructed on a $1' \times 1'$ grid with a $1'$ Gaussian kernel. The strong-selection mask removes regions with missing-source fractions greater than 20%. The effect of this mask can be seen in the background, which shows the sum image of this field's Y -band data after masking: areas that show the strong selection effect (i.e. red dots) have a zero value in the sum image (i.e. white), and are therefore masked.

6.5. Mosaic mask construction

Finally, we constructed mosaic masks for science use and calculation of the fiducial survey area for KiDS. Mosaic masks are constructed using SWARP combining all individual ten-band masks using a 'minimum' combination. Masks are constructed on a predefined WCS, using an Aitoff projection and a $6''$ pixel resolution.

Table 12. Mosaic areas computed from the 6'' mosaic masks.

Selection name	Selection bits	Selection sum	North abrea (deg ²)	South area (deg ²)	Total area (deg ²)
WCS	14	16 384	681.098	649.883	1 330.981
WCS+ASTROWISE	4,5,6,7,8,14	16 880	608.668	588.015	1 196.682
WCS+NIR	9,10,11,12,13,14	32 256	645.583	608.975	1 254.558
WCS+AW+NIR	4,5,6,7,8,9,10,11,12,13,14	32 752	577.319	550.902	1 128.220
WCS+THELI	1,2,3,14	16 398	583.059	562.348	1 145.407
WCS+THELI Cons.	0,1,2,3,14	16 399	524.231	508.495	1 032.726
WCS+AW+THELI	1,2,3,4,5,6,7,8,14	16 894	544.850	529.992	1 074.842
WCS+THELI+NIR	1,2,3,9,10,11,12,13,14	32 270	553.478	527.155	1 080.633
WCS+THELI+AW+NIR	1,2,3,4,5,6,7,8,9,10,11,12,13,14	32 766	517.342	496.689	1 014.031
WCS+TH.Cons.+AW+NIR	0,1,2,3,4,5,6,7,8,9,10,11,12,13,14	32 767	468.845	452.350	921.195
WCS+AW <i>u</i>	4,14	16 400	672.796	639.317	1 312.113
WCS+AW <i>g</i>	5,14	16 416	647.942	622.772	1 270.714
WCS+AW <i>r</i>	6,14	16 448	626.232	607.595	1 233.827
WCS+AW <i>i</i> ₁	7,14	16 512	636.101	610.085	1 246.186
WCS+AW <i>i</i> ₂	8,14	16 640	644.359	619.996	1 264.355
WCS+NIR <i>Z</i>	9,14	16 896	649.522	612.493	1 262.016
WCS+NIR <i>Y</i>	10,14	17 408	648.805	612.570	1 261.375
WCS+NIR <i>J</i>	11,14	18 432	650.089	612.589	1 262.678
WCS+NIR <i>H</i>	12,14	20 480	647.457	610.845	1 258.302
WCS+NIR <i>K</i> _s	13,14	24 576	647.737	610.384	1 258.120

Notes. The fiducial KiDS DR5 and KiDS-Legacy MASK selection is given in bold.

Using these mosaic masks, we are able to calculate the area of the KiDS DR5 data when selecting specific mask bits. Table 12 presents a compilation of the available survey areas when masking specific bits, and combinations of bits. We note a few particular examples. First, the area of the survey when considering all available tiles after removing overlaps (‘WCS’, bit 14) is 1331.0 deg². This area is then further reduced when considering masking of artefacts, stars, stellar reflections, and missing chips in the THELI-Lens *r*-band imaging (‘WCS+THELI-Lens’, bits {1–4, 14}) to 1145.4 deg². Considering only the area available to all of the *ugri₁i₂*-bands (‘WCS+THELI-Lens+ASTROWISE’, bits {1–8, 14}), the survey area reduces to 1074.8 deg². After removal of area that does not have overlap with the VISTA *ZYJHK_s*-bands, the survey area is further reduced to the ‘fiducial’ survey area of 1014.0 deg² (‘WCS+THELI+AW+NIR’, bits {1–14}).

7. Shape measurement and legacy sample construction

In this section we detail the construction of the lensing sample that will be used for the fiducial KiDS DR5 cosmological analyses. This ‘KiDS-Legacy’ sample is a subset of the full DR5 sample, determined (primarily) by the availability of reliable shape measurements. We first describe the new version of our fiducial shape measurement code *lensfit* (Miller et al. 2013, Sect. 7.1), followed by the selection of the KiDS-Legacy sample (Sect. 7.2).

7.1. DR5 *lensfit*

Fiducial shape measurement in KiDS DR5 was performed with *lensfit* v321. Demonstration of the properties of this *lensfit* version are provided in Li et al. (2023), where the accuracy of the

measured shapes were tested with simulations. Regarding the use of *lensfit* for selecting the DR5 lensing sample, there are two important changes with respect to DR4 (which was produced using *lensfit* version v309): the values of the *fitclass* flags have been modified, and the calibration procedure for the lensing weight has been modified.

7.1.1. *fitclass*

The new definitions of the *fitclass* flags in *lensfit* v321 are presented in Table 13. These new flags require the selection to be updated, and the choice of which flags to reject is also presented in the table. A total of 31.31% of the available MASK ≤ 1 sample is flagged by the various *fitclass* selections, the most severe of which is the ‘insufficient data’ flag, which is assigned to 20.69% of the MASK ≤ 1 sources. However, the majority of these sources are not modelled by *lensfit* because they are too bright (see Sect. 7.2). In practice, the *fitclass* selection removes 14.03% of sources that would otherwise exist in the sample; the majority of which are either stars (9.07%) or have multiple sources detected within a single segmentation region (and so are deemed to be blended, 3.85%).

7.1.2. Weight recalibration

Galaxy shapes are able to be expressed as ellipticities ϵ , quantified using moments of the light distribution or using direct model fits to galaxy images, and take the form of complex quantities combining the source axis ratio (q) and the major axis orientation angle (θ):

$$\epsilon \equiv \epsilon_1 + i\epsilon_2 = \frac{1-q}{1+q} \exp(2i\theta). \quad (16)$$

Table 13. `fitclass` definitions and statistics output by `lensfit v321`.

Value	Flag meaning	Source Excl.	f_{MASK} (%)	f_{avail} (%)
-10	Multiple objects	Yes	2.91	3.85
-9	Object too-large	No	0.41	0.55
-8	deprecated	No	0.00	–
-7	Centroid mismatch	Yes	0.06	0.05
-6	deprecated	No	3.85	–
-5	deprecated	–	–	–
-4	Poor model fit	Yes	0.32	0.41
-3	Unable to fit	Yes	0.57	0.66
-2	deprecated	–	–	–
-1	Insufficient data	Yes	20.69	–
0	Well modelled galaxy	No	64.43	85.42
1	Star or point-source	Yes	4.59	6.17
2	Star or point-source	Yes	2.16	2.90
Total removed:			31.31	14.03

Notes. The percentage of the full MASK ≤ 1 dataset that is assigned to each of the `fitclass` values is shown (f_{MASK}). The percentage of course that are actually present in the sample prior to the `fitclass` selection (see Sect. 7.2 and Table 14) is also shown (f_{avail}). Classes with – were never assigned to a source within the relevant sample.

The magnitude of any source’s ellipticity is therefore simply

$$|\varepsilon| = \sqrt{(\varepsilon_1)^2 + (\varepsilon_2)^2}. \quad (17)$$

Ellipticities are measured by `lensfit` using a model-fitting approach that is described at length in Miller et al. (2013). Source ellipticities are reported at their posterior-informed maximum likelihood (see Sect. 3.5 of Miller et al. 2013), and the 2D measurement variance of the likelihood in ellipticity is provided as $v_\varepsilon \equiv \sigma_\varepsilon^2$. Faint sources have inherently noisier ellipticity estimates than their brighter counterparts, and this is reflected in them having a broader likelihood surface and a larger variance v_ε . To translate this uncertainty in shape measurement to subsequent scientific analyses, each source i is therefore tagged with a shape measurement weight w_i , which is related to the inverse of the 2D variance estimate. Individual weight estimates are computed as

$$w_i = \left[\frac{v_{\varepsilon,i} \varepsilon_{\text{max}}^2}{2\varepsilon_{\text{max}}^2 - 4v_{\varepsilon,i}} + v_{\varepsilon,\text{pop}} \right]^{-1}, \quad (18)$$

where $|\varepsilon|_{\text{max}}$ is the maximum allowed ellipticity (determined by the intrinsic thickness of edge-on disk galaxies), and $v_{\varepsilon,\text{pop}}$ is the population variance in ellipticity of all galaxies in the sample.

Unfortunately, shapes measured by the `lensfit` algorithm are not free of systematic biases, and therefore require recalibration. Of particular concern is the systematic imprint of the PSF shape on the estimated properties of galaxies, as such systematic contamination of estimated ellipticities and weights would ultimately introduce bias into shear correlation functions (which are a primary probe of cosmology). As such, our recalibration procedure (described in Li et al. 2023, and with minor modifications explained below) is designed to remove residual imprints of the PSF on the data (although the resulting shape estimates nonetheless require a correction for multiplicative shear bias; see Sect. 7.1.4).

Although the KiDS PSF in the r -band is relatively round, any anisotropy can lead to inferred galaxy shapes that are correlated with the orientation of the PSF. This ‘PSF leakage’ affects both the ellipticities per source, and the associated weights. PSF leakage in measured source weights refers to the tendency of sources to be preferentially up-weighted when their on-sky orientation happens to be aligned with the PSF. PSF leakage in the measured source shapes refers to the tendency of detected sources to be preferentially aligned with the PSF, as they have higher effective surface brightnesses in this regime (and thus are less likely to have been missed during source detection).

In KiDS-Legacy, we corrected for PSF leakage in the weights of sources by measuring the linear relationship between the variance of the shape measurement per source v_ε and the scalar projection S of the PSF ellipticity ε_{PSF} in the direction of the galaxy ellipticity ε per source:

$$S_\varepsilon = \frac{\text{Real}(\varepsilon_{\text{PSF}} \varepsilon^*)}{|\varepsilon|}. \quad (19)$$

The relationship between v_ε and S_ε is highly dependent both on the resolution of the source and on the S/N of the `lensfit` model. Here resolution is defined as the ratio of the squared PSF radius (standard deviation) r_{PSF}^2 to the quadrature sum of the PSF radius and the circularised galaxy effective radius $r_{\text{ab}} = r_e \sqrt{q}$ (where r_e is the effective radius) per source:

$$\mathcal{R} \equiv \frac{r_{\text{PSF}}^2}{r_{\text{PSF}}^2 + r_{\text{ab}}^2}. \quad (20)$$

We therefore opted to fit for a linear relationship between v_ε and S_ε in bins of \mathcal{R} and S/N, which we constructed to each contain an equal number of sources. We fitted our linear regression (between all galaxies i that reside within a bin) as

$$v_{\varepsilon,i} = \alpha_S S_{\varepsilon,i} + \langle S_\varepsilon \rangle + \mathcal{N}(0, \sigma_S), \quad (21)$$

where $\langle S_\varepsilon \rangle$ is the additive component of the model, and the sources are expected to be normally distributed about the regression given a standard deviation σ_S ; that is, assuming homoskedasticity of the residuals in the fitting process. This regression provides an estimate of the PSF leakage into the shape measurement variance for sources in this bin of \mathcal{R} and S/N, parameterised by the multiplicative coefficient α_S . Panel a of Fig. 28 shows the value of α_S in our bins of \mathcal{R} and S/N, demonstrating the non-negligible systematic PSF leakage into the weights over this plane.

One development that we have made with respect to the method presented in Li et al. (2023) regards these linear fits. Inspection of the data within each bin demonstrates that the assumption of homoskedasticity is poor: measurement variance is significantly larger for sources with small intrinsic ellipticities, and therefore the residuals in the fit balloon in the middle of the fitting region. Additionally, the variance is a truncated variable, which can lead to further heteroskedasticity in the regression residuals. To combat the influence of these two effects, we implemented a conservative clipping of data with exceptionally high variance, and a correction for heteroskedasticity in the form of iterative regression with residual weighting. The result of these processes (compared to simple linear regression) is presented in panels b and c of Fig. 28.

Once the value of α_S is computed per bin, the original measurement variances (for sources in that bin) are corrected by subtracting the inferred leakage amplitude per source i :

$$v'_{\varepsilon,i} = v_{\varepsilon,i} - \alpha_S S_{\varepsilon,i}. \quad (22)$$

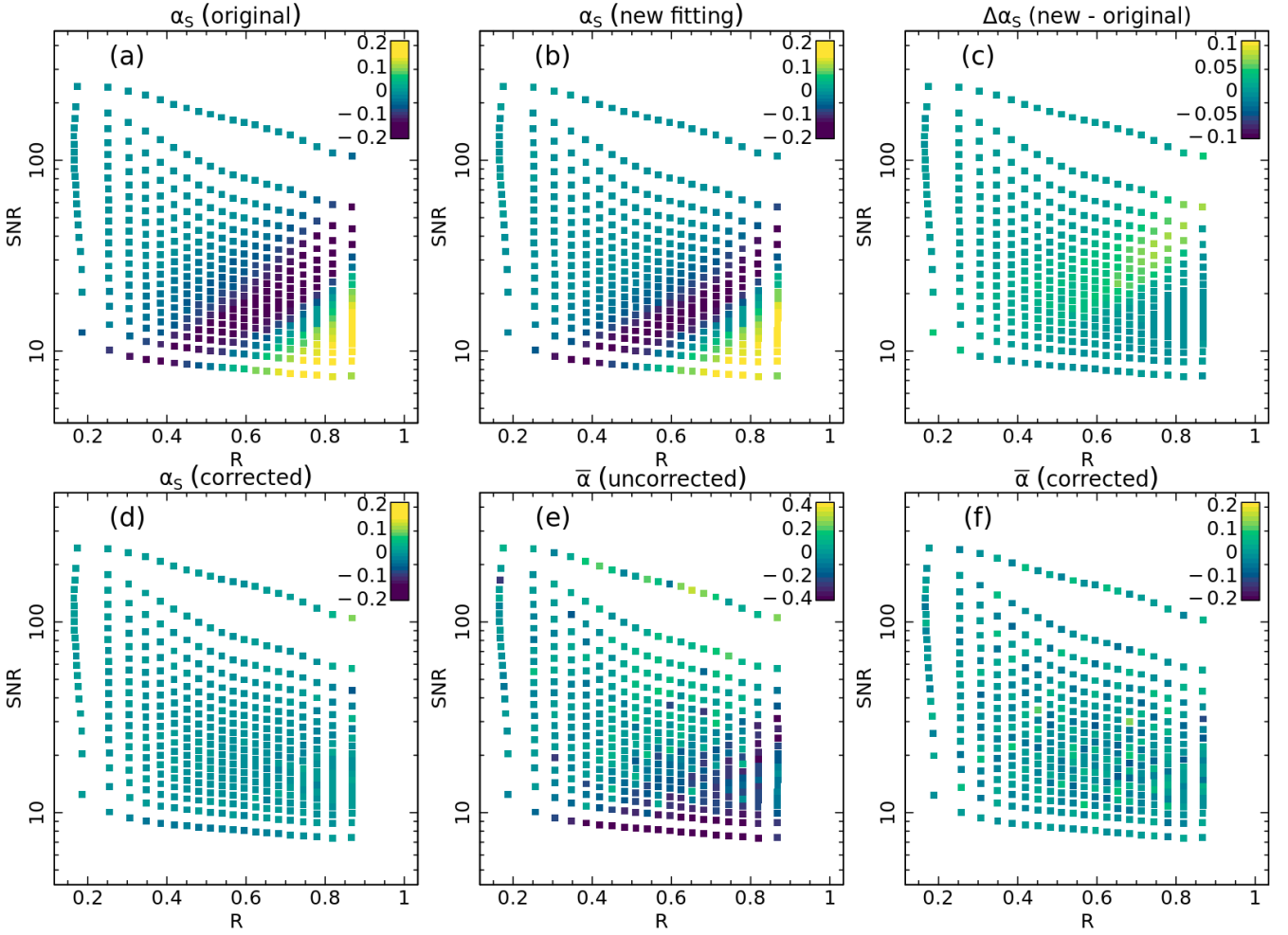


Fig. 28. Sources in the resolution-S/N plane that is used for recalibration of *lensfit* weights and shape estimates. Top row: estimates of the PSF leakage into the shape measurement variance, estimated with simple linear regression (left) and with clipped, iterative, residual-weighted linear regression (centre). Right: difference between the leakage estimates. Bottom row: PSF leakage estimates after correction of shape variances and weights (left). The estimates of PSF leakage into the source shape distributions is shown before (centre) and after (right) correction of estimated ellipticities.

The distribution of α_S values measured after this correction (i.e. re-fitting Eq. (21) with $v_{\varepsilon,i} \rightarrow v'_{\varepsilon,i}$) are shown in panel d of Fig. 28. The corrected *lensfit* measurement variances $v'_{\varepsilon,i}$ are then used to define a corrected shape measurement weight per source w'_i using a modified version of Eq. (18):

$$w'_i = \left[\frac{v'_{\varepsilon,i} \varepsilon_{\max}^2}{2\varepsilon_{\max}^2 - 4v'_{\varepsilon,i}} + v_{\varepsilon,\text{pop}} \right]^{-1}. \quad (23)$$

7.1.3. Ellipticity correction

The final step in our recalibration process is the correction of source ellipticities, designed to remove any residual PSF leakage into the distribution of measured shapes. This process is similarly performed in bins of resolution and S/N, but also in bins of photo- z (which are used for cosmic shear tomography). As such, this process formally does not contribute to the sample statistics calculated here, as the choice of tomographic bins for KiDS-Legacy is decided during the cosmological analysis. Nonetheless, we document the process here for posterity. Furthermore, the ellipticity correction does not fully encapsulate

biases inherent to the shape measurement process, as multiplicative shear biases remain and must be corrected for (see Sect. 7.1.4).

