Noname manuscript No.
(will be inserted by the editor)

The Kilo-Degree Survey

Jelte T. A. de Jong, Gijs Verdoes Kleijn, Konrad Kuijken, Edwin Valentijn, KiDS and Astro-WISE consortiums

Received: date / Accepted: date

Abstract The Kilo Degree Survey (KiDS) is a 1500 square degree optical imaging survey with the recently commissioned OmegaCAM wide-field imager on the VLT Survey Telescope (VST). A suite of data products will be delivered to ESO and the community by the KiDS survey team. Spread over Europe, the KiDS team uses Astro-WISE to collaborate efficiently and pool hardware resources. In Astro-WISE the team shares, calibrates and archives all survey data. The data-centric architectural design realizes a dynamic 'live archive' in which new KiDS survey products of improved quality can be shared with the team and eventually the full astronomical community in a flexible and controllable manner.

Keywords wide-field imaging · survey system · VLT/VST · weak gravitational lensing · photometric redshifts

1 Introduction

One of the radical advances that optical astronomy has seen in recent years is the advent of wide-field CCD-based surveys. Key ingredients for these surveys are the availability of, on the one hand, instruments with sufficiently large arrays of high quality CCD's and, on the other hand, information systems with sufficient computing and data storage capabilities to process the huge data flows. The formidable scientific relevance and importance of such surveys, particularly when freely available to researchers, is clearly demonstrated by the

F. Author
first address
Tel.: +123-45-678910
Fax: +123-45-678910
E-mail: fauthor@example.com

S. Author
second address

impact that the Sloan Digital Sky Survey (SDSS, 1), has had in several fields in astronomy.

While most large-area surveys so far have operated from the Northern hemisphere, a similar large scale survey has not been performed from the South. For European astronomy, however, the Southern hemisphere is especially important, due to the presence of ESO's Very Large Telescope (VLT) and its large array of instruments. This is now remedied with the arrival of ESO's own two dedicated survey telescopes: VISTA in the (near-)infrared and VST (VLT Survey Telescope) in the optical. Both have become operational during the past two years. The lion's share of the observing time on both survey telescopes will be invested in a set of 'public surveys'. The largest of the optical surveys is the Kilo-Degree Survey (KiDS), which is imaging 1500 square degrees in four filters (u,g,r,i) over a period of 3–4 years. Combined with one of the VISTA surveys, VIKING, which is observing the same area in ZYJHK, this will provide a sensitive, 9-band multi-colour survey.

Specifically for the handling of surveys from the VST the Astro-WISE system (16) has been designed. It allows processing, quality control and public archiving of surveys using a distributed architecture. The KiDS survey team, that is spread over different countries, performs these survey operations in Astro-WISE as a single virtual team making intensive use of web-based collaborative interfaces. This paper will discuss both the observational set-up of the KiDS survey and its primary scientific goals, as well as how the Astro-WISE system will be used to achieve these goals.

2 The Kilo-Degree Survey

2.1 Observational survey set-up

KiDS will cover 1500 square degrees, some 7% of the extragalactic sky. It consists of two patches, ensuring that observations can take place year-round. The Northern patch lies on the celestial equator, while the Southern area straddles the South Galactic Pole (Fig. 1). Together the two patches cover a range of galactic latitudes from 40 to 90 degrees, and the 10 degree width of the strips ensures that the full 3D structure of the universe is sampled well. These specific areas were chosen because they have been the target of massive spectroscopic galaxy surveys already: the 2dF redshift survey (5) covers almost the same area, and KiDS-N overlaps with the SDSS spectroscopic and imaging survey as well. This means that several 100,000 galaxy spectra and redshifts are already known in these fields, and hence that the cosmological foreground mass distribution in these fields is well mapped out. Extinction in the fields is low. The exposure times for KiDS and VIKING have been chosen to yield a median galaxy redshift of 0.8, so that the evolution of the galaxy population and matter distribution over the last $\sim$ half of the age of the universe can be studied. They are also well-matched to the natural exposure times for efficient VST and VISTA operations, and balanced over the astro-climate conditions

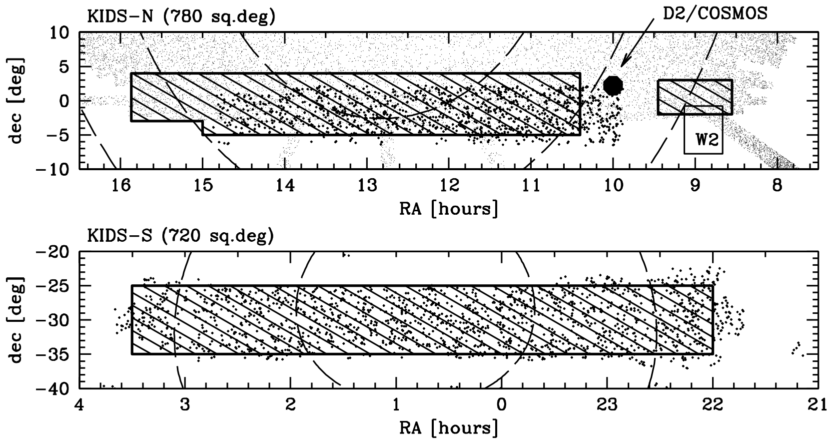

Fig. 1 Lay-out of the KIDS-North (top) and KIDS-South (bottom) fields, shown by the hatched areas. Also shown are the areas where 2DF spectra are available, indicated by the large dots, and the area covered by DR7 of the SDSS survey, indicated by the small dots. The CFHTLS-W2 field and the DS/COSMOS deep field are overplotted on the top panel.

