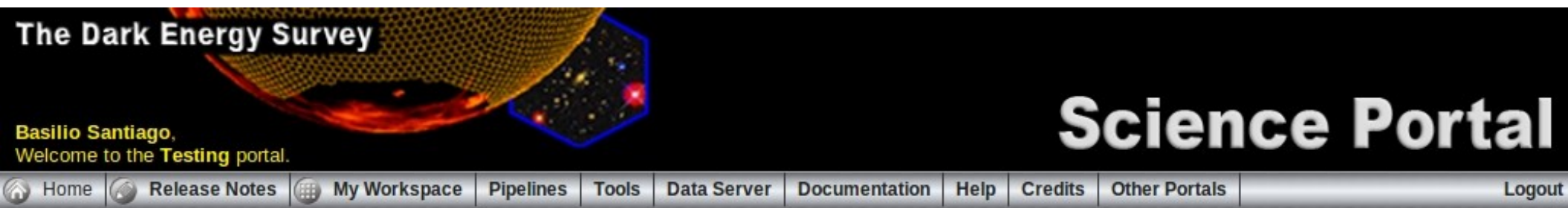




Laboratório Interinstitucional de e-Astronomia

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





DARK ENERGY
SURVEY

QA for the Dark Energy Survey



DES-BRAZIL

- 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



DARK ENERGY
SURVEY

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



DARK ENERGY
SURVEY

Navigating through the DES Portal



The Dark Energy Survey

Basilio Santiago,
Welcome to the **Santiago** portal.

Science Portal

Home Release Notes My Workspace Pipelines Tools Data Server Documentation Help Credits Other Portals Logou

>>

Introduction to the Science Portal

This portal is a web-based service that gives the user access to the data, documentation are all easily accessible in an interactive way. The portal also provides tools for power-users and system administrators to oversee the entire system.

The services currently available under the menu folder include:

- **My Workspace:** provides access to the profile, processes, configurations, products from process and uploaded data. Besides that, the user can require a **developer** status, allowing the access to: 1) the portal machines through ssh; 2) the code repository in order to pull and push code;
- **Pipelines:**
 - **Data Reduction:** provides access to Quick Reduce, a tool to reduce DECam data in the mountain. Also an adapted version to CTIO 1m runs is currently available and will turn into the PRECam reduction pipeline.
 - **Advanced Products:** here the user can produce and export catalogs, either regular or value-added, and validate them using the tools available in the portal.
 - **Science:** provides access to the available scientific pipelines, grouped in sub-categories (sub-folders) representing different working and study groups. The available pipelines (processes) in each group varies in the degree of sophistication and scope but all consist of concatenated full configurable modules chosen according to the user-specified workflow and algorithms, that query for data, process one or more jobs, depending on the number of input data sets, and generate scientific products that are stored in a temporary directory, which can later be saved or exported by the user. The processes can be run offline with the users being notified by e-mail upon completion.
- **Tools :** provides a large variety of tools, like processes running in the cluster, code manager, cluster environmental information, enabling the access to the administrative layer of the portal for those with permission. Here is also available the means to monitor the Survey Coverage
- **Data Server :** provides the infrastructure and the tools to upload data to the portal and to export the data from within the portal framework to a user specified target directory (in different formats depending on the product type) or to publish a result to the collaboration at large (such as exemplified by the value-added catalog) or in the public domain. A data release must be authorized by the management of the



DARK ENERGY
SURVEY

QA initial page



DES-BRAZIL

The Dark Energy Survey

Basilio Santiago,
Welcome to the **Santiago** portal.

Science Portal

[Home](#) [QA Data](#) [Data](#) [Configuration](#) [Documentation](#) [Help](#)

[Logout](#)

>>

QA Data

The Quality Assurance (QA) data tool is meant to evaluate the quality of both real and mock (data challenge, DC) data. QA is divided into several modules, each one dealing with a particular type of quality assurance test. Currently, QA data includes the following modules: catalog preparation, object classification, depth assessment, completeness and purity estimates, data masks, photometry, astrometry, image quality, photo-z. In the case of DES DC data, all these modules test for different DES science requirements, whose results are summarized in the Science Requirements module. Details about how these tests are carried out and how the results are presented are given below.

Navigating through the QA pipeline

As in other pipelines and tools in the DES-Brazil portal, QA Data needs to be configured in a sequence of steps, which involve:

- **Data:** click on this tab in the upper menubar in order to select the location (origin), database name and data type you want as input to the QA process. This is done in the provenance tab. The search tab allows you to refine your data query, using a coordinate range for instance.
- **Configuration:** this tab opens a window where you can configure the main parameters governing the QA process. These parameters, as the QA workflow itself, are split according to the QA module which they are relevant to. A wrench symbol next to each module allows you to view and edit the parameter values for that module.
- **Job submit:** once you go through the data selection, you will see a job submit button, which will also allow you to recheck your configuration parameters. Once you are satisfied with your data and configuration choices, you are then ready to press the submit button and start your process.
- **The actual processing:** you may logout of the portal once you submit a process and come back latter. An email will also be sent to you once the process is finished. If you resiliently decide to stay logged on, you will see a window showing the process status. This window, which is renewed every 30s, gives a lot of information, including how many processes have been run in parallel, how many were successfull, how many failed and at which modules.

QA results



DARK ENERGY
SURVEY

Data Selection



The Dark Energy Survey

Basilio Santiago,
Welcome to the **Santiago** portal.

Home QA Data **Data** Configuration Documentation Help Logout

Provenance Search

Location	Input Data	Type
ON ▼	DC6B ▼	Catalog ▼

Update

The Dark Energy Survey

Basilio Santiago,
Welcome to the **Santiago** portal.

Home QA Data **Data** Configuration Documentation Help Logout

Provenance Search

Refine by	Hint	Option	Add/Remove
Coordinates ▼	Plot Footprint	RA: 335.0 deg	-
		DEC: -43.0 deg	
		Boxsize: 1.0 deg	

Execute Query



DARK ENERGY
SURVEY

QA configuration



DESBRAZIL

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](#)



DARK ENERGY
SURVEY

Examples of configuration sets



DES-BRAZIL

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

QA Preparation

Analysis

Use analysis threshold?	<input checked="" type="radio"/> Yes <input type="radio"/> No
Analysis threshold	3 ▼
Reference band for analysis	i ▼
Magnitude type	MAG_AUTO ▼
Clean catalog using flags_r?	<input checked="" type="radio"/> Yes <input type="radio"/> 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



DARK ENERGY
SURVEY

Configurable parameters for Science Requirements



DE5-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



DARK ENERGY
SURVEY

Job submit



DESBRAZIL

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



DARK ENERGY
SURVEY

Monitoring process



DES-BRAZIL

Home	QA Data	Data	Configuration	Job Submit	Monitor	Documentation	Help	Logout
------	---------	------	---------------	------------	---------	---------------	------	--------

Pipeline: QA Data

User

Basilio Santiago

Process

10008337

Start

Sun Mar 25 07:54:16 2012

Algorithm

Configuration

[View](#)

Monitor

[Refresh](#) Update in 14 seconds

STOP

Input Data

Input data

DC6B

Type

Catalog

RA(deg)

335.0

DEC(deg)

-43.0

Boxsize(deg)

1.0

Objects

143566

Area(square deg)

0.68

Density(objects/square arc)

58.5

Block 1

Queue : 1

Module	Duration	Success	Error
Data Organizer	---	0	0

Block 2

Queue : 0

Module	Duration	Success	Error
Data Retriever	---	0	0
QA preparation	---	0	0
Classification	---	0	0



DARK ENERGY
SURVEY

Final process view



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	✓



DARK ENERGY
SURVEY

Process results (Sub)Product Log



DES-BRAZIL

earlier slide

Product log: global QA

QA Data

View Processes

Results

Comments

Summary

Data Organizer

Consolidator

Science Requirements

Input data

?