The first step in our shape recalibration process is to perform a weighted linear fit (using our recalibrated shape weights) between the distribution of source ellipticities $\varepsilon = (\varepsilon_1, \varepsilon_2)$ and measured PSF ellipticities $\varepsilon_{\text{PSF}} = (\varepsilon_{\text{PSF},1}, \varepsilon_{\text{PSF},2})$, per source i , assuming the model

$$\varepsilon_i = \alpha \varepsilon_{\text{PSF},i} + c + \mathcal{N}(0, \sigma_\varepsilon), \quad (24)$$

where $c = (c_1, c_2)$ are the additive components of the regression models for each ellipticity component, and $\sigma_\varepsilon = (\sigma_{\varepsilon,1}, \sigma_{\varepsilon,2})$ are the standard deviations of the residuals in the fits to each ellipticity component, which again are assumed to be Gaussian. These fits provide us with an estimate of the residual PSF leakage in the individual ellipticity components, $\alpha = (\alpha_1, \alpha_2)$, and an associated fit uncertainty $\sigma_\alpha = (\sigma_{\alpha,1}, \sigma_{\alpha,2})$.

The distribution of α prior to any correction, expressed as the arithmetic mean of the two components $\bar{\alpha} = (\alpha_1 + \alpha_2)/2$, is shown in panel e of Fig. 28. In most bins we observe modest biases: amplitudes are typically $\bar{\alpha} \lesssim 0.2$, and vary smoothly over the plane of resolution and S/N. However, in low S/N cases the

Table 14. Statistics of the additional selections used for KiDS-Legacy sample definition.

Selection name	Formula	N_s^r	f_{tot}^r	$N_{s,\text{run}}^r$	f_{run}^r	$N_{s,\text{run}}^k$	f_{run}^k
MASK	Eq. (29)	–	–	–	–	100 744 685	100.00%
FLAG_GAAP	Eq. (30)	85 192	0.08%	85 192	0.08%	100 659 493	99.92%
ASTEROIDS	Eq. (31)	951 580	0.94%	951 411	0.95%	99 708 082	98.97%
UNMEASURED	Eq. (32)	20 844 833	20.69%	20 196 904	20.26%	79 511 178	78.92%
BLENDING	Eq. (33)	26 486 769	26.29%	5 560 307	6.99%	73 950 871	73.40%
FITCLASS	Eq. (34)	31 545 314	31.31%	10 376 131	14.03%	63 574 740	63.10%
BINARY	Eq. (35)	143 790	0.14%	14 221	0.02%	63 560 519	63.09%
MAG_AUTO	Eq. (37)	2 658 703	2.64%	751 064	1.18%	62 809 455	62.35%
RESOLUTION	Eq. (38)	43 904 820	43.58%	12 658 667	20.15%	50 150 788	49.78%
WEIGHT	Eq. (39)	–	–	6 945 632	13.85%	43 205 156	42.89%
Full legacy sample	–	–	–	–	–	43 205 156	42.89%
$0.1 < z_B \leq 2.0$	–	–	–	478 444	1.11%	42 726 712	42.41%
$0.1 < z_B \leq 1.2$	–	–	–	5 608 698	12.98%	37 596 458	37.32%

Notes. Superscript indices specify whether a number refers to sources that are removed from the sample ('r') or kept in the sample ('k'). N_s^r is the total number of sources (from the full MASK ≤ 1 sample) that are flagged for removal by the selection s . f_{tot}^r is N_s^r expressed as a fraction of the full MASK ≤ 1 sample (N_{MASK}^k). $N_{s,\text{run}}^r$ is the number of sources from the currently un-flagged population that are newly flagged after applying selection s (assuming flags are applied in row-order). f_{run}^r is $N_{s,\text{run}}^r$ expressed as a fraction of the currently un-flagged population. $N_{s,\text{run}}^k$ is the number of sources remaining un-flagged after the flagging of selection s . f_{run}^k is $N_{s,\text{run}}^k$ expressed as a fraction of the full MASK ≤ 1 sample.

biases drop sharply to negative values, resulting in significant within-bin variation. As a result, the direct per-bin correction approach introduced in Sect. 7.1.2 does not satisfactorily correct for all biases within each bin. Instead, we invoked a hybrid approach that first removes a continuous polynomial bias over the \mathcal{R} and S/N plane, followed by a direct bin-wise correction to account for remaining (incoherent) biases.

To correct for the overall trend, we first estimated the amplitude of PSF contamination by fitting a polynomial to the estimated distribution of α (per component) and resolution vs S/N:

$$\alpha = c_0 + c_1 S/N^{-2} + c_2 S/N^{-3} + c_3 \mathcal{R} + c_4 \mathcal{R} S/N^{-2}, \quad (25)$$

where the coefficients $c_x = (c_{x,1}, c_{x,2})$ are fit for using weighted least squares, per component and with inverse variance weights (σ_α^{-2}). The result of this polynomial fit is a contiguously estimated $\hat{\alpha}$ per source that is used to remove the overall PSF leakage into the source ellipticities:

$$\hat{\epsilon} = \epsilon - \hat{\alpha} \epsilon_{\text{PSF}}. \quad (26)$$

In practice $\hat{\epsilon}$ is an intermediate product because, as mentioned previously, the ellipticities after this correction still contain residual biases due to higher-order bin-wise fluctuations in the distribution of α . These residual fluctuations are removed on a per-bin basis, using a direct correction for the residual bias in the ellipticities per bin,

$$\hat{\epsilon} = \alpha_r \epsilon_{\text{PSF}} + c + \mathcal{N}(0, \sigma_{\hat{\epsilon}}), \quad (27)$$

resulting in the final ellipticity measurements used for science:

$$\epsilon' = \hat{\epsilon} - \alpha_r \epsilon_{i,\text{PSF}}. \quad (28)$$

The distribution of measured biases after both corrections is also shown in panel f of Fig. 28.

After this recalibration procedure, sources with non-physical ellipticities (i.e. $|\epsilon_i| > 1$) are removed from the sample; however, this number is typically very small (Li et al. 2023, found

four sources out of more than 26 million were removed by this selection in their mock analyses).

Finally, we include the on-sky distribution of shape measurement statistics in Appendix G. These diagnostics are useful for understanding the variability of source shapes, PSF sizes, and possible systematic effects present in the lensing catalogue.

7.1.4. Multiplicative bias

The ellipticity estimates from *lensfit* suffer from small percent-level multiplicative biases. These biases originate from different physical effects such as source detection, blending, noise, model imperfections, and so on, and strongly depend on galaxy properties such as size, S/N, shape of the light profile, and so on. Ultimately, an estimate of the total multiplicative bias requires the choice of a source sample and the careful simulation of the shape measurement process on that sample. Such simulations are presented in Li et al. (2023) for the KiDS-1000 source sample and the tomographic bins used in the KiDS-1000 cosmic shear analysis (Asgari et al. 2021). We expect similar multiplicative biases for the KiDS-DR5 source sample presented here and refer the reader to Li et al. (2023) for orientation. However, the actual estimation of these biases will differ slightly due to the differences between DR4 and DR5 and possible differences in the upcoming KiDS-Legacy analyses compared to KiDS-1000. Hence, we do not present actual numbers for the multiplicative bias here but defer this to the forthcoming science papers.

7.2. KiDS-Legacy sample selection

The final KiDS-Legacy sample is defined as the lensing sources that satisfy a number of quality criteria. These criteria are summarised here, in Table 14, and in Fig. 29. The table presents details of the selections in terms of the number of sources removed (and remaining) in the KiDS-Legacy sample after each successive selection. Similarly, Fig. 29 presents the fraction of sources that are retained in the sample as a function of r -band magnitude after each selection.

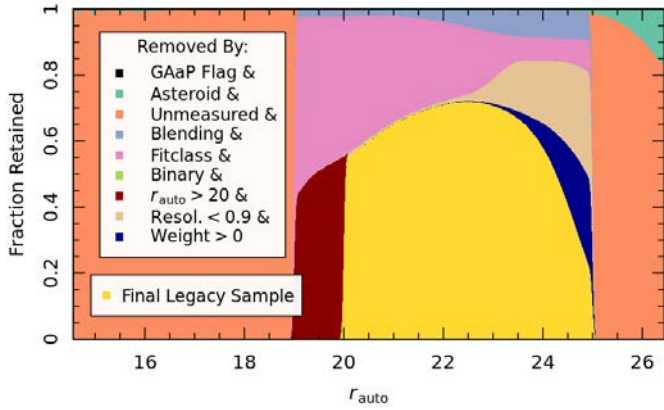


Fig. 29. Selections applied to the KiDS DR5 MASK ≤ 1 sample, and how they modify the available number of sources as a function of r -band magnitude, relative to the total number counts. The number of sources that are removed by each selection is given in Table 14.

The initial sample was constructed by applying the fiducial mask selection (Sect. 6.4):

$$\text{MASK} \leq 1. \quad (29)$$

Next, we removed all sources that do not have successful GAAP measurements in all ten photometric bands $x \in ugri_1i_2ZYJHK_s$:

$$N_{\text{flag}} = \sum_{\mathbf{x}} \text{Flag_GAAP_}\mathbf{x} == 0. \quad (30)$$

This selection removes less than 0.1% of the total MASK ≤ 1 sample.

We then removed asteroids and other transients present in the detection imaging, using a combination of extreme colour selections (Kuijken et al. 2015):

$$(g - r \leq 1.5) \text{ or } (i_1 - r \leq 1.5) \text{ or } (i_2 - r \leq 1.5). \quad (31)$$

The asteroid mask removes 0.94% of sources from the MASK ≤ 1 sample, consisting of both bright objects and a portion of the faintest sources in the sample (due to high photometric noise).

As the KiDS-Legacy sample is primarily intended for lensing science, we next selected on lensing-related quantities. First, we removed sources for which *lensfit* does not produce a shape estimate. Such sources retain a placeholder-value of exactly zero for the PSF moments (among other parameters), and are therefore able to be removed by selecting only galaxies with

$$(\text{PSF_Q11} \neq 0.0) \text{ and } (\text{PSF_Q22} \neq 0.0). \quad (32)$$

This causes a significant reduction of the sample, removing over 20 million sources. Figure 29 demonstrates, however, that this reduction is caused by a blanket cut of sources in r -band magnitude that are brighter than $r_{\text{auto}} = 19$ or fainter than $r_{\text{auto}} = 25$. These limits are internally imposed by *lensfit*. Within this magnitude window, very few sources are removed by this selection.

The next lensing-related selection removes sources that are considered to be blended. This selection is performed using the *lensfit* `contamination_radius` parameter,

$$\text{contamination_radius} > 4.25 \text{ pix}, \quad (33)$$

and removes sources that are contaminated by extracted neighbours. This selection removed 26.3% of the MASK ≤ 1 sample;

however, most of these sources are already removed by previous selections (particularly the ‘unmeasured’ selection). As such, this blending selection only removes a further 6.99% of the available sample.

As described in Sect. 7.1, the *lensfit* code provides a range of flags related to the quality of the fitted model, called `fitclass`. `fitclass` selection is particularly useful for the rejection of point-like sources from the sample, which might otherwise contaminate our lensing measurements. For the lensing sample, we applied the fiducial `fitclass` selection:

$$\text{fitclass} \notin \{-10, -7, -4, -3, -1, 1, 2\}, \quad (34)$$

where the meanings of the various flags are given in Table 13. This `fitclass` selection removes more than 31% of the MASK ≤ 1 sample from consideration; however, more than half of these sources are already removed by previous flags. The remaining 14% of sources that are removed are preferentially bright, as can be seen from Fig. 29. This is consistent with the majority of these sources being stellar contaminants; a conclusion that is supported by the `fitclass` selection producing r -band number counts that more consistent with a simple power-law.

Next we applied a further selection to remove unresolved binary stars from the sample. This binary star selection has also been updated from that used in DR4. Figure 30 shows the distribution of objects from the MASK ≤ 1 sample in r -band magnitude and (a parameter akin to) angular extent on-sky. The centre and right panels of the figure show the sample prior to binary rejection, split by ellipticity. The highly elliptical sample ($|\epsilon| > 0.8$) is used for the identification of unresolved binaries, using a cut in the size-brightness diagram, quantified using the r -band scale-length measured by *lensfit* ($R_s = \text{autocal_scalelength_pixels}$), and the r -band magnitude measured by GAAP. In DR4, the binary selection was made using a simple linear cut in this plane (the dashed red line). For DR5, we revised the cut to now perform a slightly modified linear rejection, coupled with a brightness cut:

$$\log_{10}(2R_s) \geq (24.9 - r_{\text{GAAP}})/3.5 \quad (35)$$

and

$$r_{\text{GAAP}} > 21.5. \quad (36)$$

This selection removes only 0.14% of sources in total, and only 0.02% of sources that would otherwise be currently in the lensing sample. Nonetheless, visual inspection of the $r-i$ versus $g-r$ colour-colour-space demonstrates that these selected sources are well localised on the stellar locus.

Next we applied a blanket magnitude in the detection band:

$$\text{MAG_AUTO} > 20.0. \quad (37)$$

This selection causes a 2.64% reduction in the total source counts, and a 1.18% reduction in the available sample for lensing. The origin of this selection lies in the previous *lensfit* version, which did not model the shapes for sources brighter than $r_{\text{auto}} = 20$. The fraction of sources lost to this selection is relatively small (1.18%), and therefore we opted to keep this selection to be consistent with previous releases.

Next we applied a resolution selection that is designed to improve weight and shape recalibration (Sect. 7.1.2):

$$\mathcal{R} < 0.9, \quad (38)$$

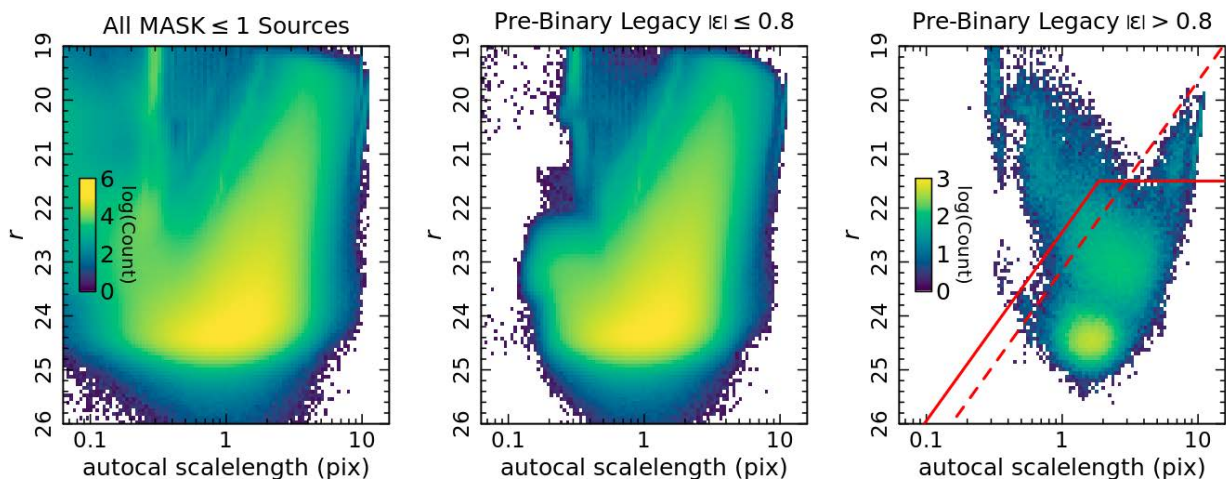


Fig. 30. Selection of binary stars in construction of the KiDS-Legacy lensing sample. Left: distribution of all DR5 sources in size and apparent magnitude. Centre: distribution of all sources remaining prior to the rejection of binary stars, which have ellipticities $|e| \leq 0.8$. The colour scale in the centre panel is the same as in the left panel. Right: as in the centre panel, but for sources with ellipticities $|e| > 0.8$. This is the space in which the binary rejection is performed. The binary rejection criteria used in previous KiDS releases is shown as the dashed red line. The new binary rejection criteria for sources in DR5 is shown as the solid red line (sources above the line are discarded).

where \mathcal{R} is defined in Eq. (20). This selection is a significant one, and removes 43.58% of the total $\text{MASK} \leq 1$ sample. Of sources otherwise in the lensing sample, this selection removes 20.15% of the sources. However, these sources are (by definition) the least resolved sources in KiDS, which means that they also have the lowest shear-responses and thus the lowest lensing weight. The sources are also preferentially faint, as can be seen from Fig. 29. Therefore, while this selection seems pathological (causing a $>20\%$ reduction in source number), Li et al. (2023) found (for DR4) that the cut only reduced the effective number density of lensing sources (i.e. including lensing weights) by $\sim 2\%$. In KiDS-Legacy the selection has a slightly larger effect, reducing the effective number density by 4.6% (see Table H.1).

Finally, we applied a selection on whether sources have a non-zero shape measurement weight:

$$\text{AlphaRecalC_weight} > 0. \quad (39)$$

This selection is highly correlated with the resolution selection, but can only be applied after the resolution selection has been applied (because the resolution selection is invoked in the weight recalibration, Sect. 7.1.2). Similarly, because of the weight recalibration, the weight selection can only be applied to the subset of the data that also satisfy the resolution selection, and so it is not meaningful to quote the fraction of the $\text{MASK} \leq 1$ dataset that satisfy the weight requirement. Furthermore, as can be seen in Fig. 29, these are primarily faint sources. Of the currently available dataset, the weight selection removes the 13.85% of the sources. We note, though, that the removal of these sources has no influence on the effective number density of lensing sources, because they carry zero lensing weight (see Appendix H).

The final KiDS-Legacy sample therefore consists of 43 205 156 sources drawn from 1014.0 deg^2 of sky. Using the effective number density formulation of Heymans et al. (2012), the sample has an effective number density for lensing of $n_{\text{eff}} = 8.92 \text{ arcmin}^{-2}$. Limiting this sample to the photo- z range expected for the KiDS-Legacy cosmological analyses ($0.1 < z_B \leq 2.0$) reduces the sample size to effective number density only slightly, to $n_{\text{eff}} = 8.82 \text{ arcmin}^{-2}$. This is the fiducial effective number density for the KiDS-Legacy sample, and is what we show for KiDS in Fig. 1. Finally, for direct comparison with

previous KiDS releases, the effective number density of sources in the photo- z range $0.1 < z_B \leq 1.2$ is $n_{\text{eff}} = 8.00 \text{ arcmin}^{-2}$. This represents a modest increase from the measured effective number density in KiDS DR4, which was $n_{\text{eff}} = 7.66 \text{ arcmin}^{-2}$ (Hildebrandt et al. 2021) in the same photo- z window. We attribute this increase in effective number primarily to the application of the strong-selection mask (Sect. 6.4), which masks areas of the survey with systematically reduced galaxy number densities.

