ANNUAL REVIEWS

Annual Review of Statistics and Its Application

Statistical Methods for Exoplanet Detection with Radial Velocities

Nathan C. Hara[1] and Eric B. Ford[2]

[1]Département d'Astronomie, Université de Genève, Versoix, Switzerland; email: nathan.hara@unige.ch

[2]Department of Astronomy, Center for Exoplanets and Habitable Worlds, Institute for Computational and Data Sciences, and Center for Astrostatistics, The Pennsylvania State University, University Park, Pennsylvania, USA; email: ebf11@psu.edu

Annu. Rev. Stat. Appl. 2023. 10:623–49

First published as a Review in Advance on November 22, 2022

The *Annual Review of Statistics and Its Application* is online at statistics.annualreviews.org

https://doi.org/10.1146/annurev-statistics-033021-012225

Keywords

correlated noise, exoplanets, model comparison, model misspecification, multivariate time series, uneven sampling

Abstract

Exoplanets can be detected with various observational techniques. Among them, radial velocity (RV) has the key advantages of revealing the architecture of planetary systems and measuring planetary mass and orbital eccentricities. RV observations are poised to play a key role in the detection and characterization of Earth twins. However, the detection of such small planets is not yet possible due to very complex, temporally correlated instrumental and astrophysical stochastic signals. Furthermore, exploring the large parameter space of RV models exhaustively and efficiently presents difficulties. In this review, we frame RV data analysis as a problem of detection and parameter estimation in unevenly sampled, multivariate time series. The objective of this review is two-fold: to introduce the motivation, methodological challenges, and numerical challenges of RV data analysis to nonspecialists, and to unify the existing advanced approaches in order to identify areas for improvement.

1. INTRODUCTION

Until recently, our Solar System was the only laboratory in which to study theories of planetary formation and evolution and to search for life beyond Earth. The presence of planets outside the Solar System, or exoplanets, was uncertain until the detection of planets orbiting a pulsar (Wolszczan & Frail 1992). The study of exoplanets as a scientific field accelerated with the detection of 51 Peg b, a planet of minimum ∼0.5 Jupiter mass orbiting a Sun-like star in 4.2 days (Mayor & Queloz 1995).

The discovery of 51 Peg b, and hundreds of additional exoplanets, was based on the radial velocity (RV) method, which is the focus of the present work, and relies on the following principle. A star around which a planet revolves has a periodic reflex motion, so it moves back and forth toward the observer with a certain velocity: its RV. By acquiring spectra of a given star at different times, the observer can measure a time series of RV through the Doppler effect. The amplitude of the variations in RV is proportional to the planetary mass, and its shape depends on the orbital eccentricity.

Mass is one of the most fundamental parameters of a planet. When the radius is measured separately via photometry, the combination of mass and radius gives the density, essential to characterizing the internal structure of the planet, and the surface gravity, key to interpreting measurements of its atmosphere (Kempton et al. 2018, Batalha et al. 2019). The eccentricity is important to understand the formation of planetary systems, especially in multiplanetary systems (e.g., Jurić & Tremaine 2008).

The RV method plays a key role in understanding the demographics of planetary systems, by allowing the detection of planets spanning a much wider range of orbital periods than other detection techniques. Moreover, it does not require a precise orientation of the orbital plane, thereby allowing RVs to characterize planetary systems in which the planets have significant mutual inclinations. As such, RV measurements have a unique potential to reveal the architecture of planetary systems with periods of up to ∼10 years (Fulton et al. 2021, Rosenthal et al. 2021).

One long-term goal of the scientific community working on exoplanets is to detect and characterize the atmospheres of a population of potentially Earth-like planets and, in particular, to search for evidence of life. Characterizing the atmosphere of Earth-like planets is beyond the reach of current facilities, but projects of ground based instruments such as PCS (Planetary Camera and Spectrograph) (Kasper et al. 2021) for ELT (Extremely Large Telescope), ANDES (ArmazoNes high Dispersion Echelle Spectrograph) for ELT, and direct imaging space mission concepts [LUVOIR (Large Ultraviolet Optical Infrared Surveyor) (**https://asd.gsfc.nasa.gov/luvoir/resources/**), HabEx (Habitable Exoplanet Observatory) (**https://www.jpl.nasa.gov/habex/documents/**), or LIFE (European Space Agency Large Interferometer For Exoplanets) (Quanz et al. 2021)] hope to achieve this by 2050.

Measuring the mass of terrestrial planets to an accuracy of 20% or better is essential to interpreting direct imaging and atmospheric observations (Batalha et al. 2019). Additionally, detecting Earth-like planets with RV measurements prior to the start of a future direct imaging mission would substantially increase its yield (Crass et al. 2021). The RV technique is poised to play a key role in the detection and characterization of Earth-like planets around Sun-like stars, but we do not yet know if it can achieve the required accuracy. The Earth produces a velocity variation of the Sun of $RV_\oplus = 9$ cm/s, which we use as a standard unit. The latest generation of spectrographs [e.g., ESPRESSO (Echelle SPectrograph for Rocky Exoplanets and Stable Spectroscopic Observations), EXPRES (EXtreme PREcision Spectrometer), NEID (NN-EXPLORE Exoplanet Investigations with Doppler spectroscopy)] have demonstrated a precision of ∼27 cm/s or 3 $RV_\oplus$ on a few nights. If the measurement noise were independent, then 200 measurements would suffice to measure the mass of an Earth twin with 20% precision, but intrinsic stellar variability causes complex, temporally correlated RV signal of the order of at least 5 $RV_\oplus$ (the RV of

Radial velocity (RV): velocity of a given star projected onto the direction observer–star

Stellar spectrum: number of photons collected from a given star per wavelength

Photometry: number of photons collected on a celestial body, here a star, in a broad spectral band as a function of time

Spectrograph: instrument used to measure the light flux as a function of its wavelength; can be used to measure the radial velocity of stars

Transit: passage of a planet between a star and an observer; results in a periodic dip in the flux received

the Sun as measured with the HARPS-N (High Accuracy Radial velocity Planet Searcher for the Northern hemisphere) spectrograph has a 22 $RV_\oplus$ standard deviation). Characterizing Earth-like planets requires better methods for mitigating stellar variability and instrument systematics. This can be viewed as a particular instance of a general problem: detecting and characterizing periodic signals in multivariate, unevenly sampled time series corrupted by complex noise.

The present work overviews the efforts undertaken to make Earth twins detectable. The RV data products are presented in Section 2. We present an overview of the different challenges in Section 3. We then present RV analysis in a recursive manner: Given summary statistics extracted from raw data and a statistical model of these, the question of how one detects planets and estimates their orbital elements is discussed in Section 4. Given the summary statistics, the question of how one builds a statistical model of the data is treated in Section 5. How to exploit information in lower data products to extract accurate RVs and useful summary statistics is discussed in Section 6.

2. DATA AND MODEL

2.1. Doppler Shifts

The RV of a star, its velocity projected onto the line of sight, can be measured due to the Doppler effect. If a source emits a photon with wavelength λ_0 and has a velocity $\mathbf{v}$ of modulus v relative to an observer, then the wavelength of the photon received is given by

$$\lambda = \lambda_0 \frac{1 + \frac{1}{c}\mathbf{k} \cdot \mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}}, \qquad 1.$$

where $\cdot$ is the scalar product and $\mathbf{k}$ is the unit vector from the observer to the source (Einstein 1905). The spectrum of a star contains thousands of absorption lines, short intervals of wavelengths for which photons are absorbed in the upper parts of stellar atmospheres. As the star moves, the Doppler effect causes the apparent wavelength of spectral lines to change (see **Figure 1**). An

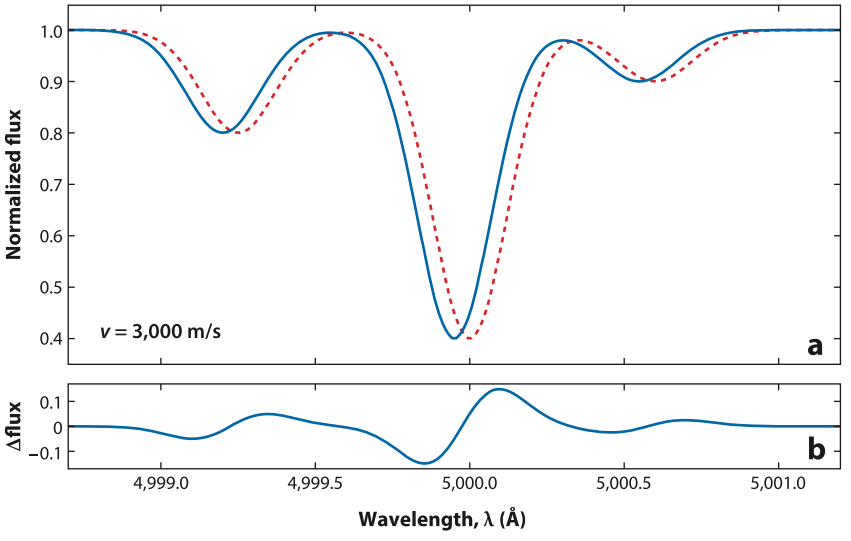

Figure 1

(*a*) Small portion of a spectrum in rest frame (*red*, *dotted*) and in observer frame (*blue*, *solid*). (*b*) Difference in flux between the Doppler-shifted and rest frame spectra. For illustration, we show a large Doppler shift typical of a binary star system. Abbreviation: v, relative velocity of the stars.

observer who acquires the spectrum of a given star multiple times can measure changes in the apparent wavelength of each spectral line as a function of time, t. This maps to measuring $\mathrm{RV}(t) \equiv \mathbf{k} \cdot \mathbf{v}(t)$ in Equation 1. For more details on the definition of RV, readers are directed to Lindegren & Dravins (2003) and Lovis & Fischer (2010).

Keplerian signal: RV signal as a function of time resulting from the orbit of a planet, as given in Equation 2 and shown in **Figure 2**

Measuring the RV from a spectrum has many steps, including the correction of instrumental effects. In this review, we start with the estimate of the true stellar spectrum, debiased from instrumental effects, and transformed to be equivalent to what an observer at the Solar System's center of mass would see. The earlier steps of analysis are briefly presented in the **Supplemental Appendix**.

Supplemental Material >

2.2. Forward Model of Planetary Effects

An orbiting planet causes a reflex motion of the star, thus creating RV variations. More precisely, the star and the planet periodically traverse elliptical orbits. The star's RV projected onto the line of sight, or RV, at time t depends on the planet's mass and orbital parameters and is given by the so-called Keplerian signal,

$$f(t; K, P, e, \omega, M_0) = K\,[\cos(\omega + \nu(t; e, P, M_0)) + e\cos\omega], \qquad 2.$$

where K, P, e, ω, and M_0 are the velocity amplitude, the orbital period, the orbital eccentricity, the argument of periastron, and the mean anomaly describing the motion of the star. The true anomaly, ν, is an angle that parameterizes the position of the star on its orbit (see Murray & Correia 2010).

The eccentricity of the ellipse is confined to [0, 1) for a bound system undergoing periodic motion. Orbits with $e \simeq 0$ are nearly circular and can be well approximated by the leading terms of a series expansion in the mean anomaly. Orbits with $e \geq 0.3$ become obviously elongated and the motion becomes noticeably uneven in time, typically leading to a sharp rise and slow fall (or vice versa), as shown in **Figure 2**. Very high eccentricities ($0.9 < e < 1 - R_\star/a$, where $R_\star$ is the stellar radius and a is the mean star-planet separation) are very rare but can create numerical difficulties. The signal amplitude depends on the mass of the star and the planet, as well as on the orbital

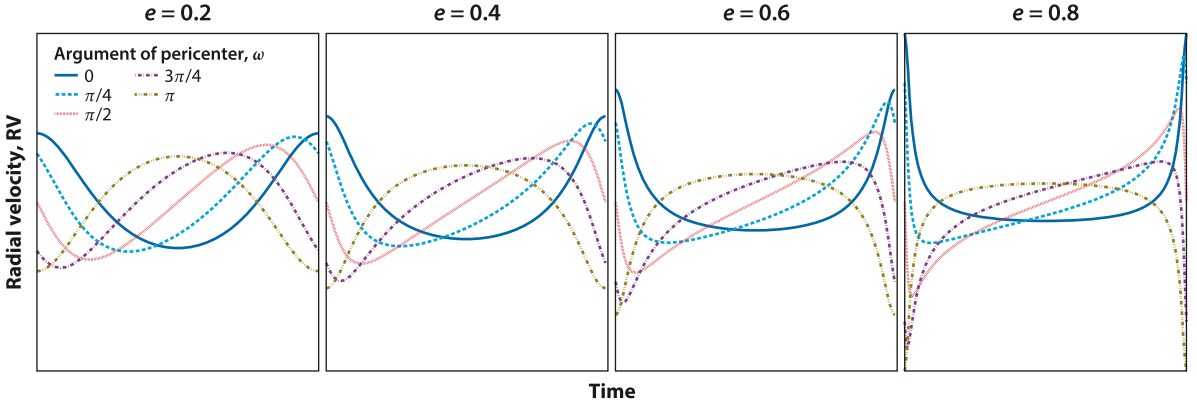

Figure 2

Radial velocity (RV) signature of a single planet. Each panel shows the RV signature over one orbital period for a planet with a given orbital eccentricity e. All curves are computed with the same RV semiamplitude K, and all panels have the same y-axis scale.

eccentricity, according to the formula

$$K = \left(\frac{2\pi G}{P}\right)^{\frac{1}{3}} \frac{m \sin i}{(m+M)^{\frac{2}{3}}} \frac{1}{\sqrt{1-e^2}}, \quad 3.$$

where M is the stellar mass, m is the planet mass, G the gravitational constant, and i is the angle between the plane in which the planet orbits and the plane perpendicular to the line of sight (Perryman 2011). Equation 3 shows that the RV semiamplitude is proportional to $m/P^{1/3}$: The more massive the star, and the smaller and farther from its star a planet is, the more difficult the planet is to detect.

