

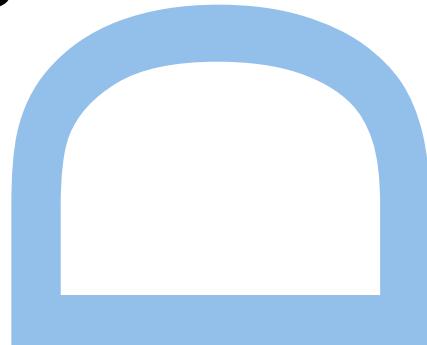
Advanced statistical data analysis methods for the detection of other Earths

João David Ribeiro Camacho

Programa Doutoral em Astronomia

Departamento de Física e Astronomia

2021

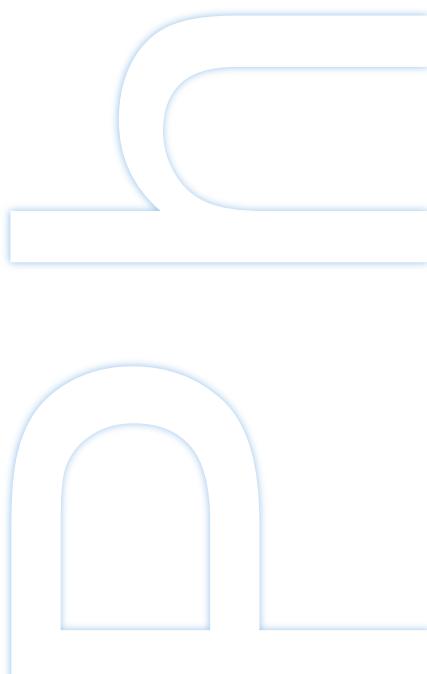


Orientador

Pedro T. P. Viana, Professor Associado, Faculdade de Ciências

Coorientador

João P. S. Faria, Investigador Auxiliar, Instituto de Astrofísica e Ciências do Espaço



Title page illustration:
Radial velocity curve of 51 Peg from Mayor & Queloz (1995)

Acknowledgements

Abstract

Resumo

viii Advanced statistical data analysis methods for the detection of other Earths

Contents

1	Introduction	19
1.1	Exoplanets search overview	19
1.1.1	Detection methods	21
1.1.2	Milestones	28
1.1.3	Statistical analysis	29
1.2	This thesis	33
2	Doppler spectroscopy	35
2.1	Mathematical background	35
2.1.1	Keplerian orbits	36
2.1.2	Orbital elements	38
2.1.3	Calculating the Keplerian function	41
2.2	Spectroscopic observations	45
2.2.1	The past, present and future of spectroscopy	45
2.2.2	Statistical and systematic errors	46
2.2.3	The cross-correlation function	48
2.3	The Stellar activity problem	52
2.3.1	Stellar activity features	52
2.3.2	Activity proxies	58
3	Gaussian Process Regression Network	63
3.1	Gaussian processes	63
3.1.1	The Gaussian distribution	64
3.1.2	The Gaussian process	65
3.1.3	The Covariance function	69
3.2	Gaussian processes frameworks	80

3.2.1	Rajpaul et al. (2015) framework	80
3.2.2	Jones et al. (2017) framework	81
3.2.3	artGPN framework	81
3.3	Gaussian processes regression networks framework	84
3.3.1	Variational inference	85
3.3.2	Simulating a quasi-periodic GP	89
3.3.3	Changing the weights	91
4	Application on radial velocity observations	95
4.1	Priors definition	95
4.1.1	Individual analysis	96
4.1.2	Combined analysis	97
4.2	Solar data	100
4.2.1	Sun observations	100
4.2.2	Results obtained on the solar data	105
4.2.3	Interpretation of the results	125
4.3	EXPRES data	133
4.3.1	HD 10700	133
4.3.2	HD 34411	136
4.3.3	HD 101501	138
4.3.4	HD 26965	140
4.3.5	Final thoughts on the EXPRES results	142
5	Conclusions	145
5.1	Future work	147
Appendices		149
A	Datasets summary	151
B	Corner plots	153

List of Figures

1.1	Mass-period distribution of all detected exoplanets. Source: http://exoplanet.eu/ To be updated until submission	20
1.2	Orbital motion of a star and a planet around a common center of mass, and the blue and red shifts on its light caused by orbit of the star.	22
1.3	Transit of a planet in front of a star and respective flux variation ΔF by it. The transit time duration is represented as t_T and the time of totality as t_F	24
1.4	The brown dwarf 2M1207 (centre) and its planetary companion (bottom left). Source: ESO	25
1.5	Gravitational microlensing amplified by an extrasolar planet.	27
1.6	Left: Hot-Neptune desert derived by Mazeh et al. (2016) with the dash-dotted lines delimiting the boundaries of this desert. Right: Radius gap from Fulton et al. (2017) where the gap around 1.8 Earth radii creates a bi-modal distribution of planets around Sun-like stars.	32
2.1	The two-body problem scheme.	37
2.2	Elliptical orbit of a celestial body in three dimensions. Source: Wikimedia Commons.	39
2.3	ESPRESSO's first light spectrum. Source: ESO/ESPRESSO team.	47
2.4	Simulated structure of a CCF.	50
2.5	Example of the equivalent width of a spectral line.	51
2.6	Solar granulation over an area of 19000 squared kilometers. Source: Big Bear Solar Observatory/New Jersey Institute of Technology.	54
2.7	A Sunspot capture by the Inouye Solar Telescope on January 28, 2020. Source: NSO/NSF/AURA.	56

2.8	Monthly mean of the number of sunspot observed since 1749. Source: https://wwwbis.sidc.be/silso	58
2.9	Simplified illustration of the impact of a spot on the CCF.	59
2.10	Expected behavior of the FWHM measurements on the example I gave on figure 2.9. On the x-axis are represented the five panes.	60
3.1	Three 1-dimensional Gaussian distributions. At solid blue a dis- tribution with $\mu = 0$ and $\sigma = 1$, at dotted red one with $\mu = 2$ and $\sigma = 0.5$, and at checkered green one with $\mu = -1$ and $\sigma = 2$.	65
3.2	Left: Three samples drawn from a prior distribution. Right: Three samples drawn from the posterior distribution after two observa- tions with the prior conditioned to these observations.	66
3.3	Samples of a squared-exponential kernel with different length- scales. Solid blue $\eta_2 = 1$, dashed red $\eta_2 = 10$, and dotted green $\eta_2 = 10$.	72
3.4	Samples of a Matérn 5/2 kernel with different length-scales. Solid blue $\eta_2 = 1$, dashed red $\eta_2 = 10$, and dotted green $\eta_2 = 10$.	74
3.5	Samples of a periodic kernel with $\eta_3 = 1$ but varying η_4 . Solid blue $\eta_4 = 0.1$, dashed red $\eta_4 = 1$, and dotted green $\eta_4 = 10$.	76
3.6	Scheme of our framework	82
3.7	Example of a GPRN diagram	86
3.8	Top: Samples of a GPRN using a node with periodic kernel and weight with a squared exponential kernel. Middle: Samples of a GPRN using a node with a quasi-periodic kernel and a weight with a squared exponential kernel. Bottom: Samples of a GP defined by a quasi-periodic kernel.	90
3.9	Samples of a GPRN using a node with a quasi-periodic kernel and a weight with a squared exponential kernel. The only different parameter being the η_2 of the weights. Top: Solid blue a GPRN using a weight with $\eta_2 = 10$ days. Middle: Dashed red a GPRN using a weight with $\eta_2 = 100$ days. Bottom: Dotted green a GPRN using a weight with $\eta_2 = 1000$ days.	92

4.1	Left: Binned RV, BIS, FWHM, and $\log R'_{hk}$ measurements of the Sun obtained with the HARPS-N solar telescope used in our analysis. Right: GLS periodograms of the binned measurements after removing a long-term trend to the data. The vertical solid lines indicate the period of 27 days. The horizontal dashed line show the false alarm probability (FAP) of 1%. These measurements are available at https://github.com/jdavidrcamacho/datasets . 101
4.2	Daily number of sunspot observed from the start of 2015 til the end of Solar cycle 24 on late 2019 (e.g. McIntosh et al., 2014; Nandy et al., 2020). Source: https://wwwbis.sidc.be/silso 103
4.3	Discrete cross-correlation functions between the RV and the BIS (solid blue), FWHM (dashed red), and $\log R'_{hk}$ (dash-dotted green). 105
4.4	Top: Superimposed fits of MAP values for the GP (solid red) and the GPRN (dashed blue) on the RV measurements. Bottom: Residuals obtained of the GP (red stars) and the GPRN (blue circles). 108
4.5	Top: Superimposed fits for the MAP values of the GP (solid red) and the GPRN (dashed blue) models on the BIS time series. Bottom: Residuals obtained of the GP (red stars) and the GPRN (blue circles). 110
4.6	Top: Superimposed fits of the MAP values for the GP (solid red) and the GPRN (dashed blue) models on the FWHM time series. Bottom: Residuals obtained of the GP (red stars) and the GPRN (blue circles). 113
4.7	Top: Superimposed fits of the MAP values for the GP (solid red) and the GPRN (dashed blue) models on the $\log R'_{hk}$ time series. Bottom: Residuals obtained of the GP (red stars) and the GPRN (blue circles). 115
4.8	Superimposed fits of the MAP values for the GP (solid red) and the GPRN (dashed blue) models on the RV and the BIS time series. On the bottom of each fit are the residuals obtained of the GP (red stars) and the GPRN (blue circles). 117

4.9	Superimposed fits of the MAP values for the GP (solid red) and the GPRN (dashed blue) models on the RV and the FWHM time series. On the bottom of each fit are the residuals obtained of the GP (red stars) and the GPRN (blue circles).	120
4.10	Superimposed fits of the MAP values for the GP (solid red) and the GPRN (dashed blue) models on the RV and the $\log R'_{hk}$ time series. On the bottom of each fit are the residuals obtained of the GP (red stars) and the GPRN (blue circles).	124
4.11	Left: Plots of the weight and node using the respective MAP values from the RV analysis. Middle: Product on the weight and node. Right: GPRN fit on the RV data points with the obtained parameters.	126
4.12	Top: ELBO and $\log \mathcal{Z}$ values given by the variational inference and DYNESTY, respectively. Bottom: Difference between the $\log \mathcal{Z}$ and the ELBO. The average value of the difference, of approximately 1.852, is plotted by the dashed horizontal line. To be updated until submission	127
4.13	Left: Structure of each component of the GPRN. Middle: Product of the node by each respective weight. Right: GPRN fit to the RV and FWHM when using the MAP values obtained on the analysis.	130
4.14	Left: Superimposed normalized predictives from GP and the GPRN. Right: Discrete CCF between the predictives of the RV and the activity indicator.	132
4.15	Left: RV observations of HD 10700. Right: Periodogram of the data after subtracting a long term trend. Also on it is marked, as a red vertical line, the expected rotation period of 34.5 days, and the FAP of 1% as a dashed horizontal line.	135
4.16	Left: Components of the GPRN. The weight ₁ connects the node to the RV, while the weight ₂ connects it to the FWHM. Right: Final fits of the GPRN to the RV and the FWHM.	135
4.17	Left: RV observations of HD 34411. Right: Periodogram of the data after subtracting a long term trend. Also on it is marked the FAP of 1% as a dashed horizontal line. No rotation period is known for this star.	137

4.18	Left: GPRN fit to the RV measurements of HD 34411. Right: Contribution of the node (top) and weight (bottom) to the final fit.	137
4.19	Left: RV observations of HD 101501. Right: Periodogram of the data after subtracting a long term trend. Also on it is marked, as a red vertical line, the expected rotation period of 17.1 days, and the FAP of 1% as a dashed horizontal line.	139
4.20	Left: GPRN fit to the RV measurements of HD 101501. Right: Contribution of the node (top) and weight (bottom) to the final fit.	139
4.21	Left: RV observations of HD 26965. Right: Periodogram of the data after subtracting a long term trend. Also on it is marked, as a red vertical line, the expected rotation period of 40 days, and the FAP of 1% as a dashed horizontal line.	141
4.22	Left: GPRN fit to the RV measurements of HD 26965. Right: Contribution of the node (top) and weight (bottom) to the final fit.	142
B.1	Posterior distributions of the parameter of the GP used in the individual analysis of the Sun's RV measurements.	154
B.2	Posterior distributions of the parameter of the GPRN used in the individual analysis of the Sun's RV measurements.	155
B.3	Posterior distributions of the parameter of the GP used in the individual analysis of the Sun's BIS measurements.	156
B.4	Posterior distributions of the parameter of the GPRN used in the individual analysis of the Sun's BIS measurements.	157
B.5	Posterior distributions of the parameter of the GP used in the individual analysis of the Sun's FWHM measurements.	158
B.6	Posterior distributions of the parameter of the GPRN used in the individual analysis of the Sun's FWHM measurements.	159
B.7	Posterior distributions of the parameter of the GP used in the individual analysis of the Sun's $\log R'_{hk}$ measurements.	160
B.8	Posterior distributions of the parameter of the GPRN used in the individual analysis of the Sun's $\log R'_{hk}$ measurements.	161
B.9	Posterior distributions of the parameter of the GP used in the combined analysis of the Sun's RV and BIS measurements.	162
B.10	Posterior distributions of the parameter of the GPRN used in the combined analysis of the Sun's RV and BIS measurements.	163

B.11 Posterior distributions of the parameter of the GP used in the combined analysis of the Sun's RV and FWHM measurements.	164
B.12 Posterior distributions of the parameter of the GPRN used in the combined analysis of the Sun's RV and FWHM measurements.	165
B.13 Posterior distributions of the parameter of the GP used in the combined analysis of the Sun's RV and $\log R'_{hk}$ measurements.	166
B.14 Posterior distributions of the parameter of the GPRN used in the combined analysis of the Sun's RV and $\log R'_{hk}$ measurements.	167
B.15 Fit and respective residuals of the MAP values from the GP on the RV and the FWHM time series of HD 10700.	168
B.16 Fit and respective residuals of the MAP values from the GP on the RV and the FWHM time series of HD 34411.	169
B.17 Fit and respective residuals of the MAP values from the GP on the RV and the FWHM time series of HD 101501.	170
B.18 Fit and respective residuals of the MAP values from the GP on the RV and the FWHM time series of HD 26965.	171

List of Tables

4.1	Parameters and prior distribution of the parameters in the individual analysis of both GP and GPRN. The y represents the input while t the time. On the individual analysis the input could either be the RVs, FWHM, BIS, or $\log R'_{hk}$	96
4.2	Parameters and prior distribution of the parameters when considering the analysis of the RVs with an activity indicator. y_{RV} represents the RV time series while y_{AI} the activity indicator (FWHM, BIS, or $\log R'_{hk}$) time series, as previously t represents the time.	98
4.3	Parameters obtained by the GP and the GPRN on the analysis of the RV time series.	107
4.4	Parameters obtained by the GP and the GPRN on the analysis of the BIS time series.	109
4.5	Parameters obtained by the GP and the GPRN on the analysis of the FWHM time series.	112
4.6	Parameters obtained by the GP and the GPRN on the analysis of the $\log R'_{hk}$ time series.	114
4.7	Parameters obtained by the GP and the GPRN on the analysis of the RV and the BIS time series.	116
4.8	Parameters obtained by the GP and the GPRN on the analysis of the RV and the FWHM time series.	119
4.9	Parameters obtained by the GP and the GPRN on the analysis of the RV and the $\log R'_{hk}$ time series.	123
4.10	Evidence values comparison, including old and new GPRN values.	128
A.1	Summary with the number of points, time spans, peak-to-peak amplitude, average error, and initial RMS of the datasets analysed on this thesis.	152

18 Advanced statistical data analysis methods for the detection of other Earths

Chapter 1

Introduction

How many extra-solar planets, also known as exoplanets, exist? That is a simple yet increasingly tricky question to answer. It is possible to say for a given day, but that same day a planet is found that surpass any expectations astronomers might have had. Twenty-six years after the first exoplanet detection we have seen exponential growth in the number of known exoplanets with no apparent decreasing rate. In this first chapter, I will summarise the main milestones that marked the field, the methods most commonly used, and how the field has changed our perception of the Universe.

1.1 Exoplanets search overview

If one asks what was the first exoplanet discovered, the answer might, surprisingly, depend on which side of the Atlantic the person is. As in any other field, earlier claims will push the boundaries of technology and scientific knowledge to the limit. That means such claims will likely be received with disbelief. Sometimes to independently verify the results years are required. Examples of these early claims are, for example, a companion to Barnard's Star by van de Kamp (1982) later disproved by Choi et al. (2013), or the Jupiter mass planet Gamma Cephei Ab made by Campbell et al. (1988) that was only confirmed decades later by Hatzes et al. (2003).

As of the 30th of September 2021, 4843 planets have been discovered¹. These populate a wide variety of orbital and mass ranges as shown by figure 1.1. Visually it is easy to identify three main groups. Giant planets at short periods were discovered using transit photometry, while giant planets at long periods were using radial velocities, and lower-mass planets at short periods where both transit and radial velocity methods are currently focusing their efforts. Next, I will focus on describing the main methods currently used for these exoplanets.

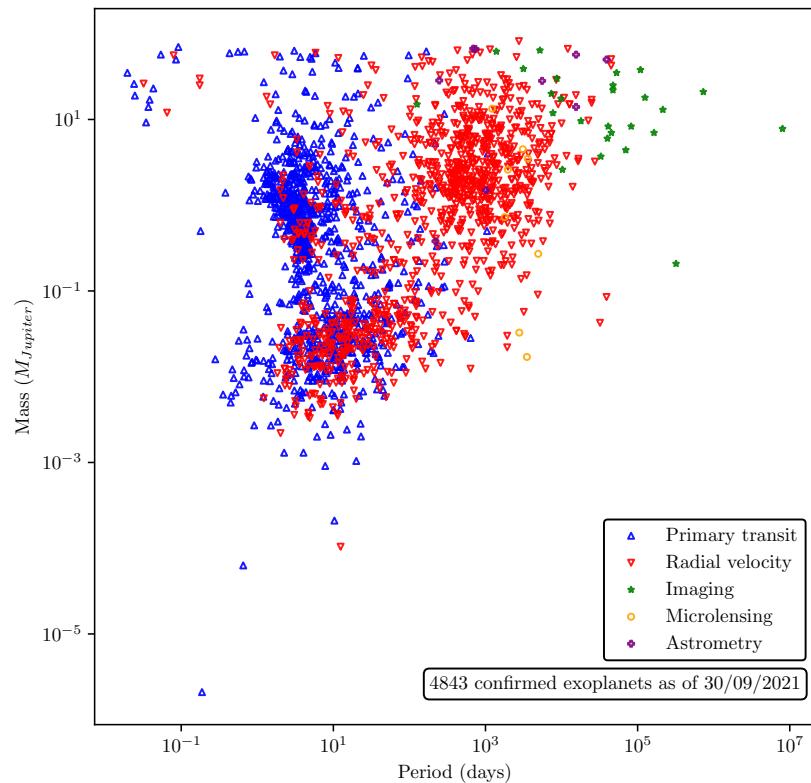


Figure 1.1: Mass-period distribution of all detected exoplanets. Source: <http://exoplanet.eu/> To be updated until submission

¹Source: <https://http://exoplanet.eu/>

1.1.1 Detection methods

There are several methods to detect exoplanets. Each one has unique strengths and limitations. The two most efficient for this task was the radial velocity and transits methods. These two discovered more than 4000 planets alone. Other successful methods include the use of direct imaging, astrometry, and gravitational microlensing. In this section, I will give a background on these five methods, how we can combine them to better characterize exoplanets, and what might be the future for these methods.

Radial velocities

The radial velocity method, also known as Doppler spectroscopy, was used by Mayor and Queloz (1995) to discover 51 Pegasi b. This method uses the shifts in the spectral lines of a star caused by the presence of a planet. When a planet exists, both will orbit a common centre of mass. This orbital motion will cause a blue and redshift in the star spectrum as it approaches or moves away from our line of sight, respectively. That is shown in figure 1.2. Precise spectrographs on Earth can then measure these periodic shifts.

If the orbits were perfect circles, the shift observed would have a sinusoidal shape. That would mean we could use a sine function to interpret the observations. Unfortunately, orbits are ellipses. That implies the orbits are elongated that is measured by their eccentricity. To explain this, a function a bit more complex than a sine is required. This function is commonly known as a Keplerian function.

Usually enough to translate the radial velocity shifts into physical meaningful information can be defined as

$$v_r = K[\cos(\omega + \nu) + e \cos(\omega)] + \gamma. \quad (1.1)$$

On this equation, K represents the radial velocity semi-amplitude. ν represents the true anomaly, ω as the argument of the pericenter, e the orbital eccentricity of the planet, and γ is the proper motion of the centre of mass (Murray and Correia, 2010; Perryman, 2011). See chapter 2 for a detailed derivation of this formula.

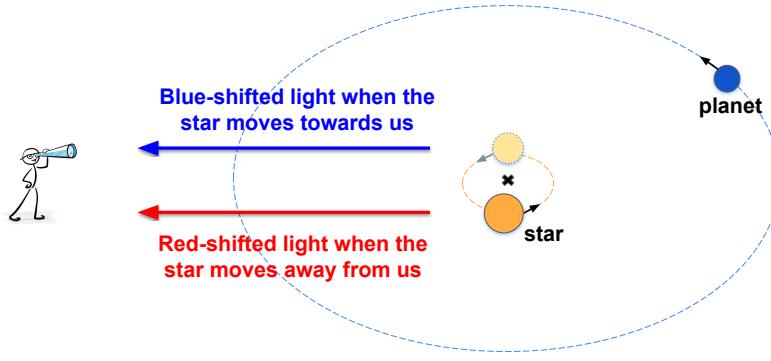


Figure 1.2: Orbital motion of a star and a planet around a common center of mass, and the blue and red shifts on its light caused by orbit of the star.

With some approximations, the radial velocity semi-amplitude is written as

$$K = 28.4 \text{ m.s}^{-1} \left(\frac{P}{\text{1 year}} \right)^{-1/3} \left(\frac{M_{\text{planet}} \sin i}{M_{\text{Jupiter}}} \right) \left(\frac{M_{\text{star}}}{M_{\text{Sun}}} \right)^{-2/3}, \quad (1.2)$$

where we have the orbital inclination i , the orbital period P in years, and the star mass M_{star} in solar masses. $M_{\text{planet}} \sin i$ is called minimum mass (Perryman, 2011). In Doppler spectroscopy, we are unable to obtain the real mass of the planet, but its minimum mass, which is statistically very close to the real planetary (Santos and Faria, 2018). I leave more details about this method in chapter 2.

That means, of course, that with this method is not possible to determine many other planetary parameters. For example, we have the mass but not the radius. Meaning a density for the planet cannot be estimated. That is where other methods come into play, as it allows us to combine the different information we obtain with each and better characterize an exoplanet. Next, I present the transits method, many times used in combination with the radial velocity method with this exact purpose.

Transits

We can define a transit as the passage of a celestial body directly between another body and the line of sight, but not in the correct position to create an eclipse. That, for example, can occur in the Solar system when we observe the transits of Mercury and Venus. It can also happen when a planet is in orbit of another star.

If the planet has an orbital axis that is perpendicular to the line of sight, then there will be a period in time where it will pass between the star and us as shown by figure 1.3. As the planet passes in front of the star, obscures a small percentage of it, diminishing the flux we receive from the star. in the process by an amount ΔF . Assuming spherical shapes and negligible flux from the planet, this will be related to the ratio of the radius of planet R_p and star R_s by (Deeg and Alonso, 2018)

$$\frac{\Delta F}{F} = \left(\frac{R_p}{R_s} \right)^2 = k^2, \quad (1.3)$$

where F is the total observed flux, and k represents the radius ratio.

A transit will carry valuable information about the exoplanet performing it. If transit is observed, equation 1.3 allows us to, knowing the stellar radius, estimate the radius of the planet. Two consecutive transits will give us the orbital period P of the planet. Using Kepler's laws, it is possible to estimate the semi-major axis of the orbit a .

If the stellar radius is not known, it is possible to combine the information from the total time of transit t_T , the time for which the total planet disk is in front of the stellar disk t_F , the period P , the radius ratio, and the orbital inclination. This relation was derived by Seager and Mallén-Ornelas (2003). One of the solutions allow us to write the ratio between a and R_s as

$$\frac{a}{R_s} = \frac{2P}{\pi} \frac{\Delta F^{1/4}}{(t_T^2 - t_F^2)^{1/2}} \quad (1.4)$$

and with all the required parameters, estimate the radius of the star. Similarly and if necessary, it is possible to estimate the density of the star with

$$\frac{\rho_s}{\rho_{Sun}} = \frac{32}{G\pi} P \frac{\Delta F^{3/4}}{(t_T^2 - t_F^2)^{3/2}}. \quad (1.5)$$

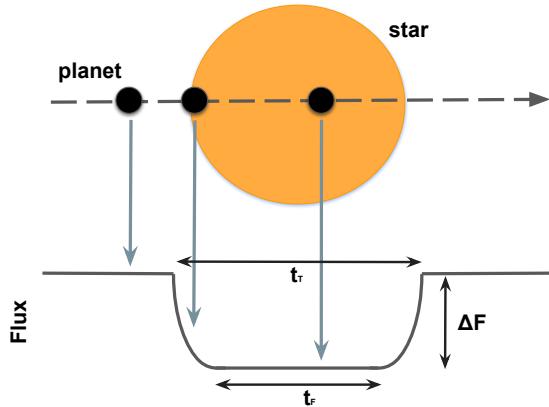


Figure 1.3: Transit of a planet in front of a star and respective flux variation ΔF by it. The transit time duration is represented as t_T and the time of totality as t_F .

The information obtained from the transit method is many times complemented with information obtained using the radial velocity method. Firstly, detecting an exoplanet using RV and transits, two independent methods, confirms the existence of such an exoplanet. Secondly, the planetary mass obtained using the RV method can be combined with the planetary radius obtained by the transit method to estimate the density of the exoplanet.

Additional, if the planetary system has more than one planet, the gravitational interactions between them will affect the time and duration of the transits. These transit timing variations (TTV) and transit duration variations (TDV) are useful to detect additional planets as small as the Earth (Holman and Murray, 2005).

Imaging

The idea of taking the picture of a planet around another star might sound like a straightforward idea. That, however, does not take into consideration many technical and physical limitations. Already in 1990 Brown and Burrows (1990) was developing the principles to search for exoplanets with this method. However, the first detection was only possible more than a decade later.

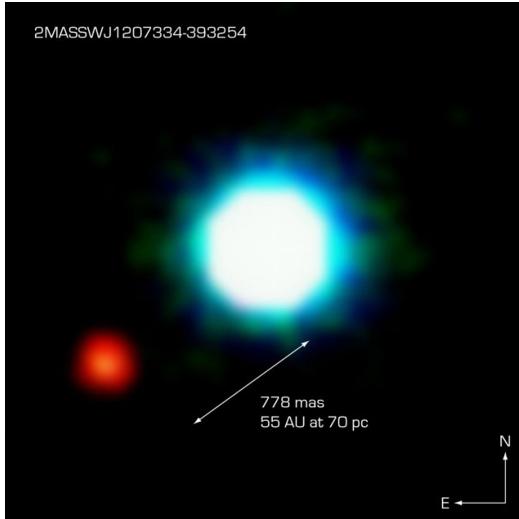


Figure 1.4: The brown dwarf 2M1207 (centre) and its planetary companion (bottom left). Source: ESO

The main limiting factor comes due to the star being many orders of magnitude brighter than a planet. Due to this, it is of no surprise that, and looking back at figure 1.1, that the planets detected through imaging are of high mass and long orbital period. For example, the first exoplanet detected with imaging is 2M1207-39 b, shown in figure 1.4 by Chauvin et al. (2004). This planet with four times the mass of Jupiter orbits a brown dwarf from a distance of around 46 AU (Ducourant et al., 2008).

One advantage of direct imaging is that is a direct detection method. That creates a unique opportunity to study the chemical and atmospheric properties of exoplanets. That is not achievable with any other method. Imaging processing techniques like the angular differential imaging (Marois et al., 2006) and the reference star differential imaging (Ruane et al., 2019), together with instruments like SPHERE (Beuzit et al., 2008) are used to explore this opportunity. With future instruments like the Nancy Grace Roman Space Telescope (Johnson et al., 2020) under development, this method will be fundamental to the study of exoplanets.

Astrometry

Astrometry is probably one of the oldest branches of astronomy and can be traced back to the time of Hipparchus, around 2200 years ago. The principles of using astrometry to detect an exoplanet are similar to the ones used on radial velocities. If a star has a planet, both will orbit a common centre of mass. As such, if we measured the star's position every day, an offset would eventually appear as it crosses the sky.

This astrometric offset, that I call α , is created due to a planet with a mass of M_p Earth-masses on a star of mass of M_s Solar-masses, at a distance of d parsecs from us. It can be measured with (Malbet and Sozzetti, 2018)

$$\alpha = 3.3 \frac{a_P}{d} \frac{M_p}{M_s}. \quad (1.6)$$

The parameter a_P is the semi-major axis of the planet's orbit. Also important to mention is that α is in units on micro arc-second (μas).

Unfortunately, only now the challenges on measuring this offset were surpassed with the ESA GAIA mission. The main difficulties arose due to the small astrometric offset expected for most cases. This is expected to be of the order of 1 μas (Santos and Faria, 2018). Even before the launch of GAIA, exoplanets detected by other methods had their discovery confirmed using astrometry. The first exoplanet discovered using only astrometry only occurred in 2010. This planet, in orbit around the star HD 176051, has a mass of around 1.5 Jupiter masses and a semi-major axis of 1.76 AU (Muterspaugh et al., 2010).

Microlensing

In my opinion, the gravitational microlensing method is one of the most fascinating methods to detect an exoplanet. This method uses an effect called the gravitational lens effect that occurs due to the fact mass bends the space around it (e.g. Sauer, 2008). To understand how the gravitational lens effect works we need to use figure 1.5. If we have two stars, the closer one to us labelled *lens star*, and the further way labelled *source star*. As the stars move in the galaxy, and under the right conditions, the source star will pass behind the lens star. As this occurs,

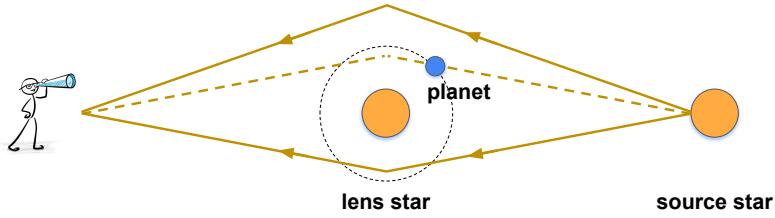


Figure 1.5: Gravitational microlensing amplified by an extrasolar planet.

the light is not blocked from us. The mass of the lens star bends the space around it. That means the light emitted from the source star, instead of being blocked by the lens star, is bent as well. If we are observing this event we will observe a change in the light we receive due to the lens effect.

Using this method to detect exoplanets is even more interesting. If the planet exists, then there will be an extra mass that will bend the light of the source star. This extra bend will occur on only one side of the lens where the planet is located, but not on the opposite side. This is represented in figure 1.5. Due to the planet, the light suffers an extra bend as it travels, instead of taking the dashed line trajectory. The physics behind this method implies that we are observing neither the planet nor the star it orbits, but the effect they cause on the light of other stars. That makes it useful for detecting low-mass planets on wide orbits. I refer to Gaudi (2010) and Gaudi (2012), for a more in-depth review of the physics of this method.

The first successful detection with this method occurred in 2003. Using the Optical Gravitational Lensing Experiment (Udalski et al., 1997), a 2.6 Jupiter-mass planet was detected on a 4.3 AU wide orbit (Bond et al., 2004). Unfortunately, this method has an important limitation. These events are a one-time event and, a second detection of the same planet using microlensing is impossible.

1.1.2 Milestones

The early 90s marked a change in the previous paradigm when technology finally allowed the first confirmed discoveries to be made. Wolszczan and Frail (1992) presented the discovery of a planetary system around the millisecond pulsar PSR1257+1. This discovery not only of one but two 66 and 98 days period planets, respectively. That was followed by the discovery of a third planet in 1994 (Wolszczan, 1994). If the existence of these three extra-solar planets is not questioned today, one might claim these discoveries are not the "true" first exoplanets due to the nature of the system host, a pulsar.

The 6th of October 1995 marked, without a doubt, the beginning of a revolution. Analysing periodic Doppler shifts in the spectra of 51 Pegasi, Mayor and Queloz (1995) confirmed the first planet around another star. This $0.47M_{Jup}$ (Jupiter masses) planet not only impacted the community for being the first exoplanet confirmed around a star. This hot Jupiter, classification given to gas giants with very short orbital periods, challenged the theories of planet formation.

At the time, and with only our Solar System as a case study, it was believed gas giants would only be formed and orbit beyond the snow-line (e.g. Pollack et al., 1996). Not only did 51 Peg b not follow these characteristics, but its existence was also later confirmed by Marcy and Butler (1995). In 1996 Lin et al. (1996) showed 51 Peg b most likely had formed by the gradual accretion of solids and capture of gas at distance around 5 AU from the star and gradually migrated inwards. The creation and migration of hot Jupiters were later coined as Type II disk migration, This occurs when the planet mass is large enough to cause density gaps along its orbit (for more see e.g. Lin and Papaloizou, 1986; D'Angelo et al., 2003). The occurrence of planet migration allowed the development of the Nice Model in 2005, which explains the evolution into their current position of the gas giants of the Solar System (Gomes et al., 2005; Morbidelli et al., 2005; Tsiganis et al., 2005).

The theories against 51 Peg b were slowly proven wrong with the discovery of other hot Jupiters (e.g. Marcy and Butler, 1996; Butler et al., 1997). The definitive proof for this class of planets came in the year 2000. Charbonneau et al. (2000), using the transit method, confirmed the existence of HD 209458 b, also detected by Henry et al. (2000) using radial velocity measurements. The detection of this

$0.69 M_{Jup}$ planet on a 3.52 days orbit using two independent methods refuted any argument against hot Jupiters. The following years saw an explosion in numbers of, not only these gas giants but of new classes of planets, see figure 1.1. The increasing number of instruments and methods available to astronomers also led to the discoveries of Neptune and Super-Earth planets (e.g. Santos et al., 2004b; Rivera et al., 20050).

1.1.3 Statistical analysis

With the number of planets discovered it is impossible not to be tempted to classify and group them according to their characteristics. The previously mentioned *Hot-Jupiters* populate the region with periods inferior to 10 days and masses above $0.1M_{Jup}$ (Udry and Santos, 2007). The discovery of Gliese 436b, a Neptune-size planet, led to the definition of *Hot-Neptunes* for Neptune-like exoplanet with short periods (Butler et al., 2004; Gillon et al., 2007).

With the same mass ranges but longer periods ($P < 100$ days) we can have cold gas giant planets, often called *Cold-Jupiters*. A famous example of these types of planets is Kepler-16b, a Saturn size and mass planet. This was the first exoplanet discovered in a circumbinary orbit around a binary star system (Doyle et al., 2011).

Another class of exoplanets we can call as *Mini-Neptunes*, for gas planets smaller than Neptune. Two examples of this type of planets are Gliese 1214b (Charbonneau et al., 2009) and Kepler-138d (Rowe et al., 2014). These planets are often also labelled as *Super-Earths*. This other class is used for rocky planets with a mass higher than the Earth, e.g., 55 Cancri e (McArthur et al., 2004), Gliese 876d (Rivera et al., 2005), and Corot-7b (Léger et al., 2009). The broad definition of these two classes, and the scientific and instrumental limitations of the discovery methods employed, can end with a planet being labelled in either class by different studies.

With a large number of exoplanets, one can also make a statistical study of their main features. These studies are possible by the surveys performed over the years such as the Californian Planet Search survey (Marcy et al., 2005), the Geneva survey of FGK stars (Mayor et al., 2011; Sousa et al., 2011) and the Kepler space telescope survey (Borucki et al., 2009). Some of the main statistical results are

listed below.

Ocurrence rate

It is believed that most Sun-like stars have planets. The quantization of this statement however is not a trivial task. One could try to calculate the average number of planets per star n . But the range of properties (size, mass, and orbital parameters) the planets display, as well as the host star properties (mass, age, and metallicity) makes it impossible to derive a function capable of it.

The most recent surveys show an important difference in the occurrence between giant and small planets (Winn, 2018). Several studies (e.g. Cumming et al., 2008; Santerne et al., 2016) agree that $\sim 10\%$ of the Sun-like stars will have a giant planet with an orbital period of a few years. Studying the long-period giants this percentage increase to $\sim 52\%$ (e.g Bryan et al., 2016), while the occurrence of Hot-Jupiter decreases to 1% or lower depending on the survey (e.g. Wright et al., 2012; Fressin et al., 2013). For smaller planets, it is believed that around 50% of all Sun-like stars have a planet with an orbit shorter than 100 *days* decreasing to around 27% for periods shorter than 50 *days* (Mayor et al., 2011).

Stellar metallicity and planets

When a star forms within a molecular cloud it is natural that the element abundance of such cloud will be reflected in the star. It is also natural to consider that such element abundance will exist in the protoplanetary disk and will influence the planets to be formed. This can be observed in the correlation between gas-giants and stellar metallicity.

It has been observed that gas-giant planets occur more frequently around metal-rich stars (e.g Santos et al., 2004a; Fischer and Valenti, 2005). With the increase number of planets discovered it was also concluded that the correlation is not observed on low-mass planet (e.g. Sousa et al., 2008). For Metal-poor stars have been suggested however to host longer period planets and less eccentric orbits (e.g. Adibekyan et al., 2013).

These discoveries, correlations, or lack thereof, are crucial for the develop-

ment of planet formation theories. For example, the relation between gas-giants and the metal-rich star seems to have origin on the primordial cloud. This cloud, being metal-rich, allow the core to form faster, reaching sufficient mass to gravitational attract gas more rapidly. I suggest the reader to Udry and Santos (2007) and Adibekyan (2019) for a more in-depth review on this issue.

Hot Neptune desert

One surprising fact about the exoplanets discovered so far is the lack of short-period Neptune-sized planets (e.g. Szabó and Kiss, 2011). This result also called *sub-Jovian desert*, is defined as the lack of exoplanets with masses around $0.1M_{Jup}$ that was initially expected to be easy to find in short-period orbits of $P < 4$ days (West et al., 2019). This anomaly, shown in figure 1.6, is believed to be associated with a combination of photo-evaporation of the atmospheres and tidal disruptions due to gas giants migrations (West et al., 2019). The existence of this desert also seems to indicate that Super-Earths and Hot-Jupiters have distinct formation processes (Owen and Lai, 2018).

Exoplanet radius valley

The exoplanet radius valley is a gap in the radius of small planets observed in the California Kepler Survey (Petigura et al., 2017; Fulton et al., 2017). This gap (see figure 1.6) known for planets with radius around the 1.8 Earth radius had already been predicted by numerical models due to photoevaporation (e.g Lopez and Fortney, 2013; Owen and Wu, 2013). Another explanation is that the luminosity of the cooling rocky core erodes the envelopes, only preserving heavy ones. Such a mechanism would produce a lack of intermediate-sized planets (Ginzburg et al., 2018).

While the gap was first identified for FGK stars, more recent works also included M stars. On it, the results indicate that this gap occurs at a lower planets' equilibrium temperature than expected makes M dwarfs stars more effective at reducing the size of their planets (Hirano et al., 2018). That seems to confirm that the driving process at creating this gap is photoevaporation.

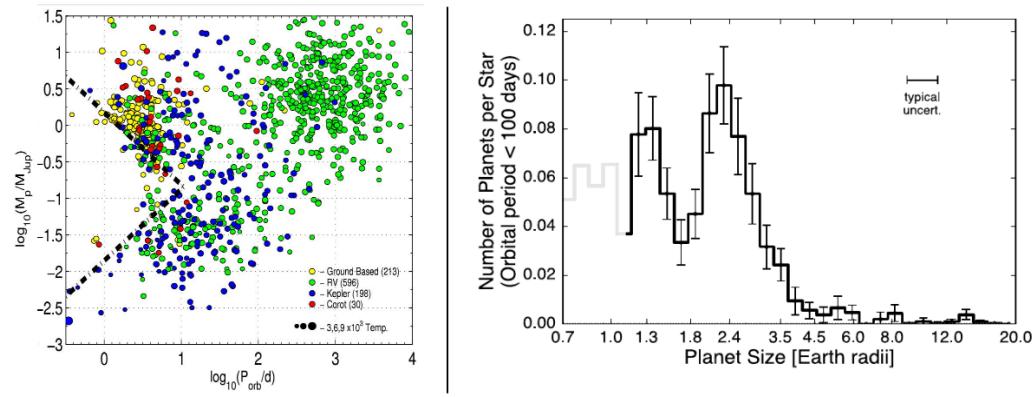


Figure 1.6: Left: Hot-Neptune desert derived by Mazeh et al. (2016) with the dash-dotted lines delimiting the boundaries of this desert. Right: Radius gap from Fulton et al. (2017) where the gap around 1.8 Earth radii creates a bi-modal distribution of planets around Sun-like stars.

1.2 This thesis

Now more than ever is necessary powerful methods to improve our planet detection capabilities. The most recent instruments are capable of reaching a precision never before seen. Meaning every day we are closer to detect a second Earth.

My work on this thesis will focus on using radial velocity data. Combining the flexibility of Gaussian processes with the structure of a neural network, I will adopt a framework known as a Gaussian process regression network to improve the detection limits using radial velocity observation.

This thesis is organized into four more chapters. Next, I will briefly explore the important aspects of the radial velocity method. Starting with the mathematical background that supports this method. I will follow by seeing how to obtain these observations and the obstacles we need to tackle when analysing radial velocity data.

In the third chapter, I will present the Gaussian process regression network. I will start with a basic introduction to Gaussian processes, explore the covariance functions of interest when performing a Gaussian process regression. The chapter ends with an explanation of the mathematics behind a Gaussian process regression network and how we intend to use it on radial velocity data.

The fourth chapter is dedicated to exploring a fully functioning Gaussian process regression network in radial velocity data. I will focus on exploring its capabilities on measurements obtained from the Sun. I will also present some of the results obtained on four stars from the EXPRESS Stellar Signal Project. The fifth and last chapter is left for the conclusions we achieved with our new framework, its pros and cons, and the future possibilities and improvements still required.

Chapter 2

Doppler spectroscopy

In this chapter, I describe the key aspects behind the measurement of precise radial velocities. I start by presenting the mathematical/physical principles behind Doppler spectroscopy. I follow this with a brief discussion of the radial velocity calculation techniques. I conclude with the challenges stellar activity create on these observations and the current mitigation strategies currently employed.

2.1 Mathematical background

As seen in the previous chapter, a planet around a star causes both to orbit a common centre of mass. If we observe the spectrum of said star, we will see that its lines shift with the star radial motion. These shifts in wavelength or relative radial velocity between the star and the observer can be measured using the relativistic Doppler shift z (Sher, 1968)

$$z \equiv \frac{\lambda_{\text{observed}} - \lambda_{\text{rest}}}{\lambda_{\text{rest}}} = \frac{\gamma}{1 + \frac{v_r}{c}} - 1, \quad (2.1)$$

where λ is either the line wavelength observed or the one measured in a rest frame. γ is a relativistic factor, c is the speed of light in the vacuum, and v_r is the radial velocity. Considering that for our problem $\gamma \approx 1$ and $c \gg v_r$ we can simplify

equation 2.1 and obtain

$$v_r = c \frac{\lambda_{\text{observed}} - \lambda_{\text{rest}}}{\lambda_{\text{rest}}}. \quad (2.2)$$

To understand how we use this measure to determine if a set of RV measurements includes one or more planets, we need to understand how we reach the keplerian function from equation 1.1 and understand its main parameters. For it I will use the works of Murray and Correia (2010); Perryman (2011), and Wright and Gaudi (2013) to derive all the necessary components.

2.1.1 Keplerian orbits

The two-body problem allows us to understand the motion of a planet and a star around its barycenter. If we considered Newton's law of universal gravitation, a body attracts another body with a force directly proportional to the product of their masses m_1 and m_2 , and inversely proportional to the square of their distances r (Newton and Halley, 1687). Mathematically this force F_g is

$$F_g = G \frac{m_1 m_2}{r^2}, \quad (2.3)$$

m_1 and m_2 , with $m_2 > m_1$, is the mass of the two bodies, r the distance separating them, and G is the universal gravitational constant valued at $6.67430 \times 10^{-11} m^3 kg^{-1} s^{-2}$ (NIST, 2019). To simplify the notation, and unless specified otherwise, I will use m_1 and m_2 to name the respective bodies.

Now if we consider the two bodies m_1 and m_2 (or a planet and a star), with position vectors \mathbf{r}_1 and \mathbf{r}_2 in relation to an origin O (figure 2.1). The relative motion of m_1 with respect to m_2 is then $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$. This allow us to write the vector representation¹ of the gravitational force each body apply as

$$\begin{aligned} \mathbf{F}_1 &= -G \frac{m_1 m_2}{r^3} \mathbf{r}, \\ \mathbf{F}_2 &= +G \frac{m_2 m_1}{r^3} \mathbf{r}. \end{aligned} \quad (2.4)$$

¹ The vector representation of Newton's law of universal gravitation is $\mathbf{F} = G \frac{m_1 m_2}{r^2} \hat{\mathbf{r}}$, where $\hat{\mathbf{r}}$ is a unit vector with the direction of the force.

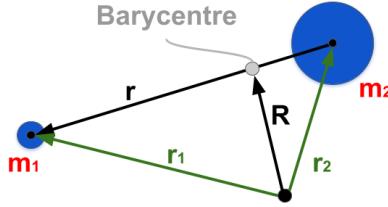


Figure 2.1: The two-body problem scheme.

Now considering Newton's second law of motion² we can re-write it as

$$\begin{aligned} m_1 \ddot{\mathbf{r}}_1 &= -G \frac{m_1 m_2}{r^3} \mathbf{r}, \\ m_2 \ddot{\mathbf{r}}_2 &= +G \frac{m_2 m_1}{r^3} \mathbf{r}. \end{aligned} \quad (2.5)$$

Since we now have $\ddot{\mathbf{r}}_1$ and $\ddot{\mathbf{r}}_2$ lets not forget that $\ddot{\mathbf{r}} = \ddot{\mathbf{r}}_1 - \ddot{\mathbf{r}}_2$. This allow us to combine the terms from 2.5 into

$$\ddot{\mathbf{r}} = -\frac{G}{r^3} (m_1 + m_2) \mathbf{r}. \quad (2.6)$$

This being the equation we need to solve to obtain the formula for a closed elliptical orbit.

The following step requires the usage of the cross product³. For an orbiting body, the position and acceleration are collinear vectors while the velocity is perpendicular to both. This means

$$\begin{aligned} \ddot{\mathbf{r}} \times \mathbf{r} &= 0 \\ \ddot{\mathbf{r}} \times \dot{\mathbf{r}} &= \text{constant}_1 \\ \mathbf{r} \times \dot{\mathbf{r}} &= \text{constant}_2 = h \end{aligned} \quad (2.7)$$

To solve equation 2.6 we now need to change our coordinate system into the polar coordinate system (r, θ) . The origin being centered on the body m_2 , and an arbitrary reference line defined for $\theta = 0$.

² Newton's second law of motion relates the force applied to a body to its change of momentum. It is defined as $\mathbf{F} = \frac{d\mathbf{p}}{dt} = m\mathbf{a}$. Where \mathbf{F} is the force, \mathbf{p} the momentum, m the mass, and \mathbf{a} the acceleration of the body.

³ The cross product is defined $\mathbf{v}_1 \times \mathbf{v}_2 = ||\mathbf{v}_1|| ||\mathbf{v}_2|| \sin \theta \mathbf{n}$, where θ is the angle between the vectors \mathbf{v}_1 and \mathbf{v}_2 , and \mathbf{n} is a unit vector perpendicular to the plane created by \mathbf{v}_1 and \mathbf{v}_2 .

Considering the equations of motion in polar coordinates⁴, and with the appropriate substitutions, we first obtain

$$h = r^2\dot{\theta}, \quad (2.8)$$

and manage to transform equation 2.6 into a second-order, linear differential equation. This equation has a general solution given by

$$\frac{1}{r} = \frac{G(m_1 + m_2)}{h^2} [1 + e \cos(\theta - \omega)], \quad (2.9)$$

where e and ω are two constants of integration. In reality, this equation is the general equation for a conic section where e is the eccentricity.

Following some more algebraic manipulations, we reach the Keplerian orbit for an ellipse in polar coordinates

$$r = \frac{a(1 - e^2)}{1 + e \cos(\theta - \omega)} = \frac{a(1 - e^2)}{1 + e \cos \nu}, \quad (2.10)$$

where a is the semi-major axis of the ellipse, e its eccentricity, and ν a quantity called a true anomaly. This solution is nothing else but the Newtonian derivation of Kepler's second law of planetary motion Kepler's first law of planetary motion.

2.1.2 Orbital elements

On equation 2.10, I defined the quantity $\nu = \theta - \omega$ but did not explained its meaning. That and other quantities of interest are shown in figure 2.2 and arise from the various angles in the orbit plane.

