



Detecting targets with radial velocities for their atmospheric characterisation with the next flagship instruments

Nathan Hara
Université de Genève

Sf2A 2023, breakout session *Origin of water and life: What astrophysical and instrumental challenges for tomorrow?*
21 June 2023

Key message

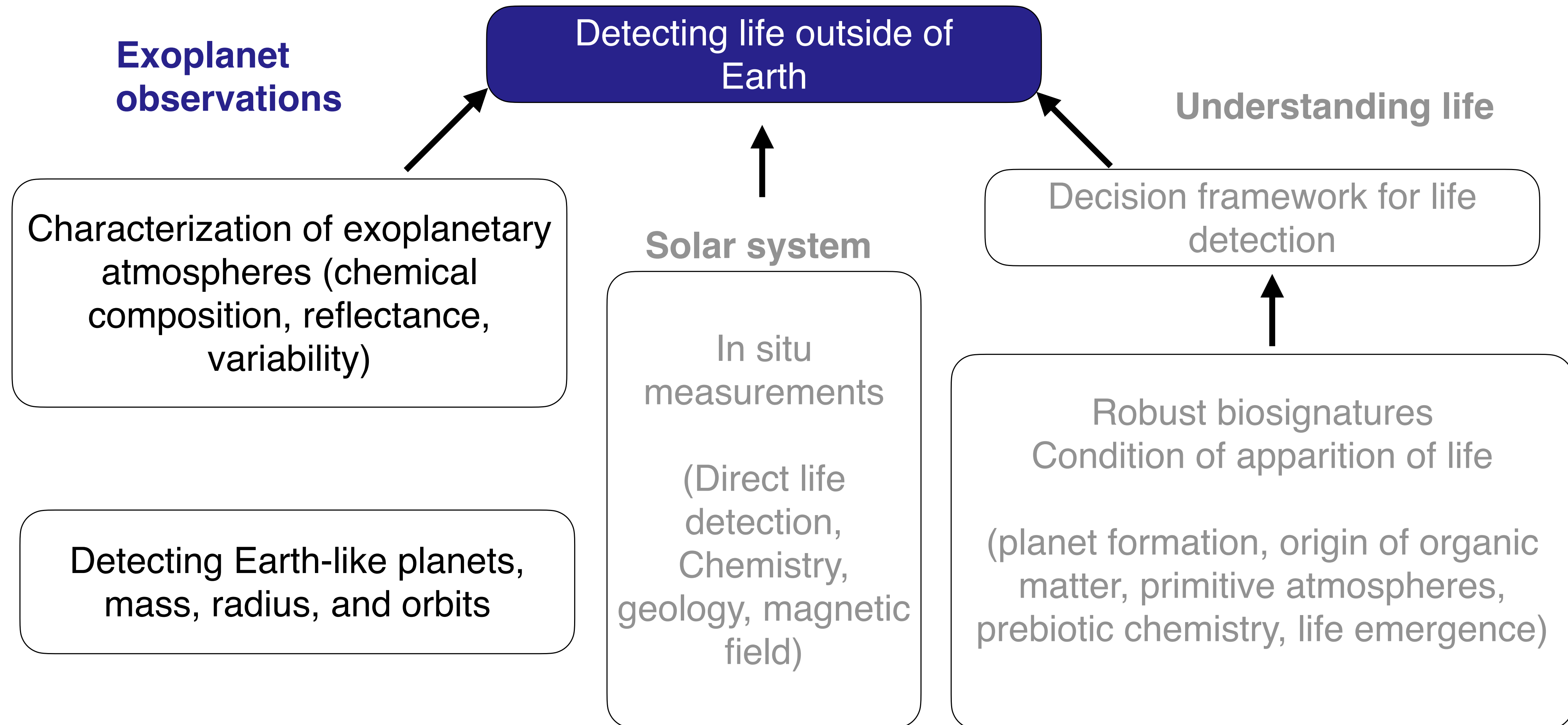
A core objective of flagship ESA and NASA mission at the 2050 horizon is to characterise the atmosphere of Earth-like exoplanets

Radial velocity measurements can pre-detect Earth-like targets and will be essential in any case to measure planetary mass

The radial velocity signal of the Earth on the Sun is 9 cm/s and current instruments are limited at ~ 1 m/s because of complex noises

Reaching 9 cm/s precision is the primary goal of the radial velocity community, it is largely a data analysis problem, and there has been recent progress

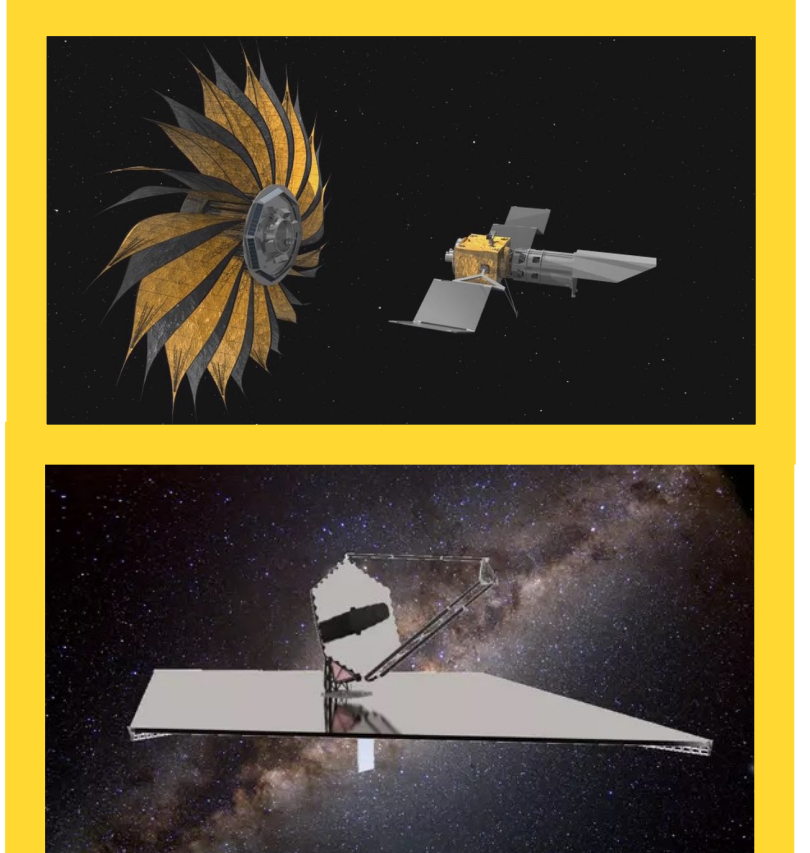
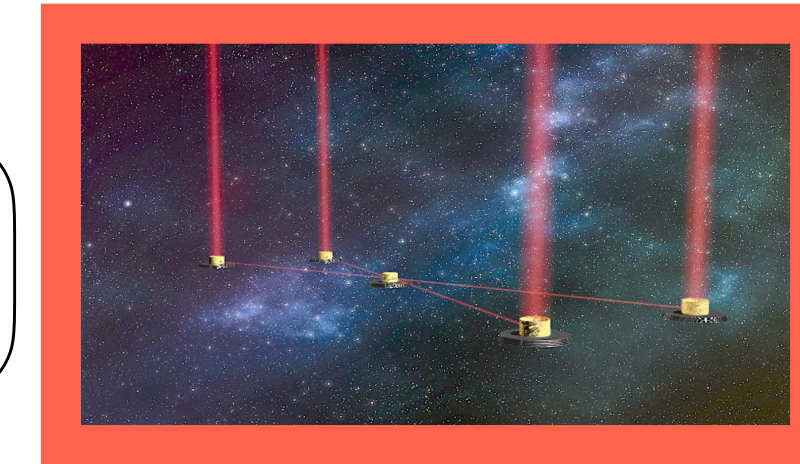
Objectives



Objectives

Exoplanet observations

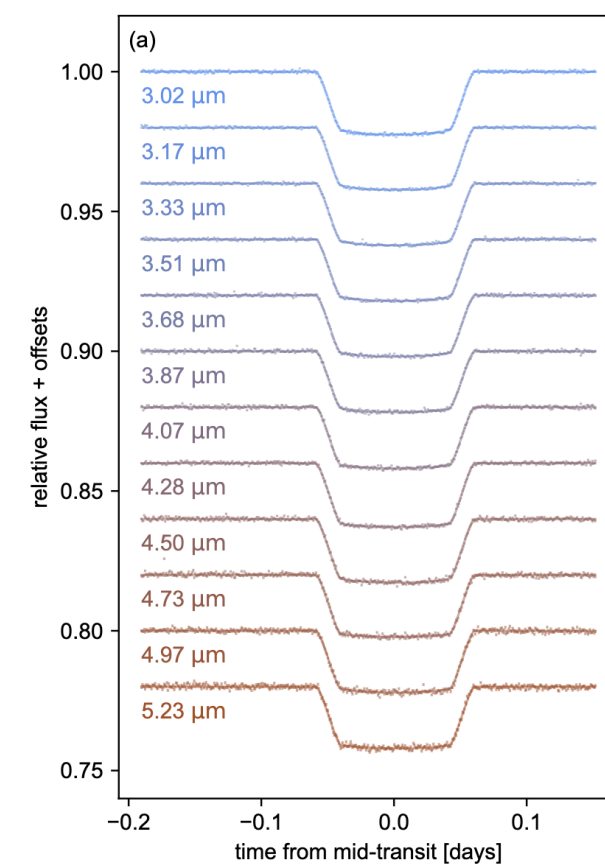
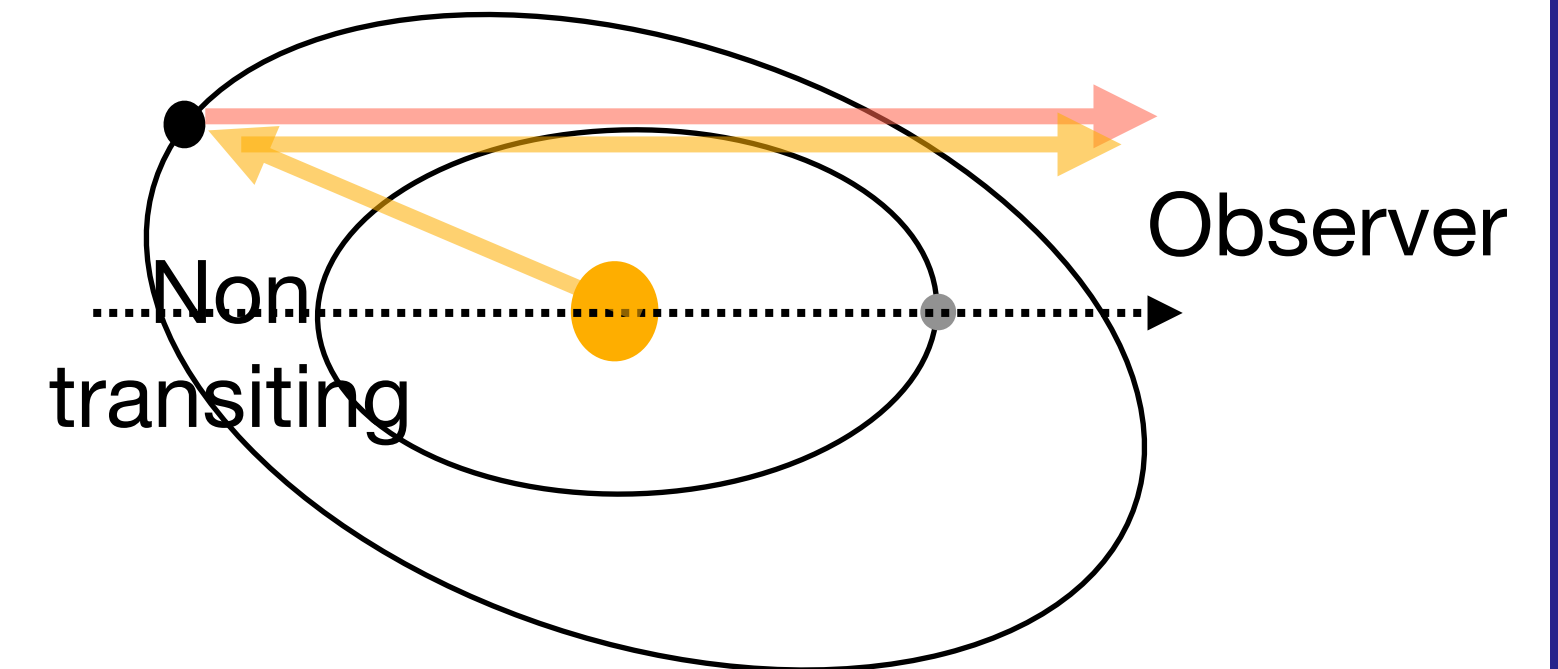
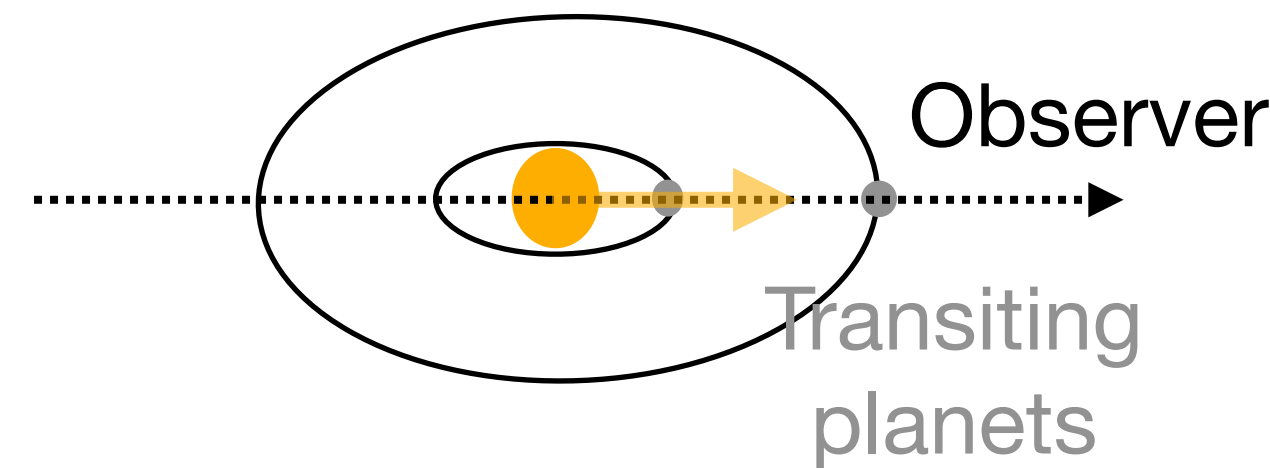
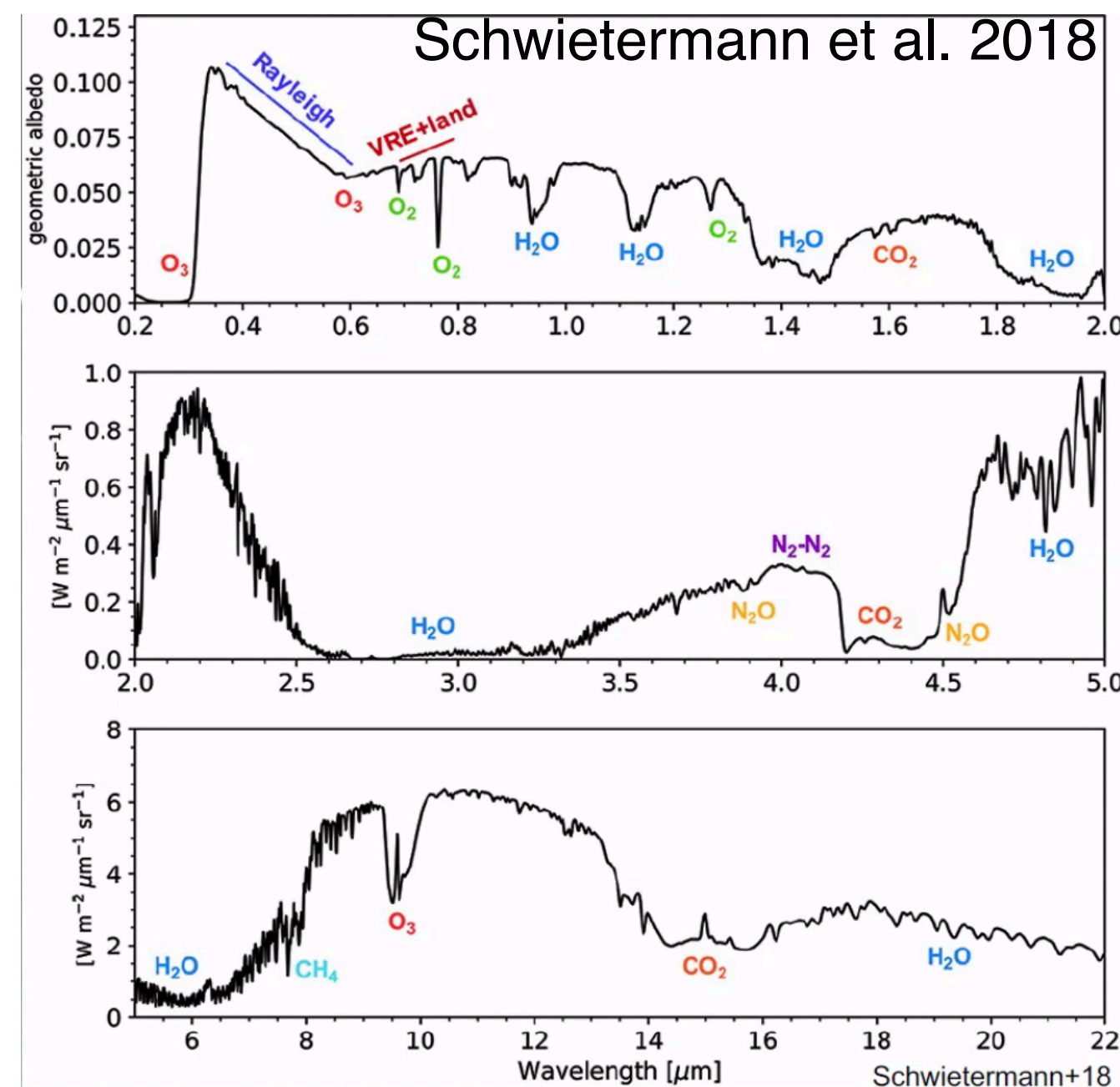
Characterizing Earth-analogs
Detecting life outside of Earth



Characterization of exoplanetary atmospheres (chemical composition, reflectance, variability)

Transmission spectroscopy:
Ariel, ANDES@ELT,
ESPRESSO

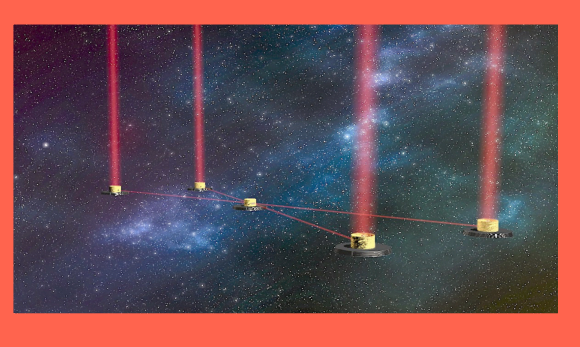
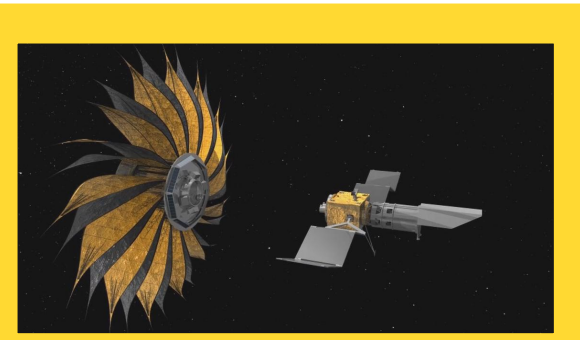
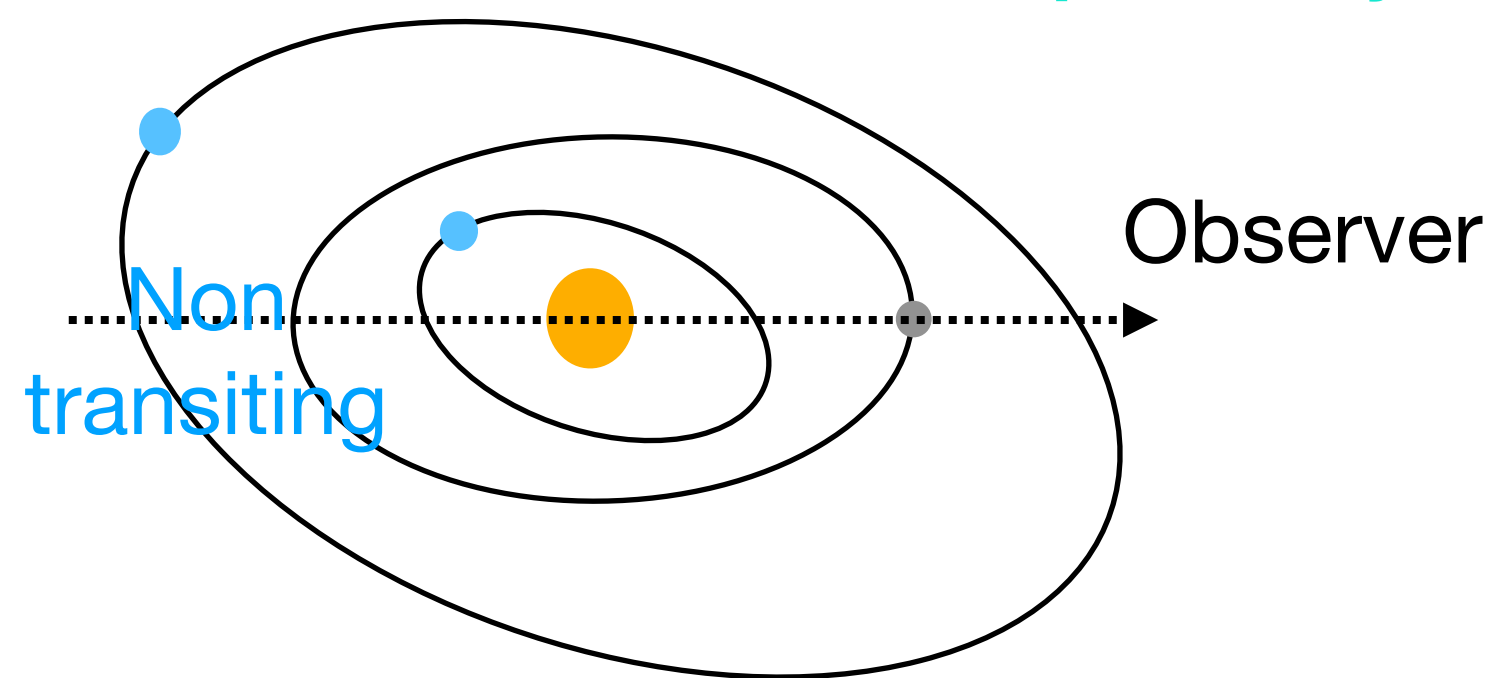
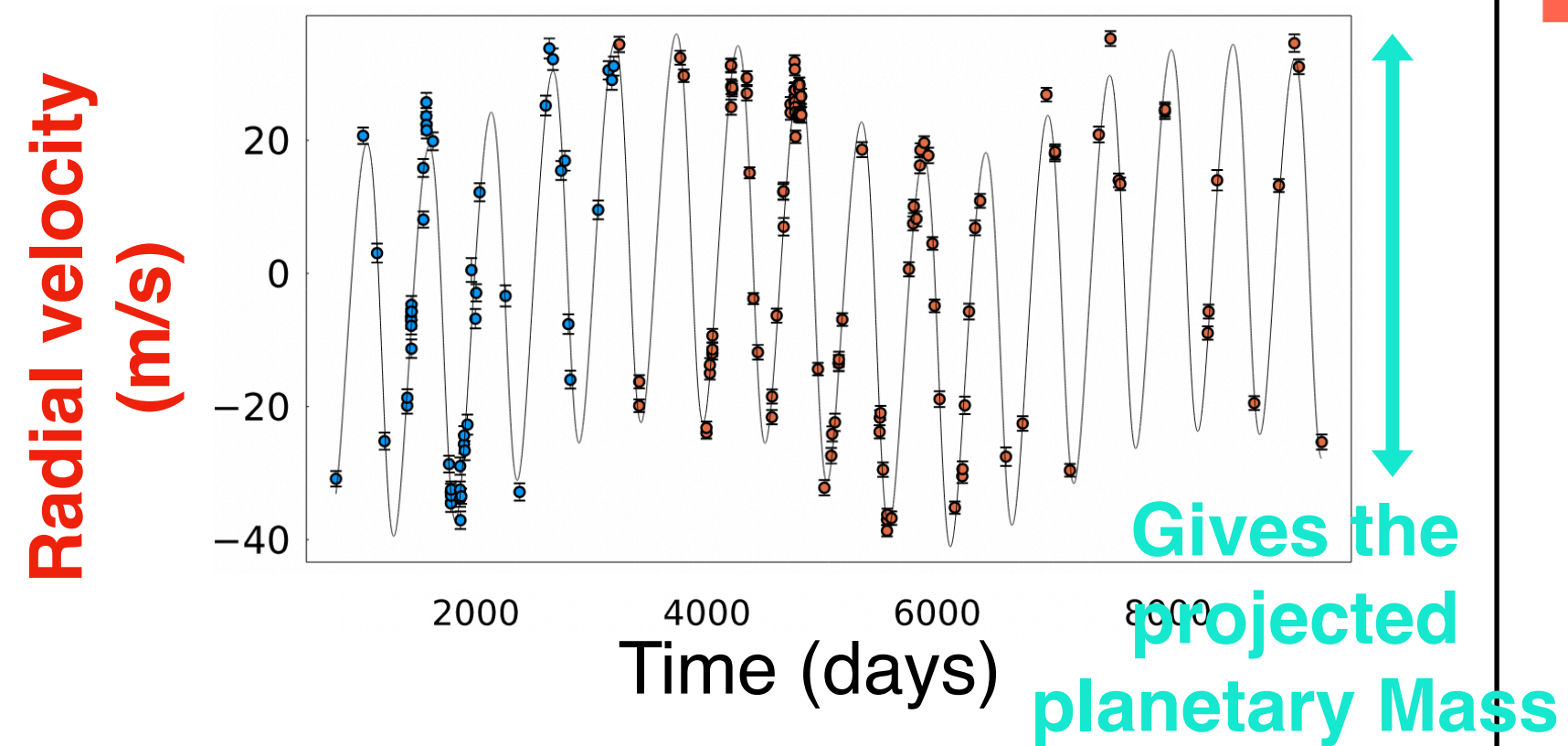
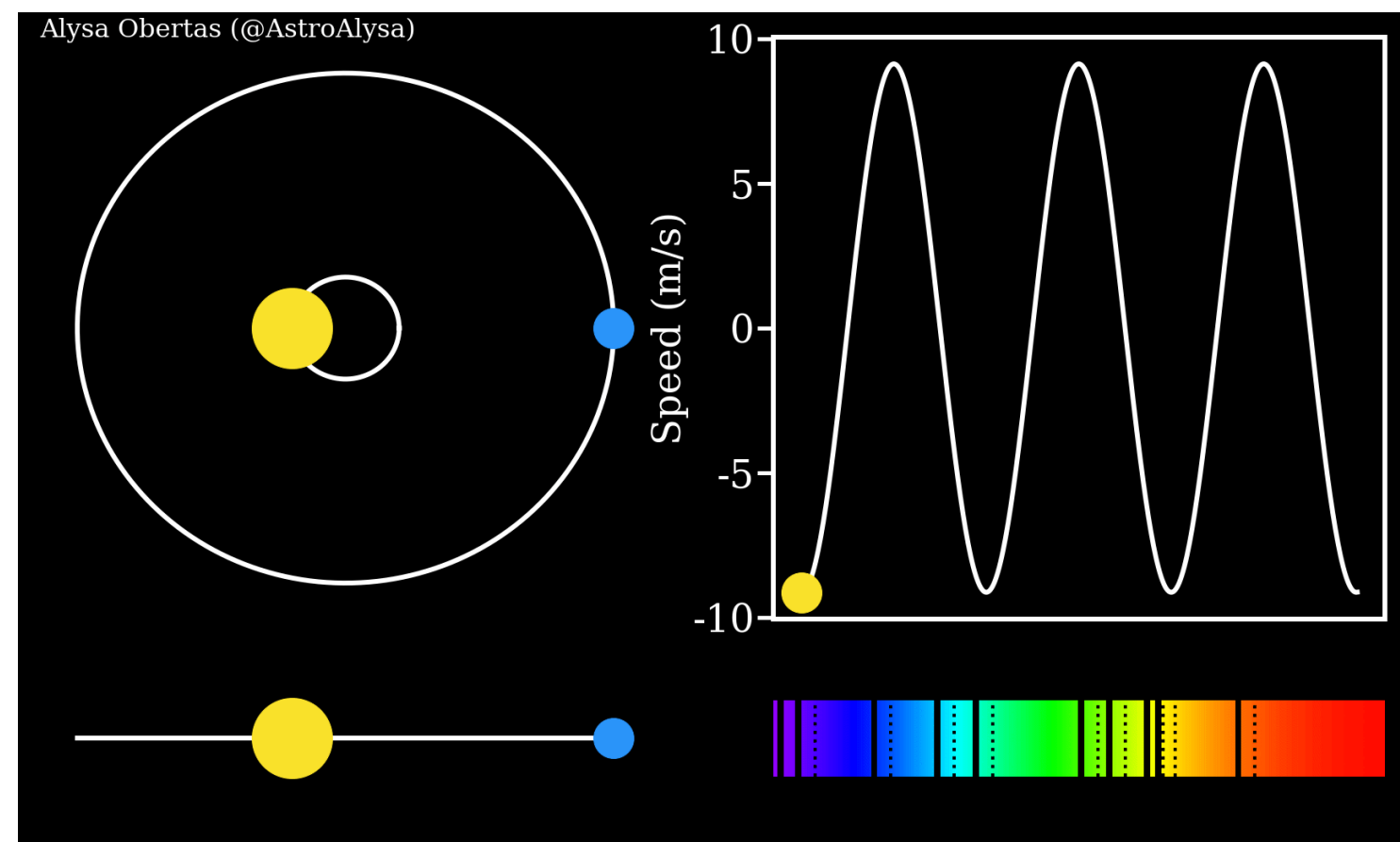
JWST, PCS@ELT
LIFE (ESA), **HWO** (NASA),



