NEW SKILLS OF RADIO ASTRONOMY DATA CENTER (RADC) AT PRAO ASC LPI

V.A.Samodurov [1,2], M.A.Kitaeva [1], E.A.Isaev [1,2], D.V.Dumsky [1], V.D.Pugachev [1], S.V.Logvinenko [1], A.U.Zaitsev [1]

[1] Pushchino Radio Astronomy Observatory ASC LPI, Pushchino, Russia
[2] National research university Higher school of economics, Moscow, Russia
sam@prao.ru

ABSTRACT. Now a day the task of comparative analysis of sample sources from different astronomical catalogs is becoming very topical for astronomers. It is useful both for the same spectral range data and for their cross-analysis for different spectral ranges. It is also important, that the real data from astronomical instruments is placed on-line. For the above purposes we are developing a website for the PRAO ASC LPI (*www.prao.ru*) and, based upon it, the Radio Astronomy Data Center (RADC). The RADC is a system for storage and processing of radio astronomical data. Here you can find a collection of tools for preparing radio astronomical observations, which includes a base set of main astronomical catalogues useful to radio astronomers, as well as a database of observation results from the Observatory.

Key words: radio survey, catalogues, database

The Radio Astronomy Data Center (RADC) at last three year was created and developed at PRAO of LPI. RADS consists of:

a) a database of celestial sources on the basis of the most important astronomical catalogs;

b) a base of observation data of Pushchino radio astronomy observatory.

The Astronomical Catalog Database *http://astro.prao.ru/db/* contains a most important astronomical catalogs which are necessary for planning of observations. Since 2011 a database of astronomical catalogs actively equipped with graphical tools for data visualization and cross-analysis of catalogues between one with other. The base of observation data of observatory *http://observations.prao.ru/* continuously gets results from the majority of observational devices and radio telescopes PRAO. This database provides access to observation instruments and telescopes descriptions, techniques of making data samples per instruments, information about types of observations, observers and dates of observations and so on. The database contains information about more than 270 thousands of files of the observational data of PRAO ASC LPI radio telescopes since 1981. The total amount of information is about one terabyte.

At present the database contains pulsar data (more than 131 thousand profiles for more than three hundred pulsars for the last 7 years), spectral data for more than hundreds of space masers (for a number of them in the database are stored perennial series data since 1981) and the data of radio all-sky surveys on 102.5 and 111 MHz.

From the middle of 2013 the database gets the most volume observation data from the third 128-beam diagram of BSA (it is developed in 2011). The total size of the data is more than 680 Gb at the beginning of August, 2013.

The total flow of data, daily saved and described in the database of Observatory, has reached more than 2.2 Gb per day. All databases of observatory work under control of Postgresql and are written into special storage of data RAID-array with a general capacity of 24 Terabytes.

Database of astronomical catalogs

In order to support and preparation of the observations on the website PRAO the site "Working environment of the radio astronomer" *http://astro.prao.ru/* is created. The database of astronomical catalogs at PRAO ASC LPI site (*http://astro.prao.ru/db/*) is one of the part of this work space. This database contains the most important catalogs with common record count nearly 1,500,000 of sources.

Programs of filling of data base catalogs allow to expand the list of necessary catalogs easily at any time. Programs of data selection on any parameters of any catalog work for all catalogs. It allows to conduct statistics of data on the chosen parameters and to prepare selections of data for observations.

Thus for each catalog two versions of a request form are given: simple and complex. In a simple form the most important parameters of sources (for example, coordinates and fluxes) with already entered borders on them are automatically entered. Therefore it is enough to press the button «Select data» and all catalogs completely, with these basic parameters for each source, will be output. The complex form allows to select any source parameters (shown in the description of the catalog), and the necessary restrictions on them.

The database contains now several major astronomical catalogs, which are necessary for planning of observations of radio astronomers. The database is consisted from several summary tables of astronomical catalogs and the

some descriptions of the actual catalog tables. The adding of catalogs and dates in catalogs tables and its maintenance is doing by special programs written in Perl and PHP. Since 2011 a database of astronomical catalogs actively equipped with graphical tools for data visualization and cross-analysis of catalogues between one with other. This will allow us is producing of statistical and cross-sectional analysis for various astronomical catalogs. So, catalogues of radio sources at different frequencies can be widely used in further theoretical and experimental research of the properties of extragalactic radio sources and objects of our Galaxy.

