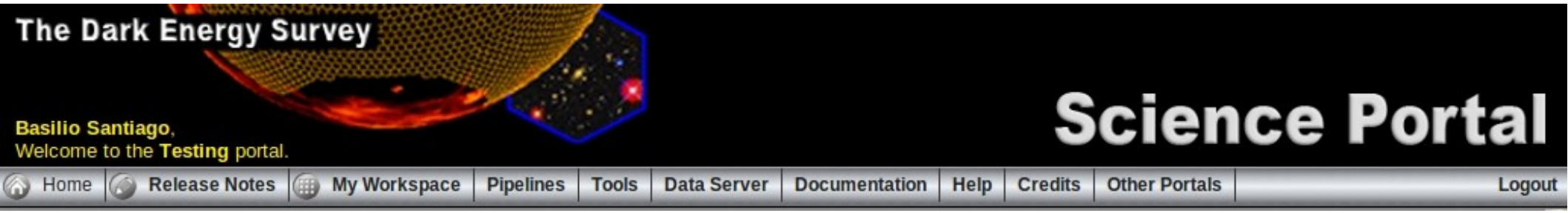
Laboratório Interinstitucional de e-Astronomia

A Quality Assurance tool for DES and beyond



Basílio Santiago

IF/UFRGS, LINEA

Brazil-LSST workshop

Campos do Jordão, March 2012

What is Quality Assurance?

- In the case of Astronomical data:

→ checking data quality in terms of photometric and astrometric precision, sample depth, object classification, image quality, etc

→ confronting quantitative data quality parameters to science requirements

Laboratório Interinstitucional de e-Astronomia

QA for the Dark Energy Survey

- Implemented in the DES Science Portal
- Implementation preceded by elaboration of scope and design documents
- Currently used to evaluate many aspects related to the quality of DES simulations: Data Challenge 6

QA Main characteristics

- transparent - availability, sharability, documentation

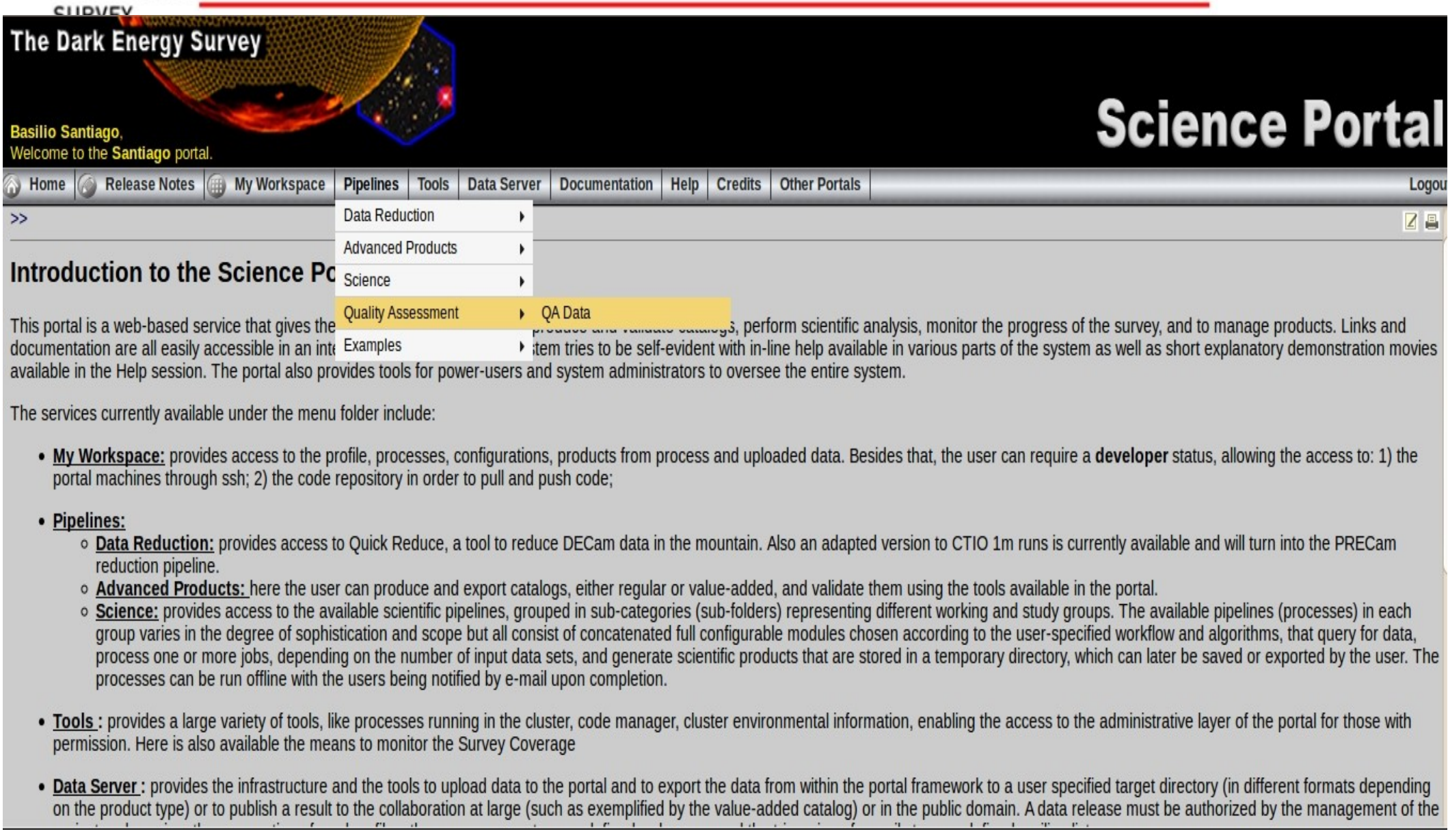
- versionable - codes, processes, people

- efficient - modular, parallelized, visual

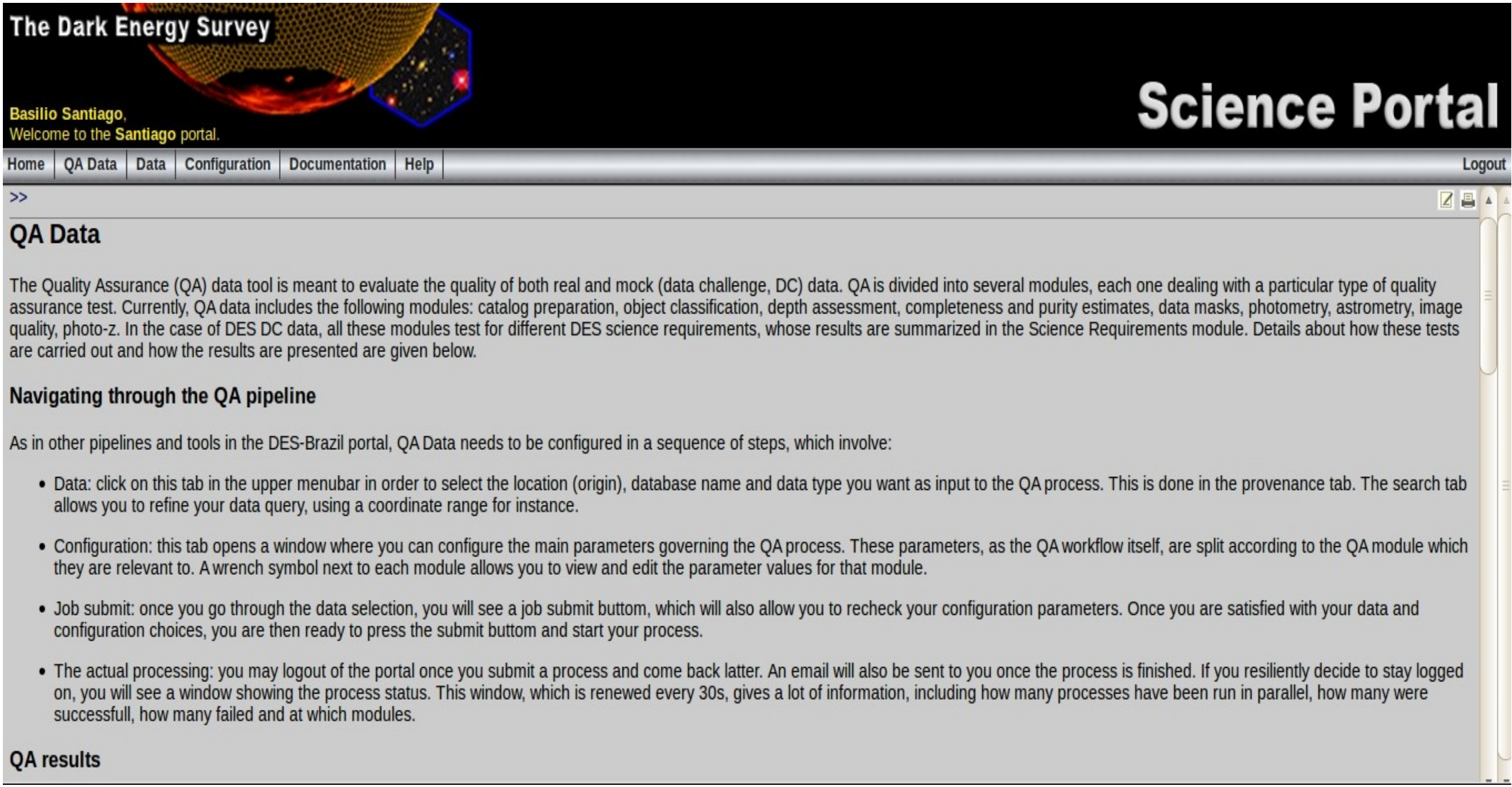
- flexible – easy to incorporate new modules, new science requirement tests, new plots and tables

