

# Upstream knowledge of targets with radial velocities

**Nathan Hara**  
Laboratoire d'astrophysique de Marseille

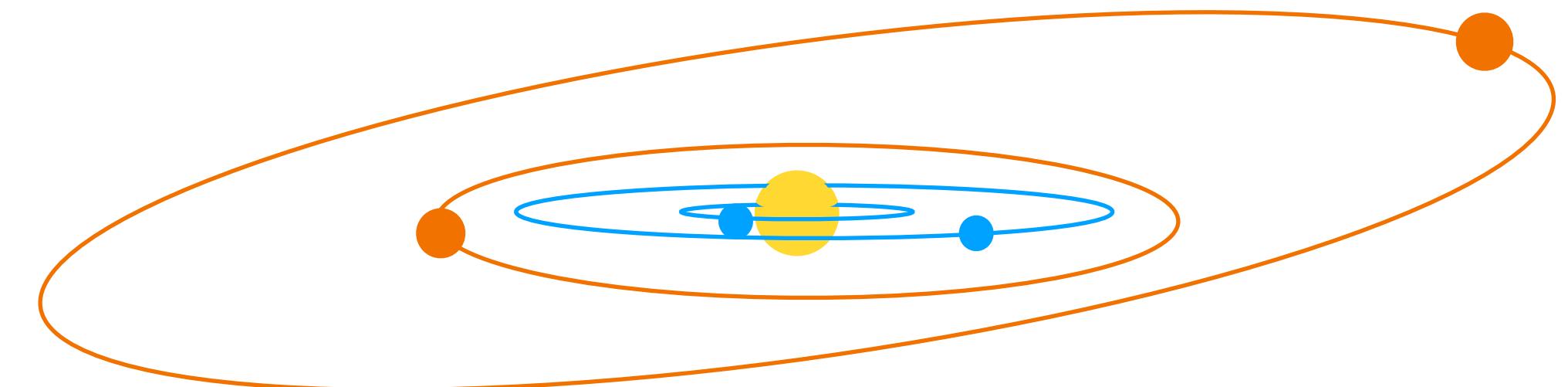
HALO workshop  
5 December 2024

# Outline

(1) The output of *HWO*, *PCS* and *LIFE* would be greatly enhanced by a pre-detection of suitable targets (precursor survey)

**Transits** would help very marginally

(2) **Radial velocity** or **astrometry** could do it

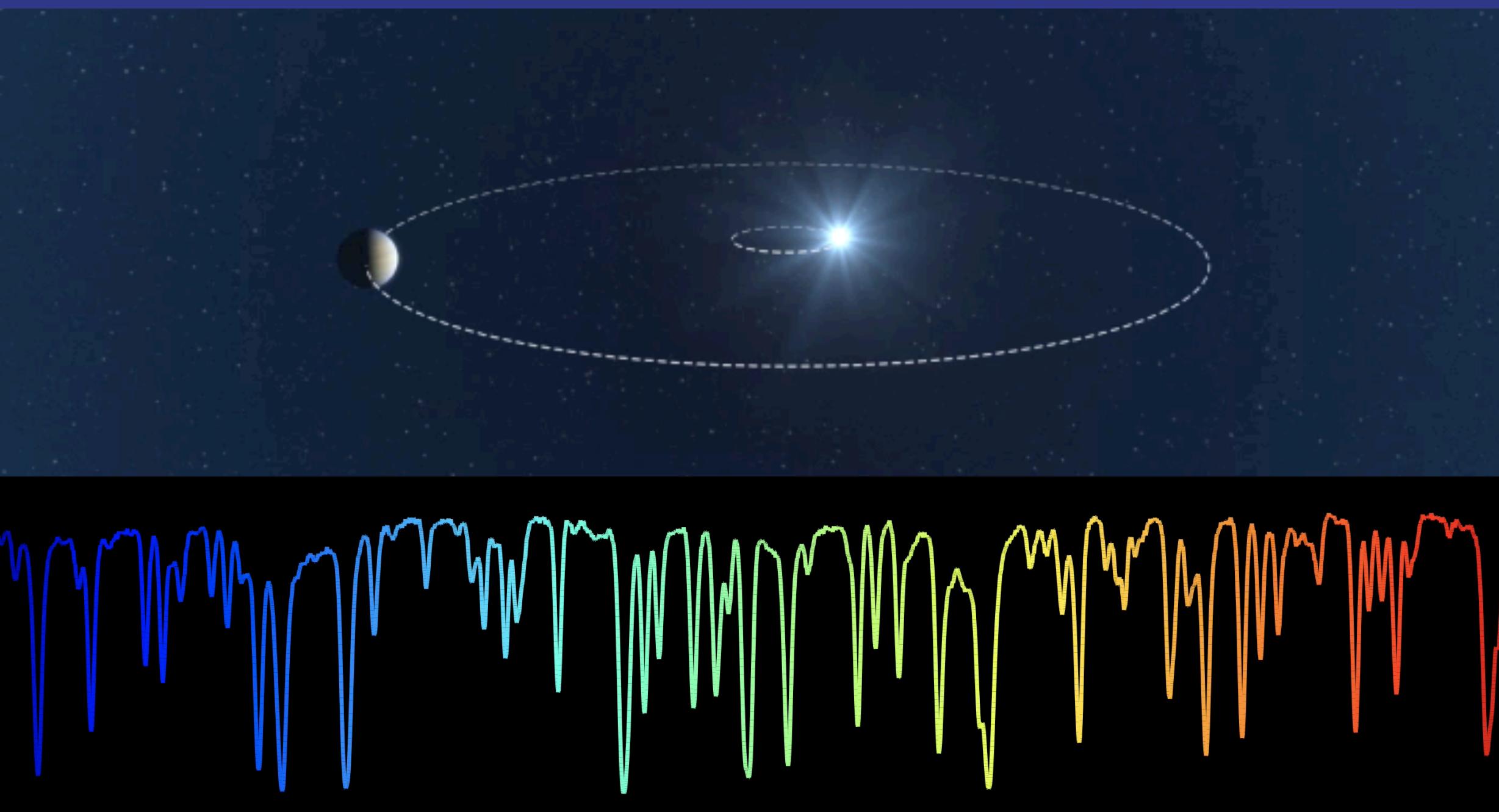


(3) Earth-analog detection with radial velocity is hindered by stellar variability and systematics: mostly not a hardware problem, needs **new data analysis methods** and appropriate **observation strategy**

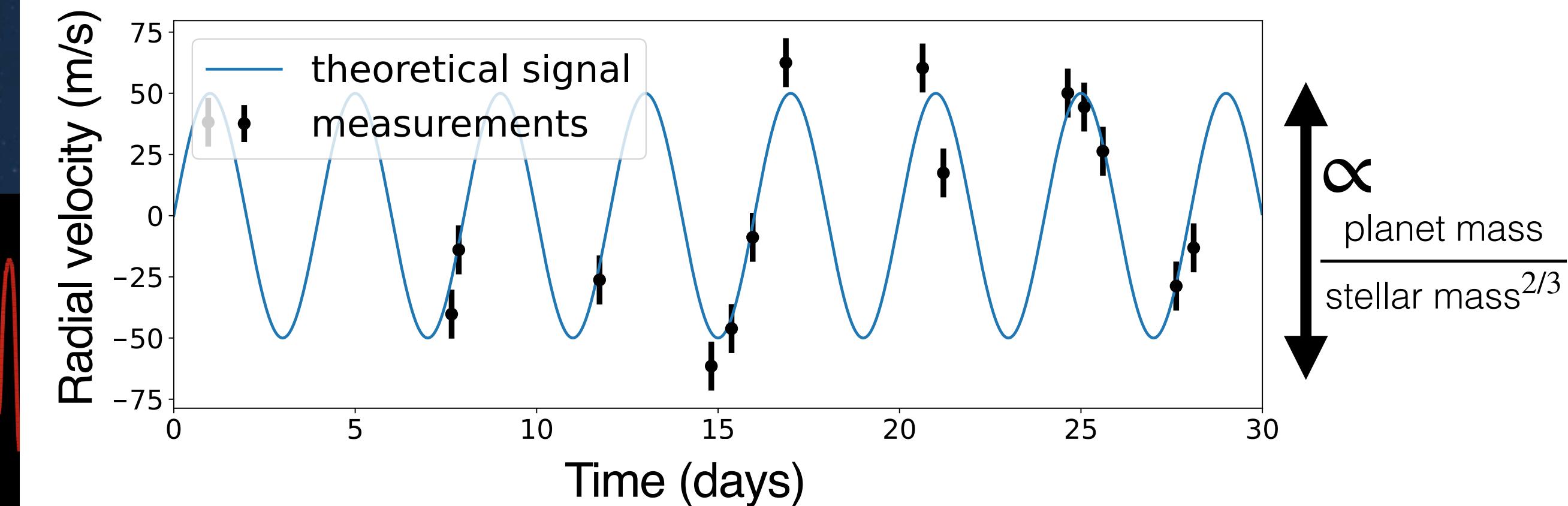
Detection with astrometry would require a new space mission (concept in study)

(4) A community effort is needed to make precursor surveys happen

# The Radial Velocity (RV) technique

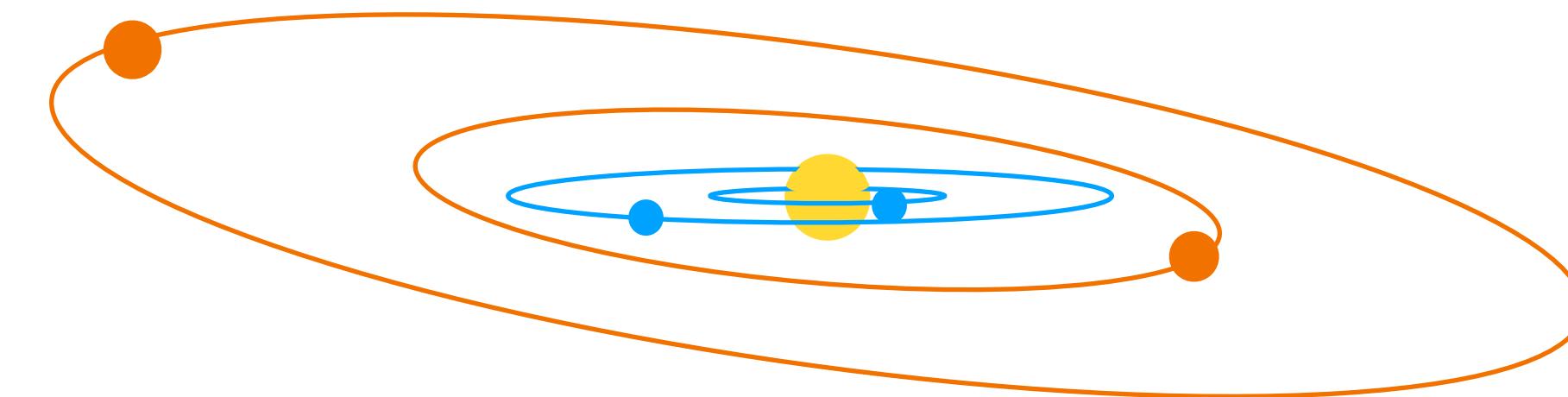


Radial velocity (RV) = velocity of the star projected onto the line of sight



## Radial velocities (RVs) are essential

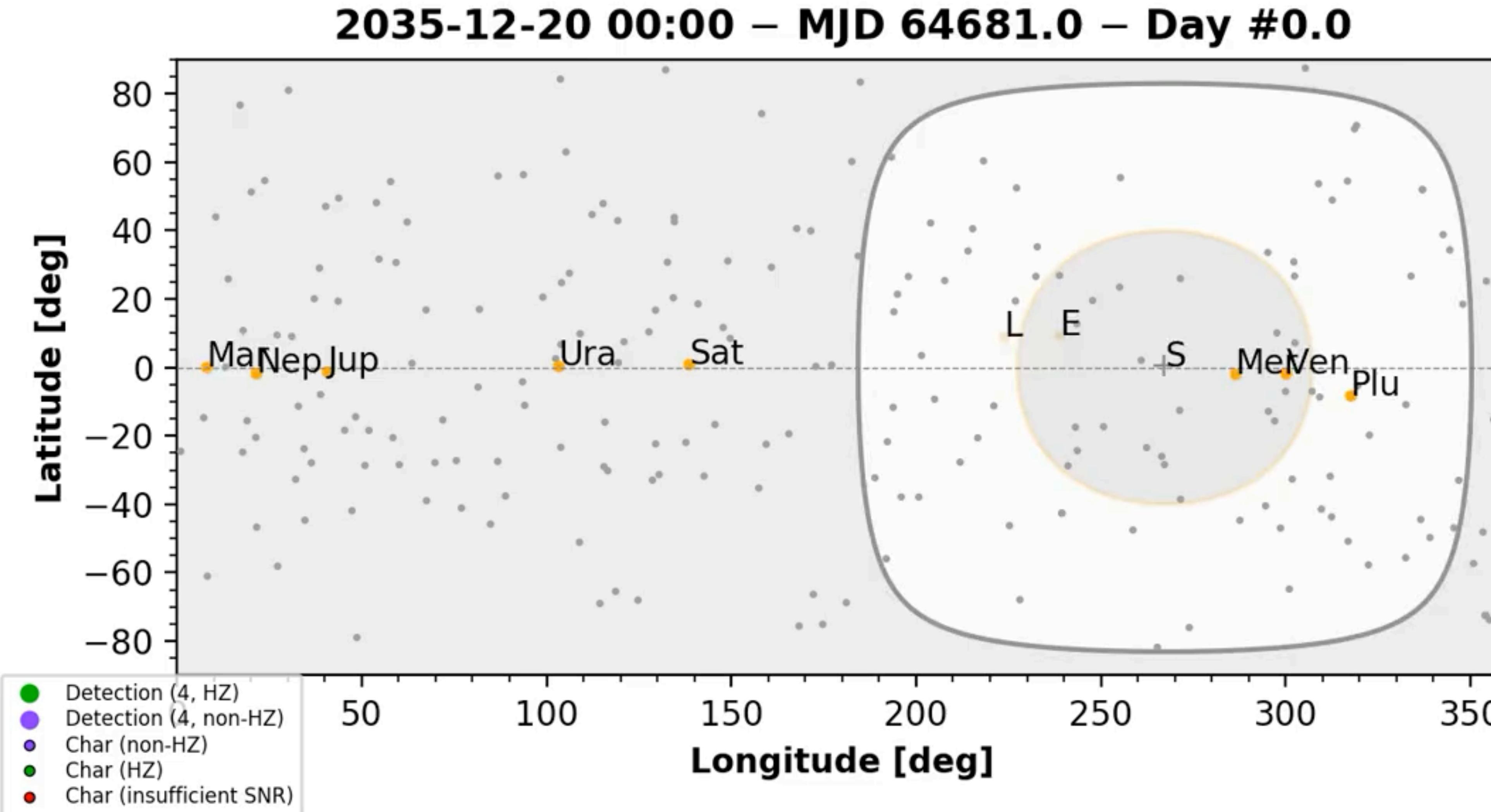
To probe outer regions of planetary systems with low transit probability



To measure exoplanetary masses: their most fundamental parameter

# Impact of a precursor survey

# Simulating the yield of HWO



Morgan et al. 2021, Savransky et al. 2020

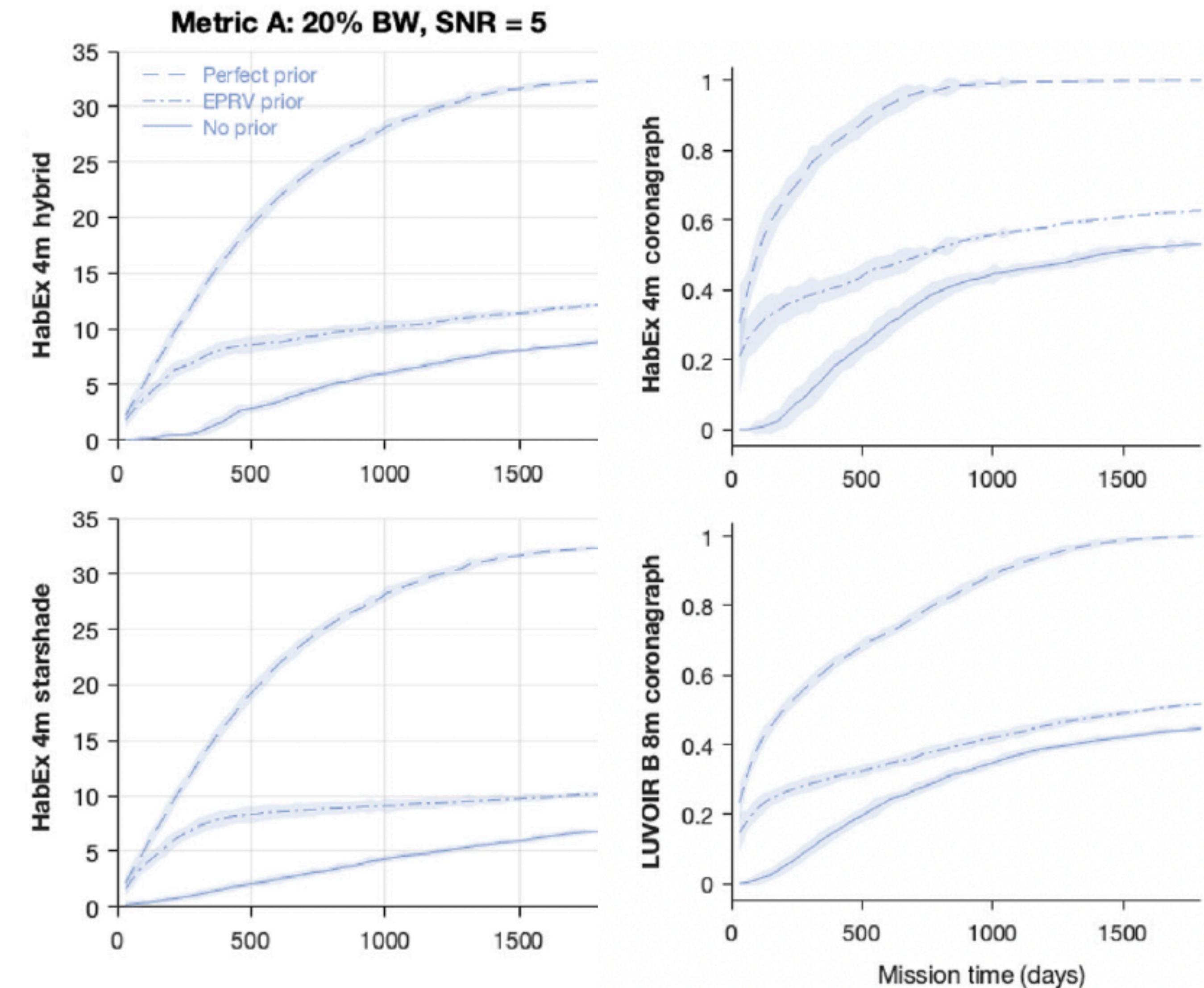
# Impact of a radial velocity precursor survey (Morgan et al. 2021)