Inclination

The orbit inclination, denoted as i , is only the angle between the orbit plane of the planet and the reference plane. It varies between 0 and π .

⁴ These are $\mathbf{r} = r\hat{\mathbf{r}}$, $\dot{\mathbf{r}} = \dot{r}\hat{\mathbf{r}} + r\dot{\theta}\hat{\theta}$, and $\ddot{\mathbf{r}} = (\ddot{r} - r\dot{\theta}^2)\hat{\mathbf{r}} + [\frac{1}{r}\frac{d}{dt}(r^2\dot{\theta})]\hat{\theta}$

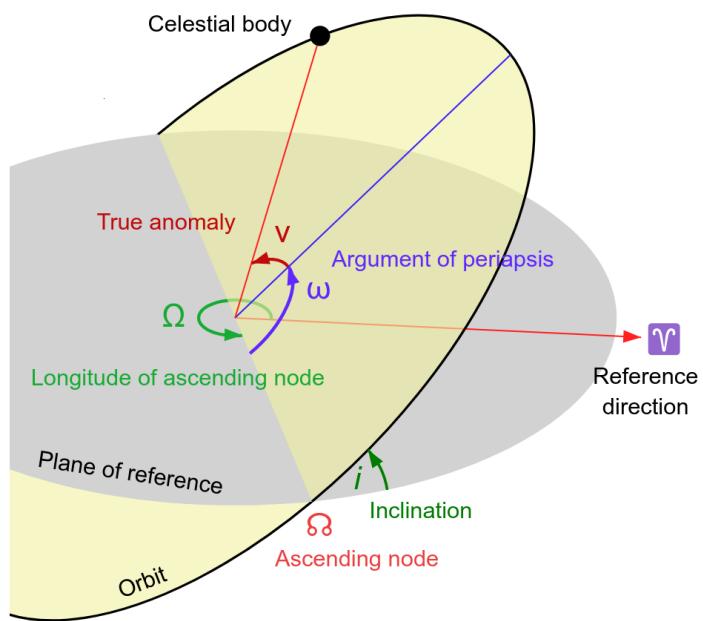


Figure 2.2: Elliptical orbit of a celestial body in three dimensions. Source: Wikimedia Commons.

Ascending node

The ascending node \varnothing is known as the point where the planet traverses the reference plane. We consider it when occurring from its southern side to the northern side.

Argument of periapsis

The argument of periapsis is also designated periastron for orbits around stars. Usually represented by an ω , it is the angle between the ascending node and the periapse of the orbit. When calculated, it is done considering the direction of motion and varies between 0 and 2π .

Longitude of ascending node

The longitude of the ascending node Ω , estimated in the reference plane, is the angle where the object moves away from the observer into the direction of the ascending node.

True anomaly

The previously mentioned true anomaly v , or seldom expressed as $v(t)$ as it is dependent on the time, is the angle between the direction of the periapsis and the position of the planet at time t .

Eccentric anomaly

The eccentric anomaly $E(t)$ is an angle obtained from the circumcircle of ellipse⁵ as the planet moves along its elliptic orbit. The true and eccentric anomalies are related by

$$\cos v(t) = \frac{\cos E(t) - e}{1 - e \cos E(t)}, \quad (2.11)$$

⁵ The circumcircle of an ellipse is an auxiliary circle with centre on the centre of the ellipse and radius equal to the semi-major axis.

that after using some trigonometric properties give us

$$\tan \frac{\nu(t)}{2} = \left(\frac{1+e}{1-e} \right)^{1/2} \tan \frac{E(t)}{2}. \quad (2.12)$$

Mean anomaly

We use the mean anomaly $M(t)$ to determine the true anomaly. Over an orbit, a planet does not move at a constant angular rate. Even so, we can calculate an average. That allows us to calculate the mean anomaly at time $t t_p$, with t_p being the moment after the passage of the periapsis and define the mean anomaly as

$$M(t) = \frac{2\pi}{P} (t - t_p), \quad (2.13)$$

with P the period of the orbit. Using orbital mechanics we can relate the mean and eccentric anomalies with

$$M(t) = E(t) - e \sin E(t). \quad (2.14)$$

2.1.3 Calculating the Keplerian function

Having defined the main orbital elements, we can finally determine the Keplerian function to model a planetary signal. Let us start by remembering Kepler's third law⁶ writing it as

$$P^2 \propto a^3, \quad (2.15)$$

where the proportionality constant to be inserted will take slightly different values accordingly to the type of orbit. If, for example, we consider the orbit of the body m_2 around the system barycenter, its period and semi-major axis relate as

$$P_{m_2}^2 = \frac{4\pi^2}{GM} a_{m_2}^3 \quad \text{with} \quad M = \frac{m_1^3}{(m_1 + m_2)^2}. \quad (2.16)$$

⁶ Kepler's third law states that the squares of the orbital period of a planet are proportional to the cube of its semi-major axis

Omitting some trigonometry manipulations, the position of the body m_2 among the line-of-sight $z(t)$, perpendicular to the plane of reference of figure 2.2, is

$$z(t) = r(t) \sin i \sin(\omega + \nu(t)), \quad (2.17)$$

with $r(t)$ being the distance to the barycenter. To obtain the velocity is then necessary to derive the expression concerning t , giving us

$$\dot{z}(t) = \sin i [\dot{r}(t) \sin(\omega + \nu(t)) + r(t) \dot{\nu}(t) \cos(\omega + \nu(t))]. \quad (2.18)$$

This velocity $\dot{z}(t)$ in no more than the projected motion along the line-of-sight, thus being a radial velocity ($v_r \equiv \dot{z}$). Even so, we still need some more steps to obtain a similar expression to equation 1.1. For it, we make some algebraic manipulations with $r(t)$ and $\dot{r}(t)$ to reach

$$v_r(t) = K[\cos(\omega + \nu(t)) + e \cos(\omega)] + \gamma, \quad (2.19)$$

with

$$K = \frac{2\pi}{P_{m_2}} \frac{a_{m_2} \sin i}{\sqrt{1 - e^2}}. \quad (2.20)$$

Using Kepler's third law from equation 2.16 we can re-write the semi-amplitude as

$$K = \left(G \frac{2\pi}{P} \right)^{1/3} \frac{m_1 \sin i}{(m_1 + m_2)^{2/3}} \frac{1}{(1 - e^2)^{1/2}}. \quad (2.21)$$

That allows us to obtain two other equations of interest. First, for the minimum mass where we have

$$m_1 \sin i = 4.919 \times 10^{-3} K \left(1 - e^2 \right)^{1/2} P^{1/3} (m_1 + m_2)^{2/3}. \quad (2.22)$$

On it $m_1 \sin i$ is given in Jupiter masses, K in meters per second, P in days, and $m_1 + m_2$ in Solar masses.

The second equation is identical to expression 1.2. For it, we need to consider circular orbits and the masses $m_1 \ll m_2$ to reach

$$K = 28.4 \frac{m_1 \sin i}{P^{1/3} m_2^{2/3}}. \quad (2.23)$$

In this equation we have K in meters per second, P in years, $m_1 \sin i$ in Jupiter masses, and m_2 in Solar masses⁷.

We now have all the necessary equations to characterize a set of RV points. That, however, is not the end. Equation 2.19 allow only a single planetary signal. For the multiple-planet fitting, we can do a linear sum for an n number of planets

$$v_r(t) = \sum_{i=1}^n K_i [\cos(\omega_i + \nu_i(t)) + e_i \cos(\omega_i)] + \gamma. \quad (2.24)$$

Another common issue with RV observations arises when we have multiple instruments observations. These observations will most likely include an instrument-dependent RV offset. That will require the use of different offsets. One, for different sections of the measurements, accordingly to the instruments.

Also quite common in RV analysis is to include a linear trend parameter d . Done to fit any long term trends that might appear in our data not necessarily related to planetary signals. That means using a Keplerian function of the type

$$v_r(t) = \sum_{i=1}^n K_i [\cos(\omega_i + \nu_i(t)) + e_i \cos(\omega_i)] + \gamma + d(t - t_0), \quad (2.25)$$

where d is the slope of the long term signal, and t_0 is the average of t .

The last major setback to fully characterize any planetary signal appears on the relation between the true, eccentric, and mean anomalies. To obtain the value of true anomaly is necessary to use its relation to the eccentric anomaly. That then requires the use of the mean anomaly to calculate the eccentric anomaly. The setback on these steps is on equation 2.14 being transcendental⁸. To calculate, we require the use of iterative methods, like the Newton-Raphson method. We start by calculating an initial value for $M(t)$ and $E(t)$ with

$$\begin{aligned} E_0 &= M + e \sin M + \frac{e^2}{2} \sin 2M, \\ M_0 &= E_0 - e \sin E_0, \end{aligned} \quad (2.26)$$

⁷ Usually equations 2.22 and 2.23 are written with the units in the denominator. That is helpful to avoid using the wrong units but, I dislike such notation and removed it in this chapter.

⁸A transcendental function is a function that can not be express by a finite sequence of the algebraic operations

calculating M using the time of the periapsis passage. We then calculate any new term iteratively with

$$\begin{aligned} E_1 &= E_0 + \frac{M - M_0}{1 - e \cos E_0}, \\ M_1 &= E_1 - e \sin E_1. \end{aligned} \quad (2.27)$$

Convergence, within a few iterations, should be achieved. Following it, we are ready to calculate the true anomaly, and with it the Keplerian function.

2.2 Spectroscopic observations

In the previous section, we saw how to mathematics being a radial velocity measurement. That, however, is only possible after having observed the spectra of a given star and appropriately processed the data. That implies transforming the spectra observed into a single measurement. Of course, this is not an easy task. It all starts in the spectrograph that, in simple terms, will receive the light from the telescope, disperse it to create a spectrum, and then record this spectrum on a detector.

2.2.1 The past, present and future of spectroscopy

When Mayor and Queloz (1995) discovered 51 Pegasi b, spectroscopic observations were done using the ELODIE spectrograph (Baranne, A. et al., 1996). This instrument entered into operation in 1993 and was capable of an RV precision of 7 m/s, later 3 m/s after an upgrade on 2014 (Pepe et al., 2018). Over the years the developed spectrographs pushed the limits of the precision achieved. For example, HARPS (Pepe et al., 2002) and HARPS-N (Cosentino et al., 2012), both spectrographs with an ever-increasing number of exoplanets discovered when began operations managed to achieve a precision of 1 m/s.

That, however, is not enough when searching for a second Earth. For it the "barrier" of detecting a signal of 9 cm/s needed to be broken. Spectrographs like ESPRESSO (Pepe et al., 2014) and EXPRES (Jurgenson et al., 2016), with a precision of 10 cm/s, are the most recent bet to achieve this objective. Although the higher precision brings several challenges, both technical and scientific, the future objective is to be capable of detecting signals of around 2 cm/s with the instruments like CODEX Pasquini et al. (2010).

Developing and building all these instruments, however, is only the first step. With it, we aim at obtaining a spectrum with as many lines as possible. If successful, we get something similar to what appears in figure 2.3. In this figure, we see many dark absorption spectral lines. Later we use it to obtain a radial velocity measure. A high number of lines, however, is not the only requirement for high precision measurements. Radial velocity measurements are also affected

by several potential error sources. Next, I will mention two sources of this noise.

2.2.2 Statistical and systematic errors

Let consider, for example, an Earth-like planet orbiting a sun-like star at $1AU$. The radial velocity expected is approximately $0.1m/s$. If we look at equation 2.2, and use lines on the visible part of the spectrum, we can calculate the wavelength shift $\Delta\lambda$ required to detect such planet. Under some approximations⁹ this implies a shift of

$$\Delta\lambda = 5 \times 10^{-7} \frac{0.1}{3 \times 10^8} \approx 2 \times 10^{-16} m. \quad (2.28)$$

We enter into a regime where we need to address the statistical and systematic errors when measuring such shifts. For example, the spectrograph needs to be stabilized and precisely calibrated in wavelength to be photon-efficient and precise on a sub-metre-per-second level (Pepe et al., 2014). For more on instrument stability errors and other instrumental factors we needed to take into account, I recommend, for example, Pepe and Lovis (2008).

Another issue arises from the shape of spectral lines. More specifically, we need to address the errors due to photon noise. Bouchy et al. (2001) determined how much this noise contribute to the final RV error when assuming a Gaussian-shaped line. Using the expression from Figueira (2018) we can calculate with

$$\sigma_{RV} = \frac{(\pi \log 2)^{-1/4}}{2} \frac{\sqrt{FWHM}}{SNR} \frac{\sqrt{PXLSC}}{C} F(C_{eff}), \quad (2.29)$$

where C is the contrast of the Gaussian line, SNR the signal-to-noise ratio of the spectrum, $PXLSC$ the dimension of the pixel as measured in velocity, $F(C_{eff})$ is a polynomial function of the effective contrast. On the previously mentioned HARPS, for example, this noise can have levels of $0.5 m/s$ with appropriate exposure times (Haywood, 2016b). I will not further explore the photon noise error. Instead I refer to Butler et al. (1996) and Bouchy et al. (2001), for more information on this issue.

⁹ The visible spectrum have a wavelength range from 3.8×10^{-7} to 7.5×10^{-7} meters, to simplify the calculations, I used the value of 5×10^{-7} meters. Another approximation done was for the speed of light in vacuum (c), I consider it as roughly $3 \times 10^8 m/s$.

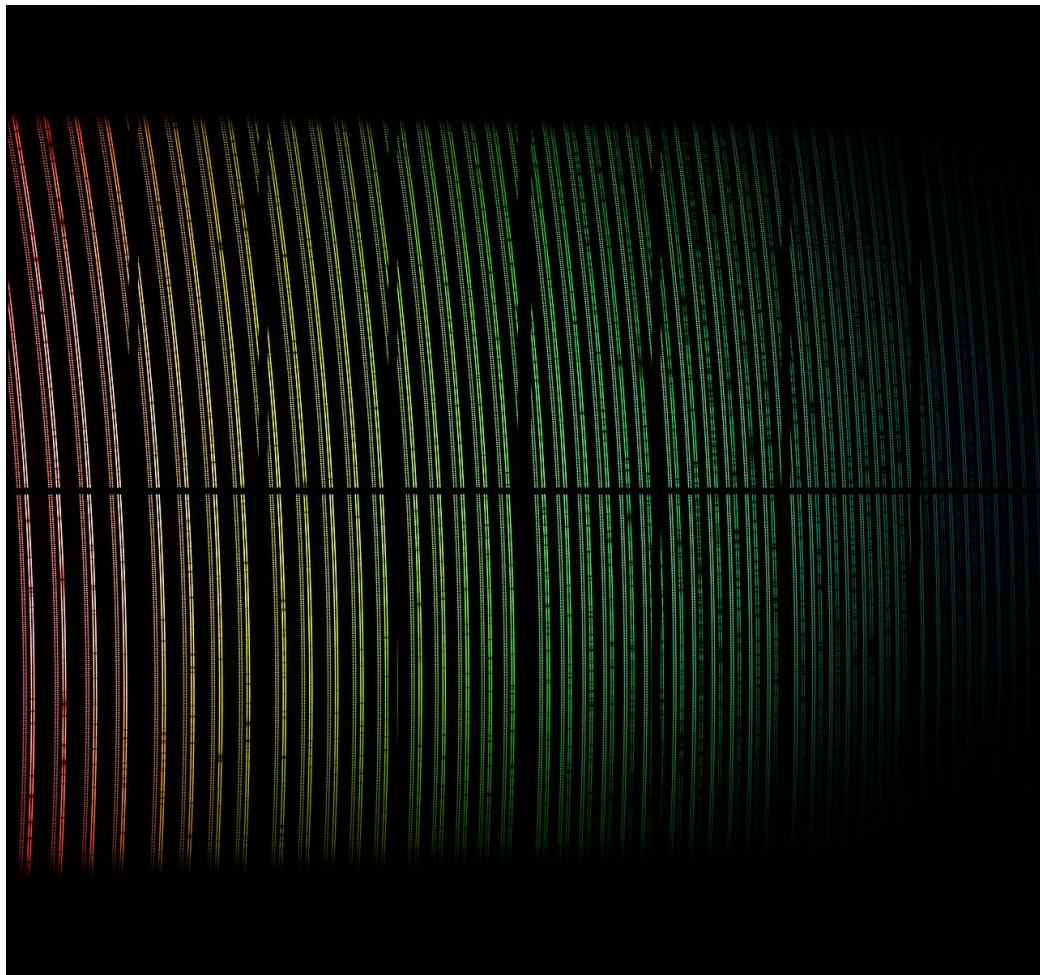


Figure 2.3: ESPRESSO's first light spectrum. Source: ESO/ESPRESSO team.

2.2.3 The cross-correlation function

As mentioned at the start of this chapter, we use absorption spectral lines to calculate the radial velocity. If we identified the lines and compare them with the theoretical value from the Solar System rest frame we can measure its radial velocity. This being represented as $\lambda_{\text{observed}} - \lambda_{\text{rest}}$ in equation 2.2.

One of the objectives when gathering RV measurements is to increase the precision as much as possible. Selecting a single line and calculate its RV shift will not guarantee us that. Instead, we use the mean spectral line obtained from an average of over thousands of spectral lines. To accomplish that, we use the cross-correlation method Queloz (1995).

Let us start by a simple definition for cross-correlation from Allende Prieto (2007). Let us consider two arrays of size k . One is from the stellar spectrum (S), and another from a template or mask (M). Cross-correlating them, we obtain the relative velocities between them. That is done using, for example, equation

$$CCF_i = \sum_k M_k S_{k+1}. \quad (2.30)$$

If the spectrum S and M are identical but shifted by an integer amount of p pixels, the maximum value on C will be the element where $i = p$.

In our case, we will use a mean spectral line and cross-correlate it with a line mask to determine the wavelength shift. It will thus be of interest to use a weighted cross-correlation. That is calculated using (Pepe et al., 2002)

$$CCF(v_r) = \int S(\lambda) M(\lambda_{v_r}) d\lambda, \quad (2.31)$$

where $\lambda_{v_r} = \sqrt{\frac{c-v_r}{c+v_r}}$, In this equation $S(\lambda)$ is the measured spectrum, and $M(\lambda_{v_r})$ the Doppler-shifted numerical mask.

The expression becomes simpler when representing M as the sum of masks M_i , for each stellar absorption line i . Considering the different relative depths of each spectral line used, we define a respective weight w . That allows us to reach the expression

$$CCF_{\text{weighted}}(v_r) = \sum_i S(\lambda_i) M_i(\lambda_{v_r}) w_i. \quad (2.32)$$

Fitting the CCF with a Gaussian function allows us to determine the radial velocity (see figure 2.4). This measure, obtained considering the location of the peak of the CCF, will be shifted from zero by the star's radial velocity. This however is not the only measure of interest we can obtain from the CCF. Other physical properties are also present.

For example, the stellar rotation velocity $v \sin i$, can be inferred from the CCF width (e.g Delfosse et al., 1998; Browning et al., 2010; Reiners et al., 2012). Another useful measure is estimated from the equivalent width. Knowing beforehand the effective temperature of the star and using the CCF equivalent width, one can estimate the star's metalicity (e.g. Mayor, 1980; Santos et al., 2002).

Besides theses physical properties, with the shape of the CFF being close to that of a Gaussian, characteristics measured on a Gaussian curve can be found on the CCF. Properties like the equivalent width and full width at half maximum (FWHM) are relative easy to obtain using the mathematics behind a Gaussian curve. I present some of these properties next.

Full width at half maximum

We can define the FWHM as the span within two points on the Gaussian curve, at which its function reaches half of its maximum value. Knowing that the probability density function of a Gaussian distribution is (Patel and Read, 1996)

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right), \quad (2.33)$$

where we have a mean μ , standard deviation σ , and variance σ^2 , we can calculate the FWHM with

$$FWHM = 2\sigma\sqrt{2\log 2}. \quad (2.34)$$

This measure will be sensitive to the variations in the CCF caused by activity on the surface of a star. With it, we can correlate any variation in the FWHM with the RV variations caused by the presence of starspots (Haywood, 2016a). Soon I will explore more about the influence of stellar activity on the FWHM.

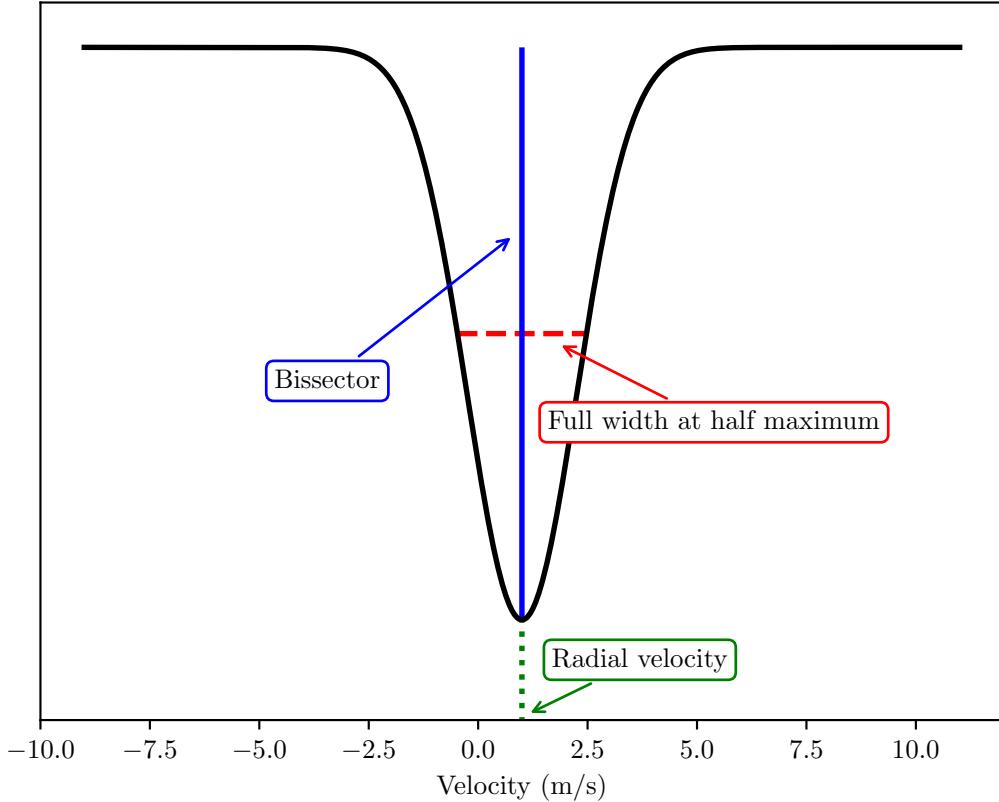


Figure 2.4: Simulated structure of a CCF.

Bisector

A property not exclusive to only Gaussian curves is the bisection. A bisection is the division of any geometric figure into two others with the same shape and size. The line dividing the two new figure is called a bisector. Shown by the blue vertical line of figure 2.4, this property will be useful later for the definition of some activity indicators.

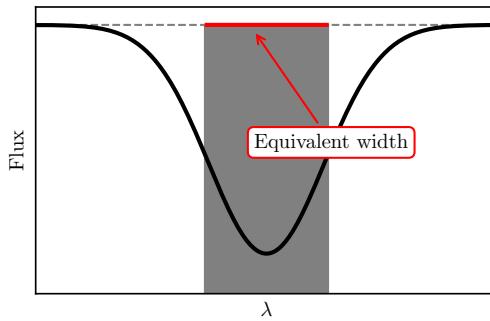


Figure 2.5: Example of the equivalent width of a spectral line.

Equivalent width

As mentioned earlier, we can use the the equivalent width (EW) to estimate the metalicity. In a spectral line, this measure, is obtained from the rectangle with a height equal to that of continuumn, and a width that creates a rectangle with area equal to the area in the spectral line (see figure 2.5).

On a CCF we can calculated it with (Borgniet et al., 2019)

$$EW = \int \left(1 - \frac{CCF(v_r)}{\max(CCF(\Delta))} \right) dv_r. \quad (2.35)$$

On it, ($CCF(\Delta)$) is the the area centred on the CCF main peak and below the CCF lower wing. On a normalized CCF, the EW will thus be a rectangle with a height equal to 1 and surface equal to the area given by $CCF(\Delta)$.

TODO: check the definition of EW better

2.3 The Stellar activity problem

Having seen how we obtained RV measurements from a star. All measures obtained still require careful analysis. One of the main challenges in exoplanets detection is due to the activity from the star contaminating the observation. These stellar signals come from physical processes that induce radial velocity variations on the stars, besides that of a planet, of course. This is known to be able to hide or mimic planetary signals in RV measurements (e.g. Santos et al., 2014; Robertson et al., 2015).

As previously mentioned, a planet induces a Doppler shift on the star spectral lines, which will be measured by a shift of the CCF peak (Figueira, 2018). This implies that any physical process affecting the shape of the spectral lines will have repercussions on the shift measured from the CCF. Already in 1997 Saar and Donahue (1997) showed this would occur when inhomogeneities on a star surface was present. These would affect the shape of the spectral lines, thus creating Doppler shifts that could be mistaken by planetary signals. With spectrographs aiming to achieve a precision of 10 cm/s, the expected signal for an Earth-like planet, is fundamental to correctly deal with any signal capable of originating a false detection.

2.3.1 Stellar activity features

The stellar activity has a wide range of sources and timescales ranging from a few minutes to dozens of years. That will of course affect short and long period planets differently. Understanding each individual feature is important to implement the necessary strategies required to be deal with each (Dumusque, X. et al., 2011a; Dumusque, X. et al., 2011b).

On timescales of minutes to hours

- *Oscillations*

Pressure or acoustic waves (p-modes for short), propagate at the surface of stars, causing the exterior envelopes to expand and contract. Usually, on G,

K, and M class stars reach timescales less than 15 minutes. In the case of the Sun, these are known to last from 5 to 15 minutes (Schrijver and Zwaan, 2000; Arentoft et al., 2008; Broomhall et al., 2011).

These oscillations originate when the plasma flows originated from the stellar convective envelope create acoustic waves. These combine, either constructively or destructively, to create stationary waves, whose periodicity of the effect last around five minutes (Cegla, 2019). Individually these waves only create an RV signal on the order of the centimetre per second, but when dozens of modes create a combined signal capable of reaching a few meters per second (Bedding and Kjeldsen, 2003; Bedding and Kjeldsen, 2007).

The induced signals of oscillation can be dealt with when spectroscopic observations are planned. Observations with 10 to 15 minutes exposures are capable of suppressing these signals on solar-like stars, (Dumusque, X. et al., 2011a; Dumusque, X. et al., 2011b). On other stars Chaplin et al. (2019) showed these exposures range from 4 minutes for low mass stars to 100 minutes for hotter higher mass stars.

With the increased precision of spectrographs, another type of oscillation is now of concern. Recently studies showed the existence of Rossby waves, known as r-modes, in the Sun (Löptien et al., 2018). These originated from toroidal modes on a rotating star, can produce RV signals of a few dozens of centimetres per second, and periods between 10 to 20 days (Lanza et al., 2019). As a recent discovery, the r-modes is a future activity source to consider.

- *Flares*

Flares are a sudden increase on the brightness of a star near its surface. In the case of the Sun, often associated with coronal mass injections. These sudden and random events generate spikes on a RV observation than can range from 0.5 to 10 m/s (Saar, 2009; Reiners, 2009). These, however, being rare are easily identifiable on, for example, the H α emission profile and are easily chunked out.

- *Granulation*

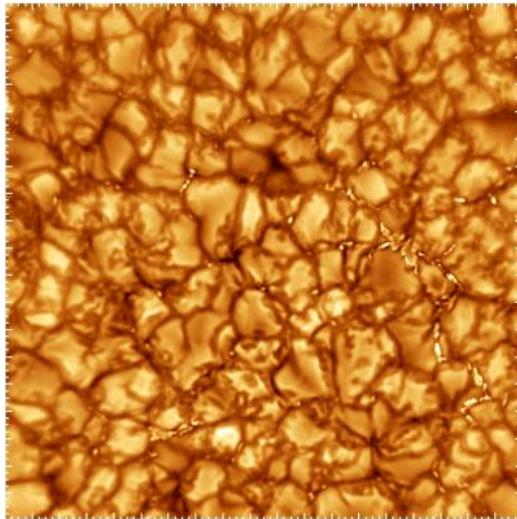


Figure 2.6: Solar granulation over an area of 19000 squared kilometers. Source: Big Bear Solar Observatory/New Jersey Institute of Technology.

The next source of noise on a spectroscopic observation I present is due to a phenomenon known as granulation (see figure 2.6). These phenomena occur when the convection in the external layers of solar-type stars, transport bright and hot material from deeper layers, flowing horizontally, cooling down and descending through the darker boundary regions known as intergranular lanes (e.g. Rieutord and Rincon, 2010; Rutten and Severino, 2012). This upward and downward movement of material does not occur equally. With the granules covering more surface area than the intergranular regions, the absorption lines of the spectrum end with an overall net blueshift (Cegla, 2019).

The average lifetime for a granule on the Sun, to reach up to 25 minutes (Bahng and Schwarzschild, 1961). The evolution of these granules and their ratio with the intergranular lines generates a constantly changing RV signal that can reach the meter per second on the Sun (Cegla, 2019). The individual granules also group together, creating supergranules where plasma flows at a much slower rate, generating RV signals capable of reaching lifetimes of 1.5 days (Del Moro, 2004; Rieutord and Rincon, 2010).

Recently, supergranulation phenomena was shown to pose a bigger than expected challenge on Earth-like planets detection (Meunier, N. and Lagrange, A.-M., 2019; Meunier and Lagrange, 2020).

To eliminate some effects of this phenomena is recommended to use the appropriate observational strategy. For example, Dumusque, X. et al. (2011a) concluded that three observations per night, separated by 1 to 2 hours, reduce its effects on RV observations. That work was also determined that, while the Doppler signal is of 1 m/s on the Sun, it can reach values of 30 m/s on stars of other spectral classes.

On timescales of days to months

- *Spots*

Spots are darker regions on the surface of the stars (see figure 2.7). These originate when the magnetic field of the star interacts with the plasma. It inhibits the outgoing convective heat flux of a region and reduces its brightness and temperature (Solanki, 2003). First identified and studied on the Sun, where are known as sunspots, are closely linked to the solar magnetic cycle (Hathaway, 2015). Records of its number have been systematically gathered for the last centuries. Its properties being extrapolated to other stars, where are known as starspots. For a more in-depth review on sunspots and their connection to the solar magnetic cycle, I recommend the reader to both Solanki (2003) and Hathaway (2015).

Since a spot emits less light than the surrounding surface, it will break the balance between the blue-shifted and red-shifted parts of a star as it rotates originating a Doppler shift on its spectrum deforming the CCF shape (Saar and Donahue, 1997; Queloz et al., 2001; Dumusque, X. et al., 2011b). This can originate signals from 0.2 to 5 m/s on solar age stars with periodicities close to the stellar rotation (Lagrange, A.-M. et al., 2010; Meunier, N. et al., 2010). Younger, more active, and rapidly rotating stars shows that its activity signal is mostly dominated by spots (e.g Reinhold et al., 2019).

To deal with these signals, there are several approaches. The first uses activity indicators, where the activity but not the planetary signals will be

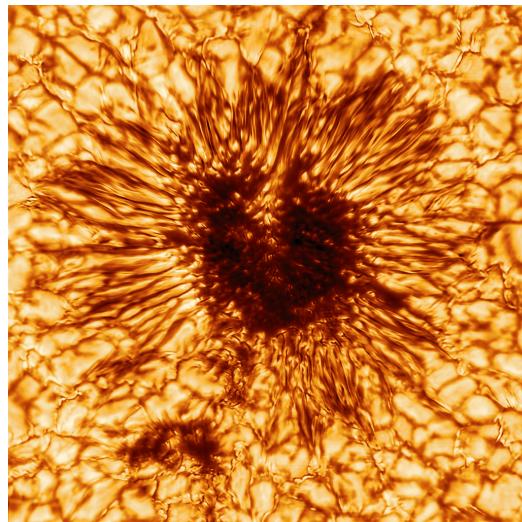


Figure 2.7: A Sunspot capture by the Inouye Solar Telescope on January 28, 2020.
Source: NSO/NSF/AURA.

present. I will discuss more of these soon. One can treat the spots signals as correlated noise and fit it with, for example, Gaussian processes (e.g Haywood et al., 2014; Faria et al., 2016). Another approach consists in using numerical simulations of the active regions to create synthetic RV measurements that are the compare it to the observed RVs, see (Oshagh, 2018) for a bit more about this.

- *Faculae and plages*

Faculae are bright structures on the photosphere of a star. In the Sun, these are formed by concentrations of magnetic field lines becoming hotter than the surrounding photosphere (Spruit, 1976). They are formed in the space between the granules. Also, they appear close to the solar limb in plages or active regions (Keller et al., 2004). Their contribution to the total solar irradiance varies with the sunspot cycle. It is higher during the sunspots maximum (Berger et al., 2007).

A plage is a bright structure of the chromosphere of a star. These are formed in regions of higher temperature and density near the spots. Similarly to

spots and faculae are associated with magnetic fields and are created when these fields compress the chromosphere (Solanki, 1993).

Being magnetic in origin, both plage and faculae, are known to suppress the convective blueshift inside active regions and lead to RV variations up to 10m/s (Meunier, N. et al., 2010). These are the processes that dominate the activity signal on slow rotating and older stars. (e.g Reinhold et al., 2019).

On timescales of years

- *Magnetic cycles*

Spots, plages and faculae have a close relationship with the solar magnetic cycle. This 11-years is produced by dynamo processes in the Sun, where the number of spots, plages and faculae increase or decrease as the cycle progress (Hathaway, 2015; Charbonneau, 2020). For example, the number of sunspots has been registered since 1749 showing a clear increase and decrease as one cycle evolves (see figure 2.8). This magnetic cycle have been observed in other stars (e.g Lovis et al., 2011; Dumusque et al., 2011).

With the search of long orbital period planets, there is now the need to deal with any signal related to this cycle. On other stars the magnetic cycle can last between 7 up to 30 years, depending on its class and age (Baliunas et al., 1998). On the Sun studies showed it can create a Doppler signal up to 16.4 m/s, and reach values as high as 25m/s on other stars(Lovis et al., 2011).

Luckily these magnetic cycles correlate well with the variation present on activity indicators. That means the signal can be identified using, for example, a linear relationship between the RV and the indicator (e.g Dumusque et al., 2012). Even so, this can bring problems if the orbital period of the planet is close to that of the cycle. In the future, better methods to deal with the magnetic cycles signal will most likely be required.

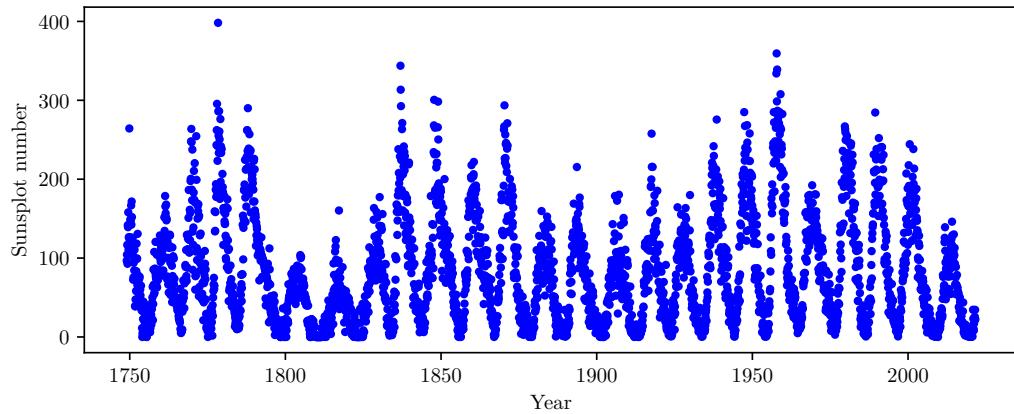


Figure 2.8: Monthly mean of the number of sunspot observed since 1749.
Source: <https://wwwbis.sidc.be/silso>

2.3.2 Activity proxies

Earlier, I mention the use of activity indicators to deal with some stellar activity but did not explain it. An activity indicator or activity proxy is an auxiliary time series usually with the same timestamps as the RV. These might be obtained from the same CCF we use to calculate the RV or from the spectroscopic observations.

Let us, for example, consider the case of only one spot on the surface of a star. For it, I use a simple scheme I made in figure 2.9 simulating it. As the spot crosses our field of view, going from 1 to 5, its signal affects the CCF shape. In our case, as the spot enters our field of view (2) it distorts the CCF. This distortion "evolves" until the spot is out of the field of view (5) and the CCF returns to normal.

To focus only on the activity indicator let us not consider the impact this would have on the RV measurement. It is easy to notice in figure 2.9 that any change in the CCF, will have an impact on both the FWHM (represented on the red horizontal line) and the bisector (represented by the blue vertical line). In the case of the FWHM, if the CCF represented by (1) has the value y_1 . The asymmetry on (2) does not make the FWHM change position, but its value will be smaller than y_1 . On pane (3) the spot affects the peak of the CCF and, as consequence, the

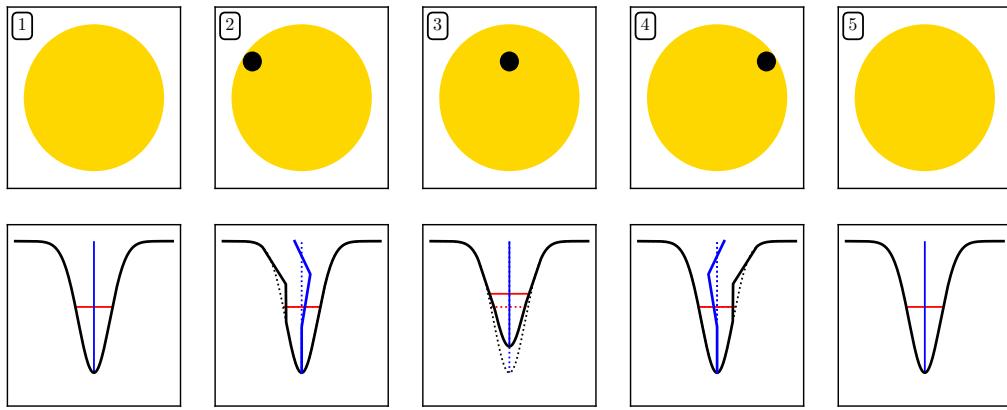


Figure 2.9: Simplified illustration of the impact of a spot on the CCF.

FWHM changes position. If the wings are left unaffected¹⁰, the value should be bigger than y_1 , as the width where the FWHM is now located is also bigger. Then on the pane (4), the FWHM decreases to a value equal to (2), increasing again to the value y_1 on (5). This is roughly plotted on figure 2.10. All this being occurring due to non-planetary signals as a planet would only shift all of the CCF from the 0 m/s mark.

Even if the previous paragraph was a gross oversimplification on the impact of a spot. As the star rotates, the spot, enter and leaves our line of sight, always deforming the CCF. Due to this, a periodicity will appear for values close to the rotation period from the changes the spot causes on the FWHM. Using this periodic information, we can associate it with the same periodic signals on the RV and reject them as a spot signal. Similarly, we can try to interpret the bisector behaviour of figure 2.9 and associate its changes with the impact of the spot. Is this that allows the creation of time-series like the bisector inverse slope.

It is the information contained on these other time-series that, for example, Rajpaul et al. (2015) and Jones et al. (2017) use to better characterize planetary signals. Similarly, in my thesis, I create another framework using Gaussian processes that use auxiliary time series to analyse any RV measurements. Next, I

¹⁰ It is a big simplification on the impact of a spot on the CCF. From pane (1) to (3) I am stating that only the peak of the Gaussian was "compressed" and the rest left unaffected.

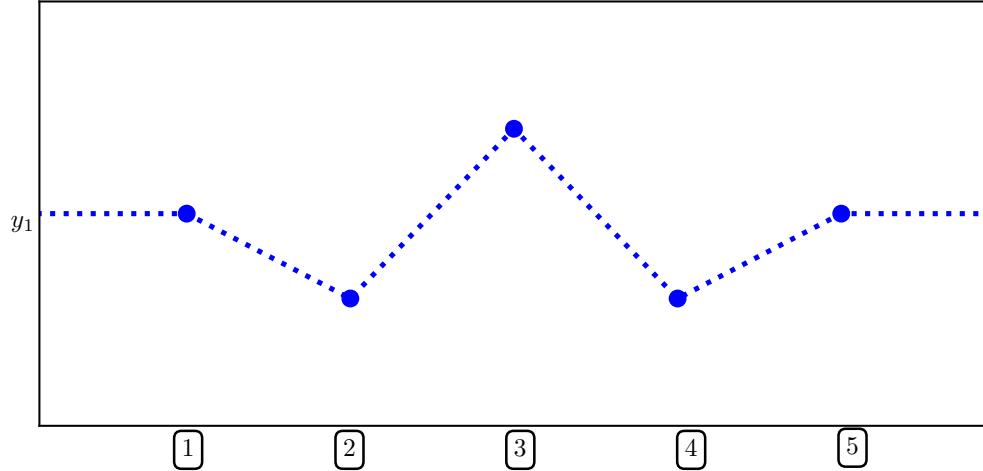


Figure 2.10: Expected behavior of the FWHM measurements on the example I gave on figure 2.9. On the x-axis are represented the five panes.

will define some of the activity proxies time series we can use.

Bisector inverse slope

The bisector inverse slope (BIS) is an asymmetry measure of the CCF. An early use can be seen, for example, on Gray (1997). On it, the authors suggested using the overall slope of the bisector, measured between bisector points at 0.50 and 0.90 of the continuum. As these measures will be different between the top and bottom parts, it is sensitive to any signal contaminating the line profile of the CCF. (Gray and Hatzes, 1997).

Improvement exists to this measurement. Queloz et al. (2001) used the difference between the top and bottom regions of the bisector, \overline{V}_t and \overline{V}_b respectively. That done to increase the accuracy of the BIS means defining it as

$$BIS = \overline{V}_t - \overline{V}_b. \quad (2.36)$$

One can define different boundaries for the top and bottom regions. That, however, requires some caution when comparing results (Queloz et al., 2001;

Figueira et al., 2013; Santerne et al., 2015). As with the FWHM, it is possible to connect the contamination of stellar activity in our data to the measures obtained for the BIS. Similarly, I will explore this better up ahead.

V_{span}

This asymmetry measurement of the CCF was proposed by Boisse et al. (2010). We can describe the V_{span} as a two-step fit of the CCF, using two Gaussian. First, we fit the top part of the CCF (RV_{high}), and the lower part (RV_{low}). The V_{span} can then be calculated as

$$V_{\text{span}} = RV_{\text{high}} - RV_{\text{low}}. \quad (2.37)$$

This proxy is less sensitive to noise, meaning this measurement can be advantageous over the BIS for observations with low S/N (Figueira et al., 2013).

V_{asy}

This indicator was developed by Figueira et al. (2013) and is known as V_{asy} . It compares the asymmetry between the red and the blue wings of the line profile. On it, we compute the spectral information between the red and the blue wings of the line profile

$$V_{\text{asy}} = \frac{\sum_{\text{flux}} (W_i(\text{red}) - W_i(\text{blue})) \times \overline{W_i}}{\sum_{\text{flux}} \overline{W_i}}, \quad (2.38)$$

where W_i is the weight for the point calculated at the flux level i. To address some limitations, Figueira et al. (2015) defined W_i as

$$W_i = c^2 \frac{\left(\frac{\partial CCF(i)}{\partial RV(i)} \right)^2}{\sigma_{CCF}(i)^2}. \quad (2.39)$$

$\sigma_{CCF}(i)$ is the error on the RVs induced by the flux measurement in pixel i, and c is the speed of light in the vacuum.

This indicator shows significant correlations with RV for active stars. it is shown to be correlated with the RV for signals due to line deformations, making it more advantageous than the BIS (Figueira et al., 2013).

S index

The S index is a magnetic activity proxy developed for the HK Project by the Mount Wilson Observatory (Vaughan et al., 1978). Created from the regularly observed CaII H and K emission lines of bright dwarf stars to characterize their magnetic activity (Wilson, 1978). It is expressed as

$$S = \alpha \frac{N_H + N_K}{N_R + N_V}, \quad (2.40)$$

where N_H and N_K are the fluxes in 1.09Å triangular bands centred on the CaII H and K lines, N_R . N_V are 20Å reference bandpasses in the continuum, and α is a calibration constant determined nightly (Vaughan et al., 1978).

Some limitations in this index led to the development of the $\log(R'_{HK})$.

$\log(R'_{HK})$

The R'_{hk} index was developed by Noyes et al. (1984). This index uses the fluxes of the H and K lines cores to calculate a star's bolometric luminosity radiated as chromospheric H and K emission (Hall, 2008). It is defined as

$$R'_{HK} = \frac{F_H + F_K}{\sigma T^4}, \quad (2.41)$$

where F_H and F_K are respectively the fluxes of the H and K lines subtracted by the chromospheric fluxes of a reference star, σ the Stefan-Boltzmann constant and T the star's temperature (Martínez-Arnáiz, R. et al., 2010). Most commonly this index is used in its logarithmic form $\log(R'_{hk})$.

This index is widely used to identify the level of activity a star possesses. Generally the boundary between active and inactive lies at $\log(R'_{hk}) = -4.75$, a limit known as Vaughan-Preston gap (Vaughan and Preston, 1980). This implying more active stars will have a $\log(R'_{hk}) > -4.75$, and less active a $\log(R'_{hk}) < -4.75$.

Chapter 3

Gaussian Process Regression Network

In this chapter, I start by presenting the mathematical background for Gaussian processes analysis. I follow it with the presentation of some frameworks currently used on radial velocity analysis. I also give some frameworks currently used on radial velocity analysis. I conclude the chapter with a detailed explanation of the framework I used on my PhD project. I show the principles behind a Gaussian process regression network and how to use it on exoplanet search.

3.1 Gaussian processes

Currently Gaussian processes is one of the most advanced fields in statistics, used various disciplines such as Finance (e.g. Nirwan and Bertschinger, 2020), pharmacology(e.g. Sahli Costabal et al., 2019), and more recently COVID-19 outbreaks prevention (e.g. Ketu and Mishra, 2020). This tool also proved do be successful in astronomy in fields such as cosmology (e.g. Melia and Yennapureddy, 2018), and of course, in exoplanet detection to model the stellar activity (e.g. Faria et al., 2016; Serrano et al., 2018; Barros et al., 2020). A GP is often seen as an esoteric tool that is used because it works, while the user ignores the mathematical background. Hoping to change this I will start with a brief introduc-

tion on the Gaussian distribution before diving into a GP per se. This chapter will only be a brief introduction to GPs and their properties. I recommend the reader to Rasmussen and Williams (2006) for a more mathematical approach.

3.1.1 The Gaussian distribution

In nature many random processes follow a Gaussian probability distribution function¹. This function is known for its "bell" shape and has a probability density function² (PDF) given by

$$p(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]. \quad (3.1)$$

This 1-dimensional (1-d) Gaussian distribution is thus defined by a mean μ and standard deviation σ (Patel and Read, 1996). This and its "bell shape" means it will have its maximum when $x = \mu$, and width proportional to σ (Sivia and Skilling, 2006). In figure 3.1 it is shown that it is a symmetric distribution that shifts on the x-axis according to its mean and increases in width with the increase of its standard deviation. If a random variable x follows a Gaussian distribution with a mean μ and standard deviation σ , we can thus write that

$$x \sim \mathcal{N}(\mu, \sigma^2). \quad (3.2)$$

The same way we define a 1-d Gaussian distribution, we can define a 2-d, 3-d, or n-d with $n \in \mathbb{N}^+$. A multivariate Gaussian distribution has a PDF defined as

$$p(X|\mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2} (X - \mu)^T \Sigma^{-1} (X - \mu)\right) \quad (3.3)$$

If $X = [X_1, X_2, \dots, X_k]$ is a k-dimensional random vector, then μ is now defined as a mean vector of size k, and Σ as a $k \times k$ covariance matrix containing the covariance between each pair of elements of the k-dimensional random vector (Gut, 2009). Similarly to the 1-d case, if a random vector X follows a

¹ A Gaussian distribution is often known as a *Normal distribution*.

² The probability density function of a continuous random variable, is a function whose integral gives us the probabilities associated with such random variable (Grinstead and Snell, 2012).