8. Data release and catalogues

The public release of KiDS data has been made through the ESO archive. As with previous releases, the data products made available via the archive include optical imaging (include science, weight, flag, and sum images), single-band detection catalogues, and multi-band catalogues (including NIR information from VIKING). However, changes to the data that are available within the survey (specifically regarding the two i -band passes) require changes to the format of the data products that are stored at ESO. Columns that are contained within the multi-band catalogues are presented in Table I.1, as well as specific columns that have changed names and/or meanings (Table I.3).

9. Summary

In this paper we present the fifth and final data release of KiDS. The data release presents optical and NIR photometry over 1370 square degrees of sky: 1347 square degrees of wide-field imaging designed for use in weak lensing studies, and 27 square degrees covering deep spectroscopic survey fields (with 4 deg^2 overlap). The release improves upon previous KiDS data releases in more ways than simply the 34% increase in survey area: multi-epoch photometry in the i -band allows for improved photo- z and temporal science; improved calibration samples from the larger spectroscopic survey overlap allow for a better quantification of survey performance and systematic effects; the improved calibration of photometry and astrometry produces higher-quality imaging and derived data products, improved masking reduces unnecessary loss of data (and the remaining data are of higher

quality); and the updated shape measurement (and calibration) leads to reductions in systematic biases in downstream data products.

The data forming this release include images (science, weight, flag, sum, and mask frames), independent catalogues (single-band source extractions per tile), multi-band catalogues (forced photometry and photo- z catalogues per tile), and mosaic catalogues (the KiDS-Legacy catalogue for weak-lensing science and the KiDZ catalogue for calibration efforts). These data are made publicly available via the ESO archive and the KiDS collaboration web pages.

The KiDS-Legacy sample consists of 42 974 574 sources drawn from 1014.0 deg² of sky, with an effective number density of $n_{\text{eff}} = 8.87 \text{ arcmin}^{-2}$. After limiting the sample to the expected tomographic limits for KiDS-Legacy cosmological analyses ($0.1 < z_B \leq 2.0$), the effective number density decreases slightly to $n_{\text{eff}} = 8.77 \text{ arcmin}^{-2}$. The KiDZ sample consists of 126 085 sources extracted from KiDS-depth imaging, with spectroscopic and photometric redshift estimates. This dataset represents a significant increase in the calibration sample available for future cosmological analyses with KiDS.

This paper is intended as a one-stop reference for current and future members of the community wishing to utilise KiDS data for their science. It is the sincere hope of the authors that this release will cement the legacy value of this unique dataset for years to come.

Acknowledgements. We thank the anonymous referee for their comments, which have undoubtedly improved the quality of the manuscript. A.H.W., H.H., A.D., C.M., R.R., and J.L.v.d.B. are supported by an European Research Council Consolidator Grant (No. 770935). A.H.W. and H.H. acknowledge funding from the German Science Foundation DFG, via the Collaborative Research Center SFB1491 “Cosmic Interacting Matters - From Source to Signal”. A.H.W. is supported by the Deutsches Zentrum für Luft- und Raumfahrt (DLR), made possible by the Bundesministerium für Wirtschaft und Klimaschutz. H.H. is also supported by a Heisenberg grant of the Deutsche Forschungsgemeinschaft (Hi 1495/5-1). K.K. acknowledges support from the Royal Society and Imperial College. M.R. acknowledges financial support from the INAF mini-grant 2022 “GALCLOCK”. M.B., P.J., and G.K. are supported by the Polish National Science Center through grant no. 2020/38/E/ST9/00395. M.B. is also supported by the Polish National Science Center through grant no. 2018/30/E/ST9/00698, 2018/31/G/ST9/03388 and 2020/39/B/ST9/03494, and by the Polish Ministry of Science and Higher Education through grant DIR/WK/2018/12. C.H. acknowledges support from the European Research Council under grant number 647112, and the UK Science and Technology Facilities Council (STFC) under grant ST/V000594/1. C.H., B.S., Z.Y., and M.Y. acknowledge support from the Max Planck Society and the Alexander von Humboldt Foundation in the framework of the Max Planck-Humboldt Research Award endowed by the Federal Ministry of Education and Research. H.H. acknowledges support from Vici grant 639.043.512, financed by the Netherlands Organisation for Scientific Research (NWO). S.S.L. is supported by NOVA, the Netherlands Research School for Astronomy. L.M. acknowledges support from the UK STFC under grant ST/N000919/1. M.B. acknowledges the INAF PRIN-SKA 2017 program I.05.01.88.04 and the funding from MIUR Premiale 2016: MITIC. P.B. acknowledges support from the German Academic Scholarship Foundation. G.C. acknowledges the support from the grant ASI n.2018-23-HH.0. J.T.A.d.J. is supported by the NWO through grant 621.016.402. B.G. acknowledges the support of the Royal Society through an Enhancement Award (RGF/EA/181006) and the Royal Society of Edinburgh for support through the Saltire Early Career Fellowship (ref. number 1914). C.G. thanks the support from INAF theory Grant 2022: Illuminating Dark Matter using Weak Lensing by Cluster Satellites. J.H.D. acknowledges support from an STFC Ernest Rutherford Fellowship (project reference ST/S004858/1). B.J. acknowledges support by STFC Consolidated Grant ST/V000780/1. L.L. has been supported by the Deutsche Forschungsgemeinschaft through the project SCHN 342/15-1 and DFG SCHN 342/13. L.M. acknowledges support from the grants PRIN-MIUR 2017 WSCC32 and ASI n.2018-23-HH.0. S.J.N. is supported by the Polish National Science Center through grant UMO-2018/31/N/ST9/03975. L.P. acknowledges support from the DLR grant 50QE2002. C.S. acknowledges support from the Agencia Nacional de Investigación y Desarrollo (ANID) through FONDECYT grant no. 11191125 and BASAL project FB210003. H.Y.S. acknowledges the support from CMS-CSST-2021-A01 and CMS-CSST-2021-B01, NSFC of China under grant

11973070, and Key Research Program of Frontier Sciences, CAS, Grant No. ZDBS-LY-7013. T.T. acknowledges funding from the Swiss National Science Foundation under the Ambizione project PZ00P2_193352. G.V.K. acknowledges financial support from the Netherlands Research School for Astronomy (NOVA) and Target. Target is supported by Samenwerkingsverband Noord Nederland, European fund for regional development, Dutch Ministry of economic affairs, Pieken in de Delta, Provinces of Groningen and Drenthe. J.Y. acknowledges the support of the National Science Foundation of China (12203084), the China Postdoctoral Science Foundation (2021T140451), and the Shanghai Post-doctoral Excellence Program (2021419). Y.H.Z. acknowledges support from the UK STFC. Based on data obtained from the ESO Science Archive Facility with DOI: <https://doi.org/10.18727/archive/37>, and <https://doi.eso.org/10.18727/archive/59> and on data products produced by the KiDS consortium. The KiDS production team acknowledges support from: Deutsche Forschungsgemeinschaft, ERC, NOVA and NWO-M grants; Target; the University of Padova, and the University Federico II (Naples). This work was performed in part at Aspen Center for Physics, which is supported by National Science Foundation grant PHY-1607611. Author Contributions: All authors contributed to the development and writing of this paper. The authorship list is given in three groups: the lead authors (A.H.W., K.K., H.H., M.R., and M.B.), followed by two alphabetical groups. The first alphabetical group includes those who are key contributors to both the scientific analysis and the data products of this manuscript and release. The second group covers those who have either made a significant contribution to the preparation of data products or to the scientific analyses of KiDS since its inception.

References

- Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, *ApJS*, **235**, 42
 Ahumada, R., Allende Prieto, C., Almeida, A., et al. 2020, *ApJS*, **249**, 3
 Aihara, H., Arimoto, N., Armstrong, R., et al. 2018, *PASJ*, **70**, S4
 Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, [arXiv:astro-ph/0609591]
 Andrews, S. K., Driver, S. P., Davies, L. J. M., et al. 2017, *MNRAS*, **464**, 1569
 Asgari, M., Lin, C.-A., Joachimi, B., et al. 2021, *A&A*, **645**, A104
 Balestra, I., Mainieri, V., Popesso, P., et al. 2010, *A&A*, **512**, A12
 Benítez, N. 2000, *ApJ*, **536**, 571
 Bertin, E. 2010, Astrophysics Source Code Library [record ascl:1010.068]
 Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, 393
 Bilicki, M., Dvornik, A., Hoekstra, H., et al. 2021, *A&A*, **653**, A82
 Blake, C., Amon, A., Childress, M., et al. 2016, *MNRAS*, **462**, 4240
 Capaccioli, M., Mancini, D., & Sedmak, G. 2005, *The Messenger*, **120**, 10
 Colless, M., Dalton, G., Maddox, S., et al. 2001, *MNRAS*, **328**, 1039
 Comparat, J., Richard, J., Kneib, J.-P., et al. 2015, *A&A*, **575**, A40
 Cooper, M. C., Griffith, R. L., Newman, J. A., et al. 2012, *MNRAS*, **419**, 3018
 Dalton, G. B., Caldwell, M., Ward, A. K., et al. 2006, *SPIE Conf. Ser.*, **6269**, 62690X
 Damjanov, I., Zahid, H. J., Geller, M. J., Fabricant, D. G., & Hwang, H. S. 2018, *ApJS*, **234**, 21
 Davies, L. J. M., Driver, S. P., Robotham, A. S. G., et al. 2015, *MNRAS*, **447**, 1014
 Davies, L. J. M., Robotham, A. S. G., Driver, S. P., et al. 2018, *MNRAS*, **480**, 768
 de Jong, J. T. A., Kuijken, K., Applegate, D., et al. 2013, *The Messenger*, **154**, 44
 de Jong, J. T. A., Verdoes Kleijn, G. A., Boxhoorn, D. R., et al. 2015, *A&A*, **582**, A62
 de Jong, J. T. A., Verdoes Kleijn, G. A., Erben, T., et al. 2017, *A&A*, **604**, A134
 Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, *MNRAS*, **413**, 971
 Driver, S. P., Wright, A. H., Andrews, S. K., et al. 2016, *MNRAS*, **455**, 3911
 Driver, S. P., Bellstedt, S., Robotham, A. S. G., et al. 2022, *MNRAS*, **513**, 439
 Edge, A., Sutherland, W., Kuijken, K., et al. 2013, *The Messenger*, **154**, 32
 Eifler, T., Miyatake, H., Krause, E., et al. 2021, *MNRAS*, **507**, 1746
 Euclid Collaboration (Guglielmo, V., et al.) 2020, *A&A*, **642**, A192
 Euclid Collaboration (Stanford, S. A., et al.) 2021, *ApJS*, **256**, 9
 Euclid Collaboration (Scaramella, R., et al.) 2022, *A&A*, **662**, A112
 Gaia Collaboration (Brown, A. G. A., et al.) 2018, *A&A*, **616**, A1
 Gaia Collaboration (Brown, A. G. A., et al.) 2021, *A&A*, **649**, A1
 Garilli, B., McLure, R., Pentericci, L., et al. 2021, *A&A*, **647**, A150
 Gong, Y., Liu, X., Cao, Y., et al. 2019, *ApJ*, **883**, 203
 González-Fernández, C., Hodgkin, S. T., Irwin, M. J., et al. 2018, *MNRAS*, **474**, 5459
 Guinot, A., Kilbinger, M., Farrens, S., et al. 2022, *A&A*, **666**, A162
 Hasinger, G., Capak, P., Salvato, M., et al. 2018, *ApJ*, **858**, 77
 Heydenreich, S., Schneider, P., Hildebrandt, H., et al. 2020, *A&A*, **634**, A104
 Heymans, C., Van Waerbeke, L., Miller, L., et al. 2012, *MNRAS*, **427**, 146
 Hikage, C., Ouri, M., Hamana, T., et al. 2019, *PASJ*, **71**, 43
 Hildebrandt, H., Viola, M., Heymans, C., et al. 2017, *MNRAS*, **465**, 1454

- Hildebrandt, H., Köhlinger, F., van den Busch, J. L., et al. 2020, *A&A*, **633**, A69
- Hildebrandt, H., van den Busch, J. L., Wright, A. H., et al. 2021, *A&A*, **647**, A124
- Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, *A&A*, **457**, 841
- Ivezić, Ž., Lupton, R. H., Schlegel, D., et al. 2004, *Astron. Nachr.*, **325**, 583
- Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, **873**, 111
- Jarvis, M. J., Bonfield, D. G., Bruce, V. A., et al. 2013, *MNRAS*, **428**, 1281
- Jee, M. J., Tyson, J. A., Hilbert, S., et al. 2016, *ApJ*, **824**, 77
- Kafle, P. R., Robotham, A. S. G., Driver, S. P., et al. 2018, *MNRAS*, **479**, 3746
- Kuijken, K. 2006, *A&A*, **456**, 827
- Kuijken, K. 2008, *A&A*, **482**, 1053
- Kuijken, K. 2011, *The Messenger*, **146**, 8
- Kuijken, K., Heymans, C., Hildebrandt, H., et al. 2015, *MNRAS*, **454**, 3500
- Kuijken, K., Heymans, C., Dvornik, A., et al. 2019, *A&A*, **625**, A2
- Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv e-prints [arXiv:1110.3193]
- Le Fèvre, O., Vettolani, G., Garilli, B., et al. 2005, *A&A*, **439**, 845
- Le Fèvre, O., Cassata, P., Cucciati, O., et al. 2013, *A&A*, **559**, A14
- Le Fèvre, O., Tasca, L. A. M., Cassata, P., et al. 2015, *A&A*, **576**, A79
- Lewis, J. R., Irwin, M., & Bunclark, P. 2010, *ASP Conf. Ser.*, **434**, 91
- Li, R., Napolitano, N. R., Feng, H., et al. 2022, *A&A*, **666**, A85
- Li, S.-S., Kuijken, K., Hoekstra, H., et al. 2023, *A&A*, **670**, A100
- Liang, S., & von der Linden, A. 2023, *MNRAS*, **519**, 2281
- Lidman, C., Tucker, B. E., Davis, T. M., et al. 2020, *MNRAS*, **496**, 19
- Lilly, S. J., Le Brun, V., Maier, C., et al. 2009, *ApJS*, **184**, 218
- Masters, D. C., Stern, D. K., Cohen, J. G., et al. 2017, *ApJ*, **841**, 111
- Masters, D. C., Stern, D. K., Cohen, J. G., et al. 2019, *ApJ*, **877**, 81
- McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, *A&A*, **544**, A156
- McFarland, J. P., Verdoes-Kleijn, G., Sikkema, G., et al. 2013, *Exp. Astron.*, **35**, 45
- Miller, L., Heymans, C., Kitching, T. D., et al. 2013, *MNRAS*, **429**, 2858
- Nakoneczny, S. J., Bilicki, M., Pollo, A., et al. 2021, *A&A*, **649**, A81
- Newman, J. A., Cooper, M. C., Davis, M., et al. 2013, *ApJS*, **208**, 5
- Parkinson, D., Riemer-Sørensen, S., Blake, C., et al. 2012, *Phys. Rev. D*, **86**, 103518
- Pickles, A. J. 1998, *PASP*, **110**, 863
- Popesso, P., Dickinson, M., Nonino, M., et al. 2009, *A&A*, **494**, 443
- Raichoor, A., Mei, S., Erben, T., et al. 2014, *ApJ*, **797**, 102
- Refregier, A. 2003, *MNRAS*, **338**, 35
- Riello, M., De Angeli, F., Evans, D. W., et al. 2021, *A&A*, **649**, A3
- Scodeggio, M., Guzzo, L., Garilli, B., et al. 2018, *A&A*, **609**, A84
- Scoville, N., Aussel, H., Brusa, M., et al. 2007, *ApJS*, **172**, 1
- Sevilla-Noarbe, I., Bechtol, K., Carrasco Kind, M., et al. 2021, *ApJS*, **254**, 24
- Silverman, J. D., Kashino, D., Sanders, D., et al. 2015, *ApJS*, **220**, 12
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, *AJ*, **131**, 1163
- The LSST Dark Energy Science Collaboration (Mandelbaum, R., et al.) 2018, arXiv e-prints [arXiv:1809.01669]
- Trump, J. R., Impey, C. D., Elvis, M., et al. 2009, *ApJ*, **696**, 1195
- Vakili, M., Hoekstra, H., Bilicki, M., et al. 2023, *A&A*, **675**, A202
- van den Busch, J. L., Hildebrandt, H., Wright, A. H., et al. 2020, *A&A*, **642**, A200
- van den Busch, J. L., Wright, A. H., Hildebrandt, H., et al. 2022, *A&A*, **664**, A170
- van der Wel, A., Noeske, K., Bezanson, R., et al. 2016, *ApJS*, **223**, 29
- Venemans, B. P., Verdoes Kleijn, G. A., Mwebaze, J., et al. 2015, *MNRAS*, **453**, 2259
- Wright, A. H., Robotham, A. S. G., Bourne, N., et al. 2016, *MNRAS*, **460**, 765
- Wright, A. H., Driver, S. P., & Robotham, A. S. G. 2018, *MNRAS*, **480**, 3491
- Wright, A. H., Hildebrandt, H., Kuijken, K., et al. 2019, *A&A*, **632**, A34
- York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, *AJ*, **120**, 1579
- ⁴ Center for Theoretical Physics, Polish Academy of Sciences, al. Lotników 32/46, 02-668 Warsaw, Poland
- ⁵ Istituto Nazionale di Astrofisica (INAF) - Osservatorio Astronomico di Capodimonte, Via Moiariello 16, 80131 Napoli, Italy
- ⁶ Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ, UK
- ⁷ Dept of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK
- ⁸ School of Physics and Astronomy, Sun Yat-sen University, Guangzhou 519082, Zhuhai Campus, PR China
- ⁹ E.A Milne Centre, University of Hull, Cottingham Road, Hull, HU6 7RX, UK
- ¹⁰ Centre of Excellence for Data Science, AI, and Modelling (DAIM), University of Hull, Cottingham Road, Kingston-upon-Hull, HU6 7RX, UK
- ¹¹ Department of Physics E. Pacini, University of Napoli Federico II, Via Cintia 6, 80126 Napoli, Italy
- ¹² Institut für Astrophysik, Türkenschanzstrasse 17, 1180 Wien, Austria
- ¹³ Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany
- ¹⁴ Dipartimento di Fisica e Astronomia “Augusto Righi”, Alma Mater Studiorum Università di Bologna, Via Gobetti 93/2, 40129 Bologna, Italy
- ¹⁵ Istituto Nazionale di Astrofisica (INAF) - Osservatorio di Astrofisica e Scienza dello Spazio (OAS) di Bologna, via Piero Gobetti 93/3, 40129 Bologna, Italy
- ¹⁶ Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands
- ¹⁷ Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK
- ¹⁸ Instituto de Ciencias del Cosmos (ICC), Universidad de Barcelona, Martí i Franqués, 1, 08028 Barcelona, Spain
- ¹⁹ Istituto Nazionale di Astrofisica (INAF) - Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
- ²⁰ School of Mathematics, Statistics and Physics, Newcastle University, Herschel Building, NE1 7RU, Newcastle-upon-Tyne, UK
- ²¹ Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
- ²² Waterloo Centre for Astrophysics, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada
- ²³ Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada
- ²⁴ Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA
- ²⁵ Universität Innsbruck, Institut für Astro- und Teilchenphysik, Technikerstr. 25/8, 6020 Innsbruck, Austria
- ²⁶ Institute of Theoretical Astrophysics, Heidelberg University, Philosophenweg 12, 69120 Heidelberg, Germany
- ²⁷ Istituto Nazionale di Fisica Nucleare (INFN) - Sezione di Bologna, viale Berti Pichat 6/2, 40127 Bologna, Italy
- ²⁸ Division of Physics, Mathematics and Astronomy, California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125, USA
- ²⁹ Department of Astrophysics, National Centre for Nuclear Research, Pasteura 7, 02-093 Warsaw, Poland
- ³⁰ Shanghai Astronomical Observatory (SHAO), Nandan Road 80, Shanghai 200030, PR China
- ³¹ University of Chinese Academy of Sciences, Beijing 100049, PR China
- ³² Instituto de Física, Pontificia Universidad Católica de Valparaíso, Casilla 4059, Valparaíso, Chile
- ³³ Institute for Particle Physics and Astrophysics, ETH Zürich, Wolfgang-Pauli-Strasse 27, 8093 Zürich, Switzerland