Each observed star has a proper motion due to the motion of the star (and the Sun) around the galactic center. On the timescale of RV observations, ∼1 to 20 years, this usually appears as constant, in some cases as a linear trend that must be added to Equation 2. Furthermore, most Sun-like stars host multiple planets (He et al. 2021). In principle, the planet-planet interactions cause deviations from the Keplerian model, but for the vast majority of planetary systems, the motion of the star can be well approximated over the timescale of RV surveys as the linear superposition of the Keplerian orbit due to each planet. Hence, our nominal physical model for a time series of RV measurements of a star hosting n planets with parameters $(K_j, P_j, e_j, \omega_j, M_{0j})_{j=1,\ldots,n}$ is

$$\mathrm{RV}(t) = c_0 + c_1 t + \sum_{j=1}^{n} f(t; K_j, P_j, e_j, \omega_j, M_{0j}). \quad 4.$$

The RV amplitude of planet j, K_j, is proportional to $m \sin i$ (see Equation 3). In the absence of information on i, the planetary mass m cannot be determined unambiguously. Constraints on i can be obtained if the planet transits or, in the case of systems with multiple high-mass short-period planets, due to detecting planet-planet interactions (e.g., Laughlin & Chambers 2001, Correia et al. 2010, Nelson et al. 2016, Rosenthal et al. 2019).

When the data are acquired with different instruments, it is important to account for the fact that they might have different zero velocity references. Hence, if there are m different instruments, c_0 in Equation 4 should be replaced by $\sum_{j=1}^{m} c_0^j \chi_j(t)$, where $\chi_j(t) = 1$ if the measurement at t is taken by instrument j and 0 otherwise.

2.3. Basic Model of Observed Data

To detect the variations due to planets, astronomers measure the RV of a given star at several irregularly spaced epochs, usually from 20 to 1,000. Each star is only observable when the star is high above the horizon, typically for only a few hours per night, and weather can prevent stars from being observed. Furthermore, a star is typically observable for ∼4–8 months, as it must not be too close to the Sun. Finally, most observatories support multiple science programs, so observations on a given star might be interrupted for a period from hours to weeks.

Each RV measurement has a nominal uncertainty depending on the number of photons received, as well as the properties of the star, spectrograph, and data reduction technique (Bouchy et al. 2001). Astronomers typically assume independent Gaussian measurement noise at each time t_i, but in order to reduce the risk of underestimating uncertainties (e.g., unmodeled planets, instrumental noise, or stellar variability), an extra Gaussian noise term, often called jitter, is added to the model (e.g., Ford 2006), leading to

$$y(t) = \mathrm{RV}(t) + \epsilon(t), \quad \text{and} \quad 5.$$

$$\epsilon(t_i) \sim N(0, \sigma_{t_i}^2 + \sigma_j^2), \quad 6.$$

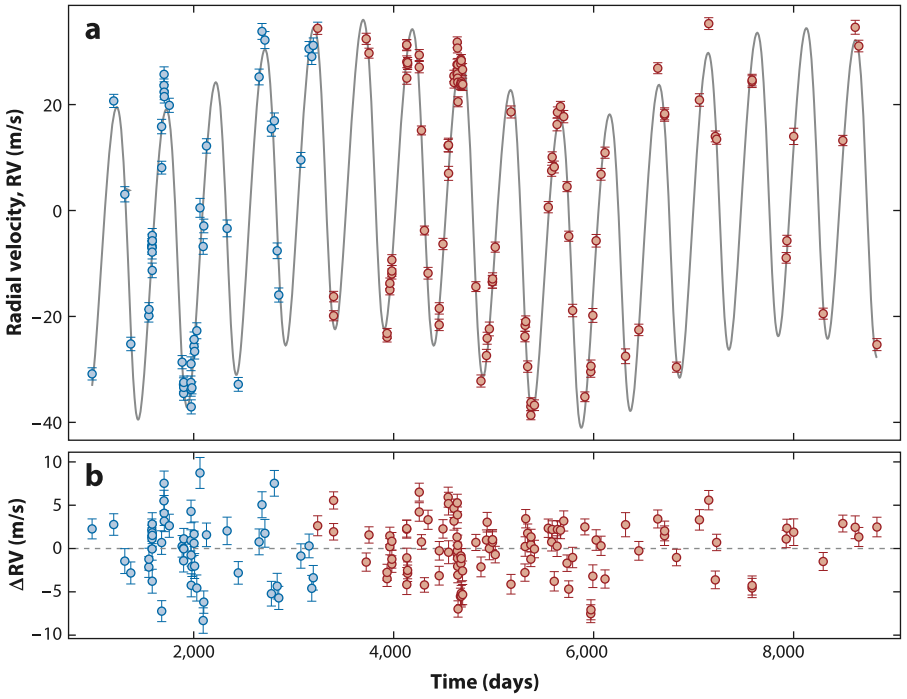

Figure 3

(*a*) Radial velocity (RV) measurements of the star HD 114783. The gray curve shows the maximum likelihood 2-planet model. (*b*) RV residuals relative to the 2-planet model above. Error bars reflect reported measurement uncertainties (σ_{t_i}) and do not reflect the 3 m/s of jitter (σ_J). Points are color-coded to indicate whether they are before or after an instrument upgrade that resulted in an unknown RV offset. Data are from Rosenthal et al. (2021).

where RV(t) is defined in Equation 4 and the value of the assumed noise for the measurement made at t_i, $\epsilon(t_i)$, follows a Gaussian distribution of variance $\sigma_{t_i}^2 + \sigma_J^2$. **Figure 3** shows an example of a two-planet fit on the RV data of HD114783.

2.4. Other Processes Affecting Radial Velocities

The model in Equations 4–6 has been extensively used to estimate the masses and orbital parameters of exoplanets (e.g., Wright & Howard 2009, Bonomo et al. 2017). For favorable stars, it is useful for analyzing RV signals with amplitudes greater than 3–10 m/s = 33–110 $RV_\oplus$, enough to characterize short-period planets with masses greater than Neptune. However, detecting less massive planets requires very precise observations and/or many observations. Accurately interpreting their RV signatures requires greater sophistication because multiple potential noise sources (listed below) become relevant and require more sophisticated models for ϵ_t.

2.4.1. Stellar effects. The principle of Doppler spectroscopy is to measure the velocity of a light source with respect to an observer, so if the hot gas in different parts of the stellar surface has different brightness and velocities, it has different impacts on the data. Several physical processes on the surface of stars have an RV signature.

Convection near the surface of the star creates outward and inward motion of the gas. This creates a pattern of evolving convection cells. This process, referred to as granulation, has

an effect on RV measurements that can be described as a stationary noise with a Lorentzian or super-Lorentzian power spectrum (Harvey 1985, Dumusque et al. 2011b, Cegla et al. 2019, Guo et al. 2022). Granulation effects on the RV have been simulated in detail (Meunier et al. 2015, Cegla et al. 2019, Dravins et al. 2021).

Local enhancement of the magnetic field at the surfaces of stars might result in regions darker or brighter than their surroundings, called spots and faculae, respectively (Schrijver 2002, Strassmeier 2009). This creates an imbalance in flux from the approaching and receding sides of the rotating star, and reduces upward convection, which has an additional net RV effect. The effect of magnetic activity on RV has been studied by Saar & Donahue (1997), Dumusque et al. (2011a), and Haywood et al. (2016) and simulated by Lagrange et al. (2010), Boisse et al. (2012), Dumusque et al. (2014), and Gilbertson et al. (2020). Active regions grow in size rapidly and disappear more gradually. The visible stellar surface is similar but not identical to itself after one stellar rotation, which results in quasiperiodic RV variations.

The rate of appearance, size and location of spots and faculae varies over a ∼10 year timescale. This affects the amplitude of short-term RV variations due to magnetic activity [up to 25 m/s (Lovis et al. 2011)] and changes the net RV effect over these timescales. Cegla (2019) and Meunier (2021) provide reviews of additional stellar processes that affect the RV (see also the **Supplemental Appendix**).

CCF Mask: function of the wavelength equal to unity except on wavelength ranges corresponding to spectral lines (see Pepe et al. 2002)

Ancillary indicator: time series of N scalars extracted from time series of N spectra, summarizing a shape variation of the spectrum

Supplemental Material >

2.4.2. Other effects. Residual instrumental systematics and absorption of certain wavelengths by Earth's atmosphere can result in spurious RV signatures of the order of a few $RV_\oplus$ (e.g., Artigau et al. 2014, Bertaux et al. 2014, Cunha et al. 2014, Smette et al. 2015). As these effects are not amenable to a concise statistical description, we refer the reader to Halverson et al. (2016) and Cretignier et al. (2021) as well as the **Supplemental Appendix**.

3. STATISTICAL FRAMEWORK

3.1. Model of Radial Velocity

The planets only affect the Doppler shift, which must be estimated and converted to RV measurements. We can express the RV extracted from the spectra, $\widehat{\mathrm{RV}(t)}$, as the sum of $\mathrm{RV}(t)$, the RV at time t due to a motion of the center of mass of the star (in particular due to planets); $\mathrm{RV}_{\mathrm{contam}}(t)$, the RV caused by stellar variability and instrument systematics; and $\epsilon(t)$, the inevitable photon noise. Then, we have

$$\widehat{\mathrm{RV}(t_i)} = \mathrm{RV}(t_i) + \mathrm{RV}_{\mathrm{contam}}(t_i) + \epsilon(t_i), \quad \text{and} \tag{7.}$$

$$\mathrm{RV}(t_i) = \mathrm{RV}_{\mathrm{planets}}(t_i) + \mathrm{RV}_{\mathrm{g}}(t_i), i = 1, \ldots, N, \tag{8.}$$

where $\mathrm{RV}_{\mathrm{planets}}(t_i)$ and $\mathrm{RV}_{\mathrm{g}}(t_i)$ are, respectively, induced by planets and by other gravitational effects, such as companion stars or proper motion in the galaxy. In Equation 4, $\mathrm{RV}_g(t)$ is defined as $\mathrm{RV}_g(t) = c_0 + c_1 t$.

In order to interpret RV measurements properly, we must adequately model $\mathrm{RV}_{\mathrm{contam}}(t)$. Unlike planetary RV signals, RV variations induced by the star and instrument do not have a constant frequency, phase, and amplitude (see Section 4.4), and stellar processes and systematics not only cause a Doppler shift in the spectrum but also cause the shape of the spectrum to vary with time (see **Figure 4**) and do not affect all lines identically (Wise et al. 2018). The shape variations of the spectra can be used to predict the associated $\mathrm{RV}_{\mathrm{contam}}$ signal, either in a machine learning approach with a training set of spectra, or in a statistical framework. In the following section, we show how these ideas can be expressed in a general statistical formalism.

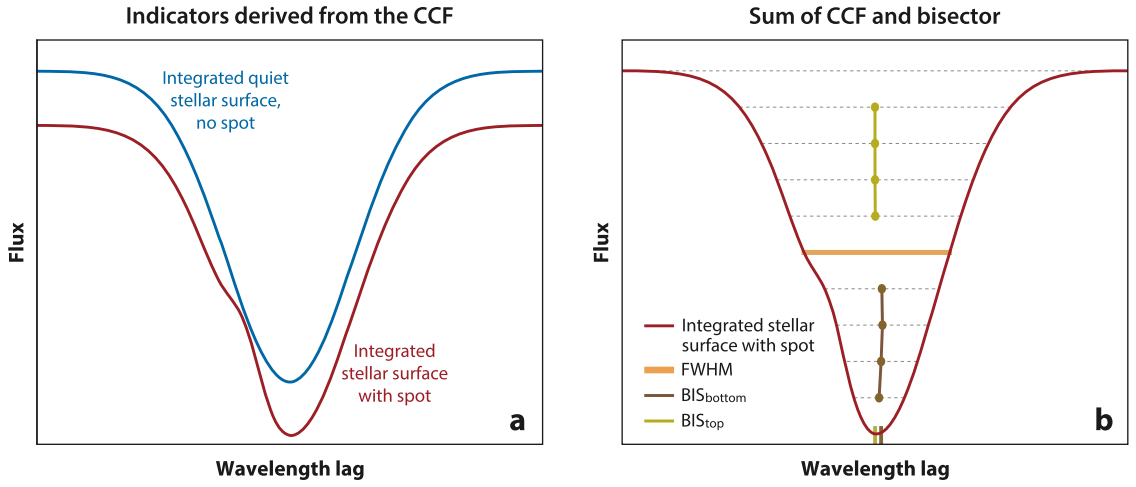

Figure 4

Schematic cross-correlation function (CCF). The scalar product of a mask and the spectrum as a function of the wavelength offset between the two, which can be visualized as an average spectral line. (*a*) The blue curve shows the CCF of the quiet stellar surface, and the red curve shows the CCF when a dark spot is present on the surface. (*b*) The CCF when the spot is present has a different full width at half maximum (FWHM) than the CCF of the quiet stellar surface and is asymmetric. The difference between the average position of the lower and upper parts of the CCF, known as bisector inverse slope (BIS) (Queloz et al. 2001), is nonzero. As the star rotates, the spot CCF will have a different Doppler shift: The bump in the red CCF moves from left to right, which also causes the shape of the spectra to change with time, in particular the BIS and FWHM. The spot in this figure covers 20% of the surface, whereas spots on the sun cover ∼0.01–0.1% of the surface.

3.2. Challenges of Radial Velocity Data Analysis

Our understanding of exoplanets can be thought of as a language where each planet is a word, and each planetary system is a sentence. As in information theory (Shannon 1948), our knowledge of this language comes through a noisy and biased communication channel. In the long term, we want to understand the full language, but first we must learn to interpret each sentence: What planets orbit a given star? The Bayesian formalism is well suited to describe such situations, and indeed, it offers a compact way to present the different problems of RV data analysis (see also Sandford et al. 2019).

