Radial Velocity Strategies for the Orbital Refinement of Exoplanet Direct Imaging Targets

Zhexing Li,[1] Stephen R. Kane,[1] Sarah Blunt,[2] and Caleb K. Harada[3, *]

[1]*Department of Earth and Planetary Sciences, University of California, Riverside, CA 92521, USA*
[2]*Department of Astronomy and Astrophysics, University of California, Santa Cruz, Santa Cruz, CA, 95064, USA*
[3]*Department of Astronomy, 501 Campbell Hall #3411, University of California, Berkeley, CA 94720, USA*

ABSTRACT

Many potential direct imaging candidates suffer from large orbital period uncertainties, leading to challenges in accurate predictions of future orbital positions and imprecise direct imaging measurements of planetary parameters. To improve the precision in orbital properties, precursor radial velocity (RV) follow-up observations for selected candidates are essential. This study examines the impact of three variables on the orbital period uncertainties of long-period giant planets: the number of future observations, the temporal gap between past and future data, and the temporal coverage of upcoming observations. Our simulations indicate that the orbital phases at which future RV observations are acquired play a significant role in reducing period uncertainties. Additionally, observing too frequently within a given time frame adds limited value to the program once a certain number of observations has been achieved. The temporal gap proves to be the most important factor when there is no strict end time to the observing campaign. However, if a strict end time is set, starting observations earlier yields improved reductions in orbital period uncertainty. These insights offer practical guidance for planning efficient RV follow-up campaigns to maximize the science yield of future space-based direct imaging missions.

Keywords: planetary systems – techniques: radial velocity, direct imaging

1. INTRODUCTION

One of the exoplanet community's primary goals is to directly image terrestrial planets within the habitable zone (HZ) of nearby stars, defined as the region around a star where the energy balance of the planet may enable the presence of surface liquid water (Kasting et al. 1993; Kane & Gelino 2012; Kopparapu et al. 2013, 2014; Kane et al. 2016; Hill et al. 2018, 2023). Many ongoing efforts are prioritizing the development of a next-generation space-based telescope capable of achieving this, such as the Habitable Worlds Observatory (HWO) recommended by the 2020 Astronomy and Astrophysics decadal survey (National Academies of Sciences, Engineering, and Medicine 2021). To maximize the science yield of such missions, previous mission concept reports highlight the need for precursor science using alternative detection methods, including radial ve-

locity (RV) observations, to provide the up-to-date information on the planetary systems of interest (The LUVOIR Team 2019; Gaudi et al. 2020). Within the RV community, efforts are underway that push the detection limits of extreme precision radial velocity (EPRV) instruments (Fischer et al. 2016; Jurgenson et al. 2016; Schwab et al. 2016; Pepe et al. 2021; Gibson et al. 2024). These RV/EPRV data may be applied to planning observations for imaging terrestrial planets (Newman et al. 2023), understanding the orbital dynamics of known systems (Kane et al. 2024; Kane & Burt 2024), addressing correlated noise in EPRV data (Luhn et al. 2023; Gupta & Bedell 2024), improving mass measurements through RV follow-ups (Burt et al. 2018; Cloutier et al. 2018; Meunier et al. 2023), and constraining the presence of additional planets around target stars (Kane et al. 2019; Dalba et al. 2021; Laliotis et al. 2023; Harada et al. 2024a,b). However, an area that remains underexplored is the extraction of precise orbital information for direct imaging (DI) candidates using RV data. While this may not pose significant issues for characterizing habitable terrestrial planets as these planets have typically been

zli245@ucr.edu

* NSF Graduate Research Fellow

observed over numerous orbital periods, long-period giant planets in distant orbits, which will be key targets for the Nancy Grace Roman Space Telescope (hereafter Roman; Spergel et al. 2015; Akeson et al. 2019), present unique challenges.

Roman is set to launch no later than May 2027 and will conduct a science demonstration mission highly relevant to the DI community. Equipped with the Coronagraph Instrument (CGI; Riggs et al. 2021; Bailey et al. 2023), Roman will demonstrate several key technologies in space for the first time (Savransky et al. 2024), which are essential for future DI missions such as the HWO in the search for habitable planets around nearby Sun-like stars. The performance and outcomes of Roman's CGI mission will have a significant impact on the design of the HWO, underscoring the importance of precursor science to support the success of this demonstration mission. Considering that Roman will likely be able to conduct deep imaging and spectroscopy for only a handful of Jovian analogs in reflected light (Wolff et al. 2024), selecting the right candidates with the best-refined orbital parameters is critical to delivering high-quality science products for Roman.

Detecting a planet through DI requires it to be sufficiently bright and separated enough from its star to fall outside the instrumental inner working angle (Kane 2013; Stark et al. 2014; Brown 2015; Kane et al. 2018; Li et al. 2021; Vaughan et al. 2023). These targets, ranging from distant gas giant planets to terrestrial planets in the HZ, will mostly have been observed by the RV method. However, previous work shows that, for long-period gas giants with years or decades-long orbits, the orbital uncertainty increases with the orbital period, mostly reaching around 10% of the period value (Kane 2007; Kane et al. 2009). In extreme cases, RV-derived orbital information can be nearly lost, as the uncertainty approaches the length of the orbital period itself (Kane et al. 2009), even for short-period transiting planets (Dragomir et al. 2020; Kane et al. 2021). Additionally, as pointed out in Spohn et al. (2022), existing RV solutions for planetary systems can grow "stale" over time, losing their predictive accuracy for future DI observability. This presents a challenge for DI because planet–star flux ratios are sensitive to the planetary orbital positions, and prior knowledge about a planet's detectability is essential before DI observations (Turnbull et al. 2021). Accurate extraction of planetary parameters from DI data relies on well-constrained orbital parameters, often derived from RV. Without precise orbital knowledge, there is a risk of scheduling observations at suboptimal epochs when the planet's brightness is reduced, or worse, when it falls below the instrument's

contrast limit or within the inner working angle. Even when robust detections are possible despite orbital imprecision, large uncertainties in orbital periods can complicate estimations of planet–star separation, phase angle, and orbital phase, affecting albedo estimations and, ultimately, planetary characterization quality. However, valuable observing time on precision RV facilities is in high demand across many science programs, and the limited time allocated for RV programs supporting future DI missions is thus in urgent need of the most efficient observing planning (Harada et al. 2024b; Kane & Burt 2024). It is therefore critical to investigate how orbital ephemerides can be effectively refined such that precursor ground-based RV observations may be carried out most efficiently.

Previous works have attempted to investigate optimal RV scheduling using Fisher information (Baluev 2008, 2009; Cloutier et al. 2018; Gupta & Bedell 2024; Lam et al. 2024). This method is mathematically elegant and presents a relatively easy and efficient way to estimate the performance of uncertainties without the need for heavy computation. However, conditions assumed by this analytical method may not hold in all cases and, when violated, could lead to an underestimation of errors or provide unreliable results. For example, Fisher information requires the linearity of equations of the RV model and parameters of interest, sufficient observations such that estimations are unbiased, and the final joint distribution can be estimated by a Gaussian profile (Baluev 2008). All of these conditions may not be satisfied in the case of cold gas giant planets that have relatively large orbital period uncertainties, short observational baselines compared to their respective orbital periods, sparse temporal sampling, and mild to high eccentricities. Numerical methods, despite the heavy computation required, do not require these conditions to be satisfied and can provide a more reliable error estimate. Because of these factors, we opted to conduct numerical simulations to estimate the uncertainties.

In this work, we present the results of a study on how the orbital period of distant giant planets can be effectively refined with respect to three observing variables: number of future observations (Num), temporal gap between past and future observations (Gap), and temporal coverage of future observations (Cov). It is intuitive to expect an inverse relationship between period uncertainty and these variables. Indeed, such relations were derived by Lucy & Sweeney (1971) for the number of total observations and the temporal baseline for all these observations, suggesting that more observations and a longer baseline are preferred. Here, instead of treating all observations as one dataset, the setup is slightly

different, where we treat past and future data as two separate datasets and investigate only how these three variables collectively influence the orbital period uncertainty estimates. Specifically, we aim to include the mutual interactions among the three variables, examine the optimal observation timing for eccentric orbits, determine the conditions at which the uncertainty reduction reaches the point of diminishing return, and identify the relative importance of the variables in the model.

The paper is organized as follows: Section 2 outlines the experiment setup, simulation variables, and validation steps we took to ensure robustness of our methods. Section 3 presents exploratory and a full suite of simulations on the influence of these variables on orbital period uncertainty. Short simulations using similar methods are presented in Section 4 to demonstrate how these results could inform observing strategies. Finally, Section 5 provides a discussion of the results of our simulations, and concluding remarks are contained in Section 6.

2. TESTING SCENARIOS AND SYNTHETIC RV GENERATION

2.1. Test System

For the purpose of this study, we focus on systems with at least one planet of interest to the DI community, particularly those with a distant orbit around their host star and relatively large orbital period uncertainties. The HD 134987 system is one such case and we used the original publication data as a test object. The discovery data spans ∼5000 days and comprises 138 observations acquired via the HIRES and UCLES spectrographs (Jones et al. 2010). The system hosts two known planets, with orbital periods of $P_b = 259.19 \pm 0.07$ days and $Pc = 5000 \pm 400$ days, for the b and c planets, respectively (Jones et al. 2010). The system lies at a distance of 26.20 pc (Gaia Collaboration et al. 2018) and the outermost planet (planet c) is one of the potential candidates for DI using Roman's CGI.

2.2. Test Scenarios

Given the relatively large orbital period uncertainty of planet c in the HD 134987 system, we aimed to conduct a suite of RV observing simulations to study the long-term behavior of errors associated with the best-estimated orbital period. The orbital period uncertainty behavior of distant planets was explored by generating synthetic RVs at future epochs and observing their long-term changes in uncertainty with respect to three variables: *Num*, *Gap*, and *Cov*. The synthetic RVs in this case are based on the best estimates of orbital parameters derived from modeling the original RVs provided by

Jones et al. (2010). Since RV follow-up campaigns often have limited knowledge of the true orbital parameters of planets of interest, observations are typically acquired at a pre-determined cadence to maximize temporal coverage along the RV phase curve throughout the observing season. This approach assumes that the best estimates of orbital components, which may initially deviate from the true parameters, gradually converge toward the true values over time. Similarly, the best estimates used to generate synthetic future RVs for planet c carries an unknown offset from the true orbital parameters. To validate the reliability of using best-estimated parameters for synthetic RV generation and the robustness of the simulation outcomes, we conducted test runs on the three variables mentioned above, analyzing each across three independent synthetic data creation methods:

1. Future RVs are calculated using the true parameters and the RVs are sampled at epochs according to those parameters. This represents an ideal (though unlikely) scenario in which the true parameters are known.

