

# Small but mighty: High-resolution spectroscopy of ultra-hot Jupiter atmospheres with compact telescopes

## Transmission spectrum of KELT-9 b with Wendelstein's FOCES spectrograph

N. W. Borsato<sup>1,4</sup> , H. J. Hoeijmakers<sup>1</sup>, D. Cont<sup>2,7</sup> , D. Kitzmann<sup>3</sup> , J. Ehrhardt<sup>2</sup>, C. Gössl<sup>2</sup> , C. Ries<sup>2</sup>,  
B. Prinoth<sup>1</sup> , K. Molaverdikhani<sup>2</sup>, B. Ercolano<sup>2</sup>, H. Kellerman<sup>2</sup>, and Kevin Heng<sup>5,6</sup>

<sup>1</sup> Lund Observatory, Division of Astronomy and Theoretical Physics, Lund University, Box 43, 221 00 Lund, Sweden  
e-mail: [n.borstato@astro.lu.se](mailto:n.borstato@astro.lu.se)

<sup>2</sup> Universitäts-Sternwarte München, Fakultät für Physik, Ludwig-Maximilians-Universität München, Scheinerstrasse 1,  
81679 München, Germany

<sup>3</sup> University of Bern, Center for Space and Habitability, Gesellschaftsstrasse 6, 3012, Bern, Switzerland

<sup>4</sup> Research Centre for Astronomy, Astrophysics and Astrophotonics, Macquarie University, Balaclava Road, Sydney, NSW 2109,  
Australia

<sup>5</sup> ARTORG Center for Biomedical Engineering Research, Murtenstrasse 50, 3008, Bern, Switzerland

<sup>6</sup> Astronomy & Astrophysics Group, Department of Physics, University of Warwick, Coventry CV4 7AL, UK

<sup>7</sup> Exzellenzcluster Origins, Boltzmannstrasse 2, 85748 Garching, Germany

Received 9 August 2023 / Accepted 20 November 2023

### ABSTRACT

When observing transmission spectra produced by the atmospheres of ultra-hot Jupiters (UHJs), large telescopes are typically the instrument of choice given the very weak signal of the planet's atmosphere. The aim of the present study is to demonstrate that, for favourable targets, smaller telescopes are fully capable of conducting high-resolution cross-correlation spectroscopy. We apply the cross-correlation technique to data from the 2.1 m telescope at the Wendelstein Observatory, using its high-resolution spectrograph FOCES, in order to demonstrate its efficacy in resolving the atmosphere of the UHJ KELT-9 b. Using three nights of observations with the FOCES spectrograph and one with the HARPS-N spectrograph, we conduct a performance comparison between FOCES and HARPS-N. This comparison considers both single-transit and combined observations over the three nights. We then consider the potential of 2 m class telescopes by generalising our results to create a transit emulator capable of evaluating the potential of telescopes of this size. With FOCES, we detected seven species in the atmosphere of KELT-9 b: Ti II, Fe I, Fe II, Na I, Mg I, Na II, Cr II, and Sc II. Although HARPS-N surpasses FOCES in performance thanks to the mirror of the TNG, our results reveal that smaller telescope classes are capable of resolving the atmospheres of UHJs given sufficient observing time. This broadens the potential scope of such studies, demonstrating that smaller telescopes can be used to investigate phenomena such as temporal variations in atmospheric signals and the atmospheric loss characteristics of these close-in planets.

**Key words.** instrumentation: spectrographs – methods: observational – techniques: spectroscopic – telescopes – planets and satellites: atmospheres

## 1. Introduction

In ground-based high-resolution spectroscopic transit observations of exoplanet atmospheres, telescope size often determines the scale of the scientific aims, with larger telescopes typically favoured for these types of observations. Atmospheric observations of exoplanets suffer from photon scarcity, and as such, larger mirrors prove indispensable, collecting more photons and compensating for the feeble planet signal, which can be as low as a few parts per million relative to the flux of the host star in the case of hot Jupiters, and even lower for smaller planets (Brogi et al. 2012). This scarcity often results in the subtle absorption of the exoplanet signal becoming indistinguishable within the noise of the observation. These challenges limit detailed observational studies to certain subclasses of exoplanets, which include the hottest exoplanets known to date, ultra-hot Jupiters (UHJs; Parmentier et al. 2018).

UHJs orbit close to their host star and serve as excellent targets for atmospheric studies thanks to their typically

inflated atmospheres, which result in large transit depths (Baraffe et al. 2010; Merritt et al. 2021). The extreme irradiation these planets undergo as a result of their proximity to their host star leads to unique atmospheric effects, such as molecular dissociation (Kitzmann et al. 2018; Arcangeli et al. 2018; Parmentier et al. 2018), cloud-free day sides (Helling et al. 2021), equatorial jets moving at velocities of kilometres per second (Showman & Polvani 2011), and mass-loss through photo-evaporation (Lowson et al. 2023).

Owing to their unique atmospheric properties, UHJs serve as laboratories for exoplanet science. Their atmospheres are generally simpler to interpret due to relatively stable equilibrium conditions at pressures of higher than  $10^{-4}$  bars (Kitzmann et al. 2018; Fossati et al. 2021; Lee et al. 2022). Though non-local thermodynamic equilibrium (NLTE) effects and deviations from chemical equilibrium can occur in the upper and lower layers, respectively (Fossati et al. 2021; Arcangeli et al. 2021), the majority of the transmission spectra of UHJs are predicted to probe regions where equilibrium assumptions hold

**Table 1.** Summary information of the transit observations used in this study.

Date	Spectrograph	Wavelength range (nm)	Resolution	$S/N$ (300 s)	Exp-time (s)	Exp-number	Baseline fraction %
31-07-2017	HARPS-N	387–690	115 000	56	600	49	57
18-07-2022	FOCES	380–885	70 000	34	300	41	18
02-09-2022	FOCES	380–885	70 000	26	300	41	33
05-09-2022	FOCES	380–885	70 000	30	300	39	17

**Notes.** This table summarises the details of the four observations used in this study. Columns denote the date of observation, the spectrograph used, the operational wavelength range in nanometres, the spectral resolution, the  $S/N$  ratio per 300 s of exposure time, with the HARPS-N scaled through Poisson noise assumptions, the exposure time in seconds, the total number of exposures, and the percentage of observations outside of transit.

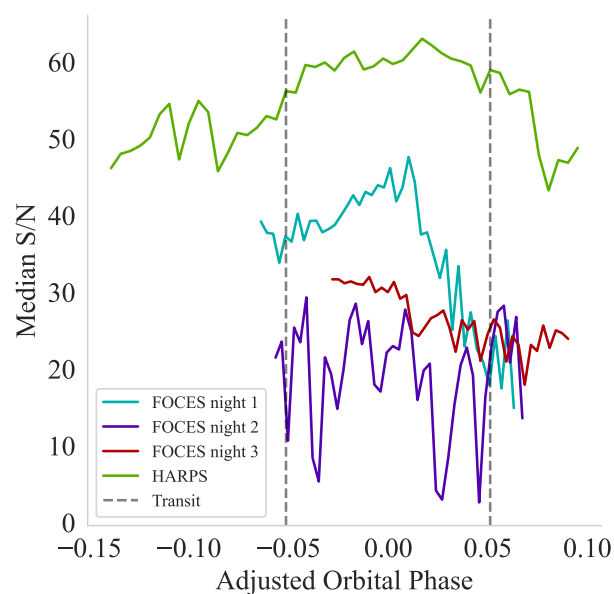
(Kitzmann et al. 2018; Lee et al. 2022). Indeed, they are ideal test targets for exploring the extent to which telescope size can limit atmospheric studies. Small telescopes are the workhorses of astronomy, enabling vital research that may be considered ‘high risk’ or may only yield results in the long-term. The lower subscription factors of these classes of telescope also enable the possibility of multiple observations of the same target, which may produce scientific results of the same quality as larger telescopes when combining these observations, and would also enable studies of the time evolution of the atmospheres of these exoplanets.

There is a growing list of telescopes that have successfully conducted high-resolution spectrographic observations of UHJs, including the Nordic Optical Telescope (with a mirror diameter of 2.5 m; Bello-Arufe et al. 2022), as well as the Fred Lawrence Whipple Observatory (Lowson et al. 2023), both of which have produced detections within a single transit. In the present work, we aim to add another 2 m class instrument to the pool of telescopes able to produce an atmospheric detection of a UHJ atmosphere in a single transit, namely the FOCES spectrograph on the 2 m Wendelstein telescope located in the Bavarian alps (Pfeiffer et al. 1998). Additionally, we aim to expand the discussion to consider the full potential of 2 m telescopes in general, and specifically in observing known UHJs. The target of choice is KELT-9 b, an UHJ with an equilibrium temperature of approximately 4000 K that has proven to be a highly observable target for benchmark cases such as that presented here (e.g. Hoeijmakers et al. 2018; Yan et al. 2019; Bello-Arufe et al. 2022; Lowson et al. 2023).

Furthermore, we aim to compare the performance of FOCES to that of the HARPS-North (HARPS-N) high-resolution spectrograph on the Telescopio Nazionale *Galileo* (TNG), which is an instrument used as standard in the study of exoplanet atmospheres (e.g. Casasayas-Barris et al. 2019; Stangret et al. 2020, 2021; Cont et al. 2021, 2022; Borsa et al. 2022; Pino et al. 2022). Following this comparison, we generalise and extrapolate our results to comment on the potential of 2 m class telescopes and their ability to resolve exoplanet atmospheres.

## 2. Observations

We used FOCES to observe three transits of KELT-9 b on 18 July 2022, 2 September 2022, and 5 September 2022 (only a partial transit). Each transit was observed in time series, taking continual exposures of 300 s in length throughout the transit, plus some additional exposures out of transit to obtain a stellar baseline without the planet signal. The other data set used is a single transit observation taken using the HARPS-N spectrograph on the TNG, which is available in a public data archive and was initially



**Fig. 1.** Time trend of the median  $S/N$  for each transit observation of KELT-9 b as a function of orbital phase. The HARPS-N  $S/N$  values have been scaled to values that would be expected to be produced for the exposure times taken by the FOCES spectrograph. The HARPS-N transit covers the entire planet, and a substantial amount of the baseline. Additionally, the  $S/N$  is consistently higher throughout the entire observation compared to the observations taken by FOCES. Signal quality also drops off for all the FOCES nights towards the end of each night.

published in Hoeijmakers et al. (2018). The transit data were obtained on 31 July 2017, with an integration time of 600s per exposure frame. The observation encompassed the entire transit and baseline measurements. Table 1 provides a summary of the observations taken with each spectrograph, including the relevant characteristics of the instruments and the relative amount of stellar baseline taken compared to the length of the transit. Airmass plots for each transit are shown in Fig. B.1; these depict where on the sky the transit occurs and during what epoch of the night.

Using the approach described by Borsato et al. (2023), we measured the average quality of the data by taking the median signal-to-noise ratio ( $S/N$ ) over all the echelle orders; this allows us to track the progression of the data quality over time throughout the transit. The progression of the  $S/N$  over time for each observation is shown in Fig. 1 as a function of orbital phase, with the beginning and end of the transit marked. The figure provides a good overall picture of how the observation progressed

throughout each night. So as to be able to compare the observations, we divided the S/N values for the HARPS-N night by a factor of  $\sqrt{2}$  (assuming a Poisson noise distribution) in order to correct for the time difference in the exposures.

