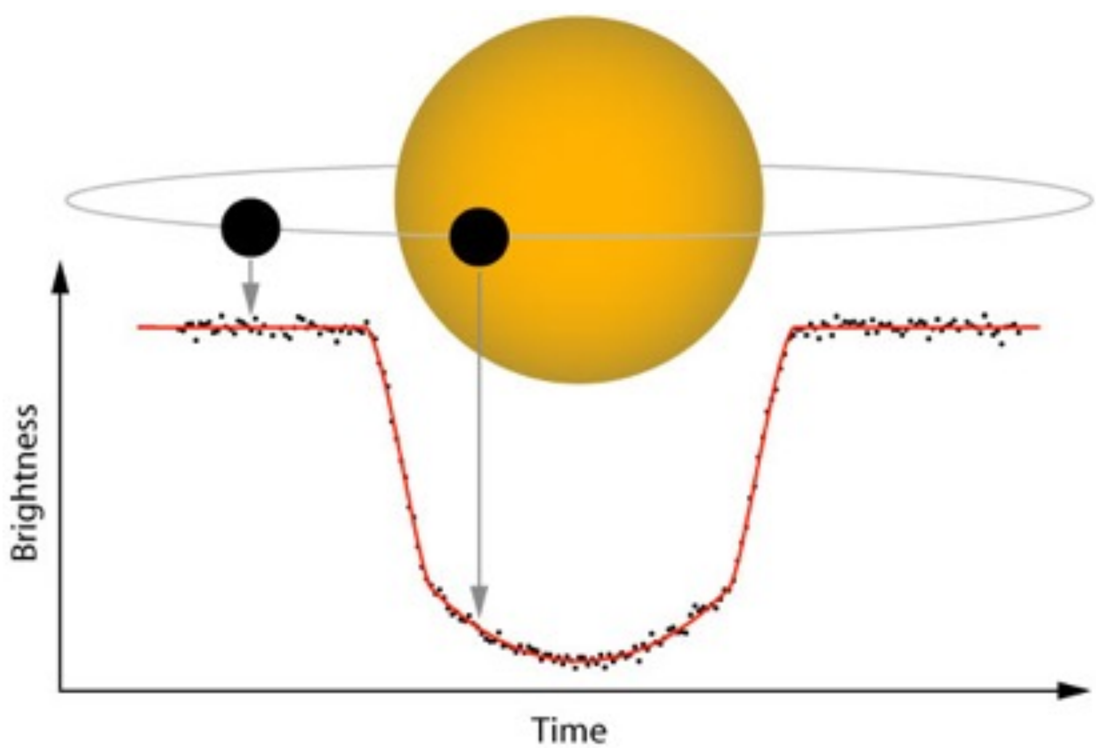
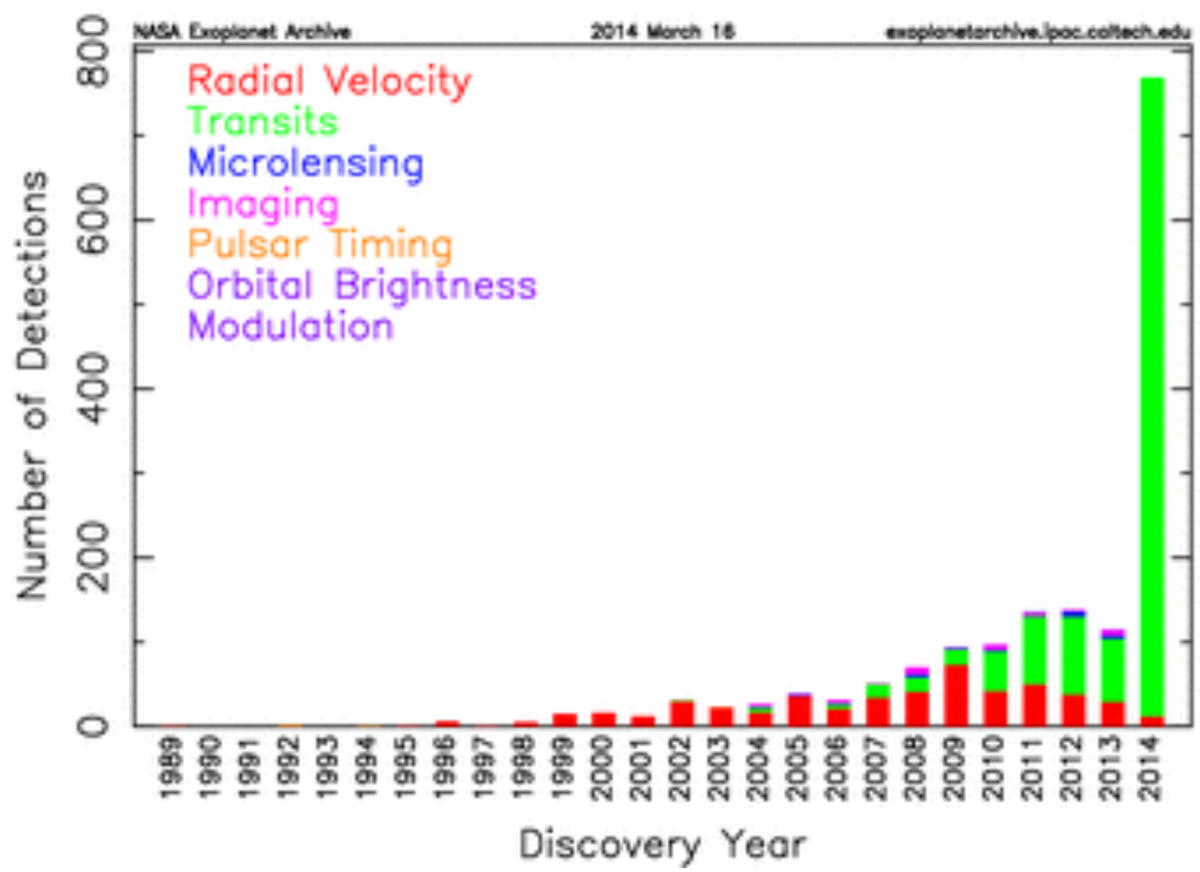
MACHINE LEARNING WORKSHOP UTSC