Table 1 KiDS Exposure Times and Observational Constraints

Filter	Exposure time (seconds)	Mag limit (AB 5σ 2")	Seeing (arcsec)	Moon phase	ADC used	Airmass
u	900	24.8	0.9–1.1	Dark	no	1.2
g	900	25.4	0.7–0.9	Dark	no	1.6
r	1800	25.2	<0.7	Dark	no	1.2
i	1080	24.2	<1.1	Any	no	2.0

on Paranal (seeing and moon phase) so that all bands can be observed at the same average rate. This strategy makes optimal use of the fact that all observations are queue-scheduled, making it possible to use the best seeing time for deep r-band exposures, for example, and the worst seeing for u. All exposure times and observing constraints are listed in Table 1.

To avoid holes in the KiDS images, observations will use 5 dithered observations per field in g, r and i and 4 in u. The dithers form a staircase pattern with dither steps of 25" in X and 85" in Y. These offsets bridge the inter-CCD gaps of OmegaCAM. The survey tiling is derived using a tiling strategy that can tile the full sky efficiently for the OmegaCAM instrument. Neighboring tiles have an overlap in RA of 1.2% and in DEC of 4% to derive the internal accuracy on photometry and astrometry required for the most stringent science cases. The Atmospheric Dispersion Corrector (ADC) of OmegaCAM could be used for all KiDS bands except u. However, KiDS does not make use of the ADC to avoid the small losses in sensitivity. Instead, the dispersion is limited by constraining the maximum airmass (see Table 1). Particularly for r this is important since this band will be used for weak lensing analyses and therefore requires a well-behaved PSF. By constraining the maximum airmass in r to 1.2, the spectral dispersion will be <0.2".

After completion of the main survey of 1500 square degrees, the whole survey area will be imaged once more in g-band. The observational set-up of this repeat pass is the same is for the main survey observations in g, with the additional requirement that it should provide at least a 2 year baseline over the whole survey area. This repeat pass will allow for accurate proper motion measurements.

2.2 Scientific goals

The central science case for KiDS and VIKING is mapping the matter distribution in the universe through weak gravitational lensing and photometric redshift measurements. However, the enormous data set that KiDS will deliver, will have many more possible applications. The main research topics that the KiDS team members will explore are outlined below.

2.2.1 Dark energy

Dark energy manifests itself in the expansion history of the universe, as a repulsive term that appears to behave like Einstein's cosmological constant. Understanding its properties more accurately is one of the central quests of cosmology of recent years. With KiDS we intend to push this question as far as possible, while recognising that the limiting factor may well be systematic effects rather than raw statistical power. Measurements done with KiDS will therefore also serve as a learning curve for future (space-based?) experiments.

The use of weak gravitational lensing as a cosmological probe is nicely summarized in (11) and (14). Essentially, its power relies on two facts: gravitational lensing is a very geometric phenomenon, and it is sensitive to mass inhomogeneities along the lines of sight.This makes it a good probe of the growth of structure with time (redshift), as well as being a purely geometric distance measure. As it happens, the distance-redshift relation and the speed with which overdensities grow with cosmic time are the two most fundamental measures of the energy content of the universe: both depend directly on the rate at which the universe expands. Making such a measurement, for which weak lensing is an excellent method, is therefore of great interest. These lensing measurements are not easy, as they require systematics better than 1% accuracy, and photometric redshifts unbiased at a similar level. However, with KiDS we have put ourselves in the optimal position to attempt this, by ensuring the best image quality in our instrument, by choosing a survey depth and area appropriately, and having a wide wavelength coverage that will make the photometric redshifts as free of error as is possible with wide-band photometry (see Fig. 2). As noted above, the expansion history can be deduced from lensing tomography in several ways, and requiring consistency is a powerful

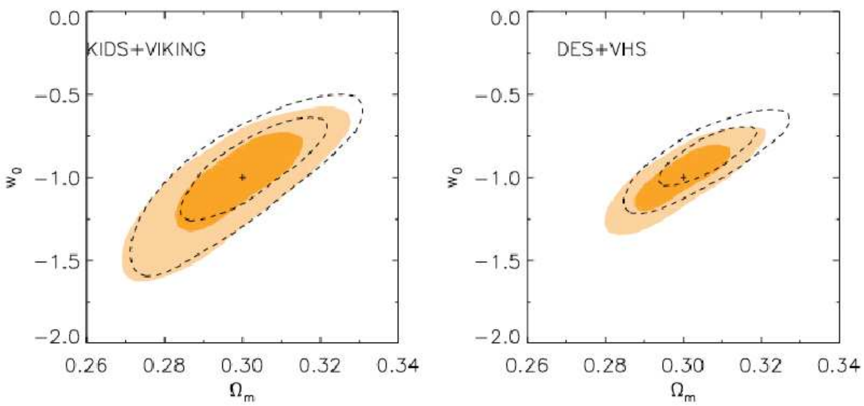

Fig. 2 Comparison of the formal statistical power and sensitivity to systematic errors in the photometric redshifts for the KiDS/VIKING and DES/VHS surveys in cosmological parameter estimation (here the matter density Ω_m and the dark energy equation of state w_0), based on simulated photometry for each of the surveys (Szomoru, Hildebrandt and Hoekstra, private comm.). The '+' represents the input truth. The coloured contours assume perfect redshift information, while the dashed contours show the effect of redshift errors. Flat geometry was assumed here, but otherwise no external information were included. Once external information is folded in, the constraints tighten and systematic effects become even more significant, demonstrating the greater robustness of the KiDS survey to this type of systematic error.

check – as well as, further down the line, an interesting test of Einstein gravity theory (e.g. 8).

An independent way to study the expansion history of the universe is by measuring the baryon acoustic oscillations (BAO). BAO is the clustering of baryons at a fixed co-moving length scale, set by the sound horizon at the time that the universe recombined and photons decoupled from baryonic matter. This scale length, which has been measured accurately in the cosmic microwave background (e.g. 15), is therefore a standard ruler, whose angular size on the sky provides a direct measurement of the angular diameter-redshift relation and hence of the expansion history. Using photometric redshifts from KiDS we can make an independent measurement of the BAO scale. Comparison of the results with ongoing spectroscopic BAO surveys provides a potent test of systematics. Based on simulations of photometric redshift measurements using *ugri* from KiDS and YJHK from VISTA we are confident that this accuracy can be achieved. Tests of the detectability of the BAO with particle and Monte-Carlo simulations, provided by Peter Schuecker, have shown that imaging surveys of the size and sensitivity of KiDS can yield values of w with $\sim$5% accuracy.

