WHISP
GIPSY-Phase Reduction Cookbook

Dolf Sijbring

May 8, 1998

Contents

1 Prologue to the GIPSY Reduction Phase

The objective of the WHISP project is to observe a large number of carefully selected galaxies in the 21-cm line with the WSRT, to reduce the data in a uniform way, and to store the results in a database for general use.
In this cookbook I will describe the procedures by which the WSRT data are reduced, analyzed and archived using the WHISP cola scripts. I do this by following the steps in the reduction with the user's point of view in mind. I will refer to the reduction process that takes place in the GIPSY environment as the *xy-reduction phase* of the WHISP process.
This cookbook can be found in the HOWTO directory of the whisp account at the Kapteyn Institute in Groningen.

2 Starting up an xy-reduction session

2.1 First step

The first question you should ask yourself is: is there a set of 6 FITS-files available reduced by the uv-reduction phase of WHISP. These 6 FITS-files are the raw map cubes at three resolutions (full, 30" and 60") with their corresponding antenna-patterns. These FITS-files can be either on tape or on disk. The WHISP Database contains the FITS-files on tapes labeled WF####, where #### is a number. The first 6 files on these tapes are the 6 FITS-files. At the end of the GIPSY data reduction phase we will be adding our results to the same tape, but this time in the format of two TAR-files.
You can search for uv-reduced FITS-files on the WHISP-site (url: www.astro.rug.nl/~whisp/Database/DatabaseStatus/Reduction/UV-plane/make_uv_list.cgi), but, most likely if it is there, it will also be GIPSY reduced. You can find that out by looking in the Gipsy Reduced List (url: www.astro.rug.nl/~whisp/Database/DatabaseStatus/Reduction/Image-plane/make_xy_list.cgi).
The most common situation will be that you have reduced a galaxy observation yourself in the NEWSTAR (= uv) phase, and now you want to continue with the GIPSY reduction phase.

2.2 Beginning

When there are raw image datacubes available in three resolutions with their antenna-patterns, you can now start the datareduction in the GIPSY phase.
To do this you login under the whisp account. The whisp account contains the WHISP-site (with database-administration and the WHISP Atlas), the necessary reduction environment (with scripts) and disk space for the reduction of a galaxy observation.
After you logged in you should go to the xy directory under the whisp directory, i.e. type: cd whisp/xy. You can not define your own directory here as in the case of the uv-reduction, since the reduction needs certain (sub)directories for it's processing. When you have finsished your reduction the directory will be emptied again for the next observation.
A matter that needs a bit of consideration is when you have logged in via rlogin. In that case you will have to set the DISPLAY environment variable, otherwise the GIDS display can not be started.
Now that all preliminaries are set you can start up GIPSY.

3 The general build-up of the xy-reduction process

The xy-reduction process in WHISP consists of four general phases,

1. preparing for cleaning (interactive)
2. finishing cleaning (batch)
3. analysis (interactive)
4. adding the reduction of the observation(s) to the WHISP database

of which the first three phases are done in GIPSY. The astronomer is led through the reduction by starting up some (master)scripts, that will present him/her with questions and graphical displays, that will enable him/her to reduce the raw radio observations in a manner that he/she sees fit and at the same time riddens him/her from the burden of keeping a detailed administration of all the steps that are taken. *This way the reduction process can be speeded up by a factor 20 or more.* It also enables the astronomer to reduce far more observations than was possible ever before, by lowering the psychological stress considerably. The WHISP scripts bring the position of the radio astronomer on a par with that of the optical astronomer.

It should be remembered, however, that the scripts were meant to be used within the framework of the WHISP program and not for a general usage, therefore some scripts are dealing with extracting information for the WHISP database.

As mentioned before, the astronomer starts up some master-script that will lead him/her through the reduction process. These master-scripts are simply modular cola-scripts. Every module executes a functional task in the reduction process. A functional task is a conglomerate of tasks that, in the eye of the script author, belong together. There are 23 functional tasks distinguished in the xy-reduction phase and they will be described seperately in the next section.