From JWST
Transiting Exoplanet
Community Early
Release Science
Team et al. 2022

Characterizing the atmosphere of **Earth-analogs** from space: priority for the horizon 2050

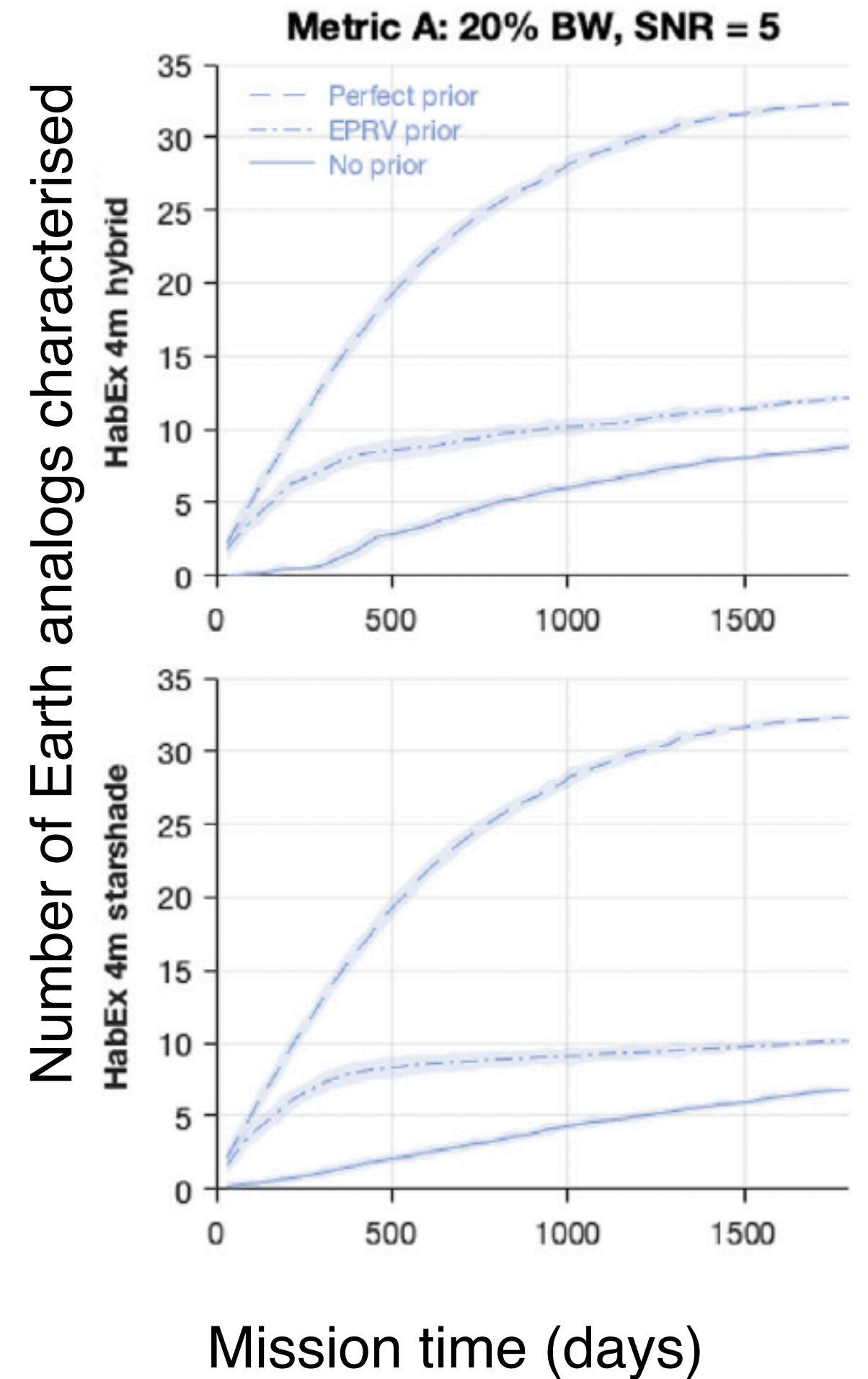
Radial velocities are key



Mass is essential to interpret spectroscopic observations of atmospheres (Batalha et al. 2017, 2019)

The yield of atmospheric characterization is greatly improved if Earth like planets are detected in advance, **require targets $\lesssim 20$ parsec**

- + Results obtained faster, more time for characterisation
- + Mitigates the risk of the missions
- + True for other kinds of planets



Morgan et al. 2021

NASA Extreme precision radial velocity report Crass et al. 2021

Roadmap to the detection of Earth-like planets with radial velocities to prepare HWO

« **Musts** » = requirements

- Determine by 2025 the **feasibility** to detect Earth-mass planets in the habitable zone of solar-type stars
- Demonstrate by 2025 **on sky** precision of 30 cm/s
- Conduct precursor survey: now-2035 on 100 stars on the « **green target list** »

« **Wants** »

- Survey as many stars as possible (« Yellow list: 100 stars »)
- Least estimated cost

Green target list (106 stars)

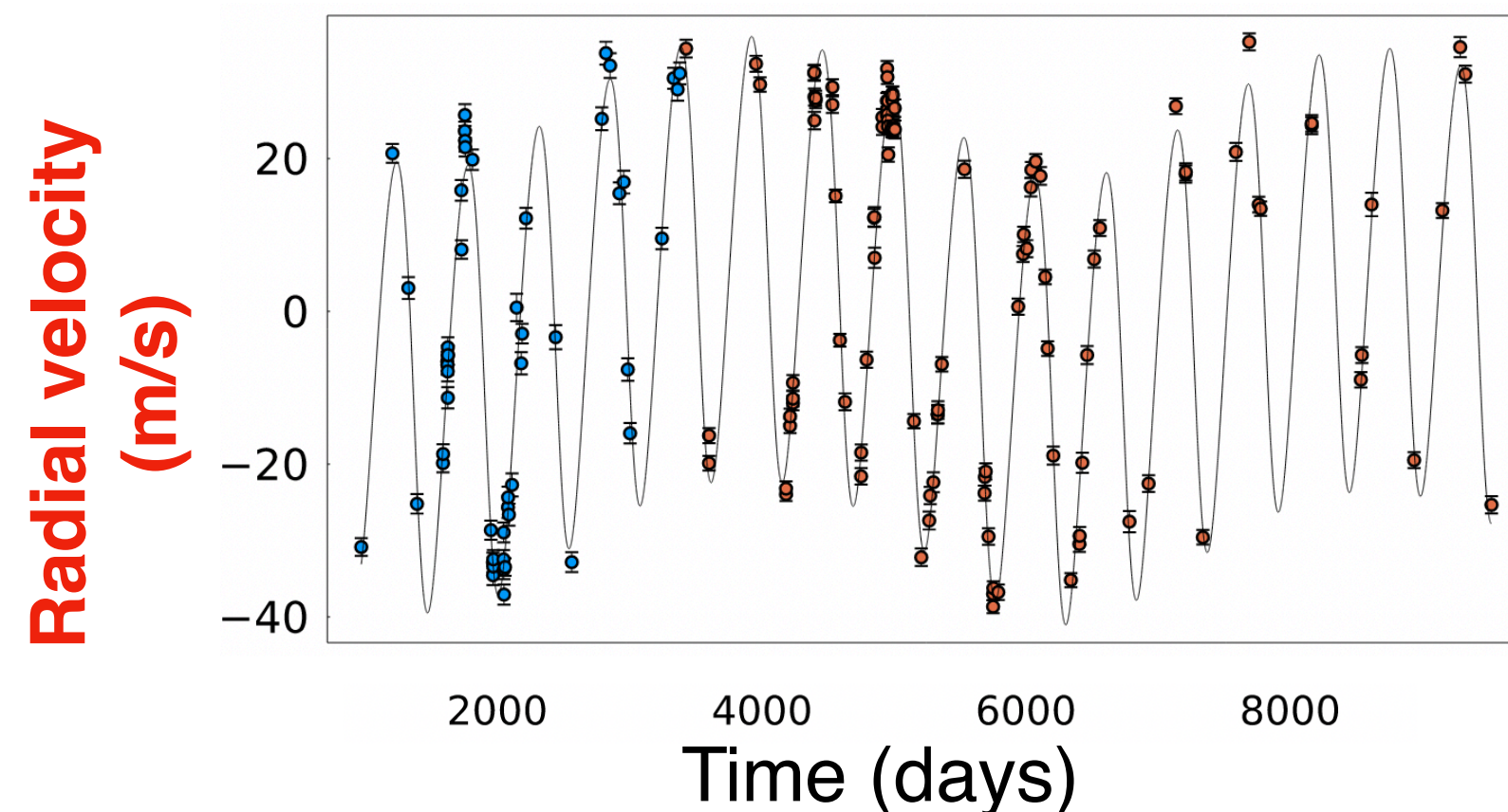
- Spectral types F7–K9 and
- $v \sin(i) < 5$ km/s
- Close (On the HabEx ‘deep survey’ or ‘50 highest priority stars’ lists (Gaudi et al. 2020), or on at least 2 other mission concept target lists (including LUVOIR-A, LUVOIR-B, HabEx ‘master list,’ Starshade Rendezvous))

NASA Extreme precision radial velocity report

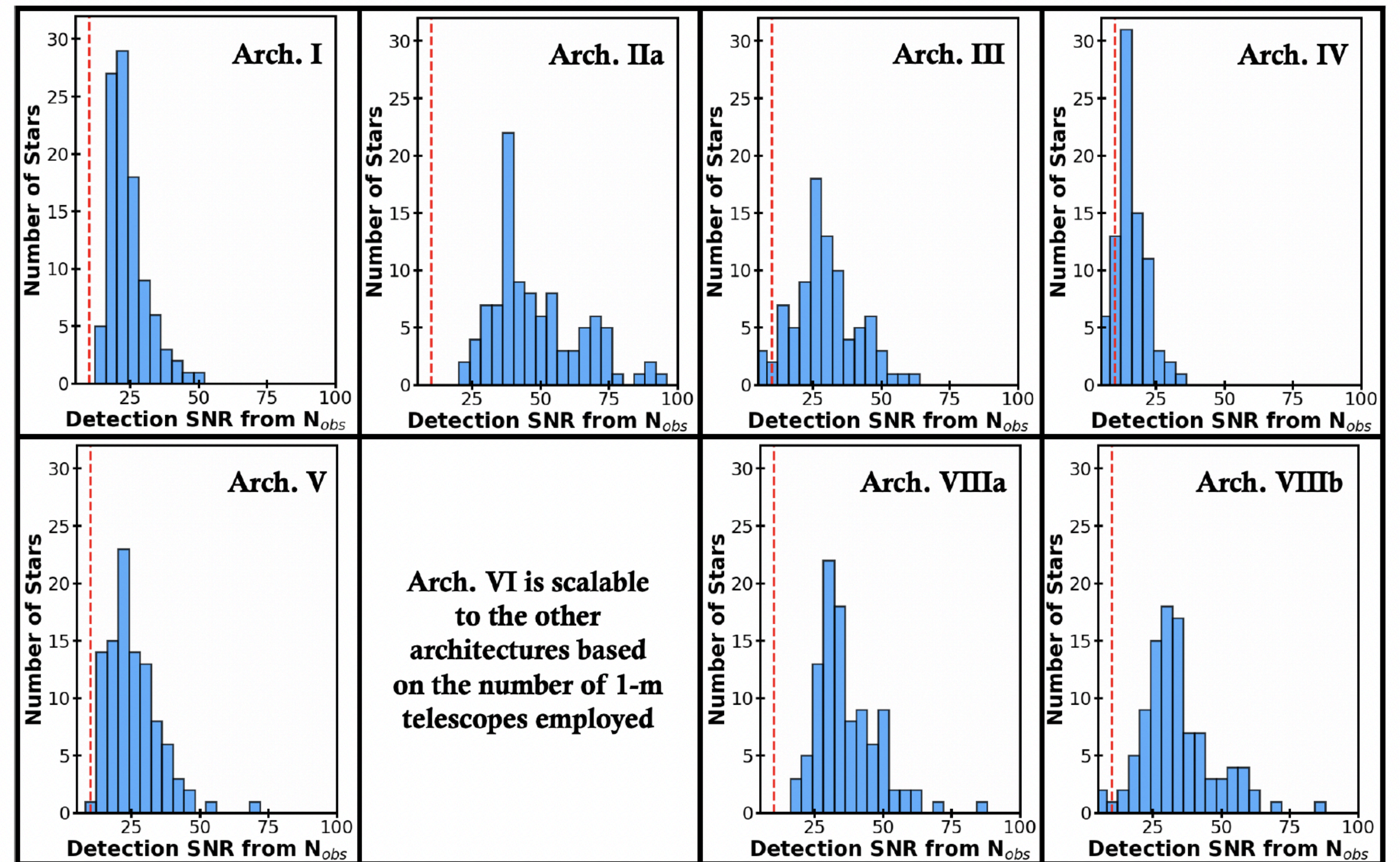
Simulation with different telescope/instrument associations

In principle the 100 prioritised targets can all be characterised

Simulation done in the best case scenario: White noise



Histogram of obtained signal to noise ratio for 100 priority stars assuming they all have an Earth



Crass et al. 2021

Excerpts from the EPRV report

Telluric line contamination is assumed to be limited and correctable

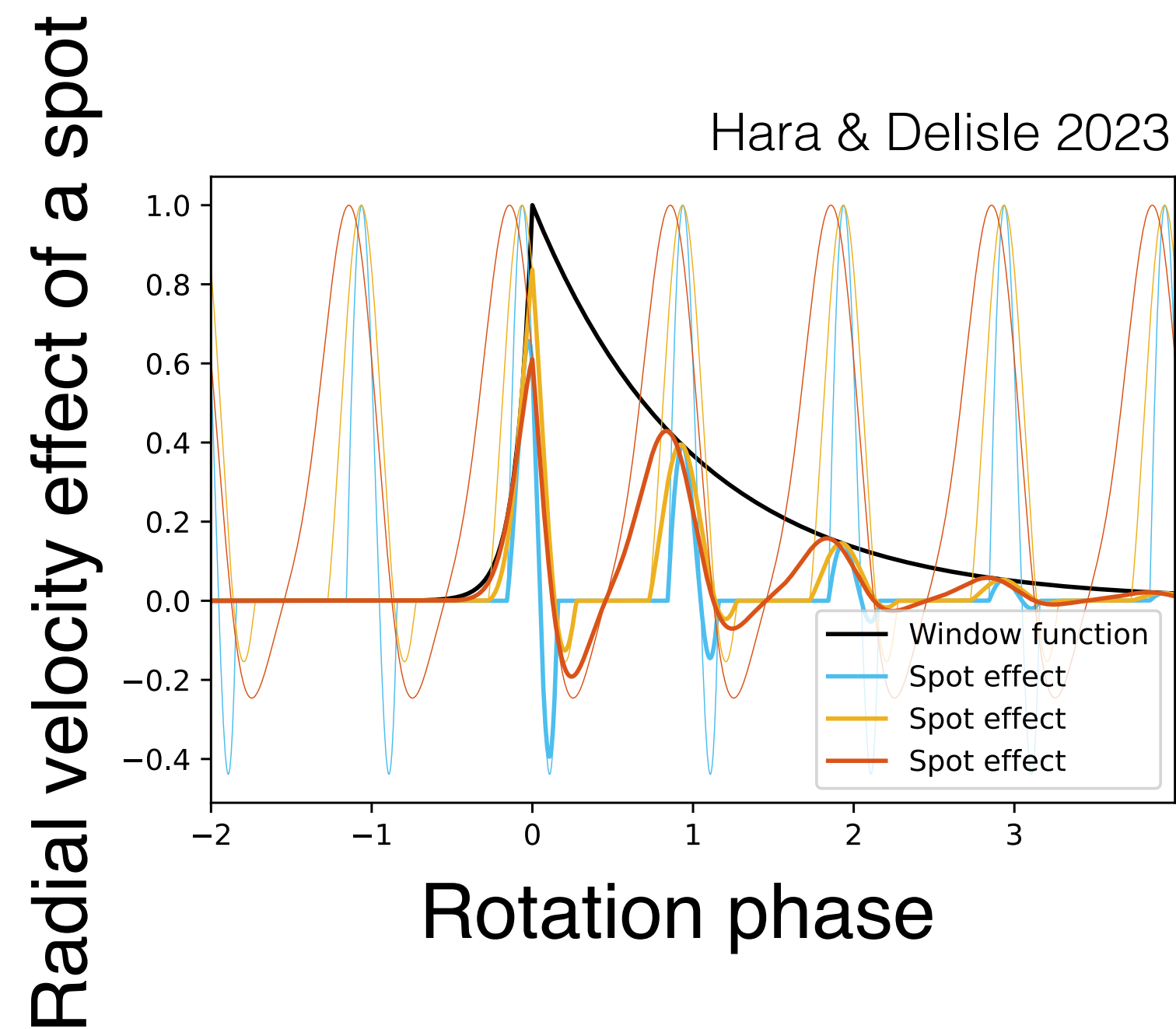
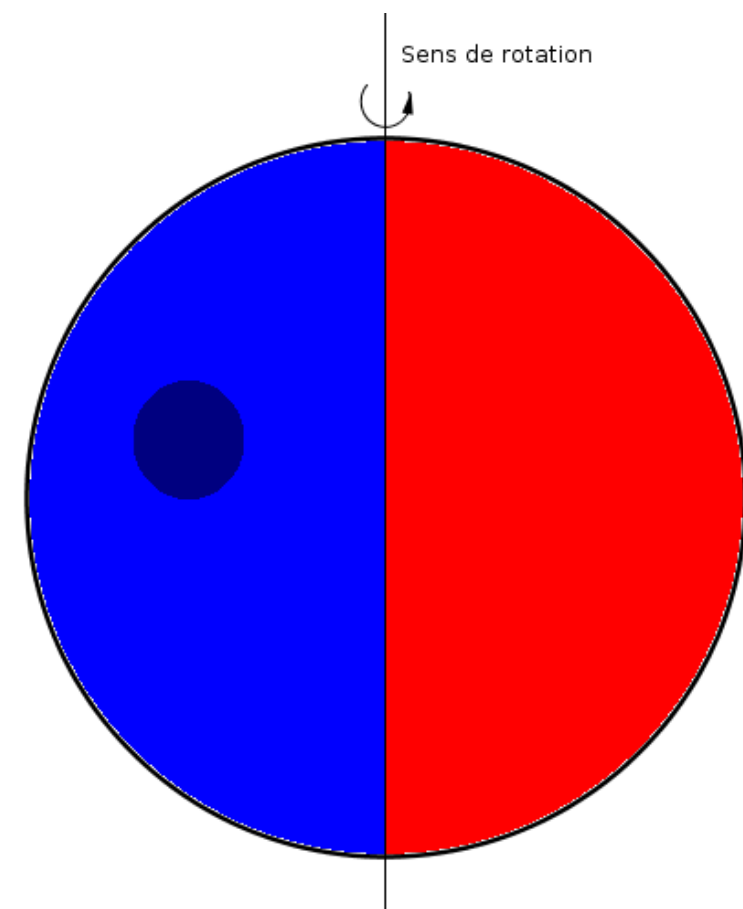
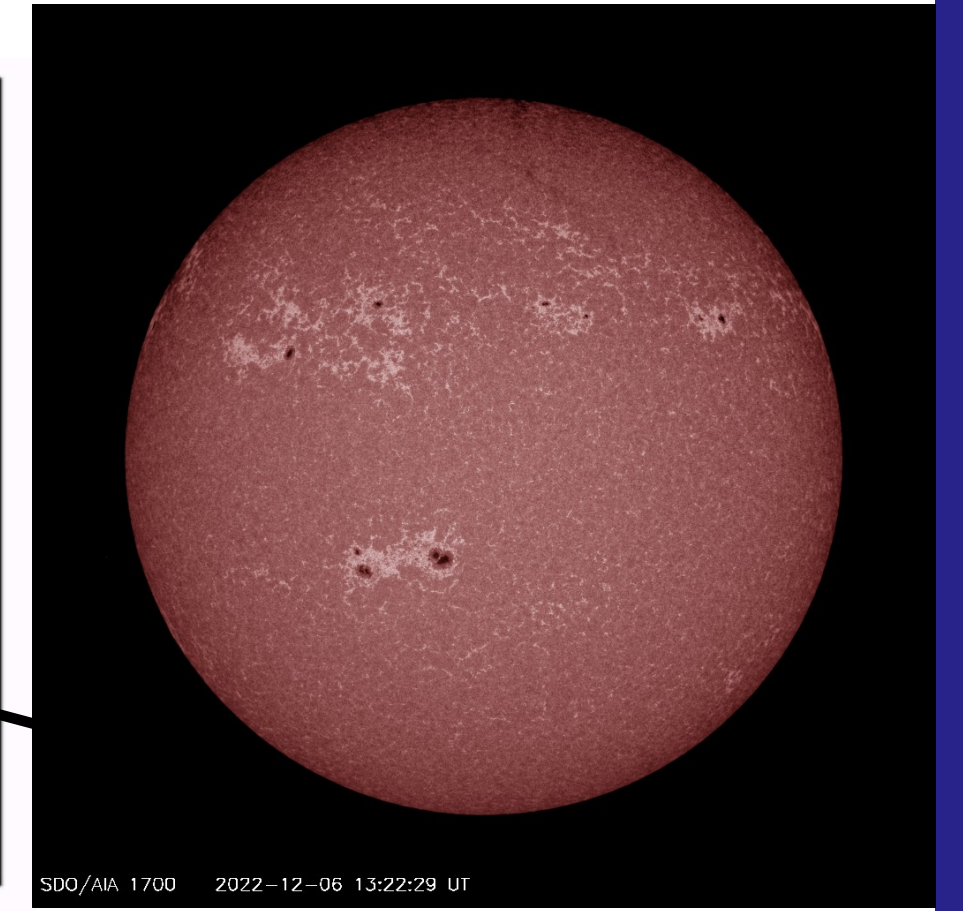
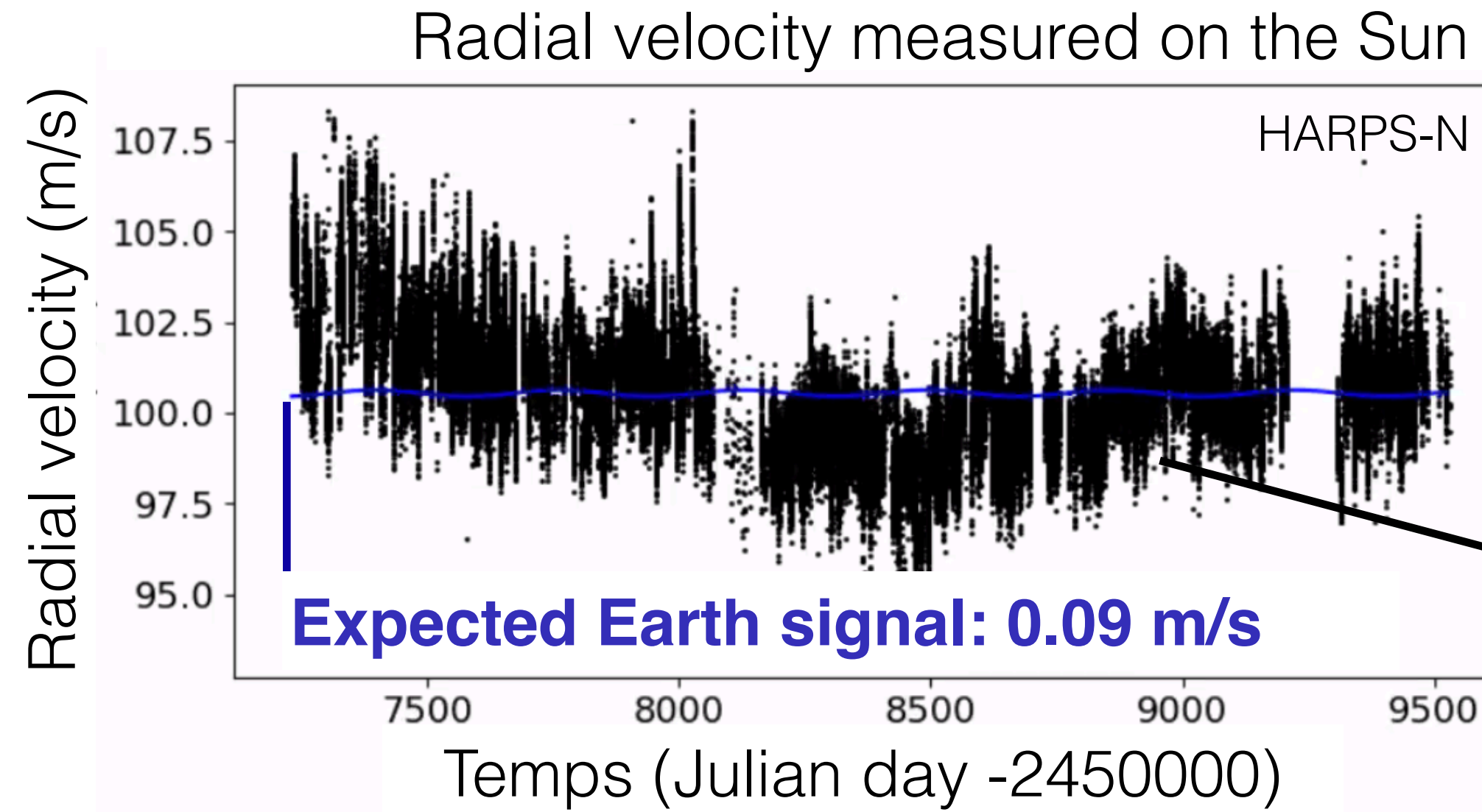
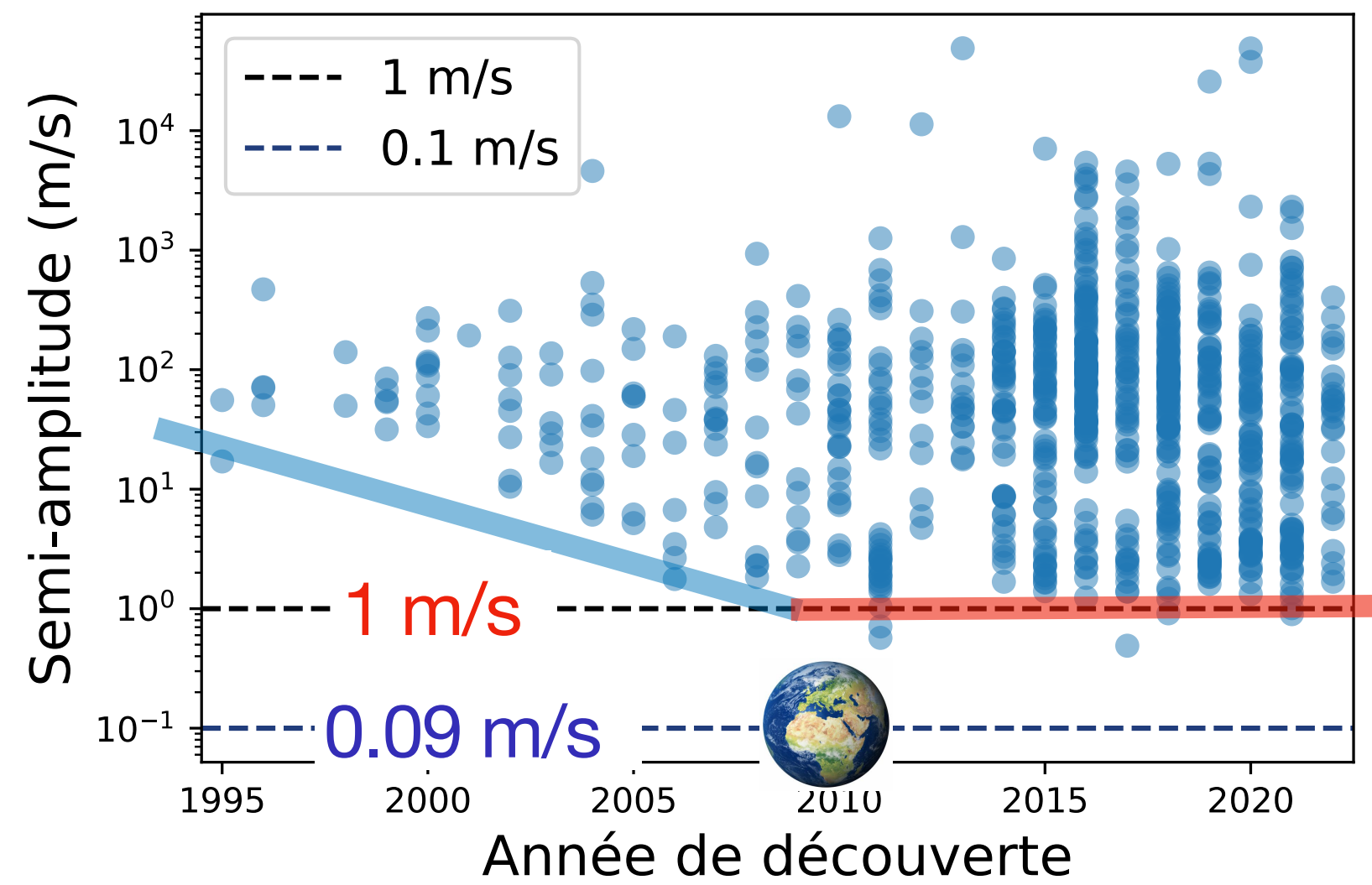
Weather conditions are uncorrelated.

