DETERMINATION OF RADIO SPECTRA FROM CATALOGUES AND IDENTIFICATION OF GIGAHERTZ PEAKED SOURCES USING THE VIRTUAL OBSERVATORY

B. Vollmer, S. Derrière, T. Boch, B. Gassmann, P. Dubois, F. Genova, F. Ochsenbein[1], T. Krichbaum[2], E. Davoust[3], and W. van Driel[4]

[1]CDS, Observatoire astronomique de Strasbourg, UMR 7550, 11 rue de l'université, F-67000 Strasbourg
[2]Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn
[3]UMR 5572, Observatoire Midi-Pyrénées, 14 avenue E. Belin, F-31400 Toulouse
[4]Observatoire de Paris, Section de Meudon, GEPI, CNRS UMR 8111 and Université Paris 7, 5 place Jules Janssen, F-92195 Meudon

ABSTRACT

We have used the 20 largest radio continuum catalogues contained in VizieR (Ochsenbein et al. 2000) to determine radio continuum spectra between wavelengths of 2cm and 1m. For 67,000 out of the 3.5 million catalogueed sources we could extract radio spectra with measurements at at least 3 independent frequencies. These have been validated by comparison with existing spectral indices from the literature. Our radio spectra data base was searched for Gigahertz peaked source candidates, which we then observed quasi-simultaneously with the Effelsberg 100-m radio telescope at 6 cm (4.85 GHz), 2.8 cm (10.45 GHz), and 9 mm (32 GHz). This is an efficient procedure to discover new Gigahertz peaked sources, which are believed to be AGNs at the beginning of their radio evolution. We are generalizing the method by using VO capabilities to (i) identify pertinent radio catalogues in the VO registry using Unified Content Descriptors (UCDs), (ii) extract relevant data, and (iii) normalize these for the determination of radio spectra. Software allowing semi-automated information retrieval is being developed within the framework of the European VO-TECH project. The potential usage of all available radio catalogues using these methods will strongly increase the number of independent radio source cross-correlations and radio spectra. The results are progressively being made available to the community through the CDS services.

Key words: Radio Continuum: general; Technique: spectroscopic; Virtual Observatory.

1. INTRODUCTION

Radio source cross-identification in the centimeter to meter wavelength domain are particularly difficult to obtain, because the underlying radio surveys can have huge differences in sensitivity and/or spatial resolution. Since the resolution depends on the observed frequency and the telescope diameter, low frequency surveys made with a small single-dish telescope can have resolutions up to tens of arcmins, while high-frequency observations with a large single dish telescope or an interferometer can have resolutions of a few arcsecs to tens of arcsecs. On the other hand, the cross-identification of radio sources at different frequencies is made easier by using the fact that, in the vast majority of sources, the spectral energy distribution has a power-law distribution. The radiation mechanism is either synchrotron emission from relativistic electrons gyrating in a magnetic field, or emission of hot thermal electrons. Synchrotron emission produces a power law spectrum with a possible cut-off or reversal of the spectral index at low frequencies due to self-absorption or comptonization. The spectrum of thermal electrons is flat, at least in the optically thin domain. Over the frequency range in which the majority of radio surveys were made, the spectra are thus well defined by a power law, i.e. as a straight line in the $\log$(flux density)–$\log$(frequency) plot commonly used in radio continuum astronomy.

2. THE PROJECT

The aim of our project is to extract radio cross-identifications/radio spectra from a heterogeneous set of radio catalogues. The input catalogues for the cross-identification can contain different degrees of information. The required minimum set parameters are a position (RA, DEC) and a flux at a frequency between 100 MHz and 20 GHz. Some catalogues provide additional information on the errors, source sizes, source names, and various flags. Moreover, the catalogue columns can have different units. Before the actual cross-identification of the radio sources, the information from the input catalogues has to be homogenized. This can be quite time consuming if it has to be done by hand for a large number of catalogues (more than 100). Therefore, we have developed Virtual Observatory (VO) tools for this task. The actual cross-identification is then performed by the existing tool SPECFIND (Vollmer et al. 2005a,b).