In the middle of the reduction process you can stop, either because you made a mistake or because you want to go home. The progress of the reduction is recorded in a file with the name: u<ugcnumber>.story. It simply records the numbers of the tasks that have been completed. If you stop the reduction process by typing <masterscript-name>, stop=y, you will stop after a functional module. When you restart the masterscript the process will continue with the next module. If you stop within a functional module, the process will redo the module after restart of the master-script. You can edit the story file to start at every, meaningfull, position in the reduction process by putting the number of the last task before you want to resume in the story file.

Let's go to the real reduction process now.

4 Notes about the individual modular steps

After the astronomer has completed the aforementioned steps, he/she can now start the WHISP xy-reduction process.

4.1 Preparing for cleaning

The first masterscript is concerned with the preparations for the cleaning of the raw WSRT observations. This process is started up by typing:

masterxy_step1

The cola-file will ask for a UGC number, e.g. 11861. After this is given each functional module will be started up in turn. In the following I will shortly describe what each functional task comprises and note some things about the required input from the astronomer.

- XYLOAD: the first functional task loads the raw WSRT observations in three resolutions (full, 30" and 60") with their corresponding antenna-patterns, either from tape or from disk. It checks whether the UGC number given by the astronomer corresponds to the UGC number in the header of the FITS-files. It also will show if there are tar-files following the FITS-files on tape, meaning that the observations have been reduced before.
- XYRAWINSPECT: this task will display first the three antenna-patterns to give an impression of the uv-coverage of the observations and whether any cleaning problems can be expected. Then it will load the entire datacube at 60 arcsecond resolution in GIDS. In most of the times the WHISP scripts will load maps with a color wedge running from -4 to +5 r.m.s. noise. The r.m.s. noise at this moment is only a 'guessed' value, but will be changed to the real value later on. This display can be used in the next step to determine the extend of the HI emission in velocity. In the background it builds up a list of galaxies from the PGC (in the file list.gal) that are within a distance of 17.5 arcminutes from the pointing center. This list will be used later for blotting, the global profiles and xv-diagrams. Also an optical image from the PSS with dimensions 35' x 35' is prepared.
- XYCONTSUB: in this task we make the radio continuum subtracted maps. It will first ask whether there is any foreground HI emission (from our own Galaxy) visible in the cube. If not, the highest and lowest channel numbers, used for the definition of the radio continuum subset range, are calculated and only the inner channel numbers have to be supplied by the astronomer. He/she can do this by making a slice through the GIDS datacube and then placing the cursor over the slice image to determine which channels are *the upper boundary to the low velocity side continuum* and *the lower boundary to the high velocity side continuum*. Just use the GIDS values for the channel numbers, they will be recalculated to GIPSY subset values. Now the range of continuum channels is detemined and the continuum emission can be subtracted.

If there is foreground emission also the extreme velocity ends of the continuum range will have to be defined.

There is a check possible at the end by subtracting the mean continuum channel from the low velocity side from the mean continuum channel from the high velocity side. No significant differences should remain.

- XYBLOTCH: in this task we will define the first order masks at 60" resolution that will be used for later cleaning. First the datacube is Hanning smoothed to improve the S/N ratio in the maps. Also all galaxies from the list of galaxies from the PGC that are in list.gal and within or near the velocity range of the observations will be displayed as crosses on the maps to be blotched. The datacube will be displayed in GIDS one by one subset and can accordingly be blothched, i.e. one has to selected the mask areas.

- XYCLEAN60HAN: this task will clean the Hanning smoothed 60" resolution maps. Since there are only 5 antenna-patterns per datacube, the antenna-pattern closest to the range of subsets to be cleaned will be used. The cleaned datacube is shown in GIDS.

- XYRESTORE60: here the clean-components will be restored after a beam has been fitted to the antenna-pattern. This beam should be near 60" x 60". The restored datacube is shown in GIDS.

- XYFINALBLOTCH: in this task we will define the second order masks at 60" for the cleaning process. To do this we first clip the cleaned and restored Hanning smoothed maps at a level of 1.8 times the r.m.s. noise level. Then the maps are displayed for blotching again. Only the areas that show clear HI emission should be selected.

- XYCLEAN60HAN: again the raw Hanning smoothed maps are cleaned using the second order masks defined in the previous step.

- XYRESTORE60: and again the clean-components are restored to the cleaned maps. If you are still not happy about the cleaning you can redo the definition of the maps as many times as you like. But if you're happy you can continue.

- XYCLEAN60: now the uniform tapered 60" maps are cleaned and displayed in GIDS.

