The present and the future of modeling eclipsing binary systems

A. Prša[iD]

*Villanova University, Department of Astrophysics and Planetary Science, 800
Lancaster Ave, Villanova PA 19085, U.S.A. (E-mail: aprsa@villanova.edu)*

Abstract. In September 2024, eclipsing binary star practitioners gathered in
Litomyšl, Czech Republic, the birth town of Zdeněk Kopal, one of the most
celebrated pioneers of our field, to discuss the latest developments and state-
of-the-art. I was invited to present my own biased view of the present and
the future of modeling eclipsing binary stars. In this contribution I attempt
to make a clear distinction between approaches that are suited to individual
objects and approaches that aim to deliver bulk results for large datasets.
I stress that our motivation should be different: individual system analysis is
warranted whenever there is potential to propose or improve our understanding
of the underlying physics, while bulk analysis should be used to probe stellar
formation and evolution channels. I briefly discuss two examples of tools to
achieve the goals: PHOEBE for individual system analysis, and PHOEBAI for
bulk analysis.

Key words: binary stars – eclipsing binaries – fundamental parameters –
artificial intelligence

1. Introduction

Eclipsing binary stars (EBs) have long been recognized as *the* calibrators of stellar astrophysics. Most of what we know about the fundamental stellar properties stems from EB studies (Torres et al., 2010). Their favorable orbital alignment with the line of sight, and consequent eclipses, make them ideal astrophysical laboratories: a simple geometry coupled with well-understood dynamical laws allow us to obtain fundamental parameters without a-priori assumptions. Fig. 1 showcases the power of EBs: the masses, radii and luminosities of stars inferred from EB observations constitute the most accurate set of constraints for the theories of stellar formation and evolution. The uncertainties are generally below 2-3%, making the error-bars in Fig. 1 smaller than the symbols.

Binary systems do not discriminate on the spectral type or luminosity class; be it main sequence stars, red giants, or compact objects, be it flaring M-dwarfs or pulsating variables, be it tight mass-transferring systems or exoplanets, all these components are found in binary stars. Once we find them in an eclipsing
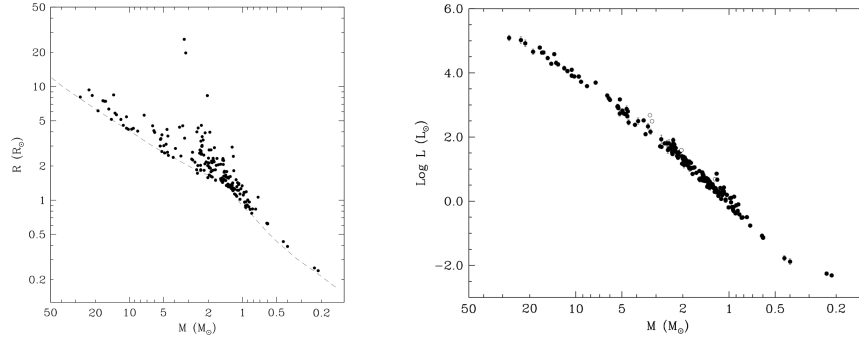
Figure 1. Radius-mass and luminosity-mass relationships for a sample of 94 EBs (Torres et al., 2010). The dashed line is the theoretical zero-age main sequence. Uncertainties in R, L and M are smaller than symbol sizes. Depicted with open symbols are stars classified as giants.

system, the path to fundamental parameters is straight-forward. For obvious reasons, colleagues who study these objects rely on masses and radii provided from EB analyses, which earns EB practitioners an occasional bottle of wine for their efforts. Further augmenting the importance of binaries is the observation that around 20% of binaries have tertiary components (Orosz, 2015), including circumbinary planets (Welsh et al., 2015); when all three components undergo eclipses, the uncertainties in fundamental parameters can reach ∼0.1% (Carter et al., 2011). It is thus of no surprise that Henry Norris Russell, in 1948, pointed out EBs as "the Royal Road" to stellar astrophysics.

