

## Are there Spectral Features in the MIRI/LRS Transmission Spectrum of K2-18b?

JAKE TAYLOR <sup>1</sup>

<sup>1</sup>*Department of Physics, University of Oxford, Parks Rd, Oxford, OX1 3PU, UK*

### ABSTRACT

Determining the composition of an exoplanet atmosphere relies on the presence of detectable spectral features. The strongest spectral features, including DMS, look approximately Gaussian. Here, I perform a suite of Gaussian feature analyses to find any statistically significant spectral features in the recently published MIRI/LRS spectrum of K2-18b (N. Madhusudhan et al. 2025). In N. Madhusudhan et al. (2025), they claim a 3.4- $\sigma$  detection of spectral features compared to a flat line. In 5 out of 6 tests, I find the data preferred a flat line over a Gaussian model, with a  $\chi^2_{\nu}$  of 1.06. When centering the Gaussian where the absorptions for DMS and DMDS peak, I find  $\ln(B) = 1.21$  in favour of the Gaussian model, with a  $\chi^2_{\nu}$  of 0.99. With only  $\sim 2\text{-}\sigma$  in favour of Gaussian features, I conclude no strong statistical evidence for spectral features.

### 1. INTRODUCTION

The search for signs of life is a key goal of the exoplanet community. The James Webb Space Telescope (JWST) has improved our ability to study the atmospheres of temperate worlds. Recent theoretical and observational work have determined H<sub>2</sub>-rich sub-Neptunes could be our current best opportunity to find life elsewhere in the universe (S. Seager et al. 2025). One hypothesised class of sub-Neptune is “Hycean” worlds (N. Madhusudhan et al. 2025), defined by a H<sub>2</sub> rich atmosphere above a global liquid water ocean.

The planet K2-18b ( $T_{\text{eq}} = 254.9 \pm 3.9$  K,  $R_p = 2.610 \pm 0.087 R_{\oplus}$ ,  $M_p = 8.63 \pm 1.35 M_{\oplus}$  (R. Cloutier et al. 2019; B. Benneke et al. 2019)) has been extensively studied, with competing hypotheses on whether it is a Hycean world (N. Madhusudhan et al. 2025), mini-Neptune (N. F. Wogan et al. 2024), or a gas dwarf with a surface magma ocean (O. Shorttle et al. 2024). K2-18b has now been observed in transmission using 3 instrument modes of JWST: NIRISS/SOSS, NIR-Spec/G395H, and MIRI/LRS. The initial near-infrared observations suggested the presence of dimethyl sulphide (DMS) (N. Madhusudhan et al. 2023), but this molecule was not confirmed in a re-analysis of the same data (S. P. Schmidt et al. 2025). Recently, mid-infrared observations of K2-18b suggested the presence of DMS and/or dimethyl disulphide (DMDS) at 3- $\sigma$  significance (N. Madhusudhan et al. 2025). The potential presence of DMS or DMDS has been raised as evidence for a bio-

sphere on K2-18b, given that these molecules are produced by phytoplankton in Earth’s oceans.

N. Madhusudhan et al. (2025) claims to reject a flat line at 3.4- $\sigma$  significance. They compare a 16-parameter “canonical” model to a 1-parameter flat line model. However, this is not a nested analysis, as their flat line model is not a subset of their “canonical” model. It is possible to compare the Bayesian evidence of these models to determine whether one is favoured over the other, however, careful consideration of the prior ranges need to be taken, as the prior space can greatly impact the Bayesian evidence (R. Trotta 2008). The authors do observe this behaviour; they comment that when increasing the prior range for their flat line model, they achieve a stronger significance. When using their 32-parameter “maximal” model<sup>2</sup>, they find they can only reject a flat line at 3- $\sigma$ , the significance has gone down because the prior space between the “maximal” and “canonical” models are different.

The 3- $\sigma$  detection of DMS/DMDS claimed by N. Madhusudhan et al. (2025) was established using a curated “canonical” model as opposed to their “maximal” model. However, the “canonical” model neglects opacity contributions from other potential molecules, which consequently boosts the detection significance by neglecting alternative species in the Bayesian evidence. They performed the same tests with their “maximal” model and find they can only reject a flat line model compared to

<sup>2</sup> These models are being used to fit 29 data points. One should generally avoid fitting data with more free parameters than data points to avoid overfitting the data.

their atmosphere models without DMS/DMDS at less than the  $2\text{-}\sigma$  level. Furthermore, the authors postulate that the abundance of DMS/DMDS they obtain are consistent with  $\sim 20\times$  Earth’s biogenic flux, as predicted by [S.-M. Tsai et al. \(2024\)](#). However, in [S.-M. Tsai et al. \(2024\)](#) they show that ethane ( $\text{C}_2\text{H}_6$ ) would be more readily detectable with this level of biogenic flux compared to DMS, yet no ethane is seen in the MIRI/LRS observations. These concerns motivate a careful reassessment of the reliability of the claimed detection of DMS and/or DMDS in K2-18b’s atmosphere.

