OPEN EUROPEAN JOURNAL ON VARIABLE STARS
http://oejv.physics.muni.cz
January 2023 DOI: 10.5817/OEJV2020-XXX ISSN 1801–5964

NEW FEATURES IN SOFTWARE OCFIT

Gajdoš, Pavol[1]

1) Institute of Physics, Faculty of Science, Pavol Jozef Šafárik University, Košice, Slovakia,
pavol.gajdos@upjs.sk

Abstract: Software OCFit is designed for fitting O-C diagrams of eclipsing binaries and transiting exoplanets by standard analytical models (e.g. light-time effect or apsidal motion). The program is written in the modern language Python. A very intuitive graphic user interface is also a part of this software. Moreover, OCFit could be used as a Python package in your own code. Its development started in 2015. GUI was published three years later. Here, we present changes and the newly implemented features in this software since its publishing in 2018 (mainly during the year 2022).

1 Introduction

The O-C diagram as the relation between differences of observed and calculated times of the minima of eclipsing binaries (EBs) and the epoch of the minima is a relatively simple but extremely valuable tool. From an analysis of such a diagram, many exciting processes in the studied system (e.g. mass transfer, apsidal motion, or presence of the third body) could be detected.

However, the exact physical models of most effects presented on the O-C diagrams are strongly non-polynomial. Fitting these models to real data is not a simple task and therefore finding suitable software to analyse O-C diagrams is hardly possible.

Consequently, the development of the own software called OCFit for fitting O-C diagrams of EBs and transiting exoplanets was started in 2015. Code is fully written in a modern programmatic language python using standard scientific packages, e.g. numpy (Harris et al., 2020) and matplotlib (Hunter, 2007). The Genetic Algorithms (GA) and Markov chain Monte Carlo (MCMC) method are used for fitting O-C diagrams. These methods are designed also for fitting similar complicated non-polynomial functions. Using GA does not require any initial values of fitting parameters, which minimizes the user influence on the results. Only their intervals are needed. However, their size could be relatively large and it is easy to guess them. The software OCFit implements simple linear and quadratic models of O-C diagrams together with models of the light-time effect (Irwin, 1952), apsidal motion (Giménez & Bastero, 1995), and some mutual effects in multi-planetary systems (Agol et al., 2005). The program consists of two parts – python package which could be separately used by anyone in his/her code and a graphical user interface (GUI) for simplifying the analysing of O-C diagrams.

The separate python package was introduced at the conference KOLOS (The international meeting about variable stars research) in 2016. The main features of GUI were presented in 2018 at the KOLOS conference and also at the 50$^{\text{th}}$ Conference on Variable Star Research in Brno. Detail description of the code, the used models and the GUI is

OPEN EUROPEAN JOURNAL ON VARIABLE STARS
http://oejv.physics.muni.cz
January 2023 DOI: 10.5817/OEJV2020-XXX ISSN 1801–5964

given in our paper Gajdoš & Parimucha (2019). The source code is fully available on the GitHub web page[1] and also on the PyPi repository[2], which enable its easy installation inside the PYTHON environment using the command "pip install OCFit".

Over the years, many changes in the code and GUI were made. Some new features were implemented and some parts of the code were improved. This paper presents the most important changes in the OCFIT package (Sec. 2) and belonging GUI (Sec. 3) in the last five years.

2 Code changes and new functions

Many of the changes in the own PYTHON package are not visible to the standard users (mainly if they work only with GUI). However, there are all the more important (including fixed bugs). A few of them reflect the progress and innovations in PYTHON language and its packages. The first version of OCFIT was written in PYTHON 2.7 which support finished in 2020. Switching to new PYTHON 3 required many code-style changes. The latest version of OCFIT supports only PYTHON 3. It was successfully tested on various platforms (WINDOWS or UNIX/LINUX) with different subversion of PYTHON (from 3.6 to 3.10).

