OPEN EUROPEAN JOURNAL ON VARIABLE STARS
March 2008 http://var.astro.cz/oejv ISSN 1801-5964

The Extended Catalog of Red AGB Variable Stars found in the NSVS Database

MAXIM USATOV[1], ARTEM NOSULCHIK[2]
1) Ceskoslovensky Armady 371, 16000, Praha, Czech Republic – maxim.usatov@bcsatellite.net
2) Dniepropetrovsk, Ukraine – artem.nosulchik@bcsatellite.net

---

Abstract

In this paper we present the catalog of new Mira-type and SR+L variable stars in the Northern Sky Variability Survey (NSVS) database, found after complementing the original NSVS object tables with a number of additional parameters. By extending the amount of criterias search queries can take advantage of and filtering known objects out, it became convenient to apply better constrained filters to search new variable star candidates. As an example of presented approach, the quest for red variables has resulted in 77 Mira-type and 717 SR+L stars found that have no identification in General Catalogue of Variable Stars, SIMBAD and VSX databases. Most likely the stars presented are new discoveries. The classification of these stars is based on the near-infrared colors from Two Micron All Sky Survey (2MASS) photometry, amplitude and period with the accuracy of the classification given ~90%, estimated basing on the current GCVS classification scheme.

Methodology

The Northern Sky Variability Survey (Wozniak et al, 2004), with approximately 19 million objects recorded, provides a wealth of data to process and even five years after publication, new discoveries are still made. The original NSVS database contains 19,995,106 objects with more than 3 billions of individual light curve measurements. One of the standard approaches to search objects within the database is to use an SQL query from the SkyDOT website. This provides the ability to constraint searches with basic criteria such as object's coordinates, median object magnitude, standard magnitude deviation, median error and available number of observations. Obviously it gets rather time consuming to find a targeted object type manually using available criterias in such an extensive database, especially when coordinates are unknown. Additionally, since data contains so many semiregular or slow irregular variables of late spectral type, we decided to process the database automatically to extend each object record with a number of additional parameters such as period and it's significance index, Stetson's variability statistics, color indices and amplitude of light variations. We decided to narrow down the search field in a few steps, described below.

Filtering the database: We required the resulting database objects to have at least 100 observational records and the following properties of light-curve parameters: $\sigma > 0.1$ mag and $\sigma/E > 3.0$, where σ is the RMS of magnitude scatter and E is the median of error estimate of all "good" magnitude measurements without problem flags. Similar criterias applied have been described in (Wozniak et al, 2004), however we use softer requirements for the quality of observations. Once all objects not matching the required criteria have been removed, the original NSVS database has been reduced down to 81,358 objects. All objects from the reduced database have been further processed against the combined table of GCVS4 Vols. I-III and NL 67-78 (Kholopov et al, 1985-1988) with improved coordinates to filter out variable stars that are already discovered. The matching box in this proceure was chosen as 14".4 that corresponds to ROTSE-I pixel size. After this operation, we remove all known stars in the Catalog of Red Variables from NSVS according to NSVS object IDs, published on the SkyDot website; detected variables from the 3rd All Sky Automated Survey (ASAS-3) and the list of MISAO project variable stars. The remaining database objects were filtered to exclude so called low quality data (TPLSTATS flag) and only measurements with flags ≤ 2 remained to exclude possible artifacts. This assures that analyzed light curve data does not contain errors such as pixel saturation, truncation by image boundaries, memory overflows, uncorrected photometry, mount flips and similar. After this, we select only those objects that show slow variability pattern, following similar procedure as described in (Wozniak et al, 2004). For this we estimate an AOV ratio of 15-day binned light curve as $R_{AOV} = (<m^2> - <m>^2)^{1/2} / <s>$, where m_i and s_i are i-th bin mean and standard deviation. Only bins with at least five valid points and objects with five or more bins were considered in this analysis. Objects with AOV ratio less than 1.6 are filtered out. More information on such an application of the AOV ratio can be obtained in (Wozniak et al, 2004). Multiple NSVS objects representing the same stars were joined into single records with the first NSVS ID mentioned and followed by it's synonyms. Finally, the new variable candidates database has been reduced down to 7,338 objects after all the processing and filtering out already discovered variable stars and low quality data.