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



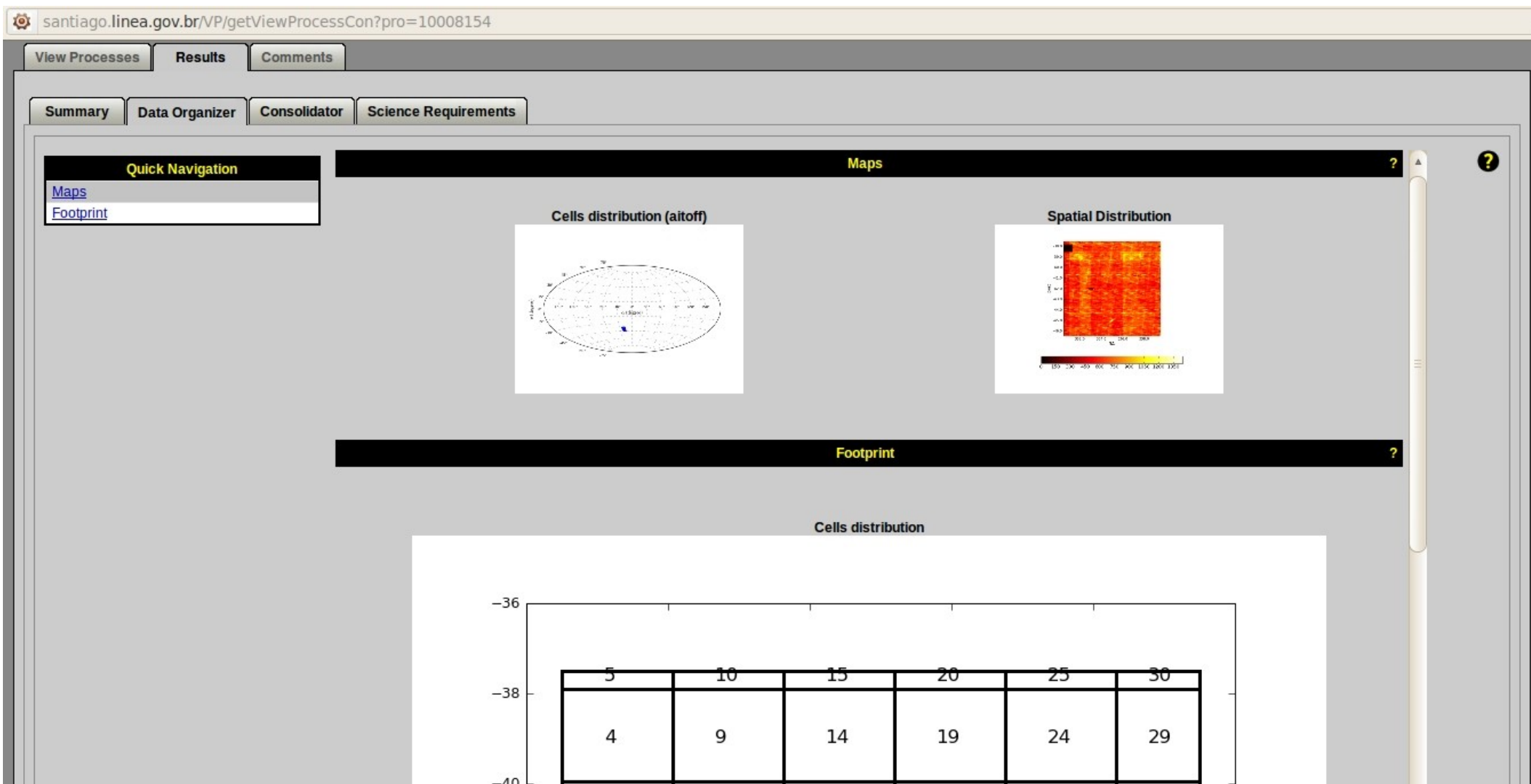
DARK ENERGY
SURVEY

Data Organizer: passage way to cell analysis



DES-BRAZIL

Where on the sky? How many parallel cells?





DARK ENERGY
SURVEY

Sub product log Looking into a single cell



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	3 σ
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



DARK ENERGY
SURVEY

Catalog preparation



DES-BRAZIL

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

Cell 12

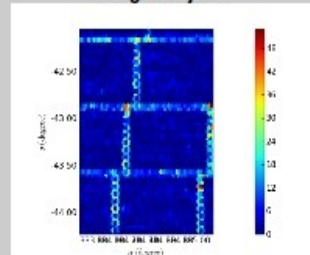
Summary Preparation Classification Masks Properties Depth Completeness Photometry Astrometry Image Quality Photo-z Science Requirements

Quick Navigation

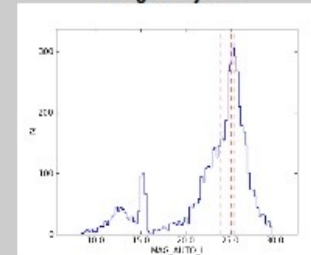
[Flag>3 objects](#)
[True galaxies in catalog](#)
[True stars in catalog](#)
[Unmatched objects](#)

