Recovering variable stars in large surveys: EA$_{up}$ Algol-type class in the Catalina Survey

A. Carmo$^1$, C. E. Ferreira Lopes$^1$, A. Papageorgiou$^{2,3}$, F. J. Jablonski$^1$, C. V. Rodrigues$^1$, A. J. Drake$^4$, N. J. G. Cross$^5$, M. Catelan$^{2,3}$

$^1$Instituto Nacional de Pesquisas Espaciais (INPE), Av. dos Astronautas, 1758 – São José dos Campos – SP, 12227-010, Brazil
$^2$Pontificia Universidad Católica de Chile, Facultad de Física, Instituto de Astrofísica Facultad de Física Av. Vicuña Mackenna 4860, 7820436 Macul, Santiago, Chile
$^3$Millennium Institute of Astrophysics, Santiago, Chile
$^4$California Institute of Technology, 1200 E. California Blvd.Pasadena, CA 91125, USA
$^5$SUPA (Scottish Universities Physics Alliance) Wide-Field Astronomy Unit, Institute for Astronomy, School of Physics and Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

Accepted 2020 August 13. Received 2020 July 26; in original form 2019 December 26

ABSTRACT

The discovery and characterization of Algol eclipsing binaries (EAs) provide an opportunity to contribute for a better picture of the structure and evolution of low-mass stars. However, the cadence of most current photometric surveys hinders the detection of EAs since the separation between observations is usually larger than the eclipse(s) duration and hence few measurements are found at the eclipses. Even when those objects are detected as variable, their periods can be missed if an appropriate oversampling factor is not used in the search tools. In this paper, we apply this approach to find the periods of stars cataloged in the Catalina Real-Time Transient Survey (CRTS) as EAs having unknown period (EA$_{up}$). As a result, the periods of $\sim 56\%$ of them were determined. Eight objects were identified as low-mass binary systems and modeled with the Wilson & Devinney synthesis code combined with a Monte-Carlo Markov Chain optimization procedure. The computed masses and radii are in agreement with theoretical models and show no evidence of inflated radii. This paper is the first of a series aiming to identify suspected binary systems in large surveys.

Key words: methods: data analysis – techniques: photometric – astronomical databases: miscellaneous – stars: variables: general – stars: late-type – stars: low-mass

1 INTRODUCTION

Eclipsing binary systems (EBs) give us important clues about the fundamental basis of stellar evolution since stellar quantities such as mass, radius, and temperature of the components can be directly assessed (Andersen 1991; Torres et al. 2010). Until the end of the twentieth century, EBs were almost exclusively studied on a case-by-case basis (Mowlavi et al. 2017). However, in recent years, the quality and quantity of astronomical data have significantly improved. Projects such as MACHO (Alcock et al. 1996), OGLE (Udalski et al. 1992), WFCAM (Hambly et al. 2008), Kepler (Borucki et al. 2010), CoRoT (Deleuil et al. 2018), Gaia (?), VVV (Minniti et al. 2010), TESS (Ricker et al. 2015), NEOWISE (Mainzer et al. 2011), and in the next few years, PLATO (Rauer et al. 2014) and LSST (LSST Science Collaboration et al. 2009; Ivezić et al. 2019), detected and will detect a large number of variable sources and a great effort is being made to provide tools for the analysis of such huge data sets.

EBs are grouped into three main branches according to the General Catalog of Variable Stars (CGVS; Samus et al. 2017): Algol (EA), β Lyrae (EB) and W Ursae Majoris (EW). In particular, several astrophysical processes like interaction between components, mass transfer and magnetic braking can be investigated using EA systems (e.g. Qian et al. 2018). Another important feature of EA systems is that they may contain low-mass stars whose radii and masses can be known to better than 5% accuracy (often better than 3%; see Feiden 2015). There is a suggestion that radii of low-mass stars are inflated by more than 10% in comparison with theoretical models for isolated stars (Kraus et al. 2011; Birkby et al. 2012; Feiden & Chaboyer 2012; Garrido et al. 2019). This may be associated with non-solar metallicity (Berger et al. 2006; López-Morales 2007), increased magnetic activity (Chabrier et al. 2007; Kraus et al. 2011), or be an observational effect related to distortions in the light curves caused by spots or flares (Morales et al. 2019).
Only a small number of low-mass binaries that are detached EBs with components of late-K or M types present accurate radii determinations in the literature (Garrido et al. 2019).

Most of the known EAs are catalogued in databases such as the OGLE (Soszyński et al. 2016); the GCVS (Samus et al. 2017); the International Variable Star Index (VSX) by the American Association of Variable Star Observers (AAVSO; Watson et al. 2006); the All-Sky Automated Survey (ASAS; Richards et al. 2012b); the LAMOST survey (Qian et al. 2018); the LINEAR survey (Palavera et al. 2013); and the Catalina Real-Time Transient Survey (CRTS) (Drake et al. 2014). Although the number of catalogued EAs is fairly large, the observational strategies in the surveys vary a lot, and the methodologies for finding periods may fail in cases having narrow eclipses or few data points during these events. Ferreira Lopes et al. (2018) discussed the influence of the resolution of the frequency grid in the search for periodicities in EAs. They have succeeded in finding periods for 4 objects previously identified by Drake et al. (2014) as having insufficient number of observations in the eclipses.

This work is the first of a series in an attempt to determine the periods of EAs having narrow eclipses adopting the methodology proposed by Ferreira Lopes et al. (2018). Here we focus on EAs from CRTS with unknown period (EA_u0 class). We have a particular interest in investigating the relatively rare low-mass binary systems. This paper is organized as follows. Section 2 presents the sample of objects analyzed. Section 3 presents a general discussion on the methods of periodic signals search. In section 4, three methods of period search are applied to the sample of interest. Section 5 shows the characteristics of the EAs found in this work, the criteria to select low-mass stars, the method to obtain the individual stellar parameters and an analysis of low-mass stars in the radius inflation context. Finally, the conclusions are presented in Section 6.

