SF2A 2006
D. Barret, F. Casoli, T. Contini, G. Lagache, A. Lecavelier, and L. Pagani (eds)

CFHT LEGACY SURVEY (CFHTLS) : A RICH DATA SET

J.-C. Cuillandre[1] and E. Bertin[2]

Abstract. The CFHTLS, a 500 nights over 5 years observing program started mid-2003 on the Canada-France-Hawaii Telescope (CFHT) using the 1 square degree optical imager MegaPrime equipped with the MegaCam camera, has entered its fourth year of data gathering. The topics covered in this article are 1) sky coverage and current data set, 2) observing process, 3) data processing at CFHT and TERAPIX, 4) data distribution and access through the Canadian Astronomy Data Centre (CADC).

1 Introduction

Wide field imaging has proven to be a valuable scientific niche for the CFHT as early as 1996 when larger telescopes (8-10m) primarily focused on spectroscopy became fully operational. While CFHT was operating a 64 Mpx camera (UH8K), and then a 100 Mpx camera (CFH12K), the next generation imager was being designed and built by the Canadian and French communities. MegaPrime is a new top-end prime focus providing a full square degree field of view for the camera it houses, MegaCam, a 340 Mpx camera built by CEA (Boulade et al. 2003), sensitive from the near ultra-violet to the near infrared. At the time this new instrumentation was being designed and built, a reflexion took place within the CFHT scientific communities on the use of this unique instrument. A condition set by CEA for building MegaCam was that a general survey servicing a large scientific community had to be executed. Through the enthusiastic response to the call for proposals, the scale of the CFHT Legacy Survey became highly ambitious, with a commitment from the Canadian and French agencies to allocate a total of 500 nights over 5 years to answer burning astrophysical questions: the nature of the Dark Energy (the Deep/SNLS survey), the dark matter power spectrum (the Wide survey), and the solar system formation history through the study of Kuiper Belt Objects (the Very Wide Survey). These specific scientific topics had to be investigated while building a data set with a true legacy value (both scientific and archival) for use by the astronomical scientific community at large.

2 Sky coverage and current data set

The CFHTLS was approved at the time CFHT was already fully operating in queued service mode (Martin et al. 2002), an operational model essential for the success of the survey components, specially the supernovae and KBOs programs with their strong time constraints. To ensure continuous observing throughout the year, four fields were chosen for both the Deep and Wide surveys, while the Very Wide survey focused on a $+/-2$ degrees wide stripe along the ecliptic.

2.1 Deep/SNLS survey

The Deep survey strategy consists in having two fields observable each night throughout the year. Each field is a single pointing observed again and again in various filters (g', r', i', z', while the u* band is exclusively acquired for the Deep part of the survey) every four nights during each of the monthly 14 to 17 nights MegaPrime observing runs. For the SNLS program, this provides a fantastic time sample of SNe Ia type throughout the year, and for the Deep program, the stacking of all these exposures brings impressive depth. Obviously, the

[1] Canada-France-Hawaii Telescope, USA & Observatoire de Paris, France
[2] Institut d'Astrophysique de Paris, France & Observatoire de Paris, France

fields were also chosen with considerations of overlaps with present and future multi-wavelengths surveys (e.g. COSMOS/ACS), as well as low dust absorption by the Galaxy and visibility from the southern hemisphere (VLT follow-ups). The following table proposes the current statistics (as of fall 2006) on the Deep/SNLS data sample. The size of the field is given in square degree, and for each filter the information proposed is: 1) approximate depth reached (in magnitude, estimated using the MegaCam exposure time calculator) for a signal-to-noise ratio of 5 (detection : photometry error = 20%) on a point source, 2) cumulated light integration (hours), 3) median image quality of all exposures (arcsec.).

Field	RA/DEC	Size	u*	g'	r'	i	z'
D1	02hr/-04deg	1×1	27.5/14/0.92	28.2/17/0.86	28.2/37/0.79	27.9/74/0.72	26.7/39/0.75
D2	10hr/+02deg	1×1	26.6/2/0.81	28.0/11/0.89	27.8/22/0.80	27.6/48/0.76	26.4/22/0.72
D3	14hr/+52deg	1×1	27.3/9/0.94	28.2/17/0.88	28.1/32/0.79	27.9/69/0.76	26.6/34/0.73
D4	22hr/-17deg	1×1	27.5/17/0.97	28.4/20/0.90	28.2/39/0.79	28.0/81/0.76	26.7/44/0.75

2.2 *Wide survey*

Just like the Deep survey, the Wide survey has several fields distributed in right ascension throughout the year, with the goal of covering a total of 170 square degrees. The survey consists of four patches ranging from 25 to 49 square degrees, integrated in all five broadband filters, each square degree pointing being integrated for: u*=1.6 hr, g'=0.7 hr, r'=0.6 hr, i'=1.2 hr, z'=2.0 hr. The depth reached for each pointing is approximately: u*=26.4 mag., g'=26.6 mag., r'=25.9 mag., i'=25.5 mag., z'=24.8 mag. The following table summarizes the status as of fall 2006. The information proposed is: 1) The size of the patch in square degrees covered in at least one filter, 2) cumulated light integration (hours) over the whole patch, 3) median image quality of all exposures (arcsec.) over the whole patch, 4) overlap with other surveys.