2.2.2 Structure of galaxy halos

Simulations of structure formation provide detailed information about the shape of dark matter halos on large scales. However, at small scales such as the inner parts of galaxy halos, complex physics that these simulations can not represent realistically (e.g. star formation, cooling, feedback etc.) starts to play an important role. The relation between light (baryons) and mass (dark matter) is crucial for our understanding of the influence of the dark matter on galaxy formation, and vice-versa. Galaxy-galaxy lensing (GGL) provides a unique way to study this relation between galaxies and their dark halos.

The gravitational lensing effect of foreground galaxies on the images of background galaxies is very weak, and can only be measured statistically. This is done by stacking large numbers of foreground galaxies and measuring the net image distortion of the background galaxies. On small scales (3–30 arcsec) the signal is dominated by the profile of the foreground galaxies' inner dark matter halos, at radii of 10 to 100s of kpc. At scales of several arcminutes GGL probes the galaxy–mass correlation and the bias parameter, while at even larger scales the distribution of the foreground galaxies in their parent group halos dominates. GGL can therefore be used to probe halos over a large range of scales and help to test the universality of the dark halo profile.

The strength of KiDS for GGL is again twofold. The shear size and thus enormous numbers of available galaxies, makes it possible to split the foreground galaxies in bins and study different galaxy types separately. The accurate photometric redshifts also allow splitting up the samples in redshift bins, thus enabling the redshift dependence to be analyzed. Furthermore, the fact that KiDS targets areas where wide-field redshift surveys have already been carried out, means that the foreground large scale structure is known, enabling the measurement of the galaxy–mass correlation for galaxy groups, clusters, and even filaments. Compared to earlier GGL studies with, for example, SDSS (e.g. 10) or CFHTLS (e.g. 13), the image quality and sensitivity of KiDS will provide many more foreground–background pairs, more accurate shape measurements, and the ability to probe the galaxy population at higher redshifts.

2.2.3 Evolution of galaxies and clusters

Within the current cosmological paradigm of Cold Dark Matter (CDM), structure formation is hierarchical and the profiles of CDM halos are universal, i.e. the same at all scales. Several of the ramifications of this picture have so far alluded rigorous observational testing. For example, various observational constraints on the influence of galaxy mergers on the evolution of the galaxy population at redshifts higher than $\sim$0.5 differ up to an order of magnitude (see e.g. (9)). The observational studies targeted small numbers of galaxies (< 1000) at high spatial resolution (e.g., (12)) or small areas (< 1 square degree, e.g. (6)). Also, galaxy clusters are probes of the highest mass peaks in the universe, but at redshifts of $z > 1$ the number of known galaxy clusters is yet too small to constrain cosmological models.

KiDS can play a major role in this field. The sensitivity of the KiDS photometry will result in the detection of an estimated 10^8 galaxies. This galaxy sample will have a median redshift of $z = 0.8$, with $\sim$20% having $1 < z < 1.5$.

Based on this sample the evolution of the galaxy luminosity function, the build-up of stellar mass and the assembly of early-type stellar systems can be traced back to unprecedented look-back times.

Cluster finding will be possible directly from the multi-colour KiDS catalogues. In total we expect KiDS to provide $1 - 2 \times 10^4$ clusters, and with the red sequence detectable out to $z \sim 1.4$ approximately 5% of these will be located at redshifts beyond 1. This will be a very important sample to further constrain cosmological parameters, provided that the relation between cluster richness and cluster mass can be calibrated. This calibration is possible since the weak lensing measurements that will be done as part of KiDS will probe the cluster mass distribution, demonstrating the pivotal advantage of combining high image quality with uniform multi-band photometry.

A different perspective of galaxy evolution will be provided by virtue of the fact that that KiDS-S overlaps two nearby superclusters (Pisces-Cetus and Fornax-Eridanus). Thus, the relation between galaxy properties (e.g. star formation rate) and environment, can be studied all the way from cluster cores to the infall regions, and to the filaments that connect clusters in the cosmic web.

2.2.4 Stellar streams and the Galactic halo

Detailed studies of the stellar halo of the Milky Way require photometry of faint stars over large areas of sky. The SDSS, although primarily aimed at cosmology and high-redshift science, has proved a milestone in Milky Way science as well, unveiling many stellar streams and unknown faint dwarf Spheroidal galaxies (e.g. 2; 3). While KiDS will image a smaller area than SDSS, it is deeper and thus will provide a view on more distant parts of the halo. But more importantly, SDSS only covered the Northern sky, leaving the Southern hemisphere as uncharted territory. Particularly in the KiDS-S area, new discoveries are bound to be made in the direct vicinity of our own Galaxy.

2.2.5 Proper motions

Proper motions of precision $\sim 0.4'' yr^{-1}$ will be available in the KiDS area, owing to the planned repeat pass that will provide a 2-year baseline. Several applications are possible, among others the detection and study of high proper-motion white dwarfs. "Ultracool" white dwarfs (<4000 K), relics from the earliest epochs of star formation, are among the oldest objects in the Galaxy and can be used to trace the very early star formation history of our Galaxy. Due to its multicolour photometry combined with proper motion information, KiDS will be able to increase the sample of known, ultracool white dwarfs significantly.

2.3 Data products

Being a Public Survey, all KiDS data will be made publicly available. The KiDS catalogue will contain some 100,000 sources per square degree (150 million sources over the full survey area), and for each square degree there will be 10GB of final image data, 15TB for the whole survey. These data will be of no use if they are not uniformly and carefully calibrated and made available in an easily accessible archive. Basic data products will be made public, both through ESO and through the Astro-WISE database, within a year after any part of the survey area has been observed in all filters. This set of basic data products includes the following:

- astrometrically and photometrically calibrated coadded and regridded images with weight maps;
- calibration images: twilight flats, dome flats, biases, fringe maps, etc.;
- single-band source catalogues;
- associated catalogues of these single-band catalogues.

