Highlights of Spanish Astrophysics XI, Proceedings of the XV Scientific Meeting of the Spanish Astronomical Society held on September 4–9, 2022, in La Laguna, Spain. M. Manteiga, L. Bellot, P. Benavidez, A. de Lorenzo-Cáceres, M. A. Fuente, M. J. Martínez, M. Vázquez- Acosta, C. Dafonte (eds.), 2023

The KOBE experiment - `KOBEsim`: improving RV detection through efficient scheduling.

O. Balsalobre-Ruza[1], J. Lillo-Box[1], and A. Berihuete[2]

[1] Centro de Astrobiología (CAB), CSIC-INTA, Camino Bajo del Castillo s/n, 28692, Villanueva de la Cañada, Madrid, Spain
[2] Depto. Estadística e Investigación Operativa, Universidad de Cádiz, Avda. República Saharaui s/n, 11510 Puerto Real, Cádiz, Spain

Abstract

At present, detecting new rocky planets within the habitable zone and the radial velocity follow-up of transiting candidates are priority objectives of the exoplanetary field. Both, require a great effort including high-precision instruments and state-of-the-art analysis techniques. Additionally, a proper observing strategy is crucial to ensure the effectiveness of the observations, avoiding unnecessary measurements that waste invaluable telescope time. In this talk, we present the `KOBEsim` algorithm, a Bayesian-based strategy for the detection of planets in radial velocity surveys. It is developed within the KOBE experiment, aspiring at maximizing the detection of potential habitable exoplanets orbiting late K-dwarfs. The algorithm uses the first data obtained for a given star to choose a target orbital period (usually the highest power periodicity) and uses Bayesian inference to propose the optimum next observing date, thus accelerating the detection/rejection of such period. This new approach has demonstrated to improve the detection efficiency in comparison with a conventional strategy of monotonic cadence, reaching a detection in ∼50% less observations and timespan. `KOBEsim` has the potential to save expensive telescope time in current and upcoming instruments, and to allow the detection of light planets further away from their host star in reasonable timespans.

1 Introduction

Over the last decades, exoplanetary exploration has been focused on filling the census of exoplanet properties. Currently, the number of confirmed planets according to the NASA Exoplanet Archive overtakes the figure of 5200 [1]. The reason for this rush is that we need statistics. Only if we have a representative sample of the planet diversity we could create reliable models, for instance on planet formation, migrations, or dependencies with the host star properties. But it is undeniable that the planets that most yearn to detect are potentially habitable. Finding life beyond the Solar System is what guides the actions of the field.

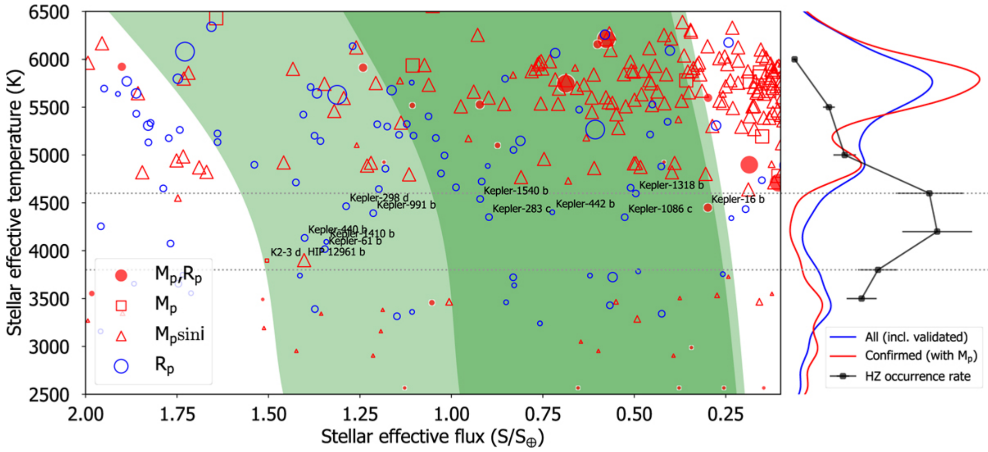

Figure 1: Exoplanet population within the HZ for different stellar types. Source [4].