2. The present

The 21st has been marked by an unprecedented advancement in data quality. NASA's *Kepler* mission, for example, routinely achieved photometric precision of the order of few tens of ppm (Borucki et al., 2010). *Kepler* observed ∼150,000 stars in the 105 square degree patch of the sky for 4 years with a 30-min cadence (and ∼1,000 stars with 2-min cadence) with 92% temporal completeness. In effect, for the first time we obtained a near-uninterrupted, ultra-precise set of observations for a few thousand EBs as well (Kirk et al., 2016). *Kepler*'s successor, *TESS* (Ricker et al., 2015), continued the census with thousands of targeted EBs (Prša et al., 2022) and ∼150,000 EBs in full frame images across the sky. While the majority of *TESS* EBs are observed continuously for about a month at a time, the dataset still represents a glimpse into EB complexities that remained hidden from ground-based observatories before. The latest space survey, Gaia (Gaia Collaboration et al., 2021), provides a stereoscopic vision of the Galaxy with simultaneous photometric, radial velocity, and astrometric ob-

Table 1. A table of EB parameters that can be inferred from a specific combination of observables. "Cal. LC" stands for flux-calibrated light curve, "SB1" stands for a single-lined spectroscopic binary, and "SB2" for the double-lined spectroscopic binary. The (✓) symbol corresponds to quantities that can be inferred from some but not all light curves. Adapted from Wrona & Prša (2024).

Parameter	Cal. LC	Cal. LC+ distance	2+ Cal. LCs	2+ Cal. LCs+dist.	2+ Cal. LCs+SB1	2+ Cal. LCs+SB2
t_0	✓	✓	✓	✓	✓	✓
P_0	✓	✓	✓	✓	✓	✓
i	✓	✓	✓	✓	✓	✓
$e\cos\omega$	✓	✓	✓	✓	✓	✓
$e\sin\omega$	✓	✓	✓	✓	✓	✓
r_1+r_2	✓	✓	✓	✓	✓	✓
T_2/T_1	✓	✓	✓	✓	✓	✓
r_2/r_1	(✓)	(✓)	(✓)	(✓)	(✓)	✓
$q=M_2/M_1$	(✓)	(✓)	(✓)	(✓)	(✓)	✓
l_3	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)
$\dot{P}, \dot{\omega}, \ldots$	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)
L_1+L_2		✓		✓		✓
L_1, L_2		✓		✓		✓
T_1, T_2			✓	✓	✓	✓
v_γ					✓	✓
M_1, M_2						✓
R_1, R_2				✓		✓
a				✓		✓

servations, with impressive first EB results for 400,000 systems (Mowlavi et al., 2023).

This explosion in photometric precision and temporal coverage put EB models to the test. At this level of precision, the models could no longer reliably reproduce observations, in part because of the missing physics, and in part because of the lack of the required numerical precision. PHOEBE (Prša & Zwitter, 2005) is one of the many codes that exhibited these inadequacies; this prompted a code rewrite in 2016 to remediate the situation.

Irrespective of the choice for the utilized model, the new decade brought on new requirements for modeling. EB practitioners have long realized that – depending on the types of observables available – the parameter space is non-linear and severely degenerate. Table 1 provides rule-of-thumb information content for available types of observables: it informs us what parameters we can infer from what types of observations. Thus, in general, to infer fundamental parameters (masses, radii, temperatures and luminosities), we need multi-band photometric and double-lined spectroscopic data.

Further, gone are the days where "solving" an EB means running an optimizer that minimizes the sum of squares of the residuals between the data

and the model, providing (unrealistically small) formal errors and publishing the results based on poorly sampled data with large scatter. Instead, the typical solution process involves several steps (Conroy et al., 2020):

Solution estimation: in order to ascertain convergence for the optimizers, the starting point in the parameter space needs to be reasonably close to a solution[1]. This is a task for estimators: methods that quickly analyze the data and provide a cursory solution. An example is an estimate of $e\cos\omega$ from eclipse separation in phase, or $r_1/a + r_2/a$ from eclipse widths, or $e\sin\omega$ from the eclipse duration ratios.

Solution optimization: once close, the solution needs to converge to a minimum as robustly as possible. This is usually achieved by finding the parameters that minimize the sum of squares of the residuals. Optimization is an iterative process that will typically follow the locally steepest slope towards the minimum. Once there, it will estimate formal errors from the covariance matrix. optimizers differential corrections (Wilson & Devinney, 1971), Nelder and Mead's simplex method (Kallrath & Linnell, 1987), and Powell's direction set method (Prša & Zwitter, 2007).