In addition, and on a longer time-scale, we intend to provide more refined and advanced data products. In the context of the lensing project for KiDS several innovative image processing techniques have been developed, and to the extent possible these will be used to generate high-level data products in the KiDS database. Many of the parameters developed for the SDSS survey will be provided. Furthermore we will look at including:

- Images with Gaussianized PSF. Versions of all images convolved with kernels chosen to result in a homogenized and Gaussianized PSF, to ease comparison with images taken at different times or with different filters or instruments.
- Aperture-matched colour catalogues. Catalogues with colours measured only from the high S/N inner regions of sources, for applications that only require flux ratios, rather than total fluxes.
- Unsharp masked images. A wealth of underlying galaxy structure can be obtained from images in which the low frequencies have been removed (dust structures, disks, etc.). We plan to provide images filtered in various ways.
- Morphological parameters. Popular galaxy profile fitting programmes have been implemented in AstroWISE, and will be run on the sources and published in the KiDS database.

3 Internet-based survey collaboration for a geographically distributed team

The KiDS survey team is an international collaboration with ~ 35 team members at institutes spread around Europe and beyond. The team uses Astro-WISE for KiDS survey handling: data processing (image calibration, stacking, cataloging), data quality control and data management (on-line archiving and publishing in the Virtual Observatory). Team members log on to a single system, Astro-WISE to do the survey work irrespective of where they are physically located. A team member can make use of a distributed pool of storage, compute and database resources spread over Europe, irrespective of where this person is. Day-to-day survey handling is done via webservices. Thus, a web browser with internet connection is what is required to start doing KiDS survey work.

This 'whoever, wherever' approach is possible because all aspects of survey handling in Astro-WISE are implemented from a data-centric viewpoint. For example, calibration scientists provide information to the data directly, not to other users. Calibration scientists transfer their insight on time-validity of a calibration to the data item itself: the information becomes part of the data item itself. Thus, calibration control is implemented as science survey data finding out which calibration data should be applied to them. Requirements on quality control are about data knowing themselves which quality they have. Human and automatic quality assessors supply to data items their verdict on its quality. In other words, human insight on quality accumulates in Astro-WISE via the data items. Data management is implemented as data knowing itself which users it can reach. Processing is implemented as data reaching compute clusters. Not as users reaching compute clusters. All survey data can access all compute clusters/databases/dataservers pooled by the KiDS team.

Moreover, a survey data item also knows itself how it was made in the past and also how it could be made better using improved calibrations, present now. This is possible because all survey processing operations are implemented as actions by data objects acting upon itself and/or other data objects. Any type of survey product, from raw to final, is represented by a class of data objects. Survey products are framed as survey objects: informational entities consisting of pixel and/or metadata. Metadata is defined here as *all* non pixel data. The survey objects carrying the information of final survey products also carry the information on how they can be created out of intermediate survey objects. This backward chaining procedure is recursively implemented up to the raw data (see Figure 3). Thus, a request by a KiDS team member for a survey product, a target, triggers a backward information flow, in the direction of the raw data. The net effect is a forward work flow description, towards the target, that is then executed. The backward information flow is implemented as queries to database initiated by the requested target itself. The database is queried for survey objects on which the target depends with the right characteristics including validity and quality. Either they exist and are returned or the query is 'backwarded' to the next level survey objects closer to the raw survey data. In conclusion, in Astro-WISE survey handling is realized by backward information flows that results in a forward processing flow. A single copy of each data item is shared by the KiDS team. This data pool is physically distributed. Only for performance reasons does the system create temporary local pixeldata copies, as needed. See the WISE paper in this special issue for more information on Astro-WISE itself.

In conclusion, Astro-WISE takes a data-centric approach to survey handling and control. Attributes of data objects solely determine which calibration data are applied to which science data, know which survey products have been qualified and know which products should be considered experimental and which ones baseline. Table 2 lists the complete set of attributes used for survey handling and control.

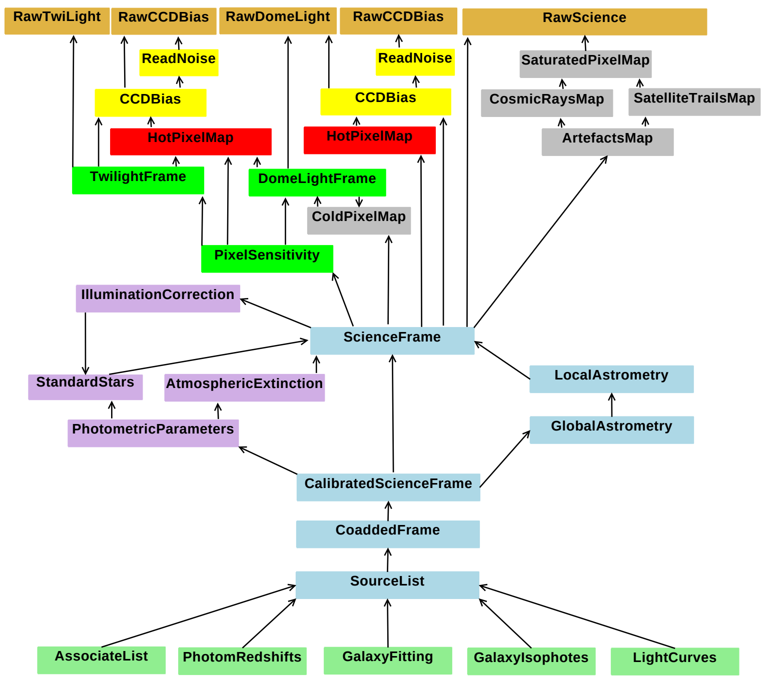

Fig. 3 Each box in this target diagram represents a class of survey objects. Survey objects not only contain the survey products denoted by familiar names in wide-field imaging. They also carry the information how they, as requested target, can be created out of other survey objects, illustrated by the arrows. Underlying is an object model that captures the relationship between requested information and the physics of the atmosphere-to-detector observational system.

