









Gaussian Processes and Nested Sampling Applied to Kepler’s Small Long-period Exoplanet Candidates

Michael R. B. Matesic^{1,2} , Jason F. Rowe² , John H. Livingston^{3,4,5} , Shishir Dholakia⁶ , Daniel Jontof-Hutter⁷ , and Jack J. Lissauer⁸ 

¹ Département de Physique, Trottier Institute for Research on Exoplanets, Ciela Institute for Computation & Astrophysical Data Analysis, Université de Montréal, 1375 Thérèse-Lavoie-Roux Avenue, Montréal, QC, H2V 0B3, Canada; michael.matesic@umontreal.ca

² Department of Physics & Astronomy, Bishop’s University, 2600 Rue College, Sherbrooke, QC, J1M 1Z7, Canada

³ Astrobiology Center, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

⁴ National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

⁵ Department of Astronomical Science, School of Physical Sciences, The Graduate University for Advanced Studies (SOKENDAI), 2-21-1, Osawa, Mitaka, Tokyo, 181-8588, Japan

⁶ Centre for Astrophysics, University of Southern Queensland, 487-535 West Street, Darling Heights, QLD 4350, Australia

⁷ Department of Physics, University of the Pacific, 3601 Pacific Avenue, Stockton, CA 95211, USA

⁸ Space Science & Astrobiology Division, MS 245-3, NASA Ames Research Center, Moffett Field, CA 94035, USA

Received 2023 September 9; revised 2023 November 20; accepted 2023 November 25; published 2024 January 19

Abstract

There are more than 5000 confirmed and validated planets beyond the solar system to date, more than half of which were discovered by NASA’s Kepler mission. The catalog of Kepler’s exoplanet candidates has only been extensively analyzed under the assumption of white noise (i.i.d. Gaussian), which breaks down on timescales longer than a day due to correlated noise (point-to-point correlation) from stellar variability and instrumental effects. Statistical validation of candidate transit events becomes increasingly difficult when they are contaminated by this form of correlated noise, especially in the low-signal-to-noise (S/N) regimes occupied by Earth–Sun and Venus–Sun analogs. To diagnose small long-period, low-S/N putative transit signatures with few (roughly 3–9) observed transit-like events (e.g., Earth–Sun analogs), we model Kepler’s photometric data as noise, treated as a Gaussian process, with and without the inclusion of a transit model. Nested sampling algorithms from the Python UltraNest package recover model evidences and maximum a posteriori parameter sets, allowing us to disposition transit signatures as either planet candidates or false alarms within a Bayesian framework.

Unified Astronomy Thesaurus concepts: [Exoplanet detection methods \(489\)](#); [Exoplanets \(498\)](#); [Transit photometry \(1709\)](#); [Nested sampling \(1894\)](#); [Bayesian statistics \(1900\)](#); [Gaussian Processes regression \(1930\)](#)

1. Introduction

The NASA Kepler space telescope (Borucki et al. 2010; Koch et al. 2010; Borucki 2016) launched in 2009 and observed $\sim 200,000$ stars within its primary field of view over the course of roughly 4 yr. With instrumental error budgets capable of detecting Earth-sized planets in year-long orbits around Sun-like stars, Kepler aimed to directly measure the occurrence rate of such objects, otherwise known as eta-Earth (η_{\oplus} ; Borucki et al. 2010). Although the fulfillment of this objective was impeded by greater noise contamination from both stellar and instrumental effects than initially anticipated (Gilliland et al. 2011, 2015), significant progress has still been made. To aid in this effort, our work debuts a novel Bayesian framework employing nested sampling (Skilling 2004, 2006) alongside simultaneous correlated noise modeling with Gaussian processes (GPs; Stein 1999; Rasmussen & Williams 2006) to more accurately conduct dispositioning and characterization between planet candidates (PCs) and false alarms (FAs). As an aside, this study distinguishes FAs, being instrumental or astrophysical variability which mimic transit events, from astrophysical false positives (FPs), being transit-like events produced by eclipsing binary stars (EBs) and blends.

Currently, no potential Earth–Sun or Venus–Sun analog system from the Kepler sample has been shown to be reliable. Moreover, the occurrence rates for planets with $0.5 < R_p < 1.75 R_{\oplus}$ and $64 < P < 500$ days, as shown in Figure 2 of Hsu et al. (2019), are either upper bounds or detections with statistical significance less than 2 standard deviations, so extrapolation to regions of parameter space with fewer candidates would incur large statistical uncertainties. Thus, the estimate of η_{\oplus} (and η_{\oplus}) can be improved via more robust reliability estimates not only in the Earth–Sun and Venus–Sun analog bins but also in adjacent bins containing few verified planets. FAs, not astrophysical FPs such as EBs, become the primary issue for Kepler Object of Interest (KOI; Thompson et al. 2018a) discrimination in these regions (see Figure 6 of Thompson et al. 2018a). Note that published KOI catalogs do not distinguish between FAs and FPs, dispositioning both classes of objects as FPs, because their purpose is to distinguish planet candidates from noncandidates.

Having undergone thorough preconditioning via the Pre-search Data Conditioning (PDC; Twicken et al. 2010; Smith et al. 2012; Stumpe et al. 2012) module of the Kepler Science Operations Center (SOC) Science Processing Pipeline (Jenkins et al. 2010a) in an attempt to mitigate instrumental trends common among all stars on the detector, the data products of KOIs should ideally only contain intrinsic stellar variability (granulation, spots, flares, oscillations, etc.) and transiting exoplanet/eclipsing stellar binary signatures; however, instrumental systematics (sudden pixel sensitivity dropout, rolling

Table 1
Model Parameter Descriptions

Transit Model Parameters
• ρ_* : Mean stellar density (gcm^{-3}).
• q_1 : Kipping (2013) limb darkening (unitless).
• q_2 : Kipping (2013) limb darkening (unitless).
• T_0 : Transit time-series epicenter (days).
• P : Orbital period of the exoplanet (days).
• b : Impact parameter (unitless).
• R_p/R_* : Ratio of planetary and stellar radii (unitless).
• F_0 : Relative photometric zero-point offset (unitless).
Noise Model Parameters
• σ_w : Multiplicative factor applied to the photometric errors reported by DR25 (white noise; unitless).
• σ_c : Amplitude scale of Equation (6) (unitless).
• l_c : Length scale of Equation (6) (unitless).

band, bad pixels, cosmic rays, etc.)—which impact lightcurves nonuniformly—can also persist (Caldwell et al. 2010; Gilliland et al. 2011, 2015; Clarke et al. 2014; Van Cleve & Caldwell 2016; Kawahara & Masuda 2019).

As previous studies, such as Data Release 25 (DR25; Twicken et al. 2016; Mathur et al. 2017; Thompson et al. 2018a), do not model the transit event and correlated noise simultaneously *or* compute the individual reliability for any single target, their results are left susceptible to FA misidentification (Foreman-Mackey et al. 2015); instead, interpolation is performed across orbital period and multiple event statistic (MES) using population-level injection results (Christiansen 2017; Bryson et al. 2020). Another example, Caceres et al. (2019), statistically classifies Earth-sized Kepler PCs in the presence of correlated noise. However, their approach is frequentist, does not reveal any long-period PCs, and does not robustly estimate transit parameters. By analyzing individual lightcurves on a per-target basis, our work better safeguards against FAs while also improving the accuracy and robustness of PC characterization in comparison to previous population-level approaches.

Accordingly, the data of any given KOI can be interpreted as having originated from a transiting PC with some noise contamination or as a purely noise FA. To assess the probability that small long-period, low-signal-to-noise ratio (S/N) patterns of photometric dips with few (roughly 3–9) observed transit-like events (i.e., the regime that includes Earth–Sun and Venus–Sun analogs) are of astrophysical origin (i.e., represent true PCs or background/hierarchical EBs which induce transit-like dips), we model Kepler’s photometric data as noise, treated as a GP, with and without the inclusion of a transit model. These are hereby denoted as the transit plus Gaussian process (TGP) and GP models, representing PC and FA interpretations, respectively; model parameters are described in Table 1. Here, two qualitatively different models are being compared: one with a pattern of transit-shaped dips (TGP) and the other without (GP). The former yields more degrees of freedom and accordingly will fit the data more closely, but we must ask whether these additional parameters are justified. To provide a principled answer, we employ Bayesian model comparison.

Rooted in Bayes’s theorem, nested sampling algorithms from the `Python` (Van Rossum 1995a, 1995b, 1995c, 1995d; Dubois et al. 1996; Oliphant 2007) `UltraNest` (Buchner 2016, 2019, 2021)

package recover maximum a posteriori (MAP) parameter sets and evidences of each model, allowing for transit signatures to be dispositioned in terms of PC and FA probabilities within a Bayesian framework. It is important to clarify that this work does not attempt to qualify KOIs beyond PC or FA status; this is in sharp contrast to Kepler planet catalogs, which disposition FAs together with FPs.

The simultaneous modeling of correlated noise additionally provides more robust constraints on transit model parameters. Thus, the analysis that we present herein also improves the characterization of PCs, most significantly in terms of their radii.

We describe our proposed methodology herein and apply it to select Kepler targets, including potential Earth–Sun and Venus–Sun analogs (see Figure 1). In Section 2, we lay the statistical foundation for Bayesian model comparison, GPs, and nested sampling (Section 2.1) before proceeding with the construction of our TGP and GP models (Section 2.2), an overview of the software architecture (Section 2.3), and ending with a summary of how we obtain derived parameters from fitted solutions (Section 2.4). We identify a sample population of small long-period, low-S/N KOIs—including Kepler’s most Earth–Sun-like exoplanet systems, Kepler-62f (KOI-701.04; Borucki et al. 2013), Kepler-442b (KOI-4742.01; Torres et al. 2015), and Kepler-452b (KOI-7016.01; Jenkins et al. 2015)—whose preceding Markov Chain Monte-Carlo (MCMC; Metropolis et al. 1953; Hastings 1970) solutions indicate potential for Earth–Sun and/or Venus–Sun analog candidacy in Section 3. The subsections of Section 4 present and analyze TGP and GP `UltraNest` solutions relative to each other in the context of Bayesian evidences and potential biases. In Section 4.1, we interpret `UltraNest` solutions of Kepler-62f and KOI-5227.01 to establish expected behavior from strong/weak PCs. The widths of phased photometric data windows and priors have the potential to influence the recovered logged Bayes factor between models; the effects of this are explored throughout Section 4.2. This section is closed with a comparison of the Bayesian evidence against the standard metrics of MES and S/N in Section 4.3. We conclude with a summary of this paper’s leading results in Section 5, and outline future work in Section 6. A list of terminology and acronyms can be found in Appendix A.

2. Methodology

In this section, we introduce the reader to fundamental methodology upon which we base our analysis, beginning with a summary of Bayesian statistics and evidence-based model comparison in Section 2.1. Our combined treatment of white and correlated noise by use of a Gaussian distribution and Matérn 3/2 kernel GP is established next. Following this, Section 2.2 provides a breakdown of each model (TGP and GP) in terms of their parameters. A step-by-step outline of our `UltraNest` software architecture and model-fitting process for any given KOI can be read in Section 2.3. Derived parameter calculations are detailed in Section 2.4.

