

UC Berkeley

UC Berkeley Previously Published Works

Title

A catalogue of over 10 million variable source candidates in ZTF Data Release 1

Permalink

<https://escholarship.org/uc/item/6gm640br>

Journal

Monthly Notices of the Royal Astronomical Society, 499(4)

ISSN

0035-8711

Authors

Ofek, Eran O
Soumagnac, Maayane
Nir, Guy
[et al.](#)

Publication Date

2020-11-07

DOI

10.1093/mnras/staa2814

Peer reviewed

A catalog of over ten million variable source candidates in ZTF data release 1

Eran O. Ofek^{★1}, Maayane Soumagnac^{1,2}, Guy Nir¹, Avishay Gal-Yam¹,
Peter Nugent², Frank Masci³, Shrinivas R. Kulkarni⁴

¹ Department of Particle Physics and Astrophysics, Weizmann Institute of Science, 76100 Rehovot, Israel

² Computational Cosmology Center, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

³ Cahill Center for Astronomy and Astrophysics, California Institute of Technology, Pasadena, CA 91125, USA

Accepted ? Received ? in original form ?

ABSTRACT

Variable sources probe a wide range of astrophysical phenomena. We present a catalog of over ten million variable source candidates found in Data Release 1 (DR1) of the Zwicky Transient Facility (ZTF). We perform a periodicity search up to a frequency of 160 day^{-1} , and we classify the light curves into erratic and smooth variables. We also present variability indicators and the results of a periodicity search, up to a frequency of 5 day^{-1} , for about 1 billion sources in the ZTF-DR1 light curve database. We present several new short-period ($< 90 \text{ min}$) candidates, and about 60 new dwarf nova candidates, including two candidate eclipsing systems. Both the 10 million variables catalog and ~ 1 billion source catalog are available online in `catsHTM` format.

Key words: catalogues – stars: variables – stars: novae, cataclysmic variables

1 INTRODUCTION

Source variability allows us to probe a wide variety of astrophysical phenomena, from binary stars and exoplanets to stellar pulsations, cataclysmic and catastrophic events, and mass accretion by massive black holes. In addition, stellar variability is of interest as it contaminates searches for transients. Therefore, comprehensive lists of variable stars are of great importance.

Several productive variable star searches have been carried out. The Catalina Real Time Survey (CRTS) published over 60,000 periodic variables (e.g., Drake et al. 2014a, Drake et al. 2014b, Drake et al. 2014c, Drake et al. 2017). Other searches include the OGLE survey (e.g., Wozniak et al. 2002), All Sky Automated Survey (e.g., Pojmanski 1997, Pojmanski 2000), Vista Vía Láctea survey (VVV; e.g., Minniti et al. 2010), SDSS stripe 82 (e.g., Sesar et al. 2007), and more. Recently, Chen et al. (2020) use the ZTF-DR2 to search periodic variables up to frequency of 40 day^{-1} . They identified about 781,000 periodic variable stars and classified them into several classes.

Here we present a search for variable star candidates in the Zwicky Transient Facility (ZTF; Bellm et al. 2019a) Data Release 1 (DR1). Variable stars are selected based on their large photometric flux root mean square (rms), compared to other stars in the field, or on some indication of periodicity in their light curves. We provide the ZTF/DR1 photometric light curve catalog of about 1.6 billion

light curves, of about a billion unique sources, in `catsHTM` format (Soumagnac & Ofek 2018). For each source, we also list some variability attributes, calculate its periodogram up to a frequency of 5 day^{-1} , and provide the periodogram highest peak significance and frequency. Next, we present a catalog of about 10.7 million variable star candidates. For each variable candidate we provide a list of variability indicators and a periodicity search up to a frequency of 160 day^{-1} .

The structure of this paper is as follows. In §2, we describe the conversion of the ZTF light curves to binary HDF5 files, and the `catsHTM` catalog of sources and their variability properties. In §3, we describe the selection of variable star candidates, while in §4 we present some selected results. We summarize in §5.

2 ZTF-DR1 LIGHTCURVES IN CATSHTM FORMAT

ZTF (Bellm et al. 2019a; Graham et al. 2019) utilizes the 48-inch Schmidt telescope on Mount Palomar, equipped with a 47 deg^2 camera. The ZTF data processing is described in Masci et al. (2019). The ZTF pipeline produces `SExtractor` (Bertin & Arnouts 1996) and `DAOPHOT` (Stetson 1987) catalogs of sources in each image. Here we use only the light curves based on the `DAOPHOT` pipeline. The ZTF photometry is calibrated against the Pan-STARRS1 catalog (Chambers et al. 2016). A color term is fitted, but the magnitudes are calculated assuming the sources have a color of $g - r = 0 \text{ mag}$.

As part of its DR1, the ZTF collaboration released ascii files

★ e-mail: eran.ofek@weizmann.ac.il

Table 1. Columns in the ZTF/DR1 light curve database.

| Column | Description |
|-----------|----------------------------------|
| HMJD | Heliocentric modified Julian Day |
| Mag | ZTF Magnitude |
| MagErr | Magnitude error |
| ColorCoef | Color coefficient |
| Flags | 16bit flag |