Extending the database with periods and variability indices: Per each remaining object in the reduced database, an automatic period search was done using self-written scripts and a slightly modified TS 1.2.01 FORTRAN time-series analysis program distributed by AAVSO. The TS has been modified to run the light curves through a pipeline of automatic routines. With TS, we applied CLEANest method (Foster, 1995) to determine the true dominant period or light variation timescale. It is shown to be an effective technique for detecting and describing multiperiodic signals, and at least in some cases is capable of divining the true nature of a multiperiodic signal, even when the single strongest peak in a periodogram occurs at a spurious frequency. Period uncertainities ΔP were estimated according to method described in (Rockenfeller et al, 2005):

$$\Delta P \approx \frac{P^2}{2T}, \qquad (1)$$

where T is the total time span of available data and P is the period in question. This approximation is valid for periods with small errors, however we anticipate that for Mira stars with strongly defined periodicities this will be an adequate approach at least for preliminary uncertainity indication. It worths to mention that compared with a 1σ confidence interval on P calculation as per the method described by (Schwarzenberg-Czerny, 1999), equation (1) used in this work is, in many cases, up to a few times more pessimistic. If period found is more than the time span of observational data T, period in the results is then indicated as $>T$, and ΔP field is left blank.

Since running a full False Alarm Probability analysis with multiple hundreds of perturbations per each object would be excessively time consuming, the period significance index has been introduced in order to be able to filter out stars with weak or poorly defined periodicities. The significance index of the strongest period peak has been estimated similarly to a method proposed by (Benko et al, 2007) to filter out noisy data:

$$s = (A_{peak} - < sp >) / \sigma_{sp}, \qquad (2)$$

where A_{peak} is the amplitude of the strongest peak in CLEANest spectrum, $< sp >$ and σ_{sp} respectively are the average and the RMS of the whole frequency spectrum peaks. In addition, Stetson's J, K and L indices (Stetson, 1996) have been automatically estimated using VARTOOLS (Hartman et al, 2007) light curve analysis program, per each object in the reduced database, in order to assist with further manual searches as a measure of coherent intrinsic light curve variability. A light curve that varies steady in time is rated with a higher Stetson indices rather than a curve with Gaussian noise. In this work, we define observations as pair if they are separated by less than 1 day. The use of Stetson's indices to reduce the database has to be done with extreme caution. According to the stars observed by (Pepper & Burke, 2006), J<<1 corresponds to non-variable sources, 1<J<10 to obvious variables and J≈1 to indistinguishable objects. We decided to remain objects in the database for future use independently on their Stetson indices, however will adopt a cutoff at a softer J<0.55 criteria, similar to (Carpenter et al, 2001) for further manual searches.

Photometry and cross-identificators: Each object in the reduced database is then automatically cross-identified with Two Micron All-Sky Survey (2MASS) Point Source Catalog for near-infrared J, H and K$_s$ photometry and coordinates. Similar to GCVS4 filtering procedure, the match radius used is 14".4 and the closest 2MASS object is selected if multiple are found. According to (Wozniak et al, 2004), this matching criteria produces less than 0.4% false matches to 2MASS objects.

Mira-type variables search criteria: Once the database is prepared, we follow the GCVS variability classification scheme introduced in NL 67-72 and GCVS vol. V and define the Mira candidates search criteria as indicated in Table 1. In order to photometrically constraint the selection, we analyze a set of 2162 known

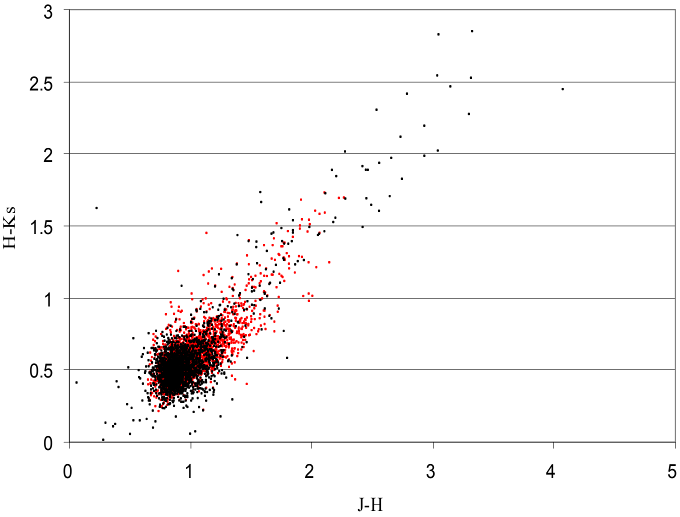

Fig. 1 – Colour-colour relation plot of known Mira stars: black – classified in GCVS catalogue, red – classified in NSVS Red Variables catalogue. Total of 4726 stars displayed.

