Multi-wavelength studies of the statistical properties of active galaxies using Big Data

A. M. Mickaelian, H. V. Abrahamyan, M. V. Gyulzadyan, G. A. Mikayelyan and G. M. Paronyan

Byurakan Astrophysical Observatory (BAO), Byurakan 0213, Armenia
email: `aregmick@yahoo.com`

Abstract. Statistical studies of active galaxies (both AGN and Starburst) using large multi-wavelength data are presented, including new studies of Markarian galaxies, large sample of IR galaxies, variable radio sources, and large homogeneous sample of X-ray selected AGN. Markarian survey (the First Byurakan Survey) was digitized and the DFBS database was created, as the biggest spectroscopic database by the number of objects involved ($\sim$20 million). This database provides both 2D images and 1D spectra. We have carried out a number of projects aimed at revealing and multi-wavelength studies of active galaxies among optical, X-ray, IR and radio sources. Thousands of X-ray sources were identified from ROSAT, including many AGN (52% among all identified sources). IRAS PSC/FSC sources were studied having accurate positions from WISE and a large extragalactic sample was created for further search for AGNs. The fraction of active galaxies among IR-selected galaxies was estimated as 24%. Variable radio sources at 1.4 GHz were revealed by cross-correlation of NVSS and FIRST catalogues using the method introduced by us for optical variability. Radio-X-ray sources were revealed from NVSS and ROSAT for detection of new active galaxies. Big Data in astronomy is described that provide new possibilities for statistical research of active galaxies and other objects.

Keywords. catalogs, surveys, virtual observatory tools, galaxies: active, quasars: general, galaxies: Seyfert, galaxies: starburst, infrared: galaxies, radio continuum: galaxies, X-rays: galaxies.

---

1. Introduction: Byurakan surveys for active galaxies

Byurakan surveys for active galaxies are well-known. The Byurakan Astrophysical Observatory (BAO) has always been one of the centres for surveys and studies of active galaxies. In 1960s B. E. Markarian started the first systematic survey for active galaxies (Markarian 1967), using ultra-violet-excess (UVX) method, First Byurakan Survey (FBS) or Markarian Survey. 1515 UVX galaxies (Markarian galaxies) were revealed (Mazzarella & Balzano 1986, Markarian *et al.* 1989). Lists of candidate active galaxies were compiled also by M. A. Arakelian (1975; 621 high surface brightness galaxies – Arakelian galaxies) and M. A. Kazarian (Kazarian *et al.* 2010) using similar methods. Thousands of active galaxies were revealed and studied. The Second Byurakan Survey (SBS; Markarian *et al.* 1983; Stepanian 2005) was deeper survey for UVX and emission line galaxies and quasi-stellar objects, QSOs ($\sim$3600 objects, including 1800 galaxies and 600 QSOs). The Second Part of the FBS revealed 1103 Blue Stellar Objects (BSOs), including QSOs and Seyferts (Mickaelian 2008). Optical identification and study of the Infra-Red Astronomical Satellite (IRAS) sources led to revelation of 1278 Byurakan-IR galaxies (BIG; Mickaelian & Sargsyan 2004), including many ultra-luminous infra-red galaxies (ULIRG), Active Galactic Nuclei (AGN) and Starbursts (SB). At last, the optical identification and study of the Röntgen Satellite (ROSAT) sources led to revelation of 4253 ROSAT AGN (Mickaelian *et al.* 2006). A number of related conferences were held in Byurakan: IAU S304 Multiwavelength AGN Surveys and Studies (Mickaelian &

32

Sanders 2014), the Symposium Astronomical Surveys and Big Data (Mickaelian *et al.* 2016), and others.

Markarian Survey was accomplished with a goal of revealing new active galaxies and this was achieved by eye selection. We have digitized this survey plates and created the Digitized First Byurakan Survey (DFBS; Mickaelian *et al.* 2007; Massaro *et al.* 2008; www.aras.am/Dfbs/dfbs.html). These low-dispersion spectra may be rather efficient for optical identifications and search for new bright AGN. The Armenian Virtual Observatory (ArVO; Mickaelian *et al.* 2009; www.aras.am/arvo.htm) was created based on the DFBS, Digitized Second Byurakan Survey (DSBS), and other digitization projects in BAO. The first science projects with DFBS/ArVO were the optical identifications of Spitzer Boötes sources in 2005 (Hovhannisyan *et al.* 2009). The development of ArVO project includes Armenian astronomical archives and present telescope data preservation, cross-correlations of direct images and low-dispersion spectra, creation of joint low-dispersion spectral database (DFBS / DSBS / HQS / Hamburg-ESO Survey (HES) / Case) and a number of other technical and scientific projects.