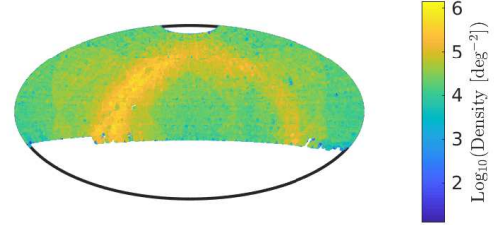
containing 1,681,215,104 light curves of about 10^9 non-unique sources detected in the g and r bands. The sources are non-unique, as they may be detected on several bands, and in several ZTF fields. Each ascii file contains all the g - and r -bands light curves of all the stars in one ZTF field (see [Bellm et al. 2019a](#), [Bellm et al. 2019b](#)). These files are available online¹. We first converted these files into fast-access binary files in HDF5 format², and have made this HDF5 catalog public³. Python and MATLAB programs to access this catalog are available as part of the `catsHTM` toolbox. For each light curve, we also provide some basic variability indicators and list the power and frequency of the highest peak in the classical periodogram (e.g., [Deeming 1975](#)) as calculated up to a frequency of 5 day^{-1} . The HDF5 file names are `ztfLCDR1_<field>.hdf5`, where `<field>` is a zero padded six digit ZTF field number. Each file contains two datasets⁴. The first dataset, named `/AllLC`, is a matrix of all the light curves of all the stars in the field. The light curves are stored one after another. The matrix columns are listed in Table 1 (see also the ZTF-DR1 online documentation⁵). The second dataset, named `/IndAllLC`, is a matrix in which each line corresponds to a light curve of one source/filter in the first dataset. For each source/filter, we store some basic information and also calculate some light curve statistics, and the frequency and power of the highest peak in its periodogram. The matrix columns are listed in Table 2. Among the columns are the indices of the first and last line of the source's light curve in the first dataset (i.e., `/AllLC`).

Next, we generated a `catsHTM`-format ([Soumagnac & Ofek 2018](#)) catalog of all the sources/filters with their properties (as listed in Table 2). For each source/filter, we also stored the exact location of its light curve in the HDF5 light curves file. This combination of `catsHTM` and HDF5 files provides rapid search capabilities by coordinates and fast access to the entire light curve data. Figure 1 shows the sky source density of this catalog.

The reformatted catalogs are available from the `catsHTM` website⁶. `catsHTM` access tools are available in MATLAB⁷ and Python⁸.

3 VARIABLE STAR SELECTION AND PROPERTIES

Here we describe the variable star candidates selection, and their variability properties.

**Figure 1.** The sky surface density of sources in the ZTF-DR1 light curve catalog, in Aitoff projection. The source density is not uniform within a ZTF field/CCD quadrant, where the number of sources is slightly lower toward the edges of the CCD quadrants.

3.1 Candidate Selection

We first estimated the sky-averaged magnitudes rms of sources as a function of magnitude. In order to estimate this, we randomly selected 40,000 pointings across the sky, and for each pointing, we retrieved all the sources within 100 arcsec from its position. In total, about 3.9×10^6 sources were selected within the ZTF footprints. Figures 2 and 3 show, in gray points, the robust⁹ standard deviation (StD) of the light curves of these stars as a function of their mean magnitude in the g - and r -band, respectively. For the region between magnitude 13 and 20, in magnitude bins of 0.1 mag, we calculated the median and the robust StD of these points. The solid lines show the 4-th degree polynomial fitted to the median-plus-six-times-the-robust-StD for the respective band, while the dashed line is the same but for the other band. The approximate 4-th degree polynomials are given in Table 3. We selected variable-star candidates based on their high value of robust StD compared to other stars with similar magnitudes. A caveat of this approach is that we assumed the StD vs. magnitude does not depend on sky position. In practice, this is likely not accurate.

Next, we selected variable-candidate light curves, based on two criteria: (i) Variable stars whose robust StD is larger than a band-dependent, magnitude-dependent threshold, and whose number of data points (in the selected field/band) is larger than 24. The magnitude-dependent threshold was selected to be the black-solid lines in Figures 2 and 3 (see Table 3); or (ii) Stars for which the highest peak in the classical periodogram calculated up to 5 day^{-1} is above 12 (normalized to the noise) and the number of data points is larger than 24. These two criteria left us with 12,761,565 (non-unique sources) candidate light curves.

Next, we merged the sources with light curves into unique sources, if the angular distance between the sources was smaller than 1.5 arcsec. This merging process left us with 10,790,224 variable star candidates. Out of the ≈ 10.8 million stars, 9,805,990 have one or more counterparts within $1.5''$ in the GAIA-DR2 catalog, while 8,530,535 have exactly one counterpart within $1.5''$ in

⁹ Calculated from the central 50% percentile of data and normalized to $1-\sigma$ assuming a Gaussian distribution.

¹ https://irsa.ipac.caltech.edu/data/ZTF/lc_dr1/

² <https://www.hdfgroup.org/solutions/hdf5/>

³ <https://euler1.weizmann.ac.il/catsHTM/>

⁴ In the HDF5 terminology a dataset can be regarded as a table of data.

⁵ <https://www.ztf.caltech.edu/page/dr1>

⁶ <https://euler1.weizmann.ac.il/catsHTM/>

⁷ <https://webhome.weizmann.ac.il/home/efek/matlab/doc/catsHTM.html>

⁸ <https://github.com/maayane/catsHTM>

| Column | Description |
|--------------|--|
| RA | J2000.0 Right Ascension [radians] |
| Dec | J2000.0 Declination [radians] |
| I1 | First line number of source light curve in the light curve HDF5 file |
| I2 | Last line number |
| Nep | Number of epochs |
| ID | Source ID |
| FilterID | Filter ID (1- <i>g</i> ; 2- <i>r</i>) |
| Field | Field ID |
| RcID | CCD/quadrant ID |
| MeanMag | Mean magnitude of source over all epochs |
| StdMag | StD of source magnitude |
| RStdMag | Robust StD of source magnitude |
| MaxMag | Maximum magnitude minus mean magnitude |
| MinMag | Mean magnitude minus minimum magnitude |
| Chi2 | χ^2 , where <i>Nep</i> is the number of degrees of freedom |
| MaxPower | The power (in units of sigma) of the highest peak in the periodogram |
| FreqMaxPower | Frequency of maximal periodogram peak |

Table 2. Columns in the `/IndAllLC` dataset, in the HDF files of all the ZTF-DR1 light curves, and in the `catsHTM` catalog of all ZTF-DR1 sources. χ^2 is the total chi-squared, calculated using the ZTF errors, where the number of degrees of freedom is *Nep* minus one. The classical periodogram was calculated in frequency range of 0 to 5 day⁻¹, with frequency steps of 1/600 day⁻¹.

