Weak Lensing Data processing for the CFHTLenS and KiDS surveys

Thomas Erben, Argelander-Institut für Astronomie

LSST school & workshop - June, 15th 2017

Ground-based Wide-Field Weak Lensing Surveys (I)

During the past 15 years we developed in Bonn the necessary infrastructure (Hardware and Software) to routinely perform lensing and multi-band studies on Wide-Field Imaging Surveys:

- 2000 The VIMOS imaging survey ($\approx$ 1 sq. deg. one-band imaging survey for one of the first detections of Cosmic-Shear)
- 2001–2005 First version of the image processing pipeline THELI to routinely go from raw sceince optical images to lensing quality data prodcuts
- 2003–2007 The Garching-Bonn-Deep Survey (inhomogeneous archival data of $\approx$ 20 sq. degrees with WFI@ESO2.2m).
- 2007–2009 The CFHTLS-Archive-Research Survey (extensions to multi-band data; later merged with the CFHTLenS effort)
- 2009–2013 CFHTLenS ($\approx$150 sq. degrees five-band with MEGAPRIME@CFHT)

Ground-based Wide-Field Weak Lensing Surveys (II)

- 2010–2017 NGVS (≈104 sq. degrees five-band survey with MEGAPRIME@CFHT around Virgo)
- 2011–2016 Stripe 82 CFHT (≈170 sq. degrees *i*-band survey with MEGAPRIME@CFHT)
- 2012–2016 RCSLenS (≈800 sq. degrees four survey with MEGAPRIME@CFHT)
- 2012–2019(?) KiDS/VIKING (≈1500 sq. degrees nine-band survey with OMEGACAM@VST and VISTA)
- 2017–2020(?) CFIS (≈5000 sq. degrees *r* and *u*-band survey with MEGAPRIME@CFHT)

You cannot evaluate data products without doing science!

The image processing pipeline THELI

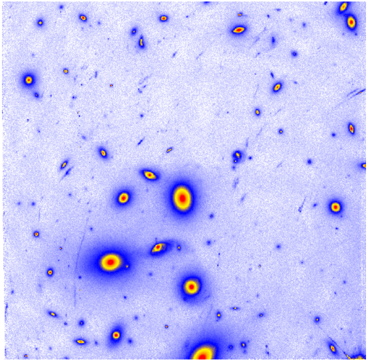
- performs all steps from raw data to imaging pixel products for multi-chip instruments
- multi-instrument processing pipeline
- fully parallelised processing
- processing with `bash`-shell scripts gluing together command line tools (remote processing)
- strongly builds on state-of-the-art software modules (Emmanuel Bertins software suite (esp. `SExtractor`, `scamp` and `swarp`), `(FLIPS)`; `eclipse`; `imcat`)
- new programs are developed in `python` (previously in `C`).
- modules can be exchanged or added easily
- Can be used by PhD students with a few days of training
- GUI version supporting more than 70(!) instruments (Mischa Schirmer)

Our current hardware

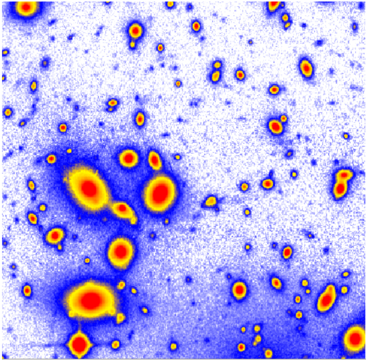
- Old setup: Cluster with 32 nodes
- 1–2 64-core machines, 512GB of memory. Suitable to work on all multi-chip instruments we currently use.
- ultra fast 20 TB of Solid state disk space for pixel processing. Can process simutaneously 62 chips of DECam with 100% processor load! I/O is not a limiting factor!
- $\approx$ 400 TB of HDD space for data storage.

The Challenge - Measuring Object Parameters

space-based



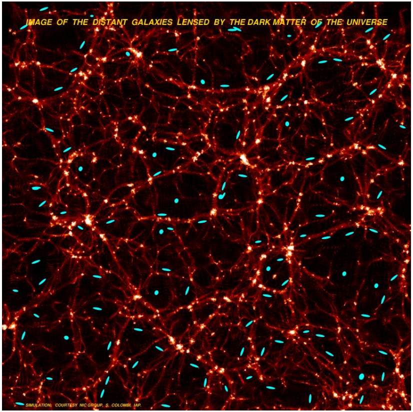
ground-based; 0.6” seeing

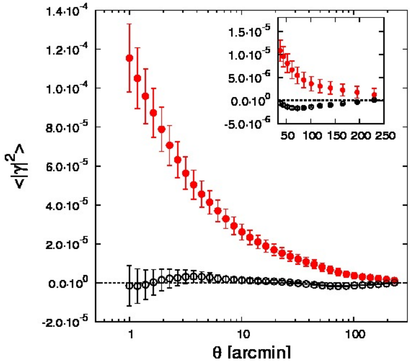


It is amazing that we can do lensing from the ground!

