Extrasolar ring-like candidates detection

Data analysis and statistical methods implementation

Student : Marco Cilibrasi
Supervisors : Dr Joel Hartman , Prof. Gaspar Bakos

Princeton University - Department of Astrophysical Sciences

October 17, 2016 - Summer Student Exam

PRINCETON UNIVERSITY

Table of contents

Introduction

Transiting planets

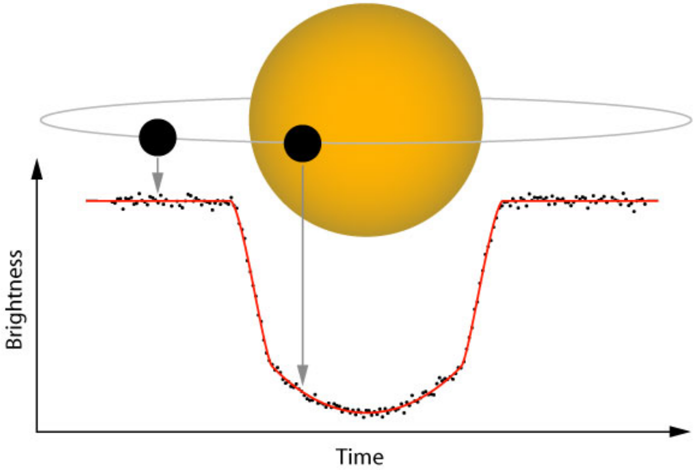


Figure: Scheme of an exoplanet transit.

HATNet & HATSouth

Figure: HATNet telescopes at FLWO, open for servicing during the day, and the HATNet telescope mount, camera, and lens. Photo credit: Gaspar Bakos.

Principal problems

- non-professional instruments lead to noisy and imperfect data

  $\rightarrow$ lot of data analysis work and algorithms

- we can detect only big ring-systems, but within the *Hill sphere*

  $\rightarrow$ $R_H = a\left(\frac{M_p}{3M_\star}\right)^{1/3}$ $\Rightarrow$ a very big, very long period

$\Rightarrow$ long period means no periodicity:

$\rightarrow$ we don't have repeated events to distinguish from random error dips

$\rightarrow$ we can't predict next transits for better instruments $\rightarrow$ literature

First lightcurve analysis

Data collection, FITMag & EPDMag

Typical data file:

- sky divided into 200 fields $\rightarrow$ star sorted by brightness
- # of the instrument and # of the image $\rightarrow$ time in HJD
- 3 different apertures $\rightarrow$ magnitude, standard error and flag (G,X)
- 3 different stage of filtering: FITMag, EPDMag, TFAMag

FITMag $\Rightarrow$ zero-point correction based on all the stars

EPDMag $\Rightarrow$ remove variations correlated with changes in instrumental parameters

