Draft version November 20, 2020
Typeset using LaTeX **twocolumn** style in AASTeX62

Characterisation of Transit Timing Variations in Overlapping Fields of the K2 Mission

Michael Porritt[1] and Ben Montet[1]

[1]*School of Physics, University of New South Wales, Australia*

ABSTRACT

The overlapping fields of campaigns 5, 16 and 18 of the K2 mission allowed many confirmed exoplanets and promising candidates to be observed for a second or even a third time, over a period of around three years. The transit times of such exoplanets can be quantified and can inform about the nature of the systems they occupy, especially over such a long timescale. Here, we analyse a set of 181 systems flagged by previous investigations and characterise the transit timing variations (TTVs) of their exoplanets. We find a substantial proportion of candidates that exhibit potentially meaningful TTVs. We also produce a smaller list of candidates with very strong evidence of meaningful, periodic TTVs, which indicates the presence of a nearby perturber.

Keywords: Transit Timing Variations

1. INTRODUCTION

1.1. *Background*

The K2 mission was developed following failure of the reaction wheels on board the Kepler Spacecraft. The new mission observed fields in the ecliptic plain for approximately 75 days at a time (Howell et al. 2014). Campaigns 5, 16 and 18 of the K2 mission observed fields that significantly overlapped, meaning that a large number of stars were observed two or three times between April 2015 and July 2018 (K2- 2018). Such long timescales are rare in the realms of exoplanet detection, but repeated viewing permits the detection and confirmation of many more subtle signals and variations in stars and their satellites.

1.2. *Theory*

A notable signal that is usually only detectable over long time-scales is the change in the precise time of transit of a given body: In a perfect Keplerian system, the times that a body transits it's host star should be separated by exactly the period of the planet, a constant. However, the universe is chaotic, and satellites are constantly being pushed and pulled by nearby bodies, and stars themselves can move quite dramatically relative to the centre of mass of it's system. In some planets, these effects can be quite significant, and their exact times of transit can vary by hours from what we'd expect given only Keplerian mechanics. Thus it is useful to define the term transit timing variation (TTV) as the difference between the observed transit time of an exoplanet and the expected time of transit according to a Keplerian model.

TTVs were first proposed as a means of detecting non-transiting exoplanets as early as 2005, as well as a means of constraining the masses, radii and orbital elements of a system in the case of more than one transiting exoplanet (Agol et al. 2005; Holman & Murray 2005). More recently, a strong precedent for this kind of analysis is in Carter et al. (2012) and in Hamann et al. (2019). The Hamann paper in particular is the basis for the methods in much of this report: it's analysis centres entirely around the K2-146 system, a star with two transiting exoplanets with very short periods. The TTVs that are discovered are very dramatic - on the scale of hours - due to the proximity of the planets and the strong harmony of their orbits. Consequentially, they were able to determine the masses, radii and orbital parameters of this system with an otherwise unobtainable level of accuracy.

Transit timing variations can reveal aspects of a stellar system undetectable other methods, and are thus an important tool for exoplanetary analysis. The ways that TTVs can be utilised are well outlined in papers such as Lithwick et al. (2012), Agol & Fabrycky (2018) and the previously mentioned K2-146 paper (Hamann et al. 2019). Transit depth variations are mentioned in some of these, which is a related form of analysis but will not be considered in this report.

1.3. *Aims*

The main goals of this project will be to develop an algorithm to evaluate the transit timing variations in

exoplanets and to identify a list of systems that are good candidates for a more thorough analysis of their TTVs.

Specifically, we will be looking at a relatively small subset of systems that were observed in at least two of campaigns 5, 16 and 18. The list we will use is the union of two sets of targets approved for the K2 mission as part of the Guest Observer Program. The two programs are GO16011 and GO18048 (See K2 approved targets). See appendix A for the combined list of EPIC IDs. These targets are good candidates for TTVs because they were requested explicitly for the purpose of having multiple campaigns of data. Most of the systems have known or probable transits and were requested with the study of TTVs in mind. The list is also not too large, which is good for the sake of computational viability.

2. DATA AND METHODS

2.1. *Retrieval and Processing of Data*

The data I used is from the EVEREST pipeline, a distribution of Kepler data that performs pixel-level decorrelation to correct for variations in the Kepler spacecraft's precise orientation (Luger et al. 2016). We will use the general, long-cadence (30 minute) data for all our targets. Some of our candidates likely have short-cadence (1 minute) data however we will use only long-cadence for the sake of consistency and computational efficiency. Short-cadence data would be recommended to be used for candidates with strong TTVs.

Before the light curves for a given EPIC can be meaningfully interpreted, it is necessary to process the data. The light curves from EVEREST are quite variable over long time-scales, due to a combination of long-term fluctuations in the telescope and stellar variations. Conveniently, our only objective of interest in this report is the proportional decrease of intensity due to transits of exoplanets. Thus, we only care about the relative flux at any given time. Therefore, if we can quantify the long-term trends in the data, we can then divide by that trend and end up with flat light curves centred around a relative flux of 1. We use the `lightkurve` module for python to determine the long-term trends and subsequently flatten our data (Lightkurve Collaboration et al. 2018). An example of the raw data trend and the final result of flattening are shown in Figures 1 and 2.

