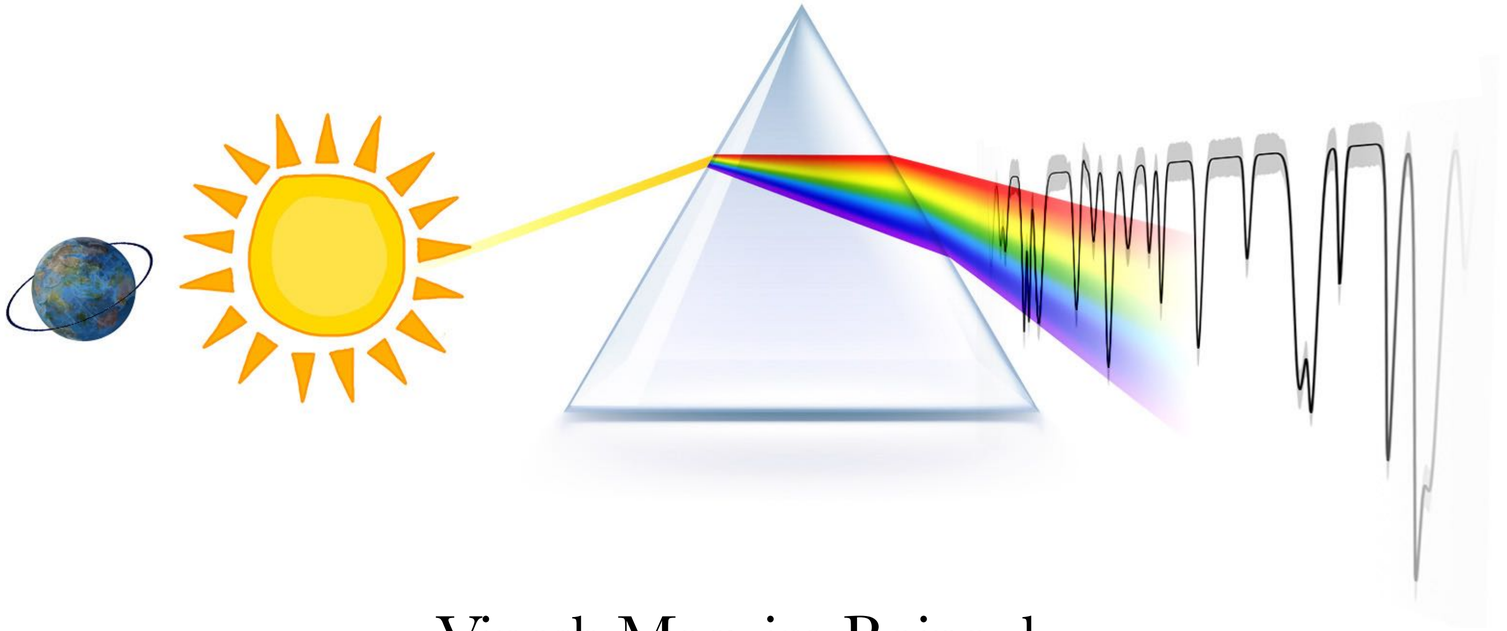
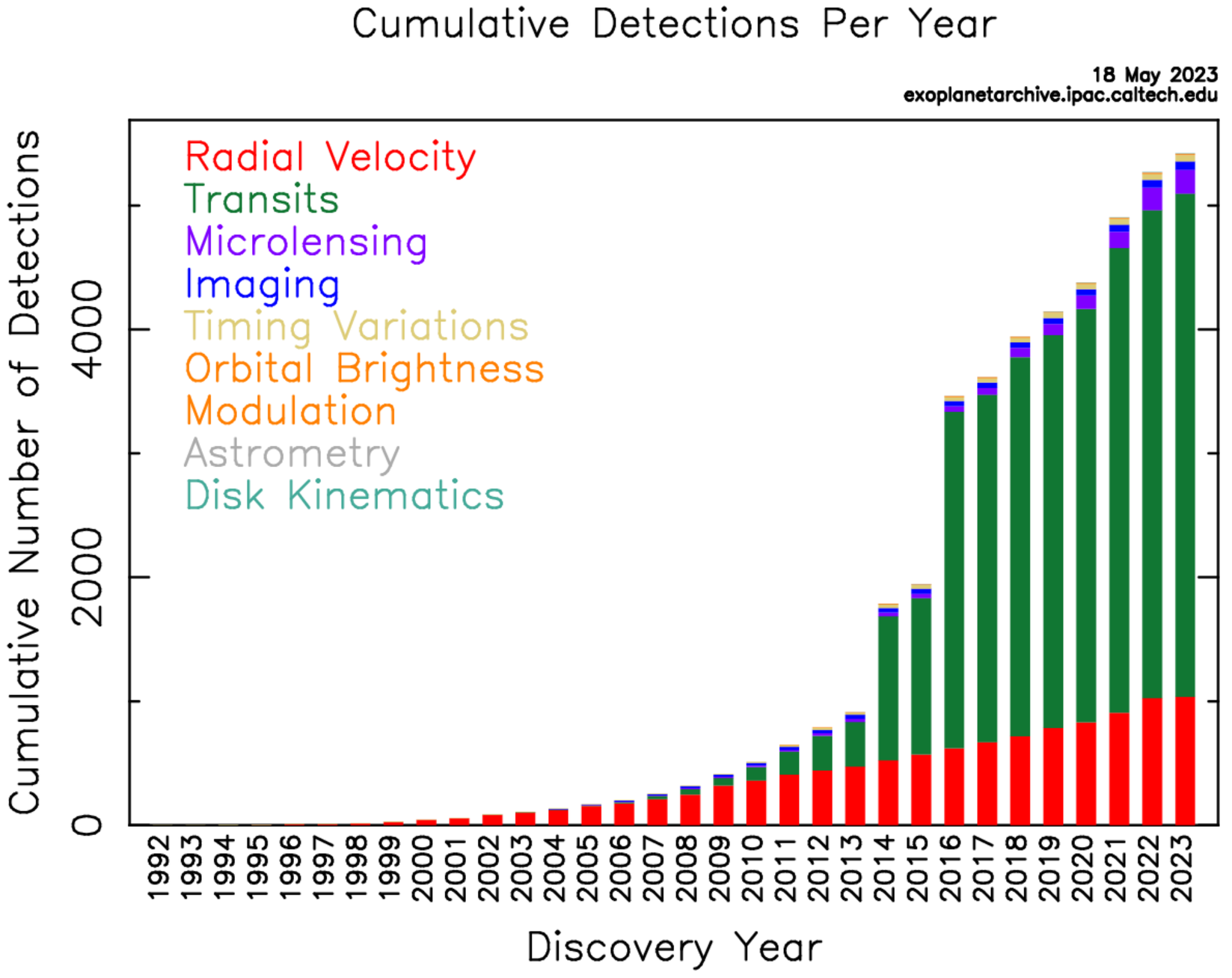
ASTRO-STATISTICAL BARRIERS TO THE FIRST EARTH-TWIN EXOPLANETS