Weak Lensing and the Large Scale Structure

Gravitational Lensing (shear studies) directly traces mass!

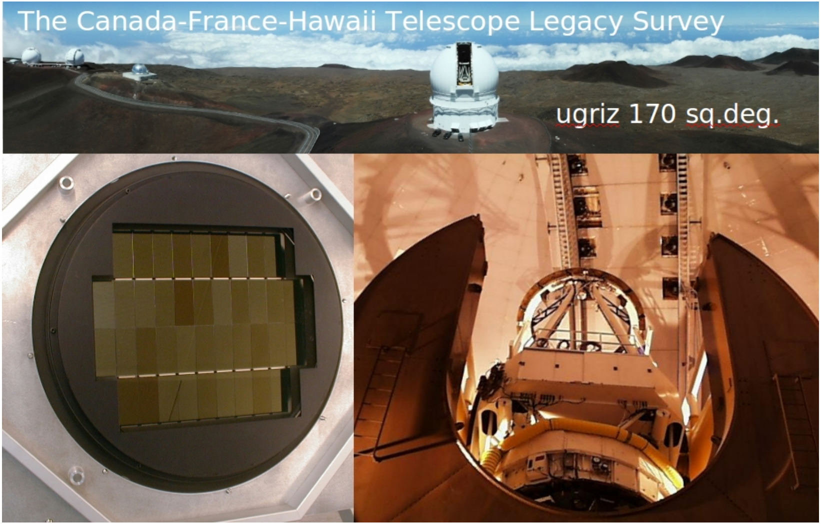




Tophat Shear from Fu et al. (2008)

State-of-the art with one-band surveys around 2008.

The CFHTLS

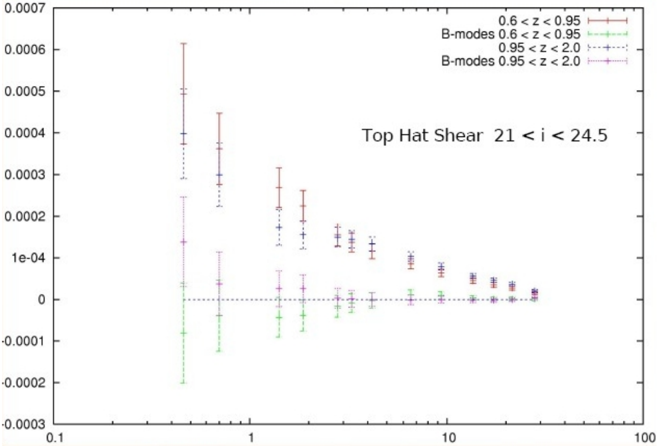


The CFHTLS was a unique lensing survey at the time: Wide Area, superb image quality, optical five colour coverage for photometric redshifts

The CFHTLS - Steps forward from old analyses

- Photometric Redshifts:
  Redshifts were no longer a free parameter that you could play with (mean redshift, tomography)
- Area and Data Quality:
  The field-to-field variation from individual pointings is expected to be much smaller than from previous surveys.
- CFHTLS-Deep:
  Many exposures of the same field with varying conditions.
  You can (and need to) test your methods.
  Analyses of the same objects from different stacks often yielded very different results.

First steps towards tomography

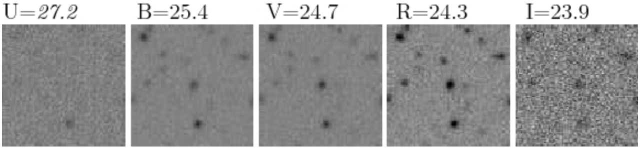


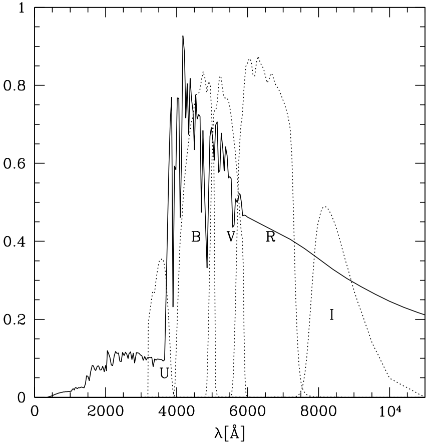
Joel Benjamin (unpublished)

The cosmic shear signal should scale as $\langle\gamma^2\rangle \propto z^2$

The CFHTLS revealed that many *details* do not yet work! To resolve them was the task of the CFHTLenS team!

