

TECHNICAL COMMENT

EXOPLANET DETECTION

Comment on “Stellar activity masquerading as planets in the habitable zone of the M dwarf Gliese 581”

Guillem Anglada-Escudé^{1,2*} and Mikko Tuomi²

Robertson *et al.* (Reports, 25 July 2014, p. 440) claimed that activity-induced variability is responsible for the Doppler signal of the proposed planet candidate GJ 581d. We point out that their analysis using periodograms of residual data is inappropriate and promotes inadequate tools. Because the claim challenges the viability of the method to detect exo-Earths, we encourage reanalysis and a deliberation on what the field-standard methods should be.

GJ 581d was the first planet candidate of a few Earth masses reported in the circumstellar habitable zone of another star (1). It was detected by measuring the radial velocity variability of its host star using High Accuracy Radial Velocity Planet Searcher (HARPS) (1, 2). Doppler time series are usually modeled as the sum of Keplerian signals plus additional effects (e.g., correlations with activity). Detecting a planet candidate consists of quantifying the improvement of a merit statistic when one signal is added to the model. Approximate methods are often used to speed up the analyses, such as computing periodograms on residual data. Even when models are linear, correlations exist between parameters. Similarly, statistics based on residual analyses are biased quantities and cannot be used for model comparison.

A golden rule in data analysis is that the data should not be corrected but that the model is what needs improvement. The inadequacy of residual analyses can be illustrated using a simple example (Fig. 1). Assume 16 measurements of the position (x) of an object as a function of time (t) and no uncertainties. We are interested in its velocity and must decide whether a constant offset x_0 is needed to model the motion. Model A (null hypothesis) consists of $x_A(t) = vt$, where v is the only free parameter, and the alternative model B is $x_B(t) = x_0 + vt$. The question is whether including x_0 is justified, given the improvement of a statistic that we define as $\chi^2 = \frac{1}{N} \sum_{i=1}^N (x_i - x(t_i))^2$. The left panel in Fig. 1 illustrates a flawed procedure that consists of adjusting Model A and then deciding whether a constant x_0 is required to explain the residuals (bottom left panel).

Because such residuals are far from a constant shift, the reduction of χ^2 is not maximal, and the fit to a constant offset is unsatisfactory. By subtracting model A from the data, we have created a new time series that is no longer representative of the original one. A more meaningful procedure consists of comparing model A to the global fit to all the parameters of model B (top right panel) to achieve the maximum improvement of our statistic.

Similarly, the analysis in Robertson *et al.* (3) only shows that the signal of GJ 581d is not present

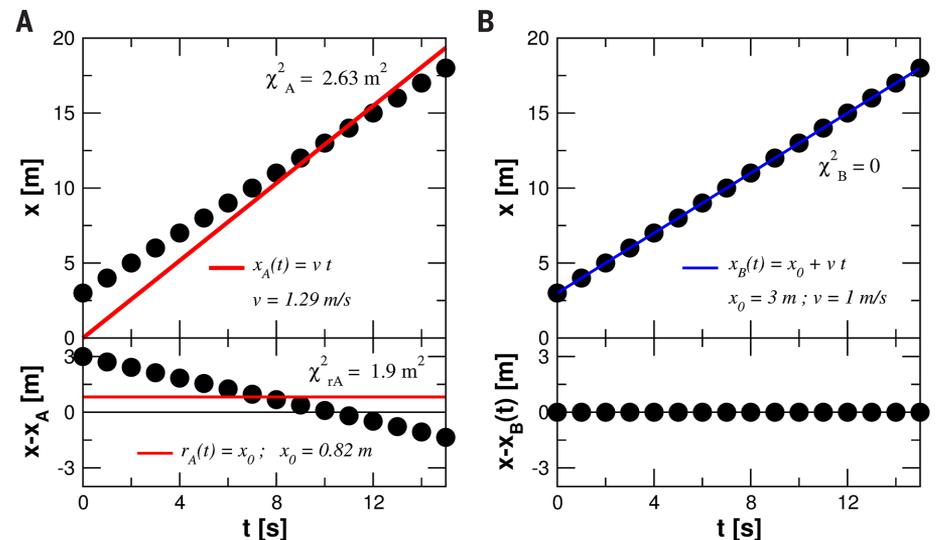


Fig. 1. This example illustrates why residual statistics must not be used to assess significance in multiparametric fits to data. We want to know if a constant x_0 is needed to model the position x of a body as a function of time t (black dots). (A) (Top) represents the fit to the null hypothesis (model A), which only includes a velocity term (red line). (A) (Bottom) shows the residuals $r_A = x - x_A(t)$ to model A. An attempt to fit a model with a constant offset to r_A is shown as a red line. (B) (Top) represents a fit adjusting all the parameters simultaneously (model B, top right panel, blue line), producing the largest possible reduction in χ^2 .

in their new residual time series. Their procedure is summarized as follows: Figure 1 and figure S3 in (3) were used to suggest significant correlations between Doppler and stellar activity measurements (chromospheric emission from the H α line). After subtracting those correlations and the first three planets, periodograms (4–6) were applied to the residuals to show that GJ 581d fell below the detectability threshold. Although the semi-amplitude K of the signal of GJ 581d is large (~ 1.6 m/s), the apparent variability induced by the RV/H α correlations is 5 m/s peak to peak, and the scatter around the fits is at the 1.5 to 2 m/s level. Subtracting those correlations biased the residuals by removing a model that likely included contributions from real signals, and additional noise was added due to the scatter in the RV/H α relations. All things considered, the disappearance of GJ581d in such residual data is not surprising. Following Fig. 1, a simultaneous fit of the 30+ parameters involved would be needed to reach meaningful conclusions. Although there may be substantial RV/H α correlations, a global optimization analysis may not support that GJ 581d is better explained by activity. A complete analysis will be presented elsewhere.

