The Astronomical Journal, 160:168 (15pp), 2020 October https://doi.org/10.3847/1538-3881/abada4

Astraea: Predicting Long Rotation Periods with 27 Day Light Curves

Yuxi(Lucy) Lu[1,2], Ruth Angus[1,2,3], Marcel A. Agüeros[1], Kirsten Blancato[1], Melissa Ness[1], Danielle Rowland[2], Jason L. Curtis[2], and Sam Grunblatt[2,3]
[1] Department of Astronomy, Columbia University, 550 West 120th Street, New York, NY, USA; lucylulu12311@gmail.com
[2] American Museum of Natural History, Central Park West, Manhattan, NY, USA
[3] Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, Manhattan, NY, USA
Received 2020 June 24; revised 2020 August 7; accepted 2020 August 7; published 2020 September 21

Abstract

The rotation periods of planet-hosting stars can be used for modeling and mitigating the impact of magnetic activity in radial velocity measurements and can help constrain the high-energy flux environment and space weather of planetary systems. Millions of stars and thousands of planet hosts are observed with the Transiting Exoplanet Survey Satellite (TESS). However, most will only be observed for 27 contiguous days in a year, making it difficult to measure rotation periods with traditional methods. This is especially problematic for field M dwarfs, which are ideal candidates for exoplanet searches, but which tend to have periods in excess of the 27 day observing baseline. We present a new tool, `Astraea`, for predicting long rotation periods from short-duration light curves combined with stellar parameters from Gaia DR2. Using `Astraea`, we can predict the rotation periods from Kepler 4 yr light curves with 13% uncertainty overall (and a 9% uncertainty for periods >30 days). By training on 27 day Kepler light-curve segments, `Astraea` can predict rotation periods up to 150 days with 9% uncertainty (5% for periods >30 days). After training this tool on these 27 day Kepler light-curve segments, we applied `Astraea` to real TESS data. For the 195 stars that were observed by both Kepler and TESS, we were able to predict the rotation periods with 55% uncertainty despite the wild differences in systematics.

Unified Astronomy Thesaurus concepts: Stellar rotation (1629); Main sequence stars (1000); Random Forests (1935)

1. Introduction

The rotation period of a star is one of the most direct observables one can measure. It is closely linked with its physical parameters such as magnetic activity, surface gravity, and even stellar age (e.g., Skumanich 1972; Barnes 2007; van Saders & Pinsonneault 2013; McQuillan et al. 2014; Davenport et al. 2019). Rotation periods can be used to age-date stars via "gyrochronology" (e.g., Barnes 2003, 2007), study the internal structures of stars, learn about stellar magnetic fields, and improve the precision of exoplanet detection.

In the field of exoplanet detection, additional astrophysical signals tied to stellar rotation can often complicate the process. For example, the effects of stellar magnetism in rotating stars can negatively affect exoplanet detection or characterization using radial velocity (RV) measurements. Dark spots and bright plages on the surface of a rotating star can alter the profiles of spectral absorption lines and introduce signals into RV time series. These effects are normally weak and can be treated as background noise in pipelines for discovering exoplanets. However, in the case of a planet orbiting an active star, the RV signal from the planet can be embedded within that from the host star, thus making planet signal extraction difficult (e.g., Haywood et al. 2014; Hillenbrand et al. 2015; Rajpaul et al. 2015). Modeling both the stellar activity from the host star and the orbital parameters of the planet simultaneously is essential in these scenarios. Furthermore, knowing the rotation period of the star can assist in improving the model (e.g., Grunblatt et al. 2015).

M dwarfs are also the most suitable host stars for finding rocky planets in the habitable zone as these stars are small (so the transit signal is larger) and dim (so the habitable zone is closer). This means the transit and radial velocity signals from small planets orbiting an M dwarf are stronger compared to those orbiting more massive, large host stars. However, the rotation periods of M dwarfs are often longer than the typical observing window of TESS (27.4 days), so nonstandard methods must be used to measure their rotation periods.

The most common tools used to measure rotation periods are Lomb–Scargle periodograms (e.g., Reinhold & Gizon 2015), auto-correlation functions (ACFs) (e.g., McQuillan et al. 2014), and Gaussian processes (e.g., Foreman-Mackey et al. 2017; Angus et al. 2018). These methods typically require the observed light curve to contain continuous data for more than one rotation period of the star in order to get an accurate estimate. Long rotation periods can be measured precisely for stars observed by Kepler (Borucki et al. 2010) that show periodic signals. However, long rotation periods for stars observed by TESS, especially those with only 27 days of observations per year (fall in this category; Ricker et al. 2015), are extremely hard to measure directly. Even more challenging, low-mass stars (e.g., M dwarf stars) usually have long rotation periods (>25–30 days; McQuillan et al. 2014). Because of this, traditional methods will not be able to provide accurate or precise rotation period measurements for most M dwarfs using TESS single-sector light curves.

As we know from empirical gyrochronology studies (e.g., Barnes 2003, 2007; van Saders et al. 2016), the rotation period of a star is mainly determined by its age and color. Therefore, if it were possible to measure the ages of stars precisely, we could accurately predict their rotation periods. However, the relation between stellar rotation, age, and color could break down at a

high Rossby number (rotation period divided by the local convective turnover time). van Saders et al. (2016) pointed out that magnetic weakening may cause stalling in stellar spin down for Rossby numbers greater than ∼2 and will cause gyrochronology relations to break down approximately halfway through the stars' main-sequence lifetime. This effect means we may not be able to predict rotation periods for stars that have already gone through half of their main-sequence lifetime. Fortunately, this effect is not significant for the catalogs we used in this study from McQuillan et al. (2014), Santos et al. (2019), and García et al. (2014).

However, the ages of stars, especially low-mass dwarfs, are extremely difficult to measure (see, e.g., Soderblom 2010 for a review of stellar ages). Fortunately, there are many easily observable, indirect age proxies that can be used in lieu of directly measured ages (the relations are very complex and thus ages are very hard to predict for main-sequence stars). For example, stellar velocity, radius, and surface gravity are all related, albeit weakly, to stellar age. Therefore, we expect to be able to extract information about rotation periods from these stellar properties. However, because the relationship between these properties and rotation period is "weak" and potentially nonlinear, a machine-learning (ML) approach can be used to combine these properties with observables such as color, surface temperature, or mass information to accurately predict stellar rotation periods.

In addition, there are some other potential indirect age proxies we can measure:

Flicker—the brightness variation on timescales of 8 hr and less caused by convection-driven fluctuations on the stellar surface (Bastien et al. 2013). By comparing flicker with asteroseismic $\log g$, Bastien et al. (2013) concluded that $\log g$ can be estimated from flicker with ∼0.1 dex uncertainty. If we are able to measure flickers for main-sequence stars, these measurements should be able to provide information about the surface gravity, which decreases as a star ages. As a result, it is possible to predict rotation periods by combining flicker with other stellar properties. One of the advantages of using this method is that flicker occurs on very short timescales. Therefore, we can extract granulation signals from light curves that are only 27 days long. However, flicker can be hard to measure in the light curves of M dwarfs due to the granulation signal being weak.

Flaring activities—both a star's flare energy and frequency of young, active stars are associated with their ages and rotation periods (Davenport et al. 2019). Because low-mass stars have deeper convective envelopes that are associated with stronger magnetic fields, flares are more commonly detected in these stars (Ilin et al. 2019). Therefore, flare rates could potentially be an indicator of the rotation periods of M dwarfs. However, one major limitation is that for inactive stars, which are typically older and have longer rotation periods, the rate of flaring is often too low to be detected within the short 27 day light curves of TESS.

Stellar kinematics—The kinematic properties of a star is shown to be related to the age of a star. For example, the vertical velocity dispersion of stars increase over time at a rate that can be quantified with an age–velocity dispersion relation (AVR). Strömberg (1946) and Spitzer & Schwarzschild (1951) first discovered older stars have higher vertical velocity dispersion, and this relation has been confirmed by further observations (e.g., Nordström et al. 2004; Holmberg et al. 2007, 2009; Aumer & Binney 2009; Yu & Liu 2018; Ting & Rix 2019). Two possible theories can explain these observations. One such theory is that all stars formed kinematically "cool," and as the Milky Way evolved, older stars were scattered to higher galactic latitudes by the giant molecular clouds and spiral arms. Therefore, these older stars have a higher velocity dispersion (e.g., Barbanis & Woltjer 1967; Lacey 1984; Sellwood & Carlberg 1984; Sellwood 2014). Another theory is that these older stars were born kinematically "hot" in the first place (e.g., Bird et al. 2013). Because the age and rotation period of a main-sequence star are correlated, the velocity dispersion, or other kinematic information (e.g., vertical velocity, galactic latitude, etc.), could also be useful in determining the rotation periods of stars.

Although there are many stellar properties closely tied to the rotation period of a star, it is hard to model the relations between stellar rotation and other physical properties. Low-order polynomial fits are often used to interpolate these relationships, but it is clear that the correlations are not simple. ML algorithms are particularly good at modeling complex, nonlinear relations. An ML model is normally trained on a large training data set for it to learn the complex relations between features and labels in the data. In this project, the features are the stellar properties (e.g., surface temperature, radius, color, etc.) and the label is the rotation period. After being trained, the ML model will be able to predict labels from features at a very fast speed. In addition, the same ML models can be adapted to different missions fairly straightforwardly by using the right training data. As a result, ML algorithms are likely to become more popular as astronomers march into the big data era. In particular, current and future missions observing stars, such as Kepler (Borucki et al. 2010), Gaia (Gaia Collaboration et al. 2016, 2018), TESS (Ricker et al. 2015), the Vera C. Rubin Observatory (LSST Science Collaboration et al. 2009), and the Planetary Transits and Oscillations of stars (PLATO; Rauer et al. 2014) will require rapid data-processing algorithms to accommodate the large data flow. It is essential to analyze data quickly and efficiently in order to maximize the information usage of these missions. Another benefit of using data-driven ML algorithms is that we can get insight on the data set itself. For example, a trained ML model can identify interesting anomalies or outliers in the data. We will describe briefly how the ML model we trained could potentially be used as a binary identifier in Section 5.