HWO yield is greatly improved if targets are detected in advance

Universe + mission simulated with EXOSIMS (Delacroix et al. 2015, Savransky et al. 2020)

With a RV survey

- Results obtained faster, more time for characterisation
- Mitigates the risk of the missions



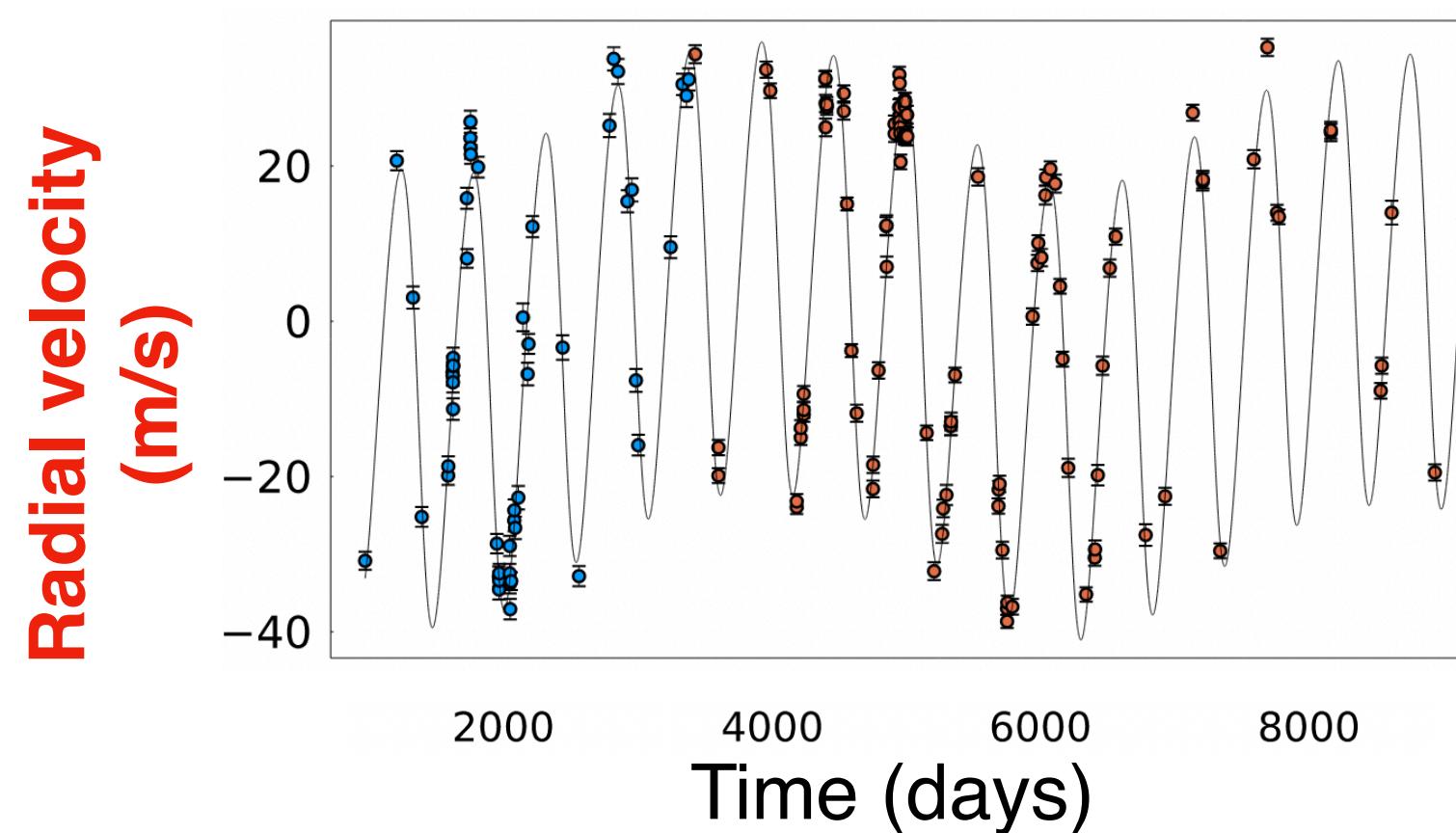
(3 cm/s) RV instrument on a 10-m class telescope surveying ~53 HabEx targets in a 5 year, 25% time survey

# RV precursor survey is feasible in principle

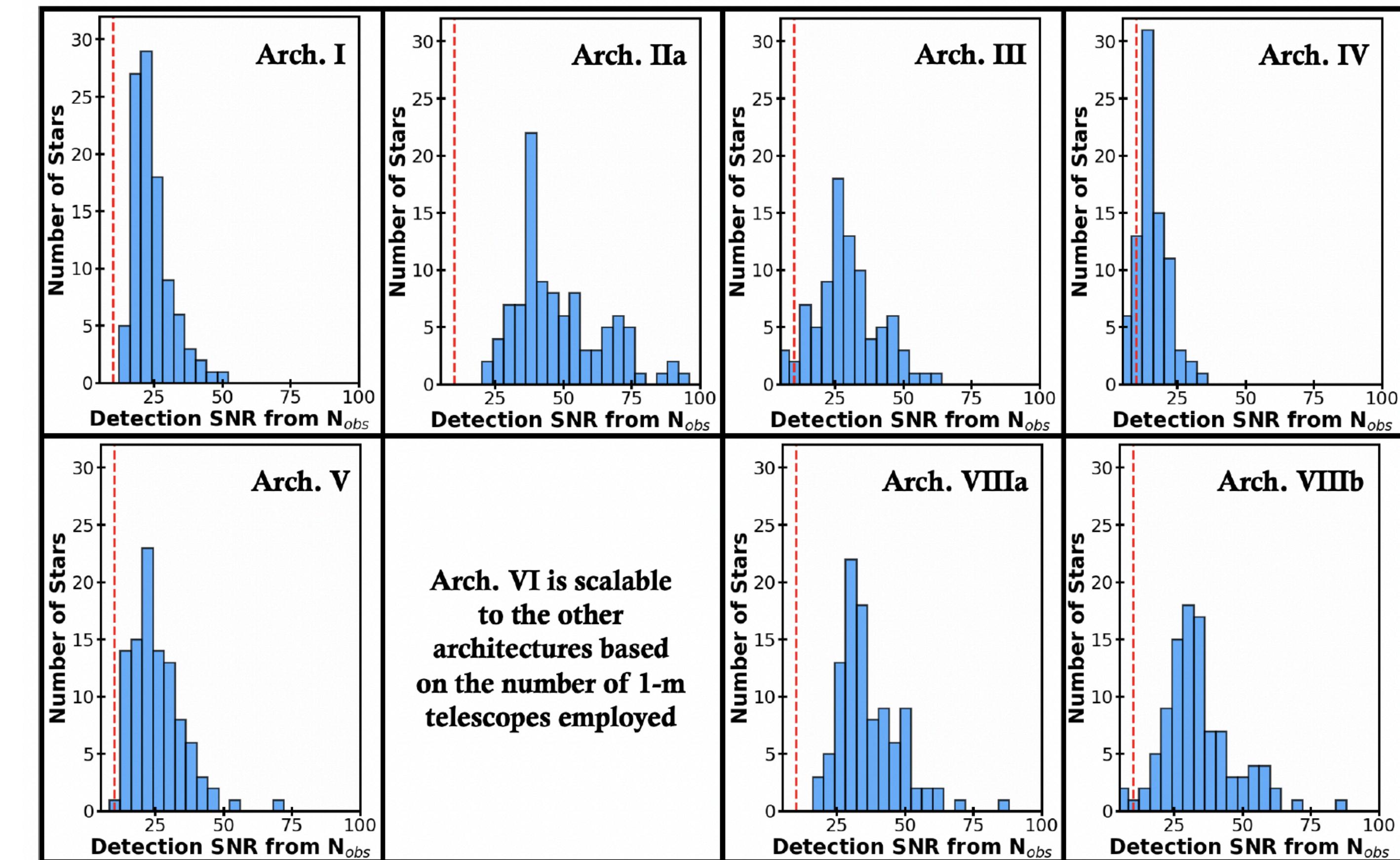
Simulation with different telescope/instrument associations

In principle the 100 prioritised targets can all be characterised

**Simulation done in the best case scenario:  
White noise**

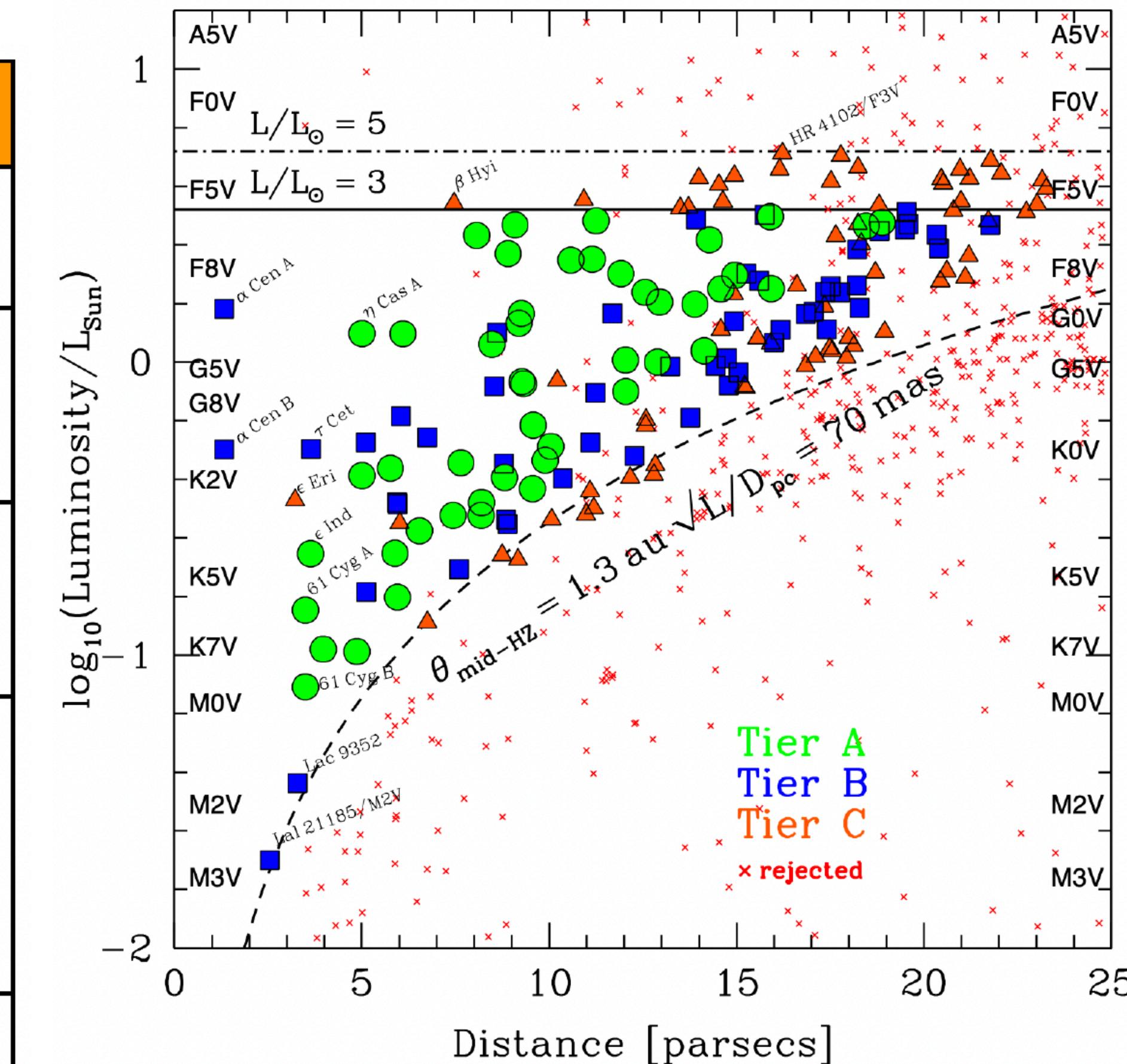


Histogram of obtained signal to noise ratio for 100 priority stars assuming they all have an Earth



# HWO prioritised targets (Mamajek et al. 2024)

Parameter	Tier A	Tier B	Tier C
IWA constraint	83 mas	72 mas	65 mas
Exoplanet brightness limit ( $R_c$ )	30.5 mag	31.0 mag	31.0 mag
Exoplanet-star Brightness ratio limit	4e-11	4e-11	2.5e-11
Disk criterion	No known dust disks of any kind	No disk, or KB disks OK if $L_{\text{disk}}/L_* \leq 10^{-4}$	All disks OK, even if $L_{\text{disk}}/L_* \geq 10^{-4}$ or detected HZ warm dust disk
Treatment of binaries	Single or binary companion $> 10''$ sep	Single or binary companion $5'' - 10''$ sep	Single or binary companion $3'' - 5''$ sep
Number of Stars	47	51	66

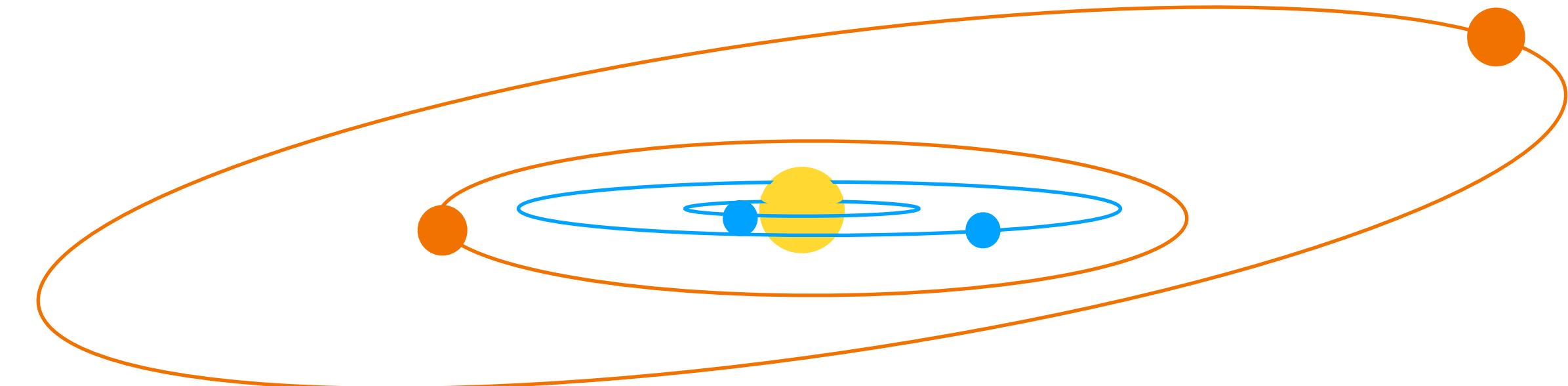


Sample	F	G	K	M
Tier A	14	15	17	1
Tier B	15	23	11	2
Tier C	37	17	12	0
Total (A+B+C)	66	55	40	3

# Further arguments

## For a RV survey of nearby stars

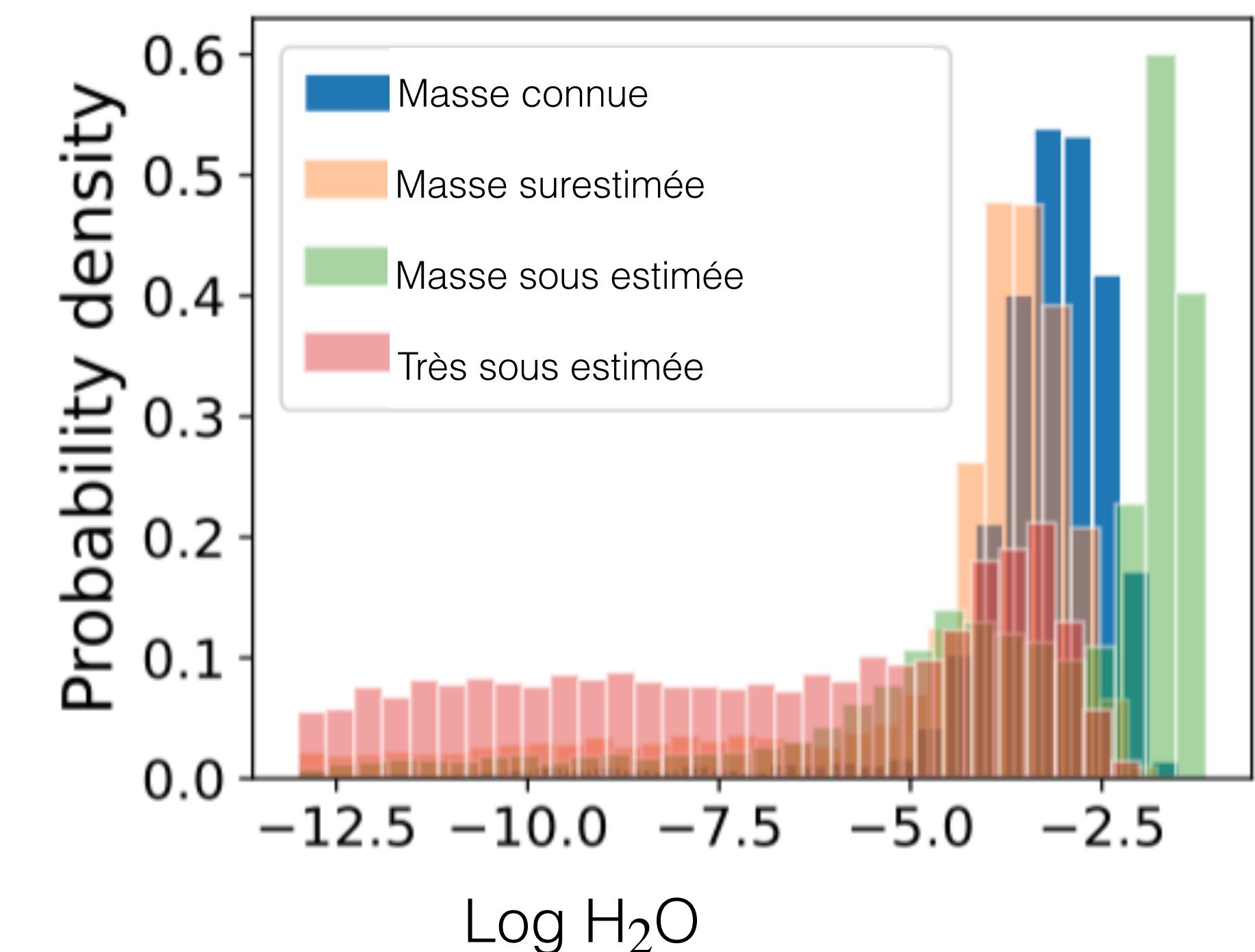
- Even if Earth-like planets are not detected, information on the system architecture
- Presence more or less likely of certain



