MNRAS **000**, 1–4 (2020) Preprint 23 January 2020 Compiled using MNRAS LATEX style file v3.0

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

Ely Doodson,[1,2????,3???⋆] Cecilia Garraffo,[2,3] Pavlos Protopapas[3] and Jeremy J. Drake[2]
[1] *School of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, United Kingdom*
[2] *Harvard-Smithsonian Center for Astrophysics, 60 Garden St, Cambridge, MA 02138, United States*
[3] *Institute for Applied Computational Science, Harvard University, Cambridge, MA 02138, United States*

23 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 starat 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 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. This cloud, though being made up of many molecules with their own random velocities, can be said to have an overall spin which, when the star collapses, will be the axis of rotation when the star is born. However, this spin rate is not static and will descrease over a stars lifetime, assuming no companion or outside influence.

Cool stars are classified as having a convective envelope, meaning the outer-most layer of the star is moving. They have a mass, $m \lesssim 1.3\mathrm{M}_{\odot}$. Due to the convection, moving ionic material in the star generates massive magnetic fields which stretch in orders of stellar radii. (SOMEONE I NEED TO CITE?) Ejected stellar material travels along these lines, forming large arms which effectively co-rotate with the star. When material at the end of these arms breaks free, the loss in angular momentum (AM) is much grater than if the same material was lost at the stars surface. This loss in AM causes the star to lessen its rotation period and spin down. This is called *magnetic breaking* and is a very effecient way for a star to spin down. (CITE HERE ALSO??? WHO?)

Understanding the evolution of stellar rotation is crucial for understanding stellar magnetic activity as this 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 dervided from our study of our own Sun so our understanding of their origin is therefore incomplete; stellar spin-down provides a powerful probe of the physics of these winds. Stellar activity and wind create an energetic photon and particle rich, radiation enviroment that has significant impact on the formation of planets and their subsequent atmospheric evolution. Understanding this will be crucial to the development of the rising field of exoplanet habitability.

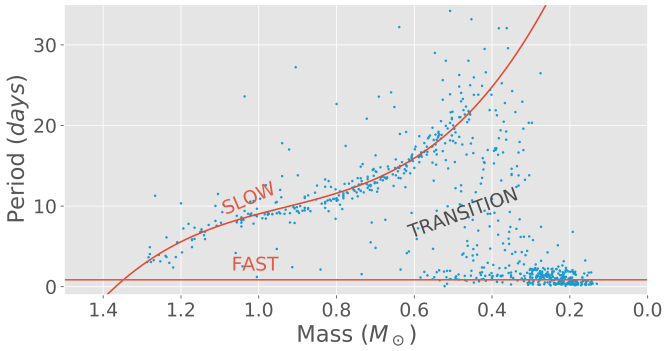

Figure 1. Plot of period vs mass for Praesepe, with visulisation of the slow, fast and transitional rotators.

It was shown in Garraffo et al. (2018) that, when accounting for the breaking due to these magnetic winds and AM loss, this bimodal population could be seen with an evolved simulation of stars. This was an analytical approach to the problem, however, with a wealth of new data from Kepler, K2 and TESS, this has provided the opportunity to approach gyrochronology from a more data driven approach and form an even greater understanding of the parameters that contrain such systems and how they will affect(effect?) the afformentioned fields. 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

Open clusters are coeval groups of stars and, when obeserved in period and mass space, two distinct populations of fast and slow rotators can be seen after ≈ 10Myrs; before this, the stars have not had time to transition and the stellar disc effects make it very difficult to paramterise and model. In addition there are transitional stars between these two populations, which are less numerous than the two populations, showing the time to move between the populations must be rapid. This can be seen in Figure 1.

Though there are lots of field stars with known masses and periods, the difficulty lies in having a correct value for their ages. This makes OC, with known ages, ideal for collecting data on period and mass, which can be used readily in the expansion of gyrochronological models. The vast majority of OC data for this project was gathered from Beuther et al. (2014) as well as some new additions to already existing OC M37,**?**. Although all these catalogs contained values for period, mass was not often provided. The magnitude and waveband(frequency maybe?) of light is corralated heavily with the mass of a star, so this is the parameter one observes when measuring the mass of a star. Each catalog presented their data in a different band of light, which make it difficult to convert them all to mass as different conversions were needed. These conversions are complex for each band and vary greatly as a function of mass, with some conversions not covering our lowest masses.