Regarding the end of support of PYTHON 2.7, the package used for MCMC fitting also had to be changed. The previously used package PYMC 2.3.8 is only partially supported under PYTHON 3. There were many complications during its building and installation, mainly under WINDOWS. Moreover, installation on the latest versions of PYTHON starting from 3.10 is impossible. Although there are newer versions of PYMC (version 3 and also 4), their implementation is significantly different and it is not a simple update from the used old version. Therefore, it did not matter whether the new version of PYMC or a completely different package would be used. The package EMCEE (Foreman-Mackey et al., 2013) was selected for the MCMC fitting. This package is widely used and well-supported. Its implementation is relatively simple and straightforward without requiring massive changes over different versions of the package. In addition, this package provides better and more detailed control of the fitting process.

From the early versions of OCFIT, storing important data (input data, values of models parameters, etc.) in a file is possible using the function Save (or button SaveClass in GUI). The saved file with all values could be reloaded and used, e.g. for creating new figures. The standard option for similar tasks in PYTHON is the package PICKLE which was already used in the previous versions of the software. The data are saved to a binary file with a very small disk size. On the other hand, one can raise some problems loading PICKLE files coming from different versions of PYTHON (2 vs 3) and/or OS (mainly WINDOWS and LINUX) but they could be prevented in many cases. Although, the binary nature of these files makes their inspection or modification outside PYTHON environment impossible. Mainly due to the possibility of analysing and conceivably also fixing some issues, the data are saved into a common JSON file in the latest version of this program.

[1] https://github.com/pavolgaj/OCFit
[2] https://pypi.org/project/OCFit/

January 2023 DOI: 10.5817/OEJV2020-XXX ISSN 1801–5964

These files are human-readable and could be opened in any text editor. The backward compatibility is included by the automatic loading of old PICKLE files and the option to save data also in the old format. Moreover, a few additional parameters (e.g. mass and an orbital inclination of EB) are stored and thus do not need to be entered repeatedly.

Furthermore, some new functions were implemented in the OCFIT package. Probably, the most interesting one is the possibility of setting the phase of secondary minima. This feature is especially important for studying O-C diagrams of eccentric eclipsing binaries (details in Sec. 2.1).

The new method for the initial fitting of the O-C diagram was included from SCIPY package (Virtanen et al., 2020). This method called differential evolution (DE; Storn & Price, 1997) is similar to already used genetic algorithms. It could be faster than GA in some cases.

The list of indirect parameters in the summary was extended to a few new ones - mainly parameters related to parabolic trend on the O-C diagram and mass transfer (e.g. $\dot{P}$, $\dot{P}/P$ and $\dot{M}$). One new model of O-Cs ("ApsidalQuad") was added. It combines a model of apsidal motion with a parabolic trend.

In the new version of this software, more different plots can be generated to analyse the process of MCMC (and also GA or DE) fitting calling function InfoMCMC (or InfoGA for GA or DE, respectively). For example, the corner plot (Foreman-Mackey, 2016) visualizes the final distribution of fitting parameters and the correlation between each other.

2.1 Eccentric orbits

It is generally known that the phase of secondary minima of EB with eccentric orbit differs from the value 0.5 valid for a circular orbit. In other words, the difference between epochs of primary and secondary minima is not exactly 0.5. Neglecting this fact could cause artificial features on O-C diagrams – mainly the separation of primary and secondary minima. Figure 1 shows the result of the wrong calculation of epochs of secondary minima. For this hypothetical EB with an orbital period of 15 days and eccentricity of only 0.05, the phase of secondary minima is 0.51. Neglecting eccentricity causes a shift between primary and secondary minima for about 4 hours! Assuming an eccentric orbit, this shift disappears.

The program OCFIT allows using the optional value of the difference between epochs of primary and secondary minima (marked dE in the software and ΔE in the following equations). This value should be given during class initialization or written in a relevant field in GUI. The default value (0.5) represents a circular orbit. This section describes how to obtain the correct value of this difference and how it is used in determining types of minima.

