
Observational Astronomy: Exoplanet Transit Parameters



L-Università
ta' Malta

Daniele Caruana

Group 8

PHY2235

B.Sc.(Hons) in Physics

23rd March 2023

Abstract

The transit method describes one of the most effective methods to detect exoplanets using the brightness of stars. In this paper, an overview of the batman library was shown to calculate light curves for different planets. The MCMC method was shown to be able to be effective in obtaining accurate estimates for the transit parameters for Kepler-1b, given observational data.

1 INTRODUCTION & THEORETICAL BACKGROUND

1.1 Exoplanets and the Transit Method

Exoplanets are planets residing in systems outside of our own solar system, orbiting a star or a stellar remnant, or rogue planets orbiting the center of the galaxy. There are several methods to detect the existence of an exoplanet, the most effective being the radial-velocity method, and the Transit method.

The Transit method uses an exoplanet's transit around its observed star in order to discover the exoplanet. When a planet's orbit is at the points at which it is in front of the star, from Earth's perspective, the brightness of the star is observed to dim. This change in brightness (or relative flux) can be measured using photometric instruments, more specifically charged-couple devices, and illustrated using light curves. Several estimates of the characteristics of a planet can be determined by observed light curves, such as the orbital period, the radius of the planet, and the inclination.

1.2 Limb Darkening Effect

Limb darkening is the visually apparent diminished brightness at the edge of a star, when compared to its centre. This effect significantly alters the shape of the light curves of an observed star, hence showing the true depth of the transit, giving a better estimate of the radius of the exoplanet. There are several laws (Espinoza & Jordán, 2015) describing limb darkening, used for different types of stars and light-curves, some of which are given as follows:

$$\begin{aligned}
 \text{Uniform law : } I(\mu) &= I_0 \\
 \text{Linear law : } I(\mu) &= I_0(1 - u_1[1 - \mu]) \\
 \text{Quadratic law : } I(\mu) &= I_0(1 - u_1[1 - \mu] - u_2[1 - \mu]^2) \\
 \text{Logarithmic law : } I(\mu) &= I_0(1 - l_1(1 - \mu) - l_2\mu \ln\mu),
 \end{aligned} \tag{1}$$

where $\mu = \cos\theta$, and θ is the angle between the observer's perspective plane, and the normal to the stellar surface.

1.3 Hot Jupiters

Hot Jupiters are a type of exoplanet located outside our solar system which have a similar size to Jupiter, however orbit their star at a very close distance, giving them extraordinarily high temperatures which can exceed a thousand degrees Celsius. Due to this, the interaction between Hot Jupiters' atmosphere and the intense solar radiation is extremely strong, causing strong atmospheric winds that can reach speeds of up to several kilometers per second, and high levels of atmospheric ionization.

Due to Hot Jupiters' massive size and short semi-major axis, their detection and viability of research is much easier when compared to smaller planets. This also facilitates the practice of transit spectroscopy as it allows more light to pass through the planets' atmosphere. Further more, due to their rapid orbit, their transit can be measured more frequent in a set time.

2 MARKOV CHAIN MONTE CARLO (MCMC) METHODS

MCMC describes a type of computation used to obtain accurate samples from a selected distribution. This is used in scenarios where simply obtaining independent samples does not compute results accurately.

Monte Carlo methods are statistical computations which are used to estimate the parameters of a given model. This is done by obtaining random samples from a probability distributions. While this can work for certain simple functions, this will not be sufficient for higher-order functions. Markov Chains are memory-less systems that produce a sequence of states, each one depending only on the active probabilistic properties. MCMC is a Bayesian process which uses a combination of these two methods, meaning that it is dependent on a probability distribution which is based on a set of ranges, called priors, assumed to be true by the analyst. The process works by having a number of walkers which start from an initial set of parameters, θ . Each walker then jumps to a new set of parameters, θ_{trial} , within the previously set priors. The new values of θ_{trial} are then compared to the experimentally observed/measured data using a function which calculates a value for the likeliness of the new trial parameters, such as a χ^2 -type check given by:

$$\ln\ell(\theta|x_1, x_2, \dots, x_n) = \frac{1}{2} \sum_i^n \left(\frac{y_{obs_i} - y_{model_i}}{y_{err_{obs,i}}} \right)^2. \tag{2}$$

If the new parameters satisfy the likelihood function, the current state's parameters will be considered as the current parameters, and the probabilistic determination for the next jump will only depend on these current parameters. This process is iterated for a number of times until it is determined that a convergence has been achieved within the distribution.

Besides using the samples to obtain accurate estimates of some parameters, samples can also be used in order to obtain a posterior probability distribution. Posterior distributions describe the uncertainty of our achieved model based on the likelihood and the prior distributions. From the known cumulative distribution function of a normal distribution, one can obtain three values for sigma or standard deviations, 68%, 95%, and 99.7%, in order to compare obtained values with the interval estimate of the obtained normal distribution.

3 DATA ANALYSIS AND RESULTS

This paper displays three different analyses: light curve modelling of Kepler-90 planets using the batman package, fitting a curve to a dataset using MCMC, and fitting a transit curve to an observed dataset of Kepler-1b.

3.1 Exoplanet transit light curve modelling for Kepler-90 planets

Provided data containing transit and physical parameters of all Kepler-90 planets was imported using the pandas library. These parameters consist of the planetary radius (R_p), the orbital period (P), the orbital inclination (i), the semi-major axis (a), and the eccentricity (e). It was made sure that the values for the planetary radius, and the semi-major axis were converted into units of stellar radii, $\frac{R_p}{R_\star}$ and $\frac{a}{R_\star}$ respectively, where $R_\star \approx 1.25 R_\odot$ (Stassun et al., 2019).

Planet	R_p / R_\star	P /days	a / R_\star	$i / ^\circ$	e
Kepler-90 b	9.61E-03	7.008	1.273E+01	89.400	0.000
Kepler-90 c	8.65E-03	8.719	1.531E+01	89.680	0.000
Kepler-90 d	2.11E-02	59.737	5.505E+01	89.710	0.000
Kepler-90 e	1.96E-02	91.939	7.225E+01	89.790	0.000
Kepler-90 f	2.12E-02	124.914	8.257E+01	89.770	0.010
Kepler-90 g	5.96E-02	210.607	1.221E+02	89.920	0.049
Kepler-90 h	8.30E-02	331.601	1.737E+02	89.927	0.011
Kepler-90 i	9.68E-03	14.449	1.841E+01	89.200	0.000

The batman (Kreidberg, 2015) package was first used in order to generate light curves for all planets, using a uniform limb darkening profile, in which the time of inferior conjunction (t_0), and the longitude of periastron (w) were set to be $t_0 = w = 0$. Then, the light curves for Kepler-90 e were plotted using both a linear and a quadratic limb darkening profile at different values of the limb darkening coefficients u_1 for the linear profile, and u_1 and u_2 for the quadratic profile.