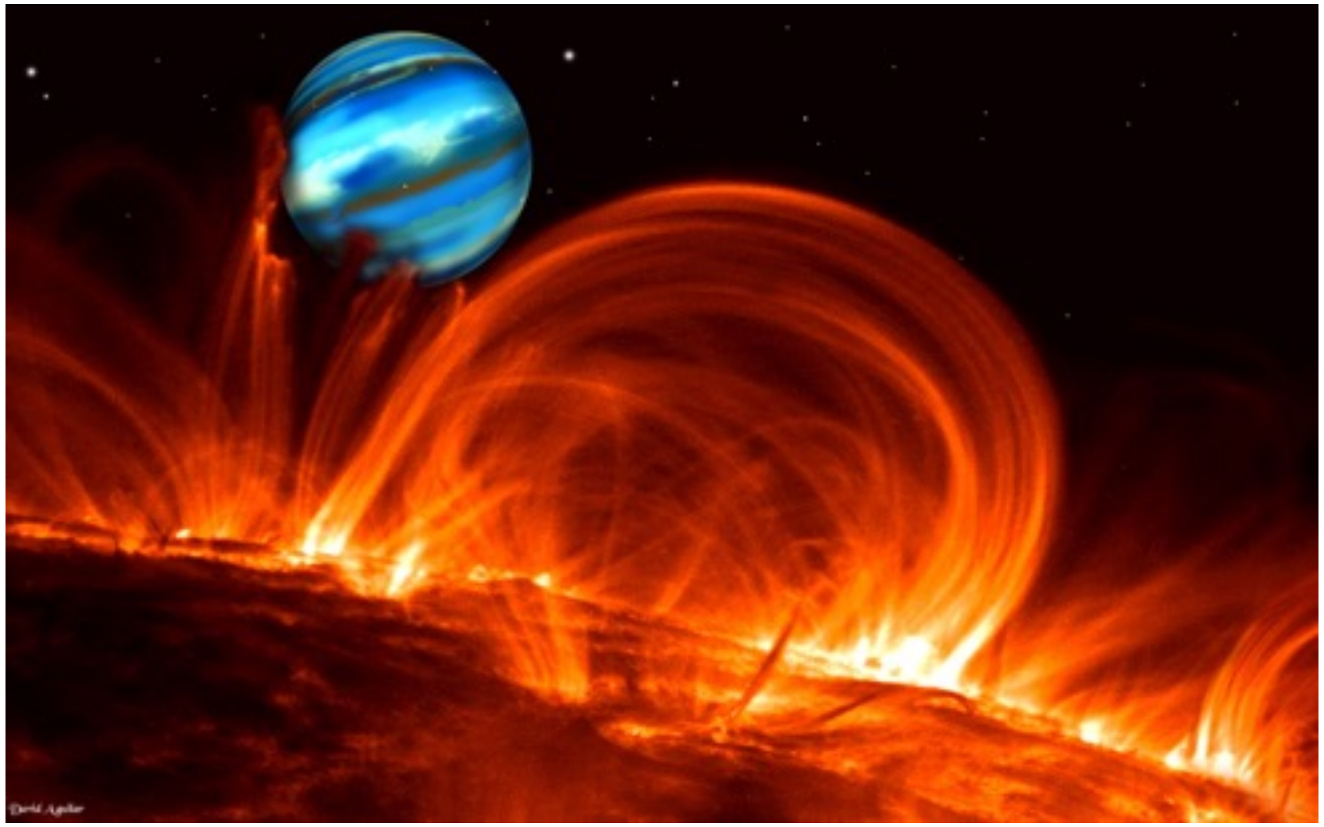
AUTOMETIC SELECTION OF TRANSIT SIGNALS

THE TRANSIT TECHNIQUE



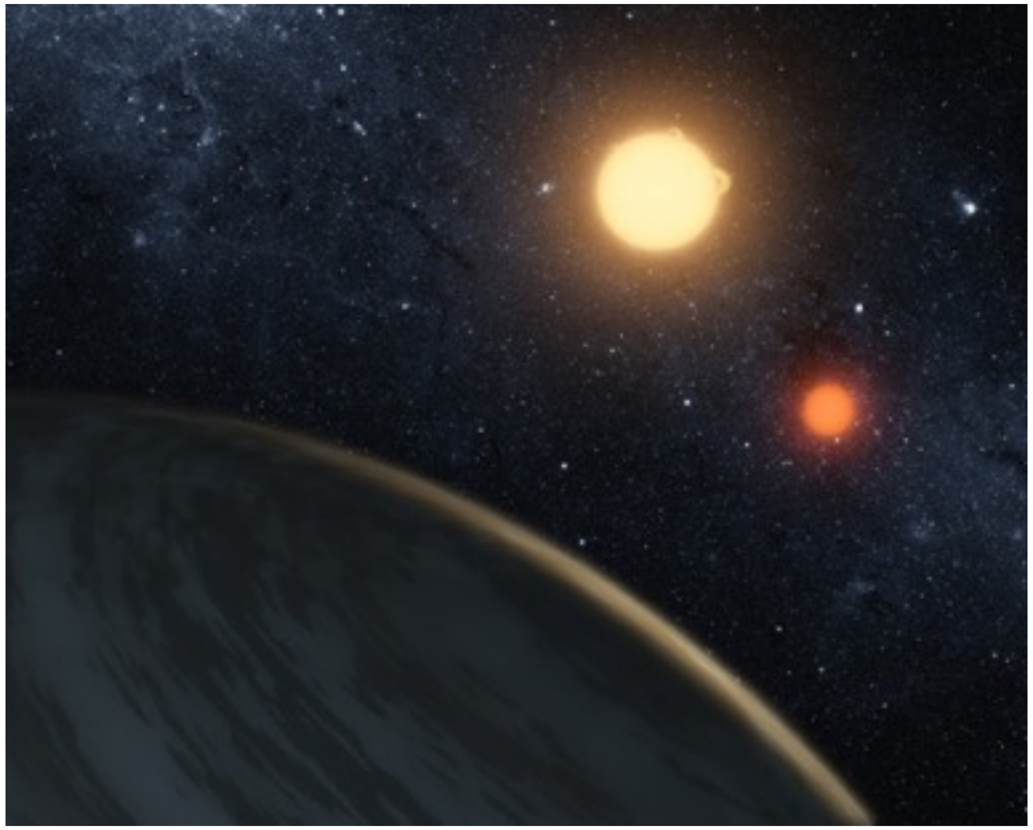


HIGHLIGHTS OF TRANSITING PLANETS



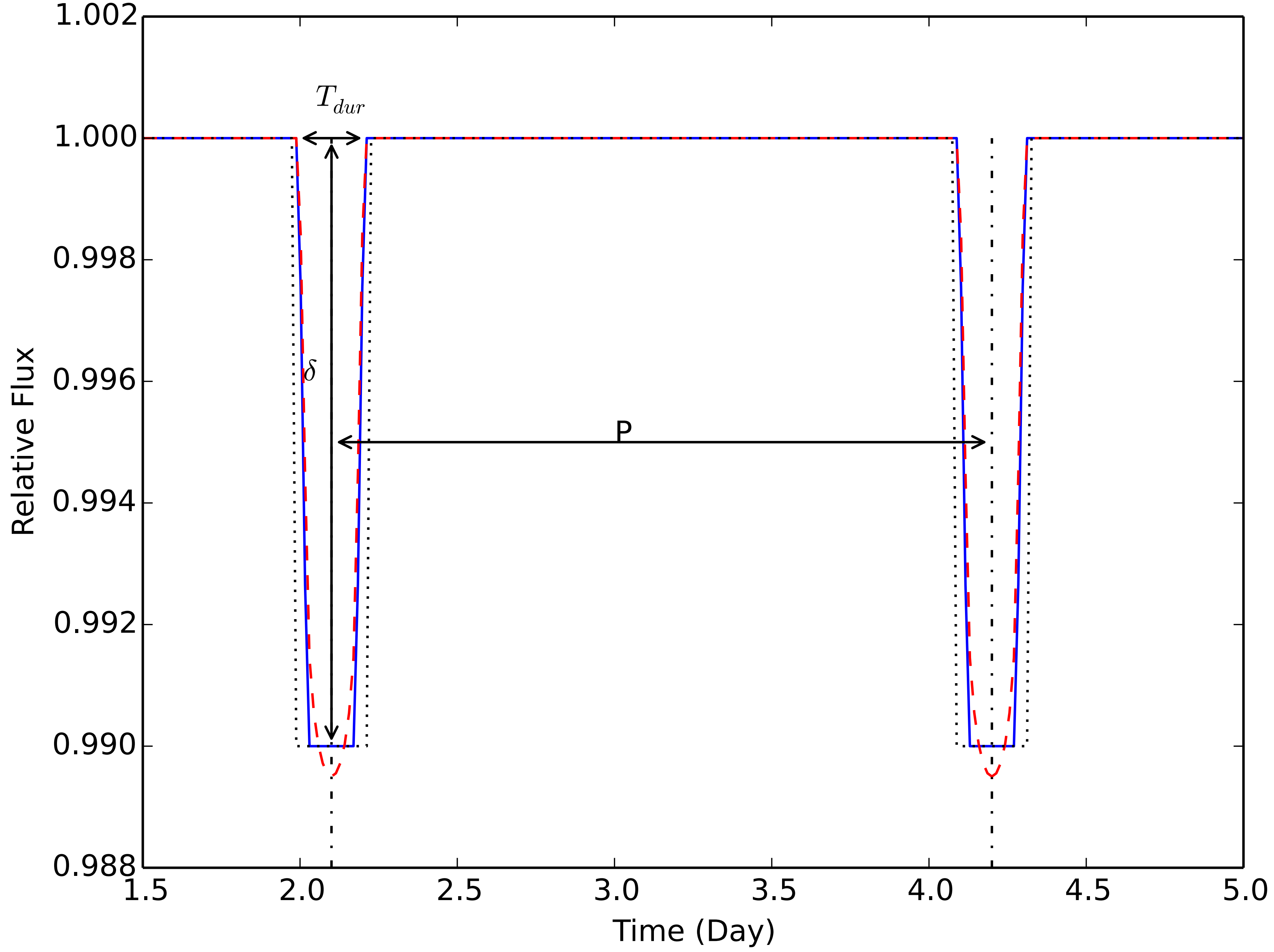
Credit: http://www.cfa.harvard.edu/pao/wallpaper.ht