2.1. Model Comparison

Bayes’s theorem (Bayes & Price 1763; Laplace 1774), which forms the basis of Bayesian statistics and probability theory, describes the process by which our knowledge of an event (posterior) is probabilistically updated according to existing

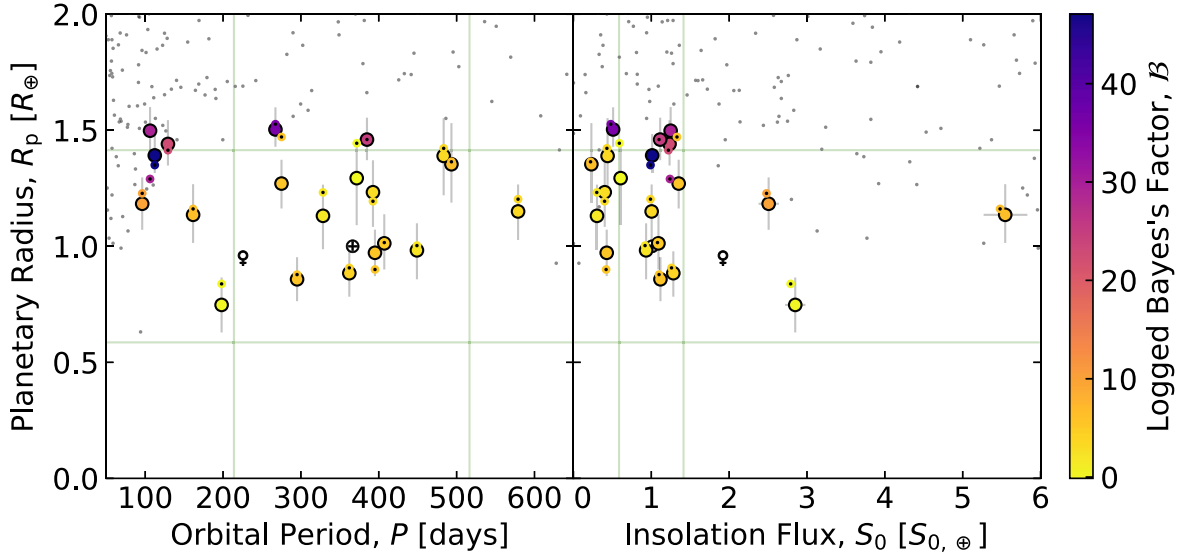


Figure 1. Our sample population of KOIs (big colored circles with black outlines) and the remaining KOI background population (small black dots), distributed according to planetary radius, R_p , orbital period, P (left), and insolation flux, S_0 (right). Note that our sample uses the newly fitted/derived results of this work, whereas background KOIs draw from the preceding MCMC solutions of Lissauer et al. (2023). Our KOIs are colored by their logged Bayes factor, \mathcal{B} —as recovered by the modeling of each individual KOI under PC (transit plus Gaussian process, TGP) and FA (Gaussian process only, GP) hypotheses—such that greater positive values indicate strong planet candidacy and vice versa for FAs, while those near zero can be interpreted as possessing inconclusive/weak significance either way (see Section 2.1). The \mathcal{B} values corresponding to our solutions are also used to outline associated Lissauer et al. (2023) results (small black dots with colored outlines) in order to facilitate a visual comparison of the *physical* parameters recovered for each KOI analyzed by both studies. Green lines outline the range within which a KOI may be deemed sufficiently “Earth–Sun-like”; these are defined according to nominal Earth values for R_p , and either P or S_0 as $x \in [1 \pm (\sqrt{2} - 1)]x_{\oplus}$. Note that all KOIs were also uniformly filtered by R_p with respect to solar values according to these same bounds. We are complete in both boxes drawn by these lines. For ease of reference, Earth (\oplus) and Venus (♀) are also plotted.

information (prior) and new observations (likelihood). In other words, it allows us to adjust our understanding of the world in order to make better informed decisions/predictions. From a statistical perspective, we can model observed data, \mathcal{D} , via the inference of model parameters, θ , using Bayes’s theorem:

$$\mathcal{P}_{\mathcal{D}}(\theta) = \frac{\mathcal{L}_{\mathcal{D}}(\theta)\pi(\theta)}{\mathcal{Z}_{\mathcal{D}}}, \quad (1)$$

where the posteriors, $\mathcal{P}_{\mathcal{D}}(\theta)$, are represented in terms of the likelihood, $\mathcal{L}_{\mathcal{D}}(\theta)$, priors, $\pi(\theta)$, and Bayesian evidence (marginal likelihood integral), $\mathcal{Z}_{\mathcal{D}}$. Here, $\mathcal{Z}_{\mathcal{D}}$ gives the probability associated with observing this realization of \mathcal{D} and is defined as

$$\mathcal{Z}_{\mathcal{D}} \equiv \int \mathcal{L}_{\mathcal{D}}(\theta)\pi(\theta)d\theta. \quad (2)$$

The $\mathcal{Z}_{\mathcal{D}}$ encodes both $\mathcal{L}_{\mathcal{D}}(\theta)$ and $\pi(\theta)$ information, so it is often employed as a metric of model suitability. Should one know the most suitable model for a given problem, the computationally expensive $\mathcal{Z}_{\mathcal{D}}$ can be readily discarded in favor of obtaining only $\mathcal{P}_{\mathcal{D}}(\theta)$ of modeled θ (e.g., likelihood-driven techniques such as MCMC). However, it is uncommon in real-world problems to possess the most suitable model with which \mathcal{D} is described in totality. As such, $\mathcal{Z}_{\mathcal{D}}$, and by extension the Bayes factor of any two models, A and B ,

$$\mathcal{B}_{A,B} \equiv \frac{\mathcal{Z}_{\mathcal{D},A}}{\mathcal{Z}_{\mathcal{D},B}}, \quad (3)$$

play a crucial role in determining the most suitable model for \mathcal{D} . This statistically robust process of model selection is known as Bayesian model comparison (Jeffreys 1939; Kass & Raftery 1995; Mackay 2003; Dunstan et al. 2020, 2022). Our study applies the

logged form of the Bayes factor:

$$\log(\mathcal{B}_{A,B}) = \log(\mathcal{Z}_{\mathcal{D},A}) - \log(\mathcal{Z}_{\mathcal{D},B}); \quad (4)$$

this quantity is defined such that positive values favor model A and negative values model B . We replace A by the TGP (PC hypothesis) model and B by the GP (FA hypothesis) model to obtain $\log(\mathcal{B}_{\text{TGP,GP}})$, whose notation we condense to \mathcal{B} :

$$\mathcal{B} \equiv \log(\mathcal{B}_{\text{TGP,GP}}). \quad (5)$$

Accordingly, increasingly positive values of \mathcal{B} promote the existence of the transiting PC while their negative counterparts suggest a FA signal originating purely from noise. Values of \mathcal{B} near zero indicate no statistically significant improvement given by the addition of transit parameters to the fit with respect to the null (noise) hypothesis; that is not to say that these are definitively PCs or FAs or that either fit is necessarily less robust, but that no statistically significant difference exists between hypotheses.

Given an informed choice of kernel (covariance function; see Rasmussen & Williams 2006), a GP may target specific behavior or systematics within a given data set—this is particularly useful when attempting to fit correlated noise present in photometric time-series observations. The squared-exponential (radial basis function; Cheney 1966; Davis 1975; Powell 1981) and Matérn (Matérn 1960) kernel families have become popular for the treatment of systematics in astronomy (Gibson et al. 2012; Roberts et al. 2012; Aigrain et al. 2015, 2016; Barclay et al. 2015; Czekala et al. 2015; Evans et al. 2015; Foreman-Mackey et al. 2017; Littlefair et al. 2017; Angus et al. 2018; Livingston et al. 2019; Brahm et al. 2023; Aigrain & Foreman-Mackey 2023; etc.); while both are

generally well suited to smooth signal applications (e.g., stellar variability), the latter is also capable of handling rougher interference (e.g., sudden pixel sensitivity dropout, as illustrated by Figure 19 of Thompson et al. 2018a). Given the known characteristics of stellar and instrumental FA sources which contaminate Kepler photometry (Gilliland et al. 2011, 2015; Van Cleve & Caldwell 2016; Van Cleve et al. 2016; Thompson et al. 2018a), we adopt the Matérn 3/2 kernel:

$$k \equiv \sigma_c^2 \left(1 + \frac{\sqrt{3}|t - t'|}{l_c} \right) \exp \left(-\frac{\sqrt{3}|t - t'|}{l_c} \right). \quad (6)$$

Here, σ_c and l_c describe the amplitude and length scales of the correlated noise with which every pair of data points, t and t' , is conditioned. White noise is incorporated as a scaling factor to the error bars belonging to each photometric observation and is obtained by fitting the standard deviation, σ_w , of a zero-mean Gaussian.

Nested sampling is a popular class of algorithm which approximates Equation (2) and provides posterior inference(s) as byproducts given \mathcal{D} , $\mathcal{L}_{\mathcal{D}}(\theta)$, and $\pi(\theta)$. Our current infrastructure makes use of UltraNest, which requires user-defined $\mathcal{L}_{\mathcal{D}}(\theta)$, and prior transforms or quantile functions mapping between physical parameter and unit hypercube sampling spaces. Uniform priors are used for all TGP/GP parameters excluding limb-darkening parameters, q_1 and q_2 , which instead use Gaussian priors (Kipping 2013).

2.2. Summary of Models

Photometric exoplanet transits were modeled using `transitfit5` (Rowe 2016). The lightcurve model uses the analytic limb-darkening transit from Mandel & Agol (2002) and assumes noninteracting Keplerian orbits. The model is parameterized with ρ_* , q_1 , q_2 , T_0 , P , b , R_p/R_* , F_0 , $\sqrt{e} \cos(\omega)$, $\sqrt{e} \sin(\omega)$, and a photometric dilution factor (see Table 1). The model can additionally include the effects of geometric albedo, ellipsoidal variations, and secondary eclipses. The calculation of Keplerian orbits derives the scaled semimajor axis, a/R_* , based on ρ_* ; this calculation assumes that the planetary mass, M_p , is much less than the mass of the host star, M_* . For all presented models in this paper, we assume (i) circular orbits (i.e., zero eccentricity, e) such that $\sqrt{e} \cos(\omega) = \sqrt{e} \sin(\omega) = 0$, (ii) no dilution (unresolved binaries),⁹ (iii) no star–planet interactions, and (iv) that the planet is completely dark (no reflection or emission). Limb-darkening parameters are from the tables of Claret & Bloemen (2011) for the Kepler bandpass. The shape information of low-S/N putative transit signatures within our regime of interest (i) leaves e and ω very weakly constrained, and (ii) makes uninformative limb-darkening priors superfluous. These assumptions and inputs to the modeling approach are similar to transit model results presented in DR25.

2.3. Software Architecture

The general step-by-step outline for the fitting of an individual KOI is detailed in this section.

1. Since the preceding MCMC architecture of Lissauer et al. (2023) modeled the transit events of pre-whitened data

⁹ DR25 lightcurves already include a crowding correction for other stars that contribute to the photometric aperture.

- rather than simultaneously fitting correlated noise and transit events as performed in this study, we treat their transit solutions as initial guesses to define focused prior widths for our transit model in UltraNest; this mitigates computationally wasteful exploration of uninformative/unlikely parameter space.
2. For noise model hyperparameters, we define physically motivated prior widths accordingly:
 - (a) While Kepler photometry typically falls within 10%–20% of the predicted white-noise budget (Gilliland et al. 2011, 2015; Van Cleve & Caldwell 2016; Van Cleve et al. 2016; Thompson et al. 2018a), we err on the side of caution with a wide uninformative prior on the scaled photometric error of $\sigma_w \in [0.2, 3.8]$.
 - (b) The amplitude scale of the correlated noise has prior width set as $\sigma_c \in [0, (F_{\max} - F_{\min})/2]$; this should not exceed the maximum flux semi-amplitude, $(F_{\max} - F_{\min})/2$, observed in a given KOI’s data set.
 - (c) The length scale of the correlated noise has prior width set as $l_c \in [2\Delta t, 2\sigma_d t_T]$; this should not fall below the measurement cadence, Δt , or exceed the phased photometric data window’s timescale, $2\sigma_d t_T$.
3. Define likelihood and prior cube transformation functions for UltraNest.
4. Set free parameters σ_d and σ_p .
5. Initialize and precompute all relevant values (i.e., GP kernel).
6. Conduct photometric time-series data preprocessing/reduction, including the following:
 - (a) Removal of other threshold-crossing events (TCEs)/KOIs associated with the same host star.
 - (b) σ_d phased photometric data window width specification.
 - (c) Linear regression removal of ramps/slopes.
 - (d) Median-based zero-point correction.
 - (e) Removal of data $\pm 5\sigma$ beyond the MCMC data-residual median value to deal with uncorrected cosmic rays, flares, or uncorrected instrumental effects following Thompson et al. (2018a).
7. Run UltraNest’s `ReactiveNestedSampler` with `RegionSliceSampler` enabled once per model (TGP and GP).