The user has three options to obtain the value of ΔE:

1. Already known value; e.g. from fitting light-curve as a phase of secondary minima.

2. Estimate it from the shift between the primary and secondary minima on the O-C diagram. This option is only implemented in GUI. However, this guess is simple. The difference between the mean level of O-Cs for secondary minima and primary ones

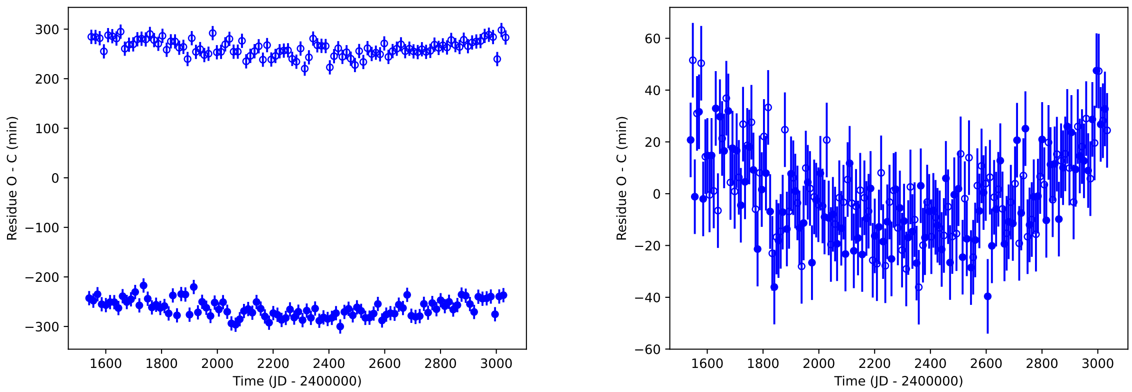

Figure 1: Example O-C diagram with neglected orbital eccentricity (*left*) and the same one with correctly calculated epochs (*right*).

divided by an orbital period gives the offset of ΔE from a value of 0.5. Although, this approach assumes that there is no real shift of mean levels (e.g. small part of the O-C diagram with apsidal motion) or asymmetry between different types of minima.

3. Calculate it from the known value of orbital eccentricity e and the argument of pericenter ω (marked w in code). This calculation implemented in the function DeltaEpoch is based on the following equations.

It can be shown these relations for true anomalies of primary (ν_0) and secondary minimum (ν_1):

$$\Delta\nu \equiv \nu_1 - \nu_0 = \pi \tag{1}$$

$$\nu_0 = \frac{\pi}{2} - \omega\,. \tag{2}$$

Then the difference in eccentric anomalies between minima is

$$\Delta\varepsilon \equiv \varepsilon_1 - \varepsilon_0 = -2\arctan\frac{\sqrt{1-e^2}}{e\cos\omega} \tag{3}$$

and the sum of eccentric anomalies in both minima is

$$\varepsilon_0 + \varepsilon_1 = -2\arctan\left(\sqrt{1-e^2}\tan\omega\right)\,. \tag{4}$$

These equations could be used for a difference in mean anomalies:

$$\Delta M \equiv M_1 - M_0 = \Delta\varepsilon - 2e\sin\frac{\Delta\varepsilon}{2}\cos\frac{\varepsilon_0+\varepsilon_1}{2} \tag{5}$$

and obtain

$$\Delta M = \Delta\varepsilon + \frac{2e\sqrt{1-e^2}\cos\omega}{1-e^2\sin^2\omega}\,. \tag{6}$$

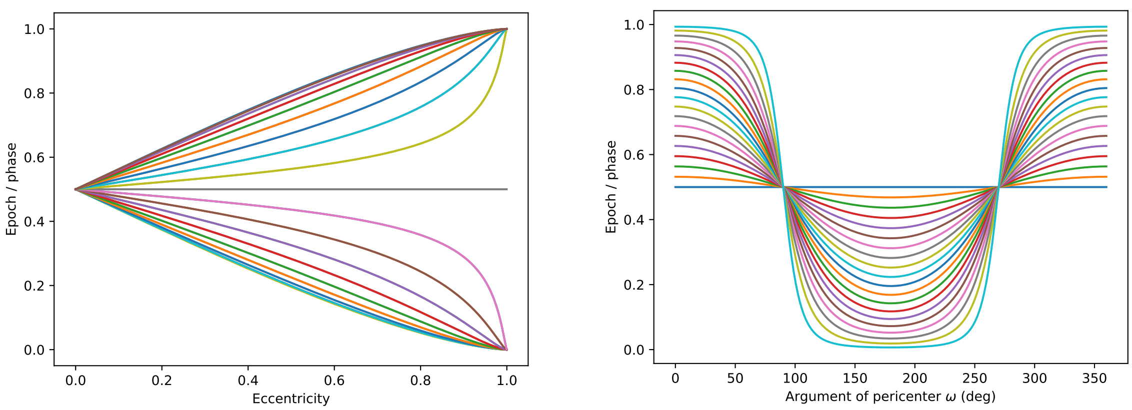

Figure 2: The difference in epoch (phase) of secondary minima as a function of eccentricity (*left*) and the argument of pericenter (*right*) for different values of the argument of pericenter or eccentricity.