Solution sampling: the minimum reached must not be construed as the final solution: the topology of the parameter space near that minimum needs to be properly explored. That is a task for samplers: methods that traverse the vicinity of the parameter space and evaluate the probabilities in order to infer realistic parameter values. This is a computationally expensive and arduous process as parameter space needs to be traversed heuristically. Typical samplers include Markov Chain Monte Carlo methods (Foreman-Mackey et al., 2019), differential evolution (Storn & Price, 1997), etc. Inference is usually done within the Bayesian formalism, and the results are typically presented in a corner plot, with individual parameter posterior probability density functions and two-dimensional parameter cross-correlations (Conroy et al., 2020).

These components need to be utilized in unison in order to derive a compelling, credible solution for EB parameters given the data. The time cost of a typical modeling process is weeks to months, depending on the complexity of the parameter space, types of observations, and data quality and quantity.

3. The future

The 21st has also been marked by an unprecedented advancement in data quantity. Table 2 provides a rough census of observed or expected numbers of EBs in recent, currently ongoing, and imminent surveys.

---

[1] We deliberately stress *a* solution rather than *the* solution because, in most circumstances, there will not be a single, unequivocal model that uniquely explains the data.

Table 2. A census of recent, ongoing, and imminent surveys, along with their EB yield.

Survey	Data type	Years of operation	EB yield
K2	photometry	2014-18	∼1,000
Kepler	photometry	2009-13	∼3,000
NGTS	photometry	2015–	∼5,000
OGLE-II	photometry	1997-2000	∼6,000
Plato	photometry	2026?	∼8,000
ASAS	photometry	1997-2010	∼10,000
Galah	spectroscopy	2013–	∼15,000
OGLE-III	photometry	2001-09	∼40,000
ASAS-SN	photometry	2014–	∼40,000
TESS	photometry	2018–	∼150,000
OGLE-IV	photometry	2010–	∼500,000
ZTF	multi-band photometry	2018–	∼500,000
Gaia	photometry, astrometry, RVs	2013–	∼1.8 million
CSST	photometry, spectroscopy	2026?	∼2 million
LSST	multi-band photometry	2025–	∼10 million

It seems clear that, over the next ten years or so, we will have well over ten million EBs in our hands. In the previous section we have established that, typically, we need ∼weeks of computer time to robustly analyze EB observables. Thus, unless we are willing to accept that EB modeling remains a boutique[2] operation, we need faster approaches.

One such approach utilizes artificial intelligence (AI). We have been inundated with AI in our everyday lives, and for good reason. It powers technologies like virtual assistants, personalized recommendations, and smart home devices, making daily tasks more efficient and tailored to individual needs. In healthcare, AI enables early diagnosis through advanced imaging and predictive analytics (Secinaro et al., 2021), while in education, it supports personalized learning experiences (Chen et al., 2020). AI also enhances industries like transportation with autonomous vehicles and navigation systems, and it revolutionizes customer service through chatbots and automated support (Mohamad Suhaili et al., 2021).

AI has been proposed for EBs as well, for automated classification (see, for example, Čokina et al. 2021) and for bulk analysis (e.g., Prša et al. 2008). The appeal is undeniable: if AI could "look at" observables, recognize them as EBs, and deduce principal parameters in a fraction of a second, it would allow us to rapidly analyze large datasets. The question, then, is how accurately can this be done? Prša et al. (2008) trained a backpropagating neural network on syn-

---

[2]The task of solving 10 million EBs in the next 100 years would require ∼4000 astronomers and ∼8 million computer cores, along with a moratorium on further observations.

thetic light curves generated by PHOEBE (Prša & Zwitter, 2005; Prša et al., 2016) where the network was shown ∼32,000 light curves with 5 corresponding parameters: temperature ratio (T_2/T_1), sum of fractional radii $(R_1/a+R_2/a)$, radial eccentricity $(e \sin \omega)$, tangential eccentricity $(e \cos \omega)$, and inclination $(\sin i)$. They demonstrated that the trained network could attain the accuracy better than 10% for 90% of the sample. This allowed the authors to deploy the network on 2500+ EBs from OGLE and CALEB databases.