¹ Ruhr University Bochum, Faculty of Physics and Astronomy, Astronomical Institute (AIRUB), German Centre for Cosmological Lensing, 44780 Bochum, Germany
e-mail: awright@astro.rub.de

² Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands

³ Istituto Nazionale di Astrofisica (INAF) - Osservatorio Astronomico di Padova, via dell’Osservatorio 5, 35122 Padova, Italy

Appendix A: Magnitude limit estimates with

LAMB-DAR

Background estimates in KiDS+KiDZ were made using LAMB-DAR (Wright et al. 2016). LAMB-DAR provides a convenient tool for estimating the correlated and uncorrelated noise present in the various sets of imaging that are present in KiDS-DR5, as it can be ran on images of arbitrary pixel scale. This allows us to simply specify the on-sky locations that we wish to sample with our sky-estimates, and allow the code to report back the noise estimates at each location on-sky for these positions. For the grid of on-sky sampling positions we chose a simple rectangular grid that sampled each of the $\sim 1 \text{ deg}^2$ pointings with a spacing of 0.02 degrees in both RA and declination. We estimate the background level of imaging data using the two methods built into the LAMB-DAR code: randoms estimation and sky estimation.

The randoms estimate of noise is the fiducial one that we use for quantifying our imaging depth in DR5, as it includes the influence of correlated noise. We summarise the computation of the randoms noise estimation process here, and direct the interested reader to Sect. 3.6 of Wright et al. (2016) for a complete description of the randoms measurement process.

The randoms noise estimate is made by measuring the flux contained within the chosen aperture (i.e. circular with a $2''$ diameter), shifted to multiple random location within the local area of the image under analysis. An example of this is shown in Fig. A.1. After randomly shifting the location of the source aperture, the flux measured in the aperture at this new location is measured and recorded. After many realisations, the distribution of random fluxes therefore encodes the expected flux contained within an aperture of this specific size for a source randomly placed in this area of imaging. This measurement is performed on the image after application of the fiducial mask, as can be seen in Fig. A.1 where a diffraction spike has been masked at the bottom of the image. However, this estimate does not include source-masking: this means that the distribution of measured random fluxes also includes contamination from detectable sources. As such, the expectation and scatter of the random flux distribution is computed with median statistics, to suppress bias caused by outliers in the tails of the flux distribution (shown as stars in the right panel of Fig. A.1).

The second estimate of the background noise is made without considering the on-sky correlation present in the background. We leverage this sky estimation method as a robustness check for our fiducial noise estimation method, and to estimate the implied degree of on-sky correlation present in the KiDS imaging. We briefly describe the sky estimation process here, and direct the interested reader to Sect. 3.5 of Wright et al. (2016) for a complete description of the measurement process.

The sky-estimate routine in LAMB-DAR assumes that the pixel-noise can be computed in annular radii around the input source, without consideration of the relative correlation between adjacent pixels. The sky estimate is made on the masked image (but, again, without source masking in our application). Figure A.2 demonstrates the estimation of the uncorrelated sky, for the same location shown in Fig. A.1. The sky estimate is made on a smaller scale than in the randoms case, and is similarly constructed to be robust to contamination: estimates of the mean/median/RMS sky are made in annuli around the requested location, and annuli with anomalously high/low values are discarded. In the case of Fig. A.2, all bins were acceptable. The resulting mean/median sky estimates, and the associated RMS scatter, are provided for this location. The values of the sky at

the requested locations are therefore tabulated in this method without consideration of pixel-to-pixel noise correlation.

Previous data releases in KiDS have leveraged background estimates similar to the uncorrelated method used by LAMB-DAR. This means that the background estimates (and therefore reported magnitude limits) in previous KiDS releases should be similar to those computed using LAMB-DAR's uncorrelated estimate. The magnitude limits estimated using this method for DR5 are $\{24.26 \pm 0.10, 25.15 \pm 0.12, 25.07 \pm 0.14, 23.66 \pm 0.25, 23.73 \pm 0.3\}$ for the $\{u, g, r, i_1, i_2\}$ -bands, respectively, in very good agreement with the estimates presented in previous releases.

An example of the estimation of these two background levels for a single pointing of KiDS is shown in Fig. A.3. The figure shows the estimate of the correlated sky noise with randoms, the estimate of the uncorrelated noise, and ratio between these two noise estimates. This ratio is an estimate of the degree of correlation present in the imaging at that location on sky. The distribution of uncorrelated sky values can be seen to closely trace the dither pattern of an observation in KiDS: the lowest noise values occur where pixels were observed by all five exposures, and the highest noise values occur in the upper and lower extremes of the image where only one exposure is available per pixel. The correlated noise estimate shows much less of this structure, suggesting that the expected $\sqrt{N_{\text{exp}}}$ reduction in noise is being hampered by other effects. The cause of this is suggested by the third panel, which shows the noise correlation factor:

$$C_n = \sigma_{\text{rand}} / \sigma_{\text{uncorr}}. \quad (\text{A.1})$$

The correlation factor can be seen to increase in regions of the image where there is considerable background contamination from, for example, residual reflections from bright stars.

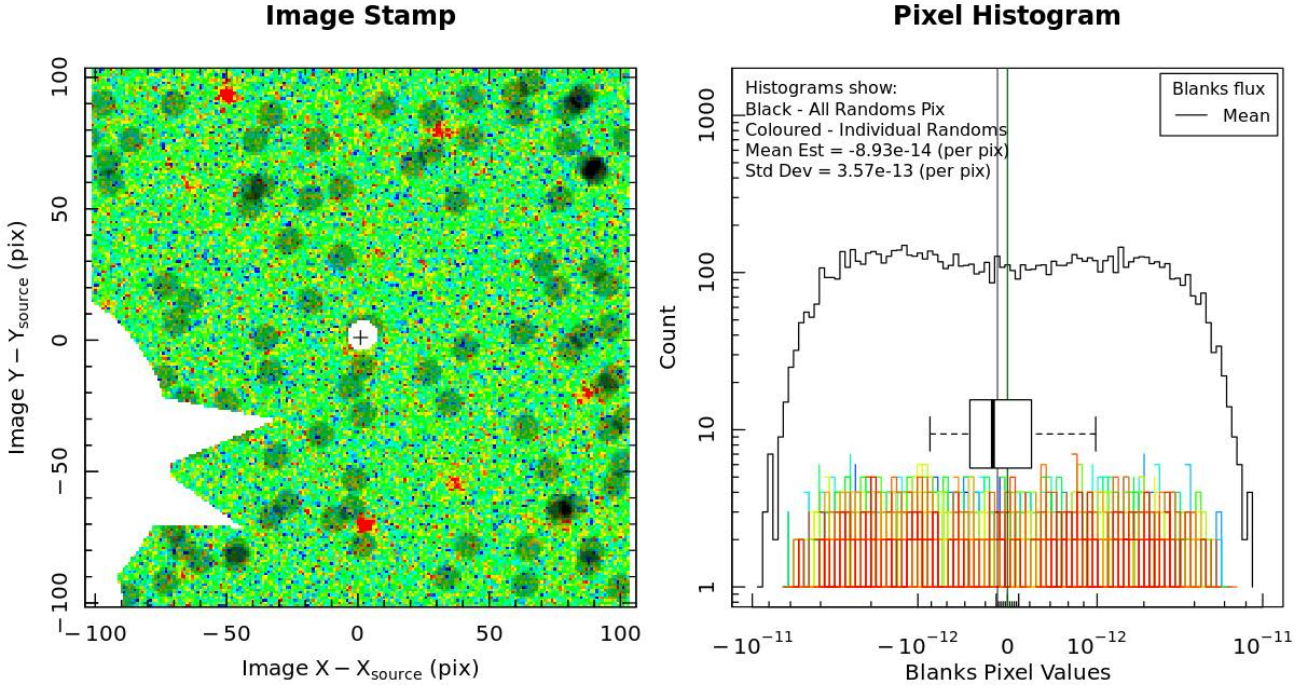


Fig. A.1. Correlated background estimated using the LAMBDAR randoms routine. *Left:* Local region of the image under analysis, with the 100 realisations of the random aperture shown as translucent black circles. The image has been masked using the fiducial mask, seen as the missing triangular region (a diffraction spike) at the bottom of the image. *Right:* Distribution of pixel values measured within the random apertures, for all apertures (black) and the individual apertures (colours). There are some apertures that are coincident with sources, resulting in large positive fluxes. These are limited in the computation of the random noise estimate through the use of median statistics. The distribution of aperture fluxes from the randoms is shown by the box-and-whisker plot, with the outlier fluxes shown as stars. This figure is a direct output of the LAMBDAR code.

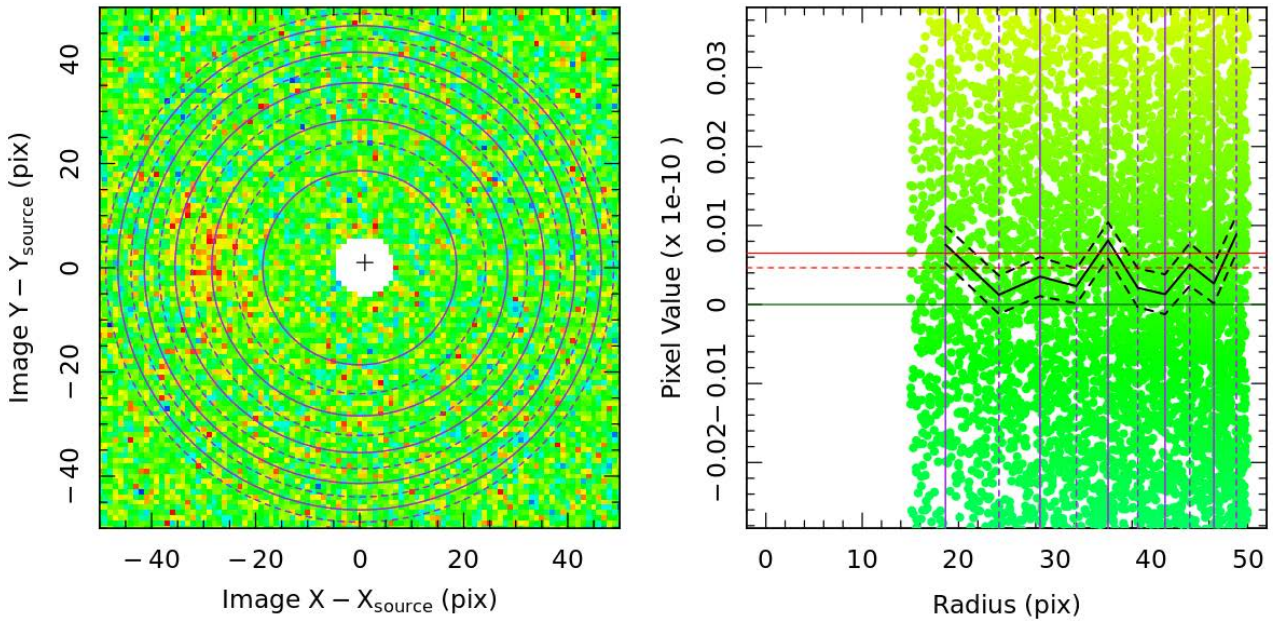


Fig. A.2. Uncorrelated background estimated using the LAMBDAR sky-estimate routine. *Left:* Annuli used to estimate the background estimates, shown against the input image (colour mapping for pixels in the left panel is given by the point colouring in the right panel). *Right:* Distribution of pixel values (points) as a function of radius from the random location chosen. The black lines show the median sky value per annulus (solid) and the uncertainty on the median (dashed). The final mean (median) sky estimate is shown as the solid (dashed) red line. This figure is a direct output of the LAMBDAR code.

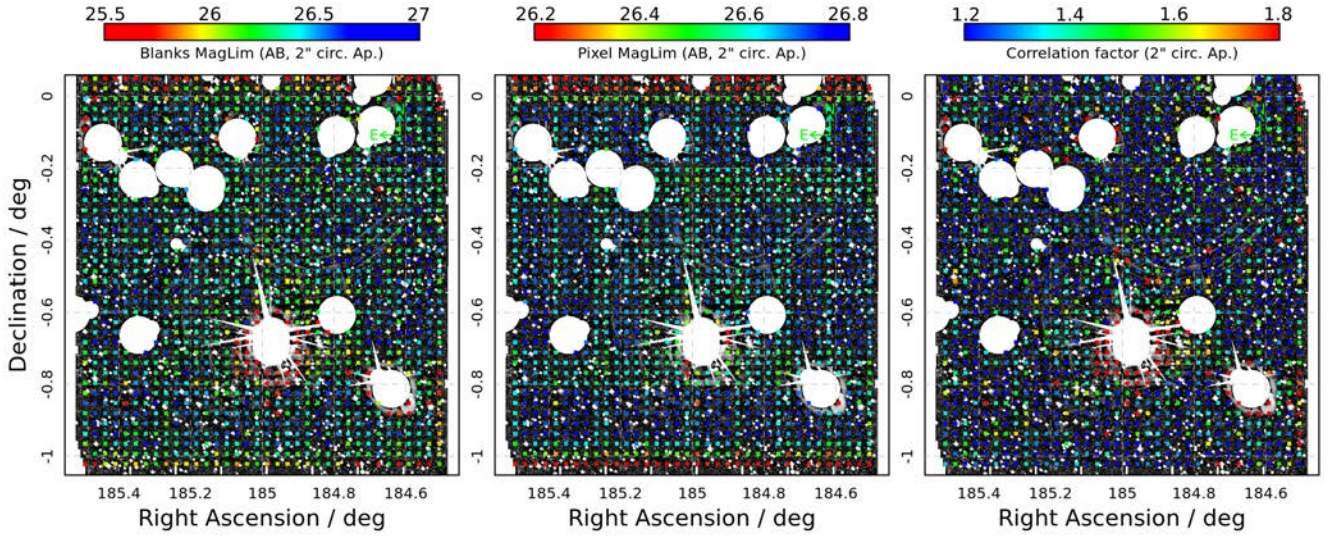


Fig. A.3. Demonstration of the background estimates produced by LAMB DAR for pointing KIDS_185p0_m0p5 in the r -band. Each panel shows a grey-scale image of the pointing, smoothed with a $2''$ Gaussian kernel and after masking. Overlaid on this image are various estimates of the image noise level, shown with rectangles covering the extent of the area used for the estimate. *Left:* Magnitude limits estimated with blanks. *Centre:* Magnitude limits estimated with the pixel RMS. *Right:* Correlation factor, estimated as the ratio between the aperture noise RMSs estimated with blanks and pixels. This factor clearly recovers the regions of the image with increased noise correlation caused by residual bright artefacts.

Appendix B: VST on-sky quality metrics

For many scientific applications, an understanding of the systematic differences in imaging quality on-sky is relevant. For imaging in KiDS the primary parameters of relevance are the magnitude limit (estimated using the method described in Appendix A), and the PSF size in each band. As such, in Fig. B.1 we show the distribution of magnitude limits estimated in each band on-sky, and in Fig. B.2 we show the distribution of average PSF sizes (per pointing, as reported by ASTROWISE). Each figure shows the parameters per filter, with the two i -band passes shown separately.