In our analogy, the message received is the raw data. The most fundamental data product is the number of electrons counted on each CCD detector of the spectrograph.[1] The sentence to be decoded is represented by vectors $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$. The vector $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n)$ represents the planetary system, where n is the number of planets (which is also a variable) and $\boldsymbol{\theta}_j$ the orbital elements of planet indexed by j (see Equation 2). The vector $\boldsymbol{\eta}$ includes all other relevant nonplanetary parameters, such as an offset, a drift, the stellar rotation period, noise amplitude, and timescale.

Let us denote the lowest-level data by $\mathbf{D}$. Our goal is to obtain a meaningful joint posterior distribution for $\boldsymbol{\eta}$ and $\boldsymbol{\theta}$ knowing $\mathbf{D}$, but the latter is corrupted by hundreds of instrumental effects, and computing the posterior directly is currently unmanageable. To be usable, $\mathbf{D}$ needs to be reduced to an estimate of the RV time series $\widehat{\mathbf{RV}} = (\widehat{\mathrm{RV}}(t_i))_{i=1,\ldots,N}$ and ancillary indicators $\mathbf{I} = (\mathbf{I}_j(t_i))_{j=1..p,i=1..N}$, also time series of length N, summarizing the variations of the shape of the

[1] We could go even further and consider data from an entire survey of spectral time series of different stars as a population using a hierarchical model to better constrain instrumental effects and potentially stellar variability.

spectrum (see **Figure 4*b***):

$$p(\boldsymbol{\theta},\boldsymbol{\eta} \mid \mathbf{D}) \approx p(\boldsymbol{\theta},\boldsymbol{\eta} \mid \mathbf{I}, \widehat{\mathbf{RV}}) = \frac{p(\mathbf{I}, \widehat{\mathbf{RV}} \mid \boldsymbol{\theta},\boldsymbol{\eta}) p(\boldsymbol{\theta},\boldsymbol{\eta})}{p(\mathbf{I}, \widehat{\mathbf{RV}})} = \frac{p(\mathbf{I}, \mathbf{RV}_{\text{contam}} + \epsilon \mid \boldsymbol{\theta},\boldsymbol{\eta}) p(\boldsymbol{\theta},\boldsymbol{\eta})}{p(\mathbf{I}, \widehat{\mathbf{RV}})}. \quad 9.$$

Reduced data: vector $\mathbf{y}$ concatenating the RV time series and ancillary indicators, extracted from the raw data $\mathbf{D}$

Likelihood: probability distribution of the reduced data $\mathbf{y}$ knowing the model parameters $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$; denoted by $p(\mathbf{y} \mid \boldsymbol{\theta},\boldsymbol{\eta})$

The last equality comes from the fact that RV(t) depends deterministically on $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$. Below, we denote the reduced data as $\mathbf{y} = (\widehat{\mathbf{RV}}, \mathbf{I})$, potentially with $\mathbf{I} = \emptyset$. We refer to $p(\boldsymbol{\theta},\boldsymbol{\eta})$ as the prior distribution and $p(\mathbf{y} \mid \boldsymbol{\theta},\boldsymbol{\eta})$ as the likelihood. In some cases, to emphasize that the likelihood is restricted to models with exactly n planets, we denote the likelihood by $p(\mathbf{y} \mid \boldsymbol{\theta},\boldsymbol{\eta}, n)$.

Based on Equation 9, we can identify several choices to be made: (*a*) We must choose a reduction method transforming the observational raw data into summary statistics $\widehat{\mathbf{RV}}$ and $\mathbf{I}$ that estimate the RV and provide useful indicators of spectral variability. (*b*) We must choose a model, that is, a likelihood $p(\mathbf{y} \mid \boldsymbol{\theta},\boldsymbol{\eta})$ and prior distribution $p(\boldsymbol{\theta},\boldsymbol{\eta})$. The key difference with step *a* is that the effect of potential planets is now explicitly included in the likelihood definition. (*c*) We must choose a decision method for confidently claiming the detection of planetary signals and estimating their masses and orbital elements for a given model or a collection of models. In the Bayesian formalism above, the decision is based on the posterior distribution $p(\boldsymbol{\theta},\boldsymbol{\eta} \mid \mathbf{y})$ and must include a discussion of its sensitivity of the result to the choices made in *a* and *b*. The Bayesian formalism only serves here to compactly present points *a*–*c* and is not always used in practice. Points *a*, *b*, and *c* are, respectively, the topics of Sections 4, 5, and 6 and must be completed by (*d*) a set of practical and reliable numerical methods for performing the necessary calculations. Numerical aspects are highlighted when relevant.

4. DETECTING PLANETS AND ESTIMATING THEIR ORBITAL ELEMENTS

4.1. Likelihood

Suppose that we have at our disposal an RV time series, a few ancillary indicators, and a statistical model of them. We want to determine how many planets can be confidently detected and what their orbital elements are.

In Section 3, we presented the different steps of exoplanet detection, particularly a likelihood function describing the distribution of data $\mathbf{y}$ as a function of the model parameters $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$, describing, respectively, the planets and all other effects. The likelihood is commonly assumed to be Gaussian:

$$\mathcal{L} \equiv p(\mathbf{y} \mid \boldsymbol{\theta},\boldsymbol{\eta}) = \frac{\mathrm{e}^{-\frac{1}{2}[\mathbf{y}-\mathbf{g}(\mathbf{t};\boldsymbol{\theta},\boldsymbol{\eta})]^T \mathbf{V}(\boldsymbol{\eta})^{-1}[\mathbf{y}-\mathbf{g}(\mathbf{t};\boldsymbol{\theta},\boldsymbol{\eta})]}}{\sqrt{(2\pi)^N |\mathbf{V}(\boldsymbol{\eta})|}}, \quad 10.$$

where $|\mathbf{V}(\boldsymbol{\eta})|$ is the determinant of the covariance matrix $\mathbf{V}(\boldsymbol{\eta})$. We treat the case where $\mathbf{y}$ is a concatenation of RV and ancillary indicator time series in Section 5. Using the model in Equations 4–6, $\boldsymbol{\theta} = (K_j, P_j, e_j, \omega_j, M_{0j})_{j=1,..,n}$, $\mathbf{g}$ is the sum of Keplerians and an affine function (see Equation 4), and $\eta = \sigma_J, c_0, c_1$; $\mathbf{V}$ is diagonal with ith element $\sigma_{i_i}^2 + \sigma_J^2$. If $\mathrm{RV}_{\text{contam}}$ is nonzero, it might be modeled by a nondiagonal matrix $\mathbf{V}(\boldsymbol{\eta})$ and/or a linear model of stellar activity indicators in $\mathbf{g}$. We return to how to specify further $\mathbf{g}(\mathbf{t}; \boldsymbol{\theta})$ and $\mathbf{V}(\boldsymbol{\eta})$ in Section 5.

Assessing the statistical significance of a putative RV signal is remarkably challenging. We review three broad approaches: periodogram-based methods (Section 4.2), Bayesian model comparison (Section 4.3), and approaches that do not aim to specify an explicit model for nuisance signals but aim to obtain robust detections (Section 4.4).

4.2. Planet Detection via Periodograms

The decision on the planets detected and their orbital elements should make use of the full posterior of orbital elements or, in a frequentist setting, the full likelihood. However, computing the posterior distribution of elements given in Equation 9 or exploring exhaustively the parameter space to evaluate the likelihood is only possible with the latest numerical tools, and it is computationally intensive. Realistic data sets often have millions of local likelihood maxima. Historically, the exploration of likelihood modes was done with periodogram methods, which are still extensively used due to their speed, numerical stability, and ability to unveil the dominating frequencies in the data.

4.2.1. Definition. Given a time series, periodograms consist in comparing the log-likelihoods of two models: a base model H_0 and a model K_ω including H_0 plus a periodic component at frequency ω, for a grid of frequencies. In periodograms, the Keplerian model is often replaced with an approximation that yields a convex model once conditioned on a small number of parameters.

The first and simplest case is the Lomb–Scargle periodogram (Lomb 1976, Scargle 1982), where the base model H_0 is Gaussian white noise, and K_ω is the same white noise plus a sine function. Both H_0 and K_ω are described with likelihoods such as that in Equation 10, with

$$H_0 : \mathbf{g} = \mathbf{0}, \tag{11.}$$

$$K_\omega : \mathbf{g}(A, B, \omega) = A \cos \omega \mathbf{t} + B \sin \omega \mathbf{t}, \tag{12.}$$

and $\mathbf{V} = \mathrm{diag}((\sigma_i^2)_{i=1,\ldots,N})$ in both models. In K_ω, $\mathbf{g}$ is the exact RV signature of a planet on a circular orbit ($e = 0$) with orbital period $P = 2\pi/\omega$ and is a good approximation when $eK \leq \sigma_{\mathrm{RV,ideal}}$. For a given ω, maximizing the likelihood with respect to A and B is equivalent to minimizing the sum of squares, and it is a linear problem. The Lomb–Scargle periodogram is, then, the difference of the log-likelihoods of models H_0 and K_ω as a function of ω.

The principle of the periodogram can be extended with more complex definitions of H_0 and K_ω and/or by interpreting the periodogram in a Bayesian context; the null hypothesis can be complexified (e.g., Baluev 2008), the periodic signals might be chosen as nonsinusoidal (e.g., Baluev 2013b, 2015), and the assumption that the noise is uncorrelated can be dropped (e.g., Delisle et al. 2020a). **Figure 5** shows an example of the application of different types of periodograms. A comprehensive list of existing periodograms is given in the **Supplemental Appendix**.

Supplemental Material >

4.2.2. Periodogram-based model comparison. The usual way to determine whether a periodogram supports the detection of a periodic signal is to compute $p(\max_\omega \mathcal{P}(\omega) \mid H_0)$, the probability distribution of the maximum value of the periodogram, $\max_\omega \mathcal{P}(\omega)$, under a null hypothesis, H_0. We compute the maximum of the periodogram for the data to be analyzed, $\mathcal{P}_\mathrm{d}$, on a grid of frequencies, $\Omega = (\omega_i)_{i=1..N}$, and define a false alarm probability (FAP) as $p(\max_{\omega\in\Omega} \mathcal{P}(\omega) \geqslant \mathcal{P}_\mathrm{d} \mid H_0)$.

The FAP is most robustly estimated by generating signals under the null hypothesis, and the FAP estimated is the fraction of simulations with a maximum peak greater than $\mathcal{P}_\mathrm{d}$. This is computationally expensive, as it requires $\sim n$ simulations for a precision of $100/n\%$ on the FAP. Since astronomers often aim for a FAP of $<0.1\%$ to 0.01%, $n > 10^4$ is needed.

A semianalytical approach allows one to simulate a reduced number of signals and to fit a generalized extreme value distribution (Süveges 2014) to the empirical distribution. The analytical approach approximates the FAP thanks to the theory of extreme values of stochastic processes using the Rice formula (Baluev 2008, 2009, 2013a), even for complex periodic shapes

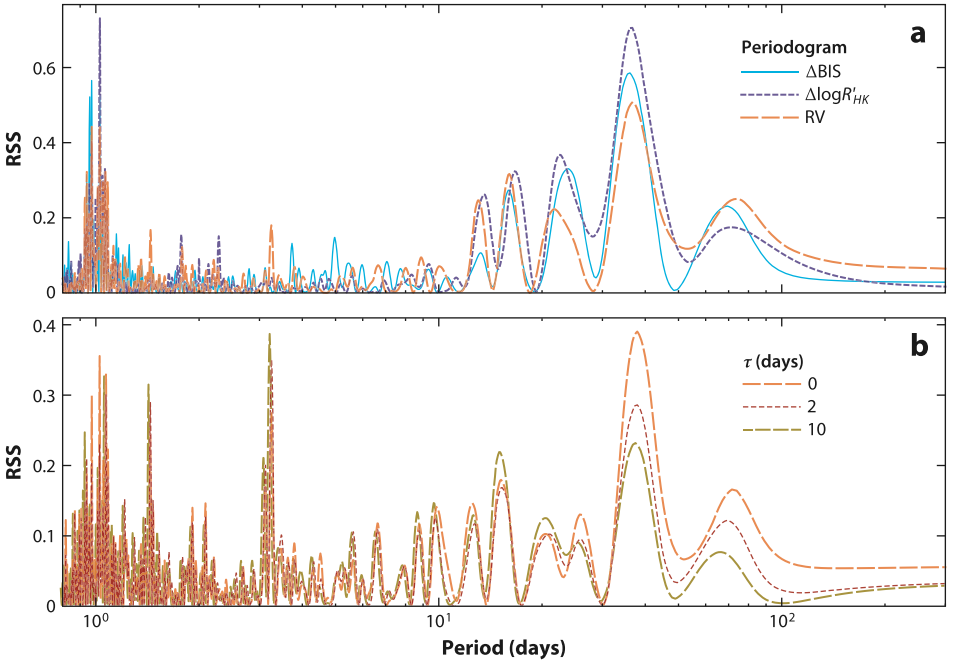

Figure 5

(*a*) Generalized Lomb–Scargle periodograms of the time series of two ancillary indicators: bisector inverse slope, ΔBIS (Queloz et al. 2001), and ancillary indicator derived from each spectrum, $\Delta \log R'_{\rm HK}$ (Noyes 1984), and radial velocity (RV) time series of the ESPRESSO (Echelle SPectrograph for Rocky Exoplanets and Stable Spectroscopic Observations) data of the star TOI 178 (Leleu et al. 2021). The time series of BIS, $\Delta \log R'_{\rm HK}$, and RV are shown in **Figure 7**. Periodograms are expressed in reduction sum of squares (RSS), normalized—that is, the difference between residual sum of squares after fitting the base model H_0 and the model including a periodic component K_ω, divided by that of the H_0 residuals (see Equations 11 and 12). The periodogram presents peaks around a period of 1 day, due to the repetition of observations every 24 h ± 2–3 h (Dawson & Fabrycky 2010). All three time series present power at ∼40 days, due to stellar rotation. (*b*) In orange, periodograms of the RV time series when the $\Delta \log R'_{\rm HK}$ and ΔBIS are included as linear predictors in the H_0 and K_ω models. This damps the amplitude of the rotational signal, and a peak at 3.2 days emerges. The red and light brown curves correspond to adding a correlated noise Gaussian model in the covariance with a Gaussian kernel of amplitude 1 m/s and timescale τ of 2 or 10 days. Adding correlated noise models always damps power at low frequencies (for details, see Delisle et al. 2020a).