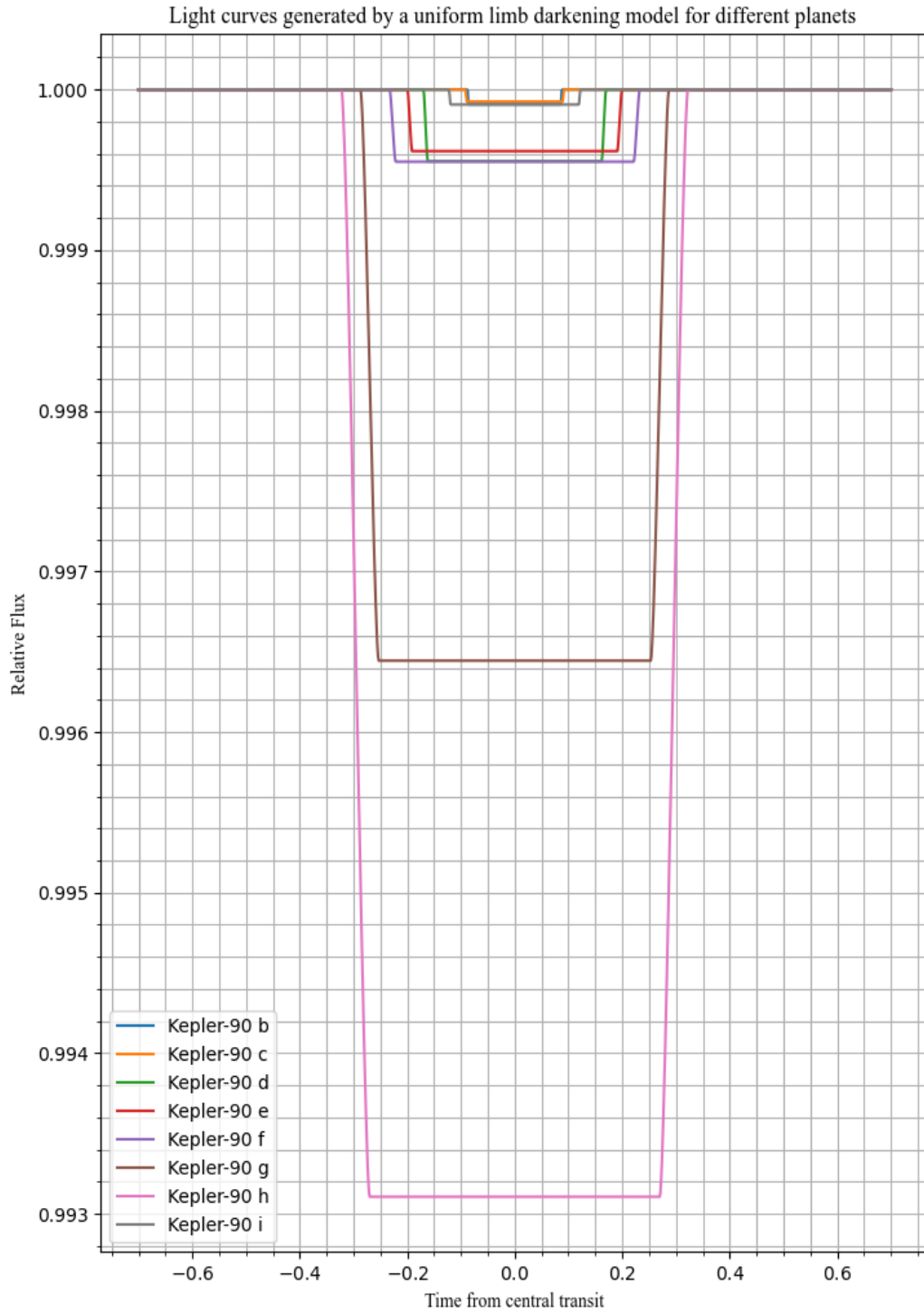


Fig. 1: Light curves of the planets of Kepler-90 using a uniform limb darkening profile

