Discovery of 36 eclipsing EL CVn binaries found by the Palomar Transient Factory

J. van Roestel,[1][*] T. Kupfer,[2] R. Ruiz-Carmona,[1] P. J. Groot,[1] T. A. Prince,[2] K. Burdge,[2] R. Laher,[3] D. L. Shupe[4] and E. Bellm[5]

[1]*Department of Astrophysics/IMAPP, Radboud University Nijmegen, PO Box 9010, 6500 GL, Nijmegen, the Netherlands*
[2]*Cahill Center for Astronomy and Astrophysics, California Institute of Technology, Pasadena, CA 91125, USA*
[3]*Spitzer Science Center, California Institute of Technology, Pasadena, CA 91125, USA*
[4]*Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA*
[5]*Department of Astronomy, University of Washington, Box 351580, Seattle, WA 98195, USA*

ABSTRACT

We report on the discovery and analysis of 36 new eclipsing EL CVn-type binaries, consisting of a core helium-composition pre-white dwarf (pre-He-WD) and an early-type main-sequence companion. This more than doubles the known population of these systems. We have used supervised machine learning methods to search 0.8 million light curves from the Palomar Transient Factory (PTF), combined with Sloan Digital Sky Survey (SDSS), Panoramic Survey Telescope and Rapid Response System (Pan-STARRS) and Two-Micron All-Sky Survey (2MASS) colours. The new systems range in orbital periods from 0.46 to 3.8 d and in apparent brightness from $\sim$14 to 16 mag in the PTF R or g' filters. For 12 of the systems, we obtained radial velocity curves with the Intermediate Dispersion Spectrograph at the Isaac Newton Telescope. We modelled the light curves, radial velocity curves and spectral energy distributions to determine the system parameters. The radii (0.3–0.7 $R_\odot$) and effective temperatures (8000–17 000 K) of the pre-He-WDs are consistent with stellar evolution models, but the masses (0.12–0.28 $M_\odot$) show more variance than models have predicted. This study shows that using machine learning techniques on large synoptic survey data is a powerful way to discover substantial samples of binary systems in short-lived evolutionary stages.

Key words: binaries: close – binaries: eclipsing – stars: individual: EL CVn – white dwarfs.

1 INTRODUCTION

EL CVn binaries are eclipsing binaries containing a low-mass ($\sim$0.15–0.33 $M_\odot$) pre-helium white dwarf (pre-He-WD) and an A/F-type main-sequence (MS) star. The prototype system, EL CVn, is part of a sample of 17 EL CVn systems (Maxted et al. 2014a) discovered by the Super Wide Angle Search for Planets (SWASP; Pollacco et al. 2006) with magnitudes in the range of $9 < V < 13$. All light curves show boxy, shallow eclipses ($\lesssim$0.1 mag depth) with periods between $\sim$0.5 and $\sim$3 d, and in most cases ellipsoidal variation due to the deformation of the A/F star. The low radial velocity amplitudes ($\sim$15–30 km s^{-1}) of the primaries confirm the low-mass nature of the pre-He-WDs.

A total of 10 EL CVn systems were found in the Kepler survey: KOI-74 (van Kerkwijk et al. 2010; Bloemen et al. 2012); KOI-81 (van Kerkwijk et al. 2010; Matson et al. 2015); KOI-1375 (Carter,

Rappaport & Fabrycky 2011); KOI-1224 (Breton et al. 2012); KIC-9164561, KIC-10727668 (Rappaport et al. 2015); KIC-4169521, KOI-3818, KIC-2851474 and KIC-9285587 (Faigler et al. 2015). All these systems were studied in great detail, and by modelling the Kepler light curves in combination with radial velocity curves, all system parameters have been determined. Four of these systems contain small pre-He-WDs ($<$0.05 $R_\odot$) and, as a consequence, their light curves feature shallow eclipses only detectable from space. The fact that 10 EL CVn-like systems are found in the Kepler field suggests that there should be many more in our Galaxy, in line with an estimate of the local space density from stellar evolution and population synthesis models, 4–10×10^{-6} pc^{-3} (Chen et al. 2017).

Besides the samples found by Kepler and SWASP, there have been serendipitous discoveries of binaries related to EL CVn systems. The star V209 in ω Cen is likely an EL CVn binary (Kaluzny et al. 2007), but the primary does not seem to be a typical main-sequence star: its mass is 0.95 $M_\odot$ but it has a temperature of 9370 K. OGLE–BLG–RRLYR–02792 is an eclipsing binary that

[*] E-mail: j.vanroestel@astro.ru.nl

contains a large pre-He-WD, which seems to be pulsating like an RR-Lyrae star (Pietrzyński et al. 2012). A possible non-eclipsing variant of an EL CVn binary is the star Regulus (α Leo). Gies et al. (2008) and Rappaport, Podsiadlowski & Horev (2009) found that Regulus A is a single-lined spectroscopic binary with a period of 40 d, consisting of an A-type primary and a companion with a mass of $>0.3\,M_{\odot}$, at the upper end of the pre-He-WD mass range.

EL CVn binaries share many characteristics with a new type of binary: R CMa-type binaries are Algol binaries with a bloated, low-mass, donor (e.g. Budding & Butland 2011; Lee et al. 2016). They are very similar to EL CVn systems, except that they are semidetached, and therefore considered the progenitors of EL CVn systems. Two detached R CMa systems have been identified using Kepler photometry and these are now considered to be newly born EL CVn binaries: KIC-10661783 (Lehmann et al. 2013) and KIC-8262223 (Guo et al. 2017).

EL CVn systems are part of a larger family of binaries where one component of the binary is an extremely low-mass white dwarf (ELMWD). The majority of ELMWD-containing binaries without a main-sequence companion have white dwarf or neutron star companions instead (e.g. Marsh, Dhillon & Duck 1995; van Kerkwijk et al. 2005). In these systems, the ELMWD dominates the luminosity, making them identifiable with a single spectrum. The ELM survey (Brown et al. 2010) uses this approach and has been successful in finding many ELMWDs in binary white dwarf systems.

In this paper, we present system parameters for 36 new EL CVn systems, all eclipsing, discovered using the Palomar Transient Factory (PTF). In Section 2, we describe the identification of the systems using supervised machine learning classifiers. In Section 3, we discuss the spectroscopic follow-up of 12 of the new systems. In Section 4, we discuss the analysis of the light curves, spectra and spectral energy distributions (SEDs), and we present the results in Section 5. In Section 6, we compare our results with theoretical prediction and we compare our sample with already known EL CVn binaries. We end with a summary and conclusion in Section 7.

2 TARGET SELECTION

2.1 The Palomar Transient Factory

The Palomar Transient Facility (PTF) used the 1.2-m Oschin Telescope at the Palomar Observatory with a mosaic camera consisting of 11 CCDs. The CCDs have 4×2 K pixels and the camera has a pixel scale of 1.02 arcsec pixel^{-1}, giving it a total field of view of 7.26 deg^2. The PTF uses an automated image processing pipeline, which performs bias and flat-field corrections, source finding and photometry. All data are automatically processed; see Rau et al. (2009) and Law et al. (2009) for further information.

2.2 Data

For all objects detected by the PTF, light curves are automatically generated (see Laher et al. 2014) and light-curve statistics are calculated. These statistics include, among others, the mean, root mean square (rms), percentiles and χ^2-statistic; see Masci & Bellm (2016) for a full list. These light-curve statistics are based on the light-curve features used in Richards et al. (2011, 2012), which are useful to distinguish different types of variable stars. It is important to note that we do not use features related to any periodicity in the light curve. This is for a practical reason; it is very difficult to automatically obtain a reliable period for all the PTF light curves because they are sparsely sampled and span many years.

Table 1. The number of objects after our initial selection with the PTF light curves (>40 epochs, $\chi^2_{\rm reduced} > 10$, <16 mag). The percentages for which additional colour information is available are shown.

Filter	No. of objects	SDSS *ugriz*	NOMAD *JHK*	Pan-STARRS *grizy*
R	532 477	43.65%	97.58%	98.92%
g'	257 918	55.45%	98.69%	99.26%
$R \cap g'$	36 943	64.39%	96.48%	98.66%

For this study, we used all available light-curve data obtained between the start of the PTF in 2008 December and 2016 March. We treat the data for the R and g' filters as two separate data sets in the subsequent analysis. These data sets are very substantial: R, $\sim$250 million objects; g', $\sim$50 million objects. We make an initial cut and select only objects that are variable by requiring that $\chi^2_{\rm reduced} > 10$, that light curves have more than 40 epochs, and that objects are brighter than 16 mag in either PTF R or PTF g'. This still leaves more than $\sim 10^5$ candidates (see Table 1 and Fig. 1).

We match the objects in these data sets to the latest Sloan Digital Sky Survey (SDSS) catalogue Data Release 13 (DR13, *ugriz* bands; Albareti et al. 2016), the Naval Observatory Merged Astrometric Dataset (NOMAD, *JHK* bands; Zacharias et al. 2004) and the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS) catalogue (*grizy* bands; Chambers et al. 2016). Table 1 gives an overview of the total number of objects and the colour availability.

2.3 Machine learning classification

To cut back on the number of candidates for an initial visual light-curve inspection, we use supervised machine learning classifiers to make a pre-selection. The idea is that instead of finding EL CVn binaries by using fixed, pre-defined, user-supplied selection criteria, a sample of known EL CVn binaries and objects that are not EL CVn binaries (i.e. a training set) is provided and a machine learning code (i.e. classifier) decides what is the best way to separate the two groups given the characteristics (called features; e.g. $g - r$ colour or the light curve's rms value). There are many different types of classifiers, and the behaviour of each classifier can be adjusted by changing so-called hyperparameters. Setting the correct hyperparameters is required to avoid overfitting or underfitting of the data. For an introduction to machine learning in astronomy, see Ivezić et al. (2014); for a practical guide to machine learning (with PYTHON), see Andreas C. Müller (2016).

Because supervised machine learning classifiers can process huge amounts of data very quickly, they have become a popular tool to handle the large amount of light curves produced by survey telescopes. Many different techniques have been tried for light-curve classification (e.g. Debosscher 2009; Palaversa et al. 2013; Angeloni et al. 2014; Peters et al. 2015; Mackenzie, Pichara & Protopapas 2016; Armstrong et al. 2016; Sesar et al. 2017). In recent years, the Random Forest method (Breiman 2001) has become very popular as it typically performs the best and is also easy to interpret (e.g. Richards et al. 2011; Masci et al. 2014).

To find EL CVn binaries, we have experimented with three different supervised machine learning classifiers based on combining decision trees: the standard Random Forest and an 'Extra-Trees' classifier (Geurts, Ernst & Wehenkel 2006), both implemented in the PYTHON package SKLEARN (Pedregosa et al. 2011), and the more sophisticated 'Gradient boosted decision tree' classifier, implemented
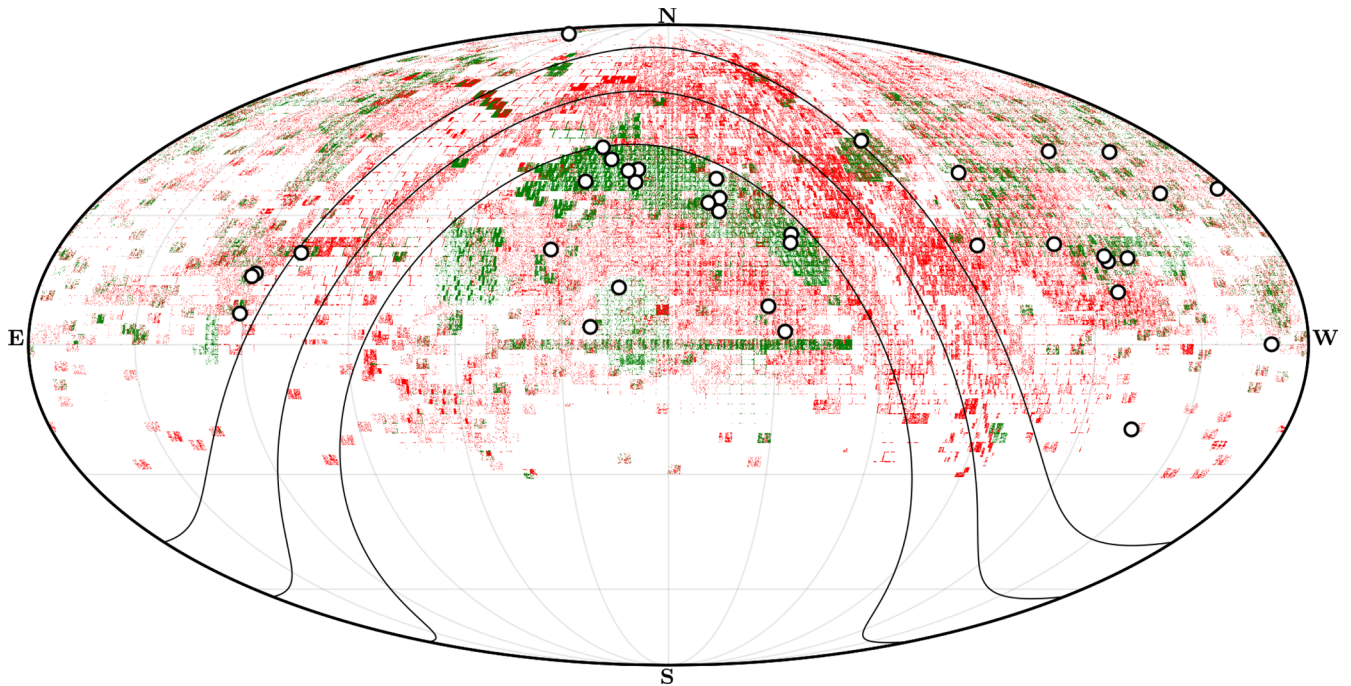
Figure 1. All objects in the PTF sample after our initial cut (>40 epochs, $\chi^2_{\mathrm{reduced}} > 10$, <16 mag), with red denoting PTF R and green denoting PTF g'. The EL CVn binaries we discovered in the data are shown as white dots. The black lines show Galactic latitudes of $-15°$, $0°$ and $15°$.

in XGBOOST (Chen & Guestrin 2016). All three classifiers combine many randomized decision trees, which are a sequence of binary decisions.

Here we briefly discuss the differences between the methods. Both Random Forest and Extra-Trees combine the prediction of many independent, randomized decision trees. The larger the number of trees the better, but at the cost of increased computation time. For both methods, each tree is built using only a subset of all features (the rule of thumb is the square root of the total number of features). Random Forest uses the best possible split of the data given the available features and uses that to separate the different classes. Extra-Trees differs from Random Forest as it does not use the best split, but a random split. This extra randomization step has the consequence that decision boundaries are smoother compared to Random Forest. Both methods are relatively simple; they have only a few hyperparameters and are relatively robust against overfitting. XGBOOST also uses many randomized decision trees. However, instead of combining many independent trees, new trees are created to optimally complement the existing trees. This is done by giving samples that were wrongly classified by the previous trees a larger weight when building the next tree. The next tree is therefore more likely to classify these objects correctly. The disadvantage of this method is that it is more sensitive to overfitting compared to Random Forest. The XGBOOST implementation has many hyperparameters, which can be set to counteract this, but it can be difficult to determine the best values for these parameters. The advantages of all three methods are that they are insensitive to uninformative features, do not require scaling of the data and are easy to interpret: they automatically determine the importance of features.

2.4 EL CVn identification

Because supervised machine learning algorithms require a training sample, we first need to identify EL CVn binaries in our data. There are no known EL CVn binaries in the PTF magnitude range, so we

need to find new ones in an old-fashioned way. We do this by selecting a sample of A-type main-sequence stars using SDSS colours ($0.8 < u - g_{\mathrm{SDSS}} < 1.5$ and $-0.5 < g - r_{\mathrm{SDSS}} < 0.2$) and we require that Stetson-K > 0.6 (one of the light-curve statistics; see also Stetson 1996). To limit the sample size and to increase post-facto confidence in the selected objects, we further require that the light curve is significantly variable ($\chi^2 > 40$) and with more than 150 epochs in R and 100 in g'. EL CVn binaries are characterized by their $\lesssim 0.1$ mag, flat-bottomed primary eclipse and slightly shallower secondary eclipse. Therefore, we carry out a period search using both Analysis-of-Variance (AoV; Schwarzenberg-Czerny 1989; Devor 2005) and Boxed-Least-Square (BLS; Kovács, Zucker & Mazeh 2002, VARTOOLS implementation, Hartman & Bakos 2016) methods on each of the light curves and we inspect each folded light curve for these criteria. If in doubt, the candidate was included in the light-curve modelling (see Section 4.1). If the light-curve fitting showed a V-shaped, non-total, eclipse, we rejected it from the sample, as these systems could also be regular MS–MS binaries. In other words, we require our systems to be totally eclipsing.

Using this method, we found six EL CVn binaries, which we then used as a training set for a Random Forest classifier, combined with a sample of 4000 randomly chosen objects (which we confirmed were not EL CVn binaries). Because the training set is so small, we do not attempt any parameter optimization, but we use the default hyperparameters (500 trees, the square root of the total number of features as the number of features per tree, and no limits on the tree depth). We applied the classifier to the data (the PTF R and PTF g' light-curve statistics combined with SDSS colours) and we inspected the ∼100 best candidates identified by the classifiers. We added newly found EL CVn candidates to the training sample and repeated the procedure an additional two times. This resulted in the discovery of an additional 11 systems, bringing the total to 17.

Because we required that SDSS colours were available, we only inspected roughly half of the data at this point (see Table 1). Therefore, we replaced the SDSS colours with the *BVRHJK* colours from
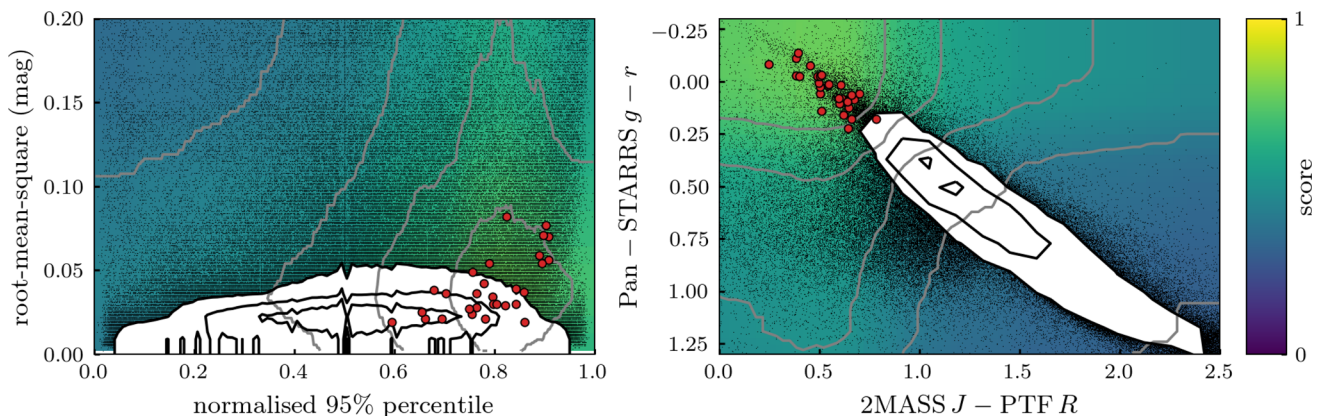
Figure 2. The left panel shows the weighted rms of the PTF light curve versus the normalized 95th percentile of the light curve (percentile 95 minus median, divided by 90 percentile range; see Table A1). The right panel shows 2MASS J-PTF R versus Pan-STARRS $g - r$ colour–colour space. The red dots show all EL CVn binaries and the black contours show all samples in the PTF R data set, with the black contours containing 25, 50 and 75 per cent of the data. Samples outside the contours are indicated with black points. The background colour indicates the EL CVn score by the Extra-Trees classifier with grey lines at every 0.1 score interval. The score is calculated by assuming the median values of the EL CVns for all parameters, except the parameters on the *x*- and *y*-axes.

the NOMAD catalogue (Pan-STARRS colours were not yet available at this time). We again checked the best 100 candidates in an iterative way, adding the new EL CVn systems to the training sample. The combined SDSS and NOMAD process resulted in a total of 30 EL CVn binaries.

