# Proposal for the use of Gaussian Processes in creating a model for Gyrochronology

**Author**: Henry Shaw
**Project partner**: Nikolas Hadjiantoni
**Project supervisor**: Dr Guy R. Davies
(Dated: October 23, 2020)

A proper measure of time is the greatest determinant for many systems that define our understanding of the universe. However, many existing approaches rely on specific system arrangements and confer a high degree of uncertainty if these conditions are not met. Gyrochronology poses a convenient and accurate method that uses the deceleration of the rotational motion of a star as a measure of its age. We propose the use of a cutting-edge machine learning method to generate a model that will predict the age of low-mass stars applying this approach. By training our algorithm on existing well understood systems, we aim to be able to make precise predictions of age with a well defined measure of uncertainty, given just the rotational velocity and a proxy for the mass, with the opportunity for other parameters of the star to be considered.

## I. CONTEXT

Gyrochronology is an emerging method in astronomy wherein the age of sun-like and late sequence stars can be determined from their rotational period and a few other easily measurable quantities. However it has not seen much application since Skumanich determined the relation between rotational period and loss of angular momenta in 1972 [1]. In recent years it has resurfaced and looks to be a promising method in determining the age of stars. The rotational period of a star is a very measurable quantity that can often be measured to a high degree of precision and has a well understood time dependence. As we observe a star spinning upon its axis, we observe periodic dips in its apparent brightness. This is a result of sunspots, less bright areas where the surface is cooler, reducing the total amount of observed light leaving the star. NASA's Kepler mission has provided the scientific community with photo-metric data on a large range of stars within its view, wherein the period of many stars has been determined through these periodic dips.

Barnes expanded on Skumanich's work, where he found that the rotational velocity of a sun-like star is proportional to the inverse square root of its age $v \propto t^{1/2}$. He proposed that the spin down (reduction in rotational velocity over time) of stars, due to magnetic breaking, coupled with its colour, or some other proxy for mass, can be used to accurately determine the age of stars stating

> "Gyrochronology transforms a rotating star into a clock which is set using the Sun and keeps time well" [2].

Barnes has since expanded on his initial work and has used this method to determine the age of stars with an error of only $\approx 15\%$ for late F, G, K, and early M stars and stated that this method works for individual stars as well as those in clusters [3]. Contrarily, existing methods for determining age require stars to be within a cluster. Therefore, more study into this area may hopefully lead to providing the astronomical community with a precise and easy to use method for determining stellar ages. This will have a myriad of practical uses, as well as possibly providing better theoretical understanding of magnetic activity and angular momentum transport in stars.

The current best understanding for why stars spin-down is magnetic breaking, for which a full explanation can be found in Saders' *et al.* 2016 paper [4]. This paper, along with work completed by Meibom *et al.* [5] bridged a broad gap in our understanding of gyrochronology. Prior to it, rotational data had only been analysed for our sun, aged at $4.6Gyrs$, and two much younger systems, NGC 6811 and The Hyades with ages of $1Gyr$ and $0.6Gyrs$ respectively. Meibom *et al.* analysed 30 stars within the NGC 6819 cluster and determined age to be $2.5Gyrs$ and in doing so, they better defined the relation between rotational period and age with a precision on the order of 10%.

The spin down of a star is the process of it losing angular momenta (and therefore rotational velocity) due to a torque created between the magnetic field of the star and its magnetic winds. The magnetic field arises in low mass stars from the motion of gas within convective envelopes, which leads to a magnetic field by dynamo action [6]. This leads to a torque acting between the fast moving magnetic field and the charged particles in the solar winds of the star, causing the star's rotation to decline. Stars with a greater mass $M \gtrapprox 1.2_{\odot}$ (which approximately corresponds to the temperature associated with the Kraft break [7]) have a very narrow convective envelope and thus the effect of magnetic breaking is heavily dampened.

An accurate prediction on a star's age is integral to many areas of astronomy as it allows us to better understand the time evolution of many astronomic phenomena. To properly study exo-planets, the age of a host star is needed to understand how a planetary system evolves over time. Studying our own planet suggests that the success of a bio-marker (chemical bi-products that are suggestive of life) will heavily depend on the age of the exo-planet and it's host star,as over a life-bearing planet's lifetime it's chemical composition changes drastically [8] [9]. Stellar aging is a key determinant in stellar archaeology [10], thus better predictions of the ages of stars may lead to advancements into how the early universe and our own galaxy formed.