A conversion was possible, however, it did not use the conventional functions to map photometry to mass. Instead 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 converted all OC and display them in Figure 2.

These tracks start at discrete mass steps (e.g 0.1, 0.15...1.35, 1.40 $M_\odot$ etc). These masses, evolving at differnt rates, are very likely to have a degeneracy in their photomerty, meaning one star can have the same photomerty value as another at a different age and mass, and as a consequence the conversion was not as simple as choosing the two closest photometries and interpolating. Instead I restricted the values of the various mass tracks to the clostest ages of that star and from this pool, interpolated between the two clostest photometries for that given age. The discontinuity of the tracks, due to discrete time steps, means there will be an inherent error in choosing the pool of ages, this has not currently been addressed, however may be implemented into error propagation in a future model if a Bayesian network approach is used.

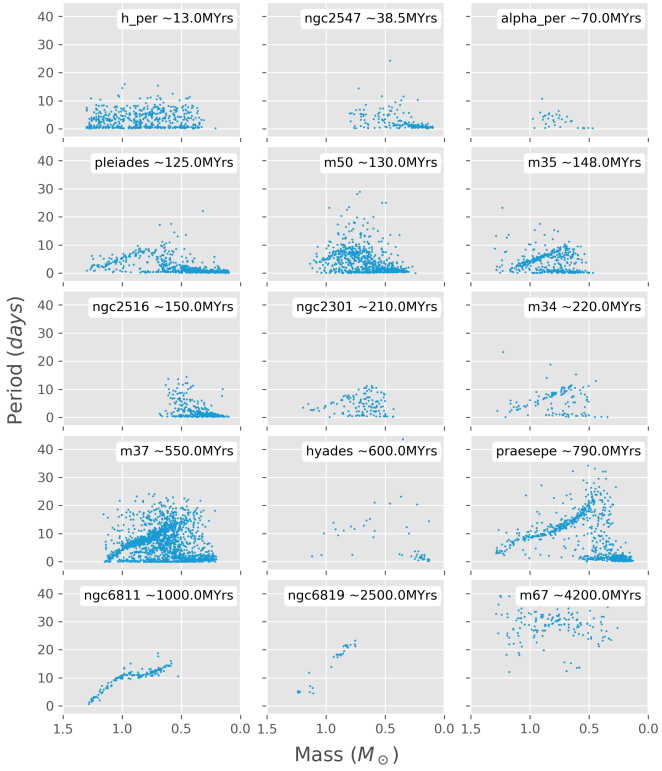

Figure 2. Plot showing all converted OC and their respective ages

3 METHODS

3.1 Unsupervised Clustering

Our inital approach to the problem was to use *Unsupervised Clustering*. Clustering is the process of grouping data of similar properties together. 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 clustering would split the data into "fast" and "slow" rotators and we could then fit a polynomial regression to each of these groups. A polynomial regression is like linear regression, where you fit a straight line to you data $\hat{y} = b_0 + b_1 x$, but with polynomial you allow n more higher order terms such as $... + b_2 x^2 + b_3 x^3 + b_4 x^4 + ... + b_n x^n$ to be considered and becomes...

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

This allows your predicted value, $\hat{y}$, to be predicted closer to the true value, y.

After a polynomial regression of the two groups cycle through each star in the cluster and assign it to the other group, if the overall fit is better with the star in the opposing group, it remained there, otherwise it is transfered back and the next star was assessed. After each star was assessed, a new polynomial fit was generated for each cluster the process repeated until stars no longer changed groups. To measure the score, mean squared error(MSE), Equation

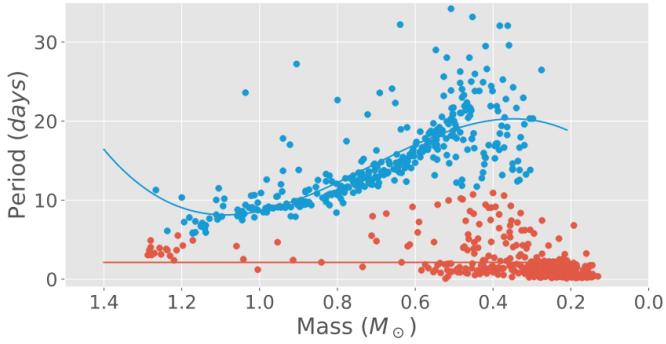

Figure 3. shows the results of unsupervised clustering of Praesepe. The "fast" rotators in blue, "slow" rotators in orange and their respective polynomial fits.

