Quasi-periodic Gaussian processes for stellar activity: From physical to kernel parameters

B. A. Nicholson [1,2]* and S. Aigrain[1]

[1]*Sub-Department of Astrophysics, University of Oxford, Keble Rd, Oxford OX13RH, UK*
[2]*University of Southern Queensland, Centre for Astrophysics, Toowoomba 4350, Australia*

ABSTRACT

In recent years, Gaussian Process (GP) regression has become widely used to analyse stellar and exoplanet time-series data sets. For spotted stars, the most popular GP covariance function is the quasi-periodic (QP) kernel, whose hyperparameters of the GP have a plausible interpretation in terms of physical properties of the star and spots. In this paper, we test the reliability of this interpretation by modelling data simulated using a spot model using a QP GP, and the recently proposed quasi-periodic plus cosine (QPC) GP, comparing the posterior distributions of the GP hyperparameters to the input parameters of the spot model. We find excellent agreement between the input stellar rotation period and the QP and QPC GP period, and very good agreement between the spot decay time-scale and the length scale of the squared exponential term. We also compare the hyperparameters derived from light and radial velocity (RV) curves for a given star, finding that the period and evolution time-scales are in good agreement. However, the harmonic complexity of the GP, while displaying no clear correlation with the spot properties in our simulations, is systematically higher for the RV than for the light-curve data. Finally, for the QP kernel, we investigate the impact of noise and time-sampling on the hyperparameters in the case of RVs. Our results indicate that good coverage of rotation period and spot evolution time-scales is more important than the total number of points, and noise characteristics govern the harmonic complexity.

Key words: methods: data analysis – techniques: photometric – techniques: radial velocities – stars: activity.

1 INTRODUCTION

Gaussian Process (GP) regression has become an increasingly popular method for analysing stellar activity signals in photometric and radial velocity (RV) time-series data. In particular, GP models based on quasi-periodic kernel functions are able to reproduce the rotationally modulated signatures of evolving magnetically active regions remarkably well (see e.g. Aigrain, Pont & Zucker 2012; Haywood et al. 2014; Angus et al. 2018). Quasi-periodic GP models have been applied extensively to stellar time-series data, both to mitigate the impact of stellar activity on planet detection or characterization (see e.g. Haywood et al. 2014; Barclay et al. 2015; Grunblatt, Howard & Haywood 2015; Rajpaul et al. 2015, and numerous later references), and to study stellar activity signals per se (see e.g. Angus et al. 2018; Barragán et al. 2021).

There are many ways to construct a quasi-periodic kernel function for a GP model, but one such kernel has become especially popular for modelling stellar activity signals, to the extent that, in the exoplanet community, it is often referred to simply as the Quasi-Periodic (QP) kernel. The QP kernel is constructed by multiplying a periodic term, consisting of the exponential of a sine-squared function, with a decaying envelope, consisting of a squared exponential function. With only four parameters, this kernel function can produce signals with different amplitudes, periods, evolution time, and degree of harmonic complexity. Random samples drawn from a QP GP prior bear a striking resemblance to observed stellar light curves (see e.g. Rajpaul et al. 2015).

A QP kernel is not always the best choice of kernel to model stellar activity signals. For example, Gilbertson et al. (2020) show that simpler, a-periodic kernels can in some circumstances perform somewhat better in terms of activity mitigation in RV time-series. However, the QP kernel explicitly encodes the commonly held belief that the underlying signal should be quasi-periodic, even if its periodic nature is not necessarily apparent in noisy and/or sparsely sampled data. Furthermore, most of the parameters of the QP kernel (known as the hyperparameters, or HPs, of the GP) admit a fairly natural interpretation in terms of the physical properties of the active regions, such as their rotation rates and lifetimes. The primary goal of this study is to investigate, using simulated data, the extent to which this 'natural' interpretation is robust and can be used to guide modelling decisions.

Observed light and RV curves of active stars often contain significant power not only at the stellar rotation period, but at its first few harmonics (see e.g. Aigrain et al. 2012), a feature which is not naturally reproduced by the standard QP kernel. Consequently, in recent years a number of other quasi-periodic kernels have been investigated, which explicitly generate power at half the rotation period, including the 'QP plus cosine' (hereafter QPC) kernel introduced by Perger et al. (2021, hereafter P21), given by equation (2), and the 'rotation term' implemented in the CELERITE package (Foreman-Mackey et al. 2017). After experimenting with both of these, we decided to include the former in this study. Achieving convergence with, and robustly interpreting the parameters of, GP

* E-mail: belinda.annette.nicholson@gmail.com

models based on the CELERITE rotation term is somewhat more challenging with our model data, and we defer that to a separate paper.

While GP models using the QP and QPC kernels provide a useful description of common observables, they are still very much phenomenological. A more physically motivated family of GP models for active star light curves was recently introduced by Luger, Foreman-Mackey & Hedges (2021). Their STARRY-PROCESS framework places a GP prior on the distribution of active regions on the stellar surface, which is described using a spherical harmonic decomposition. However, this relatively new approach has not yet been widely used, nor has it been extended to RV data, placing it outside the scope of this work.

The fact that a QP GP can be used to recover stellar rotation periods from light curves has already been demonstrated (see e.g. Angus et al. 2018). However, the behaviour of the other HPs, in particular the parameter controlling the evolution of the signal, has not yet been investigated in detail. A number of recent studies have shown that it is possible to measure active region lifetimes from photometry, for example by fitting a function to the decaying envelope autocorrelation function (ACF) of the light curve (Giles, Collier Cameron & Haywood 2017; Basri et al. 2021; Santos et al. 2021). These studies have shown that, provided the light-curve duration significantly exceeds the active region lifetime, the ACF decay time is roughly proportional to the active region lifetime, but the constant of proportionality is not unity. We would expect the results obtained with a GP model, which effectively is a more direct way of modelling the covariance, to be consistent with those obtained by fitting the ACF, but this has yet to be demonstrated.

P21 analysed simulated RVs with four different spot distributions with GP models using QP, QPC, and CELERITE rotation kernels. They were able to recover the input spot evolution time, but their study considered only a single value of the rotation period and evolution time. In this work, we extend this to a much broader parameter space, varying the rotation period, spot lifetime, and number of spots.

A common application of the QP GP is in filtering stellar activity from RV time-series. When following up planet candidates detected in transit, it is common practice to 'train' the GP on the light-curve data first (see e.g. Haywood et al. 2014; Rajpaul et al. 2015; Fulton et al. 2018). The posterior distribution over the HPs of the QP GP is then used as a prior for the RV analysis. However, this may not always be advisable, for two distinct reasons. First, the photometric data, if available, is rarely contemporaneous with the RVs, and can be taken months or even years apart. Using a 70-yr compilation of solar irradiance data, Kosiarek & Crossfield (2020) showed that all the HPs of a QP GP fit to the light curve of a Sun-like star can vary significantly over the solar cycle. This implies that it may be dangerous to train a QP GP on light-curve data before applying it to RV data, if the two were taken too far apart. Furthermore, their analysis of simultaneous SORCE photometry and HARPS-N RVs found consistent periods and evolution time-scales, but significant differences in the length scale parameter that controls the complexity of the signal within a period (often referred to as the harmonic complexity). This effect was also seen in simulated data by Barragán et al. (2022). It is related to the fact that the RV variations depend not only on the projected area of the active regions and their contrast, but also on their local line-of-sight velocity. This underpins the so-called FFprime framework for predicting RV variations from photometry (Aigrain et al. 2012), as well as latent variable models where RVs are modelled jointly with activity indicators extracted from the same spectra, as linear combinations of an underlying GP and its time derivative(s) (see e.g. Rajpaul et al. 2015; Jones et al. 2017; Delisle et al. 2022). However, as it is still quite common in the literature

to model RVs alone using a QP GP trained on photometry, it is important to test the robustness of this approach.

In this work, we first use a simple spot model to simulate simulate tightly sampled, noise-less photometric and RV time-series, and then model each independently with a QP or QPC GP. This allows us to test the extent to which the HPs of these models relate to the physical parameters of the simulations, and how closely the results based on photometry match each other. Doing this using highly idealized data enables us to explore the differences between these two types of observations at a fundamental level, removing the effects of noise and time-sampling. As a second step, we then degrade the time-sampling of the RV data, and add white noise, to make the simulations more realistic (since RV observations are typically ground-based). This allows us to investigate how time-sampling in particular affects the recovery of the GP HPs, in the case when space-based photometry is not available (i.e. for 'pure RV' surveys, rather than for transit follow-up).

The remainder of this paper is organized as follows. In Section 2, we define the QP and QPC kernels used in this work, and fit simulated time-series generated from the GP prior itself. This allows us to test for degeneracies inherent to the GP models. In Section 3, we define the spot model used to simulate stellar light and RV curves, and examine the properties of the HPs obtained from the simulated light curves. The GP fits to the idealized RV data sets are discussed in Section 4, and the effect of degrading the time-sampling of the RVs are explored in Section 5. Finally, we summarize and discuss the findings of this work in Section 6.

2 QUASI-PERIODIC GAUSSIAN PROCESS MODELS

GP regression is a powerful and flexible method for analysing data containing correlated noise. Rather than fitting a function to data, GP regression instead fits the correlation between data points. At its heart, a GP is defined by its kernel covariance function, k, which describes the extent to which two observations are correlated with each other, usually as a function of their distance in some input space (for time-series, the time interval separating them). For the purpose of modelling stellar activity signals in astronomical time-series, a quasi-periodic covariance function is commonly chosen. In the most general sense, a quasi-periodic kernel is usually obtained by multiplying a periodic term with a decaying envelope.

2.1 The quasi-periodic kernel

In the standard QP kernel most commonly used for stellar activity modelling, the periodic term is given by the exponential of a sine-squared function, while the evolution is described by a squared exponential term. Formally, the QP kernel is defined as:

$$k_{QP}(t, t') = A \exp\left[-\Gamma \sin^2\left(\frac{\pi(t - t')}{P}\right) - \frac{(t - t')^2}{2l^2}\right], \qquad (1)$$

where A is the variance, Γ is the scale factor (termed 'harmonic complexity'), and P the period of the sine-squared term, and l is the 'evolution time-scale' of the squared exponential term. These terms are referred to as the 'hyperparameters' of the GP so as to differentiate this process from fitting a parametric function to the data. For further information of the quasi-periodic kernel GP and on GP's in general, we refer the reader to Rasmussen & Williams (2006) and Roberts et al. (2013).

The QP GP used in this work is implemented using the GEORGE python package (Ambikasaran et al. 2015) by combining the EX-

PSINE2KERNEL and EXPSQUAREDKERNEL built-in kernel functions. The GEORGE kernels use the hyperparameters values in natural log as input, with the exception of the Γ term, and reparameterize the denominator of the squared exponential term as $m = 2l^2$, called the 'metric'. We also solve for the mean function of the data, in this case a constant, as this reflects the practice of fitting underlying functions (such as a Keplerian orbit in the case of RVs) to the data.

2.2 The quasi-periodic plus cosine kernel

P21 introduced the idea of adding a cosine term to the periodic component of the QP kernel, with a period equal to half that of the sine-squared term, to better handle cases where the distribution of spots on the stellar surface gives rise to signal the first harmonic of the stellar rotation period (evidenced by a secondary peak in the autocorrelation function at half the rotation period). Mathematically, this 'Quasi-Periodic plus Cosine' (QPC) kernel is defined as:

$$k_{QPC}(t, t') = A \exp\left[-\frac{(t-t')^2}{2l^2}\right] \times$$
$$\left(\exp\left[-\Gamma \sin^2\left(\frac{\pi(t-t')}{P}\right)\right] + f \cos\left(\frac{4\pi(t-t')}{P}\right)\right), \quad (2)$$

where A, Γ, P, and l retain the same definitions as in the QP kernel, and we have introduced an additional hyperparameter f, which controls the amplitude of the cosine term relative to the sine-squared term. As f tends to zero, the behaviour of k_{QPC} tends to that of k_{QP}.

