Draft version January 7, 2026
Typeset using LaTeX twocolumn style in AASTeX7.0.1

# Cross Validation with eurydice Yields Reliable Planet Mass Constraints in the HD 63433 System

Kaylee Barrera [1,2] Sarah Blunt [3] Andrew Vanderburg [4,2] Andrew W. Mann [5]
Michael C. Cushing [6] and Joe Llama [7]

[1] Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[2] Department of Physics and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[3] Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064, USA
[4] Center for Astrophysics — Harvard and Smithsonian, Cambridge, MA 02138, USA
[5] Department of Physics and Astronomy, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
[6] Ritter Astrophysical Research Center, Department of Physics and Astronomy, University of Toledo, Toledo, OH 43606, USA
[7] Lowell Observatory, Flagstaff, AZ 86001, USA

## ABSTRACT

Understanding young exoplanets is critical for constraining models of planetary evolution, yet obtaining precise masses from radial velocity (RV) measurements is challenging due to high stellar variability in young stars. Recent progress has been made using Gaussian Process (GP) models to mitigate these stellar signals, but assessing the predictive reliability of these models remains difficult, limiting their effectiveness. We introduce `eurydice`, an open-source software package designed to evaluate the predictive performance of GP models using cross-validation techniques. We apply `eurydice` to new and previously published RV datasets of the young planetary system HD 63433, demonstrating that our GP models produce planet mass estimates consistent with previously published results. By testing model reliability via cross-validation, `eurydice` supports more reliable mass measurements for active young stars. We release `eurydice` publicly to encourage its use in the community as a tool for future RV studies.

*Keywords:* exoplanets, radial velocity, stellar activity

## 1. INTRODUCTION

Planet properties are thought to be largely shaped during the earliest stages of their evolution. Core accretion models predict that planets that form with extended volatile envelopes may be eroded by mass loss processes over ∼100 Myr, eventually contracting into the older population of planets observed today (see e.g., J. L. Bean et al. 2021, H. E. Schlichting 2018, and references within). However, many details of this evolutionary process remain unclear, including the dominant mechanisms driving envelope loss. Mechanisms like photoevaporation (S. Ginzburg et al. 2018), core-powered mass loss (J. E. Owen & Y. Wu 2013), and boil-off (J. E. Owen & Y. Wu 2016) may all play a role in shaping the mature exoplanet population, though they likely operate on different timescales. Moreover, planetary systems experience significant dynamical evolution early in their history, including orbital migration (e.g., T. J. David et al. 2016; A. W. Mann et al. 2016; M. G. Barber et al. 2024; D. R. Alves et al. 2025) and planet–planet scattering or collision events (e.g., R. Frelikh et al. 2019), yet the relative importance of these processes in shaping the final characteristics of mature exoplanet populations are not well understood.

A promising way forward is to observe young planets that are still experiencing the processes that sculpt mature exoplanet populations. Several surveys are pursuing this approach, including the TESS Hunt for Young and Maturing Exoplanets Survey (THYME; E. R. Newton et al. 2019), Cluster Difference Imaging Photometric Survey (CDIPS; L. G. Bouma et al. 2019), and PSF-based Approach to TESS High Quality Data Of Stellar Clusters (PATHOS; D. Nardiello et al. 2020). Searching for planets around stars younger than 1 Gyr and studying how their properties change as a function of stellar age can reveal the key timescales on which the dynam-

Email: kaybar@mit.edu

ical and atmospheric properties of planets evolve (J. E. Owen 2020).

Nevertheless, detecting and characterizing planets around young stars remains observationally challenging, particularly with the radial velocity (RV) method. The RV signals of young stars are dominated by spot-induced variability, which can be up to hundreds of times larger than the variability of Gyr-old stars (e.g., B. L. Cale et al. 2021, A. Suárez Mascareño et al. 2021). Historically, these large activity signals have hindered RV searches for young planets, as even the RV signal of a giant, close-in planet can be masked by the stellar noise. As a result, young and/or active stars were often avoided in RV surveys (see C. K. Harada et al. 2024 for empirical evidence of this). Furthermore, the limited sample of young RV planets remains controversial, with some cases challenging the reported planetary signals (e.g., V830 Tau b in M. Damasso et al. 2020; CI Tau b in J.-F. Donati et al. 2024).

Despite these observational challenges, the RV community is beginning to explicitly target young stars in RV campaigns as statistical and observational techniques for managing stellar activity improve (e.g., Q. H. Tran et al. 2021). For example, advances in transit search algorithms in the presence of stellar noise have marked a major improvement (e.g., M. G. Barber et al. 2024). Rather than blind searches, RV studies can follow up young planets that have been independently validated and inform orbit models based on planet properties measured by transits.

Gaussian process (GP) regression is becoming a widespread technique for modeling Keplerian RV signals in the presence of stellar activity signals (R. D. Haywood et al. 2014, S. Aigrain & D. Foreman-Mackey 2023). By treating stellar activity as correlated noise with known or expected covariance properties, we can bypass the need for a functional model for stellar variability. GPs have been used very successfully on well-sampled and relatively well-understood datasets, such as Sun-as-a-star observations (B. Klein et al. 2024).

Nevertheless, as with any modeling approach, extending existing GP methods to a new context (in this case, to young star RV time series) requires some caution. S. Blunt et al. (2023) found evidence for severe overfitting in a GP model of the V1298 Tau system, a $\sim$20 Myr star with a $\sim$500 m/s stellar activity signal and 4 transiting planets, raising concerns about the reliability of the model's inferred planet masses. Subsequent analyses of this system have validated these concerns by deriving significantly lower planet masses than initially reported (S. Barat et al. 2024, Livingston et al. *in press*). In light of these issues, S. Blunt et al. (2023) emphasized the need for developing tools to assess the predictiveness of all RV models, including but not limited to GP models, a gap we aim to address here.

HD 63433 is a young (414 ± 23 Myr; J. Jones et al. 2015) nearby ($\sim$22 pc) planetary system hosting three planets orbiting a Sun-like star (spectral type G5V) in the Ursa Major moving group (B. K. Capistrant et al. 2024). Planets HD 63433 b and c, first discovered by A. W. Mann et al. (2020), are estimated to be sub-Neptunes, both with radii between 2 and $3R_\oplus$. The innermost planet HD 63433 d was recently reported by B. K. Capistrant et al. (2024) with a size of $1.073^{+0.046}_{-0.044}R_\oplus$ on a short, 4.2-day orbit. As such, HD 63433 d is the closest Earth-size planet orbiting a young star, offering a unique opportunity to understand a young Earth analog.

Given its brightness (V$\simeq$6.9 mag), HD 63433 is an attractive target for ground-based RV campaigns aimed at measuring planetary masses. Independent RV analyses by M. Damasso et al. (2023) and M. Mallorquín et al. (2023) have provided mass estimates for the outer planets HD 63433 c and d using GP frameworks to model stellar activity, but these estimates have not been assessed for reliability using held-out datasets. Given the high-amplitude activity signal, it is important to assess the predictive ability of the GP models, and therefore how reliable the inferred masses may be.

In this work, we revisit the analyses of M. Damasso et al. (2023) and M. Mallorquín et al. (2023) and add additional data to place further constraints on the masses of HD 63433 b and c, present the first mass estimates for planet d, and test the reliability of the results.

To support this analysis and larger efforts to evaluate GP models in RV studies, we introduce `eurydice`, a Python package developed to evaluate GP models using cross-validation techniques. As the usage of GP models continues to grow in the exoplanet community, so too does the need for standardized tools to assess their predictive performance across different kernel choices and modeling assumptions. `eurydice` provides a consistent framework for evaluating fitted models on held-out data, helping to improve the reliability of planetary signal detections. We use HD 63433 as the first test case for `eurydice`, demonstrating its utility in validating GP models for RV-based mass estimation.

In Section 2, we present the radial velocity data used in our analysis of HD 63433, using both previously reported data from M. Damasso et al. (2023) and M. Mallorquín et al. (2023) along with new EXPRES observations. In Section 3, we detail our GP and model-fitting framework. We introduce the `eurydice` package and suite of cross-validation tests in Section 4 and discuss modeling choices, the impacts on inferred parameters, and future work with `eurydice` in Section 5. We end with our conclusions in Section 6.

## 2. DATA

In this work, we use three RV time series derived from spectra taken by three different spectrographs. All RV data are shown in Figure 1 and each dataset is detailed below.

### 2.1. *HARPS-N*

We used the published dataset from M. Damasso et al. (2023) of HD 63433 observations from the HARPS-N spectrograph (R. Cosentino et al. 2012, 2014) installed at the Telescopio Nazionale Galileo (TNG) at the Observatory of the Roque de los Muchachos in La Palma, Spain. HARPS-N is a fiber-fed, cross-dispersed echelle spectrograph with an average spectral resolution of $\mathcal{R} =$ 115,000 and covers a wavelength range of 383-690 nm. A total of 103 spectra were collected between 26 February 2020 and 17 April 2022, and the derived RV time series was calculated from the `DRS` pipeline (version 3.7.1).

### 2.2. *CARMENES*

We used the published dataset from M. Mallorquín et al. (2023) of HD 63433 observations from the high-resolution CARMENES spectrograph (A. Quirrenbach et al. 2014, 2016) located on the 3.5 m telescope at the Calar Alto Observatory in Almería, Spain. A total of 157 spectra were collected between 19 September 2020 and 23 February 2022 using both the visual and near-infrared channels of CARMENES. The spectra were reduced using `CARACAL` (J. A. Caballero et al. 2016) and the RVs were derived using `SERVAL` (M. Zechmeister et al. 2018). The near-infrared RVs exhibited a mean uncertainty of 14.3 m/s, significantly higher than the 3.9 m/s mean uncertainty measured in the visual channel. Consequently, we excluded the near-infrared arm RVs from this analysis and only utilized the 150 RV points from the visual arm. The visual light arm has a measured spectral resolution of $\mathcal{R}$ = 93,400 and spans the wavelength range of 520-960 nm.

### 2.3. *EXPRES*

Between 16 November 2021 and 30 October 2022, we obtained 52 spectra using the EXtreme PREcision Spectrograph (EXPRES; C. Jurgenson et al. 2016) on the 4.3-m Lowell Discovery Telescope (LDT; S. E. Levine et al. 2012) located in Flagstaff, Arizona. EXPRES is a high-resolution (R$\sim$ 137,500) optical (380 - 822 nm) stabilized spectrograph designed for extreme precision RV. The spectrograph is equipped with both a ThAr lamp and Laser Frequency Comb (LFC) for precise wavelength calibration. Exposures of both calibrator sources were obtained every 30 minutes during our observations. The data were reduced using the standard EXPRES pipeline as described in R. T. Blackman et al. (2020) and telluric removal was applied using SELENITE (C. Leet et al. 2019). Wavelength calibration was performed using the *excalibur* method (L. L. Zhao et al. 2021) and the RVs were computed using the pipeline described in R. R. Petersburg et al. (2020). All EXPRES RVs used in this work are provided in Table A1.

## 3. RADIAL VELOCITY ANALYSIS

In this analysis, we have opted to exclude the EXPRES data from our modeling due to its smaller sample size (52 observations) compared to the HARPS-N and CARMENES datasets (103 and 150 observations, respectively). The EXPRES data was instead held as a test set for later cross-validation tests.

### 3.1. *Rotation Period Analysis*

Although the stellar rotation period of host star HD 63433 has been well characterized in previous studies (see A. W. Mann et al. 2020 and B. K. Capistrant et al. 2024), we performed an independent check to assess the power of the rotation signal in our combined RV dataset. We computed a Lomb-Scargle periodogram using the `astropy` LombScargle class (J. T. VanderPlas & Ž. Ivezić 2015; J. VanderPlas et al. 2012), shown in Figure 2. While the fundamental rotation period at 6.4 days was only weakly recovered, its first and second harmonics at 3.2 and 2.1 days appear more prevalent. This behavior is consistent with results from A. Vanderburg et al. (2016), who found that RV signals from stellar activity often manifest most strongly at the rotation period and its first two harmonics. Interestingly, CARMENES and HARPS-N detect different harmonics, with the former recovering the 2.1 day second harmonic and the latter recovering the 3.2 day first harmonic. This discrepancy may be due to different observation cadences that are sensitive to different parts of the stellar rotation signal. Based on this analysis, we conclude that the true rotation period is 6.4 days, consistent with previous studies (A. W. Mann et al. 2020; B. K. Capistrant et al. 2024) and use this as a prior in the GP models described below.