We argue that the results of Robertson *et al.* come from the improper use of periodograms on residual data, because they implement the same flawed procedure illustrated in Fig. 1 in this Comment. Despite the utility of periodograms for providing quick-look analyses, their inadequacy to the task has been abundantly discussed in the literature (7–12). Explicitly, derived false-alarm probabilities would be representative only if a model with one sinusoid and one offset is a sufficient description of the data, measurements are uncorrelated, noise is normally distributed, and

¹School of Physics and Astronomy, Queen Mary University of London, London E1 4NS, UK. ²Centre for Astrophysics Research, University of Hertfordshire, AL10 9AB, Hatfield, UK. *Corresponding author. E-mail: guillem.anglada@gmail.com

uncertainties are fully characterized (5). All of these hypotheses break down when dealing with Doppler residuals: The number of signals in not known a priori, fits to data correlate with residuals, and formal uncertainties are not realistic. Proposed alternatives, such as Monte Carlo bootstrapping of periodograms (5), do not help either, because those methods also ignore correlations. Resulting biases can lead to significance assessments off by several orders of magnitude. These issues were irrelevant when Doppler amplitudes abundantly exceeded uncertainties. For example, an amplitude larger than three times the uncertainties and more than 20 measurements easily leads to false-alarm probabilities smaller than 10^{-6} , which is much smaller than usual thresholds at 1 to 0.1%. For this reason, large biases were not problematic in the early detection of gas giants ($K \sim 50$ m/s and $\sigma \sim 5$ m/s) (13), and it is the main reason that periodograms of residual data are still widespread tools in Doppler analyses despite their inadequacy for the task.

In summary, analysis of statistical significance using residual data statistics leads to incorrect assessments. Although this has been a common practice in the past, the problem is now exacerbated

with signals closer to the noise and increased model complexity. The properties of the noise can be included in the model but can never be subtracted from the data. This discussion directly affects the viability of the Doppler method to find Earth-like planets. Although Earth causes a 0.1 m/s wobble around the Sun, the long-term stability of the most quiet stars is not better than 0.8 m/s (2). That is, activity-induced variability can be 5 to 10 times as large as the signal. Although global optimization does not provide an absolute guarantee of success, analyses based on residual statistics are bound to fail. If activity poses an ultimate barrier to the detection of small planets, strategic long-term plans concerning large projects will need serious revision (14). It is thus of capital importance that analysis and verification of multiplanet claims are properly done using global-optimization techniques and by acquiring additional observations.

REFERENCES AND NOTES

1. M. Mayor *et al.*, *Astron. Astrophys.* **507**, 487–494 (2009).
2. F. Pepe *et al.*, *Astron. Astrophys.* **534**, A58 (2011).
3. P. Robertson, S. Mahadevan, M. Endl, A. Roy, *Science* **345**, 440–444 (2014).
4. N. R. Lomb, *Astrophys. Space Sci.* **39**, 447 (1976).
5. A. Cumming, *Mon. Not. R. Astron. Soc.* **354**, 1165–1176 (2004).
6. M. Zechmeister, M. Kürster, *Astron. Astrophys.* **496**, 577–584 (2009).
7. P. C. Gregory, *Astrophys. J.* **631**, 1198G (2005).
8. R. P. Baluev, *Mon. Not. R. Astron. Soc.* **385**, 1279–1285 (2008).
9. M. Tuomi, *Astron. Astrophys.* **528**, L5 (2011).
10. G. Anglada-Escudé, M. Tuomi, *Astron. Astrophys.* **548**, A58 (2012).
11. R. D. Haywood *et al.*, *Mon. Not. R. Astron. Soc.* **443**, 2517–2531 (2014).
12. R. Baluev, *Mon. Not. R. Astron. Soc.* **393**, 969–978 (2009).
13. M. Mayor, D. Queloz, *Nature* **378**, 355–359 (1995).
14. Several instruments aimed at precision better than 1 m/s are being proposed and/or in construction (e.g., ESPRESSO/VLT at the European Southern Observatory, TPF at the Hobby-Eberly Telescope CARMENES at the Calar Alto Observatory, and SPIRou at the Canadian France Hawaii Telescope), because they are considered essential to detect any Earth-like planets or confirm/characterize those detected by next planet-hunting space missions (K2/NASA, TESS/NASA and PLATO/ESA).

ACKNOWLEDGMENTS

This work has been mostly supported by The Leverhulme Trust through grant RPG 2014-281 – PAN-Disciplinary algORithms for data Analysis. We thank H. R. A. Jones and R. P. Nelson for useful discussions and support.

3 September 2014; accepted 16 January 2015
10.1126/science.1260796

Comment on "Stellar activity masquerading as planets in the habitable zone of the M dwarf Gliese 581"

Guillem Anglada-Escudé and Mikko Tuomi

Science **347** (6226), 1080.
DOI: 10.1126/science.1260796

ARTICLE TOOLS

<http://science.sciencemag.org/content/347/6226/1080.2>

RELATED CONTENT

<http://science.sciencemag.org/content/sci/345/6195/440.full>
<http://science.sciencemag.org/content/sci/347/6226/1080.3.full>

REFERENCES

This article cites 12 articles, 1 of which you can access for free
<http://science.sciencemag.org/content/347/6226/1080.2#BIBL>

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)