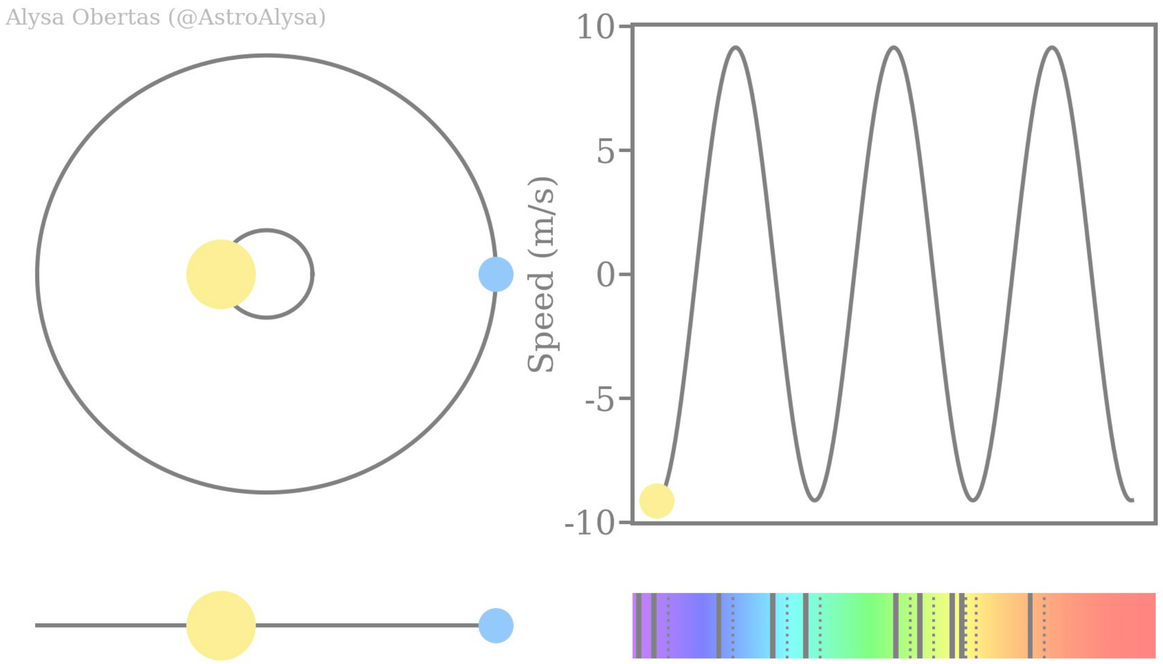


Vinesh Maguire-Rajpaul

UNIVERSITY OF CAMBRIDGE
Cavendish Laboratory


Cumulative Detections Per Year
18 May 2023
exoplanetarchive.ipac.caltech.edu
Radial Velocity
Transits
Microlensing
Imaging
Timing Variations
Orbital Brightness
Modulation
Astrometry
Disk Kinematics
Cumulative Number of Detections
0
2000
4000
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
Discovery Year

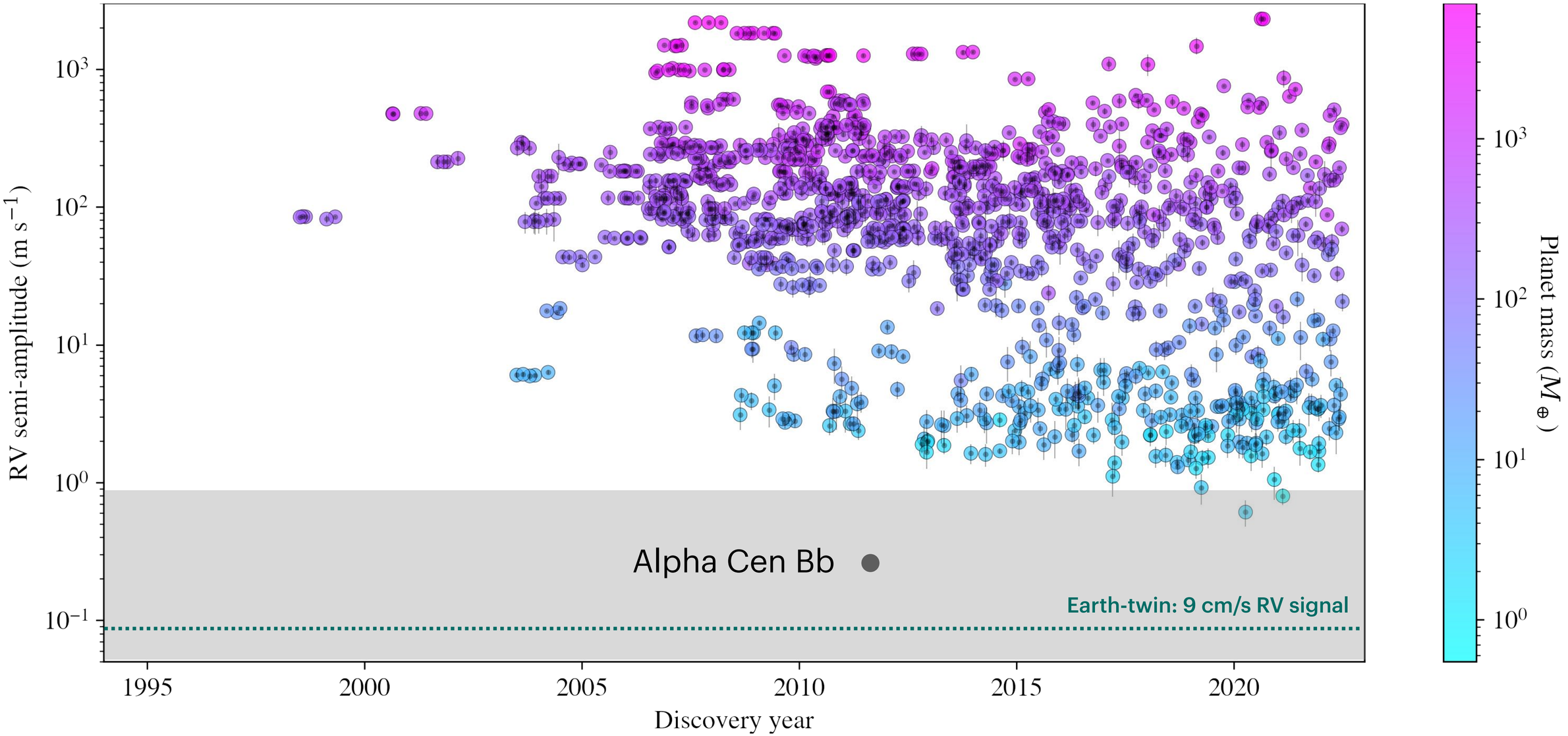
THE RV METHOD



Movie credit **@AstroAlysa**

- Measures **masses** and orbital **eccentricities**
- Constraints **density** & **surface gravity** of transiting planets
- Supports **imaging** & **atmosphere** observations
- Reveals **architecture** of planetary systems

RV detections vs time



Article metrics | Last updated: Sat, 28 Jan 2023 16:45:14 Z

An Earth-mass planet orbiting α Centauri B

Access & Citations

7355	297	315
Article Accesses	Web of Science	CrossRef

Online attention

- 106 tweeters
- 2 Google+ users
- 202 Mendeley
- 18 blogs
- 46 news outlets
- 3 Citeulike
- 5 Facebook pages
- 32 Wikipedia page