2. Future RVs are calculated using the true parameters, but the sampling of RVs is based on the best-estimate parameters. This reflects a realistic scenario for RV follow-up campaigns, where observational planning relies on the current best estimates, even though RVs are ultimately generated from a ground-truth source.

3. Future RVs are calculated using best-estimate parameters, and the sampling of RVs is also based on those estimates. This scenario represents our simulation setup, where both RV generation and sampling depend on the estimated values.

In each scenario, test runs were conducted by varying one variable at a time while keeping the other two constant. We adjusted the *Num* variable from 6 to 200 in steps of 1, with *Gap* and *Cov* fixed at the full period of the outermost planet (P_{max}); The *Gap* variable, measured in units of P_{max}, was varied from 0.1 to $20 \times P_{max}$ in steps of 0.1, while *Num* was fixed to 50 and *Cov* to $1 \times P_{max}$. The same range and step size were applied to *Cov*, with *Num* and *Gap* set to 50 and $1 \times P_{max}$, respectively. Future RV sampling epochs were determined by these three variables. For example, if the last archival RV observation occurred on 2,455,500 BJD in a system where P_{max} has a true orbital period of 1000 days and a best-estimate period of 1500 days, with the three variables set to 50, $1 \times P_{max}$, and $1 \times P_{max}$, the starting times for future RVs would be $2{,}455{,}500 + 1 \times 1000 = 2{,}456{,}500$ BJD for Scenario 1, and $2{,}455{,}500 + 1 \times 1500$

= 2,457,000 BJD for Scenario 2 and 3 (with *Gap* set to $1 \times P_{max}$, true values for Scenario 1 and estimated values for Scenario 2 and 3). The corresponding ending times would be 2,456,500 + 1× 1000 = 2,457,500 BJD for Scenario 1, and 2,457,000 + 1×1500 = 2,458,500 BJD for Scenario 2 and 3 (with *Cov* set to $1 \times P_{max}$). With the start and end times defined, synthetic RVs were populated roughly evenly between these timestamps, with some randomness introduced (as described below). Scenario 1 represents the most ideal, or "ground truth" case, where the results from this scenario are used as a control case against which results from other scenarios can be compared. Scenario 2 uses true parameters but bases its sampling on the best-estimate values, leading to minor deviations in sampling phases from the ground truth, therefore less optimal. Scenario 3 is the least ideal case, using only best-estimate parameters for both RV generation and sampling, and reflects the practical limitations due to estimation error. Our main interest lies in whether the orbital period and its uncertainty behavior observed in Scenario 3 can approximate those in Scenario 1.

For the test runs on HD 134987, we adopted the orbital parameters derived from our modeling of the original RV data in Jones et al. (2010) as the ground truth, where $P_{max} = 5227$ days. Again, these parameters are generally unknown to observers. Using these parameters, we generated synthetic RVs with the same measurement errors and RV scatter as the original RVs. The modeling of the synthetic RVs produced the best-estimate parameters for the system, where $P_{max} = 5333$ days. Given a set of values for the three variables, as well as the orbital parameters (true or estimated, depending on the scenarios), future synthetic RV timestamps were then evenly populated between the calculated starting and ending times. The number of timestamps created was set to be 50% more than the intended, with the additional RVs later removed. These timestamps were then passed through a Gaussian filter with a sigma of 10 days to further increase the variation among created times. The synthetic RV at each timestamp was then calculated for each planet in the system by solving Kepler's equations using either true or estimated orbital parameters (again depending on the scenarios). RVs were calculated using Equation (1), where V_0 is the RV offset (zero in this case), while ω, f, K, and e are argument of periastron, true anomaly, semiamplitude, and orbital eccentricity of the planet, respectively. The total RVs for the system were simply the superposition of the individual velocity contributions from each planet. These total RVs were then passed through another Gaussian filter with a sigma equal to the rms of the fit to the

original Jones et al. (2010) data to mimic the combined noise from stellar and instrumental sources. Finally, the additional 50% of the RVs introduced earlier were removed from the synthetic time series at randomly selected timestamps to mimic uneven temporal sampling. We note that this process ignores many important factors that are crucial for RV observations, such as correlated noise and the seasonal visibility. These are discussed in Section 5. In short, factors like these are highly target dependent. In order to provide an unbiased and uniform approach for any target, we opted to not include these factors and therefore our results will likely represent a lower limit on the uncertainty.

$$V = V_0 + K[\cos(\omega + f) + e \cos \omega] \qquad (1)$$

The synthetic RVs creation process was conducted at each iteration of each variable under each scenario within our automated pipeline. At each iteration, the generated dataset was fed into the RV modeling toolkit RADVEL, which employs maximum a posteriori optimization for RV fitting and Markov Chain Monte Carlo (MCMC) for confidence interval estimation. Further details about RADVEL can be found in Fulton et al. (2018) and on its documentation page. Our automated pipeline logs each step of the MCMC process and saves the final derived Keplerian orbital parameters along with their associated error estimates from the MCMC posteriors in separate files. Depending on the characteristics of the system and the available RVs, the errors estimated from the posterior distributions can be large if the system is poorly constrained. In such cases, results from each MCMC run may vary due to the probabilistic nature of the MCMC process and the broader peaks in the posteriors. To minimize the impact of any single run skewing the result, we repeated the RADVEL fitting process five times per iteration and then averaged the results from these five MCMC runs to obtain a final set of derived parameters and associated errors. The final error used to gauge the orbital period uncertainty performance was simply the average of its upper and lower bound error.

In Figure 1, we present the results from the testing simulations. Each row contains results for one variable, with the left column showing the behavior of the orbital period and and the right column displaying the associated uncertainty for each variable. In the left column across all three variables, the orbital periods in Scenario 1 and 2 both converge to the true value at a similar rate over time, as expected, given that the synthetic RVs were generated using the "ground truth" values despite minor differences in the starting and ending times due to offsets in the true and estimated period values. For

Scenario 3, the periods exhibit similar behavior, converging with comparable scatter to the other two scenarios as *Num*, *Gap*, or *Cov* variables increase. However, in this case, the periods converge to the estimated orbital period used for generating the synthetic future RVs, which aligns with expectations. In the right column, it is notable that, despite the different methods of generating and sampling future RVs, the uncertainties for P_{max} follow nearly identical trajectories across all three variables: a rapid initial decrease in error estimates that gradually levels off. This consistent behavior indicates that, even with the less optimal RV creation and sampling approach of Scenario 3, we were able to replicate the long-term uncertainty behavior observed in real RV data. Consequently, we adopt the approach of Scenario 3 to create and sample RVs in all subsequent simulations.
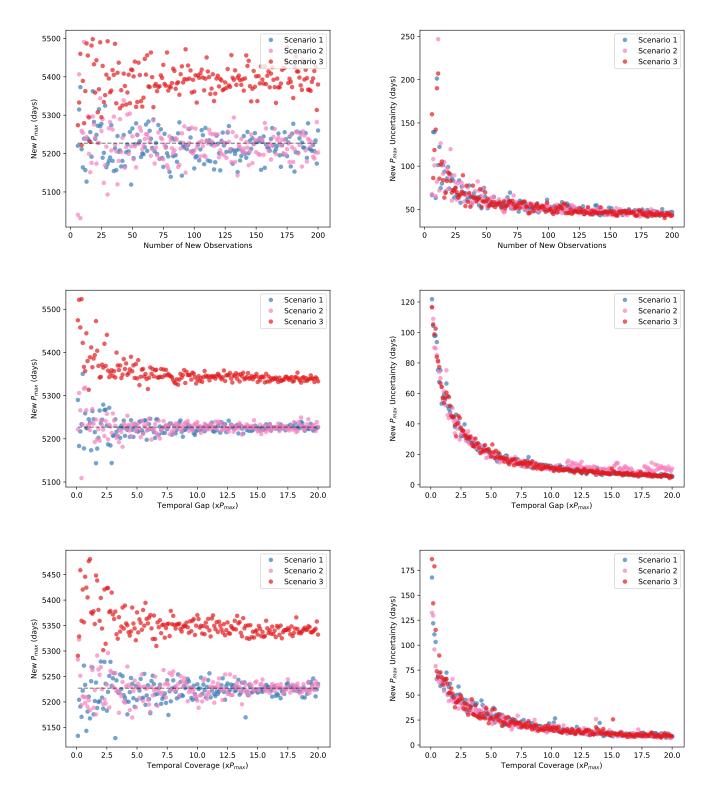
3. SIMULATIONS

3.1. *Exploratory Simulations*

Having selected the strategy for generating synthetic RVs, we proceeded with a set of exploratory simulations on the HD 134987 system. The procedure closely followed the testing process described in Section 2.2, with the exception that we adopted the Scenario 3 strategy and used real archival RVs from the original RV dataset provided by Jones et al. (2010). The P_{max} uncertainty behavior was tested in response to changes in the variables, *Num*, *Gap*, and *Cov*. Additionally, we introduced a fourth variable, instrumental precision (*Inst*), to examine its impact on the uncertainty of P_{max}. We conducted four exploratory runs, each varying one of these variables while holding the others constant: *Num* was fixed at 50, *Gap* at $0.75 \times P_{max}$, and *Cov* at $0.35 \times P_{max}$, except for their respective variable-specific runs. Instrumental precision was set to the average RV error from Jones et al. (2010), which is 1.26 m s^{-1}, except in the fourth run. Specifically, in run 1, *Num* was varied from 6 to 200 with a step of 1; in run 2, *Gap* was varied from $0.01 \times P_{max}$ to $10 \times P_{max}$ with a step of $0.01 \times P_{max}$; in run 3, *Cov* was varied from $0.01 \times P_{max}$ to $10 \times P_{max}$ with a step of $0.01 \times P_{max}$; and in run 4, *Inst* ranged from 0.05 m s^{-1} to 10 m s^{-1} with a step of 0.05 m s^{-1}. As with the previous testing runs, the modeling process was repeated five times at each iteration for each variable to average out variations from the MCMC output. Results were then modeled and the method used for modeling the best-fit curve is described in Section 4. Both the raw results and best fits for the HD 134987 exploratory runs are shown in Figure 2.