- team work - contributions from many people

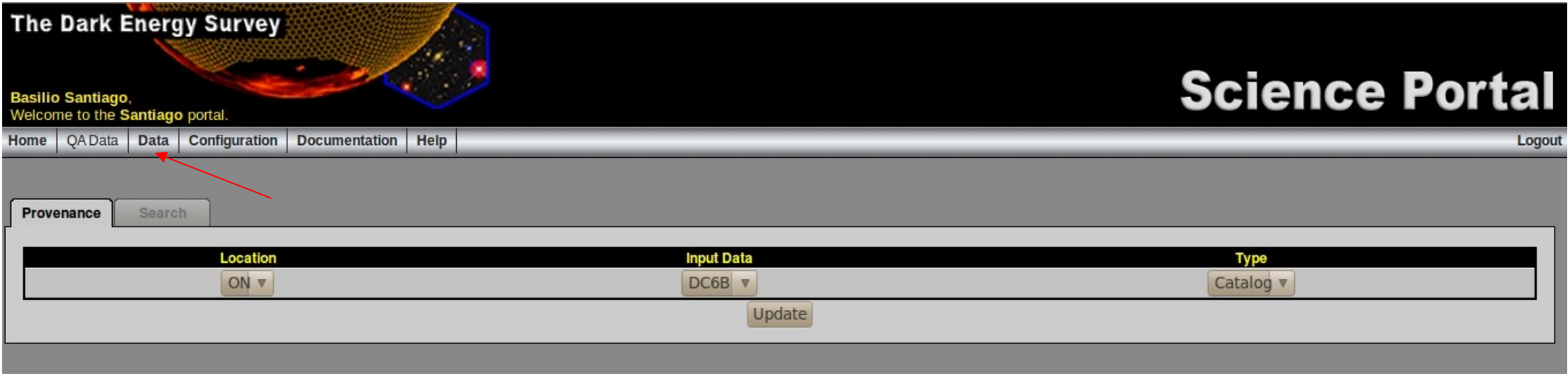
Navigating through the DES Portal

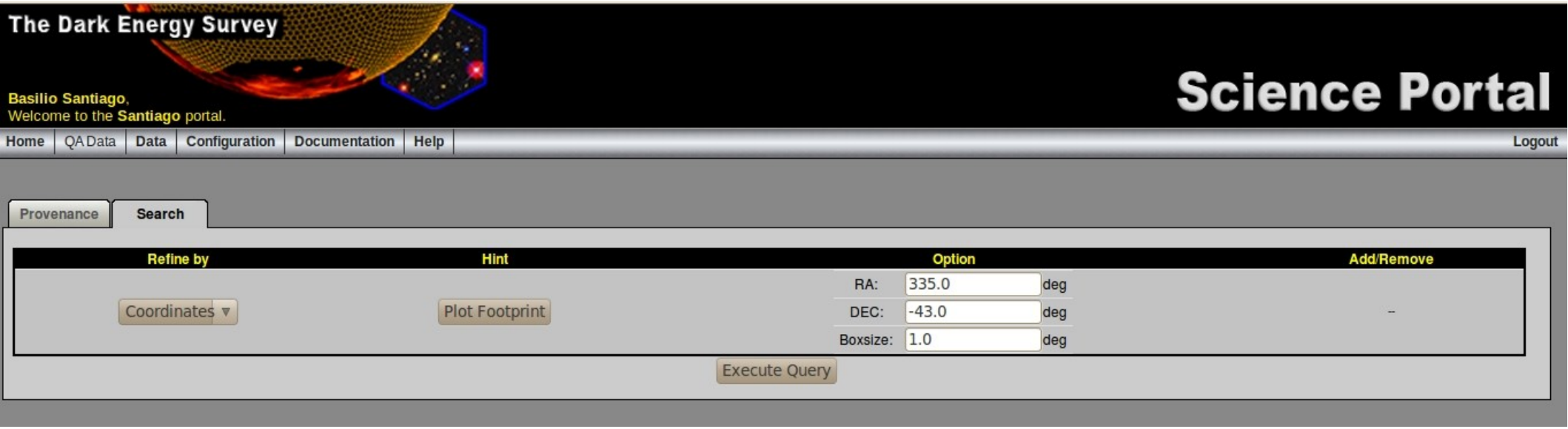


QA initial page

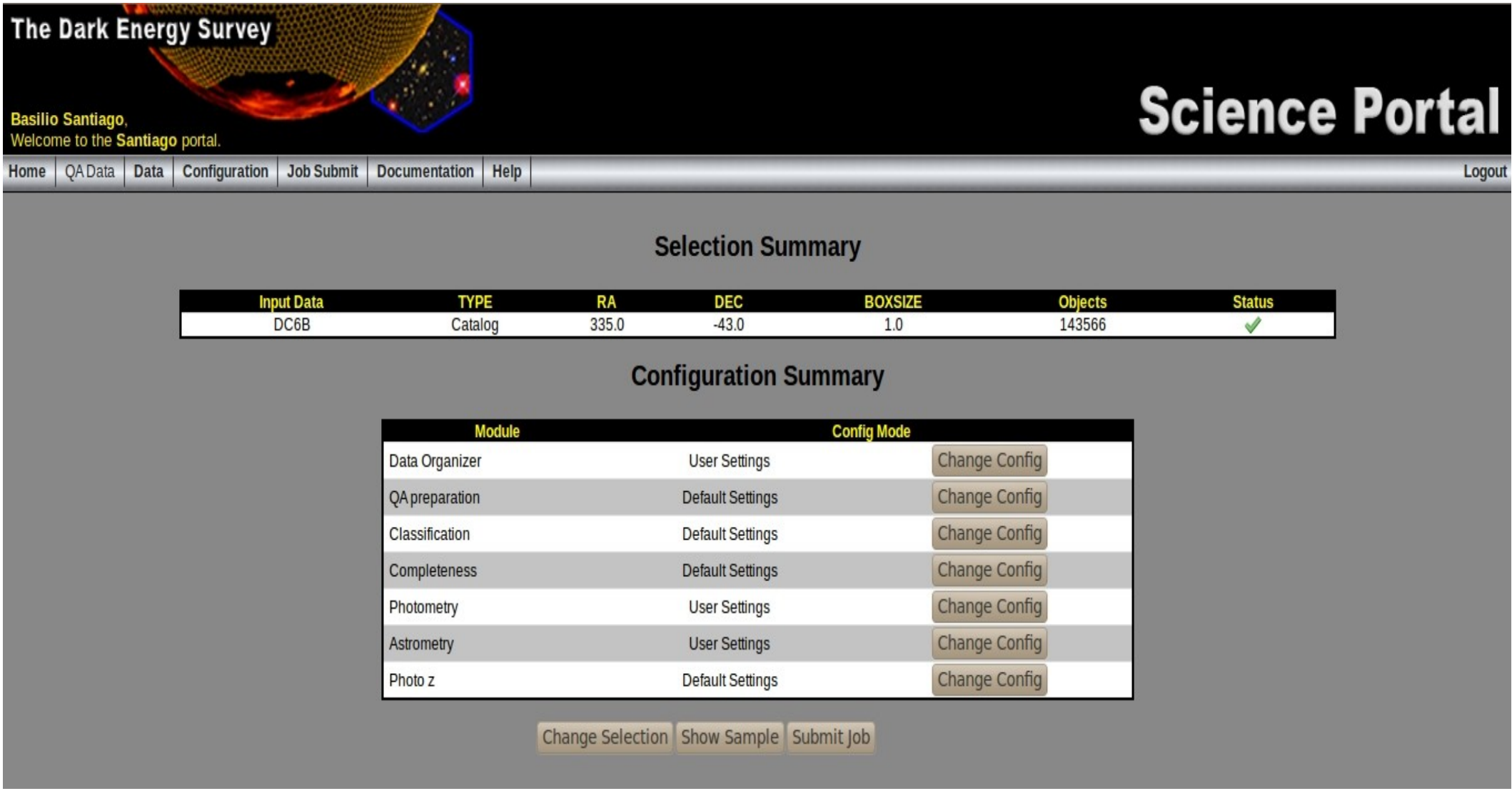


Data Selection





QA configuration



Examples of configuration sets

DARK ENERGY SURVEY

Home | QA Data | Data | Configuration | Job Submit | Documentation | Help

QA Preparation

Analysis

Use analysis threshold?	Yes / No
Analysis threshold	3
Reference band for analysis	i
Magnitude type	MAG_AUTO
Clean catalog using flags_r?	Yes / No
Limit of SEx flags_r (< 4). Full cleanse is 0.	4