## For extreme precision RVs

- Direct measurements of the mass is critical (Batalha+ 2019, Kempton+ 2018)
- Number of measurements needed proportional to individual RV uncertainty squared

$$N \propto \sigma_{\text{RV}}^2$$

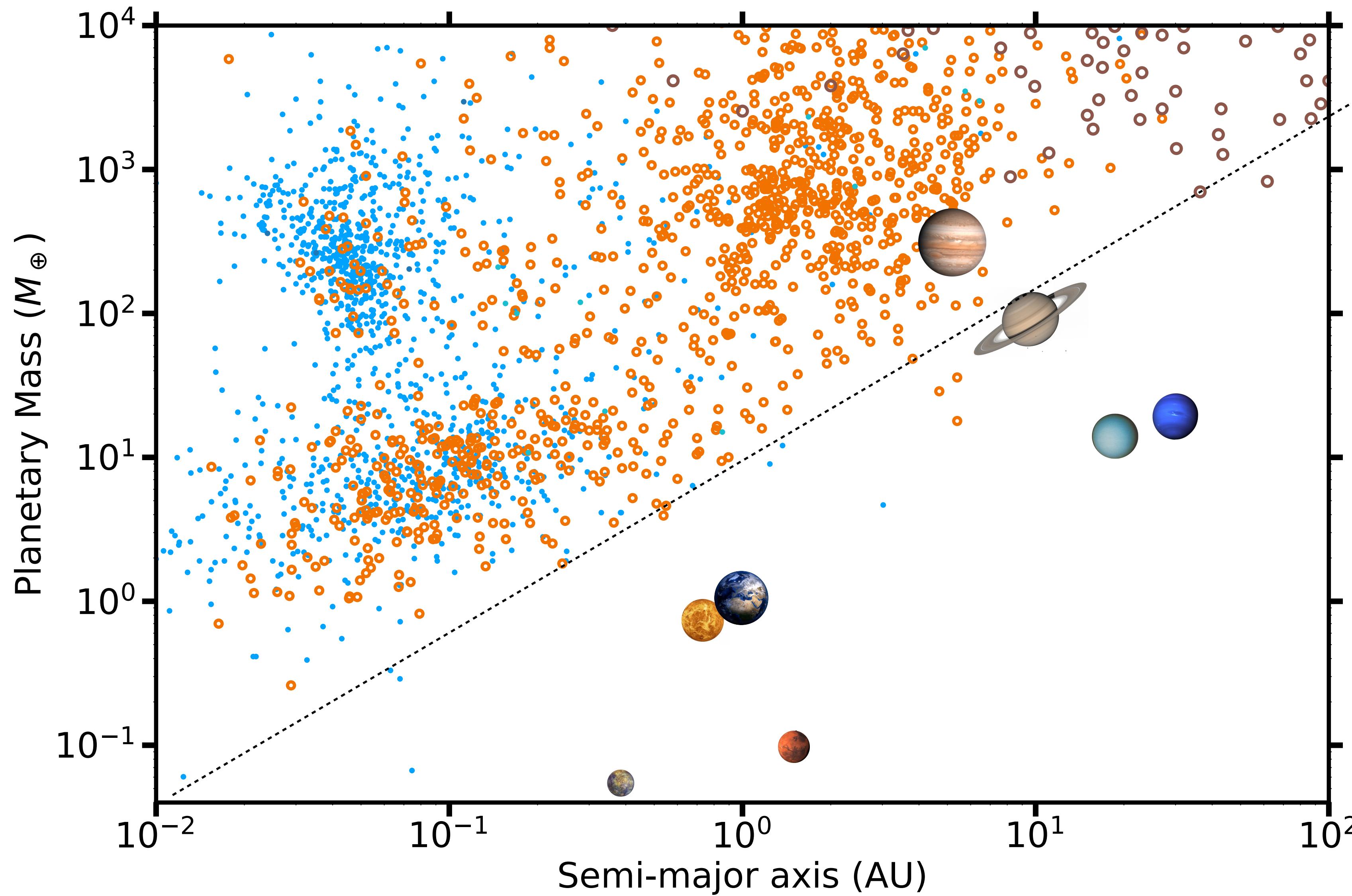


# NASA Extreme precision radial velocity report (Crass+ 2021)

« There exist multiple plausible system architectures in terms of telescope size, longitude and latitude distribution, and dedication that could, if **stellar variability mitigation, telluric mitigation, and instrumental precision goals are met**, successfully acquire a set of measurements with the statistical precision required to detect Earth analogs. »

# Difficulties of RV data analysis

# Known exoplanets



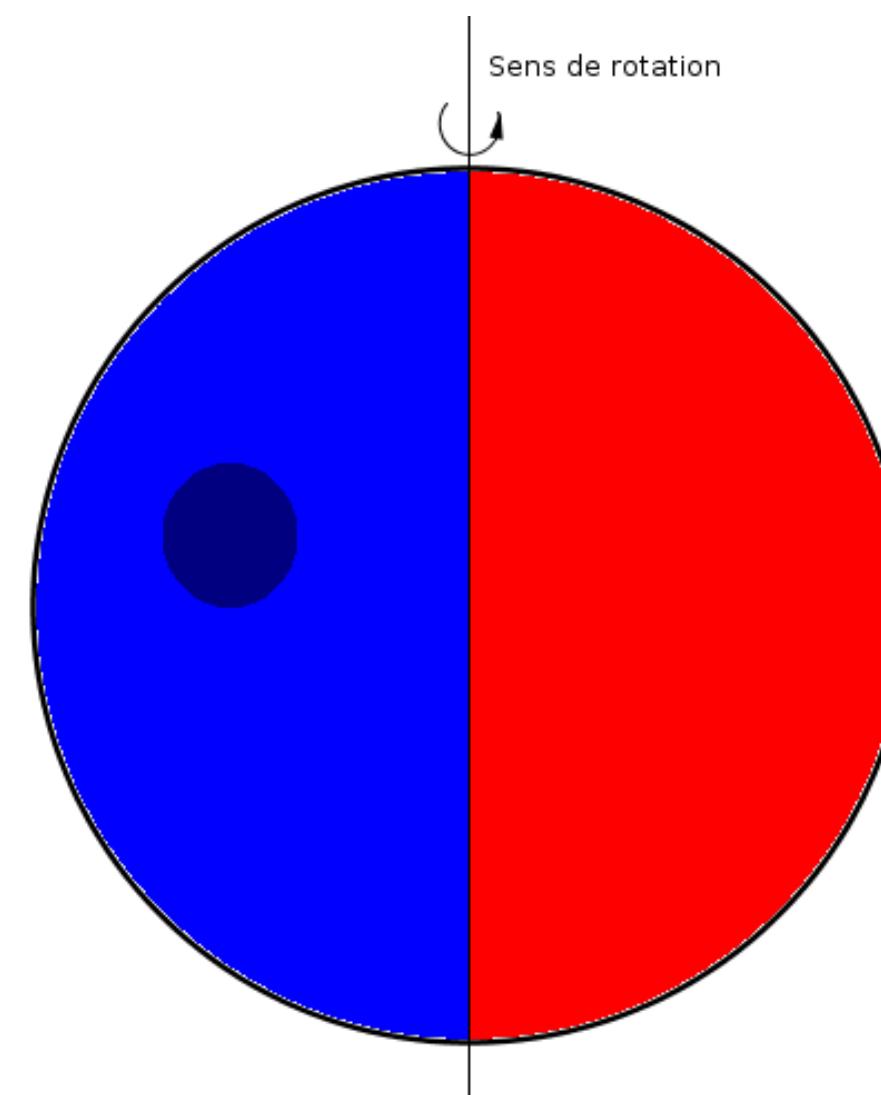
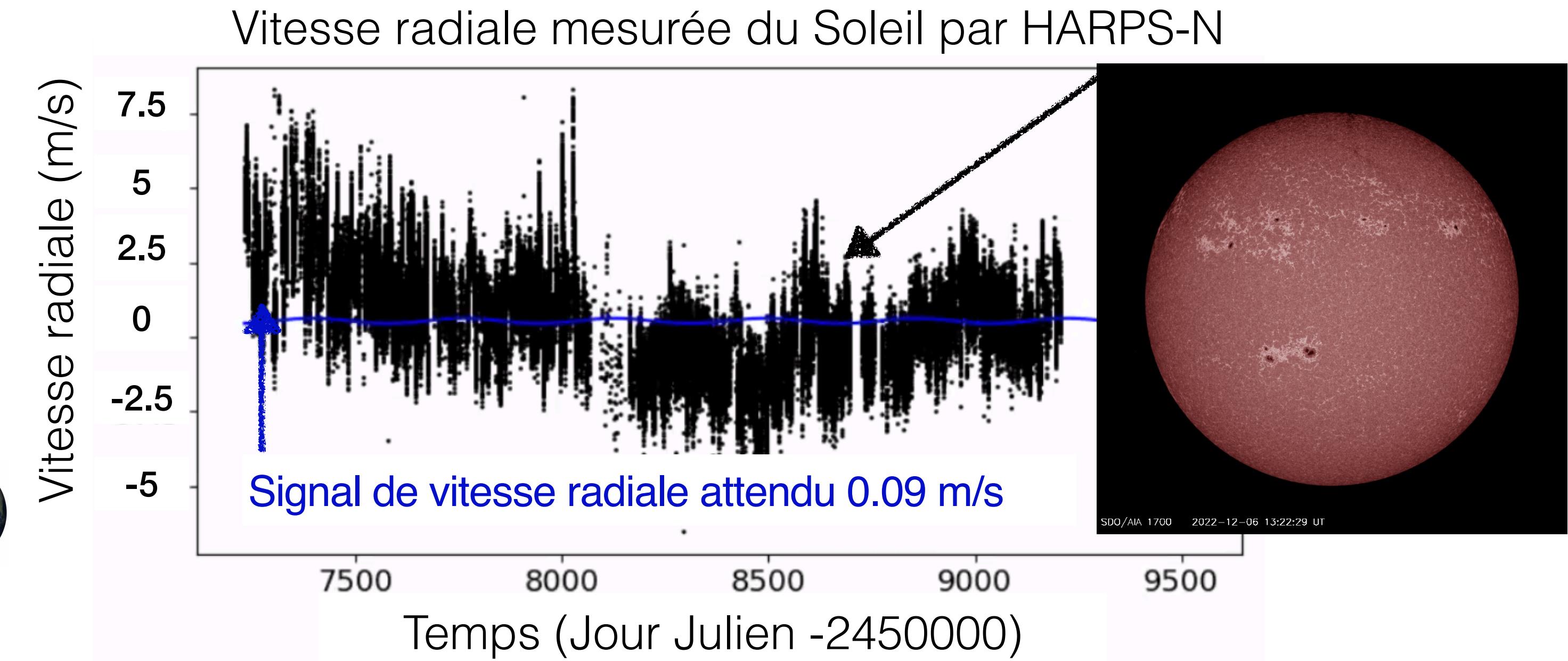
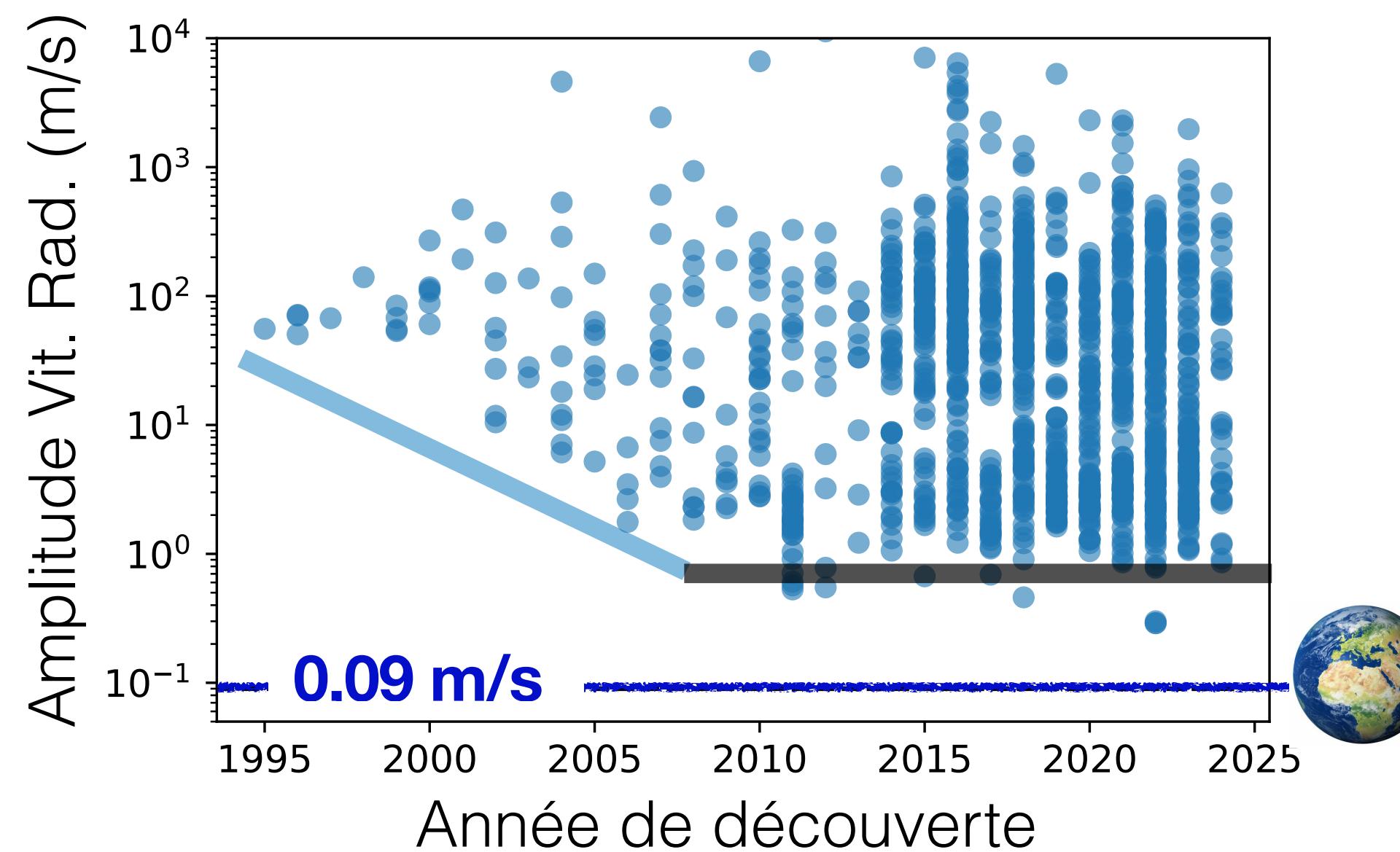
Transits