For comparison, the orbital period uncertainty for the HD 134987 c planet from our modeling of the Jones et al. (2010) data was 410 days. Results from the exploratory runs indicate that the orbital period uncertainty of this outer planet followed a similar uncertainty reduction pattern across all the three variables: *Num*, *Gap*, and *Cov*, as in the testing runs from the previous section. In all three cases, uncertainty dropped significantly initially, with a slower rate of decrease as the magnitude of each variable increased. The uncertainty behavior of each of the individual variables is expected since Lucy & Sweeney (1971) pointed out the inverse relationship between period error estimates and both number of observations and temporal baseline. For the *Inst* variable, it is unsurprising that orbital period uncertainty increased with larger instrumental precision at future epochs. However, the improvement in precision from meter-per-second instruments to extreme-precision instruments such as those discussed in Crass et al. (2021) was minimal in terms of uncertainty reduction. This suggests that, for reducing orbital period uncertainty of distant gas giant planets, there is little gain from observations using EPRV instruments and therefore utilizing the meter-per-second facilities will suffice for this task. For this reason, we do not explore instrumental precision as another variable and instead assume that all our future RVs would have a similar precision level to past data.

One noticeable difference here is that the uncertainty reduction curve for the *Gap* variable shows periodic humps and dips. This occurs because, in the exploratory runs, the *Cov* variable was set to only 0.35 times the orbital period of planet c. As a result, at each iteration, the generated synthetic RVs at future epochs covered only a portion of planet c's orbital phases. In contrast, for the testing cases in the previous section, *Cov* was set to $1 \times P_{max}$ and covered the entire period space and therefore avoided any bias in the sampled RV phases. Depending on the *Gap* variable provided, the generated data may cover either the maximum or minimum phases of planet c's RV curve or, in some cases, the zero-crossing phases. In the case of HD 134987, humps and dips appear periodically with alternating amplitudes: one large hump and dip followed by a smaller one within each temporal gap interval. Initially, one might expect such a hump and dip pair to occur when future RVs are observed near zero-crossing and maximum/minimum phases of the RV curve, with the new data providing weaker and stronger constraints on the orbital parameters, respectively. However, upon closer examination of the exploratory simulation data, we observe that orbital period uncertainty reduction appears to be greatest when future RVs were populated near the zero-crossing phases rather than near the maxi-

Figure 1. Testing result for orbital period and uncertainty evolution of the HD 134987 c planet based on synthetic RVs. Three scenarios, or data creation methods, are color coded in each panel. The left column shows orbital period and the right column shows associated uncertainty. The horizontal dashed line in the left panels represent the presumed ground truth period mentioned in the text. Top, middle, and bottom panels are results for the *Num*, *Gap*, and *Cov* variables, respectively.
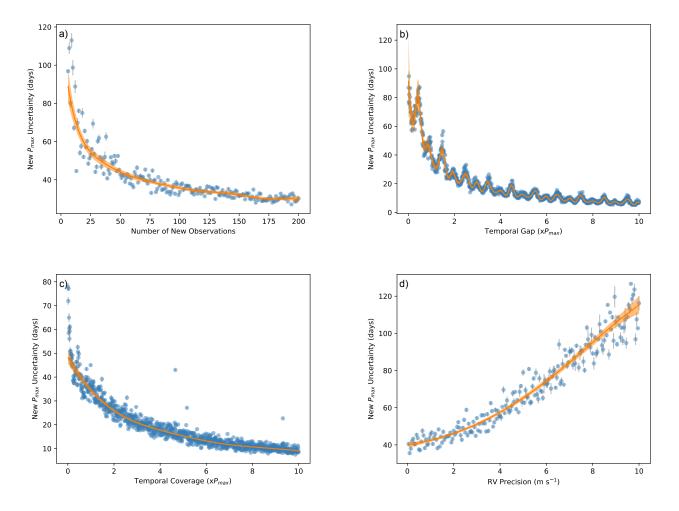
Figure 2. Simulation results from HD 134987 exploratory runs with synthetic data created at future epochs using four variables and real data from Jones et al. (2010). The P_{max} uncertainty before the simulations is 410 days. From panel (a) to panel (d) are P_{max} uncertainty with respect to: *Num*, *Gap*, *Cov*, and *Inst* variables, respectively. Raw data are in blue and the best fit model with confidence interval for each variable is in orange. Uncertainty humps and dips present in the (b) panel was due to the fact that sampling at certain future epochs/RV phases are preferred for orbital period uncertainty reduction.

mum/minimum phases of the planet c's RV curve. In Figure 3, we highlight two groups of humps and dips, separated by one period. The upper panels represent runs for HD 134987, with the left upper panel resembling the *Gap* variable result from Figure 2 but with a shortened *Gap* range on the horizontal axis to focus on the first few groups of features. Specific *Gap* values at which future data begins are indicated by arrows in the left panel. In the upper right panel, both the total RV and planet c's contribution, along with archival and synthetic RVs are shown. Here, P_{max} uncertainty humps at $Gap = 0.47 \times P_{max}$ and $Gap = 1.47 \times P_{max}$ in the upper left panel align with phases near the RV maximum for planet c, while synthetic RVs populated near zero-crossing phases correspond to uncertainty dips at $Gap = 0.22 \times P_{max}$ and $Gap = 1.22 \times P_{max}$. This suggests that, for optimal orbital period constraint, future observations should target phases where the gradient of the RV curve is steepest; that is, when the rate of RV change over time is greatest.

To verify this, we conducted another simulation for the *Gap* variable on HD 219077, a system with a single known planet exhibiting an extremely high eccentricity of 0.769 and an argument of periastron of 57 degrees for the host star (Marmier et al. 2013). The setup was similar to that for HD 134987, except we only used HARPS data and limited the *Gap* variable to a range of $0.01 \times P_{max}$ to $4 \times P_{max}$, with a step of $0.01 \times P_{max}$. The *Cov* variable was set to only $0.01 \times P_{max}$ to make sure the simulation data was able to capture the fast and short RV turnaround. The results are shown in the bottom panels of Figure 3. The bottom left panel shows a similar decline in P_{max} uncertainty with periodic humps and dips occurring once per period. Examining the location of those humps and dips highlighted by the arrows in the RV plot in the bottom right panel reveals that, once again, the greatest uncertainty reductions (e.g., at $Gap = 0.95 \times P_{max}$ and $Gap = 1.95 \times P_{max}$) were achieved by RVs covering the steepest part of the RV curve, where stellar acceleration is the highest. Due to HD 219077's high eccentricity, only one portion of the orbit experiences this high acceleration, resulting in a single pair of humps and dips in the uncertainty reduction. In contrast, the near circular orbit of HD 134987 c has both approaching and receding portions of the orbit contributing similar acceleration along our line of sight. Thus, data covering either portion produces similar uncertainty reduction patterns, leading to two groups of uncertainty humps and dips per orbital period. This confirms that future observations should ideally be conducted at phases of maximum stellar acceleration to achieve the best orbital period uncertainty reduction.

3.2. *Full Simulations*

The exploratory simulations were conducted independently for each variable, with the uncertainty analysis performed without considering the relationships among the three variables. Here, we conducted a suite of simulations to capture a comprehensive view of P_{max} uncertainty reduction in relation to all three variables for HD 134987. Each variable's value was varied within a specified range and step size as before, except that for each iteration, two other variables were varied as well. This setup created a 3D grid with each variable represented on one of the axes. We then ran full simulations to extract P_{max} uncertainty for each combination of the three variables at every grid point in the 3D space, resulting in a 4D space with uncertainty values across all grid points.

We used the original RVs for the HD 134987 system as an example of an archival dataset with which new RVs at future epochs would be created. The procedure of generating the synthetic data was similar to before, but with adjusted step sizes and ranges to better capture relationships among all three variables and their long-term behaviors. To reduce computational cost, we lowered the simulation resolution by increasing the step size for each variable. The 3D grid was formed with the *Num* varying from 6 to 200, in steps of 1, *Gap* from $0.25 \times P_{max}$ to $12 \times P_{max}$ in steps of $0.25 \times P_{max}$, and *Cov* from $0.5 \times P_{max}$ to $12 \times P_{max}$, in steps of $0.5 \times P_{max}$. As in the exploratory and testing runs, we repeated the modeling process five times at each grid point and averaged the results to minimize scatter due to broad MCMC posteriors. In total, 1,123,200 scenarios were run and fit using RadVel for the full simulation.

