
Extraction of Properties from a Sample of Galaxies

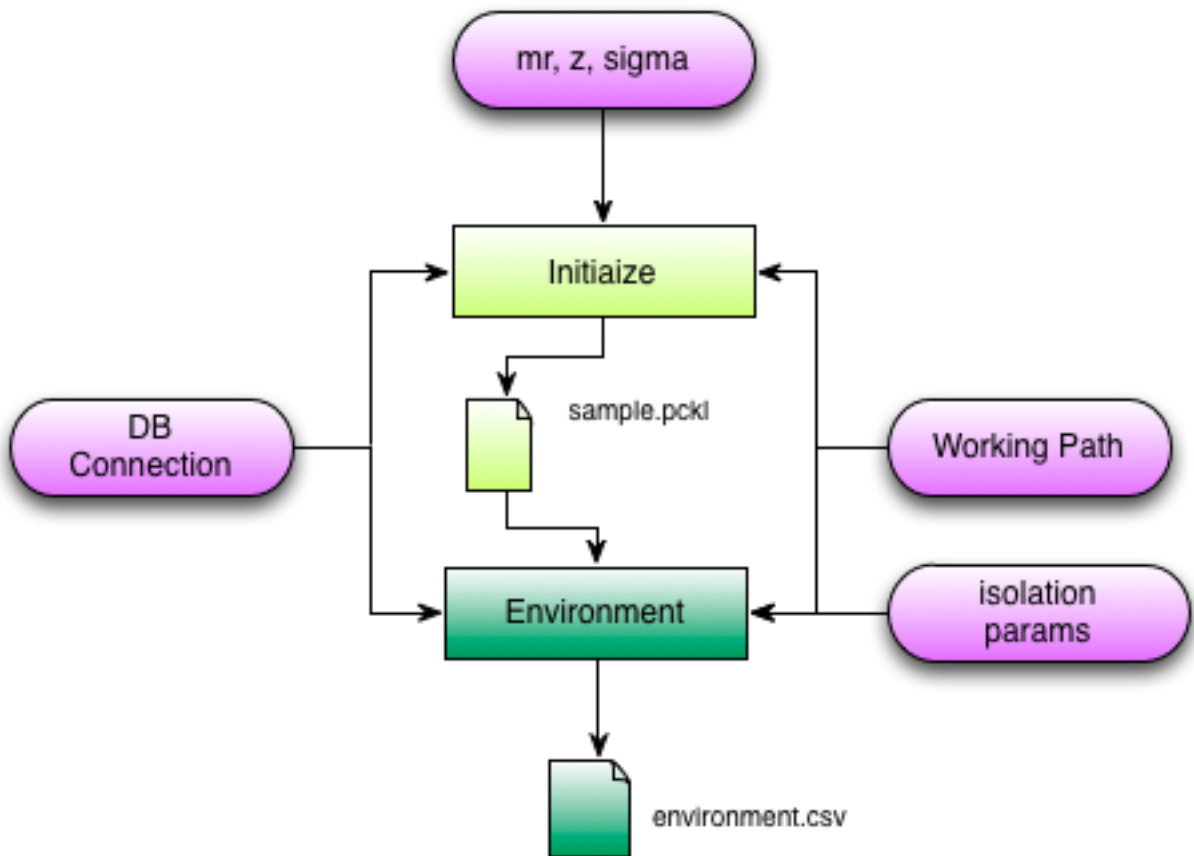


Table of contents

1. Materials and methods3

1.1 The SLOAN Digital Survey.....3

1.2 Extraction and characterization of a sample of galaxies3

1.3 Execution environment.....4

1.4 Workflows.....4

1.4.1 Initialize4

1.4.2 Environment.....5

2. Results6

2.1 Characterization of the environment for a sample of galaxies6

Appendix A – Graphical representation of workflows7



1. Materials and methods

1.1 The SLOAN Digital Survey

The Sloan Digital Sky Survey¹ (SDSS) is an ambitious project that consists on observing more than a quarter of the sky using the 2.5-meter telescope at Apache Point Observatory, New Mexico. Since data collection began in 2000, it has obtained multi-colour images of around 500 million objects, and spectra for more than 1 million objects. SDSS data have been released to the scientific community and the general public in annual increments, with the final public data release from SDSS occurring in August 2012. All the images, catalogues of measurements, and spectra are accessible online through web forms and web services querying efficient databases, and a lot of tools have been designed to browse through sky images, look up data for individual objects, or search for objects anywhere in the sky based on any criteria. SDSS data actually supports an enormous range of scientific investigations by astronomers around the world, and have become one of the main tools in astronomy research.