2 DATA: THE CRTS EA SAMPLE

The Catalina Real-Time Transient Survey consists of a collaboration in which three telescopes are used aiming to discover Near-Earth Objects (NEOs) and Potential Hazardous Asteroids (PHAs). The project covers the sky in the range of declinations $\delta = [-75^\circ, +65^\circ]$ and avoids crowded regions near to the Galactic plane ($|b| < 15^\circ$). The images are unfiltered to maximize the throughput and the photometry is carried out using the SExtractor photometry package (Bertin & Arnouts 1996). Based on data from the 0.7-m Catalina Schmidt Survey (CSS) telescope, Drake et al. (2014) identified 47,000 variable sources from the public Catalina Data Release 1 (CSDR1; Drake et al. 2012), which together with objects from the other surveys produced an on-line catalog of 61,000 variable objects. Among them, 4680 objects are classified as EA binary systems. From the re-analysis of Papageorgiou et al. (2018), there are 3456 bona fide EA detached systems in that sample; they will be used for run time evaluations in a follow-up subsection in this paper. For the scope of this work, we focus on the 153 EA_u0 systems listed in Drake et al. (2014), which were classified as unknown-period eclipsing binary candidates. In other words, these are objects that present variability typical of eclipsing variables (i.e., excursions to lower states of brightness) but had an insufficient number of observations in the eclipses for full characterization.

1 http://catalina.lpl.arizona.edu/
2 https://catalina.lpl.arizona.edu

3 METHODOLOGY FOR PERIODIC SIGNALS SEARCH

A first step in mining variable stars in large photometric surveys is the detection of changes in a source’s brightness. Once the objects presenting variability have been found, a second step is the search for periodicities (e.g. Wozniak 2000; Shin et al. 2009; Ferreira Lopes et al. 2015a; Ferreira Lopes & Cross 2016, 2017).

The identification of variability does not guarantee that the object is periodically variable. In this sense, EA-type stars can be missed both in target selection or in the periodicity search. The latter happens when the number of measurements at the eclipses is small and because outside the eclipses the variations are essentially due to noise.

The periodicity search methods usually applied to astronomical time series rely on figures of merit. In terms of the associated phase diagram at each frequency grid point, these figures of merit measure correlations or some sort of ordering. Several works discuss the efficiency in finding periodic signals in astronomical time series (e.g. Heck et al. 1985; Swingler 1989; Schwarzenberg-Czerny 1999; Shin & Byun 2004). Graham et al. (2013) tested 11 different methods for 78 types of variable stars. They find that the phase dispersion-based techniques give the best results, but there are clear dependencies on object class and light-curve quality.

In this work, three common period search methods were used: the Generalized Lomb-Scargle periodogram (GLS; Lomb 1976; Scargle 1982; Zecheister & Küster 2009), the String Length method (STR; Dworetsky 1983), and the Phase Dispersion Minimization method (PDM; Stellingwerf 1978, 2011). The three methods mentioned above have figures of merit based on different premises. The performances of these methods to detect EA periodicities are compared in this section.

For unevenly spaced data, as is the case for most of the present surveys, we face an additional difficulty: the Nyquist frequency, $f_N$, has not a precise definition anymore, and the frequency grid is consequently not well defined as well. Ferreira Lopes et al. (2018) derived an expression (see Eq. 1) that parameterizes the frequency resolution in terms of an oversampling factor with respect to the ideal, equally-spaced times series case. The number of frequencies $N_f$ is given by:

$$N_f = \left( \frac{f_{\text{max}} - f_{\text{min}}}{\delta_f} \right) \times T_{\text{tot}}$$

where $f_{\text{max}}$ and $f_{\text{min}}$ are the maximum and minimum search frequencies, $T_{\text{tot}}$ is the total time baseline of the observations, and $\delta_f$ is a parameter that measures meaningful phase variations in the phase diagram when considering changes in frequency. We see that $\delta_f = 1$ corresponds to the minimum frequency sampling for an equally-spaced times series when $f_{\text{min}} = 0$ and $f_{\text{max}} = f_N$. The quantity $1/\delta_f$ is called oversampling factor and has been used in expressions similar to Eq. 1 to define the frequency grid (Schwarzenberg-Czerny 1996a; Debusscher et al. 2007; Richards et al. 2012a; VanderPlas & Ivezić 2015; VanderPlas 2018). The $f_{\text{min}}$ value, even though being formally zero for equally spaced data, is usually defined as $2/T_{\text{tot}}$ to include at least two cycles of the longest period searched in the time series. $f_{\text{max}}$ is the upper limit in frequency, and for an equally spaced times series with step $\delta t$, $f_{\text{max}} = f_N = 0.5/\delta t$. Prior knowledge on the shape of the periodic signal allows us to go far beyond the Nyquist limit, even for the ill-defined case of unevenly spaced data. Examples are the empirical values such as $f_{\text{max}} = 10^4$ (Debusscher et al. 2007; Richards et al. 2012a; De Medeiros et al. 2013) or even larger (Schwarzenberg-
The shape of the light curve has an important role in defining the grid of frequencies. Since eclipsing binaries in general have a strong first harmonic of the fundamental frequency, even in the ideal case of equally spaced data, the description of the light curve (e.g., in terms of Fourier components) would need at least twice the frequency sampling limit needed for the fundamental frequency alone. Figure 3 in Ferreira Lopes et al. (2018) illustrates how the choice of the oversampling factor \((1/\delta\phi)\) is important for detecting periodicities. It contains five phase diagrams of different types of variable stars. Detection of EAs is the most dependent on the oversampling factor. For instance, the use of a value of \(\delta\phi = 0.2\), i.e., an oversampling of a factor of 5, still produces a blurred phase diagram. This means that the frequency grid spacing should be finer to unambiguously identifying the correct orbital period of EAs.