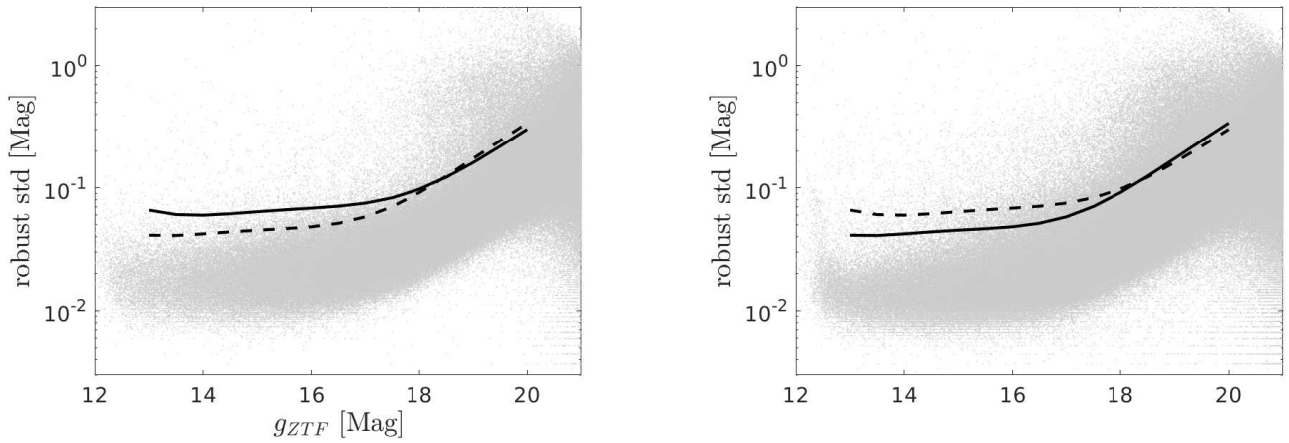


Figure 2. The *g*-band robust rms of magnitude measurements vs. source mean magnitude for about 3.9×10^6 random source entire ZTF footprints. The solid line is the 4th degree polynomial six times the rms, calculated in 0.1 magnitude bins. The dashed line is the same as the solid line but for the *r*-band data. The lines-data is given in

the GAIA-DR2 catalog. The stars without a counterpart (e.g., ghosts, bad pixels) or transient detections (e.g., ghosts, bad pixels) or transient detections are shown in Figure 4. Figure 4 shows the surface density of variable candidate stars in the Galactic plane. There are some regions outside the Galactic plane with anomalous variable-candidates surface density. This is likely due to some issues with the catalog (e.g., underestimation of the noise in some specific fields).

For these ≈ 10.8 million stars, we calculated several properties, discussed in §3.2. This catalog of variable candidates along with their variability properties is available in `catsHTM` format¹⁰ (Soumagnac & Ofek 2018).

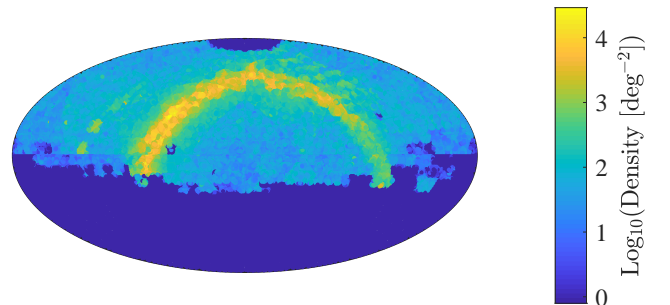


Figure 4. The sky surface density of variable candidates in the ZTF-DR1 light curve catalog, in Aitoff projection. There are some regions outside the Galactic plane with anomalous variable candidates surface density. This is likely due to some issues with the catalog (e.g., underestimation of the noise in some specific fields).

¹⁰ <https://euler1.weizmann.ac.il/catsHTM/>

| Mag (mag) | <i>g</i> -band threshold (mag) | <i>r</i> -band threshold (mag) |
|--------------|-----------------------------------|-----------------------------------|
| 13 | 0.0658 | 0.0411 |
| 13.5 | 0.0605 | 0.0409 |
| 14 | 0.0598 | 0.0421 |
| 14.5 | 0.0615 | 0.0437 |
| 15 | 0.0639 | 0.0452 |
| 15.5 | 0.0662 | 0.0464 |
| 16 | 0.0683 | 0.0481 |
| 16.5 | 0.0708 | 0.0514 |
| 17 | 0.0751 | 0.0581 |
| 17.5 | 0.0831 | 0.0705 |
| 18 | 0.0976 | 0.0915 |
| 18.5 | 0.1222 | 0.1245 |
| 19 | 0.1609 | 0.1736 |
| 19.5 | 0.2189 | 0.2435 |
| 20 | 0.3016 | 0.3393 |

Table 3. The thresholds used for the selection of variable candidates. These are the six sigma lines in the *g*- and *r*-band filters shown in Figures 2 and 3.