### 3.2. *Gaussian Process Modeling Choices*

GPs are flexible, non-parametric models that use a covariance function, or kernel, to describe the correlation between pairs of observations. For further discussion on GPs, a broad review is provided in C. E. Rasmussen

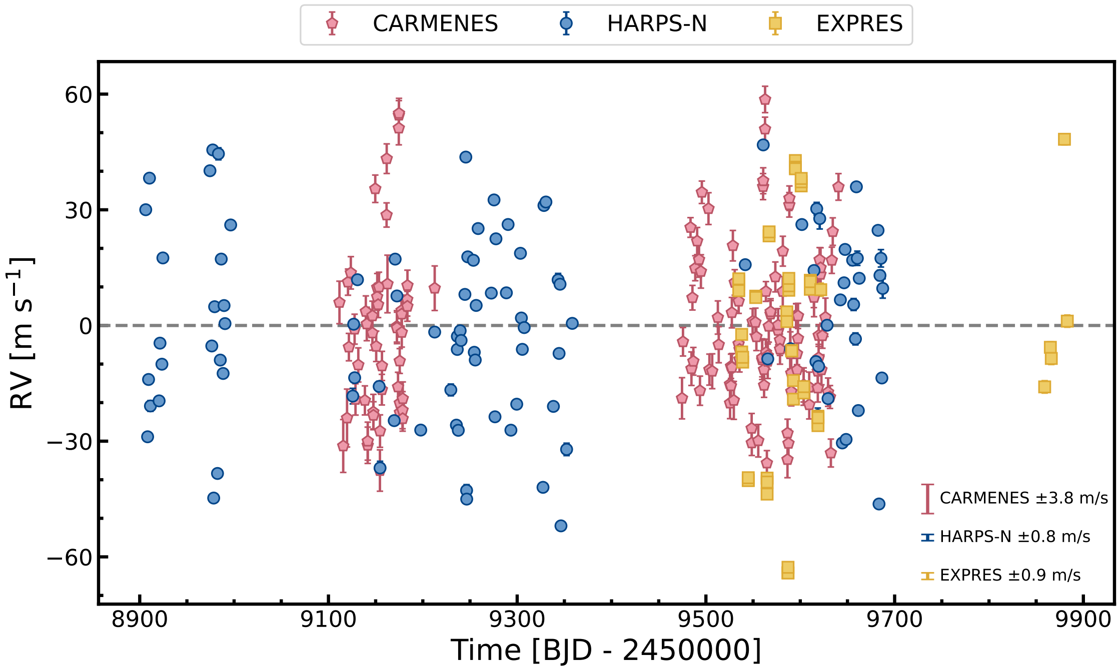


**Figure 1.** A survey of the RVs used in this analysis. The CARMENES and HARPS-N RVs are published in M. Mallorquín et al. (2023) and M. Damasso et al. (2023), respectively, while the EXPRES RVs are newly introduced in this work. Across all datasets, the RVs show a large spread despite low observational uncertainties (median error bars for each dataset are shown in the bottom right), highlighting the strong stellar variability of HD 63433 present in this data.

& C. K. I. Williams (2006) and applications of GPs to astronomical timeseries are detailed in S. Aigrain & D. Foreman-Mackey (2023).

As HD 63433 is a Sun-like star (A. W. Mann et al. 2020), we expect its p-mode oscillation, granulation, and supergranulation activity amplitudes to be similar to the Sun's, i.e., below ∼1 m/s (see N. K. O'Sullivan et al. 2025 and references therein). Therefore, the stellar activity-induced RVs are expected to be dominated by rotation or the evolution of active regions. In this regime, a quasi-periodic kernel is commonly used in GP models of RVs (S. Aigrain et al. 2012; R. D. Haywood et al. 2014):

$$k(t,t') = \eta_1^2 \exp(-\frac{(t-t')^2}{2\eta_2^2} - \frac{2\sin^2(\frac{\pi(t-t')}{\eta_3})}{\eta_4^2}) \tag{1}$$

where $t$ and $t'$ are the times of the two observations, $\eta_1$ is the GP amplitude, $\eta_2$ is the evolution timescale of active regions, $\eta_3$ is the stellar rotation period, and $\eta_4$ is the harmonic complexity, which dictates how the signal varies within a rotation period (B. A. Nicholson & S. Aigrain 2022).

When combining RV measurements from different spectrographs, an unresolved question is whether or not to treat datasets from different instruments as correlated measurements (see S. Blunt et al. 2023 section 4.1, as well as Figures 11 and 12). Since stellar activity arises from physical processes, we can predict that multiple instruments taking data at the same time will observe the same stellar signal, with potentially distinct amplitudes due to the chromaticity of stellar activity (S. Blunt et al. 2023; B. L. Cale et al. 2021). This assumption motivated our two approaches in GP modeling: a joint covariance matrix framework with a shared kernel functional form and hyperparameters that assume a correlated signal across instruments, and an added log-likelihood approach, commonly used in literature, that models the datasets with separate independent GPs but shares the same hyperparameters.

The joint covariance matrix framework uses a multi-instrument quasi-periodic Gaussian Process kernel in which the GP hyperparameters are shared across the model, while allowing for separate amplitude hyperparameters ($\eta_1$) for each instrument $i$:

$$k(t,t') = \eta_{1,i}\eta_{1,j} \exp(-\frac{(t-t')^2}{2\eta_2^2} - \frac{2\sin^2(\frac{\pi(t-t')}{\eta_3})}{\eta_4^2}). \tag{2}$$

Here, $\eta_{1,i}$ is the GP amplitude associated with the instrument that took the observation at time $t$, and $\eta_{1,j}$ is the amplitude associated with the instrument that took the observation at time $t'$. All other hyperparameters retain the same definitions as in Equation 1.

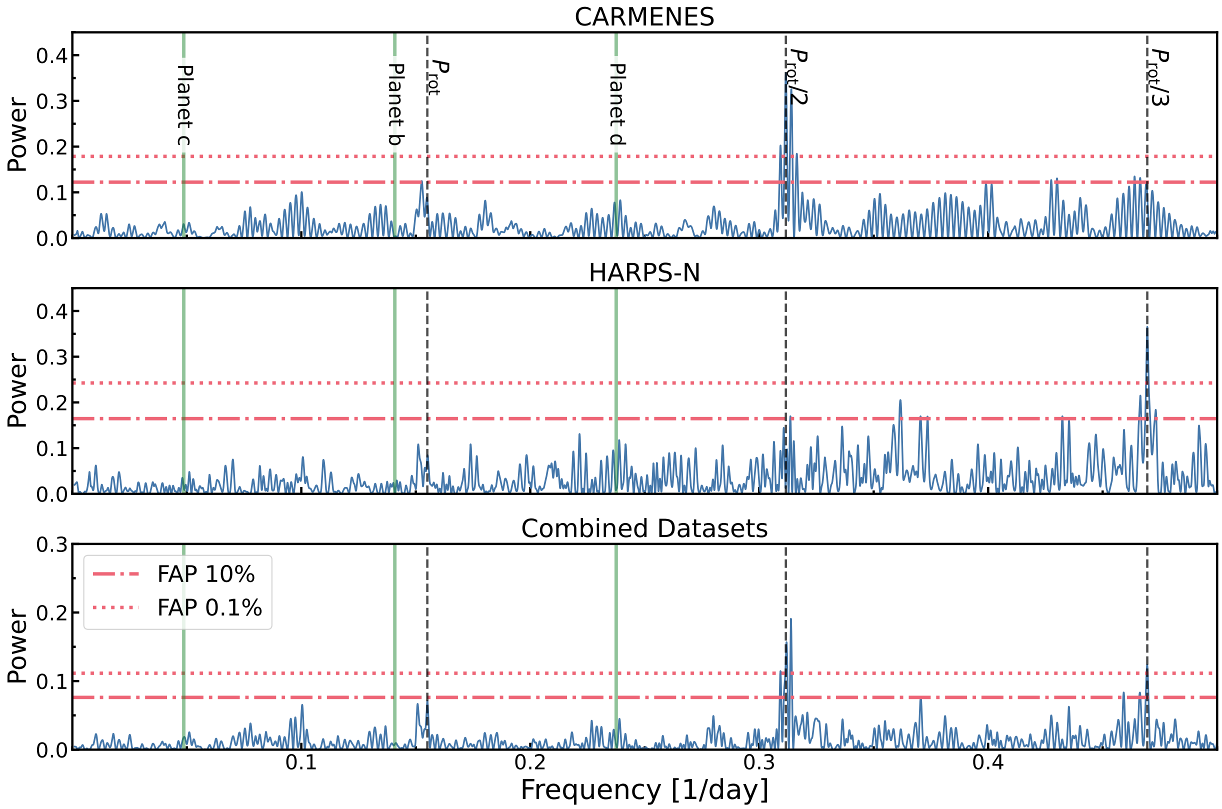


**Figure 2.** Lomb-Scargle periodograms of the CARMENES and HARPS-N RVs. **Top**: CARMENES RVs; **middle**: HARPS-N RVs; **bottom**: combined RVs. Black dashed lines indicate HD 63433's rotation period (6.4 days) and its harmonics, while green lines mark the orbital periods of the three planets. Orange dotted and dash-dotted lines correspond to false-alarm probabilities (FAPs) of 0.1% and 10%, respectively. None of the periodograms show a strong signal at the 6.4 day rotation period, but both the first and second harmonics (3.2 and 2.1 days) are detected.

The added log-likelihood approach applies a separate GP to each dataset, and the model is optimized using the sum of the individual log-likelihoods. While this framework still assumes a stellar noise model with the same GP hyperparameters across the instruments, it assumes measurements taken by different instruments are independent. Following the discussion in S. Blunt et al. (2023), modeling the different datasets as independent effectively introduces more free parameters to this model, which may increase the risk of overfitting. As this aspect is not fully understood in the literature, we sought to explore it in our analysis.

### 3.3. *Model Fitting*

To model both stellar activity and planetary signals, we defined the GP mean function as a sum of Keplerian orbits using `jaxoplanet` (S. Hattori et al. 2024) and constructed the GP models with `tinygp` (D. Foreman-Mackey et al. 2024). For transparency, our GP fitting pipeline code is publicly available on GitHub.[8]

[8] https://github.com/kayleebarrera/HD63433

We explored several model configurations, including two-planet and three-planet fits. The two-planet model was first implemented to reproduce previous analyses (M. Damasso et al. 2023; M. Mallorquín et al. 2023), which were carried out before the discovery of HD 63433 d, and to test whether our joint modeling of CARMENES and HARPS-N data would result in consistent planetary parameters. Once this baseline was established, we expanded to a three-planet model. The expected RV semi-amplitude of HD 63433 d is small ($<$ 1 m/s for an Earth-sized planet of a comparable rocky composition at a 4.2-day orbit). Thus, its signal is below the median RV uncertainties of our data, but we included it in the three-planet model to evaluate its effect on the derived parameters of the two previously known planets.

In addition to testing two- and three-planet fits and different joint modeling strategies for the combined datasets, we tested both circular and eccentric orbital models, resulting in a total of eight GP model configurations. For comparison, we also constructed a three-planet circular model assuming only white noise without a GP component. A summary of all models is presented in Table 1.

As this work did not utilize photometric data, we adopted narrow Gaussian priors on the orbital periods and transit times of the three planets, based on published values from B. K. Capistrant et al. (2024). As M. Mallorquín et al. (2023) found no evidence of transit-timing variations, we consider these priors reliable for our analysis. Uniform priors were used for all remaining parameters and summarized in Table A2. For the eccentric orbital models, we used the $\sqrt{e}\cos(\omega)$ and $\sqrt{e}\sin(\omega)$ parameterization of eccentricity and argument of periastron.

Parameter distributions were sampled using the affine invariant Markov chain Monte Carlo (MCMC) sampler implemented in `emcee` (D. Foreman-Mackey et al. 2019). For each model, the number of walkers was set to twice the number of free parameters. Circular models were run for $10^6$ steps, while eccentric models were run for $2 \times 10^6$ to ensure adequate sampling of the larger parameter space. The walkers were initialized at parameter values near those reported by B. K. Capistrant et al. (2024), M. Damasso et al. (2023), and M. Mallorquín et al. (2023). Burn-in was discarded based on the maximum autocorrelation time ($\tau$) across all parameters, and chains were thinned according to the minimum $\tau$. Convergence was assessed by visually inspecting the walker trace plots and checking that the chain lengths exceeded $50\tau$ for all parameters, as recommended by `emcee`.

## 4. EURYDICE: A SOFTWARE PACKAGE FOR CROSS-VALIDATION

As investigated in S. Blunt et al. (2023) and M. MacLeod et al. (2025), GP models applied to RV datasets, like all models, are susceptible to overfitting. Overfitting occurs when a model reproduces noise or localized trends present in its training data rather than the true underlying signal. Cross-validation is a method for evaluating model generalization by testing how well a model can predict unseen data.

To evaluate our suite of GP models, we developed `eurydice`,[9] a new open-source Python package designed to perform cross-validation on GP models applied to radial velocity (RV) data. This package provides a streamlined and standardized framework for comparing different GP configurations and assessing their predictive performance to improve the reliability of modeling planetary signals.

[9] https://eurydice.readthedocs.io/en/latest/

Currently, `eurydice` supports a simple hold-out cross-validation method,[10] implemented as follows:

1. Divide the dataset into a training set and a test set (e.g., by instrument or using a predefined split).
2. Condition the GP on the training set.
3. Evaluate the GP predictions on the held-out test set.
4. Compare the scatter of the residuals in the test set to the predicted uncertainty distribution.