In this work, we define this kernel again in GEORGE by with the addition of a COSINEKERNEL term to our GEORGE QP kernel.

2.3 Exploring inherent degeneracies in the hyperparameters

The first step in assessing the ability of a GP to recover physical parameters of a star is to investigate the stability of the GP fit itself and explore any inherent degeneracies in fitting a QP or QPC kernel. To test this, we generate a number of 'light curves' by drawing samples from our GP prior, and fit these sample curves to see if we recover the input GP hyperparameters. This step also serves to validate our fitting and posterior exploration procedure. The bulk of this analysis was carried out using the QP kernel, but we also discuss the QPC kernel and its extra hyperparameter towards the end of the section.

Our sample curves were generated over a range of hyperparameters, with 20 sample curves per set of hyperparameters. Each sample curve spans 100 d, with a sampling of three, evenly spaced points per day, and a fixed amplitude of 1. The grid of hyperparameters contain period values of 2, 10, 30, and 60 d, Γ values of 0.5 and 2 and l values of a 1, 3, and 10 times the period. These values were chosen as they represent the range of values typically found in the analysis of stellar light curves across a range of activity levels.

Markov Chain Monte Carlo (MCMC) was used to sample the posterior distribution of the hyperparameters for each sample curve, using the EMCEE python package (Foreman-Mackey et al. 2013). The MCMC was set up with 80 walkers in a Gaussian distribution about the sample curve input values, with a broad uniform prior used for each hyperparameter across all sample curves (see Table 1). There was an initial burn in of 100 steps, from which a tight Gaussian distribution around the highest probability state was used to re-initialize the walkers. The main MCMC was then run until the walkers had converged on a solution or a maximum of 8000 iterations

Table 1. Priors placed on the parameters and QP GP hyperparameters for the MCMC fit of the sample curves drawn from sets of QP GPs.

Input Parameter	Prior[a]
$\ln(A)$	$\mathcal{U}[-30, 30]$
$\ln(P)$	$\mathcal{U}[-15, 15]$
Γ	$\mathcal{U}[0, 15]$
$\ln(2l^2)$	$\mathcal{U}[-30, 30]$
Mean	$\mathcal{U}[-15, 15]$

Note.[a] $\mathcal{U}[a, b]$ refers to uniform priors between a and b.

was reached. The convergence of the MCMC walkers was checked every 100 steps, and a sample was deemed to have converged if the autocorrelation length of all chains was more than 60 times the number of elapsed iterations and had not changed by more than 1 per cent in the last 100 steps. For each of the MCMC solutions a 'burn-in' of twice the maximum chain autocorrelation length was removed, and the chains were thinned by half the minimum chain autocorrelation length to ensure an unbiased sample. The optimum solution for each hyperparameter was taken as the median of the posterior distribution, and the uncertainty as difference between the 50th and 16th percentile for the lower bound and the 84th and 50th percentile for the upper.

Fig. 1 shows how closely the input values of the period (left-hand panel), l (middle), and Γ (right-hand panel) were recovered by the GP, relative to the uncertainties calculated from the MCMC posterior distributions. Performing a Kolmogorov–Smirnov test on the distribution of (input–output)/uncertainty for each of period, l and Γ, we find D statistic values of 0.081, 0.068, and 0.058, and p values of 0.016, 0.071, and 0.18, respectively. The D statistic values are close to 0 (normal distribution), and the p values are not significant enough to reject the null hypothesis that these distributions are drawn from a standard normal distribution. In other words, all three distributions are consistent with a Gaussian with a mean around zero and a standard deviation of 1. This means that, within the limits of statistical uncertainty, the QP GP is non-degenerate and stable: the parameters that go in are the parameters that will come out.

We then tested in the same manner a set of light curves generated from a QPC kernel. The input values for P, Γ, and l were the same as in the QP case. We tested two values of f, 0.1 and 0.9, and adopted a uniform prior on $\log(f)$ between $\log(10^{-4})$ and 0 when fitting the samples. We find the same results for period, l and Γ as with the pure QP case: the input hyperparameters are all well-recovered. The fraction term, f, was only recovered in seven cases within 3-sigma, all of which had high f. Fig. 2 shows histograms of the recovered f values for the low (solid) and high (hatched) input values. For the low values of f, the QPC GP recover recovers the f values, but not as well as it does for the other parameters. The high-f value cases are not well recovered, despite these cases all having well recovered period, l and Γ values. As seen in 2, the recovered f for these high-f cases range from 0.2 to 0.9, with a median value of 0.7. This indicates that the QPC GP is only weakly sensitive to the true fraction term on the cosine component.

3 MODELLING SPOTTED STAR LIGHT CURVES WITH QUASI-PERIODIC GPS

In this section, we describe how we simulate light curves for rotating stars with evolving active regions, and then fit them with quasi-periodic GPs using the kernels described in Section 2.
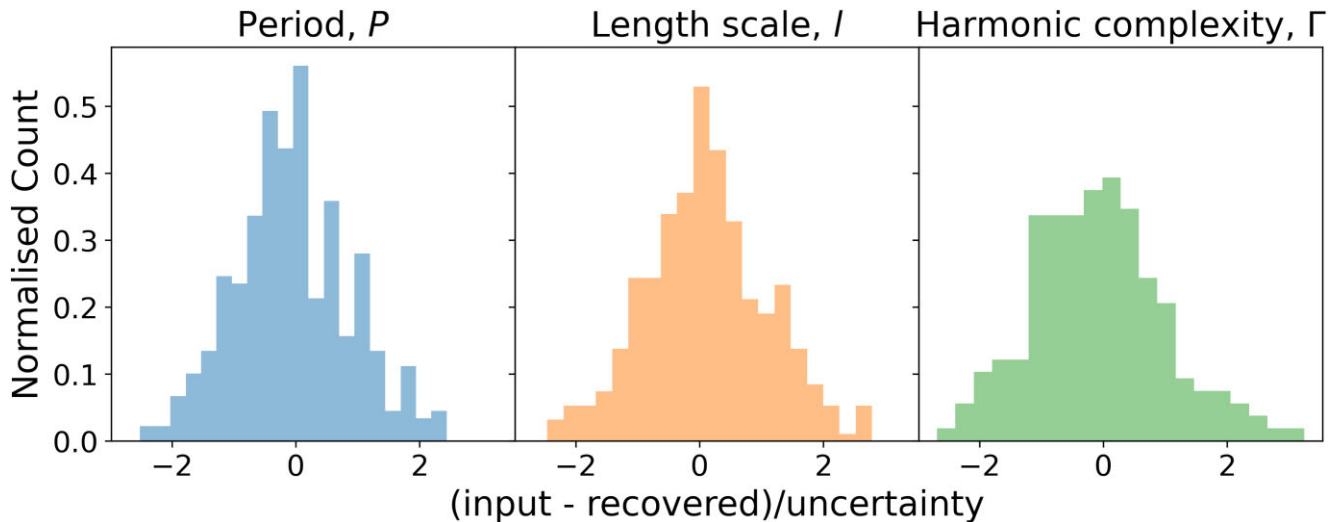
Figure 1. Recovery of input QP GP hyperparameters: each panel shows a histogram of the difference between the input and output parameter values divided by the uncertainty of the output solution for each of the varied hyperparameters: Period (P, left-hand panel), length scale (l, middle), and harmonic complexity (Γ, right-hand panel).
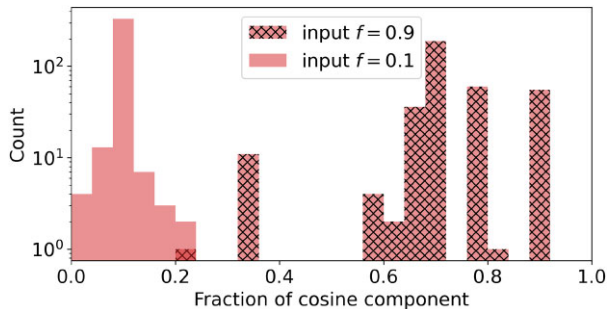

Figure 2. Histograms of the recovered cosine term fraction f from a QPC GP for sample curves drawn from a QPC GP with a high f value of 0.9 (hatched), and low f value of 0.1 (solid).

3.1 Simulating spotted star light curves

To test the correlation of the QP GP hyperparameters with physical characteristics of a star, we generate 100 simulated light curves using a simple spot model implemented in the `PySpot` package[1] (Aigrain 2021), which is a simplified version of the model used in the *Kepler* rotation and differential rotation blind exercise presented in Aigrain et al. (2015). We first define the parameters of the star and the spots in each simulation, then use these simulations to produce light and RV curves.

All simulations last $T = 250$ d, with two equally spaced observations per day, and all the stars are seen equator on, and have the same radius as the Sun. For each simulation, we draw the stellar rotation period, P_{rot}, the average number of spots present on the stellar surface at any given time, N_{Spot}, and the ratio of the characteristic lifetime of the spots to the rotation period, τ/P_{rot}, from random distributions which are listed in Table 2. Each spot grows exponentially, peaks in size at a specific time and size, and then decays again exponentially. The times at which the spot sizes peak were drawn at random times from $-T$ to $2T$. This is to avoid a situation where, particularly for long lifetimes, the spot coverage is systematically lower at the

Table 2. Stellar model input parameter distributions.

Parameter	Parameter range (uniform distribution)
Average spot number, N_{Spot}	$0 < \log_{10}(N_{Spot}) < 2$
Stellar rotation period, P_{rot}	$0 < \log_{10}(P_{rot}) < 2$
Spot decay timescale, τ	$1\,P_{rot} < \tau < 10\,P_{rot}$
Spot longitude, ϕ	$0 < \phi < 2\pi$
Spot latitude, θ	$0 < \sin(\theta) < 1$
Spot maximum area, A_{max}	$-3 \leq N_{Spot}\log_{10}(A_{max}) < -2$

beginning and end of the simulation than in the middle. The spots are assumed to be perfectly dark (contrast of 1 with respect to the stellar photosphere) and their peak sizes A_{max} range from 0.1 to 1 per cent of the visible hemisphere. The spots are randomly distributed on the stellar surface (uniform distribution in longitude, and in $\sin(\theta)$ where θ is spot latitude). This is not particularly realistic, but limiting the spots to specific latitude bands (as done for example in Aigrain et al. 2015) would not significantly alter the results of this study, because the stars are seen equator-on. After reaching its peak size, each spot decays according to an exponential with half-life τ. Following Aigrain et al. (2015), the spots grow five times more rapidly than they decay (though we also performed some simulations with equal growth and decay time, to see if this has an effect on the recovery of the spot lifetime, as discussed in Section 3.3.2). We then simulate light curves using the analytic spot model of Dorren (1987), using a linear limb-darkening law with $u = 0.5$. Only idealized, noise-free data were considered in this study, so the simulated fluxes have no associated uncertainties.

3.2 Fitting the light curves using a GP

When analysing a light curve simulated using a spot model, the GP is no longer the 'correct' model. Although the QP and QPC GP kernels described in Section 2 are very flexible, samples from GPs cannot reproduce the simulated light curves exactly (to machine precision). To account for this, we include a jitter term when modelling the light curves with the QP and QPC GPs. This jitter term absorbs any deviations between the simulated observations and the GP model. In

[1] See https://github.com/saigrain/pyspot.

Table 3. Priors placed on the parameters and QP GP hyperparameters for the MCMC fit of the model time-series data.