Home | QA Data | Data | Configuration | Job Submit | Documentation | Help

Classification

Star-Galaxy classification

Star/Galaxy separation classifier	SPREAD_MODEL
S/G classifier value. CLASS_STAR>0.9, FLUX_RADIUS std. dev.<1-3, |SPREAD_MODEL|<0.002	0.002
Limiting magnitude for star/galaxy classification using CLASS_STAR	24.0

Configurable parameters for Science Requirements

DARK ENERGY SURVEY

DES-BRAZIL

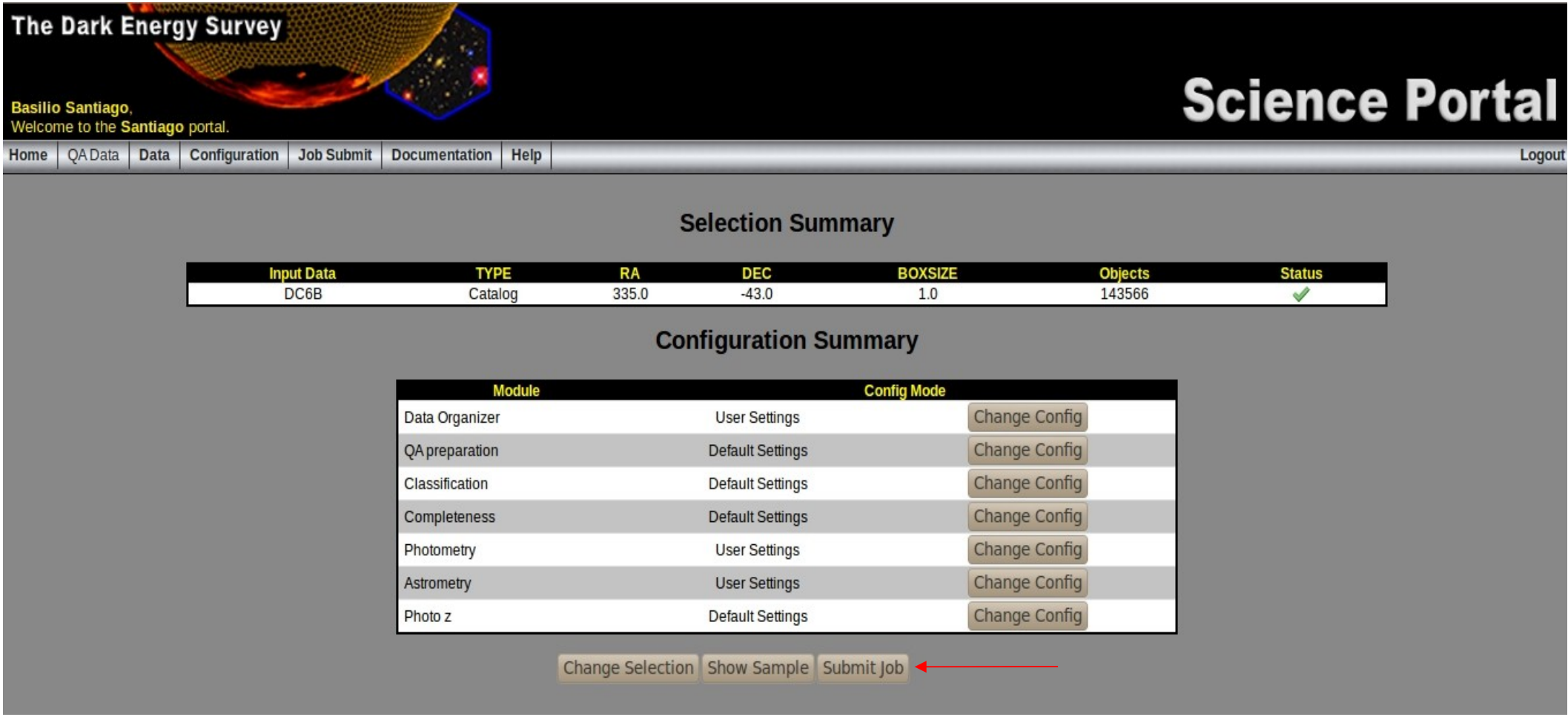
Home | QA Data | Data | Configuration | Job Submit | Documentation | Help

Photometry

R4a	
Magnitude type to use in estimating mag limits	MAG_APER2
Bright magnitude limit for mag distributions	16.0
Faint magnitude limit for mag distributions	25.0
Magnitude increment for mag distributions	0.2
Magnitude error level to use in mag limit determination	0.11
Minimum S/N	0.0
Maximum S/N	100.0
Increment in S/N	2.0

R6 mags	
Magnitude type to use in estimating mag limits	MAG_PSF
Bright magnitude limit for mag distributions	17.0
Faint magnitude limit for mag distributions	20.0
Magnitude increment for mag distributions	0.2
Magnitude error level to use in mag limit determination	0.11

Job submit


The Dark Energy Survey
Basilio Santiago,
Welcome to the Santiago portal.
Science Portal
Home
QA Data
Data
Configuration
Job Submit
Documentation
Help
Logout
Selection Summary
Input Data
TYPE
RA
DEC
BOXSIZE
Objects
Status
DC6B
Catalog
335.0
-43.0
1.0
143566
Configuration Summary
Module
Config Mode
Data Organizer
User Settings
Change Config
QA preparation
Default Settings
Change Config
Classification
Default Settings
Change Config
Completeness
Default Settings
Change Config
Photometry
User Settings
Change Config
Astrometry
User Settings
Change Config
Photo z
Default Settings
Change Config
Change Selection
Show Sample
Submit Job

Monitoring process

Final process view

DARK ENERGY SURVEY

DES-BRAZIL

santiago.linea.gov.br/VP/getViewProcessCon?pro=10008154

Automatically Saved	No
Algorithm	---
Configuration	View
Code Version	View

Input Data	
Input data	DC6B
Type	Catalog
RA (deg)	335.0
DEC (deg)	-42.
Selected Region (deg)	9.0
Objects	12299491
Area (sq deg)	54.14
Density (objects/sq arc min)	63.1

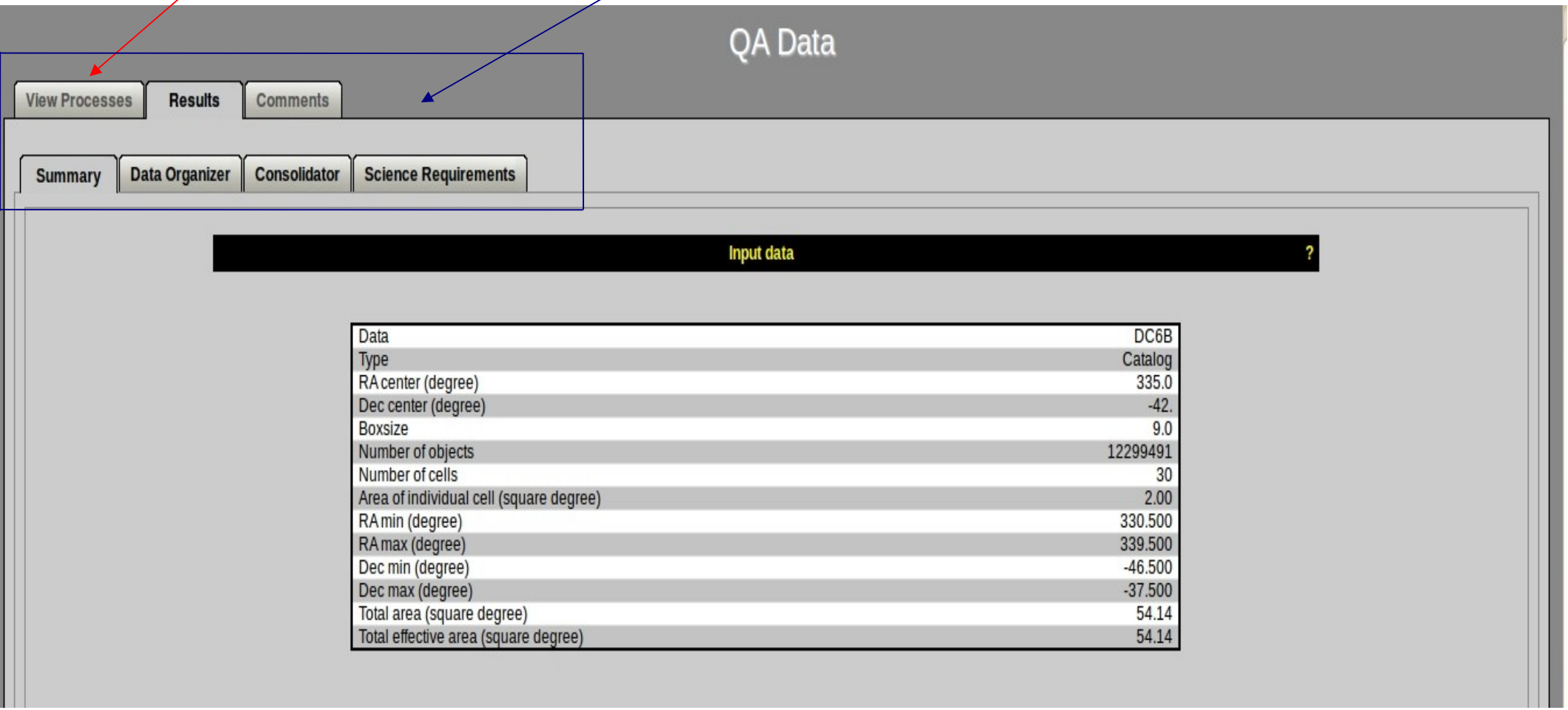
Module	Duration	Products	Config	Error Log	Pipeline Out	Log	Condor Log	NC	Status
dataorganizer	00:04:30	---						nc10	
dataretriever	00:06:38	---	---	---	---	---	---		
QA preparation	00:00:37	---		---	---	---	---		
Classification	00:00:13	---		---	---	---	---		
Masks Properties	00:05:54	---	---	---	---	---	---		
qa_basic_depth	00:00:05	---	---	---	---	---	---		
Completeness	00:00:30	---		---	---	---	---		
qa_basic_photometry	00:02:37	---		---	---	---	---		
Astrometry	00:00:10	---		---	---	---	---		
qa_basic_imagequality	00:00:03	---	---	---	---	---	---		
qa_basic_photoz	00:00:25	---		---	---	---	---		
qa_subarea_consolidator	00:00:02	---	---	---	---	---	---		
qa_area_consolidator	00:01:36	---	---					nc10	