Predictive performance is quantified using normalized residuals, which measure the deviation between model predictions and the actual RV measurements, scaled by observational errors, GP uncertainties, and instrumental jitters:

$$\text{norm. residuals} = \frac{\mathrm{RV}_{\mathrm{obs}} - \mathrm{RV}_{\mathrm{model}}}{\sqrt{\sigma^2_{\mathrm{obs}} + \sigma^2_{\mathrm{GP}} + \sigma^2_{\mathrm{jit}}}} \tag{3}$$

We report the standard deviation of these normalized residuals as metrics for our cross-validation tests in Table 1. We distinguish between the standard deviation of observations used to condition the GP ($\sigma_{\mathrm{condition}}$) and the observations used to test the model ($\sigma_{\mathrm{test}}$). Assuming an ideal model that perfectly captures all correlated noise, the normalized residuals should follow a standard normal distribution with $\sigma_{\mathrm{test}} \approx 1$. If the residuals from the test set are much broader than the expected standard normal distribution with $\sigma_{\mathrm{test}} \gg 1$, this may indicate that the model is overfitting to the training data.

Importantly, `eurydice` itself does not perform any parameter training or optimization. Instead, it operates using the results of an external GP model parameter estimation procedure (such as posterior inference from an MCMC sampler). Users first optimize the GP model parameters on the training set using their preferred parameter-fitting tools, and `eurydice` takes these fitted parameters to generate predictions on the test set. Consequently, `eurydice` offers flexibility in working with a wide range of GP models and fitting approaches. Because it does not perform parameter optimization itself, `eurydice` is not only compatible with different inference methods, but can evaluate any GP model with arbitrary kernel choices. There are a handful of excellent tools for

[10] Although termed "cross-validation," we recognize that this is a single hold-out validation rather than full k-fold cross-validation. It assesses generalization on one held-out subset but does not provide the repeated-resampling statistics of formal cross-validation.

building and fitting GP models that can be used alongside `eurydice`, including `tinygp` (D. Foreman-Mackey et al. 2024, used in this work), `george` (S. Ambikasaran et al. 2015), `celerite` (D. Foreman-Mackey et al. 2017), and `pyaneti` (O. Barragán et al. 2022), which is specifically designed for jointly modeling transits and radial velocity data.

Often, GP models in the literature are selected using Bayesian model comparison techniques, such as the Bayesian log-evidence ($\ln \mathcal{Z}$) or the Bayesian Information Criterion (BIC). While these criteria provide a relative measure of model quality based on model fit and complexity, they do not offer an absolute assessment of how well a model predicts unseen data. In contrast, cross-validation provides a robust complementary method for model selection by directly evaluating generalization on held-out data. Although cross-validation is not yet widely adopted in the radial velocity community, we aim to promote its broader use by providing tools like `eurydice` for accessible model evaluation of GPs for RV datasets. We hope that `eurydice` can be seamlessly integrated into existing RV analysis pipelines, enabling cross-validation tests without requiring heavy modification.

### 4.1. *Cross-Validation Tests on HD 63433*

Our cross-validation tests focused on three key modeling choices: (1) circular versus eccentric orbital assumptions, (2) two-planet versus three-planet system configurations, and (3) the treatment of instrument systematics using either the joint GP covariance matrix model or an added likelihood approach across instruments.

During training, all models were fit using the combined HARPS-N and CARMENES datasets to infer global model parameters. In our cross-validation tests, however, we conditioned the model on data from only one instrument and validated predictions against the other. For example, in test CV1 (see Table 1) for model M1, the GP would be conditioned solely on HARPS-N data using the median of the globally trained parameter posteriors to predict the CARMENES observations, and vice versa for CV2. In this context, training refers to determining the overall GP model structure using MCMC, while conditioning uses the trained model to generate predictions for a particular dataset. A model can generate predictions for both the data it has been trained on and the data it has not. Although both the HARPS-N and CARMENES datasets were used in training, conditioning on only one dataset isolates whether the trained GP can reproduce the signal observed by the other instrument, allowing for instrument-specific differences such as unique jitter values. This tests the model's ability to generalize across instruments without retraining.

Additionally, the EXPRES dataset, which was not included in the training set, was used for further cross-validation tests of the joint covariance matrix approach. To account for instrumental differences between the EXPRES data and the other two datasets, we fit for an RV zero-point offset and amplitude that minimized the negative log-likelihood with respect to the GP model predictions to minimize the sum of normalized residuals. We did not perform cross-validation with the EXPRES data in the added likelihood case, since there is currently no straightforward way to combine their predictive distributions to generate predictions for a third unseen instrument. To evaluate how well the GP mitigated correlated noise from stellar activity, we also performed a cross-validation test of a white-noise-only model.

In addition to these cross-validation tests, we also derived posterior distributions of planetary masses for each model configuration to illustrate how these different models impact the inferred planet properties. These posteriors are presented in Figures 3-4 and discussed in detail in the following section. All model-predicted planetary masses, including the white noise model, were consistent with each other within approximately $1\sigma$.

## 5. RESULTS AND DISCUSSION

### 5.1. *Reliability of GP Models*

Across all GP configurations, the models reproduce the radial velocity signals similarly, with no single approach strongly favored by cross-validation metrics. All inferred planetary properties remain largely unchanged, with only a minor sensitivity to model assumptions.

Adding a third planet to the model does not substantially improve the overall fit or change the other two planetary parameters. However, only one of the three planets, HD 63433 c, is confidently detected by any of our models; we are only able to place upper limits on the masses of the other two planets. Since the small third planet does not significantly affect the derived properties of the system, we adopt the three-planet model to reflect the current state of knowledge of the system.

The cross-validation tests show that circular and eccentric orbital models perform similarly, with no clear preference for either. The eccentricity posterior for HD 63433 c is roughly Gaussian and consistent with zero, reflecting a moderately well-constrained orbit. Yet, the posteriors for planets b and d indicate that their eccentricities are extremely under-constrained, with broad distributions spanning nearly the full allowed range and favoring unphysical values near e $\approx 0.9$ (Figures A1-A2). Given the closely-packed architecture of the system and

**Table 1.** Cross-validation results and derived planetary masses for each model. All models were trained on both HARPS-N AND CARMENES data. **CV1:** Model conditioned on HARPS-N data and tested on CARMENES data. **CV2:** Model conditioned on CARMENES data and tested on HARPS-N data. **CV3:** Model conditioned on both HARPS-N and CARMENES data and tested on EXPRES data. Note that CV3 is only available for the joint covariance matrix and white noise models. For all CV tests, $\sigma_{\rm condition}$ and $\sigma_{\rm test}$ describe the standard deviations of the residual distributions of the conditioning and test sets, respectively (as defined in Equation 3). A $\sigma_{\rm test}$ value closer to 1 indicates a more predictive model. All upper limits for $m_b$ and $m_d$ are reported at $3\sigma$. The adopted GP model (M3) is bolded.

| Model | Description | $\sigma_{\rm condition}$ (CV1 / CV2 / CV3) | $\sigma_{\rm test}$ (CV1 / CV2 / CV3) | Derived Planetary Masses ($M_\oplus$) |
|---|---|---|---|---|
| M1 | Joint covariance, 2-planet circular | 0.34 / 0.10 / 0.33 | 0.92 / 1.07 / 1.05 | $m_b < 22.85$<br>$m_c = 18.60^{+4.45}_{-4.26}$ |
| M2 | Joint covariance, 2-planet eccentric | 0.35 / 0.11 / 0.34 | 0.92 / 1.08 / 1.03 | $m_b < 31.96$<br>$m_c = 27.42^{+7.37}_{-6.65}$ |
| **M3** | **Joint covariance, 3-planet circular** | **0.34 / 0.11 / 0.32** | **0.92 / 1.06 / 1.04** | $\mathbf{m_b < 23.61}$<br>$\mathbf{m_c = 18.78^{+4.44}_{-4.26}}$<br>$\mathbf{m_d < 10.33}$ |
| M4 | Joint covariance, 3-planet eccentric | 0.34 / 0.12 / 0.34 | 0.94 / 1.03 / 0.99 | $m_b < 34.83$<br>$m_c = 27.45^{+7.64}_{-6.34}$<br>$m_d < 38.49$ |
| M5 | Added likelihood, 2-planet circular | 0.34 / 0.08 / - | 1.34 / 1.02 / - | $m_b < 19.39$<br>$m_c = 16.71^{+4.28}_{-4.14}$ |
| M6 | Added likelihood, 2-planet eccentric | 0.35 / 0.09 / - | 1.35 / 1.03 / - | $m_b < 34.59$<br>$m_c = 24.95^{+9.08}_{-6.59}$ |
| M7 | Added likelihood, 3-planet circular | 0.34 / 0.08 / - | 1.35 / 1.01 / - | $m_b < 20.22$<br>$m_c = 16.71^{+4.29}_{-4.15}$<br>$m_d < 10.04$ |
| M8 | Added likelihood, 3-planet eccentric | 0.34 / 0.08 / - | 1.35 / 1.02 / - | $m_b < 37.31$<br>$m_c = 24.77^{+8.83}_{-6.47}$<br>$m_d < 37.44$ |
| M9 | White-noise only, 3-planet circular | - / - / 0.98 | - / - / 30.53 | $m_b < 13.75$<br>$m_c = 18.55^{+7.85}_{-7.72}$<br>$m_d < 19.72$ |

constraints from B. K. Capistrant et al. (2024), such extreme eccentricities are likely unstable over the system lifetime. Instead, including eccentricity in the model provides additional degrees of freedom in the model, allowing the fits to absorb noise or sparse sampling by artificially inflating the semi-amplitude.

For models in which e is allowed to vary, we observe a degeneracy between e and K, resulting in slight but systematic increases in the inferred planetary masses. This degeneracy is most noticeable for planets c and d, where the posterior distributions shift toward higher masses and develop longer right-hand tails relative to the circular fits (Figures 3 and 4). As the eccentricities of the weaker signals cannot be reliably measured and allowing them to vary introduces small but consistent mass biases, we consider the circular model to provide the more reliable and physically motivated description of the system. We expect that more precise eccentricity measurements from future transits or transit timing variations could break this degeneracy and allow the planetary masses to be determined more accurately.

For the same planetary configurations, models using a joint covariance matrix are generally more predictive compared to those using an added likelihood approach. When predicting HARPS-N data conditioned on CARMENES observations, the training data standard deviation remains unchanged between the two strategies while the test data standard deviation decreases for the joint covariance models (∼1.3 for added

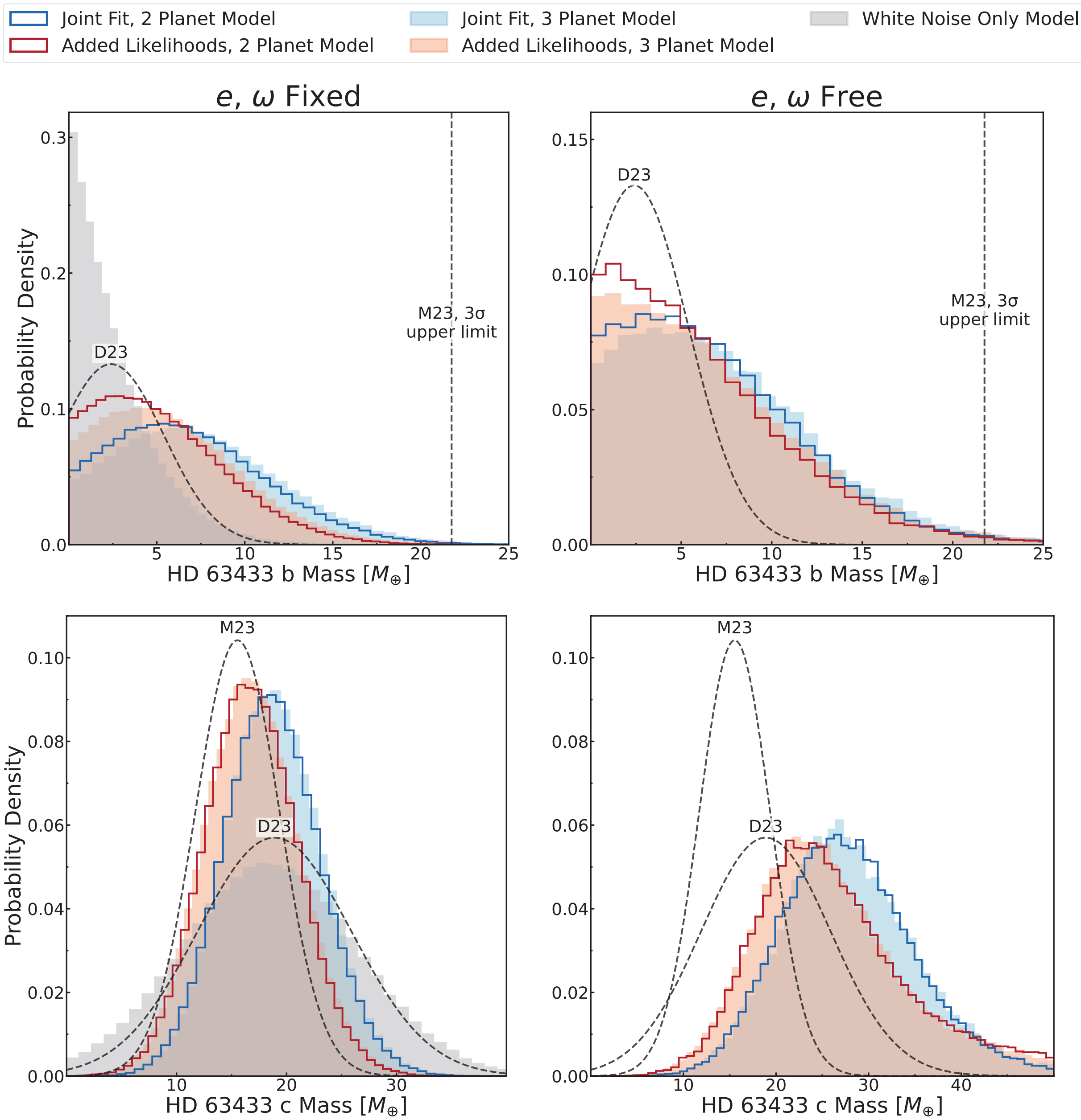


**Figure 3.** Mass posteriors for HD 63433 b **(top)** and c **(bottom)**. Circular models are shown in the left column, and eccentric models are shown in the right. The dashed curves labeled M23 and D23 show the mass estimates from M. Mallorquín et al. (2023) and M. Damasso et al. (2023), respectively, plotted as Gaussians using the reported median and $\pm 1\sigma$ uncertainties for illustration. The derived planetary masses from our models remain consistent across all tested models within their $1\sigma$ uncertainties. The circular models agree better with literature estimates from M23 and D23, as the eccentric models tend to shift the inferred masses higher. (See planetary mass comparisons in Table 1).

likelihoods versus ~0.9 for the joint covariance matrix), indicating slightly improved predictive performance.