Radial velocity

Imaging

Source: [exoplanet.eu](http://exoplanet.eu)  
April 2024

# The stellar variability problem

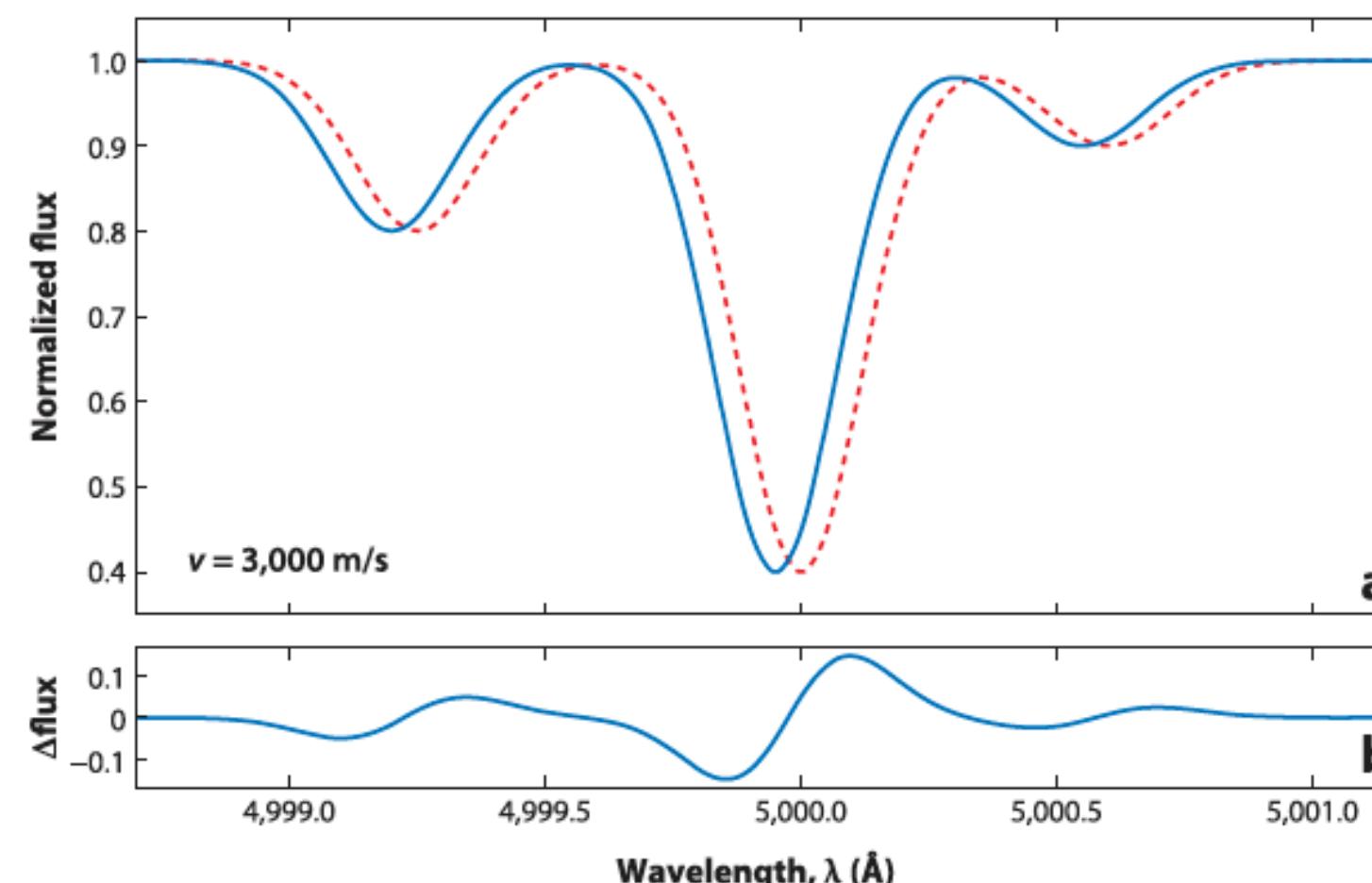


No exo-Earth detection possible with radial velocities so far

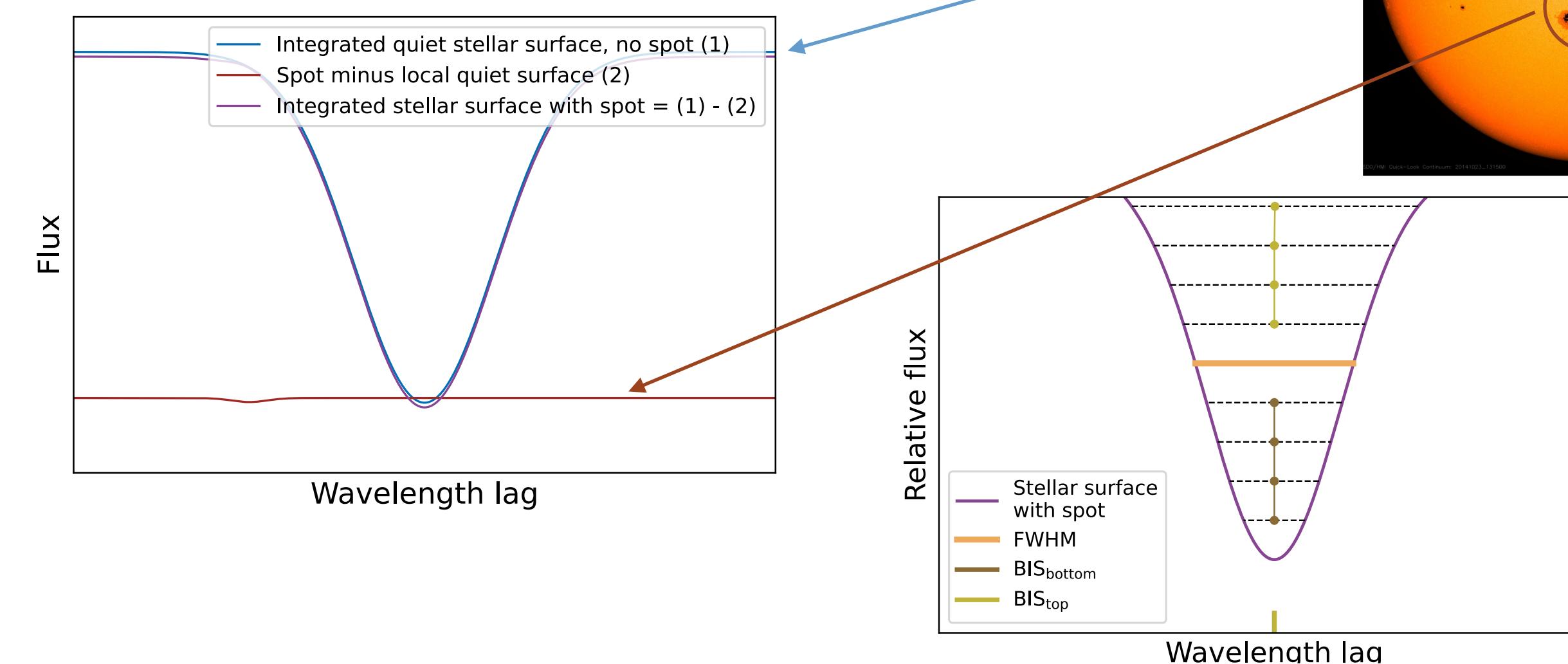
Calibration is still to be improved but instrument precision is not the main limitation  
Only solution: new data analysis methods  
(>250 ref. Hara & Ford 2023, *Annual Reviews*)

# Key ideas to correct stellar variability

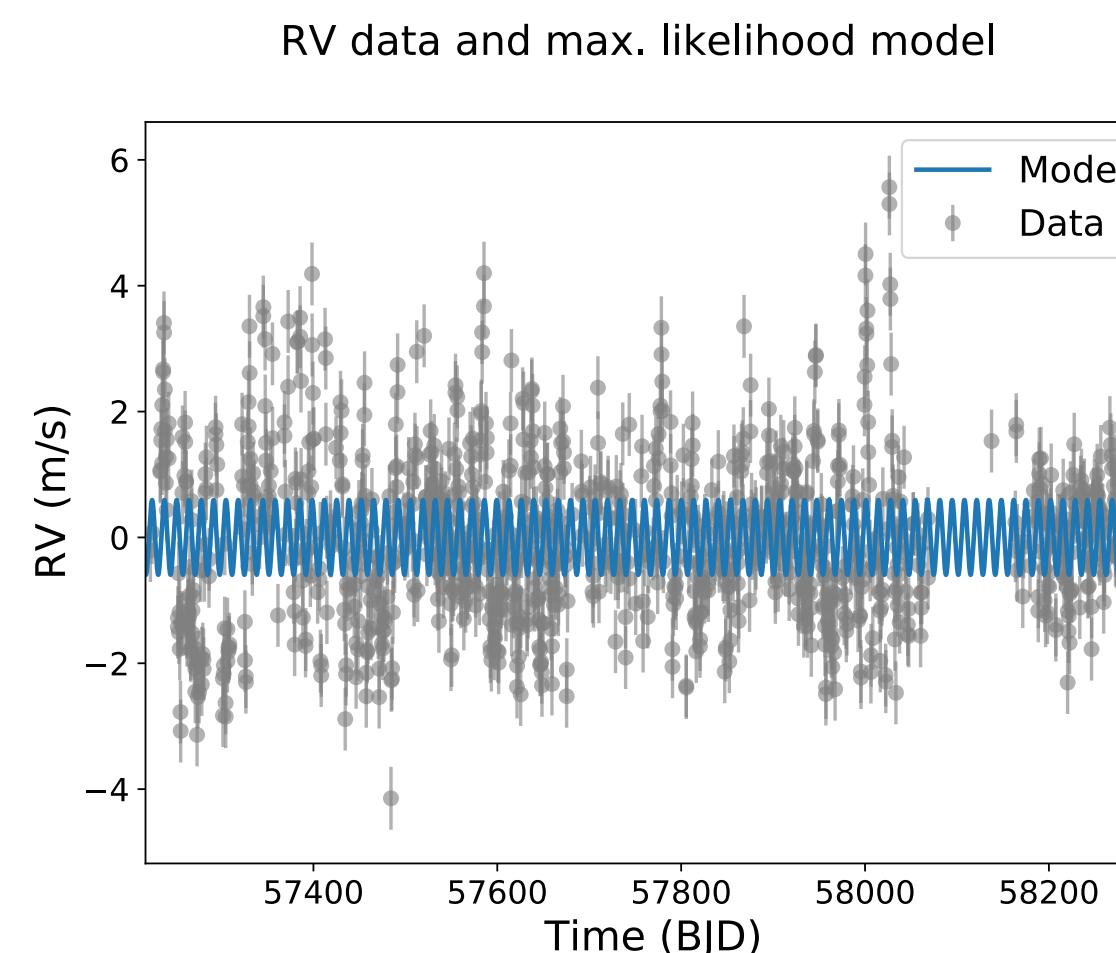
Planets induce a pure Doppler shift



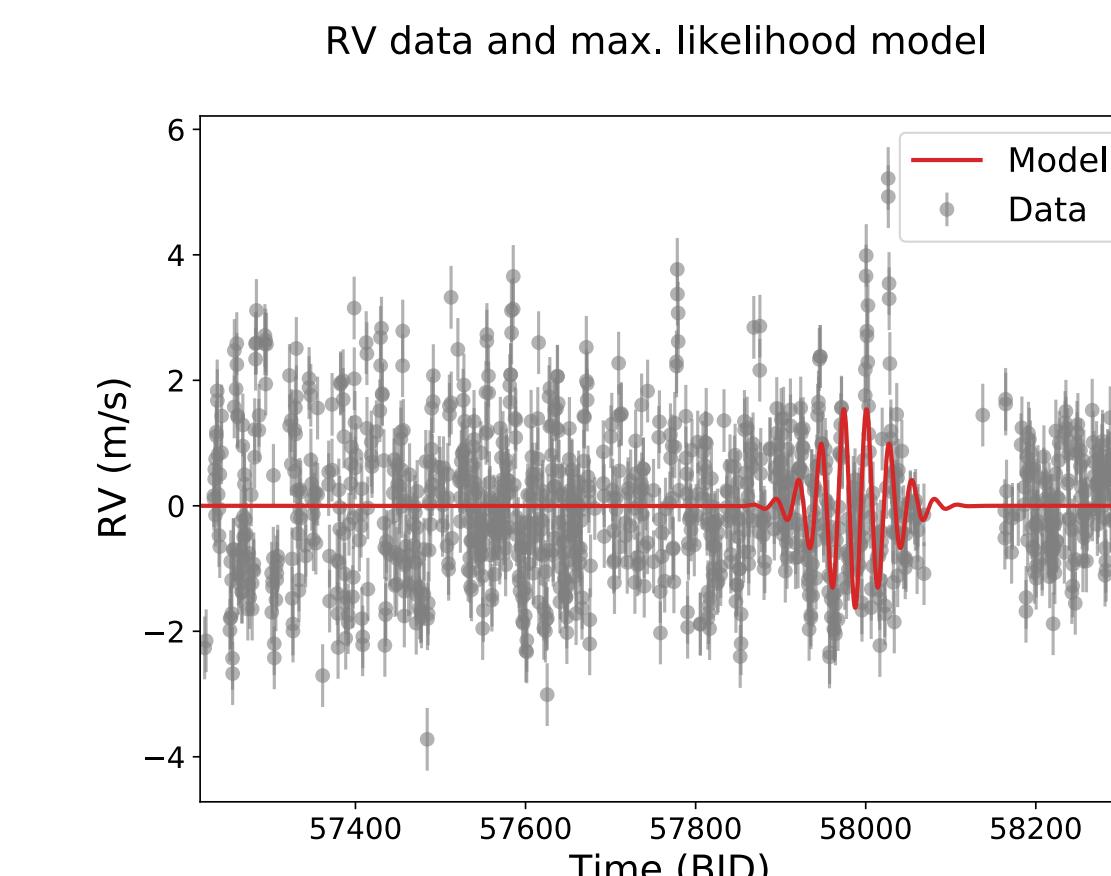
stellar and instrumental effects change  
the shape of the spectrum



Planets induce a periodic signal



Stellar and  
instrumental  
effects are  
(usually) not  
strictly periodic



# The extreme precision RV problem in a nutshell

$$RV_{\text{measured}} = RV_{\text{center of mass}} + RV_{\text{contam}}$$

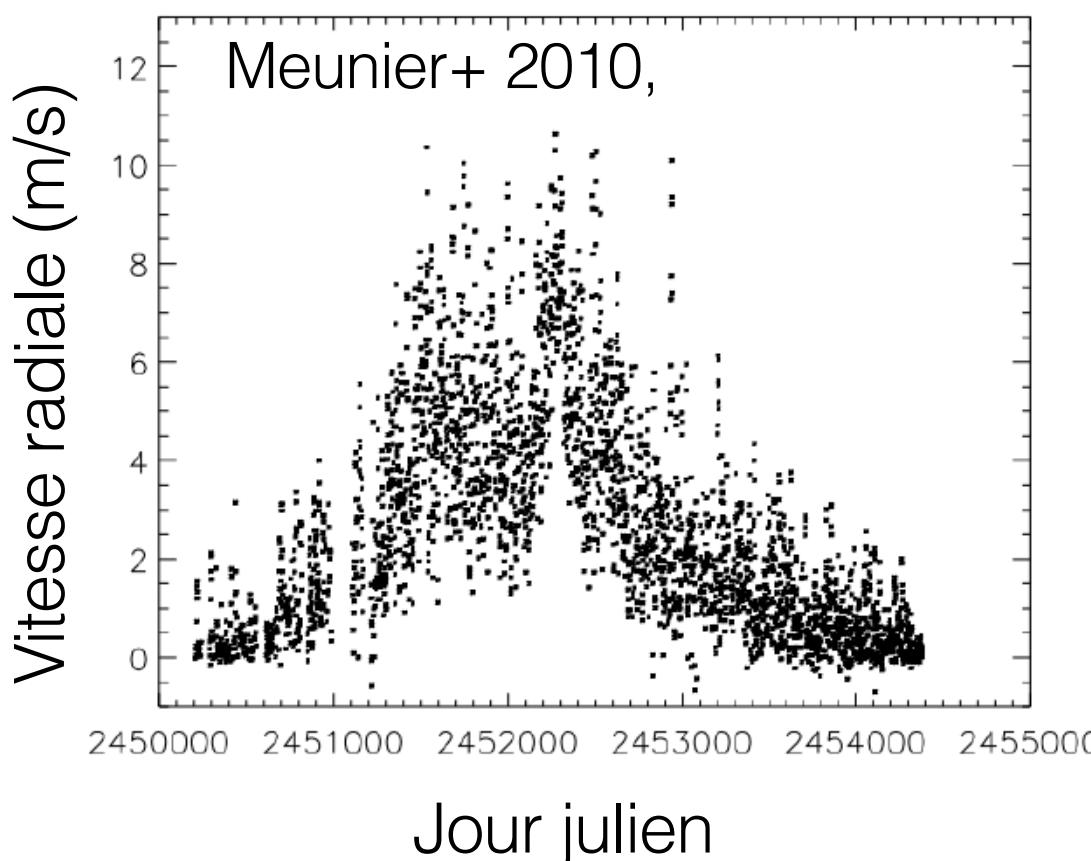
We want

- 1) An RV whose definition is as close as possible as a global Doppler shift
- 2) To leverage the shape of the spectra to estimate the contamination
- 3) To estimate the uncertainties on  $RV_{\text{contam}}$

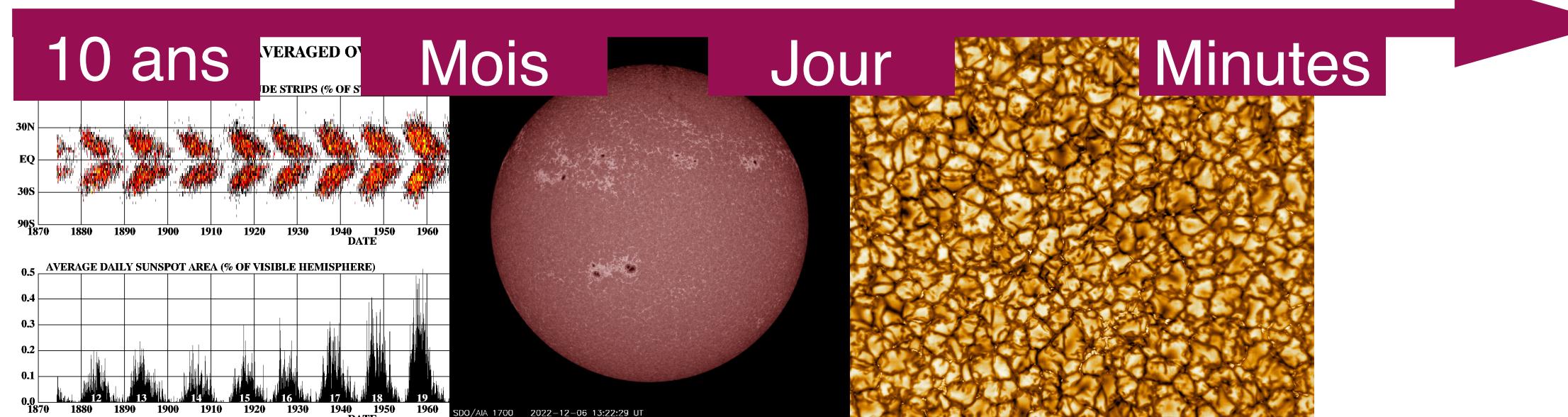
Can be seen as a statistical or machine learning problem

# Understanding stellar variability and modelling it in practice

Detailed physical models  
Analysis of solar data



Numerous processes at  
Different timescales



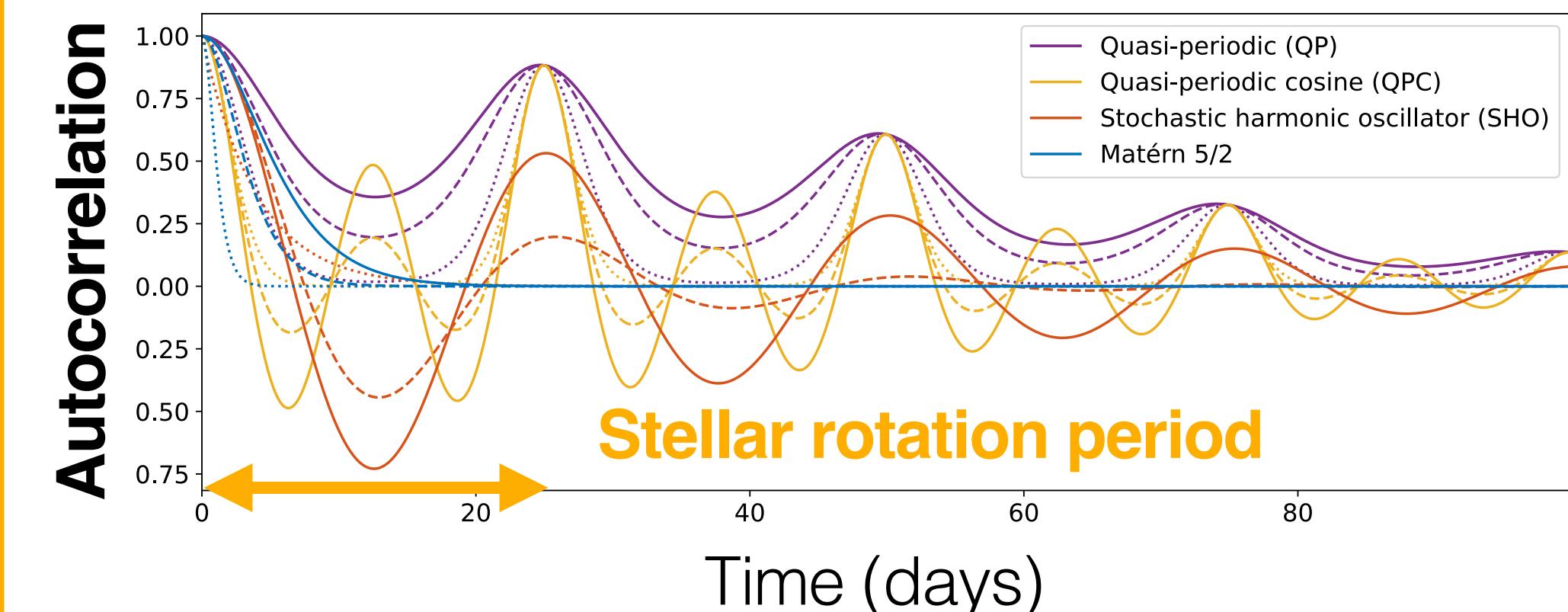
Magnetic activity  
p.ex. Meunier+ 2010, 2012, 2019  
Boisse+ 2012, Dumusque+ 2014,  
Haywood+ 2016, Al Moulla 2023

Granulation and  
super granulation  
Cegla+2013, 2019  
Dravins+ 2021,

+ oscillations,  
winds,  
gravitational  
redshift...

