Detection Limits of Low-mass, Long-period Exoplanets Using Gaussian Processes Applied to HARPS-N Solar Radial Velocities

N. Langellier[1,2], T. W. Milbourne[1,2], D. F. Phillips[2], R. D. Haywood[2,3,21], S. H. Saar[2], A. Mortier[4,5],
L. Malavolta[6], S. Thompson[4], A. Collier Cameron[7], X. Dumusque[8], H. M. Cegla[8,9,22], D. W. Latham[2],
J. Maldonado[10], C. A. Watson[11], N. Buchschacher[8], M. Cecconi[12], D. Charbonneau[2], R. Cosentino[12], A. Ghedina[12],
M. Gonzalez[12], C-H. Li[2], M. Lodi[12], M. López-Morales[2], G. Micela[10], E. Molinari[12,13], F. Pepe[8], E. Poretti[12,14],
K. Rice[15,16], D. Sasselov[2], A. Sozzetti[17], S. Udry[8], and R. L. Walsworth[2,18,19,20]

[1] Department of Physics, Harvard University, 17 Oxford Street, Cambridge MA 02138, USA; nlangellier@gmail.com
[2] Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[3] Astrophysics Group, University of Exeter, Exeter EX4 2QL, UK
[4] Astrophysics Group, Cavendish Laboratory, J.J. Thomson Avenue, Cambridge CB3 0HE, UK
[5] Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[6] Dipartimento di Fisica e Astronomia "Galileo Galilei," Università di Padova, Vicolo dell'Osservatorio 3, I-35122 Padova, Italy
[7] Centre for Exoplanet Science, SUPA, School of Physics and Astronomy, University of St Andrews, St Andrews KY16 9SS, UK
[8] Observatoire de Genève, Université de Genève, 51 chemin des Maillettes, 1290 Versoix, Switzerland
[9] Department of Physics, University of Warwick, Coventry, CV4 7AL, UK
[10] INAF-Osservatorio Astronomico di Palermo, Piazza del Parlamento 1, I-90134 Palermo, Italy
[11] Astrophysics Research Centre, School of Mathematics and Physics, Queen's University Belfast, BT7 1NN, Belfast, UK
[12] INAF-Fundacion Galileo Galilei, Rambla Jose Ana Fernandez Perez 7, E-38712 Brena Baja, Spain
[13] INAF-Osservatorio Astronomico di Cagliari, Via della Scienza 5, I-09047 Selargius CA, Italy
[14] INAF-Osservatorio Astronomico di Brera, Via E. Bianchi 46, I-23807 Merate (LC), Italy
[15] SUPA, Institute for Astronomy, Royal Observatory, University of Edinburgh, Blackford Hill, Edinburgh EH93HJ, UK
[16] Centre for Exoplanet Science, University of Edinburgh, Edinburgh, UK
[17] INAF-Osservatorio Astrofisico di Torino, via Osservatorio 20, I-10025 Pino Torinese, Italy
[18] Department of Physics, University of Maryland, College Park, MD 20742, USA
[19] Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742, USA
[20] Quantum Technology Center, University of Maryland, College Park, MD 20742, USA

Abstract

Radial velocity (RV) searches for Earth-mass exoplanets in the habitable zone around Sun-like stars are limited by the effects of stellar variability on the host star. In particular, suppression of convective blueshift and brightness inhomogeneities due to photospheric faculae/plage and starspots are the dominant contribution to the variability of such stellar RVs. Gaussian process (GP) regression is a powerful tool for statistically modeling these quasi-periodic variations. We investigate the limits of this technique using 800 days of RVs from the solar telescope on the High Accuracy Radial velocity Planet Searcher for the Northern hemisphere (HARPS-N) spectrograph. These data provide a well-sampled time series of stellar RV variations. Into this data set, we inject Keplerian signals with periods between 100 and 500 days and amplitudes between 0.6 and 2.4 m s^{-1}. We use GP regression to fit the resulting RVs and determine the statistical significance of recovered periods and amplitudes. We then generate synthetic RVs with the same covariance properties as the solar data to determine a lower bound on the observational baseline necessary to detect low-mass planets in Venus-like orbits around a Sun-like star. Our simulations show that discovering planets with a larger mass ($\sim$0.5 m s^{-1}) using current-generation spectrographs and GP regression will require more than 12 yr of densely sampled RV observations. Furthermore, even with a perfect model of stellar variability, discovering a true exo-Venus ($\sim$0.1 m s^{-1}) with current instruments would take over 15 yr. Therefore, next-generation spectrographs and better models of stellar variability are required for detection of such planets.

Unified Astronomy Thesaurus concepts: Gaussian Processes regression (1930); Radial velocity (1332); Exoplanets (498); Solar activity (1475)

1. Introduction

State-of-the-art radial velocity (RV) searches for low-mass, long-period exoplanets are limited by signals produced by stellar magnetic variability. An Earth-like planet in orbit around a Sun-like star in its habitable zone induces a reflex RV signal on the order of 0.1 m s^{-1}. However, the presence of acoustic oscillations, magnetoconvection, large-scale magnetic structures, and other stellar processes induce RV perturbations that

can exceed 1 m s^{-1} (see Fischer et al. 2016 and references therein; Cegla 2019). This stellar variability can conceal and even mimic planetary signals in RV surveys (Robertson et al. 2020), and has resulted in many false detections (e.g., CoRoT-7d, Haywood et al. 2014; GJ 581d and g, Robertson et al. 2014; and Alpha Centauri Bb, Rajpaul et al. 2016). Furthermore, these stellar processes act on timescales between minutes and months (Kjeldsen & Bedding 1995; Giles et al. 2017). For Sun-like stars, the dominant contributions to these intrinsically driven RV perturbations are from the suppression of convective blueshift and brightness inhomogeneities modulated at the

rotation period (Meunier et al. 2010; Dumusque et al. 2014). The wide range of timescales and nontrivial correlations between these processes require a sophisticated statistical framework to decouple stellar activity processes from planetary signals. Hall et al. (2018) and Haywood et al. (2020) study these effects by recovering injected planets of known properties into real data.

In this work, we use state-of-the-art Gaussian process (GP) regression to account for the temporal correlations of rotationally modulated stellar activity (Dumusque et al. 2017; Damasso et al. 2019). This GP regression is trained on solar data, as measured by a purpose-built solar telescope feeding the High Accuracy Radial velocity Planet Searcher for the Northern hemisphere (HARPS-N) spectrograph (Cosentino et al. 2014) operating at the Telescopio Nazionale Galileo (TNG) in the Canary Islands (Dumusque et al. 2015; Phillips et al. 2016). In Section 2, we first present the solar data and the GP regression along with the resulting fit on the daily averaged solar RVs. In Section 3, we introduce synthetic planets of varying semi-amplitude and orbital period in order to determine the sensitivity of the GP regression to temperate, low-mass planet searches. We conclude in Section 4 with an analysis of the baseline of RV observations and model assumptions that are necessary for a true exo-Earth detection.