The statistical cross-analysis of these various catalogs is suitable for research of properties of separate objects, and for the statistical analysis of properties of various classes of objects, and for research of properties of catalog data (completeness, reliability, calibration of catalogs, etc.).

For the tasks we develop a facility to display the data by graphic tools (generally through the PHP language) both one and a few catalogues within a chosen area in the sky; to display of data and statistical analysis of main parameters of each catalogue as a whole; to show statistics of cross-identifications of the catalogs chosen by the user.

At present already works the system of automatic mapping for 50 thousand the brightest sources for each catalog. Fluxes of all sources are divided on a logarithmic scale of fluxes on 6 equal "stellar magnitude". Such general mapping of catalogs gives a simple, but very effective instrument of visual quality control of completeness of catalog data and possibility of rough comparison of populations of sources in different catalogs. On each of these catalogs it is possible to look at maps on our site of catalogs, under "map" links. It is interesting that on these maps difficulties in processing of each of catalogs are well shown also. For example, for the catalog on 1400 MHz [3] difficulties of processing of brighter (right) part of a area of the Milky Way become insuperable. As result we can see a gap in the catalog on this place.

Thus it is well traced weaker (and therefore easier for processing) the left part of the Milky Way, being traced by bright sources. As we can see, even a fluent comparative analysis of common graphical data for each catalog can provide food for thought and an evaluation of the quality of the catalog data.

Cross-identification and joint data visualization of several catalogs.

For more complex tasks we develop tools of the graphic cross-analysis of catalog data. By means of these tools will be displayed some catalogs within a chosen platform in the sky; will be generated the statistics of cross-identifications of the catalogs chosen by the user, etc.

In fig. 1 – an example of imposing of three catalogs within the chosen area: 4C [2] (178 MHz, green crosses) and catalogs on 1400 MHz [1] (pink circles) and 4850 MHz [3] (blue crosses). Avoiding of catalog data from high-frequency catalogs round a bright source from 4C of the catalog (3C144 – in the center) is noticeable.

It is obvious that such evident check of coincidence of sources of different catalogs gives a reliable and simple way of the analysis of catalog data. The most promising part of work which is done while only out of a database of catalogs: flux identification of sources in various catalogs about mutual completeness of catalogs, the analysis of various selections of sources, the spectral indexes determined by various couples of the catalogs etc. Already fulfilled algorithms now are transferred under a database. That will allow to take a considerable step forward in development of the comparative statistical analysis of various radio-astronomical catalogs (and further - catalogs of astronomical objects in other spectral ranges), and in the on-line mode. Similar technologies is used in the world still slightly.

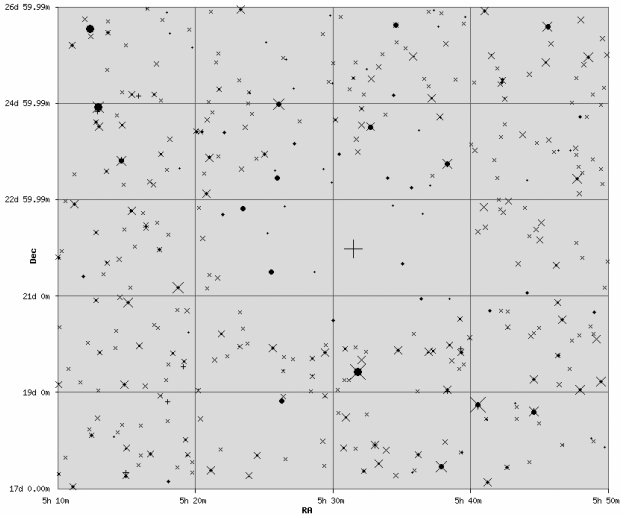

Figure 1: Example of the comparative analysis of the data field for 3 catalogs.

Database of observation results

After preparation of observations and forming observations schedule this observations are carried out at different PRAO ASC LPI radio telescopes. And we have quite differently data as results of this observation: pulsar, survey, observations of scintillate sources, spectral observation and other. The common database of radio sources observations at PRAO ASC LPI (meter and millimeter radio wavelength spectral diapason) was created for systematization and analysis of observation data. By this database of the site "Electronic database of observation results from radio telescopes of PRAO ASC LPI" was launched in 2006 year: *http://observations.prao.ru/*.