**Likelihood** supposed Gaussian and stationary with qualitative parameters

$$p(\text{data} | \theta, \eta) = \frac{1}{\sqrt{2\pi}^N \sqrt{|V(\eta)|}} e^{-\frac{1}{2} (\text{data} - f(\theta, \eta))^T V(\eta)^{-1} (\text{data} - f(\theta, \eta))}$$



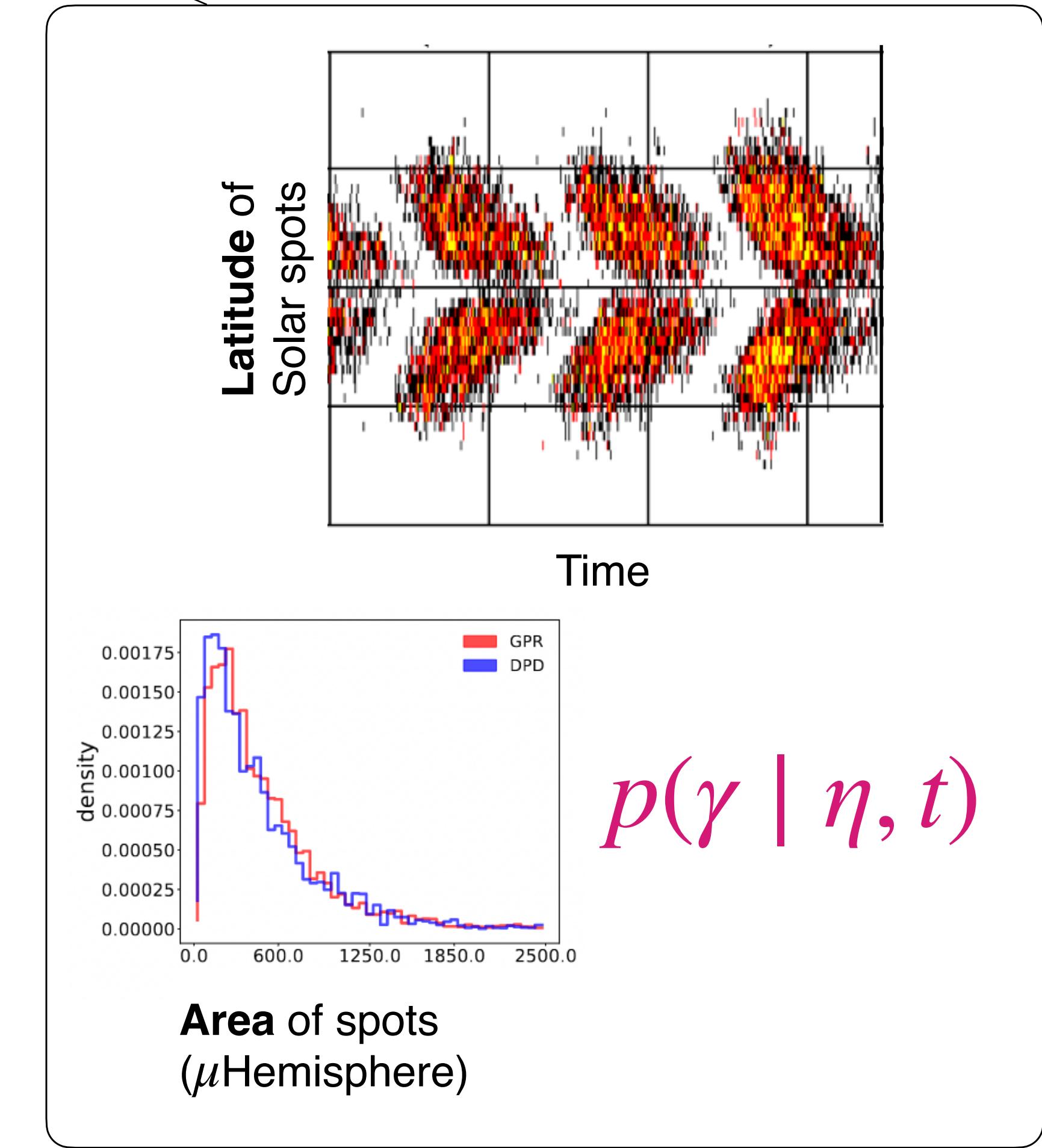
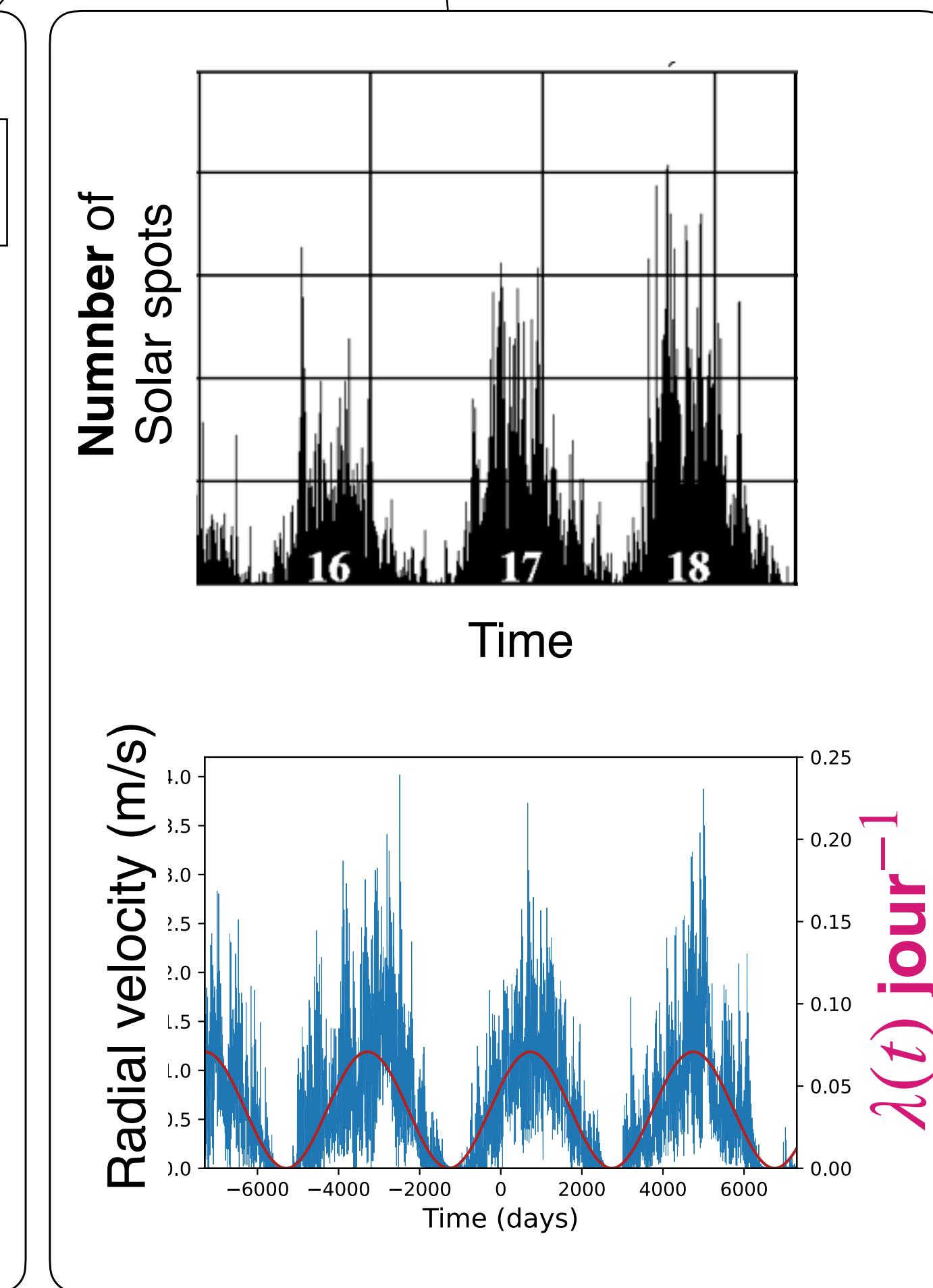
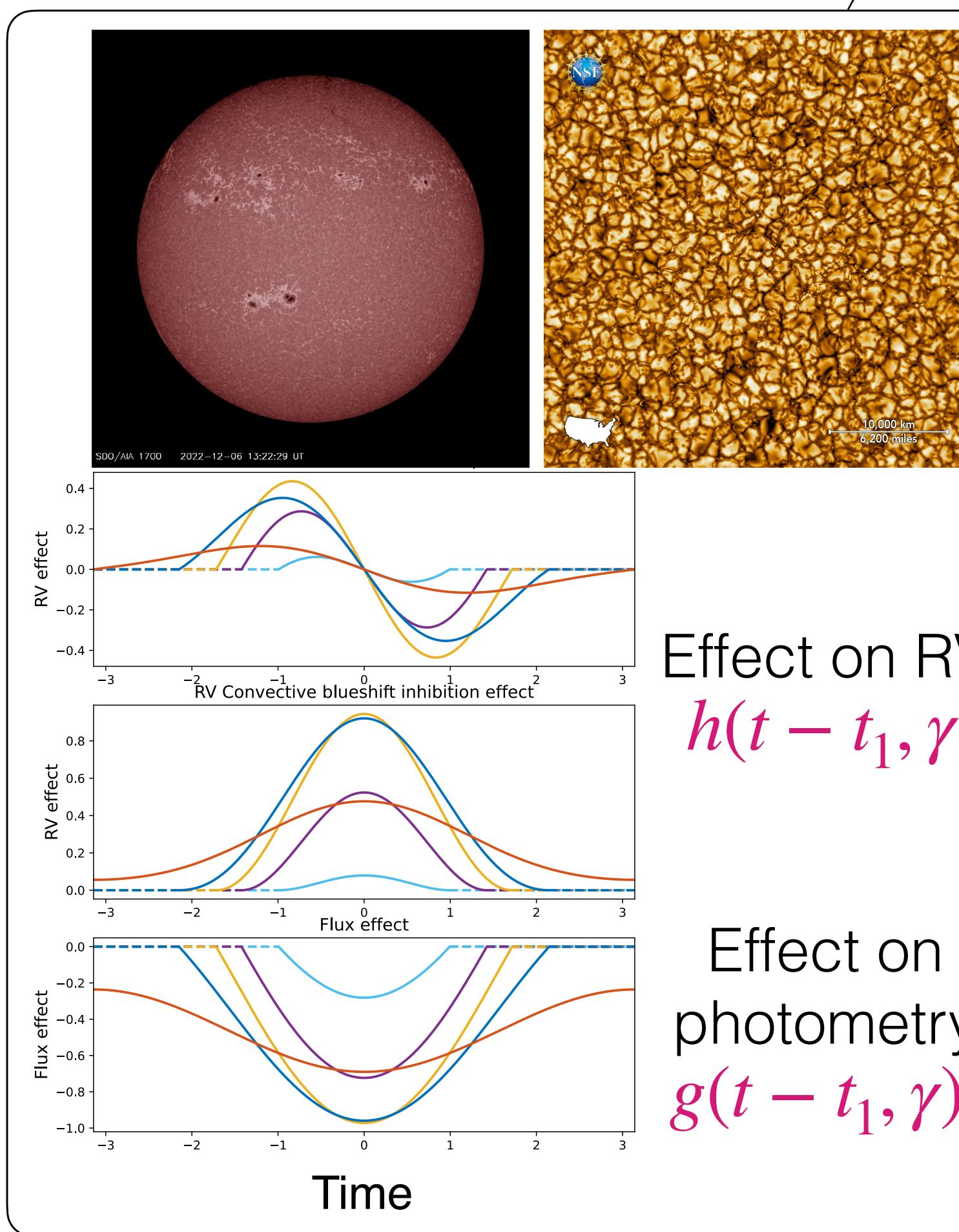
Aigrain et al. 2012, Haywood et al. 2014, Foreman Mackey et al. 2017, Rajpaul et al. 2015, Perger et al. 2021, Jones et al. 2022...

