MNRAS **000**, 1–5 (2020) Preprint 27 January 2020 Compiled using MNRAS LATEX style file v3.0

PHYS6013 Midterm Report: A Machine Learning Approach to Cool Star Spin-Down

Ely Doodson, Supervisor: Cecilia Garraffo
Additional supervision from: Pavlos Protopapas, and Jeremy J. Drake

27 January 2020

ABSTRACT

Observations of young open clusters have shown a bimodal distribution in the rotation periods of cool stars. This bi-modality stems from stars having fast or slow rotation periods. The evolution of this trend through time suggests a fast transition from fast to slow rotating. Our current understanding of cool star spin down, through magnetic braking, accounts for the slow rotators branch, while the fast rotators remain somewhat of a mystery.

Our goal is to build a predictive probabilistic spin-down model that links the period of a star at any given mass and age. We use machine learning to predict the age at which each star transitions from fast to slow-rotation. Using a graphical model we will translate the distribution of initial periods into a rotation period probability distribution for a given mass and age.

1 INTRODUCTION

Stars are born from the collapse of clouds made of dust and gas. Such a cloud, though being made up of many molecules with their own random velocities, can be said to be spinning in one direction overall. Some of this angular momentum (AM) is sequestered in newly-formed stars which consequently rotate. However, their rotation rates are not static and decrease over a stars lifetime as a result of AM lost through stellar winds.

Cool stars, which are those classified as having a convective envelope (a turbulent layer in which convection carries a significant fraction of the outward energy flow) and a mass $\lesssim 1.3M_{\odot}$, have an interior magnetic dynamo powered by rotation which drives a magnetized wind. Ejected stellar material travels along magnetic field lines, effectively co-rotating with the star out to several stellar radii and carrying away much more AM than if the material was lost at the stellar surface.

The rate at which a star rotates drives the stellar strength of this magnetic breaking, the faster a star rotates, the stronger its magnetic dynamo and therefore the more prominent the winds that slow the star. This self-regulating magnetic breaking mechanism was expected to cause all the stars in an evolving population to converge to a unique relationship between rotation period and age that depends on mass. This is the basis for Gyrochronology, a technique that can be used to estimate the age of cool stars. However, when observing groups of coeval stars in open clusters (OC) with ages $< 1Gyrs$, a branch of fast rotating stars is also observed, see Figure 1. This fast rotators branch is a mystery and is not encompassed by current gyrochronological models, leaving a gap in their predictive capabilities. There is a model that attempts to explain spin-down using core-envelope decoupling in Barnes & Kim (2010), but fully convective stars have challenged this by showing the same spin-down (Newton et al. (2016); Douglas et al. (2017).) This work hopes to fill this gap.

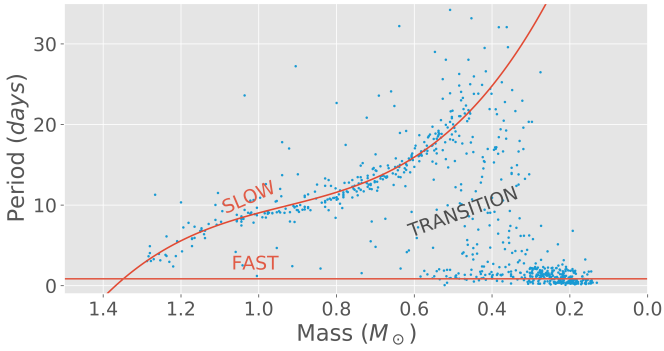

Figure 1. Plot of rotation period vs mass for the Praesepe young open cluster (age ≈ 790 Myrs), with visulisation of the slow, fast and transitional rotators.

Understanding the evolution of stellar rotation is crucial for understanding stellar magnetic activity. The magnetic dynamo generated by stellar rotation is responsible for many observed stellar characteristics such as star spots, UV/X-ray chromospheric and coronal emission, and the aforementioned magnetised winds driving stellar spin down.

So far our understanding of these magnetised winds is purely from the Sun. Stellar activity and wind create an energetic photon and particle radiation environment that has significant impact on the formation of planets and their subsequent atmospheric evolution. This is important for planets orbiting close to their host star which constitute the vast majority of detected rocky planets orbiting in the "habitable

zone". Understanding this will be crucial to the development of the rising field of exoplanet habitability.