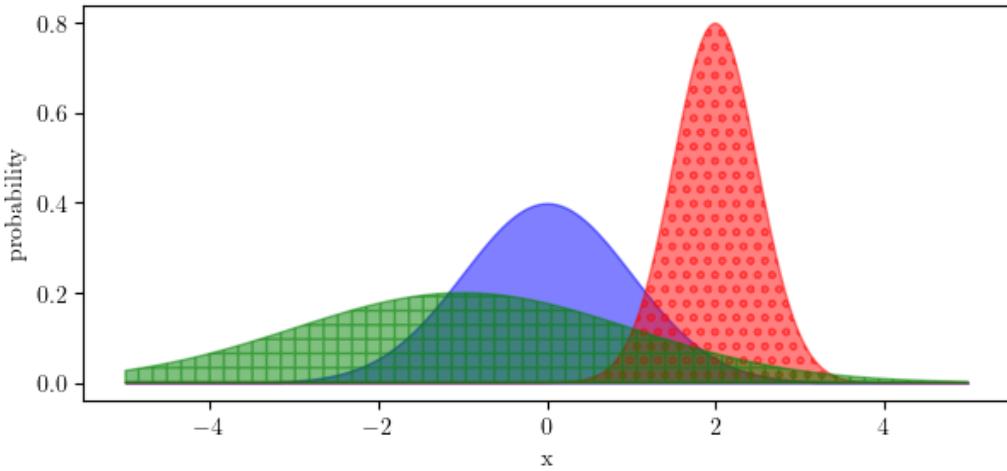


Figure 3.1: Three 1-dimensional Gaussian distributions. At solid blue a distribution with $\mu = 0$ and $\sigma = 1$, at dotted red one with $\mu = 2$ and $\sigma = 0.5$, and at checkered green one with $\mu = -1$ and $\sigma = 2$.

multivariate Gaussian distribution with a mean μ and covariance Σ , we can write it as

$$X \sim \mathcal{N}(\mu, \Sigma). \quad (3.4)$$

3.1.2 The Gaussian process

Now that we defined the multivariate Gaussian distribution, we can look at a GP as the generalization of the multivariate Gaussian distribution. It is characterized by a mean function $m(t)$ and a covariance function $k(t, t')$ ³.

A Gaussian process provides a useful Bayesian tool for modelling using a non-parametric model. On a GP the parameters we are trying to learn are functions. We first create a prior distribution over it and then condition it to the data. With this, we combine information from a sample function of our data with the prior distribution and encapsulate it with a posterior distribution (see figure 3.2).

³ Note that, for here on, I will use t as input instead of the usual x . That is due to, in my work, the input usually defines the time of a given observation.

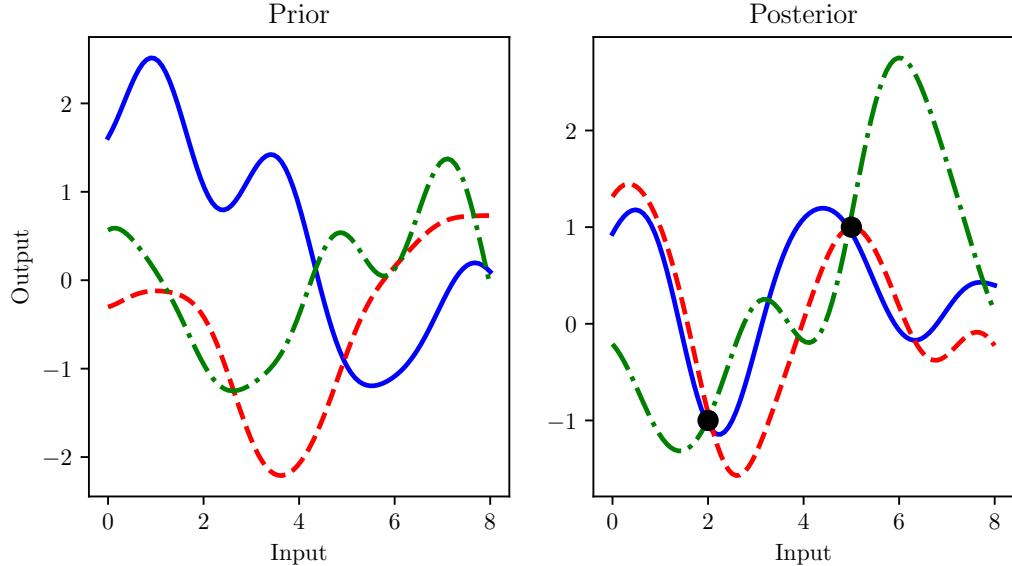


Figure 3.2: Left: Three samples drawn from a prior distribution. Right: Three samples drawn from the posterior distribution after two observations with the prior conditioned to these observations.

Formally one can define a GP $f(t)$ as a collection of random variables, any finite number of which have a joint Gaussian distribution (Rasmussen and Williams, 2006). This implies we can write our GP as

$$f(t) \sim \mathcal{GP}(m(t), k(t, t')), \quad (3.5)$$

where we have

$$\begin{aligned} m(t) &= E[f(t)], \\ k(t, t') &= E[(f(t) - m(t))(f(t') - m(t'))], \end{aligned} \quad (3.6)$$

where E is the expected value.

As previously mentioned, a GP is a non-parametric model. That means the model structure is determined from data instead. In exoplanet search, this flexibility allows astronomers to not only constrain a model to the data but to deal with the correlated stellar activity component of the measurements (e.g. Haywood

et al., 2014; Faria et al., 2016; Cloutier, R. et al., 2019). A reader most likely will learn that a GP is, for simplicity, defined as having a zero mean. In exoplanet search, this is not true. The most common mean function used is a sum of keplerian functions similar to equation 1.1.

That designates the kernel to be the element in a GP that will deal with the stellar activity component of our data while the mean will fit the planetary signals. Since this activity is correlated and quasi-periodic, the most used kernel is known as the quasi-periodic kernel, see section 3.1.3 for a more detailed description of this kernel. This kernel became the go-to at exoplanet search using GPs due to its parameters often associated with physical properties of the stellar activity like the stellar rotation and decaying timespan of the active regions.

Having defined mean and covariance functions, the next step is to optimize all the parameters in our model accordingly to the given data. For this, we can make use of two paradigms. Either Minimizing or maximizing objective functions or using Markov chain Monte Carlo (MCMC) methods to obtain samples from the parameter space.

For the first one can use, for example, `scipy.optimize.minimize`, a set of mathematical algorithms built on the SciPy library of Python⁴ (Virtanen et al., 2020). Minimizing an objective function, although straightforward, has its challenges. These methods are known to be sensitive to the existence of local minima. Meaning that the optimized result will sometimes not correspond to the true global minimum desired by the user.

More reliable are the sampling methods. Some widely used libraries in astronomy include, for example, `emcee` (Foreman-Mackey et al., 2019), `dynesty` (Speagle, 2020), and `zeus` (Karamanis et al., 2021). These libraries use different sampling methods, each with its weak and strong points. `emcee` uses an affine invariant MCMC ensemble sampler developed by Goodman and Weare (2010) useful to sample badly scaled distributions but not well suited for high dimensional or multi-modal probability surfaces. `dynesty` uses a nested sampling method developed by Higson et al. (2019) capable of calculating Bayesian poste-

⁴ Since the computational part of my work used the Python programming language, I will limit the examples given to this language. Many other programming languages exist, such as the R programming language or the MATLAB numeric computing environment, with similar algorithms.

priors and evidences, but as with other nested sampling methods, it can often be very inefficient to explore the parameter space in progressively higher likelihoods. Lastly, `zeus` implements the ensemble slice sampling method, an extension of slice sampling developed by Neal (2000) capable of sampling from highly correlated distributions.
maybe add refs from works that used emcee, dynesty and zeus

Independently of the method used, when we try to optimize the model parameters is common to use the log marginal likelihood ($\log \mathcal{L}$) as evaluation criteria. This quantity takes the form of

$$\log \mathcal{L} = -\frac{1}{2} (y - m)^T K^{-1} (y - m) - \frac{1}{2} \log |K| - \frac{n}{2} \log (2\pi), \quad (3.7)$$

where $y - m$ is the RV residuals (the subtraction of RV measurements by the mean function), K the covariance matrix of the GP, and n the number of RV measurements. One can divide equation 3.7 into three terms. $-\frac{1}{2} (y - m)^T K^{-1} (y - m)$ is the only one that includes our data and is responsible for the fit. The term $-\frac{1}{2} \log |K|$ will measure and penalize models according to their complexity⁵. The last term $-\frac{n}{2} \log (2\pi)$ is only a normalization constant (Rasmussen and Williams, 2006).

The log marginal likelihood value alone has no useful interpretation. Its advantage comes when used in a model with two different sets of parameters values. The value obtained give us an indication of which set is a better representation of the data. The set with a higher value is considered the best solution. For example, a set of parameters whose log marginal likelihood is -100 is better than one that obtained -200.

To compare GP models we can use the marginal likelihood \mathcal{Z} , that in my work I will call evidence. This measurements is the fully marginalized likelihood that acts as normalization factor in parameter estimation (Sivia and Skilling, 2006). We can define it as

$$\mathcal{Z} \equiv p(D | M_i) = \int p(\theta | M_i) p(D | \theta, M_i) d\theta, \quad (3.8)$$

⁵ We can call this term as the *Occam's Razor* in GP model selection. Named after William of Occam (1285-1347) famous for the quote "The explanation requiring the fewest assumptions is most likely to be correct".

where M_i is our model, D the observed data, and θ the parameters vector. In practice, this quantity is challenging to compute, but approximations are known, such as the one derived by Perrakis et al. (2014).

From the evidence of two different models M_1 and M_2 , we can then obtain the likelihood ratio know as Bayes factor (K_B)

$$K_B = \frac{\mathcal{Z}_{M_1}}{\mathcal{Z}_{M_2}}. \quad (3.9)$$

The natural logarithm of the Bayes factor is then used as model selection criteria (Jeffreys, 1983; Kass and Raftery, 1995). For example, when comparing a 0-planet model with evidence against a 1-planets model, a Bayes factor of 150 between them indicates that there is enough evidence to claim a planetary detection (e.g. Faria et al., 2016; Faria et al., 2020).

3.1.3 The Covariance function

The covariance function, most commonly know as kernel, is the most important element in a GP. It is the kernel that controls all the modelling features. As such, it is crucial to choose an appropriate one to properties present in the data (e.g. periodicity, smoothness, or stationarity).

Some important aspects of a kernel are that if it is a function of $t - t'$, we call it *stationary*. That implies it is invariant to translations, meaning that the kernel value remains the same independent of how the inputs shifts among t (Duvenaud, 2014). If the kernel is a function of $|t - t'|$ we classify it as *isotropic* whose value does not depend on direction (Rasmussen and Williams, 2006).

A kernel $k(t, t')$ will create a Gram matrix K containing the covariance between any two points t and t' . This Gram matrix will be positive semidefinite matrix, that is, a symmetric matrix with non-negative eigenvalues (Rajpaul, 2017). Taking into consideration the set of input points $\{t_1, t_2, \dots, t_n\}$, K will take the form

$$K = \begin{bmatrix} k(t_1, t_1) & k(t_1, t_2) & \dots & k(t_1, t_n) \\ k(t_2, t_1) & k(t_2, t_2) & \dots & k(t_2, t_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(t_n, t_1) & k(t_n, t_2) & \dots & k(t_n, t_n) \end{bmatrix}. \quad (3.10)$$

Since this Gram matrix contains the covariance of the input points, we call it a covariance matrix.

We can now look at two of the most important properties of a GP kernel, addition and multiplication, both proofs for them can be found in Rasmussen and Williams (2006). These are:

- The sum of two kernels $k_1(t, t')$ and $k_2(t, t')$ is a valid kernel. This means that if we have two independent GPs $f_1(t) \sim \mathcal{GP}(0, k_1(t, t'))$ and $f_2(t) \sim \mathcal{GP}(0, k_2(t, t'))$, we can created a new GP $g(t)$

$$g(t) \sim \mathcal{GP}(0, k_1(t, t') + k_2(t, t')). \quad (3.11)$$

- The product of two kernels $k_1(t, t')$ and $k_2(t, t')$ is also a valid kernel. Again, if we have two independent GPs $f_1(t) \sim \mathcal{GP}(0, k_1(t, t'))$ and $f_2(t) \sim \mathcal{GP}(0, k_2(t, t'))$, we can thus created a GP $h(t)$ defined by

$$h(t) \sim \mathcal{GP}(0, k_1(t, t') \times k_2(t, t')). \quad (3.12)$$

These two properties allow the creation of new kernels capable of modelling data with a more complex structure. These properties are fundamental if we want to determine what kernels we might want to use in RV data analysis. See Duvenaud (2014) for a more detailed work on the structures obtained by the sum and/or multiplication of different kernels.

In the following pages, I present the main kernels of interest in GP analysis. Each kernel is characterized by a set of parameters (commonly called hyperparameters) that I represent using the Greek alphabet plus an index. To avoid repeating definitions, for different kernels, the parameter represented by the same letter and index share the same interpretation unless I mention it otherwise on the text.

Squared-exponential kernel

The first kernel I present is *squared-exponential kernel*⁶, one of the most widely used kernels in machine learning. It is given by

$$k(t, t') = \eta_1^2 \exp\left(-\frac{(t - t')^2}{2\eta_2^2}\right), \quad (3.13)$$

where η_1 represents the amplitude of the signal and η_2 is a parameter called *length-scale*. One way of interpreting this length-scale is as being the time required for the function change significantly (Rasmussen and Williams, 2006). The higher this parameter, the smoother the squared-exponential kernel becomes, see figure 3.3. In the limit, if $\eta_2 \rightarrow \infty$ the squared exponential kernel becomes the constant kernel. Although widely used, this kernel is known for being too smooth to realistically model a considerable number of physical processes (Stein, 1999). As a consequence, I will next present two alternatives to the squared-exponential.

Rational quadratic kernel

A possible replacement of the squared exponential kernel is the *rational quadratic kernel* given by

$$k(t, t') = \eta_1^2 \left(1 + \frac{(t - t')^2}{2\alpha\bar{\eta}_2^2}\right)^{-\alpha}. \quad (3.14)$$

While a squared exponential assumes the data evolves at one particular time scale, it might be useful to consider a sum of squared exponential kernels, each with different time scales. The rational quadratic allows this. This kernel is an infinite sum of squared exponential kernels of distinct time scales, controlled by the parameter α . We reach this kernel by expressing the scale mixture of different squared exponential kernels from a gamma distribution with shape α and mean $\bar{\eta}_2^{-2}$ (Wilson, 2014).

⁶ The squared-exponential kernel is commonly called as the radial basis function or RBF kernel in machine learning.

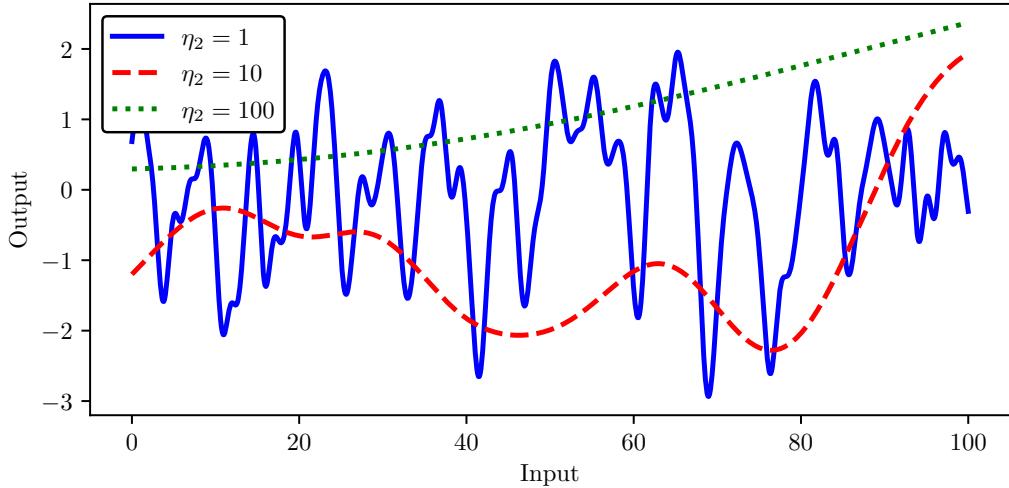


Figure 3.3: Samples of a squared-exponential kernel with different length-scales. Solid blue $\eta_2 = 1$, dashed red $\eta_2 = 10$, and dotted green $\eta_2 = 100$.

This deduction allows $\bar{\eta}_2$ to be the mean (thus the bar over this term) time scale of the infinite sum of different squared exponential kernels, each with a different decaying timescale. When $\alpha \rightarrow \infty$, the rational quadratic kernel converges to the squared exponential kernel (Rasmussen and Williams, 2006).

Matérn family of kernels

The Matérn family of covariance functions are defined as

$$k(t, t') = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2}(t-t')}{\eta_2} \right)^\nu K_\nu \left(\frac{\sqrt{2\nu}(t-t')}{\eta_2} \right), \quad (3.15)$$

where Γ is the gamma function (Artin and Butler, 2015), K_ν is the modified Bessel function of the second kind (Abramowitz and Stegun, 1964), and ν is a positive parameters. If $\nu \rightarrow \infty$ we obtain the squared exponential kernel (Rasmussen and Williams, 2006).

Of this family is of interest to consider ν as a half-integer. The three most

common in machine learning are known as the *Matérn 1/2*⁷, *Matérn 3/2* and the *Matérn 5/2* kernels. This last one was already been used in RV analysis by Gilbertson et al. (2020) alone. They are defined as

$$k(t, t') = \eta_1^2 \left(-\frac{(t - t')}{\eta_2} \right), \quad (3.16)$$

$$k(t, t') = \eta_1^2 \left(1 + \frac{\sqrt{3}(t - t')}{\eta_2} \right) \exp \left(-\frac{\sqrt{3}(t - t')}{\eta_2} \right), \quad (3.17)$$

and

$$k(t, t') = \eta_1^2 \left(1 + \frac{\sqrt{5}(t - t')}{\eta_2} + \frac{5(t - t')^2}{3\eta_2^3} \right) \exp \left(-\frac{\sqrt{5}(t - t')}{\eta_2} \right), \quad (3.18)$$

respectively.

The η_2 parameters share the same interpretation as the one given for the squared exponential kernel. Some works use the Matérn kernels as a solution to deal with the smoothness of the squared exponential kernel (e.g. Abdessalem et al., 2017). For comparison, and given the same parameters, we see from samples taken with a Matérn 5/2 kernel (figure 3.4), it is far less smooth than the ones drawn for a squared exponential kernel (figure 3.3).

Periodic kernel

Another kernel of interest is the *periodic kernel* used to model data that, as the name gives say, has a periodic behaviour. It is usually written as

$$k(t, t') = \eta_1^2 \exp \left[-\frac{2}{\eta_4^2} \sin^2 \left(\frac{\pi(t - t')}{\eta_3} \right) \right]. \quad (3.19)$$

This kernel have two new parameters. The period of the repetitions η_3 and a length-scale η_4 that models the structure inside a period. Despite the fact that it is

⁷This kernel is also be called as exponential kernel in other works.

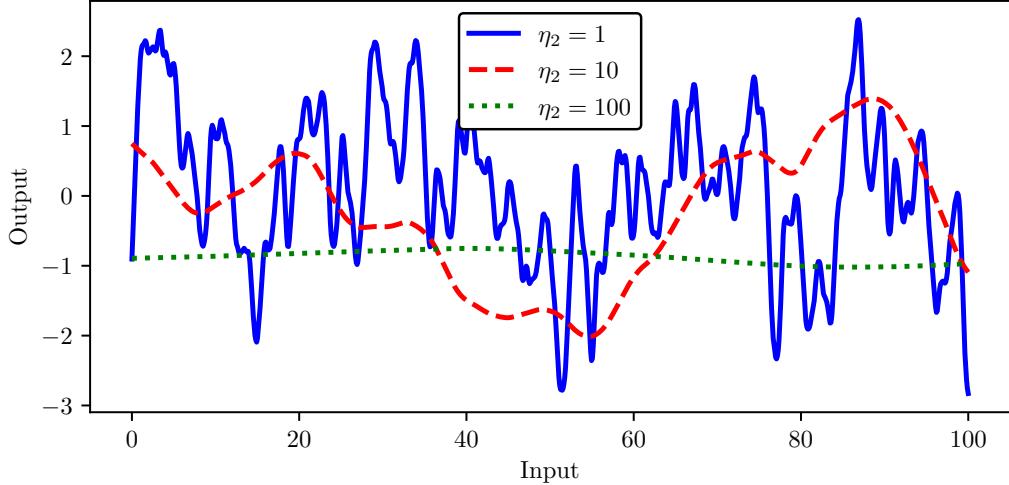


Figure 3.4: Samples of a Matérn 5/2 kernel with different length-scales. Solid blue $\eta_2 = 1$, dashed red $\eta_2 = 10$, and dotted green $\eta_2 = 10$.

widely used, its origins are sometimes ignored. Mackay (1998) derived this kernel by mapping(or warping) the squared exponential kernel into the two-dimensional space $u(t) = \left[\cos\left(\frac{2\pi}{\eta_3}t\right), \sin\left(\frac{2\pi}{\eta_3}t\right) \right]$.

The deduction is trivial, quick, and of interest to show. We start from the squared exponential kernel of equation 3.13, and map it into the two-dimensional space⁸. With this we obtain a new kernel given by

$$k(t, t') = \exp \left[-\frac{(\cos t - \cos t')^2 + (\sin t - \sin t')^2}{2\eta_2^2} \right].$$

Using the binomial theorem (Abramowitz and Stegun, 1964) we can expand the squared terms into

$$\begin{aligned} (\cos t - \cos t')^2 &= \cos^2 t - 2 \cos t \cos t' + \cos^2 t', \\ (\sin t - \sin t')^2 &= \sin^2 t - 2 \sin t \sin t' + \sin^2 t', \end{aligned}$$

allowing us to use the Pythagorean trigonometric identity $\cos^2 X + \sin^2 X = 1$, to

⁸To simplify the notation I omitted the $\frac{2\pi}{\eta_3}$ term as it does not affect the deduction.

simplify $k(t, t')$ into

$$k(t, t') = \exp\left(-\frac{2 - 2 \cos t \cos t' - 2 \sin t \sin t'}{2\eta_2^2}\right).$$

We now make use of the prosthaphaeresis formulas (Kung, 1996)

$$2 \cos X \cos Y = \cos(X - Y) + \cos(X + Y),$$

$$2 \sin X \sin Y = \cos(X - Y) - \cos(X + Y),$$

and transform the kernel into

$$k(t, t') = \exp\left[-\frac{1 - \cos(t - t')}{\eta_2^2}\right].$$

The last steps require the use of the double-angle formula $\cos 2X = \cos^2 X - \sin^2 X$ to give us

$$k(t, t') = \exp\left[-\frac{1 - \cos^2\left(\frac{t-t'}{2}\right) + \sin^2\left(\frac{t-t'}{2}\right)}{\eta_2^2}\right].$$

Using again the Pythagorean identity we obtain

$$k(t, t') = \exp\left[-\frac{2}{\eta_2^2} \sin^2\left(\frac{t-t'}{2}\right)\right].$$

Thus ending the deduction. Not forgetting the $\frac{2\pi}{\eta_3}$ term we obtain the periodic kernel

$$k(t, t') = \exp\left[-\frac{2}{\eta_2^2} \sin^2\left(\frac{\pi}{\eta_3}(t - t')\right)\right].$$

The deduction allows me to mention two details. First, the η_4 of the periodic kernel and the η_2 of the squared exponential kernel are the same or at least should share the same characteristics. This shows that η_4 is a length-scale that will determine the smoothness, or lack of it, inside a period given by eta_3 (see figure 3.5). The higher the value of η_4 the more smooth is the structure inside our

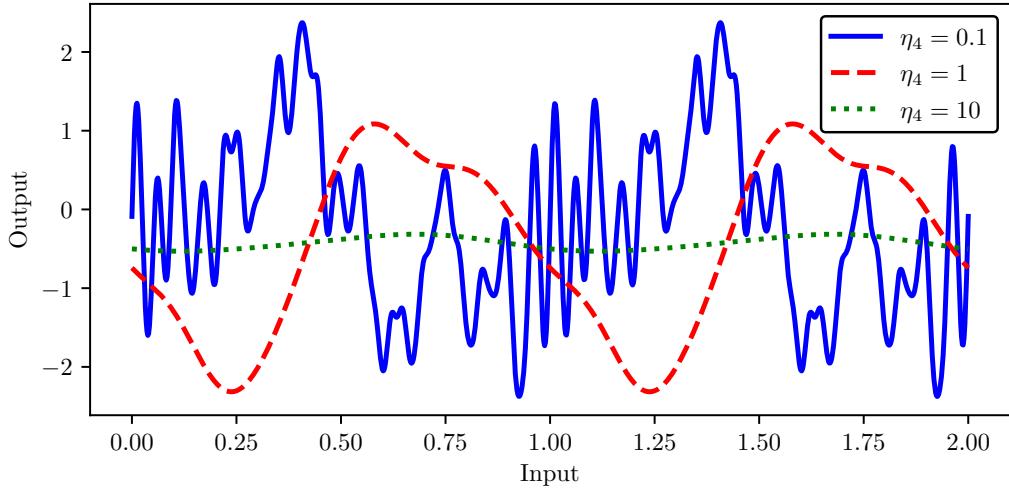


Figure 3.5: Samples of a periodic kernel with $\eta_3 = 1$ but varying η_4 . Solid blue $\eta_4 = 0.1$, dashed red $\eta_4 = 1$, and dotted green $\eta_4 = 10$.

period. In other words, as $\eta_4 \rightarrow \infty$, the periodic kernel tends to a smooth sinusoid. The other important detail out of this deduction is the possibility of new periodic kernels. To create new periodic kernels, the same mapping can be done using other kernels. I will explore this more further ahead.

Quasi-periodic kernel

The most common kernel in exoplanet search is the called *quasi-periodic kernel*. Since the signal induced by the active regions on an RV set is not strictly periodic, that is, it evolves over time, we need to use a quasi-periodic function capable of having this into consideration. We attain this by using the property discussed in equation 3.12 and multiplying two "simpler" kernels, the periodic and the squared-exponential kernels. This multiplication creates a model that decays away from the detected periodicity, allowing a locally periodical structure (Duvenaud et al., 2013). The quasi-periodic kernel is thus defined as

$$k(t, t') = \eta_1^2 \exp \left[-\frac{(t - t')^2}{2\eta_2^2} - \frac{2}{\eta_4^2} \sin^2 \left(\frac{\pi}{\eta_3} (t - t') \right) \right]. \quad (3.20)$$

In this kernel, η_2 is often interpreted as being the decaying time scale of the active regions on the surface of a star. η_3 gives the rotation period of the observed star, while η_4 is a parameter that controls the sinusoidal variations of the periodic signal, that is, the structure within η_3 . When analysing RV data, we also add a white noise term σ to account for any uncorrelated noise present in the data⁹.

Piecewise polynomial kernel

The Wendland compactly supported functions are a family of piecewise polynomial functions with compact support (Wendland, 2004). By definition, a piecewise polynomial kernel with compact support is a kernel that allows the covariance between two points to be zero when its distance exceeds an imposed threshold, creating a sparse covariance matrix in the process (Rasmussen and Williams, 2006). This characteristic implies that by multiplying a *piecewise kernel* by a squared exponential kernel, we obtain a new covariance function suited to model short-scale correlations. In this family, increasing the polynomial order creates increasingly smoother kernels (Davies, 2015). We can, for example, use a third-order piecewise polynomial and defined the kernel

$$k(t, t') = \eta_1^2 \exp\left(-\frac{(t - t')^2}{2\eta_2^2}\right) (3r_p + 1)(1 - r_p)^3, \quad (3.21)$$

where we have $r_p = 0$ if $|t - t'| > \frac{\eta_3}{2}$ and $r_p = (t - t')$ otherwise. I chose to show a third order polynomial due to having a similar behavior to a squared exponential for $0 < \frac{2|r|}{\eta_3} < 1$. It will effectively truncate the squared exponential to small length correlations.

One challenge while analysing RV data with a *QP* kernel is the effect that small length correlations might have. These might be responsible for lower η_2 and higher η_4 values when these parameters try to model correlated noise variations. Having defined this new kernel we can added it to a quasi-periodic kernel. This

⁹It is usual to include a white noise term in most GP analysis but I have omitted this term for simplicity.

means having a new kernel defined as

$$k(t, t') = \eta_1^2 \exp \left[-\frac{(t - t')^2}{2\eta_2^2} - \frac{2}{\eta_4^2} \sin^2 \left(\frac{\pi}{\eta_3} (t - t') \right) \right] + \theta_1^2 \exp \left(-\frac{(t - t')^2}{2\theta_2^2} \right) (3r_p + 1)(1 - r_p)^3. \quad (3.22)$$

The idea behind this new kernel is to allow the quasi-periodic term to model other more significant signals while the piecewise part focuses on the 2 to 3 days length variations.

From rational quadratic to periodic

As mentioned earlier, having deduced the periodic kernel opens the door to the creation of new periodic kernel, and as consequence, new quasi-periodic kernels. Following similar steps we can achieve this using the rational quadratic kernel. In this case we map the rational quadratic kernel into the two-dimensional space $u(t) = [\cos(\frac{2\pi}{\eta_3}t), \sin(\frac{2\pi}{\eta_3}t)]$. Meaning, starting from

$$k(t, t') = \left[1 + \frac{(t - t')^2}{2\alpha\bar{\eta}_2^2} \right]^{-\alpha},$$

we obtain

$$k(t, t') = \left[1 + \frac{(\cos t - \cos t')^2 + (\sin t - \sin t')^2}{2\alpha\bar{\eta}_2^2} \right]^{-\alpha}.$$

We need to work this kernel using the expansion of the squared terms, followed by the use of the Pythagorean identity, prostaphaeresis and double-angle formulas. In the end we get

$$k(t, t') = \left[1 + \frac{2}{\alpha\bar{\eta}_2^2} \sin^2 \left(\frac{\pi}{\eta_3} (x - x') \right) \right]^{-\alpha}. \quad (3.23)$$

This new kernel has some interesting characteristics. In the rational quadratic kernel, η_2 was the mean time-scale of the infinite sum of squared exponential kernels controlled by α . Similarly, when $\alpha \rightarrow \infty$ this new kernel becomes the

periodic kernel from 3.19. An advantage of creating a periodic kernel from a rational quadratic is due to the latter being differentiable in mean square (MS differentiable) for every α . That does not occur with the Matérn kernels (Rasmussen and Williams, 2006). This implies, for example, $\frac{\partial^2 k(t, t')}{\partial t \partial t'}$ exists (Grigoriu, 2002).

To think: Other kernels of interest

3.2 Gaussian processes frameworks

The idea of using the information contained in activity indicators to better characterize RV measurements is not a new one. Other frameworks already exist that extend the use of GPs and allow the introduction of the information contained in other time series. These frameworks are an improvement concerning a standard GP regression framework using only RVs but are still not able to tackle the non-stationary behaviour of the stellar activity. The models assume the parameters stay constant over the whole input space (input independent). That means, for example, to assume a constant decaying timescale over the whole input space (Plagemann et al., 2008).

3.2.1 Rajpaul et al. (2015) framework

Developed as an extension of the FF' framework (Aigrain et al., 2012), this framework includes the $\log(R'_{hk})$ index and the BIS to disentangle the activity component from the planetary signals. The RVs and activity indicators are expressed as a linear combination of GPs given by

$$\begin{cases} \Delta RV = V_c G(t) + V_r \dot{G}(t), \\ \text{BIS} = B_c G(t) + B_r \dot{G}(t), \\ \log(R'_{HK}) = L_c G(t). \end{cases} \quad (3.24)$$

V_c , V_r , B_c , B_r , and L_c are free parameter that will control the amplitudes of the GPs. $G(t)$ is a GP with kernel $k_G(t, t)$. Each time-series had a respective mean function as well. For the RVs they used a second-order polynomial function while for the other time-series a constant function to fit the DC offset sufficed. This does not imply that other mean functions can not be implemented. The novel term $\dot{G}(t)$, is the derivative of $G(t)$. Considering that the derivative of a GP with kernel $k_G(t, t)$ is also a GP the only step left is to find its kernel. Osborne (2010) showed that for this new GP it is given by

$$k_{\dot{G}}(t, t') = \frac{\partial^2 k_G(t, t')}{\partial t \partial t'}. \quad (3.25)$$

That implies that if the kernel of $G(t)$ is not MS differentiable, it can not be used in this framework.

3.2.2 Jones et al. (2017) framework

This framework extend the work of Rajpaul et al. (2015). On it we have a class of models were multiple stellar activity indicators (more than two) can be taken into account. Formally it defined as

$$\begin{cases} u(t) = m_0 + a_{01}G(t) + a_{02}\dot{G}(t) + a_{03}\ddot{G}(t) + a_{04}Z_0(t) + \epsilon_0(t), \\ q_1(t) = m_1 + a_{11}G(t) + a_{12}\dot{G}(t) + a_{13}\ddot{G}(t) + a_{14}Z_1(t) + \epsilon_1(t), \\ \vdots \\ q_p(t) = m_p + a_{p1}G(t) + a_{p2}\dot{G}(t) + a_{p3}\ddot{G}(t) + a_{p4}Z_p(t) + \epsilon_p(t). \end{cases} \quad (3.26)$$

While Rajpaul et al. (2015) included only the first derivative of a GP, Jones et al. (2017) included a aditional second derivative to deal with the effects of the stellar spots area. They added zero mean GPs given by $Z_0(t), \dots, Z_p(t)$ to allow the fit of other structures besides the those that the model given by $X(t), \dot{X}(t)$, and $\ddot{X}(t)$ are capable of. Also included are white noise terms given by ϵ . Recently Gilbertson et al. (2020) used this framework to model stellar activity.

3.2.3 artGPN framework

Early in my PhD, I focused on modelling HARPS-like datasets containing the RV measurements (ΔRV) and three activity indicators (FWHM, BIS, and $\log R'_{HK}$). This lead to the creation of the GP framework shown in figure 3.6. On it, the outputs are independent GPs fitted using a linear combination of the nodes (f_1, f_2, \dots, f_q) and weights $(W_{11}, W_{22}, \dots, W_{3q}, W_{4q})$ kernels. Each output have the covariance function

$$\begin{cases} k_{\Delta\text{RV}} = k_{w_{11}}k_{f_1} + k_{w_{12}}k_{f_2} + \dots + k_{w_{1q}}k_{f_q}, \\ k_{\text{FWHM}} = k_{w_{21}}k_{f_1} + k_{w_{22}}k_{f_2} + \dots + k_{w_{2q}}k_{f_q}, \\ k_{\text{BIS}} = k_{w_{31}}k_{f_1} + k_{w_{32}}k_{f_2} + \dots + k_{w_{3q}}k_{f_q}, \\ k_{\log R'_{HK}} = k_{w_{41}}k_{f_1} + k_{w_{42}}k_{f_2} + \dots + k_{w_{4q}}k_{f_q}. \end{cases} \quad (3.27)$$

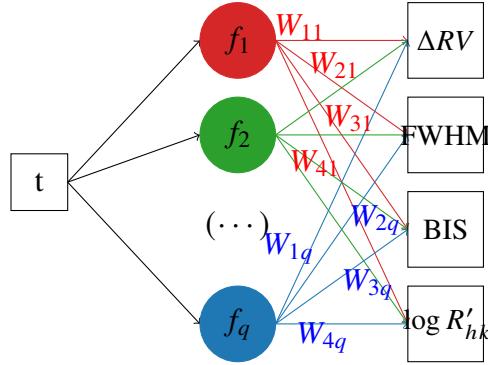


Figure 3.6: Scheme of our framework using HARPS-like datasets. Our regression network will be formed by q nodes (f_1, f_2, \dots, f_q) and $4 \times q$ weights ($w_{11}, w_{21}, w_{31}, \dots, w_{3q}, w_{4q}$) kernels connected similar to a neural network.

This allows to consider the outputs as GPs given by

$$\begin{cases} \Delta RV \sim \mathcal{GP}(m_{\Delta RV}(t), k_{\Delta RV}(t, t')), \\ \text{FWHM} \sim \mathcal{GP}(m_{\text{FWHM}}(t), k_{\text{FWHM}}(t, t')), \\ \text{BIS} \sim \mathcal{GP}(m_{\text{BIS}}(t), k_{\text{BIS}}(t, t')), \\ \log R'_{HK} \sim \mathcal{GP}\left(m_{\log R'_{HK}}(t), k_{\log R'_{HK}}(t, t')\right), \end{cases} \quad (3.28)$$

each with a mean $m_i(t)$ of any kind (constant, linear, etc...), and kernel $k_i(t, t')$ defined accordingly to the nodes and weights.

Equation 3.27 implies that this network can have a q number of nodes, each of a different kind (squared exponential, periodic, etc..). If necessary, the nodes can share parameters among themselves. For example, I could define two nodes, one periodic, and another quasi-periodic, with the same η_3 but the remaining parameters free. The weights, I usually defined as constant kernels to work as amplitudes, but can also take any other form. We can have, for example, squared exponential weights, all with the same η_2 value but different values of η_1 .

This condition allows the fit of the same signals present in different datasets. My model matches the semi-parametric latent factor model of Teh et al. (2005) if the weights are defined to be constant kernels, meaning it only scales the magnitude

of the nodes.

Depending on the output, the mean functions can take the form of a Keplerian, linear, quadratic, cubic, or sinusoidal function. The covariance functions are given by equation 3.27.

We can make use of the log marginal likelihood to evaluate this model. Assuming that each GP in the model is independent, the total log marginal likelihood ($\log p(\mathbf{y}|t)$) can be obtained but the sum of each output log marginal likelihood. Meaning we will have a log marginal likelihood given by

$$\log \mathcal{L}_{\text{total}} = \log \mathcal{L}_{\Delta RV} + \log \mathcal{L}_{FWHM} + \log \mathcal{L}_{BIS} + \log \mathcal{L}_{\log R'_{HK}}. \quad (3.29)$$

Individually we calculate the log marginal likelihood with equation 3.7. For example, the log $\mathcal{L}_{\Delta RV}$ is expressed as

$$\log \mathcal{L}_{\Delta RV} = -\frac{1}{2}\Delta RV^T K_{\Delta RV} \Delta RV - \frac{1}{2} \log |K_{\Delta RV}| - \frac{n}{2} \log 2\pi. \quad (3.30)$$

Due to its design I named this framework *artificial Gaussian process network* and a python implementation called artGPN can be found on Github¹⁰.

¹⁰<https://github.com/jdavidrcamacho/artgpn>

3.3 Gaussian processes regression networks framework

A Gaussian processes regression network (GPRN) was developed by Wilson et al. (2012). Unlike the previously GP frameworks, it is built in a way capable of handling non-stationary time series. While previous frameworks only account for fixed correlations among time series, a GPRN has an adaptive structure capable of capturing input dependent correlations across the input space. This is done by combining the properties of GP regression with the properties of single-layer artificial neural networks (Rumelhart et al., 1986) to create a Bayesian model for multi-output regression.

Wilson et al. (2012) created this framework to model a P -dimensional function $y(t)$ containing signal and noise correlations that vary with t . To this, they proposed modelling $y(t)$ as

$$y(t) = W(t) [f(t) + \sigma_f \epsilon(t)] + \sigma_y z(t). \quad (3.31)$$

On this network the $Q \times 1$ vector of latent functions (or nodes) $f(t)$ and $P \times Q$ matrix of weight functions (or weights) $W(t)$ are independent GPs defined as

$$\begin{aligned} f_q(t) &\sim \mathcal{GP}(0, k_{f_q}) \text{ for } q = 1, \dots, Q, \\ W_{pq}(t) &\sim \mathcal{GP}(0, k_{w_{pq}}) \text{ for } p = 1, \dots, P \text{ and } q = 1, \dots, Q, \end{aligned} \quad (3.32)$$

and $\epsilon(t)$ and $z(t)$ are white noise processes. This structure creates a linear combination of independent GPs, that by it turn creates the adaptive mixture of GPs capable of dealing with nonlinear correlations between the outputs. This also creates heavy-tail predictive distributions that does not easily over-fits the model (Li et al., 2020).

In the GPRN, the input dependent correlations are dealt with nodes $f(t)$, while the final amplitudes are given by the weights $W(t)$. These weights, being independent GPs, make the GPRN framework input dependent and non-stationary (Heinonen et al., 2015). What this mean is that, while on a GP, its kernel as amplitude η_1 that is a constant value this does not occur on a GPRN. On

it the kernel of each $y(t)$ is given by

$$k_{y_p}(t_a, t_b) = \sum_{q=1}^Q \left[W_{pq}(t_a) \left(k_{f_q}(t, t') + \sigma_f^2 \delta_{ab} \right) W_{pq}(t_b) \right] + \sigma_y^2 \delta_{ab}, \quad (3.33)$$

with δ_{ab} being the Kronecker delta. This kernel has an amplitude $W_{pq}(t)$, that being a GP, makes the GPRN structure non-stationary. If the weights use the same kernel, with the same parameters, the connections in the network will vary with t at an equal rate (Wilson et al., 2012).

Another advantage of a GPRN comes from the latent functions being able to take any form. That means we can have q GPs with any kind of kernel (periodic, quasi-periodic, squared exponential, etc...). This creates a mixture of the different kernels that is again input dependent and thus non-stationary, that continuously shifts within regions of completely different covariance structures (Wilson, 2014). This overall flexible structure of the GPRN is, of course, ideal to use on non-stationary signals coming from, for example, stellar activity.

One final detail of the overall structure is that, while both the nodes $f(t)$ and weights $W(t)$ are independent GPs, their product $f(t)W(t)$ will most likely not be a GP (Rasmussen and Williams, 2006). Not to be confused with the property 3.12 that says that the product of two kernels creates a valid kernel.

3.3.1 Variational inference

The GPRN makes use of RVs together with a p number of activity indicators (AI), with the same timestamps, in a scheme exemplified in the diagram of figure 3.7. The main objective of this framework is to use the information contained in the activity proxies to better separate activity from planetary signals in the RVs. The weight connections in figure 3.7 being dependent on t , will increase or decrease in strength with t . This allows the best nodes to fit the data accordingly, and if necessary, the time-varying weights will make a node have a different impact throughout the data.

Unlike a GP, the non-stationary nature of a GPRN makes inference intractable. This means that the posterior for the GPRN cannot be calculated ana-

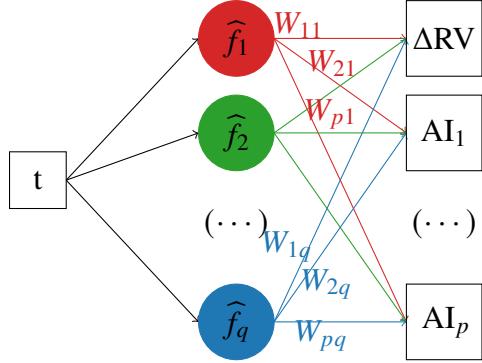


Figure 3.7: Example of a GPRN diagram considering a typical RV time-series with P activity indicator (AI). The regression network will be formed by Q nodes $[\widehat{f}_1(t), \widehat{f}_2(t), \dots, \widehat{f}_q(t)]$ and $P \times Q$ weights $[W_{11}(t), W_{21}(t), \dots, W_{p1}(t), \dots, W_{1q}(t), W_{2q}(t), \dots, W_{pq}(t)]$. For notation simplification we wrote each node as $\widehat{f}_i(t) = f_i(t) + \sigma_{f_i}\epsilon(t)$ and dropped the dependence on t in the diagram.

lytically as in a standard GP, and thus an approximation is required. To overcome this issue I made use of variational inference (Jordan et al., 1999).

For this it is used a family of factorized distributions $q(\mathbf{f}, \mathbf{w})$ to obtain the closest approximating distribution to the posterior $p(\mathbf{f}, \mathbf{w}|D)$ defined by a given set of outputs D . This approximation is measured with the Kullback-Leiber (KL) divergence (Kullback, 1959),

$$KL(q(\mathbf{f}, \mathbf{w})||p(\mathbf{f}, \mathbf{w}|D)) = E \left[\log \frac{q(\mathbf{f}, \mathbf{w})}{p(\mathbf{f}, \mathbf{w}|D)} \right]. \quad (3.34)$$

With this measure we are able of transforming an inference problem into an optimization problem (Blei et al., 2017).

To obtain the factorized distributions $q(\mathbf{f}, \mathbf{w})$ we use mean-field inference for a GPRN developed by Nguyen and Bonilla (2013). In mean-field inference we have the family of factorized distribution

$$q(\mathbf{f}, \mathbf{w}) = \prod_{q=1}^Q q(f_q) \prod_{p=1}^P q(w_{pq}) \quad (3.35)$$

where our distributions $q(f_q) = \mathcal{N}(\mu_{f_q}, \Sigma_{f_q})$ and $q(w_{pq}) = \mathcal{N}(\mu_{w_{pq}}, \Sigma_{w_{pq}})$.

For each distribution it is then required to determine its parameters μ and Σ . For the nodes the parameters are obtained with

$$\begin{aligned}\Sigma_{f_q} &= \left[K_{f_q}^{-1} + \frac{1}{\sigma} \sum_{p=1}^P \text{diag} \left(\mu_{w_{pq}} \circ \mu_{w_{pq}} + \text{diag}(\Sigma_{w_{pq}}) \right) \right]^{-1}, \\ \mu_{f_p} &= \frac{1}{\sigma} \Sigma_{f_q} \sum_{p=1}^P \sum_{q=1}^Q \left(\mathbf{y}_p - \sum_{k \neq q}^Q \mu_{w_{pk}} \circ \mu_{f_k} \right) \circ \mu_{w_{pq}},\end{aligned}\quad (3.36)$$

while for the weight they are given by

$$\begin{aligned}\Sigma_{w_{pq}} &= \left[K_{w_{pq}}^{-1} + \frac{1}{\sigma} \sum_{p=1}^P \text{diag} \left(\mu_{f_q} \circ \mu_{f_q} + \text{diag}(\Sigma_{f_q}) \right) \right]^{-1}, \\ \mu_{w_{pq}} &= \frac{1}{\sigma} \Sigma_{w_{pq}} \sum_{p=1}^P \sum_{q=1}^Q \left(\mathbf{y}_p - \sum_{k \neq q}^Q \mu_{f_k} \circ \mu_{w_{pk}} \right) \circ \mu_{f_q}.\end{aligned}\quad (3.37)$$

On these equations we have $\sigma = \sum_{p=1}^P \sum_{n=1}^N \left(\sigma_{yerr_{np}}^2 + \sigma_{y_p}^2 \right)$ as the sum of all measurements errors σ_{yerr} and jitters σ_y . y_p is the N-dimensional vector with the measurements of dataset p , K is the covariance matrix of the respective node or weight. The $\text{diag}()$ function turns the diagonal elements of a matrix in a vector, or a vector into a diagonal matrix. For last we also have \circ representing the Hadamard product (Styan, 1973).

Our final approximation is measured by minimizing the KL divergence. This minimization is equivalent to maximize the evidence lower bound (ELBO) (Blei et al., 2017). For the GPRN Nguyen and Bonilla (2013) defined the ELBO as

$$ELBO(q) = E_q [\log p(D|\mathbf{f}, \mathbf{w})] + E_q [\log p(\mathbf{f}, \mathbf{w})] + \mathcal{H}_q [q(\mathbf{f}, \mathbf{w})], \quad (3.38)$$

where the first term is known as the expected log-likelihood, the second the expected log-prior and the last is known as entropy. To finalize we need the definition of these three terms. Their deduction can be seen in Nguyen (2015).

The expected log-likelihood is given by

$$\begin{aligned}
 E_q [\log p(D|\mathbf{f}, \mathbf{w})] = & -\frac{1}{2} \sum_{n=1}^N \sum_{p=1}^P \left[\log 2\pi \left(\sigma_{y_{err,np}}^2 + \sigma_{y_p}^2 \right) \right] \\
 & - \frac{1}{2\sigma^2} \sum_{n=1}^N \left(\mathbf{y}_n^T - \mathcal{W}_{w_n} \mathcal{F}_{f_n} \right)^T \left(\mathbf{y}_n^T - \mathcal{W}_{w_n} \mathcal{F}_{f_n} \right) \\
 & - \frac{1}{2\sigma^2} \sum_{p=1}^P \sum_{q=1}^Q \left[\text{diag} \left(\Sigma_{f_q} \right)^T \left(\mu_{w_{pq}} \circ \mu_{w_{pq}} \right) \right. \\
 & \quad \left. + \text{diag} \left(\Sigma_{w_{pq}} \right)^T \left(\mu_{f_q} \circ \mu_{f_q} \right) + \text{Tr} \left(\Sigma_{f_q} \Sigma_{w_{pq}} \right), \right]
 \end{aligned} \tag{3.39}$$

where the $\text{Tr}()$ function returns the trace of the matrix. \mathbf{y}_n^T is a $1 \times P$ vector containing all observations at entry n , \mathcal{W}_{w_n} is a $P \times Q$ matrix containing the variational means μ_w at entry n , and \mathcal{F}_{f_n} is a $P \times 1$ vector containing the variational means μ_f at entry n .

The expected log-prior is given by

$$\begin{aligned}
 E_q [\log p(\mathbf{f}, \mathbf{w})] = & -\frac{1}{2} NQ (P+1) \log 2\pi \\
 & - \frac{1}{2} \sum_{q=1}^Q \left[\log |K_{f_q}| + \mu_{f_q}^T K_{f_q}^{-1} \mu_{f_q} + \text{Tr} \left(K_{f_q}^{-1} \Sigma_{f_q} \right) \right] \\
 & - \frac{1}{2} \sum_{p=1}^P \sum_{q=1}^Q \left[\log |K_{w_{pq}}| + \mu_{w_{pq}}^T K_{w_{pq}}^{-1} \mu_{w_{pq}} + \text{Tr} \left(K_{w_{pq}}^{-1} \Sigma_{w_{pq}} \right) \right],
 \end{aligned} \tag{3.40}$$

where K is the covariance matrix of the respective node or weight.