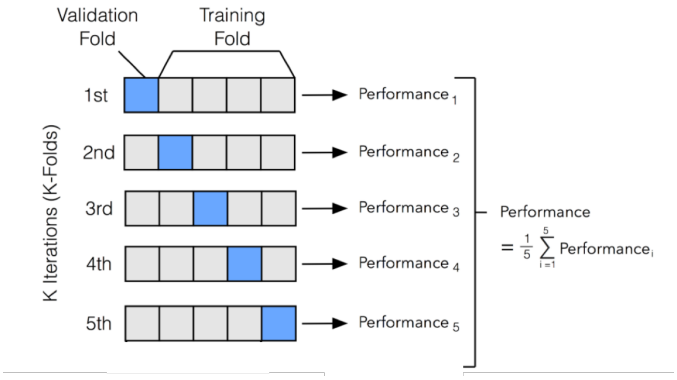

Figure 4. shows how cross validation splits a data set and produces a score of that model, from github.io (2020).

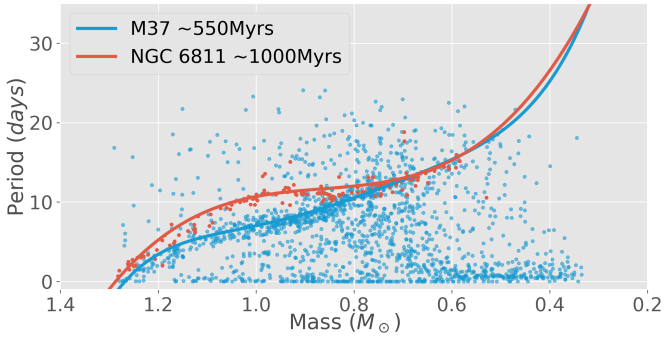

Figure 5. shows M37 in blue and NGC6811 in orange with their respective best fits.

2, was minimised and reduced.

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

subbing in Equation 1

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

where y_i is the true value, and $\hat{y}_i$ is the predicted value.

A reduction in MSE means the predicted values are close to the true values and a good fit is being generated, if we minimise this MSE then we are rewarding the model with a fit closer to the true data.

Figure 3 shows the results off this approach. The fits generated were a polynomial $\mathcal{O}(x^3)$ for the blue "slow" rotators, and $\mathcal{O}(x)$ for the orange "fast" rotators. As it can be seen, this approach does not generate an accurate representation of the two groups. This is because the transition between these two clusters is effectively instantaneous, however, in reality this is not the case. As we also wanted to try understand how these stars transitioned throughout their lifetime, this approach would not extract this kind of information and so clustering was deemed an uneffective method.

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 5, to have the transition between the fast and slow rotators. The sigmoid function, shown in Equation 6, maps any value of x to the range $0 \leq \Phi(x) \leq 1$, this can be thought of as switching between the two lines in Figure 1.

The method was to remove the "slow" and as many transitional rotators for each cluster as possible, by eye. From this subset of stars I created 10 bins according to mass and reduced this subset to the central 80% to remove outliers and and remaining transitional stars.(Is a figure needed to show the reduction?) This reduced data set was then fit using *Ridge Regression*, which scores using MSE, Equation 2, with the addition of 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*.

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

Cross validation is a method to choose the best model from a selection, such as we have here with varying polynomial degree and regularisation term. For each model, your 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 your current model. The remaining subsets are then used to train your 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 your 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 450Myrs has allowed the stars on the "slow" branch to reduce their spin further. It can also be seen that there is a lack of a "fast" population for NGC6811 in the region 0.5 $M_\odot \leq m \leq$ 1.3 $M_\odot$. This is because all these stars have had time to transition.