2. Methods

2.1. Data

We take 5 minute disk averaged exposures of the Sun using the solar telescope and the HARPS-N spectrograph at the TNG, and use a baseline of around 800 days of near-continuous, daytime solar spectra (Dumusque et al. 2015; Phillips et al. 2016). Neutral density filters in the HARPS-N calibration system are used to match the solar flux to the exposure time that is set to integrate over solar p-modes. For each exposure, the HARPS-N Data Reduction Software (Baranne et al. 1996; Sosnowska et al. 2012) computes the barycenter-corrected RV with 0.4 m s^{-1} single exposure precision and the Mt. Wilson S index (Wilson 1968), a measure of stellar magnetic activity derived from chromospheric re-emission in the core of the singly ionized Ca H and K line cores (Linsky & Avrett 1970). The resulting RVs are further reduced, as described in Collier Cameron et al. (2019), to remove the RV signatures of the solar system planets, effects of differential extinction across the solar disk in Earth's atmosphere, and other systematic effects due to Earth's orbit around the Sun. Additionally, we eliminate exposures contaminated by clouds or other significant atmospheric losses. We realize signal-to-noise ratios (S/N) above 300 in most exposures. After cutting exposures contaminated by clouds, we observe variations in the S/N below the 10% level over many months with statistical RV precision varying between 0.45 and 0.6 m s^{-1}. We have not observed correlations between extracted RVs and S/N. Finally, in order to mimic the sampling of a typical stellar observing schedule while preserving the exquisite S/N of the solar telescope (Phillips et al. 2016), we compute daily averaged values of each quantity including the RVs, the S index, and the corresponding mean Julian date. While realistic stellar observing schedules would not allow for this level of averaging, which integrates over variability on minute and hour timescales, we wish to assess the best-case scenario using an ultra-high S/N data set.

2.2. Gaussian Process Kernel

Magnetic variability can be the dominant source of variance for stellar RVs of nearby, bright Sun-like and low-mass stars, reaching levels surpassing 1 m s^{-1} (Isaacson & Fischer 2010; Motalebi et al. 2015). Intrinsic stellar variability introduces correlations into the RV time series that are difficult to model deterministically. GP regressions have emerged as a powerful statistical technique that relaxes the assumption of uncorrelated noise by adding nonzero terms to the off-diagonal of the data covariance matrix. Usually a kernel function is chosen to describe these covariances as a function of measurement separation time (Rasmussen & Williams 2006; Haywood et al. 2014; Rajpaul et al. 2015; Faria et al. 2016; Damasso et al. 2019).

Since magnetic activity is well described by the S index and is modulated at the rotation period of the star, we fit the S index to a GP with a quasi-periodic (QP) covariance kernel function, k_{QP}. This kernel function is chosen heuristically to model the known properties of the magnetic activity, yielding the correlation between two measurements at times t_i and t_j given by

$$k_{QP}(t_i, t_j) = \exp\left(-\frac{(t_i - t_j)^2}{2\tau^2} - \frac{1}{2\eta^2}\sin^2\left(\frac{\pi(t_i - t_j)}{P_{rot}}\right)\right),$$
(1)

where τ is related to the average lifetime of active regions on the Sun, η is related to the average distribution of activity in the photosphere of the Sun, and P_{rot} is approximately the synodic rotation period of the Sun. Grunblatt et al. (2015) show that this kernel function performs the best among three common GP kernel functions. The GP regression also includes a white noise (WN) term to account for additional uncorrelated noise sources such as instrumental effects. Amplitudes of both the correlated and uncorrelated terms are allowed to vary, leading to a covariance matrix, $\boldsymbol{K}_S$, with components

$$K_{S_{ij}} = \sigma_{QP,S}^2 k_{QP}(t_i, t_j) + \sigma_{WN,S}^2 I_{ij},$$
(2)

where $\sigma_{QP,S}$ and $\sigma_{WN,S}$ are the two amplitudes, and $\boldsymbol{I}$ is the identity matrix.

2.3. Fitting the S Index

With a single, time-independent parameter representing the mean value of the S index, μ_S, the set of fit parameters contains six variables:

$$\boldsymbol{\theta}_S = \{\mu_S, \sigma_{QP,S}, \sigma_{WN,S}, \tau, \eta, P_{rot}\}$$
(3)

and the GP likelihood function (Rasmussen & Williams 2006) is given by

$$\mathcal{L}(\boldsymbol{\theta}_S | \boldsymbol{s}, \boldsymbol{t}) = \frac{1}{\sqrt{\det(2\pi \boldsymbol{K}_S)}}\exp\left(-\frac{1}{2}\Delta\boldsymbol{s}^T \boldsymbol{K}_S^{-1}\Delta\boldsymbol{s}\right),$$
(4)

where $\Delta\boldsymbol{s} = \boldsymbol{s} - \mu_S$ and $\boldsymbol{s}$ is a vector containing the daily averaged S index at times $\boldsymbol{t}$ for all the S index values shown in Figure 1(a). Each parameter is assigned a uniform prior probability with upper and lower bounds encompassing realistic values, shown in Table 1. The posterior probability distribution, $p_S(\boldsymbol{\theta}_S | \boldsymbol{t})$, is then proportional to $\mathcal{L}(\boldsymbol{\theta}_S | \boldsymbol{s}, \boldsymbol{t})$ within the prior bounds and zero otherwise.
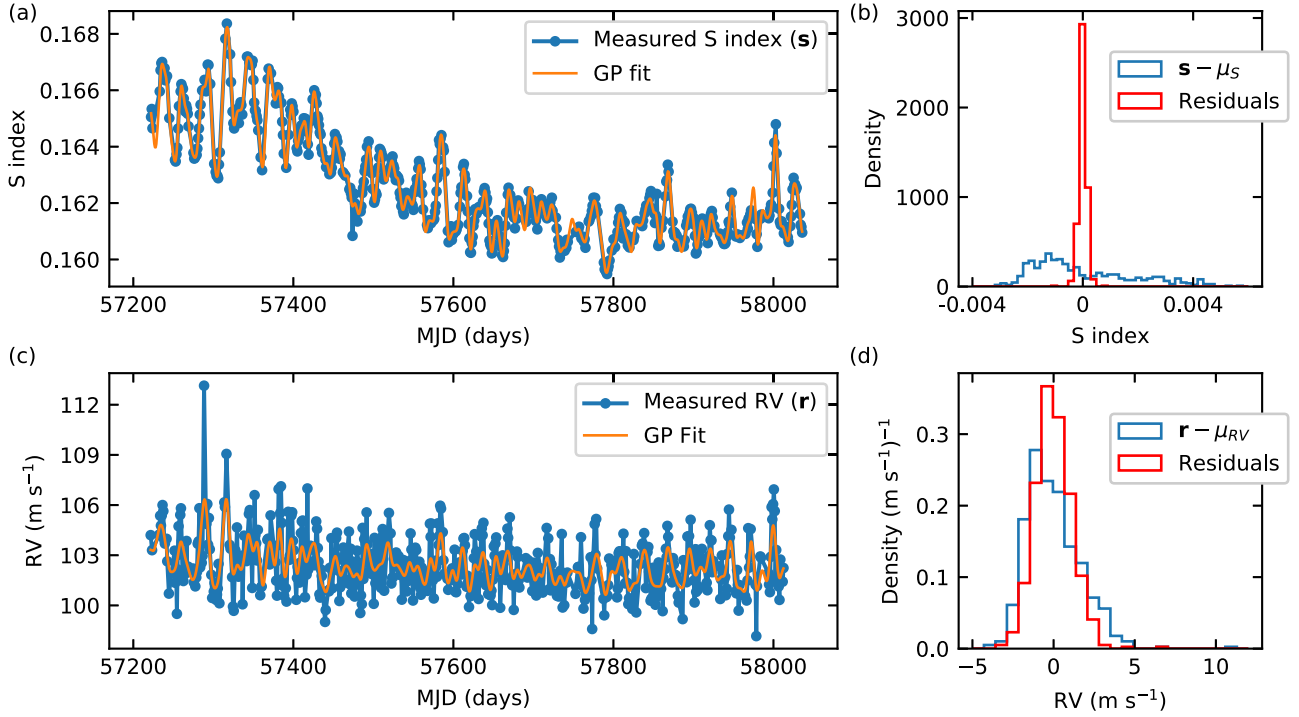
Figure 1. GP regression fits to ∼800 days of solar S index and RV data from the solar telescope and HARPS-N. (a) S index (blue) and GP regression fit (orange). (b) Histogram of mean-subtracted S index values (blue) and histogram of the residuals of the GP regression fit (red). (c) Daily average RVs (blue) and GP regression fit (orange). The mean value represents the HARPS-N instrumental offset and carries no physical meaning. (d) Histogram of mean-subtracted RVs (blue) and histogram of the residuals of the GP regression fit (red). The rms variation is reduced from 1.65 to 1.14 m s^{-1}. The relative size of the WN compared to the correlated noise is higher in the RVs than in the S index, emphasizing the need for a more sophisticated model to account for RV variation.