(Baluev 2013b, 2015) or correlated noise (Delisle et al. 2020a). These analytical approximations are very accurate in practice if at least a few tens of data points are available (Süveges et al. 2015).

4.2.3. Pros and cons. Periodograms are fast and numerically stable, and the advent of accurate analytical estimates of the FAP, even in the case where signals are searched for simultaneously (Baluev 2013a) or with Gaussian correlated noise models (Delisle et al. 2020a), further simplifies their use. They provide a useful visual diagnostic of the frequency content of the signal. However, they have drawbacks.

First, most periodograms scan for the orbital period of one planet at a time. For multiple-planet systems, periodograms can be applied iteratively, first to identify the most conspicuous planet from the data, then to look for a $(j + 1)$th planet in the residuals of the best-fit model using j planets. However, periodograms are sensitive to aliasing or spectral leakage (Dawson & Fabrycky 2010). Aliases of different signals can add destructively or coherently, so even if the highest peak of a

periodogram is statistically significant, its period may not match that of any physical signal. This can be avoided by searching for multiple signals simultaneously (Ford et al. 2011, Baluev 2013a). A brute force search for two signals at once is expensive, and allowing for searching more rapidly becomes prohibitive. Other approaches based on sparse recovery allow one to search for several signals at low computational cost (Hara et al. 2017).

Stellar activity causes low-frequency signals that can be mistaken for planets. It is better not only to search for several planets simultaneously but also to fit the parameters of stellar activity and planetary signals jointly. As discussed in Section 4.1, stellar variability is often modeled as correlated RV noise using a nondiagonal covariance matrix $\mathbf{V}$ that requires additional parameters that must be inferred from the data. This means that the log-likelihood is nonconvex for a given orbital period. Marginalizing (or even optimizing) over the kernel parameters dramatically increases the computational cost of periodograms, although it can be done (Delisle et al. 2018).

4.3. Bayesian Approach to Planet Detection

A more principled approach to comparing models with different numbers of planets is to directly use the Bayesian formalism of Equation 9. The first method consists in computing the marginal likelihood, or the Bayesian evidence, of the n-planet model. We let Θ_n denote the parameter space of all possible combinations of n planets, and

$$p(\mathbf{y} \mid n) = \iint_{\Theta_n} p(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\eta}, n) p(\boldsymbol{\theta}, \boldsymbol{\eta} \mid n) \, \mathrm{d}\boldsymbol{\theta} \, \mathrm{d}\boldsymbol{\eta}, \qquad 13.$$

where $p(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\eta}, n)$ is given by Equation 10. If one can compute the Bayesian evidence for models with $n + 1$ and n planets, then their ratio gives the Bayes factor (Kass & Raftery 1995). If the noise model is correct, then when one considers more planets than are justified for the given data set, the evidence decreases as more planets are added. For instance, when the priors on the different parameters are considered independent, the prior term decreases geometrically with the number of planets but the model does not result in a significantly higher likelihood. Bayesian model comparison for exoplanet detection was suggested by Gregory (2005a, 2007), Ford & Gregory (2007), Tuomi & Kotiranta (2009), Ford et al. (2011), and Díaz et al. (2016), and it has become one of the primary methods for establishing the statistical significance of detections.

Astronomers compute the Bayes factors for increasing n, starting at $n = 0$, until they are below a certain threshold. The literature contains various heuristics for the interpretation of Bayes factors (e.g., Jeffreys 1961). Unfortunately, following guidelines blindly can be very dangerous. For important scientific discoveries, astrophysicists routinely demand that a frequentist test rejects a null hypothesis test with a p-value of $\sim 10^{-3}$ or even $\sim 10^{-7}$ before publishing a result. In such situations, a Bayesian would demand a posterior odds ratio (i.e., prior odds ratio times the Bayes factor) exceeding 10^3 or even 10^7 before publishing a discovery.

Bayes factors compare different numbers of planets. It is insufficient to claim a confident detection of a planet, which requires knowledge of the orbital elements to a certain accuracy: There is little value in a planet detection without estimates of its size and period. Well-defined orbital elements correspond to a sharp posterior mode, whose existence can be checked with the posterior samples of orbital elements. This can be done with periodograms, or directly with the posterior distribution of orbital elements if it can be estimated reliably. If so, the detection criterion can be constructed to convey information about the planets' location. For instance, a detection claim can be defined as "there is a planet with orbital elements in Θ," where Θ is a region of the parameter space (Brewer & Donovan 2015, Handley et al. 2015, Hara et al. 2022c). In that case, the criterion minimizing the number of missed detections for a certain tolerance to false ones as well as

maximizing a certain utility function is the false inclusion probability (FIP) (Hara et al. 2022a), i.e., the probability of not having a planet with orbital elements in Θ.

A Bayes factor strongly favoring an $(n + 1)$-planet model over an n planet model does not necessarily imply that there must be an $(n + 1)$th planet, if the physical or statistical models used are inaccurate. It is important to check the dependence of the results on the adopted model.

A significant barrier to more routine adoption of Bayesian model comparison is the difficulty of computing the Bayesian evidence accurately. Nelson et al. (2020) compared different methods to compute the Bayesian evidence and found that the agreement between them goes from a factor of $\sim$1 for 0 planet models to $\sim 10^2$ for 2 planet models. Additionally, the observed dispersion of Bayes factor estimates from multiple runs of the same method was often significantly greater than the reported uncertainties for some methods. As a consequence, they recommend evaluating the numerical uncertainty based on several independent runs. When a fast estimate is needed, a Laplace approximation is advised over heuristics such as the Akaike information criterion or Bayesian information criterion. Numerical details are discussed further in the **Supplemental Appendix**.

Supplemental Material >

4.4. Qualitative Approach to Planet Detection

The methods presented in Sections 4.2 and 4.3 rely on a complete model of the signal. All the alternative hypotheses are made explicit and are compared with one another. This gives meaningful statistical significance indicators and measures of uncertainty. However, if none of the noise models capture stochastic variation in the data, the inferences are unreliable. A second approach consists, in the formalism of Equation 8, of extracting meaningful information about $\mathrm{RV}_{\mathrm{planets}}(t)$ without explicitly specifying a model for $\mathrm{RV}_{\mathrm{contam}}(t)$. Such approaches, which are the object of this section, also might serve another purpose: diagnosing unanticipated effects.

On the timescale of RV observations, unless there are strong gravitational interactions between the planets, planetary signals are purely periodic, unlike stellar and instrumental signals. To diagnose whether a signal is truly periodic, Schuster (1898) and Mortier & Collier Cameron (2017) compute classical periodograms by adding one point at a time and checking that the amplitude of the peak corresponding to the candidate planet increases steadily. Alternatively, Gregory (2016) and Hara et al. (2022b) use the Bayesian framework described in Section 4.3 and add an apodization factor to the Keplerians: The model of Equation 2 is multiplied by a Gaussian term $e^{-(t-t_0)^2/(2\tau^2)}$, where t_0 and τ are free parameters. If the signal is consistent, the probability that τ exceeds the total observation time span should be high.

Another line of work consists of searching for periodic signals in the data, without specifying a parametric form. Zucker (2015) suggests using a Hoeffding test. Zucker (2018) applies the formalism of distance correlation to evaluate the statistical dependence of a cyclic variable with period P and the data, and this is further explored by Binnenfeld et al. (2021), who aim to find RV variations that are statistically independent from the spectral shape variations. This work uses spectral information through all pairwise distances between two spectra measurements and shows promising results.

4.5. Parameter Estimation and Uncertainty Quantification

Once the dominant local modes of the likelihood or posterior have been identified by a periodogram analysis or with an random sampler with good convergence properties, more computationally intensive methods can be employed for finer parameter estimation in this region. The posterior distribution of the elements can be evaluated with Markov chain Monte Carlo (MCMC) algorithms (Ford 2005, Gregory 2005b). Brute-force random walk MCMC requires carefully chosen proposal distributions for systems with up to four planets (Ford 2006). Modern studies typically use ensemble samplers (Foreman-Mackey et al. 2013, Nelson et al. 2014), adaptive

Metropolis sampling (Delisle et al. 2018), Hamiltonian and/or geometric MCMC samplers (e.g., Papamarkou et al. 2021), or premarginalization with a Laplace approximation of the evidence over the linear parameters (Price-Whelan et al. 2017). In each of these methods, the orbital periods are initialized at multiple values very near the dominant signals found by the periodogram analysis. The results are generally reliable if the likelihood is dominated by a single mode and the initial estimate of the period falls into that mode. Some nested samplers have shown good performances in blindly locating different likelihood maxima (Brewer & Donovan 2015, Faria et al. 2018).

5. MODEL: SPECIFYING THE PRIORS AND LIKELIHOODS

Different choices of priors and likelihoods have a very strong impact on the detection of planets. Priors have a strong influence both on the detection and orbital elements estimation of low amplitude signals, where the likelihood is less constraining (Hara et al. 2022c). To avoid priors that are unrealistically diffuse, a possibility is to use well tested reference priors and to exclude orbital configurations that are unstable (Tamayo et al. 2020, Stalport et al. 2022). A discussion of the influence of priors and a list of commonly used ones is provided in the **Supplemental Appendix**, along with a more detailed presentation of the concepts used in the definition of the likelihood, which is outlined in the current section.

We have seen in Section 2.4 that RV measurements are contaminated by several stellar and instrumental processes (RV_{contam} in Equation 8). The first family of methods to deal with these signals uses the ancillary indicators to predict RV variations based on the indicators. In the formalism of Section 3.2, this consists of building a likelihood $p(\mathbf{RV} \mid \mathbf{I}, \mathbf{D})$. Historically, the first method was to represent RV variations induced by the instrument and star as a linear combination of the ancillary indicators (Queloz et al. 2001, 2009; Dumusque et al. 2017). The efficiency of this method is also limited due to the fact that there can be phase shifts between activity-induced changes in RVs and ancillary indicators, reducing the correlation (Santerne et al. 2015, Lanza et al. 2018, Collier Cameron et al. 2019). To mitigate this issue, Simola et al. (2022) use a so-called change point detection algorithm to diagnose changing correlation patterns between the ancillary indicators and the RV. However, carefully selected indicators can be used as linear indicators (see Section 6.2).

Another way to represent the signals corrupting RVs is to represent them as correlated Gaussian noise. One can still follow the formalism of Section 4.1, but now specifying the matrix $\mathbf{V}$ with a kernel, which gives the correlation between the value of the stellar RV signal at t and $t + \Delta t$. A list of common kernels is provided in the **Supplemental Appendix**. The correlation of RVs due to stellar variability can be expressed as a sum of the correlations due to each of the processes described in Section 2.4 (see **Figure 6**).

Aigrain et al. (2012) showed that when nearly contemporaneous photometric and RV observations of a star are available, the effect of one stellar spot on RV can be predicted through the photometric flux and its derivative. For most targets, continuous photometric measurements are unavailable. One can estimate the derivative of the flux at any time using Gaussian processes (GPs). These are stochastic processes $G(\mathbf{t})$, a function of a variable $\mathbf{t}$ such that for any n values of $\mathbf{t}, \mathbf{t}_1, \ldots, \mathbf{t}_n$, $(G(\mathbf{t}_1), \ldots, G(\mathbf{t}_n))$ follows a multivariate Gaussian distribution. In general, $\mathbf{t}$ can be a vector, but for our purposes, we define it as the time t. This implies that the GP is defined by two quantities: its mean, $m(t)$, and kernel, $k(t, t')$, equal to the covariance of $G(t)$ and $G(t')$ (Rasmussen & Williams 2005).

Aigrain et al. (2012) use photometric data to predict the RV variation. We can go a step further and analyze RV and ancillary indicators simultaneously by building a likelihood $p(\mathbf{RV}, \mathbf{I} \mid \mathbf{D})$, typically in the GP framework. RVs and ancillary indicators are expressed as linear combinations of a latent GP and its derivatives (Rajpaul et al. 2015, Gilbertson et al. 2020, Gordon et al. 2020,

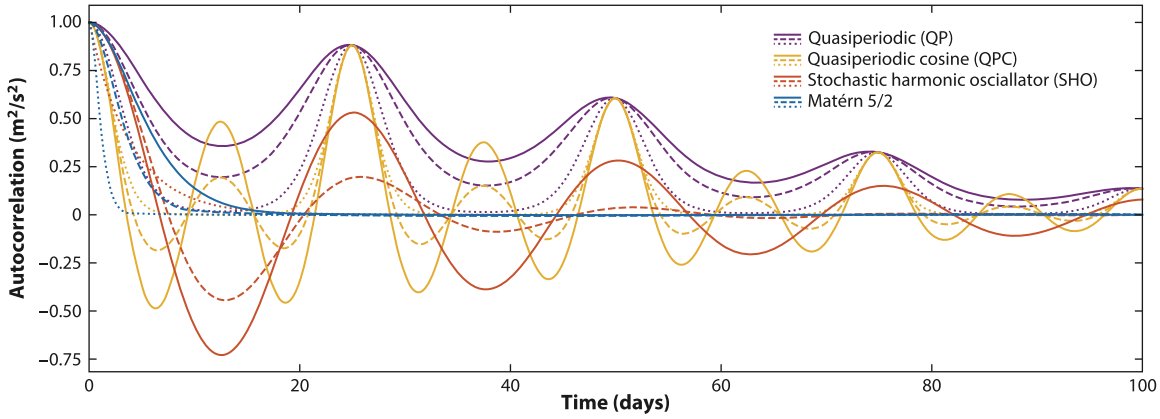

Figure 6

Kernel functions: functions of the time lag Δt between two observations. The covariance matrix $\mathbf{V}$ in Equation 10 is such that its element i, j is a sum of kernel functions evaluated in $\Delta t = t_i - t_j$, where t_i and t_j are the epochs of the ith and jth observations, respectively. Each color corresponds to a different family of kernels: quasiperiodic (QP) (Aigrain et al. 2012, Haywood et al. 2014), quasiperiodic cosine (QPC) (Perger et al. 2021), stochastic harmonic oscillator (SHO) (Foreman-Mackey et al. 2017), and Matérn 5/2 (Foreman-Mackey et al. 2017, Gilbertson et al. 2020). Each family has several parameters, including an amplitude, a timescale, and, for all but Matérn 5/2, a period. They also have shape parameters, and plain, dashed, and dotted lines correspond to different values of these shape parameters. The periods are taken as 25 days, close to the Sun's rotational period, and all kernels are normalized so that they are equal to 1 in $\Delta t = 0$. These different kernels model different phenomena (see Section 2.4.1). The QP kernel is well suited to represent stellar rotational effects due to convective blueshift inhibition with various stellar inclinations. QPCs model the fact that spots appear preferentially on stellar longitudes with a 180° phase shift in Sun-like stars (Borgniet et al. 2015, Spergel et al. 2015). SHO kernels take negative values; this type of anticorrelation arises due to the breaking of imbalance of flux of the approaching and receding limbs, and certain parameter values correspond to a super-Lorentzian power spectral density used to model granulation (Foreman-Mackey et al. 2017, Gilbertson et al. 2020). Matérn 5/2 kernels are efficient for modeling RV induced by spots and faculae with a lifetime lower than the rotation period (Gilbertson et al. 2020). Explicit kernels' expressions and values of the parameters are given in the **Supplemental Appendix**.

