Unsigned Magnetic Flux as a Proxy for Radial-velocity Variations in Sun-like Stars

R. D. Haywood[1,2] , T. W. Milbourne[3,4], S. H. Saar[3] , A. Mortier[5] , D. Phillips[3] , D. Charbonneau[3] , A. Collier Cameron[6] ,
H. M. Cegla[7,8] , N. Meunier[9], and M. L. Palumbo III[10]
[1] Astrophysics Group, University of Exeter, Exeter EX4 2QL, UK; R.D.Haywood@exeter.ac.uk
[2] STFC Ernest Rutherford Fellow
[3] Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[4] Department of Physics, Harvard University, 17 Oxford Street, Cambridge MA 02138, USA
[5] KICC & Cavendish Laboratory, University of Cambridge, J.J. Thomson Avenue, Cambridge CB3 0HE, UK
[6] Centre for Exoplanet Science, SUPA School of Physics and Astronomy, University of St Andrews, North Haugh, St Andrews KY16 9SS, UK
[7] University of Warwick, Department of Physics, Gibbet Hill Road, Coventry, CV4 7AL, UK
[8] Observatoire de Genève, Université de Genève, 51 Chemin des Maillettes, CH-1290 Versoix, Switzerland
[9] Université Grenoble Alpes, CNRS, IPAG, F-38000 Grenoble, France
[10] Pennsylvania State University, State College, PA 16801, USA
Received 2020 May 27; revised 2022 June 23; accepted 2022 June 23; published 2022 August 8

Abstract

A major obstacle to detecting and characterizing long-period, low-mass exoplanets is the intrinsic radial-velocity (RV) variability of host stars. To better understand RV variability, we estimate disk-averaged RV variations of the Sun over its magnetic cycle, from the Fe I line observed by SDO/HMI, using a physical model for rotationally modulated magnetic activity that was previously validated against HARPS-N solar observations. We estimate the unsigned magnetic flux and show that a linear fit to it reduces the RMS of RV variations by 62%, i.e., a factor of 2.6. We additionally apply the FF' method, which predicts RV variations based on a star's photometric variations. At cycle maximum, we find that additional processes must be at play beyond suppression of convective blueshift and velocity imbalances resulting from brightness inhomogeneities, in agreement with recent studies of RV variations. By modeling RV variations over the magnetic cycle using a linear fit to the unsigned magnetic flux, we recover injected planets at a period of $\approx$300 days with RV semi-amplitudes down to $0.3 \mathrm{~m~s}^{-1}$. To reach $0.1 \mathrm{~m~s}^{-1}$, we will need to identify and model additional phenomena that are not well traced by $|\hat{B}_{\mathrm{obs}}|$ or FF'. This study motivates ongoing and future efforts to develop observation and analysis techniques to measure the unsigned magnetic flux at high precision in slowly rotating, relatively inactive stars like the Sun. We conclude that the unsigned magnetic flux is an excellent proxy for rotationally modulated, activity-induced RV variations, and could become key to confirming and characterizing Earth analogs.

Unified Astronomy Thesaurus concepts: Solar cycle (1487); Active sun (18); Quiet sun (1322); Sunspots (1653); Solar faculae (1494); Exoplanet detection methods (489); Radial velocity (1332); Astronomy data analysis (1858)

1. Introduction

The main obstacle we face in detecting, confirming, and characterizing Neptune- to Earth-mass exoplanets via radial-velocity (RV) monitoring is the intrinsic variability of the host stars themselves (see the EPRV Working Group's Final Report, Crass et al. 2021; National Academies of Sciences, Engineering, and Medicine 2018, and references in both). RV monitoring is the most widely applicable technique to determine the masses of the small planets to be discovered by TESS and PLATO. Mass is the most fundamental parameter of a planet: it is central to theoretical models of planet composition and structure (e.g., Zeng & Sasselov 2013). Planetary mass dictates the amount of observing time required to characterize a planet's atmosphere, so it is essential that we know masses reliably to plan atmospheric follow-up observations (Morley et al. 2017; Batalha et al. 2019), e.g., with JWST and ARIEL. To determine accurate and precise planetary masses, we need to develop robust, physically motivated models for stellar variability. We still lack a complete and detailed understanding of how the interplay between magnetic fields and granulation gives rise to RV variations on the Sun and other stars (Crass et al. 2021, Section A2).

The Sun is the only star whose surface we can image directly and at high resolution, making it an ideal test bench to examine the physical phenomena responsible for intrinsic RV variability. It is also the only star whose RV we know independently of spectroscopic measurements (e.g., from the High Accuracy Radial-velocity Planet Searcher in the Northern hemisphere, HARPS-N; Cosentino et al. 2012).

On timescales of several rotation periods (weeks–months), RV variability is driven by magnetic activity in the photosphere. The manifestations of magnetic activity relevant to the present analysis are sunspots and faculae. Sunspots are relatively large, dark areas of strong magnetic fields (e.g., Foukal 2004). Faculae are small, bright magnetic flux tubes (Spruit 1976). They tend to be located in the lanes between supergranular cells and are spread all over the solar surface, thus forming the photospheric magnetic *network* (e.g., Foukal 2004). In regions of enhanced magnetic activity, faculae cluster into areas of *plage* (e.g., Schrijver & Zwaan 2000, Figure 1.1). The total surface area covered by the network varies throughout the Sun's 11-yr cycle, although to a much smaller extent than spots and plage (Meunier 2018, 2003). At low activity levels, magnetic elements tend to be spread throughout the solar surface in the form of the network. Plage coverage increases with magnetic activity, and at high activity levels, the majority of
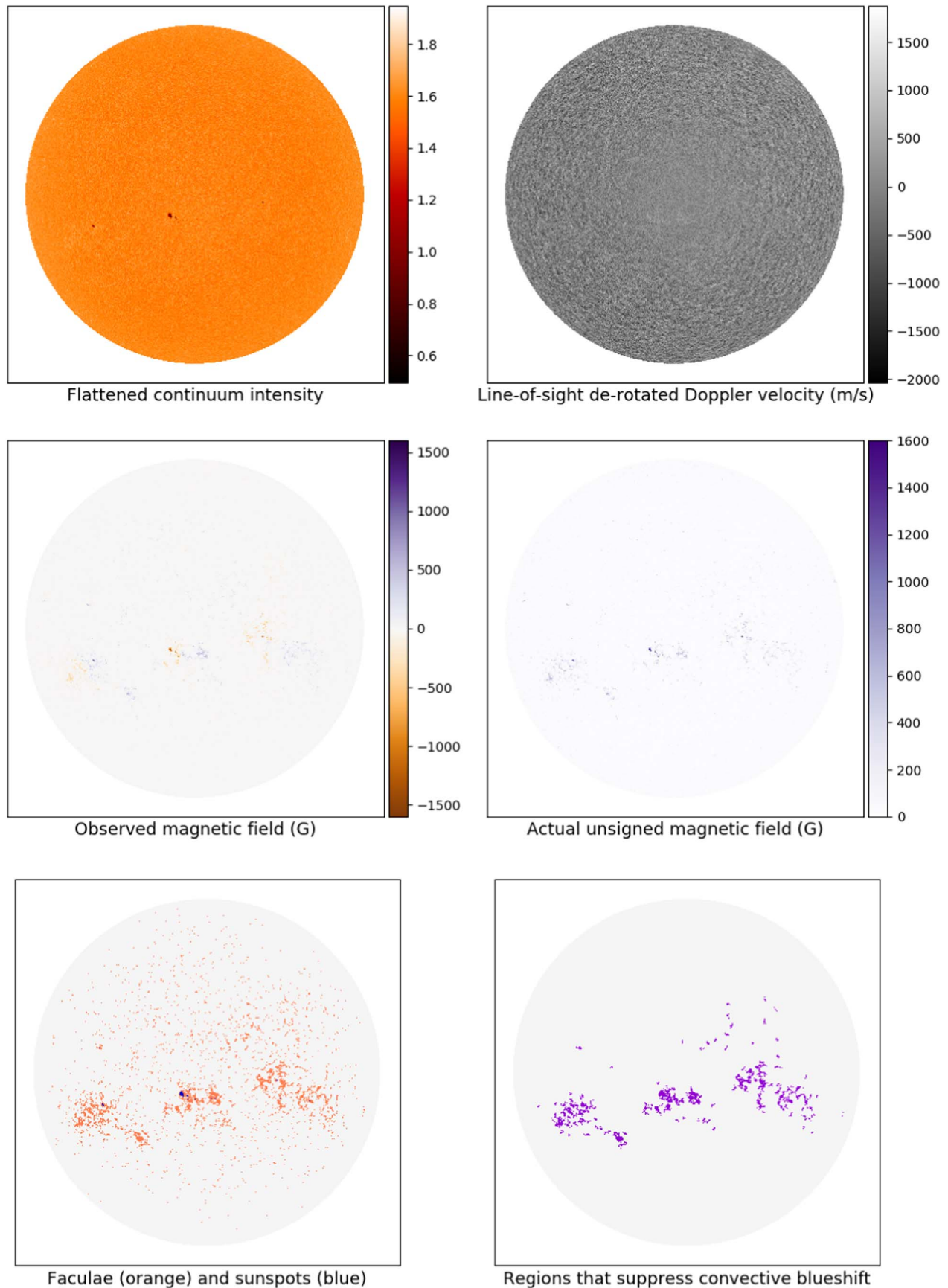
Figure 1. SDO/HMI data products for an observation set taken on 2015 November 28 at 20:00:00 UTC. This set of images is representative of the Sun during high activity levels. Faculae and sunspots fill 3.25% and 0.03% of the solar disk, respectively. The areas that suppress convective blueshift, predominantly faculae in regions of plage, shown in the last panel, fill 1.59% of the solar disk. The disk-averaged unsigned magnetic flux is 9.99 G. *Notes:* The flattened intensity is normalized to the mean intensity. The line-of-sight velocity is shown after subtracting the solar rotation profile and the velocity of the spacecraft.

Table 1
Uncertainties for the Quantities Estimated in this Analysis, Namely the Line-of-sight (LOS) Velocity and the Unsigned LOS Magnetic Flux

Quantity	Measured to Precision	Notes and References
LOS velocity per pixel, $v_{\rm pix}$	7 m s^{-1}	Photon noise at disk center in Dopplergram (hmi.V_720s). From (Couvidat et al. 2016, Table 1).
LOS velocity over full solar disk	0.002 m s^{-1}	Average over the full time series of: $v_{\rm pix}/\sqrt{n_{\rm pix}}$, where $n_{\rm pix}$ is the number of pixels within $\mu = 0.3$ in a given Dopplergram.
Spacecraft velocity	0.01 m s^{-1}	From Couvidat et al. 2016.
Pipeline to measure ΔRV from HMI images	0.1 m s^{-1}	Systematic uncertainty in our analysis, particularly in the classification of active region areas.
LOS velocity (disk averaged)	0.1 m s^{-1}	We add the instrument, pipeline, and astrophysical uncertainties listed above in quadrature.
LOS unsigned magnetic flux per pixel, $B_{\rm pix}$	3 G	Photon noise at disk center in magnetogram (hmi.M_720s). From (Couvidat et al. 2016, Table 1).
LOS unsigned magnetic flux (disk averaged)	0.0009 G	Average over the full time series of: $B_{\rm pix}/\sqrt{n_{\rm pix}}$.

**Note.** See details in Section 2.1.2.

magnetic elements on the surface are concentrated in plage, rather than in network (e.g., see plage and network filling factors shown in Figure 4). Plage generally decays into network over timescales of several rotations, which explains the larger network coverage when activity is high.

Magnetic elements inhibit convective motions, thereby suppressing some of the convective blueshift that results from granulation (e.g., Dravins et al. 1981). This suppression of convective blueshift is the dominant contributor to RV variations in the Sun (Saar & Donahue 1997). Solar observations show RV variations, modulated by the Sun's rotation and evolving over timescales of days to weeks, with amplitudes of several meters per second (Meunier et al. 2010a, 2010b; Dumusque et al. 2014; Haywood et al. 2016; Milbourne et al. 2019). Convective blueshift is suppressed by faculae in concentrated areas of plage, while faculae in the network, being more spatially diffuse, do not perturb convective flows significantly (Milbourne et al. 2019, Section 4.4). Sunspots contribute little to observed suppression of convective blueshift, as they are dark (and thus contribute relatively less to observed spectra) and cover a very little area in comparison to faculae (e.g., Lagrange et al. 2010; Haywood et al. 2016). At high solar activity levels, sunspots do contribute significantly to RV variations. Because they are much darker than the surrounding photosphere, they produce significant inhomogeneities in surface brightness, which result in RV variations with an RMS of 60 cm s^{-1} and frequent peak-to-peak amplitudes of 2 m s^{-1}, with variations of up to 5 m s^{-1} (Lagrange et al. 2010). On the other hand, faculae (both in the network and plage) are only 10% brighter than the quiet photosphere at optical wavelengths; moreover, they are distributed much more uniformly longitudinally on the solar disk, so their brightness-induced RV contribution mostly cancels out on rotational timescales. On longer (magnetic cycle) timescales, large-scale changes in the number of faculae can produce long-term, bulk RV shifts (e.g., Saar & Fischer 2000; Meunier et al. 2010b, Figure 8).

Magnetic elements enhance chromospheric column density, which strengthens the emission reversals in the Ca II H&K cores. Thus, the filling factor of magnetic elements correlates strongly with the amount of emission in the cores of the Ca II H&K lines (e.g., Meunier 2018, Figure 1) as measured by the $\log R'_{\rm HK}$ index (Vaughan et al. 1978; Noyes et al. 1984). Solar observations show that RV variations and $\log R'_{\rm HK}$ correlate strongly over the Sun's 11-yr magnetic cycle (e.g., Meunier et al. 2010a, Figure 13). However, the $\log R'_{\rm HK}$ does not trace RV variations below the meter-per-second barrier; a linear fit to the

RV–Ca index relation over a full solar magnetic cycle brings the RV RMS to 1.05 m s^{-1} (e.g., Meunier & Lagrange 2013, Figure 12). This RMS is of order 2 m s^{-1} during high magnetic activity levels. During low activity levels, over timescales of a few solar rotations, the RV–Ca correlation is relatively low (e.g., Meunier et al. 2010a, Figure 13; see also Figure 5 in this paper). This can be explained by the fact that magnetic elements are predominantly found in the network, which contributes to Ca II H&K emission but does not significantly affect RV variations (Milbourne et al. 2019). At high activity levels (on the rotation timescale), RV variations do not correlate well with Ca II H&K emission, e.g., in (Lagrange et al. 2010, Figure 12) and (Haywood et al. 2016, Figure A1). Part of this discrepancy may be that the Ca II H&K emission forms in the chromosphere, and therefore follows a different limb-darkening law and projection effects than RVs, which are measured from photospheric absorption lines (see Appendix A in this paper). To summarize, solar RV variations induced by magnetic activity on timescales of the order of a few rotation periods are not directly correlated and in phase with Ca II H&K emission, in part because faculae affect RV variations differently depending on whether they are in a sparse network or in concentrated regions of plage, and due to changes in spectral line profiles.