Table 1
Summary of Parameters Used for Markov Chain Monte Carlo Sampling for the S Index and RV Fits of Solar Telescope and HARPS-N Data

	Fit Parameter	Description	Units	Prior	MCMC Initial Guess	Fit Value
θ_S (S index parameters)	μ_S	S index mean		$U(0.15, 0.17)$	mean(s) = 0.1626	$0.1627^{+0.0002}_{-0.0002}$
	$\sigma_{QP,S}$	correlated noise amplitude		$U(10^{-5}, 10^{-1})$	std(s) = 1.83×10^{-3}	$1.43^{+0.10}_{-0.09} \times 10^{-3}$
	$\sigma_{WN,S}$	WN amplitude		$U(10^{-5}, 10^{-1})$	std(s) = 1.83×10^{-3}	$0.178^{+0.007}_{-0.007} \times 10^{-3}$
	τ	active region lifetime	days	$U(5, 100)$	22	$23.6^{+1.1}_{-1.1}$
	η	smoothing parameter		$U(0.1, 0.9)$	0.5	$0.58^{+0.04}_{-0.04}$
	P_{rot}	solar rotation period	days	$U(24, 32)$	27	$28.1^{+0.4}_{-0.5}$
θ_{RV} (RV parameters)	μ_{RV}	RV mean	m s^{-1}	$U(97, 113)$	mean(r) = 102.4	$102.4^{+0.3}_{-0.3}$
	$\sigma_{QP,RV}$	correlated noise amplitude	m s^{-1}	$U(0.01, 10)$	std(r) = 1.65	$1.44^{+0.16}_{-0.15}$
	$\sigma_{WN,RV}$	WN amplitude	m s^{-1}	$U(0.01, 10)$	std(r) = 1.65	$1.25^{+0.04}_{-0.04}$

Note. The upper and lower bounds of the uniform priors are given in addition to the initial guesses used for each parameter. The last column shows the resulting median value of the MCMC samples and their corresponding 16% and 84% quantiles as error bars.

We estimate the posterior distribution using an affine-invariant Markov chain Monte Carlo (MCMC) method, implemented with the Python (van Rossum 1995) packages NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), emcee (Foreman-Mackey et al. 2013), and George (Ambikasaran et al. 2015). Following Foreman-Mackey et al. (2013) 32 walkers are used to sample the parameter space, and are initialized with a normal distribution around mean values from a maximum-likelihood fit and known properties of the Sun. These values are summarized in Table 1. The first 100 samples are discarded allowing the walkers to converge to the posterior distribution before evaluating an additional 50,000 samples per walker. To remove correlations between samples, they are thinned by keeping one out of every n_S samples for each walker, where $n_S = 21$ is the average correlation length of the walkers as estimated by the autocorrelation. This yields a total of 76,032 uncorrelated samples for each parameter.

The resulting parameter estimates, $\hat{\theta}_S$, are displayed in the last column of Table 1 as the median value of these uncorrelated samples. The error bars are reported as the corresponding 16% and 84% quantiles. The values obtained are consistent with known solar properties. Of note is the active region lifetime, τ, which is less than one rotation period despite lifetimes of photospheric faculae typically being greater than six rotation periods. Variability of the distribution of faculae on the solar surface is expected to drive this value down below the lifespan of an individual facular region. The GP regression fit to the S index and a histogram of the residuals are shown in
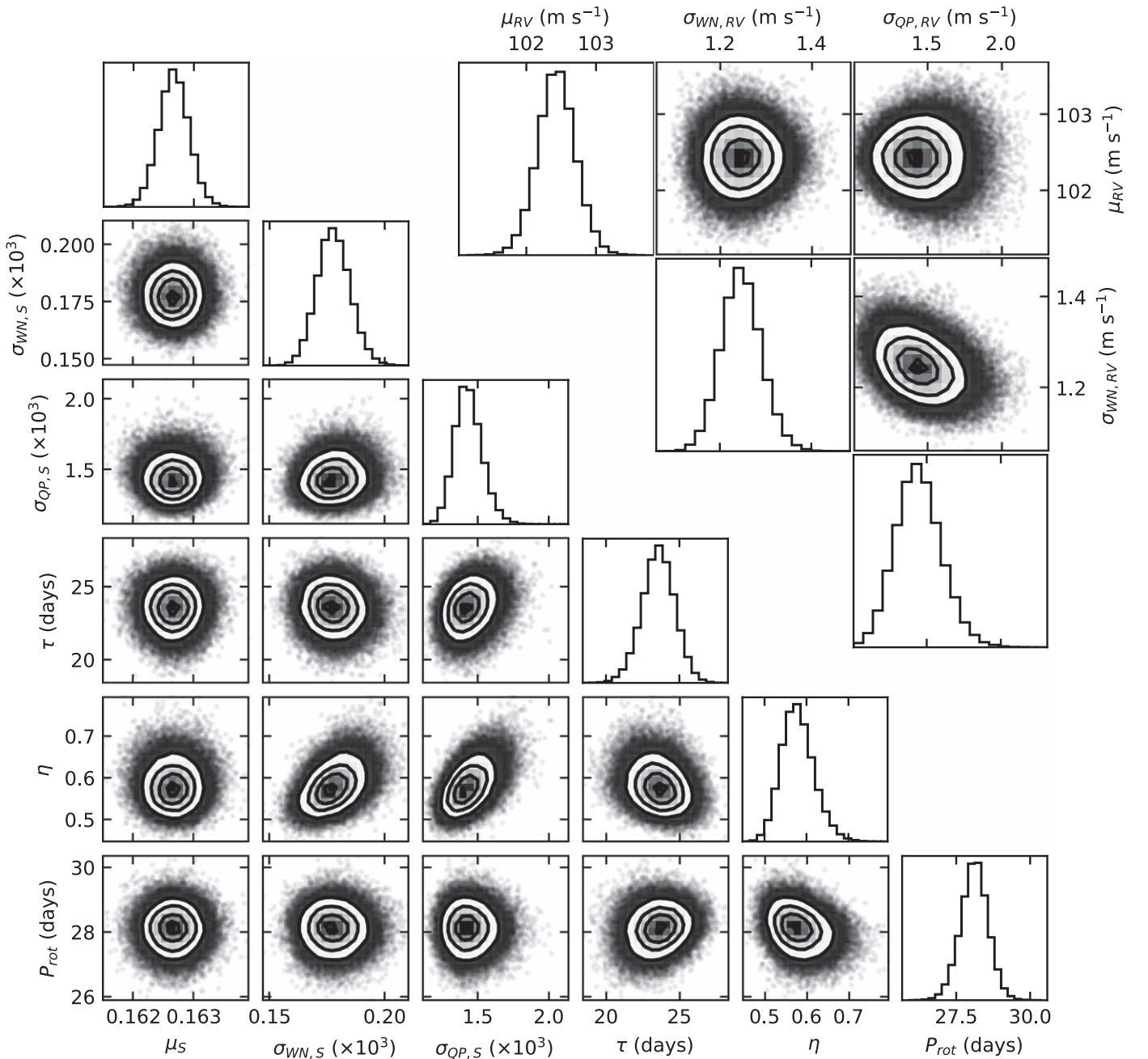
Figure 2. Marginalized distributions of the posterior distribution given by the MCMC samples. The lower left triangle shows samples (and contour lines) from the S index fit and the upper right triangle shows samples (and contour lines) from the RV fit. Each parameter is unimodal and only slight correlations exist between the parameters. Thus both posterior distributions are well behaved.