3.2 Properties of variable star candidates

For each variable star candidate, we calculated additional variability properties. The full list of attributes in our variable-star candidates catalog is presented in Table 4. Here we discuss selected properties and present their distributions.

The five highest peaks in the periodogram: We recalculated the classical periodogram, this time up to frequency of 160 day^{-1} . We choose to calculate the classical periodogram, and not the Lomb-Scargle periodogram (Lomb 1976, Scargle 1982), due to speed considerations. We kept in the catalog the height and frequency of the five highest peaks in the periodogram.

In Figure 5, we show the first highest-peak frequency vs. second highest-peak frequency in the periodogram for 5,400,501 sources for which the highest periodogram peak is above 12. The vertical, and horizontal lines that are over abundant in sources are mainly due to the window function of the data and aliases with the window function. The lines near frequency of 0.004 to 0.006 day^{-1} corresponds to several months, which is the typical yearly observation period of ZTF per source. The line near the frequency of 0.03 day^{-1} corresponds to the lunar synodic period. Next, we can see a series of lines at about 0.33, 0.5, 1, 2, ..., and up to 13 day^{-1} . These lines correspond to the whole multiplicities of the sidereal day period (i.e., 0.997 day). Finally, the thick line near 48 day^{-1} corresponds to the ZTF minimal mean cadence of about 30 min. The concentration of sources near the whole multipliers of the sidereal day period is mainly due to periodic variables in which the true period (typically with long periods) generated a strong alias with the sidereal period. In fact, in such cases, two peaks are generated, one below and one above the sidereal period, and the frequencies of these two peaks are $1.013 \pm 1/P \text{ day}^{-1}$, where P is the true periodicity and 1.013 day^{-1} is the sidereal day frequency. This explains the bifurcation of many of the lines clearly seen in this diagram, and it allows us to estimate the true periodicity of such sources.

Figure 6 presents the histogram of the highest peak frequency, for sources in which the power of the highest peak is larger than 12. The bins are per 0.003 day^{-1} frequency, but are plotted in logarithmic scale. The peaks discussed in the context of Figure 5 are clearly visible. The jump in the abundance (and also the power, see Figure 6) of sources around frequency of 10 day^{-1} is likely due to

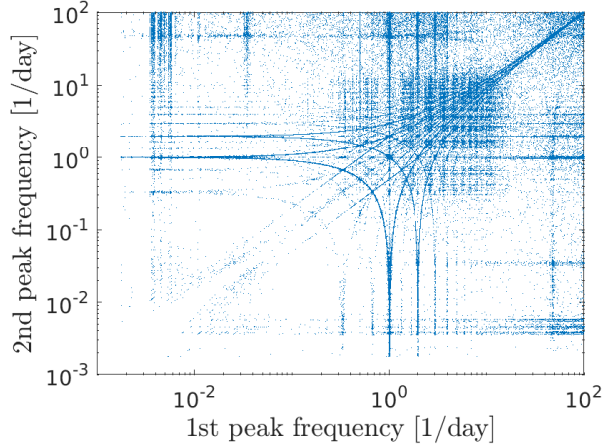


Figure 5. The first highest peak frequency vs. second highest peak frequency in the periodogram, for sources in which the highest periodogram peak is above 12. As discussed in the text, the over abundance in specific frequencies is due to the aliasing of a true period with the window function of the data.

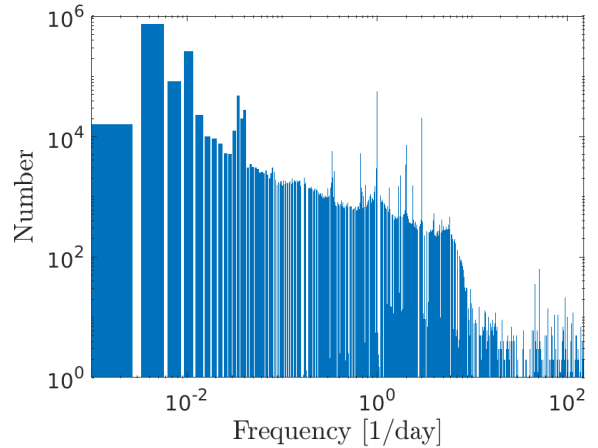


Figure 6. The histogram of the highest peak frequency, for sources in which the power of the highest peak is larger than 12.

loss of efficiency above this frequency (due to the ZTF cadence), and the fact we are using the classical periodogram, which, unlike the Lomb-Scargle periodogram (Lomb 1976), does not normalize the power to the frequency response of the periodogram to a sine wave at that frequency.

Figure 7 shows the periodogram highest peak vs. power. The solid black line marks our periodicity threshold (12), while prominent frequencies are marked at the bottom.