We see a similar behavior in observations of Sun-like stars. As part of the Mount Wilson HK Project, several dozen slowly rotating, old Sun-like stars were monitored in optical photometry and in Ca II emission over the past several decades (Baliunas et al. 1995; Wilson 1978). Observations showed that as these stars became more magnetically active (as indicated via their S-index), they also got brighter. Their surfaces are therefore dominated by bright faculae rather than dark spots, just like the Sun (Lockwood et al. 2007; Radick et al. 2018). On timescales of several years, the RMS of stellar RV variations increases as the $\log R'_{\rm HK}$ increases (Saar & Fischer 2000; Lovis et al. 2011). Aigrain et al. (2012) developed a model to estimate stellar activity-induced RV variations based on the star's optical photometric variations. The model accounts for RV variations produced by dark spots and bright plage *that are spatially associated with spots*, both through rotational flux imbalance and suppression of convective blueshift. Haywood et al. (2014) tested this model on the solar analog CoRoT-7 using simultaneous photometric and RV observations taken at high cadence over a rotation period. While the model of Aigrain et al. (2012) captures a significant part of the rotationally modulated RV signal (with an RMS scatter of order 2 m s^{-1}), it leaves out rotationally modulated variations with an RMS of
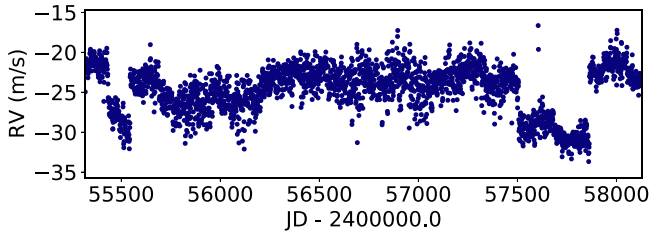
Figure 2. Disk-averaged radial velocity of the Sun as estimated from SDO/HMI images, before subtracting the disk-averaged velocity of the quiet Sun. The jumps in RV are instrument systematics, as HMI is not calibrated for long-term stability.

$3.95 \, \mathrm{m \, s^{-1}}$. This additional RV signal likely originates from magnetic regions that have low-intensity contrast, and whose brightness-induced RV variations are therefore low. Hojjatpanah et al. (2020) examined correlations between contemporaneous HARPS RV variations and TESS photometry of 171 GKM stars and found two distinct regimes, where faculae-dominated stars exhibit a weaker correlation between RV and photometry compared to spot-dominated stars. These findings are in agreement with precursor studies by Bastien et al. (2014) and Oshagh et al. (2017). These spectroscopic and photometric observations of CoRoT-7 and other stars are consistent with the Sun's behavior, i.e., that the solar surface is dominated by low-contrast plage, which drives RV variations.

At present, we cannot yet routinely measure a proxy for activity-induced RV variations that enables us to confirm and characterize long-period exoplanets with RV semi-amplitudes of $30 \, \mathrm{cm \, s^{-1}}$ and below (Crass et al. 2021). On both the Sun and other Sun-like stars, Ca II H&K emission does not systematically correlate as strongly with activity-induced RV variations. For the Sun, a strong correlation has been observed between activity-induced RV variations and the unsigned, full-disk magnetic flux, over the magnetic cycle (Deming & Plymate 1994; Lanza et al. 2016; Meunier 2018) and on the rotation timescale (Haywood et al. 2016). The disk-averaged RV time series estimated from the Michelson Doppler Imager on board the Solar and Heliospheric Observatory (SoHo/MDI) by (Meunier et al. 2010b, and subsequent papers) is well sampled and spans over 4 yr of Cycle 23, prior to SDO's launch; however, it cannot be compared to *ground-truth*, direct disk-integrated RV observations. To date, systematic RV campaigns of the Sun as a star have been carried out spectroscopically, with HARPS via sunlight reflected from asteroids (Haywood et al. 2016; Lanza et al. 2016) and more recently with dedicated solar feeds at HARPS-N (Dumusque et al. 2015; Collier Cameron et al. 2019; Dumusque et al. 2021), HARPS (Dumusque 2019, HELIOS) and NEID (Lin et al. 2022). The longest running spectroscopic solar RV campaign, by HARPS-N, began in 2015, when the Sun's variability levels were already past maximum and declining.

In the present analysis, we estimate rotationally modulated RV variations of the Sun from SDO/HMI images over 8 yr, at a daily cadence (Section 2). We estimate RVs from HMI images using a technique that has been previously validated against direct Sun-as-a-star HARPS-N observations (Section 3). Our resulting time series spans almost a full magnetic cycle and covers the cycle's maximum, which was not observed by HARPS-N. We present time series of RV variations and unsigned magnetic flux in Section 3. We model our RV time series using a linear fit in unsigned magnetic flux in Section 4 and with the FF' model of Aigrain et al. (2012) in Section 5. In Section 6, we identify potential additional physical effects giving rise to rotationally

modulated RV variations that are not well traced by either model, and identify limitations of the FF' model. We perform simple planet injections to assess the performance of the unsigned magnetic flux for mitigating rotationally modulated RV variations in Section 7. We discuss future prospects, including ways to measure the unsigned magnetic flux in other stars in Section 8, and present our conclusions in Section 9.

2. Data

2.1. SDO/HMI Images

We use 720 s HMI exposures of continuum intensity (both uncorrected and corrected for limb darkening by the HMI team), Dopplergrams, and magnetograms, as represented in Figure 1. The HMI instrument takes six measurements of intensity across a narrow wavelength range centered on the Fe I line at 6173 Å (see Figure 6 of Schou et al. 2012). These points are fitted with a Gaussian profile to generate the main HMI data products, to generate the main HMI data products including velocity (line shift), intensity (depth), magnetic field strength (width due to Zeeman broadening), and continuum intensity (Schou et al. 2012, Section 3.3). The line shifts and magnetic field values extracted for each pixel should be independent, physical quantities, obtained from combinations of the different intensities at different points on the measured Fe I line. While the line asymmetry stemming from convection within each pixel is not preserved, we expect this technique to capture asymmetries due to physical processes occurring over scales larger than a pixel. We refer the reader to Schou et al. (2012) for further details on how these images are extracted from the raw filtergrams.

2.1.1. Temporal Sampling

SDO/HMI has operated almost continuously since the start of the mission, except for spacecraft operations and calibrations, and eclipses that happen due to the geosynchronous orbit of the SDO spacecraft (Hoeksema et al. 2018). There have been very few anomalies requiring interruptions, none of which have been prolonged compared to the seasonal eclipses. We take a set of images every 4 hr (6 times per 24 hr period) from 2010 April 7, 04:00:00 UTC up to 2018 January 12 20:00:00 UTC, amounting to a total of 16,855 sets of images spanning 2811 days. We take daily averages to minimize the contribution of short-term processes, namely oscillations and magnetoconvection. We choose not to use the SDO/HMI images at their highest cadence in order to maintain the relevance of this analysis to stellar studies, while still sampling the Sun multiple times a day. Indeed, the cadence of current and planned stellar RV surveys is 1–3 observations per night at most.

2.1.2. Instrument Precision and Stability

All sources of uncertainty that affect our RV estimates are listed in Table 1. Hoeksema et al. (2018) recently assessed the performance of HMI and reported that the instrument continues to work according to its original specification. The data products are corrected on a regular basis as the calibrations improve (e.g., instrument thermal environment, focus, image distortions, optics alignment, cosmic-ray corrections, etc.). They report that the quality of the data is very uniform with time. The HMI data products are well calibrated (Hoeksema et al. 2018), with the exception of the long-term stability of the Dopplergrams. The Doppler velocity maps were designed for

helioseismology investigations, so they are not calibrated to be stable over timescales longer than a few hours or days (Schou et al. 2011). This is shown in Figure 2, which shows the disk-averaged velocity of the Sun taking irregular jumps of several meters per second over the course of the SDO mission. To correct for this effect, we perform all of our velocity calculations relative to the disk-averaged velocity of the magnetically inactive, quiet Sun for each Doppler image (as in Meunier et al. 2010b; Haywood et al. 2016; Milbourne et al. 2019). We estimate the velocity of the quiet Sun by summing over all pixels identified as nonmagnetic (see Section 3), and excluding pixels considered magnetically active (even at the peak of the Sun's magnetic cycle, fewer than 5% of all pixels within the solar disk are magnetically active). We are therefore only considering RV *variations* (ΔRV). Importantly, this subtraction cancels out all velocity flows from the quiet Sun. This means that our RV estimates are free from pressure-mode (p-mode) oscillations, granulation, and supergranulation motions, which would otherwise induce uncorrelated noise at the $1 \, \mathrm{m \, s^{-1}}$ level (e.g., Meunier et al. 2015).

Beyond the lack of long-term calibration, Couvidat et al. (2016) report that the most significant source of instrument-related uncertainty that remains in individual HMI Doppler-grams is the orbital velocity of the spacecraft, which is uncertain to $0.01 \, \mathrm{m \, s^{-1}}$. Indeed, we see a systematic sinusoidal shift with a periodicity of 12 and 24 hr due to the spacecraft's orbit. This systematic should mostly average out through our sampling, but to be on the safe side, we add Couvidat et al. (2016)'s uncertainty of $0.01 \, \mathrm{m \, s^{-1}}$ in quadrature to our RV uncertainties (see Table 1). The total number of pixels inside the solar disk as it appears on the HMI image varies by $\pm 3\%$ over the course of each year, primarily due to Earth's eccentric orbit. To correct for this effect, we normalize all our quantities by the total number of pixels in each image, i.e., we estimate disk-averaged quantities. We ran tests to assess the potential effect of hypothetical spurious pixels from the continuum, Doppler, and magnetic images. For example, we find that setting the velocity to zero for a large patch of a 20×20 pixel square incurs changes in RV of $0.04 \, \mathrm{m \, s^{-1}}$. The RV effect of spurious pixels distributed randomly on the solar disk is less than $0.01 \, \mathrm{m \, s^{-1}}$. To be conservative, we assign a constant RV uncertainty of $0.1 \, \mathrm{m \, s^{-1}}$ for all RV estimates, to account for uncertainties arising from our pipeline to estimate RVs (detailed in Section 3). Our choice of magnetic and continuum intensity thresholds are based on previous studies, but they remain somewhat arbitrary, and changing them slightly will impact RV estimates at the centimeter-per-second level. We conservatively remove sets of images with $v \sin i$ values more than 3σ deviant from the mean (6 out of 16,855), and those with focal length values more than 3σ deviant (2 out of 16,855).

The instrument systematics are independent of those that HARPS-N and other stellar spectrographs suffer from (e.g., Dumusque et al. 2021); therefore the HMI-RV time series complements spectroscopic data sets for the study of solar variability.

2.2. SORCE Total Solar Irradiance Observations

To apply the FF' method of Aigrain et al. (2012) in Section 5, we use Total Solar Irradiance (TSI) (akin to a Kepler light curve for the Sun) observations taken by the Total Irradiance Monitor (TIM) on board the SOlar Radiation and Climate Experiment (SORCE) satellite (Kopp et al. 2001; Lean et al. 2005).[11] We use the TIM's daily average TSI measurements, which span the full duration of our SDO/HMI time series. The TIM takes observations every 50 s when the spacecraft faces the Sun, and these observations are then combined to produce daily averages. We concatenate this time series with our SDO/HMI time series of daily images. Because there is a gap in the SORCE time series around 1200–1400 days into the SDO mission, we are left with a combined time series of 2535 daily observations, spanning 2811 days. The SORCE time series is plotted in the last panel of Figure 4. The TIM achieves a precision of 4–17 ppm per observation (Kopp 2014). For comparison, Kepler achieved the same level of precision on a 7 to 9th V magnitude star in a long-cadence observation[12] (30-minute integration time). TIM has a long-term stability of about 10 ppm yr^{-1} (Kopp 2014).

2.3. S-index from Mount Wilson and HARPS-N

We compare our SDO/HMI-derived quantities (ΔRV, $|\hat{B}_{\mathrm{obs}}|$) against Ca II H&K emission observations. For this, we use overlapping S-index observations of the Sun seen as a star at the Mount Wilson Observatory as part of the HK Project, fully homogenized and calibrated by Egeland et al. (2017). Their observations run from 1966 until 2015. To cover the 2015–2018 period, we use daily averaged S-index observations taken by the solar telescope that fed the HARPS-N spectrograph since 2015 July (Milbourne et al. 2019). The Mount Wilson and HARPS-N data sets overlap for around 45 days in 2015, which we can use to stitch the two S-index time series together. We do this by rescaling the overlapping part of the HARPS-N S-index time series so it has the same variance as the Mount Wilson S-index time series and subtracting the offset between the two data sets. The full time series is shown in the second to last panel of Figure 4.

3. Estimating the Full-disk RV Variations and Magnetic Flux of the Sun

We estimate disk-averaged active region filling factors, RV variations, and unsigned (unpolarized) magnetic fluxes of the Sun using spatially resolved images from SDO/HMI images according to the same method as Milbourne et al. (2019), adapted from that of Haywood et al. (2016), which builds on the techniques originally developed by Meunier et al. (2010b) and Fligge et al. (2000). Ervin et al. (2022) produced a publicly available Python pipeline, SolAster, which can estimate RV and unsigned magnetic flux time series similar to those presented in this present analysis.[13]
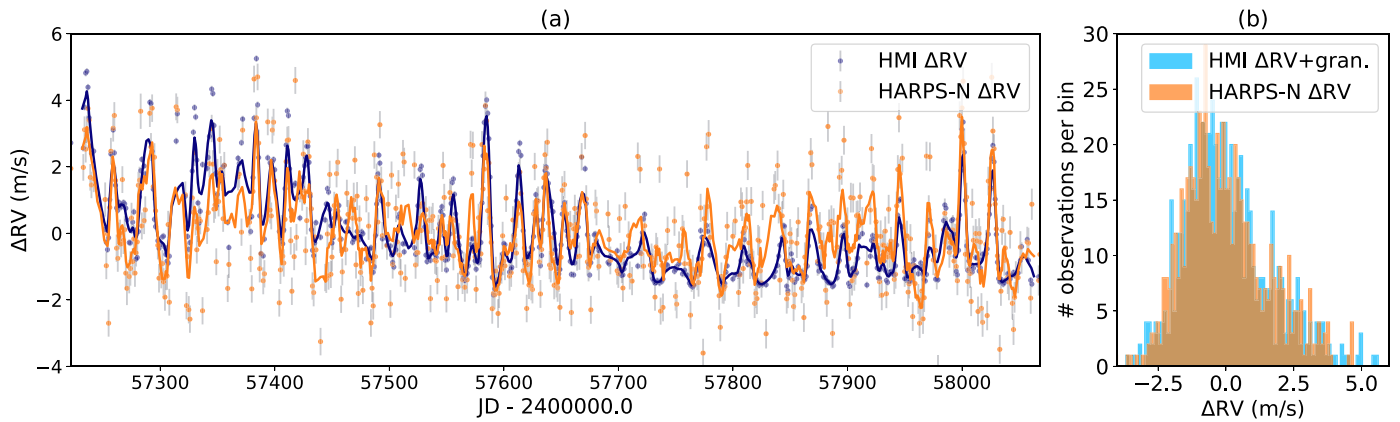
3.1. Separating Magnetically Active Regions from the Quiet Sun

We separate magnetically active regions from quiet-Sun regions by applying a threshold in unsigned radial magnetic field strength for each pixel according to the cutoff found by

---

[11] SORCE TSI data available at: http://lasp.colorado.edu/home/sorce/data/tsi-data/.

[12] Source: https://keplergo.arc.nasa.gov/CalibrationSN.shtml

[13] SolAster (Ervin et al. 2022) is available publicly at: https://tamarervin.github.io/SolAster/.

Figure 3. Panel (a) shows RV time series estimated from HMI images and observed directly by HARPS-N in the interval where they overlap. The lines are 15 day smoothed running averages. In panel (b), we add random Gaussian noise ($\sigma = 0.8$ m s^{-1}) to the HMI RVs to approximate the impact of granulation, for better comparison with the HARPS-N RVs.