## 2. Current Projects on multi-wavelength search and studies of active galaxies

Current projects on search and study of active galaxies using multi-wavelength (MW) big data and statistical methods include:

- FBS QSOs and Seyferts (1986-2001)
- IRAS BIG sample; AGN, SB, ULIRGs (1994-2010)
- ROSAT Bright Source Catalogue (BSC) / Faint Source Catalogue (FSC) sources; search for new AGN (2002-2006)
- Statistical analysis of bright AGN (2001-present)
- Spectral and MW study Markarian galaxies (2010-present)
- IRAS Point Source Catalogue (PSC) / Faint Source Catalogue (FSC); large IR galaxies sample (2011-present)
- Spitzer ULIRGs; high IR/opt flux ratio galaxies (2003-2010)
- HRC (Hamburg ROSAT Catalogue) / BHRC (Byurakan-Hamburg ROSAT Catalogue) AGN sample (2010-present)
- NVSS (The NRAO VLA Sky Survey) – FIRST (Faint Images of the Radio Sky at Twenty cm) cross-correlation; study of radio variability (2013-present)
- ROSAT-NVSS cross-correlation; search for new AGN and statistics (2013-present)
- MW study of blazars (2014-present)
- Fine analysis of spectra and study of physical properties of AGN (2001-2007)
- DFBS new bright active galaxies (2002-pres.)
- Fine classification; accurate AGN types and subtypes (2006-present)

Here we give a brief overview of the most important ones.

2.1. *Markarian galaxies*

After the digitization of Markarian Survey and creation of the DFBS database, we carried out a number of other studies related to Markarian galaxies. An important work was the homogeneous classification of Markarian galaxies for activity types based on SDSS spectra. Spectra for 779 Mrk galaxies were available and these objects obtained their new classes and subclasses. We give in Fig. 1 examples of Markarian (Mrk) galaxies classified as Narrow Line Seyfert 1 (NLS1) based on its relatively narrow permitted lines, presence
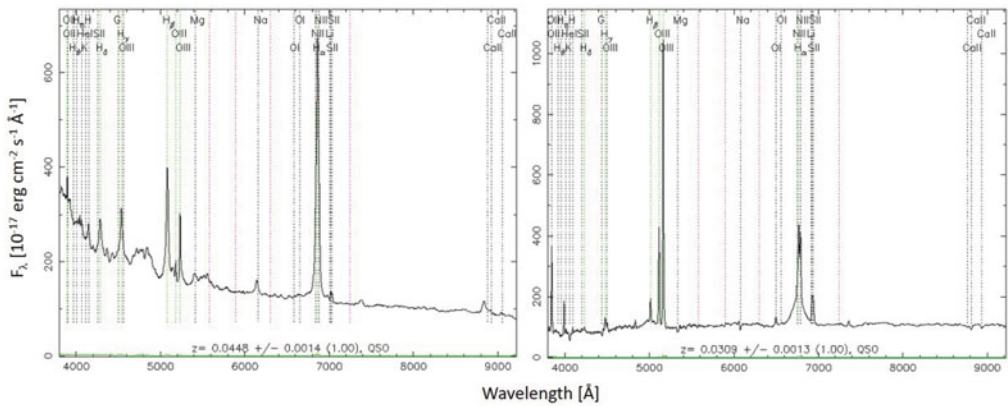
Figure 1. Examples of SDSS spectra of Mrk galaxies and their fine classification; Mrk 142 (Narrow Line Seyfert Galaxy, NLS1) and Mrk 1469 (S1.8 showing broad emission components both for Hα and Hβ).
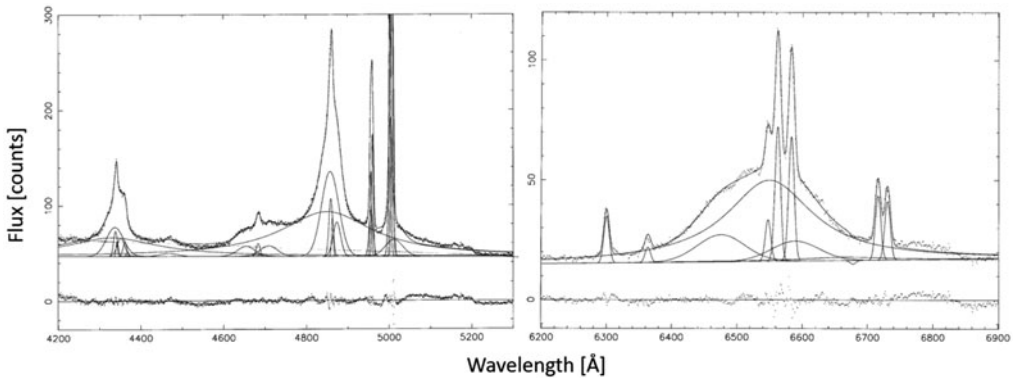
Figure 2. Fine analysis of the emission-line spectra of active galaxies with the dedicated software SPECTRAI. Examples of the results of the fitting in ranges around Hβ (left panel) and Hα (right panel). Left panel: Mrk 509, right panel: Mrk 926.

of intense FeII lines and soft X-ray and S1.8 based on its broad emission line components both for Hα and Hβ.