Release Date: Wed Jan 11 2012 - 17:22:56 [santiago portal]

Process results (Sub)Product Log

earlier slide

Product log: global QA



Data	DC6B
Type	Catalog
RA center (degree)	335.0
Dec center (degree)	-42.
Boxsize	9.0
Number of objects	12299491
Number of cells	30
Area of individual cell (square degree)	2.00
RA min (degree)	330.500
RA max (degree)	339.500
Dec min (degree)	-46.500
Dec max (degree)	-37.500
Total area (square degree)	54.14
Total effective area (square degree)	54.14

Data Organizer: passage way to cell analysis

Where on the sky? How many parallel cells?

santiago.linea.gov.br/VP/getViewProcessCon?pro=10008154

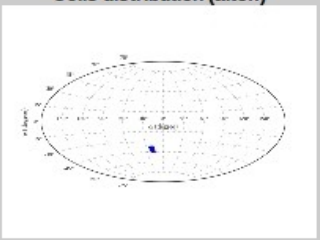
View Processes | Results | Comments

Summary | Data Organizer | Consolidator | Science Requirements

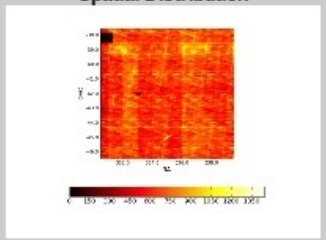
Quick Navigation
- Maps
- Footprint

Maps ?

Cells distribution (aitoff)

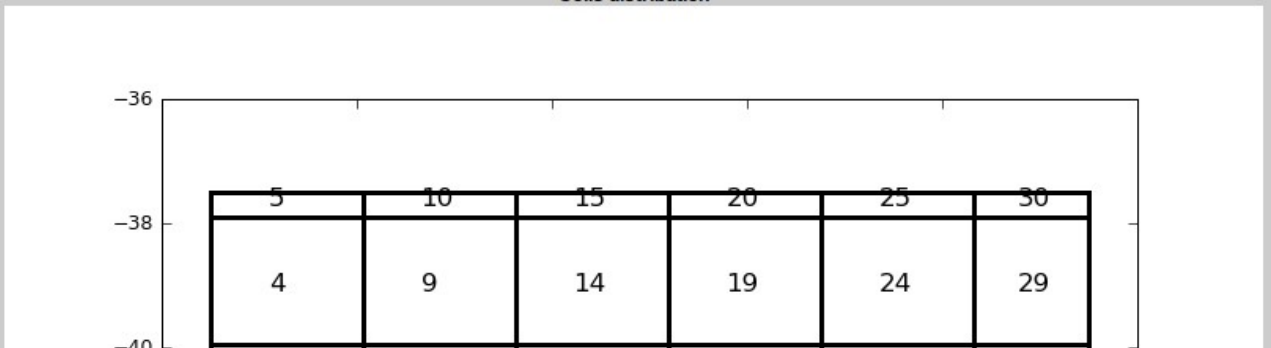


Spatial Distribution



Footprint ?

Cells distribution



Sub product log
Looking into a single cell

DARK ENERGY SURVEY

DES-BRAZIL

Chosen cell in DO

santiago.linea.gov.br/showFile/000010008154/productLog_1.12.xml

Cell 12

Summary | Preparation | Classification | Masks Properties | Depth | Completeness | Photometry | Astrometry | Image Quality | Photo-z | Science Requirements

Quick Navigation

- Input data
- Catalog preparation
- DM Matching
- Truth table properties
- Magnitude limit
- 99mag
- Classification
- Depth
- Completeness and Purity
- Completeness SRs
- Photometry
- Photometry SRs
- Astrometry
- Astrometry SRs
- Image Quality
- Photo-z

Input data

RA min (deg)	333.629
RA max (deg)	335.192
Dec min (deg)	-44.226
Dec max (deg)	-42.047
RA center (deg)	334.411
Dec center (deg)	-43.137
Total area (sq deg)	2.33

Catalog preparation

Use magnitude error limit to filter catalog?	Yes
Magnitude error threshold	3σ
Reference band	i
Magnitude type	MAG_AUTO
Use flags_r to filter catalog?	Yes
flags_r <	4
Original number of objects	585497
Number of objects after mag err cut off and filtering	223639

DM Matching

Number of matched stars	45339
Number of matched galaxies	100802
Number of unmatched objects	77498

Truth table properties

Catalog preparation

DARK ENERGY SURVEY

DES-BRAZIL

santiago.linea.gov.br/showFile/000010008154/productLog_1.12.xml

Cell 12

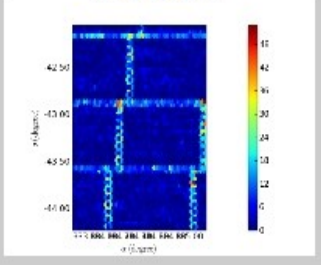
Summary | Preparation | Classification | Masks Properties | Depth | Completeness | Photometry | Astrometry | Image Quality | Photo-z | Science Requirements

Quick Navigation

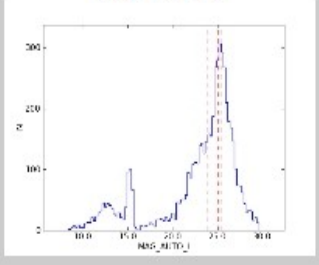
- Flag>3 objects
- True galaxies in catalog
- True stars in catalog
- Unmatched objects

Flag>3 objects ?

Flag>3 objects

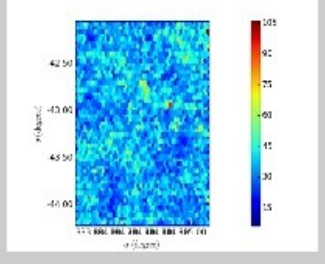


Flag>3 objects

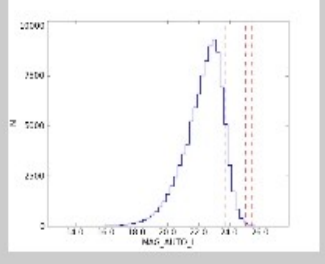


True galaxies in catalog ?

True galaxies in catalog



True galaxies in catalog



True stars in catalog ?

Object classification

DARK ENERGY SURVEY

DES-BRAZIL

santiago.linea.gov.br/showFile/000010008154/productLog_1.12.xml

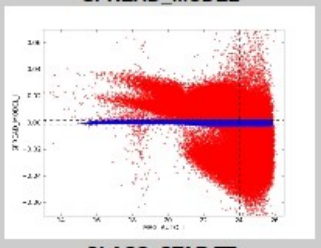
Summary | Preparation | Classification | Masks Properties | Depth | Completeness | Photometry | Astrometry | Image Quality | Photo-z | Science Requirements

Quick Navigation

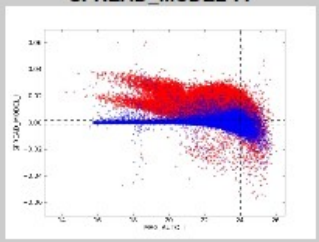
- Plots
- Statistics

Plots ?

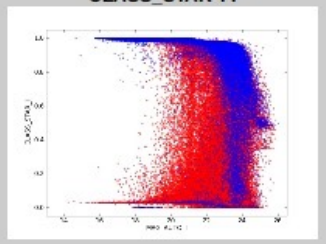
SPREAD_MODEL



SPREAD_MODEL TT



CLASS_STAR TT



Statistics ?

Algorithm	SPREAD_MODEL
SPREAD_MODEL	0.002
Number of classified objects	223639
Number of classified stars	46990
Number of true matched stars	45339
Number of classified galaxies	176649
Number of true matched galaxies	122837

Online help on each tab

Any questions?