Statistical models are not quantitatively related to the physical model

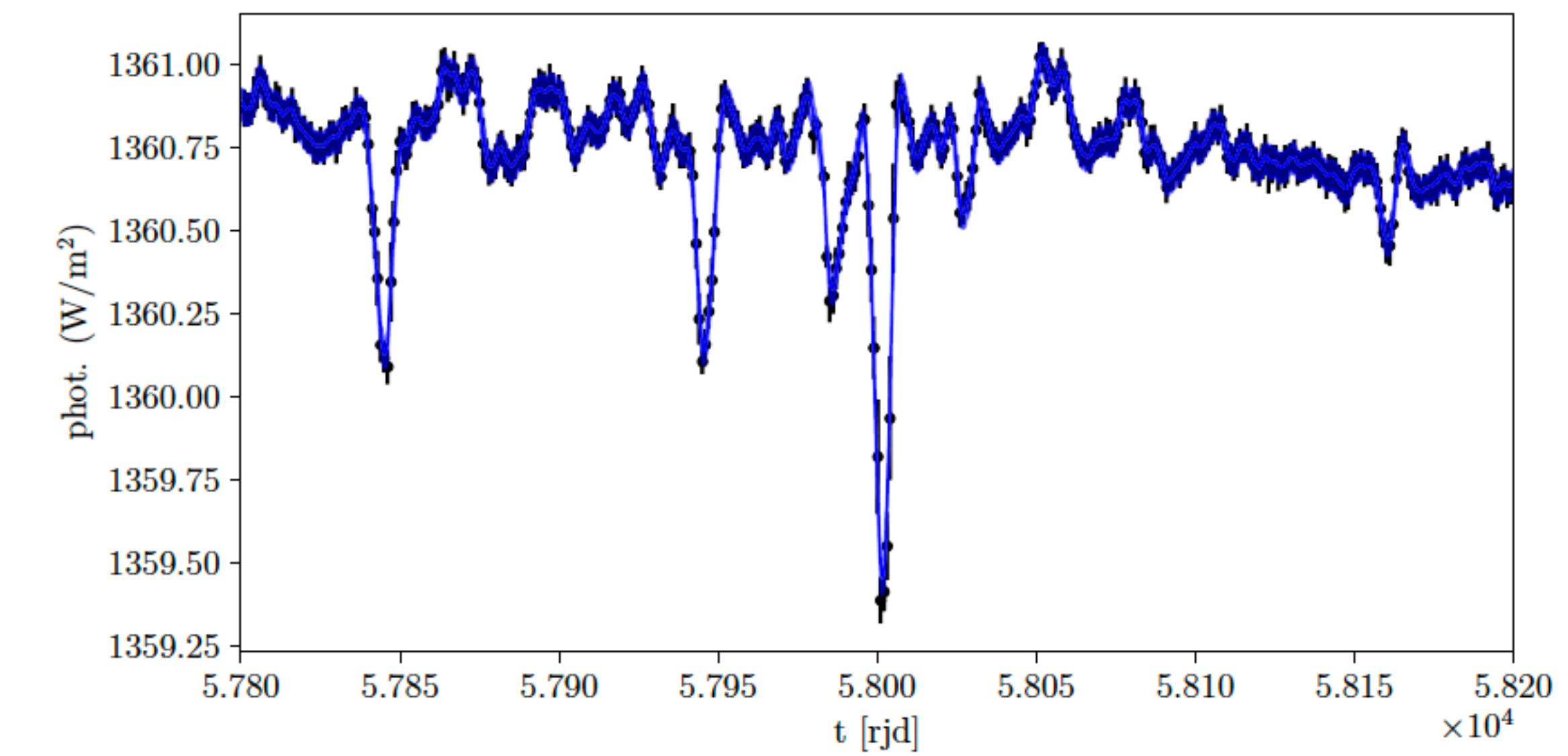
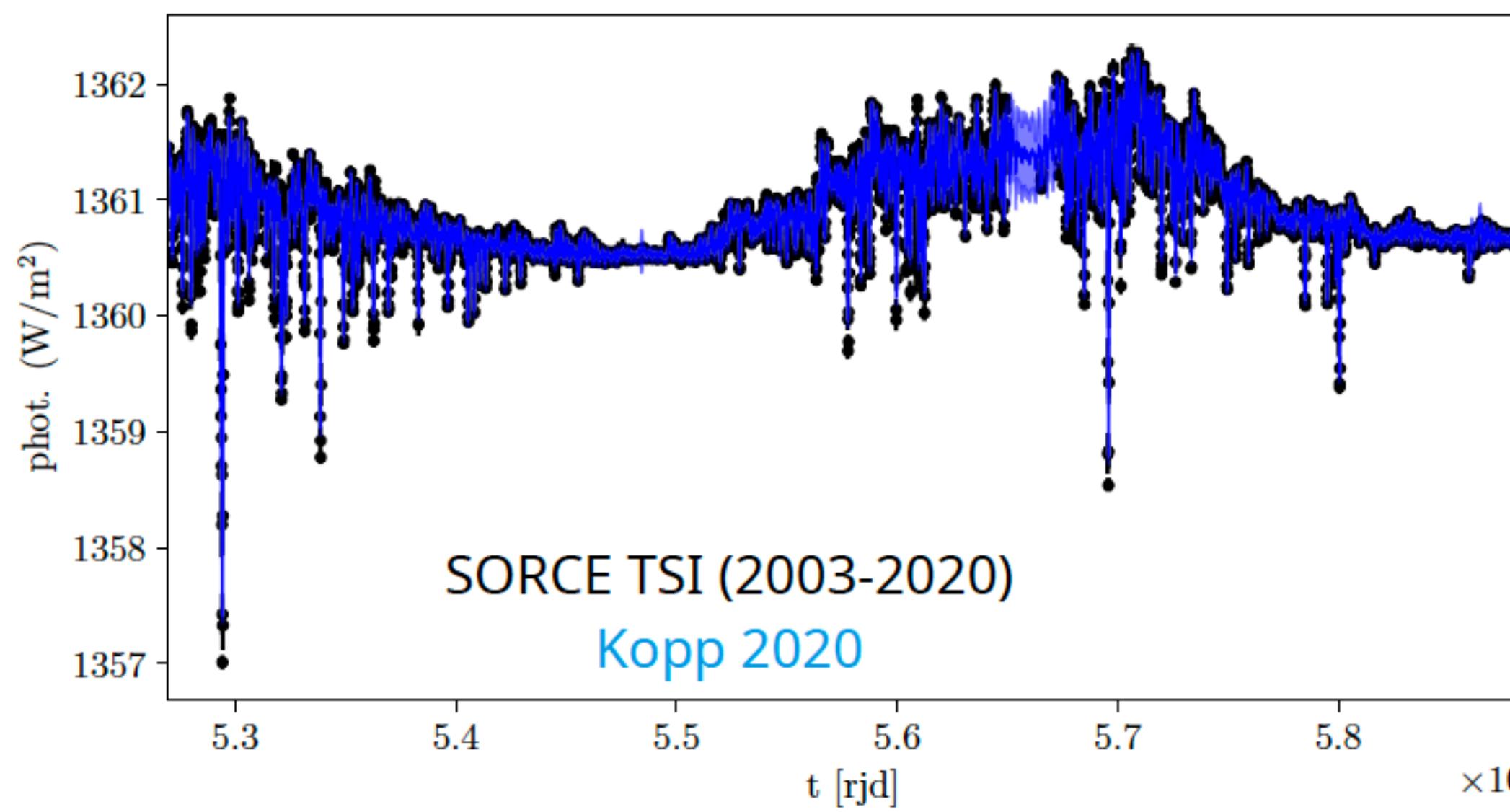
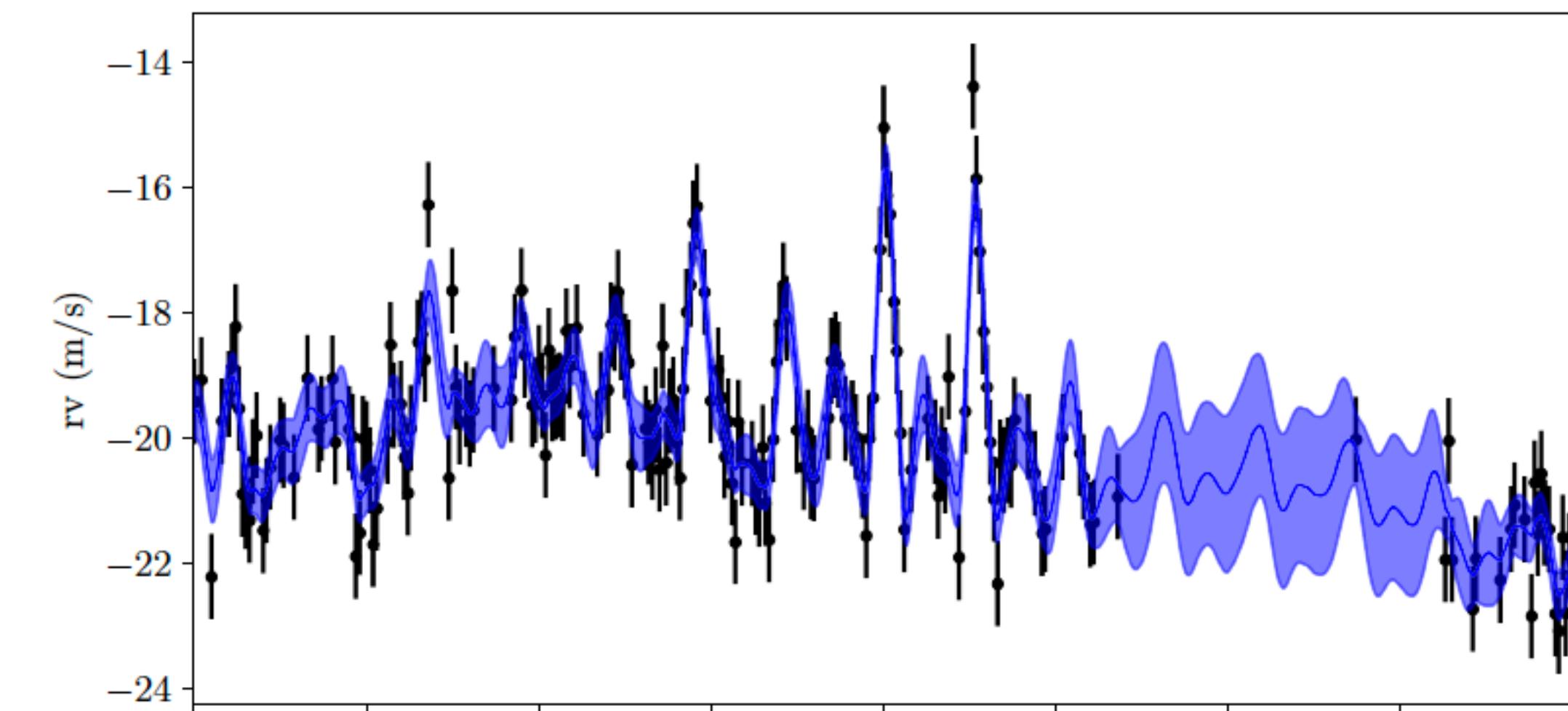
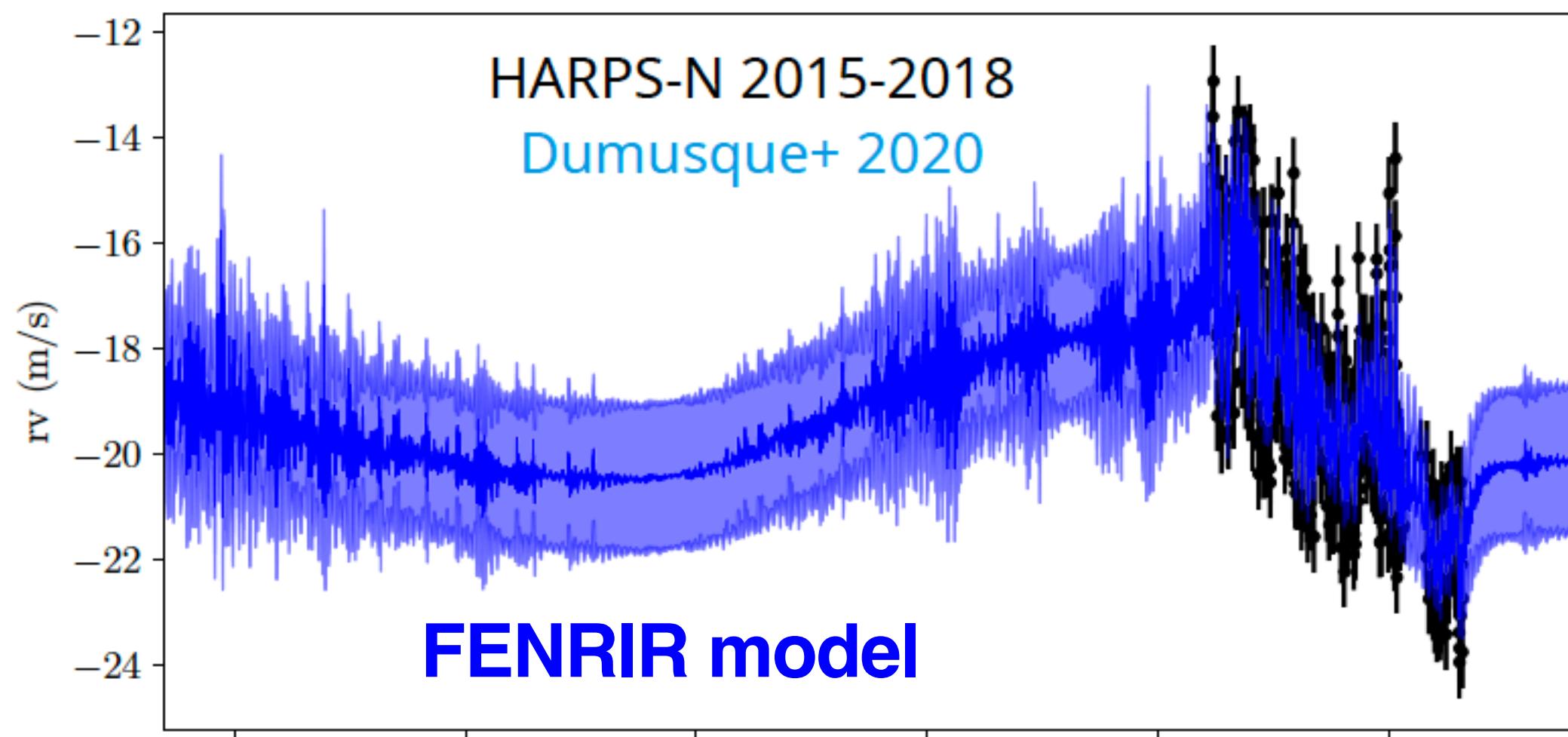
# From the physical model, we compute the covariance

$$\iint g(t - t_1, \gamma) h(t - t'_1, \gamma) \lambda(t) p(\gamma | t, \eta) dt d\gamma = \kappa_\eta(P(t_1), RV(t'_1))$$

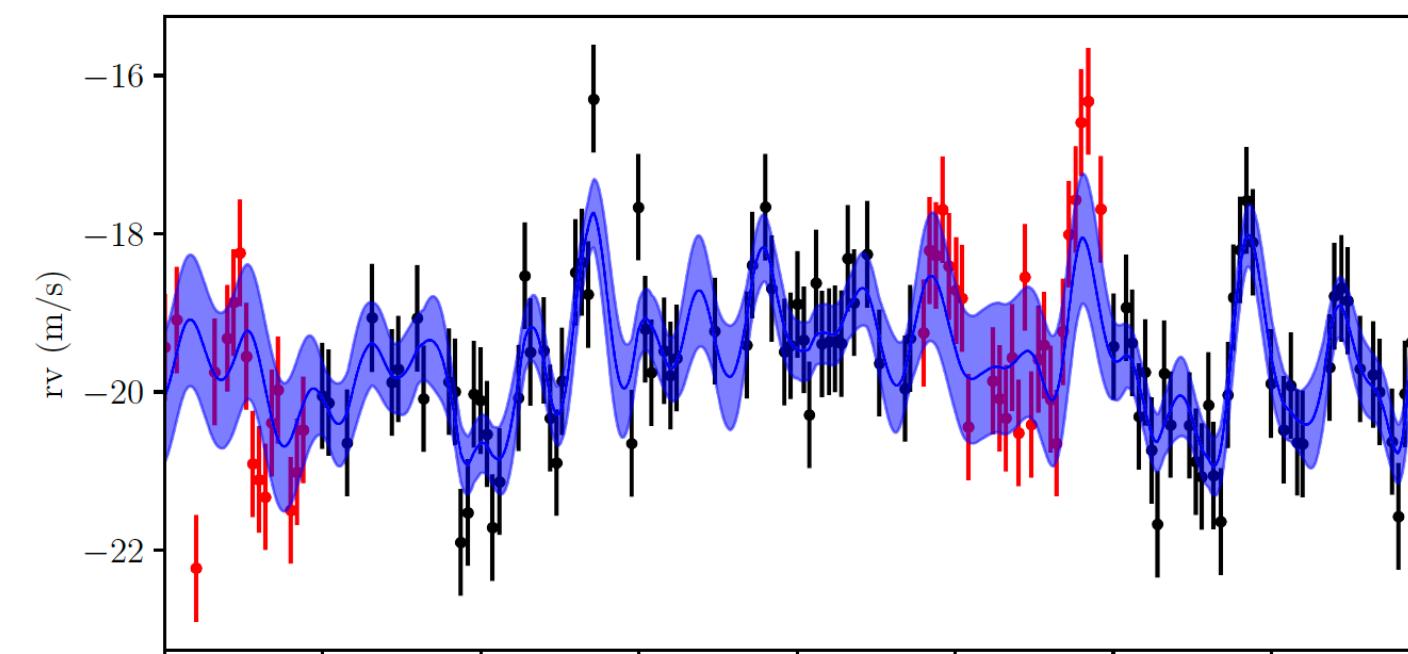
Hara & Delisle  
2023



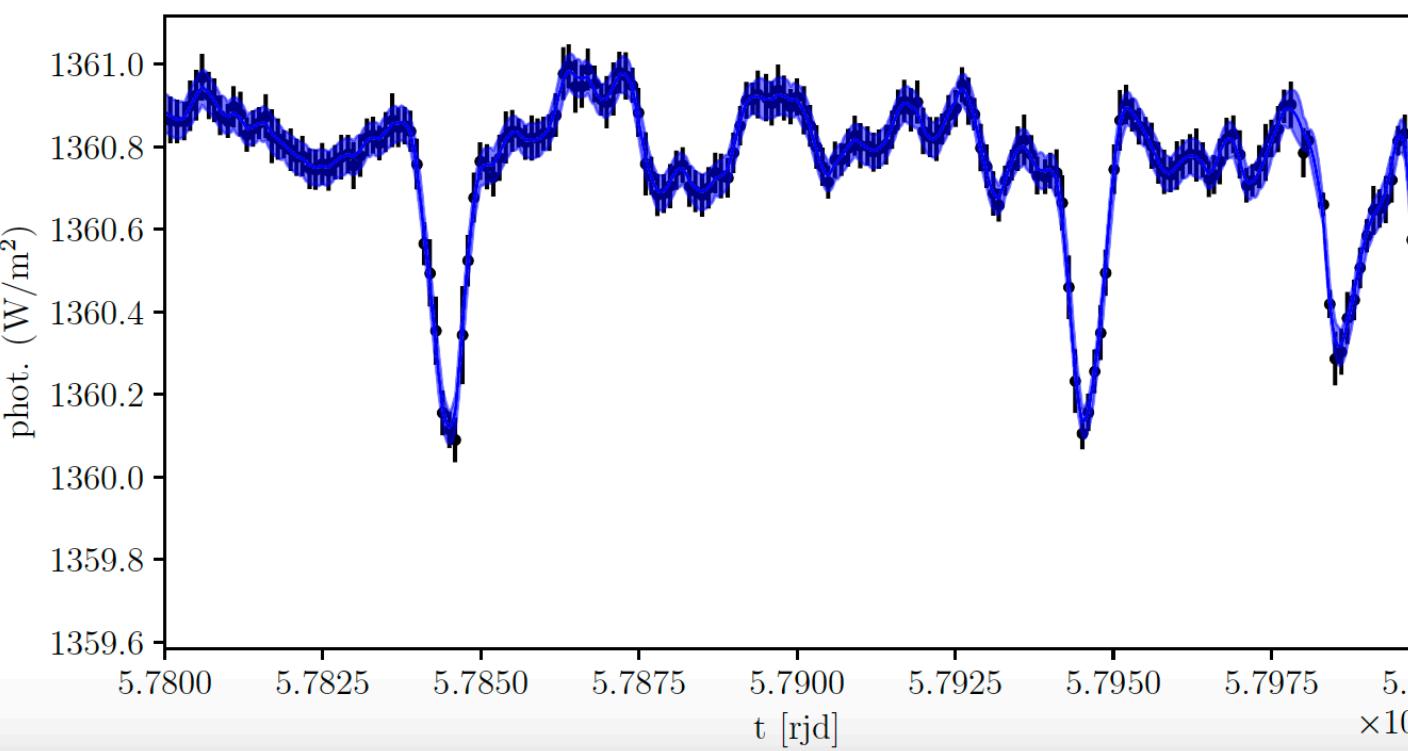
# Performance of the stellar activity model



# Performance of the stellar activity model



We remove a the **test set**



We fit the parameters to the remaining points  
(the **training set**)

We compare the prediction of the **model** with the **test set**

We compute the likelihood of the predictive distribution

Model	modes	Period (d)	Inclination (deg)	Cross-validation score
data-driven	2	$27.03^{+0.12}_{-0.11}$	–	-177.15 ( $-178.15^{+1.51}_{-1.74}$ )
physical	sep., lat. dist.	$26.82^{+0.07}_{-0.07}$	$2.59^{+2.90}_{-1.78}$	-181.54 ( $-182.06^{+1.09}_{-1.25}$ )
physical	sep., opp. lat.	$26.82^{+0.08}_{-0.08}$	$3.61^{+4.71}_{-2.56}$	-181.64 ( $-182.16^{+1.07}_{-1.32}$ )
latent	2	$22.44^{+0.41}_{-0.39}, 26.78^{+0.21}_{-0.21}$	–	-184.15 ( $-185.97^{+1.81}_{-1.83}$ )
data-driven	1	$27.00^{+0.12}_{-0.12}$	–	-187.24 ( $-188.89^{+2.02}_{-2.56}$ )
physical	mix., lat. dist.	$27.07^{+0.12}_{-0.12}$	$29.12^{+2.25}_{-2.09}$	-187.69 ( $-188.11^{+1.02}_{-1.29}$ )
latent	1	$24.33^{+0.67}_{-0.63}$	–	-191.72 ( $-196.61^{+4.70}_{-12.59}$ )
physical	mix., opp. lat.	$27.90^{+0.16}_{-0.16}$	$1.20^{+1.26}_{-0.85}$	-216.24 ( $-218.74^{+2.49}_{-2.94}$ )