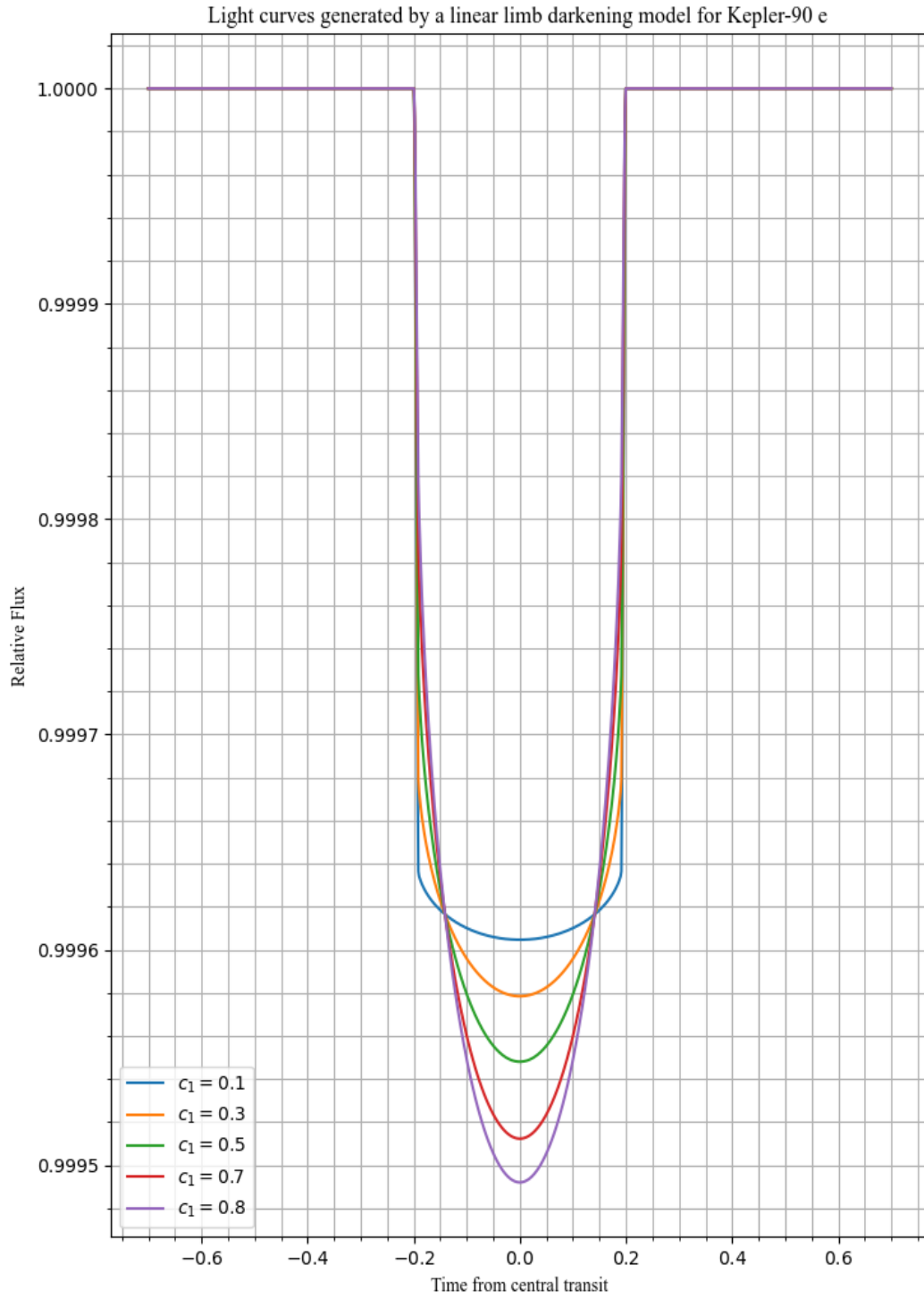


Fig. 2: Light curves of Kepler-90 e using a linear limb darkening profile

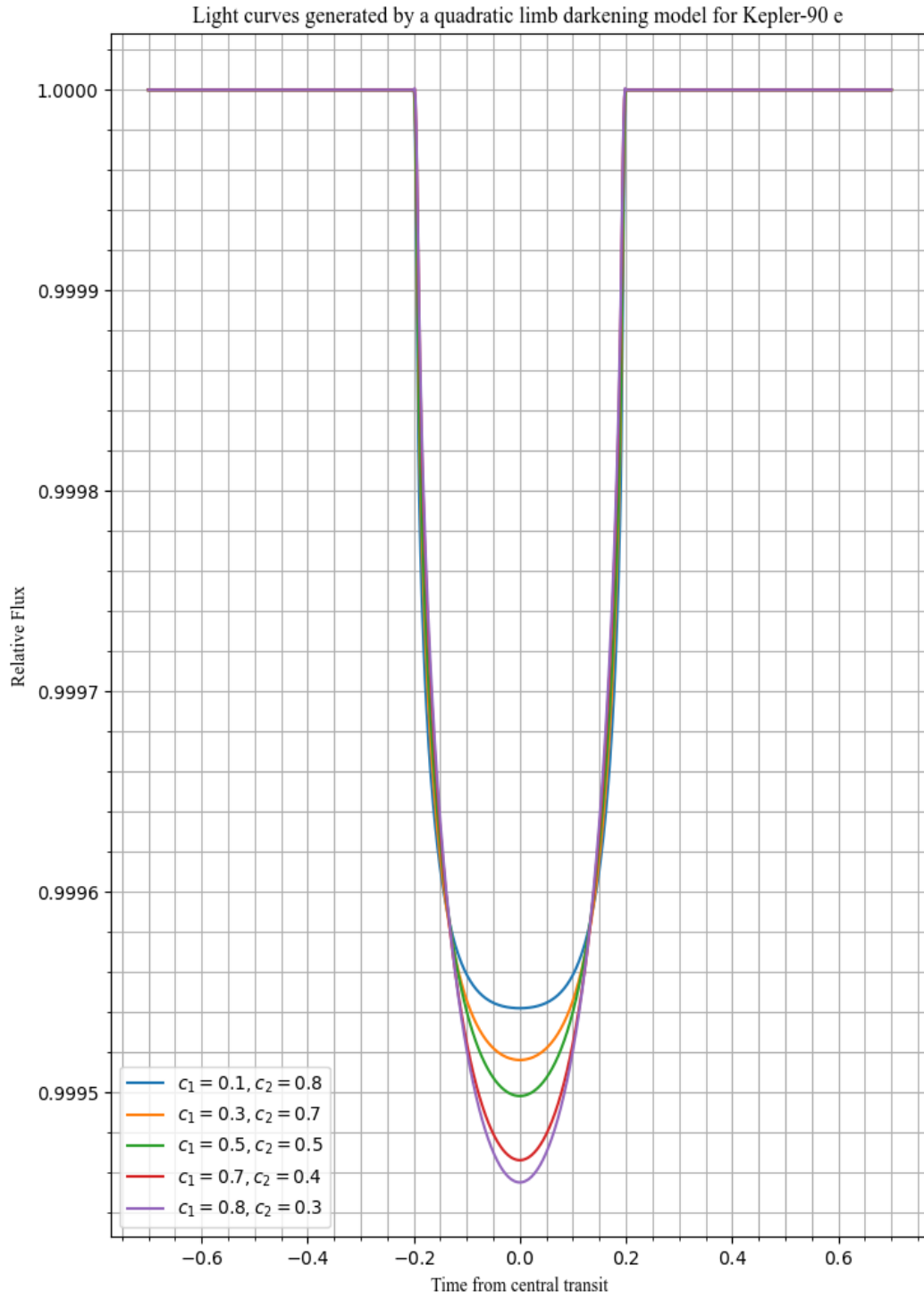


Fig. 3: Light curves of Kepler-90 e using a quadratic limb darkening profile

3.2 Using MCMC to fit a sine curve model to data

The dataset for the X and Y axes was first imported using the pandas library, and a model for the general function of a sine curve,

$$y = a + b\sin(cx + d), \quad (3)$$

was defined. The imported data was then imported in order to obtain a prior range for the y-intercept a , and the amplitude b .

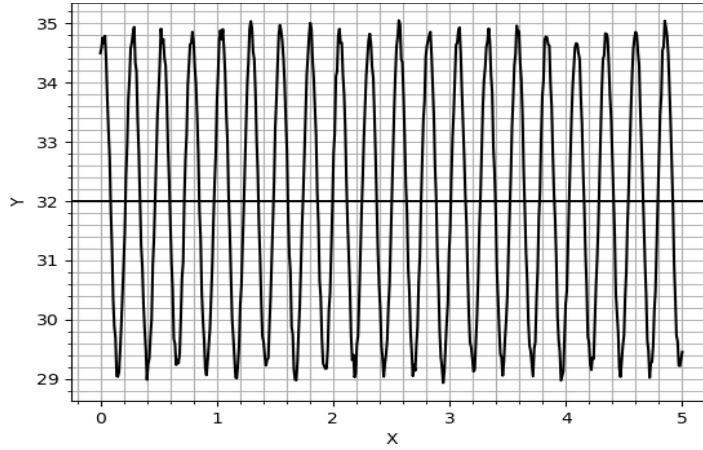


Fig. 4: Plot of the curve data

Using this plot and referencing to the provided values, a function was defined which checks whether the arguments of θ , where $\theta = [a, b, c, d]$, are within their prior ranges. This was done by outputting $-\text{inf}$ if one of the inputted parameters does not satisfy its respective range, and 0 if all parameters are within their respective ranges. The prior ranges for these parameters were set to be

$$\begin{aligned} 30 &\leq a \leq 33.5, \\ 2.5 &\leq b \leq 3.5, \\ 23 &\leq c \leq 26, \\ 6 &\leq d \leq 11. \end{aligned} \quad (4)$$

Another function was defined in order to calculate the likelihood of each parameter using equation 2. This was then referred to another function, called the probability function, which returns the value outputted by the prior function added to the value returned by the likelihood function, such that if all parameters are within their respective prior ranges, the value outputted is the value returned by the likelihood function.

Finally, a function to calculate the step methodology across the parameter space, p_0 was created, followed by the function used to run the MCMC algorithm, using the emcee library, and all the other previously defined functions. This was done using 100 walkers with 3000 iterations.