Cell 12

Summary | Preparation | Classification | Masks Properties | Depth | Completeness | Photometry | Astrometry | Image Quality | Photo-z | Science Requirements

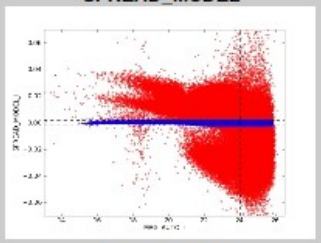
Quick Navigation

Plots

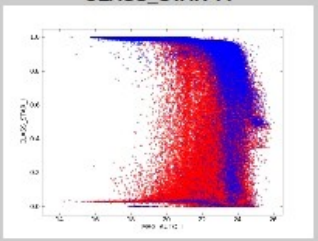
Statistics

Plots ?

SPREAD_MODEL



CLASS_STAR TT



Des-Brazil - Mozilla Firefox

twiki.linea.gov.br/bin/view/PortalDocumentation_des/QADataClassificationHelp?sk

Classification

This module shows some basic plots and statistics on object classification. QA is currently splitting the sample in only two categories: stars and galaxies. They are separated using a classifier parameter and a cut-off in this quantity as defined in the configuration stage. There are two subsections to this module: Plots and Statistics.

Plots

The Plots subsection shows scatter plots of the chosen classifier (default is the new spread_model) against the selected magnitude at configuration. Stars and galaxies are represented by blue and red colours, respectively. One of the plots shows them as classified. The other plots represents the distribution of true stars and galaxies in the classifier vs. magnitude plane. A similar plot for the true objects is provided using the class_star parameter, for comparison.

Statistics

The Statistics subsection lists the basic configuration parameters and the number of classified and true objects of each type. Notice that the number of true objects is always smaller than the number of classified ones, since there are always objects in the catalog for which no truth table counterpart was found.

Algorithm	
SPREAD_MODEL	
Number of classified objects	
Number of classified stars	
Number of true matched stars	
Number of classified galaxies	176649

Sample Depth

DARK ENERGY SURVEY

DES-BRAZIL

santiago.linea.gov.br/showFile/000010008154/productLog_1.12.xml

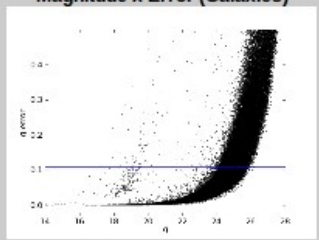
Summary | Preparation | Classification | Masks Properties | **Depth** | Completeness | Photometry | Astrometry | Image Quality | Photo-z | Science Requirements

Quick Navigation

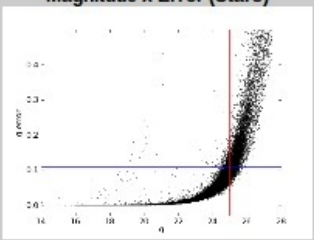
- g
- r
- i
- z
- Y

g ?

Magnitude x Error (Galaxies)



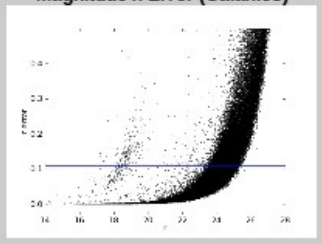
Magnitude x Error (Stars)



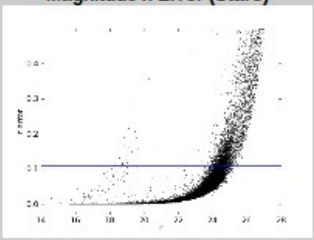
Galaxies magnitude limit for 10σ	24.74
Stars magnitude limit for 10σ	24.96

r ?

Magnitude x Error (Galaxies)



Magnitude x Error (Stars)



Galaxies magnitude limit for 10σ	24.20
Stars magnitude limit for 10σ	24.39

i ?

Completeness and Purity

DARK ENERGY SURVEY

DES-BRAZIL

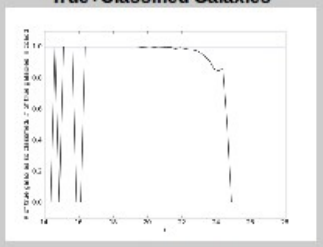
Summary | Preparation | Classification | Masks Properties | Depth | **Completeness** | Photometry | Astrometry | Image Quality | Photo-z | Summary of SRs

Quick Navigation

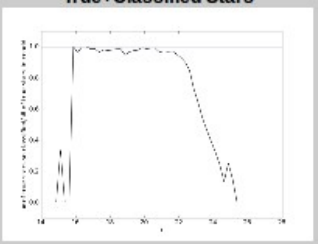
- Completeness wrt TT
- Completeness wrt CAT
- Purity
- Science Requirements

Completeness wrt CAT

True+Classified Galaxies



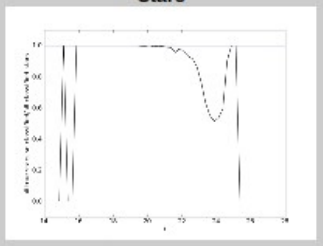
True+Classified Stars



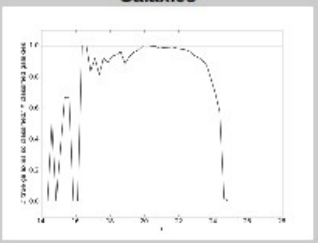
Median completeness for True+Classified Galaxies	0.99
Maximum completeness for True+Classified Galaxies	1.00
Median completeness for True+Classified Stars	0.93
Maximum completeness for True+Classified Stars	1.00

Purity

Stars



Galaxies



Median purity for Stars	0.99
Maximum purity for Stars	1.00
Median purity for Galaxies	0.89
Maximum purity for Galaxies	1.00

Photometric precision

DARK ENERGY SURVEY

DES-BRAZIL

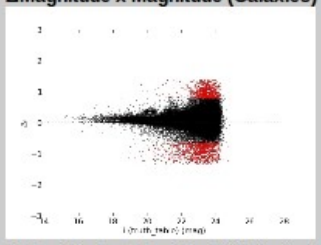
Summary | Preparation | Classification | Masks Properties | Depth | Completeness | Photometry | Astrometry | Image Quality | Photo-z | Science Requirements

Quick Navigation

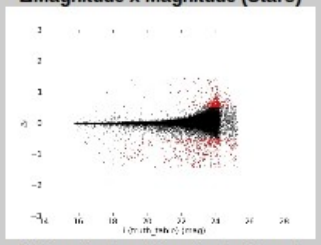
- g
- r
- i
- z
- Y
- Input Parameters
- R4a
- R6mag
- R6

i

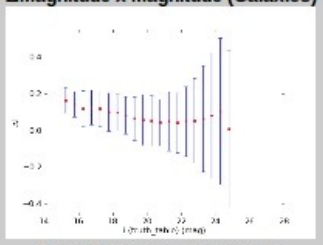
ΔMagnitude x Magnitude (Galaxies)



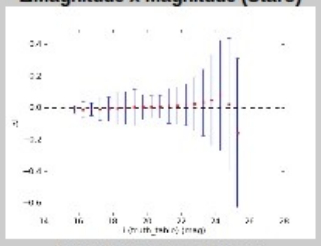
ΔMagnitude x Magnitude (Stars)



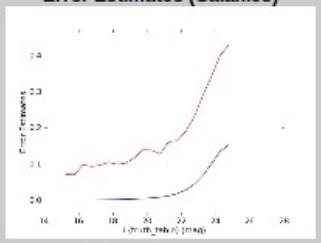
ΔMagnitude x Magnitude (Galaxies)



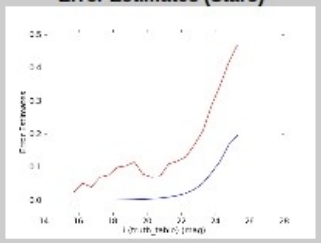
ΔMagnitude x Magnitude (Stars)



Error Estimates (Galaxies)



Error Estimates (Stars)



Mean mag residual for true galaxies	0.06
Median mag residual for true galaxies	0.05
RMS mag residual for true galaxies	0.24

Photometric SRs

DARK ENERGY SURVEY

DES-BRAZIL

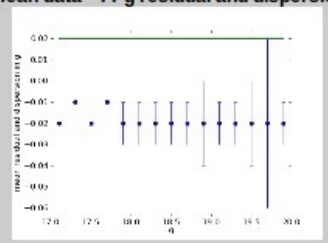
Summary | Preparation | Classification | Masks Properties | Depth | Completeness | Photometry | Astrometry | Image Quality | Photo-z | Science Requirements

Quick Navigation

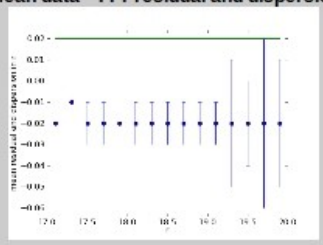
- g
- r
- i
- z
- Y
- Input Parameters
- R4a
- R6mag
- R6

R6mag ?