Fig. 1 shows the plane \(N_f\) (number of frequencies in the frequency grid) versus total observation time for a number of recent surveys, given the median sampling time, \(\delta t\), of each survey. The dotted line shows the locus of Eq. 1 for a target \((f_{max}−f_{min}) = 30 d^{-1}\) and \(\delta\phi = 1\). Values of the oversampling factor \(1/\delta\phi\) of 10 and 100 are shown for reference. Surveys such as Kepler, CoRoT and TESS have relatively small values of \(T_{tot}\), but a good cadence, so they can even surpass the exemplified goal of having \((f_{max}−f_{min}) = 30 d^{-1}\). The other surveys have poorer cadences and require extending the natural frequency grid, besides oversampling it to probe timescales of variability of the order of hours. One might ask what are the effects of integration time on the frequency domain. Since the measurements are mathematically equivalent to the convolution of Dirac-\(\delta\) functions (the sampling) with a boxcar (the integration time), the result in the frequency domain is suppression of the high frequencies, as in the case for an equally spaced series. For the CRTS, the integration time is typically 30 sec and the median sampling is ~ 20 min, meaning that the suppression of high frequencies is small.

### 3.1 Run time and hit rate

To test with different values of \(\delta\phi\), we called the GLS, PDM, and STR procedures in a single loop for which the total run time of an object is the sum of the times spent to run each of the three methods. Obviously, the run time is directly proportional to the product \(N_f \times N_p\), where \(N_p\) is the number of points in the light curve. A fiducial mark for this is a run time of 13 sec to explore a dataset with \(N_p = 312\) points, and \(N_f = 2 \times 10^5\) frequencies. We also define an efficiency rate, \(E_{rate}\), which is calculated as follows.

We used the sample of 3456 detached EAs cataloged by Papageorgiou et al. (2018) as a reference since those systems have well determined periods. We ran experiments with an oversampling factor of 100 (or \(\delta\phi=0.01\)), and \(N_f = 2 \times 10^6\). \(E_{rate}\) is the fractional number of systems for which we can recover the correct orbital period with a difference less than 1% from the cataloged value. Table 1 summarizes the results for the different methods. The second column represents direct detections of the fundamental frequency and the third column corresponds to harmonics and sub-harmonics.

| Method | Hits (%) | Multiples (%) | Misses (%) |
|--------|----------|---------------|------------|
| GLS    | 2        | 84            | 14         |
| PDM    | 11       | 82            | 7          |
| STR    | 50       | 33            | 17         |

Periodicity search methods often find values that are half or multiples of the correct period when we take into account that the shape of the light curve of an eclipsing binary (specially detached systems) is very well defined. Cases of exactly equal components are relatively rare compared to the spurious cases. Hits at 1/2 times the correct period are expected, since Fourier-based methods are sensitive to a single harmonic frequency, and the first harmonic of the fundamental orbital frequency may have a larger amplitude than the fundamental. This problem is known and discussed in the literature (e.g., Richards et al. 2012a; Drake et al. 2014). Even the better behaved methods as STR and PDM may be fooled by the presence of particular configurations in the folded light curve or by chance arrangements due to noise. We see that the most reliable method in terms of direct hits is STR, which found the cataloged periodicity for 50% of the objects in the test sample. If we add to this the detections of multiples, \(E_{rate}=83\%\). The count of successes in this form is even better for PDM, with \(E_{rate}=93\%\). GLS also performs well, with \(E_{rate}=85\%\). The three methods combined allow recovering 98% of the periods.

The most common reasons for misses are strong levels of noise, outliers that deviate substantially from the mean noise characteristics and cases with period close to multiples of one day. The GLS method has only ~ 2% of hits, but this is not surprising con-
Figure 2. Phase diagrams and periodograms for four EA\textsubscript{up} CRTS objects with new period determinations. For each object, the bottom panels show the GLS, PDM, and STR periodograms for the main variability period $P$ (left panels) and the first harmonic, $0.5 \times P$ (right panels). A vertical yellow line highlights the main variability period. The periodograms in black use the frequency grid adopted in this work ($\delta\phi = 0.01$). The blue crosses show the periodograms using a coarser grid. The object name and period found are shown at the top of each diagram.

considering that we are treating highly non-sinusoidal light curves with poor sampling in time. The PDM method has a better performance compared to GLS, finding the correct periodicity for 11% of the objects in the test sample, but we have to recall that this method suffers in the case of small number of data points and big gaps in the phase diagram.

4 SEARCHING FOR THE PERIODS OF EA\textsubscript{up} STARS

Guided by the benefits of oversampling and by careful examination of the phase diagrams once a signal is detected, we examined the folded light curves of the 153 EA\textsubscript{up} listed in Drake et al. (2014). Table A1 in the Appendix A shows the parameters related to the period search in the EA\textsubscript{up} sample. Convincing light curves were obtained for 87 sources, i.e., 56% of the EA\textsubscript{up} sample. From now on, we will refer to this set of objects as "iEA": a shortcut for "sample of newly identified eclipsing binaries among the EA\textsubscript{up} of Drake et al. (2014)".

Figure 2 shows periodograms and folded light curves for four iEA objects. The lower panels illustrate how a fine frequency grid (in black, $\delta\phi = 0.01$) improves the detection of signals since the frequency grid points are closer to the true frequency. An inadequate, coarse frequency grid (in blue) may miss completely the correct peaks.

5 RESULTS AND DISCUSSION

To have a broad view about the characteristics of the sample of newly identified objects, we compare in Fig. 3 the properties of iEA systems with those of the EAs previously identified by Drake et al. (2014). The main results for each panel are summarized below:

(a) The time windows of the iEA and EA samples did not present any significant difference, as expected;

(b) The iEAs had a smaller number of observations in comparison with the EA sample, the medians being 200 and 332 observations, respectively. In other words, the iEAs had 40% less observations than the EA sample;

(c) The median magnitudes were 15.08 mag and 15.85 mag for the iEA and EA samples, respectively. It means that the iEAs were slightly brighter than the EAs, on average by 0.77 mag;

(d) Our approach tends to be more efficient in EAs with depth of eclipse shallower than 0.25 mag;

(e) The period distributions of iEAs and EAs showed median values of 2.12 days and 0.87 days, respectively. This indicates we were identifying a number of long-period systems, which have a poor sampling due to the small number of points in the light curve;

(f) The same behavior found in panel (e) was seen when other catalogs are considered. The median periods within the range of 0-8 days were 1.49 days and 1.50 days, for the AAVSO respectively (Watson et al. 2006) and ASAS (Richards et al. 2012b) catalogs.