Yeo et al. (2013):

$$|B_{r,ij}| > 3\,\sigma_{B_{obs,ij}}/\mu_{ij}, \qquad (1)$$

The factor μ accounts for foreshortening, and is equal to $\cos\theta_{ij}$, where θ_{ij} is the angle between the outward normal to the feature on the solar surface and the direction of the line of sight of the SDO spacecraft. The term $\sigma_{B_{obs,ij}}$ represents the noise in the observed magnetic field, for each pixel at position i, j on the image. Yeo et al. (2013) estimated $\sigma_{B_{obs,ij}}$ to be 8 G (photon dominated), so the magnetic field threshold is 24 G. We exclude *isolated* pixels that are above this threshold as they are likely to be false positives (typically a handful of pixels per image).

3.2. Filling Factors of Sunspots and Plage

To identify faculae and sunspots, we apply the intensity threshold of Yeo et al. (2013), at 0.89 times the mean flattened intensity over quiet-Sun regions. We further identify faculae in concentrated regions of *plage*, as opposed to faculae dispersed in the *network* (see, e.g., Introduction, Section 3), according to the area threshold estimated by Milbourne et al. (2019). We identify plage as magnetically active facular regions whose area on the flattened solar disk exceeds 20 microhemispheres (μhem), corresponding to about 60 Mm^2.

We estimate the disk-averaged filling factors of sunspots and plage as follows:

$$f_{spot, plage} = \frac{1}{N_{pix}} \sum_{ij} W_{ij}, \qquad (2)$$

where N_{pix} is the total number of pixels in the solar disk and the weight W_{ij} is set to 1 in sunspot (or plage) pixels, and 0 in quiet-Sun pixels.

3.3. RV Variations

Milbourne et al. (2019) derived solar RV variations from SDO/HMI images for an 800 day period overlapping disk-integrated RV observations of the Sun with the HARPS-N spectrograph. They reproduced rotation-modulated RV variations in good agreement with the HARPS-N observations, down to an RMS level of 1.21 m s^{-1}, which is consistent with residual motions that are expected from granulation and

supergranulation (Meunier et al. 2015). Their model, which we apply here, accounts for the suppression of convective blueshift from magnetic regions, and the velocity imbalances resulting from brightness inhomogeneities. We refer the reader to the appendix of Milbourne et al. (2019), which fully describes the model. Estimating RV variations according to this technique and model allows us to determine solar RV variations that we can compare directly with spectroscopic measurements of other stars, which are derived from thousands of spectral lines, not just the Fe I line measured by SDO/HMI.

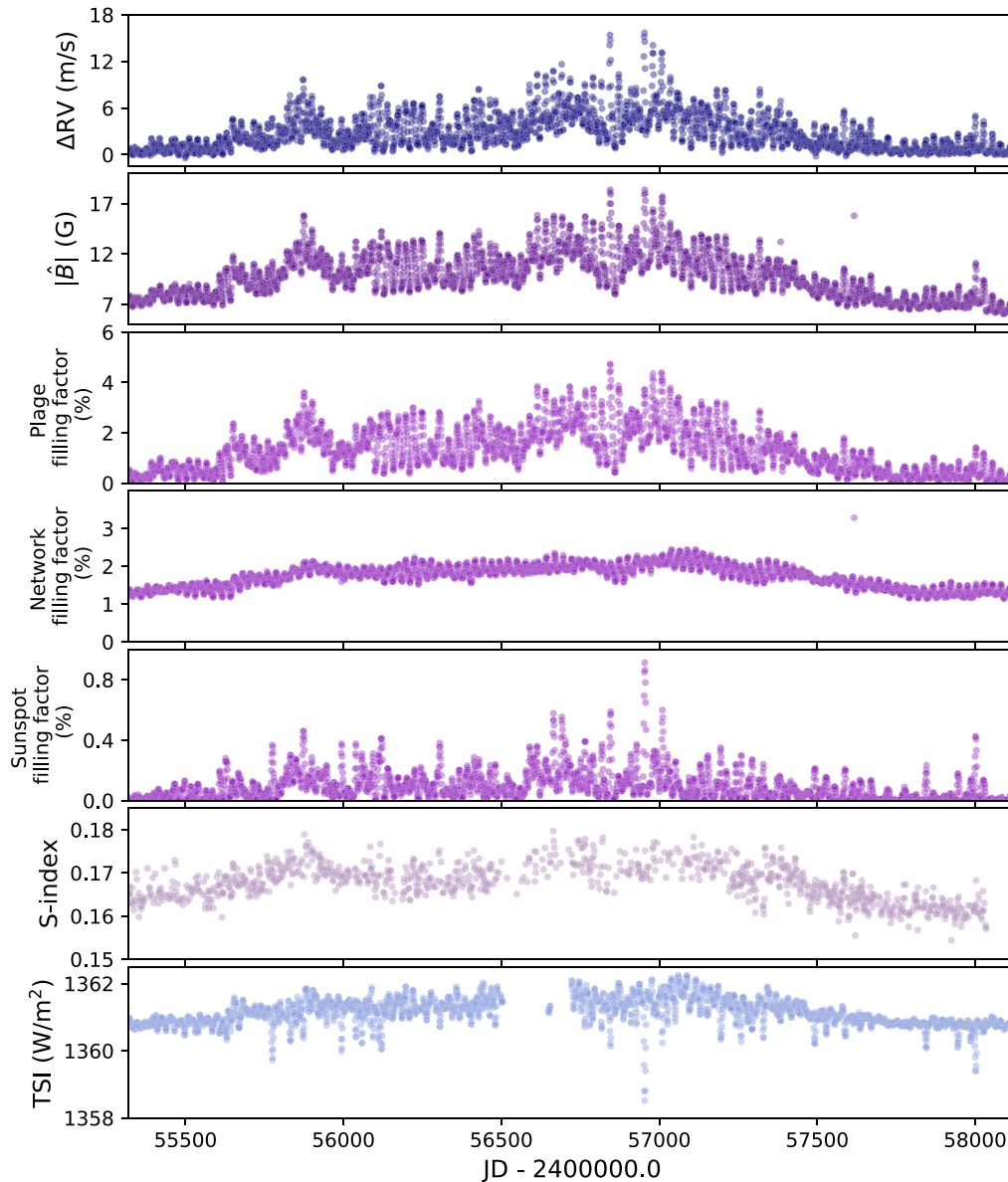### 3.4. Unsigned Magnetic Flux $|\hat{B}_{obs}|$

We compute the disk-averaged, line-of-sight unsigned (i.e., unpolarized) magnetic flux of the Sun, by summing the intensity-weighted line-of-sight absolute magnetic flux in each pixel according to Haywood et al. (2016):

$$|\hat{B}_{obs}| = \frac{\sum_{ij}|B_{obs,ij}|\,I_{ij}}{\sum_{ij}I_{ij}}, \qquad (3)$$

where I_{ij} is the observed, non-flattened continuum intensity of the Sun. We do not flatten the intensity continuum in order to obtain the *observed* unsigned magnetic flux.

3.5. Comparison of ΔRV with Overlapping HARPS-N Observations

To confirm that our method does indeed provide reliable ΔRV estimates at the meter-per-second level, we compare our HMI-derived time series with direct observations of the Sun observed as a star by the high precision, long-term stable spectrograph HARPS-N (Phillips et al. 2016). Milbourne et al. (2019) previously compared RVs from HMI images with direct RV observations by HARPS-N spanning 800 days. We refer the reader to Milbourne et al. (2019) for a detailed analysis. Figure 3 shows RVs from both instruments between 2015 July 18 and 2018 January 12, during which there are 624 days with observations from both SDO and HARPS-N. Both time series are mean subtracted. The HARPS-N RVs are daily averages of 30–40 exposures of 5 minutes each, published by Dumusque et al. (2021). Granulation is treated differently in the two data sets: HARPS-N contains residual granulation signals (averaged over 1 day), whereas in the HMI time series the granulation is subtracted altogether. For this reason, we see uncorrelated

Figure 4. Top to bottom: SDO/HMI full-disk quantities—total RV *variations* (after subtracting quiet-Sun velocity), observed unsigned magnetic flux; filling factors of faculae in concentrated areas of plage, faculae in the diffuse network, and sunspots; Ca II H&K emission S-index from Mount Wilson (Egeland et al. 2017) and HARPS-N (Milbourne et al. 2019); TSI from SORCE.

noise in the HARPS-N RVs that is not present in the HMI RVs. To allow for a clearer comparison, in panel (a) of Figure 3 we overplot 15 day averages of the two RV time series. Both ΔRV time series exhibit very similar variations over the rotational and magnetic cycle timescales. Panel (b) shows the HARPS-N ΔRV compared with the HMI ΔRV inflated by random Gaussian noise to approximate the granulation signal. We use a distribution centered at $0\,\mathrm{m\,s^{-1}}$ with standard deviation $\sigma = 0.8\ \mathrm{m\,s^{-1}}$ in accordance with the residual RMS previously estimated by Meunier et al. (2015). The two time series show a very similar distribution. The Spearman correlation coefficient between the HMI and HARPS-N ΔRV is 0.6. In the histograms for both the HARPS-N and HMI data, we see a tail toward positive ΔRV, arising from two factors. First, the tail is caused by the asymmetric nature of suppression of convective blueshift, whereby active regions lead to a

redshift, i.e., positive ΔRV. As expected, the tail at positive ΔRV and the comparatively steep drop-off at negative ΔRV are slightly more apparent in HMI ΔRV not inflated with white noise. Second, a positive tail also arises from the long-term magnetic cycle that creates an initial downward slope and is flat after JD ≈ 2457600.

In summary, we see good agreement between HMI and HARPS-N RV variations (in agreement with previous studies by Milbourne et al. 2019; Haywood et al. 2016), which gives us confidence that our method to estimate RV variations from HMI images is indeed reliable. We do not see large outliers that could induce spatial systematics (e.g., those identified by Löhner-Böttcher & Schlichenmaier 2013). We note that we do not expect the two data sets to be completely identical since we know that we are missing some physics in our model to estimate HMI RVs (see, e.g., Section 6).
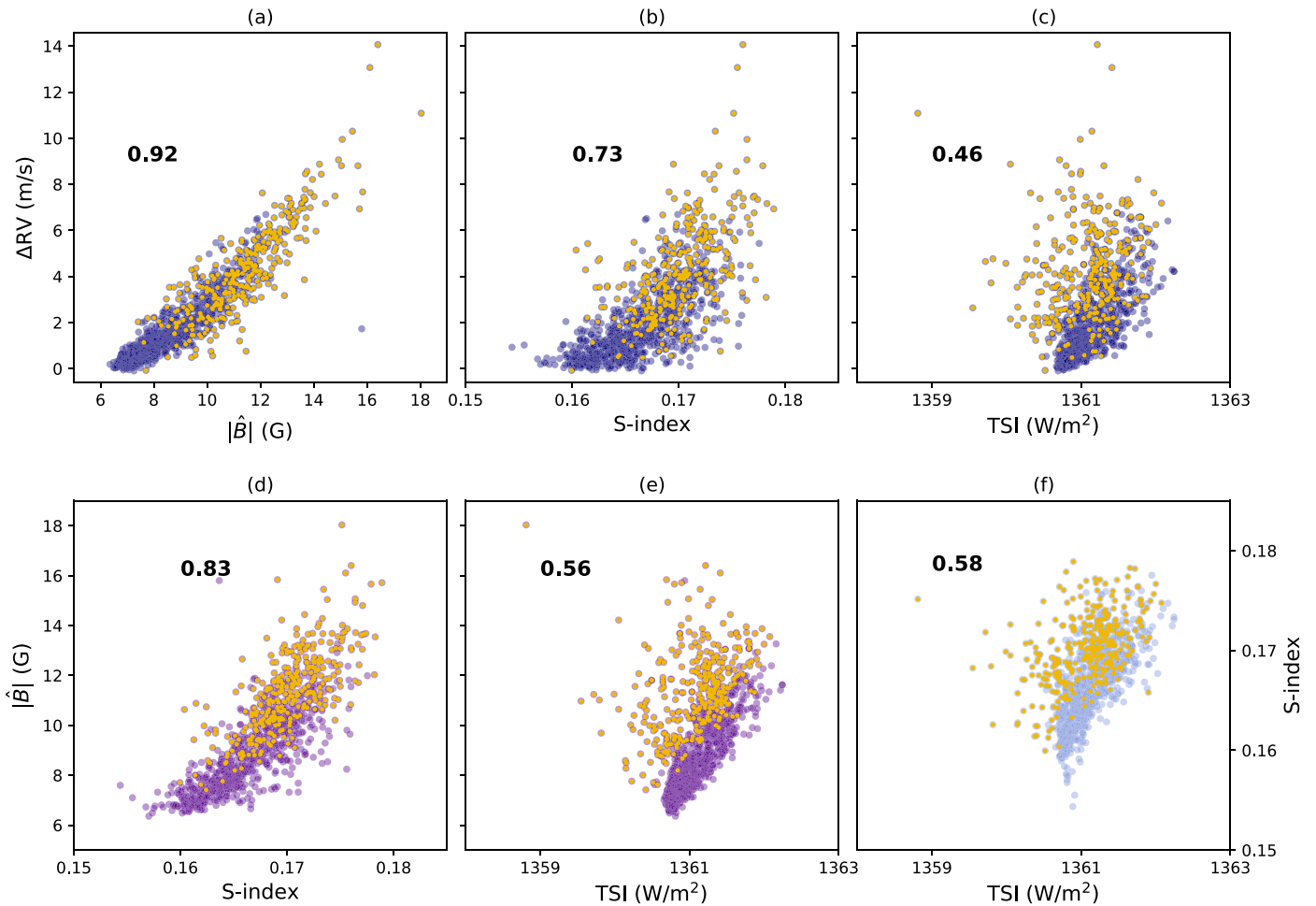
Figure 5. Top: SDO/HMI-derived RV variations plotted as a function of (a) unsigned magnetic flux $|\hat{B}_{\rm obs}|$, (b) S-index and (c) TSI. Bottom: unsigned magnetic flux as a function of (d) S-index and (e) TSI; and (f): S-index and TSI against one another. Observations highlighted in yellow have a *high* sunspot filling factor ($f_{\rm spot} > 0.15\%$). Spearman correlation coefficients are given for each pair of correlates.

3.6. Correlations between ΔRV, $|\hat{B}_{\rm obs}|$, Photometry, and S-index

The time series of RV variations, unsigned magnetic flux, plage, and sunspot filling factors are plotted alongside coeval TSI and S-index observations in Figure 4. The SDO/HMI disk-averaged quantities shown in Figure 4 are at the maximum cadence considered in this study (1 observation every 4 hr, i.e., 6 per day). We then average the SDO/HMI quantities over daily bins and concatenate this data set with the time series of S-index and TSI, resulting in time series of 1036 points each. We show the RV variations as a function of unsigned magnetic flux, S-index, and TSI in the top row of Figure 5. The bottom row shows the three activity indicators plotted as a function of each other. Figure 5 shows that the RV variations correlate much better with the unsigned magnetic flux ($R = 0.92$) than the S-index ($R = 0.73$) or optical photometry ($R = 0.46$). Observations of high sunspot coverage are highlighted in yellow. We see that the Sun is dominated by spots at high activity levels. At the peak of the activity cycle, the Sun's photometric variations are anticorrelated with Ca II H&K variations (Radick et al. 2018, Figure 1), so the Sun is spot dominated; at lower activity levels, they are positively correlated, implying that the solar surface is dominated by faculae/plage. This is in agreement with previous studies (Fröhlich & Lean 1998; Krivova et al. 2007; Shapiro et al. 2014).