We use the particularly well-studied ML approach of random forest (RF; Breiman 2001) to predict the rotation periods for stars in the TESS 27 day observing fields, based on their stellar properties (Table 1 shows the list of properties used to predict rotation periods). RF is an ML algorithm that combines multiple decision trees to prevent overfitting and a suitable algorithm to learn complex nonlinear relations between different stellar properties. Decision trees use multiple parameters (e.g., effective temperature, radius, luminosity, etc.), which are often called "features," to split the data into different subsets (where the data split is called a "node.") and predict the "label" (e.g., rotation period). An RF algorithm trains a number of decision trees on different subsets of the data and predicts the label by averaging the resulting predictions from each decision tree. This ML approach has huge potential to automate the delivery of rotation period from observation data. RF, compared to neural networks or deep learning, is relatively easier to interpret as the input features are selected by the user and the user can calculate the feature importance and gain insight into the data itself. This method can be used to capture

Table 1
Final Training Features Used in This Project Sorted into Four Categories

Feature Name	Description	Categories
bp_g (c, r)	Integrated BP mean magnitude—*G*-band mean magnitude.	Direct Gaia observations (gaia-kepler.fun).
phot_g_mean_flux_over_error (c, r)	Mean flux in the *G* band divided by its error.	
parallax (c)	Parallax.	
T_{eff} (r)	Estimate of effective temperature from Apsis-Priam (Andrae et al. 2018).	Stellar properties derived from Gaia observations (gaia-kepler.fun).
lum_val (r)	Estimate of luminosity from Apsis-FLAME (Andrae et al. 2018).	
radius_val (r)	Estimate of radius from Apsis-FLAME (Andrae et al. 2018).	
r_lo/r_hi (c)	68% confidence interval on distances from Bailer-Jones et al. (2018).	
v_tan (r)	Velocity tangential to the celestial sphere ($\sqrt{v_{ra}^2 + v_{decl.}^2}$).	Kinematic derived from Gaia proper motion, R.A., decl. and parallax using astropy.
v_b (r)	Velocity in the direction of galactic latitude.	
b (r)	Galactic latitude of the object at the reference epoch (Butkevich & Lindegren 2014).	
LG_peaks (c)	Maximum peak height from the Lomb–Scargle periodogram.	Light-curve statistics.
R_{var} (c, r)	Photometric variability of the light curve (95th percentile—5th percentile of the normalized flux).	
flicker (r)	Brightness variation on timescales of 8 hr and less calculated with the FLICKER software.	

Note. Other than the radius value itself, we also included the 68% confidence interval of distances (r_hi/r_low) from Bailer-Jones et al. (2018) as training features. "c" and "r" represent whether the feature was used in training the classifier and regressor, respectively.

and effectively model the relationships between stellar rotation, stellar age, and stellar parameters including temperature, radius, and surface gravity. RFs are already used in astronomy, both in classification and regression problems. For example, Richards et al. (2011) classified variable stars with sparse and noisy time-series data with a ∼20% error, and Miller et al. (2015) inferred fundamental stellar parameters for ∼54,000 known variables with an RF regressor.

In this paper, we exploit the relationships between rotation periods and other fundamental stellar parameters, which occur as a result of stellar evolution. We predict rotation periods without requiring long time-series observations using an RF algorithm. The features we used to train the model and their origins are described in Section 2. We first classify stars to determine whether their rotation periods are measurable and then use an RF regressor to predict the rotation periods of those classified as "measurable." The details of how we trained and optimized the models are described in Section 3; the testing results for Kepler and TESS stars are described in Section 4. Limitations and reasons why we are able to predict long rotation periods from short-duration light curves are discussed in Section 5.

2. Data and Methods

In order to train and test an ML model, we need both a training data set (Section 2.1) and a testing data set (Section 2.2). The purpose of a training data set is to train the model to learn the complex relations between a number of "features" (stellar properties; Section 2.3) and the "label" (rotation period). The purpose of a testing set is to have a number of stars that are not from the training set to validate the trained model.

After constructing a reasonable training and testing set, we selected the useful stellar properties that are important to predict the rotation period in Section 2.3. One of the features we focused on is the variability of the light curve, which normally is the flux variation (range or standard deviation) averaged over one or multiple rotation period(s). Because we do not have information on the rotation periods, we will discuss how we can use the flux variation over the entire observing period to approximate the variability of the light curve (see Section 2.3).

The training process for the RF classifier and the RF regressor is described in Section 2.4. A classifier is used to identify a group(s) of data that are similar. A regressor is used to model the complex relationships between features and labels in order to predict new labels from new features (the simplest regressor is a linear regressor).

By combining a classifier and a regressor, we are able to classify whether or not a star has a "measurable" stellar rotation period and predict its period if the period is "measurable." Figure 1 shows the pipeline of `Astraea`,[4] the RF package (classifier + regressor) used to predict the rotation period from stellar parameters. Details of how it is built will be described in this section.

2.1. Kepler Training Set

We selected our training data from the Kepler field because there already exist catalogs for the rotation periods of these stars and long rotation periods measured from 4 yr Kepler light curves are more reliable. The majority of the rotation periods we used to train our models were from McQuillan et al. (2014). They analyzed 133,030 main-sequence Kepler targets and measured rotation periods (between 0.2 and 70 days) for 34,030 stars by using an automated ACF-based method. The ACF-based method has its advantages over a Fourier-based method or Lomb–Scargle periodogram because the rotation

[4] Available at https://astraea.readthedocs.io/en/latest/.

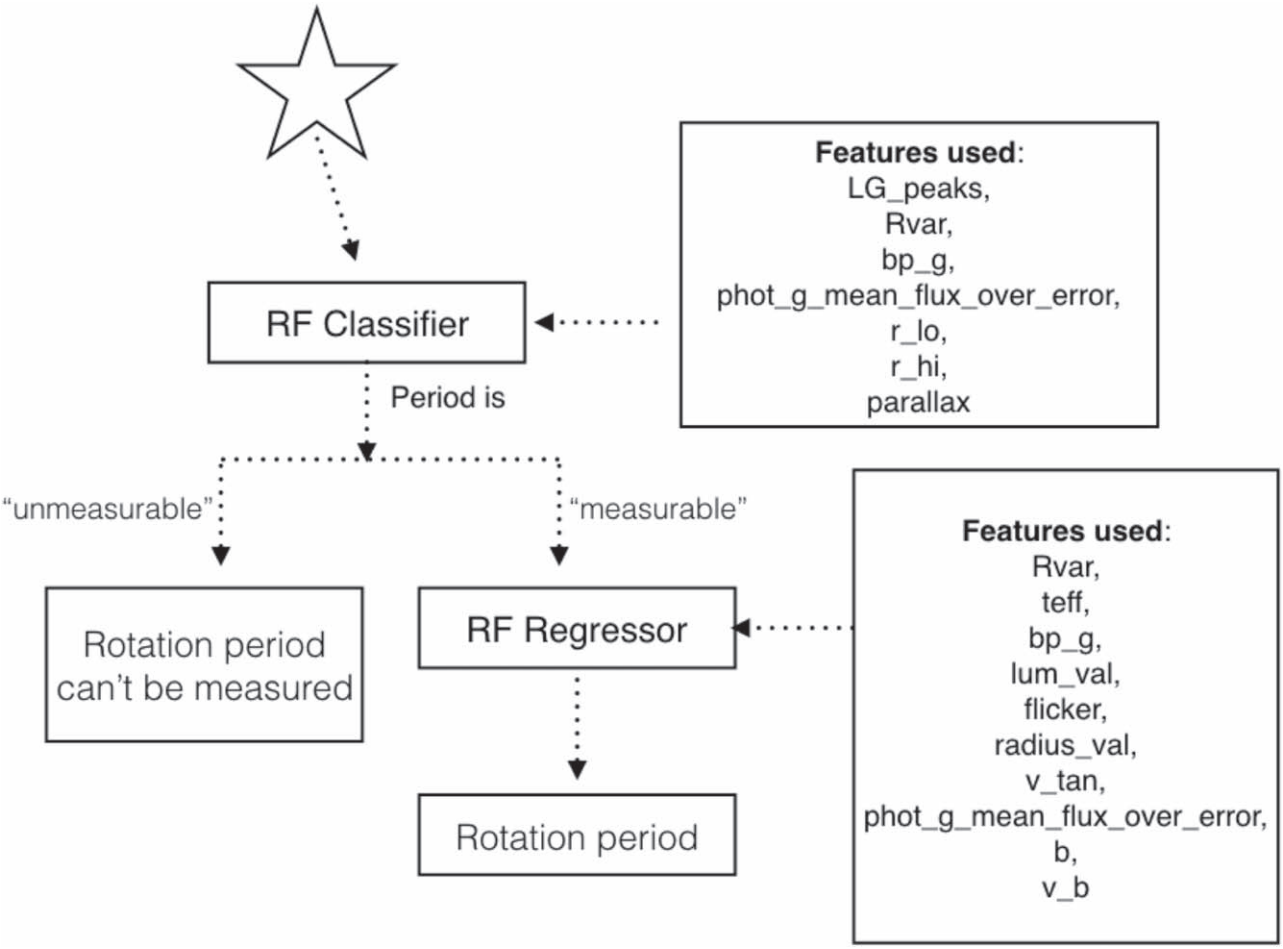

Figure 1. Pipeline of `Astraea`, our open-source software developed in this project. Features (stellar properties and light-curve statistics) of the star are first passed through the RF classifier to identify whether or not the period is "measurable." If the period is "measurable," the features will then be passed through the RF regressor to predict its stellar rotation. The feature descriptions are provided in Table 1.

period signals in the light curves are neither purely sinusoidal nor strictly periodic.

We utilized all the 133,030 main-sequence stars analyzed in McQuillan et al. (2014) to train a model to determine whether the rotation period for a star can be obtained. Because our main goal was to predict long rotation periods from short-duration light curves, in addition to the 34,030 stars with rotation period measurements from McQuillan et al. (2014), we also added 4637 stars that have rotation periods up to ∼150 days from Santos et al. (2019) and García et al. (2014), in which they used a combination of wavelet analysis and the ACF to measure the periods. Within these added rotation periods, 70 of them have rotation period >70 days. This provided us with ∼38,000 Kepler stars. Figure 2 shows a histogram of the rotation periods in our training set.

We split the data into the training data set and a validation data set so we can train our model on the training set and tune our trained model on the validation set. The difference between a validation data set and a testing set is subtle, but the validation data are typically used to tune the hyperparameters (parameters relate to the ML model; see Section 3) and the testing set is used to test the optimized model. The validation and testing sets are both important because although the validation set can be used to optimize the model in order to make sure the ML model is not overfitting the validation data, a testing set is needed to test the final optimized model. The training set is composed of 80% random selection of stars from McQuillan et al. (2014) and the 4637 stars from García et al. (2014) and Santos et al. (2019). The validation set is the remaining 20% stars.