3.1 Managing the KiDS survey data

A KiDS team member starts survey handling by filtering the pool of survey data on which the survey handling should act. This is called setting a 'Context' in Astro-WISE. Table 3 lists the three configuration parameters that define Context. Upon logging into Astro-WISE, the user parameter is automatically set to the numerical value for this Astro-WISE user. Then the team member selects the project to work in. All results will be part of this project. The team member selects also a minimum privileges level. KiDS survey data are managed

Table 2 Set of data-item attributes used for data management, calibration control, quality assessment and control in Astro-WISE.

attribute	content description	possible values
_creator	Astro-WISE user that created the data object	any Astro-WISE user name
_project	project data object belongs to	any Astro-WISE project name
_privileges	operational level at which data object resides	1,2,3,4,5
is_valid	validity indicator set by user	0(bad),1(no verdict),2(good)
quality_flags	quality flag set by system	any integer (bitwise), with 0==good
timestamp_start timestamp_end	start and end of validiy range in time for a calibration object	
creation_date	time of creation of data object	

by having 5 different survey operational levels named privileges levels, see Table 4 for an overview. Astro-WISE's data-centric viewpoint leads to the term privileges. The object has increasing privileges to access users with numerical increase of the privileges level. The _privileges attribute of a data object tells to which level it belongs.

The baseline KiDS survey products reside at privileges level 2 (named PROJECT). These data can be accessed only by KiDS survey team members. Each KiDS team member can experiment in her/his own privileges level 1 (MYDB) to create improved versions of these baseline products. Only the single team member can access survey data at MYDB ad promote it to the PROJECT level. A KiDS project manager can promote baseline survey data from privileges level 2 to all higher levels. Survey data at privileges level 3 (ASTRO-WISE) can be accessed by all Astro-WISE users. At 4 (privileges level WORLD) the data become accessible additionally to the astronomical community without an Astro-WISE account (anonymous users). Finally at 5 (privileges level VO), the data are accessible also from the Virtual Observatory. It is the combination of Context parameters user, project and privileges level that determine what data is filtered to be accessible for survey handling. This achieved by comparing the context parameters to their counterpart attributes of data objects (see Table 2). Figure 3.1 shows schematic diagrams of data pools after filtering. Additional customization of the data input pool is implemented as additional query constraints (for examples see Table 5).

The three data attributes involved in Context establish the data pool on which any survey handling acts. Together with the other attributes in Table 2 they allow quality control and calibration control operations on the survey data. We first give a more detailed description of the attributes before describing quality and calibration control in Astro-WISE.

1. **Creator.** Creator is the Astro-WISE user that created the data object. An Astro-WISE user is a person with an Astro-WISE database account consisting of an id number and name (e.g., AWJJOPLIN). Each person has only one account and therefore a single identity within the Astro-WISE system. Each data object is associated with a single Astro-WISE user, the data object creator. This is the Astro-WISE user that created/ingested the

data. Once established the creator of a data object cannot be changed. At the awe-prompt the creator of data object `exampleObject` can be printed to the screen with

```
awe> exampleObject._creator # Returns user id
990
awe> from common.database.Database import Database
awe> Database.users[exampleObject._creator] # Returns user name
'AWJJOPLIN'
```

2. **Project.** A project in Astro-WISE is a set of Astro-WISE users that want to share a set of data objects. A project has a project id, a name, a description, project members and optionally an instrument. One or more Astro-WISE users can be member of a project, and an Astro-WISE user can be member of more than one project. Each data object belongs to one and only one project. The project to which a data object belongs is chosen upon the creation/ingestion of the data entity and can not be changed after that. Some projects have all Astro-WISE users as members: these are called public projects. Some projects have a subset of Astro-WISE users as members: these are called private projects. At the awe-prompt the project to which data object `exampleObject` belongs can be printed to the screen with

```
awe> exampleObject._project #Returns project id
14
awe> from common.database.Database import Database
awe> Database.projects[exampleObject._project] # Returns project name
'KIDS'
```

An overview of all projects in the database, their members etc. can be found on the Astro-WISE webpages at the following address: http://process.astro-wise.org/Projects.
3. **Privileges.** Survey data management is facilitated by having pools of data at five different levels named privileges levels. Astro-WISE's data-centric viewpoint leads to the term privileges. The survey object has increasing privileges to access users with numerical increase of the privileges level. The more privileged data is, the more Astro-WISE users it can reach. The five privileges levels are listed in Table 4. The initial privileges of a data object are set by the creator upon the creation/ingestion of the data entity. They can be changed after that by creator and project managers. This is called publishing. The publishing webservice is at http://pubs.astro-wise.org/. At the awe-prompt the privileges of data object `exampleObject` can be printed to the screen with

```
awe> exampleObject._privileges
1
awe> context.publish(exampleObject,privileges=2,commit=1)
....
awe> exampleObject._privileges
2
```

4. **is_valid.** This data attribute is the validity indicator as set by users. It stores the quality assessment performed by a survey team member. Its default value upon creation of a data item is is_valid= 1, meaning no user assessment has taken place. The team member can change this to is_valid= 0, meaning bad quality, or to is_valid= 2, meaning data is qualified to be good. The calibration control webservice (http://calts.astro-wise.org/) and quality-WISE webinterface (part of the Database Viewer (DBViewer) at http://dbview.astro-wise.org/) allow changing is_valid values. At the command-line one uses:

```
awe> thisobject.is_valid
1
awe> context.update_is_valid(thisobject,0)
awe> thisobject.is_valid
0
```