We do not account for errors due to RV contamination from additional planets in the same planetary system

Stellar variability is assumed to be adequately characterized and mitigated during the data extraction of RV signals such that the remaining signal due to stellar variability is uncorrelated in time.

Extreme precision radial velocity is very useful if these problems can be mitigated

Noise in radial velocities is not white

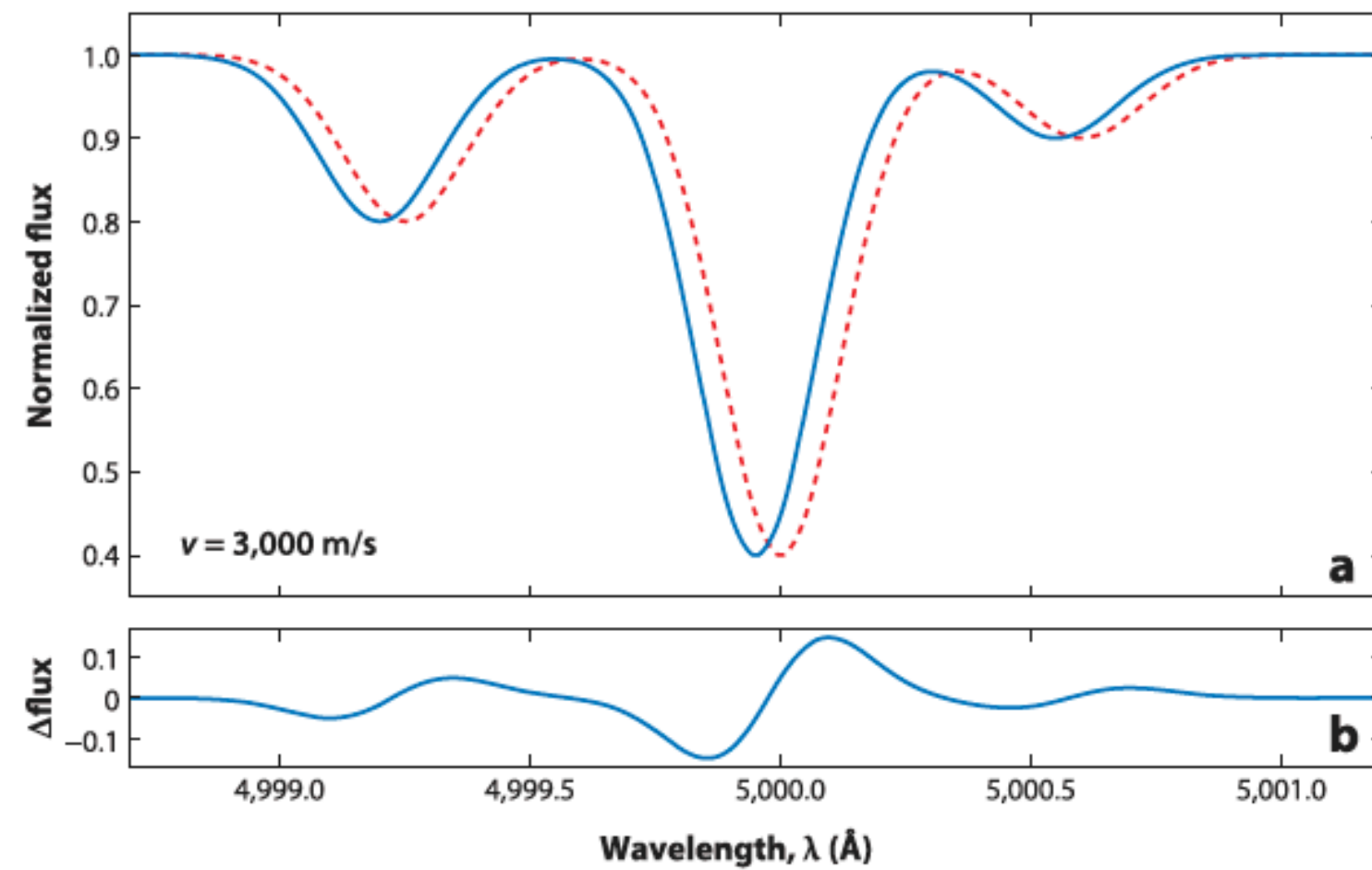


Stellar variability (magnetic activity, granulation, oscillations, meridional winds...) + systematics + telluric lines absorption

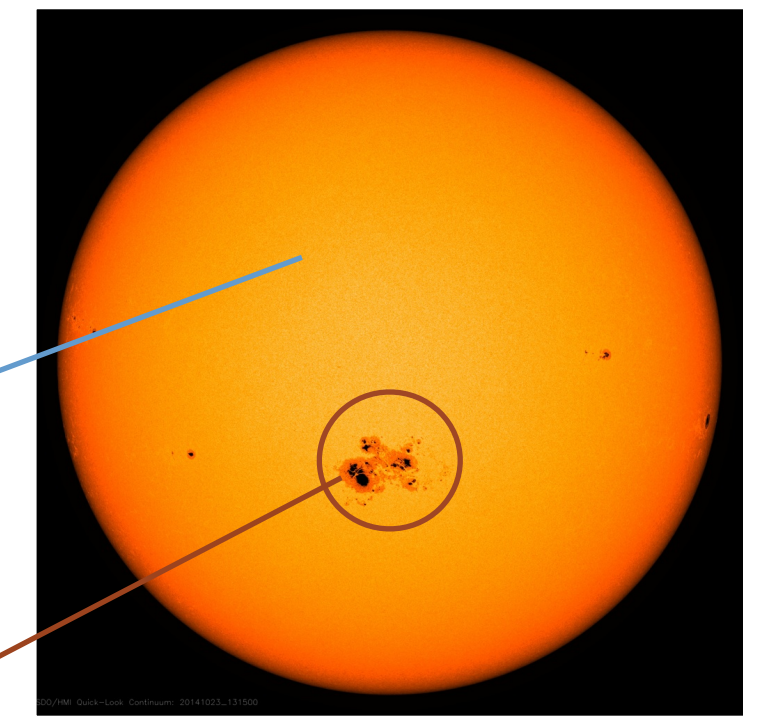
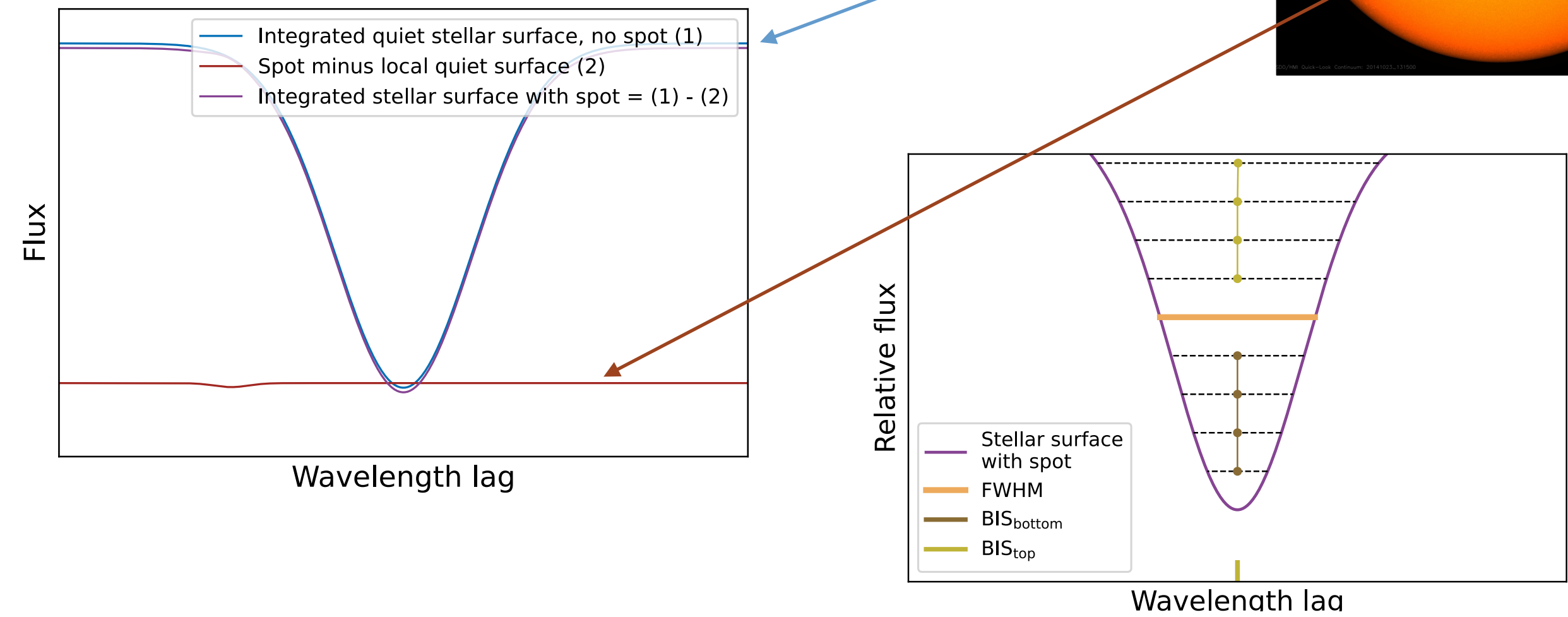
understanding and correcting these signals is crucial, data analysis methods play a key role

Key ideas

Planets induce a pure Doppler shift

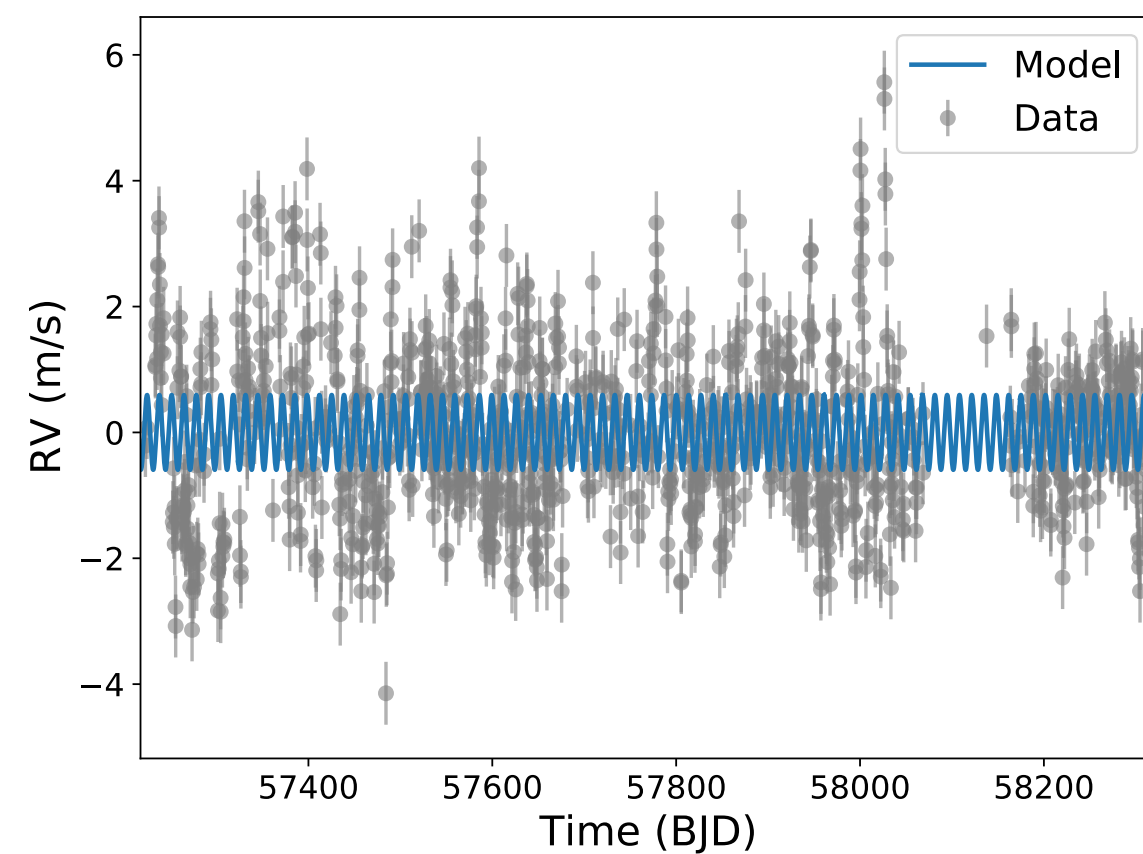


stellar and instrumental effects change the shape of the spectrum



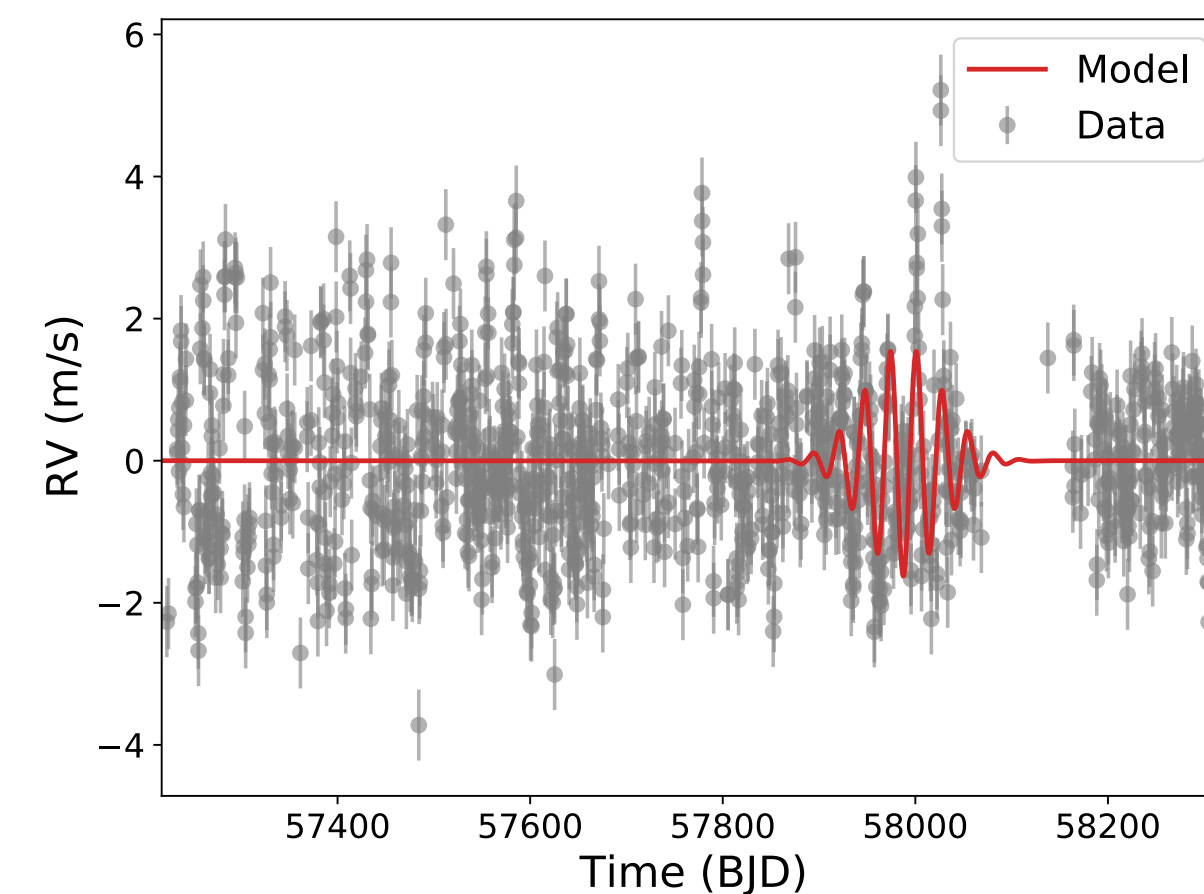
Planets induce a periodic signal

RV data and max. likelihood model



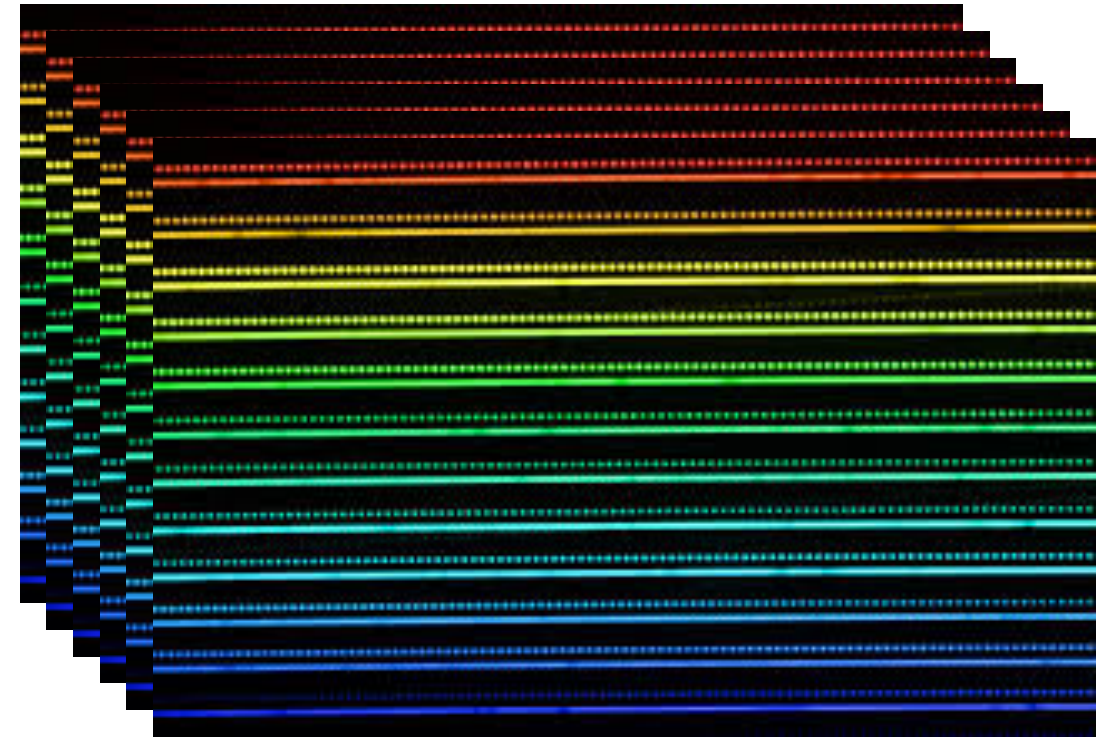
Stellar and instrumental effects are (usually) not strictly periodic