When predicting CARMENES data from HARPS-N observations, the test data standard deviation is marginally higher for the joint covariance models. In all cases, however, conditioning on HARPS-N alone appears to under-constrain the model, with the training data standard deviation for all configurations shrinking to ~0.1 compared to ~0.3 for the CARMENES-only cases. Under this scenario, the joint covariance models show slightly better predictive accuracy in the training data standard deviations (~0.11 versus ~0.08 for the added likelihoods).

When comparing the derived mass distributions, the joint covariance approach produces marginally higher masses for planets b and c, although the differences re-

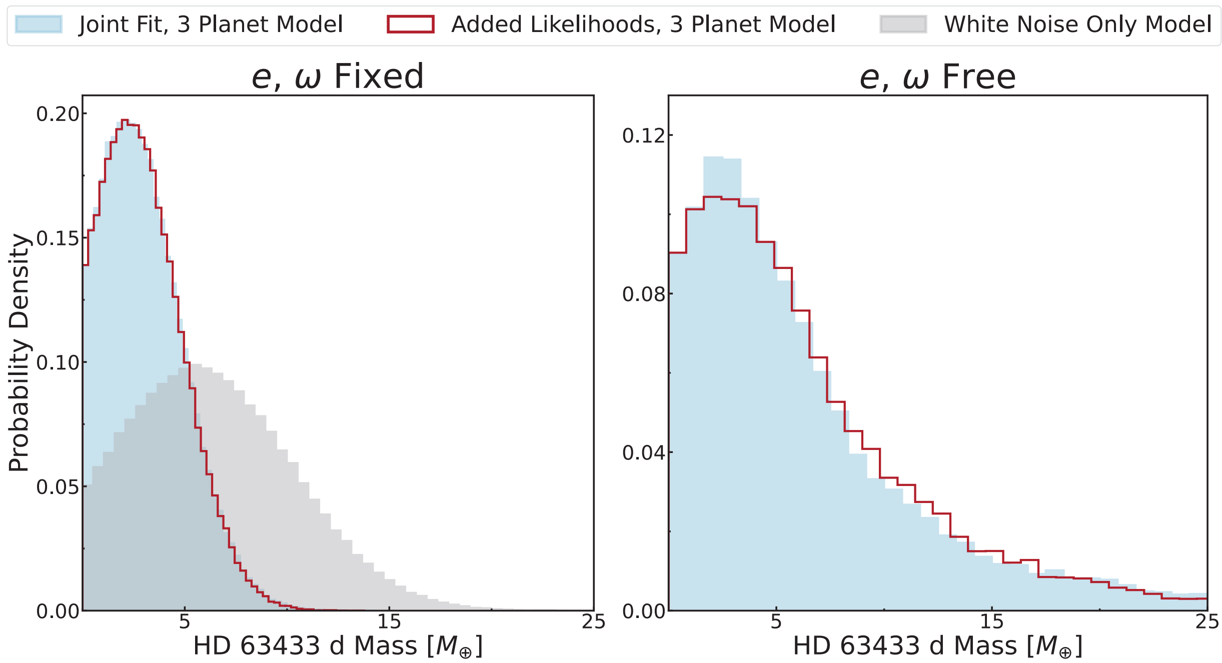


**Figure 4.** Mass posteriors for HD 63433 d. The joint covariance model and the added-likelihood model produce nearly identical distributions. Both predict lower masses than the white-noise-only model, indicating that modeling stellar noise with GPs can push the detection barrier to smaller masses. As with the other planets in Figure 3, allowing eccentricity (right) shifts the posterior toward higher masses, producing an extended right-hand tail.

main well within $1\sigma$ between the two methods. For the smaller planet d, the mass posteriors are essentially unchanged, reflecting that neither method has a clear advantage in detecting its low amplitude.

Both the joint-fit and added-likelihood modeling approaches produce reasonable results, but we ultimately adopt the three-planet circular model using the joint covariance matrix approach (M3). The model and its predictions are presented in Figure 5, and a histogram of the normalized residuals in Figure 6. The best-fit parameters for M3 are reported in Table 2. We believe that treating data points from different instruments taken around the same time as physically correlated provides a more realistic representation of the underlying noise structure, though this assumption should be further tested. Overall, the consistency of cross-validation results across all model configurations reinforces our confidence in the robustness of the inferred planetary parameters.

As an additional test, we compared the adopted GP model to the white-noise only model. The jitter values of both instruments increase significantly in the white noise case, from $3.52^{+2.46}_{-2.34}$ to $23.76^{+1.78}_{-1.59}$ m/s for HARPS-N, and from $0.73^{+0.78}_{-0.51}$ to $19.15^{+1.22}_{-1.11}$ m/s for CARMENES. As shown in Figure 7, it is clear that the jitter in the white-noise only model is dominated by stellar activity and dominates over the phase-folded Keplerian signal of HD 63433 c. However, the mass distribution for HD 63433 c derived from the white-noise model agrees with that found with a GP, albeit with broader uncertainties. Without modeling the stellar noise, the white-noise model also inflates the mass of HD 63433 d and detects a limited signal from HD 63433 b, although both are also consistent with the results of the GP model. Comparisons to a white-noise model mainly serve as a check that managing stellar activity with a GP is appropriate for properly estimating the masses of this system.

### 5.2. *Properties of HD 63433 Planets*

With our new fit to the full set of radial velocity observations of HD 63433, we are now able to update our knowledge of the planets' properties. In general, we find results that are consistent with those from previous studies by M. Damasso et al. (2023) and M. Mallorquín et al. (2023).

We detect the radial velocity signal of the outermost planet HD 63433 c with a mass of $18.78^{+4.44}_{-4.26} M_\oplus$ at approximately $4.2\sigma$ confidence. Combined with the known radius of $2.521 \pm 0.1 R_\oplus$ (B. K. Capistrant et al. 2024), this yields a density of $\rho_c = 6.43^{+1.79}_{-1.58}$ g/cm$^3$, considerably higher than Neptune ($\rho \approx 1.64$ g/cm$^3$; R. Helled et al. 2020). Observations of atmospheric escape from HD 63433 c revealed excess Ly $\alpha$ absorption, suggesting that some hydrogen may still be escaping, although additional searches in H$\alpha$ and He I show much weaker

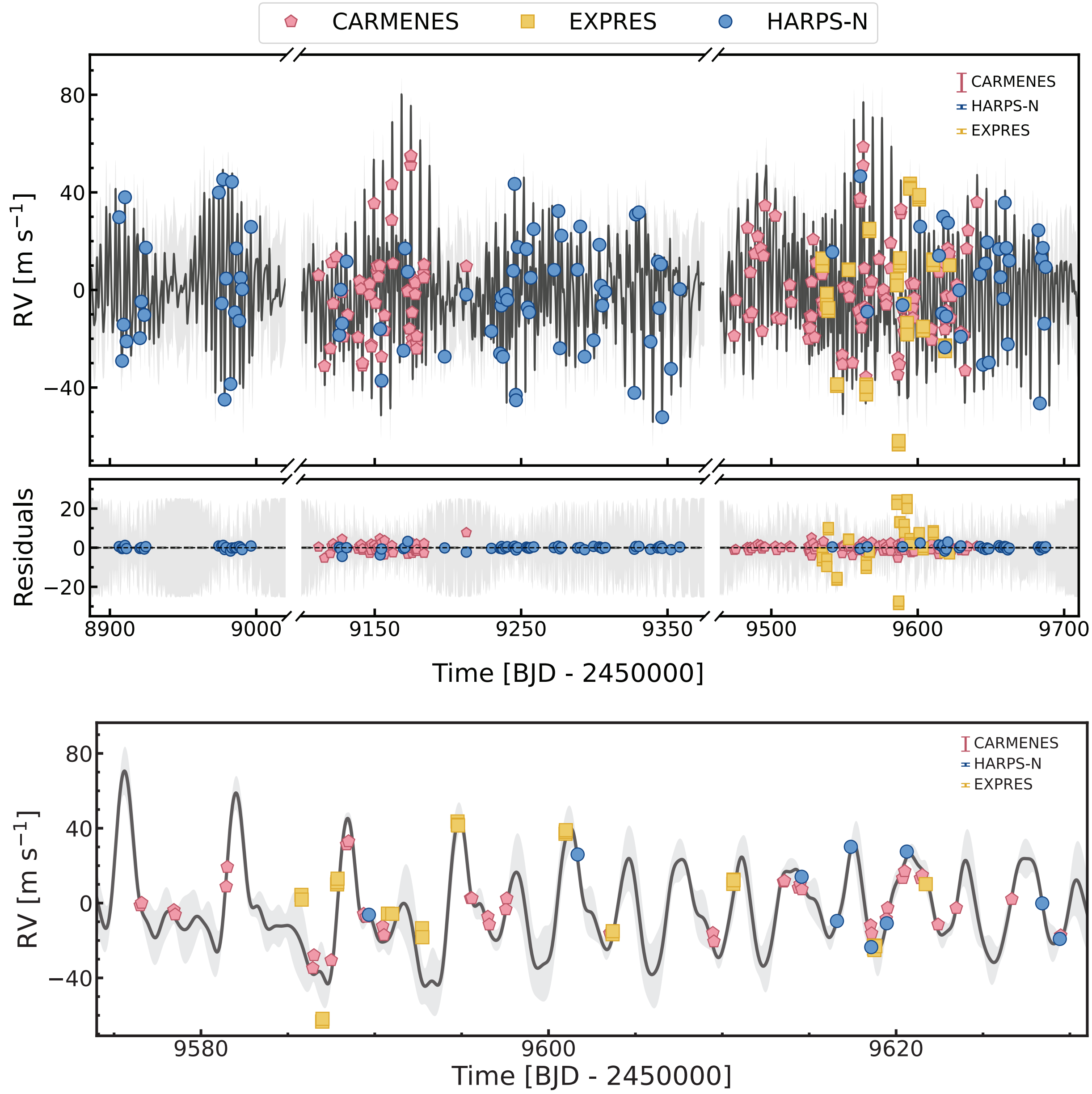


**Figure 5. Top:** The full GP model M3 (see Table 1 and Section 5.1), which we adopt for our analysis. The black line shows the mean GP prediction, and the shaded grey region corresponds to the $1\sigma$ model uncertainty. The bottom panel displays the residuals between the observed data and the GP prediction. **Bottom:** Zoomed-in view of the adopted GP model (M3). The GP, which is conditioned only on the HARPS-N and CARMENES data, accurately reproduces the structure of those datasets, but some features of the held-out EXPRES data are not fully captured. Although the model performs reasonably well, the GP may still miss patterns in new data. Overall, the GP provides a reliable fit while maintaining good predictive performance, even if some small-scale variability in the test set is not reproduced.

signals, with $3\sigma$ upper limits of 0.4–0.5% excess He I absorption (M. Zhang et al. 2022; J. Orell-Miquel et al. 2024). While the current escape is minimal, the planet's higher density at a young age suggests that HD 63433 c must have lost the majority of its primordial H/He earlier in its history or never had a significant atmosphere. According to theoretical structure models (L. Zeng et al. 2019), HD 63433 c falls in regimes consistent with compositions of 50% $H_2O$ and 50% rock or a mostly Earth-like core with $H_2$ envelopes (Figure 8). This scenario is consistent with evolution models in which a lower-density progenitor planet underwent rapid loss of its primordial envelopes within the first ~100 Myr after formation from high-intensity X-ray or extreme UV radiation (A. P. Jackson et al. 2012; J. E. Owen & Y. Wu 2017). We conclude that HD 63433 c is likely a sub-