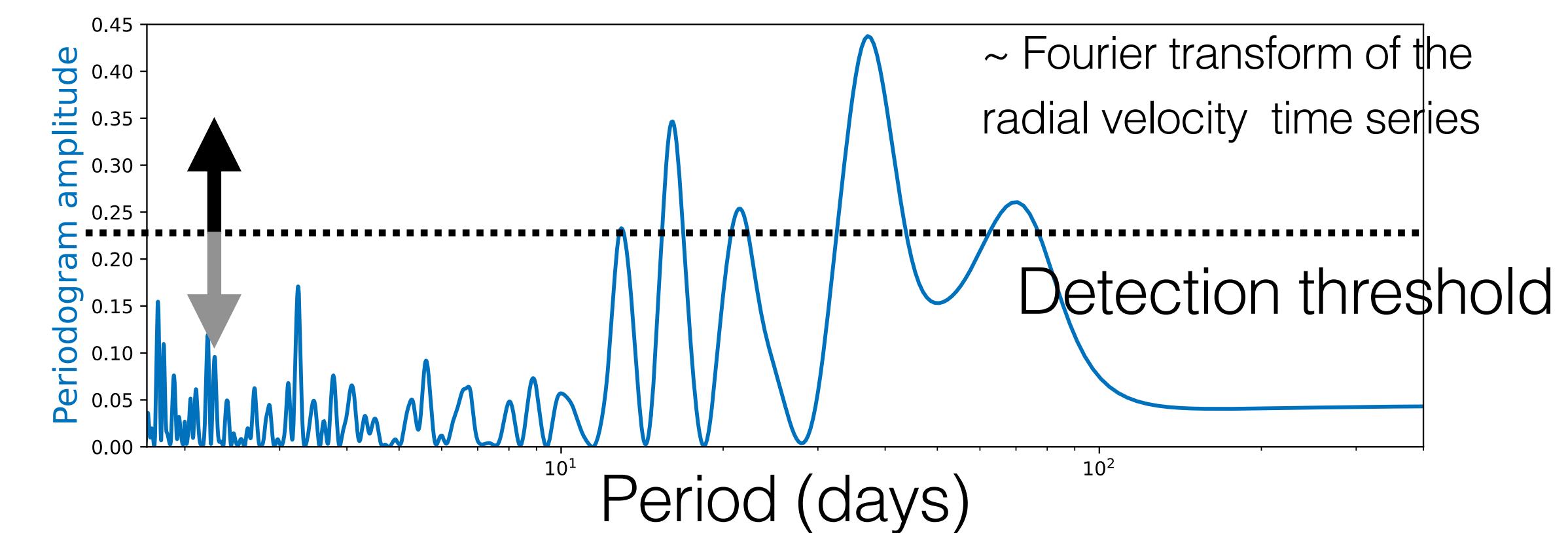
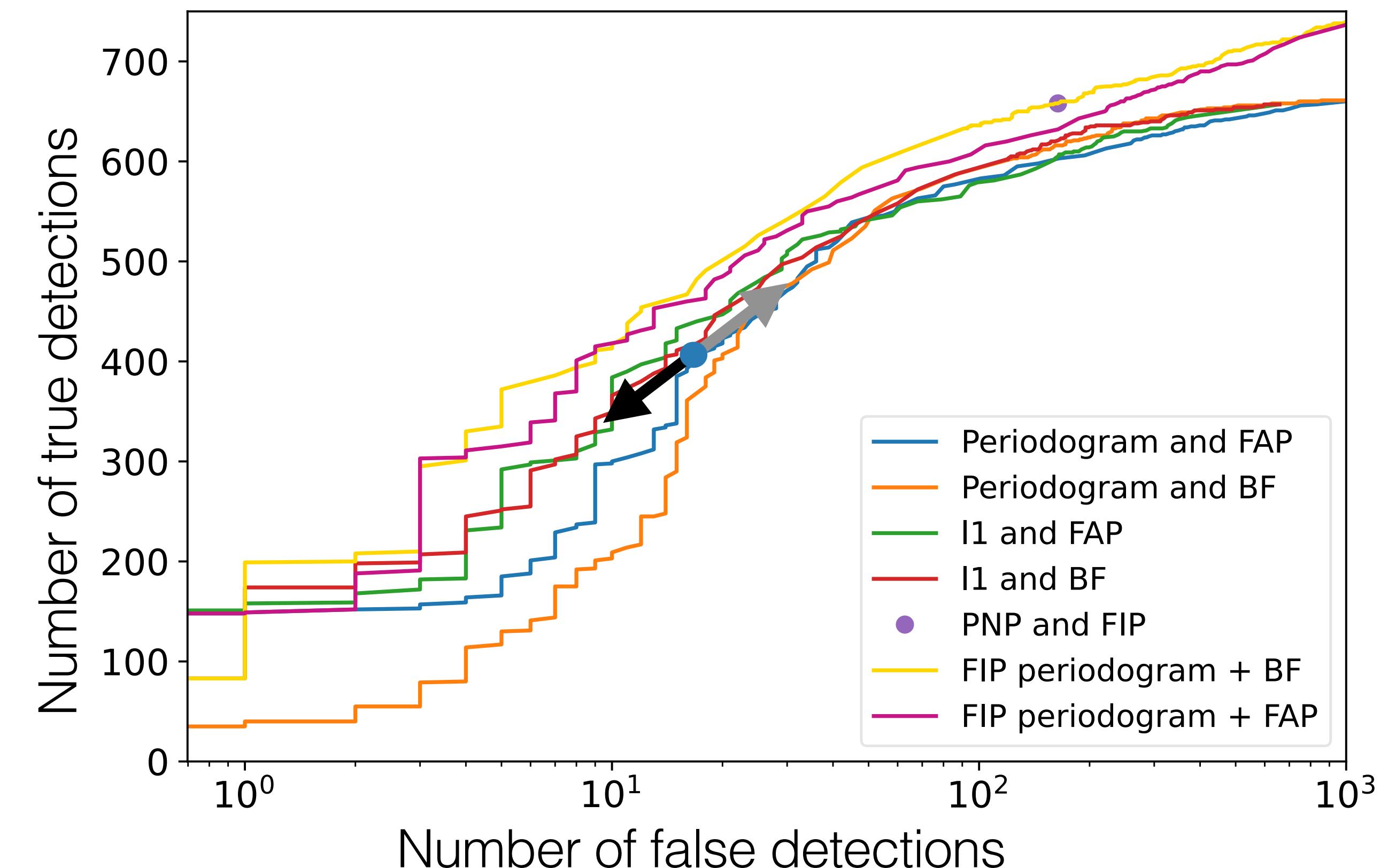
# Optimal criterion to detect exoplanets (1)

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



# Optimal criterion to detect exoplanets (2)

parameters are such that  $\theta_i$  belongs to  $\Theta_1 \cap \Theta_2$  and  $\theta_2$  whether we associate  $\theta_1$  to  $\Theta_1$  or  $\Theta_2$ , we have two or only case, we choose the injection which leads to as many corr

We denote by  $n$  the maximum number of different  $\theta_i$ s t note by  $A_n^k$  the region of parameter space with  $k$  component in each of the  $\Theta_i$ ,  $i = 1..m$ ,  $m \leq n$ .

If  $k$  planets are truly present in the data,  $n$  detections are c means that the true detections of  $\min(k, n) - i$  are missed. V adding a term  $-\beta(\min(k, n) - i)$  whenever it happens. The e

(19)

$$E_{\theta, \eta} [U \{a, (\theta, \eta)\}] = -n \alpha p(0 | y)$$

$$+ [-(n-1)\alpha I_{A_1^k} - (n\alpha + \beta)(1 - I_{A_1^k})] p(1 | y)$$

$$+ [-(n-2)\alpha I_{A_2^k} - ((n-1)\alpha + \beta)I_{A_2^k} - (n\alpha + 2\beta)(1 - I_{A_2^k})]$$

$\vdots$

$$(23) \quad + \sum_{i=1}^k [-(n-i)\alpha - (k-i)\beta] I_{A_i^k} - (n\alpha + k\beta) \left(1 - \sum_{i=1}^k I_{A_i^n}\right)$$

(24)

$$\vdots$$

$$(25) \quad + \sum_{i=1}^n [-(n-i)\alpha + \beta] I_{A_i^n} - n(\alpha + \beta) \left(1 - \sum_{i=1}^n I_{A_i^n}\right)$$

$$(26) \quad + \sum_{i=1}^n [-(n-i)\alpha + \beta] I_{A_i^{n+1}} - n(\alpha + \beta) \left(1 - \sum_{i=1}^n I_{A_i^{n+1}}\right)$$

(27)

$$\vdots$$

26

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

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

while the maximum utility problem can be

$$\min_n \left( \dots \right)$$

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 constraint  $n < n_0$ , we can choose  $\gamma$  such that the solution of the maximization satisfying the constraint. We show th to a larger value of  $v_n + \frac{1}{\gamma} u_n$ .

From our hypotheses, we see that the rat  $n \leq n_0$ , we will have

$$v_n + \frac{1}{\gamma} u_n$$

if we take

$$(47) \quad \gamma \geq \frac{u}{v}$$

Note that if  $v_{n_0} - v_{n_0+1} = 0$ , since  $v_n - v_{n-1}$  also  $v_n - v_{n+1} = 0$ , and we can always re such that  $v_n - v_{n+1} \neq 0$ . For  $n > n_0$ , we w

$$v_n + \frac{1}{\gamma} u_n$$

if

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

These two conditions can be satisfied sin

$$\frac{u_{n_0} - u_{n_0+1}}{v_{n_0+1} - v_{n_0}}$$

Choosing  $\gamma$  between those two bounds g increasing function of  $x$ .

As long as the sequence  $(u_{n+1}^y - u_n^y)_{n \in \mathbb{N}}$  the constrained problem have the same solt but can be ensured under the following con

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

$$+ \sum_{i=1}^n -(n-i)(\alpha + \beta) I_{A_i^n} - n(\alpha + \beta) \left(1 - \sum_{i=1}^n I_{A_i^{n+1}}\right) - (n_{max} - n)$$

Re-arranging the terms, we have

$$(29) \quad E_{\theta, \eta} [U \{a, (\theta, \eta)\}] = -n\alpha + (\alpha + \beta) \sum_{i=1}^n i I_{A_i^n} - \beta \sum_{k=1}^{n_{max}} k p(k)$$

$$(30) \quad = -(\alpha E[\text{FD}] + \beta E[\text{MD}])$$

where  $E[\text{FD}]$  and  $E[\text{MD}]$  are the expected numbers of false detections and when claiming the detection of components with parameters in  $\Theta_1, \dots, \Theta_n$ ,

$$(31) \quad E[\text{FD}] = n \sum_{i=1}^n i I_{A_i^n}$$

$$(32) \quad E[\text{MD}] = \bar{n} - \sum_{i=1}^n i I_{A_i^n},$$

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

$$(33) \quad E_{\theta, \eta} [U \{a, (\theta, \eta)\}] = -n + (1 + \gamma) \sum_{i=1}^n i I_{A_i^n} - \gamma \bar{n}.$$

Where  $\bar{n}$  do not depend on the number of planets

CONTINUOUS MULTIPLE HYPOTHESIS TESTING FOR OPTIMAL EXOPLANET DETECTION

LEMMA E.1. Denoting  $I_{\Theta_i}$  where  $\llbracket 1, n \rrbracket_j$  is a draw of  $j$  indices without replacement from  $\llbracket 1, n \rrbracket$ , we have

$$(34) \quad I_{A_j} = \sum_{k_1, \dots, k_j \in \llbracket 1, n \rrbracket_j} I_{\Theta_{k_1} \wedge \Theta_{k_2} \wedge \dots \wedge \Theta_{k_j}}$$

where  $I_{\Theta_i}$  is

Then

$$(37) \quad \sum_{i=1}^n I_{\Theta_i} = \sum_{i=1}^{n-1} \sum_{j=k_1, \dots, k_j \in \llbracket 1, n \rrbracket_j \setminus \{i\}} I_{\Theta_i \wedge \Theta_{k_1} \wedge \dots \wedge \Theta_{k_j}}$$

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

ENCE OF

be guaran s denote b:

THE arabilian d

(13) ha

The and mi

and mi

PROOF. Let us suppose that there exists  $n$  such that  $u_{n+1} < u_n$

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

ent to

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

by  $i_0$  an index such that  $i_0 = \arg \max_{i=1..n+1} I_{\Theta_i^{n+1}}$

28

The term  $I_{\Theta_{i_0}^{n+1}} + \sum_{i=1}^{n-1} I_{\Theta_i^n}$  is a sum of  $I_{\Theta_i}$  with disjoint  $\Theta_i$ . By d

hand side of the inequality is less than or equal to 0 and the left hand greater than or equal to 0, which is absurd.

If  $\forall i = 1..n+1$ ,  $\exists j = 1..n-1$ ,  $\Theta_i^{n+1} \cap \Theta_j^{n-1} \neq \emptyset$ . In that case  $\llbracket 1, n \rrbracket \Theta_i^n \cap \Theta_j^{n-1} \neq \emptyset$ , otherwise due to lemma 3.2 this would lead to

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

THEOREM E.5. Let us consider a dataset  $y$  and suppose that it

arability at all orders,  $n = 1..n_{max}$  then there exists an increasing that (63) and (64) have the same solution, and a function  $\gamma'(x) > 0$  have the same solution.

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 DE

In this appendix, we show that the optimal procedure is similar i defined as in Hara et al. (2022b).

DEFINITION F.1 (Missed detections: other definition). If  $n$  com

in fact there are  $n' > n$  components truly present in the data, we cot

onsider

an incre

problem (

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

$$(61) \quad E[\text{MD}] = \sum_{k=n+1}^{n_{max}} (k-n)p(k | y)$$

Assuming that  $\alpha \neq 0$  (or equivalently  $\alpha > 0$ , since  $\alpha$  is non negative), we can divide Eq. (59) by  $\alpha$ . Denoting by  $\gamma = \beta/\alpha$ , without loss of generality we can maximize

$$(62) \quad E_{\theta, \eta} [U \{a(\Theta_1, \dots, \Theta_n), (\theta, \eta)\}] = -n + \sum_{j=1}^n j I_{A_j} - \gamma \sum_{k=n+1}^{n_{max}} (k-n)p(k | y).$$

The expected utility is

$$(63) \quad E_{\theta, \eta} [U \{a(\Theta_1, \dots, \Theta_n), (\theta, \eta)\}] = -n + \sum_{i=1}^n i I_{A_i} - \gamma \sum_{i=1}^{n_{max}} (k-n)p(k | y)$$

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

$$(38) \quad \arg \max_{\Theta_1 \in T \setminus \Theta_{n+1}, \dots, \Theta_n \in T \setminus \Theta_{n+1}} I_{\Theta_{n+1}} + \sum_{i=1}^n I_{\Theta_i}$$

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

$$(39) \quad \arg \max_{\Theta_1 \in T \setminus \Theta_{n+1}, \dots, \Theta_n \in T \setminus \Theta_{n+1}} \sum_{i=1}^n I_{\Theta_i} = (\Theta_i^n)_{i=1..n}$$

As a consequence,

$$(40) \quad \arg \max_{\Theta_{n+1} \in T \setminus \Theta_{n+1}} \arg \max_{\Theta_1 \in T \setminus \Theta_{n+1}, \dots, \Theta_n \in T \setminus \Theta_{n+1}} \max_{\forall i, j \in \llbracket 1, n \rrbracket, i \neq j, \Theta_i \cap \Theta_j = \emptyset} I_{\Theta_{n+1}} + \sum_{i=1}^n I_{\Theta_i} = (\Theta_i^n)_{i=1..n}$$

up to a permutation of the indices (see remark B.2). If  $\exists i \in \llbracket 1, n+1 \rrbracket, \forall j$ ,  $\Theta_j^n = \emptyset$  then the same argument applies and the solution to  $(P_{n+1})$  is  $(\Theta_1^n, \dots,$

Let  $i \in \llbracket 1, n \rrbracket$ , CONINUOUS MULTIPLE HYPOTHESIS TESTING FOR OPTIMAL cases at

Theorem E.1 assumes component separability (Definition 3.3)

than necessary for some of the lemmas, and is not made by de

throughout the appendix that the  $\Theta_i$ s are pairwise disjoint. Tl

Eq. (4), we can write  $n - \sum_{j=1}^n j I_{A_j} = \sum_{i=1}^n \text{FIP}_{\Theta_i}$ . We conside

and define

$$(41) \quad u_n := \sum_{i=1}^n \text{FIP}_{\Theta_i}$$

$$(42) \quad v_n := \sum_{k=n+1}^{n_{max}} (k-n)p(k | y); \quad v'_n = \bar{n} - n +$$

where  $\bar{n} = \sum_{k=1}^{n_{max}} k p(k | y)$ . The sequence  $u_n$  is the expected n

and  $v'_n$  are the expected number of missed detections for the mi

F.1 and 2.3, respectively (see Appendix A for details). Note that

but we chose not to make that dependence explicit to simplify n

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

PROOF. Let us suppose that there exists  $n$  such that  $u_{n+1} < u_n$

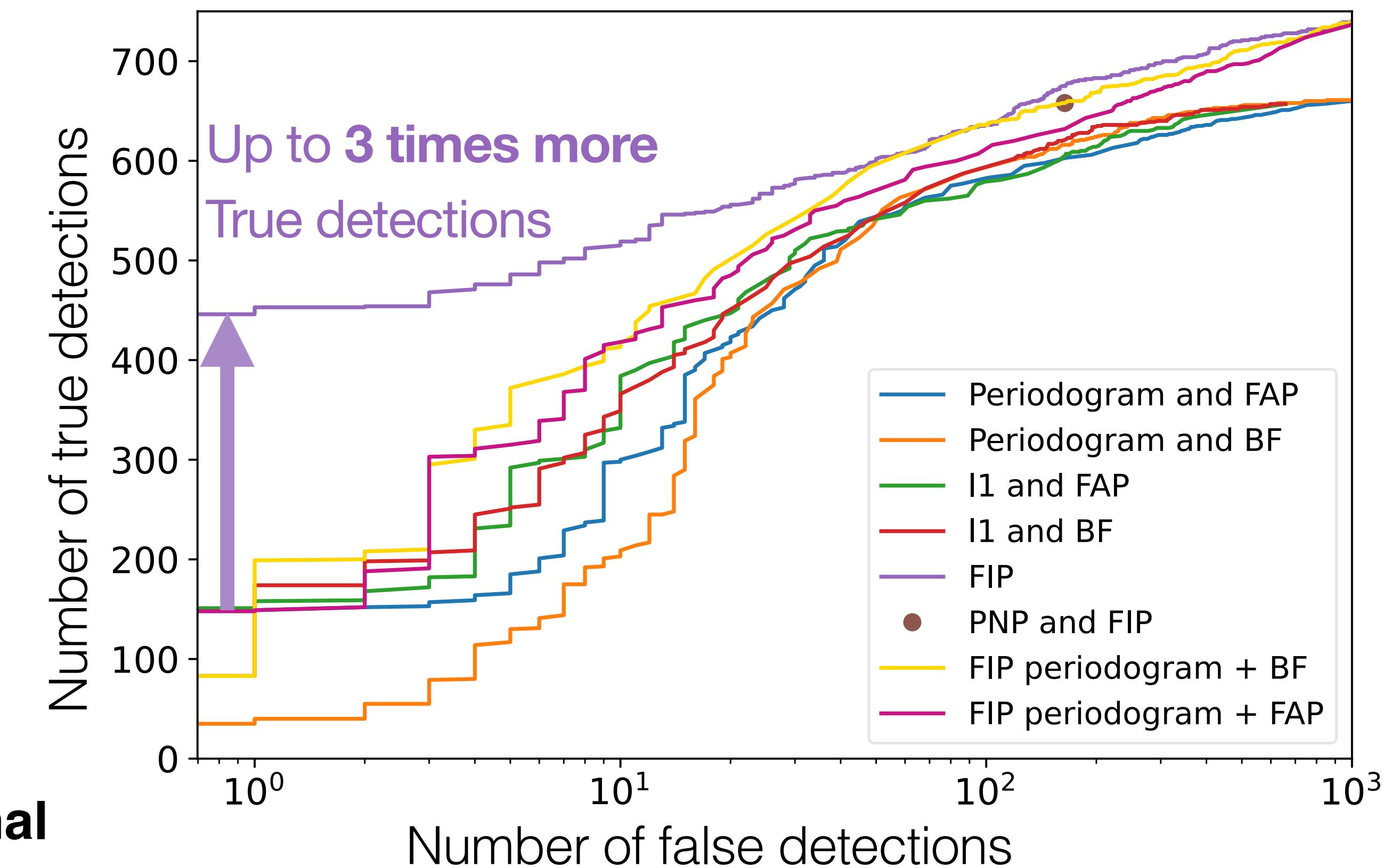
# Optimal criterion to detect exoplanets (3)

- **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**
- **Optimal for all exoplanet detection methods**