Among the three FOCES nights, the first exhibited the highest S/N, which later dropped by a factor of two towards the end of the observation. The night began with an overcast sky of cirrus clouds. However, these clouds began to dissipate just before observations began. Approximately midway through observations, the signal quality dropped due to a slow build-up of clouds towards the end of the night. The S/N on the second night of observations showed significant variation throughout the entire transit. The entire night was characterised by the presence of cirrus clouds of varying opacity. Nevertheless, these clouds did not completely obscure the sky, allowing some measurements to be obtained. Throughout the night, the signal dropped steadily and oscillated periodically. The third night was mainly clear, with occasional thin and narrow cirrus clouds passing by, but due to twilight timing constraints, only a partial transit could be obtained. As the night progressed, the signal dropped slightly during the exposure times, likely due to fast-moving clouds partially obscuring the observations. However, these clouds were transient, leading to a patchwork drop in night flux rather than a consistent decrease. The HARPS-N spectrograph outperforms all three FOCES observations by a factor of approximately two in terms of S/N, and no significant problems with the weather were noted.

### 3. Data processing

FOCES spans a wavelength range of between 383 and 885 nm at a spectral resolution of  $R \sim 70\,000$  (Pfeiffer et al. 1998; Brucalassi et al. 2012; Wang et al. 2017). The primary science products derived from the spectrograph's pipeline include two-dimensional echelle data, recorded errors, and blaze profiles. We stitched the overlapping pixels of the echelle orders by averaging the overlapping flux values and using the recorded errors as weights, thereby generating a one-dimensional spectrum of each exposure. The HARPS-N spectrograph spans a wavelength range from 387 to 690 nm with a resolution of  $R \sim 115\,000$ . The scientific products available from these observations include the two-dimensional echelle orders and the blaze-corrected, stitched, and resampled one-dimensional spectra.

The gathered data necessitated the removal of systematic signals extraneous to the planet. We adopted the methodology outlined in Hoeijmakers et al. (2020) for this process, which entailed the elimination of atmospheric telluric contamination, velocity offsets, the interstellar medium (ISM), the host star signal, and residual contamination in the continuum. The same process for the removal of systematic signals was uniformly applied to both the FOCES and HARPS-N datasets, despite their differing S/N levels. To remove the telluric lines produced by water and carbon dioxide, we used the one-dimensional spectra. We were able to model the telluric lines using the software package `molecfit` (Smette et al. 2015; Kausch et al. 2015), thereby creating an atmospheric telluric model across the full wavelength range of the data analysed. After modelling, we remove the telluric lines from the observation by dividing by a telluric model produced for each exposure.

Furthermore, there are two velocity offsets that require correction; these are caused by two Doppler shifts, those arising from the Earth–Sun barycentric velocity and the radial velocity of the star produced by the planet's gravitational pull. The

correction leaves the spectra at a constant radial velocity shift set by the systemic velocity of approximately  $-18\text{ km s}^{-1}$  (Gaudi et al. 2017). The barycentric velocity correction is calculated using the Python library `astropy`, while the radial velocity offset is available from the discovery paper (Gaudi et al. 2017).

To remove the signal of the host star, we compute the average spectrum over all the exposures of KELT-9 b, and divide each exposure by this average; this removes the static components of the spectrum while maintaining the time transient properties, such as the signal from the planet. Dividing by the average produces a residual continuum shape, which needs to be flattened using a median filter.

Finally, any bad pixels remaining are flagged as not a number (NaN) using a running standard deviation, and any values exceeding  $5\sigma$  are flagged. The data are further cleaned by applying a Gaussian high-pass filter with a bandwidth of  $80\text{ km s}^{-1}$ , and once again marking any outliers beyond the  $5\sigma$  limit. Finally, any residual ISM effects are manually flagged as NaNs. The NaNs generated from this process create an outlier mask that enables us to bypass problematic pixels during the cross-correlation process. All these steps are applied through the cross-correlation software package `tayph`<sup>1</sup>, which is under development at Lund Observatory and forms the basis for similar recent cross-correlation studies (e.g. Hoeijmakers et al. 2020, 2024; Prinoth et al. 2022, 2023; Borsato et al. 2023).

### 4. The cross-correlation technique

To search for atmospheric signals, we applied the cross-correlation method. This innovative approach was first introduced by Snellen et al. (2010), enabling the detection of CO in the transmission spectrum of HD 209458 b. The essence of this method is to combine the flux contributions of specific spectral lines within a planet's transmission spectrum, thereby effectively generating a weighted average of the spectral line profile for a given species based on its known line positions.

However, the variable radial velocity of the planet introduces a shift in the spectral line positions within the wavelength space. To accommodate this variability, the technique employs Doppler shifting to adjust the known line positions to different radial velocity values. This step allows the spectrum to be scanned continuously for these shifted lines.

The real strength of this method becomes evident when the predicted line positions align with the planet's absorption spectrum. This alignment leads to a correlation peak, which signifies the successful detection of a species within the planet's atmosphere.

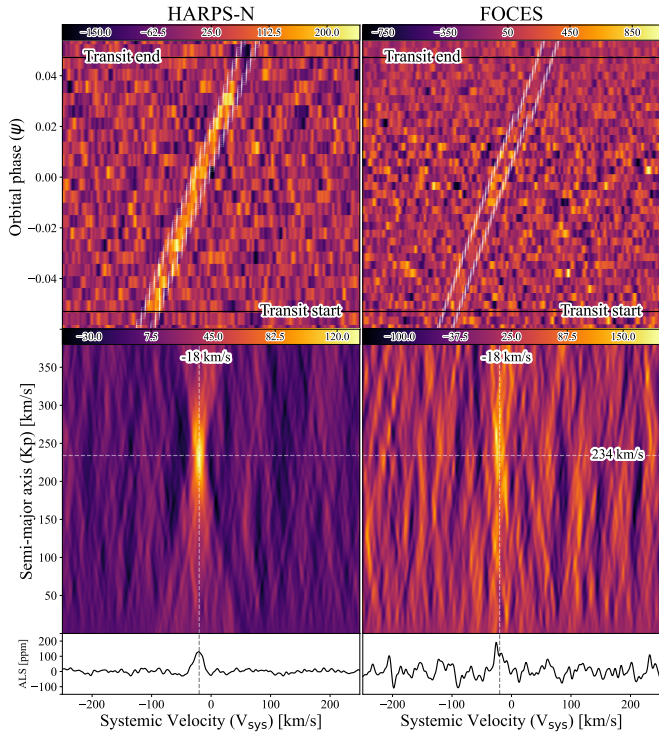
In order to carry out this method, it is necessary to construct a spectral template,  $T$ . This template calculates the transmission spectrum of either the entire planet or an individual species. It is then used to generate a set of weights, which are assigned to the anticipated line positions of the species under investigation:

$$\hat{T}_i(v) = \frac{T_i(v)}{\sum T_i(v)}. \quad (1)$$

Here, the template function,  $T_i(v)$ , is dependent on the radial velocity, where  $i$  refers to the species under investigation, and  $v$  denotes the radial velocity of the template. The weights are scaled so that the total sum of the templates is equal to one, with large weight values allocated to the line positions. This weighted

<sup>1</sup> <https://github.com/Hoeijmakers/tayph>





**Fig. 2.** Cross-correlation results for searching for Fe I in a single transit. On the left are the results for the HARPS-N spectrograph on the TNG, and on the right we show the results for the FOCES spectrograph for the night of 18 July 2022. Top panel: cross-correlation maps for Fe I. The pixels covering the trace of the planet are outlined to help guide the eye to where to expect the signal. Middle panel:  $K_p V_{\text{sys}}$  maps for each Fe I detection. The detections are the bright features in the centres of the plots. The dashed cross-hairs show the row where the one-dimensional CCF (bottom panel) has been extracted and the expected systemic velocity of KELT-9 b ( $-20 \text{ km s}^{-1}$ ). Bottom panels: one-dimensional CCFs for the single transits, showing the average line depth for the given signal. The dashed line shows the expected systemic velocity of KELT-9 b, which is the same as in  $K_p V_{\text{sys}}$  map.

average is then employed to calculate the mean line depth, as illustrated in the following equation:

$$c(v) = \sum_{i=0}^N x_i \hat{T}_i(v), \quad (2)$$

where  $c(v)$  represents the computed mean depth of absorption lines, where each line is weighted relative to its position at a specific Doppler shift. The dependence on the velocity term implies that the final value obtained from the sum will change depending on the Doppler shift applied to the calculation. The total quantity of evaluated wavelength bins is represented by  $N$ . Each  $x_i$  corresponds to the spectral data that align with the same wavelength segment  $i$  as the reference template.

The cross-correlation function (CCF) leverages the fact that the radial velocity of a planet changes throughout its transit, leading to a shift in the line positions of the transmission spectrum. By applying Eq. (2), a peak will form at the correct velocity position if the species is present in the planet’s atmosphere. This process results in an absorption trace on a two-dimensional grid of time versus planetary radial velocity, as shown in the top panel of Fig. 2.

After applying the CCF, further processing steps are necessary to detect an atmospheric signal. The process consists

of several sequential actions. Initially, residual broadband variations stemming from the removal of the stellar signal are eliminated. Subsequently, the Rossiter–McLaughlin (RM) effect (Rossiter 1924; McLaughlin 1924), a phenomenon caused by the planet temporarily distorting the line-of-sight velocity of the star by obscuring part of the surface of the rotating star (commonly known as the Doppler shadow) is modelled and removed. The methodology adopted for removal of broadband variation and compensation for the Rossiter–McLaughlin effect follows that of Hoeijmakers et al. (2020), who successfully conducted a chemical inventory of WASP-121 b. Finally, we remove any residual stellar aliases by applying the same vertical detrending adapted from Prinoth et al. (2022) and implemented by Borsato et al. (2023) to remove residual cross-correlation signatures coming directly from the star’s spectra.

It is also possible to enhance the S/N of a cross-correlation function by co-adding all pixels along the trace of the planet. This is achieved by shifting all exposures into the time-independent rest frame of the planet using its orbital phase and inclination. This process can be executed with the following equation:

$$v_{rv} = v_{\text{orb}} \sin 2\pi\phi \sin i + v_{\text{sys}}, \quad (3)$$

where  $v_{\text{orb}}$  indicates the exoplanet’s orbital velocity around its host star,  $\phi$  represents the orbital phase of the exoplanet,  $i$  refers to the inclination of the system relative to the observer’s line of sight, and  $v_{\text{sys}}$  corresponds to the systemic velocity of the host star.

We shift the cross-correlation values according to the equations for a presumed semi-major amplitude ( $K_p$ , which is  $v_{\text{orb}} \times \sin i$  in Eq. (3)) by applying the known orbital phase of the planet. To find the optimal alignment, we sample a range of  $K_p$  values. Each one is then collapsed along the time axis by computing a vertical average. When the shift leads to vertical alignment of the cross-correlation function, a peak becomes evident. However, if the alignment is not accurate, the cross-correlation functions will not constructively interfere with each other. As a result, we expect to see a peak at the correct  $K_p$  value.

We construct a grid, or a  $K_p V_{\text{sys}}$  map (Brogi et al. 2012), from all the collapsed cross-correlation functions corresponding to the various shifts. This map is identified by a peak at the correct  $K_p$  and  $v_{\text{sys}}$  values. The row corresponding to this peak provides an optimised one-dimensional CCF, where all exposures are combined to yield the most robust signal. For this study, we used a grid of semi-major amplitudes spanning from 0 to  $400 \text{ km s}^{-1}$ , incremented in steps of  $1 \text{ km s}^{-1}$ .