The other operation that lends itself to the identification of transits is to trim the data. Extreme data points can cause problems down the line, and so we can remove some outliers. As discussed prior, exoplanet transits manifest as dips in the relative flux. We therefore do not care about any physical phenomena that cause significant increases in the flux - namely stellar flares (Maehara et al. 2012). Consequently, we remove any data points that are above some displacement from the median. Similarly, issues with the Kepler spacecraft drifting - particularly prevalent in later campaigns - are known to cause substantial drops in flux (K2- 2018). Therefore I also employ a lower limit, although this must be greater than the upper limit, since the last thing thing we want is to remove data points from the transits themselves. The upper and lower limits I chose are 5 standard deviations and 10 standard deviations respectively. Through some experimentation I found that these are more than accommodating of any transits due to exoplanets. Otherwise, any mistakenly removed data points should be easily identifiable further down the line. An example of the final result of both flattening and removal of outliers is shown in figure 2.

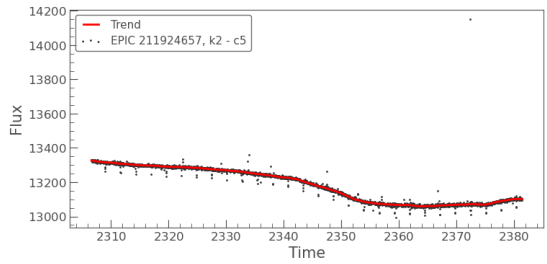

Figure 1. The light curve for K2-146 campaign 5. The data (in black) shows a clear long-term trend that is probably not physical, and is unwanted for the purposes of this report. In red is the trend computed by `lightkurve`.

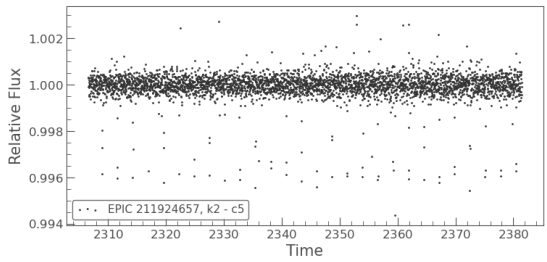

Figure 2. The light curve for K2-146 campaign 5. Flattened and outliers removed using `lightkurve`.

2.2. *Period Determination*

The period of a planets orbit is the first and foremost parameter needed to be able to reliably identify exoplanet transits. Therefore considerable thought must be given to the methods used. One of the most common techniques for finding transit periodicities is the Box Least Squares method. A detailed explanation of how the BLS method works can be found here. In short, BLS finds periodic transits by assuming a naive box-shaped transit. This simplifies the entire transit model

down to four parameters: the time of the first transit, the transit depth, the transit duration, and the period. This way we can very quickly compute the relativity probabilities of different transit periods as well as these other parameters. The result is a periodogram, a plot of a range of periods versus the associated probabilities (See figure 3). Therefore, all that we need to do to determine the period is to take the maximum-likelihood period.

In practice, we limit the available periods to a reasonable range; the limits used in this report are a minimum period of 0.7 days and a maximum of roughly half the total duration of the given campaigns. Greater bounds can be given to accommodate a greater range of periods, however since the aim of this campaign is to analyse TTVs, singular transits are not of interest. Similarly, as the period gets small, the number of transits we will have to fit for becomes very large, thus the lower bound. In this report, we compute Box Least Squares using the `Astropy` package for python (Astropy Collaboration et al. 2013, 2018).

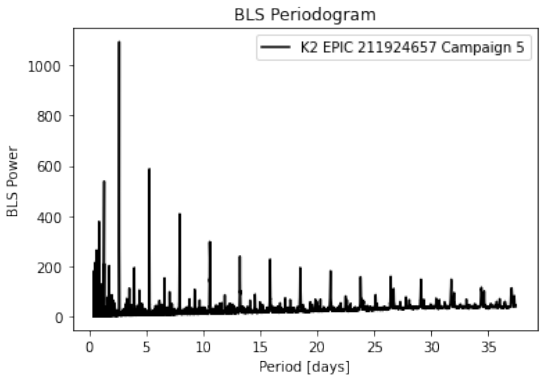

Figure 3. An example of a BLS periodogram for K2-146 campaign 5. The periodogram shows a clear maximum-likelihood period at around 2.6 days, which is consistent with previous knowledge of this system (Hamann et al. 2019). Thus, with this knowledge we can commence the fitting of a proper transit shape.

This is a fairly reliable method to use for period determination, and we very well might just proceed to transit fitting. However, the targets of this report are already fairly well known and well analysed, and many of them have already been analysed with much higher-calibre and more robust systems. One such system is Kruse et al. (2019); part of the results of this paper is a list of EPICs and associated exoplanetary parameters. Of these systems, exactly 107 out of our 181 candidates had associated data. Therefore the process to determine periods is modified, such that if a given system contains data in this set, instead of searching a range of periods from 0.7 to around 35 days we search a much smaller range of $\pm 10\%$ of the period found in Kruse et al. (2019). We search a range since the given parameters are gathered from K2 campaigns 0 through 8, and period variations over campaigns are very much possible, particularly since that is precisely what we are looking for in this report.

