Astronomical Journal of Azerbaijan, 2022, Vol. 17, No. 2

STATISTICALLY OPTIMAL PHENOMENOLOGICAL MODELLING OF VARIABLE STARS

I. L. Andronov [a*], ***V. V. Breus*** [a], ***L. L. Chinarova*** [a],
S. V. Kolesnikov [a], ***L. S. Kudashkina*** [a], ***M. Yu. Pyatnytskyy*** [b]

[a] *Department "Mathematics, Physics and Astronomy", Odessa National Maritime University, Odessa, Ukraine*

[b] *Private Observatory "Osokorky", Kyiv, Ukraine*

A set of algorithms and programs is reviewed, which are most effective for statistically optimal mathematical modelling of various types of variability. Although these methods were proposed and applied for photometry and polarimetry of 2000+ variable stars, as excellent laboratories to study various processes, they may be applied to signals of any nature, e.g. technical, medical or social. For "mono-periodic" stars (with high level of coherence), "global" approximations are most effective - either a trigonometric polynomial of statistically optimal order (mainly for pulsating stars or for eclipsing binaries with smooth curves (EW, EB), or for stars with elliptic variability. However, for Algol-type variables or transits of exoplanets, a "special shape/pattern" approximations superimposed onto the second-order trigonometric polynomial are recommended. For quasi-periodic variations, the wavelet analysis may be used (in some programs, with an adaptive effective width determination), or the scalegram analysis with further local weighted approximation with an optimal width. This method may be effective even in a case of strong flickering, fractal variations, red noise. Wavelet and (especially) scalegram analysis (some modifications are called "multi-resolution", "multi time-scale") may separate variability at different time scales. Particularly, some types suggest long-term aperiodic variability, which may be effectively modelled by an algebraic or trigonometrical polynomial. In a case of trends, the periodogram analysis is to be made using a complete mathematical model instead of super-simplified "detrending" or "pre-whitening", which may cause wrong (biased) results. Some algorithms are illustrated by application to the observations.

Keywords: methods: data analysis – methods: statistical – astronomical databases – AstroInformatics – stars: variables – stars: binaries

* E-mail: tt_ari@ukr.net

Dedicated to the memory of
Prof. Vladimir Platonovich Tsesevich
(11.10.1907 – 28.10.1983),
A great Scientist, Lecturer,
Popularizer and Organizer of Science

1. INTRODUCTION

Variable stars show a wide variety of types of variability. According to the "General Catalogue of Variable Stars" [1], there are 70+ "main" types of variability and hundreds of "mixed" types with two or more types of variability.

This makes impossible to apply a single algorithm to all these signals, e.g. the most popular Fourier Transform (FT) and few its variations (FFT, DFT, DCFT etc.) and improvements based on the mono- or multi-harmonic approximations. It is necessary to have a set of numerous algorithms to apply one or few of them to a concrete object.

For the illustration of some algorithms, we have used the data from the AAVSO International database [2] and the "Open European Journal on Variable Stars".

Recent reviews on these and other methods are in [3], [4], whereas some highlights of the results on pulsating [5] and interacting binary [6] stars are published.

The studies by our group (there are many other co-authors of other published papers) were named in 1983 as an "Inter-Longitude Astronomy" (ILA) project [7], [8] in a collaboration with the "Ukrainian Virtual Observatory" (UkrVO) and "AstroInformatics" projects [9].

In this paper, we show references for original papers with a complete description of the algorithms and/or applications to concrete objects.

2. ALGORITHMS AND SOFTWARE FOR THE DATA ANALYSIS

2.1. General approach

We prefer to use (linear or non-linear) least squares (LSQ) approximations

$$x_C(t) = \sum_{\alpha=1}^{m} C_alpha \cdot f_alpha(t) \qquad (1)$$

with an improved version of the test function [10]

$$\Phi(x_k; C; f) = \sum_{\alpha,\beta=1}^{m} p(t_k, t_j) \cdot w_{kj} (O-C)_k (O-C)_j, \qquad (2)$$