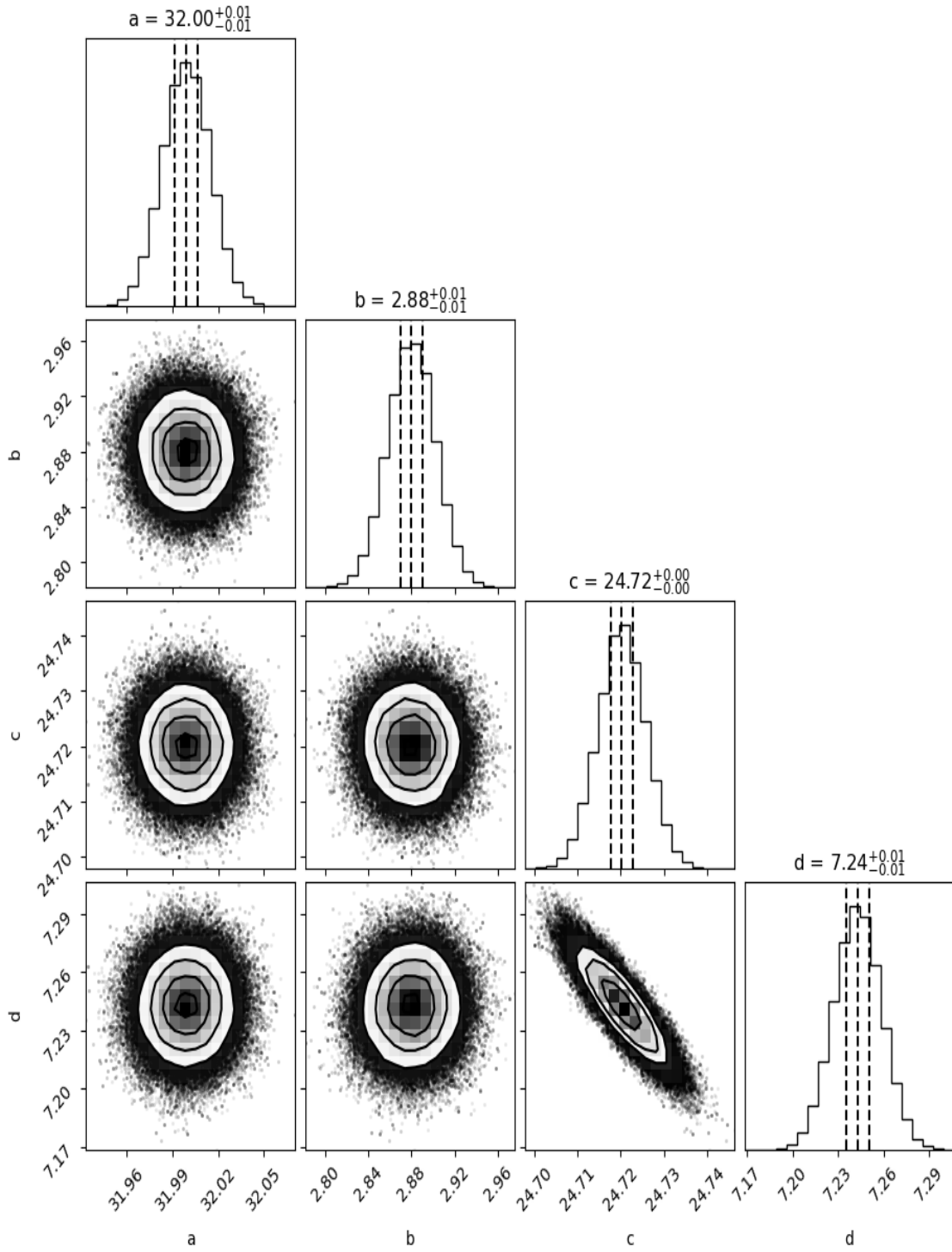


Fig. 5: A corner plot for the parameters of the sine curve

Hence, it was determined that the provided data can be fit using the model in eq. 3, where $a = 32.00 \pm 0.01$, $b = 2.88 \pm 0.01$, $c = 24.72 \pm 0.00$, and $d = 7.24 \pm 0.01$.

3.3 Fitting a transit curve to an observed dataset of Kepler-1b

The data for the relative flux and the respective Heliocentric Julian Date was imported using the pandas library. This was then plotted in order to get an assumption for the time of inferior conjunction, t_0 . Moreover, an estimate for the radius of the planet was obtained from the plot and hence,

$$\Delta\phi = \frac{R_p^2}{R_\star^2} \quad (5)$$

$$\frac{R_p^2}{R_\star^2} \approx 0.12.$$

As per (Holman et al., 2007), the limb darkening coefficients were taken to be $c_1 = 0.22$, and $c_2 = 0.32$. It was provided that the orbital period of this planet is $P \approx 2.16 \times 10^5$ s, and the mass of the star is $M_\star \approx 1.96 \times 10^{30}$. Hence, using the Kepler's third law, the estimated value of the semi-major axis was found as follows:

$$P^2 = \frac{4\pi^2}{G \cdot (M_\star + M_{planet})} a^3 \quad (6)$$

$$a \approx 7.713R_\odot,$$

where the gravitational constant was considered to be $G = 6.67 \times 10^{-11} \text{m}^3 \text{kg}^{-1} \text{s}^2$. MCMC functions were then defined in a similar manner to those in section 3.2, using the batman parameters as a model as described in section 3.1. Prior ranges were then set with reference to the previously obtained values. The range for the orbital inclination was taken to be $60^\circ \leq 90^\circ$.

```

1 def model(theta, x):
2     t0, per, rp, a, inc = theta
3     params = bat.TransitParams()
4     params.w = 0
5     params.ecc = 0
6     params.limb_dark = 'quadratic'
7     params.u = [0.22, 0.32]
8     params.t0 = t0
9     params.per = per
10    params.rp = rp
11    params.a = a
12    params.inc = inc
13
14    m = batman.TransitModel(params, x)
15    rel_flux = m.light_curve(params)
16    return rel_flux