Flag>3 objects

Flag>3 objects

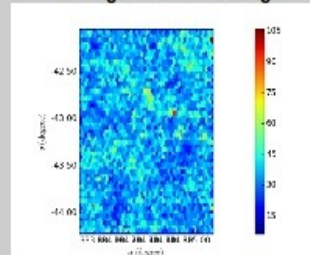


Flag>3 objects

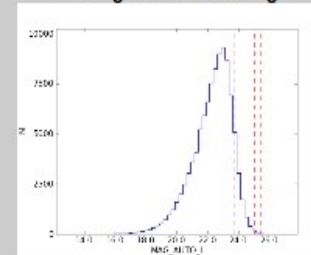


True galaxies in catalog

True galaxies in catalog



True galaxies in catalog



True stars in catalog

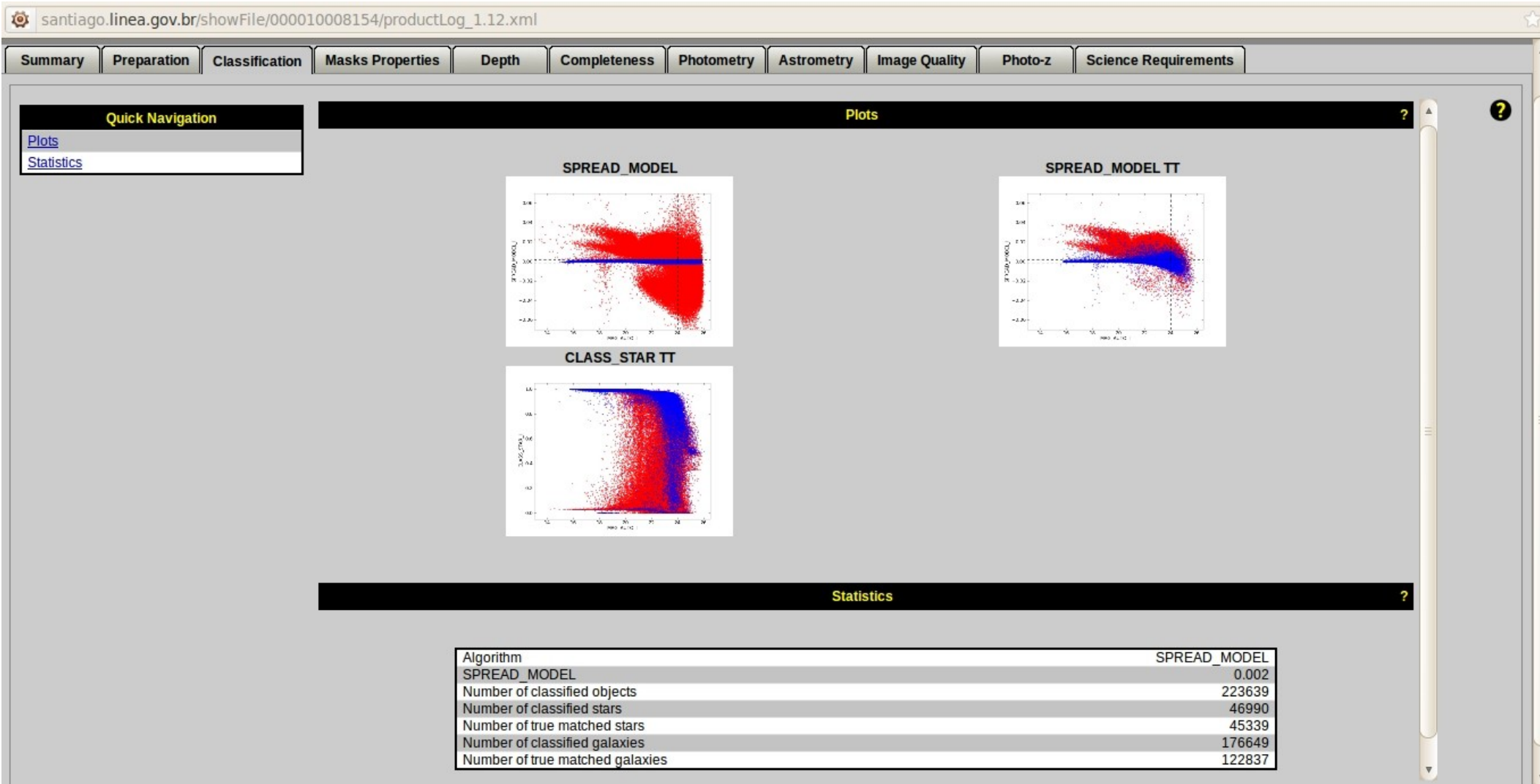


DARK ENERGY
SURVEY

Object classification



DES-BRAZIL





DARK ENERGY
SURVEY

Online help on each tab



DES-BRAZIL

Any questions?

Cell 12

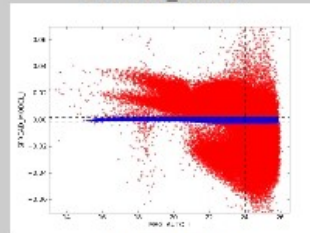
Summary Preparation Classification Masks Properties Depth Completeness Photometry Astrometry Image Quality Photo-z Science Requirements

Quick Navigation

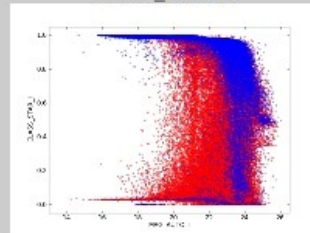
[Plots](#)
[Statistics](#)

Plots

SPREAD_MODEL



CLASS_STAR TT



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

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.