where $x_k,\ k = 1..n$ are observed ("O") values of the signal at the moment (argument) t_k, $x_C(t_k)$ are calculated (C) values at the same moment, weights $w_{kj} = \sigma_0^2 \cdot \mu_{ij}^{-1}$, μ_{ij} is a covariation matrix of errors of observations. The classical approach is to use only diagonal matrices $w_{kj} = w_{kk} \cdot \delta_{kj}$. In this case, $w_{kk} = \sigma_0^2/\sigma_k^2$ (σ_0 is called "the unit weight error", σ_k is a statistical error estimate of x_k. A simplified version $w_{kk} = 1$ is used very often in electronic tables and many papers. The weight (or "filter", or "kernel") function is dependent on the dimensionless "distance" $z = (t - T_0)/\Delta t$, where, according to the "wavelet terminology" [11], T_0 is "shift" and Δt is scale. In the "scalegram" terminology, they are called "trial time" and "filter half-width", respectively [10].

2.2. Periodic Signals

The Fourier Transform (FT) as well as its modifications suggest that the data are distributed "evenly" ("equdistantly", $t_k = t_m + (k - m) \cdot \delta,\ k, m = 1..n,\ \delta$ is time resolution.The duration of observations $D = (t_n - t_1) \cdot n/(n - 1)$, as one may formally extend $t_{k+l \cdot n} = t_k + l \cdot P$ for any integer l. The test periods are resticted to $P_j = D/j = P_1/j$, $j = 1..\text{int(n/2)}$. However, the observations of a star during the night (D) do not generally cover an integer number of periods. This causes a lot of peaks at the periodogram [12], which is computed at a grid of trial frequencies differing with a step of $\Delta f = \Delta\phi/(t_n - t_1)/s$, where $Delta\phi$ is a phase shift between the first and last observations, s is the degree of the trigonometric polynomial ($s = 1$ for a sinusoid), and is recommended to be in the interval [0.04-0.06] (anyway, < 0.1). The Fortran program [12] was later realized in the program MCV ("Multi-Column Viewer") [13], [14]. Moreover, the program MCV has additional features like a periodogram analysis using a polynomial trend superimposed on a trigonometric polynomial of power s, a multi-periodic multi-harmonic approximation with trend, an algorithm of the "artificial comparison star", et al.

Instead of using trigonometric polynomials, one may split observations into L subintervals (for each trial period) and use approximations with spline functions of different order and defect (see [15] for a review and a comparative study).

An alternate group of methods for searching for periodicity is called the "non-parametric" (or "point-point") methods, contrary to "parametric (or "point-curve") methods mentioned above (see [16] for a comparative study).

2.3. Quasi-Periodic and-Aperiodic Signals

[16] There are three main methods to study such signals. The autocorrelation methods are widely used for regularly spaced data (e.g. [17]). Detailed modelling

of the autocorrelation function bias owed to a limited number of de-trended observations was presented by [18], with applications to autoregressive models with noise.

Four-component model of the auto-correlation function for the X-Ray (Chandra) data of the prototype polar AM Her lead to the discovery of the two-component shot noise from the accretion column above the white dwarf. This was interpreted by cutting of the column into 17 pieces due to the magneto-hydrodynamic instability [19].

Wavelet analysis of the irregularly spaced time series by the least-squares method with supplementary weights was introduced by [20], who described spectral and statistical properties of the trial and smoothing functions are studied, and optimal argument and frequency steps are suggested.

This method was applied to almost two hundred semi-regular variables [21]. The decay coefficient of the exponent in the weight function significantly influences of results of the data with long seasonal gaps. Some stars show drastic changes of the amplitude, like transitions between the Mira-type variations and constant stars.

The "scalegram analysis" may be based on any weight function. E.g. the Morlet-type wavelet, which is mentioned above. However, for a strong cycle-to-cycle variability, one has to choose shorter effective intervals. So sines of small angles may be replaced by a line, and cosines - to parabola.

The method of "Running Parabolae" was proposed in [22] as a generalization of the method of "running approximation" ("running mean", "moving average". The weight function for the observation x_k at time t_k was chosen as $p_k = (1 - [|t_k - t_0|/\Delta t]^\alpha)^\beta$. The weight function is zero outside the interval $[t_0 - \Delta t, t_0 + \Delta t]$.

The values $\alpha = 2$ and $\beta = 2$ are sufficient to avoid the discontinuity of the smoothing function and its derivative. The systematic error for the method of "running parabolae" is much smaller as compared with that for "running mean" with the same filter half=width Δt.