KOI-500 – the most packed multi-systems
Credit: Ian Steadman

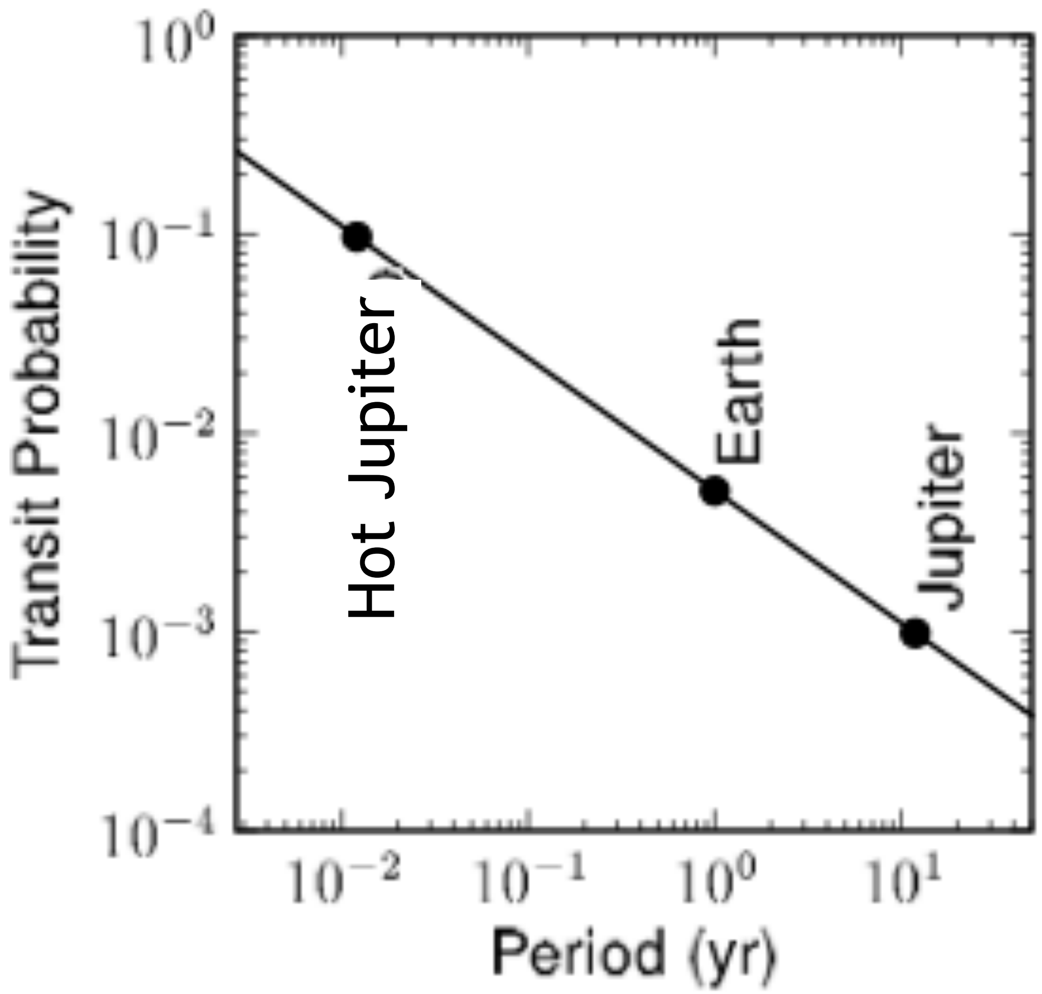


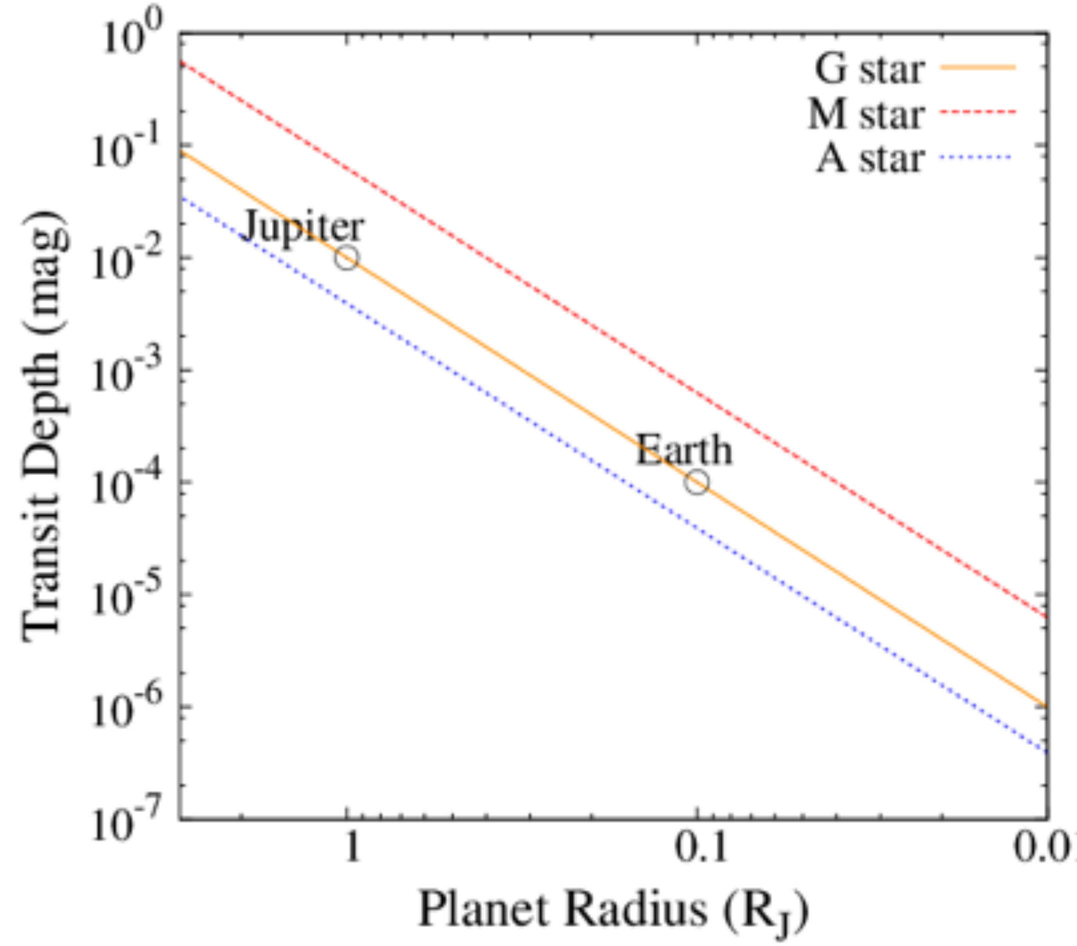
Kepler-16b circumbinary planet
Credit: NASA/JPL-Caltech/T.Pyle

BASIC TRANSIT GEOMETRY



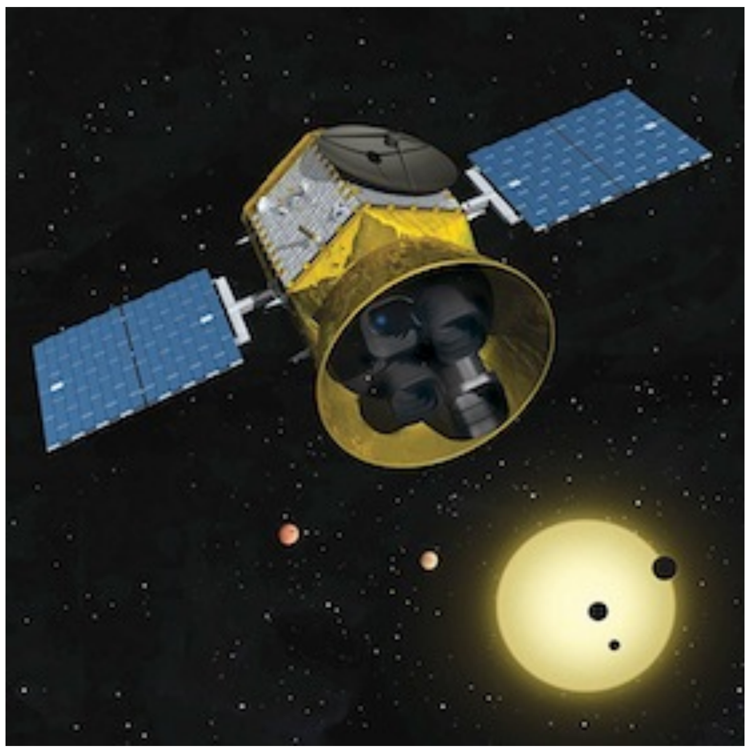
Transit Surveys



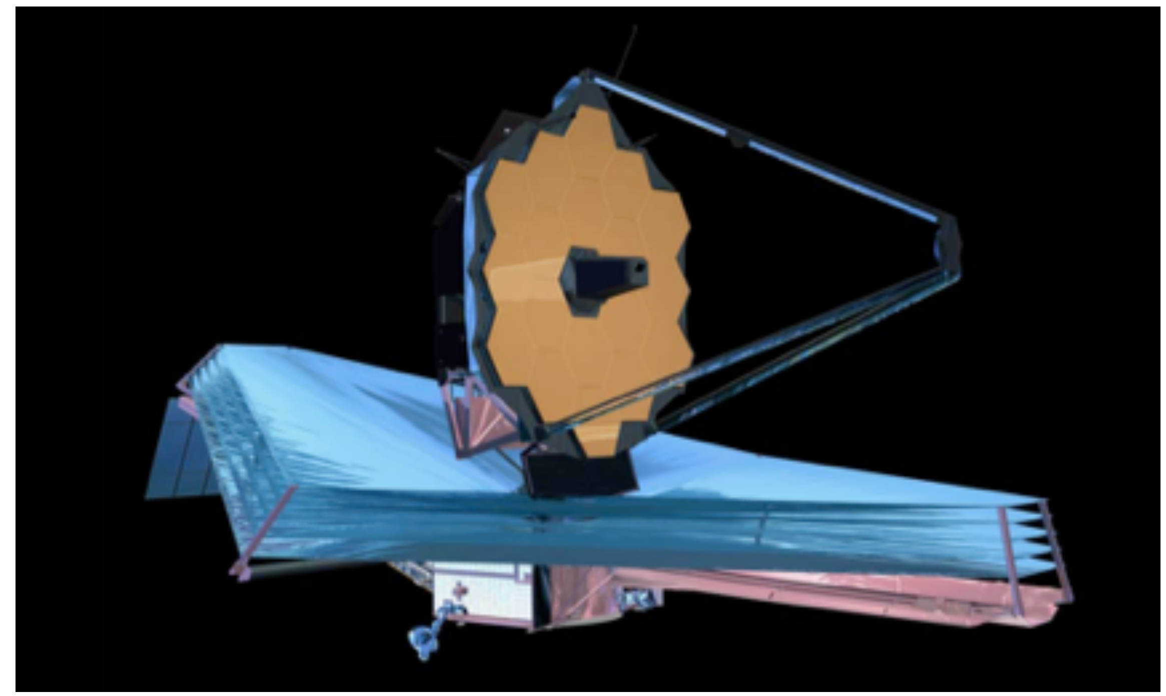


- Survey **tens of thousands of** stars.
- Obtain photometry precision of at least 1%.