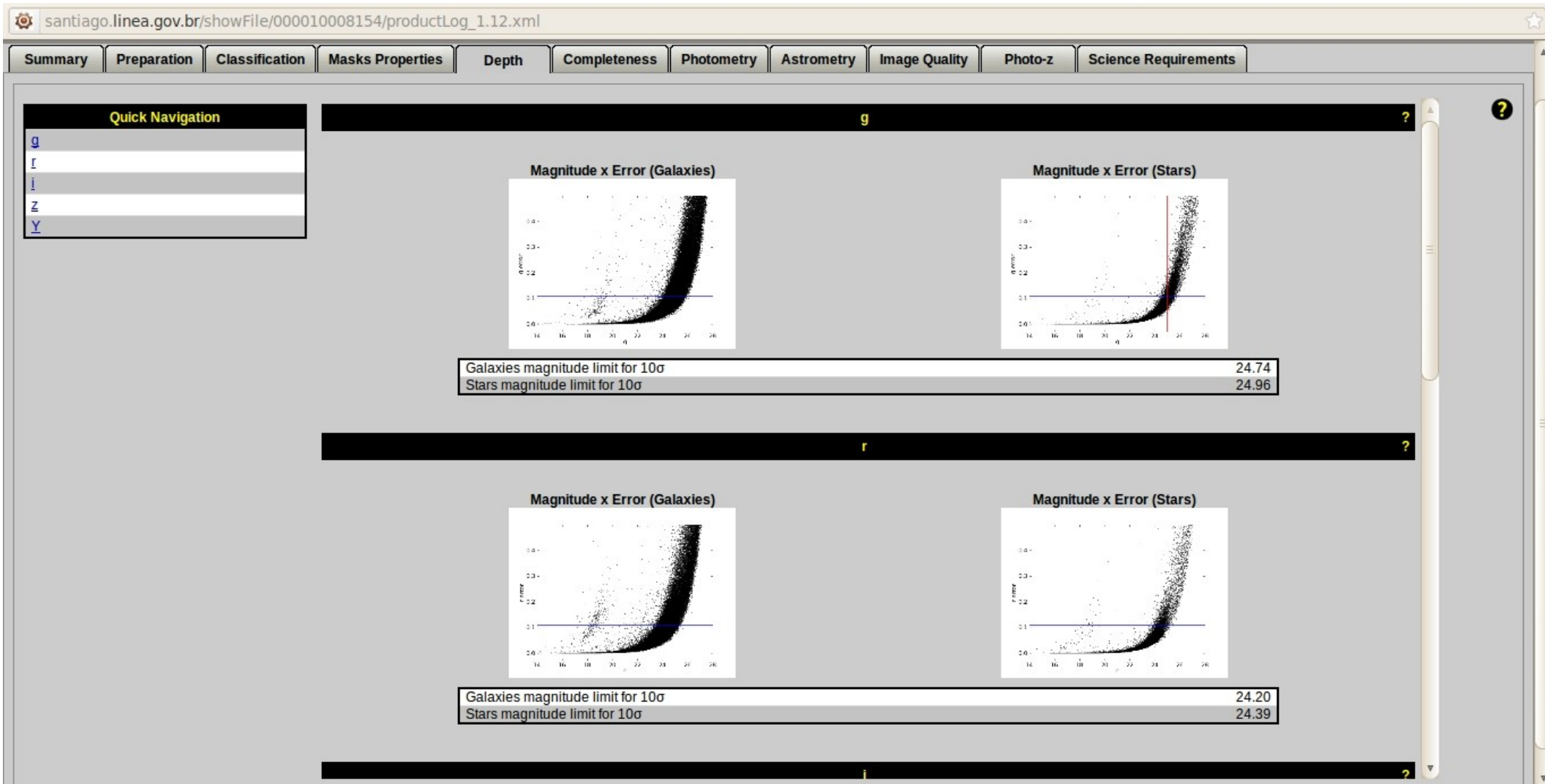


DARK ENERGY
SURVEY

Sample Depth



DES-BRAZIL





DARK ENERGY
SURVEY

Completeness and Purity



DES-BRAZIL





DARK ENERGY
SURVEY

Photometric precision



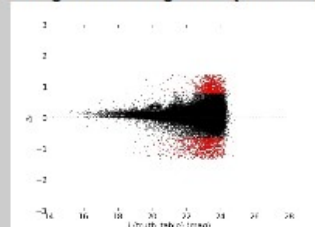
DES-BRAZIL

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

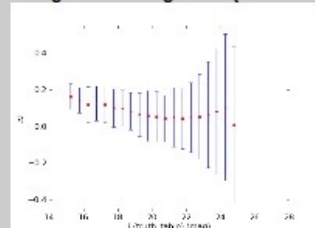
Quick Navigation

[g](#)
[r](#)
[i](#)
[z](#)
[y](#)
[Input Parameters](#)
[R4a](#)
[R6mag](#)
[R6](#)

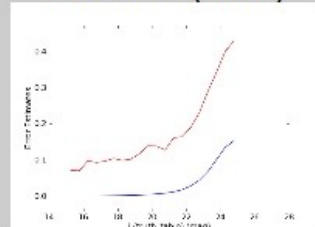
Δ Magnitude x Magnitude (Galaxies)



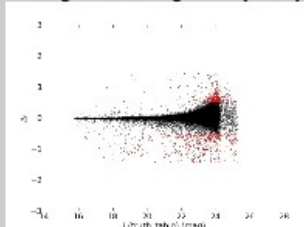
Δ Magnitude x Magnitude (Galaxies)



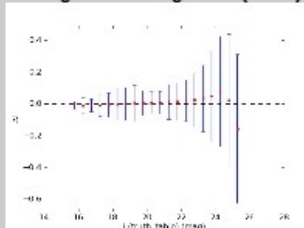
Error Estimates (Galaxies)



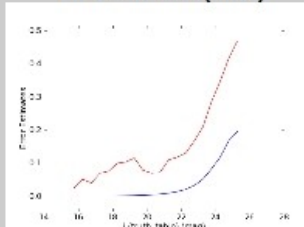
Δ Magnitude x Magnitude (Stars)



Δ Magnitude x Magnitude (Stars)



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



DARK ENERGY
SURVEY

Photometric SRs



DES-BRAZIL

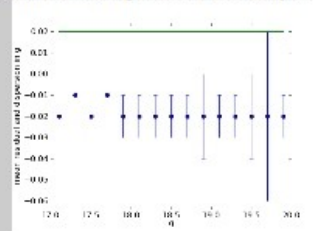
Summary Preparation Classification Masks Properties Depth Completeness Photometry Astrometry Image Quality Photo-z Science Requirements

Quick Navigation

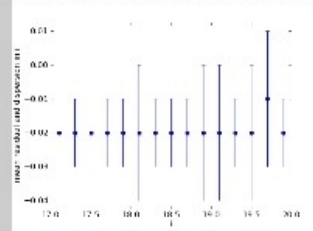
[g](#)
[r](#)
[i](#)
[z](#)
[Y](#)
[Input Parameters](#)
[R4a](#)
[R6mag](#)
[R6](#)

R6mag

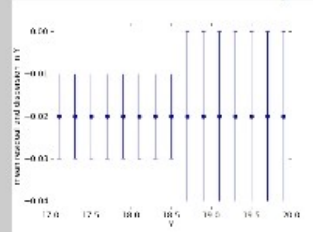
mean data - TT g residual and dispersion



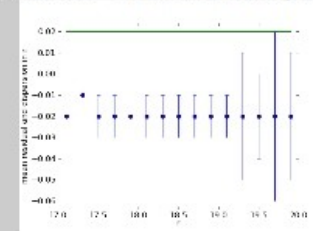
mean data - TT i residual and dispersion



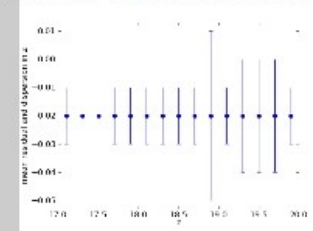
mean data - TT Y residual and dispersion