To model the simulation results, we employed a Random Forest Regressor model to simultaneously analyze the impact of all three variables on P_{max} uncertainty using the Python machine learning framework Scikit-learn (Pedregosa et al. 2011). The modeling allows us to find the best fit through three dimensions, gain insight into the mutual dependence of the modeled variables, and their relative importance in the model. Random forest is a tree-based ensemble learning method that operates by constructing multiple decision trees during training and aggregating their results to make predictions, either for classification or regression tasks. Each decision tree itself is a single, independent modeling unit built on a random subset of the data and predictor variables (in our case, only three: *Num*, *Gap*, and *Cov*). This randomness helps reduce overfitting, as individual trees may capture noise, but the aggregation of results from all the decision trees smooths out the overall prediction. At each iteration, each decision tree decides
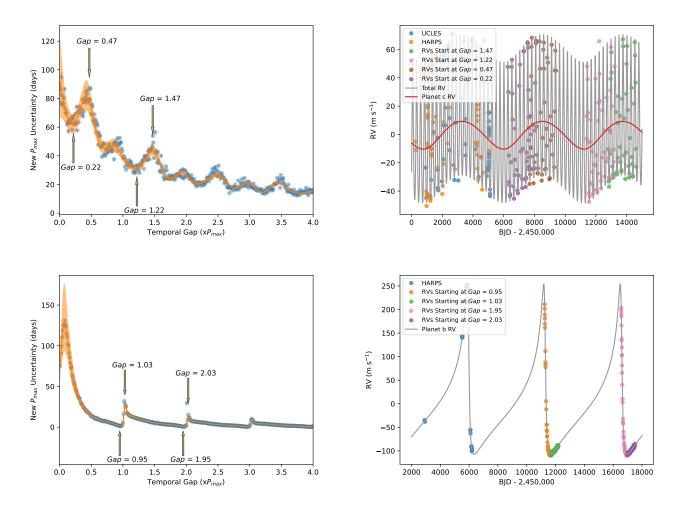
Figure 3. Simulation results of HD 134987 (upper panels) and HD 219077 (lower panels) for the Gap variable. Left panels show the P_{max} uncertainty results from the simulation. Runs highlighted by the arrows near the humps and dips in the left panel corresponds to runs displayed in the right panel. Archival data, synthetic data starting at different Gap values, total RV, and the outer most planet's RV are shown in the right panels. Synthetic RVs were populated according to the orbit of the outer most planet. The P_{max} uncertainties before the simulations for HD 134987 and HD 219077 are 410 and 140 days, respectively

on the best available variable to use at the current tree node and splits the subset of data assigned to each tree in a way that minimize the variance within that subset. Such a splitting strategy occurs iteratively at all nodes within a tree for all trees in the model until it meets certain criteria such as the maximum number of tree depth, minimum number of samples left in the each node, minimum reduction in the chosen performance metric such as the mean squared error (MSE) between the current level within the tree to the previous level, and many others. For regression tasks, each tree outputs a prediction by averaging the values in the relevant leaf node (tree nodes after which no more splitting occurs) and the Random Forest model finally takes the average of the predictions from all the individual trees to provide the final output. Since our dataset is relatively large but not too compli-

cated, this model suits our modeling purpose perfectly as it is able to scale well to large datasets, prevent overfitting to the data, and yield accurate predictions with a relatively simple algorithm.

To train the model, all samples in our dataset were reshuffled and divided into 75% training set and 25% testing set. The model was then trained on the training data using 10-fold cross-validation while performing a grid search for hyperparameter tuning and MSE as the criterion to measure the quality of the split at each node. The hyperparameters tuned along with the values selected for the best model, are included in Table 1. The best-fit model yields a root mean squared error of 2.08 and 1.85 days on the training and testing data, respectively. The training and testing data errors yield similar low values which suggests the model

Table 1. Values Used For Hyperparameter Tuning.

Hyperparameters	Values
max depth	2, 4, 6, 8, 10, **20**, 30, None
min samples leaf	1, **2**, 5, 10, 15, 20
min samples split	2, 4, 6, 8, **10**, 12, 15, 20
max features	**1**, 2, 3
n estimators	100, 200, 300, **400**, 500, 600, 700
min impurity decrease	**0**, 0.01

Note—Hyperparameter values selected for the best-fit model are highlighted in bold. Details of each hyperparameter see scikit-learn documentation page.

has been well-trained. Once the model was trained, we applied the entire dataset to the best-fit model, with results smoothed by our model shown in Figure 4. Since our data is in four dimensions, we plot the *Num* variable on the horizontal axis and the *Gap* variable on the vertical axis, with the *Cov* variable represented in each panel, and the derived P_{max} uncertainty color coded in each panel. Once again, the three variables are expressed in units of the orbital period of the outermost planet, P_{max}. Due to the entirely data-driven nature of the Random Forest algorithm, the final smoothed model of the simulation data still exhibits some localized fluctuation due to noise and therefore does not produce a completely smoothed contour as a traditional parametric model would. However, the model yields a relatively smoothed contour without outliers, showing the general relationship among the three variables. We additionally include two plots in Appendix A, with Figure 8 and Figure 9 showing the *Gap* variable and the *Num* variable in the panels, respectively.

Given that the simulation resolution for the *Gap* variable is much lower compared to that in the exploratory runs, the effects of high and low P_{max} uncertainties for the *Gap* variable, which indicate different starting timestamps of simulated runs, are not as clearly visible as before. In addition, the P_{max} uncertainty fluctuation is greatest when the *Num* variable is at a low value and becomes smoother as *Num* increases, as can be seen in all three figures. This makes sense since a smaller number of future observations would not be able to constrain the orbit as well, and the results from the MCMC process would be highly variable compared to cases with large amount of RVs at future epochs. In all three figures, the overall P_{max} uncertainty decreases monotonically with increasing values of all three variables. Specifi-

cally, the rate of uncertainty reduction happens fastest when the *Num*, *Gap*, and *Cov* variables are all small, and then gradually slows to an almost static state when only two of the variables have large values. This is evident in all three figures, as there is little change in the color and distribution of the contour in the bottom half of the panels. This suggests that there is a stopping threshold for the three variables when planning observations. Once the threshold is crossed, further delaying the start of an observational campaign, increasing the baseline coverage, or greatly increasing the number of observations or cadence, would not yield a similar amount of P_{max} uncertainty reduction as before, and the value in waiting longer and utilizing more observing resources would reach a point of diminishing returns. Such a threshold can be identified at the "elbow" point of the uncertainty reduction curve in a 2D plot, or the "cliff" in a 3D contour plot, beyond which the gradient of uncertainty drastically reduces.

The location of this stopping threshold can be seen more clearly with the help of partial dependence plots (PDP). A PDP shows the dependence between the model output, in this case P_{max} uncertainty, and the input variable of interest, marginalizing over the values of all other variables. If two variables are selected for a PDP, then the interaction between the two variables and their effects on the model output can be revealed with a 2D contour. Here in Figure 5, we show 2D PDPs between the three variables from the full simulation. The top, middle, and bottom panels show the PDPs for *Num* and *Gap*, *Num* and *Cov*, and *Gap* and *Cov*, respectively. Evidently, P_{max} uncertainty reduces towards the upper right corner as values for both variables increase, but the rate decreases with the magnitude of both variables. In the top and the middle panels, P_{max} experiences a larger range of uncertainty reduction at lower *Gap* and *Cov* values as *Num* increases, and the reduction curve becomes flatter as *Gap* and *Cov* increase towards high values, making observing scenarios with high *Gap* or *Cov* less rewarding. For example, at $Gap = 4 \times P_{max}$ in the top panel, it is apparent that observing with 200 observations is no better than observing with only 80 observations. At $Gap = 5 \times P_{max}$, 40 observations should be sufficient. Similar conclusions can be drawn for the *Cov* variable in the middle panel: for the $Cov = 2 \times P_{max}$ scenario, there is no need to carry out large number of observations, as 60 is enough. In the bottom panel, where the interaction between the *Gap* and the *Cov* variables are displayed, the P_{max} uncertainty reduction behaves similarly, except in a more linear fashion in this parameter space. The stopping threshold, where further increasing the
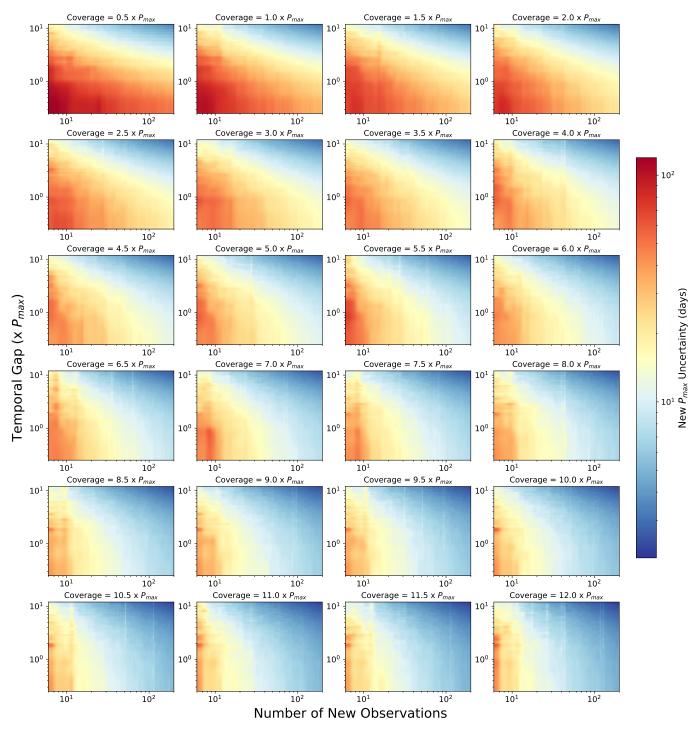
Figure 4. The best-fit random forest model of the full simulation data for HD 134987 c, with the *Num* variable on the horizontal axis, the *Gap* variable on the vertical axis, and the *Cov* variable represented across the panels. The *Gap* and the *Cov* variables are expressed in unit of P_{max}, the orbital period of the outermost planet. Derived P_{max} uncertainties are color coded in each panel, with higher and lower ends of the uncertainty range coded by red and blue, respectively. Fluctuations within the contour are due to the final resolution on both axes and the nature of the random forest model that is completely data informed and therefore produces a relatively rough surface. The P_{max} uncertainty before the simulations is 410 days.
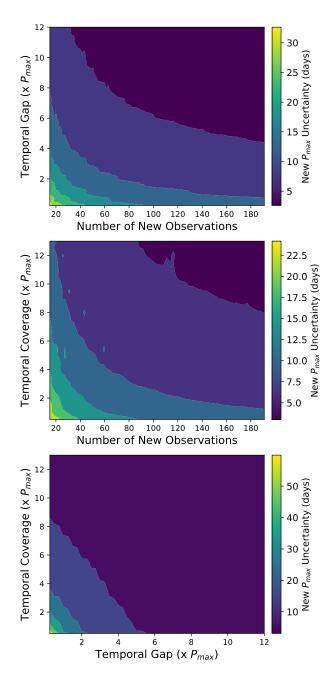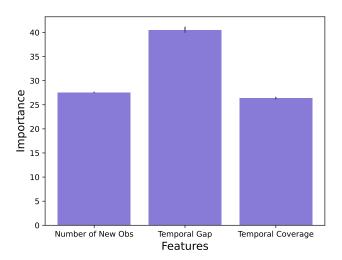
Figure 5. 2D Partial Dependence Plots (PDP) of the three variables from the full simulation. A PDP with two variables shows the mutual variable interaction and effects on the model output while marginalizing other variables. Top panel: *Gap* vs *Num*; middle panel: *Cov* vs *Num*; bottom panel: *Cov* vs *Gap*. Darker color represents lower P_{max} uncertainties values. The P_{max} uncertainty before the simulations is 410 days.

variable yield little additional value, is more obvious: beyond $Gap = 6 \times P_{max}$, for example, increasing the temporal gap gains little, and the P_{max} uncertainty becomes insensitive to *Cov* variables at all.