In this research note, I explore whether a flat line can be rejected by the K2-18b MIRI transmission spectrum, and therefore determine if an atmospheric signal can be claimed.

## 2. METHODS

When using the Bayesian evidence to compute whether one model is favoured by a set of data compared to another model (e.g. an atmospheric model compared to a flat line), one model should ideally be a subset (‘nested’) of the other model with the same priors on the common parameters. In this section, I perform a nested approach to determine if a Gaussian model is statistically preferred relative to a flat line.

Following previous work (e.g. [E. M. May et al. 2023](#)), I search for spectral features, while remaining agnostic to the particular gas causing the feature, with a Gaussian function:

$$\delta_\lambda(A, \mu, \sigma_m, c) = A \exp\left(-\frac{(\lambda - \mu)^2}{2\sigma_m^2}\right) + c \quad (1)$$

where  $\delta_\lambda$  is the model transit depth, and  $\mu$ ,  $\sigma_m$ , and  $A$  are the centre, standard deviation, and amplitude of the Gaussian, respectively, and  $c$  is a vertical offset. Given that the spectral shapes of DMS and DMDS are Gaussian-like (see Figure 3 of [N. Madhusudhan et al. \(2025\)](#)), this is a reasonable approximation. I fit for the flat line using the following form:

$$\delta_\lambda(c) = c. \quad (2)$$

For the analysis, I only consider the observations obtained with the *JexoRes* pipeline, which is the same data shown in Figure 2 of [N. Madhusudhan et al. \(2025\)](#). I performed the same analysis with the *JexoPipe* data, finding a similar conclusion.<sup>3</sup>

I perform 3 different Gaussian feature tests: 1) the standard form (4 free parameters), 2) a negative Gaussian (4 free parameters), 3) a combination of the two

(7 free parameters). For each of these tests, I perform an additional analysis where I make informed choices of where the spectral features should be (motivated by Figure 6 of [N. Madhusudhan et al. \(2025\)](#), which shows that features at  $\sim 7\ \mu\text{m}$  and  $\sim 8.8\ \mu\text{m}$  drive the claimed detections), which therefore fixes the  $\mu$  parameter. I plot the results in Figure 1 with the corresponding  $\chi^2$ , degrees of freedom, reduced  $\chi^2$ , and the  $\ln$  Bayesian evidence obtained using *MultiNest* ([F. Feroz et al. 2009](#)).

## 3. RESULTS

I find that all the tested models have similar  $\ln$ -Bayesian evidences, indicating that a Gaussian model is not statistically preferred over a flat line. Since the difference in the  $\ln$ -Evidence of a Gaussian model and the flat line model is the  $\ln$  Bayes Factor ( $\ln(\text{B})$ ), one can assess which model can best describe the observations using the Jeffrey’s scale<sup>4</sup>, which defines ‘no evidence’ ( $\ln(\text{B}) < 1$ ), ‘weak evidence’ ( $1 \leq \ln(\text{B}) < 2.5$ ), ‘moderate evidence’ ( $2.5 \leq \ln(\text{B}) < 5$ ), and ‘strong evidence’ ( $\ln(\text{B}) \geq 5$ ). Ultimately, 5 out of 6 of the Gaussian model tests fall into the ‘no evidence’ category according to the Jeffrey’s scale, with the models with the best  $\ln$ -Evidences having a Gaussian fixed to specific wavelengths. The only model marginally preferred over a flat line is the combination of a positive and negative Gaussian fixed at  $7\ \mu\text{m}$  and  $8.8\ \mu\text{m}$ , respectively, with  $\ln(\text{B}) = 1.21$ . However, even in this case, the Gaussians are only in the ‘weak evidence’ category. When using ‘equivalent sigma’ values (e.g. [R. Trotta 2008](#)), as commonly done in the exoplanet retrieval literature (and in [N. Madhusudhan et al. \(2025\)](#)), this would be a ‘ $2\text{-}\sigma$  detection’ compared to the flat line model, even though there is little statistical evidence favouring the two fixed Gaussians over a flat line.