2 The KOBE experiment

The majority of the efforts dedicated to searching for habitable planets have preferred two spectral types as host stars: M and G-dwarfs. The M-dwarfs are an interesting target from the detectability point of view. Since they are cold and less massive stars, their habitable zone (HZ) is closer and thus inducing higher radial velocity (RV) semiamplitudes and requiring less time to infer their presence. As an example, the high-resolution spectrograph CARMENES [2] is optimized to find planets orbiting M-dwarf stars. In parallel, since the only planet that we are certain to be habitable is the Earth, G-dwarfs (solar-type stars) seem the perfect target from an astrobiological point of view. One of the most fruitful missions, Kepler, was specifically designed for finding Earth-type worlds around solar-type stars [3].

Nonetheless, we are missing what might be the perfect target. Putting the focus on M and G stars has resulted in an observational bias around K-dwarfs (see Fig. 1). This spectral type is in the middle of the other two, thus finding a trade-off between habitability and detectability. Moreover, theoretical studies have shown that the highest occurrence rate of habitable planets is around the quiet late-K-dwarfs [15]. These reasons have inspired to start the first RV survey devoted specifically to late-type K-dwarfs. The K-dwarfs Orbited By habitable Exoplanets (KOBE) [4] experiment, a legacy program of the Calar Alto Observatory (CAHA; Almería, Spain), making use of CARMENES at the 3.5 m telescope. Its observations began in January 2021 and will be monitoring 50 late K-dwarf stars over five consecutive semesters.

The aim of KOBE is to detect a handful of habitable planets around K-type stars to ease the paucity. Reaching such an ambitious goal requires an observational strategy designed accordingly to the needs. `KOBEsim` [5] is an open-source algorithm written in Python language that we have developed to make more efficient the detection of planets in blind-RV surveys.

3 Observational strategy of KOBEsim

The proposed observational strategy consists of two steps. First, the star is monitored with the usual survey strategy until the periodogram shows a predominant periodicity ($P_{\rm peak}$). For example, a conventional strategy would be following a monotonic cadence that consists of gathering data every N days. Second, $P_{\rm peak}$ is targeted by the KOBEsim algorithm to speed up the confirmation or rejection of a planetary origin. The algorithm uses the prior knowledge (the observations gathered so far) to predict future measurements and makes a ranking of the dates available to observe the target again. The criterion for the ranking is based on the expected knowledge gain. In the following sections, we see in more detail the method.

3.1 Statistical framework

The mathematical tool we use is the Bayes factor. Basically, it compares two models to assess from which one is more likely that our data come[1]. We compute this quantity by comparing 1) the null hypothesis in which we assume there is no planet in the system (H_0), and 2) a 1-planet model orbiting with a periodicity near to $P_{\rm peak}$ (H_1). The Bayes factor is frequently used in planetary searches to claim a detection when it is higher than a threshold [6], [7], [8]. In our case, we opt for a conservative criterion considering a detection at $\ln(B_{10}) > 6$ [9].

3.2 Code methodology

The algorithm is composed of three steps: estimation of the orbital parameters, simulation of the next RV measurement at different orbital phases, and selection of the optimum observing date according to the expected increase of the Bayes factor.

1. **Parameter inference and evidence of the models:** In order to infer the orbital parameters from the observations, `KOBEsim` explores the parameter space and samples the posterior distribution by using the Markov chain Monte Carlo (MCMC) affine invariant ensemble sampler `emcee` [12]. The model used for the 1-planet hypothesis is a single Keplerian from the python module `RadVel` [13], meanwhile for the null-hypothesis the RV takes a constant value equal to the systemic velocity. Next, it calculates the Bayes factor metric employing the `bayev` [14] code to quantify how much evidence we have with the current observations on the existence of the planet at the selected period.

2. **Simulation of the future dates:** To find the candidate dates we divide the period under study into a total of $N_{\rm phases}$ orbital sub-phases. We choose the next assigned date from the schedule at the telescope that matches each sub-phase. Next, from the posterior probability distributions inferred through the MCMC algorithm in the previous step, `KOBEsim` predicts the RVs at each potential observing date. Running again `emcee` and `bayev` over each of the datasets (our already gathered RV plus one additional datapoint corresponding to each predicted RV at a proposed date), we end

[1] B_{10} is the notation for a Bayes factor comparing a hypothesis H_1 over another H_0.