RV data and max. likelihood model



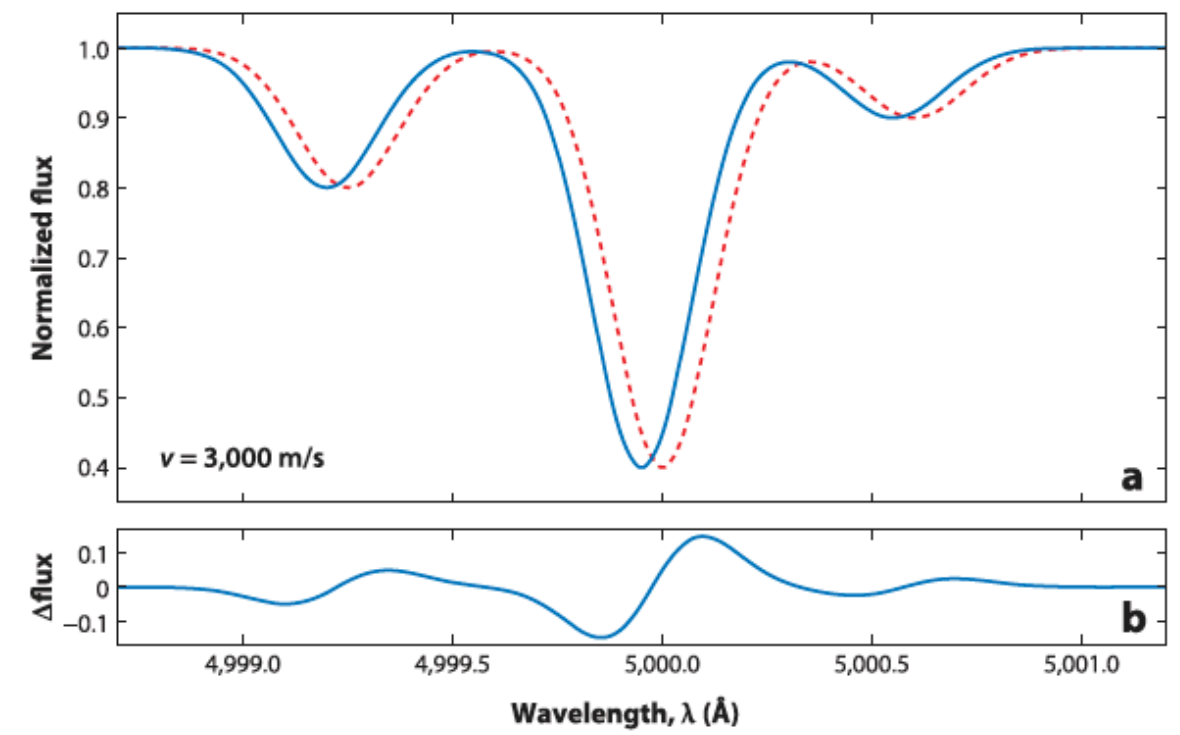
Steps of RV analysis

Time series of Raw spectra



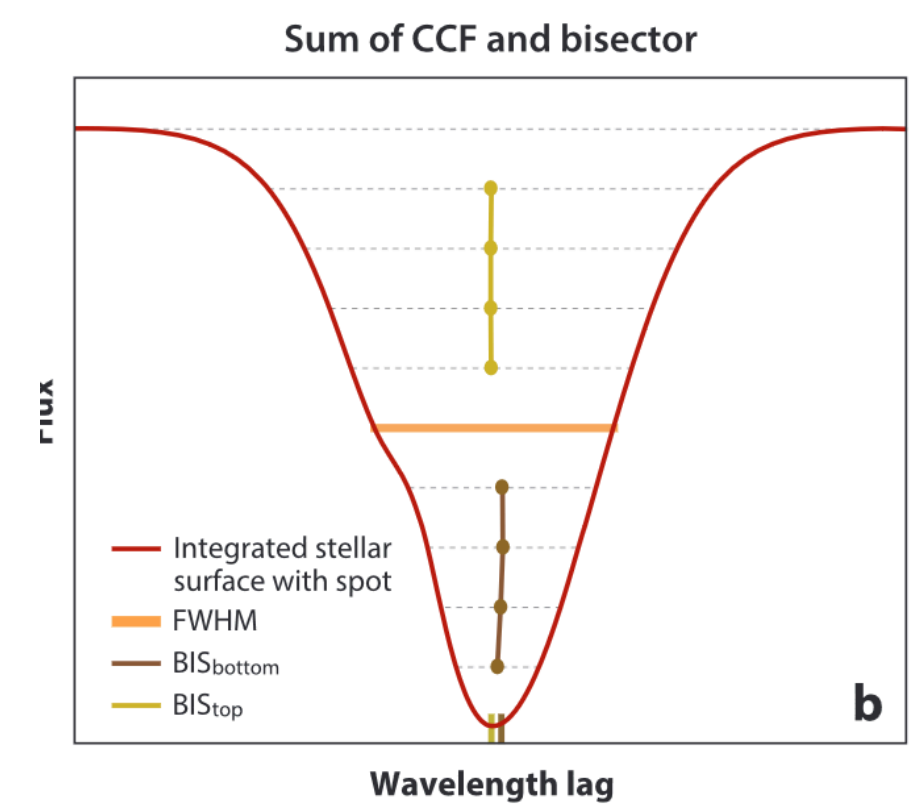
Reduce to

1 dimensional spectrum



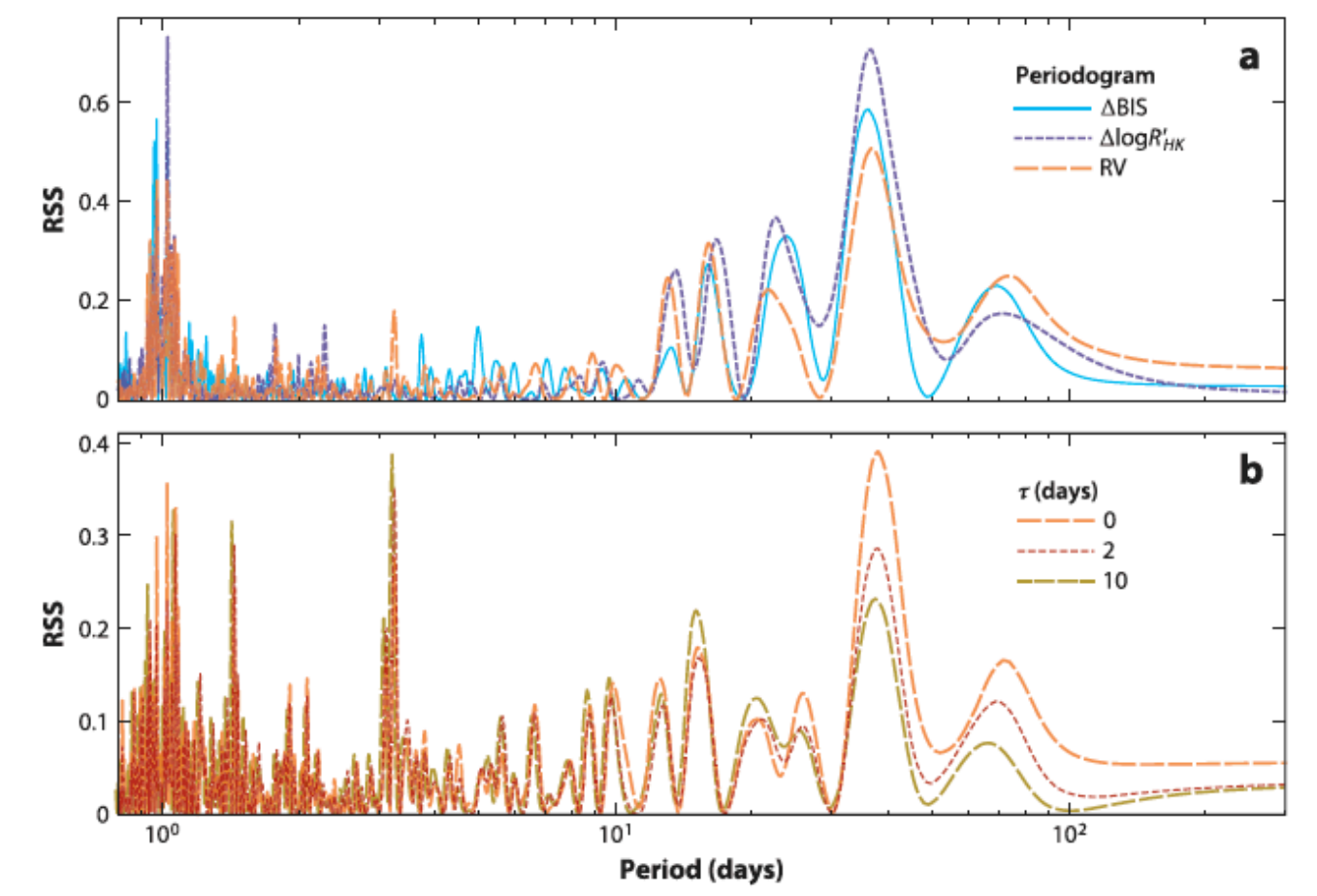
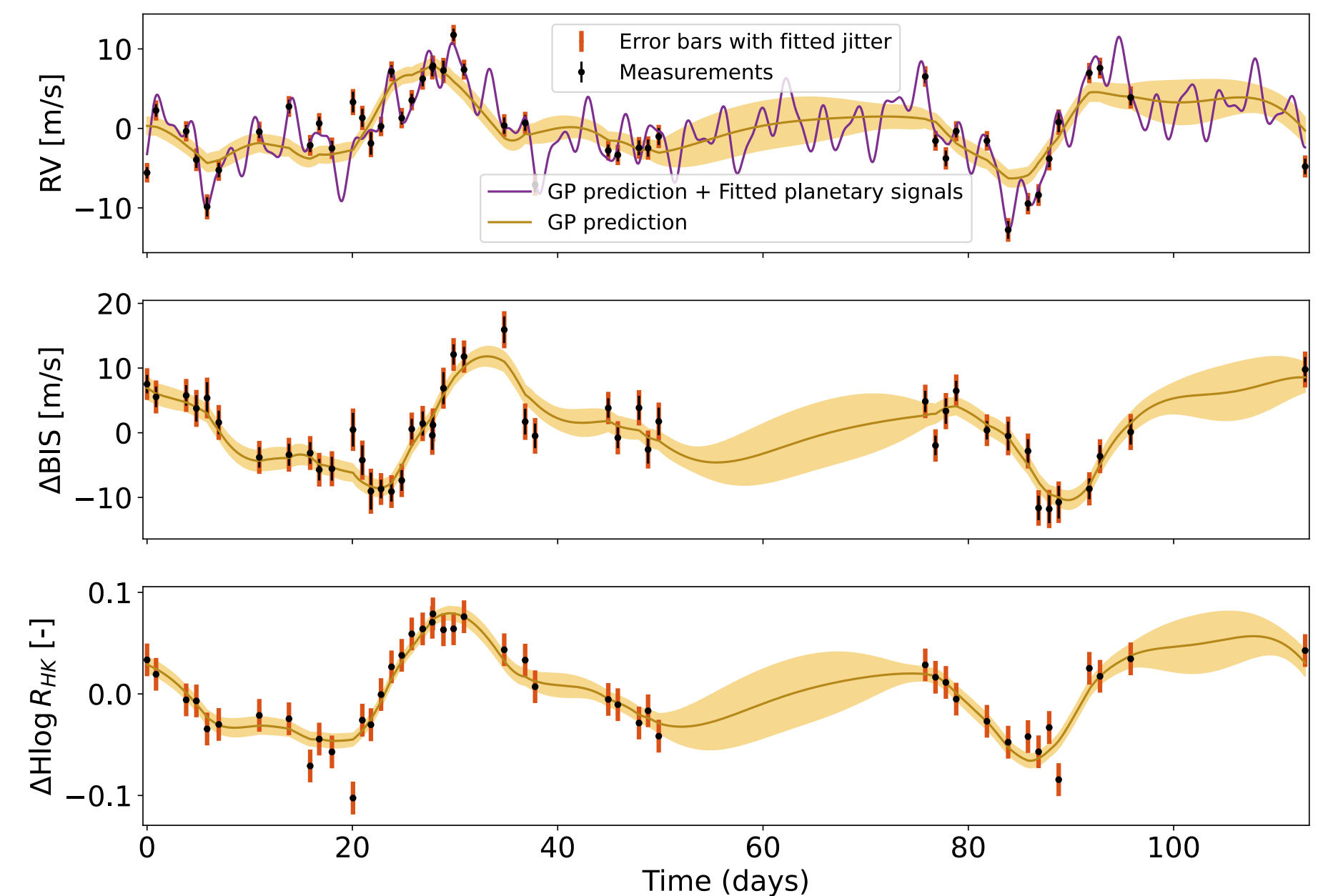
Extract

Radial velocity Indicators



Modelling: which noise properties, priors?

Are planets detected? Which ones?



Hara & Ford, Annual Reviews of statistics and Its Application: what are the problems to solve?

An optimal exoplanet detection criterion

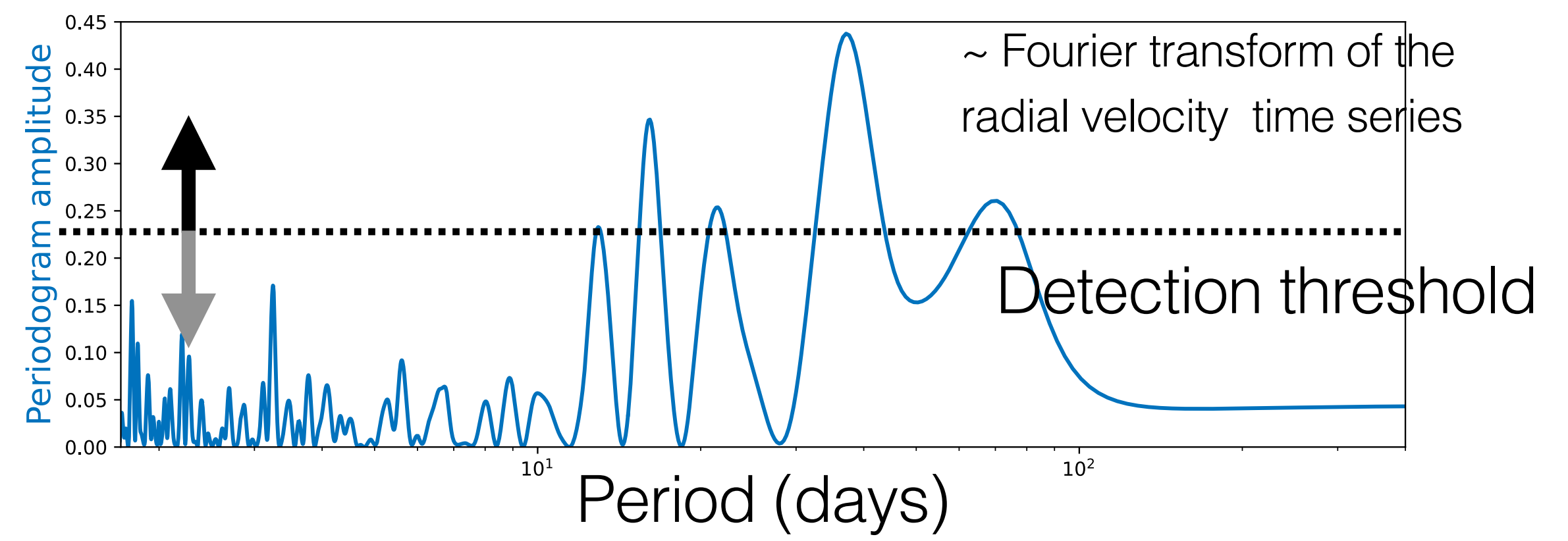
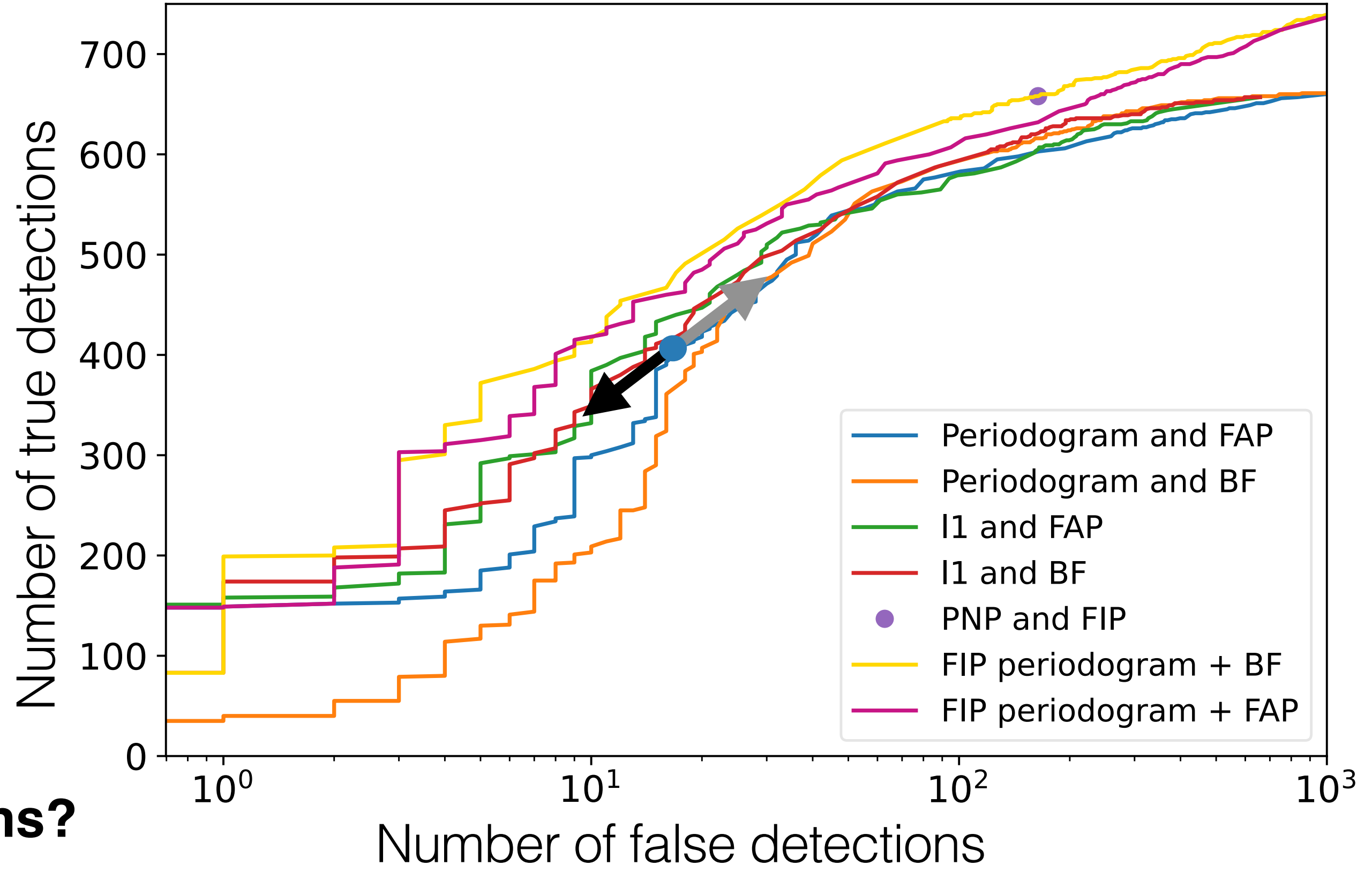
1000 radial velocity datasets with 0,1 or 2 planets

- Analysed with the same model
- With different detection criteria

Bayes factors and FAPs

- **Optimal?**
Which criterion maximises true detections?
- Do not encode where the planet is

- Are not defined on a very intuitive scale



When claiming the detection of $n > 0$ components with parameters in $\Theta_i, i = 1, \dots, n$, the Θ_i s can be considered as n "boxes". We want to evaluate the quality of this claim if the true components have parameters in $\Theta_1, \dots, \Theta_n$, where Θ_i might be different from Θ_j for $i \neq j$. To put it in a more formal way, we consider the following problem: given n boxes $\Theta_1, \dots, \Theta_n$ and a set of n components $\theta_1, \dots, \theta_n$, we want to find the maximum number of correct detections. Furthermore, if a component θ_i belongs to several boxes, there might be several ways to yield different number of correct detections. Suppose components, one in Θ_1 , one in Θ_2 , with a non empty intersection $\Theta_1 \cap \Theta_2$ and θ_1 belongs to $\Theta_1 \cap \Theta_2$ and θ_2 belongs to Θ_2 , we have two or only one case, we choose the injection which leads to as many correct detections as possible. We denote by m the maximum number of different θ_i s that belong to A_m^k , the region of parameter space with k component components in each of the $\Theta_i, i = 1, \dots, m, m \leq n$.

If k planets are truly present in the data, n detections are missed. Adding a term $-\beta(\min(k, n) - i)$ whenever it happens. The expected utility $E[U\{a, (\theta, \eta)\}]$ is given by

$$E_{\theta, \eta}[U\{a, (\theta, \eta)\}] = -n\alpha p(0|y) + [-(n-1)\alpha I_{A_1} - (n\alpha + \beta)(1 - I_{A_1})] p(1|y) + [-(n-2)\alpha I_{A_2} - ((n-1)\alpha + \beta) I_{A_2} - (n\alpha + 2\beta)] p(2|y) + \dots + \left[\sum_{i=1}^k -(n-i)\alpha - (k-i)\beta \right] I_{A_k} - (n\alpha + k\beta) \left(1 - \sum_{i=1}^k I_{A_i} \right)$$

$$E_{\theta, \eta}[U\{a, (\theta, \eta)\}] = -n\alpha + (\alpha + \beta) \sum_{i=1}^n i I_{A_i} - \beta \sum_{k=1}^{n_{max}} k p(k|y) = -(\alpha E[FD] + \beta E[MD])$$

$$E_{\theta, \eta}[U\{a, (\theta, \eta)\}] = -n + (1 + \gamma) \sum_{i=1}^n i I_{A_i} - \gamma \bar{n}$$

$$E_{\theta, \eta}[U\{a, (\theta, \eta)\}] = -n + \alpha \sum_{i=1}^n i I_{A_i} - \beta \sum_{k=1}^{n_{max}} k p(k|y)$$

PROOF. With the notation above, we have seen that u_n is increasing, v_n is decreasing, and $v_{n-1} - v_n$ is decreasing. Furthermore, by hypothesis $u_n - u_{n-1}$ is increasing, which by definition of v'_n means that $v'_{n-1} - v'_n$ is decreasing. In the following, we reason on v_n but the argument is identical if v_n is replaced by v'_n . Let us fix $x > 0$. The constrained problem is

$$\min_n v_n \quad \text{subject to } u_n \leq x$$

Since u_n is increasing, there is a highest n_0 constraint, i.e. such that $u_{n_0} \leq x$. Since problem is found for $n = n_0$, for any $n > n_0$, we choose γ such that the solution of the maximization satisfying the constraint. We show that to a larger value of $v_n + \frac{1}{\gamma} u_n$. From our hypotheses, we see that the ratio $v_n + \frac{1}{\gamma} u_n$ is increasing.

$$\gamma \geq \frac{u}{v}$$

Note that if $v_{n_0} - v_{n_0+1} = 0$, since $v_n - v_{n+1}$ is increasing, we can always find $n > n_0$ such that $v_n - v_{n+1} > 0$. For $n > n_0$, we want to maximize $v_n + \frac{1}{\gamma} u_n$.

$$\gamma \leq \frac{u}{v}$$

These two conditions can be satisfied simultaneously if $\frac{u_{n_0} - u_{n_0+1}}{v_{n_0} - v_{n_0+1}} > \frac{u_{n_0+1} - u_{n_0+2}}{v_{n_0+1} - v_{n_0+2}}$.

As long as the sequence $(u_{n+1}^y - u_n^y)_{n \geq n_0}$ is strictly increasing, the constrained problem has the same solution but can be ensured under the following condition.

LEMMA E.4. If $\forall n > 0, \exists i_0, \forall j = 1, \dots, n$

$$E_{\theta, \eta}[U\{a, (\theta, \eta)\}] = -n\alpha + (\alpha + \beta) \sum_{i=1}^n i I_{A_i} - \beta \sum_{k=1}^{n_{max}} k p(k|y)$$

$$E[FD] = n - \sum_{i=1}^n i I_{A_i}$$

$$E[MD] = \bar{n} - \sum_{i=1}^n i I_{A_i}$$

where $\bar{n} := \sum_{k=1}^{n_{max}} k p(k|y)$ does not depend on the number of components n (or equivalently $\alpha > 0$, since α is non negative), we can denote by $\gamma = \beta/\alpha$, without loss of generality we can maximize

$$E_{\theta, \eta}[U\{a, (\theta, \eta)\}] = -n + (1 + \gamma) \sum_{i=1}^n i I_{A_i} - \gamma \bar{n}$$

LEMMA E.1. Denoting $[1, n]_j$ where $[1, n]_j$ is a draw of j indices without replacement

$$I_{A_j} = \sum_{k_1, \dots, k_j \in [1, n]_j} I_{\Theta_{k_1} \wedge \Theta_{k_2} \wedge \dots \wedge \Theta_{k_j}}$$

$$\sum_{i=1}^n I_{\Theta_i} = \sum_{i=1}^{n-1} \sum_{j=0}^n \sum_{k_1, \dots, k_j \in [1, n] \setminus \{i\}} I_{\Theta_{k_1} \wedge \dots \wedge \Theta_{k_j} \wedge \Theta_i}$$

In this sum, the term $I_{\Theta_i \wedge \Theta_{k_1} \wedge \dots \wedge \Theta_{k_j}}$ appears n times, $n-1$ times, so we obtain the desired result.

PROOF. Let us suppose that $\exists n \geq 1$ such that $u_{n+1} - u_n < u_n - u_{n-1}$. Replacing by the explicit expression of u_n , the inequality is equivalent to

$$\sum_{i=1}^n I_{\Theta_i} - \sum_{i=1}^{n+1} I_{\Theta_i} < \sum_{i=1}^{n-1} I_{\Theta_i} - \sum_{i=1}^n I_{\Theta_i}$$

$$\sum_{i=1}^n I_{\Theta_i} - \sum_{i=1}^{n+1} I_{\Theta_i} < \sum_{i=1}^{n-1} I_{\Theta_i} - \sum_{i=1}^n I_{\Theta_i}$$

The term I_{Θ_i} is a sum of n I_{Θ_i} , with disjoint Θ_i . By definition of u_n , the left hand side of the inequality is less than or equal to 0 and the left hand side is greater than or equal to 0, which is absurd.

If $\forall i = 1, \dots, n+1, \exists j = 1, \dots, n-1, \Theta_i \cap \Theta_j \neq \emptyset$. In that case $[1, n] \cap \Theta_j \neq \emptyset$, otherwise due to lemma 3.2 this would lead to a contradiction.

If the condition of Lemma E.4 is not satisfied one can find a counterexample $u_n^y < u_{n-1}^y - u_{n-2}^y$ and the equivalence of utility maximisation and optimisation is not guaranteed. Finally, we have the desired result.

THEOREM E.5. Let us consider a dataset y and suppose that it is separable at all orders, $n = 1, \dots, n_{max}$ then there exists an increasing function $\gamma(x)$ and a function $\gamma'(x) > 0$ such that $\gamma(x) > \gamma'(x)$ is the solution of the constrained problem.

PROOF. Under the hypothesis of separability, by lemma E.4, (increasing, and by lemma E.3, we have the desired result.

APPENDIX F: OTHER DEFINITION OF MISSED DETECTIONS. In this appendix, we show that the optimal procedure is similar to the one defined in Hara et al. (2022b).

DEFINITION F.1 (Missed detections: other definition). If n components are truly present in the data, we define the missed detections as $m = n - \sum_{i=1}^n I_{A_i}$.

Let us consider $\Theta_{n+1} \in T$. The solution to (P_{n+1}) can be written as

$$\arg \max_{\Theta_i \in T, \forall i \in [1, n], \Theta_i \cap \Theta_{n+1} = \emptyset} I_{\Theta_{n+1}} + \sum_{i=1}^n I_{\Theta_i}$$

Either $\forall i \in [1, n], \Theta_i \cap \Theta_{n+1} = \emptyset$ then thanks to (P1), for $E = T^n$ and $T^n, \forall i, x_i \notin \Theta_{n+1}$

$$\arg \max_{\Theta_i \in T, \forall i \in [1, n], \Theta_i \cap \Theta_{n+1} = \emptyset} \sum_{i=1}^n I_{\Theta_i} = (\Theta_i^n)_{i=1, \dots, n}$$

$$\arg \max_{\Theta_i \in T, \forall i \in [1, n], \Theta_i \cap \Theta_{n+1} = \emptyset} \sum_{i=1}^n I_{\Theta_i} + \sum_{i=1}^n I_{\Theta_i} = (\Theta_i^n)_{i=1, \dots, n}$$

up to a permutation of the indices (see remark B.2). If $\exists i \in [1, n+1], \forall j \in [1, n], \Theta_i \cap \Theta_j \neq \emptyset$ then the same argument applies and the solution to (P_{n+1}) is $(\Theta_i^n)_{i=1, \dots, n}$.

LEMMA E.2. $\forall y$ in the sample space the sequence $(u_n)_{n=1, \dots, n_{max}}$ and $(v'_n)_{n=1, \dots, n_{max}}$ are decreasing.