JOURNAL ARTICLE

Ghost in the time series: no planet for Alpha Cen B

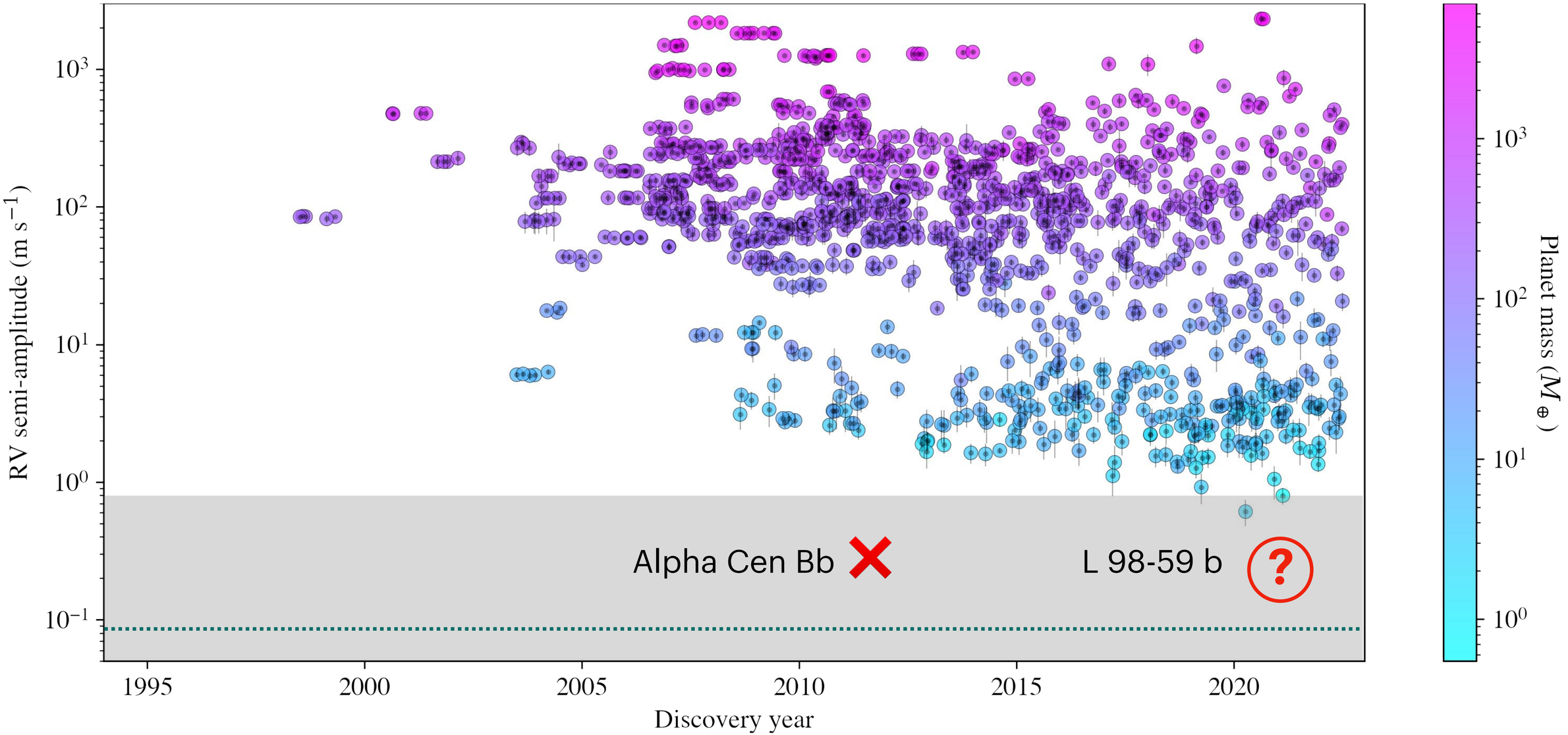
FREE

V. Rajpaul ✉, S. Aigrain ✉, S. Roberts

Monthly Notices of the Royal Astronomical Society: Letters, Volume 456, Issue 1, 11 February 2016, Pages L6–L10, https://doi.org/10.1093/mnrasl/slv164

Published: 20 November 2015 **Article history** ▾

RV detections vs time



Warm terrestrial planet with half the mass of Venus transiting a nearby star

Overview of attention for article published in Astronomy and Astrophysics, September 2021

SUMMARY | News | Blogs | Twitter | Wikipedia | Reddit | Dimensions citations

Title	Warm terrestrial planet with half the mass of Venus transiting a nearby star
Published in	Astronomy and Astrophysics, September 2021
DOI	10.1051/0004-6361/202140728
Authors	O. D. S. Demangeon, M. R. Zapatero Osorio, Y. Alibert, S. C. C. Barros, V. Adibekyan, H. M... [show]

A non-zero Doppler amplitude is not enough: revisiting the putative radial velocity detection of sub-Venus exoplanet L 98-59 b

Vinesh M. Rajpaul,[1]★ Oscar Barragán,[2] and Norbert Zicher[2]
[1]*Astrophysics Group, Cavendish Laboratory, JJ Thomson Avenue, Cambridge CB3 0HE , UK*
[2]*Sub-department of Astrophysics, Department of Physics, University of Oxford, Oxford, OX1 3RH, UK*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

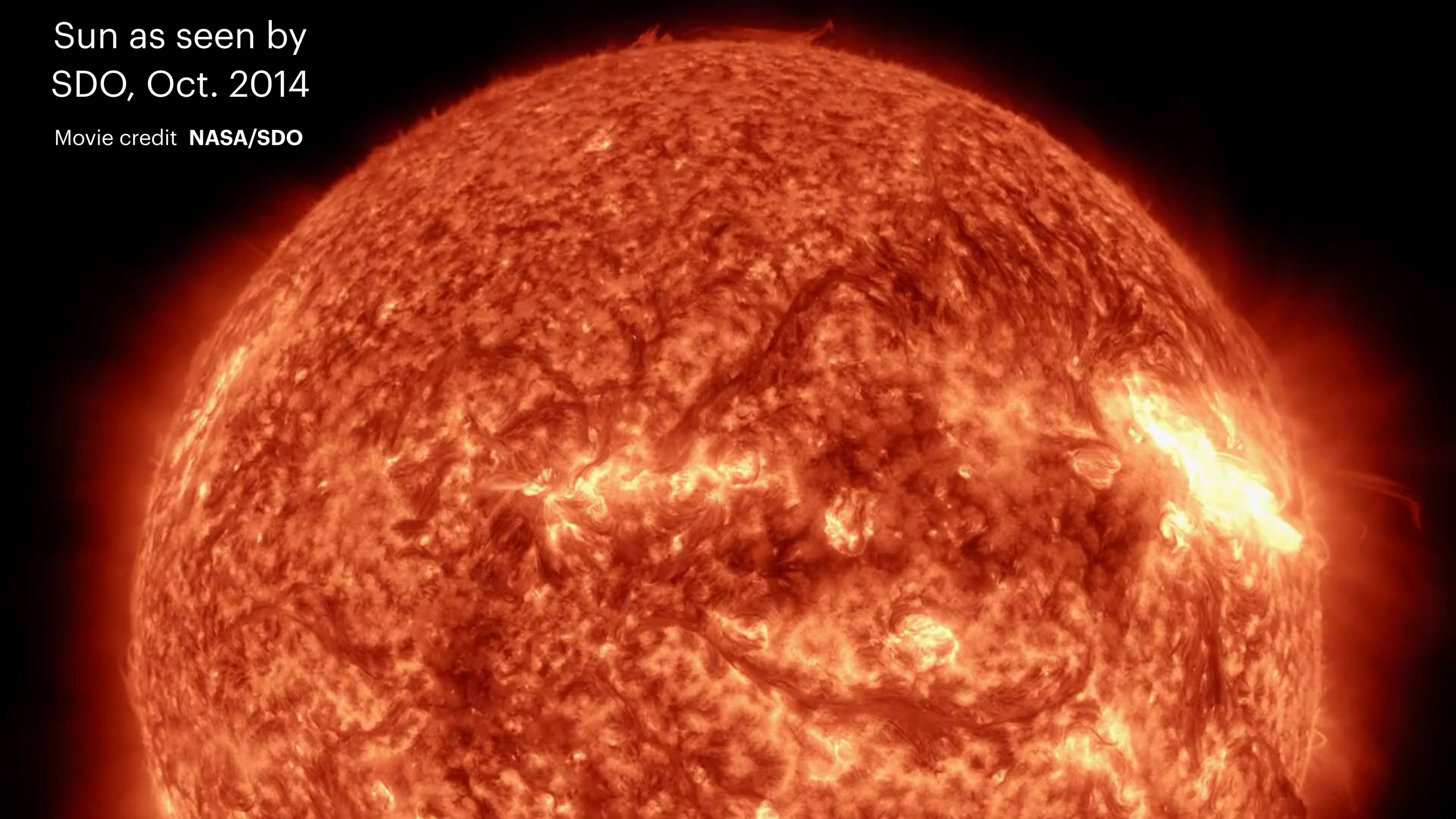
L 98-59 b is an exoplanet with radius 0.85 $R_\oplus$ and orbital period 2.25 d, discovered transiting a nearby M-dwarf by TESS in 2019. In 2021, a Doppler detection of L 98-59 b was announced, with radial velocity (RV) semi-amplitude $K_b = 0.46^{+0.20}_{-0.17}$ m s^{-1} inferred from 164 HARPS and 66 ESPRESSO RVs. This translated into a mass of $M_b = 0.40^{+0.16}_{-0.15}$ $M_\oplus$, or half the mass of Venus, making L 98-59 b (by far) the lowest-mass planet with a putative RV detection. To illuminate the difficulties of making secure Doppler detections of < 1 m s^{-1} signals, we argue here that the published RV detection of L 98-59 b is not supported by sufficient statistical evidence. We show that under the published modelling assumptions, Bayesian model comparison (BMC) suggests a $\sim 50\%$ false-detection probability. We also often infer ~ 40 cm s^{-1} semi-amplitudes for Keplerians with periods not corresponding to any known planet; importantly, though, these 'detections' are all rejected by BMC. By implementing a more sophisticated stellar activity model and more realistic parameter priors than in the published model, we infer a mass of $M_b = 0.47 \pm 0.14$ $M_\oplus$ from semi-amplitude $K_b = 0.56 \pm 0.16$ m s^{-1}, with BMC now indicating substantial albeit still not decisive evidence for the detection. It remains unclear whether such a detection is consistent with the data quality, and we suggest that definitive detection and characterisation may require many additional RVs, alongside careful modelling and stringent statistical analysis. Our case study has implications for current extreme-precision RV campaigns and the longer-term quest to discover the first Earth-twins.