The analytical accuracy estimates for the generalized smoothing function, its derivative and the extrema for any basic and weight functions are presented by [23]. Particularly, this formalism was extended to the wavelet analysis [20].

There was introduced a "σ"$-$ scalegram anslysis, where σ is an unbiased weighted mean squared deviation of the observations from the approximation. For nearly periodic signals, it has two characteristic "standstills" at $\Delta t << P$ and and $\Delta t >> P$ with a linking part inbetween. This allows to determine an effective period or a quasi-period (="cycle length") and a corresponding semi-amplitude.

The statistically optimal value of Δt may be determined using the "x_C- scalegram, i.e. the dependence on Δt of the r.m.s. accuracy estimate of the observations at times of these approximations. However, for noisy observations, there may be a local minimum at $\Delta t \approx 0.5P$, then a hump and decrease to a lower asymptotic value $\Delta t/P \to \infty$, which corresponds to global parabolic approximation using a rectangular weight function. This is an often case for visual observations of semi-regular variables and low-amplitude Mira-type stars. Then one may recommend to use the "S/N=SNR" (amplitude "signal-to-noise" ratio. The value of Δt, which corresponds to a maximum of SNR, is typically close (within a few per cent) to a (global or local) minimum of x_C.

In the case of multi-periodic variations, there may be corresponding "stairs". Extremely interesting is a case of fractal variations, which was discovered in a cataclysmic variable AM Her [24]. One may see similar variations in flickering of cataclysmic or symbiotic variables, but for much smaller intervals of Δt.

In the majority of our programs dealing with the "weight" function, the "bisquare" ($\alpha = 2$, $\beta = 2$) version is used. Obviously, there may be other pairs for the approximation, starting from a rectangular ($\alpha \to \infty$, $\beta > 0$, or, alternately, $\alpha > 0$, $\beta = 0,$), triangular ($\alpha = 1$, $\beta = 1$). In the algorithm LOWESS, different sets of parameters (α, β and the degree of polynomial) may be used [25]. E.g. [26] recently used ($\alpha = 3$, $\beta = 3$).

2.4. Times of Extrema

The studies of period variations are done using either the complete light curve (rare) or the moments of the characteristic events (minima, maxima of brightness, or crossing some value (e.g. the $\gamma-$ velocity by a radial velocity). Recently, there is an often abbreviation ToM (Time Of Maximum/Minimum). The deviation $(O-C)$ from a linear ephemeris is often interpreted by the presence of the third (etc. body) around the eclipsing binary (LiTE - Light Time Effect), apsidal motion, mass transfer between the components of eclipsing systems or periodic (the Blazhko effect) or aperiodic variations ([27]). Such studies use only one ToM from a complete night (or even a season), thus needs much smaller number of observations - only a small part of the light curve near an extremum.

An old method was a graphic one based on the hand-drawn light curve. It is not accurately repeating, and there is no correct error estimate. Kwee & van der Woerden [28] proposed a numerical method, based on a search of the point of "symmetry" (or "reflection"). This method was polular for decades. However, it produces a "saw-like" test function, so may produce an underestimated accuracy of ToM, which corresponds to a local minimum (which may be different from the global one because of fluctuations of the test functions).

A parabolic approximation was mentioned in [29]. For asymmetric curves (e.g. of the pulsating variables), a high-order polynomials (of a statistically optimal order) may be recommended [12]. They were used e.g. for determination of the 6509 extrema of 147 semi-regular variables [30].

High-order polynomials may be good at the center of the interval, but may produce formal waves closer to the borders of the interval, and gaps in the observations. Cubic splines seem better than the polynomials with the same number of parameters [15].

Local "asymptotic parabola" fits of signals with almost linear ascending and descending branches connected by relatively short parabolic transitions at maximum or minimum were proposed by [31], [32] and further realized by [33]. This is a kind of spline with a changing fixed order $(1 + 2 + 1)$.

Non-polynomial approximations for generally asymmetric shapes were reviewed by [34]. New "Wall-Supported Polynomial" algoritms [35] may be effective for modelling flat eclipses and exoplanet transitions. Finally, these and other algorithms for statistically optimal determination of phenomenological parameters of extrema were realized in the software MAVKA [36] – totally, 11 classes of functions with 21 individual functions, with automatic determination of the degree of symmetrical or (generally) asymmetrical polynomials.