Following Meunier et al. (2019) who reported a hysteresis between ΔRV and Ca II H&K in solar Cycle 23, we investigate the hystereses between ΔRV, S-index, and $|\hat{B}_{\rm obs}|$ in Cycle 24. We observe hystereses between all quantities (shown in Figure A1) and discuss their physical origins in detail in Appendix A.

Time lags of 1–3 days between RV variations and the bisector span and FWHM have been reported previously in spectroscopic HARPS-N observations of the Sun (Collier Cameron et al. 2019, Figure 15). We cross correlate ΔRV, $|\hat{B}_{\rm obs}|$, and the S-index against each other to look for time shifts between them. We do not find any significant time shifts between any of our observables.

3.7. Periodogram Analysis

Figure 6 shows generalized Lomb–Scargle periodograms (Zechmeister & Kürster 2009) of the RV variations (panel (a)), the filling factors of plage (panel (b)) and sunspots (panel (c)) and the unsigned magnetic flux (panel (d)). Most of the periodicity below 100 days is confined to periods close to the rotation period and its first harmonic. It is worth noting that these peaks are in fact forests of peaks, in which several peaks are significant above the 0.001% confidence level. This means that depending on when or for how long we might observe the Sun, we may measure rotation periods differing by several days (e.g., Mortier & Collier Cameron 2017; Nava et al. 2020). In
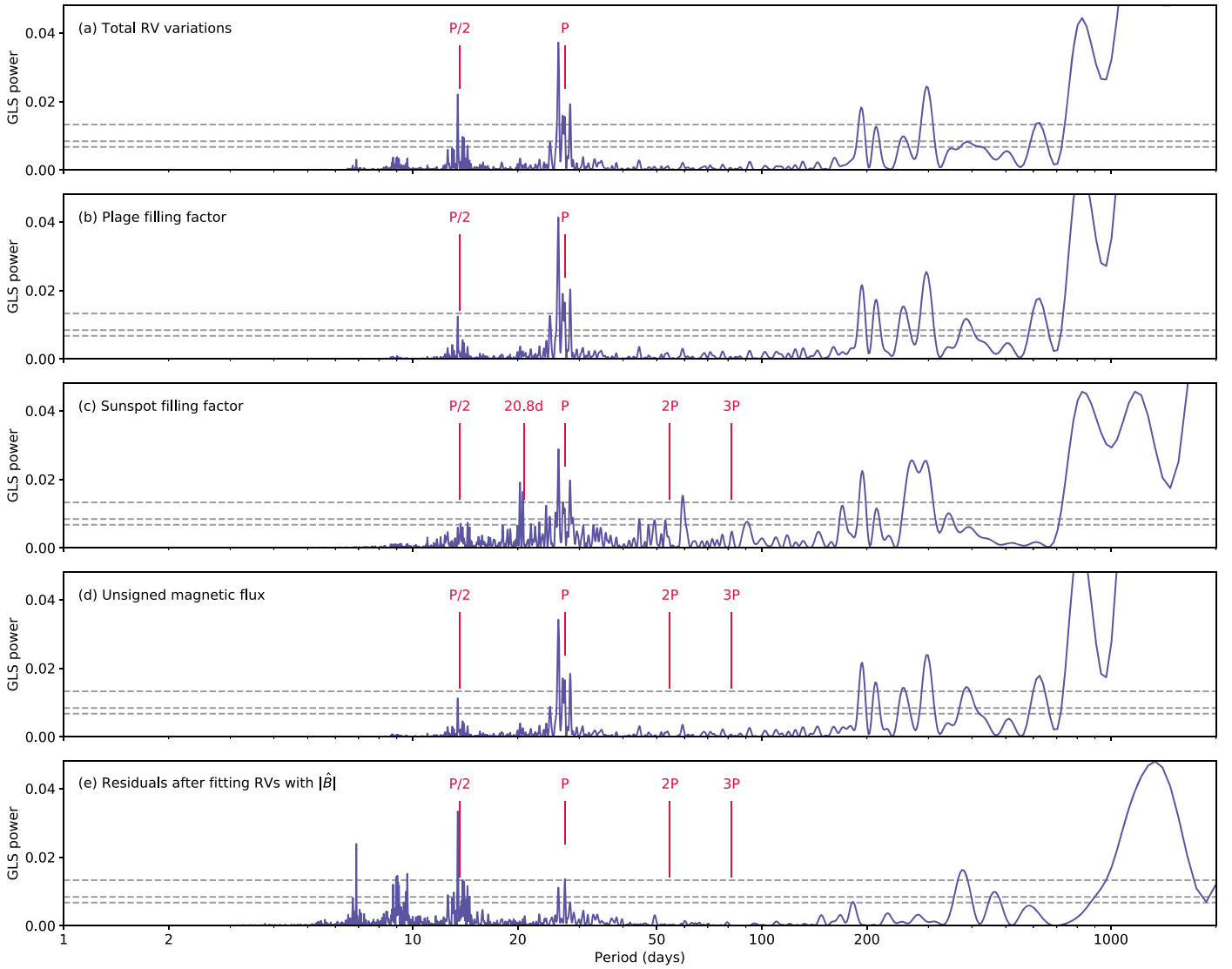
Figure 6. Generalized Lomb–Scargle periodograms of (a) the total SDO/HMI-derived RV variations, (b) the filling factor of plage, (c) the filling factor of sunspots, (d) the unsigned magnetic flux $|\hat{B}_{\rm obs}|$, and (e) residuals after modeling ΔRV with a linear fit of $|\hat{B}_{\rm obs}|$. The horizontal gray dashed lines represent the false alarm probability levels (from bottom to top: 10%, 1%, and 0.001%).

the periodogram of the sunspot filling factor, we also detect a significant peak consistent with the 20.8 day peak detected by Lagrange et al. (2010). This peak is possibly related to the lifetime of the spots. Alternatively, it may be associated with global-scale equatorial Rossby waves (r-modes) that produce oscillations on a 19 day recurrence timescale (Lanza et al. 2019). In this periodogram, we see many significant peaks around the rotation period and its harmonics. Additionally, some peaks are significantly different from the rotation period or its harmonics; as Lagrange et al. (2010) previously emphasized, we should be careful when attributing these signals to non-activity processes. The strong signals in the spot filling factor do not necessarily translate into signals in the total RV variations, because the Sun is faculae-dominated for the majority of its cycle (e.g., Fröhlich & Lean 1998; Krivova et al. 2007; Shapiro et al. 2014). However, in younger, faster rotating Sun-like stars whose behavior has been observed to be spot dominated (see Lockwood et al. 2007; Radick et al. 2018), we would certainly expect the RV variations to show a more *spot-like* periodogram structure.

4. Modeling Δ RV Using $|\hat{B}_{\rm obs}|$

We model the RV variations estimated in Section 3, $\Delta RV(t)$ as a linear model of $|\hat{B}_{\rm obs}|$:

$$\Delta RV(t) = \alpha \frac{|\hat{B}_{\rm obs}|(t)}{\langle |\hat{B}_{\rm obs}| \rangle} + \Delta RV_0, \qquad (4)$$

where α is a constant scaling factor and ΔRV_0 is a constant zero-point offset. $\langle |\hat{B}_{\rm obs}| \rangle$ is the mean of $|\hat{B}_{\rm obs}|$ over the full time series. We optimize the parameters α and ΔRV_0 via a least-squares procedure. We model the full $\Delta RV(t)$ time series of daily averages (plotted in the top panel of Figure 4). The fit over the full magnetic cycle is shown in the top panel of Figure 7. To examine the performance of the linear $|\hat{B}_{\rm obs}|$ model as a function of magnetic activity levels, we model two separate stretches, at activity maximum and minimum, spanning 600 days each. The low activity, quiet epoch ranges from 2016 May 20 (JD = 2457529) to 2018 January 10 (JD = 2458129). The high activity epoch spans 600 days from 2014 March 1 (JD = 2456718) to
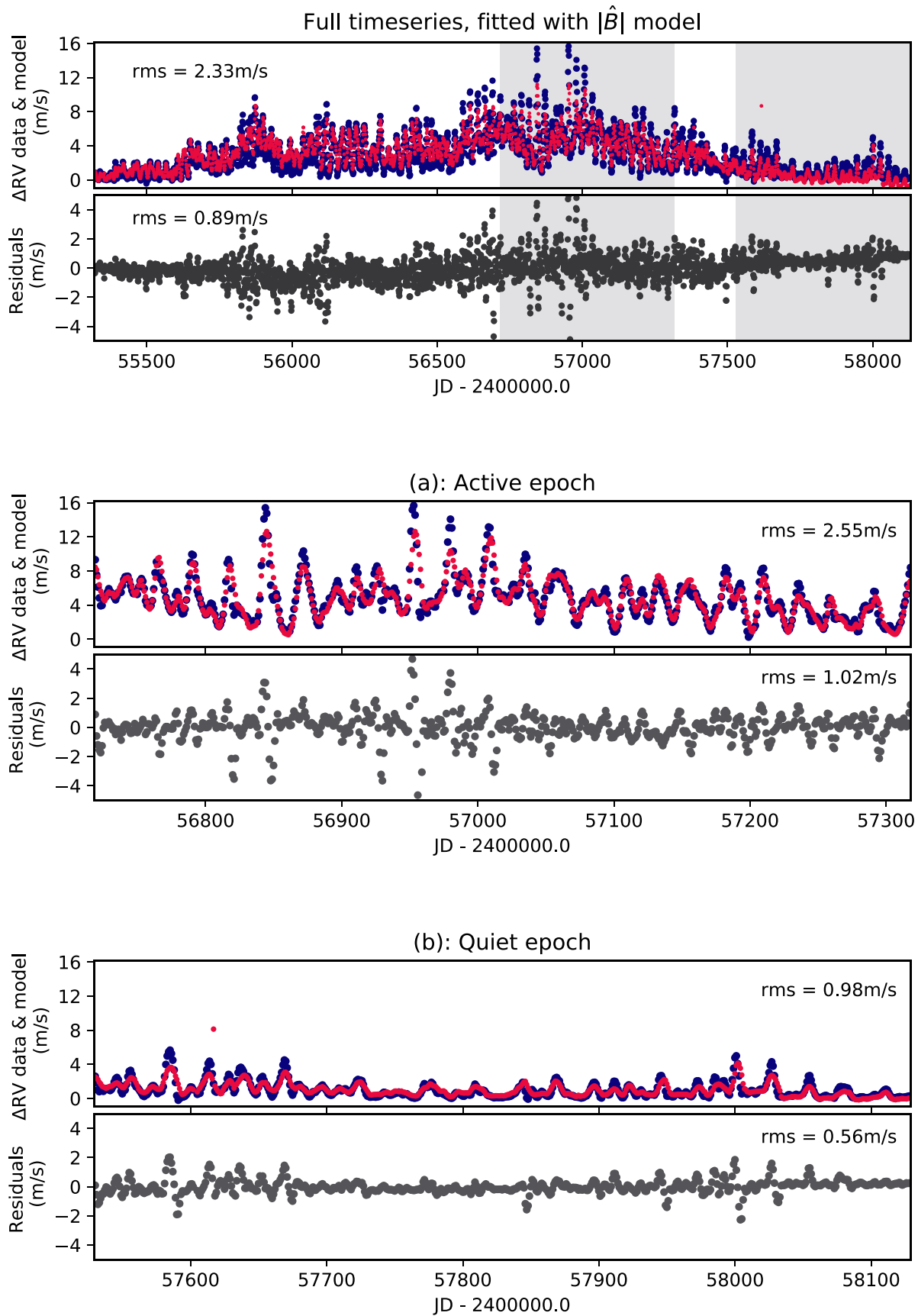
Figure 7. Top panel: estimated ΔRV (in blue) for the full (non-concatenated) SDO/HMI-RV data set, modeled with a linear fit in unsigned magnetic flux (red); residuals of the fit (gray). Panel (a): same as top, but for the 600 day epoch of high magnetic activity. Panel (b): same as top, but for the quiet 600 day epoch of low magnetic activity.

Table 2
Best-fit Parameter Estimates and RMS Values for Each of the Models Tested in Sections 4 and 5

Model	Span of Data Modeled	Parameter Estimates (m s^{-1})			RMS Values (m s^{-1})		% Reduction in RV RMS Amplitude
		α	β	RV$_0$	Data	Residuals	
$|\hat{B}_{\rm obs}|$ model	Full time series	9.62 ± 0.01	$\cdots$	-6.78 ± 0.01	2.33	0.89	62%
$|\hat{B}_{\rm obs}|$ model	Quiet epoch	8.38 ± 0.04	$\cdots$	-5.32 ± 0.03	0.98	0.56	43%
$|\hat{B}_{\rm obs}|$ model	Active epoch	11.49 ± 0.02	$\cdots$	-8.81 ± 0.02	2.55	1.02	60%
$|\hat{B}_{\rm obs}|$ with FF' model	Full time series	9.36 ± 0.01	-2.16 ± 0.02	-6.68 ± 0.01	2.27	0.85	63%
$|\hat{B}_{\rm obs}|$ with FF' model	Quiet epoch	8.24 ± 0.04	-1.97 ± 0.04	-5.46 ± 0.03	0.98	0.51	47%
$|\hat{B}_{\rm obs}|$ with FF' model	Active epoch	11.31 ± 0.02	-2.70 ± 0.04	-8.82 ± 0.02	2.55	0.97	62%

Note. The quiet and active epochs span 600 days each. Note that the $|\hat{B}_{\rm obs}| + FF'$ model can only be run on days for which both SDO/HMI and SORCE TSI data are available (hence the slightly different data RMS estimates).

2015 October 22 (JD = 2457318). The *active* and *quiet* fits are plotted in panels (a) and (b) of Figure 7. The estimated parameters for all fits are reported in Table 2. The RMS of the full data set is 2.33 m s^{-1}, and that of the residuals is 0.89 m s^{-1}. Overall, a simple $|\hat{B}_{\rm obs}|$ model reduces the RMS of the RV variations by 62%, i.e., a factor of 2.6. Although we do see correlated residuals at times of high activity, the residuals over the full cycle are flat.

Rotationally modulated RV variations at times of low activity —The dominant process at play is the suppression of convective blueshift incurred by areas of plage (Milbourne et al. 2019). $|\hat{B}_{\rm obs}|$ correlates extremely well with their presence, as can be seen in Figure B1 in Appendix B. There are very few spots; the maximum filling factor of sunspots in this 600 day stretch is 0.03% (compared with 0.14% in the active stretch and 0.09% overall). We therefore expect $|\hat{B}_{\rm obs}|$ to correlate well with RV variations. Indeed, this model is an excellent fit during activity minimum, as evidenced in Figure 7(b). Over the quiet epoch, the RV residuals have an RMS of 0.56 m s^{-1}.