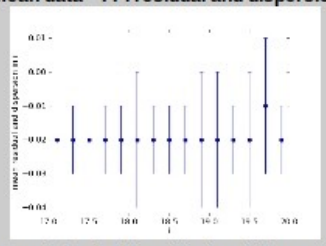
mean data - TT g residual and dispersion



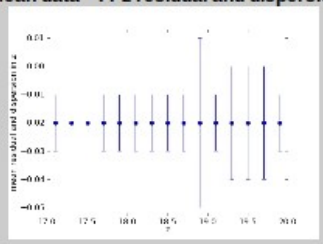
mean data - TT r residual and dispersion



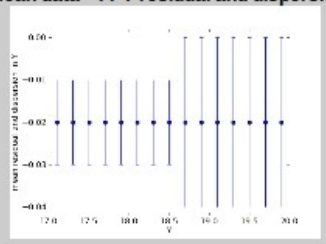
mean data - TT i residual and dispersion



mean data - TT z residual and dispersion



mean data - TT Y residual and dispersion



R6mag: g band <cat - TT> offset:	-0.02
R6mag: g band cat - TT dispersion:	0.00
R6mag: r band <cat - TT> offset:	-0.02

Colour colour diagrams

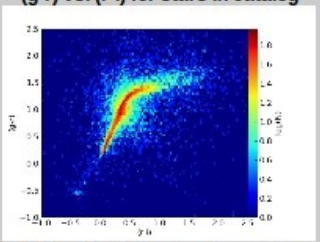
DARK ENERGY SURVEY

DES-BRAZIL

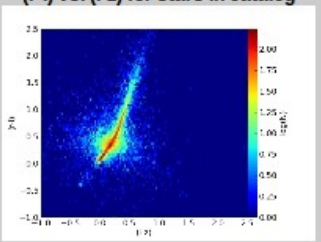
Summary | Preparation | Classification | Masks | Depth | Photometry | Color-color | Astrometry | Image Quality | Photo-z | Summary of SRs

Stars | Galaxies

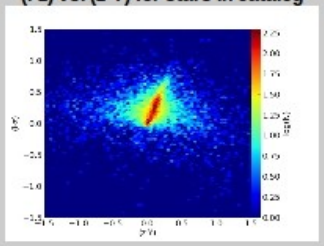
(g-r) vs. (r-i) for stars in catalog



(r-i) vs. (i-z) for stars in catalog



(i-z) vs. (z-Y) for stars in catalog



Checking the Astrometry

DARK ENERGY SURVEY

DES-BRAZIL

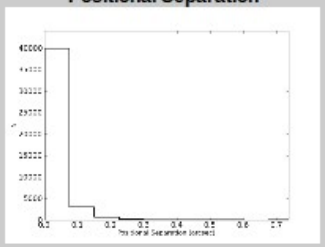
Summary | Preparation | Classification | Masks Properties | Depth | Completeness | Photometry | Astrometry | Image Quality | Photo-z | Science Requirements

Quick Navigation

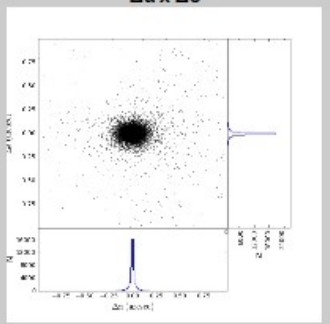
- Stars
- Galaxies
- Science Requirements

Stars

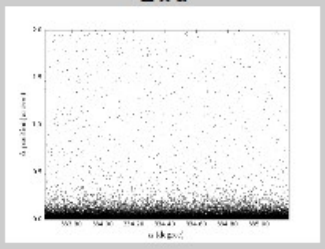
Positional Separation



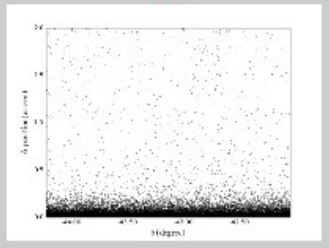
Δα x Δδ



Δ x α



Δ x δ



mean RA residual	0.011
RMS RA residual	0.236
mean DEC residual	-0.005
RMS DEC residual	0.172
Number of points used	45339
Positional separation (in arcsec) above which there are 2% of the objects	0.7368
Percentage of objects with positional separation equal to 0.0000 arcsec	3.6

Astrometry SRs

DARK ENERGY SURVEY

DES-BRAZIL

santiago.linea.gov.br/showFile/000010008154/productLog_1.12.xml#

Cell 12

Summary | Preparation | Classification | Masks Properties | Depth | Completeness | Photometry | Astrometry | Image Quality | Photo-z | Science Requirements

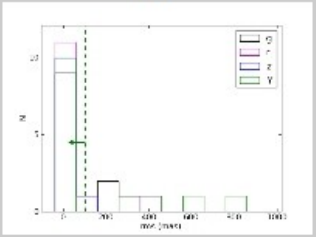
Quick Navigation

- Stars
- Galaxies
- Science Requirements

Science Requirements ?

R14

R15

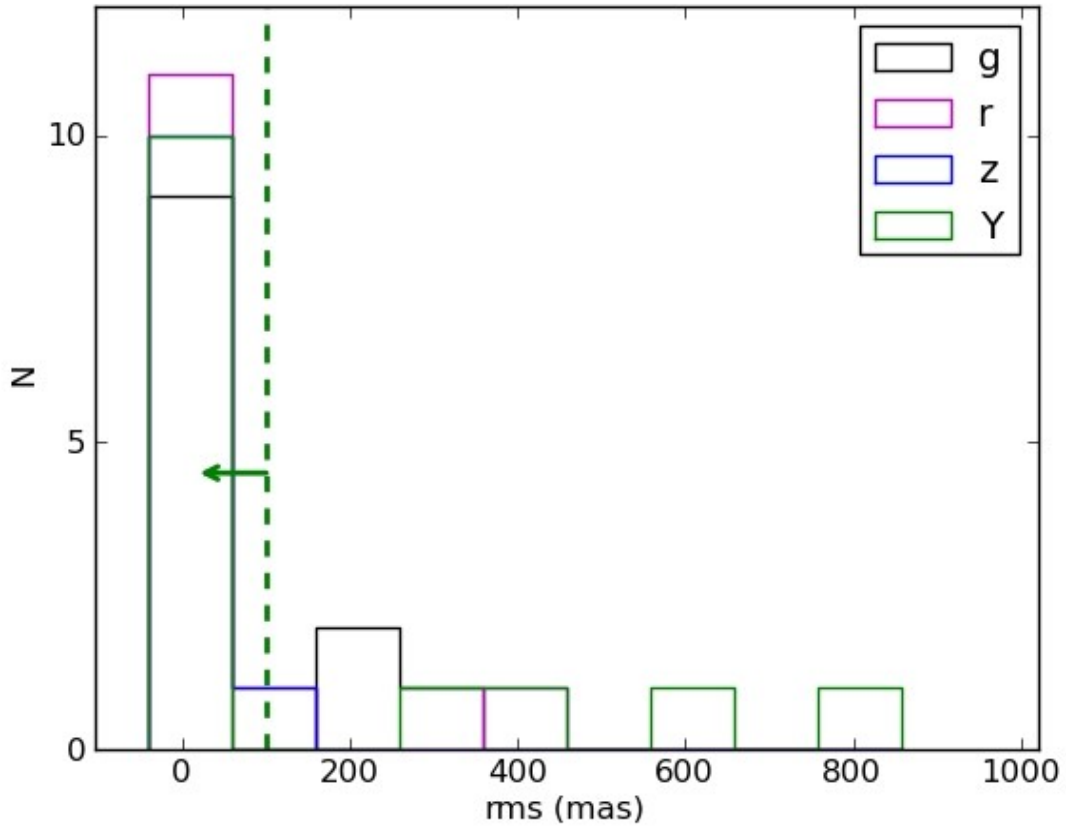


RMS residual wrt UCAC in scale:	
2.570 sq. deg.	457.760
0.286 sq. deg.	454.758
0.161 sq. deg.	456.969
0.643 sq. deg.	456.707
0.071 sq. deg.	452.937
0.103 sq. deg.	456.082

Region Used	
Ramin1	333.600000
Ramax1	335.200000
Decmin1	-44.200000
Decmax1	-42.000000
Number of cells used	16

Filter	g
1,5,50,95,99 Percentile positions	
1%	7.28
5%	7.28
50%	49.34
95%	2693.16
99%	2693.16

Cross filter astrometry comparison



R15. Distribution of RMS residual between i band and other bands for the different subcells. Different filter (grzY) are displayed as indicated. The dashed red line indicates requirement value.