Overall, our approach detected EA systems having longer periods (lower number of cycles in the total time span of the obser-
Panels (a) to (e) compare the iEA sample with the catalog of EAs with known periods from Drake et al. (2014). Panel (f) compares the period distribution of the iEA with objects having periods up to 8 days in other catalogs from the literature (AAVSO and ASAS).

vations) and smaller number of observations. We used the Mann-Whitney U-test statistics to evaluate if the iEA and EA samples belong to the same parent population. For the relevant parameters (panels c-f in Fig. 3), the null hypothesis that the samples belong to the same parent population is not accepted at the 99% confidence level. Given the sizes of the samples, this suggests that the entanglement of poor sampling in terms of cadence and small number of samples may have an important impact on any attempt to derive statistical properties from time-sparse surveys.

5.1 Low-mass eclipsing binaries

Given the importance of low-mass stars to study theoretical models of stellar structure and evolution, we used color criteria to look for low-mass stars within the iEA sample. The first step was to transform $V_{\text{CSS}}$ to Johnson $V$. We used the expression presented in Drake et al. (2013):

$$V = V_{\text{CSS}} + 0.31 \times (B - V)^2 + 0.04. \quad (2)$$

The $(B - V)$ index came from the APASS catalog (Henden et al. 2016). The photometric uncertainties of the original CSS data were determined using an empirical relationship between source flux and the observed photometric scatter. Graham et al. (2017) presented a correction for the estimated error (see their Fig. 1). An analytical fit to this correction is shown in Papageorgiou et al. (2018), which we used here.

We adopted the color criteria of Papageorgiou et al. (2018) to select low-mass stars candidates. They were: $V - K_s > 3.0$, according to Hartman et al. (2011); $0.35 < J - H < 0.8$ and $H - K_s < 0.45$ mag, based on Lépine & Gaidos (2011) and Zhong et al. (2015). Infrared colors were obtained from the 2MASS $JHK_s$ photometry (Cutri et al. 2003). For both 2MASS and APASS catalogs, we performed a conservative search within a radius of 2″ at the position of each iEA object. Twenty-two objects do not have APASS $(B - V)$ color information; for these, we apply the transformation from 2MASS colors to the Johnson-Cousins system, as provided by Bilir et al. (2008, their Eq. 16).

The literature data were collected from Vizier catalogs using the astroquery modules that handle Vizier and CDS catalogs and Astropy tools$^3$ (Astropy Collaboration et al. 2013; Price-Whelan et al. 2018). The reddening value in the $V$ band was calculated using $E(B - V)$ values by Green et al. (2018) and a total-to-selective extinction ratio $R_V = 3.1$. The Python package dustmaps$^4$ (Green et al. 2015) allows us to use the best-fit (maximum probability density) line-of-sight reddening corresponding to each distance modulus derived from the Gaia DR2. For the $JHK_s$ bands, the interstellar extinction was calculated using the Python package mwdust$^5$ (Bovy et al. 2016), which supports the combination of the following catalogs: Marshall et al. (2006), Green et al. (2015) and Drimmel et al. (2003). Fig. 4 shows the result of the color criteria applied for iEAs as a color-color diagram, in which 8 low mass system candidates are located.

We used the Gaia data to build the color-absolute magnitude

---

$^3$ http://www.astropy.org

$^4$ https://doi.org/10.21105/joss.00695

$^5$ https://github.com/jobovy/mwdust
diagram of EAs as shown in Fig. 5. The extinction values $A_V$ and $E(G_RP - G_P)$ are obtained directly from Gaia DR2. For objects that do not have these values in the catalog, we use a correction based on values of $A_V = 3.1E(B-V)$, as given by Gaia Collaboration et al. (2018, see Eq. 1). As we can see, the 8 iEAs low-mass candidates share the same region in the diagram as the low-mass systems reported by Papageorgiou et al. (2018). They have K-M spectral types with temperatures of about 4,000 K.

5.2 Low-mass stellar parameters

The light curves used in this step present long-term variations which were removed using a linear, parabolic, or sine function as in Papageorgiou et al. (2018, see section 5). Such variations can be related to stellar magnetic activity or to the presence of a third body in the system (e.g. Applegate 1992; Morales et al. 2010b; Ferreira Lopes et al. 2015c; Bours et al. 2016; Almeida et al. 2019). Besides that, the presence of a non-visible third body in an eccentric orbit can cause a rapid orbital precession (Soderhjelm 1975). Therefore, in observations with many cycles, the eclipses may become shallower and shallower a long time and even disappear (Graczyk et al. 2013; Jurysek et al. 2018). For the 8 low-mass stars found, objects CSS J084835.7 $+ 253917$, CSS J020021.5 $+ 213412$, and CSS J071357.2 $+ 342138$ showed eclipses only in the first half part of the time series. Examining in detail, we see that the absence of eclipses in the second half of the light curves is consistent with the combination of the cadence of the observations and the orbital period itself.

The light curves of the low-mass iEA candidates were modeled with the Wilson & Devinney (WD; 1971) code, which is widely used for the analysis of eclipsing binary data. The Markov Chain Monte Carlo (MCMC; Metropolis et al. 1953) method was used to find the modal values of the fitted parameters of the light curve synthesis code from the probability distribution of the parameters, and consequently, an estimate of the errors associated to each parameter. Parameter convergence happens in up to ~ 50,000 iterations. The model had four free parameters: secondary temperature ($T_2$), modified Kopal potentials ($\Omega_1$, $\Omega_2$), and orbital inclination ($i$). The primary temperature $T_1$ was assumed to correspond to the $T_{eff}$ value given in Gaia DR2. Only CSS J080549.6$+403108$ does not have an estimated temperature. In this case, we set $T_1$ based on the Gaia absolute magnitude (combined for the two stars) and the depth of the eclipses, as explained below. The mass ratio $q$ and eccentricity $e$ were fixed such that $q = 1$ and $e = 0$.