What gives?

1. Stellar nuisance signals

Sun as seen by SDO, Oct. 2014

Movie credit **NASA/SDO**



STARS ARE ACTIVE

Earth to Scale

Image credit **NASA/SDO**

STELLAR ACTIVITY

STELLAR OR PLANETARY SIGNAL...?

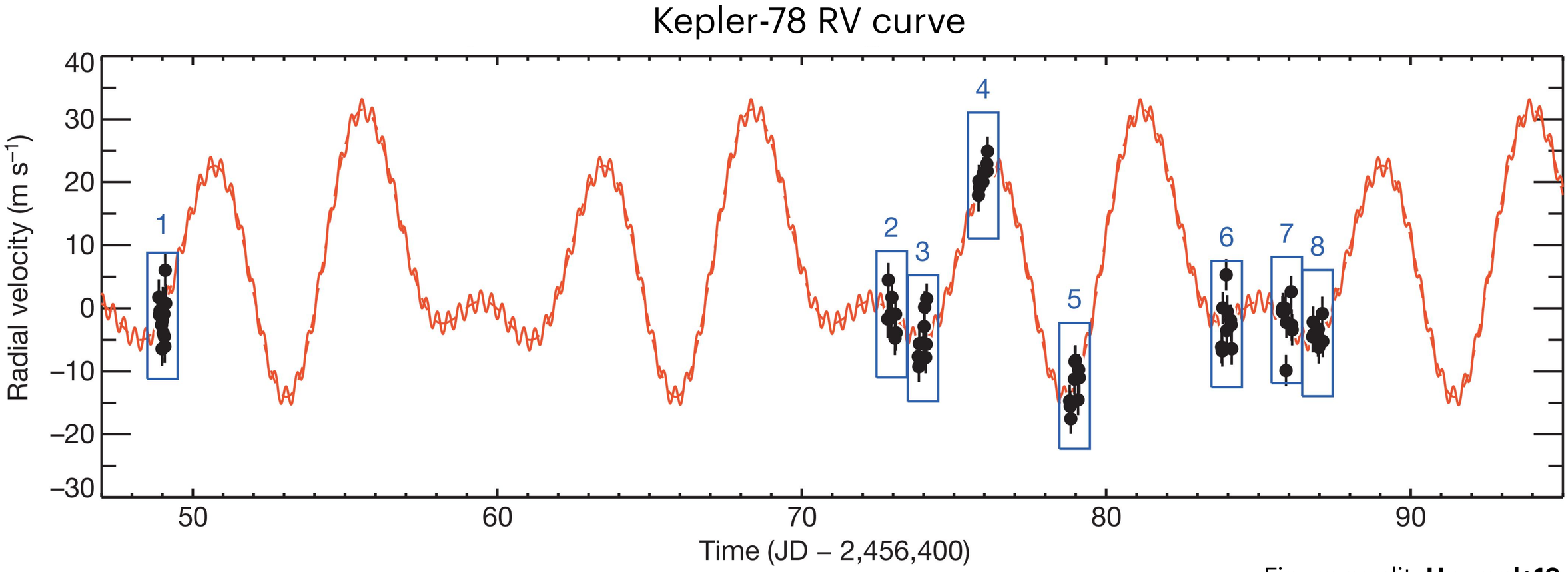


Figure credit **Howard+13**

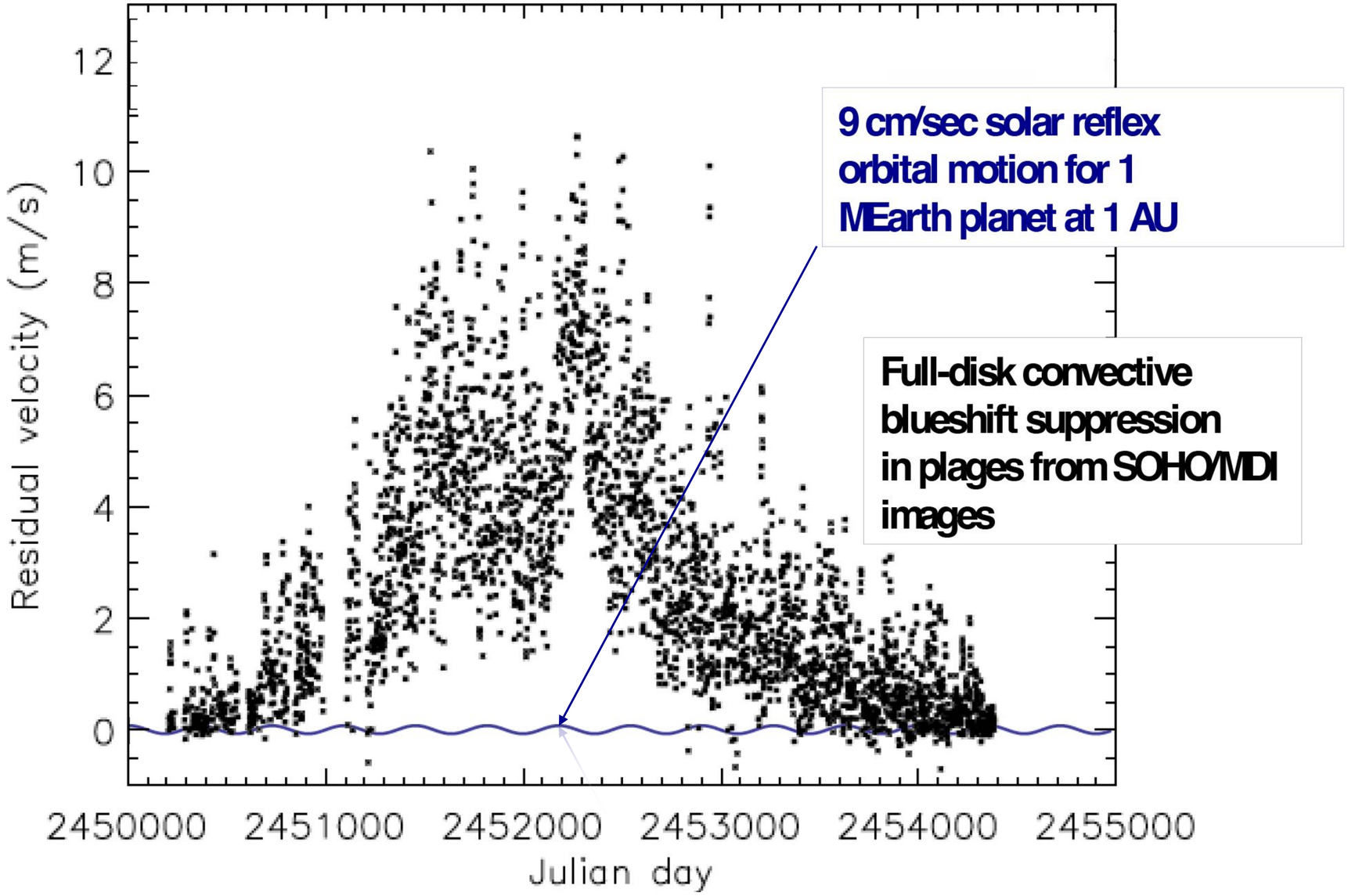


Figure credit **Meunier+10, Mortier 15**

STELLAR ACTIVITY

PROPERTIES OF ROTATIONALLY-MODULATED SIGNALS

- **Time scales similar** to those associated with **planets** (days to years)
- **Quasi-periodic** (periodic rotation + evolving active regions + activity cycles)
- **Somewhat smooth** (active regions don't change instantaneously)
- **Stochastic** (active regions seem to appear randomly)

STELLAR ACTIVITY

TRYING TO MODEL IT PARAMETRICALLY

[...] fitting sine waves at the rotational period of the star and the significant harmonics [...]