Finally, we extract the one-dimensional CCF by selecting the  $K_p$  that best aligns the CCF functions. In order to determine the amplitude, width, and centre locations of the signal, we employ a least-squares fit to the one-dimensional CCF under the assumption of a Gaussian profile. Uncertainties in the amplitude, width, and centre locations are derived via a bootstrap method, which involves sampling the noise floor of the one-dimensional CCF.

This process begins by removing a region in the one-dimensional CCF where the peak is located, specifically for CCF values ranging from  $-50 \text{ km s}^{-1}$  to  $10 \text{ km s}^{-1}$ . We then compute the standard deviation of the remaining values to estimate the standard deviation of the noise. Assuming a normally distributed noise profile based on this computed standard deviation, we create a Gaussian noise model. We then bootstrap all points in the CCF, randomly altering each point using this noise model, applying the same least-squares Gaussian fit, and recording the measured parameters each time. After repeating this process for

5000 iterations, we calculate the standard deviation of the fitted values, treating these as their uncertainties. The statistical significance of the signal is then evaluated as the ratio of the measured amplitude to its corresponding uncertainty.

#### 4.1. Combining the FOCES nights

To amalgamate the three FOCES observations into a single  $K_p V_{\text{sys}}$  map, we employed an approach adapted by Borsato et al. (2023), which involves computing a weighted average of the  $K_p V_{\text{sys}}$  maps. Given that only one spectrograph is under consideration, we do not need to factor in the instrument's wavelength range, which is part of the process used in this latter publication. We therefore proceed by constructing appropriate weights to combine the nights.

Firstly, we mask out the signal from the planet in the  $K_p V_{\text{sys}}$  map. Our mask was devised to screen all data values ranging from  $-50 \text{ km s}^{-1}$  to  $10 \text{ km s}^{-1}$  for the systemic velocity, and  $100 \text{ km s}^{-1}$  to  $300 \text{ km s}^{-1}$  for the orbital velocity, which is the same region applied in Borsato et al. (2023). The selected window intentionally spans a broad velocity range to capture all known atmospheric detections of exoplanets, while allowing for possible offsets arising from atmospheric and dynamical effects (Prinoth et al. 2022). Additionally, changes in the region boundaries do not have a large effect on results provided the signal is contained within it. Upon applying the mask, we calculate the standard deviation of the remaining pixels, providing a measure of any residual noise and systematic uncertainties affecting the overall noise floor.

Next, we invert the mask, thereby isolating the region where we anticipate the signal. Any signal within this location means that the region will yield a higher average value compared to the background. By taking the ratio of this mean to the calculated standard deviation, we generate a pseudo-S/N, which offers a general assessment of the quality of the night in terms of S/N and can then be converted into the desired weight. Lastly, we normalise the pseudo-S/N values for all observations so that they add up to one, using these resultant values as the weights.

#### 4.2. Cross-correlation templates

In this study, we used the standard templates provided by the Mantis network (Kitzmann et al. 2023), which comprises more than 750 high-resolution templates for various species across a grid of different atmospheric temperatures. These templates are publicly available<sup>2</sup> and are computed using a forward modelling approach, taking into account the transmission spectrum of the target planet and the absorption lines of the species under scrutiny. The computation process includes opacity functions, continuum opacity contributions, and assumptions about metal abundance and temperature profiles. These templates provide a standardised and efficient alternative to creating new templates for each individual study. Mantis templates have been successfully applied in the detection of a range of species using the ESPRESSO spectrograph on the VLT (Azevedo Silva et al. 2022). Our search only focused on the species reported in Hoeijmakers et al. (2019) – which have an assumed equilibrium temperature of 4000 K – to simplify the comparison between telescopes.

<sup>2</sup> <https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/669/A113>

#### 4.3. Model injection and comparing instruments

In order to make a comparative analysis of the data from FOCES and HARPS-N, our initial step was to create a forward model of KELT-9 b's transmission spectrum, taking inspiration from similar studies (e.g. Hoeijmakers et al. 2019). We used the same model as described by Kitzmann et al. (2023) to compute the transmission spectrum but changed the planet's radius and surface gravity to that of KELT-9 b. The temperature profile of the atmosphere was assumed to be isothermal, with a value of 4000 K. We used all opacity species listed in Table 2 of Kitzmann et al. (2023). The corresponding line absorption coefficients were calculated with the HELIOS-K<sup>3</sup> opacity calculator (Grimm & Heng 2015; Grimm et al. 2021). The chemical composition of the atmosphere was obtained using the equilibrium chemistry code FASTCHEM 2<sup>4</sup> (Stock et al. 2018, 2022), assuming solar element abundances. The resultant transmission spectrum, which is representative of the observed atmospheric characteristics, is shown in Fig. C.1.

Following the methodology outlined in Hoeijmakers et al. (2015; which searched for TiO in HD 209458 b, and used model injection to explain its non-detection), we take the predicted transmission spectrum, inject it into the pipeline-reduced spectra obtained from the observations of both instruments, and rotationally broaden the modelled lines to account for the planet's rotation. Subsequently, we perform a Gaussian convolution to match the spectral resolution of each spectrograph. The wavelength axis is interpolated in order to align with the spectrograph grid. Finally, we perform cross-correlation on the spectra a second time. This approach facilitates a comparison between the expected cross-correlation signal and the actual signal generated by the spectrograph. By subtracting the cross-correlation results for the case with an injected signal from the cross-correlation profile without an injected signal we obtain a line depth and shape that are predicted by the model. These are shown in Fig. 3. We then use this predicted line to calculate the model discrepancy, as detailed in Table 1.

### 5. The detection feasibility of a 2 m telescope

While our observations focus on the results from observations made at the Wendelstein Observatory, we aim to expand our analysis to consider 2 m class telescopes in general. Our goal in this section is to develop a transit emulator<sup>5</sup> that generates mock cross-correlation results of transit spectra. To do this, we use measurements of the S/N of Wendelstein as a baseline to extrapolate the capability of 2 m class telescopes in observing exoplanet atmospheres. For the best Wendelstein night, the  $S/N \sim 35$  for the host-star KELT-9, which has a visual magnitude of 7.6. Taking this as an ideal S/N, we can create an approximate analytical equation that estimates the S/N for a specific stellar magnitude:

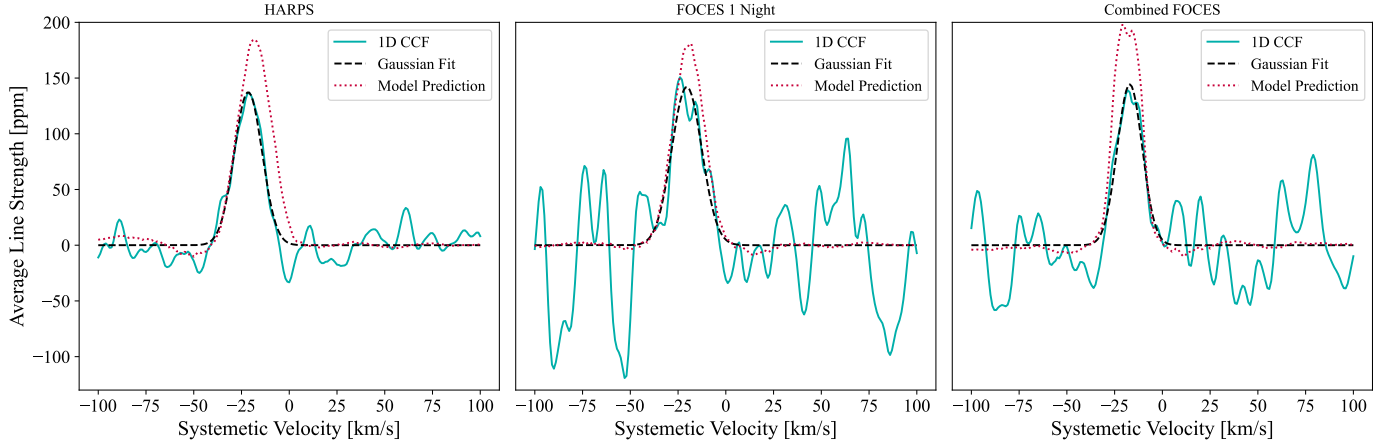
$$S/N \approx 35\sqrt{10^{(m-7.6)}}, \quad (4)$$

where  $m$  is the magnitude of the star in question. We require a model in order to evaluate the detectability of the atmosphere of a UHJ. In this case, we selected spectral templates for Fe I from the Mantis template library of Kitzmann et al. (2023) and used these as models of the atmosphere of UHJs in general. We chose these templates over our own generated model to simplify the approach and increase reproducibility. To simulate the

<sup>3</sup> <https://github.com/exoclimate/HELIOS-K>

<sup>4</sup> <https://github.com/exoclimate/FastChem>

<sup>5</sup> [https://github.com/nborsato/transit\\_emulator](https://github.com/nborsato/transit_emulator)



**Fig. 3.** Comparison of 1D CCFs for the FeI signal between different datasets extracted from the peak of each  $K_p V_{\text{sys}}$  plot. Each panel represents the data from a different source: HARPS-N, FOCES Night 1, and FOCES nights 1, 2, and 3 combined, respectively. In each plot, the teal line represents the one-dimensional cross-correlation function, the black dashed line is the Gaussian fit to the peak in the plot, and the red dotted line is the model prediction, as discussed in Sect. 4.3. There is a small discrepancy in the model peak height for each result; this is likely driven by the larger wavelength range of FOCES, which changes the predicted cross-correlation result. The model-predicted CCF has values of slightly below zero near the peak of the function. This observation is attributed to cross-correlation aliases, a phenomenon explained in Borsato et al. (2023).

transit of a UHJ, we first reduced the templates down to a resolution of 70 000 by applying Gaussian convolution. For this general case, we used the KELT-9 b transit duration of 3.91 h (Gaudi et al. 2017), which we divided into steps of 330 s to simulate an exposure time of 300 s plus 30 s of overhead readout.

With each step, we calculate the phase angle of the planet:

$$\phi = \frac{2\pi(t)}{p}, \quad (5)$$

where  $t$  is the time step in question, which has a value range of  $t \in \left[-\frac{3.91}{2}, \frac{3.91}{2}\right] h$ , and  $p$  is the orbital period of the planet. With orbital velocity, phase angle, and systemic velocity obtained from Gaudi et al. (2017), we apply

$$v_r = v_{\text{orb}} \sin \phi + v_{\text{sys}} \quad (6)$$

to obtain the radial velocity of the planet at each exposure time, and we then create a time series to emulate the exoplanet transit by Doppler shifting the template for each calculated radial velocity value. This results in 42 exposures for the entire transit. We also calculate the differences in the radial velocity steps, and further broaden the spectral lines of exposure to account for the line broadening caused by the Doppler shift in radial velocity space that the planet undergoes through each exposure.

Finally, we add Gaussian noise with a mean of zero to each spectrum using the S/N estimate provided by Eq. (4), converting it to a standard deviation. This process creates a set of mock transit exposures for which we can perform cross-correlation. We then follow the procedure outlined in Sect. 4 to constrain detection significance.