In this particular case, the P_{max} uncertainty reduction does not level off until the *Gap* and *Cov* variables reach large values. Given the long period of the HD 134987 c planet (~5000 days), the concern for diminishing returns for these two variables is practically irrelevant, as it is highly unlikely that the next RV monitoring campaign would start nearly a century later or last for a similar duration. However, this might not hold true for other targets with shorter periods. It is important to highlight that the threshold for reaching the uncertainty plateau for the *Num* variable is a key factor to consider when planning observations, as one should aim for the most efficient observing strategy and avoid oversubscribing valuable time on precision RV instruments unnecessarily.

Another way to examine each variable's impact on our model is to look at the importance of each variable. Here, we calculate the feature importance for *Num*, *Gap*, and *Cov* within the selected best-fit model using permutation feature importance. This model inspection technique measures the importance of each variable to the overall model performance by randomly shuffling data of the variables under inspection to break relationships between target and variables. If the permutation result shows a large decrease in model performance for a particular variable compared to a baseline value, then it suggests the model is highly dependent on that variable and the variable has high importance within the model. The higher the feature importance value of a variable, the greater its impact within the model. This technique is more robust than other methods, such as impurity-based feature importance, as its calculation is model-agnostic and not biased toward features with high cardinality. The feature importance results of our best-fit model are shown in Figure 6. Each bar represents one of the variables, with the standard deviation of each importance value indicated by a vertical line at the top of each bar. The figure suggests that all three variables contribute meaningfully to the model, given the non-negligible feature importance values, which is reassuring. Furthermore, it appears that the *Gap* variable has the greatest importance, while the *Num* and *Cov* variables have comparable importance. Because of this, the *Gap* variable possesses the highest predictive power within our model and has the highest influence on the model performance, followed by *Num* and *Cov* variable. For instance, the model indicates that, on average, holding *Num* and *Cov* constant while varying *Gap* leads to a larger reduction in P_{max} uncertainty than other scenarios where *Cov* is varied while holding *Num* or *Gap* constant, or *Num* is varied with *Gap* and *Cov* fixed. This finding indicates that, all else being equal, delaying fu-

Figure 6. Feature importance of the three variables, *Num*, *Gap*, and *Cov* variables in the best-fit random forest model on the full simulation data. Each bar represents the mean importance from all permutations carried out, with the vertical line at the top of each bar showing the standard deviation of the importance for that variable/feature. The higher the bar is, the more importance that variable is to our model.

ture observations to increase the temporal gap between the last observation and the upcoming campaign is relatively more effective in reducing P_{max} uncertainties.

4. ALTERNATIVE SIMULATIONS

Of course, knowing the optimal observing strategy and the corresponding expected P_{max} uncertainty reduction performance before an observing campaign is of particular importance to observers. The full simulation in earlier sections provides valuable insights into P_{max} uncertainty behavior concerning all three variables simultaneously, as well as the relative importance of each variable. However, such a full simulation is computationally expensive and typically impractical to conduct for all targets before the campaign begins. In addition, the full simulation does not take into account realistic constraints, such as future mission start times before which observations have to be taken. In this section, we therefore offer examples of shorter simulations with realistic constraints on the start and end times of future observations in the context of upcoming space-based direct imaging missions, such as the Roman telescope.

4.1. *Simulation Objects*

For the short simulations, we selected three targets that are currently high-priority candidates for the Ro-

man CGI[1]: HD 134987, the same target as before (Jones et al. 2010); 47 UMa, a 3-planet system popular in the direct imaging community (Butler & Marcy 1996; Fischer et al. 2002; Gregory & Fischer 2010); and 14 Her, a 2-planet system with one very long period planet (Butler et al. 2003; Wittenmyer et al. 2007; Bardalez Gagliuffi et al. 2021; Rosenthal et al. 2021). Similar to the previous simulations, estimates of orbital parameters are required to create synthetic RVs. Instead of using parameters from a single prior publication, we compiled all publicly available RVs from various sources and modeled them together to derive the most up-to-date orbital parameters for all three targets using RadVel. The most current orbital parameters for the HD 134987, 47 UMa, and 14 Her systems are provided in Table 2.

4.2. *Simulations and Result*

Although 47 UMa has three planets of high interest to the direct imaging community, we focus only on the outer most planet, d, for the observing strategy simulation for P_{max} reduction, and apply the same focus to the other two systems as well. The synthetic RV generation follows the same method as before, but the simulation setup is slightly modified. Instead of creating a full grid of simulations based on the three variables *Num*, *Gap*, and *Cov* to test uncertainty changes without enforcing a fixed starting and ending time for the observing campaign at future epochs, these shorter simulations were all set to end at the same future date. This design mimics the need to complete precursor observations before future mission begins. The primary difference between these short simulations is their varying starting times. For each target, we conducted eight different runs under the scenario of precursor RV observations for Roman targets, with all runs ending around February 1st, 2028, assuming a launch date of May 2027 and accounting for the engineering time before the Roman's science observations commence. The first run begins on February 1st, 2024, with subsequent runs starting 6 months after the previous run's start date. All other simulation setup details remained consistent with prior simulations. Because of the fixed ending time, each successive run has a larger *Gap* and smaller *Cov* value than the previous runs. However, given the short observing baseline and large P_{max} values in this case, these differences are quite small. Details of the eight runs for the three targets are

---

[1] The complete high-priority target list for Roman CGI can be found at the Imaging Mission Database: https://plandb.sioslab.com/index.php.

Table 2. Most Up-to-date System Parameters of HD 134987, 47UMa, and 14 Her.

System	Planet	P (days)	e	ω (deg)	T_p (BJD)	K (m s^{-1})	a (au)	$M_p\sin i$ ($M_{\rm J}$)
HD 134987[a]	b	258.25 ± 0.025	0.2318 ± 0.0051	355.2 ± 1.4	$2,459,110.96 \pm 0.96$	50.03 ± 0.29	0.818 ± 0.012	$1.62^{+0.047}_{-0.048}$
	c	6348^{+71}_{-53}	$0.099^{+0.031}_{-0.032}$	291^{+17}_{-14}	$2,458,512^{+420}_{-240}$	11.03 ± 0.29	6.92 ± 0.11	1.061 ± 0.042
47 UMa[b]	b	$1075.61^{+0.79}_{-0.67}$	$0.0312^{+0.0071}_{-0.0073}$	335^{+15}_{-14}	$2,458,379^{+46}_{-44}$	$47.31^{+0.50}_{-0.46}$	$2.06^{+0.032}_{-0.033}$	2.395 ± 0.079
	c	2290 ± 11	$0.255^{+0.057}_{-0.070}$	86 ± 11	$2,457,882^{+64}_{-58}$	7.61 ± 0.45	$3.407^{+0.054}_{-0.055}$	$0.478^{+0.031}_{-0.030}$
	d	16288^{+950}_{-340}	$0.376^{+0.075}_{-0.081}$	66^{+13}_{-12}	$2,451,397^{+280}_{-370}$	$11.87^{+1.00}_{-0.85}$	$12.64^{+0.49}_{-0.34}$	$1.38^{+0.16}_{-0.13}$
14 Her[c]	b	$1766.71^{+0.58}_{-0.59}$	$0.3652^{+0.0035}_{-0.0034}$	$22.60^{+0.57}_{-0.56}$	$2,456,673.4^{+2.3}_{-2.3}$	89.47 ± 0.42	$2.835^{+0.040}_{-0.041}$	4.86 ± 0.14
	c	19751^{+4700}_{-590}	$0.372^{+0.058}_{-0.026}$	$21.3^{+3.3}_{-1.5}$	$2,451,427^{+160}_{-140}$	$43.8^{+2.7}_{-1.8}$	$14.26^{+2.1}_{-0.52}$	$5.34^{+0.53}_{-0.31}$

NOTE—The ω values in each row of the table are those of the star, not of the planets. Stellar mass for each system was taken from Rosenthal et al. (2021).

[a] Stellar mass $M_* = 1.093 \pm 0.047\ M_\odot$. RVs included: HARPS (pre- and post-upgrade; Trifonov et al. 2020), HIRES (pre- and post-upgrade; Rosenthal et al. 2021), and UCLES (Jones et al. 2010). Current best estimate for planet c's period from NASA Exoplanet Archive is 5960^{+170}_{-150} days from Rosenthal et al. (2021). A separate work by Wittenmyer et al. (2020) reports a period of 5358 ± 31 days for planet c.

[b] Stellar mass $M_* = 1.005 \pm 0.047\ M_\odot$. RVs included: APF-Levy (Rosenthal et al. 2021), ELODIE (Naef et al. 2004), HIRES (post-upgrade; Rosenthal et al. 2021), HJS-Coude (Wittenmyer et al. 2009), HET-HRS (Wittenmyer et al. 2009), and Lick-Hamilton (Rosenthal et al. 2021). Current best estimate for planet d's period from NASA Exoplanet Archive is 14002^{+4018}_{-5095} days from Gregory & Fischer (2010).

[c] Stellar mass $M_* = 0.969 \pm 0.042\ M_\odot$. RVs included: APF-Levy (Rosenthal et al. 2021), ELODIE (Naef et al. 2004), HIRES (pre- and post-upgrade; Rosenthal et al. 2021), HJS-Coude (Wittenmyer et al. 2009). Current best estimate for planet c's period from NASA Exoplanet Archive is 15732^{+2654}_{-1896} days from Feng et al. (2022).

summarized in Table 3 in Appendix B. Once the simulations were complete, we compiled the results from all the runs together for each target and created a best-fit smoothed model around each run's data along with a confidence interval, similar to what was done in the left panels of Figure 3. To fully capture the P_{max} uncertainty behavior, including the initial rapid decline and other nonlinear features, we used the Generalized Additive Model (GAM) to model the short simulation results.