3. CATACLYSMIC AND SYMBIOTIC VARIABLES

Cataclysmic variables were intensely studied in our group polarimetrically in 1989-2014, based on observations at the Shain 2.6m telescope and the telescope AZT-11 of the Crimean Astrophysical Observatory, by S.V.Kolesnikov and N.M.Shakhovskoy (1931-2011). The history of polarimeters were reviewed by [37], [38], the program for the data reduction was described by [39] and the highlights were briefly listed in [40], [6].

CCD observations were performed within an international campaign ILA ("Inter-Longitude Astronomy") in many countries. The orbital and spin variability of the intermediate polars BG CMi [41], MU Cam [42], V405 Aur [43], V2306 Cyg [44] were studied. All these systems show a spin-up of the magnetic white dwarf due an accretion torque.

"The Noah" project ("40 nights of observations") was performed to study switches of the accretion from one pole to another in an asynchronous polar BY Cam [45], as was theoretically suggested by [46]. During the "Noah-2" project, a third bright accretion region was detected [47]. It may be located near the threat point, similar to another synchronizing polar V1432 Aql [48].

The low-field magnetic dwarf nova DO Draconis exhibited a new type of variability - "transient periodic oscillations" TPO [49], which are distincly different from the "quasi-periodic oscillations" seen in "non-magnetic" Nova-like stars like TT Ari [50].

Long-term photometry of the symbiotic nova V1329 Cyg [51] and V1016 Cyg [52] was analyzed.

4. PULSATING VARIABLES

Long-period pulsating variables (LPV) were studied on old photographic plates from the Odessa "7-camera" astrograph, as well as on published photometric surveys, mainly from the AFOEV and AAVSO databases.

The catalogue of main characteristics of pulsations of 173 semi-regular stars was published by [30] and based on the observations of the AFOEV [53] database, which were previously cleaned from obvious outliers. The catalogue contains results of the periodogram, scalegram and the wavelet analyses.

A series of papers was on Mira Ceti-type variables with H_2O maser emission [54], [55], [56], [57]. For relatively stable light curves of M-type variables, the trigonometric polynomials of statistically optimal order were computed, and the Fourier-related coefficients were published [58], [59].

Phase plane analysis of the photometrical variations of long-period variables was proposed by [60]. Some asymptotic giant branch stars showed a linear decrease of the period [61], [62].

Periodogram analysis of RV Tau-type stars and their classification was presented by [63].

5. ECLIPSING AND "HEART-BEAT" VARIABLES

Phenomenological modeling of the light curves of Algol-type (EA) eclipsing binary stars may be optimally done using the NAV ("New Algol Variable") algorithm instead of using (very high order) trigonometrical polynomials with dozens of parameters [64]. However, it may be applied also to β Lyrae and W UMa-type stars with much smoother light curves [65], which are classically supposed not to have distinct begin and end of eclipses [1]. The accuracy of the period and other parameters may be better by a factor of few times (or even dozens times) than compared to harmonic or multi-harmonic approximations.

The analysis of the $(O-C)$ diagrams was performed to study effects of the mass transfer and presence of the third components in eclipsing systems with circular or elliptic orbits [66], [67], [68],

In the binary system 14 Pegasi a reflection effect due to a highly eccentric orbit was detected [69] "Heart-Beat" variable.

An extreme "Direct Impactor" model was proposed for the system V361 Lyr, where the accretion stream impacts the star, which nearly fills it's Roche lobe without formina a bright accretion disk [70].

6. NEWLY AND RECENTLY DISCOVERED VARIABLE STARS

A new software FVSE (Flexible Variable Star Extractor) for variable stars detection using various variable detection indices was elaborated [71]. The phenomenological approximations of newly discovered eclipsing binaries using the NAV algorithm allows to make estimates of the parameters of these systems, which may be used for further physical modelling [72], [73], [74]. A new variable TIC 230386284 combines EA+UV types [75].

7. CONCLUSIONS

We presented some results of our studies of variable stars of different types with an extended list of references to original papers.