Finally, the wanted difference in epochs is

$$\Delta E = \frac{\Delta M}{2\pi} \,. \tag{7}$$

Figure 2 shows the calculated values of ΔE for different values of eccentricity (from 0 to 1) and the argument of pericenter (from 0° to 360°). It is clear that it can acquire any value from 0 to 1.

Determining the type of observed minimum (at the time of observation t_O) and calculating its real epoch is straightforward if the reference time T_0, orbital period P and difference in epochs ΔE are known. Firstly, calculate the "observed epoch":

$$E_O = \frac{t_O - T_0}{P} \,. \tag{8}$$

It could have any value (i.e. also non-integer value for primary minimum) as it contains also observational effects – such as errors and trends on the O-C diagram. From it, the observed phase could be calculated:

$$\phi_O = E_O - \mathrm{round}(E_O) \tag{9}$$

where "round" means standard mathematical rounding. Thus calculated observed phase goes from -0.5 to 0.5. Now, the minimum is primary (i.e. $type = 0$) if

$$|\phi_O| \leq \min\left(|\phi_O - \Delta E|, |\phi_O - \Delta E + 1|\right) \,. \tag{10}$$

This condition means that the distance (in phase) to the primary minimum is smaller than to the closest secondary one. Otherwise, the minimum is secondary (i.e. $type = 1$). The exact epoch of minimum is

$$E = \mathrm{round}(E_O - type \times \Delta E) + type \times \Delta E \,. \tag{11}$$

OPEN EUROPEAN JOURNAL ON VARIABLE STARS
http://oejv.physics.muni.cz
January 2023 DOI: 10.5817/OEJV2020-XXX ISSN 1801–5964

This epoch could be used in standard calculation of the predicted time of minimum t_C by a well-known equation:

$$t_C = T_0 + P \times E\,. \tag{12}$$

The presented workflow is already implemented in the OCFit and used for calculation epochs by a function called Epoch.

3 Changes in GUI

Most of the changes in OCFit were performed in GUI as the most used part of the software – only minimum users adopt the basic package in their scripts. Many changes reflect modifications in the main package – a few new buttons and windows were added (see Fig. 3). Other ones are focused mainly on the better comfort of the users. For example, the size of some buttons was very tiny on big high-resolution displays (4K monitors). This issue is fixed now by the dynamic and proportional size of all windows, buttons, input boxes etc. User input is better handled and controlled to avoid possible mistakes.