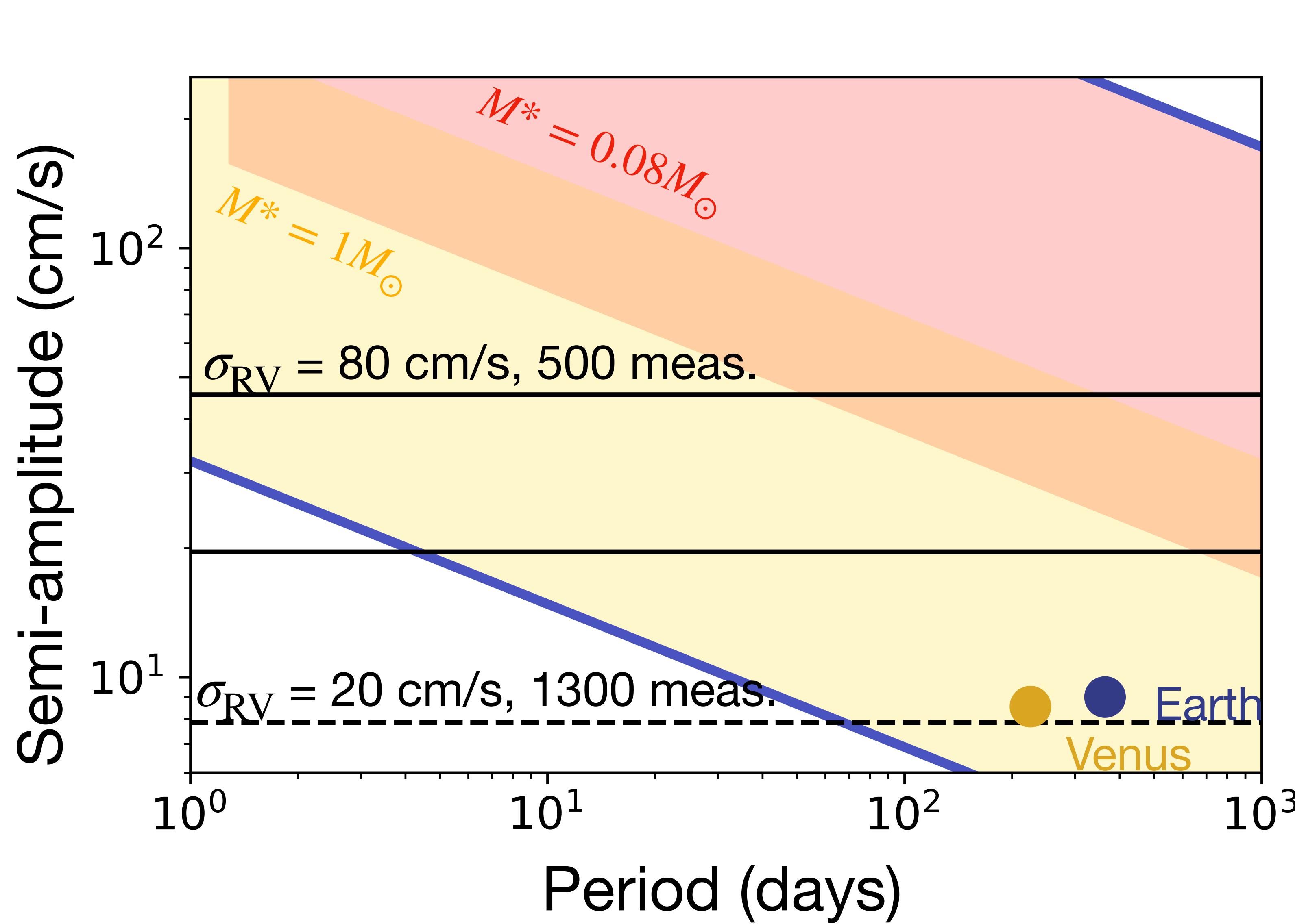


In a collection of independent detections made with TIP 99%, on average 99% are correct

90%  
50%

90%  
50%

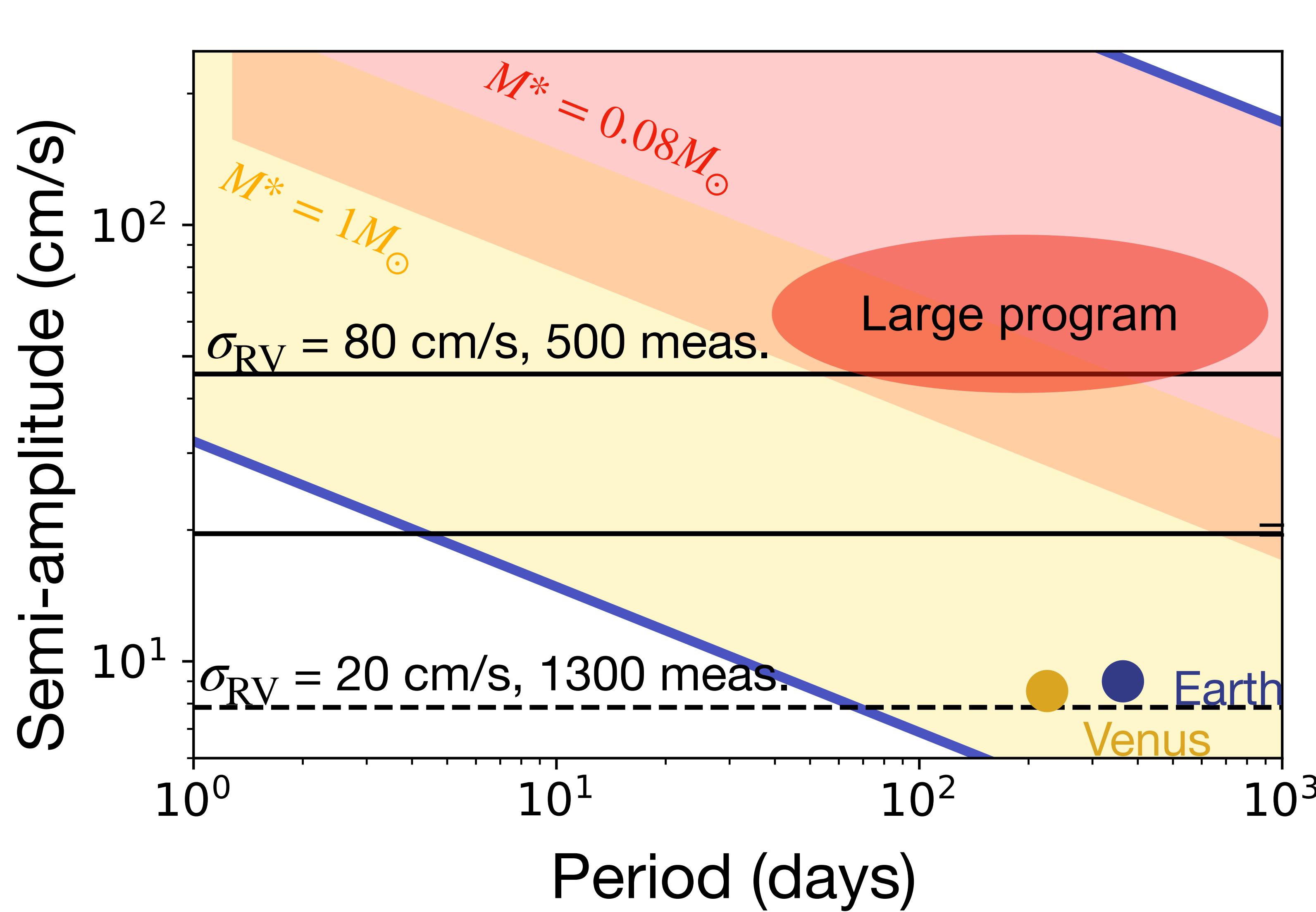
# Motivation for a precursor survey



Shaded areas:  
semi amplitude of RV signals  
planets from 0.5 to 5 Earth masses, 0.01-10 AU for a  $1M_\odot$  star in yellow, and a  $0.08 M_\odot$  in red.

Horizontal lines:  
semi amplitudes measurable with a 10% precision on mass assuming **uncorrelated**, Gaussian error of amplitude  $\sigma_{RV}$

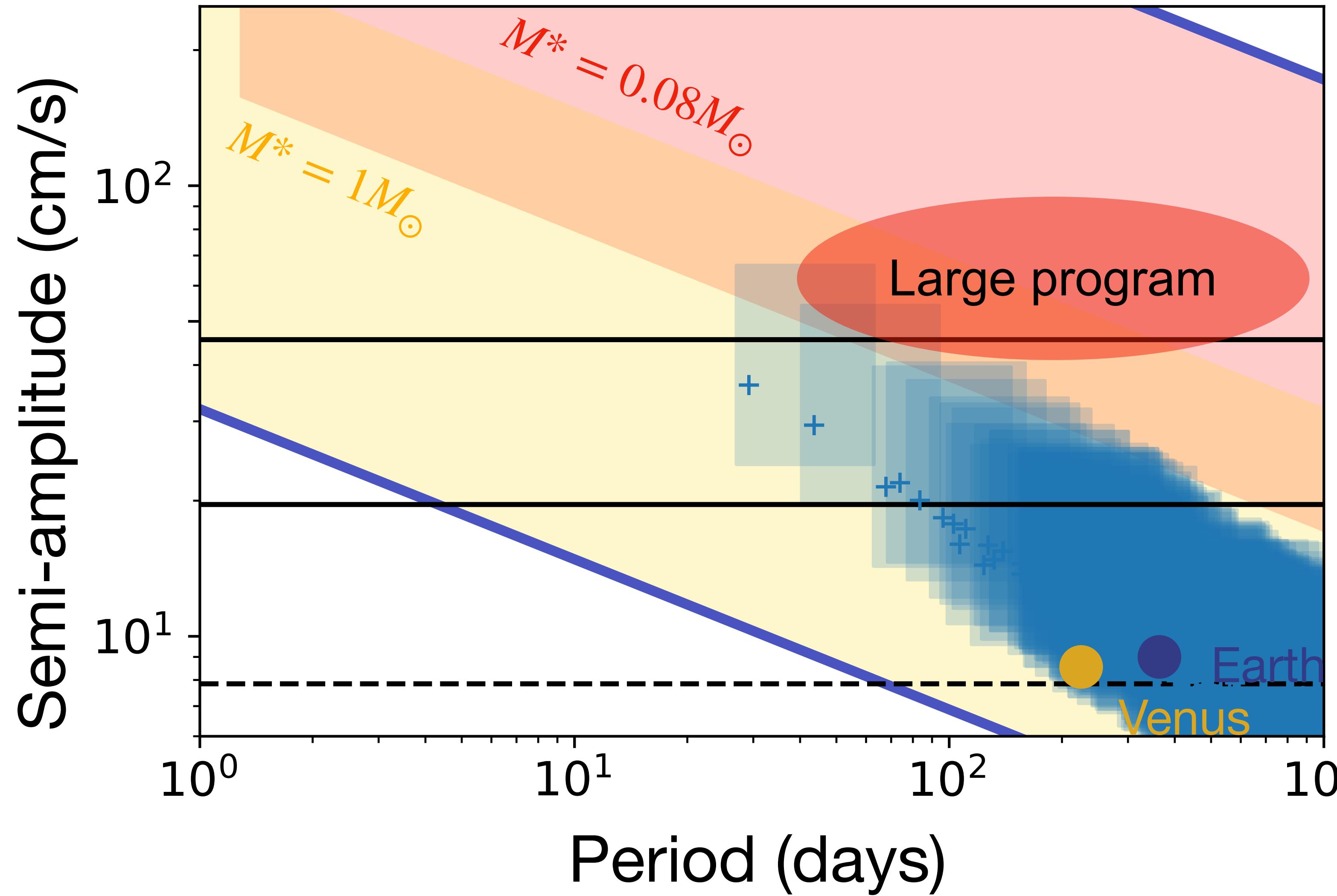
# Motivation for a precursor survey



Shaded areas:  
semi amplitude of RV signals  
planets from 0.5 to 5 Earth masses, 0.01-10 AU for a  $1M_\odot$  star in yellow, and a  $0.08 M_\odot$  in red.

Horizontal lines:  
semi amplitudes measurable with a 10% accuracy assuming **uncorrelated**, Gaussian error of amplitude  $\sigma_{RV}$

# Motivation for a precursor survey



Expected RV amplitude for the  
Mamajek+ 2024 HWO target list  
[0.95, 1.67] AU, [0.8, 1.4]  $M_\oplus$

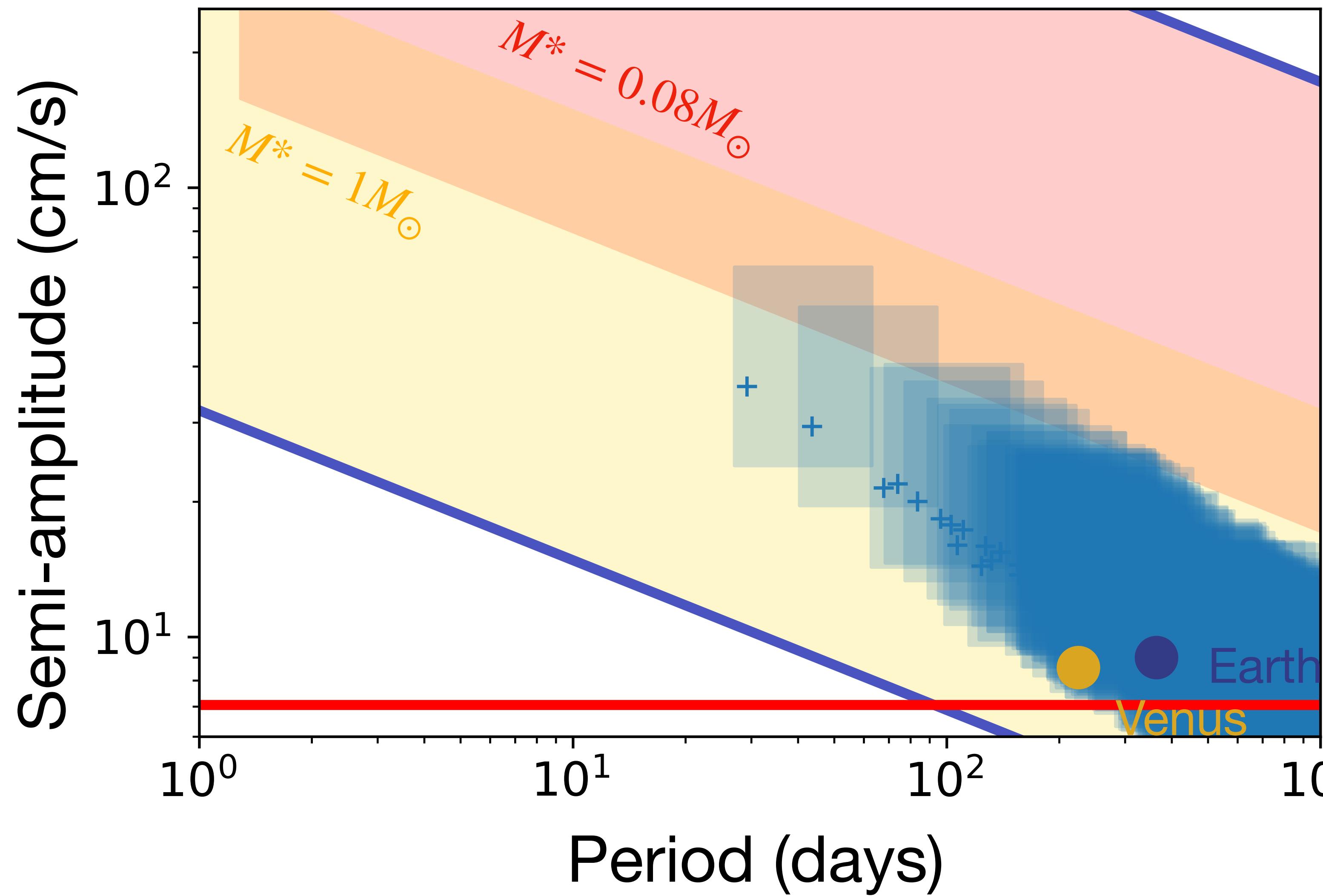
$\sigma_{\text{RV}} = 80 \text{ cm/s, 500 meas.}$

$\sigma_{\text{RV}} = 50 \text{ cm/s, 1300 meas.}$   
or  $\sigma_{\text{RV}} = 20 \text{ cm/s, 208 meas.}$

$\sigma_{\text{RV}} = 20 \text{ cm/s, 1300 meas.}$

10% precision on mass =  
10 sigma detection

# Motivation for a precursor survey



Expected RV amplitude for the  
Mamajek+ 2024 HWO target list  
[0.95, 1.67] AU, [0.8, 1.4]  $M_\oplus$

$\sigma_{\text{RV}} = 30 \text{ cm/s, 1300 meas, 6}$   
sigma detection

# NASA Extreme precision radial velocity report (Crass+ 2021)

Existing high precision spectrograph:

HARPS, HARPS-N, **ESPRESSO**, NIRPS (for M dwarves)/ SOPHIE, SPIRou (Europe-led)  
NEID, KPF (US)

Existing network: NASA Extreme precision RV initiative

Upcoming:

**HARPS3 (Cambridge)**, G-Clef (Harvard)

Some key findings of the EPRV report:

- Establishment of a NASA EPRV Research Coordination Network and Standing Advisory
- More interdisciplinary collaboration (with stellar physicists in particular)
- Collaborative work on standard datasets (especially Sun-fed spectrographs, coordinate observations with major instruments on a small set of bright standard stars)

# Other topics

Achieving <1 cm/s calibration accuracy with laser frequency combs

Observation strategy: how to most efficiently spend observation time?

Telluric contamination: how to remove it?

Is spectropolarimetry interesting to correct stellar variability?

Is there an advantage to using adaptive optics?

Target list for PCS: NIRPS spectrograph in the infrared

HARPS3 spectrograph dedicated to HWO target scanning

Impact of RV survey on *LIFE*

# Conclusion

RV precursor surveys could greatly improve the yield of *HWO*, *LIFE*, PCS

Needs work on

- High resolution spectrograph hardware (better calibration source, adaptive optics)
- Data analysis methods (telluric, systematics, stellar variability)
- Lobbying: getting funds to push precursor surveys
- Quantifying the impact on *LIFE*
- What is the required precision on mass measurements to interpret atmosphere observations?
- Are there data analysis techniques that could be transferred between RV, transits and imaging?



# References

**Target list:** Mamajek et al. 2024, Gaudi et al. 2020

**Impact of RV survey,** Morgan et al. 2021, Crass et al. 2021

**Roadmap to Extreme Precision Radial Velocity:** Crass et al. 2021

**Statistical methods for exoplanet detection with radial velocities:** Hara & Ford 2023

# By-product: statistical Doppler imaging