With this sample, we trained the three different classifiers (Random Forest, Extra-Trees and XGBOOST) and determined the best hyperparameters. We used the PTF variability features combined with the Pan-STARRS colours. The goal of our classifier is not to classify all samples correctly (high precision), but instead to rank the candidates according to 'EL CVn' likeness. Therefore, we do not optimize the precision of our classifier, but instead we use the area-under-curve (AUC) for the receiver operating characteristic (ROC). We do this using stratified K-fold cross-validation to calculate the ROC-AUC score. For more details on classifier metrics (such as ROC-AUC) and model optimization, see Ivezić et al. (2014); Andreas C. Müller (2016).

For both Random Forest and Extra-Trees, we find similar optimal hyperparameters. Using more than 600 trees does not improve performance significantly. The number of features per tree influences the ROC-AUC score marginally, but there is a clear preference for only using two features per tree. We checked different hyperparameters that limit the depth or complexity of the tree, but we find that the ROC-AUC score only decreases when the tree depth or complexity is limited using any of the hyperparameters. As an example, Fig. 2 shows how the score of the Extra-Trees classifier varies in two cross sections of the parameters space.

For XGBOOST there are more hyperparameters to tune. We start by optimizing the most important three (i.e. the number of estimators, the learning rate and the tree depth) while setting the other parameters to typical values. After finding the optimal combination of these hyperparameters, we continue to optimize the minimum child weight, subsampling fraction and the column subsample fraction.

After training all classifiers, we selected the top 1000 candidates (in both R and g' data sets) from the three classifiers and visually inspected their light curves. We found an additional six EL CVn binaries, bringing the final number to 36 systems, as listed in Table 2.

A quick comparison between the classifiers shows that both the Extra-Trees classifier and the XGBOOST classifier perform equally well while the Random Forest performs slightly worse. This is confirmed by the ranking of the last six discovered EL CVn

Table 2. Overview of the EL CVn binaries we discovered in the PTF data. In the rest of the paper, we use the PTF name. The PTF R column lists the median magnitude of the light curve in the R band.

PTF name	IAU name (PTF1 J...)	P (d)	PTF R
PTFS1600y	J004040.23+412521.61	1.184	13.7
PTFS1600ad	J004300.75+381537.26	1.084	14.7
PTFS1700do	J005424.06+411126.98	3.051	15.7
PTFS1600aa	J005659.72+130920.66	0.693	15.9
PTFS1601p	J011909.91+435907.11	1.222	15.3
PTFS1501bh	J012814.72+040551.90	0.620	13.9
PTFS1601q	J013336.92+470600.18	1.252	16.2
PTFS1601cl	J014839.10+382314.56	0.892	13.6
PTFS1402de	J021913.15+215921.98	0.619	15.0
PTFS1607aa	J071207.01+211654.98	0.846	15.0
PTFS1607v	J075310.42+835154.79	0.720	15.3
PTFS1607t	J075642.49+162143.99	0.876	14.2
PTFS1607ab	J075950.03+154319.09	0.773	14.0
PTFS1608ab	J080425.26+070845.24	0.610	14.6
PTFS1612al	J121254.27+363341.76	0.637	15.7
PTFS1512bf	J124154.58+001333.06	0.607	14.2
PTFS1613s	J133220.59+352847.28	1.142	14.3
PTFS1613u	J133929.37+455055.64	0.564	15.3
PTFS1615ag	J150041.84−191417.23	0.681	14.3
PTFS1615v	J150327.61+460322.78	0.559	15.9
PTFS1515ay	J150336.10+195842.16	0.464	14.8
PTFS1615w	J152726.81+120453.54	1.441	14.9
PTFS1615ao	J152758.90+190751.63	0.895	15.0
PTFS1615u	J153005.01+202157.06	0.778	15.8
PTFS1616cr	J162342.13+231456.58	0.565	14.0
PTFS1617n	J173257.98+403600.93	2.337	15.3
PTFS1617m	J175433.50+230041.83	3.773	14.7
PTFS1619l	J191826.08+485302.94	1.160	13.7
PTFS1521ct	J213318.98+254126.30	1.172	15.8
PTFS1621ax	J213534.11+233313.86	1.018	15.0
PTFS1521cm	J214858.33+030417.50	0.685	15.1
PTFS1622by	J220719.56+085415.66	0.749	15.8
PTFS1522cc	J225539.41+342137.72	0.572	14.7
PTFS1622aa	J225652.53+390822.70	0.766	15.6
PTFS1622bt	J225755.64+310133.67	0.688	15.1
PTFS1723aj	J231010.08+331249.78	1.109	14.8

binaries that were all further down the list for the Random Forest method. Although the performances of Extra-Trees and XGBOOST are comparable, tuning the XGBOOST classifier took significantly more time and effort. Because of the combination of yield versus investment, we deem the Extra-Trees classifier to be best (in this particular case).

The PTF observed the Kepler field and has thus observed the EL CVn binaries found by Kepler. None of these binaries was recovered by our search and we investigated the reason why. First of all, most Kepler systems feature eclipses much shallower than the PTF can detect. The Kepler EL CVn systems with deep enough eclipses to be detected by the PTF were not recovered because either the star was saturated in the PTF data, or the object was not observed at a sufficient number of epochs.

3 SPECTROSCOPY

For 19 of our EL CVn systems, we obtained phase-resolved spectroscopy with the Isaac Newton Telescope (INT). We used the Intermediate Dispersion Spectrograph (IDS) with the R632V grating (0.90 Å pixel^{-1}, 3800–5800 Å) for eight bright nights and the R900V grating for nine bright nights on three separate runs (0.63 Å pixel^{-1}, 4000–5500 Å). Conditions were good with seeing of $\lesssim$ 1 arcsec, except for the last four nights. During these nights, the seeing was 2–5 arcsec and two nights were mostly clouded. An overview of the spectroscopic runs, the set-up and the weather quality is given in Table A2.

Because the orbital period and phase for all systems is determined very precisely from the photometry (see Section 5), we timed the observations such that we observed the systems around orbital phases 0.25 and 0.75. The signal-to-noise per pixel of each spectrum ranges between 40 and 80, sufficient to detect the weaker metal lines in the A/F star's spectrum. Spectra were taken in pairs, and before or after each stellar spectrum a calibration lamp spectrum (CuAr) was obtained to make sure the wavelength calibration was stable.

The data were reduced using IRAF. We used L.A.COSMIC (van Dokkum 2001) to remove cosmic rays and we performed standard bias and flat calibrations. For the wavelength calibration, we used ∼40 arc lines, which resulted in a typical rms uncertainty on the wavelength solution of $\lesssim$0.1 pixels (4–6 km s^{-1}).

4 METHODS AND ANALYSIS

4.1 Light curves

By modelling the light curves, we put strong constraints on the system parameters. To construct a model light curve given a set of binary star parameters, we use LCURVE (by T. R. Marsh and collaborators; see Copperwheat et al. 2010 and also Copperwheat et al. 2011 and Parsons et al. 2011). The LCURVE code uses grids of points to model the two stars. The shape of the stars in the binary is set by a Roche potential. We assume that the orbit is circular and that the rotation periods of the stars are synchronized to the orbital period. We discuss the validity of the latter assumption in Section 6.1. We calculate the light curves by assuming the effective wavelength of the PTF filters: 4641 Å for the g' filter and 6516 Å for the R filter. In this section (and in the rest of the paper), we refer to the A/F-type main sequence as the primary (subscript '1') and the pre-He-WD as the secondary (subscript '2').

The free parameters of the model are the orbital period (P) and mid-eclipse time (t_0), both in BMJD$_{TDB}$ (the barycentric Julian date in the terrestrial dynamic time frame, minus 2400000.5), the

effective temperatures of both stars ($T_{1,2}$), the scaled radii of both stars ($r_{1,2} = R_{1,2}/a$, where a is the binary separation), the inclination angle (i), the mass ratio ($q = M_2/M_1$), an albedo (absorption) for both stars, a linear limb darkening coefficient ($x_{1,2}$) and a gravity darkening coefficient ($y_{1,2}$) in the relation $I \propto g^y$ (where g is the local surface gravity; von Zeipel 1924). Not all these parameters are well constrained by the data and therefore we fix or set an allowed range for some parameters. We constrain the temperature of the primary star (T_1) to 6500–10 000 K, which is the temperature range of A/F-type main-sequence stars. This is necessary because with only a light curve the temperature ratio is well constrained, but the absolute values of the temperatures of each star are not. We do not use the resulting temperatures of the light-curve fit, but instead determine the effective temperatures of both stars by modelling the SED (see Section 4.2). We fix the limb darkening coefficient of star 2 (x_2) to 0.5, as the effect on the light curve is minimal. We allow the limb darkening coefficient of the A/F star (x_1) to vary between 0.08 and 1.05, which are the lowest and highest values for stars in the allowed temperature range (Giannnas et al. 2013).

To determine the uncertainty on the parameters, we use the Markov chain Monte Carlo (MCMC) method as implemented by EMCEE (Foreman-Mackey et al. 2013). The standard method to determine the uncertainties on the parameters is to use the least-squares (χ^2) statistic. However, this assumes that the uncertainty estimates of the data are correct and Gaussian distributed. This is not the case for the PTF light curves (as in many observational data sets). Ignoring this problem leads to an underestimate of the uncertainties in the derived parameters, and can in some cases also change the solution. To solve this problem, we add additional white noise[1] to our model; see section 8 in Hogg, Bovy & Lang (2010) and see Foreman-Mackey (2013) for a simple example. This means that our model has the white noise amplitude as an extra parameter, which we can simply optimize over, exactly the same as for the light-curve parameters.

This method requires the following modification to the standard least-squares function:

$$\tilde{\chi}^2 = \sum_n \frac{[y_n - m_n(p)]^2}{\sigma_n^2 + f^2 m_n(p)^2} + \log\left[\sigma_n^2 + f^2 m_n(p)^2\right]. \quad (1)$$

Here, y_n is the data, m_n is the light-curve model as a function of the light-curve parameters p, σ_n are the uncertainties and f is a factor that adds an extra noise source. Note that the first term of the equation is almost the same as in a regular least-squares (χ^2) regression, except for the additional noise term, $f^2 m_n(p)^2$. The first term can be minimized by letting f go to infinity, instead of minimizing the difference between the data and the model, $y_n - m_n(p)$. Therefore, the second term is needed to penalize models with a large value of f. Using this equation, the optimal amount of white noise is added to account for any difference between the data and the model. To obtain the best model, we simply minimize $\tilde{\chi}^2$ over the light-curve parameters p and the parameter f, just like regular least-squares regression.

For each of the systems, we first find the approximate solution using a simple simplex minimizer of the modified least-squares function. We then use EMCEE to find the best set of parameters of all the available light curves for that system. For each filter,

---

[1] If the noise cannot be treated as white noise, but the noise is correlated (red noise), Gaussian process regression can be used. For a simple example, see Foreman-Mackey (2014) and an example of this method used to model flickering in a cataclysmic variable by McAllister et al. (2017).

we use different values for x_1, y_1 and 'absorb', while the rest of the parameters are filter independent. We use 256 parallel MCMC chains (called 'walkers') and we use at least 2000 generations or more if needed. Any further calculations are done using the last 2560 positions of the walkers.

4.2 Effective temperature

To determine the temperatures of both components, we fit the SED of the target with model spectra, similar to Maxted et al. (2011). We use data from the *Galaxy Evolution Explorer* (*GALEX*; far-ultraviolet and near-ultraviolet; Bianchi, Conti & Shiao 2014), SDSS DR13 (*ugriz*; Albareti et al. 2016), Pan-STARRS (*grizy*; Chambers et al. 2016), 2MASS (*HJK*; Skrutskie et al. 2006) and the *Wide-field Infrared Survey Explorer* (*WISE*; *W*1 and *W*2; Wright et al. 2010) for each target (where available). We used as model spectra the BaSeL3.1 spectral library (Westera et al. 2002). To calculate the flux per band, we convolve the model spectra with each band's response curve.

The overall spectrum is the sum of two model spectra of a given temperature and metallicity, created using bilinear interpolation from the BaSeL library. With only an SED, it is not possible to measure the metallicity of the stars reliably. However, as metallicity and temperature are correlated, we treat the metallicity of both stars as free parameters and marginalize over them in the final result. For the surface gravity, we assume $\log g = 4$ for the A/F star and $\log g = 5$ for the pre-He-WD. We set the relative contribution to the total light by the ratio between r_1 and r_2 obtained from the light curve. At first, we also used the temperature ratio obtained from the light curve, but we learnt that this gave inconsistent predictions for the eclipse depths. This is likely as a result of the use of blackbody spectra by LCURVE. Instead, we directly use the eclipse depth of the primary eclipse instead of the temperature ratio. The final variable is the extinction, set by $E(B - V)$. To calculate the reddening following from the extinction, we used the reddening law by Cardelli, Clayton & Mathis (1989) with $R_V = 3.1$ (as implemented by PYSYNPHOT).

To determine the temperatures of both stars we minimized the function

$$\tilde{\chi}^2 = \sum_n \frac{[y_n - m_n(p)]^2}{\sigma_n^2 + f^2 m_n(p)^2} + \log\left[\sigma_n^2 + f^2 m_n(p)^2\right]$$
$$+ \text{prior}[r_1/r_2, E(B - V)], \tag{2}$$

where y is the data, m is the model and f is an additional noise factor. We used a value for $E(B - V)$ according to Schlafly & Finkbeiner (2011), with an uncertainty of 0.034 (as in Maxted et al. 2011). For some added flexibility in our model, we added an extra term of uncertainty to the magnitudes (f), similar to the way it was applied in equation (1). We again use EMCEE to determine the best values and uncertainties, as in Section 4.1.

4.3 Radial velocity

To obtain the radial velocity curve of the primary star, we cannot use the Balmer absorption lines in the spectrum because these are present both in the A/F star and the pre-He-WD. Using these would not yield reliable results. Instead, we use the metal lines present in the spectra of the A/F-type stars. We cross-correlate the spectra with a template: a high-resolution spectrum of the A5 star HD 145689 (Bagnulo et al. 2003). We first interpolate the target spectrum to the (much higher) sampling of HD 145689. We then remove the continuum with a low-order polynomial and determine the radial

velocity shift using cross-correlation. To estimate the uncertainty on the radial velocity shift, we add random Gaussian noise to the target spectra according to the uncertainty per pixel and measure the radial velocity shift. We repeat this process 11 times and use the standard deviation of the results as the uncertainty. We use the metal lines in the ranges 4150–4301, 4411–4791 and 4941–5400 Å to obtain three separate measurements of the radial velocity shift. The radial velocity measurements are corrected to the heliocentric velocity frame with the RVCORRECT task in IRAF.

To determine the radial velocity amplitude, we fit a sinusoidal curve with a fixed value for the period and phase to the measurements. This leaves only the amplitude and systemic velocity as free parameters. We again use a modified least-squares objective function, which can also take into account underestimated uncertainties (similar to equation 1):

$$\tilde{\chi}^2 = \sum_n \frac{[y_n - m_n(p)]^2}{\sigma_n^2 + f^2} + \log\left(\sigma_n^2 + f^2\right). \tag{3}$$

Here, y_n are the radial velocity measurements, σ is the statistical uncertainty on the cross-correlation results and f is the extra white noise. Fitting the data shows (Table 3) that f ranges between 7 and 12 km s^{-1}, a factor of 2 higher than the statistical noise σ. This extra noise is partially a result of the uncertainties in the wavelength calibration (4–6 km s^{-1}) but does not account for all residual variance. This means either that we underestimate the uncertainties in the cross-correlation procedure (e.g. by the normalization of the spectra) or that we underestimate the uncertainties in the calibration process of the spectra. This could be a result of the instabilities of the optical elements in the IDS/INT combination, which typically changes for each observing run. As we combine data from four different observing runs, this could result in minor differences in the set-up. A potential method to verify this is to check the wavelength of sky emission lines, but these are not available in the spectral range we use.

4.4 Galactic kinematics

For the 12 systems for which we have obtained a radial velocity measurement, we calculate their Galactic location and velocity. We determine the distance to the systems by using the *K*-band magnitude and absolute radius, combined with the *K*-band surface brightness calibration by Kervella & Fouqué (2008). The proper motions of the systems are taken either from the USNO CCD Astrograph Catalog Release 5 (UCAC5; Zacharias et al. 2017) or from the Initial *Gaia* Source List (Smart & Nicastro 2013). Combined with right ascension and declination, we calculate velocity in the direction of the Galactic Centre (V_ρ) and the Galactic rotation direction (V_ϕ), the Galactic orbital eccentricity (e) and the angular momentum in the Galactic *z*-direction (J_z). The Galactic radial velocity V_ρ is negative towards the Galactic Centre, while stars that are revolving on retrograde orbits around the Galactic Centre have negative V_ϕ. Stars on retrograde orbits have positive J_z. Thin disc stars generally have very low eccentricities e. Population membership can be derived from the position in the V_ρ–V_ϕ diagram and the J_z–e diagram (Pauli et al. 2006).

4.5 Masses and radii

To fully solve for the elements of the binary system, we need to combine the information from the light-curve fit with an additional piece of information to set the scale of the system. This is typically
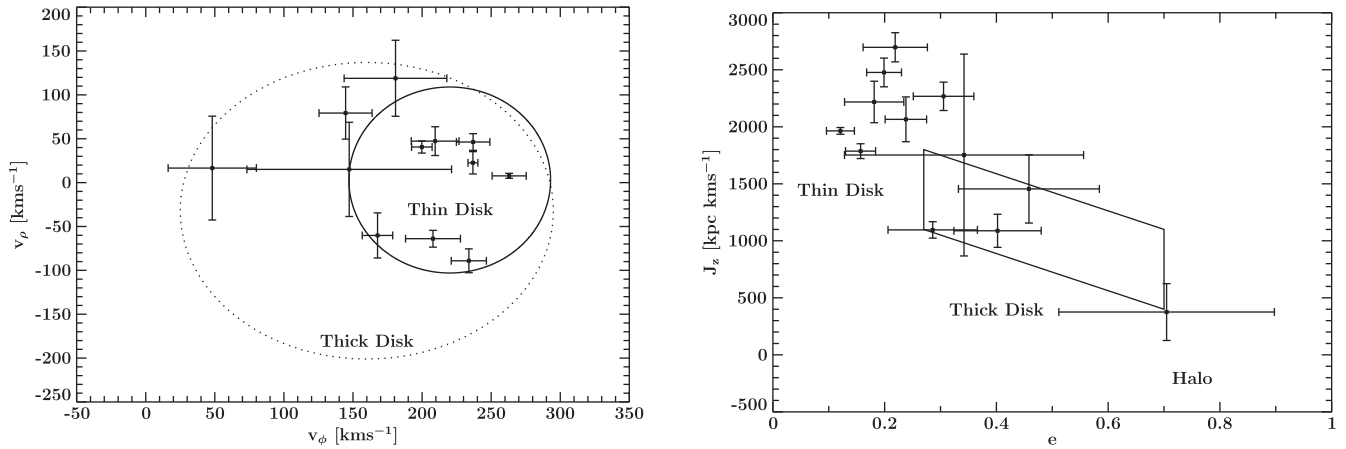
Figure 3. $V_\phi - V_\rho$ (left) and e–J_z (right) diagrams. The solid and dotted ellipses render the 3σ thin and thick disc contours in the $V_\phi - V_\rho$ diagram, while the solid box in the e–J_z marks the thick disc region as specified by Pauli et al. (2006).

Table 3. Radial velocity amplitude of the A/F star, the systemic velocity, the residual variance of the fit, the derived distance, the measured proper motions and the associated stellar population for the 12 EL CVn systems with radial velocity curves.