With our mock-transit spectrum generator, we first emulated a transit of KELT-9 b using the setup presented above to compare how well our process matches the observations. Further, we took an average of four emulated nights to determine the extent to which we should expect the statistical significance to increase when combining nights, and to determine whether or not combining four transits will produce the same results as a 4 m class telescope. Finally, with our understanding of the performance of this emulator, we emulated observations using FeI templates at equilibrium temperatures of 2500 K, 2000 K,

and 4000 K for stars of different magnitude, ranging from 6 to 11.5 in steps of 0.5 mag. We used three templates in order to consider the effect that equilibrium temperature has on detection significance and to gain a better idea of the potential observability of the planets. We emulated 30 transits for each magnitude and took combined averages over an increasing number of nights, up to 30, observing the increase in statistical significance with the addition of each new transit.

## 6. Results

We first present a comparative visual representation of the FeI detection from a single night using HARPS-N and the first night with FOCES. We illustrate these results in Fig. 2, where in the upper portion of the plots we show the cross-correlation function before conversion to a  $K_p V_{\text{sys}}$  map. In the centre of these plots, we present the  $K_p V_{\text{sys}}$  map, which shows the signal of Fe I for both spectrographs; though the HARPS-N signal is much stronger. Finally, in the lower portion of these figures, we present the extracted one-dimensional CCF, which shows the idealised signal produced from the cross-correlation process. We are able to detect atomic iron using both spectrographs, as reflected in the  $K_p V_{\text{sys}}$  plots and the one-dimensional CCFs. However, a prominent difference emerges in the trace of the planet within the CCF: HARPS-N data clearly reveal the trace, whereas it is only partially discernible in the FOCES output. A closer look at the one-dimensional function illustrates the dominance of noise in the FOCES data for that night. Nevertheless, the emergence of the signal confirms the feasibility of resolving the atmosphere of KELT-9 b with a single transit, even when baseline out-of-transit observations are minimal.

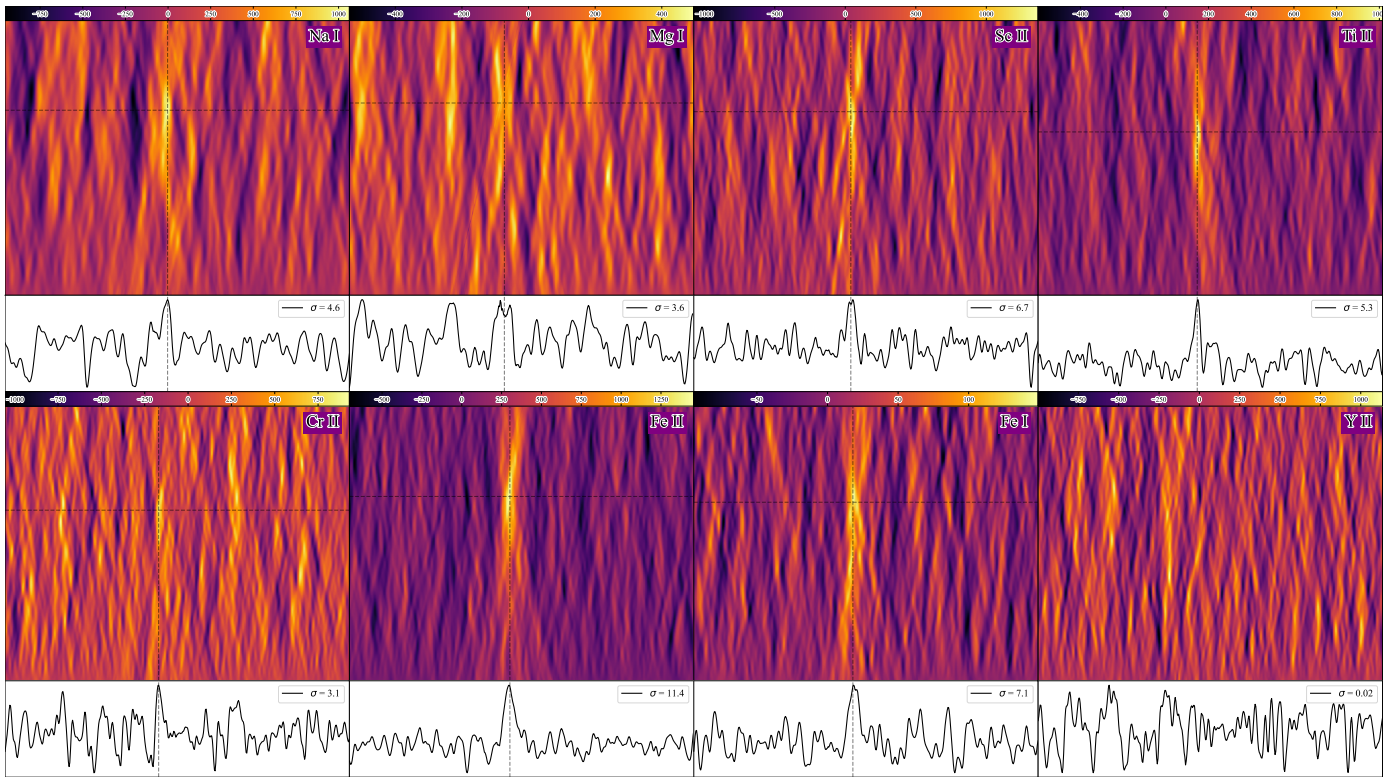
In Fig. 3, we compare the performance of the FOCES to that of the HARPS-N spectrograph for a single night of observation, and when all three nights are combined. We do this by plotting the one-dimensional CCFs with their corresponding Gaussian fits, and the signal predicted by the model as described in Sect. 4.3. The values and statistics for each fit are provided for comparison in Table 2, including how the model compares to the actual data. At first glance, it is clear that HARPS-N provides



**Table 2.** Comparative statistics of CCF from the HARPS-N and FOCES observations for the Fe I detection.

Parameter	HARPS-N	FOCES	Combined FOCES
Amplitude (ppm)	$132.7 \pm 6.9$	$142.1 \pm 24.4$	$123.1 \pm 17.2$
Centre ( $\text{km s}^{-1}$ )	$-21.3 \pm 0.5$	$-20.4 \pm 1.6$	$-16.4 \pm 1.7$
Width ( $\text{km s}^{-1}$ )	$7.0 \pm 0.4$	$7.9 \pm 1.5$	$8.2 \pm 1.2$
Detection significance	19.2	5.8	7.2
Model discrepancy	$0.3 \pm 0.04$	$0.1 \pm 0.1$	$0.3 \pm 0.1$
Discrepancy significance	7.5	1.0	3.0

**Notes.** This tables shows the summary statistics from the bootstrap fits of the one-dimensional cross-correlation function for Fe I. The column labelled ‘FOCES’ showcases the fit for the most successful observation night, which was on 18 July 2022. Furthermore, the ‘Combined FOCES’ column presents a fit applied to results generated by following the methodology described in Sect. 4.1. Detection significance is the ratio of the amplitude fits to their measured standard deviation. Model discrepancy is the measured difference between the predicted signal of the model and the actual signal. It is measured as one minus the ratio of the fitted amplitude of the cross-correlation function against the height of the model’s peak.



**Fig. 4.** Cross-correlation results for the combined FOCES observations across the eight species investigated. Ordered left to right, then top to bottom are: Na I, Mg I, Sc II, Ti II, Cr II, Fe II, Fe I, and Y II. The top panels of each detection contain the  $K_p V_{\text{sys}}$  maps. The cross-hairs locate the points where the signal peak has been extracted, and the bottom plots show the one-dimensional cross-correlation slice after extracting the rows containing the peak. The detection significance ( $\sigma$ ) is indicated in the legend. We have managed to successfully recover seven of the eight detections reported in Hoeijmakers et al. (2019). Only Y II has not been recovered, but it is known to be a weak signal and difficult for the HARPS-N spectrograph to resolve. While the signals are present in most plots they exhibit a strong degree of noise, indicative of using a 2 m class telescope.

superior signal quality even when combining the transits. However, the line depths are comparable across all three nights. When comparing to the model injection with the observations depicted in Fig. 3, the discrepancy, computed as the difference between the signal and the model, is approximately equivalent for both the HARPS-N observation and the combined FOCES observations, although the uncertainties associated with the FOCES discrepancy are nearly twice as large as those produced by HARPS-N. This implies that, when juxtaposed with our model, the signals derived from both spectrographs diverge from the model to a similar extent.

The  $K_p V_{\text{sys}}$  maps and one-dimensional CCFs for the combined FOCES nights, across all eight detections, are displayed in Fig. 4. The amplitude, width, centre locations, and detection significance are detailed in Table A.1. Using our method, we were able to successfully replicate seven out of the eight detections previously mentioned in Hoeijmakers et al. (2019). The only exception was Y II, which we did not find in our results. This is consistent with the difficulties encountered in reanalysing the HARPS-N data to confirm the presence of Y II, as noted by Borsato et al. (2023). If Y II is indeed present, its relatively weak signal is likely due to its low abundance in the planet’s

atmosphere. More definitive conclusions about the presence of Y II would require further observations, either through combined efforts or with the use of a larger telescope.

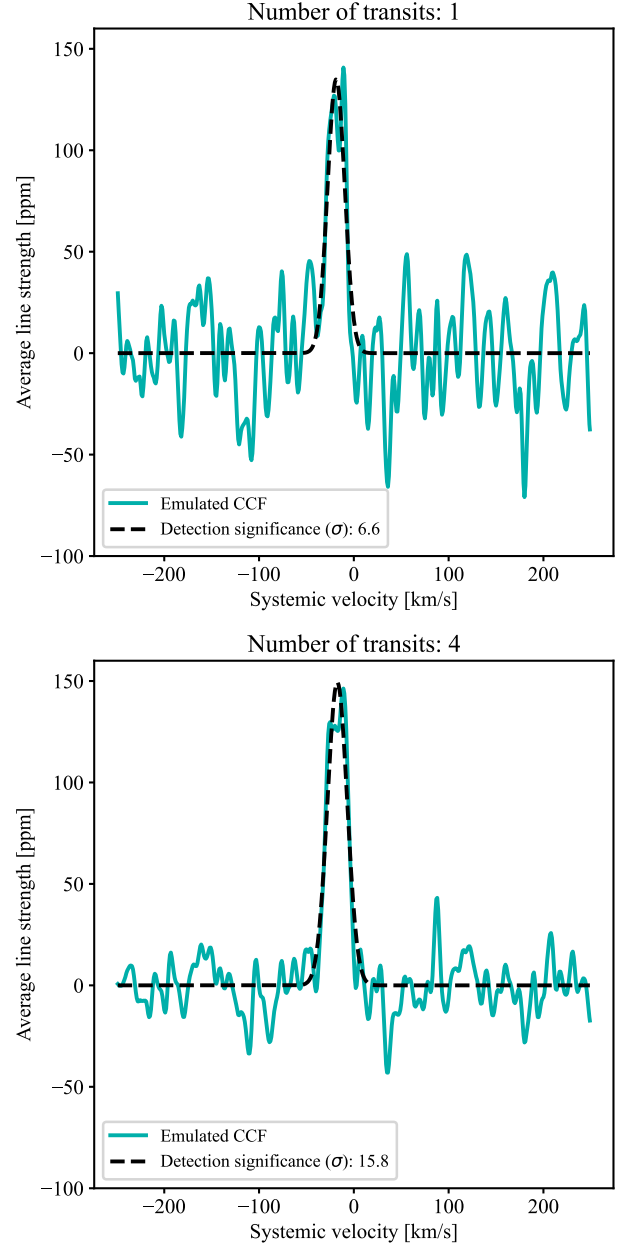
Among the detected signals, Fe II appeared to be the strongest. However, we also noted substantial detections of Fe I and Ti II, along with weaker detections of Na I, Mg I, Na II, Cr II, and Sc II. The central peaks of these signals were measured using a least-squares Gaussian fit, which estimates the best fit for the centre, width, and amplitude of the cross-correlation peak. The derived parameters aligned with the findings of Hoeijmakers et al. (2019), with the uncertainties of our measurements overlapping with those reported by these authors. However, the tentative Mg I detection merits particular attention; it exhibits considerable smearing and is significantly delocalised from its actual signal, implying that there may be additional issues with this detection, and that further exploration is warranted.