For last we have the entropy that is given by

$$\begin{aligned}
 \mathcal{H}_q [q(\mathbf{f}, \mathbf{w})] = & \frac{1}{2} NQ (P+1) (1 + \log 2\pi) \\
 & + \frac{1}{2} \sum_{q=1}^Q \left[\log |\Sigma_{f_q}| \right] + \frac{1}{2} \sum_{p=1}^P \sum_{q=1}^Q \left[\log |\Sigma_{w_{pq}}| \right]
 \end{aligned} \tag{3.41}$$

3.3.2 Simulating a quasi-periodic GP

A GPRN is a construct of several GPs. This condition can make us question if, under certain conditions, we can mimic a standard GP. Or, in the case of RV analysis, we can ask if a GPRN can simulate the traditional quasi-periodic GP.

For simplification, let us consider the analysis of a single time series ($P = 1$). If we were modelling RV measurements, in the traditional GP, we would use a quasi-periodic kernel given by 3.20. To try and simulate that, my GPRN will only use one node ($Q = 1$) and, as consequence, a single weight ($P \times Q = 1$). As I mentioned earlier, the product of two GPs does not necessarily create a GP, but we can still create samples of a GPRN and from a GP and compare them.

The first GPRN setup relies on a node with a GP using a periodic kernel, and a weight defined by a GP using a squared exponential kernel. On both GP and the GPRN, the parameter will be the same. That is, I have defined $\eta_1 = 1$ m/s, $\eta_2 = 50$ days, $\eta_3 = 20$ days, and $\eta_4 = 0.75$.

The second setup uses a node with a quasi-periodic kernel with the same parameters as the GP's kernel. The weight, although a squared exponential kernel, will now have an $\eta_1 = 1$ m/s and an $\eta_2 = 10000$ days. Such large η_2 will make the squared exponential behave as a constant. The main idea is to multiply the quasi-periodic node by a "constant" weight¹¹.

For each, I obtained three samples¹², see figure 3.8. When obtaining samples for the GPRN with a periodic node, the main visible feature is a stable periodic signal. This meaning that its structure for a given period does not change, only the amplitude increases or decreases accordingly to the variation of the weight. This is expected, considering that, on a GPRN, the final amplitudes are given by the weights $W(t)$. As such this weight should not change the periodic features of the node. That creates a framework that is not flexible to model changes within a given period.

¹¹ One could think of a simpler setup and use constant kernels in the weights. Indeed Wilson (2014) mentions this possibility. Unfortunately, the mean-field inference algorithms require the inversion of the covariance matrix of the weights. That makes impossible the use of constant kernels.

¹² I only plotted three samples for easier visualization. More could be easily done and plotted.

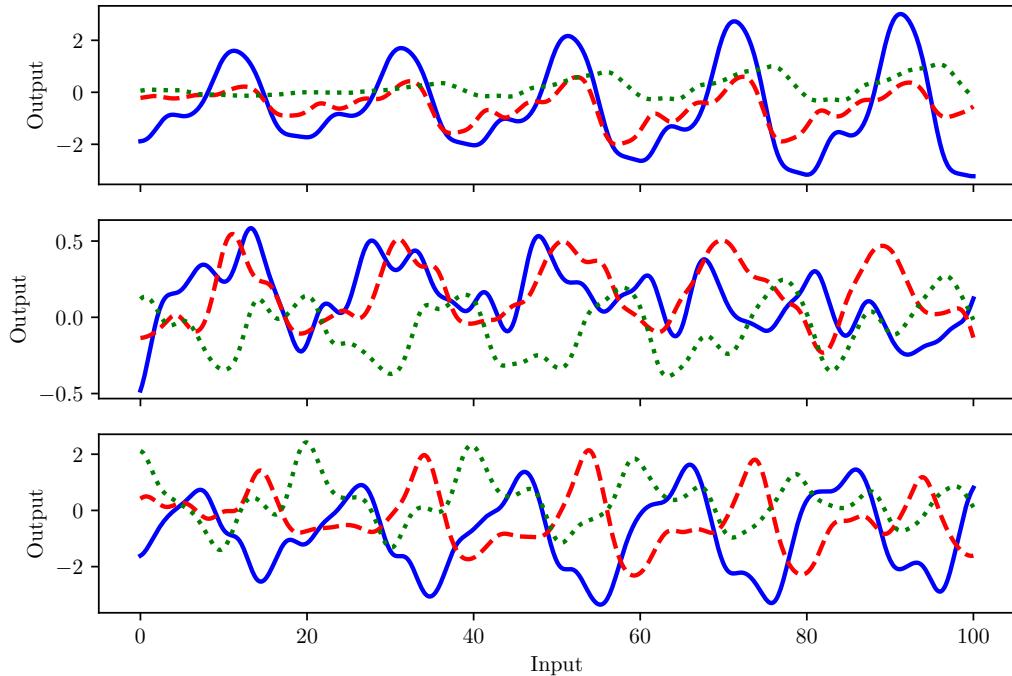


Figure 3.8: Top: Samples of a GPRN using a node with periodic kernel and weight with a squared exponential kernel. Middle: Samples of a GPRN using a node with a quasi-periodic kernel and a weight with a squared exponential kernel. Bottom: Samples of a GP defined by a quasi-periodic kernel.

The second GPRN setup shows a more quasi-periodic behaviour. Again not surprisingly, considering I was using a quasi-periodic kernel in the node. It is also apparent that the weight shows a constant behaviour as the amplitude of the samples does not change drastically, unlike the first GPRN setup. These setup samples, when compared with the samples of a quasi-periodic GP, have similar behaviour. As such, in my tests of chapter 4, I will compare the result obtained by a quasi-periodic GP with one obtained using a GPRN with a quasi-periodic node. That does not mean other, more useful, GPRN configurations may not exist.

3.3.3 Changing the weights

In this thesis I will focus on using weights given by a GP with a squared exponential kernel. As such, understanding how this weight will influence the nodes of a GPRN is very important. To show how the paper of the weights in the GPRN, let us look at it with a quasi-periodic node and a squared exponential weight.

Changing η_1

For a weight using a squared exponential kernel, the first parameter to have a look at is, of course, the amplitude η_1 . That is the easiest parameter to interpret as its impact is proportional to its value. If the amplitude of the weight increases or decreases, so does the overall amplitude of the GPRN. For example, if the GPRN node has a $\eta_1 = 2$ m/s a weight with $\eta_1 = 5$ m/s, the GPRN amplitude will approximately be the product of these, that is 5 m/s. This excluding, of course, the impact of the other parameter of the weight.

Changing η_2

Next we have the impact of η_2 on the overall structure of the GPRN. To check this I have, similar to subsection 3.3.2, defined a node with $\eta_1 = 1$ m/s, $\eta_2 = 50$ days, $\eta_3 = 20$ days, and $\eta_4 = 0.75$. The weight had $\eta_1 = 1$ m/s, while I gave three different values for η_2 .

In figure 3.9 we can see the influence the η_2 of the weight has in the overall structure. For an easier interpretation, I used the same node sample on the three cases. There are clear characteristics we can see it is influenced by the weight. The less smooth the squared exponential, in the example the weight with $\eta_2 = 10$ days, the less similar the GPRN output becomes from the node output. The output sign for the GPRN depends on both weight and node signs, meaning equal signs return a positive output while different returns a negative output. The GPRN amplitude is also quite sensitive to the weight value as it decreases as the weight output tends to zero, meaning the GPRN structure disappears.

It is clear also that the higher the value of the η_2 of the weight, the more

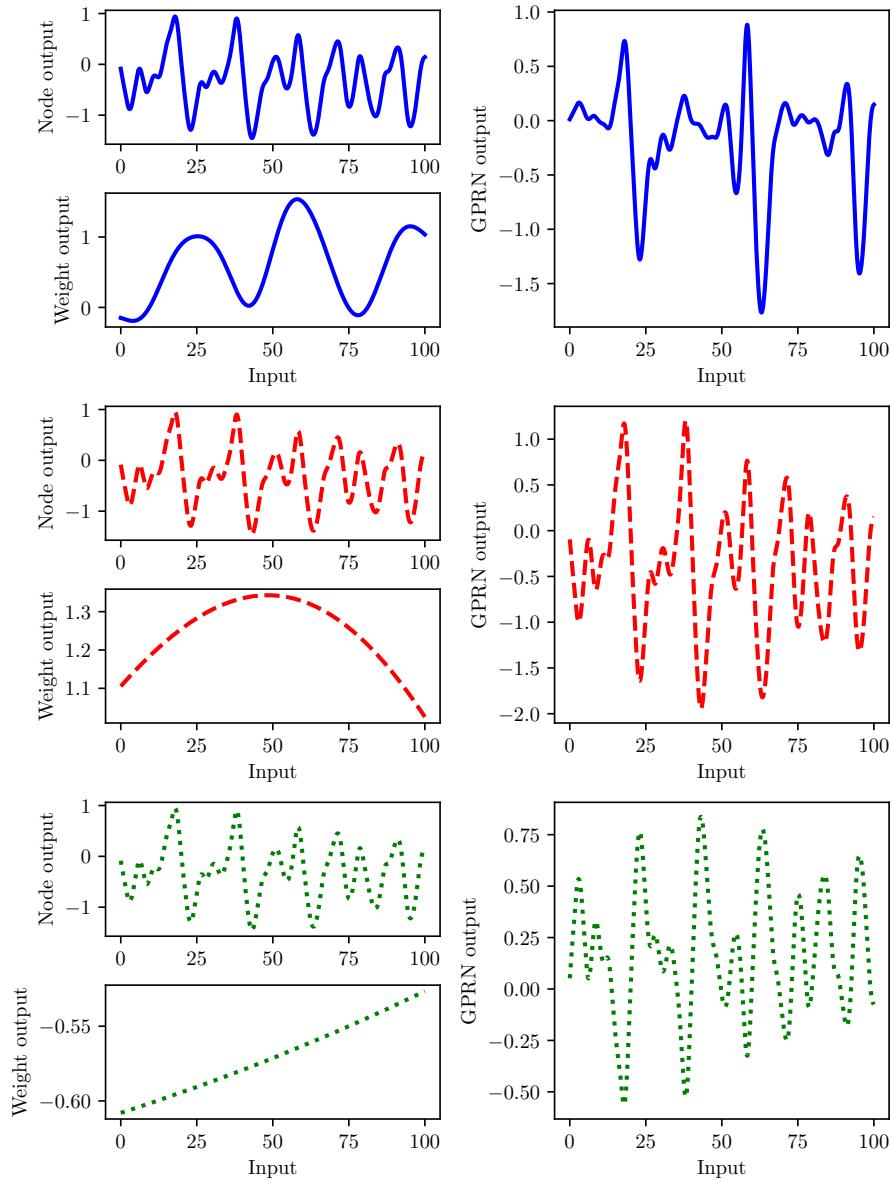


Figure 3.9: Samples of a GPRN using a node with a quasi-periodic kernel and a weight with a squared exponential kernel. The only different parameter being the η_2 of the weights. Top: Solid blue a GPRN using a weight with $\eta_2 = 10$ days. Middle: Dashed red a GPRN using a weight with $\eta_2 = 100$ days. Bottom: Dotted green a GPRN using a weight with $\eta_2 = 1000$ days.

constant its GP becomes, as expected. Curious are the weights signs for $\eta_2 = 100$ days and $\eta_2 = 1000$ days. The samples having opposite sign have a very similar structure but with opposite concavity. This indicates that not only the value of η_2 will determine the shape of the GPRN the respective sign as well.

Chapter 4

Application on radial velocity observations

In this chapter, I show the results I've obtained with a GPRN. I start by showing its performance on radial velocity observations of the Sun. I not only use these observations as the test case on our GPRN. I test its performance and main results that might lead to future improvements of this framework.

Also in this chapter, I present the performance of the GPRN on four stars from the EXPRES Stellar-Signals Project. Following our results with the Sun, these four other stars are ideal to measure the performance of the GPRN on stars with different activity levels.

4.1 Priors definition

I start by presenting by priors I used on the MCMC used to obtain my results. They were defined thinking on using as much information as possible from the time series, e.g. using its peak-to-peak amplitude (PTP) as a limit. I avoided, whenever possible undefined or unexplained values as prior parameters. I also defined the priors to allow that the same kernel, independently if it is on a GP or GPRN, uses the same prior. For example, all the amplitudes η_1 uses the same prior. The parameters notation will be the same as the one shown in chapter 3.1.3.

Variables	Parameter (units)	Prior
η_1	Amplitude (m/s)	$\mathcal{MLU}(y_\sigma, 2 \times y_{\text{PTP}})$
η_2	Decaying timespan (days)	$\mathcal{LU}(t_{AV}, 10 \times t_{\text{PTP}})$
η_3	Period (days)	$\mathcal{U}(10, 50)$
η_4	Length scale	$\mathcal{LU}(0.1, 5)$
s	White noise amplitude (m/s)	$\mathcal{MLU}(y_\sigma, 2 \times y_{\text{PTP}})$
slope	Slope of the mean function	$\mathcal{N}\left(0, \frac{y_\sigma}{t_\sigma}\right)$
offset	Offset of the mean function (m/s)	$\mathcal{U}(y_{\min}, y_{\max})$

Table 4.1: Parameters and prior distribution of the parameters in the individual analysis of both GP and GPRN. The y represents the input while t the time. On the individual analysis the input could either be the RVs, FWHM, BIS, or $\log R'_{hk}$.

To simplify the reading of the notation I used, I divided it into two. First, the priors used on the individual analysis of the time series. Succeeded with the combined analysis of RV plus activity indicator.

4.1.1 Individual analysis

For each of these analyses, let us set y as the measurements and t as the time. Each prior is defined on table 4.1. On it, σ stands for standard deviation. PTP is the peak-to-peak amplitude for y or the timespan for t . \min and \max stand for the minimum and maximum values of the measurements, respectively. For last, t_{AV} is the average time between consecutive observations. That value is 2.205 days.

As for the distributions, I used a modified log-uniform (\mathcal{MLU}) on the amplitudes¹. This allows the support of the distribution to include zero. On η_2 and η_4 (the time scales) I choose a log-uniform distribution (\mathcal{LU}) instead². For the η_3 I defined a uniform distribution (\mathcal{U}) between 10 and 50, acceptable margins for the rotation period of a Sun-like star. The white noise (or jitter) s used the same

¹ This distribution is also known as the *modified Jeffreys prior*.

² This distribution is also known as the *Jeffreys prior*.

prior as η_1 .

Important to mention is that the value of η_2 was constrained. This parameter being closely related to the evolution of the active regions on the surface of the stars have been suggested to be constrained to values bigger than η_3 (e.g. Kosiarek and Crossfield, 2020). When in a node, I constrain its value to be higher than $0.5\eta_3$. When in a weight I constrain it to be higher than the η_2 of the node ($0.5\eta_3 < \eta_{2_{\text{node}}} < \eta_{2_{\text{weight}}}$). For the GP analysis I simply constrained its η_2 to be higher than $0.5\eta_3$.

For last, I also consider a linear function as the mean function on each time series. This function is defined using a slope and an offset. The slope used a normal distribution (\mathcal{N}) with mean zero and variance defined as the ratio of the standard deviations of the measurements and the time. The offset used a uniform distribution between the minimum and maximum values to the given input.

4.1.2 Combined analysis

The priors for the analysis of RV plus an activity indicator had similar priors (see table 4.2). One difference is, when considering a GPRN, we now have one node and two weights. The distributions are the same as before, with each weight amplitude using its respective input to define the parameters of the distributions. The parameters of the white noise s are defined similarly to the η_1 of the weights. There is a white noise per time series, whose distribution parameters will depend on its respective time series.

The η_1 of the node has a slight difference. To account both RV and indicator, the knee or the point where the distribution changes from uniform to log-uniform is defined as the lowest standard deviation of the two time series that I represent as $[y_{RV}, y_{AI}]_{\sigma}^{\min}$. The upper bound of the distribution is specified to be twice the highest of the two peak-to-peak amplitudes obtained from the RVs and the indicator. I represent this value by $[y_{RV}, y_{AI}]_{\text{PTP}}^{\max}$.

On the GP analysis, the framework used is created by combining the kernels on the nodes and the weights. That implies that, for each time series, the final amplitude is the product of the node and weight amplitudes. That allows us to simplify the parameter space and fix one amplitude to one while the other is free.

Variables	Parameter (units)	Prior
node η_1	Amplitude (m/s)	$\mathcal{MLU}([y_{RV}, y_{AI}]_{\sigma}^{\min}, 2 \times [y_{RV}, y_{AI}]_{\text{PTP}}^{\max})$
node η_2	Decaying timespan (days)	$\mathcal{LU}(t_{AV}, 10 \times t_{\text{PTP}})$
node η_3	Period (days)	$\mathcal{U}(10, 50)$
node η_4	Length scale	$\mathcal{LU}(0.1, 5)$
weight η_1	Amplitude (m/s)	$\mathcal{MLU}(y_{\sigma}, 2 \times y_{\text{PTP}})$
weight η_2	Decaying timespan (days)	$\mathcal{LU}(t_{AV}, 10 \times t_{\text{PTP}})$
s	White noise amplitude (m/s)	$\mathcal{MLU}(y_{\sigma}, 2 \times y_{\text{PTP}})$
slope	Slope of the mean function	$\mathcal{N}\left(0, \frac{y_{\sigma}}{t_{\sigma}}\right)$
offset	Offset of the mean function (m/s)	$\mathcal{U}(y_{\min}, y_{\max})$

Table 4.2: Parameters and prior distribution of the parameters when considering the analysis of the RVs with an activity indicator. y_{RV} represents the RV time series while y_{AI} the activity indicator (FWHM, BIS, or $\log R'_{hk}$) time series, as previously t represents the time.

In this case, I set the η_1 of the node, while the weights being constants only had a η_1 with a prior distribution equal to the weight η_1 of table 4.2. The remaining parameters of the quasi-periodic kernel used the corresponding distributions to the ones defined for the node on table 4.2.

The means are again linear functions. Again, there is one mean function per time series. The distributions are identical to the ones used on the individual analysis. With the parameters depending on the time series, the mean function is related.

4.2 Solar data

The Sun is the nearest star to us and consequently been studied and monitored over millennia. That enabled a determination of its parameters with very high precision. Based on its spectral class, it is a G2V-type main-sequence star. Its interior is characterized by a radiative zone from up to 0.7 of its radius covered by a convective shell (Del Zanna and Mason, 2013). It is in this convective region that the magnetic activity responsible for the stellar signals contaminating our observations are produced by a physical process known as solar dynamo (e.g. Pagano, 2013; Charbonneau, 2014). This process creates the well know 11-year sunspot cycle, in which, for example, the number of spots, flares and coronal mass ejections varies, and with it, the respective magnetic activity of Sun (for a review on the solar cycle see Hathaway, 2015).

4.2.1 Sun observations

The main results I have obtained for the GPRN used the three years of solar RV, FWHM and BIS measurements taken by Dumusque et al. (2015) and publicly available at (Dumusque et al., 2021). They used a solar telescope fibre-fed into the HARPS-N spectrometer (Cosentino et al., 2012). The available observations were executed between 2015 and 2018. That happened during the end of the solar cycle 24, with the Sun approaching its less active period. With observations being done during this minimum activity period, it is fair to expect less prominent signal contamination due to spots or flares. The observations were corrected of any signal due to solar system dynamics, including planetary signals (Collier Cameron et al., 2019).

We binned our data to remove any oscillations and granulation signals (Dumusque, X. et al., 2011a). For each day, if available, I binned three consecutive measurements if not, I disregarded the day. The final error I calculated by the quadratic sum of the picked measurements. In total, we obtained 497 measurements for each time series, spanning 1094 days (see figure 4.1). A small summary of these time series exist on appendix A.

The final data set obtained offered an ideal test subject. Each measurement

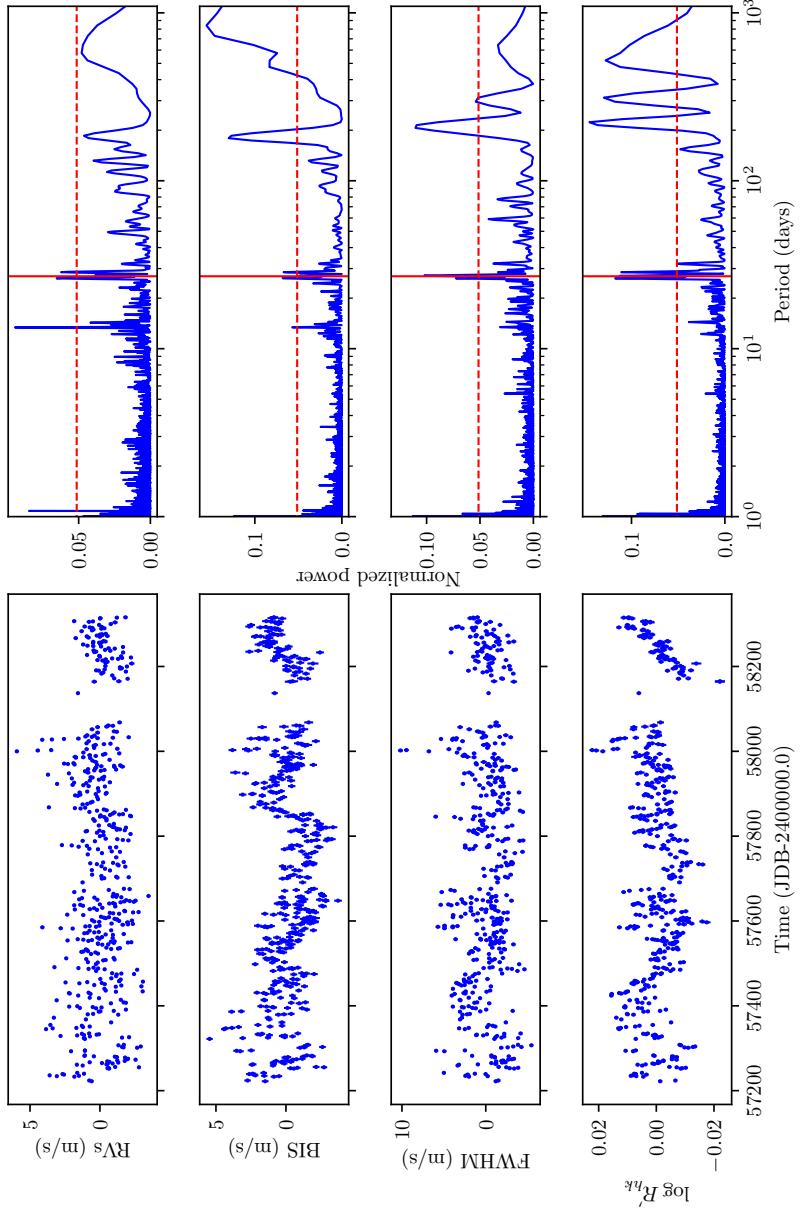


Figure 4.1: Left: Binned RV, BIS, FWHM, and $\log R'_{hk}$ measurements of the Sun obtained with the HARPS-N solar telescope used in our analysis. Right: GLS periodograms of the binned measurements after removing a long-term trend to the data. The vertical solid lines indicate the period of 27 days. The horizontal dashed line show the false alarm probability (FAP) of 1%. These measurements are available at <https://github.com/jdavidrcamacho/datasets>.

simulated a usual RV 15-minutes exposure spectroscopic observation. Without any planetary signal, one can determine the effects of the stellar activity on RV observations and how to best use the activity indicators to improve planetary signal detection.

Periodogram analysis

Before using a GPRN framework on the Sun data, we can look at the periodogram of the time series. I computed the generalized Lomb–Scargle (GLS) periodograms as it is implemented on Astropy (Astropy Collaboration et al., 2018) for the RV, FWHM, BIS, and the $\log R'_{hk}$ (figure 4.1). I only did this with the intent of helping me to interpret any periodic signals in the future results. I will not use the periodogram information to define any priors I will express in section 4.1 nor limit any parameter search space.

All the time series show a peak in power around the 27 days synodic rotation period marked by the red vertical line (Wilcox, 1972). The RV time series has a collection of peaks around the 27 days mark higher than the 1% false alarm probability (FAP). This multiple peak "behaviour" is also present on the activity indicators. This indicates that these peaks are indeed activity related. If they are the result of, for example, spots on different latitudes and/or differential rotation is a subject I will not approach on this thesis. The RV data has an even more significant peak on the second harmonic, around 13.5 days. It is the only time series with this behaviour.

On the activity indicators part, the BIS is the proxy with the periodogram structure closer to what is observed on the RV. This indicator also has a peak around 13.5 days higher than the 1% FAP. This peak however is not more significant to the ones present around the rotation period. The second harmonic is either not present or has no significant power for the remaining indicators. Both FWHM and $\log R'_{hk}$ show a collection of peak around the rotation period significantly higher than the 1% FAP level.

Considering the observations occurred with the Sun approaching its less active phase these are interesting results. Although approaching the minima of the magnetic cycle it was not the minima itself. This is seen by the daily sunspots

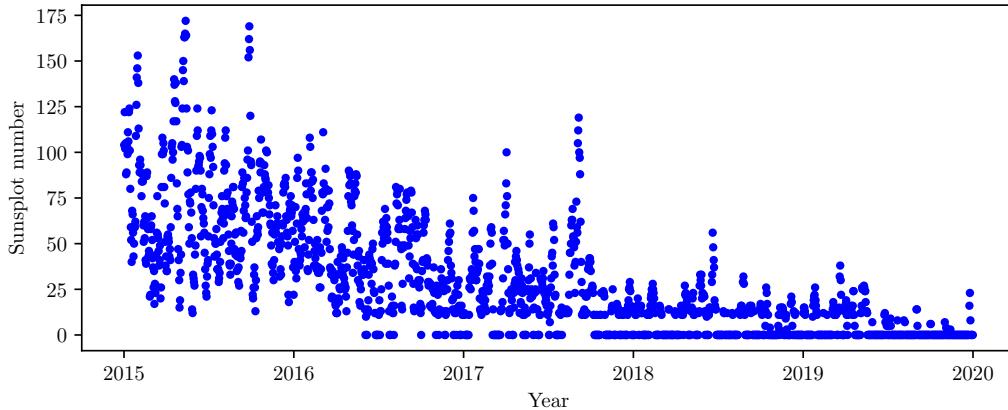


Figure 4.2: Daily number of sunspot observed from the start of 2015 till the end of Solar cycle 24 on late 2019 (e.g. McIntosh et al., 2014; Nandy et al., 2020). Source: <https://wwwbis.sidc.be/silso>

number recorded of figure 4.2. It shows a very considerable number of spots present on the surface of the Sun during the observations. It can also be seen a clear decrease in the number of spots as the Sun approached the cycle's minima. The number of days with spots however is significantly higher than the days without. From the start of 2015 till the end of 2018, there were 1130 days with spots against 331 days without.

Outside the rotation period and harmonics "region", all the time series show a strong signal around the 200 days mark. Although with a significant power on the indicators, this signal should not interfere with the rotation period estimation. The RV also shows some structure below the two-days mark that is most likely an artefact of the sampling.

The periodograms structure on both RV and indicators clearly signals a periodic signal of around 27 days. Based on this, it is of no surprise the periodograms of all the time-series to show a prominent peak on the rotation period. With this in mind I would not be surprised if both GPs and GPRN could converge to a solution close to the rotation value. The only major unknown is if the 13.5 days harmonic will interfere in some cases.

Time lags

Temporal shift for the Sun measurements between the RV and the activity proxies were reported by Collier Cameron et al. (2019). In their work they determined a lag of 1 and 3 days between the RV and the FWHM and BIS, respectively.

The observations we used have some significant differences to the ones presented there. The original data we used use different data reduction software. To correct for the instrumental systematics, Dumusque et al. (2021) adapted the ESPRESSO data reduction software to HARPS-N. That allowed them to obtain more precise RV measurements. Besides that, we binned the available measurements to simulate a 15-minutes exposure spectroscopic observation.

With that in mind, obtaining the exact same cross-correlation functions is unlikely. However, if they exist, similar time lags should be observable in our time series. Using the same discrete correlation function developed by Edelson and Krolik (1988) we thus calculated our own CCFs. The results are showed on figure 4.3. In our observations the RV indeed lead the FWHM by 1 to 2 days. This time shift increased to 2 to 3 days when comparing RV and BIS. For last, the chromospheric indicator $\log R'_{hk}$ shows no lag between it and the RV.

The correlation of figure 4.3 showed the FWHM and the $\log R'_{hk}$ to have a long period signal. Signal that however was not present on the BIS. Removing it with a linear trend did not change the location of the shifts, but decreased the correlation signal to what was similar observed by Collier Cameron et al. (2019).

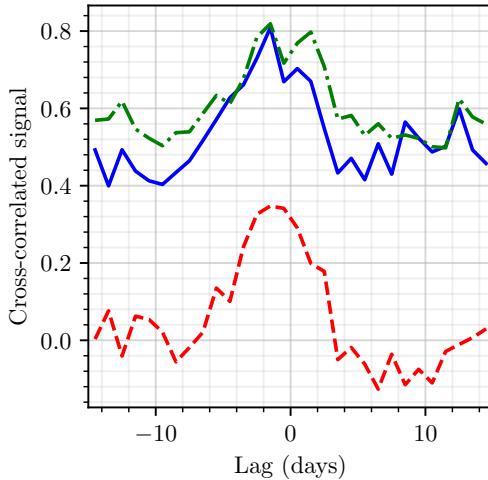


Figure 4.3: Discrete cross-correlation functions between the RV and the BIS (solid blue), FWHM (dashed red), and $\log R'_{hk}$ (dash-dotted green).

4.2.2 Results obtained on the solar data

To understand the capabilities of the GPRN, I will always compare it with a similar GP. I implemented the GPRN presented on chapter 3.3 on a python package currently available on Github³. To analyse the Sun data, the GPRN was defined by one node and, as consequence, one weight per time series. The node is a quasi-periodic GP, with each weight being a squared exponential GP.

When using a GP for the individual analysis, I used a package also developed called `tedi`⁴. Each time series analysed used the quasi-periodic kernel of 3.20 plus a white noise term. The combined GP analysis of the RV measurements plus an activity indicator used the artGPN framework I mentioned in 3.2.3. Again, only with one node and one weight per time series. Unlike what is allowed on the GPRN, I used a quasi-periodic kernel on the node and constant kernels for the weights.

All these experiment were done together with `emcee` (Foreman-Mackey et

³ <https://github.com/jdavidrcamacho/gprn>

⁴ <https://github.com/jdavidrcamacho/tedi>

al., 2019). That allowed the creation of Markov chains Monte Carlo (MCMC) simulations to sample the parameter space and obtain the posterior distributions. The MCMC was left running until the convergence criteria of having a sample size of, at least, 25 times the integrated autocorrelation time (τ), was achieved. This criterion defines τ as the number of steps needed for a chain that forgets where it began or reached its equilibrium (Sokal, 1997). Every 5000 iterations, this criterion was evaluated to determine the convergence of the MCMC. I then used 2τ as the number of burn-in iterations of the MCMC. As a metric to determine the best model I used the Bayesian evidence $\log \mathcal{Z}$. I also considered how much the model diminished the root mean squared (RMS) of the residuals to compare the different fits.

In the following pages, I will present the GP and GPRN results obtained while analysing the solar data. I will describe the obtained results, one by one, accordingly to the time series used. For each model, I will present the *maximum a posteriori* (MAP) as well as the 50% quantile (median⁵) estimates. I will also use the 16% and 84% quantiles values to calculate the error bars. The complete corner plots obtained are on an appendix, located in appendix B. I will start by an interpretation of the results obtained and leave a more careful examination of the behaviour of the GPRN and GPs for section 4.2.3.

Analysis of the RV time series

The first results I obtained were on the RV data alone. Both GP and GPRN frameworks converged to similar results. These are shown in table 4.3. On both models, we have parameters consistent within a 1-sigma interval. It is noteworthy to mention the obtained rotation period η_3 . As previously seen, the periodogram of these measurements gave significant power to the 13-days harmonic. Our analysis got no samples for the 13-days harmonic. That meant the two models were capable of obtaining values close to the expected rotation period instead. Both the GP and the node of the GPRN have similar a decaying timescale (η_2). Another interesting result is the MAP values of the η_4 parameter. While the values are consistent within a 1-sigma interval, the higher MAP value for the GP indicates its fit (figure 4.4)

⁵ I use the median value due to, unlike the mean, be more robust against outliers and skewed distributions.

Framework	Variables	MAP value	Median value
GP	η_1	1.145	$1.179^{+0.088}_{-0.081}$
	η_2	20.374	$20.342^{+2.149}_{-2.042}$
	η_3	26.031	$26.239^{+0.618}_{-0.562}$
	η_4	0.625	$0.632^{+0.073}_{-0.063}$
	s	0.797	$0.796^{+0.040}_{-0.039}$
	slope	-0.005	$-0.005^{+4 \times 10^{-4}}_{-4 \times 10^{-4}}$
	offset	-19.124	$-19.117^{+0.114}_{-0.148}$
GPRN	node η_1	0.720	$1.412^{+4.977}_{-1.097}$
	node η_2	19.972	$20.258^{+2.015}_{-1.931}$
	node η_3	25.921	$26.188^{+0.657}_{-0.559}$
	node η_4	0.584	$0.610^{+0.068}_{-0.063}$
	weight η_1	1.038	$1.556^{+5.178}_{-1.220}$
	weight η_2	1001.229	$3569.232^{+4025.923}_{-2123.239}$
	s	0.781	$0.793^{+0.040}_{-0.040}$
	slope	-0.005	$-0.005^{+4 \times 10^{-4}}_{-4 \times 10^{-4}}$
	offset	-19.032	$-19.131^{+0.142}_{-0.140}$

Table 4.3: Parameters obtained by the GP and the GPRN on the analysis of the RV time series.

has a periodicity with a slightly smoother structure.

With both models obtaining similar parameters, the similar fit of figure 4.4 is not surprising. Even so, the evidence gives a slight advantage to the GP. This model obtained a $\log \mathcal{Z}$ of -789.106 ± 0.075 , while the GPRN obtained a $\log \mathcal{Z}$ equal to -789.460 ± 0.307 . This evidence is supported by the achieved RMS reduction. Of the two, the GP decreased it slightly more, reducing the RMS by 3.075 versus a reduction of 2.876 on the GPRN.

The fits and residuals are presented in figure 4.4. Visually, it can be observed that the residuals of the GPRN tend to be worse on measurements that deviate more from the rest of the observations. For example, on the 58000 days mark. On the

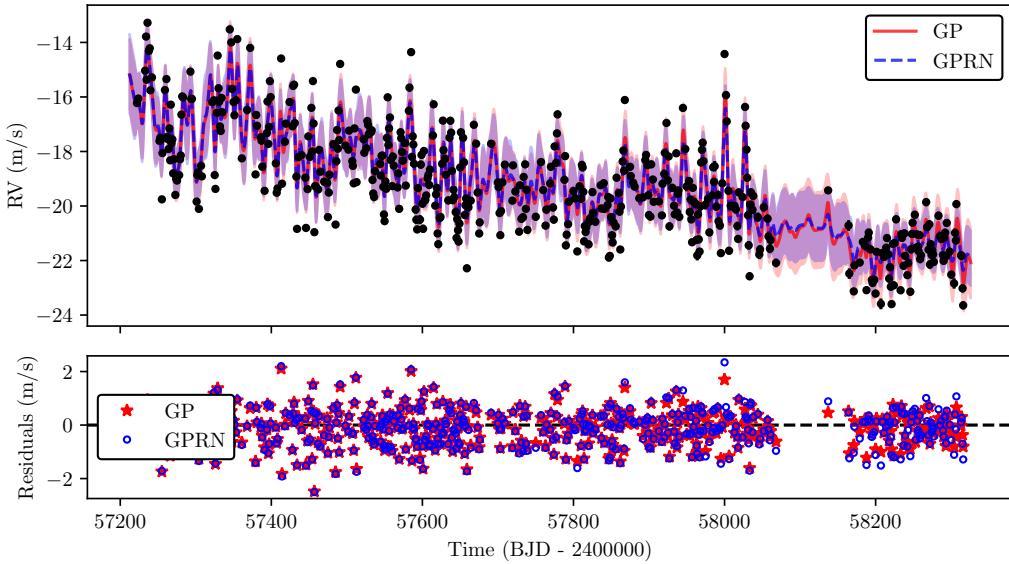


Figure 4.4: Top: Superimposed fits of MAP values for the GP (solid red) and the GPRN (dashed blue) on the RV measurements. Bottom: Residuals obtained of the GP (red stars) and the GPRN (blue circles).

measurement that corresponds to an RV of almost 14 m/s. There is a clear difference between the GP and the GPRN. It seems the GP is more flexible to reach these measurements. Also interesting is the behaviour of the two models following the gap around the 58100 days-mark days-mark. After it, the residuals show the fit of the models become drastically different. I will discuss more this behaviour in section 4.2.3.

The final aspect I want to point out in this analysis is the convergence speed of the models. I saw the GP proved to be considerably faster to achieve the convergence criteria, requiring only 10000 iterations. The GPRN, on the other hand, required 40000 iterations. Considering the mathematical complexity of the GPRN that slows the MCMC used, the need for more iterations to achieve convergence might constrain any future use of a GPRN, as time is always an issue.

Framework	Variables	MAP value	Median value
GP	η_1	1.342	$1.360^{+0.098}_{-0.086}$
	η_2	24.004	$23.739^{+1.584}_{-1.558}$
	η_3	27.064	$27.073^{+0.365}_{-0.360}$
	η_4	0.713	$0.726^{+0.051}_{-0.051}$
	s	0.406	$0.405^{+0.031}_{-0.029}$
	slope	-0.005	$-0.005^{+0.001}_{-0.001}$
	offset	-91.695	$-91.663^{+0.190}_{-0.182}$
GPRN	node η_1	1.568	$1.557^{+4.511}_{-1.203}$
	node η_2	23.235	$23.575^{+1.588}_{-1.570}$
	node η_3	26.853	$26.993^{+0.364}_{-0.381}$
	node η_4	0.724	$0.735^{+0.056}_{-0.050}$
	weight η_1	0.855	$1.492^{+4.804}_{-1.130}$
	weight η_2	2431.229	$4909.158^{+3588.740}_{-2517.306}$
	s	0.414	$0.418^{+0.030}_{-0.030}$
	slope	-0.005	$-0.005^{+0.001}_{-0.001}$
	offset	-91.689	$-91.669^{+0.177}_{-0.173}$

Table 4.4: Parameters obtained by the GP and the GPRN on the analysis of the BIS time series.

Analysis of the BIS time series

Similar to the RV analysis, the results obtained for the BIS observations show two models with similar behaviour. The parameters obtained are in table 4.4. Again these show to be consistent within a 1-sigma interval of each other. The rotation period using the BIS measurements was easily obtained on both models. An interesting result is seen for the two amplitudes of the GPRN. The product of their MAP value is consistent with the GP amplitude. That is unlike with was seen with the RV analysis. Also very similar is also the value of the η_4 parameter. If my interpretation of this parameter is correct, this seems to indicate both models have a similar structure inside the obtained period.

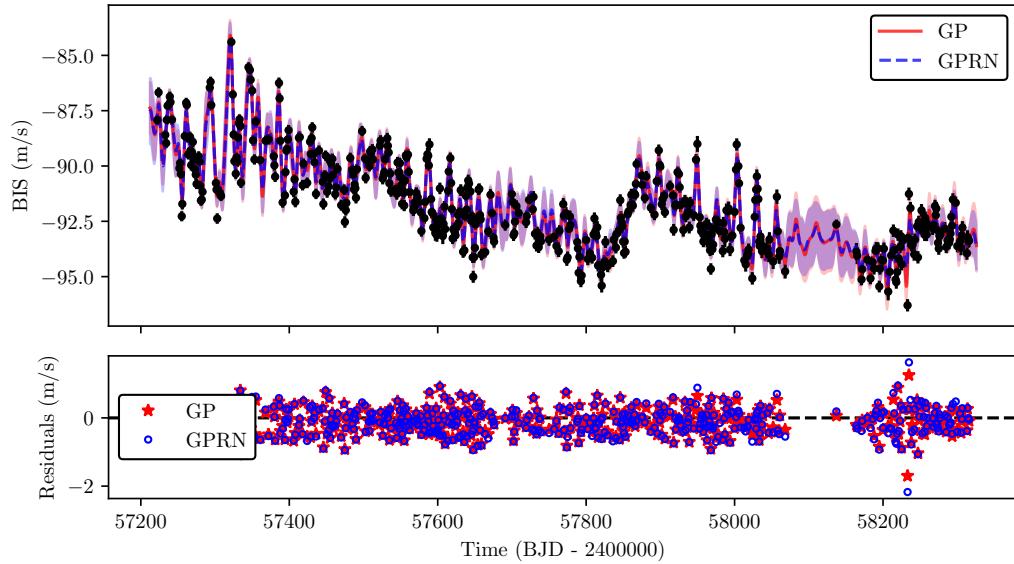


Figure 4.5: Top: Superimposed fits for the MAP values of the GP (solid red) and the GPRN (dashed blue) models on the BIS time series. Bottom: Residuals obtained of the GP (red stars) and the GPRN (blue circles).

The parameters obtained imply, with no surprise, that the fits of the posteriors are very similar. I show it in figure 4.5. Even so, the residuals show two points, on the 58200-day mark, with a noticeable difference. As I mentioned with the RV results, the models show their modelling limits when an observation deviate too much from the remaining ones. These two BIS datapoints on the 58200-day are slightly different from the trend observed on the other measurements. If this implies they are behaving as outliers is an analysis I will not make. But again the GPRN seem to show more difficulties to model them. Again, I will explore this behaviour in section 4.2.3.

The evidence I have calculated shows the GP to be the preferable model. This model obtained a $\log \mathcal{Z}$ of -639.606 ± 0.064 against a $\log \mathcal{Z}$ of -643.512 ± 0.212 for the GPRN. The RMS reduction obtained support this preference. It is also higher on the GP. That model decreased the RMS by 5.355 times, while the GPRN decreased it slightly less, by 4.866 times. Problematic again is the speed of the GPRN. The GPRN model showed to be not only slower per iteration but requiring

more iterations to converge again. The GP only needed 10000 to achieve it against the 25000 iterations required by the GPRN.

Analysis of the FWHM time series

As previously, the GP and GPRN obtained consistent results on the FWHM analysis. I show these in table 4.5. With this time series, both models found a rotation period around the 27-days mark. Interestingly the η_3 MAP values were slightly higher than in the previous two analyses. While the decaying timespan η_2 was lower, obtaining values around the 19 days mark. The η_4 parameter is consistent with what was obtained previously. That indicates a periodic structure similar to the RV and the BIS. That parameter had a higher MAP value for the GPRN node. If anything, this indicates a slightly smoother quasi-periodic fit on the GPRN.

The values obtained on the weight, especially for its η_2 indicate an almost constant weight. Such will not change the quasi-periodic behaviour of the node significantly. I did not mention this earlier, but it was observed something similar on the RV and the BIS. As such, these time series seem to have a very quasi-periodic GPRN.

One difference from the previous results exists in the slope of the mean function. The FWHM slope is closer to zero than the values obtained on the RV and the BIS. Just by observing the data in figure 4.6, it is not surprising. Unlike the previous two analysis, this time-series seem to not have a long term decreasing trend.

The previously mentioned flexibility of the models, or lack thereof, is again observed. The residuals show that the GP is, again, more capable of modelling data points that deviate more. That is observed in the 58000 days mark. In this region, there are two points with an FWHM of around 7795 m/s. These two points have an FWHM higher than any other and are more easily fit by the GP.

When comparing the two models, the evidence again is advantageous for the GP. For it, we obtained a $\log \mathcal{Z}$ of -977.182 ± 0.657 . The GPRN, although far behind, got a $\log \mathcal{Z}$ of $-979.329.606 \pm 0.496$. The evidence difference is, once more, backed by the respective RMS reduction. This meaning, of course, a higher

Framework	Variables	MAP value	Median value
GP	η_1	1.905	$1.947^{+0.142}_{-0.126}$
	η_2	19.449	$19.123^{+1.732}_{-1.728}$
	η_3	27.903	$28.014^{+0.628}_{-0.582}$
	η_4	0.644	$0.720^{+0.091}_{-0.088}$
	s	1.093	$0.720^{+0.091}_{-0.088}$
	slope	-0.001	$-0.001^{+0.001}_{-0.001}$
GPRN	offset	7785.100	$7785.016^{+0.244}_{-0.247}$
	node η_1	0.647	$1.777^{+5.770}_{-1.391}$
	node η_2	19.001	$19.056^{+1.636}_{-1.776}$
	node η_3	28.244	$27.982^{+0.613}_{-0.553}$
	node η_4	0.731	$0.723^{+0.091}_{-0.089}$
	weight η_1	1.898	$1.733^{+5.978}_{-1.351}$
	weight η_2	1812.053	$4662.489^{+3829.389}_{-2618.007}$
	s	1.136	$1.129^{+0.060}_{-0.058}$
	slope	-0.001	$-0.001^{+0.001}_{-0.001}$
	offset	7785.094	$7785.015^{+0.237}_{-0.226}$

Table 4.5: Parameters obtained by the GP and the GPRN on the analysis of the FWHM time series.

RMS reduction on the GP. This model reduced the RMS by 2.528 times, while the GPRN reduced it only by 2.231 times.

A surprise, however, appeared when determining the converge of the models. The GPRN, again slower per iteration, required only 25000 steps to achieve convergence. The GP required 30000 iterations. If we calculate the integrated autocorrelation time, the longer convergence was due to two parameters. The η_2 and the offset of the GP showed more resilience to converge. The τ values for them was 1054 and 1050, respectively. That was almost twice as long as the remaining parameters. The why for this to happen is not observable by their respective posterior distributions as they show a nice Gaussian-like behaviour (figure B.5 of the appendix).

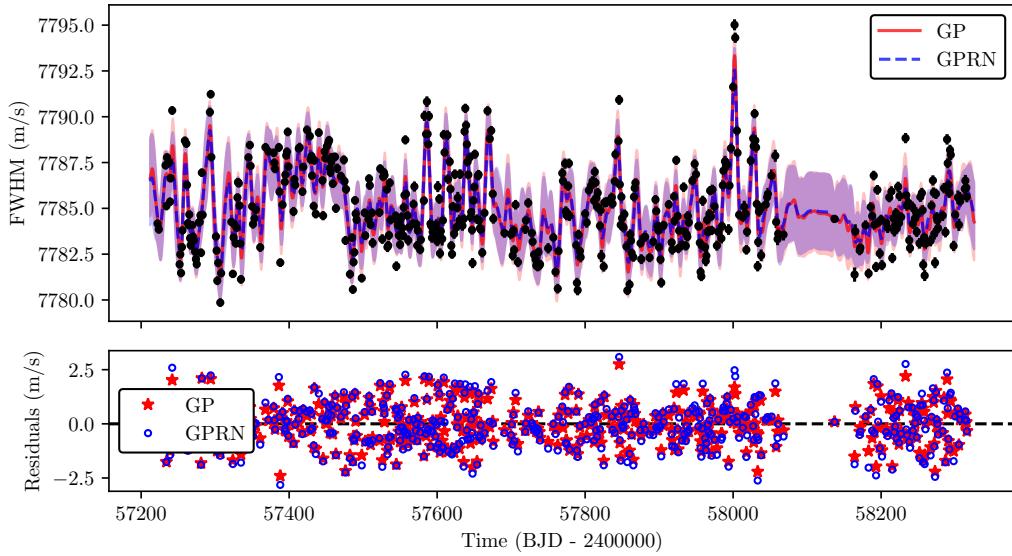


Figure 4.6: Top: Superimposed fits of the MAP values for the GP (solid red) and the GPRN (dashed blue) models on the FWHM time series. Bottom: Residuals obtained of the GP (red stars) and the GPRN (blue circles).

Analysis of the $\log R'_{hk}$ time series

In the last individual analysis I made for the Sun, I used the $\log R'_{hk}$. Once more, both the GP and GPRN obtained consistent results. That can be seen in table 4.6. Looking at the MAP values, both rotation periods obtained are higher than the 27 days mark, not far from the synodic rotation period. The η_4 length-scale was also higher than in the previous analysis. On both models, it surpassed the unit, implying a far smoother structure inside the detected rotation period.