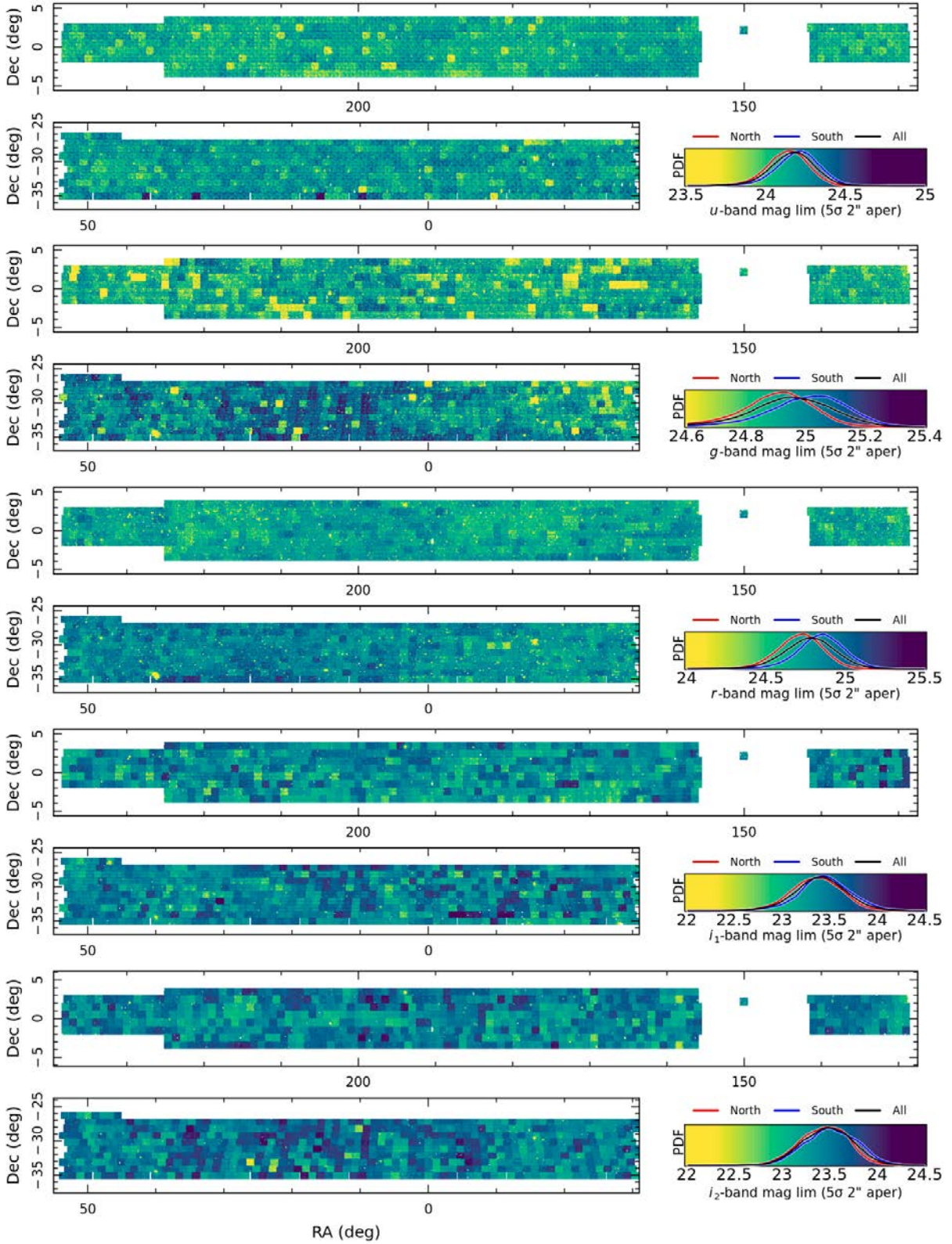


Fig. B.1. Distribution of limiting magnitudes in the optical bands for the KiDS fields. The background levels measured using 2'' circular apertures, as described in Appendix A. To construct the on-sky mosaic, estimates per-pointing are combined using the same WCS cuts as are applied to the data (Sect. 6.5).

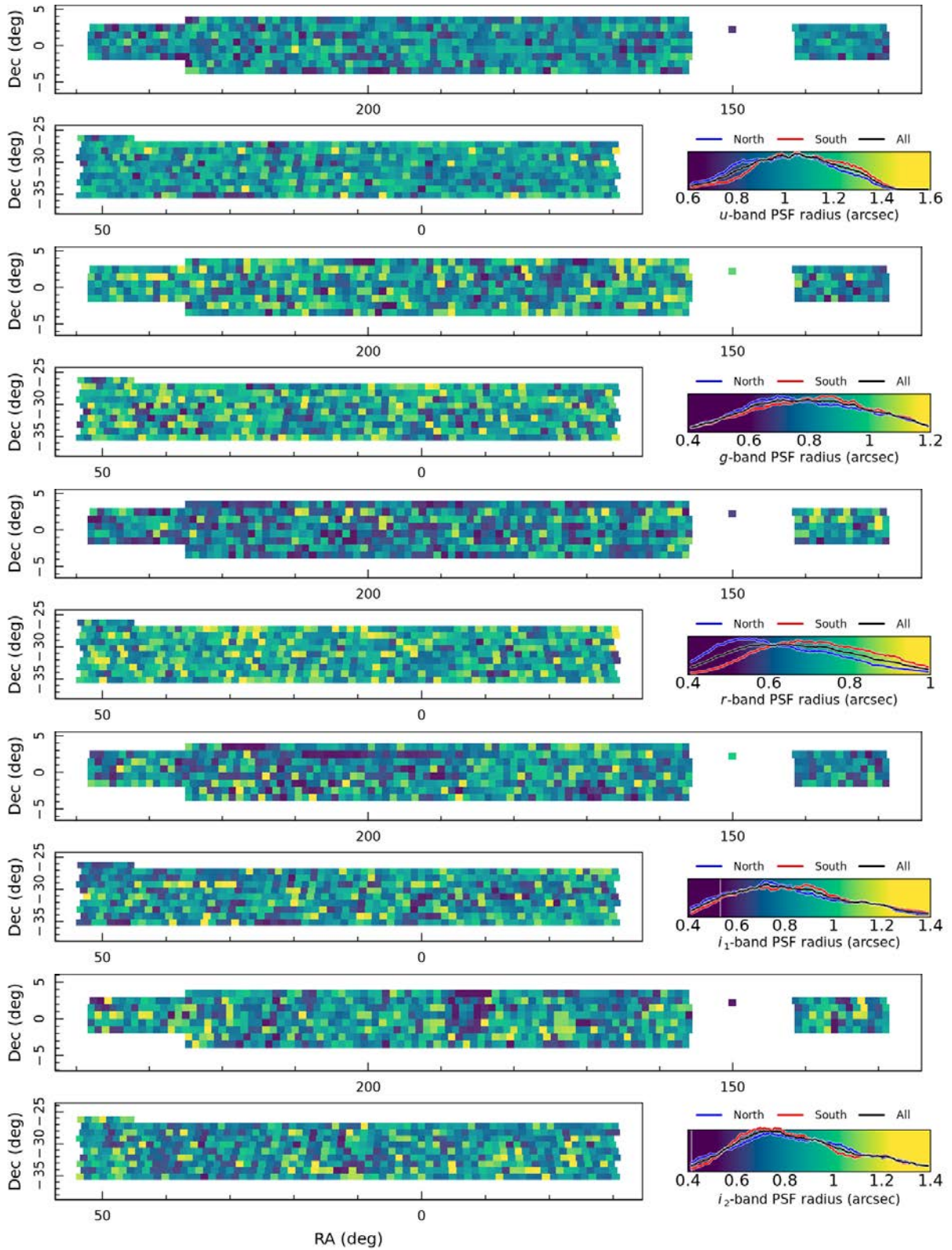


Fig. B.2. Distribution of PSF FWHM sizes (in arcsec) as reported by ASTROWISE for each pointing within KiDS.

Appendix C: Astrometric residuals compared to *Gaia*

As mentioned in Sect. 3.2, the KiDS data are calibrated to *Gaia* stars at the J2000 epoch. In practice, this means that the positions of stars used to calibrate KiDS have been shifted (using their observed positions and estimated proper motions) back to where they would have been in the year 2000. Crucially, though, this does not just have a random effect per star: as the Solar System moves through the Milky Way, an additional coherent shift is imprinted on the stellar distribution. The purpose of calibrating to J2000 epoch stellar positions is to remove this coherent drift.

However, KiDS observations were not taken in the year 2000. Rather, the typical KiDS observation was taken around 2015–2016. As a result, the observed positions of stars in KiDS imaging are actually much more similar to the observed positions of *Gaia* stars at epoch J2016.

The fact that KiDS is calibrated to J2000 epoch stars, but that KiDS images were taken around epoch J2016 (similar to *Gaia*) leads to an apparent systematic bias when performing a direct sky match between the KiDS DR5 catalogues and *Gaia* stars at their epoch J2016 positions: there appears to be a systematic bias in the two hemispheres, at the level of $|\Delta RA| \sim 0''.1$, which has different sign in the northern and the southern hemispheres. This can be seen in Fig. C.1. Of particular note, though, is the low scatter in each of the systematically offset clouds. This is demonstrative that, modulo the difference in overall coordinates between epoch J2000 and epoch J2016 positions, the relative position of stars in KiDS and *Gaia* are very similar (because the observations were actually taken at similar epochs).

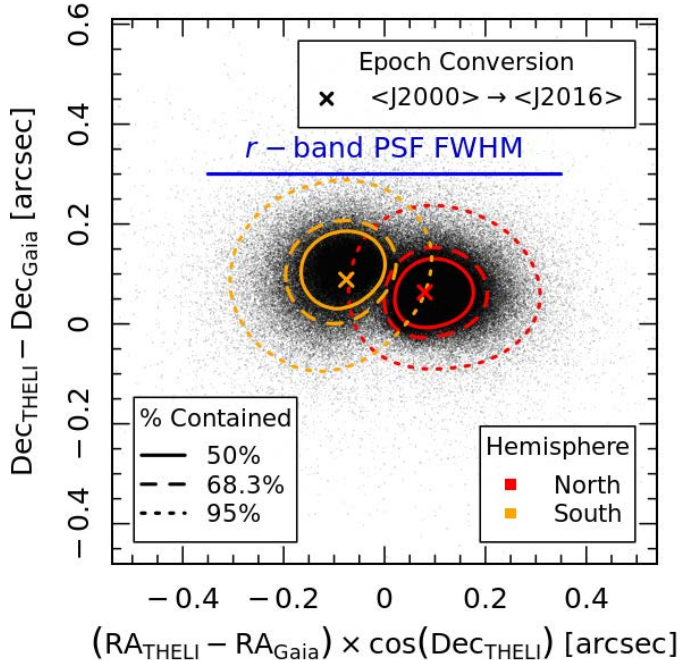


Fig. C.1. Astrometric residuals between KiDS and *Gaia* when comparing to J2016 epoch stellar positions. The systematic difference is caused by the systematic shift in the positions of stars between 2000 and 2016, caused by the motion of the Solar System through the Milky Way.

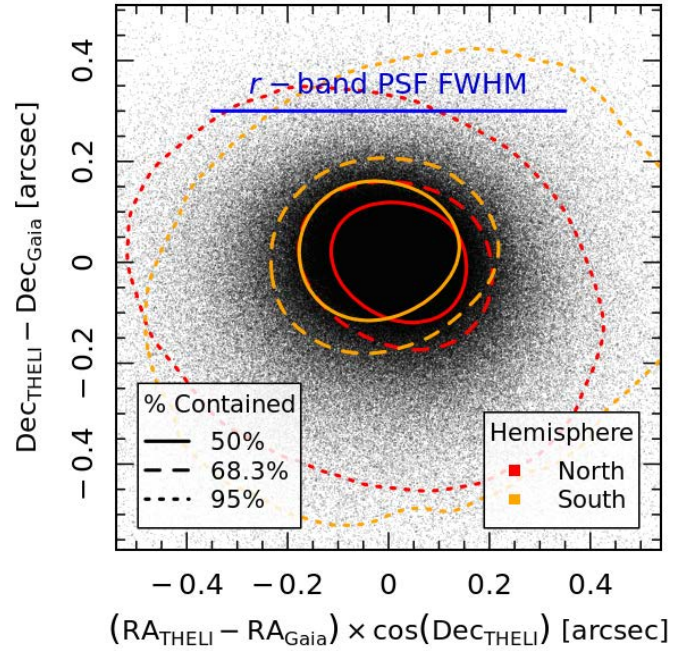


Fig. C.2. Astrometric residuals between KiDS and *Gaia* when comparing to J2000 epoch stellar positions. The systematic difference is no longer present; however, a considerable increase in scatter is visible. This is due to the KiDS stars being relatively positioned according to the J2016 epoch (determined by when the images were taken).

Figure C.2 shows the same as Fig. C.1, but now using J2000 epoch positions. When performing the comparison of per-star astrometry between the KiDS data and the J2000 epoch *Gaia* stars, one expects that the systematic offsets would disappear, which it does. However, we also see a considerable increase in scatter: this is due to the proper motions of stars introducing noise, as they shuffle positions between epoch J2000 and epoch J2016.

Appendix D: List of reprocessed *u*-band co-adds

As discussed in Sect. 3.5.1, *u*-band data were flagged for reprocessing with a modified astrometric distortion polynomial order, to reduce pathological errors in co-add construction. The full list of reprocessed *u*-band co-adds is given in Table D.1.

Table D.1. Co-adds for which the u -band was reprocessed with linear polynomial order for the astrometric distortion definition.

KiDS_0.0_-27.2	KiDS_0.0_-28.2	KiDS_0.0_-30.2	KiDS_0.0_-31.2	KiDS_1.1_-27.2	KiDS_1.2_-30.2
KiDS_10.3_-29.2	KiDS_10.4_-30.2	KiDS_10.6_-32.1	KiDS_11.3_-28.2	KiDS_11.4_-29.2	KiDS_11.5_-30.2
KiDS_11.9_-33.1	KiDS_12.0_-34.1	KiDS_12.4_-28.2	KiDS_12.5_-29.2	KiDS_12.8_-31.2	KiDS_13.2_-34.1
KiDS_13.5_-27.2	KiDS_13.5_-28.2	KiDS_13.7_-29.2	KiDS_14.1_-32.1	KiDS_14.4_-34.1	KiDS_14.6_-35.1
KiDS_14.7_-28.2	KiDS_14.8_-29.2	KiDS_15.0_-30.2	KiDS_15.1_-31.2	KiDS_15.3_-32.1	KiDS_15.4_-33.1
KiDS_15.7_-27.2	KiDS_15.8_-28.2	KiDS_15.8_-35.1	KiDS_158.4_3.5	KiDS_16.3_-31.2	KiDS_16.5_-32.1
KiDS_16.8_-27.2	KiDS_16.8_-34.1	KiDS_16.9_-28.2	KiDS_161.0_0.5	KiDS_163.0_-1.5	KiDS_164.0_-0.5
KiDS_164.0_1.5	KiDS_165.5_3.5	KiDS_167.5_-3.5	KiDS_167.5_2.5	KiDS_168.0_-1.5	KiDS_168.5_3.5
KiDS_17.0_-35.1	KiDS_17.1_-29.2	KiDS_17.3_-30.2	KiDS_17.5_-31.2	KiDS_17.6_-32.1	KiDS_17.8_-33.1
KiDS_170.0_-1.5	KiDS_171.0_0.5	KiDS_171.5_-3.5	KiDS_173.0_-0.5	KiDS_173.5_3.5	KiDS_175.0_-0.5
KiDS_175.0_1.5	KiDS_175.5_2.5	KiDS_177.0_0.5	KiDS_177.5_2.5	KiDS_177.5_3.5	KiDS_178.0_1.5
KiDS_179.0_-1.5	KiDS_179.5_3.5	KiDS_18.0_-34.1	KiDS_18.1_-28.2	KiDS_18.2_-35.1	KiDS_18.4_-30.2
KiDS_18.6_-31.2	KiDS_18.8_-32.1	KiDS_181.0_0.5	KiDS_181.0_1.5	KiDS_181.5_-2.5	KiDS_183.0_-1.5
KiDS_185.0_-0.5	KiDS_186.0_1.5	KiDS_186.5_2.5	KiDS_187.0_0.5	KiDS_19.2_-28.2	KiDS_19.2_-34.1
KiDS_19.4_-29.2	KiDS_19.5_-35.1	KiDS_190.0_-1.5	KiDS_192.5_3.5	KiDS_193.0_-0.5	KiDS_193.0_0.5
KiDS_193.0_1.5	KiDS_194.5_3.5	KiDS_196.0_0.5	KiDS_2.4_-32.1	KiDS_20.0_-32.1	KiDS_20.2_-27.2
KiDS_20.2_-33.1	KiDS_20.3_-28.2	KiDS_20.4_-34.1	KiDS_20.5_-29.2	KiDS_201.6_-3.5	KiDS_204.0_-0.5
KiDS_207.6_2.5	KiDS_21.0_-31.2	KiDS_21.2_-32.1	KiDS_21.3_-27.2	KiDS_21.4_-28.2	KiDS_21.4_-33.1
KiDS_21.6_-29.2	KiDS_21.6_-34.1	KiDS_21.9_-30.2	KiDS_219.6_2.5	KiDS_22.1_-31.2	KiDS_22.4_-32.1
KiDS_22.6_-33.1	KiDS_22.8_-29.2	KiDS_22.8_-34.1	KiDS_23.0_-30.2	KiDS_23.6_-27.2	KiDS_23.7_-28.2
KiDS_23.8_-33.1	KiDS_23.9_-29.2	KiDS_24.5_-31.2	KiDS_24.7_-32.1	KiDS_24.8_-28.2	KiDS_25.0_-33.1
KiDS_25.1_-29.2	KiDS_25.2_-34.1	KiDS_25.3_-30.2	KiDS_25.5_-35.1	KiDS_25.6_-31.2	KiDS_25.8_-27.2
KiDS_25.9_-32.1	KiDS_26.0_-28.2	KiDS_26.1_-33.1	KiDS_26.2_-29.2	KiDS_26.4_-34.1	KiDS_26.8_-31.2
KiDS_27.1_-28.2	KiDS_27.1_-32.1	KiDS_27.3_-33.1	KiDS_27.6_-30.2	KiDS_27.6_-34.1	KiDS_28.0_-27.2
KiDS_28.0_-31.2	KiDS_28.2_-32.1	KiDS_28.5_-29.2	KiDS_28.8_-34.1	KiDS_29.1_-31.2	KiDS_29.2_-27.2
KiDS_29.2_-35.1	KiDS_29.3_-28.2	KiDS_29.4_-32.1	KiDS_29.6_-29.2	KiDS_3.4_-28.2	KiDS_3.5_-32.1
KiDS_3.6_-33.1	KiDS_3.6_-34.1	KiDS_30.0_-34.1	KiDS_30.3_-31.2	KiDS_30.5_-28.2	KiDS_30.6_-32.1
KiDS_30.8_-29.2	KiDS_31.1_-30.2	KiDS_31.4_-27.2	KiDS_31.6_-28.2	KiDS_31.6_-35.1	KiDS_31.8_-32.1
KiDS_31.9_-29.2	KiDS_32.1_-33.1	KiDS_32.2_-30.2	KiDS_32.5_-27.2	KiDS_32.8_-35.1	KiDS_32.9_-32.1
KiDS_33.0_-29.2	KiDS_33.4_-30.2	KiDS_33.8_-31.2	KiDS_33.9_-28.2	KiDS_335.5_-31.2	KiDS_336.5_-32.1
KiDS_34.2_-29.2	KiDS_34.8_-27.2	KiDS_346.3_-29.2	KiDS_346.5_-27.2	KiDS_347.6_-28.2	KiDS_348.7_-28.2
KiDS_349.2_-34.1	KiDS_349.6_-30.2	KiDS_35.0_-31.2	KiDS_35.3_-35.1	KiDS_35.6_-33.1	KiDS_35.9_-27.2
KiDS_350.6_-32.1	KiDS_352.1_-28.2	KiDS_355.3_-32.1	KiDS_355.4_-30.2	KiDS_355.5_-27.2	KiDS_356.4_-33.1
KiDS_356.4_-35.1	KiDS_356.5_-31.2	KiDS_356.6_-27.2	KiDS_356.6_-29.2	KiDS_357.6_-33.1	KiDS_358.8_-35.1
KiDS_358.9_-28.2	KiDS_36.0_-34.1	KiDS_36.1_-28.2	KiDS_36.1_-31.2	KiDS_36.5_-29.2	KiDS_36.5_-32.1
KiDS_36.5_-35.1	KiDS_37.2_-34.1	KiDS_37.6_-29.2	KiDS_37.6_-32.1	KiDS_37.7_-35.1	KiDS_38.0_-30.2
KiDS_38.0_-33.1	KiDS_38.1_-27.2	KiDS_38.4_-28.2	KiDS_38.4_-31.2	KiDS_38.4_-34.1	KiDS_38.9_-35.1
KiDS_39.2_-33.1	KiDS_39.3_-27.2	KiDS_39.6_-31.2	KiDS_39.9_-29.2	KiDS_4.5_-28.2	KiDS_4.6_-29.2
KiDS_4.7_-31.2	KiDS_4.7_-32.1	KiDS_4.9_-35.1	KiDS_40.0_-32.1	KiDS_40.1_-35.1	KiDS_40.3_-30.2
KiDS_40.4_-33.1	KiDS_40.8_-31.2	KiDS_40.8_-34.1	KiDS_41.0_-29.2	KiDS_41.5_-27.2	KiDS_41.8_-28.2
KiDS_41.9_-31.2	KiDS_42.2_-29.2	KiDS_42.4_-32.1	KiDS_42.6_-27.2	KiDS_42.6_-35.1	KiDS_42.8_-33.1
KiDS_43.1_-31.2	KiDS_43.3_-29.2	KiDS_43.5_-32.1	KiDS_43.7_-27.2	KiDS_43.8_-35.1	KiDS_44.0_-28.2
KiDS_44.4_-29.2	KiDS_44.7_-32.1	KiDS_44.9_-27.2	KiDS_45.0_-35.1	KiDS_45.1_-28.2	KiDS_45.1_-33.1
KiDS_45.4_-31.2	KiDS_45.6_-26.2	KiDS_45.6_-29.2	KiDS_46.0_-27.2	KiDS_46.3_-28.2	KiDS_47.1_-32.1
KiDS_47.5_-33.1	KiDS_47.8_-29.2	KiDS_48.0_-34.1	KiDS_48.2_-32.1	KiDS_48.9_-26.2	KiDS_48.9_-31.2
KiDS_49.2_-34.1	KiDS_49.3_-27.2	KiDS_49.9_-33.1	KiDS_49.9_-35.1	KiDS_5.6_-27.2	KiDS_5.7_-29.2
KiDS_5.8_-31.2	KiDS_50.8_-28.2	KiDS_51.1_-35.1	KiDS_51.3_-31.2	KiDS_52.2_-26.2	KiDS_52.3_-33.1
KiDS_52.9_-32.1	KiDS_53.0_-28.2	KiDS_53.3_-26.2	KiDS_53.5_-29.2	KiDS_6.1_-35.1	KiDS_6.8_-28.2
KiDS_6.8_-29.2	KiDS_7.3_-35.1	KiDS_7.9_-27.2	KiDS_7.9_-28.2	KiDS_8.0_-29.2	KiDS_8.1_-30.2
KiDS_8.2_-32.1	KiDS_8.4_-34.1	KiDS_8.5_-35.1	KiDS_9.0_-27.2	KiDS_9.0_-28.2	KiDS_9.1_-29.2
KiDS_9.3_-31.2	KiDS_9.4_-32.1	KiDS_9.5_-33.1	KiDZ_209.4_5.1	KiDZ_31.6_-4.6	KiDZ_32.6_-4.6