```

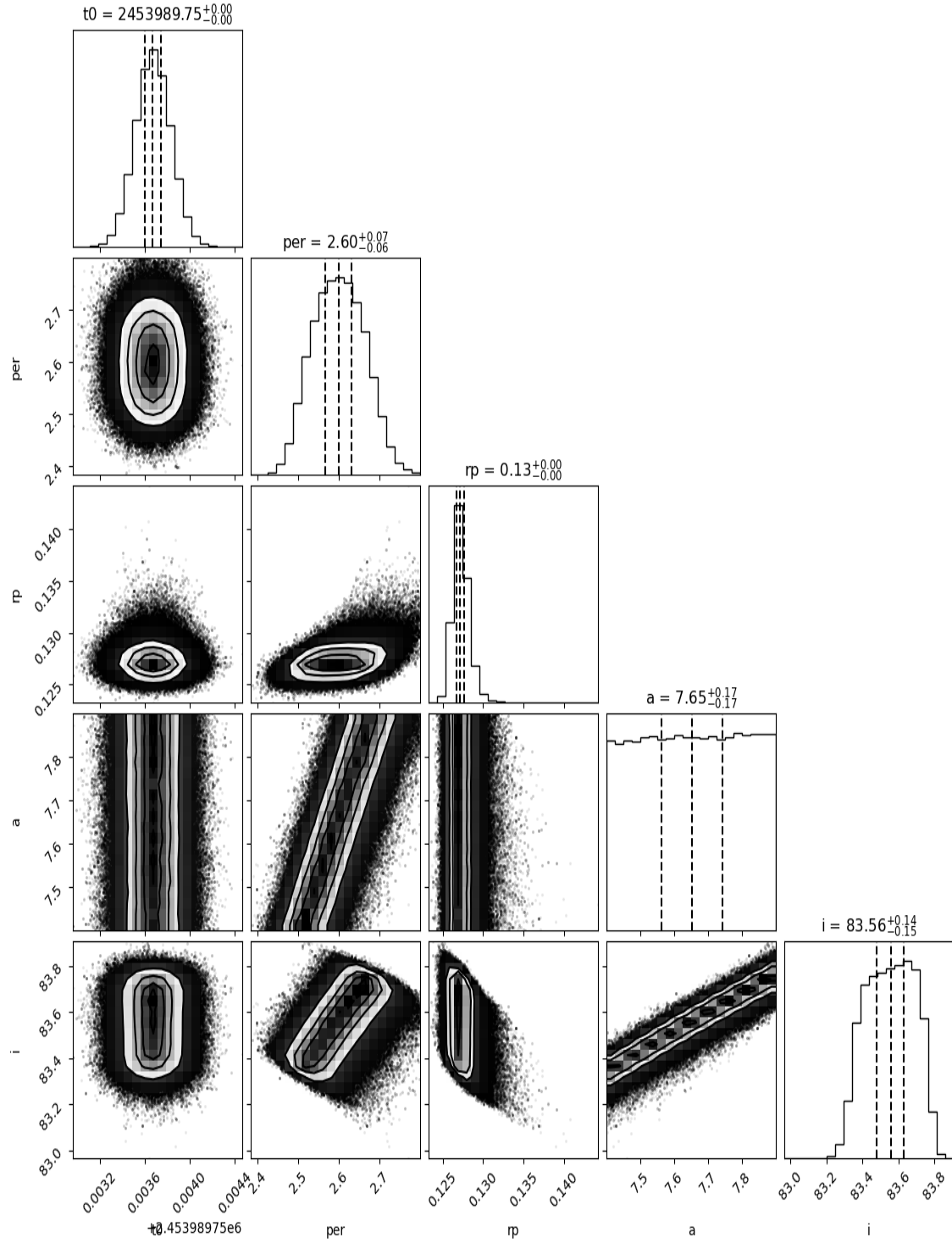


Fig. 6: A corner plot for the parameters of the transit of Kepler-1b

4 DISCUSSION

4.1 Exoplanet transit light curve modelling for Kepler-90 planets

In this section, a light curve was modelled according to different planetary and transit parameters for all the planets in the Kepler-90 system. As can be seen in figure 1, the light curves for the different planets have different shapes, where the dips in relative flux represent the transits of the planet when it is between the observer and the star. Hence it can be said that the planets that are represented to have a wider transit, spend a longer time between the star and the observer. Furthermore, the depth of the dips represent by how much the brightness of the star was observed to decrease, which hints towards the radius of the planet in transit. In the linear and quadratic curves, fig. 2 and fig. 3 respectively, clearly show that an increase in the values of c_1 , increase the depth of the dips in relative flux. Furthermore, the shapes and accuracies of the light curve change significantly according to different values of the limb darkening coefficients, describing better the transit of exoplanets.

4.2 Using MCMC to fit a sine curve model to data

In this section, a model for a sine curve with the general form $y = a + b(\sin(cx + d))$, was fitted to provided data for the X and Y axes, using the emcee library to run the MCMC algorithm. A corner plot was then plotted showing the values of the parameters obtained at convergence, as well as the correlation between them. It can be seen by the the correlation plots in 5, that the only correlation present is an inverse relation between c and d , which can be seen by the lack of circularity in the c-d plot.

4.3 Fitting a transit curve to an observed dataset of Kepler-1b

In this section, MCMC and batman were used in order to obtain accurate estimates for different planetary and orbital parameters of Kepler-1b. A corner plot was then plotted showing the values of the different obtained parameters. By referring to fig. 6, it can be seen that the following results were obtained:

$$t_0 \approx 2453989.75 \pm 0.00, \text{in HJD},$$

$$P \approx 2.54 \leq 2.60 \leq 2.67, \text{in days},$$

$$R_p \approx 0.13 \pm 0.00, \text{in } R_\star,$$

$$a \approx 7.65 \pm 0.17, \text{in } R_\star,$$

$$i \approx 83.42 \leq 83.56 \leq 83.71, \text{in } ^\circ.$$

However, as can be seen in the plot for a , the value didn't converge. This could be due to inaccurate priors, insufficient walkers, or insufficient iterations. Furthermore, the correlation plots show how the different parameters affect each other. While most parameters seem to have little correlation to each other, a positive relation seems to exist between a and P , and a and i . However, due to the lack of convergence for a , this could prove to be inaccurate. A 1σ posterior plot was then plotted to show the deviation of the results.