On the GPRN, the η_2 timescale of the weight obtained the highest values of the four analyses. The previous time series obtained values between 1000 and 2500 days, roughly between 2 and 7 years. In this analysis, however, the MAP value was of 4124.169 days, roughly 11.3 years. That is quite an interesting result. The weight with such a high η_2 in without a doubt becoming constant. However, it will still be decaying on a timescale of 11 years. It is known that the $\log R'_{hk}$ is very sensitive to stellar activity. If this result is somewhat related to the 11 years

Framework	Variables	MAP value	Median value
GP	η_1	0.006	$0.006^{+0.001}_{-1\times10^{-4}}$
	η_2	21.211	$20.823^{+1.367}_{-1.357}$
	η_3	28.397	$28.414^{+0.661}_{-0.610}$
	η_4	1.052	$1.095^{+0.091}_{-0.079}$
	s	0.001	$0.001^{+8\times10^{-5}}_{-8\times10^{-5}}$
	slope	-4×10^{-5}	$-4\times10^{-5}^{+3\times10^{-6}}_{-3\times10^{-6}}$
GPRN	offset	-4.989	$-4.989^{+0.001}_{-0.001}$
	node η_1	0.010	$0.050^{+0.042}_{-0.024}$
	node η_2	20.530	$20.468^{+1.351}_{-1.426}$
	node η_3	28.273	$27.425^{+0.720}_{-0.666}$
	node η_4	1.074	$1.049^{+0.087}_{-0.079}$
	weight η_1	0.102	$0.050^{+0.042}_{-0.028}$
	weight η_2	4124.169	$5518.676^{+3541.486}_{-2991.382}$
	s	0.001	$0.001^{+9\times10^{-5}}_{-8\times10^{-5}}$
	slope	-4×10^{-5}	$-4\times10^{-5}^{+2\times10^{-6}}_{-2\times10^{-6}}$
	offset	-4.989	$-4.989^{+0.001}_{-0.001}$

Table 4.6: Parameters obtained by the GP and the GPRN on the analysis of the $\log R'_{hk}$ time series.

magnetic cycle is a hypothesis that will require more careful analysis for the future.

A very noticeable difference exists between the values of the evidence and RMS reduction. The evidence shows once more the GP to be the best model obtaining a $\log Z$ of 2188.970 ± 0.115 and the GPRN getting a $\log Z$ of 2168.673 ± 0.384 . These values are quite different from those of the other analyses. I cannot exclude the hypothesis that such high values are occurring due to an over-fitting of the models.

If we look at the fitted models, shown in figure 4.7, they show a near-perfect fit. The GP obtained an RMS reduction of 10.101, and the GPRN obtained an RMS reduction of 9.756. I will explore some hypotheses for this soon, in section 4.2.3.

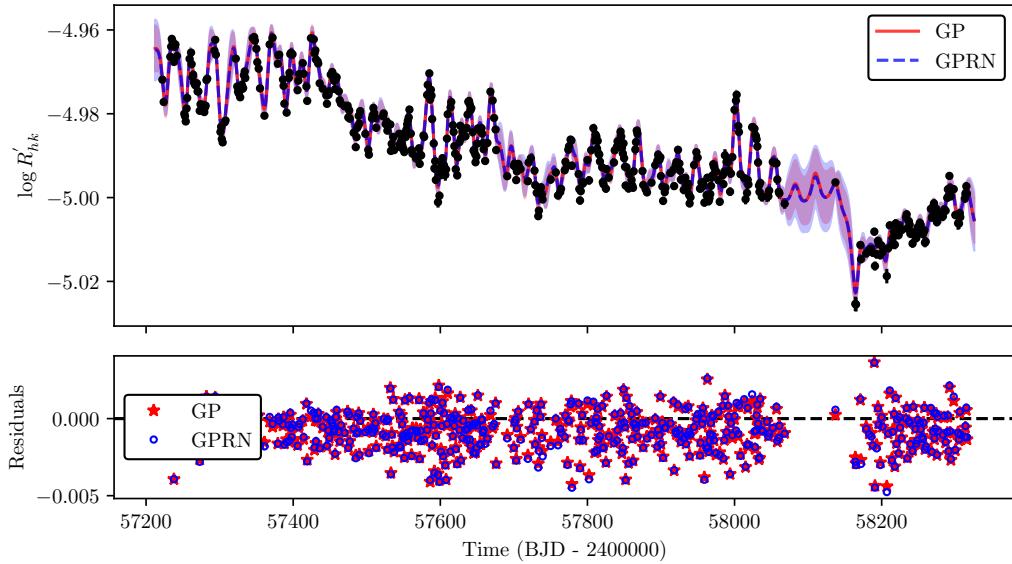


Figure 4.7: Top: Superimposed fits of the MAP values for the GP (solid red) and the GPRN (dashed blue) models on the $\log R'_{hk}$ time series. Bottom: Residuals obtained of the GP (red stars) and the GPRN (blue circles).

Computationally the GP proved to be more efficient. It required 10000 iterations to achieve our convergence criteria. The GPRN achieved it after 30000 iterations.

Analysis of the RVs and BIS

I now present the more interesting part of this work. That is, the combined analysis of the RV and a activity indicator. The first being an analysis of the RV and BIS measurements. The GPRN model now has two η_2 time-scales absent in the GP model we use. That means we now have a GPRN with 14 parameters while the GP has 11 parameters. However, under the right circumstances these two models show be equivalent.

The parameters easy to compare between the models are, by now without much of a surprise, are consistent. I mean that, for example, both models manage

Framework	Variables	MAP value	Median value
GP	η_2	22.478	$22.711^{+1.317}_{-1.241}$
	η_3	26.961	$26.849^{+0.341}_{-0.331}$
	η_4	0.696	$0.705^{+0.043}_{-0.041}$
	$\eta_{1_{\text{RV}}}$	1.185	$1.214^{+0.095}_{-0.088}$
	$\eta_{1_{\text{BIS}}}$	1.315	$1.336^{+0.086}_{-0.079}$
	s _{RV}	0.801	$0.816^{+0.040}_{-0.037}$
	s _{BIS}	0.374	$0.397^{+0.023}_{-0.028}$
	slope _{RV}	-0.005	$-0.005^{+0.001}_{-5 \times 10^{-3}}$
	offset _{RV}	-19.047	$-19.105^{+0.164}_{-0.160}$
	slope _{BIS}	-0.005	$-0.005^{+0.001}_{-0.001}$
GPRN	offset _{BIS}	-91.618	$-91.660^{+0.174}_{-0.175}$
	node η_1	1.163	$0.547^{+1.450}_{-0.371}$
	node η_2	24.319	$23.647^{+1.640}_{-1.603}$
	node η_3	26.745	$27.055^{+0.408}_{-0.390}$
	node η_4	0.783	$0.744^{+0.058}_{-0.056}$
	weight $\eta_{1_{\text{RV}}}$	0.413	$2.209^{+4.748}_{-1.626}$
	weight $\eta_{2_{\text{RV}}}$	281.890	$2057.797^{+4117.223}_{-1677.437}$
	weight $\eta_{1_{\text{BIS}}}$	0.669	$3.049^{+6.617}_{-2.247}$
	weight $\eta_{2_{\text{BIS}}}$	1676.794	$5074.099^{+3661.307}_{-2668.754}$
	s _{RV}	1.187	$1.204^{+0.042}_{-0.039}$
	s _{BIS}	0.423	$0.438^{+0.032}_{-0.030}$
	slope _{RV}	-0.005	$-0.005^{+3 \times 10^{-3}}_{-3 \times 10^{-3}}$
	offset _{RV}	-19.239	$-19.177^{+0.108}_{-0.106}$
	slope _{BIS}	-0.005	$-0.005^{+0.001}_{-0.001}$
	offset _{BIS}	-91.780	$-91.780^{+0.172}_{-0.106}$

Table 4.7: Parameters obtained by the GP and the GPRN on the analysis of the RV and the BIS time series.

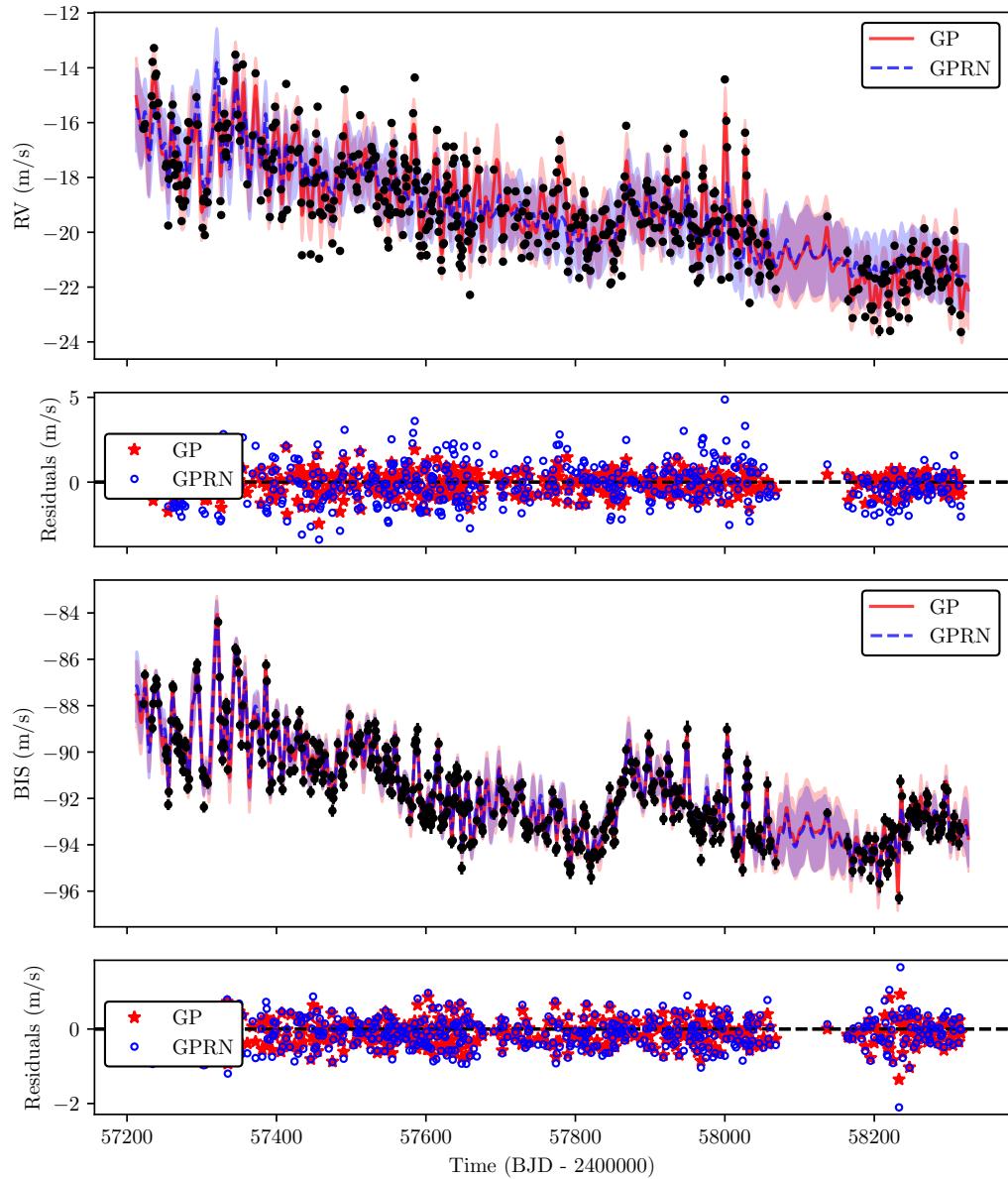


Figure 4.8: Superimposed fits of the MAP values for the GP (solid red) and the GPRN (dashed blue) models on the RV and the BIS time series. On the bottom of each fit are the residuals obtained of the GP (red stars) and the GPRN (blue circles).

to obtain the expected rotation period of the Sun given by η_3 . These results are shown on table 4.7. Curiously if we take a look at the GPRN plot of the posterior distributions on figure B.10. We can see some samples for η_3 for values around 36-37 days and for 44-45 days. These unexpected samples correspond to smaller values of η_4 , and have no impact on the values of the remaining ones. These samples however have a minimal impact in the final solution. The values of the parameters of GP and of the node of the GPRN show these are indeed showing a quasi-periodic behaviour. The weights on our GPRN however, seem to behave differently. If for the BIS, its weight η_2 indicate its tending to a constant function, this is different on the RV. The smaller η_2 value on the RV weight indicate trying to fit some structure. The fits presented on figure 4.8 help us explain why. While the BIS fits show to be reasonable good, for both GP and GPRN, the same does not happen on the RV data points. The residuals show the GPRN is having a clear difficulty on fitting these data points. One explanation is the node being sufficiently good to fit the BIS but not the RV. If this happens the RV weight might try to compensate and try to fit the data, thus becoming less constant and more wiggly. On section I will explore more of this important issue.

Seeing a worse GPRN fit on the RV data is of no surprise its impact on the evidence and RMS. The GP has an explicit better evidence with a $\log \mathcal{Z}$ of -1422.486 ± 0.195 . This is comparable to the sum of the evidences of the GP analysis of the RV and the BIS. The GPRN obtained a $\log \mathcal{Z}$ of -1471.618 ± 0.724 . This values is far lower to the sum of the evidences of its corresponding individual analysis. The RMS calculation also reflect the problems the GPRN fit. If for the BIS the GP reduced the RMS by 5.573 and the GPRN by 4.700. On the RV, the GP reduced it by 2.950 while the GPRN only reduced it by 1.685. This is a far lower, even when comparing to the RMS of the individual analysis.

Computationally the joint GP analysis was faster in achieving convergence. This is a trend I can now say the GP show, in relation to the GPRN. The GP converged after 15000 iterations. The GPRN, that is now a even more complex models with two weights, achieved convergence after 100000 iterations.

Framework	Variables	MAP value	Median value
GP	η_2	19.154	$19.674^{+1.302}_{-1.301}$
	η_3	27.336	$27.370^{+0.454}_{-0.425}$
	η_4	0.622	$0.663^{+0.056}_{-0.054}$
	$\eta_{1_{\text{RV}}}$	1.152	$1.174^{+0.087}_{-0.077}$
	$\eta_{1_{\text{FWHM}}}$	1.908	$1.939^{+0.130}_{-0.122}$
	s_{RV}	0.804	$0.805^{+0.038}_{-0.037}$
	s_{FWHM}	1.082	$1.098^{+0.057}_{-0.056}$
	slope _{RV}	-0.005	$-0.005^{+4 \times 10^{-4}}_{-5 \times 10^{-4}}$
	offset _{RV}	-19.080	$-19.116^{+0.148}_{-0.145}$
	slope _{FWHM}	-0.001	$-0.001^{+0.001}_{-0.001}$
GPRN	offset _{FWHM}	7785.114	$7785.028^{+0.237}_{-0.234}$
	node η_1	1.951	$0.511^{+1.147}_{-0.340}$
	node η_2	22.183	$20.779^{+1.934}_{-2.000}$
	node η_3	27.643	$27.680^{+0.621}_{-0.581}$
	node η_4	1.008	$0.861^{+0.120}_{-0.109}$
	weight $\eta_{1_{\text{RV}}}$	0.440	$1.809^{+3.691}_{-1.257}$
	weight $\eta_{2_{\text{RV}}}$	37.319	$66.105^{+56.031}_{-24.361}$
	weight $\eta_{1_{\text{FWHM}}}$	0.675	$4.812^{+9.103}_{-3.383}$
	weight $\eta_{2_{\text{FWHM}}}$	1660.697	$4673.599^{+3920.324}_{-2729.219}$
	s_{RV}	1.074	$1.129^{+0.056}_{-0.0057}$
	s_{FWHM}	1.229	$1.200^{+0.058}_{-0.056}$
	slope _{RV}	-0.004	$-0.004^{+4 \times 10^{-3}}_{-4 \times 10^{-3}}$
	offset _{RV}	-19.686	$-19.467^{+0.220}_{-0.265}$
	slope _{FWHM}	-0.001	$-1 \times 10^{-3}^{+0.001}_{-0.001}$
	offset _{FWHM}	7783.414	$7784.166^{+0.789}_{-0.716}$

Table 4.8: Parameters obtained by the GP and the GPRN on the analysis of the RV and the FWHM time series.

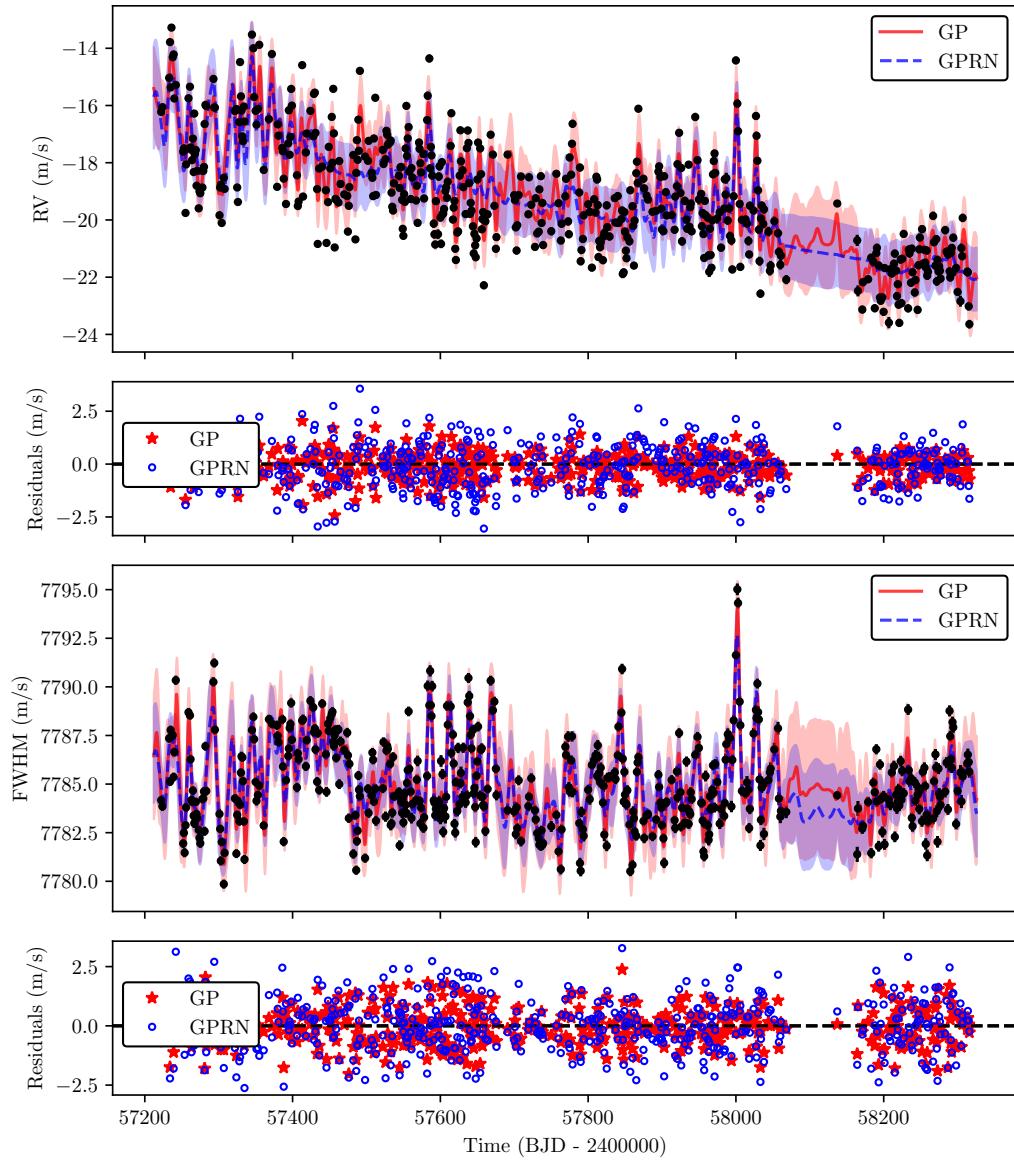


Figure 4.9: Superimposed fits of the MAP values for the GP (solid red) and the GPRN (dashed blue) models on the RV and the FWHM time series. On the bottom of each fit are the residuals obtained of the GP (red stars) and the GPRN (blue circles).

Analysis of the RVs and FWHM

Our second combined analysis uses the RV and FWHM measurements. On it, some common parameters show a different result. All are presented on table 4.8. On the GPRN, the η_4 length-scale of the node obtained values higher than expected. That on both for the MAP and median values. Considering the interpretation this parameter usually has, it makes the node considerably more sinusoidal than its counterpart on the GP analysis. That might explain the behaviour of other parameters on the GPRN.

First are the weights and their parameters. The weight connected to the FWHM behaves seen on other tests. That is obtaining a high value on the decaying time scale. That indicates an almost constant term. The weight connected to the RV however, has a η_2 smaller than 100 in both MAP and median values. That indicates a weight with a considerable structure.

The η_4 of the GPRN also has an interesting relationship with offsets on the mean functions. That is seen on the posterior distributions of appendix B.12. Both offsets show a bimodal distribution. These correspond to two different regions of the η_4 parameter space. Curiously, the MCMC converged to the region of significantly higher η_4 and rejected the one corresponding to values of $\eta_4 \sim 0.7$.

Looking at the fits obtained with the models in figure 4.9, the GPRN showed issues again. If we consider the residuals obtained and the RMS reduction we can determine where. The GPRN reduced the RMS by 1.917 times on the RV measurements, and 2.068 on the FWHM measurements. While the values corresponding to the FWHM is similar to the respective individual analysis, it is not for the RV. That indicates that while the node of GPRN was sufficient to model the FWHM data points, it was incapable of doing the same on the RV. That led to an extra effort to the weight connected to the RV to tackle this, thus having significantly more structure. The weight being a squared-exponential GP will not have a quasi-periodic behaviour, limiting its fitting capabilities. The inability of the node to fit the RV, and the fitting limitations of the squared exponential weight also explains why the white noise term converged to a significantly higher value.

On the combined GP analysis, none of this is observed. The figure 4.9 shows, through the residuals obtained, a far better fit. That will, of course, reflect

the reduction of RMS. On the RV is was reduced 3.073 times and on the FWHM 2.584 times. Unsurprisingly this is reflected in the evidence of the models. The GPRN obtained a $\log \mathcal{Z}$ of -1812.882 ± 3.339 , and the combined GP a $\log \mathcal{Z}$ of $-17.67.184 \pm 0.105$.

Computationally the joint GP analysis managed to converge after 15000 iterations. The GPRN took 105000 iterations.

Analysis of the RVs and $\log R'_{hk}$

Our last analysis consisted on the combination of the RV and $\log R'_{hk}$ time series. The final parameters are shown on table 4.9. The quasi-periodic parameters are fairly similar between the models. That is, the decaying timespan η_2 and the period η_3 , are consistent within a $1 - \sigma$ interval. Only the length-scale η_4 is slightly different. The GPRN model with a higher η_4 seem to imply a smoother structure inside the detected periodicity.

The weights of the GPRN reveal a behaviour similar to the previous analysis. While the weight associated to the FWHM is behaving as a quasi constant kernel ($\eta_2 \sim 3000\text{days}$). The weight connected to the RV measurements has, again, a η_2 value that creates a significant structure on the weight. With a MAP value of 86.682 days, this created a squared exponential weight with a clear decayling structure inside our 1094 days span time-series. The reason seem can be infered from the fits of figure 4.10. It seems that the GPRN fits the activity indicator more easily, leaving the RV weight to compensate the bad fit on the RV data points. Similarly to the previous, RV plus FWHM analysis, the bad fit on the RV is also reflected on a increase white noise. This better or worse fit will be more clearly explained in the next section.

It is with no surprise all this to be reflected on the RMS and evidence of the models. The residuals of figure 4.10 show the GPRN with higher residuals on the RV data, while the difference of the $\log R'_{hk}$ are less noticeable. The RMS reduction however indicate the GP fitted better both datasets. For the RV, the combined GP decreased the RMS by 2.745 times. For the $\log R'_{hk}$ this decrease was of 10.757 times. The GPRN decreased the RV and the $\log R'_{hk}$ by 1.797 and 9.628 times, respectively. The evidence obtained on the GP was of 1394.419 ± 0.141 , far

Framework	Variables	MAP value	Median value
GP	η_2	21.318	$21.135^{+1.059}_{-1.076}$
	η_3	27.559	$27.651^{+0.427}_{-0.419}$
	η_4	0.904	$0.924^{+0.059}_{-0.057}$
	$\eta_{1_{RV}}$	1.257	$1.294^{+0.110}_{-0.105}$
	$\eta_{1_{\log R'_{hk}}}$	0.006	$0.005^{+4 \times 10^{-4}}_{-4 \times 10^{-4}}$
	s_{RV}	0.864	$0.850^{+0.039}_{-0.037}$
	$s_{\log R'_{hk}}$	0.001	$0.001^{+9 \times 10^{-5}}_{-8 \times 10^{-5}}$
	$slope_{RV}$	-0.005	$-0.005^{+0.001}_{-0.001}$
	$offset_{RV}$	-19.167	$-19.102^{+0.193}_{-0.187}$
	$slope_{\log R'_{hk}}$	-4×10^{-5}	$-4 \times 10^{-5}^{+3 \times 10^{-6}}_{-3 \times 10^{-6}}$
GPRN	$offset_{\log R'_{hk}}$	-4.989	$-4.989^{+0.001}_{-0.001}$
	node η_1	0.001	$0.073^{+0.053}_{-0.027}$
	node η_2	19.654	$20.498^{+1.362}_{-1.392}$
	node η_3	28.615	$28.316^{+0.706}_{-0.651}$
	node η_4	1.057	$1.078^{+0.088}_{-0.082}$
	weight $\eta_{1_{RV}}$	5.939	$10.086^{+5.870}_{-4.432}$
	weight $\eta_{2_{RV}}$	86.682	$88.988^{+21.555}_{-19.312}$
	weight $\eta_{1_{\log R'_{hk}}}$	0.024	$0.059^{+0.037}_{-0.027}$
	weight $\eta_{2_{\log R'_{hk}}}$	3041.504	$5898.438^{+3312.965}_{-3012.180}$
	s_{RV}	1.126	$1.123^{+0.039}_{-0.037}$
	$s_{\log R'_{hk}}$	0.001	$0.001^{+9 \times 10^{-5}}_{-9 \times 10^{-5}}$
	$slope_{RV}$	-0.005	$-0.005^{+3 \times 10^{-4}}_{-3 \times 10^{-4}}$
	$offset_{RV}$	-19.400	$-19.374^{+0.098}_{-0.099}$
	$slope_{\log R'_{hk}}$	-4×10^{-5}	$-4 \times 10^{-5}^{+2 \times 10^{-6}}_{-2 \times 10^{-6}}$
	$offset_{\log R'_{hk}}$	-4.990	$-4.990^{+0.039}_{-0.037}$

Table 4.9: Parameters obtained by the GP and the GPRN on the analysis of the RV and the $\log R'_{hk}$ time series.

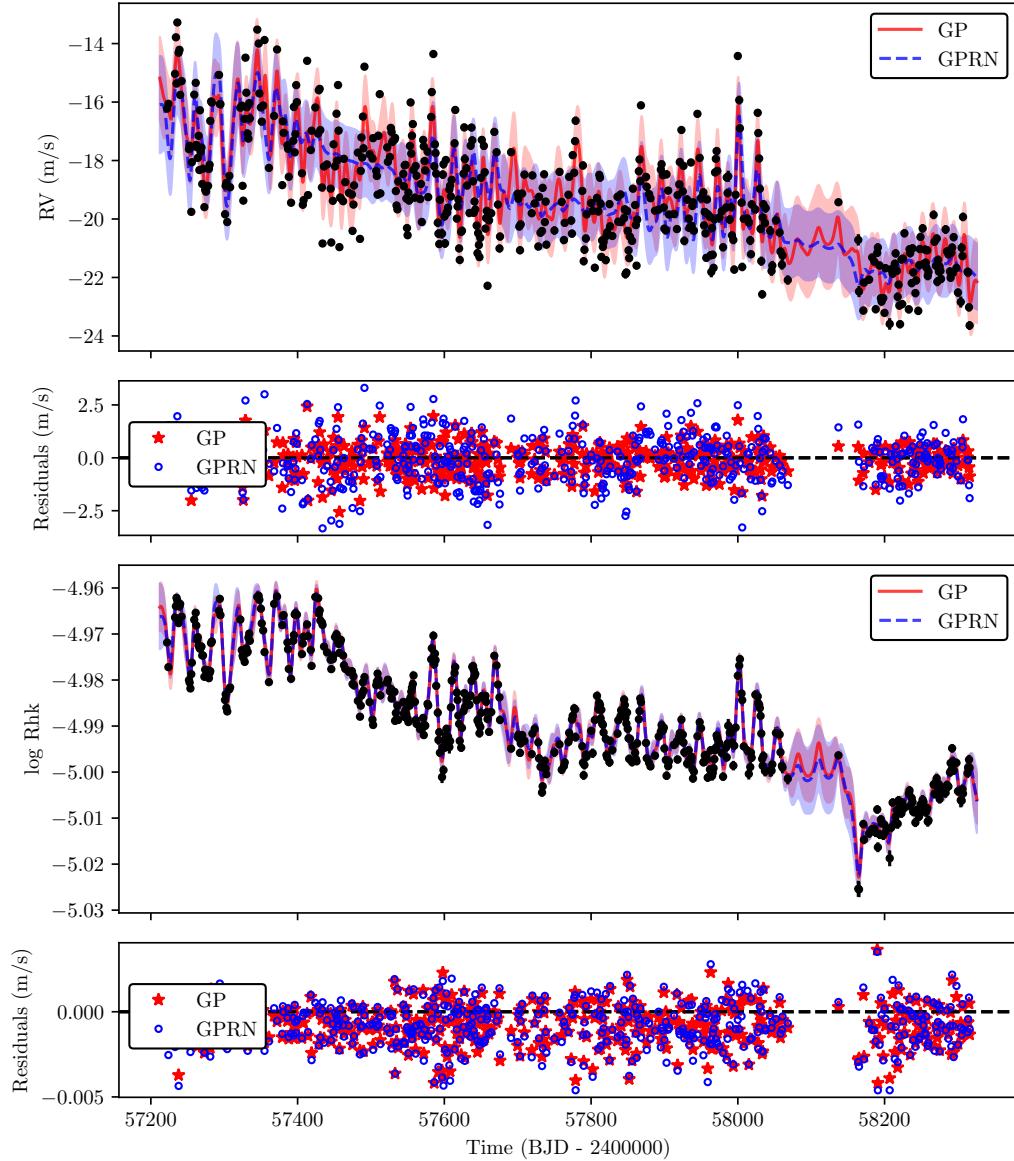


Figure 4.10: Superimposed fits of the MAP values for the GP (solid red) and the GPRN (dashed blue) models on the RV and the $\log R'_{hk}$ time series. On the bottom of each fit are the residuals obtained of the GP (red stars) and the GPRN (blue circles).

higher than the one of the GPRN. For our framework $\log \mathcal{Z}$ had the value of 1368.207 ± 5.843 .

Similar to the previous analysis, the joint GP analysis was computationally more efficient. It managed to converge after 15000 iterations, while the GPRN required 100000 iterations.

4.2.3 Interpretation of the results

The previous pages focused only in presenting the performance of the GPRN and compare it to a traditional GP. I mentioned the influence of parameters like, for example, the decaying timespan η_2 would have. That however was not backed by any graphical explanation. I left it for this section since a more in dept analysis. I will explore result on the individual RV analysis, as well as the combined RV and FWHM analysis. I only do it for these two as the conclusions can be generalized for the others analysis.

The complexity of a GPRN is important to understand. That is not only due to its poor performance. It is important to understand if comparing it to a GP is the correct approach. A GPRN is constituted by weights and nodes defined as independent GPs. It is only normal to question what will be the impact each term will have on the final fit.

Returning to the individual analysis of the RV time series from 4.2.2. Someone familiar with the properties of the kernels used in GP regression, will easily determine how a given GP is behaving. For example, the high values, for both MAP and median, obtained on the weight η_2 parameter. That, on a GP, indicates the squared exponential as converging to a quasi-constant function. Or at least a squared exponential with an long decaying time span.

Generating the plots of the node and weight this is exactly what is happening. From figure 4.11 we can see a quasi-constant weight and a extremely wiggly node. On the figure we see that it will be the node to model the structure present on the RV data points. The weight only increases or decreases the amplitude of the node to better represent the RV signal. If we include the linear function used as a mean, we obtain the final fit to the data plotted of the right side of figure 4.11. This behaviour is expected as, per equation 3.33, the amplitude and non-stationarity of

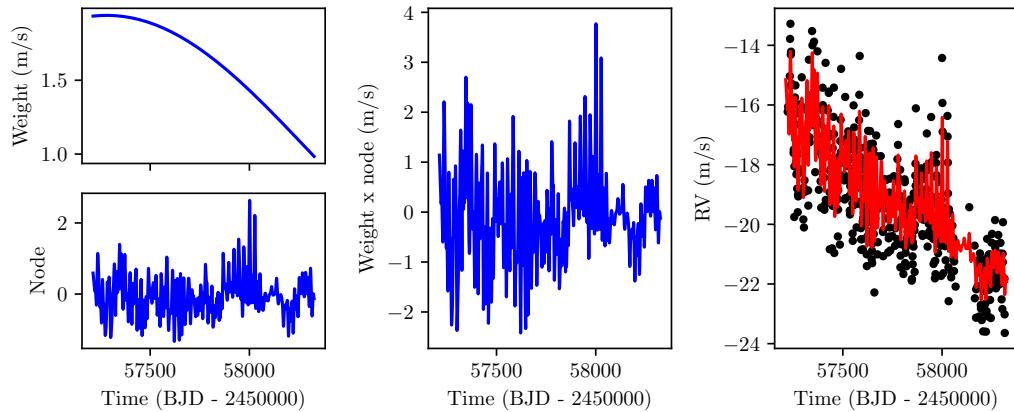


Figure 4.11: Left: Plots of the weight and node using the respective MAP values from the RV analysis. Middle: Product on the weight and node. Right: GPRN fit on the RV data points with the obtained parameters.

the GPRN are defined by the weight.

Although not shown here, this relation between the node and weight is common in all individual analysis. The node always adapts it self to the structure of the time series. That is complemented by a quasi-constant weight that only alters the node's amplitude through out time.

Considering the non-stationary structure of the GPRN that is being produced. Knowing the stellar activity from our data has a non-stationary behavior. Why have the individual analysis with the GPRN performed worse than a simple quasi-periodic GP? Two hypothesis exist that might explain it.

As mentioned in chapter 3.3, inference in a GPRN is intractable. The solution we found for this issue was to use mean-field inference to find an approximation to the posterior. And as a measure of that approximation we used the ELBO, also known as, evidence lower bound. As the name indicates, this measure is a lower bound approximation we will use to evaluate the posterior. I do not prove here but other works showed the ELBO to be in fact a lower bound to the true log-evidence (e.g Bishop, 2006).

Knowing this, implies that the GPRN evidences calculated earlier with the

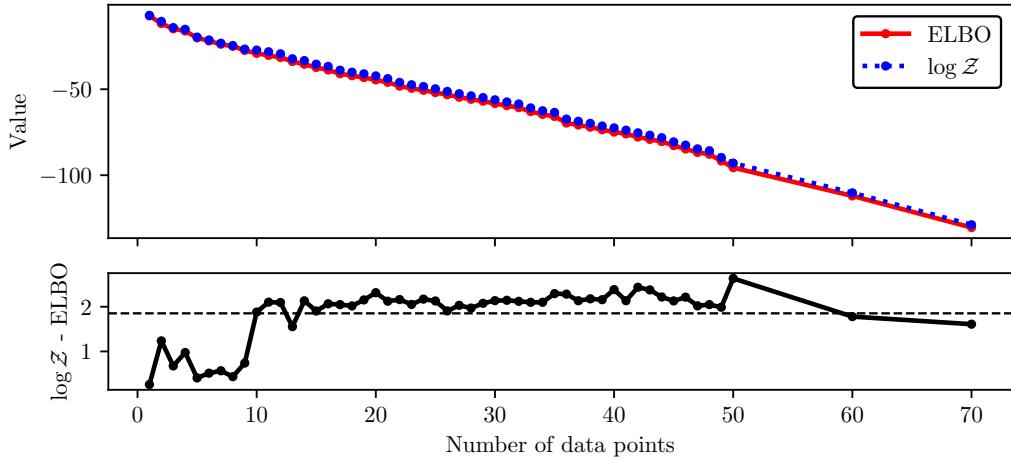


Figure 4.12: Top: ELBO and $\log \mathcal{Z}$ values given by the variational inference and DYNESTY, respectively. Bottom: Difference between the $\log \mathcal{Z}$ and the ELBO. The average value of the difference, of approximately 1.852, is plotted by the dashed horizontal line. **To be updated until submission**

Perrakis method are also a lower bound. This of course means that, for some cases the true evidence of the GPRN might be bigger than that of the GP. Although not feasible to implement on the GPRN efficiently. The true evidence can be calculated, for a given set of data points and a given set of GPRN parameter, using nested sampling (Skilling, 2006).

To test this hypothesis I used of the MAP parameter from table 4.3 and calculated both ELBO and $\log \mathcal{Z}$ for different number of RV data points. As nested sampling algorithm I used the functions implemented on DYNESTY (Speagle, 2020). The result of this experiment is plotted on figure 4.12. Considering we are using an approximation obtained from mean-field inference the question was on how far off we were from the real value. Also important to know is if this difference was constant or random for the different tests.

With no surprise all the ELBO values were smaller than the $\log \mathcal{Z}$ from DYNESTY. The difference between them was also not large. On average the ELBO and $\log \mathcal{Z}$ differ by 1.852. This seemed to be fairly constant difference with the increasing number of data points. With the non-randomness of this difference

	GP	GPRN	GPRN+1.852
RV	-789.106	-789.460	-787.608
BIS	-639.606	-643.512	-641.660
FWHM	-977.182	-979.329	-977.477
$\log R'_{hk}$	2188.970	2168.673	2170.525
RV and BIS	-1422.486	-1471.618	-1469.766
RV and FWHM	-1767.184	-1812.882	-1811.03
RV and $\log R'_{hk}$	1394.419	1368.207	1370.059

Table 4.10: Evidence values comparison, including old and new GPRN values.

we can, with confidence, include this value and recalculate the evidences for the GPRN. I show the old and new evidences on table 4.10.

This correction still does not give all the answers. On the individual analysis we now see the GPRN with higher evidence for the RV. We also see a evidence value indicating both models equally good on the FWHM. But it still leave the GP a better model on most cases.

A problem not solved is, even if the evidence is higher on the GPRN, we still have lower residuals on the GP models. That however might also be explained. On chapter 3 I stated that a GPRN will no easily over-fit a model. That being due to, and as stated on Wilson et al. (2012), a GPRN having a heavy-tailed predictive distribution. The reason behind that statement is on how a GPRN is created. On a GPRN we will always have at least two latent variables. On the simplest case we have one node and one weight. That setting will cause each output (RV, FWHM, etc...) to have a marginal distribution given by two independent marginal distributions. Important to remember, each node and weight is an independent GP. That means that, for each, we have a independent marginal distribution that is Gaussian.

Knowing this, it implies that each output marginal distribution is the product of two Gaussian distributions. It can be proved, but not shown here, that this product is not Gaussian (e.g. Cui et al., 2016). The product will have a distribution the given by a re-normalized Bessel function of the second kind of order zero.

For more on these functions I refer to Arfken et al. (2013). This type of Bessel function has indeed heavier tails than a Gaussian distribution. Also important to note is that this distribution occurs independently of the type covariance functions used on the node and the weight GPs.

That is what makes the posterior distributions of a GPRN to have bigger tails, and as consequence, less prone to over-fitting. That also causes its predictive does not require to approximate itself to the data since the tail does it. That means the RMS will, most likely, be worse of a GPRN, even when the evidence supports this model.

These two hypothesis are valid in all the analysis I made. However, when we analysed the RV plus as activity indicator the behavior became even more complex. To explain it I will use the results we obtained on 4.2.2.

On that analysis I used the RV and FWHM. As I mentioned, the parameters obtained on the GPRN indicated a quasi-constant weight connected to the FWHM, while the one of the RV had more structure. That is easily observable of figure 4.13 when I plot each component of the GPRN. The different structures of the weights seem related to the structure of the node. Looking again to figure 4.13. On the FWHM case, the node is relatively well design to fit the data. That leave the weight will only needing to fix the final amplitude. For the RV data points however, the node with a structure not good enough to represent the data. The weight connecting the node to the data has no other choice but to try compensate this bad fit. For it the MCMC converged to smaller values of η_2 . That would allow the function to change more significantly, thus more "wiggly".

However, this extra effort from the weight is not enough to overcome the bad fit. A first logical explanation is that using a squared-exponential kernel is not enough. Other kernels like the rational quadratic or Matérn kernel might indeed be better. Unfortunately due to time constrains, I had no possibility to test this. Even with a sufficiently malleable weight, that does not explain why the node adapts itself more easily to the FWHM measurements and not the RV.

One explanation is that there might be more information on the FWHM than on the RV. That begin, a quasi-periodic signal more clear on the FWHM. This behaviour was also recurrent on the BIS and $\log R'_{hk}$. The node always adapted itself better to the activity indicator than to the RV.

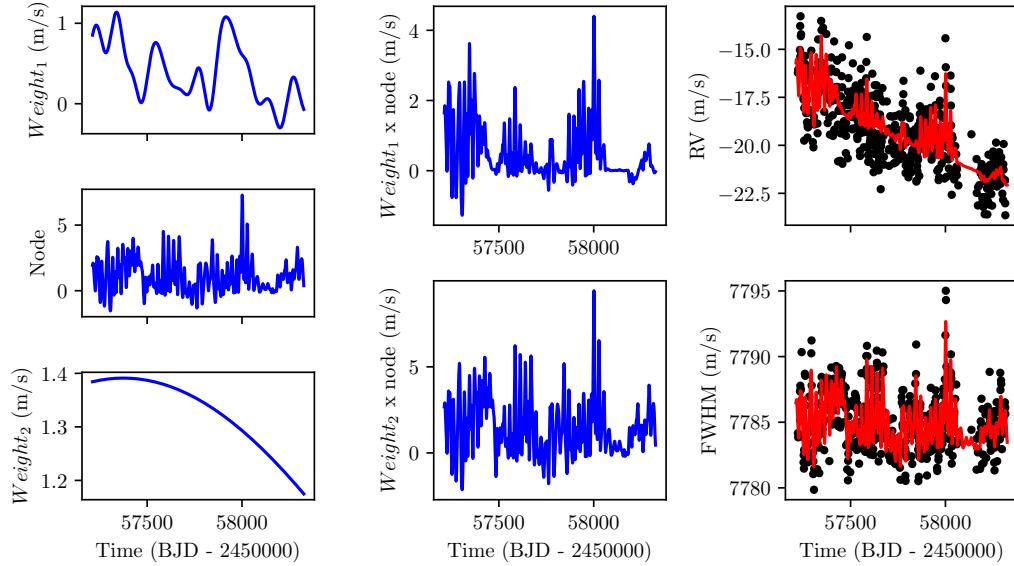


Figure 4.13: Left: Structure of each component of the GPRN. Middle: Product of the node by each respective weight. Right: GPRN fit to the RV and FWHM when using the MAP values obtained on the analysis.

Time lags

As said on section 4.2.1, the activity indicators showed a delay of 2 to 3 days on the BIS and of 1 to 2 days on the FWHM in relation to the RV. One aspect of interest is if this time lag was being taken into consideration in our analysis.

To determine it we used the predictives of our models. The objective being comparing the predictive of the RV with the one of the indicator and calculate if any shift among them existed. Figure 4.14 shows the results.

On the right are plotted the predictives of the RV and the indicator accordingly the framework used. The CCF was calculated using the obtained values of each predictive. The plots on the right however were adjusted to better visualise the predictives. To achieve it, I removed the mean and the divided each predictive by the respective η_1 . That allowed to normalize each predictive and more easily compare them visually. With it, it became possible, for example, to visualize a shift

on the GP predictives of the RV and BIS. These however required to be measured mode precisely. For it, we used again the discrete correlation function of Edelson and Krolik (1988).

On the left of figure 4.14 are the cross-correlation function plots. Each of these plots are between a GP and a GPRN predictives that they represent. With these it is clear that the GPRN is incapable of modelling any time lag between the time series. On the other hand the GP framework shows a shift between 2 to 3 days on the predictives of the RV and BIS. The GP framework also finds a shift around 1 days between the RV and FWHM. For last, the GP results show no significant shift between the $\log R'_{hk}$ and RV.

The lack of flexibility on the GPRN to model these shifts help explain the evidence differences from table 4.10. The framework not being capable of shifting one time series in relation to the other creates a worse fit. That added to the fact the GPRN seem to prefer to fit one time series with the node. Trying to fit the remaining time series with the weight compromises the framework.

However knowing that the time lags between the RV and activity indicators exist. Knowing that the GPRN seem to fit the activity indicator with the node and tries to mold the RV with the weight. Such lack of time lag should be of no surprise between the predictives. That issue needs to be solved. With either using a GP on the weight capable of addressing lags between time series. Or adapting our variational inference algorithms to detecting any sort of time lag and shift the inputs accordingly. Since on the combined GP analysis, the GPs are independent, sharing only the parameters, might explain why this analysis is capable of dealing with phase shifts. Unfortunately, it does not help to find a solution for this lack of flexibility on a GPRN.

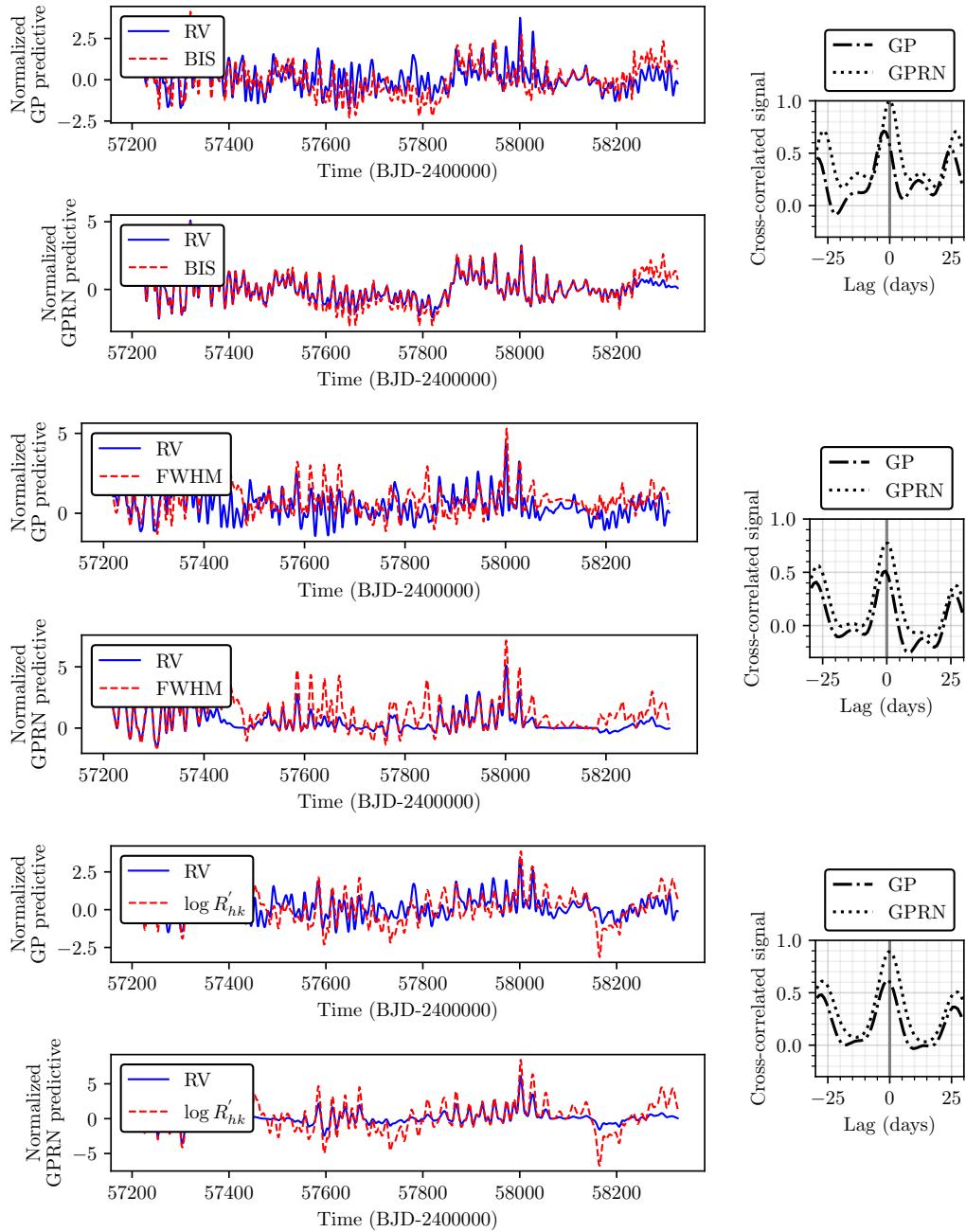


Figure 4.14: Left: Superimposed normalized predictives from GP and the GPRN. Right: Discrete CCF between the predictives of the RV and the activity indicator.

4.3 EXPRES data

In this section, I present the analysis of four stars from the EXPRES Stellar-Signals Project (ESSP). The ESSP aimed at comparing several spectroscopic data analysis methods. Thanks to it, we had access to spectroscopic observations carried by the Extreme Precision Spectrograph (EXPRES) (Jurgenson et al., 2016). In total, we had four targets to apply our GPRN framework on it. Those were HD 10700, HD 26965, HD 101501, and HD 34411.

Our GPRN would then use both RV and FWHM to analyse the observations and determine the scatter return from the RV fit. The RMS reduction of our framework would then be compared with the other methods to determine the level of success on mitigating the stellar variability. All the methods and results obtained with them are to be published by Zhao et. al. (in prep).