To limit the range of values for the parameters in a model fit, we used the total magnitude of the system as an additional constraint. The predicted total apparent magnitude outside the eclipses is estimated taking into account the interstellar extinction and the distance. As in the previous section, we adopted the distances from Baier-Jones et al. (2018) with the revisions from Green et al. (2018). Here the morphological parameters of the phase-folded light curves do not need a physical modeling and they were obtained with the procedure LMinpy (Newville et al. 2016), which provides a least-squares minimization routine for analysis of the data. Two Gaussians for the primary and secondary eclipses plus a constant out-of-eclipses baseline were used as the model.

We assumed that the ratio of the depths of eclipses is proportional to the ratio of the individual luminosities: $\Delta L_1/\Delta L_2 \propto L_1/L_2$. Using this information, we estimated the spectral type of each component. We made use of the temperature-dependent bolometric corrections for stars in the main sequence of Pecaut & Mamajek (2013). This provides a starting point to find a solution, and helps to define the range in which each parameter is searched for. The depth ratio is also used by other authors to estimate temperatures in eclipsing systems (e.g., Armstrong et al. 2014).

Regarding other model parameters, they were fixed according to the following assumptions. The albedos were assumed to be $A_1 = 0.42$ mag.
The lines represent theoretical models from Baraffe blue, and the sample from Garrido et al. (2019) in black and gray.

The parameter ratios obtained from the WD code fitting of the light curves are very reliable. However, there is a large degeneracy in the determination of the absolute values of each component. To circumvent this problem, we followed the procedure described in Coughlin et al. (2011, Sec. 5). It combines the observed effective temperature of the entire system, parameter ratios from the WD fitting, and theoretical values of radii and temperatures of low-mass stars. We adopted the effective temperature from Gaia, which has a typical accuracy of 324 K, for sources brighter than $G = 17$ mag and having $T_{\text{eff}}$ in the range $3,000 - 10,000$ K. We used this error for all eight low-mass candidates, including CSSJ003441.8 – 135033 and CSSJ080549.6 + 403108, which have $G \sim 17.3$ mag. From the WD fitting, we obtained the temperature ratio and $R_{\text{sum}}$, which is defined by Coughlin et al. (2011) as the sum of the stellar radii divided by the binary’s semi-major axis. The theoretical values of radius and temperature are those presented in Baraffe et al. (1998). We assumed an age of 5.0 Gyr and $[M/H] = 0.0$ for $0.075 < M(M_{\odot}) < 1.0$. The resulting absolute parameters are shown in Table 2.

Having fixed the temperatures, masses, and radii of the binary components, a second WD fit was performed to refine the inclination and potential values. The results are shown in Table B2. Figure 6 shows the final fit obtained from Table 2 and Table B2. We verified whether the fit parameters produce a distance consistent with that provided by Bailer-Jones et al. (2018). To calculate the distances, we corrected the Johnson $V$ magnitude using the extinction provided by Green et al. (2018) and $R_V = 3.1$. The calculated distances (see Table B2) are in agreement with those found by Bailer-Jones et al. (2018). For CSS J162549.3+102124, the distance is underestimated in about 22%, while for the other systems the distances differ by no more than 10% from Bailer-Jones et al. (2018).

These differences can be related to the $T_{\text{eff}}$ estimate since the luminosity is proportional to $T_{\text{eff}}^4$. For instance, the $Gaia$ distance for CSS J162549.3+102124 is recovered if we assume a temperature $\sim 250$ K higher than that used by us.

### 5.3 Low-mass binaries and the radius inflation problem

Garrido et al. (2019) characterized a sample of 230 detached close-orbiting eclipsing binaries with low-mass main-sequence components ($M < 1M_{\odot}$), orbital periods shorter than 2 days, and temperatures below 5720 K. They suggest a trend according to which low-mass stars would have inflated radii. Besides, they found that around 61% of the sample has the secondary star more inflated in radius than the primary.

In addition, Coughlin et al. (2011) reported 95 low-mass eclipsing binaries in the initial Kepler data release with periods as long as 10 days. They presented evidence that the radius inflation of low-mass stars in binary systems decreases for longer periods: for $P < 1.0$ days, the median value of the difference between the fits and theoretical radii was about 13.0%, whereas for $1.0 < P(\text{days}) < 10.0$ and $P > 10.0$ days, the value was 7.5% and 2.0%, respectively.

Figure 7 presents a mass-radius diagram for low-mass stars in binaries, where the objects found in this work are shown in red and blue, and the sample from Garrido et al. (2019) in black and gray. The lines represent theoretical models from Baraffe et al. (1998) with isochrones of 1, 5 and 8 Gyr for solar metallicity and helium abundance $Y = 0.275$. The data from Garrido et al. (2019) have most points systematically above the theoretical tracks, for both primary and secondary stars.

Our sample is well distributed around the theoretical model predictions and the median value of the difference between our results and the theoretical radii was about 4.8%. The error bars were symmetric, in the sense that they were larger for negative residuals (Table 2). Only three systems of our sample had periods shorter than two days, unlike the periods considered by Garrido et al. (2019). Even though we have a small sample, our results do not confirm the trend of inflated radii for the components of low-mass eclipsing binary systems.

### 6 CONCLUSIONS

We reported the first determination of orbital periods for 87 EA-type eclipsing binary systems cataloged in Drake et al. (2014) as “unknown period eclipsing binaries” (EA$_{\text{up}}$), using the approach proposed by Ferreira Lopes et al. (2018). This was $\sim 56\%$ of the total number of objects in the (EA$_{\text{up}}$) sample. We recovered periods in data sets of poorer quality compared with previous attempts.