ID	K_1 (km s^{-1})	γ (km s^{-1})	f (km s^{-1})	Distance (pc)	$\mu_\alpha \cos(\delta)$ (mas yr^{-1})	μ_δ (mas yr^{-1})	Pop.
PTFS1600y	22.8 ± 0.9	-88.7 ± 0.9	6.9	2340 ± 70	-7.3 ± 2.0^a	-4.2 ± 1.3^a	Thin
PTFS1600ad	29.7 ± 1.4	-23.3 ± 1.2	7.2	3770 ± 180	1.1 ± 1.5^a	-1.8 ± 1.5^a	Thin
PTFS1601p	18.4 ± 2.1	-45.9 ± 1.8	12.4	4960 ± 500	1.4 ± 4.6^a	-7.6 ± 4.2^a	Thin/thick
PTFS1501bh	24.0 ± 1.5	16.6 ± 1.2	8.2	1280 ± 70	10.1 ± 1.5^a	-4.9 ± 1.5^a	Thin
PTFS1601cl	35.2 ± 2.1	-14.4 ± 1.2	9.1	2890 ± 90	2.0 ± 1.3^a	2.2 ± 1.3^a	Thin
PTFS1607t	26.7 ± 1.7	31.0 ± 1.2	5.5	2160 ± 70	-2.1 ± 0.3^a	1.6 ± 1.6^a	Thin
PTFS1607ab	32.7 ± 1.3	-37.6 ± 1.1	6.3	1810 ± 70	-3.1 ± 2.6^b	-8.2 ± 3.1^b	Thin
PTFS1512bf	31.4 ± 1.9	70.3 ± 1.5	11.8	1820 ± 50	-19.2 ± 6.4^a	4.7 ± 5.7^a	Thick
PTFS1617n	17.7 ± 1.7	-198.7 ± 1.9	7.2	5700 ± 380	-1.8 ± 2.6^a	-2.9 ± 2.7^a	Thick/halo
PTFS1617m	13.1 ± 2.4	-40.4 ± 2.1	10.9	4060 ± 180	-0.9 ± 1.5^a	-9.2 ± 1.7^a	Thin/thick
PTFS1619l	22.7 ± 1.3	-12.5 ± 1.0	5.8	2040 ± 140	-1.9 ± 1.6^a	-2.0 ± 1.6^a	Thin
PTFS1521cm	34.7 ± 2.0	-6.5 ± 1.9	7.9	2870 ± 110	9.6 ± 2.8^a	-7.8 ± 2.7^a	Thick

^{a}The proper motion is taken from the UCAC5 catalogue (Zacharias, Finch & Frouard 2017).
^{b}The proper motion is taken from the Initial *Gaia* Source List (Smart & Nicastro 2013).

done by measuring the radial velocity amplitude of both stars. We only have the radial velocity amplitude of one of the stars in the binary. In principle, we can combine this with the mass ratio q, but the uncertainties on the mass ratio derived from the light-curve fitting are high and the uncertainties on the masses scale with a high power of q (for low q, $M_1 \propto K_1^2 q^{-3}$ and $M_2 \propto K_1^3 q^{-2}$), and therefore they are not constraining.

To circumvent this problem, we make use of the assumption that the primary star is a main-sequence star. Using only the light-curve parameters, we can calculate the average density of the main-sequence component:

$$\langle \rho \rangle = \frac{3\pi}{G P^2 r_1^3 (1+q)}. \tag{4}$$

To propagate the uncertainties correctly, we calculate the stellar density for each point in the MCMC chain and we assign a random temperature according to our measurement of the SED. With the average density and temperature of the main-sequence star, we can use stellar models to determine its mass. We use the Yale–Potsdam stellar models (Spada et al. 2017) and follow the same procedure as in Breton et al. (2012) to make a continuous mapping of the mass in T–$\langle \rho \rangle$ space. We convolve each track with a Gaussian probability function with a standard deviation of 200 K in temperature and 0.1

dex in density. For each point in the temperature–density grid, we assign the mass with the highest probability. We use this mapping to calculate the primary mass for the posterior distribution of the light-curve fits (see Section 4.1). As can be seen in Fig. 4, most but not all measurements agree with the models. Two systems, PTFS1612al and PTFS1615u, have slightly higher densities than would be the case for a solar metallicity composition for zero-age main-sequence models. For these two systems, we extrapolate the models to determine the mass.

With the mass of the primary (M_1) combined with q, i and P, we calculate the semimajor axis (a) using Kepler's law:

$$a^3 = GM_1(1+q)\left(\frac{2\pi}{P}\right)^2. \tag{5}$$

Note that in both equations the mass ratio is present in the form of $1 + q$, and because the mass ratio is small ($q \sim 0.1$), the high uncertainty on q only mildly affects the accuracy on a and M_1. However, the uncertainty on the pre-He-WD mass ($M_2 = qM_1$) is proportional to the uncertainty on q, which means that the uncertainty on M_2 is too high to be constraining.

This can be solved by including the measured radial velocity (K_1) in our calculation, which is available for 12 systems. We use an iterative approach to find the optimal solution as in Rappaport
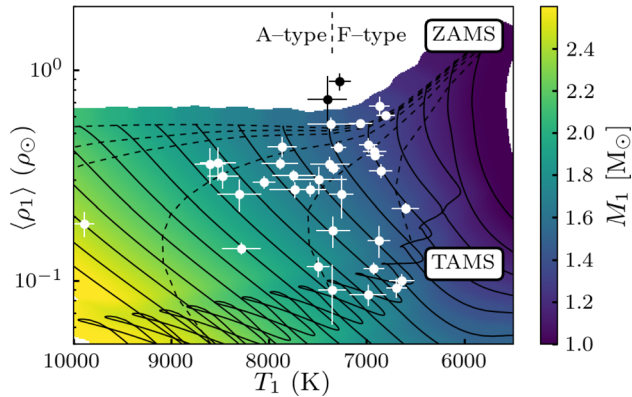
Figure 4. The temperature versus the average density of the main-sequence stars of the binary system, indicated with a black or white dot. The black lines show main-sequence evolution tracks (solar metallicity) from Spada et al. (2017) between 1.0 and 2.5 $M_\odot$ with 0.1 $M_\odot$ intervals. The colour map shows the mass of the star according to the interpolation method by Breton et al. (2012). The dashed lines are isochrones of 0, 0.1, 0.2, 0.5, 1 and 2 Gyr since the start of the main sequence.

et al. (2015), again for each sample from the light-curve fit posterior distribution. This involves calculating M_1, q and $\langle \rho \rangle$ until the solution converges, which it does after two iterations.

5 RESULTS

For the 36 EL CVn binaries in the PTF data, we fit the light curves with a binary star model (see Fig. A1 and Table A4 in the Appendix). The best model fits to the light curves all show a flat-bottomed primary eclipse and a round-bottom secondary eclipse. The orbital period of the binary, the radii of both stars and the orbital inclination are typically well constrained, but the uncertainty on the mass ratios of the systems is typically $\gtrsim 10$ per cent. The extra noise term in the fit for the light curves is typically $\lesssim 1$ per cent. This is consistent with the expected uncertainty in the absolute photometric calibration, which is not part of the error bars of the light curves. The orbital periods of the binaries range from 0.46 to 3.8 d, with inclinations between 74° and 90°. The radii of the primary stars divided by the semimajor axis (r_1) are typically 0.2–0.5, and the primary stars fill about 0.4–0.9 of their Roche lobe. The average density derived from the light curve is typically between 10 and 70 per cent of Solar density, consistent with A/F-type main-sequence stars. The mass ratios, as determined from the light curves, are typically between 0.08 and 0.2, but there are outliers to larger ratios. However, the uncertainties on the outliers are high. The mass ratio is determined from the amplitude of the inter-eclipse variability, which is in some cases not significant (e.g. 1700do) and explains the high uncertainty on the value for the mass ratio in some cases. From the light curve, we determined the temperature ratio of the two stars, assuming blackbody SEDs of typically 0.5–0.95.

The results of the SED fitting are shown in Table 4 and Fig. A2. The temperatures of the A/F stars in the EL CVn systems range between 6600 and 10 000 K, consistent with temperatures for A/F-type main-sequence stars (F-type, 6000–7350 K; A-type, 7350–10 000 K; Pecaut & Mamajek 2013). The temperatures of the pre-He-WDs range from 7900 to 17 000 K. In all systems, the pre-He-WDs are hotter than the A/F star companion. The uncertainty on the A/F star's temperature is typically 100–200 K. The temperature of the pre-He-WD is less well constrained (100–1400 K), because it depends on the availability of ultraviolet data and on how accu-

rately the eclipse depth can be measured from the light curves. The rms scatter between the data and model is typically $\lesssim 5$ per cent, with a few outliers to 10 per cent (see Table A3). This residual scatter might be the result not only of calibration differences between telescopes, but also of observations being taken at a random phase. An observation taken in-eclipse results in a ~ 10 per cent lower flux than out-of-eclipse.

For 12 of the EL CVn systems, we obtain usable radial velocity curves and measure the radial velocity amplitude (see Fig. A3). For the remaining seven systems, there were not enough measurements obtained or these were observed at unfavourable orbital phases, precluding an accurate radial velocity measurement. However, all radial velocity amplitudes are low, in the range of 20–40 km s^{-1}. This confirms that the secondary stars in these binaries are indeed low-mass stars.

Using all information available, we determine the stellar parameters of the stars in the EL CVn systems, summarized in Table 4. The masses of the A/F-type stars range between 1.3 and 2.4 $M_\odot$. The radii of these stars (1.15–2.7 $R_\odot$) are consistent with these stars being regular main-sequence stars.

The radii of the pre-He-WDs range between 0.17 and 0.65 $R_\odot$. To calculate the mass of the pre-He-WD (M_2), we include the measured radial velocity amplitude if available, which 'replaces' the uncertain mass-ratio measurement from the light curve. For most of the systems, we do not have a radial velocity amplitude measurement, so we do depend on the mass ratio to determine the mass of the pre-He-WDs, which range between 0.12 and 0.5 $M_\odot$. As discussed in Sections 4.1 and 4.5, the mass determination of the secondary using only the mass ratio is very uncertain because of the high uncertainty on the mass ratio. If we limit ourselves only to systems for which we have a radial velocity amplitude, the mass range is 0.12–0.28 $M_\odot$, which is significantly smaller.

For the sample for which we have radial velocity curves, we determine the motion in the Galactic plane and derive their population membership as described in Section 4.4 and shown in Table 3. Fig. 3 shows that more than half of the systems are part of the thin disc population. A few are part of the thick disc, and PTFS1617n could also be a halo object.

6 DISCUSSION

6.1 Co-rotation

In the light-curve modelling (Section 4.1), we assume that both stars are synchronized with the orbit. Previous studies of EL CVn binaries have made the same assumption, but all authors acknowledge that it might not be correct, as mass accretion can spin up the A/F star significantly (see Section 6.2). van Kerkwijk et al. (2010) extensively discuss how all parameter estimates are affected by incorrectly assuming co-rotation. Because the precision of PTF light curves is far lower than the precision of the Kepler light curves, the only significant effect that this assumption has in our analysis is on the estimate of the mass ratio. If a star is rotating faster than the orbital period, the mass ratio (q) is overestimated. We quantify this by simulating a typical EL CVn light curve with a primary star that is rotating two and four times faster than the orbital period, while keeping all other parameters the same. Fitting these light curves with the model assuming co-rotation results in values for q of 0.02 and 0.10 higher than the initial value of $q = 0.17$. All other light-curve parameters do not change significantly. Therefore, we conclude that for mildly faster-than-synchronous rotating primary stars ($P_{rot}/P_{orb} > 0.5$), the effect on the mass ratio is similar

Table 4. System parameters of all EL CVn systems and the uncertainty (standard deviation) on the parameters. Systems for which a radial velocity measurement is used to calculate the parameters are indicated with the 'RV' superscript. This mainly affects the reliability and systematics on the mass and surface gravity of the pre-He-WD (M_2 and $\log g_2$).

Name	P (d)	i (°)	M_1 (M$_\odot$)	M_2 (M$_\odot$)	R_1 (R$_\odot$)	R_2 (R$_\odot$)	T_1 (K)	T_2 (K)	$\log g_1$	$\log g_2$
1600y[RV]	1.1838920	84.5	1.62	0.17	2.41	0.46	6930	8900	3.88	4.33
	0.0000008	2.7	0.04	0.01	0.07	0.02	100	110	0.02	0.03
1600ad[RV]	1.0840448	86.5	1.76	0.23	1.83	0.35	8050	10490	4.16	4.72
	0.0000010	2.2	0.04	0.01	0.05	0.02	120	200	0.02	0.04
1700do	3.0507582	87.4	2.40	0.81	2.34	0.33	9890	17100	4.08	5.31
	0.0000278	1.8	0.06	0.25	0.13	0.03	90	1400	0.04	0.15
1600aa	0.6934558	78.7	1.67	0.50	1.67	0.55	7880	9300	4.21	4.67
	0.0000006	0.9	0.05	0.09	0.04	0.02	190	400	0.02	0.09
1601p[RV]	1.2215885	83.8	1.82	0.14	1.65	0.34	8600	11700	4.26	4.54
	0.0000051	3.2	0.06	0.02	0.14	0.04	160	500	0.06	0.10
1501bh[RV]	0.6204144	78.4	1.30	0.12	1.23	0.20	6870	11100	4.38	4.91
	0.0000005	1.9	0.04	0.01	0.07	0.01	110	400	0.04	0.06
1601q	1.2515058	80.5	1.85	0.30	1.93	0.46	8300	10700	4.13	4.58
	0.0000051	2.9	0.08	0.18	0.15	0.04	200	700	0.05	0.40
1601cl[RV]	0.8917354	82.9	2.02	0.28	2.44	0.52	8290	10100	3.97	4.45
	0.0000005	2.9	0.06	0.01	0.07	0.02	200	300	0.02	0.03
1402de	0.6189694	87.0	1.61	0.36	1.56	0.45	7860	9300	4.27	4.69
	0.0000011	2.5	0.04	0.13	0.07	0.02	150	300	0.03	0.23
1607aa	0.8463120	84.6	1.85	0.30	1.81	0.38	8470	10300	4.19	4.76
	0.0000016	3.4	0.05	0.07	0.08	0.02	160	300	0.03	0.13
1607v	0.7198356	82.6	1.58	0.20	1.83	0.16	7260	10900	4.11	5.32
	0.0000020	5.8	0.06	0.05	0.16	0.03	120	500	0.06	0.24
1607t[RV]	0.8759507	76.6	1.40	0.16	1.87	0.38	6600	8600	4.04	4.48
	0.0000004	1.0	0.05	0.01	0.05	0.01	140	200	0.02	0.03
1607ab[RV]	0.7730986	83.8	1.40	0.19	1.45	0.32	6980	8810	4.26	4.71
	0.0000002	2.3	0.03	0.01	0.05	0.01	100	80	0.03	0.04
1608ab	0.6101718	86.8	1.50	0.11	1.39	0.52	7400	7900	4.32	4.04
	0.0000014	1.9	0.06	0.10	0.04	0.02	200	400	0.02	0.24
1612al	0.6369260	86.8	1.38	0.16	1.16	0.37	7280	10300	4.45	4.50
	0.0000006	2.0	0.08	0.09	0.05	0.02	110	300	0.03	0.22
1512bf[RV]	0.6074343	87.2	1.39	0.17	1.53	0.32	6910	9740	4.21	4.65
	0.0000002	1.9	0.02	0.01	0.02	0.01	90	180	0.01	0.04
1613s	1.1420695	76.2	1.83	0.17	2.72	0.25	7350	13700	3.83	4.88
	0.0000024	6.0	0.08	0.05	0.30	0.05	140	800	0.08	0.28
1613u	0.5644902	81.6	1.52	0.27	1.65	0.37	7340	9690	4.19	4.73
	0.0000003	2.7	0.02	0.04	0.04	0.01	70	160	0.02	0.08
1615ag	0.6806897	85.7	1.52	0.27	1.63	0.32	7370	10200	4.20	4.87
	0.0000046	3.3	0.05	0.07	0.05	0.02	200	400	0.02	0.13
1615v	0.5594054	73.7	1.39	0.13	1.50	0.33	6920	9400	4.23	4.54
	0.0000003	1.4	0.03	0.03	0.04	0.01	120	300	0.02	0.17
1515ay	0.4642873	89.0	1.33	0.15	1.30	0.46	6800	7930	4.33	4.27
	0.0000001	1.1	0.03	0.04	0.02	0.01	100	150	0.01	0.10
1615w	1.4407151	77.7	1.61	0.24	2.59	0.40	6690	10300	3.82	4.63
	0.0000024	2.7	0.05	0.05	0.14	0.03	110	200	0.04	0.14
1615ao	0.8954515	77.6	1.64	0.41	1.82	0.64	7580	8700	4.13	4.43
	0.0000007	0.8	0.05	0.12	0.05	0.03	170	160	0.02	0.15
1615u	0.7777349	82.4	1.50	0.24	1.27	0.16	7400	12200	4.40	5.41
	0.0000011	4.2	0.09	0.10	0.14	0.02	200	600	0.08	0.27
1616cr	0.5649690	82.5	1.40	0.07	1.36	0.46	7060	8000	4.32	3.93
	0.0000002	0.9	0.03	0.02	0.02	0.01	120	170	0.01	0.07
1617n[RV]	2.3367776	87.3	1.80	0.18	2.41	0.38	7500	11600	3.93	4.55
	0.0000052	2.1	0.04	0.02	0.12	0.03	110	400	0.04	0.07
1617m[RV]	3.7728999	87.8	1.68	0.14	2.57	0.69	6990	9320	3.84	3.89
	0.0000083	1.5	0.06	0.03	0.08	0.03	190	190	0.02	0.10
1619l[RV]	1.1599993	83.2	1.56	0.17	2.13	0.34	6870	9200	3.97	4.60
	0.0000017	4.1	0.05	0.01	0.14	0.04	120	150	0.05	0.09
1521ct	1.1724964	83.0	1.82	0.36	1.72	0.56	8520	9800	4.23	4.50
	0.0000014	1.4	0.06	0.29	0.11	0.04	190	300	0.04	0.28
1621ax	1.0181522	84.0	1.69	0.30	2.14	0.17	7350	11800	4.00	5.48
	0.0000045	4.6	0.06	0.07	0.14	0.03	170	700	0.04	0.22
1521cm[RV]	0.6854774	80.0	1.49	0.21	1.49	0.43	7290	9240	4.27	4.49
	0.0000002	1.0	0.02	0.01	0.03	0.01	70	90	0.02	0.03

Table 4 – *continued*

Name	P (d)	i (°)	M_1 (M$_\odot$)	M_2 (M$_\odot$)	R_1 (R$_\odot$)	R_2 (R$_\odot$)	T_1 (K)	T_2 (K)	$\log g_1$	$\log g_2$
1622by	0.7486683	85.8	1.69	0.31	1.84	0.33	7700	11100	4.13	4.88
	0.0000016	3.4	0.07	0.07	0.07	0.02	300	1400	0.03	0.13
1522cc	0.5717853	81.2	1.40	0.26	1.62	0.27	6860	9570	4.17	4.99
	0.0000003	3.1	0.04	0.04	0.05	0.01	120	190	0.02	0.08
1622aa	0.7661291	84.7	1.60	0.16	1.74	0.26	7500	10900	4.16	4.85
	0.0000038	4.0	0.08	0.05	0.10	0.03	300	1300	0.04	0.19
1622bt	0.6884160	79.2	1.65	0.29	1.74	0.29	7700	12200	4.18	4.97
	0.0000004	2.0	0.06	0.04	0.05	0.01	200	1000	0.02	0.08
1723aj	1.1088064	85.6	1.57	0.18	2.51	0.23	6640	11000	3.84	4.98
	0.0000009	3.3	0.05	0.03	0.07	0.02	130	400	0.02	0.11

or smaller than the statistical uncertainty on the mass ratio. If the primary star is rotating faster, then the mass ratio is overestimated.

This overestimate propagates through to the rest of our parameter estimates: the average density of the primary is overestimated, and therefore the mass of the primary is underestimated (Fig. 4) and the semimajor axis is overestimated. However, the effect is small as these parameters only weakly depend on the mass ratio (see equations 4 and 5). For a large part of our sample, we do not have any radial velocity amplitudes, and for these systems we rely on the mass ratio to calculate the mass of the pre-He-WD. As mentioned in Section 4.5, the mass of the pre-He-WD depends on the mass ratio to the third power. Combined with a high statistical and systematic uncertainty on the mass ratio, this makes the calculations of M_2 (without a radial velocity amplitude) unreliable.

