Do you want to detect exoplanet atmospheric signals?
Don't forget to clean up your spectra!
Check out the best detrending method for you!

Marcelo Aron Fetzner Keniger*, Matteo Brogi, David Armstrong, Siddharth Gandhi
*E-mail: **Marcelo-Aron.Fetzner-Keniger@warwick.ac.uk**

UNIVERSITY OF WARWICK

Visit my webpage and try Astroclimes for yourself!

What is the problem?

Signals from exoplanet atmospheres are much weaker than the telluric and the host star's spectral lines.

What is the solution?

Removing the telluric and stellar lines before searching for the planetary signal, a process often called "detrending".

And how do we do that?

There are many ways proposed in the literature, for example:

1. Directly modelling telluric lines using algorithms such as Molecfit [5] and Astroclimes [3]. The stellar lines are removed with a master stellar template defined by the nightly average of the (telluric removed) observational spectra (see Figure 1).
2. Blind detrending using Principal Component Analysis (PCA) [4]. PCA finds common modes over time for each pixel in the spectral time series, so the (roughly) stationary telluric and stellar lines are removed, while the (moving) planetary signal is not. Which common modes or features are removed in the spectra is determined by the number of PCA components (N_C).

Step-by-step detrending process with Astroclimes

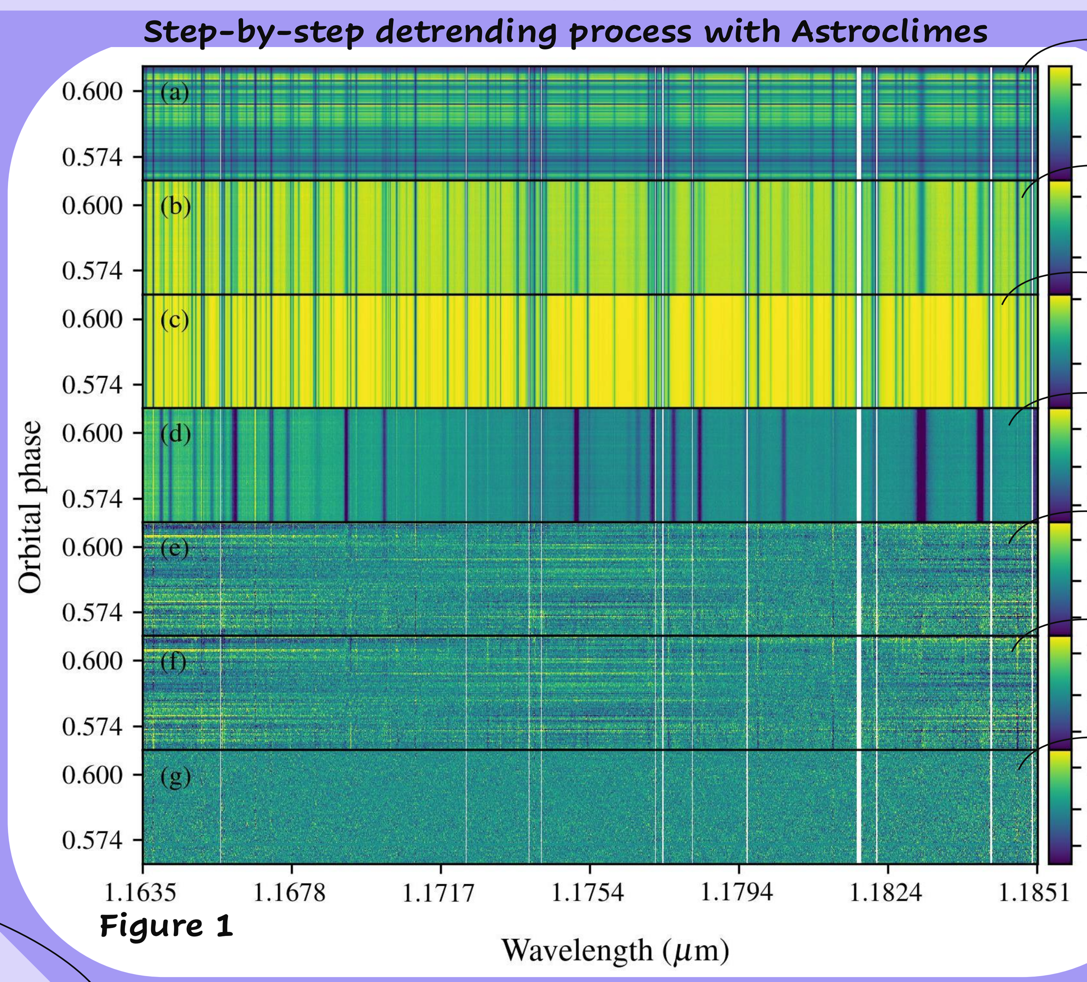

Figure 1

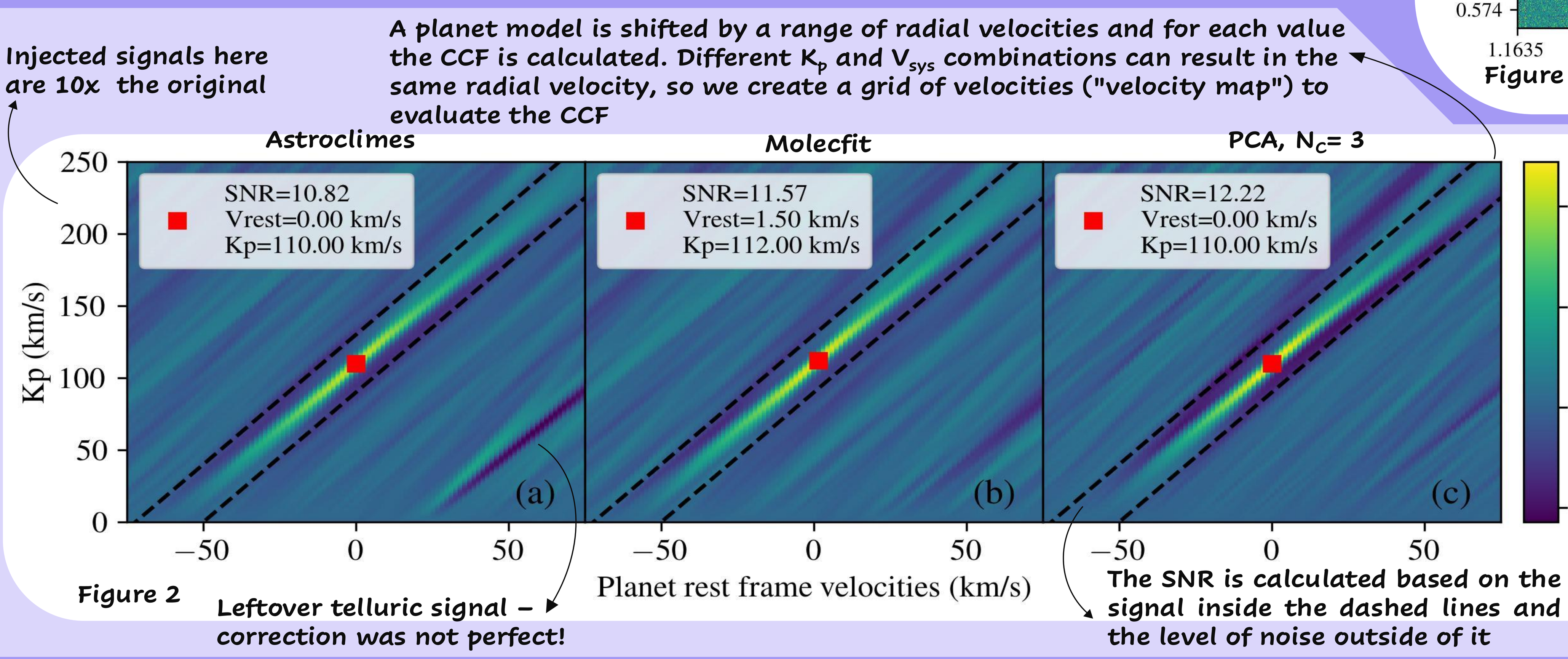

Figure 2

Then, how is the signal detected?

Even after detrending, the signal is still buried under the noise! High resolution cross-correlation spectroscopy (HRCCS) [1] can be used to combine the signal from all resolved spectral lines originating from the planet's atmosphere, thus strengthening the detection. From the cross-correlation function (CCF), a signal-to-noise ratio (SNR) can be calculated (Figure 2). To claim a detection, one typically needs an SNR of at least 4!

But what is the catch?

Detrending methods often affect the planet signal being studied. In the case of Astroclimes and Molecfit, part of the signal is removed when dividing by the master stellar template. For PCA, we cannot control what features are modelled by each component, so choosing too few components may result in a poor correction, but choosing too many may remove part of the planetary signal as well. Figure 3 shows how the injected signal changes after detrending with Astroclimes (panel b), Molecfit (panel c) and PCA (panel d).

The difference between the residuals after detrending with and without an injected planetary signal should yield precisely the injected planetary signal!