To solve for the Matérn 3/2 GP’s hyperparameters prior to each iteration’s likelihood evaluation, matrix inversion must be performed. Since the transit events are effectively isolated in time, their correlated noise components can be approximated to share negligible covariance; we represent this by means of a block-diagonal approximation to the kernel, drastically decreasing the computational burden of matrix inversions. Naturally, benefits in performance scale with the number of transit events in a given KOI’s data set. For the task of inversion, we use Cholesky decomposition—a method roughly twice as efficient as lower–upper (LU) decomposition (Cholesky 1924; Banachiewicz 1942; Press et al. 1986; Schwarzenberg-Czerny 1995). Nonetheless, each iteration is still expensive.

2.4. Derived Parameters

Formulas for derived parameters can be found within this section. We compute transit duration according to Equation (16) of Seager & Mallén-Ornelas (2003), rewritten here using

Kepler’s third law (Kepler 1619) as

$$t_{\text{T}} = \left(\frac{3P}{\pi^2 G \rho_{\star}} \right)^{\frac{1}{3}} \sqrt{\left(1 + \frac{R_{\text{p}}}{R_{\star}} \right)^2 - b^2}. \quad (7)$$

The insolation flux, $S_0 \equiv L/4\pi a^2$, can be combined with Kepler’s third law to yield

$$\frac{S_0}{S_{0,\oplus}} = \frac{L_{\star}}{L_{\odot}} \left[\frac{M_{\star}}{M_{\odot}} \left(\frac{P}{P_{\oplus}} \right)^2 \right]^{-\frac{2}{3}}. \quad (8)$$

Isochrone-derived stellar luminosity and mass from Berger et al. (2020) are used alongside our fitted orbital period to compute insolation flux via Equation (8).

Although we do not model eccentricity directly (e.g., Van Eylen et al. 2019), a minimum eccentricity may be estimated following Kipping (2014) and Torres et al. (2015) via the comparison of stellar density recovered by our model against independent estimates:

$$e_{\text{min}} \equiv \frac{|1 - (\rho_{\star,\text{model}}/\rho_{\star,\text{indep}})^{2/3}|}{1 + (\rho_{\star,\text{model}}/\rho_{\star,\text{indep}})^{2/3}}. \quad (9)$$

To obtain distributions of independent parameters, we use the reported value and lower/upper uncertainties to sample from a two-piece normal distribution (Wallis 2014). For example, the R_{p} distribution is derived via a convolution between the fitted R_{p}/R_{\star} posterior distribution and a two-piece Gaussian distribution of R_{\star} constructed from Berger et al. (2020).

3. Candidates Modeled Herein

Targets were selected from a comprehensive catalog of Kepler candidates with revised lightcurve analyses (Lissauer et al. 2023). Included within our sample are 12 PCs whose planetary and host stellar radii, R_{p} and R_{\star} , plus either orbital period, P , or insolation flux, S_0 , nominally (neglecting uncertainties) lie within $\pm(\sqrt{2} - 1)$ of Earth and solar values (e.g., $R_{\text{p}} \in [1 \pm (\sqrt{2} - 1)]R_{\oplus}$). There are just 13 KOIs identified by these criteria: 2719.02, 4742.01, 4878.01, 5554.01, 5755.01, 5971.01, 6971.01, 7179.01, 7470.01, 7591.01, 7923.01, 8107.01, and 8174.01; we have not yet analyzed KOI-5755.01 because it lacks a converged preceding MCMC solution.

The nature of this regime places candidates at significant risk of being FAs; MES $\lesssim 8$ and S/N $\lesssim 10$ are predominantly exhibited within our sample. Furthermore, KOIs-5044.01, 5971.01, 7621.01, and 7923.01 yield suspect derived transit durations, t_{T} , which diverge significantly from those expected for equatorial transits of planets on circular orbits, $t_{\text{T},c}$; regardless, similar performance here between TGP and GP models (or PC and FA hypotheses) has resulted in none of these KOIs possessing high \mathcal{B} .

In addition to those candidates listed above, we also included other small long-period PCs; among these are three of Kepler’s validated exoplanets whose characteristics most closely approach the Earth–Sun analog regime—excellent targets against which we may baseline our framework: Kepler-62f (KOI-701.04), Kepler-442b (KOI-4742.01), and Kepler-452b (KOI-7016.01). The comprehensive target list is given in Table 2 alongside fitted and derived parameters, as well as Berger et al. (2020) stellar property inputs for the free parameter choice of $\sigma_{\text{d}} = 8$ and $\sigma_{\text{p}} = 5$.

4. Numerical Results

There are three leading results to be discussed in this section, beginning with an overview of recovered TGP and GP solutions for cases of strong and weak PC evidence in Section 4.1. Here, the former is demonstrated by the baseline target, Kepler-62f, and the latter by a member of our sample KOI population, KOI-5227.01. This is followed by Section 4.2, which investigates the influence that varying the free model parameters, σ_{d} and σ_{p} , has on the \mathcal{B} ; these control phased photometric data window and prior widths, respectively. We conclude with a heuristic evaluation of the fitted \mathcal{B} against reported DR25 (or Q1–16) MES and derived S/N scores for our target population in Section 4.3. A summary of results for the complete KOI sample can be found in Figure 1 and Table 2.

4.1. Strong and Weak Cases

For the demonstration of strong and weak PC evidence, we compare the recovered UltraNest TGP and GP solutions given free parameter choices of $\sigma_{\text{d}} = 8$ and $\sigma_{\text{p}} = 5$ for Kepler-62f and KOI-5227.01. Figure 2 shows the photometric data—preprocessed according to Section 2.3—overlaid by TGP and GP solutions in unfolded original and folded GP-corrected states. Although we know Kepler-62f to be a bona fide exoplanet, both the phase (Figure 2) and corner (Figure 7) plots demonstrate competitive performance between PC and FA hypotheses; similar model performance is shown for KOI-5227.01 (see Figures 8 and 9). While it is nontrivial to individually discriminate or relatively rank Kepler-62f and KOI-5227.01 by eye, our statistical analysis places them among the strongest and weakest PCs of the $\sigma_{\text{d}} = 8$ and $\sigma_{\text{p}} = 5$ subset, with recovered \mathcal{B} values of $35.8_{-1.1}^{+1.2}$ and $-0.1_{-1.2}^{+1.1}$, respectively. This translates to strong favor for the former’s PC status, whereas the latter can be said to fall within the range of values which we deem indistinguishable, in that sufficiently strong evidence supporting either hypothesis is lacking. Our example also serves to highlight the importance of both joint-transit-noise modeling and robust Bayesian model comparison techniques—especially when working within lower-S/N regimes such as these, without which we would be more susceptible to target misclassification.

Generally, we have seen (Figures 4 and 5) two subpopulations reflective of the above comparison emerge from our analysis (strong PCs and inconclusive/weak PCs and/or FAs). Included in the strong PC group are reassuringly three known exoplanets, Kepler-62f, Kepler-442b, and Kepler-452b, in addition to two more promising but less potentially Earth–Sun-like PCs, KOI-2719.02 and KOI-6971.01. The remainder of Table 2’s candidates fall within the currently indistinguishable range of $\sim \pm 10 \mathcal{B}$; further observations and/or deeper probabilistic analyses are likely required before more definitive conclusions may be made.

It should be noted that we do not yet observe candidates whose \mathcal{B} strongly favors the FA hypothesis. We suspect this to be the result of (i) survivor bias potentially introduced by our lack of targets with MES or S/N below ~ 7 (i.e., we have yet to analyze sufficiently poor photometry) and/or (ii) a miscalibrated \mathcal{B} scale (i.e., the magnitude of the $\mathcal{Z}_{\mathcal{D}}$ penalty incurred by the TGP model’s additional parameters and, by extension, the \mathcal{B} floor are currently unknown). To obtain statistically robust conclusions to these hypotheses, future work will implement large-scale injection-recovery testing.

Table 2
Key Parameters Summary

KOI	P (days)	t_T (hr)	$t_{T,c}$ (hr)	S_0 ($S_{0,\oplus}$)	R_p (R_\oplus)	R'_p (R_\oplus)	R_* (R_\odot)	T_{eff} (K)	MES	S/N'	S/N	\mathcal{B}
701.04	267.282	7.8 ^{+0.3} _{-0.3}	7.7 ^{+4.1} _{-1.7}	0.51	1.50 ^{+0.10} _{-0.07}	1.52 ^{+0.22} _{-0.11}	0.70	4967	14.3	19	19 ⁺² ₋₂	35.8
2719.02	106.261	6.6 ^{+0.3} _{-0.3}	7.6 ^{+0.2} _{-0.2}	1.25	1.50 ^{+0.10} _{-0.10}	1.29 ^{+0.16} _{-0.13}	0.69	4601	10.0	7.7	15 ⁺² ₋₂	29.4
4742.01	112.303	5.9 ^{+0.3} _{-0.3}	7.2 ^{+0.1} _{-0.1}	1.01	1.39 ^{+0.09} _{-0.08}	1.35 ^{+0.12} _{-0.10}	0.62	4602	12.9	13	14 ⁺¹ ₋₁	47.1
4878.01	449.016	13.4 ^{+2.1} _{-1.4}	16.6 ^{+0.6} _{-0.6}	0.93	0.98 ^{+0.12} _{-0.12}	1.01 ^{+0.10} _{-0.10}	1.05	5906	7.5 ⁺	8.3	8.0 ^{+1.8} _{-1.9}	2.1
5044.01	161.533	2.1 ^{+0.4} _{-0.5}	12.7 ^{+0.5} _{-0.4}	5.55	1.14 ^{+0.13} _{-0.12}	1.16 ^{+0.16} _{-0.15}	1.11	6344	8.4 ⁺	6.9	5.9 ^{+1.1} _{-1.1}	6.4
5227.01	371.653	11.2 ^{+3.1} _{-1.8}	12.9 ^{+0.4} _{-0.3}	0.61	1.29 ^{+0.17} _{-0.20}	1.45 ^{+0.15} _{-0.11}	0.83	5486	8.4 ⁺	11	8.2 ^{+2.5} _{-2.6}	-0.1
5554.01	362.181	18.7 ^{+3.5} _{-2.6}	15.5 ^{+0.6} _{-0.5}	1.28	0.88 ^{+0.09} _{-0.10}	0.91 ^{+0.07} _{-0.07}	1.04	5945	7.3 ⁺	12	13 ⁺³ ₋₃	4.1
5704.01	96.167	2.7 ^{+0.3} _{-0.3}	7.9 ^{+0.2} _{-0.2}	2.51	1.18 ^{+0.11} _{-0.11}	1.23 ^{+0.15} _{-0.17}	0.75	5075	7.6 ⁺	7.7	6.9 ^{+1.1} _{-1.2}	9.8
5971.01	493.328	3.1 ^{+0.9} _{-0.8}	12.0 ^{+0.5} _{-0.4}	0.23	1.35 ^{+0.18} _{-0.17}	1.36 ^{+0.22} _{-0.15}	0.76	4847	7.6 ⁺	7.8	5.9 ^{+1.1} _{-1.1}	6.3
6971.01	129.222	6.9 ^{+0.4} _{-0.4}	8.0 ^{+0.2} _{-0.2}	1.23	1.44 ^{+0.10} _{-0.09}	1.42 ^{+0.13} _{-0.11}	0.68	4921	8.1	12	14 ⁺² ₋₂	22.6
7016.01	384.847	10.2 ^{+0.5} _{-0.5}	15.7 ^{+0.5} _{-0.4}	1.11	1.46 ^{+0.09} _{-0.09}	1.46 ^{+0.13} _{-0.09}	1.07	5900	7.6	12	12 ⁺¹ ₋₁	24.9
7179.01	407.093	13.9 ^{+2.6} _{-5.1}	16.1 ^{+0.6} _{-0.5}	1.09	1.01 ^{+0.13} _{-0.11}	1.02 ^{+0.11} _{-0.11}	1.06	5946	7.8	8.4	7.4 ^{+1.7} _{-1.8}	5.2
7330.01	198.141	12.2 ^{+3.6} _{-3.1}	14.7 ^{+0.5} _{-0.5}	2.85	0.75 ^{+0.12} _{-0.12}	0.84 ^{+0.08} _{-0.09}	1.19	5490	8.0	7.2	6.0 ^{+1.9} _{-1.7}	0.2
7470.01	392.553	8.2 ^{+1.4} _{-1.4}	13.3 ^{+0.4} _{-0.4}	0.40	1.23 ^{+0.14} _{-0.15}	1.20 ^{+0.16} _{-0.15}	0.82	5016	7.2	7.2	6.9 ^{+1.6} _{-1.5}	3.0
7591.01	328.339	15.7 ^{+6.1} _{-4.3}	10.6 ^{+0.2} _{-0.2}	0.30	1.13 ^{+0.13} _{-0.14}	1.24 ^{+0.12} _{-0.16}	0.65	4786	7.4	8.5	7.5 ^{+2.1} _{-1.9}	1.9
7621.01	275.075	2.9 ^{+0.2} _{-0.2}	13.4 ^{+0.5} _{-0.4}	1.35	1.27 ^{+0.10} _{-0.11}	1.47 ^{+0.18} _{-0.22}	0.97	5662	8.0	7.5	18 ⁺³ ₋₃	5.7
7716.01	483.346	15.8 ^{+2.6} _{-1.9}	14.5 ^{+0.8} _{-0.6}	0.44	1.39 ^{+0.16} _{-0.17}	1.42 ^{+0.19} _{-0.15}	0.85	5451	7.1	8.3	7.7 ^{+1.7} _{-1.8}	3.5
7923.01	395.122	21.4 ^{+0.9} _{-0.8}	13.5 ^{+0.2} _{-0.2}	0.43	0.97 ^{+0.10} _{-0.10}	0.90 ^{+0.12} _{-0.07}	0.82	5064	10.0	14	19 ⁺⁴ ₋₄	5.6
8107.01	578.916	18.8 ^{+1.5} _{-0.9}	22.8 ^{+0.9} _{-0.8}	1.00	1.15 ^{+0.12} _{-0.12}	1.20 ^{+0.09} _{-0.07}	1.35	5832	7.6	15	14 ⁺³ ₋₃	3.5
8174.01	295.061	16.5 ^{+1.3} _{-1.5}	14.9 ^{+0.4} _{-0.4}	1.12	0.86 ^{+0.10} _{-0.09}	0.88 ^{+0.09} _{-0.07}	1.04	5284	7.4	10	11 ⁺² ₋₂	5.7