mean data - TT r residual and dispersion



mean data - TT z 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

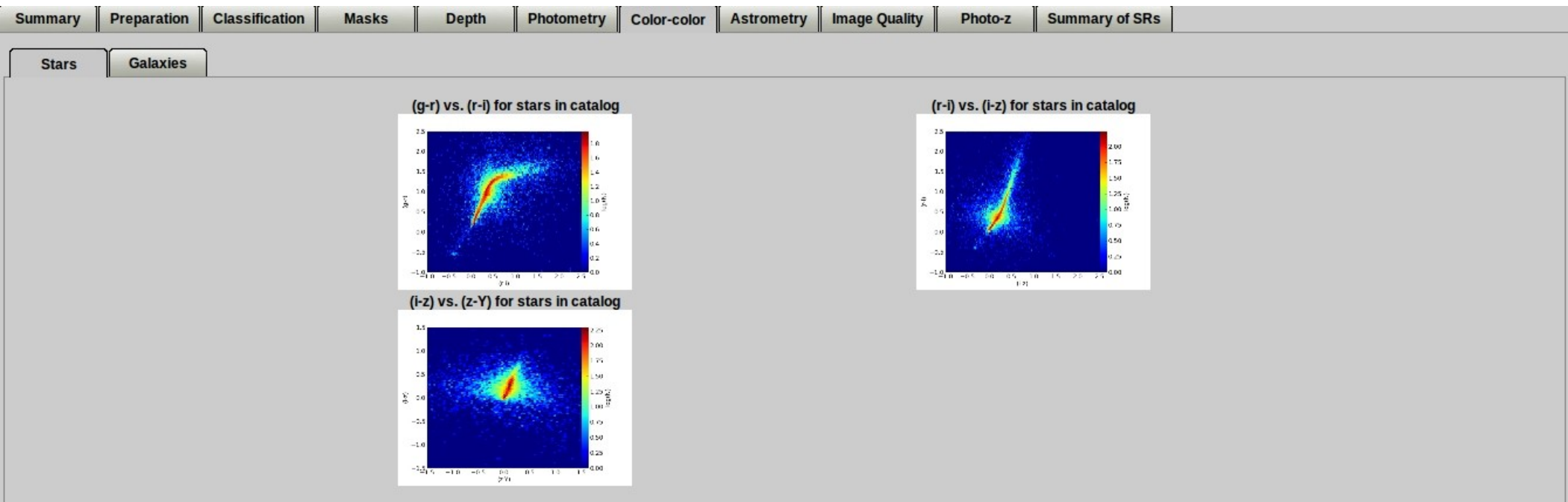


DARK ENERGY
SURVEY

Colour colour diagrams



DES-BRAZIL



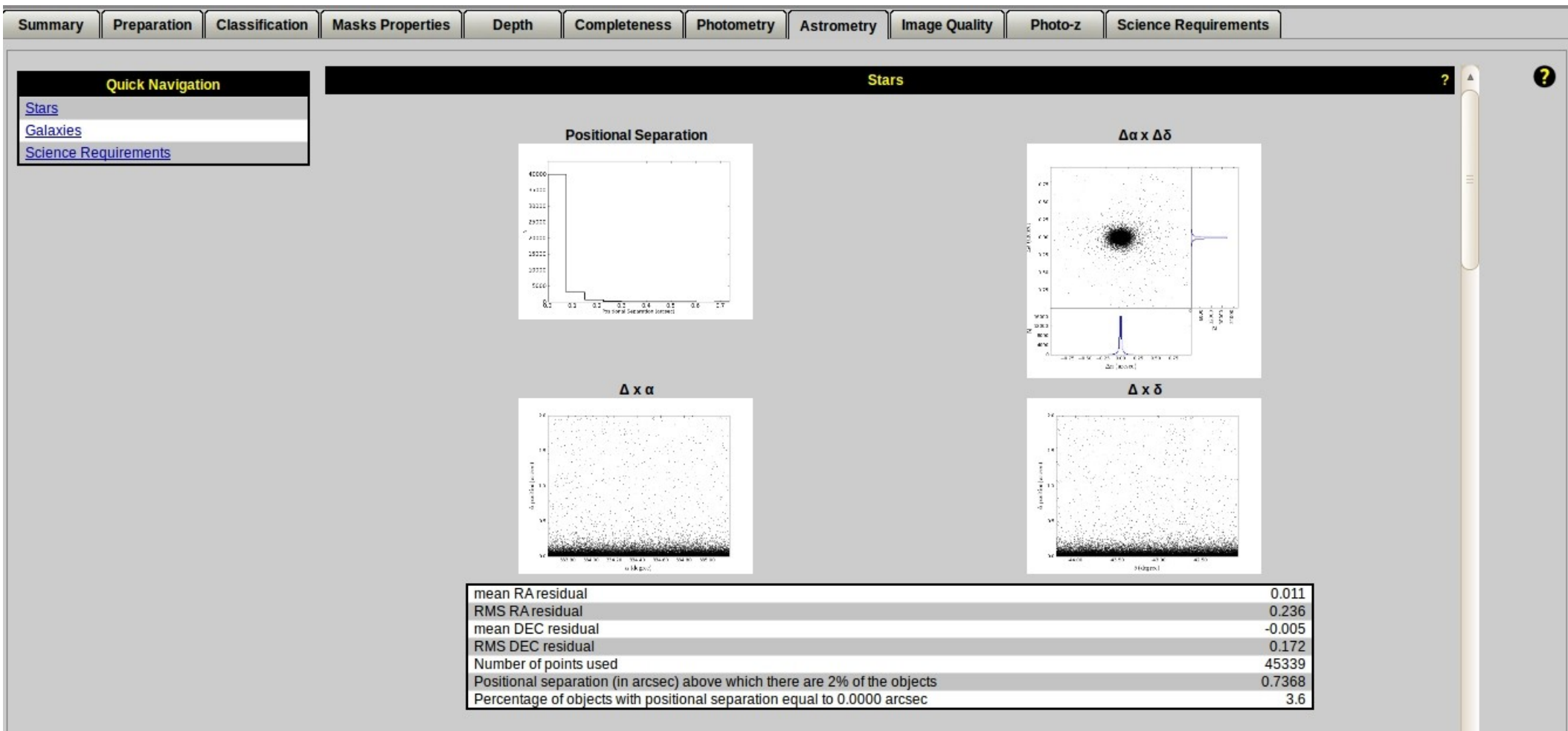


DARK ENERGY
SURVEY

Checking the Astrometry



DES-BRAZIL





DARK ENERGY
SURVEY

Astrometry SRs



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

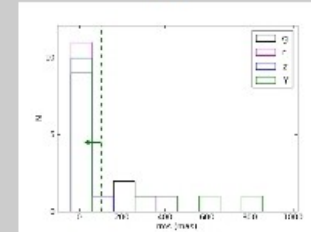
[Stars](#)
[Galaxies](#)
[Science Requirements](#)

