

## ANALYSING KEPLER STELLAR SURFACE ROTATION AND ACTIVITY WITH ROOSTER

S. N. Breton<sup>1</sup>, A. R. G. Santos<sup>2</sup>, S. Mathur<sup>3,4</sup>, R. A. García<sup>1</sup>, L. Bugnet<sup>5</sup> and P. L. Pallé<sup>3,4</sup>

**Abstract.** It is crucial for our knowledge of stellar evolution to be able to efficiently determine stellar surface rotation periods in large stellar samples. Random forest (RF) learning abilities are exploited to automate the extraction of rotation periods in *Kepler* light curves. We train three different classifiers: one to detect if rotational modulation is present in the light curve; one to select the rotation period among estimates provided by ACF and wavelet analysis methods; and finally one to flag classical pulsators or close binary candidates that can bias our rotation-period determination. We test our machine learning pipeline, ROOSTER, on the *Kepler* K and M dwarf sample using the most up-to-date reference catalog. We show that we are able to detect rotational modulations with an accuracy of 94.2% and to retrieve final rotation periods with an accuracy of 95.3%. This value is raised to 99.5% after visually inspecting 25.2% of the stars. Over the two main analysis steps, the pipeline yields a global accuracy of 92.1% before visual checks, 96.9% after. The method is then applied to analyse the F and G stars observed by *Kepler*. The methodology presented here can be adapted to extract surface rotation periods for stars observed by other missions, like K2, TESS, and PLATO.

Keywords: Methods: data analysis – Stars: solar-type – Stars: activity – Stars: rotation – starspots

### 1 Introduction

Corotating dark spots and bright faculae on the stellar surface lead to brightness variations (e.g. Berdyugina 2005; Strassmeier 2009). Therefore, the long-term photometric surveys performed by a space instrument like *Kepler* (Borucki et al. 2010) provides ideal datasets to measure stellar surface rotation periods and build stellar rotation catalogs. These rotation catalogs can then be used to constrain gyrochronology models (e.g. Barnes 2003, 2007; Mamajek & Hillenbrand 2008; Meibom et al. 2011; García et al. 2014) in order to provide an estimate of the age of each considered target. Understanding the origin of the discrepancies between the different methods used to estimate stellar ages (for example between asteroseismology and gyrochronology) is a key issue in stellar physics (see e.g. Angus et al. 2015; van Saders et al. 2016). They are also of greatest interest to study the interplay between rotation and magnetic activity (e.g. Mathur et al. 2014) and yield information directly related to the dynamics at a given time for planetary system, which is crucial when considering star-planet interactions (e.g. Zhang & Penev 2014; Mathis 2015; Bolmont & Mathis 2016; Strugarek et al. 2017; Benbakoura et al. 2019). One of the main challenge that we face today on the observational side is finding efficient and reliable methods to analyse the large amount of collected data in large-scale surveys. Over the last years, automatic methods have been implemented to perform classification tasks related to stellar physics (e.g. Blomme et al. 2011; Armstrong et al. 2016; Bass & Borne 2016; Bugnet et al. 2019; Kuzlewicz et al. 2020; Audenaert et al. 2021). We focus here on the possibilities and outcomes offered by random forest classification methods to deal with stellar surface rotation in a photometric survey like the one completed by *Kepler*

<sup>1</sup> AIM, CEA, CNRS, Université Paris-Saclay, Université de Paris, Sorbonne Paris Cité, F-91191 Gif-sur-Yvette, France

<sup>2</sup> Department of Physics, University of Warwick, Coventry, CV4 7AL, UK

<sup>3</sup> Instituto de Astrofísica de Canarias, 38200 La Laguna, Tenerife, Spain

<sup>4</sup> Universidad de La Laguna, Dpto. de Astrofísica, 38205 La Laguna, Tenerife, Spain

<sup>5</sup> Flatiron Institute, Simons Foundation, 162 Fifth Ave, New York, NY 10010, USA