GAM is a semi-parametric model based on the Generalized Linear Model, which models the relationship between the target, or dependent variable Y and one or more predictor, or independent variables x_n by allowing for nonlinear relationships. GAM takes the form of:

$$g(E(Y)) = \beta_0 + f_1(x_1) + f_2(x_2) + ... + f_n(x_n) \quad (2)$$

where $E(Y)$ is the expected value of Y, $g()$ is the link function relating the expected value of the target variable to predictor variables, β_0 is the intercept term; and $f_n()$ represents feature functions of the predictor variables, which can take on parametric or non-parametric forms. In this case, the feature functions used here are penalized B splines that would allow us to model nonlinear relationships without specifying a particular functional form. GAM additively combines the effects of

individual smoothing functions and includes regularization to help prevent overfitting. This flexibility and ability to model highly nonlinear relationship makes GAM an excellent tool for data with unknown degree of nonlinearity. Here, we used the PyGAM package (Servén & Brummitt 2018) to model the short simulation results. For the spline terms, a penalty on the second derivative was applied to regularize the model and produce smoother results. The target variable Y must be specified from a family of exponential distributions; in our case, a gamma distribution with a log transformation link function was chosen. A grid search was performed to find the best penalizing values as well as the optimal number of splines for the spline terms using generalized cross validation (GCV). The smoothing penalty was varied across 30 values spaced logarithmically between -2 and 2, with larger values imposing stronger penalties. The number of splines was varied between 5 and 30 in steps of 5, with higher numbers able to fit more complex structures but also increases the risk of overfitting. The grid search selects the model with the lowest GCV score as the best-fit model for each run. We present the modeling results for all eight runs of all three targets in Figure 7.
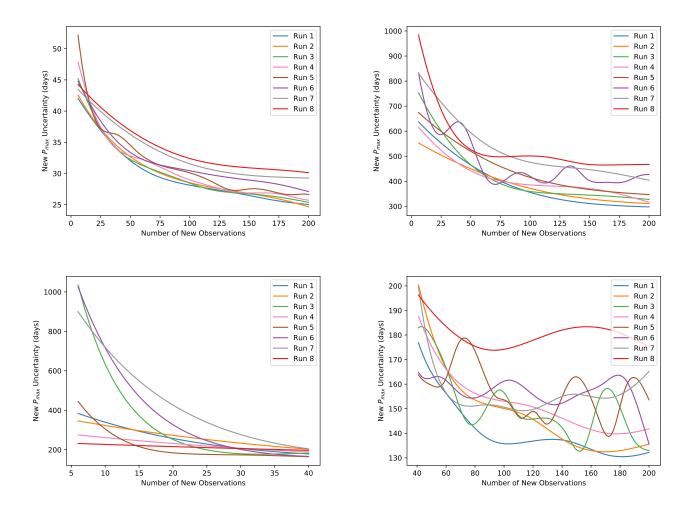
Figure 7. Short simulation results for HD 134987 (top left), 14 Her (top right), and 47 UMa (bottom panels) using GAM. All eight runs for each target are included with confidence intervals omitted to avoid figures being too cluttered. Results for 47 UMa was split into left ($Num \leq 40$) and right ($Num > 40$) panels to reveal differences between runs at larger Num values. The P_{max} uncertainties before the simulations for HD 134987, 14 Her, and 47 UMa are 62, 2600, and 640 days, respectively.

In all runs for the three targets, the P_{max} uncertainty measurement improves as expected. However, the rate of reduction is less drastic compared to the exploratory or full simulations due to the much shorter temporal baseline. The P_{max} uncertainty behaviors differ across the targets due to variations in the original datasets (e.g., total previous baseline, number of previous observations) and the unique characteristics of each planet. For instance, the uncertainty reduction effect for HD 134987 c is not as significant as for 47 UMa d, or 14 Her c, since HD 134987 c has a relatively shorter orbit and therefore its orbital uncertainty is already fairly well constrained. Several key details can be seen from the figures. First, as in the exploratory and full simulations, the uncertainty curve for each run flattens at larger Num values, indicating that additional observations may not necessarily add much more values in con-

straining the period. For example, in all cases, there is little gain in reducing P_{max} uncertainty once the number of new observations exceeds 100. Second, due to the data-driven nature of the spline terms, fluctuations in data will be reflected in the model. Such variations, especially at higher Num values as seen in Run 6 of 14 Her (in the upper right panel of Figure 7) and runs of 47 UMa (in the lower right panel of Figure 7), do not come from sampling along different parts of the phase curve as we saw in Figure 3, but rather arise from the relatively noisier MCMC solutions. This is likely due to the complexity of the planetary system. For example, the three-planet 47 UMa system exhibited higher level fluctuations than the two-planet 14 Her system. This suggests model unpredictability at larger Num values for more complex systems, where P_{max} uncertainty has already reached the performance plateau, further sug-

gesting limited value in additional observations in this case. Finally, for targets with results that do not exhibit high level of uncertainty fluctuations, it can be seen that runs with earlier start dates tend to outperform those with later start dates. Runs that began their observations sooner, with smaller *Gap* values, generally show better performance than those that started later. This trend holds even for the noisier case of 47 UMa, though the performance margin between the best and the worst runs vary by target and dataset characteristics. This result suggests that, for shorter simulations with a fixed end time, beginning observations as early as possible improves performance by increasing the temporal coverage of future data and minimizing the temporal gap between old and new data. It is worth noting that this conclusion does not contradict results from the full simulations, where the *Gap* variable was the most influential factor under the assumption that adjusting the *Gap* or *Cov* variables does not reduce the total temporal baseline (*Gap* plus *Cov*). However, we do not have such an assumption here for the short simulations due to the constraint of having a fixed observing end time. A later start date would result in a shorter total temporal baseline, resulting in poorer performance.

4.3. *An RV Data and Orbits Repository for Direct Imaging*

To address the growing need for the most refined orbital parameters of direct imaging candidates from RV observations for purposes such as mission planning and RV observing simulations presented in this paper, we are establishing a Github repository. Our goal is to host all publicly available RV data for relevant direct imaging candidates, along with fitted RV models, providing the community with the most up-to-date orbital parameters estimates. Many publications model and publish only the RVs observed within their specific observing programs, resulting in partial updates on a target's orbital parameters. This practice often misses the opportunity to provide the most refined parameters by not incorporating all previously available data. Our repository aims to offer the most comprehensive RV models of direct imaging targets with all publicly available data aggregated in one place. We plan to automate the modeling process using RadVel, employing a shell script to organize and concatenate all input data before feeding it into the modeling setup file. This will expedite follow-up modeling when new data are added to the repository and ensure that all data are modeled in a consistent manner. The repository is currently in its early stages, and collecting all data and performing RV modeling on all the targets will take time. Currently, we plan to include

all targets on the Roman CGI candidate list and, if resources allow, incorporate astrometry data and expand the repository to include targets for missions such as the Habitable Worlds Observatory, or even non-imaging missions. We welcome community members interested in joining and contributing to this effort. All data and the latest RV models for the three targets discussed in Section 4.1 are already included in the repository.

5. DISCUSSION

We performed simulations to explore the orbital period uncertainty behavior of the outermost planet in a system, focusing on the influence of three RV observational variables: the number of future observations, the temporal gap between the old and the new data, and the temporal coverage of the new observations. Synthetic RVs were generated to mimic real-world data, including observational white noise and uneven temporal sampling. However, some key observational constraints were not included, and we briefly discuss their potential effects on our simulations here.

Temporally correlated red noise can significantly contribute to the total error budget for RV observations if the stars in question exhibits strong intrinsic stellar variability. Mitigating correlated noise is an active area of research within the RV community, particularly for the EPRV observations aimed at detecting Earth-like planets with small RV signals (Dumusque et al. 2011; Fischer et al. 2016; Hall et al. 2018; Chaplin et al. 2019; Crass et al. 2021; Luhn et al. 2023; Gupta & Bedell 2024). However, for our simulations, which focus on long-period, high-amplitude gas giant planets orbiting mature, non-active stars, short-term low-amplitude stellar activity signals are less relevant here, and the treatment of correlated noise is generally unnecessary. We refer readers to EPRV studies for more on the impact of correlated noise on RV observations. We expect that the exclusion of red noise in our simulated data may slightly underestimate the derived errors of the orbital parameters and therefore the P_{max} uncertainty curves or contours would be slightly inflated or have a less dramatic initial decline. However, this would not affect the overall results.

Seasonality is another common issue for observations, as targets may be unobservable in the night sky for part of the year. The loss of observing opportunities due to target visibility will effect uncertainty outcomes. In our simulations, we assumed a simplified approach where observations are not impacted by target visibility. This assumption was made to focus on the effects of the three main variables *Num*, *Gap*, and *Cov*, without introducing another confounding variable that could overcomplicate

the model. We expect that excluding seasonality may slightly underestimate derived parameter errors. This is especially true if the target has an orbit similar to the high eccentricity case of Figure 3, in which case, missing the optimal observing time due to seasonal visibility may have a significant impact on the uncertainty reduction. In these cases, we recommend conducting observations at the next best time, such as when Gap is around $0.90 \times P_{max}$ of the HD 219077 case.

Disruption to observing programs due to weather, forest fire, instrument failures, or interruptions from other observing programs could negatively impact an observing campaign. Since our simulations did not assume a specific observing facility, we did not account for annual average weather conditions at any particular observatory. The randomness we introduced into the temporal sampling of our synthetic RV data partially mimics the weather unpredictability, though its effect may vary based on actual conditions at a given location. We do not expect this to have a noticeable impact on our simulation results.

6. CONCLUSIONS

Long-period gas giant planets that are potential candidates for future direct imaging missions, such as Roman, typically have poorly constrained orbital parameters due to the long observational baseline required to fully characterize their orbits. The large uncertainties associated with these parameters will have an impact on the quality of future imaging missions, which will depend on the available orbital information extracted from existing RV data. Given the limited time available on precision RV facilities, there is therefore an urgent need for the RV community to plan an efficient precursor observing campaign to provide the best constraints on orbital parameters of candidates before the start of these imaging missions. In this work, we explored the effects of the Num, Gap, and Cov variables on the behavior of P_{max} uncertainty through a suite of full simulations and provided optimal observing solutions under realistic constraints of these three variables through a set of short simulations. We summarize our findings below:

1. For the purpose of refining orbits of cold gas giant planets that are candidates of future direct imaging missions, the use of EPRV facilities is unnecessary and the use of RV facilities with precision on the order of meter-per-second suffices.