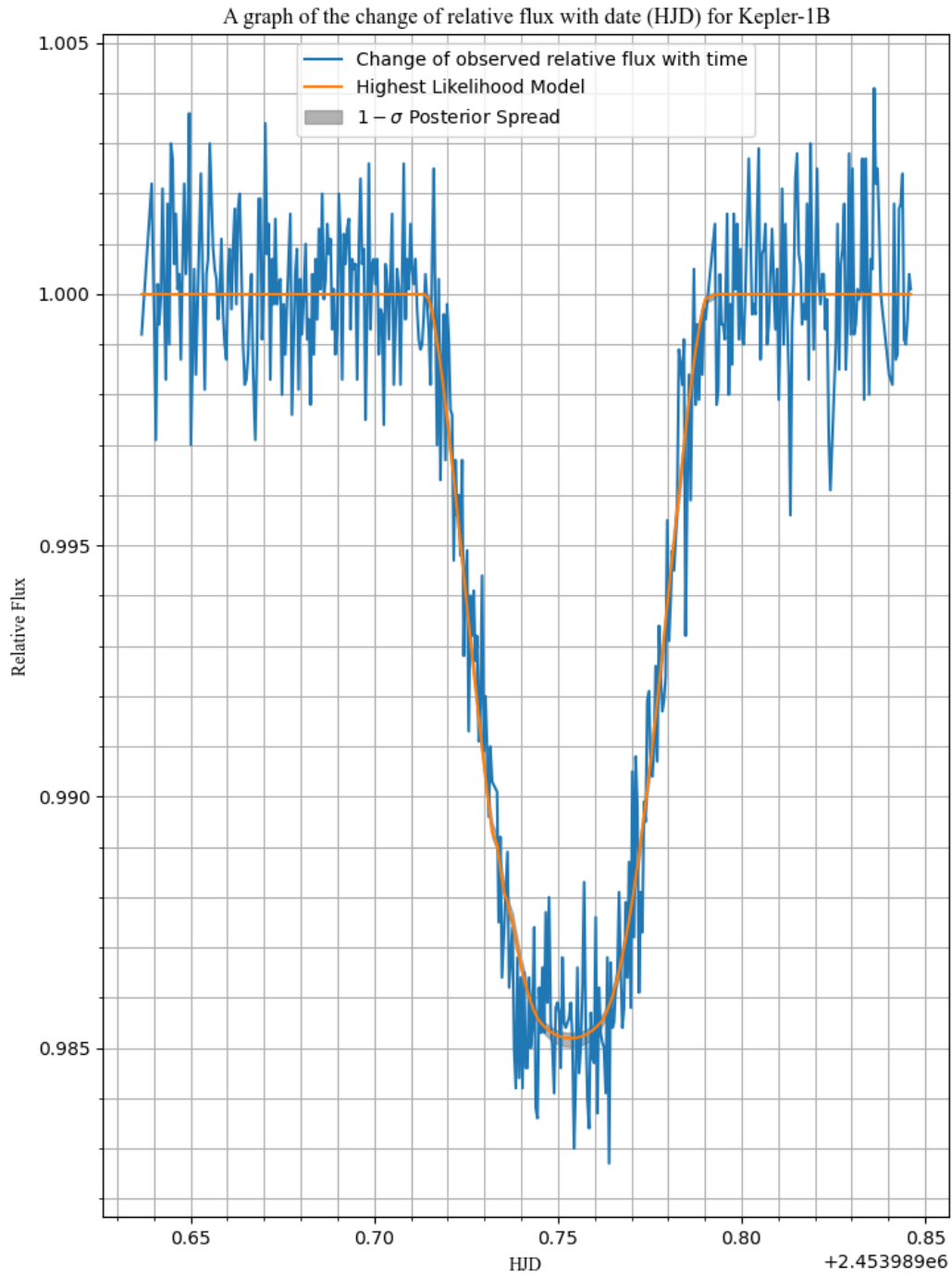


Fig. 7: A 1σ spread with the highest likelihood model

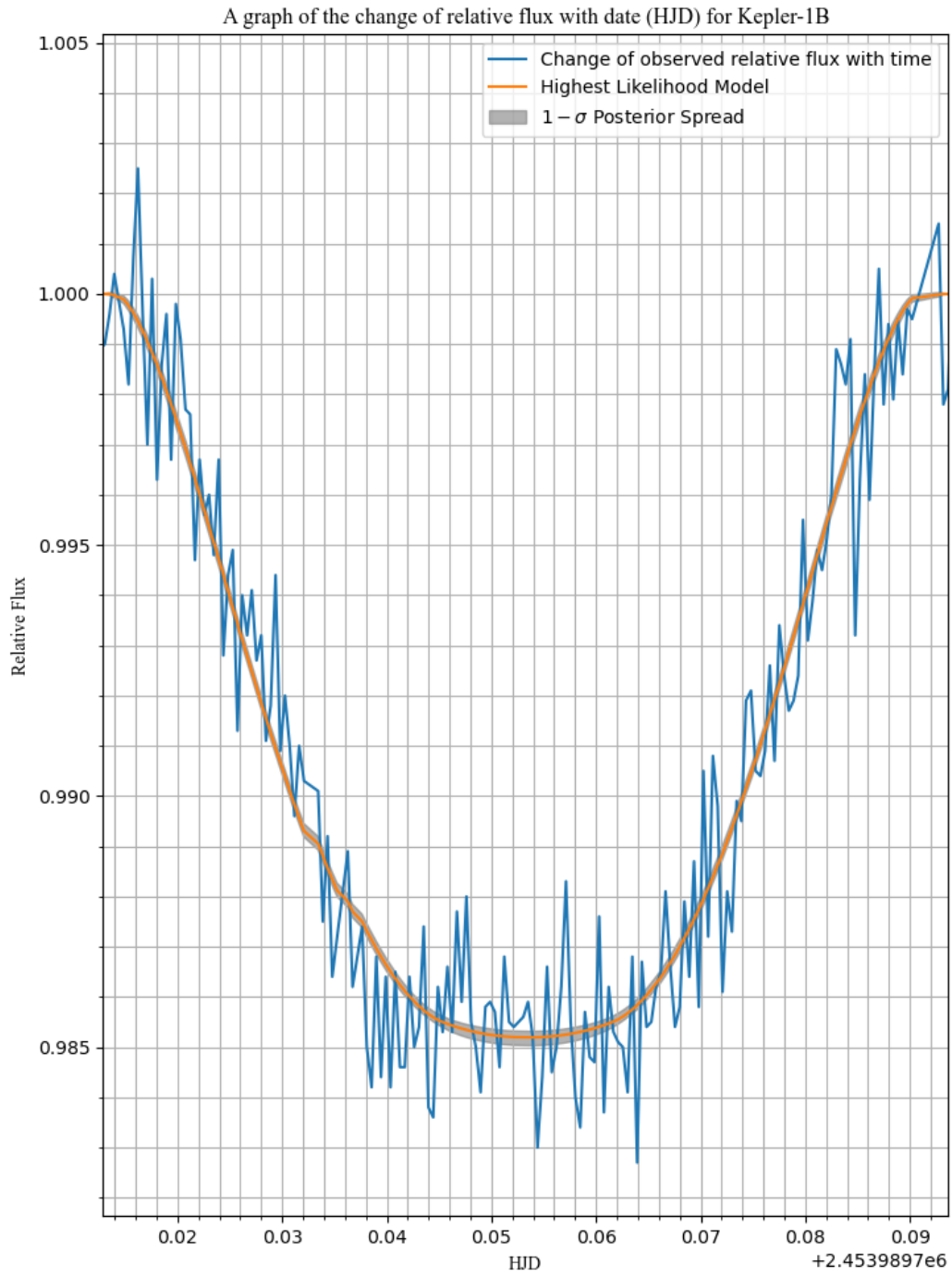


Fig. 8: A partial view of the 1σ spread with the highest likelihood model

5 CONCLUSION

In this experiment, an overview of the batman package was shown by calculating light curves for the planets of the Kepler-90 system and how different parameters affect their light curves. Furthermore a brief overview of MCMC was shown by showing how a model can be fitted to some observed data. Lastly, these two methods were combined in order to calculate estimates of the planetary and orbital parameters of Kepler-1b by fitting a batman model to some observed relative flux data. This is useful as more accurate observed data can be fitted to these models in order to show better properties of far away systems.