Barragán et al. 2022, Delisle et al. 2022). Since $G(t + \Delta t) \approx G(t) + \dot{G}(t)\Delta t$, using the derivative can account for small phase shifts between RV and indicators. The data $\mathbf{y}$ in Equation 10 are then the concatenation of the RV vector and ancillary indicators. For example, this framework can be used to model simultaneously the RVs derived in different spectral bands (Cale et al. 2021). Camacho et al. (2022) model the RV and ancillary indicator time series in a GP regression network. It allows, in particular, consideration of noise with heavy tails and offers ways to account for nonstationarity. **Figure 7** shows an example of RV and ancillary indicators modeled simultaneously with GPs.

Supplemental Material >

Evaluating the likelihood in Equation 10 for certain values of the parameters requires the log determinant of $\mathbf{V}(\boldsymbol{\eta})$ and computing $[\mathbf{y} - \mathbf{g}(\mathbf{t}; \boldsymbol{\theta})]^T \mathbf{V}^{-1}(\boldsymbol{\eta})[\mathbf{y} - \mathbf{g}(\mathbf{t}; \boldsymbol{\theta})]$. With standard algorithms the computational cost scales as $O(N^3)$ with the size N of the data set $\mathbf{y}$, and it becomes prohibitive for $N \gtrsim 10^3$. This prevents using many indicators on stars with hundreds of points, or analyzing the tens of thousands of RV measurements on the Sun. Fortunately, for certain choices of the kernel, the covariance matrix has a semiseparable expression, and then the computational cost scales as $O(N)$ (the `celerite` framework of Foreman-Mackey et al. 2017). Similarly, `TemporalGPs.jl` (**https://github.com/JuliaGaussianProcesses/TemporalGPs.jl**) allows for $O(N)$ computation of a broad class of 1D kernel functions in Julia. The `celerite` framework is generalized by Delisle et al. (2020b) to the `S+LEAF` framework, which models covariances as a sum of semiseparable and LEAF matrices (**https://gitlab.unige.ch/Jean-Baptiste.Delisle/spleaf**). Quasiseparable kernels provide even greater flexibility for modeling multivariable time series

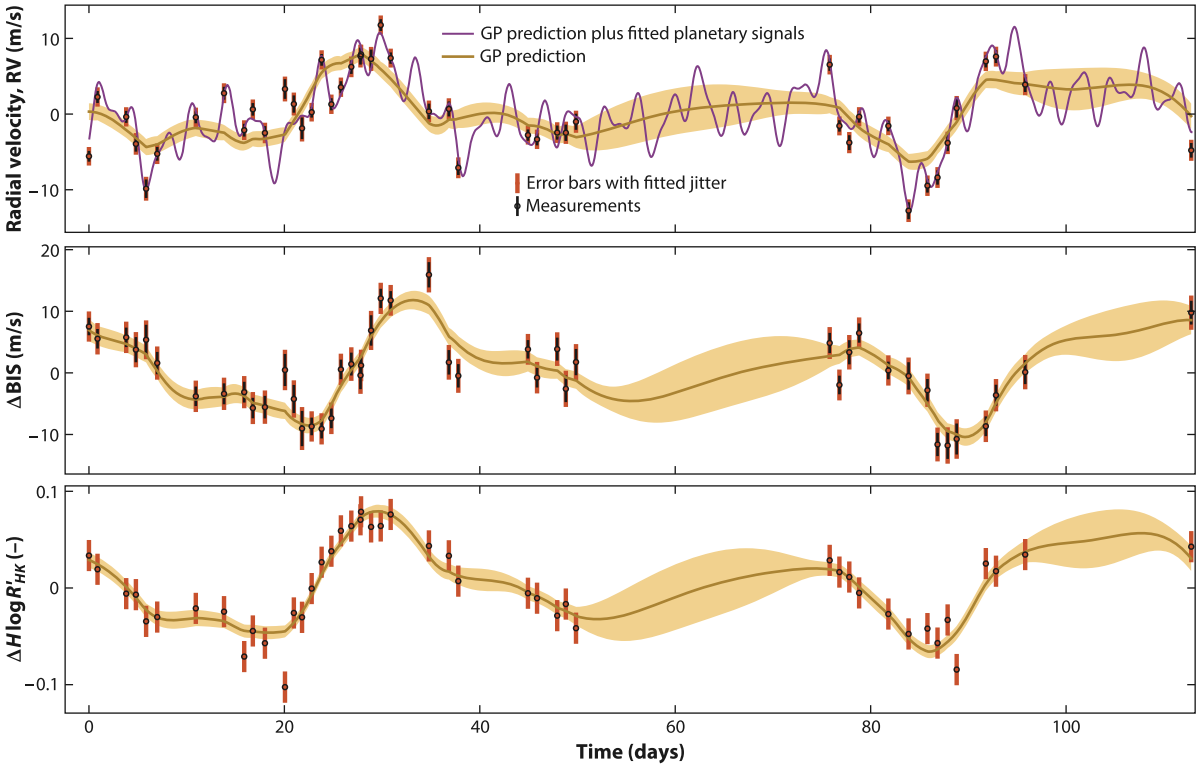

Figure 7

Gaussian process (GP) regression applied to the time series of radial velocity (RV), and ancillary indicators ΔBIS (Queloz et al. 2001) and $\Delta \log R'_{\rm HK}$ (Noyes 1984) index of 46 measurements obtained on the star TOI 178 with the ESPRESSO (Echelle SPectrograph for Rocky Exoplanets and Stable Spectroscopic Observations) instrument (Leleu et al. 2021). Each of the three time series $y_i(t)$, $i = 1, 2, 3$, is represented by a model $y_i(t) = \alpha_i G(t) + \beta_i \dot{G}(t) + \epsilon_m(t) + \epsilon_J(t)$. The GP G has a stochastic harmonic oscillator kernel (Foreman-Mackey et al. 2017) with three parameters, fitted with the `S+LEAF` package (Delisle et al. 2022). The processes $\epsilon_i^m(t)$ and $\epsilon_i^J(t)$ represent the measurement error and jitter. We represent with a brown line the mean of the posterior predictive distribution for each time series, and the shaded yellow areas represent ±1 standard deviation of $\alpha_i G(t) + \beta_i \dot{G}(t)$. The purple curve represents the variations predicted by the best-fit Keplerian model for the planet, whose presence is confirmed by transits (Leleu et al. 2021). Other abbreviations: BIS, bisector inverse slope; $\Delta \log R'_{\rm HK}$, ancillary indicator derived from each spectrum.

(`tinygp`) (**https://tinygp.readthedocs.io/en/stable/index.html**). Delisle et al. (2022) show that if RV and ancillary indicators are linear combinations of a GP $G(t)$ and its derivatives, and the covariance of $G(t)$ has an `S+LEAF` form, the computation of the likelihood of the augmented data is still $O(N)$ (**https://gitlab.unige.ch/Jean-Baptiste.Delisle/spleaf**).

The merits of the different noise models can be evaluated through Bayesian model comparison (see Section 4.3) (e.g., Ahrer et al. 2021, Suárez Mascareño et al. 2021, Faria et al. 2022); alternatively, Bayesian model averaging can be used (Hara et al. 2022a).

6. ANALYSIS METHODS: DEEPER LEVELS

So far, we have seen different methods to detect planets and estimate their orbital elements given priors and likelihood in Section 4. We have seen how the model can be built based on a certain

reduction of the data in Section 5. We now present techniques to extract physical information from the spectra.

Our goal is to identify conceptual similarities between different methods, rather than to assess their performances. While the community has begun to compare methods (Zhao et al. 2022), proper evaluations are challenging since one does not know the true RV for real stars due to potential undetected planets (except for our Sun; see Collier Cameron et al. 2019, Dumusque et al. 2021, Lin et al. 2022).

6.1. Estimating Radial Velocity

For a fixed stellar spectrum, the RV is well defined and can be extracted from the spectrum using multiple techniques described below. However, at the sub-m/s (or a few $RV_\oplus$) level, the shape of the spectrum changes, so there is no model-independent definition of RV, and each method to estimate it relies on more or less explicit assumptions. Depending on how this shape change is modeled, the estimate of the Doppler shift changes.

Three main families of methods for extracting the velocity in each spectrum separately are in use. As discussed in Section 3, the historic method consists of cross-correlating the spectrum with an idealized mask (Baranne et al. 1979, Pepe et al. 2002). Denoting by $f(\lambda)$ the flux as a function of the wavelength Λ, the cross-correlation function (CCF) is

$$\mathrm{CCF}(\Delta\lambda) = f * \mathrm{Mask}(\Delta\lambda), \qquad 14.$$

where $*$ is the convolution operator. This method can be viewed as computing an average line shape. Each line contributes to improving the signal-to-noise ratio (SNR) but also contributes a bias since the true line wavelengths are unknown. Additionally, the CCF loses information about differences in shapes of each line. Often, many lines are excluded from the CCF mask to reduce contamination of lines likely to contribute significant bias. Alternatively, Lienhard et al. (2022) proposes a least-square deconvolution technique to estimate a common spectral line profile.

A second approach, template matching, consists of measuring the RV based on a template or model spectrum and a Taylor expansion for the spectrum as a function of velocity (Connes 1985, Bouchy et al. 2001, Anglada-Escudé & Tuomi 2012, Astudillo-Defru et al. 2015, Jones et al. 2020). Related work has been done by Cretignier et al. (2022). This approach is based on the approximation

$$f(\lambda) \simeq f_0(\lambda) + \frac{RV}{c}\frac{\mathrm{d}\, f_0}{\mathrm{d}\log\lambda}(\lambda), \qquad 15.$$

where $f_0(\lambda)$ is a reference spectrum and c is the speed of light. Both the CCF and template matching approaches aim to reduce the impact of stellar variability by careful selection of lines/wavelengths for inclusion but lack a mechanism to recognize stellar variability. Some authors measure the RV of each line (or $\sim$2 Å chunk of the spectrum) separately (Dumusque 2018) and compute a weighted average RV, or they do independent RV estimates in several bandwidths (Feng et al. 2017, Zechmeister et al. 2018). The rationale is to disentangle signals that affect all wavelengths identically (planet candidates) from other signals.

A third approach builds a forward model of the spectrum,

$$f(\lambda^o) = \left[f_0(\lambda^{\mathrm{bc}}) + \frac{RV}{c}\frac{df_o}{d\log\lambda}(\lambda^{\mathrm{bc}}) \right] T(\lambda^o) * IP(\lambda^o), \qquad 16.$$

where $f(\lambda^0)$ is the spectrum in the observer frame; $\lambda^{\mathrm{bc}}(t) = \lambda^o\{1 + v_{\mathrm{bc}}(t)/c\}/\sqrt{1 - v_{\mathrm{bc}}^2(t)/c^2}$ is the wavelength in a frame that accounts for $v_{\mathrm{bc}}(t)$, the known motion of the observatory relative to the Solar System barycenter (the barycentric correction); $T(\cdot)$ is an atmospheric transmission profile;

and $IP(\cdot)$ is the instrument response. Variations on forward modeling have focused on telluric effects (Butler et al. 1996, Hirano et al. 2020) or on stellar variability (Bedell et al. 2019, Gilbertson et al. 2020, Jones et al. 2020). Ongoing research is developing computational tractable approaches for including both simultaneously (e.g., the `StellarSpectraObservationFitting` Julia package; **https://github.com/christiangil/StellarSpectraObservationFitting.jl**).

As methods become more sophisticated, the estimation of the velocity and fitting of the spectral shape tend to be done simultaneously on the time series of spectra. In principle, this mitigates the chance that the Doppler shift is contaminated by shape changes and uses temporal structure to further constrain activity.

6.2. Estimating Nuisance Radial Velocity Signals

To put Equation 9 into practice, one must specify a statistical model for $p(\mathbf{I}, \mathbf{RV}_{\rm contam} \mid \boldsymbol{\theta}, \boldsymbol{\eta})$. First, one chooses $\mathbf{I}$, the method of dimensional reduction, and then a form for the likelihood that describes $\mathbf{I}$ and $\mathbf{RV}_{\rm contam}$. Several studies have explored a simple linear model where the nuisance signal is a linear combination of indicators, $\mathbf{RV}_{\rm contam} = \sum_j \alpha_j \mathbf{I}_j$. For rotational-linked variability, it can be useful to generalize to models that predict $\mathbf{RV}_{\rm contam}$ at time t using spectra taken at times near t (Collier Cameron et al. 2019, Zhao & Ford 2022).