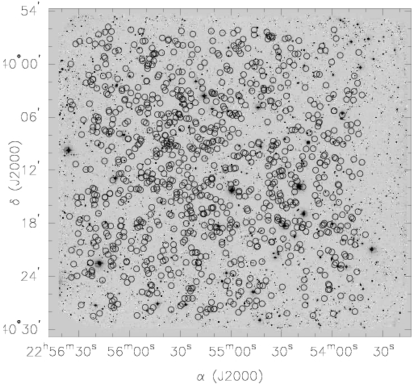
Scientific Project: Lyman Break Galaxies

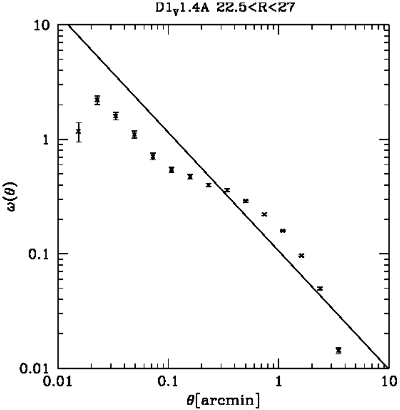




Scientific Project: Lyman Break Galaxies

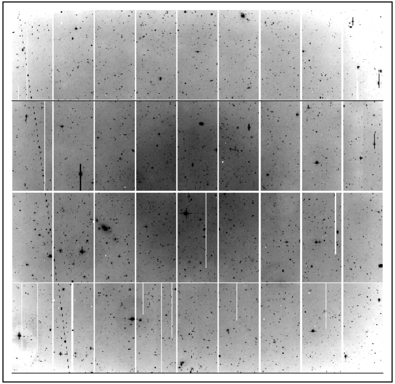
Correlation function of 16000 u^* dropout candidates (first try):

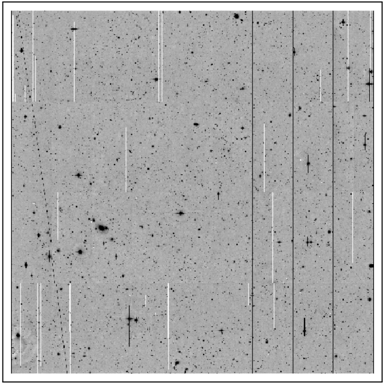




Detrended images

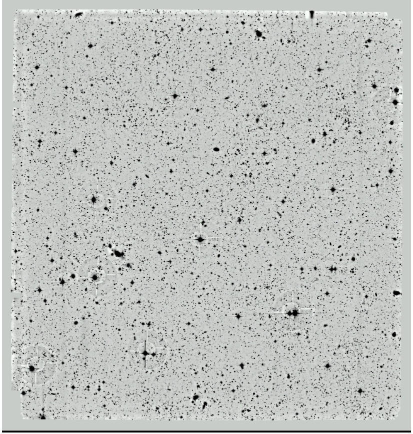
A detrended i-band image: The photometric homogenization leads to a non-uniform background.

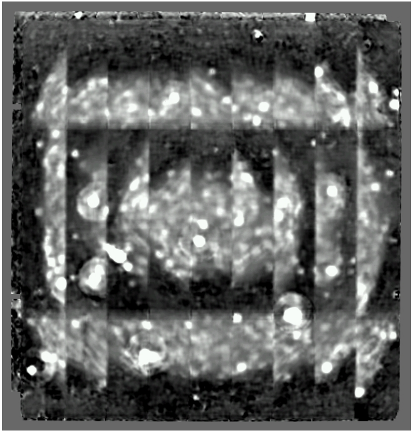




Scientific Project: Lyman Break Galaxies

One- and two-pass sky-background subtraction:

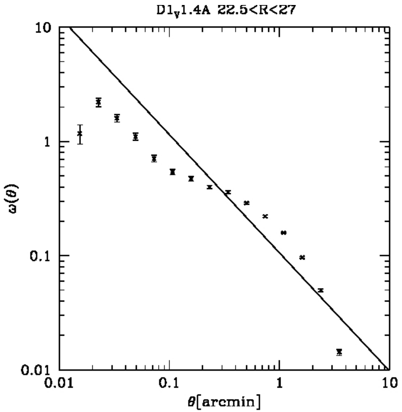


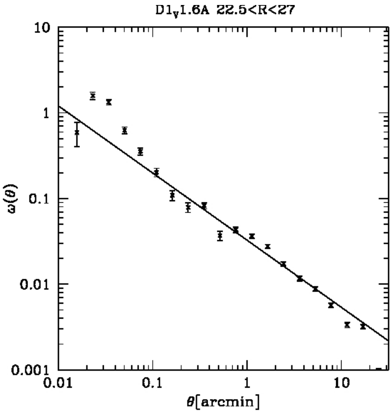


Differences occur because of the non-flat preprocessed images (illumination correction).
The amplitude of the differences are on the 1σ noise level of the original image.