We now move on to review the results of our transit emulator. Referring to Fig. 5, we see that our emulation yields a significance value of  $6.6\sigma$  for a single emulated transit, which compares favourably with the FOCES value of  $5.8\sigma$  in Table 2, indicating a slightly higher significance. This discrepancy likely stems from the fact that the S/N drops towards the end of the observation, while in the emulated case, it remains consistent. When averaging over four nights, the significance increases to  $15.8\sigma$ , which aligns well with a single HARPS-N night; although the significance is slightly lower, it represents a much larger improvement compared to the stacked observations produced by the FOCES spectrograph. Our results from running the emulator with the 2500 K template are depicted in Fig 6. We note that single-transit detections are only possible at a  $5\sigma$  significance level for magnitudes of 7.5 or lower, while dimmer targets necessitate multiple observations. However, even planets orbiting around stars of 10.5 mag attain a constrained  $5\sigma$  detection if approximately 20 transits are captured, suggesting that all the UHJs listed in Table 3 could feasibly be observed with a 2 m class telescope. A final point to note is that detection significance appears to plateau, particularly for the brighter targets; for instance, the magnitude 6.0 case stabilises around a detection significance of  $20\sigma$  after about 7 transits. We discuss this point further in the following section.

## 7. Discussion

In ideal observing conditions, the HARPS-N spectrograph on the TNG can resolve the atmosphere of KELT-9b with impressively high fidelity with a single transit observation. However, our results – shown in Fig 2 – demonstrate that FOCES, on a telescope half the size of the TNG, can also resolve the atmosphere of this source even when observation conditions are poor. We therefore propose that by using multiple transits, results comparable to those achieved with HARPS-N can also be realised using moderately sized telescopes provided that the stellar signal can be resolved and that a sufficient number of transits can be obtained.

The HARPS-N spectrograph outperforms the FOCES spectrograph by a factor of two to three considering all three combined nights of FOCES observations. This disparity is larger than anticipated based solely on telescope size; we would have expected four nights to be sufficient to yield detection statistics comparable to those of HARPS-N from one night. Therefore, there are other factors degrading the quality of these observations.

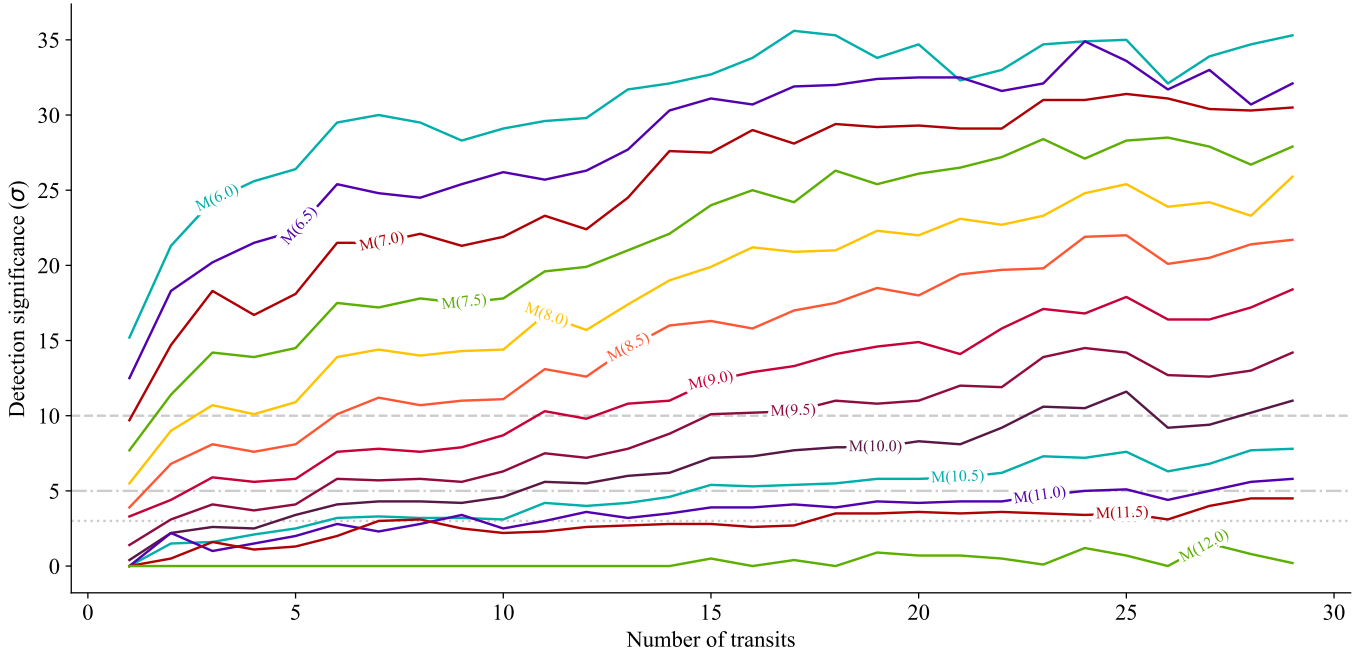


**Fig. 5.** Signal enhancement through transit stacking as demonstrated by the emulator developed in this study. Both panels depict the one-dimensional cross-correlation function (teal) alongside the Gaussian fit to the line profile (black dashed lines) for a single night, with the significance of the fit specified in the legend. The top panel illustrates the results for one night, whereas the bottom panel shows the amplified signal clarity achieved when four transits are stacked together, underscoring a marked reduction in the noise profile. This comparison demonstrates the benefits of transit stacking for improved signal detection.

The prevailing weather conditions, particularly cloud cover, were likely the most substantial factors influencing the results. The cloud cover during all FOCES observations likely had a substantial impact on the detection statistics used for the comparison. The presence of clouds at some point during each of the three observations hindered the process of making a robust comparison between the two telescopes and their respective spectrographs.

An alternative explanation for the deteriorated data quality could be the air mass during the observations. However, when





**Fig. 6.** Detection significance trends when stacking transit observations based on the S/N of Wendelstein. Each line represents the number of transits required to reach the 3, 5, and 10 $\sigma$  significance thresholds marked as the grey dashed horizontal lines. These predictions are based on the 2500 K template, which has shallower lines, and therefore these predictions are conservative and could be more robust on a per-transit basis if the equilibrium temperatures of the planets were higher.

**Table 3.** Summary of the orbital and physical characteristics that could be observable with FOCES.

Name	RA (deg)	Dec (deg)	$V_{\text{mag}}$	$T_{\text{eq}}$ (K)	Period (days)	Semi-major axis (AU)	Mass ( $M_{\text{Jup}}$ )	Inclination (deg)
KELT-20 b	294.663	31.219	3.0	2230.0	3.5	0.054	3.5	86.2
WASP-189 b	225.687	-3.031	6.6	3353.0	2.7	0.051	2.0	84.03
KELT-9 b	307.86	39.939	7.6	3921.0	1.5	0.034	2.9	87.2
TOI-1431 b	316.204	55.588	8.0	2370.0	2.7	0.047	3.1	80.4
KELT-7 b	78.296	33.318	8.5	2048.0	2.7	0.045	1.2	83.9
HD 202772 A b	319.7	-26.616	8.3	2132.0	3.3	0.052	1.0	84.2
KELT-17 b	125.617	13.735	9.3	2087.0	3.1	0.049	1.3	84.87
HAT-P-70 b	74.55	9.998	9.5	2562.0	2.7	0.047	6.8	96.5
HD 2685 b	7.329	-76.304	9.6	2061.0	4.1	0.057	1.2	89.4
KOI-13 b	286.971	46.868	10.0	2550.0	1.8	0.036	9.3	86.77
TOI-2109 b	253.188	16.58	10.0	3631.0	0.7	0.018	5.0	70.74

**Notes.** The table includes right ascension (RA), declination (Dec), visual magnitude ( $V_{\text{mag}}$ ), equilibrium temperature ( $T_{\text{eq}}$ ), orbital period, semi-major axis, mass (in terms of Jupiter mass  $M_{\text{Jup}}$ ), and orbital inclination.

comparing all the observations in Fig. B.1, we observe that all were conducted in conditions of low air mass, suggesting this variable likely did not contribute significantly to the discrepancy in the results. A portion of all the three transits observed by FOCES took place during astronomical twilight, potentially introducing contamination from diffuse background light. However, as this contamination would not consistently affect all exposures, its overall impact is likely to be minimal.

An alternative factor that may have played a smaller role is the limited baseline measurements of the host star. The single night of HARPS-N has baseline observations that almost match the total baseline duration of all three FOCES observations. Despite the absence of additional planetary flux, this increase in the measurement of the stellar flux could permit

a more precise constraint on the stellar component subtracted during the cross-correlation process. To draw more definitive conclusions, additional controls are necessary, including consistent baseline timings, similar observing windows, and most importantly comparable weather conditions.

Despite these issues, our findings show that FOCES at the Wendelstein Observatory is capable of detecting signals from a multitude of chemical species present in the atmosphere of KELT-9 b. As detailed in Table A.1 and Fig. 4, we successfully retrieved spectral signals of seven atomic and ionised species, each with a detection significance exceeding 3 $\sigma$ . Moreover, the enhancements in detection significance, indicated in Table 2 and Fig. 3, suggest that multiple transit observations can lead to results that are of comparable quality to those from HARPS-N.

We note that HARPS-N has successfully conducted day-side observations of KELT-9 b (Pino et al. 2020). Based on our findings, we believe that it is reasonable to hypothesise that analogous observations could be carried out using FOCES. These results suggest that smaller telescopes could play a pivotal role in broadening the scientific scope of exoplanet atmospheric studies.

The distinctive advantage that small telescopes offer lies in the abundance of their observational time. These telescopes are capable of revisiting targets repeatedly, which not only places them in a competitive position with larger classes of telescopes, but also allows them to track the temporal evolution of these planetary atmospheres. Larger telescopes, due to intense competition for observation time and the drive for unique discoveries, tend to impose limitations on the scope of potential scientific questions that may require substantial amounts of observing time. Therefore, small telescopes could play a crucial role in the systematic monitoring of such planetary atmospheres, especially considering their known dynamic evolution. For example, recent observations suggested that KELT-9 b is gradually losing its atmosphere (Wytenbach et al. 2020). Therefore, the persistent and repetitive observational capabilities of small telescopes could provide invaluable insights into the evolution of the atmospheres of exoplanets.