Dimension reduction has been applied to real and synthetic spectra. Davis et al. (2017) applied principal component analysis (PCA) to spectra generated with SOAP 2.0. Since real stellar spectra are more complicated, their analysis provides a lower limit on the number of PCA components needed to accurately reconstruct solar spectra (ranging from 1 to 4 depending on the spectral resolution and SNR). Analysis of solar data suggests that only 6–13 basis vectors are necessary to model solar variability at the resolution and SNR of HARPS-N observations, and at least four of those are clearly linked to effects due to the instrument or unique to Sun-as-a-star observations (Collier Cameron et al. 2021). Together, these suggest that reducing spectra to an RV and two to six indicators is a fruitful direction for future research.

Dimension reduction and estimation of contaminating RV can be done simultaneously. Existing methods extract ancillary indicators either in each spectrum separately, from a time series of CCFs (or other summaries of the spectra), or from the time series of spectra themselves. In the following subsections we describe the associated methods.

6.2.1. Extraction spectrum by spectrum. Some methods extract activity indicators from each spectrum separately. For example, one can measure line shapes (or deviations from their time-average). Shape line indicators were first extracted from the CCF, in particular its asymmetry (Queloz et al. 2001) and width (Queloz et al. 2009). Other indicators involve fitting Gaussian distributions on each side of the CCF (Figueira et al. 2013). Holzer et al. (2021) fit a superposition of Gauss–Hermite functions to each line to build shape indicators.

Another approach is to measure properties of specific absorption lines of interest. The utility of different lines depends on the effective temperature of the host star. For Sun-like stars, S and $\log R'_{HK}$ are popular magnetic activity indicators based on emission in the core of the Calcium II H and K lines (Noyes 1984). For cooler stars, the Hα line is a more useful indicator (Kürster et al. 2003, Robertson et al. 2016). The unsigned magnetic field estimated from the spectrum can be a powerful indicator of $\mathbf{RV}_{\rm contam}$ (Haywood et al. 2020) and might generalize better than any individual line.

Santerne et al. (2015) and Lanza et al. (2018) find that activity indicators can have relatively weak correlations with RV signals contaminating the data. One contributing factor is the expected time lag between RVs and other indicators, noted in the case of RV and photometry (Queloz et al. 2001, Santos et al. 2003, Queloz et al. 2009) and other indicators (Suárez Mascareño et al.

2017). This time lag can be partially mitigated in models where the derivative of the latent GP appears (Rajpaul et al. 2015, Delisle et al. 2022) or with adaptive correlation Simola et al. (2022).

6.2.2. Methods using stellar variability indicators. More generally, spectra can be reduced to a set of summary statistics that serve as stellar variability indicators, and their time series can be analyzed jointly with the RVs. By utilizing temporal information, we may obtain less noisy indicators.

The most common reduction is to compute the CCF. However, good indicators should be insensitive to true Doppler shifts, so Collier Cameron et al. (2021) proposed the Scalpels method to analyze the autocorrelation of the CCF. The resulting time series are analyzed with a PCA, and the PCA scores are used as variability indicators. Since the autocorrelation function is insensitive to shifts, true Doppler shifts will not affect the resulting variability indicators. de Beurs et al. (2020) explored more flexible supervised learning approaches to predicting $\mathrm{RV}_{\mathrm{contam}}$ from a training data set of either simulated or solar data CCFs. It is unclear whether the greater flexibility of neural networks will outweigh their added complexity and difficulty of interpretation relative to linear regression or Scalpels, particularly given the limited size of data sets available for stars other than the Sun.

Cretignier et al. (2022) transform the spectra into shells instead of CCFs, in an effort to reduce the amount of information lost when averaging lines of different depths. The PCA scores for the shell are used as spectral indicators but are first orthonormalized with respect to the shell corresponding to a pure Doppler shift.

More research is needed to develop effective means of comparing choices of summary statistics, variability indicators, and corresponding likelihoods. Several methods were explored in the EXPRES Stellar Signals Project, where several teams analyzed observations of four stars by EXPRES (Zhao et al. 2022). One key finding was that while many methods could reduce the root mean square RV of observations, the estimates of $\mathbf{RV}_{\mathrm{contam}}$ differ significantly across methods. Without knowing the true velocity, it was impossible to determine which methods are best. In Section 7.2, we present several ideas for a robust comparison of the different methods leveraging the existing and upcoming RV observations of the Sun with different instruments.

Sun-as-a-star observations allow the testing and validation of methods for mitigating stellar variability. Three methods, Scalpels (Collier Cameron et al. 2021), linear regression on CCFs, and neural networks (de Beurs et al. 2020), have been shown to significantly reduce the level of RV variability in solar observations. Existing and upcoming comparisons of Sun-as-a-star observations from multiple instruments, including a new generation of more highly stabilized spectrographs, will help disentangle solar RV variations due to solar variability and from instrument specific signals, as well as to refine the comparisons of different analysis methods.

6.2.3. Methods using the time series of spectra. Rajpaul et al. (2020) propose an alternative approach of measuring a ΔRV from each pair of spectra and reconstructing the RV time series (modulo a constant), while giving greater weight to pairs that are more similar. In order to reduce the computational cost, they split the spectra into many small chunks to be analyzed separately.

In principle, one could perform inference on the entire spectroscopic time series. Jones et al. (2020) introduced Doppler-constrained PCA to characterize stellar variability and provided a proof of concept on simulated solar observations. Bedell et al. (2019) applied a similar method to real observations, emphasizing modeling of telluric contamination and neglecting stellar variability. They regularized the variables using ℓ_1- and ℓ_2-norm constraints. More recent work has developed computationally efficient implementations that simultaneously model both stellar and telluric variability, while also accounting for the instrumental profile and allowing for physically informed regularization schemes.

7. CONCLUSION

7.1. Summary

RV is poised to play a key role in the study of exoplanets in the next decade as a primary way to measure their masses; to detect interesting, nontransiting planets; to have a more complete view of planetary system architectures; and to provide targets whose atmospheres will be further characterized by spectro-imaging. Currently, characterizing an Earth twin is out of reach, and improving data analysis techniques will play a fundamental role.

We presented the different steps of RV data analysis, separating them in three problems (see Section 3): (*a*) reducing the information of the spectrum into an RV time series, (*b*) modeling nuisance signals and the prior information on planetary and nuisance parameters, and (*c*) deciding how many planets are present and what their orbital elements are. Each of these steps requires numerical methods, where convergence should be carefully checked. Since there are multiple reasonable choices for *a*, *b*, and *c*, researchers should perform multiple analyses with different assumptions to determine whether key conclusions are sensitive to these choices, especially of the likelihood and prior.

7.2. Future Work

The exoplanet community is now well equipped to analyze RV observations when planetary signals dominate stellar and instrumental variability. When these planetary signals and corrupting ones are of similar amplitude, the fact that different methods yield different results (Zhao et al. 2022) shows that more work remains to be done. We believe that further research in steps *a* and *b* is crucial.

The ability of RVs to detect and measure the mass of potential Earth twins is critically linked to how precisely the contaminating RV signals, the RV signal not due to the motion of the center of mass of the star (RV_{contam} in Equations 7–8), can be predicted and how accurately the uncertainty on these predictions can be quantified. Correcting these signals can be done either in a statistical framework by building a likelihood function or with supervised learning techniques trained to predict the RV contamination signal from spectral shapes. The new models may be data-driven or physics-driven, especially for RV signals originating from the star, for which there is a substantial modeling effort (see Section 2.4.1) yet to be translated to data analysis techniques. Further research in step *a* is also required to ensure that all the information contained in the spectra is used and that the RV and indicators derived are reliable summary statistics.

We believe that one of the most critical questions is how to validate choices for each processing step as effective tools for detecting and characterizing low-mass planets in the presence of complex noise. We propose three ways forward, the first two concerning the validation of reduction and modeling (steps *a* and *b*) and the third the whole process (*a–c*).

The first approach is Bayesian model comparison, to evaluate the relative merits of different stellar activity and instrument systematics models. For example, since we know the true solar velocity, we can evaluate the signal RV_{contam} of each method applied to Sun-as-a-star observations. A similar approach may be possible for some other stars hosting a massive planet on an eccentric orbit, since one can exclude the possibility of additional planets for a wide range of orbital periods based on orbital stability considerations (Brewer et al. 2020, Stalport et al. 2022). Large and computationally demanding computer models could provide data for training and testing models, though they must be validated beforehand. The observations of the Sun by different high precision spectrographs can be leveraged to disentangle instrument specific noise from stellar variability.

Second, for stars with a sufficiently large number of observations, one could evaluate the accuracy and precision of a model mitigating stellar variability via its ability to predict RV_{contam} (see Equation 7) and stellar variability indicators at times not used in the training of the model. Given multiple processes operating on a variety of timescales, one must carefully design the training, testing, and validation procedure. For example, if one obtained multiple spectra per night and stellar variability were operating on timescales of weeks, then one could trivially predict the RV_{contam} from another observation on the same night, without learning how to recognize stellar variability in line shapes or depth ratios. Additionally, one must be careful that cross-validation is not undermined by researchers effectively trying many strategies and reporting results from those that appear to work best.

The two approaches above pertain to validating models for contaminating signals. The most convincing method for validating models may be via planet injection-recovery tests. One group of researchers would work on injecting planets in real or simulated data sets. Another group of researchers would blindly analyze large ensembles of simulated data sets. This would allow each proposed data reduction method and likelihood to be evaluated differentially, i.e., comparing the inferred velocities from multiple simulated data sets generated from the same true data set. Given the inevitably arbitrary nature of labeling some signals as confident exoplanet detections, it is not sufficient for teams to label which signals they believe are due to exoplanets. Instead, we recommend that they provide a list of all putative signals, along with quantitative measures of the signals' statistical significances (FAP, Bayes factor, FIP; see Section 4). Then, the number of false detections can be computed as a function of the number or properties of missed planets. To generate the data, one approach would be to remove telluric and instrumental effects, introduce a larger number of artificial Doppler shifts, reinject the telluric and instrumental effects, and generate many new synthetic data sets (perhaps adding additional noise along the way). One must be careful in implementing this approach, as an error in the injection process could become a feature learned by data-driven methods that would not be available for realistic data sets.

In this review, we adopted the viewpoint that RV data analysis should be viewed as a whole from the most basic data products to the final decisions. It will be crucial to build accurate instrument systematics and stellar noise models, to leverage as much as possible the information contained in the spectra, and to build reliable metrics to validate the different methods. Progress will most likely stem from the combination of a deep knowledge of the instruments and astronomical context combined with formal approaches, which will provide adequate tools to represent and analyze the data.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The authors thank Michaël Crétignier, who suggested presenting the RV reduction methods in three broad classes, for his feedback; Sahar Shahaf and Xavier Dumusque for their insightful suggestions; and David Hogg for suggesting to compare the prediction errors of different methods. N.C.H. acknowledges the financial support of the National Centre for Competence in Research PlanetS of the Swiss National Science Foundation (SNSF). E.B.F. acknowledges the financial support from the Heising-Simons Foundation grant #2019-1177. The Center for Exoplanets and Habitable Worlds is supported by the Pennsylvania State University and the Eberly College of Science.

LITERATURE CITED

Ahrer E, Queloz D, Rajpaul VM, Ségransan D, Bouchy F, et al. 2021. The HARPS search for southern extra-solar planets—XLV. Two Neptune mass planets orbiting HD 13808: a study of stellar activity modelling's impact on planet detection. *Mon. Not. R. Astron. Soc.* 503(1):1248–63

Aigrain S, Pont F, Zucker S. 2012. A simple method to estimate radial velocity variations due to stellar activity using photometry. *Mon. Not. R. Astron. Soc.* 419:3147–58

Anglada-Escudé G, Tuomi M. 2012. A planetary system with gas giants and super-Earths around the nearby M dwarf GJ 676A. Optimizing data analysis techniques for the detection of multi-planetary systems. *Astron. Astrophys.* 548:A58

Artigau É, Astudillo-Defru N, Delfosse X, Bouchy F, Bonfils X, et al. 2014. Telluric-line subtraction in high-accuracy velocimetry: a PCA-based approach. In *Observatory Operations: Strategies, Processes, and Systems V*, ed. AB Peck, CR Benn, RL Seaman, p. 914905. Bellingham, WA: SPIE

Astudillo-Defru N, Bonfils X, Delfosse X, Ségransan D, Forveille T, et al. 2015. The HARPS search for southern extra-solar planets. XXXVI. Planetary systems and stellar activity of the M dwarfs GJ 3293, GJ 3341, and GJ 3543. *Astron. Astrophys.* 575:A119

Baluev RV. 2008. Assessing the statistical significance of periodogram peaks. *Mon. Not. R. Astron. Soc.* 385:1279–85

Baluev RV. 2009. Accounting for velocity jitter in planet search surveys. *Mon. Not. R. Astron. Soc.* 393:969–78

Baluev RV. 2013a. Detecting multiple periodicities in observational data with the multifrequency periodogram – I. Analytic assessment of the statistical significance. *Mon. Not. R. Astron. Soc.* 436(1):807–18

Baluev RV. 2013b. Detecting non-sinusoidal periodicities in observational data: the von Mises periodogram for variable stars and exoplanetary transits. *Mon. Not. R. Astron. Soc.* 431:1167–79

Baluev RV. 2015. Keplerian periodogram for Doppler exoplanet detection: optimized computation and analytic significance thresholds. *Mon. Not. R. Astron. Soc.* 446:1478–92

Baranne A, Mayor M, Poncet JL. 1979. CORAVEL – a new tool for radial velocity measurements. *Vistas Astron.* 23:279–316