To get an idea of how accurate our method is for systems without given parameters, we can compute periods for the systems with given parameters, and compare to the results of Kruse et al. (2019). This comparison is shown in figure 4. The results suggest that our program correctly calculates the period roughly nine times out of ten, just as a crude estimation. Now, for these 107 systems in particular, we do not have to worry since we are already given the period, but this provides some insight into how accurate the 74 other systems' periods will be.

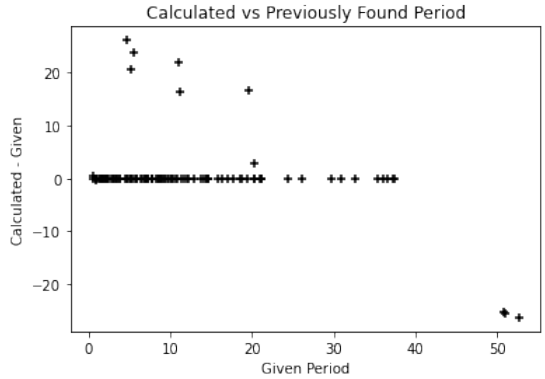

Figure 4. A comparison of calculated and previously determined values of period. The x axis contains the "True" value for period given by Kruse et al. (2019), and the y axis specifies the error in our calculations compared to that. Therefore, if a value is near zero on the vertical axis, then our predictions are accurate, and are at least close to being correct. Consequently, points far from the vertical zero are bad predictions. We find that 95 out of 107 of our calculated periods are within 5% of the given values.

2.3. *Fitting Light Curves*

Once we know the period and reference time of the first transit, we can fit the data to a transit model. The model we will use is based on the `exoplanet` module, with modifications to also fit transit timing variations (Foreman-Mackey et al. 2020). We will use a simple optimisation routine from `exoplanet` to find the maximum-likelihood model in each case.

The light curve resulting from a transiting exoplanet is a complex model, with many potential variables. The minimum set of parameters that can uniquely describe the transits are outlined in table 1. These are the pa-

rameters we will fit for. Notably, we will not be fitting for the masses of planets or stars, nor the precise nature of the orbit, namely semi-major axis, eccentricity and angle of periastron. Although these parameters are essential to fully describing a stellar system, they are not necessary for describing transit shape and transit timing variations, which is our goal.

Table 1. The set of parameters that will be used in the transit model along with their physical interpretations. The limb-darkening parameters are utilised as part of the `exoplanet` model to describe stellar brightness, attributed specifically to Kipping (2013).

Parameter	Physical Interpretation
p	Period of transit
t_0	Time of the first transit
$mean$	Average flux out of transit (≈ 1)
r/R_*	Radius proportional to host star
b	Impact parameter of transit
u_1&u_2	Stellar limb-darkening parameters
$TTVs$	List of TTVs, one for each transit

Theoretically, we could simply put all our parameters into this transit model and optimise it to the given light curve. The problem with this approach, however would be that in some systems we can expect to have up to 100 individual transits in a single campaign, which would need to be fitted simultaneously. Therefore, we use a sequential approach to evaluating the parameters.

1. First of all we fold the transits. This means that we locate the centre of each transit, then cut each transit out and stack them all on top of one another. We end up with a plot with a time-base exactly one period long, with the centres of the transits roughly aligned. See figure 5.

2. Then we fit all the transit parameters except the TTVs to that plot. The result is a model that specifies the shape of the transit, and is based solely on the folded data. See figure 6.

3. Then, using the transit model determined above, we take each individual transit and fit independently for the time of transit. This gives us the timing variation for each transit. See figures 7 and 8.

4. Then, we go back and repeat the first three steps several times, however when we fold the data in step 1, we also move each transit according to the TTV we have just found to make the folded plot a better fit.

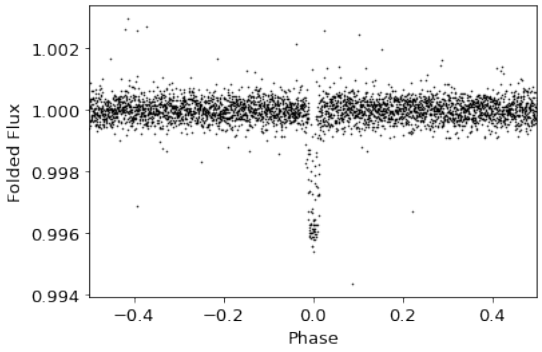

Figure 5. The folded flux plot for K2-146 campaign 5 as an example. The horizontal axis specifies the phase from the transit midpoint, that is, the time from the transit midpoint as a proportion of the period. Thus why the horizontal axis is exactly between ± 0.5. This plot shows the transits of the 2.6 day period planet in the K2-146 system.

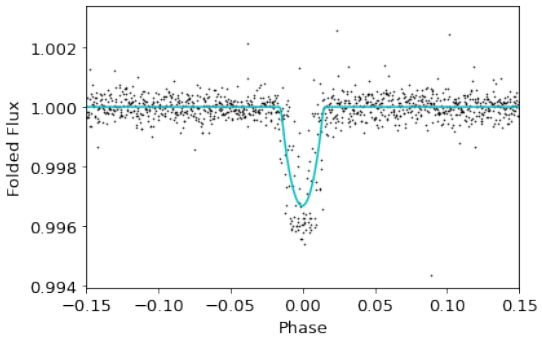

Figure 6. The folded flux for K2-146 campaign 5 and the first-attempt transit model in cyan. We see that the model is reasonably good at modelling the behaviour of the flux during the transit. After this fitting, we get preliminary values for r/R_*, b, u_1 and u_2 (Also for the $mean$ but this is a fairly redundant variable). The data itself however is not very consistent over each transit and that is because there are large TTVs, and so many transits here are being represented inaccurately, since there actual midpoints are slightly earlier or later than we have assumed.