This database provides access to observation instruments and telescopes descriptions, techniques of making data samples per instruments, information about types of observations, observers, dates of observations and so on. For example, it is possible to view all data for the last few days or view changes of single chosen source. With the help of data analysis we can calibrate received signal from standard sources, monitoring states of antennas and noise level and levels of interference and so on. Below we propose to consider examples of the most common at the Observatory observations.

Pulsar observations

Currently, the database contains observational data for more than 300 pulsars, 131 thousand profiles of pulsars since the end 2006 on the end of March, 2013.Thus it is

possible to see both changes of profiles of a concrete pulsar day after day, and schedules of changes of brightness of a pulsar (measured in relation units signal/noise).It is noticeable that various pulsars possess a different behavior for all the time of observations: some practically don't change the brightness, in behavior of others seasonal (?) changes and even long-term trends are traced. All these changes are apparently partially connected with a condition of the effective area, partially reflect changes in the interstellar environment (long-period blinking, etc.).The part from them, probably, reflects a physical condition of the pulsar. These effects demand further studying and division into the specified components.

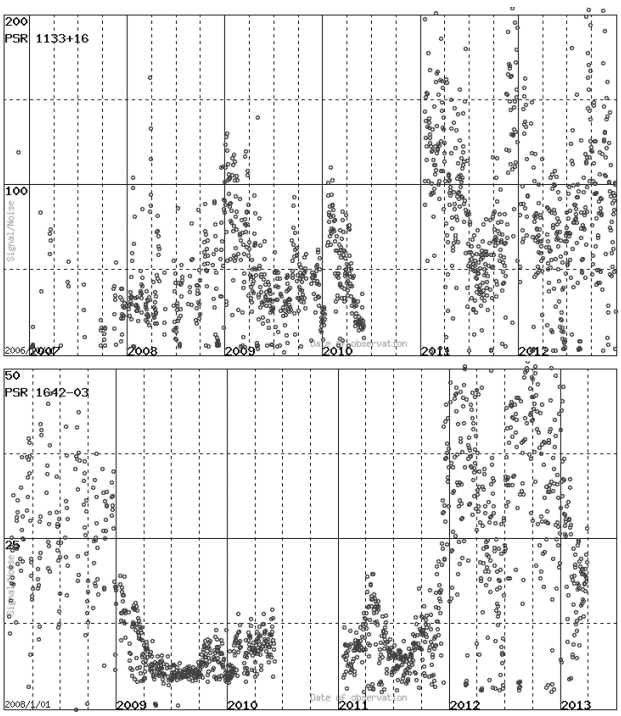

Figure 2 a,b: The signal/noise ratio for PSR 1133+16 (upper) and for PSR 1642-03 (bottom) in 2007-2013.

On the figure 2 a,b are graphs of the signal/noise ratio for PSR 1133+16 and for PSR 1642-03. For these pulsars systematization of changes is observed. Here are probable as the seasonal changes generated by change of characteristics of the antenna, as well as real changes in pulsars or in the environment of distribution of signals from them. There is evidence of the similarity of the changes some times, but identity isn't quite obvious. For PSR 1642-03 to appear there are considerable long-term changes of brightness between two levels.

The maser observations in water vapor line at 22 GHz

Similar long-term changes can be easily obtained from databases and for other types of sources. From 2009 observational database was supplemented by cosmic masers many years data (Fig. 5). The majority of the water maser sources (at 22 GHz) introduced its data for 1994-2010, but some sources - observations there are since 1981.

In 2010, the possibility of a database on data output and analysis of observations of the maser sources have been expanded (the new version of the database is located at *http://observations.prao.ru/cgi/new/index.pl* Now it is possible to output summary information of a set of parameters depending on time on each of sources. Below we can see an example of such analysis for one of sources – W Hya. Because of the low situation in the sky (α_{1950}= 13h 46m 12s δ_{1950} = -28° 7' 9") this source is very low and inconvenient for observation in middle latitudes. The height of the source of the close horizon, weather, and, of course, the condition of receiving equipment affect on measurements of the flux density of this source. All of this is clearly seen in figure 3 a,b where showed parameters of several observation sessions of this source in the April 2010.