The sample of iEA binaries selected among the EA$_{\text{up}}$ objects in this work tends to show longer orbital periods, slightly brighter objects, shallower eclipses, and, of course, fewer data points than the EAs identified in the original work of Drake et al. (2014). The results exemplify how oversampling the natural frequency grid provided by the median sampling time can be effective in finding new objects having narrow eclipses in sparse time series. In future efforts, we intend to analyze other large surveys in an attempt to retrieve objects hidden by poor cadence and a small number of measurements.

A sub-sample of 8 eclipsing binaries with K and M spectral types was selected. We have determined the stellar parameters of the binary components by modeling their light curves with the Wilson-Devinney code. Our sample did not show clear evidence for inflated radii, either in primary or secondary components. As the radius inflation is expected to be stronger for systems with shorter periods, we cannot be sure that there are only 3 objects with periods less than 2 days in our sample of low-mass binaries. Additional observations with radial velocities would help to constrain and refine the present individual solutions.

### ACKNOWLEDGEMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. C. E. F. L. acknowledges a post-doctoral fellowship from the CNPq. N. J. G. C. acknowledges support from the UK Science and Technology Facilities Council. The authors thank to MCTIC/FINEP (CT-INFRA grant 0112052700) and the Embrace Space Weather Program for the computing facilities at INPE. C. V. Rodrigues thanks the grant #2013/26258-4 from São Paulo Research Foundation (FAPESP) and CNPq (Proc. 303444/2018-5). Support for A. P. and M. C. is provided by Proyecto Basal AFB-170002; by the Chilean Ministry for the Economy, Development, and Tourism’s Millennium Science Initiative through grant IC120009, awarded to the Millennium Institute of Astrophysics (MAS); and by FONDECYT project #1171273.
Table 2. Absolute parameters of the stellar components of the low-mass binary system candidates. The errors were estimated considering an accuracy of the effective temperature from \textit{Gaia} of 324 K.

| Name                  | ID                | Period (days) | $T_\text{eff}$ (K) | $T_\text{p}$ (K) | $T_\text{k}$ (K) | $M_1$ (M$_\odot$) | $M_2$ (M$_\odot$) | $R_1$ (R$_\odot$) | $R_2$ (R$_\odot$) | $\alpha$ (R$_\odot$) |
|-----------------------|-------------------|---------------|------------------|------------------|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| CSS J003441.8-135033  | 1012004004843     | 1.97420921    | 3936             | 3938             | 3844             | 0.61$^{+0.16}_{-0.14}$ | 0.58$^{+0.08}_{-0.06}$ | 0.56$^{+0.05}_{-0.04}$ | 0.53$^{+0.03}_{-0.02}$ | 7.03$^{+0.58}_{-0.49}$ |
| CSS J145100.7+052843  | 1104080006168     | 5.06655354    | 3954             | 3974             | 3932             | 0.62$^{+0.13}_{-0.11}$ | 0.61$^{+0.09}_{-0.08}$ | 0.43$^{+0.02}_{-0.01}$ | 0.42$^{+0.02}_{-0.01}$ | 13.33$^{+0.62}_{-0.60}$ |
| CSS J162549.4+102124  | 1109080703294     | 2.07041048    | 3828             | 3853             | 3800             | 0.59$^{+0.20}_{-0.19}$ | 0.57$^{+0.10}_{-0.09}$ | 0.51$^{+0.03}_{-0.02}$ | 0.49$^{+0.03}_{-0.02}$ | 7.17$^{+0.38}_{-0.34}$ |
| CSS J020021.5+213412  | 112101041164      | 2.32385716    | 3819             | 3819             | 3819             | 0.54$^{+0.20}_{-0.19}$ | 0.56$^{+0.10}_{-0.09}$ | 0.57$^{+0.03}_{-0.02}$ | 0.57$^{+0.03}_{-0.02}$ | 7.74$^{+0.41}_{-0.39}$ |
| CSS J084835.7+253917  | 112604006161      | 2.43505230    | 4510             | 4587             | 4406             | 0.74$^{+0.08}_{-0.08}$ | 0.74$^{+0.08}_{-0.08}$ | 0.64$^{+0.02}_{-0.02}$ | 0.60$^{+0.02}_{-0.02}$ | 8.76$^{+0.72}_{-0.70}$ |
| CSS J071357.2+342138  | 113502018057      | 4.87375954    | 3868             | 3896             | 3835             | 0.60$^{+0.17}_{-0.16}$ | 0.58$^{+0.10}_{-0.09}$ | 0.62$^{+0.03}_{-0.02}$ | 0.59$^{+0.03}_{-0.02}$ | 12.79$^{+0.95}_{-0.90}$ |
| CSS J080549.6+603108  | 114003402721      | 0.66901017    | 3512             | 3514             | 3510             | 0.39$^{+0.23}_{-0.21}$ | 0.39$^{+0.19}_{-0.18}$ | 0.30$^{+0.04}_{-0.04}$ | 0.30$^{+0.04}_{-0.04}$ | 2.96$^{+0.42}_{-0.40}$ |
| CSS J090355.4+533132  | 115201059450      | 0.66327468    | 3962             | 3963             | 3961             | 0.62$^{+0.13}_{-0.11}$ | 0.62$^{+0.09}_{-0.08}$ | 0.61$^{+0.05}_{-0.04}$ | 0.61$^{+0.05}_{-0.04}$ | 3.44$^{+0.18}_{-0.16}$ |

Figure 6. Phase diagrams (top panels) and residuals of the fit (bottom panels) for the eight low-mass binary system candidates (for more detail see Sect. 5.2). The blue line shows the best fit solution using the parameters found in tables 2 and B2.
Figure 7. Radius as function of mass for low-mass stars in detached binary systems. Blue and red squares represent the primary and the secondary components of our low-mass candidates, respectively. Black and gray dots represent data from Garrido et al. (2019). The theoretical models from Baraffe et al. (1998) for 1, 5, and 8 Gyr are represented as blue, orange, and green solid lines, respectively.