Mira-type stars classified in the GCVS catalogue and 2564 Mira-type stars classified in the NSVS Red Variables catalogue. For the GCVS catalogue stars, we have processed a coordinate target query in the 2MASS PSC catalogue through VizieR to obtain JHK photometry. A colour-colour relation plot is displayed on Fig. 1.

Table 1 – Mira Candidate Criteria

Parameter	Range of Values
J-H (2MASS)	0.6 to 6
H-K$_s$ (2MASS)	0.2 to 3
Period	80-1000 days
Amplitude (m_{ROTSE})	>1.3 for "M:" >2.0 for "M"

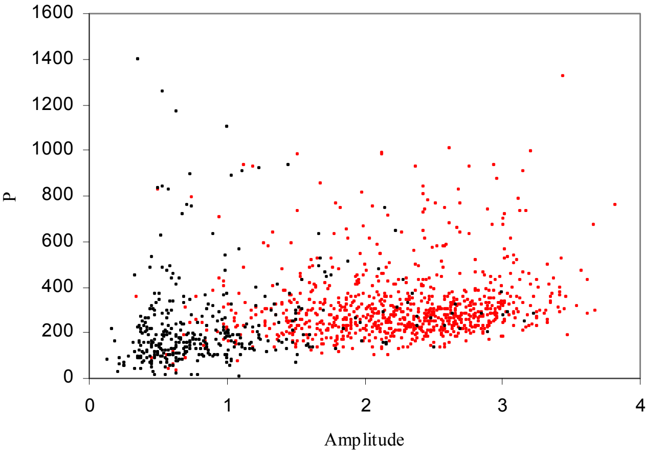

Fig. 2 – Amplitude (m_{ROTSE}) versus Period plot of known Miras, represented red, and SR+L variable stars, shown black, using data from the NSVS database and classified in GCVS catalogue

The light variation amplitudes in ROTSE-1 system may decrease by a factor of few for the AGB stars if compared to standard V-Band measurements. A further conversion from the standard V-band Mira-type light variation amplitude to instrumental magnitudes is thus required as ROTSE-I did not have a filter. For this purpose, we analyze a set of known 873 Mira and 376 semiregular/irregular stars in the NSVS database, present and already classified in the GCVS catalogue, as shown in Fig. 2. All the stars in this set satisfy the $\sigma > 0.1$ mag and $\sigma/E > 3.0$ light curve quality requirements. Obviously, the time span of NSVS data is not enough to have a clear classification cutoff for those classes, as well as there is a high possibility that these classes overlap or some of the stars in the analyzed set are misclassified. Applying cutoff for Mira stars at amplitude greater than 1.3 magnitudes ROTSE, the confusion matrix for the analyzed data results in accuracy as indicated in the Table 2 below.

Table 2 – Confusion Matrix

Actual Class	Predicted Class	
	M (%)	SR+L (%)
M	92.6	7.4
SR+L	18.6	81.4

The accuracy results obtained for the amplitude criteria are in line with the confusion matrix for the two-class problem of red variable stars classification published in (Wozniak et al, 2004) where a Support Vector Machine trained on a few thousand stars has been applied, although amplitude criteria alone is up to a few percents less accurate. At the end we adopt a cutoff at >1.3m for M: classification and >2.0m for M – the latter requirement misclassifies only 6.9% out of all SR+L stars in the analyzed star set. Thus we count the above mentioned criteria as acceptable for such a preliminary classification, if combined with additional color and period constraints. We admit that a further manual check is required to confirm that the light curves of the candidates represent Mira-type stars. As an example, there are 13 stars originally classified as Mira type and manually reclassified to SR+L as in all these objects there was only a single high amplitude observation record responsible for the classification. The NSVS IDs of these objects are: 13606440, 11119349, 9660191, 6806231, 3239390, 2004083, 1377573, 13649055, 13769599, 13812508, 13830647, 14097523 and 15209765. It is also worth to mention that we have tried to apply a method of separating red variables described by (Mattei et al, 1997) which is based on the spectrum instability criteria, however, with the available data, the confusion matrix for the M and SR+L class separation was approximately a few percents worse than with the amplitude criteria used alone, even with the best possible fit of the cutoff line, thus we remained the amplitude as the key criteria.