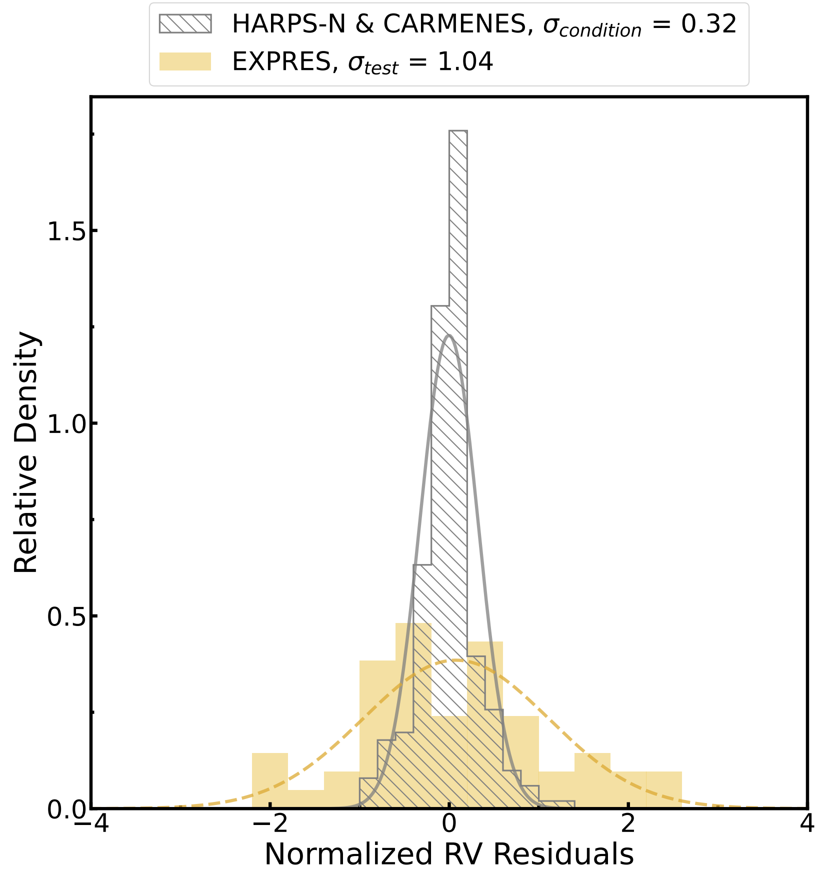


**Figure 6.** Histogram of cross-validation results for the adopted GP model (M3). The conditioning data (HARPS-N and CARMENES) histogram is narrow, with a standard deviation of $\sigma_{\text{condition}} = 0.32$, reflecting that the GP is closely fitting to the training points. The testing data (EXPRES) histogram, modeled with a Gaussian, has a standard deviation of $\sigma_{\text{test}} = 1.04$, signaling that the model generalizes reasonably to unseen data.

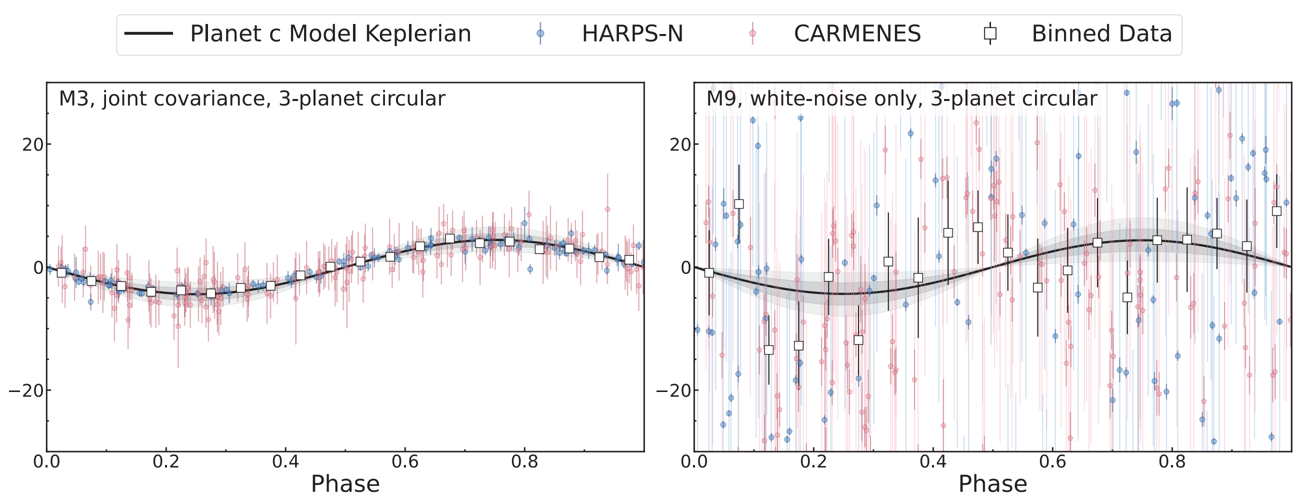


**Figure 7.** Phase-folded RV models of HD 63433 c, with (left) and without (right) a GP. The data are binned in the white squares, and the best-fit Keplerian model is plotted in black with $1\sigma$ and $2\sigma$ uncertainties in grey. On the left, the stellar activity component of the adopted GP model (M3) is subtracted from the data, greatly improving the agreement between the observations and the Keplerian model. The larger, lighter error bars on the right represent the model jitter added in quadrature to the data uncertainties, displaying the increased scatter ($\sim$50 m/s) when stellar activity is not accounted for.

Neptune that has reached its final radius and will not undergo significant contraction in the future.

We are unable to confidently detect the RV signal of the middle planet HD 63433 b. We are able to place a

**Table 2.** Best-fit (median) parameters for adopted model M3.

| Parameter | Best-fit value[a] |
|---|---|
| **Fitted** | |
| *GP terms:* | |
| $\eta_{1,HARPS-N}$ [m s$^{-1}$] | $25.44^{+2.10}_{-1.96}$ |
| $\eta_{1,CARMENES}$ [m s$^{-1}$] | $18.71^{+1.62}_{-1.39}$ |
| $\eta_2$ [d] | $19.31^{+2.00}_{-1.66}$ |
| $\eta_3$ [d] | $6.39 \pm 0.02$ |
| $\eta_4$ | $0.21 \pm 0.02$ |
| *Instrumental terms:* | |
| $\sigma_{jit,HARPS-N}$ [m s$^{-1}$] | $3.52^{+2.46}_{-2.34}$ |
| $\gamma_{HARPS-N}$ [m s$^{-1}$] | $-15805.51^{+3.18}_{-3.16}$ |
| $\sigma_{jit,CARMENES}$ [m s$^{-1}$] | $0.73^{+0.78}_{-0.51}$ |
| $\gamma_{CARMENES}$ [m s$^{-1}$] | $2.62^{+2.40}_{-2.39}$ |
| *Planet b parameters:* | |
| $T_{\rm conj,b}$ [BJD - 2450000] | $8916.45286 \pm 0.00042$ |
| $P_b$ [d] | $7.107934 \pm 0.000005$ |
| $K_b$ [m s$^{-1}$] | $< 7.90$ |
| *Planet c parameters:* | |
| $T_{\rm conj,c}$ [BJD - 2450000] | $8844.05971 \pm 0.00052$ |
| $P_c$ [d] | $20.543784 \pm 0.000016$ |
| $K_c$ [m s$^{-1}$] | $4.41^{+1.04}_{-1.00}$ |
| *Planet d parameters:* | |
| $T_{\rm conj,d}$ [BJD - 2450000] | $9373.82337 \pm 0.00103$ |
| $P_d$ [d] | $4.209078 \pm 0.000022$ |
| $K_d$ [m s$^{-1}$] | $< 4.12$ |
| **Derived** | |
| $m_b$ [$M_\oplus$] | $< 23.61$ |
| $\rho_b$ [g cm$^{-3}$] | $< 15.18$[b] |
| $m_c$ [$M_\oplus$] | $18.78^{+4.44}_{-4.26}$ |
| $\rho_c$ [g cm$^{-3}$] | $6.43^{+1.79}_{-1.58}$[b] |
| $m_d$ [$M_\oplus$] | $< 10.33$ |
| $\rho_d$ [g cm$^{-3}$] | $< 50.60$[b] |

[a] The uncertainties are given as the 16th and 84th percentiles of the posterior distributions and upper limits are reported at $3\sigma$.

[b] Mean densities derived using radii reported in B. K. Capistrant et al. (2024).

$3\sigma$ upper limit on the mass of this planet of 23.61 $M_\oplus$, in agreement with the upper limit of $21.76 M_\oplus$ found in M. Mallorquín et al. (2023). Similar to planet c, HD 63433 b has non-detections of atmospheric escape (M. Zhang et al. 2022; J. Orell-Miquel et al. 2024; M. K. Alam et al. 2024). Hydrodynamical models from M. Zhang et al. (2022) estimated a short atmospheric mass-loss timescale of 80 Myr for this planet, suggesting that its primordial H/He atmosphere may have already evaporated at its current age of 414 Myr (J. Jones et al. 2015; B. K. Capistrant et al. 2024).

L. A. Rogers (2015) argues that most planets larger than $\sim 1.6 R_\oplus$ are too low-density to be composed only of iron and silicates. As HD 63433 b has a radius of $2.112^{+0.093}_{-0.086} R_\oplus$ (B. K. Capistrant et al. 2024), it is still large enough to be expected to have a gaseous envelope. Given our current observational constraints, we are unable to distinguish between a pure rock composition and a small H/He atmosphere. Thus, additional RV observations are needed to provide better constraints on the mass in order to rule out rocky compositions.

Finally, we place the first constraints on the mass of HD 63433 d, the innermost rocky planet recently detected by B. K. Capistrant et al. (2024). This planet is slightly larger than Earth at $1.073^{+0.046}_{-0.044} R_\oplus$, and we place a $3\sigma$ upper limit on the mass of $10.33 M_\oplus$. As visualized in Figure 8, the current $3\sigma$ upper limit places HD 63433 beyond the mass limits of a pure iron planet. A pure iron planet with HD 63433 d's radius would have a mass less than $2.72 M_\oplus$ based on models from L. Zeng et al. (2019), which would induce a signal with a semi-amplitude of at most 1.23 m/s, compared to the $3\sigma$ upper limit semi-amplitude we found of 4.12 m/s.

Given its near Earth-like size and proximity to its host star, HD 63433 d is unlikely to host any substantial atmosphere (L. A. Rogers 2015). Our mass limit is consistent with a terrestrial composition of this planet, but actually detecting the RV amplitude of HD 63433 d with this composition will be challenging, since the semi-amplitude is only expected to be about $\sim$55 cm/s. Placing further constraints would require a coordinated observing campaign with dense sampling and advanced stellar activity mitigation methods.