DATA AVAILABILITY

The data underlying this article are available in the article and in its online supplementary material.

REFERENCES

Alcock C., et al., 1996, ApJ, 461, 84
Almeida L. A., et al., 2019, AJ, 157, 150
Andersen J., 1991, A&ARv, 3, 91
Applegate J. H., 1992, ApJ, 385, 621
Armstrong D. J., Gómez Maqueo Chew Y., Faedi F., Pollacco D., 2014, MNRAS, 437, 3473
Astropy Collaboration et al., 2013, A&A, 558, A33
Bailer-Jones C. A. L., Rybizki J., Fouesneau M., Mantelet G., Andrae R., 2018, AJ, 156, 58
Baraffe I., Chabrier G., Allard F., Hauschildt P. H., 1998, A&A, 337, 403
Berger D. H., et al., 2006, ApJ, 644, 475
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bilir S., Ak S., Karaali S., Cabrera-Lavers A., Chonis T. S., Gaskell C. M., 2008, MNRAS, 384, 1178
Birkby J., et al., 2012, MNRAS, 426, 1507
Borucki W. J., et al., 2010, Science, 327, 977
Bours M. C. P., et al., 2016, MNRAS, 460, 3873
Bovy J., Rix H.-W., Green G. M., Schlafly E. F., Finkbeiner D. P., 2016, ApJ, 818, 130
Chabrier G., Gallardo J., Baraffe I., 2007, A&A, 472, L17
Coughlin J. L., López-Morales M., Harrison T. E., Ule N., Hoffman D. I., 2011, AJ, 141, 78
Cutri R. M., et al., 2003, 2MASS All Sky Catalog of point sources.
Damerdey Y., Klotz A., Boer M., 2007, AJ, 133, 1470
De Medeiros J. R., et al., 2013, A&A, 555, A63
Debosscher J., Sarro L. M., Aerts C., Cuypers J., Vandenbussche B., Garrido R., Solano E., 2007, A&A, 475, 1159
Deleuil M., et al., 2018, A&A, 619, A97
Drake A. J., et al., 2012, in American Astronomical Society Meeting Abstracts #219. p. 428.20
Drake A. J., et al., 2013, ApJ, 763, 32
Drake A. J., et al., 2014, ApJS, 213, 9
Drimmel R., Cabrera-Lavers A., López-Corredoira M., 2003, A&A, 409, 205
Dwo etsky M. M., 1983, MNRAS, 203, 917
Feiden G. A., 2015, in Rucinski S. M., Torres G., Zejda M., eds, Astronomical Society of the Pacific Conference Series Vol. 496, Living Together: Planets, Host Stars and Binaries. p. 137 (arXiv:1506.02715)
Feiden G. A., Chaboyer B., 2012, ApJ, 757, 42
Ferreira Lopes C. E., Cross N. J. G., 2016, A&A, 586, A36
Ferreira Lopes C. E., Cross N. J. G., 2017, A&A, 604, A121
Ferreira Lopes C. E., Dékány I., Catelan M., Cross N. J. G., Angeloni R., Leão I. C., De Medeiros J. R., 2015a, A&A, 573, A100
Ferreira Lopes C. E., Dékány I., Catelan M., Cross N. J. G., Angeloni R., Leão I. C., De Medeiros J. R., 2015b, A&A, 573, A100
Ferreira Lopes C. E., Leão I. C., de Freitas D. B., Canto Martins B. L., Catelan M., De Medeiros J. R., 2015c, A&A, 583, A134
Ferreira Lopes C. E., Cross N. J. G., Jablonski F., 2018, MNRAS, 481, 3083
Ferreira Lopes C. E., et al., 2020, MNRAS, 496, 1730
Gaia Collaboration et al., 2018, A&A, 616, A13
Garrido H. E., Cruz P., Diaz M. P., Aguilar J. F., 2019, MNRAS, 482, 5379
Graczyk D., et al., 2011, Acta Astron., 61, 103
Graham M. J., Drake A. J., Djorgovski S. G., Mahabal A. A., Donalek C., Duan V., Maker A., 2013, MNRAS, 434, 3423
Graham M. J., Djorgovski S. G., Drake A. J., Stern D., Mahabal A. A., Glikman E., Larson S., Christensen E., 2017, MNRAS, 470, 4112
Green G. M., et al., 2015, ApJ, 810, 25
Green G. M., et al., 2018, MNRAS, 478, 651
APPENDIX A: PERIODS FOR THE IEA SAMPLE

Table A1 shows the periods and the corresponding uncertainties for the IEA sample obtained in this work. These objects were found within the EA$_{w}$ sample of objects with unknown periods from Drake et al. (2014). The methodology used is described in detail in Section 3. The CRTS name and the periods are in the first and second columns, respectively. $\sigma_p$ is the error in the period considering the region of the eclipses in the phase diagram. $V$(mag) is the CSS magnitude, $A$ is the amplitude and $\sigma_A$ is the amplitude error considering the region outside the eclipses in the phase diagram. $\log(T_{\text{rel}}/P)$ shows the number of cycles in the total time range of the observations in logarithmic scale. Large values of $\log(T_{\text{rel}}/P)$ represent better period estimates. For more information, see Ferreira Lopes et al. (2018).

APPENDIX B: PARAMETERS FOR LIGHT CURVE SOLUTION

As described in Section 5.2, the low-mass star candidates were modeled with the WD light curve synthesis code combined with an MCMC optimization procedure. Table B1 shows the parameters and the ratio of parameters found in the light curve modeling. These values were obtained as described by Coughlin et al. (2011) and were used to find the absolute parameters shown in Table 2. The temperatures obtained by Coughlin et al. (2011) were used in the WD code in a second fit, with $T_1/T_2$ fixed in order to refine $i$, $\Omega_1$ and $\Omega_2$ for a final solution (see Table B2). The first and second fits made with WD code produce $T_2/T_1$ and $(R_1 + R_2)/a$ consistent to within 9%. Using the fit parameters we can recover the distances, which are in agreement with those found in the literature. The potentials of CSS J145100.7+052843 and CSS J071357.2+342138 are relative large; this is due to the fact that there are few photometric points in the eclipse, which probably leads to underestimated values of the sum of the relative radii.

