The Westerbork Survey of HI in Spiral Galaxies

(WHISP)

Project Report

13 March 1993

1 Introduction

WHISP represents the efforts of astronomers in Groningen and Dwingeloo to map the distribution and velocity structure of neutral hydrogen in some 500 to 1000 galaxies with the WSRT in the next decade. These efforts will result in a catalog of HI distributions and velocity fields that will serve as a basis for research in many areas, for example: dark halo's, the effects of environment on the structure and growth of HI disks and galaxy distances.

In this report we describe in some detail our views concerning the execution of this survey, data reduction, man power needs, hard- and software requirements, etc.

2 Scientific Motivation

Maps of the neutral hydrogen distribution and velocity fields are now available for about one hundred spiral galaxies (see below). In our first observing proposal (WH 1103, dated 12-3-92) and in our ASTRON proposal for program support (N93 B01, not yet funded), we argue in detail why a major effort should be undertaken to increase this number drastically, i.e. by about one order of magnitude. The main reasons are:

- The need to cover a range in morphological type, luminosity and environment, e.g. for the study of dark halos and the study of the distribution of stars and gas.
- The need for HI velocity fields, to be used in observational cosmology, in particular to determine accurate distances with the Tully-Fisher relation.

- The need to find more good candidates for detailed studies of specific problems.
- The need to assemble a well documented data base with information on the distribution and kinematics of neutral hydrogen in galaxies.

3 Existing WSRT data

A summary of existing WSRT HI observations for the period 1981 - 1992 for individual galaxies is given in Table 1 (observing time > 4 hours); see Appendix. Well-known clusters, voids and objects with $\delta < 0°$ have been omitted. The list contains 267 objects, 222 of these have been observed for at least 10 hours (40 in the context of WHISP in 1992, project 1103). For several galaxies multiple observations exist. In these cases the most recent year of observation and the corresponding project number are given. Note that Table 1 does not contain all galaxies in the local group and in the M81 group because of the velocity cutoff of 200 km/s.

For the galaxies in Table 1 with UGC numbers we have searched the HI catalog of Richter and Huchtmeier (version June 1992) for bibliographical references. The statistics resulting from this are given in Table 2.

Table 2. Galaxies observed with the WSRT (> 10 hrs)

Years	1981-1989	1990-1992	1981-1992
Total number	102	120	222
Number of UGC galaxies	99	110	209
Number of UGC galaxies with one or more WSRT publications in Richter and Huchtmeier (June 1992)	44	15	59

Some comments on this table are in order. The bibliography included with the Richter and Huchtmeier catalog *does* contain dutch PhD theses. Thus, the relatively small number of observations carried out in the years 1981 to 1989 that has lead to one or more publications is a matter of concern. Some discussion of the observational results for the galaxies not referenced in Richter and Huchtmeier may have appeared in the literature, but no extensive presentation has appeared. On the other hand, the number of observations for the period 1990 to 1992 given in Table 2 includes 40 galaxies observed for WHISP in 1992 (project 1103), $\sim$ 25 galaxies observed by M.Verheijen (PhD student in Groningen, project 1046), and $\sim$ 10 galaxies observed by A. Broeils (PhD thesis september 1992).

4 Existing VLA data

A summary, similar to that for the WSRT given in section 3, has been made for galaxies observed with the VLA (prior to March 1992). Statistics for galaxies observed with the VLA for at least 2 hours are included in Tables 3a and 3b.

5 Selection of galaxies

The primary catalog from which galaxies will be selected for observation with the WSRT is the Uppsala General Catalogue of Galaxies (UGC, Nilson 1973). This catalog lists the galaxies north of $\delta = -2°30'$ with blue diameters ≥ 1 arcmin; it also includes smaller galaxies to the limiting magnitude 14.5. Data in the RC3 (de Vaucouleurs et al. 1991) have been added to the UGC and the HI magnitudes and profile widths listed in the RC3 have been used to calculate mean fluxdensities. (This flux-density scale agrees well with that of Broeils 1992, thesis.)

Tables 3a and 3b give the flux-density distribution for a sample of 3148 galaxies with $\delta \geq +20^\circ$ and blue diameters ≥ 1.5 arcmin. From Tables 3a and 3b we find that 195 galaxies have already been observed with the WSRT in the period 1981 - 1992, and about 100 with the VLA (with some overlap). It is to be expected that a fairly large number of the approximately 250 galaxies already observed with the WSRT or VLA will have to be re-observed, either because the existing data are not good enough or because the instrumental specifications are unsuitable for our purposes. Some of the data may also be difficult to recover. The number of galaxies that will have to be re-done may well be of the order of 100. Note that in Tables 3a and 3b all morphological types have been included.

Experience shows that for galaxies with flux densities in excess of 100 mJy HI maps of good quality can be obtained in a single 12 hour observation (For examples see thesis A. Broeils 1992, e.g. NGC 2998, $\simeq 60$ mJy, 2×12 hrs). A typical value for the profile width is 250 km/s. Thus a flux-density level of 100 mJy corresponds roughly to a total HI flux of 25 Jy·km/s. Depending on ones taste, a single 12 hour observation would also be adequate for somewhat fainter galaxies.

