# Working with the SDSS images (Part 1):

(The following is valid for the Data Release 7.)

1) In the page http://das.sdss.org/www/html/post_coords.html, put the coordinates of the galaxy. Note that the coordinates must be given in degrees and in decimal format. For example, in the case of NGC3245, the coordinates are 156.826633,28.507378 instead of 10h27m18.4s +28d30m27s. Don't leave blank spaces. Click "Submit request".
To get the coordinates of a galaxy, use the NED database (http://nedwww.ipac.caltech.edu/forms/byname.html) and use always J2000 coordinates.

2) After "Submit", a new panel will show the image parameters. For our example,
run, rerun, camcol, field
5079, 40, 2, 45

(To know the meaning of these parameters go to http://www.sdss.org/dr7/algorithms/dataProcessing.html.) Clicking the link it is possible to go to a page where all the available images are.

3) In this case, select the files drField-005079-2-40-0045.fit (this file has the observing and calibration parameters), fpC-005079-r2-0045.fit.gz (r-band), fpC-005079-g2-0045.fit.gz (g-band) (and any other desired bands) saving the files in a directory with a name allusive to the galaxy. For this galaxy, the directory name should be "ngc3245". This directory would be inside another directory with a name as "sdss". (Naming the directories in this way is useful to build scripts.)

4) Unpack the gz files with "gunzip" and rename the files. In this case, suggested names would be

fpC-005079-g2-0045.fit ---> n3245g.fit <--- take care of the letter that indicates the filter
fpC-005079-r2-0045.fit ---> n3245r.fit
drField-005079-2-40-0045.fit ---> drField-n3245.fit

maintaining the files in the directory .../sdss/ngc3245/.

5) It is possible to display the images with the DS9 utility. It's important to show one of the images to verify that the galaxy is not in the border of the image. If the galaxy is in the border, it is possible to merge two or three images, but this will be described in a more advanced guide. Move the mouse pressing the right button to change the colour scaling. Observe that the star halos extend away the star centre, especially when the star is very bright. The case of NGC 3245 is easy because there are not very bright stars in the surrounding. Pick a multi-colour scheme, as the SLS scale, to enhance the low levels of the image.

6) Now it is time to "mask" the galaxy. For that, first could be convenient to get a median-smoothed image. To make a median-smoothed image, use the "median" task in IRAF, with x and y sizes of, for example, 9 pixels.

e.g.
cl> median n3245r.fit n3245rsmooth.fit 9 9

Display the median-smoothed image in DS9, and then pick a multi-colour scheme as SLS and play with the contrast to see how far out some of the stellar halos extend. Over this image mark some regions with the left button of the mouse enclosing the bright stars and their halos in the images, as well as other bright zones in the image. Use circles, ellipses, boxes, and polygons (Region-->Shape in DS9) where needed, trying to optimize the enclosed zones. The masks should be large enough to enclose the bright zones, but small enough to leave useful space around the galaxy. Save the regions using the "Ciao" format (Region-->File format-->Ciao; Region--

>Save Regions). It is convenient to give also to the regions file a name indicating to which galaxy is associated with. In this case, a good name could be n3245r.reg.

7) The pixel values of the image we get from the SDSS archive are coded as 16-bits integers. It is necessary to convert those values to floating point. This can be done simply multiplying the original image by 1.0. In IRAF, for this case

cl> imarith n3245r.fit * 1.0  n3245r.fit

8) Now, we have here some notes about sky subtraction, focused on SDSS images.
What we are trying to do is two things:
a. Remove the sky background from the image, first flattening it if it isn't flat to begin with.
b. Record the distribution of sky values for later analysis -- in particular, we want to (eventually) compute an uncertainty for our sky background measurement.

The SDSS images are easier than other images because they are almost always flat: that is, the sky background level does not change (aside from pixel-to-pixel noise variations) from one side of the image to another.  So we only need to measure the mean background level and subtract that number from the image.

The first thing to do is to inspect an image to see if the background is flat, or if there are some large-scale variations that will have to be removed.  The way to do this is displaying the smoothed  image and look for possible large-scale variations.  This can be done using a multicolour colour map in the image display program as was indicated before.

For the median-smoothed images use the r-band images.

Once we have an image with a "flat" background (we'll discuss later how to deal with non flat backgrounds) we need to measure the overall background level.

There is no one "best" (or at least agreed-upon) way to do this; the way we have been doing it for these images seems to work well, and is described below.