1.2 Extraction and characterization of a sample of galaxies

The formation and evolution of galaxies is related to the environment and the interactions with other galaxies. Galaxies form in slightly over-dense regions after the Big Bang. These proto-galaxies are the building blocks that assemble to form the galaxies that we can observe currently. The properties of a galaxy depend on its history of merging with companions and on the presence and intensity of interactions with companion galaxies. Galaxies located in dense groups or clusters are usually more massive than isolated galaxies and often show an elliptical shape (instead of spiral). On the other hand, galaxies that are interacting usually show signs of recent star formation and a higher prevalence of active supermassive black holes.

A careful quantification of the environment and interactions is fundamental to draw conclusions about the formation and evolution of the galaxy. A correlation between these environmental parameters and the properties of the galaxy can explain which of the latter ones are intrinsic to the galaxy (nature) and which are induced by the interaction with companions (nurture), however this is not an easy task. The quantification of the environment requires knowing the position of the companion galaxies to estimate its association with the target galaxy. It is not possible to measure accurately the real position of a galaxy in 3D space, but its projected position on the sky, and this makes very difficult to check if two nearby galaxies are really associated or not. The use of spectroscopy combined with the Hubble law allow to estimate the distance to a galaxy (with some error).

An approach to tackle with the nature vs. nurture topic is to use robust estimators of the environment and tidal forces (interaction with companions) in a big sample of galaxies. The SDSS provides an ideal sample to estimate these derived properties. Galaxies having a redshift between 0.03 and 0.1 covered by the SDSS spectroscopic survey compose the sample to be studied by this first RO. A catalogue of potential nearby companions is obtained for each target galaxy. The properties of the potential companions were

¹ <http://www.sdss.org>

obtained from the SDSS photometric survey public archive. The workflows developed evaluate in a first step the real or fake association of a potential companion with the target galaxy, based on its optical properties (magnitude in r-band, distance to the target galaxy, spectroscopic redshift and photometric redshift). After that, two environmental estimators are derived for each target galaxy, taking into account only the companions that have been inferred to be really physically associated. The first parameter accounts for the density of companions of similar mass in a region around the target galaxy and the second parameter estimates the gravitational forces exerted by these companions. The results allow the study of the correlation of these parameters with other properties of the galaxies.

1.3 Execution environment

The first requirement to run the workflows provided by both ROs is Taverna Workbench² 2.4 or higher. AstroTaverna³ (Taverna plugin) is also needed in order to get functionalities related with Virtual Observatory web services queries and management of standard VOTable data formats.

In general, the execution environment is a Linux distribution including Python⁴ 2.x and a bash shell. Access to a PostgreSQL⁵ database storing the physical parameters provided by SDSS is also needed; a dump file of database may be downloaded from the AMIGA web server and in order to be deployed and accessible from a local execution environment. In particular, the following list shows other specific dependencies for the main workflows:

- Initialize
 - Python with *psycopg*⁶ package
- Environment
 - Python with *psycopg* and *numpy*⁷ package
 - AstroTaverna

1.4 Workflows

1.4.1 Initialize

This workflow saves a tabular *.pck/ Python pickle dataset in the local file system, containing values calculated on physical parameters associated to potential companions of a sample of target galaxies. These original physical parameters are extracted from a postgresQL database, containing information of all galaxies covered by the SDSS spectroscopic survey. The workflow first access the external database