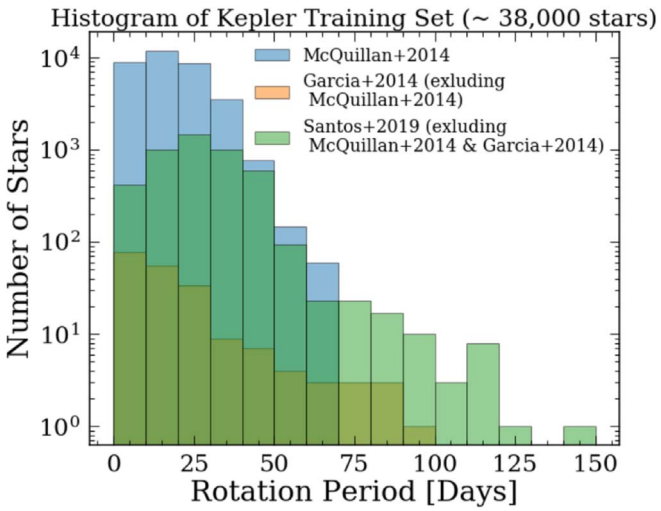

Figure 2. Histogram of the rotation periods in the Kepler training set. 34,030 stars are from García et al. (2014) and 4637 stars from Santos et al. (2019) and García et al. (2014). Supplementing the main set with these later catalogs increased the number of long-period rotators, including 70 stars with periods longer than >70 days. Note that the *y*-axis is in log scale.

2.2. TESS Test Set

After cross-matching with the TESS light-curve database hosted by the Mikulski Achieve Space Telescopes (MAST),[5] we were able to find 205 Kepler targets with TESS 2 minute cadence PDCSAP light curves. We excluded 10 stars from the equal-mass binary sequence by performing a magnitude cut on the color–magnitude diagram (CMD). A star in a unresolved close binary system with an equal-mass companion will not affect its color but double its luminosity due to the starlight from its equal-mass companion. As a result, these stars will lie on the CMD above the main-sequence stars and form a "binary sequence." To exclude these stars, we first fitted a sixth-order polynomial to the McQuillan data sample in CMD and shifted the function up by ∼0.27 dex in absolute G magnitude and excluded any stars lying above the shifted polynomial function. After the cut, we were able to obtain a testing set of 195 stars.

2.3. Feature Selection

Measuring variability—The brightness variation due to magnetic activity on the surface of a star has been shown to correlate with stellar activity and therefore should be related to the rotation period (e.g., Pizzolato et al. 2003; Hartman et al. 2011; McQuillan et al. 2014; Davies et al. 2015; Santos et al. 2019). However, brightness variation from a light curve includes more than the magnetic activity from the surface. Granulation, instrumental noise, p-mode oscillations, etc. can also modulate a light curve. As a result, ideally, we would measure the light-curve variability by taking into account the stellar rotation period. Two popular measurements are the average amplitude of variability within one period and the standard deviation of a subseries of length five times the rotation period, and these can be parameterized by $R_{\rm per}$ or $S_{\rm ph}$, respectively. These two variables take into account the rotation period of a star and are shown to be closely correlated with the magnetic activity and rotation period of a star. However, in order to measure these quantities accurately, the stars would have had gone through more than one full revolution in the observation window. For stars observed by TESS, most slow rotators have not gone through even one full revolution within the 27 day observing period. Therefore, it is almost impossible to get accurate measurements for either quantity, especially for the slow rotators.

Fortunately, $R_{\rm var}$ (95th percentile—5th percentile of the normalized flux) is a good estimator for $R_{\rm per}$ and $S_{\rm ph}$, and its measurement does not require information on the stellar rotation. $R_{\rm per}$ is calculated by computing the 5th–95th percentile range of the flux of each full stellar revolution and then taking the average of these quantities. On the other hand, $R_{\rm var}$ is the 5th–95th percentile flux range of the entire light curve. $R_{\rm var}$ and $R_{\rm per}$ are therefore most similar when the stellar rotation period is long, because fewer full revolutions take place. Stars with long periods usually have smaller variability amplitudes and therefore smaller $R_{\rm var}$ and $R_{\rm per}$ values (e.g., see Figure 3). This is why the two quantities are similar at small values. $R_{\rm var}$ is more sensitive to long-term light-curve systematics than $R_{\rm per}$, and this is particularly true for rapid rotators where $R_{\rm per}$ is calculated over many short time intervals and averaged. This is also why $R_{\rm var}$ is slightly larger than $R_{\rm per}$ at large values (i.e., for rapid rotators): long-term light-curve systematics slightly increase the variance in the light curve, which inflates $R_{\rm var}$ relative to $R_{\rm per}$. This could potentially mean we will be able to predict long rotation periods better than short rotation periods because $R_{\rm var}$ is a better proxy for $R_{\rm per}$ and $S_{\rm ph}$ for slow rotators.

Features used—The features used to train/test the models are (i) three measurements directly from the light curves, (ii) all the Gaia columns (including error columns), in which nine were later found useful in predicting stellar rotations, and (iii) two kinematic statistics derived from Gaia parameters.

To obtain Gaia parameters of our sample of 133,030 Kepler stars, we used the publicly available Kepler–Gaia DR2 cross-matched catalog.[6] The majority of stellar features used for rotation period prediction were obtained from the Gaia DR2 catalog, and the distance measurements were obtained from Bailer-Jones et al. (2018).

In addition to the features from Gaia, we also calculated three variables directly from the light curves and two additional kinematic features. These features have been shown to be related to the rotation period of a star (details described later in this section).

The features measured from the light curves are: (i) $R_{\rm var}$, the range of variability in the light curve, which was calculated as the difference between the flux values at the 95th percentile and the 5th percentile, (ii) flicker, brightness variation on timescales of 8 hr and less, calculated with `FLICKER`, our new open-source software we developed to calculate flicker using the method described in Bastien et al. (2013; detailed description in Section Appendix), and (iii) the Lomb–Scargle periodogram maximum peak height.

Additional kinematic features we calculated are v_tan, the velocity tangential to the celestial sphere, and v_b, the velocity in the direction of the galactic latitude from Gaia R.A. and decl. coordinates, proper motion, and parallax.

Selecting training features—To start with, our full set of features consisted of every column in the Gaia DR2 catalog, plus the three light-curve statistics and the two velocities described above, making a total of 148 features. However, we did not expect that every feature in the Gaia DR2 data set would be useful. For example, it seems unlikely that the R.A. and decl. would be strongly related to stellar rotation period. Thus, we performed feature selection (selecting the important features to speed up the training process and potentially increase model performance) for both the classifier and the regressor using the method described in the next paragraph to isolate features that provide significant information about rotation periods.

We selected these features by first training the RF models on all columns from Gaia, the kinematic features and the light-curve statistics calculated from 4 yr stitched Kepler light curves. We then calculating the "*gini*" feature importance (Breiman et al. 1984) using `scikit-learn` (Pedregosa et al. 2011). This importance was determined by calculating the mean decrease in impurity (MDI), which indicates whether a single feature alone can predict the outcome. For example, if one can predict the rotation period of a star just by the effective temperature, then the node, where the data split (refer to Section 1 for how RF works), is considered pure because the model will only split the data into different subsets based on the effective temperature. On the other hand, if the rotation period is also related to other parameters (the data are split based on

[5] https://archive.stsci.edu/access-mast-data

[6] Available at gaia-kepler.fun.

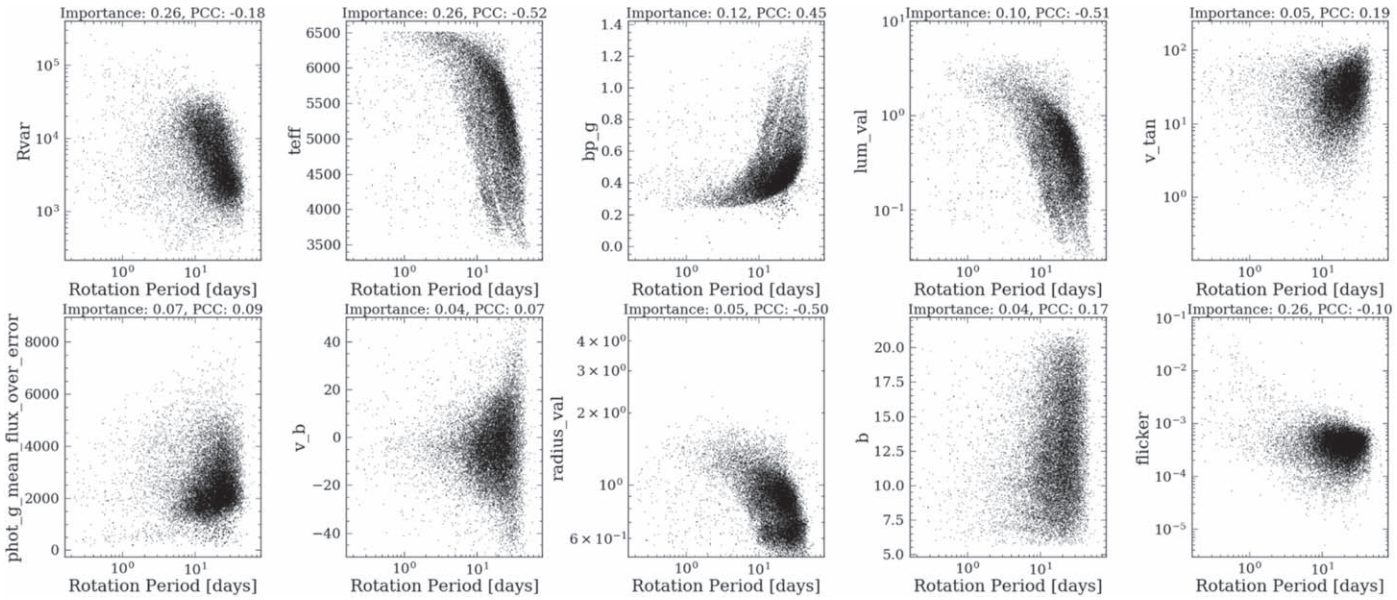

Figure 3. Relationships between features used to train the RF regressor and rotation periods for all 34,030 stars in McQuillan et al. (2014), Santos et al. (2019), and García et al. (2014). It is clear that these stellar quantities are related to the rotation period of a star, but the correlations are complex and often cannot be described with simple low-order polynomials. The *y*-axis labels are described in Table 1. The title of each plot shows the *gini* importance (0 to ∼1, where 0 is not important at all) and the Pearson correlation coefficient (PCC) for each of the features in the final training phase with the optimized hyperparameters, which we will describe in the later sections.

more than the effective temperature), then there is an impurity in the node. The *gini* importance is normalized over all features and ranges from 0 to 1. A *gini* importance of 1 for a feature means the prediction of rotation period can be determined solely by this one feature. Typically, feature values with wildly different ranges need to be normalized to a common scale in order to ensure the feature importance does not appear to be higher/lower than they should because of their range. However, the RF algorithm does not require feature normalization as it splits the data based on the feature values and the splitting is independent of the feature range. Calculating this importance is a good way to eliminate irrelevant features—features that do not contribute significantly to the prediction of stellar rotation (*gini* importance of 0). We sorted the features by decreasing the *gini* importance and performed cross-validation tests using RF regression with an increasing number of features and selected the smallest number of features that led to a converged accuracy (for classifier) or χ^2 value (for regressor). The accuracy/χ^2 value converges when the change is smaller than 5%. The accuracy is a way to estimate the performance of a classifier and the χ^2 value is a way to estimate the deviation between the target rotation period and the predicted rotation period for a regressor. Cross-validation tests are often used to maximize the performance of the model. We trained the RF model on the training set, and by maximizing the model performance on the cross-validation set, we will be able to optimize the model. To perform the cross-validation tests, we randomly excluded 20% of the data in the training phase and predicted the rotation periods for these stars using the trained model. For each set of features, we performed the cross-validation test 10 times and took the median of the average χ^2 values. Figure 3 shows the relationships between these features and rotation periods for the 34,030 stars in McQuillan et al. (2014), Santos et al. (2019), and García et al. (2014) as well as their *gini* importance.

