Contrib. Astron. Obs. Skalnaté Pleso **55/3**, 370 – 373, (2025)
https://doi.org/10.31577/caosp.2025.55.3.370

Prediction of photometric parameters of overcontact eclipsing binaries using deep-learning model

Š. Parimucha[1], **P. Gajdoš**[1,2], **Y. Markus**[1] and **V. Kudak**[3]

[1] *Institute of Physics, Faculty of Science, Šafárik University, Košice, Slovakia*
[2] *Astronomical Institute of the Czech Academy of Sciences 251 65 Ondřejov, The Czech Republic*
[3] *Laboratory of Space Researches, Uzhhorod National University, Uzhhorod, Ukraine*

Received: November 14, 2024; Accepted: December 13, 2024

Abstract. This paper presents a deep-learning model for predicting the photometric parameters of W UMa-type eclipsing binaries without spots. The model assumes that the light curve (LCs) can be described by four parameters: the orbital inclination, photometric mass ratio, temperature ratio, and common potential. The training dataset comprises 500 000 simulated LCs in the Gaia G passband. The best results were obtained using the random forest predictor, achieving a Mean Absolute Percentage Error (MAPE) of 6.1%. The study concluded that the quality of parameter prediction strongly depends on the quality of the analyzed LCs curves, requiring careful data preprocessing.

Key words: eclipsing binaries – machine-learning – deep-learning

1. Introduction

Eclipsing binary stars of the W UMa type (overcontact binaries) are a type of close binary system in which both stars are so close that they share a common envelope of material. Their shapes are strongly distorted by tidal forces and the frame of Roche geometry allows us to describe their common surface by its potential (Prša, 2018). These systems are among the most frequently found variable stars because of their short orbital period (up to 1 d) and typical light-curve (LC) shape. Space and ground-based surveys discovered several tens of thousands of such systems; however, only a small fraction of them determined their photometric parameters. These parameters can be obtained from the solutions of the LCs. However, several software packages dedicated to LCs solutions for eclipsing binaries exist, such as the PHOEBE (Conroy et al., 2020), JKTEBOP (Southworth et al., 2004), and ELISa (Čokina et al., 2021), this procedure is not straightforward. It requires strong interaction with the user and a qualified estimate of the initial parameters. This paper presents a deep-learning model for predicting the photometric parameters of W UMa-type eclipsing binaries that are unaffected by spots. This prediction model assumes that the LC can be described by four parameters: the orbital inclination i, photometric mass ratio q, the temperature ratio t_1/t_2 and common potential Ω. These parameters determine the shapes of stars, their relative dimensions, and their radiation properties.

2. Training and evaluation dataset

The training dataset was created using the ELISa code (Čokina et al., 2021). The LCs of the overcontact binaries were simulated from parameters covering a wide range of physically correct star values, using a random uniform distribution in the intervals used (Tab. 1). In total, 500 000 LCs were created in the Gaia G passband.

Table 1. Parameters and intervals used for the simulated LCs.

Parameter	Interval
i	40-90°
q	0.05-2.0
Ω	2.0-4.0
t_1/t_2	1.0-1.3

To evaluate our model, we randomly selected 100 LCs of a set of overcontact eclipsing binaries from the Gaia catalogue of eclipsing binaries (Mowlavi et al., 2023), as classified by Parimucha et al. (2024). All LCs were manually fitted using the ELISa code to determine the basic parameters of the systems. The Gaia LCs were preprocessed (outlier removal), rebinned to 100 data points, folded, and normalized to the median flux.

3. Deep learning model and its performance

To train our models, we used all the simulated LCs and the dataset was expanded using Gaussian noise augmentation. The 20% of them were randomly selected for validation of the trained model. We tested several models using different machine-learning algorithms including gradient boosting, random forest, linear, quadratic, and logistic regressions, as well as several convolution networks. The quality of the model was tested on the validation dataset and quantified by the MAPE (Mean Absolute Percentage Error) parameter. The best results were obtained by the random forest predictor, which achieved a MAPE value of 6.1%. Examples of predicted and fitted LCs are shown in Fig. 1.