Note. Key parameters summary for our KOI sample. From left to right, are KOI numbers (KOI-701.04, KOI-4742.01, and KOI-7016.01 correspond to Kepler-62f, Kepler-442b, and Kepler-452b, respectively; none of the other KOIs are validated Kepler planets), orbital period, P , transit duration, t_T , central transit duration, $t_{T,c}$, insolation flux, S_0 , planetary and stellar radii, R_p and R_* , stellar effective temperature, T_{eff} , multiple event statistic, MES, signal-to-noise, S/N, and logged Bayes factor, \mathcal{B} . Of these, the P and \mathcal{B} are fitted (inputs), t_T , $t_{T,c}$, S_0 , R_p , and S/N are derived (outputs), R_* and T_{eff} are given by Berger et al. (2020), R'_p and S/N' are given by Lissauer et al. (2023), and MES is given by DR25 where available; those KOIs not found by DR25 were also missed in Data Release 24, so their MES are instead taken from Q1–16 (Mullally et al. 2015) and identified using “*.” All fitted/derived parameters are in bold. These values are reported for the free parameter choice of $\sigma_d = 8$ and $\sigma_p = 5$.

4.2. Varying Free Parameters

There are two free parameters required to initialize our modeling pipeline, these being σ_d and σ_p . Of concern to us is their influence on the recovered posteriors and \mathcal{B} . We begin with σ_d , which acts as a multiplicative factor to the width of the phased photometric data window, defined as $2\sigma_d t_T$, where t_T is the transit duration as found in the preceding MCMC solution of Lissauer et al. (2023). In order to better model the correlated noise present within each transit event, we may leverage out-of-transit observations, which locally share noise characteristics with those in transit. Naturally, the question then arises as to how much out-of-transit data should be included when defining the phased photometric data window of any given fit? While there exists an abundance of available out-of-transit data, two constraining factors must be considered: (i) computational cost and (ii) information gain. While only so much can be done in terms of computing power and software optimization, we can more deeply consider the notion that correlation between in-transit and out-of-transit noise decreases with increasing distance from the transit midpoint. To simplify things, the upper bound on the GP’s l_c prior can be set as the width of the phased photometric data window. It follows that a suitable choice of phased photometric data window width then preserves the GP’s ability to accurately model timescales relevant to the transit event(s) and subsequent statistical meaning of TGP-GP model comparison; in other words, the possibility of solutions preferring longer-timescale fluctuations with little in-transit information is mitigated.

Core to the Bayesian approach of statistical model comparison is the consideration of a priori knowledge, as seen in Equation (2). While intended to be used with informative (nonuniform) priors, it is not uncommon to lack the independent parameter constraints and distributions necessary for this. Such is the case in this study for all transit and noise model parameters, excluding q_1 and q_2 , whose prior distributions are precalculated with well-defined support according to Kipping (2013). Varying searchable parameter space does not pose much of an issue apart from sampling inefficiencies when employing informative priors, as exploration beyond their regions of significant probability density with naturally well-defined support returns little to no information. The same cannot be said for uniform priors, whose normalization biases measurements of the \mathcal{B} via Equation (2) (for a detailed study regarding priors and caveats such as this within the context of Bayesian inference and model selection, see Llorente et al. 2022).

To explore this implication, we propose a two-step method of quantile standardization to assess the potential bias induced by choices of σ_p with fixed $\sigma_d = 2$. First, overly wide priors are cast using $\sigma_p = 25$ in an attempt to fully capture the desired empirical posteriors. Once recovered, these can be used to provide accurate quantiles as

$$q(\sigma_p) \equiv \frac{1}{2} \left[1 + \text{erf} \left(\frac{\sigma_p}{\sqrt{2}} \right) \right] \times 100\%, \quad (10)$$

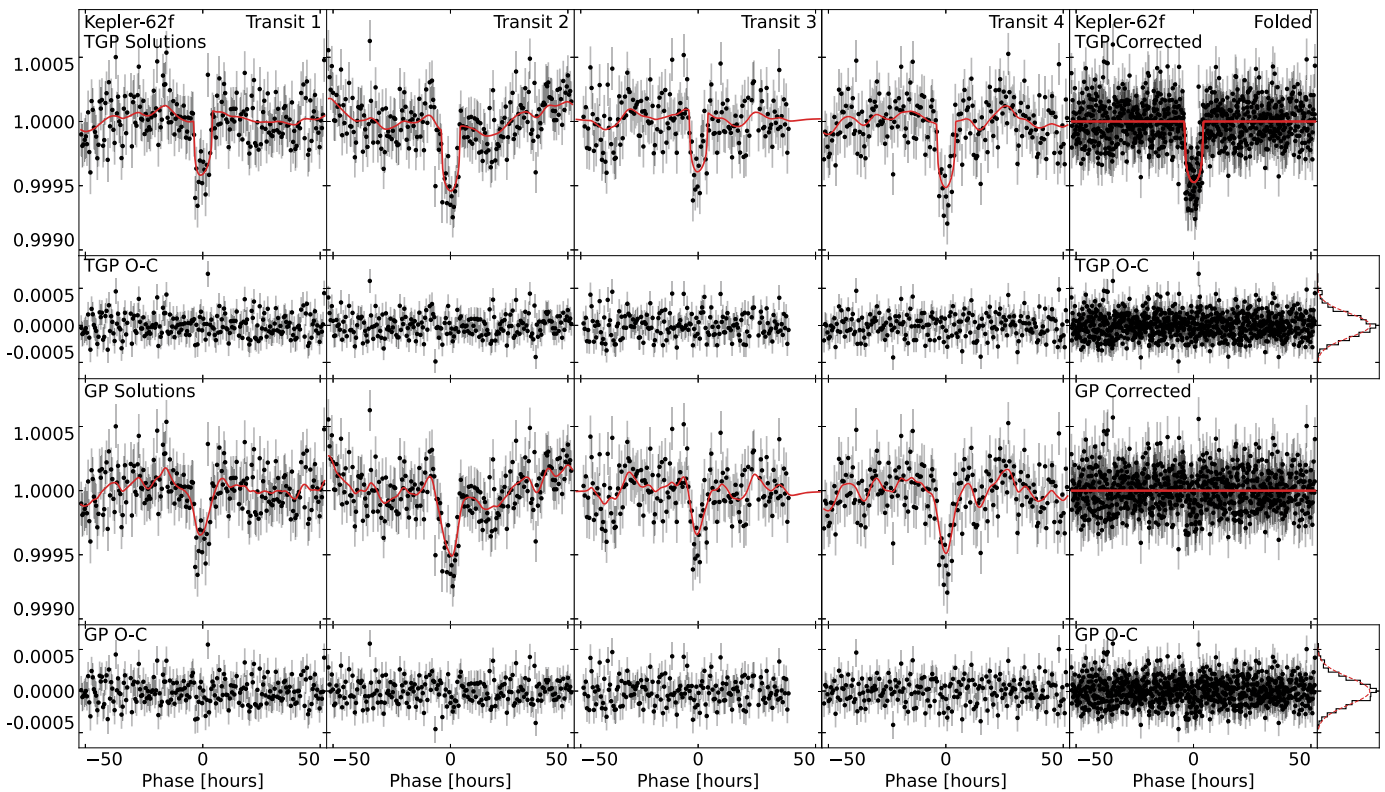


Figure 2. Kepler’s photometric PDC data (black) for all observed transit events of Kepler-62f with $\sigma_d = 8$ and $\sigma_p = 5$, overlaid by MAP (red) TGP (PC hypothesis; top) and GP (FA hypothesis; bottom) model solutions alongside corresponding O-C (data – MAP) residuals. GP-corrected phase-folded results are shown in the rightmost column and are accompanied by O-C residuals. O-C residual histograms are overlaid by Gaussian distributions with zero mean and TGP or GP MAP-scaled median photometric error standard deviations (dashed red) in order to help identify signs of overfitting (i.e., non-Gaussian O-C residuals).

with which the prior widths of subsequent runs can be defined; this notation is not to be confused with the limb-darkening parameters, q_1 and q_2 . In our code, these quantiles are set by the corner quantile(x , q , weights) function, with x and weights arguments given by $\sigma_d = 2$, $\sigma_p = 25$ UltraNest empirical posteriors and q by $1 - q(\sigma_p)$ or $q(\sigma_p)$ for lower and upper bounds, respectively. A detailed example of this process with accompanying visuals can be found in Figure 3.

To assess correlation between the \mathcal{B} and σ_d , the sample population of KOIs was fit for $\sigma_d \in [2, 4, 8, 16]$ and fixed $\sigma_p = 5$. The resulting emergence of two KOI subpopulations was immediately apparent (see Figure 4), these being strong PCs with promising follow-up potential (red) and inconclusive/weak PCs and/or FAs (blue). Generally, the PC group demonstrates positive trajectory with respect to σ_d across all metrics, whereas the latter group evolves in a relatively flat fashion. Here, the \mathcal{B} exhibits a clearly defined subpopulation boundary at a value of approximately 10. Overall, this behavior suggests that the dispositioning of PCs from FAs is largely independent of the chosen σ_d , meaning that smaller phased photometric data windows may be favored to reduce the computational burden while retaining sufficient information; this additionally supports a transit-like timescale upper bound on the GP’s l_c .