Input (hyper)parameter	Prior[a]
$\ln(A)$	$\mathcal{U}[-30, 30]$
$\ln(P)$	$\mathcal{U}[\log(\max[2T_{min}, 1]), T_{max}]$[b]
Γ	$\mathcal{U}[0, 15]$
$\ln(2l^2)$	$\mathcal{U}[-30, 30]$
Mean	$\mathcal{U}[-15, 15]$
$\ln(J)$	$\mathcal{U}[-20, 20]$

Note.[a] $\mathcal{U}[a, b]$ refers to uniform priors between a and b. [b] T_{max} and T_{min} refer to the maximum and minimum length of time between data points within a given time-series.
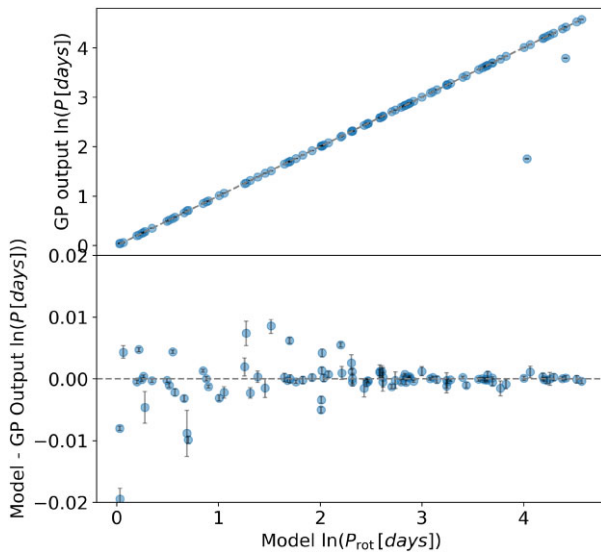
Figure 3. The top panel compares the input stellar rotation period to the output QP GP period (both in natural log) for the model light-curve data, with the unity relation indicated by the grey dashed line. The bottom panel shows the residuals to the unity relation, with axes scaled to exclude the larger outliers so as to more clearly show the scatter around the zero (grey dashed line).

practice, a jitter term is included by adding a constant value J to the diagonal of the covariance matrix. We fit for J alongside the other hyperparameters, resulting in more robust uncertainties on the latter. The full set of parameters and GP hyperparameters that are solved for, and their associated prior distributions are given in Table 3.

Initial guesses for each hyperparameter were defined as follows. The GP period P was set to the stellar rotation period P_{rot}, and the GP length scale l to the spot decay time-scale τ. The GP mean and variance A were set to the sample mean and variance of the simulated light curve. Since there is no single parameter of our stellar spot model that relates directly to Γ, we adopt an initial guess of $\Gamma = 1$ across all models. This value marks the transition between high and low harmonic complexity cases (Barragán et al. 2022), and was chosen because it allows easy exploration of parameter space in either direction.

The optimum GP hyperparameters were found using Markov Chain Monte Carlo (MCMC) with the python package EMCEE (Foreman-Mackey et al. 2013). The MCMC was run with 80 walks for to a maximum of 10 000 steps, unless the convergence was reached sooner. MCMC analyses struggle with 'perfect', error-free data, so to help with convergence we specify a small uncertainty

of 10^{-4} for the simulated data. The convergence of the MCMC walkers was checked and measured using the same method outlined in Section 2.3. For the solutions that met this convergence criteria, a 'burn-in' of twice the maximum chain autocorrelation length was removed, and the chains were thinned by half the minimum autocorrelation length.

Given their greater complexity compared to the GP sample curves, the posterior distribution of the GP analysis of the light curves contains more local minima in which walkers can get 'lost'. We mitigate this in two ways: First we set the walkers to follow a weighted combination of differential evolution (80 per cent probability) and differential evolution snooker (20 per cent probability) moves, as this is suited for multimodal posterior distributions. Secondly, we implement a three stage burn in: the MCMC was run for 300 iterations and then reinitialized from the highest probability solution, repeating twice before the main MCMC run. If by the end of the main run there was not a converged solution for any of the reasons outlined above, a 'burn in' of 5000 steps was removed from the start of the chains, the same test for convergence run again. If a solution passed, the same procedure for thinning and removing the 'burn-in' was used as for the MCMC solutions that converged normally. If the solution still failed convergence, then no further burn-in or thinning was applied and a flag allocated to note it as a non-converged solution.

3.3 Results

For all converged MCMC solutions, the final value was taken as the median value of the posterior distribution, and we estimate an uncertainty by calculating the difference between the 50th and 16th percentile for the lower bound, and the 84th and 50th percentile for the upper. Any non-converged solutions were excluded from the analysis below. See Appendix A for further discussion on the handful of model light curves for which the QP GP did not converge to a single set of solutions.

3.3.1 Stellar rotation period and QP GP period

A comparison between the stellar model rotation periods and the output QP GP period is shown in Fig. 3. As seen in the top panel, the QP GP recovers the stellar rotation period for all but one of the model light cures. For the two outlier points, the recovered period values are approximately a half and a tenth of the input model stellar rotation periods, owing to the particular distribution and evolution of the spots creating signals at those periods.

The bottom panel shows the residuals of the GP period from the input model rotation period, excluding the outlier point to better see the scatter about zero. This scatter around zero (perfect recovery of the rotation period) is evenly distributed above and below, and increases for shorter rotation periods. This increased scatter is due to the finite sampling, with the longer periods being better sampled by the two points per day in our model light curve, and hence are better recovered by the GP.

3.3.2 Stellar spot evolution time, τ, and QP GP evolution time-scale, l

Fig. 4 compares spot evolution time-scale, τ, of the model stellar light curves with the QP GP length scale, l. The grey dashed line highlights the 1:1 correlation, the purple vertical dashed line shows the total time span, T, of the model light curves of 250 d, and the
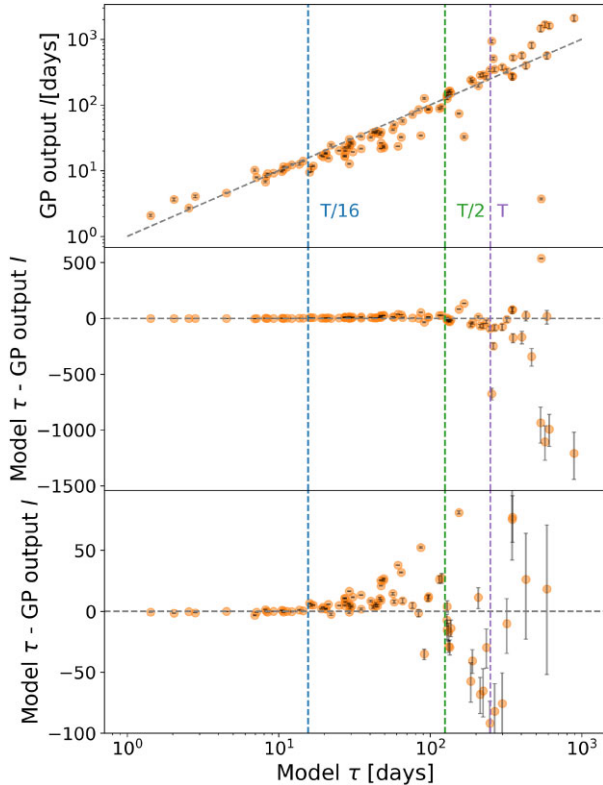
Figure 4. This figure compares the input stellar model spot evolution time, τ, to the solutions QP GP evolution time-scale, l, for the set of model light curves (top panel), the residuals (middle), and the residuals with y-axis adjusted to highlight features at smaller τ (bottom). The grey dashed line indicates the unity relation, and the vertical dashed lines indicate the total (purple, T), half (green, T/2) and sixteenth (blue,T/16) time span of the data.

green and blue vertical dashed lines indicate a half and sixteenth of the time span, respectively.

These results show a strong correlation between the spot evolution time-scale and the QP GP length scale hyperparameter. In the case of the specific way we have defined the spot behaviour in our stellar models, this is nearly a 1:1 relation, particularly at smaller values of τ. The deviation from 1:1 recovery increases with increasing time-scale τ as the values of τ approach the total time-span of the data, T. The QP GP length scale underestimates τ for values of τ between approximately T/16 and T/2, past which the GP overestimates τ with increasing scatter beyond τ greater than the time span of the data.

In the stellar models used for this work, the spots emerge on the stellar surface five times faster than they decay, a choice motivated by observations of spot emergence and decay times on the Sun. To see whether this has an impact on the recovered QP GP length scale, we generated another set of simulated light curves with equal emergence and decay times, and modelled them in the same way. The results are shown in Fig. 5. Comparing this to Fig. 4, the scatter around the 1:1 line is smaller for equal emergence and decay times. Again we see increasing scatter with increasing τ, but this only becomes marked for values of τ approximately equal to or greater than the total time span of the data set (purple dashed line). This result is unsurprising: when spot emergence is much faster than decay, the light curves are still dominated by changes occurring on the spot decay time, but this is occasionally perturbed by the emergence of a large spot causing a significant change on a much more rapid time-scale. This is more
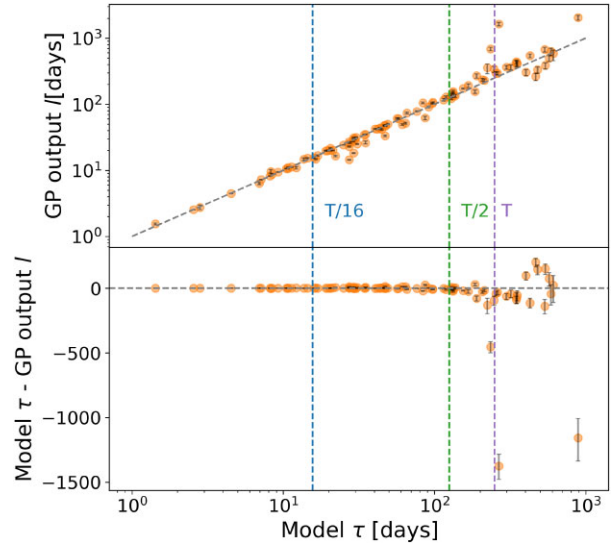

Figure 5. Same as Fig. 4 (direct comparison of values in the top panel, and residuals from the unity relation in the bottom panel), but for stellar models with equal spot emergence and decay times.

challenging for our GP model, which has a single evolution length scale, to reproduce.

3.3.3 Stellar model parameters and QP GP harmonic complexity, Γ

There is no intuitive physical interpretation for the harmonic complexity QP GP hyperparameter, so we examine how this hyperparameter behaves across all the input stellar model properties. Fig. 6 shows the solutions for the QP GP Γ with total spot number, stellar rotation period, spot evolution time-scale, and the ratio of spot evolution time to the stellar rotation period. We calculate Pearson product-moment correlation coefficients of -0.36 with total spot number, 0.66 with stellar rotation period, 0.62 with spot evolution, and 0.12 with the ratio of spot evolution time to stellar rotation period.

Since the Γ term governs how complex the curve is within one rotation period, the correlation with stellar rotation period results from longer period light curves being better sampled within one period, and so more complex behaviour can be resolved, giving a higher value for gamma.

The correlation with evolution time, however, is spurious, as demonstrated by the bottom panel. By dividing through by the period, we remove all correlation of the evolution time-scale with Gamma: at any given period, there is no correlation between τ and Γ.

The slight, though not significant, correlation with total spot number is likely due to the difference in distribution of spots (see e.g. Aigrain et al. 2015), which occurs when there are very few compared to many spots on the stellar disc. Given the uniform distribution imposed on the spot latitudes across all stellar models, models with a large number of spots will have an equally uniform spot distribution, but asymmetries in spot distribution will occur when there are very few spots.