## II. AIM

The aim of this project is to create a model that can accurately predict the age of low mass stars given an input of their rotational period, mass (or a proxy for mass) and possibly some other parameters such as metallicity. We aim for our model to be correct within at least 30% of it's true value. The model may be extended for all masses of stars, if it is deemed appropriate later on in the project, however for now we will be focused on sun like stars with masses $m$ in the range $0.8M_{\odot} < m < 1.15M_{\odot}$, for the reason mentioned above.

## III. METHOD

Gaussian processes (GP) will be used to create the model for our data. A GP is an infinite dimensional multivariate normal over all possible functions [11]. It is possible to marginalise over the GP to collapse the infinitely dimensional multivariate normal to an N dimensional multivariate normal, where N is the number of dimensions we are concerned with. For a multivariate normal with two vector inputs $x$ and $y$ the posterior is given by

$$P(x,y) = N\left(\begin{bmatrix}\mu_x \\ \mu_y\end{bmatrix}, \begin{bmatrix}\Sigma_x & \Sigma_{xy} \\ \Sigma_{xy}^T & \Sigma_y\end{bmatrix}\right) \tag{1}$$

where $\mu_x, \mu_y$ are the mean values of $x$ and $y$ and $\Sigma_x, \Sigma_y, \Sigma_{xy}$ are the covariance matrices of $xx$, $yy$ and $xy$ respectively. The probability of x is given by

$$p(x) = \int p(x,y)dy = N(\mu_x, \Sigma_x) \tag{2}$$

This can be generalised such that $y$ is an infinite set of variables we are not concerned with and $x$ is the finite set of variables we are interested it. In a physical interpretation, $y$ is the data and $x$ is the point to make predictions at. This leads to the ability to be able to make predictions about x given y through Bayes rule. The full derivation can be found in Rasmussen's 2003 book on the subject and leads to the result [12]:

$$P(x|y) = N(\mu_x + \Sigma_{xy}\Sigma_y^{-1}(y-\mu_y), \Sigma_x - \Sigma_{xy}\Sigma_y^{-1}\Sigma_{xy}^T) \tag{3}$$

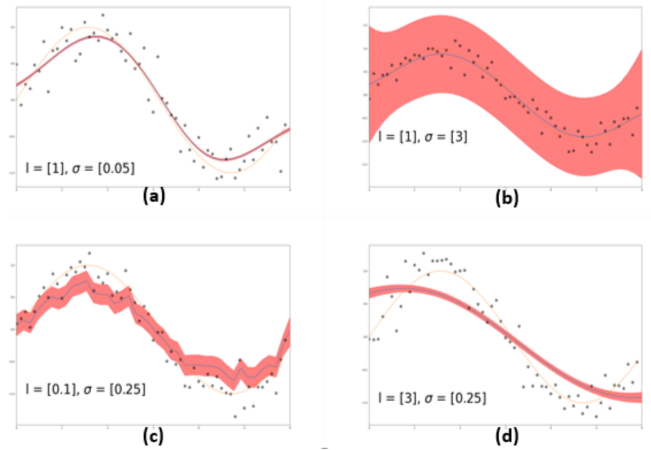


FIG. 1: Gaussian process fitting using various values of the length and signal variance applied to the same data set.

Gaussian processes are a form of lazy machine learning, as no generalisation is made on the data until a query to the system is made, that measures the similarity between points within the training data to make predictions on the value of other points [13]. Due to the Bayesian nature of GPs any prediction made has an intrinsic uncertainty to it in the form of a normal distribution. A GP can be fully described by its mean function $m(x)$ and it's covariance function $k(x, x')$. $m(x)$ is a user-defined function that determines the value that the Gaussian process will tend to in the absence of data and in most cases is set to zero. A suitable mean function is necessary when predicting values outside of the range of training data. The covariance function plays a greater role in shaping the model as it is what relates one point to another $(x \rightarrow x')$ and is defined by the kernel. There are several common kernels one might choose, the most common of which being the squared exponential [11]

$$k(x,x') = \sigma^2 exp\left[-\frac{(x-x')^2}{2l^2}\right] \tag{4}$$

where the signal variance $\sigma^2 > 0$ and the length scale $l > 0$. These two parameters are known as the hyper parameters of the kernel. A hyper parameter is a variable whose value is used to control the process. $l$ describes how smooth the function will be; if the length scale is large, the function can only change over a large period, whereas a smaller length scale will allow the function to change very quickly. $\sigma^2$ is a scaling factor that determines how far the function can deviate from the mean. Examples of different values of $l$ and $\sigma^2$ can be seen in Figure 1.