Rotationally modulated RV variations at times of high activity— The best-fit estimates of the model parameters (α and ΔRV_0) differ significantly from the fit at low magnetic activity (see Table 2). This is because convective blueshift is more suppressed by larger magnetic structures (which are more prevalent in periods of high activity), i.e., the faculae in concentrated areas of plage, as previously found by Meunier et al. (2010b, Figure 7) and Milbourne et al. (2019, Figure 6). The $|\hat{B}_{\rm obs}|$ model accounts for RV variations down to a residual RMS of 1.02 m s^{-1}. As can be seen in Figure 7(a), significant rotationally modulated RV variations remain unaccounted for. In Figure 8(a), we zoom-in further on the RV variations over 4–5 solar rotations at activity maximum. This stretch spans 120 days from 2014 October 7 (JD = 2456938) to 2015 February 4 (JD = 2457058). The smooth, rotationally modulated signal that remains in the RV residuals is very clear on this timescale. At activity maximum, we expect suppression of convective blueshift to produce significant RV variations (Meunier et al. 2010a, 2010b; Haywood et al. 2016). Additionally, there are sunspots, which now produce significant RV variations by blocking Doppler-shifted flux on the rotating solar surface (Saar & Donahue 1997; Lagrange et al. 2010; Meunier et al. 2010b; Haywood et al. 2016). The relationship between this sunspot flux-blocking RV term and $|\hat{B}_{\rm obs}|$ is more complex than for the RV due to suppression of convective blueshift. In fact, they do not correlate with each other ($R = 0.08$). When a spot crosses the central meridian, this RV contribution is zero, while $|\hat{B}_{\rm obs}|$ would be at its maximum. Therefore, a simple, linear $|\hat{B}_{\rm obs}|$ model cannot adequately capture RV variations from sunspot flux blocking.

To investigate this hypothesis, we apply the technique developed by Aigrain et al. (2012). Their FF' term accounts for RV variations incurred by brightness inhomogeneities on a rotating disk.

5. Modeling ΔRV Using Both $|\hat{B}_{\rm obs}|$ and the FF' Model

A simple $|\hat{B}_{\rm obs}|$ model fits RV variations well overall, but becomes insufficient at the peak of the solar magnetic cycle (see Section 4), where significant, correlated residuals remain, as visible in Figure 7(a). To account for RV variations from flux blocking from sunspots and plage, we apply the FF' technique of Aigrain et al. (2012, Equation (10)) that models the RV perturbation due to a spot crossing the disk. Since both dark inhomogeneities (from spots) and bright inhomogeneities (from faculae, mainly in plage) produce RV perturbations (e.g., Meunier et al. 2010a, Figure 7), we use the total magnetic filling factor $f_{\rm all}$ rather than $f_{\rm spot}$. We use the TSI as F (Section 2.2). We compute its first time derivative F' by interpolating the TSI observations onto an evenly, over-sampled array and then fit them using Gaussian-process regression using a basic square exponential kernel (Rasmussen & Williams 2006). We then compute the derivative using second-order accurate central differences. We scale this basis function with a constant β that we fit for in our least-squares optimization, along with α and ΔRV_0. Our resulting basis function has an RMS amplitude of 0.2 m s^{-1}. The overall fits with $|\hat{B}_{\rm obs}| + FF'$ are very similar to those resulting from modeling the RVs with $|\hat{B}_{\rm obs}|$ only (as is done in Section 4). The residuals still display correlated behavior. We conclude that the FF' model does not fully account for the signals leftover from the $|\hat{B}_{\rm obs}|$ model.

6. Rotationally Modulated RV Variations Not Traced by $|\hat{B}_{\rm obs}|$ or FF'

The residuals from both models tested ($|\hat{B}_{\rm obs}|$ and $|\hat{B}_{\rm obs}|+FF'$) display correlated behavior during times of high magnetic activity. We propose two explanations.

First, the FF' term does not adequately fit the RV variations resulting from sunspot flux blocking (Section 5). Several studies have previously found that the FF' method cannot fully account for RV variations (e.g., Oshagh et al. 2017; Bastien et al. 2014; Haywood et al. 2014). We note that the FF' term (or the F^2 term that can be used to account for suppression of convective blueshift) is not *expected* to match RV variations perfectly because F' includes the derivative of the limb darkening of F, which should
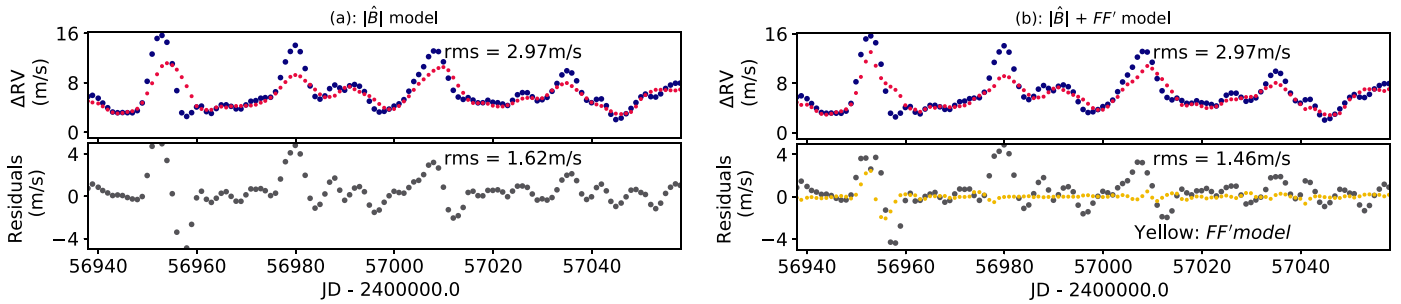
Figure 8. Zoom-in on ΔRV over 4–5 solar rotations (120 days) at activity maximum, to highlight the smooth, rotationally modulated signals in both the observations (blue) and the residuals (gray) after fitting the model (red). Panel (a): $|\hat{B}_{obs}|$ model of Section 4. Panel (b): $|\hat{B}_{obs}| + FF'$ model of Section 5; the FF' term is shown (yellow) in the residuals panel.

not be a part of the RV model. We demonstrate this in detail in Appendix C.

Another possible explanation for the RV residuals is that there are additional processes at play, which are either missing from the RV model of Milbourne et al. (2019) and Haywood et al. (2016) used to estimate RV variations, or that do not correlate directly with $|\hat{B}_{obs}|$ (or FF'). This finding is consistent with that of Miklos et al. (2020), who investigated the activity sensitivity of spectral lines observed by HARPS-N and concluded that there must be additional factors, not yet accounted for by current state-of-the-art models such as in Meunier et al. (2017). Other types of surface velocity fields not included in our model may give rise to rotationally modulated RV variations, such as the following:

Evershed flows—Sunspots are made of umbral and penumbral regions. Evershed flows, which are contained within penumbral regions and flow radially outward from the central umbra to the outer edge of the sunspot, are tangential to the surface (Evershed 1909). They will be most visible in RV for sunspots located away from the disk center, where their flows are more directed along our line of sight (e.g., Haywood et al. 2016, Figure 6).

Moat flows—Outside of the penumbra, but within the active region, are so-called moat flows (e.g., Solanki 2003 and references therein). These are also tangential to the surface but weaker than penumbral flows (of order 1/4 to 1/2 as strong), but as they are brighter (since they are typically in plage) and normally cover an area $\approx 3\times$ larger than the spot. They may contribute significantly to the total RV signal (e.g., Iampietro et al. 2019).

Active region inflows—Gizon et al. (2001) report the presence of inflows toward active regions on the Sun's surface, with amplitudes up to $50\,\mathrm{m\,s}^{-1}$ (see Gizon et al. 2010, Section 7.1.).

Unresolved flows—We could be seeing residual effects arising from unresolved flow motions and magnetic processes taking place within magnetically active SDO pixels. HMI samples the Fe I line profile (at 6173 Å) at only six points in wavelength (see, e.g., Section 2.1). The line shift (velocity), depth, width, magnetic field, and continuum intensity are then determined by fitting a symmetric Gaussian to these points. This coarse sampling at the pixel level, and the fact that pixels are of a size comparable to that of granules mean that we are missing spectral-line asymmetries from processes taking place below the pixel resolution, e.g., due to convection. Saar (2009) makes some initial attempts (expanding on Saar 2003) at including convective line asymmetries in plage models.

Zeeman broadening—A final item not included in our model is the direct effect of magnetic fields on the line profiles themselves (Reiners et al. 2013). Due to Zeeman broadening in magnetic regions, the line profiles originating there are wider and shaped differently. If the lines are stronger, they can show enhanced equivalent widths as well. These differences lead to subtle RV changes as active regions rotate and change in number and size. Reiners et al. (2013) showed that RV amplitudes resulting from Zeeman broadening were of the order $\Delta v_B \approx 300\,f\,(B/1\,\mathrm{kG}\,)^2(\lambda/1\mu\,\mathrm{m})^2\,\mathrm{m\,s}^{-1}$, where f is the magnetic filling factor, B is the local magnetic field strength, and λ is the wavelength. Note that HMI measures fluxes (i.e., B times area) and the actual B in resolved plage flux tubes is ~ 1.5 kG (e.g., Buehler et al. 2015). Most plage contains a mix of flux tubes and field-free areas. Thus, the true solar magnetic f is close to 1%–2% (note that we are ignoring the ubiquitous weak turbulent fields in this estimate). Adopting $\langle\lambda\rangle = 0.5\,\mu$m, we can estimate $\Delta v_B \approx 1.7$–$3.4\,\mathrm{m\,s}^{-1}$. Note that Zeeman broadening is only partly and imperfectly removed by the fit to B since the actual RV dependence is $\propto f\lambda^2 B^2$. Since proper treatment would require a different calculation of the filling factors, and due to the wavelength dependence, a different computation of RVs, we leave exploring this to a future paper. We note, however, that the Zeeman broadening effect would follow B in phase, and could reduce residuals coincident with large B concentrations during active epochs.

7. Using $|\hat{B}_{obs}|$ to Confirm and Characterize Long-period, Low-mass Planets

Previous studies have shown that we *must* account for RV variability in order to detect and characterize low-mass, long-period planets (e.g., Saar 2009; Hall et al. 2018; Meunier & Lagrange 2019, and others). Here, we test whether the unsigned magnetic flux could, in principle, be used to mitigate rotationally modulated RV variations to characterize small planets accurately and precisely (e.g., to better than 10% precision in mass; Zeng & Sasselov 2013). The present analysis is not intended to be a comprehensive exploration of parameter space, nor is it meant to reflect realistic ground-based observing conditions for stellar RV surveys. When facing reality, the most important factor to consider will be the precision to which $|\hat{B}_{obs}|$ can be measured; we discuss this in Section 8.1. The effects of magnetoconvection, i.e., granulation and supergranulation will also need to be accounted for as they produce RV variability at the meter-per-second level (e.g., Meunier et al. 2015).

Table 3
Best-fit Parameter Estimates for Each of the Planet Injection Tests Made in Section 7

Activity Model	Injected Planet Parameters			Best-fit Parameter Estimates and 1σ Uncertainties					
	K (m s^{-1})	$P_{\rm orb}$ (days)	t_0 (JD)	α (m s^{-1})	β (m s^{-1})	ΔRV_0 (m s^{-1})	K (m s^{-1})	$P_{\rm orb}$ (days)	t_0 (days)
None	2.0	300.38	56819.92	⋯	⋯	2.79 ± 0.03	1.87 ± 0.04	301.6 ± 0.4	$56{,}831 \pm 1$
None	1.5	300.38	56819.92	⋯	⋯	2.78 ± 0.03	1.39 ± 0.04	305.5 ± 0.8	$56{,}846 \pm 3$
None	1.0	300.38	56819.92	⋯	⋯	2.76 ± 0.03	1.22 ± 0.04	318.8 ± 0.6	$56{,}897 \pm 2$
None	0.5	300.38	56819.92	⋯	⋯	2.75 ± 0.02	1.20 ± 0.03	320.9 ± 0.5	$56{,}911 \pm 2$
$\lvert \hat{B}_{\rm obs} \rvert$	0.5	300.38	56819.92	9.5 ± 0.1	⋯	-6.8 ± 0.1	0.48 ± 0.03	297 ± 1	$56{,}830 \pm 3$
$\lvert \hat{B}_{\rm obs} \rvert + FF'$	0.5	300.38	56819.92	9.5 ± 0.1	-2.5 ± 0.2	-6.8 ± 0.1	0.47 ± 0.03	297 ± 1	$56{,}833 \pm 3$
$\lvert \hat{B}_{\rm obs} \rvert$	0.3	300.38	56819.92	9.6 ± 0.1	⋯	-6.8 ± 0.1	0.29 ± 0.03	296 ± 1	$56{,}838 \pm 5$
$\lvert \hat{B}_{\rm obs} \rvert + FF'$	0.3	300.38	56819.92	9.5 ± 0.1	-2.5 ± 0.2	-6.8 ± 0.1	0.29 ± 0.03	295 ± 2	$56{,}844 \pm 6$
$\lvert \hat{B}_{\rm obs} \rvert$	0.1	300.38	56819.92	9.8 ± 0.1	⋯	-7.0 ± 0.1	0.32 ± 0.04	335 ± 2	$56{,}715 \pm 6$
$\lvert \hat{B}_{\rm obs} \rvert + FF'$	0.1	300.38	56819.92	9.7 ± 0.1	-2.4 ± 0.3	-7.0 ± 0.1	$0.24^{+0.04}_{-0.06}$	334^{+3}_{-42}	$56{,}712^{+152}_{-8}$

Note. Note that 0.5 and 0.3 m s^{-1} injected planets are recovered while the 0.1 m s^{-1} injected planet is not recovered (see Section 7 for details). The times of transit t_0 have 2400000 JD subtracted from them.
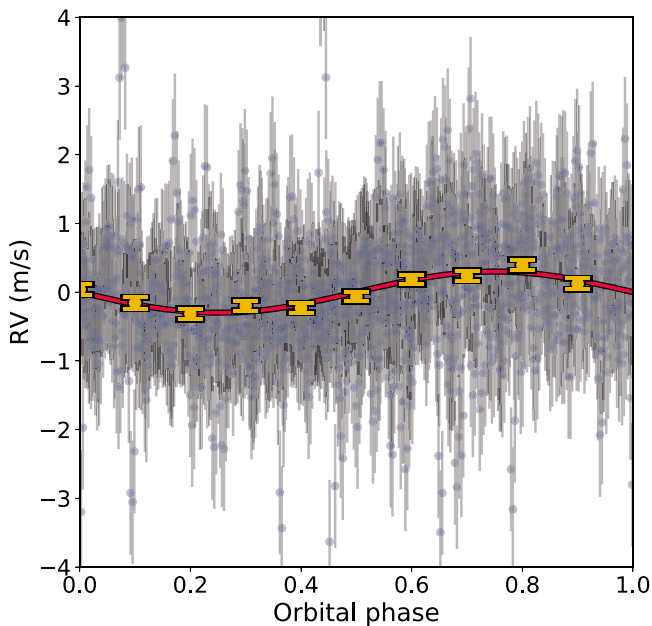
Figure 9. Phase plot of the orbit of an injected planet with $K = 0.3$ m s^{-1}, at $P \simeq 300$ days (red line) recovered when modeling the RV observations with a linear fit of $\lvert \hat{B}_{\rm obs} \rvert$. Blue points: RV observations after subtracting the $\lvert \hat{B}_{\rm obs} \rvert$ model and zero-point RV offset. Yellow points: inverse-variance weighted averages of the blue points in bins of size 0.1 in phase. The fits work well down to $K = 0.3$ m s^{-1} but fail at 0.1 m s^{-1} (see Section 7 for details).