2. For future observations with partial temporal coverages ($Cov < 1 \times P_{max}$), the start time of an observing campaign, determined by the Gap variable, is especially important since the RV phases covered by observations influence the extent of P_{max} uncertainty reduction. Specifically, future RVs that cover the phase where the RV curve's gradient is steepest provide better P_{max} uncertainty reduction compared to observations along other RV phases.

3. All three variables have significant contributions to P_{max} uncertainty reduction, with the Gap and Cov variables having the most and least predictive power, respectively, on the model's performance. On average, the Gap variable has the largest influence on the P_{max} uncertainty model performance, and fixing Num and Cov while varying Gap tends to result in the greatest reduction in P_{max} uncertainty. With all else being equal, it is preferable to delay the start of future observations, allowing for an increased temporal gap between the previous and upcoming campaigns, which could deliver better uncertainty reduction results, given that there is no strict end time for future observations. This would be the case for low priority stars that are not part of the target lists of any future missions. Removing them from the observing queue and saving time for higher priority targets would not necessarily harm their orbit refinement progress.

4. If the future observing campaign is constrained by the strict program end time, it can achieve a better P_{max} uncertainty reduction by beginning observations as soon as possible, rather than delaying as in cases without a hard end time. Within the given observing window, increasing the temporal coverage of future observations as much as possible enhances P_{max} uncertainty reduction performance. This would be the case for precursor observations for future direct imaging missions where orbit refinement needs to be performed before mission target lists can be finalized. Therefore, such observations should be assigned with high priorities and commenced as early as possible. In this case, when the total temporal baseline (including all old and new data) was fixed, smaller Gap and therefore larger Cov on average contributed to better P_{max} uncertainty performance. This highlights Gap and Cov being crucial variables, and how Gap can affect model performance differently under different conditions.

5. Obtaining too many observations within a limited time frame, delaying future observations for too long, or conducting an observing campaign for too long is unnecessary. The P_{max} uncertainty eventually plateaus past a certain stopping threshold,

where further increases in these variables do not significantly improve uncertainty reduction. One should look for the "elbow" point of the uncertainty curve beyond which uncertainty decrease starts to level off. Careful observational planning is necessary to avoid oversubscribing valuable RV facilities for observations that offer limited gains in value.

6. We are developing a Github repository[2] aimed at housing all publicly available RV data and their latest models to provide the best orbital parameter estimates for direct imaging candidates. Currently, priority is given to Roman targets, with plans to expand the repository to include other high-priority targets for various missions. The code for the simulations presented in this paper are available as well[3].

Refinement of orbital period uncertainties for long-period planets within known planetary systems that will become direct imaging candidates is a critical task that cannot be overlooked. The orbital information provided by RV data not only affects the precision of predicted planetary locations at future epochs, but also influences the scientific yield of DI, which heavily depends on RV measurement precision. A good understanding of how future RV observations could enhance orbital parameter measurement precision, and to what extent each underlying variable impacts will be crucial to the planning of efficient ground-based RV follow-up campaigns without oversubscribing the valuable telescope time. Designing target follow-up campaigns, collecting data, modeling, and maintaining results in a consistent way to provide the most refined orbital parameter estimates requires a community effort and should begin as early as possible to maximize the science return of future DI missions.

REFERENCES

Akeson, R., Armus, L., Bachelet, E., et al. 2019, arXiv e-prints, arXiv:1902.05569, doi: 10.48550/arXiv.1902.05569

Bailey, V. P., Bendek, E., Monacelli, B., et al. 2023, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 12680, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 126800T, doi: 10.1117/12.2679036

Baluev, R. V. 2008, MNRAS, 389, 1375, doi: 10.1111/j.1365-2966.2008.13656.x

—. 2009, MNRAS, 393, 969, doi: 10.1111/j.1365-2966.2008.14217.x

Bardalez Gagliuffi, D. C., Faherty, J. K., Li, Y., et al. 2021, ApJL, 922, L43, doi: 10.3847/2041-8213/ac382c

Brown, R. A. 2015, ApJ, 799, 87, doi: 10.1088/0004-637X/799/1/87

Burt, J., Holden, B., Wolfgang, A., & Bouma, L. G. 2018, AJ, 156, 255, doi: 10.3847/1538-3881/aae697

Butler, R. P., & Marcy, G. W. 1996, ApJL, 464, L153, doi: 10.1086/310102

Butler, R. P., Marcy, G. W., Vogt, S. S., et al. 2003, ApJ, 582, 455, doi: 10.1086/344570

Chaplin, W. J., Cegla, H. M., Watson, C. A., Davies, G. R., & Ball, W. H. 2019, AJ, 157, 163, doi: 10.3847/1538-3881/ab0c01

Cloutier, R., Doyon, R., Bouchy, F., & Hébrard, G. 2018, AJ, 156, 82, doi: 10.3847/1538-3881/aacea9

Crass, J., Gaudi, B. S., Leifer, S., et al. 2021, arXiv e-prints, arXiv:2107.14291, doi: 10.48550/arXiv.2107.14291

Dalba, P. A., Kane, S. R., Howell, S. B., et al. 2021, AJ, 161, 123, doi: 10.3847/1538-3881/abd6ed

---

[2] The Github repository for storing all RV data and modeling of imaging targets can be accessed here: https://github.com/sblunt/roman-orbits.

[3] The Github repository for code used to run the simulations can be accessed here: https://github.com/zhexingli/RVRefine.

Dragomir, D., Harris, M., Pepper, J., et al. 2020, AJ, 159, 219, doi: 10.3847/1538-3881/ab845d

Dumusque, X., Udry, S., Lovis, C., Santos, N. C., & Monteiro, M. J. P. F. G. 2011, A&A, 525, A140, doi: 10.1051/0004-6361/201014097

Feng, F., Butler, R. P., Vogt, S. S., et al. 2022, ApJS, 262, 21, doi: 10.3847/1538-4365/ac7e57

Fischer, D. A., Marcy, G. W., Butler, R. P., Laughlin, G., & Vogt, S. S. 2002, ApJ, 564, 1028, doi: 10.1086/324336

Fischer, D. A., Anglada-Escude, G., Arriagada, P., et al. 2016, PASP, 128, 066001, doi: 10.1088/1538-3873/128/964/066001

Fulton, B. J., Petigura, E. A., Blunt, S., & Sinukoff, E. 2018, PASP, 130, 044504, doi: 10.1088/1538-3873/aaaaa8

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1, doi: 10.1051/0004-6361/201833051

Gaudi, B. S., Seager, S., Mennesson, B., et al. 2020, arXiv e-prints, arXiv:2001.06683, doi: 10.48550/arXiv.2001.06683

Gibson, S. R., Howard, A. W., Rider, K., et al. 2024, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 13096, Ground-based and Airborne Instrumentation for Astronomy X, ed. J. J. Bryant, K. Motohara, & J. R. D. Vernet, 1309609, doi: 10.1117/12.3017841

Gregory, P. C., & Fischer, D. A. 2010, MNRAS, 403, 731, doi: 10.1111/j.1365-2966.2009.16233.x

Gupta, A. F., & Bedell, M. 2024, AJ, 168, 29, doi: 10.3847/1538-3881/ad4ce6

Hall, R. D., Thompson, S. J., Handley, W., & Queloz, D. 2018, MNRAS, 479, 2968, doi: 10.1093/mnras/sty1464

Harada, C. K., Dressing, C. D., Kane, S. R., & Ardestani, B. A. 2024a, ApJS, 272, 30, doi: 10.3847/1538-4365/ad3e81

Harada, C. K., Dressing, C. D., Kane, S. R., et al. 2024b, 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

Hill, M. L., Bott, K., Dalba, P. A., et al. 2023, AJ, 165, 34, doi: 10.3847/1538-3881/aca1c0

Hill, M. L., Kane, S. R., Seperuelo Duarte, E., et al. 2018, ApJ, 860, 67, doi: 10.3847/1538-4357/aac384

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

Jones, H. R. A., Butler, R. P., Tinney, C. G., et al. 2010, MNRAS, 403, 1703, doi: 10.1111/j.1365-2966.2009.16232.x

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

Kane, S. R. 2007, MNRAS, 380, 1488, doi: 10.1111/j.1365-2966.2007.12144.x

—. 2013, ApJ, 766, 10, doi: 10.1088/0004-637X/766/1/10

Kane, S. R., & Burt, J. A. 2024, AJ, 168, 279, doi: 10.3847/1538-3881/ad8a68

Kane, S. R., & Gelino, D. M. 2012, PASP, 124, 323, doi: 10.1086/665271

Kane, S. R., Li, Z., Turnbull, M. C., Dressing, C. D., & Harada, C. K. 2024, arXiv e-prints, arXiv:2408.00263, doi: 10.48550/arXiv.2408.00263

Kane, S. R., Mahadevan, S., von Braun, K., Laughlin, G., & Ciardi, D. R. 2009, PASP, 121, 1386, doi: 10.1086/648564

Kane, S. R., Meshkat, T., & Turnbull, M. C. 2018, AJ, 156, 267, doi: 10.3847/1538-3881/aae981

Kane, S. R., Hill, M. L., Kasting, J. F., et al. 2016, ApJ, 830, 1, doi: 10.3847/0004-637X/830/1/1

Kane, S. R., Dalba, P. A., Li, Z., et al. 2019, AJ, 157, 252, doi: 10.3847/1538-3881/ab1ddf

Kane, S. R., Bean, J. L., Campante, T. L., et al. 2021, PASP, 133, 014402, doi: 10.1088/1538-3873/abc610

Kasting, J. F., Whitmire, D. P., & Reynolds, R. T. 1993, Icarus, 101, 108, doi: 10.1006/icar.1993.1010

Kopparapu, R. K., Ramirez, R. M., SchottelKotte, J., et al. 2014, ApJL, 787, L29, doi: 10.1088/2041-8205/787/2/L29

Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, ApJ, 765, 131, doi: 10.1088/0004-637X/765/2/131

Laliotis, K., Burt, J. A., Mamajek, E. E., et al. 2023, AJ, 165, 176, doi: 10.3847/1538-3881/acc067