In order for the GP to best fit the data, the *maximum*

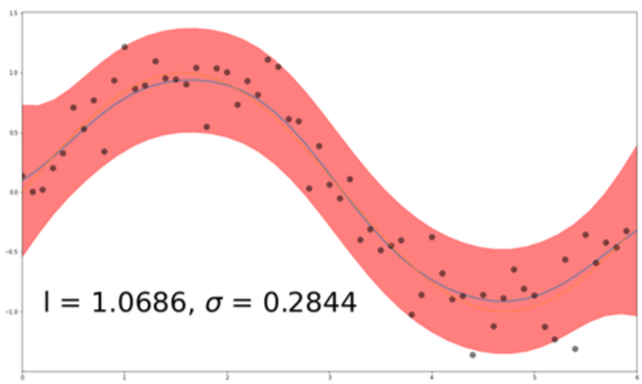


FIG. 2: Fitted Gaussian process with MAP hyper parameters

*a posteriori* (MAP) values of the hyper parameters must be determined. This is found by maximising the marginal likelihood given a distribution of the all combinations of the hyper parameters. In Figure 2 the GP with the MAP values for the original data set can be seen. For our application we expect the rotation of stars to change smoothly over large length scales, therefore larger values of the length parameter will be favoured.

The drawback of computing GPs is their $O\left(n^3\right)$ computational complexity which comes through having to invert matrices (see Equation 3). With data sets of several thousand points, this is still computable. However, for large data sets, $n > 10000$, a typical desktop computer will struggle to compute this in a practical amount of time. Fortunately, there are several workarounds to this issue. Firstly pymc4, the python module chosen to compute the GPs, which is set to be released at some point during the run of this project, will use a TensorFlow back end and therefore will be heavily optimised for use on GPUs [14]. Secondly, Google Colab will be used as the main development environment as it offers the use of fast CPUs, GPUs and TPUs freely available for public use [15]. Therefore, once pymc4 is released computation will be incredibly efficient and the $O\left(n^3\right)$ should not impede progress significantly. Finally, there are approximations that can be made to reduce the complexity such as sparse inference [16] [17]. The current version of pymc, pymc3, allows for sparse GP methods to be used .

GPs have been chosen as the primary tool to analyse our data set as opposed to neural nets (NN) as they possess several properties that better suit them to this task. It is well know that neural nets are over-confident as they give no measure of uncertainty when making decisions, hence uncertainty in NN predictions must be determined through other methods. When the data set they are studying is incredibly large they will tend towards the correct model, but for small data sets they are not reliable and even a well trained model may make incorrect decisions. Contrarily, GPs handle small data sets well, as they intrinsically provide a measure of uncertainty and as a result they are far better at making predictions of the model in areas where no data exists. To ensure our model works correctly we will only fit to data where there is a known value of star's age so that we can check that the model is making correct predictions. As the amount of available data may be small, GPs will excel over NNs in this environment.

## IV. TIMELINE AND PLAN

Outlined in Appendix A is the proposed timescale for the project. Week 1 is w/c 28/09/2020 and week 24 relates to w/c 08/03/2021. We are currently at the end of week four (as of the date of submission) in the plan and are on schedule. The bulk of the time has been attributed to developing our model, with several weeks for optimisation. We believe this large span of time to be justified as the success of the model is what will dictate the success of the entire project. Throughout the plan, time has been allocated to review the literature, this is to ensure our grasp of the subject is strong and to account for any scientific advancements that may occur over the lifetime of the project. Two weeks have been assigned at the end of the plan to account for any unforeseen circumstances. This plan will be used as a benchmark to measure progress, yet is subject to deviation if we believe it is called for. A private Github repository will be used to track progress as well as acting as a backup of all project work in case of unforeseen events.

There are several options for what the over-arching aim for this project could be. At this stage the two main ideas are application to exo-planets or stellar archaeology. Although limited research into these areas has been completed so far, we expect to independently explore these routes and use them as the focus of our individual projects. Although GPs form the body of the method described, there is a chance we may not be able to apply them to a data set as large as what is available - hence we may turn to neural nets as an alternative if GPs prove to be inadequate.