The key step here is that the second time we fold the light curve data, we take into account the TTVs just calculated, and consequentially our new folded data will be a much better representation of the true shape of the transit. Therefore we will achieve a much more accurate transit shape. And therefore we will attain more accurate TTVs. Then the whole algorithm feeds through becoming more accurate every time and the model should converge after a few iterations. An example of the results after three iterations are in figures 9 and 10.

2.4. *MCMC*

The algorithm we have developed so far seems to be working very well. However, in our final results it would

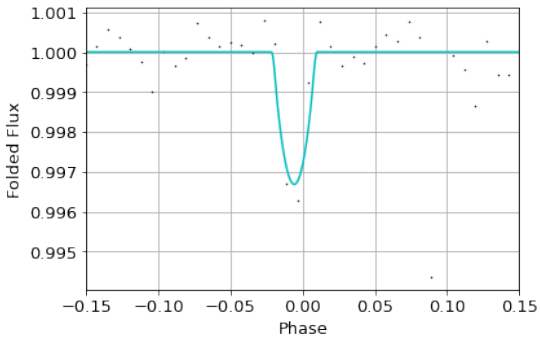

Figure 7. The data for the 19th transit of K2-146 campaign 5 and the first-attempt at fitting the transit timing variation. The result of this attempt is that this transit is earlier than we expect by about 22 minutes.

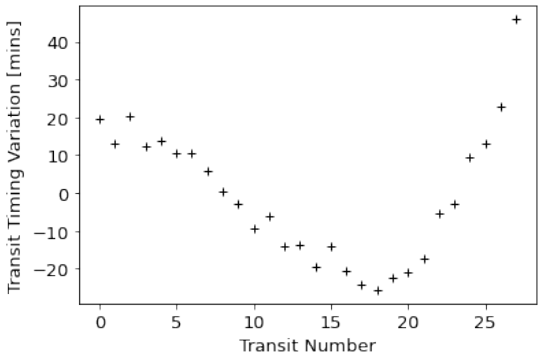

Figure 8. The TTVs determined using the first-pass transit shape. Each data point represents the transit timing variation found for each transit. There seems to be some coherent trend despite this being a fairly naive model.

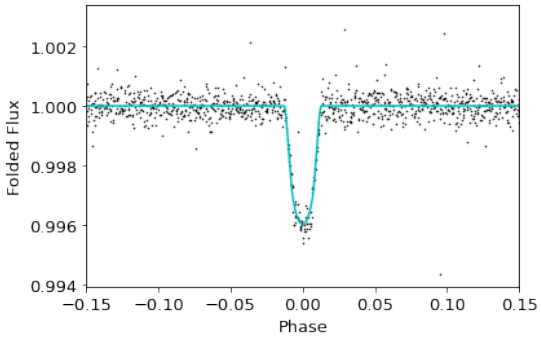

Figure 9. The folded flux for K2-146 campaign 5 adjusted for transit timing variations, and the third iteration transit model in cyan. We see that finding and accounting for TTVs means that the folded plot exhibits a much more definite transit shape. Therefore we can expect our resultant model to be far more accurate than in the first attempt

be best to be able to express some form of uncertainty in each of our results. This would give us a good idea of the general uncertainty of our model as well as visualising

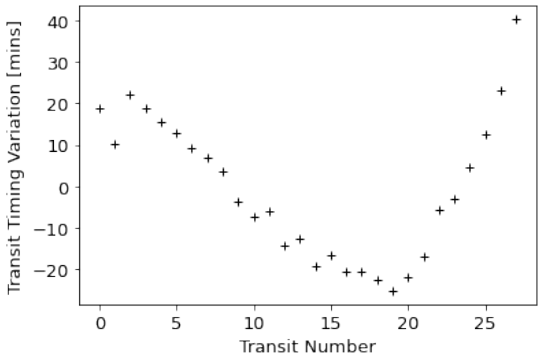

Figure 10. The TTVs for K2-146 campaign 5 after the third iteration. We see only very slight changes compared to the first iteration, which suggests that the model converged very quickly after the the first iteration.

which data points are more certain than others. We would also like to be more certain that the model is correctly converging, since in the previous section we use a fairly simple optimisation function from `exoplanet`.

The solution to these problems is to use Markov chain Monte Carlo (MCMC) methods. MCMC is a method of Bayesian statistical modelling comprising of walkers and which optimises based on probability, which is exactly what we want. Using MCMC allows us to extract meaningful uncertainties - which is practically impossible by analytical methods for such a dynamical system - and it gives us great confidence in our results. Specifically, we will be using the `PyMC3` package for python (Salvatier et al. 2016a)

Before we perform MCMC analysis, we must first specify the priors we will place on our parameters, since we understandably want to limit our parameters to values that are both physical and contextually feasible. See table 2.