7.1. Procedure

Here, we inject synthetic planet signals in the 8 yr long, daily averaged SDO/HMI-derived ΔRV time series (see Figure 4), down-sampled to 6 months per year, to simulate the visibility pattern of a star observed from the ground (e.g., Hall et al. 2018, Section 3.3). For simplicity, we do not remove observations to mimic weather losses, as Hall et al. (2018) estimate that targets visible during the summer only experience a 6% loss of nights due to weather at optimal observing sites. We are left with 1456 daily observations spread over 8 consecutive seasons. We consider scenarios both with and without modeling solar activity. In all cases, we consider a planet with a circular orbit, with period $P_{\rm orb} = 300.38$ days and

time of transit $t_0 = 2456819.92$ JD (close to the mid-time of the data set). We choose $P_{\rm orb}$ close to but not exactly at 300 days to avoid overlap with any potential aliases arising from SDO's geosynchronous orbit. We consider three RV semi-amplitudes K: 0.5, 0.3, and 0.1 m s^{-1}. Assuming an orbital inclination of $90°$ and a $1 M_\odot$ host star with a typical uncertainty for a bright, solar analog of $0.03 M_\odot$, an RV semi-amplitude of 0.5 m s^{-1} corresponds to a planet with a mass of $5.2 M_\oplus$, 0.3 m s^{-1} corresponds to $3.1 M_\oplus$, and 0.1 m s^{-1} corresponds to $1.04 M_\oplus$. We also consider larger semi-amplitudes K: 2.0, 1.5, and 1.0 m s^{-1} for simulations with no activity modeling.

For each scenario, we fit models consisting of a circular Keplerian and zero-point offset, and either the linear function of $\lvert \hat{B}_{\rm obs} \rvert$ (of Section 4), or the $\lvert \hat{B}_{\rm obs} \rvert + FF'$ combination (of Section 5), or no activity modeling. We account for the remaining residuals by adding the residual RMS values (0.89 and 0.85 m s^{-1} for the $\lvert \hat{B}_{\rm obs} \rvert$ and $\lvert \hat{B}_{\rm obs} \rvert + FF'$ models, respectively; see Table 2) in quadrature to the 1σ RV uncertainties (0.1 m s^{-1}, see Table 1). Rounding up, we obtain an effective RV uncertainty of 0.9 m s^{-1} for both models. In a real-case scenario, one would implement a correlated noise framework (e.g., Gaussian-process regression) to ensure that the parameter estimates are as accurate and precise as they can be in the presence of correlated noise. This statistically demanding analysis is beyond the scope of the present analysis, whose primary purpose is to determine whether the planets can be recovered. We assume prior indication of a planet in this range of orbits, e.g., through the detection of one or more transits, so we impose broad Gaussian priors of 10 and 20 days on $P_{\rm orb}$ and t_0, respectively. We maximize the likelihood of each model and determine the best-fit parameter values through a Markov chain Monte Carlo (MCMC) procedure similar to the one described in Haywood et al. 2014, in an affine-invariant framework (Goodman & Weare 2010). In all cases, the MCMCs reveal a parameter space with multiple local maxima in likelihood. As is done routinely in exoplanet analyses, if the majority of the MCMC chains give the same solution, we remove deviant chains before estimating the model parameters and uncertainties. However, if no single area of parameter space is clearly preferred, we deem the MCMC outcome as a non-detection.

7.2. Outcomes

The results of all scenarios are presented in Table 3. In all cases where solar activity has been taken into account, the estimates for α, β, and ΔRV_0 match those determined via the least-squares optimization procedures of Sections 4 and 5 within 1σ–2σ (see Table 2).

No activity modeling—We injected and tried to recover planets with relatively large RV semi-amplitudes (injected $K = 2$, 1.5, and $1 \, \mathrm{m \, s^{-1}}$) without modeling solar activity. We retrieved injected planets with $K = 2 \, \mathrm{m \, s^{-1}}$ with K and P_orb parameters accurate to about 3σ from the injected value. The phase (given by t_0) is off from the injected value by 11 days, well outside of the 1σ range of 1 day. For injected planets with $K = 1.5 \, \mathrm{m \, s^{-1}}$, we recover the amplitude K within 3σ from the injected value, but not the orbital period, which does still lie within $\approx 6\sigma$ of the true value. We do not retrieve accurate parameters for injected planets with $K = 1$ and $0.5 \, \mathrm{m \, s^{-1}}$. In both cases, the MCMC settles on a signal at 320 days with an amplitude of $1.2 \, \mathrm{m \, s^{-1}}$. Looking at the periodograms in Figure 6(a), we see a strong, broad peak around 300–330 days in the RVs, which disappears after fitting with $|\hat{B}_\mathrm{obs}|$ (Figure 6(e)). We conclude that the MCMC is settling for the nearest significant signal in the RVs, as the planet signal does not produce a strong enough local maximum in likelihood space. Modeling activity-induced RV variations is thus necessary to retrieve planets with amplitudes $K = 1 \, \mathrm{m \, s^{-1}}$ and below.

Injected planets with $K = 0.5$ and $0.3 \, \mathrm{m \, s^{-1}}$—Fitting solar activity with a linear $|\hat{B}_\mathrm{obs}|$ term, we recover the planet signal for both of these K amplitudes. The RV amplitudes are estimated accurately within 1σ. The orbital periods are systematically underestimated by up to 3σ, as was found by (Hall et al. 2018, Figure 5) who performed very similar simulations. The orbital phases t_0, too, are off from their correct value by up to 3σ, due to P_orb being offset. The planet amplitude is recovered with the same significance using either the $|\hat{B}_\mathrm{obs}|$ model or the $|\hat{B}_\mathrm{obs}|$ and FF' combination. We show the phase-folded orbit of the injected $K = 0.3 \, \mathrm{m \, s^{-1}}$, recovered using the $|\hat{B}_\mathrm{obs}|$ model in Figure 9.

Injected planet $K = 0.1 \, \mathrm{m \, s^{-1}}$—Neither activity models are sufficient to recover signals of this amplitude. The parameter space explored by the MCMC chains shows multiple solutions with similar likelihoods at the 300 day period of the injected planet, as well as at 330 days and near 400 days. When we inject a planet at 300 days with $K = 0.1 \, \mathrm{m \, s^{-1}}$, we systematically find that the most likely solution is for a signal with a period of 335 days with an RV amplitude of $0.3 \, \mathrm{m \, s^{-1}}$. The 335 day signal does not appear to be caused by uneven sampling, as we see complete and uniform coverage at all phases.

Stellar activity signals in the 300 day range—We see several significant peaks at 200–400 days in the periodogram of the RVs (Figure 6(a)). $|\hat{B}_\mathrm{obs}|$, too, exhibits significant and similar (but nonidentical) periodicities in the same range (Figure 6(d)). Panel (e) shows that the RV residuals (after applying the $|\hat{B}_\mathrm{obs}|$ model) exhibit comparatively less power in this period range, but we still see two significant peaks in the 350–500 day range. Meunier et al. (2010b, Figure A1) also detect significant peaks at 300–400 days in solar RV variations of Cycle 23. Their RV variations are estimated using an independent method, using catalogs of sunspot and plage records and magnetograms from SoHo/MDI (which is in a different orbit than SDO). The most likely explanation for the nature of the 335 day signal is that it is a long-term signature of magnetic activity.

7.3. From This Idealized Scenario to Stellar Observations

In this idealized setup, we successfully retrieve 300 day planet orbits with RV amplitudes down to $0.3 \, \mathrm{m \, s^{-1}}$, while $0.1 \, \mathrm{m \, s^{-1}}$ signals remain out of reach. To break the $0.1 \, \mathrm{m \, s^{-1}}$ barrier, additional RV signals not well traced by either or FF' will need to be modeled adequately. More generally, these planet injection tests show that in order to access planets with periods of a few hundred days, we will need to model all stellar signals that have similar periods and amplitudes much larger than K. The $|\hat{B}_\mathrm{obs}|$ and the combined $|\hat{B}_\mathrm{obs}| + FF'$ models perform equally well at this orbital period range. This is expected since we obtain very similar RV residuals with both models (see Sections 4 and 5). The planet retrievals carried out here are a best-case scenario that only considers *rotationally modulated* RV variations from magnetically active regions. In stellar observations, there will be additional intrinsic variability from magnetoconvection. For this data set, it would have an RMS of $1.1 \, \mathrm{m \, s^{-1}}$ given our cadence and sampling strategy (Meunier et al. 2015). Long-baseline stellar observations are expected to feature RV variations from large-scale meridional circulation, which varies with the magnetic cycle (Komm et al. 1993; Meunier 1999; Ulrich 2010). Although meridional flows have not yet been clearly identified in stars other than the Sun, they produce peak-to-peak RV amplitudes of 1–$1.4 \, \mathrm{m \, s^{-1}}$ when viewing the Sun edge-on, and 2.3–$3.3 \, \mathrm{m \, s^{-1}}$ in a pole-on scenario (Meunier & Lagrange 2020). Meunier & Lagrange (2020) extend their results to other Sun-like stars and predict peak-to-peak amplitudes of up to $4 \, \mathrm{m \, s^{-1}}$ in stars with strong magnetic cycles seen pole-on. The significance of a planet detection will depend strongly on how precisely we can measure the unsigned magnetic flux in Sun-like stars; see further discussion in Section 8.1. Additionally, there will be instrumental systematics of order 0.1–$1 \, \mathrm{m \, s^{-1}}$ for current-generation spectrographs (Fischer et al. 2016). In particular, wavelength calibration and long-term stability remain challenging even in current state-of-the-art spectrographs (e.g., Cosentino et al. 2012). Our RV and $|\hat{B}_\mathrm{obs}|$ time series are sampled simultaneously, and every night for 8 seasons; ground-based surveys will suffer losses from poor weather (e.g., Hall et al. 2018). These caveats will impact the performance of the technique presented here and diminish the significance of the mass determinations. We note that a systematic investigation of the detectability of low-mass, long-period planets is underway (Langellier et al. 2021). They explore a broad range of parameter space using HARPS-N RV observations of the Sun, so their analysis is highly complementary to the one here. Langellier et al. (2021) inject a wide variety of planet signals into solar observations and recover them by treating magnetic activity using Gaussian-process regression. For an in-depth investigation of the impact of observation strategies on the detectability of Earth-mass, long-period planets, we refer the reader to Hall et al. (2018).

8. Future Prospects

8.1. Prospects for Measuring $|\hat{B}_\mathrm{obs}|$ for Other Stars

The significance of an exoplanet detection will depend strongly on how precisely we can measure the unsigned magnetic flux in stars other than the Sun. Zeeman Doppler

imaging (ZDI; Donati & Brown 1997) has long been used to image the large-scale (polarized) magnetic field structures of stars, particularly those that are fast rotating and much more active than the Sun (see Reiners 2012, and references therein). However, RV variations stem from magnetic fields taking place on much smaller spatial scales than those probed by ZDI. In principle, it is possible to measure small-scale, unsigned magnetic flux by examining Zeeman broadening in magnetically sensitive spectral lines of stars (Robinson 1980; Saar 1988). It has been detected, for example, in the younger, moderately active, faster rotating ($P_{\rm rot} \approx$ 12 days) K2 dwarf Epsilon Eridani, which has an average unsigned magnetic flux in the range of 125–200 G (Valenti et al. 1995; Lehmann et al. 2015). The Sun's average unsigned magnetic field is 20 times smaller (10 G). Several studies have attempted to measure Zeeman broadening in slowly rotating, relatively quiet late-type stars (e.g., Saar & Linsky 1986; Saar et al. 1986; Basri & Marcy 1988; Rueedi et al. 1997; Anderson et al. 2010). More recently, Kochukhov et al. (2020) developed a method employing multiple lines with different Zeeman splitting patterns and making use of differential magnetic intensification of line equivalent widths (Basri et al. 1992; Saar et al. 1992). This work looks very promising for accurately pushing $|\hat{B}_{\rm obs}|$ detections to lower levels: using only eight spectral lines, Kochukhov et al. (2020) detect filling factors as low as $f = 7\%$ and unsigned average fields as low as 220 G. Modeling more lines should yield further improvements in determining $|\hat{B}_{\rm obs}|$ at low levels. Mortier (2016) combine multi-profile least-squares deconvolution (originally proposed by Kochukhov et al. 2010) with singular value decomposition to extract the unsigned magnetic flux from thousands of spectral lines. Their preliminary application to high-resolution spectra from HARPS and HARPS-N of an inactive K3 dwarf gives encouraging results.

The aforementioned studies identify several avenues to improve our prospects of detecting Zeeman broadening in old, quiet-Sun-like stars. These include improving modeling of spectral lines (better atomic data and better understanding of line broadening), particularly in the (near-)IR where Zeeman broadening is stronger as it has a squared dependence on wavelength; improving our constraints on convection and turbulence in the stellar atmosphere; and improving our understanding of the impact of line blending and telluric contamination. The magnetic-region filling factor and magnetic field strength are, to some extent, degenerate, but their product, $B.f$ is more easily measurable (e.g., Gray 1984; Saar 1988; Reiners 2012, and references therein). For example, the SPIRou spectrograph at CFHT provides meter-per-second precision spectroscopic observations that extend into the near-IR (Artigau et al. 2011) where the Zeeman effect is stronger, and will therefore provide excellent observations to improve our techniques to measure unsigned magnetic flux in slowly rotating, quiet-Sun-like stars.

8.2. Prospects for Improving the FF′ and F² Methods

In Section 6 and Appendix C, we noted that the FF' and F^2 terms of Aigrain et al. (2012) have intrinsic limitations, largely because they introduce extra limb-darkening-like damping to the predicted ΔRV. It may be possible to improve these methods by applying appropriate *anti-limb-darkening* functions to each, timed to the central meridian passage a magnetic feature. We provide further details in Appendix D. There are

several difficulties associated with our proposed correction, but we plan to experiment with these ideas in the near future.

8.3. Future Improvements to the SDO/HMI-RV Pipeline

The HMI instrument is not stable over timescales longer than a few days, because it was originally designed for helioseismic observations (see Section 2.1.2 and Figure 2). To correct for this, we look at RV *variations*, which we obtain by subtracting the velocity of the quiet Sun. In the present analysis, we estimate this quiet-Sun velocity by excluding magnetically active pixels. This is a reasonable approximation since active regions occupy only a few percent of the solar disk (5% at the peak of the magnetic cycle). Ideally, one should replace each active pixel by a non-active pixel from the same spatial location (e.g., from a few days before or after the active region's presence). This could be achieved by compiling a quiet-Sun *template* image. Such a fix is unlikely to account for the rotationally modulated RV residuals we are seeing, but may improve the accuracy of ΔRV and ultimately enable us to probe deeper into the physical processes at play in RV variations.

9. Conclusions

At present, we cannot yet routinely measure a proxy for activity-induced RV variations that enables us to confirm and characterize long-period exoplanets with RV semi-amplitudes of 30 cm s^{-1} and below (Crass et al. 2021). In this paper, we estimate the disk-averaged, rotationally modulated RV variations of the Sun as a star over magnetic Cycle 24 from spatially resolved images of the Sun taken by the HMI on board SDO. To do so, we apply a model that was previously validated against overlapping HARPS-N solar observations by Milbourne et al. (2019). We also estimate the disk-averaged, unsigned (i.e., unpolarized) magnetic flux. Our findings and conclusions are summarized here:

1. The SDO/HMI-derived RV data set presented here covers nearly an entire magnetic cycle, including maximum variability levels that have not been observed by HARPS-N or other precise, long-term stable spectrographs. Furthermore, its systematics are independent of such spectroscopic data sets. Our HMI-RV data set has a high cadence and high signal-to-noise ratio (Figure 4). It thus provides a testbed to identify and probe the underlying physical processes that are responsible for rotationally modulated RV variations. We refer the reader to the publicly available Python pipeline, SolAster, which can estimate the same time series as those presented in this analysis.[14]

2. Periodograms of the Sun's RV (Figure 6) show that the majority of the power is shared between the rotation period (P) and its first harmonic ($P/2$). Both peaks are significant, and each is in fact broad forests of significant peaks that lie up to 2–3 days away from P and $P/2$.