## V. RESULTS

It is expected that gyrochronology will be useful in determining the age of stars, as preliminary analysis on simulated data has proven that GPs are a sensible approach to such non-linear regression tasks. We analysed the relation between rotational period and age with data that was generated such that the rotational age was proportional to the square root of the period

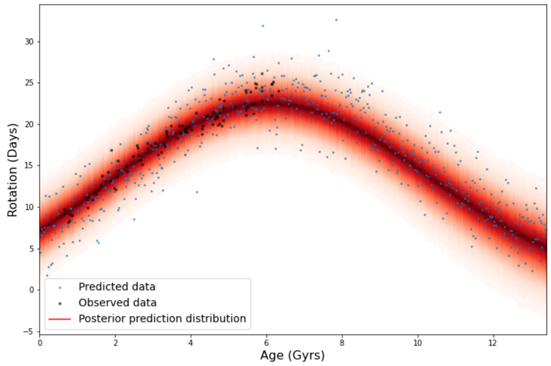


FIG. 3: Posterior predictions on sample gyrochronology data with MAP hyper parameters and zero mean function. Data was spaced across 100 points and samples were taken at 600.

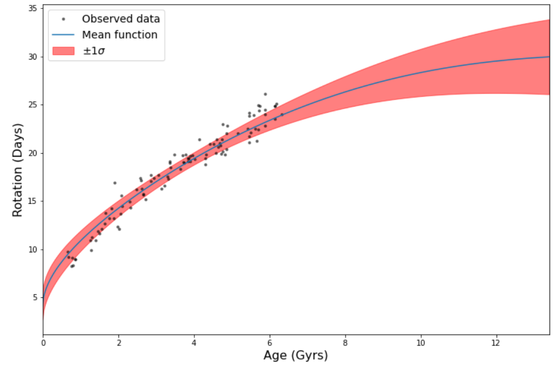


FIG. 4: Fitted function to data using GPs to sample gyrochronology data with MAP hyper parameters and square root mean function

and had Gaussian noise. The data only contained values for rotational period for ages $0.5Gyrs \rightarrow 6.5Gyrs$. The squared exponential kernel, defined in Equation 4, was chosen due to its versatility. As stated in the method, we expected a large length scale and such the value for this was determined from a Normal distribution with $\mu = 12, \sigma = 5$. We were uncertain what the value of $\sigma$ would be thus it was chosen from a Half-Cauchy distribution such that a wide range of positive values were possible. Initially a mean function of 0 was tried on the data; the model fitted well in regions where there was data, but was poor at predicting in regions where there was none. As the sampled ages went further from the data, predictions of the rotational period tended to zero, which did not reflect the expected behaviour, as seen in Figure 3.

To counter this, a square root mean function was used which resulted in a superior model that appeared to predict data better outside the region of data. The map values of the length and sigma parameters for the kernel were found to be $l = 4.52, \sigma = 1.36$. The resulting function fits the data well and appears to sensibly predict values for rotational period for ages in the entire range. The analysed data set contained 100 data points in the range $0.5Gyrs \rightarrow 6.5Gyrs$ and the posterior was sampled at 600 points in the range $0 \rightarrow 13.5Gyrs$ and the computation time was still short ($\approx 15minutes$).However, it is expected that for larger data sets or data sets with more than one input variable, the computational time may exceed what is reasonable. Therefore, preliminary analysis of sparse GPs was conducted which produced similar results to what is seen in Figure 4 considerably faster.

## VI. CONCLUSION

Gaussian processes proved successful in analysing basic simulated gyrochronology data in the preliminary tests. We hope that by analysing real world data with additional variables we will be able to determine whether the current understanding of gyrochronology is accurate. If no apparent relation is found, it may suggest that our current understanding of the spin-down of stars is incorrect, in which case the results of this project would be a strong argument for more research into this area. However, if this method is successful and can accurately predict the ages of stars to a reasonable accuracy, it can be reproduced and used to determine or check the ages of stars. It is expected that our findings will suggest that the truth lies somewhere between these two possibilities. There is already evidence to suggest that the effect of magnetic breaking is reduced for older stars [4] and we hope to draw a conclusion that will give strong insight into whether or not this is the case.