Barragán O, Aigrain S, Rajpaul VM, Zicher N. 2022. PYANETI – II. A multidimensional Gaussian process approach to analysing spectroscopic time-series. *Mon. Not. R. Astron. Soc.* 509(1):866–83

Batalha NE, Lewis T, Fortney JJ, Batalha NM, Kempton E, et al. 2019. The precision of mass measurements required for robust atmospheric characterization of transiting exoplanets. *Astrophys. J.* 885(1):L25

Bedell M, Hogg DW, Foreman-Mackey D, Montet BT, Luger R. 2019. WOBBLE: a data-driven analysis technique for time-series stellar spectra. *Astron. J.* 158(4):164

Bertaux JL, Lallement R, Ferron S, Boonne C, Bodichon R. 2014. TAPAS, a web-based service of atmospheric transmission computation for astronomy. *Astron. Astrophys.* 564:A46

Binnenfeld A, Shahaf S, Anderson RI, Zucker S. 2021. New periodograms separating orbital radial velocities and spectral shape variation. arXiv:2111.02383 [astro-ph.IM]

Boisse I, Bonfils X, Santos NC. 2012. SOAP – a tool for the fast computation of photometry and radial velocity induced by stellar spots. *Astron. Astrophys.* 545:A109

Bonomo AS, Desidera S, Benatti S, Borsa F, Crespi S, et al. 2017. The GAPS programme with HARPS-N at TNG. XIV. Investigating giant planet migration history via improved eccentricity and mass determination for 231 transiting planets. *Astron. Astrophys.* 602:A107

Borgniet S, Meunier N, Lagrange AM. 2015. Using the Sun to estimate Earth-like planets detection capabilities. V. Parameterizing the impact of solar activity components on radial velocities. *Astron. Astrophys.* 581:A133

Bouchy F, Pepe F, Queloz D. 2001. Fundamental photon noise limit to radial velocity measurements. *Astron. Astrophys.* 374:733–39

Brewer BJ, Donovan CP. 2015. Fast Bayesian inference for exoplanet discovery in radial velocity data. *Mon. Not. R. Astron. Soc.* 448(4):3206–14

Brewer JM, Fischer DA, Blackman RT, Cabot SHC, Davis AB, et al. 2020. EXPRES. I. HD 3651 as an ideal RV benchmark. *Astron. J.* 160(2):67

Butler RP, Marcy GW, Williams E, McCarthy C, Dosanjh P, Vogt SS. 1996. Attaining Doppler precision of 3 m s^{-1}. *Publ. Astron. Soc. Pac.* 108:500

Cale BL, Reefe M, Plavchan P, Tanner A, Gaidos E, et al. 2021. Diving beneath the sea of stellar activity: chromatic radial velocities of the young AU Mic planetary system. *Astron. J.* 162(6):295

Camacho JD, Faria JP, Viana PTP. 2022. Modelling stellar activity with Gaussian process regression networks. arXiv:2205.06627 [astro-ph.EP]

Cegla H. 2019. The impact of stellar surface magnetoconvection and oscillations on the detection of temperate, Earth-mass planets around Sun-like stars. *Geosciences* 9(3):114

Cegla HM, Watson CA, Shelyag S, Mathioudakis M, Moutari S. 2019. Stellar surface magnetoconvection as a source of astrophysical noise. III. Sun-as-a-star simulations and optimal noise diagnostics. *Astrophys. J.* 879(1):55

Collier Cameron A, Ford EB, Shahaf S, Aigrain S, Dumusque X, et al. 2021. Separating planetary reflex Doppler shifts from stellar variability in the wavelength domain. *Mon. Not. R. Astron. Soc.* 505(2):1699–717

Collier Cameron A, Mortier A, Phillips D, Dumusque X, Haywood RD, et al. 2019. Three years of Sun-as-a-star radial-velocity observations on the approach to solar minimum. *Mon. Not. R. Astron. Soc.* 487(1):1082–100

Connes P. 1985. Absolute astronomical accelerometry. *Astrophys. Space Sci.* 110:211–55

Correia ACM, Couetdic J, Laskar J, Bonfils X, Mayor M, et al. 2010. The HARPS search for southern extra-solar planets. XIX. Characterization and dynamics of the GJ 876 planetary system. *Astron. Astrophys.* 511:A21

Crass J, Gaudi BS, Leifer S, Beichman C, Bender C, et al. 2021. Extreme Precision Radial Velocity Working Group final report. arXiv:2107.14291 [astro-ph.IM]

Cretignier M, Dumusque X, Hara NC, Pepe F. 2021. YARARA: significant improvement in RV precision through post-processing of spectral time series. *Astron. Astrophys.* 653:A43

Cretignier M, Dumusque X, Pepe F. 2022. Stellar activity correction using PCA decomposition of shells. *Astron. Astrophys.* 659:A68

Cunha D, Santos NC, Figueira P, Santerne A, Bertaux JL, Lovis C. 2014. Impact of micro-telluric lines on precise radial velocities and its correction. *Astron. Astrophys.* 568:A35

Davis AB, Cisewski J, Dumusque X, Fischer DA, Ford EB. 2017. Insights on the spectral signatures of stellar activity and planets from PCA. *Astrophys. J.* 846:59

Dawson RI, Fabrycky DC. 2010. Radial velocity planets de-aliased: a new, short period for super-Earth 55 Cnc e. *Astrophys. J.* 722:937–53

de Beurs ZL, Vanderburg A, Shallue CJ, Dumusque X, Collier Cameron A, et al. 2020. Identifying exoplanets with deep learning. IV. Removing stellar activity signals from radial velocity measurements using neural networks. arXiv:2011.00003 [astro-ph.EP]

Delisle JB, Hara N, Ségransan D. 2020a. Efficient modeling of correlated noise. I. Statistical significance of periodogram peaks. *Astron. Astrophys.* 635:A83

Delisle JB, Hara N, Ségransan D. 2020b. Efficient modeling of correlated noise. II. A flexible noise model with fast and scalable methods. *Astron. Astrophys.* 638:A95

Delisle JB, Ségransan D, Dumusque X, Diaz RF, Bouchy F, et al. 2018. The HARPS search for southern extra-solar planets. XLIII. A compact system of four super-Earth planets orbiting HD 215152. *Astron. Astrophys.* 614:A133

Delisle JB, Unger N, Hara NC, Ségransan D. 2022. Efficient modeling of correlated noise. III. Scalable methods for jointly modeling several observables' time series with Gaussian processes. arXiv:2201.02440 [astro-ph.EP]

Díaz RF, Ségransan D, Udry S, Lovis C, Pepe F, et al. 2016. The HARPS search for southern extra-solar planets. XXXVIII. Bayesian re-analysis of three systems. New super-Earths, unconfirmed signals, and magnetic cycles. *Astron. Astrophys.* 585:A134

Dravins D, Ludwig HG, Freytag B. 2021. Spatially resolved spectroscopy across stellar surfaces. V. Observational prospects: toward Earth-like exoplanet detection. *Astron. Astrophys.* 649:A17

Dumusque X. 2018. Measuring precise radial velocities on individual spectral lines. I. Validation of the method and application to mitigate stellar activity. *Astron. Astrophys.* 620:A47

Dumusque X, Boisse I, Santos NC. 2014. SOAP 2.0: a tool to estimate the photometric and radial velocity variations induced by stellar spots and plages. *Astrophys. J.* 796:132

Dumusque X, Borsa F, Damasso M, Díaz RF, Gregory PC, et al. 2017. Radial-velocity fitting challenge. II. First results of the analysis of the data set. *Astron. Astrophys.* 598:A133

Dumusque X, Cretignier M, Sosnowska D, Buchschacher N, Lovis C, et al. 2021. Three years of HARPS-N high-resolution spectroscopy and precise radial velocity data for the Sun. arXiv:2009.01945 [astro-ph.SR]

Dumusque X, Santos NC, Udry S, Lovis C, Bonfils X. 2011a. Planetary detection limits taking into account stellar noise. II. Effect of stellar spot groups on radial-velocities. *Astron. Astrophys.* 527:A82

Dumusque X, Udry S, Lovis C, Santos NC, Monteiro MJPFG. 2011b. Planetary detection limits taking into account stellar noise. I. Observational strategies to reduce stellar oscillation and granulation effects. *Astron. Astrophys.* 525:A140

Einstein A. 1905. Zur Elektrodynamik bewegter Körper. *Ann. Phys.* 322(10):891–921

Faria JP, Santos NC, Figueira P, Brewer BJ. 2018. kima: Exoplanet detection in radial velocities. *J. Open Source Softw.* 3(26):487

Faria JP, Suárez Mascareño A, Figueira P, Silva AM, Damasso M, et al. 2022. A candidate short-period sub-Earth orbiting Proxima Centauri. *Astron. Astrophys.* 658:A115

Feng F, Tuomi M, Jones HRA. 2017. Agatha: disentangling periodic signals from correlated noise in a periodogram framework. *Mon. Not. R. Astron. Soc.* 470:4794–814

Figueira P, Santos NC, Pepe F, Lovis C, Nardetto N. 2013. Line-profile variations in radial-velocity measurements. Two alternative indicators for planetary searches. *Astron. Astrophys.* 557:A93

Ford EB. 2005. Quantifying the uncertainty in the orbits of extrasolar planets. *Astron. J.* 129:1706–17

Ford EB. 2006. Improving the efficiency of Markov chain Monte Carlo for analyzing the orbits of extrasolar planets. *Astrophys. J.* 642:505–22

Ford EB, Gregory PC. 2007. Bayesian model selection and extrasolar planet detection. In *Statistical Challenges in Modern Astronomy IV*, ed. GJ Babu, ED Feigelson, p. 189

Ford EB, Moorhead AV, Veras D. 2011. A Bayesian surrogate model for rapid time series analysis and application to exoplanet observations. *Bayesian Anal.* 6(3):475–599

Foreman-Mackey D, Agol E, Ambikasaran S, Angus R. 2017. Fast and scalable Gaussian process modeling with applications to astronomical time series. *Astron. J.* 154:220

Foreman-Mackey D, Hogg DW, Lang D, Goodman J. 2013. emcee: The MCMC hammer. *Publ. Astron. Soc. Pac.* 125:306

Fulton BJ, Rosenthal LJ, Hirsch LA, Isaacson H, Howard AW, et al. 2021. California Legacy Survey. II. Occurrence of giant planets beyond the ice line. *Astrophys. J. Suppl. Ser.* 255(1):14

Gilbertson C, Ford EB, Jones DE, Stenning DC. 2020. Toward extremely precise radial velocities. II. A tool for using multivariate Gaussian processes to model stellar activity. *Astrophys. J.* 905(2):155

Gordon TA, Agol E, Foreman-Mackey D. 2020. A fast, two-dimensional Gaussian process method based on `celerite`: applications to transiting exoplanet discovery and characterization. *Astron. J.* 160(5):240

Gregory PC. 2005a. A Bayesian analysis of extrasolar planet data for HD 73526. *Astrophys. J.* 631:1198–214

Gregory PC. 2005b. *Bayesian Logical Data Analysis for the Physical Sciences*. Cambridge, UK: Cambridge Univ. Press

Gregory PC. 2007. A Bayesian Kepler periodogram detects a second planet in HD208487. *Mon. Not. R. Astron. Soc.* 374:1321–33

Gregory PC. 2016. An apodized Kepler periodogram for separating planetary and stellar activity signals. arXiv:1601.08105 [astro-ph.SR]

Guo Z, Ford EB, Stello D, Luhn JK, Mahadevan S, et al. 2022. Modeling stellar oscillations and granulation in radial velocity time series: a Fourier-based method. arXiv:2202.06094 [astro-ph.SR]

Halverson S, Terrien R, Mahadevan S, Roy A, Bender C, et al. 2016. A comprehensive radial velocity error budget for next generation Doppler spectrometers. In *Ground-Based and Airborne Instrumentation for Astronomy VI*, ed. CJ Evans, L Simard, H Takami, p. 99086p. Bellingham, WA: SPIE

Handley WJ, Hobson MP, Lasenby AN. 2015. POLYCHORD: next-generation nested sampling. *Mon. Not. R. Astron. Soc.* 453:4384–98

Hara NC, Boué G, Laskar J, Correia ACM. 2017. Radial velocity data analysis with compressed sensing techniques. *Mon. Not. R. Astron. Soc.* 464:1220–46

Hara NC, de Poyferré T, Delisle JB, Hoffmann M. 2022a. A continuous multiple hypothesis testing framework for optimal exoplanet detection. arXiv:2203.04957 [astro-ph.IM]

Hara NC, Delisle JB, Unger N, Dumusque X. 2022b. Testing whether a signal is strictly periodic. Application to disentangling planets and stellar activity in radial velocities. *Astron. Astrophys.* 658:A177

Hara NC, Unger N, Delisle JB, Díaz R, Ségransan D. 2022c. Improving exoplanet detection capabilities with the false inclusion probability: comparison with other detection criteria in the context of radial velocities. *Astron. Astrophys.* 663:A14

Harvey J. 1985. High-resolution helioseismology. In *Future Missions in Solar, Heliospheric & Space Plasma Physics*, ed. E Rolfe, B Battrick, p. 199. Noordwijk, Neth.: ESA

Haywood RD, Collier Cameron A, Queloz D, Barros SCC, Deleuil M, et al. 2014. Planets and stellar activity: hide and seek in the CoRoT-7 system. *Mon. Not. R. Astron. Soc.* 443:2517–31

Haywood RD, Collier Cameron A, Unruh YC, Lovis C, Lanza AF, et al. 2016. The Sun as a planet-host star: proxies from SDO images for HARPS radial-velocity variations. *Mon. Not. R. Astron. Soc.* 457:3637–51

Haywood RD, Milbourne TW, Saar SH, Mortier A, Phillips D, et al. 2020. Unsigned magnetic flux as a proxy for radial-velocity variations in Sun-like stars. arXiv:2005.13386 [astro-ph.SR]

He MY, Ford EB, Ragozzine D. 2021. Friends and foes: conditional occurrence rates of exoplanet companions and their impact on radial velocity follow-up surveys. *Astron. J.* 162:5

Hirano T, Kuzuhara M, Kotani T, Omiya M, Kudo T, et al. 2020. Precision radial velocity measurements by the forward-modeling technique in the near-infrared. *Publ. Astron. Soc. Jpn.* 72(6):93

Holzer PH, Cisewski-Kehe J, Zhao L, Ford EB, Gilbertson C, Fischer DA. 2021. A stellar activity F-statistic for exoplanet surveys (SAFE). *Astron. J.* 161(6):272

Jeffreys H. 1961. *Theory of Probability*. Oxford, UK: Clarendon. 3rd ed.