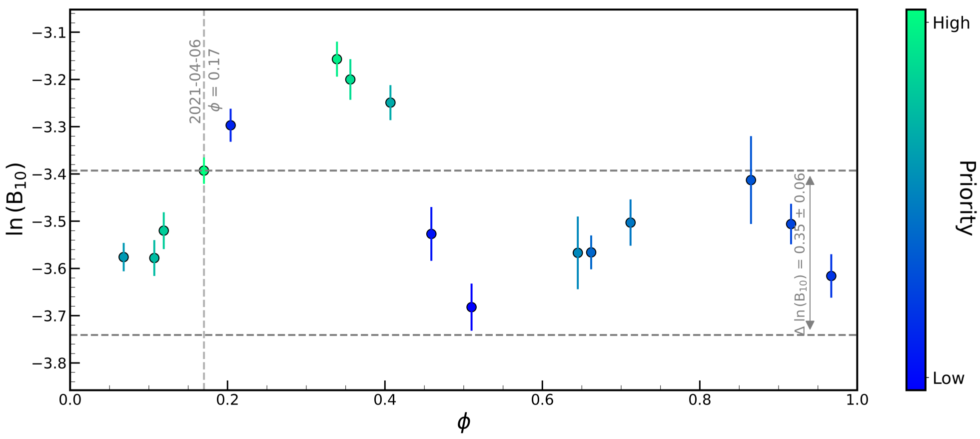

Figure 2: KOBEsim output figure. Predicted Bayes factor for each candidate date (orbital phase). Source [5].

up with an estimation of the expected increase in the Bayes factor ($\Delta \ln(B_{10})$) for each of the proposed future dates.

3. **Ranking of the candidate observing dates:** `KOBEsim` sorts the tested dates according to the Bayes factor giving the maximum priority to the highest $\Delta \ln(B_{10})$. However, the largest $\ln(B_{10})$ increase may occur at a very distant date, which is against the efficiency of the observations. To prevent this situation, we introduce a weight to the increase of the Bayes factor with the shape of a density function of a beta distribution. Hence, the ranking is done according to the Bayes factor weighted with this function to find a trade-off between the number of observations and timespan. In Fig. 2 we show an example of the output figure of `KOBEsim`. Additionally, it delivers a table with these results sorted by the priority of the the candidate dates.

4 `KOBEsim` efficiency test

In this section, we show a test with simulated data to quantify the efficiency improvement when using this strategy. We use ten mock RVs for a $5\,M_\oplus$ planet orbiting within the HZ of a late-K-dwarf with a 59-days period, corresponding to an RV semiamplitude of $1.2\,\mathrm{m\,s^{-1}}$. We estimate how long it would take to detect this planet following different strategies: with a monotonic cadence (MC), following the recommendations of `KOBEsim` (`K`) every new observation, and `KOBEsim` using the beta function ($\mathtt{K}_\beta$, see step 3 of Sect. 3.2).

In Fig. 3 we show the evolution of the Bayes factor (y-axis) as a function of the number of observations (x-axis) for each strategy (color-code). From a quick visual inspection, it is easy to perceive that the number of observations is greatly reduced when using the algorithm in comparison with the MC strategy. It is also highlighting the timespan (see legend at

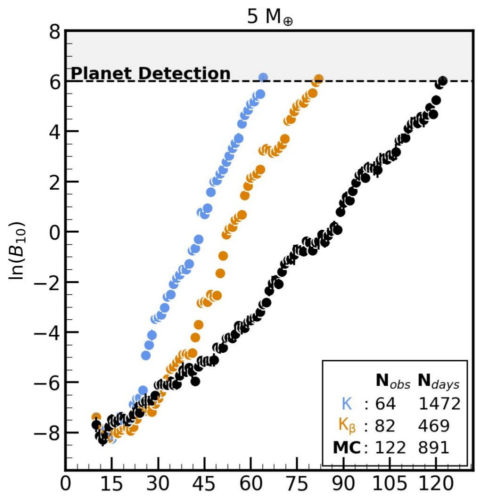

Figure 3: Prediction in the evolution of the logarithm of the Bayes factor for a 5 $M_\oplus$ simulated planet. Source [5].

the right bottom of Fig. 3) is only improved when using the $\mathtt{K}_\beta$ strategy, thus finding the optimum trade-off in sake of efficiency. The improvement is, therefore, a 33 % in the number of observations and a 47 % in timespan.