At least ~ 300 galaxies with $\delta \geq +20^\circ$ and blue diameters ≥ 1.5 arcmin, not yet observed with WSRT or VLA, have flux densities in excess of 100 mJy; see Tables 3a and 3b. In future, with increased sensitivity, it should be possible to go down to ~ 30 mJy in one 12 hr observation.

Table 3a. Statistics for UGC galaxies with $\delta > 20^\circ$ and blue diameter ≥ 2.0 arcmin

Fluxdensity RC3 (mJy)	Number	Observed with WSRT (> 10 hrs)	Observed with VLA (> 2 hrs)	Not observed with WSRT or VLA
not–listed	429	8	6	416
0–20	109	9	2	98
20–50	222	27	6	192
50–100	206	37	6	165
100–200	159	33	11	120
200–500	116	33	14	77
500–1000	43	16	12	21
1000–>	41	11	22	14
Total–	1325	174	79	1103

Table 3b. Statistics for UGC galaxies with $\delta > 20^\circ$ and $1.5 \leq$ blue diameter < 2.0 arcmin

Fluxdensity RC3 (mJy)	Number	Observed with WSRT (> 10 hrs)	Observed with VLA (> 2 hrs)	Not observed with WSRT or VLA
not–listed	1159	2	2	1155
0–20	261	-	1	260
20–50	263	7	3	253
50–100	90	6	1	83
100–200	36	3	1	32
200–500	14	3	1	11
500–1000	-	-	-	-
1000–>	-	-	-	-
Total–	1823	21	9	1794

6 Observations

In the first few years WHISP will consist mainly of single 12 hr observations of objects selected from the sample of $\sim$ 300 galaxies with $\delta \geq +20°$, blue diameters ≥ 1.5 arcmin, and fluxdensity > 100 mJy, not yet observed with the WSRT or VLA (see previous section), and of repeat observations where necessary. Multiple 12 hour observations may be needed for large objects and for objects of special interest. Some short observations will also be needed to determine rough fluxdensities and velocities for galaxies without HI information. Prior to 1996 the limiting sensitivity in one 12 hour observation will be about 1×10^{20} HI atoms/cm^2 ($1\ \sigma$). For more extended structures we can go down to 2×10^{19} HI atoms/cm^2 ($1\ \sigma$). From 1996 onwards the sensitivity will be at least 2 times better.

Table 4. Distribution of Right Ascension for 275 galaxies with $\delta > 20°$, blue diameter ≥ 1.5 arcmin, and fluxdensity > 100 mJy, not yet observed with the WSRT (> 10 h) or VLA (> 2 h)

Right Ascension	Number
$0^{\rm h}$–$3^{\rm h}$	31
$3^{\rm h}$–$6^{\rm h}$	15
$6^{\rm h}$–$9^{\rm h}$	36
$9^{\rm h}$–$12^{\rm h}$	78
$12^{\rm h}$–$15^{\rm h}$	73
$15^{\rm h}$–$18^{\rm h}$	22
$18^{\rm h}$–$21^{\rm h}$	5
$21^{\rm h}$–$24^{\rm h}$	15
Total–	275

Our aim is to present a 'block schedule' for the WSRT observations, such that no further scheduling will be necessary. For the first 100 galaxies or so this can be done fairly easily, but because the distribution in right ascension of the candidate galaxies is quite non-uniform, see Table 5, it will not be possible to do this at later times. We will propose a special HI project of clusters near 24 hours RA that can be used to fill the hour angle gaps in the block schedule for WHISP galaxies.

7 Data management

Figure 1 gives a schematic representation of the main phases that are involved in HI observations with the WSRT for WHISP. A key person in the project will be the data-manager. This person will be responsible for:

- Monitoring of the progress of the survey.
- The preparation of observing requests to be sent to ASTRON.
- Quality assessment and editing of the calibrated UV data coming from ASTRON.
- Map making (Fourier transforms).
- Inspection of the resulting data cubes.
- The overall management of the data base that will result from the survey.

We expect that a data manager will be appointed at the Kapteyn Institute in the summer of this year.

The members of the project team will share the responsibility for the execution of the survey, up to and including the production of the data base with ‘end products’: HI distributions, channel maps, velocity fields, etc. Presently we are discussing the division of labour such that each member has a well defined task. Together with promovendi and postdocs the members of the project team will also take a large share in the astronomical analysis.

Because we wish to proceed with the survey in 1993, and it may take some time before an experienced data manager is available, we have organized the various organizational tasks as follows:

1) Initiatives for new observations will be collected by *van Albada*, who will prepare the final lists of objects to be observed.

2) Block scheduling will be done by *van der Hulst*.

3) *Braun* will act as informant on WHISP in Dwingeloo and will keep an eye on progress made with observations and data handling in Dwingeloo.

4) *Begeman*, *van Gorkom*, *van der Hulst* and *Sancisi* will supervise the data manager in the areas of quality assessment and editing of UV data, map making and the production of data cubes.