2.2. *Fine analysis of emission-line spectra of AGN*

Most of type 1 AGN have contaminated spectral lines so that the broad component is overlapped on the narrow ones. A programme of fine analysis of emission line profiles and detailed spectroscopic classification of bright S1 galaxies (including NLSl ones) was conducted in collaboration with Véron & Véron-Cetty (Observatoire de Haute-Provence, OHP, France). It was aimed at studying the differences between the classical Broad Line Seyfert 1 (BLS1) and NLS1 galaxies (Osterbrock & Pogge 1985) and correlations between FeII and Balmer lines, Fe II and [OIII], and better understanding of the physics of AGN central regions. However, it appears that there is not a strict separation between the BLSl and NLSl galaxies (the linewidth limit at 2000 km/s is arbitrary), and the intermediary objects may fill the gap. High-dispersion spectroscopy of 90 S1s having intermediate FWHMs was carried out at OHP 1.93m telescope with 0.9 Å/pix dispersion, the spectral resolution being 3.2Å. S1s from Véron-Cetty & Véron (2010) with V<16 and z<0.1 were selected, including well-known objects (Fig. 2).

2.3. *ROSAT sources and HRC/BHRC extragalactic sample*

Two large works on optical identifications of ROSAT sources were carried out, using ROSAT BSC and FSC (Véron-Cetty *et al.* 2004; Mickaelian *et al.* 2006). A collaboration between Armenian (BAO), French (OHP), Hamburg quasar survey (HQS) and ROSAT teams with participation of scientists from Mexico and China was conducted to study the Hamburg-ROSAT identifications and to select NLS1 galaxies for further detailed investigation. Moreover, sources from ROSAT FSC were identified and were observed as well. Beside NLS1 galaxies, X-ray sources identified with AGN contain other interesting subsamples: QSOs, BLS1 galaxies, BL Lacertae objects (BLL), etc. An X-ray and optically flux-limited sample of ∼1000 Seyfert 1 galaxies was created.

The joint catalogue of AGN selected from optical identifications of X-ray sources was created as a combination of two samples: Hamburg-ROSAT Catalogue (HRC; Zickgraf *et al.* 2003) and Byurakan-Hamburg-ROSAT Catalogue (BHRC; Mickaelian *et al.* 2006). Using the recent optical and MW catalogues we have revised both samples excluding false AGN and adding new genuine ones. Thus, a new large homogeneous complete sample of 4253 X-ray selected AGN was created. 3352 of them are listed in the Catalogue of QSOs and Active Galaxies (Véron-Cetty & Véron 2010) and 387 also are in Roma BZCAT (Massaro *et al.* 2015). 901 candidate AGN are subject for further study. We classified 173 of these objects using their SDSS DR12 spectra. Various activity types were revealed. A special emphasis is made on narrow-line Sy1.0–Sy1.5 galaxies and QSOs, as many of them have soft X-ray, strong FeII lines, and relatively narrow lines coming from BLR. We have retrieved MW data from recent catalogues and carried out statistical investigations for the whole AGN sample. We have analyzed X-ray properties of these sources to find a limit between normal galaxies and X-ray AGN (Paronyan *et al.* 2014a; 2014b; Mickaelian & Paronyan 2014; Mickaelian *et al.* 2016).

After using SDSS12 spectral classification and other available data, our sample of active galaxies and other extragalactic X-ray sources now consists of 4745 objects, including 333 BL Lac and HPQ, 1309 QSOs, 1876 Sy1.0-1.5, 111 Sy1.8-2.0.