Jones DE, Stenning DC, Ford EB, Wolpert RL, Loredo TJ, Dumusque X. 2020. Improving exoplanet detection power: multivariate Gaussian process models for stellar activity. arXiv:1711.01318 [astro-ph.IM]

Jurić M, Tremaine S. 2008. Dynamical origin of extrasolar planet eccentricity distribution. *Astrophys. J.* 686:603–20

Kasper M, Cerpa Urra N, Pathak P, Bonse M, Nousiainen J, et al. 2021. PCS – a roadmap for exoearth imaging with the ELT. arXiv:2103.11196 [astro-ph.IM]

Kass RE, Raftery AE. 1995. Bayes factors. *J. Am. Stat. Assoc.* 90(430):773–95

Kempton EMR, Bean JL, Louie DR, Deming D, Koll DDB, et al. 2018. A framework for prioritizing the TESS planetary candidates most amenable to atmospheric characterization. *Publ. Astron. Soc. Pac.* 130(993):114401

Kürster M, Endl M, Rouesnel F, Els S, Kaufer A, et al. 2003. The low-level radial velocity variability in Barnard's star (= GJ 699). Secular acceleration, indications for convective redshift, and planet mass limits. *Astron. Astrophys.* 403:1077–87

Lagrange AM, Desort M, Meunier N. 2010. Using the Sun to estimate Earth-like planets detection capabilities. I. Impact of cold spots. *Astron. Astrophys.* 512:A38

Lanza AF, Malavolta L, Benatti S, Desidera S, Bignamini A, et al. 2018. The GAPS Programme with HARPS-N at TNG. XVII. Line profile indicators and kernel regression as diagnostics of radial-velocity variations due to stellar activity in solar-like stars. *Astron. Astrophys.* 616:A155

Laughlin G, Chambers JE. 2001. Short-term dynamical interactions among extrasolar planets. *Astrophys. J.* 551(1):L109–13

Leleu A, Alibert Y, Hara NC, Hooton MJ, Wilson TG, et al. 2021. Six transiting planets and a chain of Laplace resonances in TOI-178. *Astron. Astrophys.* 649:A26

Lienhard F, Mortier A, Buchhave L, Collier Cameron A, López-Morales M, et al. 2022. Multi-mask least-squares deconvolution: extracting RVs using tailored masks. *Mon. Not. R. Astron. Soc.* 513(4):5328–43

Lin ASJ, Monson A, Mahadevan S, Ninan JP, Halverson S, et al. 2022. Observing the Sun as a star: design and early results from the NEID solar feed. *Astron. J.* 163(4):184

Lindegren L, Dravins D. 2003. The fundamental definition of "radial velocity." *Astron. Astrophys.* 401:1185–201

Lomb NR. 1976. Least-squares frequency analysis of unequally spaced data. *Astrophys. Space Sci.* 39:447–62

Lovis C, Dumusque X, Santos NC, Bouchy F, Mayor M, et al. 2011. The HARPS search for southern extra-solar planets. XXXI. Magnetic activity cycles in solar-type stars: statistics and impact on precise radial velocities. arXiv:1107.5325 [astro-ph.SR]

Lovis C, Fischer D. 2010. Radial velocity techniques for exoplanets. In *Exoplanets*, ed. S Seager, pp. 27–53. Tucson: Univ. Ariz. Press
Mayor M, Queloz D. 1995. A Jupiter-mass companion to a solar-type star. *Nature* 378:355–59
Meunier N. 2021. Stellar variability in radial velocity. arXiv:2104.06072 [astro-ph.SR]
Meunier N, Lagrange AM, Borgniet S, Rieutord M. 2015. Using the Sun to estimate Earth-like planet detection capabilities. VI. Simulation of granulation and supergranulation radial velocity and photometric time series. *Astron. Astrophys.* 583:A118
Mortier A, Collier Cameron A. 2017. Stacked Bayesian general Lomb-Scargle periodogram: identifying stellar activity signals. *Astron. Astrophys.* 601:A110
Murray CD, Correia ACM. 2010. Keplerian orbits and dynamics of exoplanets. In *Exoplanets*, ed. S Seager, pp. 15–23. Tucson: Univ. Ariz. Press
Nelson BE, Ford EB, Buchner J, Cloutier R, Díaz RF, et al. 2020. Quantifying the Bayesian evidence for a planet in radial velocity data. *Astron. J.* 159(2):73
Nelson BE, Ford EB, Wright JT, Fischer DA, von Braun K, et al. 2014. The 55 Cancri planetary system: fully self-consistent N-body constraints and a dynamical analysis. *Mon. Not. R. Astron. Soc.* 441:442–51
Nelson BE, Robertson PM, Payne MJ, Pritchard SM, Deck KM, et al. 2016. An empirically derived three-dimensional Laplace resonance in the Gliese 876 planetary system. *Mon. Not. R. Astron. Soc.* 455:2484–99
Noyes RW. 1984. The study of stellar global oscillations by CAII H and K variations. In *Space Research in Stellar Activity and Variability*, ed. A Mangeney, F Praderie, pp. 113–15. Paris: Obs. Paris
Papamarkou T, Lindo A, Ford EB. 2021. Geometric adaptive Monte Carlo in random environment. *Found. Data Sci.* 3(2):201–24
Pepe F, Mayor M, Galland F, Naef D, Queloz D, et al. 2002. The CORALIE survey for southern extra-solar planets. VII. Two short-period Saturnian companions to HD 108147 and HD 168746. *Astron. Astrophys.* 388:632–38
Perger M, Anglada-Escudé G, Ribas I, Rosich A, Herrero E, Morales JC. 2021. Auto-correlation functions of astrophysical processes, and their relation to Gaussian processes. Application to radial velocities of different starspot configurations. *Astron. Astrophys.* 645:A58
Perryman M. 2011. *The Exoplanet Handbook*. Cambridge, UK: Cambridge Univ. Press
Price-Whelan AM, Hogg DW, Foreman-Mackey D, Rix HW. 2017. The Joker: a custom Monte Carlo sampler for binary-star and exoplanet radial velocity data. *Astrophys. J.* 837(1):20
Quanz SP, Ottiger M, Fontanet E, Kammerer J, Menti F, et al. 2021. Large Interferometer For Exoplanets (LIFE): I. Improved exoplanet detection yield estimates for a large mid-infrared space-interferometer mission. arXiv:2101.07500 [astro-ph.EP]
Queloz D, Bouchy F, Moutou C, Hatzes A, Hébrard G, et al. 2009. The CoRoT-7 planetary system: two orbiting super-Earths. *Astron. Astrophys.* 506:303–19
Queloz D, Henry GW, Sivan JP, Baliunas SL, Beuzit JL, et al. 2001. No planet for HD 166435. *Astron. Astrophys.* 379:279–87
Rajpaul V, Aigrain S, Osborne MA, Reece S, Roberts S. 2015. A Gaussian process framework for modelling stellar activity signals in radial velocity data. *Mon. Not. R. Astron. Soc.* 452(3):2269–91
Rajpaul VM, Aigrain S, Buchhave LA. 2020. A robust, template-free approach to precise radial velocity extraction. *Mon. Not. R. Astron. Soc.* 492(3):3960–83
Rasmussen CE, Williams CKI. 2005. *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press
Robertson P, Bender C, Mahadevan S, Roy A, Ramsey LW. 2016. Proxima Centauri as a benchmark for stellar activity indicators in the near-infrared. *Astrophys. J.* 832(2):112
Rosenthal LJ, Fulton BJ, Hirsch LA, Howard A. 2019. An expanded catalog of long-period exoplanets, discovered with HIRES, Lick-Hamilton, and APF. *Bull. Am. Astron. Soc.* 51(6):303.03
Rosenthal LJ, Fulton BJ, Hirsch LA, Isaacson HT, Howard AW, et al. 2021. The California Legacy Survey. I. A catalog of 178 planets from precision radial velocity monitoring of 719 nearby stars over three decades. *Astrophys. J. Suppl. Ser.* 255(1):8
Saar SH, Donahue RA. 1997. Activity-related radial velocity variation in cool stars. *Astrophys. J.* 485(1):319–27
Sandford E, Kipping D, Collins M. 2019. The multiplicity distribution of Kepler's exoplanets. *Mon. Not. R. Astron. Soc.* 489(3):3162–73

Santerne A, Díaz RF, Almenara JM, Bouchy F, Deleuil M, et al. 2015. PASTIS: Bayesian extrasolar planet validation – II. Constraining exoplanet blend scenarios using spectroscopic diagnoses. *Mon. Not. R. Astron. Soc.* 451:2337–51

Santos NC, Udry S, Mayor M, Naef D, Pepe F, et al. 2003. The CORALIE survey for southern extra-solar planets. XI. The return of the giant planet orbiting HD 192263. *Astron. Astrophys.* 406:373–81

Scargle JD. 1982. Studies in astronomical time series analysis. II – Statistical aspects of spectral analysis of unevenly spaced data. *Astrophys. J.* 263:835–53

Schrijver CJ. 2002. Solar spots as prototypes for stellar spots. *Astron. Nachr.* 323:157–64

Schuster A. 1898. On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena. *Terr. Magn.* 3(1):13–41

Shannon C. 1948. A mathematical theory of communication. *Bell Syst. Tech. J.* 27:379–423, 623–56

Simola U, Bonfanti A, Dumusque X, Cisewski-Kehe J, Kaski S, Corander J. 2022. Accounting for stellar activity signals in radial-velocity data by using change point detection techniques. *Astron. Astrophys.* 664:A127

Smette A, Sana H, Noll S, Horst H, Kausch W, et al. 2015. Molecfit: a general tool for telluric absorption correction. I. Method and application to ESO instruments. *Astron. Astrophys.* 576:A77

Spergel D, Gehrels N, Baltay C, Bennett D, Breckinridge J, et al. 2015. Wide-Field InfrarRed Survey Telescope-astrophysics focused telescope assets WFIRST-AFTA 2015 report. arXiv:1503.03757 [astro-ph.IM]

Stalport M, Delisle JB, Udry S, Matthews EC, Bourrier V, Leleu A. 2022. A general stability-driven approach for the refinement of multi-planet systems. *Astron. Astrophys.* 664:A53

Strassmeier KG. 2009. Starspots. *Astron. Astrophys. Rev.* 17:251–308

Suárez Mascareño A, Damasso M, Lodieu N, Sozzetti A, Béjar VJS, et al. 2021. Rapid contraction of giant planets orbiting the 20-million-year-old star V1298 Tau. *Nat. Astron.* 6:232–40

Suárez Mascareño A, Rebolo R, González Hernández JI, Esposito M. 2017. Characterization of the radial velocity signal induced by rotation in late-type dwarfs. *Mon. Not. R. Astron. Soc.* 468(4):4772–81

Süveges M. 2014. Extreme-value modelling for the significance assessment of periodogram peaks. *Mon. Not. R. Astron. Soc.* 440(3):2099–114

Süveges M, Guy LP, Eyer L, Cuypers J, Holl B, et al. 2015. A comparative study of four significance measures for periodicity detection in astronomical surveys. *Mon. Not. R. Astron. Soc.* 450(2):2052–66

Tamayo D, Cranmer M, Hadden S, Rein H, Battaglia P, et al. 2020. Predicting the long-term stability of compact multiplanet systems. *PNAS* 117(31):18194–205

Tuomi M, Kotiranta S. 2009. Bayesian analysis of the radial velocities of HD 11506 reveals another planetary companion. *Astron. Astrophys.* 496:L13–16

Wise AW, Dodson-Robinson SE, Bevenour K, Provini A. 2018. New methods for finding activity-sensitive spectral lines: combined visual identification and an automated pipeline find a set of 40 activity indicators. *Astron. J.* 156:180

Wolszczan A, Frail DA. 1992. A planetary system around the millisecond pulsar PSR1257 + 12. *Nature* 355:145–47

Wright JT, Howard AW. 2009. Efficient fitting of multiplanet Keplerian models to radial velocity and astrometry data. *Astrophys. J. Suppl. Ser.* 182:205–15

Zechmeister M, Reiners A, Amado PJ, Azzaro M, Bauer FF, et al. 2018. Spectrum radial velocity analyser (SERVAL). High-precision radial velocities and two alternative spectral indicators. *Astron. Astrophys.* 609:A12

Zhao J, Ford EB. 2022. FIESTA II. Disentangling stellar and instrumental variability from exoplanetary Doppler shifts in Fourier domain. arXiv:2201.03780 [astro-ph.EP]

Zhao LL, Fischer DA, Ford EB, Wise A, Cretignier M, et al. 2022. The EXPRES stellar signals project II. State of the field in disentangling photospheric velocities. *Astron. J.* 163(4):171

Zucker S. 2015. Detection of periodicity based on serial dependence of phase-folded data. *Mon. Not. R. Astron. Soc.* 449:2723–33

Zucker S. 2018. Detection of periodicity based on independence tests – III. Phase distance correlation periodogram. *Mon. Not. R. Astron. Soc.* 474(1):L86–L90

Annual Review of Statistics and Its Application

Volume 10, 2023

Contents

Errata

An online log of corrections to *Annual Review of Statistics and Its Application* articles may be found at http://www.annualreviews.org/errata/statistics