HAT-144-0000490.tfalc

```
10-82908    54521.9526236 11.32249  0.00586  G 11.06558  0.00569  G 10.96616  0.00590  G  10.8429   10.524  10.4061 10.88871 10.88776 10.89038
10.88605        10.88706        10.88662
10-82909    54521.9565627 11.17933  0.00503  G 10.92483  0.00481  G 10.82089  0.00490  G  10.8401  10.5238  10.4008 10.89393 10.88957 10.88861
10.89049        10.88611        10.88510
10-82910    54521.9604747 11.26497  0.00566  G 11.03258  0.00559  G 10.93098  0.00578  G  10.8155  10.5109  10.3869 10.87186 10.87510 10.87437
10.88354        10.87826        10.87214
10-82911    54521.9643753 11.15917  0.00504  G 10.90547  0.00484  G 10.80519  0.00495  G  10.8524   10.533  10.4108 10.90126 10.89844 10.89649
10.89496        10.89869        10.89631
10-82912    54521.9682874 11.14182  0.00490  G 10.90015  0.00471  G 10.79288  0.00478  G  10.8057  10.5073   10.384 10.86921 10.87094 10.87081
10.87027        10.87384        10.87310
10-82913    54521.9721997 11.40891  0.00675  G 11.16120  0.00673  G 11.06266  0.00708  G   10.823  10.5152  10.3982 10.87351 10.88230 10.88758
10.88394        10.89039        10.89117
10-82914    54521.9761117 11.40852  0.00668  G 11.16710  0.00667  G 11.06308  0.00698  G   10.822  10.5205  10.3996 10.88131 10.88433 10.88482
10.88478        10.88040        10.88023
10-82915    54521.9800123 11.20629  0.00515  G 10.97644  0.00501  G 10.87859  0.00514  G   10.812  10.5114    10.39 10.86838 10.87250 10.87397
10.87261        10.87371        10.88038
10-82916    54521.9839126 11.04593  0.00475  G 10.82107  0.00463  G 10.72770  0.00476  G  10.8107  10.5112  10.3926 10.87301 10.87571 10.87733
10.87888        10.87634        10.87827
10-82917    54521.9878252 11.20761  0.00509  G 10.94314  0.00483  G 10.83604  0.00491  G   10.828  10.5127  10.3932 10.87383 10.87738 10.87867
10.88282        10.87910        10.88054
10-82918    54521.9917254 11.29551  0.00554  G 11.06969  0.00546  G 10.98099  0.00567  G  10.8409  10.5367  10.4213 10.89429 10.89449 10.89644
10.90088        10.90084        10.90239
10-82919    54521.9956376 11.46176  0.00663  G 11.23874  0.00668  G 11.14451  0.00702  G  10.8255  10.5275  10.4072 10.87476 10.88683 10.89128
10.88153        10.88720        10.89687
10-82920    54521.9995382 11.20571  0.00492  G 10.97659  0.00474  G 10.87767  0.00481  G  10.8112  10.5096  10.3877 10.86911 10.87116 10.87155
10.86727        10.87398        10.87247
10-82921    54522.0034502 11.57808  0.00734  G 11.34502  0.00740  G 11.24658  0.00779  G   10.841   10.536  10.4132 10.89337 10.89120 10.88984
10.88998        10.88879        10.88448
10-82922    54522.0073508 11.61879  0.00707  G 11.35867  0.00692  G 11.26056  0.00723  G  10.8363  10.5147  10.3959 10.87518 10.87373 10.87342
10.88435        10.88311        10.88089
```

Figure: Typical data file, this is the 490^{th} brightest star in the 144^{th} field

TFAMag

- target time-series $Y(i); i = 1, 2, ..., N$
- current best estimate (average first) $A(i); i = 1, 2, ..., N$
- template set $X_j(i); i = 1, 2, ..., N; j = 1, 2, ..., M$

  $\Rightarrow \quad F(i) = \sum_{j=1}^{M} c_j X_j(i)$

$$\text{minimize} \quad D = \sum_{i=1}^{N} [Y(i) - A(i) - F(i)]^2$$

- $g_{j,k} = \sum_{i=1}^{N} X_j(i) X_k(i) \;\; ; \;\; j,k = 1,2,...,M$
- $h_j = \sum_{i=1}^{N} \bar{Y}(i) X_j(i)$ where $\bar{Y}(i) \equiv Y(i) - A(i)$
- Minimizing D $\Rightarrow$ $c_j = \sum_{k=1}^{M} g_{j,k}^{-1} h_k$

$$\Rightarrow \quad A(i) = Y(i) - \sum_{k=1}^{M} c_k X_k(i) \quad \text{untill} \quad \sigma^2 < 10^{-3}$$