## 2 Analysing surface rotation with ROOSTER

The rotational analysis method combining the auto-correlation function (ACF; McQuillan et al. 2013) and the time frequency analysis of the wavelet power spectrum (WPS; Torrence & Compo 1998; Mathur et al. 2010; García et al. 2014) has shown to be the method able to provide the most complete and reliable set of rotation period estimates during the hare and hounds exercise performed by Aigrain et al. (2015). However, to be applied to the much larger sample of K and M stars, this required to perform a lot of visual inspections as it was done in Santos et al. (2019, hereafter S19). The rotation catalog of S19 was built after visually inspecting about 60% of the considered sample. The Random fOrest Over STEllar Rotation pipeline (ROOSTER, Breton et al. 2021) was designed to reduce the number of light curves and respective rotation diagnostics that have to be visually inspected when analysing large samples. ROOSTER is composed of three random forest (RF) classifiers. The first one is dedicated to assess the presence of a rotational modulation in the considered star. The second one selects the period linked to the rotational modulation among the different estimates yielded by the ACF/WPS rotational analysis. The performance of those classifiers have been assessed with the KM sample of S19. We obtain an accuracy of 94.2% and 95.3% for the first and second step, respectively. Designing a strategy that required the visual inspection of only 25.2% of the stars in the sample, a significant increase from what had to be done for S19, we were able to raise the score of the second classifier to 99.5%. The accuracy estimation over the two steps of analysis was therefore of 92.1% and 96.9%, before and after visual inspection, respectively. The last classifier of the ROOSTER framework is an auxiliary RF classifier which flags close-binary or classical-pulsator candidates, that is targets for which signals in the light curves can be confused with rotational modulation (see S19).

## 3 First outcomes from the new rotation catalog

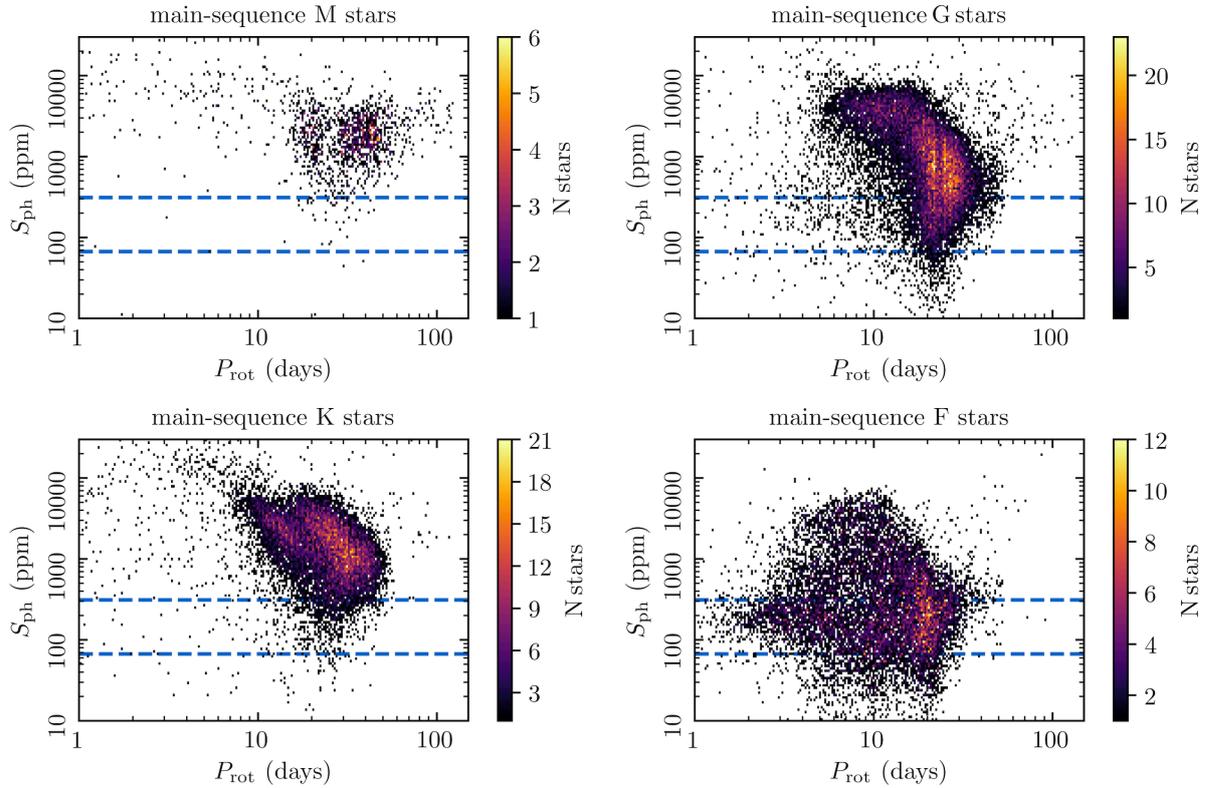
Having been properly evaluated, ROOSTER abilities were exploited when analysing the FG main-sequence and subgiant sample of *Kepler* (Santos et al. 2021, hereafter S21). In total, 159,442 *Kepler* targets were analysed in S19 and S21, yielding rotation period  $P_{\text{rot}}$  and photometric magnetic activity index  $S_{\text{ph}}$  (see Mathur et al. 2014, for the definition) for 55,232 stars. For each spectral type, F, G, K and M,  $P_{\text{rot}}$  and  $S_{\text{ph}}$  values are represented together in the panels of Fig 1. In the four panels, the existence of a saturation regimes at high activity and short rotation period is apparent, particularly for G stars. Magnetic activity intensity then decreases as the star spins down due to stellar rotational braking. The rotation period for the slowest, less active rotators is difficult to constrain from the data collected by an instrument like *Kepler* due to the small amplitude of the long-term rotational modulation. This might explain why this region of the diagram is not populated with the data presented here. While G and K stars experience magnetic braking along their evolution, this mechanism is less efficient for earliest F-type stars, which explains why the F-star population present a larger proportion of fast rotators than cooler stars. The dynamo is also expected not to be very vigorous for those stars, hence the small  $S_{\text{ph}}$  values.