The principal problem with this type of approach is that, in the absence of post-festum tests, we do not know *which* 10% are not correctly characterized. Given that the neural network maps input (light curves) to output (parameters) opaquely, not only do we lack the ability to evaluate per-light curve accuracy, we also cannot attribute meaningful uncertainties to any obtained parameters. The remedy is to run post-festum tests, but these can be both computationally and temporally expensive, which defeats the initial purpose at least to some extent.

There exists an alternative, though: we can exchange the neural network's inputs and outputs. Instead of providing light curves as input, we provide parameters; instead of collecting parameters as output, we collect light curves. The network becomes an *emulator* of the physical engine. Assuming that the network can be trained adequately to encompass the non-linear properties of our complex, multivariate parameter space, each computation that takes minutes to complete using the physical model can take milliseconds using a trained network. Thus, the AI-powered engine becomes a drop-in replacement for the physical model. In our particular case, PHOEBAI replaces PHOEBE (cf. Fig. 2).
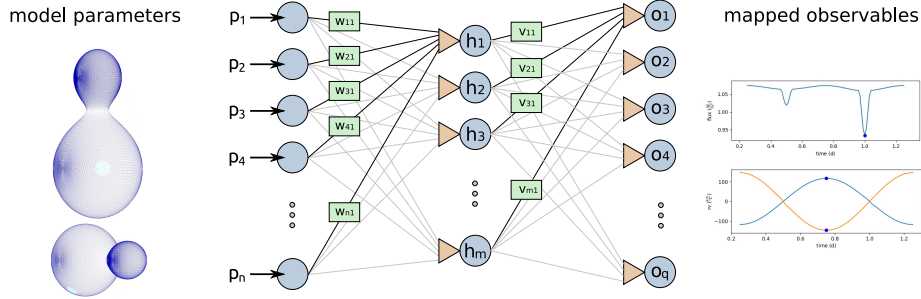

Figure 2. A schematic representation of the feed-forward neural network that acts as an emulator to the physical engine. Parameters p_i are mapped through connection weights w_{ij} across hidden layers h_j to the output units o_k, representing synthetic observables.

Let us first estimate the potential speed-up and establish a good motive for this effort. In Section 2 we estimate that, for the completion of the estimation, optimization, and sampling stages, we require $\sim 10^6$ forward model computa-

tions. If each model takes ~ 2 minutes to compute , we are talking about 3.8 *years* of compute time on a single processor core. To handle this volume of computations in a reasonable time (say, 2 weeks), we need 100 cores – i.e., a computer cluster. For the neural network, on the other hand, the model can be "computed" within a millisecond. In that case, the net compute time becomes ~ 1.5 hours *on a single processor*. Thus, AI holds promise of a $\sim 10^6$-fold speed-up.

There are certainly limitations to such an emulator, even if the underlying network is perfectly trained.

– The inputs and the outputs are fixed, determined by the network topology. If the network it trained with, say, 10 input parameters and it outputs, say, 500 fluxes in equidistant phases, we are restricted to using the same 10 parameters and the emulator will yield fluxes in those exact 500 phase points. The observations will thus need to be interpolated to the corresponding phase array for comparison with the model, and any other parameters present in the physical model will not play a role in the emulator.

– The network emulates phased light curves, so there can be no temporal variation in the data. Any variability that is not captured by the parameters used for training the network will inevitably skew the results to the extent of the degeneracy between included and omitted parameters. Any trends in the data, or more complex noise models, need to be accounted for either before (i.e., by detrending) or concurrently with sampling (i.e., by modifying the maximized probability function).

– Neural networks cannot be used on input that deviates from the training set: while they are well suited for interpolation, they are notoriously bad at extrapolation (Freeman & Skapura, 1991). Thus, if the training set does not cover the parameter space adequately, sampling will not be done correctly. It also means that, if the *density* of the covered parameter space is not representative of actual distributions, the results may be biased or suffer from undersampling systematics.

– The parameter space is inherently non-linear and highly degenerate. The choice of input parameters is thus crucial for emulator performance and result fidelity. Should any of the principal parameters be omitted, or if the chosen parameters are not orthogonal, the emulator will be misspecified and the results will be biased.

These limitations can be controlled with a careful optimization of the neural network topology and the construction of the training set. Once optimized and trained, the emulator is ready to be used as a drop-in replacement of the physical model. That includes the optimization stage and the sampling stage: given the data, the emulator is able to minimize the cost function and to sample the

probability density. It is able to do so in a $\sim$millionth of the physical model's time cost.