Figures 1(a), (b). The variances and covariances of the MCMC samples are shown graphically in the bottom left half of Figure 2. The largest correlations are the (η, $\sigma_{QP,S}$) and (η, $\sigma_{WN,S}$) pairs with statistically significant Pearson correlation coefficients (Student 1908) of 0.49 and 0.42, respectively. This could indicate an unaccounted for spatial dependence in the magnetic activity. Furthermore, significant correlations exist in the (η, τ) and (η, P_{rot}) pairs, potentially reflecting the migration of active regions from higher latitudes (and thus longer rotation period) to lower latitudes (and thus shorter rotation period) over the 2.2 yr of data. Additionally, $\sigma_{QP,S}$ increases slightly with τ, lending further evidence to a time dependence in the fit parameters as τ will be longer for the on average larger,

longer-lived active regions earlier in the solar magnetic cycle. These correlations suggest a more sophisticated model may more closely represent the physical mechanisms in the photosphere of the Sun and thus capture more of the RV variation due to activity. Further investigation of these correlations are left to future study.

2.4. Fitting the RVs

We assume that the magnetic activity driving the S index also affects the observed RVs with equivalent correlated variability and thus may be described with equal GP regression parameters: τ, η, and P_{rot}. Unlike Rajpaul et al. (2015) who fit the S index and RVs simultaneously, we model the S index and

RVs consecutively in order to reduce computational complexity inherent to GPs and MCMC. Although a small improvement would be expected from simultaneous fitting, we do not anticipate this would change the overall conclusion of this work. We fit the RVs using the same QP kernel function of Equation (1) with the spot lifetime, τ, spot distribution, η, and the rotation period, P_{rot}, fixed at the median value from the S index fit. Because the effects of solar system planets have already been removed from the RVs, we are left with only three fit parameters:

$$\boldsymbol{\theta}_{RV} = \{\mu_{RV}, \sigma_{QP,RV}, \sigma_{WN,RV}\} \quad (5)$$

and the likelihood function becomes

$$\mathcal{L}(\boldsymbol{\theta}_{RV}|\boldsymbol{r}, \boldsymbol{t}) = \frac{1}{\sqrt{\det(2\pi \boldsymbol{K}_{RV})}} \exp\left(-\frac{1}{2}\Delta \boldsymbol{r}^{\mathrm{T}} \boldsymbol{K}_{RV}^{-1} \Delta \boldsymbol{r}\right), \quad (6)$$

where $\Delta \boldsymbol{r} = \boldsymbol{r} - \mu_{RV}$, $\boldsymbol{r}$ is a vector of the RVs at times $\boldsymbol{t}$, and $\boldsymbol{K}_{RV}$ is populated with $\boldsymbol{\theta}_{RV}$ by substituting all "S" subscripts with "R" in Equation (2). Priors are shown in Table 1.

The posterior probability, $p_{RV}(\boldsymbol{\theta}_{RV}|\boldsymbol{r}, \boldsymbol{t})$, is proportional to the likelihood function within the bounds of the prior and zero otherwise. It is sampled using the same MCMC protocol as with the S index and we observe a correlation length of $n_{RV} = 11$. Retaining one of every n_{RV} samples leads to a total of 145,312 uncorrelated samples. The parameter estimates, $\hat{\boldsymbol{\theta}}_{RV}$, are summarized in Table 1 and the resulting GP regression fit and residuals are shown in Figures 1(c) and (d). The fit reduces the rms scatter in the data from 1.65 to 1.14 m s^{-1}, consistent with Milbourne et al. (2019), Miklos et al. (2020), and other solar analyses (see Section 5). The marginalized distributions are shown in Figure 2 and again display a well-behaved posterior distribution. The only significant correlation exists between $\sigma_{QP,RV}$ and $\sigma_{WN,RV}$ and is negative, which is expected as these parameters will trade off the amount of variation seen in the RVs.

3. Sensitivity Map

3.1. Synthetic Planet Model

To explore the limits of our GP regression for detecting low-mass, long-period exoplanets, we inject synthetic planets with varying Keplerian parameters into the solar RVs. The general Doppler-induced RV, $v_{RV}(t)$, of a host star by a companion planet is given by

$$v_{RV}(t) = K_{pl}[\cos(\omega + \nu(t|t_p, P_{orb}, e)) + e\cos(\omega)], \quad (7)$$

where K_{pl} is the semi-amplitude, ω is the argument of periastron, e is the eccentricity, t_p is the time of pericenter passage, P_{orb} is the orbital period, and $\nu(t|t_p, P_{orb}, e)$ is the true anomaly (Perryman 2011). As a best-case scenario, we restrict ourselves to circular orbits (i.e., $e = 0$), which simplify the Doppler shift to the sine function

$$v_{RV}(t) = K_{pl}\sin\left(\frac{2\pi t}{P_{orb}} + \phi\right), \quad (8)$$

where ϕ is an arbitrary phase. A synthetic planet is then generated by choosing values for the semi-amplitude, orbital period, and phase followed by adding $v_{RV}(t)$ to the vector of measured solar RVs, $\boldsymbol{r}$, observed at times, $\boldsymbol{t}$. The vector of the

RVs thus undergoes the transformation

$$\boldsymbol{r} \rightarrow \boldsymbol{r} + v_{RV}(\boldsymbol{t}). \quad (9)$$

The GP regression now contains a mean function given by the addition of the overall mean value of the RVs, μ_{RV}, along with the Keplerian parameters required to describe the injected circular planetary orbit. This list of fit parameters, $\boldsymbol{\theta}_{RV}$, then becomes

$$\boldsymbol{\theta}_{RV} = \{K_{pl}, P_{orb}, \phi, \mu_{RV}, \sigma_{QP,RV}, \sigma_{WN,RV}\} \quad (10)$$

and the likelihood function is still given by Equation (6), with the exception that the vector of the fit residuals becomes $\Delta \boldsymbol{r} = \boldsymbol{r} - (\mu_{RV} + v_{RV}(\boldsymbol{t}))$ with $\boldsymbol{t}$ as the vector of observation times and v_{RV} as the function defined in Equation (8).