To check if the A/F stars are rotating faster than synchronous, we compare the rotation periods to the orbital period for stars in known EL CVn systems. The orbital period of the main-sequence star has been determined for five Kepler EL CVn binaries by measuring the projected rotational velocity ($v \sin i$): 1.79(60) d, 0.79(14) d, 5.0(2.4) d and 1.71(62) d (Faigler et al. 2015) and 0.93 d (Lehmann et al. 2013). In addition, the rotational period has tentatively been identified from a frequency analysis for KOI-81 (0.48 d; Matson et al. 2015), KOI-74 (0.59 d; Bloemen et al. 2012), KOI-1224 (3.49 d; Breton et al. 2012) and KIC-8262223 (0.62 d; Guo et al. 2017). All rotational periods are of the same order as the orbital period. A detailed comparison of the rotational and orbital periods shows that most stars rotate faster than synchronous, but in one case the rotation period is longer than the orbital period. For three cases (all from Faigler et al. 2015) the rotation period is consistent with the orbital period of the binary, but because the uncertainties on the rotational periods are large, it is difficult to say if they are synchronized. Thus, the data indicate that at least some (or maybe most) of the A/F stars are not synchronized with the orbital period.

However, there is an important difference between the PTF sample and the sample of EL CVn systems with known rotation periods of the primary (all found by Kepler). The relative size of the A/F star (r_1) is a factor of ∼3 larger in the PTF sample, which strongly affects the synchronization time-scale of the star ($\propto r_1^{-8.5}$; Zahn 1977). We used the equation by Zahn (1977) and tabulated values for E_2 from Claret (2004) to calculate the synchronization time-scale for each of the EL CVn systems in our sample. This shows that the synchronization time-scale of the A/F-type star is less than 10 Myr in 20 systems, and less than 100 Myr for 32 systems. For these 32 systems, the synchronization time is shorter than the time since mass transfer (0–260 Myr; see Fig. 5). The remaining four systems (with the smallest values for r_1) have synchronization time-scales that are significantly longer than the estimated age. Based on this

theoretical prediction, we can assume that most of the A/F stars are rotating synchronously. We would need to obtain an independent measurement of the rotation period to find whether this is actually the case.

6.2 Binary evolution and stellar parameters

In the canonical formation channel of EL CVn binaries (e.g. Chen et al. 2017), two main-sequence stars of similar mass are born at a short orbital period of a few days. The more massive star evolves faster and increases in radius. Before it ascends the red giant branch (RGB), it fills its Roche lobe and starts stable mass transfer to the lower-mass secondary star. This process continues until almost the complete outer envelope is transferred (identified as R CMa-type binaries; e.g. Lee et al. 2016). The remnant of the initially more massive star has become a pre-WD with a helium core and a thick hydrogen envelope ($\approx$0.01–0.04 M$_\odot$; see Istrate et al. 2016a; Chen et al. 2017). The accretor has become a rejuvenated main-sequence star of spectral type A or F, which dominates the luminosity of the system. If present in a specific stellar population as, for example, found in clusters, such a system would be identified as a 'blue straggler'. If the orbital inclination is such that it shows eclipses, we identify it as an EL CVn binary.

The structure and evolution of pre-He-WDs have been extensively studied, as they also occur at more advanced binary evolutionary stages with either a white dwarf (Marsh et al. 1995) or a neutron star (van Kerkwijk et al. 2005) as companions. Modelling of the formation process of binaries with a low-mass pre-He-WD (Althaus et al. 2013; Istrate et al. 2014a; Istrate, Tauris & Langer 2014b; Istrate et al. 2016a; Chen et al. 2017) shows that there are strong correlations between the binary orbital period, and the mass, temperature, radius and age of the pre-He-WD. First, higher-mass pre-He-WDs are formed at longer orbital periods. This is a direct result of the mass accretion process. This relation was already found in pre-He-WD–neutron star binaries and has been parametrized by Lin et al. (2011). Binary evolution studies by Istrate et al. (2014b) and Chen et al. (2017) also predict this period–mass relation for EL CVn binaries. The latter shows that the relation between orbital period and mass is very robust, but at the low-mass end of the relation (0.16–0.20 M$_\odot$) there is some spread.

Pre-He-WDs (of a given mass) are also predicted to follow a particular evolutionary track, corresponding to a particular combination of radius and temperature as a function of age. The temperature and radius are directly related to the envelope mass and core mass of the white dwarf. Directly after the mass accretion process ends, the pre-He-WD is large ($\gtrsim 0.5$ R$_\odot$) and has a low ($\lesssim 8000$ K) surface temperature. While the hydrogen envelope is slowly being
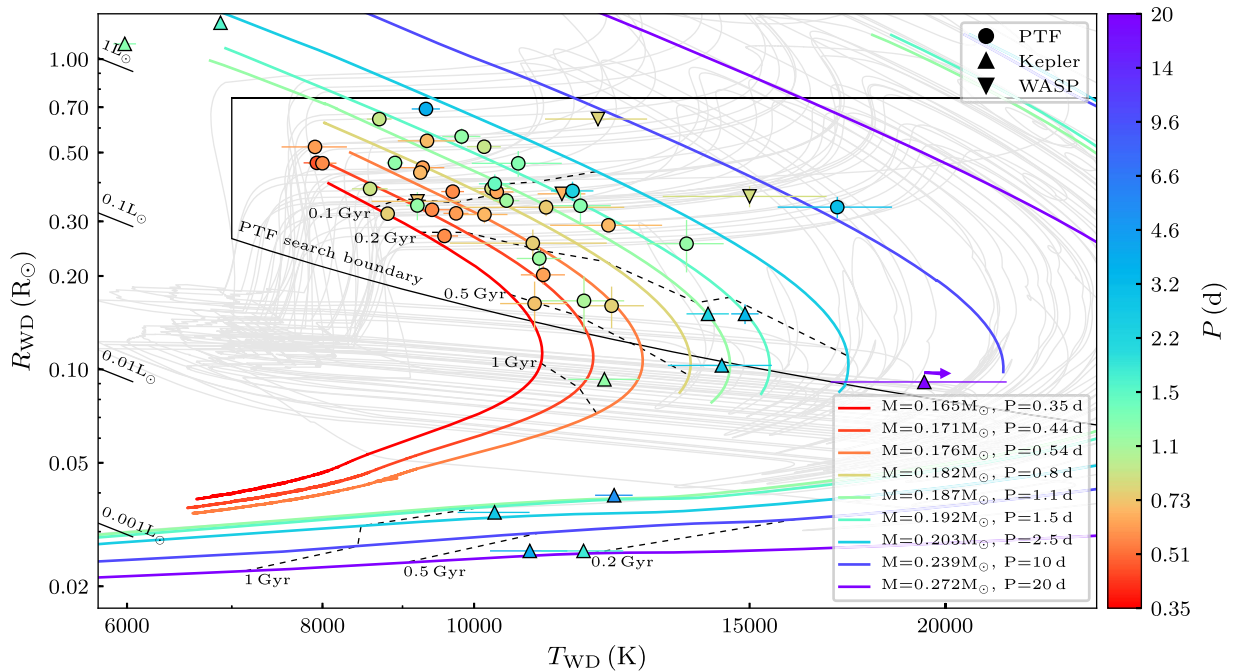
Figure 5. Temperature versus radius of the pre-He-WDs, with the colours indicating different orbital periods. The coloured lines indicate evolution tracks by Althaus et al. (2013) for different masses. The dots indicate pre-He-WDs from this study and triangles are other low-mass pre-He-WDs in EL CVn systems: upward-pointing triangles indicate Kepler discoveries and downward-pointing triangles indicate SWASP discoveries (for references, see Section 1). During the evolution of pre-He-WDs, multiple hydrogen shell flashes can occur, indicated as grey lines. The tracks before the first H-flash and after the last H-flash are shown as coloured lines. Isochrones are shown as dashed lines, counted from the end of mass transfer in the binary. The solid black line indicates the approximate detection limits, estimated by assuming a $T_1 = 7000$ K, $R_1 = 1.5$ R$_\odot$ primary star. The bottom boundary is set by an eclipse depth of 0.03 mag in the R band, the top boundary is set by the requirement of a flat-bottom eclipse ($R_1 > 2R_2$) and the left limit is set by the requirement that the flat-bottom eclipse is deeper than the secondary eclipse ($T_1 < T_2$).

consumed by shell burning, the pre-He-WD shrinks and increases in temperature while maintaining an approximately constant luminosity (this phase is often referred to as the constant luminosity phase). When almost the entire envelope has been consumed, the pre-He-WD starts to cool down while the radius keeps decreasing (the cooling track). At the beginning of the cooling track, multiple hydrogen shell flashes (H-flash; e.g. Driebe et al. 1998) can occur in the more massive pre-He-WDs. These flashes briefly increase the temperature and radius of the star, after which the white dwarf settles back on the cooling track. The exact mass boundary at which this starts to occur is uncertain. Models by Althaus et al. (2013) show shell flashes for masses above $\sim$0.18 M$_\odot$, while Istrate et al. (2014b) put this boundary at $\sim$0.21 M$_\odot$.

Fig. 5 shows the temperature versus radius of the pre-He-WD, with the colour of the points indicating the orbital period of the system. This shows that the temperatures and radii we find are consistent with predictions for pre-He-WDs in the constant luminosity phase, and before the occurrence of any H-flash. While some of the measurements are also consistent with pre-He-WDs undergoing an H-flash (grey lines), the short time spent in this phase makes this extremely unlikely. For the PTF sample, the orbital period (indicated by the colours) follows the same trend as the models, with long-period systems containing larger and hotter pre-He-WDs. To test if the data match the models, we interpolate the models in orbital period, which allows us to test directly how well the radius and temperature match the model for a given orbital period. The fraction of measurements within 1, 3 and 5 standard deviations is 25, 75 and 86 per cent. Given the fact that we interpolate, and the uncertainties on radius and temperature could contain some sys-

tematic uncertainties, we conclude that most of the measurements agree with the models. This comparison to the models also allows us to infer the time since the end of mass transfer, which ranges from 0 to 260 Myr with an average of 110 Myr.

Fig. 6 shows the orbital period of the binary versus the mass of the pre-He-WD. The left panel shows the measured values, while the right panel shows the expected masses using the models inferred from the measured temperature and radius (Fig. 5). The measured values indicate that all pre-He-WDs are low-mass systems, but the PTF sample scatters around the model predictions. The right panel shows that if the radius and temperature measurements and models are used to derive the mass, the results fall within 10 per cent of the prediction of the mass–period relation for pre-He-WDs.

There are two possible explanations for this discrepancy. Either we have underestimated the uncertainties on the mass measurements, or there is some additional intrinsic spread on the predicted mass versus period, radius and temperature not properly modelled. There are a number of possible systematic uncertainties that could affect the mass determination. First of all, we have assumed that the A/F star is a regular main-sequence star with a solar metallicity to estimate its mass (see Fig. 4). In Fig. 6, we have indicated how the mass estimate changes if we assume a different metallicity. If the real metallicity is lower than assumed, the mass is overestimated. This could explain a part of the inconsistency with the theory, but extreme metallicities would be needed to explain the largest outliers. Because thick disc systems generally have a lower metallicity, the masses for these systems could be overestimated. However, the thick disc systems do not show any particular trend,
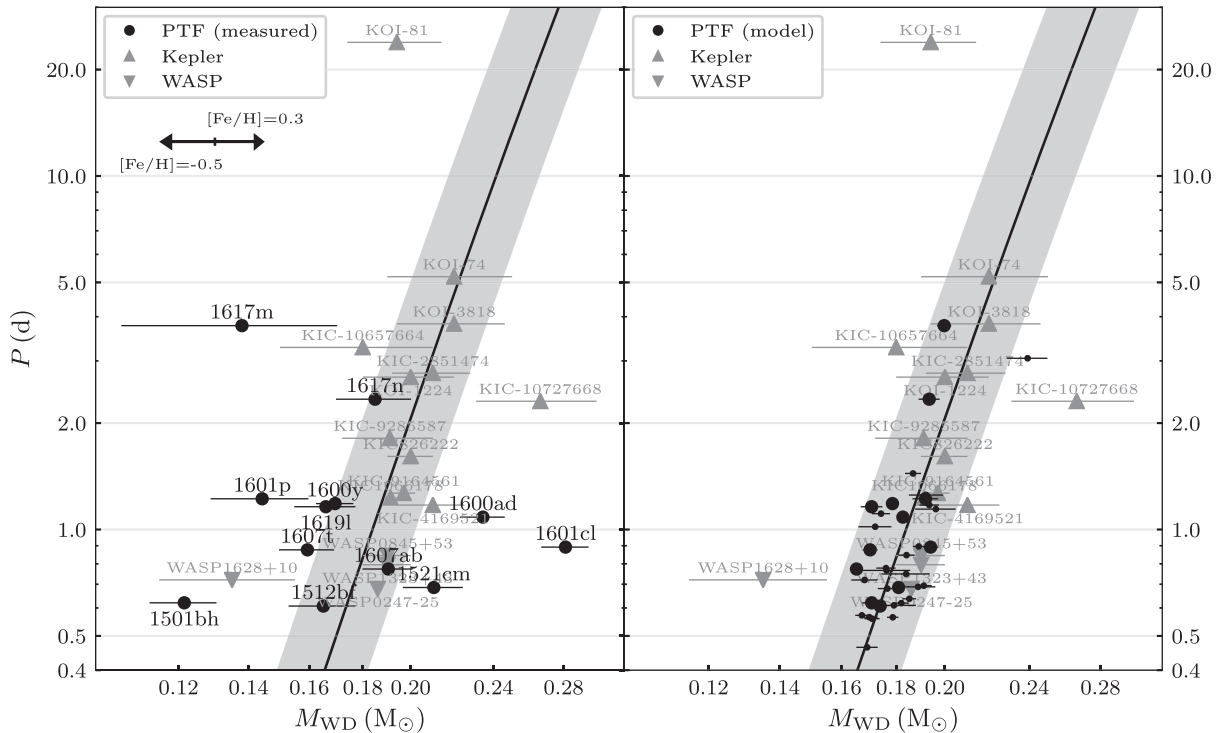
Figure 6. Pre-He-WD mass versus the orbital period of the binary. The left panel shows the measured mass (using the radial velocity amplitude) of the PTF sample and the right panel shows the mass derived from the radius and temperature models (see Fig. 5). Large points indicate PTF systems with a radial velocity measurement and small points indicate systems without a radial velocity measurement (not shown in the right panel because of the high uncertainties). Grey triangles indicate mass–period measurements of other EL CVn systems; upward-pointing triangles indicate Kepler discoveries and downward-pointing triangles indicate SWASP discoveries (for references, see Section 1). The black line shows the period–mass relation by Lin et al. (2011), and the shaded area indicates a 10 per cent uncertainty on this relation.

indicating that this assumption might not be the dominant cause of the inconsistency with the model predictions.

Another possibility is that we have underestimated our measurement uncertainties. The mass of the pre-He-WD is mostly determined by the radial velocity measurement. As shown in Fig. A3 in the Appendix, we need to add an additional uncertainty to the formal uncertainties in order to explain all variance in the radial velocity measurements. For PTFS1601cl (one of the outliers), where we did not obtain radial velocity measurements at the quadratures, small systematic offsets between measurements can have a large impact on the radial velocity amplitude. We did check the radial velocity amplitude measurements of PTFS1512bf by obtaining a few high-resolution spectra with the Echellette Spectrograph and Imager (ESI) on the Keck telescope. The resulting radial velocity amplitude measured from these spectra is consistent with the result from the IDS spectra, which leads to the conclusion that uncertainties resulting from an unstable detector are most likely very small.

An alternative explanation is that there is some intrinsic variance between mass and period, radius and temperature. For example, Istrate et al. (2016b) show that assumptions about rotation, diffusion and metallicity give different results when modelling the mass, radius and temperature of pre-He-WDs. The magnitude of the effect is estimated to be low, about 10 per cent. This would be enough to explain the variance in the right panel but it cannot explain the outliers on the left panel.

To solve this ambiguity, a measurement of both the main-sequence and pre-He-WD radial velocity is needed. This allows the mass of both stars to be calculated by only using Kepler's law (combined with the period and inclination measurement from the

light curve). This has been done for SWASP J0247-25 (Maxted et al. 2013), KOI-81 (Matson et al. 2015), KIC-10661783 (Lehmann et al. 2013) and KIC-8262223 (Guo et al. 2017). For SWASP J0247-25, KIC-10661783 and KIC-10661783, the mass of the pre-He-WD agrees well with the period–mass relation, but for KOI-81, the mass is significantly lower ($0.10\,\mathrm{M_\odot}$) than the period–mass relation predicts. This hints that there is more scatter in the period–mass relation than models estimated.

6.3 Galactic population and space density

Using stellar evolution and population synthesis codes, Chen et al. (2017) predict a space density of 4–$10 \times 10^{-6}\,\mathrm{pc^{-3}}$ for EL CVn binaries (including non-eclipsing ones) with orbital periods less than 2.2 d. In addition, they predicted that EL CVn binaries should mainly be found in young stellar populations, and therefore should be more abundant in the thin disc. We use the Galaxy model based on SDSS data by Jurić et al. (2008) to estimate how many EL CVn binaries we would expect to see given this space density and in what ratios between thin disc, thick disc and halo. We populate our model Galaxy with stars with absolute magnitudes according to a normal distribution with a mean and standard deviation of $R = 2.46 \pm 0.54\,\mathrm{mag}$, values determined from our sample of 36 systems. We simulate the PTF coverage by using (overly) simple requirements: $|b| > 15$, $\delta > 0$, $13.5 < R < 16$ (see Fig. 1). We ignore the Galactic plane because these fields tend to be observed only $\sim$50 times. The minimum number of epochs in our uncovered sample is 58, indicating that at least 58 observations are needed to identify an EL CVn binary. Using the 58 epoch limit, we derive

an effective coverage of 32.8 per cent for the remaining area. We also correct for the requirement that the systems must be eclipsing. This decreases the number of observable EL CVn systems by a factor of 0.307, which is determined from our sample using radii and inclination. Even if the binaries are eclipsing, if the pre-He-WD is too small (and thus old), we would not find it in the PTF data. To correct for this, we assume a lifetime of EL CVn binaries of 1 Gyr (the main-sequence lifetime of a 2-$M_\odot$ star) and we compare this to the typical age of PTF EL CVn binaries (0–260 Myr). Therefore, we assume that the PTF can only detect 26 per cent of all EL CVn binaries.

According to the model and the assumed selection criteria, 26 per cent of the PTF sample should be from the thin disc. If we assume that our model is correct, then there is only a 1.8 per cent chance, $\sum_{n=7}^{12} \binom{12}{n} 0.26^n (1 - 0.26)^{12-n}$, of finding $\geq$ seven thin disc systems out of a total of 12 EL CVn systems. If any of the ambiguous cases are from the thin disc, this probability drops well below 1 per cent. This indicates that our model is unlikely to be correct, and confirms that EL CVn systems are more abundant in the thin disc compared to the average stellar population, as was already suggested by Chen et al. (2017).

Using the model and estimated PTF detection efficiency, we also predict that we should have found ∼300–750 EL CVn systems, a factor of ∼10–25 higher than we actually recovered. This could simply be because we are overpredicting the contributions of the thick disc and halo. However, even if we assume a factor of 4 higher contribution from the thin disc (to bring the model in line with the ratio of thin to thick disc systems), the model still predicts at least a factor of 5–12 more EL CVn systems than we have found. Another uncertain estimate that could explain the discrepancy is the assumed efficiency of the PTF at finding EL CVn binaries. The PTF observing cadence and coverage are highly inhomogeneous, and the assumptions we have used are very simple approximations. Assuming that we can find all EL CVn systems that are observed more than 58 times and are younger than 200 Myr is overly optimistic, and could explain the discrepancy of a factor of 5 (or more).

The inhomogeneity of the PTF data set makes it difficult to perform a proper study of the Galactic distribution and space density of EL CVn systems. We do find tentative results that EL CVn systems occur more often in the thin disc, as was predicted by Chen et al. (2017). We also find that the space density is at the lower bound, or even below the prediction of 4–10×10^{-6} pc^{-3}. To properly measure the properties of the population of EL CVn systems, a larger sample of EL CVn binaries is needed, preferably from a more uniform variability survey.

