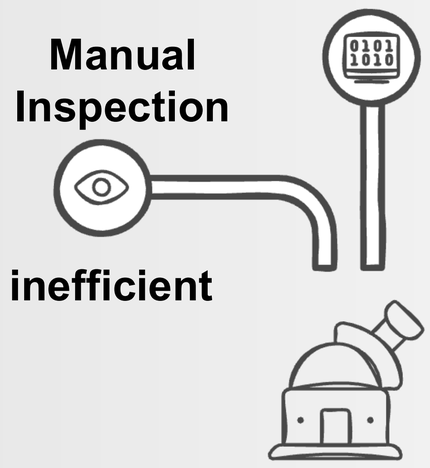
Identifying Binary Asteroids Using Machine Learning with Simulated Lightcurves and FOSSIL Survey Data

Cheng-An Hsieh[1*],Chan-Kao Chang (章展誥)[2], Ying-Tung Chen (陳英同)[2] , Matthew J. Lehner[2, 3], Shiang-Yu Wang (王祥宇)[2], TAOS II team
[1]Department of Physics, National Taiwan University
[2]Institute of Astronomy and Astrophysics, Academia Sinica, No.1, Sec. 4, Roosevelt Rd, Taipei 10617, Taiwan
[3]Department of Physics and Astronomy, University of Pennsylvania, 20 South 33rd Street, Philadelphia, PA 19125, USA

*Email: b12202042@ntu.edu.tw

Why use ML to identify binary asteroids?

- Binary asteroids provide valuable insights into the formation and evolution of the solar system. They can be detected by analyzing lightcurves that exhibit eclipsing features and an **inverse U** shape. In the past, identifying binaries through lightcurves required manual inspection.
- The Formation of the Outer Solar System: an Icy Legacy (FOSSIL) survey offers tens of thousands of lightcurves of small solar system bodies, serving as a valuable resource for identifying binary systems. We trained a ML model to identify binary lightcurves using a dataset of simulated lightcurves.



ML Approach
New method to identify binary asteroids.

Large Sky Survey
A lot of Lightcurves awaiting analysis.

How to train the ML model?

1. Asteroids light curve model

We use ellipsoid model with three axes, a, b and c to represent the two asteroids in the binary system where $a \geqq b \geqq c$. Following the binary classifications of Walsh et al. (2015)[3], we adopted diameter ratio and separation to generate synthetic lightcurves and added photometric noise according to the FOSSIL survey.

2. Feature Engineering Random forest classifier (RFC)

We extracted features from the lightcurves and employed a RFC[4] to develop a model for identifying binary systems. The RFC was trained on features such as FWHM, skewness, and dip depth. Preliminary results showed an accuracy of approximately 0.76 on the testing set.

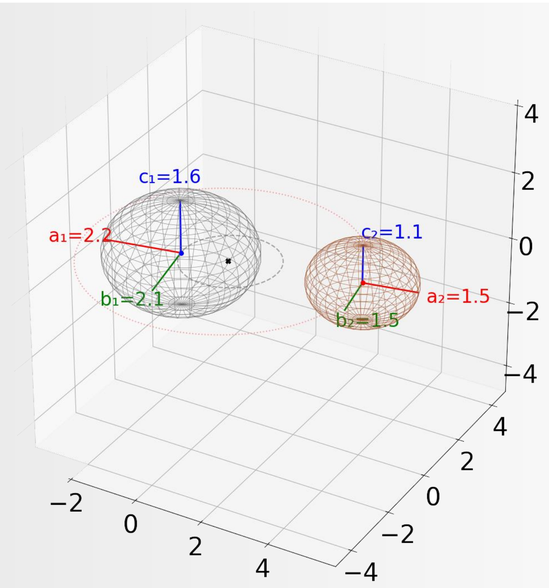


Fig 1. Binary Model
Defined by corresponding semi-axes for its two asteroids: a_1, a_2 (red), b, b_2 (green) and c_1, c_2 (blue).

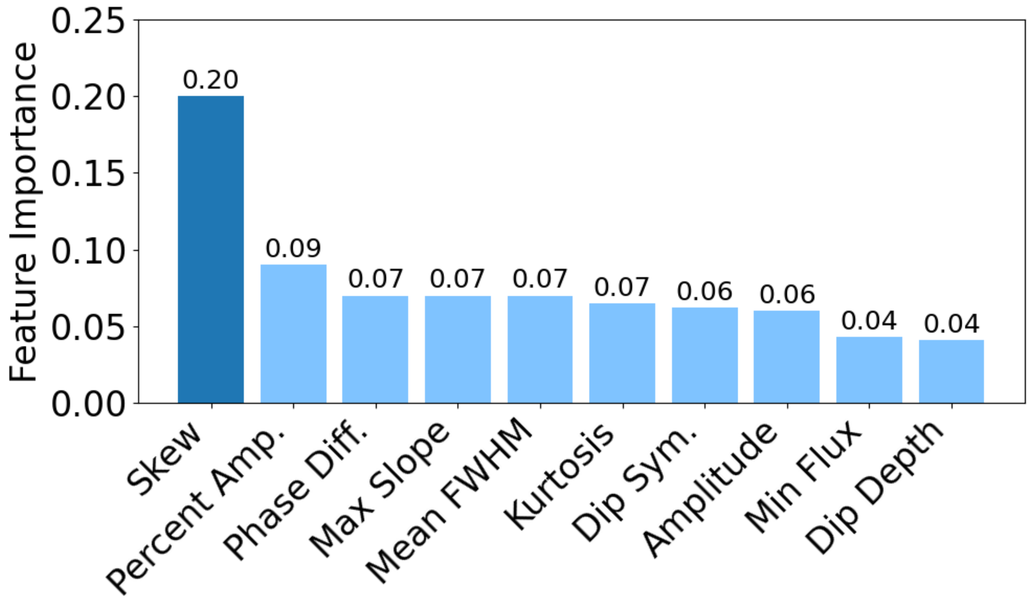


Fig 2. Important Features
Top 10 most important features from our RFC for identifying binary asteroids.