PROOF. Let us suppose that there exists n such that $u_{n+1} < u_n$ and $v'_n > v'_{n-1}$. Then $v'_n - v'_{n-1} > 0$ and $u_n - u_{n-1} < 0$. Let us denote by i_0 an index such that $i_0 = \arg \max_{i=1, \dots, n+1} I_{\Theta_i}$.

$$u_n := \sum_{i=1}^n \text{FIP}_{\Theta_i}$$

$$v'_n := \sum_{k=n+1}^{n_{max}} (k-n)p(k|y); \quad v'_n = \bar{n} - n + \sum_{k=n+1}^{n_{max}} k p(k|y)$$

The sequence u_n is the expected number of true detections and v'_n is the expected number of missed detections for the procedure \mathcal{P}_n and \mathcal{P}'_n , respectively (see Appendix A for details). Note that u_n and v'_n are decreasing.

LEMMA E.2. $\forall y$ in the sample space the sequence $(u_n)_{n=1, \dots, n_{max}}$ and $(v'_n)_{n=1, \dots, n_{max}}$ are decreasing.

$$u_{n+1} - u_n = \sum_{i=1}^{n+1} \text{FIP}_{\Theta_i} - \sum_{i=1}^n \text{FIP}_{\Theta_i} = \text{FIP}_{\Theta_{n+1}} - \sum_{i=1}^n I_{\Theta_i}$$

$$v'_n - v'_{n-1} = \sum_{k=n+1}^{n_{max}} k p(k|y) - \sum_{k=n}^{n_{max}} k p(k|y) = (n_{max} - n) p(n|y) + \sum_{k=n+1}^{n_{max}} k p(k|y) - n p(n|y)$$

$$v'_n - v'_{n-1} = (n_{max} - n) p(n|y) + \sum_{k=n+1}^{n_{max}} k p(k|y) - n p(n|y)$$

$$v'_n - v'_{n-1} = (n_{max} - n) p(n|y) + \sum_{k=n+1}^{n_{max}} k p(k|y) - n p(n|y)$$

$$v'_n - v'_{n-1} = (n_{max} - n) p(n|y) + \sum_{k=n+1}^{n_{max}} k p(k|y) - n p(n|y)$$

$$E_{\theta, \eta}[U\{a, (\theta, \eta)\}] = -n\alpha + \alpha \sum_{i=1}^n i I_{A_i} - \beta \sum_{k=n+1}^{n_{max}} (k-n)p(k|y)$$

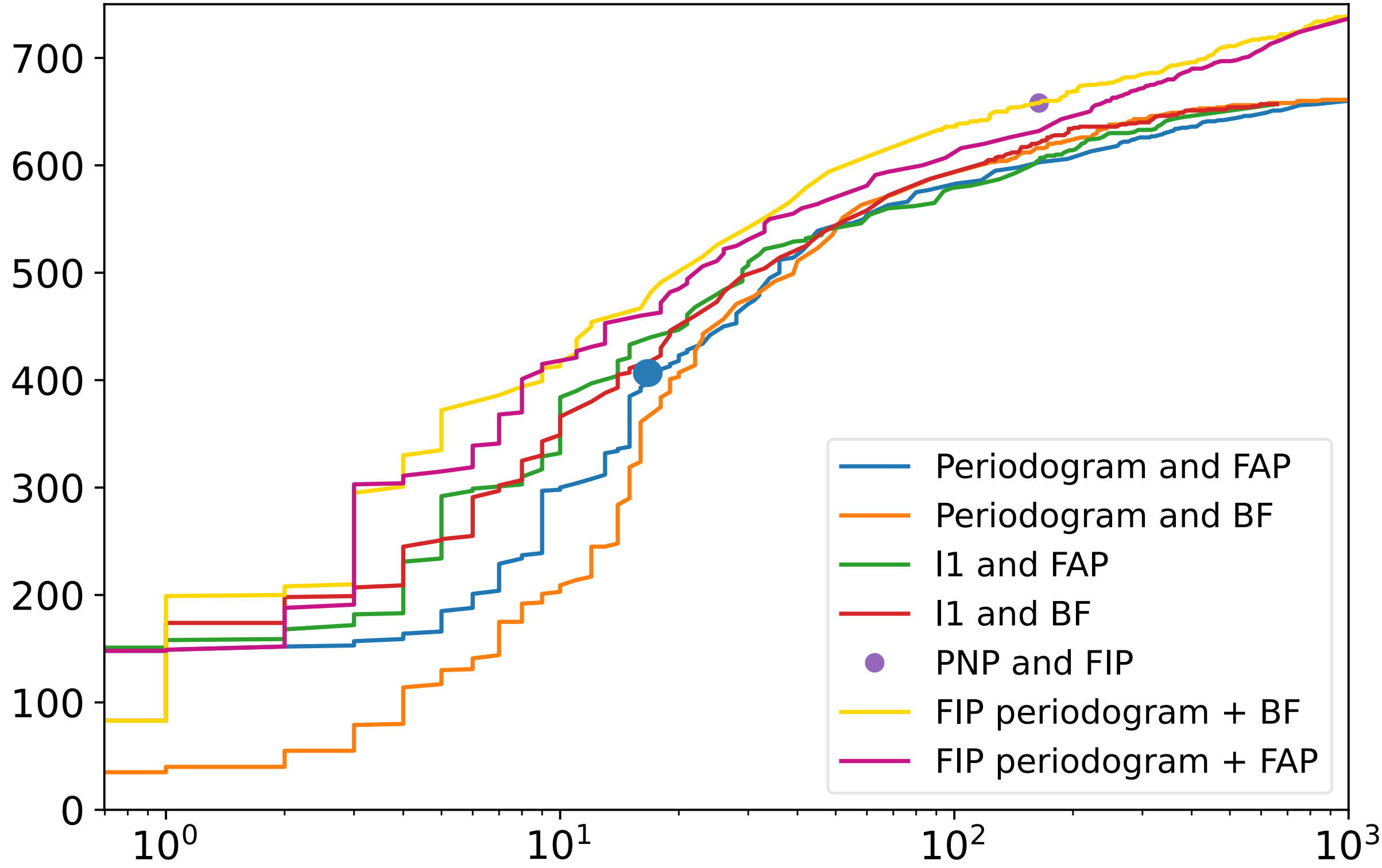
$$E[FD] = n - \sum_{i=1}^n i I_{A_i}$$

$$E[MD] = \sum_{k=n+1}^{n_{max}} (k-n)p(k|y)$$

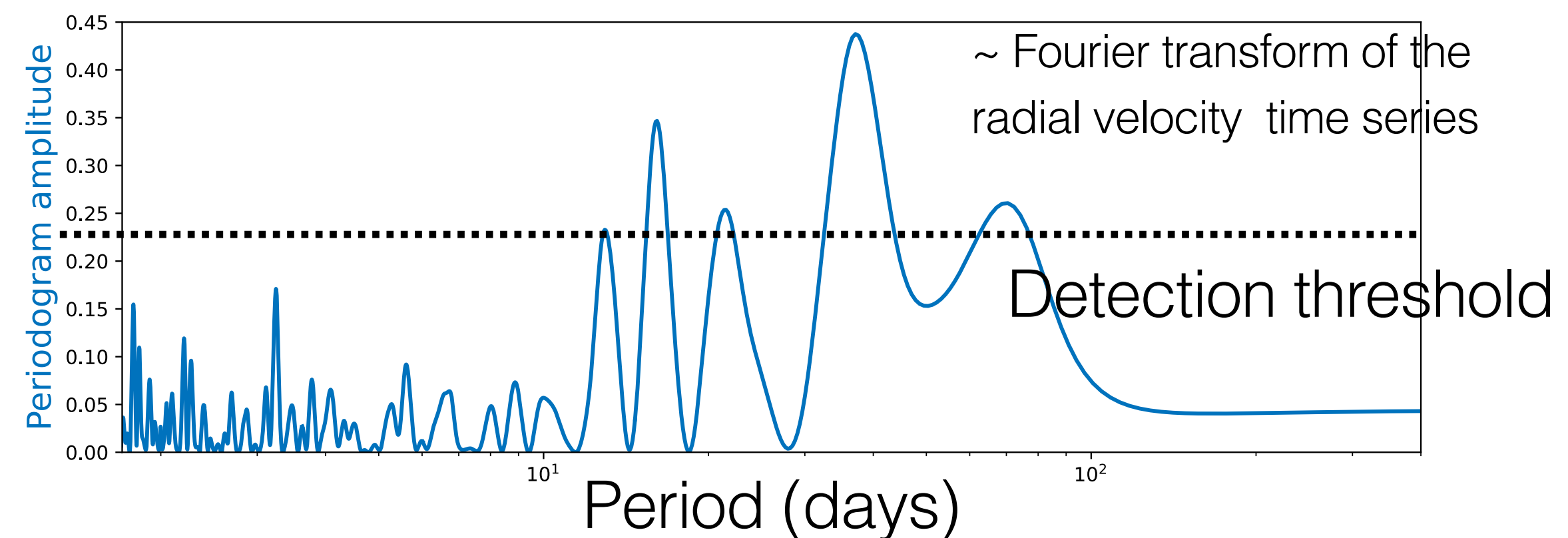
$$E_{\theta, \eta}[U\{a(\Theta_1, \dots, \Theta_n), (\theta, \eta)\}] = -n + \sum_{j=1}^n j I_{A_j} - \gamma \sum_{k=n+1}^{n_{max}} (k-n)p(k|y)$$

$$E_{\theta, \eta}[U\{a(\Theta_1, \dots, \Theta_n), (\theta, \eta)\}] = -n + \sum_{j=1}^n j I_{A_j} - \gamma \sum_{k=n+1}^{n_{max}} (k-n)p(k|y)$$

Number of true detections



Number of false detections

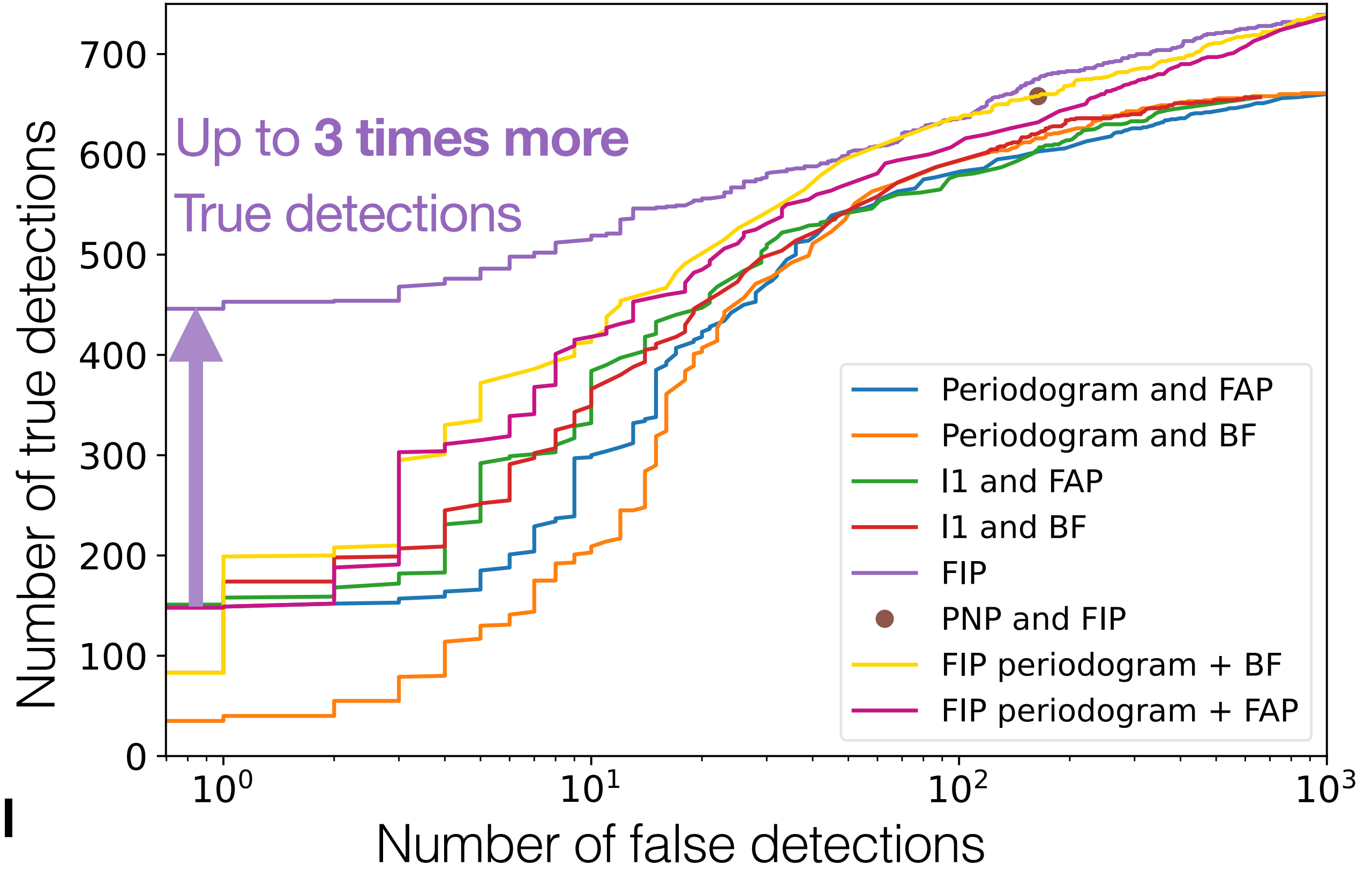


An optimal exoplanet detection criterion

- **Mathematical proof of optimality** of a new detection criterion called « True inclusion probability » (TIP)
 - Optimal in a general case
- Hara et al. 2023, Annals of Applied Statistics (in revision)
 Hara, Unger, Delisle, Díaz, Ségransan 2022

Bayes factors and FAPs

- Optimal?
 —→ **New criterion demonstrably optimal**
- Do not encode where the planet is
 —→ **Encoded in new criterion**
- Are not defined on a very intuitive scale
 —→ **New criterion is an actual probability**



In a collection of independent detections made with TIP 99%, on average 99% are correct

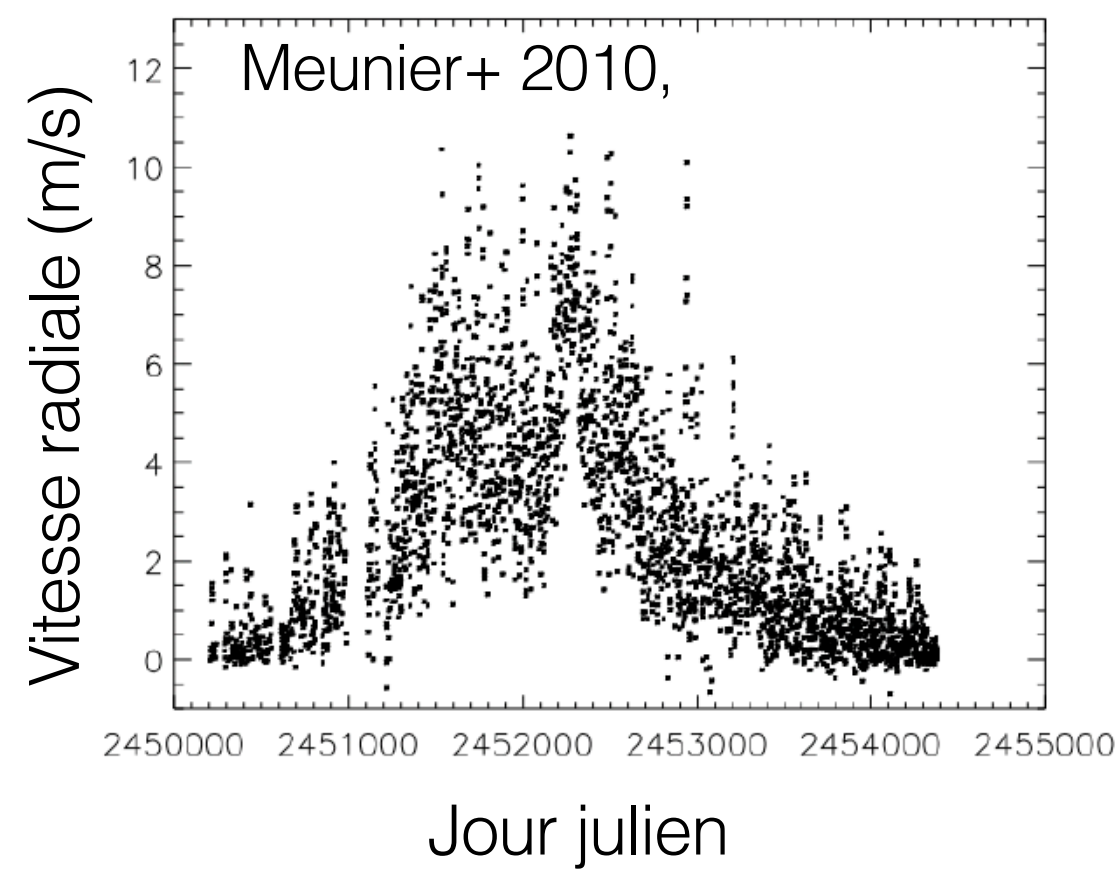
90%	90%
50%	50%

Understanding stellar variability to correct it

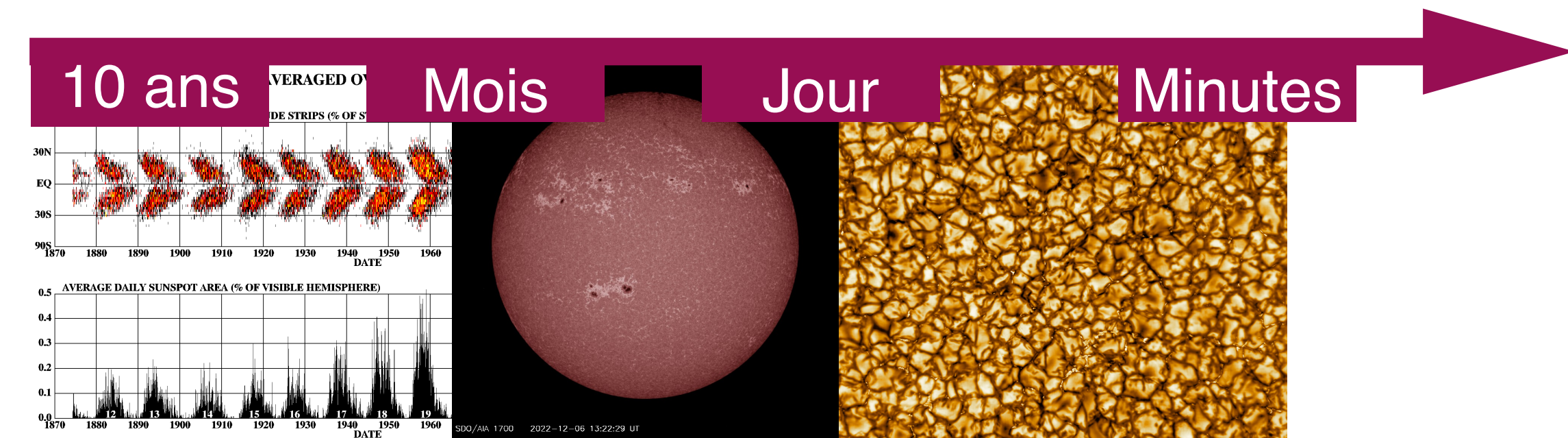
Detailed physical models
Analysis of solar data

?

Data analysed with a statistical model supposed Gaussian and stationary with qualitative parameters



Numerous processes at
Different timescales



Magnetic activity

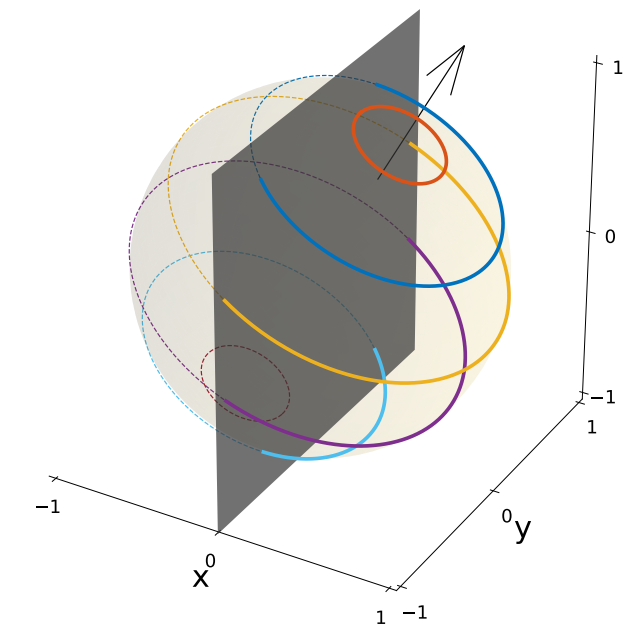
p.ex. Meunier+ 2010, 2012, 2019
Boisse+ 2012, Dumusque+ 2014,
Haywood+ 2016, Al Moulla 2023

Granulation and
super granulation

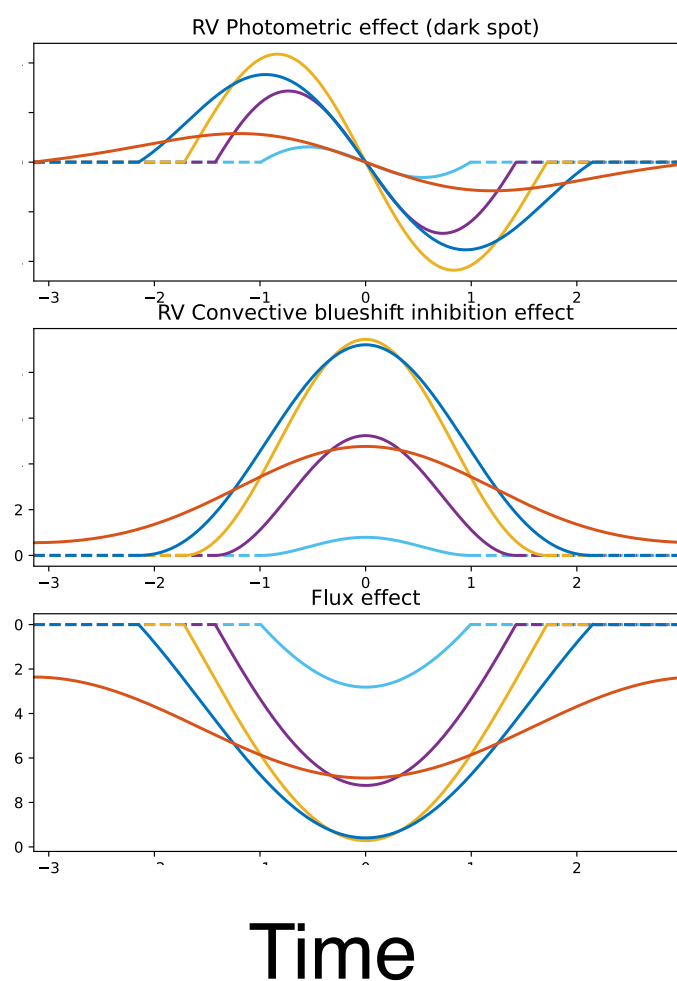
Cegla+2013, 2019
Dravins+2021,

+ oscillations,
winds,
gravitational
redshift...