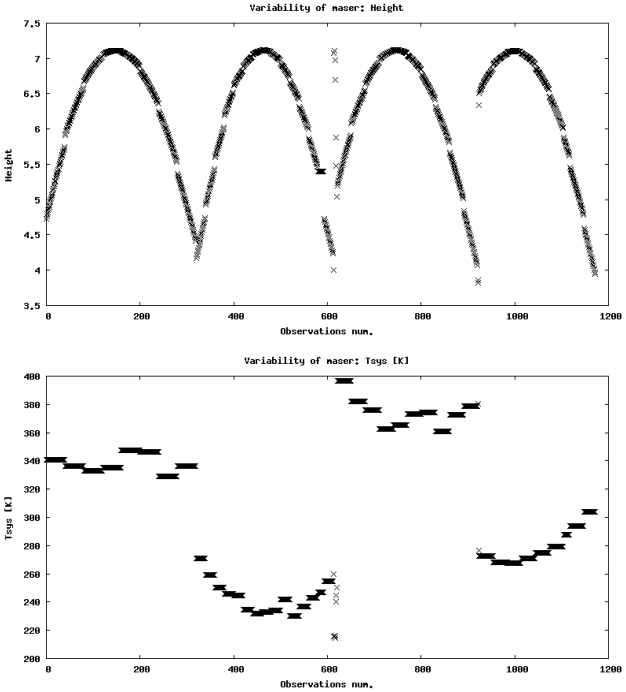

Figure 3 a,b: The statistic of 4 observational sessions of source W Hya:

a) (upper) Dependence of the height of the source above the horizon from file number in April, 2010.

b) (bottom) Dependence of T_{sys} from file number in April, 2010. It is well visible that in T_{sys} is lower, when source above.

The multi-beam survey observations of the northern hemisphere of the sky at 110 MHz with LPI's BSA.

PRAO ASC LPI has several world class radio telescopes. One of them is the Big Scanning Antenna (BSA). The BSA Radio Telescope is a phased array comprising 16384 dipoles and covering an area of 384x187 m (geometrical area is more than 70,000 m^2; effective area of about 30,000 m^2). The BSA's operating frequency is 109-112 MHz (since 1996 and 101-104 MHz pre-1996) and it is most sensitive telescope of this range in the world, as well as one of the most sensitive in the world in the meter wave band. The most important feature of the BSA is that it works in full-power receiving mode. It allows us to detect, besides discrete radio sources, the background radiation of our Galaxy and extended radio sources (with a typical size of up to 2-3 degrees). Another important feature of the BSA is that was originally designed with possibility to generate a multi-beam pattern. Until 2007 the

polar pattern of the BSA consisted of 16 beams, then a second, independent 16-beam chart was created. Finally, in 2010-11 a third beam former system was developed.

The new beam former system forms a 128-beam pattern with a field-of-view of 500 square degrees and requires a multi-channel registration system. Data channels from this multi-beam pattern are equipped by receivers in some stages. The first 48 beams were equipped with multi-beam receivers conducting 24/7 data recording since July 7 2012 (fig 4 a,b,c). Since April 1 2013 data is being recorded from 96 beams in the $-8°<\delta<+43°$ declination sector.