Looking at the relationships in Figure 3, the *gini* importance, and the Pearson correlation coefficient (PCC, a statistical value to measure the linear correlation between two variables), there exist strong correlations between rotation period and $R_{\rm var}$, effective temperature, Gaia color ($G_{BP} - G$, also called bp_g), luminosity, and radius. There also exist weak correlations between rotation period and the other features plotted.

$R_{\rm var}$ is known to be strongly correlated with rotation period (Pizzolato et al. 2003; Hartman et al. 2011; Walkowicz & Basri 2013; McQuillan et al. 2014; Santos et al. 2019). It is also proven that the rotation period is a strong function of effective temperature and age (this is the principle behind gyrochronology; e.g., Skumanich 1972; Kawaler 1988; Barnes 2003, 2007), and age is weakly correlated with multiple stellar parameters such as luminosity, surface gravity, and kinematics. It makes sense, therefore, that $R_{\rm var}$, effective temperature, and color would have the strongest correlation with rotation period and the other stellar parameters would have weaker relations with rotation period.

There exist both strong and weak correlations between rotation period and a number of other stellar parameters. These relationships are difficult to reproduce using physical or simple empirical models. However, an ML algorithm like RF is effective at predicting properties from a large number of weakly correlated features, and this is why it is so well suited to predicting rotation periods from other stellar parameters.

2.4. Random Forest Classification and Regression

The RF algorithm merges multiple decision trees to get a more accurate and stable prediction. This algorithm is also known to reduce overfitting, which is a common problem in single decision trees. RF can be used in both classification and regression. It also requires less computational time compared to deep learning and is able to handle outliers. However, RFs are not capable of extrapolating data so we could only in theory

predict rotation periods up to ∼150 days, which was determined by the upper limit for rotation periods in our training set. We used the Python `scikit-learn` package to train our RF classifier and regressor. The hyperparameters were set to default for the classifier for simplicity, and we explored the hyperparameters used in our regressor later on in this section.

Random forest classification—RF models are not good at extrapolating data. This means we are only able to predict rotation periods in the same parameter space as the training set, and this is the main reason we need a classifier—to determine whether a star lies in the same parameter space as the stars in the catalog from McQuillan et al. (2014). Another motivation for a classifier is that not all stars with rotation periods show detectable signals in their light curve. For example, a star could be inactive and therefore not have detectable spot modulations. It could have starspots distributed homogeneously on the surface that cancel out any variations in its light curve. We could also be viewing the star pole on and, therefore, not be able to detect any azimuthal variations. Both of these factors require us to train a classifier to first determine if it is possible to predict a reliable rotation period. The labels were created using stars in the McQuillan et al. (2014) catalog, where the 34,030 stars that have rotation periods were labeled "measurable" and the remaining 99,000 stars were labeled "unmeasurable." Because the method of McQuillan et al. (2014) was conservative, our classifier trained on this data set was also on the conservative side, i.e., it is possible that the periods of some stars with periodic brightness variations in their light curves were classified as "unmeasurable" with our classifier. So this classifier is not a perfect tool to determine whether a star has a detectable period but rather a way to classify whether the star would appear in the McQuillan et al. (2014) catalog.

Features used to train the classifier are LG_peaks, $R_{\rm var}$, bp_g, phot_g_mean_flux_over_error, r_lo, r_hi, and parallax (refer to Table 1 for a detailed description of each variable).

Random forest regression—To predict rotation periods, the regressor was used if a star's period was labeled as "measurable" by the classifier. Here we used an RF regressor because a star's rotation period is correlated with its other stellar properties, and RF regression is useful for predicting continuous values from various features. An RF regressor trains multiple independent decision trees on a different subset of the data where each tree could give a slightly different period prediction. The model then takes the average of all the predictions from all the trees and their uncertainties to determine the final predicted rotation period.

Features used to train the regressor are $R_{\rm var}$, teff, bp_g, lum_val, flicker, radius_val, v_tan, phot_g_mean_flux_over_error, b, and v_b (refer to Table 1 for a detailed description for each variable).

3. Optimizing and Assessing the Performance of the Random Forest Models

We trained both the classifier and the regressor on 80% of the data and used the remaining 20% to perform cross-validation tests, which is a good way to prevent overfitting. The features used for each model and their permutation feature importance are shown in Figure 5.

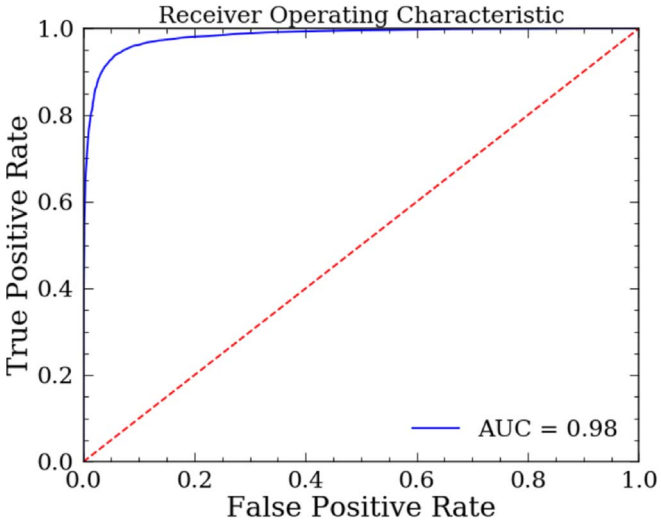

Figure 4. The receiver operating characteristic curve for the classifier. The area under the curve (AUC) shows how well the model can distinguish between classes. It is often used to estimate the accuracy or performance of the classifier. The model reached maximum accuracy of 98% with a threshold of 0.4.

3.1. Random Forest Classifier

The outputs of the classifier were numbers from 0 to 1 for each star, where 0 means the period was 100% "unmeasurable" and 1 means it was 100% "measurable." One can simply say if the predicted number was greater than 0.5 (which means the threshold was 0.5), then the period was "measurable." However, the best way to determine the threshold is to maximize the area underneath a receiver operating characteristic curve (ROC) as shown in Figure 4. An ROC curve shows the predicted false-positive rate (FPR) against the true-positive rate (TPR) for various thresholds. The FPR is the total number of false-positive cases (e.g., the number of stars where their rotation periods can be measured and are predicted "unmeasurable") divided by the total number of negative cases (e.g., the number of stars where their rotation periods are predicted "unmeasurable"). The TPR is the total number of true-positive cases (e.g., the number of stars where their rotation periods can be measured and are predicted "measurable") divided by the total number of positive cases (e.g., the number of stars where their rotation periods are predicted "measurable"). These statistics are useful especially in cases where the training data set is overflowed by one label (positive or negative). For example, if 98% of the stars in the training set has a "measurable" rotation period, then an incorrect model that predicts every star has a "measurable" rotation period will reach an accuracy of 98%. However, this model is clearly wrong, and by calculating the TPR and FPR, one can get a better understanding of the true accuracy of the model. A perfect model would have a false-positive rate of 0 and a true-positive rate of 1, and the curve would go straight up the TPR axis until it reaches 1 then go horizontal on the FPR axis. Thus, the closer the ROC curve approaches (0, 1), the more accurate the classifier is. We determined the accuracy and threshold by finding the point along the curve where TPR-FPR was maximized. This yielded a 98% accuracy with a threshold of 0.4.

3.2. Random Forest Regressor

Hyperparameter optimization. To achieve the best performance of the model, we optimized the hyperparameters

Table 2
Optimal Hyperparameters in the RF Regressor Model that Minimized the Average χ^2 or rMAD

Hyperparameter Name	Description	Grid-search Range	Value that Minimizes the Average χ^2	Value that Minimizes the Average rMAD
n_estimator	No. of decision trees used in the RF model	1–100	20	1
max_depth	Maximum depth of the tree	1–150	50	100
max_features	No. of features to consider when looking for the best split	1–10	6	3

Note. By minimizing the average χ^2, we had low variance but high bias in the model, and by minimizing the MAD, high variance but low bias was achieved.

(parameters describing the model) using a grid-search method. The hyperparameters we considered and their optimal values are shown in Table 2. For each set of hyperparameters, we performed a Monte Carlo cross-validation test 10 times with 20% of the data, chosen randomly each time, left out during the training process. For each of these tests, we calculated the average $\chi^2 = \frac{\Sigma_i^N (y_i - y_{\rm predict})^2 / \sigma_{y_i}^2}{N}$ and the relative median absolute deviation (rMAD) $= \text{median}\left(\frac{|\ y_i - y_{i\rm predict}\ |}{y_i}\right)$, where ($i = 1, 2, 3, ..., N$), y_i is the expected rotation period value, $y_{predict}$ is the predicted rotation period value, and N is the number of data points. The overall average χ^2 and rMAD of the model for each set of hyperparameters were then represented by the median values of these 10 tests.

Two sets of optimal hyperparameters were obtained by minimizing the average χ^2 or minimizing the rMAD. Minimizing the χ^2 reduced the spread of the data (variance) and minimizing the rMAD reduced the systematic bias in the data (bias). In order to get a more precise result, we selected the model that minimized the average χ^2.