It was shown in Garraffo et al. (2018) that, when accounting for the effects of surface magnetic field complexity on AM loss, the bimodal rotation population could be reproduced in simulations of stellar spin down. This was an analytical approach to the problem, however, a wealth of new data from Kepler, K2 and TESS, has provided the opportunity to approach gyrochronology from a more data-driven approach and form an even greater understanding of the parameters that constrain such systems.

The aim of this project is to utilize this growing database of observed stellar rotation periods and new numerical methods of machine learning to make a model capable of predicting a star's period, for a given mass and age.

2 OBSERVATIONS

2.1 Data Reduction

Though there are lots of field stars with known masses and periods, the difficulty lies in having a correct value for their ages. OC are coeval groups of stars, making them ideal for collecting data on period and mass when their ages are known. When observed in period and mass space, two distinct populations of fast and slow rotators can be seen after $\approx$ 10Myrs; before this time, stars are magnetically coupled with their stellar discs which modulate stellar rotation through "disc locking". In addition there are transitional stars between these two populations, which are less numerous, showing that the time to move between the populations must be rapid, shown in Figure 1.

Stellar rotation periods for OC were gathered from Beuther et al. (2014) and combined with our new additions to already existing OC measurements, such as for the M37 cluster Chang et al. (2015). Although all these catalogs contained values for period, mass was not often provided. Stellar masses were derived from catalogued multi-band photometry converted using stellar model photometric predictions. I used the Modules for Experiments in Stellar Astrophysics (MESA) Isochrones and Stellar Tracks (MIST) tables from Choi et al. (2016), which are simulations that provide information on the properties of stars, for a range of masses, evolved through time. I have obtained masses for all OC stars in our master catalogue and display them in Figure 2.

3 ANALYSIS METHODS

3.1 Unsupervised Clustering

Our initial approach was to model the two branches (fast and slow rotators) as different populations. In order to separate them we use *Unsupervised Clustering*. Clustering is the process of grouping data of similar properties together such as stars on one or the other rotation branch. The process being unsupervised means the data did not have a "correct" label telling us which group it belongs to, which one can use to reinforce and encourage certain clustering. This method would split the data into clusters that could then be fit with a polynomial regression. A polynomial regression is similar to linear regression, where you fit a straight line is fitted to the

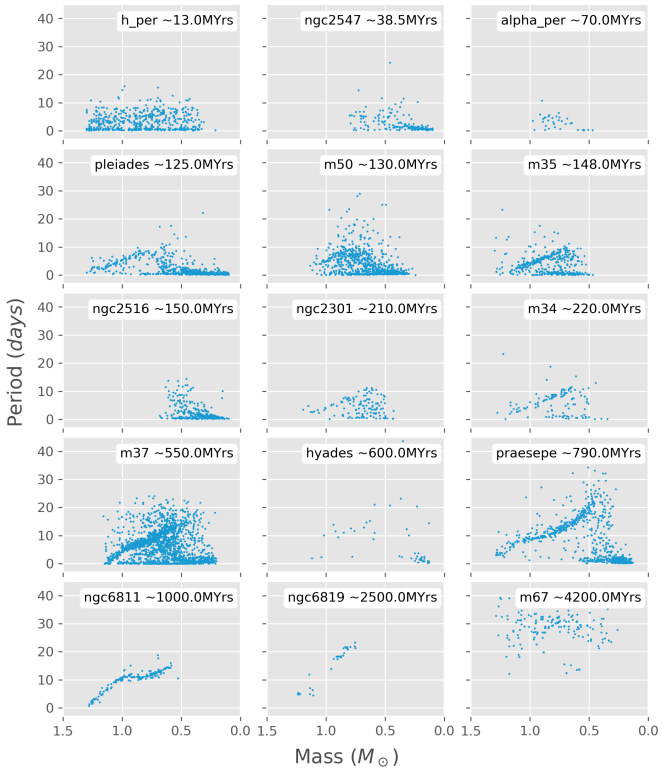

Figure 2. Plot of period versus mass of all converted OC and their respective ages

data $\hat{y} = b_0 + b_1 x$, but with a polynomial you allow p more higher order terms such as $...+b_2x^2+b_3x^3+b_4x^4+...+b_px^p$ to be considered such that

$$\hat{y} = b_0 + \sum_{j=1}^{p} b_j x^j \tag{1}$$

This allows, $\hat{y}$, to be a closer approximation to the true value, y, because the higher order terms allow the regression to change more rapidly, fitting the data more closely.