Image quality

santiago.linea.gov.br/showFile/000010008154/productLog_1.20.xml#

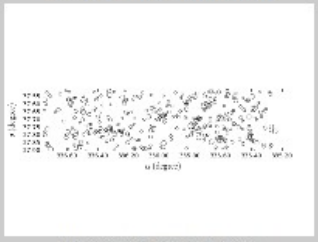
Summary | Preparation | Classification | Masks Properties | Depth | Completeness | Photometry | Astrometry | Image Quality | Photo-z | Summary of SRs

Quick Navigation

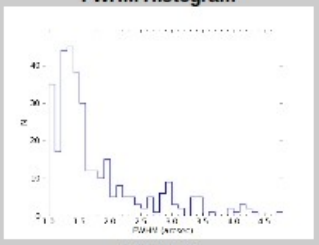
- Plots
- Statistics

Plots

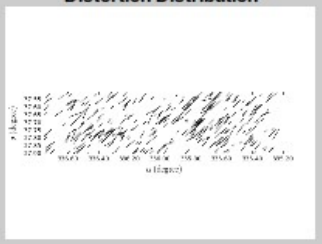
Scaled Stars Distribution



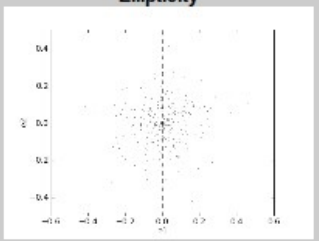
FWHM Histogram



Distortion Distribution



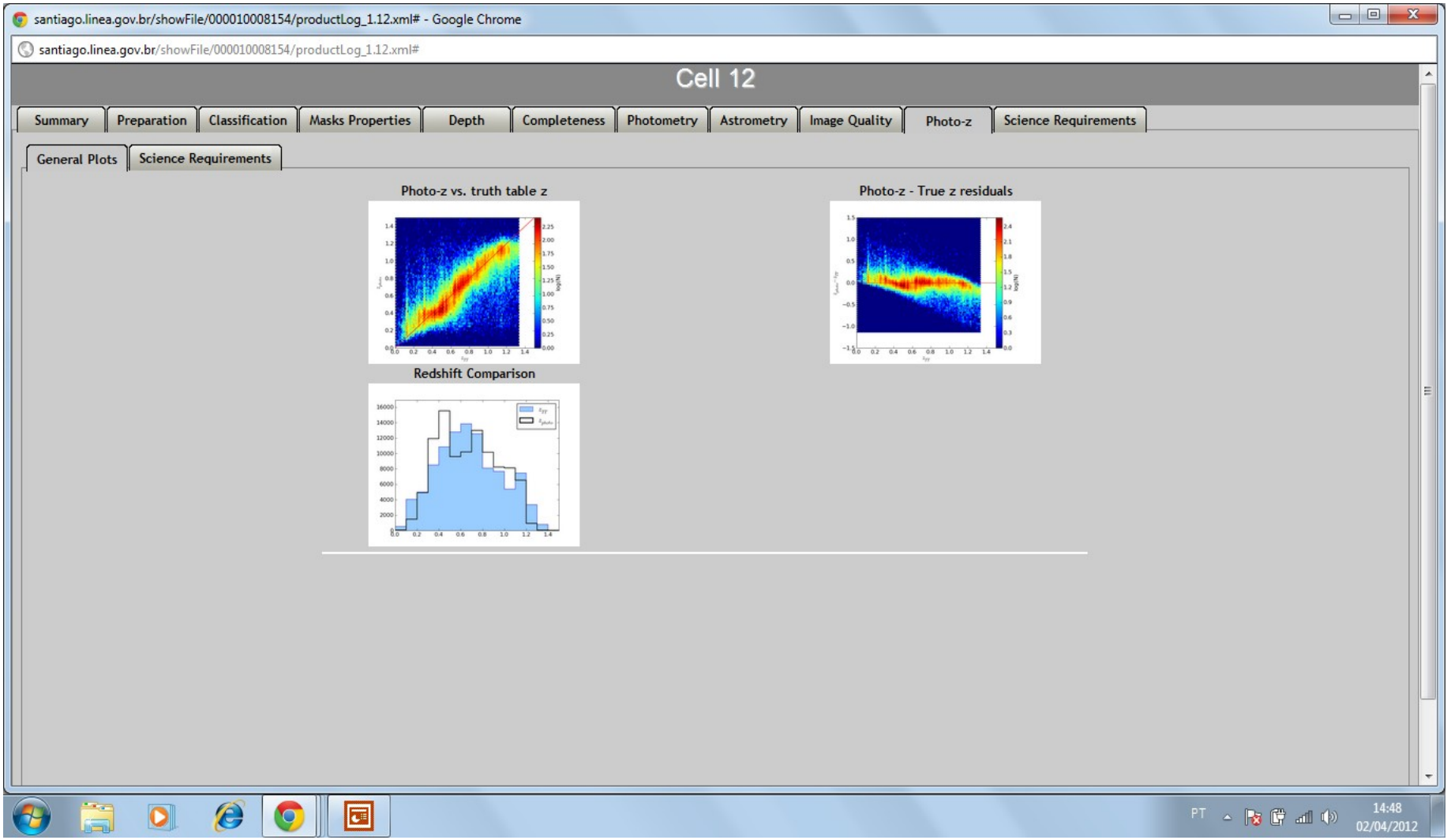
Ellipticity



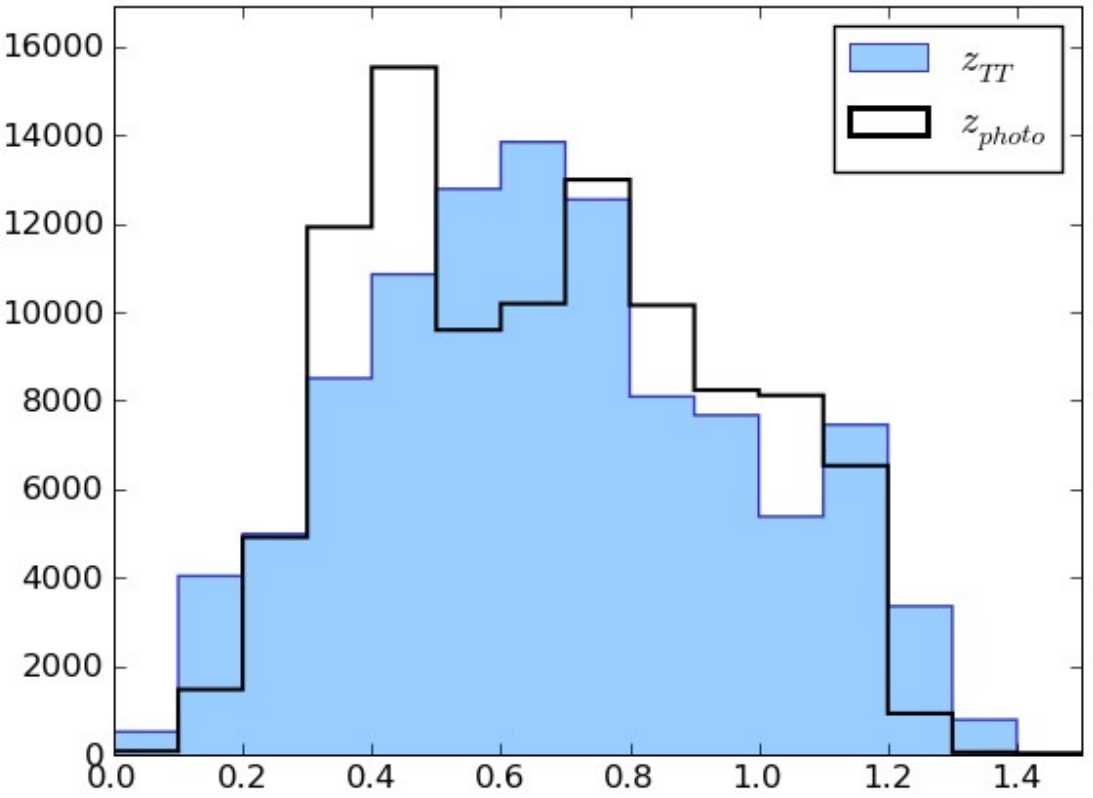
Statistics

Number of objects	44455
Mean of e1	0.0256
Standard deviation of e1	0.1533
Mean of e2	0.0014
Standard deviation of e2	0.1185
Mean of FWHM	1.4064
Median of FWHM	1.2417
Standard deviation of FWHM	1.2417

Photometric redshifts



True z vs Photo-z



Distribution of both truth table and photometric redshifts in 0.1 bins.

Summary of SRs

DARK ENERGY SURVEY

DES-BRAZIL

Summary | Preparation | Classification | Masks | Depth | Photometry | Color-color | Astrometry | Image Quality | Photo-z | Summary of SRs

Requirement	Value	Pass/fail
Completeness		
R27.0: completeness (%) > 0.93	NA	-
R28.0: Nstar/Ngals deviation (%) < 0.01	NA	-
Photometry		
R4.0: g mag lim > 24.6	24.98	✔
R4.1: r mag lim > 24.1	24.87	✔
R4.2: i mag lim > 24.3	24.49	✔
R4.3: z mag lim > 23.8	23.93	✔
R4.4: Y mag lim > 21.5	23.25	✔
R4.5: g mag lim > 24.6	-99.99	✘
R4.6: r mag lim > 24.1	-99.99	✘
R4.7: i mag lim > 24.3	-99.99	✘
R4.8: z mag lim > 23.8	24.50	✔
R4.9: Y mag lim > 21.5	23.19	✔
R6.0: g calibration accuracy (%) < 0.02	-0.011	✔
R6.1: r calibration accuracy (%) < 0.02	-0.015	✔
R6.2: i calibration accuracy (%) < 0.02	-0.020	✔
R6.3: z calibration accuracy (%) < 0.02	-0.020	✔
R6.4: Y calibration accuracy (%) < 0.02	-0.021	✔
R6.5: (g-r) calibration accuracy (%) < 0.02	0.000	✔
R10.0: short sci req 10 < 1.0	NA	-
Astrometry		
R14.0: RMS wrt UCAC3 (mas) < 100.0	485.18	✘
R15.0: RMS g vs i (mas) < 100.0	16.26	✔
R15.1: RMS r vs i (mas) < 100.0	11.48	✔
R15.2: RMS z vs i (mas) < 100.0	7.92	✔
R15.3: RMS Y vs i (mas) < 100.0	9.21	✔
Photo-z		
R8.0: galaxies photo-z dispersion < 0.12	0.170	✘
R22.0: photo-z bias uncertainty < 0.001	-0.0015	✔
R22.1: photo-z bias covariance < 0.0005	0.00800	✘
R23.0: fraction zbias beyond 2sig (%) < 0.1	0.10	✔
R23.1: fraction zbias beyond 3sig (%) < 0.015	0.0400	✘