Another influence in the outcome of these results are the input values and prior distributions used. Unlike the other hyperparameters that have initial values derived from the properties of the individual light curves, Γ is given a starting value of 1 for all light curves. The walkers thus have further to explore to find an optimum solution, and may become 'stuck' in a local preferred solution instead of finding the global optimum solution. However, as explained in Section 3.2,
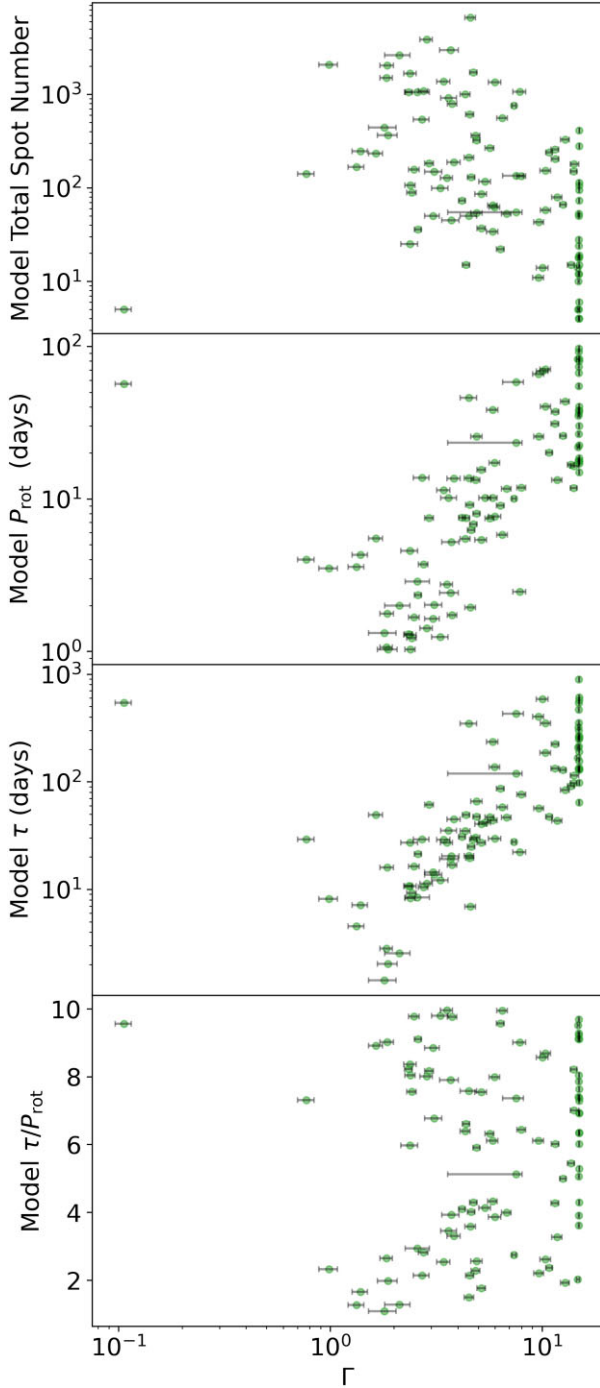
Figure 6. Light-curve MCMC solutions for the QP GP harmonic complexity, Γ, against the input stellar model parameters of total spot number, stellar rotation period, spot evolution time-scale, and the ratio of the spot evolution time-scale to the stellar rotation period.

this is mitigated by the multiple burn-in stages and setting a high maximum number of iterations to allow time for the walkers to find the best solution. The clustering of points near $\Gamma = 15$ is a result of the walkers reaching the upper bound of the uniform prior. Increasing this bound, however, does not change this behaviour, as it results from the GP trying to fit every small feature of the data that are due to the high-cadence sampling. This demonstrates the need for having a reasonable upper bound on Γ so as to prevent overfitting the data.
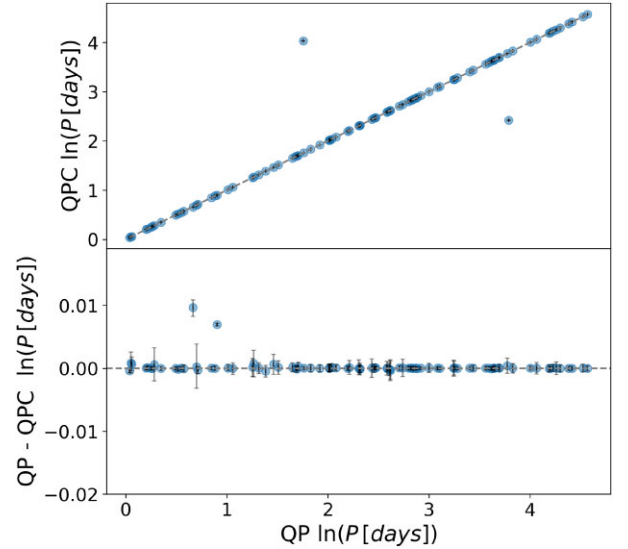
Figure 7. This figure compares the QP and QPC solution for ln (Period). Figure description is same as Fig. 3.

3.4 QPC versus QP GP analyses of light curves

The same analysis was repeated with the QPC kernel, with the same MCMC framework and convergence criteria. There were two non-converged solutions, one of which had also failed to converge in the QP case (see Appendix A for further discussion). When comparing the two analyses, all models that had a non-converged solution in either QP or QPC GP analyses were removed.

Figs 7, 8, and 9 show the comparison between the hyperparameters that are common to the QP and QPC kernels. Across all three hyperparameters the values QP and QPC values are almost identical.

For period, the agreement between the QP and QPC solutions is excellent, even at shorter period where scatter between the input rotation period and output QP GP period is higher. Additionally, there are one fewer outliers in the QPC kernel results. The QPC kernel is expected to better handle P/2 aliases in the data, but this is not seen in this case. The outlier in the QP case at $\sim 0.5P$ remains an outlier here, and has been underestimated ever further to ~ 0.1 of the rotation period. The outlier that was at $\sim 0.1P$ in the QP case, on the other hand, has converged to the input rotation value in the QPC case. The models with that are the largest outliers in period are also the largest outliers in l and Γ.

The results for evolution time-scale l between the QP and QPC analyses are also in excellent agreement. They show increasing scatter at longer l, though all agree within the MCMC solution uncertainties. This scatter begins to increase at T/2, and increases significantly past T, as seen most clearly in bottom panel of Fig. 8.

Fig. 9 demonstrates the excellent agreement between the QP and QPC solutions for Γ. In particular this is highlighted in the lower panel of this figure, which has been scaled to exclude the largest residuals and better see this scatter about zero. Above $\Gamma \sim 2$ (and ignoring the points clustered at the edge of the uniform prior boundary values of 15), the scatter in the values is smaller than the MCMC uncertainties. Within those uncertainties, all residual points can be said to fall on zero, although all QPC points are systematically slightly larger than the QP values.

The near identical results seen between the QP and QPC kernels presented above can be explained by the distribution of f values. As shown in Fig. 10, all but a few of the converged solution have solution
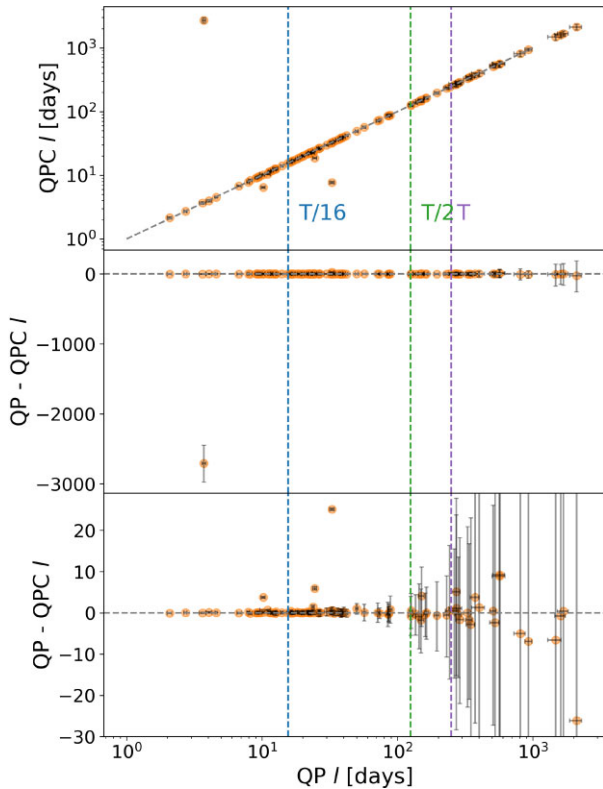
Figure 8. This figure compares the solutions of the QP and QPC GP evolution time-scale, *l*. Figure description is same as Fig. 4.
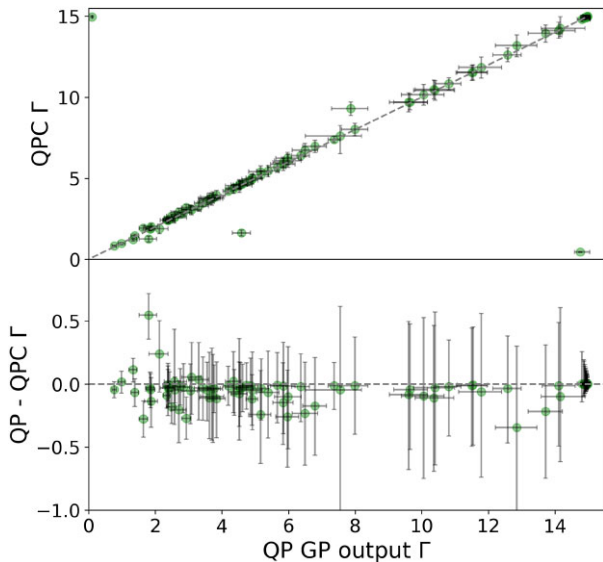

Figure 9. Top panel of this figure compares the QP and QPC kernel solutions for the harmonic complexity, Γ, with grey dashed line indicating the unity relation. Bottom panel shows the residuals around the one-to-one relation, with y-axis scaled to exclude the four largest outliers.

for *f* that are very small and tend towards zero. As such, these GPs are behaving more like the QP kernel, and hence giving near identical results for period, *l* and Γ. This is surprising, as more than a third of our model curves have some signal at half the rotation period (seen as a peak in their autocorrelation function at P/2), and thus are expected to have values of *f* greater than zero. It is possible that the cosine
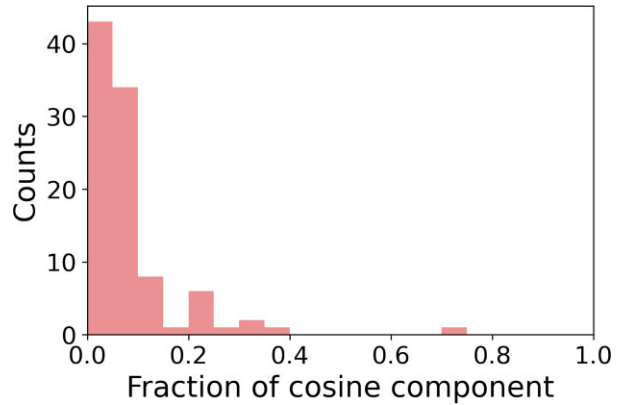
Figure 10. This figure shows a histogram of the fraction of cosine component, *f*, values for the QPC GP for our model light curves.

term in the QPC kernel only requires a small amplitude to have a significant effect, but our results may also indicate that QP GP alone handles signals at P/2 almost as effectively as the QPC kernel.

4 APPLYING THE QP KERNEL TO MODEL RV TIME-SERIES

It is common practice to use photometric data to inform the priors for QP GP analyses of RV data, often used to mitigate stellar activity signals in the search for Keplarian signals from exoplanets. To explore how the hyperparameters of a QP GP fit to photometric data will compare to that of RV data, we use the same spot models used above to compute RV time-series using the following, very simple procedure (Aigrain et al. 2012): at each time-point, we compute the local RV of the stellar surface at the centre of each spot, and multiply it by the spot's contribution to the light curve (in relative flux units).