TESS AND JWST



Target Feed

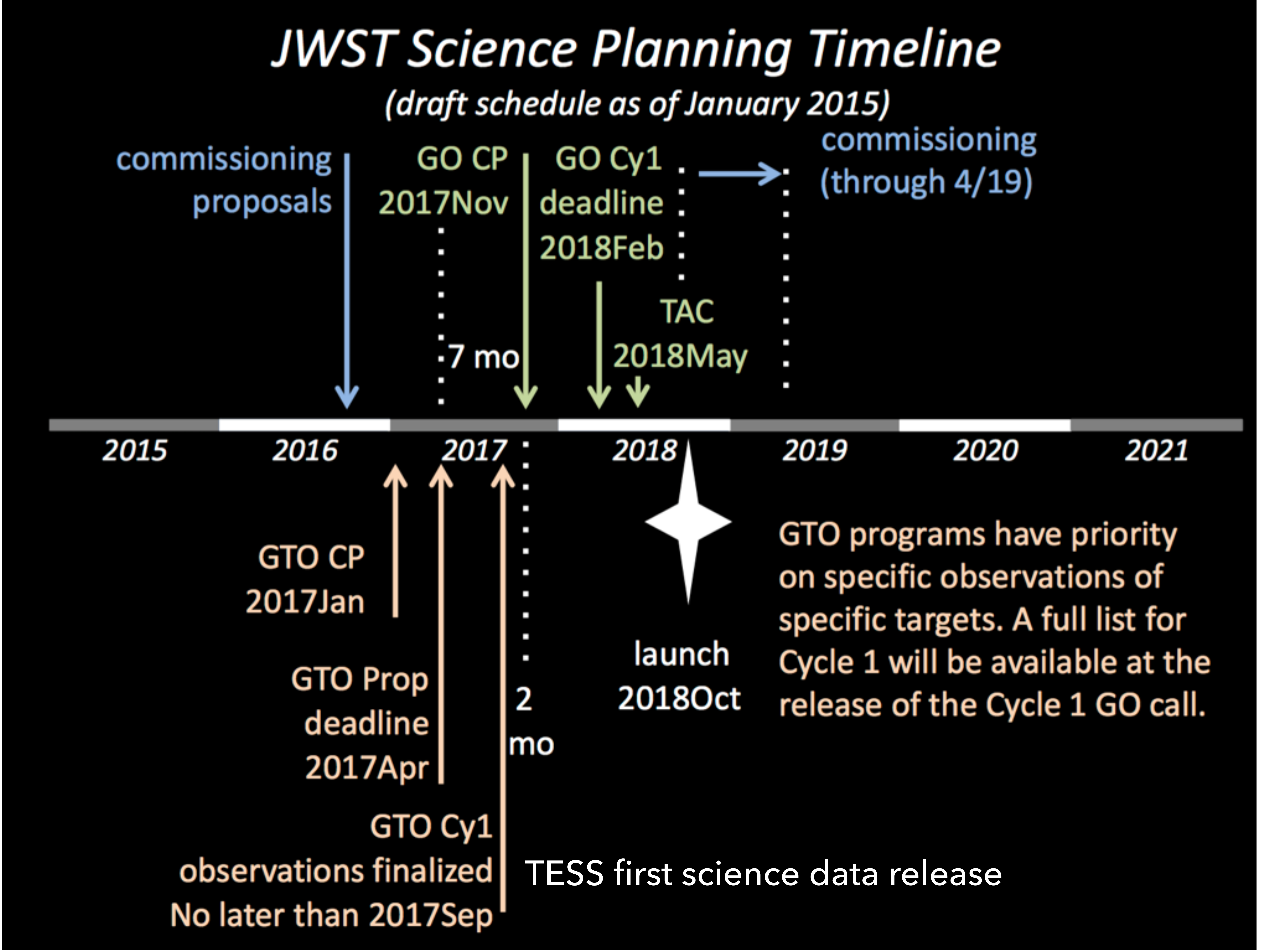


TESS:

Transit exoplanet survey satellite aiming for discovery of super earth around bright stars.

JWST:

8 billion Infrared Satellite, measuring the atmosphere properties of exoplanets.


JWST Science Planning Timeline
(draft schedule as of January 2015)
commissioning proposals
GO CP 2017Nov
GO Cy1 deadline 2018Feb
commissioning (through 4/19)
7 mo
TAC 2018May
2015
2016
2017
2018
2019
2020
2021
GTO CP 2017Jan
GTO Prop deadline 2017Apr
2 mo
launch 2018Oct
GTO programs have priority on specific observations of specific targets. A full list for Cycle 1 will be available at the release of the Cycle 1 GO call.
GTO Cy1 observations finalized No later than 2017Sep
TESS first science data release

WHY TESS DATA?

- Data quality in between HATNet and Kepler.
- Needed rapid/accurate selection of candidates for JWST observations.
- Much more data to deal with than Kepler.

WHAT'S THE AIM OF THE PROJECT

- Step 1: Build up a frame work for transiting planetary candidate selection.
- Step 2: Improve on the previous result making use of what we learnt.
- Main Question: what to optimize?

WHAT IS THE TRAINING SET

- Kepler data
- Kepler data + TESS Noise
- Simulated TESS data
- What fraction of Planet + False positives + False Alarms to input?

WHAT ALGORITHMS TO USE

- Classification problem
- SVM, random forest, KNN, logistic regression, …

WHAT FEATURES TO USE

- Observable of individual light curves.
    - shape of the signal compare to a transit
    - significance of the signal
    - noise properties of the light curve
    - other features in the light curve
- Other observables (centroid time series of the star).
- Stellar Properties.
- Signal correlation with other detected signals.

READING MATERIALS

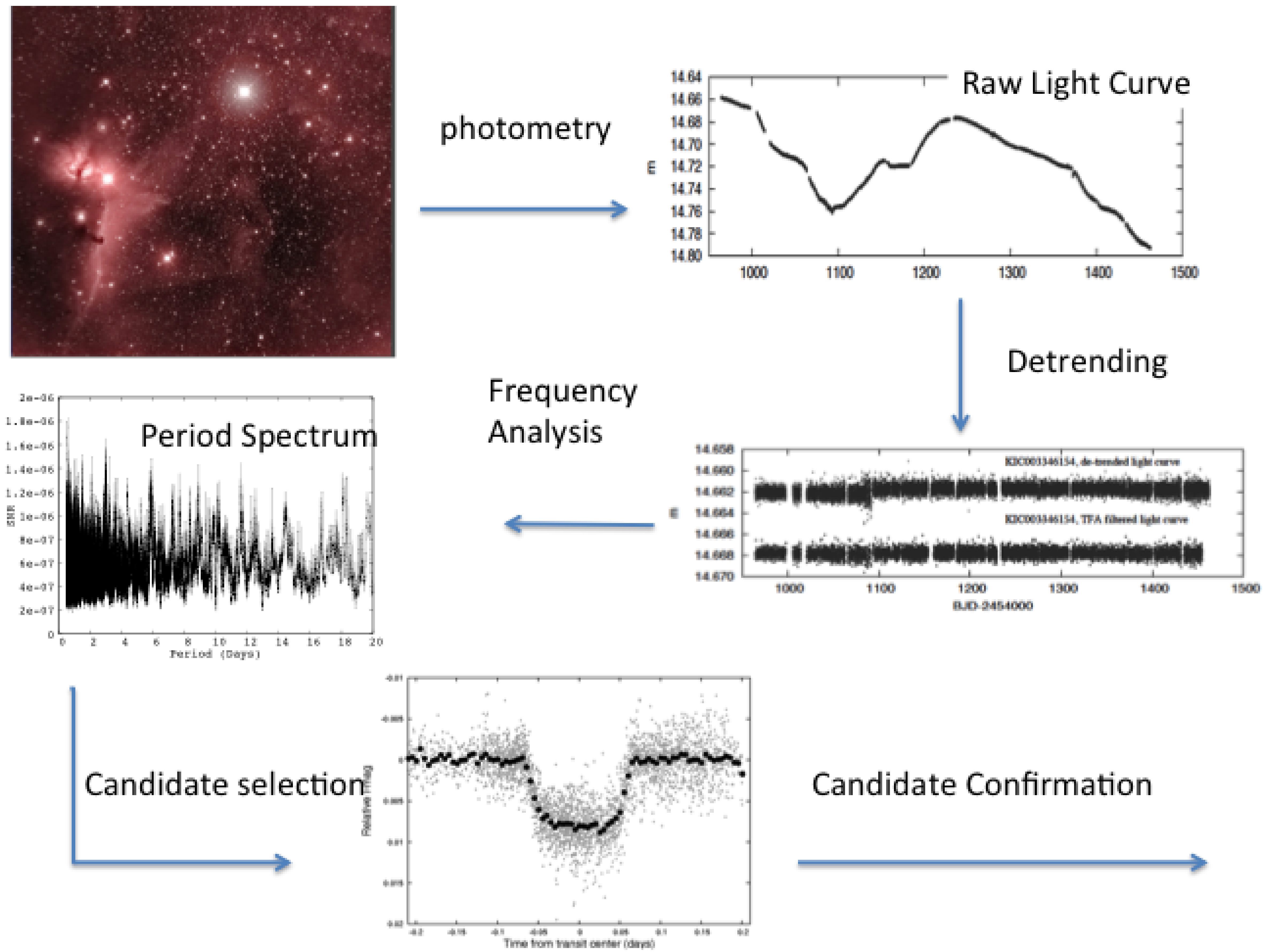
- http://arxiv.org/pdf/1408.1496v2.pdf
- http://arxiv.org/pdf/1512.06149v1.pdf
- Exoplanet handbook by Michael Perryman, Section 6, 6.1, 6.2.1, 6.2.5, 6.2.7, 6.4.1
- http://www.planethunters.org