3.2. Retrieval of Injected Signals

Using the techniques of the previous section, we construct a map of detection sensitivities for a range of synthetic, low-mass, long-period planets. Our grid contains 210 injected planets with semi-amplitudes from 0.6 to 2.4 m s^{-1} in 0.2 m s^{-1} steps and orbital periods from 100 to 500 days in 20 days steps. The phase of each planet is drawn from a uniform distribution, $\phi \sim U(-0.1, 0.1)$ radians, and the prior on the phase is uniform from $-\pi$ to π radians. This is done to avoid numerical instabilities associated with the phase occurring near the boundary of the prior. The prior on the semi-amplitude allows only positive values less than 10 m s^{-1} and the prior on the orbital period is in the range $(\frac{1}{2}P_{orb}, 2P_{orb})$. For planets detected with a high degree of statistical significance, the priors are uninformative and do not affect the results. However, for the lowest mass planets with 1σ detections or less, the priors do constrain the results as described below. These planets set the lower bound on the range of injected semi-amplitudes used in this analysis.

We draw MCMC samples, exploring the semi-amplitude and orbital period linearly, in the same fashion as the previous section with the non-Keplerian parameter priors unchanged. We again define the parameter estimates, $\hat{\boldsymbol{\theta}}_{RV}$, as the median value of the uncorrelated samples and the corresponding lower and upper bounds, $\hat{\boldsymbol{\theta}}_{RV,lower}$ and $\hat{\boldsymbol{\theta}}_{RV,upper}$, as the 16% and 84% quantiles. The statistical significance, SS_i, is defined as

$$SS_i = \frac{|\hat{\theta}_i|}{\frac{1}{2}(\hat{\theta}_{i,upper} - \hat{\theta}_{i,lower})}, \quad (11)$$

where θ_i is a given parameter from the vector defined in Equation (10). We plot the statistical significance of the recovered orbital period and semi-amplitude of the 210 synthetic planets in Figure 3. The orbital period is determined with a high degree of statistical significance, though for semi-amplitudes below 1 m s^{-1} this should be taken as an upper bound as the MCMC samples begin to encounter the edges of the uniform prior. The semi-amplitude, however, is much less certain. For a 5σ "discovery" threshold, planets with a semi-amplitude less than 1 m s^{-1} would require more observations than the 800 days of solar telescope data used in this analysis. The structure of the contours in Figure 3(a) is likely due to our noncontinuous observing schedule. In particular, Nava et al. (2019) find that activity-driven signals at orbital periods unrelated to either planetary orbital periods or the stellar
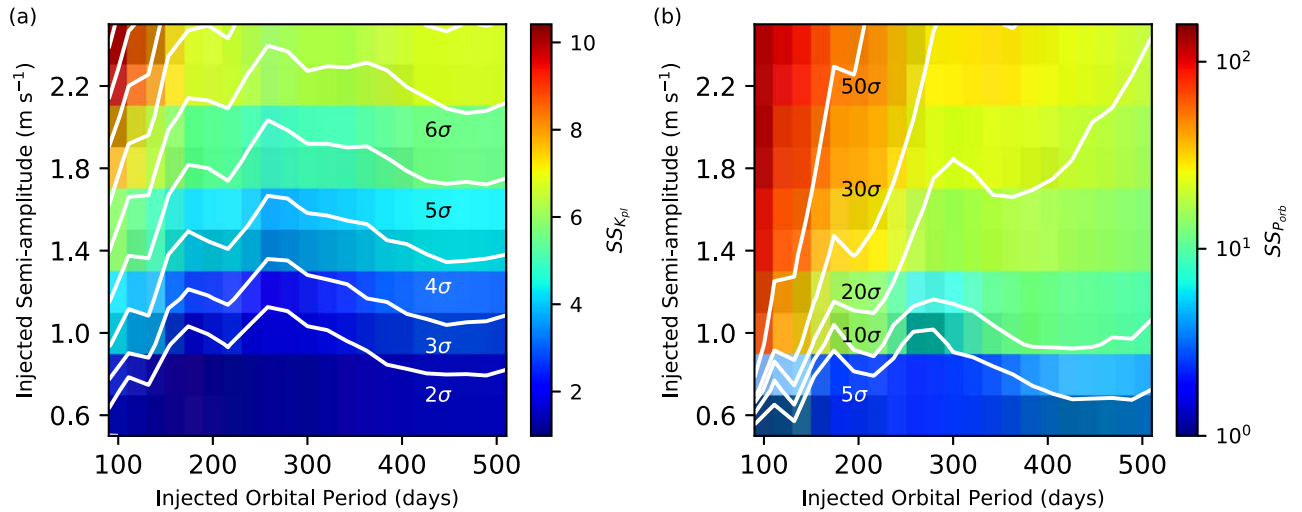
Figure 3. Sensitivity maps of the recovered (a) semi-amplitude, K_{pl}, and (b) orbital period, P_{orb}, of synthesized planets using the GP regression with synthetic planetary signals injected into the 800 days of solar RVs. Color bars show the statistical significance (SS) of the recovered parameters as defined in Equation (11). The white lines in panel (a) show contours of 2σ, 3σ, ... statistical significance and the white lines in panel (b) show contours of 5σ, 10σ, 20σ, 30σ, and 50σ statistical significance. Each "pixel" in either image represents one of the 210 simulated planets. The orbital period is generally recovered with a high degree of confidence but the semi-amplitude is only recovered at the 5σ level for planets with semi-amplitude above 1 m s^{-1}.

rotation period can arise from uneven sampling. Thus even the near-daily, long-baseline observing schedule of the solar telescope decreases the semi-amplitude sensitivity (Hall et al. 2018). We find agreement within parameter uncertainties between the injected and extracted orbital periods and semi-amplitudes, suggesting no systematic effects induced by the GP regression. In the next section we explore the baseline of observations required to detect sub-m s^{-1} planets using completely synthetic RVs.

4. Synthetic RVs

To study longer observing baselines than in the previous section, we synthesize not only Keplerian Doppler shifts, but also the solar RVs themselves. We thus extend the observing baseline to determine the requirements for detecting a low-mass, long-period analog. A synthetic planet with the orbital period of Venus (225 days) is injected with varying semi-amplitude. We use the orbital period of Venus to avoid systematic biases associated with measuring any periodic signal at 365 days.

4.1. Observing Schedule

The first step to synthesize solar telescope data is to create a realistic observing schedule to account for inclement weather and telescope downtime. We assume a naive model where each calendar month, m, is assigned an observation probability, $p_{obs}(m)$, which represents the fraction of days that have observations in month m. These probabilities are estimated using historical solar telescope data with the corresponding fraction of days with 10 or more observations. These probabilities are shown in red in Figure 4(a).