6.4 Comparison with the SWASP sample

To better understand the biases of our survey, we compare the PTF sample to the sample found by SWASP (Maxted et al. 2014a). While we have both used the light-curve characteristics to identify EL CVn binaries, there are some intrinsic differences between the surveys, and therefore different biases in finding EL CVn systems. The most obvious difference between the surveys is the magnitude range; SWASP probes magnitudes between 9 and 13 mag, while our sample is fainter, between 13.5 and 16 mag. A second major difference is the cadence and the number of epochs in a light curve; PTF light curves have an irregular cadence and a low number of epochs (∼100) compared to the regular cadence and better-sampled SWASP light curves (∼4000–13 000 epochs).

There are indeed differences between the two samples. First, the distance range for the SWASP sample is 100–1200 pc, while the PTF
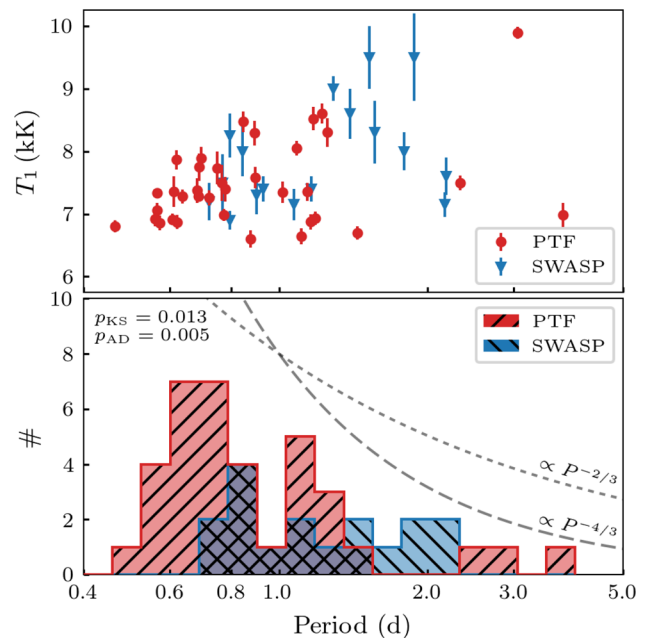
Figure 7. Top: the temperature of the A/F star versus the orbital period of the EL CVn system. Bottom: the distribution of orbital periods of the PTF sample (red) and the SWASP sample (blue). In the top left, the results of the KS test and AD test are shown (see text). The top panel shows that there is a strong correlation between orbital period and temperature, which is a result of the binary evolution process (see text). The histograms show that the PTF data are more biased to short-period systems compared to the SWASP sample. It also shows a possible gap at periods of 1 d, caused by a detection bias against these systems. For comparison, we also plotted the detection probability of an eclipsing population with well-sampled light curves ($\propto P^{-2/3}$, dotted line), and for light curves with a limited amount of epochs ($\propto P^{-4/3}$, dashed line).

sample reaches 1200–5000 pc. This is expected given the different magnitude range of the two surveys. Therefore, we also expect to find relatively more thick disc and halo systems compared to thin disc systems in the PTF sample. However, using the one-sided Fisher's exact test (Fisher 1934), we find no significant difference in the relative number of thin disc systems. This is consistent with our finding that EL CVn systems are more numerous in the thin disc compared to the average stellar population and it explains why at larger distances it is still the most dominant population.

The two samples are also different with regards to orbital period and temperature of the A/F star (which dominates the luminosity); see Fig. 7. We performed a Kolmogorov–Smirnov (KS) test and an Anderson–Darling (AD) test (e.g. section 3.1 in Feigelson & Babu 2012) to compare the distribution of orbital periods. Both tests show that it is unlikely that the samples are drawn from the same distribution ($p_{\mathrm{KS}} = 0.013$ and $p_{\mathrm{AD}} = 0.005$). The histogram in Fig. 7 shows that we find more short orbital period systems and the top panel in Fig. 7 shows that, at short orbital periods, the temperature of the primary star is low. This correlation can be understood because the mass of the A/F star is correlated with the orbital period. High-mass main-sequence stars ($2\,\mathrm{M_\odot}$) begin their main-sequence lifetime at a temperature of 9500 K, and cool down to 7500 K towards the end of their main-sequence lifetime (see Fig. 4). Main-sequence stars of $1.3\,\mathrm{M_\odot}$ start at a temperature of 7500 K and only cool by 500 K during their time on the main sequence. Therefore, the PTF's sensitivity for lower-luminosity (lower-temperature) EL CVn systems

(partially) explains why the PTF sample contains more short-period systems.

A second explanation is that, because of the sparse sampling, it is harder to find long-period systems with the PTF compared to SWASP. Both surveys use eclipses to find EL CVn binaries, and therefore are biased towards short-period systems ($Prob_{\rm ecl} \propto R_1 P^{-2/3}$). In addition, short-period systems spend a larger fraction of their orbit in eclipse ($\tau_{\rm ecl} \propto R_1 P^{-2/3}$). This does not bias the SWASP search as the light curves are well sampled. With the PTF, however, a lack of observations during the eclipse can hinder the identification of a system as an EL CVn binary.

The difference between the two samples shows that selection effects make it difficult to determine the intrinsic properties of EL CVn binaries. To do so requires an integrated approach: stellar evolution and population synthesis models should be used to simulate a sample of EL CVn binaries, which are then 'observed' by simulating the variability survey that was used to find the real sample. Such a calculation is difficult given the inhomogeneity of the PTF sample, and it is beyond the scope of this work.

6.5 Pulsations

Pre-He-WDs are predicted to exhibit both p- and g-mode pulsations (e.g. Córsico & Althaus 2014, 2016; Córsico et al. 2016; Istrate et al. 2016b). Pulsations have been found in two of the SWASP EL CVn binaries: WASP 0247-25 (Maxted et al. 2013; Istrate, Fontaine & Heuser 2017) and WASP 1628+10 (Maxted et al. 2014b). The pulsation periods are 5–10 min and the amplitudes ~1–2 per cent of the pre-He-WD luminosity. Models of pre-He-WDs predict that in a large area in T–$\log g$ space, pre-He-WDs should feature pulsations; see fig. 10 in Córsico et al. (2016) and see also Istrate et al. (2016b). Many of the pre-He-WDs in the PTF sample lie in this region, making them useful to test the general predictions for pulsation theory. In addition, because stellar parameters can be measured very precisely, a pulsating pre-He-WD in an eclipsing binary is extremely useful to test evolutionary and seismic models in great detail (e.g. Istrate et al. 2017).

Unfortunately, the very sparse sampling of the PTF light curves makes it very difficult to identify such pulsations. We did attempt to find pulsations by using a Lomb–Scargle algorithm (Lomb 1976; Scargle 1982, implementation by VanderPlas & Ivezić 2015) on the residuals of the light curves. We found periodic behaviour in the residuals at predicted periods of ~10 min for a number of the systems, but because of the sparse sampling and low amplitude, it is difficult to determine if these are real or not. High-cadence follow-up photometry is needed to establish the reality of these pulsations.

7 SUMMARY AND CONCLUSION

In this paper, we report on the discovery and analysis of 36 new EL CVn systems extracted from PTF data. With this sample, we more than double the number of known EL CVn systems. To find the EL CVn systems, we used machine learning classifiers to make a pre-selection of candidates. This has proven to be an efficient method to minimize the number of light curves that have to be visually inspected.

The radii (0.16–0.7 $R_\odot$) and temperatures (8000–17 000 K) of the pre-He-WDs in the sample indicate that they are all young systems in the constant luminosity phase (0–250 Myr) of their evolution. The masses of the pre-He-WDs are all low (<0.3 $M_\odot$), but

our measurements show a large spread around the predicted mass–period relation, which remains unexplained. If we use the measured radii and temperatures combined with models, we do find masses consistent with the mass–period relation. Either this discrepancy is due to systematic or underestimated uncertainties in our measurements, or there is more variance in the masses than the stellar evolution models predict. This problem can be resolved by obtaining more accurate radial velocity measurements (ideally for both stars in the binary to obtain an independent mass ratio measurement), and by more extensively testing the effect on the mass–period relation of, for example, different metallicities and rotation rates.

Although a detailed population study is difficult with the PTF data set, we find that EL CVn binaries occur more often in the thin disc than an average Galactic stellar population. In addition, we find that the space density is most likely lower than the predicted value of 4–10×10^{-6} pc^{-3}. To properly determine the properties of the EL CVn population, a more systematic search combined with stellar and Galactic modelling is required.

This new sample of young pre-He-WDs will be useful to put many theoretical models to the test, including stellar structure models for low-mass white dwarfs, pulsation models and binary evolution models. In addition, the methods we have used to identify EL CVn systems can easily be adapted to find other rare types of variable stars and these (and similar machine learning methods) will be vital to fully utilize (future) variability surveys such as the ZTF (Bellm 2014), NGTS (Wheatley et al. 2018), GOTO (Steeghs & Galloway 2017), BlackGEM (Bloemen et al. 2015), the *Transiting Exoplanet Survey Satellite* (*TESS*; Ricker et al. 2015), *PLATO* (Rauer et al. 2014) and the Large Synoptic Survey Telescope (LSST; Ivezić et al. 2008).

REFERENCES

Albareti F. D., (SDSS Collaboration) et al., 2016, ApJS, 233, 25

Althaus L. G., Miller Bertolami M. M., Córsico A. H., 2013, A&A, 557, A19

Andreas C. Müller S. G., 2016, Introduction to Machine Learning with Python, O'Reilly, Sebastopol, CA

Angeloni R. et al., 2014, A&A, 567, A100

Armstrong D. J. et al., 2016, MNRAS, 456, 2260

Bagnulo S., Jehin E., Ledoux C., Cabanac R., Melo C., Gilmozzi R., ESO Paranal Science Operations Team, 2003, The Messenger, 114, 10

Bellm E., 2014, in Wozniak P. R., Graham M. J., Mahabal A. A., Seaman R., eds, The Third Hot-wiring the Transient Universe Workshop. Los Alamos National Laboratory, Santa Fe, NM, p. 27 (arXiv:1410.8185)

Bianchi L., Conti A., Shiao B., 2014, VizieR Online Data Catalog, 2335

Bloemen S. et al., 2012, MNRAS, 422, 2600

Bloemen S., Groot P., Nelemans G., Klein-Wolt M., 2015, in Rucinski S. M., Torres G., Zejda M., eds, ASP Conf. Ser. Vol. 496, Living Together: Planets, Host Stars and Binaries. Astron. Soc. Pac., San Francisco, p. 254

Breiman L., 2001, Mach. Learn., 45, 5

Breton R. P., Rappaport S. A., van Kerkwijk M. H., Carter J. A., 2012, ApJ, 748, 115

Brown W. R., Kilic M., Allende Prieto C., Kenyon S. J., 2010, ApJ, 723, 1072

Budding E., Butland R., 2011, MNRAS, 418, 1764

Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245

Carter J. A., Rappaport S., Fabrycky D., 2011, ApJ, 728, 139

Chambers K. C. et al., 2016, preprint (arXiv:1612.05560)

Chen T., Guestrin C., 2016, preprint (arXiv:1603.02754)

Chen X., Maxted P. F. L., Li J., Han Z., 2017, MNRAS, 467, 1874

Claret A., 2004, A&A, 424, 919

Copperwheat C. M., Marsh T. R., Dhillon V. S., Littlefair S. P., Hickman R., Gänsicke B. T., Southworth J., 2010, MNRAS, 402, 1824

Copperwheat C. M. et al., 2011, MNRAS, 410, 1113

Córsico A. H., Althaus L. G., 2014, A&A, 569, A106

Córsico A. H., Althaus L. G., 2016, A&A, 585, A1

Córsico A. H., Althaus L. G., Serenelli A. M., Kepler S. O., Jeffery C. S., Corti M. A., 2016, A&A, 588, A74

Debosscher J., 2009, PhD thesis, Institute of Astronomy, Katholieke Universiteit Leuven, Belgium

Devor J., 2005, ApJ, 628, 411

Driebe T., Schoenberner D., Bloecker T., Herwig F., 1998, A&A, 339, 123

Faigler S., Kull I., Mazeh T., Kiefer F., Latham D. W., Bloemen S., 2015, ApJ, 815, 26

Feigelson E. D., Babu G. J., 2012, Modern Statistical Methods for Astronomy. Cambridge Univ. Press, Cambridge

Fisher R., 1934, Statistical Methods for Research Workers, 5 edn. Oliver and Boyd, Edinburgh.

Foreman-Mackey D., 2013, Example: Fitting a Model to Data. Available at: http://dan.iel.fm/emcee/current/user/line/#maximum-likelihood-estimation

Foreman-Mackey D., 2014, Tutorial: model fitting with correlated noise. Available at: http://dfm.io/george/current/user/model/

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Geurts P., Ernst D., Wehenkel L., 2006, Mach. Learn., 63, 3

Gianninas A., Strickland B. D., Kilic M., Bergeron P., 2013, ApJ, 766, 3

Gies D. R. et al., 2008, ApJ, 682, L117

Guo Z., Gies D. R., Matson R. A., García Hernández A., Han Z., Chen X., 2017, ApJ, 837, 114

Hartman J. D., Bakos G. Á., 2016, Astron. d Comput., 17, 1

Hogg D. W., Bovy J., Lang D., 2010, preprint (arXiv:1008.4686)

Istrate A. G., Tauris T. M., Langer N., Antoniadis J., 2014a, A&A, 571, L3

Istrate A. G., Tauris T. M., Langer N., 2014b, A&A, 571, A45

Istrate A. G., Marchant P., Tauris T. M., Langer N., Stancliffe R. J., Grassitelli L., 2016a, A&A, 595, A35

Istrate A. G., Fontaine G., Gianninas A., Grassitelli L., Marchant P., Tauris T. M., Langer N., 2016b, A&A, 595, L12

Istrate A. G., Fontaine G., Heuser C., 2017, ApJ, 847, 130

Ivezić v. et al., 2008, Serbian Astronomical Journal, 176, 1

Ivezić Ž., Connolly A., Vanderplas J., Gray A., 2014, Statistics, Data Mining and Machine Learning in Astronomy. Princeton University Press, Princeton, NJ

Jurić M. et al., 2008, ApJ, 673, 864

Kaluzny J., Rucinski S. M., Thompson I. B., Pych W., Krzeminski W., 2007, AJ, 133, 2457

Kervella P., Fouqué P., 2008, A&A, 491, 855

Kovács G., Zucker S., Mazeh T., 2002, A&A, 391, 369

Laher R. R. et al., 2014, PASP, 126, 674

Law N. M. et al., 2009, PASP, 121, 1395

Lee J. W., Kim S-L., Hong K., Koo J-R., Lee C-U., Youn J-H., 2016, AJ, 151, 25

Lehmann H., Southworth J., Tkachenko A., Pavlovski K., 2013, A&A, 557, A79

Lin J., Rappaport S., Podsiadlowski P., Nelson L., Paxton B., Todorov P., 2011, ApJ, 732, 70

Lomb N. R., 1976, Ap&SS, 39, 447

Mackenzie C., Pichara K., Protopapas P., 2016, ApJ, 820, 138

Marsh T. R., Dhillon V. S., Duck S. R., 1995, MNRAS, 275, 828
Masci F., Bellm E., 2016, PTF light curve statistics. Available at: http://www.ptf.caltech.edu/system/media_files/binaries/30/original/Objects_SourcesTable_cols_v3.html
Masci F. J., Hoffman D. I., Grillmair C. J., Cutri R. M., 2014, AJ, 148, 21
Matson R. A., Gies D. R., Guo Z., Quinn S. N., Buchhave L. A., Latham D. W., Howell S. B., Rowe J. F., 2015, ApJ, 806, 155
Maxted P. F. L. et al., 2011, MNRAS, 418, 1156
Maxted P. F. L. et al., 2013, Nature, 498, 463
Maxted P. F. L. et al., 2014a, MNRAS, 437, 1681
Maxted P. F. L., Serenelli A. M., Marsh T. R., Catalán S., Mahtani D. P., Dhillon V. S., 2014b, MNRAS, 444, 208
McAllister M. J. et al., 2017, MNRAS, 464, 1353
Palaversa L. et al., 2013, AJ, 146, 101
Parsons S. G., Marsh T. R., Gänsicke B. T., Drake A. J., Koester D., 2011, ApJ, 735, L30
Pauli E.-M., Napiwotzki R., Heber U., Altmann M., Odenkirchen M., 2006, A&A, 447, 173
Pecaut M. J., Mamajek E. E., 2013, ApJS, 208, 9
Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825
Peters C. M. et al., 2015, ApJ, 811, 95
Pietrzyński G. et al., 2012, Nature, 484, 75
Pollacco D. L. et al., 2006, PASP, 118, 1407
Rappaport S., Podsiadlowski P., Horev I., 2009, ApJ, 698, 666
Rappaport S., Nelson L., Levine A., Sanchis-Ojeda R., Gandolfi D., Nowak G., Palle E., Prsa A., 2015, ApJ, 803, 82
Rau A. et al., 2009, PASP, 121, 1334
Rauer H. et al., 2014, Experimental Astronomy, 38, 249
Richards J. W. et al., 2011, ApJ, 733, 10
Richards J. W. et al., 2012, ApJ, 744, 192
Ricker G. R. et al., 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003
Robitaille T. P. et al., (Astropy Collaboration) 2013, A&A, 558, A33
Scargle J. D., 1982, ApJ, 263, 835
Schlafly E. F., Finkbeiner D. P., 2011, ApJ, 737, 103
Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
Schwarzenberg-Czerny A., 1989, MNRAS, 241, 153
Sesar B. et al., 2017, AJ, 153, 204
Skrutskie M. F. et al., 2006, AJ, 131, 1163
Smart R. L., Nicastro L., 2013, VizieR Online Data Catalog, 1324
Spada F., Demarque P., Kim Y-C., Boyajian T. S., Brewer J. M., 2017, ApJ, 838, 161
Steeghs D., Galloway D., 2017, Gravitational wave Optical Transient Observatory. Available at: https://www2.warwick.ac.uk/fac/sci/physics/research/astro/research/goto/
Stetson P. B., 1996, PASP, 108, 851
Tody D., 1993, in Hanisch R. J., Brissenden R. J. V., Barnes J., eds, ASP Conf. Ser. Vol. 52, Astronomical Data Analysis Software and Systems II. Astron. Soc. Pac., San Francisco, p. 173
van Dokkum P. G., 2001, PASP, 113, 1420
van Kerkwijk M. H., Bassa C. G., Jacoby B. A., Jonker P. G., 2005, in Rasio F. A., Stairs I. H., eds, ASP Conf. Ser. Vol. 328, Binary Radio Pulsars. Astron. Soc. Pac., San Francisco, p. 357 (astro-ph/0405283)
van Kerkwijk M. H., Rappaport S. A., Breton R. P., Justham S., Podsiadlowski P., Han Z., 2010, ApJ, 715, 51
VanderPlas J. T., Ivezić Ž., 2015, ApJ, 812, 18
von Zeipel H., 1924, MNRAS, 84, 665
Westera P., Lejeune T., Buser R., Cuisinier F., Bruzual G., 2002, A&A, 381, 524
Wheatley P. J. et al., 2018, MNRAS, in press
Wright E. L. et al., 2010, AJ, 140, 1868
Zacharias N., Monet D. G., Levine S. E., Urban S. E., Gaume R., Wycoff G. L., 2004, Bulletin of the American Astronomical Society, 36, 1418
Zacharias N., Finch C., Frouard J., 2017, VizieR Online Data Catalog, 1340
Zahn J.-P., 1977, A&A, 57, 383

APPENDIX A: ADDITIONAL TABLES AND FIGURES

Table A1. List of the features used by the machine learning classifiers.