Lam, C., Bedell, M., Zhao, L. L., Gupta, A. F., & Ballard, S. A. 2024, AJ, 168, 200, doi: 10.3847/1538-3881/ad739b

Li, Z., Hildebrandt, S. R., Kane, S. R., et al. 2021, AJ, 162, 9, doi: 10.3847/1538-3881/abf831

Lucy, L. B., & Sweeney, M. A. 1971, AJ, 76, 544, doi: 10.1086/111159

Luhn, J. K., Ford, E. B., Guo, Z., et al. 2023, AJ, 165, 98, doi: 10.3847/1538-3881/acad08

Marmier, M., Ségransan, D., Udry, S., et al. 2013, A&A, 551, A90, doi: 10.1051/0004-6361/201219639

Meunier, N., Pous, R., Sulis, S., Mary, D., & Lagrange, A. M. 2023, A&A, 676, A82, doi: 10.1051/0004-6361/202346218

Naef, D., Mayor, M., Beuzit, J. L., et al. 2004, A&A, 414, 351, doi: 10.1051/0004-6361:20034091

National Academies of Sciences, Engineering, and Medicine. 2021, Pathways to Discovery in Astronomy and Astrophysics for the 2020s, doi: 10.17226/26141

Newman, P. D., Plavchan, P., Burt, J. A., et al. 2023, AJ, 165, 151, doi: 10.3847/1538-3881/acad07

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Pepe, F., Cristiani, S., Rebolo, R., et al. 2021, A&A, 645, A96, doi: 10.1051/0004-6361/202038306

Riggs, A. J. E., Bailey, V., Moody, D. C., et al. 2021, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 11823, Techniques and Instrumentation for Detection of Exoplanets X, ed. S. B. Shaklan & G. J. Ruane, 118231Y, doi: 10.1117/12.2598599

Rosenthal, L. J., Fulton, B. J., Hirsch, L. A., et al. 2021, ApJS, 255, 8, doi: 10.3847/1538-4365/abe23c

Savransky, D., Bailey, V. P., Wolff, S. G., et al. 2024, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 13092, Space Telescopes and Instrumentation 2024: Optical, Infrared, and Millimeter Wave, ed. L. E. Coyle, S. Matsuura, & M. D. Perrin, 130921I, doi: 10.1117/12.3020514

Schwab, C., Rakich, A., Gong, Q., 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, 99087H, doi: 10.1117/12.2234411

Servén, D., & Brummitt, C. 2018, pyGAM: Generalized Additive Models in Python, doi: 10.5281/zenodo.1208723

Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv e-prints, arXiv:1503.03757. https://arxiv.org/abs/1503.03757

Spohn, C., Savransky, D., & Morgan, R. 2022, AJ, 163, 163, doi: 10.3847/1538-3881/ac5049

Stark, C. C., Roberge, A., Mandell, A., & Robinson, T. D. 2014, ApJ, 795, 122, doi: 10.1088/0004-637X/795/2/122

The LUVOIR Team. 2019, arXiv e-prints, arXiv:1912.06219. https://arxiv.org/abs/1912.06219

Trifonov, T., Tal-Or, L., Zechmeister, M., et al. 2020, A&A, 636, A74, doi: 10.1051/0004-6361/201936686

Turnbull, M. C., Zimmerman, N., Girard, J. H., et al. 2021, Journal of Astronomical Telescopes, Instruments, and Systems, 7, 021218, doi: 10.1117/1.JATIS.7.2.021218

Vaughan, S. R., Gebhard, T. D., Bott, K., et al. 2023, MNRAS, 524, 5477, doi: 10.1093/mnras/stad2127

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

Wes McKinney. 2010, in Proceedings of the 9th Python in Science Conference, ed. Stéfan van der Walt & Jarrod Millman, 56 – 61, doi: 10.25080/Majora-92bf1922-00a

Wittenmyer, R. A., Endl, M., & Cochran, W. D. 2007, ApJ, 654, 625, doi: 10.1086/509110

Wittenmyer, R. A., Endl, M., Cochran, W. D., Levison, H. F., & Henry, G. W. 2009, ApJS, 182, 97, doi: 10.1088/0067-0049/182/1/97

Wittenmyer, R. A., Wang, S., Horner, J., et al. 2020, MNRAS, 492, 377, doi: 10.1093/mnras/stz3436

Wolff, S. G., Wang, J., Stapelfeldt, K., et al. 2024, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 13092, Space Telescopes and Instrumentation 2024: Optical, Infrared, and Millimeter Wave, ed. L. E. Coyle, S. Matsuura, & M. D. Perrin, 1309255, doi: 10.1117/12.3020205
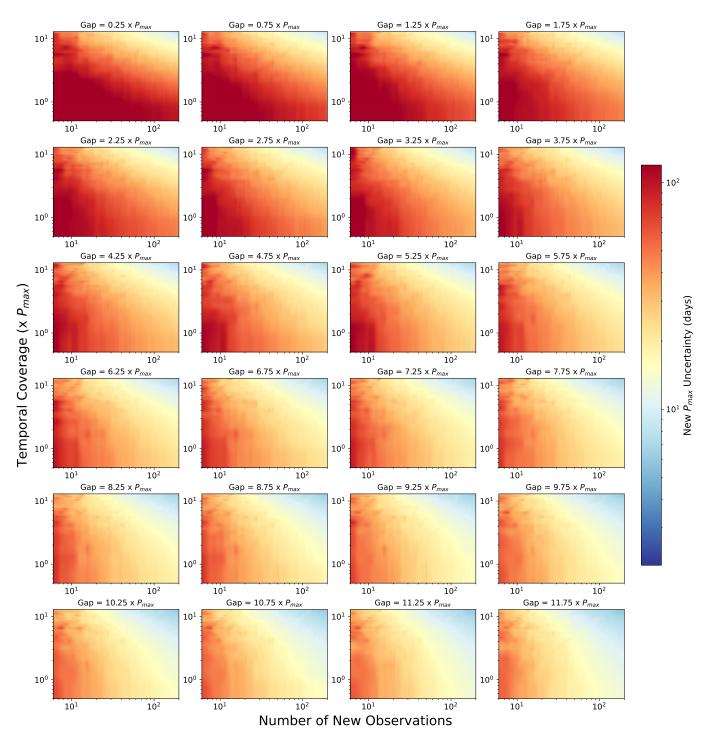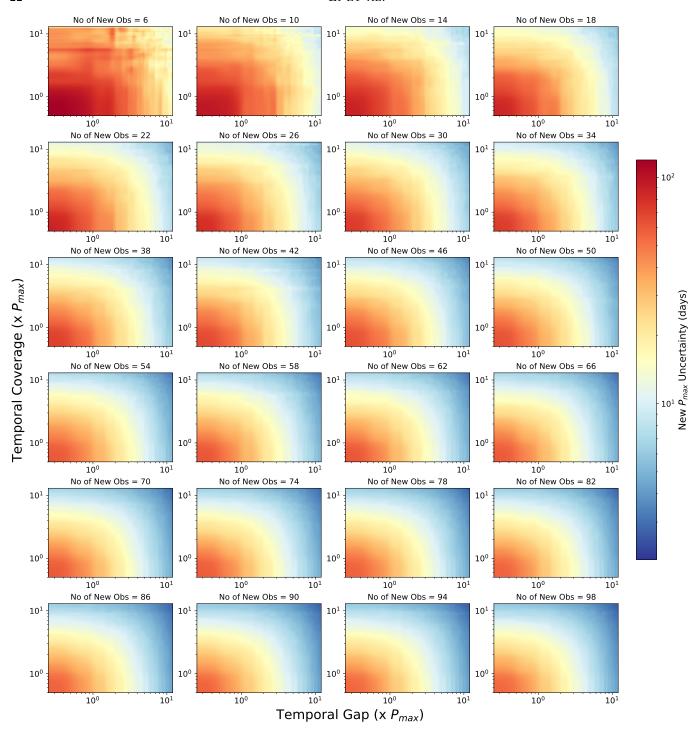
APPENDIX

A. ADDITIONAL FIGURES

Figure 8. Similar to Figure 4, except that the *Cov* variable is now represented on the vertical axis and the *Gap* variable is shown across the panels, with a step of 0.5 xP_{max}.

Figure 9. Similar to Figure 4, except that the *Gap* variable is now represented on the horizontal axis, the *Cov* variable on the vertical axis, and the *Num* variable is represented by the panels, with a step of 4 and up to *Num* = 98. Higher values of *Num* are not shown here as there is little change in the shape and color of the contour.

B. ADDITIONAL TABLES

Table 3. Details of Short Simulation Runs

Target	Run	Starting Time	Gap	Cov
HD 134987[a]	1	2024-02-01	0.2561	0.2302
	2	2024-08-01	0.2848	0.2015
	3	2025-02-01	0.3138	0.1725
	4	2025-08-01	0.3423	0.1440
	5	2026-02-01	0.3713	0.1150
	6	2026-08-01	0.3998	0.0865
	7	2027-02-01	0.4288	0.0575
	8	2027-08-01	0.4573	0.0290
47 UMa[b]	1	2024-02-01	0.0875	0.0898
	2	2024-08-01	0.0987	0.0786
	3	2025-02-01	0.1100	0.0673
	4	2025-08-01	0.1212	0.0562
	5	2026-02-01	0.1325	0.0449
	6	2026-08-01	0.1436	0.0337
	7	2027-02-01	0.1549	0.0224
	8	2027-08-01	0.1660	0.0113
14 Her[c]	1	2024-02-01	0.0720	0.0732
	2	2024-08-01	0.0811	0.0641
	3	2025-02-01	0.0903	0.0549
	4	2025-08-01	0.0994	0.0458
	5	2026-02-01	0.1086	0.0366
	6	2026-08-01	0.1177	0.0275
	7	2027-02-01	0.1269	0.0183
	8	2027-08-01	0.1360	0.0092

NOTE—The Gap and the Cov variables are in units of P_{max}.

[a] Timestamp of last real RV observation at BJD 2,458,715.728759 (2019-08-20).

[b] Timestamp of last real RV observation at BJD 2,458,916.894692 (2020-03-08).

[c] Timestamp of last real RV observation at BJD 2,458,906.164844 (2020-02-26).