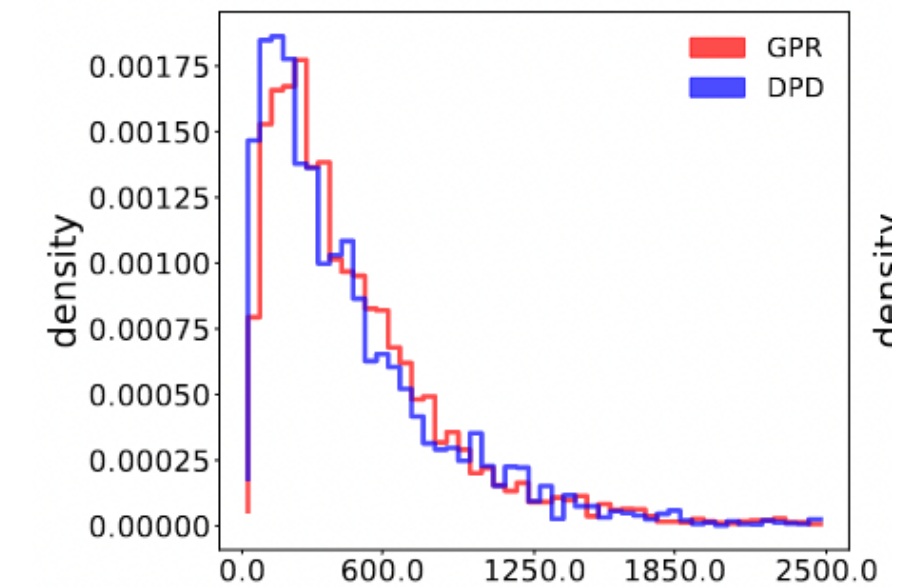
Build the statistical model from a physical one



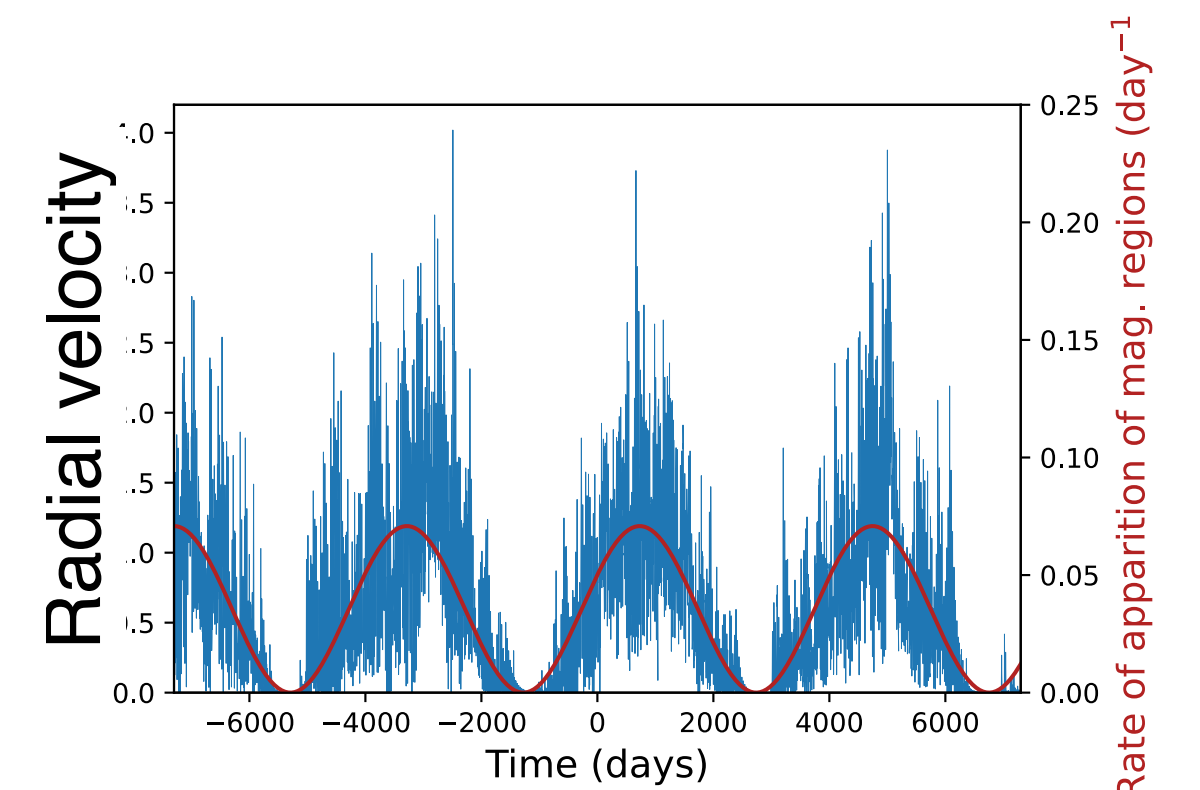
Photometry Radial velocity



Hara & Delisle 2023



Area of spots
(μ Hemisphere)

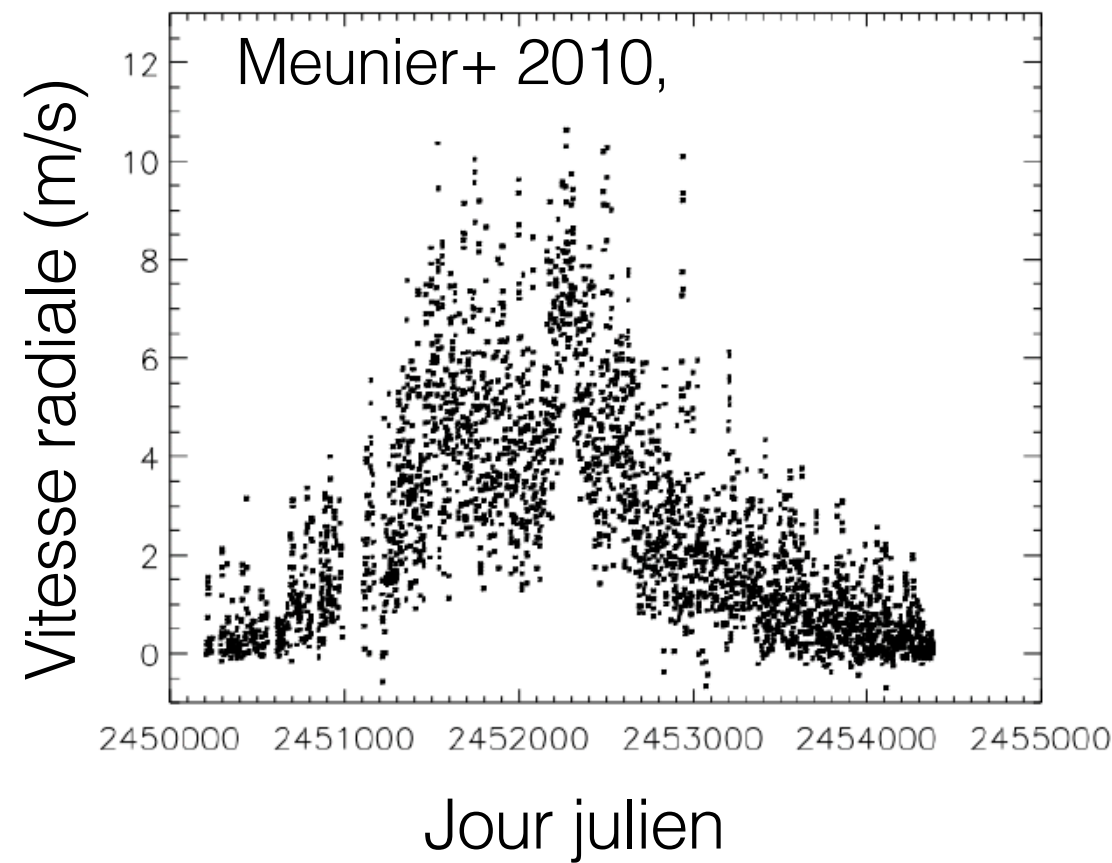


Understanding stellar variability to correct it

Detailed physical models
Analysis of solar data

?

Building the statistical model



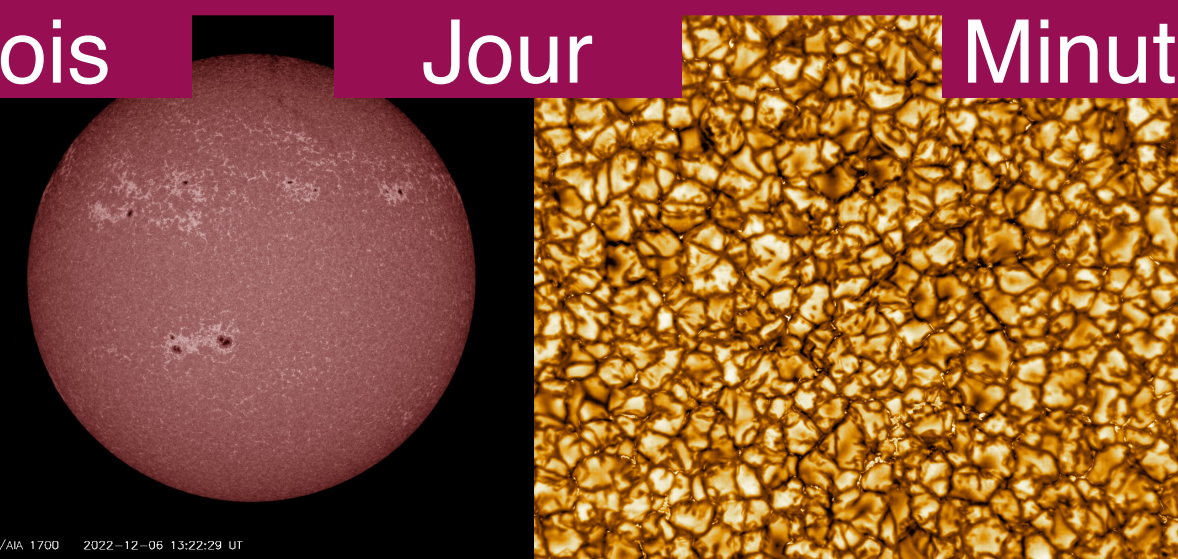
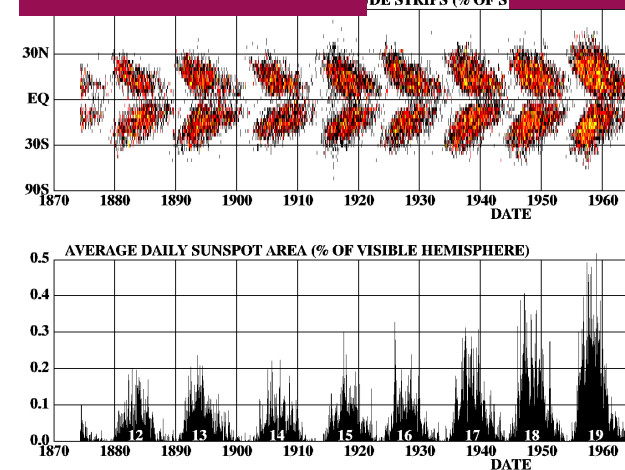
Numerous processes at
Different timescales

10 ans

Mois

Jour

Minutes



Magnetic activity

p.ex. Meunier+ 2010, 2012, 2019
Boisse+ 2012, Dumusque+ 2014,
Haywood+ 2016, Al Moulla 2023

Granulation and
super granulation

Cegla+2013, 2019
Dravins+2021,

+ oscillations,
winds,
gravitational
redshift...

Appendix A: Covariance and cumulants

Appendix A.1: Covariance and cumulants
In this appendix, we consider the covariance and autocovariance functions of a random process $X(t)$. These functions are defined as follows: $Cov(X(t), X(t')) = \langle X(t)X(t') \rangle - \langle X(t) \rangle \langle X(t') \rangle$ and $Cov(X(t), X(t')) = \langle X(t)X(t') \rangle - \langle X(t) \rangle \langle X(t') \rangle$. The covariance function is a symmetric function of the time difference $\tau = t - t'$. The autocovariance function is a symmetric function of the time difference $\tau = t - t'$. The power spectrum density (PSD) is the Fourier transform of the autocovariance function. The PSD is a symmetric function of the frequency ω . The PSD is a symmetric function of the frequency ω . The PSD is a symmetric function of the frequency ω .

Appendix A.2: Cumulants
The cumulants of a random process $X(t)$ are defined as follows: $\kappa_1 = \langle X(t) \rangle$, $\kappa_2 = Cov(X(t), X(t'))$, $\kappa_3 = \langle X(t)X(t')X(t'') \rangle - 3\langle X(t) \rangle \langle X(t') \rangle \langle X(t'') \rangle - 2\langle X(t) \rangle \langle X(t')X(t'') \rangle - 2\langle X(t') \rangle \langle X(t)X(t'') \rangle - 2\langle X(t'') \rangle \langle X(t)X(t') \rangle + 6\langle X(t) \rangle \langle X(t') \rangle \langle X(t'') \rangle$. The cumulants are symmetric functions of the time differences. The cumulants are symmetric functions of the time differences. The cumulants are symmetric functions of the time differences.

Appendix B: CCF-related quantities
The cross-correlation function (CCF) is defined as $CCF(v) = \langle X(t)X(t+v) \rangle - \langle X(t) \rangle \langle X(t+v) \rangle$. The CCF is a symmetric function of the time difference v . The CCF is a symmetric function of the time difference v . The CCF is a symmetric function of the time difference v . The CCF is a symmetric function of the time difference v . The CCF is a symmetric function of the time difference v .

Appendix C: Granulation covariance
The covariance function of granulation is defined as $Cov(X(t), X(t')) = \langle X(t)X(t') \rangle - \langle X(t) \rangle \langle X(t') \rangle$. The covariance function is a symmetric function of the time difference $\tau = t - t'$. The covariance function is a symmetric function of the time difference $\tau = t - t'$. The covariance function is a symmetric function of the time difference $\tau = t - t'$. The covariance function is a symmetric function of the time difference $\tau = t - t'$. The covariance function is a symmetric function of the time difference $\tau = t - t'$.

Kernel type	rank
Sudden appearing and exp. decay	1
Asymmetric exp. (dis)appearing	2
Symmetric exp. (dis)appearing	3

Table C.1: Kernels (k^W) obtained for different window shapes (W), normalized such that $k^W(0) = 1$.

Phase shifts and faculae/spot ratios

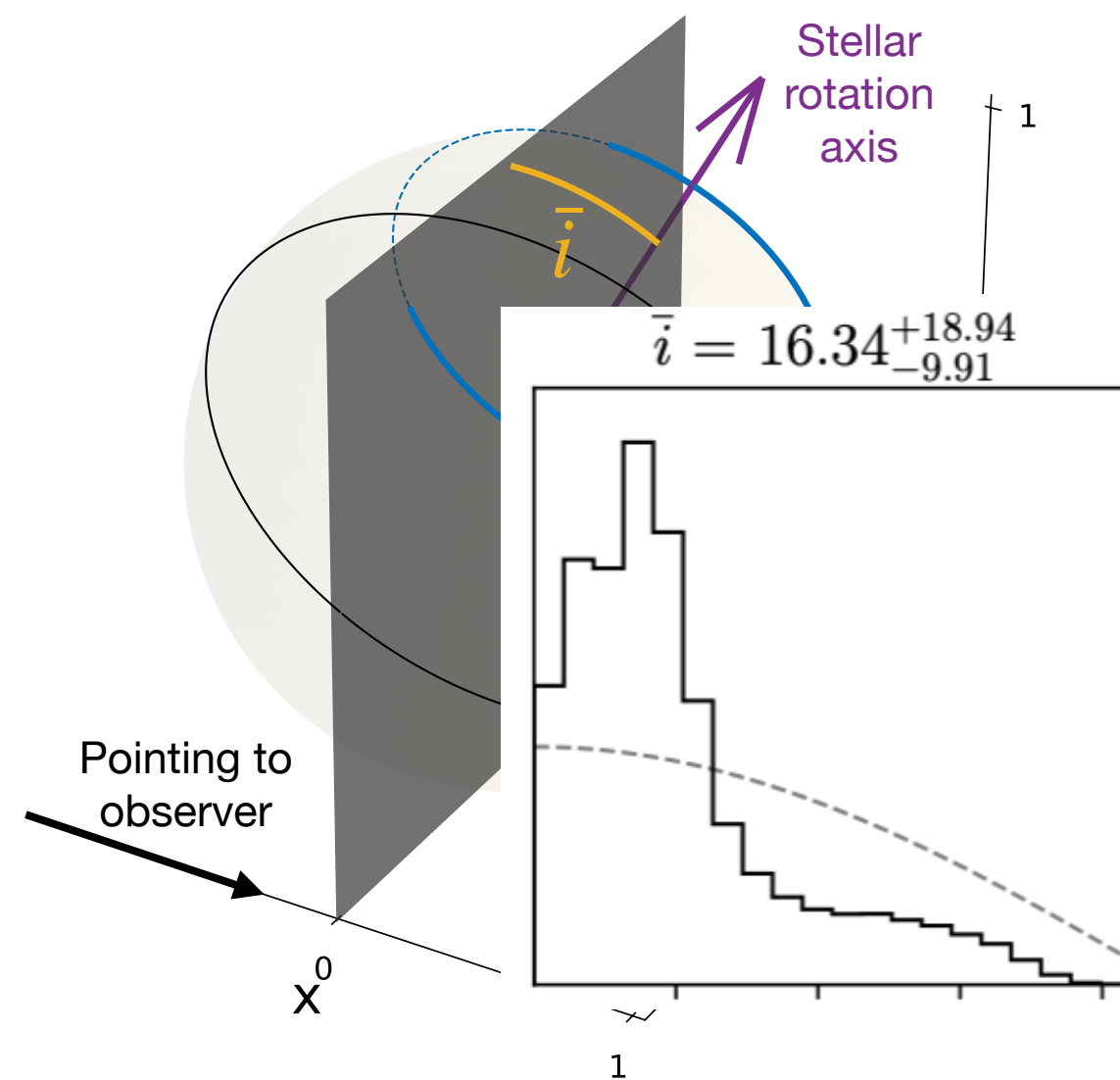
Detailed physical models
Analysis of solar data

Hara & Delisle 2023

Statistical model with
physical parameters

(In review)

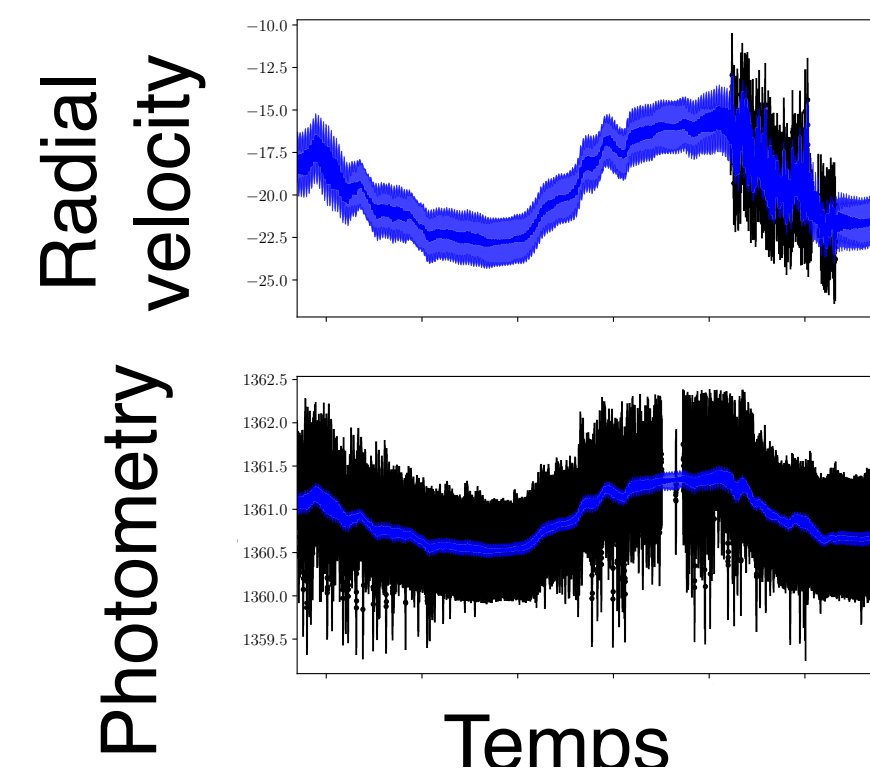
Statistical Doppler imaging



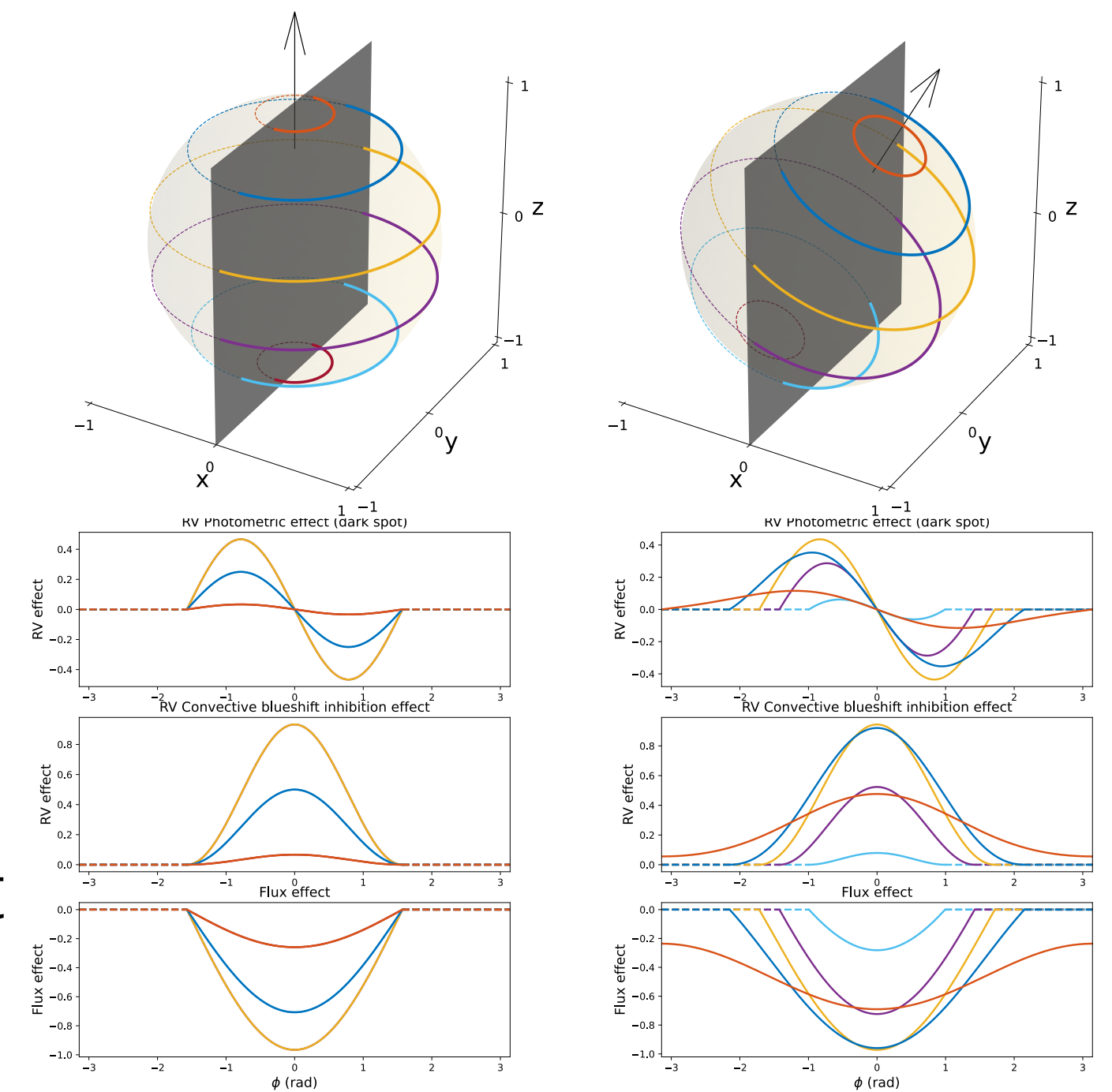
Finding the solar inclination from RVs (HARPS-N) and photometry (SORCE)