PTF variability	
wRMS (mag)	Weighted rms of the PTF light curve
skew (mag)	Skewness of the PTF light curve
medAbsDev (mag)	Median absolute deviation of the PTF light curve
StetsJ	Stetson-J statistic of the PTF light curve
StetsK	Stetson-K statistic of the PTF light curve
Neumann	Von Neumann ratio statistic of the PTF light curve
MBR	Median buffer range: fraction of points more than 20 per cent of the light-curve amplitude from the weighted mean magnitude, divided by total number of epochs
AMBS{1,2,3}	Fraction of light-curve points that are # standard deviation above the mean magnitude
BMBS{1,2,3}	Fraction of light-curve points that are # standard deviation below the mean magnitude
prange{#} (mag)	Range containing {90,80,65,50,35,20} per cent of the data.
percentile{#}	#th percentile minus the median of the PTF light curve, divided by prange90, with # in {5,10,17.5,25,32.5,40,60,67.5,75,82.5,90,95} .
PAN-STARRS colours	
PSr (mag)	Pan-STARRS r – median of the light curve
PSgr (mag)	Pan-STARRS $g - r$
PSri (mag)	Pan-STARRS $r - i$
PSiz (mag)	Pan-STARRS $i - z$
PSzy (mag)	Pan-STARRS $z - y$
2MASS colours	
J (mag)	2MASS J – median of the light curve
JH (mag)	2MASS $J - H$
HK (mag)	2MASS $H - K$

Table A2. Overview of the nights at the INT with the IDS.

Date	Grating	CCD	Seeing (arcsec)	Weather
2016-09-07	R632V	RED+2	0.6	Excellent
2016-09-08	R632V	RED+2	0.7	Excellent
2016-09-09	R632V	RED+2	0.6	Excellent
2016-09-10	R632V	RED+2	0.7	Good
2016-09-11	R632V	RED+2	0.6–1.0	Good
2016-09-12	R632V	RED+2	0.7–1.0	Good
2016-09-13	R632V	RED+2	0.8–1.2	Good
2016-09-14	R632V	RED+2	1.0	Good
2016-12-14	R900V	RED+2	0.8–1.4	Good
2016-12-15	R900V	RED+2	1.4	OK
2017-01-09	R900V	RED+2	0.8	Good
2017-01-10	R900V	RED+2	1.2–2.6	OK–Bad
2017-03-10	R900V	EEV10	1.5–3.0	Bad
2017-03-11	R900V	EEV10	–	Clouds
2017-03-12	R900V	EEV10	2–4	Bad
2017-03-13	R900V	EEV10	1.5	Bad

Table A3. The temperatures of the A/F star (T_1) and pre-He-WD (T_2) determined from the SED of the binary stars (see Fig. A2). The $E(B - V)$ values are taken from Schlegel, Finkbeiner & Davis (1998) and Schlafly & Finkbeiner (2011), with an uncertainty of 0.034 (as in Maxted et al. 2011). The 'rms' column indicates the additional uncertainty added to account for all variance, which is achieved by the parameter f in equation (2).

ID	T_1 (K)	T_2 (K)	$E(B - V)$	rms
1600y	6930 ± 100	8900 ± 110	0.047	0.05
1600ad	8050 ± 120	10490 ± 200	0.024	0.04
1700do	9890 ± 110	17100 ± 1400	0.015	0.03
1600aa	7890 ± 190	9300 ± 400	0.107	0.02
1601p	8600 ± 160	11700 ± 500	0.030	0.03
1501bh	6870 ± 110	11100 ± 400	0.035	0.03
1601q	8300 ± 230	10700 ± 700	0.081	0.03
1601cl	8280 ± 200	10100 ± 300	0.030	0.04
1402de	7870 ± 150	9300 ± 300	0.100	0.03
1607aa	8470 ± 160	10300 ± 300	0.087	0.04
1607v	7260 ± 130	10900 ± 500	0.090	0.03
1607t	6600 ± 140	8600 ± 200	0.009	0.10
1607ab	6980 ± 100	8810 ± 80	0.005	0.09
1608ab	7360 ± 240	7900 ± 400	0.037	0.05
1612al	7280 ± 110	10300 ± 300	0.039	0.01
1512bf	6920 ± 90	9740 ± 180	0.022	0.03
1613s	7350 ± 140	13700 ± 800	0.051	0.04
1613u	7340 ± 70	9690 ± 160	0.006	0.04
1615ag	7380 ± 200	10200 ± 400	0.093	0.05
1615v	6920 ± 120	9400 ± 300	0.030	0.03
1515ay	6800 ± 100	7930 ± 150	0.029	0.04
1615w	6690 ± 110	10300 ± 200	0.046	0.04
1615ao	7580 ± 170	8700 ± 160	0.070	0.03
1615u	7400 ± 200	12200 ± 600	0.069	0.06
1616cr	7060 ± 120	8000 ± 170	0.095	0.10
1617n	7500 ± 110	11600 ± 400	0.022	0.03
1617m	6990 ± 190	9320 ± 190	0.071	0.04
1619l	6870 ± 120	9200 ± 150	0.031	0.04
1521ct	8520 ± 180	9800 ± 300	0.090	0.02
1621ax	7340 ± 170	11800 ± 700	0.108	0.04
1521cm	7290 ± 80	9240 ± 90	0.051	0.02
1622by	7730 ± 260	11100 ± 1400	0.079	0.07
1522cc	6860 ± 120	9570 ± 190	0.042	0.06
1622aa	7490 ± 290	10900 ± 1300	0.160	0.11
1622bt	7750 ± 220	12200 ± 1000	0.061	0.07
1723aj	6640 ± 130	11000 ± 400	0.061	0.04

Table A4. The parameters of the light-curve models shown in Fig. A1. This table shows the median and rms of the final 5120 points in our MCMC chains. Note that these distributions are not normally distributed and parameters can be strongly correlated.

ID band	P (d)	t_0 (BMJD$_{tdb}$)	i (°)	q	$\langle r_1 \rangle$	$\langle r_2 \rangle$	T_2/T_1	absorb$_R$	absorb$_{g'}$	$\log(f_R)$	$\log(f_{g'})$	Fill	ρ ($\rho_\odot$)
1600y	1.1838920	55570.2084	84.5	0.12	0.421	0.081	0.73	0.8	0.8	−2.1	−2.5	0.748	0.114
R+g	0.0000008	0.0006	2.7	0.02	0.010	0.003	0.03	0.2	0.2	0.0	0.1	0.019	0.007
1600ad	1.0840448	56247.4677	86.5	0.11	0.328	0.063	0.80	1.3	1.0	−2.3	−2.4	0.575	0.292
R+g	0.0000010	0.0006	2.2	0.03	0.008	0.003	0.03	0.5	0.5	0.0	0.1	0.025	0.019
1700do	3.0507595	55556.8044	87.4	0.33	0.178	0.025	0.77	2.8		−2.3		0.380	0.187
R	0.0000281	0.0014	1.8	0.10	0.010	0.002	0.04	1.2		0.0		0.020	0.024
1600aa	0.6934558	56892.6368	78.7	0.29	0.391	0.128	0.94	3.6	3.2	−2.5	−2.6	0.801	0.362
R+g	0.0000006	0.0003	0.8	0.05	0.007	0.004	0.01	0.8	0.8	0.2	0.4	0.026	0.021
1601p	1.2215885	57152.5232	84.0	0.21	0.274	0.056	0.84	1.4	1.1	−2.2	−2.3	0.534	0.358
R+g	0.0000051	0.0010	3.1	0.09	0.019	0.005	0.03	1.0	1.1	0.0	0.3	0.036	0.059
1501bh	0.6204144	55097.3927	78.4	0.14	0.356	0.058	0.64	0.5	0.4	−2.1	−2.1	0.646	0.674
R+g	0.0000005	0.0008	1.9	0.04	0.017	0.004	0.04	0.3	0.2	0.1	0.0	0.023	0.085
1601q	1.2515054	57190.1373	80.6	0.12	0.308	0.074	0.91	1.6	2.2	−2.3	−2.5	0.535	0.266
R+g	0.0000051	0.0012	2.7	0.07	0.021	0.007	0.03	1.1	1.2	0.1	0.3	0.049	0.051
1601cl	0.8917354	56063.3087	82.9	0.11	0.475	0.101	0.95	2.3	1.9	−2.1	−2.3	0.830	0.142
R+g	0.0000005	0.0004	2.9	0.02	0.011	0.004	0.02	0.8	0.7	0.0	0.0	0.021	0.009
1402de	0.6189694	55768.8152	87.0	0.17	0.405	0.119	0.91	2.5		−2.9		0.757	0.454
R	0.0000011	0.0008	2.4	0.07	0.015	0.006	0.03	1.1		0.3		0.052	0.049
1607aa	0.8463124	56246.6579	84.4	0.14	0.375	0.079	0.92	2.2	3.7	−2.6	−2.6	0.678	0.314
R+g	0.0000017	0.0007	3.6	0.03	0.017	0.005	0.03	1.0	0.9	0.3	0.2	0.024	0.035
1607v	0.7198355	55769.1206	82.7	0.13	0.447	0.039	0.66	2.0	2.2	−2.4	−2.3	0.808	0.256
R+g	0.0000020	0.0014	5.7	0.04	0.037	0.007	0.08	1.0	1.0	0.1	0.0	0.038	0.048
1607t	0.8759507	56158.7102	76.6	0.09	0.417	0.085	0.74	0.7	0.5	−2.3	−2.3	0.712	0.220
R+g	0.0000004	0.0004	1.0	0.02	0.009	0.003	0.03	0.2	0.2	0.0	0.0	0.020	0.014
1607ab	0.7730986	55151.7862	83.8	0.18	0.351	0.077	0.77	1.9	1.5	−2.3	−2.2	0.665	0.439
R+g	0.0000002	0.0004	2.3	0.03	0.010	0.003	0.03	0.4	0.4	0.1	0.0	0.020	0.035
1608ab	0.6101718	57034.9178	86.8	0.07	0.390	0.146	0.90	0.4		−3.0		0.656	0.551
R	0.0000014	0.0003	1.9	0.07	0.006	0.004	0.02	0.3		0.3		0.058	0.044
1612al	0.6369260	55782.6928	86.8	0.12	0.322	0.103	0.69	0.1		−2.3		0.574	0.880
R	0.0000006	0.0007	2.0	0.06	0.009	0.004	0.03	0.1		0.1		0.043	0.084
1512bf	0.6074343	56100.9311	87.2	0.10	0.438	0.091	0.66	0.1	0.1	−2.1	−2.2	0.762	0.391
R+g	0.0000002	0.0002	1.9	0.02	0.005	0.003	0.03	0.1	0.1	0.0	0.0	0.017	0.012
1613s	1.1420695	56511.1762	76.2	0.10	0.470	0.044	0.60	1.2	1.2	−2.2	−2.2	0.812	0.091
R+g	0.0000024	0.0014	6.0	0.03	0.050	0.008	0.05	0.5	0.5	0.1	0.0	0.050	0.023
1613u	0.5644902	56787.5255	81.6	0.18	0.472	0.107	0.77	1.4	1.6	−2.4	−2.5	0.889	0.340
R+g	0.0000003	0.0003	2.7	0.03	0.012	0.004	0.03	0.4	0.4	0.1	0.2	0.023	0.024
1615ag	0.6806898	55380.8645	85.8	0.17	0.410	0.079	0.76	2.2	1.9	−2.2	−2.4	0.772	0.358
R+g	0.0000046	0.0010	3.3	0.04	0.013	0.005	0.05	1.0	0.8	0.3	0.1	0.026	0.029
1615v	0.5594054	54962.6621	73.7	0.10	0.457	0.099	0.71	0.2	0.1	−2.2	−2.9	0.784	0.410
R+g	0.0000003	0.0007	1.4	0.02	0.012	0.004	0.03	0.1	0.1	0.1	0.3	0.029	0.031
1515ay	0.4642873	56138.1751	89.0	0.11	0.452	0.161	0.83	0.0	0.1	−2.2	−1.9	0.794	0.606
R+g	0.0000001	0.0002	1.1	0.03	0.003	0.002	0.02	0.0	0.1	0.1	0.0	0.027	0.018
1615w	1.4407151	56530.5582	77.7	0.15	0.393	0.060	0.60	0.7	0.4	−2.9	−2.5	0.723	0.092
R+g	0.0000024	0.0010	2.7	0.04	0.021	0.004	0.04	0.3	0.2	0.3	0.1	0.026	0.015
1615ao	0.8954515	56308.1400	77.6	0.24	0.368	0.129	0.94	2.1	1.9	−2.4	−2.5	0.729	0.273
R+g	0.0000007	0.0005	0.8	0.06	0.007	0.005	0.02	1.1	1.0	0.3	0.2	0.038	0.020
1615u	0.7777349	56185.6219	82.4	0.16	0.297	0.038	0.59	1.4	1.8	−2.1	−2.2	0.554	0.726
R+g	0.0000011	0.0011	4.2	0.07	0.034	0.006	0.05	0.7	0.8	0.1	0.0	0.041	0.191
1616cr	0.5649690	55972.1677	82.5	0.05	0.416	0.141	0.84	0.4	0.5	−2.3	−3.0	0.657	0.558
R+g	0.0000002	0.0002	0.9	0.01	0.005	0.002	0.02	0.2	0.1	0.2	0.2	0.021	0.025
1617n	2.3367776	55591.4345	87.3	0.21	0.258	0.040	0.63	2.6	1.5	−2.1	−2.4	0.501	0.117
R+g	0.0000052	0.0027	2.1	0.07	0.012	0.003	0.04	0.8	0.5	0.1	0.1	0.025	0.013
1617m	3.7728999	56584.2146	87.8	0.24	0.206	0.056	0.68	0.6	0.7	−2.1	−2.2	0.410	0.085
R+g	0.0000083	0.0010	1.5	0.15	0.005	0.002	0.03	0.5	0.4	0.1	0.1	0.046	0.011
1619l	1.1599993	56560.2644	82.7	0.14	0.385	0.061	0.67	1.2	3.6	−2.2	−2.3	0.704	0.153
R+g	0.0000017	0.0016	4.2	0.04	0.022	0.006	0.04	0.4	0.8	0.1	0.3	0.029	0.021
1521ct	1.1724964	56907.7345	83.2	0.17	0.282	0.093	0.92	1.2	2.2	−2.3	−2.9	0.528	0.373
R+g	0.0000013	0.0005	1.4	0.09	0.009	0.004	0.02	0.8	1.1	0.1	0.3	0.046	0.044
1621ax	1.0181525	56741.7492	83.1	0.17	0.401	0.029	0.57	2.9	2.2	−2.2	−2.2	0.758	0.171
R+g	0.0000044	0.0017	4.8	0.04	0.028	0.006	0.07	0.9	1.1	0.0	0.1	0.031	0.027