- XYCONTSUBREST: here the radio continuum emission from the full and 30" datacubes is removed, using the settings defined for the 60" resolution maps.

- XYMASKS: in this task the masks defined at 60" resolution are regridded to be fitted to the full and 30" resolution datacubes. After this task has finished you will be asked whether you want to continue with the cleaning of the full and 30" resolution datacubes. If you type Y or give a return the cleaning will be done in batch mode. You can leave GIPSY now if you want.

4.2 Finishing the cleaning

This second phase is already started up by typing Y at the end of the previous phase in the reduction. It is run totally in batch mode and can take from 15 minutes to 2 hours, depending on the size of the galaxy and the speed of the computer. The task of this phase is to clean the full resolution and 30 arcsecond radio maps of the observation. In the third and next phase the clean components of the three resolutions will be restored.

4.3 Analysis

The third masterscript is concerned with finishing the cleaning process of the raw WSRT observations and analyzing the results in a first order manner. This process is started up by typing:

masterxy_step3

The cola-file will again ask for a UGC number, e.g. 11861. First it will check whether you have completed the cleaning process. After this each functional module will be started up in turn. In the following I will shortly describe what each functional task comprises and note some things about the required input from the astronomer.

- XYRESTOREREST: this task will restore the clean-components to the three resolutions after fitting a beam to the antenna-patterns. It also will display every datacube in turn in GIDS. The cleaned and restored datacubes only contain the maps that are outside the range of maps used to define the radio continuum emission.

- XYRADIOCONTIN: in this task all channel maps used to define the radio continuum emission are averaged to make a radio continuum map of the observation. This is done for all three resolutions. These maps will be placed in the 2-dimensional analysis cube later on.

- XYGLOBALPROF: this task will extract the information for the HI global profile from the datasets. It will place this information in the files flux11861.dat for the full resolution if the UGC number is 11861, flux11861s30.dat for the 30”, flux11861s60.dat for the 60” and flux11861s60han.dat for the 60” Hanning smoothed data. The files contain 5 columns with the velocity of the channel, the primary beam corrected (PBC) flux in the clean-components, the primary beam UNcorrected residual flux, the primary beam UNcorrected dirty flux and an estimate for the noise in the flux values. This estimate is calculated by multiplying the squareroot of the number of beams in the mask area by the noise per beam area.

  Now an important issue arises. Sometimes more than one galaxy can be distinguished in an observation. The XYGLOBALPROF cola-file is then asking you to define masks around each individual galaxy. It will come up with the GIDS display and put a cross where it thinks will be a galaxy. This cross is retrieved from the list.gal file. If no galaxy is at the position of the cross just select no area around it and it will disappear from further analysis. If there is a galaxy at a position not known to the PGC you should add that position to the list.gal file with an appropriate velocity. As a PGC number give 99999. If more than one galaxy is unknown give the second unknown galaxy a PGC number 99998, etc. This PGC number will be added to later set-names and files to distinguish the different galaxies.

If you have finished the first galaxy masks you can blank this galaxy before continuing to the second galaxy.

For every galaxy the global profile is determined and written to a file

flux<ugcnumber>s60hanp<pgcnumber>.dat

It contains the same columns as the previous flus-files, except that the third and fourth colomn are set to 0, since the fluxes are no longer derived from the clean-components. The masks belonging to the galaxies are stored seperately, their name containing the PGC number.

- XYTOTALHI: this tasks forms the PBC total HI maps at all three resolutions. It selects only the areas within the masks and then sums all channels, after which a PBC is applied. Also a S/N map is calculated.

- XYVELOCFIELD: the primary task of this step is to make a velocity field of the observations by taking the first moment of the data within the masked areas. However, all options of the program MOMENTS are used. The results are stored in the set u<ugcnumber>s<resolution>-2dim. In this set the zeroth moment is replaced by the total HI map from the previous step, the number of line channels used in moments is replaced by the one derived in the total HI step, and the S/N map and the radio continuum map are added to the set.

- XYCOMBINE: this task 'glues' the continuum subtracted radio continuum channels (i.e. empty noise channels) to the cleaned HI line channels, so that the datacubes contain their standard 63 or 127 channels.