With an age of 414 Myr (J. Jones et al. 2015; B. K. Capistrant et al. 2024), the HD 63433 planets are old enough to no longer be rapidly contracting, and should be roughly their "final" radii; for example, J. E. Owen 2020 calculates that 500 Myr sub-Neptunes should contract by $\sim$5% before achieving their $\sim$Gyr radii. However, the difference between expected current and final radii is significant enough for future extreme precision RV campaigns to measure and use precise mass measurements to further constrain planetary evolution models. Although the radius difference for these moderately young sub-Neptunes relative to their Gyr radii is smaller than for very young (<10 Myr) planets, their lower stellar activity makes precise RV measurements more feasible. Stellar activity, and not RV precision, remains the major hurdle our field will need to overcome to make

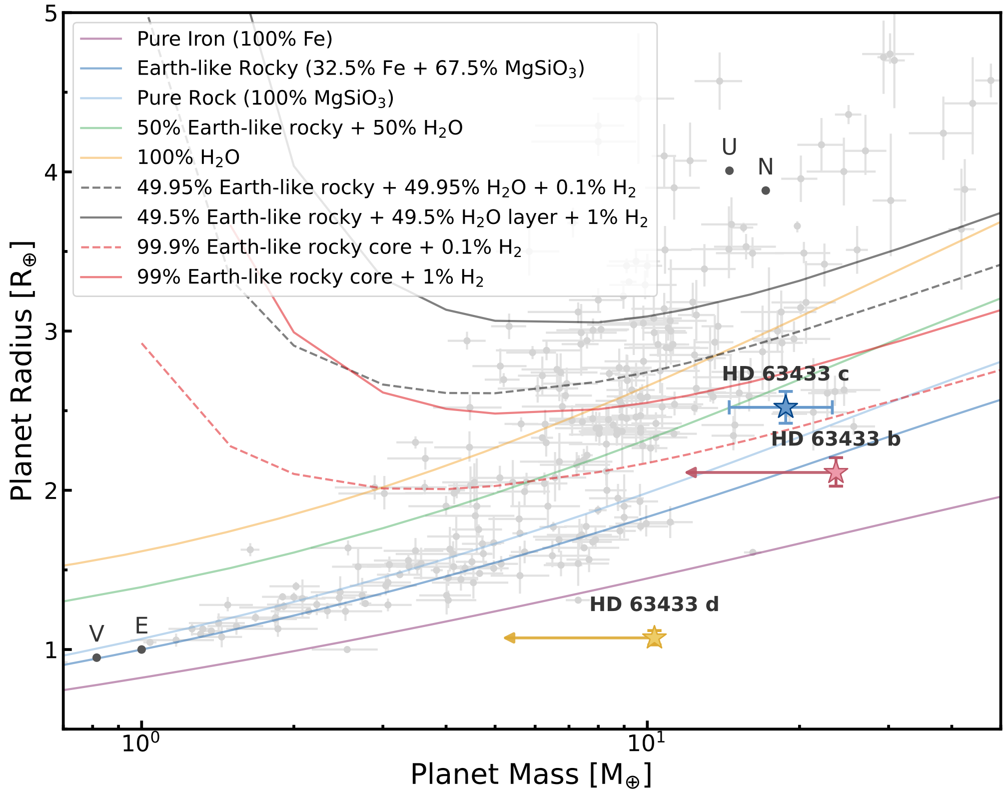


**Figure 8.** Mass-radius diagram for small exoplanets. The HD 63433 planets are shown as colored stars and labeled. Other planets from the NASA Exoplanet Archive (J. L. Christiansen et al. 2025) are shown as grey points, and composition curves from L. Zeng et al. (2019) are shown as solid and dashed curves. HD 63433 c has a reliably measured mass incompatible with an Earth-like rocky composition, suggesting it has retained a low-density volatile envelope. HD 63433 b and c have only upper limits on their masses that cannot uniquely constrain their compositions.

these measurements. This work, which has studied the activity of the star HD 63433, as well as the impact of various modeling assumptions on the inferred planet properties, is a necessary first step toward making planetary mass measurements with sufficient precision to distinguish between these evolution models.

### 5.3. *Future Work with `eurydice`*

Future work with `eurydice` aims to expand the framework to handle multi-dimensional GPs that jointly model photometry or activity indicators. These multi-dimensional GPs are expected to provide an alternative treatment of correlated stellar noise and are already implemented in other RV modeling frameworks like `pyaneti` (O. Barragán et al. 2022), so bridging cross-validation to those more elaborate models will be an important path forward. Additionally, `eurydice` currently implements a simple hold-out cross-validation method. In the future, more advanced cross-validation techniques, such as k-fold or leave-one-out cross-validation, will be implemented to validate models at different levels of robustness.

## 6. CONCLUSIONS

In this work, we introduced `eurydice` and applied it to both existing and new observations of the HD 63433 system. Our main conclusions can be summarized as follows:

1. `eurydice` is an open-source Python package for performing hold-out cross-validation on GP models of stellar activity. It provides a standardized, flexible framework to evaluate how well these models predict unseen RV data, making it an effective tool for assessing predictive performance and identifying signs of overfitting.

2. We applied `eurydice` to models of the young HD 63433 planetary system. We derived mass measurements for the outermost planet HD 63433 c at $19^{+4}_{-4} M_\oplus$, a 4.2$\sigma$ detection. We placed 3$\sigma$ upper limit masses for inner planets HD 63433 b and d at $< 22 M_\oplus$ and $< 10 M_\oplus$, respectively. These masses are consistent with past work reported in M. Damasso et al. (2023) and M. Mallorquín et al.

(2023), and cross-validation affirms the reliability of these measurements.

3. We directly compared two GP approaches for handling multi-instrument RV data. In the first, a joint covariance matrix encodes the assumption of correlated signals across instruments, while the added likelihood approach, much more prevalent in the literature, treats each instrument independently. Our cross-validation tests found marginally better performance for the joint covariance matrix models, but there is no significant difference in the planetary parameters derived. We adopt the joint covariance matrix model as a more physically motivated description of stellar activity, but future work should rigorously test this assumption.

GP modeling is becoming increasingly standard for RV studies, yet the reliability of GP fits remains a source of uncertainty in understanding young exoplanet systems. Going forward, we hope cross-validation becomes a standard tool in the exoplanet community, providing more reliable measurements of planet properties and guiding future studies toward a clearer understanding of how young planets mature and evolve.

## ACKNOWLEDGMENTS

K.B. acknowledges support from the MIT Undergraduate Research Opportunities Program (UROP). This work was catalyzed during the Lamat REU program supported by NSF grant 2150255. AV and KB acknowledge support from the Sloan Research Fellowship.

These results made use of the Lowell Discovery Telescope (LDT) at Lowell Observatory. Lowell is a private, non-profit institution dedicated to astrophysical research and public appreciation of astronomy and operates the LDT in partnership with Boston University, the University of Maryland, the University of Toledo, Northern Arizona University and Yale University. This work used the EXtreme PREcision Spectrograph (EXPRES) that was designed and commissioned at Yale with financial support by the U.S. National Science Foundation under MRI-1429365 and ATI-1509436 (PI D. Fischer). The authors acknowledge the MIT Office of Research Computing and Data for providing high-performance computing resources that have contributed to the research results reported within this paper. This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program.

*Software:* emcee (D. Foreman-Mackey et al. 2019), tinygp (D. Foreman-Mackey et al. 2024), jaxoplanet (S. Hattori et al. 2024), numpy (C. R. Harris et al. 2020), corner (D. Foreman-Mackey 2016), matplotlib (J. D. Hunter 2007), pandas ( The Pandas Development Team 2020), scipy (P. Virtanen et al. 2020)

## REFERENCES

Aigrain, S., & Foreman-Mackey, D. 2023, ARA&A, 61, 329, doi: 10.1146/annurev-astro-052920-103508

Aigrain, S., Pont, F., & Zucker, S. 2012, MNRAS, 419, 3147, doi: 10.1111/j.1365-2966.2011.19960.x

Alam, M. K., Kirk, J., Dos Santos, L. A., et al. 2024, AJ, 168, 102, doi: 10.3847/1538-3881/ad50d4

Alves, D. R., Jenkins, J. S., Vines, J. I., et al. 2025, MNRAS, 536, 1538, doi: 10.1093/mnras/stae2582

Ambikasaran, S., Foreman-Mackey, D., Greengard, L., Hogg, D. W., & O'Neil, M. 2015, IEEE Transactions on Pattern Analysis and Machine Intelligence, 38, 252, doi: 10.1109/TPAMI.2015.2448083

Barat, S., Désert, J.-M., Vazan, A., et al. 2024, Nature Astronomy, 8, 899, doi: 10.1038/s41550-024-02257-0

Barber, M. G., Mann, A. W., Vanderburg, A., et al. 2024, Nature, 635, 574, doi: 10.1038/s41586-024-08123-3

Barragán, O., Aigrain, S., Rajpaul, V. M., & Zicher, N. 2022, MNRAS, 509, 866, doi: 10.1093/mnras/stab2889

Bean, J. L., Raymond, S. N., & Owen, J. E. 2021, Journal of Geophysical Research (Planets), 126, e06639, doi: 10.1029/2020JE006639

Blackman, R. T., Fischer, D. A., Jurgenson, C. A., et al. 2020, AJ, 159, 238, doi: 10.3847/1538-3881/ab811d

Blunt, S., Carvalho, A., David, T. J., et al. 2023, AJ, 166, 62, doi: 10.3847/1538-3881/acde78

Bouma, L. G., Hartman, J. D., Bhatti, W., Winn, J. N., & Bakos, G. Á. 2019, ApJS, 245, 13, doi: 10.3847/1538-4365/ab4a7e

Caballero, J. A., Guàrdia, J., López del Fresno, M., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9910, Observatory Operations: Strategies, Processes, and Systems VI, ed. A. B. Peck, R. L. Seaman, & C. R. Benn, 99100E, doi: 10.1117/12.2233574

Cale, B. L., Reefe, M., Plavchan, P., et al. 2021, AJ, 162, 295, doi: 10.3847/1538-3881/ac2c80

Capistrant, B. K., Soares-Furtado, M., Vanderburg, A., et al. 2024, AJ, 167, 54, doi: 10.3847/1538-3881/ad1039

Christiansen, J. L., McElroy, D. L., Harbut, M., et al. 2025, PSJ, 6, 186, doi: 10.3847/PSJ/ade3c2

Cosentino, R., Lovis, C., Pepe, F., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, ed. I. S. McLean, S. K. Ramsay, & H. Takami, 84461V, doi: 10.1117/12.925738

Cosentino, R., Lovis, C., Pepe, F., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, ed. S. K. Ramsay, I. S. McLean, & H. Takami, 91478C, doi: 10.1117/12.2055813

Damasso, M., Lanza, A. F., Benatti, S., et al. 2020, A&A, 642, A133, doi: 10.1051/0004-6361/202038864

Damasso, M., Locci, D., Benatti, S., et al. 2023, A&A, 672, A126, doi: 10.1051/0004-6361/202245391

David, T. J., Hillenbrand, L. A., Petigura, E. A., et al. 2016, Nature, 534, 658, doi: 10.1038/nature18293

Donati, J.-F., Finociety, B., Cristofari, P. I., et al. 2024, MNRAS, 530, 264, doi: 10.1093/mnras/stae675

Foreman-Mackey, D. 2016, The Journal of Open Source Software, 1, 24, doi: 10.21105/joss.00024

Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, AJ, 154, 220, doi: 10.3847/1538-3881/aa9332

Foreman-Mackey, D., Farr, W., Sinha, M., et al. 2019, The Journal of Open Source Software, 4, 1864, doi: 10.21105/joss.01864

Foreman-Mackey, D., Yu, W., Yadav, S., et al. 2024, dfm/tinygp: The tiniest of Gaussian Process libraries, v0.3.0 Zenodo, doi: 10.5281/zenodo.10463641

Frelikh, R., Jang, H., Murray-Clay, R. A., & Petrovich, C. 2019, ApJL, 884, L47, doi: 10.3847/2041-8213/ab4a7b

Ginzburg, S., Schlichting, H. E., & Sari, R. 2018, MNRAS, 476, 759, doi: 10.1093/mnras/sty290

Harada, C. K., Dressing, C. D., Turtelboom, E. V., et al. 2024, arXiv e-prints, arXiv:2409.10679, doi: 10.48550/arXiv.2409.10679

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357, doi: 10.1038/s41586-020-2649-2

Hattori, S., Garcia, L., Murray, C., et al. 2024, exoplanet-dev/jaxoplanet: Astronomical time series analysis with JAX, v0.0.2 Zenodo, doi: 10.5281/zenodo.10736936

Haywood, R. D., Collier Cameron, A., Queloz, D., et al. 2014, MNRAS, 443, 2517, doi: 10.1093/mnras/stu1320

Helled, R., Nettelmann, N., & Guillot, T. 2020, Space Science Reviews, 216, doi: 10.1007/s11214-020-00660-3

Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90, doi: 10.1109/MCSE.2007.55

Jackson, A. P., Davis, T. A., & Wheatley, P. J. 2012, MNRAS, 422, 2024, doi: 10.1111/j.1365-2966.2012.20657.x

Jones, J., White, R. J., Boyajian, T., et al. 2015, ApJ, 813, 58, doi: 10.1088/0004-637X/813/1/58

Jurgenson, C., Fischer, D., McCracken, T., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, ed. C. J. Evans, L. Simard, & H. Takami, 99086T, doi: 10.1117/12.2233002

Klein, B., Aigrain, S., Cretignier, M., et al. 2024, MNRAS, 531, 4238, doi: 10.1093/mnras/stae1313

Leet, C., Fischer, D. A., & Valenti, J. A. 2019, AJ, 157, 187, doi: 10.3847/1538-3881/ab0d86

Levine, S. E., Bida, T. A., Chylek, T., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8444, Ground-based and Airborne Telescopes IV, ed. L. M. Stepp, R. Gilmozzi, & H. J. Hall, 844419, doi: 10.1117/12.926415

MacLeod, M., Blunt, S., De Rosa, R. J., et al. 2025, ApJ, 978, 50, doi: 10.3847/1538-4357/ad93c8

Mallorquín, M., Béjar, V. J. S., Lodieu, N., et al. 2023, A&A, 671, A163, doi: 10.1051/0004-6361/202245397

Mann, A. W., Newton, E. R., Rizzuto, A. C., et al. 2016, AJ, 152, 61, doi: 10.3847/0004-6256/152/3/61

Mann, A. W., Johnson, M. C., Vanderburg, A., et al. 2020, AJ, 160, 179, doi: 10.3847/1538-3881/abae64

Nardiello, D., Piotto, G., Deleuil, M., et al. 2020, MNRAS, 495, 4924, doi: 10.1093/mnras/staa1465

Newton, E. R., Mann, A. W., Tofflemire, B. M., et al. 2019, ApJL, 880, L17, doi: 10.3847/2041-8213/ab2988