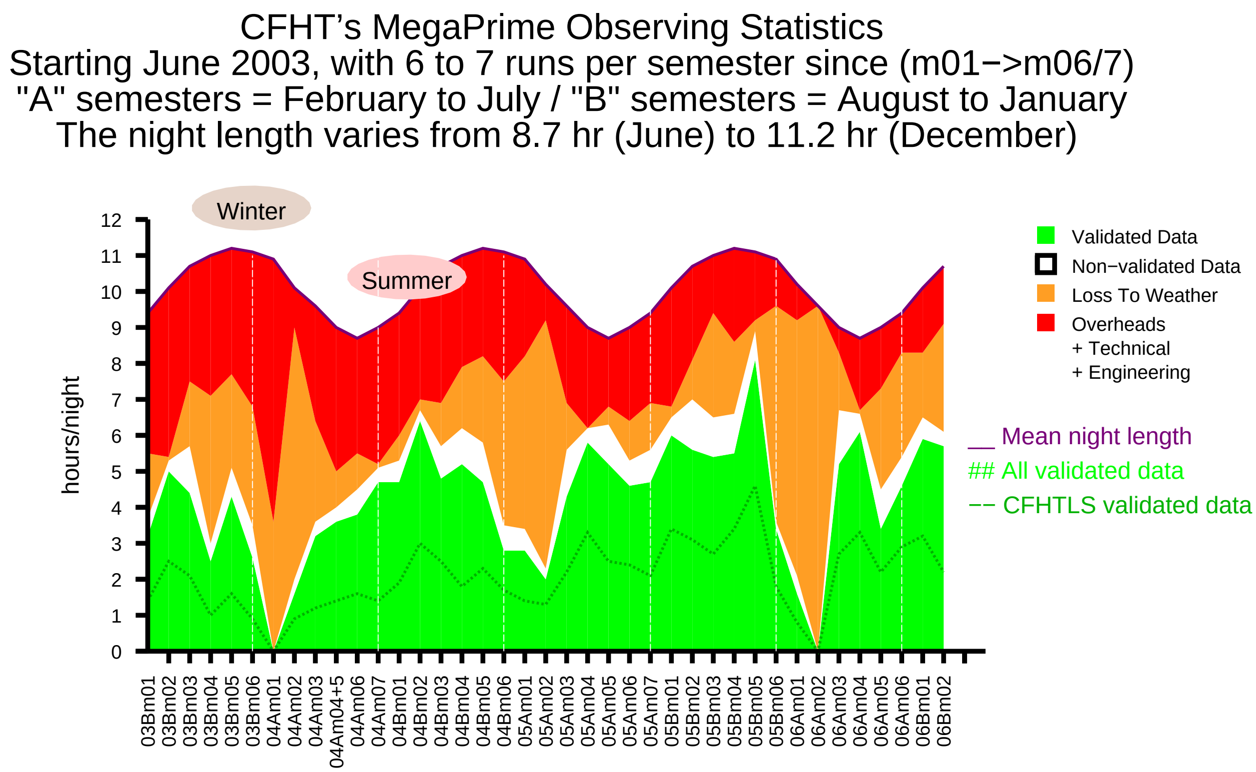
Field	RA/DEC	Size	u*	g'	r'	i	z'	Overlap
W1	02hr/-07deg	6×7	17/0.91	40/0.81	36/0.78	53/0.70	17/0.72	XMM LSS
W2	08hr/-04deg	4×5	0/0.00	15/0.83	16/0.81	39/0.67	2/0.71	
W3	14hr/+54deg	7×7	4/0.76	31/0.86	24/0.83	54/0.70	10/0.64	Groth Strip
W4	22hr/+1deg	5×5	9/0.85	14/0.81	4/0.68	23/0.68	23/0.65	VVDS & UKIDSS

2.3 *Very Wide survey*

At the time of the mid-term review, the Very Wide component was downscaled to allow the cosmology drivers, considered of higher priority, to proceed at the required pace. This implied stopping the discovery fields, but the entire recovery and confirmation observing strategy was preserved in order to validate scientifically the time already spent on this survey. The data set is nonetheless quite considerable with 410 square degrees on the ecliptic covered in a diversity of three filters: g', r', and i' to a depth of respectively 25.5, 25.0 and 24.4 magnitudes (approximately 500 seconds total per filter).