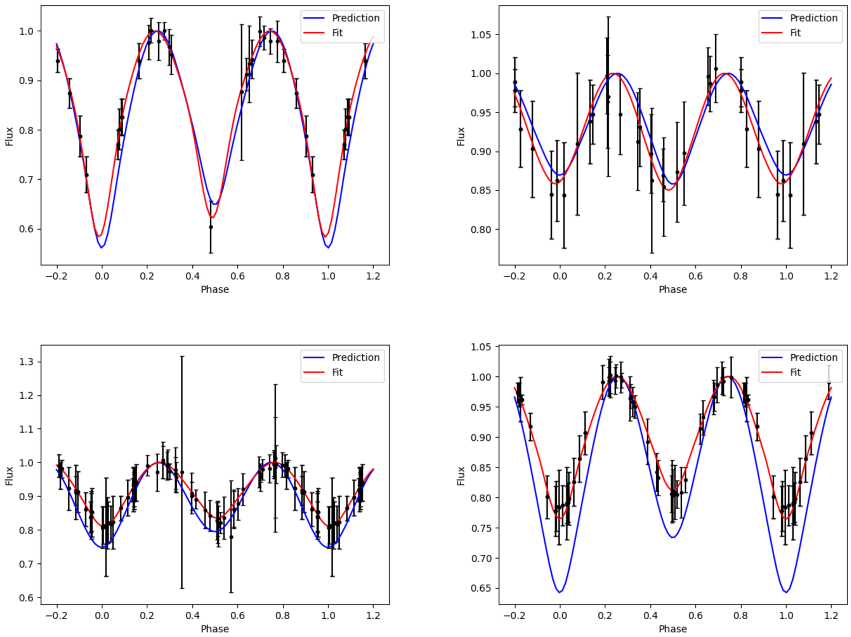

Figure 1. The examples of the predicted and fitted LCs of four overcontact eclipsing binaries from the Gaia catalogue.

4. Discussions

Our study leads to several conclusions. Most importantly, the quality of prediction of the parameters strongly depends on the quality of the analyzed LCs and careful data preprocessing is required. We require good data coverage, remove outliers, and ensure the curve is phase-folded well. The prediction is sensitive mainly to outliers, which is evident for curves with few data points. We also note that the space of the simulated LCs must be sufficiently dense, and a random uniform distribution of parameters appears to be a good solution. The parameters predicted by our model are good starting points for analysis using other methods such as LSQ and/or MCMC fitting to obtain precise results with parameter errors. Subsequent analysis of the residua can reveal bad classifications and/or LC anomalies such as spots and pulsations.

Acknowledgements. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (`https://www.cosmos.esa.int/gaia`). This research was supported by the Slovak Research and Development Agency under Contract No. APVV-20-0148. P.G. was supported by internal grant No. VVGS-2023-2784 of the

P. J. Šafárik University in Košice and funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I03-03-V05-00008 and The Ministry of Education, Youth and Sports of the Czech Republic by grant LTT-20015. Y.M. was funded by the EU NextGenerationEU through The Recovery and Resilience Plan for Slovakia under project No. 09I03-03-V01-00011 and grant UPJŠ VVGS 2024-3096 V.K. were supported by the Ukrainian national grant 0122U000937.

References

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

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

Parimucha, Š., Gajdoš, P., Markus, Y., & Kudak, V., Deep-learning classification of eclipsing binaries. 2024, *Contributions of the Astronomical Observatory Skalnaté Pleso*, **54**, 167, DOI:10.31577/caosp.2024.54.2.167

Prša, A. 2018, *Modeling and Analysis of Eclipsing Binary Stars*, 2514-3433 (IOP Publishing), DOI:10.1088/978-0-7503-1287-5, ISBN: 978-0-7503-1287-5

Southworth, J., Maxted, P. F. L., & Smalley, B., Eclipsing binaries in open clusters - II. V453 Cyg in NGC 6871. 2004, *Monthly Notices of the RAS*, **351**, 1277, DOI:10.1111/j.1365-2966.2004.07871.x

Čokina, M., Fedurco, M., & Parimucha, Š., ELISa: A new tool for fast modelling of eclipsing binaries. 2021, *Astronomy and Astrophysics*, **652**, A156, DOI:10.1051/0004-6361/202039171