Nicholson, B. A., & Aigrain, S. 2022, MNRAS, 515, 5251, doi: 10.1093/mnras/stac2097

Orell-Miquel, J., Murgas, F., Pallé, E., et al. 2024, A&A, 689, A179, doi: 10.1051/0004-6361/202449411

O'Sullivan, N. K., Aigrain, S., Cretignier, M., et al. 2025, MNRAS, 541, 3942, doi: 10.1093/mnras/staf1168

Owen, J. E. 2020, MNRAS, 498, 5030, doi: 10.1093/mnras/staa2784

Owen, J. E., & Wu, Y. 2013, ApJ, 775, 105, doi: 10.1088/0004-637X/775/2/105

Owen, J. E., & Wu, Y. 2016, ApJ, 817, 107, doi: 10.3847/0004-637X/817/2/107

Owen, J. E., & Wu, Y. 2017, ApJ, 847, 29, doi: 10.3847/1538-4357/aa890a

Petersburg, R. R., Ong, J. M. J., Zhao, L. L., et al. 2020, AJ, 159, 187, doi: 10.3847/1538-3881/ab7e31

Quirrenbach, A., Amado, P. J., Caballero, J. A., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, ed. S. K. Ramsay, I. S. McLean, & H. Takami, 91471F, doi: 10.1117/12.2056453

Quirrenbach, A., Amado, P. J., Caballero, J. A., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, ed. C. J. Evans, L. Simard, & H. Takami, 990812, doi: 10.1117/12.2231880

Rasmussen, C. E., & Williams, C. K. I. 2006, Gaussian Processes for Machine Learning

Rogers, L. A. 2015, ApJ, 801, 41, doi: 10.1088/0004-637X/801/1/41

Schlichting, H. E. 2018, in Handbook of Exoplanets, ed. H. J. Deeg & J. A. Belmonte, 141, doi: 10.1007/978-3-319-55333-7_141

Suárez Mascareño, A., Damasso, M., Lodieu, N., et al. 2021, Nature Astronomy, 6, 232, doi: 10.1038/s41550-021-01533-7

The Pandas Development Team. 2020, pandas-dev/pandas: Pandas, latest Zenodo, doi: 10.5281/zenodo.3509134

Tran, Q. H., Bowler, B. P., Cochran, W. D., et al. 2021, AJ, 161, 173, doi: 10.3847/1538-3881/abe041

Vanderburg, A., Plavchan, P., Johnson, J. A., et al. 2016, MNRAS, 459, 3565, doi: 10.1093/mnras/stw863

VanderPlas, J., Connolly, A. J., Ivezic, Z., & Gray, A. 2012, in Proceedings of Conference on Intelligent Data Understanding (CIDU, 47–54, doi: 10.1109/CIDU.2012.6382200

VanderPlas, J. T., & Ivezić, Ž. 2015, ApJ, 812, 18, doi: 10.1088/0004-637X/812/1/18

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261, doi: 10.1038/s41592-019-0686-2

Zechmeister, M., Reiners, A., Amado, P. J., et al. 2018, A&A, 609, A12, doi: 10.1051/0004-6361/201731483

Zeng, L., Jacobsen, S. B., Sasselov, D. D., et al. 2019, Proceedings of the National Academy of Science, 116, 9723, doi: 10.1073/pnas.1812905116

Zhang, M., Knutson, H. A., Wang, L., et al. 2022, AJ, 163, 68, doi: 10.3847/1538-3881/ac3f3b

Zhao, L. L., Hogg, D. W., Bedell, M., & Fischer, D. A. 2021, AJ, 161, 80, doi: 10.3847/1538-3881/abd105

# APPENDIX

**Table A1**. RVs derived from EXPRES spectra.

| Time [BJD] | RV [m $s^{-1}$] | $\sigma_{RV}$ [m $s^{-1}$] |
|---|---|---|
| 2459534.810552 | 10.760 | 0.886 |
| 2459534.823394 | 11.530 | 0.755 |
| 2459534.837843 | 9.019 | 0.731 |
| 2459534.853980 | 12.103 | 0.752 |
| 2459537.900158 | -2.315 | 0.846 |
| 2459537.914665 | -6.865 | 0.931 |
| 2459538.961198 | -9.500 | 0.808 |
| 2459538.969824 | -8.240 | 0.730 |
| 2459544.845671 | -40.111 | 1.075 |
| 2459544.860033 | -40.224 | 0.907 |
| 2459544.874006 | -39.475 | 0.885 |
| 2459552.789574 | 7.608 | 0.772 |
| 2459552.797160 | 7.251 | 0.845 |
| 2459564.828333 | -43.729 | 0.804 |
| 2459564.842154 | -39.574 | 0.758 |
| 2459564.856901 | -40.602 | 0.786 |
| 2459566.953363 | 23.235 | 0.945 |
| 2459566.967272 | 24.278 | 0.866 |
| 2459585.779930 | 3.534 | 0.962 |
| 2459585.794421 | 1.015 | 0.843 |
| 2459586.973151 | -64.115 | 0.862 |
| 2459586.983498 | -64.193 | 0.789 |
| 2459586.993271 | -62.701 | 0.839 |
| 2459587.832862 | 9.138 | 0.868 |
| 2459587.846746 | 10.292 | 0.888 |
| 2459587.861261 | 12.213 | 1.033 |
| 2459590.757694 | -6.478 | 1.064 |
| 2459591.006895 | -6.619 | 0.954 |
| 2459592.721409 | -14.282 | 0.999 |
| 2459592.738293 | -19.110 | 1.002 |
| 2459594.762732 | 42.774 | 1.064 |
| 2459594.777590 | 41.057 | 1.156 |
| 2459594.793006 | 40.632 | 1.205 |
| 2459600.976240 | 36.196 | 0.857 |
| 2459600.986720 | 37.261 | 0.810 |
| 2459600.998656 | 38.135 | 0.887 |

**Table A1** *continued*

**Table A1** *(continued)*

| Time [BJD] | RV [m s$^{-1}$] | $\sigma_{RV}$[m s$^{-1}$] |
|---|---|---|
| 2459603.683994 | -17.507 | 0.771 |
| 2459603.693087 | -15.851 | 0.771 |
| 2459610.630149 | 9.387 | 0.781 |
| 2459610.638326 | 11.722 | 0.742 |
| 2459610.646550 | 11.372 | 0.792 |
| 2459618.713927 | -23.526 | 0.800 |
| 2459618.726495 | -25.140 | 0.731 |
| 2459618.739595 | -23.704 | 0.819 |
| 2459618.752535 | -25.968 | 0.830 |
| 2459618.803671 | -23.913 | 0.796 |
| 2459621.705842 | 9.282 | 1.687 |
| 2459858.997210 | -15.924 | 1.603 |
| 2459865.014861 | -5.659 | 1.423 |
| 2459866.004055 | -8.548 | 1.572 |
| 2459879.951915 | 48.288 | 1.436 |
| 2459882.973990 | 1.128 | 1.579 |

**Table A2.** Prior parameters of all models.

| Parameter | Prior |
|---|---|
| $\eta_{1,HARPS-N}$ [m s$^{-1}$] | $\mathcal{U}(0, 500)$ |
| $\eta_{1,CARMENES}$ [m s$^{-1}$] | $\mathcal{U}(0, 500)$ |
| $\eta_2$ [d] | $\mathcal{U}(0, 500)$ |
| $\eta_3$ [d] | $\mathcal{U}(5.4, 6.6)$ |
| $\eta_4$ | $\mathcal{U}(0, 1)$ |
| $\sigma_{jit,HARPS-N}$ [m s$^{-1}$] | $\mathcal{U}(0, 3\sigma_{HARPS-N})$ |
| $\gamma_{HARPS-N}$ [m s$^{-1}$] | $\mathcal{U}(-16000, -15500)$ |
| $\sigma_{jit,CARMENES}$ [m s$^{-1}$] | $\mathcal{U}(0, 3\sigma_{CARMENES})$ |
| $\gamma_{CARMENES}$ [m s$^{-1}$] | $\mathcal{U}(-3\sigma_{CARMENES}, 3\sigma_{CARMENES})$ |
| $T_{0,b}$ [BJD] | $\mathcal{N}(2458916.45286, 0.00042)^a$ |
| $P_b$ [d] | $\mathcal{N}(7.1079342, 0.0000049)^a$ |
| $\sqrt{e}\sin(\omega)_b$ | $\mathcal{U}(-1, 1)$ |
| $\sqrt{e}\cos(\omega)_b$ | $\mathcal{U}(-1, 1)$ |
| $K_b$ [m s$^{-1}$] | $\mathcal{U}(0, 30)$ |
| $T_{0,c}$ [BJD] | $\mathcal{N}(2458844.05971, 0.00052)^a$ |
| $P_c$ [d] | $\mathcal{N}(20.543784, 0.000016)^a$ |
| $\sqrt{e}\sin(\omega)_c$ | $\mathcal{U}(-1, 1)$ |
| $\sqrt{e}\cos(\omega)_c$ | $\mathcal{U}(-1, 1)$ |
| $K_c$ [m s$^{-1}$] | $\mathcal{U}(0, 30)$ |
| $T_{0,d}$ [BJD] | $\mathcal{N}(2459373.82337, 0.00104)^a$ |
| $P_d$ [d] | $\mathcal{N}(4.209078, 0.000022)^a$ |
| $\sqrt{e}\sin(\omega)_d$ | $\mathcal{U}(-1, 1)$ |
| $\sqrt{e}\cos(\omega)_d$ | $\mathcal{U}(-1, 1)$ |
| $K_d$ [m s$^{-1}$] | $\mathcal{U}(0, 10)$ |

[a] Transit timing and periods taken from B. K. Capistrant et al. (2024)

**Table A3.** Best-fit (median) parameters for two planet models.

| Parameter | Joint-fit (e=$\omega$=0) | Joint-fit (e,$\omega$ free) | Added-likelihoods (e=$\omega$=0) | Added-likelihoods (e,$\omega$ free) |
|---|---|---|---|---|
| $\eta_{1,HARPS-N}$ [m s$^{-1}$] | $25.50^{+2.11}_{-1.94}$ | $25.43^{+2.17}_{-2.05}$ | $24.03^{+2.12}_{-1.87}$ | $23.93^{+2.10}_{-1.87}$ |
| $\eta_{1,CARMENES}$ [m s$^{-1}$] | $18.71^{+1.60}_{-1.37}$ | $18.75^{+1.66}_{-1.42}$ | $20.11^{+2.02}_{-1.74}$ | $20.05^{+1.98}_{-1.74}$ |
| $\eta_2$ [d] | $19.50^{+2.15}_{-1.74}$ | $19.48^{+1.97}_{-1.66}$ | $20.59^{+2.07}_{-1.73}$ | $20.39^{+2.06}_{-1.71}$ |
| $\eta_3$ [d] | $6.38 \pm 0.02$ | $6.39 \pm 0.02$ | $6.38 \pm 0.02$ | $6.38 \pm 0.02$ |
| $\eta_4$ | $0.21 \pm 0.02$ | $0.22 \pm 0.02$ | $0.22 \pm 0.02$ | $0.23 \pm 0.02$ |
| $\sigma_{jit,HARPS-N}$ [m s$^{-1}$] | $3.27^{+2.38}_{-2.18}$ | $3.53^{+2.32}_{-2.31}$ | $1.73^{+2.17}_{-1.23}$ | $1.77^{+2.28}_{-1.26}$ |
| $\gamma_{HARPS-N}$ [m s$^{-1}$] | $-15805.40^{+3.18}_{-3.15}$ | $-15804.90 \pm 3.25$ | $-15806.30^{+3.51}_{-3.50}$ | $-15806.04^{+3.51}_{-3.52}$ |
| $\sigma_{jit,CARMENES}$ [m s$^{-1}$] | $0.76^{+0.82}_{-0.53}$ | $0.81^{+0.87}_{-0.57}$ | $0.72^{+0.78}_{-0.50}$ | $0.73^{+0.76}_{-0.51}$ |
| $\gamma_{CARMENES}$ [m s$^{-1}$] | $2.60^{+2.38}_{-2.39}$ | $2.72^{+2.39}_{-2.48}$ | $1.89^{+3.62}_{-3.68}$ | $1.92^{+3.69}_{-3.68}$ |
| $T_{\rm conj,b}$ [BJD - 2450000] | $8916.45286 \pm 0.00042$ | $8916.45261^{+0.00064}_{-0.00063}$ | $8916.45286^{+0.00042}_{-0.00041}$ | $8916.45262^{+0.00065}_{-0.00064}$ |
| $\sqrt{e}\sin(\omega)_b$ | 0 (fixed) | $-0.18^{+0.50}_{-0.48}$ | 0 (fixed) | $-0.25^{+0.57}_{-0.46}$ |
| $\sqrt{e}\cos(\omega)_b$ | 0 (fixed) | $0.04^{+0.42}_{-0.77}$ | 0 (fixed) | $0.09^{+0.44}_{-0.80}$ |
| $K_b$ [m s$^{-1}$] | $2.17^{+1.64}_{-1.33}$ | $2.08^{+3.32}_{-1.45}$ | $1.60^{+1.44}_{-1.06}$ | $1.83^{+3.30}_{-1.32}$ |
| $T_{\rm conj,c}$ [BJD - 2450000] | $8844.05971 \pm 0.00052$ | $8844.05973^{+0.00059}_{-0.00060}$ | $8844.05971 \pm 0.00052$ | $8844.05971 \pm 0.00060$ |
| $P_c$ [d] | $20.543784 \pm 0.000016$ | $20.543791 \pm 0.000018$ | $20.543784 \pm 0.000016$ | $20.543791 \pm 0.000018$ |
| $\sqrt{e}\sin(\omega)_c$ | 0 (fixed) | $0.41^{+0.21}_{-0.34}$ | 0 (fixed) | $0.42^{+0.28}_{-0.52}$ |
| $\sqrt{e}\cos(\omega)_c$ | 0 (fixed) | $-0.30^{+0.26}_{-0.17}$ | 0 (fixed) | $-0.20^{+0.46}_{-0.25}$ |
| $K_c$ [m s$^{-1}$] | $4.37^{+1.04}_{-1.00}$ | $5.33^{+1.14}_{-1.13}$ | $3.93^{+1.01}_{-0.97}$ | $4.88^{+1.33}_{-1.15}$ |