This challenge proved a valuable opportunity for us. It allows us to test the GPRN on stars with different levels of stellar activity and test its detection limits. It also grants us the opportunity to compare our framework with other time-domain methods. The GPRN created used a quasi-periodic GP on the node and a squared exponential GP on each weight. The priors where defined similarly to the ones used on the Sun analysis. These are shown on section 4.1.2.

4.3.1 HD 10700

HD 10700, or commonly known as τ Ceti, is a G8V type star. The $\log R'_{hk}$ value of -4.976 shows it to be chromospherically inactive, with a rotation period of 34.5 days (Keenan and McNeil, 1989; Pizzolato et al., 2003; Isaacson and Fischer, 2010). Four planet dynamically packed have been confirm to orbit this star (Feng et al., 2017).

The 174 RV measurements available are shown on figure 4.15. The observations have a RMS of 1.863 m/s and a time span of 470 days. Performing a periodogram analysis shows that most of the peaks are below the 1% false alarm probability level (figure 4.15). That seems to demonstrate that this method does not detect significant periodic signals. Its low activity level might imply there will not be significant activity contaminating the RV measurements with spurious

periodic signals. That being true, it might difficult the GPRN of obtaining any meaningful result.

This was indeed what happened. After running an MCMC for 225000 iterations, results obtained are important to analyse. The GPRN was not capable of, for example, find the expected rotation period. It instead obtained a period of $5.809^{+1.545}_{-0.562}$ days, with a MAP value of 5.350 days. Considering the lack of activity it might have contributed to this results. If spots or other stellar activity processes aren't present on the surface of the star, the rotation period print on the observation would be minimum.

The fit on figure 4.20 also indicated an incapacity of the GPRN to model the RV data. The node, with a extremely low decaying time span (~ 2.715 days) adapts itself exclusively to the activity indicator, leaving no flexibility to fit the RV. This resulted on a insignificant reduction on the RMS. That meant a reduction of only 2.652% in relation to the initial RV measurements. On the FWHM we had a far more significant reduction on the RMS. It was reduced by almost four times, from 11.761 m/s to 2.944 m/s.

These results at first sight seem like a bad performance of the GPRN. However these might be the expected behaviour for observations with no significant activity signals. The main idea behind a GPRN is to have a non-stationary model capable of modelling the non-stationary signals from the stellar activity. Considering the low activity level of HD 10700, there will be a limited activity signal contribution on the RV. That implying of course no quasi-periodic signal due to stellar activity on the measurements to be detected.

Considering the previous performance of the GPRN on the Sun observations I felt the need to confirm the results of HD 10700. To determine if the lack of activity was the reason behind the bad fits on the GPRN I redid the analysis with the artGPN framework from 3.2.3.

The new analysis achieved a similar performance. It failed to obtain a rotation period as well. The MAP value of it was of 5.624 days, similar to the GPRN. I left the fit of these GP to appendix B.15. These are very similar to the GPRN ones with the RV again badly modelled. That makes me confirm the results of the GPRN are indeed due to no quasi-periodic signal on the RV.

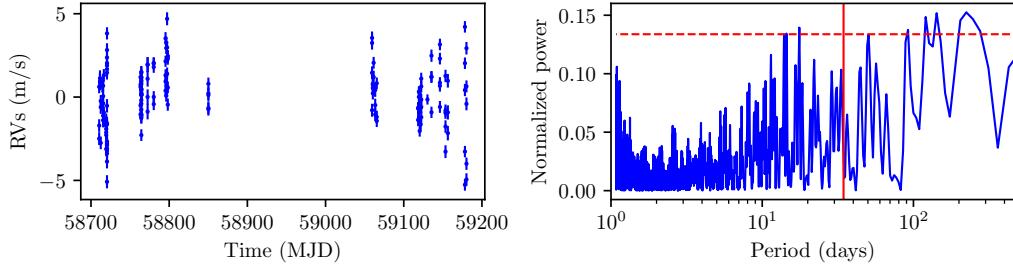


Figure 4.15: Left: RV observations of HD 10700. Right: Periodogram of the data after subtracting a long term trend. Also on it is marked, as a red vertical line, the expected rotation period of 34.5 days, and the FAP of 1% as a dashed horizontal line.

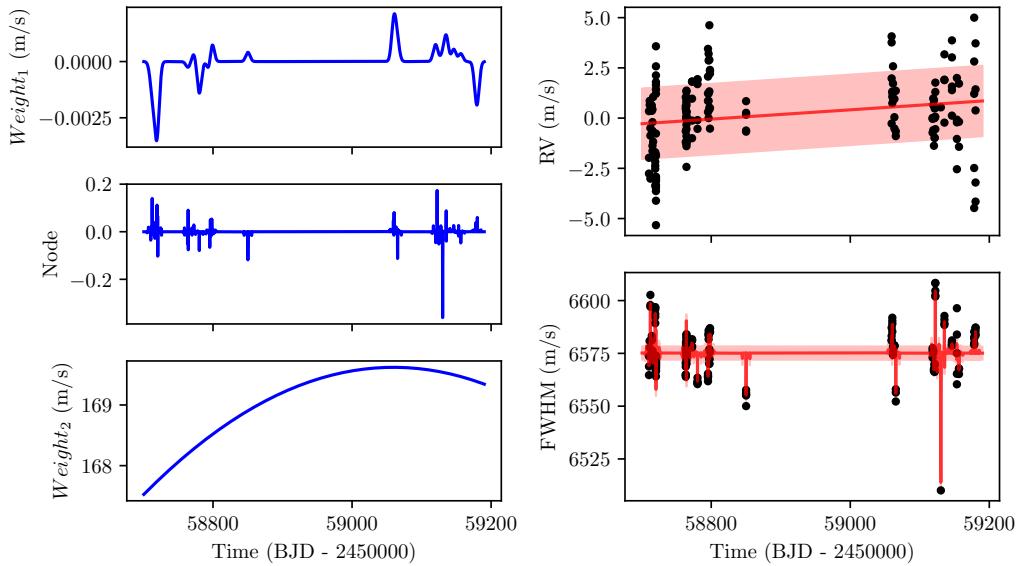


Figure 4.16: Left: Components of the GPRN. The weight₁ connects the node to the RV, while the weight₂ connects it to the FWHM. Right: Final fits of the GPRN to the RV and the FWHM.

4.3.2 HD 34411

HD 34411, also known as λ Aurigae, is a solar-type star classified as a G1V (Boyajian et al., 2012). This star is the least active of the four in analysis. It has a $\log R'_{hk}$ of -5.085 dex, and no known rotation period (Isaacson and Fischer, 2010). Such low value for the $\log R'_{hk}$, indicative of a very low magnetic activity, makes it a candidate star for a Maunder minimum event on its activity cycle (Lubin et al., 2012).

The low activity level can be observed on the periodogram of figure 4.17. On it no peak below 100 days is capable of surpassing the 1% FAP threshold. Not shown here, but similar lack of significant peak is seen on the FWHM periodogram.

Similar to HD 10700, the lack of activity had an impact on the parameters found by the GPRN. It was unable to converge to a well-defined period. By its marginal posterior probability distribution, the η_3 remained unconstrained until convergence, obtaining a final value of $29.425^{+13.165}_{-17.776}$ days. Its MAP value was as of 44.104 days although most likely not related to the true rotation period. However the GPRN still managed to model some structure on the RV. This resulted on a slight decrease of the RMS from 1.768 m/s to 1.421 m/s. On the FWHM the fit was, again, more successful managing to decrease the RMS from 5.244 m/s to 1.538 m/s.

The GPRN fit of the RV and the FWHM are shown at figure 4.18. The node is again preferring to adapt itself to the indicator, leaving the weight₁ to fit the RV as best as possible. However the limited modeling of the RV is also observed when analysing these measurements with a combined GP. See appendix B.16 for the obtained fits with this framework. The GP managed a slightly better RV RMS of 1.181 m/s and an converged value of η_3 of $10.543^{+13.711}_{-1.146}$ days. If this value is related to the rotation period is unknown.

The result with both GP and GPRN however, seems to support that, like as with HD 10700 the GPRN is not behaving wrongly. Meaning that, its bad fit seem to results of the lack of activity signals on the measurements. That of course does not excuse the GPRN to be a worse model in relation to a combined GP framework.

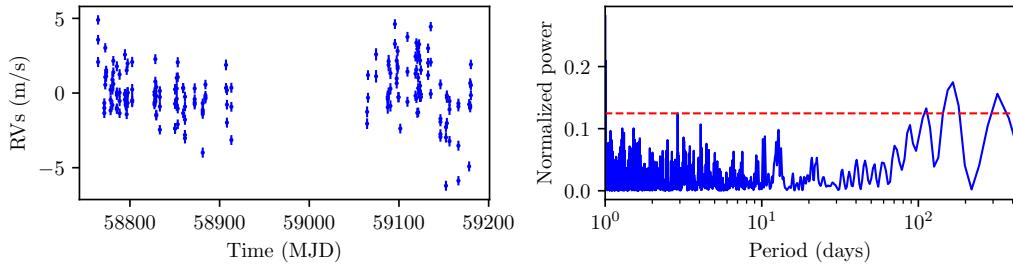


Figure 4.17: Left: RV observations of HD 34411. Right: Periodogram of the data after subtracting a long term trend. Also on it is marked the FAP of 1% as a dashed horizontal line. No rotation period is known for this star.

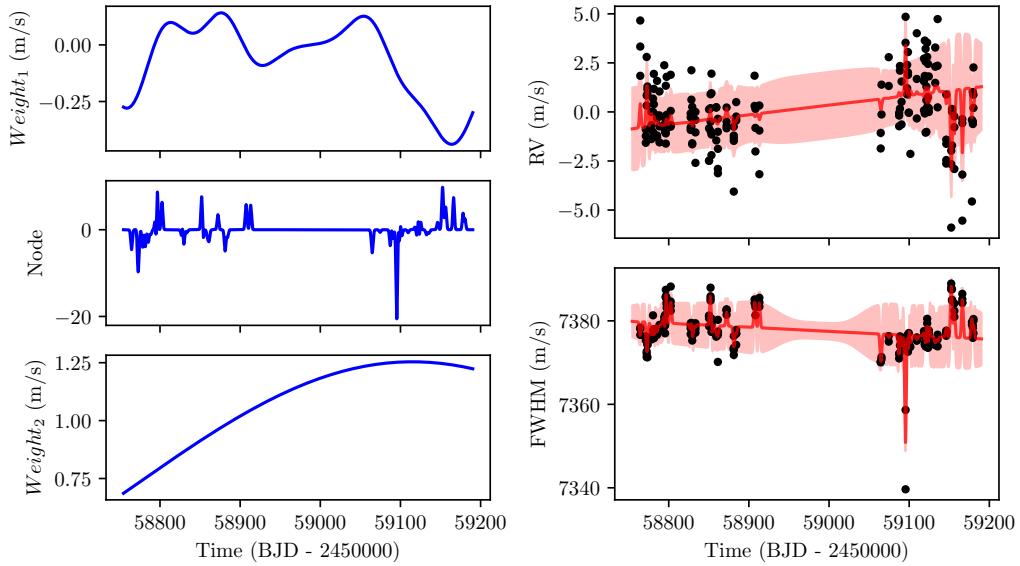


Figure 4.18: Left: GPRN fit to the RV measurements of HD 34411. Right: Contribution of the node (top) and weight (bottom) to the final fit.

4.3.3 HD 101501

HD 101501, also known as 61 Ursae Majoris, is a late main-sequence star classified as a G8V type star (Wilson, 1962). Of the four stars in this analysis is the most active one. It has a $\log R'_{hk} = -4.483$ and a rotation period of 17.1 *days* (Isaacson and Fischer, 2010; Maldonado et al., 2010). No known planet exist on this star.

We were given 45 observations scattered through a time span of 655 days. These are plotted on figure 4.19. On the periodogram of the RV observations the peak due to the rotation period is noticeable as having the highest power. However another significant peak exists at 12 days of unknown origin. Although not shown here, these peaks are also present on the FWHM periodogram. Whatever the origin for the 12 days peak, we can relate it to activity and not to a planetary signal.

The RV data had RMS scatter of 4.887 *m/s*, the highest value of the four stars. That supporting the high activity level of HD 101501. Being true that the low activity levels of the previous stars made the convergence of our model inefficient. The GPRN should have better results for HD 101501.

Our framework converged at the end of 500000 iterations. The η_3 determine was of $18.557^{+21.571}_{-4.835}$ *days*, consistent with the rotation period. A long "tail" for higher period was present, related to higher values of the decaying times span of the weight connected to the FWHM. For value of η_2 on this weight around 800 days constrains the η_3 to values around the rotation period. For higher values of η_2 , the η_3 becomes unconstrained creating a "tail" on the final posterior distribution. This "tail" however also represented regions of lower log-posterior values. The MAP value of the η_3 was of 14.981 *days*. The MAP solution of the η_2 on the node was similar, with a value of 14.888 *days*. This is possibly related to the growth and decay of the active regions.

The RV RMS saw a decrease from 4.920 *m/s* to just 2.156 *m/s*. That represents a decrease of 2.279 times, making clear the GPRN is performing better on this star. The fits of both RV and FWHM are shown on figure 4.20. On it we can see the usual GPRN choice of fitting better the activity indicator with the node. On the indicator the RMS decrease from 24.780 *m/s* to 1.753 *m/s*, a decrease of almost 14 times of its initial value. This might be the first proof that the GPRN

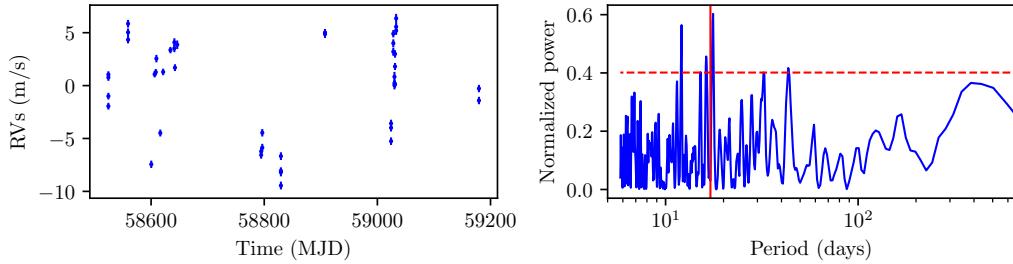


Figure 4.19: Left: RV observations of HD 101501. Right: Periodogram of the data after subtracting a long term trend. Also on it is marked, as a red vertical line, the expected rotation period of 17.1 days, and the FAP of 1% as a dashed horizontal line.

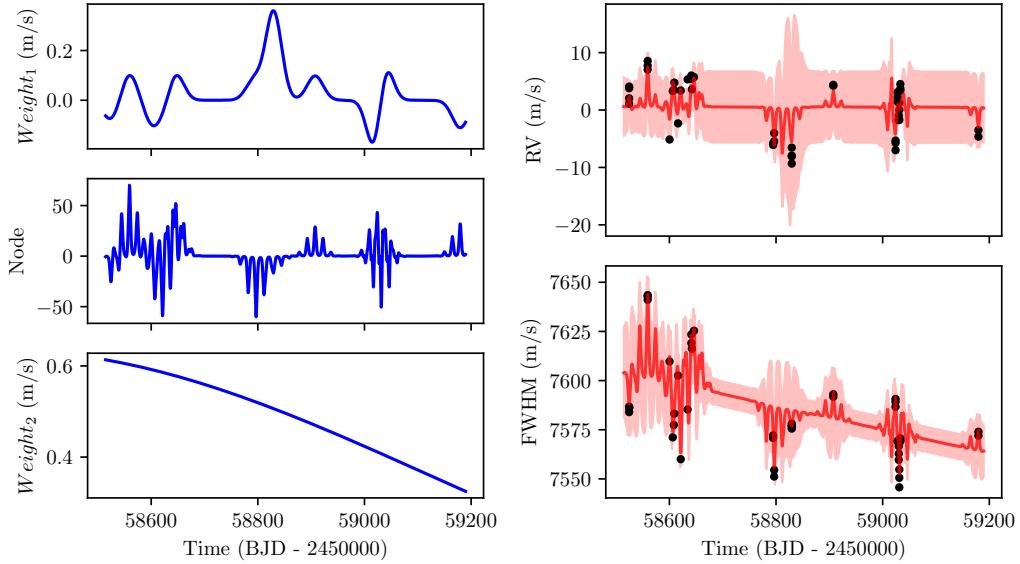


Figure 4.20: Left: GPRN fit to the RV measurements of HD 101501. Right: Contribution of the node (top) and weight (bottom) to the final fit.

performs better in more active stars.

We compared again the GPRN with a combined GP analysis. Curiously the GP reach a rotational period smaller to what was expected. On it the median

value was of $14.986_{-0.884}^{+11.853}$ days, and the MAP value of 14.374 days. The "tail" present on this solution is this time related with the η_4 value. Higher values of η_3 correlates with smaller values of η_4 , however also represent solutions of lower log-posterior values. Although with a lower solution for the rotation period, the GP managed to decrease the RMS to the value of 0.692 m/s.

4.3.4 HD 26965

For last I left the more interesting star. HD 26965, also known as 40 Eridani, is a K1V star with a rotation period between 39 and 44.5 days (Ma et al., 2018). It has a $\log R'_{hk}$ of -4.928 dex, making it the second most active star of the four (Pizzolato et al., 2003; Isaacson and Fischer, 2010).

Still debatable is the existence of an exoplanet. A planetary signal with a period of approximately 42 days was detected (Díaz et al., 2018; Ma et al., 2018). However having an orbital period so close to the rotation period of the star makes its existence suspicious.

We were given 114 observations with a time span of 465 days. The RV measurements shown an RMS scatter of around 3.234 m/s. The high RMS and increasing RV amplitude over time (see figure 4.21) show an increase in stellar activity.

The periodogram shows the region of the expected rotation period populated by a group of peaks with a significant power. That might explain the difficulties found in determining an accurate rotation period and the still-open debate of if the planetary signal discovered of 42 days is due to stellar activity or not. With a slightly higher power on the periodogram are a set of peaks around the 17 days mark. That is likely related to harmonics due to the periodic signals around the rotation period. The same behaviour, although not shown here, is seen on the periodogram of the FWHM measurements.

Our GPRN was left running for 120000 iterations until the MCMC achieved convergence. In the end we obtained a rotation period of $31.146_{-13.171}^{+5.818}$ days. The corresponding MAP value was of 32.785 days. This although lower than expect is still inside the region dominated by significant peaks on the periodogram. The node η_2 obtained a MAP value of 24.112 days, likely related to the evolution of

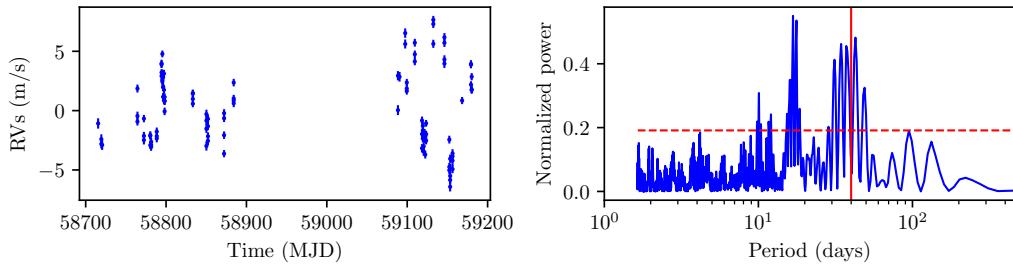


Figure 4.21: Left: RV observations of HD 26965. Right: Periodogram of the data after subtracting a long term trend. Also on it is marked, as a red vertical line, the expected rotation period of 40 days, and the FAP of 1% as a dashed horizontal line.

the active regions. As such it makes me believe this η_3 value is indeed related to the rotation period.

The fit, shown on figure 4.22, resulted on a RMS decrease from 3.234 m/s to just 0.815 m/s . That signified an RMS reduction of almost four times its original value. The fit of the GPRN also showed an unexpected result. The behavior of the node and weight connected to the RV data seems to indicate it is choosing to fit the RV. To compensate for this, the weight connected to the FWHM show a rougher structure to adapt itself to the activity indicator. If the GPRN is fitting all periodic signals, including the claimed planetary signals, is left for a future analysis. More about this star the meaning of this RV RMS will be discussed in Zhao et. al. (in prep).

However, if the GPRN is absorbing the planetary signal, the same occurs with a combined GP analysis. That analysis obtained a rotation period more inline with previous studies. The median value was of $36.983^{+8.402}_{-9.701} \text{ days}$. The GP decreased even more the RMS of the RV, achieving the value of 0.750 m/s . Again if this signifies an absorption of any planetary signals is left for future work.

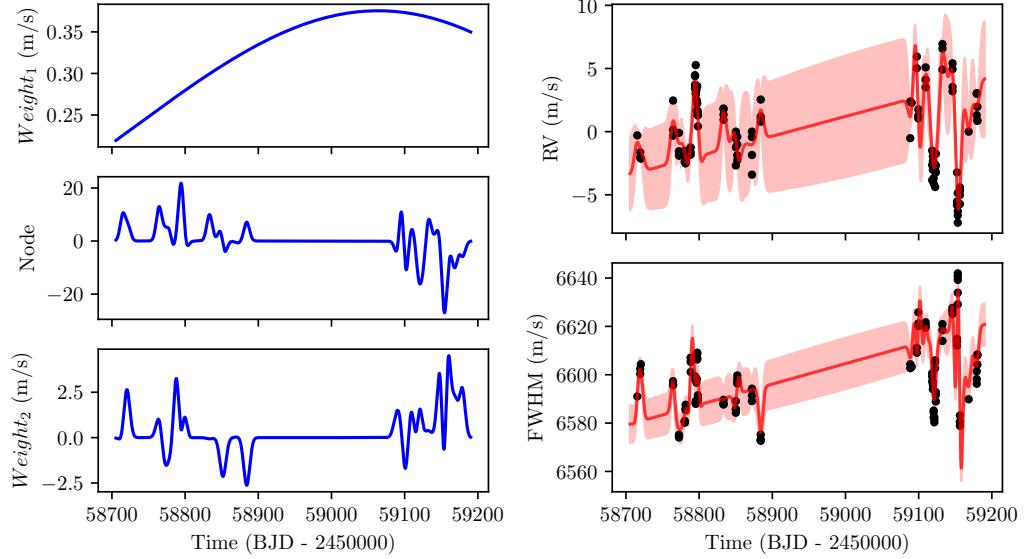


Figure 4.22: Left: GPRN fit to the RV measurements of HD 26965. Right: Contribution of the node (top) and weight (bottom) to the final fit.

4.3.5 Final thoughts on the EXPRES results

The four stars analysed for the ESSP proved to be an unique opportunity to test the GPRN on stars of different activity levels. These results indeed show the GPRN to have different performances accordingly to the activity level of the star. With its non-stationary structure it would be expected the GPRN would have a better performance on high activity levels. This is indeed the case. However, a even better performance was observed when a combined GP analysis was done.

Similarly to the case study of the Sun, the GPRN shows an unexpected lack of flexibility that need to be addressed. However I recommend the reading of the work of Zhao et. al. (in prep). On it the GPRN results are compared to 20 other methods. Considering the limitations showed by the GPRN, addressing them will only improve its performance. Thus, the results of the GPRN can be seen as a lower bound of its capabilities.

Regarding the stars known to have planets, the next obvious step would be

to include one or more keplerian functions on the GPRN analysis. On the current GPRN implementation that is easily included on the mean function. Unfortunately time constrains didn't allows a more careful analysis with it.

144 Advanced statistical data analysis methods for the detection of other Earths

Chapter 5

Conclusions

When I started this thesis, its main objective was to use advanced statistical data analysis methods to detect and characterize exoplanets using RV observations. In particular, aiming at detecting exoplanets with mass and orbital period similar to the Earth, on solar-type stars. To achieve it, we aimed at adapting the work of Wilson et al. (2012) to be used on radial velocity and respective activity indicators data.

A GPRN, being a non-stationary framework, is an improvement when compared with frameworks like the ones of Rajpaul et al. (2015) or Jones et al. (2017). Its application on radial velocity data proved to be a greater challenge than originally anticipated. The framework defined by Wilson et al. (2012) did not meet all our requirements, and some changes were needed. In the end, we opted to use the mean-field inference for GPRN developed by Nguyen and Bonilla (2013). That, however, didn't come without its constraints. The main being the variational inference used and the approximations used on its algorithms. Even so, I am secure the results I showed in this thesis are robust and in agreement with other methods.

In chapter 3, and before introducing the GPRN, I give a brief explanation on which kernels are usually used on GP regression. There I introduce an idea that I believe is still fairly unexplored when using GP in our field. New kernels can be created with interesting properties. Experimenting with new periodic or quasi-periodic kernels might lead to results not achieved before when using the "traditional" quasi-periodic kernel. Also in this chapter, I presented all the mathe-

matical structures needed to create a GPRN. I tried to do this without exaggerating the number of deductions needed to reach the necessary equations. However, I recognize that some of the expressions I use would be easier to understand knowing all the steps I used to reach them. The appendix of Nguyen (2015) is probably the best source to better understand some of them.

The main results I achieved in chapter 4 using the GPRN are somewhat frustrating. Unfortunately, the vast majority of my PhD was spent adapting the GPRN and writing and the necessary algorithms into a python package. Only upon beginning my last year of the PhD, I managed to have a fully functional GPRN. That lead to the immediate use of the Sun observations from Dumusque et al. (2021). However, the sheer number of observations and the computational complexity of the GPRN meant we only had one shot at obtaining results for the time frame we had. In the end, many experiments were left out due to this. We opted by applying a GPRN on its simplest form of one node and one weight per dataset. That does not take full advantage of the capabilities of a GPRN. Every time the GPRN results were compared with a GP, the GP came out on top. Considering this occurred using a very limited GPRN in terms of nodes and weights, the future of the GPRN is still promising.

Although the GP models showed to always obtain a higher evidence and lower RMS. I cannot ignore that the GPRN results were always in agreement with those of the GP. For example, the GPRN managed to recover the Sun's rotation period the same way the GP did. Considering the GPRN has still a lot of room for improvement, I would not be surprised that, with a different combination of nodes and weights, it surpassed the traditional GP framework.

5.1 Future work

The results clearly show the GPRN implemented, although functional, needs to be improved. Using a GPRN with just one node and one weight is clearly not enough to compete against a GP. In this form, the GPRN cannot, for example, successfully model or compensate for time lags between the time series.

I do not know if this should be addressed by using weights with a different kernel and/or different priors. I do not know if increasing the number of nodes would solve it. However, knowing these time lags exist, we need to find a way for the GPRN to deal with them.

When I initially started working on this project, one of the main objectives was to use it on a set of observations and determine how well we would capture a planetary signal. The use of a keplerian on the mean function of the GPRN is easy, however untested. If the GPRN will detect planetary signals more easily than a standard GP is work still to be done.

Another issue that is a major setback in using a GPRN is its implementation. Some computational improvements I believe need to be addressed. Due to its intractability, a GPRN approximates the posterior distributions via mean-field inference. I cannot exclude the possibility that other variation inference methods could not improve our results. Another variation method could improve our current model (one node and one weight per dataset) and help us obtain more precise model evidences. The current GPRN python package is also highly inefficient. The mathematical complexity and amount of matrices inversions required is a major setback. This is not necessarily easy to solve, but one way of speeding things would be writing the GPRN with another programming language (e.g. C++).

Appendices

Appendix A

Datasets summary

Star	Dataset	N	Time span (days)	PTP amplitude (m/s)	Average error (m/s)	RMS (m/s)
Sun	RV	497	1094	10.364	0.133	1.982
Sun	BIS	497	1094	11.900	0.265	1.954
Sun	FWHM	497	1094	15.160	0.265	2.244
Sun	$\log R'_{hk}$	497	1094	0.064	0.001	0.0127
HD 10700	RV	174	470	10.328	0.370	1.884
HD 10700	FWHM	174	470	98.257	0.739	11.758
HD 34411	RV	188	416	10.744	0.337	1.767
HD 34411	FWHM	188	416	49.294	0.674	5.244
HD 101501	RV	45	655	17.881	0.353	4.920
HD 101501	FWHM	45	655	97.777	0.707	24.480
HD 26965	RV	114	465	14.144	0.334	3.234
HD 26965	FWHM	114	465	69.244	0.667	15.792

Table A.1: Summary with the number of points, time spans, peak-to-peak amplitude, average error, and initial RMS of the datasets analysed on this thesis.

Appendix B

Corner plots

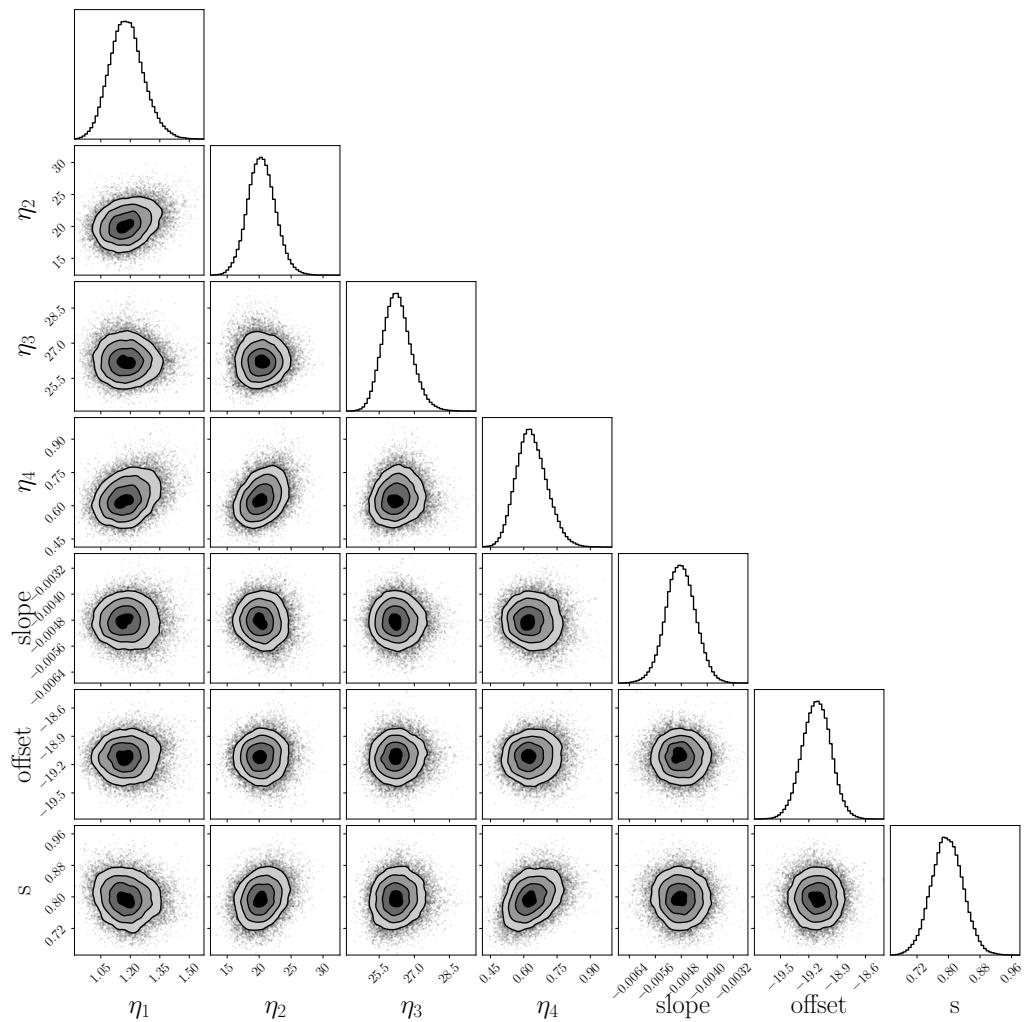


Figure B.1: Posterior distributions of the parameter of the GP used in the individual analysis of the Sun's RV measurements.

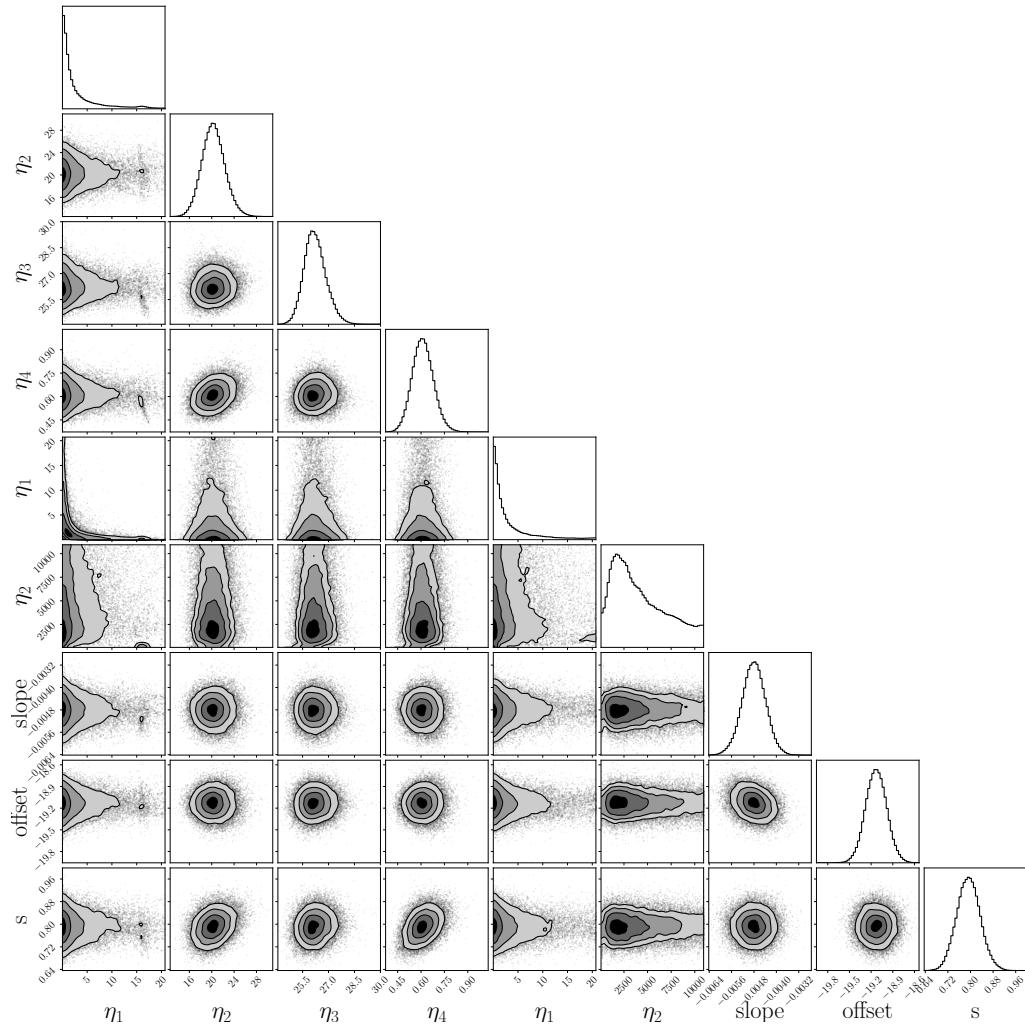


Figure B.2: Posterior distributions of the parameter of the GPRN used in the individual analysis of the Sun's RV measurements.

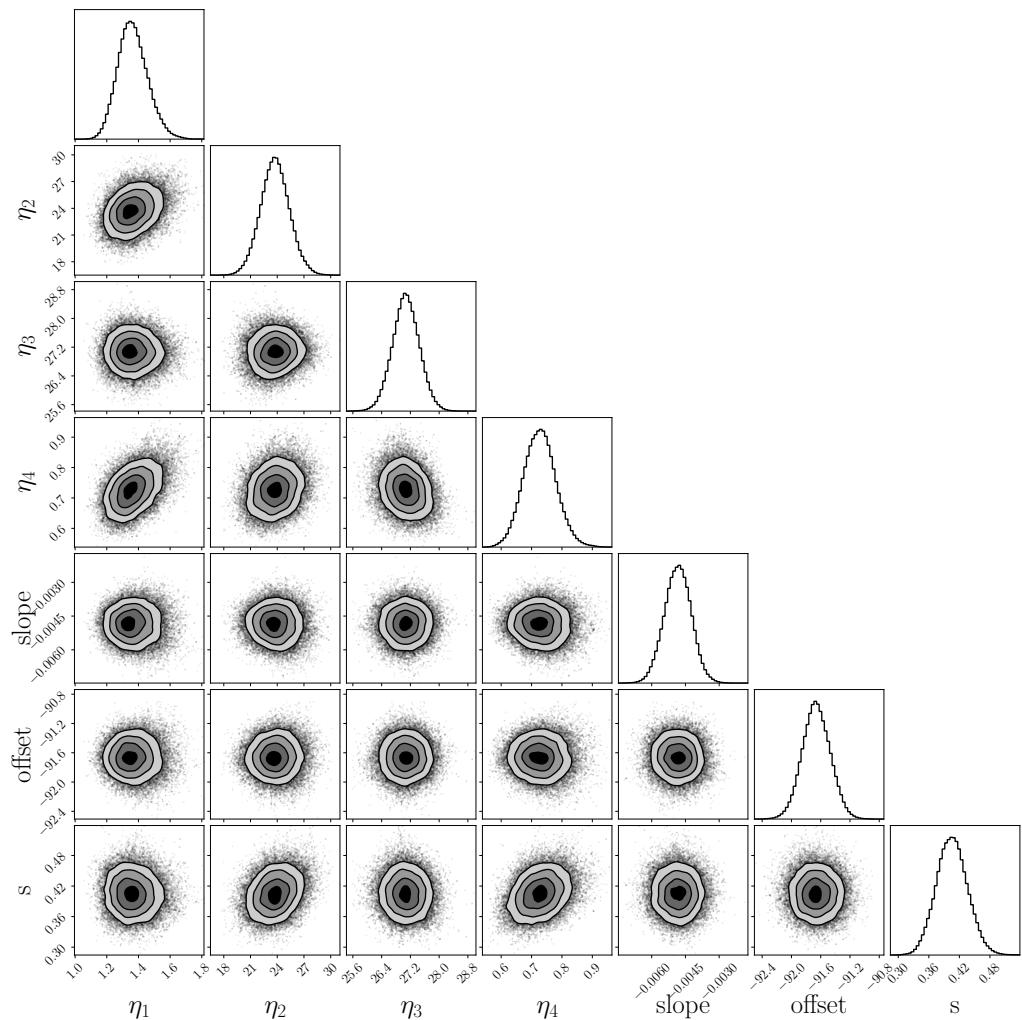


Figure B.3: Posterior distributions of the parameter of the GP used in the individual analysis of the Sun's BIS measurements.

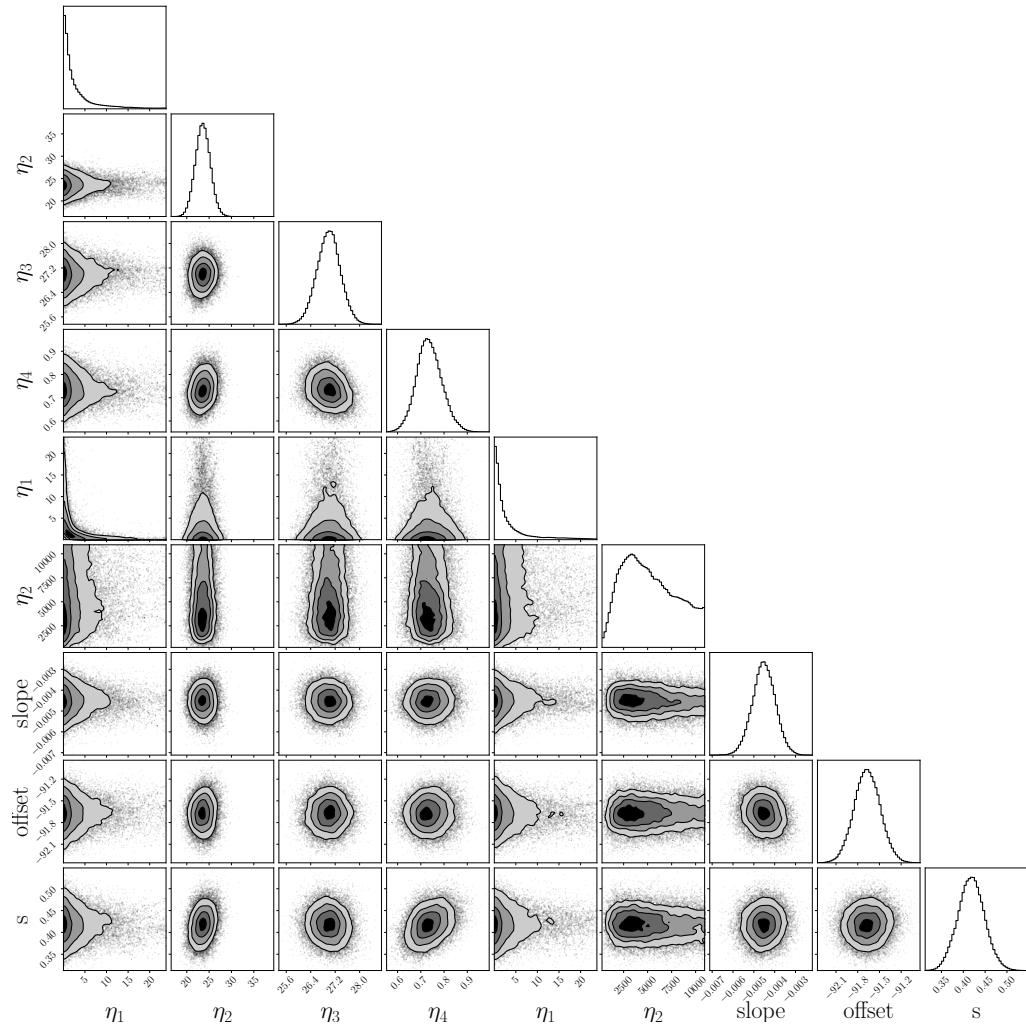


Figure B.4: Posterior distributions of the parameter of the GPRN used in the individual analysis of the Sun's BIS measurements.

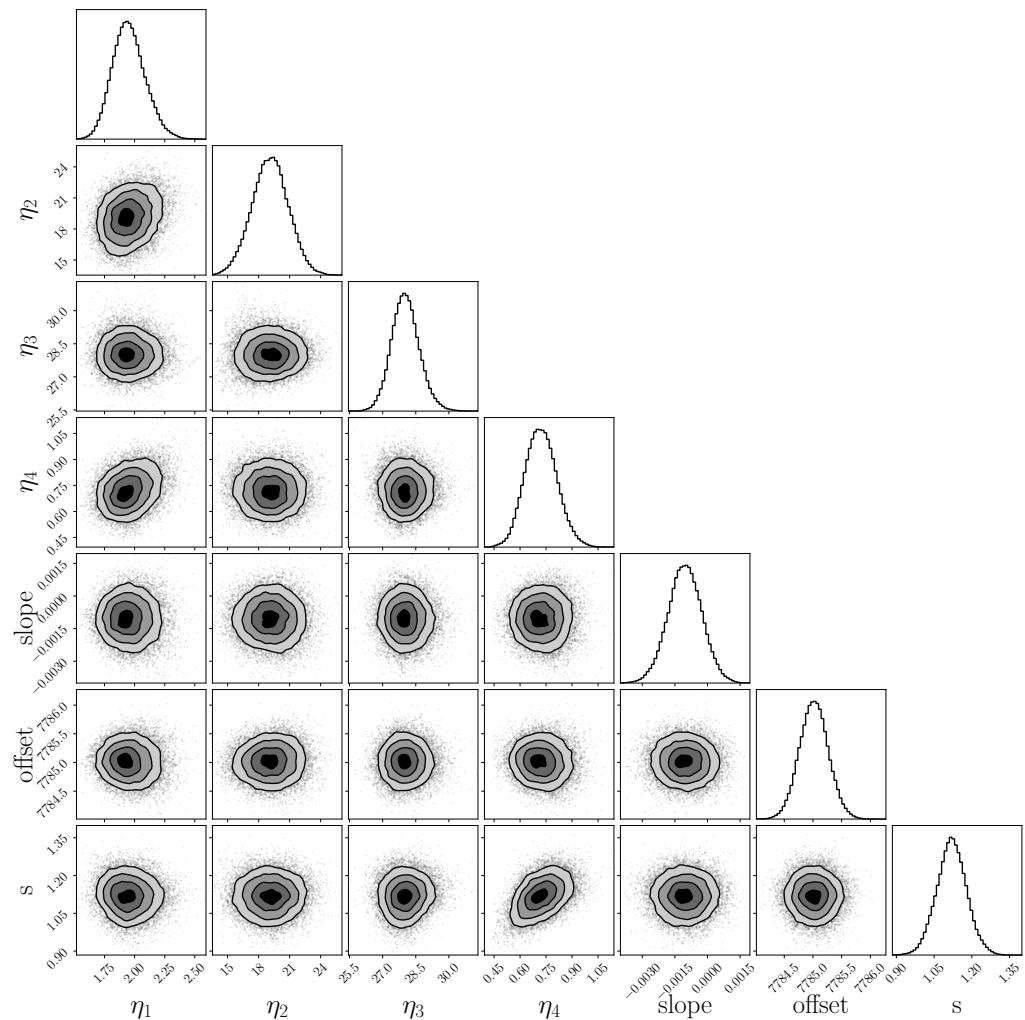


Figure B.5: Posterior distributions of the parameter of the GP used in the individual analysis of the Sun's FWHM measurements.

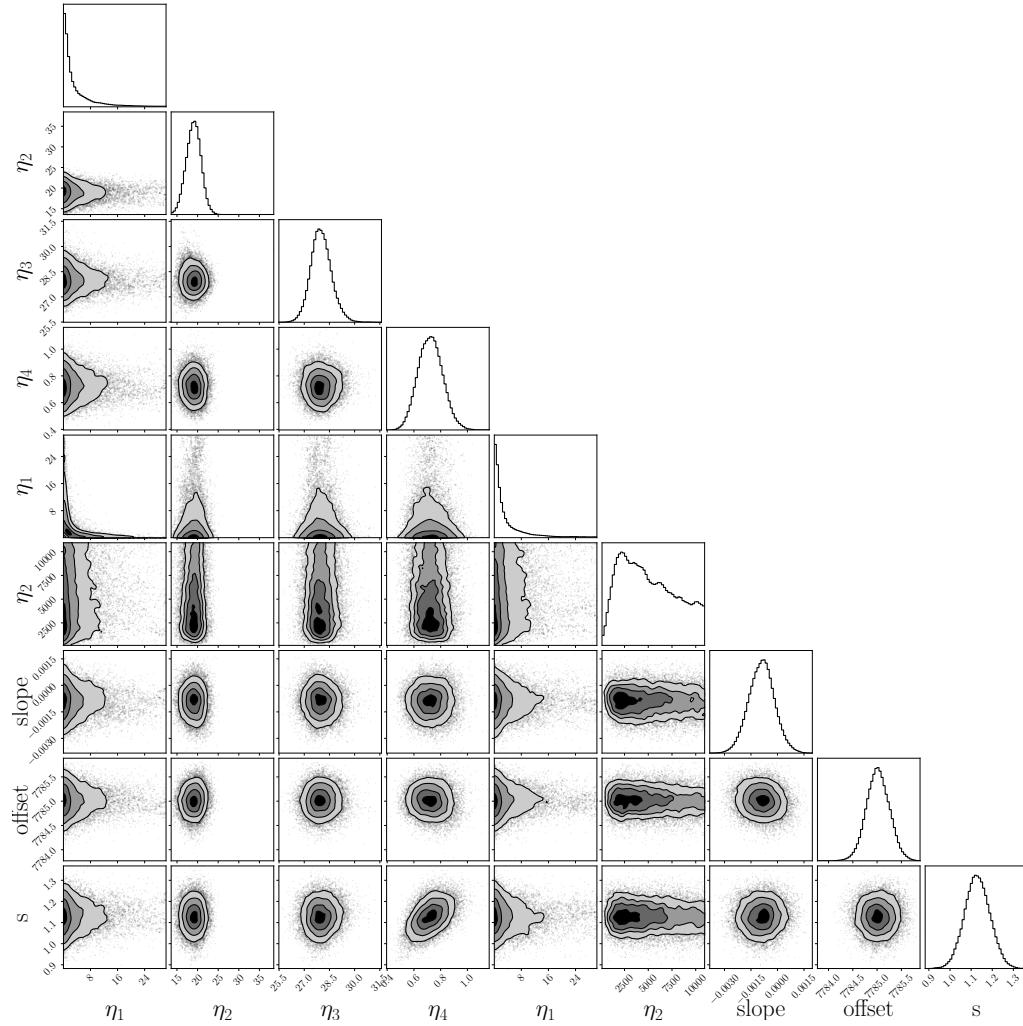


Figure B.6: Posterior distributions of the parameter of the GPRN used in the individual analysis of the Sun's FWHM measurements.