REFERENCES

1. Samus' N.N., Kazarovets E.V., Durlevich O.V., Kireeva N.N., Pastukhova E.N., 2017,Astronomy Reports, **61**, 80

2. Kafka S., 2021, Observations from the AAVSO International Database, https://www.aavso.org

3. Andronov I.L., 2020, Knowledge Discovery in Big Data from Astronomy and Earth Observation, 1st Edition. Edited by Petr Skoda and Fathalrahman Adam. ISBN: 978-0-128-19154-5. Elsevier, 2020, **191**

4. Andronov I.L., Breus, V.V., Kudashkina L.S., 2020, "Development Of Scientific Schools Of Odessa National Maritime University"Collective monograph. Riga, Baltija Publishing,**3**

5. Andronov I. L., Marsakova V.I., Kudashkina L. S., Chinarova L. L., 2014, AASP, **4**, 3

6. Andronov I. L., Andrych K. D., Antoniuk K.A., et al., 2017ASPC..**511**...43

7. Andronov I.L., Antoniuk K.A., Augusto P., et al., 2003, A&AT, **22,** 793

8. Andronov I. L., Antoniuk K.A., Baklanov A.V., et al, 2003, OAP, **23**, 8

9. Vavilova I.B., Pakulyak L. K., Shlyapnikov A. A., et al., 2012, KPCB, **28,** 85

10. Andronov I. L. 1997, A&AS, **125,** 207

11. Andronov I. L., 1998, Kinematics Phys. Celest. Bodies, **14,** 374

12. Andronov I.L., 1994, OAP, **7,** 49

13. Andronov I.L., Baklanov, A.V., 2004, AstSR, **5,** 264

14. Andronov I.L., Baklanov, A.V., 2004, *http://uavso.org.ua/mcv/MCV.ZIP*

15. Andronov I.L., 1987, Publications of the Astronomical Institute of the Czechoslovak Academy of Academy of Sciences, **70**, 161

16. Andronov I.L., Chinarova L.L., 1997, Kinematics and Physics of Celestial Bodies 13 **(6)**, 55-65

17. Jenkins G.M., Watts D.G., 1968, Spectral Analysis and its applications. San Francisco: Holden-Day, **548**

18. Andronov I.L. 1994, AN, **315**, 353

19. Andronov I.L., Burwitz V., Reinsch K., et al. 2003, OAP, **16,** 7

20. Andronov I.L. 1998, Kinematics Phys. Celest. Bodies, **14,** 374

21. Chinarova L.L., 2010, OAP, **23**, 25

22. Andronov I.L., 1990, Kinematika Fiz. Nebesn. Tel, **6**, 87

23. Andronov I.L. 1997, A&AS, **125**, 207

24. Andronov I.L., Kolesnikov S.V., Shakhovskoy N.M. 1997, Odessa Astronomical Publications, **10**, 15

25. Cleveland W., 1979, Journal of the American Statistical Association, **74**, 829

26. Kovtyukh V., Lemasle B., Bono G., et al., 2022, MNRAS, **510**,1894

27. Tsesevich V.P., 1970, Peremennye zvezdy i sposoby ikh issledovaniia (Variable stars and methods of their study). Moskva: Pedagogika, **239**

28. Kwee K.K., van Woerden H., 1956, Bulletin of the Astronomical Institutes of the Netherlands, **12**, 327

29. Tsesevich V.P. 1971, Instationary stars and methods of their investigation. Eclipsing variables (Nestatsionarnye zvezdy i metody ikh issledovaniya), Moskva: Nauka, **352**

30. Chinarova L.L., Andronov I.L., 2000, Odessa Astronomical Publications, **13,** 116

31. Marsakova V.I., Andronov I.L. 1996, Odessa Astronomical Publications, **9,** 127

32. Andronov I.L., Marsakova V.I., 2006, Astrophysics, **49**, 370

33. Andrych K.D., Andronov I.L., Chinarova L.L., Marsakova V.I. 2015, Odessa Astronomical Publications **28**, 158

34. Andronov I.L., 2005, ASPC, **335**, 37

35. Andrych K.D., Andronov I.L., Chinarova L.L., 2017, Odessa Astronomical Publications, **30**, 57

36. Andrych K.D., Andronov I.L., Chinarova L.L., 2020, JPhSt, **24**, 1902

37. Kolesnikov S.V., 2019, JPhSt, **23**, 3901

38. Kolesnikov S.V., 2019, JPhSt, **23,** 4901

39. Breus V.V., Andronov I.L., Kolesnikov S.V., Shakhovskoy N.M., 2007, Astronomical and Astrophysical Transactions, **26,** 241