3. CROSS-IDENTIFICATION OF RADIO SOURCES

SPECFIND is a hierarchical code. It classifies a source j as parent, sibling, or child with respect to a given source i using a procedure which has different stages where stage 2 and 3 are refinements of stage 1.
stage 1: taking into account proximity criteria:
stage 2: taking into account flux densities at the same frequency:
stage 3: depending on flux densities at different frequencies, based on the expected radio spectral index:
At the end of this procedure source i and its siblings are considered the same source. If source j is identified as a parent, source i might be a resolved sub-source of source j. If source i has children, it might be extended, and the children represent its resolved sub-sources. SPECFIND also adds the flux densities of the children of source i at the same frequency. If the sum equals the flux density of source i within the errors, then source i is considered as extended. At the end of the cross-identification the self-consistency of the hierarchy and the uniqueness is tested. A radio source cannot be assigned to more than one physical object.

4. SPECFIND

SPECFIND is an efficient tool to identify radio spectra using radio catalogues of different formats. It can handle radio surveys of very different resolutions and sensitivities. It has been applied to 22 survey catalogues at 11 different frequencies. These catalogues were downloaded from the VizieR database. They contain a total of 3.5 million sources, leading to more than 700 000 independent radio cross-identifications and $\sim 67\,000$ independent radio spectra (Fig. 1) with more than two independent frequencies. The code was tested and its results validated by a comparison between the spectral indices found by SPECFIND and those determined by other authors. The determined spectral indices have an error of about ± 0.3. The radio spectra are available in VizieR at CDS.

5. SCIENTIFIC FOLLOW-UP PROJECT: GIGAHERTZ PEAKED SOURCES (GPS)

For the scientific use of the SPECFIND database in VizieR we have searched for sources with peculiar spectra, which are flat or inverted. These sources represent Gigahertz Peaked Sources (GPS) candidates or future mm-VLBI targets. Gigahertz Peaked Sources are supposed to be young radio-loud AGNs, and are not well studied (for a review see O'Dea 1998). Existing GPS samples are small (< 200, see, e.g. Labiano et al. 2007) and an extension of this sample is needed to investigate whether these sources are the precursor of the local FRI and FRII radio galaxies. We extracted 220 objects with flat or inverted spectra from the SPECFIND catalogue and observed them quasi-simultaneously at 4.8 GHz, 10.4 GHz, and 32 GHz with the Effelsberg 100m telescope. Since the SPECFIND spectra are made from multi-epoch data, we can assess in this way the percentage of variable sources in our sample. These variable sources represent potential candidates for intraday variable (IDV) sources. The preliminary results are that more than 50% of the sources show a peak above 1 GHz, $\sim 50\%$ are variable, $\sim 20\%$ have flat spectra up to 9mm, and less than 10% of the sources have an uncertain classification.

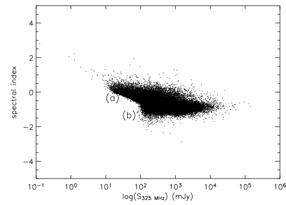

Figure 1. The spectral index as a function of the flux density at 325 MHz (from Vollmer et al. 2005)

6. GOING FURTHER WITH VO CAPABILITIES

The next step in our project is to include other available radio catalogues (more than 100). To do so, we first have to identify the catalogues in the VO world. Then, the information from these catalogues has to be homogenized. This homogenization, and the search for relevant radio catalogues, are time consuming. Moreover, we realized that the descriptions of the catalogues as they are given, e.g. in VizieR, are not sufficient for our needs: the survey resolution, frequency, size of the largest object are missing. For an efficient process of catalogue preparation we developed 3 VO tools within the framework of VO-TECH at the CDS:
- a tool to search for useful catalogues in the Virtual Observatory (Registry query tool),
- a tool to extract relevant information and to uniformize the catalogueue information (Data homogenization tool), and
- a tool to characterize the data (Characterization tool).