After a polynomial regression of the two groups, I cycle through each star in a cluster and assign it one or the other of the groups. After each star was assessed, a new polynomial fit was generated for each clustered group and the process repeated until stars no longer changed groups. The performance of this clustering is scored using the mean squared error (MSE), Equation 2, was minimised and reduced.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2}$$

and substituting in Equation 1

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \left(b_0 + \sum_{j=1}^{p} b_j x_i^j\right)\right)^2 \tag{3}$$

where n is the number of values, y_i is the true value, and $\hat{y}_i$ is the predicted value. A reduction in MSE means the predicted values are closer to the true values and a better fit is being generated. Minimising this MSE rewards the model that fits closer to the true data.

Figure 3 shows the results of this approach. The fits

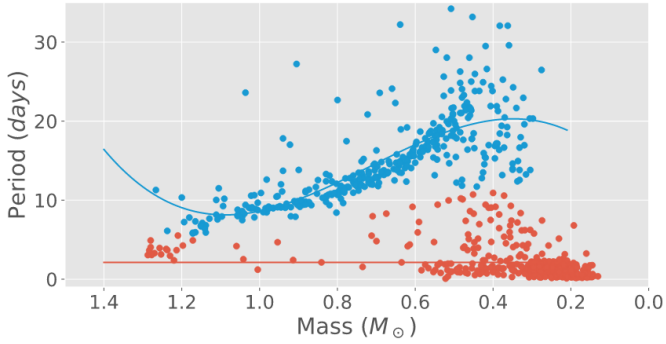

Figure 3. The results of unsupervised clustering on period versus mass data on the Praesepe cluster in which stars are assigned to either the fast or slow rotating groups. The fast rotators are shown in orange, slow rotators are shown in orange and their respective polynomial fits.

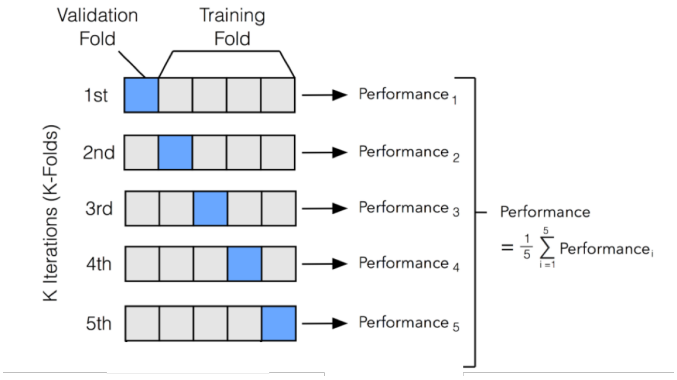

Figure 4. An illustration of how cross validation splits a data set and produces a score of that model, from github.io (2020).

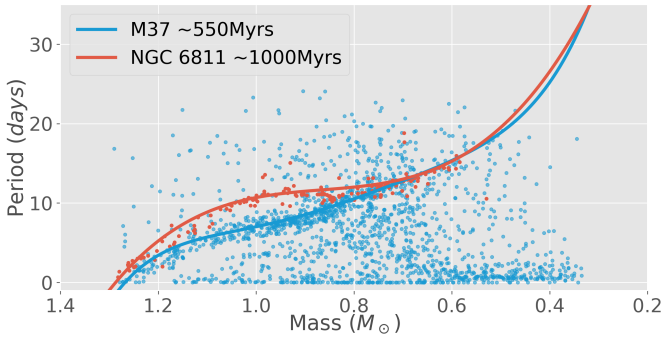

Figure 5. The rotation period as a function of stellar mass for stars in M37 in blue and NGC6811 in orange with their respective minimised polynomial fits.

generated were a polynomial $\mathcal{O}(x^3)$ for the slow rotators, and $\mathcal{O}(x)$ for the fast rotators. As shown, this approach does not generate an accurate representation of the two groups. This is because this modelling technique splits the data into two groups, that has an abrupt transition between them, however, in reality this is not the case and shown in the large transitional region in 1. As a result, this method neglects all the information concerning the transition of the stars from one regime to the other. For that reason, we decided to explore methods that would capture that.

3.2 Polynomial Ridge Regression

Our next approach was to fit a polynomial to a portion of the slow rotators and combine this polynomial with a sigmoid function, $\Phi(x)$ shown in Equation 6 below, to represent the transition between the fast and slow rotators. The sigmoid function maps any value of x to the range $0 \leq \Phi(x) \leq 1$, and can be thought of as switching between the two lines in Figure 1.