The global model fitted on the RVs is therefore:

$$
\begin{aligned}
\text{subset 2008} \quad : \quad & lin0 + lin1 \cdot JDB_{2008} + lin2 \cdot JDB_{2008}^2 + A_{RV-Rhk} \cdot RHK_{low\,freq,2008} \\
\text{subset 2009} \quad : \quad & lin0 + lin1 \cdot JDB_{2009} + lin2 \cdot JDB_{2009}^2 + A_{RV-Rhk} \cdot RHK_{low\,freq,2009} \\
& + A11s \cdot sin(\frac{2\pi}{P1}) \cdot JDB_{2009} + A11c \cdot cos(\frac{2\pi}{P1}) \cdot JDB_{2009} \\
& + A12s \cdot sin(\frac{2\pi}{P1/2}) \cdot JDB_{2009} + A12c \cdot cos(\frac{2\pi}{P1/2}) \cdot JDB_{2009} \\
\text{subset 2010} \quad : \quad & lin0 + lin1 \cdot JDB_{2010} + lin2 \cdot JDB_{2010}^2 + A_{RV-Rhk} \cdot RHK_{low\,freq,2010} \\
& + A21s \cdot sin(\frac{2\pi}{P2}) \cdot JDB_{2010} + A21c \cdot cos(\frac{2\pi}{P2}) \cdot JDB_{2010} \\
& + A23s \cdot sin(\frac{2\pi}{P2/3}) \cdot JDB_{2010} + A23c \cdot cos(\frac{2\pi}{P2/3}) \cdot JDB_{2010} \\
& + A24s \cdot sin(\frac{2\pi}{P2/4}) \cdot JDB_{2010} + A24c \cdot cos(\frac{2\pi}{P2/4}) \cdot JDB_{2010} \\
\text{subset 2011} \quad : \quad & lin0 + lin1 \cdot JDB_{2011} + lin2 \cdot JDB_{2011}^2 + A_{RV-Rhk} \cdot RHK_{low\,freq,2011} \\
& + A31s \cdot sin(\frac{2\pi}{P3}) \cdot JDB_{2011} + A31c \cdot cos(\frac{2\pi}{P3}) \cdot JDB_{2011} \\
& + A32s \cdot sin(\frac{2\pi}{P3/2}) \cdot JDB_{2011} + A32c \cdot cos(\frac{2\pi}{P3/2}) \cdot JDB_{2011} \\
& + A33s \cdot sin(\frac{2\pi}{P3/3}) \cdot JDB_{2011} + A33c \cdot cos(\frac{2\pi}{P3/3}) \cdot JDB_{2011}
\end{aligned}
$$

23 parameters just for stellar activity!

What about Gaussian processes?

a **GP** is the **infinite-dimensional** version of a **Gaussian distribution**

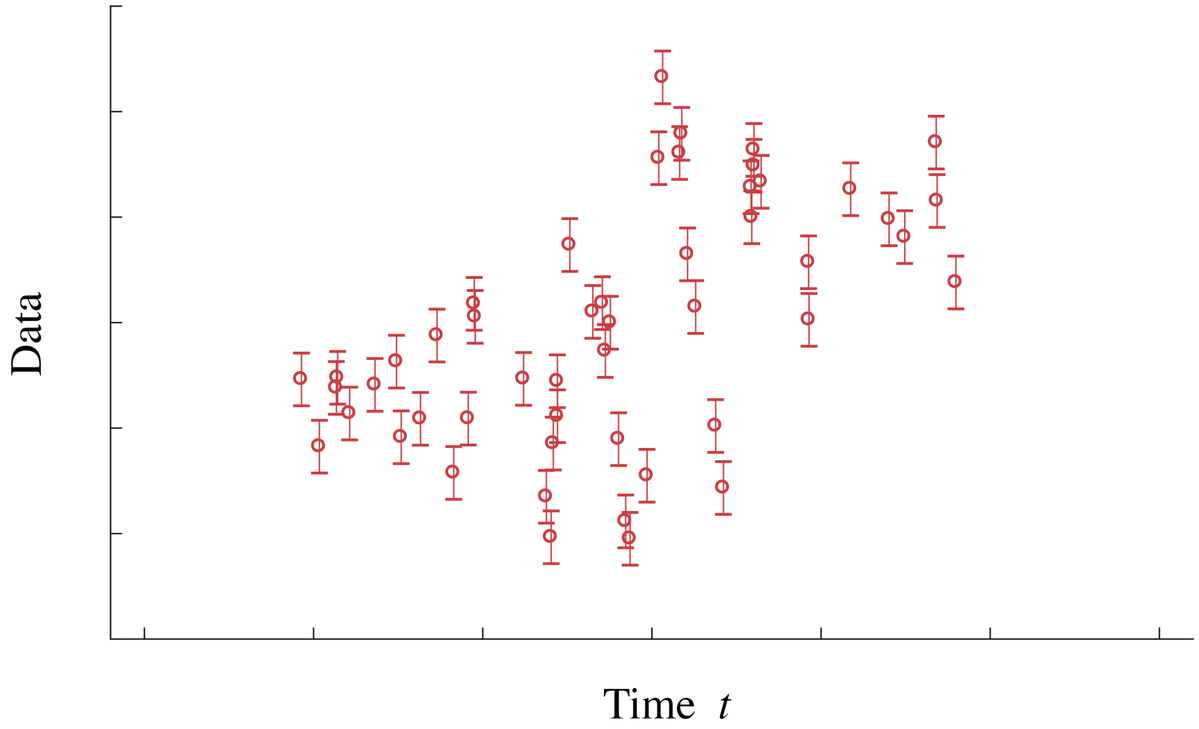
INTRODUCING GPs

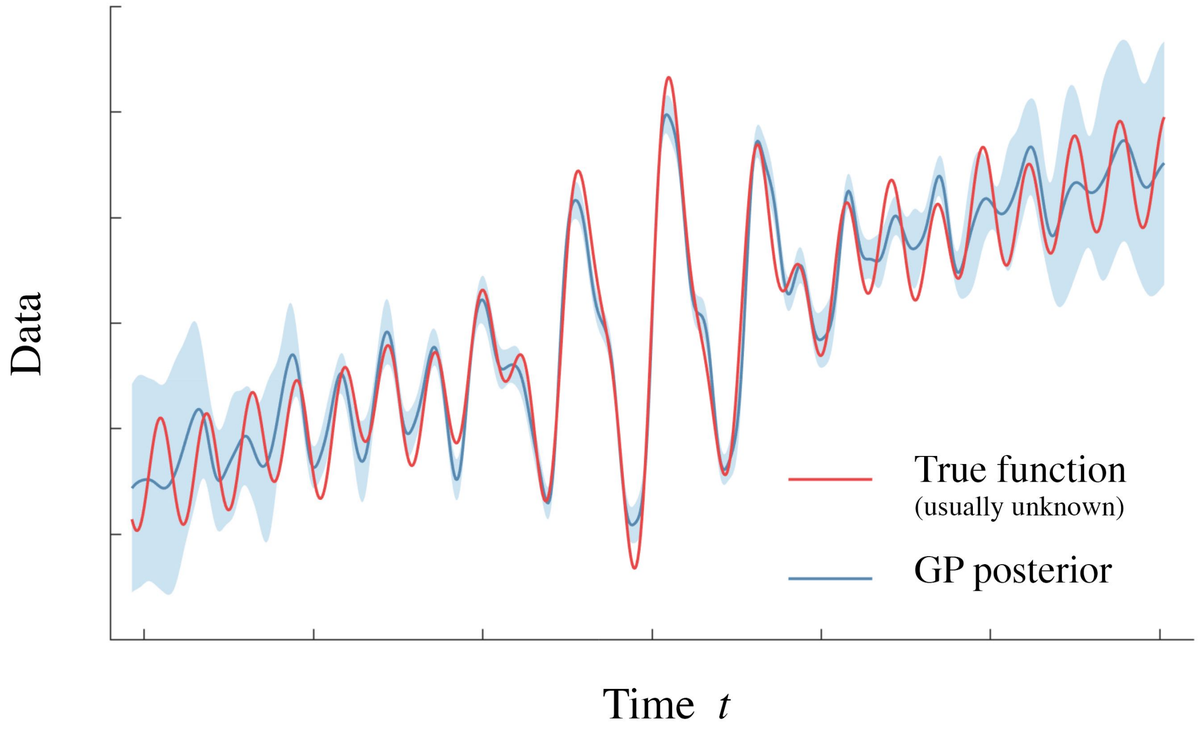
THE REMARKABLE PROPERTIES OF GAUSSIANS

- **Marginalisation property**: can compute marginals & conditionals for arbitrary subsets of variables
- **Conjugacy**: Gaussian prior + likelihood → posterior that is also a Gaussian
- In practice: **GP prior + data → GP posterior distribution** that can be evaluated **analytically** (via some linear algebra)
- So: GPs empower us to **learn unknown functions (+ error bars) from data**

INTRODUCING GPs

WHAT DOES THIS LOOK LIKE IN PRACTICE?