5. **quality_flags.** This data attribute collects the quality flags as set by the system. Automatically the quality of objects (i.e., survey products) is verified upon creation. If the quality is compromised the quality_flags are set to a value $\neq 0$. It is a bitwise flag. Each type of object (raw science frames, calibrated science frames, astrometric solutions) has its own definition of what each bit means. Both DBViewer and its quality-WISE component facilitate easy access to quality_flag information for survey control. At the command-line:

```
awe> astrometricParametersObject.quality_flags
10
awe> astrometricParametersObject.get_qcflags_set_dict()
{'NREF_TOO_HIGH': (1, 'There are too many reference stars.'),
'RMS_TOO_HIGH': (3, 'RMS too high.')}
```

6. **timestamp_start, timestamp_end.** These attributes of a calibration data item define the range in time for which it is applicable. Calibration data get default timestamp ranges upon creation. These can be modified by survey team members. For example, it might be decided that a zeropoint might apply to one or more nights, or just a few hours instead. The calibration control webservice (CalTS, http://calts.astro-wise.org/) provides an interface for this.
7. **creation_date.** Upon creation every object that results from processing has an attribute that stores the moment of its creation. This information is relevant as 'Newer is better' is a general rule for objects to determine which calibration data item should be applied to them from the pool of applicable calibration data items.

3.2 Quality assessment and control

Three types of quality assessment are facilitated in Astro-WISE:

Table 3 Context configuration parameters

Context configuration parameter	Commands
user	automatic upon Astro-WISE log in
project	context.set_project() print context.get_current_project()
privileges	context.set_privileges() print context.get_current_privileges()

Table 4 Sharing KiDS data in Astro-WISE: privilege levels

privileges level: name	data is shared with
1: MYDB	only the creator
2: PROJECT	every member of the project to which the data object belongs
3: ASTRO-WISE	all Astro-WISE users
4: WORLD	the world: Astro-WISE users and persons without an Astro-WISE account (latter via webservice DBViewer)
5: VO	the whole world and data also available through the Virtual Observatory webservices

Table 5 Query limitations to context subsets

purpose	query postfix at awe-python prompt
limit query to data created by user only	query.user_only()
limit query to data from project only	project_only()
limit query to supplied privileges level only	privileges_only(1/2/3/4/5)
limit query to myDB data from current project only	query.current_mydb_only()
limit query to public(privileges¿=3) data or private data(privileges¡=2)	query.public_data_only(self,public=True)
limit the query to public,private or customized set of projects	query.public_project_only(self, public=True, ids=[])
limit the query to public,private or customized set of projects	query.public_project_only(self, public=True/False, ids=[])
favour project calibration data over data of other projects	query.project_favourite()

1. **’verify’**: automatic verification by the system upon creation of a data object
2. **’compare’**: comparison of a processing result to an earlier similar result
3. **’inspect’**: manual inspection by a human user of a data object

This approach is generic: all classes of data objects resulting from processing (i.e., ’processing targets’) have the three types of methods implemented. The actual content of the method is based on the type of data. Being applied to a data item, the verdict of the assessment is stored as a data attribute, again with generic names: either quality_flags or is_valid)(see Table 6). Additional quality control information can be stored also in a Comment object as a free string that links to the process target. Tracing back the dependence of the quality of end products on pipeline configuration and (calibration) data at all

Table 6 Types of quality assessment. The associated methods are implemented for each class of data objects with the generic naming convention listed in column 2. The verdict is stored in attributes of the data object, as listed in column 3.

assessment type	method	attribute
user inspect	processTarget.inspect()	is_valid
system verify	processTarget.verify()	quality_flags
compare	processTarget.compare()	quality_flags

processing stages is facilitated by the combination of this generic approach to quality assessment and the data lineage in Astro-WISE (see paper on data lineage in this issue for more information). This is illustrated in the QualityWISE webservice that bundles quality assessment information (verdicts, inspection figures, numbers) for human inspection (see Figure 3.2 for an impression and the QualityWISE paper in this issue for more detail). The QualityWISE page of each data item provides links to the QualityWISE page of data on which it depends. To zoom in on quality issues, the user can querying using the webservice for database querying and viewing (DBViewer) which provides links to the QualityWISE page of returned data items. Finally, at the commandline the user can customize with maximum freedom the quality assessment to particular needs. For example by using batch scripts for mass quality control with switches to an interactive mode when needed.

For quality control one must easily distinguish and compare baseline and experimental versions of survey data. The privileges levels serve to distinguish between these versions. A KiDS member experiments to improve quality of data at the MYDB level from where the member has access to KiDS data at all privileges levels. The results are only acccessible to the KiDS team member. Bad outcomes are discarded by invalidating the data. Promising outcomes can be shared with the team by publishing the survey object to the PROJECT level (see Figure 3.2). Fellow team members can then inspect the data and provide feedback (e.g., using Comments objects in Astro-WISE). The final verdict is set in the is_valid attribute (0=bad, 2=good). Upon team acceptance the survey object becomes baseline and can be published higher up, eventually for delivery to the outside world. Compromises in data quality identified only after publishing can be handled adequately. For example, all data derived using a specific calibration file can be isolated via database queries with few lines of code exploiting data lineage. Derived data can then be invalidated at all privileges levels.