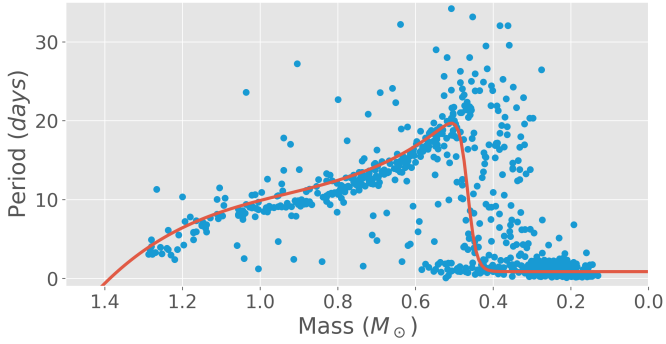

Figure 6. shows an optimised fit using ridge regression and sigmoid for the OC Praesepe

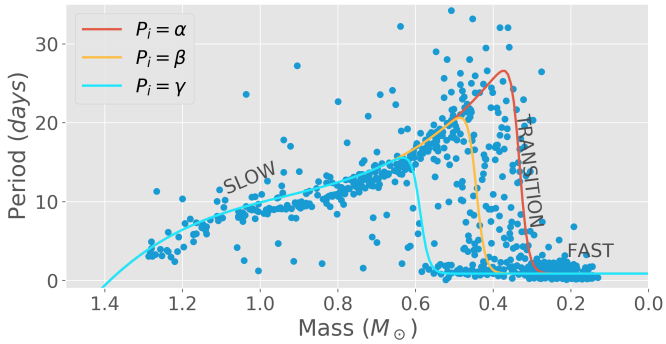

Figure 7. shows what we hypothesise the OC star distribution would follow for 3 different initial periods, α, β & γ.

This optimised polynomial can then be combined with the sigmoid function($\Phi(x)$), as seen in Equation 5, and minimised to produce our current model shown in Figure 6

$$\text{Period} = \left(b_0 + b_1 m + b_2 m^2 ... b_i m^j\right) * \Phi(x) \tag{5}$$

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

The coefficients of these polynomials were thought to be the key to understanding the progression of the system. However, it was found that there was no obivous trend to these coefficients through time. The fit in Figure 6 is able to predict the majority of the "slow" and "fast" rotator branch, however, becomes problematic for the region 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.

A potential solution to this would be to calculate an error associated with each point of the fit. A region of uncertainty around the prediction would be produced, which would be low if the data spread was minimal and high when approaching the overlapped area.

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 of a single initial rotation period, we think the distribution would look like 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. The different lines in Figure 7 represent theoterical distributions, provided different values of initial period. When taking an average of all these initial periods, we will get a distribution which looks like the observed OC, which has a degeneracy between the two branches.

To try and address this, an estimate for inital 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.

4 FUTURE WORK?

The degeneracy discussed in previous sections is a large motivator to approch this model probabilistically. Instead of generating a polynomial regression and modelling its evolution through time, a heat map of the likelihood of a star being in that spot for a given age and mass could be generated. One could then sample stars of random masses for a given age to generate a synthetic population of stars, whos distribution looks like that of an observed OC.

ACKNOWLEDGEMENTS

My amazing supervisors < 3

The Acknowledgements section is not numbered. Here you can thank helpful colleagues, acknowledge funding agencies, telescopes and facilities used etc. Try to keep it short.

REFERENCES

Beuther H., Klessen R. S., Dullemond C. P., Henning T. K., 2014, Protostars and planets VI. University of Arizona Press

Choi J., Dotter A., Conroy C., Cantiello M., Paxton B., Johnson B. D., 2016, The Astrophysical Journal, 823, 102

Garraffo C., et al., 2018, The Astrophysical Journal, 862, 90

github.io E., 2020, K-fold Cross Validation Visual Description, http://ethen8181.github.io/machine-learning/model_selection/model_selection.html

APPENDIX A: SOME EXTRA MATERIAL

If you want to present additional material which would interrupt the flow of the main paper, it can be placed in an Appendix which appears after the list of references.

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