This paper has been typeset from a TeX file prepared by the author.
### Table A1. Parameters for the IEsA identified in this work. The period in days and amplitude in magnitudes are presented with their respective errors, $\sigma_P(d)$ and $\sigma_A(mag)$. The $V$ is in magnitudes and number of cycles in logarithmic scale.

| Name                  | $P$    | $\sigma_P$ | $V$    | $A$    | $\sigma_A$ | $\log(T_P)$ | $\Omega_1$ | $\Omega_2$ | $\Omega_{tot}$ | $\Omega_i$ | $\Delta T$ (day) | $d$ (days) | $\sigma_P(d)$ | $\sigma_A(mag)$ |
|-----------------------|--------|------------|--------|--------|------------|--------------|------------|------------|----------------|------------|-----------------|------------|---------------|----------------|
| CSS J041816.3-362427  | 1.2694 | 0.0047     | 15.17  | 0.79   | 0.04        | 3.37         |            |            |                |            |                 |            |               |                |
| CSS J045516.1-040733  | 3.7448 | 0.0703     | 15.30  | 1.18   | 0.02        | 3.29         |            |            |                |            |                 |            |               |                |
| CSS J083114.1+041938  | 3.1869 | 0.0077     | 14.18  | 0.07   | 0.01        | 3.95         |            |            |                |            |                 |            |               |                |
| CSS J100837.0-073021  | 1.8364 | 0.0100     | 14.53  | 0.65   | 0.30        | 2.03         |            |            |                |            |                 |            |               |                |
| CSS J045508.6-030953  | 1.8227 | 0.0170     | 12.16  | 1.52   | 0.05        | 2.73         |            |            |                |            |                 |            |               |                |
| CSS J060321.6-040517  | 1.3587 | 0.0471     | 16.00  | 0.79   | 0.02        | 3.29         |            |            |                |            |                 |            |               |                |
| CSS J085748.6+034448  | 3.2070 | 0.0250     | 15.48  | 0.37   | 0.03        | 2.95         |            |            |                |            |                 |            |               |                |
| CSS J163922.0+330106  | 1.1870 | 0.0474     | 14.78  | 0.71   | 0.01        | 3.29         |            |            |                |            |                 |            |               |                |
| CSS J060449.8-072110  | 1.3208 | 0.0628     | 14.47  | 0.78   | 0.01        | 3.29         |            |            |                |            |                 |            |               |                |
| CSS J035318.6+043558  | 0.3031 | 0.0036     | 13.35  | 0.75   | 0.02        | 3.45         |            |            |                |            |                 |            |               |                |
| CSS J084552.6-044118  | 0.3419 | 0.0074     | 15.70  | 1.34   | 0.03        | 2.98         |            |            |                |            |                 |            |               |                |

---

### Table B1. Parameters obtained from the model fit to the light curves. The first one presents the name of the object, followed by the CSS ID, temperature ratio $T_i/T_{tot}$, sum of fractional radii ($R_i/R_{tot}$), inclination $i$, and modified Kopal potential (dimensionless) $\Omega_1$ and $\Omega_2$.

| Name                  | CSS ID | $\Omega_1$ (scale) | $\Omega_2$ | $\Omega_{tot}$ | $\Omega_i$ | $\Delta T$ (day) | $d$ (days) | $\sigma_P(d)$ | $\sigma_A(mag)$ |
|-----------------------|--------|-------------------|------------|----------------|------------|-----------------|------------|---------------|----------------|
| CSS J03441.8-315035   | 1012004004434 | 3896 | 0.97 ± 0.01 | 0.15 ± 0.02 | 88.88 ± 0.29 | 11.73 ± 1.45 | 14.31 ± 0.83 |
| CSS J145100.7+052843  | 11040004468 | 3954 | 0.96 ± 0.01 | 0.06 ± 0.01 | 89.67 ± 0.13 | 32.95 ± 1.77 | 31.85 ± 3.18 |
| CSS J165249.4+100124  | 1109007062934 | 3828 | 0.95 ± 0.01 | 0.14 ± 0.02 | 89.83 ± 0.10 | 32.16 ± 3.56 | 31.70 ± 3.99 |
| CSS J205021.5+234112  | 11012011464 | 3819 | 1.00 ± 0.01 | 0.15 ± 0.01 | 89.58 ± 0.24 | 14.45 ± 0.62 | 14.44 ± 0.67 |
| CSS J084835.7+233917  | 110200400616 | 4510 | 0.87 ± 0.01 | 0.14 ± 0.02 | 87.81 ± 0.55 | 16.22 ± 2.34 | 14.28 ± 0.58 |
| CSS J073157.2+241238  | 11020010067 | 3668 | 0.95 ± 0.01 | 0.09 ± 0.01 | 89.77 ± 0.17 | 22.32 ± 1.70 | 21.66 ± 1.27 |
| CSS J080549.6+4030108 | 110400302771 | 3512 | 0.98 ± 0.01 | 0.20 ± 0.02 | 89.42 ± 0.20 | 11.05 ± 0.83 | 10.74 ± 0.64 |
| CSS J093555.4+533132  | 1102001095400 | 3962 | 1.00 ± 0.01 | 0.35 ± 0.02 | 88.72 ± 0.78 | 6.72 ± 0.45 | 6.73 ± 0.44 |

---

The temperatures and masses of the components were defined as described in the Section 5.1 and the model fit to the light curve was redone in order to improve parameters such as $\Omega_1$, $\Omega_2$, and $i$. The last column is the estimated distance from the fitted parameters.