GP ACTIVITY MODELLING

QUASI-PERIODIC COVARIANCE KERNEL

$$k(\tau) = h^2 \exp\left(\frac{-\tau^2}{2\lambda_e^2}\right) \exp\left(\frac{-\sin^2(\pi\tau / P)}{2\lambda_p^2}\right),$$

characteristic period

active region coverage

evolution time scale

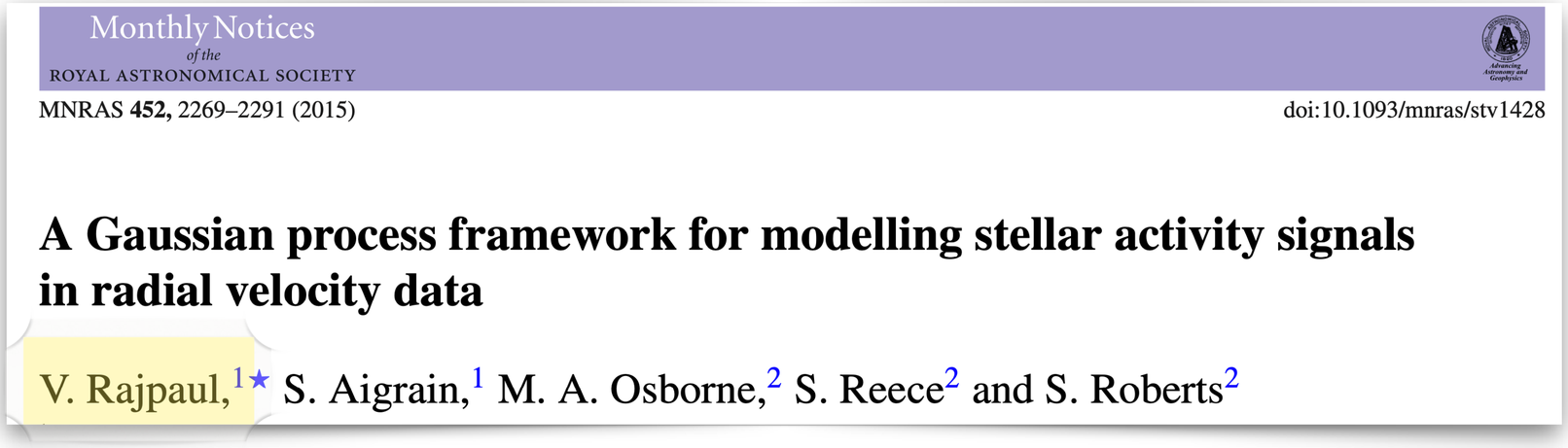
roughness/structure per period

where $\tau := t - t'$

time between any 2 points

GP ACTIVITY MODELLING

- Enable **lower-mass planets** to be detected around a **wider variety of stars**
- Further improvements via **multi-dimensional GP** modelling (Rajpaul+2015), even when $P_{\rm orb} = P_{\rm rot}$. Many significant exoplanets detected via this technique.


Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

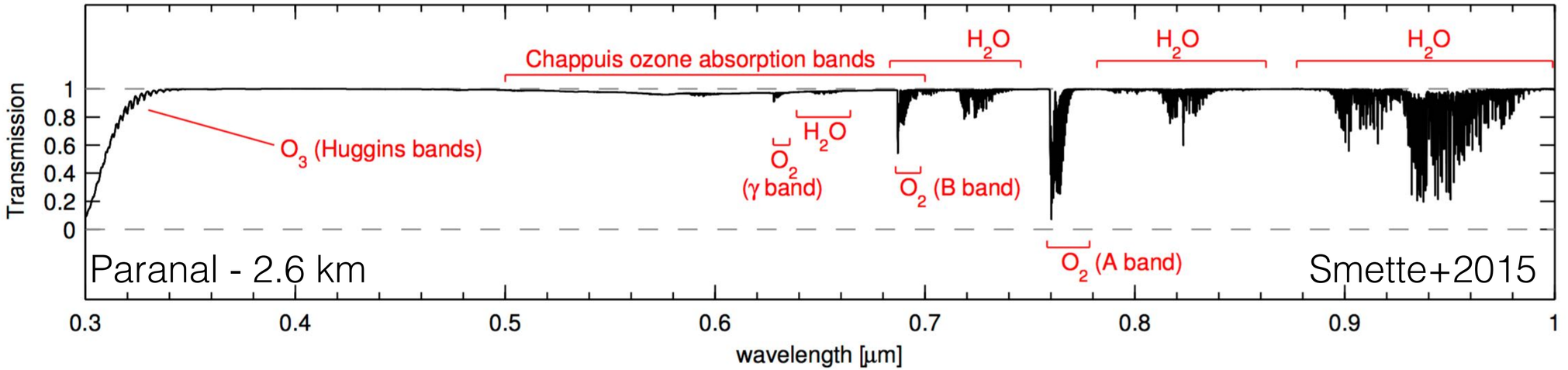
MNRAS **452,** 2269–2291 (2015) doi:10.1093/mnras/stv1428

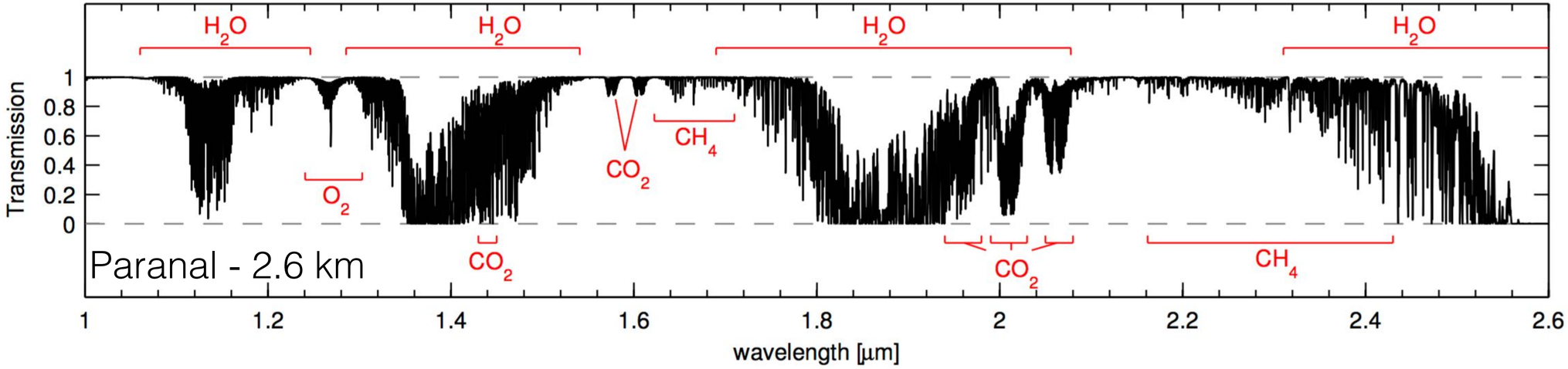
A Gaussian process framework for modelling stellar activity signals in radial velocity data