The method consisted of isolating just the slow rotators population for each cluster. In order to do this I binned the data according to mass and selected the closest 80% of the data to the mean to remove outliers and remaining transitional stars. This reduced data set was then fit using *Ridge Regression*, which scores using MSE, Equation 2, with the addition of a regularisation term, λ, which stops the coefficients becoming too large, shown in Equation 4. This regularisation coefficient, λ, and the degree of polynomial, p, was chosen via *Cross Validation*.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - b_0 - \sum_{j=1}^{p} b_j x_i^j\right)^2 + \lambda \sum_{j=1}^{p} b_j^2 \tag{4}$$

Cross validation is a method to choose the best model from a selection of models with varied parameters, such as we have here with varying polynomial degree and regularisation term. For each model, the data is split into a number of random subsets which maintain the original data distribution. One of these subsets is chosen as the validation set, and used to score the current model. The remaining subsets are then used to train the model on. The validation set used is then changed and the same process repeated until every subset has been a validation set. This ensures that all data has been used to train and validate the models. All these scores are then averaged to give a performance for the model. Figure 4 shows this for 5 subsets.

In Figure 5 I show the results of only the polynomial ridge regression for two clusters of different ages. It can be seen from the data and fits that the 450Myr difference has allowed the stars on the slow branch to reduce their spin further. A lack of the fast rotator population for NGC6811 can be seen in the region 0.5 $M_\odot \leq m \leq$ 1.3 $M_\odot$. This is because the higher mass stars transition first and have had time to transition. This optimised polynomial can then be combined with the sigmoid function, as seen in Equation 5, and minimised to produce our current model shown in Figure 6

$$\text{Period} = \left(b_0 + b_1 m + b_2 m^2 + \cdots + b_p m^p\right) * \Phi(x) \tag{5}$$

$$\Phi(x) = \frac{1}{1 + \exp^{-(\beta x + \alpha)}} \tag{6}$$

where β is the slope of the sigmoid and α is the displacement along the x-axis.

Originally, we aimed to interpret the time evolution of these polynomials as the evolution of spin. However, we

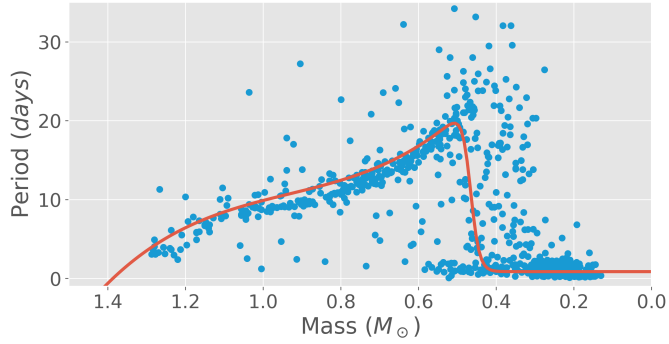

Figure 6. The period versus mass data of Praesepe and the resulting minimised polynomial ridge regression.

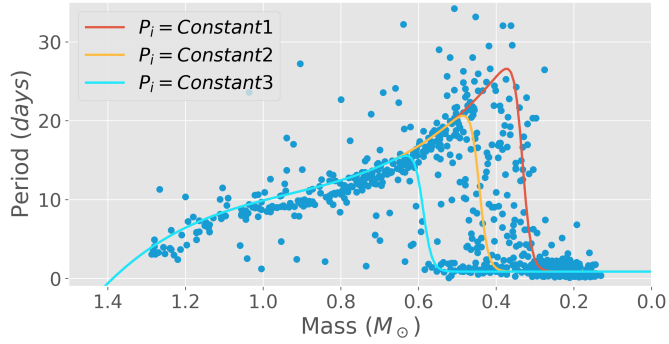

Figure 7. Theoretical lines that Praesepe's period versus mass data would follow if it was born with an initial period equal to a constant e.g. Constant1/Constant2/Constant3

found that the shape of the slow rotators changed in a way that required to change the order of the functional dependence (Curtis et al. (2019)).

The fit in Figure 6 is able to predict the majority of the slow and fast rotator branch, however, it becomes problematic for the region in which they overlap, which varies in size for different OC. We think this degeneracy is caused by a spread of initial periods and is discussed further in Section 3.3.