In testing how the variance of prior width affects our results, we use quantiles of $\sim 68.2689\%$, 95.4410% , 99.7300% , 99.9937% , and 99.9999% for prior widths of subsequently independently computed nested sampling runs corresponding to $\sigma_p \in [1, 5]$ and fixed $\sigma_d = 2$. As expected, the \mathcal{B} possesses a

general inverse proportionality to σ_p resulting from Equation (2) (see Figure 5). Conveniently, however, this trend is roughly similar across the entire KOI sample population, indicating that sample populations with prior widths set by a consistent/shared choice of quantile will experience a population-wide shift in \mathcal{B} plus some small variance of $\sim \pm 5$ such that relative ranking between targets are still valid. Interestingly, it seems that this variance is reduced to stochastic order with greater probability of PC status, meaning that we can be relatively more confident in our findings for the red subpopulation. While these tests must be performed on a much larger KOI population to draw any statistically robust conclusions, if the general inverse trend is to hold, one could theoretically fit and recover an empirical correction relating the \mathcal{B} and σ_p . For now, there does not seem to be a preferred σ_p in terms of bias reduction, but we can recommend quantiles corresponding to at least $\sigma_p = 3$ in order to promote sufficient prior-posterior information gain and properly recovered posteriors.

Finally, Figures 4 and 5 both suggest the presence of five strong PCs: three known exoplanets, Kepler-62f, Kepler-442b, and Kepler-452b, and two new additions residing within the habitable zones of their host stars, KOI-2719.02 and KOI-6971.01. KOI-5704.01 and KOI-7621.01 are also noteworthy: The former consistently crests the \mathcal{B} boundary of ~ 10 , whereas the latter displays anomalous behavior in that it draws a steep upward trajectory in \mathcal{B} characteristic of the strong PC subpopulation across σ_d while only escaping the otherwise clearly defined boundary at $\sigma_d = 16$; this is not observed for any other member of the inconclusive/weak subpopulation.

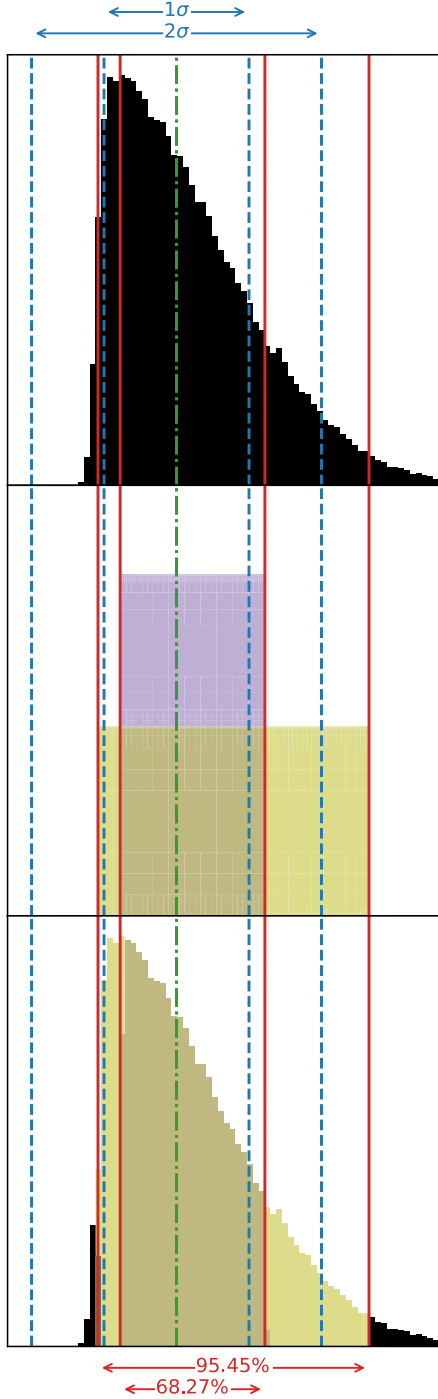


Figure 3. A single-parameter mock example of the process used to obtain quantile-based prior widths for standardization tests conducted in Figure 5. In the first nested sampling run, a wide net is cast to ensure the parameter’s empirical posterior distribution (black) is captured in its entirety (top). In subsequent runs, quantiles (solid red) of this complete posterior are used to define subsequent prior widths (middle), with which associated fractions of the complete posterior are recovered (bottom). While demonstrated with uniform priors, this is not a constraining factor; we apply this methodology to both uniform and nonuniform priors. In this example, skewed, non-Gaussian posterior behavior causes standard deviations (dashed blue) from the complete posterior’s median (dashed-dotted green) to be inefficient in capturing parameter space information compared to quantiles. Since recovered posteriors are not guaranteed to be Gaussian in nature, we ensure consistent information gain by focusing only on relevant regions of parameter space via quantiles. The middle panel also visualizes how the normalized amplitude of a uniform prior is dictated by its width; this is a known origin of bias in Equation (2).

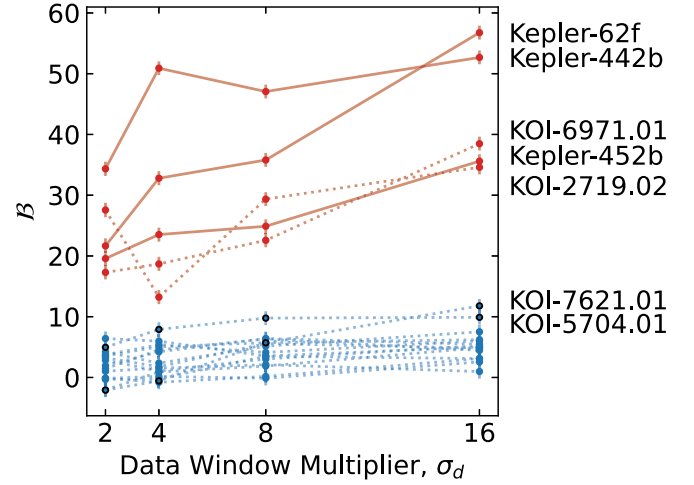


Figure 4. The logged Bayes’ factor, \mathcal{B} , of each KOI listed in Table 2 as they vary with the phased photometric data window multiplier, σ_d . Red and blue subpopulations correspond to strong PCs and periodic transit events possessing inconclusive/weak evidence-based model preference with respect to the PC and FA hypotheses, respectively; solid lines indicate previously validated planets.

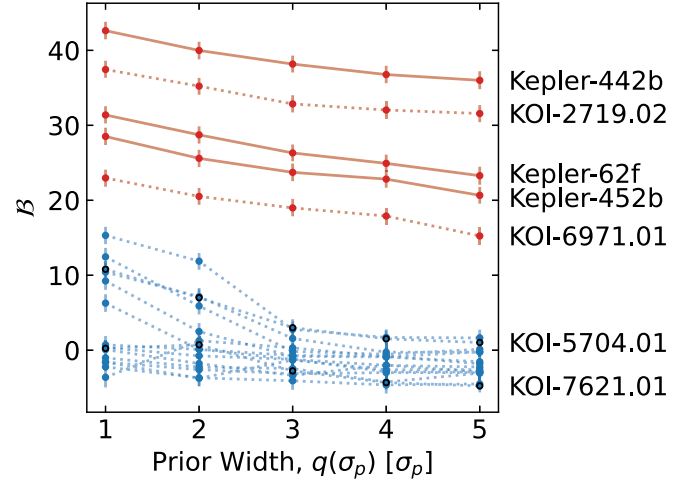


Figure 5. Same as Figure 4, but with respect to σ_p , which sets quantile-based prior widths.

If verified with current fitted parameters (see Table 2), KOI-7621.01 would rank alongside Kepler-62f, Kepler-442b, and Kepler-452b in terms of Earth–Sun analog candidacy. That being said, some points of contention must be addressed regarding this candidate: (i) we derive a nearly parabolic minimum eccentricity of $e_{\min} = 0.89^{+0.02}_{-0.04}$, and (ii) the photometric data contain long-timescale fluctuations of considerable amplitude, likely caused by spot modulation. Since eccentricity is degenerate with the mean stellar density and impact parameter—found to be $\rho_* = 76^{+26}_{-31} \text{ g cm}^{-3}$ and $b = 0.33^{+0.28}_{-0.23}$, whereas Berger et al. (2020) obtained $\rho_* = 1.00 \pm 0.11 \text{ g cm}^{-3}$ —we cannot readily conclude whether this is a highly eccentric orbit or the consequence of grazing transits.

In terms of obtaining best-fitting model parameters, fluctuations caused by spot modulation affect a greater number of data points than those caused by transit-timescale events such that our GP’s l_c is motivated toward longer timescales, thereby foregoing the ability to represent transit events in favor of

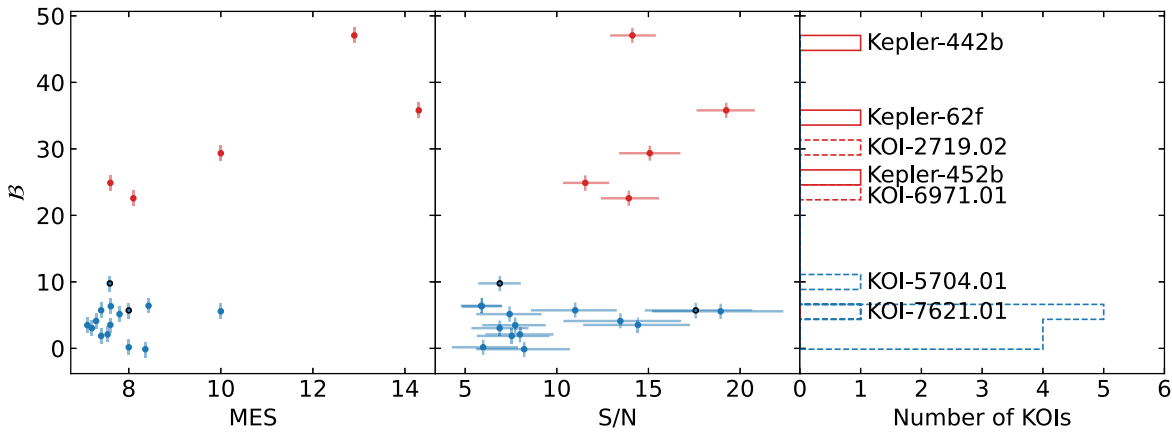


Figure 6. DR25 MES (left) and our S/N (middle) metrics compared against our logged Bayes factor, \mathcal{B} , and their histogram (right) for the KOI sample given in Table 2. Red and blue subpopulations correspond to strong PCs and periodic transit events possessing inconclusive/weak evidence-based model preference with respect to PC and/or FA hypotheses, respectively; solid lines indicate previously validated planets. Note the ample separation between strong PCs and inconclusive/weak PCs (or FAs) revealed by the \mathcal{B} , which MES and S/N are otherwise blind to.

producing an overall “better” fit. It follows in subsequent TGP-GP model comparison that the TGP model will always outperform the GP-only solution. We can then expect the \mathcal{B} to become artificially inflated with increasing σ_d in the presence of long-timescale correlated noise. Of the strong PC subpopulation, KOI-2719.02 is the only target to experience this spot modulation inflation effect; its strong PC status is not invalidated, however, as even with a subunity GP l_c from the $\sigma_d=2$ and $\sigma_p=5$ solution, its \mathcal{B} remains significant in magnitude.

Our future work will adopt more aggressive data preconditioning techniques (e.g., application of low-frequency bandpass filters) in an effort to mitigate \mathcal{B} inflation by focusing the GP on transit-like timescales. All things considered, both KOI-5704.01 and KOI-7621.01 certainly warrant further investigation in future studies.

4.3. Comparisons to Multiple Event Statistic and Signal-to-noise Ratio

For our final result, we empirically compare the Bayesian evidence approach against Q1–16 (or DR25 where available) MES and our derived S/N according to Equation (5) of Rowe et al. (2015); the latter two being standard metrics for candidate discrimination in TCE/KOI searches and catalogs. Figure 6 shows that our novel methodology presents the ability to clearly distinguish between strong PCs (red) and periodic transit events possessing inconclusive/weak evidence-based model preference with respect to PC and FA hypotheses (blue) in regimes contaminated by high levels of white and correlated noise, as demonstrated by the ample separation of these two subpopulations. This is in contrast to both the MES and S/N, which completely mix populations, save for Kepler-62f and Kepler-442b, as the only KOIs that we analyzed with $\text{MES} > 10$. Although we observe no strong PCs with derived $\text{S/N} < 10$, the appearance of inconclusive/weak PCs or FAs beyond this threshold illustrates a potential deficiency of S/N when used as a PC-FA discrimination metric in comparison to the \mathcal{B} . In the context of choosing MES or S/N cutoffs for searches, this means that strong PCs are likely to be lost and/or inconclusive/weak PCs and/or FAs included. Should this clear separation between populations hold across larger KOI samples, the \mathcal{B} could substantially reduce this blind spot.