For any given day d in month m, a uniform random number, $r(d) \sim U(0, 1)$, is drawn and compared to $p_{obs}(m)$. If $r(d) \leqslant p_{obs}(m)$, an observation occurs. This process is repeated for as many consecutive days as needed. Using this process, we create a synthetic 30 yr observing schedule mimicking the solar telescope seasonal variations. The resulting synthetic observing schedule is shown in blue in Figure 4(a).

4.2. Observation Times

The second step for simulating data is to generate an observation time for each day with an observation. We begin by computing the mean observation time of all exposures in each day from historical solar telescope data, limiting only to days with at least 10 exposures. This distribution of mean observation times is histogrammed and fit to a Gaussian kernel density estimate (KDE; Scott 2015). Synthetic observation times are drawn from this KDE. Figure 4(b) shows the distribution of mean telescope exposure times and the resulting observation times drawn from the KDE. These observation times coupled with the observing days from the previous section completely determine the observing schedule of the simulated data.

4.3. Synthesized RVs

Finally we synthesize the RVs in two steps. The first step repeats the procedure from Section 3 whereby a Keplerian term given by Equation (8) is generated with several test semi-amplitudes and a phase drawn randomly in the interval $(-0.1, 0.1)$ radians. The orbital period, however, is fixed at 225 days. The second step involves modeling the magnetic activity by using a GP with the same kernel function given in Equation (2). The parameter values used to populate the covariance matrix are taken from the GP regression fit of the solar data as given in Table 1. A random sample of RVs is then drawn from the GP using these parameter values and the synthetic observation times generated in the previous section. The sum of the Keplerian term and the GP term constitute the full RV synthesis. Note that the solar telescope RVs only cover a fraction of the solar magnetic cycle during solar minimum. As such we do not model RV variability during high activity levels. Thus, our resulting synthetic RVs represent a best-case scenario, as other Sun-like stars tend to have higher levels of magnetic activity than the Sun (Reinhold et al. 2020). However, we emphasize that many recent studies successfully treated stellar variability by applying GP regression in RV variations of stars much more active than the Sun, such as the
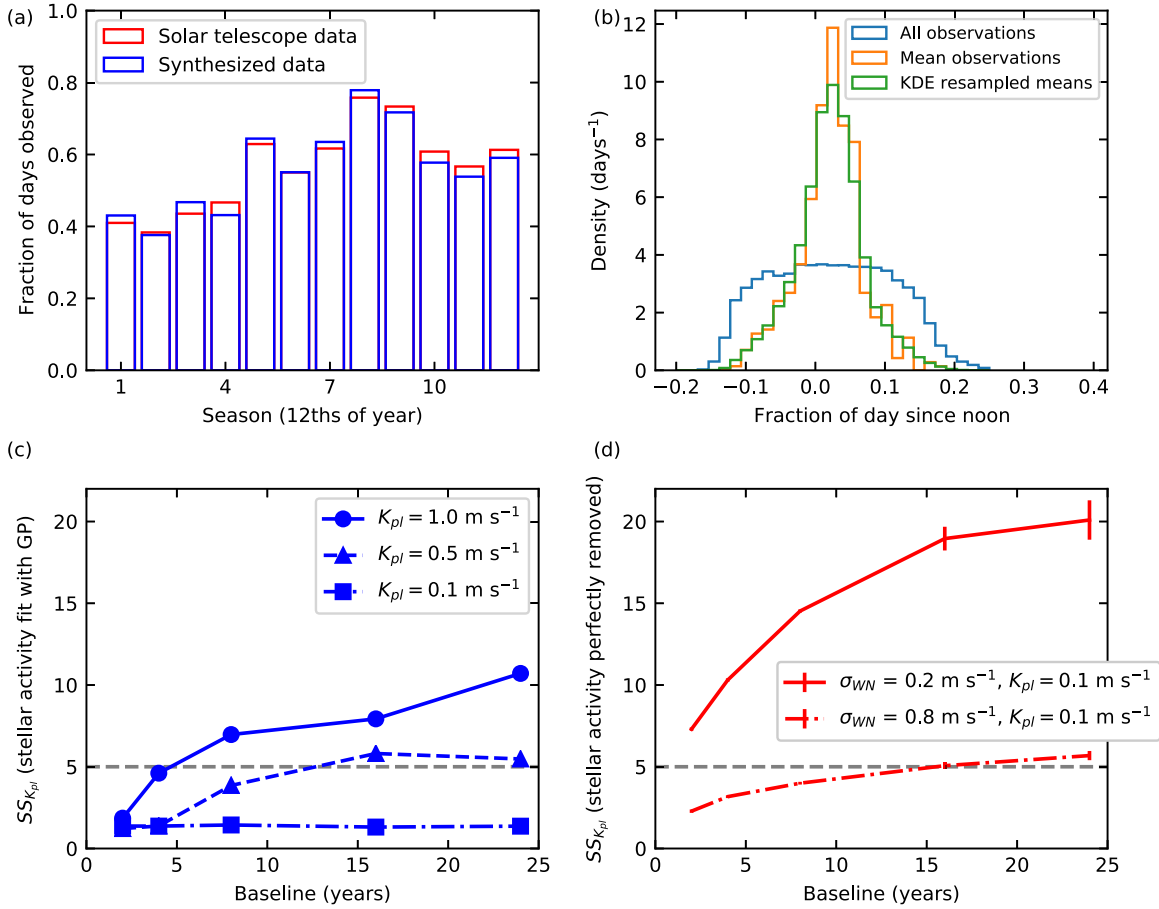
Figure 4. Generating synthetic RVs based on historical solar telescope observations. (a) Seasonal variation of solar telescope and synthetic observations. Red histogram shows the fraction of days with at least 10 5-minute exposures for 12 evenly spaced sections of a year (does not correspond to calendar months). Blue histogram shows the analogous ratio of seasonal observations from dates synthesized from the red histogram. (b) Histogram of RV observation times given as a fraction of a day since noon UTC: blue—all 5 minute solar telescope exposures, orange—daily averaged observation times for days with at least 10 5-minute exposures, and green—synthesized observation times using a Gaussian kernel density estimate. (c) Statistical significance (SS), defined in Equation (11), of the recovered semi-amplitude, K_{pl}, using a GP regression for an exo-Venus (225 day orbital period) as a function of baseline of data. Solid (circles) and dashed (triangles) curves show the SS for an exo-Venus with semi-amplitudes of 1.0 m s^{-1} and 0.5 m s^{-1}, respectively. Dashed–dotted (squares) curve shows an upper bound for a semi-amplitude of 0.1 m s^{-1}. Dashed gray line shows the 5σ detection threshold. (d) SS for a 0.1 m s^{-1} exo-Venus with stellar variability perfectly removed (i.e., only WN). Solid and dashed curves show the SS for WN amplitudes of 0.2 m s^{-1} and 0.8 m s^{-1}, respectively. Error bars are shown as vertical lines and are too small to be seen for baselines less than 15 yr. Again, the dashed gray line shows the 5σ detection threshold.

young pre-main-sequence stars V830 Tau (Donati et al. 2016) and AU Mic (Klein et al. 2020). In both studies, they were able to detect and characterize planetary signals with amplitudes five times smaller than the intrinsic RV variations. Therefore, we expect that an analysis including a detailed model of the solar magnetic cycle would not appreciably change the conclusions of this work, but confirmation is left to future study when precise RVs have been measured across a full solar cycle.