The resulting RV time-series have the same sampling and duration as the light curves. While they are not representative of real RV data sets, which are typically ground-based and have much sparser time-sampling, we use these RV simulations to explore the differences between the hyperparameters of the QP GP models trained on the photometric and RV time-series. The impact of the difference in time sampling between space and ground based observations will be explored separately in Section 5.

The same MCMC procedure was used to find solutions for the hyperparameters of a QP GP applied to the RV time-series, and the treatments of non-converged solutions was the same. The two stellar models that gave multimodal (non-converged) solutions in the photometric data, had converged, single solutions in the RVs, and a handful of the stellar models that had converged solutions for the LC data had non-converged solutions here. To best compare the LC and RV results, all models that non-converged solutions in either analysis were removed.

4.1 RV versus light-curve period

The comparison of the solution to the QP GP period from light curve versus RV data is shown in Fig. 11. The top panel plots the GP solution for the light curve against those of the RVs, with the grey dashed line showing the unity relation, and the bottom panel shows the residuals from this, excluding the outlier seen in the upper panel results from the underestimated period of one light curve QP GP solution. The period solutions for the RV QP GP shows no such outliers, and recover well the model stellar rotation periods.
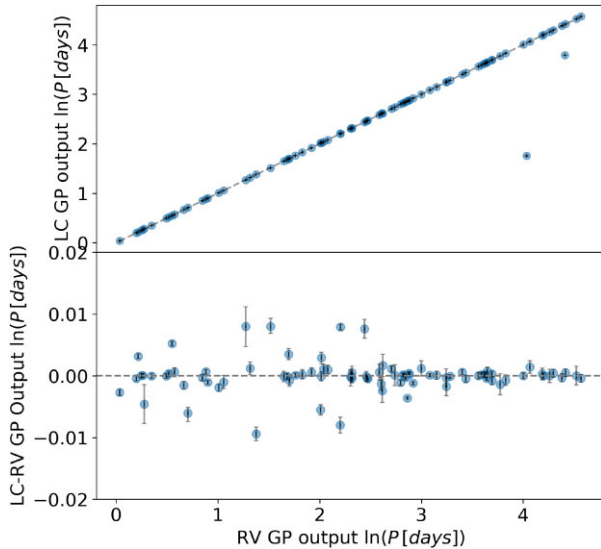
Figure 11. The top panel compares the output QP GP period (in natural log) for the radial velocity (RV) data with the light-curve (LC) data, with the unity relation indicated by the grey dashed line. The bottom panel shows the residuals to the unity relation, with axes scaled to exclude the larger outliers and more clearly show the scatter around the zero (grey dashed line).
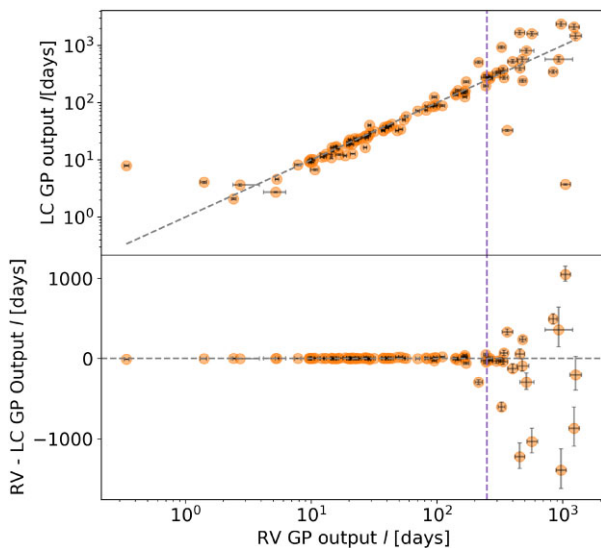

Figure 12. This figure compares the solutions QP GP evolution time-scale *l* for model light curves (LC) with the solutions for the radial velocity (RV) data. The grey dashed line indicates the unity relation, and the purple vertical dashed lines indicate the total time span of the date of 250 d.

The lower panel show a greater scatter in the residuals for shorter periods, reflecting the scatter seen in the light-curve GP solutions. Overall, the light curve and RV QP GP solutions for the period are in excellent agreement, both reflecting the stellar rotation period.

4.2 RV versus light curve *l*

Fig. 12 compares the QP GP length scale solution from RV time-series versus the solution from model light-curve data, with the grey dashed line indicating the unity relation, and the vertical purple dashed line showing the total time span of the data of 250 d. The light curve and RV solutions are in good agreement, but differ more
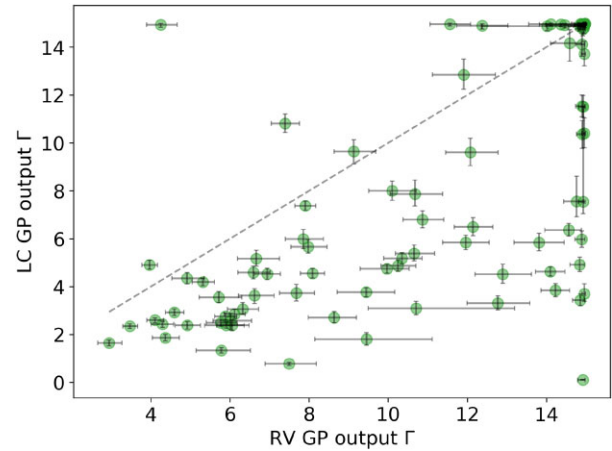

Figure 13. The RV recovered harmonic complexity term (Γ) of the QP GP, versus that from the light-curve (LC) data. The unity relation is shown by the grey dashed line.

at larger length scale, particularly for those length scales greater than the time span of the data.

4.3 RV versus light-curve Γ

Fig. 13 compares the RV and light-curve solutions of the QP GP harmonic complexity, Γ, with the 1:1 relation shown as a grey dashed line. Unlike the period and the length scale, the RV and light-curve QP GPs have very different solutions for Γ. The RV solution gives a consistently larger Γ than the light-curve data from the same stellar model. Further, the moderate correlations seen in the light-curve solutions between Γ and the stellar rotation period is reduced in the solutions for RVs, with a Pearson correlation coefficient of 0.5. There remains a slight, but weaker anticorrelation with total spot number of -0.29.

The lack of agreement between the harmonic complexity term between RV and photometric data has implications for the way we approach activity modelling in RV data sets (see Section 6).

4.4 QPC versus QP GP analyses of RVs

We repeat the GP analysis of the model RVs with the QPC kernel, again using the same MCMC framework, convergence criteria, and treatment on non-converged solutions. Only one model had non-converged solutions, and this was a different model to those that did not converge in any other GP analysis run (see Appendix A for further discussion).

A comparison of results between the QP and QPC kernels for model RV data are shown in Figs 14, 15, and 16. As was the case for the analysis of the model light curves, the QPC solutions for period, evolution time-scale and harmonic complexity are nearly identical to those of the QP kernel.

Examining the residuals of the period comparison (lower panel of Fig. 14), there are three notable outliers. These models are also the three largest outliers in Γ but not in *l*. As in the light-curve analysis, the scatter in the difference between the QP and QPC kernels is smaller than the uncertainty, and does not vary with period.

Similarly with the evolution time-scale, *l*, aside from some outliers, the scatter in the residuals about zero is smaller than the uncertainties, with the scatter and degree of uncertainty increasing for longer time-scales.
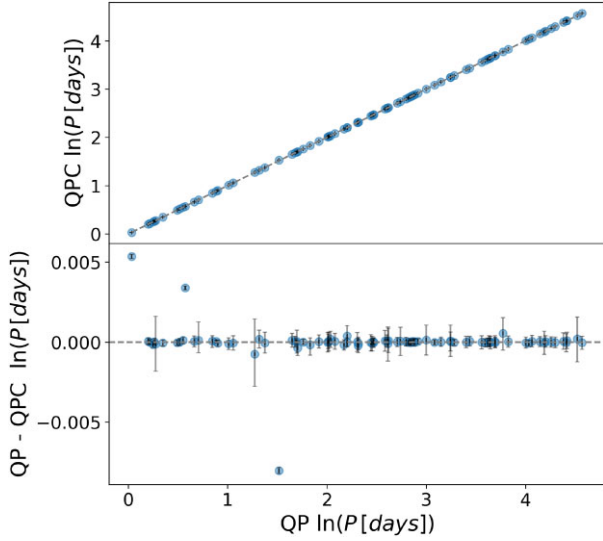
Figure 14. This figure compares the QP and QPC solution for ln (Period). Figure description is same as Fig. 3.
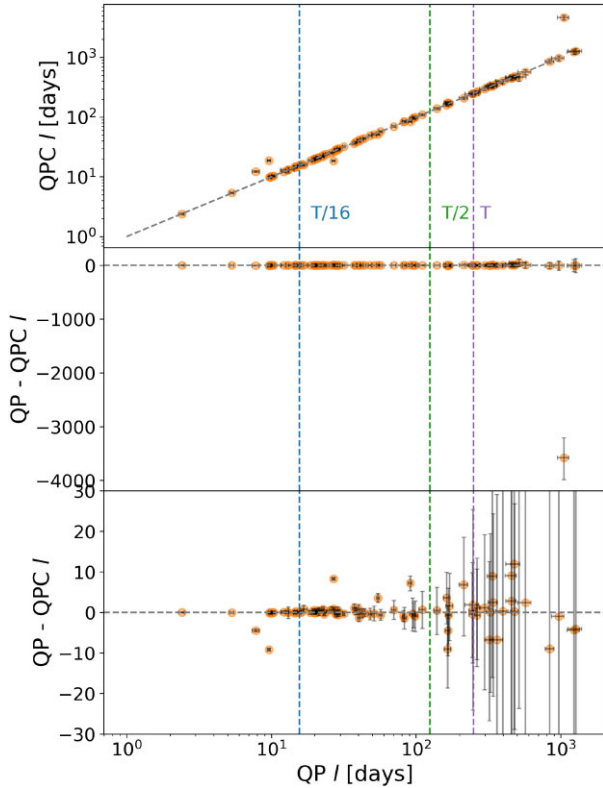

Figure 15. This figure compares the solutions of the QP and QPC GP evolution time-scale, *l*. Figure description is same as Fig. 4.

The harmonic complexity shows different behaviour here compared to the LC analysis, with the QPC Γ values being consistently higher than the QP solution for the same stellar model RV. This offset in the harmonic complexity can be explained by the higher fraction of cosine component *f* in the QPC solutions of the model RVs. As seen in Fig. 17, many models had a values of *f* above 0.1, and thus the cosine component was accounting for more of the intraperiod signal, leaving the harmonic complexity term to fit to more of the
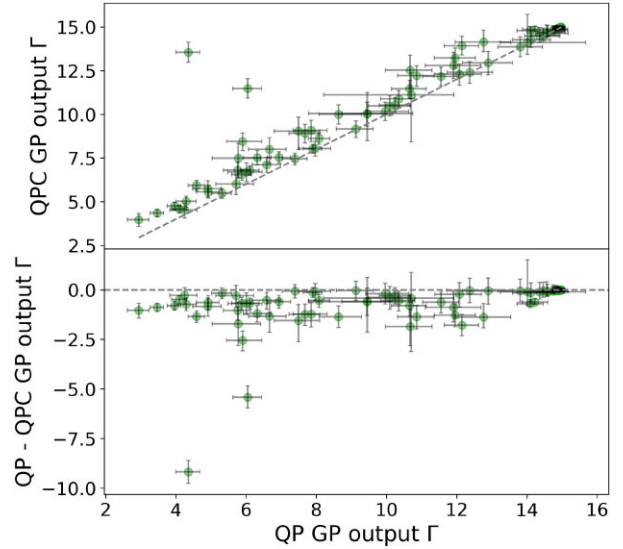

Figure 16. This figure compares the QP and QPC GP solutions for Γ for analysis of model RV data with the unity relation shown as a grey dashed line (top panel), and the residuals from the 1:1 relation with the grey line indicating zero residuals (bottom panel).
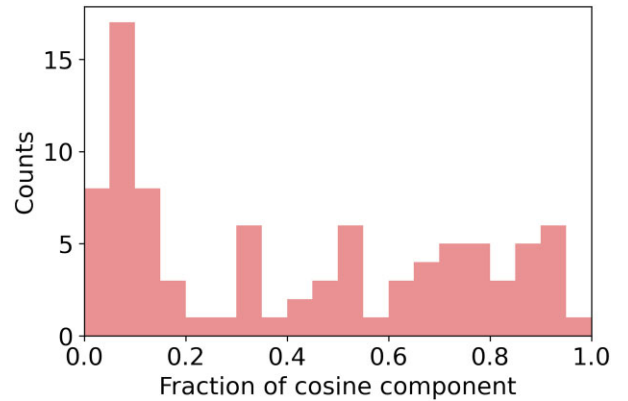

Figure 17. This figure shows a histogram of the fraction of cosine component, *f*, values for the QPC GP for our model light curves.

subtle variations from point-to-point, increasing the complexity term overall – without the larger, simpler signal to fit to, the Γ parameter was instead sensitive to the smaller scale changes.