Table 2. Priors (probability distributions) placed on the orbital model parameters. N means a Gaussian distribution and U means a uniform distribution. Note that these limits will also be used for the naive optimisation model. *Included within `exoplanet` and based on Kipping (2013)

Parameter	Priors
p	Constant
t_0	Constant
$mean$	$N(\mu = 1.0, \sigma = 0.1)$
r/R_*	$U(0.01, 0.1)$
b	$U(0, 1 + r/R_*)$
u_1	$U(0, 2)$*
u_2	$U(-1, 1)$*
TTV	$N(\mu = 0, \sigma = 36mins)$

Important things to note: first, as mention previously, p and t_0 are taken to be constant as calculated using BLS. This might seem ineffective seeing as BLS will not necessarily yield the highest level of accuracy, however we must consider that even if they are wrong, our algorithm will be able to account for it by adjusting the TTVs accordingly. Thus it simplifies our process to keep them constant. Secondly, we are only considering relative radii between 1% and 10% of the host star. There most certainly exist bodies outside those ranges such as smaller planets and brown dwarfs, but these are usually not worth consideration; Planets 1% as small as their star will only cause a maximum decrease in flux on the scale of 0.0001, which is usually well within our range of uncertainty and so these are almost impossible to detect. On the other hand, 'planets' with a relative size of 10% or greater are usually binaries or brown dwarfs, which are mostly outside the scope of this report.

Finally, the prior on TTVs might seem overly constraining, especially considering that Hamann et al. (2019) found the TTVs for K2-146 to be as large as hundred of minutes. However, that is over the timescale of all campaigns. These priors will be based upon singular campaigns, and will be entirely independent. In fact, if we look at figure10 we see that over campaign 5, the TTVs that we calculate within this single campaign are no larger than 40 minutes. This is because, even if the true TTVs over this time are very large, the BLS part of the algorithm will simply interpret the planet as transiting earlier or later, or that the period is larger or smaller than the true value. Therefore, considering that K2-146 is a system with disproportionately large TTVs, a standard deviation of 36 minutes on the TTVs priors should be sufficient to accommodate the vast majority of transit timing variations in our data set.

Therefore, this is the model we will use. To summarise the process thus far: we first prepare the data of a single campaign by flattening and trimming outliers. Then we use BLS to evaluate the period and t_0. Then we run several iterations of basic optimisation starting from the specified priors. Then finally we run MCMC and obtain our final parameters and associated uncertainties.

2.5. *Stitching TTVs*

Once we have the data for all campaigns of a given system, we then want to combine the transit timing variations of all campaigns so that we can identify behaviour that might suggest a perturber. Now the different campaigns of data may contain different values for the period, and will certainly contain different values for t_0. If we want to look at the long term behaviour then we have to unify this data to have the same reference parameters, else the results will not have any meaning. We do this by changing the parameters of all campaigns such that their TTVs are all in reference to the average period of the campaigns, and to the reference transit time of the first campaign. Such a result is shown in figure 11.

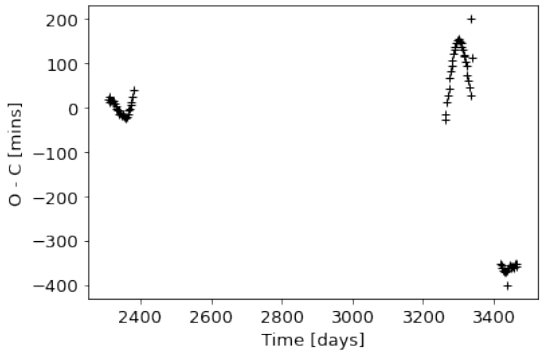

Figure 11. The TTVs for k2-146 over campaigns 5, 16, and 18. The results have been stitched and so are in reference to the same period and t_0. We can see a very distinct trend that strongly suggests an external perturber.

This will be the primary product for the set of systems chosen. The stitched TTV plot is not direct evidence of any sort of perturbations, but this is where the scope of this report ends. Ways that this could be further extended are explored in the discussion.

3. RESULTS

The entirety of the results of this algorithm as well as the code to produce it can be found at https://github.com/porrittmichael/Taste-of-Research. This is to facilitate the very large number of graphs and information associated.

The entire algorithm as described above was run for all 181 stars of interest over the course of 1-2 weeks. The process was not expected to take so long, but limited computational power is probably the main reason. Nevertheless all 181 systems were analysed and the results of every step of the algorithm were saved, including orbital parameters as well as graphs.

Specifically, for each system and for each campaign there is recorded the flattened and trimmed light curve, the best-fit folded data and the associated model, all the individual transits stacked in a single plot and the final TTVs for that campaign. Then there is also a plot of the stitched TTVs which encompasses all the campaigns. Additionally, the best fit model parameters and their uncertainties have been recorded for each campaign.

3.1. *Categorisation of Systems*

Each system had to be individually looked at and then be categorised and annotated. The annotations include the following information:

- Whether the model has a good fit - This is predominantly based off a visual inspection of the folded transits as well as the individual transits.
- Whether the system is a good candidate for having meaningful TTVs - Provided the transit model was a good fit, this was based off whether the TTVs might have any meaningful trend. This was quite lenient and likely contains a large number of systems without significant TTVs. Additionally, if a system had errors in the model but could potentially contain TTVs, it was marked as a TTV candidate.
- The most likely cause of the transit signal. The categories are those specified in table 3 and described through the rest of this section. Categorisation was based primarily off the fit and the orbital parameters.