We then draw MCMC samples and define the parameter estimates and lower and upper bounds as before to determine the statistical significance of the recovered semi-amplitude for baselines of data equal to 2, 4, 8, 16, and 24 yr. We repeat this process for semi-amplitudes equal to 1, 0.5, and 0.1 m s^{-1}. The resulting statistical significances are shown in Figure 4(c). The dashed–dotted curve representing the 0.1 m s^{-1} planet represents an upper bound, with the prior probability distribution on the semi-amplitude restricting its value to be positive. Even for a 0.5 m s^{-1} planet, we determine that between 10 and 15 yr of data are needed to reach the 5σ discovery threshold using this GP regression and data similar to that of the HARPS-N solar telescope.

As shown above, a multidecade temporal baseline is required to detect a temperate, low-mass planet orbiting a Sun-like star using this GP regression. Thus we next assume that we have a direct measurement of the magnetic variability of the target star and are able to perfectly remove the effects of variability from the RVs, leaving only the planetary signals and WN. We can then fit to a sine curve with simple least-squares methods. We expect the semi-amplitude uncertainty, $\sigma_{K_{pl}}$, to scale with the measurement sensitivity, σ_{RV}, divided by the square root of the number of RV measurements, N (Cloutier et al. 2018),

$$\sigma_{K_{pl}} = \sigma_{RV}\sqrt{\frac{2}{N}}. \quad (12)$$

To confirm this scaling, we synthesize RV time series comprising simple WN and a Keplerian Doppler shift of $K_{pl} = 0.1$ m s^{-1} for an injected planet with an orbital period of 225 days (i.e., an exo-Venus). We generate this data for varying baselines and WN levels of both 0.8 and 0.2 m s^{-1}. The statistical significance of the semi-amplitude, K_{pl}, from least-squares fits to these data sets are shown in Figure 4(d) and are in good agreement with expectations set by Equation (12).

We emphasize that we need to reduce the WN to levels approaching $0.2 \, \mathrm{m \, s^{-1}}$ to reach the 5σ detection threshold in only a few years for a true exo-Venus.

5. Discussion and Conclusions

The GP regression of Section 2 reduces the rms variation of the solar RVs from 1.65 to $1.14 \, \mathrm{m \, s^{-1}}$. However, treating stellar variability with this GP regression still requires 10–15 yr of densely sampled RV observations to detect long-period, low-mass planets. This is much longer than would be expected if the RVs contained purely uncorrelated WN. This result is in line with more physically motivated techniques. For example, Milbourne et al. (2019) used magnetograms and Dopplergrams from the Helioseismic and Magnetic Imager (HMI) on board the Solar Dynamics Observatory to derive activity-driven RV time series. By modeling the HARPS-N RVs using these activity time series, they reduced the RV rms from 1.65 to $1.21 \, \mathrm{m \, s^{-1}}$. Haywood et al. (2020) reduced the RV rms to $0.85 \, \mathrm{m \, s^{-1}}$ by modeling the RVs with a linear combination of the unsigned magnetic flux from HMI and the total solar irradiance, using the FF′ method (Aigrain et al. 2012). Fitting our GP regression to the unsigned flux may yield a smaller variation in the RV residuals, and this will be investigated in a future work. Dumusque (2018) and Cretignier et al. (2020) used logistic regression and gradient boosting on HARPS-N solar spectra to differentiate activity sensitive and insensitive lines. Computing RVs from these sets of lines reduced the RV rms to $0.9 \, \mathrm{m \, s^{-1}}$. Similarly, Miklos et al. (2020) applied the techniques of Meunier et al. (2017) to the HARPS-N solar RVs to estimate magnetoconvective RV variations. This analysis, however, was unable to reduce the observed rms RVs.

That the GP regression—a statistical tool—performs similarly to these more physically motivated analyses demonstrates its power: rather than requiring high-resolution solar images or specially tuned line lists, the GP regression assumes a correlation between the solar S index and observed RVs. However, the fact that the GP analysis and these physically motivated techniques arrive at the same approximately $1 \, \mathrm{m \, s^{-1}}$ RV uncertainty level indicates that something else—either another physical process operating on a different timescale or an instrumental systematic—is limiting the performance of these techniques. Haywood et al. (2020) discuss some physical effects missing from the current models, to be addressed in future work. However for current data analysis models, the solar data set used in this work, with nearly 2.5 yr of near-daily observations mostly in the decline phase of Cycle 24, represents a best-case scenario; we can therefore only perform 5σ recoveries of long-period planets with semi-amplitudes greater than $1 \, \mathrm{m \, s^{-1}}$, consistent with the results of the recent community-wide RV challenge of Dumusque et al. (2017).

Our work using synthetic RVs indicates that more observations will not quickly overcome this limit: as shown in Figure 4, it will take 10–15 yr to reach 5σ on a $0.5 \, \mathrm{m \, s^{-1}}$ RV signal with a 225 days period (i.e., the orbital period of Venus), and at least 25 yr for a $0.1 \, \mathrm{m \, s^{-1}}$ RV signal. The last panel of this figure also indicates that a perfect model of activity-driven correlated variations would not in itself suffice for the rapid detection of an exo-Venus or an exo-Earth; even in the absence of correlated noise, a current-generation spectrograph with a long-term stability of about $0.8 \, \mathrm{m \, s^{-1}}$ would need a 10–15 yr observing baseline to reach a 5σ detection of an exo-Earth.

Successful exo-Earth discovery therefore requires both more sophisticated models of stellar variability and the improved RV precision, long-term stability, and dense observational sampling from a next-generation spectrograph (Wright & Robertson 2017) such as the Echelle SPectrograph for Rocky Exoplanets and Stable Spectroscopic Observations (ESPRESSO; Pepe et al. 2013, 2021; Damasso et al. 2020; Suárez Mascareño et al. 2020), NEID (Allen et al. 2018), EXtreme PREcision Spectrometer (EXPRES; Jurgenson et al. 2016; Blackman et al. 2020; Brewer et al. 2020; Petersburg et al. 2020), HARPS3 (Thompson et al. 2016), or GMT Consortium Large Earth Finder (G-CLEF; Szentgyorgyi et al. 2014).

Facilities: TNG: HARPS-N (Cosentino et al. 2014), Solar Telescope (Dumusque et al. 2015; Phillips et al. 2016).

Software: HARPS-N Data Reduction Software (Baranne et al. 1996; Sosnowska et al. 2012), Python (van Rossum 1995): NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), emcee (Foreman-Mackey et al. 2013), George (Ambikasaran et al. 2015).