Standard deviation after a high-order polynomial fit: We subtracted from the time of each light curve the mean time and divided it by the time range. Each time-normalized light curve was fitted with a 10th degree polynomial. Next, we calculated the standard deviation of the polynomial-subtracted light curve (i.e., *Std-Poly10*). The motivation for obtaining this property is that variable stars with smooth and slowly evolving light curves (e.g., Mira stars) can be easily identified by their low *StdPoly10*, while eruptive variables (e.g., Dwarf novae) tend to have a high standard deviation even after the removal of such a polynomial. The response of this estimator to eruptive variables is not uniform. The reason is that

| Column | Property | Units | Explanation |
|--------|-------------------|----------------------|--|
| 1 | RA | deg | J2000.0 Right Ascension in the ZTF catalog (Epoch \approx J2018) |
| 2 | Dec | deg | J2000.0 Declination in the ZTF catalog (Epoch \approx J2018) |
| 3 | Napp | | Number of matched sources in ZTF-DR1 |
| 4 | SelBit | | Flag indicating how the source was selected (e.g., STD, periodogram) |
| 5 | NcatDR1 | | Number of ZTF-DR1 LCs within 1.5'' |
| 6 | Nobs | | Number of photometric observations in selected LCs |
| 7 | Ntotobs | | Total number of photometric observations (all fields/filters) |
| 8 | Filter | | Selected filter (1-g; 2-r). |
| 9 | MeanMag | mag | Mean magnitude in selected LC. |
| 10 | MedMag | mag | Median magnitude |
| 11 | MeanErr | mag | Mean error |
| 12 | MedErr | mag | Median error |
| 13 | StdMag | mag | StD of magnitudes |
| 14 | MinMag | mag | Minimum magnitude |
| 15 | MaxMag | mag | Maximum magnitude |
| 16 | RstdMag | mag | Robust StD of magnitudes |
| 17 | RangeT | day | Time range |
| 18 | StdT | day | StD of time |
| 19 | Chi2 | | χ^2 of light curve, where Nobs-1 is the number of d.o.f. |
| 20 | Corr_MagColorCoef | | Correlation coefficient between mag and color term |
| 21 | Mode1frac | | Fraction of measurements in the 1st most populated magnitude bin |
| 22 | Mode1mag | mag | Magnitude of the 1st most populated magnitude bin |
| 23 | Mode2frac | | Fraction of measurements in the 2nd most populated magnitude bin |
| 24 | Mode2mag | mag | Magnitude of the 2nd most populated magnitude bin |
| 25 | Mode3frac | | Fraction of measurements in the 3rd most populated magnitude bin |
| 26 | Mode3mag | mag | Magnitude of the 3rd most populated magnitude bin |
| 27 | StdPoly10 | mag | StD of residuals from fitting 10th degree polynomial |
| 28 | MaxPSpower1 | | Power of 1st highest peak in periodogram |
| 29 | MaxPSfreq1 | day ⁻¹ | Frequency of 1st highest peak in periodogram |
| 30 | MaxPSpower2 | | Power of 2nd highest peak in periodogram |
| 31 | MaxPSfreq2 | day ⁻¹ | Frequency of 2nd highest peak in periodogram |
| 32 | MaxPSpower3 | | Power of 3rd highest peak in periodogram |
| 33 | MaxPSfreq3 | day ⁻¹ | Frequency of 3rd highest peak in periodogram |
| 34 | MaxPSpower4 | | Power of 4th highest peak in periodogram |
| 35 | MaxPSfreq4 | day ⁻¹ | Frequency of 4th highest peak in periodogram |
| 36 | MaxPSpower5 | | Power of 5th highest peak in periodogram |
| 37 | MaxPSfreq5 | day ⁻¹ | Frequency of 5th highest peak in periodogram |
| 38 | Ngaia | | Number of counterparts in GAIA-DR2 within 1.5'' |
| 39 | MagG | mag | GAIA <i>G</i> -magnitude of nearest source |
| 40 | MagBp | mag | GAIA <i>B_p</i> -magnitude |
| 41 | MagRp | mag | GAIA <i>R_p</i> -magnitude |
| 42 | Plx | mas | GAIA-DR2 parallax |
| 43 | PlxErr | mas | GAIA-DR2 parallax error |
| 44 | PM | mas yr ⁻¹ | GAIA-DR2 total proper motion |
| 45 | ExcessNoise | mas | GAIA-DR2 excess noise |
| 46 | z_SDSS | | SDSS redshift |
| 47 | zErr_SDSS | | SDSS redshift error |
| 48 | class_SDSS | | SDSS-spectrum object class |
| 49 | subClass_SDSS | | SDSS-spectrum object subclass |
| 50 | objType_LAMOST | | LAMOST-spectrum object type |
| 51 | class_LAMOST | | LAMOST-spectrum object class |
| 52 | subClass_LAMOST | | LAMOST-spectrum object subclass |
| 53 | z_LAMOST | | LAMOST redshift |
| 54 | zErr_LAMOST | | LAMOST redshift error |

Table 4. Columns available in the ZTF-DR1 variable candidates catalog.

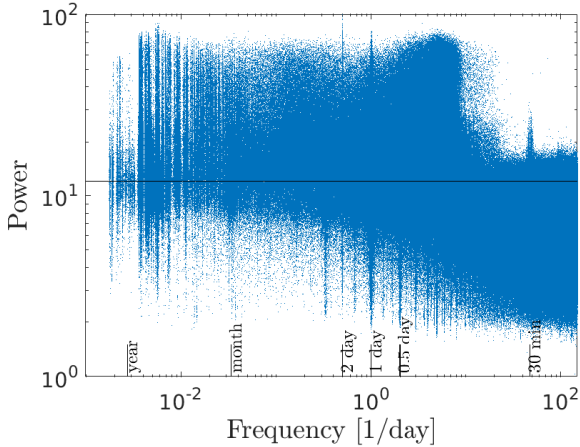


Figure 7. The frequency vs. power of the highest peak in the periodogram for the sources in the variable candidates catalog. The power is normalized by the variance, such that it roughly corresponds to the significance of the peak in units of the noise. The solid line shows the power equal 12 level.