V. Rajpaul,[1]★ S. Aigrain,[1] M. A. Osborne,[2] S. Reece[2] and S. Roberts[2]

2. Non-astrophysical nuisance signals

TELLURIC ABSORPTION





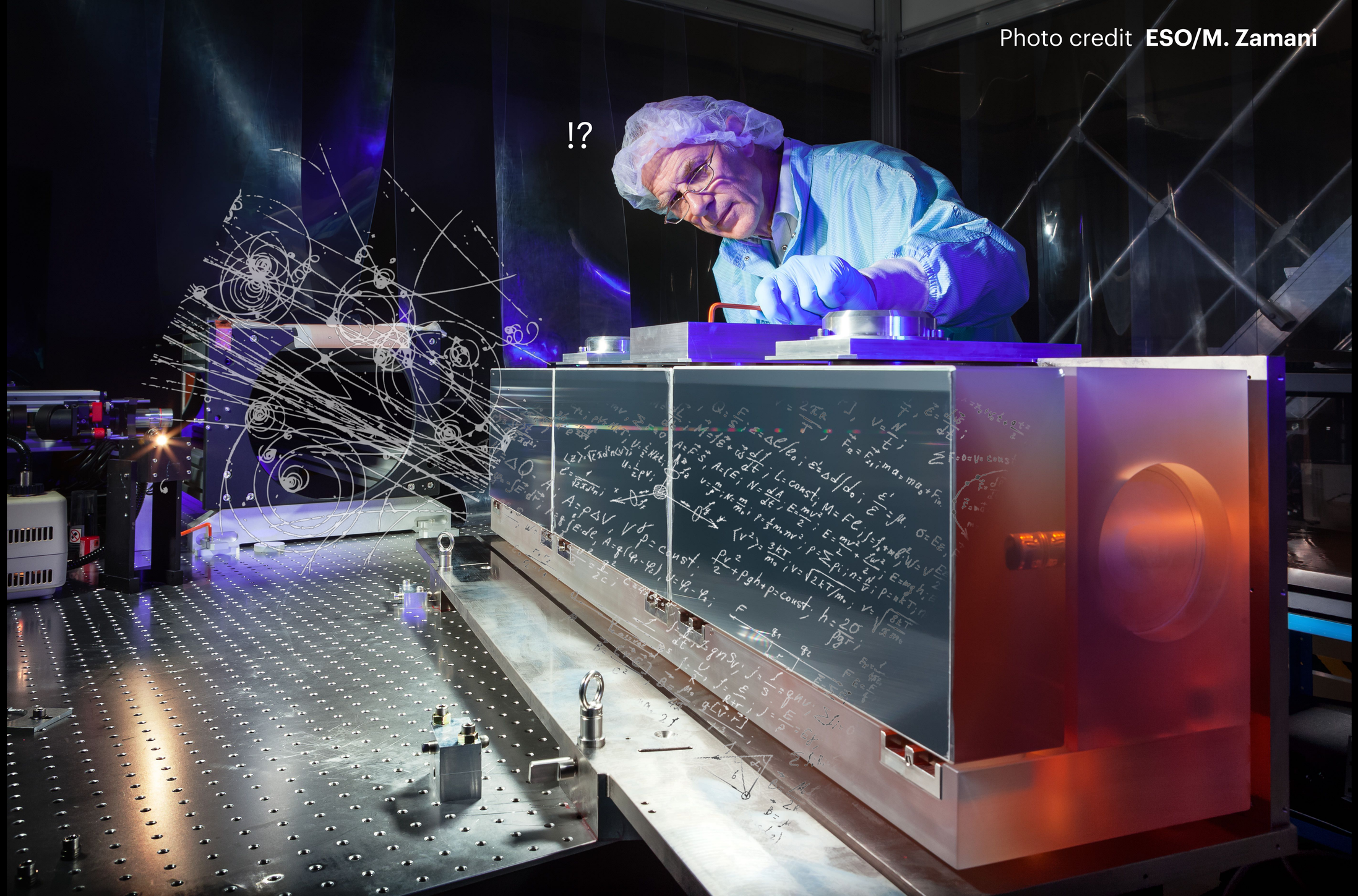
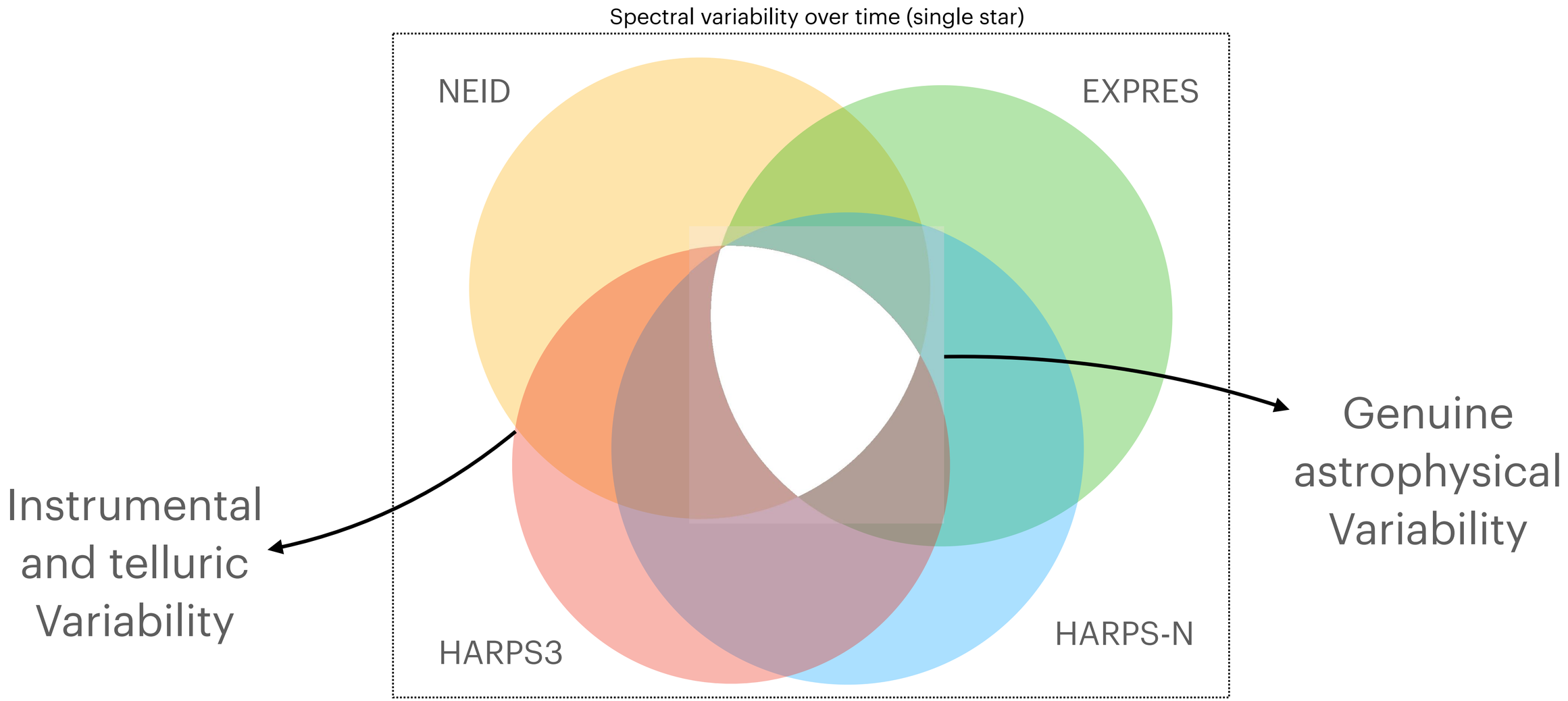


Photo credit **ESO/M. Zamani**

MULTI-SPECTROGRAPH RVs?



3. Non-trivial Bayesian inference

BAYESIAN CHALLENGES

"Computing the posterior distribution for [Keplerian orbital elements] is only possible with the latest numerical tools, and is **computationally expensive**...realistic data sets often have **millions of local likelihood maxima**"

– Hara & Ford, 2023
(*Annu. Rev. Stat. Appl.* 2023. 10:25.1–25.27)

C2D3 Cambridge Centre for Data-Driven Discovery

PolyChord/
PolyChordLite

Public version of PolyChord: See polychord.co.uk for PolyChordPro

farhanferoz/
MultiNest

MultiNest is a Bayesian inference tool which calculates the evidence and explores the parameter space which may contain multiple posterior...