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

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

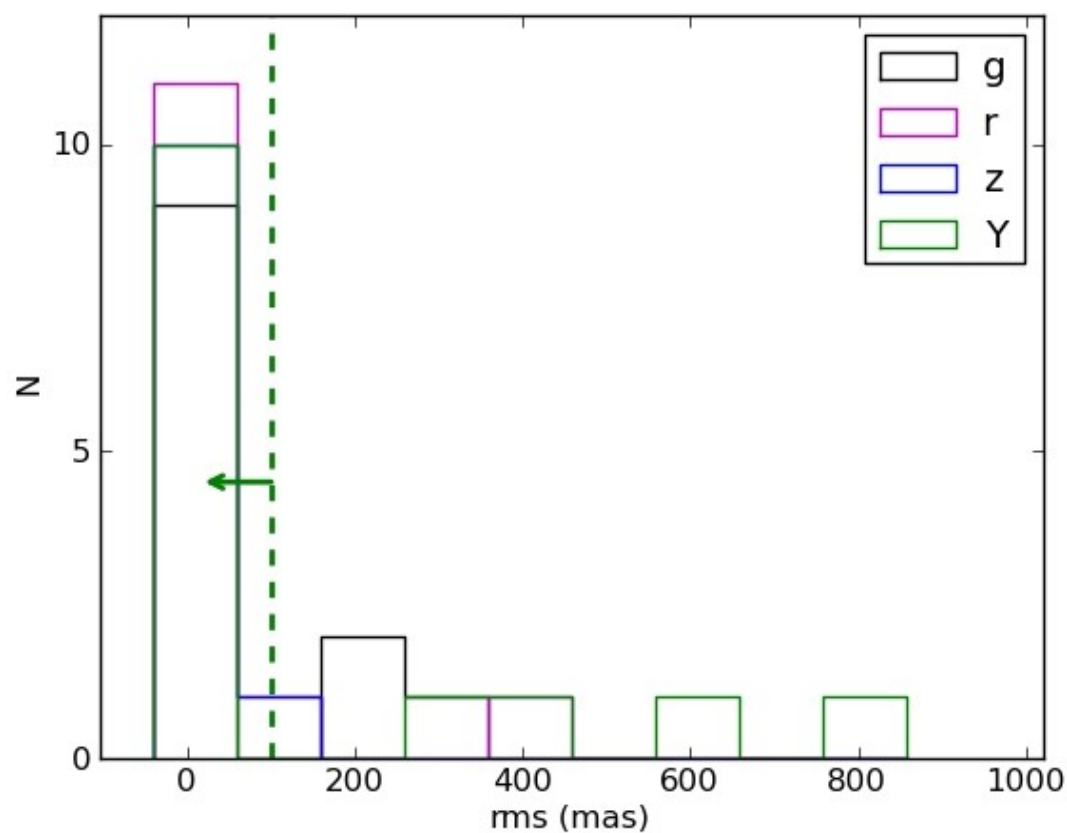


DARK ENERGY
SURVEY

Cross filter astrometry comparison



DES-BRAZIL



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.