5) As regards data analysis and astronomical processing, *Begeman* will act as consultant on technical matters, including the operation of software packages.

6) *Van Albada* will supervise the overall management of the data base.

8 Data reduction

In the organization chart given in Figure 1, calibration and first editing of UV data, e.g. for telescope failures, is foreseen to take place in Dwingeloo. These are minor tasks however. In principle therefore the calibrated UV data, or even the raw data, could be sent directly to the data manager in Groningen.

Traditionally, quality assessment, UV editing, map making, etc. are being done by the astronomer, often a promovendus, responsible for the project. This phase requires considerable know-how in the area of spectral-line radio-synthesis, and a large amount of routine work. For promovendi this can therefore be quite time consuming. In WHISP we will streamline this phase of the data reduction by having a single, experienced person, i.e. the data manager, perform these tasks. Astronomers will start their work with data cubes ready for astronomical processing.

We estimate that the production of a data cube for a single 12 hr observation can be done in one day by an experienced person (more or less full time work). For the data manager it should therefore be possible to handle at least 100 observations per year.

9 Astronomical processing

PhD projects based on WHISP would typically be based on about 100 galaxies. This is almost an order of magnitude more than is customary now, and one may wonder whether a PhD student can handle so many data. On the other hand, an experienced HI observer can make final maps and velocity fields for about 5 systems in one week. Clearly, PhD students need a lot of time to learn the techniques of 21-cm line synthesis. We believe this situation will improve greatly when PhD students no longer have to do the quality assessment, UV editing and Fourier transforms themselves, and understand all the intricacies involved. This will allow them to analyse and discuss a much larger sample of galaxies and also leave more time for interpretative work. A consequence of this approach will be that some of the new PhD's know far less about radio astronomy than presently.

10 Archiving

An important goal of WHISP is the production of a well documented data base containing HI distributions, channel maps, velocity fields, etc. This data base should also incorporate existing WSRT data. If necessary we will (re)do the data reduction for unpublished material. For each galaxy one needs about 100 to 200 Mbytes, thus several hundred Gbytes are needed for the results of the entire survey.

A cheap solution would be to store the results on DAT tapes. This permits a reasonably flexible access to the data base and it would fulfill the needs of the astronomical community in Groningen. On the other hand, this solution would not be very practical for astronomers outside Groningen.

It appears that an elegant solution for this problem is within reach, because the Computing Center of Groningen University will most likely buy a tape robot with a data storage capacity of 1000 to 4000 Gbytes. The Computing Center is planning to make this a flexible facility, offering not only archive space, but also a disk backup facility and even virtual disk space. Precise costs are not yet known but are expected to be of order *f* 1000 per Gbyte for a period of say 5 years. Storage of the HI database on this system would allow direct and flexible access to the results of WHISP for the entire astronomical community.

11 Requirements

Below we identifye a number of requirements in the area of hardware and software, and we propose a division of responsibilities between ASTRON and the Kapteyn Institute.

11.1 Hardware

For the data reduction we need fairly powerfull workstations (HP 730 or equivalent), each with $\sim$ 5 Gbyte diskspace. One twelve hour observation with 128 line channels produces a UV data base of about 100 Mbyte and a spectral line cube of typically 150 Mbyte (full resolution). One work station should be reserved for the data manager. Four PhD students can share 2 workstations and 2 X-terminals. We expect to have these facilities available at the Kapteyn Institute as part of the normal infra-structure. This expectation is based on the replacement of the Alliant at the Kapteyn Institute with a number of workstations by the Faculty W&N of the RuG. But no firm promise has been made yet!

11.2 Software

UV-data:
It will be important to have fast and flexible access to the UV data. Ideally, inspection of the progress of the observations should be possible through remote access to the WSRT HP 730. Unfortunately, the NEWSTAR system is still incomplete and does not function properly, and the display software is too primitive. At present it is not yet possible to inspect and edit the UV data in a flexible way, comparable to AIPS. For efficient reduction of the data it is essential that NEWSTAR includes a task with at least the same functionality

as the interactive AIPS-task TVFLAG. A smart, automatic flagging routine is very desirable as well. We request that these modifications and additions to NEWSTAR be carried out by ASTRON.

Maps:
Map processing and further analysis will be done with GIPSY for a period of 3 to 5 years at least. Eventually we will switch to AIPS++, once it has acquired the functionality we need for the project. But WHISP does not depend on AIPS++ and can be carried out well with the tools offered by GIPSY.

11.3 Financial

WHISP can become a success only with extensive OIO and postdoc support. We have submitted a proposal for 'programma financiering' (N93 B01).

Consultations with ASTRON are presently taking place concerning financial support for data managing.

We expect to ask support from ASTRON in the area of data archiving, in order to finance a share in the tape robot at the Computing Center of the RuG; see section 10.

We will ask support for workstations in the unlikely case that the faculty W&N does not support the replacement of the Alliant.

12 Project team

T.S. van Albada (p.i.)
K. Begeman
R. Braun
J.H. van Gorkom
J.M. van der Hulst
R. Sancisi
H. van Woerden