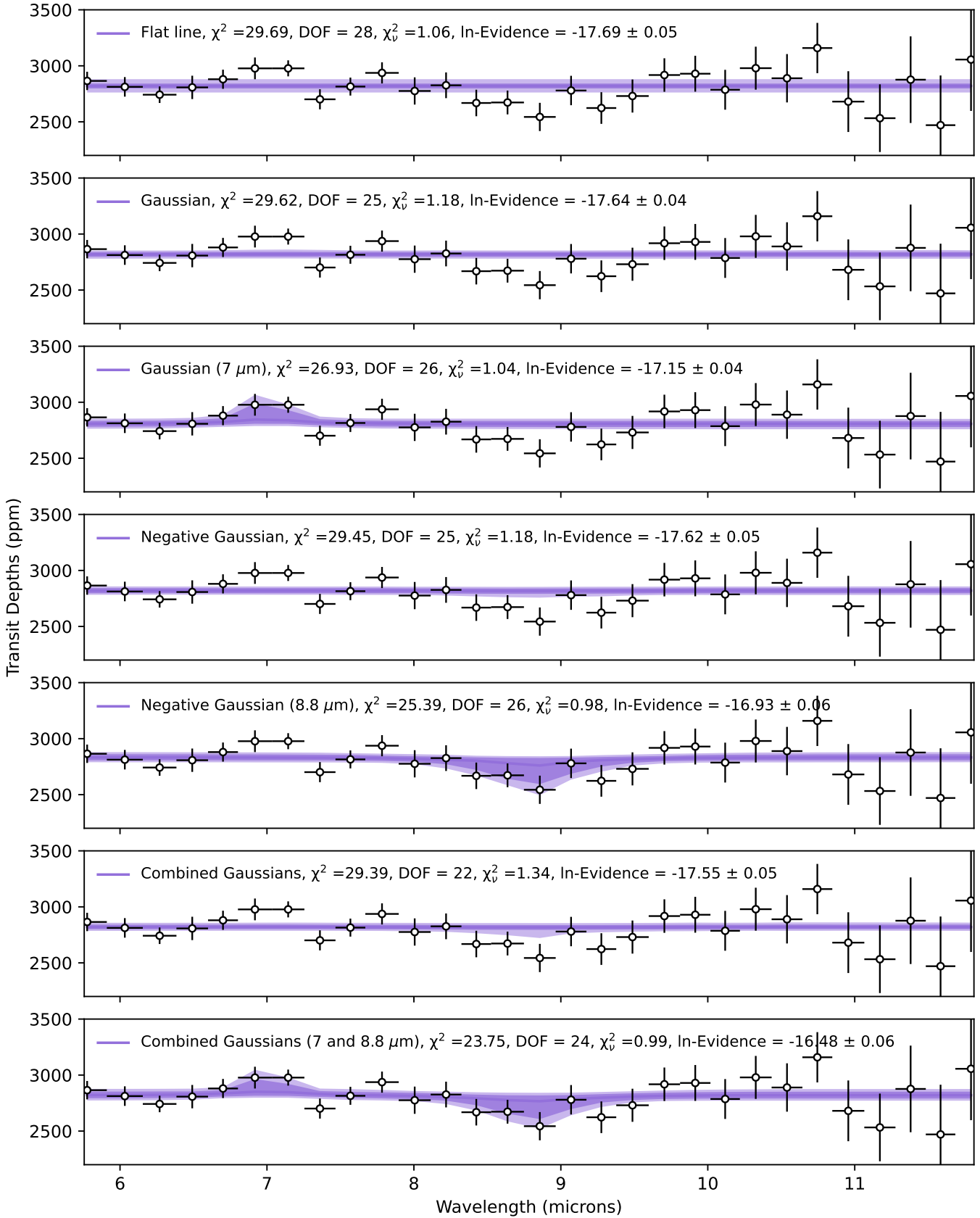
Further, when inspecting the fitted models in Figure 1, one can see that the models that agnostically search for Gaussian features result in flat spectra. Only when fixing the Gaussians to specific wavelengths do I see any deviation from a flat line. This test suggests that only considering a limited set of absorbers at fixed wavelengths, as done in [N. Madhusudhan et al. \(2025\)](#)’s ‘canonical’ model, inflates the statistical evidence for an atmospheric detection.

## 4. CONCLUSIONS

I perform an independent analysis of the MIRI/LRS observations of K2-18b recently published by [N. Madhusudhan et al. \(2025\)](#). The specific question of this research note was to determine if there were any de-

<sup>3</sup> Data obtained from [N. Madhusudhan \(2025\)](#).