Appendix E: VISTA on-sky quality metrics

As with Appendix B, the distribution of magnitude limits and PSF sizes on-sky in the NIR bands is a relevant statistics for understanding systematic variations of photometric quality over the KiDS-DR5 survey. Figure E.1 shows the distribution of depths in each of the five NIR filters, computed using the method of Appendix A and combined as described in Sect. 4.3. The dither pattern of the VISTA observations is visible, particularly

in cases where sequences of observations are particularly shallow (leading to stripes and rectangles in the image).

In Fig. E.2 we also show the distribution of VISTA PSF sizes on-sky, as reported by ESO for each paw-print. PSF sizes have been combined on-sky using the same procedure as for the magnitude limits. Again, the paw-print pattern of the VISTA telescope is visible. We note, however, that the distribution here is not necessarily representative of the information contained in the photometry, as the measurement of photometry by GAAP is

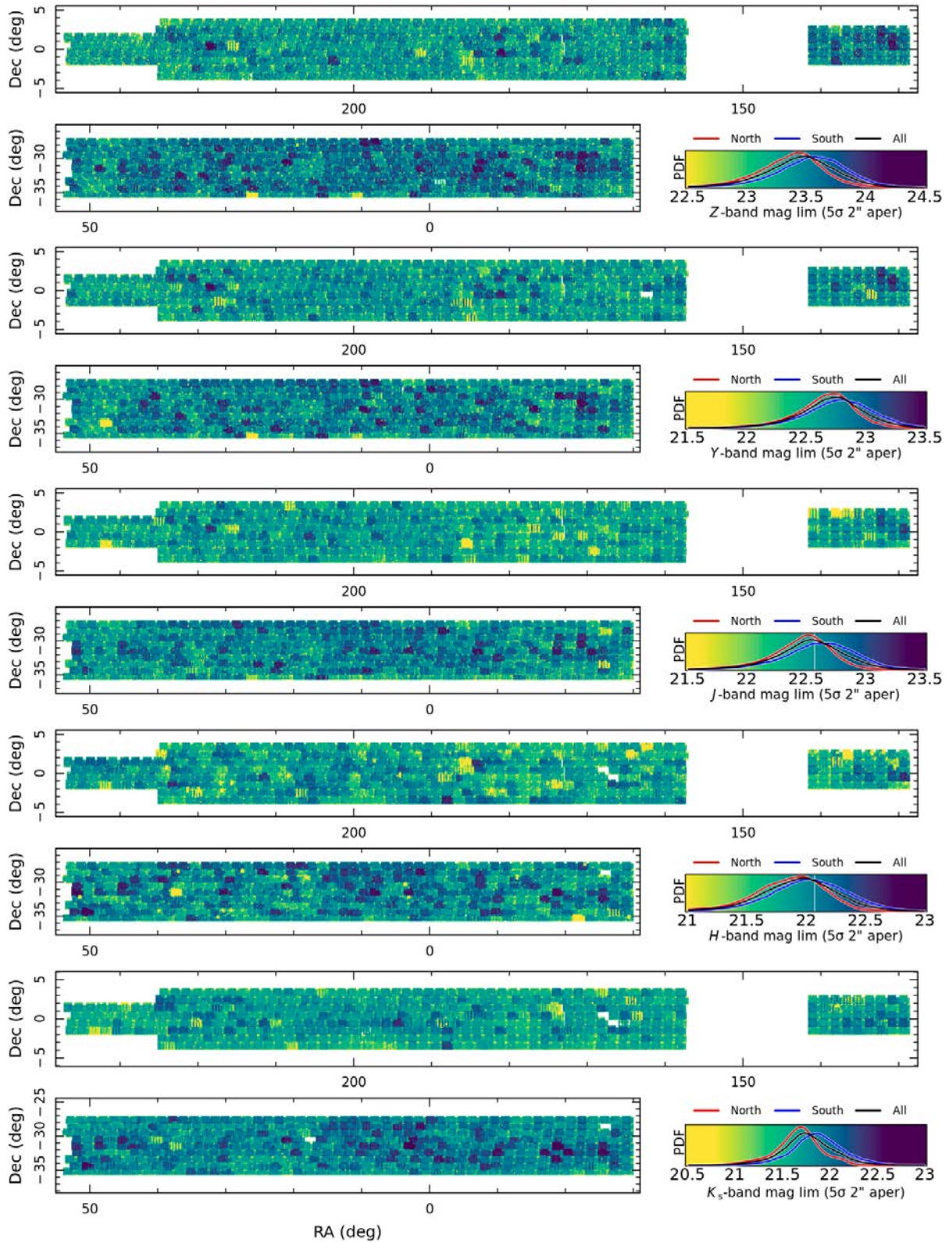


Fig. E.1. Distribution of effective NIR magnitude limits in the KiDS fields, computed using LAMBDA with 2'' circular apertures.

PSF-size dependent. In cases where the PSF is particularly poor for an observation, that observation cannot be used in the final flux calculation (as GAAP failed to produce a flux measurement in that case).

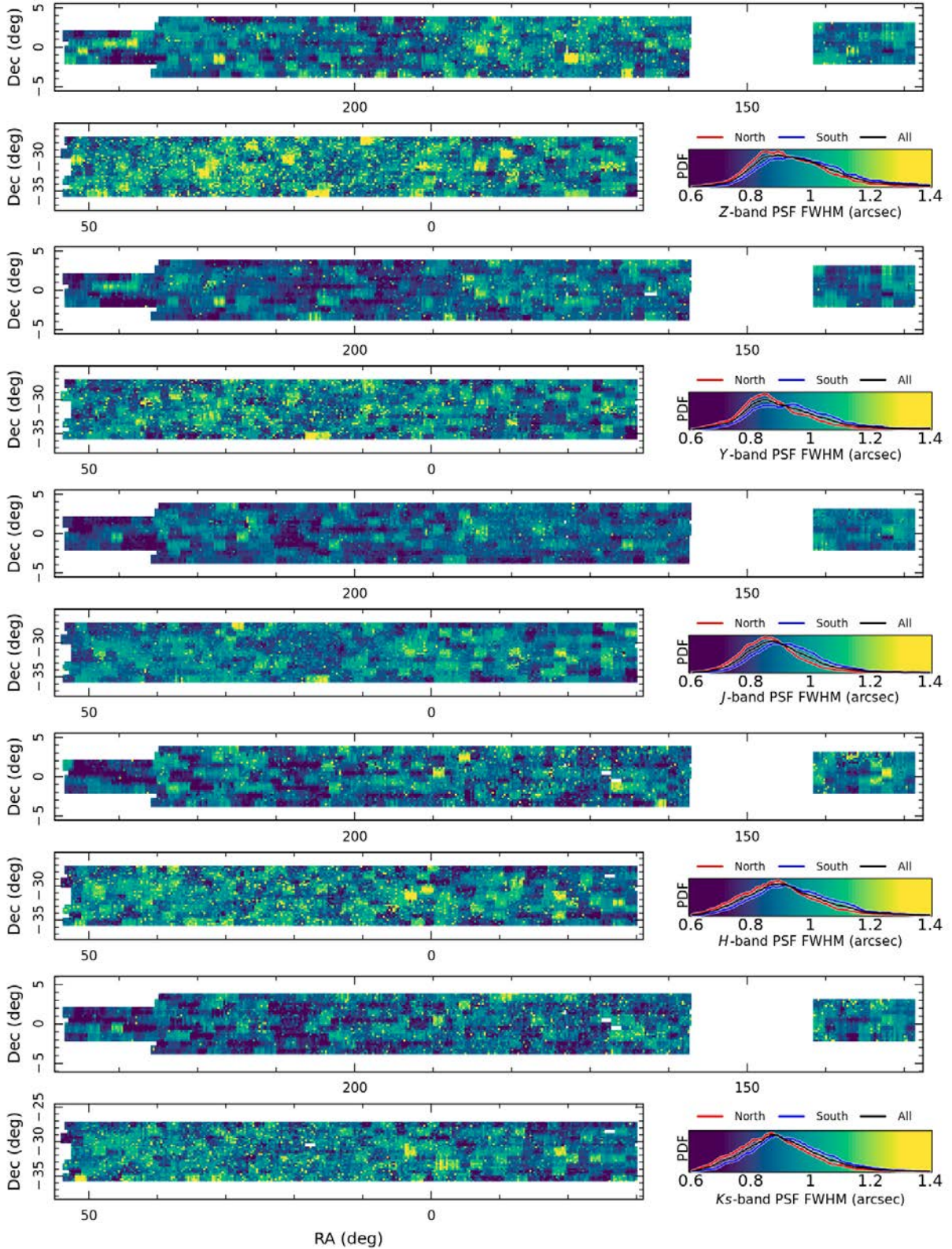


Fig. E.2. Distribution of mean PSF FWHM sizes (in arcsec) as reported by ESO for each paw-print observation of the KiDS fields. Binning is made in steps of $\{\delta RA, \delta Dec\} = \{0.25/\cos(Dec), 0.25\}$ deg, as the spacing between adjacent chips in the paw-print is ~ 0.2 deg.

Appendix F: Masking correlations

The areas masked by each of our various mask criteria are not unique. An area marked for flagging because of stellar reflections

in the r -band, for example, are likely highly correlated with areas marked for similar reflections in the g -band, because the same stars cause the artefacts in both images. Furthermore, the KiDS observing strategy exacerbates this effect because each optical

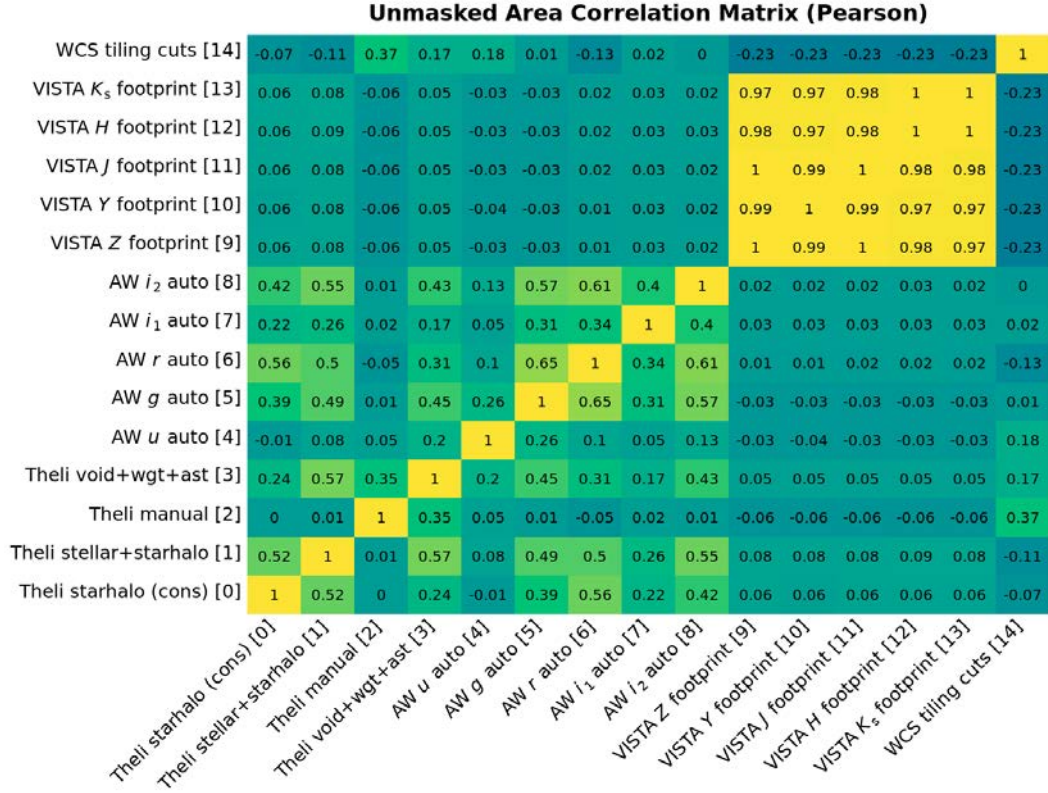


Fig. F.1. Correlations between masked area per pointing, for each of the individual mask bits. Correlations are computed between the area available per pointing after the application of the relevant mask bits, for the full KiDS-DR5 survey.

filter tiles the sky with essentially the same dither pattern. As such, it is worth noting the degree of correlation between areas masked (per pointing) by each of the bits in our mask.

Figure F.1 shows the correlation matrix of masked area per pointing, for individual masking bits. Correlations are computed between the total area masked per pointing under each condition. As such, the correlations are indirect (i.e. they do not necessarily mean the same pixels are masked in each case), but nonetheless are indicative of the approximate survey area masked in each of the cases.

The figure shows that the footprint mask in each of the NIR bands are highly correlated, as VIKING naturally observed a consistent footprint in each filter. Slight deviations from unity in the correlations of the NIR footprint masks come from the additional masking that we perform in the NIR (due to the PSF size, for example; see Sect. 6.4). Otherwise, there is a surprising lack of correlation in the masks, indicating that information contained in each of the masks is somewhat independent. This is particularly interesting in the case of the r -band, where masks constructed within ASTROWISE and THELI are maximally correlated at the level of $\sim 50\%$, despite being constructed from the same raw images. This motivates our fiducial choice of masks: individual masks from ASTROWISE and THELI flag somewhat different artefacts, and so the combination of these masks is the most appropriate (and conservative) choice.

Appendix G: Shape measurement properties

The distribution of shape measurement properties on-sky is a useful diagnostic of the level of variability in the lensing sample, after recalibration, which can influence modelling and hint

at possible systematic effects. Figure G.1 shows the on-sky distribution of PSF ellipticity in each of the five exposures (per pointing) that contribute to our shape measurement. Figure G.2 shows the distribution of measured shapes, and their uncertainty. Finally, Fig. G.3 shows the on-sky sum of the recalibrated shape-measurement weight.