Binary Asteroids Identification

1. The initial test utilized 926 lightcurves from the FOSSIL Survey with well-measured rotation periods, and manual inspection identified **21 binary candidates**.
2. ML model identified 309 potential binary lightcurves.
3. We further applied a filter: amplitude > 0.25 and dip-depth > 0.4 yielded **75 likely binaries**, and we flagged **147 candidates for review** if amplitude, dip depth, minimum flux exceeded median of all lightcurves.
4. This study shows that the ML approach can reduce manual review time by **75%**. We will improve the model's accuracy and apply it to the full FOSSIL dataset.

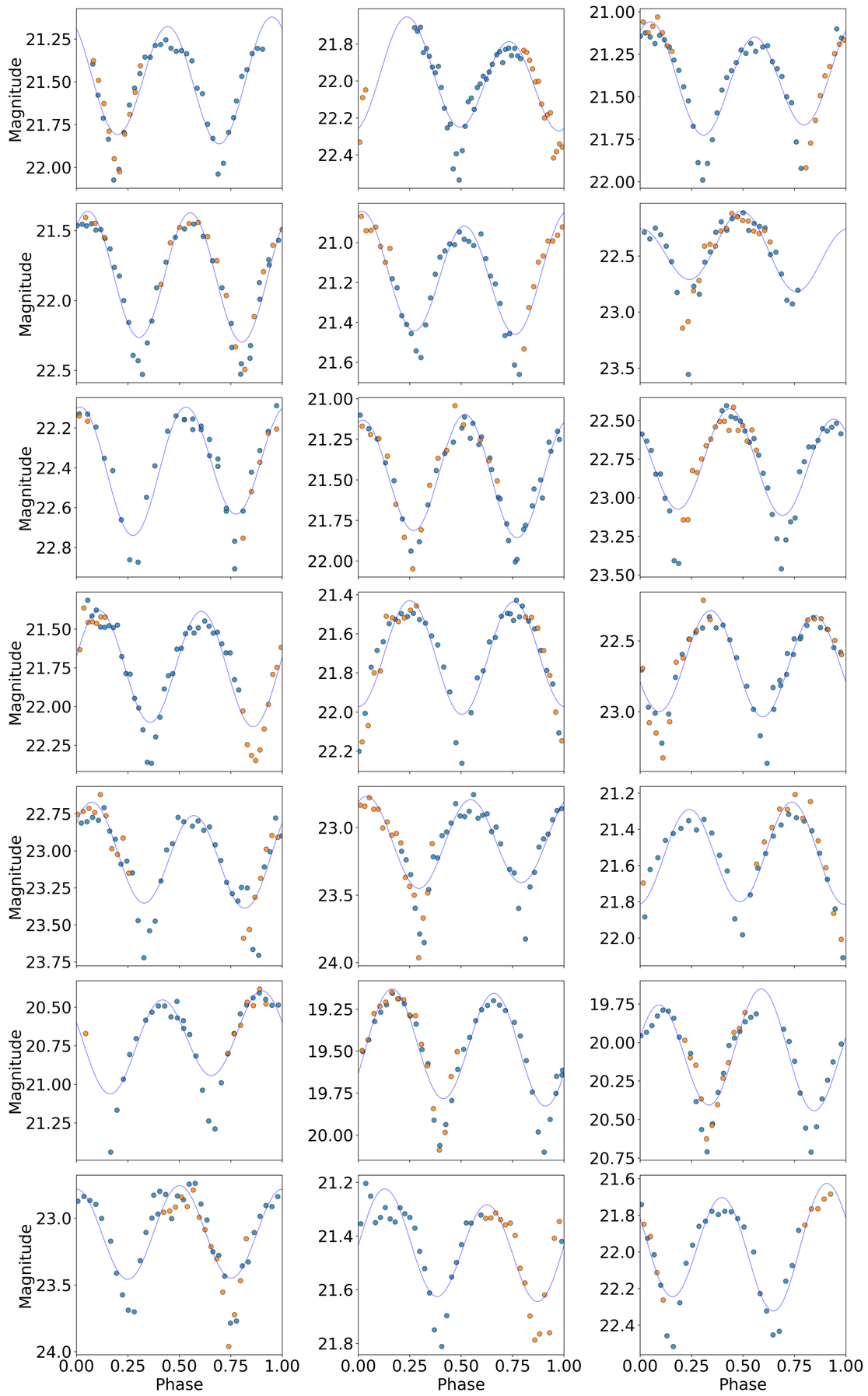


Fig. 3. Light curves of 21 binary candidates
The blue line shows the second-order Fourier fit, blue and orange dots represent observations from different days.

[1][2]Chang, et al.(2021)(2022) https://www.fossil-survey.org/ [3]Walsh, et al. (2015) [4]Breiman, L. (2001). [5]Cabral, et al.(2018)