3.3 Survey calibration control

Like science data, calibration data is represented as objects in Astro-WISE. In addition to quality parameters, these objects carry a creation date and editable timestamps that mark their validity period. A request for a target generates a database query that returns all good-quality survey objects with a validity period that covers the observation date. The newest good-quality calibration object is then selected applying the survey handling rule "newer is better". The KiDS calibration scientists spread over Europe collaborate using the calibration control webservice CalTS to manipulate this eclipsing of older calibrations by new ones (see Figure 3.3 and paper on user interfaces in this issue for details). They use Context to limit the survey calibration operations to a subset of calibration data. As the survey progresses the calibration scientists build up a coverage of calibration which covers time continuously. With this calibration build up subtle trends as a function of observational state (instrument configuration, telescope position, atmospheric state) become statistically significant. Investigation of such trends can be done generally using small scripts. As illustration, to inspect the relation between the error in the zeropoint and standard deviation in the pixel values of the flatfield applied on the standard star field, the calibration scientist enters at the command line:

```
awe>qpp=(PhotometricParameters.chip.name=='ESO_CCD_#73') &
... (PhotometricParameters.filter.name=='OCAM_g_SDSS')
awe> for pp in qpp:
awe> ... print pp.zeropnt.error,pp.photcat.frame.flat.imstat.stdev
```

The first line queries for all photometric parameters of a certain detector and filter. The loop that follows prints the zeropoint error and standard deviation of the applied flat. The deepening of the physical understanding of the OmegaCAM instrument and the Paranal atmosphere leads to better calibrations which eclipse the older ones. The final result is that the instrument plus atmosphere become continuously calibrated instead of establishing calibrations on a per dataset basis. This continuous calibration coverage can be pooled with all Astro-WISE users working on OmegaCAM data by publishing the data to privileges 3 or higher.

3.4 From 'quick look' to public survey delivery and advanced products

Quick look: initial calibration data, initial quality assessment code. Re-run, re-run: datalineage answers the question: changes in configuration, changes in calibration data. Then which subset needs to be re-processed and which one not. Advanced products: Galfit/phot, photometric redshifts

4 From 'quick look' to final public survey delivery and advanced products

KiDS survey operations have started 15 October 2011. With initial calibrations the KiDS survey fields are being calibrated. The calibration and quality control are leading to calibration data that improves on this initial version of the first KiDS survey data. The pipeline configuration and methods are fine-tuned. With the direct access to data lineage in Astro-WISE it becomes straightforward to re-process only the KiDS data affected by these changes.

The resulting good quality calibrated survey data and catalogs are then used for production o advanced products (e.g., photometric redshifts, galaxy morphometry, source variability analysis) also within Astro-WISE (see paper on galaxy morphometry in this issue). This way the KiDS team is moving from a 'quick-look' version of first survey products towards publishing of the complete KiDS Public Survey and advanced products. The team uses Astro-WISE as a 'live archive' that captures the accumulation of knowledge about OmegaCAM, VST and the KiDs survey data over the years.

Acknowledgements This work is financially supported by the Netherlands Research School for Astronomy (NOVA) and Target (www.rug.nl/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. Target operates under the auspices of Sensor Universe.

References

1. Abazajian, K. N., et al.: The Seventh Data Release of the Sloan Digital Sky Survey, ApJS, 182, 543 (2009)
2. Belokurov, V., et al.: The field of streams: Sagittarius and its siblings, ApJ, 642, L137 (2006)
3. Belokurov, V., et al.: Cats and Dogs, Hair and a Hero: A Quintet of New Milky Way Companions, ApJ, 654, 897 (2007)
4. Cole, S., et al.: The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications, MNRAS, 362, 505 (2005)
5. Colless et al.: The 2dF Galaxy Redshift Survey: spectra and redshifts, MNRAS, 328, 1039 (2001)
6. Cooper, M. C., Griffith, R. L., Newman, J. A., et al. The DEEP3 Galaxy Redshift Survey: the impact of environment on the size evolution of massive early-type galaxies at intermediate redshift, MNRAS, 419, 3018 (2012)
7. Eisenstein, D. et al.: Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies, ApJ, 633, 560 (2005)
8. Jain, B. & Zhang, P.: Observational tests of modified gravity, PhRvD, 78, 063503 (2008)
9. Lotz, J. M., Jonsson, P., Cox, T. J., et al. , The Major and Minor Galaxy Merger Rates at $z < 1.5$, ApJ, 742, 103 (2011)
10. Mandelbaum, R. et al.: Density profiles of galaxy groups and clusters from SDSS galaxy-galaxy weak lensing, MNRAS, 372, 758 (2006)
11. Mellier, Y.: Probing the Universe with Weak Lensing, ARA&A, 37, 127 (1999)
12. Man, A. W. S., Toft, S., Zirm, A. W., Wuyts, S., & van der Wel, A. The Pair Fraction of Massive Galaxies at $0 \leq z \leq 3$, ApJ, 744, 85 (2012)
13. Parker, L. C., Hoekstra, H., Hudson, M. J., van Waerbeke, L. & Mellier, Y.: The Masses and Shapes of Dark Matter Halos from Galaxy-Galaxy Lensing in the CFHT Legacy Survey, ApJ, 669, 21 (2007)

14. Peacock, J. A., et al.: ESA-ESO Working Group on Fundamental Cosmology, ESO/ESA Working Group 3 Report (2006)
15. Spergel, D. N., et al.: Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology, ApJS, 170, 377 (2007)
16. Valentijn, E. A., et al.: Astro-WISE: Chaining to the Universe, ASPC, 376, 491 (2007)

Fig. 4 Schematic representations of the data access scope for five different settings of Context (listed on top of each diagram). In all Context settings the Astro-WISE user is set to 'User X' and the project in Context to 'Project A'. A separate diagram is shown for each of the five minimim privileges levels that User X can set for the Context. The columns on the diagram are divided over three projects. Per project there is one column for each project member. Such a column represents the data objects which belong to the project and which are created by that project member. The data objects can have 5 different privileges levels which are separated in rows. Thus each box in the diagram represents data objects which belong to the same project, have the same creator and the same privileges level (i.e., have the same values for attributes _project, _privileges and _creator. Green colored boxes indicate the data access scope of User X for the chosen Context settings: the data objects to which User X has access.