5 IMPACT OF RV TIME SAMPLING

RV observations are typically ground-based, and as such their time-sampling is generally much sparser than that of the simulated data sets analysed in the previous section, and have much higher uncertainties. To test the application of QP GP regression to more realistic RV data, we add noise to these data representative of the uncertainties expected at world-leading precision radial velocity facilities with a mean uncertainty of $20\,\mathrm{cm\,s^{-1}}$, and a standard deviation of $15\,\mathrm{cm\,s^{-1}}$ to approximately reflect the night-to-night variations caused by different atmospheric conditions. We then re-sample the model RV curves with the type of sampling expected from ground-based observatories by choosing a real star at random, and calculating its observability from a chosen observatory given seasonal and airmass constraints, as one would when planning a RV observing season. We

then randomly selected *N* observations within the observable times, over a time span of 250 d, with $N = 250, 100, 50,$ and 30. While 30 to 50 points per season are typical of fairly intensive, present-day RV monitoring campaigns, 100 to 250 points per season corresponds to the planned sampling rates for future EPRV surveys aiming to detect Earth analogues, such as the Terra Hunting Experiment (THE; Hall et al. 2018). To ensure adequate sampling of the rotation periods for even the most rapidly rotating simulated stars, we added a further constraint that the shortest gap between any two data points in a given time-series should not be longer than the shortest rotation period in our simulations. This mimics what an observer would do when planning observations.

An example of this progressive degradation of the time-sampling and increased errors is shown for one of the stellar models in Fig. 18, with the data points shown in grey. The re-sampled light curves we analysed in the same way as light curves and the original RV curves, using the same MCMC sampling and convergence criteria. As before, we exclude any non-converged solutions from our analysis. The lines in each panel of 18 are samples from the QP GP conditioned on the data, with hyperparameters drawn from the MCMC posterior distribution. Larger scatter between the curves indicate that the GP model is less well constrained. The effect of reducing the number of observations by a factor up to 5 (down to $N = 100$) appears negligible in these plots, but it becomes noticeable when $N = 50$ or 30.

We then compare the input stellar rotation period and output GP period in Fig. 19. In the bottom, which shows the residuals from a one-to-one correspondence, we have excluded the clear outliers visible in the top panel, and zoomed-in to better show the behaviour of the remaining data points. With decreased sampling, the scatter around the input value increases for smaller values of the stellar rotation period, as these periods are less well resolved by the data. For d$N = 50$, only three cases with rotations periods below 4 d ($\ln(P) < 1.38$) converged. For $N = 30$ this number went down even further to one. This highlights the importance of sampling both the stellar rotation period and the orbital period(s) of any candidate planets adequately. For rapidly rotating stars, this necessarily translates into large numbers of observations per season unless the planet periods are also short.

In Fig. 20 we compare the input model spot evolution time to the QP GP solutions for the evolution time-scale in the re-sampled data sets. As seen with the rotation period solutions, there is greater scatter at smaller values of the spot evolution time. Further, in the 30 and 50 point data sets there are no converged solutions for evolutionary time-scales below 8 and 10 d, respectively. While the full sample overestimates the evaluation time-scale for mid-range values of spot evolution time, this overestimation improves at the lower sampling rate of 250 and 100 d, but then underestimates the evolution time in the 50 and 30 data point sets.

These differences between the time samples can be understood from the fact that the GP model is merely an approximation to the photometric and RV curves, a fact that becomes more noticeable with the greater sampling of the data.

The spot model used to generate the data leads to some modest but noticeable high-frequency features (sharp changes in the data due to the emergence or disappearance of individual spots) that are apparent in the LC and RV curves, and that the QP GP struggles to reproduce. At the highest (500) time-sampling this actually leads to non-convergence in the cases of the shortest spot evolution times.

At $N = 250$ and 100, the MCMC converges, but the derived *l* values are significantly lower than the injected spot lifetimes τ. This is because the GP needs as much freedom as possible to reproduce these high frequency features, and it achieves this by having a shorter *l*.
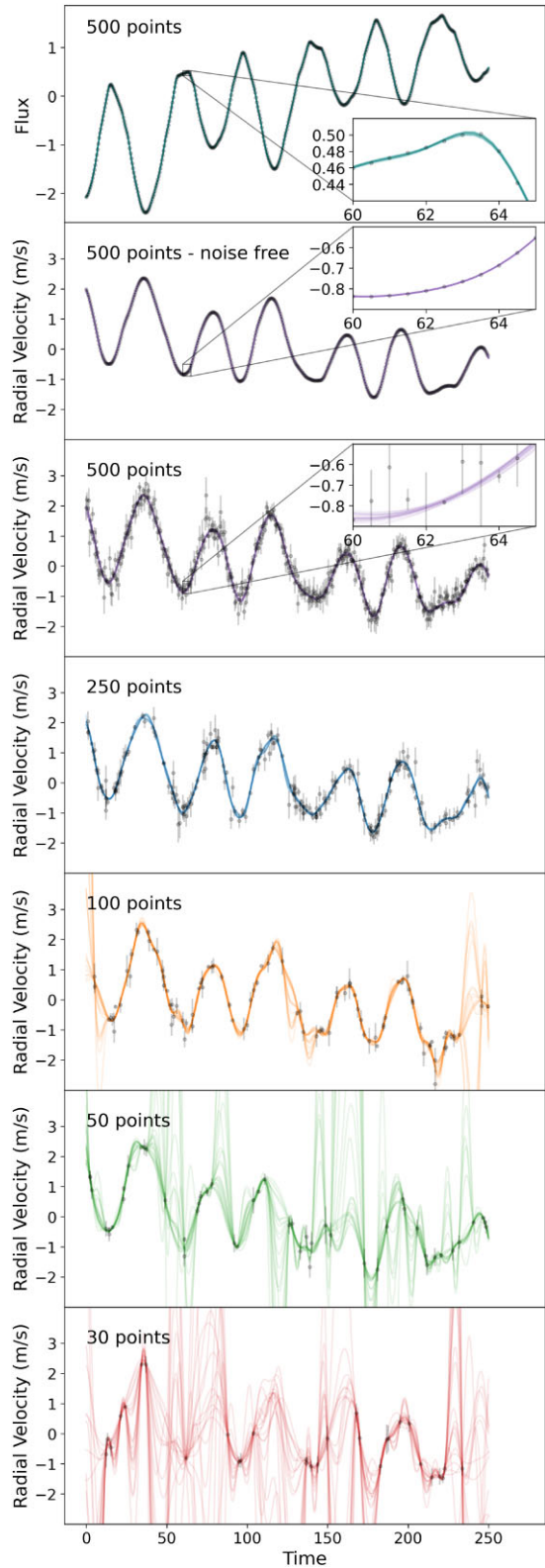

Figure 18. This figure shows an example model light curve (top panel), its associated RV data, and the RV data sets with degraded sampling. The lines plotted in each are the QP GP model with hyperparameters drawn from the MCMC posteriors.
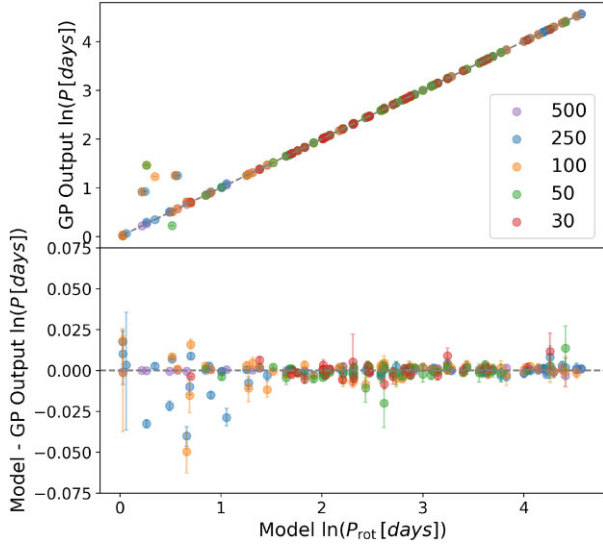
Figure 19. The top panel compares the input stellar rotation period to the output QP GP period (both in natural log) for different different time samplings of the model RV data, with the unity relation indicated by the grey dashed line. The bottom panel shows the residuals to the unity relation, with axes scaled to exclude the larger outliers and more clearly show the scatter around the zero (grey dashed line).
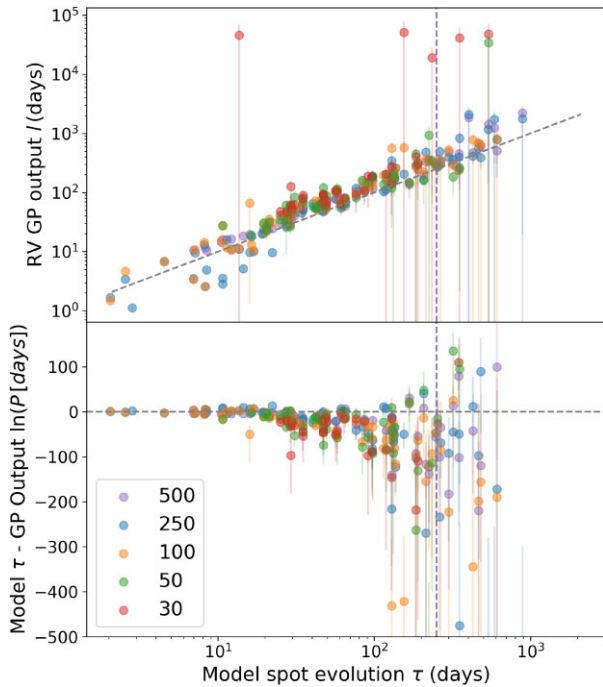

Figure 20. This figure compares the solutions QP GP evolution time-scale l for different re-samplings of the RV data. The grey dashed line indicates the unity relation, and the grey vertical dashed lines indicate the total time span of the date of 250 d.