<sup>4</sup> Table 1 and 2 in [R. Trotta \(2008\)](#)



**Figure 1.** Model fits to the MIRI/LRS observations of K2-18b from [N. Madhusudhan et al. \(2025\)](#). The median, 1- $\sigma$  and 2- $\sigma$  credible region envelopes for each model fit are overlaid. The  $\chi^2$ , degrees of freedom, reduced  $\chi^2$ , and the ln Bayesian evidence values are annotated.

tectable features within the data itself, or whether the data is consistent with a flat line. I performed a Gaussian feature analysis with varying levels of complexity, as opposed to a complete atmospheric retrieval. Despite the claim of a  $3.4\text{-}\sigma$  deviation from a flat line in [N. Madhusudhan et al. \(2025\)](#), I find a flat line is an acceptable fit.

Of the tested models, only one, where I use two Gaussians centred on fixed wavelengths, results in weak evidence over a flat line ( $\ln(B) = 1.21$ , Bayes factor = 3.35). Therefore, there is no strong evidence for detected spectral features in K2-18b’s MIRI transmission spectrum.

## ACKNOWLEDGMENTS

**Thanks to Suzanne Aigrain, Louis-Philippe Coulombe, Mark Hammond, Ryan MacDonald, Peter McGill, Ben Sutcliffe, and Luis Welbanks for their feedback, which improved the clarity of this manuscript.** J.T. was supported by the Glasstone Benefaction, University of Oxford.

## REFERENCES

- Benneke, B., Wong, I., Piaulet, C., et al. 2019, *ApJL*, 887, L14, doi: [10.3847/2041-8213/ab59dc](https://doi.org/10.3847/2041-8213/ab59dc)
- Cloutier, R., Astudillo-Defru, N., Doyon, R., et al. 2019, *A&A*, 621, A49, doi: [10.1051/0004-6361/201833995](https://doi.org/10.1051/0004-6361/201833995)
- Feroz, F., Hobson, M. P., & Bridges, M. 2009, *MNRAS*, 398, 1601, doi: [10.1111/j.1365-2966.2009.14548.x](https://doi.org/10.1111/j.1365-2966.2009.14548.x)
- Madhusudhan, N. 2025, OSF. [osf.io/gmhw3](https://osf.io/gmhw3)
- Madhusudhan, N., Constantinou, S., Holmberg, M., et al. 2025, arXiv e-prints, arXiv:2504.12267, doi: [10.48550/arXiv.2504.12267](https://doi.org/10.48550/arXiv.2504.12267)
- Madhusudhan, N., Sarkar, S., Constantinou, S., et al. 2023, *ApJL*, 956, L13, doi: [10.3847/2041-8213/acf577](https://doi.org/10.3847/2041-8213/acf577)
- May, E. M., MacDonald, R. J., Bennett, K. A., et al. 2023, *ApJL*, 959, L9, doi: [10.3847/2041-8213/ad054f](https://doi.org/10.3847/2041-8213/ad054f)
- Schmidt, S. P., MacDonald, R. J., Tsai, S.-M., et al. 2025, arXiv e-prints, arXiv:2501.18477, doi: [10.48550/arXiv.2501.18477](https://doi.org/10.48550/arXiv.2501.18477)
- Seager, S., Welbanks, L., Ellerbroek, L., Bains, W., & Petkowski, J. J. 2025, arXiv e-prints, arXiv:2504.12946, doi: [10.48550/arXiv.2504.12946](https://doi.org/10.48550/arXiv.2504.12946)
- Shorttle, O., Jordan, S., Nicholls, H., Lichtenberg, T., & Bower, D. J. 2024, *ApJL*, 962, L8, doi: [10.3847/2041-8213/ad206e](https://doi.org/10.3847/2041-8213/ad206e)
- Trotta, R. 2008, *Contemporary Physics*, 49, 71, doi: [10.1080/00107510802066753](https://doi.org/10.1080/00107510802066753)
- Tsai, S.-M., Innes, H., Wogan, N. F., & Schwieterman, E. W. 2024, *ApJL*, 966, L24, doi: [10.3847/2041-8213/ad3801](https://doi.org/10.3847/2041-8213/ad3801)
- Wogan, N. F., Batalha, N. E., Zahnle, K. J., et al. 2024, *ApJL*, 963, L7, doi: [10.3847/2041-8213/ad2616](https://doi.org/10.3847/2041-8213/ad2616)