Permutation feature importance—We calculated the permutation feature importance to study how each feature impacts the prediction results using the optimized model. By calculating the permutation feature importance, we are able to interpret the model and potentially gain insight into how the stellar properties are related to the rotation period. The permutation feature importance can be calculated by randomly shuffling values within a single feature and observing how the model performance changes. This is effectively removing each feature from the model and preventing it from being informative, and measures how good the model can still predict the data. This importance is a more accurate measurement of how much of a role each feature plays in determining the outcome, compared to the *gini* importance. We used the R^2 (coefficient of determination) regression score to measure the model performance, $R^2 = \Sigma_i (y_{\rm predict} - \bar{y})^2 / \Sigma_i (y_i - \bar{y})^2$, where $y_{\rm predict}$ is the predicted rotation period and $\bar{y}$ is the average rotation period. It provides a measure of how close the data are to the fitted regression function. The R^2 score is commonly between 0 and 1, and the higher the score, the better the fit. To obtain the importance for each feature, we calculated the R^2 score on the training set and reshuffled the values within one feature and kept the rest of the training data set unchanged. We then passed the new training data set to the model again to calculate a new score based on this modified training set. The feature importance is the difference between these two scores, normalized to sum to one across features. Figure 5 shows the permutation feature importance for both the RF classifier (using 4 yr Kepler light curves) and the RF regressor (separately calculated for 4 yr Kepler light curves and 27 day Kepler light-curve segments).

The power of the highest peak in the Lomb–Scargle periodogram of each star's light curve (LG_peaks) was the most important feature for the classifier. Because the classifier was trained on targets from McQuillan et al. (2014), the RF classifier learned the algorithm they used to determine whether the light-curve signal was periodic. McQuillan et al. (2014) determined whether a rotation period was reliable (or whether a star has a rotation period signal that can be detected) by setting a threshold for the maximum peak height from the ACF, which is similar to the maximum peak height from the Lomb–Scargle periodogram. As a result, it makes sense that LG_peaks is the most important feature in determining whether a star can be included in the catalog from McQuillan et al. (2014).

The confidence interval of the distances also determines whether or not a star's period is measurable. One potential reason is that a larger distance error (or any error from luminosity, temperature, etc.) is also associated with a larger error in the observables (photometry and parallax, etc.), suggesting a fainter or/and more distant star whose period would normally be harder to determine. Because errors from stellar properties are correlated (Andrae et al. 2018), the RF classifier would only use one of these errors as an important feature (similar to determining the rotation period, the RF regressor treated the effective temperature as one of most important features but not the color, though they are very similar).

Other features, such as $R_{\rm var}$, bp_g, g_over_error, and distances, not only determine whether or not we can recover the rotation period but also contain information about the rotation period itself. Because a shorter rotation period is easier to recover, it is not surprising that these attributes appear to be important in both classification and regression models.

The importance of the regressor was more evenly distributed over multiple features. This implies that the rotation period is closely related to multiple stellar properties, and precise rotation periods can only be predicted using multiple features. $T_{\rm eff}$ and $R_{\rm var}$ are known to be strongly correlated with rotation periods (e.g., Santos et al. 2019), and the kinematics of a star, as mentioned in Section 1, could also be used to constrain its age and, therefore, its period.

The importance trend for the model trained on 27 day Kepler light curves closely follows the model trained on full 4 yr Kepler light curves. However, the flicker feature is more important for 4 yr light curves compared to that of 27 day light curves. This suggests the flicker value encodes more information from the rotation period as we average over longer time spans or that the flicker measurement becomes more precise,

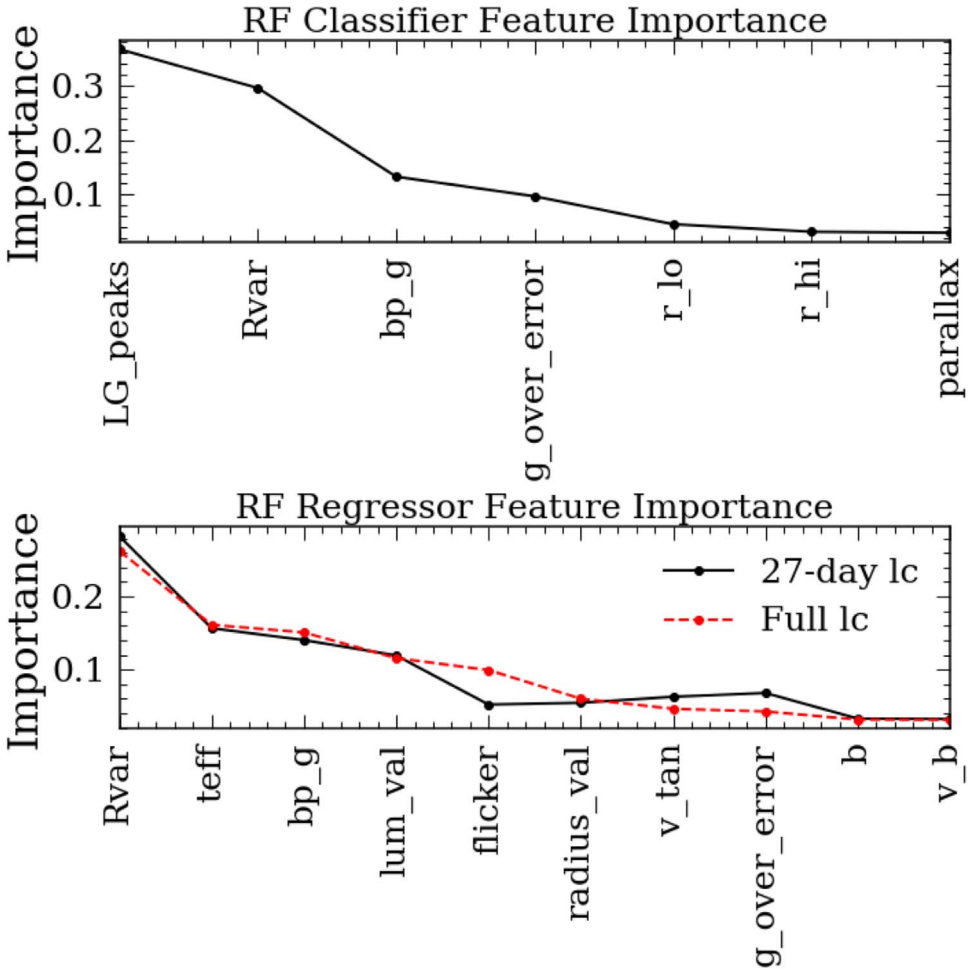

Figure 5. Permutation feature importance on the Kepler cross-validation set (∼7000 stars), where *g_over_error* is the *G*-band mean flux divided by its error. The two lines in the regressor represented the feature importance for training on 27 day light curves (solid black line) and that on full 4 yr Kepler light curves (dashed red line). $R_{\rm var}$ and flicker are measurements for Kepler light curves with their corresponding timescale (measured from the 27 day light curves for the solid black line and from 4 yr light curves for the dashed red line).

and therefore more discerning of log g, as more quarters are incorporated.

4. Results

In this section, we present the performance of our optimized model (M_χ^2) on Kepler data with 4 yr light curves, on simulated TESS data calculated by splitting full Kepler light curves into 27 day sections, and on real TESS data.

4.1. Performance on Kepler Data

Figure 6 shows the testing result on full 4 yr Kepler light curves with points colored by their effective temperature. In general, cooler stars spin more slowly because they have deeper convection zones, which mean they have stronger magnetic fields and therefore spin down faster due to magnetic braking compared to hotter stars. We picked the model with the lowest χ^2, which also minimized the scatter (variance). However, low-variance models normally have high systematic bias. It is clear from the residual shown in the bottom panel that we systematically overpredicted the short rotation periods and underpredicted the long rotation periods. We estimated the uncertainty by calculating 1.5*MAD from the residuals, and we can recover the rotation periods with an uncertainty of 13% and long rotation periods (>30 days) with an uncertainty of 9%.

4.2. Performance on 27 Day Kepler Light Curves

Testing our model on Kepler 4 yr light curves gave us promising results. However, our main goal for this model is to predict rotation periods from 27 day TESS light curves. To do this, we split each 4 yr light curve from the Kepler training set into multiple 27 day segments and calculated $R_{\rm var}$ and flicker for these short-duration light curves. Other features from Gaia remained the same for each target. Breaking up the light curves from the Kepler training set and treating each 27 day light curve as a separate star effectively expanded our number of training targets to over 1.8 million (∼34,000 4 yr light curves from Kepler, with each of these light curves splitting into ∼54 27 day light curves).

Comparing 4 yr and 27 day light curves—Figure 7 shows a comparison between $R_{\rm var}$ and flicker values from 4 yr light curves with those of 27 day light curves. We quantified the differences between these two statistics for the 4 yr light curves and 27 day light curves by calculating 1.5*MAD (a measure of the standard deviation that is robust to outliers) of the residuals. This yielded a standard deviation of 30% and 35% for flicker and $R_{\rm var}$, respectively. The scatter in these two light-curve

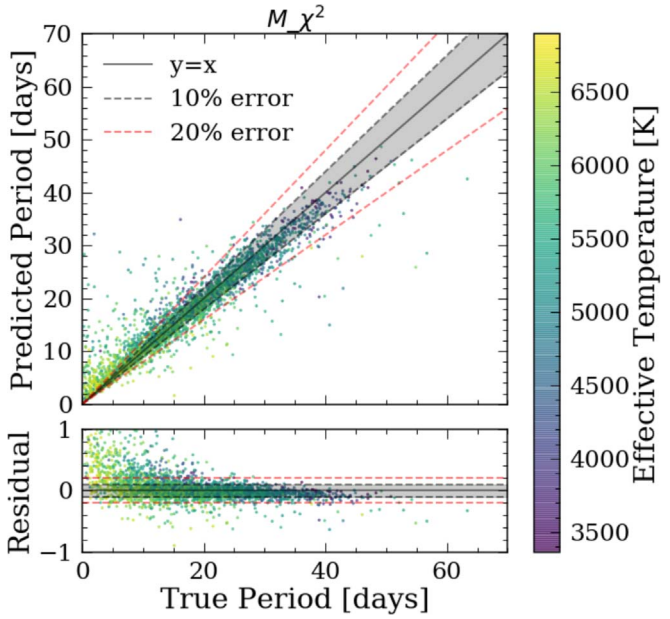

Figure 6. Periods predicted by the RF regressor vs. the rotation periods measured in McQuillan et al. (2014) for ∼3,000 stars colored by their effective temperature, where hotter stars tend to rotate faster, as expected. The gray area occupies the 10% error space. The top panel shows the predicted rotation period vs. the true rotation period from McQuillan et al. (2014), and the bottom panel shows the residual. We are able to predict rotation periods with an uncertainty of 13% and long rotation periods (>30 days) with an uncertainty of 9%.