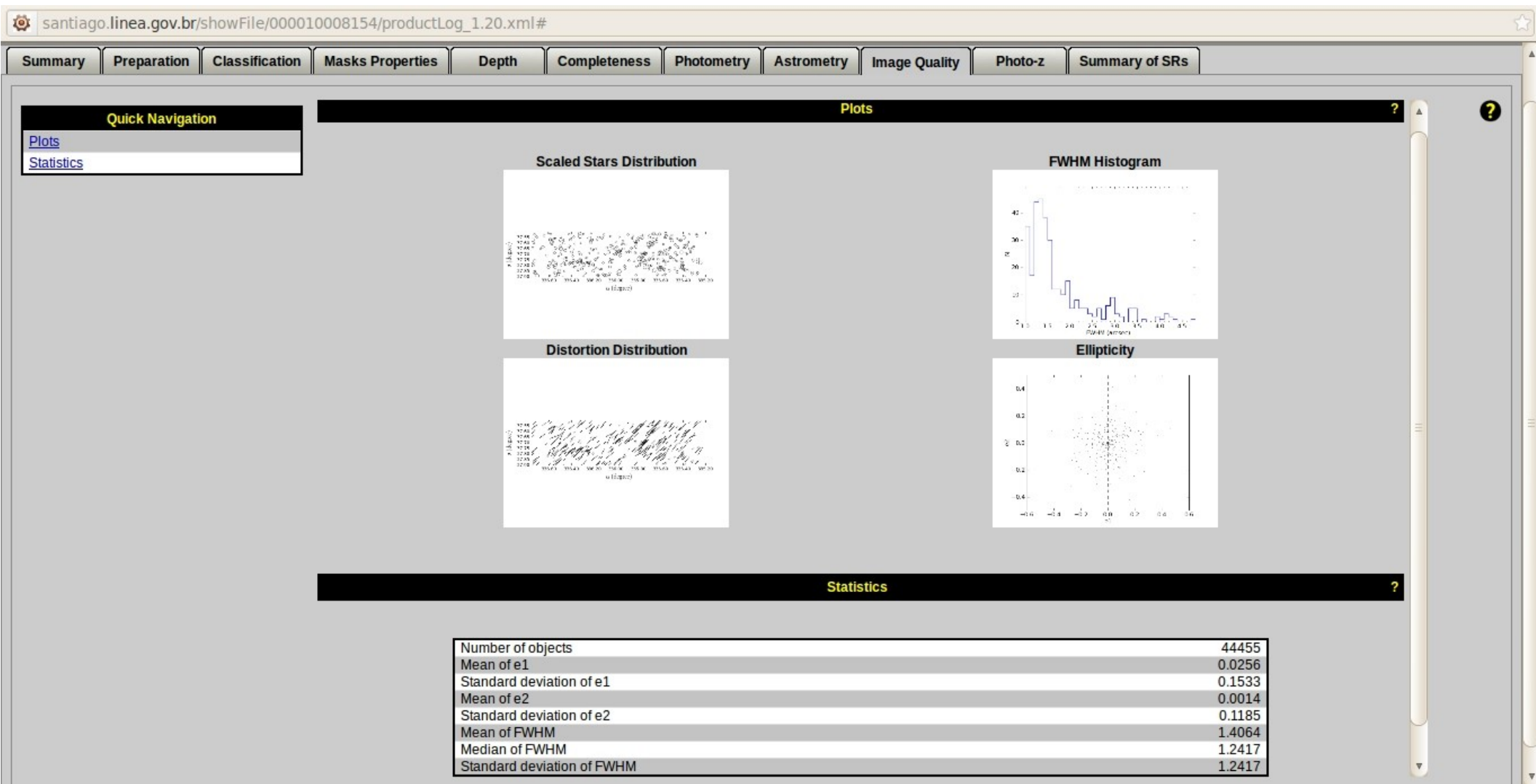


DARK ENERGY
SURVEY

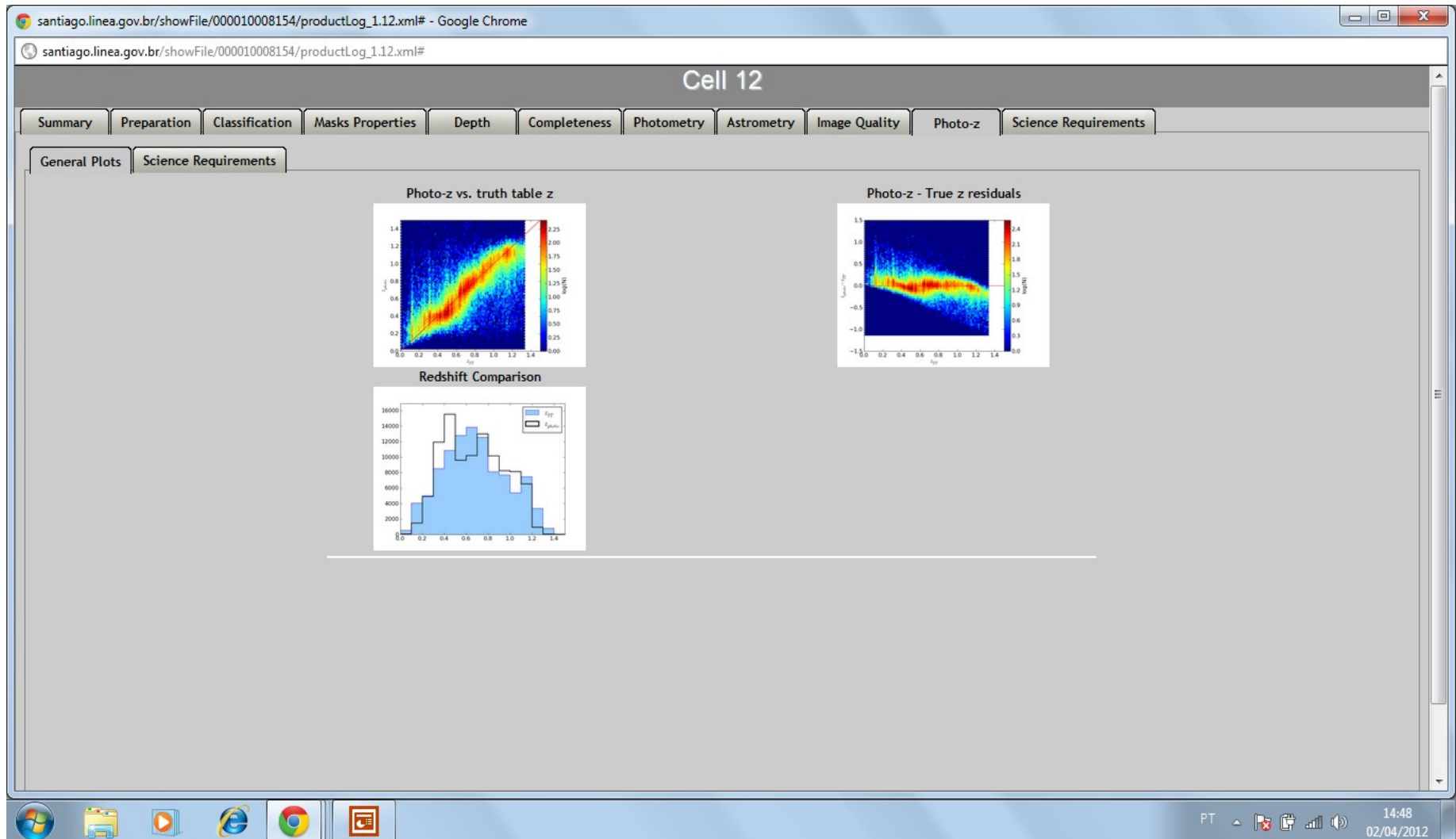
Image quality



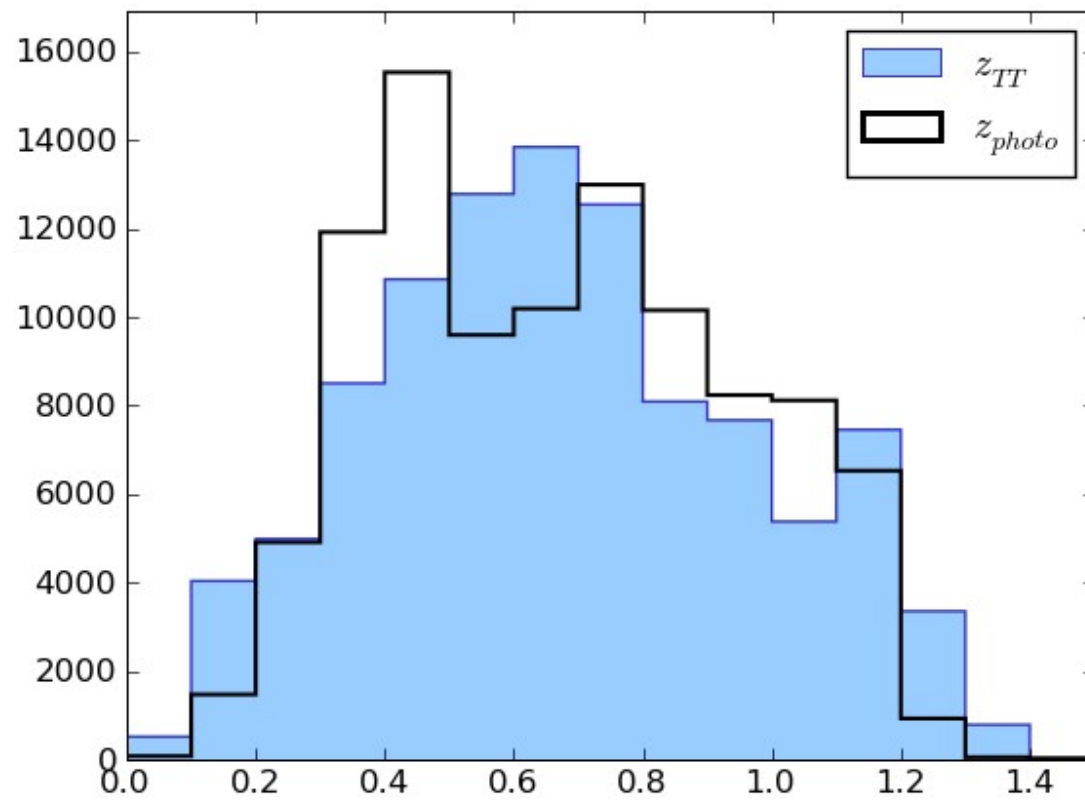
DES-BRAZIL



Photometric redshifts



True z vs Photo-z



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



DARK ENERGY
SURVEY



DES-BRAZIL

Summary of SRs

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

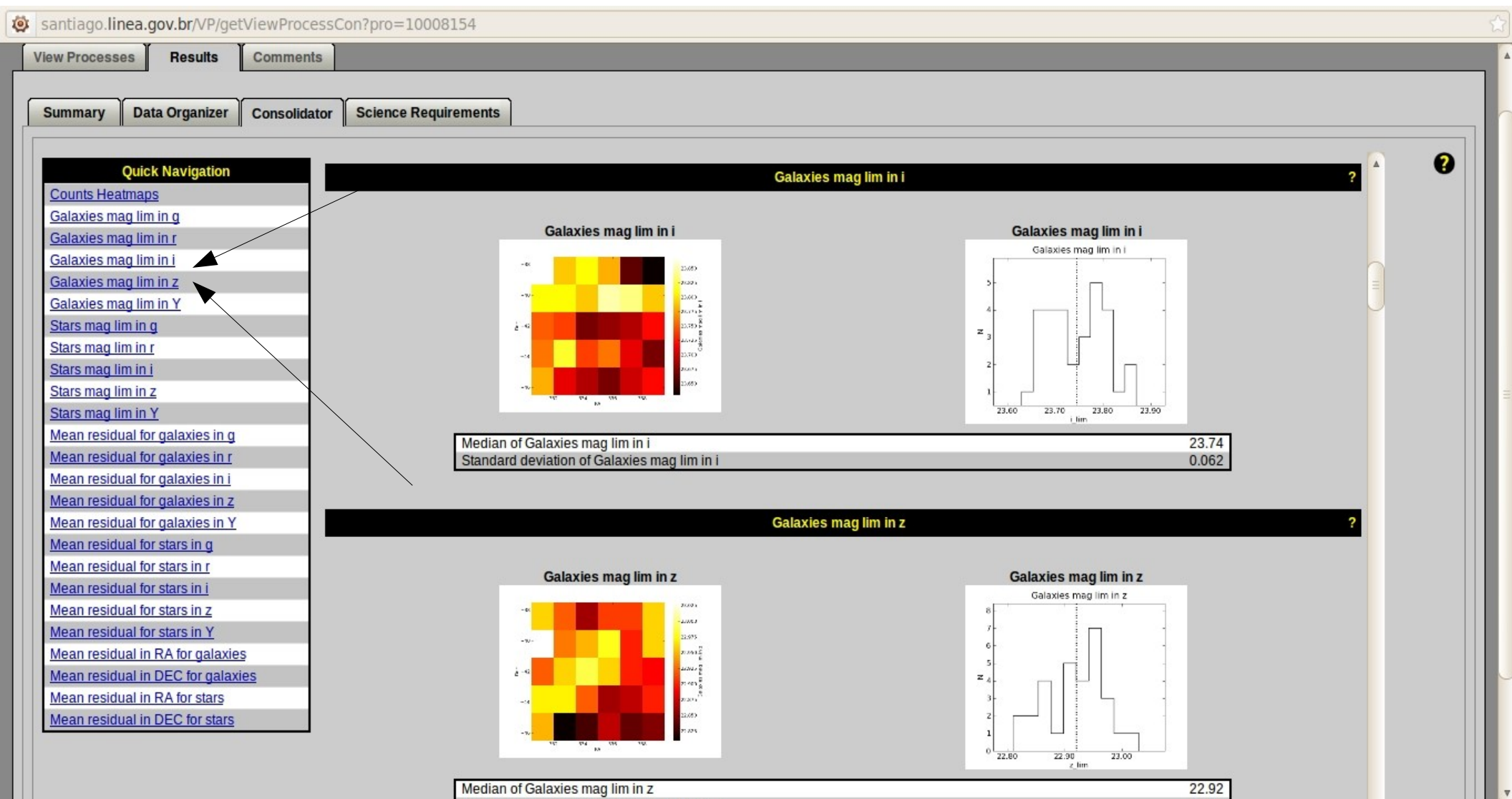


DARK ENERGY
SURVEY

Global QA analysis (Back to the Product Log)



DES-BRAZIL





DARK ENERGY
SURVEY

Global QA analysis (continuation)



DES-BRAZIL

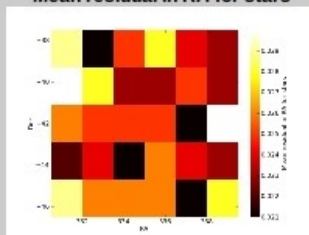
Summary Data Organizer Consolidator Science Requirements

Quick Navigation

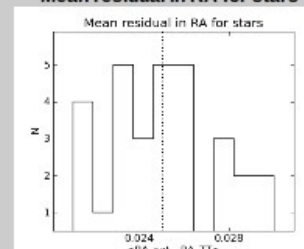
- [Counts Heatmaps](#)
- [Galaxies mag lim in g](#)
- [Galaxies mag lim in r](#)
- [Galaxies mag lim in i](#)
- [Galaxies mag lim in z](#)
- [Galaxies mag lim in Y](#)
- [Stars mag lim in g](#)
- [Stars mag lim in r](#)
- [Stars mag lim in i](#)
- [Stars mag lim in z](#)
- [Stars mag lim in Y](#)
- [Mean 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](#)

Mean residual in RA for stars

Mean residual in RA for stars



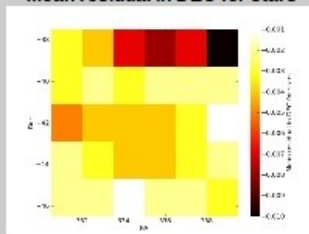
Mean residual in RA for stars



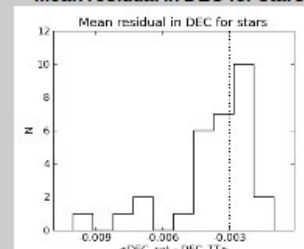