The basic idea is to make many local measurements of "clean" regions of the image using the IRAF task IMEXAM.  The IMEXAM task can print a line of statistics for a square subsection of the image using the "m" key; the value we will use is the median.  (Why the median?  Because it is generally better at ignoring the effects of extreme outliers, such as cosmic-ray hits or bad pixels, than the mean.) The total set of outputs from IMEXAM-m are saved to a text file, and the MEAN of all these median values is then our estimate of the sky background.

IMEXAM-m calculates statistics in a box centred on the cursor; the size of the box is defined by the "ncstat" and "nlstat" parameters.  We usually use:
ncstat = 10
nlstat = 10
This values can be changed using

cl> epar imexam

This defines the statistics region as a 10x10 box.

The basic approach is to display the image (iraf DISPLAY command), then call IMEXAM and move the cursor around on the image, looking for relatively clean regions of the image and typing "m" to get statistics. For SDSS images, we usually get something like 50-120 separate measurements. The statistics lines are printed out in the xgterm window; but we can save automatically the results calling IMEXAM as cl> imexam logfile=filename_nxxxxy.dat

(in this suggested mane nxxxx is the galaxy name and r is the filter; the exact name doesn't matter, except that it should reference the galaxy and the specific image used, so that there is no confusion later).

What are "relatively clean regions of image"? Basically, we need to sample the dark parts of the image, away from any bright stars and away from any galaxies.

It is critically important to note that we are able to detect galaxy light out to about twice the D_25 diameter (that is the diameter listed in RC3 and in NED), so measurements have to be made at least that far away from the galaxy. (If the galaxy is high inclined, so that its appearance is rather elliptical, then obviously we can approach closer on the semi-minor axis).

The median-smoothed image is sometimes a useful reference, since it can show how far the halos of bright stars extend, or where a faint background galaxy is, or the existence of regions where flat-fielding or sky flattening did not quite work. But DON'T use the median-smoothed image to do the statistics with IMEXAM. Use always the original non-smoothed image.

9) It is possible to calculate the mean of all the median values with several tools, like IDL, EXCEL, etc. But a Python script written by Peter Erwin is enclosed to do this. This script is "skyval_simple.py" that must be put in a directory included in the Python path. If the script is in a new directory, it is necessary to include this in the path. (For some systems, this can be done including a line like

export PYTHONPATH=/home/username/progs/python:/usr/lib/python

or

setenv PYTHONPATH /home/username/progs/python:/usr/lib/python

in the shell profile file: .bashrc, .cshrc, or any other.)

To calculate the mean of the median values, type

$ skyval_simple.py tempsky.dat

This script optionally computes a bootstrap-resampling error on the mean. To use it:

$ skyval_simple.py tempsky.dat --bootstrap

Currently, we are estimating the uncertainty by using bootstrap resampling, which means recalculating the mean many times using subsets of the median values.

The "normal" standard deviation tells the distribution of 10x10-pixel-box median values about the mean of those values. This is what we would use as the "error" on the sky background if we were doing a measurement on another (single) 10x10-pixel-box. However, that is not the case; we are trying to determine the mean sky value for the entire image, and we want to know how accurate that mean value is. For this, something like bootstrap resampling is a reasonable approach.

10) To subtract the background use the IRAF task IMARITH.

cl> imarith nxxxx.fits - 1234.56 nxxxx_ss.fit where "1234.56" must be replaced by whatever is the true sky value, and "ss"is a reminder that this is a "sky-subtracted" image.

If this image will be used as input to the ELLIPSE task it is necessary to specify ".fit" as the filename ending, because ELLIPSE knows to look for a mask image only if the main image ends in ".fit" instead of ".fits".

11) Now we need to convert the mask regions file in a file understood by ELLIPSE. A useful task for this is PLCREATE. This task is from an old IRAF edition; so it is enclosed here in the tar file plcreate_redhat.tar.gz. To install it, use the following steps:
- copy this file to the iraf directory, and go there;
- unpack the file using
  $ tar -zxvf plcreate_redhat.tar.gz
  this will create a new directory with the files plcreate.par and x_ximages.e.
- Then, in the configuration file login.cl (or loginuser.cl), include the lines:
  reset plcreate = home$plcreate_redhat/
  task plcreate = plcreate$x_ximages.e

12) Now we are ready to convert the ".reg" files to ".pl" files. For this, use in IRAF the following syntax:

cl> plcreate n3245r.reg "n3245r_ss.fit" n3245r_ss.pl

It is important to enclose the image name with quotation marks. And it is important to assign to the ".pl" file the same root name as the ".fit" file.

In a second part, we will talk about how to apply the ELLIPSE task.