6.1. Registry query tool

This tool finds the VO resources based on Unified Content Descriptors (UCDs). The UCDs are a controlled

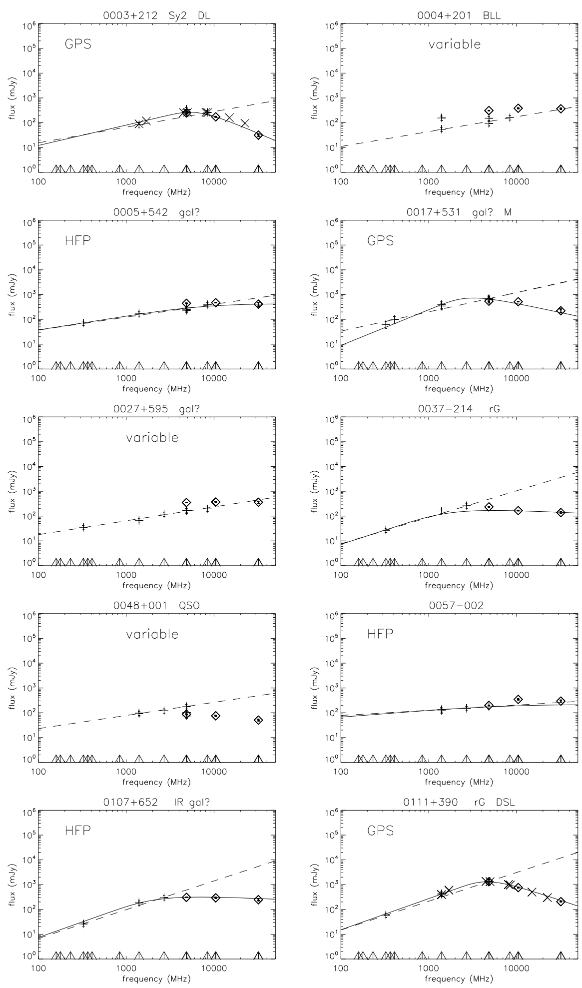

Figure 2. Examples for Spectra of Gigahertz Peaked Sources (GPS) from the SPECFIND catalogue observed at three frequencies with the Effelsberg 100m telescope. +: SPECFIND data; x: VLA quasi-simultaneous observations (Dallacasa et al. 2000); diamonds: Effelsberg observations; dashed line: SPECFIND fitted spectrum; solid line: fitted GPS spectrum.

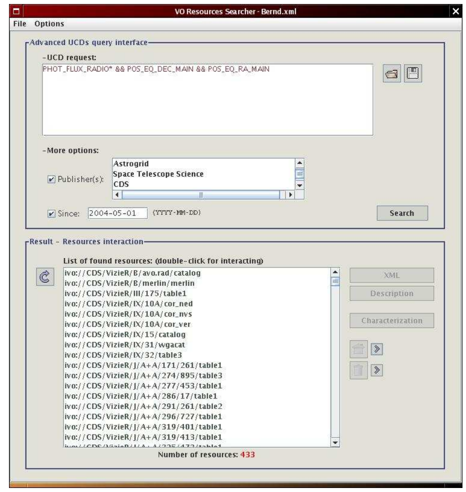

Figure 3. Interface of the registry query tool. Upper part: query based on UCDs, lower part: list of selected catalogues according to the query.

vocabulary defined by the VO to describe astronomical quantities. It is written in Java and uses XMLDV API to get data from an XML registry. The user specifies a required minimum set of UCDs. The tool searches in the VO registry for all catalogues whose descriptions contain these UCDs. For example, in our project the minimum set of parameters needed for the radio cross-identification are source coordinates and a radio flux. The corresponding UCDs are: right ascension: POS_EQ_RA, declination: POS_EQ_DEC, radio flux: PHOT_FLUX_RADIO*. Since we require all UCDs to be present in the catalogue, the request to the registry query tool is: PHOT_FLUX_RADIO* && POS_EQ_RA && POS_EQ_DEC (Fig. 3). The result of the query is a list of relevant radio catalogues (lower part of Fig. 3). The catalogues can then be sorted into useful and not useful catalogues by viewing the catalogue descriptions (VizieR Readme) in a web browser. A workspace permits to save and restore all actions performed on the catalogues. At the end, a final list of relevant catalogues is established.