ORCID iDs

N. Langellier ● https://orcid.org/0000-0003-2107-3308
T. W. Milbourne ● https://orcid.org/0000-0001-5446-7712
D. F. Phillips ● https://orcid.org/0000-0001-5132-1339
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
L. Malavolta ● https://orcid.org/0000-0002-6492-2085
S. Thompson ● https://orcid.org/0000-0002-8039-194X
A. Collier Cameron ● https://orcid.org/0000-0002-8863-7828
X. Dumusque ● https://orcid.org/0000-0002-9332-2011
H. M. Cegla ● https://orcid.org/0000-0001-8934-7315
D. W. Latham ● https://orcid.org/0000-0001-9911-7388
J. Maldonado ● https://orcid.org/0000-0002-2218-5689
N. Buchschacher ● https://orcid.org/0000-0002-3697-1541
M. Cecconi ● https://orcid.org/0000-0001-5701-2529
D. Charbonneau ● https://orcid.org/0000-0002-9003-484X
R. Cosentino ● https://orcid.org/0000-0003-1784-1431
A. Ghedina ● https://orcid.org/0000-0003-4702-5152
M. Lodi ● https://orcid.org/0000-0002-5432-9659
M. López-Morales ● https://orcid.org/0000-0003-3204-8183
G. Micela ● https://orcid.org/0000-0002-9900-4751
E. Molinari ● https://orcid.org/0000-0002-1742-7735
E. Poretti ● https://orcid.org/0000-0003-1200-0473
K. Rice ● https://orcid.org/0000-0002-6379-9185
D. Sasselov ● https://orcid.org/0000-0001-7014-1771
A. Sozzetti ● https://orcid.org/0000-0002-7504-365X
S. Udry ● https://orcid.org/0000-0001-7576-6236
R. L. Walsworth ● https://orcid.org/0000-0003-0311-4751

References

Aigrain, S., Pont, F., & Zucker, S. 2012, MNRAS, 419, 3147
Allen, L. E., Wright, J., Rajagopal, J., et al. 2018, AAS Meeting, 231, 246.08
Ambikasaran, S., Foreman-Mackey, D., Greengard, L., Hogg, D. W., & O'Neil, M. 2015, ITPAM, 38, 252
Baranne, A., Queloz, D., Mayor, M., et al. 1996, A&AS, 119, 373
Blackman, R. T., Fischer, D. A., Jurgenson, C. A., et al. 2020, AJ, 159, 238
Brewer, J. M., Fischer, D. A., Blackman, R. T., et al. 2020, AJ, 160, 67
Cegla, H. M. 2019, Geosc, 9, 114
Cloutier, R., Doyon, R., Bouchy, F., & Hébrard, G. 2018, AJ, 156, 82
Collier Cameron, A., Mortier, A., Phillips, D., et al. 2019, MNRAS, 487, 1082
Cosentino, R., Lovis, C., Pepe, F., et al. 2014, Proc. SPIE, 9147, 91478C
Cretignier, M., Francfort, J., Dumusque, X., Allart, R., & Pepe, F. 2020, A&A, 640, A42
Damasso, M., Sozzetti, A., Lovis, C., et al. 2020, A&A, 642, A31
Damasso, M., Zeng, L., Malavolta, L., et al. 2019, A&A, 624, A38
Donati, J. F., Moutou, C., Malo, L., et al. 2016, Natur, 534, 662
Dumusque, X. 2018, A&A, 620, A47
Dumusque, X., Boisse, I., & Santos, N. C. 2014, ApJ, 796, 132
Dumusque, X., Borsa, F., Damasso, M., et al. 2017, A&A, 598, A133
Dumusque, X., Glenday, A., Phillips, D. F., et al. 2015, ApJL, 814, L21
Faria, J. P., Haywood, R. D., Brewer, B. J., et al. 2016, A&A, 588, A31
Fischer, D. A., Anglada-Escude, G., Arriagada, P., et al. 2016, PASP, 128, 066001
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Giles, H. C. A., Collier Cameron, A., & Haywood, R. D. 2017, MNRAS, 472, 1618
Grunblatt, S. K., Howard, A. W., & Haywood, R. D. 2015, ApJ, 808, 127
Hall, R. D., Thompson, S. J., Handley, W., & Queloz, D. 2018, MNRAS, 479, 2968
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
Haywood, R. D., Collier Cameron, A., Queloz, D., et al. 2014, MNRAS, 443, 2517
Haywood, R. D., Milbourne, T. W., Saar, S. H., et al. 2020, arXiv:2005.13386
Isaacson, H., & Fischer, D. 2010, ApJ, 725, 875
Jurgenson, C., Fischer, D., McCracken, T., et al. 2016, Proc. SPIE, 9908, 99086T
Kjeldsen, H., & Bedding, T. R. 1995, A&A, 293, 87
Klein, B., Donati, J.-F., Moutou, C., et al. 2020, MNRAS, 502, 188
Linsky, J. L., & Avrett, E. H. 1970, PASP, 82, 169
Meunier, N., Lagrange, A.-M., & Borgniet, S. 2017, A&A, 607, A6
Meunier, N., Lagrange, A.-M., & Desort, M. 2010, A&A, 519, A66
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
Motalebi, F., Udry, S., Gillon, M., et al. 2015, A&A, 584, A72
Nava, C., López-Morales, M., Haywood, R. D., & Giles, H. A. C. 2019, AJ, 159, 23
Pepe, F., Cristiani, S., Rebolo, R., et al. 2013, Msngr, 153, 6
Pepe, F., Cristiani, S., Rebolo, R., et al. 2021, A&A, 645, A96
Perryman, M. 2011, The Exoplanet Handbook (Cambridge: Cambridge Univ. Press)
Petersburg, R. R., Ong, J. M. J., Zhao, L. L., et al. 2020, AJ, 159, 187
Phillips, D. F., Glenday, A. G., Dumusque, X., et al. 2016, Proc. SPIE, 9912, 9912
Rajpaul, V., Aigrain, S., Osborne, M. A., Reece, S., & Roberts, S. 2015, MNRAS, 452, 2269
Rajpaul, V., Aigrain, S., & Roberts, S. 2016, MNRAS, 456, L6
Rasmussen, C. E., & Williams, C. K. I. 2006, Gaussian Processes for Machine Learning (2nd ed.; Cambridge, MA: MIT Press)
Reinhold, T., Shapiro, A. I., Solanki, S. K., et al. 2020, Sci, 368, 518
Robertson, P., Mahadevan, S., Endl, M., & Roy, A. 2014, Sci, 345, 440
Robertson, P., Stefansson, G., Mahadevan, S., et al. 2020, ApJ, 897, 125
Scott, D. W. 2015, Multivariate Density Estimation: Theory, Practice, and Visualization (2nd ed.; New York: Wiley)
Sosnowska, D., Lodi, M., Gao, X., et al. 2012, Proc. SPIE, 8451, 8451
Student 1908, Biometrika, 6, 302
Suárez Mascareño, A., Faria, J. P., Figueira, P., et al. 2020, A&A, 639, A77
Szentgyorgyi, A., Barnes, S., Bean, J., et al. 2014, Proc. SPIE, 9147, 914726
Thompson, S. J., Queloz, D., Baraffe, I., et al. 2016, Proc. SPIE, 9908, 99086F
van Rossum, G. 1995, Python Tutorial, Tech. Rep. CS-R9526, https://ir.cwi.nl/pub/5007/05007D.pdf
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Wilson, O. C. 1968, ApJ, 153, 221
Wright, J. T., & Robertson, P. 2017, RNAAS, 1, 51