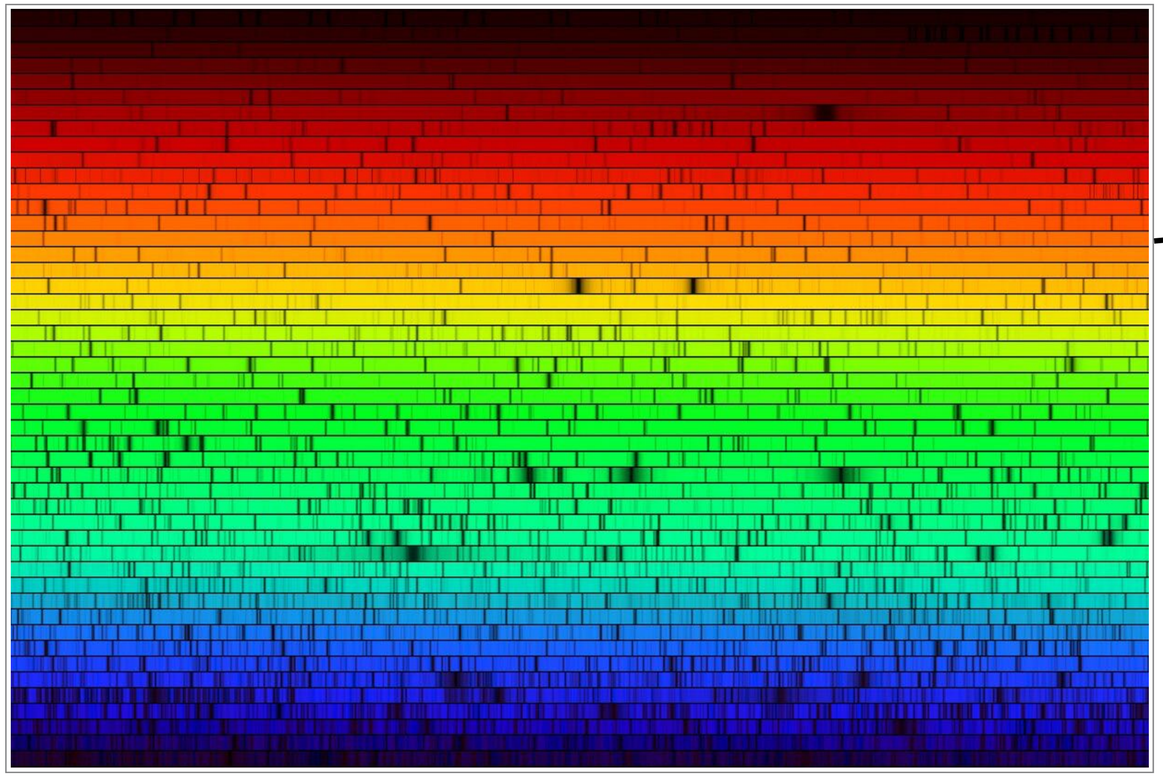
BAYESIAN CHALLENGES

Supercomputers and clever algorithms notwithstanding...

- Computing **Bayesian evidences** still ***very* challenging**
- Convenience often favoured over thoroughness
- Detections often **lack adequate statistical evidence**
- BMC alternatives: **cross-validation**, **Bayesian quadrature**, **evidence network**s ...?
- **Model averaging** rather than selection?

4. Unused spectral information

UNUSED SPECTRAL INFORMATION



RV + activity indicators

Decontaminate RVs using
indicator time series

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **452,** 2269–2291 (2015) doi:10.1093/mnras/stv1428

A Gaussian process framework for modelling stellar activity signals in radial velocity data

V. Rajpaul,[1]★ S. Aigrain,[1] M. A. Osborne,[2] S. Reece[2] and S. Roberts[2]

Time domain

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **507,** 1847–1868 (2021) https://doi.org/10.1093/mnras/stab2192
Advance Access publication 2021 July 30

A HARPS-N mass for the elusive Kepler-37d: a case study in disentangling stellar activity and planetary signals

V. M. Rajpaul,[1,2]★ L. A. Buchhave,[3] G. Lacedelli,[4,5] K. Rice,[6,7] A. Mortier,[1,8] L. Malavolta,[4,5] S. Aigrain,[2] L. Borsato,[5] A. W. Mayo,[9] D. Charbonneau,[10] M. Damasso,[11] X. Dumusque,[12] A. Ghedina,[13] D. W. Latham,[10] M. López-Morales,[10] A. Magazzù,[13] G. Micela,[14] E. Molinari,[15] F. Pepe,[12] G. Piotto,[4,5] E. Poretti,[13,16] S. Rowther,[17,18] A. Sozzetti,[11] S. Udry[12] and C. A. Watson[19]

Wavelength domain

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **492,** 3960–3983 (2020) doi:10.1093/mnras/stz3599
Advance Access publication 2020 January 3

A robust, template-free approach to precise radial velocity extraction

V. M. Rajpaul,[1]★ S. Aigrain[2] and L. A. Buchhave[3]

The Astronomical Journal, 163:171 (34pp), 2022 April https://doi.org/10.3847/1538-3881/ac5176

The EXPRES Stellar Signals Project II. State of the Field in Disentangling Photospheric Velocities

Lily L. Zhao[1,2,36], Debra A. Fischer[1], Eric B. Ford[3,4,5,6,37], Alex Wise[3,37], Michaël Cretignier[7,38], Suzanne Aigrain[8,39], Oscar Barragan[8,39], Megan Bedell[2,36], Lars A. Buchhave[9,39], João D. Camacho[10,11,40], Heather M. Cegla[12,13,41], Jessi Cisewski-Kehe[14], Andrew Collier Cameron[15,42], Zoe L. de Beurs[16,17,43], Sally Dodson-Robinson[18,19,44], Xavier Dumusque[7,38], João P. Faria[10,11,40], Christian Gilbertson[3,4,5,37], Charlotte Haley[20,44], Justin Harrell[18,44], David W. Hogg[2,21,22,23,36], Parker Holzer[24], Ancy Anna John[15,42], Baptiste Klein[8,39], Marina Lafarga[12,41], Florian Lienhard[25], Vinesh Maguire-Rajpaul[25,39], Annelies Mortier[25,26], Belinda Nicholson[8,39], Michael L. Palumbo, III[3,37], Victor Ramirez Delgado[18,44], Christopher J. Shallue[27,43], Andrew Vanderburg[28,29,43], Pedro T. P. Viana[10,40,30], Jinglin Zhao[3,37], Norbert Zicher[8,39], Samuel H. C. Cabot[1], Gregory W. Henry[31], Rachael M. Roettenbacher[1,32], John M. Brewer[33], Joe Llama[34], Ryan R. Petersburg[35], and Andrew E. Szymkowiak[1]

GPs, neural networks, PCA, linear regression, cross-correlation, line-by-line methods, etc.

RVs often discrepant at ~ 1 m s^{-1} level (!)

5. Miscellaneous

NON-GAUSSIANITY & OUTLIERS

"we used an **additive [white Gaussian noise] jitter** term, added in quadrature to the formal error bars, to account for all **additional sources of uncertainty**…"

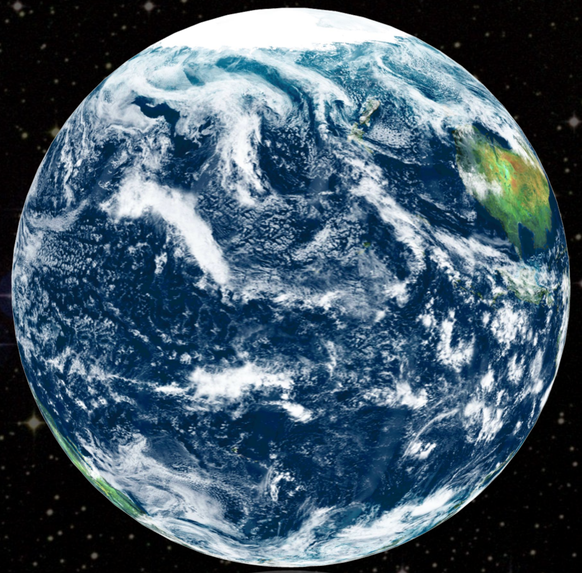
— *countless exoplanet papers*

MODEL MIS-SPECIFICATION

"Essentially, **all models are wrong, but some are useful**. However, the approximate nature of the model must always be borne in mind."

– George Box (statistician)

When we announce the
discovery of the first "Earth twin" —



we ought to be **certain**
it's not a false positive


THE
TERRA HUNTING EXPERIMENT