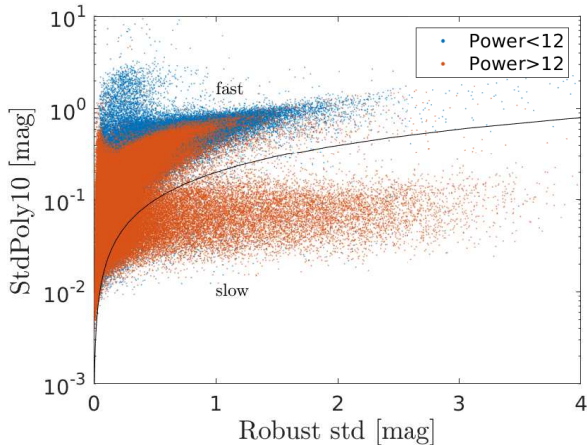


Figure 8. The robust StD vs. the *StdPoly10* (i.e., StD after fitting and subtracting a 10th degree polynomial from the light curve). Blue points correspond to sources whose highest peak in the periodogram is below 12, while red points denote sources whose highest peak in the periodogram is above 12.

different sources were observed with a different time range and sampling.

Figure 8 shows the robust StD vs. the StD after fitting and removing the 10th degree polynomial from the light curve. The solid black line corresponds to $y = 0.2x$, where x and y are the plot axes. Blue points correspond to sources in which the highest peak in the periodogram is below 12, while red points denote sources in which the highest peak in the periodogram is above 12. This plot distinguishes between fast variables (mainly above the black line), and slow variables (mainly below the black line). This plot shows that slow variables are mostly periodic. Furthermore, periodic variables that are classified as fast tend to have low amplitudes compared to the population of fast variables.

Peaks in the magnitude histogram: For each light curve, we calculated a histogram of the magnitude measurements with 0.25 mag bins. We selected the three bins with the highest number of magnitude measurements. For each one of these three

bins, we stored the magnitude at the middle of the bin (i.e., `modelmag`, `mode2mag`, `mode3mag`) and the fraction of measurements in the bin compared to the total number of measurements (i.e., `modelfrac`, `mode2frac`, `mode3frac`). These indicators may be able to assist in identifying variable stars that mainly move from one state to another; examples include eclipsing binaries and dwarf novae. It can also be used to identify stars that are likely flagged as variable candidates due to outliers or due to the mistaken matching of multiple sources as a single star.

4 SELECTED RESULTS

Here we present some selected results related to variable stars. In §4.1 and §4.2, we present a preliminary search for short-period variables, and dwarf novae, respectively.

4.1 short period variability

Given the typical minimum cadence of ZTF (about 30 min), our period search is not very effective above frequencies of 12 day^{-1} (i.e., Nyquist frequency for 30 min sampling). Since the ZTF cadence is not strict, but has some randomness, it is possible to identify periods shorter than about 1 hr. However, any such search for short-period variability will have low efficiency and a high false-alarm rate.

In order to test the possibility of identifying short-period variable candidates, we selected sources using the following criteria: (i) The power of the highest peak in the periodogram > 25 . (ii) The frequency of the first and second highest peaks in the periodogram are $> 18 \text{ day}^{-1}$. (iii) Robust StD $> 0.02 \text{ mag}$. (iv) *StdPoly10* > 1.1 . (v) Magnitude $G < 18$. This resulted in 63 candidates. We inspected all these candidates by eye and found eight promising short-period candidates. The properties of these candidates are listed in Table 5, while their light curves, periodograms, and folded light curves, are presented in Figure 9.

4.2 Dwarf novae candidates

We conducted a partial search for dwarf nova eruptions. We used the following selection criteria: (i) The power of the highest peak in the periodogram > 8 . (ii) The frequency of the highest peak in the periodogram is between 0.01 and 0.1 day^{-1} . (iii) Robust StD $> 0.2 \text{ mag}$. (iv) *StdPoly10* $> 0.3 \text{ mag}$. (v) Magnitude range $> 2.5 \text{ mag}$. This resulted in 452 candidates. We inspected the candidates light curves by eye and selected 76 dwarf novae (DN) candidates. The selection was based on the similarity of the light curves to other known classical dwarf novae. About 60 of the candidates were previously unknown. The candidates are listed in Table 6. Two candidates show some evidence of eclipses. The light curve of the best eclipsing dwarf nova candidate, ZTF J222959.198 + 521507.78, is shown in Figure 10.

5 SUMMARY

We present a catalog of over ten million variable star candidates identified in the ZTF-DR1 light curves database. We provide the entire ZTF-DR1 light curves database in HDF5 format and an index catalog in `catsHTML` format. We also generated a catalog of variable candidates in `catsHTML` format. We provide some variability