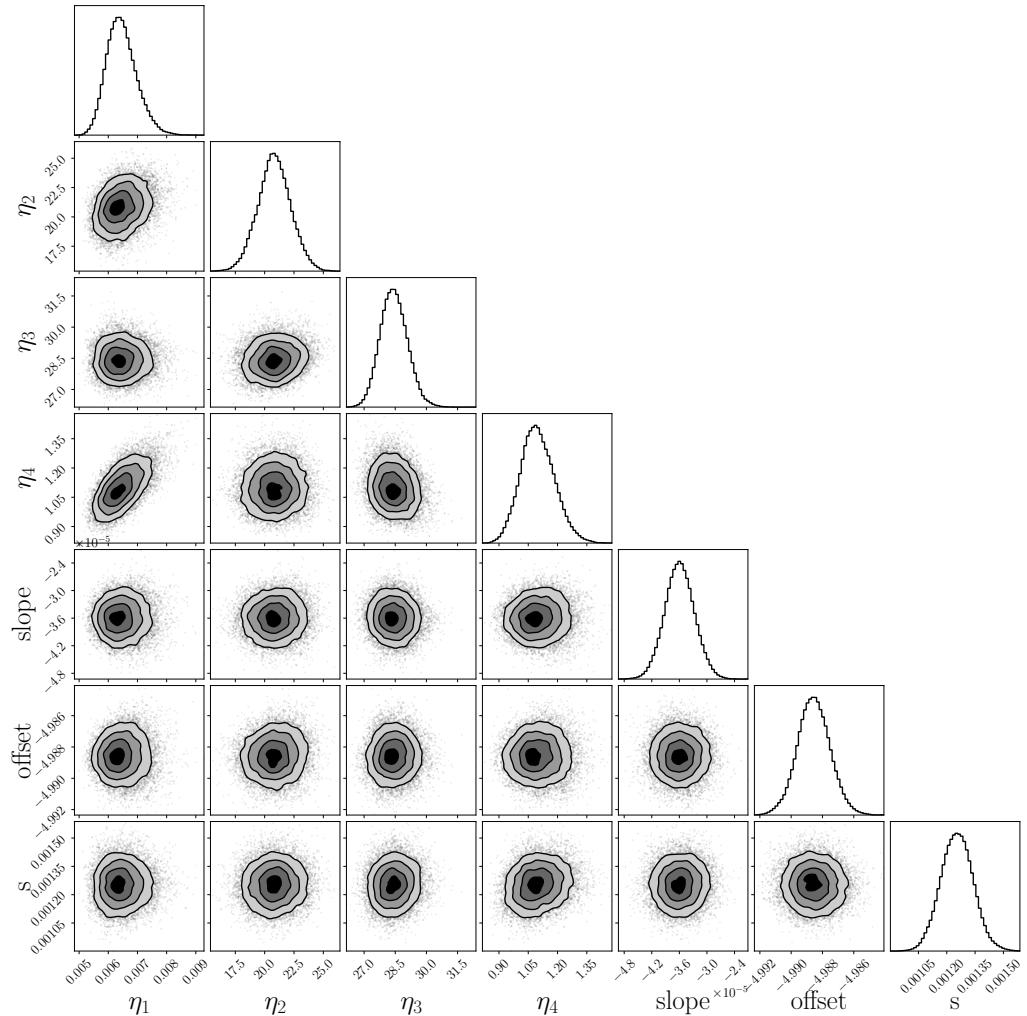


Figure B.7: Posterior distributions of the parameter of the GP used in the individual analysis of the Sun's $\log R'_{hk}$ measurements.

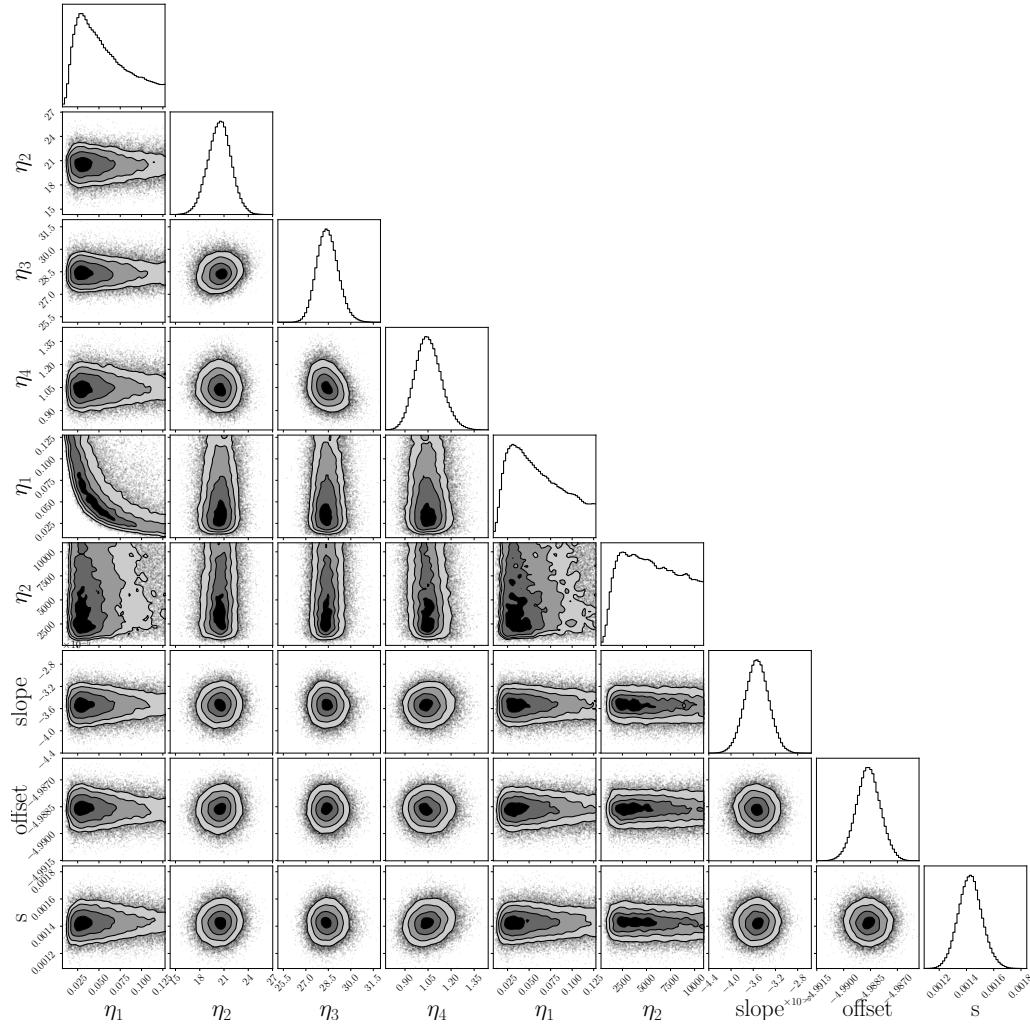


Figure B.8: Posterior distributions of the parameter of the GPRN used in the individual analysis of the Sun's $\log R'_{hk}$ measurements.

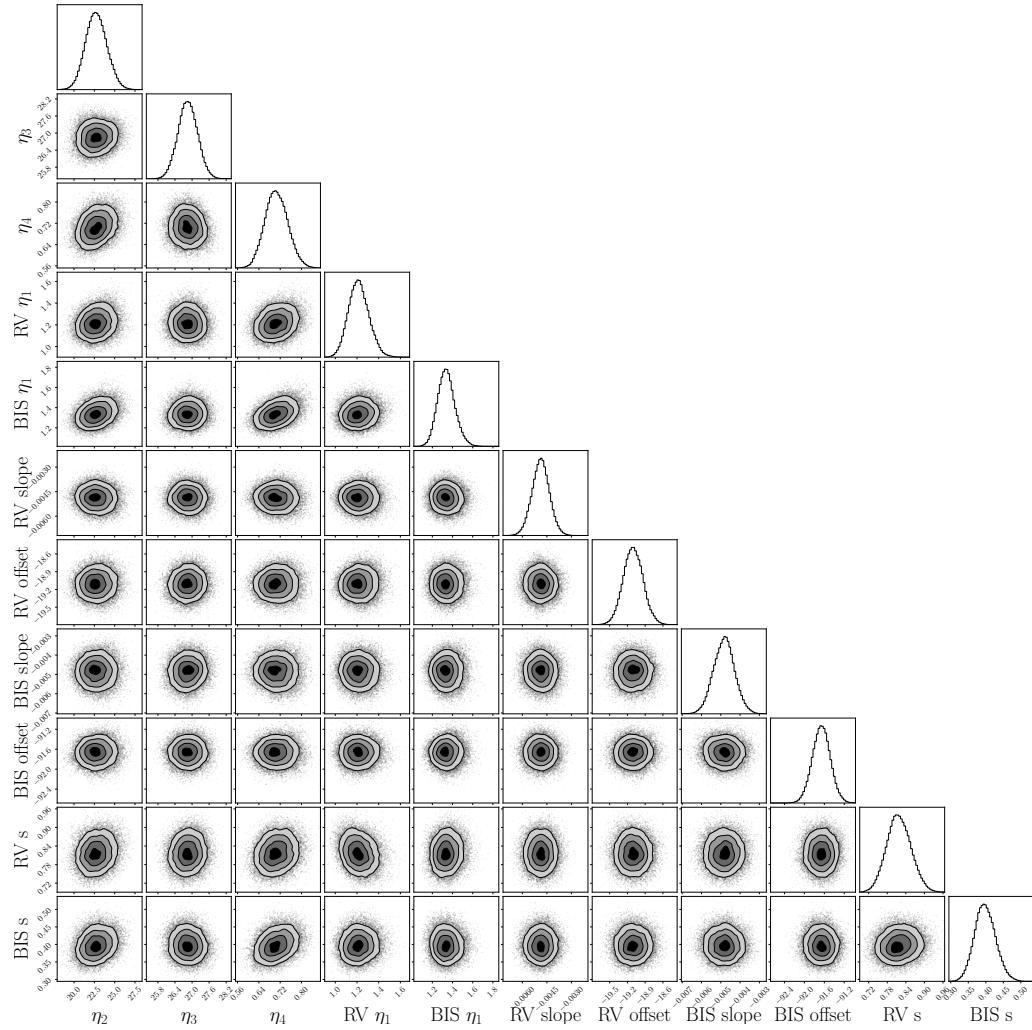


Figure B.9: Posterior distributions of the parameter of the GP used in the combined analysis of the Sun's RV and BIS measurements.

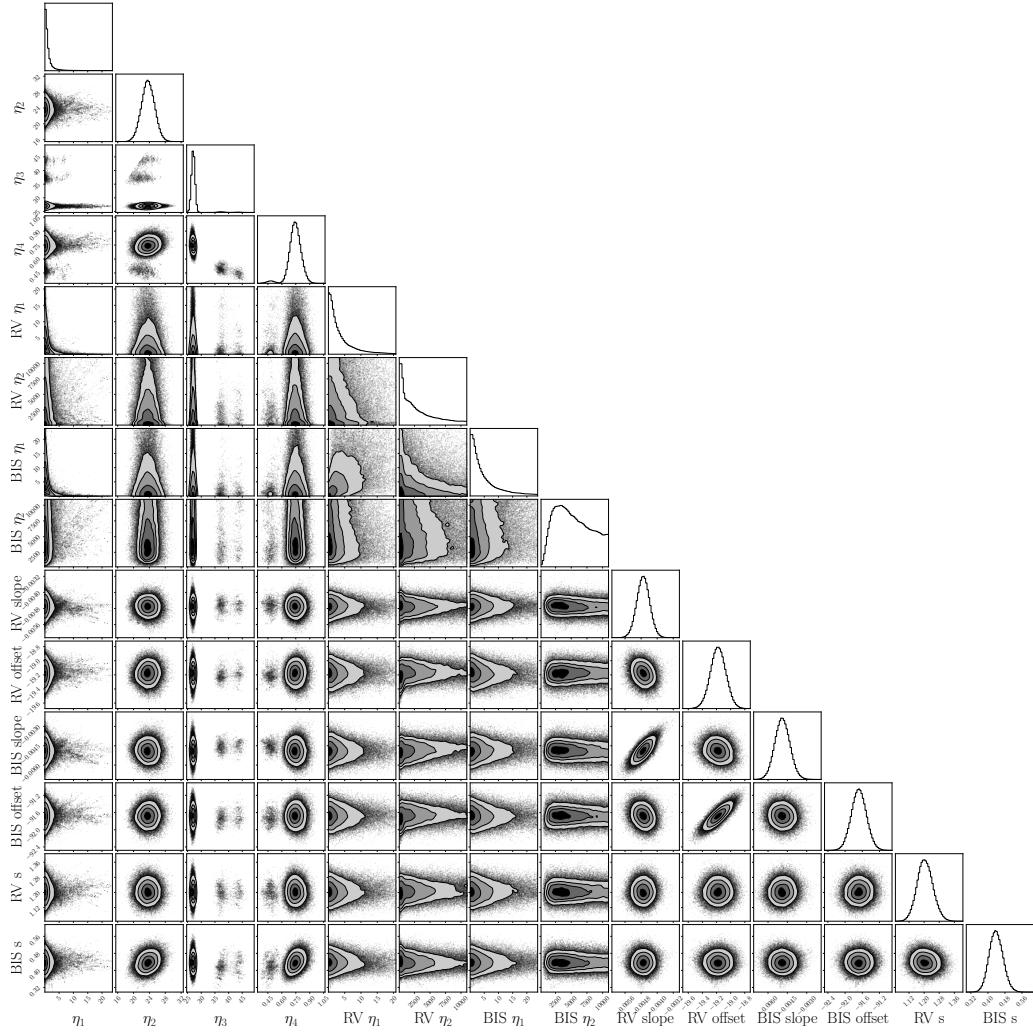


Figure B.10: Posterior distributions of the parameter of the GPRN used in the combined analysis of the Sun's RV and BIS measurements.

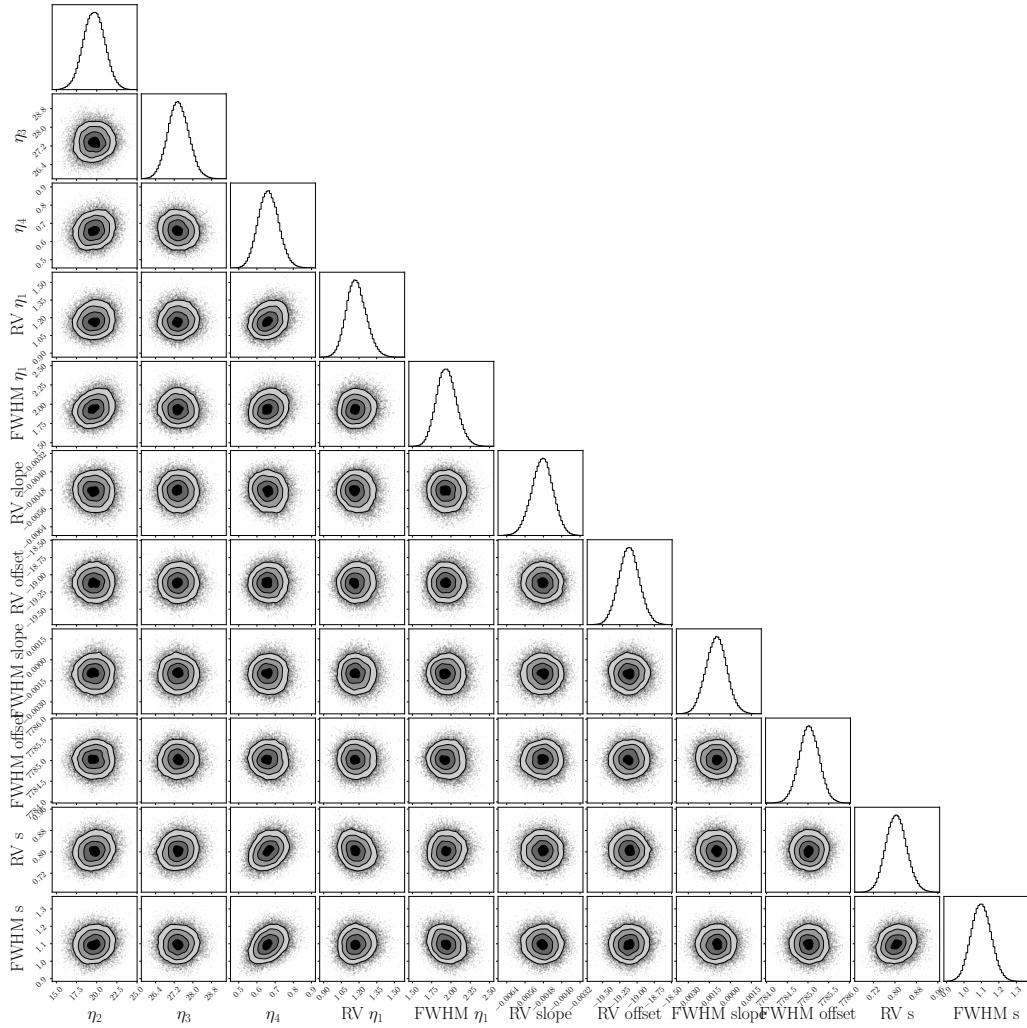


Figure B.11: Posterior distributions of the parameter of the GP used in the combined analysis of the Sun's RV and FWHM measurements.

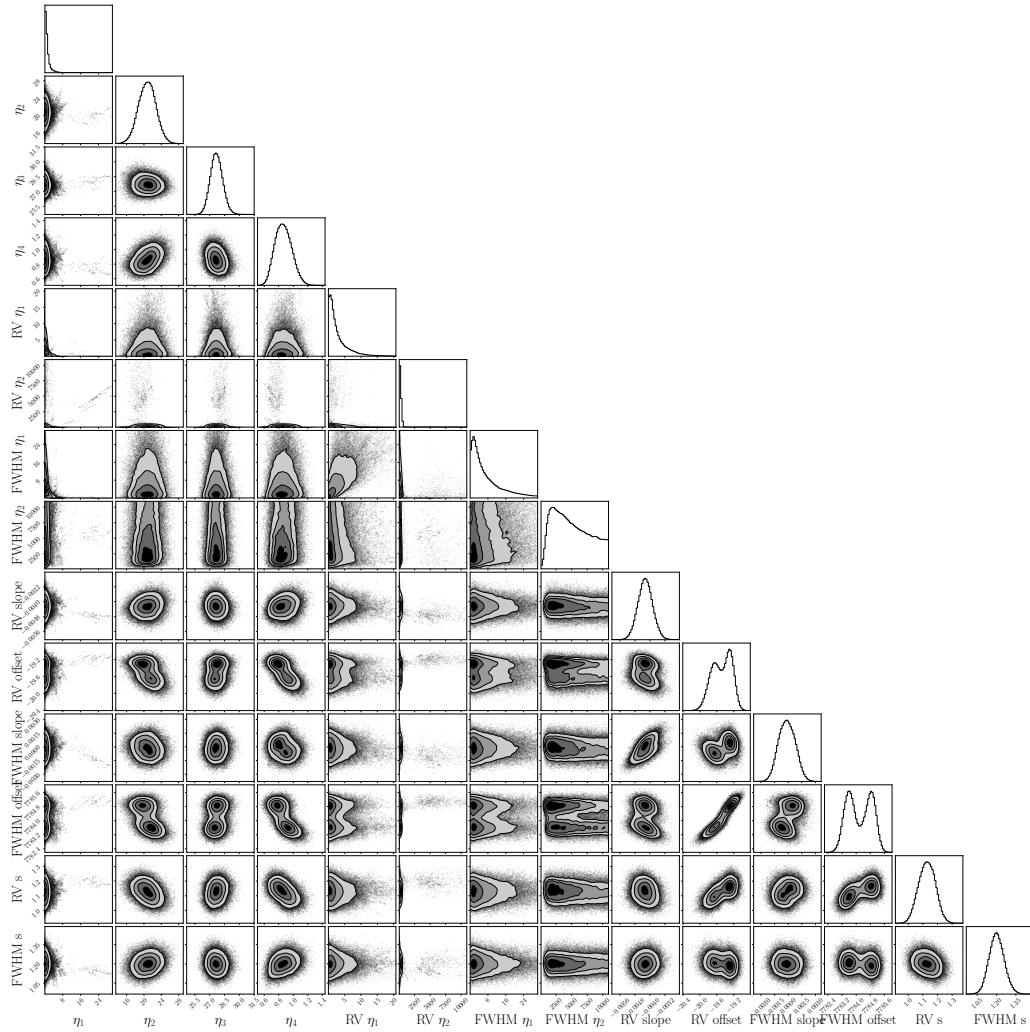


Figure B.12: Posterior distributions of the parameter of the GPRN used in the combined analysis of the Sun's RV and FWHM measurements.

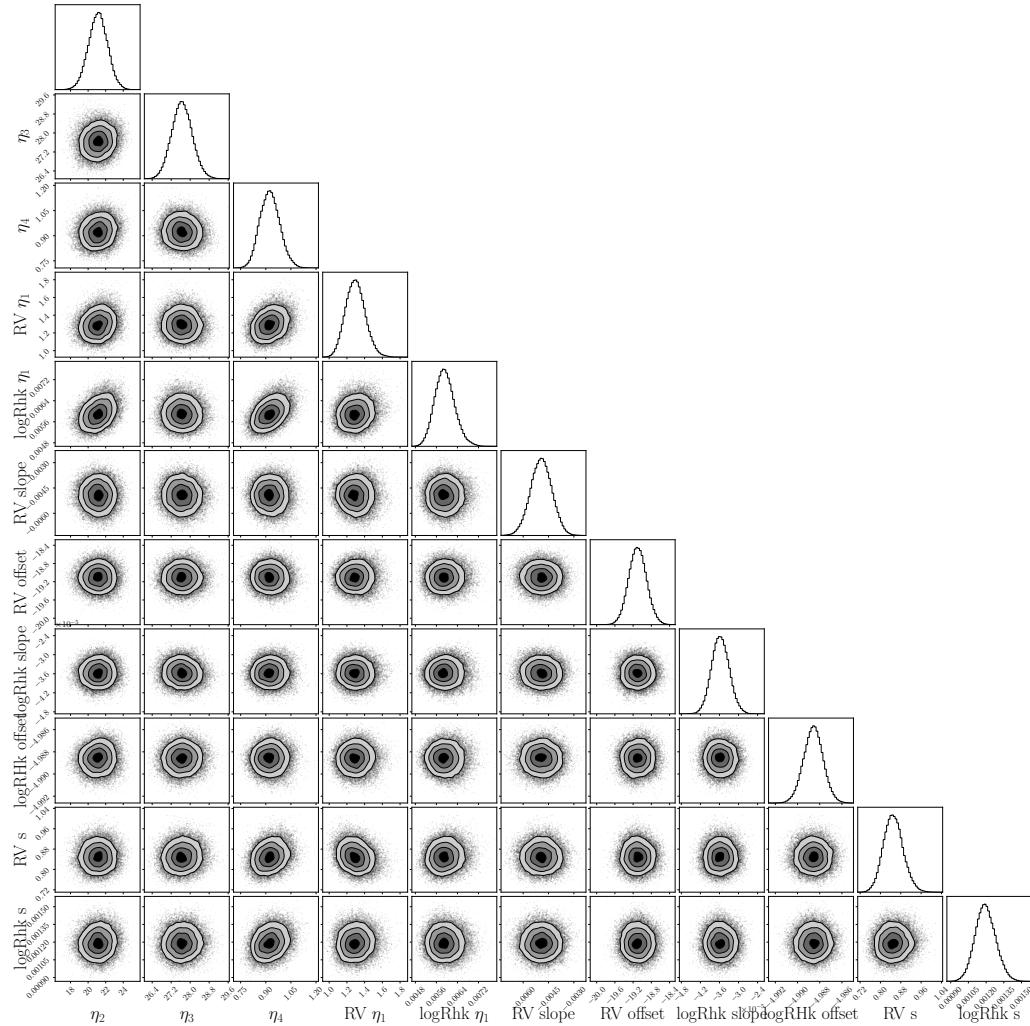


Figure B.13: Posterior distributions of the parameter of the GP used in the combined analysis of the Sun's RV and $\log R'_{hk}$ measurements.

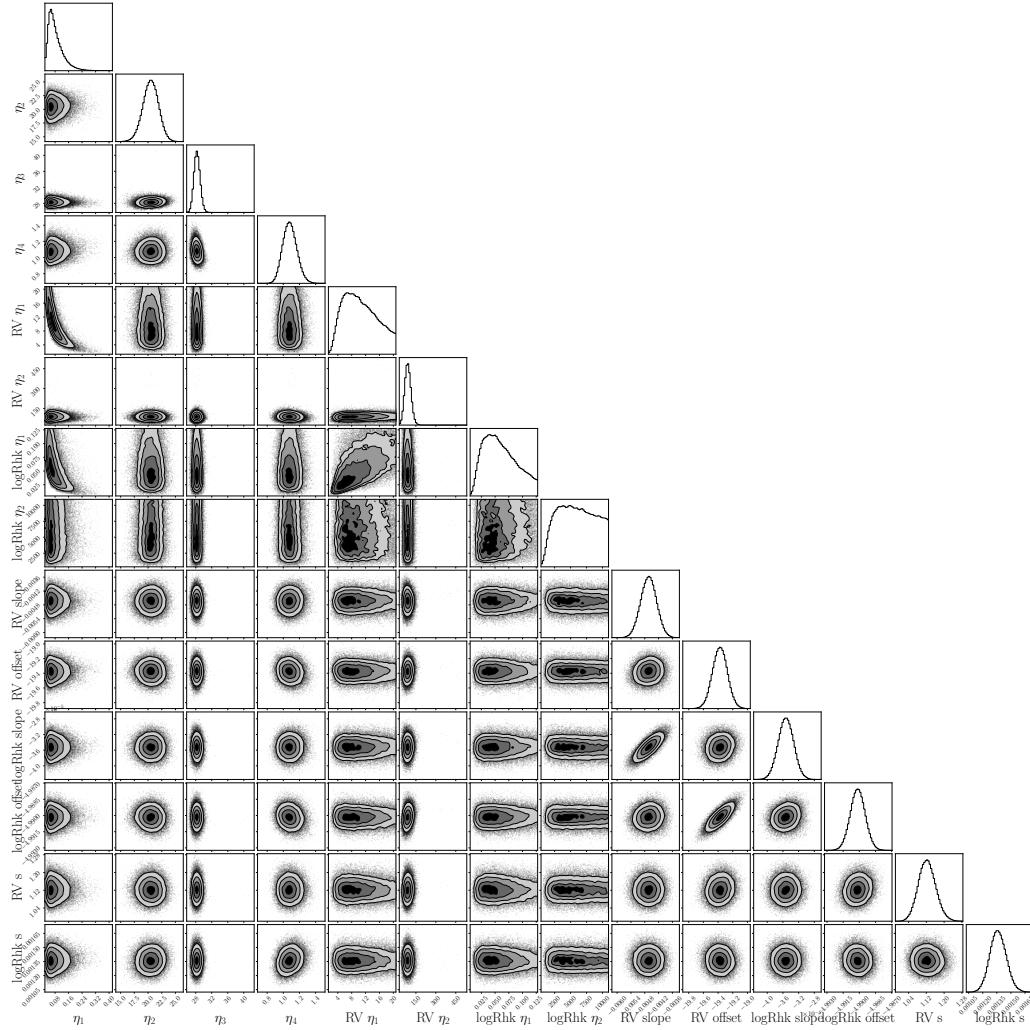


Figure B.14: Posterior distributions of the parameter of the GPRN used in the combined analysis of the Sun's RV and $\log R'_{hk}$ measurements.

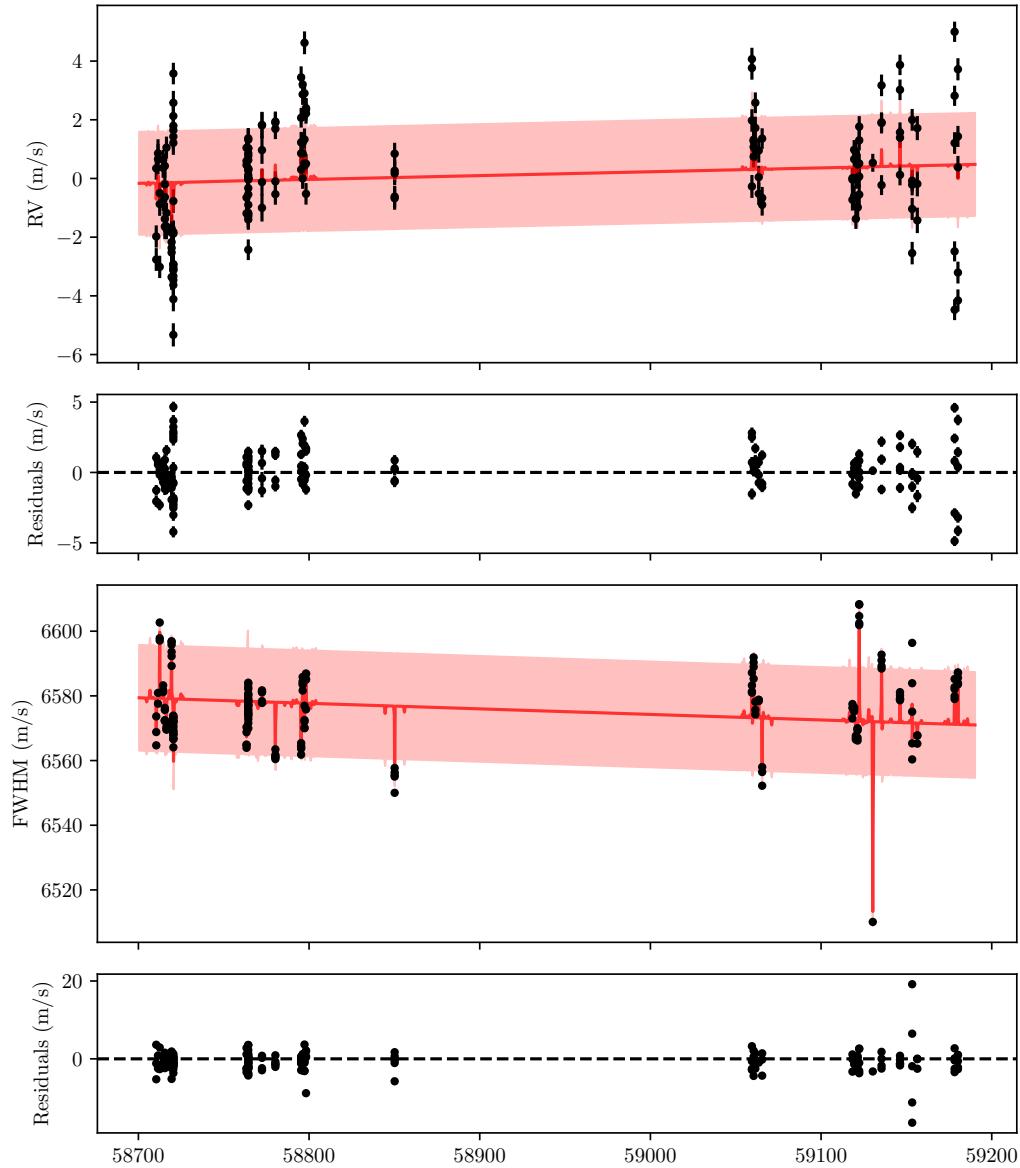


Figure B.15: Fit and respective residuals of the MAP values from the GP on the RV and the FWHM time series of HD 10700.

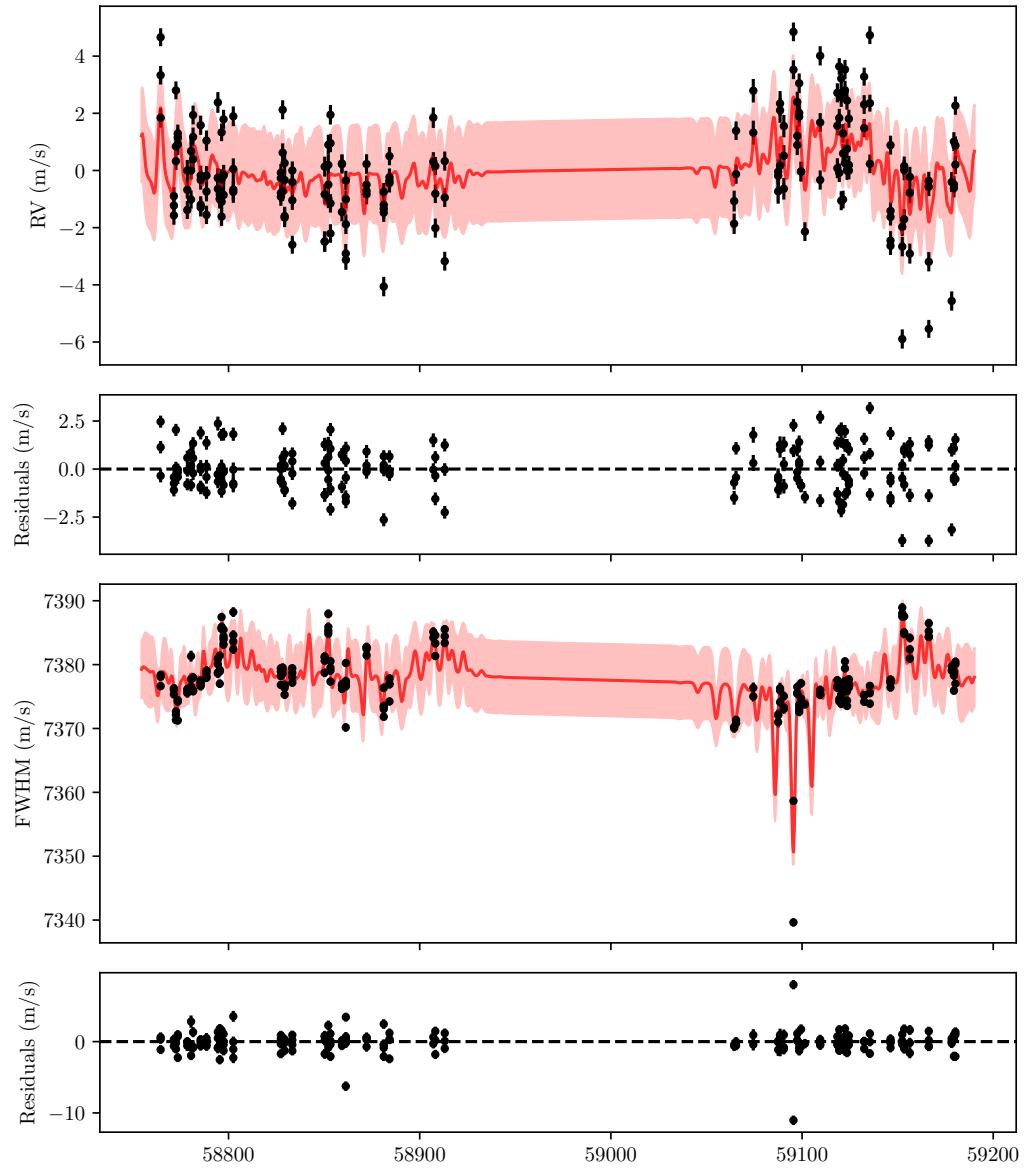


Figure B.16: Fit and respective residuals of the MAP values from the GP on the RV and the FWHM time series of HD 34411.

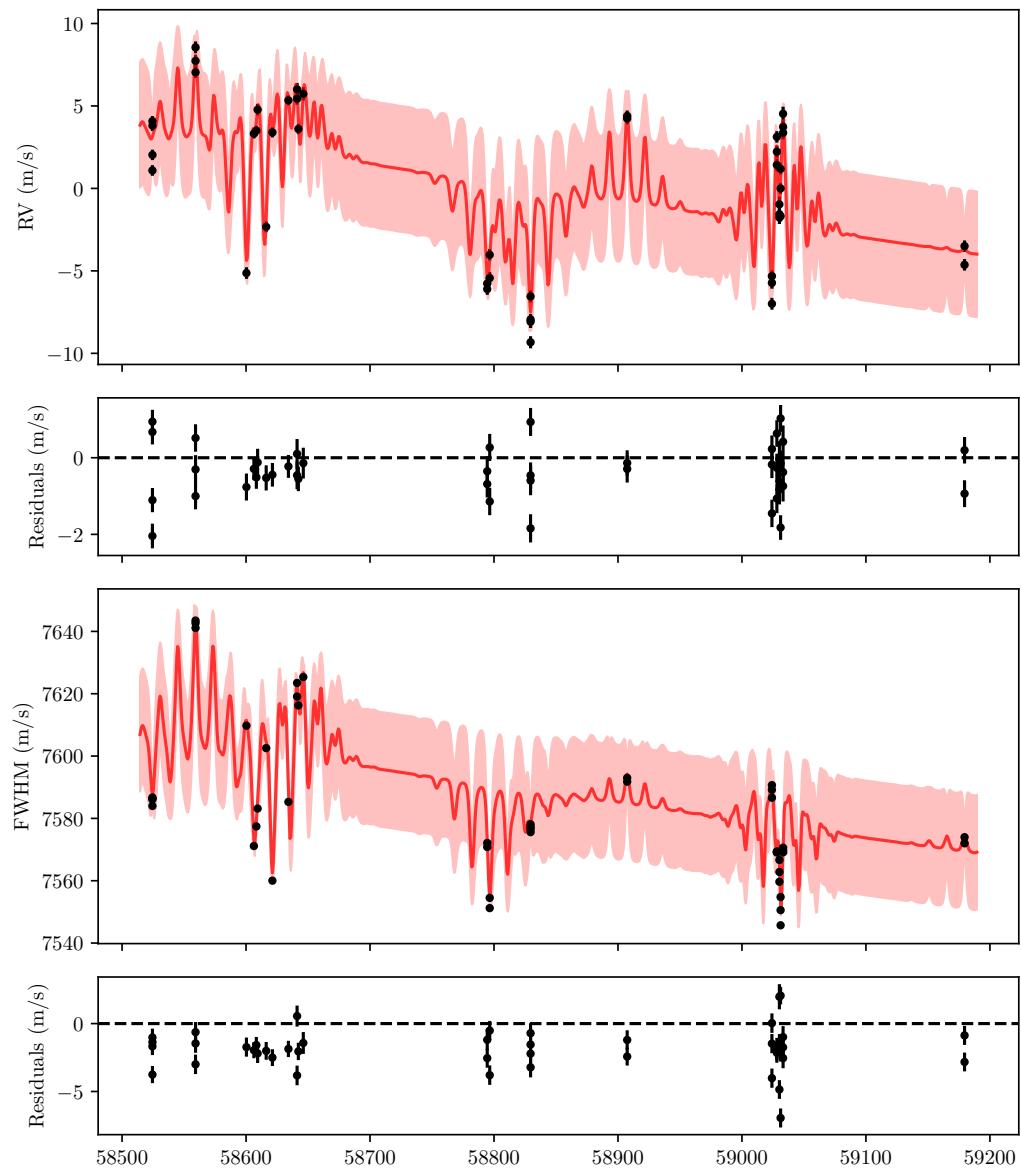


Figure B.17: Fit and respective residuals of the MAP values from the GP on the RV and the FWHM time series of HD 101501.

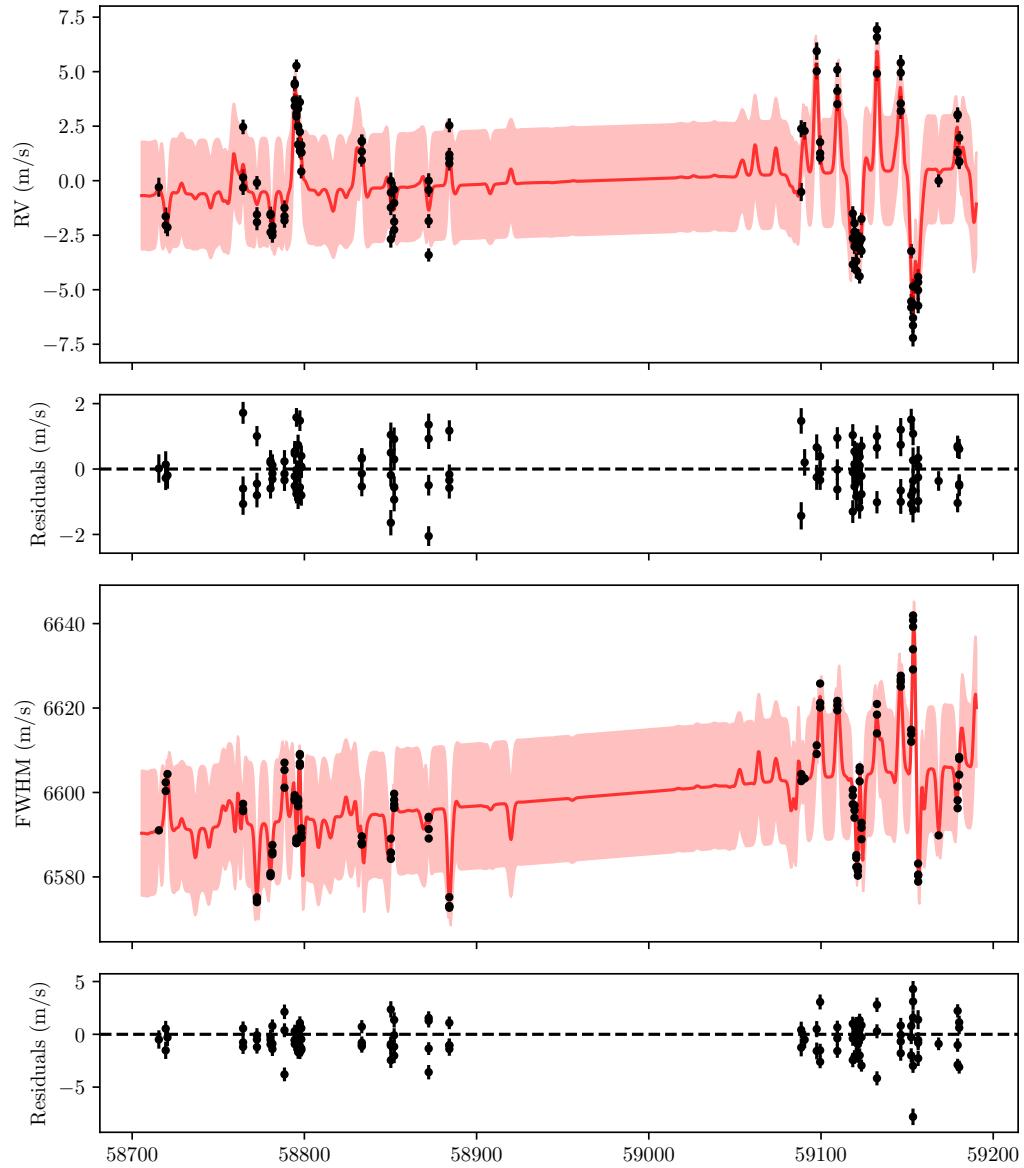


Figure B.18: Fit and respective residuals of the MAP values from the GP on the RV and the FWHM time series of HD 26965.

Bibliography

- Abdessalem, A. B. et al. (2017). “Automatic Kernel Selection for Gaussian Processes Regression with Approximate Bayesian Computation and Sequential Monte Carlo”. In: *Frontiers in Built Environment* 3. doi: [10.3389/fbuil.2017.00052](https://doi.org/10.3389/fbuil.2017.00052).
- Abramowitz, M. and I. Stegun (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Applied mathematics series. U.S. Government Printing Office.
- Adibekyan, V. Z. et al. (2013). “Orbital and physical properties of planets and their hosts: new insights on planet formation and evolution”. In: *A&A* 560, A51. doi: [10.1051/0004-6361/201322551](https://doi.org/10.1051/0004-6361/201322551).
- Adibekyan, V. (2019). “Heavy Metal Rules. I. Exoplanet Incidence and Metallicity”. In: *Geosciences* 9(3). doi: [10.3390/geosciences9030105](https://doi.org/10.3390/geosciences9030105).
- Aigrain, S. et al. (2012). “A simple method to estimate radial velocity variations due to stellar activity using photometry*”. In: *Monthly Notices of the Royal Astronomical Society* 419(4). doi: [10.1111/j.1365-2966.2011.19960.x](https://doi.org/10.1111/j.1365-2966.2011.19960.x).
- Allende Prieto, C. (2007). “Velocities from Cross-Correlation: A Guide for Self-Improvement”. In: *??jnlAJ* 134(5). doi: [10.1086/522051](https://doi.org/10.1086/522051).
- Arentoft, T. et al. (2008). “A Multisite Campaign to Measure Solar-like Oscillations in Procyon. I. Observations, Data Reduction, and Slow Variations”. In: *ApJ* 687(2). doi: [10.1086/592040](https://doi.org/10.1086/592040).
- Arfken, G. et al. (2013). *Mathematical Methods for Physicists: A Comprehensive Guide*. Elsevier Science. ISBN: 9780123846549.
- Artin, E. and M. Butler (2015). *The Gamma Function*. Dover Books on Mathematics. Dover Publications. ISBN: 9780486803005.

- Astropy Collaboration et al. (2018). “The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package”. In: ??*jnlAJ* 156(3), 123. doi: [10.3847/1538-3881/aabc4f](https://doi.org/10.3847/1538-3881/aabc4f).
- Bahng, J. and M. Schwarzschild (1961). “Lifetime of Solar Granules.” In: *ApJ* 134. doi: [10.1086/147160](https://doi.org/10.1086/147160).
- Baliunas, S. L. et al. (1998). “Activity Cycles in Lower Main Sequence and POST Main Sequence Stars: The HK Project”. In: *Cool Stars, Stellar Systems, and the Sun*. Ed. by R. A. Donahue and J. A. Bookbinder. Vol. 154. Astronomical Society of the Pacific Conference Series.
- Baranne, A. et al. (1996). “ELODIE: A spectrograph for accurate radial velocity measurements”. In: *Astron. Astrophys. Suppl. Ser.* 119(2). doi: [10.1051/aas:1996251](https://doi.org/10.1051/aas:1996251).
- Barros, S. C. C. et al. (2020). “Improving transit characterisation with Gaussian process modelling of stellar variability”. In: *A&A* 634, A75. doi: [10.1051/0004-6361/201936086](https://doi.org/10.1051/0004-6361/201936086).
- Bedding, T. R. and H. Kjeldsen (2003). “Solar-like Oscillations”. In: *Publications of the Astronomical Society of Australia* 20(2). doi: [10.1071/AS03025](https://doi.org/10.1071/AS03025).
- Bedding, T. R. and H. Kjeldsen (2007). “Observations of solar-like oscillations”. In: *Communications in Asteroseismology* 150. doi: [10.1553/cia150s106](https://doi.org/10.1553/cia150s106).
- Berger, T. E. et al. (2007). “What are ‘Faculae’?” In: *New Solar Physics with Solar-B Mission*. Ed. by K. Shibata et al. Vol. 369. Astronomical Society of the Pacific Conference Series.
- Beuzit, J.-L. et al. (2008). “SPHERE: a ‘Planet Finder’ instrument for the VLT”. In: *Ground-based and Airborne Instrumentation for Astronomy II*. Ed. by I. S. McLean and M. M. Casali. Vol. 7014. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. doi: [10.1117/12.790120](https://doi.org/10.1117/12.790120).
- Bishop, C. (2006). *Pattern Recognition and Machine Learning*. Information Science and Statistics. Springer. ISBN: 9780387310732.
- Blei, D. M. et al. (2017). “Variational Inference: A Review for Statisticians”. In: *Journal of the American Statistical Association* 112(518). doi: [10.1080/01621459.2017.1285773](https://doi.org/10.1080/01621459.2017.1285773).
- Boisse, I. et al. (2010). “Disentangling between stellar activity and planetary signals”. In: *Proceedings of the International Astronomical Union* 6(S273). doi: [10.1017/S1743921311015389](https://doi.org/10.1017/S1743921311015389).

- Bond, I. A. et al. (2004). “OGLE 2003-BLG-235/MOA 2003-BLG-53: A Planetary Microlensing Event”. In: *ApJ* 606(2). doi: [10.1086/420928](https://doi.org/10.1086/420928).
- Borgniet, S. et al. (2019). “Consistent radial velocities of classical Cepheids from the cross-correlation technique”. In: *A&A* 631, A37. doi: [10.1051/0004-6361/201935622](https://doi.org/10.1051/0004-6361/201935622).
- Borucki, W. et al. (2009). “KEPLER: Search for Earth-Size Planets in the Habitable Zone”. In: *Transiting Planets*. Ed. by F. Pont et al. Vol. 253. doi: [10.1017/S1743921308026513](https://doi.org/10.1017/S1743921308026513).
- Bouchy, F. et al. (2001). “Fundamental photon noise limit to radial velocity measurements”. In: *A&A* 374. doi: [10.1051/0004-6361:20010730](https://doi.org/10.1051/0004-6361:20010730).
- Boyajian, T. S. et al. (2012). “Stellar Diameters and Temperatures. I. Main-sequence A, F, and G Stars”. In: *ApJ* 746(1), 101. doi: [10.1088/0004-637X/746/1/101](https://doi.org/10.1088/0004-637X/746/1/101).
- Broomhall, A. M. et al. (2011). “Solar-cycle variations of large frequency separations of acoustic modes: implications for asteroseismology”. In: *MNRAS* 413(4). doi: [10.1111/j.1365-2966.2011.18375.x](https://doi.org/10.1111/j.1365-2966.2011.18375.x).
- Brown, R. A. and C. J. Burrows (1990). “On the feasibility of detecting extrasolar planets by reflected starlight using the Hubble Space Telescope”. In: 87(2). doi: [10.1016/0019-1035\(90\)90150-8](https://doi.org/10.1016/0019-1035(90)90150-8).
- Browning, M. K. et al. (2010). “Rotation and Magnetic Activity in a Sample of M-Dwarfs”. In: ??*jnlAJ* 139(2). doi: [10.1088/0004-6256/139/2/504](https://doi.org/10.1088/0004-6256/139/2/504).
- Bryan, M. L. et al. (2016). “Statistics of Long Period Gas Giant Planets in Known Planetary Systems”. In: *ApJ* 821(2), 89. doi: [10.3847/0004-637X/821/2/89](https://doi.org/10.3847/0004-637X/821/2/89).
- Butler, R. P. et al. (1996). “Attaining Doppler Precision of 3 M s-1”. In: *PASP* 108. doi: [10.1086/133755](https://doi.org/10.1086/133755).
- Butler, R. P. et al. (1997). “Three New “51 Pegasi-Type” Planets”. In: *ApJ* 474(2). doi: [10.1086/310444](https://doi.org/10.1086/310444).
- Butler, R. P. et al. (2004). “A Neptune-Mass Planet Orbiting the Nearby M Dwarf GJ 436”. In: *ApJ* 617(1). doi: [10.1086/425173](https://doi.org/10.1086/425173).
- Campbell, B. et al. (1988). “A Search for Substellar Companions to Solar-type Stars”. In: *ApJ* 331. doi: [10.1086/166608](https://doi.org/10.1086/166608).