where $\sigma^2 = \frac{D}{N-M}$

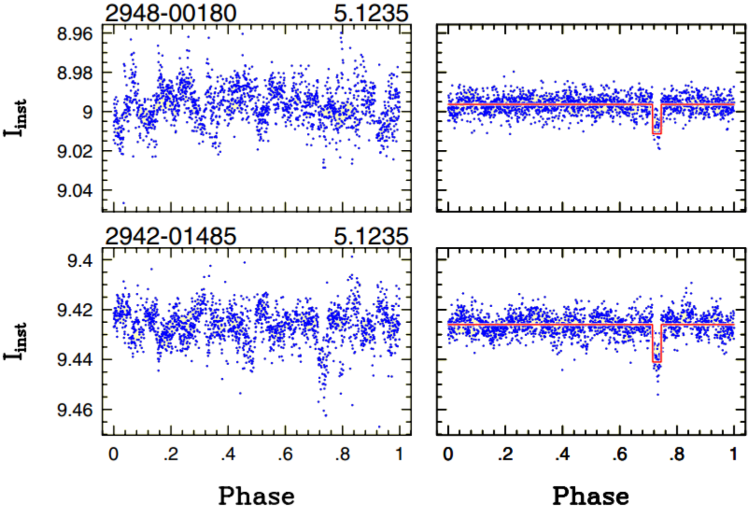


Figure: *A trend filtering algorithm for wide-field variability surveys*

Searching algorithms

BLS algorithm

- Strictly periodic signal x_i, period P_0
- Binned time series $\hat{x}_i \to n_{bin}$ bins
- Only two discrete values, H and L
- Time spent in the transit phase L is qP_0
- Zero-average time series $\Rightarrow \quad H = -L\frac{q}{1-q}$

Folded time series $\bar{x}_i$ for a given trial period P.
Step function with lower level $\hat{L}$ at $[i_1, i_2]$

Minimize reduced least square

$$D = \frac{1}{N}\left[\sum_{i=1}^{i_1}(\bar{x}_i - \hat{H})^2 + \sum_{i=i_1}^{i_2}(\bar{x}_i - \hat{L})^2 + \sum_{i=i_2}^{n}(\bar{x}_i - \hat{H})^2\right]$$

$$\text{Minimization} \quad \Rightarrow \quad \hat{L} = \frac{s}{r} \quad ; \quad \hat{H} = -\frac{s}{1-r}$$

$$\text{where} \quad r = \frac{i_2 - i_1}{N} \quad \text{and} \quad s = \frac{1}{N}\sum_{i=i_1}^{i_2} \bar{x}_i$$

$$\Rightarrow \quad D = \frac{1}{N}\sum_{i=1}^{n} \bar{x}_i^2 - \frac{s^2}{r(1-r)}$$

$$\text{then we search for} \quad SR = MAX\left\{\sqrt{\frac{s^2(i_1,i_2)}{r(i_1,i_2)[1-r(i_1,i_2)]}}\right\}$$

$$\Rightarrow \quad D_{min} = \frac{1}{N}\sum_{i=1}^{N} x_i^2 - SR_{MAX}^2$$

Interesting parameters

- $\delta \equiv H - L \quad \Rightarrow \quad SNR = \delta/\sigma(x_i)$
- $SDE = \frac{SR_{peak} - <SR>}{\sigma(SR)}$

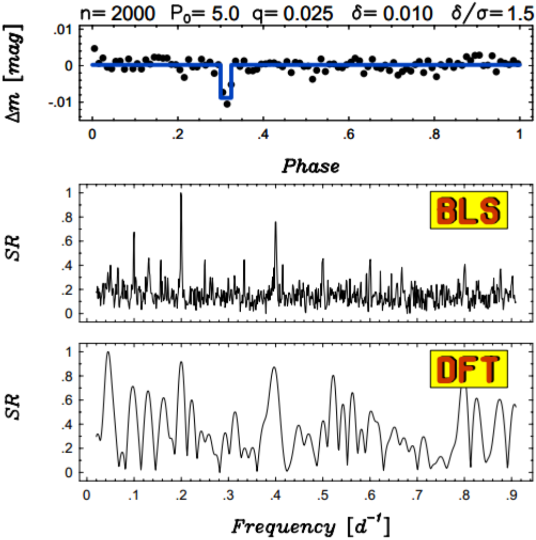
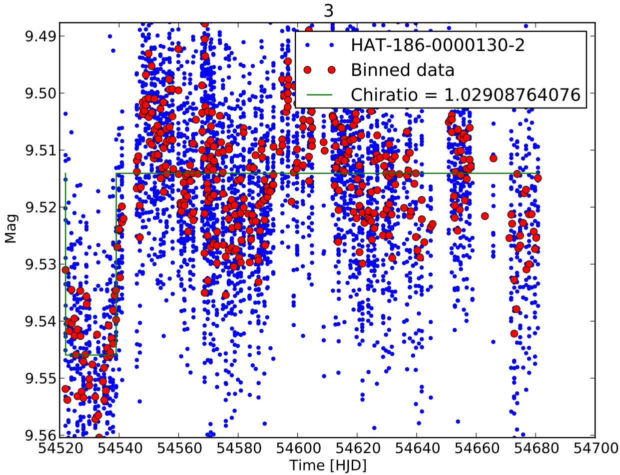