CURRENT PRACTICE

- Kepler TCERT - random forest
- HATNet - SVM

POSSIBLE APPLICATIONS

- Kepler
- HATNet
- K2
- **TESS**



photometry
Raw Light Curve
Detrending
Frequency
Analysis
Period Spectrum
Period (Days)
SNR
BJD-2454000
Candidate selection
Candidate Confirmation
Time from transit center (days)
Relative mag

Image credit: Wikimedia

The Kepler Mission

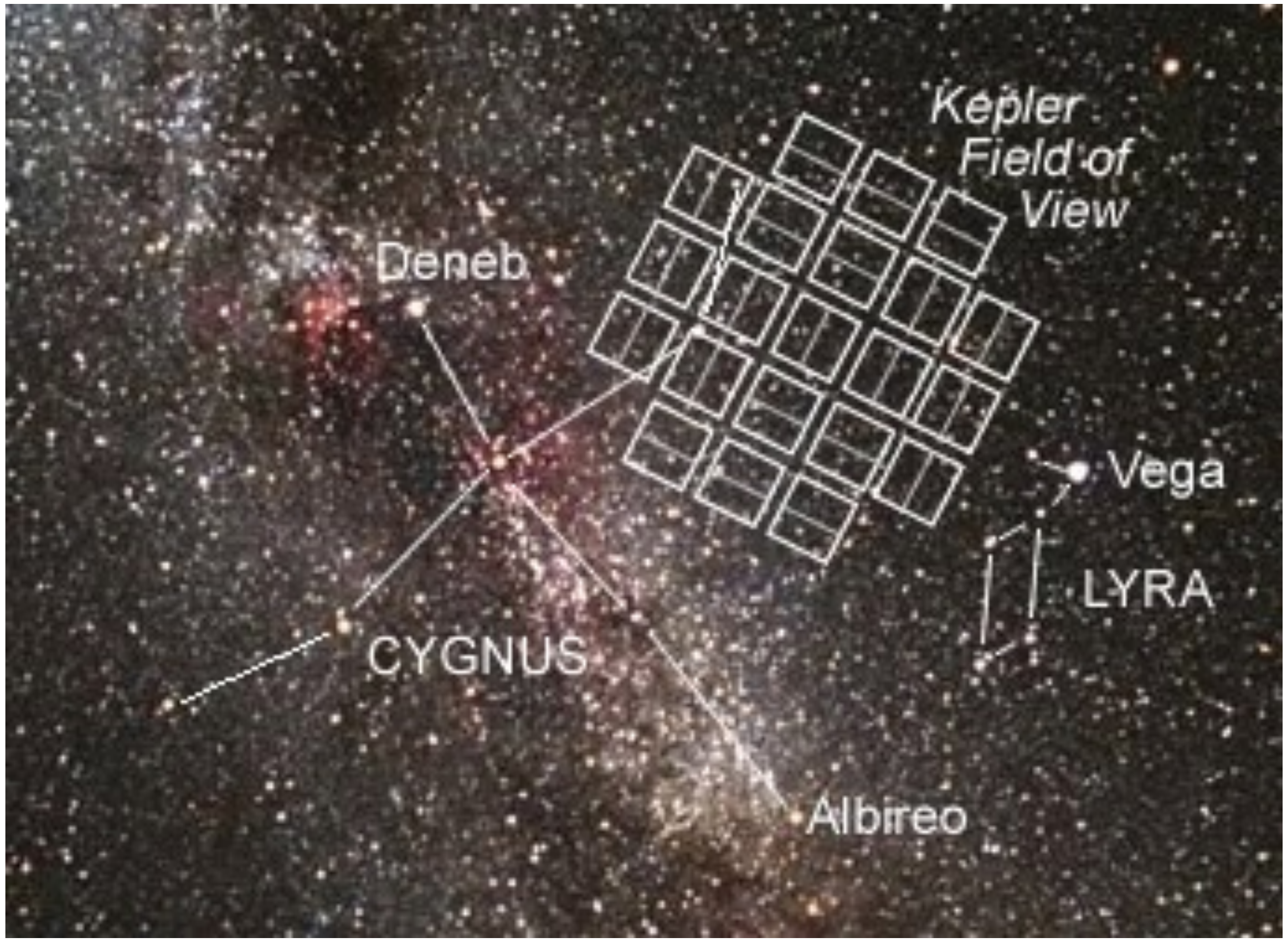
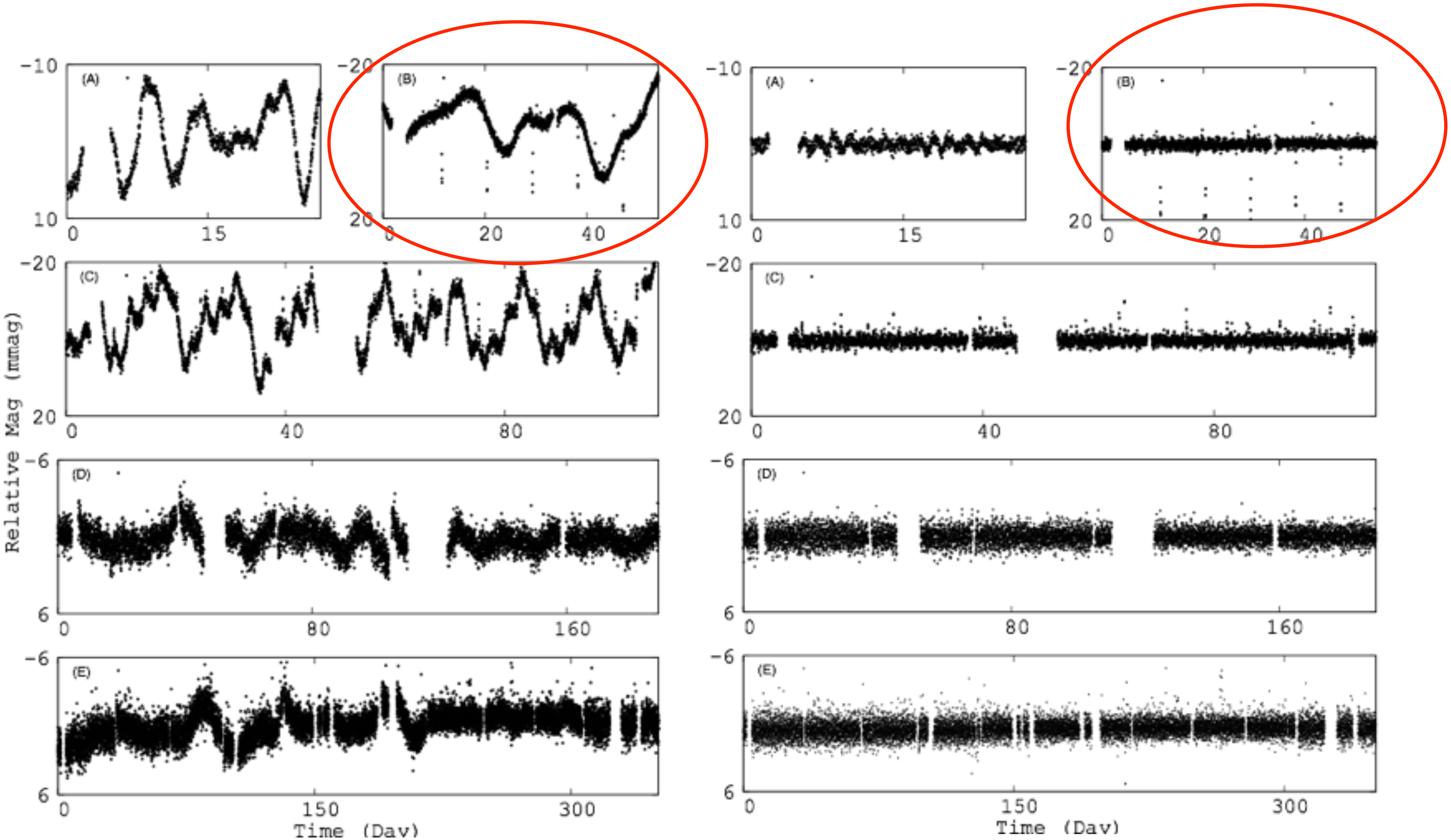


Image by Carter Roberts / Eastbay Astronomical Society

WHAT WE ARE LOOKING FOR



WHAT ARE WE LOOKING FOR

- Strictly Periodic dimming signals, box-like
- Radius derived from transit geometry similar to what is a planet
- Astrophysical origin

FALSE POSITIVE (AFP) AND FALSE ALARMS (NTP)

False Positives:

- **Often** high signal to noise;
- Have Astrophysical origin;

False Alarms:

- Often low signal to noise;
- Due to defect in data;