Leveraging the successful detections achieved with FOCES, we generalised the observational capabilities of 2 m telescopes through the creation of an emulator. The emulator reproduces a similar statistical significance to FOCES, and also creates a statistical significance comparable to HARPS-N when four emulated nights are combined, as shown in Fig. 5. This outcome suggests that resolving the atmospheres of exoplanets only requires stacking transits and can yield results comparable to those obtained with larger classes of telescopes. With this success, we generalised the average S/N of FOCES to predict how many transits would be required to resolve the atmospheres of other known UHJs with a 2 m telescope. Using the exoplanet archive, we compiled a list of potential UHJ targets that could feasibly be observed with a 2 m telescope in under 20 transits, as shown in Table 3. Such a high number of observations would be a formidable undertaking for a large telescope, but is indeed plausible given the ample observation time available with 2 m telescopes. We chose targets that exhibit a sufficiently high atmospheric equilibrium temperature to allow molecular dissociation, thereby enabling us to observe atomic and ionised line transitions at optical wavelengths. Should an observation programme be launched with 2 m telescopes, most targets would need to be observed over several transits, except for planets with magnitudes of less than 7.5 (see Figs. 6, D.2, and D.3). Admittedly, unforeseeable factors, such as weather, may lead to the requirement for additional observations in order to offset lost nights. Nevertheless, with perseverance, 2 m telescopes, like that at the Wendelstein Observatory, have significant potential for the observation of an extensive range of UHJ atmospheres.

In addition to providing arguments for the potential of 2 m telescopes, our emulator seems to suggest there is a limit to the increase in statistical significance when stacking transit observations. In Fig. 6, for stars with visual magnitudes of 6.5, significance increases with the square root of the number of transits, yet appears to reach a limit after multiple transits have been stacked; this trend can also be seen in Fig. D.3, and to a lesser extent in Fig. D.2. However, the limit lies at different levels of significance: the 4000 K template reaches a significance of 50, while the 2500 K template only reaches a significance of 25. To better illustrate this phenomenon, we present Fig. D.4, which

compares the increasing trends in detection significance for emulated stars of visual magnitudes 6.0 and 11.5. This plot distinctly reveals a divergence from the expected Poisson trend in the case of the magnitude 6.0 star. The explanation we propose is that with a sufficient number of transits, the alias floor of the CCF (Borsato et al. 2023) becomes resolved, becoming the primary contributor to the standard deviation calculation used to model the noise profile described in Sect. 5, causing the noise level to stay locked at this detection significance. In this case specifically, the aliases arise only from the self-correlation of lines in the Fe I template. In an actual case, other additional aliases can come from different species in the planetary spectrum and could have a stronger effect. The reason for this variation in level is linked to the differences in the line profiles of each template, which vary in the amount of aliasing lines present in the CCF. Figure D.1 illustrates this, showing that more lines are present in the 4000 K template than in the 2500 K and 2000 K templates, respectively. This result suggests that the HARPS-N observation of KELT-9 b is likely already resolving the alias floor, with a single observation being sufficient to constrain the noise profile. Therefore, in high-S/N regimes, stacking transits yields diminishing returns with the cross-correlation technique, unless aliasing is accounted for.

While FOCES provides a demonstration of the potential of a high-performance spectrograph on a 2 m telescope, its scientific reach is primarily limited to UHJs, because its wavelength range is limited to optical wavelengths. Recent studies, particularly those involving the GIANO-B spectrograph on the TNG (Brogi et al. 2018; Guilluy et al. 2019), have convincingly demonstrated the scientific utility of high-resolution near-infrared observations for exoplanets with lower equilibrium temperatures, such as hot Jupiters. This suggests that similar science cases—ranging from helium evaporation (Oklopčić & Hirata 2018; Nortmann et al. 2018) to the detection of various molecular species, such as CO (Yan et al. 2022) – could potentially be replicated on smaller telescopes, provided they are outfitted with comparable near-infrared capabilities. The capability for long-term baseline measurements facilitated by small telescopes could offer in-depth insights into this evaporation process, potentially shedding light on competing theories of planetary evaporation, such as photo-evaporation (Hu et al. 2015; Owen & Wu 2017) versus core-powered mass loss (Ginzburg et al. 2018; Gupta & Schlichting 2019, 2020). Given the comparable performance of FOCES with respect to that of HARPS-N, it is plausible that a similar near-infrared spectrograph like GIANO-B could achieve similar results.

Large telescopes often excel in scientific investigations that demand precision and reduced noise levels; for example when resolving individual lines (e.g. Seidel et al. 2022; Pai Asnodkar et al. 2022), analysing atmospheric dynamics (Ehrenreich et al. 2020; Prinoth et al. 2022), or hunting for fainter signals from heavier elements (Jiang et al. 2023). However, a common challenge with large telescopes is their oversubscription. It is assumed that oversubscription leads to prioritisation of the best-qualified projects (Castro & Yáñez 2003); which while generally accurate, reduces the scope for creative experimentation and diverse observational strategies. It is worth noting that the first exoplanet discovery, 51 Pegasi b, was made using a 2 m class telescope (Mayor & Queloz 1995), despite the availability of larger telescopes at the time. Our study demonstrates that when it comes to studying exoplanet atmospheres, smaller telescopes can deliver comparable performance to their larger counterparts, enabling further experimentation and possibly leading to new discoveries.

## 8. Conclusion

In this study, we successfully resolved the atmosphere of KELT-9 b in a single transit, and detected seven of the eight detections presented in Hoeijmakers et al. (2019). Additionally, we compared the performances of the high-resolution spectrograph FOCES at the Wendelstein Observatory to the HARPS-N spectrograph on the TNG and found promising evidence that, given more observations, FOCES can produce results of a comparable level of quality to those obtained with HARPS-N. UHJs can be viewed as testing grounds for observational concepts related to atmospheres. With our results it can be argued that smaller telescopes carry the potential to explore new ideas and reframe the scientific questions that we should pursue with larger telescopes. Our results demonstrate that efforts to maintain smaller, high-performance telescopes are worthwhile, that 2 m telescopes are likely to continue to deliver pioneering science, and suggest that we should even expand the instrumentation they are equipped with in order to study UHJ atmospheres in greater detail.

*Acknowledgements.* N.W.B. acknowledge partial financial support from The Fund of the Walter Gyllenberg Foundation. B.E. and K.M. are supported by the Excellence Cluster ORIGINS which is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2094-390783311. D.C. is supported by the LMU-Munich Fraunhofer-Schwarzschild Fellowship and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2094 – 390783311.

## References

- Arcangeli, J., Désert, J.-M., Line, M. R., et al. 2018, *ApJ*, **855**, L30
- Arcangeli, J., Désert, J. M., Parmentier, V., Tsai, S. M., & Stevenson, K. B. 2021, *A&A*, **646**, A94
- Azevedo Silva, T., Demangeon, O. D. S., Santos, N. C., et al. 2022, *A&A*, **666**, L10
- Baraffé, I., Chabrier, G., & Barman, T. 2010, *Rep. Progr. Phys.*, **73**, 016901
- Bello-Arufe, A., Buchhave, L. A., Mendonça, J. M., et al. 2022, *A&A*, **662**, A51
- Borsa, F., Giacobbe, P., Bonomo, A. S., et al. 2022, *A&A*, **663**, A141
- Borsato, N. W., Hoeijmakers, H. J., Prinoth, B., et al. 2023, *A&A*, **673**, A158
- Brogi, M., Snellen, I. A. G., de Kok, R. J., et al. 2012, *Nature*, **486**, 502
- Brogi, M., Giacobbe, P., Guilluy, G., et al. 2018, *A&A*, **615**, A16
- Brucalassi, A., Feger, T., Grupp, F., et al. 2012, *Proc. SPIE*, **8446**, 84462F
- Casasayas-Barris, N., Pallé, E., Yan, F., et al. 2019, *A&A*, **628**, A9
- Castro, A. I. G. D., & Yáñez, J. 2003, *A&A*, **403**, 357
- Cont, D., Yan, F., Reiners, A., et al. 2021, *A&A*, **651**, A33
- Cont, D., Yan, F., Reiners, A., et al. 2022, *A&A*, **668**, A53
- Ehrenreich, D., Lovis, C., Allart, R., et al. 2020, *Nature*, **580**, 597
- Fossati, L., Young, M. E., Shulyak, D., et al. 2021, *A&A*, **653**, A52
- Gaudi, B. S., Stassun, K. G., Collins, K. A., et al. 2017, *Nature*, **546**, 514
- Ginzburg, S., Schlichting, H. E., & Sari, R. 2018, *MNRAS*, **476**, 759
- Grimm, S. L., & Heng, K. 2015, *ApJ*, **808**, 182
- Grimm, S. L., Malik, M., Kitzmann, D., et al. 2021, *ApJS*, **253**, 30
- Guilluy, G., Sozzetti, A., Brogi, M., et al. 2019, *A&A*, **625**, A107
- Gupta, A., & Schlichting, H. E. 2019, *MNRAS*, **487**, 24
- Gupta, A., & Schlichting, H. E. 2020, *MNRAS*, **493**, 792
- Helling, C., Lewis, D., Samra, D., et al. 2021, *A&A*, **649**, A44
- Hoeijmakers, H. J., de Kok, R. J., Snellen, I. A. G., et al. 2015, *A&A*, **575**, A20
- Hoeijmakers, H. J., Ehrenreich, D., Heng, K., et al. 2018, *Nature*, **560**, 453
- Hoeijmakers, H. J., Ehrenreich, D., Kitzmann, D., et al. 2019, *A&A*, **627**, A165
- Hoeijmakers, H. J., Seidel, J. V., Pino, L., et al. 2020, *A&A*, **641**, A123
- Hoeijmakers, H. J., Kitzmann, D., Morris, B. M., et al. 2024, *A&A*, in press, <https://doi.org/10.1051/0004-6361/202244968>
- Hu, R., Seager, S., & Yung, Y. L. 2015, *ApJ*, **807**, 8
- Jiang, Z., Wang, W., Zhao, G., et al. 2023, *AJ*, **165**, 230
- Kausch, W., Noll, S., Smette, A., et al. 2015, *A&A*, **576**, A78
- Kitzmann, D., Heng, K., Rimmer, P. B., et al. 2018, *ApJ*, **863**, 183
- Kitzmann, D., Hoeijmakers, H. J., Grimm, S. L., et al. 2023, *A&A*, **669**, A113
- Lee, E. K. H., Prinoth, B., Kitzmann, D., et al. 2022, *MNRAS*, **517**, 240
- Lowson, N., Zhou, G., Wright, D. J., et al. 2023, *AJ*, **165**, 101
- Mayor, M., & Queloz, D. 1995, *Nature*, **378**, 355
- McLaughlin, D. B. 1924, *ApJ*, **60**, 22
- Merritt, S. R., Gibson, N. P., Nugroho, S. K., et al. 2021, *MNRAS*, **506**, 3853
- Nortmann, L., Pallé, E., Salz, M., et al. 2018, *Science*, **362**, 1388
- Oklopčić, A., & Hirata, C. M. 2018, *ApJ*, **855**, L11
- Owen, J. E., & Wu, Y. 2017, *ApJ*, **847**, 29
- Pai Asnodkar, A., Wang, J., Gaudi, B. S., et al. 2022, *AJ*, **163**, 40
- Parmentier, V., Line, M. R., Bean, J. L., et al. 2018, *A&A*, **617**, A110
- Pfeiffer, M. J., Frank, C., Baumüller, D., Fuhrmann, K., & Gehren, T. 1998, *A&AS*, **130**, 381
- Pino, L., Désert, J.-M., Brogi, M., et al. 2020, *ApJ*, **894**, L27
- Pino, L., Brogi, M., Désert, J. M., et al. 2022, *A&A*, **668**, A176
- Prinoth, B., Hoeijmakers, H. J., Kitzmann, D., et al. 2022, *Nat. Astron.*, **6**, 449
- Prinoth, B., Hoeijmakers, H. J., Pelletier, S., et al. 2023, *A&A*, **678**, A182
- Rossiter, R. A. 1924, *ApJ*, **60**, 15
- Seidel, J. V., Cegla, H. M., Doyle, L., et al. 2022, *MNRAS*, **513**, L15
- Showman, A. P., & Polvani, L. M. 2011, *ApJ*, **738**, 71
- Smette, A., Sana, H., Noll, S., et al. 2015, *A&A*, **576**, A77
- Snellen, I. A. G., de Kok, R. J., de Mooij, E. J. W., & Albrecht, S. 2010, *Nature*, **465**, 1049
- Stangret, M., Casasayas-Barris, N., Pallé, E., et al. 2020, *A&A*, **638**, A26
- Stangret, M., Pallé, E., Casasayas-Barris, N., et al. 2021, *A&A*, **654**, A73
- Stock, J. W., Kitzmann, D., Patzer, A. B. C., & Sedlmayr, E. 2018, *MNRAS*, **479**, 865
- Stock, J. W., Kitzmann, D., & Patzer, A. B. C. 2022, *MNRAS*, **517**, 4070
- Wang, L., Grupp, F., Kellermann, H., Hopp, U., & Bender, R. 2017, *Proc. SPIE*, **10400**, 104001V
- Wytenbach, A., Mollière, P., Ehrenreich, D., et al. 2020, *A&A*, **638**, A87
- Yan, F., Casasayas-Barris, N., Molaverdikhani, K., et al. 2019, *A&A*, **632**, A69
- Yan, F., Pallé, E., Reiners, A., et al. 2022, *A&A*, **661**, A6