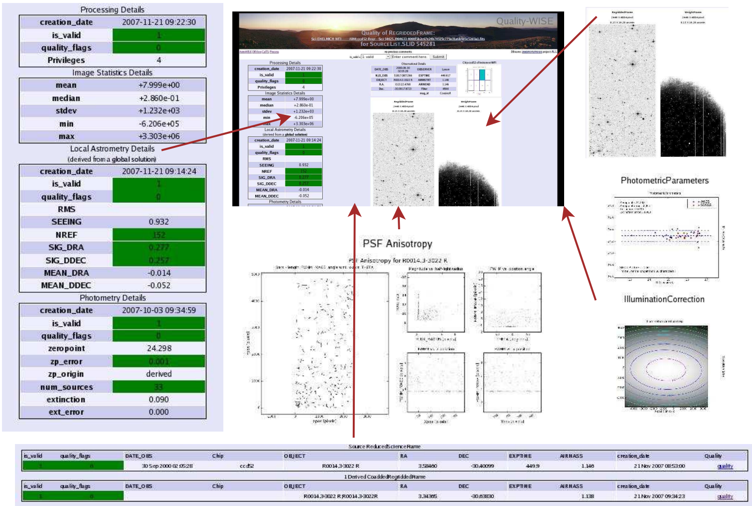

Fig. 5 The QualityWISE webservice bundles quality information for a single data object. A example page is shown partially. The arrows link to zoom ins. See QualityWISE paper in this issue for full discussion of this webservice and quality control in Astro-WISE in general.

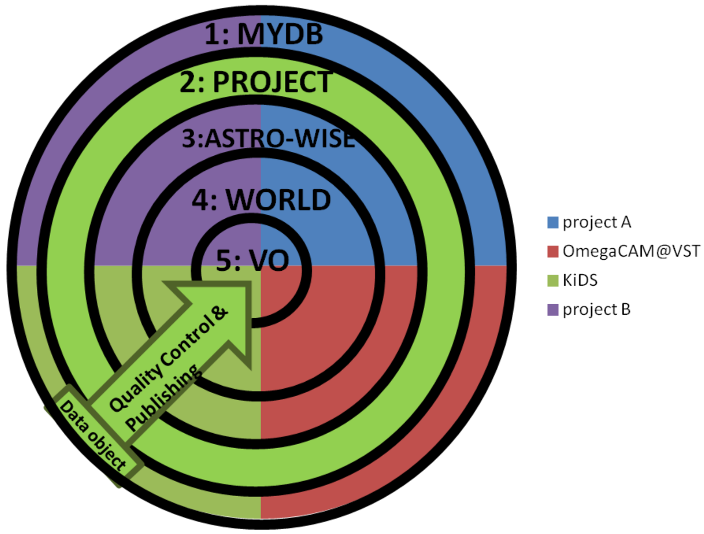

Fig. 6 This diagram shows the survey operational level at which survey handling is done. This is configured by setting the minimum privileges level in the Astro-WISE context. Each annulus represents a privileges level. The operational level includes the annulus and all levels interior to it. Each color represents a project (KiDS in denoted in green). The baseline KiDS survey products reside at 2:PROJECT. These data can be accessed only by KiDS survey team members. Each KiDS team member can experiment in her/his own level 1:MYDB to create improved versions of these baseline products. Survey data at 1:MYDB is only accessible by the single team member. If content, the member promotes the products to 2:PROJECT to share them with the team. The KiDS project manager can publish baseline survey data from 2 to levels 3 to 5. Survey data at 3:ASTRO-WISE can be accessed by all Astro-WISE users. At 4:WORLD, the data become accessible additionally to the astronomical community without an Astro-WISE account (anonymous users). At 5:VO, the data are accessible also from the Virtual Observatory.

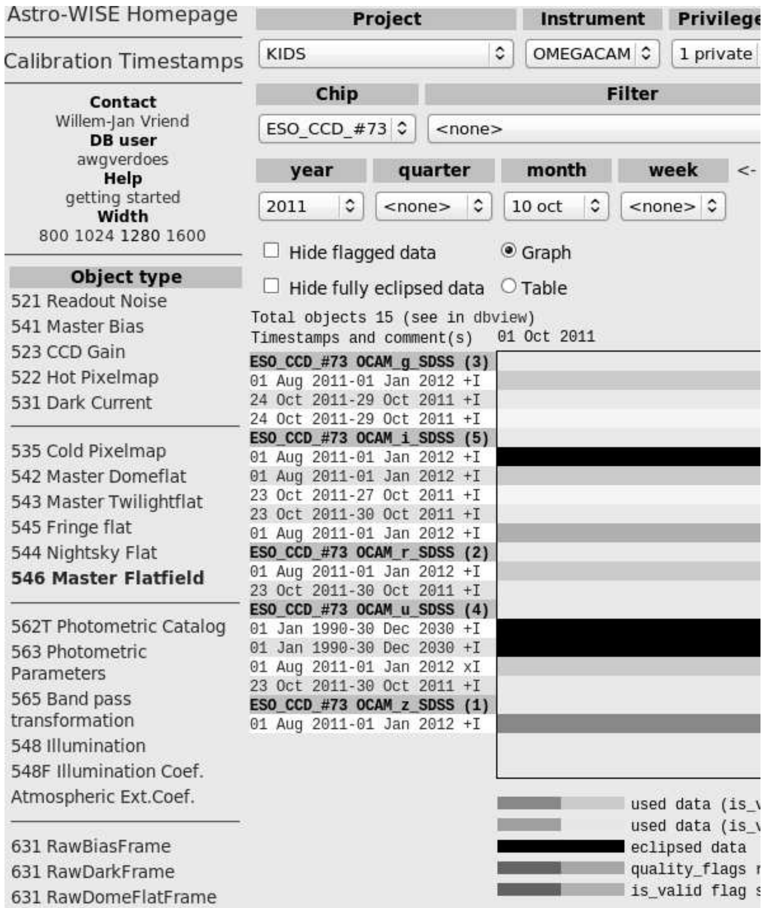

Fig. 7 Snapshot of the Calibration Time Stamp editor webservice (CalTS). All types of calibration can be selected and information on them graphically depicted. The horizontal bars show the time range validity of an object. The vertical stacking of the bars is ordered by creation date newer ones (green) eclipse older versions of calibrations (black). Other color codings depict the quality assessment verdicts. The validity time ranges and quality assessment parameters can be adjusted. See paper on User Interfaces in this issue for technical details.