2.4. *IRAS PSC/FSC Combined Catalogue and its extragalactic sample*

We have compiled a joint catalogue of IRAS PSC and FSC by means of cross-correlation of these two lists (Abrahamyan *et al.* 2015). We have used a new tool for cross-correlations and as a result we obtained 73,770 associations. In addition, we have made cross-correlations with AKARI-IRC, AKARI-FIS and Wide-Field Infrared Survey Explorer (WISE) catalogues. As a result, we created a catalogue with high positional accuracy and with 17 photometric measurements from 1.25 to 160 μm range, providing a detailed catalogue for 345,163 IRAS point sources. 145,902 (42%) of them are candidate galaxies, subject for our further studies of this large IR selected galaxy sample. The number of genuine galaxies (available from catalogs) is 78,240 and among them 18,567 are found to be AGN (from VCV-13, BZCAT v5 and SDSS QSOs catalogs), which makes their fraction 23.7%. Redshifts are available in total for 22,970 galaxies (mainly AGN and a small part of absorption galaxies).

2.5. *NVSS-FIRST cross-matching and radio variability at 1.4 GHz*

NVSS (1,773,484 sources down to sensitivity 2.5 mJy with accuracy of 15 arcsec rms; Condon *et al.* 1998) and FIRST (946,464 sources down to sensitivity 1 mJy with accuracy of 5 arcsec rms; Helfand *et al.* 2015) have been carried out at the same wavelength (21 cm) or frequency (1400 MHz), which gives a unique possibility to study radio variability of the common sources. The cross-correlation with a dedicated software (http://arvo.sci.am/crosscorrelation/crosscor.html) enabled us to find 556,282 radio

sources present in both catalogs. Using the 3σ criteria we distinguished 6,301 variable radio sources, and with certain limitations specified the 260 strongest radio variables. We cross-correlated these 260 sources with other catalogs at different wavelengths (APM, SDSS DR10, VCV-13, BZCAT, 2MASS, WISE). As a result, we obtained photometric data for optical, NIR, MIR and radio ranges for these 260 variable radio sources to study them in details (Abrahamyan & Mickaelian 2014a; 2014b). The same will be done with all 6,301 sources, which we consider as AGN. These 6301 radio variables have FIRST-NVSS flux differences $\geqslant$ 15 mJy and 1699 radio variables have flux differences $\geqslant$ 50 mJy. Interestingly, when analysing these sources by activity types, blazars do not appear to have the largest variability.

We have also studied the optical variability for these radio sources.

2.6. *ROSAT-NVSS AGN sample*

X-ray and radio radiation is typical of many active galaxies, but there are other X-ray and radio sources as well. However, the cross-correlation of these two kind of sources results in mostly AGN, especially if we use not very deep surveys. We attempted to create an X-ray/radio AGN catalog and make its MW studies. We have cross-correlated ROSAT catalogs with NVSS one with a search radius of 30". 9,193 associations have been found. To distinguish AGN from the normal bright galaxies and clusters, Véron-Cetty & Véron (2010) AGN catalog was used. 3,094 AGN were found. We were left with 6,099 X-ray/radio sources without an optical identification. Brighter objects are normal bright galaxies, while we believe that all faint ones are candidate AGN with some contamination of distant clusters. SDSS spectroscopic survey allows us classify objects by activity types, and a number of our candidate AGN is found to be present in SDSS. We attempt to find connections between the fluxes in different wavelength ranges, which will allow us to confirm AGN and blazars candidates and in some cases, find new ones (Paronyan *et al.* 2014). If we consider all associated sources as AGN (given that both X-ray and radio is present), then they make up 7.4% among 124,730 ROSAT X-ray sources and only 0.5% among 1,773,484 NVSS radio ones. However, it is obvious that these numbers are underestimated, as we consider only one private method to reveal AGN.

3. Summary

By using recent multi-wavelength big data on active galaxies and modern statistical methods, we have accomplished a number of project on search and study of such objects. A number of results were achieved by our search and studies of active galaxies at various wavelength ranges, such as:

• optical identification of 3212 ROSAT FSC X-ray sources; discovery of new bright AGN (QSOs and Seyferts) among X-ray sources, optical identification of 2791 ROSAT FSC X-ray sources (BHRC objects); estimation of the abundance of various types of objects among X-ray sources, discovery of new types of X-ray galaxies: interacting pairs and multiples,

• a new sample of 330 IRAS / 2MASS / FIRST / SDSS selected galaxies, useful for comparative study of properties of IR sources,

• based on the combined ROSAT BSC/FSC catalogue of optical identifications, we estimated AGN content of X-ray sources as 52.9%,

• we have combined IRAS PSC and FSC catalogs and compiled its extragalactic sample, which allowed to estimate AGN content among IR sources as 23.7%.

Many new projects are possible based on our studies and using our data products (DFBS database, published catalogues, etc.). We give in Table 3 the summary of our

projects on search and studies of active galaxies: short project name, years of accomplishment, objectives and number of involved objects.

