WFIRST: Microlensing Analysis Data Challenge

R.A. Street, Las Cumbres Observatory, rstreet@lco.global
and
WFIRST Microlensing Science Investigation Team

WFIRST will produce thousands of high cadence, high photometric precision lightcurves of microlensing events, from which a wealth of planetary and stellar systems will be discovered. However, the analysis of such lightcurves has historically been very time consuming and expensive in both labor and computing facilities.
This poses a potential bottleneck to deriving the full science potential of the WFIRST mission.
To address this problem, the WFIRST Microlensing Science Investigation Team designing a seriesof data challenges to stimulate research to address outstanding problems of microlensing analysis. These range from the classification and modeling of triple lens events to methods to efficiently yet thoroughly search a high-dimensional parameter space for the best fitting models.

WFIRST will survey:

~2 deg^2 in Galactic Bulge
6 observing seasons, 72d each

with cadences:
~15min in W149 'wide' filter (927-2000nm) and
~12hr in Z

~40,000 datapoints per star for ~85% of sample

The mission is expected to produce
~4 x 10^7 lightcurves total down to H_{AB}<21.6mag

Including:
~1400 bound planet events ~300 free-floating planet events

For details, see Spergel et al. 2013.

ANALYZING WFIRST LIGHTCURVES

WFIRST will produce thousands of high cadence, high photometric precision lightcurves of microlensing events, from which a wealth of planetary and stellar systems will be discovered.

However, the analysis of such lightcurves has historically been very time consuming and expensive in both labor and computing facilities. This poses a potential bottleneck to deriving the full science potential of the *WFIRST* mission.

Binary lens models typically include at least 9 free parameters, and finding the model that best fits the data is a complex minimization process, with multiple degeneracies.

Currently, modeling is conducted by a very small number of specialists worldwide and in general is a time consuming and laborious process using private software. This tends to create a backlog of data awaiting analysis, and more subtle planet signatures can be missed.

To make the most of the *WFIRST* mission, we need to develop our analysis capabilities, automating the process wherever possible.

While major progress has been made in recent years to make analysis software faster and more automated (e.g. Dominik 2007, Bozza 2010, Bachelet et al. 2017), existing code and analysis process doesn't scale to large datasets.

References:
Bozza, V., 2010, MNRAS, 408, 2188. Bachelet, E. et al., 2017, ApJ, 154, 208. Dominik, M. 2007, MNRAS, 377, 1679.

DATA CHALLENGE GOALS

- To stimulate research effort into outstanding modeling issues
- To stimulate development of algorithms to detect and classify microlensing events in *WFIRST* data
- To stimulate development of software for modeling microlensing events, capable of conducting analyses of *WFIRST*-scale datasets

HOW THE CHALLENGES WILL WORK

A series of simulated *WFIRST* lightcurve datasets will be released. The nature of the variability of each object will remain a secret until the challenge submission deadline.
The release will describe the metrics that will be used to evaluate all entries to the challenge. These will include numerical metrics comparing the accuracy of the model parameters as well as for the performance of the software used, as well as descriptive statements including the computer facilities used, techniques used etc.

Each person or team will submit a summary of their findings and metrics by the deadline.
Entries will be evaluated by a panel of experts drawn from the community, and the results published in a refereed paper.

Data releases: At the annual microlensing conference (late January/early February)
Submission deadline: 31 October

CHALLENGE DATASETS

Challenge 1 (2018):
Distinguishing single and binary lenses and variable stars
~250 *WFIRST* lightcurves
50 single and binary lenses injected
~200 lightcurves mimicking common variable types
(RR Lyrae, flares, Cephieds...)

Challenge 2 (2019):
Distinguish and model binary and triple lenses 50 binary and triple lens
event lightcurves

Challenge 3 (2020):
Whole survey analysis
100,000 lightcurves, including thousands of single and binary lens events,
tens of triple lens events and thousands of variable stars.

RESOURCES

For more information and the data releases, visit:
http://microlensing-source.org/data-challenge/
Summary of links to available software packages:
http://microlensing-source.org/software/
Work supported by a NASA Exoplanets Research Program grant, NNX15AC97G.

REFERENCES

Bozza, V., 2010, MNRAS, 408, 2188.
Bachelet, E. et al., 2017, ApJ, 154, 208.
Dominik, M. 2007, MNRAS, 377, 1679.
Spergel et al. 2013. arXiv:1305.5425