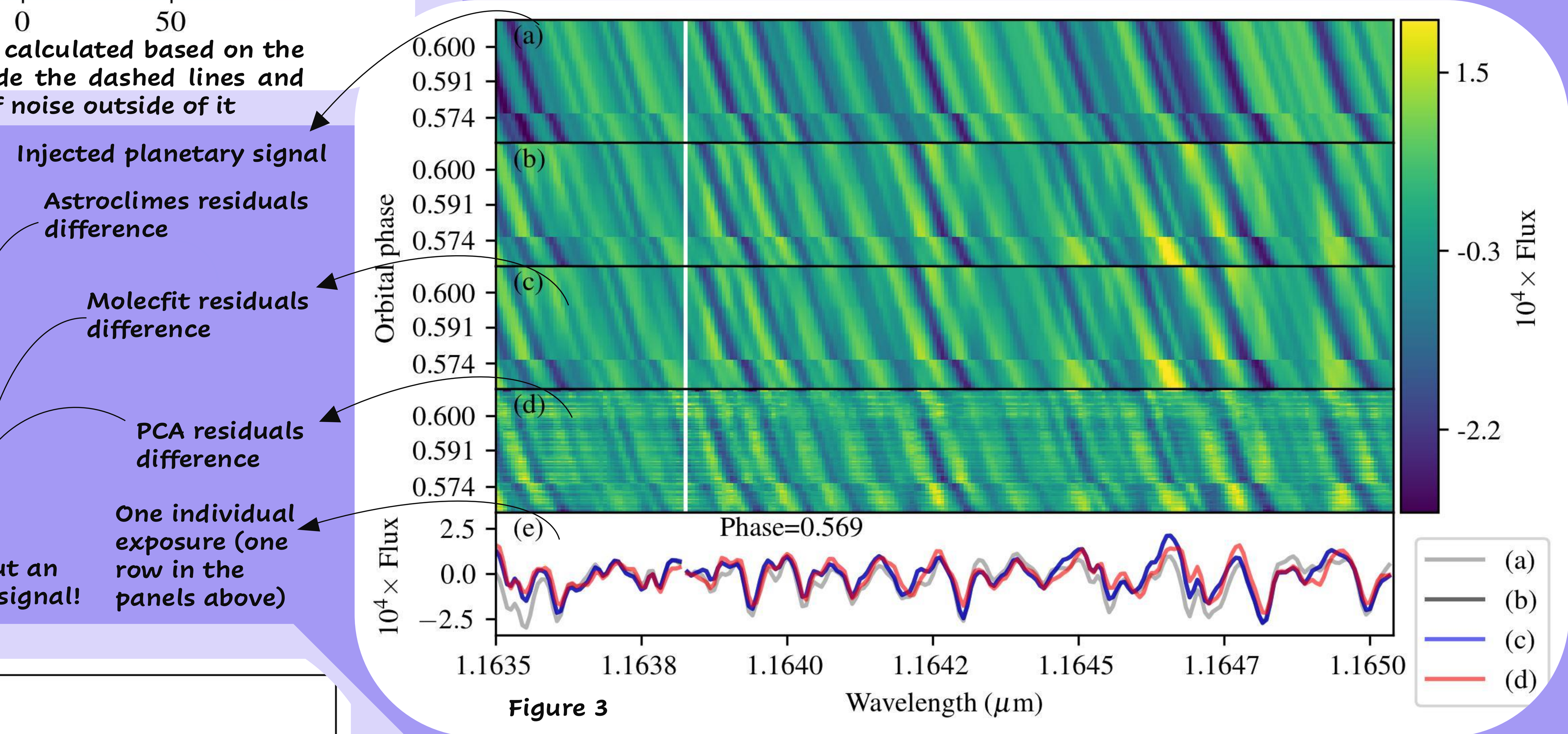

Figure 3

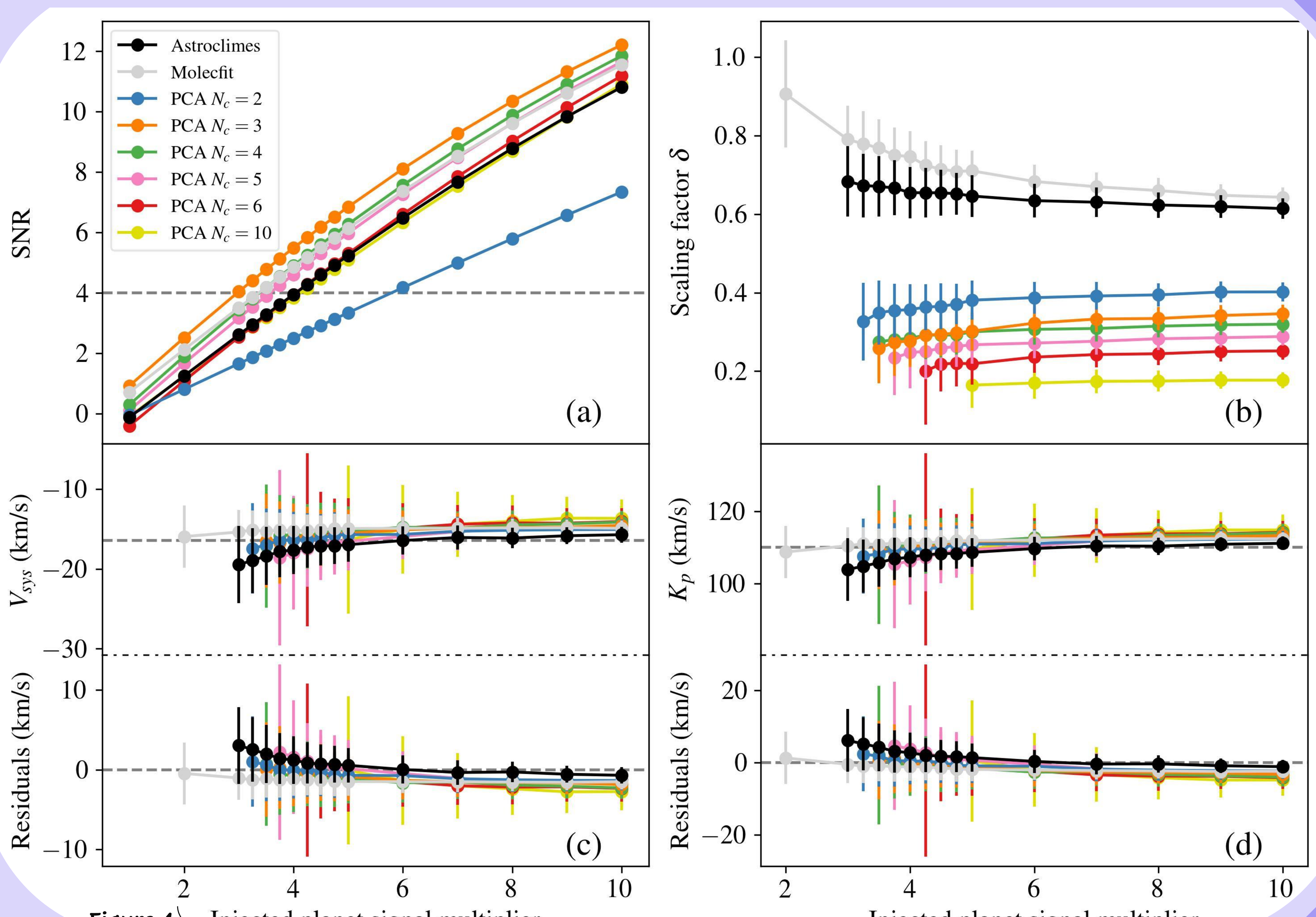

Figure 4 Injected planet signal multiplier

Planet model is scaled by a multiplier to increase signal strength

So, what is the work presented here?

We studied how the detection of a model planetary atmospheric signal was affected by three detrending methods (Astroclimes, Molecfit, PCA) with a series of injection and recovery tests. The model has H_2O only and is the best fit from [6].

After detrending the spectra with the injected signal, an MCMC was run to determine how the signal had changed both in position (based on K_p and V_{sys}) and magnitude (quantified by a scaling factor δ).

Takeaway messages

- PCA is better at removing the telluric lines and achieves higher SNR overall, but it is also more aggressive on the signal than Astroclimes or Molecfit (see Figures 3 and 4).
- PCA loses SNR and signal retention as we add more components, so it is crucial to choose the optimal number of PCA components (e.g. see [2]).

Please talk to me if…

- you need to get rid of telluric lines in your data;
- you want to know why it is important to better preserve a planet's signal albeit at lower SNR;
- you have any questions/criticism/feedback!

Additional analysis

- Here, only results from one night and K_p = 110 km/s are shown, but the full study has an extra night and explores the cases K_p = 50, 80 km/s.
- For the lower K_p values, Astroclimes/Molecfit not only achieve higher SNR, but also retain the signal better (PCA struggles more because a lower K_p means the planet signal does not move across pixels as much throughout the night).
- Extra night favours N_C = 5, highlighting the importance of choosing the optimal number of PCA components and showing that this number can change from night to night.
- Combining the data from five nights, we do not detect the signal reported in [6].

Check out the paper:

References

[1] Brogi, M., & Line, M. R. 2019, AJ, 157, 114.
[2] Cheverall, C. J., et al. 2023, MNRAS, 522, 661.
[3] Fetzner Keniger, M. A., et al. 2025. RASTI, 4, rzaf041.
[4] Giacobbe, P., et al. 2021. Nature, 592, 205.
[5] Smette, A. et al., 2015. A&A, 576, A77
[6] Webb, R., et al. 2022, MNRAS, 514, 4160-4172.