3. We fit RV variations with a linear model of the unsigned magnetic flux, and find that it reduces the RMS of RV variations by 62%, i.e., a factor of 2.6, from 2.33– 0.89 m s^{-1} (Figure 7, Section 4). The residuals of the fit display rotationally modulated behavior, particularly at times of high magnetic activity (Figure 8). To try to

---

[14] SolAster (Ervin et al. 2022) is publicly available at: https://tamarervin.github.io/SolAster/.

account for these residuals, we fit RV variations with a combination of a linear $|\hat{B}_{\mathrm{obs}}|$ term and an FF' term from the method of Aigrain et al. (2012) (Section 5). This yields only modest RMS improvements, as the combined model gives a residual RMS of $0.85\,\mathrm{m\,s}^{-1}$ for the full time series. We show that the FF' model does not adequately account for RV variations from magnetic regions because it over-accounts for limb darkening (Section 6 and Appendix C), and we propose a correction to potentially improve the performance of the FF' method (Section 8.2 and Appendix D).

4. Modeling RV variations with $|\hat{B}_{\mathrm{obs}}|$ and the FF' method allows us to identify additional physical processes responsible for rotationally modulated RV variations. These signals are either missing from the RV model of Milbourne et al. (2019) and Haywood et al. (2016) that we use to estimate RV variations from SDO/HMI images, or they are not well traced by $|\hat{B}_{\mathrm{obs}}|$ or FF', or both. Particularly at high magnetic activity levels, the residuals display significant, rotationally modulated variations at the meter-per-second level. We discuss physical processes that may contribute to these additional RV variations beyond suppression of convective blueshift and brightness inhomogeneities: horizontal flows (such as Evershed flows, moat flows, and active region inflows), flows that are not resolved by SDO/HMI's pixels, and Zeeman broadening (Section 6).

5. We inject planet signals to test the performance of the unsigned magnetic flux $|\hat{B}_{\mathrm{obs}}|$ for mitigating rotationally modulated RV variations in surveys of low-mass, long-period planets orbiting Sun-like stars (Section 7). We inject planets with orbital periods of $\simeq300$ days and RV semi-amplitudes of 0.5, 0.3, and $0.1\,\mathrm{m\,s}^{-1}$. We are unable to reliably detect planets with RV semi-amplitudes below $1\,\mathrm{m\,s}^{-1}$ without modeling activity-induced RV variations. The $|\hat{B}_{\mathrm{obs}}|$ model and the combined $|\hat{B}_{\mathrm{obs}}| + FF'$ model give very similar results. The parameters of the planets with $K = 0.3$ and $0.5\,\mathrm{m\,s}^{-1}$ are detected accurately to within 1σ of the injected parameters, when accounting for solar activity in our model. We do not retrieve injected signals with $K = 0.1\,\mathrm{m\,s}^{-1}$.

We conclude that $|\hat{B}_{\mathrm{obs}}|$ could, in principle, enable us to extract planet signals down to $0.3\,\mathrm{m\,s}^{-1}$, but we will also need to model additional RV variations to reach $0.1\,\mathrm{m\,s}^{-1}$ (Section 7.3). The significance of planet detections in stellar observations will depend crucially on how precisely we may be able to measure $|\hat{B}_{\mathrm{obs}}|$. Stellar RV observations will also be affected by (super)granulation signals at the meter-per-second level (Meunier et al. 2015), instrumental systematics, and ground-based observing schedules. The most promising avenue to measure the unsigned magnetic flux in slowly rotating, relatively inactive stars is by measuring Zeeman broadening of magnetically sensitive lines in high precision spectra (e.g., Kochukhov et al. 2020).

Appendix A
Hystereses between ΔRV, S-index, and $|\hat{B}_{\mathrm{obs}}|$

Meunier et al. (2019) showed that for a given level of Ca II H&K emission, RV variations have a comparatively lower amplitude during the descending phase of the magnetic cycle than in its ascending phase. This is likely because the average latitude $\langle|\phi|\rangle$ of active regions changes over the course of the magnetic cycle, highest in the ascending phase, and lowest in the descending. Signals produced by an active region depend on the region's position on the solar disk; RV and S-index behave differently with line-of-sight angle and follow different limb-darkening laws because they originate at different heights in the solar atmosphere (further details in Meunier et al. 2019, Section 6.1). Note that $|\hat{B}_{\mathrm{obs}}|$ is a line-of-sight observable, and only subject to foreshortening.

To examine these long-term effects, we averaged our time series with a 300 day boxcar smooth. This is long enough to smooth out both rotational modulation, and the growth-decay timescales of large active regions, thus concentrating on purely cyclic variation. As shown in Figure A1, we observe a
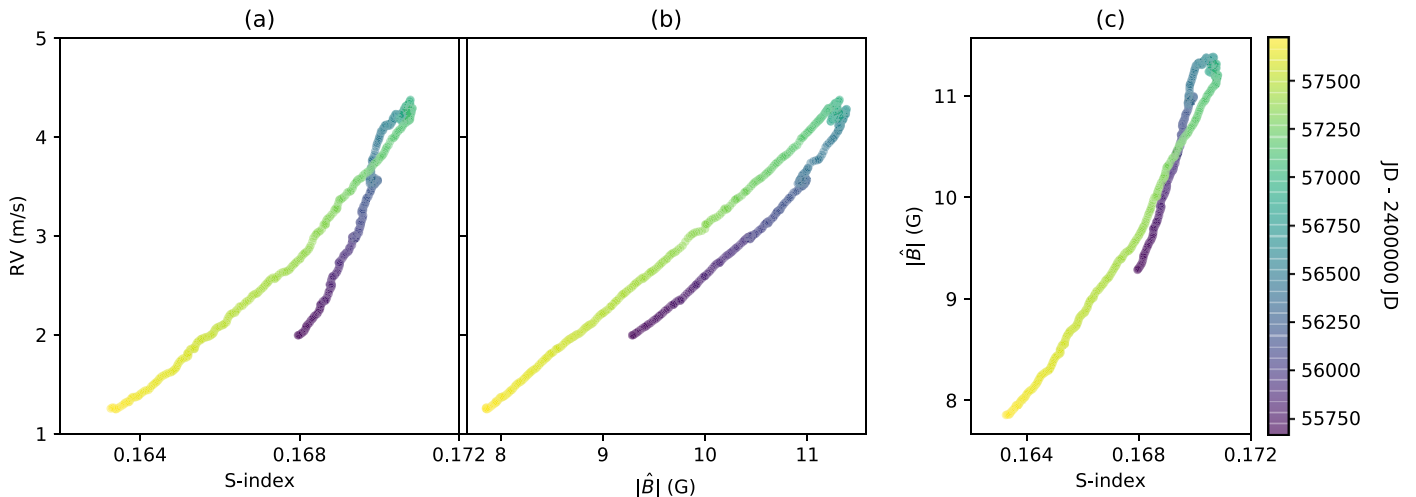
Figure A1. Hysteresis between RV variations and (a) Ca II H&K emission as measured by the S-index, (b) unsigned magnetic flux, $|\hat{B}_{\rm obs}|$, and (c) between $|\hat{B}_{\rm obs}|$ and the S-index. All time series are smoothed over 300 day bins to reveal the long-term hystereses. The points are color coded according to time: the ascending phase (increasing activity) is in purple and blue while the descending phase (decreasing activity) is in green and yellow. Note that in these plots the SDO/HMI ΔRV and $|\hat{B}_{\rm obs}|$ time series are concatenated with the S-index time series (resulting in time series of 1089 points each).

hysteresis between ΔRV and S-index (panel (a)), between ΔRV and the unsigned magnetic flux $|\hat{B}_{\rm obs}|$ (panel (b)), and between the S-index and $|\hat{B}_{\rm obs}|$ (panel (c)).

The RV–S-index hysteresis looks qualitatively similar to that observed in the previous solar magnetic cycle by Meunier et al. (2019, Figure 7). In the ascending phase (dark purple line in Figure A1), plages, at higher $\langle|\phi|\rangle$, are (more) limb brightened, and also increasing in total area, yielding a larger S-index per unit projected plage area than later in the cycle. There is also an RV effect, as the difference in RV between nonmagnetic pixels and plage pixels peaks at $\mu \sim 0.9$ (Palumbo et al. 2017). This azimuthal ring lies entirely within $\phi = \pm 25°$, the approximate *active latitude zone*. Thus, as the cycle progresses and the active region average latitude decreases, the per-pixel average RV difference with the quiet Sun *decreases*. In the ascending phase, this is more than compensated by the increasing filling factor, but once $f_{\rm plage}$ starts declining in the descending phase, ΔRV drops steadily.

The hysteresis between ΔRV and $|\hat{B}_{\rm obs}|$ is in some ways simpler to understand. At the 300 day level of smoothing, the twin cycle maxima of Cycle 24 at $t \sim 2455818$ JD (the weaker northern hemisphere peak) and $t \sim 2456918$ JD (the stronger southern hemisphere peak) are both flattened into one slow increase in $|\hat{B}_{\rm obs}|$. With an initial $\phi \sim 25°$, then mostly decreasing throughout the cycle, the average net RV per plage pixel should also be decreasing. This is counterbalanced, however, by the filling factor, which increases more quickly than RV decreases in both the ascending phase and the peak(s) of the cycle. Thus, ΔRV continues to increase with $|\hat{B}_{\rm obs}|$ in these phases, more slowly at maximum when $|\hat{B}_{\rm obs}|$ growth is also reduced, then only finally reversing in the decline phase, when both $|\hat{B}_{\rm obs}|$ and ΔRV plummet.

Panel (c) of Figure A1 shows a hysteresis between the S-index and $|\hat{B}_{\rm obs}|$, which is likely due to the different limb-darkening behaviors, with the S-index getting a boost from limb brightening when in the higher $\langle|\phi|\rangle$ ascending phase.

Further study is needed to better understand the differences in projection effects between these three observables, to correct for them and therefore obtain tighter correlations between ΔRV, S-index, and $|\hat{B}_{\rm obs}|$.

Attempt to account for the hysteresis between ΔRV and $|\hat{B}_{\rm obs}|$— To capture the information on the hysteresis shown in Figure A1(b), we fitted the ascending and descending phases of the magnetic cycle separately, using two $|\hat{B}_{\rm obs}|$ terms and two zero-point offsets. We identified the peak of the magnetic cycle, i.e., the point separating the ascending and descending phases as the maximum in unsigned magnetic flux at JD = 2456844.9, with unsigned magnetic flux 18.3G. As we expected, this timestamp falls between the double-peaked shape of the magnetic cycle, which occurs because the active region bands reach maximum activity levels at slightly different times. This timestamp also falls within the 300 day bin where the hysteresis turns over as shown in Figure A1. We obtain different model parameters for the ascending ($\alpha = 11.15 \pm 0.01$ m s^{-1}, $\Delta RV_0 = -7.98 \pm 0.01$ m s^{-1}) and descending phases ($\alpha = 8.65 \pm 0.01$ m s^{-1}, $\Delta RV_0 = -6.27 \pm 0.01$ m s^{-1}). However, the residual RMS (over the full cycle) is 0.83 m s^{-1} (64% reduction in RV variations), which is only a small improvement compared to fitting the full cycle as one (RMS = 0.89 m s^{-1}, 62% reduction). This is an improvement, but since it is small and adds more parameters and complexity, we leave this avenue open for future investigations.

Appendix B
## Correlations between Filling Factors and Activity Indicators

In Figure B1, we show the plage filling factor as a function of $|\hat{B}_{\rm obs}|$ and S-index, and the spot filling factor as a function of $|\hat{B}_{\rm obs}|$. The magnetic filling factors are estimated according to Equation (2).

Appendix C
Demonstration of the Limitations to the *FF'* Method

Here, we show why the *FF'* method cannot match RV variations perfectly. Consider an equatorial spot. If we assume, for simplicity, linear limb darkening, and solar inclination $i = 90°$, the flux from the spot of area A can be written as

$$F = A \cos\theta (1 - \epsilon_{\rm s} + \epsilon_{\rm s} \cos\theta), \qquad (C1)$$
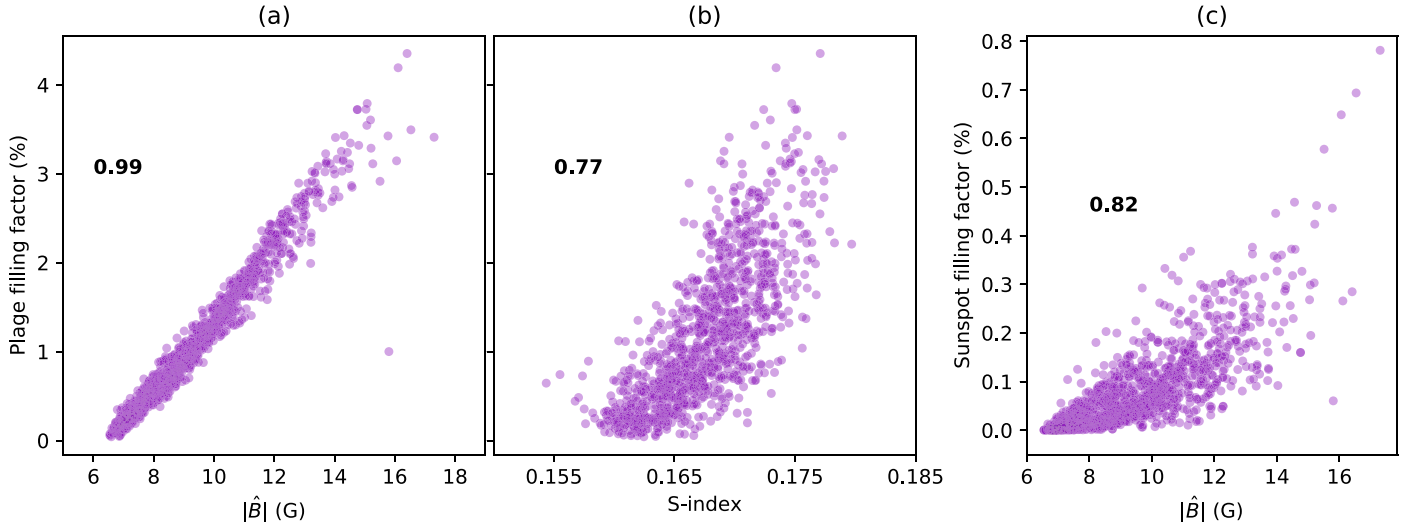
Figure B1. Plage filling factor plotted against (a) $|\hat{B}_{\mathrm{obs}}|$ and (b) S-index. (c) Shows the sunspot filling factor as a function of $|\hat{B}_{\mathrm{obs}}|$. Spearman correlation coefficients are given for each pair of correlates. Note that in these plots the SDO/HMI ΔRV and $|\hat{B}_{\mathrm{obs}}|$ time series are concatenated with the S-index time series (resulting in time series of 1089 points each).

where θ is the angle of the surface normal to the line of sight (at disk center), and ϵ_{s} is the linear limb-darkening coefficient for the spot. The derivative with respect to θ is

$$F' = -A \sin\theta (1 - \epsilon_{\mathrm{s}} + 2\epsilon_{\mathrm{s}} \cos\theta). \qquad (C2)$$

To match the RV change due to the rotation of a spot, this simple model (based on Saar & Donahue 1997) yields

$$\Delta v_{\mathrm{spot}} = -A \sin\theta \cos\theta (1 - \epsilon_{\mathrm{s}} + \epsilon_{\mathrm{s}} \cos\theta). \qquad (C3)$$

The $\sin\theta$ term captures the RV deficit, and is weighted by the projected area of the spot ($\cos\theta$) and its limb darkening, $(1 - \epsilon_{\mathrm{s}} + \epsilon_{\mathrm{s}} \cos\theta)$.