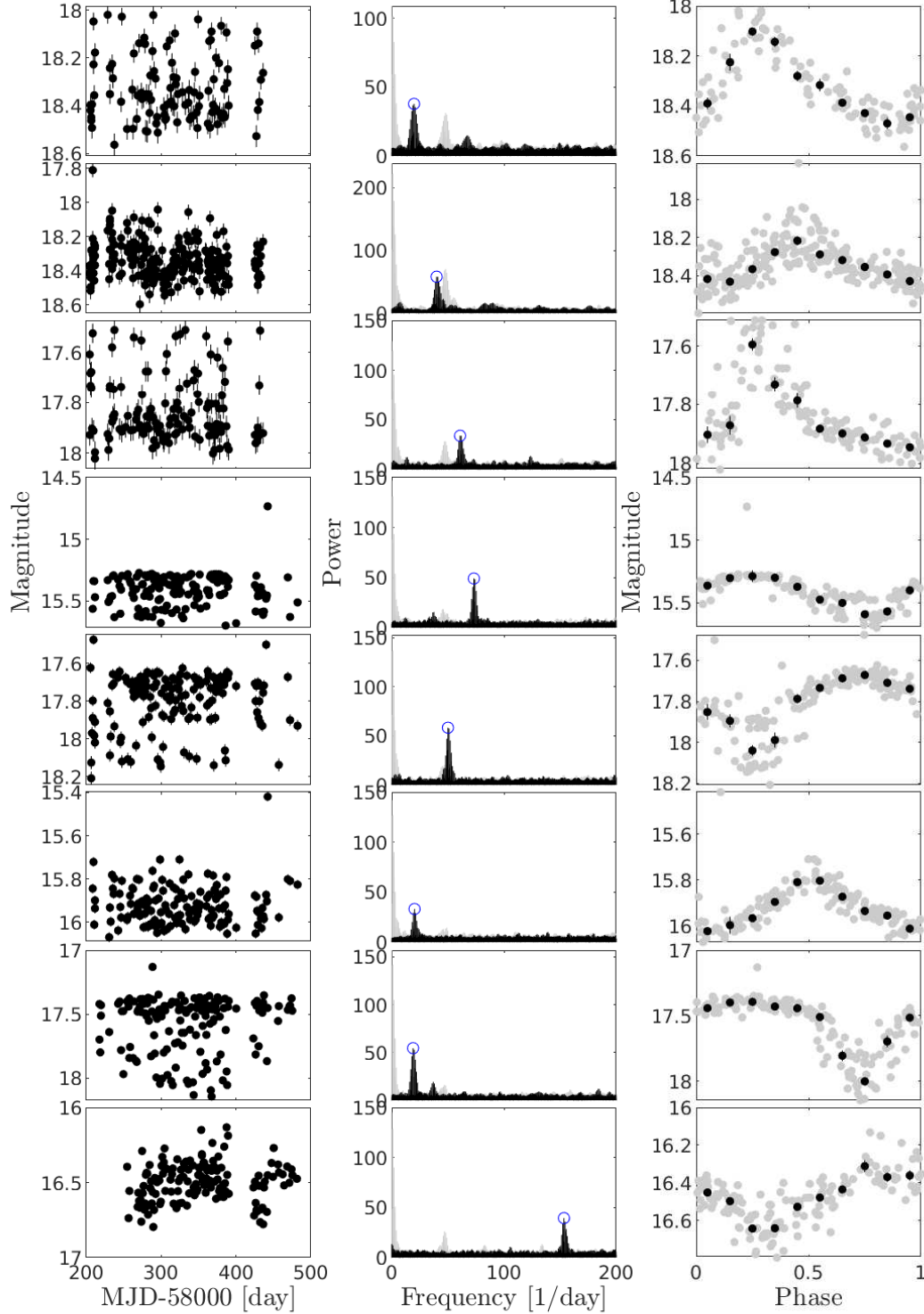


Figure 9. The light curves of eight short-period (< 90 min) variable candidates – row per source. The left column shows the light curves. The middle column presents the power spectra (black), the window functions (gray), and the strongest power-spectrum peaks (blue circles). The right column shows the light curves folded into the period (strongest peak in the periodogram). Gray points are the measurements while black dots are median magnitudes in bins of 0.1 of the period. Candidate coordinates and periods are listed in Table 5 (in the same order).

| J2000.0 R.A. (deg) | J2000.0 Dec. (deg) | Filter (\circ) | Med. Mag. (mag) | Period (min) | Comments (\circ) |
|-----------------------|-----------------------|-----------------------|--------------------|-----------------|-------------------------|
| 290.81069 | 3.89778 | <i>g</i> | 18.36 | 72.44524 | |
| 288.28656 | 12.08099 | <i>g</i> | 18.36 | 36.01894 | |
| 288.67014 | 19.64048 | <i>g</i> | 17.87 | 23.64739 | RGB star |
| 322.73628 | 44.34625 | <i>g</i> | 15.40 | 19.66984 | MGAB-V249 - AM CVn |
| 313.81656 | 46.85180 | <i>g</i> | 17.75 | 28.73726 | nearby X-ray source |
| 320.62207 | 57.32472 | <i>g</i> | 15.94 | 70.86444 | |
| 326.03118 | 58.30484 | <i>r</i> | 17.47 | 75.75160 | |
| 5.74011 | 61.68544 | <i>r</i> | 16.49 | 9.38529 | V1033 Cas (Nova) |

Table 5. List of selected short-period (< 90 min) variable candidates found in the ZTF variability catalog. Med. Mag. is the median magnitude in the ZTF band. Period accuracy is about a few milliseconds. The comments are based on information from SIMBAD (Wenger et al. 2000).

| J2000.0 R.A. (deg) | J2000.0 Dec. (deg) | Filter (\circ) | Min. Mag. (mag) | Max. Mag. (mag) | Med. Mag. (mag) | Comments (\circ) |
|-----------------------|-----------------------|-----------------------|--------------------|--------------------|--------------------|---|
| 313.82849 | -16.44455 | <i>r</i> | 16.65 | 20.19 | 18.74 | SDSS J205518.83-162640.4 |
| 109.17199 | -6.94683 | <i>g</i> | 13.98 | 19.75 | 18.95 | FQ Mon (Mira cand), but looks like a DN |
| 99.22749 | 0.03812 | <i>g</i> | 12.37 | 17.29 | 16.65 | CW Mon (DN) |
| 280.32467 | 4.15413 | <i>r</i> | 17.81 | 20.90 | 20.36 | |
| 93.38401 | 6.95214 | <i>g</i> | 17.39 | 21.46 | 18.37 | |

Table 6. List of selected dwarf novae (DN) candidates found in the ZTF variability catalog. First five lines are shown. The full table is available in the electronic version of the paper.