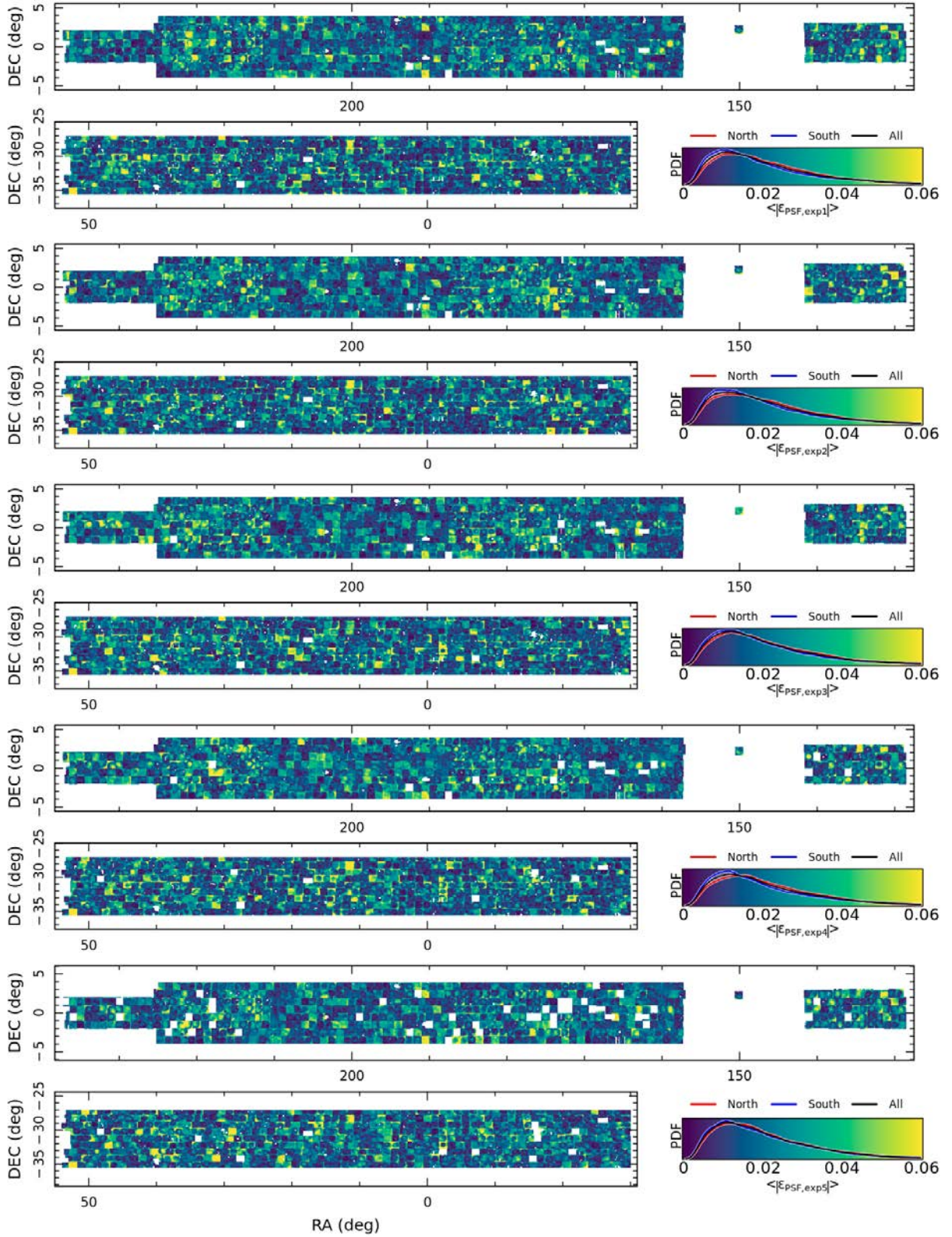


Fig. G.1. Distribution of PSF ellipticities ($|\epsilon_{\text{PSF}}|$) measured in each of the five r -band exposures used for shape measurement by *lensfit*. The panels show the relative homogeneity of the PSF measured across all exposures, and the amount of variability between exposures of the same pointing.

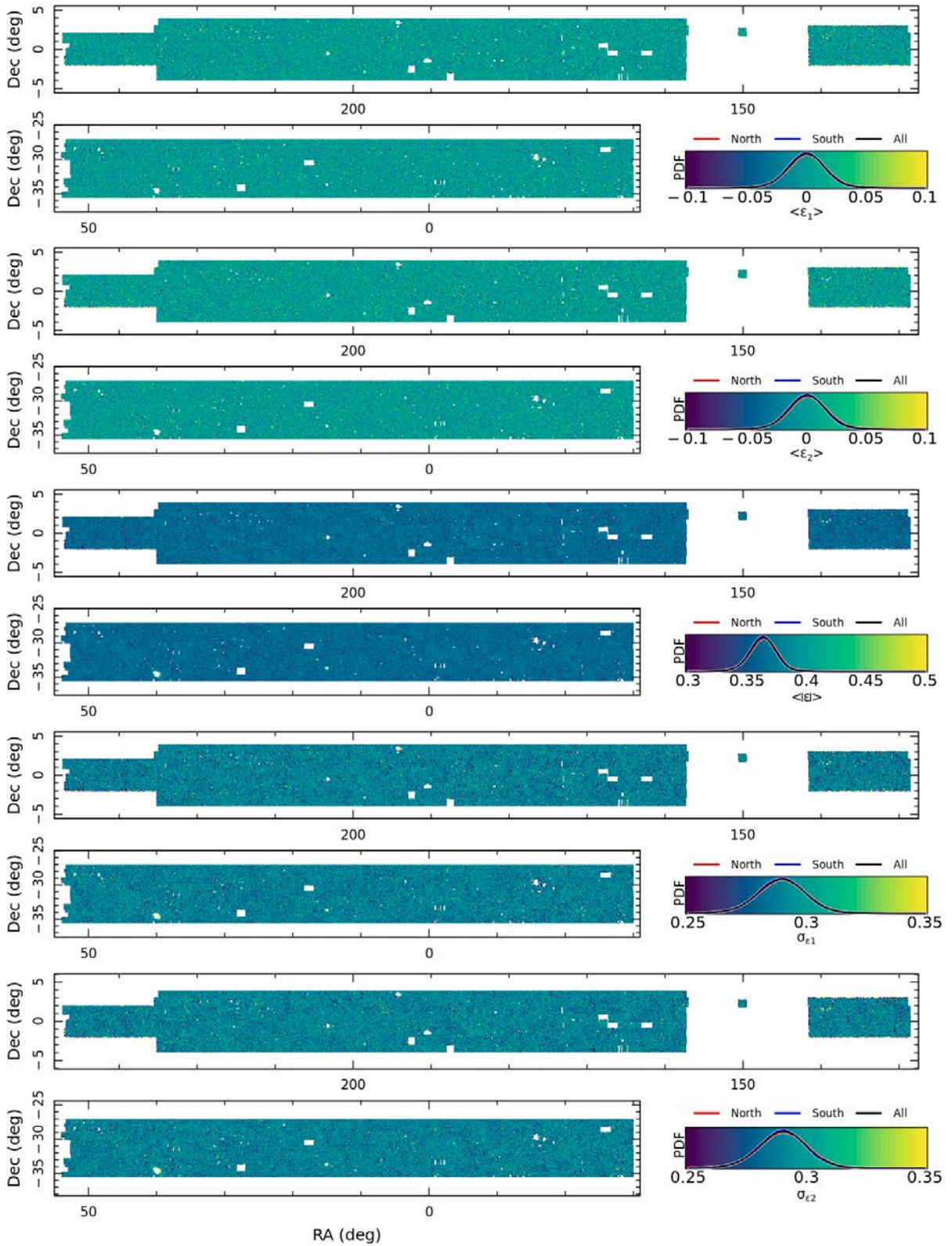


Fig. G.2. Distribution of ϵ_1 , ϵ_2 , $|\epsilon|$, σ_{ϵ_1} , and σ_{ϵ_2} measured by *lensfit* after recalibration. All estimates are unweighted.

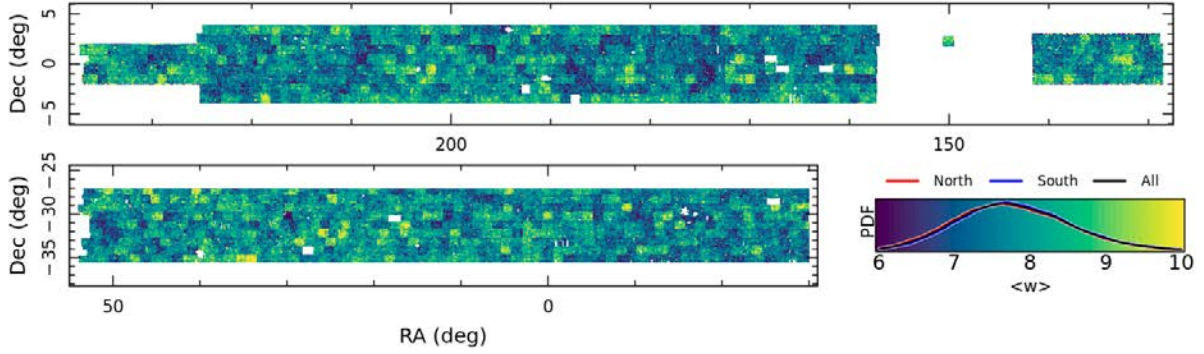


Fig. G.3. Distribution of recalibrated shape-measurement weights on-sky.

Appendix H: Sample selection with weighting

Here we provide the details of the sample selections used in KiDS-Legacy, as presented in Table 14, but computed using weighted values rather than direct source counts, and expressed in terms of the effective number density of sources $\eta = A^{-1}(\sum w)^2/(\sum w^2)$, computed using the uncalibrated lensing weight w and the survey area in square-arcminutes A . This provides an indication of the most significant selections, when considering their influence on the weighted source sample that is used for computation of, for example, shear correlation functions. The weighted selections are shown in Table H.1.

The table shows that the most significant selections, in terms of lensing weight, are the blending, fitclass, magnitude, and resolution selections. These remove, respectively: 5%, 9%, 2%, and 5% of the available effective number density.

Finally, we reproduce the fractional selection of sources as a function of r -band magnitude shown in Fig. 29, now computed as a weighted fraction in Fig. H.1. The distribution is truncated to the limits that are analysed by *lensfit* (i.e. $19 \leq r \leq 25.5$).

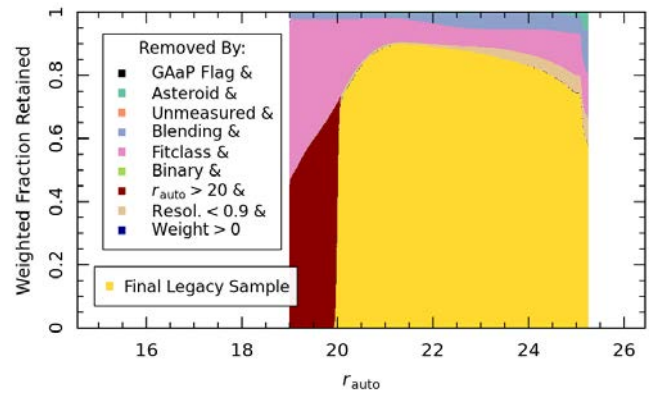


Fig. H.1. As in Figure 29, but now computed using sources weighted by their uncalibrated shape measurement weight. The distribution is truncated to the limits of analysed by *lensfit* ($19 \leq r \leq 25.5$), as the weights are not defined outside this region.

Table H.1. Statistics of the KiDS-Legacy sample definition, specified in terms of the effective number density.

Selection Name	Formula	η_s^r	f_{tot}^r	$\eta_{s,\text{run}}^r$	f_{run}^r	$\eta_{s,\text{run}}^k$	f_{run}^k
MASK	Eq. 29	-	-	-	-	10.94	100.00%
FLAG_GAAP	Eq. 30	0.01	0.08%	0.01	0.08%	10.93	99.92%
ASTEROIDS	Eq. 31	0.03	0.26%	0.03	0.26%	10.90	99.66%
UNMEASURED	Eq. 32	-	-	-	-	10.90	99.66%
BLENDING	Eq. 33	0.53	4.82%	0.52	4.75%	10.39	95.00%
FITCLASS	Eq. 34	0.97	8.88%	0.95	9.09%	9.45	86.36%
BINARY	Eq. 35	0.03	0.30%	0.00	0.04%	9.44	86.33%
MAG_AUTO	Eq. 37	0.37	3.35%	0.21	2.19%	9.26	84.60%
RESOLUTION	Eq. 20	0.82	7.47%	0.43	4.60%	8.92	81.54%
WEIGHT	Eq. 39	-	-	-	-	8.92	81.53%
Full Legacy Sample	-	-	-	-	-	8.92	81.53%
0.1 < z_B ≤ 2.0	-	-	-	-	-	8.82	80.60%
0.1 < z_B ≤ 1.2	-	-	-	-	-	8.00	73.13%

Notes. Effective number density is defined as $\eta = A^{-1}(\sum w)^2/(\sum w^2)$, computed using the uncalibrated shape measurement weight, w , and the survey area in square-arcmin, A . Superscript indices specify whether a number refers to sources that are removed from the sample ('r') or kept in the sample ('k'). η_s^r is the effective number density of sources (from the full MASK ≤ 1 sample) that are flagged for removal by the selection s . f_{tot}^r is η_s^r expressed as a fraction of the full MASK ≤ 1 sample (η_{MASK}^k). $\eta_{s,\text{run}}^r$ is the effective number density of sources from the currently un-flagged population that are newly flagged after applying selection s (assuming flags are applied in row-order). f_{run}^r is $\eta_{s,\text{run}}^r$ expressed as a fraction of the currently un-flagged population. $\eta_{s,\text{run}}^k$ is the effective number density of sources remaining un-flagged after the flagging of selection s . f_{run}^k is $\eta_{s,\text{run}}^k$ expressed as a fraction of the full MASK ≤ 1 sample.

Appendix I: ESO data products

The catalogues and images that are available within the ESO database have changed between DR4 and DR5, as the format of the data has changed. In particular, the addition of the second-pass *i*-band necessitates a modification to the catalogue (and metadata) formats. The description of the columns contained

within the ESO multi-band catalogues is provided in Table I.1, the columns contained within the single-band catalogues are provided in Table I.2. Columns that have changed meaning between DR4 and DR5 are specifically highlighted in Table I.3. Details of the image headers that are released are provided in the Table I.4.

Table I.1. Columns provided in the ten-band catalogue.

Label	Unit	Format	Description
Identifiers per source and Pointing on-sky:			
ID		30A	ESO ID
KIDS_TILE		16A	Name of the pointing in ASTRO-WISE convention
THELI_NAME		16A	Name of the pointing in THELI convention
SeqNr		1J	Running object number within the catalogue
SLID		1J	ASTRO-WISE Source list ID
SID		1J	ASTRO-WISE Source ID within the source list
Parameters derived from the THELI <i>r</i> -band detection image:			
FLUX_AUTO	Jy	1E	<i>r</i> -band flux
FLUXERR_AUTO	Jy	1E	Error on FLUX_AUTO
MAG_AUTO	mag	1E	<i>r</i> -band magnitude
MAGERR_AUTO	mag	1E	Error on MAG_AUTO
KRON_RADIUS	pixel	1E	Scaling radius of the ellipse for magnitude measurements
BackGr	Jy	1E	Background counts at centroid position
Level	Jy	1E	Detection threshold above background
MU_THRESHOLD	mag · arcsec ⁻²	1E	Detection threshold above background
MaxVal	Jy	1E	Peak flux above background
MU_MAX	mag · arcsec ⁻²	1E	Peak surface brightness above background
ISOAREA_WORLD	deg ²	1E	Isophotal area above analysis threshold
Xpos	pixel	1E	Centroid x position in the THELI image
Ypos	pixel	1E	Centroid y position in the THELI image
RAJ2000	deg	1D	Centroid sky position right ascension (J2000)
DECJ2000	deg	1D	Centroid sky position declination (J2000)
A_WORLD	deg	1E	Profile RMS along major axis
B_WORLD	deg	1E	Profile RMS along minor axis
THETA_J2000	deg	1E	Position angle (west of north)
THETA_WORLD	deg	1E	Position angle (Counterclockwise from world x-axis)
ERRA_WORLD	deg	1E	World RMS position error along major axis
ERRB_WORLD	deg	1E	World RMS position error along minor axis
ERRTHETA_J2000	deg	1E	Error on THETA_J2000
ERRTHETA_WORLD	deg	1E	Error on THETA_WORLD
FWHM_IMAGE	pixel	1E	FWHM assuming a Gaussian object profile
FWHM_WORLD	deg	1E	FWHM assuming a Gaussian object profile
Flag		1I	SOURCE EXTRACTOR extraction flags
FLUX_RADIUS	pixel	1E	Half-light radius
CLASS_STAR		1E	Star-galaxy classifier
MAG_ISO	mag	1E	<i>r</i> -band isophotal Magnitude
MAGERR_ISO	mag	1E	Error on MAG_ISO
FLUX_ISO	Jy	1E	<i>r</i> -band isophotal Flux
FLUXERR_ISO	Jy	1E	Error on FLUX_ISO
MAG_ISOCOR	mag	1E	<i>r</i> -band corrected Isophotal Magnitude
MAGERR_ISOCOR	mag	1E	Error on MAG_ISOCOR
FLUX_ISOCOR	Jy	1E	<i>r</i> -band corrected Isophotal Flux
FLUXERR_ISOCOR	Jy	1E	Error on FLUX_ISOCOR
NIMAFLAGS_ISO		1I	Number of flagged pixels over the isophotal profile
IMAFLAGS_ISO		1I	FLAG-image flags ORed over the isophotal profile
XMIN_IMAGE	pixel	1I	Minimum x-coordinate among detected pixels
YMIN_IMAGE	pixel	1I	Minimum y-coordinate among detected pixels
XMAX_IMAGE	pixel	1I	Maximum x-coordinate among detected pixels
YMAX_IMAGE	pixel	1I	Maximum y-coordinate among detected pixels
X_WORLD	deg	1D	Barycentre position along world x axis
Y_WORLD	deg	1D	Barycentre position along world y axis
X2_WORLD	deg ²	1E	Variance of position along X_WORLD (alpha)
Y2_WORLD	deg ²	1E	Variance of position along Y_WORLD (delta)
XY_WORLD	deg ²	1E	Covariance of position X_WORLD, Y_WORLD

Continued on next page

Table I.1. continued.