REFERENCES

- Espinoza, N & Jordán, A (2015) Limb darkening and exoplanets: testing stellar model atmospheres and identifying biases in transit parameters. *Monthly Notices of the Royal Astronomical Society*, 450(2) 1879–1899. <https://doi.org/10.1093/mnras/stv744>
- Holman, M J Winn, J N Latham, D W O'Donovan, F T Charbonneau, D Torres, G Sozzetti, A Fernandez, J & Everett, M E (2007) The transit light curve (tlc) project. vi. three transits of the exoplanet tres-2. *The Astrophysical Journal*, 664(2) 1185–1189. <https://doi.org/10.1086/519077>
- Kreidberg, L (2015) batman: BAsic Transit Model cAlculationN in Python., 127(957) 1161. <https://doi.org/10.1086/683602>
- Stassun, K G Oelkers, R J Paegert, M Torres, G Pepper, J De Lee, N Collins, K Latham, D W Muirhead, P S Chittidi, J Rojas-Ayala, B Fleming, S W Rose, M E Tenenbaum, P Ting, E B Kane, S R Barclay, T Bean, J L Brassuer, C E ... Winn, J N (2019) The Revised TESS Input Catalog and Candidate Target List., 158(4) Article 138, 138. <https://doi.org/10.3847/1538-3881/ab3467>

APPENDIX

1 CODE LISTING

1.1 Light curve modelling using batman for Kepler-90

```
1 import pandas as pd
2 import numpy as np
3 import batman as bat
4 import matplotlib.pyplot as plt
5 import matplotlib.ticker as mticker
6
7
8 def planet_param(arr):
9     planet_dic = {}
10    names = arr.index
11    for index, param in enumerate(arr):
12        planet_dic[names[index]] = param
13    return planet_dic
14
15
16 def transit(planet, profile, c1=0.0, c2=0.0):
17     parameters = bat.TransitParams()
18     parameters.t0 = 0
19     parameters.w = 0
20     if profile == 'u':
21         parameters.limb_dark = 'uniform'
22         parameters.u = []
23     elif profile == 'l':
24         parameters.limb_dark = 'linear'
25         parameters.u = [c1]
26     elif profile == 'q':
27         parameters.limb_dark = 'quadratic'
28         parameters.u = [c1, c2]
29     parameters.rp = planet['Rp / Rs']
30     parameters.a = planet['a / R']
31     parameters.inc = planet['Orbital Inclination (deg)']
32     parameters.ecc = planet['Eccentricity']
33     parameters.per = planet['Orbital Period (yr)'] * 365
34     return parameters
35
36
```

```
37 # Importing data
38
39 df = pd.read_csv('kepler_90_pl.csv')
40 df.set_index('Planet', inplace=True)
41 k_90_rad = 1.25 * 695700
42
43 # Modifying data and units
44
45 df['Rp / Rs'] = df['Planetary Radius (Earth Radius)'] * 6378.14 / k_90_rad
46 df['a / R'] = df['Semimajor Axis (AU)'] * 1.496e8 / k_90_rad
47
48 # Initializing arrays
49 time_arr = np.linspace(-0.7, 0.7, 1000)
50 planet_arr = np.array(df.index.values.tolist())
51 c1 = [0.1, 0.3, 0.5, 0.7, 0.8]
52 c2 = [0.8, 0.7, 0.5, 0.4, 0.3]
53
54 # Generating uniform models for all planets
55
56 fig, ax = plt.subplots(1)
57 ax.figure.set_size_inches(8.27, 11.69)
58 for i in planet_arr:
59     m = bat.TransitModel(transit(df.loc[i], 'u'), time_arr)
60     rel_flux = m.light_curve(transit(df.loc[i], 'u'))
61     plt.plot(time_arr, rel_flux, label=i)
62
63 csfont = {'fontname': 'Times New Roman'}
64 plt.ylabel(r"Relative Flux", **csfont)
65 plt.xlabel(r"Time from central transit", **csfont)
66 plt.title(r"Light curves generated by a uniform limb darkening model for different planets "
67           , **csfont)
68 plt.legend()
69 plt.minorticks_on()
70 plt.grid(which='both')
71 formatter = mticker.ScalarFormatter(useMathText=True)
72 ax.xaxis.set_major_formatter(formatter)
73 plt.savefig('uniform.png')
74 plt.show()
75 # Generating a linear model
76
```



```
77 fig, ax = plt.subplots(1)
78 ax.figure.set_size_inches(8.27, 11.69)
79 for i in c1:
80     m = bat.TransitModel(transit(df.loc['Kepler-90 e'], 'l', i), time_arr)
81     rel_flux = m.light_curve(transit(df.loc['Kepler-90 e'], 'l', i))
82     plt.plot(time_arr, rel_flux, label=r"$c_1 = {}".format(i))
83
84 csfont = {'fontname': 'Times New Roman'}
85 plt.ylabel(r"Relative Flux", **csfont)
86 plt.xlabel(r"Time from central transit", **csfont)
87 plt.title(r"Light curves generated by a linear limb darkening model for Kepler-90 e ", **
88          csfont)
89 plt.legend()
90 plt.minorticks_on()
91 plt.grid(which='both')
92 formatter = mticker.ScalarFormatter(useMathText=True)
93 ax.xaxis.set_major_formatter(formatter)
94 ax.get_yaxis().get_major_formatter().set_useOffset(False)
95 plt.savefig('linear.png')
96 plt.show()
97
98 # Generating a quadratic model
99
100 fig, ax = plt.subplots(1)
101 ax.figure.set_size_inches(8.27, 11.69)
102 for i in range(len(c1)):
103     m = bat.TransitModel(transit(df.loc['Kepler-90 e'], 'q', c1[i], c2[i]), time_arr)
104     rel_flux = m.light_curve(transit(df.loc['Kepler-90 e'], 'q', c1[i], c2[i]))
105     plt.plot(time_arr, rel_flux, label=r"$c_1 = {}, c_2 = {}".format(c1[i], c2[i]))
106
107 csfont = {'fontname': 'Times New Roman'}
108 plt.ylabel(r"Relative Flux", **csfont)
109 plt.xlabel(r"Time from central transit", **csfont)
110 plt.title(r"Light curves generated by a quadratic limb darkening model for Kepler-90 e ", **
111          csfont)
112 plt.legend()
113 plt.minorticks_on()
114 plt.grid(which='both')
115 formatter = mticker.ScalarFormatter(useMathText=True)
116 ax.xaxis.set_major_formatter(formatter)
117 ax.get_yaxis().get_major_formatter().set_useOffset(False)
```

```
116 plt.savefig('quadratic.png')
117 plt.show()
118
119 print("Debug")
```