Median of Mean residual in RA for stars	0.025
Standard deviation of Mean residual in RA for stars	0.003

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



DARK ENERGY
SURVEY

Global SR analysis



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](#)

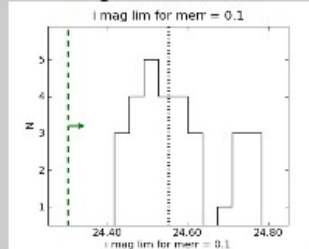
[R23.1](#)

[R27.0](#)

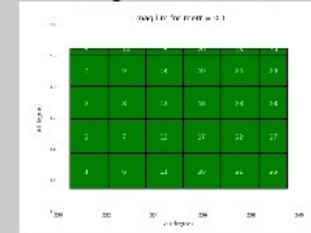
[R28.0](#)

R4.2

i mag lim for merr = 0.1



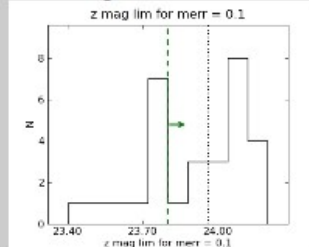
i mag lim for merr = 0.1



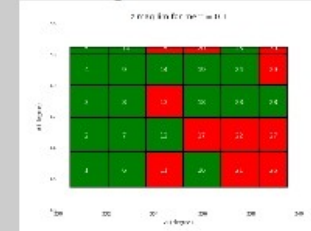
Average	24.58
Median	24.55
Standard deviation	0.105

R4.3

z mag lim for merr = 0.1



z mag lim for merr = 0.1



Average	23.97
---------	-------



DARK ENERGY
SURVEY

Global SRs QA (continuation)



DES-BRAZIL

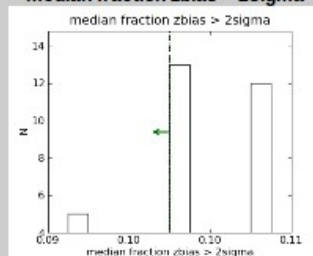
View Processes Results Comments

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](#)

median fraction zbias > 2sigma



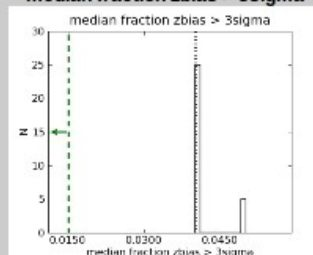
median fraction zbias > 2sigma



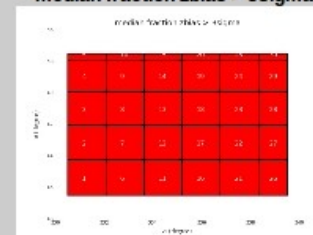
Average	0.102
Median	0.100
Standard deviation	0.007

R23.1

median fraction zbias > 3sigma



median fraction zbias > 3sigma



Average	0.042
Median	0.040
Standard deviation	0.004



DARK ENERGY
SURVEY

Closing remarks



DES-BRAZIL

- 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)