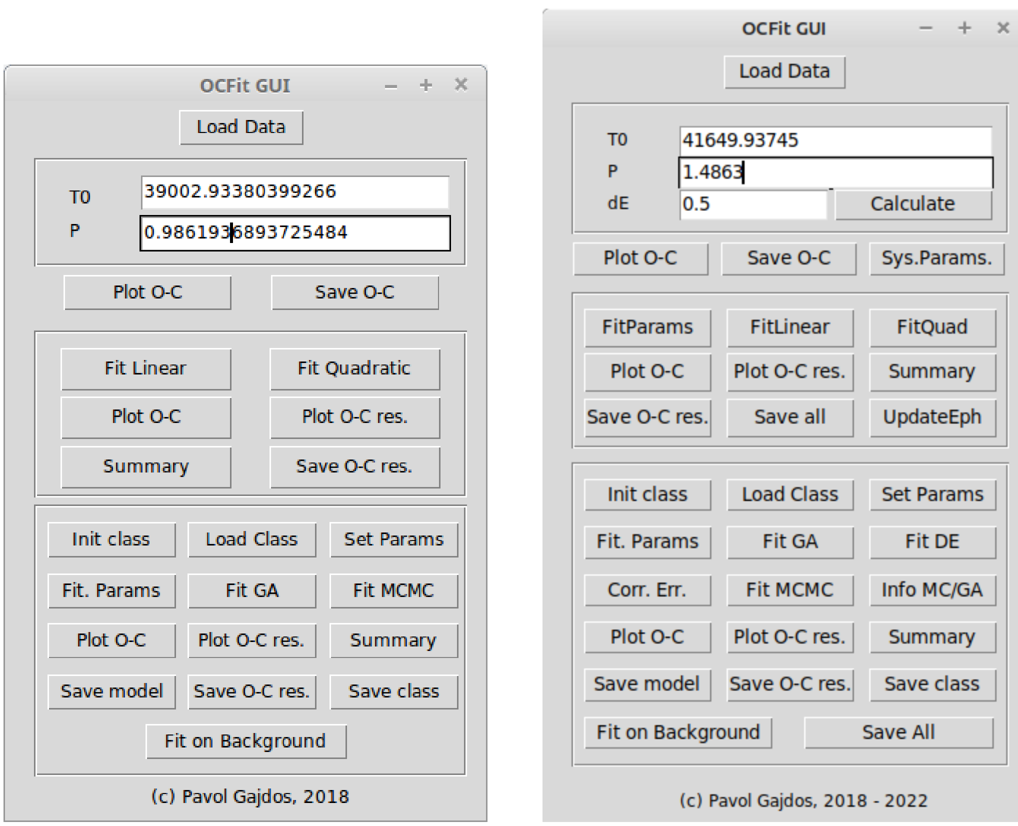

Figure 3: Comparison of the first version of GUI (*left*) and the current one (*right*).

New functions in GUI allow using the OCFit in a more powerful way. Many features which were hidden and could be used only directly from the main package in the own scripts, are now available also from GUI. Some parameters of the studied system (e.g.

January 2023 DOI: 10.5817/OEJV2020-XXX ISSN 1801–5964

masses and orbital inclination) could be added and after that used to calculate more indirect parameters of the O-C model (e.g. mass of the third body) which are displayed in the summary window. Errors of input data points (values of O-Cs) can be corrected (rescaled) after initial model fitting by GA or DE method. This is useful mainly if the input errors are underestimated. In such cases, fitting using the MCMC method could fail.

A separate button for updating linear ephemeris "UpdateEph" ensures that the original ephemeris is not replaced if the fitting using linear or quadratic function does not provide a suitable result.

The button "Save All" is extremely useful and makes the work with OCFIT much easier. It saves all possible outputs (figures, model curve, residue and parameters of the class) to individual files. Thus, they do not have to be saved separately and the chance of forgetting to save some important result disappears.

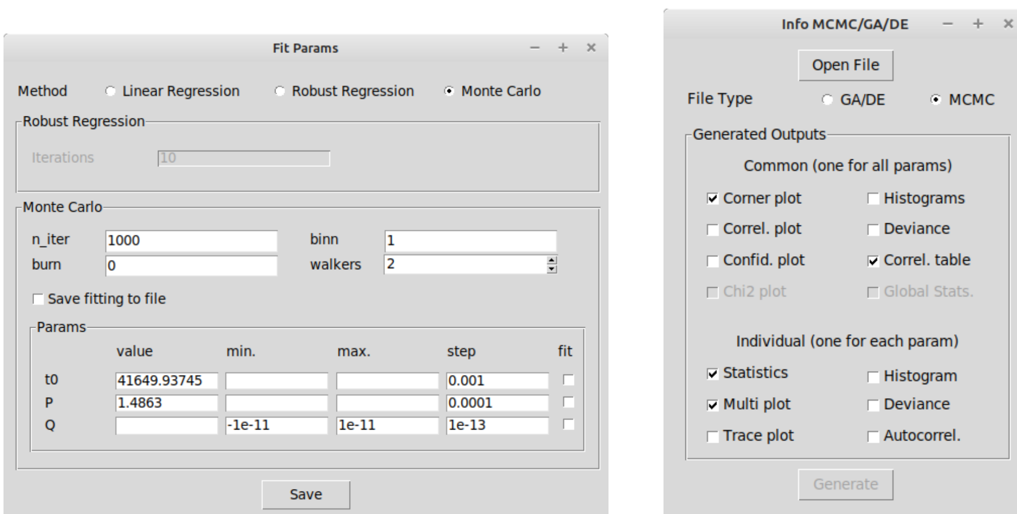

Figure 4: The setting of the fitting method and its parameters for a linear and quadratic model of O-Cs (*left*). Analysis of the process of fitting using GA/DE or MCMC method (*right*).