## Appendix A: Detections

We present the measured parameters of each cross-correlation peak for each signal recovered when combining the three FOCES observations. Table A.1 outlines the measured amplitude, centre,

width, and detection significance of each cross-correlation peak. All detection significances, except for Y II, are detected above a  $3\sigma$  threshold.

Table A.1: Summary of the detection parameters and significance for the each of the eight species in the transmission spectrum of KELT-9 b.

Element	Amplitude [ppm]	Centre [km/s]	Width [km/s]	Detection Significance
Na I	$1118 \pm 242$	$-17.1 \pm 2.49$	$5.83 \pm 2.31$	4.6
Mg I	$403 \pm 112$	$-26.8 \pm 2.68$	$7.30 \pm 1.60$	3.6
Sc II	$1281 \pm 191$	$-21.7 \pm 1.15$	$5.29 \pm 1.06$	6.7
Ti II	$976 \pm 183$	$-19.1 \pm 1.64$	$5.06 \pm 2.75$	5.3
Cr II	$986 \pm 322$	$-25.9 \pm 1.57$	$3.02 \pm 1.56$	3.1
Fe II	$1387 \pm 121$	$-17.4 \pm 0.81$	$7.65 \pm 0.87$	11.4
Fe I	$147 \pm 21$	$-16.9 \pm 1.25$	$6.62 \pm 1.09$	7.1
Y II	$20 \pm 1080$	$-20.2 \pm 5.48$	$12.4 \pm 232$	0.02

**Notes.** The table lists the amplitude, centre, width, and detection significance for each detected element

## Appendix B: Air-mass plots

Here, we present the air-mass plots for each night of observation, highlighting where the transit occurs.

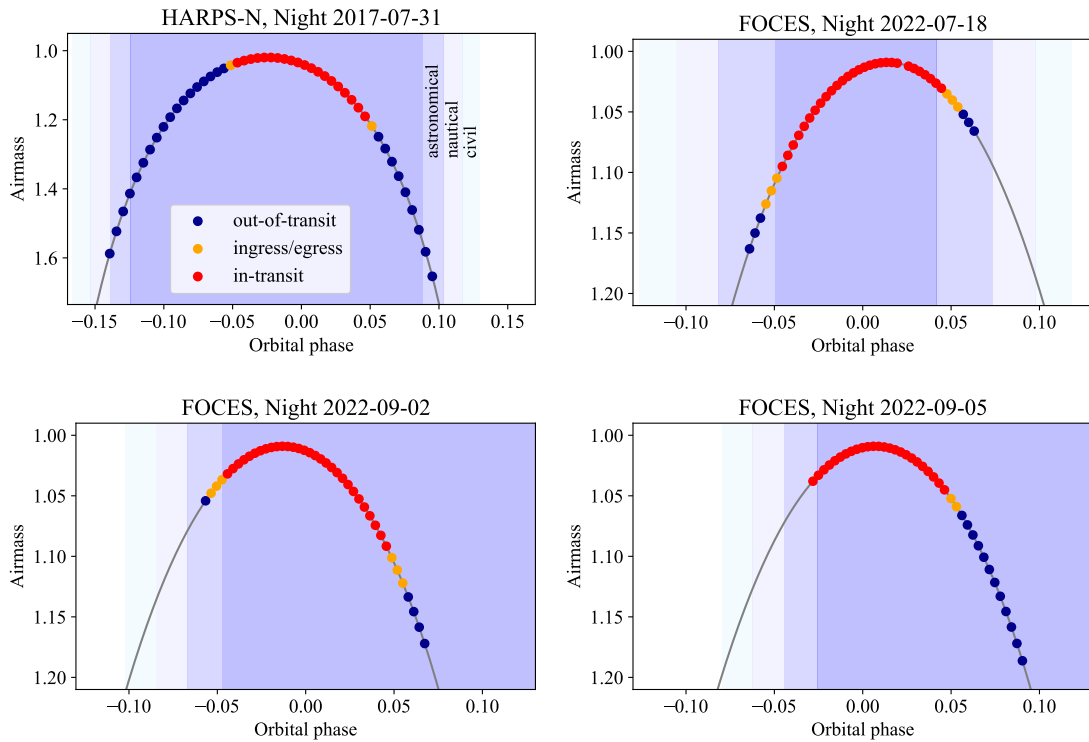


Fig. B.1: Airmass plots for KELT-9 b observations. Each data point represents an individual exposure frame. In-transit, ingress/egress, and out-of-transit exposures are indicated with red, orange, and blue. The civil, nautical, and astronomical twilight phases are indicated by the blue shaded regions. All transits were observed at low airmass. Parts of the transits observed with FOCES occurred during astronomical twilight given the short duration of the nights in summer.

## Appendix C: KELT-9 b atmospheric model

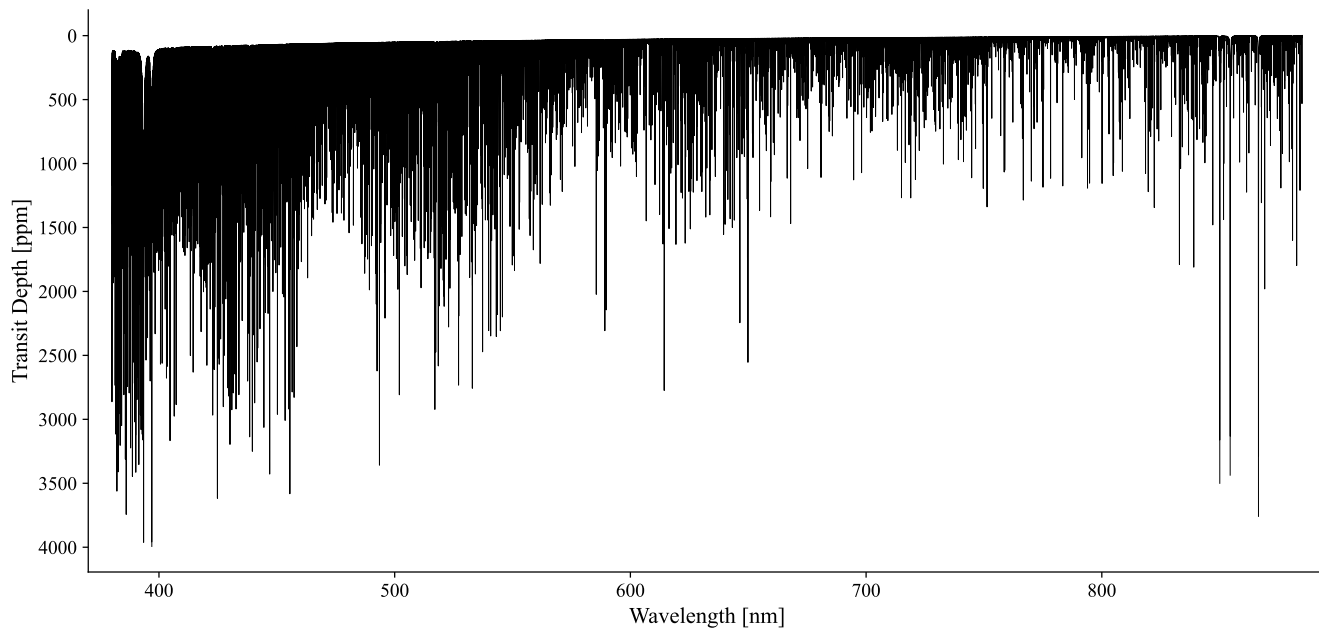


Fig. C.1: Radiative transfer model employed for model injection in this research. The model spans the wavelength range of FOCES. Line depth varies, typically on the order of a few 1000 ppm.

In this section, we present the atmospheric model utilised for the injection procedure of this study, shown in Fig. C.1. The model, on average, exhibits a relative transit depth on the order of a few tenths of a percent for the deepest lines. The high transit depth highlights why this planet is so conducive to atmospheric detections, which is a critical factor in WENDELSTEIN's success in detecting the planet's atmosphere in a single transit.

## Appendix D: Targets observable for Wendelstein

In this section, we present three plots that illustrate the temperature dependence on line depth and its subsequent impact on detection significance. The first plot displays transmission templates from the Mantis network library (Kitzmann et al. 2023), which were utilised in the emulator developed in this study. The remaining two plots depict the increase in detection significance for the 2000 K and 4000 K templates with a varying number of transits, illustrating the temperature-dependent detectability of specific species in exoplanet atmospheres. Given the aperture size of 2 m class telescopes, it is anticipated that multiple transit observations will be necessary to discern the atmospheres of targets dimmer than 7.5 magnitude; nevertheless, achieving this objective remains a viable endeavour. Furthermore, we also include a comparison of the increase in S/N for 6.0 and 11.5 magnitude stars, which depicts how well the stacking results of the CCF align with Poisson trends.