² <http://www.taverna.org.uk/download/workbench/2-4/>

³ <http://wf4ever.github.com/astrotaverna/>

⁴ <http://www.python.org/>

⁵ <http://www.postgresql.org/>

⁶ <http://initd.org/psycopg/docs/>

⁷ <http://numpy.scipy.org/>

located in the AMIGA server and selects the target galaxies from the sample (those having spectroscopic redshift between 0.03 and 0.1). It then creates a tabular gridded datacube with values associated to potential neighbours. These values are calculated for each point of a 3D space defined by the axes: *magnitude in r band*, *photometric redshift* and *sigma level of detection*. The input default values to build the parameterised datacube are:

- $14.5 < m_r < 22.5$ - step 0.5
- $0 < z < 0.11$ - step 0.01
- $0.1 < \sigma < 3.2$ - step 0.2

Auxiliary function libraries and scripts are also copied in local file system, and the PYTHONPATH environmental variable is set to a value provided by the user as the Working Path of the digital experiment. Other user provided input values are the **database** connection settings: hostname, login and password.

1.4.2 Environment

This workflow takes as input the path of the tabular **.pckl* Python pickle dataset created in the previous workflow, as well as the database connection settings and several criteria on how to filter the potential companions of the target galaxies. It provides a file with the SDSS identifiers of each target galaxy of the sample, environmental estimators and radius where the 10th companion has been found. The workflow looks for potential companions in radius ranging from 3Mpc to 11Mpc, with a step of 1Mpc. The user may modify these numbers at the input stage, as well as several limits and ranges needed in the filtering process. As in the previous workflow other provided input values are the Working Path of the digital experiment and the **database** connection settings: hostname, login and password.

2. Results

2.1 Characterization of the environment for a sample of galaxies

Two environmental parameters were derived using the Sample Selection RO. The first parameter traces the density of local companions and the second one traces the tidal forces exerted by the companions. The estimation was obtained for a sample of ~300000 galaxies with redshift between 0.03 and 0.1 covered by the SDSS spectroscopic survey.

The relation of these parameters with some properties of the galaxies was studied. A correlation between the density and tidal parameters and the presence of an Active Galactic Nuclei (AGN) was found. Galaxies located in denser environments (high value of the density parameter) present less prevalence of AGN selected based on the properties of the optical spectrum, but a higher prevalence of radio jets in their cores (radio AGN). On the other hand, a higher value of the tidal estimator is correlated with a higher prevalence of radio AGN and optically selected AGN.

Our interpretation of the results is as follows. Galaxies in denser environments do not have the supply of cold gas that is required to power optically selected AGN. However, warm gas can power radio AGN in these dense environments. On the other hand, all types of close interactions with companions (traced by a higher value of the tidal estimator) produce an enhancement on the prevalence of all types of AGN. The interaction with companion galaxies can fuel the gas to the centre of galaxies where is accreted by the central supermassive black hole producing both radio and optical AGN.