Scientific Project: Lyman Break Galaxies

Correlation function of 16000 u^* dropout candidates with two- and one-pass sky-background subtraction:





The background fluctuation (together with revised masks) lead to significantly different results !!!

Data products within the LenS collaboration

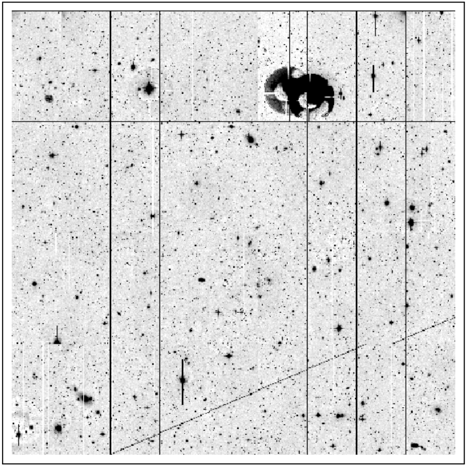
For each survey we produce:

- co-added and calibrated images for all fields and filters
- sky-subtracted single frames with astrometric and photometric information
- image masks to exclude areas from science analysis (bright stars, stellar haloes etc.); masks are verified manually
- multi-colour catalogues with mask information
- photometric redshifts (BPZ)
- catalogues with photometric redshifts and lensing quantities

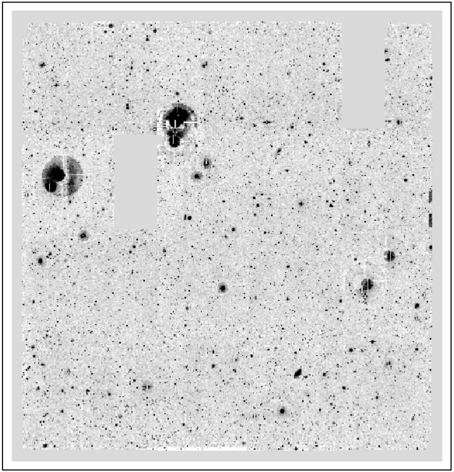
Operations are mainly done in Bonn, Vancouver, Groningen (pixel processing) and Edinburgh (catalogue production); final products are mainly checked via WWW pages and control plots

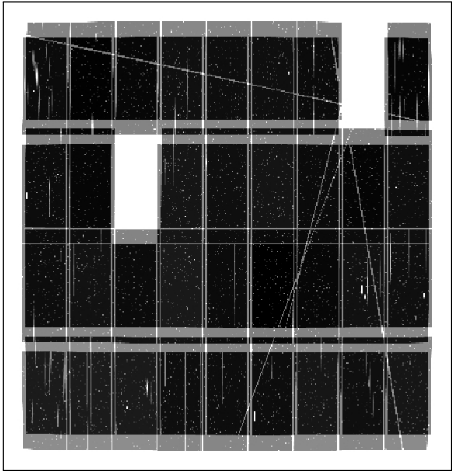
Co-added images

- We perform a *linear, weighted average* co-addition
- We identify defects on individual frames!
- Bad pixels
- Cosmic Rays (`SExtractor`, `EYE`)
- satellite tracks (THELI module developed by Patrick Hudelot)
- short asteroid tracks (masked on co-added images)



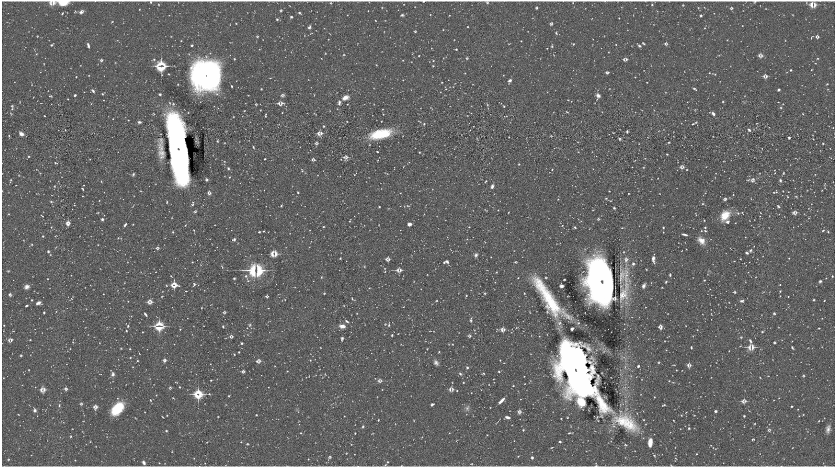
CFHTLS Wide *r*-band field with two exposures