In comparison, the FF' method delivers

$$FF'/A = -A \sin\theta \cos\theta (1 - \epsilon_{\mathrm{s}} + \epsilon_{\mathrm{s}} \cos\theta)$$
$$\times (1 - \epsilon_{\mathrm{s}} + 2\epsilon_{\mathrm{s}} \cos\theta), \qquad (C4)$$

or,

$$FF'/A = \Delta v_{\mathrm{spot}} (1 - \epsilon_{\mathrm{s}} + 2\epsilon_{\mathrm{s}} \cos\theta). \qquad (C5)$$

Thus, the FF' method captures Δv_{spot} but adds an additional limb-darkening-like term. This leads to systematic effects, underestimating Δv_{spot} progressively more and more as the spot moves away from the disk center.

Although we do not use the F^2 term proposed by Aigrain et al. (2012) to correct for the convective suppression arising primarily in plage, we note that it is similarly flawed. Following a similar analysis,

$$\Delta v_{\mathrm{plage}} = A \cos^2\theta (1 - \epsilon_{\mathrm{p}} + \epsilon_{\mathrm{p}} \cos\theta), \qquad (C6)$$

where ϵ_{p} is the linear limb-darkening coefficient for plage. But F^2 yields

$$F^2/A = A \cos^2\theta (1 - \epsilon_{\mathrm{p}} + \epsilon_{\mathrm{p}} \cos\theta)^2, \qquad (C7)$$

or,

$$F^2/A = \Delta v_{\mathrm{plage}} (1 - \epsilon_{\mathrm{p}} + \epsilon_{\mathrm{p}} \cos\theta), \qquad (C8)$$

which contains an extra limb-darkening term.

In the case of a spot, its limb darkening ϵ_{s} is unlikely to differ significantly from the quiet-Sun value ϵ_{q}. This is because limb darkening is wavelength dependent, and the strength-weighted average for HARPS-N RV lines is perhaps $\langle\lambda\rangle \sim 500\,\mathrm{nm}$, not much different from the HMI line ($\lambda = 617.3\,\mathrm{nm}$). FF' captures the RV perturbation due to a spot crossing the disk effectively, but then applies a second, additional, stronger limb darkening.

We thus warn that, while the FF' term is partially successful, it (and the related F^2 term) are also flawed, since they intrinsically add extra, unwanted limb-darkening-like corrections. Therefore, one cannot expect the FF' term (or the F^2 term) to perfectly account for the RV variations of magnetic features.

Appendix D
On Improving the FF' and F^2 Terms

For passage of a spot (detected, e.g., in photometry), the corrected FF' becomes

$$FF'_{\mathrm{corrected,spot}} = FF'/(1 - \epsilon_{\mathrm{s}} + 2\epsilon_{\mathrm{s}} \cos\theta). \qquad (D1)$$

A similar formula can similarly be applied to plage (detected, e.g., in Ca II H&K), by using the limb-darkening coefficient ϵ_{p}. We note that the equivalent correction to F^2 for convective suppression, which occurs predominantly in plage (rather than spots) is

$$F^2_{\mathrm{corrected,plage}} = F^2/(1 - \epsilon_{\mathrm{p}} + \epsilon_{\mathrm{p}} \cos\theta). \qquad (D2)$$

In order to apply the above corrections, one has to know when individual active regions cross the solar/stellar disk. To some extent, one can track the meridian passage of active regions on solar/stellar surfaces via monitoring of disk-averaged photometric and Ca II H&K emission. However, an active region often contains both spots *and* plage, resulting in a mixed, degenerate photometric signal. Also, it would be difficult to apply this correction in terms of timing when multiple active regions are present. Additionally, we note that ϵ_{s} and ϵ_{p} remain poorly known due to the lack of realistic models. Better values may be derivable from 3D-MHD simulations (e.g., Cegla et al. 2018). It is also difficult to choose an appropriate value of $\langle\epsilon\rangle$ for spectra that span hundreds of nm. To summarize, applying a

correction to the FF' term could potentially improve fits to RV variations incurred by spots and plage, but there will likely remain residual RV variations.

ORCID iDs

R. D. Haywood ⓞ https://orcid.org/0000-0001-9140-3574
S. H. Saar ⓞ https://orcid.org/0000-0001-7032-8480
A. Mortier ⓞ https://orcid.org/0000-0001-7254-4363
D. Phillips ⓞ https://orcid.org/0000-0001-5132-1339
D. Charbonneau ⓞ https://orcid.org/0000-0002-9003-484X
A. Collier Cameron ⓞ https://orcid.org/0000-0002-8863-7828
H. M. Cegla ⓞ https://orcid.org/0000-0001-8934-7315
M. L. Palumbo III ⓞ https://orcid.org/0000-0002-4677-8796

References

Aigrain, S., Pont, F., & Zucker, S. 2012, MNRAS, 419, 3147
Anderson, R. I., Reiners, A., & Solanki, S. K. 2010, A&A, 522, A81
Artigau, É, Donati, J. F., & Delfosse, X. 2011, in ASP Conf. Ser. 448, Planet Detection, Magnetic Field of Protostars and Brown Dwarfs Meteorology with SPIRou, ed. C. Johns-Krull, M. K. Browning, & A. A. West (San Francisco, CA: ASP), 771
Baliunas, S. L., Donahue, R. A., Soon, W. H., et al. 1995, ApJ, 438, 269
Basri, G., & Marcy, G. W. 1988, ApJ, 330, 274
Basri, G., Marcy, G. W., & Valenti, J. A. 1992, ApJ, 390, 622
Bastien, F. A., Stassun, K. G., Pepper, J., et al. 2014, AJ, 147, 29
Batalha, N. E., Lewis, T., Fortney, J. J., et al. 2019, ApJL, 885, L25
Buehler, D., Lagg, A., Solanki, S. K., & van Noort, M. 2015, A&A, 576, A27
Cegla, H. M., Watson, C. A., Shelyag, S., et al. 2018, ApJ, 866, 55
Collier Cameron, A., Mortier, A., Phillips, D., et al. 2019, MNRAS, 487, 1082
Cosentino, R., Lovis, C., Pepe, F., et al. 2012, Proc. SPIE, 8446, 84461V
Couvidat, S., Schou, J., Hoeksema, J. T., et al. 2016, SoPh, 291, 1887
Crass, J., Gaudi, B. S., Leifer, S., et al. 2021, arXiv:2107.14291
Deming, D., & Plymate, C. 1994, ApJ, 426, 382
Donati, J. F., & Brown, S. F. 1997, A&A, 326, 1135
Dravins, D., Lindegren, L., & Nordlund, A. 1981, A&A, 96, 345
Dumusque, X. 2019, BAAS, 51, 6
Dumusque, X., Boisse, I., & Santos, N. C. 2014, ApJ, 796, 132
Dumusque, X., Cretignier, M., Sosnowska, D., et al. 2021, A&A, 648, A103
Dumusque, X., Glenday, A., Phillips, D. F., et al. 2015, ApJL, 814, L21
Egeland, R., Soon, W., Baliunas, S., et al. 2017, ApJ, 835, 25
Ervin, T., Halverson, S., Burrows, A., et al. 2022, AJ, 163, 272
Evershed, J. 1909, MNRAS, 69, 454
Fischer, D. A., Anglada-Escude, G., Arriagada, P., et al. 2016, PASP, 128, 066001
Fligge, M., Solanki, S. K., & Unruh, Y. C. 2000, A&A, 353, 380
Foukal, P. V. 2004, Solar Astrophysics (2nd edn.; New York: Wiley)
Fröhlich, C., & Lean, J. 1998, GeoRL, 25, 4377
Gizon, L., Birch, A. C., & Spruit, H. C. 2010, ARA&A, 48, 289
Gizon, L., Duvall, T. L. J., & Larsen, R. M. 2001, in IAU Symp. 203, Recent Insights into the Physics of the Sun and Heliosphere: Highlights from Soho and Other Space Missions (Cambridge: Cambridge Univ. Press), 189
Goodman, J., & Weare, J. 2010, Commun. Appl. Math. Comput. Sci., 5, 65
Gray, D. F. 1984, ApJ, 277, 640
Hall, R. D., Thompson, S. J., Handley, W., & Queloz, D. 2018, MNRAS, 479, 2968
Haywood, R. D., Collier Cameron, A., Queloz, D., et al. 2014, MNRAS, 443, 2517
Haywood, R. D., Collier Cameron, A., Unruh, Y. C., et al. 2016, MNRAS, 457, 3637
Hoeksema, J. T., Baldner, C. S., Bush, R. I., Schou, J., & Scherrer, P. H. 2018, SoPh, 293, 45
Hojjatpanah, S., Oshagh, M., Figueira, P., et al. 2020, A&A, 639, A35
Iampietro, A., Saar, S. H., Haywood, R. D., & Milbourne, T. W. 2019, AGU Meeting, 2019, SH41F–3321
Kochukhov, O., Hackman, T., Lehtinen, J. J., & Wehrhahn, A. 2020, A&A, 635, A142
Kochukhov, O., Makaganiuk, V., & Piskunov, N. 2010, A&A, 524, A5
Komm, R. W., Howard, R. F., & Harvey, J. W. 1993, SoPh, 147, 207
Kopp, G. 2014, JSWSC, 4, A14
Kopp, G., Rottman, G., Harder, J., et al. 2001, AGU Meeting, 2001, SH52A–08
Krivova, N. A., Balmaceda, L., & Solanki, S. K. 2007, A&A, 467, 335
Lagrange, A.-M., Desort, M., & Meunier, N. 2010, A&A, 512, A38
Langellier, N., Milbourne, T. W., Phillips, D. F., et al. 2021, AJ, 161, 287
Lanza, A. F., Gizon, L., Zaqarashvili, T. V., Liang, Z. C., & Rodenbeck, K. 2019, A&A, 623, A50
Lanza, A. F., Molaro, P., Monaco, L., & Haywood, R. D. 2016, A&A, 587, A103
Lean, J., Rottman, G., Harder, J., & Kopp, G. 2005, SoPh, 230, 27
Lehmann, L. T., Künstler, A., Carroll, T. A., & Strassmeier, K. G. 2015, AN, 336, 258
Lin, A. S. J., Monson, A., Mahadevan, S., et al. 2022, AJ, 163, 184
Lockwood, G. W., Skiff, B. A., Henry, G. W., et al. 2007, ApJS, 171, 260
Löhner-Böttcher, J., & Schlichenmaier, R. 2013, A&A, 551, A105
Lovis, C., Dumusque, X., Santos, N. C., et al. 2011, arXiv:1107.5325
Meunier, N. 1999, ApJ, 527, 967
Meunier, N. 2003, A&A, 405, 1107
Meunier, N. 2018, A&A, 615, A87
Meunier, N., Desort, M., & Lagrange, A.-M. 2010a, A&A, 512, A39
Meunier, N., Lagrange, A.-M., & Desort, M. 2010b, A&A, 519, A66
Meunier, N., & Lagrange, A. M. 2013, A&A, 551, A101
Meunier, N., & Lagrange, A. M. 2019, A&A, 628, A125
Meunier, N., & Lagrange, A.-M. 2020, A&A, 638, A54
Meunier, N., Lagrange, A.-M., Borgniet, S., & Rieutord, M. 2015, A&A, 583, A118
Meunier, N., Lagrange, A. M., & Cuzacq, S. 2019, A&A, 632, A81
Meunier, N., Mignon, L., & Lagrange, A.-M. 2017, A&A, 607, A124
Miklos, M., Milbourne, T. W., Haywood, R. D., et al. 2020, ApJ, 888, 117
Milbourne, T. W., Haywood, R. D., Phillips, D. F., et al. 2019, ApJ, 874, 107
Morley, C. V., Kreidberg, L., Rustamkulov, Z., Robinson, T., & Fortney, J. J. 2017, ApJ, 850, 121
Mortier, A. 2016, in 19th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun (CS19) (Berlin: Springer), 134
Mortier, A., & Collier Cameron, A. 2017, A&A, 601, A110
National Academies of Sciences, Engineering, and Medicine 2018, Exoplanet Science Strategy (Consensus Study Report) (Washington, DC: The National Academies Press)
Nava, C., López-Morales, M., Haywood, R. D., & Giles, H. A. C. 2020, AJ, 159, 23
Noyes, R. W., Weiss, N. O., & Vaughan, A. H. 1984, ApJ, 287, 769
Oshagh, M., Santos, N. C., Figueira, P., et al. 2017, A&A, 606, A107
Palumbo, M. L. I., Haywood, R. D., Saar, S. H., Dupree, A. K., & Milbourne, T. W. 2017, AGUFM, 2017, P53B–2651
Phillips, D. F., Glenday, A. G., Dumusque, X., et al. 2016, Proc. SPIE, 9912, 99126Z
Radick, R. R., Lockwood, G. W., Henry, G. W., Hall, J. C., & Pevtsov, A. A. 2018, ApJ, 855, 75
Rasmussen, C. E., & Williams, C. K. I. 2006, Gaussian Processes for Machine Learning (Cambridge, MA: MIT Press)
Reiners, A. 2012, LRSP, 9, 1
Reiners, A., Shulyak, D., Anglada-Escudé, G., et al. 2013, A&A, 552, A103
Robinson, R. D., Jr. 1980, ApJ, 239, 961
Rueedi, I., Solanki, S. K., Mathys, G., & Saar, S. H. 1997, A&A, 318, 429
Saar, S. H. 1988, ApJ, 324, 441
Saar, S. H. 2003, in ASP Conf. Ser. 294, The Effects of Plage on Precision Radial Velocities, ed. D. Deming & S. Seager (San Francisco, CA: ASP), 65
Saar, S. H. 2009, in AIP Conf. Ser. 1094, 15th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, ed. E. Stempels (Melville, NY: AIP), 152
Saar, S. H., & Donahue, R. A. 1997, ApJ, 485, 319
Saar, S. H., & Fischer, D. 2000, ApJL, 534, L105
Saar, S. H., & Linsky, J. L. 1986, AdSpR, 6, 235
Saar, S. H., Linsky, J. L., & Duncan, D. K. 1986, in Cool Stars, Stellar Systems, and the Sun (Vol. 254) ed. M. Zeilik & D. M. Gibson (Berlin: Springer-Verlag), 275
Saar, S. H., Piskunov, N. E., & Tuominen, I. 1992, in ASP Conf. Ser. 26, Magnetic Surface Images of the BY DRA Star HD 82558, ed. M. S. Giampapa & J. A. Bookbinder (San Francisco, CA: ASP), 255
Schou, J., Scherrer, P. H., Bush, R. I., et al. 2011, SoPh, 275, 207
Schou, J., Scherrer, P. H., Bush, R. I., et al. 2012, SoPh, 275, 229
Schrijver, C. J., & Zwaan, C. 2000, Solar and Stellar Magnetic Activity (Cambridge: Cambridge Univ. Press)
Shapiro, A. I., Solanki, S. K., Krivova, N. A., et al. 2014, A&A, 569, A38
Solanki, S. K. 2003, A&ARv, 11, 153

Spruit, H. C. 1976, SoPh, 50, 269
Ulrich, R. K. 2010, ApJ, 725, 658
Valenti, J. A., Marcy, G. W., & Basri, G. 1995, ApJ, 439, 939
Vaughan, A. H., Preston, G. W., & Wilson, O. C. 1978, PASP, 90, 267

Wilson, O. C. 1978, ApJ, 226, 379
Yeo, K. L., Solanki, S. K., & Krivova, N. A. 2013, A&A, 550, A95
Zechmeister, M., & Kürster, M. 2009, A&A, 496, 577
Zeng, L., & Sasselov, D. 2013, PASP, 125, 227