Table 3 – SR+L Candidate Criteria

Parameter	Range of Values
J-H (2MASS)	0.6 to 6
H-K$_s$ (2MASS)	0.2 to 3
Period	>20 days
Amplitude (m_{ROTSE})	<1.3

SR+L variables search criteria: In this part, we concentrate on selecting the remaining asymptotic giant branch (AGB) stars. Given the relatively short time span of observations available, the most reliable classification that can be given would be SR+L, thus we agree that further research is required to improve the classification into semiregular (SR) and irregular (L) types. Using color and period charts available from (Wozniak et al, 2007), we define the SR+L candidate criteria as shown in Table 3 to the left. The <1.3 mag amplitude criteria is required to separate semiregular and irregular stars from Mira type variables.

Validating results: Once Mira candidates are found, all stars are further automatically checked against SIMBAD and VSX databases to make sure none appear as identified variable stars and to obtain additional cross-identificators. We also run a search in Google using the "NSVS *ID*" query, where *ID* is the identificator of the object in the NSVS database. If at least one result is found, we exclude this object from the publication. This prooved to be an effective technique to filter out stars that already published but are not added to the VSX or SIMBAD catalogues yet.

At last, we visually check each light curve to make sure there are no mistakes passed through the filtering algorithm. With this catalog, each of the variable stars presented is given near-infrared colors from the Two Micron All Sky Survey (2MASS) photometry, 2MASS identifier and position, USNO-A2.0 identifier if available, period derived with CLEANest method (Foster, 1995), it's uncertainity ΔP and significance index *s*, AOV ratio, ROTSE-I magnitude system amplitude, list of SIMBAD identifiers matched with 14.4" box, maximum epoch, minimum and maximum brightness.

Notes

Within this catalog, we can confirm that two stars present in the New catalogue of Suspected Variable stars (NSV) are indeed variable stars. Their IDs are: NSV 24725 and NSV 13218.

Below is the list of objects with notes regarding their respective SIMBAD cross-identificators with NSVS identificator followed by SIMBAD ID and description:

162671, [K98c] Em* 121 – Emission-line Star

2086548, GB6 B0347+5841,F3R 4214,NVSS J035200+585005,WN B0347.8+5841,BWE 0347+5841,87GB 034751.9+584119,TXS 0347+586 – Objects most probably from an overlayed radio source

5875992, 2MASS J20541430+4405257,[CP2005]7 – In North America and Pelican nebulae complex

8381329, HBHA 2203-18,PK 060+001,[KW97] 38-55,IRAS 19400+2422,PN K4-32 – Earlier classified as planetary nebula, however in the list of objects that are *not* a planetary nebulae. SIMBAD comments: NOT a PN (V/84/notpn#181) [01-Jan-2000].

11063779, IRAS 18441+2216,2MASS J18461336+2220109,[PCC93] 300 – Star with envelope of OH/IR type, galactic H2O maser

Acknowledgements

This research has made use of the SIMBAD and VizieR databases operated at the Centre operated at the Centre de Données Astronomiques (Strasbourg) and also of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

This research involved TS code, copyright of The American Association of Variable Star Observers (AAVSO) that is distributed on their website and version 1.1 of The International Variable Star Index (VSX) database, also operated by AAVSO (http://www.aavso.org/vsx/). Additionally, we made use of the General Catalogue of Variable Stars operated by Sternberg Astronomical Insitute of Moscow, Russia (http://www.sai.msu.su/groups/cluster/gcvs/). All the light curves for the NSVS objects can be obtained on the SkyDot website hosted by Los Alamos National Laboratory (http://skydot.lanl.gov/).