KEPLER AUTO CANDIDATE SELECTION PIPELINE

Prediction on training set

Training set	PC	AFP	NTP
PC	2843	8	28
AFP	98	271	24
NTP	25	12	11267

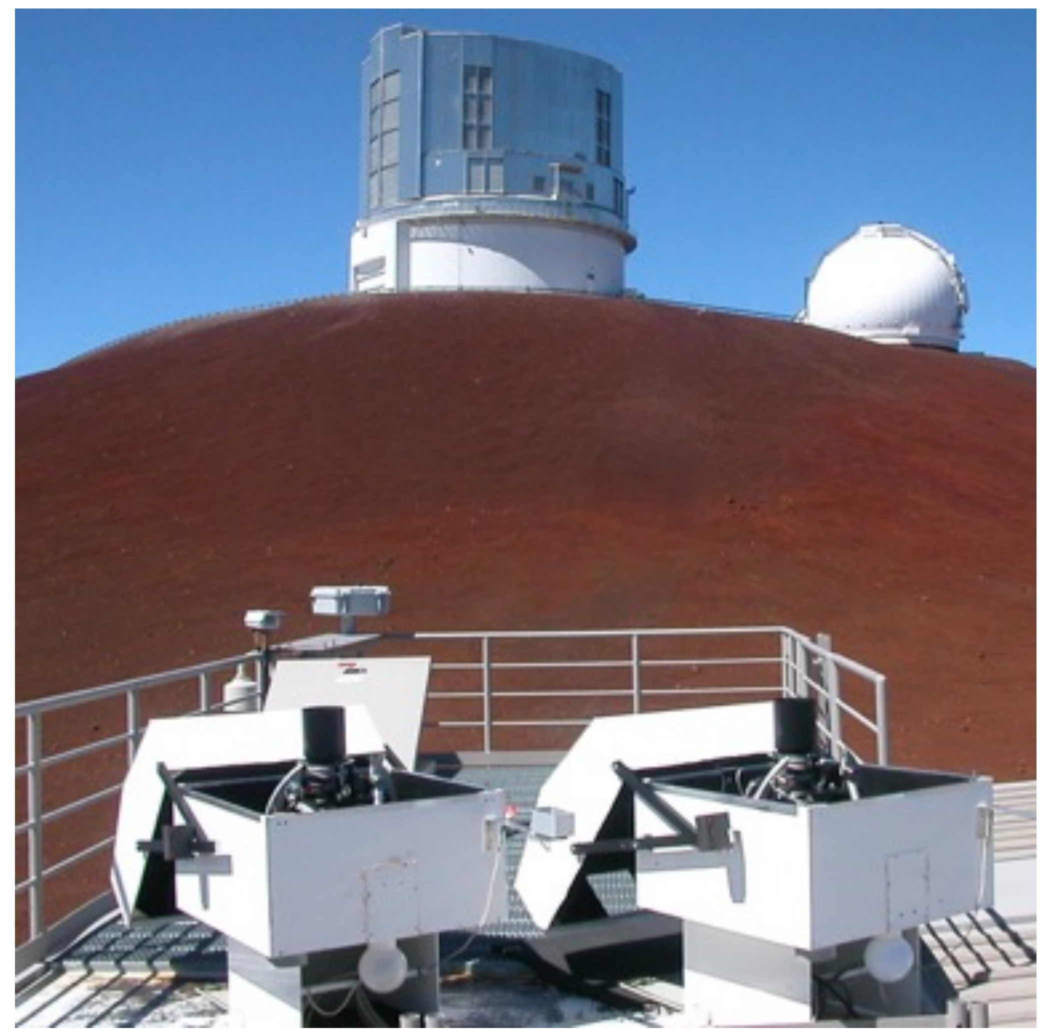
Prediction on a later sample

TCE Class	PC	AFP	NTP
PC	314	32	43
AFP	95	553	450

HATNet (2003-present)

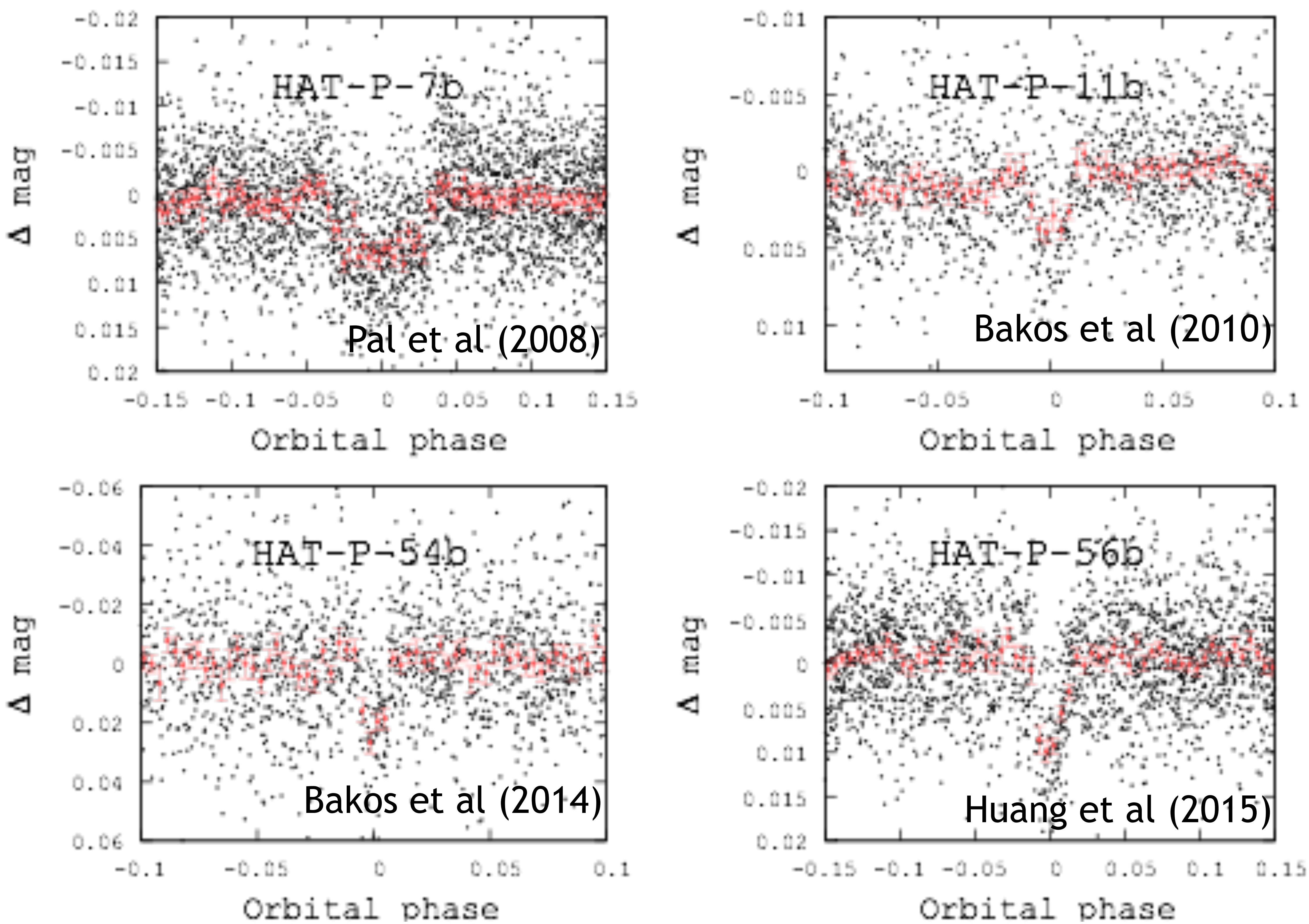
HAT – Hungarian-made Automated Telescopes

- The Network (six 11 cm diameter, wide field telescopes)
- Highlights and Discoveries (~56 published planets)