5. Conclusions

Our analysis of targets via the simultaneous modeling of transits alongside a combined white and correlated noise GP yields fundamental transit parameters (e.g., scaled planetary radius, R_p/R_*) and Bayesian evidence-driven PC-FA model comparison in the most robust approach to date. It is then important to note that there is a discrepancy between our results and those of DR25, as illustrated by the case of Kepler-452b. While DR25 yields an estimated reliability value of $\sim 40\%$ (see Figure 11 of Thompson et al. 2018a), we recover a strongly favored PC status of $\mathcal{B} = 24.9^{+1.1}_{-1.2}$. Note that the Bayesian evidence does not directly translate to a reliability percentage, so here we simply compared the interpretations of their results; however, we aim to create a mapping between the two in future works. To do so, we next plan to conduct extensive injection-recovery testing, which will also facilitate an understanding of the \mathcal{B} floor from strong FAs.

Having performed Bayesian model comparison between PC and FA hypotheses on the periodic transit-like photometric events of each KOI listed in Table 2, we report strong PC dispositions of Kepler-62f, Kepler-442b, and Kepler-452b—agreeing with preexisting studies—plus the two new additions of KOI-2719.02 and KOI-6971.01, as well as two moderately strong PCs, KOI-5704.01 and KOI-7621.01.

Preliminary testing indicates a demand for the choice of free model parameters, σ_d and σ_p , to be shared across any given sample population of KOIs in order to promote statistically sound comparisons between targets. Furthermore, smaller phased photometric data windows (lower σ_d) and consistent quantile-based prior widths likely mitigate potential biases.

The recovered posteriors of fitted/derived parameters were used to obtain a statistical description of the S/N with uncertainties on a per-target basis, rather than its point-estimate counterpart commonly reported in previous studies. The TGP approach also yields similar to significantly improved values of S/N with respect to those reported by Lissauer et al. (2023; e.g., see KOI-2719.02 and KOI-7621.01 in Table 2).

That being said, both MES and S/N exhibited vulnerability to candidate misidentification, whereas the \mathcal{B} was able to clearly distinguish strong PCs from inconclusive/weak PCs and/or FAs. Regardless of whether the \mathcal{B} is adopted as a standard metric in PC-FA dispositioning, the MES and S/N

should undergo additional investigation and be used thoughtfully.

6. Next Steps

When allocated one node (32 CPU threads) on high-performance computing clusters, our current nested sampling infrastructure sees typical per-target timescales on the order of a week. As such, our future work will instead rely on the development of a simulation-based inference (SBI; Cranmer et al. 2020) machine-learning infrastructure; these have seen great success in recent years (see Alsing et al. 2018, 2019; Miller et al. 2020; Tejero-Cantero et al. 2020; Miller et al. 2022; Legin et al. 2023b). The amortized nature of SBI will allow for computationally efficient deployment across parameter space in catalog-wide applications to current and future missions (Kepler, K2, TESS, PLATO, etc.). Cutting-edge supporting frameworks/methodologies (see McEwen et al. 2021; Legin et al. 2023a; Jeffrey & Wandelt 2023) will facilitate the core Bayesian evidence-based approach debuted here.

Acknowledgments

We thank Steve Bryson and Laurence Perreault-Levasseur for their assistance, consultation, and contributions. MRBM received financial support from a Bishop’s University Foundation Graduate Entrance Scholarship and a Fondation Arbour Master’s Scholarship. JFR acknowledges support from the NSERC Discovery program and the Canada Research Chair program. This research made use of Digital Research Alliance of Canada (DRAC; <https://alliancecan.ca/en>) computer resources from a RAC allocation to JFR. This research has made use of the NASA Exoplanet Archive (<https://exoplanetarchive.ipac.caltech.edu>), which is operated by the California Institute of Technology under contract with the National Aeronautics and Space Administration through the Exoplanet Exploration Program. The specific observations analyzed can be accessed via [10.26093/cds/vizier.22350038](https://cds.u-strasbg.fr/arc/abstract/10.26093/cds/vizier.22350038) (Thompson et al. 2018a, 2018b).

Facilities: Astrophysics Data System (NASA ADS), NASA 0.95m Kepler Satellite Mission (Borucki et al. 2010; Koch et al. 2010; Borucki 2016), [NASA Exoplanet Archive](https://exoplanetarchive.nasa.gov) (NASA 2022)

Software: ChainConsumer (Hinton 2016); corner (Foreman-Mackey 2016); Matplotlib (Hunter 2007); Numba (Lam et al. 2015); NumPy (Harris et al. 2020); pandas (McKinney 2010; The Pandas Development Team 2020); Python (Van Rossum 1995a, 1995b, 1995c, 1995d; Dubois et al. 1996; Oliphant 2007); SciPy (Virtanen et al. 2020); transitfit5 (Rowe 2016); UltraNest (Buchner 2016, 2019, 2021).

Appendix A Terminology and Acronyms

A list of terminology and acronyms alongside their corresponding definitions can be found here.

1. *Argument of periapsis*, ω . Angle from ascending node to periapsis along the direction of motion for a given orbiting body.
2. *Bayesian evidence*, $\mathcal{Z}_{\mathcal{D}}$. The model-agnostic probability of observing the data, \mathcal{D} (see Equations (1) and (2)).

3. *Central transit duration*, $t_{T,c}$. The duration, t_T , of an equatorial ($b = 0$) transit.
4. *Correlated noise amplitude scale*, σ_c . The Matérn 3/2 kernel GP unitless amplitude scale (see Equation (6); fitted noise model parameter).
5. *Correlated noise length scale*, l_c . The Matérn 3/2 kernel GP unitless length scale (see Equation (6); fitted noise model parameter).
6. *Data window width free parameter*, σ_d . The factor used to define the phased photometric data window in terms of transit durations, t_T ’s, out from the transit midpoint, as discussed in Section 4.2.
7. *Eccentricity*, e . The value describing orbital shape, ranging from circular ($e = 0$) to elliptical ($0 < e < 1$) to parabolic ($e = 1$).
8. *Eccentricity projections*, $\sqrt{e} \cos(\omega)$ and $\sqrt{e} \sin(\omega)$. Transit model parameters for noncircular orbits representing eccentricity, e , vector components projected with respect to the argument of periapsis, ω .
9. *Effective temperature*, T_{eff} . The average surface temperature of an object (e.g., star), given in this study by Berger et al. (2020).
10. *Ephemeris*, T_0 . The KOI’s time-series epicenter in units of days (fitted transit model parameter).
11. *Eta-Earth (or eta-Venus)*, η_{\oplus} (or η_{\oplus}). The occurrence rate of Earth-like (or Venus-like) planets around Sun-like stars.
12. *False alarm*, *FA*. A periodic transit-like signal caused by instrumental and/or stellar noise/variability.
13. *False positive*, *FP*. A periodic transit-like signal caused by physical sources other than a transiting exoplanet (e.g., eclipsing stellar binary signatures).
14. *Gaussian process model*, *GP*. The FA hypothesis model (see Table 1).
15. *Impact parameter*, b . The unitless projection of a/R_* with respect to the orbital inclination, i (fitted transit model parameter).
16. *Insolation flux*, S_0 . The measure of incident solar radiation on a surface or body (e.g., exoplanet; see Equation (8)).
17. *Kepler Object of Interest*, *KOI*. A periodic transit-like event or TCE that warrants further review (Thompson et al. 2018a).
18. *Likelihood*, $\mathcal{L}_{\mathcal{D}}(\theta)$. The probability which quantifies how strongly the data, \mathcal{D} , supports the modeled parameters, θ (see Equation (1)).
19. *Limb darkening*, q_1 and q_2 . The unitless Kipping (2013) reparameterization of Mandel & Agol (2002) limb darkening (fitted transit model parameters).
20. *Logged Bayes factor*, \mathcal{B} . The logged Bayes factor representing the difference in logged Bayesian evidences, $\mathcal{Z}_{\mathcal{D}}$, between any two models (e.g., TGP and GP) applied to the same data set, \mathcal{D} (see Equations (3), (4), and (5)).
21. *Maximum a posteriori*, *MAP*. The most probable set of modeled parameters as given by Bayes’s theorem (see Equation (1)).
22. *Mean stellar density*, ρ_* . The mean stellar density in units of grams per cubic centimeter (fitted transit model parameter).
23. *Multiple event statistic*, *MES*. A measure describing the combined significance of all observed transits in the detrended and whitened lightcurve with the assumption

- of a linear ephemeris, T_0 (Jenkins 2002; Thompson et al. 2018a).
24. *Orbital period*, P . The KOI's time-series orbital period in units of days (fitted transit model parameter).
 25. *Photometric zero point*, F_0 . The relative (unitless) photometric zero-point offset (fitted transit model parameter).
 26. *Planet candidate*, PC . A KOI which has passed FA vetting procedures but has yet to undergo/pass FP vetting and/or be confirmed by alternative observation techniques.
 27. *Planetary radius*, R_p . The radius of the companion exoplanet.
 28. *Posterior distribution*, $\mathcal{P}_{\mathcal{D}}(\theta)$. The updated probability of modeled parameters, θ , given new data, \mathcal{D} , and informed by combining the likelihood, $\mathcal{L}_{\mathcal{D}}(\theta)$, priors, $\pi(\theta)$, and Bayesian evidence, $\mathcal{Z}_{\mathcal{D}}$ (see Equation (1)).
 29. *Prior distribution*, $\pi(\theta)$. The initial probability or belief about given model parameters, θ , before any new data, \mathcal{D} , is taken into account (see Equation (1)).
 30. *Prior width free parameter*, σ_p . The factor used to define fitted parameter prior widths for UltraNest in terms of MCMC-recovered standard deviations with respect to their maximum-likelihood estimator values as given by Lissauer et al. (2023) before quantile-defined widths are obtained using Equation (10), as discussed in Section 4.2.
 31. *Scaled planetary radius*, R_p/R_* . The unitless ratio of companion planetary and host stellar radii (fitted transit model parameter).
 32. *Scaled semimajor axis*, a/R_* . Unitless ratio of the companion exoplanet's semimajor axis scaled with respect to the host's stellar radius.
 33. *Signal-to-noise*, S/N . The quantification of a desired signal's quality with respect to the level of unwanted noise contamination (see Equation (5) of Rowe et al. 2015).
 34. *Stellar radius*, R_* . The radius of the host star, given in this study by Berger et al. (2020).
 35. *Threshold-crossing event*, TCE . A periodic signal identified by the Transiting Planet Search (Jenkins et al. 2010b; Twicken et al. 2016; Jenkins 2020) module of the SOC Science Processing Pipeline (Jenkins et al. 2010a).
 36. *Transit plus Gaussian process model*, TGP . The PC hypothesis model (see Table 1).
 37. *Transit duration*, t_T . The total time taken for the (exoplanet) companion to occult its host (star) from ingress to egress (i.e., beginning to end; see Equation (7)).
 38. *White-noise amplitude scale*, σ_w . The unitless scaling factor to DR25-reported photometric errors (fitted noise model parameter).

Appendix B Strong and Weak Cases Supplementary Figures

Appendix B contains Figures 7–9.