- XYXVDIAGRAMS: in this task an XV-diagram, i.e. a slice along the major (and minor) axis is made. If there is more than one galaxy, this is done for every galaxy. The first thing that has to be done is to determine the position angle (PA) of the major axis. For this we use an iterative method. The astronomer has to select an area of the total HI map to which an ellipse is fitted. Then, with the new values for the ellipse as an initial estimate a second estimate is determined (the area to which the ellipse is fitted can be changed in the mean time). If the estimates do not change anymore, the astronomer is asked if he/she is content with the estimates. If so he/she is asked whether he/she wants to use any of the old (=opticallly derived PA value) or new (=derived from total HI map) PA values. If so the astronomer is asked which of the two he/she wants to use. If not so a velocity field is shown with degree markers and he/she is asked to put in a value by hand.

Now slices along the major and minor axes are made and displayed in GIDS. Also a GPLOT screen shows the values of the velocity field(s) plotted an a contour plot of the XV-diagram to show how well the velocity fields are determined.

A file u<ugcnumber>-galaxies.list is written containing the PGC number, other name, RA and Dec, magnitude, morpholgical type, velocity, major and minor axes and position angle of the galaxies found in the observation. An exclamation mark after the PA value indicates an HI determined value.

- XYCOMMENT: this task adds all kinds of comments to the GIPSY sets.

- XYEXTRACT: this major task extracts information for the WHISP database from the GIPSY sets. This information and the answers the astronomer gives to the questions is stored in a file xygal<ugcnumber>.tex or if more than one galaxy per observation is available xygal<ugcnumber>p<pgcnumber>.tex. This tex-file is stored in a directory defined by the environment variable $xyarchive. Also this task will add the galaxy or galaxies to the file xygalaxylist.txt, that is in the directory defined by the environment variable $globalarchive. This file is used in the WHISP-database site for administration. The script loops through all galaxies that could be distiguished in the XYGLOBALPROF task.

  First the cola-file wants you to determine the size of the HI emission in the total HI map. It comes up with the total HI map and asks you to determine the 'edges' of the HI emission. This edge is orange if you use the RONEKERS colorcoding. All three resolutions are asked.

  After this it will determine the total HI mass, the systemic velocity and the velocity widths W50 and W20 of the profile of the galaxy. Check the GIPSY program PROFGLOB to see how to interact manually with the global profile analysis.

  Next determine whether the maps show any defects. The possible answers are shown (yes/no).

  Then you must describe the total HI map, the shape of the velocity field, the shape of the XV-diagram along major and minor axis.

  Next the number of HI companions is asked.

  Finally, the continuum maps are roughly cleaned and restored, and it is asked whether there is any continuum emission associated with the galaxy under consideration and how many continuum sources with fluxes larger than 5 x r.m.s. are visible in the three 'cleaned' radio continuum maps.

  At the end an info-file with the name-form u<ugcnumber>p<pgcnumber>info.txt is written, that will be used in the WHISPLOT task later on.

- XYARCHIVE: this task stores the reduced and analyzed data on the same tape as the FITS-files with the raw uncleaned channel maps, but this time in a tar-format, using GTAR. The datacubes are stored in (GIPSY) GDS-format, to contain as much as possible information about the xy-reduction phase.

First the script checks whether the tape contains the same galaxy as will be written to it. If this is ok the following files will be written to tape:

In a first tar-file the 2-dimensional results are stored for 3 resolutions then the xv-diagrams along major and minor axis are stored and then the flux*.dat, u*-galaxies.list text and xygal*.tex files. Then, in a second tar-file, the cleaning masks at 60" resolution will be stored, then possible additional masks when more than 1 galaxy is found in the observation(s) and finally the cleaned cubes (+ continuum subtracted continuum channels) at three resolutions will be stored.

- WHISPLOT: this task will produce the WHISP Atlas pages, one for every galaxy, that can be viewed on the WHISP-database site. The WHISPLOT cola-script will need the file u<ugcnumber>p<pgcnumber>info.txt made in the XYEXTRACT task. You should first check if the information in this file is correct, especially the name of the galaxy.

  It will ask the following keywords:

  GPLOTFILE= the default is e.g. gplotu11861p.txt, but if there is more than one galaxy in the observation the name should be e.g. gplotu11861p38739.txt, i.e. the PGC number should be added to the name.