3 Observations

The first year of MegaPrime operations was affected primarily by an image quality issue which impeded the Wide survey from making progress because its science depends strongly on the point spread function being uniform across the field of view. Also, during the first two years, the bad weather during the winter on Mauna Kea gave a sense to the CFHTLS Steering Group of the importance of ranking the three individual surveys within the time allocation given to the CFHTLS (the mid-term review would eventually lead the CFHT Scientific Advisory Committee to downscale the Very Wide component in order to let the two cosmological surveys to get realized within the available sky time). The following illustrates the problem: while the CFHTLS had been designed for a validation rate of 6 hours per night to be shared typically 50–50 % with the normal "PI" programs (validation means that a given exposure fits within the observing specifications, i.e. image quality, sky background, sky transparency), the actual rate per night rarely goes above such value due to a combination of observing overheads, technical problems, and mostly bad weather conditions. The impact of the bad weather during the winters is clearly visible on the fields D2 and W2 (tables above), hence a fourth Wide field (W4, summer field) was open in 2006 to realize some needed observing time on that component.



The CFHT queue system is flexible enough to allow the coordinators of each survey to enter new groups of observations between observing runs. Real-time statistics are available immediately and strategies can be altered on a daily basis.

4 Data processing

4.1 CFHT's Elixir

Starting with the CFH12K data, CFHT has developed a data pipeline called Elixir (Magnier & Cuillandre 2002) meant to remove the instrumental signature at the pixel level (no geometrical ressampling) and to produce an astrometric calibration at the pixel level (approximately 0.2 arcsec), as well as providing a photometric calibration per run (zero points). The data processing includes the classic steps such as bias and flat-fields. But for MegaCam, the removal of the fringes in the z' band data, which reach an amplitude of 15% of the sky background, has proven to be a challenge: the recipe successfully derived for CFH12K using a deterministic measurement of the fringes leaves typically 10% of the z' frames poorly corrected. The future releases of the CFHTLS data set will use a refined method to salvage those exposures. Also, CFHT is committed to provide to its community images photometrically flat: this means convolving the flat-field with a function that does produce an homogeneous photometry across the entire field of view. These photometric flat-fields allow the stacking and direct comparisons of largely dithered fields.

4.2 TERAPIX

When MegaCam was designed, it was clear that by the time the first images would become available (2003) personal computing facilities for astronomers would still be insufficient to handle the fantastic amount of data (0.7 Gb per image). The TERAPIX data center, based at the Institut d'Astrophysique de Paris, has the mission of precisely calibrating, both photometrically and astrometrically, all the CFHTLS images, produce data stacks after proper data ressampling to remove the geometric signature of the instrument, and finally deliver catalogs and a suite of quality assessment products. TERAPIX has currently 100 Terabytes of disk online, and a 240 Gigaflops capacity on 24 nodes. TERAPIX also offers these services to the normal scientific MegaPrime programs.

At the heart of this complex pipeline, based exclusively on unique in-house developed programs, are the three main elements: 1) QualityFITS, a package that conducts a suite of quality assessments, 2) SCAMP (Bertin 2006), a package calibrating the photometry and astrometry for each image in respect of a large collection of associated images and external references (e.g. USNO-B1 catalog), and 3) SWARP, a package to ressample and co-add at once hundreds of images calibrated with SCAMP.

5 Data flow and data access

Defining, obtaining, archiving, pre-processing, calibrating, processing, extracting the sources, and distributing the CFHTLS data is a complex undertaking. With CFHT and TERAPIX, a third entity is involved in the CFHTLS global data flow that eventually brings the data to the users. The Canadian Astronomy Data Centre (CADC) handles data archiving and distribution of all CFHTLS products. While CADC releases both the raw and Elixir processed data, the typical end users of CFHTLS data will prefer the TERAPIX stacks. Elixir data from CFHT are available for download within two weeks after the end of an observing runs, whereas TERAPIX has elected to produce major CFHTLS releases once a year.

6 Conclusion

Despite difficulties during its first year, the CFHTLS appears today as a brisk success, technically, strategically, and scientifically. Each difficult step in the process of swiftly providing high quality products to the CFHTLS scientific community has been tackled through cooperative work, and major scientific results have already been published on a large variety of topics: cosmic shear, the dark energy, quasars, large scale structures, clusters of galaxies, galaxy morphology, stellar populations in the Galaxy, Kuiper belt objects, and GRBs afterglows, to name a few. The survey started mid-2003 and has accumulated some delays in the first two years, which should lead to a short extension (up to end of 2008) to complete the Wide survey. TERAPIX is currently preparing its fourth release (T004) for the Canadian and French communities, while the T0003 release from February 2006 will be the first release open to the world in February 2007 through CADC.

All informations, references, and links pertaining to the CFHTLS are proposed on the main site at CFHT: http://www.cfht.hawaii.edu/Science/CFHTLS

References

Bertin, E., 2006, ASPC, 351, 112

Boulade, O., Charlot, X., Abbon, P. et al. 2003, PASP, 4841, 72

Magnier, E., Cuillandre, J.-C. 2002, PASP, 116, 449

Martin, P., Savalle, R., Vermeulen, T., Shapiro, J. 2002, PASP, 4844, 74