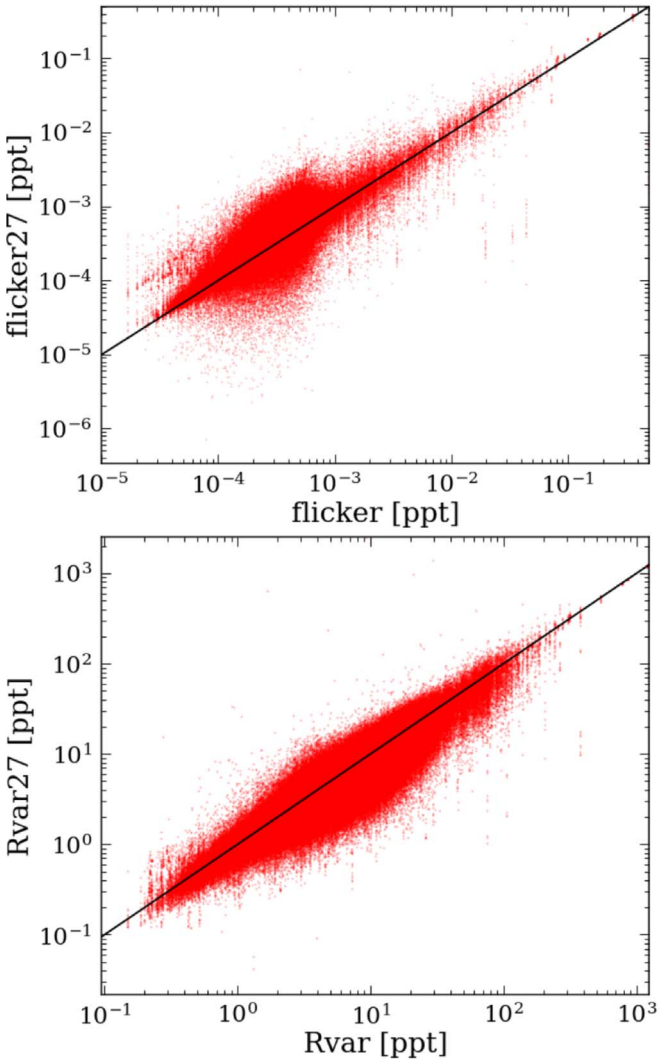

Figure 7. Flicker and $R_{\rm var}$ values for 4 yr light curves compared to those of 27 day light curves for ∼100,000 Kepler stars (∼5 million 27 day light-curve statistics from splitting up 4 yr light curves). The solid lines are the identity functions. The relative MAD between 27 day light curves and 4 yr light curves are 30% and 35% for flicker and $R_{\rm var}$ respectively.

statistics constrains how well we can predict rotation periods and is discussed in the later paragraphs.

After excluding the 195 stars observed by both Kepler and TESS, which we later tested our model on, we trained the model on ∼80% of these 1.8 million 27 day light curves and tried to recover the remaining 20%. Figure 8 shows the results for ∼20% of the targets (∼300,000) in the McQuillan et al. (2014) catalog. We did not optimize the hyperparameters again because both training sets are from Kepler, and we assumed the light-curve statistics we calculated will be similar so the optimized hyperparameters will also be similar. The general trend follows that shown in Figure 6. But with more training data (because we broke the Kepler 4 yr light curves into multiple 27 day light curves), most of the predictions have uncertainties on the order of 9%, and we are able to predict long rotation periods (>30 days) with an uncertainty of 5%. This is important because despite the measurements for flicker and $R_{\rm var}$ from 27 day light curves being worse, we were able to get a more precise result by increasing the number of training data by splitting the full 4 yr light curves. A fit could potentially be used to correct for the bias; however, this bias is subject to change. For example, the difference in the noise properties between TESS and Kepler could affect the systematic bias. More discussion is included in Section 5.

One additional feature worth pointing out is the vertical streaks in Figure 8. This is most likely due to the variation in $R_{\rm var}$ and flicker (Figure 7). After splitting the 4 yr light curve of each star into multiple 27 day light curves, there existed multiple training data that had the same values for every feature except $R_{\rm var}$ and flicker (because we recalculated these two values for every 27 day light curve). This would cause the model to have multiple different predictions for the same star even though this star only has one rotation period measured with traditional methods.

One concern is that we were not able to recover rotation periods for fast rotators with high precision when compared with the use of traditional methods. One potential reason could be that some very fast rotators are synchronized binaries. Synchronized binaries are binary stars whose tidal interactions have synchronized their rotation periods with their orbital periods, i.e., they are tidally locked with each other. There is mounting evidence to show that a large fraction of cool stars which rotate faster than 7–10 days are, in fact, synchronized binaries (e.g., Simonian et al. 2019; Angus et al. 2020). The rotation periods of stars in synchronized binary systems have been influenced by tides and will not be the same as (and will probably be shorter than) the rotation period expected for each star based on their temperatures, surface gravities, and ages.

Our main goal with this RF model was to predict long rotation periods with short TESS light curves, which is difficult to do using traditional methods. So, not being able to predict short rotation periods accurately is not a major concern for our

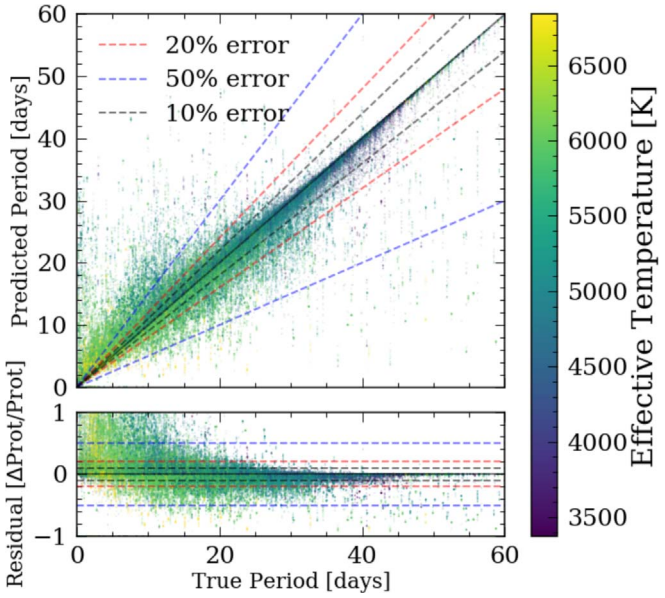

Figure 8. Results for 370,208 27 day Kepler light-curve segments, colored by Gaia effective temperature. The top panel shows the comparison between the predicted rotation periods and true rotation periods. The bottom panel shows the residual; 3% of the data that has residuals greater than 1 were cut out. There is a clear temperature gradient from fast rotators to slow rotators, where hotter stars tend to rotate faster, as expected. The period predictions have an average uncertainty of 9% for all the stars and 5% for slow rotators (>30 days).

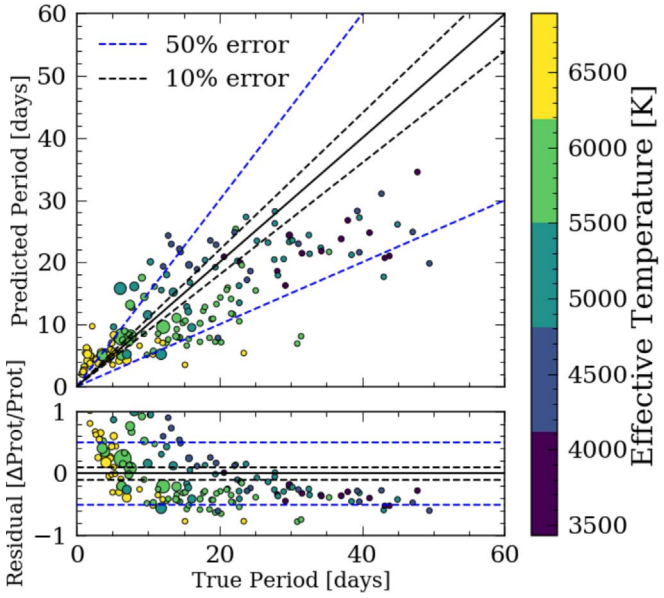

Figure 9. Testing result on 195 TESS targets in the Kepler field using the M_χ^2 model trained on 27 day Kepler light-curve segments, colored based on the effective temperature. The marker size indicates the signal-to-noise ratio (S/N). The S/N is calculated by dividing $R_{\rm var}$ by the noise floor level calculated in Figure 10. The uncertainty is 55% for all predictions as well as for slow rotators.

algorithm because we could combine both methods to measure the rotation periods of all ranges. Furthermore, we are predicting rotation periods instead of measuring them. This means that even though our results are not as accurate as periods measured with traditional methods, we can still predict rotation periods when traditional methods fail to measure them.

4.3. Performance on Real TESS Data

We downloaded 195 TESS 2 minute cadence PDCSAP light curves from MAST and calculated the Lomb–Scargle maximum peak height, flicker, and $R_{\rm var}$ from the TESS 27 day light curves. The rest of the features were acquired from Gaia. We first passed these stars through the trained classifier, and all 195 rotation periods were identified as "measurable." These targets were then fed to the trained regressor (trained on 27 day light curves) in order to predict their rotation periods.

Ideally, we would train the model on TESS targets because the variables calculated from the light curves (Flicker/$R_{\rm var}$) are expected to differ between TESS and Kepler due to their different bandpasses. Detailed discussions of the difficulties of applying a model trained on Kepler to TESS are included in Section 5. However, we do not yet have a large-enough training set for TESS that includes enough rotation periods. Because of that, the result here is a preliminary test of how well the model, trained on Kepler, can predict rotation periods from TESS short-duration light curves.

The major difference between the results for simulated and real TESS light curves (Figures 8 and 9, respectively) is that the model, tested on real TESS data, suffers from higher bias for slow rotators. This may be due to additional white noise scatter in TESS light curves, which limits measurements of $R_{\rm var}$ and flicker in real TESS light curves. The signal-to-noise ratio for $R_{\rm var}$ (indicated by the size of the marker: the larger the marker, the higher the S/N) indicated that the summary statistics calculated from the TESS light curves are not reliable and are possibly dominated by the noise (further discussion in Section 5). As a result, the predictions are most likely dominated by the temperature of the star, which is supported by the clear color gradient. However, this preliminary test shows promising results in using RF to predict long rotation periods from short-duration light curves from TESS.

5. Discussion and Future Work

In performing this analysis, we revealed a few limitations and unforeseen possibilities for our RF classification and period prediction. We outline the most important of these below.

Better long rotation period predictions for Kepler stars—It is clear from the uncertainty analysis in Figures 6 and 8 that we are able to predict long rotation periods with a higher precision (∼4% better) than short rotation periods using the RF regressor. There are a few reasons why this might be the case.