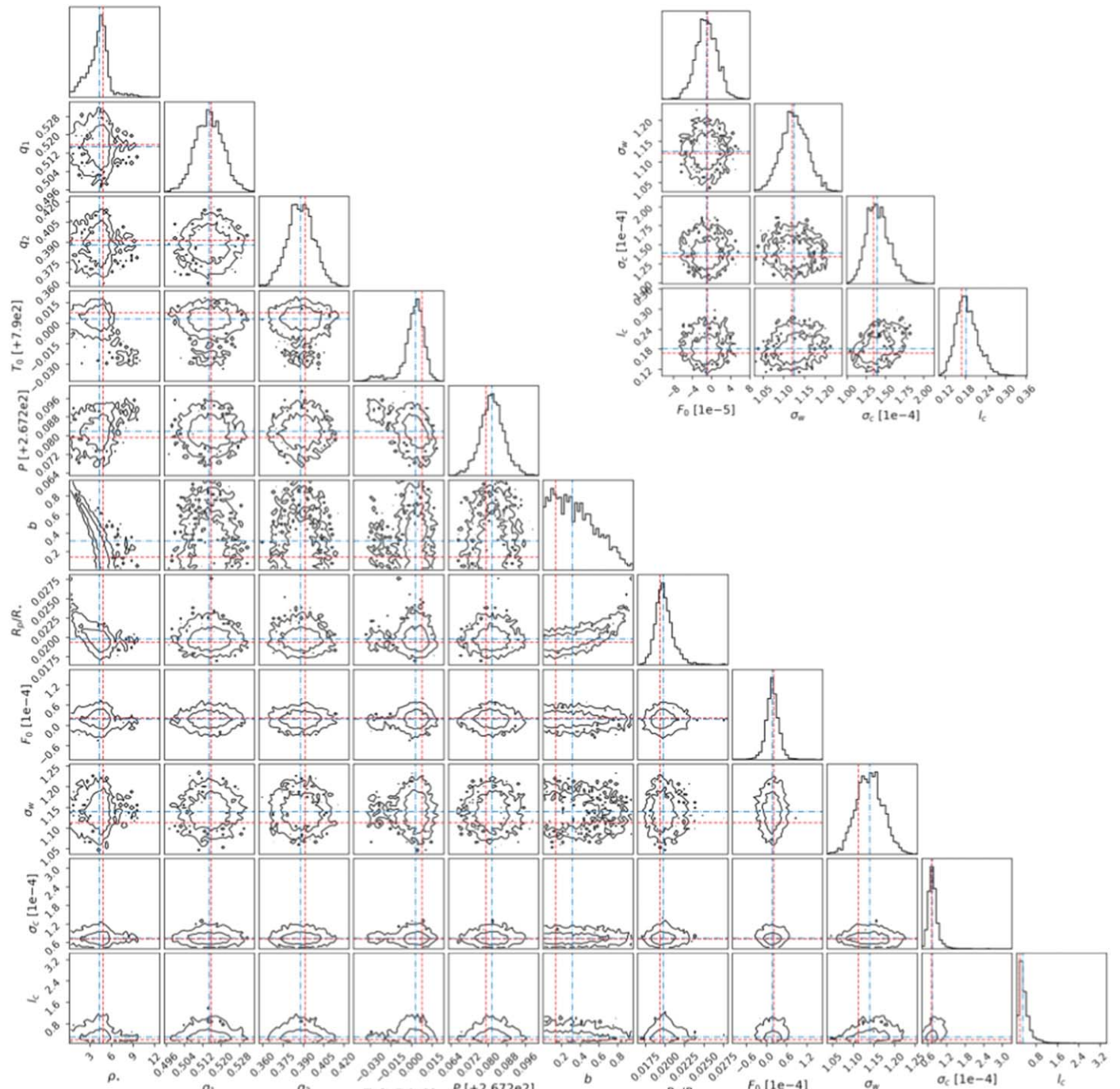


Figure 7. $\sigma_d = 8$, $\sigma_p = 5$ Kepler-62f TGP (bottom left) and GP (top right) corner plots depicting parameter behavior/covariance, with empirical posterior distributions overlaid by MAP (red) and median (blue) solutions above each column.

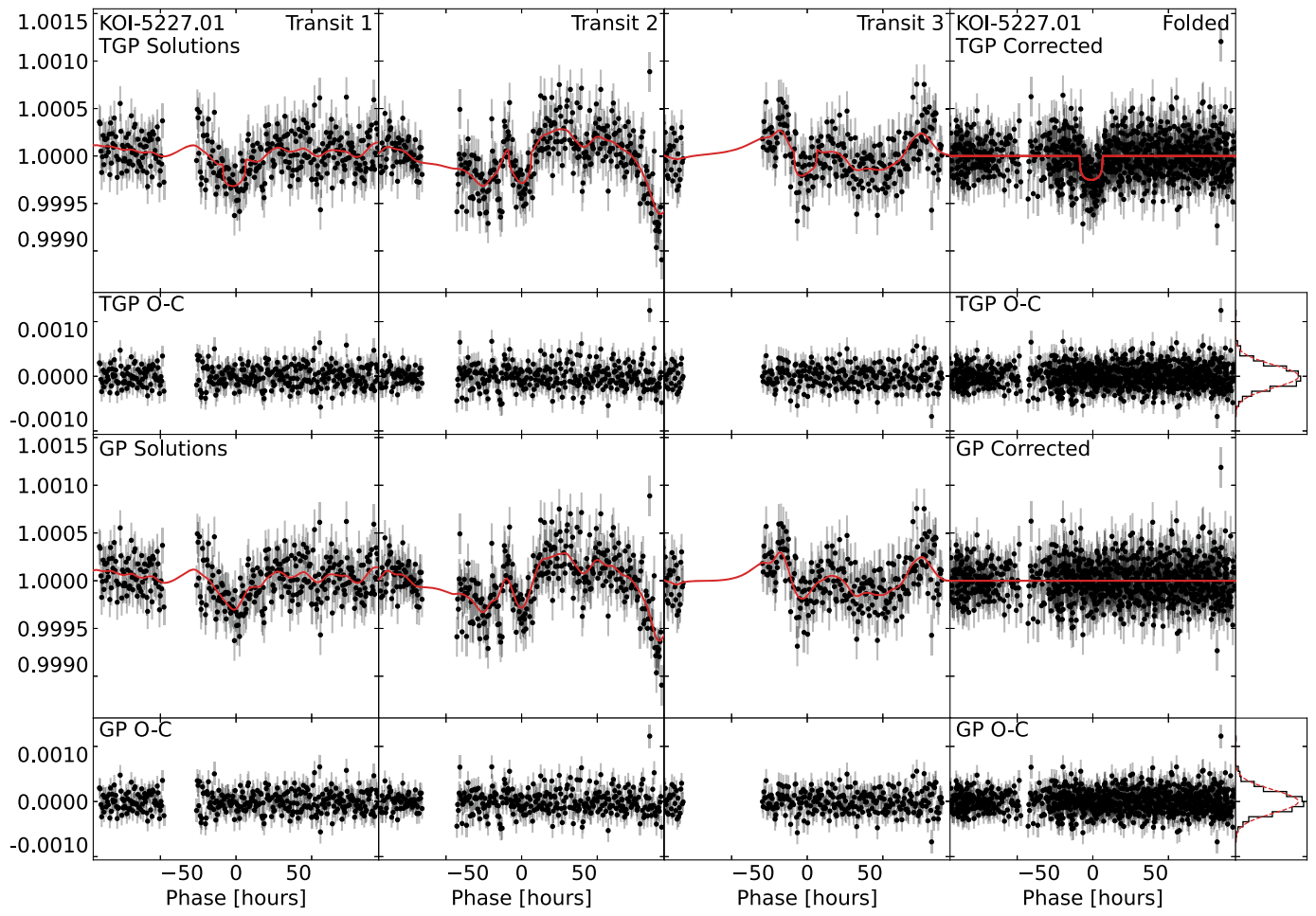


Figure 8. Same as Figure 2, but for KOI-5227.01.

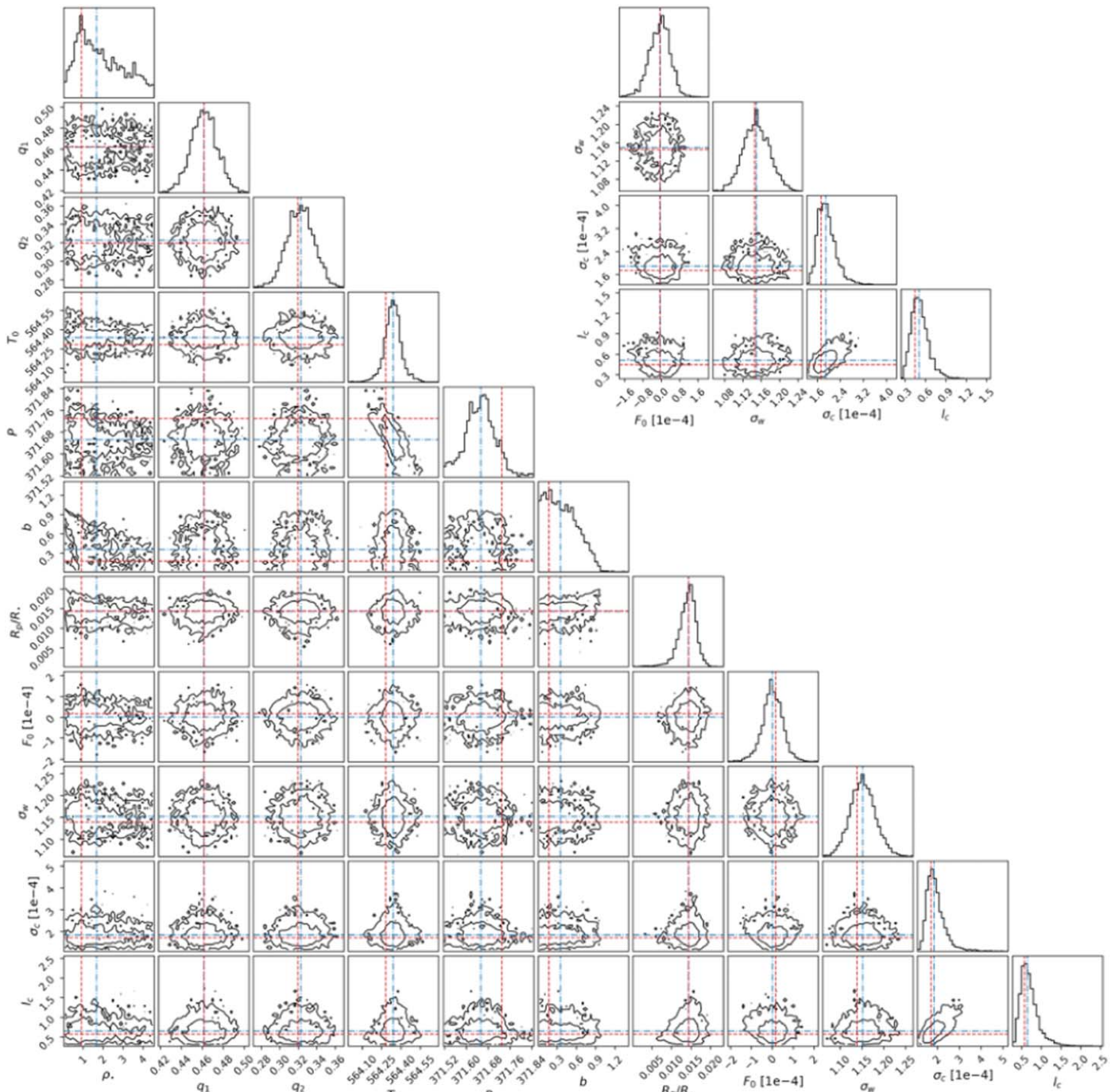


Figure 9. Same as Figure 7, but for KOI-5227.01.