Table 3. Categories and associated number of systems identified as belonging to each in the 181 systems included in this report.

Category	Count
Planet	96
Binaries	31
Starspot	5
Single Transit	17
Disappearing Planet	9
Nothing	14
No Data	9

The meanings of these categories are outlined below, along with some examples:

Planet—Planetary transits. These generally fit the transit model very well since most of our assumptions were based around our transits being planetary in nature. These are the main targets of this report, and the exact types of systems we aimed to investigate. See figure 12.

Of these candidates, 73 were determined to have a good fit and to have possible TTVs; 18 were determined to have some issues with fitting but could potentially be good TTV candidates; 5 were bad candidates.

Binaries—Eclipsing binary stars. These exhibited behaviour such as dual transits, and were typically very badly fitted since binary transits tend to result in very

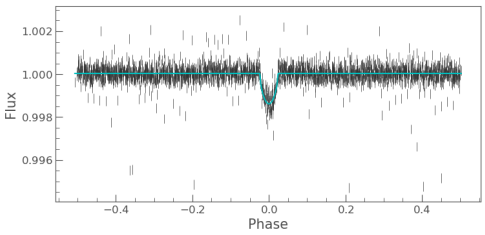

Figure 12. The folded plot from EPIC 211331236, campaign 18. An example of a well-fitted planetary transit with a period of 1.29 days.

significant drops in brightness, much more than we accounted for. Examples of light curves caused by binaries can be found in LaCourse et al. (2015). See figure 13.

Binary stars were not the intended target of this report, but despite this, and despite the best-fit model being very inaccurate, the resultant transit timing variations were actually very feasible. In most cases, the transit model was simply not deep enough, however, since the TTV fitting is based purely on probability and the transits are so definite, the transit time was able to be computed with relatively high precision.

Of all the identified binary systems, two were determined to have a reasonably good fit, and 29 were determined not to.

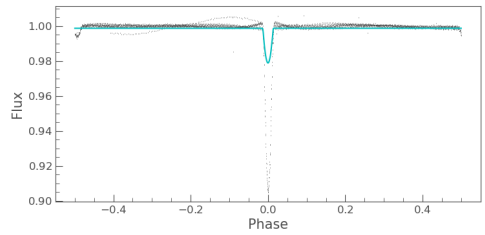

Figure 13. The folded plot from EPIC 212085740, campaign 5. We see that first of all, the transit fit is insufficiently deep to accurately model this transit, but also that there is a second transit dip roughly half a period away from the zero here; This corresponds to the secondary eclipse which occurs only in binary star systems.

Starspot—False positive due to a significant starspot. These typically manifested as roughly sinusoidal fluctuations in intensity (Giles et al. 2017). These were clearly very badly fit to a transit model. See figure 14.

Single Transit—One or very few transits of a planet in each campaign. Single transit events such as this are of great interest generally when multiple campaigns of data are available but in terms of TTVs, meaningful results are unattainable from so few transits. See figure 15.

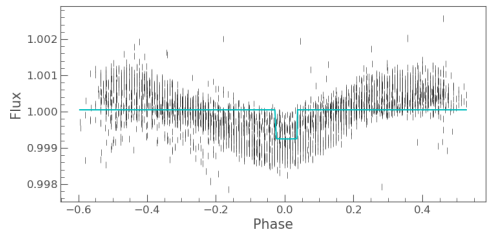

Figure 14. The folded plot from EPIC 211315836, campaign 5. An example of a significant starspot with a period of 1.38 days. The transit model is clearly useless in this case.

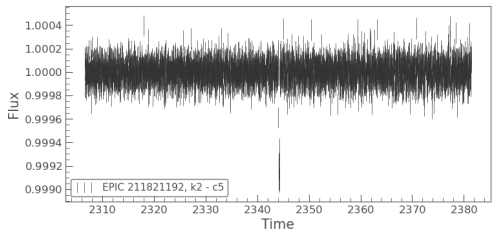

Figure 15. The full light curve from EPIC 211821192, campaign 5. We can identify a single transit occurring in this light curve at around 2344 days. Single transit events are largely discarded in this report since TTVs cannot be meaningfully extracted.

Disappearing Planet—Very strong evidence of planetary transits, but transits have disappeared in at least one campaign. These objects are not immediately useful to this report but having a planet suddenly ceasing to transit can potentially be very informative, and it suggests some significant external perturbations.

Nothing—The nothing tag is for light curves in which the best-fit transit model is questionable at best. These generally contained no convincing evidence for a planetary transit. Some of these are potentially false-negatives, but if there are any planets in this set of systems, the algorithm developed in this report was not able to find them.

No Data—This subset of systems were those that were cast out for one of two reasons: Either there was only one detected transit, which would indicated that any attempt at quantifying the period would be impossible, or else that the period evaluated by BLS was less then 0.7, which was decided in section 2.2 to be the lower limit for a reasonable period.

4. DISCUSSION

4.1. *Noteworthy Results*

Out of the 73 likely planets with good fits and possible TTVs, a subset of 15 exhibited TTVs with a very convincing periodic trend. It is important to note that the remainder of the 73 could also have meaningful TTVs: the list of 15 is just the list of most likely candidates based off of visual inspection. Notably, most of these have data from all three campaigns (5, 16, 18), a couple have data from only two but are included because they have both a strong trend and have many transits per campaign (short period).