As the time-sampling decrease further (N = 50 and 30), this no longer happens, as these high frequency features are no longer resolved, but we also lose the ability to recover short evolution time-scales due to them being undersample. These cases do not converge and are not shown on the plot. For intermediate values of τ, we observe that the values of l are systematically larger than tau for low
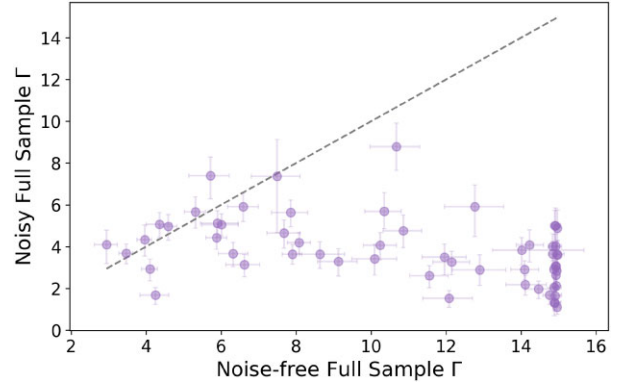

Figure 21. This figure compares the solutions QP GP harmonic complexity Γ for 'perfectly' sampled RV data with and without noise. The grey dashed line indicates the unity relation.
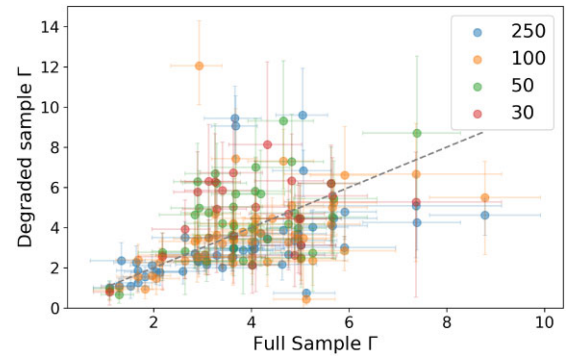

Figure 22. This figure compares the solutions QP GP harmonic complexity Γ for 'perfectly' sampled noisy RV data that with degraded time sampling. The grey dashed line indicates the unity relation.

N. This is because with limited time sampling we can only place an upper limit on the rate at which the signal evolves. Once tau is long compared to the sampling it then becomes measurable by the QP GP. Note that all these considerations would apply to light-curve data as well, but our light curves were not simulated with reduced time-sampling because typically light curve with sufficient precision are measured from space with the high cadence sampling that allows.

Fig. 21 compares the Γ values obtained for the noisy versus noise-free RV time-series, with the full 500-point sampling. The cluster of *Gamma* values at the upper end of the prior range ($\Gamma = 15$), which was present in the noise-free case, disappears in the noisy case. We speculate that, in the noise free cases, the GP model attempts to reproduce the small, sharp changes caused by the rapid emergence of individual spots, by adopting a short intraperiodic length scale (i.e. a large value of Γ). As soon as even small amounts of white noise are added in the data, these small changes are no longer discernible. There is no apparent correlation between the individual Γ values obtained in the noisy and noise-free cases. This implies that, more than any physical property of the active regions, the noise properties of the data are the key factor driving the measured Γ values.

To assess the impact of sampling on Γ, Fig. 22 compares solutions to the noise-included full data set to the degraded data. Solutions below $\Gamma \sim 2.5$ across all sampling levels roughly agree with the full sample solutions. Beyond this value, the scatter in Γ values increases dramatically, regardless of the sampling. This again confirms that the Γ parameter is predominately driven by noise, which is on the same scale across all samplings of data.

6 DISCUSSION AND CONCLUSIONS

This work seeks to find what physical meaning, if any, can be attributed to the QP GP hyperparameters derived from photometric time-series. In the context of the use of photometric data to inform stellar activity jitter models in RV data, we investigate the similarities and differences between the QP GP hyperparameters of photometric and RV data. Finally, we look at the impact of time sampling on RV data on the solution to a QP GP, and the impact this has on the inferred stellar properties.

6.1 The relation between QP GP hyperparameters and stellar properties

To investigate the physical meaning of QP GP hyperparameters, we solve for them for a set of model light curves with varying rotation periods and spot evolution times with MCMC, and compare the output hyperparameters to the input stellar properties. The results from this showed almost perfect recovery of the input stellar rotation period, with increasing scatter for shorter periods due to the finite sampling of the data. This result is expected, and has been indicated previously. Angus et al. ([2018](#)) conducted a similar study to the work presented here focusing on the recovery of rotation periods from *Kepler* light curves by a QP GP, using a similar spot model to validate their results. The results in this work form a much tighter correlations with period, and fewer cases of the QP GP finding an alias of the true stellar rotation period due to the lack of noise included in our stellar model light curves.

We also find a correlation between the spot evolution time and the QP GP length scale, consistent with behaviour demonstrated by Santos et al. ([2021](#)) between spot evolution time and light-curve autocorrelation decay time-scales. Toward longer spot evolution times the GP overestimates the evolution length scale, with increasing deviation and scatter when the spot evolution time-scales approach the time span of the data. This correlation is tighter when the spot emergence and decay times are equal. This behaviour is a natural consequence of the fact that the GP kernel has only one parameter to describe the evolution of the periodic signal over time. If the spot emergence times are considerably shorter than the spot decay times, the simulated light and RV curves contain both shorter term changes due to spot emergence, and slower changes due to spot decay. The best-fitting GP evolutionary time-scale is then a compromise between those values, and less tightly correlated to either. The longer period single is the one that is preferentially found, as it is better sampled by the data. This indicates that the QP GP length scale can be used as an indicator of spot decay time and evolution.

The harmonic complexity term, Γ does not have an intuitive physical interpretation, but does display some correlation with physical stellar properties. The Γ parameter is weakly correlated with total spot number, though not at all with the average spot number. We also find Γ to correlate moderately with rotation period and spot evolution time, although the latter is shown to be a spurious correlation. The simplicity of our spot model does not allow for a deeper exploration of the possible influences on Γ. Results from P21 indicate some influence of spot distribution on this hyperparameter. Future studies involving more complex stellar models with varying spot emergence patterns and differential rotation may reveal more about the usefulness on the Γ hyperparameter in probing the physical properties of stars from their photometric and RV time-series.

6.2 QP GP regression on photometry versus RVs

We perform the same analysis on time-series of RV data points to investigate if the same QP GP hyperparameters can be applied to both types of data. We show in this work that, although the period and the length scale hyperparameters match for both photometric and RV data, the harmonic complexity term does not: it is systematically higher in RV than photometric time-series. This can be explained as the RV time-series will behave as a combination of the photometric time-series and its first derivative, creating a more complex change in the RV with time compared to the photometry. It is important to note, however, that the model curves used in this analysis can be considered in the 'high complexity' regime for the QP GP harmonic complexity. Barragán et al. ([2021a](#)) show that this behaviour of the harmonic complexity between a time-series and its derivative is dependent on the complexity of the time-series. A high harmonic complexity time-series will have a derivative with a higher harmonic complexity. For time-series data with low harmonic complexity, however, the harmonic complexity of the time-series will be almost equally low. This explains the discrepancy between the findings here and those of Kosiarek & Crossfield ([2020](#)), who compare QP GP hyperparameters between Solar photometric and RV observations, and find a marginally lower harmonic complexity (their η_4 'length scale' is equivalent to $1/\Gamma$) in three of their RV data compared to the photometry, and the opposite in another, though in none of the sets are these values as drastically different as we observe in some of our models. We conclude from this that it is not appropriate to blindly use the harmonic complexity derived from photometry for analysis of RV time-series.

One common scenario not investigated in this work is the use of photometric data to constrain a GP on RVs taken months or years after the photometric observations, as happens in the case of RV follow-up of planet candidates from transit surveys. This investigation is left to future work, as it requires a more sophisticated stellar model than the one used here. Such an investigation would require activity-cycle-like variations in addition to individual spot evolution and rotational modulation to best estimate the efficacy of using non-contemporaneous photometry.

6.3 Comparative performance of the QP and QPC kernels

The QP and QPC kernels behave almost identically for the recovery of stellar rotation period and spot evolution times in both photometric and RV data. In particular for photometric data, the QPC tended to have very small values of cosine fraction f, and so was largely behaving as a QP kernel. Since we allowed the harmonic complexity Γ of the QPC kernel to vary in our analysis, it is possible that any variability present at half the rotation period was explained by higher Γ, rather than higher f. P21 explored this potential degeneracy by fixing the harmonic complexity and varying f only. We did the same test, repeating our QPC analysis of the model light curves with Γ set to 5.2 (corresponding to the value of $\omega = 0.31$ adopted by P21, where $\Gamma = 1/2\omega^2$). However, we found that this had almost no impact on the resulting distribution of f values, which were almost identical to those obtained when Γ was allowed to vary. In fact, the number of light curves with $f < 0.1$ increased slightly. On the other hand, fixing Γ significantly increased the scatter in the values of l for $\tau > T/2$, and led to poorly constrained solution for the mean in around 10 per cent of our simulations.

In RV data, the fraction f of the cosine component was higher on average, as was the harmonic complexity of the sine-squared term.

Even so, the ability to recover periods and evolution time-scales from RV data is largely the same with either kernel.

Both kernels gave rise to at least one significant outlier in our sample of 100 simulated stars, where the stellar period and evolution time-scale τ were very different from the GP period and length scale l. The QPC kernel gave rise to even fewer of these outliers than the QP kernel, but the total numbers are very small in either case.

We can also compare the QP and QPC kernels in terms of MCMC convergence. Both kernels led to non-converged solutions for some models: 2 and 10 in total for the QP kernel on light curves and radial velocities, respectively, compared with 2 and 1 for the QPC case. In real-world scenarios, though, the MCMC set up can be tuned to achieve convergence for each individual case.

Overall, the QP and QPC kernels are comparable in performance, both having advantages and drawbacks. The QP kernel is simpler, with fewer parameters, and in most cases we do not find sufficient evidence of superior performance with the QPC kernel to justify the additional cost. The only advantage conferred by the QPC kernel on our sample appears to be improved MCMC convergence in around 10 per cent of the RV curves, though we were unable to identify any obvious reason why convergence was more difficult to reach in those particular cases. Future work with more realistic spot models and spot distribution patterns may shed more light on this.

6.4 The impact of 'realistic' noise and time sampling on QP GP analysis of RVs

To reflect the scenario where there is only RV data for the QP GP analysis, we degrade the time sampling of our RV data set to better reflect the observations from ground-based observatories. When it comes to the correlation of these degraded QP GP, the results show that the recovery in evolution time and period is limited to both the total time span of the data and the minimum separation of the data, the former limiting the ability to reproduce long evolution times, and the latter impacting the ability to sample the shorter period and evolution's times the larger this minimum separation becomes with decreasing amounts of data. It is also apparent from the insets in Fig. 18 that the fewer data points give fewer constraints on the GP solution, making it increasingly more flexible. This means that there is more and more degeneracy between a GP-based activity model for the RVs and any potential planetary signal. This highlights the importance of adequate time sampling in RVs, both in total time span, and in sampling frequency.

ACKNOWLEDGEMENTS

DATA AVAILABILITY

The stellar model data can be found in the supplementary information to this paper, and the code used to generate them can be found here: https://github.com/saigrain/pyspot (Aigrain 2021). The code used for the QP GP analysis and time sampling is available on request.