5 Conclusions

The biggest conclusion is that there is room for improving the efficiency of observations. Reducing the observational time required to achieve our scientific goals means being capable of including additional targets or even achieving goals that otherwise would be inaccessible. In particular, our approach demonstrates speeding up detections up to a 50 % in both number of observations and timespan. The more challenging the target is, the higher the efficiency gain when using `KOBEsim`. Therefore, its use could be decisive to detect rocky planets within the HZ in reasonable timespans.

GTO programs can be highly benefited from the use of this algorithm since they enjoy wider freedom in their schedule. An opportunity for saving time and favor the detection of the most elusive planets for the upcoming generation of instruments such as HARPS3 [16] or NIRPS [17]. We want to note that, even designed for blind-search surveys, it also can be highly useful in the RV follow-up of transiting candidates since the orbital period is clear since the very beginning of the observations. Furthermore, by customizing the code, one could optimize other problems such as discerning between competing periodicities, or the number of planets in the system. Indeed, the strategy could be used to schedule the observation of time series in any other field than exoplanets just by modifying the models.

Acknowledgments

We want to thank to all KOBE team members, which includes (apart from the authors) N. C. Santos, A. Santerne, A. M. Silva, D. Barrado, J. Faria, A. Castro-González, M. Morales-Calderón, A. Saavedra, E. Marfil, S. G. Sousa, V. Adibekyan, S. C. C. Barros, E. Delgado-Mena, N. Huélamo, M. Deleuil, O. D. S. Demangeon, P. Figueira, and S. Grouffal.

References

[1] Akeson, R. L., Chen, X., Ciardi, D., et al. 2013, Publications of the Astronomical Society of the Pacific, 125, 989–999

[2] Quirrenbach, A., CARMENES Consortium, Amado, P. J., et al. 2020, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 11447, Society of Photo-Optical Instrumentation Engineers (SPIE) Confer- ence Series, 114473C

[3] Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977

[4] Lillo-Box, J., Santos, N. C., Santerne, A., et al. 2022, Astronomy & Astrophysics, 667, A102

[5] Balsalobre-Ruza, O., Lillo-Box, J., Berihuete, A., et al. 2022, A&A, accepted [arXiv:2210.11207]

[6] Lillo-Box, J., Figueira, P., Leleu, A., et al. 2020, A&A, 642, A12

[7] Mortier, A., Zapatero Osorio, M. R., Malavolta, L., et al. 2020, MNRAS, 499, 5004

[8] Faria, J. P., Suárez Mascareño, A., Figueira, P., et al. 2022, A&A, 658, A115

[9] Jeffreys, H. 1961, 3rd Edition, Clarendon Press, Oxford

[10] Mackay, D. J. C. 2003, Information Theory, Inference and Learning Algorithms Mayor, M. & Queloz, D. 1995, Nature, 378, 355

[11] Thrane, E. & Talbot, C. 2019, PASA, 36, e010

[12] Foreman-Mackey, D., Conley, A., Meierjurgen Farr, W., et al. 2013, emcee: The MCMC Hammer

[13] Fulton, B. J., Petigura, E. A., Blunt, S., & Sinukoff, E. 2018, PASP, 130, 044504

[14] Díaz, R. F., Ségransan, D., Udry, S., et al. 2016, A&A, 585, A134

[15] Kunimoto, M., & Matthews, J. M. 2020, AJ, 159, 248

[16] Thompson, S. J., Queloz, D., Baraffe, I., 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, 99086F

[17] Wildi, F., Blind, N., Reshetov, V., et al. 2017, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10400, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 1040018