See table 4.

Table 4. Final list of EPIC numbers of the systems that are very likely to have TTVs that are very promising for further investigation.

EPIC Number
211359660
211432922
211490999
211718187
211733267
211763214
211800191
211816003
211897691
211924657
211941472
212072539
212088059
212150006
212154564

Furthermore, there is also a number of binary stars that have very strong evidence for having fluctuating TTVs. We would be inclined to dismiss these for any number of reasons: The transit fits are very inaccurate for most, the model we used is designed to model the transits of planets, not stars, and also because this was not the objective of the report. However, these are being included as a recommendation for future investigation because they have very clear trends in their TTVs *and* they have a large number of transits per campaign (due to small periods) which gives greater confidence in the resultant trends. See table 5.

In many ways, transit timing variations in the transits of binary systems can be more interesting than those of exoplanets. Binary stars tend to have relatively short periods and are accordingly fairly close. Therefore, for a perturber to cause TTVs, the perturbing force has to be of a comparative magnitude to the force from the nearby star. A large cause of these eclipse timing variations is the presence of a third star in the system. In fact, it is estimated that at least 40% of all short-period

binary systems have a third star (LaCourse et al. 2015; Tokovinin et al. 2006).

Table 5. List of binary stars that exhibit significant TTVs, and are very promising for further investigation.

EPIC Number
211613886
211812160
211929937
211946007
212012387
212024647
212085740
212096658
212163353

4.2. *Potential Improvements*

There are a number of examples where the algorithm outlined in the course of this report fell short and could be most improved. There were many different ways that the light curves of different systems were incorrectly interpreted or processed, but below are the main recurring issues that were encountered:

Most of the issues with this program are related to eclipsing binaries. Now, these were at no point the primary objectives of this report, however a few alterations to the existing program would correct many of the main problems the models for binaries have. Correcting these issues would most like also make the program on the whole more reliable.

First of all is to do with the removal of outliers that is outlined in section 2.1. For most planetary transits, trimming removed data points that might otherwise cause errors in the subsequent model fitting. For several binaries, however, it cause many completely reasonable data points towards the middle of each transit to be removed with no gain. Potential solutions to this problem are to improve the way we assess the value of a data point by taking into account the quality flags associated with Kepler data. This was excluded mainly due to ignorance of its existence. Another way we might improve this is determine the rough locations of transits prior to trimming and then only remove outliers if they are away from any transits.

The other main issue with binaries was that the transit model did not allow for transiting objects larger than 10% the radius of the host star, and so most transit models for binaries ended up being nowhere near deep enough. The reason for this low upper limit was initially because we were excluding binaries, but it would most likely not change the results of any other type of system to increase this limit.

4.3. *Future Extension*

The aim of this report was achieved with considerable success, with our list of strong transit timing variable candidates. In addition, a large number of discoveries were made about systems that were not of immediate interest to this report. There are many ways the results of this report could be taken and improved upon and extended.

This report is intended to be continued in the same way as the previously mentioned investigation into K2-146 (Hamann et al. 2019). The primary product of this report is the transit timing variations for a list of promising systems. The transit timing variations alone are merely an indicator of external forces, but if they are used as part of a dynamical N-body simulation, then they can very tightly constrain the mass and orbital parameters of a stellar system. This is the method used in the Hamann paper and is the direction in which the results of this report are intended to be taken.

In conjunction with the above, it would be complimentary and helpful to perform stellar spectroscopy on the systems in the planet candidate list. This would enable the extraction of radial velocities and in turn would allow us to verify or disprove the existence of potentially non-transiting objects once an N-body analysis has been performed.

In the course of this report, it has also been shown that transit timing variations are useful tools of analysis for not only planets but also for binary stars. We somewhat unintentionally ended up with a list of very good candidates for TTVs in binary stars. This has very interesting prospects for investigation into these possible trinary systems that would have to be relatively compact to have such noticeable effects. The process for this would be very similar to that of planetary TTVs. That is, perform a dynamical N-body analysis to try to infer an object that would cause the observed timing variations. Subsequent radial velocity measurements might also be of use.

Other applications of our results are in investigating the identified *single transit* systems as well as the *disappearing planet* systems. Both of these are only really identifiable over multiple campaigns. For single transit systems we might run a modified BLS algorithm over the entire data set of multiple campaigns to discover long-period planets. For the disappearing planet systems we might develop some sort of transit depth analysis and/or N-body simulations to identify possible causes for such behaviour.

It is also worth noting that continued refinement of the methods and results of this report would likely yield a larger set of candidates worthy of further investigation.

4.4. *Conclusion*

In summary, the aims outlined at the beginning of this report were very much achieved. We developed a largely successful algorithm to model the transits of a stellar system and extract the transit timing variations. We identified 15 planetary systems with promising TTVs and we also found a list of 9 binary systems that exhibit potentially meaningful TTVs.

4.5. *Acknowledgements*

This report made use of `exoplanet` (Foreman-Mackey et al. 2020) and its dependencies (Agol et al. 2020; Kumar et al. 2019; Astropy Collaboration et al. 2013, 2018; Kipping 2013; Luger et al. 2019; Lightkurve Collaboration et al. 2018; Salvatier et al. 2016b; Theano Development Team 2016).