To be refined

Solar observations

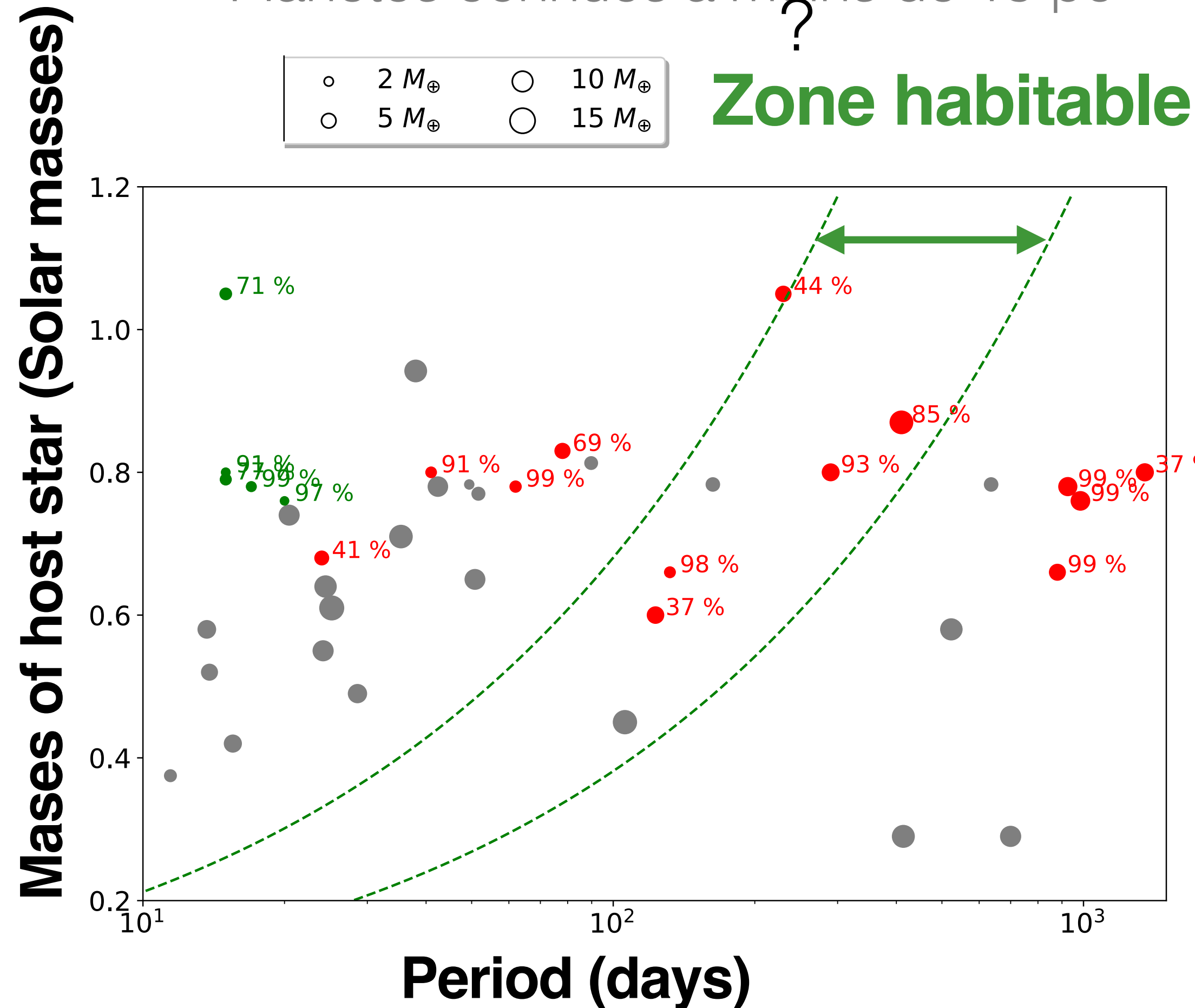


Fast,
Applicable to transit Spectroscopy

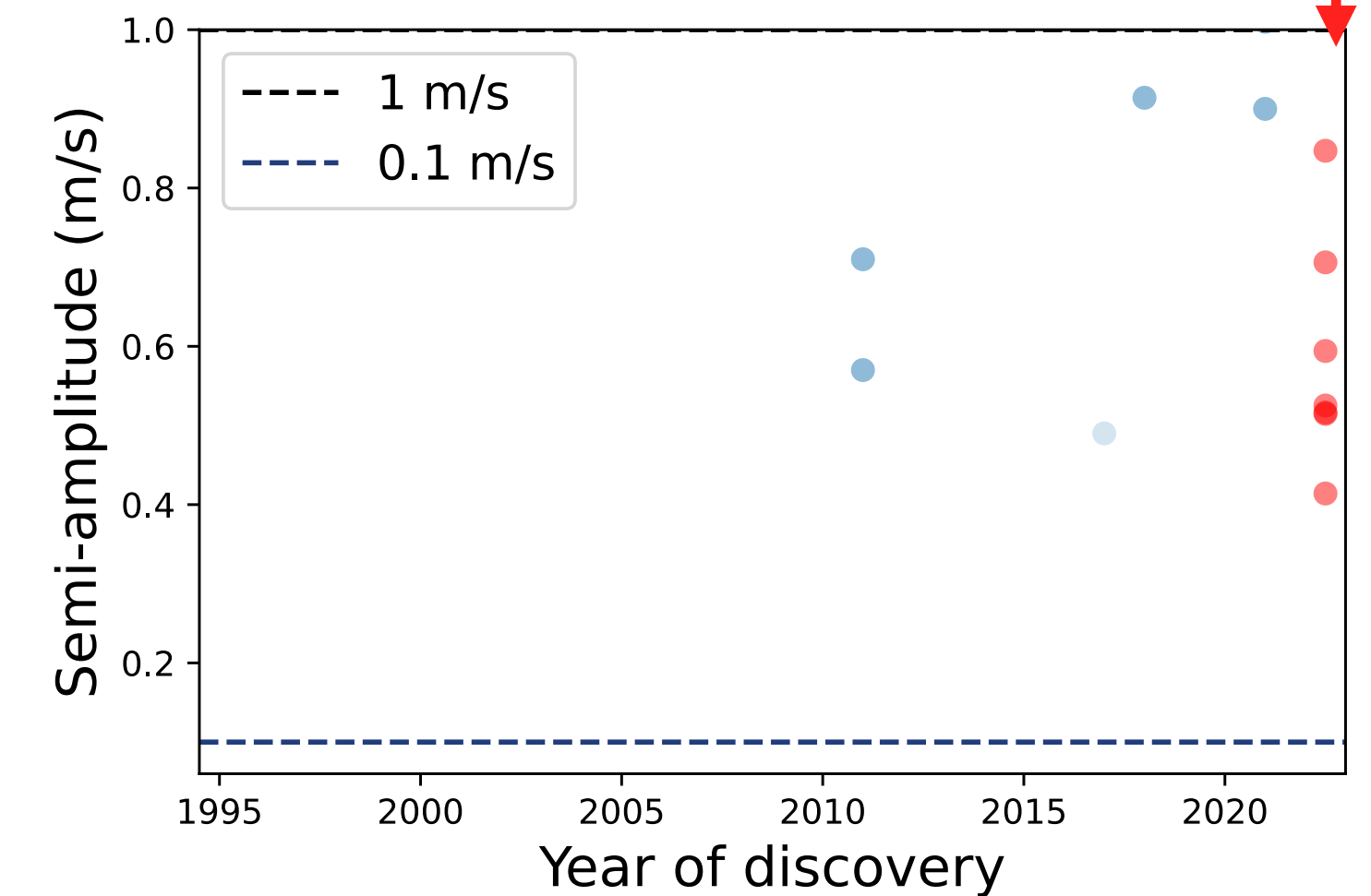
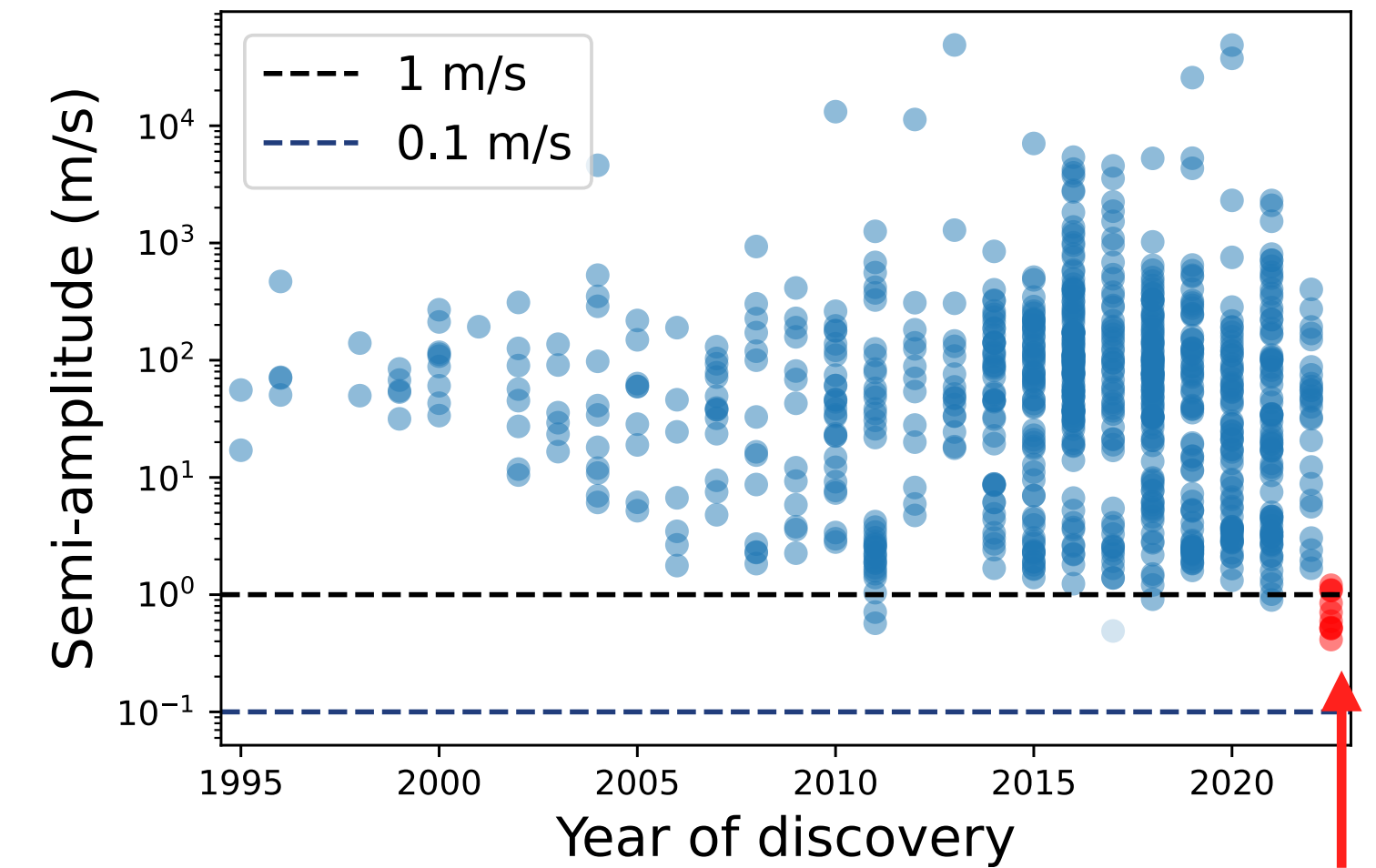


Application: super-Terres dans la zone habitable

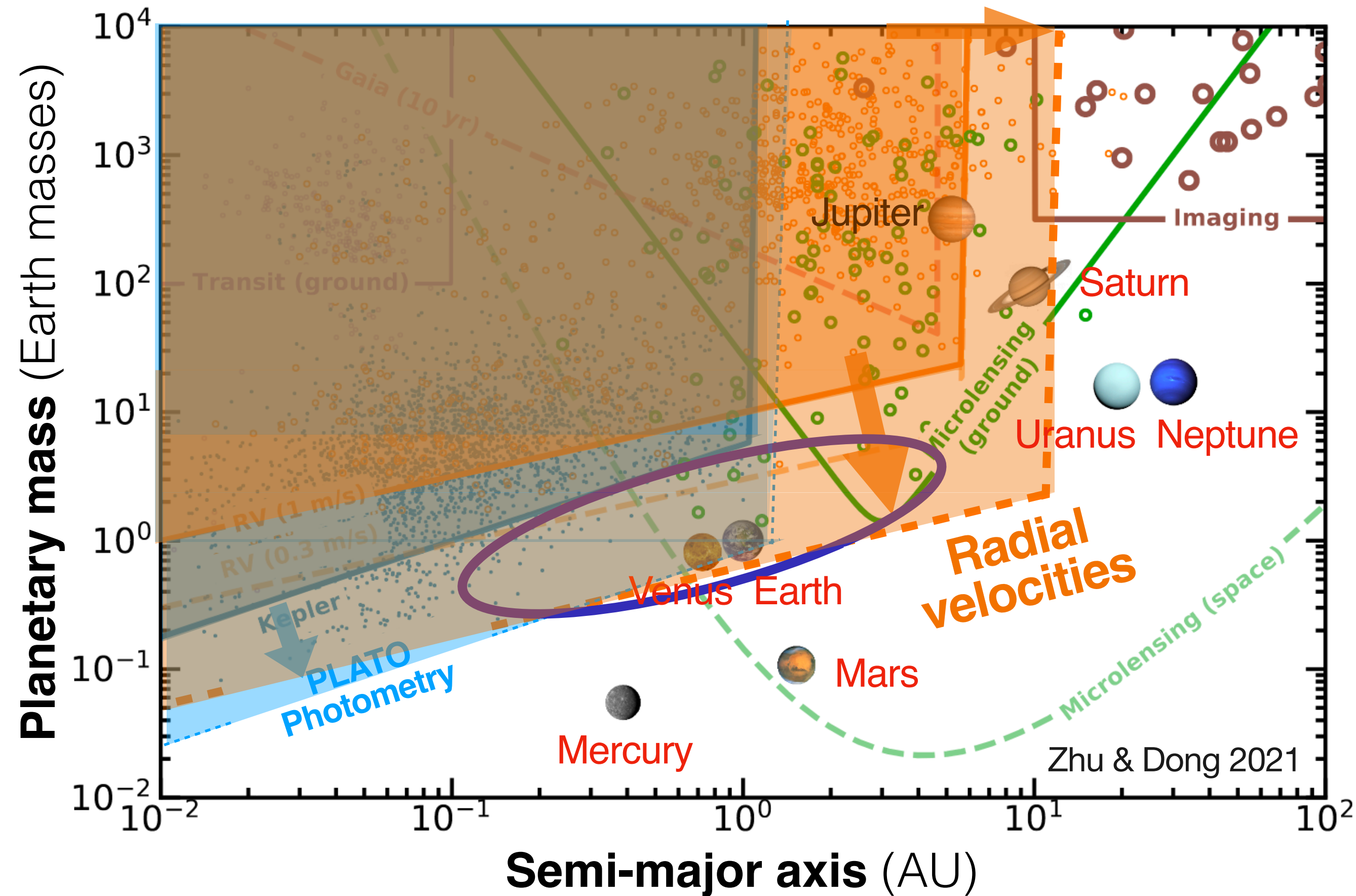
Planètes connues à moins de 15 pc



Large program HARPS (PI) 340h to confirm these detections
 (with X. Dumusque, M. Crétignier, N. Unger) Oct. 2022- Oct. 2023



Conclusion

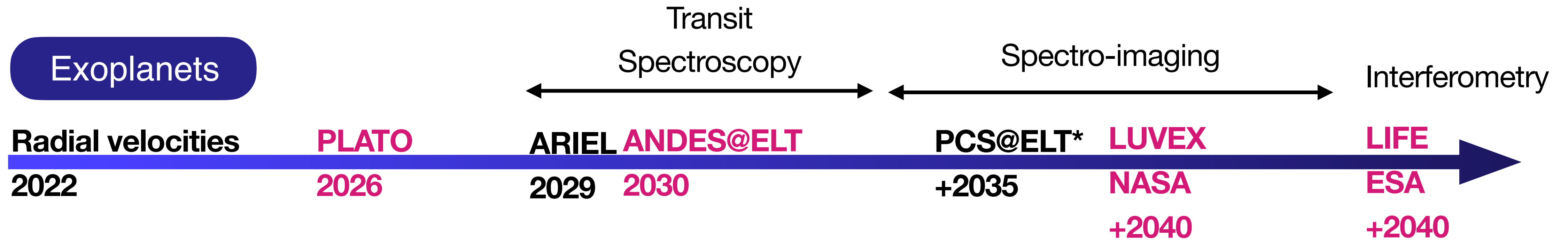
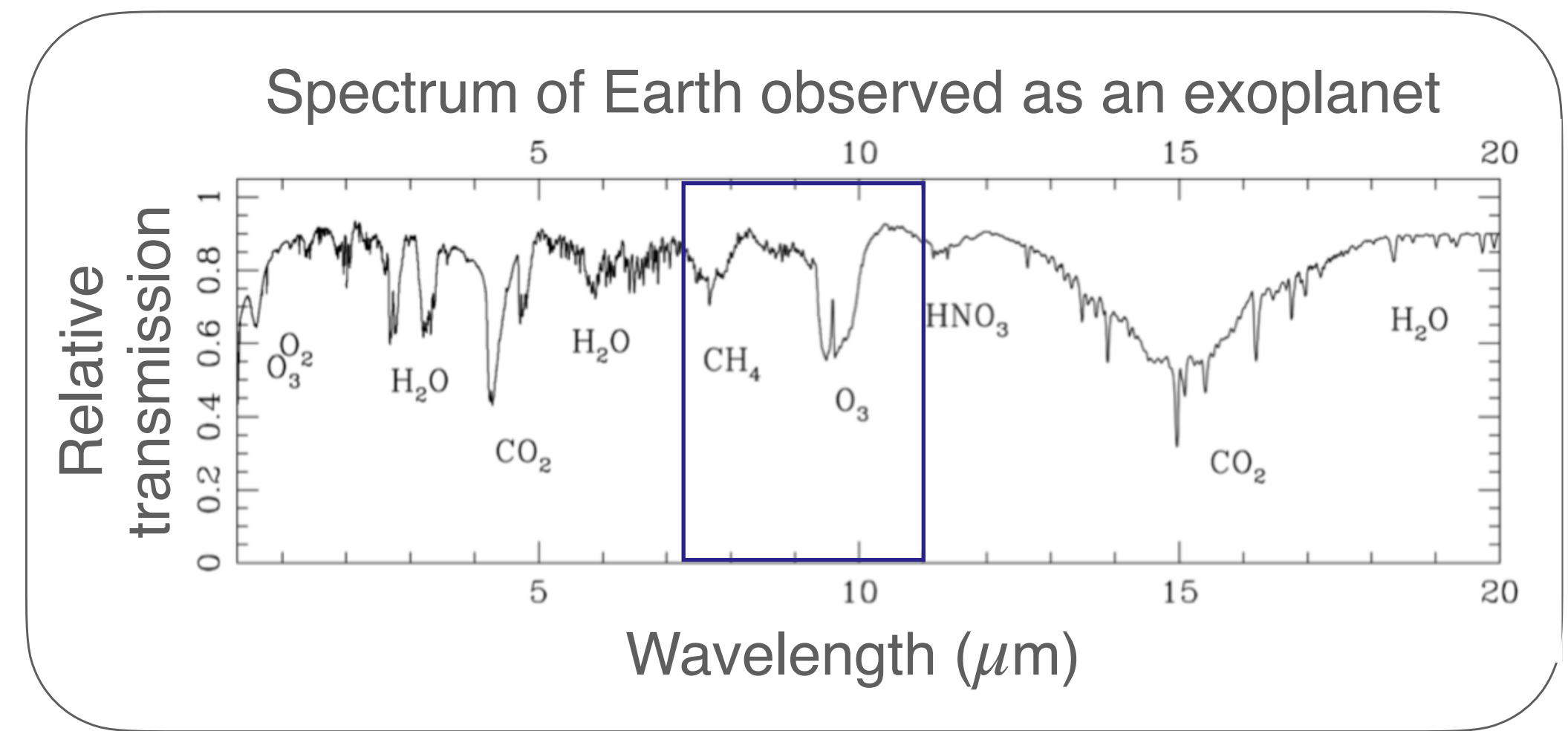


Radial velocity measurements would greatly improve the scientific yield of 2050 flagship missions

Detecting Earth like planets with radial velocities is a very difficult problem but progress is made

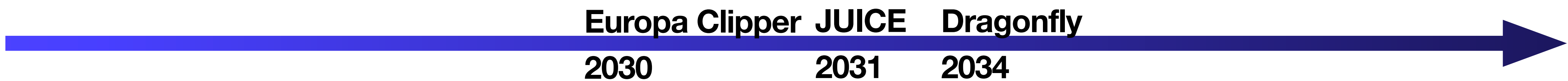
Results of the large program end of next year

Instruments for the identification of life



Lab experiments

Identification of biosignatures



Solar system

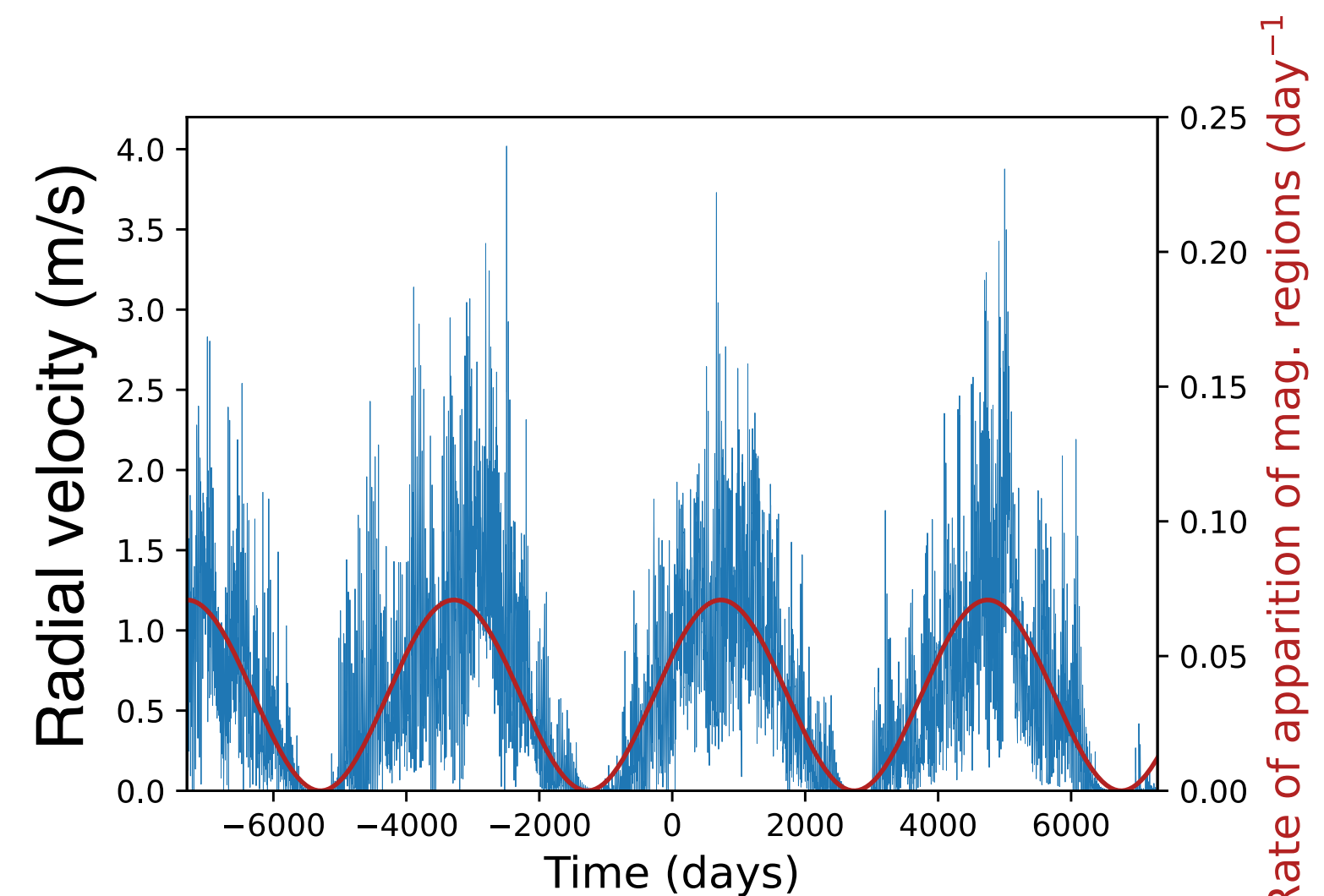
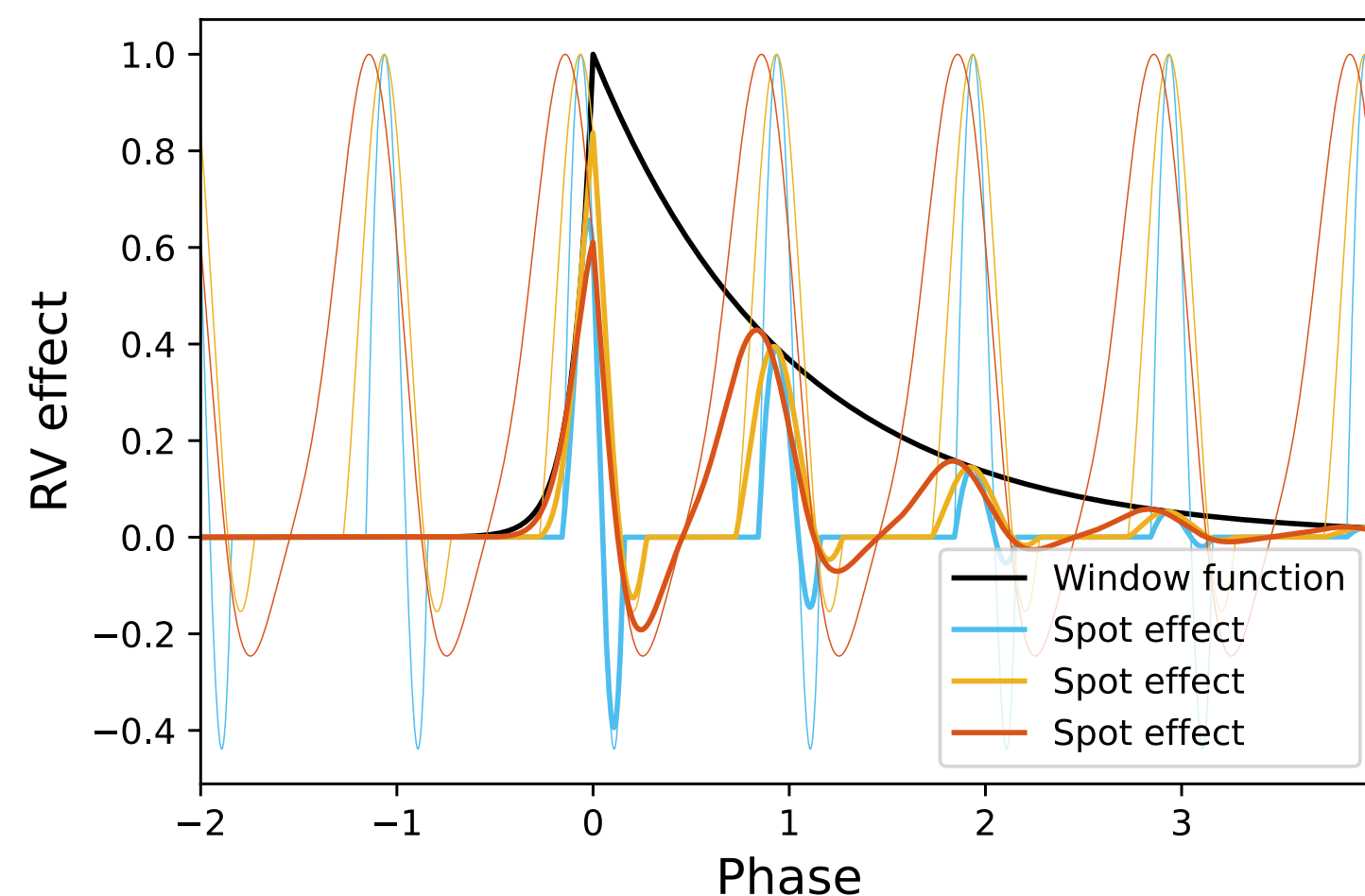
Generate realisations of the FENRIR model

(1) Draw times at which features appear according to a **variable rate** $\lambda(t)$



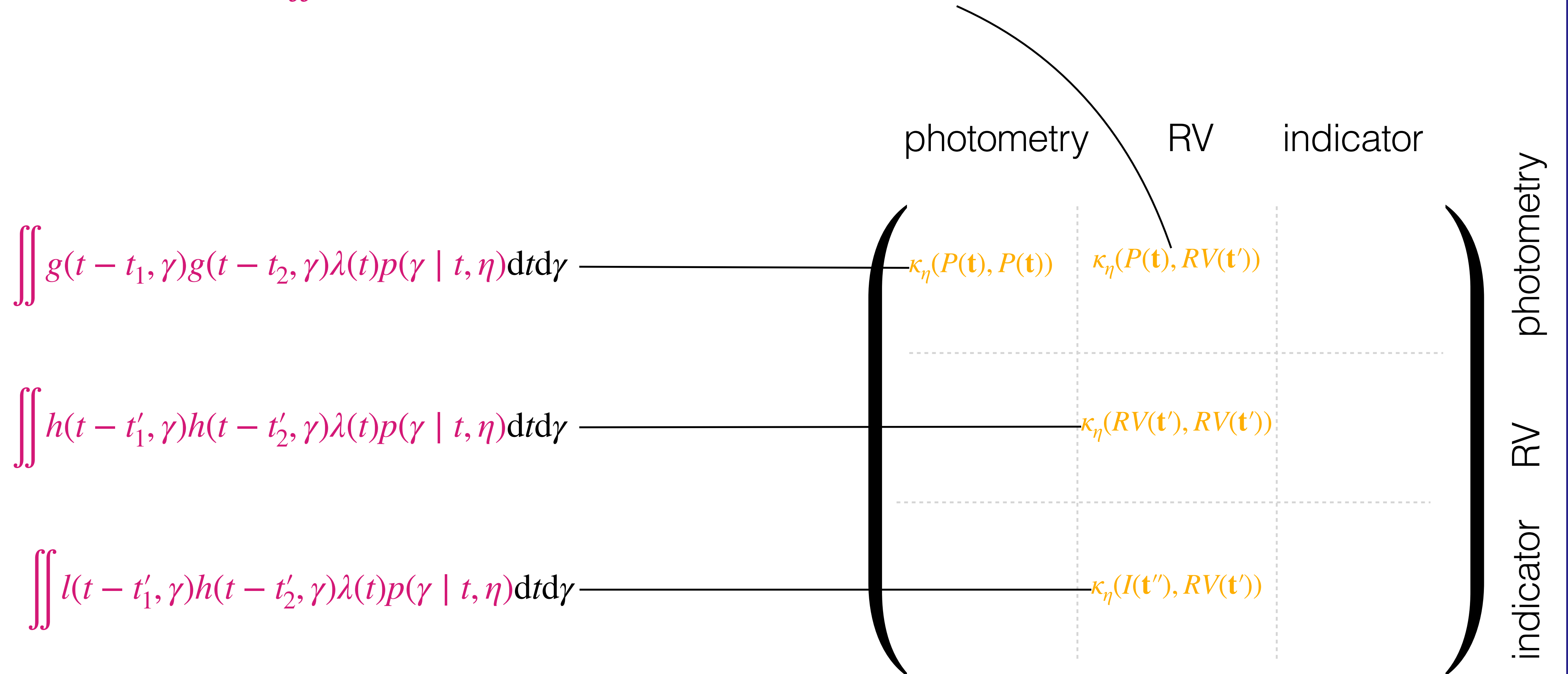
(2) If a feature appears at time t draw its parameters γ from distribution $p(\gamma | \eta, t)$
Where η are the stellar parameters

γ → (3) Add $g(t - t_1, \gamma)$ to your photometry time series and $h(t - t_1, \gamma)$ to your RV time series



From the **physical model**, compute the **covariance** of the likelihood function

$$\iint g(t - t_1, \gamma)h(t - t'_1, \gamma)\lambda(t)p(\gamma | t, \eta)dtd\gamma =$$