3.3 The Importance of Initial Period, P_i

We think the difficulty in these predictions of period, for a given age and mass, stem from the lack of information on the initial period distribution. In a hypothetical OC, whose stars are born with the same constant initial rotation period, we think the stellar distribution would follow one of the lines in Figure 7. However, in reality these observed OC are made up of a range of masses and initial periods, of which only the former can be measured. When taking an average of all these initial periods, we will get a distribution which looks like an observed OC, which has the degeneracy between the two branches due to the distribution of initial periods when the stellar population formed.

To try and address this, an estimate for initial period distribution can be made from the youngest cluster's distribution, that is no longer under the influence of disc effects, such as H Persei (h_per) in Figure 2.

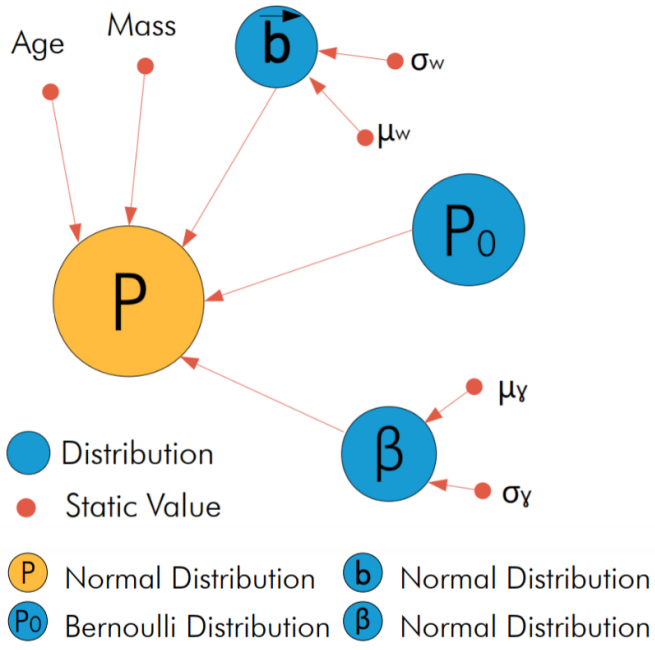

Figure 8. An illustration of how distributions of $\vec{b}$ $(= (b_0, b_1, b_2 ... b_p)$, β and static values will be combined to give a predicted period

4 FUTURE WORK

So far models have been driven heuristically. Future attempts will be derived from first principles, meaning they will be drawn from fundamental distributions rooted in a physical understanding of the system, and, using inference, will give us confidence intervals on a prediction: directly addressing the degeneracy present in earlier models. This will provide us with a probabilistic graphical model, displayed in Figure 8.

By 03/02/2020, the aim is to solidifying my understanding of the graphical model, how it applies to the problem and if it is an appropriate modeling technique. Days until 17/02/2020 will be spent understanding Markov Chain Monte Carlo (MCMC) methods and probabilistic coding modules "Pyro" and "PyTorch". By the end of March, a determined model will be obtained and test its viability. From April onwards time will be used to write the paper presenting the finding.

Ultimately we hope to obtain a paper that lays out a model that can accurately recreate and predict period of a star with a known age and mass. However, if an adequate model is not found there are many approaches with Neural Networks that can predict with a quantified error.

REFERENCES

Barnes S. A., Kim Y.-C., 2010, ApJ, 721, 675
Beuther H., Klessen R. S., Dullemond C. P., Henning T. K., 2014, Protostars and planets VI. University of Arizona Press

Chang S. W., Byun Y. I., Hartman J. D., 2015, AJ, 150, 27
Choi J., Dotter A., Conroy C., Cantiello M., Paxton B., Johnson B. D., 2016, The Astrophysical Journal, 823, 102
Curtis J. L., Agüeros M. A., Douglas S. T., Meibom S., 2019, ApJ, 879, 49
Douglas S. T., Agüeros M. A., Covey K. R., Kraus A., 2017, ApJ, 842, 83
Garraffo C., et al., 2018, The Astrophysical Journal, 862, 90
Newton E. R., Irwin J., Charbonneau D., Berta-Thompson Z. K., Dittmann J. A., West A. A., 2016, ApJ, 821, 93
github.io E., 2020, K-fold Cross Validation Visual Description, http://ethen8181.github.io/machine-learning/model_selection/model_selection.html

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