As proof of concept, we trained an emulator on 6 photometric parameters: the effective temperature ratio T_2/T_1, eccentricity e, argument of periastron ω, orbital inclination i, the sum of fractional radii $r_1 + r_2$, and their ratio r_2/r_1. For efficiency, we replaced e and ω with $e\sin\omega$ and $e\cos\omega$, and i with $\cos i$. The parameters were sampled from distributions that cover a wide enough range to encompass the case study light curves. The network was trained on $\sim$600k light curves synthesized using PHOEBE. As outputs, the network used phase-folded light curves sampled in 500 points. We denote this emulator PHOEBAI: PHOEBE via AI (Wrona & Prša, 2024).

The topology of the network was determined by the `RandomizedSearchCV` method (Bergstra & Bengio, 2012) from the `sklearn` library (Pedregosa et al., 2011). The method evaluated thousands of structures with varying numbers of hidden layers, nodes, and activation functions. The chosen shape was the simplest among the best performing in terms of minimizing the sum of residual squares between predicted and actual sets of light curves. It comprises four hidden layers with 512, 512, 512, and 1024 nodes, with the sigmoid activation function.

Application of this network was tested on a subset of EBs from the *TESS* EB catalog (Prša et al., 2022) that matched our training set distributions. The optimizer utilized differential evolution (Storn & Price, 1997) from the `scipy.optimize` library (Virtanen et al., 2020). The final results from this optimization were subsequently used as starting values for MCMC sampling using the emcee Python library (Foreman-Mackey et al., 2019).

During the sampling process, we employed 80 walkers and 4000 iterations to obtain posterior distributions. Fig. 3 compares PHOEBAI and PHOEBE performance on a detached EB, TIC 279097693. This is a typical (rather than cherry-picked) result from the test sample: it shows that PHOEBAI is able to accurately capture both the posteriors and parameter correlations.

4. Conclusions

As we continue to be inundated with high-quality data, we need to distinguish between "boutique" operation and bulk processing. When our understanding of the physical processes can be advanced by pointed investigations of individual objects, the well-established methods and codes remain indispensable. On the other hand, to understand large sample properties and parameter distributions that probe the outcomes of stellar formation and evolution principles, we need to rely on novel methods. One such method, presented in this paper, employs feed-forward neural networks. It is important to realize that it is not the only way or even the preferred way to model large swaths of data, but – as demonstrated here – PHOEBAI certainly shows promise in delivering robust results on par with
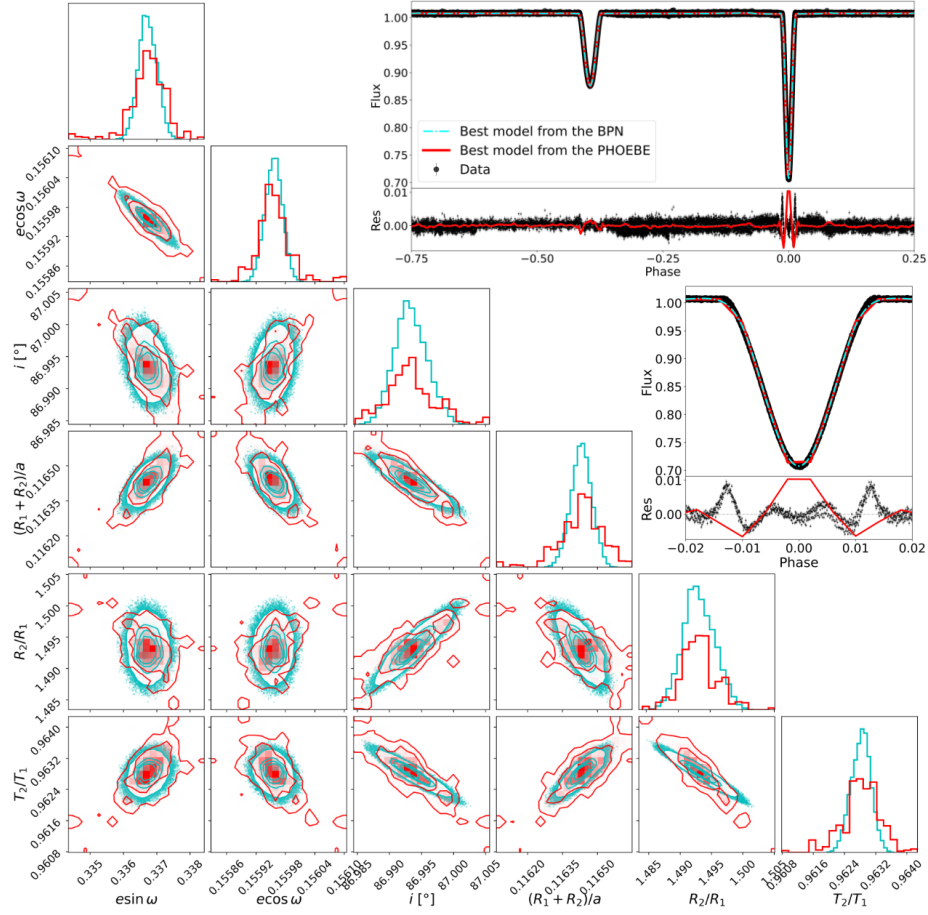
Figure 3. Sampling results for TIC 279097963, a detached EB observed by *TESS* in Sector 4. The top right panel depicts a phase-folded light curve with the best–fit PHOEBE model (red) and the best-fit PHOEBAI model (cyan), along with the residuals, plotted over data. The panel below zooms in on the primary eclipse, demonstrating that the residuals are due to the coarse phase sampling. The corner plot compares parameter posteriors (diagonal) and 2-D correlations (off-diagonal) for the 6 sampled parameters with PHOEBE (red) and PHOEBAI (cyan). PHOEBAI posteriors are better sampled but slightly narrower. Overall, there is a remarkable agreement between the two approaches, where PHOEBAI performed ∼250,000 times faster than the PHOEBE sampler.