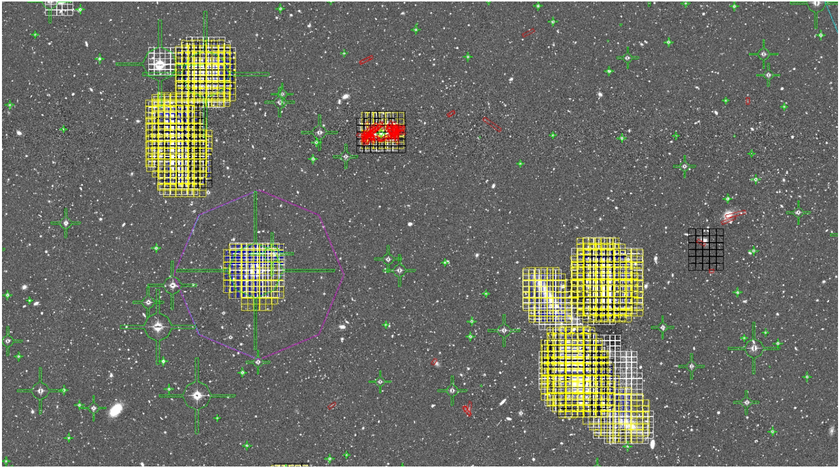
Our procedures allow us to obtain clean images with very few individual exposures; image selection on photometric quality, seeing, sky-background, stellar ellipticity.

A section from NGVSm1p1



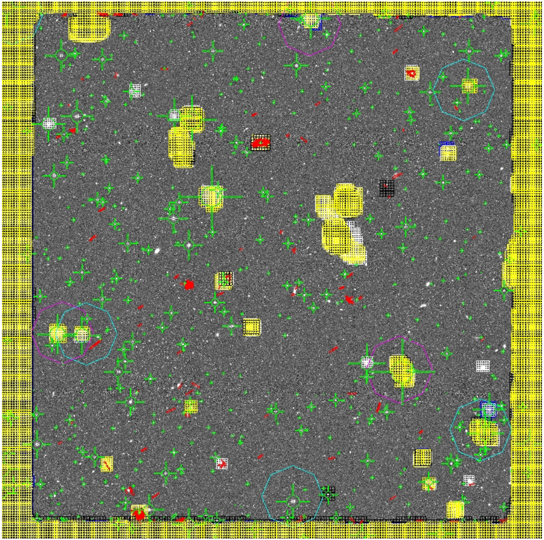
Regions around bright stars and big galaxies need to be excluded from our studies.

Semi-Automatic Masking



Moderately bright Stars are masked with template masks; large scale defects produce significant jumps in the object number density; each mask is manually adapted at the end (especially for OMEGACAM)!

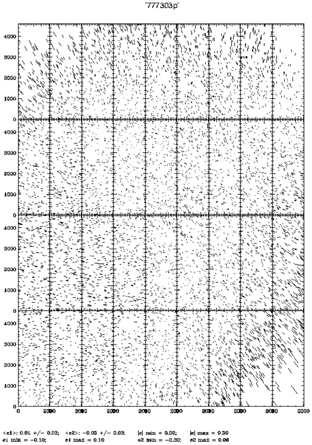
Complete Masking of NGVSm1p1

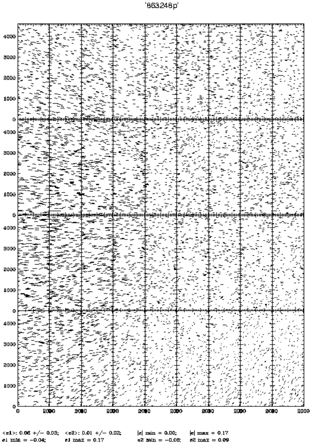


About 20%-30% of the total area are lost due to masks!

Shape measurements on individual frames

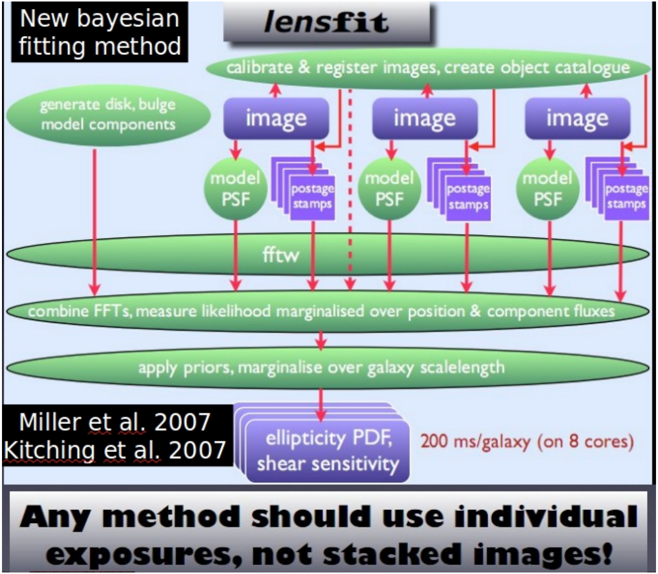
MEGAPRIME PSF anisotropy patterns:



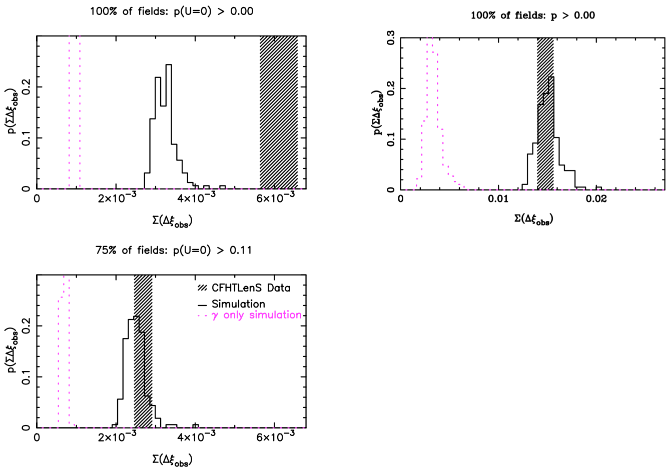


Within CFHTLenS, Lance Miller refined *lensfit* to optimally perform shape measurements on individual frames (discrete jumps of PSF, correlated noise in co-added images)!

Lensfit

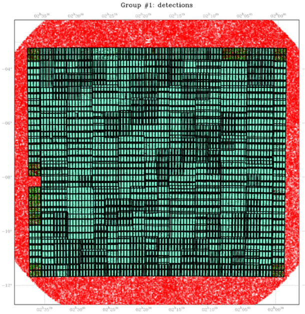

New bayesian fitting method
lensfit
calibrate & register images, create object catalogue
generate disk, bulge model components
image
image
image
model PSF
postage stamps
model PSF
postage stamps
model PSF
postage stamps
fftw
combine FFTs, measure likelihood marginalised over position & component fluxes
apply priors, marginalise over galaxy scalelength
Miller et al. 2007
Kitching et al. 2007
ellipticity PDF, shear sensitivity
200 ms/galaxy (on 8 cores)
Any method should use individual exposures, not stacked images!

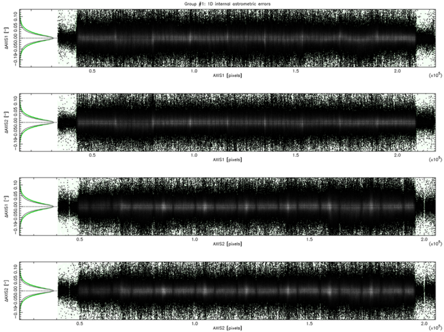
Weak Lensing field selection for CFHTLenS and KiDS



75% of CFHTLenS fields and 100% of our KiDS fields pass weak lensing quality control (better camera(?), lensfit improvements(?), CFHTLenS reanalysis is underway.)

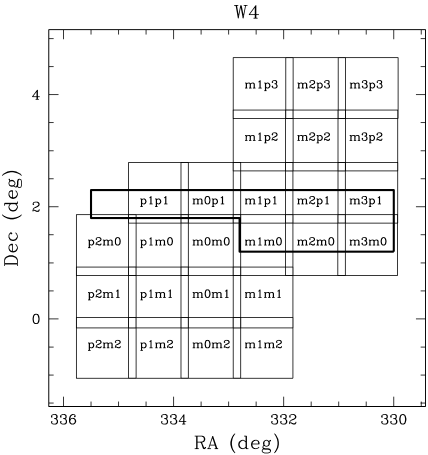
Astrometry/Photometry

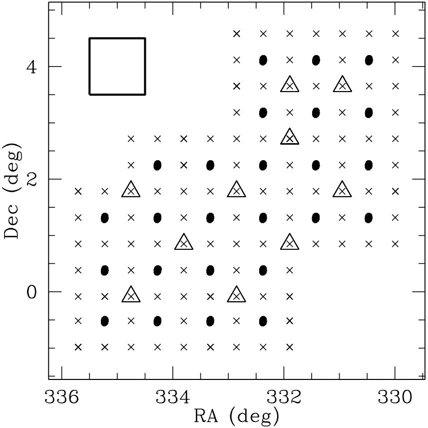




Results for photometry and astrometry are better the more data we use! CFHTLS provided the Presurvey and additional phoptometric pegs.
The internal accuracy of our astrometric solution is around 0.02 arcsec.