  PSFILE= the default is u11861plot.ps, but again if there is more than one galaxy in the observation the name should be u11861plot38693.ps.

  PGCNR= give the PGC number of the galaxy if there is no default.

  MINBOX= the user should provide the box that must be displayed in the images.

  The other keywords are mostly straightforeword. A matter that usually needs some attention is the labeling of the axes. Look in the WHISP Atlas for examples.

  Finally, I want to mention a situation that arises now and then, which can not be handeled without some extra work from the astronomer. There are cases where there are additional galaxies in the observations that are visible at the 30" and 60" resolution but not in the full resolution maps, because they are to far away from the pointing center (remember that the full resolution maps have only one quarter of the field of view of the 30" and 60" maps (see the WHISP reduction document for this matter)). WHISPLOT wants to have three resolution maps, so it will crash if it wants to select an area outside the full resolution maps. The astronomer now has to do two things. First he/she has to extend the axes of the full resolution 2-dim set. He/she can use the GIPSY task REPROJ for this. Give for the new box size -512 -512 511 511, instead of the old -256 -256 255 255. As a second step the astronomer must use the WHISP cola-file GETIMAGE to retrieve a larger area from the PSS. The default box for GETIMAGE is 35' x 35', but will now have to be extended to 60' or 80'. When these steps have been taken, WHISPLOT can proceed as usual.

- XYDELETE: This final task should not be completed before the next phase is finished. This task will delete all files from disk and has two options:

  1. You can start again to redo the analysis, i.e. all files and sets are deleted except the raw uncleaned datacubes and the antenna-patterns belonging to it.
  2. You can delete everything and start the reduction of another galaxy. After you have left GIPSY you should also delete any .LOG files in the directory.

4.4 Adding the reduced observations to the WHISP-database site

The last phase of the xy-reduction consists of a few steps that are necessary to add the galaxy information to the WHISP-database site. To complete these steps open a second terminal window under the whisp account.

- First go to the directory ~/whisp/xy/plots, i.e. a subdirectory of the directory where you did your xy-reduction. Here a file with the name u<ugcnumber>p.ps or

  u<ugcnumber>p<pgcnumber>.ps should reside. You will use the script convert.plot to convert this postscript-file to a gif-file.

  Usage : convert.plot <ugc-number> [-p pgc-number]

  The PGC number is only necessary if there is more than one galaxy in the observation. After a while the graphical program XV will come up with a screen (don't forget to set the DISPLAY environment variable). Right-click your mouse button, select save from the menu, click on gif, then on save. Now a gif-file should be saved. You can exit XV now.

- For the next step you have to go to the directory ~/whisp/xy/uv. In here is a script obs.red.html that will create an html-file (name = u<ugcnumber><pgcnumber> obsred.html) containing information about the observation and reduction of a galaxy. The input needed is from

  1. the xygal<ugcnumber>.tex or xygal<ugcnumber>p<pgcnumber>.tex file
  2. the uvgal<ugcnumber>.tex file. If this file is not present it can be made on the fly by the script, however you will have to load the SCAN-file of the galaxy from tape in the directory defined by the environment variable $uvgaldir. After you're ready delete this file.
  3. the terminal.

  Now start up the script with

  Usage : obs.red.html <ugc-number> [-p pgc-number]

Again, add the PGC number if more than one galaxy is in the observation(s). Now you will have to answer the questions asked.

After the script has finished some files will have been moved to the directory defined by the environment variable $tmp_galaxy_descr. These files will be used in the WHISP-database site.

- To add the information to the WHISP-database site go to the html-directory ~/public_html/Database/OverviewCatalog/ListByName/. Start the script copy_from_whisp:

  Usage : copy_from_whisp <ugc-number> [-p pgc-number]

  This will copy the necessary files to this directory. Then start create.galaxyref:

  Usage : create.galaxyref <ugc-number> [-p pgc-number]

  This will create a subdirectory with the name of the observation and add the galaxy to the WHSIP-database Atlas. If there is more than one galaxy in the observation, all galaxy Atlas-pages will be place in the directory of that observation but with their appropriate names and all galaxies will have their own links and pages on the site Atlas. In this case the name of the galaxy will be seperately asked by the script.

When these steps are completed, it is now save to proceed with the WHISP task XY-DELETE in GIPSY. This will clean-up the directory and prepare it for the next galaxy reduction.