REFERENCES

Aigrain S., 2021, pyspot: Simulate Light and Radial Velocity Curves of Spotted Stars. Available at: https://doi.org/10.5281/zenodo.5654179

Aigrain S., Pont F., Zucker S., 2012, MNRAS, 419, 3147

Aigrain S. et al., 2015, MNRAS, 450, 3211

Ambikasaran S., Foreman-Mackey D., Greengard L., Hogg D. W., O'Neil M., 2015, IEEE Trans. Pattern Anal. Mach. Intell., 38, 252

Angus R., Morton T., Aigrain S., Foreman-Mackey D., Rajpaul V., 2018, MNRAS, 474, 2094

Barclay T., Endl M., Huber D., Foreman-Mackey D., Cochran W. D., MacQueen P. J., Rowe J. F., Quintana E. V., 2015, ApJ, 800, 46

Barragán O., Aigrain S., Rajpaul V. M., Zicher N., 2022, MNRAS, 509, 866

Barragán O., Aigrain S., Gillen E., Gutiérrez-Canales F., 2021, Res. Notes Am. Astron. Soc., 5, 51

Basri G., Streichenberger T., McWard C., Edmond L., IV, Tan J., Lee M., Melton T., 2022, ApJ, 924, 31

Delisle J. B., Unger N., Hara N. C., Ségransan D., 2022, A&A, 659, A182

Dorren J. D., 1987, ApJ, 320, 756

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Foreman-Mackey D., Agol E., Ambikasaran S., Angus R., 2017, AJ, 154, 220

Fulton B. J., Petigura E. A., Blunt S., Sinukoff E., 2018, PASP, 130, 044504

Gilbertson C., Ford E. B., Jones D. E., Stenning D. C., 2020, ApJ, 905, 155

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

Grunblatt S. K., Howard A. W., Haywood R. D., 2015, AJ, 808, 127

Hall R. D., Thompson S. J., Handley W., Queloz D., 2018, MNRAS, 479, 2968

Haywood R. D. et al., 2014, MNRAS, 443, 2517

Jones D. E., Stenning D. C., Ford E. B., Wolpert R. L., Loredo T. J., Gilbertson C., Dumusque X., 2017, preprint (arXiv:1711.01318)

Kosiarek M. R., Crossfield I. J. M., 2020, AJ, 159, 271

Luger R., Foreman-Mackey D., Hedges C., 2021, AJ, 162, 124

Perger M., Anglada-Escudé G., Ribas I., Rosich A., Herrero E., Morales J. C., 2021, A&A, 645, A58 (P21)

Rajpaul V., Aigrain S., Osborne M. A., Reece S., Roberts S., 2015, MNRAS, 452, 2269

Rasmussen C. E., Williams C. K. I., 2006, Gaussian Processes for Machine Learning. MIT Press, Cambridge

Roberts S., Osborne M., Ebden M., Reece S., Gibson N., Aigrain S., 2013, Phil. Trans. R. Soc. A, 371, 20110550

Santos A. R. G., Mathur S., García R. A., Cunha M. S., Avelino P. P., 2021, MNRAS, 508, 267

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: THE NON-CONVERGED SOLUTIONS

Across all of the studies in this work, there was always at least one model for which the MCMC did not converge to a single set of solutions for the QP or QPC GP hyperparameters. Further, the models that did not converge were not always the same. This section explores some of the non-convergence scenarios.

Most of the non-converged MCMC solutions gave multimodal posterior distributions, an example of which is given in Fig. A1. This figure exemplifies what occurs in a majority of these multimodal
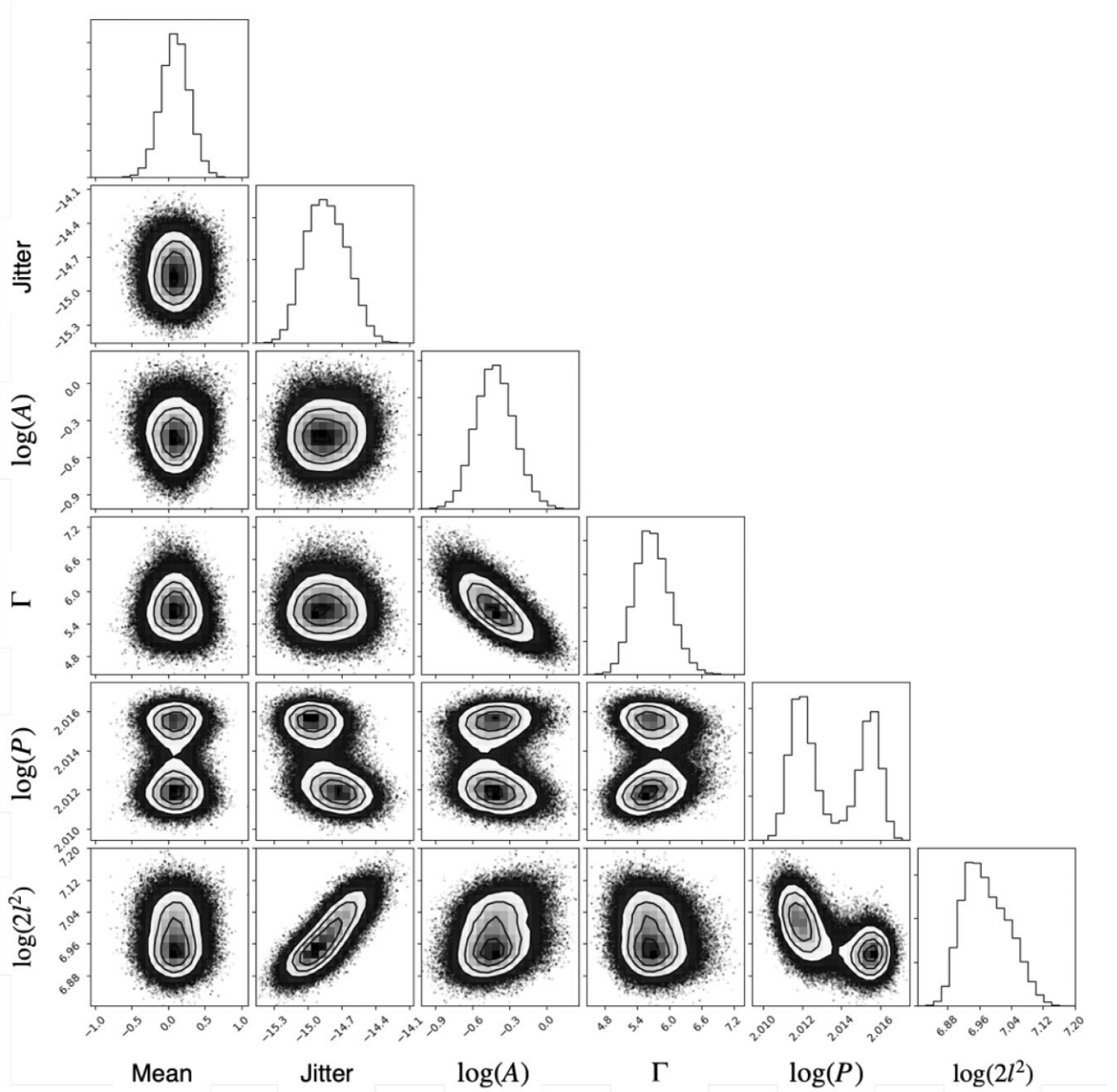
Figure A1. An example of bimodal posterior distribution of a QP GP fit to a model light curve.

cases of the QPC non-converges cases, and most of the QP non-converged cases: there are two equally probable solutions for the 'jitter' white noise term and the harmonic complexity term. Given their equal likelihood, the MCMC is unable to prefer one solution over the other. Looking at the jitter versus Γ panel in this figure we can see that they are correlated: the lower jitter values give solutions with higher harmonic complexity, and vice versa. Bimodality is seen in other hyperparameters as well, such as between l and period, or in period alone. The cases with bimodal posteriors have no similarities in their stellar parameters, and in the jitter-Γ bimodal cases the solutions to the period and evolution time-scale l still match the input model period and spot evolution time, respectively, so can be considered reliable solutions despite not strictly meeting our convergence criteria due to the bimodality in the posteriors of the other parameters. Such a bimodality in posteriors is possible in real data too, and in such circumstances, so long as the likelihood are

equal for both sets of parameters, they must be considered as equally likely given the data. In this work, however, given the lack of a single solution across all parameters these cases, for consistency across the results they are excluded from our analysis.

Some of the non-converged solutions for the QP analysis of the model RV data had unconstrained, poorly constrained posteriors, or complex posterior distributions, which meant that the convergence criteria were not met in the 10 000 step limit of our MCMC. As with the bimodal solutions, there is nothing particularly remarkable about the input stellar parameters of these models. The worst of the unconstrained solutions shown in Fig. A2. While this model did have a relative short rotation period with a short evolution time-scale, this was not an extreme case, and the QP GP was able to find converged solutions for even shorter rotations periods, and smaller ratios of rotation period to evolution time-scale. Trying an MCMC with a different 'move' setup did not alter the convergence of these cases.
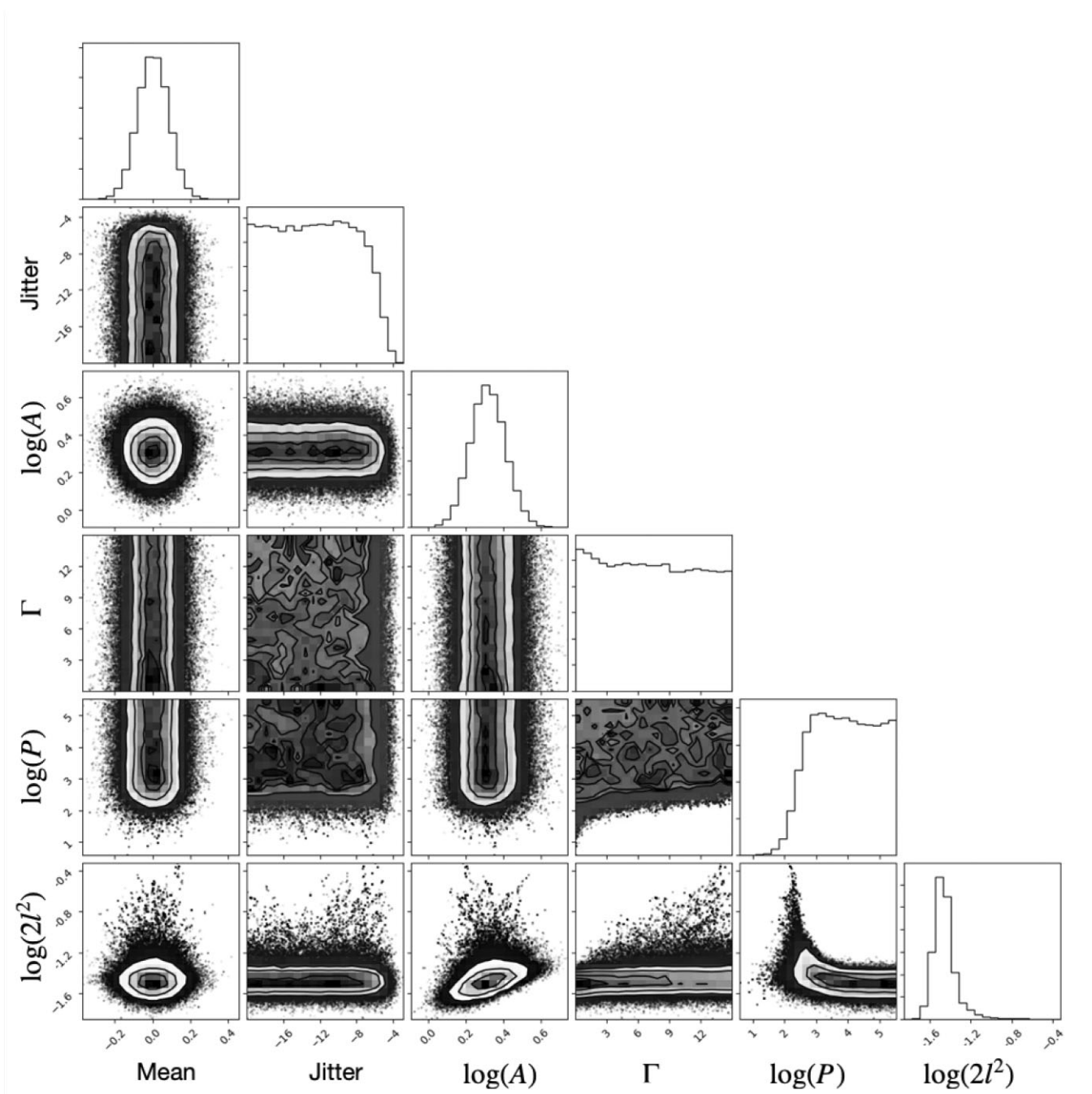
Figure A2. An example of the posterior distribution of an unconstrained MCMC solution.

This paper has been typeset from a TEX/LATEX file prepared by the author.