the physical engines. Above all, it remains vital to understand the limitations and ranges of application for both the physics-based models and the AI-based emulators and to use them appropriately.

References

Bergstra, J. & Bengio, Y., Random Search for Hyper-Parameter Optimization. 2012, *J. Mach. Learn. Res.*, **13**, 281–305

Borucki, W. J., Koch, D., Basri, G., et al., Kepler Planet-Detection Mission: Introduction and First Results. 2010, *Science*, **327**, 977, DOI:10.1126/science.1185402

Carter, J. A., Fabrycky, D. C., Ragozzine, D., et al., KOI-126: A Triply Eclipsing Hierarchical Triple with Two Low-Mass Stars. 2011, *Science*, **331**, 562, DOI:10.1126/science.1201274

Chen, L., Chen, P., & Lin, Z., Artificial Intelligence in Education: A Review. 2020, *IEEE Access*, **8**, 75264, DOI:10.1109/ACCESS.2020.2988510

Conroy, K. E., Kochoska, A., Hey, D., et al., Physics of Eclipsing Binaries. V. General Framework for Solving the Inverse Problem. 2020, *Astrophysical Journal, Supplement*, **250**, 34, DOI:10.3847/1538-4365/abb4e2

Foreman-Mackey, D., Farr, W. M., Sinha, M., et al. 2019, emcee v3: A Python ensemble sampling toolkit for affine-invariant MCMC, Zenodo

Freeman, J. A. & Skapura, D. M. 1991, *Neural networks: algorithms, applications, and programming techniques* (Redwood City, CA, USA: Addison Wesley Longman Publishing Co., Inc.), ISBN: 0-201-51376-5

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al., Gaia Early Data Release 3. Summary of the contents and survey properties. 2021, *Astronomy and Astrophysics*, **649**, A1, DOI:10.1051/0004-6361/202039657

Kallrath, J. & Linnell, A. P., A New Method to Optimize Parameters in Solutions of Eclipsing Binary Light Curves. 1987, *Astrophysical Journal*, **313**, 346, DOI:10.1086/164971

Kirk, B., Conroy, K., Prša, A., et al., Kepler Eclipsing Binary Stars. VII. The Catalog of Eclipsing Binaries Found in the Entire Kepler Data Set. 2016, *Astronomical Journal*, **151**, 68, DOI:10.3847/0004-6256/151/3/68

Mohamad Suhaili, S., Salim, N., & Jambli, M. N., Service chatbots: A systematic review. 2021, *Expert Systems with Applications*, **184**, 115461, DOI:10.1016/j.eswa.2021.115461