## 4 A look at the KOIs

ROOSTER was also used to perform an analysis of the stellar rotation of the *Kepler* Objects of Interests (KOIs Brown et al. 2011), that is to say, host stars of confirmed or candidate planetary objects. It is expected that magnetic activity can be triggered by star-planet magnetic interactions (e.g. Strugarek et al. 2019). Figure 2 shows the orbital period  $P_{\text{orb}}$ - $S_{\text{ph}}$  diagram for the confirmed planet-host stars. Due to *Kepler* detection biases (the transit method favours detection of planets with large radius and short orbital periods), it is expected that the Earth is situated in the bottom right corner of the figure with only a few other planets. No clear correlation pattern between  $P_{\text{orb}}$  and  $S_{\text{ph}}$  appears in this diagram.

## 5 Conclusions

Data analysis machine learning methods are increasingly relevant for astronomical purposes as we nowadays face the challenge of analysing data from surveys with hundreds of thousands of stars (or even millions, for an instrument like TESS). We showed here how exploiting the abilities of RF classifiers with the ROOSTER pipeline allowed us to significantly reduce the amount of visual inspections required to build the new *Kepler* FG main-sequence and subgiants catalog of S21. We commented on the first outcomes of the complete *Kepler*



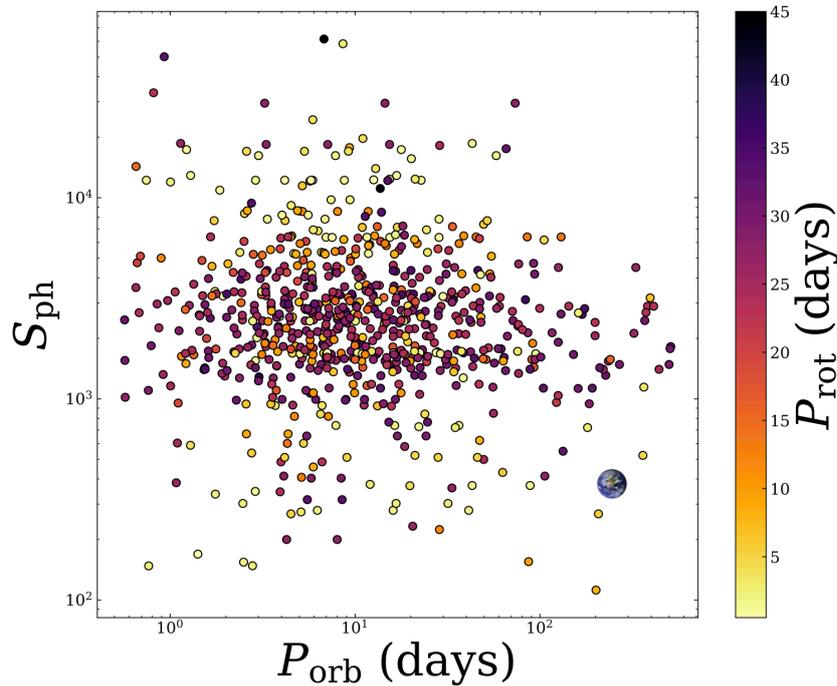
**Fig. 1.** Magnetic activity index  $S_{\text{ph}}$  as a function of the rotation period for M (*top left*), K (*bottom left*), G (*top right*) and F (*bottom right*) stars. Adapted from Santos et al. (2019, 2021).