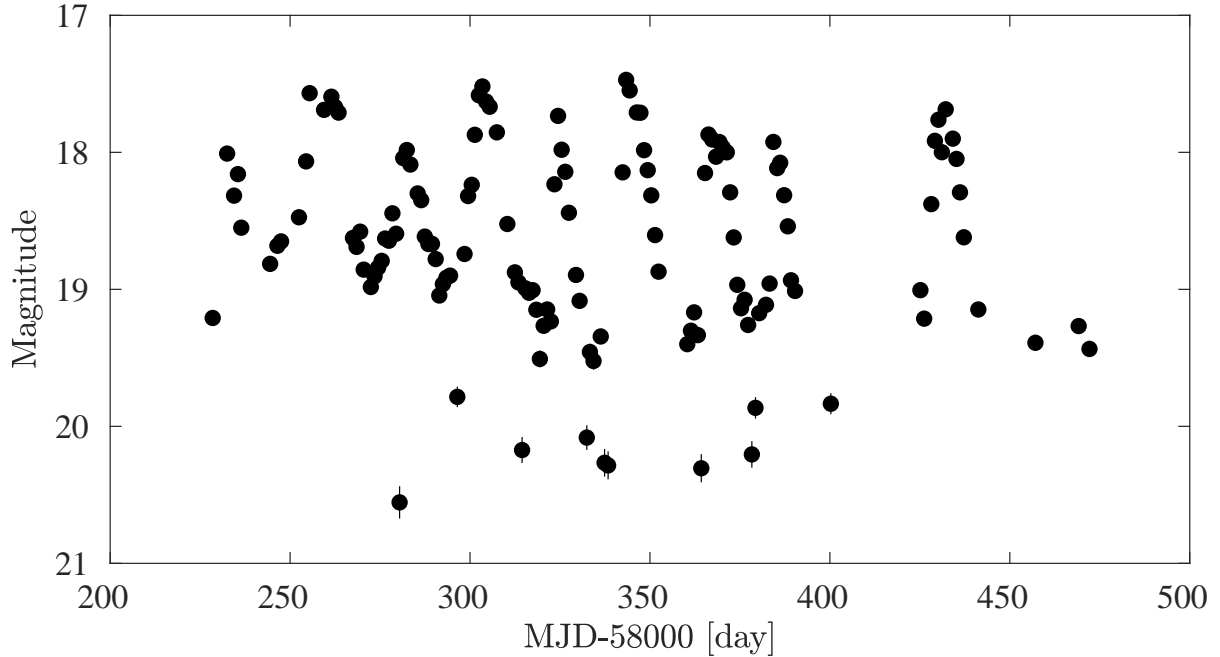


Figure 10. The light curve of the best eclipsing dwarf nova candidate, ZTF J222959.198 + 521507.78. A possible period for the eclipses is 5.9186 h. However, this is not a unique solution and further observations are required.

indicators for the variable star candidates, including a division into erratic (fast) and smooth (slow) variable sources.

We conducted a preliminary search for variable sources with a periodicity shorter than 90 min. We identified eight candidates, one of which is a known AM CVn star. We also searched for outbursting dwarf novae, and found about 60 new dwarf nova candidates. This catalog may also be useful for transient searches where a catalog of stellar variability helps in removing false alarms.

ACKNOWLEDGMENTS

E.O.O. is grateful for the support by grants from the Israel Science Foundation, Minerva, Israeli Ministry of Technology and Science, the US-Israel Binational Science Foundation, Weizmann-UK, Weizmann-Yale, and the Weizmann-Caltech grants.

REFERENCES

- Bellm E. C., et al., 2019a, *PASP*, **131**, 018002
- Bellm E. C., et al., 2019b, *PASP*, **131**, 068003
- Bertin E., Arnouts S., 1996, *A&AS*, **117**, 393
- Chambers K. C., et al., 2016, arXiv e-prints, p. [arXiv:1612.05560](https://arxiv.org/abs/1612.05560)
- Chen X., Wang S., Deng L., de Grijs R., Yang M., Tian H., 2020, arXiv e-prints, p. [arXiv:2005.08662](https://arxiv.org/abs/2005.08662)
- Deeming T. J., 1975, *Ap&SS*, **36**, 137
- Drake A. J., et al., 2014a, *ApJS*, **213**, 9
- Drake A. J., et al., 2014b, *MNRAS*, **441**, 1186
- Drake A. J., et al., 2014c, *ApJ*, **790**, 157
- Drake A. J., et al., 2017, *MNRAS*, **469**, 3688
- Graham M. J., et al., 2019, *PASP*, **131**, 078001
- Lomb N. R., 1976, *Ap&SS*, **39**, 447
- Masci F. J., et al., 2019, *PASP*, **131**, 018003
- Minniti D., et al., 2010, *New Astron.*, **15**, 433
- Pojmanski G., 1997, *Acta Astron.*, **47**, 467
- Pojmanski G., 2000, *Acta Astron.*, **50**, 177
- Scargle J. D., 1982, *ApJ*, **263**, 835
- Sesar B., et al., 2007, *AJ*, **134**, 2236
- Soumagnac M. T., Ofek E. O., 2018, *PASP*, **130**, 075002
- Stetson P. B., 1987, *PASP*, **99**, 191
- Wenger M., et al., 2000, *A&AS*, **143**, 9
- Wozniak P. R., Udalski A., Szymanski M., Kubiak M., Pietrzynski G., Soszynski I., Zebrun K., 2002, *Acta Astron.*, **52**, 129