References

- Akerlof, C., et al, 2000, AJ, 119, 1901, ROTSE All-Sky Surveys for Variable Stars. I. Test Fields (2000AJ....119.1901A)
- Benko et al, 2007, Acta Astronomica, Vol. 57, pp. 73-85,Variable Star Census in CoRoT "Eyes" (2007AcA....57...73B)
- Carpenter, J. M., et al, 2001, AJ, 121, 3160, Near-Infrared Photometric Variability of Stars toward the... (2001AJ....121.3160C)
- Durlevich O.V., et al, 1994, The list of errors in the GCVS, 4th edition. I. Volumes I-III, Bull. Inf. CDS, 45, 19
- Foster G., The Astronomical Journal vol. 109, no. 4, p. 1889-1902, 1995, The cleanest Fourier spectrum (1995AJ....109.1889F)
- Hartman, J.D., et al, 2007, AAS M. #211, #150.02, A Deep Transit Survey of the Open Cluster M37... (2007AAS...21115002H)
- Kazarovets E.V., et al, 74nd Name-List of Variable Stars, nf. Bull. Variable Stars, IBVS No.4659, 1999
- Kazarovets E.V., Kireeva N.N., Samus N.N., Durlevich O.V., The 77th Name-list of Variable Stars, IBVS No.5422, 2003
- Kazarovets E.V., Samus N.N., 1995, 72nd Name-List of Variable Stars, Inf. Bull. Variable Stars, No.4140 IBVS No.4140, 1995
- Kazarovets E.V., Samus N.N., 70th Name-List of Variable Stars, Inf. Bull. Variable Stars, No.3530 IBVS No.3530, 1990
- Kazarovets E.V., Samus N.N., 73nd Name-List of Variable Stars, Inf. Bull. Variable Stars, IBVS No.4471, 1997
- Kazarovets E.V., Samus N.N., Durlevich O.V., 75nd Name-List of Variable Stars, Inf. Bull. Variable Stars, IBVS No.4870, 2000
- Kazarovets E.V., Samus N.N., Durlevich O.V., 76nd Name-List of Variable Stars, Inf. Bull. Variable Stars, IBVS No.5135, 2001
- Kazarovets E.V., Samus N.N., Goranskij V.P., 71th Name-List of Variable Stars, Inf. Bull. Variable Stars, IBVS No.3840, 1993
- Kholopov P.N., et al, 1985-1988, General Catalogue of Variable Stars, 4rd ed., vols. I-III, Nauka, Moscow
- Kholopov P.N., et al, 67th Name-List of Variable Stars, Inf. Bull. Variable Stars, IBVS No.2681, 1985
- Kholopov P.N., et al, 68th Name-List of Variable Stars, Inf. Bull. Variable Stars, No.3058 IBVS No.3058, 1987
- Kholopov P.N., et al, 69th Name-List of Variable Stars, Inf. Bull. Variable Stars, IBVS No.3323, 1989
- Mattei, J.A., et al, 1997, ESA SP-402 (July 1997), p. 269-274, Classification of Red Variables (1997ESASP.402..269M)
- Messineo, M., et al, 2008, A&A 533, 86 GHz SiO maser survey of late-type stars in the Inner Galaxy. III…
- Pepper, J., et al, 2006, AJ, 132, 1177, Survey for Transiting Extrasolar Planets in Stellar Systems. IV... (2006AJ....132.1177P)
- Rockenfeller B., et al, 2005, A&A 448, Variability and periodicity of field M dwarfs revealed by... (2006A&A...448.1111R)
- Stetson, P. B., 1996, PASP v.108, p.851, On the Automatic Determination of Light-Curve Parameters for... (1996PASP..108..851S)
- Welch and Stetson, 1993, AJ, 105, 1813, Robust variable star detection techniques suitable for.. (1993AJ....105.1813W)
- Wozniak, P.R., et al, 2004, AJ, 127, 2436, Northern Sky Variability Survey: Public Data Release (2004AJ....127.2436W)
- Wozniak, P.R., et al, 2004, AJ, 128, 2965, Identifying Red Variables in the NSVS (2004AJ....128.2965W)