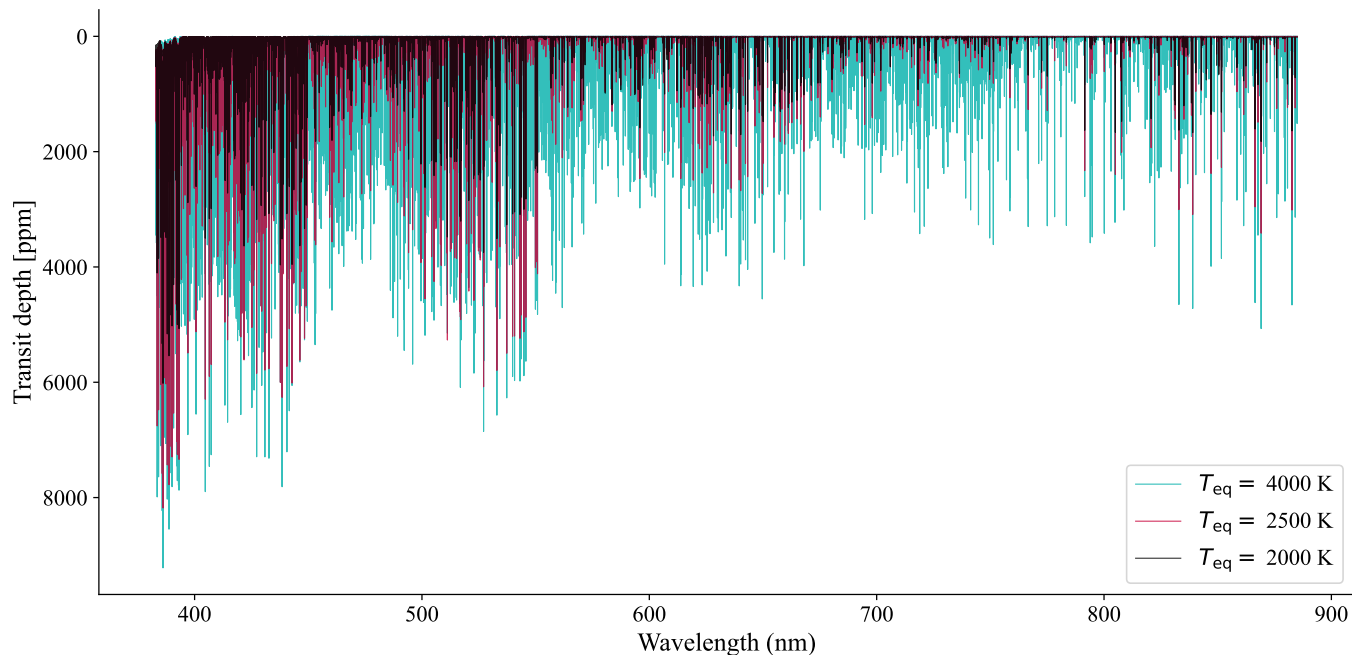


Fig. D.1: Variation in Fe I transmission with temperature. The plot illustrates three Fe I transmission templates used to emulate the transits. The 4000K template is represented in blue, the 2500 K template in red, and the 2000 K template in black. A discernible trend of increasing transmission depths is observed with escalating temperature, wherein more lines become pronounced and stronger. This enhancement in line strength with temperature increment effectively introduces more structure to the auto-correlation function, elucidating the temperature-dependent behaviour of Fe I transmission in the emulation of transits.

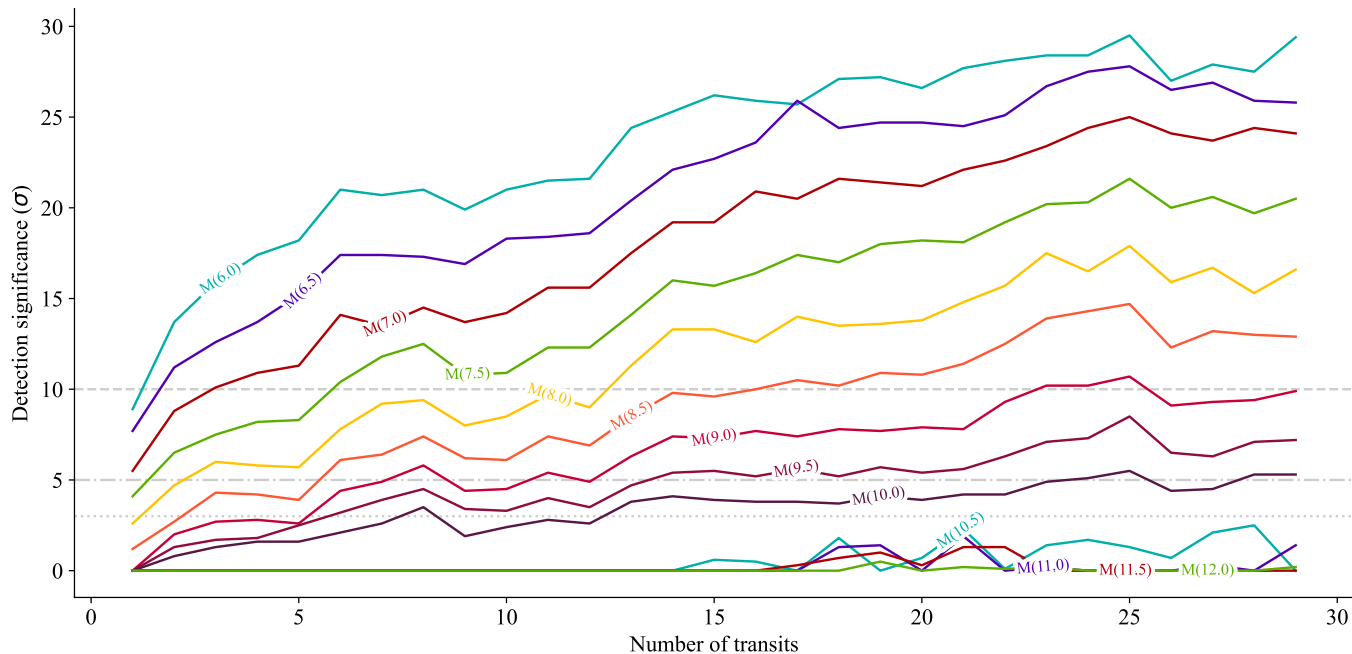


Fig. D.2: Detection significance trends for the 2000 K following the same process as Fig. 6.



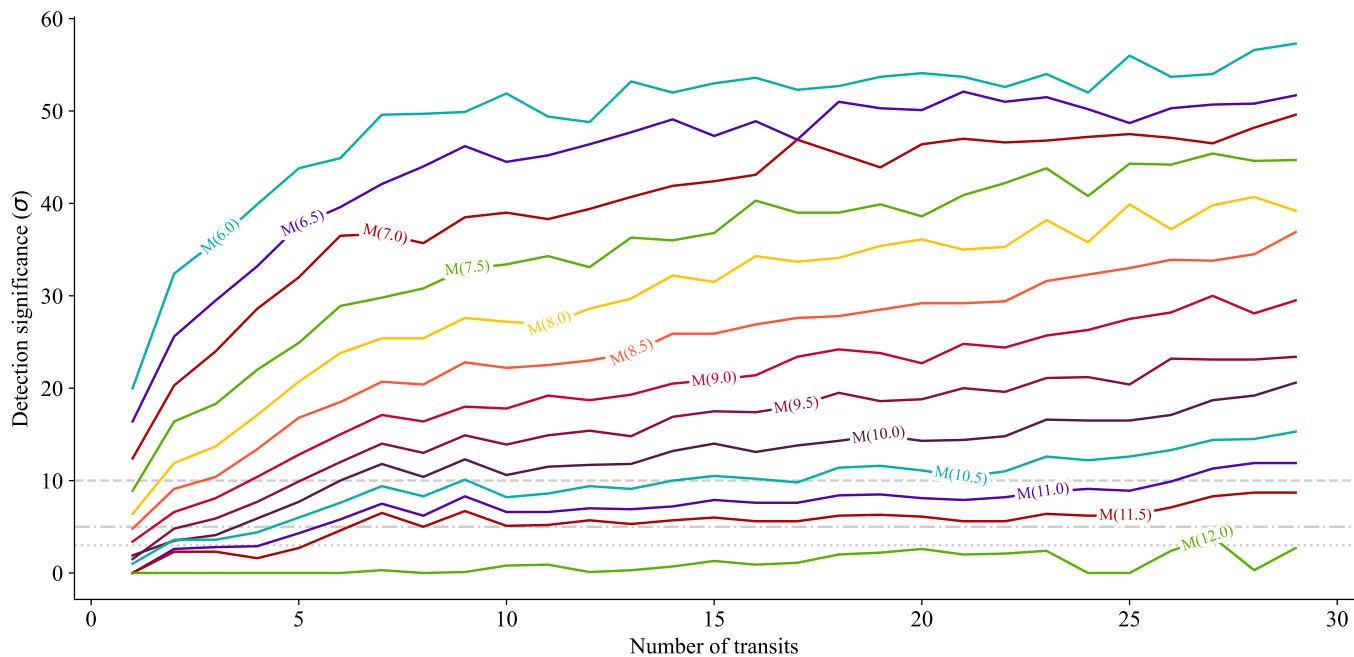


Fig. D.3: Detection significance trends for the 4000 K following the same process as Fig. 6.

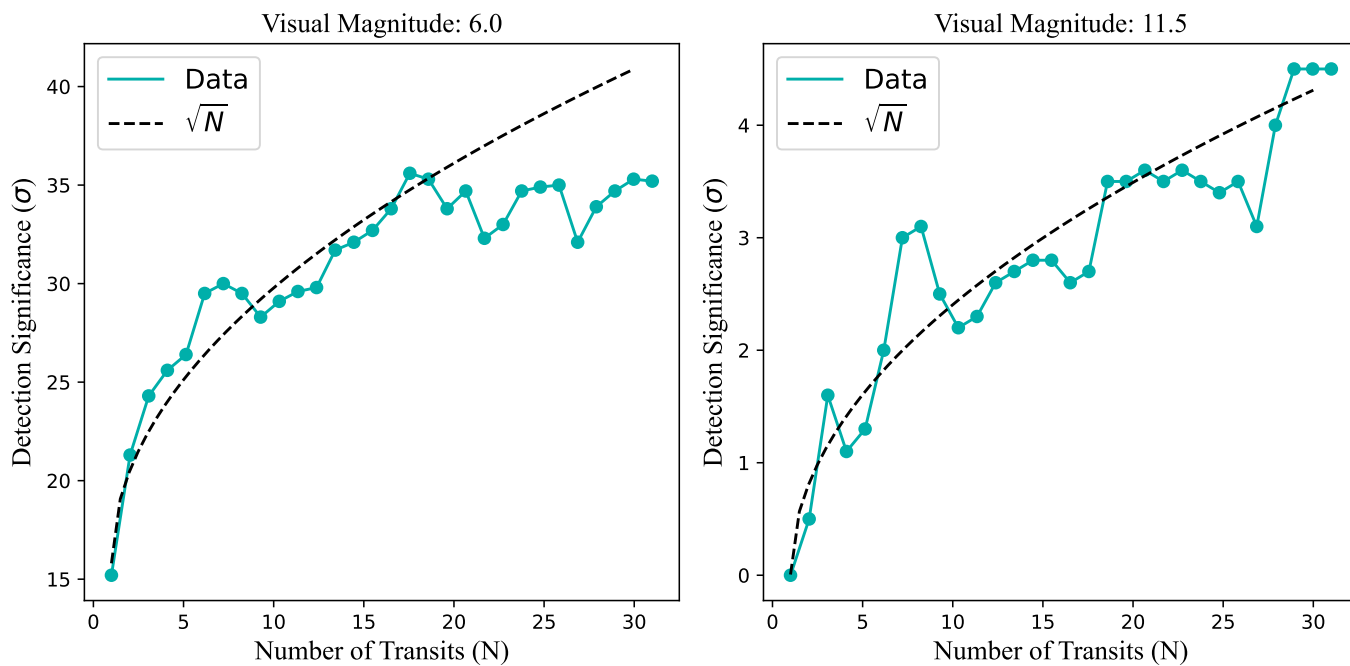


Fig. D.4: Comparison of the predicted Poisson trend with actual significance in stacked transits, with the transit emulator. Each plot features the number of transits against the detection significance, depicted in teal, alongside a dashed line representing the square root fit of the first ten stacked transits. The left panel illustrates results from stacking transits around a star of magnitude 6.0. The right panel shows a similar trend for a star of magnitude 11.5.