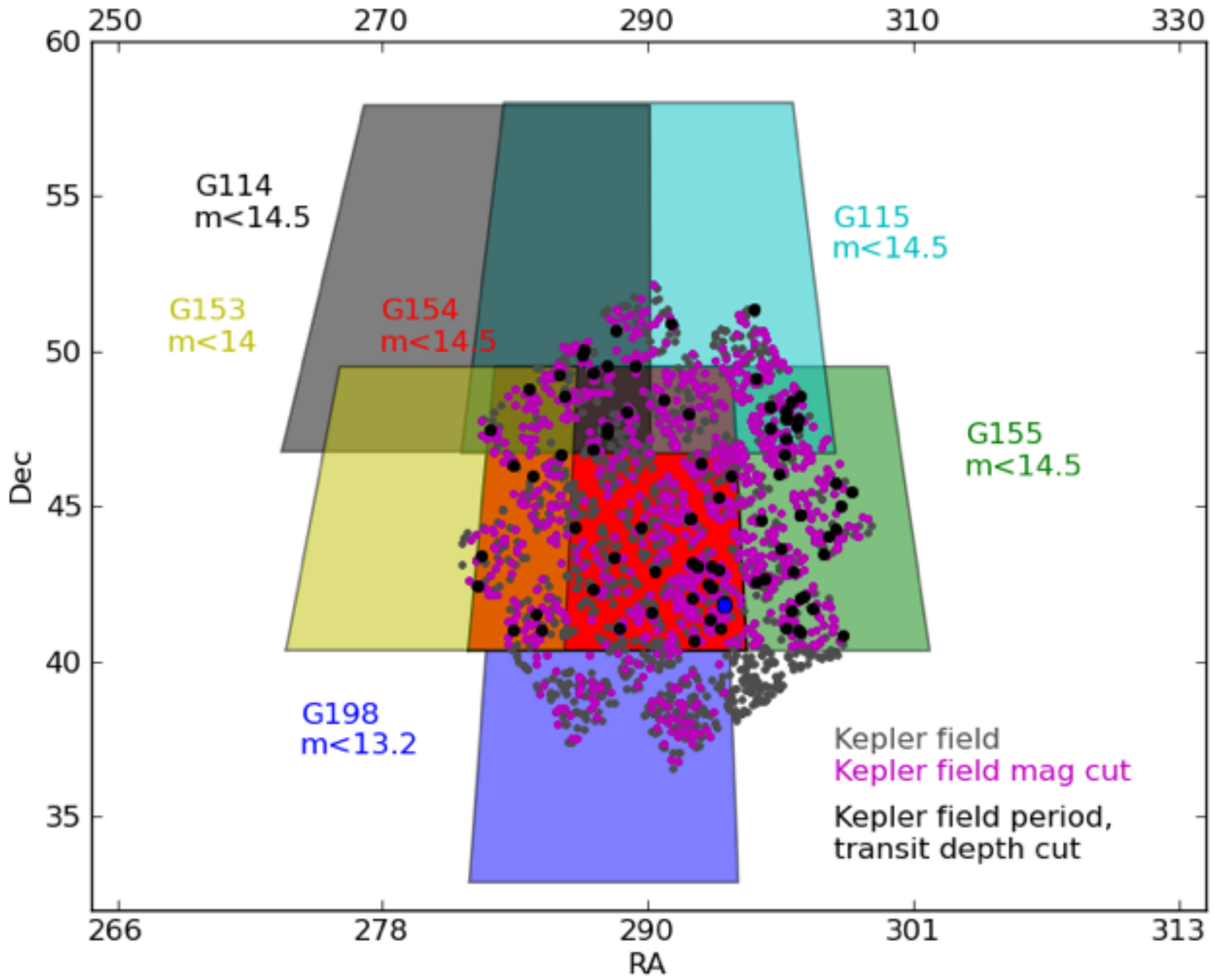


Bakos et al 2004

Examples of HATNet discoveries



Calibrate HAT with Kepler –
What can we learn from simultaneous high precision space based data?