6.2. Data homogenization tool

These relevant catalogues can be directly loaded into the data homogenization tool (upper part of Fig. 4). The aim of this tool is to create homogenized data from a heterogeneous set of catalogues. It is written in Java and works on XML tables. In a first step the user specifies a set of output columns. In our case we defined the output columns according to the needs of SPECFIND (lower part of Fig. 4). In a second step the tool generates an interface where a column of the entry radio catalogue is assigned to a user specified output column (Fig. 5). The user can always change the input column that he wants to assign to an output column. It is also possible to assign an arithmetic combination of different input columns or conditions on input columns to an output column. If, e.g.,

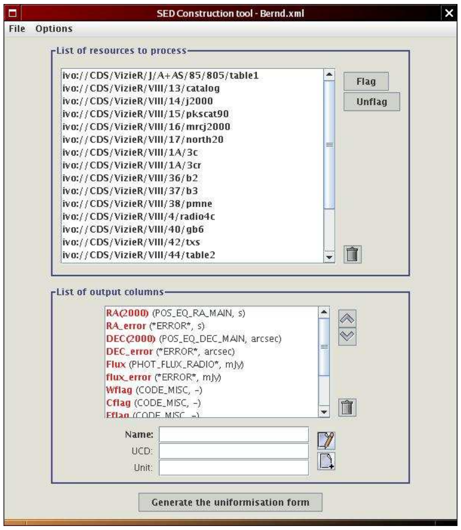

Figure 4. Data homogenization tool. Upper part: list of uploaded relevant catalogues. lower part: user specified set of output columns.

Figure 5. Homogenization interface. Each row corresponds to one catalogue. Each column is predefined by the user. An column of the input catalogue is assigned to a predefined output column (Fig. 4).

there is no explicit column for the flux density error in the input catalogue, it can be created assuming that it is 20% of the flux density. Units are converted automatically. At the end the tool generates an ASCII or VOTable output table for each input radio catalogue. These output tables can be directly used by SPECFIND.

6.3. Characterization tool

We realized that the description of the radio catalogues does not contain all necessary information for the radio cross-identification. We therefore decided to develop a third VO tool which permits to specify this missing information. More generally, it will permit to create a full description of a VO resource based on the VO characterization data model adding for instance information about the observations date, the resolution, etc.. This tool serves as input for the data homogenization tool.

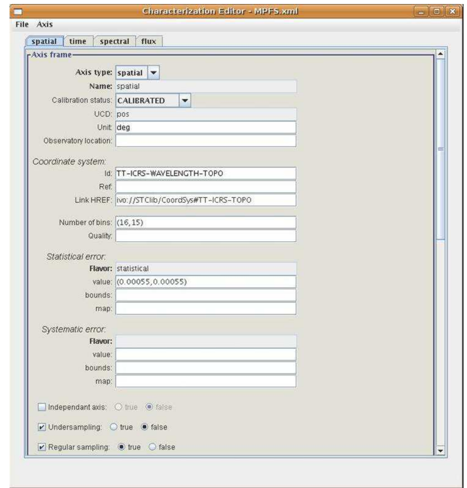

Figure 6. Characterization tool.

7. FURTHER STEPS

Within the framework of VO-TECH we have developed three VO tools at CDS: (i) a tool to search for useful radio catalogues in the Virtual Observatory, (ii) a tool to extract relevant information from these catalogues and to uniformize the catalogue information, and (iii) a tool to characterize the data. During the development we kept the tools as general as possible. In this way they can be used in other astronomical contexts. We also plan to include the extended radio catalogue description gathered by the characterization tool into VizieR. Prototypes of these tools are available at http://eurovotech.org/twiki/bin/view/VOTech.

REFERENCES

Dallacasa D., Stanghellini C., Centonza M., Fanti R. 2000, A&A, 363, 887

Labian, A., Barthel P.D., O'Dea C.P., et al. 2007, A&A, 463, 97

Ochsenbein F., Baue, P., Marcout J. 2000, A&AS, 143, 23

O'Dea C. 1998, PASP, 110, 493

Vollmer B., Davoust E., Dubois P., et al. 2005a, A&A, 431, 1177

Vollmer B., Davoust E., Dubois P., et al. 2005b, A&A, 436, 757