FGKM catalog that is now available through the work of S19 and S21. In particular, we were able to show that the shape of the rotation period-magnetic activity diagram shape is significantly different depending on the spectral type of the considered population of targets. Finally, we emphasised that our first analysis of the confirmed planet-host stars did not reveal a particular relation between planetary orbital period and magnetic activity level. This preliminary work needs to be refined and extended to include a comparison with the *Kepler* sample without detected planet.

S.N.B. and R.A.G. acknowledge the support from PLATO and GOLF CNES grants. A.R.G.S. acknowledges the support from STFC consolidated grant ST/T000252/1 and NASA grant NNX17AF27G. S.M. acknowledges the support from the Spanish Ministry of Science and Innovation with the Ramon y Cajal fellowship number RYC-2015-17697. P.L.P. and S.M. acknowledge support from the Spanish Ministry of Science and Innovation with the grant number PID2019-107187GB-I00. This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program.

## References

- Aigrain, S., Llama, J., Ceillier, T., et al. 2015, *MNRAS*, 450, 3211  
 Angus, R., Aigrain, S., Foreman-Mackey, D., & McQuillan, A. 2015, *MNRAS*, 450, 1787  
 Armstrong, D. J., Kirk, J., Lam, K. W. F., et al. 2016, *MNRAS*, 456, 2260  
 Audenaert, J., Kuzlewicz, J. S., Handberg, R., et al. 2021, arXiv e-prints, arXiv:2107.06301  
 Barnes, S. A. 2003, *ApJ*, 586, 464  
 Barnes, S. A. 2007, *ApJ*, 669, 1167  
 Bass, G. & Borne, K. 2016, *MNRAS*, 459, 3721



**Fig. 2.** Photometric magnetic activity index  $S_{\text{ph}}$  as a function of the orbital period  $P_{\text{orb}}$  of the confirmed KOIs. The surface rotation period of the star,  $P_{\text{rot}}$ , is color-coded.

- Benbakoura, M., Réville, V., Brun, A. S., Le Poncin-Lafitte, C., & Mathis, S. 2019, *A&A*, 621, A124
- Berdyugina, S. V. 2005, *Living Reviews in Solar Physics*, 2, 8
- Blomme, J., Sarro, L. M., O'Donovan, F. T., et al. 2011, *MNRAS*, 418, 96
- Bolmont, E. & Mathis, S. 2016, *Celestial Mechanics and Dynamical Astronomy*, 126, 275
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Science*, 327, 977
- Breton, S. N., Santos, A. R. G., Bugnet, L., et al. 2021, *A&A*, 647, A125
- Brown, T. M., Latham, D. W., Everett, M. E., & Esquerdo, G. A. 2011, *AJ*, 142, 112
- Bugnet, L., García, R. A., Mathur, S., et al. 2019, *A&A*, 624, A79
- García, R. A., Ceillier, T., Salabert, D., et al. 2014, *A&A*, 572, A34
- Kuszlewicz, J. S., Hekker, S., & Bell, K. J. 2020, *MNRAS*, 497, 4843
- Mamaejek, E. E. & Hillenbrand, L. A. 2008, *ApJ*, 687, 1264
- Mathis, S. 2015, *A&A*, 580, L3
- Mathur, S., García, R. A., Ballot, J., et al. 2014, *A&A*, 562, A124
- Mathur, S., García, R. A., Régulo, C., et al. 2010, *A&A*, 511, A46
- McQuillan, A., Mazeh, T., & Aigrain, S. 2013, *ApJ*, 775, L11
- Meibom, S., Barnes, S. A., Latham, D. W., et al. 2011, *ApJ*, 733, L9
- Santos, A. R. G., Breton, S. N., Mathur, S., & García, R. A. 2021, *ApJS*, 255, 17
- Santos, A. R. G., García, R. A., Mathur, S., et al. 2019, *ApJS*, 244, 21
- Strassmeier, K. G. 2009, *A&A Rev.*, 17, 251
- Strugarek, A., Bolmont, E., Mathis, S., et al. 2017, *ApJ*, 847, L16
- Strugarek, A., Brun, A. S., Donati, J. F., Moutou, C., & Réville, V. 2019, *ApJ*, 881, 136
- Torrence, C. & Compo, G. P. 1998, *Bulletin of the American Meteorological Society*, 79, 61
- van Saders, J. L., Ceillier, T., Metcalfe, T. S., et al. 2016, *Nature*, 529, 181
- Zhang, M. & Penev, K. 2014, *ApJ*, 787, 131