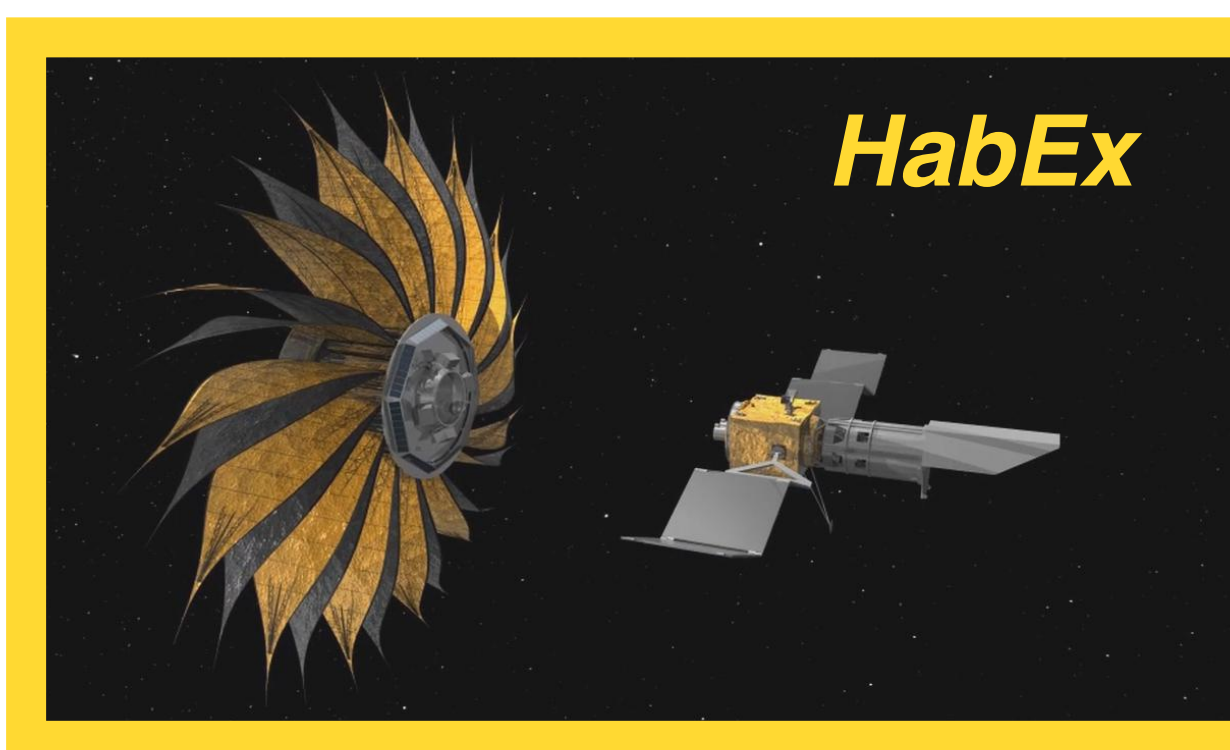
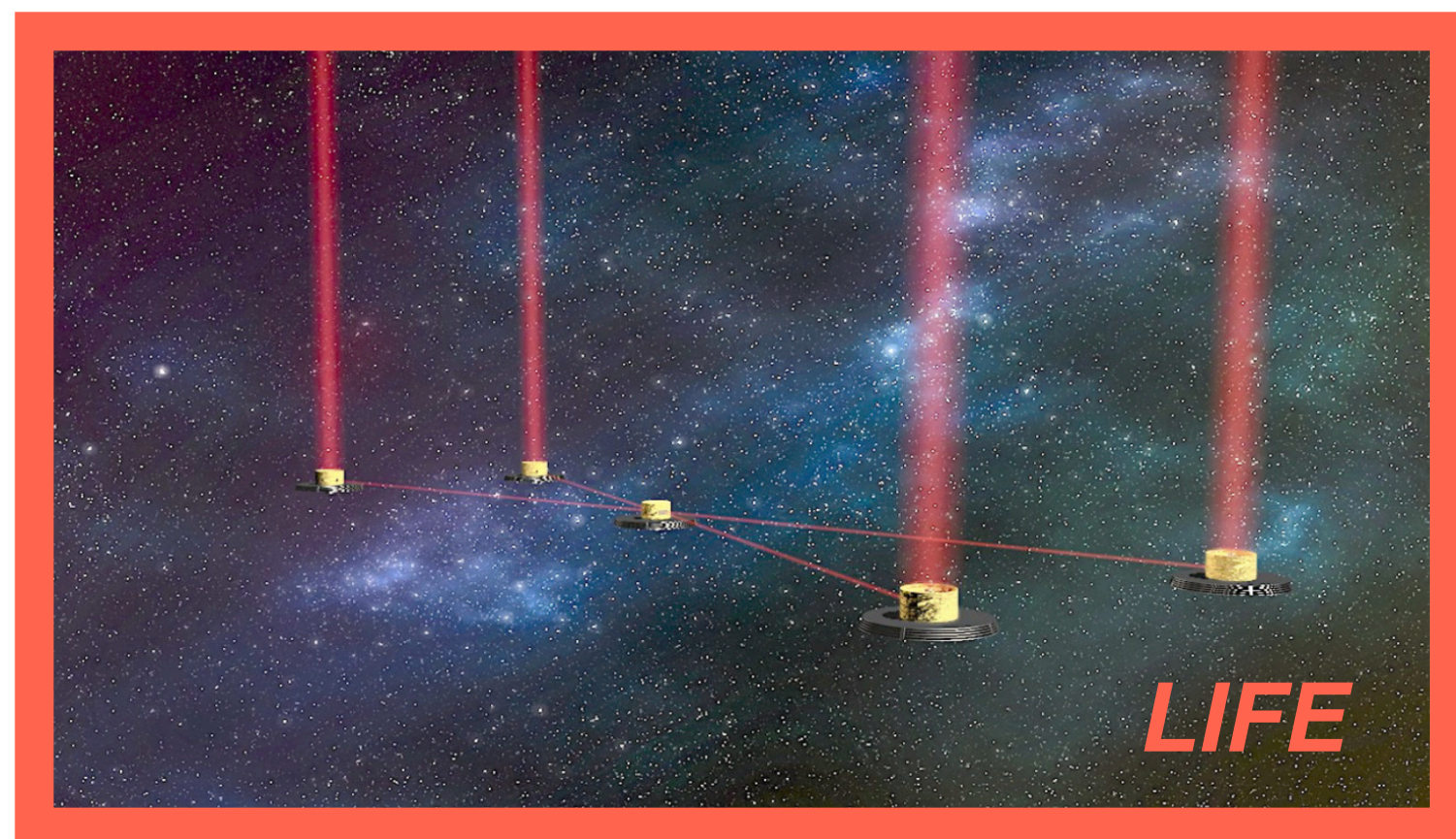
Objectives

Exoplanet observations

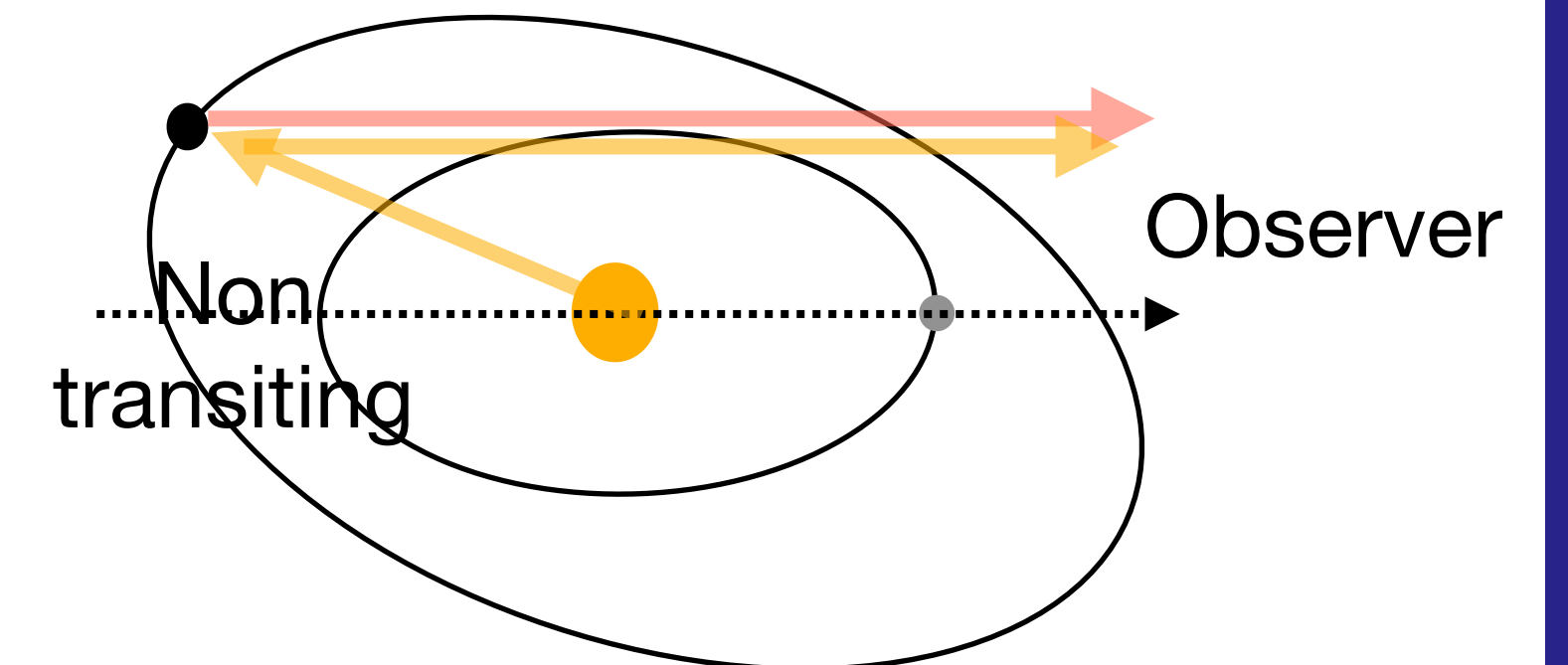
Characterizing Earth-analogs
Detecting life outside of Earth

Characterization of exoplanetary atmospheres

Detecting life outside of Earth



Spectro-imaging: LIFE (ESA), LUVEX (NASA), JWST), PCS@ELT



Thermal emission:
LIFE, ESA

Reflected light:
HWO, NASA

Objectives

Exoplanet observations

Characterizing Earth-analogs
Detecting life outside of Earth

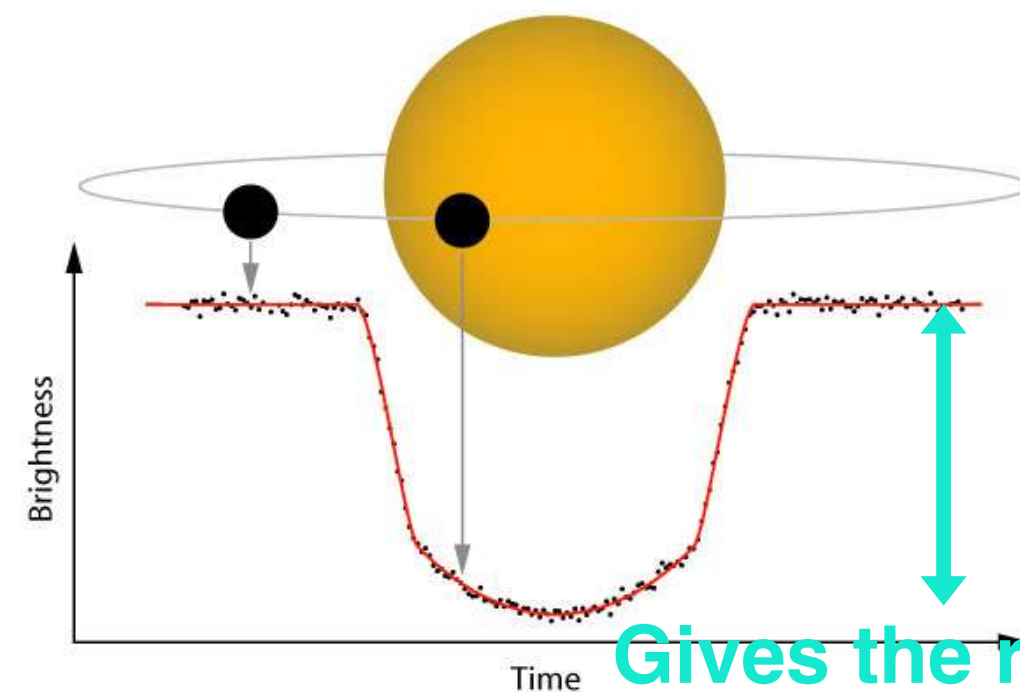
Characterization of exoplanetary atmospheres (chemical composition, reflectance, variability)

Detecting Earth-like planets, mass, radius, and orbits



Require targets $\lesssim 20$ parsec

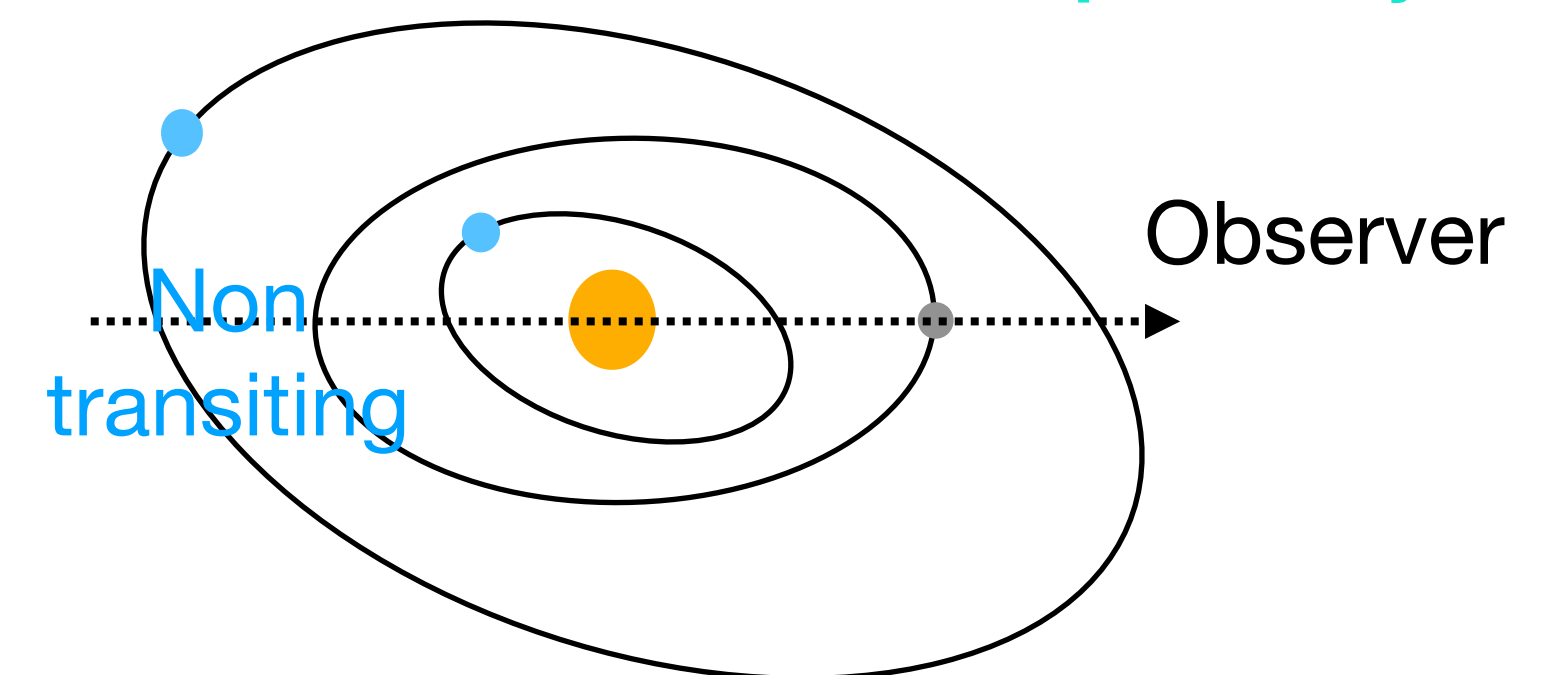
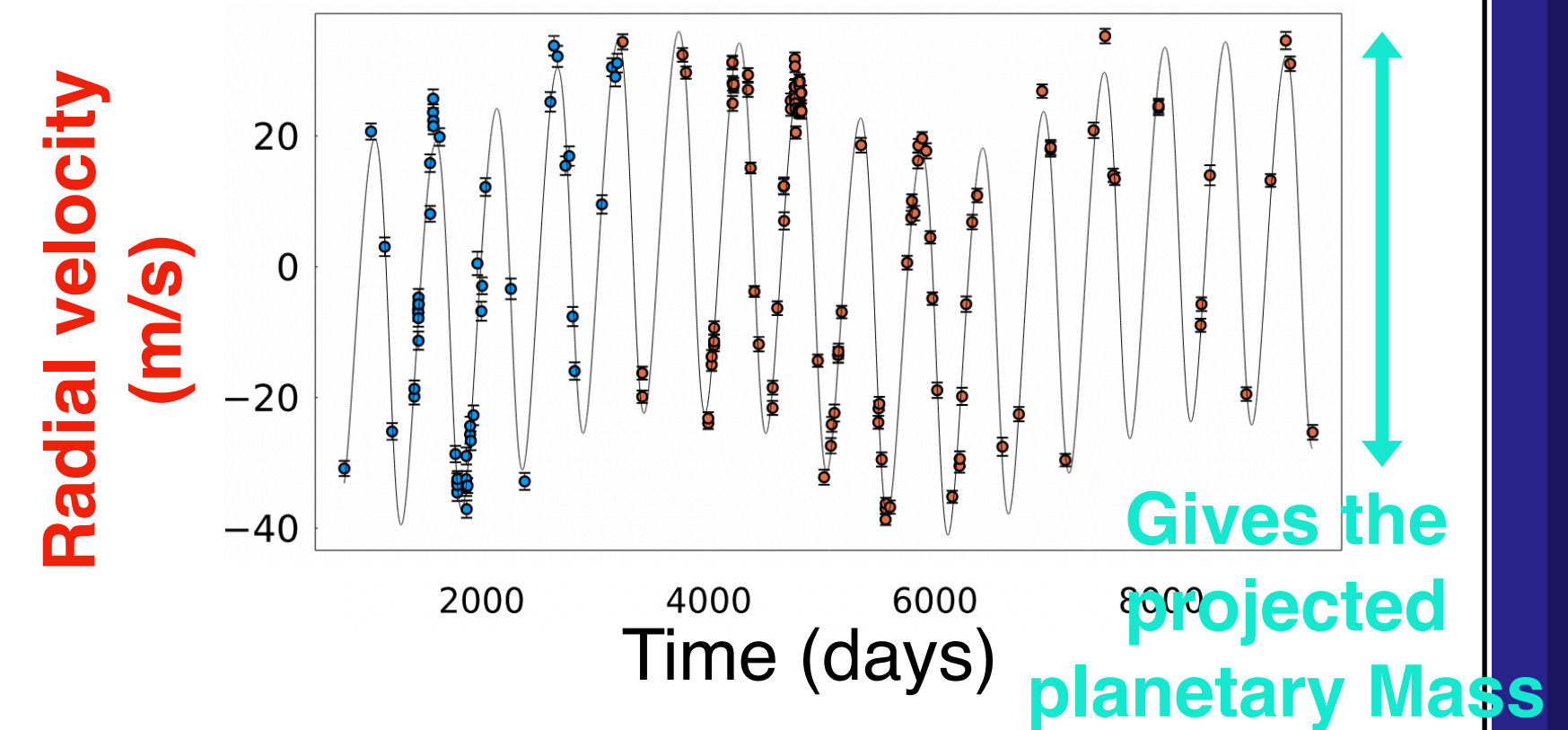
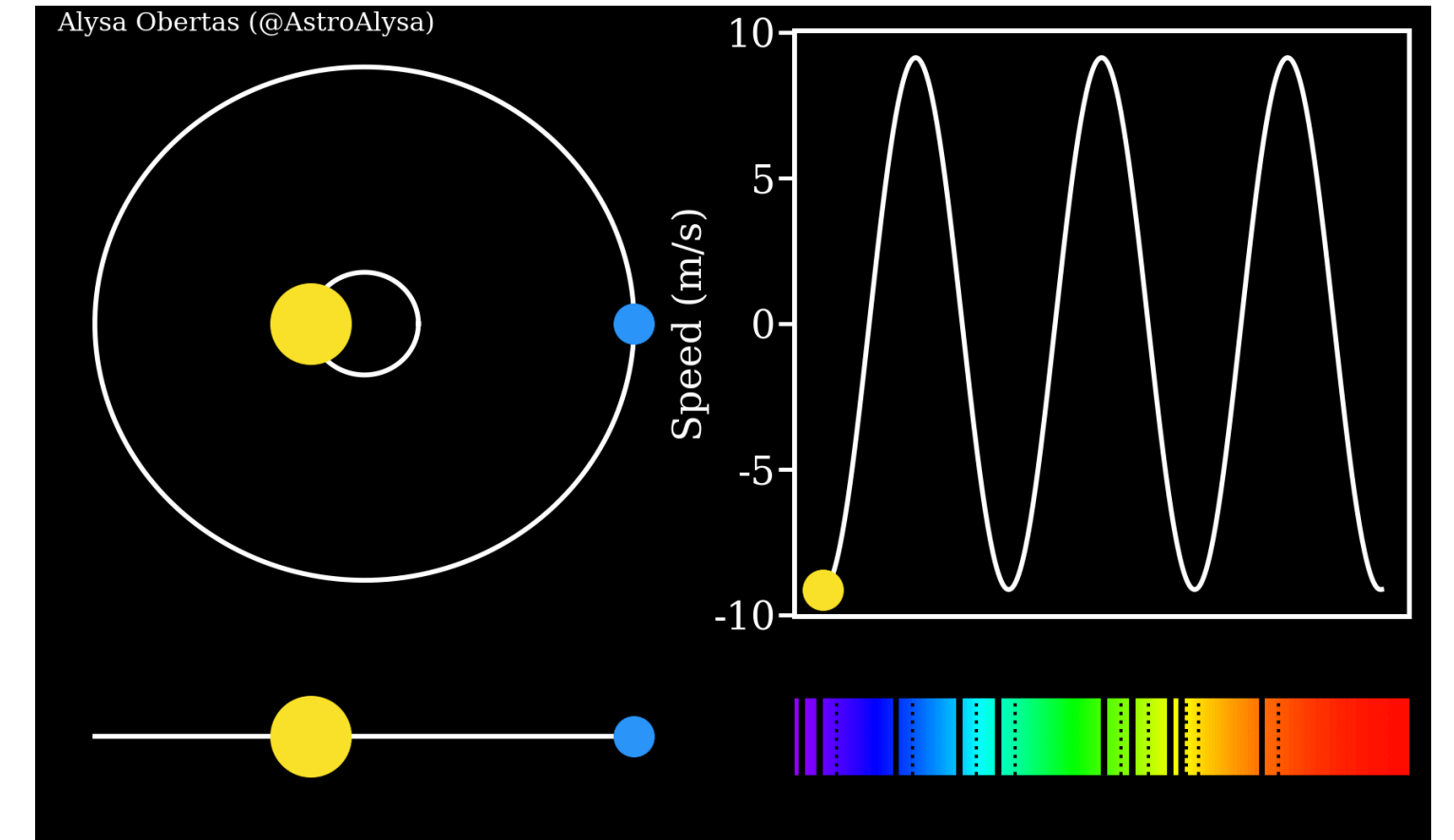
Transits



Gives the ratio
Of planetary and
Stellar radii

Requires very
Specific system
geometry

Radial velocity (RV)

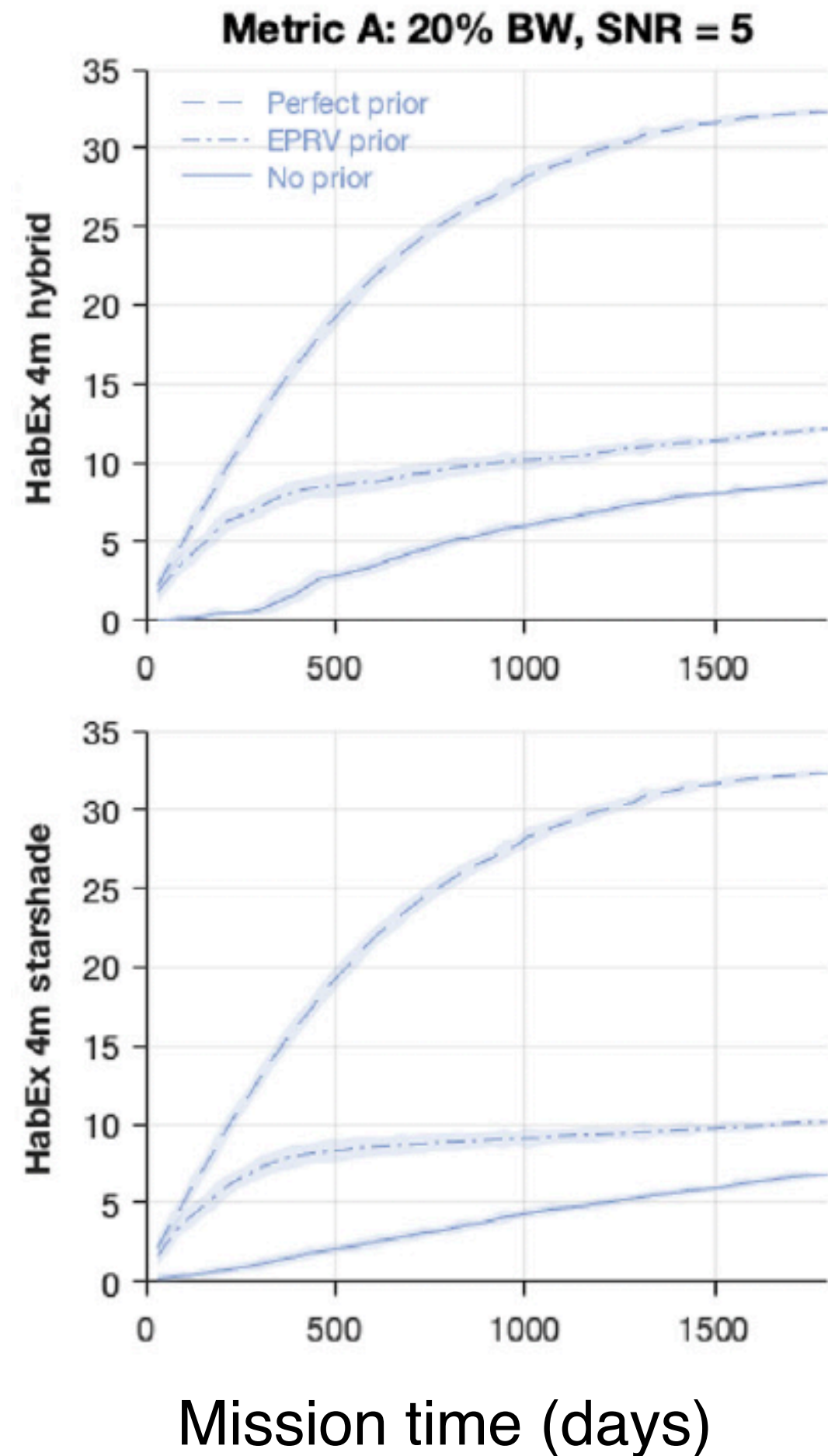


Radial velocities are key



Mass is essential to interpret spectroscopic observations of atmospheres (Batalha et al. 2017, 2019)

Number of Earth analogs characterised



Morgan et al. 2021

The yield of atmospheric characterization is greatly improved if Earth like planets are detected in advance, **require targets $\lesssim 20$ parsec**

- + Results obtained faster, more time for characterisation
- + Mitigates the risk of the missions
- + True for other kinds of planets

Radial velocity (RV)