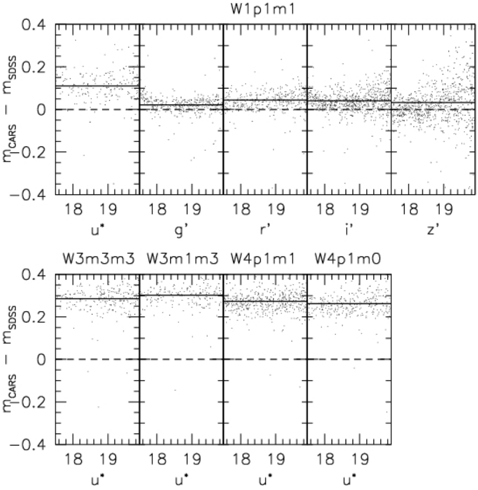
Global Astrometry





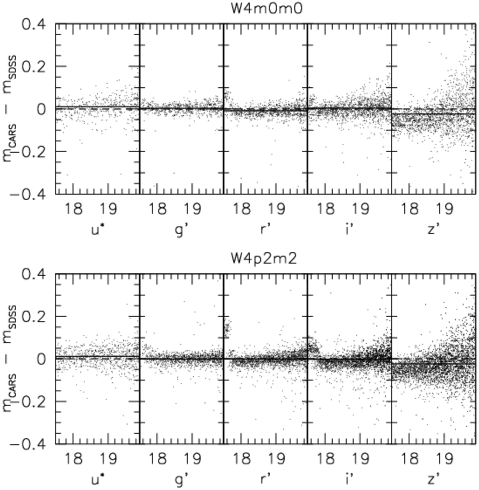
Global astrometry/photometry improves data quality remarkably. The precise effect of global astrometry is difficult to assess!

Data Evaluation: Accuracy of Phot. Calib. (CARS-I)



Internal and external checks suggest that we have a reasonable, unbiased photometric calibration in $g' r' i' z'$ and systematics in the u^*-band.

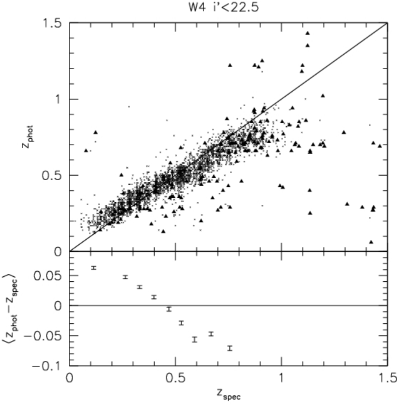
Data Evaluation: Accuracy of Phot. Calib. (all data)



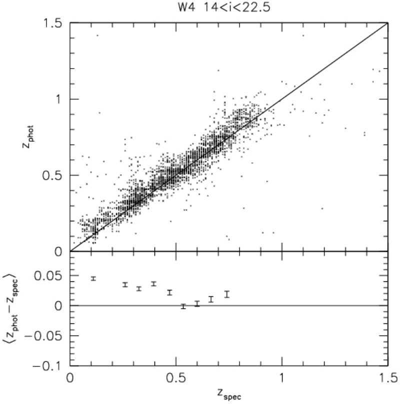
Photometric calibration became significantly better when using all available data (photometric peg images; Presurvey)

Data Evaluation: Photometric Redshifts

We can determine photo-z's with about $\sigma_{\Delta z/(1+z)} \approx 0.03 - 0.04$ and an outlier rate of about 4%.

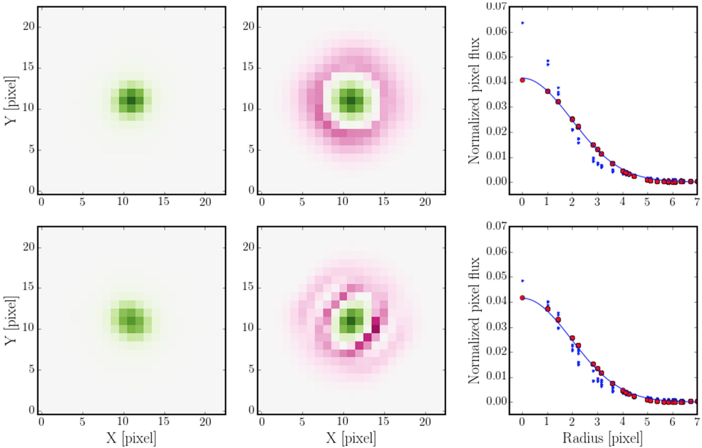


first try: photo-z with zeropoint recalibration (spectro-z)



second try: photo-z without any recalibration!