Appendix A – Graphical representation of workflows

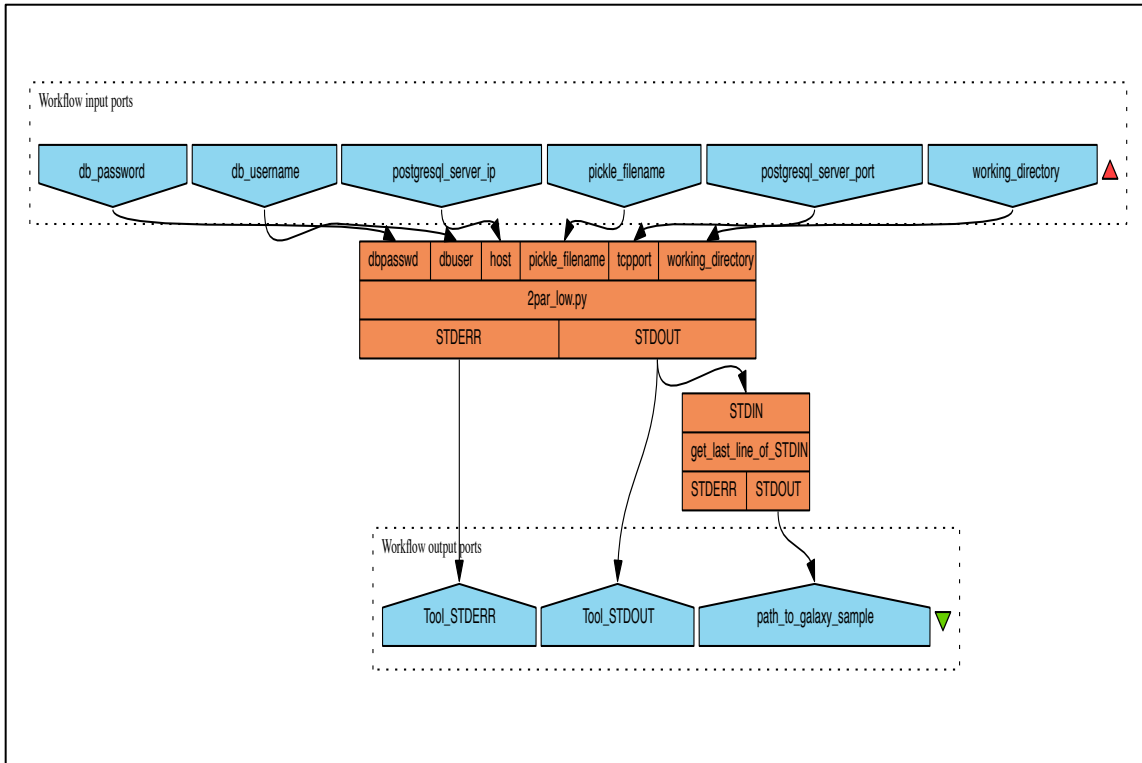


Figure 1: Initialize Workflow

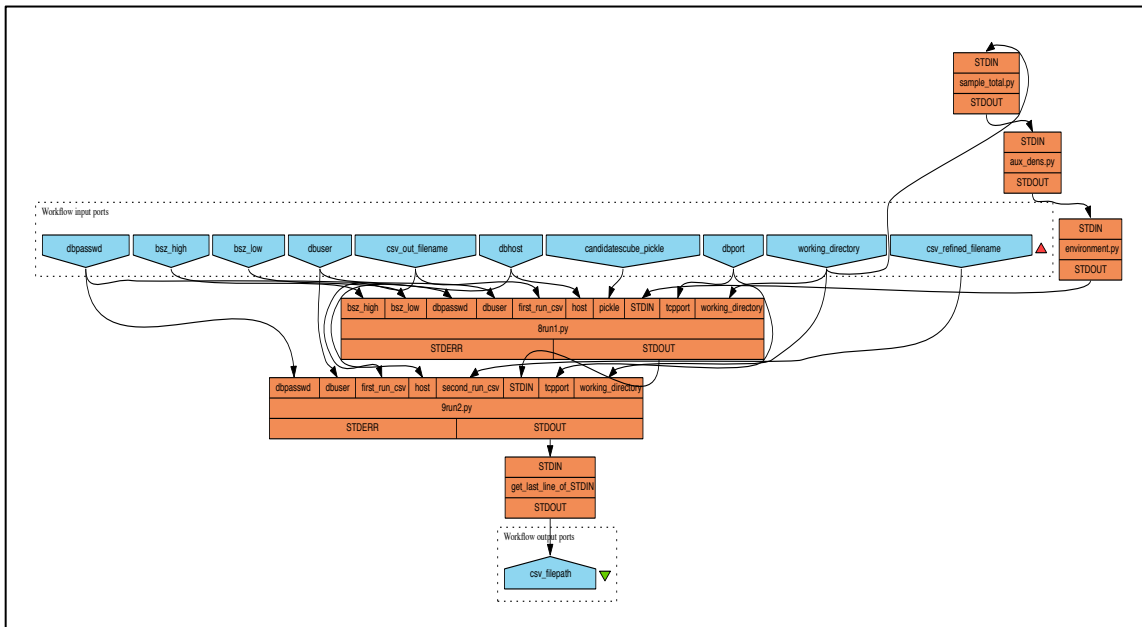


Figure 2: Environment Workflow