- Cegla, H. (2019). “The Impact of Stellar Surface Magnetoconvection and Oscillations on the Detection of Temperate, Earth-Mass Planets Around Sun-Like Stars”. In: *Geosciences* 9(3). doi: [10.3390/geosciences9030114](https://doi.org/10.3390/geosciences9030114).
- Chaplin, W. J. et al. (2019). “Filtering Solar-Like Oscillations for Exoplanet Detection in Radial Velocity Observations”. In: *??jnlAJ* 157(4), 163. doi: [10.3847/1538-3881/ab0c01](https://doi.org/10.3847/1538-3881/ab0c01).
- Charbonneau, D. et al. (2000). “Detection of Planetary Transits Across a Sun-like Star”. In: *ApJ* 529(1). doi: [10.1086/312457](https://doi.org/10.1086/312457).
- Charbonneau, D. et al. (2009). “A super-Earth transiting a nearby low-mass star”. In: *Nature* 462(7275). doi: [10.1038/nature08679](https://doi.org/10.1038/nature08679).
- Charbonneau, P. (2014). “Solar Dynamo Theory”. In: *Annual Review of Astronomy and Astrophysics* 52(1). doi: [10.1146/annurev-astro-081913-040012](https://doi.org/10.1146/annurev-astro-081913-040012).
- Charbonneau, P. (2020). “Dynamo models of the solar cycle”. In: *Living Reviews in Solar Physics* 17(1), 4. doi: [10.1007/s41116-020-00025-6](https://doi.org/10.1007/s41116-020-00025-6).
- Chauvin, G. et al. (2004). “A giant planet candidate near a young brown dwarf. Direct VLT/NACO observations using IR wavefront sensing”. In: *A&A* 425. doi: [10.1051/0004-6361:200400056](https://doi.org/10.1051/0004-6361:200400056).
- Choi, J. et al. (2013). “Precise Doppler Monitoring of Barnard’s Star”. In: *ApJ* 764(2), 131. doi: [10.1088/0004-637X/764/2/131](https://doi.org/10.1088/0004-637X/764/2/131).
- Cloutier, R. et al. (2019). “Confirmation of the radial velocity super-Earth K2-18c with HARPS and CARMENES”. In: *A&A* 621. doi: [10.1051/0004-6361/201833995](https://doi.org/10.1051/0004-6361/201833995).
- Collier Cameron, A et al. (2019). “Three years of Sun-as-a-star radial-velocity observations on the approach to solar minimum”. In: *Monthly Notices of the Royal Astronomical Society* 487(1). doi: [10.1093/mnras/stz1215](https://doi.org/10.1093/mnras/stz1215).
- Cosentino, R. et al. (2012). “Harps-N: the new planet hunter at TNG”. In: *Ground-based and Airborne Instrumentation for Astronomy IV*. Ed. by I. S. McLean et al. Vol. 8446. International Society for Optics and Photonics. SPIE.
- Cui, G. et al. (2016). “Exact Distribution for the Product of Two Correlated Gaussian Random Variables”. In: *IEEE Signal Processing Letters* 23(11). doi: [10.1109/LSP.2016.2614539](https://doi.org/10.1109/LSP.2016.2614539).
- Cumming, A. et al. (2008). “The Keck Planet Search: Detectability and the Minimum Mass and Orbital Period Distribution of Extrasolar Planets”. In: *PASP* 120(867). doi: [10.1086/588487](https://doi.org/10.1086/588487).

- D'Angelo, G. et al. (2003). "Orbital Migration and Mass Accretion of Protoplanets in Three-dimensional Global Computations with Nested Grids". In: *ApJ* 586(1). doi: [10.1086/367555](https://doi.org/10.1086/367555).
- Davies, A. (2015). "Effective Implementation of Gaussian Process Regression for Machine Learning". PhD thesis. University of Cambridge.
- Deeg, H. J. and R. Alonso (2018). "Transit Photometry as an Exoplanet Discovery Method". In: *Handbook of Exoplanets*. Ed. by H. J. Deeg and J. A. Belmonte. doi: [10.1007/978-3-319-55333-7\117](https://doi.org/10.1007/978-3-319-55333-7_117).
- Del Moro, D. (2004). "Solar granulation properties derived from three different time series". In: *A&A* 428. doi: [10.1051/0004-6361:20040466](https://doi.org/10.1051/0004-6361:20040466).
- Del Zanna, G. and H. Mason (2013). "The Sun as a Star". In: *Planets, Stars and Stellar Systems: Volume 4: Stellar Structure and Evolution*. Ed. by T. D. Oswalt and M. A. Barstow. Springer Netherlands: Dordrecht. ISBN: 978-94-007-5615-1. doi: [10.1007/978-94-007-5615-1_3](https://doi.org/10.1007/978-94-007-5615-1_3).
- Delfosse, X. et al. (1998). "Rotation and chromospheric activity in field M dwarfs". In: *A&A* 331.
- Díaz, M. R. et al. (2018). "The Test Case of HD 26965: Difficulties Disentangling Weak Doppler Signals from Stellar Activity". In: ??*jnlAJ* 155(3), 126. doi: [10.3847/1538-3881/aaa896](https://doi.org/10.3847/1538-3881/aaa896).
- Doyle, L. R. et al. (2011). "Kepler-16: A Transiting Circumbinary Planet". In: *Science* 333(6049). doi: [10.1126/science.1210923](https://doi.org/10.1126/science.1210923).
- Ducourant, C. et al. (2008). "An accurate distance to 2M1207Ab". In: *A&A* 477(1). doi: [10.1051/0004-6361:20078886](https://doi.org/10.1051/0004-6361:20078886).
- Dumusque, X. et al. (2011). "The HARPS search for southern extra-solar planets. XXX. Planetary systems around stars with solar-like magnetic cycles and short-term activity variation". In: *A&A* 535, A55. doi: [10.1051/0004-6361/201117148](https://doi.org/10.1051/0004-6361/201117148).
- Dumusque, X. et al. (2021). "Three years of HARPS-N high-resolution spectroscopy and precise radial velocity data for the Sun". In: *A&A* 648, A103. doi: [10.1051/0004-6361/202039350](https://doi.org/10.1051/0004-6361/202039350).
- Dumusque, X. et al. (2012). "An Earth-mass planet orbiting α Centauri B". In: *Nature* 491(7423). doi: [10.1038/nature11572](https://doi.org/10.1038/nature11572).
- Dumusque, X. et al. (2015). "HARPS-N OBSERVES THE SUN AS A STAR". In: *The Astrophysical Journal* 814(2). doi: [10.1088/2041-8205/814/2/121](https://doi.org/10.1088/2041-8205/814/2/121).

- Dumusque, X. et al. (2011a). “Planetary detection limits taking into account stellar noise - I. Observational strategies to reduce stellar oscillation and granulation effects”. In: *A&A* 525. doi: [10.1051/0004-6361/201014097](https://doi.org/10.1051/0004-6361/201014097).
- Dumusque, X. et al. (2011b). “Planetary detection limits taking into account stellar noise - II. Effect of stellar spot groups on radial-velocities”. In: *A&A* 527. doi: [10.1051/0004-6361/201015877](https://doi.org/10.1051/0004-6361/201015877).
- Duvenaud, D. (2014). “Automatic model construction with Gaussian processes”. PhD thesis. University of Cambridge.
- Duvenaud, D. et al. (2013). “Structure Discovery in Nonparametric Regression through Compositional Kernel Search”. In: ed. by S. Dasgupta and D. McAllester. Vol. 28. Proceedings of Machine Learning Research 3. PMLR: Atlanta, Georgia, USA.
- Edelson, R. A. and J. H. Krolik (1988). “The Discrete Correlation Function: A New Method for Analyzing Unevenly Sampled Variability Data”. In: *ApJ* 333. doi: [10.1086/166773](https://doi.org/10.1086/166773).
- Faria, J. P. et al. (2016). “Uncovering the planets and stellar activity of CoRoT-7 using only radial velocities”. In: *A&A* 588, A31. doi: [10.1051/0004-6361/201527899](https://doi.org/10.1051/0004-6361/201527899).
- Faria, J. P. et al. (2020). “Decoding the radial velocity variations of HD 41248 with ESPRESSO”. In: *A&A* 635, A13. doi: [10.1051/0004-6361/201936389](https://doi.org/10.1051/0004-6361/201936389).
- Feng, F. et al. (2017). “Color Difference Makes a Difference: Four Planet Candidates around τ Ceti”. In: ??*jnlAJ* 154(4), 135. doi: [10.3847/1538-3881/aa83b4](https://doi.org/10.3847/1538-3881/aa83b4).
- Figueira, P. et al. (2013). “Line-profile variations in radial-velocity measurements - Two alternative indicators for planetary searches”. In: *A&A* 557. doi: [10.1051/0004-6361/201220779](https://doi.org/10.1051/0004-6361/201220779).
- Figueira, P. et al. (2015). “Line-profile variations in radial-velocity measurements (Corrigendum) - Two alternative indicators for planetary searches”. In: *A&A* 582. doi: [10.1051/0004-6361/201220779e](https://doi.org/10.1051/0004-6361/201220779e).
- Figueira, P. (2018). “Deriving High-Precision Radial Velocities”. In: *Asteroseismology and Exoplanets: Listening to the Stars and Searching for New Worlds*. Ed. by T. L. Campante et al. Vol. 49. doi: [10.1007/978-3-319-59315-9_10](https://doi.org/10.1007/978-3-319-59315-9_10).

- Fischer, D. A. and J. Valenti (2005). “The Planet-Metallicity Correlation”. In: *ApJ* 622(2). doi: [10.1086/428383](https://doi.org/10.1086/428383).
- Foreman-Mackey, D. et al. (2019). “emcee v3: A Python ensemble sampling toolkit for affine-invariant MCMC”. In: *The Journal of Open Source Software* 4(43), 1864. doi: [10.21105/joss.01864](https://doi.org/10.21105/joss.01864).
- Fressin, F. et al. (2013). “The False Positive Rate of Kepler and the Occurrence of Planets”. In: *ApJ* 766(2), 81. doi: [10.1088/0004-637X/766/2/81](https://doi.org/10.1088/0004-637X/766/2/81).
- Fulton, B. J. et al. (2017). “The California-Kepler Survey. III. A Gap in the Radius Distribution of Small Planets”. In: *??jnlAJ* 154(3), 109. doi: [10.3847/1538-3881/aa80eb](https://doi.org/10.3847/1538-3881/aa80eb).
- Gaudi, B. S. (2010). “Exoplanetary Microlensing”. In: *arXiv e-prints*, arXiv:1002.0332.
- Gaudi, B. S. (2012). “Microlensing Surveys for Exoplanets”. In: 50. doi: [10.1146/annurev-astro-081811-125518](https://doi.org/10.1146/annurev-astro-081811-125518).
- Gilbertson, C. et al. (2020). “Towards Extremely Precise Radial Velocities: II. A Tool For Using Multivariate Gaussian Processes to Model Stellar Activity”. In: *arXiv e-prints*, arXiv:2009.01085.
- Gillon, M. et al. (2007). “Detection of transits of the nearby hot Neptune GJ 436 b”. In: *A&A* 472(2). doi: [10.1051/0004-6361:20077799](https://doi.org/10.1051/0004-6361:20077799).
- Ginzburg, S. et al. (2018). “Core-powered mass-loss and the radius distribution of small exoplanets”. In: *MNRAS* 476(1). doi: [10.1093/mnras/sty290](https://doi.org/10.1093/mnras/sty290).
- Gomes, R. et al. (2005). “Origin of the cataclysmic Late Heavy Bombardment period of the terrestrial planets”. In: *Nature* 435(7041). doi: [10.1038/nature03676](https://doi.org/10.1038/nature03676).
- Goodman, J. and J. Weare (2010). “Ensemble samplers with affine invariance”. In: *Communications in Applied Mathematics and Computational Science* 5(1). doi: [10.2140/camcos.2010.5.65](https://doi.org/10.2140/camcos.2010.5.65).
- Gray, D. F. (1997). “Absence of a planetary signature in the spectra of the star 51 Pegasi”. In: *Nature* 385(6619).
- Gray, D. F. and A. P. Hatzes (1997). “Non-Radial Oscillation in the Solar-Temperature Star 51 Pegasi”. In: *The Astrophysical Journal* 490(1). doi: [10.1086/304857](https://doi.org/10.1086/304857).
- Grigoriu, M. (2002). *Stochastic Calculus: Applications in Science and Engineering*. Stochastic Calculus: Applications in Science and Engineering. Birkhäuser Boston. ISBN: 9780817642426.

- Grinstead, C. and J. Snell (2012). *Introduction to Probability*. American Mathematical Society. ISBN: 9780821894149.
- Gut, A. (2009). “The Multivariate Normal Distribution”. In: *An Intermediate Course in Probability*. Springer New York: New York, NY. ISBN: 978-1-4419-0162-0. doi: [10.1007/978-1-4419-0162-0_5](https://doi.org/10.1007/978-1-4419-0162-0_5).
- Hall, J. C. (2008). “Stellar Chromospheric Activity”. In: *Living Reviews in Solar Physics* 5(1). doi: [10.12942/lrsp-2008-2](https://doi.org/10.12942/lrsp-2008-2).
- Hathaway, D. H. (2015). “The Solar Cycle”. In: *Living Reviews in Solar Physics* 12(1), 4. doi: [10.1007/lrsp-2015-4](https://doi.org/10.1007/lrsp-2015-4).
- Hatzes, A. P. et al. (2003). “A Planetary Companion to γ Cephei A”. In: *ApJ* 599(2). doi: [10.1086/379281](https://doi.org/10.1086/379281).
- Haywood, R. D. et al. (2014). “Planets and stellar activity: hide and seek in the CoRoT-7 system”. In: *Monthly Notices of the Royal Astronomical Society* 443(3). doi: [10.1093/mnras/stu1320](https://doi.org/10.1093/mnras/stu1320).
- Haywood, R. D. (2016a). “Stellar Activity as a Source of Radial-Velocity Variability”. In: *Radial-velocity Searches for Planets Around Active Stars*. Springer International Publishing: Cham. ISBN: 978-3-319-41273-3. doi: [10.1007/978-3-319-41273-3_2](https://doi.org/10.1007/978-3-319-41273-3_2).
- Haywood, R. (2016b). *Radial-velocity Searches for Planets Around Active Stars*. Springer Theses. Springer International Publishing. ISBN: 9783319412733.
- Heinonen, M. et al. (2015). “Non-Stationary Gaussian Process Regression with Hamiltonian Monte Carlo”. In: *arXiv e-prints*, arXiv:1508.04319.
- Henry, G. W. et al. (2000). “A Transiting “51 Peg-like” Planet”. In: *ApJ* 529(1). doi: [10.1086/312458](https://doi.org/10.1086/312458).
- Higson, E. et al. (2019). “Dynamic nested sampling: an improved algorithm for parameter estimation and evidence calculation”. In: *Statistics and Computing* 29(5). doi: [10.1007/s11222-018-9844-0](https://doi.org/10.1007/s11222-018-9844-0).
- Hirano, T. et al. (2018). “Exoplanets around Low-mass Stars Unveiled by K2”. In: ??*jnlAJ* 155(3), 127. doi: [10.3847/1538-3881/aaa9c1](https://doi.org/10.3847/1538-3881/aaa9c1).
- Holman, M. J. and N. W. Murray (2005). “The Use of Transit Timing to Detect Terrestrial-Mass Extrasolar Planets”. In: *Science* 307(5713). doi: [10.1126/science.1107822](https://doi.org/10.1126/science.1107822).

- Isaacson, H. and D. Fischer (2010). “Chromospheric Activity and Jitter Measurements for 2630 Stars on the California Planet Search”. In: *ApJ* 725(1). doi: [10.1088/0004-637X/725/1/875](https://doi.org/10.1088/0004-637X/725/1/875).
- Jeffreys, H. (1983). *Theory of Probability*. International series of monographs on physics. Clarendon Press. ISBN: 9780198531937.
- Johnson, S. A. et al. (2020). “Predictions of the Nancy Grace Roman Space Telescope Galactic Exoplanet Survey. II. Free-floating Planet Detection Rates”. In: *??jnlAJ* 160(3), 123. doi: [10.3847/1538-3881/aba75b](https://doi.org/10.3847/1538-3881/aba75b).
- Jones, D. E. et al. (2017). “Improving Exoplanet Detection Power: Multivariate Gaussian Process Models for Stellar Activity”. In: *arXiv e-prints*, arXiv:1711.01318.
- Jordan, M. I. et al. (1999). “An introduction to variational methods for graphical models”. In: *Machine learning* 37(2).
- Jurgenson, C. et al. (2016). “EXPRES: a next generation RV spectrograph in the search for earth-like worlds”. In: *Ground-based and Airborne Instrumentation for Astronomy VI*. Vol. 9908. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. doi: [10.1117/12.2233002](https://doi.org/10.1117/12.2233002).
- Karamanis, M. et al. (2021). “zeus: A Python implementation of Ensemble Slice Sampling for efficient Bayesian parameter inference”. In: *arXiv e-prints*, arXiv:2105.03468.
- Kass, R. E. and A. E. Raftery (1995). “Bayes Factors”. In: *Journal of the American Statistical Association* 90(430). doi: [10.1080/01621459.1995.10476572](https://doi.org/10.1080/01621459.1995.10476572).
- Keenan, P. C. and R. C. McNeil (1989). “The Perkins Catalog of Revised MK Types for the Cooler Stars”. In: 71. doi: [10.1086/191373](https://doi.org/10.1086/191373).
- Keller, C. U. et al. (2004). “On the Origin of Solar Faculae”. In: *ApJ* 607(1). doi: [10.1086/421553](https://doi.org/10.1086/421553).
- Ketu, S. and P. K. Mishra (2020). “Enhanced Gaussian process regression-based forecasting model for COVID-19 outbreak and significance of IoT for its detection”. In: *Applied Intelligence*.
- Kosiarek, M. R. and I. J. M. Crossfield (2020). “Photometry as a Proxy for Stellar Activity in Radial Velocity Analyses”. In: *The Astronomical Journal* 159(6). doi: [10.3847/1538-3881/ab8d3a](https://doi.org/10.3847/1538-3881/ab8d3a).
- Kullback, S. (1959). *Information Theory and Statistics*. Wiley: New York.
- Kung, S. H. (1996). “Proof Without Words: The Difference-Product Identities”. In: *Mathematics Magazine* 69(4). doi: [10.1080/0025570X.1996.11996453](https://doi.org/10.1080/0025570X.1996.11996453).

- Lagrange, A.-M. et al. (2010). “Using the Sun to estimate Earth-like planets detection capabilities - I. Impact of cold spots”. In: *A&A* 512. doi: [10.1051/0004-6361/200913071](https://doi.org/10.1051/0004-6361/200913071).
- Lanza, A. F. et al. (2019). “Sectoral r modes and periodic radial velocity variations of Sun-like stars”. In: *A&A* 623, A50. doi: [10.1051/0004-6361/201834712](https://doi.org/10.1051/0004-6361/201834712).
- Léger, A. et al. (2009). “Transiting exoplanets from the CoRoT space mission. VIII. CoRoT-7b: the first super-Earth with measured radius”. In: *A&A* 506(1). doi: [10.1051/0004-6361/200911933](https://doi.org/10.1051/0004-6361/200911933).
- Li, S. et al. (2020). “Scalable Gaussian Process Regression Networks”. In: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*. Ed. by C. Bessiere. Main track. International Joint Conferences on Artificial Intelligence Organization.
- Lin, D. N. C. and J. Papaloizou (1986). “On the Tidal Interaction between Protoplanets and the Protoplanetary Disk. III. Orbital Migration of Protoplanets”. In: *ApJ* 309. doi: [10.1086/164653](https://doi.org/10.1086/164653).
- Lin, D. N. C. et al. (1996). “Orbital migration of the planetary companion of 51 Pegasi to its present location”. In: *Nature* 380(6575). doi: [10.1038/380606a0](https://doi.org/10.1038/380606a0).
- Lopez, E. D. and J. J. Fortney (2013). “The Role of Core Mass in Controlling Evaporation: The Kepler Radius Distribution and the Kepler-36 Density Dichotomy”. In: *ApJ* 776(1), 2. doi: [10.1088/0004-637X/776/1/2](https://doi.org/10.1088/0004-637X/776/1/2).
- Löptien, B. et al. (2018). “Global-scale equatorial Rossby waves as an essential component of solar internal dynamics”. In: *Nature Astronomy* 2. doi: [10.1038/s41550-018-0460-x](https://doi.org/10.1038/s41550-018-0460-x).
- Lovis, C. et al. (2011). “The HARPS search for southern extra-solar planets. XXXI. Magnetic activity cycles in solar-type stars: statistics and impact on precise radial velocities”. In: *arXiv e-prints*, arXiv:1107.5325.
- Lubin, D. et al. (2012). “Frequency of Maunder Minimum Events in Solar-type Stars Inferred from Activity and Metallicity Observations”. In: *ApJ* 747(2), L32. doi: [10.1088/2041-8205/747/2/L32](https://doi.org/10.1088/2041-8205/747/2/L32).
- Ma, B. et al. (2018). “The first super-Earth detection from the high cadence and high radial velocity precision Dharma Planet Survey”. In: *MNRAS* 480(2). doi: [10.1093/mnras/sty1933](https://doi.org/10.1093/mnras/sty1933).

- Mackay, D. (1998). "Introduction to Gaussian Processes". In: *Neural Networks and Machine Learning*. Ed. by C. M. Bishop. Springer-Verlag. ISBN: 978-3540649281.
- Malbet, F. and A. Sozzetti (2018). "Astrometry as an Exoplanet Discovery Method". In: *Handbook of Exoplanets*. Ed. by H. J. Deeg and J. A. Belmonte. doi: [10.1007/978-3-319-55333-7\196](https://doi.org/10.1007/978-3-319-55333-7_196).
- Maldonado, J. et al. (2010). "A spectroscopy study of nearby late-type stars, possible members of stellar kinematic groups". In: *A&A* 521, A12. doi: [10.1051/0004-6361/201014948](https://doi.org/10.1051/0004-6361/201014948).
- Marcy, G. et al. (2005). "Observed Properties of Exoplanets: Masses, Orbits, and Metallicities". In: *Progress of Theoretical Physics Supplement* 158. doi: [10.1143/PTPS.158.24](https://doi.org/10.1143/PTPS.158.24).
- Marcy, G. W. and R. P. Butler (1995). "The Planet around 51 Pegasi". In: *American Astronomical Society Meeting Abstracts*. Vol. 187. American Astronomical Society Meeting Abstracts.
- Marcy, G. W. and R. P. Butler (1996). "A Planetary Companion to 70 Virginis". In: *ApJ* 464. doi: [10.1086/310096](https://doi.org/10.1086/310096).
- Marois, C. et al. (2006). "Angular Differential Imaging: A Powerful High-Contrast Imaging Technique". In: *ApJ* 641(1). doi: [10.1086/500401](https://doi.org/10.1086/500401).
- Martínez-Arnáiz, R. et al. (2010). "Chromospheric activity and rotation of FGK stars in the solar vicinity *** - An estimation of the radial velocity jitter". In: *A&A* 520. doi: [10.1051/0004-6361/200913725](https://doi.org/10.1051/0004-6361/200913725).
- Mayor, M. (1980). "Metal abundances of F and G dwarfs determined by the radial velocity scanner CORAVEL". In: *A&A* 87(1-2).
- Mayor, M. et al. (2011). "The HARPS search for southern extra-solar planets XXXIV. Occurrence, mass distribution and orbital properties of super-Earths and Neptune-mass planets". In: *arXiv e-prints*, arXiv:1109.2497.
- Mayor, M. and D. Queloz (1995). "A Jupiter-mass companion to a solar-type star". In: *Nature* 378(6555). doi: [10.1038/378355a0](https://doi.org/10.1038/378355a0).
- Mazeh, T. et al. (2016). "Dearth of short-period Neptunian exoplanets: A desert in period-mass and period-radius planes". In: *A&A* 589, A75. doi: [10.1051/0004-6361/201528065](https://doi.org/10.1051/0004-6361/201528065).

- McArthur, B. E. et al. (2004). “Detection of a Neptune-Mass Planet in the ρ^1 Cancri System Using the Hobby-Eberly Telescope”. In: *ApJ* 614(1). doi: [10.1086/425561](https://doi.org/10.1086/425561).
- McIntosh, S. W. et al. (2014). “Deciphering Solar Magnetic Activity. I. On the Relationship between the Sunspot Cycle and the Evolution of Small Magnetic Features”. In: *ApJ* 792(1), 12. doi: [10.1088/0004-637X/792/1/12](https://doi.org/10.1088/0004-637X/792/1/12).
- Melia, F. and M. K. Yennapureddy (2018). “Model selection using cosmic chronometers with Gaussian Processes”. In: 2018(2), 034. doi: [10.1088/1475-7516/2018/02/034](https://doi.org/10.1088/1475-7516/2018/02/034).
- Meunier, N. and A. M. Lagrange (2020). “The effects of granulation and supergranulation on Earth-mass planet detectability in the habitable zone around F6-K4 stars”. In: *A&A* 642, A157. doi: [10.1051/0004-6361/202038376](https://doi.org/10.1051/0004-6361/202038376).
- Meunier, N. and Lagrange, A.-M. (2019). “Unexpectedly strong effect of supergranulation on the detectability of Earth twins orbiting Sun-like stars with radial velocities”. In: *A&A* 625. doi: [10.1051/0004-6361/201935099](https://doi.org/10.1051/0004-6361/201935099).
- Meunier, N. et al. (2010). “Using the Sun to estimate Earth-like planets detection capabilities - II. Impact of plages”. In: *A&A* 512. doi: [10.1051/0004-6361/200913551](https://doi.org/10.1051/0004-6361/200913551).
- Morbidelli, A. et al. (2005). “Chaotic capture of Jupiter’s Trojan asteroids in the early Solar System”. In: *Nature* 435(7041). doi: [10.1038/nature03540](https://doi.org/10.1038/nature03540).
- Murray, C. D. and A. C. M. Correia (2010). “Keplerian Orbits and Dynamics of Exoplanets”. In: *Exoplanets*. Ed. by S. Seager.
- Muterspaugh, M. W. et al. (2010). “The Phases Differential Astrometry Data Archive. V. Candidate Substellar Companions to Binary Systems”. In: ??jnlAJ 140(6). doi: [10.1088/0004-6256/140/6/1657](https://doi.org/10.1088/0004-6256/140/6/1657).
- Nandy, D. et al. (2020). “Sunspot Cycle 25 is Brewing: Early Signs Herald its Onset”. In: *Research Notes of the American Astronomical Society* 4(2), 30. doi: [10.3847/2515-5172/ab79a1](https://doi.org/10.3847/2515-5172/ab79a1).
- Neal, R. M. (2000). “Slice Sampling”. In: *arXiv e-prints*, physics/0009028.
- Newton, I. and E. Halley (1687). *Philosophiae naturalis principia mathematica*. Jussu Societatis Regiae ac typis Josephi Streater, prostant venales apud Sam. Smith.
- Nguyen, T. and E. Bonilla (2013). “Efficient Variational Inference for Gaussian Process Regression Networks”. In: *Proceedings of the Sixteenth International*

- Conference on Artificial Intelligence and Statistics*. Ed. by C. M. Carvalho and P. Ravikumar. Vol. 31. Proceedings of Machine Learning Research. PMLR: Scottsdale, Arizona, USA.
- Nguyen, T. V. (2015). “Scalable and automated inference for gaussian process models”. PhD thesis. Australian National University. Research School of Computer Science.
- Nirwan, R. S. and N. Bertschinger (2020). “Applications of Gaussian Process Latent Variable Models in Finance”. In: *Intelligent Systems and Applications*. Ed. by Y. Bi et al. Springer International Publishing: Cham. ISBN: 978-3-030-29513-4.
- NIST (2019). *The NIST Reference on Constants, Units, and Uncertainty*. National Institute of Standards and Technology, USA.
- Noyes, R. W. et al. (1984). “Rotation, convection, and magnetic activity in lower main-sequence stars”. In: *ApJ* 279. doi: [10.1086/161945](https://doi.org/10.1086/161945).
- Osborne, M. A. (2010). “Bayesian Gaussian processes for sequential prediction, optimisation and quadrature”. PhD thesis. Oxford University, UK.
- Oshagh, M. (2018). “Noise Sources in Photometry and Radial Velocities”. In: *Asteroseismology and Exoplanets: Listening to the Stars and Searching for New Worlds*. Ed. by T. L. Campante et al. Springer International Publishing: Cham. ISBN: 978-3-319-59315-9.
- Owen, J. E. and D. Lai (2018). “Photoevaporation and high-eccentricity migration created the sub-Jovian desert”. In: *MNRAS* 479(4). doi: [10.1093/mnras/sty1760](https://doi.org/10.1093/mnras/sty1760).
- Owen, J. E. and Y. Wu (2013). “Kepler Planets: A Tale of Evaporation”. In: *ApJ* 775(2), 105. doi: [10.1088/0004-637X/775/2/105](https://doi.org/10.1088/0004-637X/775/2/105).
- Pagano, I. (2013). “Stellar Activity”. In: *Planets, Stars and Stellar Systems: Volume 4: Stellar Structure and Evolution*. Ed. by T. D. Oswalt and M. A. Barstow. Springer Netherlands: Dordrecht. ISBN: 978-94-007-5615-1. doi: [10.1007/978-94-007-5615-1_10](https://doi.org/10.1007/978-94-007-5615-1_10).
- Pasquini, L. et al. (2010). “CODEX: An Ultra-stable High Resolution Spectrograph for the E-ELT”. In: *The Messenger* 140.
- Patel, J. and C. Read (1996). *Handbook of the Normal Distribution, Second Edition*. 2nd. Statistics: A Series of Textbooks and Monographs. Taylor & Francis. ISBN: 9780824793425.

- Pepe, F et al. (2002). “HARPS: ESO’s coming planet searcher. Chasing exoplanets with the La Silla 3.6-m telescope”. In: *The Messenger* 110.
- Pepe, F. et al. (2002). “The CORALIE survey for southern extra-solar planets VII. Two short-period Saturnian companions to <ASTROBJ>HD 108147</ASTROBJ> and <ASTROBJ>HD 168746</ASTROBJ>”. In: *A&A* 388. doi: [10.1051/0004-6361:20020433](https://doi.org/10.1051/0004-6361:20020433).
- Pepe, F. et al. (2014). “ESPRESSO: The next European exoplanet hunter”. In: *Astronomische Nachrichten* 335(1). doi: [10.1002/asna.201312004](https://doi.org/10.1002/asna.201312004).
- Pepe, F. A. and C. Lovis (2008). “From HARPS to CODEX: exploring the limits of Doppler measurements”. In: *Physica Scripta Volume T* 130, 014007. doi: [10.1088/0031-8949/2008/T130/014007](https://doi.org/10.1088/0031-8949/2008/T130/014007).
- Pepe, F. et al. (2014). “Instrumentation for the detection and characterization of exoplanets”. In: *Nature* 513(7518). doi: [10.1038/nature13784](https://doi.org/10.1038/nature13784).
- Pepe, F. et al. (2018). “High-Precision Spectrographs for Exoplanet Research: CORAVEL, ELODIE, CORALIE, SOPHIE and HARPS”. In: *Handbook of Exoplanets*. Ed. by H. J. Deeg and J. A. Belmonte. Springer International Publishing: Cham. ISBN: 978-3-319-55333-7. doi: [10.1007/978-3-319-55333-7_190](https://doi.org/10.1007/978-3-319-55333-7_190).
- Perrakis, K. et al. (2014). “On the use of marginal posteriors in marginal likelihood estimation via importance sampling”. In: *Computational Statistics Data Analysis* 77. doi: <https://doi.org/10.1016/j.csda.2014.03.004>.
- Perryman, M. (2011). *The Exoplanet Handbook*. Cambridge University Press. doi: [10.1017/CBO9780511994852](https://doi.org/10.1017/CBO9780511994852).
- Petigura, E. A. et al. (2017). “The California-Kepler Survey. I. High-resolution Spectroscopy of 1305 Stars Hosting Kepler Transiting Planets”. In: ??jnlAJ 154(3), 107. doi: [10.3847/1538-3881/aa80de](https://doi.org/10.3847/1538-3881/aa80de).
- Pizzolato, N. et al. (2003). “The stellar activity-rotation relationship revisited: Dependence of saturated and non-saturated X-ray emission regimes on stellar mass for late-type dwarfs”. In: *A&A* 397. doi: [10.1051/0004-6361:20021560](https://doi.org/10.1051/0004-6361:20021560).
- Plagemann, C. et al. (2008). “Nonstationary Gaussian Process Regression Using Point Estimates of Local Smoothness”. In: *Machine Learning and Knowledge Discovery in Databases*. Ed. by W. Daelemans et al. Springer Berlin Heidelberg: Berlin, Heidelberg. ISBN: 978-3-540-87481-2.

- Pollack, J. B. et al. (1996). “Formation of the Giant Planets by Concurrent Accretion of Solids and Gas”. In: 124(1). doi: [10.1006/icar.1996.0190](https://doi.org/10.1006/icar.1996.0190).
- Queloz, D. et al. (2001). “No planet for HD 166435”. In: *A&A* 379(1). doi: [10.1051/0004-6361:20011308](https://doi.org/10.1051/0004-6361:20011308).
- Queloz, D. (1995). “Echelle Spectroscopy by a CCD at Low Signal-to-Noise Ratio”. In: *Symposium - International Astronomical Union* 167. doi: [10.1017/S0074180900056473](https://doi.org/10.1017/S0074180900056473).
- Rajpaul, V. et al. (2015). “A Gaussian process framework for modelling stellar activity signals in radial velocity data”. In: *Monthly Notices of the Royal Astronomical Society* 452(3). doi: [10.1093/mnras/stv1428](https://doi.org/10.1093/mnras/stv1428).
- Rajpaul, V. M. (2017). “Gaussian process tools for modelling stellar signals and studying exoplanets”. PhD thesis. University of Oxford.
- Rasmussen, C. E. and C. K. I. Williams (2006). *Gaussian Processes for Machine Learning*. eng. 1st. MIT Press. ISBN: 0-262-18253-X.
- Reiners, A. (2009). “Activity-induced radial velocity jitter in a flaring M dwarf”. In: *A&A* 498(3). doi: [10.1051/0004-6361/200810257](https://doi.org/10.1051/0004-6361/200810257).
- Reiners, A. et al. (2012). “A Catalog of Rotation and Activity in Early-M Stars”. In: ??*jnlAJ* 143(4), 93. doi: [10.1088/0004-6256/143/4/93](https://doi.org/10.1088/0004-6256/143/4/93).
- Reinhold, T. et al. (2019). “Transition from spot to faculae domination. An alternate explanation for the dearth of intermediate Kepler rotation periods”. In: *A&A* 621, A21. doi: [10.1051/0004-6361/201833754](https://doi.org/10.1051/0004-6361/201833754).
- Rieutord, M. and F. Rincon (2010). “The Sun’s Supergranulation”. In: *Living Reviews in Solar Physics* 7(1), 2. doi: [10.12942/lrsp-2010-2](https://doi.org/10.12942/lrsp-2010-2).
- Rivera, E. J. et al. (2005). “A ~ 7.5 M_{\oplus} Planet Orbiting the Nearby Star, GJ 876”. In: *ApJ* 634(1). doi: [10.1086/491669](https://doi.org/10.1086/491669).
- Robertson, P. et al. (2015). “Stellar Activity Mimics a Habitable-zone Planet around Kapteyn’s Star”. In: *ApJ* 805(2), L22. doi: [10.1088/2041-8205/805/2/L22](https://doi.org/10.1088/2041-8205/805/2/L22).
- Rowe, J. F. et al. (2014). “Validation of Kepler’s Multiple Planet Candidates. III. Light Curve Analysis and Announcement of Hundreds of New Multi-planet Systems”. In: *ApJ* 784(1), 45. doi: [10.1088/0004-637X/784/1/45](https://doi.org/10.1088/0004-637X/784/1/45).
- Ruane, G. et al. (2019). “Reference Star Differential Imaging of Close-in Companions and Circumstellar Disks with the NIRC2 Vortex Coronagraph at the W. M. Keck Observatory”. In: ??*jnlAJ* 157(3), 118. doi: [10.3847/1538-3881/aafee2](https://doi.org/10.3847/1538-3881/aafee2).

- Rumelhart, D. E. et al. (1986). “Learning representations by back-propagating errors”. In: *Nature* 323(6088).
- Rutten, R. and G. Severino (2012). *Solar and Stellar Granulation*. Nato Science Series C: Springer Netherlands. ISBN: 9789400909113.
- Saar, S. H. (2009). “The Radial Velocity Effects of Stellar Surface Phenomena”. In: *15th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun*. Ed. by E. Stempels. Vol. 1094. American Institute of Physics Conference Series. doi: [10.1063/1.3099086](https://doi.org/10.1063/1.3099086).
- Saar, S. H. and R. A. Donahue (1997). “Activity-Related Radial Velocity Variation in Cool Stars”. In: *The Astrophysical Journal* 485(1). doi: [10.1086/304392](https://doi.org/10.1086/304392).
- Sahli Costabal, F. et al. (2019). “Machine learning in drug development: Characterizing the effect of 30 drugs on the QT interval using Gaussian process regression, sensitivity analysis, and uncertainty quantification”. In: *Computer Methods in Applied Mechanics and Engineering* 348. doi: [10.1016/j.cma.2019.01.033](https://doi.org/10.1016/j.cma.2019.01.033).
- Santerne, A. et al. (2015). “pastis: Bayesian extrasolar planet validation – II. Constraining exoplanet blend scenarios using spectroscopic diagnoses”. In: *Monthly Notices of the Royal Astronomical Society* 451(3). doi: [10.1093/mnras/stv1080](https://doi.org/10.1093/mnras/stv1080).
- Santerne, A. et al. (2016). “SOPHIE velocimetry of Kepler transit candidates. XVII. The physical properties of giant exoplanets within 400 days of period”. In: *A&A* 587, A64. doi: [10.1051/0004-6361/201527329](https://doi.org/10.1051/0004-6361/201527329).
- Santos, N. C. et al. (2002). “The CORALIE survey for southern extra-solar planets. IX. A 1.3-day period brown dwarf disguised as a planet”. In: *A&A* 392. doi: [10.1051/0004-6361:20020876](https://doi.org/10.1051/0004-6361:20020876).
- Santos, N. C. et al. (2004a). “Spectroscopic [Fe/H] for 98 extra-solar planet-host stars. Exploring the probability of planet formation”. In: *A&A* 415. doi: [10.1051/0004-6361:20034469](https://doi.org/10.1051/0004-6361:20034469).
- Santos, N. C. et al. (2004b). “The HARPS survey for southern extra-solar planets. II. A 14 Earth-masses exoplanet around μ Arae”. In: *A&A* 426. doi: [10.1051/0004-6361:200400076](https://doi.org/10.1051/0004-6361:200400076).
- Santos, N. C. et al. (2014). “The HARPS search for southern extra-solar planets - XXXV. The interesting case of HD41248: stellar activity, no planets?” In: *A&A* 566. doi: [10.1051/0004-6361/201423808](https://doi.org/10.1051/0004-6361/201423808).

- Santos, N. C. and J. P. Faria (2018). “Exoplanetary Science: An Overview”. In: *Asteroseismology and Exoplanets: Listening to the Stars and Searching for New Worlds*. Ed. by T. L. Campante et al. Springer International Publishing: Cham. ISBN: 978-3-319-59315-9.
- Sauer, T. (2008). “Nova Geminorum 1912 and the origin of the idea of gravitational lensing”. In: *Archive for History of Exact Sciences* 62(1).
- Schrijver, C. J. and C. Zwaan (2000). *Solar and Stellar Magnetic Activity*. Cambridge Astrophysics. Cambridge University Press. doi: [10.1017/CBO9780511546037](https://doi.org/10.1017/CBO9780511546037).
- Seager, S. and G. Mallén-Ornelas (2003). “A Unique Solution of Planet and Star Parameters from an Extrasolar Planet Transit Light Curve”. In: *ApJ* 585(2). doi: [10.1086/346105](https://doi.org/10.1086/346105).
- Serrano, L. M. et al. (2018). “Distinguishing the albedo of exoplanets from stellar activity”. In: *A&A* 611, A8. doi: [10.1051/0004-6361/201731206](https://doi.org/10.1051/0004-6361/201731206).
- Sher, D. (1968). “The Relativistic Doppler Effect”. In: 62.
- Sivia, D. and J. Skilling (2006). *Data Analysis: A Bayesian Tutorial*. Oxford science publications. OUP Oxford. ISBN: 9780191546709.
- Skilling, J. (2006). “Nested sampling for general Bayesian computation”. In: *Bayesian Anal.* 1(4). doi: [10.1214/06-BA127](https://doi.org/10.1214/06-BA127).
- Sokal, A. (1997). “Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms”. In: *Functional Integration: Basics and Applications*. Ed. by C. DeWitt-Morette et al. Springer US: Boston, MA. ISBN: 978-1-4899-0319-8. doi: [10.1007/978-1-4899-0319-8_6](https://doi.org/10.1007/978-1-4899-0319-8_6).
- Solanki, S. K. (1993). “Smallscale Solar Magnetic Fields - an Overview”. In: 63(1-2). doi: [10.1007/BF00749277](https://doi.org/10.1007/BF00749277).
- Solanki, S. K. (2003). “Sunspots: An overview”. In: *A&A Rev.* 11(2-3). doi: [10.1007/s00159-003-0018-4](https://doi.org/10.1007/s00159-003-0018-4).
- Sousa, S. G. et al. (2008). “Spectroscopic parameters for 451 stars in the HARPS GTO planet search program. Stellar [Fe/H] and the frequency of exo-Neptunes”. In: *A&A* 487(1). doi: [10.1051/0004-6361:200809698](https://doi.org/10.1051/0004-6361:200809698).
- Sousa, S. G. et al. (2011). “Spectroscopic stellar parameters for 582 FGK stars in the HARPS volume-limited sample. Revising the metallicity-planet correlation”. In: *A&A* 533, A141. doi: [10.1051/0004-6361/201117699](https://doi.org/10.1051/0004-6361/201117699).

- Speagle, J. S. (2020). “DYNESTY: a dynamic nested sampling package for estimating Bayesian posteriors and evidences”. In: *MNRAS* 493(3). doi: [10.1093/mnras/staa278](https://doi.org/10.1093/mnras/staa278).
- Spruit, H. C. (1976). “Pressure equilibrium and energy balance of small photospheric fluxtubes.” In: *Sol. Phys.* 50(2). doi: [10.1007/BF00155292](https://doi.org/10.1007/BF00155292).
- Stein, M. (1999). “Interpolation of Spatial Data: Some Theory for Kriging”. In:
- Styan, G. P. (1973). “Hadamard products and multivariate statistical analysis”. In: *Linear Algebra and its Applications* 6. doi: [https://doi.org/10.1016/0024-3795\(73\)90023-2](https://doi.org/10.1016/0024-3795(73)90023-2).
- Szabó, G. M. and L. L. Kiss (2011). “A Short-period Censor of Sub-Jupiter Mass Exoplanets with Low Density”. In: *ApJ* 727(2), L44. doi: [10.1088/2041-8205/727/2/L44](https://doi.org/10.1088/2041-8205/727/2/L44).
- Teh, Y.-W. et al. (2005). “Semiparametric Latent Factor Models”. In: *Artificial Intelligence and Statistics 10*.
- Tsiganis, K. et al. (2005). “Origin of the orbital architecture of the giant planets of the Solar System”. In: *Nature* 435(7041). doi: [10.1038/nature03539](https://doi.org/10.1038/nature03539).
- Udalski, A. et al. (1997). “Optical Gravitational Lensing Experiment. OGLE-2 – the Second Phase of the OGLE Project”. In: 47.
- Udry, S. and N. C. Santos (2007). “Statistical Properties of Exoplanets”. In: *Annual Review of Astronomy and Astrophysics* 45(1). doi: [10.1146/annurev.astro.45.051806.110529](https://doi.org/10.1146/annurev.astro.45.051806.110529).
- van de Kamp, P. (1982). “The planetary system of Barnard’s star”. In: *Vistas in Astronomy* 26(2). doi: [10.1016/0083-6656\(82\)90004-6](https://doi.org/10.1016/0083-6656(82)90004-6).
- Vaughan, A. H. and G. W. Preston (1980). “A survey of chromospheric CA II H and K emission in field stars of the solar neighborhood.” In: *PASP* 92. doi: [10.1086/130683](https://doi.org/10.1086/130683).
- Vaughan, A. H. et al. (1978). “Flux measurements of CA II H and K emission”. In: *Publications of the Astronomical Society of the Pacific* 90. doi: [10.1086/130324](https://doi.org/10.1086/130324).
- Viryanen, P. et al. (2020). “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. In: *Nature Methods* 17. doi: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2).

- Wendland, H. (2004). *Scattered Data Approximation*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press. ISBN: 9781139456654.
- West, R. G. et al. (2019). “NGTS-4b: A sub-Neptune transiting in the desert”. In: *MNRAS* 486(4). doi: [10.1093/mnras/stz1084](https://doi.org/10.1093/mnras/stz1084).
- Wilcox, J. M. (1972). “Divers Solar Rotations”. In: *Cosmic Plasma Physics*. Ed. by K. Schindler. Springer US: Boston, MA. ISBN: 978-1-4615-6758-5.
- Wilson, A. G. (2014). “Covariance kernels for fast automatic pattern discovery and extrapolation with Gaussian processes”. PhD thesis. University of Cambridge.
- Wilson, A. G. et al. (2012). “Gaussian Process Regression Networks”. In: *Proceedings of the 29th International Conference on International Conference on Machine Learning*. ICML’12. Omnipress: Edinburgh, Scotland. ISBN: 978-1-4503-1285-1.
- Wilson, O. C. (1962). “Relationship Between Colors and Spectra of Late Main-Sequence Stars.” In: *ApJ* 136. doi: [10.1086/147437](https://doi.org/10.1086/147437).
- Wilson, O. C. (1978). “Chromospheric variations in main-sequence stars”. In: *ApJ* 226. doi: [10.1086/156618](https://doi.org/10.1086/156618).
- Winn, J. N. (2018). “Planet Occurrence: Doppler and Transit Surveys”. In: *Handbook of Exoplanets*. Ed. by H. J. Deeg and J. A. Belmonte. doi: [10.1007/978-3-319-55333-7_195](https://doi.org/10.1007/978-3-319-55333-7_195).
- Wolszczan, A. and D. A. Frail (1992). “A planetary system around the millisecond pulsar PSR1257 + 12”. In: *Nature* 355(6356). doi: [10.1038/355145a0](https://doi.org/10.1038/355145a0).
- Wolszczan, A. (1994). “Confirmation of Earth-Mass Planets Orbiting the Millisecond Pulsar PSR B1257+12”. In: *Science* 264(5158). doi: [10.1126/science.264.5158.538](https://doi.org/10.1126/science.264.5158.538).
- Wright, J. T. et al. (2012). “The Frequency of Hot Jupiters Orbiting nearby Solar-type Stars”. In: *ApJ* 753(2), 160. doi: [10.1088/0004-637X/753/2/160](https://doi.org/10.1088/0004-637X/753/2/160).
- Wright, J. T. and B. S. Gaudi (2013). “Exoplanet Detection Methods”. In: *Planets, Stars and Stellar Systems: Volume 3: Solar and Stellar Planetary Systems*. Ed. by T. D. Oswalt et al. Springer Netherlands: Dordrecht. ISBN: 978-94-007-5606-9. doi: [10.1007/978-94-007-5606-9_10](https://doi.org/10.1007/978-94-007-5606-9_10).