Table A4 – *continued*

ID band	P (d)	t_0 (BMJD$_\text{tdb}$)	i (°)	q	$\langle r_1 \rangle$	$\langle r_2 \rangle$	T_2/T_1	absorb$_R$	absorb$_{g'}$	$\log(f_R)$	$\log(f_{g'})$	Fill	ρ ($\rho_\odot$)
1521cm	0.6854774	56068.9363	80.0	0.22	0.381	0.110	0.78	0.7	0.8	−2.9	−2.6	0.740	0.428
R+g	0.0000002	0.0002	1.0	0.04	0.008	0.003	0.02	0.3	0.2	0.3	0.2	0.025	0.027
1622by	0.7486683	55718.0157	85.6	0.16	0.424	0.076	0.81	3.5	3.1	−2.4	−2.9	0.791	0.270
R+g	0.0000015	0.0010	3.6	0.03	0.017	0.005	0.04	0.9	0.8	0.3	0.3	0.024	0.025
1522cc	0.5717853	56641.9749	81.2	0.18	0.472	0.078	0.69	1.4	1.2	−2.4	−2.6	0.894	0.331
R+g	0.0000003	0.0003	3.1	0.03	0.013	0.004	0.04	0.3	0.3	0.1	0.2	0.018	0.023
1622aa	0.7661291	57136.1682	84.7	0.10	0.410	0.060	0.75	0.6	1.4	−2.4	−2.9	0.716	0.302
R+g	0.0000038	0.0014	4.0	0.03	0.022	0.006	0.05	0.5	0.7	0.2	0.3	0.026	0.038
1622bt	0.6884160	56746.2432	79.2	0.17	0.425	0.071	0.70	1.7	1.4	−2.3	−2.5	0.795	0.314
R+g	0.0000004	0.0003	2.0	0.02	0.012	0.003	0.03	0.4	0.3	0.0	0.2	0.017	0.023
1723aj	1.1088064	56733.1351	85.6	0.11	0.460	0.042	0.52	0.5	0.5	−2.4	−3.1	0.818	0.100
R+g	0.0000009	0.0004	3.3	0.02	0.011	0.003	0.05	0.2	0.2	0.0	0.2	0.022	0.006
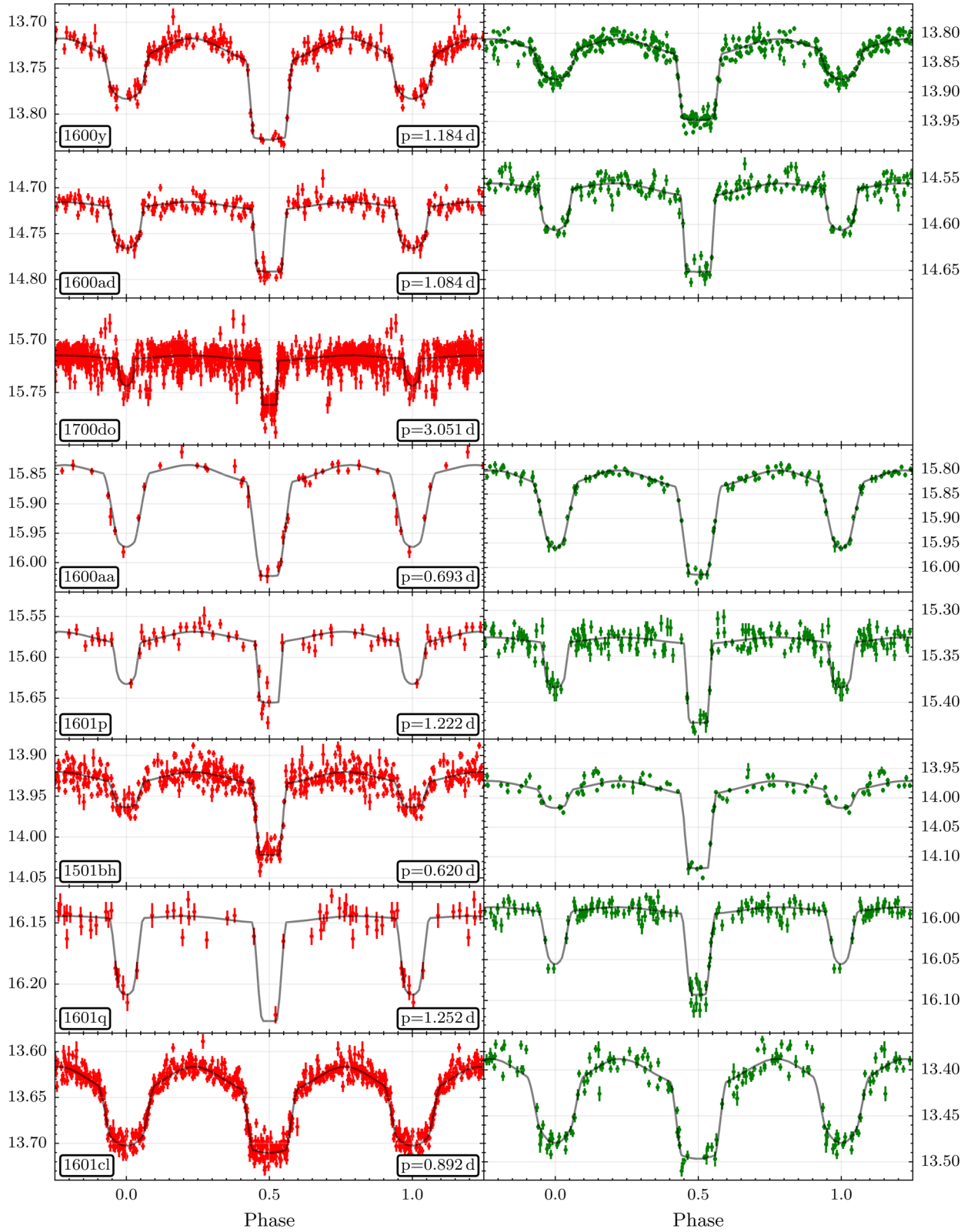
Figure A1. The PTF light curves in R (left) and g' (right) with the best model overplotted (see Table A4).
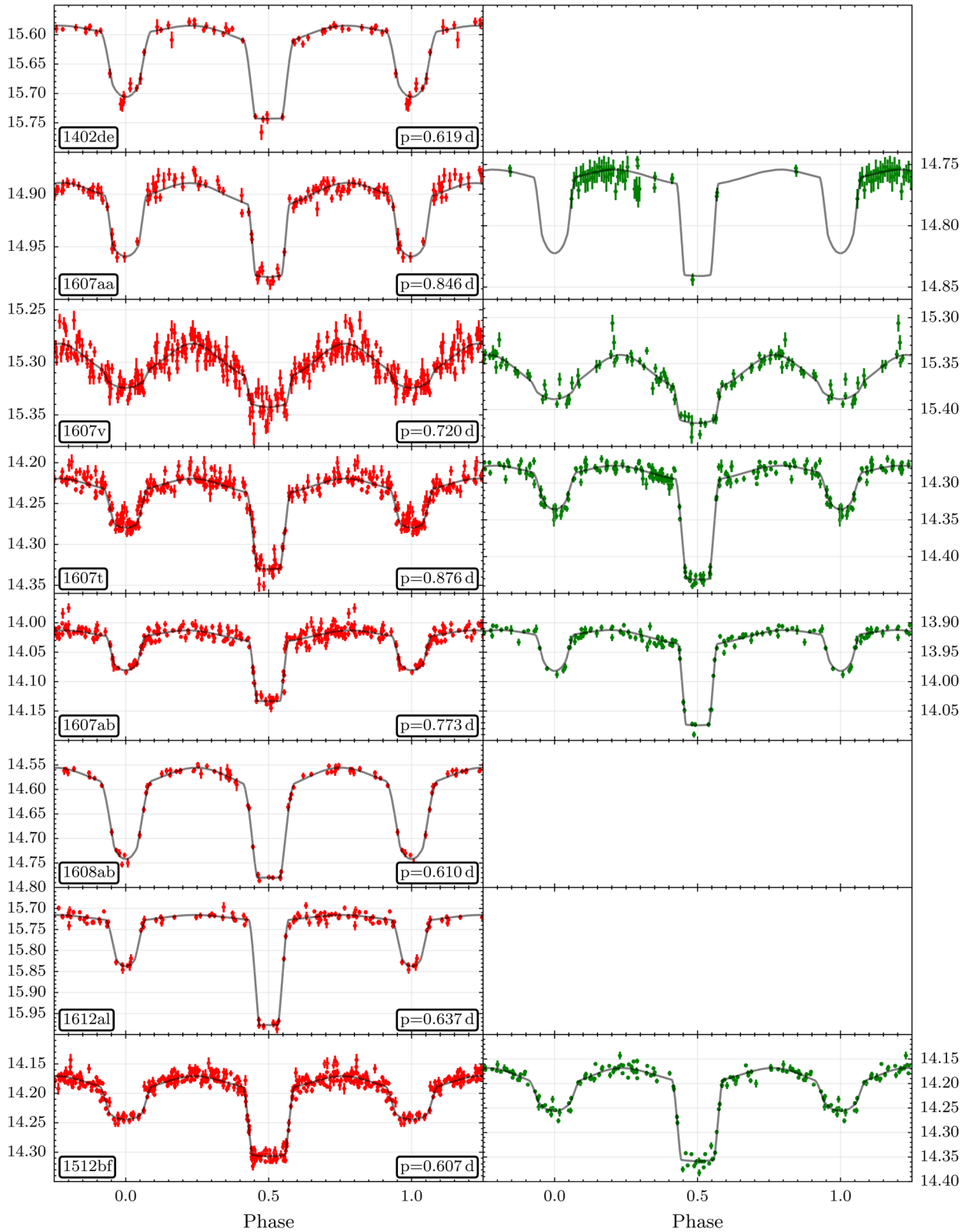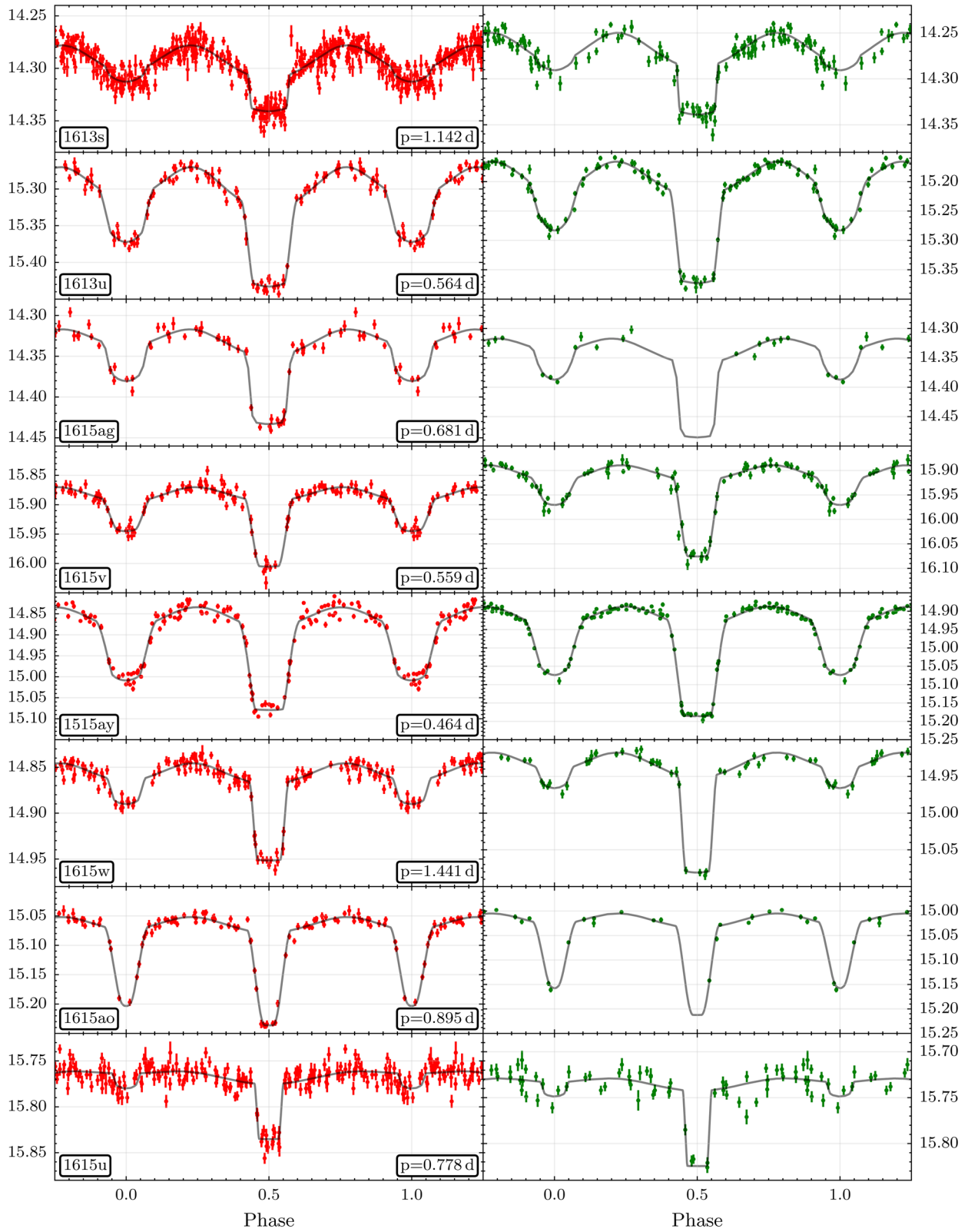
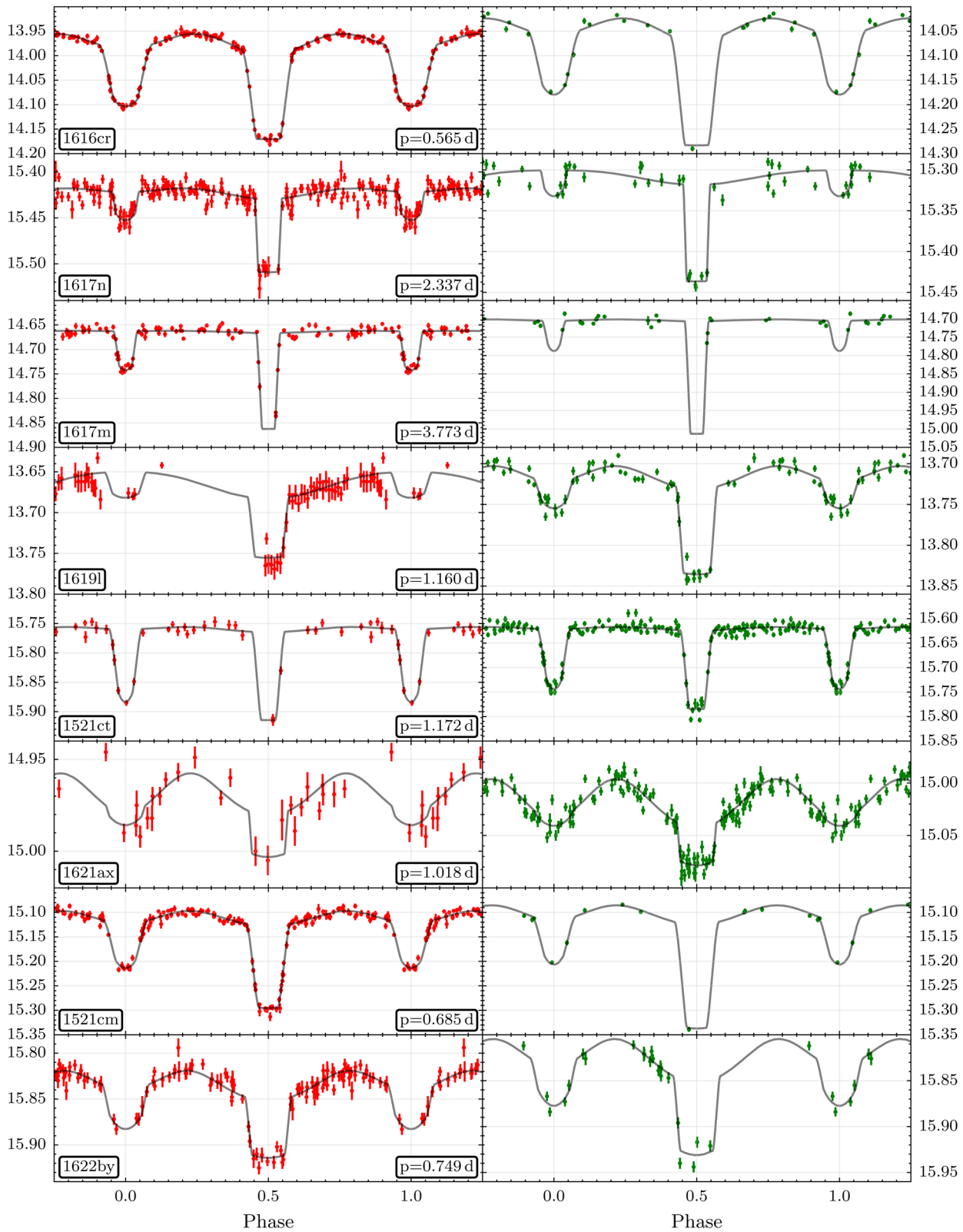**Figure A1** – *continued*

Figure A1 – *continued*
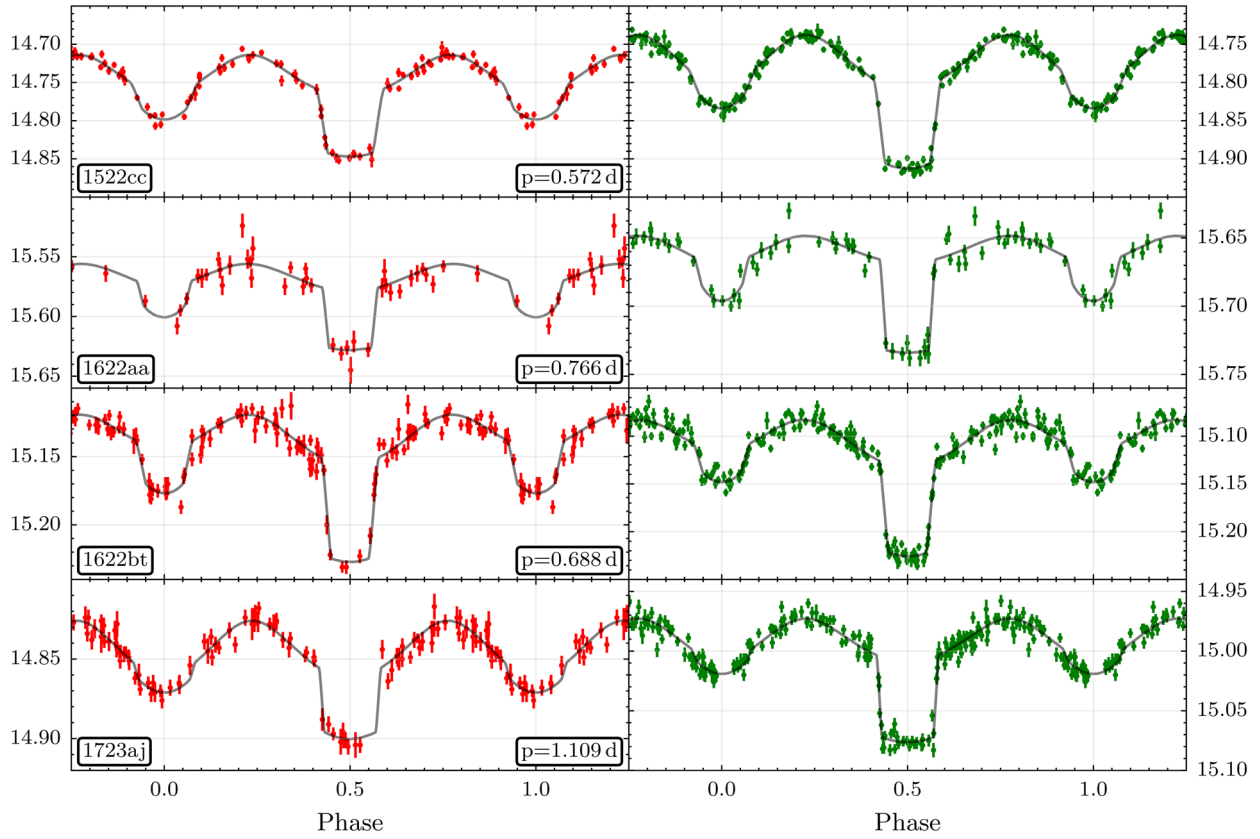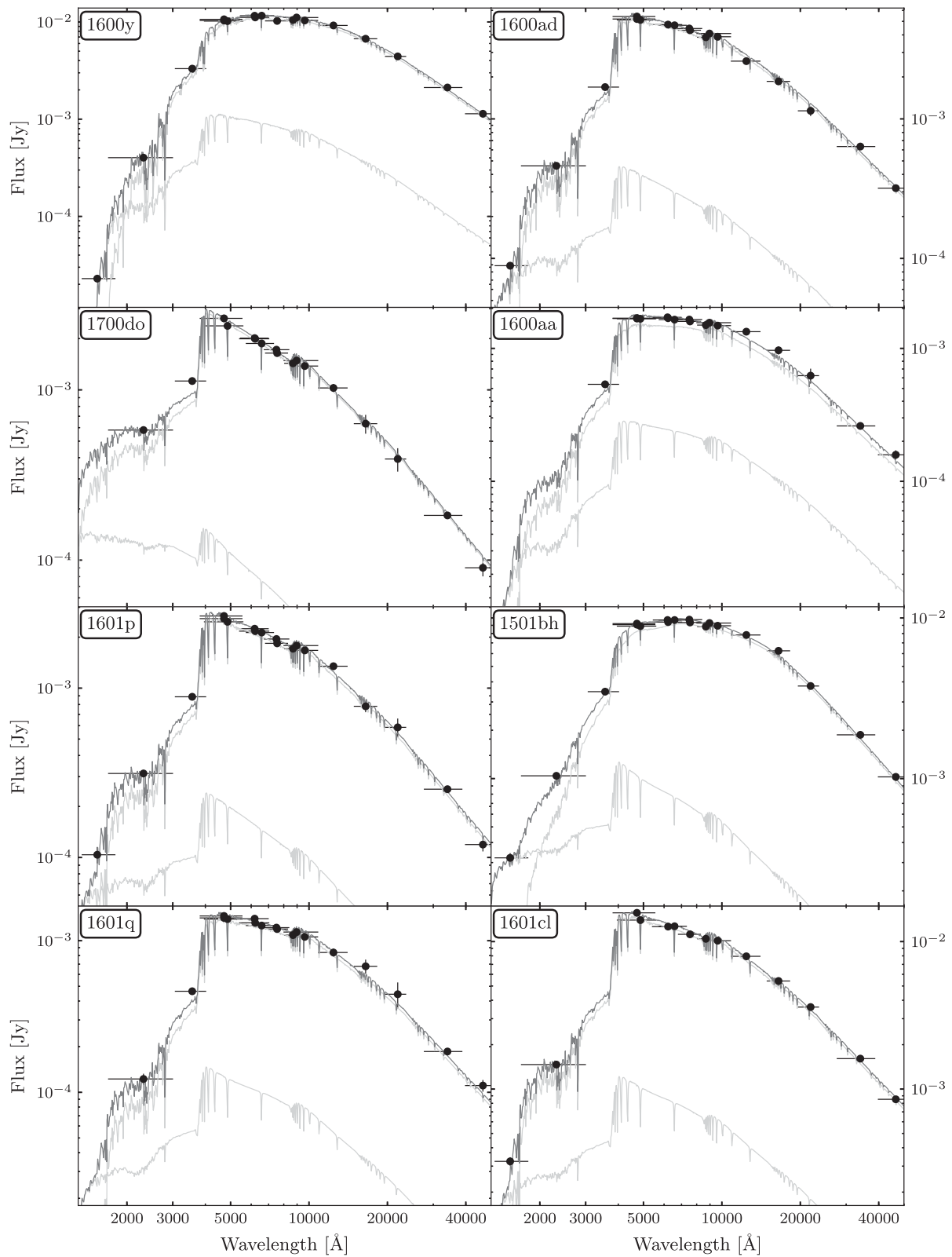
Figure A1 – *continued*