REFERENCES

2018, Kepler/K2 Guest Observer Program, NASA. https://keplerscience.arc.nasa.gov/k2-data-release-notes.html

Agol, E., & Fabrycky, D. C. 2018, Transit-Timing and Duration Variations for the Discovery and Characterization of Exoplanets, ed. H. J. Deeg & J. A. Belmonte, 7

Agol, E., Luger, R., & Foreman-Mackey, D. 2020, AJ, 159, 123

Agol, E., Steffen, J., Sari, R., & Clarkson, W. 2005, MNRAS, 359, 567

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33

Astropy Collaboration, Price-Whelan, A. M., SipHocz, B. M., et al. 2018, aj, 156, 123

Carter, J. A., Agol, E., Chaplin, W. J., et al. 2012, Science, 337, 556. https://science.sciencemag.org/content/337/6094/556

Foreman-Mackey, D., Luger, R., Czekala, I., et al. 2020, exoplanet-dev/exoplanet v0.4.1, , , doi:10.5281/zenodo.1998447. https://doi.org/10.5281/zenodo.1998447

Giles, H. A. C., Collier Cameron, A., & Haywood, R. D. 2017, MNRAS, 472, 1618

Hamann, A., Montet, B. T., Fabrycky, D. C., Agol, E., & Kruse, E. 2019, AJ, 158, 133

Holman, M. J., & Murray, N. W. 2005, Science, 307, 1288

Howell, S. B., Sobeck, C., Haas, M., et al. 2014, PASP, 126, 398

Kipping, D. M. 2013, MNRAS, 435, 2152

Kruse, E., Agol, E., Luger, R., & Foreman-Mackey, D. 2019, ApJS, 244, 11

Kumar, R., Carroll, C., Hartikainen, A., & Martin, O. A. 2019, The Journal of Open Source Software, doi:10.21105/joss.01143. http://joss.theoj.org/papers/10.21105/joss.01143

LaCourse, D. M., Jek, K. J., Jacobs, T. L., et al. 2015, MNRAS, 452, 3561

Lightkurve Collaboration, Cardoso, J. V. d. M., Hedges, C., et al. 2018, Lightkurve: Kepler and TESS time series analysis in Python, Astrophysics Source Code Library, , , ascl:1812.013

Lithwick, Y., Xie, J., & Wu, Y. 2012, ApJ, 761, 122

Luger, R., Agol, E., Foreman-Mackey, D., et al. 2019, AJ, 157, 64

Luger, R., Agol, E., Kruse, E., et al. 2016, AJ, 152, 100

Maehara, H., Shibayama, T., Notsu, S., et al. 2012, Nature, 485, 478

Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. 2016a, PeerJ Computer Science, 2, e55

—. 2016b, PeerJ Computer Science, 2, e55

Theano Development Team. 2016, arXiv e-prints, abs/1605.02688. http://arxiv.org/abs/1605.02688

Tokovinin, A., Thomas, S., Sterzik, M., & Udry, S. 2006, A&A, 450, 681

APPENDIX

All of the results of this project in addition to the code used to produce it can be found at https://github.com/porrittmichael/Taste-of-Research.

A. LIST OF TARGETS BY EPIC ID

Table 6. List of K2 targets by their EPIC number. The list is a combination of two K2 observer target lists: GO16011 and GO18048 (See K2 approved targets). There are 181 systems in total.

212088059	211526186	211821192	211503363	211743874	211442297	212119244	211498244
211422931	212152114	211946007	211428897	211541590	211510580	212083250	212009427
212012030	211519965	212006344	211924657	212159623	211516501	211718187	211945201
211897691	211439059	211896651	211925181	211925595	211733267	212157262	211817229
211413752	211713099	212021516	211972744	212082682	211731298	211319617	212154564
211351097	212088895	211910237	211959012	212020442	211763214	211770696	211818569
211744153	211391664	212008766	211969807	212037403	211417727	211843564	212085740
212132195	211736671	211418729	211315836	212110007	211837343	212024647	211355342
228682450	211430148	211897272	211682544	211606790	211469889	211611158	211525389
211509553	211645912	211736305	211886472	211407755	211642307	212150006	211418290
211509975	211822797	211913395	212069861	211705654	211431013	211442571	211539054
212088968	212120773	211351816	211770578	211579112	211399359	211885185	211409299
211791178	211913977	212099230	212069706	211383821	211535327	211822953	211800191
211432946	211492449	211613886	211825799	212141021	212096658	212110888	212006318
211972837	212011230	211772227	211916756	212012119	212072539	211564796	211894612
211732801	211357309	211796070	211619879	211491383	211419451	211431505	211788221
211770390	211987878	212009150	212161956	211919004	211401787	211812160	211762841
211783206	211834065	211680698	211906259	211816003	211551160	212163353	212109135
211786983	212157058	211331236	211563123	211359660	211797637	211505089	211529065
211411112	211432103	211978988	212130773	212019055	211579683	212136123	211432922
211562654	211923431	211327855	211814733	211586387	212164470	211991987	211995398
211703878	211997641	212009702	211633458	211594205	211754962	211929937	211941472
211490999	212012387	211770795	212066407	211779390			