The fitting process for the linear and quadratic models of O-Cs was changed and available options were extended. A new window for the selection fitting method and setting its parameters was added (see Fig. 4). There are three available methods. The most simple is a standard linear regression. Iterative robust regression handles outlier points better by using them with lower weights. The MCMC method helps to judge whether a fitted linear or quadratic trend is real and how significant it is. This method provides the best uncertainties estimation and allows to analyse of the probable distribution of the values of fitted parameters. An apparent quadratic trend could be detected as bias if the distribution of the quadratic term Q is extremely wide and nearly uniform. The setting of parameters of the MCMC method is very similar to setting them in the main part of the program during fitting more complex O-Cs models. It is required to input the parameters of MCMC sampling (e.g. the number of iterations) together with the values, limits, and steps of fitted parameters of the linear or quadratic ephemeris. The whole sampling could

January 2023 DOI: 10.5817/OEJV2020-XXX ISSN 1801–5964

be also saved to file for further analysis.

Separate and independent analysis of MCMC sampling or GA/DE fitting process is now also available directly from GUI (Fig. 4). This means that sampling of a different model fitted to other data like the one currently working with could be analysed. Moreover, this analysis could also be performed without loading any O-C data into the OCFit software. The sampling is loaded from a file that was already saved before. However, it is necessary to choose a good type of input file (GA/DE fitting or MCMC sampling). There are multiple individual and general figures which could be plotted – e.g. a corner plot, histograms, trace plot or autocorrelation plot. Additionally, textual statistics with mean value, median, standard deviation, confident intervals, etc. and correlation table could be generated. All outputs are automatically saved to the files with appropriate names. Detail analysis could help us to find out if the process of fitting successfully converged or the number of iterations was insufficient, or the fitting interval of some parameter was not set correctly. Furthermore, a non-trivial correlation between fitted parameters could be studied in this way (mostly using the corner plot). Any correlation between parameters could influence the final results and make determined uncertainties less reliable because it deforms obtained final distribution of fitted parameters and it may not be exactly Gaussian any more as assumed at the calculation of values and uncertainties of parameters.

4 Conclusion

Since the first version of OCFit in 2015, the software was extensively modified and modernized. The current one is supported by the latest version of PYTHON and uses up-to-date common packages.

From user comments, bug reports and asks for help, this program is already being used by various astronomers from around the world in their research. About 200 people downloaded any software release from the GitHub web page. More than 2500 downloads from the PyPi repository were performed (using the "pip install OCFit" command). The original paper about OCFit (Gajdoš & Parimucha, 2019) has up to now twelve citations (without self-citations) in high-impact journals. These numbers oblige to still update this software – keep it supported by the new versions of PYTHON and/or OSs, fix possible bugs and implement new features.

Acknowledgements: Big thanks to everyone using software OCFit and especially for all comments, suggestions, discovered bugs and feedback in general. This work was supported by the Slovak Research and Development Agency under contract No. APVV-20-0148 and internal grant VVGS-PF-2021-2087 of the Faculty of Science, P. J. Šafárik University in Košice.

References

Agol, E., Steffen, J., Sari, R., & Clarkson, W., 2005, *MNRAS*, **359**, 567, 2005MNRAS.359..567A

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

Foreman-Mackey, D., 2016, *JOSS*, **1**, 24, 2016JOSS....1...24F

Gajdoš, P., & Parimucha, Š., 2019, *OEJV*, **197**, 71, 2019OEJV..197...71G

Giménez, A., & Bastero, M., 1995, *Ap&SS*, **226**, 99, 1995Ap&SS.226...99G

Harris, C. R., et al., 2020, *Natur*, **585**, 357, 2020Natur.585..357H

Hunter, J. D., 2007, *CSE*, **9**, 90, 2007CSE.....9...90H

Irwin, J. B., 1952, *ApJ*, **116**, 211, 1952ApJ...116..211I

Storn, R., & Price, K., 1997, *JGlobOptim*, **11**, 341

Virtanen, P., et al., 2020, *NatMe*, **17**, 261, 2020NatMe..17..261V