1. *Inhomogeneous data*—We added stars with long rotation periods from Santos et al. (2019) and García et al. (2014), and they did not use the same methods to determine the rotation periods as McQuillan et al. (2014). Santos et al. (2019) and García et al. (2014) used the combination of wavelet analysis and the ACF, whereas McQuillan et al. (2014) only used the ACF. Because of the differences in their methods, the rotation period measurements from Santos et al. (2019) and García et al. (2014) could be slightly different from those from McQuillan et al. (2014). This could cause the data splitting in the RF regressor to be biased, causing it to find a slightly different relation between features and long rotation periods, and ending up with better predictions for slow rotators.
2. *Physics of the slow rotators*—Slow rotators might have a more straightforward relationship between their stellar

properties and their rotation periods. In Figure 3, there seems to be less scattering in $R_{\rm var}$ versus rotation period, and $R_{\rm var}$ is the most important feature in predicting rotation periods. In addition, rotation periods for fast rotators might still be affected by initial conditions from when the stars were born. As stars contract onto the main sequence, they gradually spin down. As a result, some of the fast rotators might still contain information on their birth angular momentum so their stellar properties are not closely related to their rotation periods.

Information in the light curve—The fact that we are able to predict long rotation periods (>27 days) by training on 27 day light curves, plus Gaia photometry, seems counter to intuition.

However, this is demonstrative of the utility of automated methodologies like RF regressors to learn the mapping from data to label on a data point-by-point basis. Similarly, Blancato et al. (2020) used a convolutional neural network to predict stellar properties, including rotation periods, directly from 27 day Kepler light curves. They are able to recover short rotation periods better than the method presented here for <35 day periods. This suggests that, by calculating only a couple summary statistics, we did not use all the information contained from the light curve. However, Blancato et al. (2020) are not able to predict rotation periods >35 days as accurately as our approach. The comparison could also support the idea that in order to accurately predict long rotation periods from short-duration light curves, we need more than just the information contained in the light curve themselves.

Limitations of predicting TESS rotation periods—There are a couple of important differences between Kepler and TESS that make applying a trained model on Kepler to TESS difficult. Here, we discuss differences in observing direction, bandpass, precision, and cadence.

1. Observing direction: TESS points at a different area of the sky every 27 days whereas Kepler only pointed at one direction. The kinematics used to train the model are not in the galactic coordinates system because the radial velocities are not available for most stars. Therefore, the v_tan and v_b relations with age are different in different directions. Although the kinematics were not that important for determining the rotation periods for Kepler stars (see Figure 5), we expect they may be more important for predicting stellar rotations for stars in the TESS observing field. As a result, we will only be able to predict rotation periods for stars in the direction of the Kepler field.
2. Bandpass differences: TESS and Kepler also have different observing bandpass and instrumental precision. TESS is targeting low-mass stars, which are cooler and redder, whereas Kepler is targeting Sun-like stars. As a result, TESS observes in the wavelengths of ~600–1100 nm, whereas Kepler observed between the wavelengths of ~400–900 nm. Because of this, any calculations made from the light curves (e.g., LG_peaks, $R_{\rm var}$ and flicker) are likely to be different. Figure 10 shows comparisons between the $R_{\rm var}$ and flicker calculated from the TESS and Kepler light curves for the 195 testing stars. Flicker calculated from TESS is always greater than that calculated from Kepler. This could be because the surface granulation signal of a star corresponding to flicker is louder in a redder bandpass. This would mean the flicker could potentially have more information about rotation periods in the TESS light curves and be a more important feature than it was in the Kepler training features. Alternatively, this could be because TESS light curves have higher-amplitude white noise background than Kepler light curves, which is added to the flicker estimate (see point below). One could correct these values based on TESS magnitude and obtain a better result on the TESS test set.
3. Instrumental precision: TESS has a lower instrumental precision at all magnitudes compares to Kepler. Figure 11 shows the systematic noise versus TESS magnitude for the TESS and Kepler light curves of the 195 testing stars. We calculated the systematic noise for these 195 stars by measuring the standard deviation of the flux in a 3 hr window and took the median of these values. Although following similar trends, the systematic noise in TESS light curves is one order of magnitude higher than that of Kepler for a given TESS magnitude.

In addition to the fact that TESS has a higher systematic noise in the light curves, the noise floor, especially for high TESS magnitudes, is comparable to

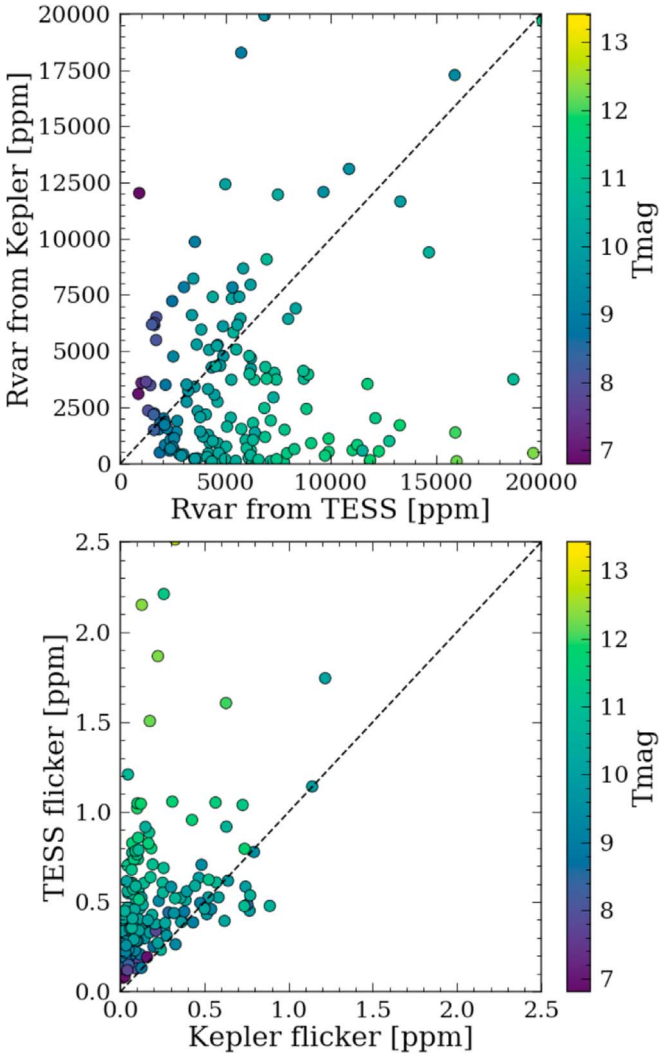

Figure 10. Comparisons between $R_{\rm var}$ [ppm] and flicker [ppm] calculated from the Kepler light curves and TESS light curves of the 195 testing stars, colored by TESS magnitude. There is a magnitude gradient in both plots, and flicker values calculated from TESS light curves are systematically higher than those from Kepler light curves.

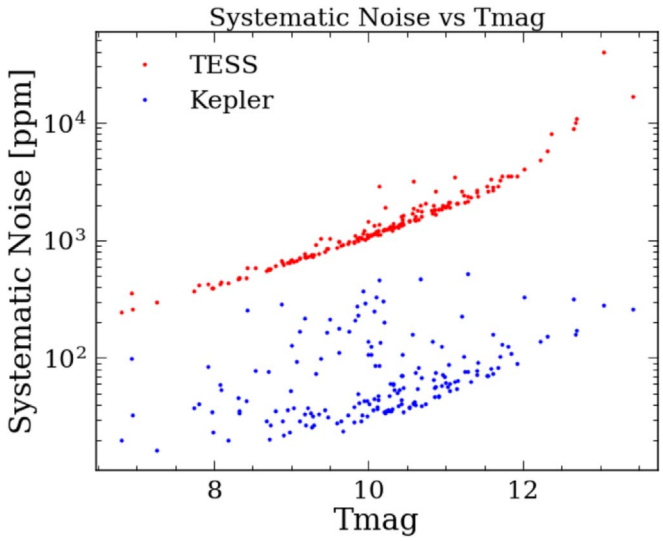

Figure 11. Systematic noise (standard deviation on a 3 hr window) vs. TESS magnitude for the 195 stars observed by both missions. At any given TESS magnitude, the systematic noise in the TESS light curve is always, on average, one magnitude higher than that of a Kepler light curve for the same star. This means any measurements extracted from the TESS and Kepler light curves are expected to be different.

the $R_{\rm var}$ and flicker measurements (see Figure 10). This could mean that these measurements are not accurate or even worse, we could be measuring the systematic noise instead of any physical quantities. The noise floor of TESS could also limit our ability to predict long rotation periods because stars with longer rotation periods typically exhibit lower $R_{\rm var}$ signals.

4. Pixel size: TESS has a pixel size of 21 arcseconds, which is large compared to Kepler, which has $3\farcs98$ pixels. This means the TESS light curves are more likely to be affected by contamination from nearby starlight.
5. Cadence: We calculated the light-curve statistics ($R_{\rm var}$ and flicker) from both the original TESS light curves (2 minute cadence) and the smoothed light curve (taking the rolling median of 30 minutes to simulate the Kepler cadence) and did not find significant changes. Therefore, the differences between the cadence in TESS and Kepler would not be significant. However, in this project, we only investigated the effects between 30 minute and 2 minute cadence data and extending the study to other cadence differences is interesting but beyond the scope of this project (Blancato et al. 2020 have done a more thorough study of the effect of cadence).

Despite the differences, we were still able to recover long rotation periods from real TESS light curves within 50% uncertainty. This means our model can potentially be applied to other surveys such as LSST and PLATO.

Potential alternative uses for this RF regressor—The main goal for this model is to predict long rotation periods (>15 days) for main-sequence stars from 27 day TESS light curves; however, it may have other applications. Because RF models are not particularly good at extrapolating data, any stars that have anomalous stellar parameters are most likely to be identified as outliers. Consequently, this model could potentially be used to gain insight on the outliers within data sets. Here, we list a couple of potential applications for this RF model:

1. If a star has a predicted rotation period much larger than the rotation period measured from traditional methods (LG, ACF, etc.), this star may have undergone tidal synchronization, resulting from a closely orbiting companion star. We could possibly create a synchronized binary detector with our regressor.
2. We could try to infer the inclination of a star by predicting $R_{\rm var}$ from the known rotation period. If a star is inclined to be almost pole on, its photometric variability measured directly from the light curve will be smaller than that predicted for the $R_{\rm var}$–stellar rotation relation.
3. We could compare the rotation periods of stars with close orbiting hot Jupiters and those without to study how these hot Jupiters might affect the rotation period and magnetic activity of their host stars.

Future work—Due to the limitations of predicting TESS rotation periods with a model trained on the Kepler data set, we will want to train our RF regressor on rotation periods measured from TESS targets across the entire observing zone using ACF. We will then want to create a catalog of TESS rotation periods that can be used by the astronomy community. It would also be interesting to investigate how the sparsity in feature parameter space affects the model prediction.