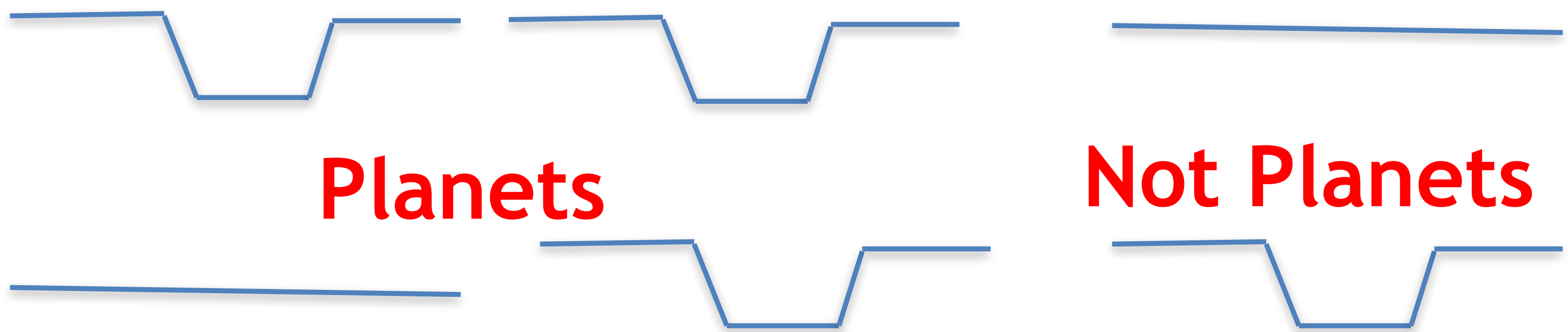


Huang et al , in preparation

Machine Learning 101

Kepler

HATNet

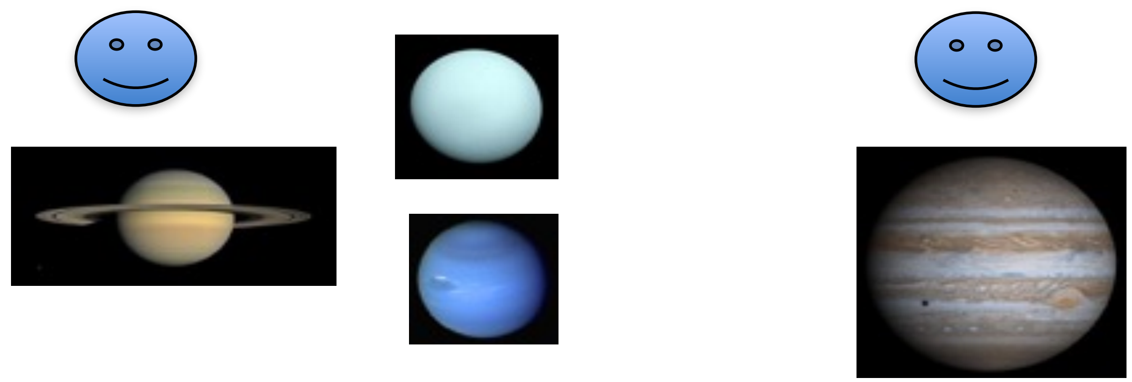


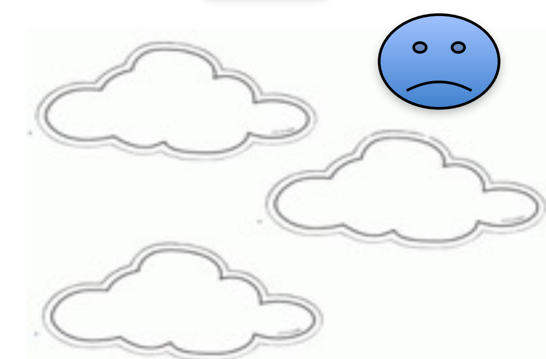
Planets

Not Planets

Data in Kepler

Property vector





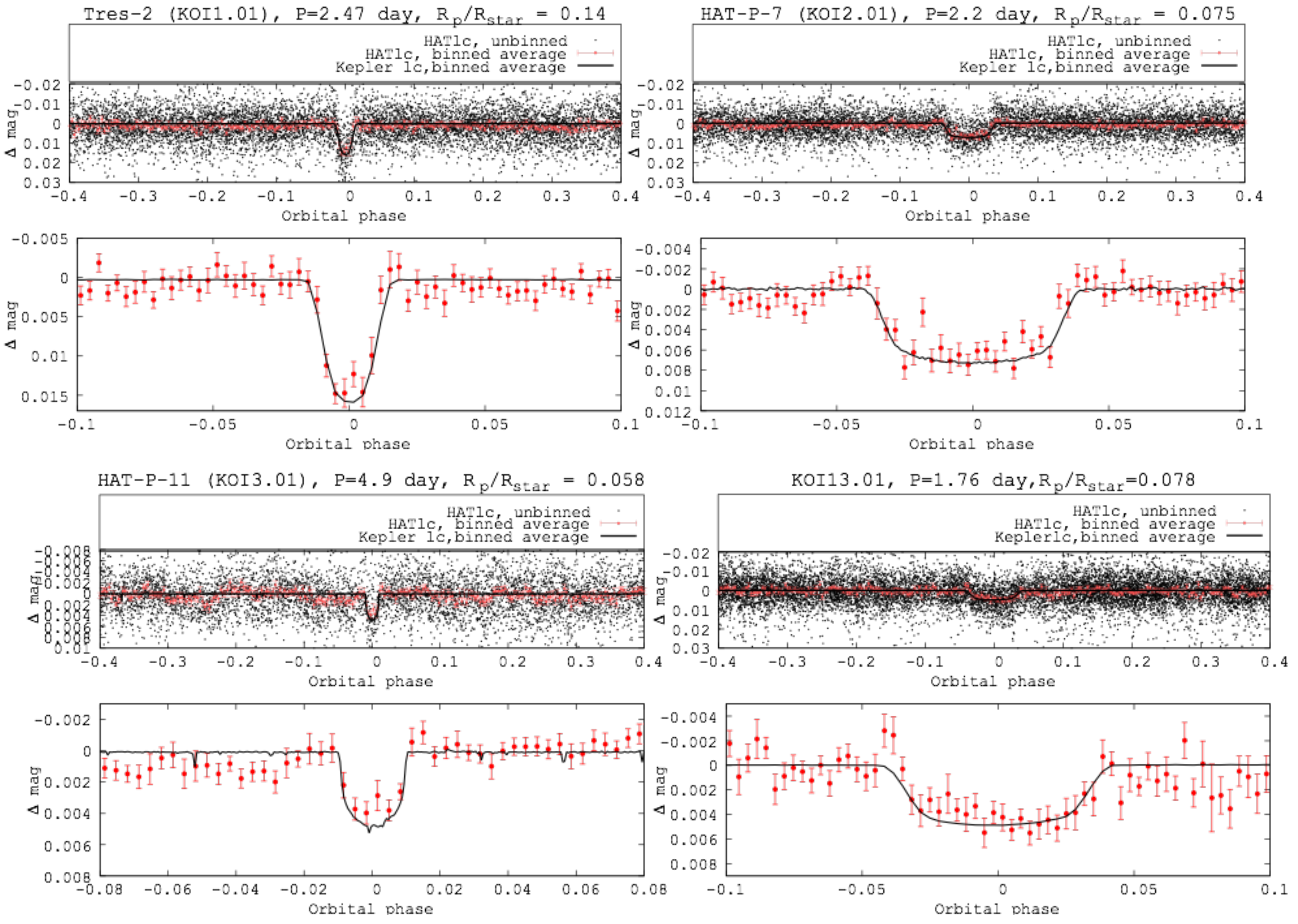
Training

Machine Learning Algorithm (Support Vector Machin

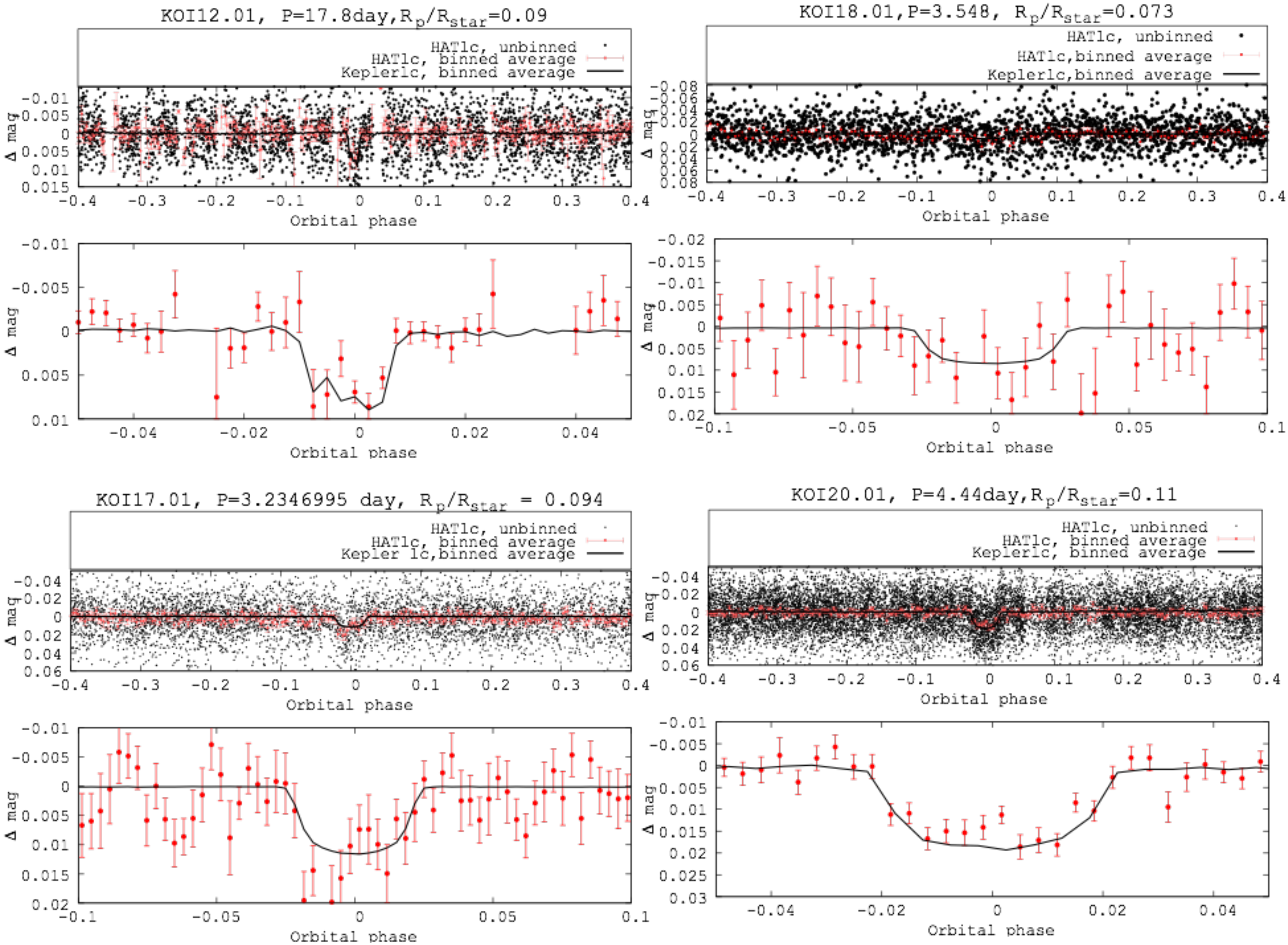
Classifying

Data not in Kepler

Potential planetary candidates subjected to follow



Tres-2 (KOI1.01), P=2.47 day, R p/R star = 0.14
HATlc, unbinned
HATlc, binned average
Kepler lc,binned average
Δ mag
Orbital phase
HAT-P-7 (KOI2.01), P=2.2 day, R p/R star = 0.075
HATlc, unbinned
HATlc, binned average
Kepler lc,binned average
Δ mag
Orbital phase
HAT-P-11 (KOI3.01), P=4.9 day, R p/R star = 0.058
HATlc, unbinned
HATlc, binned average
Kepler lc,binned average
Δ mag
Orbital phase
KOI13.01, P=1.76 day,R p/R star=0.078
HATlc, unbinned
HATlc, binned average
Keplerlc,binned average
Δ mag
Orbital phase



KOI12.01, P=17.8day,R_p/R_star=0.09
HATlc, unbinned
HATlc, binned average
Keplerlc, binned average
Δ mag
Orbital phase
KOI18.01,P=3.548, R_p/R_star=0.073
HATlc, unbinned
HATlc,binned average
Keplerlc,binned average
Δ mag
Orbital phase
KOI17.01, P=3.2346995 day, R_p/R_star = 0.094
HATlc, unbinned
HATlc, binned average
Kepler lc,binned average
Δ mag
Orbital phase
KOI20.01, P=4.44day,R_p/R_star=0.11
HATlc, unbinned
HATlc, binned average
Keplerlc, binned average
Δ mag
Orbital phase

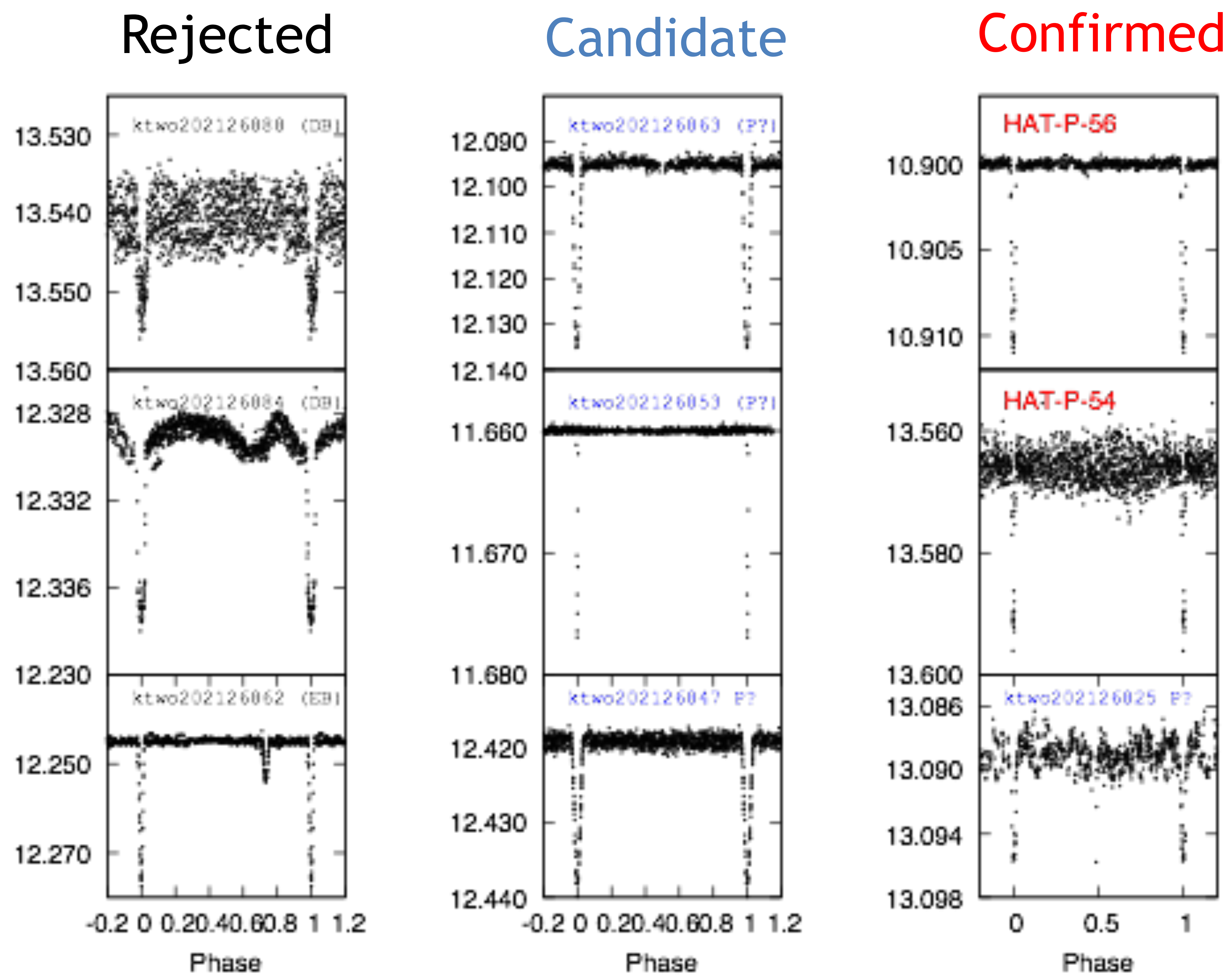
Center field (red) confusion matrix:

13.3% False Alarm rate

93.4% completeness rate (in selection process)

117 (TT)	8 (TF)
18 (FT)	1743(FF)

Follow up Selected HAT candidate with K2



Deep primary eclipse- Rejected based on K2 light curves