Label	Unit	Format	Description
ERRX2_WORLD	deg ²	1E	Error on X2_WORLD
ERRY2_WORLD	deg ²	1E	Error on Y2_WORLD
ERRXY_WORLD	deg ²	1E	Error on XY_WORLD
CXX_WORLD	deg ⁻²	1E	SOURCE EXTRACTOR Cxx object ellipse parameter
CYY_WORLD	deg ⁻²	1E	SOURCE EXTRACTOR Cyy object ellipse parameter
CXY_WORLD	deg ⁻²	1E	SOURCE EXTRACTOR Cxy object ellipse parameter
ERRCXX_WORLD	deg ⁻²	1E	Error on CXX_WORLD
ERRCYY_WORLD	deg ⁻²	1E	Error on CYY_WORLD
ERRCXY_WORLD	deg ⁻²	1E	Error on CXY_WORLD
A_IMAGE	pixel	1D	Profile RMS along x-axis
B_IMAGE	pixel	1D	Profile RMS along y-axis
ERRA_IMAGE	pixel	1E	Error on A_IMAGE
ERRB_IMAGE	pixel	1E	Error on B_IMAGE
S_ELLIPTICITY		1E	SOURCE EXTRACTOR ellipticity (1-B_IMAGE/A_IMAGE)
S_ELONGATION		1E	SOURCE EXTRACTOR elongation (A_IMAGE/B_IMAGE)
MAG_APER_4	mag	1E	<i>r</i> -band magnitude within a circular aperture of 4 pixels
MAGERR_APER_4	mag	1E	Error on MAG_APER_4
FLUX_APER_4	Jy	1E	<i>r</i> -band flux within a circular aperture of 4 pixels
FLUXERR_APER_4	Jy	1E	Error on FLUX_APER_4
...			
Similarly for radii 6, 8, 10, 14, 20, 30, 40, 60 pixels, up to			
...			
MAG_APER_100	mag	1E	<i>r</i> -band magnitude within a circular aperture of 100 pixels
MAGERR_APER_100	mag	1E	Error on MAG_APER_100
FLUX_APER_100	Jy	1E	<i>r</i> -band flux within a circular aperture of 100 pixels
FLUXERR_APER_100	Jy	1E	Error on FLUX_APER_100
IS00	pixel ²	1I	Isophotal area at level zero
IS01	pixel ²	1I	Isophotal area at level one
IS02	pixel ²	1I	Isophotal area at level two
IS03	pixel ²	1I	Isophotal area at level three
IS04	pixel ²	1I	Isophotal area at level four
IS05	pixel ²	1I	Isophotal area at level five
IS06	pixel ²	1I	Isophotal area at level six
IS07	pixel ²	1I	Isophotal area at level seven
ALPHA_J2000	deg	1D	SOURCE EXTRACTOR centroid right ascension (J2000)
DELTA_J2000	deg	1D	SOURCE EXTRACTOR centroid declination (J2000)
SG2DPHOT		1I	2DPhot StarGalaxy classifier (1 for high confidence star)
HTM		1J	Hierarchical Triangular Mesh (level 25)
FIELD_POS		1I	Reference number to field parameters
List-driven GAAP photometry on the ASTRO-WISE co-added KiDS images and the pawprint VIKING images:			
Agaper_0p7	arcsec	1E	Major axis of GAAP aperture MIN_APER 0'7
Bgaper_0p7	arcsec	1E	Minor axis of GAAP aperture MIN_APER 0'7
Agaper_1p0	arcsec	1E	Major axis of GAAP aperture MIN_APER 1'0
Bgaper_1p0	arcsec	1E	Minor axis of GAAP aperture MIN_APER 1'0
PAgaap	deg	1E	Position angle of major axis of GAAP aperture (north of west)
and then for each band $x \in \{u, g, r, i_1, i_2, Z, Y, J, H, K_s\}$: ^(a)			
FLUX_GAAP_0p7_x	Jy	1E	GAAP flux in band <i>x</i> with MIN_APER =0'7
FLUXERR_GAAP_0p7_x	Jy	1E	Error on FLUX_GAAP_0p7_x
MAG_GAAP_0p7_x	mag	1E	<i>x</i> -band GAAP magnitude with MIN_APER =0'7
MAGERR_GAAP_0p7_x	mag	1E	Error on MAG_GAAP_0p7_x
FLAG_GAAP_0p7_x		1J	GAAP Flag for <i>x</i> -band photometry with MIN_APER =0'7
FLUX_GAAP_1p0_x	Jy	1E	GAAP flux in band <i>x</i> with MIN_APER =1'0
FLUXERR_GAAP_1p0_x	Jy	1E	Error on FLUX_GAAP_1p0_x
MAG_GAAP_1p0_x	mag	1E	<i>x</i> -band GAAP magnitude with MIN_APER =1'0
MAGERR_GAAP_1p0_x	mag	1E	Error on MAG_GAAP_1p0_x
FLAG_GAAP_1p0_x		1J	GAAP Flag for <i>x</i> -band photometry with MIN_APER =1'0

Continued on next page

Table I.1. continued.

Label	Unit	Format	Description
Optimal-aperture GAAP ten-band photometry including interstellar extinction corrections			
Agaper	arcsec	1E	Major axis of GAAP aperture for optimal MIN_APER
Bgaper	arcsec	1E	Minor axis of GAAP aperture for optimal MIN_APER
and then for each band $x \in \{u, g, r, i_1, i_2, Z, Y, J, H, K_s\}$: ^(a)			
EXTINCTION_x	mag	1E	Galactic extinction in band x
MAG_GAAP_x	mag	1E	Corrected x -band GAAP magnitude with optimal MIN_APER
MAGERR_GAAP_x	mag	1E	Error on MAG_GAAP_x
FLUX_GAAP_x	Jy	1E	x -band GAAP flux with optimal MIN_APER
FLUXERR_GAAP_x	Jy	1E	Error on FLUX_GAAP_x
FLAG_GAAP_x		1I	GAAP Flag for x -band photometry with optimal MIN_APER
MAG_LIM_x	mag	1E	x -band limiting magnitude for optimal MIN_APER
Ten-band photometric redshifts (BPZ):			
Z_B		1D	Ten-band BPZ redshift estimate (posterior mode)
Z_B_MIN		1D	Lower bound of the 68% confidence interval of Z_B
Z_B_MAX		1D	Upper bound of the 68% confidence interval of Z_B
T_B		1D	Spectral type corresponding to Z_B
ODDS		1D	Empirical ODDS of Z_B
Z_ML		1D	Ten-band BPZ maximum likelihood redshift
T_ML		1D	Spectral type corresponding to Z_ML ^(b)
CHI_SQUARED_BPZ		1D	chi squared value associated with Z_B
M_0	mag	1D	Reference magnitude for BPZ prior
BPZ_FILT		1J	filters with good photometry (BPZ)
NBPZ_FILT		1J	number of filters with good photometry (BPZ)
BPZ_NONDEFILT		1J	filters with faint photometry (BPZ)
NBPZ_NONDEFILT		1J	number of filters with faint photometry (BPZ)
BPZ_FLAGFILT		1J	flagged filters (BPZ)
NBPZ_FLAGFILT		1J	number of flagged filters (BPZ)
SG_FLAG		1E	Star/Gal Classifier
MASK		1J	Ten-band mask information ^(c)
PSF estimates (<i>lensfit</i>)			
PSF_e1		1E	mean ellipticity of PSF, component one
PSF_e2		1E	mean ellipticity of PSF, component two
PSF_Strehl_ratio		1E	Pseudo-Strehl ratio of PSF
PSF_Q11		1E	model PSF moment Q11
PSF_Q22		1E	model PSF moment Q22
PSF_Q21		1E	model PSF moment Q21
Stellar Mass estimates (LePhare) ^(d)			
for each band $x \in \{u, g, r, i_1, i_2, Z, Y, J, H, K_s\}$:			
MAGABS_GAAP_x		1D	rest-frame x -band magnitude
KCOR_x		1D	x -band k -correction
LUM_GAAP_r_bestfit		1D	r -band luminosity
mstar_bestfit		1D	stellar mass
mstar_med		1D	posterior median stellar mass
mstar_lower		1D	posterior 16 th percentile stellar mass
mstar_upper		1D	posterior 84 th percentile stellar mass
sfr_bestfit		1D	star-formation rate
sfr_med		1D	posterior median star-formation rate
sfr_lower		1D	posterior 16 th percentile star-formation

Continued on next page

Table I.1. continued.

Label	Unit	Format	Description
sfr_upper		1D	posterior 84 th percentile star-formation
Dereddened colours based on optimal-aperture GAAP photometry			
COLOUR_GAAP_u_g	mag	1E	$u - g$ colour index (dereddened)
COLOUR_GAAP_g_r	mag	1E	$g - r$ colour index (dereddened)
COLOUR_GAAP_r_i1	mag	1E	$r - i_1$ colour index (dereddened)
COLOUR_GAAP_r_i2	mag	1E	$r - i_2$ colour index (dereddened)
COLOUR_GAAP_i1_Z	mag	1E	$i_1 - Z$ colour index (dereddened)
COLOUR_GAAP_i2_Z	mag	1E	$i_2 - Z$ colour index (dereddened)
COLOUR_GAAP_Z_Y	mag	1E	$Z - Y$ colour index (dereddened)
COLOUR_GAAP_Y_J	mag	1E	$Y - J$ colour index (dereddened)
COLOUR_GAAP_J_H	mag	1E	$J - H$ colour index (dereddened)
COLOUR_GAAP_H_Ks	mag	1E	$H - K_s$ colour index (dereddened)
DIFF_GAAP_i1_i2	mag	1E	$i_1 - i_2$ i -band magnitude difference

Notes. ^(a)See Sect. 6.5 for the definitions of the flux and magnitude zero-points. ^(b)Definition of the spectral types: 1=CWW-EII, 2=CWW-Sbc, 3=CWW-Scd, 4=CWW-Im, 5=KIN-SB3, 6=KIN-SB2 ^(c)For the meaning of the mask see Table 11 ^(d)Stellar masses are computed in the same manner as described in Wright et al. (2019)

Table I.2. Columns provided in the single-band source lists.

Label	Format	Unit	Description
2DPHOT	J	Source	classification (see sect. 4.5.1 in de Jong et al. 2015)
X_IMAGE	E	pixel	Object position along x
Y_IMAGE	E	pixel	Object position along y
NUMBER	J		Running object number
CLASS_STAR	E		SEXTRACTOR S/G classifier
FLAGS	J		Extraction flags
IMAFLAGS_ISO	J		FLAG-image flags summed over the iso. profile
NIMAFLAG_ISO	J		Number of flagged pixels entering IMAFLAGS_ISO
FLUX_RADIUS	E	pixel	Radius containing half of the flux
KRON_RADIUS	E	pixel	Kron apertures in units of A or B
FWHM_IMAGE	E	pixel	FWHM assuming a Gaussian core
ISOAREA_IMAGE	J	pixel ²	Isophotal area above Analysis threshold
ELLIPTICITY	E		$1 - B_IMAGE/A_IMAGE$
THETA_IMAGE	E	deg	Position angle (CCW/x)
MAG_AUTO	E	mag	Kron-like elliptical aperture magnitude
MAGERR_AUTO	E	mag	RMS error for AUTO magnitude
ALPHA_J2000	D	deg	Right ascension of barycentre (J2000)
DELTA_J2000	D	deg	Declination of barycentre (J2000)
FLUX_APER_2	E	count	Flux within circular aperture of diameter 2 pixels
...
FLUX_APER_200	E	count	Flux within circular aperture of diameter 200 pixels
FLUXERR_APER_2	E	count	RMS error for flux within aperture of diameter 2 pixels
...
FLUXERR_APER_200	E	count	RMS error for flux within aperture of diameter 200 pixels
MAG_ISO	E	mag	Isophotal magnitude
MAGERR_ISO	E	mag	RMS error for isophotal magnitude
MAG_ISOCOR	E	mag	Corrected isophotal magnitude (deprecated)
MAGERR_ISOCOR	E	mag	RMS error for corrected isophotal magnitude
MAG_BEST	E	mag	Best of MAG_AUTO and MAG_ISOCOR
MAGERR_BEST	E	mag	RMS error for MAG_BEST
BACKGROUND	E	count	Background at centroid position
THRESHOLD	E	count	Detection threshold above background
MU_THRESHOLD	E	arcsec ⁻²	Detection threshold above background
FLUX_MAX	E	count	Peak flux above background
MU_MAX	E	arcsec ⁻²	Peak surface brightness above background
ISOAREA_WORLD	E	deg ²	Isophotal area above Analysis threshold
XMIN_IMAGE	J	pixel	Minimum x-coordinate among detected pixels

Continued on next page

Table I.2. Continued.

Label	Format	Unit	Description
YMIN_IMAGE	J	pixel	Minimum y-coordinate among detected pixels
XMAX_IMAGE	J	pixel	Maximum x-coordinate among detected pixels
YMAX_IMAGE	J	pixel	Maximum y-coordinate among detected pixels
X_WORLD	D	deg	Barycentre position along world x axis
Y_WORLD	D	deg	Barycentre position along world y axis
XWIN_IMAGE	E	pixel	Windowed position estimate along x
YWIN_IMAGE	E	pixel	Windowed position estimate along y
X2_IMAGE	D	pixel ²	Variance along x
Y2_IMAGE	D	pixel ²	Variance along y
XY_IMAGE	D	pixel ²	Covariance between x and y
X2_WORLD	E	deg ²	Variance along X-WORLD (alpha)
Y2_WORLD	E	deg ²	Variance along Y-WORLD (delta)
XY_WORLD	E	deg ²	Covariance between X-WORLD and Y-WORLD
CXX_IMAGE	E	pixel ⁻²	Cxx object ellipse parameter
CYY_IMAGE	E	pixel ⁻²	Cyy object ellipse parameter
CXY_IMAGE	E	pixel ⁻²	Cxy object ellipse parameter
CXX_WORLD	E	deg ⁻²	Cxx object ellipse parameter (WORLD units)
CYY_WORLD	E	deg ⁻²	Cyy object ellipse parameter (WORLD units)
CXY_WORLD	E	deg ⁻²	Cxy object ellipse parameter (WORLD units)
A_IMAGE	D	pixel	Profile RMS along major axis
B_IMAGE	D	pixel	Profile RMS along minor axis
A_WORLD	E	deg	Profile RMS along major axis (WORLD units)
B_WORLD	E	deg	Profile RMS along minor axis (WORLD units)
THETA_WORLD	E	deg	Position angle (CCW/world-x)
THETA_J2000	E	deg	Position angle (east of north) (J2000)
ELONGATION	E	deg	A_IMAGE/B_IMAGE
ERRX2_IMAGE	E	pixel ²	RMS error on variance of position along x
ERRY2_IMAGE	E	pixel ²	RMS error on variance of position along y
ERRXY_IMAGE	E	pixel ²	RMS error on covariance between x and y position
ERRX2_WORLD	E	deg ²	RMS error on variance of position along X-WORLD (alpha)
ERRY2_WORLD	E	deg ²	RMS error on variance of position along Y-WORLD (delta)
ERRXY_WORLD	E	deg ²	RMS error on covariance between X-WORLD and Y-WORLD
ERRCXX_IMAGE	E	pixel ⁻²	Cxx error ellipse parameter
ERRCYY_IMAGE	E	pixel ⁻²	Cyy error ellipse parameter
ERRCXY_IMAGE	E	pixel ⁻²	Cxy error ellipse parameter
ERRCXX_WORLD	E	deg ⁻²	Cxx error ellipse parameter (WORLD units)
ERRCYY_WORLD	E	deg ⁻²	Cyy error ellipse parameter (WORLD units)
ERRCXY_WORLD	E	deg ⁻²	Cxy error ellipse parameter (WORLD units)
ERRA_IMAGE	E	pixel	RMS position error along major axis
ERRB_IMAGE	E	pixel	RMS position error along minor axis
ERRA_WORLD	E	deg	World RMS position error along major axis
ERRB_WORLD	E	deg	World RMS position error along minor axis
ERRTHETA_IMAGE	E	deg	Error ellipse position angle (CCW/x)
ERRTHETA_WORLD	E	deg	Error ellipse position angle (CCW/world-x)
ERRTHETA_J2000	E	deg	J2000 error ellipse pos. angle (east of north)
FWHM_WORLD	E	deg	FWHM assuming a Gaussian core
ISO0	J	pixel ²	Isophotal area at level zero
ISO1	J	pixel ²	Isophotal area at level one
ISO2	J	pixel ²	Isophotal area at level two
ISO3	J	pixel ²	Isophotal area at level three
ISO4	J	pixel ²	Isophotal area at level four
ISO5	J	pixel ²	Isophotal area at level five
ISO6	J	pixel ²	Isophotal area at level six
ISO7	J	pixel ²	Isophotal area at level seven
SLID	K		ASTRO-WISE SourceList identifier
SID	K		ASTRO-WISE source identifier
HTM	K		Hierarchical Triangular Mesh (level 25)
FLAG	K		Not used

Table I.3. Changes to variable names and definitions in DR5 compared to DR4.

Variables Name Changes		
Previous Name	New Name	Definition
MAG_AUTO	MAG_AUTO_raw	Uncorrected r -band AUTO magnitude
MAG_GAAP_i	–	Deleted
–	MAG_GAAP_i1	New column
–	MAG_GAAP_i2	New column
–	FLUX_GAAP_i1	New column
–	(etc. for errors)	

Variables Definition Changes		
Name	Old Definition	New Definition
MASK	Defined as in Kuijken et al. (2019)	Defined as in Table 11

Table I.4. Main keywords in the ten-band catalogue headers

Keyword	Description
RA	Field centre (J2000) (deg)
DEC	Field centre (J2000) (deg)
PROV[1, 2, 3, 4]	Originating $[u, g, r, i_1, i_2]$ -band co-add file name
FPRA[1, 2, 3, 4]	Footprint [SE,NE,NW,SW] corner RA (deg)
FPDE[1, 2, 3, 4]	Footprint [SE,NE,NW,SW] corner DEC (deg)
CALSTARS	Number of stars used for <i>Gaia</i> calibration (MIN_APER=0'7)
DMAG_[U, G, R, I1]	SLR+ <i>Gaia</i> $[u, g, r, i_1]$ -band offset (MIN_APER=0'7)
DMAG_[U, G, R, I2]	SLR+ <i>Gaia</i> $[u, g, r, i_2]$ -band offset (MIN_APER=0'7)
CALSTR_1	Number of stars used for <i>Gaia</i> calibration (MIN_APER=1'0)
DMAG_[U, G, R, I1]_1	SLR+ <i>Gaia</i> $[u, g, r, i_1]$ -band offset (MIN_APER=1'0)
DMAG_[U, G, R, I1]_2	SLR+ <i>Gaia</i> $[u, g, r, i_2]$ -band offset (MIN_APER=1'0)
OB[U, G, R, I1, I2]_STRT	$[u, g, r, i_1, i_2]$ -band Observing Block start
ASSON1	Associated ten-band mask
ASSON2	Associated THELI r -band detection image
ASSON3	Associated THELI r -bandweight image