**Table A4.** Best-fit (median) parameters for three planet models.

| Parameter | Joint-fit (e,$\omega$ free) | Added-likelihoods (e=$\omega$=0) | Added-likelihoods (e,$\omega$ free) |
|---|---|---|---|
| $\eta_{1,HARPS-N}$ [m s$^{-1}$] | $25.36^{+2.16}_{-1.99}$ | $24.20^{+2.15}_{-1.91}$ | $24.01^{+2.16}_{-1.89}$ |
| $\eta_{1,CARMENES}$ [m s$^{-1}$] | $18.65^{+1.68}_{-1.39}$ | $20.00^{+2.01}_{-1.73}$ | $19.94^{+2.03}_{-1.74}$ |
| $\eta_2$ [d] | $19.30^{+2.03}_{-1.63}$ | $20.44^{+1.99}_{-1.69}$ | $20.29^{+2.10}_{-1.71}$ |
| $\eta_3$ [d] | $6.39 \pm 0.02$ | $6.38 \pm 0.02$ | $6.39 \pm 0.02$ |
| $\eta_4$ | $0.22 \pm 0.02$ | $0.22 \pm 0.02$ | $0.23 \pm 0.02$ |
| $\sigma_{jit,HARPS-N}$ [m s$^{-1}$] | $3.53^{+2.38}_{-2.35}$ | $1.74^{+2.25}_{-1.25}$ | $1.71^{+2.30}_{-1.22}$ |
| $\gamma_{HARPS-N}$ [m s$^{-1}$] | $-15804.98^{+3.27}_{-3.19}$ | $-15806.25^{+3.56}_{-3.50}$ | $-15805.97^{+3.55}_{-3.48}$ |
| $\sigma_{jit,CARMENES}$ [m s$^{-1}$] | $0.79^{+0.87}_{-0.56}$ | $0.70^{+0.77}_{-0.49}$ | $0.73^{+0.81}_{-0.51}$ |
| $\gamma_{CARMENES}$ [m s$^{-1}$] | $2.71^{+2.48}_{-2.34}$ | $1.76^{+3.60}_{-3.68}$ | $1.90^{+3.65}_{-3.61}$ |
| $T_{\rm conj,b}$ [BJD - 2450000] | $8916.45263^{+0.00065}_{-0.00064}$ | $8916.45286 \pm 0.00042$ | $8916.45261 \pm 0.00065$ |
| $P_b$ [d] | $7.107938 \pm 0.000005$ | $7.107934 \pm 0.000005$ | $7.107938 \pm 0.000005$ |
| $\sqrt{e}\sin(\omega)_b$ | $-0.19^{+0.50}_{-0.51}$ | 0 (fixed) | $-0.24^{+0.54}_{-0.46}$ |
| $\sqrt{e}\cos(\omega)_b$ | $0.07^{+0.41}_{-0.74}$ | 0 (fixed) | $0.11^{+0.39}_{-0.80}$ |
| $K_b$ [m s$^{-1}$] | $2.23^{+3.62}_{-1.53}$ | $1.81^{+1.49}_{-1.16}$ | $1.89^{+2.84}_{-1.34}$ |
| $T_{\rm conj,c}$ [BJD - 2450000] | $8844.05972^{+0.00059}_{-0.00061}$ | $8844.05971^{+0.00052}_{-0.00051}$ | $8844.05971^{+0.00061}_{-0.00059}$ |
| $P_c$ [d] | $20.543792 \pm 0.000018$ | $20.543784 \pm 0.000016$ | $20.543791 \pm 0.000018$ |
| $\sqrt{e}\sin(\omega)_c$ | $0.44^{+0.20}_{-0.33}$ | 0 (fixed) | $0.39^{+0.31}_{-0.52}$ |
| $\sqrt{e}\cos(\omega)_c$ | $-0.28^{+0.25}_{-0.17}$ | 0 (fixed) | $-0.23^{+0.45}_{-0.23}$ |
| $K_c$ [m s$^{-1}$] | $5.34^{+1.14}_{-1.08}$ | $3.93^{+1.01}_{-0.97}$ | $4.89^{+1.35}_{-1.16}$ |
| $T_{\rm conj,d}$ [BJD - 2450000] | $9373.82349^{+0.00112}_{-0.00114}$ | $9373.82337 \pm 0.00103$ | $9373.82347^{+0.00109}_{-0.00113}$ |
| $P_d$ [d] | $4.209075^{+0.000023}_{-0.000022}$ | $4.209078 \pm 0.000022$ | $4.209075^{+0.000022}_{-0.000023}$ |
| $\sqrt{e}\sin(\omega)_d$ | $0.27^{+0.50}_{-0.71}$ | 0 (fixed) | $0.10^{+0.57}_{-0.66}$ |
| $\sqrt{e}\cos(\omega)_d$ | $-0.26^{+0.71}_{-0.49}$ | 0 (fixed) | $-0.28^{+0.70}_{-0.43}$ |
| $K_d$ [m s$^{-1}$] | $1.88^{+2.47}_{-1.24}$ | $1.12^{+0.89}_{-0.71}$ | $1.70^{+1.83}_{-1.12}$ |

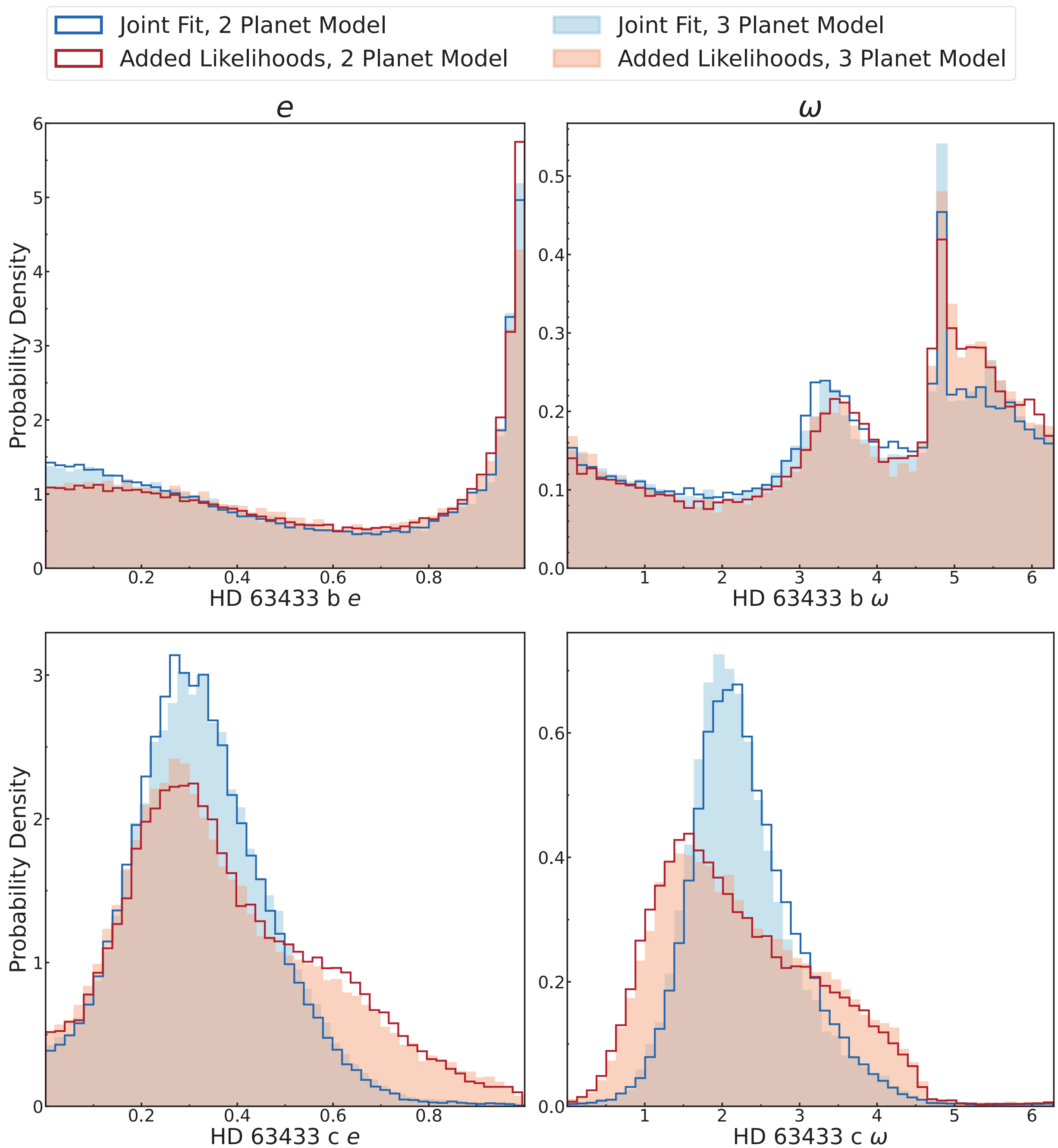


**Figure A1.** $e$ and $\omega$ posterior distributions for planets HD 63433 b and c.

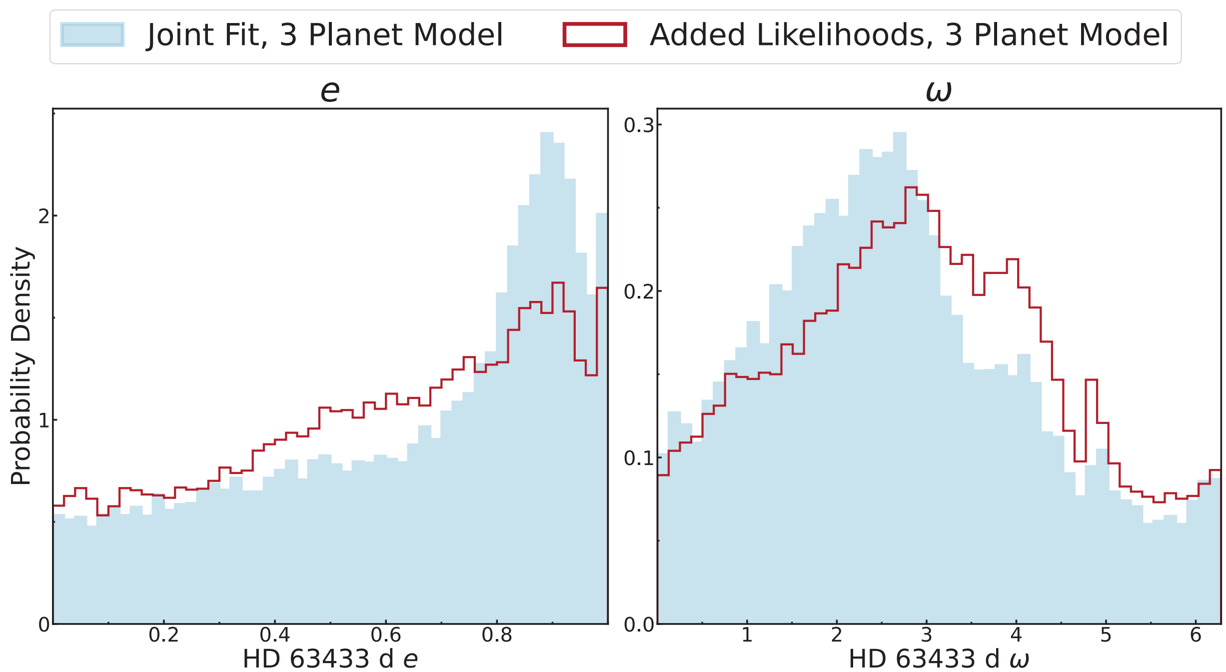


**Figure A2.** $e$ and $\omega$ posterior distributions for planet HD 63433 d.