ORCID iDs

Michael R. B. Matesic <https://orcid.org/0000-0002-1119-7473>
 Jason F. Rowe <https://orcid.org/0000-0002-5904-1865>
 John H. Livingston <https://orcid.org/0000-0002-4881-3620>
 Shishir Dholakia <https://orcid.org/0000-0001-6263-4437>
 Daniel Jontof-Hutter <https://orcid.org/0000-0002-6227-7510>
 Jack J. Lissauer <https://orcid.org/0000-0001-6513-1659>

References

Aigrain, S., & Foreman-Mackey, D. 2023, *ARAA*, 61, 329
 Aigrain, S., Hodgkin, S. T., Irwin, M. J., Lewis, J. R., & Roberts, S. J. 2015, *MNRAS*, 447, 2880

Aigrain, S., Parviainen, H., & Pope, B. J. S. 2016, *MNRAS*, 459, 2408
 Alsing, J., Charnock, T., Feeney, S., & Wandelt, B. 2019, *MNRAS*, 488, 4440
 Alsing, J., Wandelt, B., & Feeney, S. 2018, *MNRAS*, 477, 2874
 Angus, R., Morton, T., Aigrain, S., Foreman-Mackey, D., & Rajpaul, V. 2018, *MNRAS*, 474, 2094
 Banachiewicz, T. 1942, *AJ*, 50, 38
 Barclay, T., Endl, M., Huber, D., et al. 2015, *ApJ*, 800, 46
 Bayes, T., & Price, R. 1763, *RSPT*, 53, 370
 Berger, T. A., Huber, D., van Saders, J. L., et al. 2020, *AJ*, 159, 280
 Borucki, W. J. 2016, *RPPh*, 79, 036901
 Borucki, W. J., Agol, E., Fressin, F., et al. 2013, *Sci*, 340, 587
 Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Sci*, 327, 977
 Brahm, R., Ulmer-Moll, S., Hobson, M. J., et al. 2023, *AJ*, 165, 227
 Bryson, S., Coughlin, J., Batalha, N. M., et al. 2020, *AJ*, 159, 279
 Buchner, J. 2016, *S&C*, 26, 383
 Buchner, J. 2019, *PASP*, 131, 108005

- Buchner, J. 2021, *JOSS*, **6**, 3001
- Caceres, G. A., Feigelson, E. D., Jogesh Babu, G., et al. 2019, *AJ*, **158**, 58
- Caldwell, D. A., Kolodziejczak, J. J., Van Cleve, J. E., et al. 2010, *ApJL*, **713**, L92
- Cheney, E. W. 1966, *Introduction to Approximation Theory* (New York: McGraw-Hill)
- Cholesky, A.-L. 1924, *BGeod*, **2**, 67
- Christiansen, J. L. 2017, in *Planet Detection Metrics: Pixel-Level Transit Injection Tests of Pipeline Detection Efficiency for Data Release 25*, Kepler Science Document KSCI-19110-001, ed. M. R. Haas & N. M. Batalha (Moffett Field, CA: NASA), 18
- Claret, A., & Bloemen, S. 2011, *A&A*, **529**, A75
- Clarke, B., Kolodziejczak, J. J., & Caldwell, D. A. 2014, AAS Meeting Abstracts, **224**, 120.07
- Cranmer, K., Brehmer, J., & Louppe, G. 2020, *PNAS*, **117**, 30055
- Czekala, I., Andrews, S. M., Mandel, K. S., Hogg, D. W., & Green, G. M. 2015, *ApJ*, **812**, L28
- Davis, P. J. 1975, *Interpolation and Approximation* (New York: Dover)
- Dubois, P. F., Hinsen, K., & Hugunin, J. 1996, *ComPh*, **10**, 262
- Dunstan, D. J., Crowne, J., & Drew, A. J. 2020, arXiv:2007.09702
- Dunstan, D. J., Crowne, J., & Drew, A. J. 2022, *NatSR*, **12**, 993
- Evans, T. M., Aigrain, S., Gibson, N., et al. 2015, *MNRAS*, **451**, 680
- Foreman-Mackey, D. 2016, *JOSS*, **1**, 24
- Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, *AJ*, **154**, 220
- Foreman-Mackey, D., Montet, B. T., Hogg, D. W., et al. 2015, *ApJ*, **806**, 215
- Gibson, N. P., Aigrain, S., Roberts, S., et al. 2012, *MNRAS*, **419**, 2683
- Gilliland, R. L., Chaplin, W. J., Dunham, E. W., et al. 2011, *ApJS*, **197**, 6
- Gilliland, R. L., Chaplin, W. J., Jenkins, J. M., Ramsey, L. W., & Smith, J. C. 2015, *AJ*, **150**, 133
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Natur*, **585**, 357
- Hastings, W. K. 1970, *Biometrika*, **57**, 97
- Hinton, S. R. 2016, *JOSS*, **1**, 00045
- Hsu, D. C., Ford, E. B., Ragozzine, D., & Ashby, K. 2019, *AJ*, **158**, 109
- Hunter, J. D. 2007, *CSE*, **9**, 90
- Jeffrey, N., & Wandelt, B. D. 2023, arXiv:2305.11241
- Jeffreys, H. 1939, *Theory of Probability* (Oxford: Oxford Univ. Press)
- Jenkins, J. M. 2002, *ApJ*, **575**, 493
- Jenkins, J. M. 2020, in *Kepler Data Processing Handbook: Overview of the Science Operations Center*, ed. J. M. Jenkins (Moffett Field, CA: NASA), 2
- Jenkins, J. M., Caldwell, D. A., Chandrasekaran, H., et al. 2010a, *ApJL*, **713**, L87
- Jenkins, J. M., Chandrasekaran, H., McCauliff, S. D., et al. 2010b, *Proc. SPIE*, **7740**, 77400D
- Jenkins, J. M., Twicken, J. D., Batalha, N. M., et al. 2015, *AJ*, **150**, 56
- Kass, R. E., & Raftery, A. E. 1995, *JASA*, **90**, 773
- Kawahara, H., & Masuda, K. 2019, *AJ*, **157**, 218
- Kepler, J. 1619, *Ioannis Kepleri harmonices mundi libri V (Lincii Austriae: sumptibus Godofredi Tampachii ..., excudebat Ioannes Plancus)*
- Kipping, D. M. 2013, *MNRAS*, **435**, 2152
- Kipping, D. M. 2014, *MNRAS*, **440**, 2164
- Koch, D. G., Borucki, W. J., Basri, G., et al. 2010, *ApJL*, **713**, L79
- Lam, S. K., Pitrou, A., & Seibert, S. 2015, in *Proc. Second Workshop LVM Compiler Infrastructure in HPC*, Vol. 7 ed. H. Finkel (New York: ACM), 1
- Laplace, P. S. 1774, *Mémoires de Mathématique et de Physique*, **6**, 621
- Legin, R., Adam, A., Hezaveh, Y., & Perreault-Levasseur, L. 2023a, *ApJL*, **949**, L41
- Legin, R., Hezaveh, Y., Perreault-Levasseur, L., & Wandelt, B. 2023b, *ApJ*, **943**, 4
- Lissauer, J. J., Rowe, J. F., Jontof-Hutter, D., et al. 2023, *Updated Catalog of Kepler Planet Candidates: Focus on Accuracy and Orbital Periods*, arXiv:2311.00238
- Littlefair, S. P., Burningham, B., & Helling, C. 2017, *MNRAS*, **466**, 4250
- Livingston, J. H., Crossfield, I. J. M., Werner, M. W., et al. 2019, *AJ*, **157**, 102
- Llorente, F., Martino, L., Curbelo, E., Lopez-Santiago, J., & Delgado, D. 2022, arXiv:2206.05210
- Mackay, D. J. C. 2003, *Information Theory, Inference and Learning Algorithms* (Cambridge: Cambridge Univ. Press)
- Mandel, K., & Agol, E. 2002, *ApJL*, **580**, L171
- Matérn, B. 1960, *Spatial Variation (Lecture Notes in Statistics)* Vol. 36(New York: Springer)
- Mathur, S., Huber, D., Batalha, N. M., et al. 2017, *ApJS*, **229**, 30
- McEwen, J. D., Wallis, C. G. R., Price, M. A., & Docherty, M. M. 2021, arXiv:2111.12720
- McKinney, W. 2010, in *Proc. 9th Python in Science Conference*, ed. S. van der Walt & J. Millman (Austin, TX: SciPy), 56
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, *JChPh*, **21**, 1087
- Miller, B., Cole, A., Weniger, C., et al. 2022, *JOSS*, **7**, 4205
- Miller, B. K., Cole, A., Louppe, G., & Weniger, C. 2020, arXiv:2011.13951
- Mullally, F., Coughlin, J. L., Thompson, S. E., et al. 2015, *ApJS*, **217**, 31
- NASA 2022, *NASA Exoplanet Archive: A Service of NASA*, NASA Exoplanet Science Institute, <https://exoplanetarchive.ipac.caltech.edu>
- Olivant, T. E. 2007, *CSE*, **9**, 10
- Powell, M. J. D. 1981, *Approximation Theory and Methods* (Cambridge: Cambridge Univ. Press)
- Press, W. H., Flannery, B. P., & Teukolsky, S. A. 1986, *Numerical Recipes. The Art of Scientific Computing* (Cambridge: Cambridge Univ. Press)
- Rasmussen, C. E., & Williams, C. K. I. 2006, *Gaussian Processes for Machine Learning* (Cambridge: MIT Press)
- Roberts, S., Osborne, M., Ebdem, M., et al. 2012, *RSPTA*, **371**, 20110550
- Rowe, J. 2016, *Kepler: Kepler Transit Model Codebase Release, v1.0*, Zenodo, doi:10.5281/zenodo.60297
- Rowe, J. F., Coughlin, J. L., Antoci, V., et al. 2015, *ApJS*, **217**, 16
- Schwarzenberg-Czerny, A. 1995, *A&AS*, **110**, 405
- Seager, S., & Mallén-Ornelas, G. 2003, *ApJ*, **585**, 1038
- Skilling, J. 2004, in *AIP Conf. Ser. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, ed. R. Fischer, R. Preuss, & U. V. Toussaint (Melville, NY: AIP), 395
- Skilling, J. 2006, *BayAn*, **1**, 833
- Smith, J. C., Stumpe, M. C., Van Cleve, J. E., et al. 2012, *PASP*, **124**, 1000
- Stein, M. L. 1999, *Interpolation of Spatial Data: Some Theory for Kriging* (New York: Springer)
- Stumpe, M. C., Smith, J. C., Van Cleve, J. E., et al. 2012, *PASP*, **124**, 985
- Tejero-Cantero, A., Boelts, J., Deistler, M., et al. 2020, *JOSS*, **5**, 2505
- The Pandas Development Team 2020, *pandas-dev/pandas: Pandas, v2.1.0*, Zenodo, doi:10.5281/zenodo.3509134
- Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. 2018a, *ApJS*, **235**, 38
- Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. 2018b, *VizieR On-line Data Catalog, J/ApJS/235/38*
- Torres, G., Kipping, D. M., Fressin, F., et al. 2015, *ApJ*, **800**, 99
- Twicken, J. D., Chandrasekaran, H., Jenkins, J. M., et al. 2010, *Proc. SPIE*, **7740**, 77401U
- Twicken, J. D., Jenkins, J. M., Seader, S. E., et al. 2016, *AJ*, **152**, 158
- Van Cleve, J. E., & Caldwell, D. A. 2016, in *Kepler Instrument Handbook*, ed. M. R. Haas & S. B. Howell (Moffett Field, CA: NASA), 1
- Van Cleve, J. E., Christiansen, J. L., Jenkins, J. M., et al. 2016, in *Kepler Data Characteristics Handbook*, ed. D. Caldwell et al. (Moffett Field, CA: NASA), 2
- Van Eylen, V., Albrecht, S., Huang, X., et al. 2019, *AJ*, **157**, 61
- Van Rossum, G. 1995a, *Extending and embedding the Python interpreter*, Centrum Wiskunde & Informatica, <https://ir.cwi.nl/pub/5006>
- Van Rossum, G. 1995b, *Python tutorial*, Centrum Wiskunde & Informatica, <https://ir.cwi.nl/pub/5007>
- Van Rossum, G. 1995c, *Python reference manual*, Centrum Wiskunde & Informatica, <https://ir.cwi.nl/pub/5008>
- Van Rossum, G. 1995d, *Python library reference*, Centrum Wiskunde & Informatica, <https://ir.cwi.nl/pub/5009>
- Virtanen, P., Gommers, R., Oliviant, T. E., et al. 2020, *NatMe*, **17**, 261
- Wallis, K. F. 2014, arXiv:1405.4995