6. Conclusion

Rotation periods are important for studying stellar magnetic activity, improving RV measurements for exoplanet searches, and even determining stellar ages. Stellar rotation periods have been precisely measured using traditional methods, such as periodograms and ACFs, for Kepler targets. However, instead of having 4 yr light curves, most TESS stars will only have 27 day light curves for every one-year observing window. This increases the difficulty of using traditional methods to recover rotation periods, especially those of M dwarfs, which often have periods greater than 27 days (McQuillan et al. 2014).

We presented a new method to predict long rotation periods from short-duration light curves using RF, an ML algorithm. We first trained an RF classifier on stars from McQuillan et al. (2014), Santos et al. (2019), García et al. (2014), and Gaia DR2 (Gaia Collaboration et al. 2016, 2018) and distances from Bailer-Jones et al. (2018) to identify whether the rotation period of a star is "measurable." A regressor, trained on the same targets, was then used if the rotation period of a star could be predicted based on the classifier. The data set and features used to train these models were described in Section 2. We find that the most important features used to predict the rotation periods are $R_{\rm var}$, effective temperature, Gaia color, luminosity, and flicker. We calculated the uncertainties by calculating the median absolute deviation of predicted rotation periods. We were able to predict rotation periods of Kepler stars with an average uncertainty of 13% (9% for rotation periods >30 days) with 4 yr light curves and 9% (5% for rotation periods >30 days) with 27 day light curves. We found that long rotation periods were predicted more precisely than short rotation periods. When applying this regressor trained on Kepler data to TESS data, we were able to recover rotation periods of TESS stars in the Kepler field with an uncertainty of 55%. The decrease in precision was most likely due to the differences between the two missions, described in Section 5. This preliminary test on TESS stars showed promising results, and we expect to be able to predict rotation periods with smaller

errors if we can train the regressor on TESS targets. The two open-source software packages, FLICKER and Astraea, developed in this project, are available on Github and are described in Section Appendix. In the future, we hope to train the RF regressor on TESS data and create a catalog of rotation periods.

R.A. acknowledges support from NASA award 80NSSC 20K1006

We want to thank Angela Santos for sharing her data.

This work made use of the gaia-kepler.fun cross-match database created by Megan Bedell.

This paper includes data collected by the Kepler mission. Funding for the Kepler mission is provided by the NASA Science Mission directorate.

This paper includes data collected by the TESS mission. Funding for the TESS mission is provided by the NASA Explorer Program.

This work has made use of data from the European Space Agency (ESA) mission Gaia (https://www.cosmos.esa.int/gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement.

This research made use of Astropy,[7] a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013; Price-Whelan et al. 2018).

Facilities: Gaia, Kepler, TESS.

Software: Astraea (this work), Astropy (Astropy Collaboration et al. 2013; Price-Whelan et al. 2018), FLICKER (this work), Numpy (Oliphant 2006), Scikit-learn (Pedregosa et al. 2011), Scipy (Virtanen et al. 2020), Pandas (pandas development team 2020), Matplotlib (Hunter 2007).

Appendix
Software Products

This project resulted in two open-source software packages in Python: FLICKER (https://github.com/lyx12311/FLICKER) and Astraea (https://github.com/lyx12311/Astraea). FLICKER can be used to calculate flicker for one light curve or multiple light curves. It calculates the median flicker across light curves if passed through a multidimensional array. Figure A1 shows the comparison between flicker values provided in Bastien et al. (2013) and those calculated with FLICKER for 100 Kepler stars listed in their paper.

Astraea is a software package that includes the RF classifier and regressor trained on Kepler targets. It can be used to recover rotation periods for any stars observed by Kepler or TESS. However, because this model is only trained on Kepler stars, any rotation periods predicted for targets outside of the Kepler field are subject to higher uncertainties.

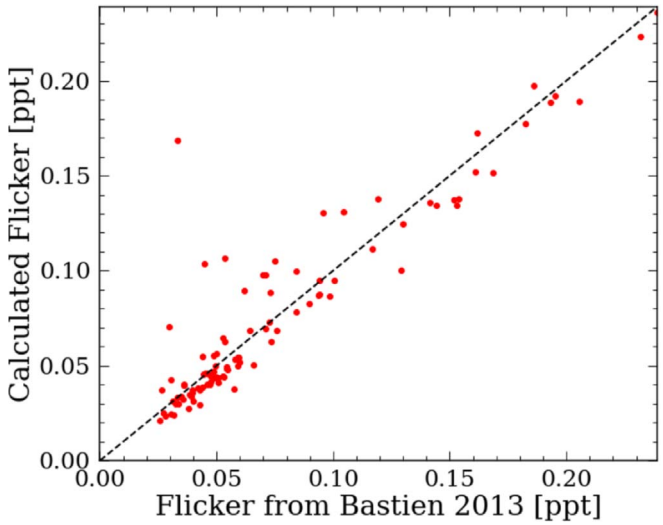

Figure A1. 8 hr flicker values calculated by Bastien et al. (2013) vs. those calculated with the software package FLICKER. The results are consistent with one another.

ORCID iDs

Yuxi(Lucy) Lu https://orcid.org/0000-0003-4769-3273
Ruth Angus https://orcid.org/0000-0003-4540-5661
Marcel A. Agüeros https://orcid.org/0000-0001-7077-3664
Melissa Ness https://orcid.org/0000-0001-5082-6693
Jason L. Curtis https://orcid.org/0000-0002-2792-134X
Sam Grunblatt https://orcid.org/0000-0003-4976-9980

References

Andrae, R., Fouesneau, M., Creevey, O., et al. 2018, A&A, 616, A8
Angus, R., Beane, A., Price-Whelan, A. M., et al. 2020, AJ, 160, 90
Angus, R., Morton, T., Aigrain, S., Foreman-Mackey, D., & Rajpaul, V. 2018, MNRAS, 474, 2094
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Aumer, M., & Binney, J. J. 2009, MNRAS, 397, 1286
Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, AJ, 156, 58
Barbanis, B., & Woltjer, L. 1967, ApJ, 150, 461
Barnes, S. A. 2003, ApJ, 586, 464
Barnes, S. A. 2007, ApJ, 669, 1167
Bastien, F. A., Stassun, K. G., Basri, G., & Pepper, J. 2013, Natur, 500, 427
Bird, J. C., Kazantzidis, S., Weinberg, D. H., et al. 2013, ApJ, 773, 43
Blancato, K., Ness, M., Huber, D., et al. 2020, arXiv:2005.09682
Borucki, W. J., Koch, D., Basri, G., et al. 2010, Sci, 327, 977
Breiman, L. 2001, Mach. Learn., 45, 5
Breiman, L., Friedman, J., Stone, C., & Olshen, R. 1984, Classification and Regression Trees (Boca Raton, FL: CRC Press)
Butkevich, A. G., & Lindegren, L. 2014, A&A, 570, A62
Davenport, J. R. A., Covey, K. R., Clarke, R. W., et al. 2019, ApJ, 871, 241
Davies, G. R., Chaplin, W. J., Farr, W. M., et al. 2015, MNRAS, 446, 2959
Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, AJ, 154, 220

[7] http://www.astropy.org

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1
García, R. A., Ceillier, T., Salabert, D., et al. 2014, A&A, 572, A34
Grunblatt, S. K., Howard, A. W., & Haywood, R. D. 2015, ApJ, 808, 127
Hartman, J. D., Bakos, G. Á, Noyes, R. W., et al. 2011, AJ, 141, 166
Haywood, R. D., Collier Cameron, A., Queloz, D., et al. 2014, MNRAS, 443, 2517
Hillenbrand, L., Isaacson, H., Marcy, G., et al. 2015, in 18th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, ed. G. van Belle & H. C. Harris (Cambridge: Cambridge Univ. Press), 759
Holmberg, J., Nordström, B., & Andersen, J. 2007, A&A, 475, 519
Holmberg, J., Nordström, B., & Andersen, J. 2009, A&A, 501, 941
Hunter, J. D. 2007, CSE, 9, 90
Ilin, E., Schmidt, S. J., Davenport, J. R. A., & Strassmeier, K. G. 2019, A&A, 622, A133
Kawaler, S. D. 1988, ApJ, 333, 236
Lacey, C. G. 1984, MNRAS, 208, 687
LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, arXiv:0912.0201
McQuillan, A., Mazeh, T., & Aigrain, S. 2014, ApJS, 211, 24
Miller, A. A., Bloom, J. S., Richards, J. W., et al. 2015, ApJ, 798, 122
Nordström, B., Mayor, M., Andersen, J., et al. 2004, A&A, 418, 989
Oliphant, T. E. 2006, A guide to NumPy, Vol. 1 (USA: Trelgol Publishing)
pandas development team 2020, pandas-dev/pandas: Pandas, v1.1.2, Zenodo, doi:10.5281/zenodo.3509134
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, J. Mach. Learn. Res., 12, 2825, https://jmlr.org/papers/v12/pedregosa11a.html
Pizzolato, N., Maggio, A., Micela, G., Sciortino, S., & Ventura, P. 2003, A&A, 397, 147
Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., et al. 2018, AJ, 156, 123
Rajpaul, V., Aigrain, S., Osborne, M. A., Reece, S., & Roberts, S. 2015, MNRAS, 452, 2269
Rauer, H., Catala, C., Aerts, C., et al. 2014, ExA, 38, 249
Reinhold, T., & Gizon, L. 2015, A&A, 583, A65
Richards, J. W., Starr, D. L., Butler, N. R., et al. 2011, ApJ, 733, 10
Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, JATIS, 1, 014003
Santos, A. R. G., García, R. A., Mathur, S., et al. 2019, ApJS, 244, 21
Sellwood, J. A. 2014, RvMP, 86, 1
Sellwood, J. A., & Carlberg, R. G. 1984, ApJ, 282, 61
Simonian, G. V. A., Pinsonneault, M. H., & Terndrup, D. M. 2019, ApJ, 871, 174
Skumanich, A. 1972, ApJ, 171, 565
Soderblom, D. R. 2010, ARA&A, 48, 581
Spitzer, L. J., & Schwarzschild, M. 1951, ApJ, 114, 385
Strömberg, G. 1946, ApJ, 104, 12
Ting, Y.-S., & Rix, H.-W. 2019, ApJ, 878, 21
van Saders, J. L., Ceillier, T., Metcalfe, T. S., et al. 2016, Natur, 529, 181
van Saders, J. L., & Pinsonneault, M. H. 2013, ApJ, 776, 67
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Walkowicz, L. M., & Basri, G. S. 2013, MNRAS, 436, 1883
Yu, J., & Liu, C. 2018, MNRAS, 475, 1093