Mowlavi, N., Holl, B., Lecoeur-Taïbi, I., et al., Gaia Data Release 3. The first Gaia catalogue of eclipsing-binary candidates. 2023, *Astronomy and Astrophysics*, **674**, A16, DOI:10.1051/0004-6361/202245330

Orosz, J. A., Triple Stars Observed by Kepler. 2015, in Astronomical Society of the Pacific Conference Series, Vol. **496**, *Living Together: Planets, Host Stars and Binaries*, ed. S. M. Rucinski, G. Torres, & M. Zejda, 55

Pedregosa, F., Varoquaux, G., Gramfort, A., et al., Scikit-learn: Machine Learning in Python. 2011, *Journal of Machine Learning Research*, **12**, 2825

Prša, A., Conroy, K. E., Horvat, M., et al., Physics Of Eclipsing Binaries. II. Toward the Increased Model Fidelity. 2016, *Astrophysical Journal, Supplement*, **227**, 29, DOI:10.3847/1538-4365/227/2/29

Prša, A., Guinan, E. F., Devinney, E. J., et al., Artificial Intelligence Approach to the Determination of Physical Properties of Eclipsing Binaries. I. The EBAI Project. 2008, *Astrophysical Journal*, **687**, 542, DOI:10.1086/591783

Prša, A. & Zwitter, T., A Computational Guide to Physics of Eclipsing Binaries. I. Demonstrations and Perspectives. 2005, *Astrophysical Journal*, **628**, 426, DOI:10.1086/430591

Prša, A., Kochoska, A., Conroy, K. E., et al., TESS Eclipsing Binary Stars. I. Short-cadence Observations of 4584 Eclipsing Binaries in Sectors 1-26. 2022, *Astrophysical Journal, Supplement*, **258**, 16, DOI:10.3847/1538-4365/ac324a

Prša, A. & Zwitter, T., Introducing Powell's Direction Set Method to a Fully Automated Analysis of Eclipsing Binary Stars. 2007, in Astronomical Society of the Pacific Conference Series, Vol. **370**, *Solar and Stellar Physics Through Eclipses*, ed. O. Demircan, S. O. Selam, & B. Albayrak, 175

Ricker, G. R., Winn, J. N., Vanderspek, R., et al., Transiting Exoplanet Survey Satellite (TESS). 2015, *Journal of Astronomical Telescopes, Instruments, and Systems*, **1**, 014003, DOI:10.1117/1.JATIS.1.1.014003

Russell, H. N., The Royal Road of Eclipses. 1948, **7**, 181

Secinaro, S., Calandra, D., & Secinaro, A. e. a., The role of artificial intelligence in healthcare: a structured literature review. 2021, *BMC Med Inform Decis Mak*, **21**, 125, DOI:10.1186/s12911-021-01488-9

Storn, R. & Price, K., Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. 1997, *J. of Global Optimization*, **11**, 341–359, DOI:10.1023/A:1008202821328

Torres, G., Andersen, J., & Giménez, A., Accurate masses and radii of normal stars: modern results and applications. 2010, *Astronomy and Astrophysics Reviews*, **18**, 67, DOI:10.1007/s00159-009-0025-1

Čokina, M., Maslej-Krešňáková, V., Butka, P., & Parimucha, Š., Automatic classification of eclipsing binary stars using deep learning methods. 2021, *Astronomy and Computing*, **36**, 100488, DOI:10.1016/j.ascom.2021.100488

Virtanen, P., Gommers, R., Oliphant, T. E., et al., SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. 2020, *Nature Methods*, **17**, 261, DOI:10.1038/s41592-019-0686-2

Welsh, W. F., Orosz, J. A., Short, D. R., et al., Kepler 453 b - The 10th Kepler Transiting Circumbinary Planet. 2015, *Astrophysical Journal*, **809**, 26, DOI:10.1088/0004-637X/809/1/26

Wilson, R. E. & Devinney, E. J., Realization of Accurate Close-Binary Light Curves: Application to MR Cygni. 1971, *Astrophysical Journal*, **166**, 605, DOI:10.1086/150986

Wrona, M. & Prša, A., The Eclipsing Binaries via Artificial Intelligence. II. Need for Speed in PHOEBE Forward Models. 2024, *arXiv e-prints*, arXiv:2412.11837