1.2 Using MCMC to fit a sine model to data

```
1 import pandas as pd
2 import matplotlib.pyplot as plt
3 import numpy as np
4 import emcee
5 import corner
6 import multiprocessing
7
8
9 def model(theta, x):
10     a, b, c, d = theta
11     model_val = a + (b * (np.sin((c * x)+d)))
12     return model_val
13
14
15 def ln_like(theta, x, y, y_err):
16     ln_like_val = -0.5 * np.sum(np.square((y - model(theta, x)) / y_err))
17     return ln_like_val
18
19
20 def ln_prior(theta):
21     a, b, c, d = theta
22     if 30 <= a <= 33.5 and 2.5 <= b <= 3.5 and 23 <= c <= 26 and 6 <= d <= 11:
23         return 0
24     else:
25         return -np.inf
26
27
28 def ln_prob(theta, x, y, y_err):
29     lp = ln_prior(theta)
30     if not np.isfinite(lp):
31         return -np.inf
32     return lp + ln_like(theta, x, y, y_err)
33
34
35 def _p0(initial_val, n_dim_val, n_walkers_val):
```

```
36     return [np.array(initial_val) + [1e-4, 75e-3, 1e-2, 1e-2] * np.random.randn(n_dim_val)
37             for i in
38                 range(n_walkers_val)]
39
40 def mcmc(p0, n_walkers, n_iter, n_dim, ln_prob, data_tup, pool):
41     sampler_val = emcee.EnsembleSampler(n_walkers, n_dim, ln_prob, args=data_tup, pool=pool)
42     print("Running burn-in...")
43     p0, _, _ = sampler_val.run_mcmc(p0, 800, progress=True)
44     sampler_val.reset()
45
46     print("Running production...")
47     posteriors_val, prob_val, state_val = sampler_val.run_mcmc(p0, n_iter, progress=True)
48     print("Production Done")
49     return sampler_val, posteriors_val, prob_val, state_val
50
51
52 df = pd.read_csv('mcmc_intro_group_5.csv')
53 data = (df['x'], df['y'], df['y_err'])
54
55 # Value Checking from plot
56
57 plt.plot(df['x'], df['y'])
58 plt.axhline(y=32)
59 plt.minorticks_on()
60 plt.grid(which='both')
61 plt.xlabel("X")
62 plt.ylabel("Y")
63 plt.savefig("Sine")
64 plt.show()
65
66 n_walkers = 100
67 n_iter = 3000
68
69 initial = np.array([32.5, 3, 24, 8])
70 n_dim = len(initial)
71
72 p0 = _p0(initial, n_dim, n_walkers)
73
74 if __name__ == '__main__':
75     with multiprocessing.Pool(4) as pool:
```

```
76     sampler, posteriors, prob, state = mcmc(p0, n_walkers, n_iter, n_dim, ln_prob, data,
77     pool)
78
79     # Corner Plot
80     labels = ['a', 'b', 'c', 'd']
81     print("Done")
82     fig, ax = plt.subplots(1)
83     ax.figure.set_size_inches(8.27, 11.69)
84     fig = corner.corner(samples, show_titles=True, labels=labels, plot_datapoints=True,
85     quantiles=[0.32, 0.5, 0.68])
86     plt.savefig("corner_sine.png")
87
88 print("Debug")
```

1.3 Using MCMC and batman to fit a light curve model to data

```
1 import batman
2 import pandas as pd
3 import numpy as np
4 import batman as bat
5 import emcee
6 import multiprocessing
7 import corner
8 import matplotlib.pyplot as plt
9
10
11 def model(theta, x):
12     t0, per, rp, a, inc = theta
13     params = bat.TransitParams()
14     params.w = 0
15     params.ecc = 0
16     params.limb_dark = 'quadratic'
17     params.u = [0.22, 0.32]
18     params.t0 = t0
19     params.per = per
20     params.rp = rp
21     params.a = a
22     params.inc = inc
23
24     m = batman.TransitModel(params, x)
```

```
25     rel_flux = m.light_curve(params)
26     return rel_flux
27
28
29 def ln_like(theta, x, y, y_err):
30     ln_like_val = -0.5 * np.sum(np.square((y - model(theta, x)) / y_err))
31     return ln_like_val
32
33
34 def ln_prior(theta):
35     t0, per, rp, a, inc = theta
36     if (0.71 + 2.453989e6 <= t0 <= 0.78 + 2.453989e6) and (2.1 <= per <= 2.8) and (0.09 <=
37         rp <= 0.15) and (
38         7.4 <= a <= 7.9) and 60 <= inc <= 90:
39         return 0
40     else:
41         return -np.inf
42
43 # 0.785, 0.72 = 0.065
44 # p squared = a cubed
45 # rp/rs = change in flux root
46 def ln_prob(theta, x, y, y_err):
47     lp = ln_prior(theta)
48     if not np.isfinite(lp):
49         return -np.inf
50     return lp + ln_like(theta, x, y, y_err)
51
52
53 def _p0(initial_val, n_dim_val, n_walkers_val):
54     return [np.array(initial_val) + [1e-3, 1e-4, 1e-3, 1e-4, 0.1] * np.random.randn(
55         n_dim_val) for i in
56         range(n_walkers_val)]
57
58 def mcmc(p0, n_walkers, n_iter, n_dim, ln_prob, data_tup, pool):
59     sampler_val = emcee.EnsembleSampler(n_walkers, n_dim, ln_prob, args=data_tup, pool=pool)
60     print("Running burn-in...")
61     p0, _, _ = sampler_val.run_mcmc(p0, 800, progress=True)
62     sampler_val.reset()
63
```

```
64     print("Running production...")
65     posteriors_val, prob_val, state_val = sampler_val.run_mcmc(p0, n_iter, progress=True)
66     print("Production Done")
67     return sampler_val, posteriors_val, prob_val, state_val
68
69
70 def posteriors_func(samples_post, chain, x):
71     models = []
72     drw = np.floor(np.random.uniform(0, len(chain), size=samples_post)).astype(int)
73     th_s = chain[drw]
74     for i in th_s:
75         mdl_post = model(i, x)
76         models.append(mdl_post)
77     spread = np.std(models, axis=0)
78     med_mod = np.median(models, axis=0)
79     return med_mod, spread
80
81
82 if __name__ == '__main__':
83     with multiprocessing.Pool(4) as pool:
84         df = pd.read_csv("kepler_lc_group_4.csv")
85
86         time_arr = np.linspace(np.min(df['HJD'].to_numpy()), np.max(df['HJD'].to_numpy()),
87                                433)
88
89         data = (time_arr, df['Rel_Flux'], df['Flux_err'])
90
91         n_walkers = 250
92         n_iter = 5000
93
94         initial = np.array([0.75 + 2.453989e6, 2.5, 0.12, 7.71, 90])
95         n_dim = len(initial)
96
97         p0 = _p0(initial, n_dim, n_walkers)
98
99         sampler, posteriors, prob, state = mcmc(p0, n_walkers, n_iter, n_dim, ln_prob, data,
100         pool)
101         samples = sampler.flatchain
102         print("Done")
103
104         # Corner Plot
```

```
103     labels = ['t0', 'per', 'rp', 'a', 'i']
104
105     fig, ax = plt.subplots(1)
106     ax.figure.set_size_inches(8.27, 11.69)
107     corner.corner(samples, show_titles=True, labels=labels, plot_datapoints=True,
108 quantiles=[0.32, 0.5, 0.68])
109     plt.savefig("cornerbatman.png")
110     plt.show()
111
112     # Plotting 1-sigma posterior
113
114     th_max = samples[np.argmax(sampler.flatlnprobability)]
115     best_fit = model(th_max, time_arr)
116     med_mod_val, spread_val = posteriors_func(600, samples, time_arr)
117
118     fig, ax = plt.subplots(1)
119     ax.figure.set_size_inches(8.27, 11.69)
120     csfont = {'fontname': 'Times New Roman'}
121     plt.plot(df['HJD'], df['Rel_Flux'], label='Change of observed relative flux with
122 time')
123     plt.plot(df['HJD'], best_fit, label='Highest Likelihood Model')
124     plt.fill_between(df['HJD'], med_mod_val-spread_val, med_mod_val+spread_val, color='
125 grey', alpha=0.6,
126 label=r'$1-\sigma$ Posterior Spread')
127     plt.xlabel("HJD", **csfont)
128     plt.ylabel("Relative Flux", **csfont)
129     plt.title("A graph of the change of relative flux with date (HJD) for Kepler-1B", **
130 csfont)
131     plt.legend()
132     plt.minorticks_on()
133     plt.grid(which='both')
134     plt.savefig("PosteriorPlot_full")
135     plt.show()
136     print('Theta:', th_max)
```