4. Acknowledgements

References

Abrahamyan, H. V. & Mickaelian, A. M. 2014a, *Proc. IAU Symp. 304. CUP*, 100

Abrahamyan, H. V. & Mickaelian, A. M. 2014b, *Proc. IAU Symp. 304. CUP*, 102

Abrahamyan, H. V., Mickaelian, A. M., & Knyazyan, A. V. 2015, *A&C*, 10, 99; VizieR On-line Data Catalog II/338

Alam, S., Albareti, F. D., Allende Prieto, C., *et al.* 2015, *ApJS*, 219, 12

Arakelian, M. A. 1975, *Commun. BAO*, 47, 3

Condon, J. J., Cotton, W. D., Greisen, E. W., *et al.* 1998, *AJ*, 115, 1693

Cutri, R. M., Wright, E. L., Conrow, T., *et al.* 2013, *AllWISE Data Release, IPAC/Caltech*, VizieR Catalog II/328

Helfand, D. J., White, R. L., & Becker, R. H. 2015, *ApJ*, 801, 26

Hovhannisyan, L. R., Mickaelian, A. M., Weedman, D. W., *et al.* 2009, *AJ*, 138, 251

Kazarian, M. A., Adibekyan, V. Zh., McLean, B., *et al.* 2010, *Ap*, 53, 57; Vizier On-line Data Catalog VII/254

Kazarian, M. A. & Mickaelian, A. M. 2007, *Ap*, 50, 127

Markarian, B. E. 1967, *Ap*, 3, 24

Markarian, B. E., Lipovetsky, V. A., Stepanian, J. A., *et al.* 1989, *Comm. SAO*, 62, 5

Massaro, E., Maselli, A., Leto, C., *et al.* 2015, *Ap&SS*, 357, 75

Massaro, E., Mickaelian, A. M., Nesci, R., & Weedman, D. (Eds.) 2008, *Digitized First Byurakan Survey, ARACNE Editrice, Rome*, 78 p.

Mazzarella, J. M. & Balzano, V. A. 1986, *ApJS*, 62, 751

Mickaelian, A. M. 1995, *Ap*, 38, 349

Mickaelian, A. M. 2008, *AJ*, 136, 946; VizieR On-line Data Catalog III/258

Mickaelian, A. M. 2014, *Proc. IAU Symp. 304. CUP*, 304, 1

Mickaelian, A. M., Abrahamyan, H. V., Paronyan, G. M., & Harutyunyan, G. S. 2013, *AN*, 334, 887

Mickaelian, A. M., Hovhannisyan, L. R., Engels, D., Hagen, H., & Voges, W. 2006, *A&A*, 449, 425

Mickaelian, A. M., Lawrence, A., & Magakian, T. Yu. (Eds.) 2016, *Proc. Symposium, ASP Conf. Ser.*

Mickaelian, A. M., Mikayelyan, G. A., & Sinamyan, P. K. 2011, *MNRAS*, 415, 1061

Mickaelian, A. M., Nesci, R., Rossi, C., *et al.* 2007, *A&A*, 464, 1177

Mickaelian, A. M., Paronyan, G. M., Harutyunyan, G. S., *et al.* 2016, *A&ApTr*, 29, 333.

Mickaelian, A. M. & Sanders, D. B. (Eds.) 2014, *Proc. IAU S304. CUP*

Mickaelian, A. M. & Sargsyan, L. A. 2004, *Ap*, 47, 213

Mickaelian, A. M. & Sargsyan, L. A. 2010, *Ap*, 53, 483

Mickaelian, A. M., Sargsyan, L. A., Gigoyan, K. S., *et al.* 2009, *Romanian Astron. J.*, 18S, 249

Paronyan, G. M., Abrahamyan, H. V., Harutyunyan, G. S. & Mickaelian, A. M. 2014, *Proc. IAU Symp. 304. CUP*, 164

Paronyan, G. M., Harutyunyan G. S., & Mickaelian, A. M. 2014, *Proc. IAU Symp. 304. CUP*, 166

Paronyan, G. M., Mickaelian, A. M., & Abrahamyan H. V. 2014, *Proc. IAU Symp. 304. CUP*, 161

Sargsyan, L. A., Mickaelian, A. M., Weedman, D., & Houck, J. 2008, *ApJ*, 683, 114
Stepanian, J. A. 2005, *RMxAA*, 41, 155
Véron, P., Lindblad, P. O., Zuiderwijk, E. J., *et al.* 1980, *A&A*, 87, 245
Véron-Cetty, M. P., Balayan, S. K., Mickaelian, A. M., *et al.* 2004, *A&A*, 414, 487
Véron-Cetty, M.-P. & Véron, P. 2010, *A&A*, 518, 10