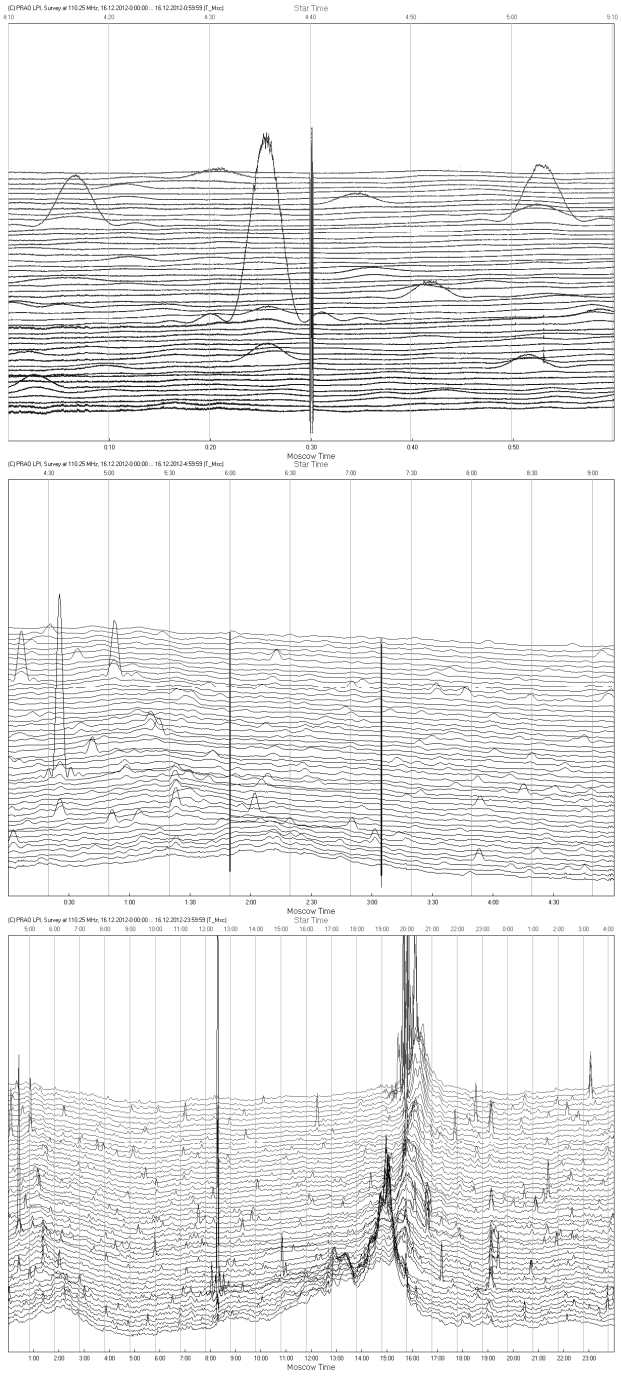

Figure 4 a,b,c: displayed observational dates for 1 hour observations (upper), 5 hour (middle) and day and night (bottom) observations.

Each of the 96 registered beams from BSA's third pattern captures data in 6 channels, 0.42 MHz wide, from 109 to 111 MHz with time resolution of 0.1 second. Each hour two 46 MB files are generated containing 48 channels from different beams. As a result 2.2 GB of data is generated every day. On fig. 4 a,b,c displayed examples 1 hour observations, 5 hour and day and night observations. We may look that dates have good and stable quality.

Additionally, test observations with time resolution of 0.02 s and 32 channel resolution (78 kHz per channel) have been conducted successfully. Daily volume of data generated by these observations was 55 GB. In 2014 all 128 channels will be operational. The 2nd radiation pattern is currently still active and is covering the $+49.1°<\delta<+55.2°$ declination sector.

As a result, scientific installations that are connected to all the diagrams of the BSA send enormous amounts of information. The volume of data from the 3rd multibeam diagram of the BSA from August 2013 has already reached 680 GB. This data presents huge resources for short and long term monitoring of various classes of radio sources (including transients), as well as observing space weather, Earth's ionosphere conditions, discovery of new radio sources etc. Currently a database is being developed for these type of observations. This database will feature an on-line graphical interface that will display current states of the BSA antenna, radio sky in the 109-111 MHz frequency band, currently monitored radio sources and more. Commissioning of functionality for full-time data output onto the PRAO ASC LPI website is planned for the end of 2013.

In view of the colossal volumes, the data requires new methods of processing. We are planning to integrate the use of supercomputers as well as distributed client network computing on the basis of BOINC technology.

Perspective applications of the database and new processing methods:

- Space weather – determining scintillation indexes of hundreds of radio sources on the scale of days and months
- Compilation of a catalogue of about 10,000 scintillation sources in our survey
- Monitoring of changes in flux of hundreds and thousands of sources
- Search for pulsars
- Search for radio transients
- And many others...

We also very hope on a possible co-operation and exchange of science dates with the Ukrainian radio astronomers working on the survey programs by radio decameter radio telescope UTR near Kharkov.

References

White R.L., Becker R.H.: 1992, *Astrophys. J. Suppl. Ser.*, **79**, 331.

Pilkington J.D.H., Scott, P.F.: 1965, *Mem. R. Astron. Soc.*, **69**, 183.

Gregory P.C., Scott W.K., Douglas K., Condon J.J.: 1996, *Astrophys. J. Suppl. Ser.*, **103**, 4272.