Figure A1 – *continued*

Figure A2. The SED and the best-fitting model spectra (see Table A3). The grey lines show the SED of the A/F star and pre-He-WD. The black line shows the sum of both components. The A/F star dominates the SED over the whole wavelength range, except in the far-UV in some of the cases.
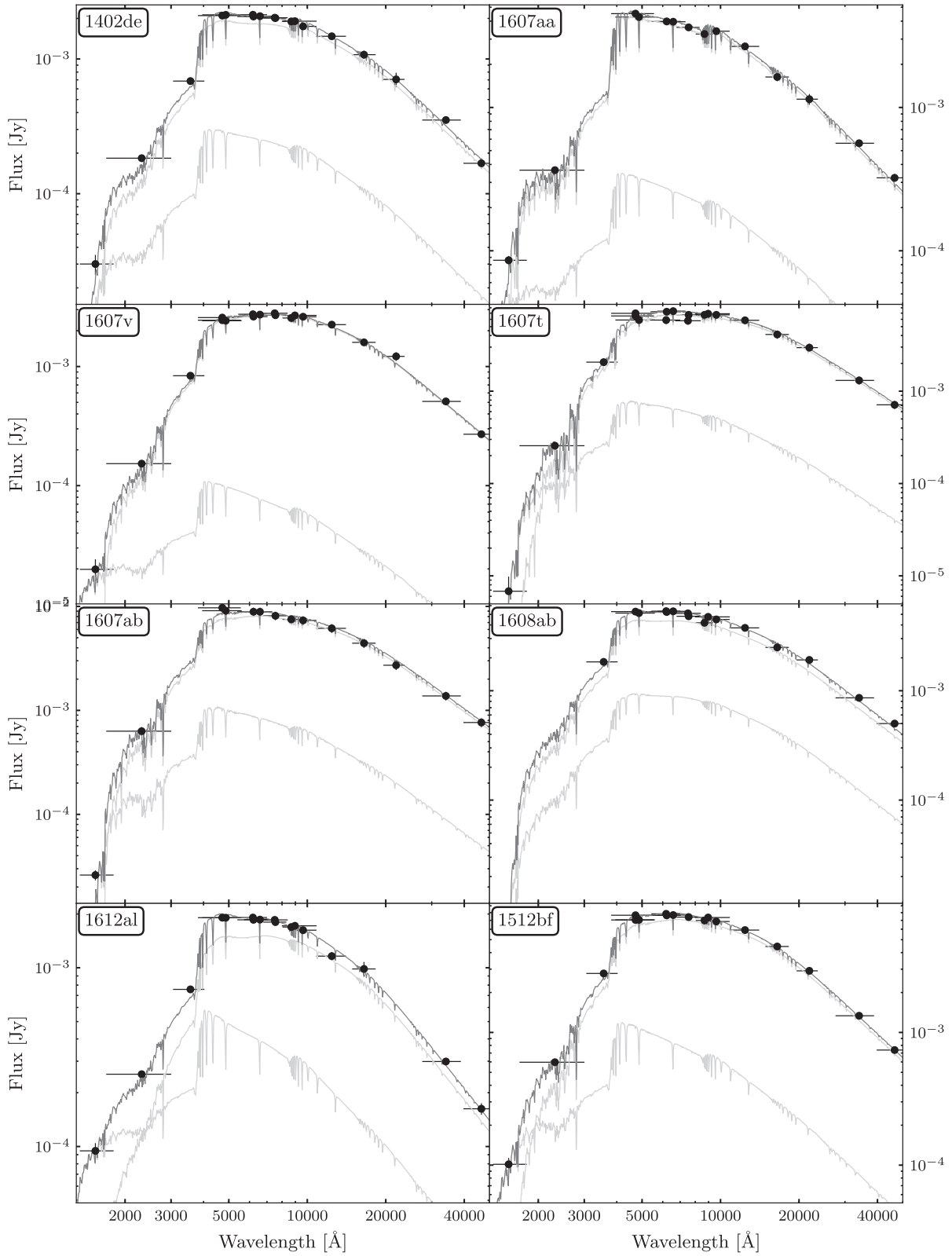
Figure A2 – *continued*
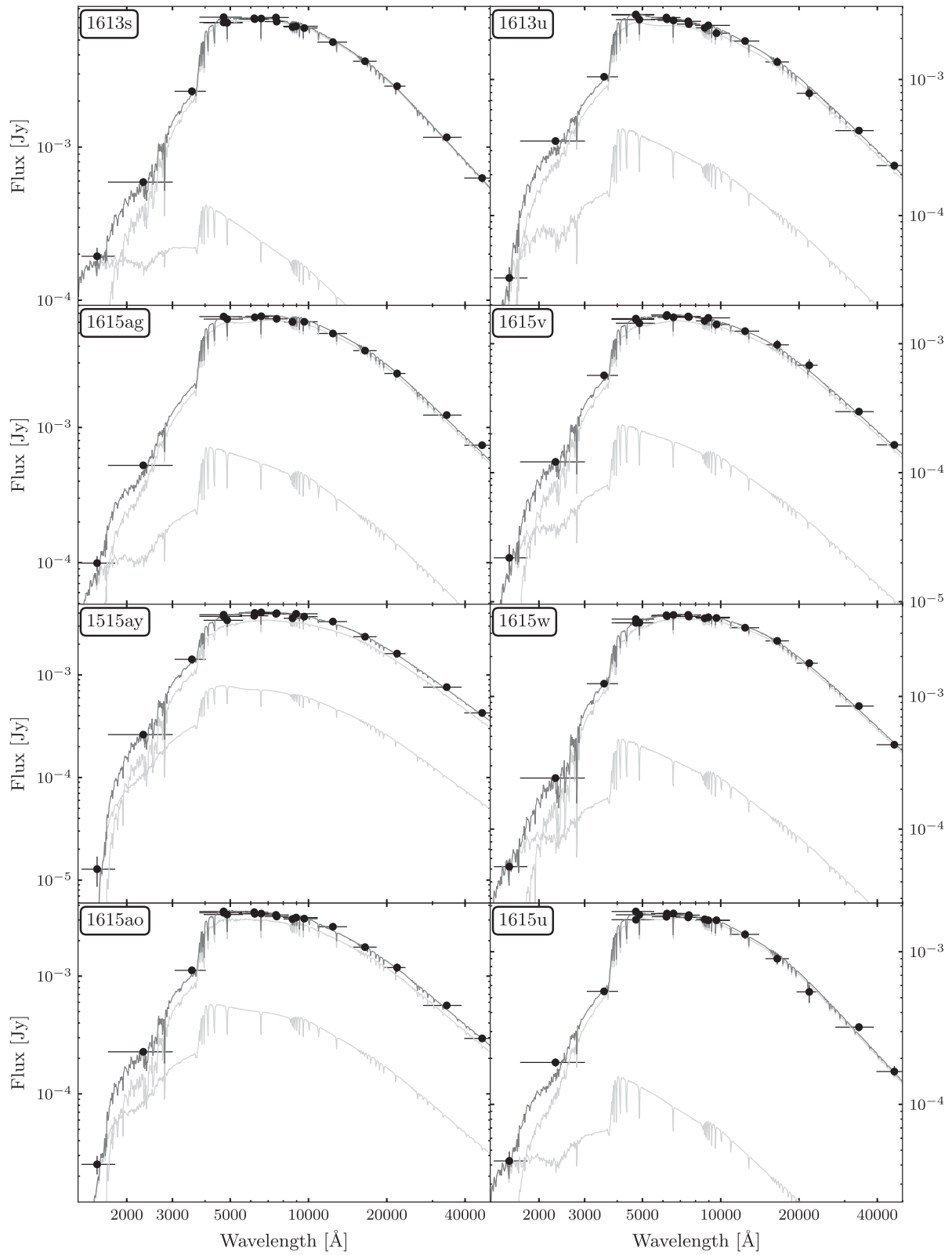
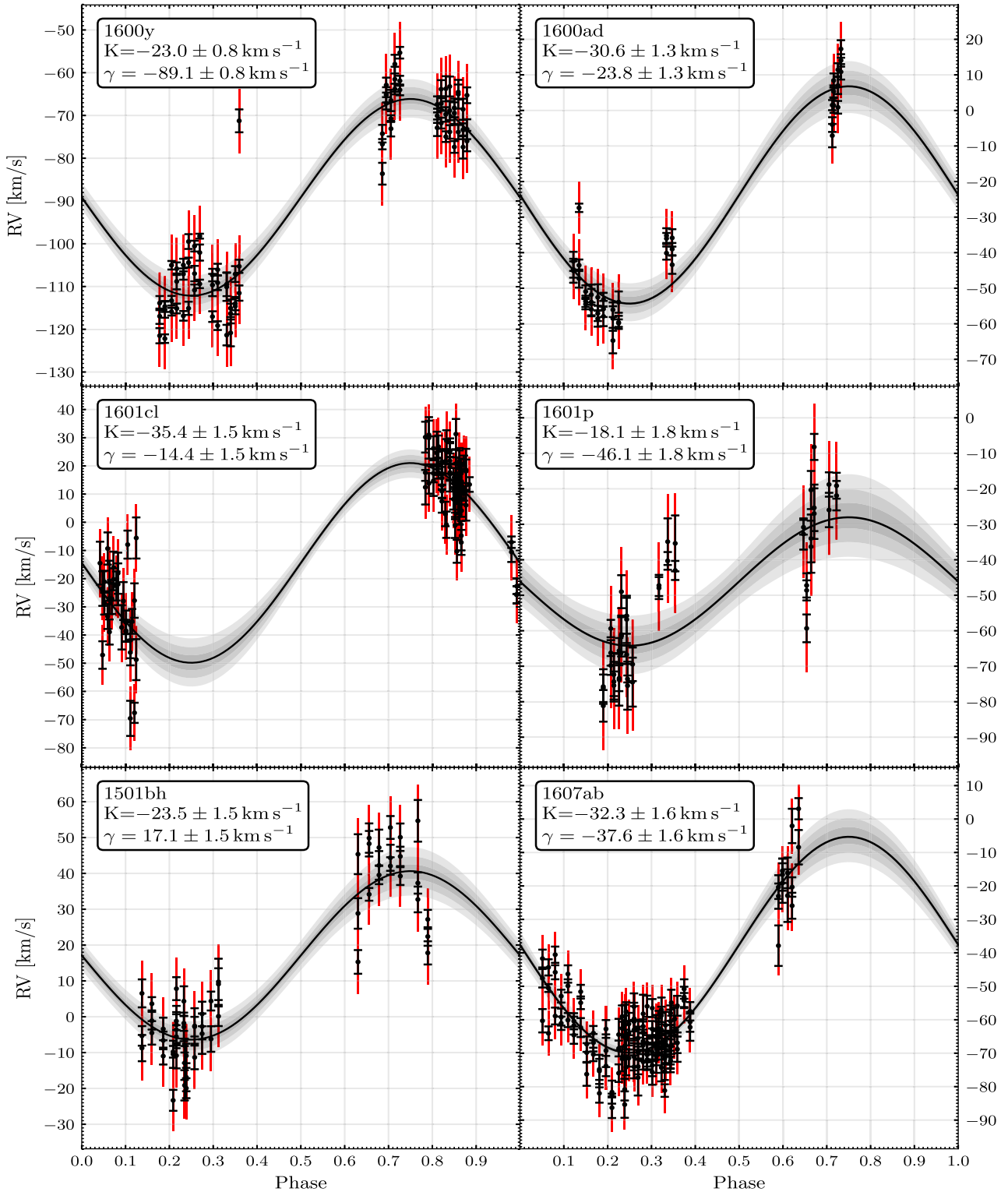**Figure A2** – *continued*
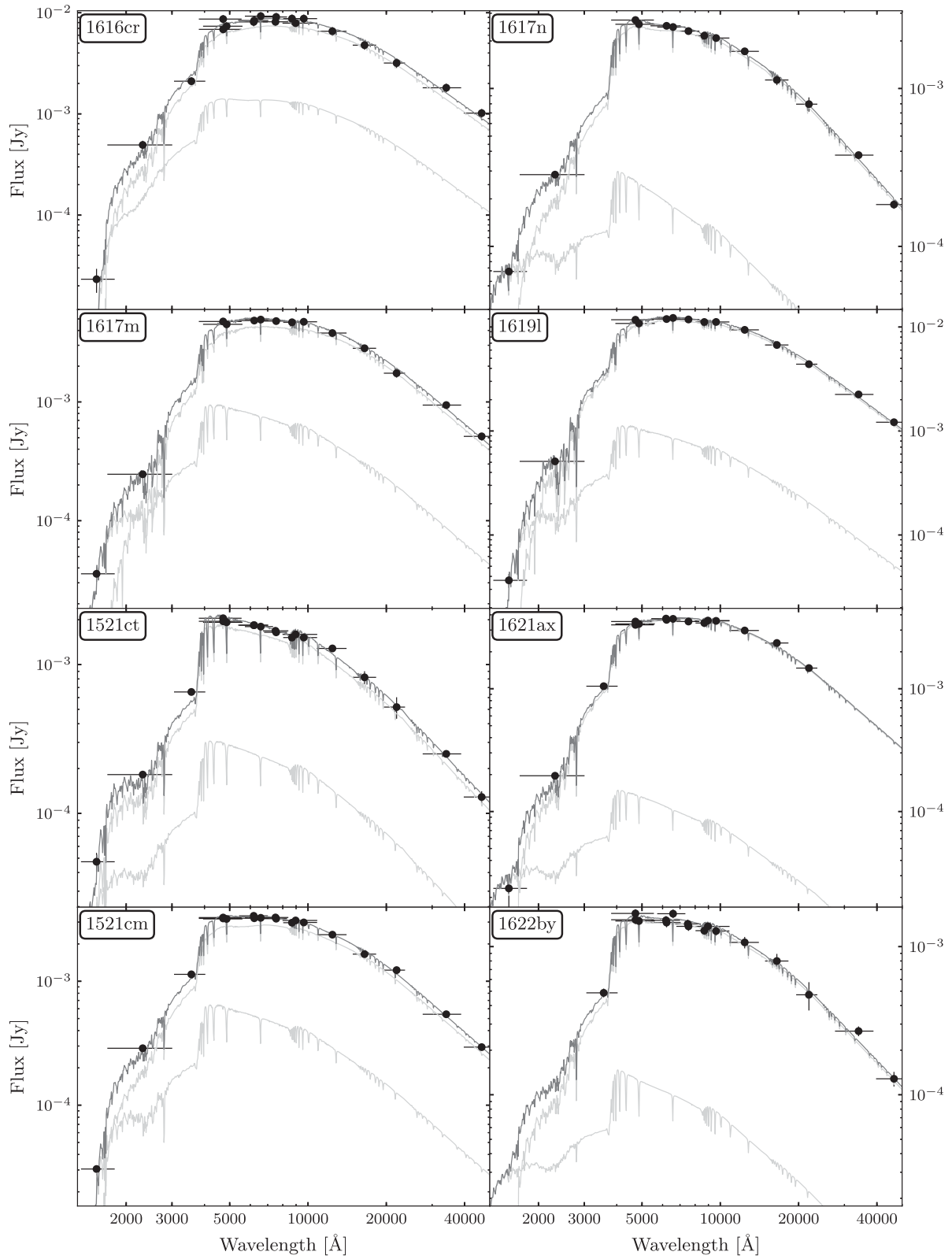
Figure A2 – *continued*
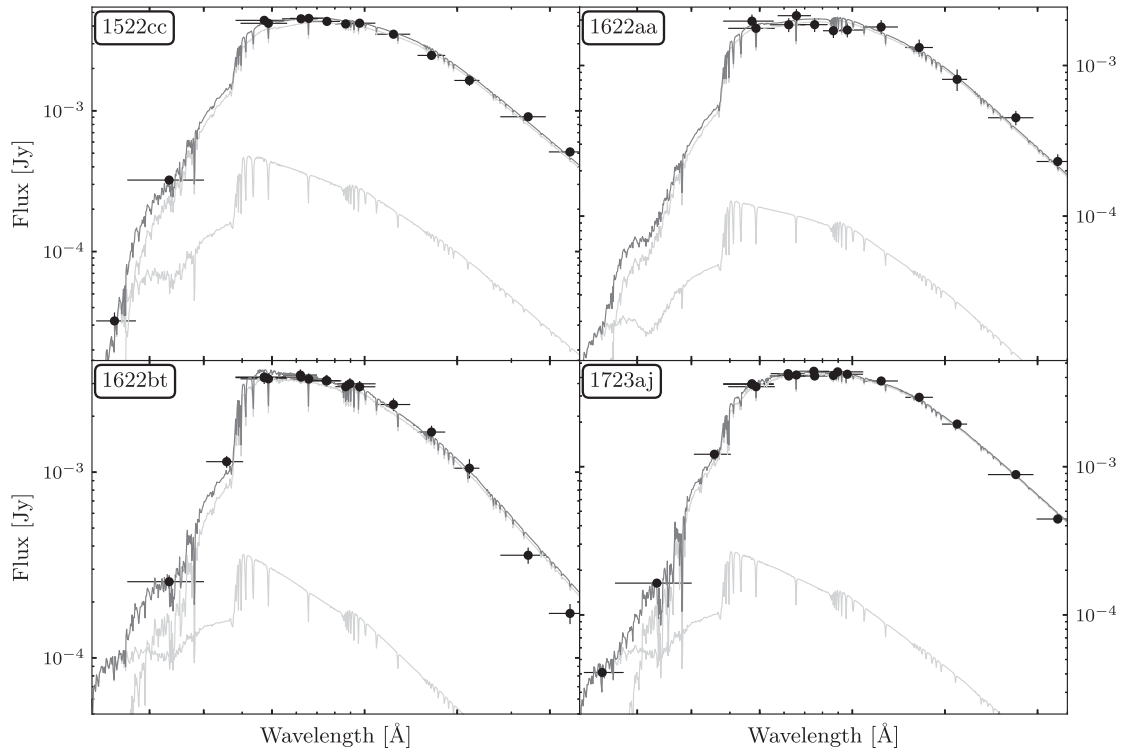
Figure A2 – *continued*

Figure A3. The radial velocity measurements with the INT and the best-fitting model. Black error bars show the estimated uncertainty from the cross-correlation procedure, while the red error bars show the uncertainties required to account for all residual variance. The shaded grey contours show the 1, 2 and 3 standard deviation intervals of model, obtained using the larger uncertainties.
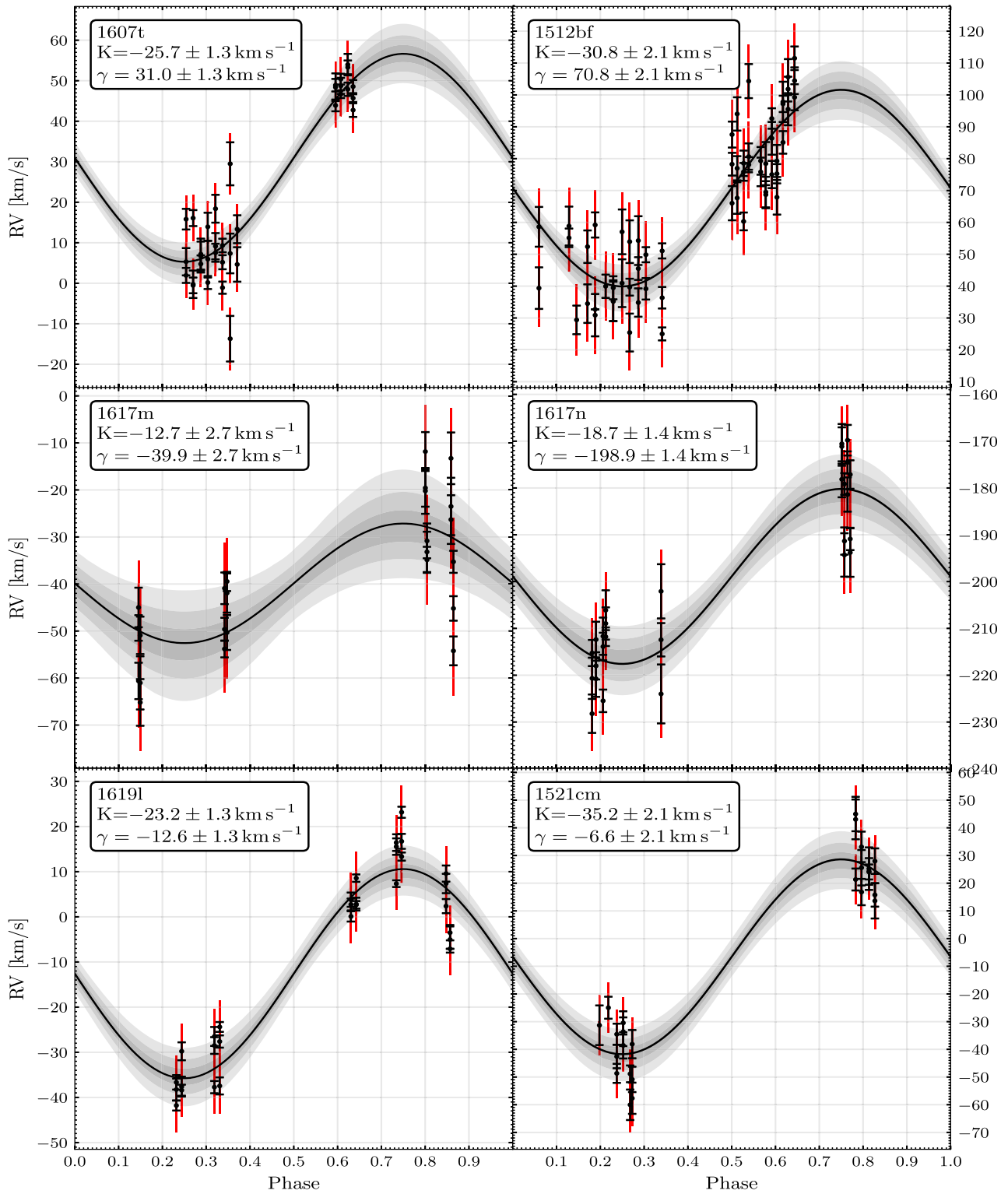
Figure A3 – *continued*

This paper has been typeset from a TEX/LATEX file prepared by the author.