Global QA analysis (Back to the Product Log)

santiago.linea.gov.br/VP/getViewProcessCon?pro=10008154

View Processes | Results | Comments

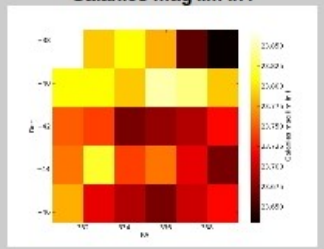
Summary | Data Organizer | Consolidator | Science Requirements

Quick Navigation

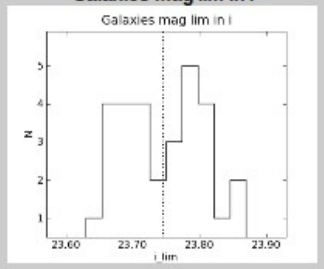
- Counts Heatmaps
- Galaxies mag lim in g
- Galaxies mag lim in r
- Galaxies mag lim in i
- Galaxies mag lim in z
- Galaxies mag lim in Y
- Stars mag lim in g
- Stars mag lim in r
- Stars mag lim in i
- Stars mag lim in z
- Stars mag lim in Y
- Mean residual for galaxies in g
- Mean residual for galaxies in r
- Mean residual for galaxies in i
- Mean residual for galaxies in z
- Mean residual for galaxies in Y
- Mean residual for stars in g
- Mean residual for stars in r
- Mean residual for stars in i
- Mean residual for stars in z
- Mean residual for stars in Y
- Mean residual in RA for galaxies
- Mean residual in DEC for galaxies
- Mean residual in RA for stars
- Mean residual in DEC for stars

Galaxies mag lim in i ?

Galaxies mag lim in i



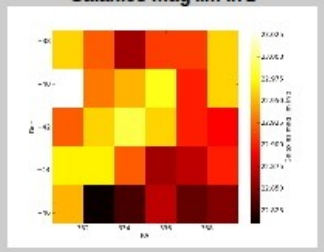
Galaxies mag lim in i



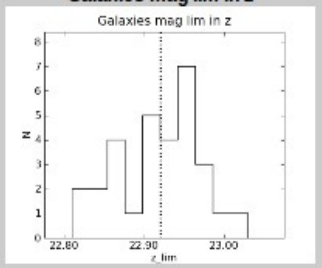
Median of Galaxies mag lim in i	23.74
Standard deviation of Galaxies mag lim in i	0.062

Galaxies mag lim in z ?

Galaxies mag lim in z



Galaxies mag lim in z



Median of Galaxies mag lim in z	22.92

Global QA analysis (continuation)

DARK ENERGY SURVEY

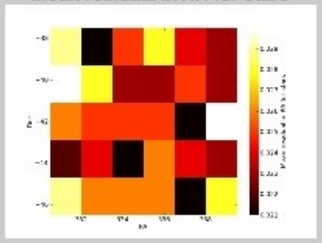
DES-BRAZIL

Summary | Data Organizer | Consolidator | Science Requirements

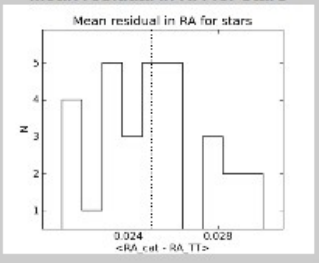
Quick Navigation

Mean residual in RA for stars

Mean residual in RA for stars



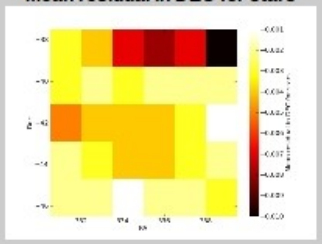
Mean residual in RA for stars



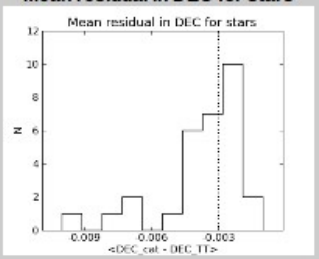
Median of Mean residual in RA for stars	0.025
Standard deviation of Mean residual in RA for stars	0.003

Mean residual in DEC for stars

Mean residual in DEC for stars



Mean residual in DEC for stars



Median of Mean residual in DEC for stars	-0.003
Standard deviation of Mean residual in DEC for stars	0.002

Global SR analysis

DARK ENERGY SURVEY

DES-BRAZIL

View Processes | Results | Comments

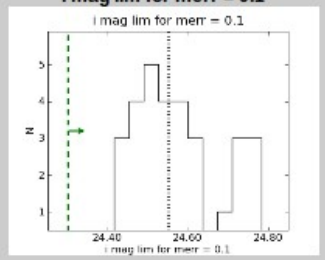
Summary | Data Organizer | Consolidator | Science Requirements

Quick Navigation

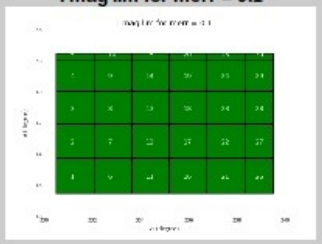
- R4.0
  - R4.1
  - R4.2
  - R4.3
  - R4.4
  - R4.5
  - R4.6
  - R4.7
  - R4.8
  - R4.9
- R6.0
  - R6.1
  - R6.2
  - R6.3
  - R6.4
  - R6.5
- R8.0
- R14.0
- R15.0
  - R15.1
  - R15.2
  - R15.3
- R22.0
  - R22.1
- R23.0
  - R23.1
- R27.0
- R28.0

R4.2 ?

i mag lim for merr = 0.1



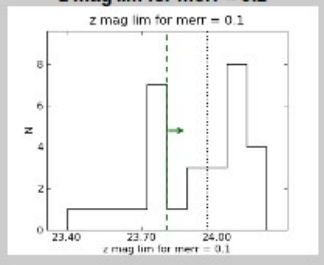
i mag lim for merr = 0.1



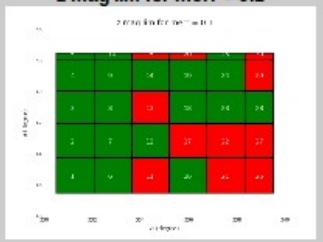
Average	24.58
Median	24.55
Standard deviation	0.105

R4.3 ?

z mag lim for merr = 0.1



z mag lim for merr = 0.1



Average	23.92

Global SRs QA (continuation)

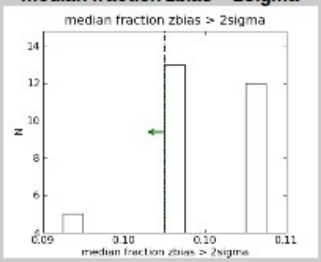
View Processes | **Results** | Comments

Summary | Data Organizer | Consolidator | **Science Requirements**

Quick Navigation

- R4.0
  - R4.1
  - R4.2
  - R4.3
  - R4.4
  - R4.5
  - R4.6
  - R4.7
  - R4.8
  - R4.9
- R6.0
  - R6.1
  - R6.2
  - R6.3
  - R6.4
  - R6.5
- R8.0
- R14.0
- R15.0
  - R15.1
  - R15.2
  - R15.3
- R22.0
  - R22.1
- R23.0
  - R23.1
- R27.0
- R28.0

median fraction zbias > 2sigma

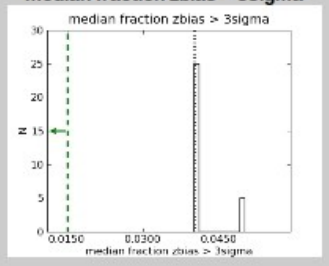


median fraction zbias > 2sigma

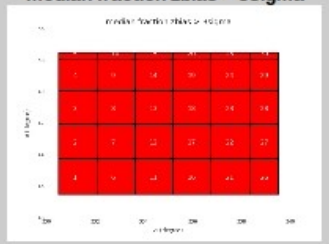
Average	0.102
Median	0.100
Standard deviation	0.007

R23.1 ?

median fraction zbias > 3sigma



median fraction zbias > 3sigma



Average	0.042
Median	0.040
Standard deviation	0.004

Closing remarks

- A 1st version of QA is now solidly implemented in the Portal → an accessible and well versioned tool available to the entire collaboration.

- We have already run it over most of the DC6 region.

- Near Future: Include advanced QA tests: number counts, correlation function,luminosity function, etc.

- Near Future: adapt quality analysis and tests to real data (substitute a truth table by golden standards)