Photometry with localized PSF gaussianization



Konrad Kuijken and Hendrik Hildebrandt developed a localized PSF gaussianization procedure for CFHTLenS and KiDS.

Photo-z from gaussianised images

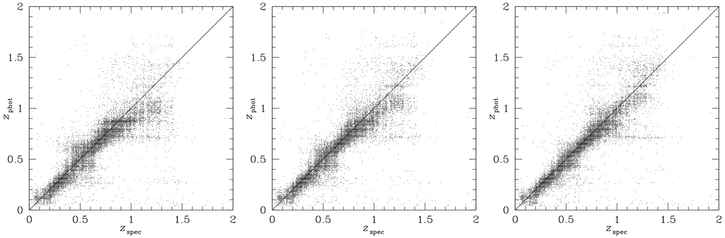
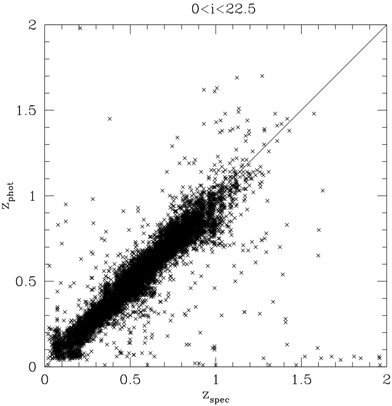
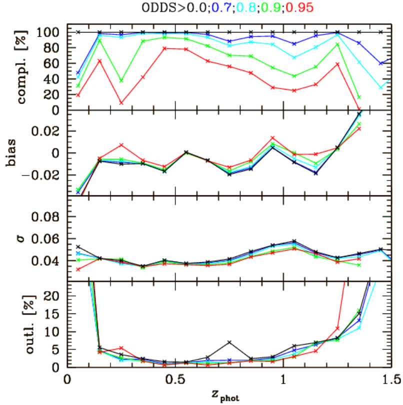


Figure fron Hildebrandt et al. (2012)

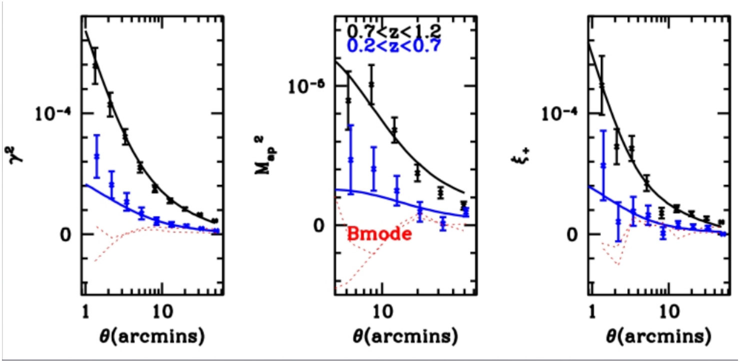
Final Photo-z accuracy for CFHTLS Wide





The result of our efforts

$\Lambda CDM : \sigma_8 = 0.88; \Omega_m = 0.28$ (no fit to the data!)



The shear tomography confirms the good quality of our shape measurements and our photometric redshifts.

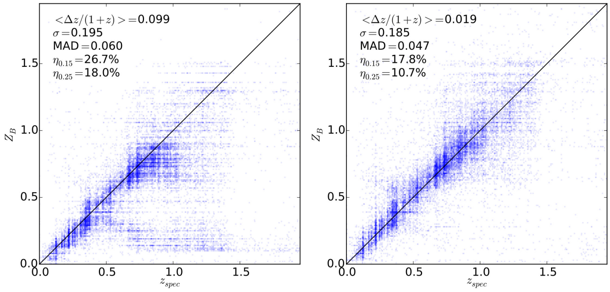
LENSING AND SHEAR STUDIES WORK!!

Achievements and Lessons Learnt by the CFHTLenS/KiDS effort

- Data Processing:
  Our methods are accurate enough to use the full potenital of current weak lensing surveys but we continuously improve them!
- Code/algorithm development/testing:
  It is best done *with* real data! The archives have plenty (CFHTLenS, KiDS, DES, (HSC), ...)!
- Slicing projects:
  The CFHT-Deep fields with hundreds of images from the same patch of the sky provide invaluable testing grounds!
- Data usage:
  Use all available data always! Precision lensing and photo-z analyses are best done on the single images.
- Team work:
  A very close collaboration between the data processing and science teams is essential!

The Next major Steps

- Update our methods to work on non pointing-based surveys (done for the pixel processing)
- Apply a second competitive shape measurement pipeline
- Extend our procedures to the infrared near-infrared domain (work led by Angus Wright) The next KiDS-analysis will include 5-band *z*, *J*, *Y*, *H*, *Ks* near-infrared coverage



Preliminary KiDS/VIKING 9-band photometric redshifts