## VII. PROPOSED EXTENSIONS

If the project runs well, there are several options for extensions to the proposed project we could make:

- Compare the run time and accuracy of GPs and Neural nets - NN are typically faster for large data sets ($n > 10000$), but does the reduction in cost outweigh the potential loss in accuracy?
- Include more parameters into the input data for the GP (metallicity, size, temperature) - this may lead to a more accurate model
- Investigate the effects of accounting for the Rossby threshold for stars of mass $M > 1.15M_\odot$ [4] [18]

### Appendix A: Gantt chart

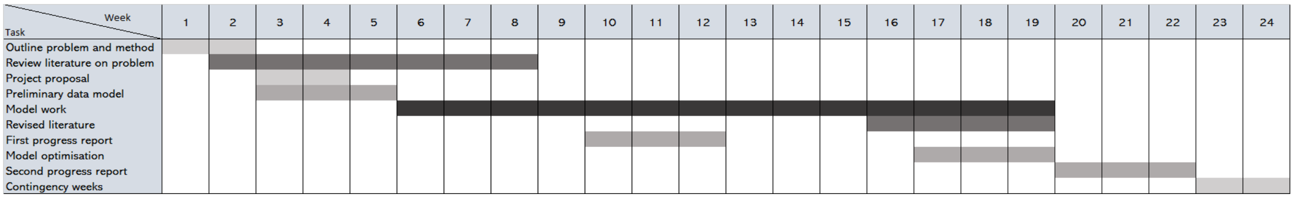


FIG. 5: Gantt chart to outline the proposed timeline of the project.


[1] A. Skumanich, Time Scales for Ca II Emission Decay, Rotational Braking, and Lithium Depletion, Astrophys. J. **171**, 565 (1972).

[2] S. A. Barnes, On the rotational evolution of solar- and late-type stars, its magnetic origins, and the possibility of stellar gyrochronology, The Astrophysical Journal **586**, 464 (2003).

[3] S. A. Barnes, Ages for illustrative field stars using gyrochronology: Viability, limitations, and errors, The Astrophysical Journal **669**, 1167 (2007).

[4] J. L. van Saders, T. Ceillier, T. S. Metcalfe, V. S. Aguirre, M. H. Pinsonneault, R. A. García, S. Mathur, and G. R. Davies, Weakened magnetic braking as the origin of anomalously rapid rotation in old field stars, Nature **529**, 181 (2016).

[5] S. Meibom, S. A. Barnes, I. Platais, R. L. Gilliland, D. W. Latham, and R. D. Mathieu, A spin-down clock for cool stars from observations of a 2.5-billion-year-old cluster, Nature **517**, 589 (2015).

[6] N. Ivanova and R. E. Taam, Magnetic braking revisited, The Astrophysical Journal **599**, 516 (2003).

[7] R. P. Kraft, Studies of stellar rotation. v. the dependence of rotation on age among solar-type stars, The Astrophysical Journal **150**, 551 (1967).

[8] J. F. Kasting, When methane made climate, Scientific American **291**, 78 (2004).

[9] L. Kaltenegger, W. A. Traub, and K. W. Jucks, Spectral evolution of an earth-like planet, The Astrophysical Journal **658**, 598 (2007).

[10] A. Frebel, Stellar archaeology: Exploring the universe with metal-poor stars, Astronomische Nachrichten **331**, 474 (2010), https://onlinelibrary.wiley.com/doi/pdf/10.1002/asna.201011362.

[11] N. Gibson, S. Aigrain, S. Roberts, T. Evans, M. Osborne, and F. Pont, A gaussian process framework for modelling instrumental systematics: application to transmission spectroscopy, Monthly notices of the royal astronomical society **419**, 2683 (2012).

[12] C. E. Rasmussen, Gaussian processes in machine learning, in *Summer School on Machine Learning* (Springer, 2003).

[13] M. Ebden, Gaussian processes: A quick introduction (2015), arXiv:1505.02965 [math.ST].

[14] J. Salvatier, T. V. Wiecki, and C. Fonnesbeck, Probabilistic programming in python using PyMC3, PeerJ Computer Science **2**, e55 (2016).

[15] Google, Google colab.

[16] L. Csató and M. Opper, Sparse on-line gaussian processes, Neural computation **14**, 641 (2002).

[17] R. Herbrich, N. D. Lawrence, and M. Seeger, Fast sparse gaussian process methods: The informative vector machine, in *Advances in neural information processing systems* (2003) pp. 625–632.

[18] B. Durney and J. Latour, On the angular momentum loss of late-type stars, Geophysical & Astrophysical Fluid Dynamics **9**, 241 (1977).