40. Andronov I.L., 2008, JPhSt, **12**, 2902

41. Kim Y.G., Andronov I.L., Park S.S., Jeon Y.B. 2005, A&A, **441**, 663

42. MU Cam Kim Yong-Gi, Andronov I.L., Park Sung-Su, et al., 2005, JASS, **22**, 197

43. Breus V.V., Andronov I.L., Dubovsky P., et al., 2013, JPhSt, **17**, 3901

44. Breus V., Petrík K., Zola S., 2019, MNRAS, **488,** 4526

45. Mason P.A., Ramsay G., Andronov I., et. al., 1998, MNRAS, **295,** 511

46. Andronov I.L. 1987Ap&SS.**131.**.557

47. Andronov I., Antoniuk K., Breus V., et al., 2008, Open Physics **6**, 385

48. Andronov I.L., Baklanov A.V., Burwitz V. 2006, A&A, **452**, 941

49. Andronov I.L., Chinarova L.L., Han W., Kim Y., Yoon J.-N., 2008, A&A, **486**, 855

50. Kim Y., Andronov I.L., Cha S.M., Chinarova L.L., Yoon J.N., 2008, A&A,**496**, 765 Nova-like cataclysmic variable TT Arietis-QPO behaviour coming back from positive superhumps

51. Chochol D., Andronov I.L., Arkhipova V.P., et al, 1999, Contributions of the Astronomical Observatory Skalnate Pleso **29**, 31

52. Parimucha S., Arkhipova V.P., Chochol D., et al., 2000 Contributions of the Astronomical Observatory Skalnate Pleso **30**, 99-116 18

53. Schweitzer E., 2000, http://cdsarc.u-strasbg.fr/afoev

54. Kudashkina L.S., Rudnitskij G.M.: 1988, Peremennye Zvezdy, **22,** 925

55. Andronov I.L., Kudashkina L.S., Romanenko T.V.: 1992, Peremennye Zvezdy, **23**, 23

56. Kudashkina L.S., Rudnitskij G.M.: 1994, Odessa Astron. Publ., 7/1, **66-69**

57. Kudashkina L.S.: 2012, Odessa Astron. Publ., **25/1**, 18-20

58. Kudashkina L.S., Andronov I.L.: 1996, Odessa Astron. Publ., **9**, 108-111

59. Kudashkina L.S., Andronov I.L.: 2010, Odessa Astron. Publ., **23**, 65-66

60. Andronov I.L., Kudashkina L.S.: 2017, Odessa Astron. Publ., **30**, 93

61. Kudashkina L.S., Marsakova V.I.: 2013, Czestochowski Kalendarz Astronomiczny, **10**, 243

62. Kudashkina L.S., Andronov I.L.: 2010, Odessa Astron. Publ., **23**, 67

63. Kudashkina L.S.: 2020, Annales Astronomiae Novae, **1**, 199-204

64. Andronov I. L., 2012, Ap, **55**, 536

65. Tkachenko M.G., Andronov I.L., Chinarova L.L. 2016, JPhSt, **20**, 4902

66. Tvardovskyi D.E., Marsakova V.I., Andronov I.L. 2017, OAP, **30**, 135

67. Tvardovskyi D.E., Marsakova V.I., Andronov I.L.. 2020, JPhSt, **24**, 3904 Third Components With Elliptical Orbits in the Eclipsing Binaries

68. Pyatnytskyy M., 2021, The Journal of the American Association of Variable Star Observers, **49**, 58

69. Pyatnytskyy M.Yu., Andronov I.L. Research Notes of the AAS,**4**, 24

70. Andronov I.L., Richter G.A., 1987, AN, **308**, 235

71. Breus V.V., 2017, **7**, 38 Advances in Astronomy and Space Physics, 7, 3-5

72. Andronov I.L., Breus V.V., Zola S., 2012, OAP, **25**, 150

73. Andronov I.L., Kim Yonggi, Kim Young-Hee,et al., 2015, Journal of Astronomy and Space Sciences **32,** 127

74. Kim Yonggi, Andronov I,L., Andrych K.D., et al., 2020, JKAS, **53**, 43

75. Pyatnytskyy M., 2020, The Journal of the American Association of Variable Star Observers, **48**, 171/