Figure: *A box-fitting algorithm in the search for periodic transits*

In our case → no periodic events ⇒ new parameter

$$\chi^2_{lin} = \frac{1}{N}\sum_{i=1}^{N}[x_i - mean(\mathbf{x})]^2 \quad \Rightarrow \quad \chi^2_{ratio} = \frac{\chi^2_{lin}}{D_{min}}$$



Histogram algorithm

BLS $\rightarrow$ flag transit points $\Rightarrow$ count transit occurences

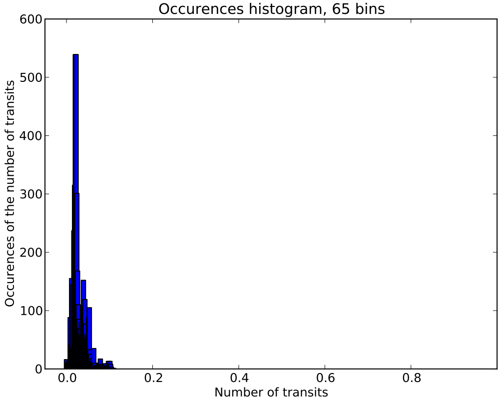


Figure: TFA histogram

- x axis : number of flagged transits in a certain image (%)
- y axis : number of images with that number of transits (tot number)

$\rightarrow$ Poisson distribution $\Rightarrow$ threshold cut (~ 0.2)

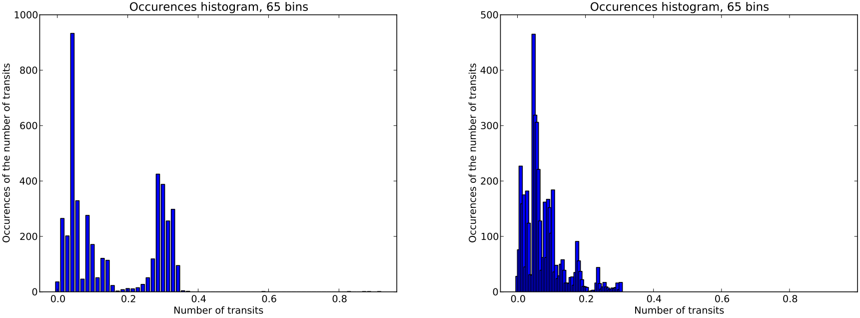


Figure: FIT & EPD histogram

PCA algorithm

X_{ij} is the j^{th} (over P) observed point of the i^{th} (over N) lightcurve

Correlation matrix $\Rightarrow$ $S_{jk} = \frac{1}{N}\sum_{i=1}^{N} X_{ji}X_{ki}$

Find eigenvectors sorted by eigenvalue $\mathbf{u}^t$ and define $PC_t(i) = \sum_{j=1}^{P} X_{ij}u_j^t$

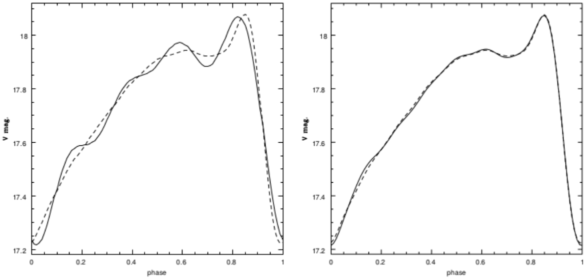


Figure: Reproduction example: Fourier vs PCA, 9^{th} order

Two possible solutions:

1. subtraction : $X_{ij}^{fin} = X_{ij} - \sum_{k=1}^{n} < \mathbf{X}_i, \mathbf{u}^k > u_j^k$
2. correlation and cut : $C_i^t =< \mathbf{X}_i, \mathbf{u}^t >= \sum_{j=1}^{P} X_{ij} u_j^t$

We adopt the second solution, cutting over an arbitrary threshold

Future developments

Ring transit modelling

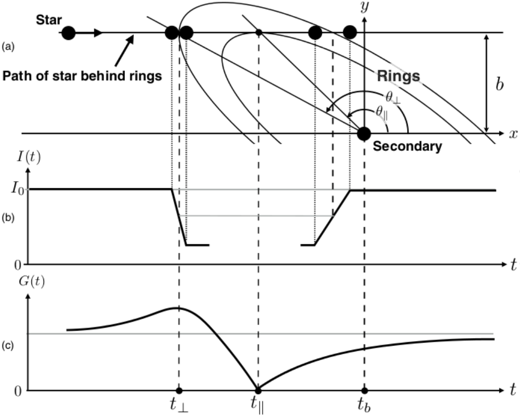


Figure: Panel (a) shows a ring system inclined with a star passing behind it. Panel (b) shows the resultant light curve $I(t)$. Panel (c) highlights the three significant epochs in the rate of change of ring radius $G(t) = \frac{dr}{dt}$. [3]

$G(t)$ as a function of geometrical parameters, measured $\frac{dI(t)}{dt} \equiv \frac{dI(r)}{dr} G(t)$

$$\text{best fit} \Rightarrow \text{geometry} \ \& \ \frac{dI(r)}{dr} \rightarrow \text{opacity } \tau(r)$$

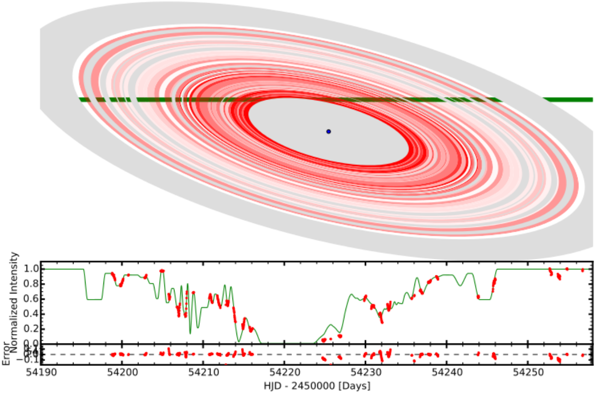


Figure: Model ring fit to $J1407$ data. The lower graph shows the model transmitted intensity $I(t)$. [3]

Thank you for your attention

Bibliography

G. Bakos et al.
Wide-Field Millimagnitude Photometry with the HAT: A Tool for Extrasolar Planet Detection (2004)
the Astronomical Society of the Pacific, **116**, 266-277

G. Kovacs et al.
A trend filtering algorithm for wide-field variability surveys (2005)
Mon. Not. R. Astron. Soc., **356**, 557-567

G. Kovacs et al.
A box-fitting algorithm in the search for periodic transits (2002)
A&A, **391**, 369-377

Bibliography

S. M. Kanbur et al.
The Use of Principal Components Analysis in Analysing Variable Star Data (2000)
The Impact of Large-Scale Surveys on Pulsating Star Research, ASP Conference Series, **203**, 56-59

S. M. Kanbur & H. Mariani
Principal Component Analysis of RR Lyrae light curves (2004)
Mon. Not. R. Astron. Soc.

M. A. Kenworthy & E. E. Mamajek
Modelling giant extrasolar ring systems in eclipse and the case of J1407B: sculpting by exomoons? (2015)
The Astrophysical Journal, **800**, 126-135